

Essentials of MOSFETs

Unit 1: Transistors and Circuits

26/5/2020

Lecture 1.1: The MOSFET as a Black Box

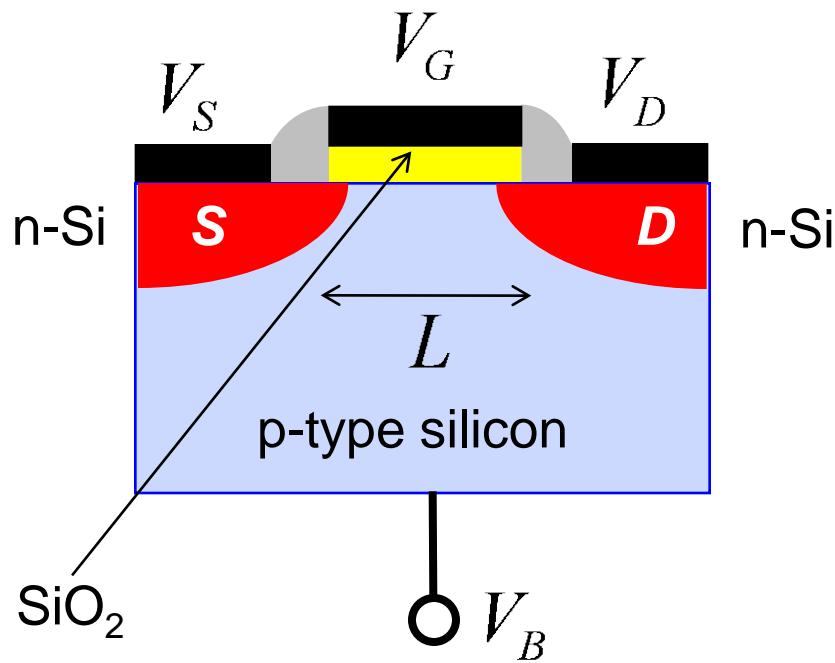
s.HARIPRAKASH

Mark Lundstrom

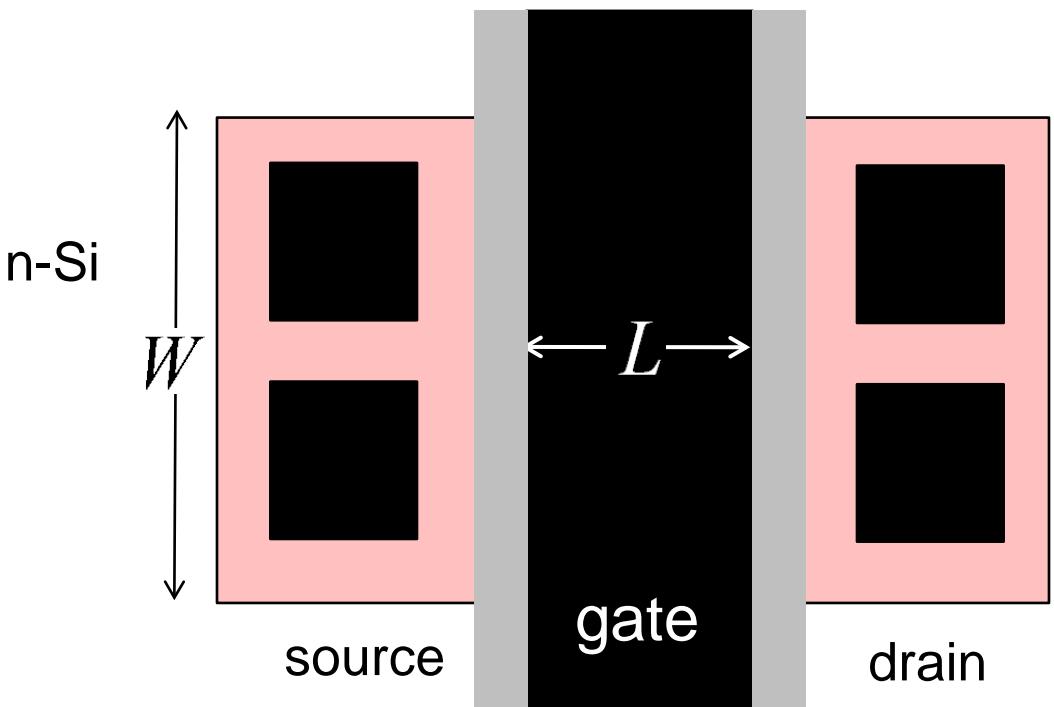
lundstro@purdue.edu
Electrical and Computer Engineering
Purdue University
West Lafayette, Indiana USA

Side and top views of a MOSFET

Metal Oxide Semiconductor Field Effect Transistor



side view

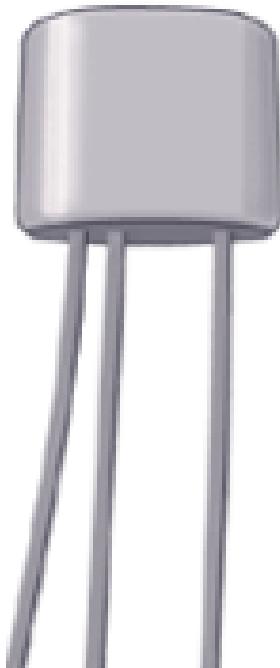


Lundstrom: 2018

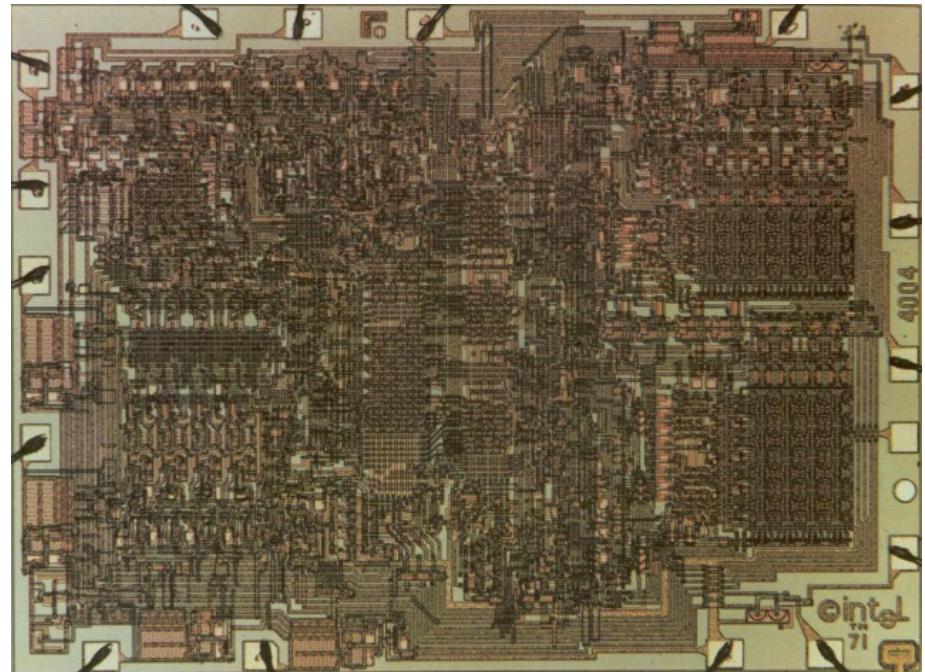
top view

Transistors

Discrete



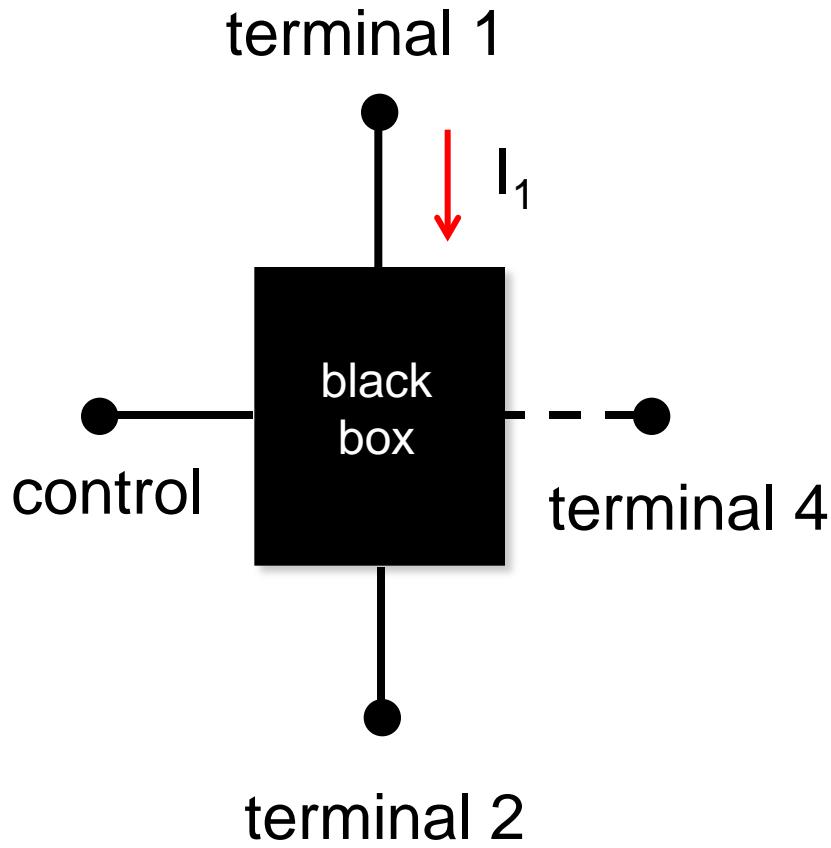
Integrated circuits



Intel 4004 (2300 transistors) 1971

<https://www.extremetech.com/computing/105029-intel-4004-the-first-cpu-is-40-years-old-today>

The transistor as a “black box”



There are many kinds of transistors:

MOSFET

SOI MOSFET

FinFETs

SB FET

FinFET

MODFET (HEMT)

bipolar transistor

JFET

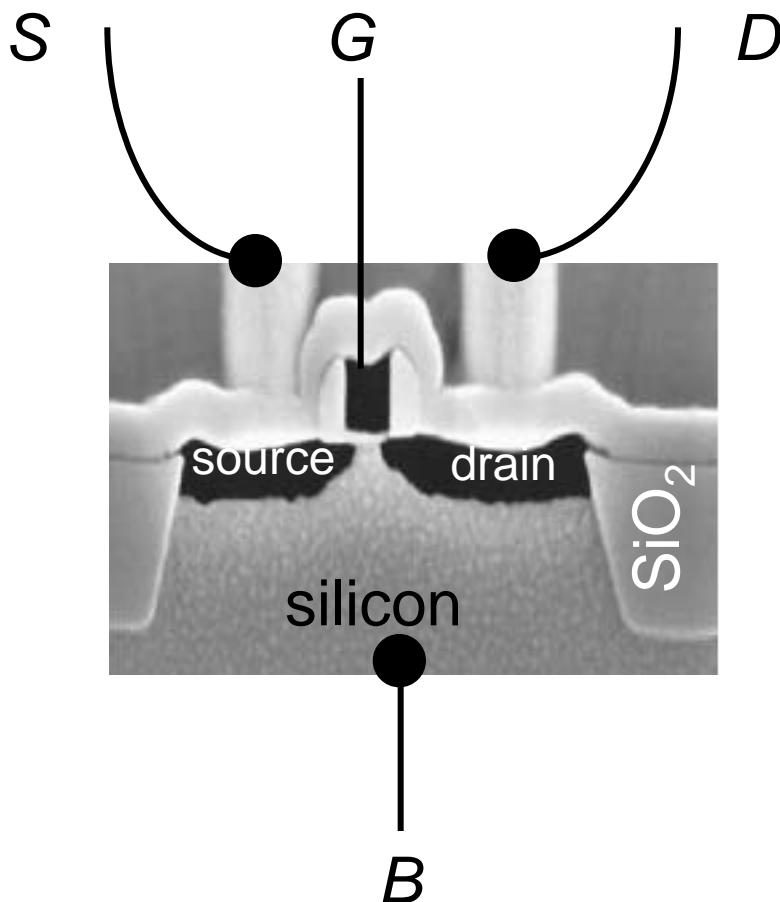
heterojunction bipolar transistor

BTBT FET

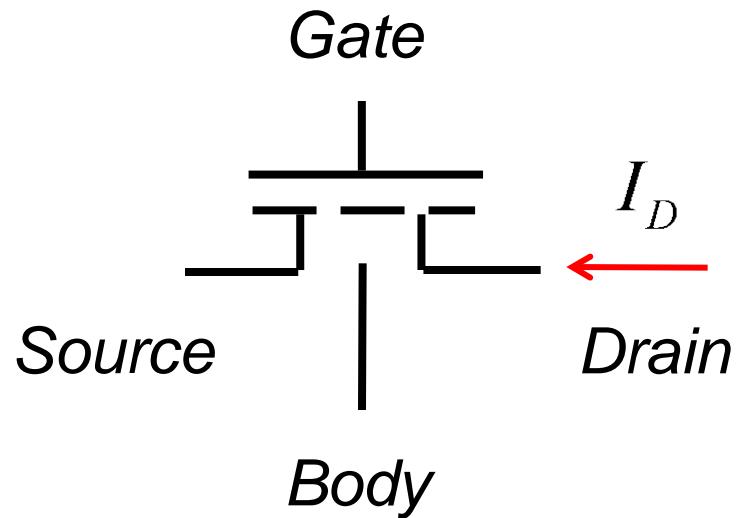
SpinFET

...

The bulk MOSFET



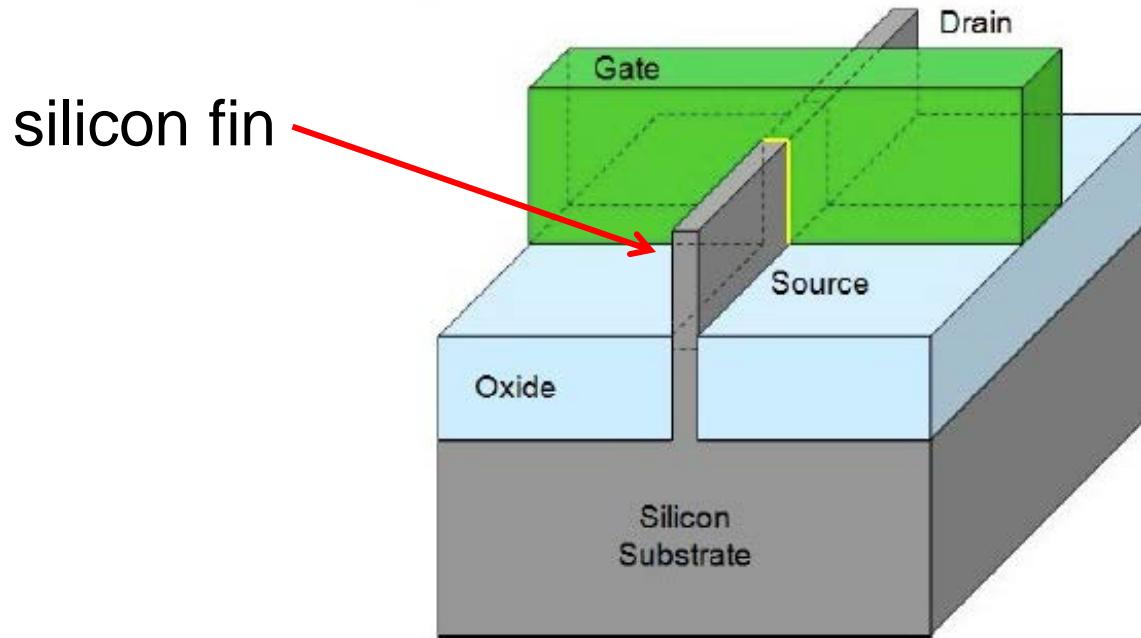
circuit symbol



(Texas Instruments, ~2000)

Lundstrom: 2018

Modern MOSFETs: The FinFET



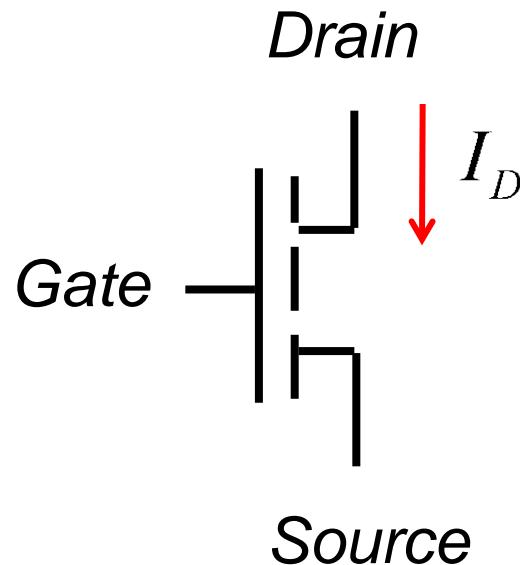
Source: Intel

Digh Hisamoto, Wen-Chin Lee, Jakub Kedzierski, Hideki Takeuchi, Kazuya Asano, Charles Kuo, Erik Anderson, Tsu-Jae King, Jeffrey Bokor, Chenming Hu, "FinFET-a self-aligned double-gate MOSFET scalable to 20 nm," *IEEE Transactions on Electron Devices*, **47**, 2320-2325, 2000.

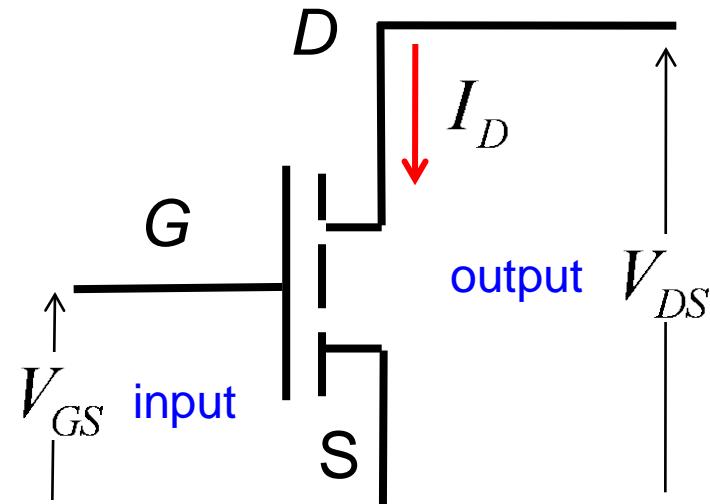
(VS)
TRANSFER CHARACTERISTICS: THE OUTPUT CURRENT IN TERMS OF INPUT VOLTAGE AT A FIXED OUTPUT VOLTAGE.

The MOSFET as a 2-port device

MOSFET circuit symbol



common source



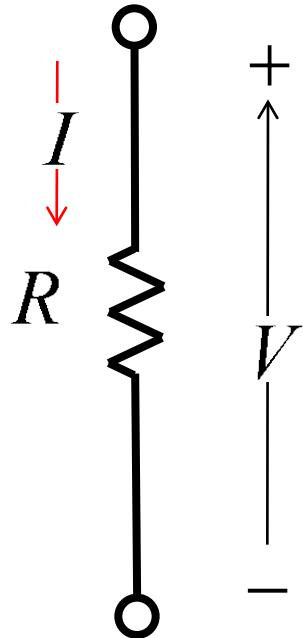
current vs. voltage (IV)
characteristics

$$I_D(V_G, V_S, V_D)$$

$I_D(V_{GS})$ at a fixed V_{DS} transfer

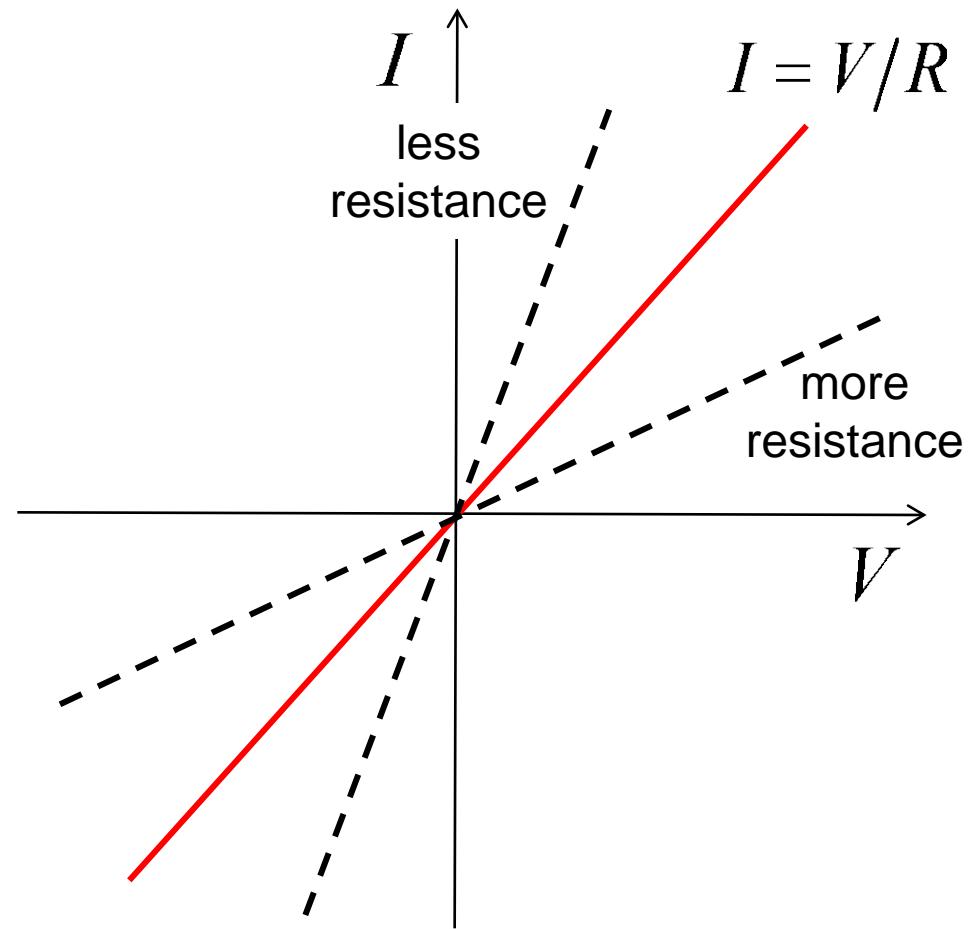
$I_D(V_{DS})$ at a fixed V_{GS} output

IV characteristics: resistor

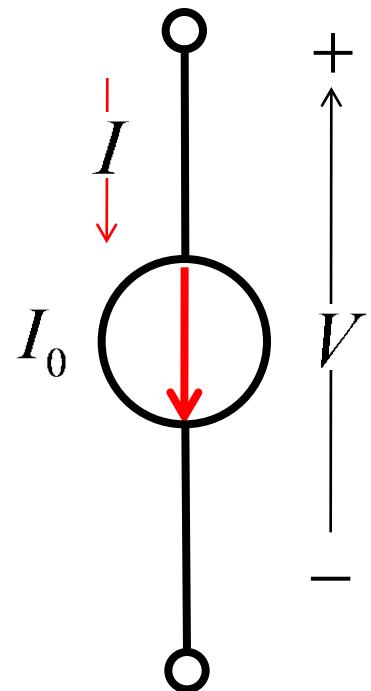


$$I = V/R \text{ Ohm's Law}$$

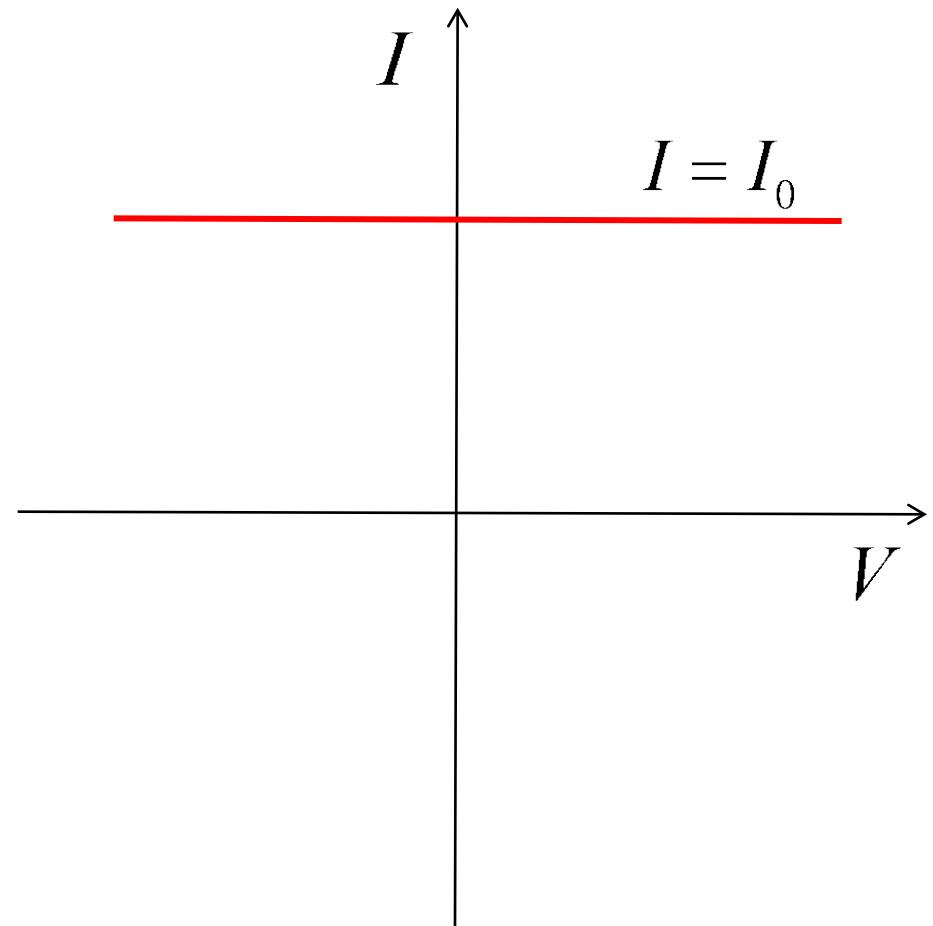
Georg Ohm, 1827



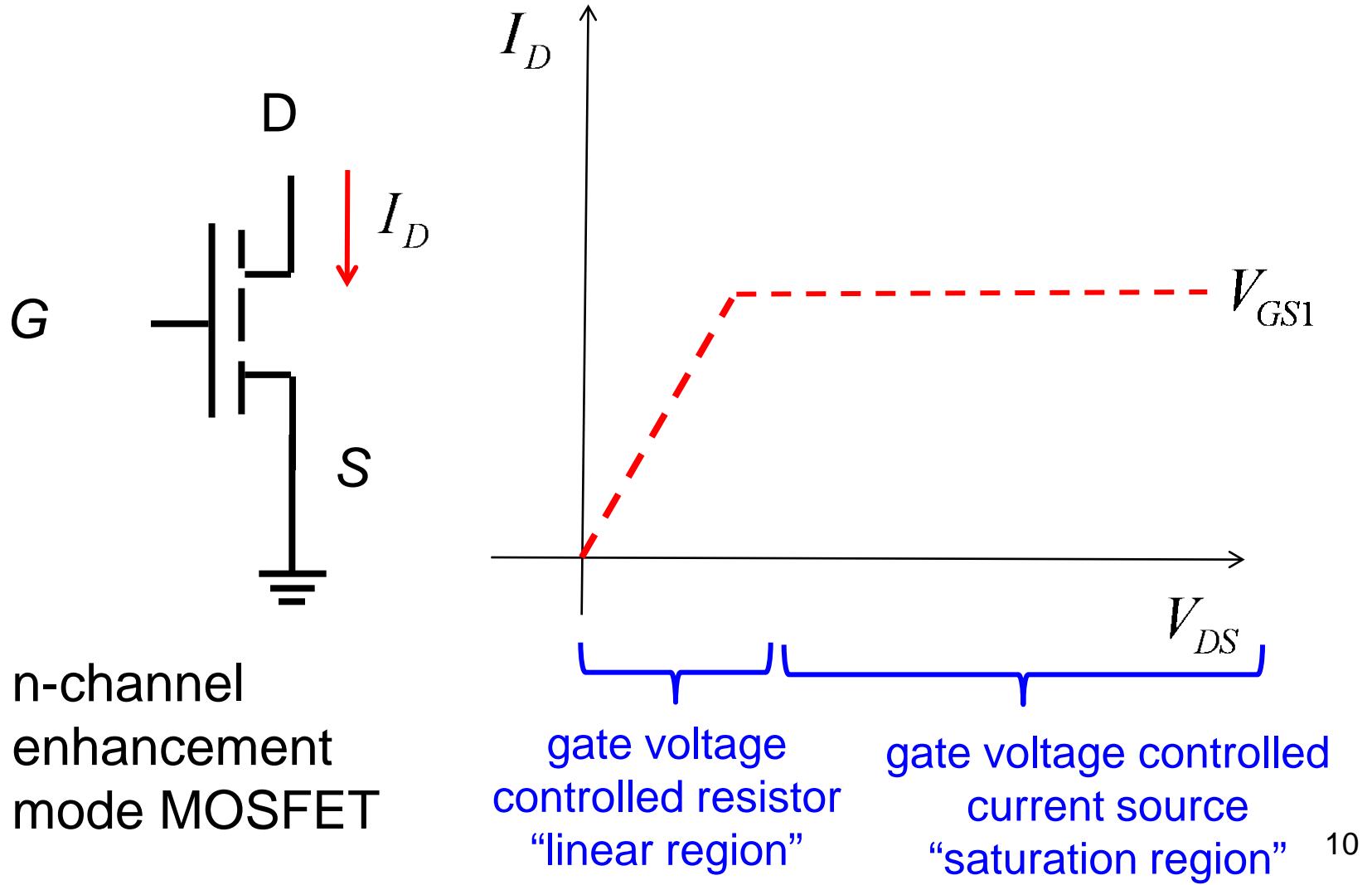
IV characteristics: ideal current source



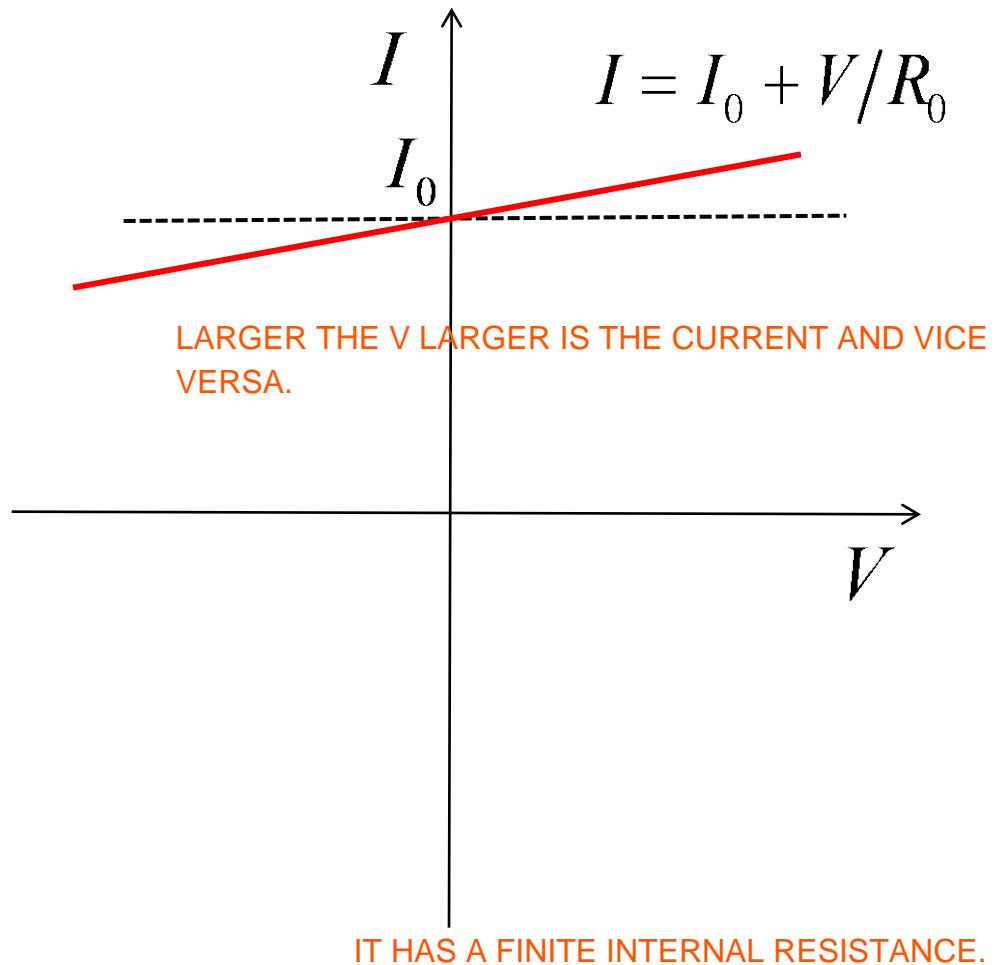
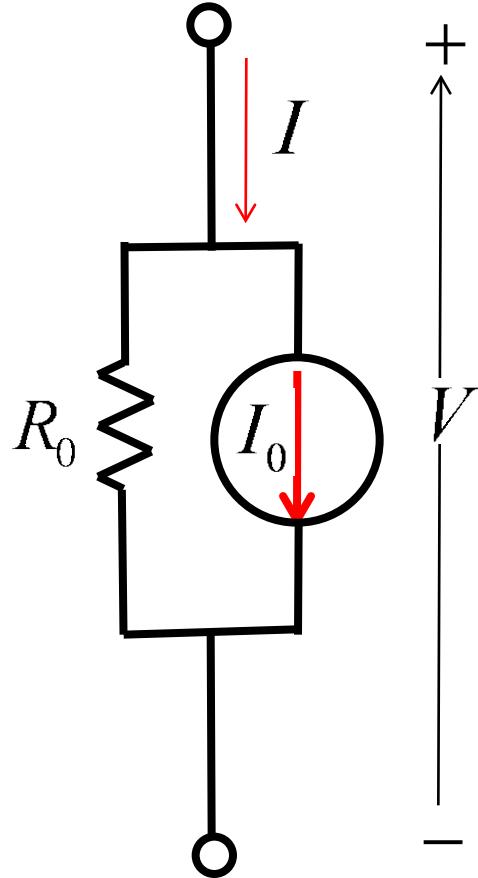
$$I = I_0$$



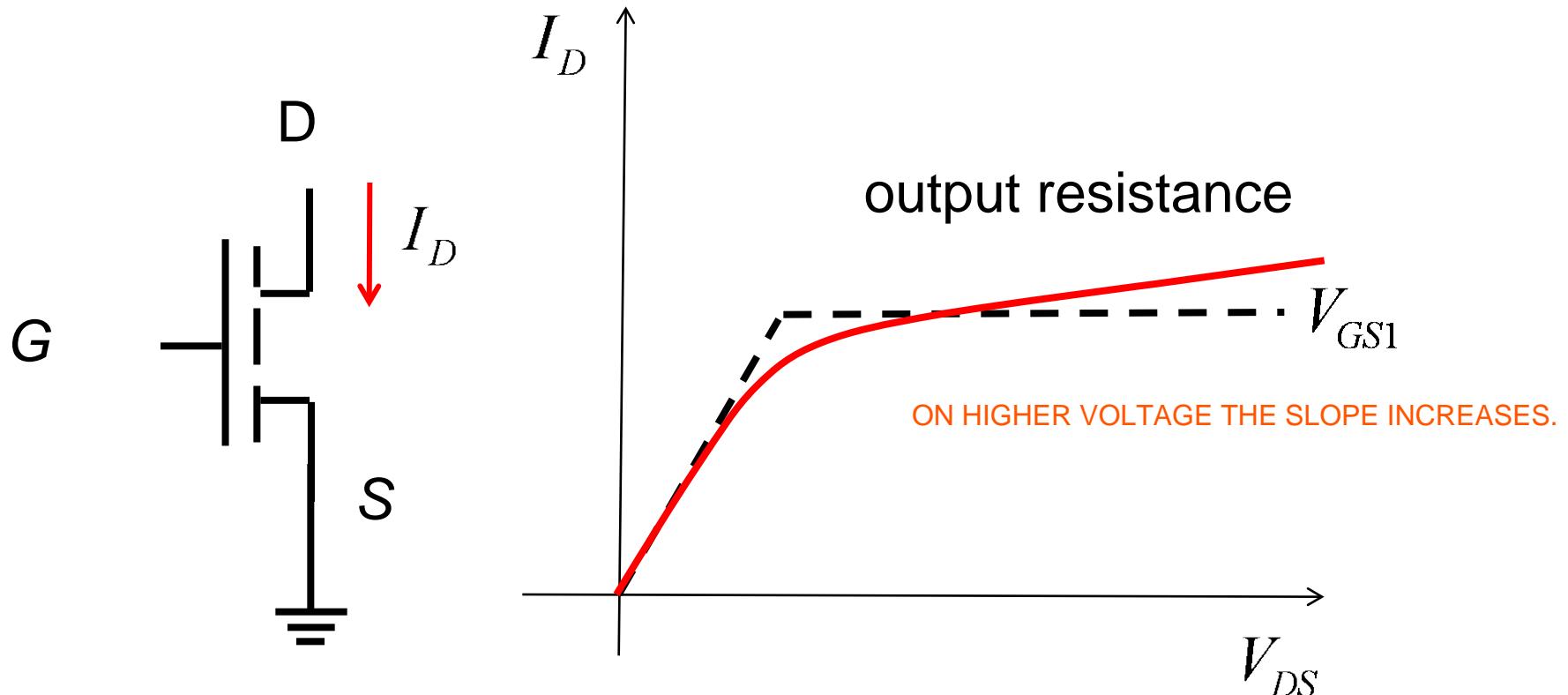
IV characteristics: transistors



IV characteristics: real current sources

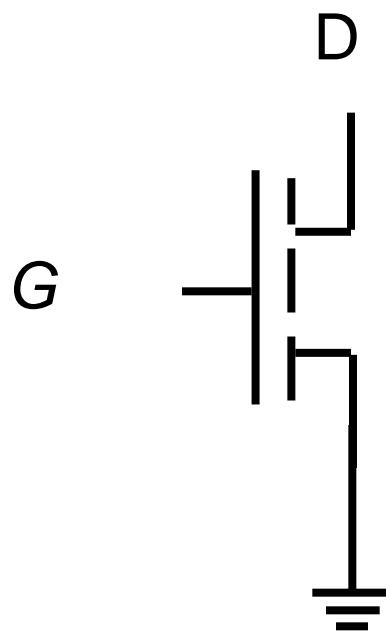


IV characteristics: transistors

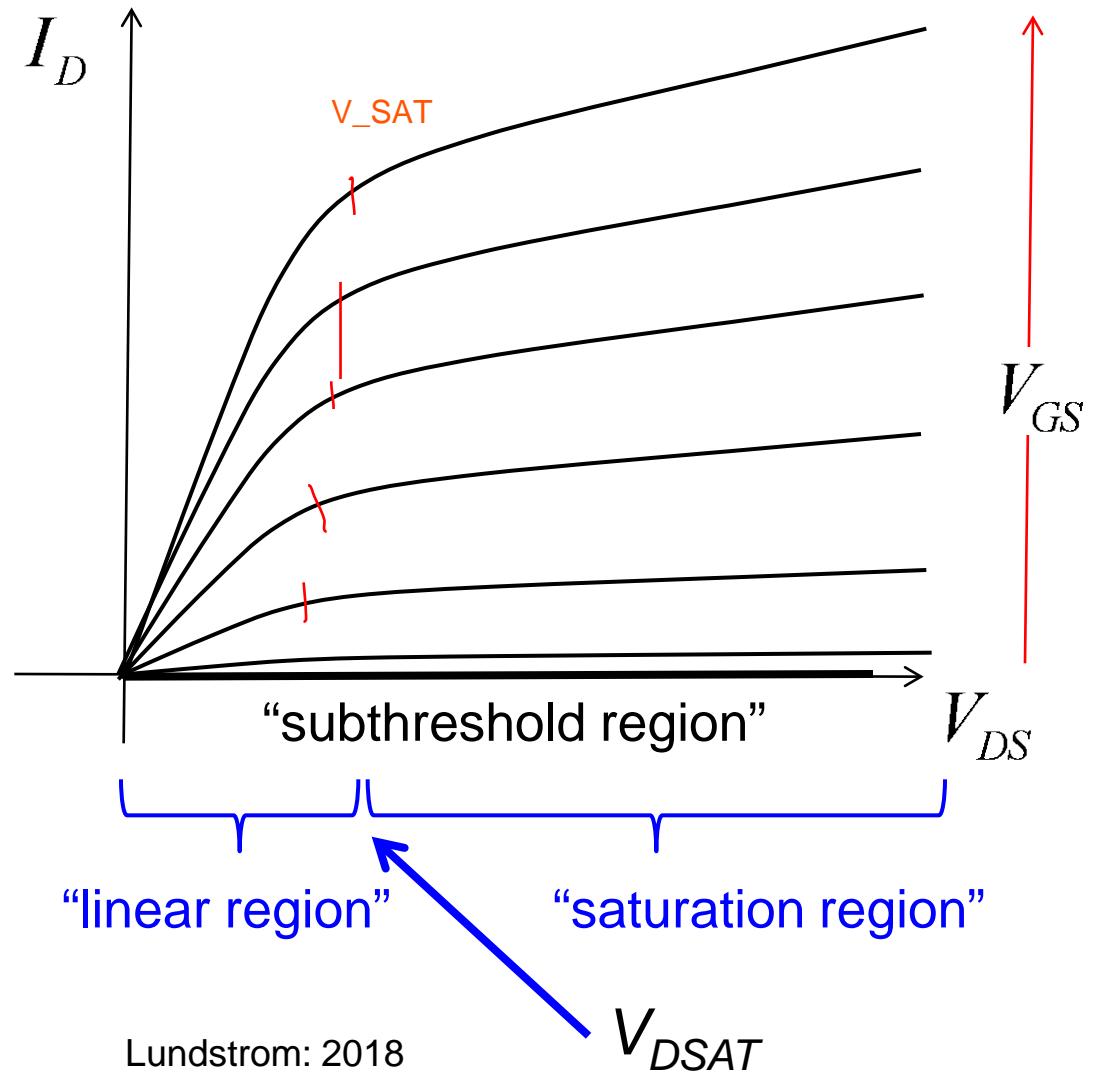


n-channel
enhancement
mode MOSFET

MOSFET IV: output characteristics

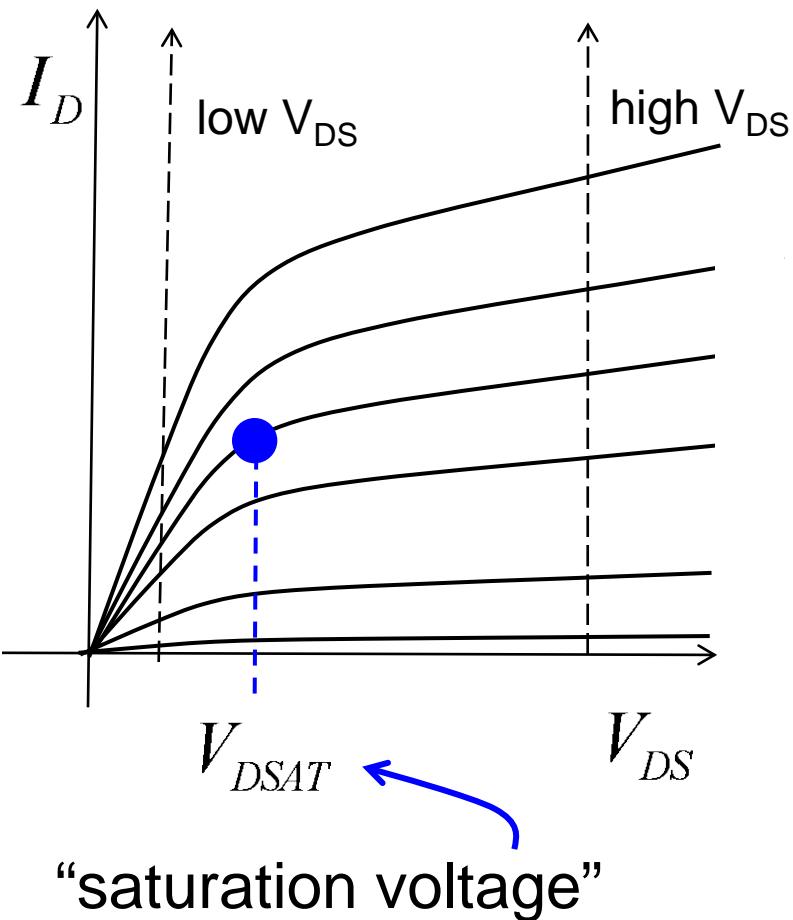


n-channel
enhancement
mode MOSFET

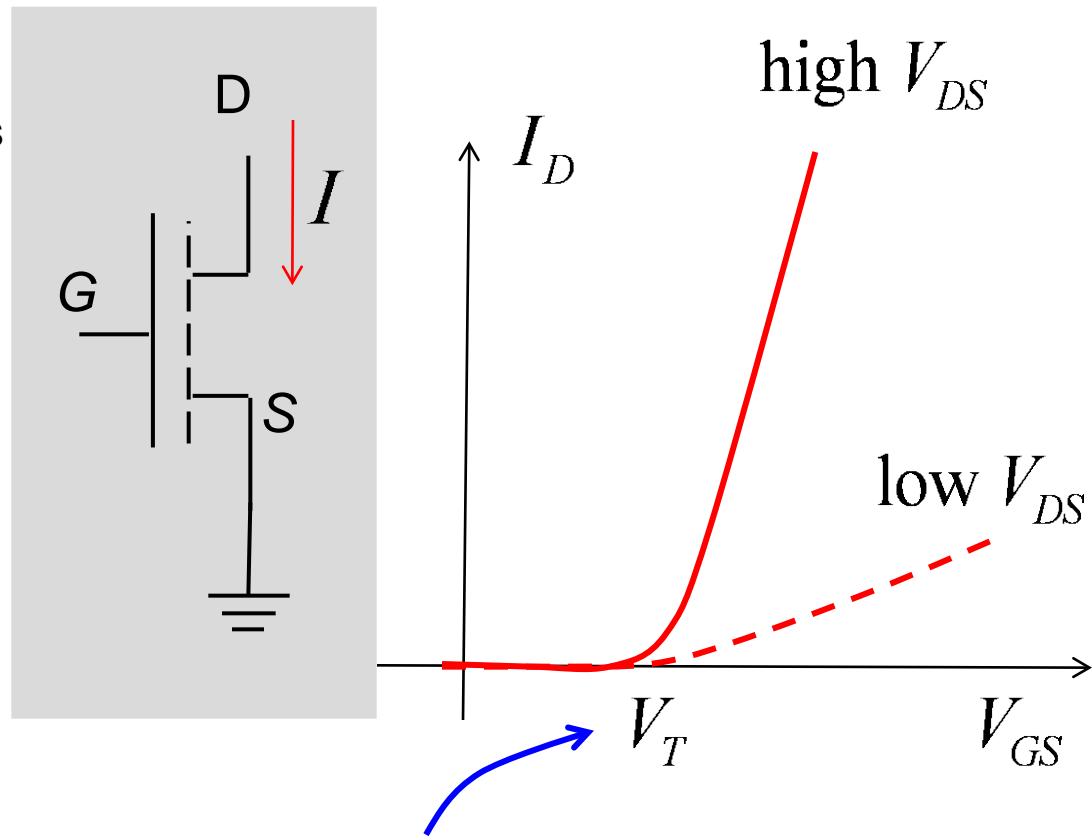


Output vs. transfer characteristics

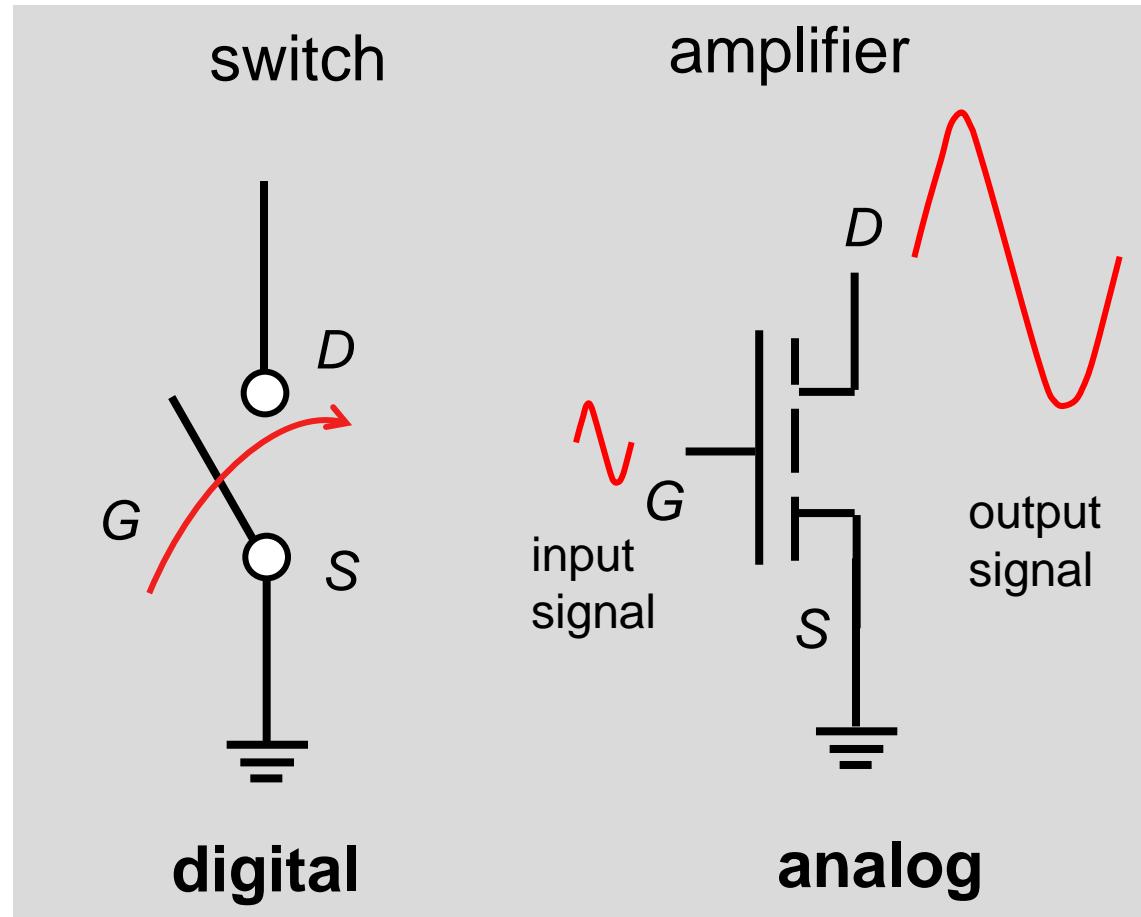
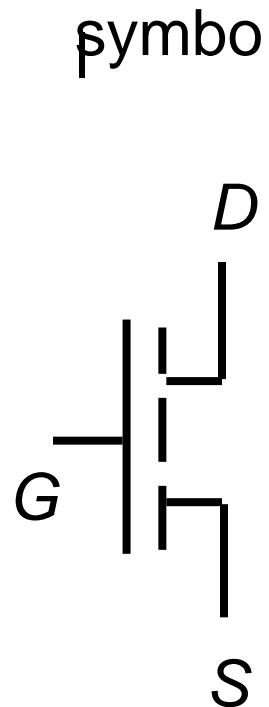
output characteristics



transfer characteristics

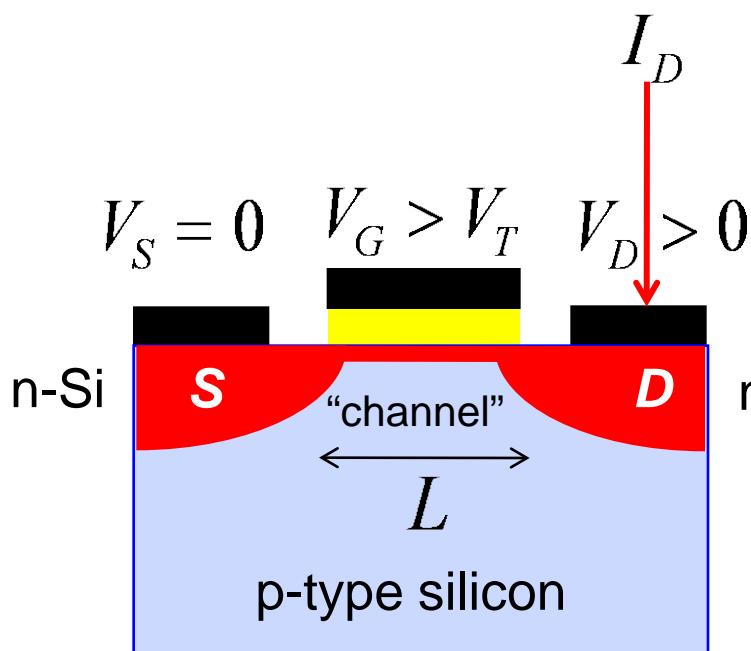


Applications of MOSFETs

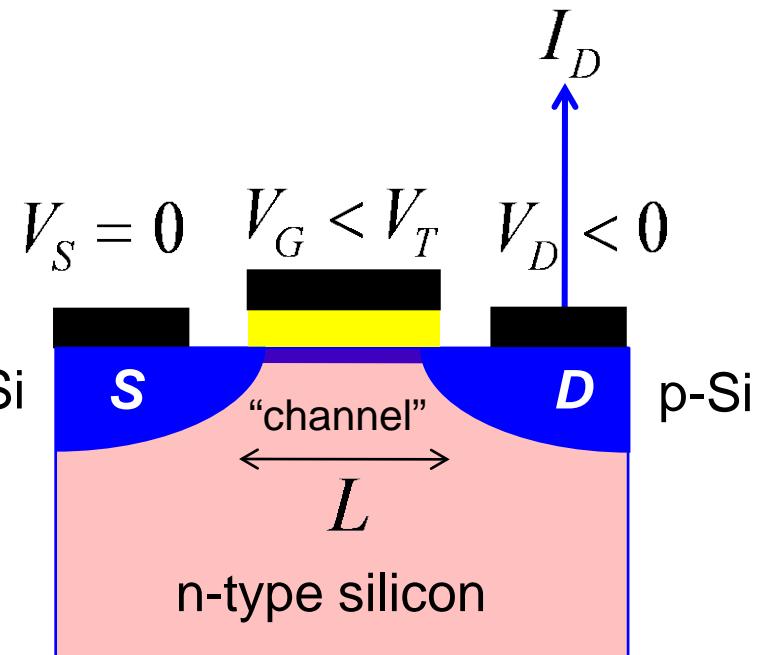


N-channel vs. P-channel MOSFETs

N-MOSFET



P-MOSFET



side view

side view

Summary

- 1) Transistors are three (or sometime four) terminal devices that control a large output current with an input voltage (or sometimes with a small input current).
- 2) Transistors can operate as a voltage controlled resistor or as a voltage controlled current source.
- 3) The shape of the IV characteristics make transistors useful in digital and analog circuits.
- 4) The shape of the IV characteristics is determined by the physics of the transistor.

Next topic: A primer on digital circuits

Device engineers assess MOSFETs in terms of a few key **device metrics**.

To understand these device metrics, we must first understand a little about digital and analog circuits.

Essentials of MOSFETs

Unit 1: Transistors and Circuits

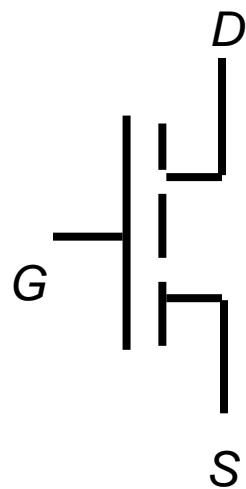
Lecture 1.2: Digital Circuits

Mark Lundstrom

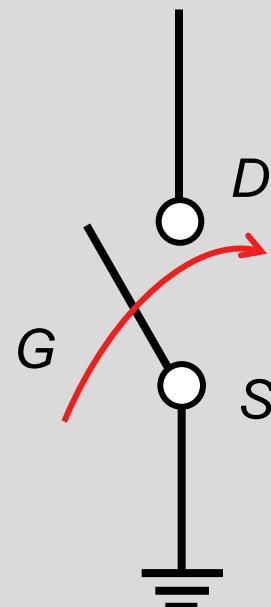
lundstro@purdue.edu
Electrical and Computer Engineering
Purdue University
West Lafayette, Indiana USA

Applications of MOSFETs

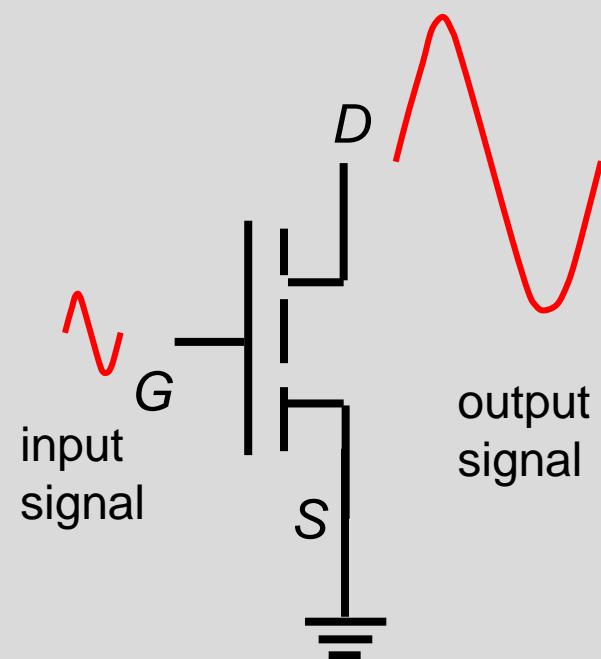
symbol



digital

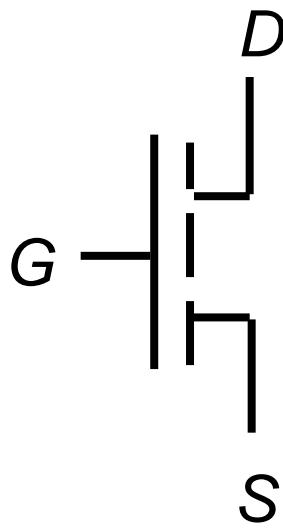


analog



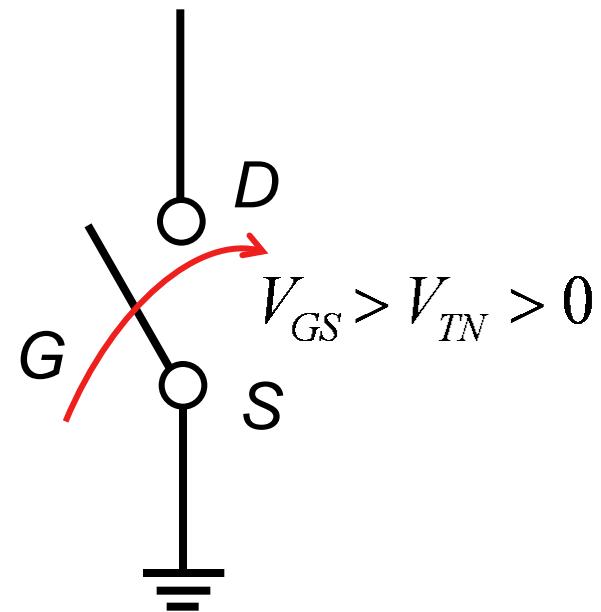
N-MOSFET as a switch

N-MOSFET
symbol



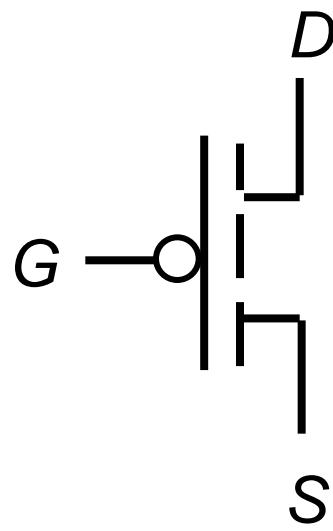
switch

$$V_{DS} > 0$$



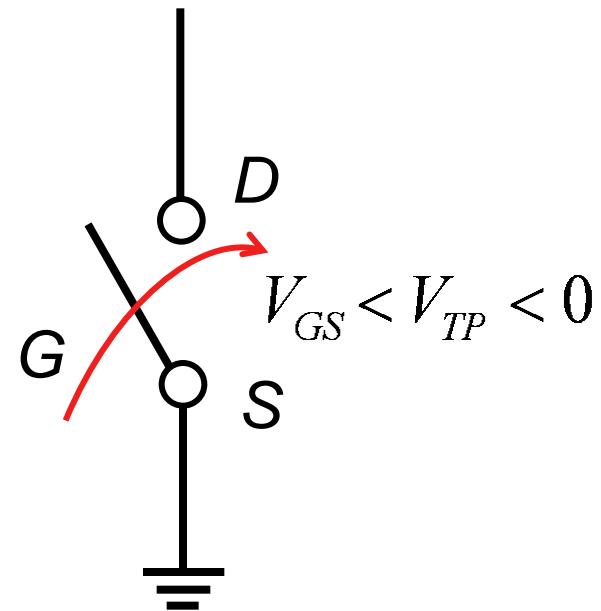
P-MOSFET as a switch

P-MOSFET
symbol

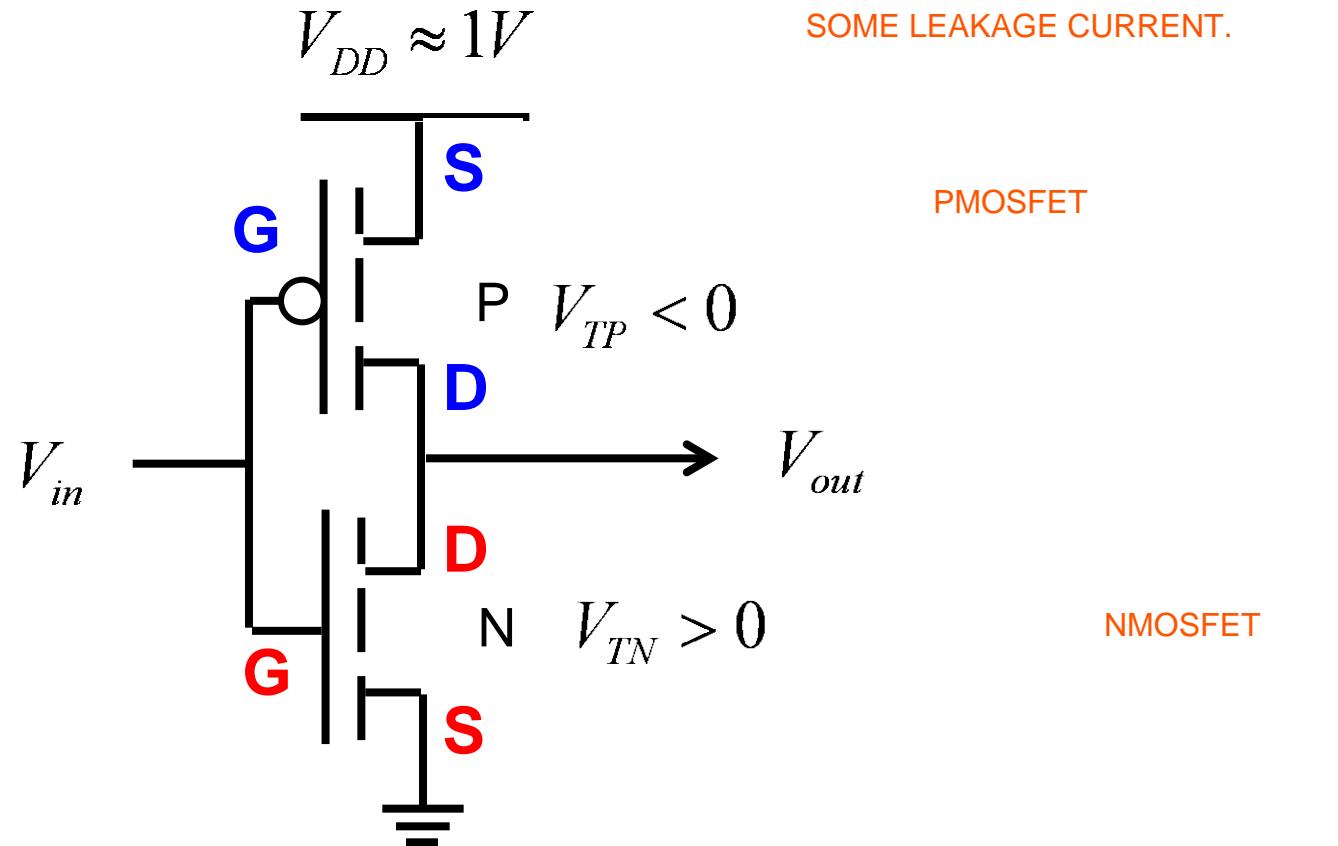


switch

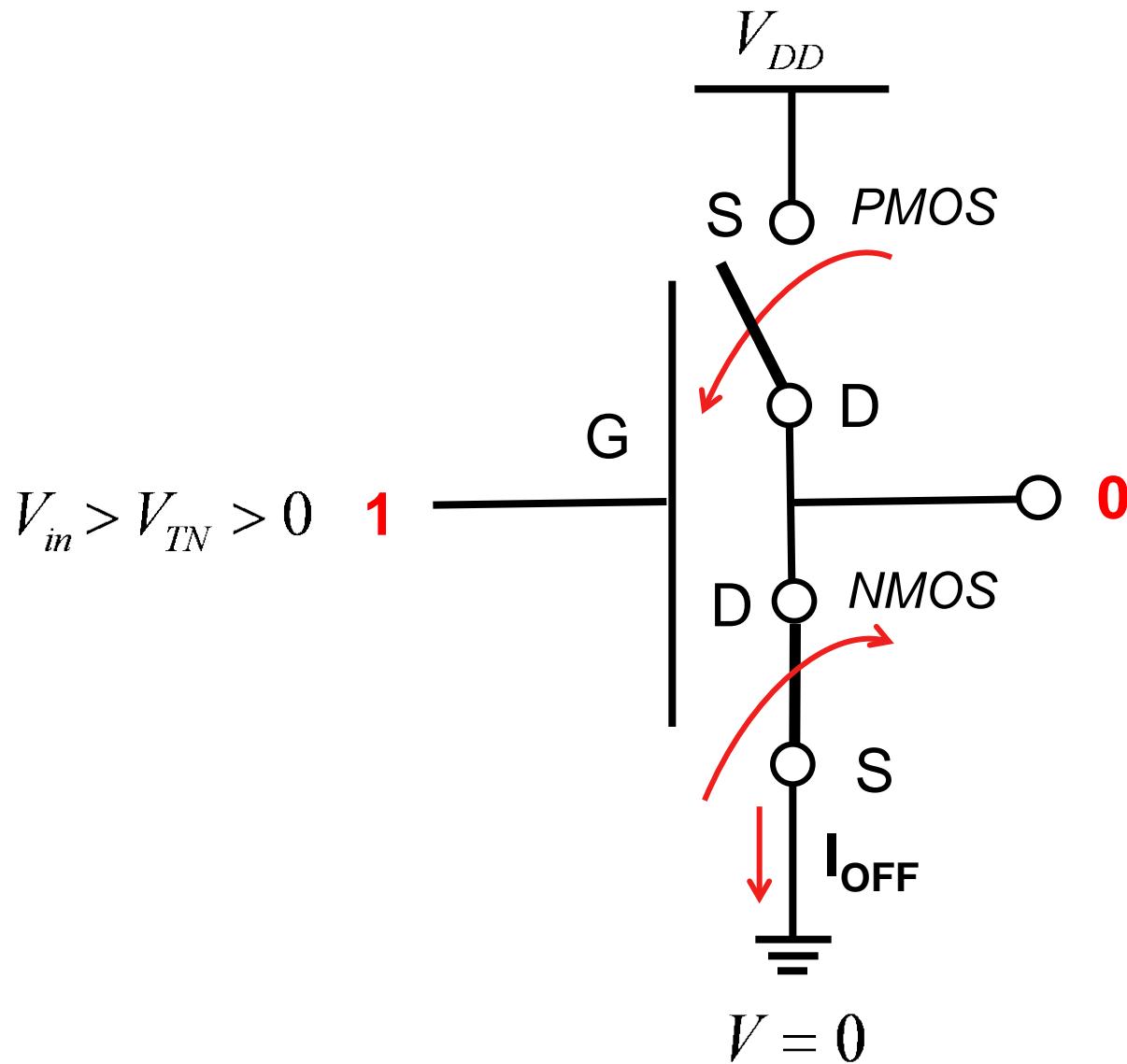
$$V_{DS} < 0$$



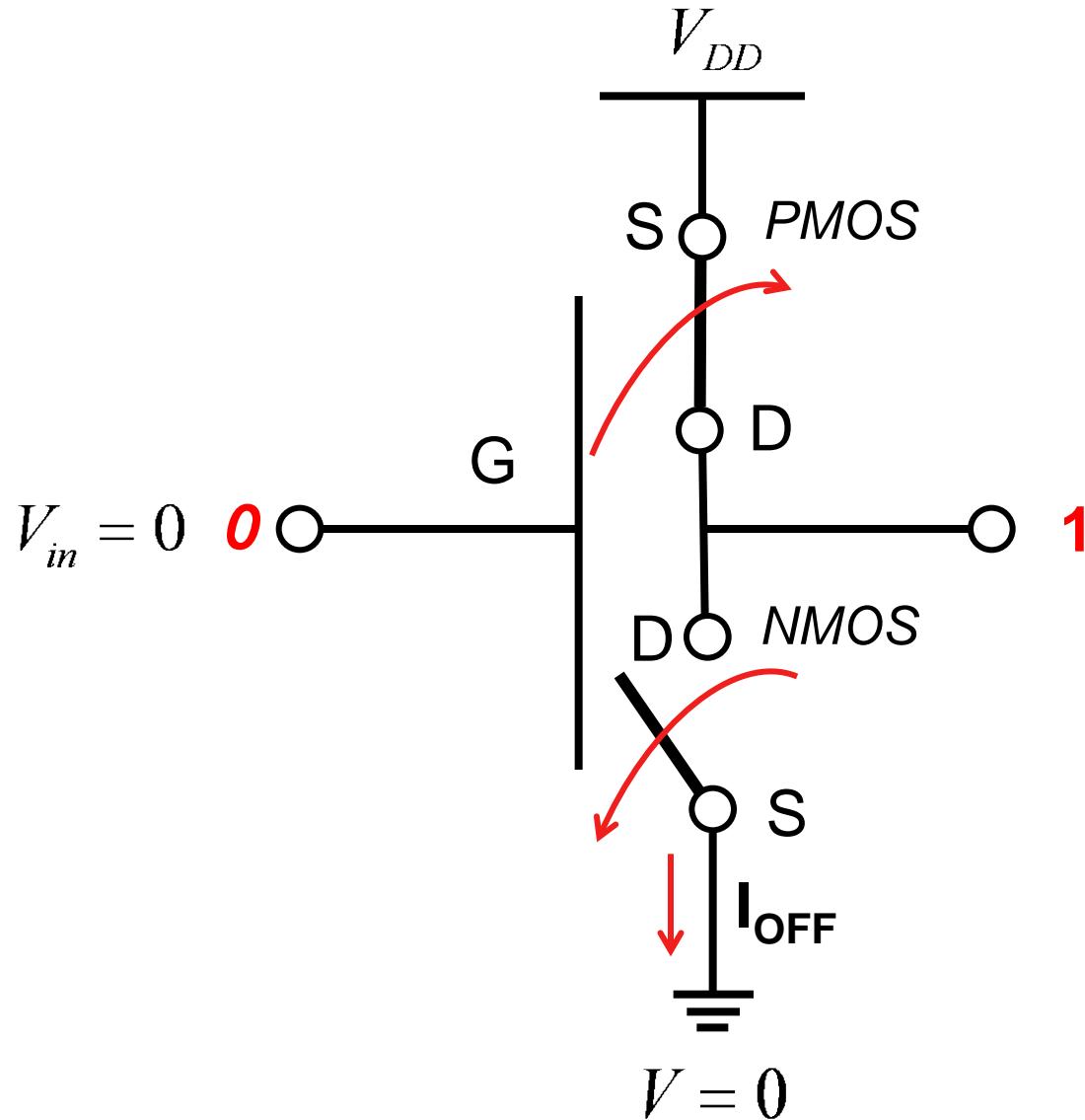
CMOS Inverter



CMOS inverter

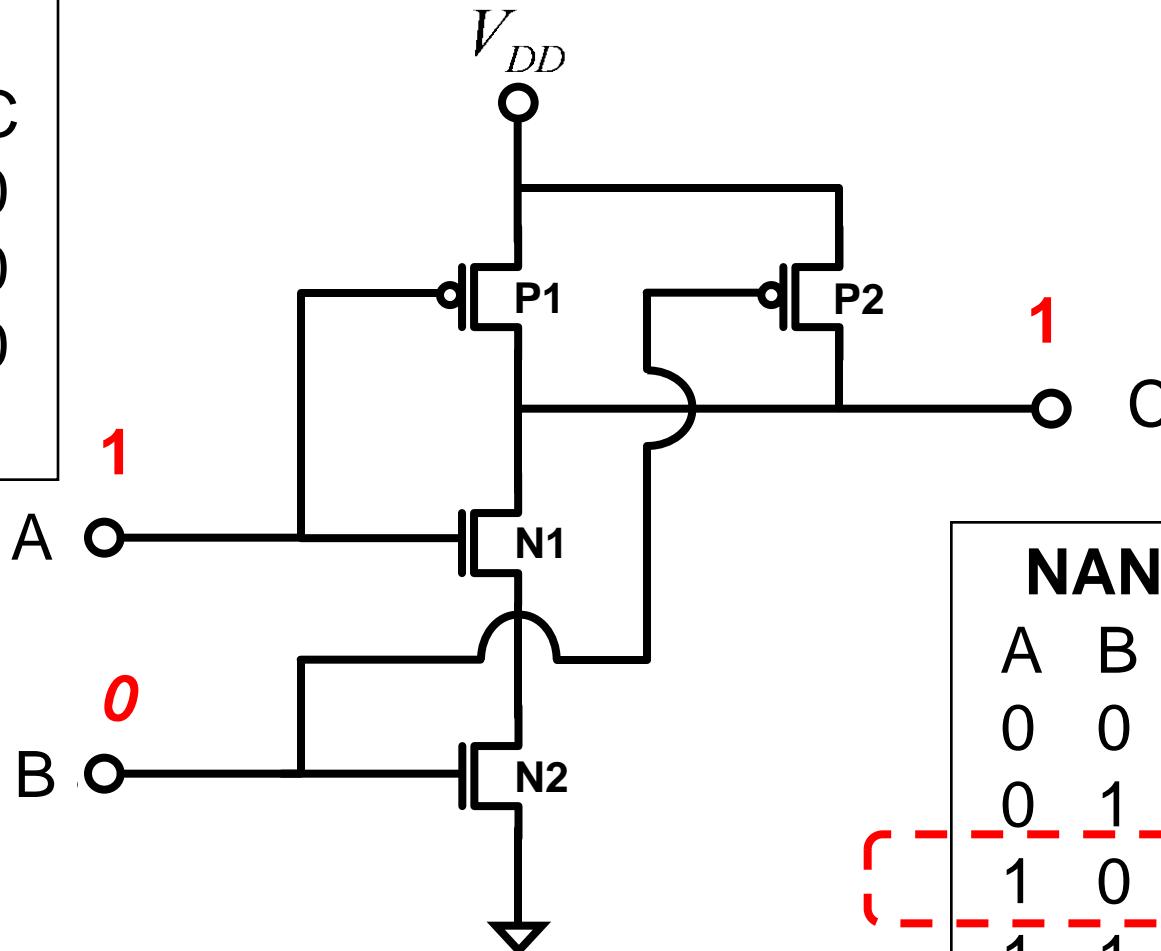


CMOS inverter



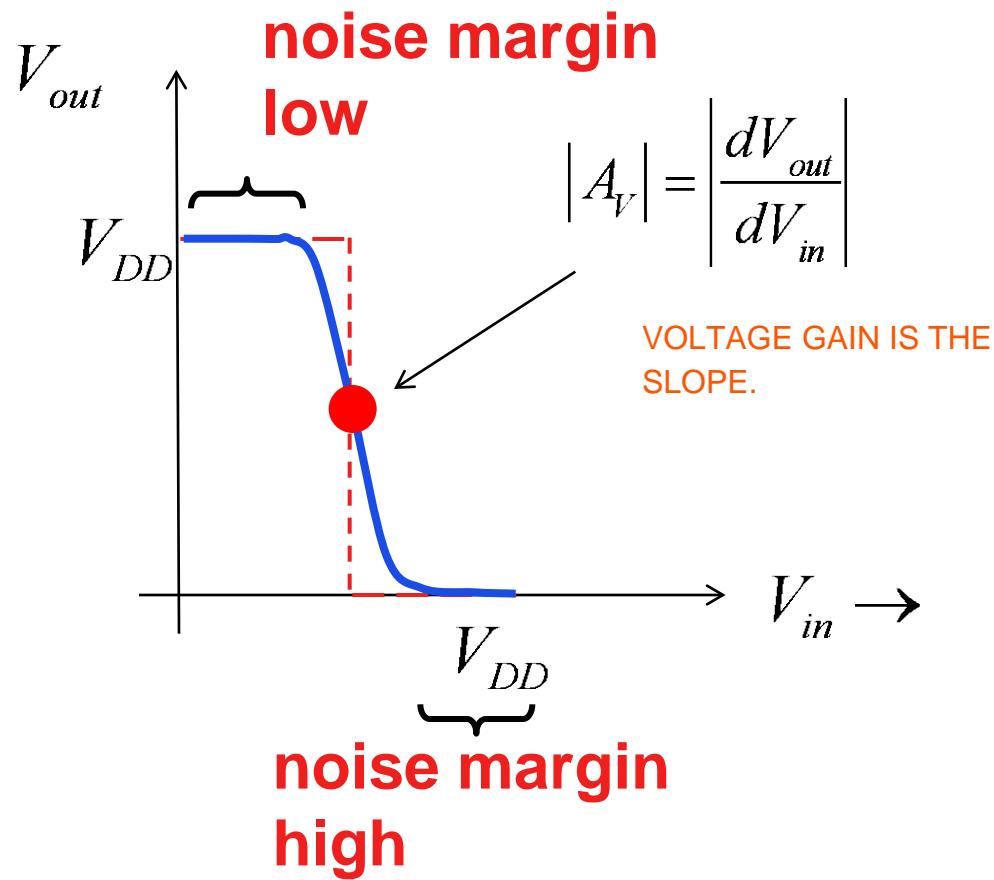
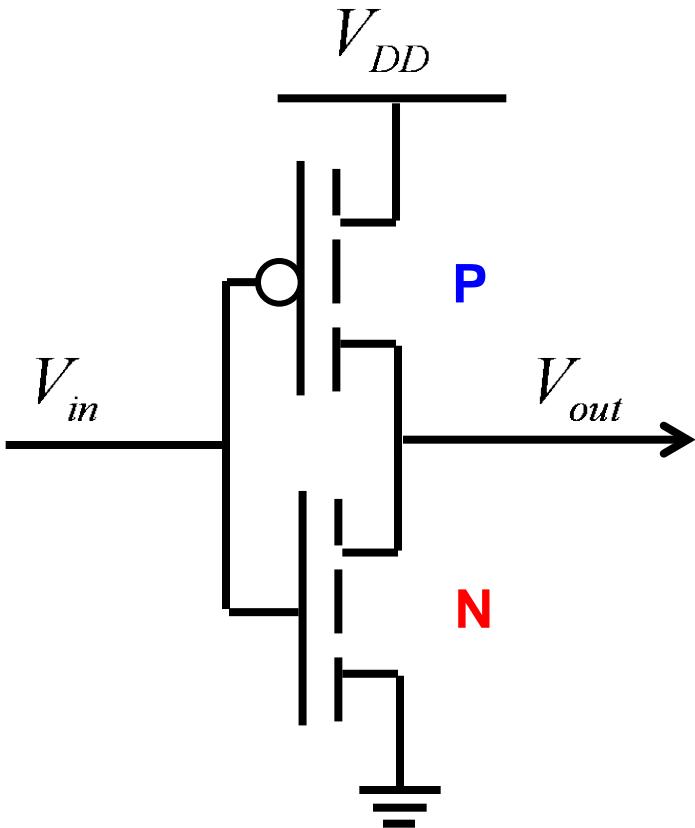
Two input NAND gate

AND		
A	B	C
0	0	0
0	1	0
1	0	0
1	1	1



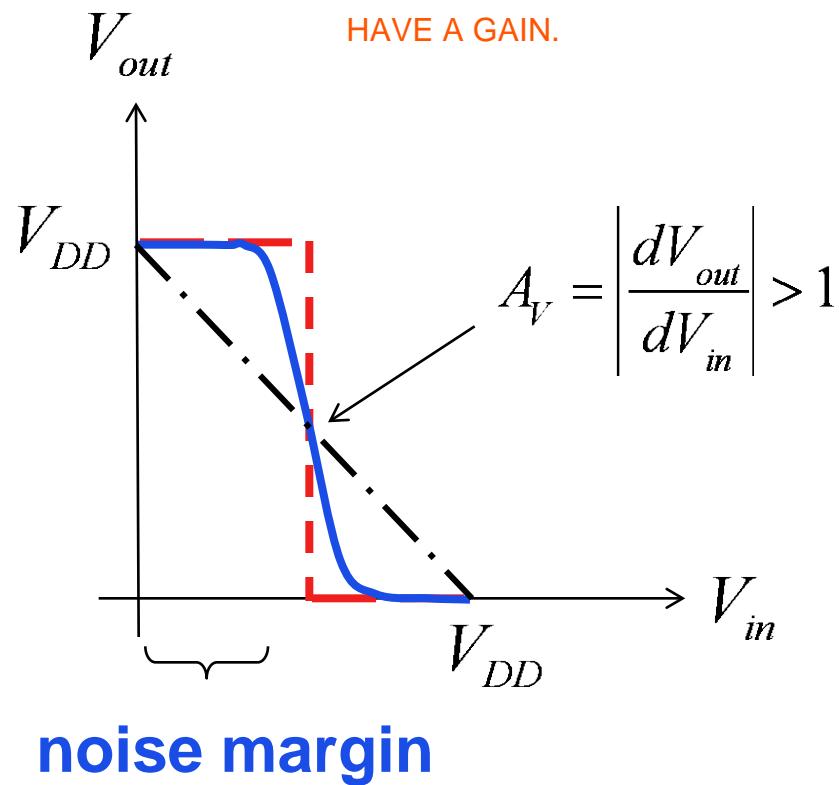
NAND		
A	B	C
0	0	1
0	1	1
1	0	1
1	1	0

Transfer characteristics



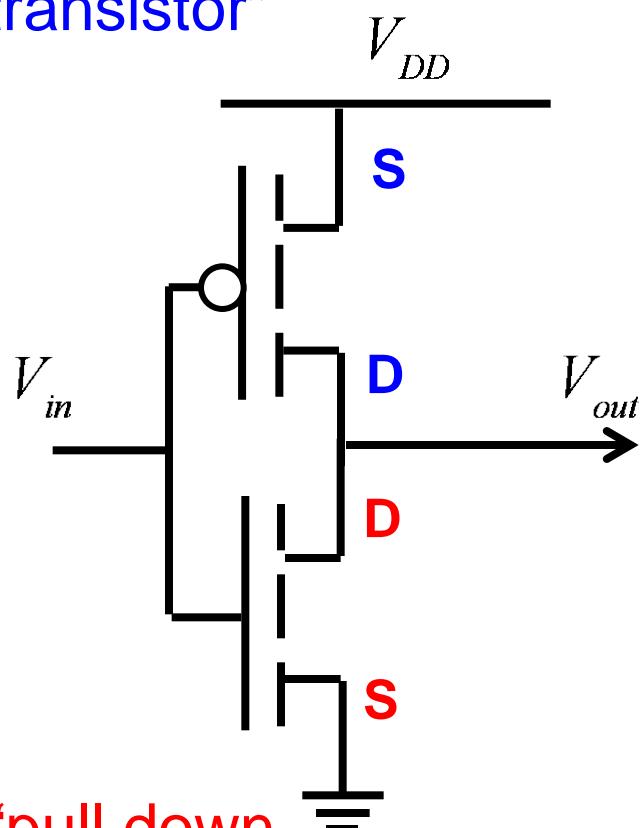
Gain restores signal levels

IF GAIN IS EQUAL TO EXACTLY ONE WE CANT ABLE TO RESIST THE NOISE OR ERROR MARGINS, SO WE NEED TO HAVE A GAIN.



CMOS inverter

“pull up
transistor”

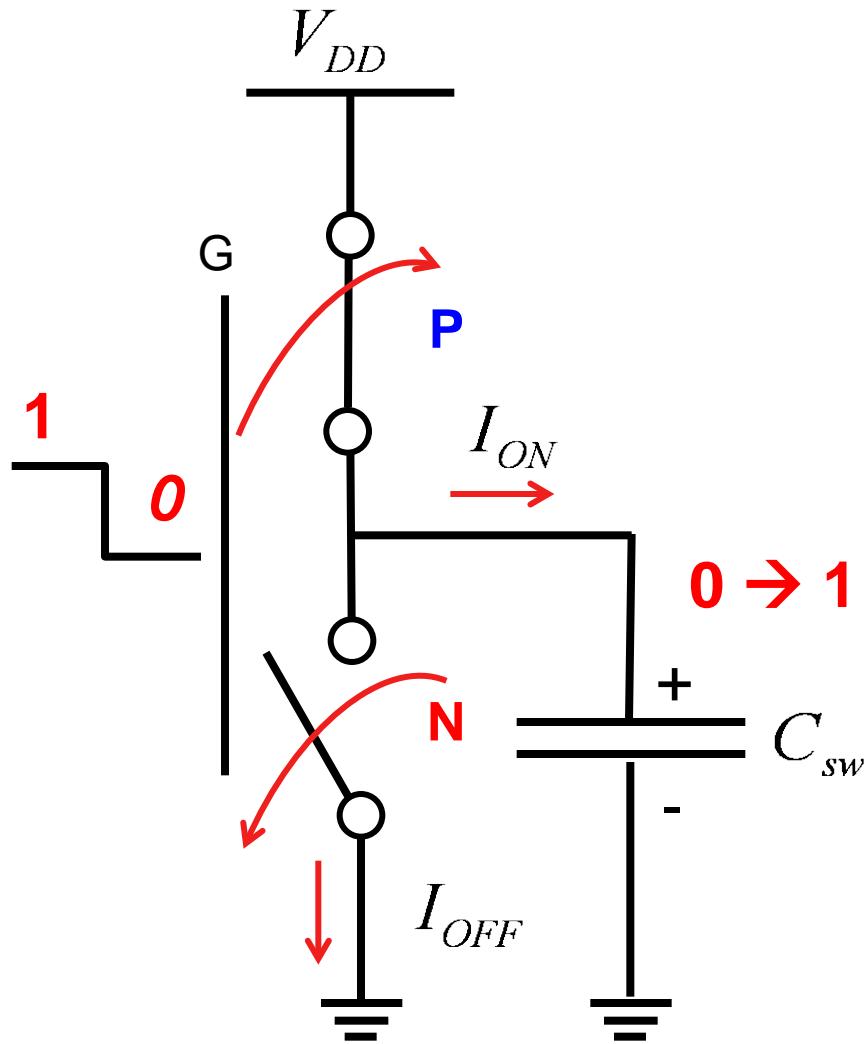


“pull down
transistor”

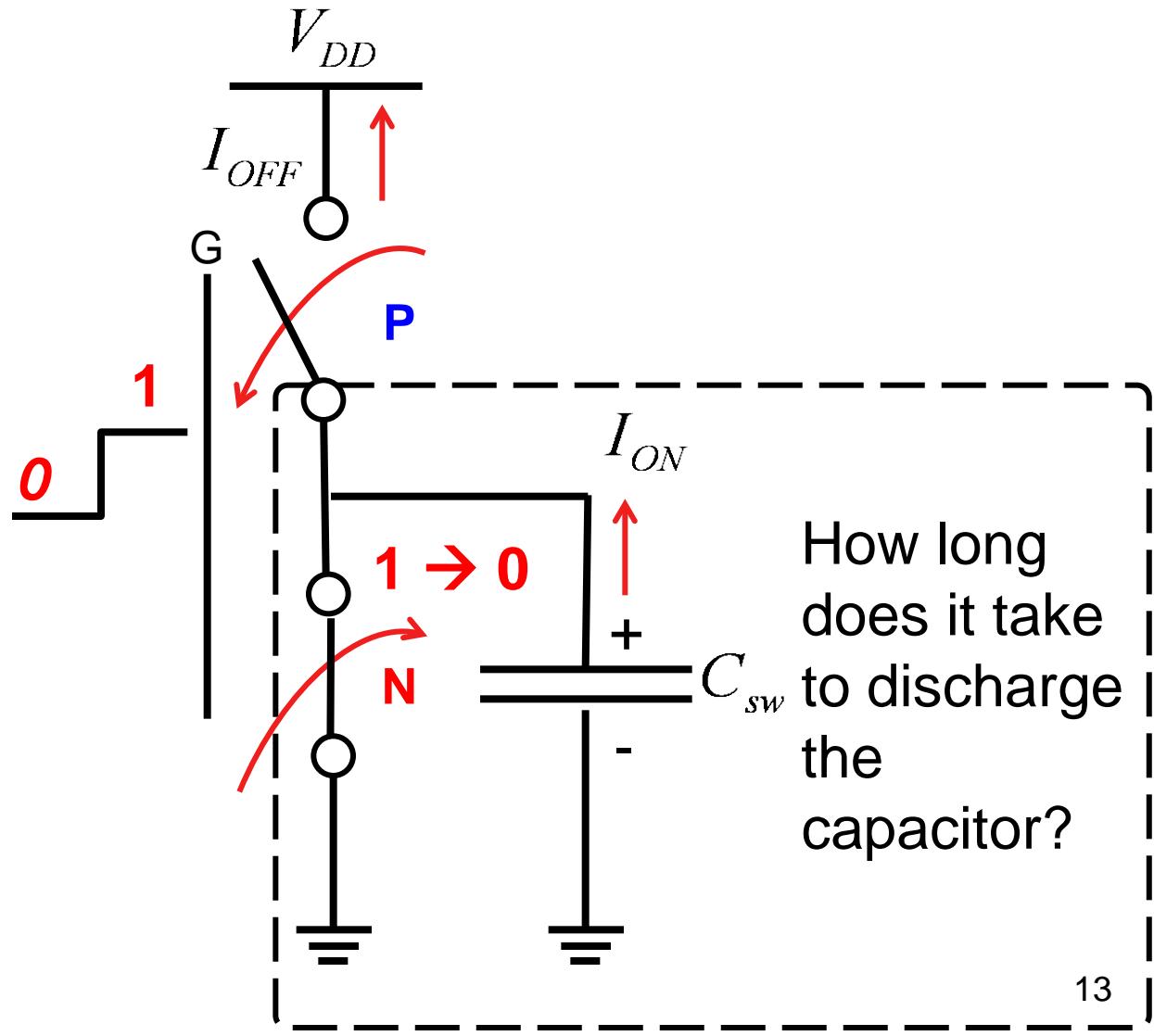
- 1) little current flows (power dissipation) unless switching
- 2) good noise margins if device has voltage gain

What determines **speed and power**?

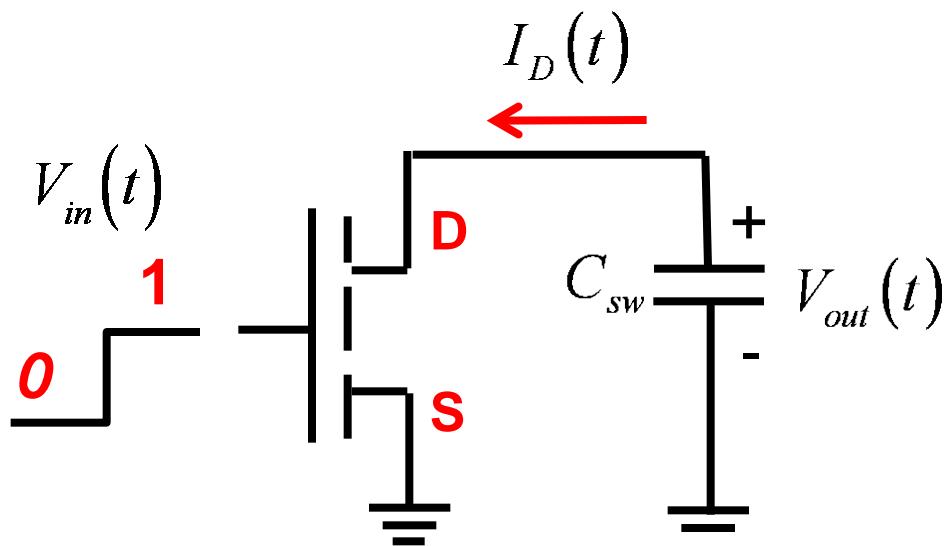
Switching (input 1 to 0)



Switching (input 0 to 1)



Discharging time



$$Q = C_{sw} V_{DD}$$

SINCE IT IS CHARGED
BY V_{DD} .

$$I_{ON} = \frac{Q}{\tau}$$

CURRENT=CHARGE/TIME

$$\tau = \frac{C_{sw} V_{DD}}{I_{ON}}$$

DC on-current controls switching time.

System speed

until ~1990:

device delay > interconnect delay

90 nm technology:

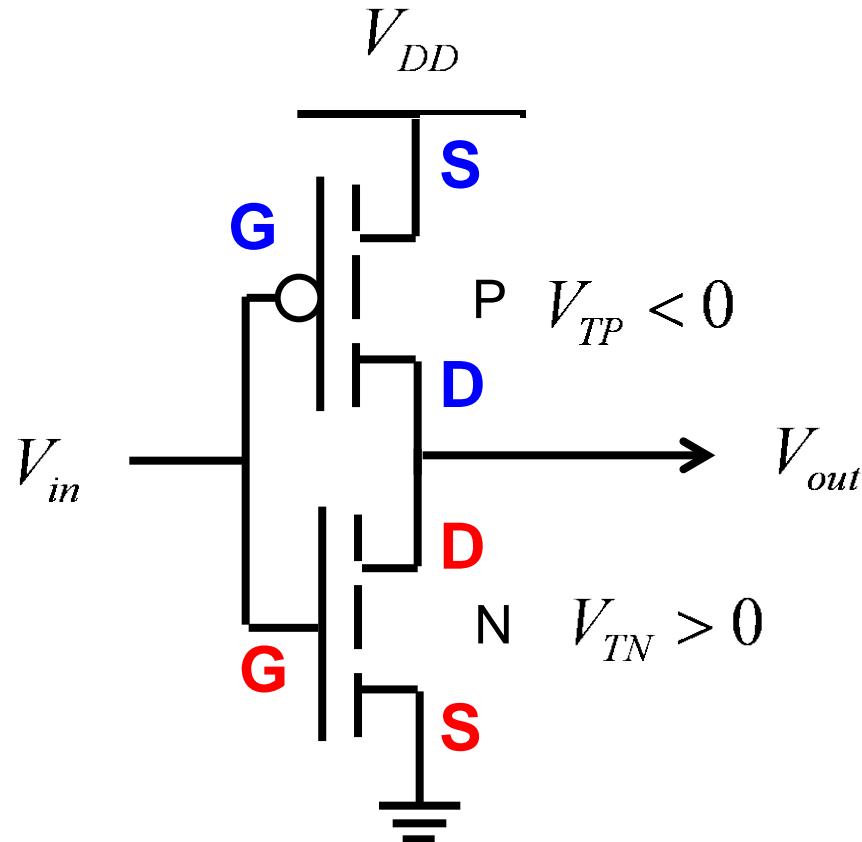
device delay: ~ 1ps
1 mm interconnect delay: ~ 6 ps

2015 (10 nm technology):

device delay: ~0.1ps
1 mm interconnect: ~30ps

J. Meindl, “Beyond Moore’s law: the interconnect era,” *Computing in Science and Engineering*, 2003

Power dissipation



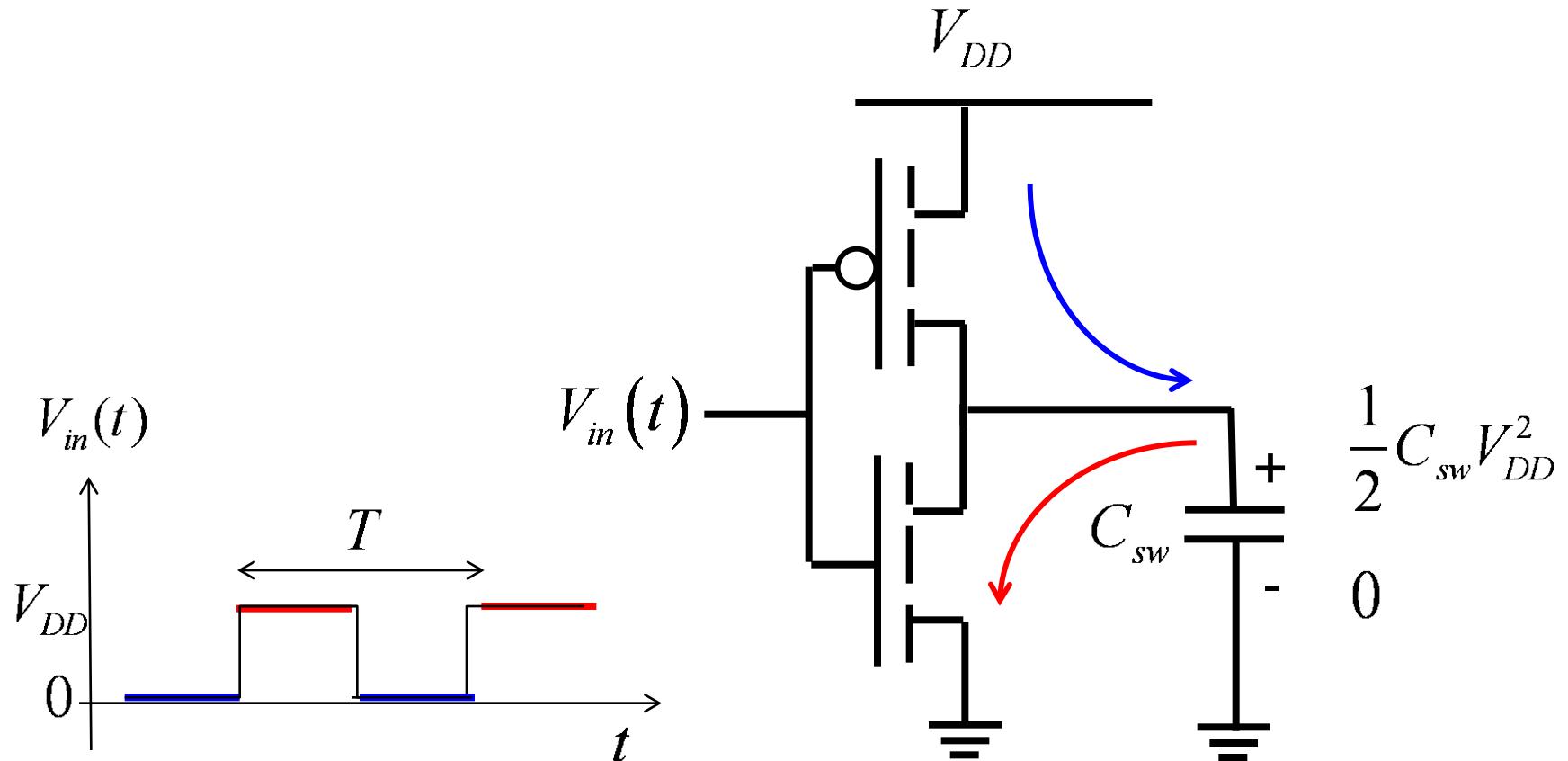
Switching speed is inversely proportional to the DC on-current.

The ideal CMOS inverter only dissipates power while switching (**dynamic power**). IDEALLY NO LEAKAGE CURRENT

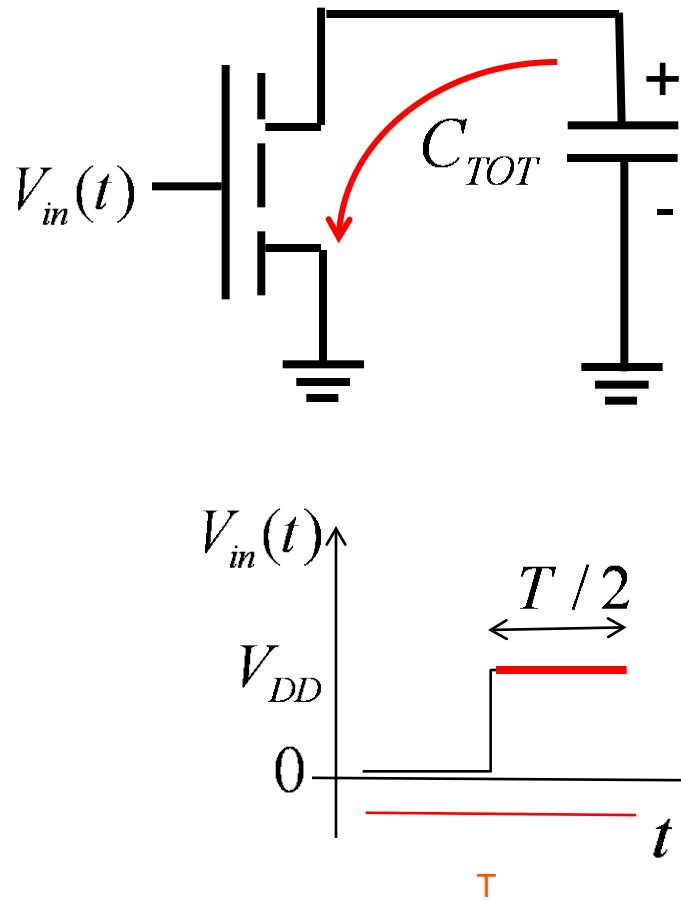
Real CMOS also dissipates power when not switching (**static power**).

DUE TO LEAKAGE CURRENT

Dynamic power



Dynamic power



$$E_C(0) = \frac{1}{2} C_{sw} V_{DD}^2$$

$$E_C(T/2) = 0$$

$$P_{dynamic} = \frac{\Delta E}{T/2} = \frac{C_{sw} V_{DD}^2}{T} = f C_{sw} V_{DD}^2$$

$$P_{dynamic} = \alpha f C_{sw} V_{DD}^2$$

switching “activity”

DYNAMIC POWER IS A POWER CONSUMED WHILE THE INPUTS ARE ACTIVE.

STATIC POWER IS THE POWER CONSUMED WHILE THERE WAS NO CURRENT ACTIVITY.

Dynamic power

$$P_{dynamic} = \alpha f C_{sw} V_{DD}^2$$

switching “activity”

At a given frequency, the dynamic power is proportional to **the power supply voltage squared** and to the **frequency**.

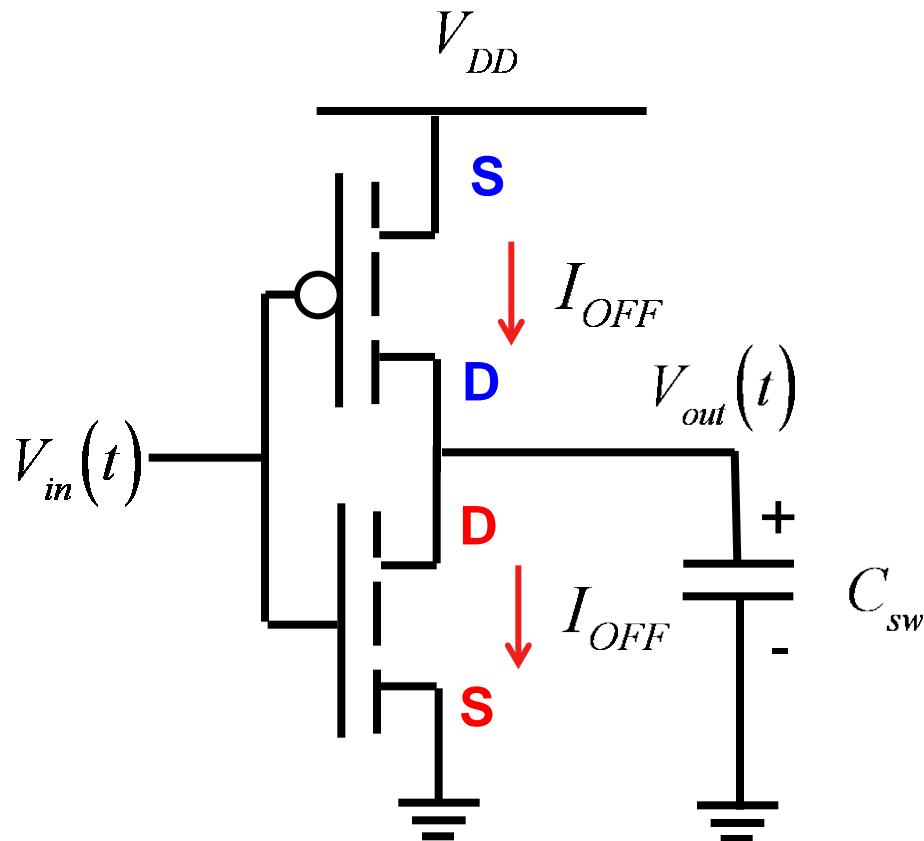
TRANSISTORS WHICH USES SMALL AMOUNT OF VOLTAGE IS APPRECIABLE.

What determines the **static power**?

AS TRANSISTORS BECOME SMALLER AND SMALLER OVER THE YEAR, THESE LEAKAGE CURRENTS ALSO GOT INCREASED. THEY ARE STILL VERY SMALL WHEN WE SEE A SINGLE TRANSISTOR BUT WHEN IT COMES TO A BILLION TRANSISTORS ITS

Static power

HUGE.



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

N_G = no. of gates

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

CMOS speed and power

$$\tau = C_{sw} V_{DD} / I_{ON}$$

$$P_{\text{dynamic}} = \alpha f C_{sw} V_{DD}^2$$

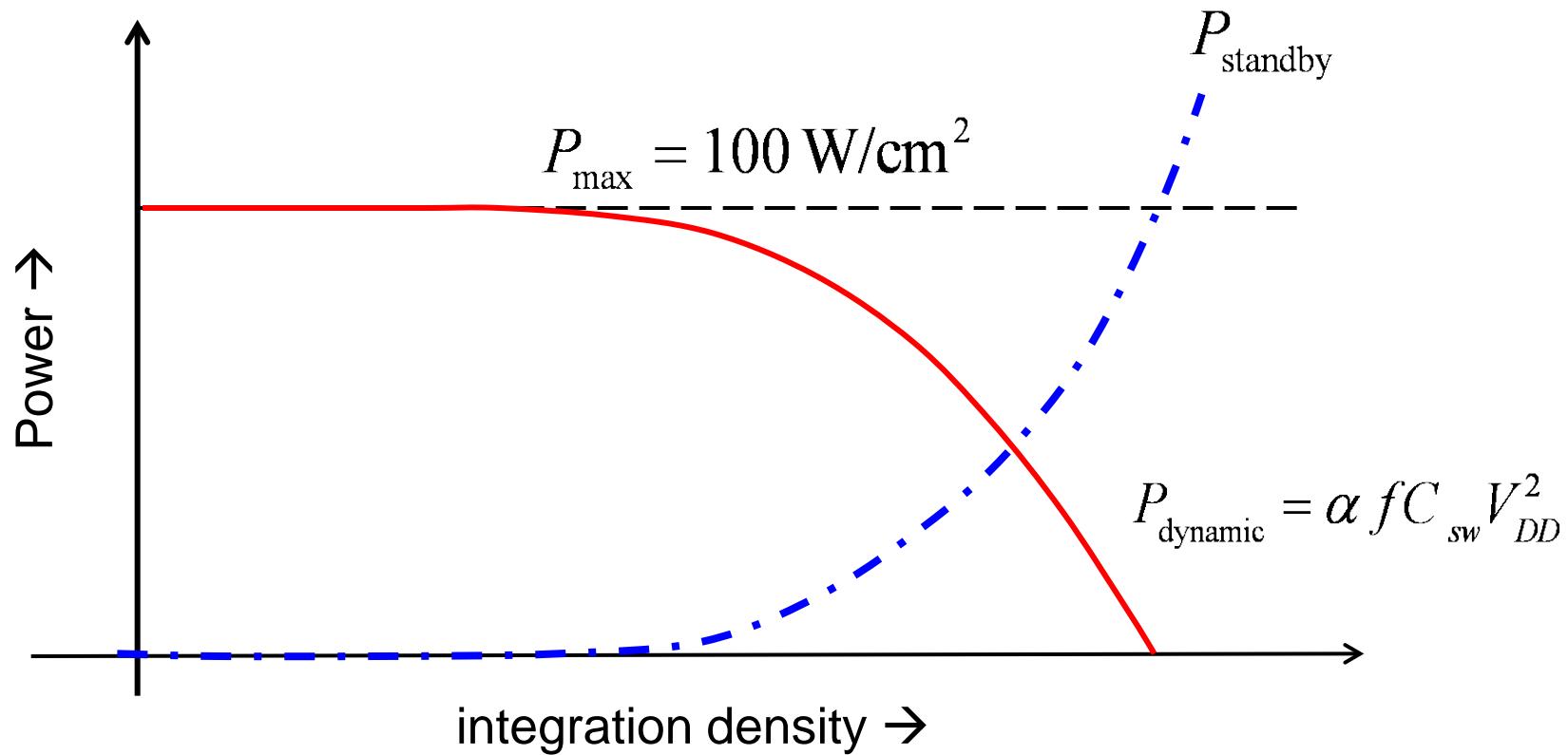
$$P_{\text{static}} = N_G I_{OFF} V_{DD}$$

- 1) Higher on-current means higher speed
- 2) Faster operation means more dynamic power
- 3) Lower V_{DD} means lower power
- 4) More leakage means more power dissipation

MOST ELECTRONIC DEVICES ARE POWER CONSTRAINED. WE HAVE A POWER BUDGET. WE HAVE POWER DISSIPATION LIMIT.

Power constrained design

SUM OF DYNAMIC AND STATIC MUST BE INNER MY BUDGET. BECAUSE TOTAL POWER CONSUMPTION IS THEIR SUM.



(after Dave Frank, IBM)

22

Summary

Complementary MOS (CMOS) makes use of NMOS and PMOS transistors.

The basic building block of CMOS logic is the CMOS inverter.

Voltage gain is required for noise margins.

On-current, off-current, and power supply voltage are key parameters.

Next topic: A primer on analog circuits

We now understand what's important for digital circuits. In the next lecture, we'll take a quick look **at analog and radio frequency** circuits.

Essentials of MOSFETs

Unit 1: Transistors and Circuits

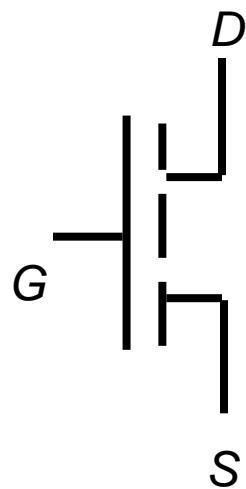
Lecture 1.3: Analog/RF Circuits

Mark Lundstrom

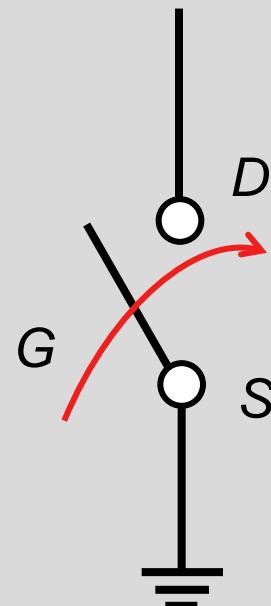
lundstro@purdue.edu
Electrical and Computer Engineering
Purdue University
West Lafayette, Indiana USA

Applications of MOSFETs

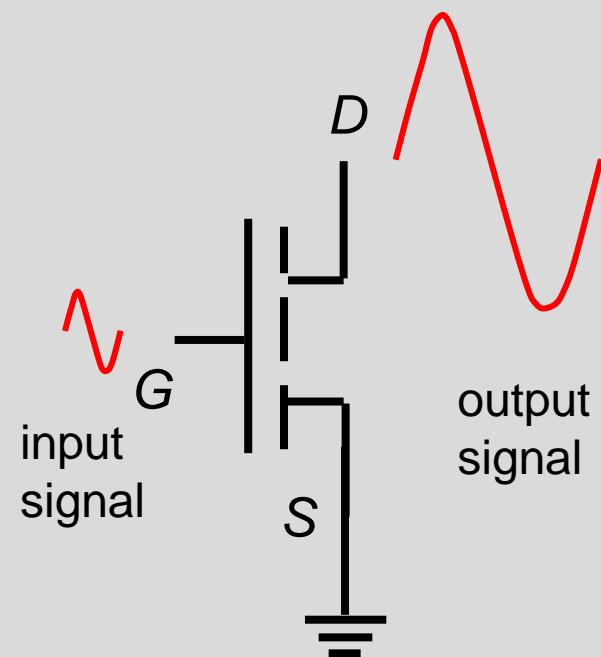
symbol



digital



analog



Why analog /RF? Why CMOS?

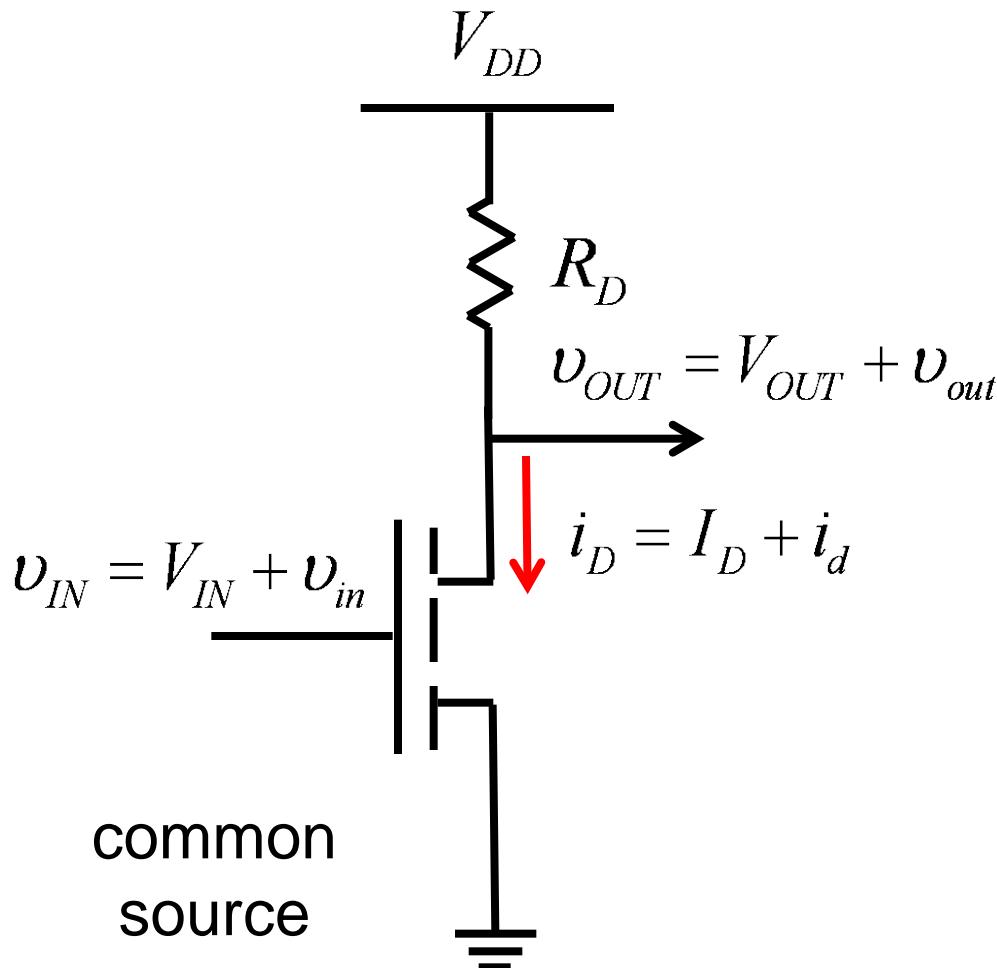
Many applications involve analog / RF signals:

- 1) many natural signals are analog (sensors)
- 2) disk drive electronics
- 3) wireless receivers and transmitters
- 4) optical receivers
- 5) microprocessors / memories

CMOS: IF WE ARE GONNA USE ANALOG ALONE- BJT IS MORE SUITABLE.

- 1) many systems are both analog and digital
- 2) CMOS dominant for digital electronics
- 3) CMOS performance is acceptable for many analog applications (but not, generally as good as bipolar)

Basic CS Amplifier



$$I_D = f(v_{GS}, v_{DS})$$

$$i_d = g_m v_{gs} \quad \begin{matrix} \text{PROPORTIONALITY CONSTANT} \\ \text{TRANSCONDUCTANCE.} \end{matrix}$$

$$g_m = \frac{i_d}{v_{gs}} = \frac{\delta I_D}{\delta V_{GS}} \approx \left. \frac{\partial I_D}{\partial V_{GS}} \right|_{V_{DS}}$$

$$g_m \equiv \left. \frac{\partial I_D}{\partial V_{GS}} \right|_{V_{DS}}$$

QUASISTATIC APPROXIMATION

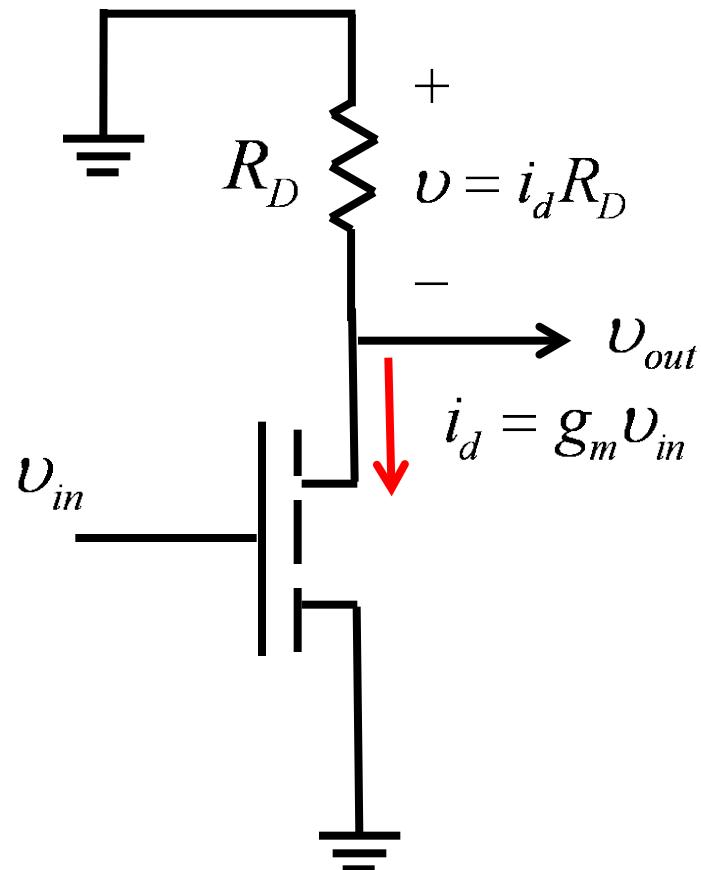
Lundstrom: 2018

WE CAN USE A DC BIAS TO DESCRIBE THE TIME VARYING SIGNAL AS LONG AS THE FREQUENCY IS NOT TOO FAST AS LONG AS THE DEVICE CAN KEEP UP WITH THE FREQUENCY.

BY SUPERPOSITION WE ARE SOLVING IT.

Basic CS Amplifier: ac analysis

AT SMALL SIGNAL ANALYSIS AC ACTS LIKE A GROUND.
ac ground



IT GOES THROUGH THE RESISTOR R_D

$$v_{out} = -i_d R_D = -g_m R_D v_{in}$$

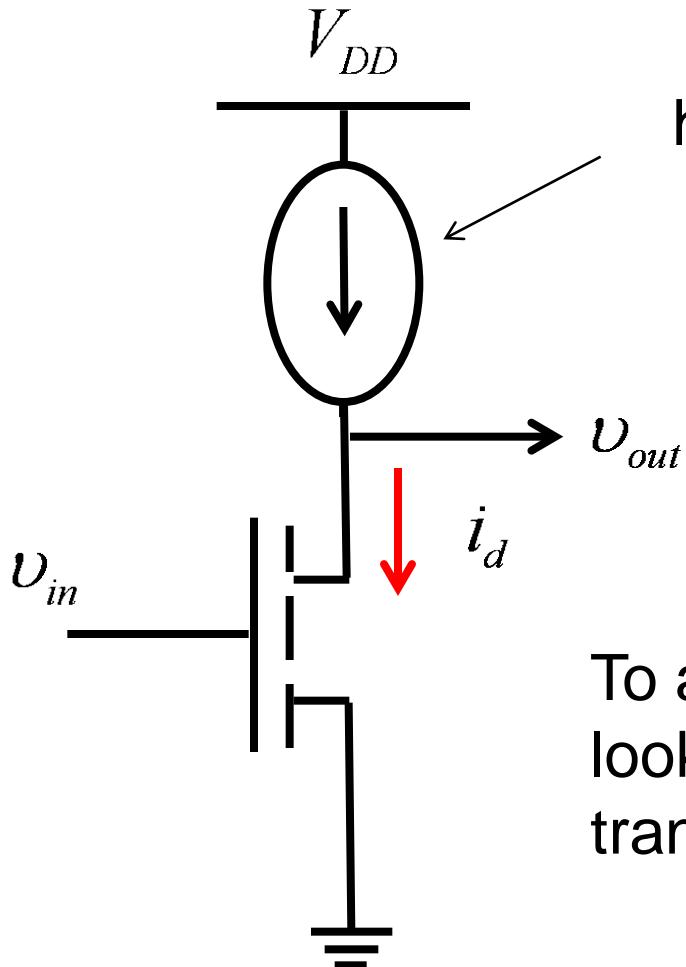
$$A_v = \frac{v_{out}}{v_{in}} = -g_m R_D$$

Transconductance is an important analog figure of merit.

It is a measure of the transistor's ability to amplify.

Maximum gain

(IDEAL CURRENT SOURCE WHICH HAS INFINITE HIGH RESISTANCE)



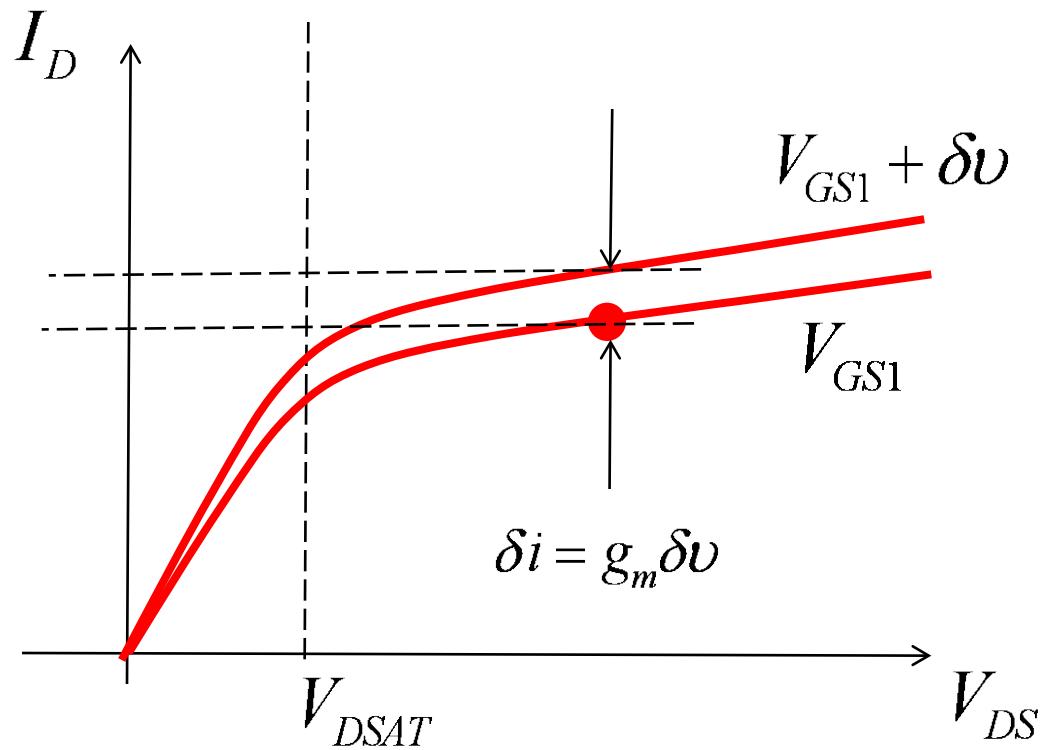
high impedance current source
DRAIN RESISTANCE IS ASSUMED AS INFINITE.

$$R_D \rightarrow \infty$$

$$A_v(\max) = ?$$

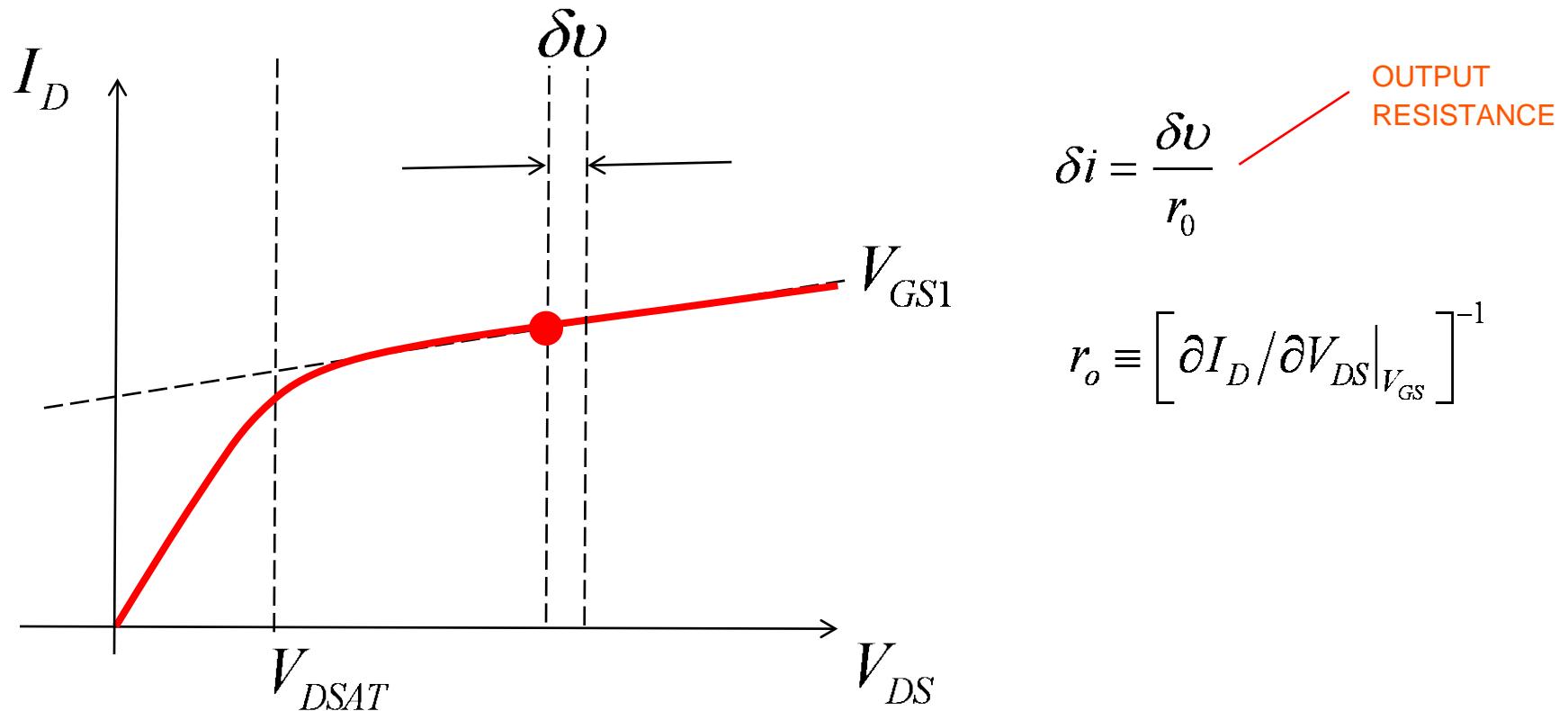
To answer this question, we need to look a little more carefully at the transistor.

Transconductance

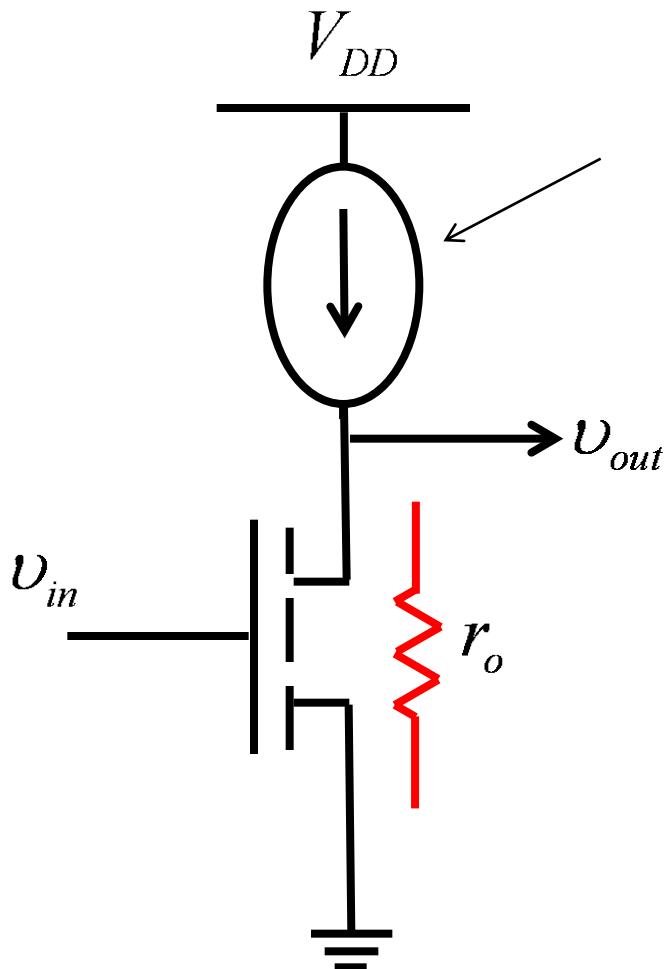


$$g_m \equiv \left. \frac{\partial I_D}{\partial V_{GS}} \right|_{V_{DS}}$$

Output resistance



Maximum gain



high impedance current source

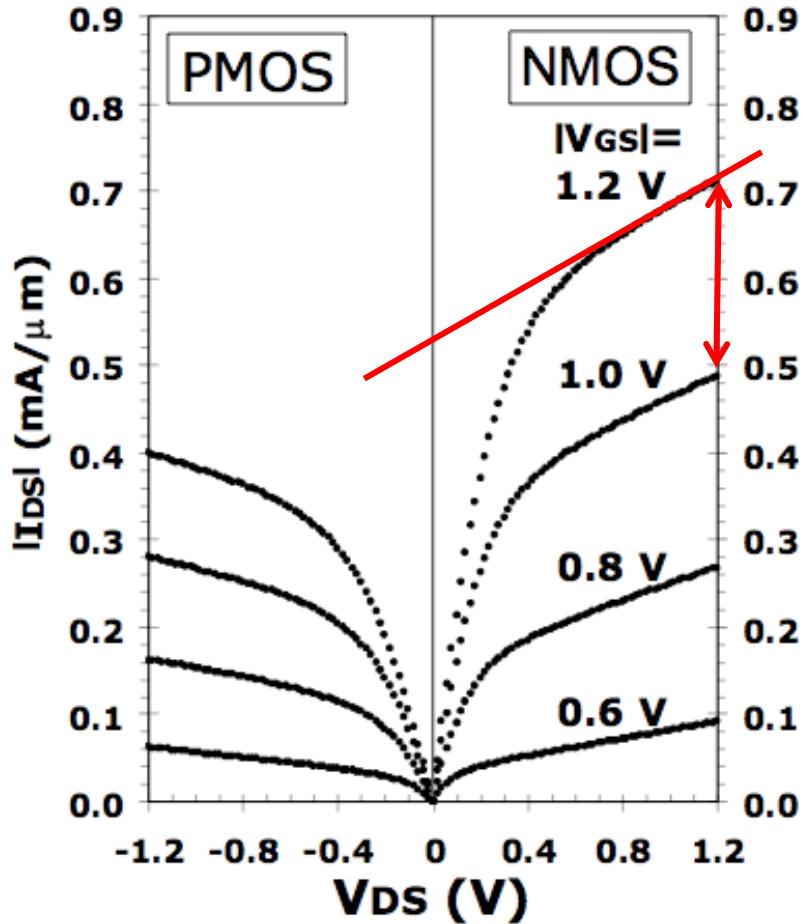
$$A_v(\max) = -g_m r_0$$

The **output resistance**, r_0 , of the MOSFET is an important figure of merit.

TRANSCONDUCTANCE

So is the **self gain**, $g_m r_0$.

Self-gain for 65 nm digital CMOS



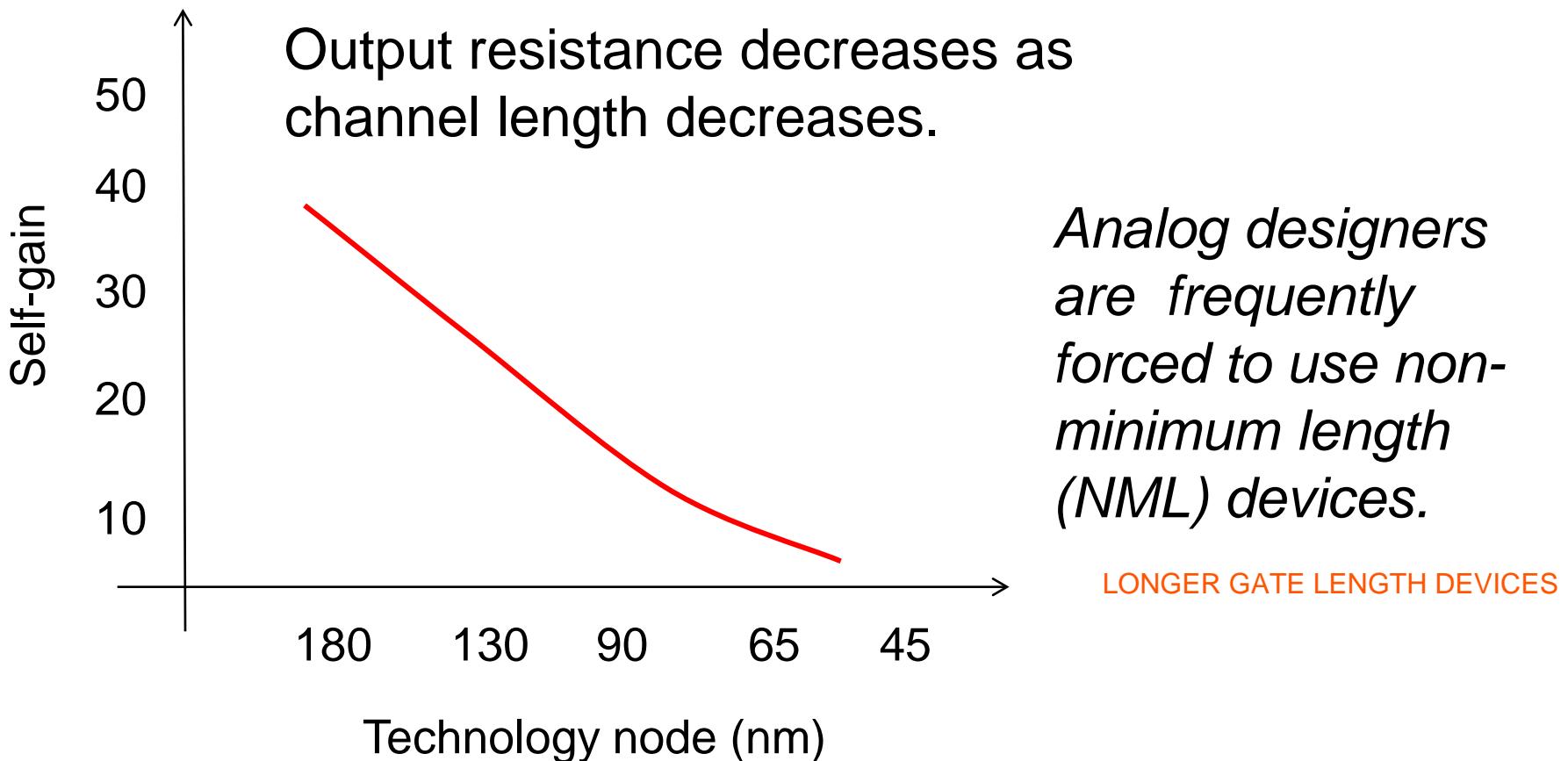
$$g_m \approx \frac{0.2 \text{ mA}/\mu\text{m}}{0.2 \text{ V}} = 1 \text{ mS}/\mu\text{m}$$

$$r_o \approx \frac{1.2 \text{ V}}{0.18 \text{ mA}/\mu\text{m}} \approx 7 \text{ K}\Omega\cdot\mu\text{m}$$

$$|A_v(\max)| = g_m r_o \approx 7$$

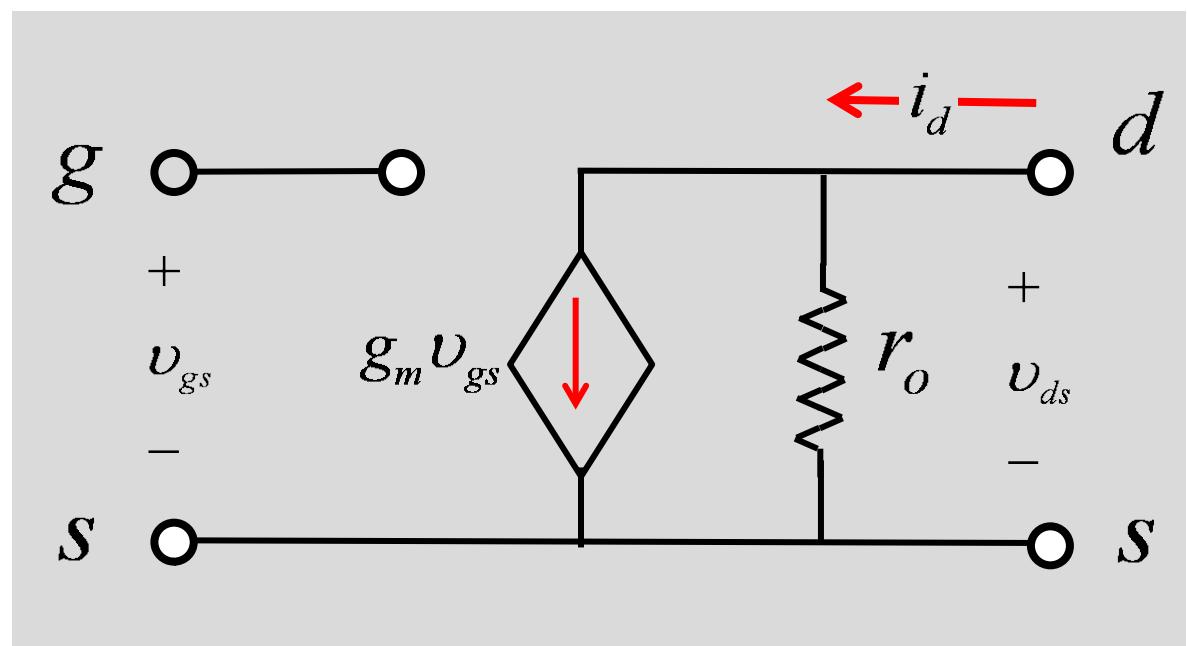
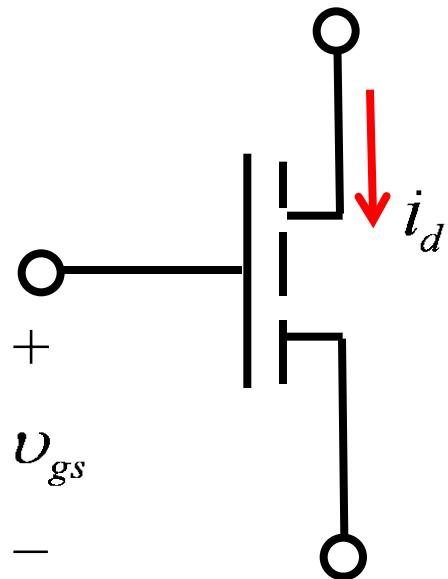
FOR BFT THE GAIN WILL BE MUCH HIGHER IN 10,20,30 TIMES

Self-gain vs. scaling



High frequency performance

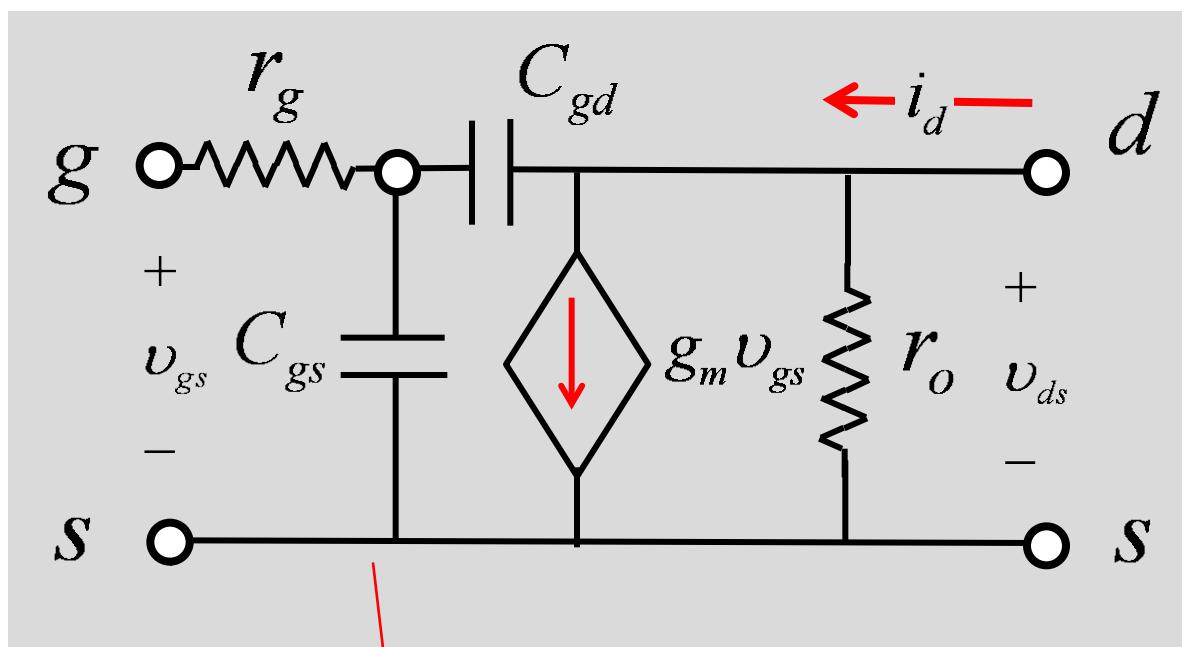
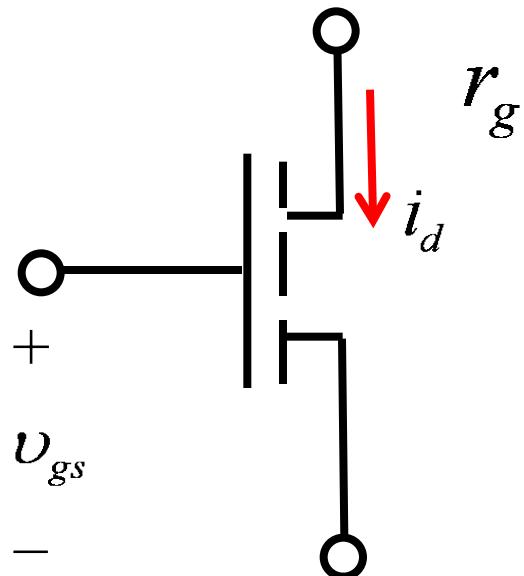
Small-signal, **low frequency** equivalent circuit



$$i_d = g_m v_{gs} + v_{ds}/r_0$$

High frequency performance

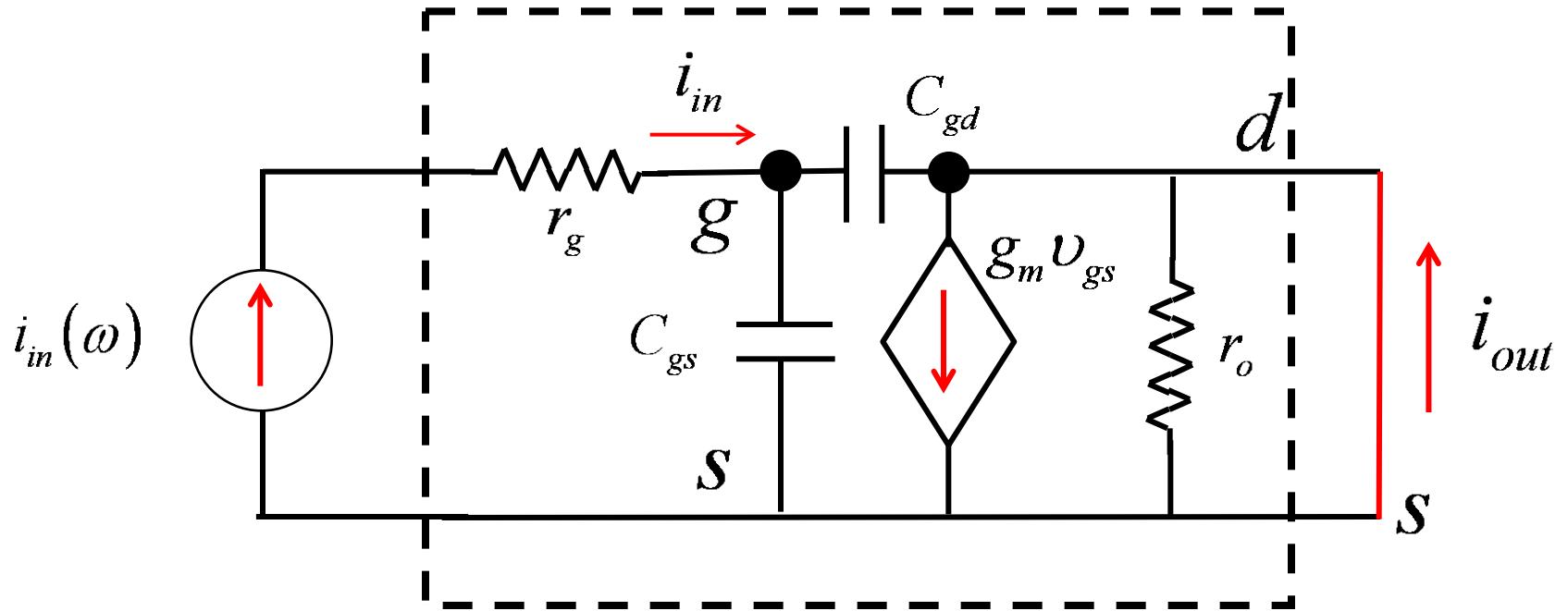
Small-signal, **high frequency** equivalent circuit



$$i_d = g_m v_{gs} + v_{ds} / r_o$$

CAPACITOR BW GATE AND SOURCE.
UNDESIRABLE-TO SMALL IN SIGNIFICANCE

Short circuit current gain



$$i_{out} + (j\omega C_{gd})v_{gs} = g_m v_{gs} \rightarrow i_{out} \approx g_m v_{gs}$$

$$v_{gs} = i_{in} \frac{1}{j\omega(C_{gs} + C_{gd})}$$

$$i_{out} \approx \frac{g_m}{j\omega C_{TOT}} i_{in}$$

Gain-bandwidth product

$$i_{out} \approx \frac{g_m}{j\omega C_{TOT}} i_{in}$$

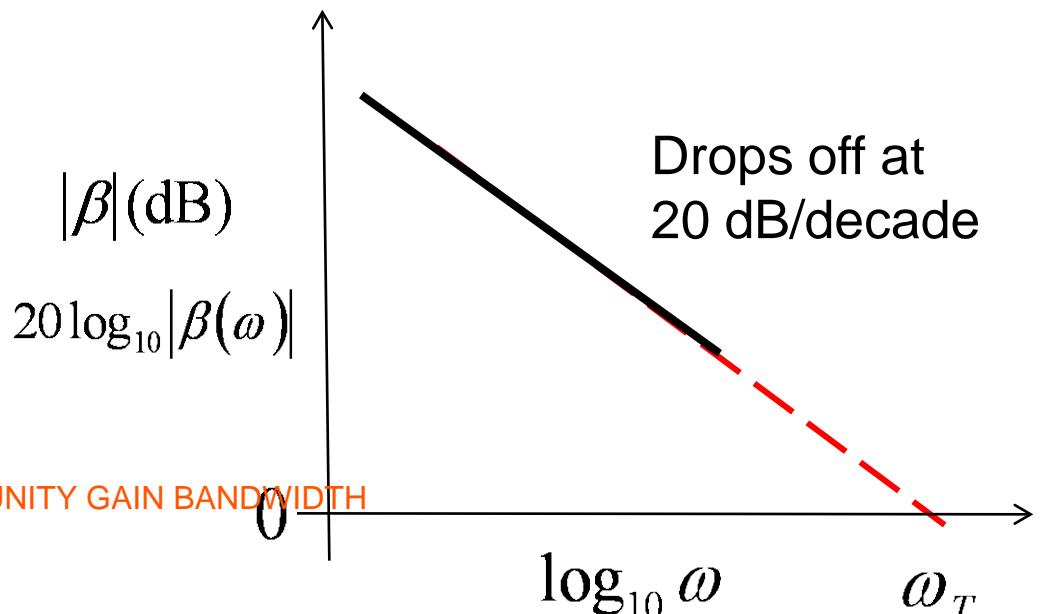
IF WE INCREASE THE FREQUENCY BY A FACTOR OF 10,WE WILL DECREASE THE GAIN BY 20 DECIBALS.

$$|\beta(\omega)| \approx \frac{g_m}{\omega C_{TOT}}$$

$$|\beta(\omega_T)| = 1 = \frac{g_m}{\omega_T C_{TOT}}$$

MAXIMUM FREQUENCY THATS CALLED UNITY GAIN BANDWIDTH PRODUCT.

$$\omega_T = \frac{g_m}{C_{TOT}} = 2\pi f_T$$



MAXIMUM POSSIBLE FREQUENCY WHERE GAIN IS 1.BEYOND THAT POINT WE DONT HAVE ANY USEFUL AMPLIFICATION.THATS THE MAXIMUM POSSIBLE FREQUENCY THAT WE COULD OPERATE.

Lundstrom: 2018

Gain-bandwidth product

$$\omega_T = 2\pi f_T = \frac{g_m}{C_{TOT}}$$

The **gain-bandwidth product** is an important figure of merit for high frequency transistors.

f_{MAX}

$$f_T = \frac{g_m}{2\pi C_{TOT}}$$

insensitive to r_g and r_o

independent of W

channel length scaling
increases f_T

$$f_{MAX} \approx \frac{\omega_T}{\sqrt{4r_g(1/r_o + \omega_T C_{gd})}}$$

sensitive to
parasitics

- r_g
- r_o
- C_{gd}

Another figure of merit is f_{MAX} , the **maximum frequency of oscillation** or the **unity power gain**.

Analog figures of merit

transconductance

device noise

self gain

device mismatch

f_T and f_{MAX}

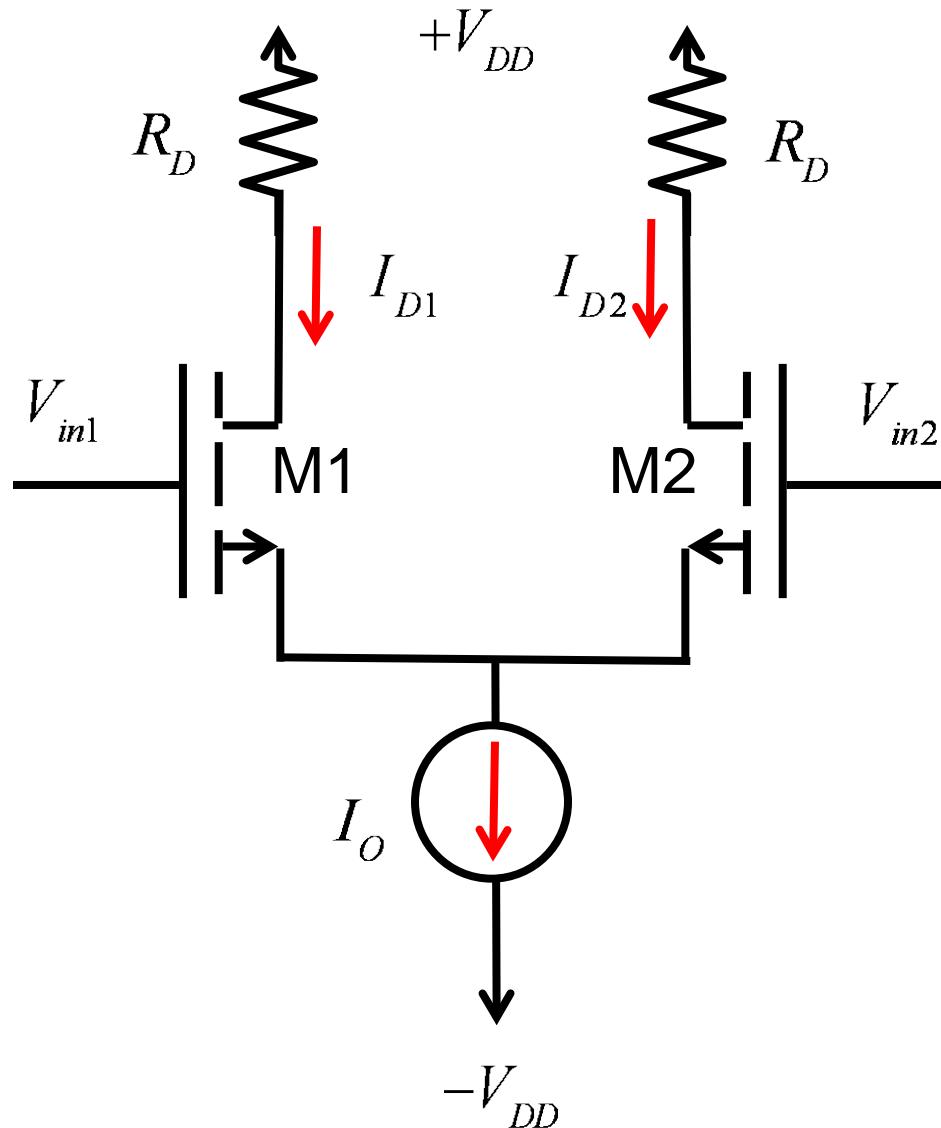
device linearity

FOR BALANCING AND HAVE IDEALITY AND SMALL DIFFERENCE BW

Source coupled pairs

CHARACTERISTICS OF ONE OTHER.

IDEAL BEHAVIOUR PROPER OPERATIONS.



Summary

Small signal transconductance is an important figure of merit for a transistor. A_{GAIN}

Self-gain ($g_m r_0$) is another important figure of merit.

f_T and f_{max} are key figure of merit for RF applications.

Other important device parameters are noise, mismatch, and linearity.

Next topic: Device metrics

Now that we understand what's important for digital and analog circuits, we can define an easily-measured, relevant set of **device metrics**.

Essentials of MOSFETs

Unit 1: Transistors and Circuits

Lecture 1.4: MOSFET Device Metrics

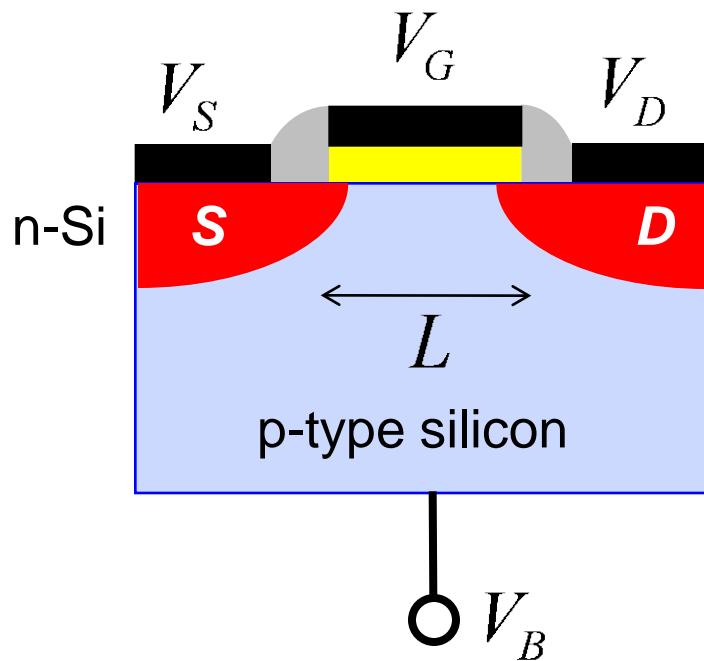
Mark Lundstrom

lundstro@purdue.edu
Electrical and Computer Engineering
Purdue University
West Lafayette, Indiana USA

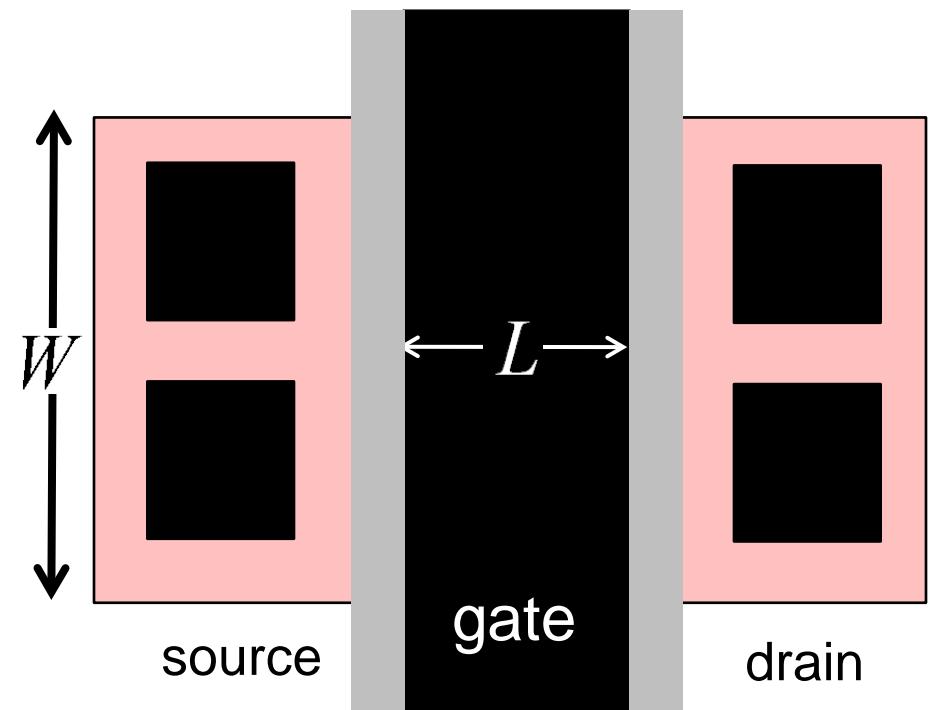
DOUBLE THE WIDTH OF THE MOSFET DOUBLE THE CURRENT

On current scales with MOSFET width

side view

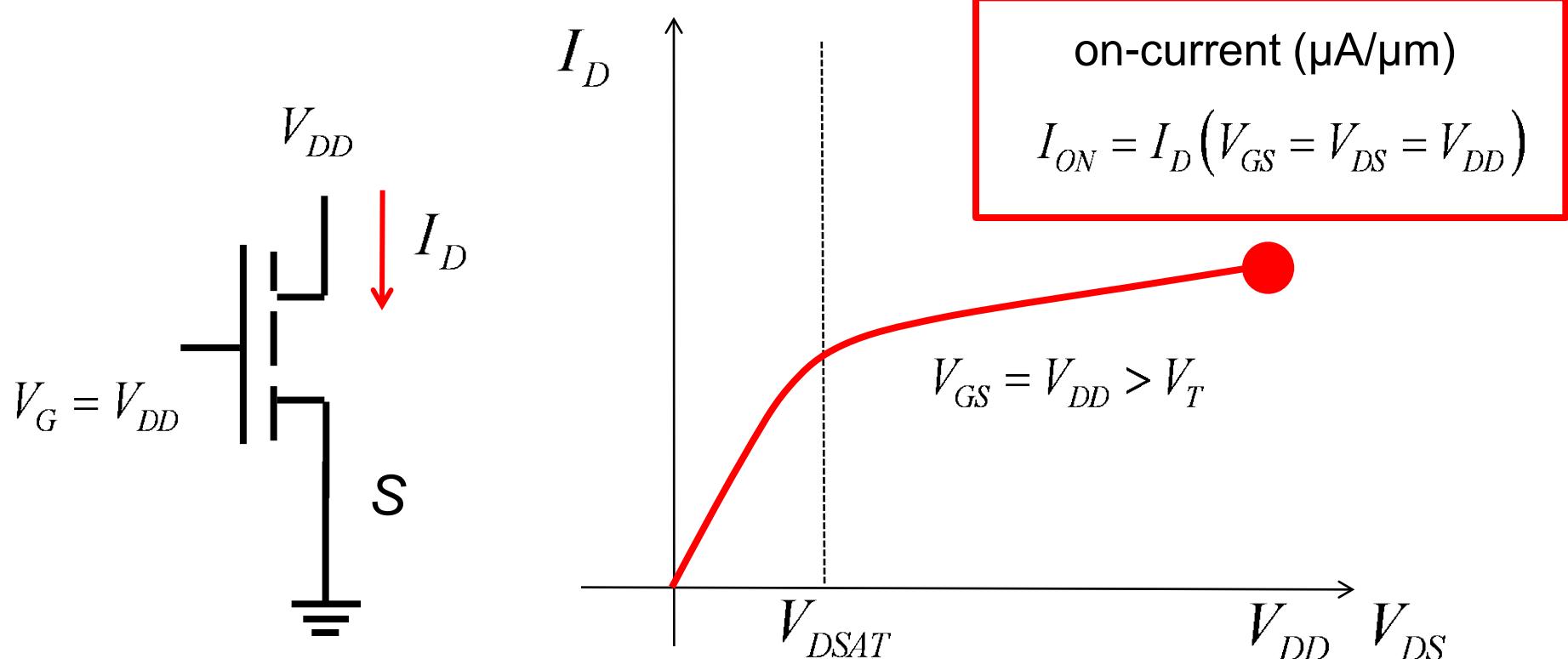


top view



Currents will be quoted in millamps per micrometer of width
(or microamps per micrometer).

1) On-current



output characteristic:
 I_D vs. V_{DS} at fixed V_{GS}

The on-current metric

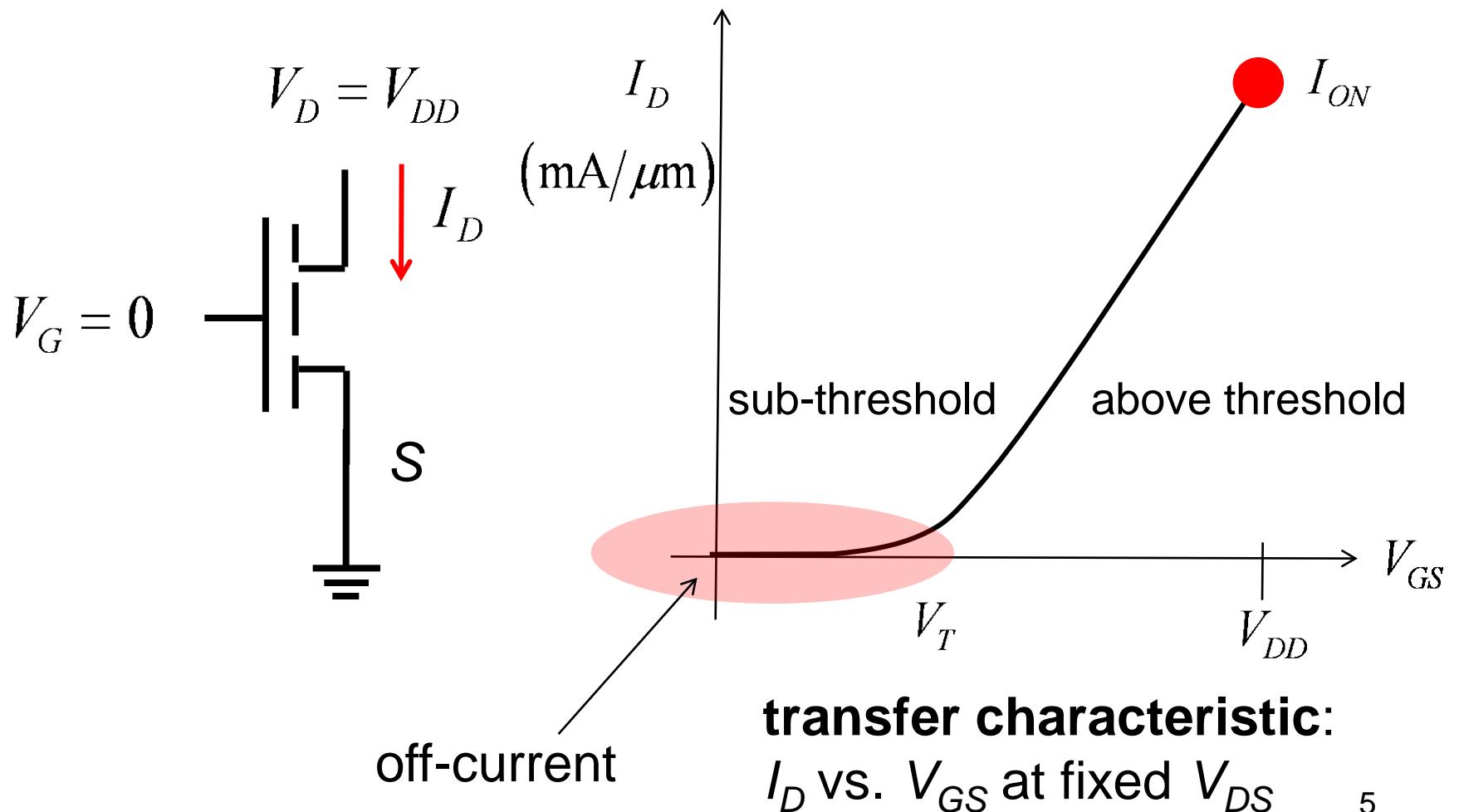
$$I_{ON} = I_D (V_{GS} = V_{DS} = V_{DD})$$

On current is an important device metric for digital electronics because it determines the maximum speed of the circuit.

$$\tau_{cir} \propto \frac{C_{sw} V_{DD}}{I_{ON}}$$

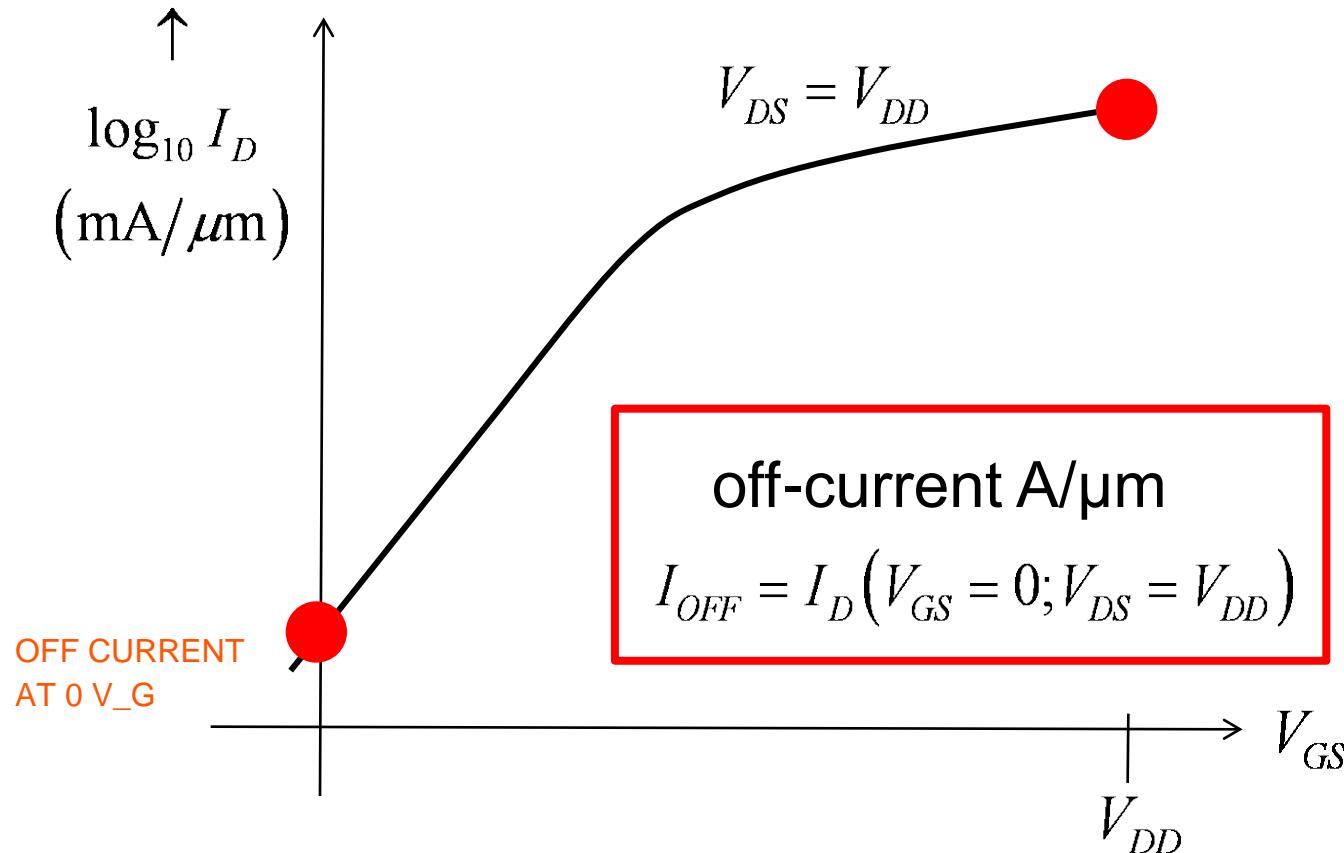
circuit speed

2) Off-current



Off-current

transfer characteristic on a semi-log plot:



The off-current metric

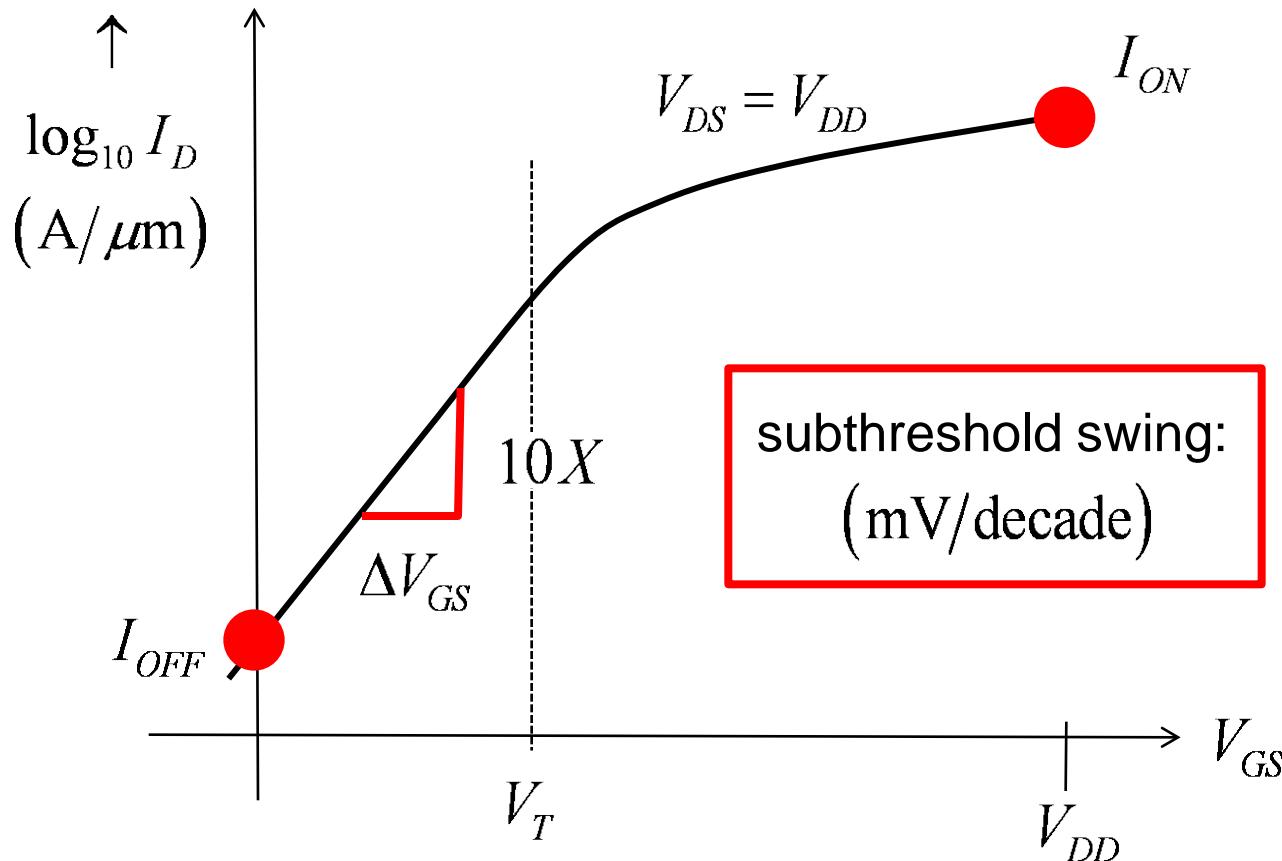
$$I_{OFF} = I_D(V_{GS} = 0; V_{DS} = V_{DD})$$

Off-current is an important device metric for static power.

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

3) Subthreshold Swing

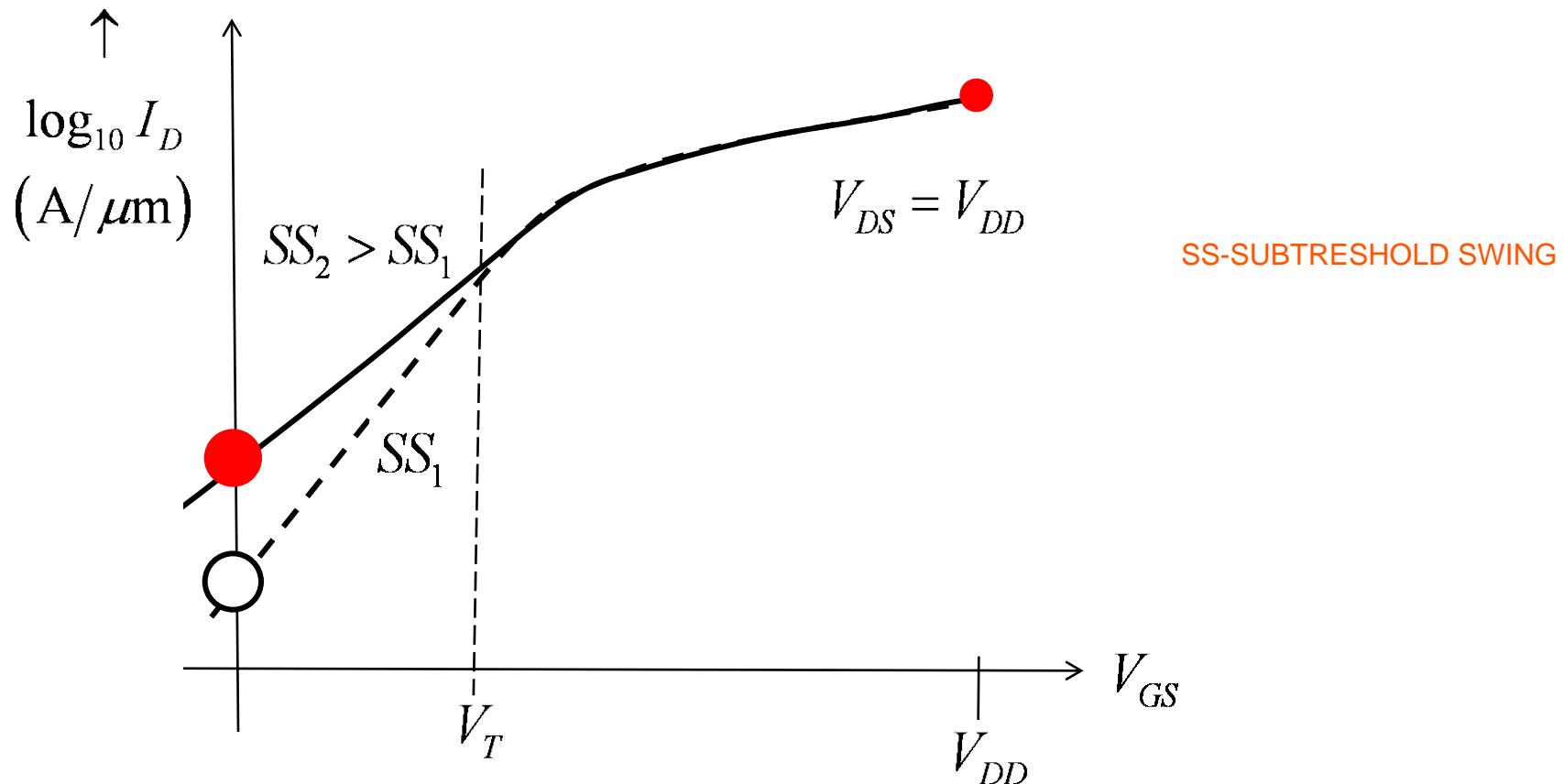
transfer characteristics:



HOW MUCH WE NEED
TO INCREASE THE GATE
VOLTAGE IN ORDER TO
INCREASE THE
SUBTHRESHOLD
REGIME BY A FACTOR
OF 10.

Higher SS increases off-current (exponentially)

transfer characteristics:



SS, V_{DD}, and Power

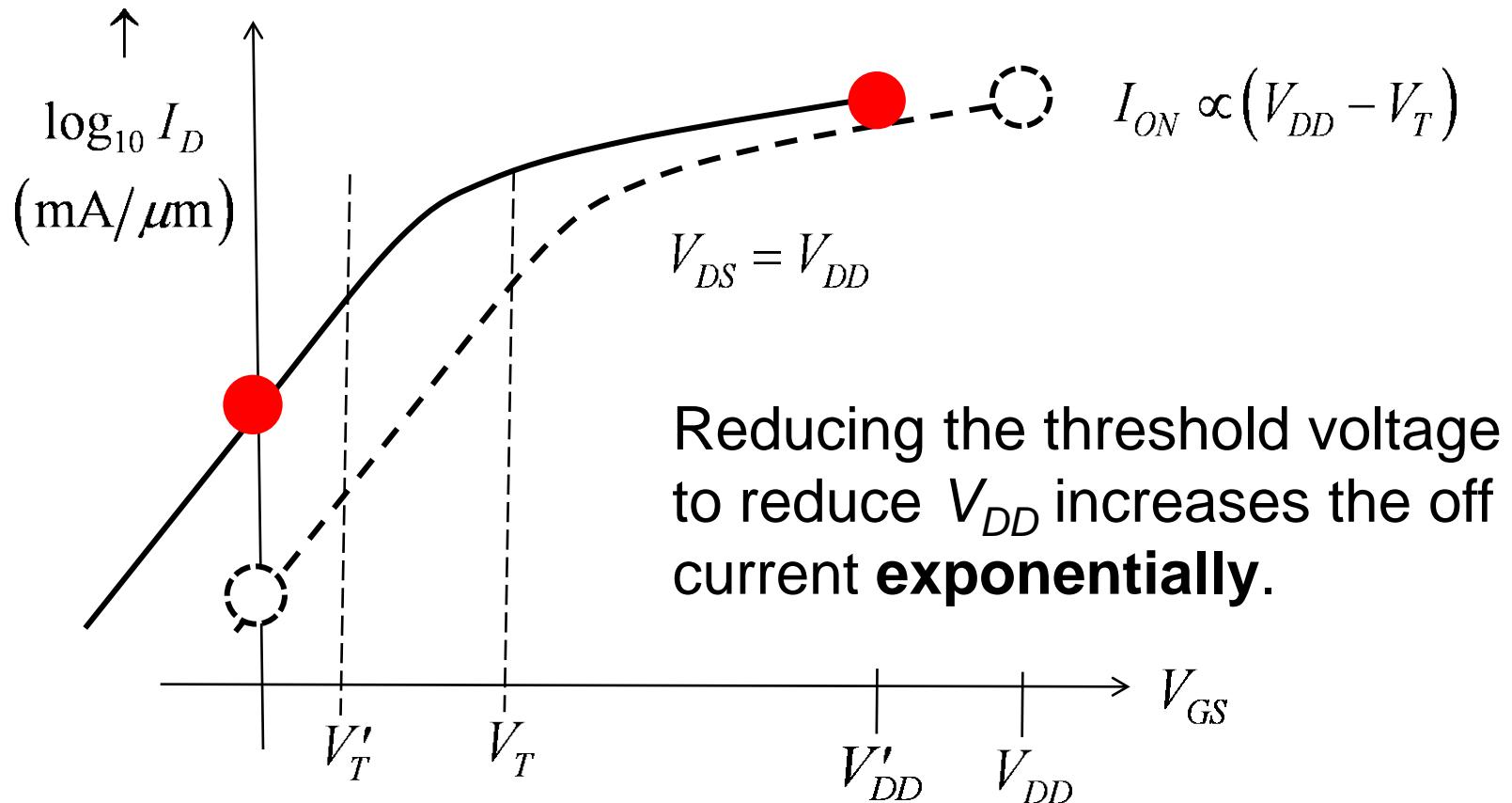
$$P_{static} = N_G I_{OFF} V_{DD}$$

$$P_{dynamic} = \alpha f C_{sw} V_{DD}^2$$

- Static power is controlled by the off-current.
- The SS determines the off-current
- To minimize dynamic power dissipation at a given frequency, we should use the lowest power supply voltage possible.
- The SS determines the minimum power supply.

Reducing V_T increases off-current (exponentially)

transfer characteristics:



The SS metric

$$SS = \left(\frac{\partial(\log_{10} I_D)}{\partial V_{GS}} \right)^{-1} = \frac{\text{mV}}{\text{decade}}$$

The number of mV that the gate voltage must increase to increase the drain current by a factor of 10.

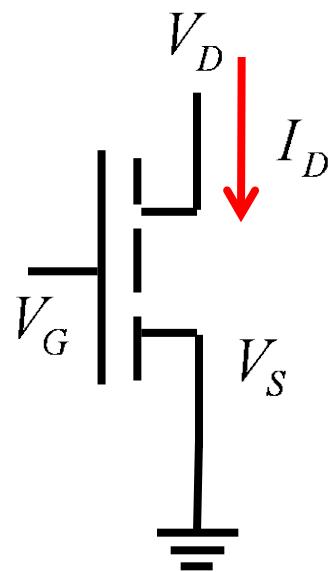
THE PHYSICS OF THE MOSFET DICTATES THAT THE SUBTHRESHOLD SWING CAN NEVER BE LESS THAN 60 mV PER DECADE.

SS must be minimized to minimize off-current.

SS must be minimized to minimize V_{DD} .

In a MOSFET, $SS > 60 \text{ mV/decade}$ at $T = 300 \text{ K}$

Effect of the drain voltage

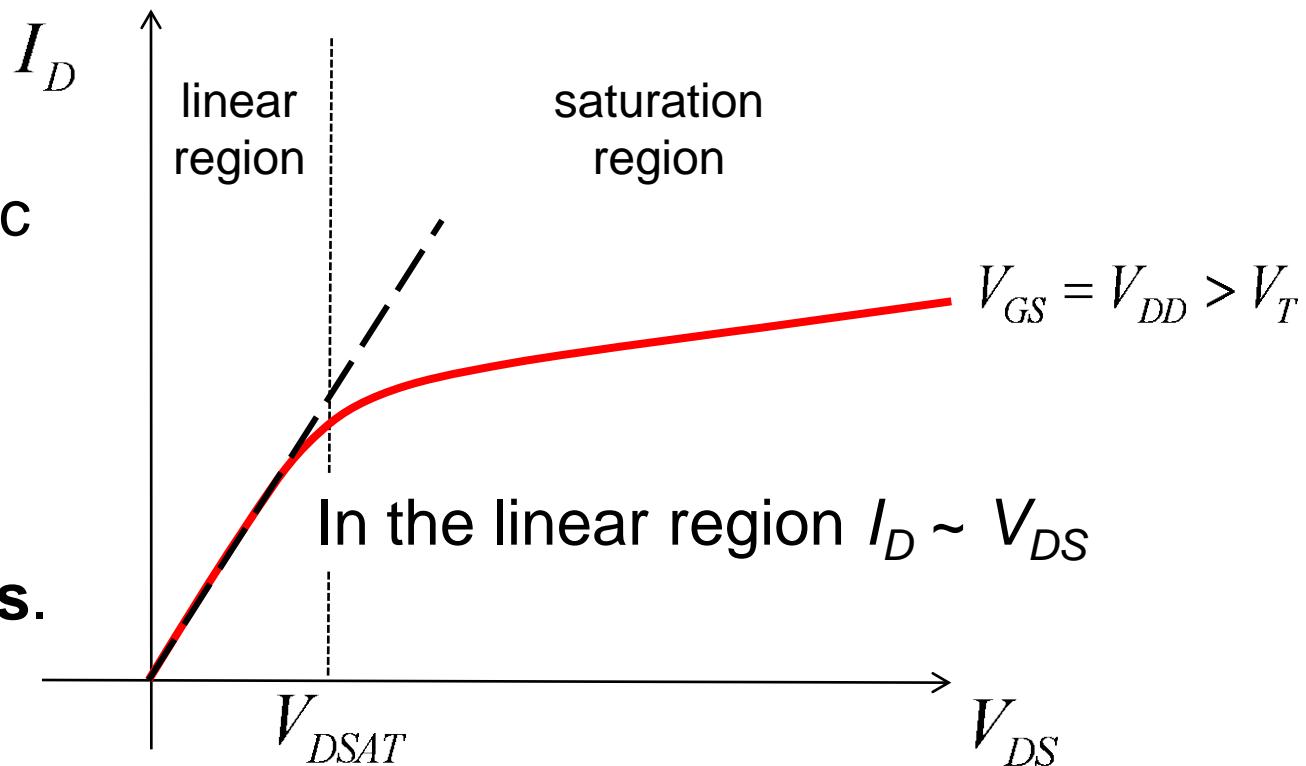


Question: How does the drain voltage affect I_D ?

Answer: In several different ways that are related to the same underlying physics.

Linear region

This is an intrinsic effect present in ideal and real MOSFETs, but there are other **non-ideal effects**.

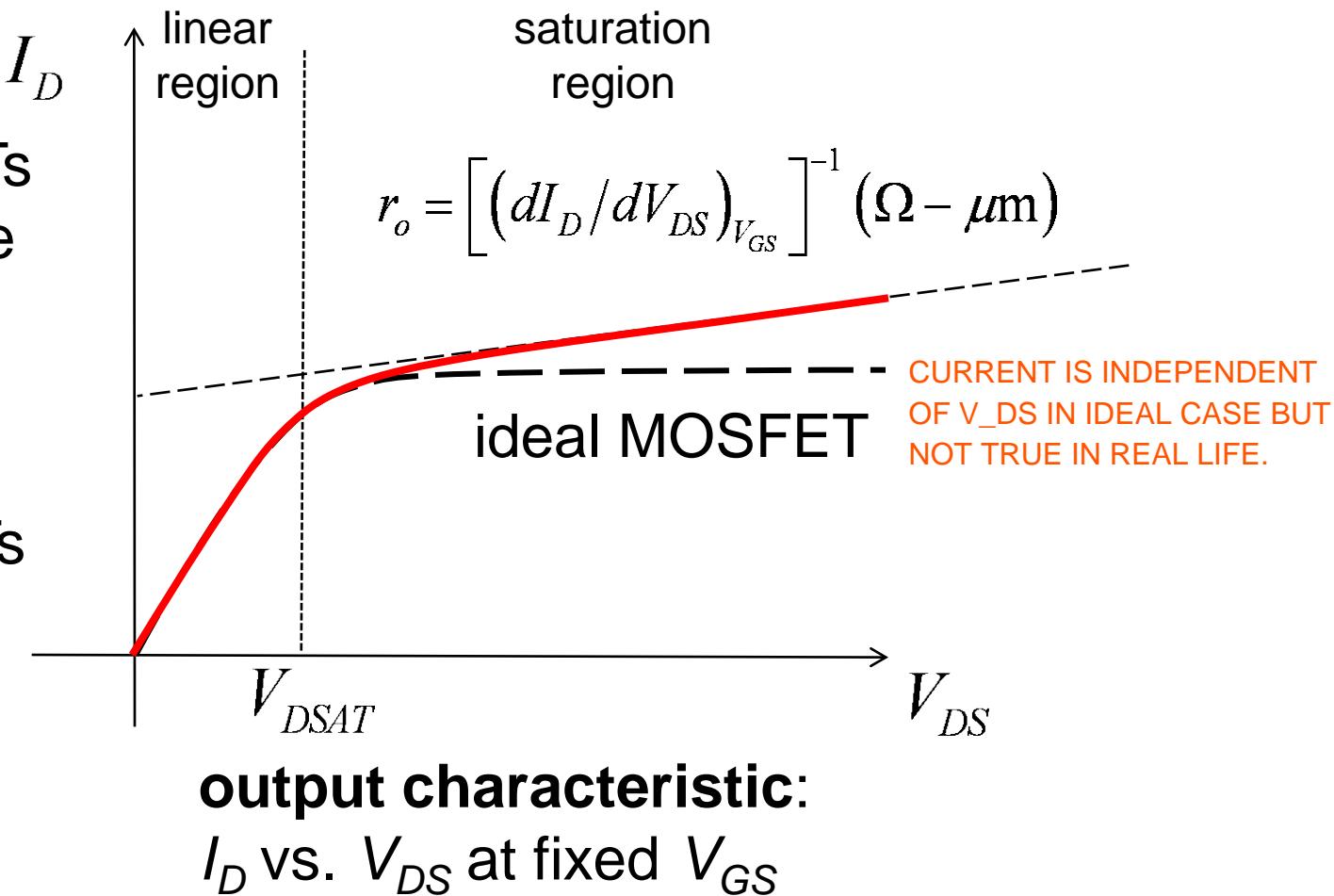


output characteristic:
 I_D vs. V_{DS} at fixed V_{GS}

Effect of V_{DS} in the saturation region

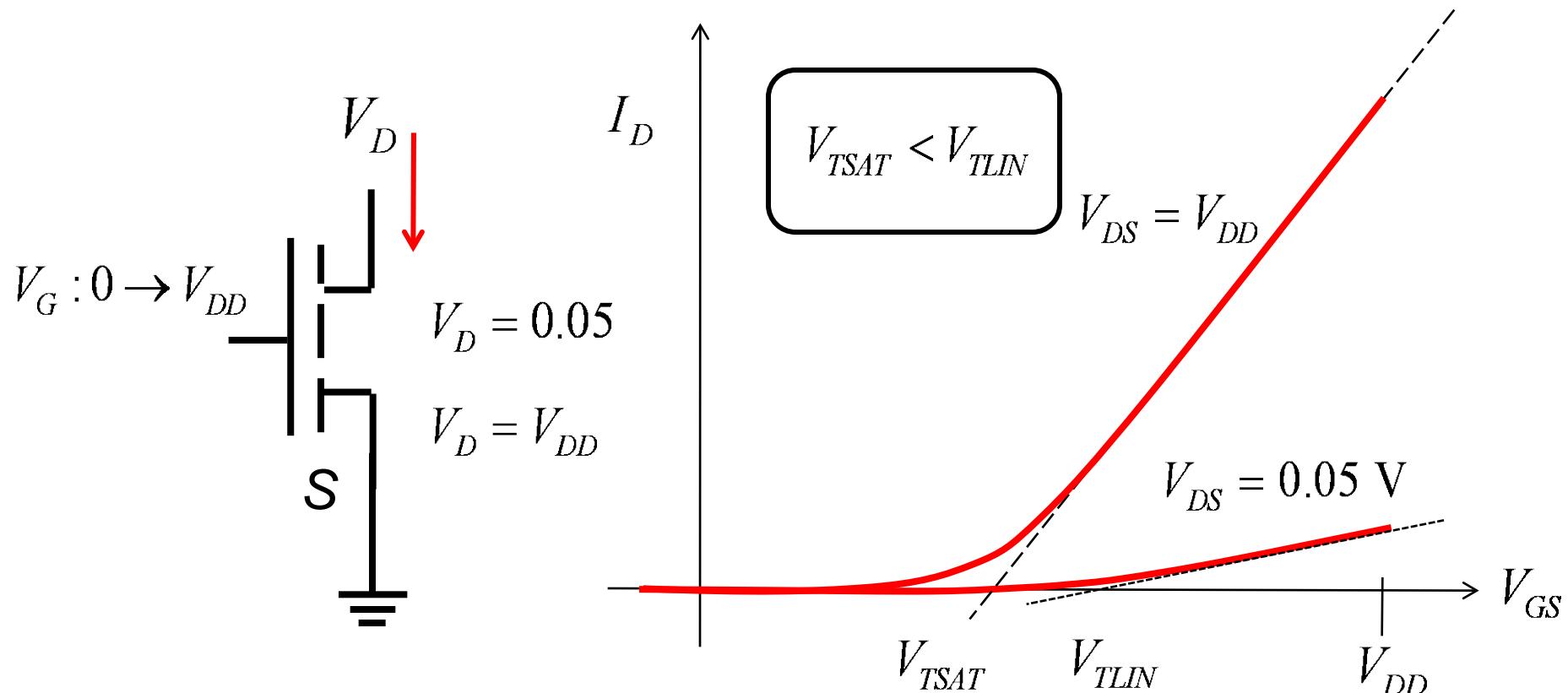
Ideal MOSFETs have an infinite output resistance.

Real MOSFETs have a finite output resistance.



THE DIFFERENCE IN THRESHOLD VOLTAGE CAN'T BE DIFFERENT ON DIFFERENT OPERATING CONDITIONS. IT'S NOT A DESIRABLE EFFECT. WE USUALLY NOTICE THAT V_T IN LINEAR REGION AND IN SATURATION REGION ARE DIFFERENT. WE NEED TO CHOOSE A TRANSISTOR WHICH HAS V_T UNCHANGED.

Effect of V_{DS} on transfer characteristic

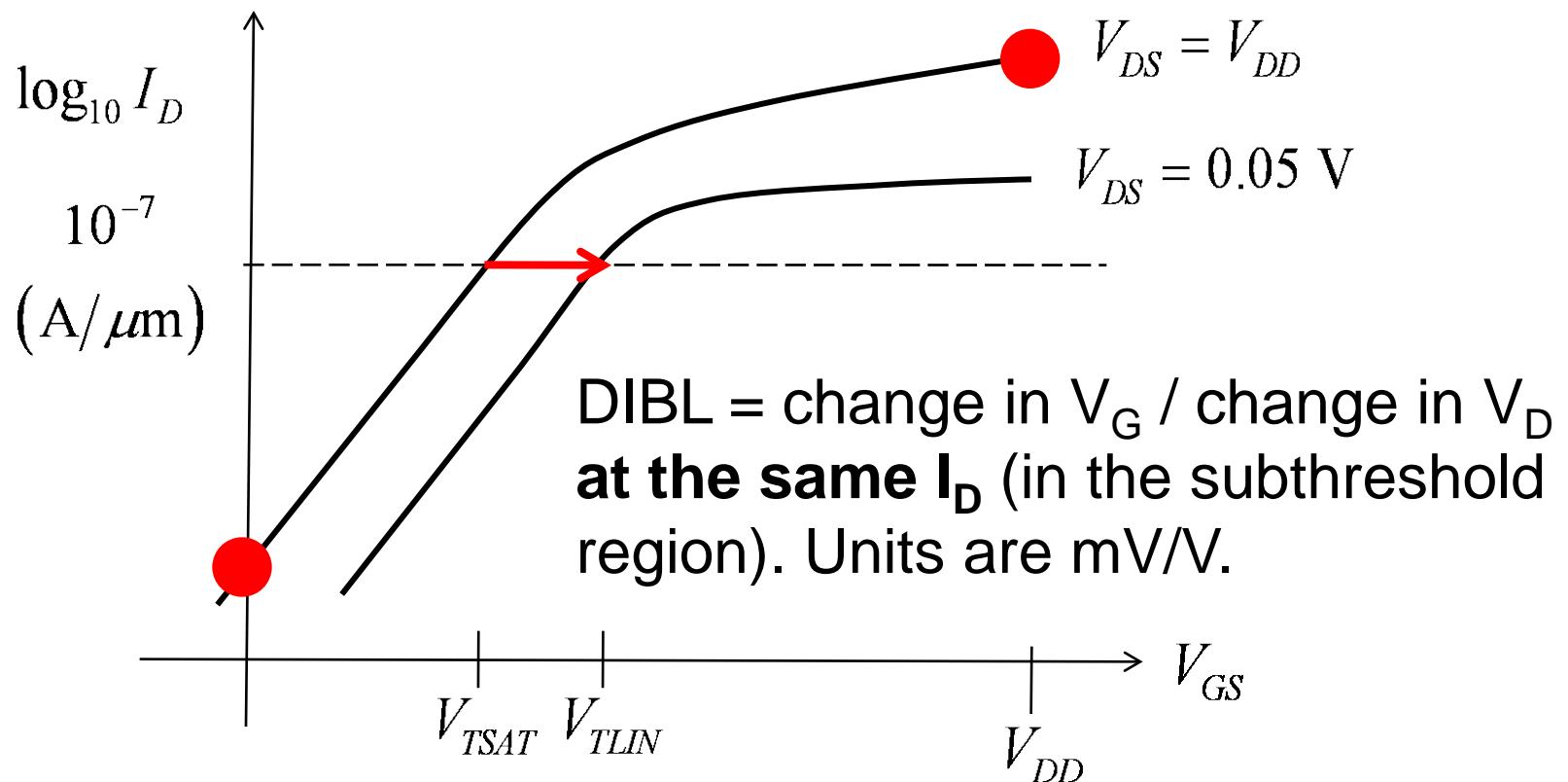


threshold voltage
in saturation region

threshold voltage
in linear region

Effect of V_{DS} in subthreshold

transfer characteristics:



Drain voltage non-idealities

All of these non-ideal effects:

- threshold voltage dependence on V_{DS}
- non-infinite output resistance
- DIBL

Are due to the same physics. A single metric is used to assess the magnitude of all these effects.

DIBL: Drain Induced Barrier Lowering

4) The DIBL metric

$$DIBL = \left. \frac{\partial V_{GS}}{\partial V_{DS}} \right|_{I_D} = \frac{mV}{V}$$

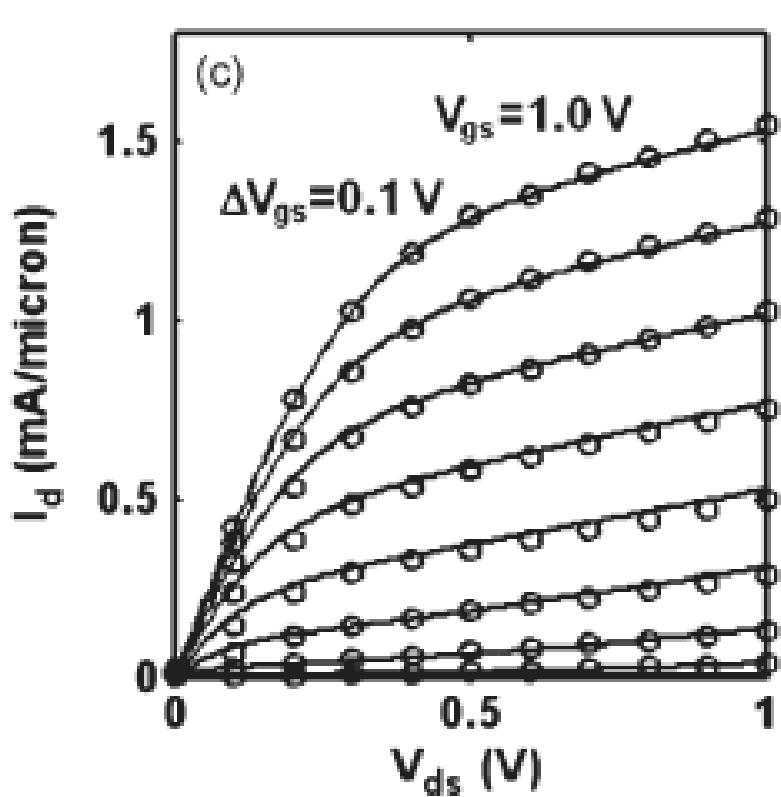
DIBL = change in V_G /
change in V_D **at the same**
 I_D (in the subthreshold
region). Units are mV/V.

The higher the DIBL, the more sensitive the
threshold voltage is to the drain voltage (and
the lower the output resistance).

Key Figures of Merit for digital applications

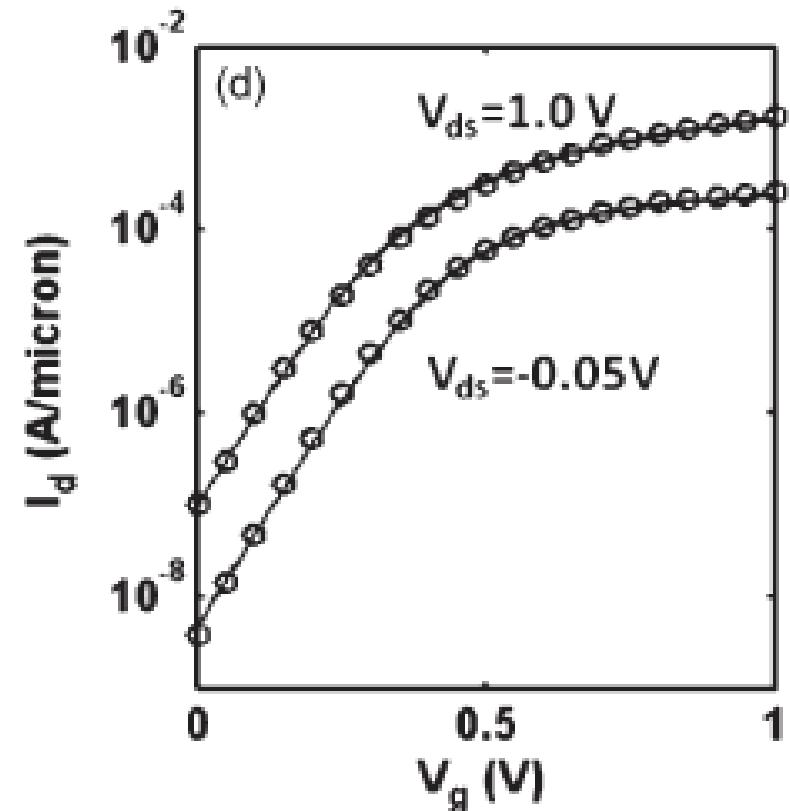
- 1) On current
- 2) Off-current
- 3) Subthreshold swing
- 4) DIBL

Example: 32 nm N-MOS technology



$$I_{ON} \approx 1.55 \mu\text{A}/\mu\text{m}$$

$$I_{OFF} \approx 10^{-7} \text{ A}/\mu\text{m}$$

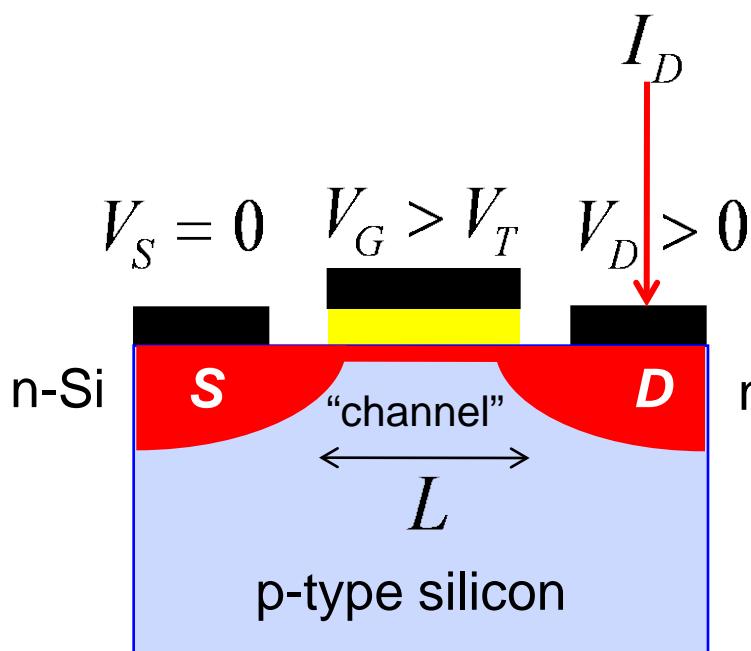


$$SS \approx 105 \text{ mV/decade}$$

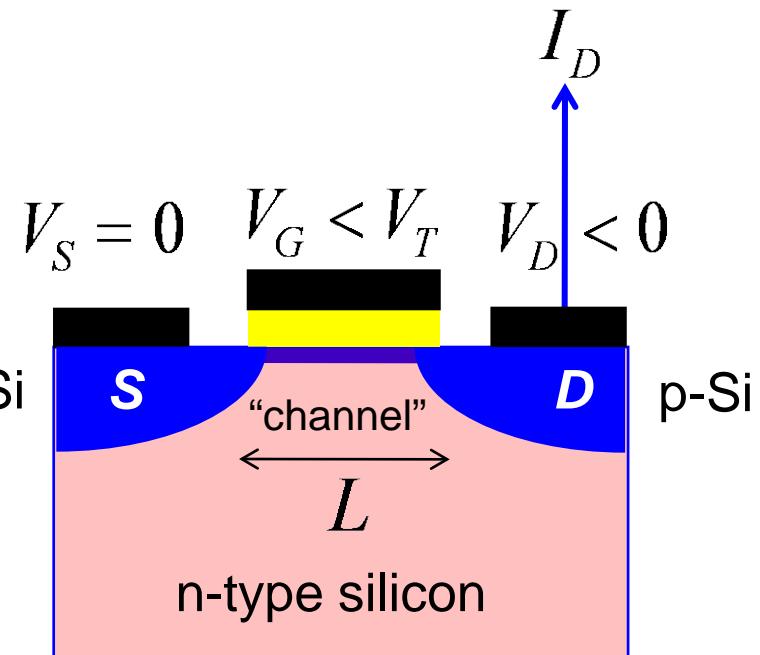
$$DIBL \approx 160 \text{ mV/V}$$

N-channel vs. P-channel MOSFET

n-MOSFET



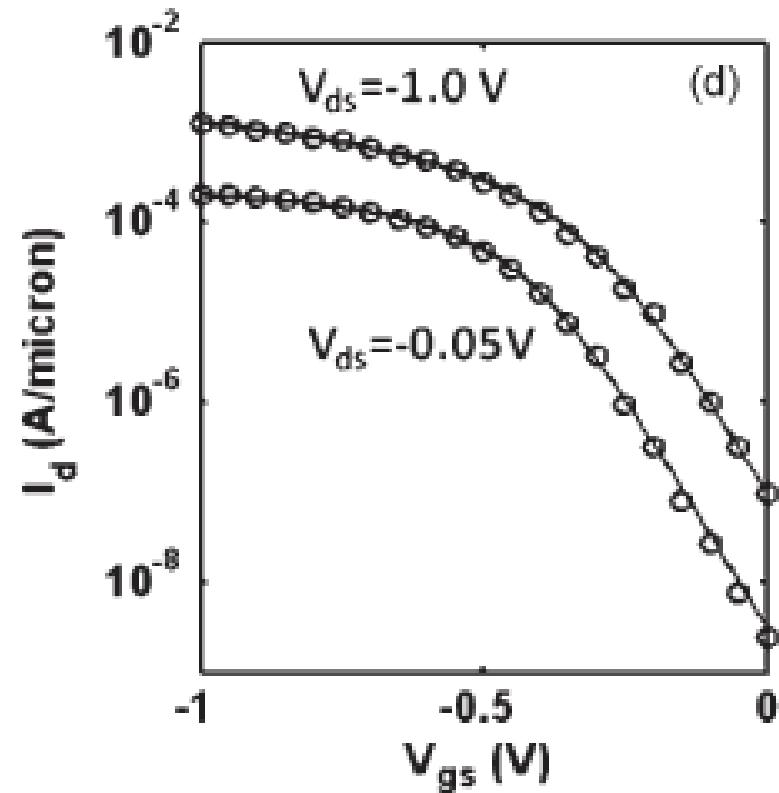
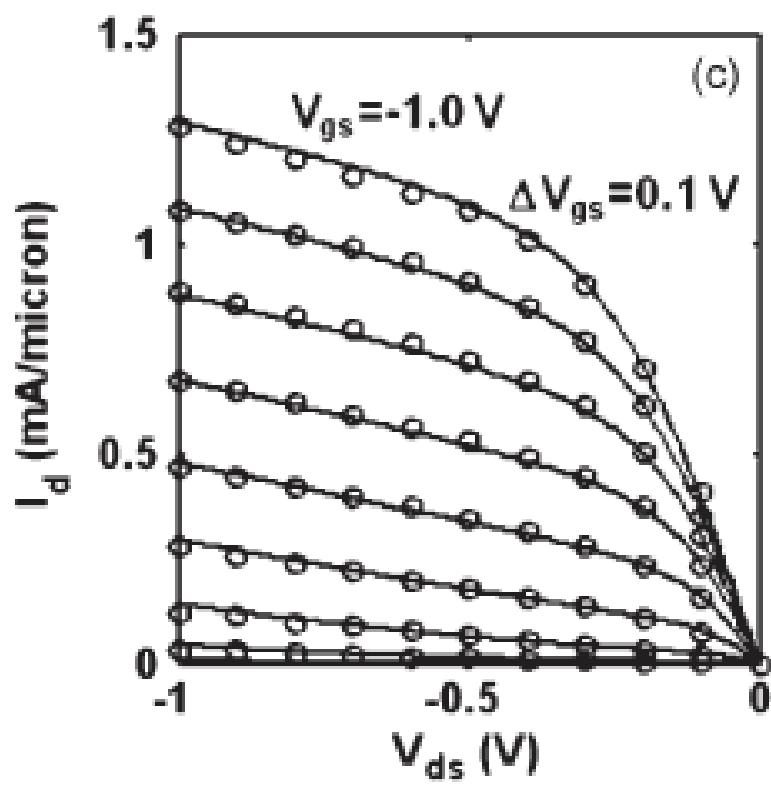
p-MOSFET



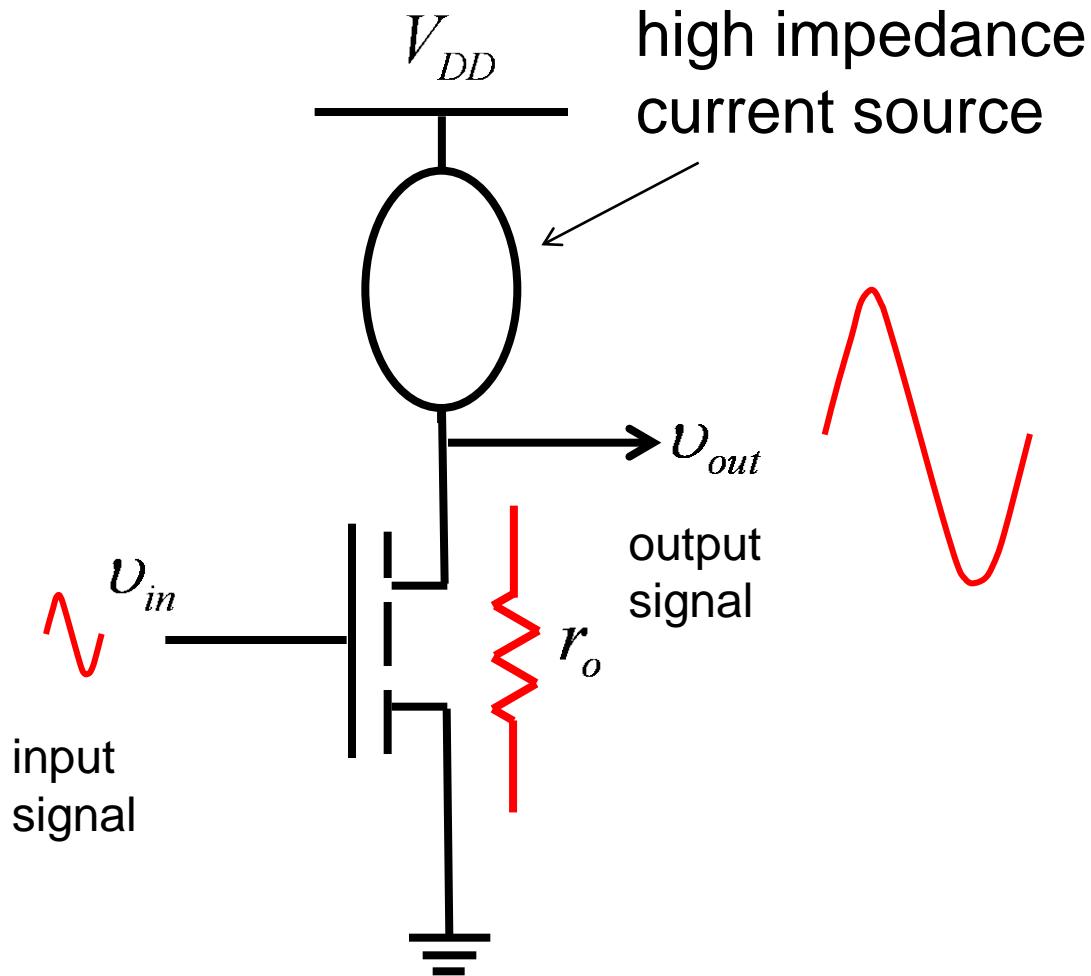
side view

side view

Example: 32 nm P-MOS technology



Analog device metrics



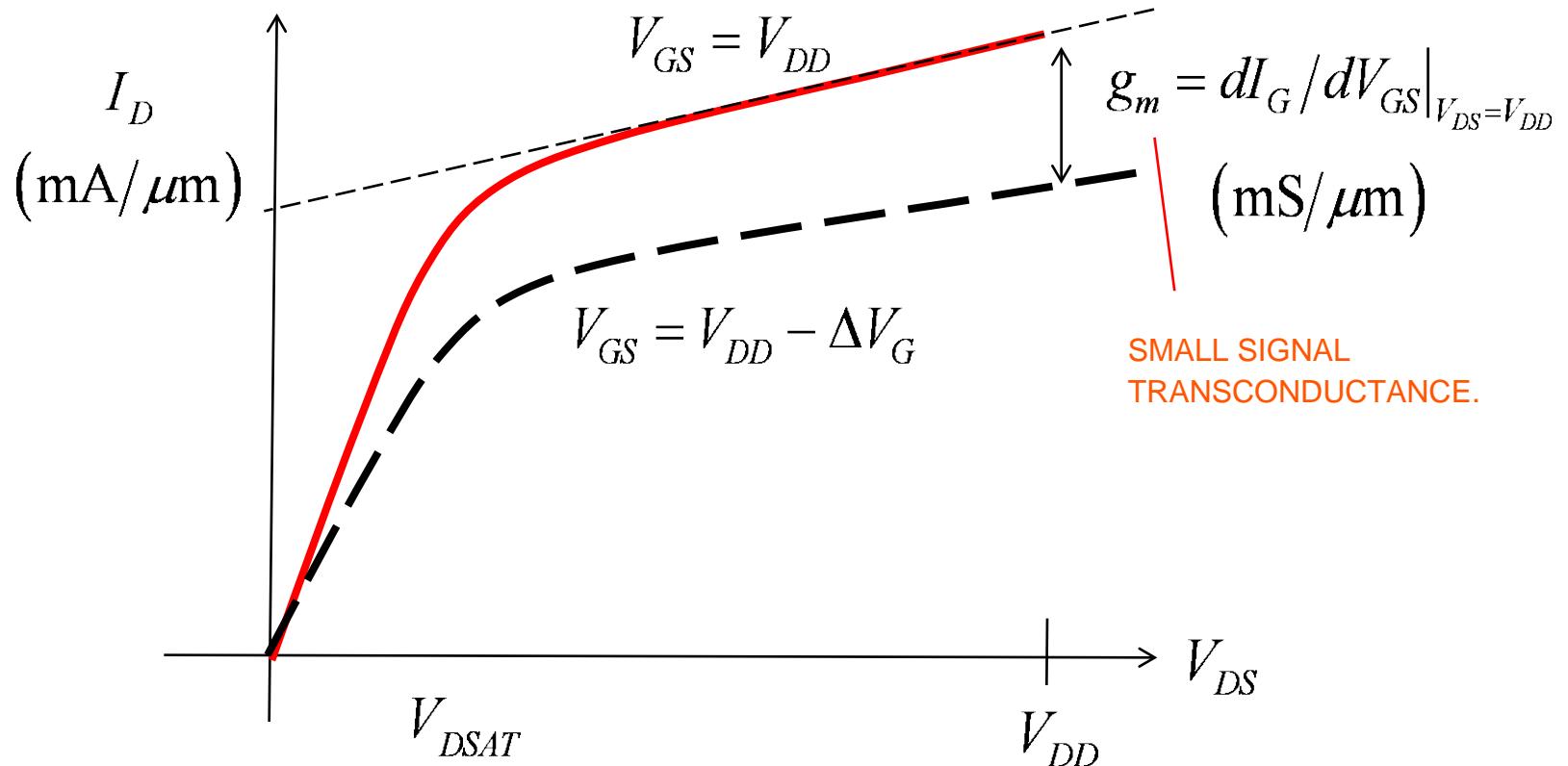
$$A_v(\max) = -g_m r_o$$

g_m and r_o are important analog device metrics

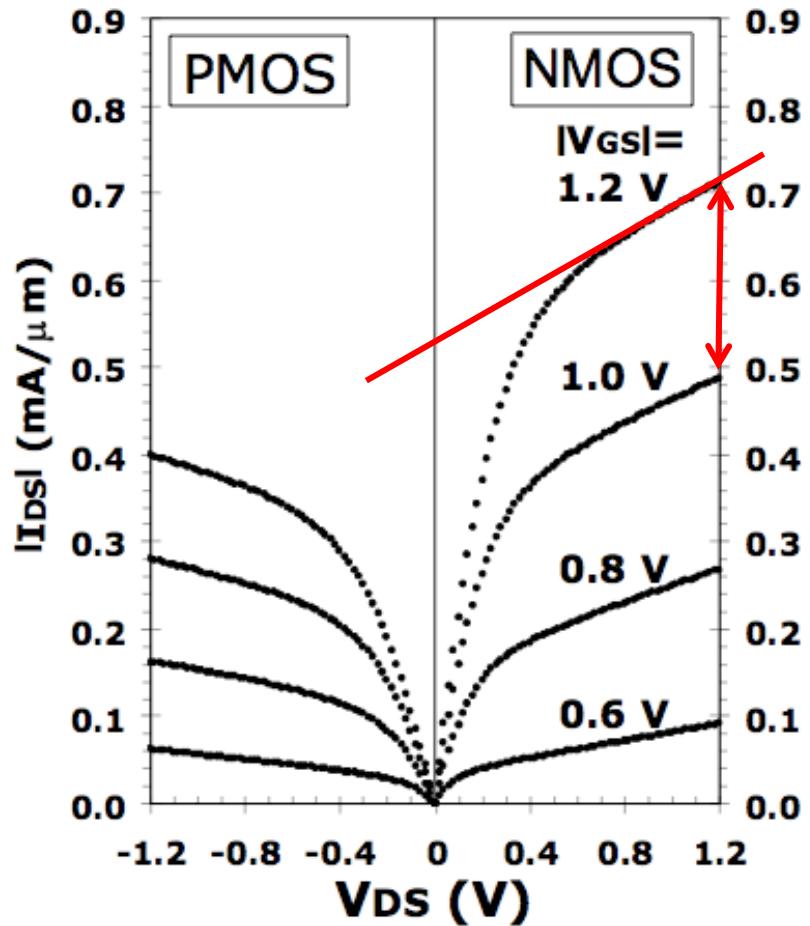
(also f_T , f_{max} , linearity, noise, mismatch, etc.)

Analog device metrics

$$r_o = \left[\left(dI_D / dV_{DS} \right)_{V_{GS}=V_{DD}} \right]^{-1} \Omega - \mu\text{m}$$



Self-gain for 65 nm technology

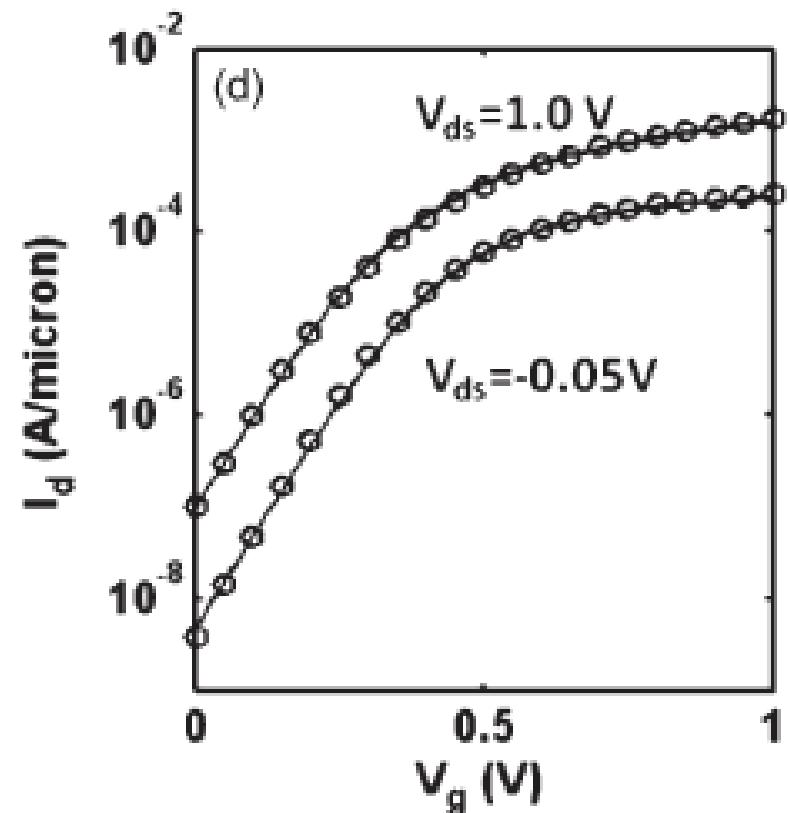
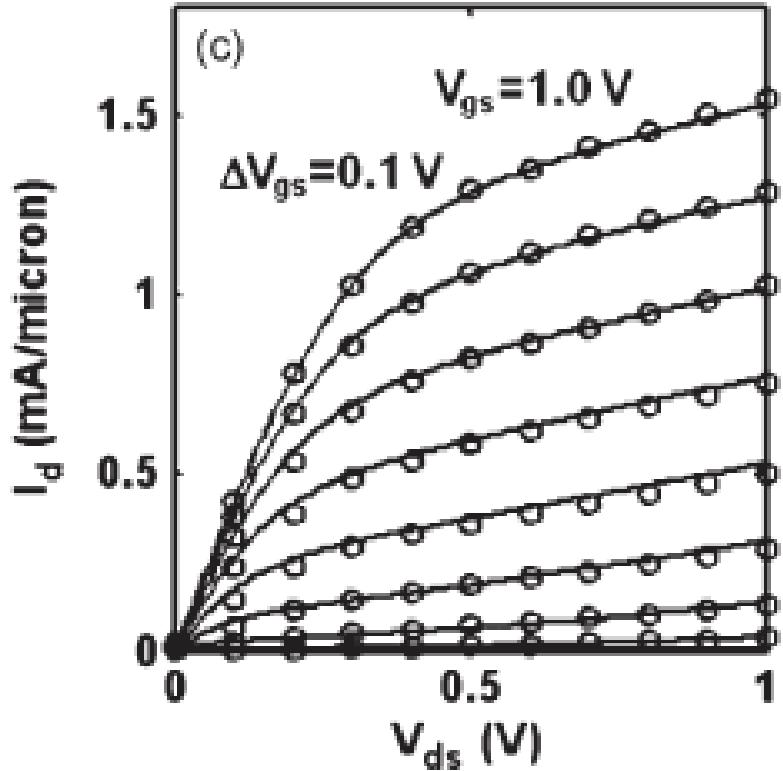


$$g_m \approx \frac{0.2 \text{ mA}/\mu\text{m}}{0.2 \text{ V}} = 1 \text{ mS}/\mu\text{m}$$

$$r_o \approx \frac{1.2 \text{ V}}{0.18 \text{ mA}/\mu\text{m}} \approx 7 \text{ k}\Omega\cdot\mu\text{m}$$

$$|A_v(\max)| = g_m r_o \approx 7$$

32 nm N-MOS technology



$$g_m = 2.5 \text{ mS}/\mu\text{m}$$

$$r_o \approx 2.2 \text{ k}\Omega \cdot \mu\text{m}$$

$$|A_v(\max)| = g_m r_o \approx 5.5$$

COMPARING WITH 65nm TECH g_m is 2.5 TIMES LARGER THAN 65 NM BUT r_o IS SMALLER.

GAIN IS ALSO LOW.

Recap

Given the measured characteristics of a MOSFET, you should be able to determine:

1. on-current: I_{ON}
2. off-current: I_{OFF}
3. subthreshold swing, SS
4. drain induced barrier lowering: DIBL

5. output resistance: r_o
6. *transconductance*: g_m

threshold voltage: $V_T(\text{lin})$ and $V_T(\text{sat})$

drain saturation voltage: V_{DSAT}

Our goal in this course is to understand these device metrics and parameters.

Summary

Key device metrics for digital applications are on-current, off-current, SS, and DIBL.

Key device metrics for analog applications are small signal transconductance, output resistance, and self-gain.

Given a set of IV characteristics, you should be able to extract these metrics.

Our focus in this course is to relate these device metrics to the underlying physics.

Next topic: Compact circuit models

Device metrics help device engineers and circuit design engineers to relate, in a general way, device performance to circuit performance.

Actual circuit design requires more. The next lecture is a short introduction to **compact circuit models**.

Essentials of MOSFETs

Unit 1: Transistors and Circuits

Lecture 1.5: Compact Models

Mark Lundstrom

lundstro@purdue.edu
Electrical and Computer Engineering
Purdue University
West Lafayette, Indiana USA

About “compact models”

The term, **compact model**, is used in many fields of science and engineering to describe a simple (usually analytical) model - as opposed to a simulation, which is typically a numerical solution to a set of equations.

In electronics, “**compact model**” means a model that describes an electronic device in a form that is suitable for use in numerical circuit simulation programs.

SUITABLE FOR NUMERICAL SIMULATIONS LIKE PSPICE ETC.

Two types of compact models

In the course, we will distinguish between two different types of compact models:

1) Compact physical models

These models aim to describe a device in terms of a few parameters with strong physical significance. These kinds of models are useful for device characterization, process monitoring, and for the conceptual understanding that guides device research.

Our focus in this course is on this type of compact model.

Second type of compact models

2) Compact device models for circuit simulation

These models accurately relate the currents that flow into a device's terminals to the voltages on the leads **in a form suitable for use in numerical circuit simulation programs**. To describe everything relevant to a circuit, these models are more complex, but the core of the model is usually a compact physics model.

These kinds of compact models play a critical role in connecting semiconductor R&D and manufacturing to product design.

Focus of course and this lecture

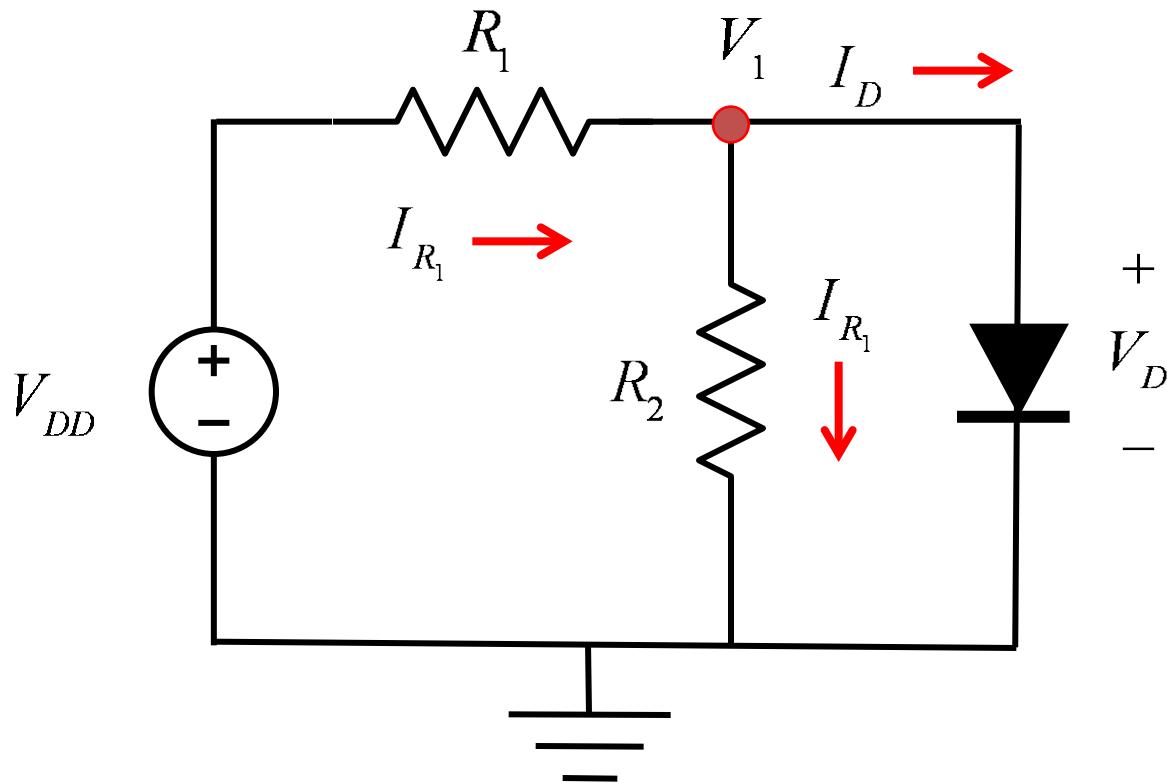
Our focus in this course is on Type 1 compact models designed to provide insight into the physics of MOSFETs.

But everyone involved in electronics should also know a bit about the Type 2 compact models for circuit simulation.

This lecture will introduce a few key considerations for Type 2 models. The course will then focus on Type 1 models for understanding, and we'll return at the end of the course to say a little more about Type 2 compact models.

Compact models for numerical circuit analysis

node 1



KCL:

$$I_{R_1} - I_{R_2} - I_D = 0$$

$$I_{R_1} = (V_{DD} - V_1) / R_1$$

$$I_{R_2} = V_1 / R_2$$

$$I_D(V_1) = ?$$

We need a
compact model!

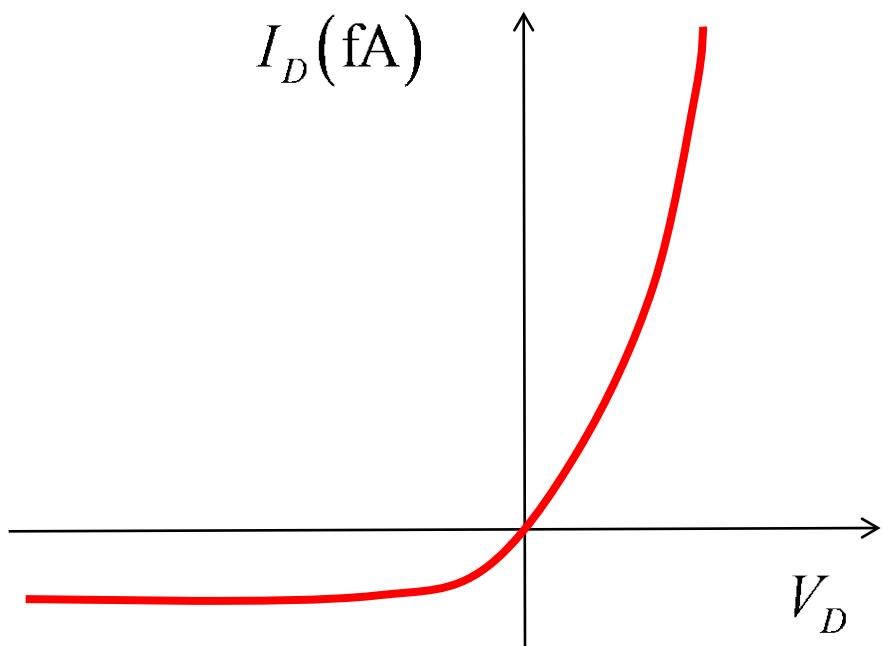
Numerical circuit analysis

Kirchoff's Current Law (KCL) gives one equation in one unknown \square , the voltage at node 1:

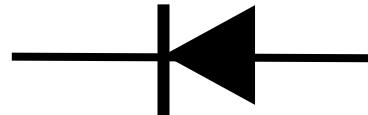
$$f(V_1) = (V_{DD} - V_1)/R_1 - V_1/R_2 - I_D(V_1) = 0$$

The diode current is a non-linear function of the voltage across its terminals; a numerical solution will be needed, but first, we need a compact model for $I_D(V_1)$.

Simple, compact model for a diode



$$I_D = I_S \left(e^{qV_D/k_B T} - 1 \right)$$



- V_D +

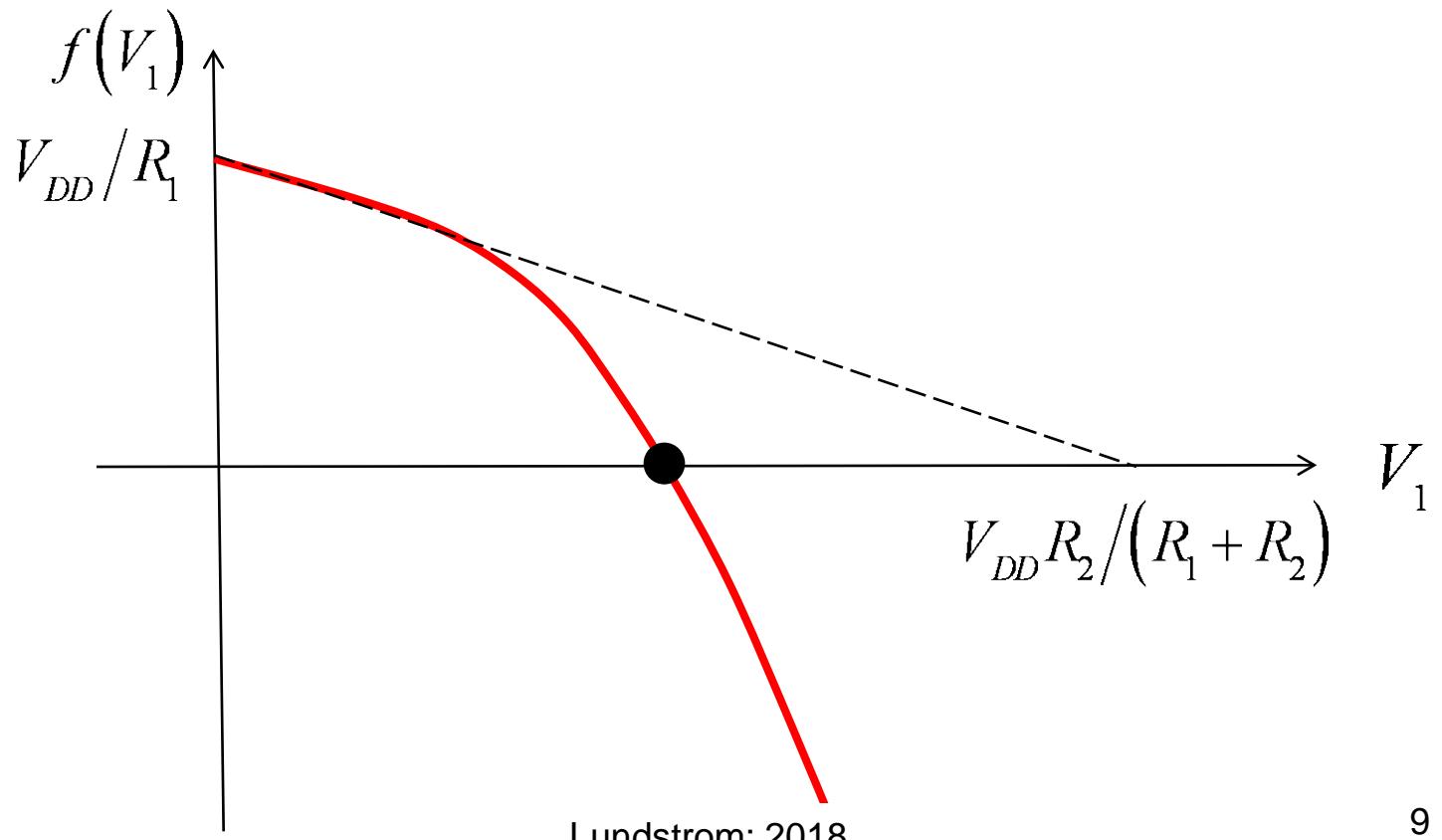
“Shockley diode equation”

$$I_S = A q \left(\frac{n_i^2}{N_D} \sqrt{\frac{D_p}{\tau_p}} + \frac{n_i^2}{N_A} \sqrt{\frac{D_n}{\tau_n}} \right)$$

(Additional physical effects would be included to describe real diodes.)

Solving the circuit equation

$$f(V_1) = (V_{DD} - V_1)/R_1 - V_1/R_2 - I_S(e^{qV_1/k_B T} - 1) = 0$$



Newton's method

$$f(V_1) = (V_{DD} - V_1)/R_1 - V_1/R_2 - I_S(e^{qV_1/k_B T} - 1) = 0$$

1) Guess the voltage at node 1 and then correct the guess:

$$V_1^1 \quad f(V_1^1) \neq 0 \quad f(V_1^1 + \delta V) = 0$$

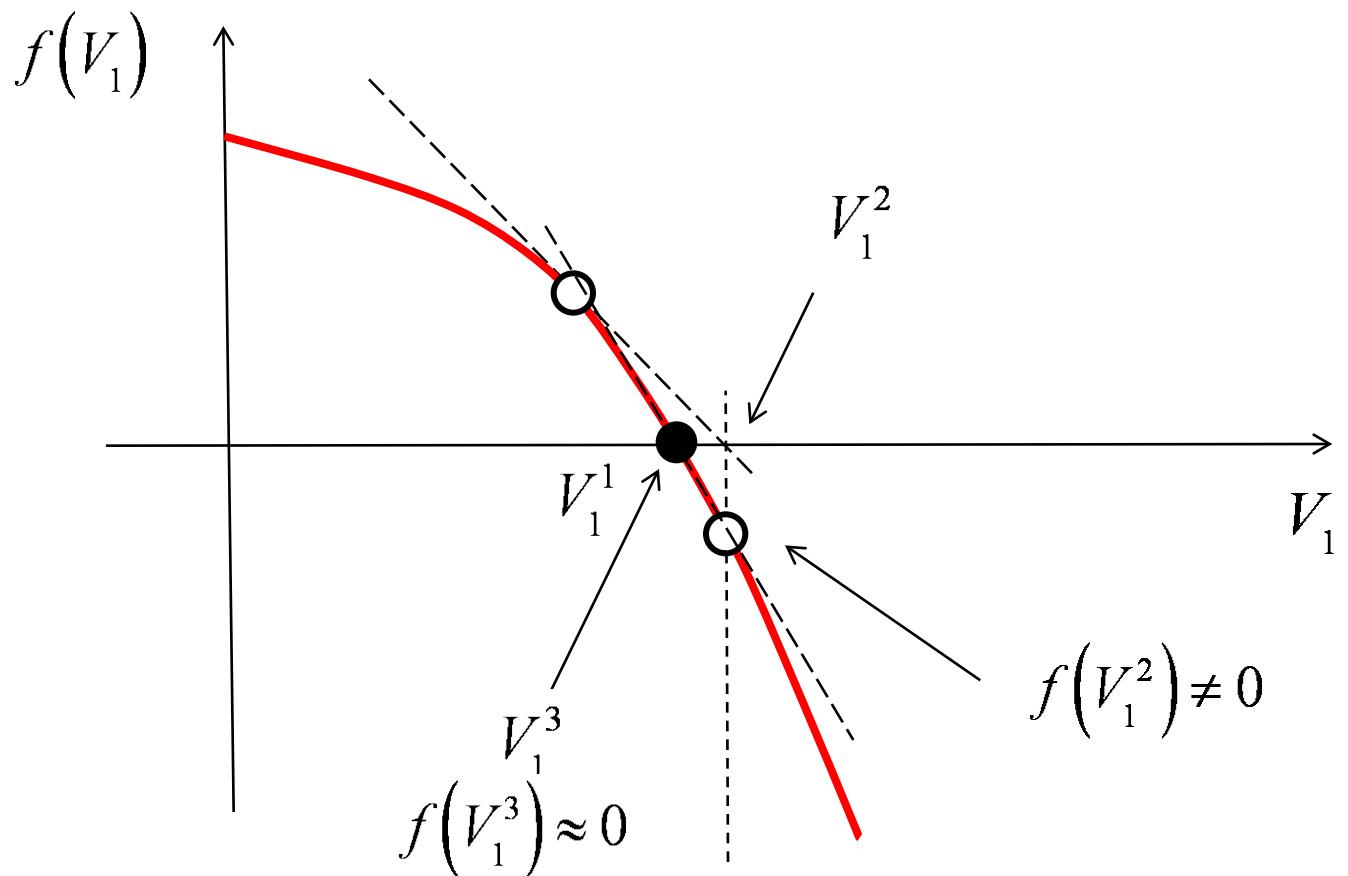
2) Expand $f(V_1)$ in a Taylor series:

$$f(V_1^1 + \delta V) \approx f(V_1^1) + \frac{df}{dV}\Bigg|_{V_1^1} \delta V = 0 \quad \delta V = -\frac{f(V_1^1)}{\frac{df}{dV}\Bigg|_{V_1^1}}$$

3) New guess:

$$V_1^2 = V_1^1 + \delta V \quad f(V_1^2) \neq 0 \quad \rightarrow ? \quad \text{iterate}$$

Graphical representation



Newton's method

$$f(V_1) = (V_{DD} - V_1)/R_1 - V_1/R_2 - I_S(e^{qV_1/k_B T} - 1) = 0$$

1) kth guess:

$$V_1^k$$

2) Correction:

$$\delta V^k = -\frac{f(V_1^k)}{df/dV|_{V_1^k}}$$

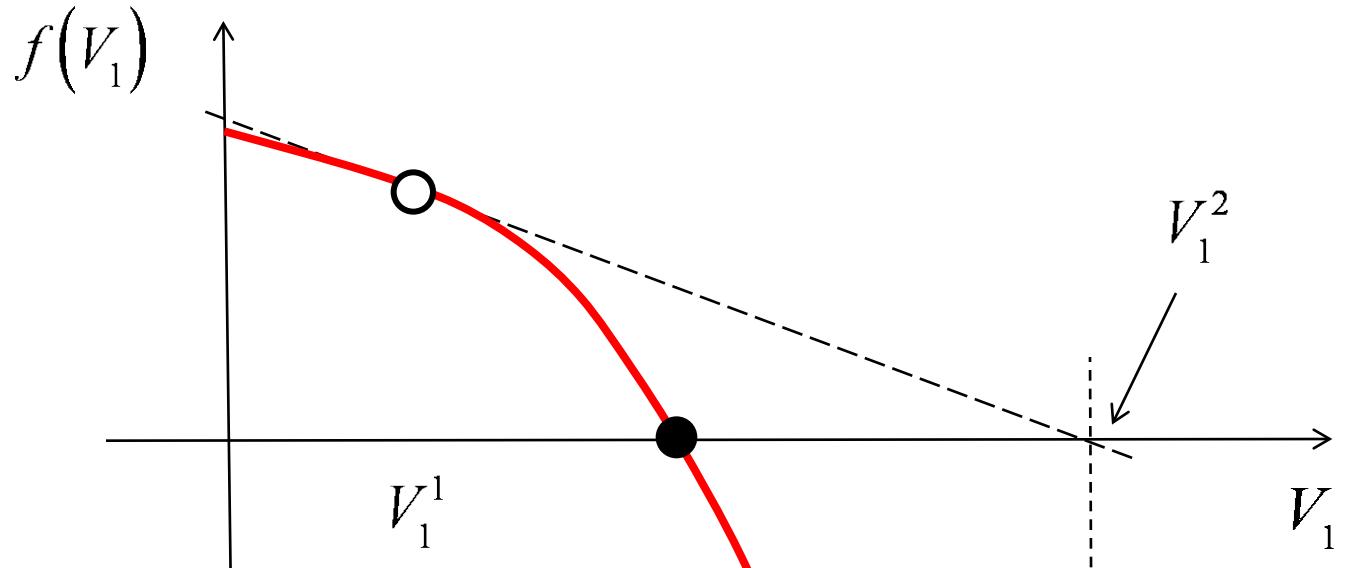
3) New guess:

$$V_1^{k+1} = V_1^k + \delta V^k$$

4) Check convergence: $f(V_1^{k+1}) \approx 0$?

5) Iterate to a pre-determined tolerance

What can go wrong?

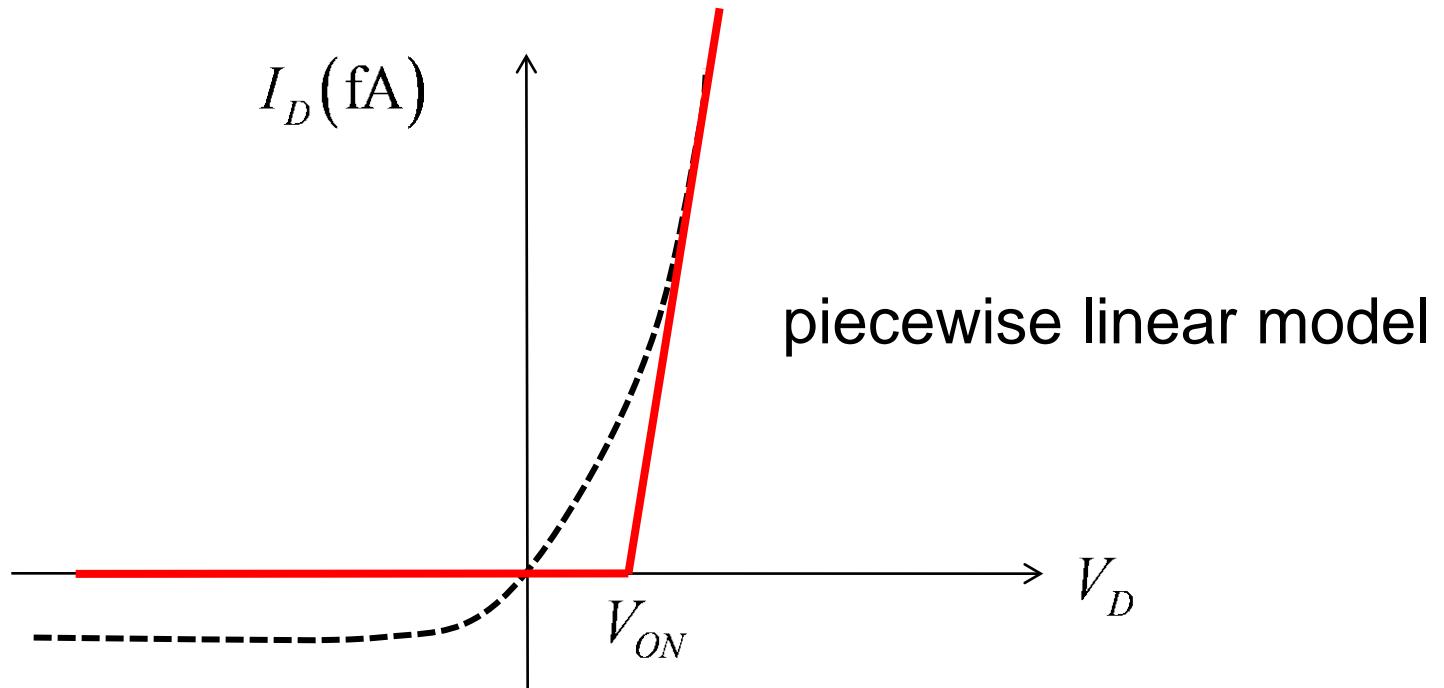


The compact model must describe the device outside the range of voltages that will occur in practice.

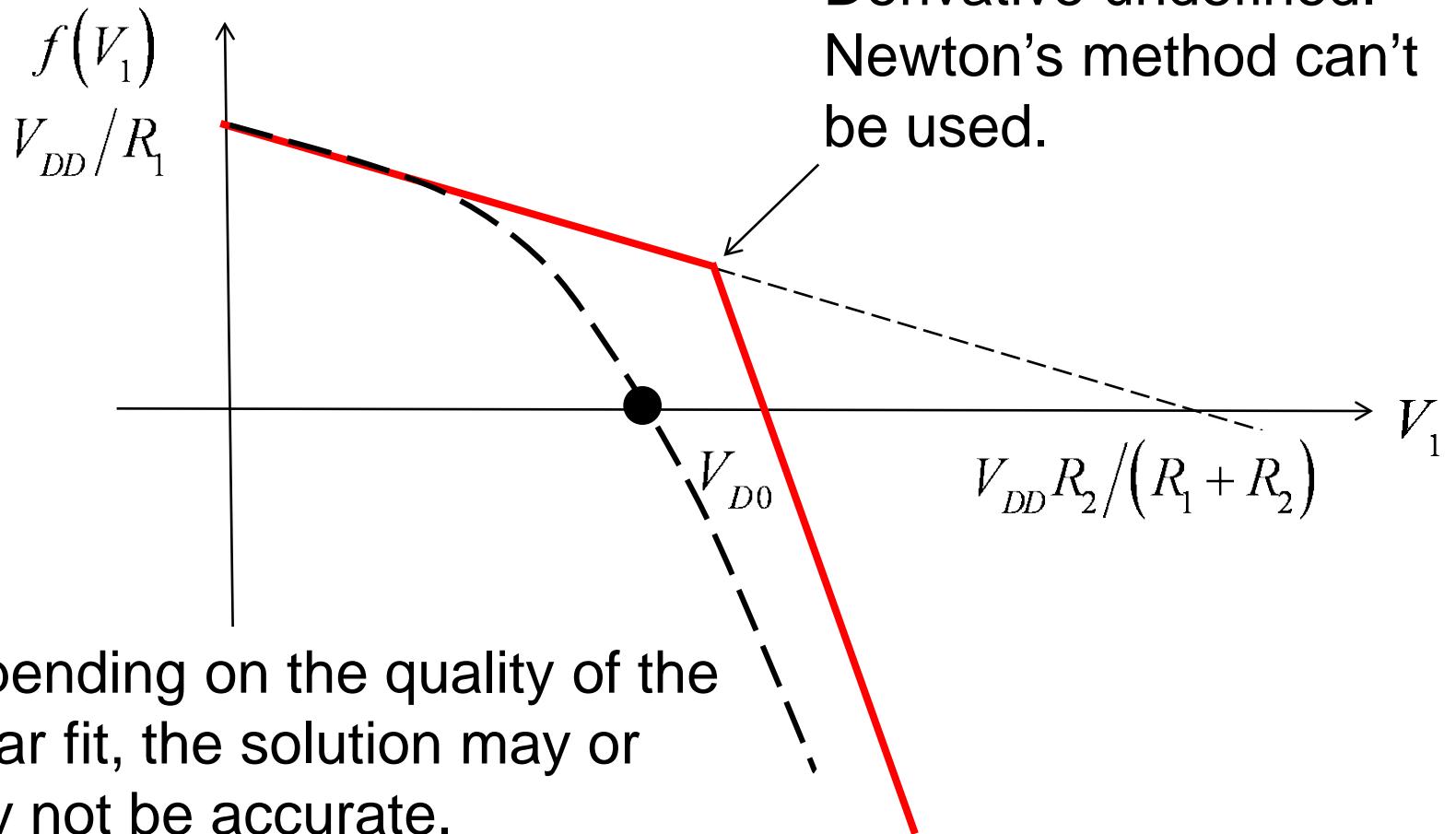
Numbers that are too large or too small for the computer to represent must be avoided.

$$V_1^2 \rightarrow \infty$$

A simpler compact model



A simpler compact model



Some requirements of a compact circuit model

Must accurately describe the electrical performance of a device for all intended applications.

- DC
- Small-signal AC
- Transient
- Noise analysis

Must do this over a wide range of biases and temperatures.

Must describe transistors of any size.

Must be computationally efficient and robust for use in numerical circuit simulation.

The need for a physics-based compact model

Physics-based model relate the needs of designers to the manufacturing process.

Generally results in the fewest number of model parameters and simplifies model calibration.

These models are suitable for statistical and predictive modeling.

In addition to a strong physical core, a compact model must satisfy certain mathematical requirements.

Mathematical requirements of a compact model

For numerical convergence, the first derivative with respect to terminal voltages must be continuous. (For small signal simulation, the second and third order derivatives must be continuous. Model with derivatives of all orders are desired.)

The model must be well-behaved outside the range of expected operating voltages.

Model simplicity is necessary to minimize solution time.

For more information

On the mathematical aspects of compact models and circuit simulation:

J. Roychowdhury, *Numerical Simulation and Modeling of Electronic and Biochemical Systems*, now Publishers, Inc., Hanover, MA, 2009.

On industrial strength compact models for MOSFETs:

Y. Tsividis and C. McAndrew, *Operation and Modeling of the MOS transistor*, 3rd ed., Oxford Univ. Press, New York, 2011.

Summary

Type 1 compact models succinctly describe the essential physics of a device.

Type 2 compact models are used by designers – to design circuits and to communicate with manufacturing.

The core of a Type 2 model should be a physics-based (Type 1) compact model.

Type 2 models must satisfy the needs of designers as well as the mathematical constraints of the circuit simulator.

Essentials of MOSFETs

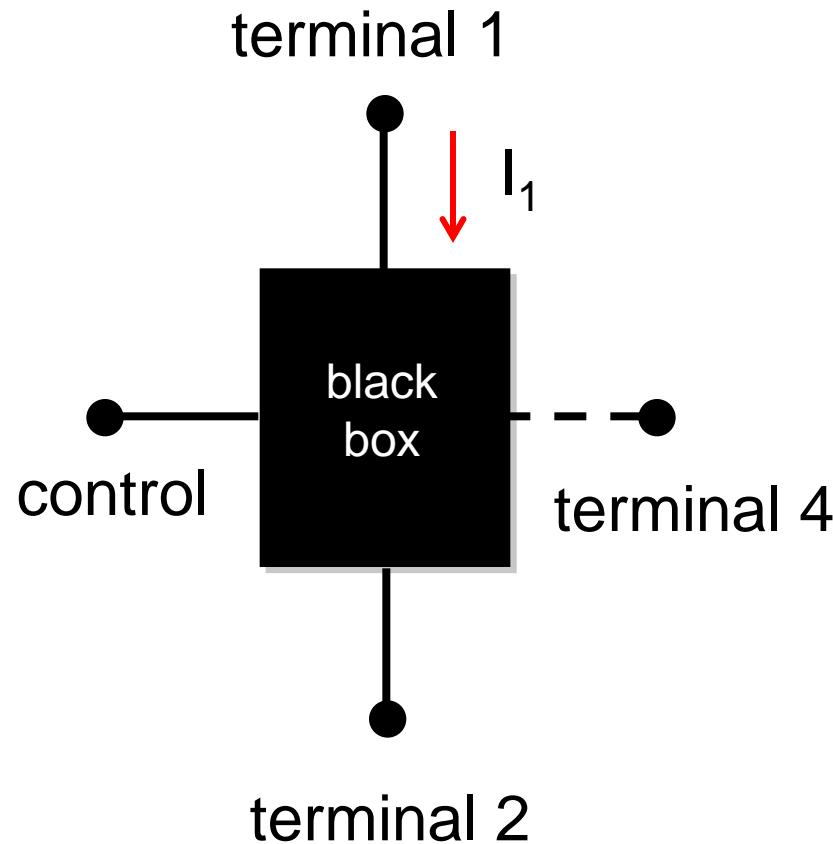
Unit 1: Transistors and Circuits

Lecture 1.6: Unit 1 Recap

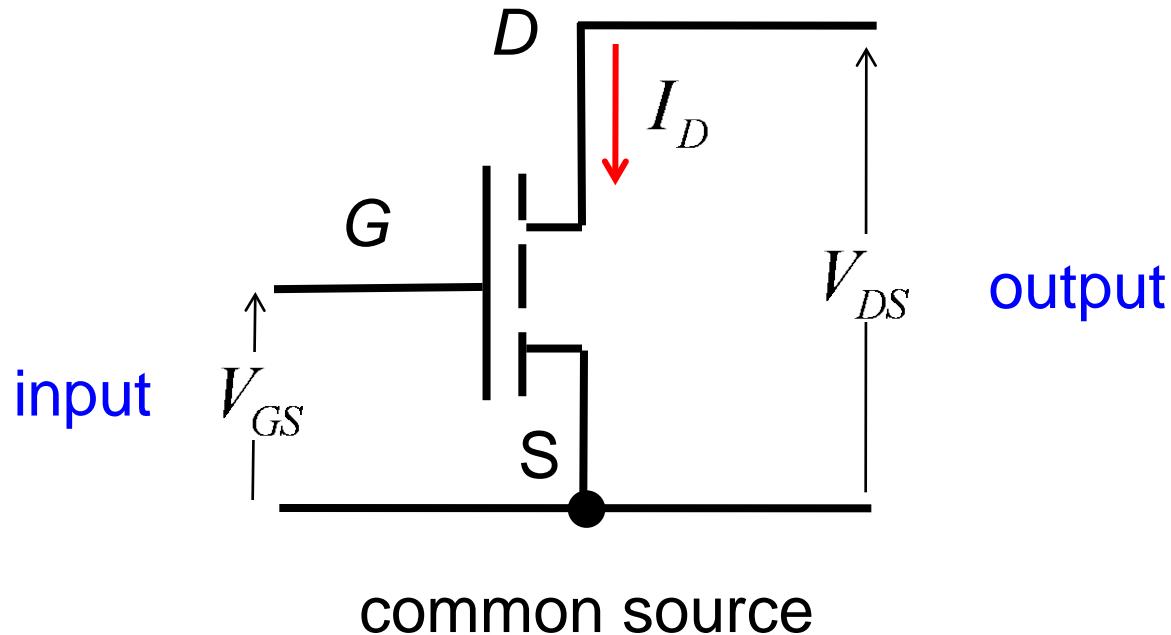
Mark Lundstrom

lundstro@purdue.edu
Electrical and Computer Engineering
Purdue University
West Lafayette, Indiana USA

The transistor as a “black box”



The MOSFET as a 2-port device

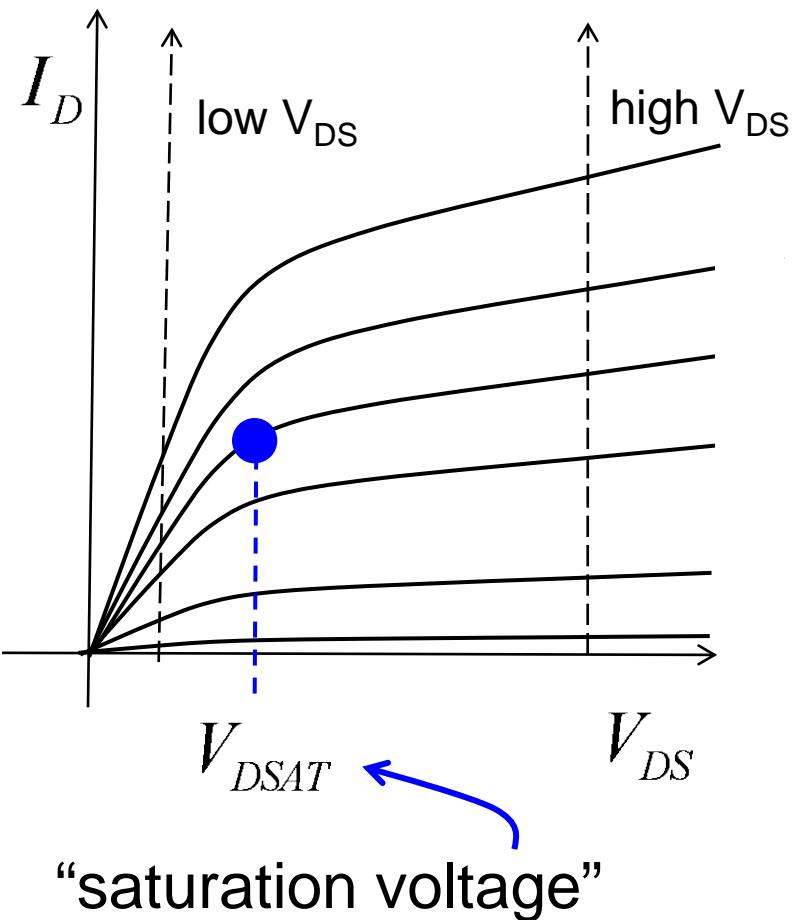


$I_D(V_{GS})$ at a fixed V_{DS} Transfer characteristics

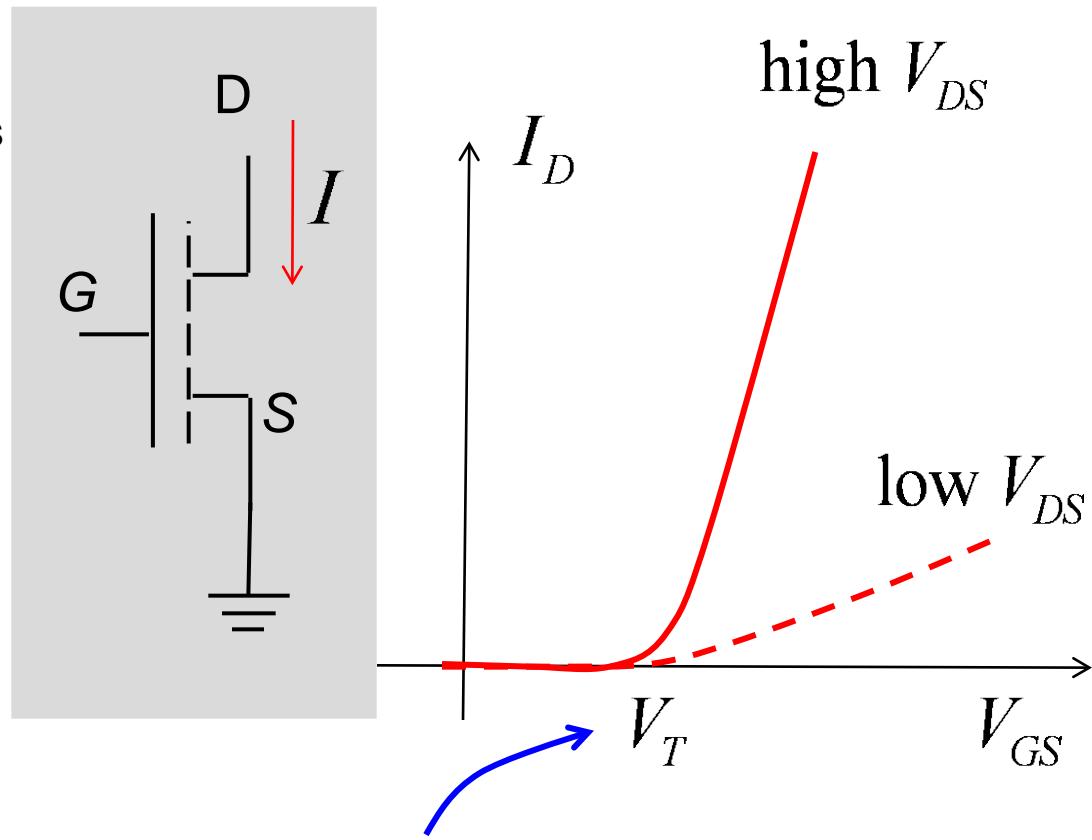
$I_D(V_{DS})$ at a fixed V_{GS} Output characteristics

Output vs. transfer characteristics

output characteristics

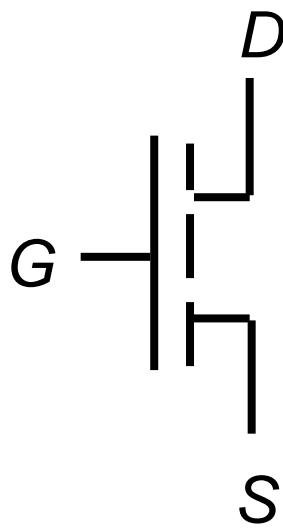


transfer characteristics



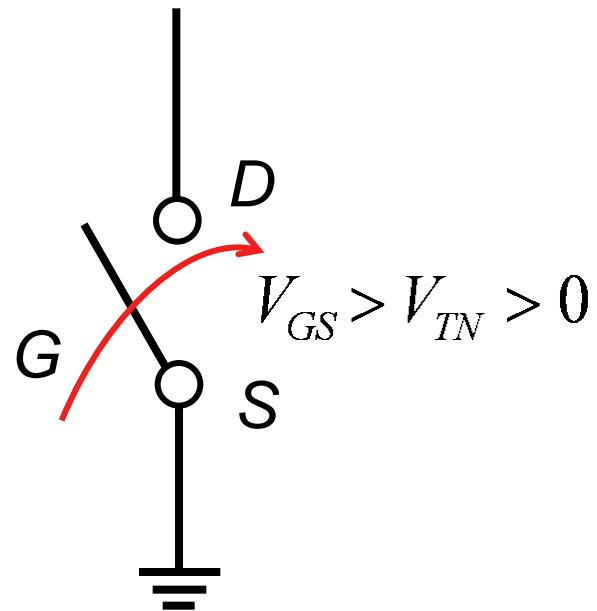
Digital electronics

N-MOSFET
symbol

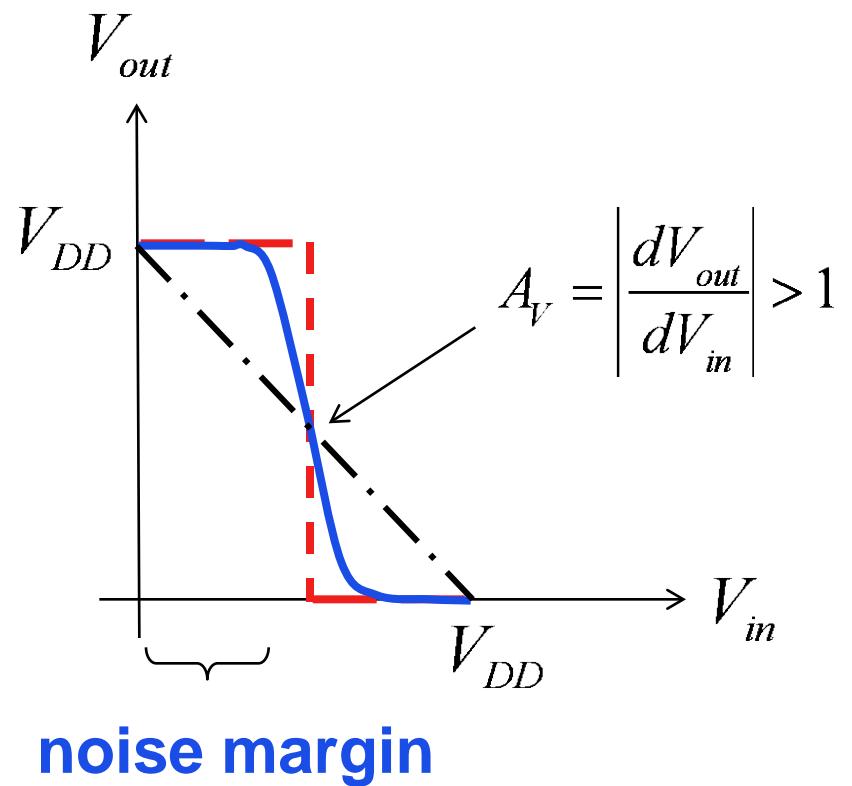
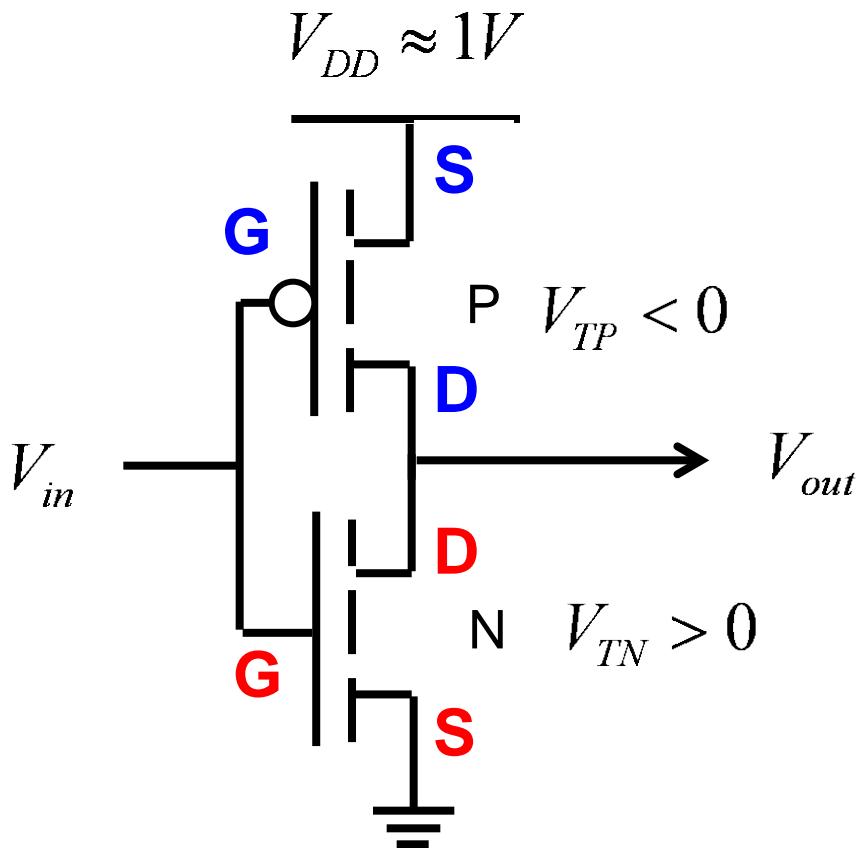


switch

$$V_{DS} > 0$$



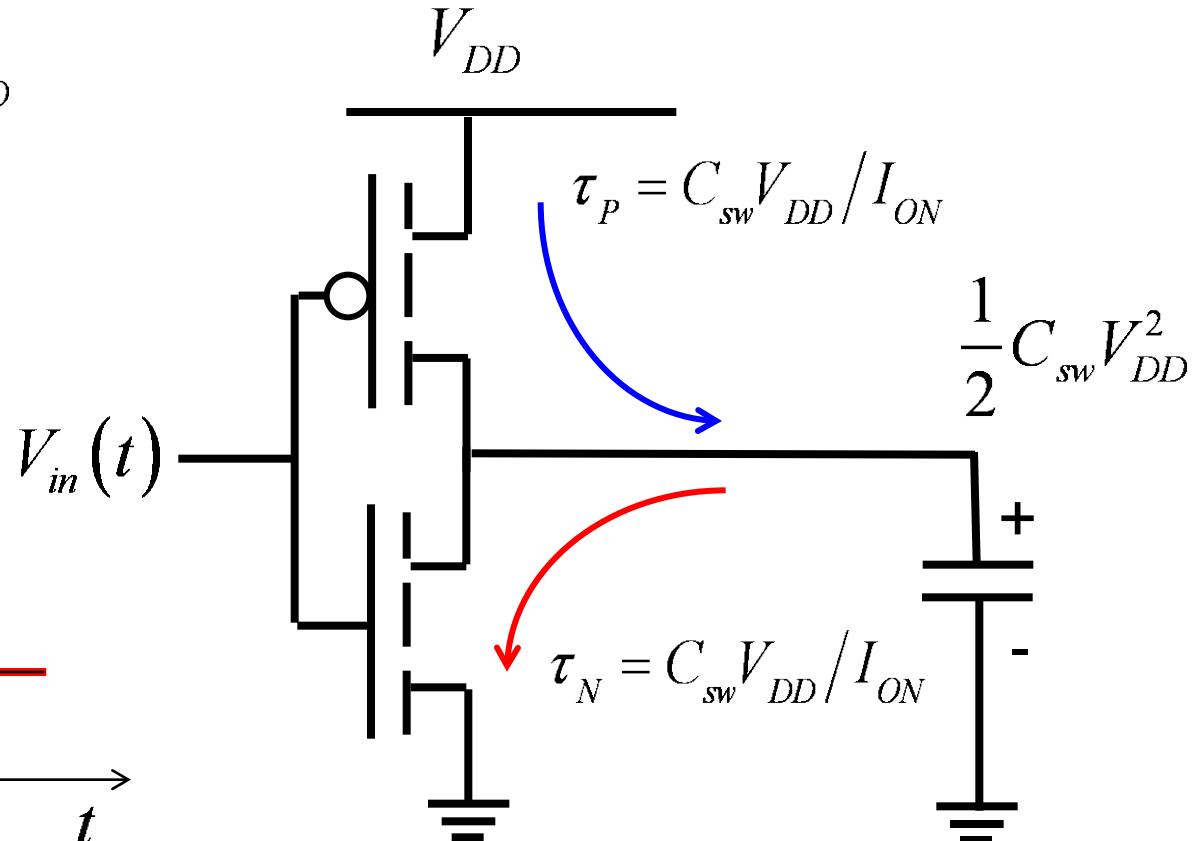
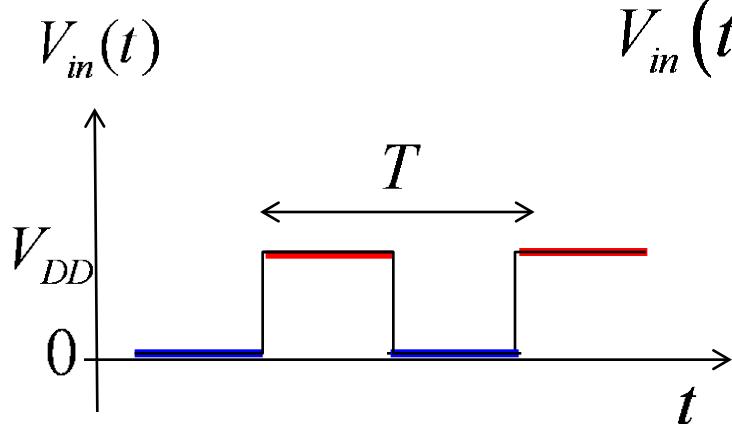
CMOS Inverter



Switching

$$P_{\text{dynamic}} = \alpha f C_S V_{DD}^2$$

$$P_{\text{static}} = N_G I_{OFF} V_{DD}$$



CMOS speed and power

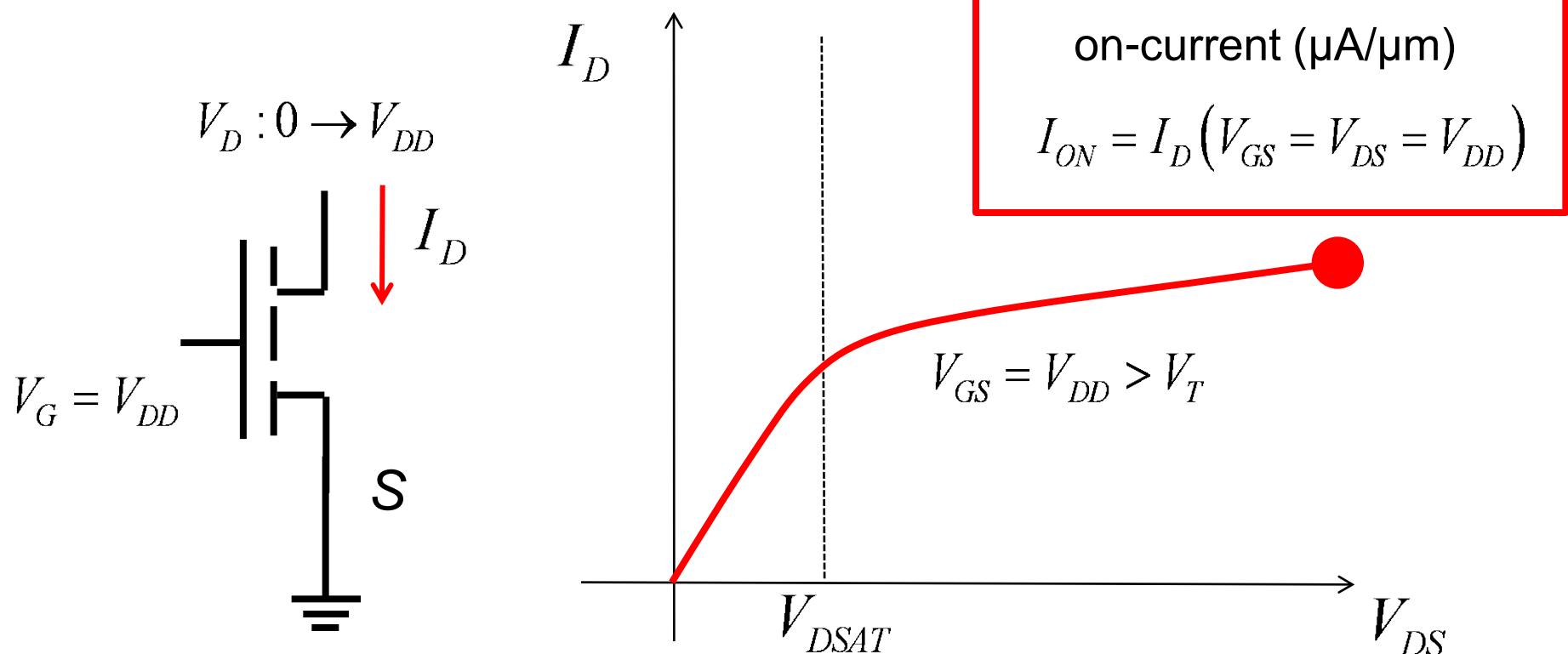
$$\tau = C_{sw} V_{DD} / I_{ON} \quad P_{\text{dynamic}} = \alpha f C_s V_{DD}^2 \quad P_{\text{static}} = N_G I_{OFF} V_{DD}$$

- 1) Higher on-current means higher speed
- 2) Faster operation means more dynamic power
- 3) Lower V_{DD} means lower power
- 4) More leakage means more power dissipation

Key Figures of Merit for digital applications

- 1) On current
- 2) Off-current
- 3) Subthreshold swing
- 4) DIBL

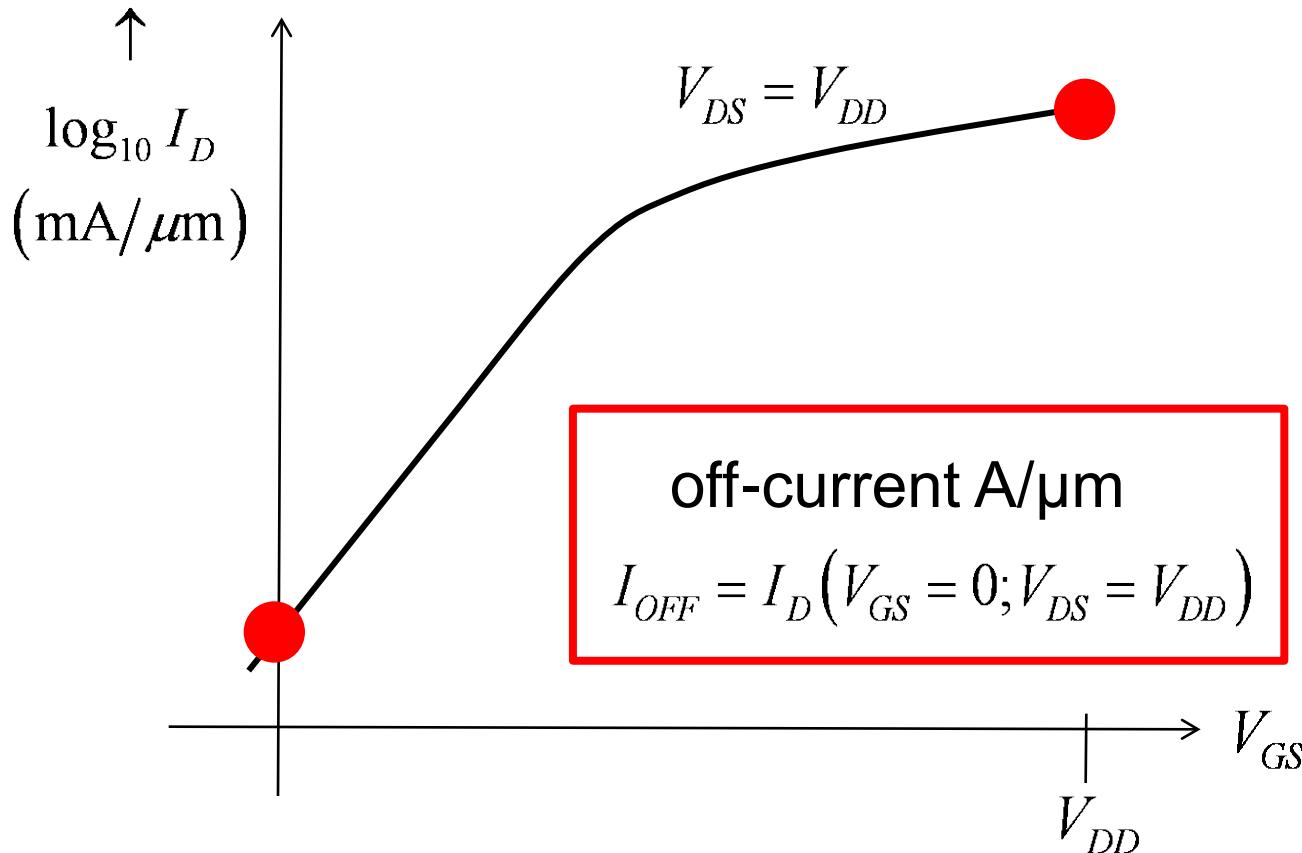
1) On-current



output characteristic:
 I_D vs. V_{DS} at fixed V_{GS}

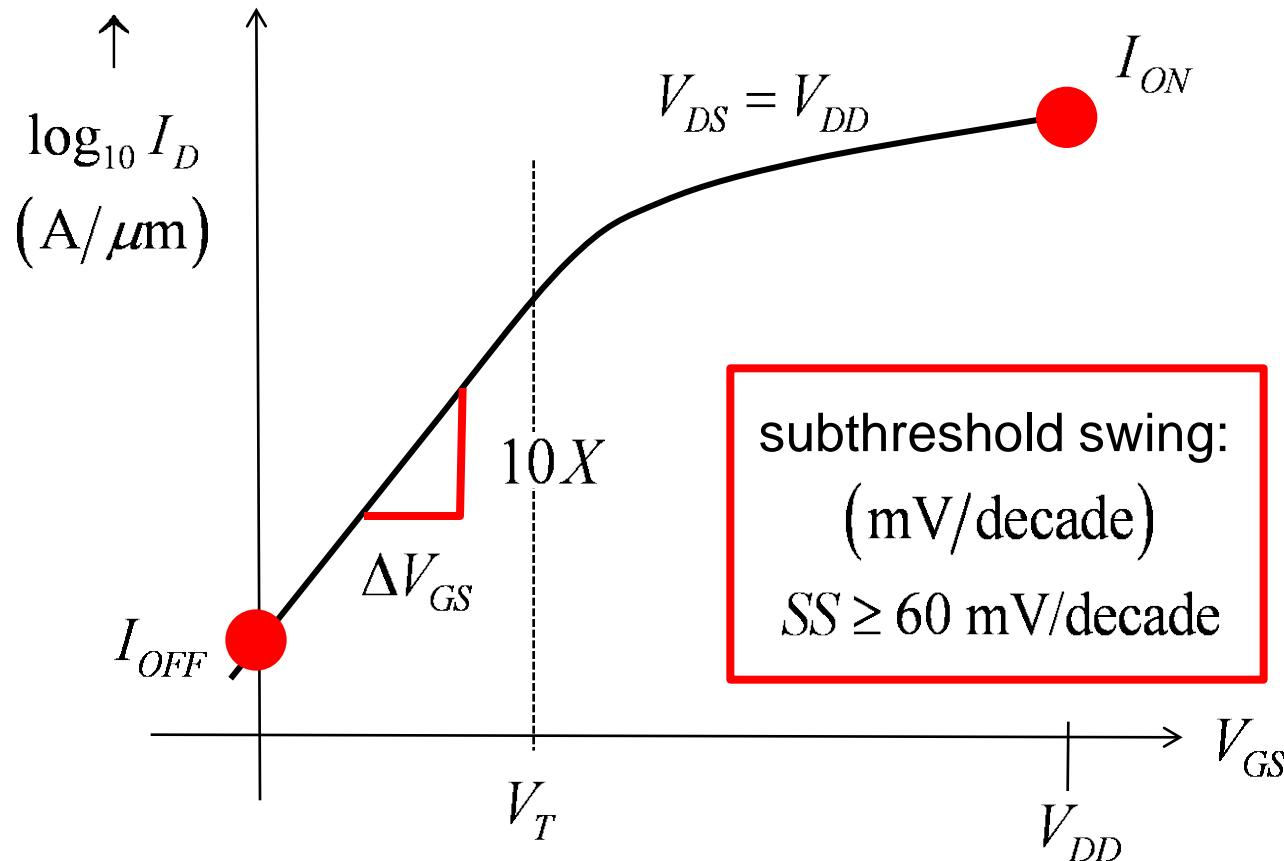
2) Off-current

transfer characteristic on a semi-log plot:



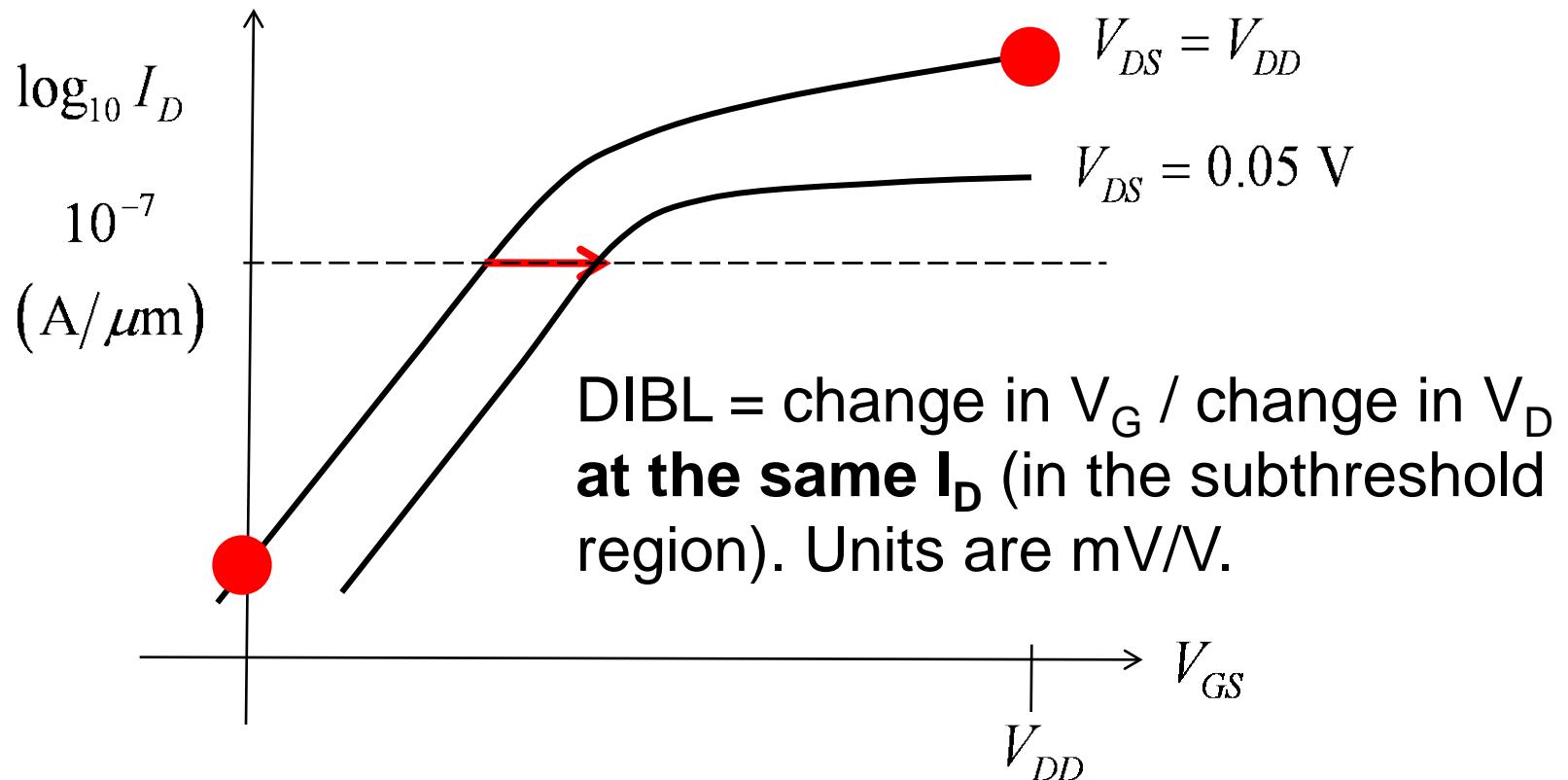
3) The Subthreshold Swing

transfer characteristics:

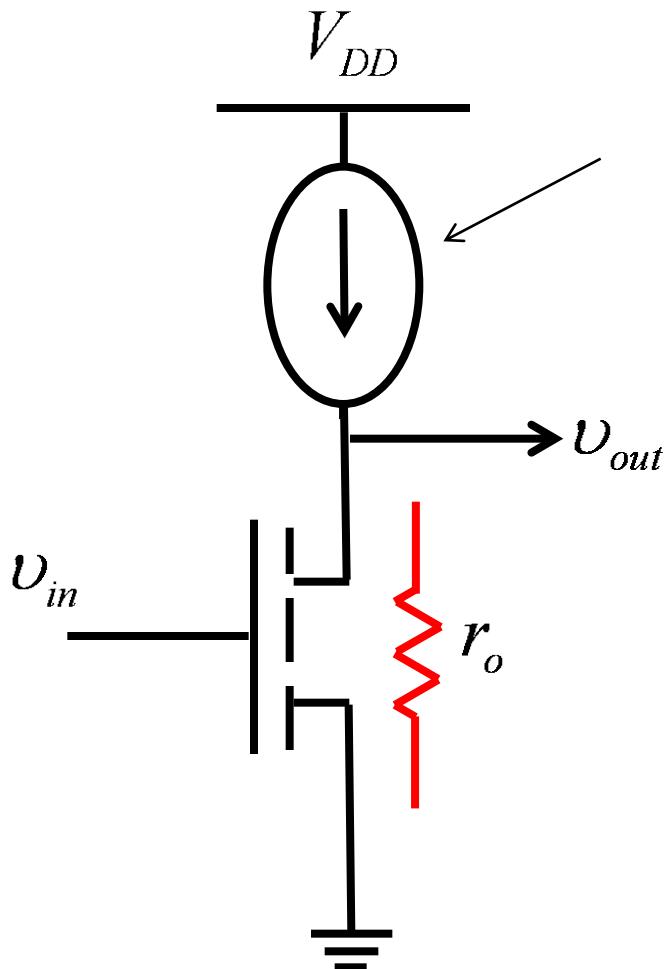


4) DIBL

transfer characteristics:



Analog / RF



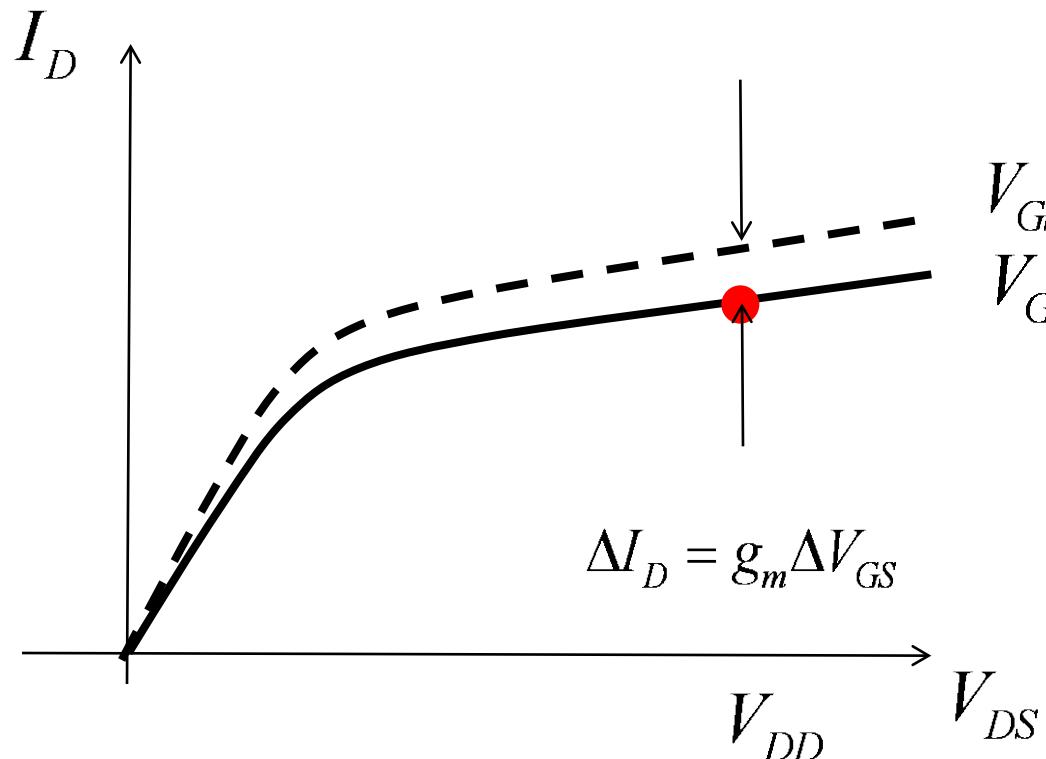
high impedance current source

$$A_v(\text{max}) = -g_m r_0$$

The **transconductance**, g_m , and **output resistance**, r_0 , of the MOSFET are important figures of merit.

So is the **self gain**, $g_m r_0$.

Transconductance

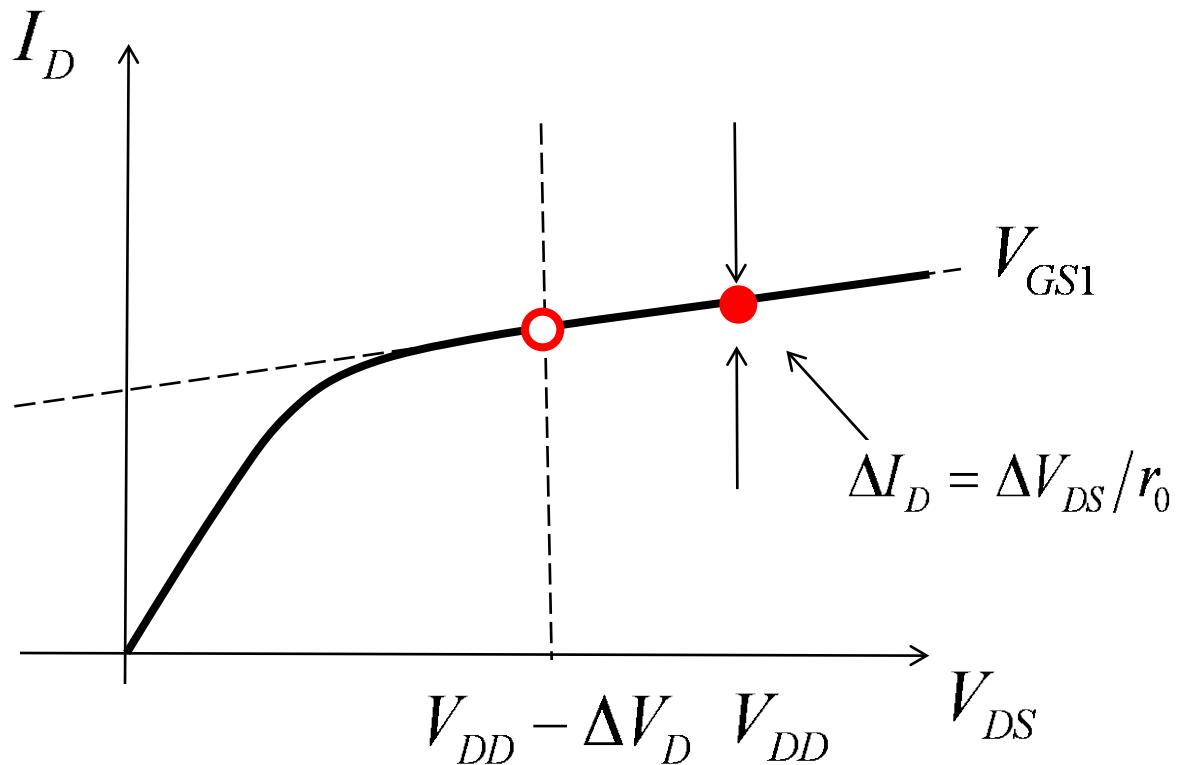


$$\begin{aligned} &V_{GS1} + \Delta V_{GS} \\ &V_{GS1} \end{aligned}$$

$$g_m \equiv \left. \frac{\partial I_D}{\partial V_{GS}} \right|_{V_{DS}}$$

$(\mu\text{S}/\mu\text{m})$

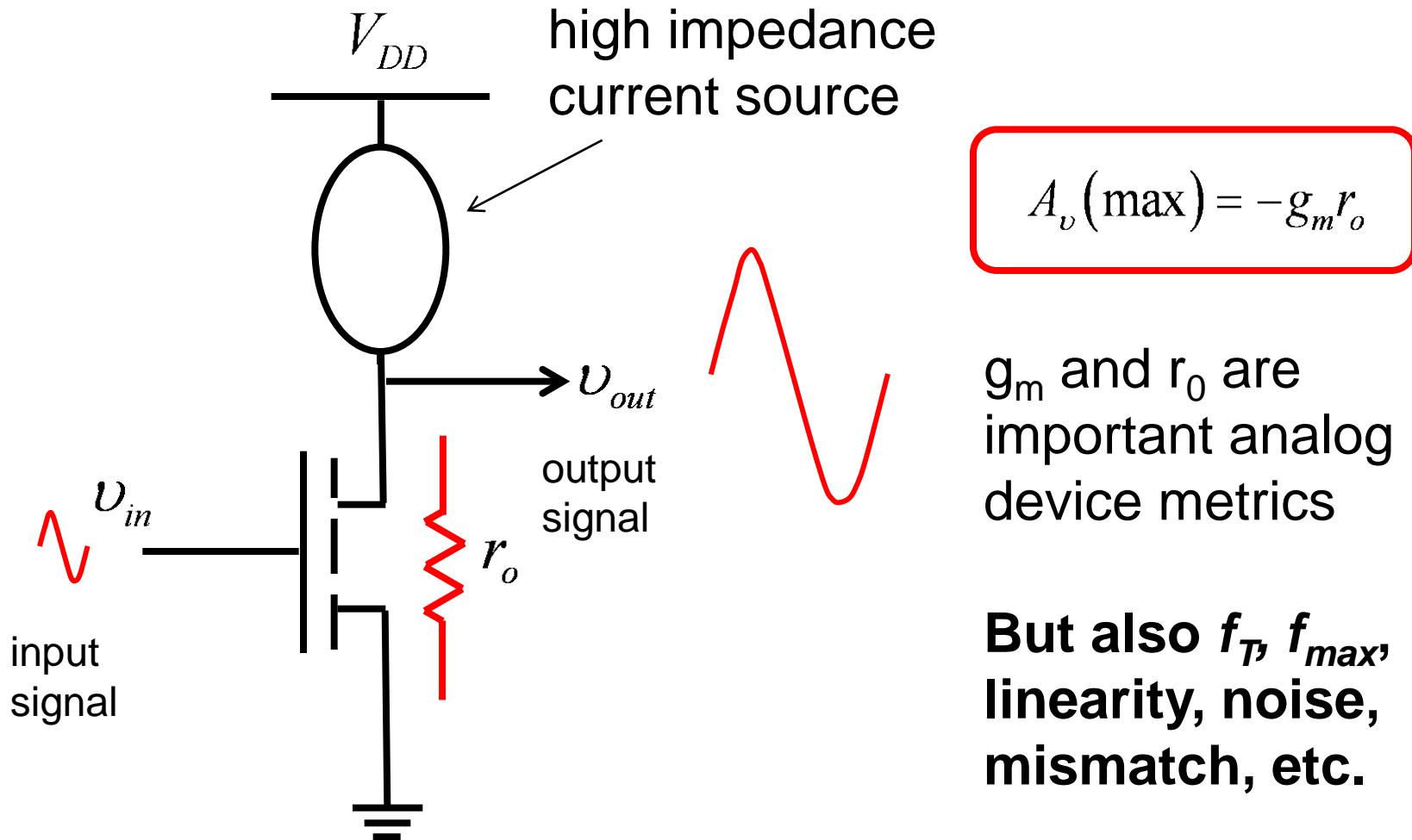
Output resistance



$$r_0 \equiv \left(\frac{\partial I_D}{\partial V_{DS}} \Big|_{V_{GS}} \right)^{-1}$$

$\Omega - \mu\text{m}$

Other analog device metrics



Recap

Given the measured characteristics of a MOSFET, you should be able to determine:

1. on-current: I_{ON}
2. off-current: I_{OFF}
3. subthreshold swing, SS
4. drain induced barrier lowering: DIBL

5. output resistance: r_o
6. *transconductance*: g_m

threshold voltage: $V_T(\text{lin})$ and $V_T(\text{sat})$

drain saturation voltage: V_{DSAT}

Our goal in this course is to understand these device metrics and parameters.

Two types of compact models

In the course, we will distinguish between two different types of compact models:

1) Compact physical models

These models aim to describe a device in terms of a few parameters with strong physical significance. These kinds of models are useful for device characterization, process monitoring, and for the conceptual understanding that guides device research.

Our focus in this course is on this type of compact model.

Second type of compact models

2) Compact device models for circuit simulation

These models accurately relate the currents that flow into a device's terminals to the voltages on the leads **in a form suitable for use in numerical circuit simulation programs**. To describe everything relevant to a circuit, these models are more complex, but the core of the model is usually a compact physics model.

These kinds of compact models play a critical role in connecting semiconductor R&D and manufacturing to product design.

Summary

Be sure that you are familiar with the shape of a transistor IV characteristic (linear and saturation regions, sub-threshold and above threshold, output and transfer characteristics, etc.)

You should also be familiar with the key metrics for digital applications (on- and off-currents, SS, DIBL) and for analogue applications (and how to extract them from measured IV characteristics).

Next, in Unit 2, we will start looking inside the transistor “black box”.

Essentials of MOSFETs

Unit 2: Essential Physics of the MOSFET

Lecture 2.1: Energy Band Diagram Review

Mark Lundstrom

lundstro@purdue.edu
Electrical and Computer Engineering
Purdue University
West Lafayette, Indiana USA

Energy band diagrams

An energy band diagram is a plot of the bottom of the conduction band and the top of the valence band vs. position.

Energy band diagrams are a powerful tool for understanding semiconductor devices because they provide **qualitative solutions to the semiconductor equations**.

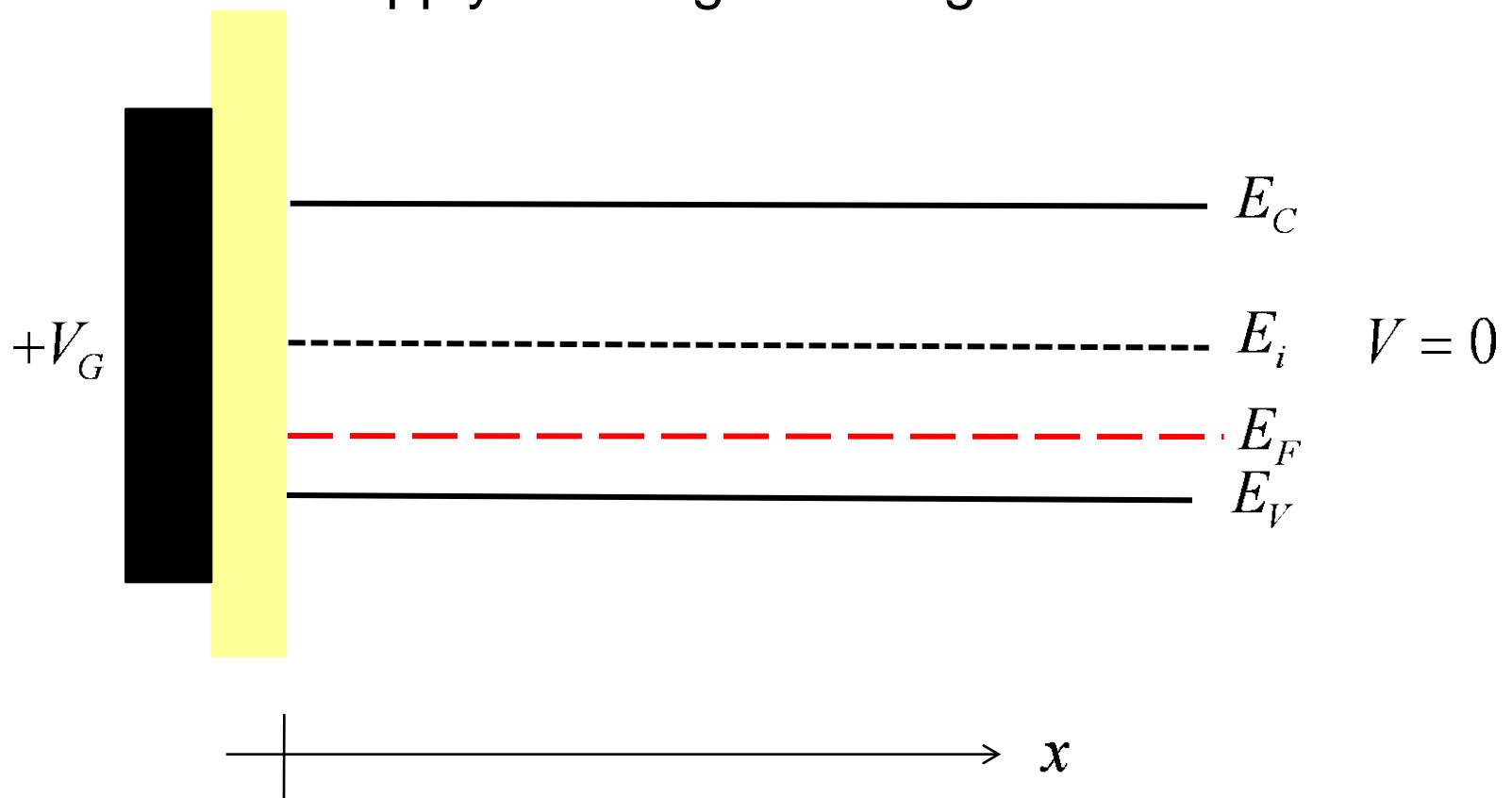
The Fermi level in equilibrium

The Fermi level is constant in equilibrium.

$$J_n = n\mu_n \frac{dF_n}{dx} = 0 = n\mu_n \frac{dE_F}{dx} \rightarrow E_F \text{ is constant}$$

Band bending

What happens when we apply a voltage to the gate?



Voltage and electron potential energy

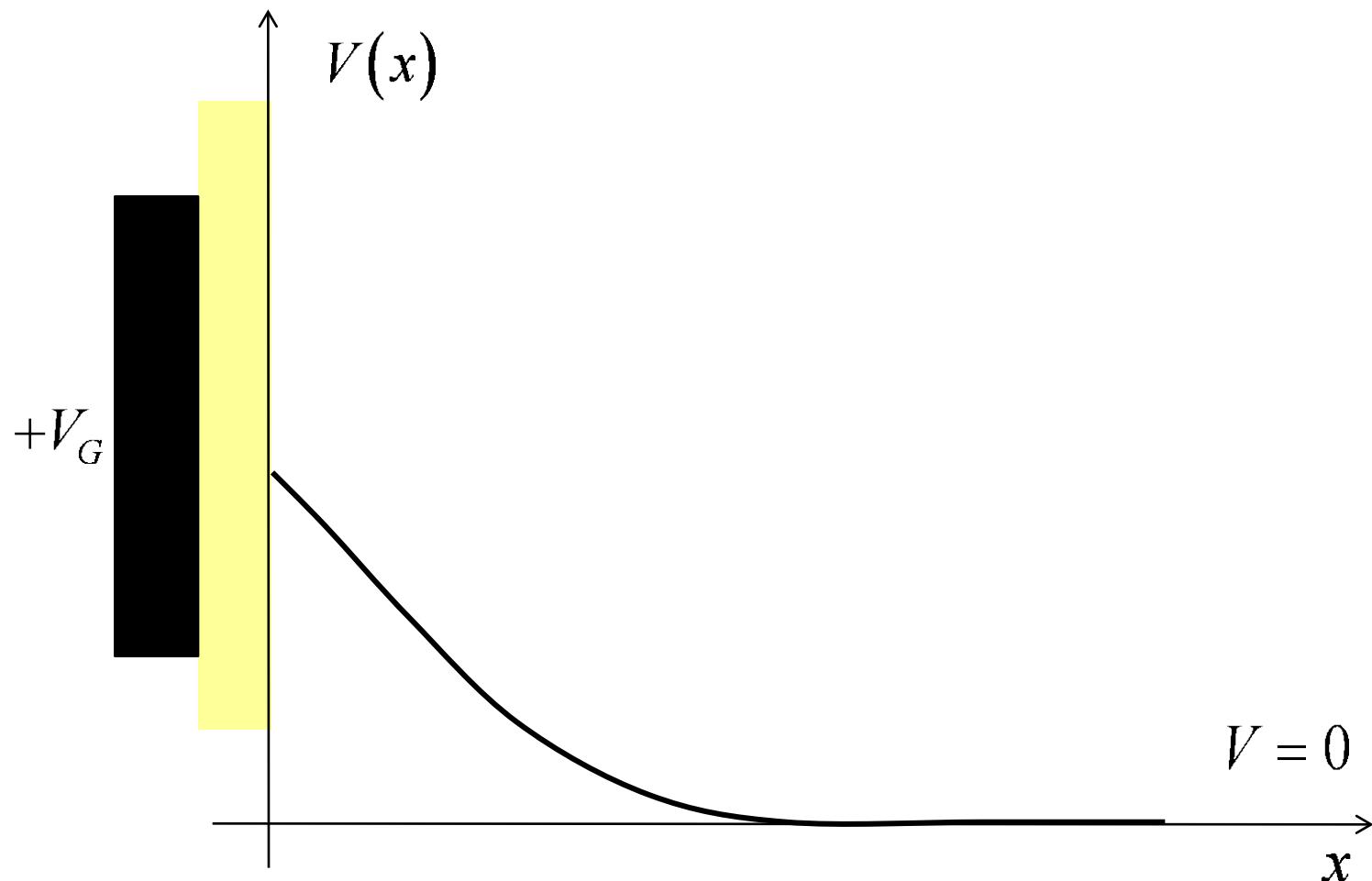
$$E = -qV$$



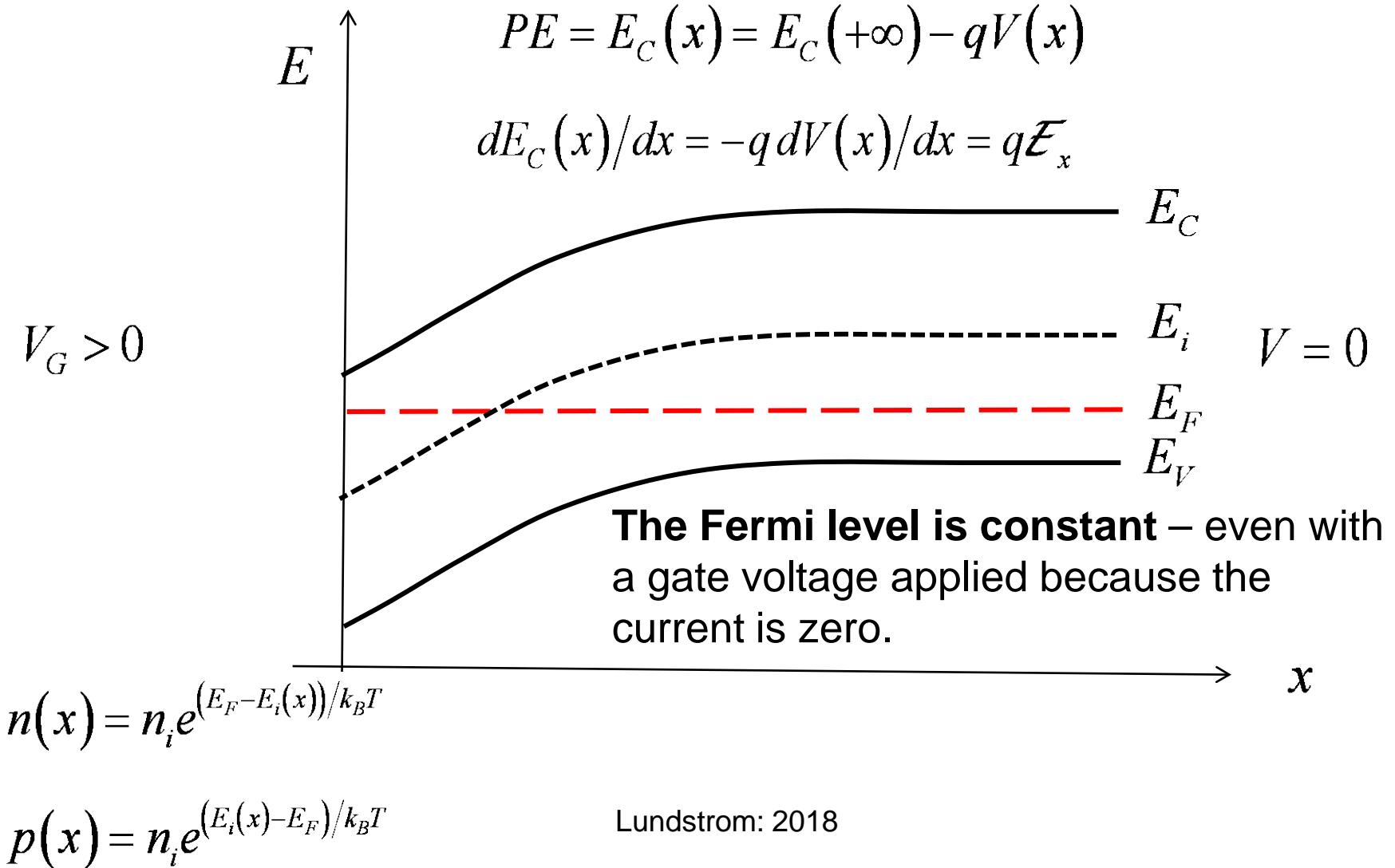
$$+V$$

A positive potential **lowers** the energy of an electron.

Electrostatic potential vs. position

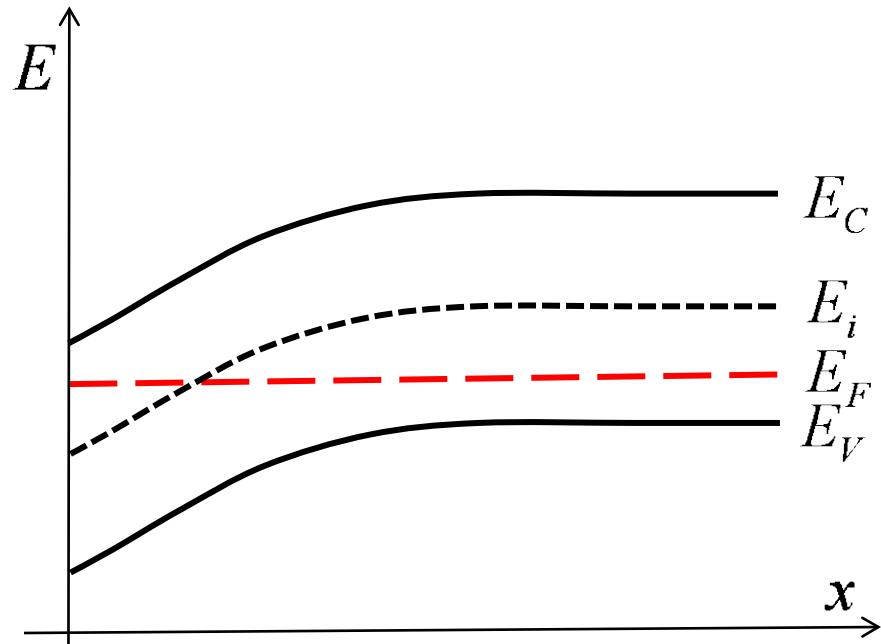


Electrostatic potential causes band bending



Band diagrams

1) Draw the band diagram



$$\frac{d\mathcal{E}}{dx} = \frac{\rho(x)}{K_S \epsilon_0}$$

2) Read the band diagram

$$V(x) \propto -E_C(x)$$

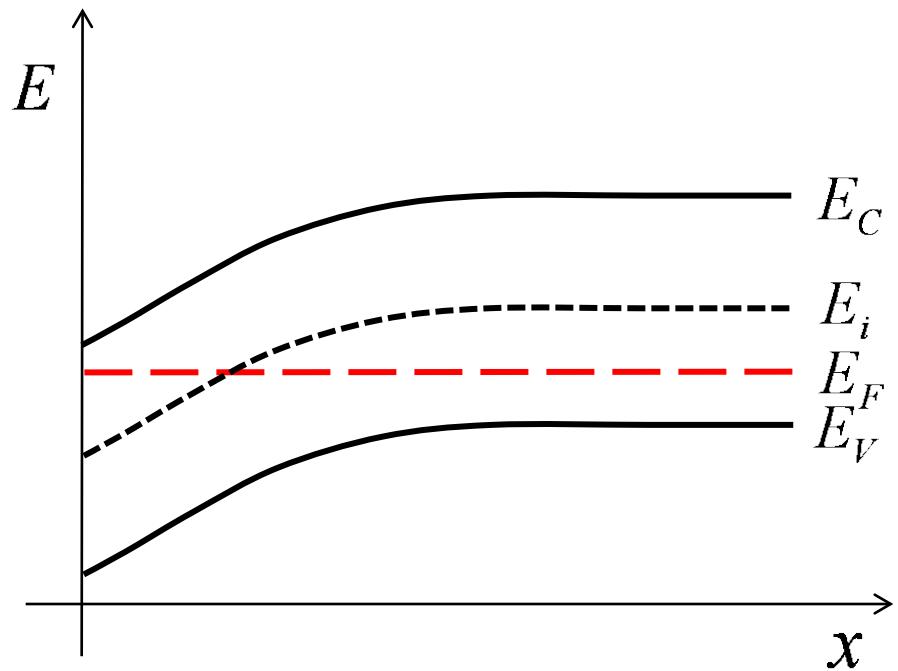
$$\mathcal{E} \propto dE_C(x)/dx$$

$$\log n(x) \propto E_F - E_i(x)$$

$$\log p(x) \propto E_i(x) - E_F$$

$$\rho(x) \propto d^2E_C(x)/dx^2$$

Practice

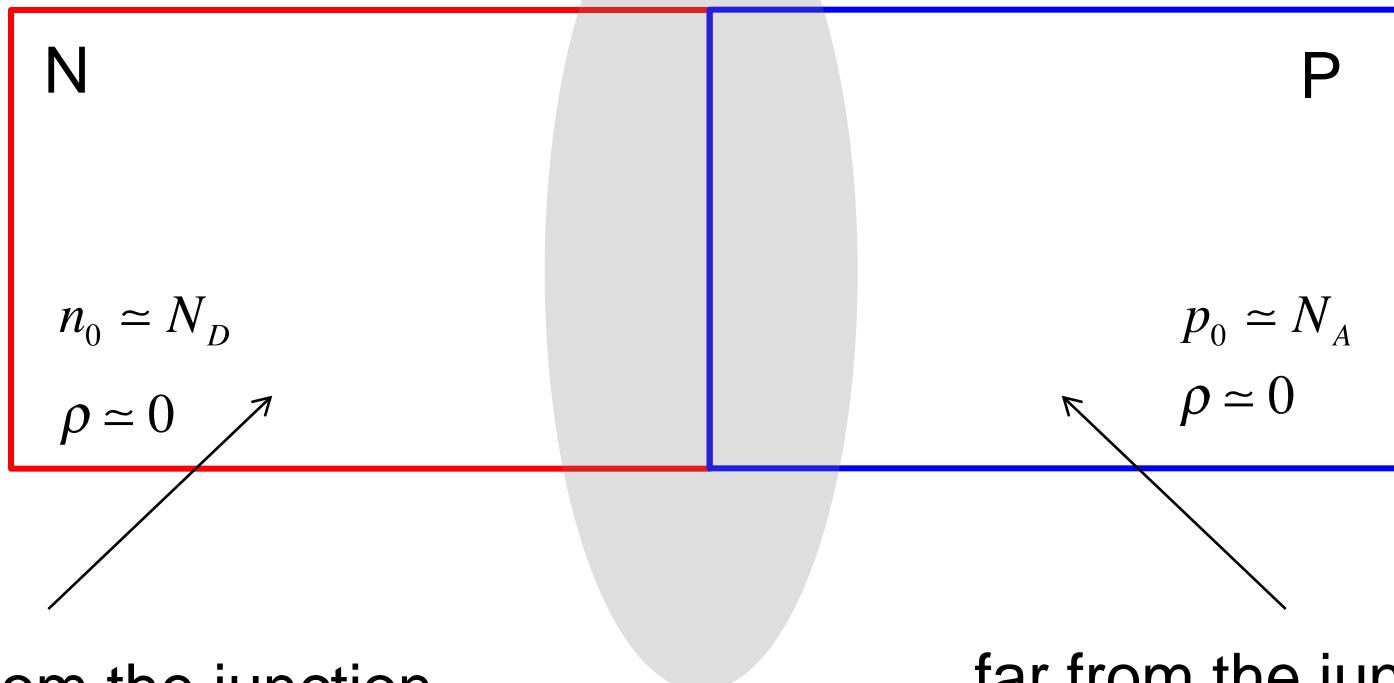


Sketch vs. position:

- Electrostatic potential
- Electric field
- Electron density
- Hole density
- Space charge density

Another example: NP junction in equilibrium

the bands will bend
near the junction



far from the junction,
the bands will be flat

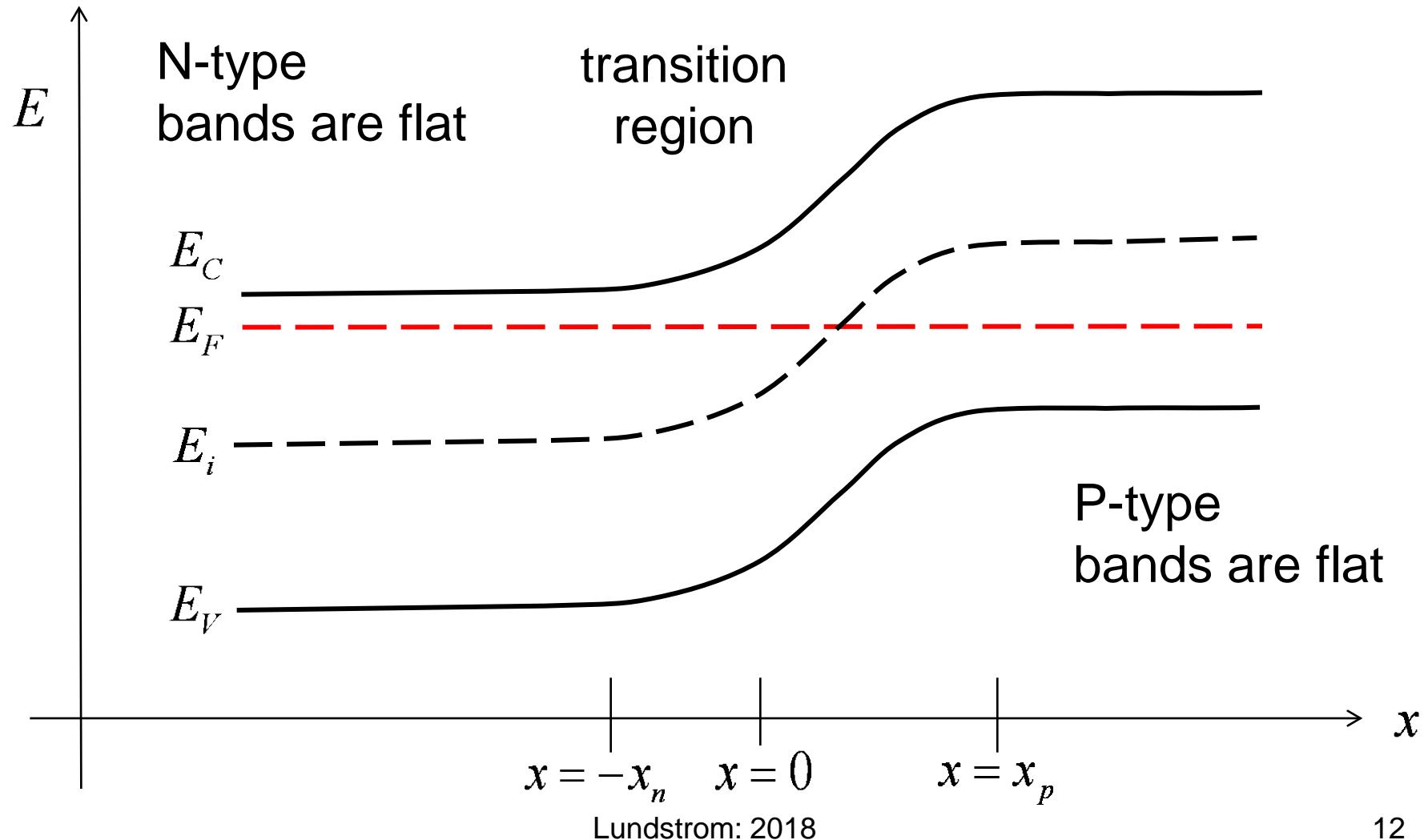
far from the junction,
the bands will be flat

Procedure: Equilibrium energy band diagram



- 1) Begin with E_F
- 2) Draw the E-bands where you know the carrier density then connect the two regions.
- 3) Then “read” the energy band diagram to obtain the electrostatic potential, electric field, carrier densities, and space charge density vs. position.

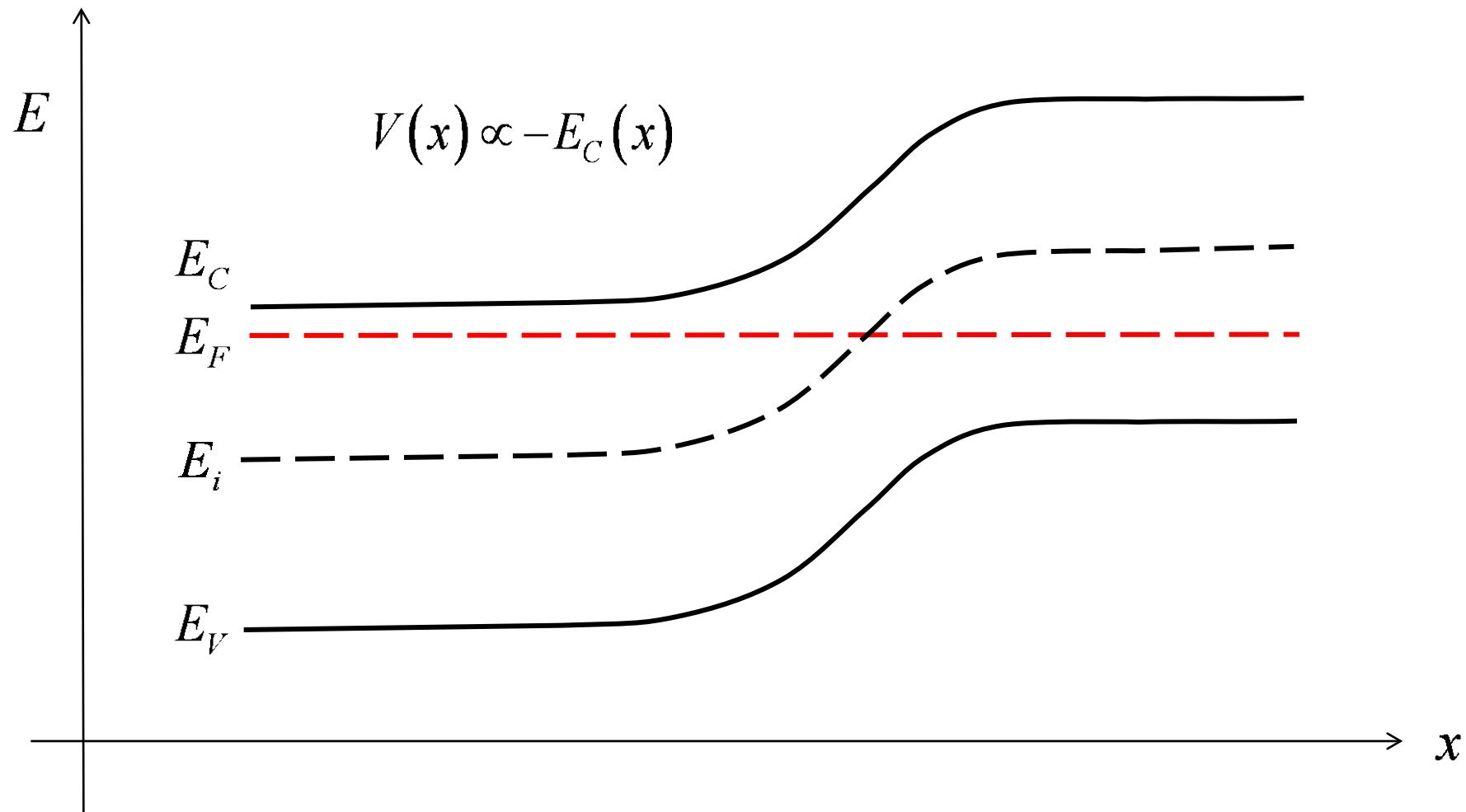
Energy band diagram



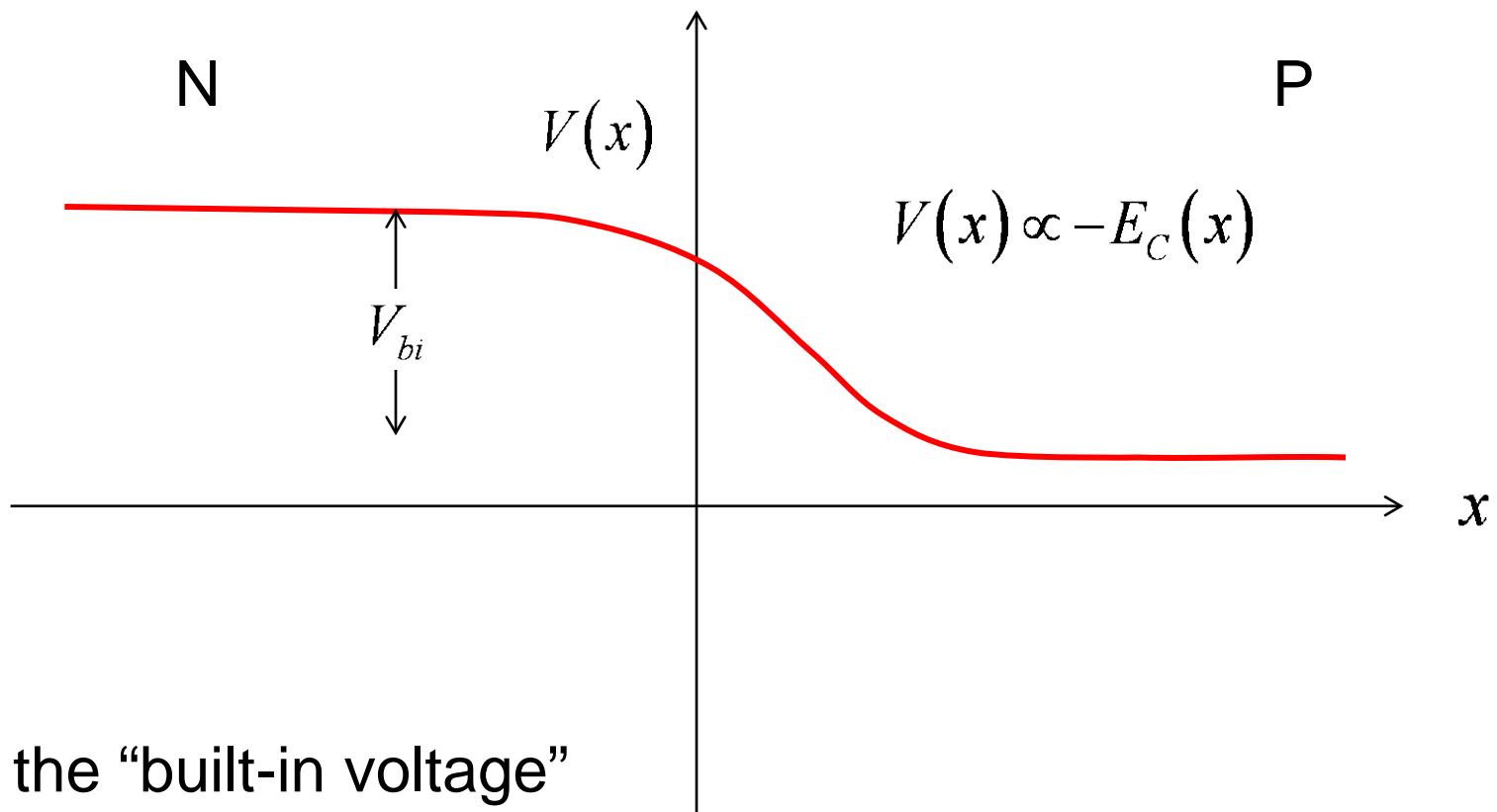
Now, “read” the e-band diagram

- 1) Electrostatic potential vs. position
- 2) Electric field vs. position
- 3) Electron and hole densities vs. position
- 4) Space-charge density vs. position

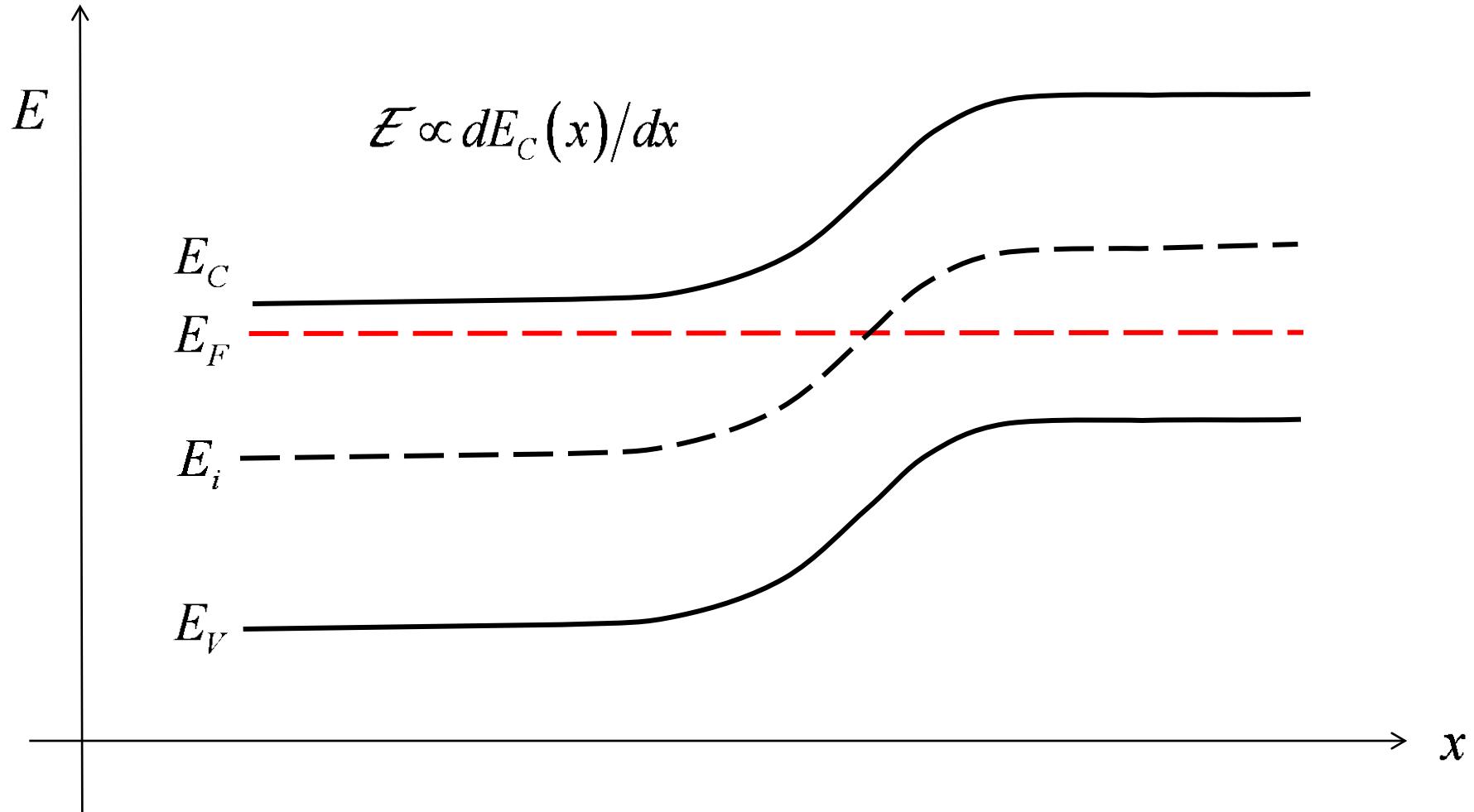
1) Electrostatic potential?



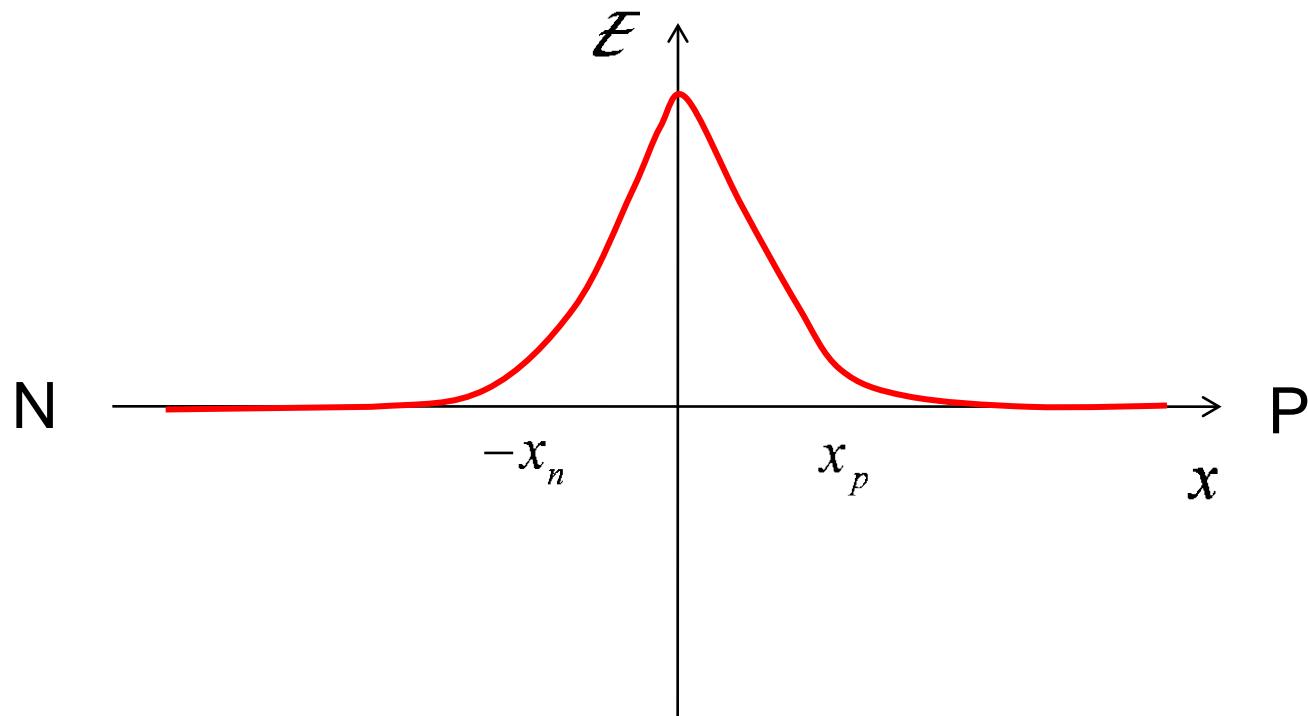
Electrostatics: $V(x)$



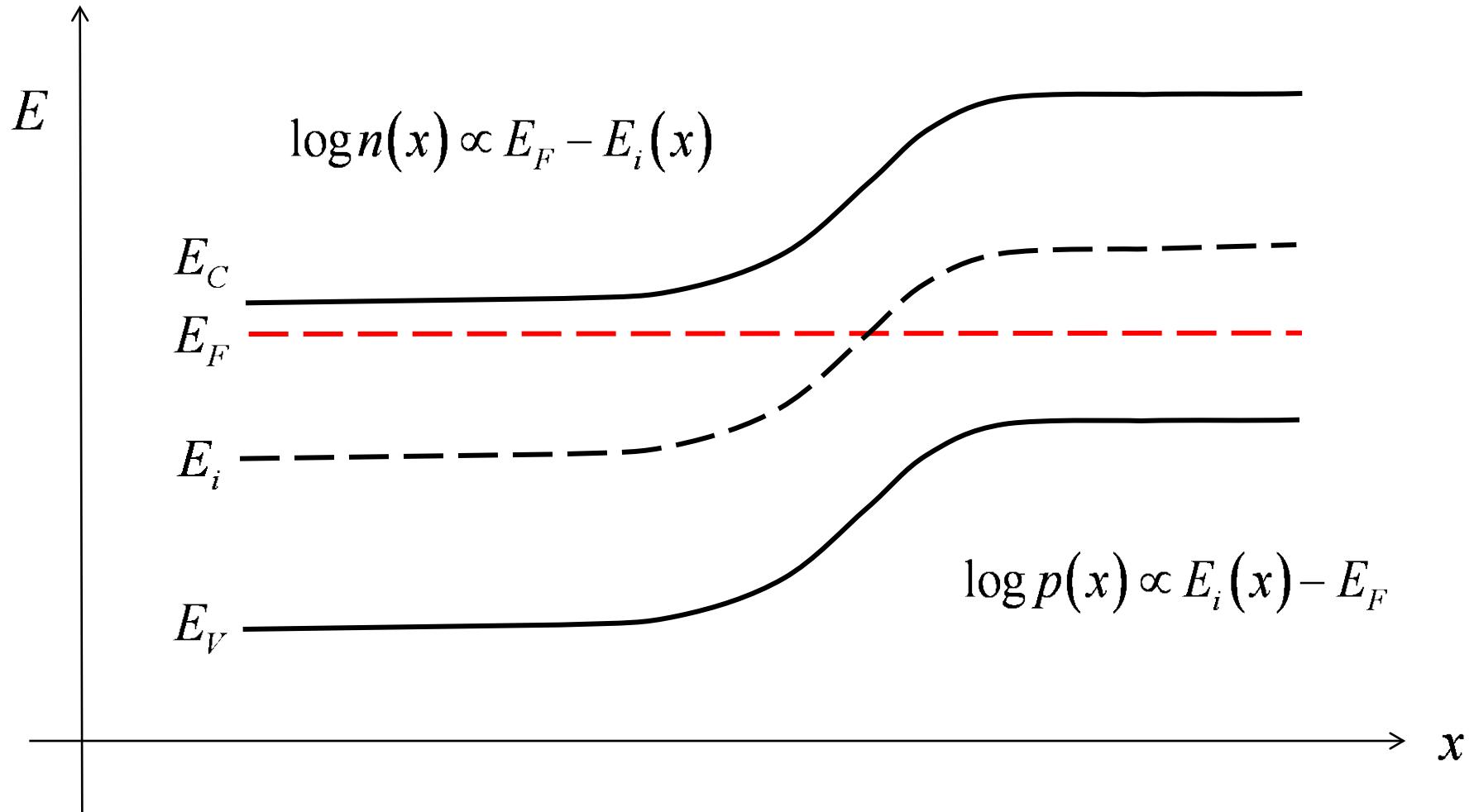
2) Electric field?



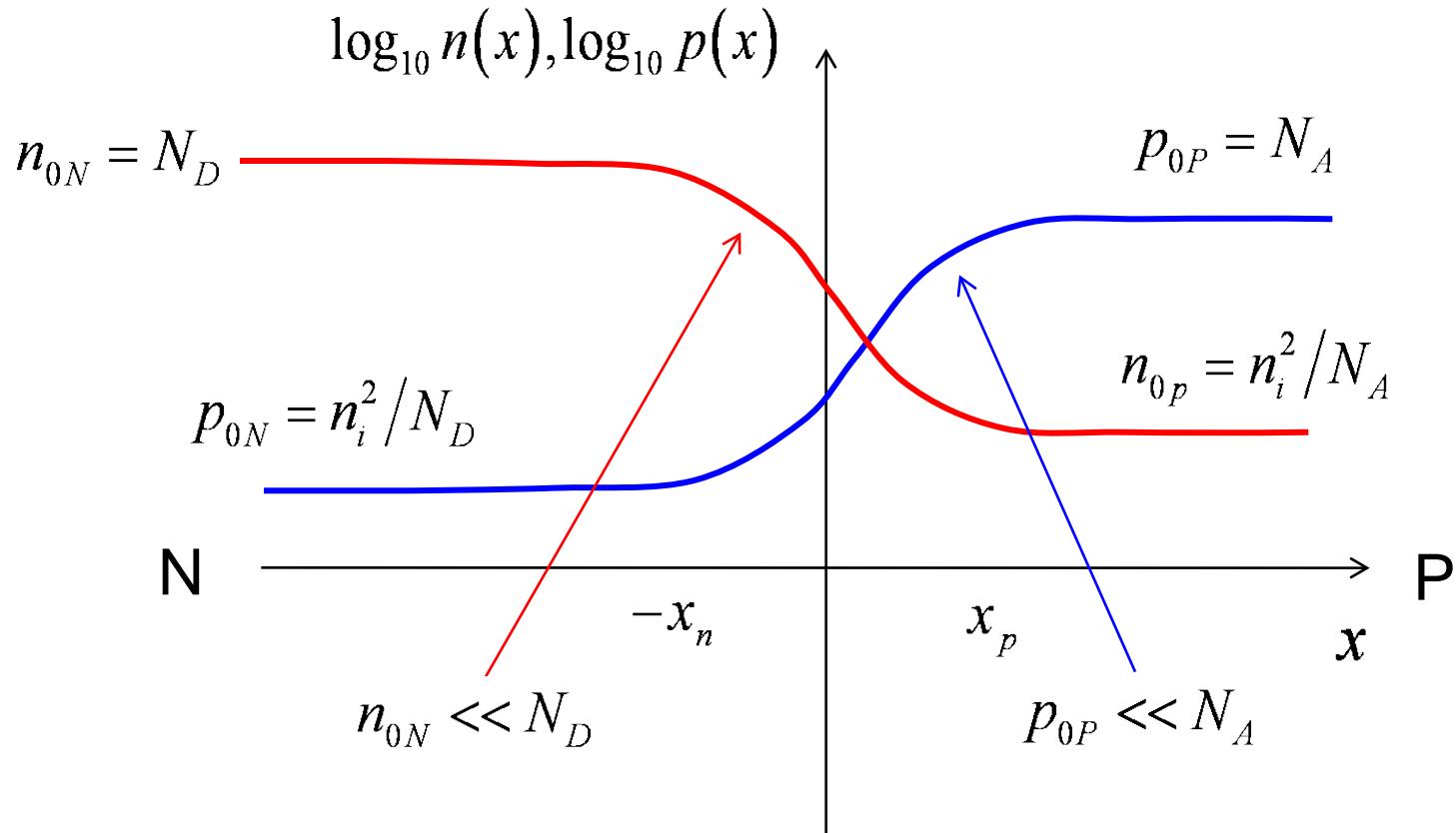
Electric field: $\mathcal{E}(x)$



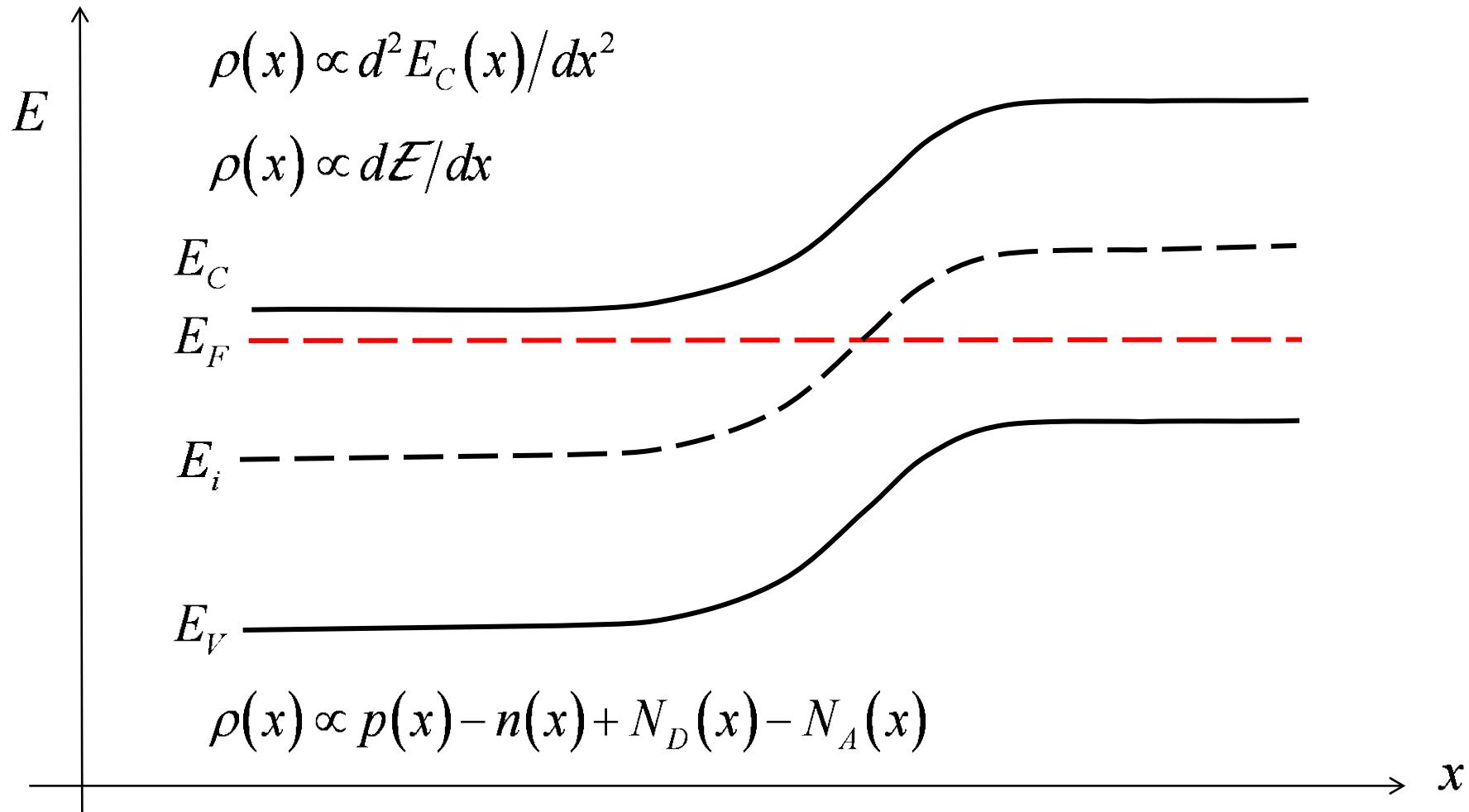
3) Carrier densities?



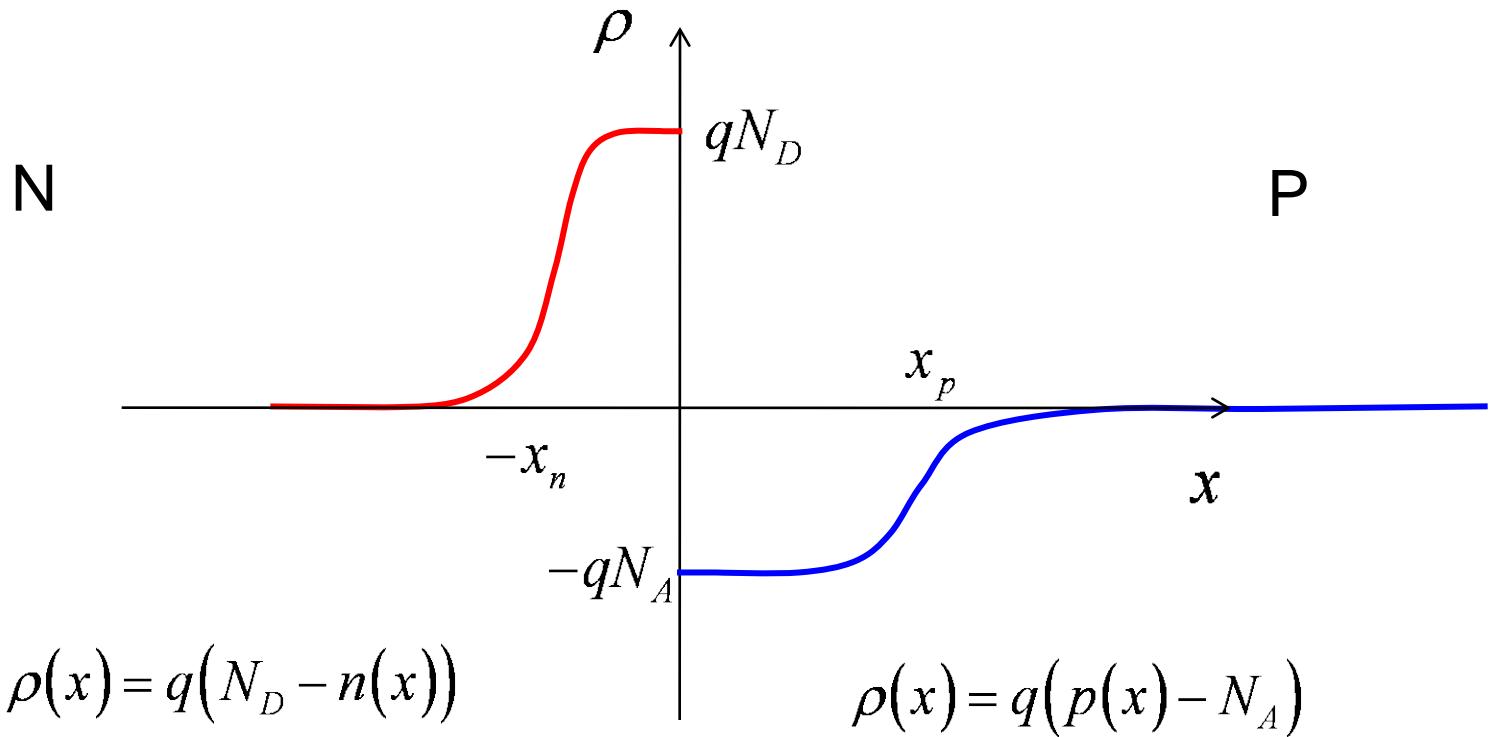
Carrier densities vs. x



4) Space charge density?



Electrostatics: $\rho(x)$



Summary

$$\frac{\partial p}{\partial t} = -\nabla \cdot \left(\frac{\vec{J}_p}{q} \right) + G_p - R_p$$

$$\frac{\partial n}{\partial t} = -\nabla \cdot \left(\frac{\vec{J}_n}{-q} \right) + G_n - R_n$$

$$\nabla \cdot (K_s \epsilon_0 \vec{\mathcal{E}}) = \rho$$

Three coupled, nonlinear PDE's in three unknowns:

$$p(\vec{r}), n(\vec{r}), V(\vec{r})$$

Drawing and then reading an E-band diagram gives us a qualitative solution to these equations.

Next topic:

In the next lecture, we will use energy band diagrams to develop a qualitative understanding of MOSFET IV characteristics.

Essentials of MOSFETs

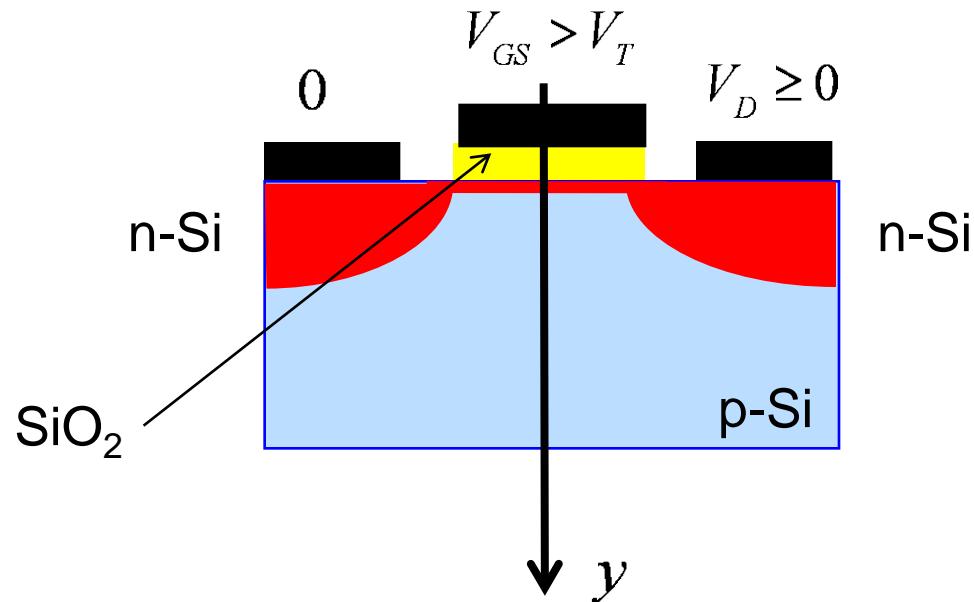
Unit 2: Essential Physics of the MOSFET

Lecture 2.2: Energy Band View of the MOSFET

Mark Lundstrom

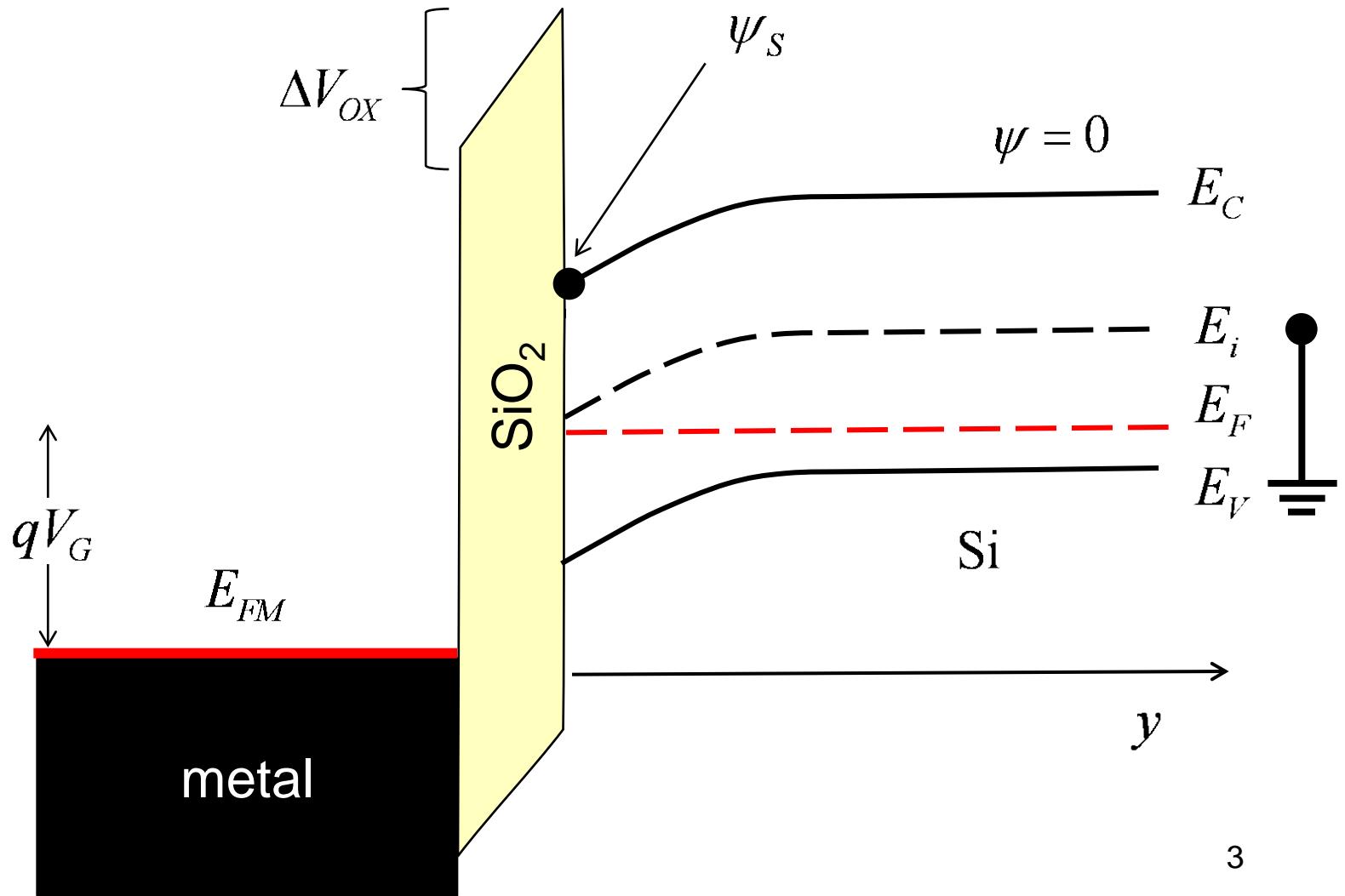
lundstro@purdue.edu
Electrical and Computer Engineering
Purdue University
West Lafayette, Indiana USA

Understanding MOSFETs

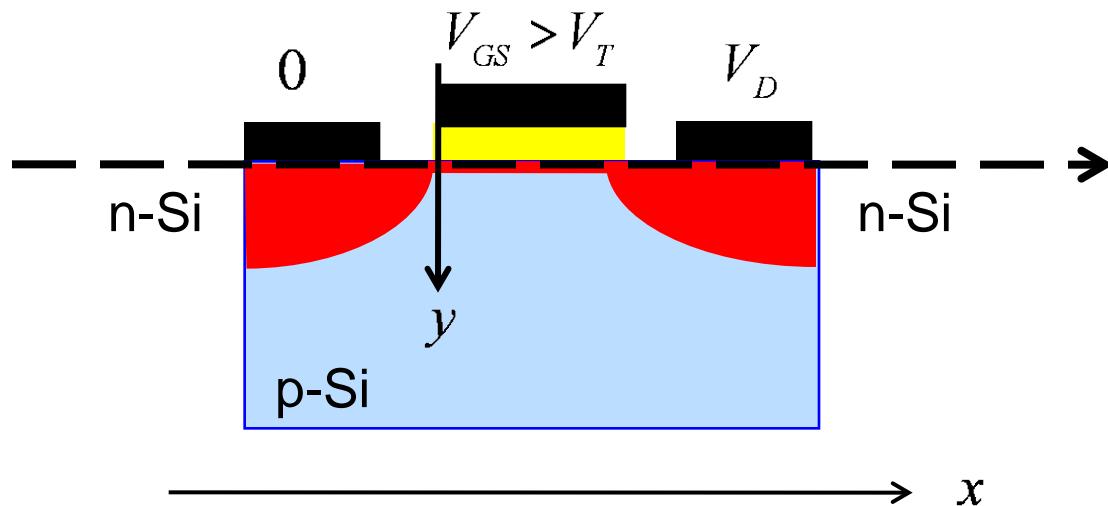


To understand any device, we should first draw an ***Energy Band Diagram***.

Normal to the channel

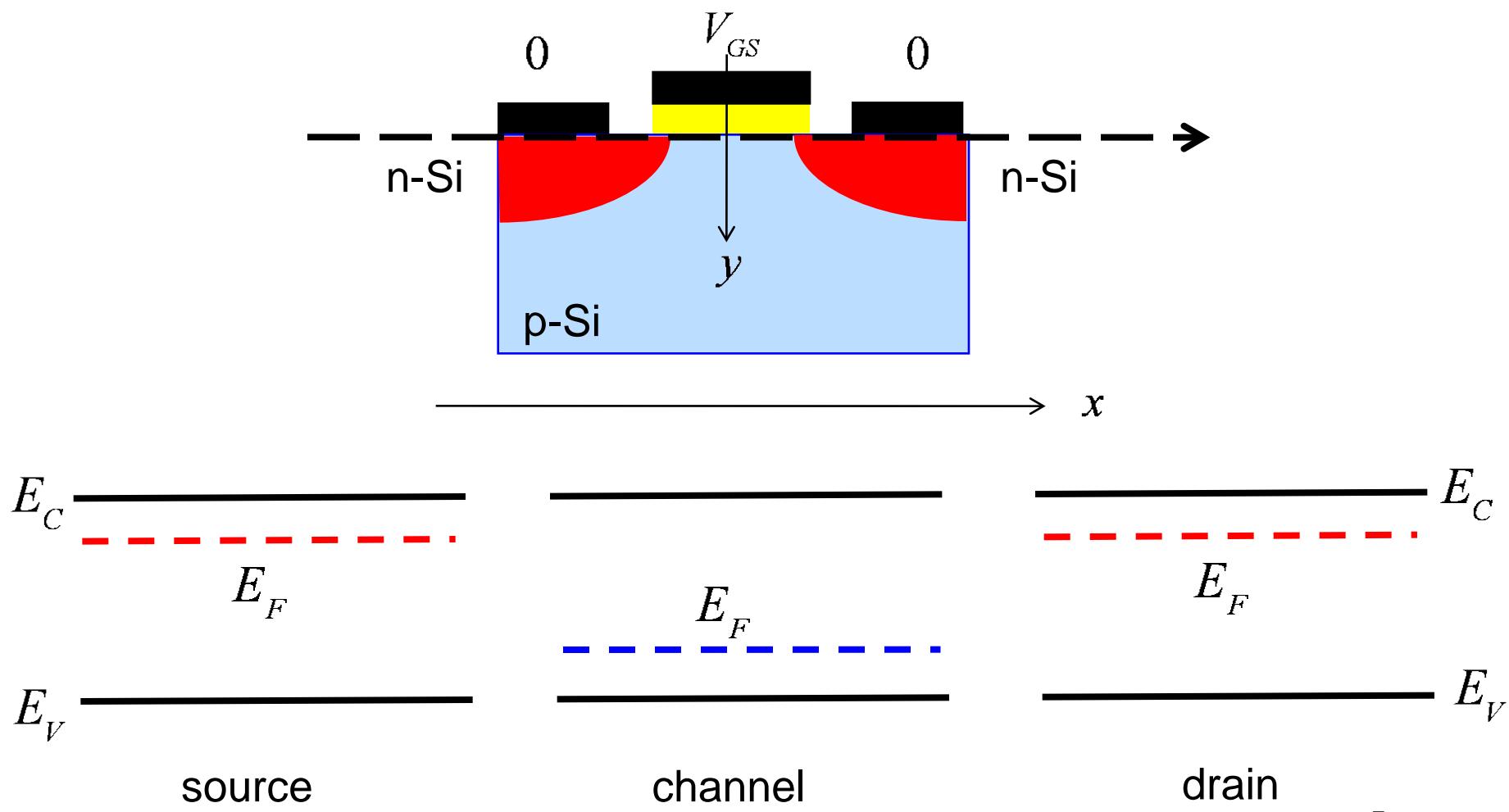


Along the channel

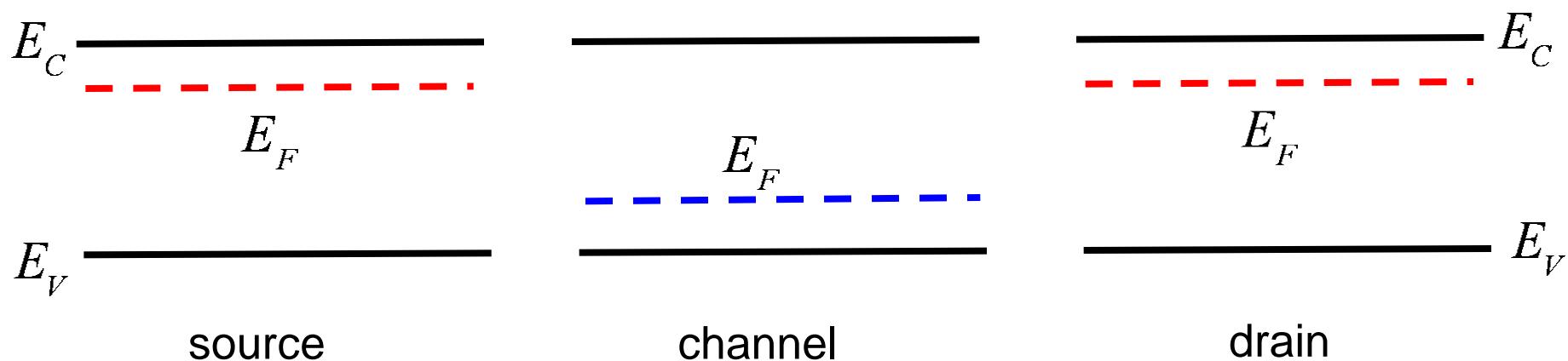


To understand this device, we should first draw an ***Energy Band Diagram***.

Equilibrium E-band diagram



Three separate semiconductors



1) Equilibrium: Fermi level is constant

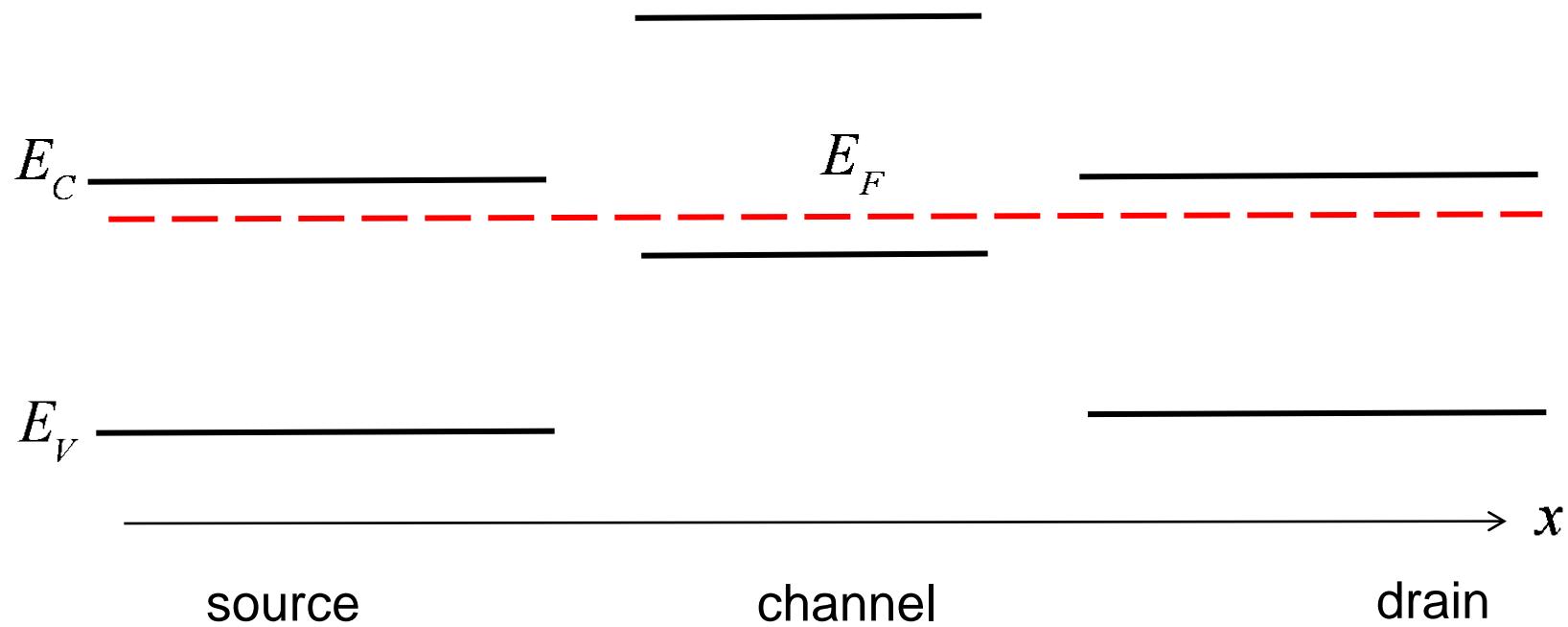
we are bringing E_F of source and drain down to create a common E_F .

2) Changes in electrostatic potential, change the electron's energy.

$$E_C(x) = E_{C0} - q\psi(x)$$

$$E_V(x) = E_{V0} - q\psi(x)$$

Putting the three pieces together

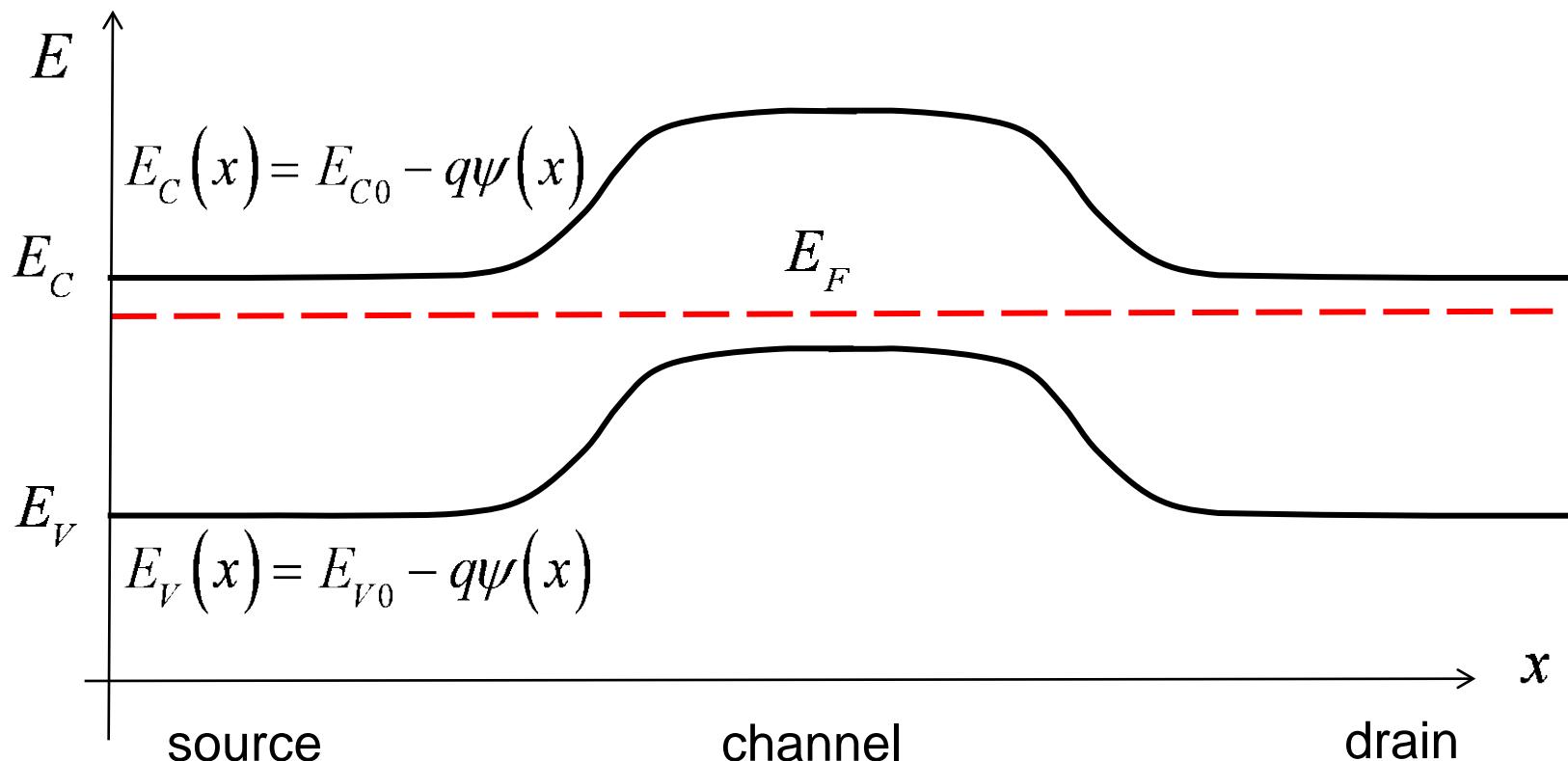


$$E_C(x) = E_{C0} - q\psi(x)$$

$$E_V(x) = E_{V0} - q\psi(x)$$

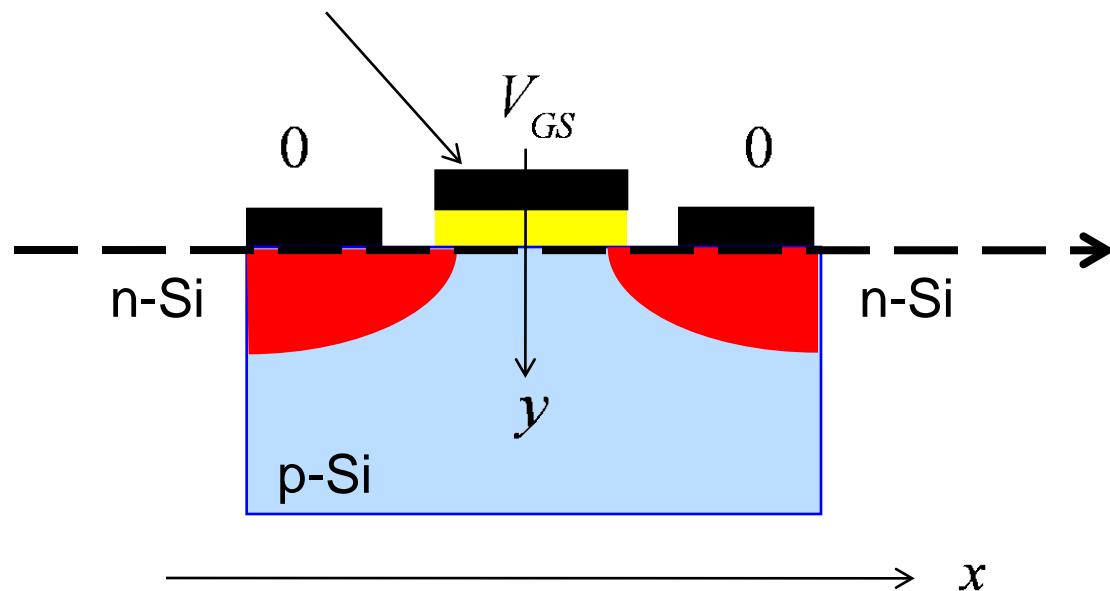
Final result

Now, what effect does a gate voltage have?



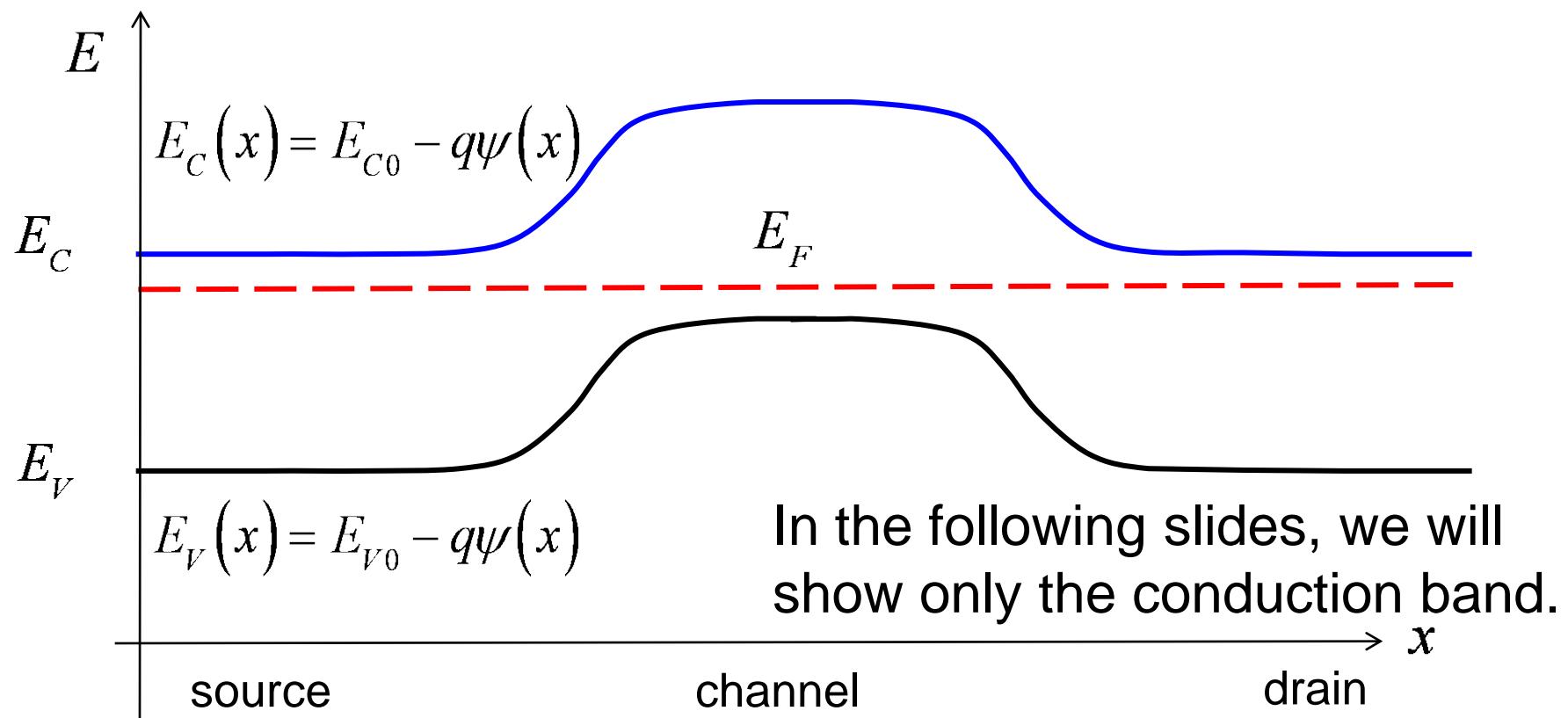
Equilibrium energy band diagram

A positive gate voltage will **increase** the electrostatic potential in the channel and therefore **lower** the electron energy in the channel.



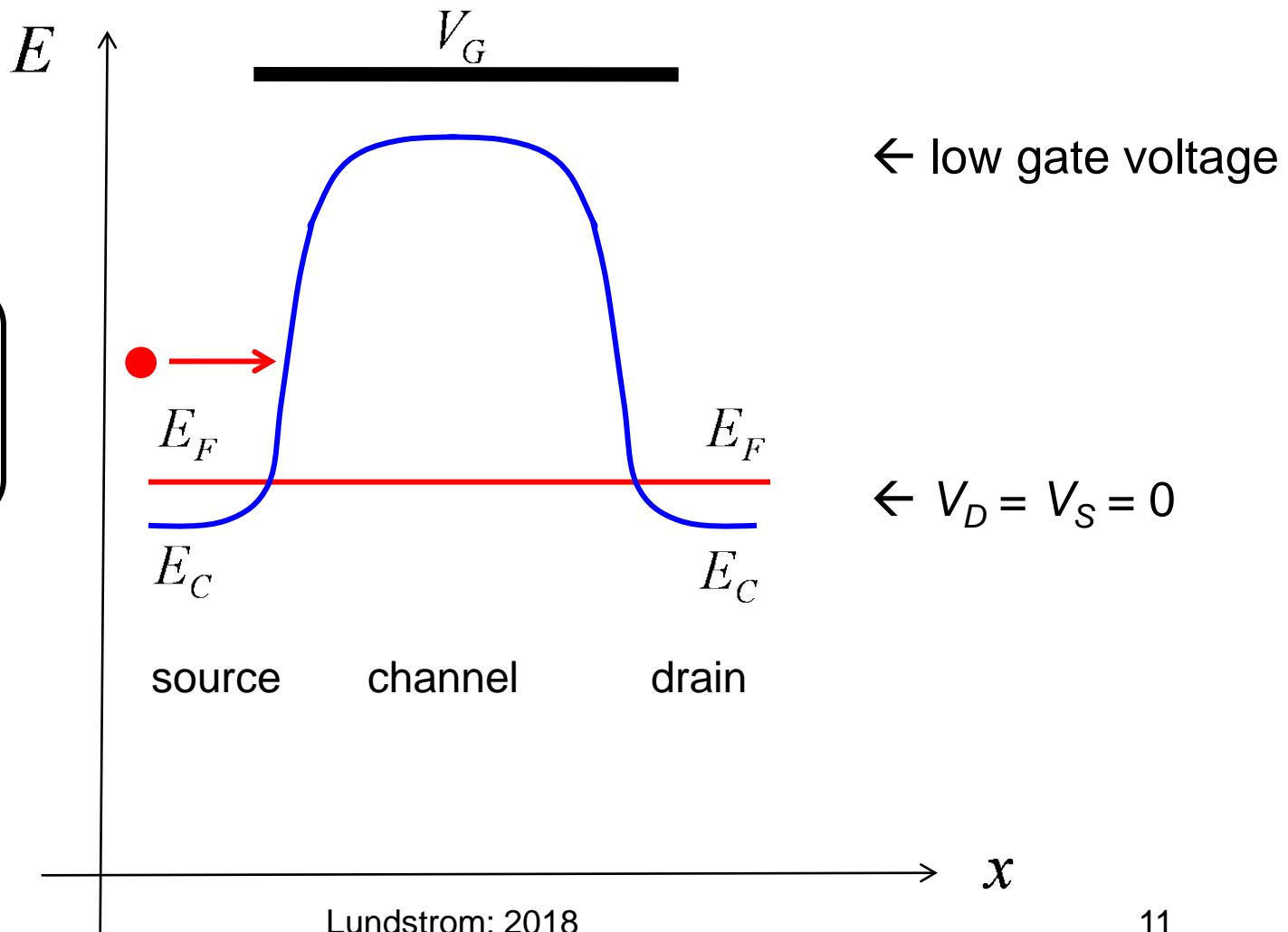
N-channel MOSFET: only electrons

Now, what effect does a gate voltage have?

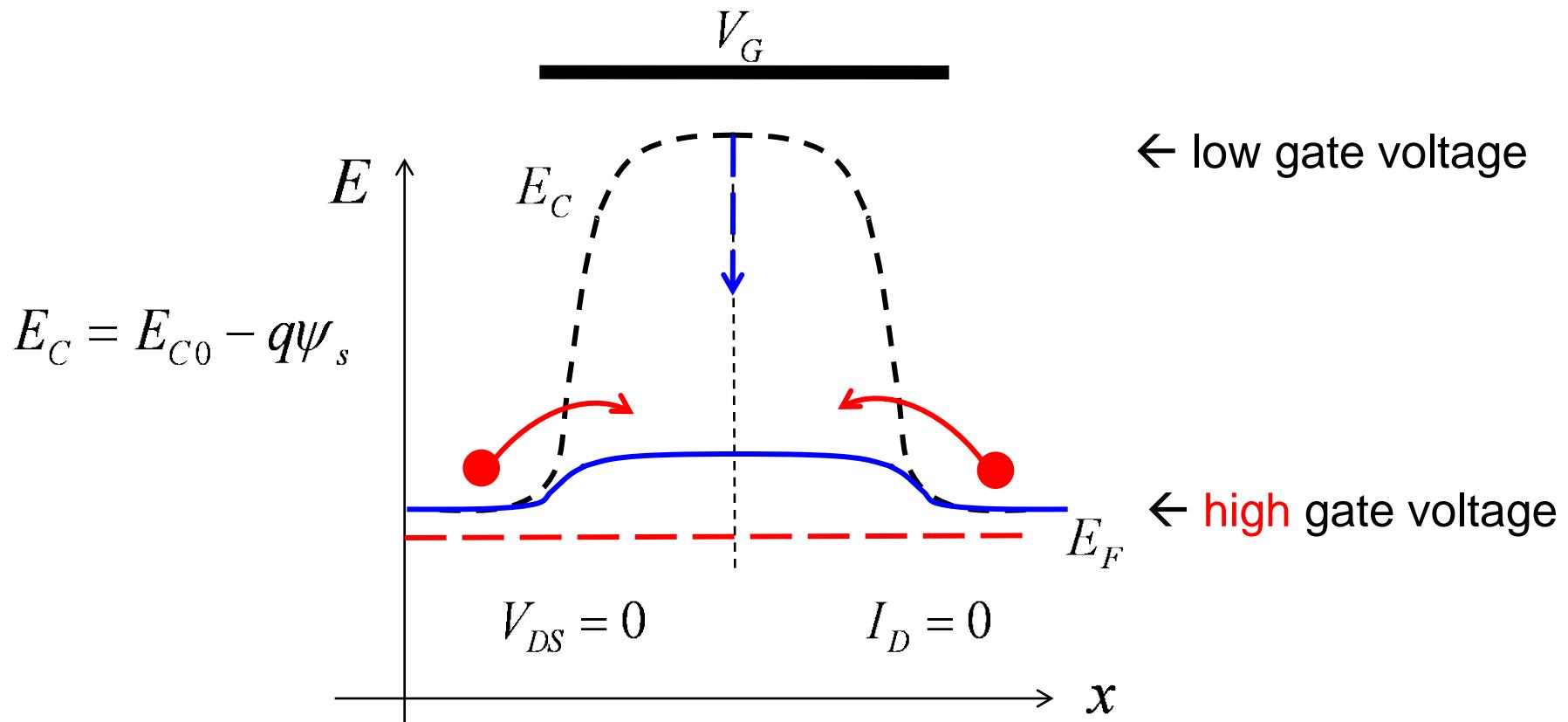


The transistor as a barrier controlled device

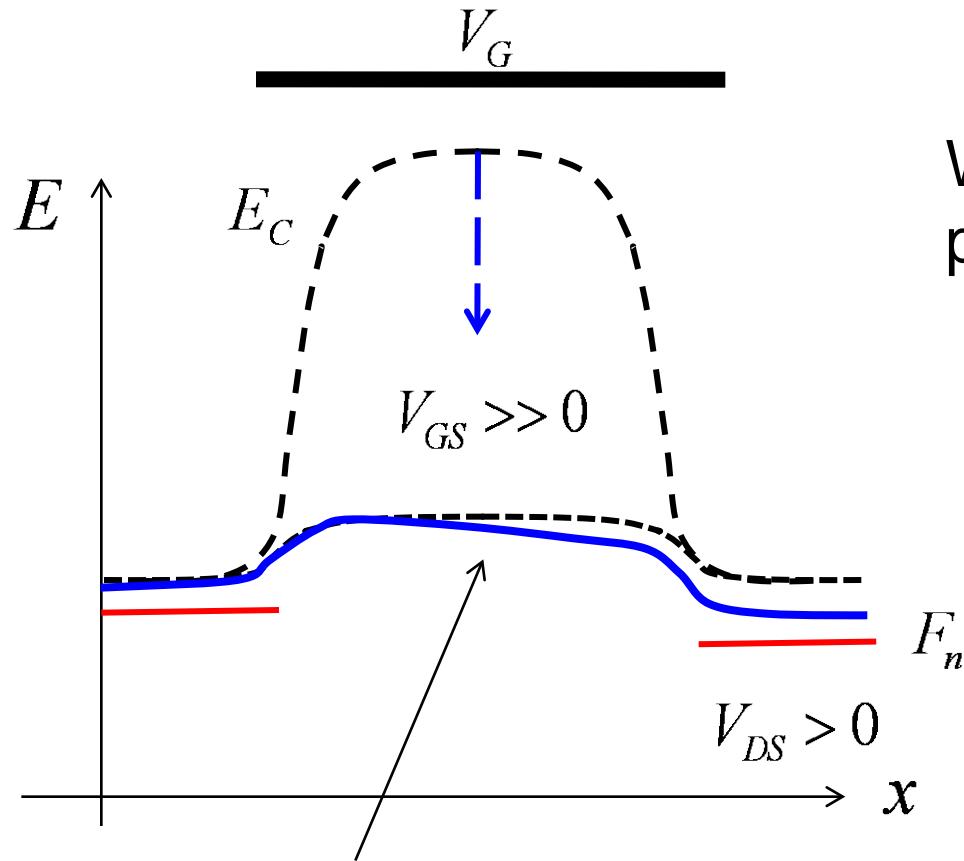
$$E = -qV$$



Examine effect of gate voltage first



Now add a small drain voltage

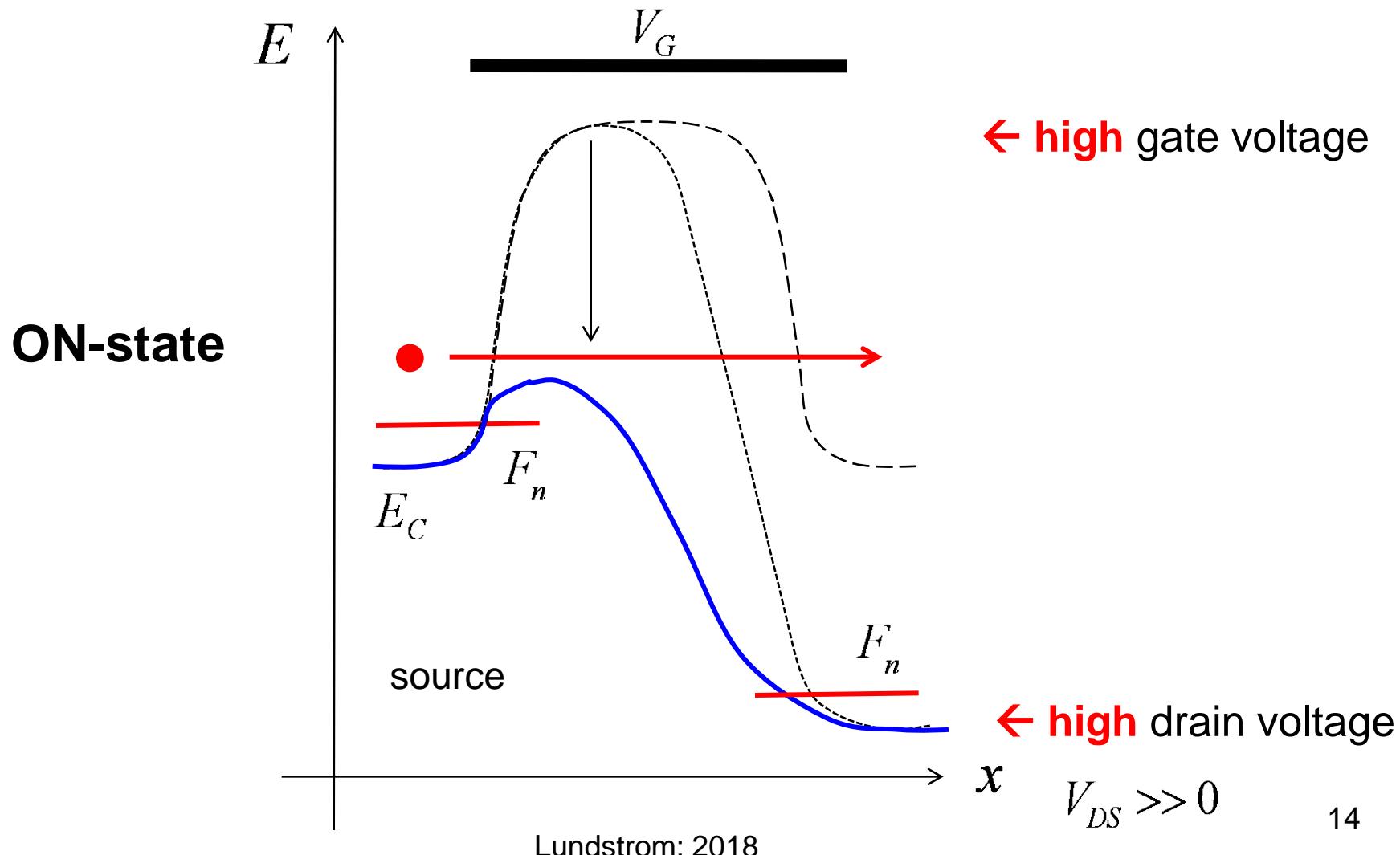


constant electric field
substantial electron density

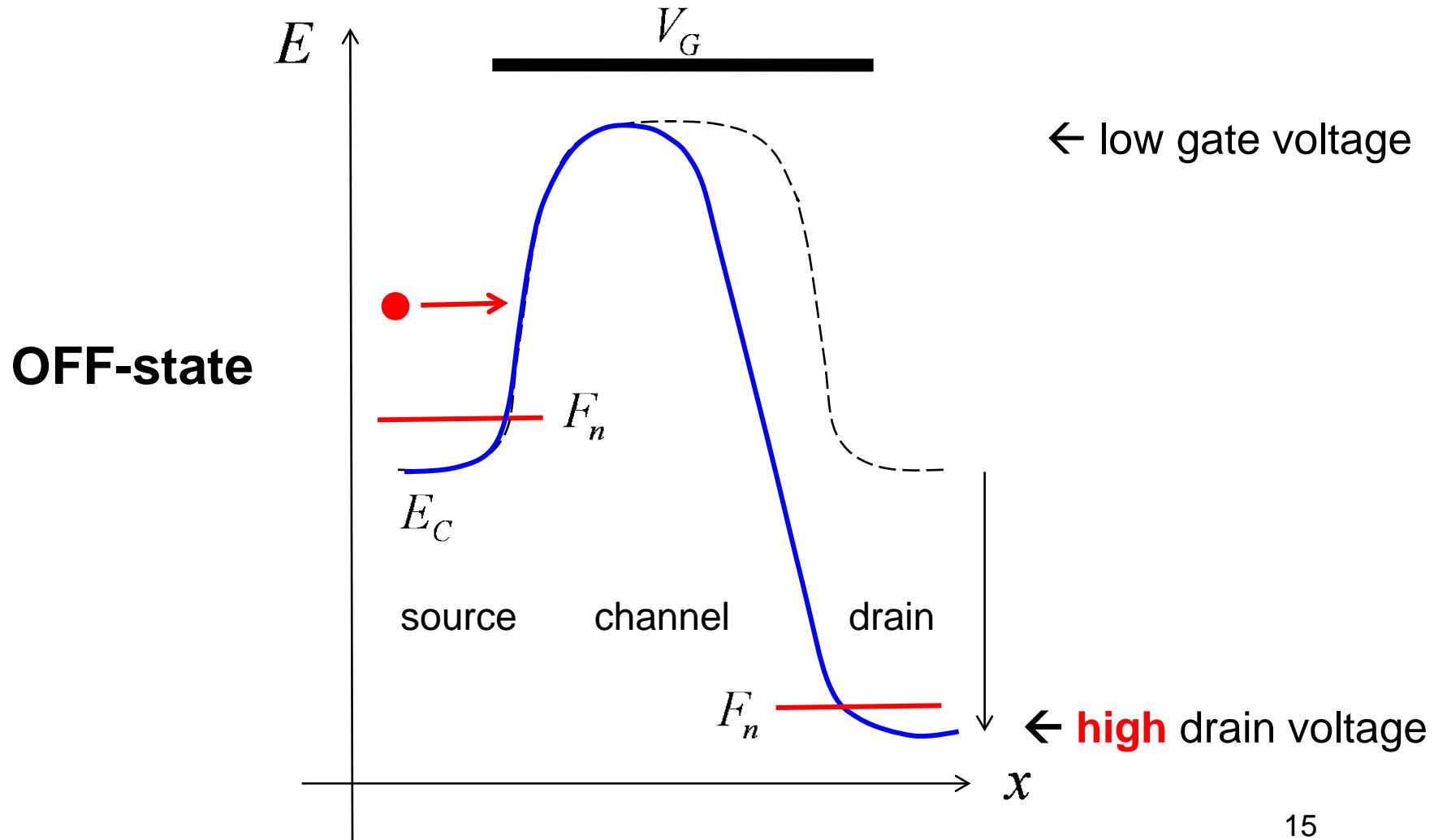
What if we apply a small positive voltage to the drain?

- 1) The Fermi level in the drain is lowered.
- 2) The conduction band is lowered too, but the electron density stays the same.

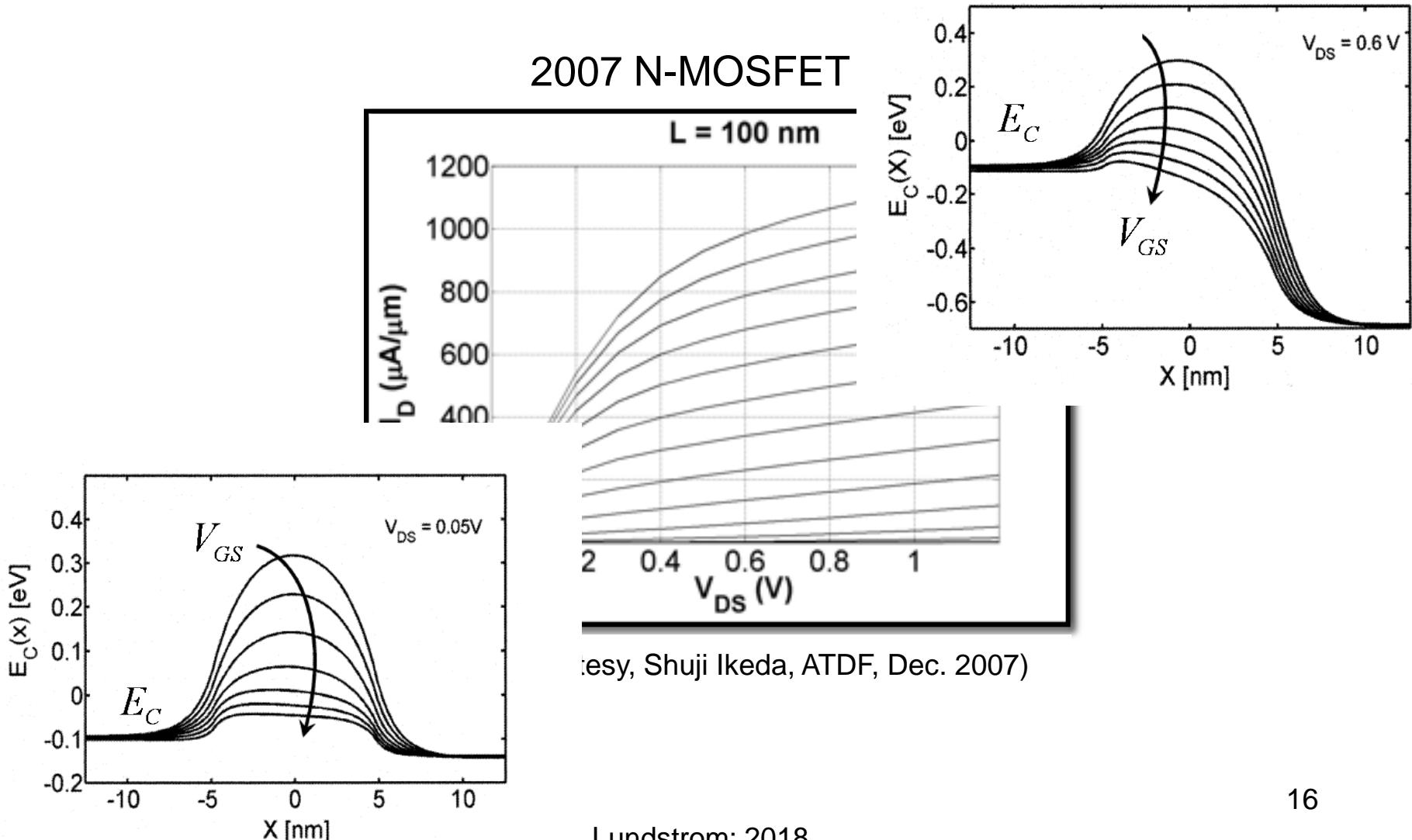
Now increase the drain voltage



Now remove the gate voltage

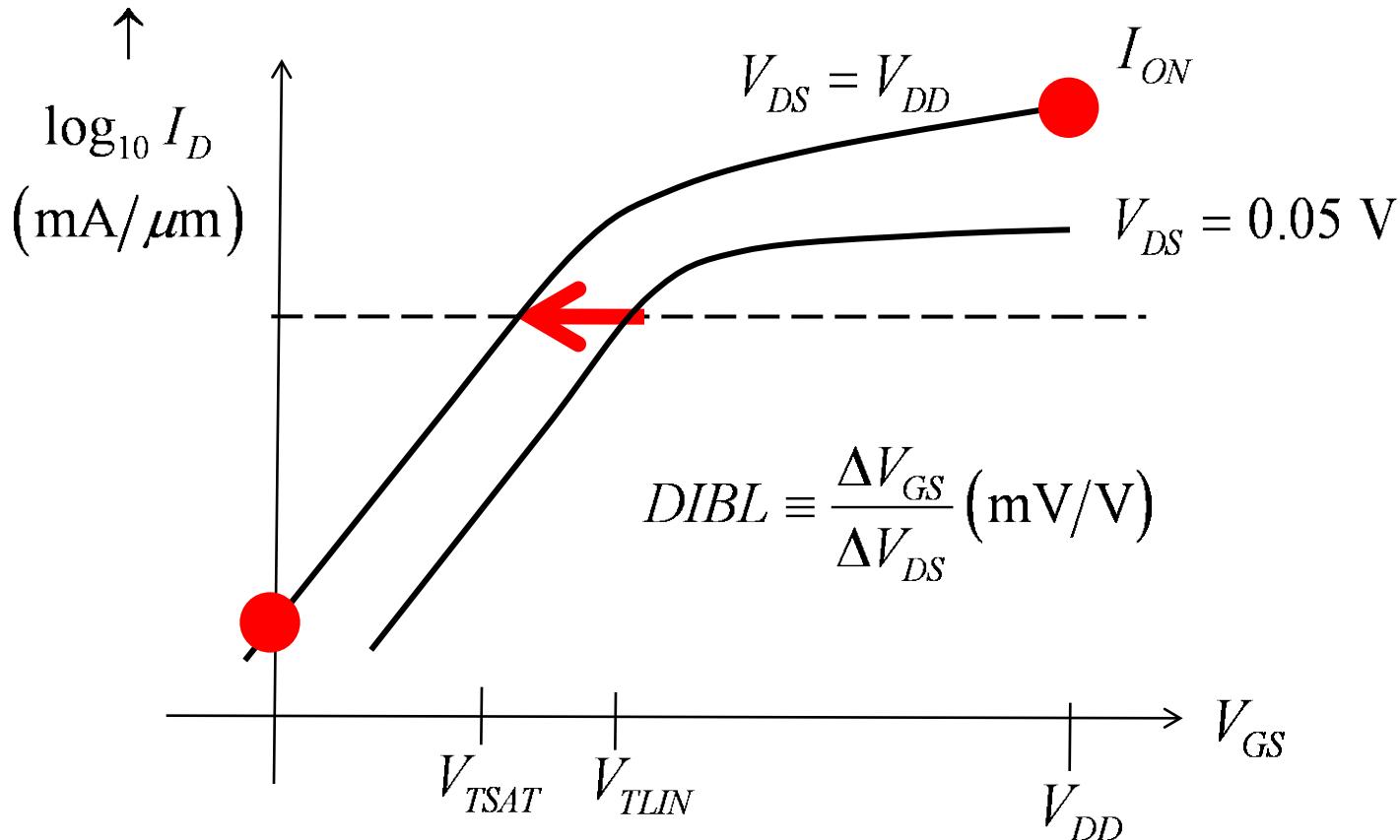


How transistors work

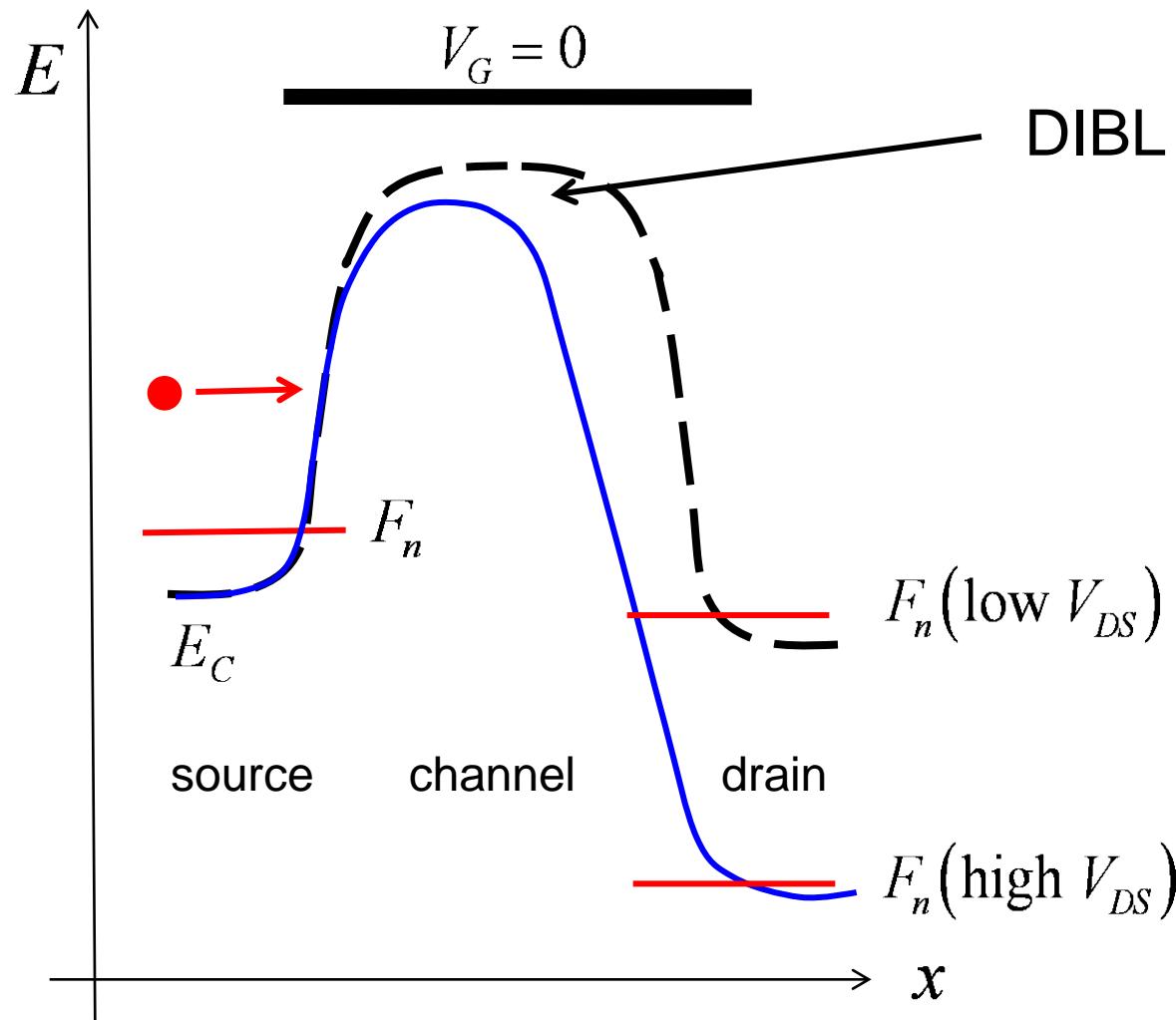


DIBL

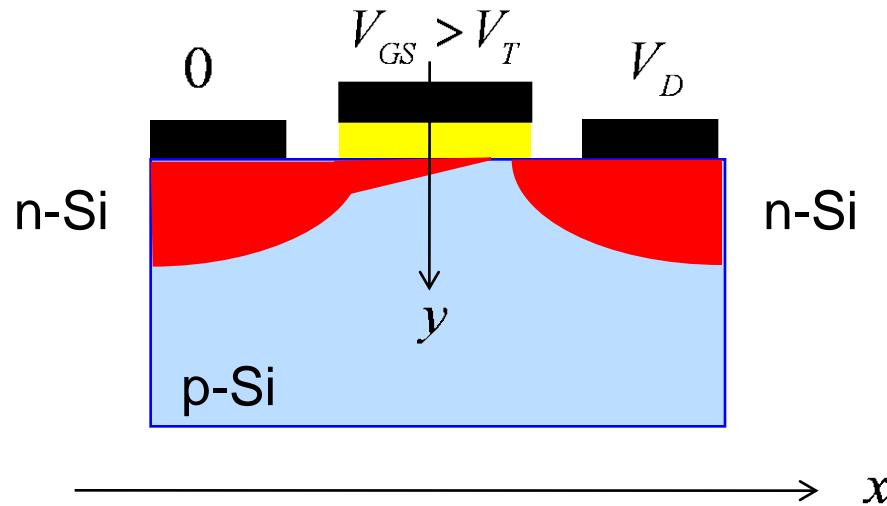
transfer characteristics:



Understanding DIBL with an e-band diagram



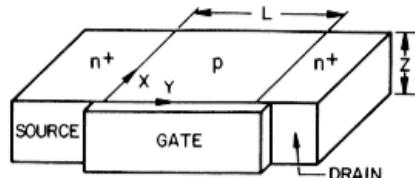
2D energy band diagrams



We have been discussing energy band diagrams from the source to the drain along the top of the Si, but more generally, we should look at the 2D energy band diagram.

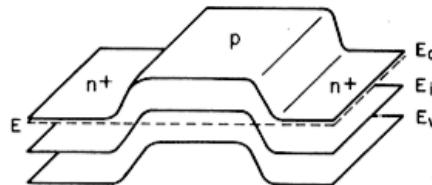
2D energy band diagram on n-MOSFET

(a)



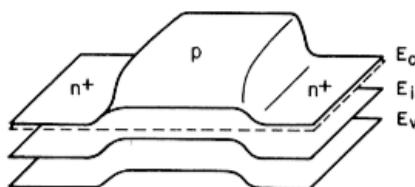
a) device

(b)



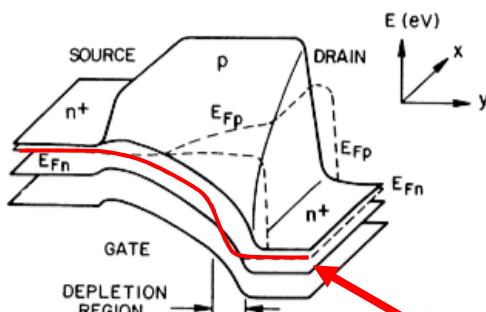
b) equilibrium (flat band)

(c)



c) equilibrium ($V_S > 0$)

(d)



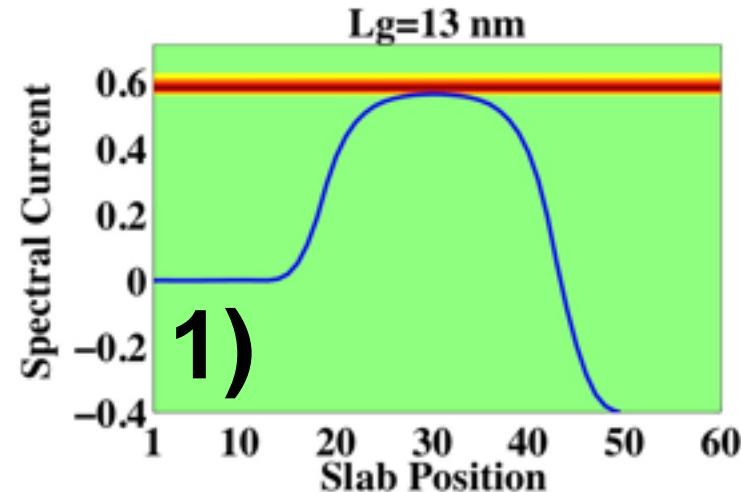
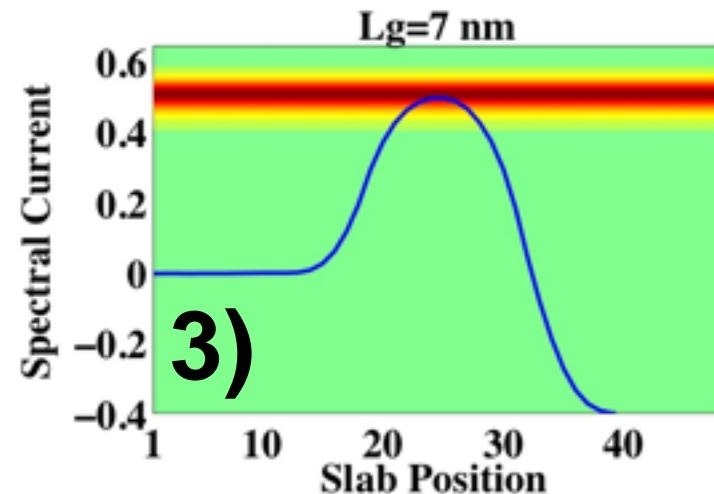
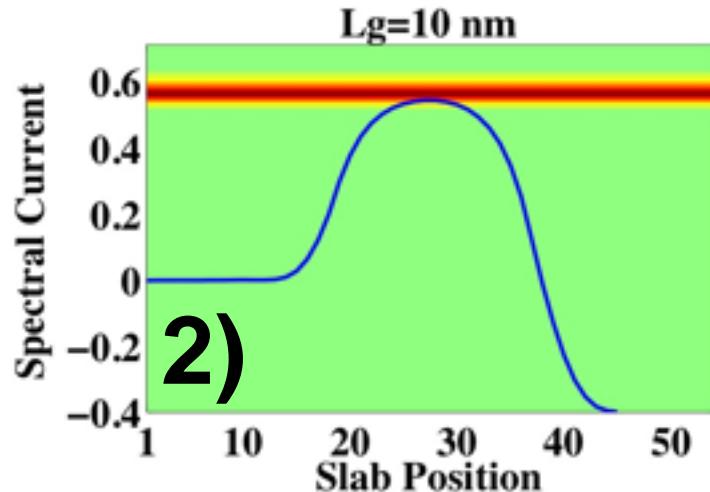
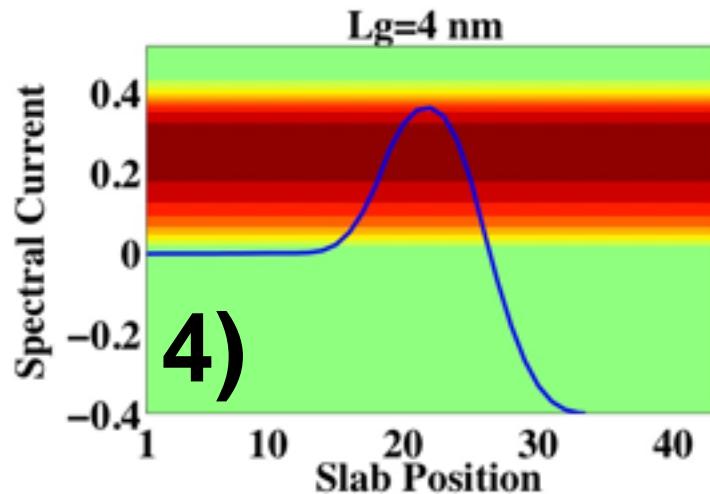
d) non-equilibrium with V_G and $V_D > 0$
applied

F_N

Essential physics of a transistor

A MOSFET (and most transistors) are barrier-controlled devices.

Limits to barrier control: quantum tunneling



Summary

- 1) Energy band diagrams provide a qualitative understanding of how MOSFETs operate.
- 2) MOSFETs are barrier controlled devices – the drain current is controlled by the height of an energy barrier between the source and channel.
- 3) In a well-designed transistor, the height of the energy barrier is strongly controlled by the gate voltage and weakly by the drain voltage (DIBL).

Next topic:

In the next lecture, we will discuss traditional MOSFET theory, which describes the IV characteristics with simple analytical expressions.

Essentials of MOSFETs

Unit 2: Essential Physics of the MOSFET

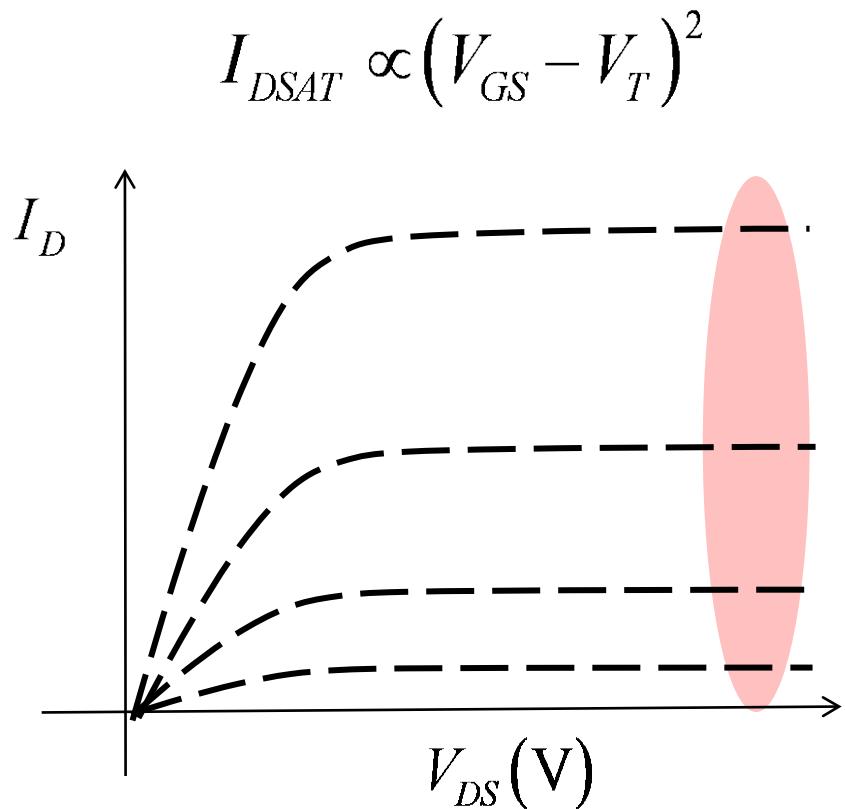
Lecture 2.3: MOSFET IV Theory

Mark Lundstrom

lundstro@purdue.edu
Electrical and Computer Engineering
Purdue University
West Lafayette, Indiana USA

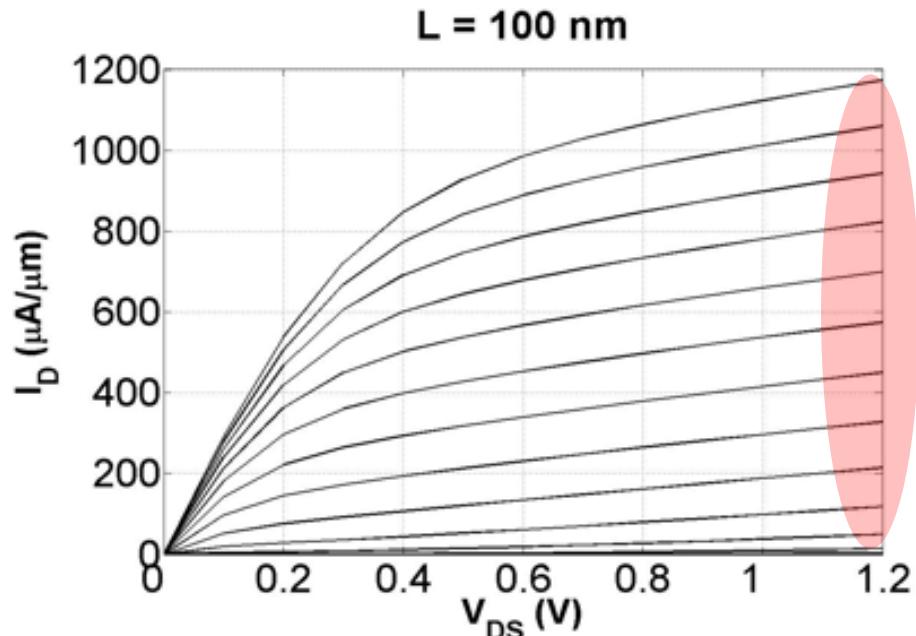
Long vs. short channel MOSFETs

Square Law

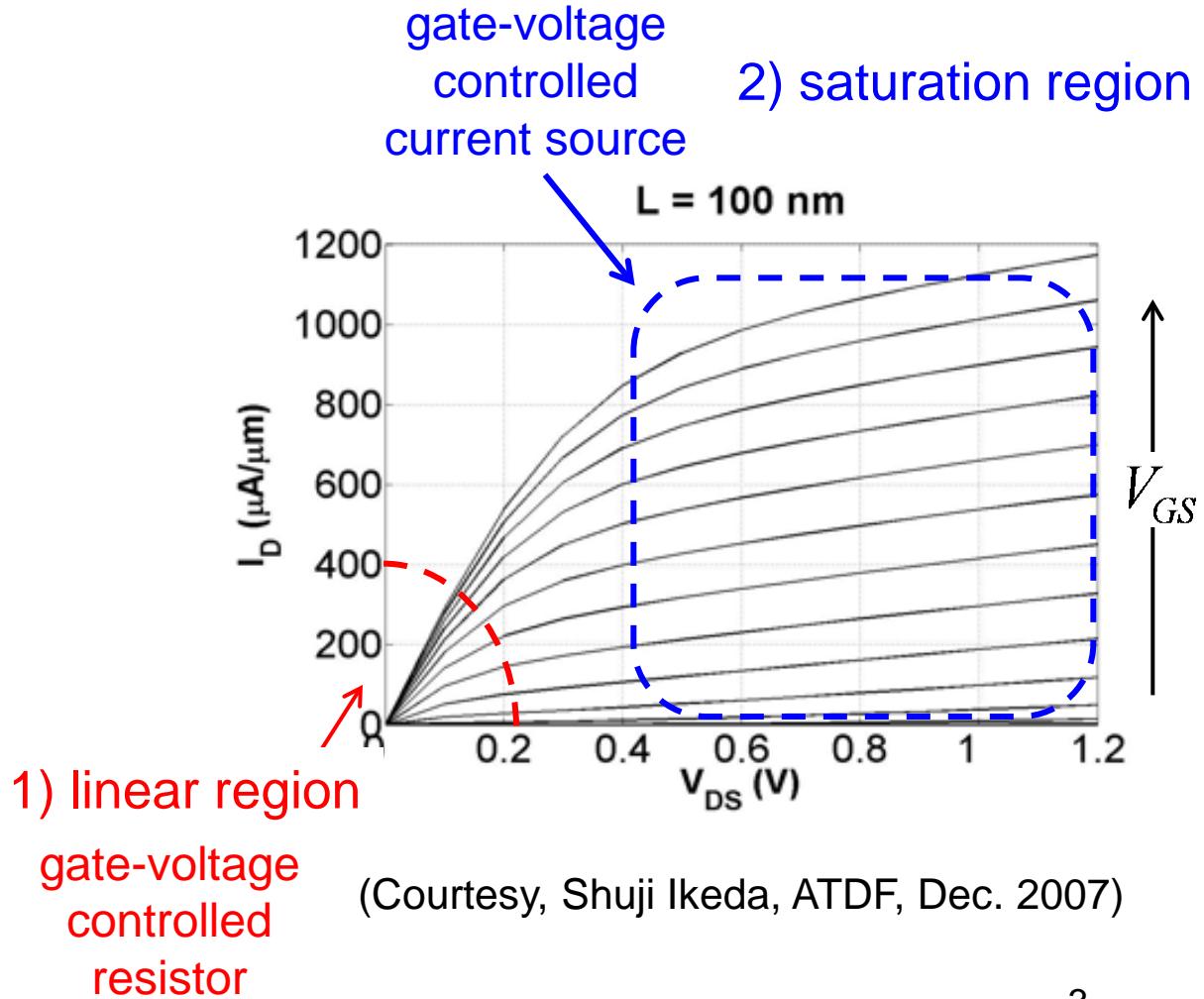
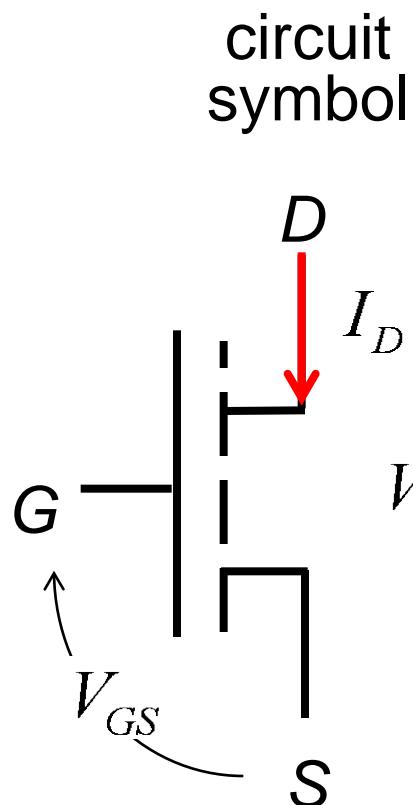


“Velocity saturated”

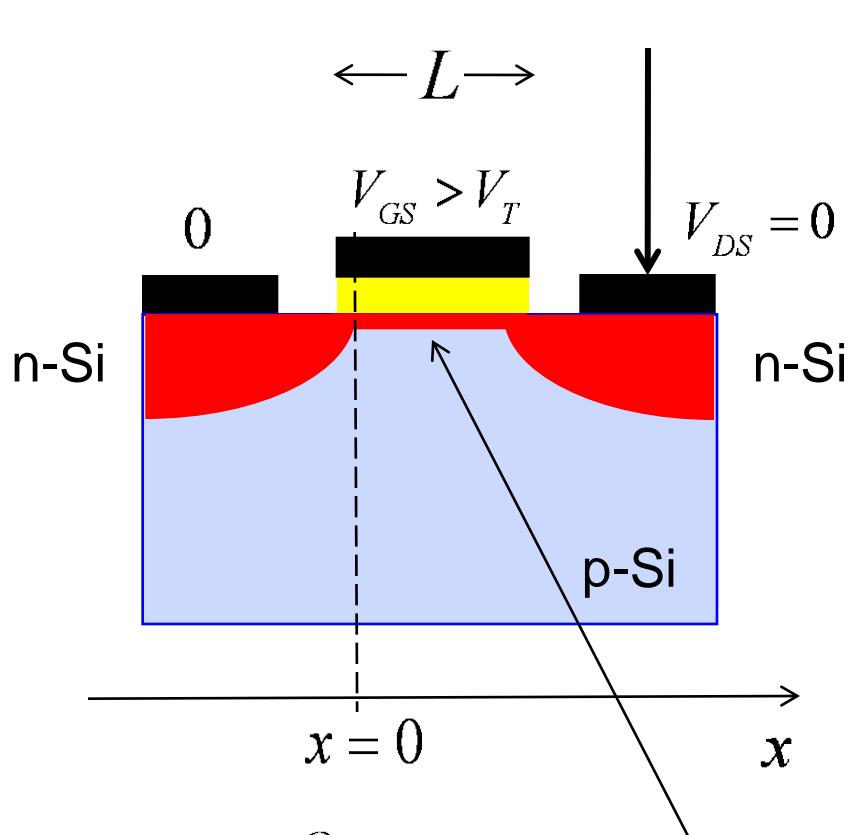
$$I_{DSAT} \propto (V_{GS} - V_T)$$



MOSFET IV characteristic

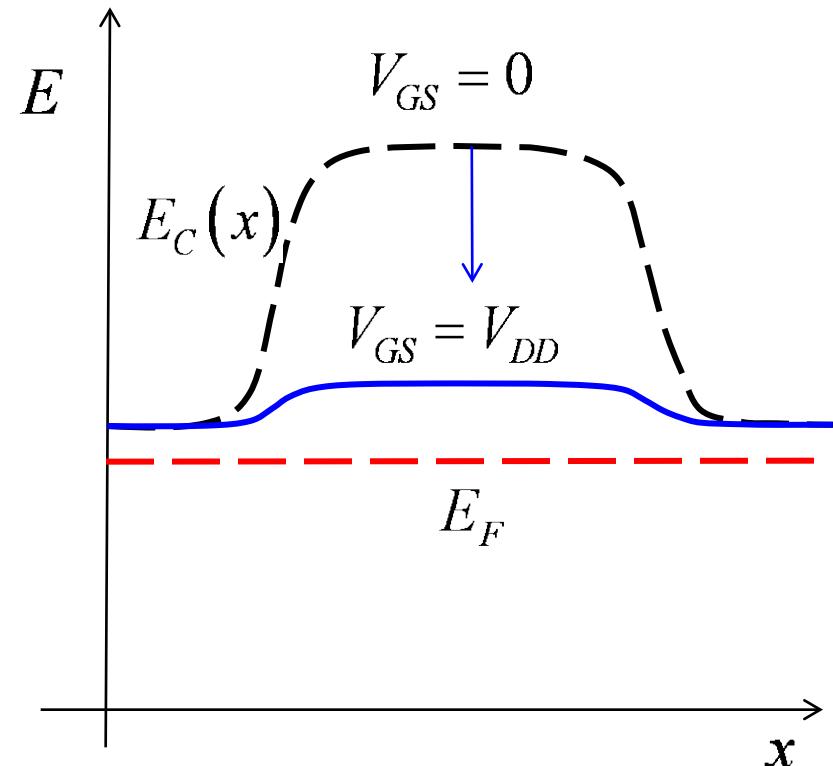


MOSFET e-band (equilibrium)

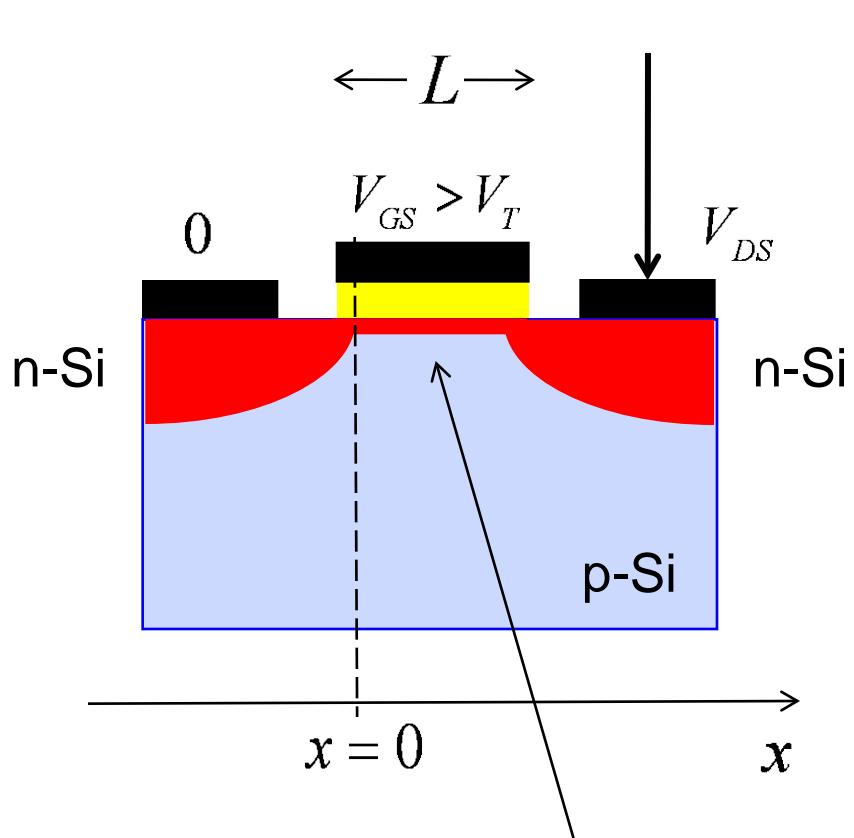


$$C \equiv \frac{Q}{V}$$

$$Q_n \approx -C_{ox}(V_{GS} - V_T) \quad \text{C/cm}^2$$

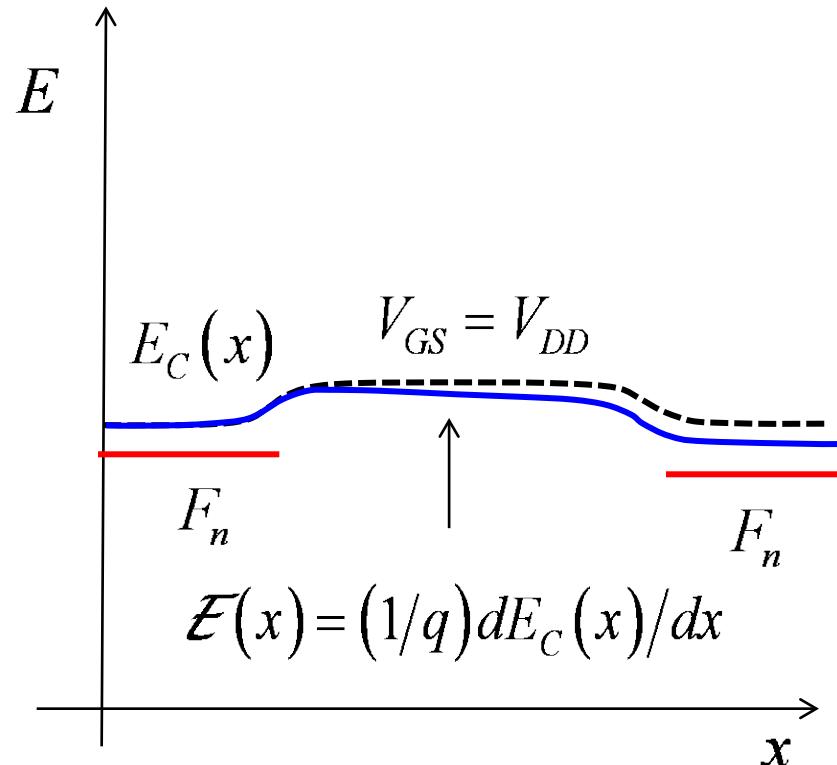


MOSFET e-band (high V_{GS} , low V_{DS})

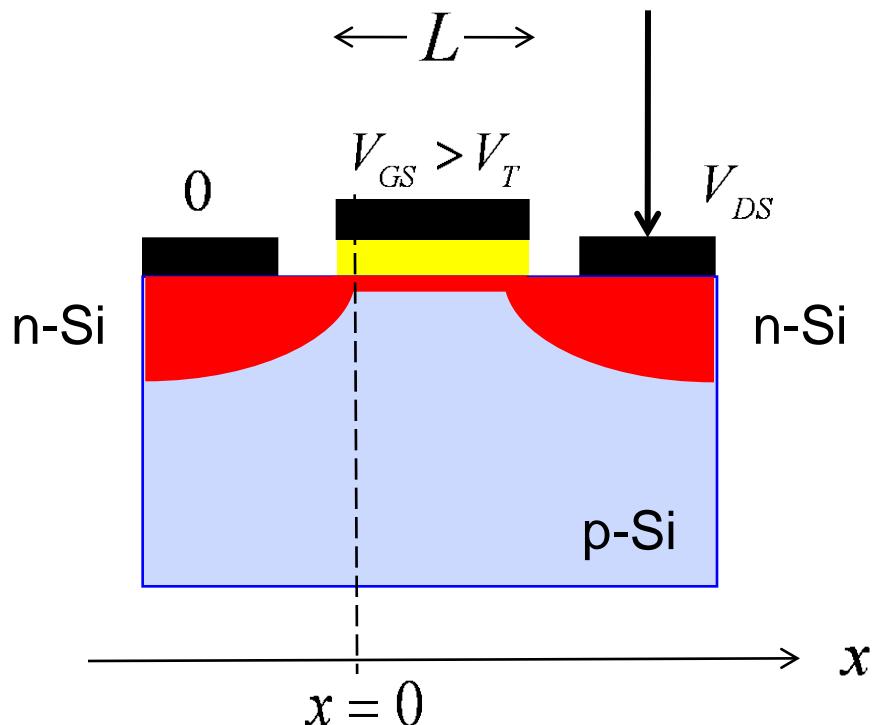


$$Q_n \approx -C_{ox}(V_{GS} - V_T) \quad \text{C/cm}^2$$

$$\mathcal{E}_x = -V_{DS}/L$$



MOSFET IV



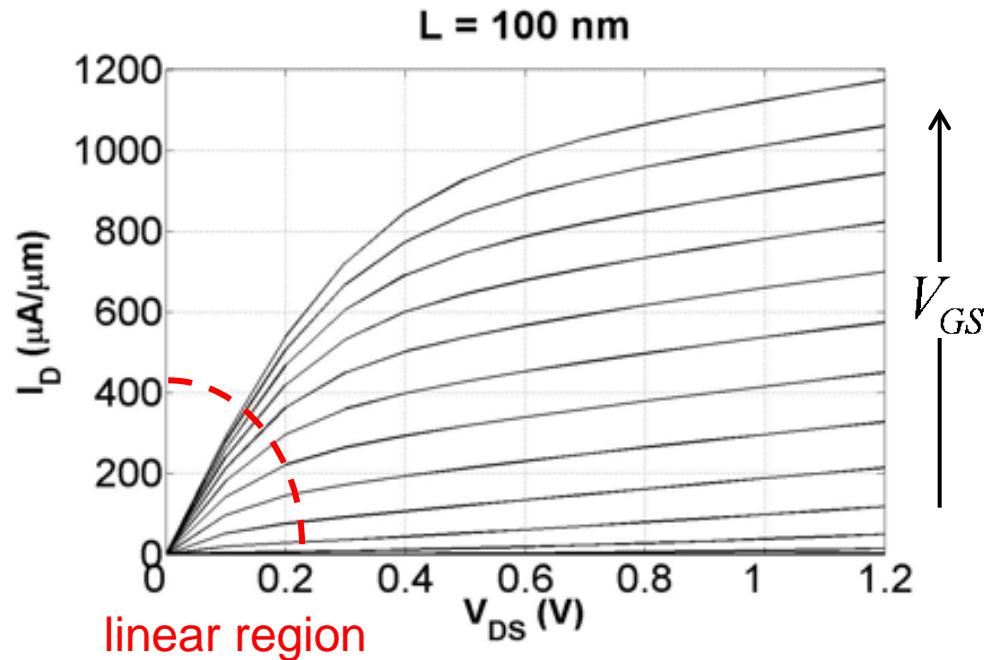
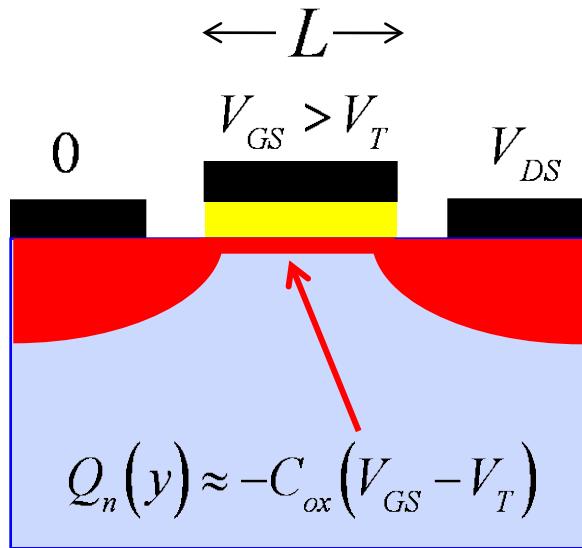
MOS electrostatics

Current is charge per unit time

$$I_D = -W Q_n(x) \langle v_x(x) \rangle$$

$$\left\{ \begin{array}{l} Q_n \approx -C_{ox}(V_{GS} - V_T) \quad \text{C/cm}^2 \\ Q_n \approx 0 \quad (V_{GS} < V_T) \\ C_{ox} = \frac{K_O \epsilon_0}{x_o} \quad \text{F/cm}^2 \end{array} \right.$$

MOSFET IV: low V_{DS}



$$I_D = -W Q_n(x) \langle v_x(x) \rangle$$

$$Q_n = -C_{ox}(V_{GS} - V_T)$$

$$\langle v_x \rangle = -\mu_n \mathcal{E}_x$$

$$\mathcal{E}_x = -V_{DS}/L$$

$$I_D = \frac{W}{L} \mu_n C_{ox} (V_{GS} - V_T) V_{DS}$$



Mobility and effective mobility

The mobility of carriers in the channel of a MOSFET is often called the effective mobility. It is lower than the mobility of electrons in bulk silicon.

Example:

Consider an ultra-thin body (UTB) Si on Insulator (SOI) MOSFET with an undoped channel.

For undoped bulk Si: $\mu_n = 1360 \text{ cm}^2/\text{V-s}$

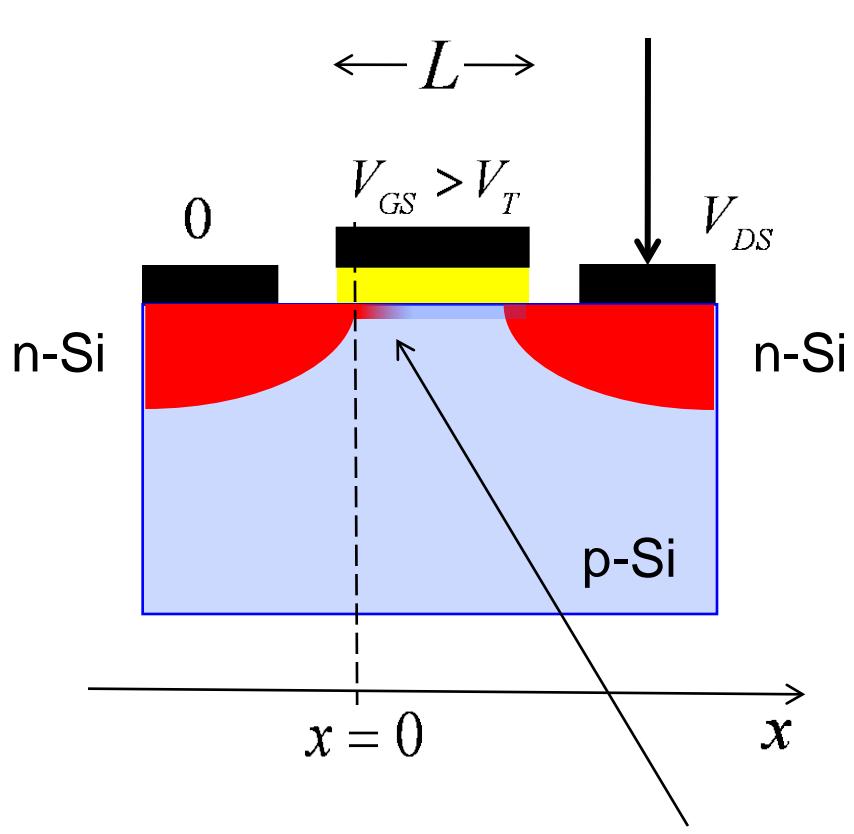
For the UTB MOSFET: $\mu_n = 200 - 300 \text{ cm}^2/\text{V-s}$

The lower mobility is due to **surface roughness scattering** at the Si/oxide interface.

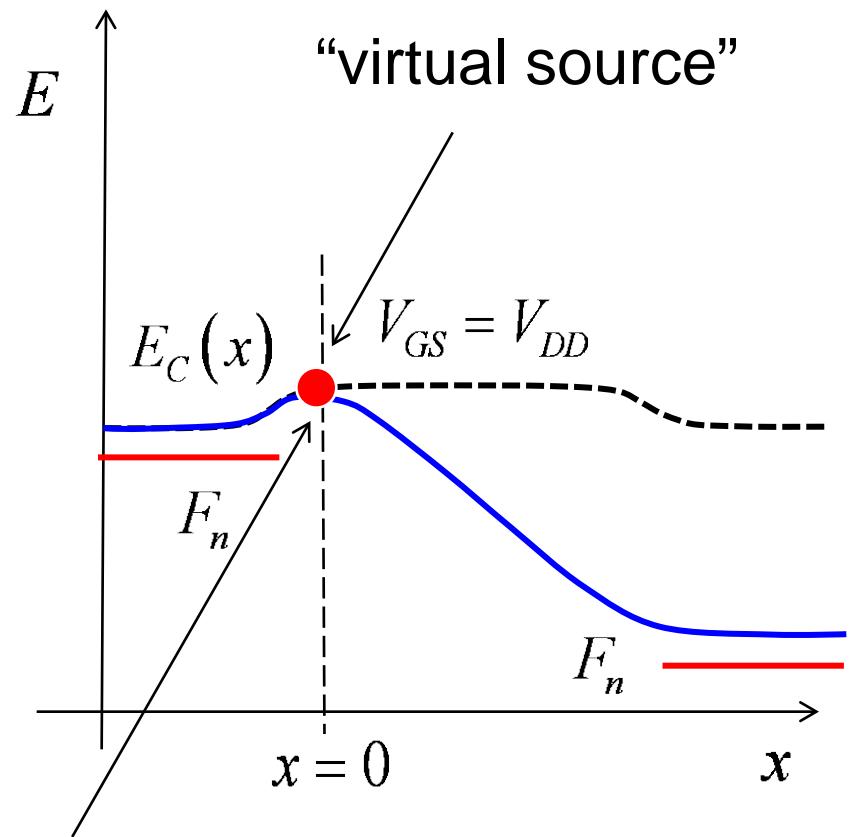
Outline

- 1) Linear region
- 2) **Saturation region**
 - classical pinch off model
 - velocity saturation model

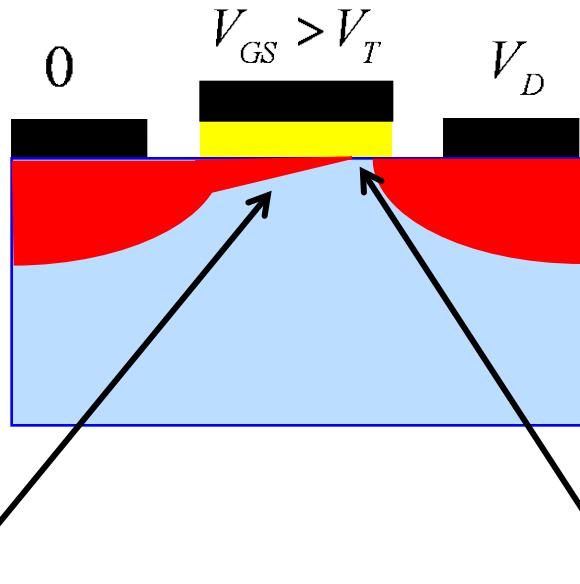
MOSFET e-band (high V_{GS} , low V_{DS})



$$Q_n(x=0) \approx -C_{ox}(V_{GS} - V_T) \quad \text{C/cm}^2$$



MOSFET IV: “pinch-off” at high V_{DS}



$$Q_n(x) = -C_{ox}(V_{GS} - V_T - V(x))$$

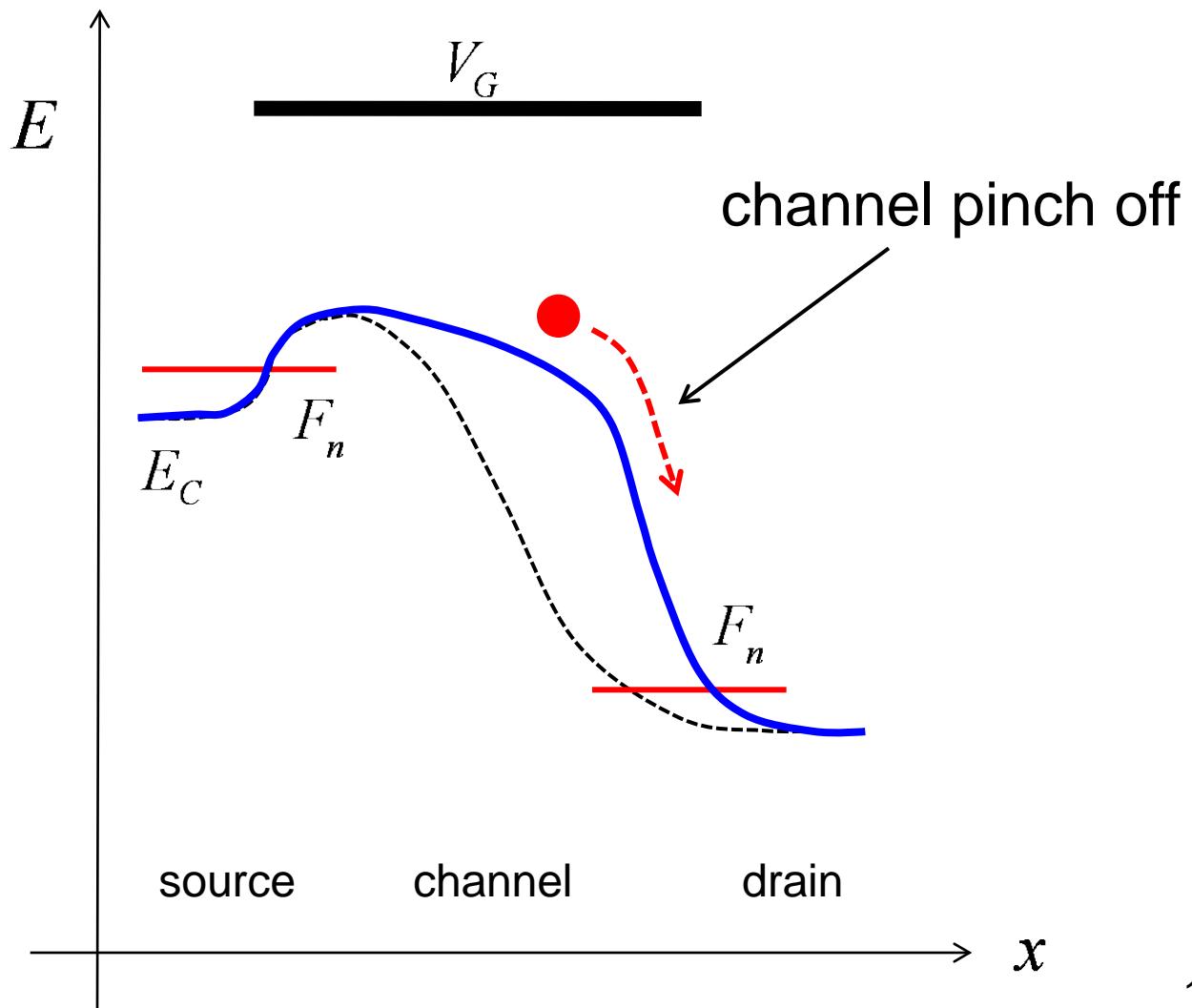
$$V(x_{pinch}) = (V_{GS} - V_T)$$

$$Q_n(x_{pinch}) \approx 0$$

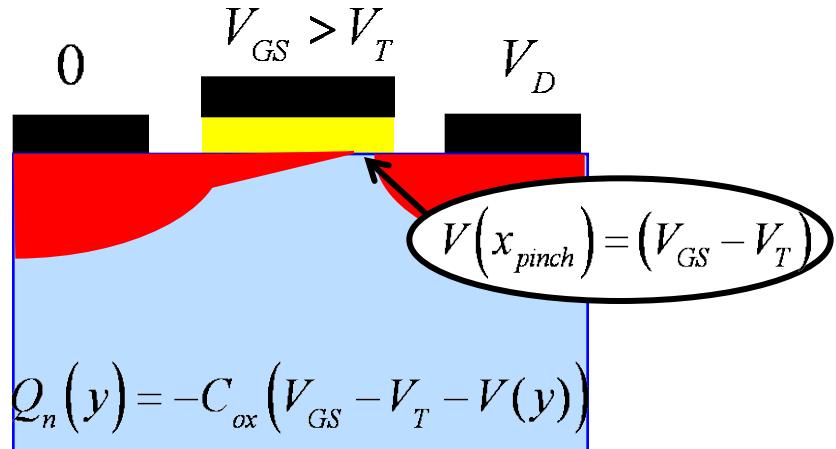
Note: thickness of channel illustrates the areal density of electrons – not the actual thickness.

Electric field is very large in the pinch-off region.

“Pinch off” on an energy band diagram

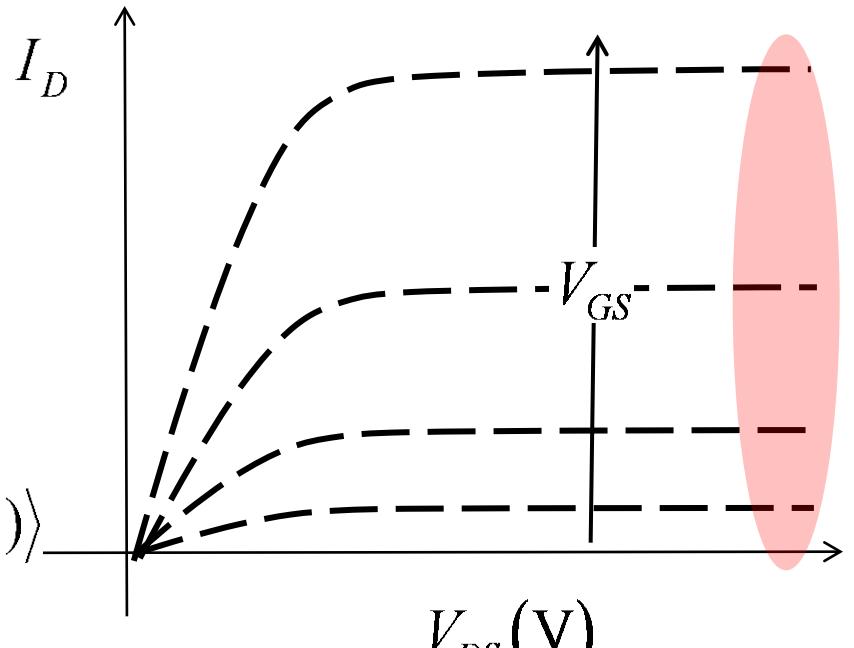


MOSFET IV: high V_{DS}



$$I_D = -W Q_n(x) \langle v_x(x) \rangle = W Q_n(0) \langle v_x(0) \rangle.$$

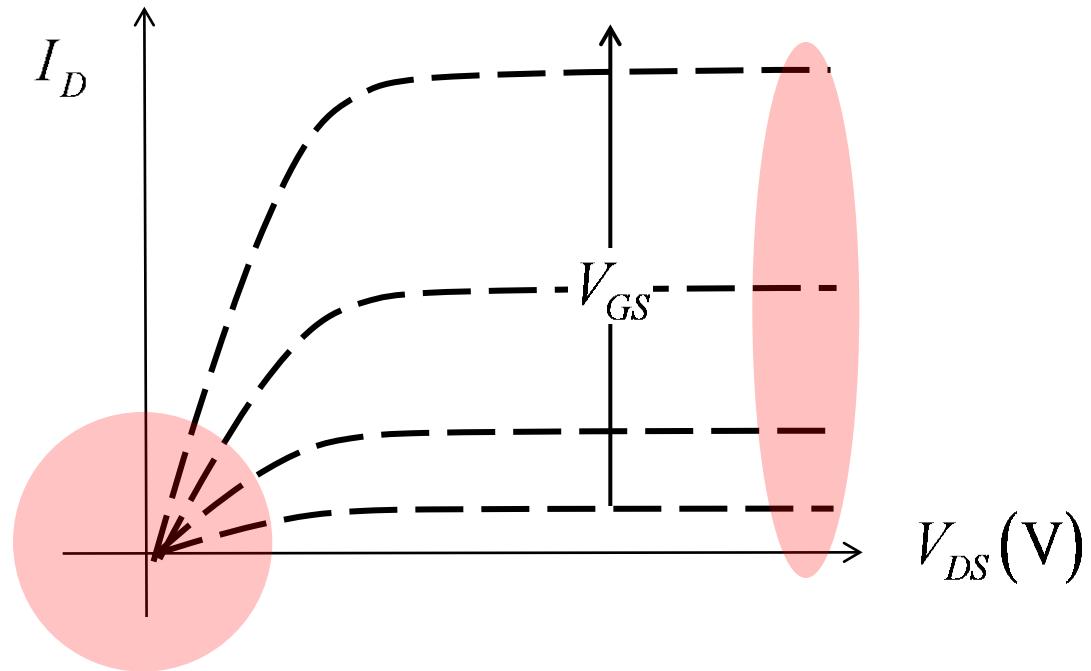
$$Q_n(0) = -C_{ox}(V_{GS} - V_T)$$



$$I_D = \frac{W}{2L} \mu_n C_{ox} (V_{GS} - V_T)^2$$

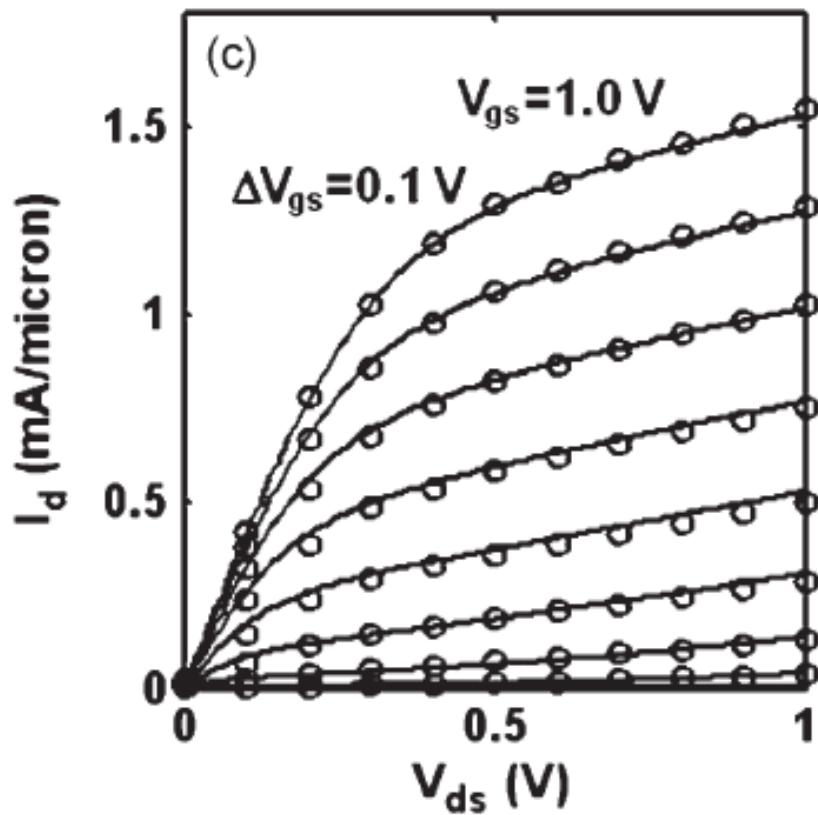
$$\mathcal{E}_x(0) \approx -V(x_{pinch})/L = -(V_{GS} - V_T)/L$$

The square law MOSFET



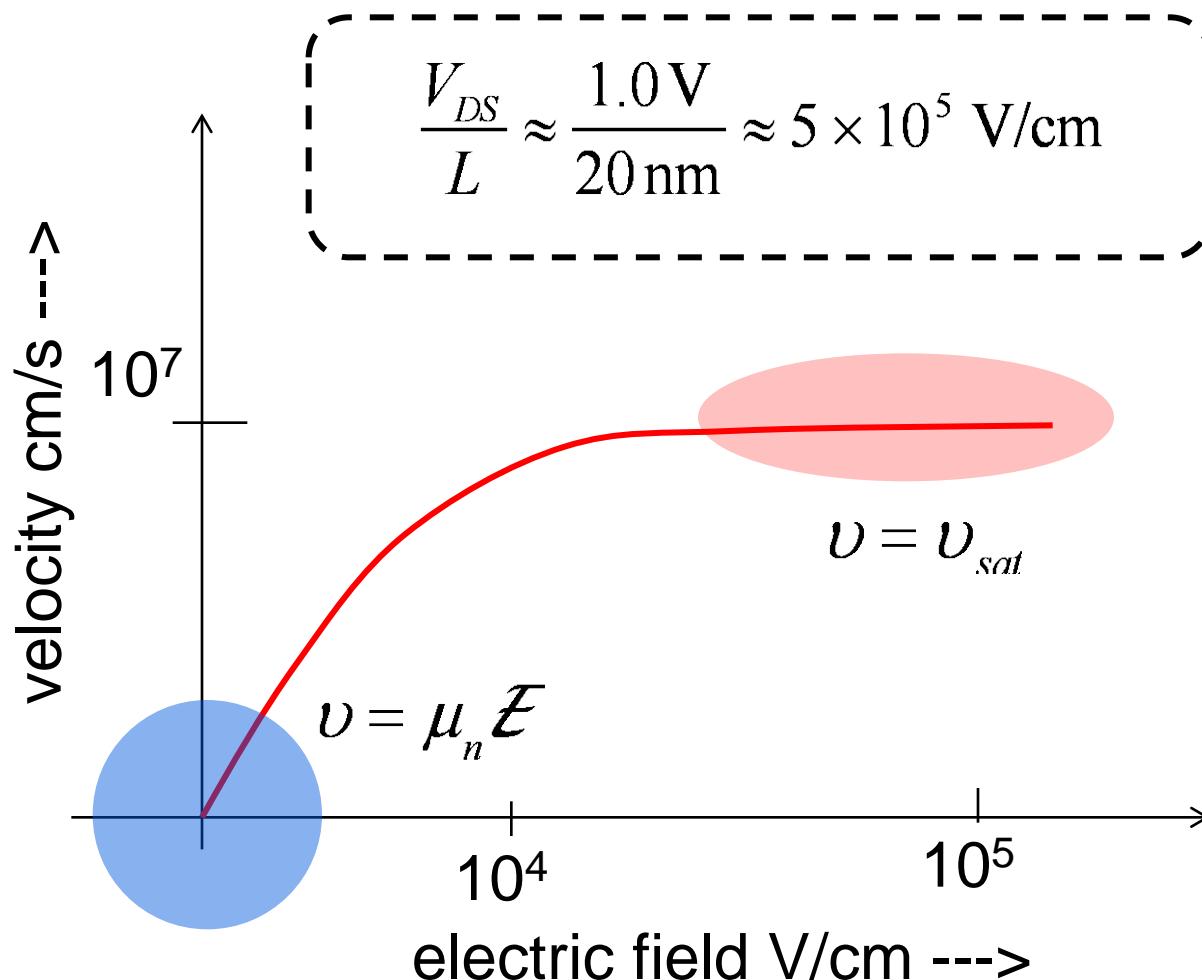
$$I_D = \frac{W}{L} \mu_n C_{ox} (V_{GS} - V_T) V_{DS} \quad I_D = \frac{W}{2L} \mu_n C_{ox} (V_{GS} - V_T)^2$$

Modern, short channel MOSFETs

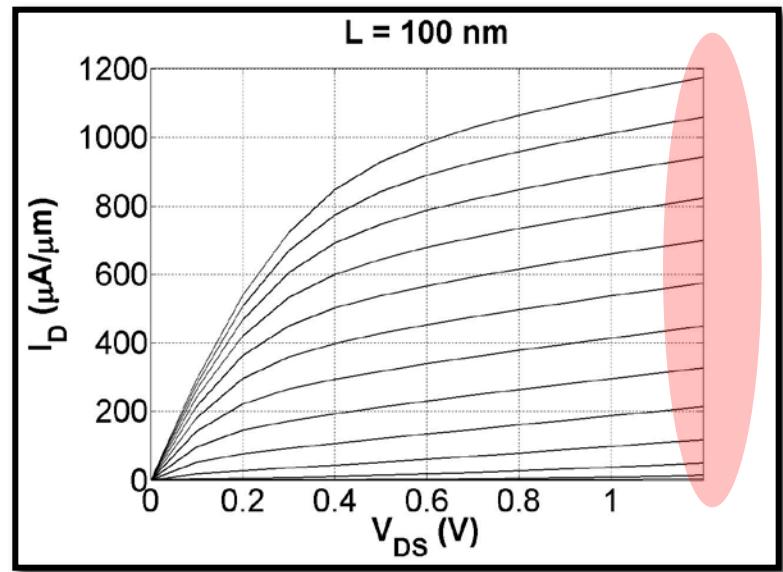
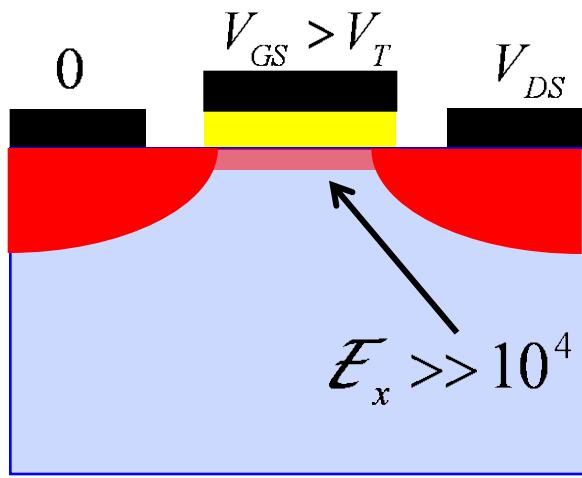


$$I_{DSAT} \propto (V_{GS} - V_T)$$

High V_{DS} : Velocity saturation



MOSFET IV: velocity saturation



$$I_D = -W Q_n(x) \langle v_x(x) \rangle$$

(Courtesy, Shuji Ikeda, ATDF, Dec. 2007)

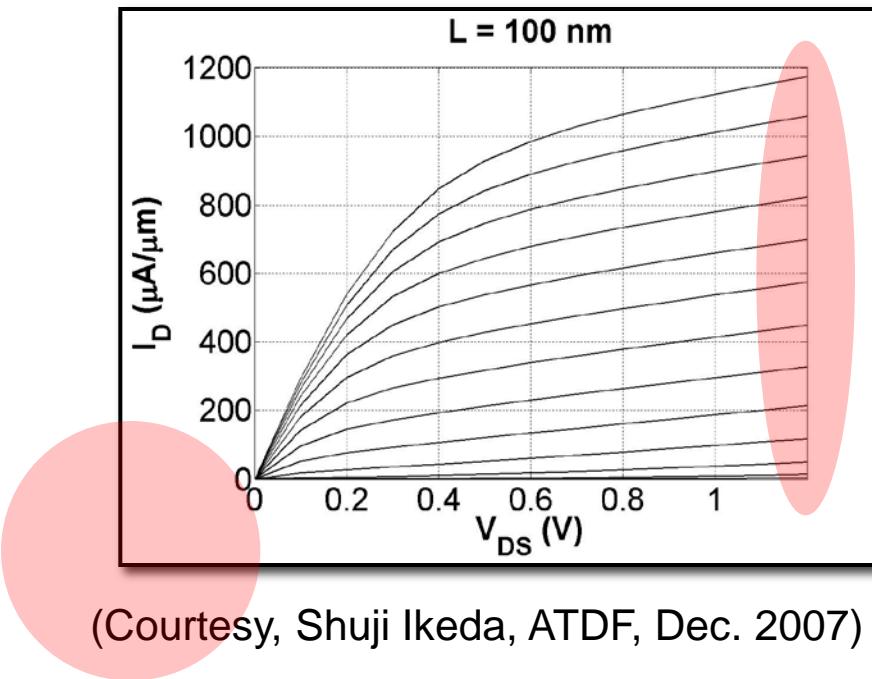
$$Q_n = -C_{ox} (V_{GS} - V_T)$$

$$\langle v_x \rangle = v_{sat}$$

$$I_D = W C_{ox} v_{sat} (V_{GS} - V_T)$$



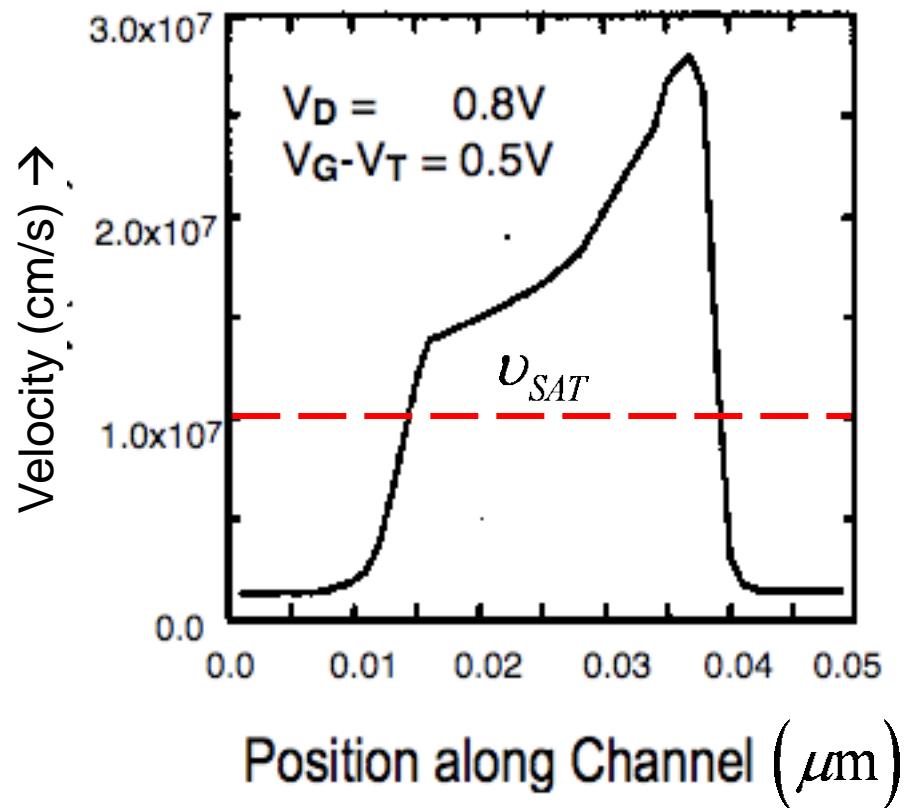
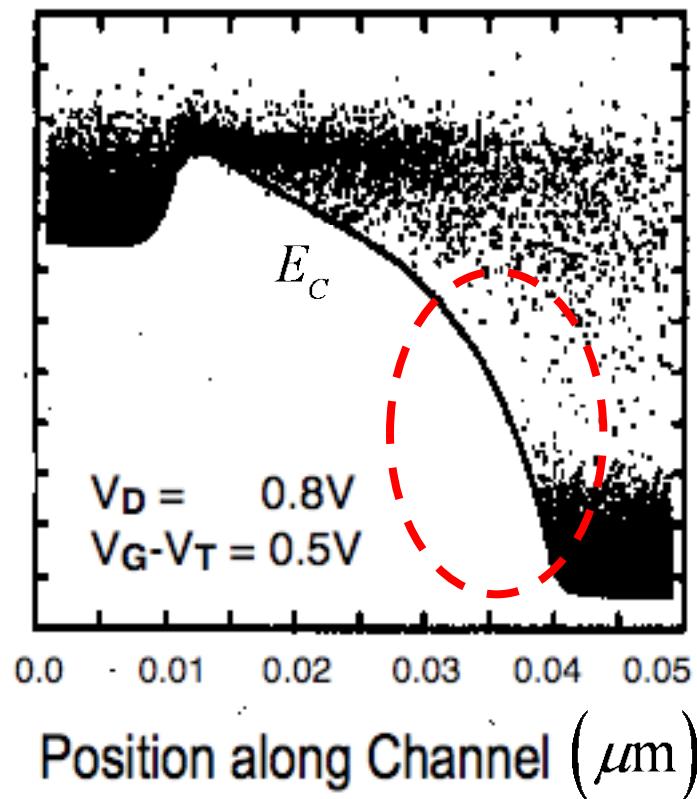
The velocity saturated MOSFET



$$I_D = \frac{W}{L} \mu_n C_{ox} (V_{GS} - V_T) V_{DS}$$

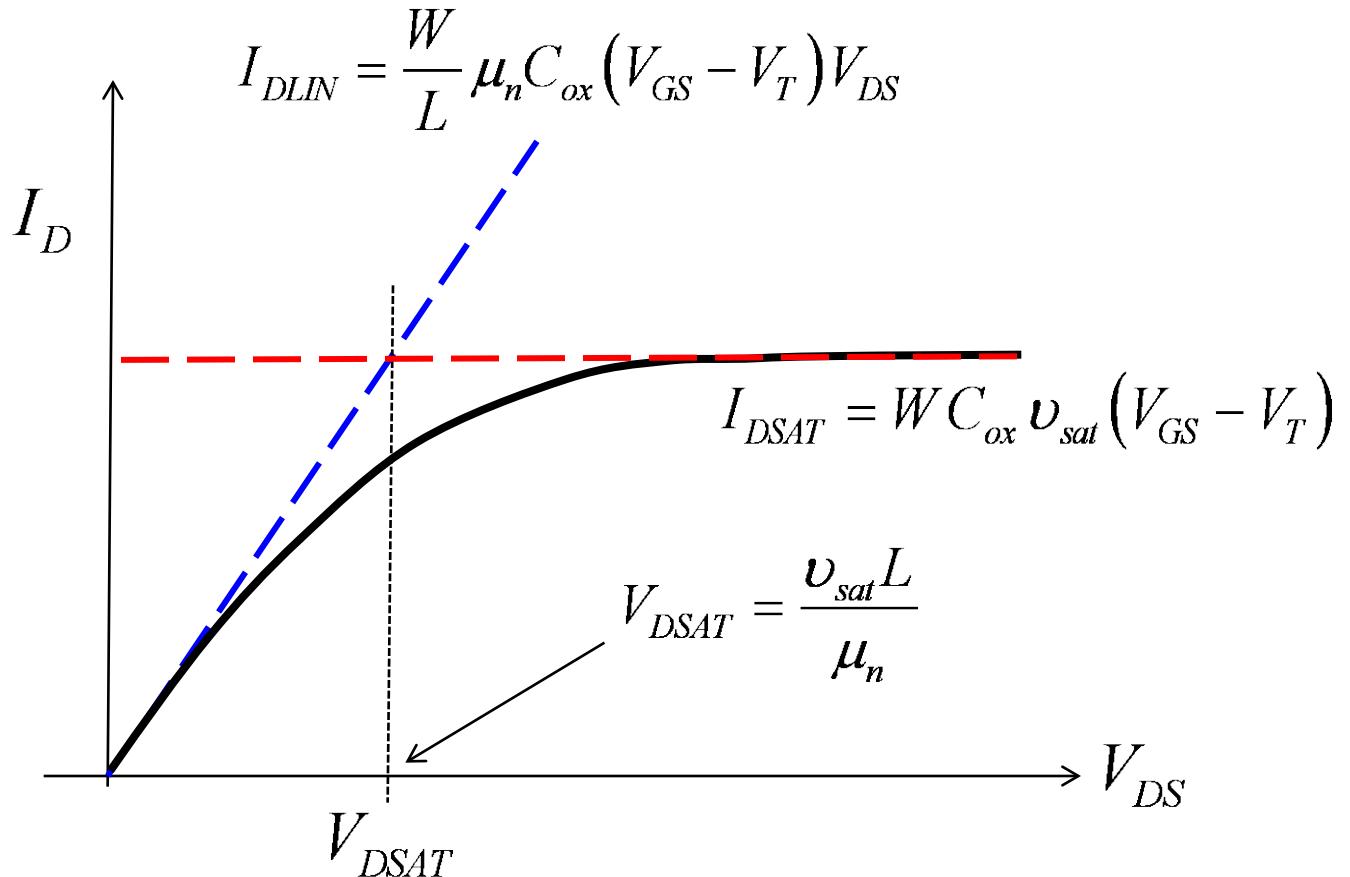
$$I_D = W C_{ox} v_{sat} (V_{GS} - V_T)$$

Velocity overshoot



D. Frank, S. Laux, and M. Fischetti, Int. Electron Dev. Mtg., Dec., 1992.

Piecewise linear model



We have developed a 2-piece approximation to the MOSFET IV characteristic.

Summary

- 1) Analytical expression that describe the linear and saturation regions of a MOSFET are easy to develop.
- 2) The velocity saturation model describes modern transistors.
- 3) Energy band diagrams, not equations, explain how transistors work.
- 4) A model that smoothly connects the linear and saturation regions is needed for circuit simulation.

Next topic:

In the next lecture, we will discuss traditional (square law) MOSFET theory in a way that describes the IV characteristics with simple analytical expressions that **smoothly connect** the linear and saturation regions.

Essentials of MOSFETs

Unit 2: Essential Physics of the MOSFET

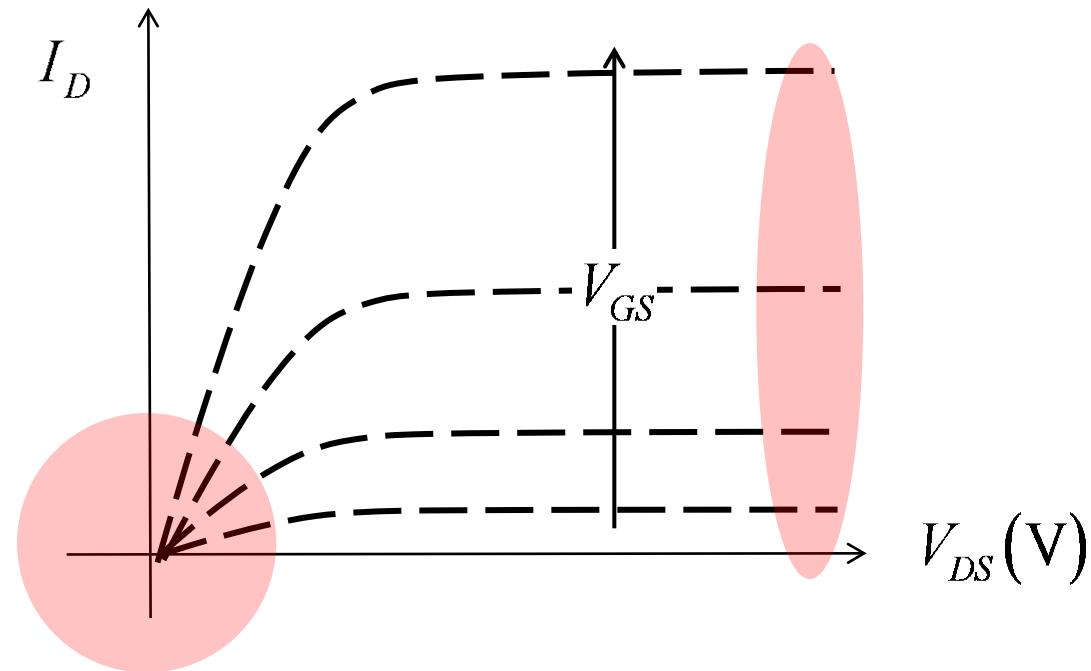
Lecture 2.4: The Square Law MOSFET

Mark Lundstrom

lundstro@purdue.edu
Electrical and Computer Engineering
Purdue University
West Lafayette, Indiana USA

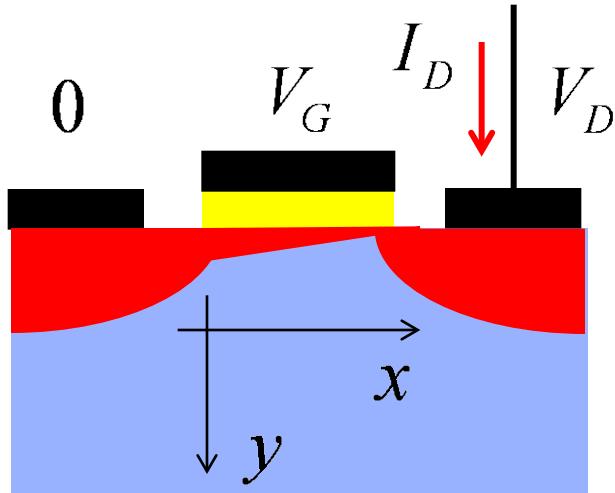
we have discussed this on a long channel or square mosfet where current is directly proportional to the voltage in its saturation reign.

Square law MOSFET theory



$$I_D = \frac{W}{L} \mu_n C_{ox} (V_{GS} - V_T) V_{DS} \quad I_D = \frac{W}{2L} \mu_n C_{ox} (V_{GS} - V_T)^2$$

I-V formulation



$$I_D = W |Q_n(x)| v_x(x) \text{ Amperes}$$

$$v_x(x) = -\mu_n \mathcal{E}_x(x) = \mu_n dV/dx$$

$$I_D = W |Q_n(x)| \mu_n \frac{dV}{dx}$$

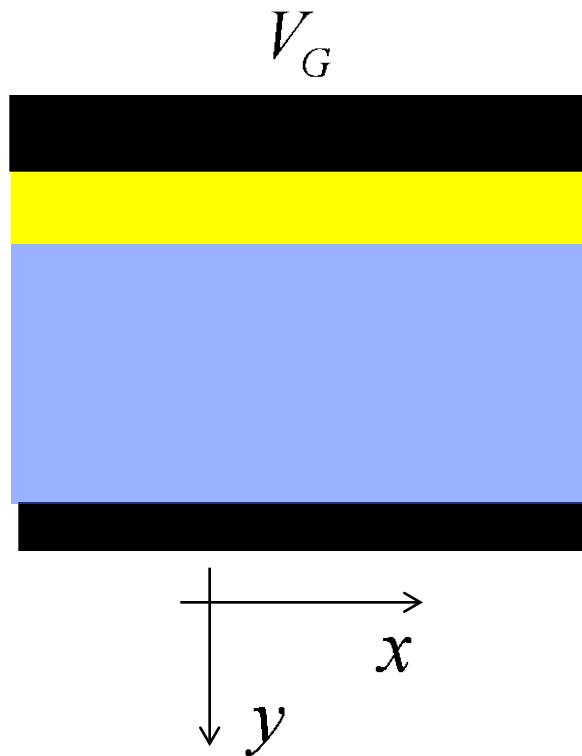
$$I_D dx = W |Q_n(V)| \mu_n dV$$

to include diffusion:

$$-\frac{dV}{dx} \rightarrow \frac{d(F_n/q)}{dx}$$

$$I_D = -\frac{W}{L} \mu_n \int_0^{V_{DS}} |Q_n(V)| dV$$

1D MOS capacitor



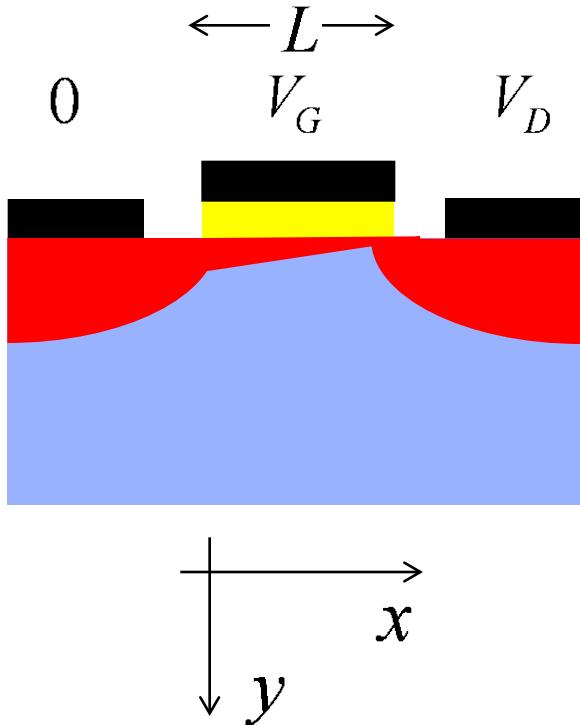
$$V_G - V_T \leq 0$$

$$Q_n \approx 0$$

$$V_G - V_T > 0$$

$$Q_n = -C_{ox} (V_G - V_T)$$

Gradual channel approximation



for $0 \leq x \leq L$

$$V = V(x)$$

$$V(0) = 0 \quad V(L) = V_D$$

1D MOS-C: $Q_n = -C_{ox}(V_G - V_T)$

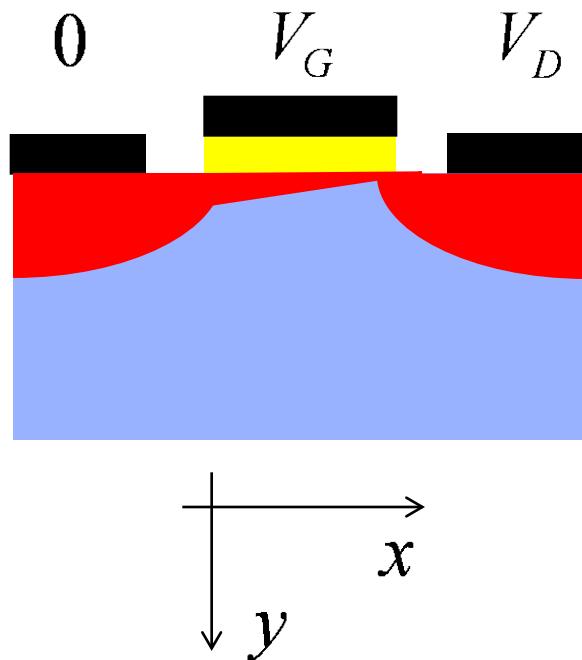
GCA: $\mathcal{E}_x \ll \mathcal{E}_y$

$$V_T \rightarrow V_T(x) = V_T + V(x)$$

$$V_G - V_T > 0$$

$$Q_n(x) = -C_{ox}[V_{GS} - V_T - V(x)]$$

IV relation



$$I_D = \frac{W}{L} \mu_n \int_0^{V_{DS}} |Q_n(V)| dV$$

$$I_D = \mu_n C_{ox} \frac{W}{L} \int_0^{V_D} [V_{GS} - V_T - V] dV$$

$$I_D = \mu_n C_{ox} \frac{W}{L} \left[(V_{GS} - V_T) V_{DS} - \frac{V_{DS}^2}{2} \right]$$

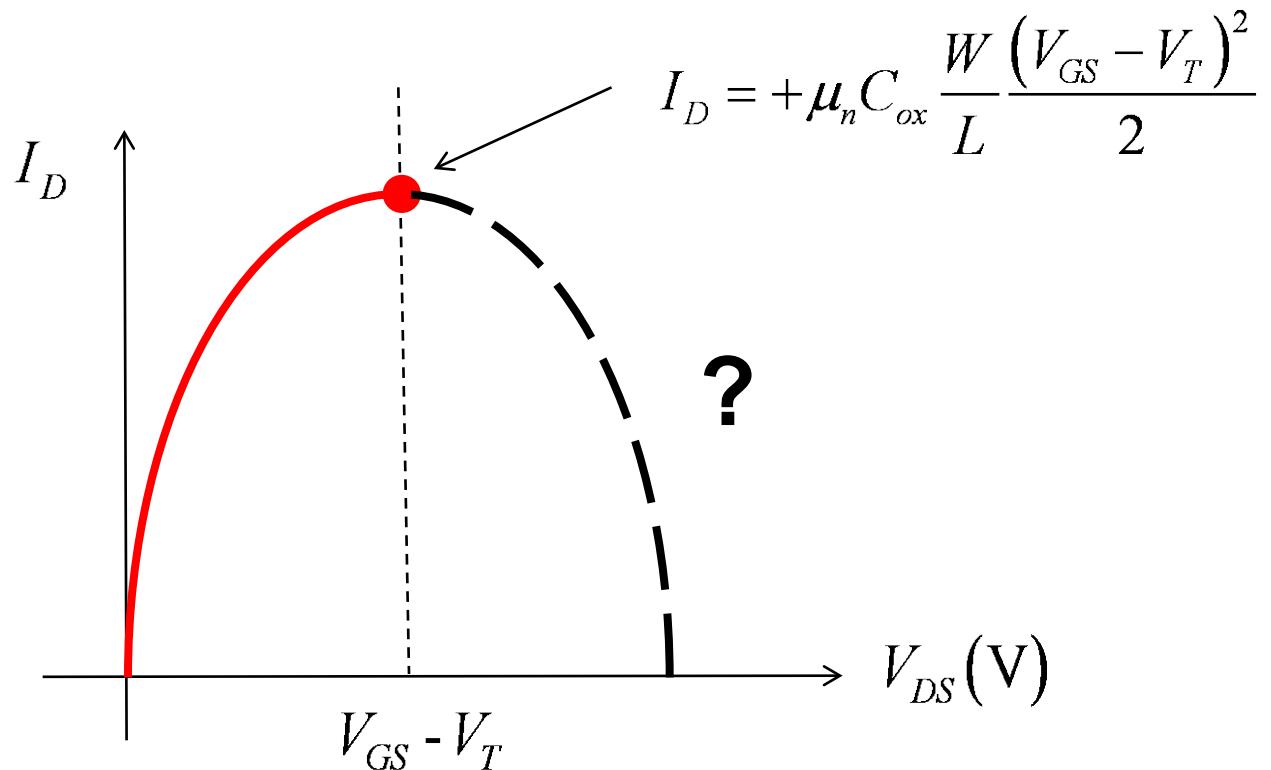
THEY ARE KNOWN AS TRIODE DEVICES BECAUSE

IN 1960 S VACUUM TUBES WERE DOMINANT

ELECTRONIC DEVICES. THE IV CHAR OF THE MOSFET
LOOKED LIKE A VACCUM TUBE TRIODE. HENCE WHY

(triode region of operation)

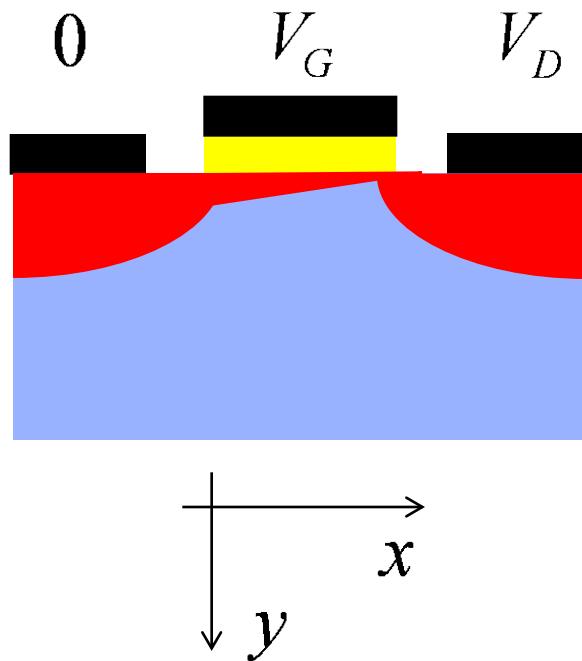
Beyond pinch-off?



$$I_D = +\mu_n C_{ox} \frac{W}{L} \left[(V_{GS} - V_T) V_{DS} - \frac{V_{DS}^2}{2} \right]$$

Lundstrom: 2018

Pinch-off



$$Q_n(L) = -C_{ox} [V_{GS} - V_T - V_D]$$

when $V_D = V_{GS} - V_T$,

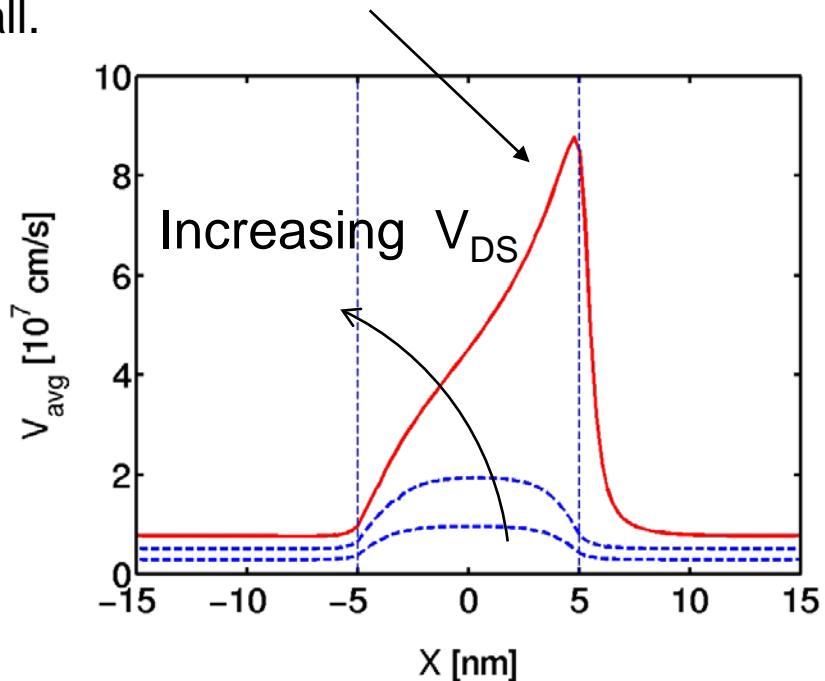
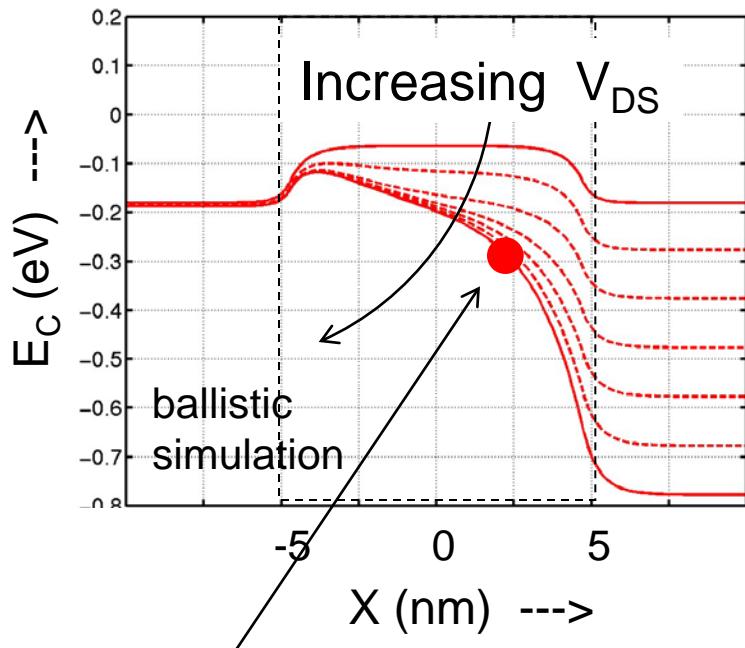
then $Q_n(L) = 0$

$\mathcal{E}_y \gg \mathcal{E}_x$ GCA fails!

but current still flows!

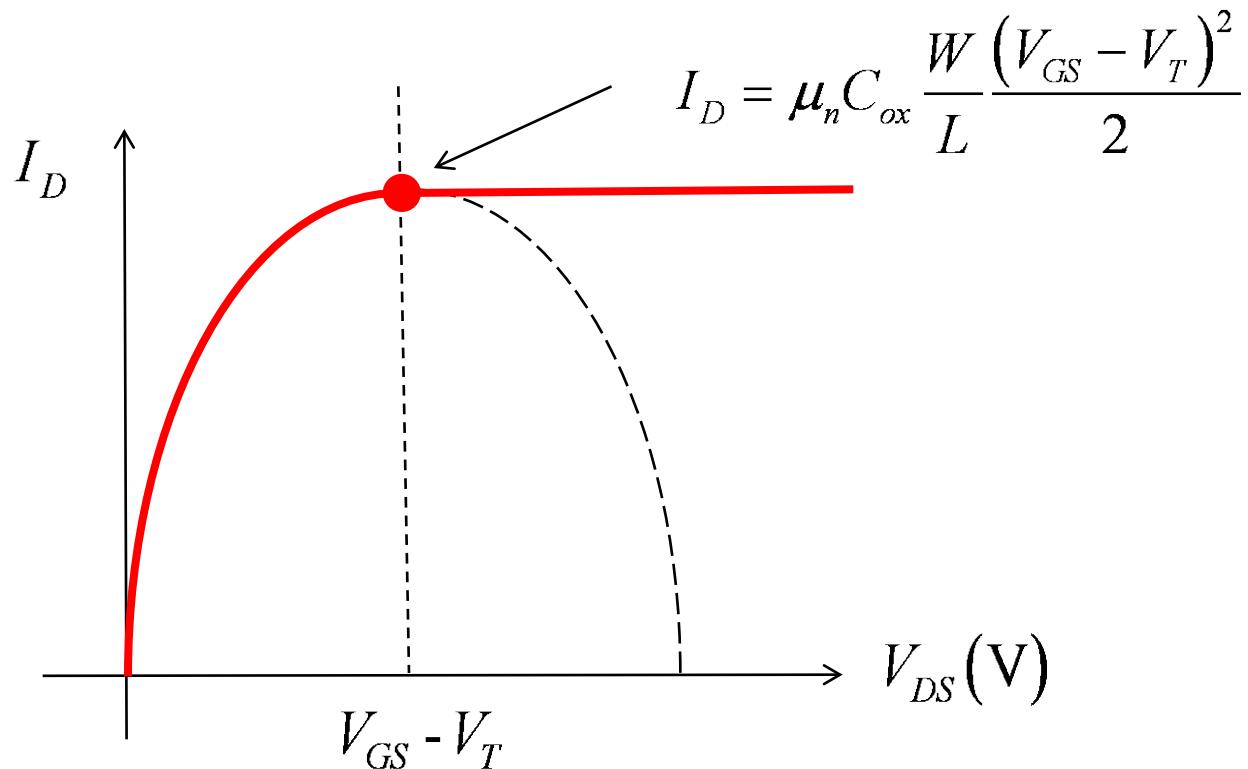
Pinch off in a MOSFET

The electron velocity is very high in the pinch-off region. High velocity implies low inversion layer density (because I_D is constant). In the textbook model, we say $Q_i \approx 0$, but it is not really zero - just very small.



pinch-off point: where the electric field along the channel becomes very large. Note that electrons are simply swept across the high-field (pinched-off) portion at very high velocity.

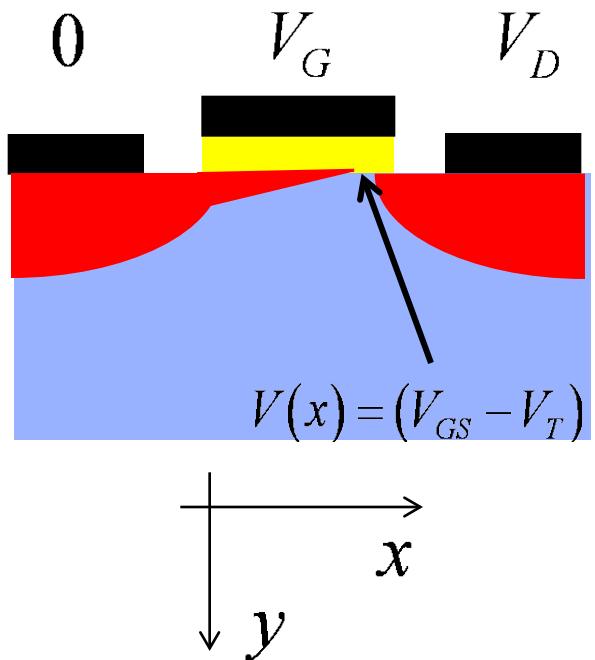
IV beyond pinch-off



$$I_D = \mu_n C_{ox} \frac{W}{L} \left[(V_{GS} - V_T) V_{DS} - \frac{V_{DS}^2}{2} \right]$$

Lundstrom: 2018

Complete IV characteristic



$$V_{GS} > V_T$$

$$V_{DS} < V_{GS} - V_T$$

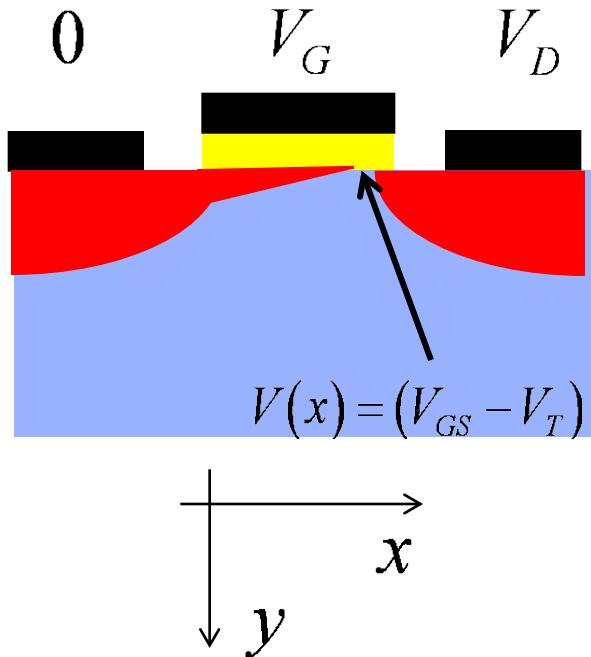
$$I_D = \mu_n C_{ox} \frac{W}{L} \left[(V_{GS} - V_T) V_{DS} - \frac{V_{DS}^2}{2} \right]$$

$$V_{GS} > V_T$$

$$V_{DS} > V_{GS} - V_T$$

$$I_D = \mu_n C_{ox} \frac{W}{2L} (V_{GS} - V_T)^2$$

Linear region



$$V_{GS} > V_T$$
$$V_{DS} < V_{GS} - V_T$$

$$I_D = \mu_n C_{ox} \frac{W}{L} \left[(V_{GS} - V_T) V_{DS} - \frac{V_{DS}^2}{2} \right]$$

$$V_{GS} > V_T$$
$$V_{DS} \ll V_{GS} - V_T$$

$$I_D = \mu_n C_{ox} \frac{W}{L} (V_{GS} - V_T) V_{DS}$$

The electric field in the channel

small V_{DS}

$$I_D = \mu_n C_{ox} \frac{W}{L} (V_{GS} - V_T) V_{DS}$$

$$I_D = W |Q_n(x=0)| \langle v_x(x=0) \rangle$$

$$I_D = W C_{ox} (V_{GS} - V_T) (-\mu_n \mathcal{E}_x(0))$$

$$\mathcal{E}_x(0) = -\frac{V_{DS}}{L}$$

large V_{DS}

$$I_D = \mu_n C_{ox} \frac{W}{2L} (V_{GS} - V_T)^2$$

$$I_D = W |Q_n(x=0)| \langle v_x(x=0) \rangle$$

$$I_D = W C_{ox} (V_{GS} - V_T) (-\mu_n \mathcal{E}_x(0))$$

$$\mathcal{E}_x(0) = -\frac{(V_{GS} - V_T)}{2L}$$

Summary

Triode region

$$V_{GS} > V_T$$

$$V_{DS} < V_{GS} - V_T$$

$$I_D = \mu_n C_{ox} \frac{W}{L} \left[(V_{GS} - V_T) V_{DS} - \frac{V_{DS}^2}{2} \right]$$

Beyond pinch-off region
(saturation region)

$$V_{GS} > V_T$$

$$V_{DS} > V_{GS} - V_T$$

$$I_D = \mu_n C_{ox} \frac{W}{2L} (V_{GS} - V_T)^2$$

Linear region

$$V_{DS} \ll V_{GS} - V_T$$

$$I_D = \mu_n C_{ox} \frac{W}{L} (V_{GS} - V_T) V_{DS}$$

Next topic:

Modern MOSFETs are not square law devices, but this example is an illustration of a model that works smoothly from the linear to saturation region.

A full range IV characteristic for velocity saturated MOSFETs can be developed, but it is a bit more complicated.

In the next lecture, we will show how our two-piece velocity saturated model can be easily extended to a full range model.

Essentials of MOSFETs

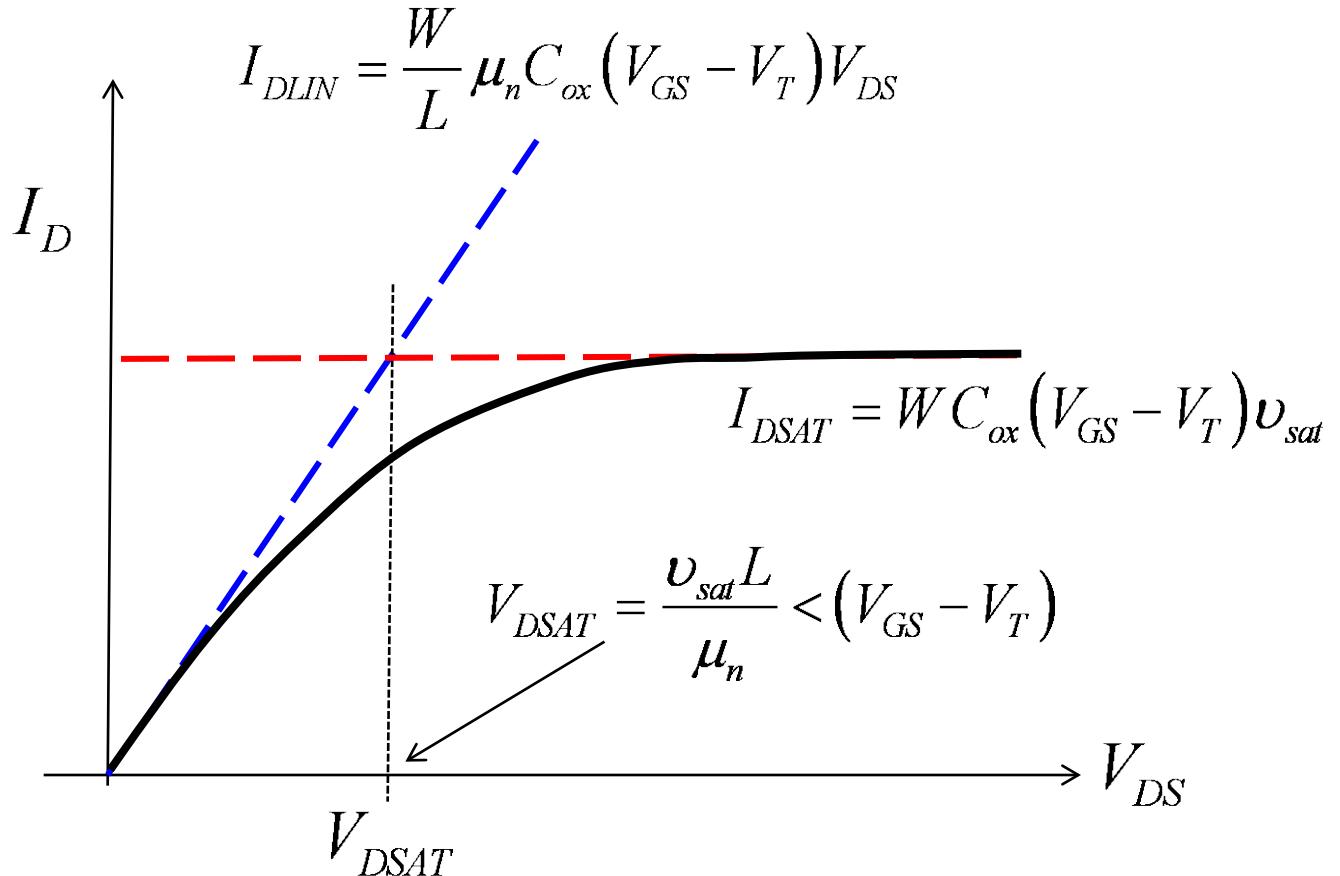
Unit 2: Essential Physics of the MOSFET

Lecture 2.5: The Virtual Source Model

Mark Lundstrom

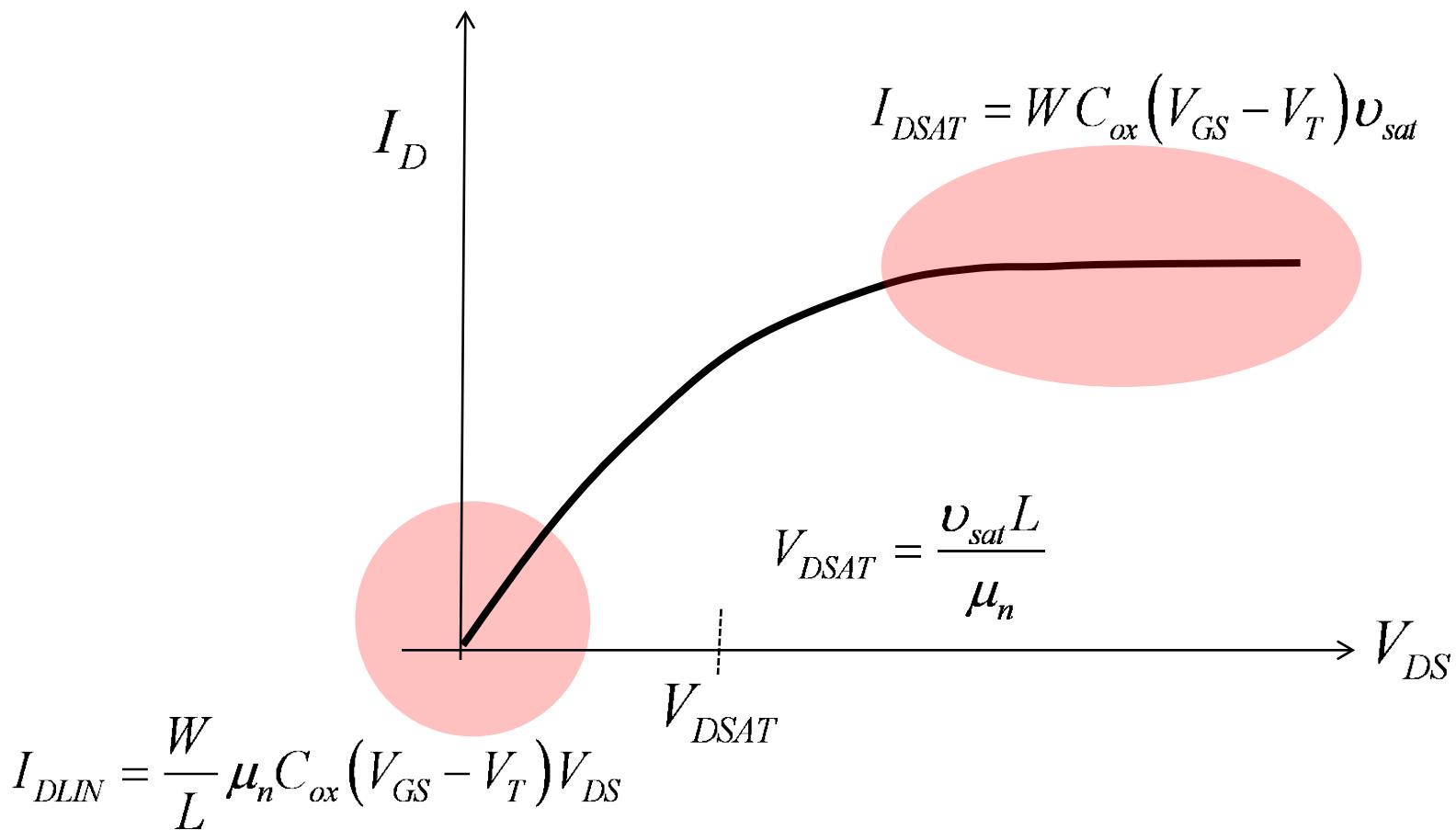
lundstro@purdue.edu
Electrical and Computer Engineering
Purdue University
West Lafayette, Indiana USA

Velocity saturated MOSFET: IV (review)



We have developed a 2-piece approximation to the MOSFET IV characteristic.

From a two-piece to continuous model



Can we produce a model that smoothly goes from the linear to saturation regions?

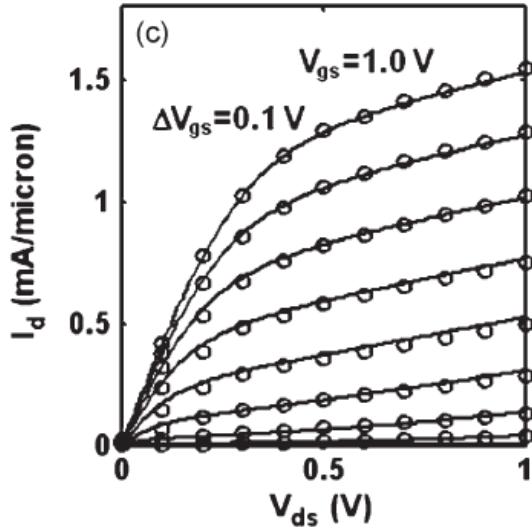
The MIT Virtual Source (VS) model

1674

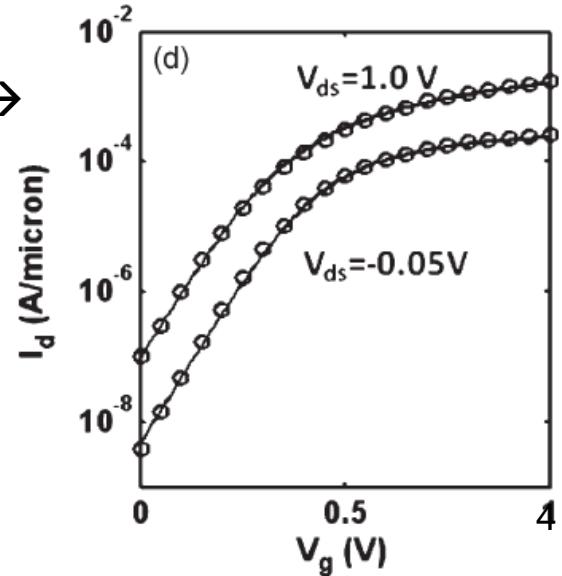
IEEE TRANSACTIONS ON ELECTRON DEVICES, VOL. 56, NO. 8, AUGUST 2009

A Simple Semiempirical Short-Channel MOSFET Current–Voltage Model Continuous Across All Regions of Operation and Employing Only Physical Parameters

Ali Khakifirooz, *Member, IEEE*, Osama M. Nayfeh, *Member, IEEE*, and Dimitri Antoniadis, *Fellow, IEEE*



← 32 nm technology →



Lundstrom: 2018

Piecewise model for $I_D(V_{GS}, V_{DS})$

$$I_D/W = |Q_n(V_{GS})| \langle v_x(V_{DS}) \rangle$$

$$V_{GS} \geq V_T : Q_n(V_{GS}) = -C_{ox}(V_{GS} - V_T) \quad V_{DS} \leq V_{DSAT} : \langle v_x(V_{DS}) \rangle = \left(\mu_n \frac{V_{DS}}{L} \right)$$

$$V_{GS} < V_T : Q_n(V_{GS}) = 0 \quad V_{DS} > V_{DSAT} : \langle v_x(V_{DS}) \rangle = v_{sat}$$

If we can make the average velocity go smoothly from the low V_{DS} to high V_{DS} limits, then we will have a smooth model for $I_D(V_{GS}, V_{DS})$ – above threshold.

From low V_{DS} to high V_{DS}

$$\frac{1}{\langle v_x(V_{DS}) \rangle} = \frac{1}{\mu_n V_{DS}/L} + \frac{1}{v_{sat}} \rightarrow \langle v_x(V_{DS}) \rangle = \left[\frac{V_{DS}/V_{DSAT}}{1 + V_{DS}/V_{DSAT}} \right] v_{sat}$$

smaller velocity having component will dominate.

$$V_{DSAT} = v_{sat} L / \mu_n$$

$$\langle v_x(V_{DS}) \rangle = F_{SAT}(V_{DS}) v_{sat}$$

$$F_{SAT}(V_{DS}) = \frac{V_{DS}/V_{DSAT}}{\left[1 + (V_{DS}/V_{DSAT})^\beta \right]^{1/\beta}}$$

The extra parameter, β , is empirically adjusted to fit the IV characteristic. Typically, $\beta \approx 1.4 - 1.8$ for both N-MOSFETs and for P-MOSFETs. (semi-empirical)

Empirical saturation function

$$\langle v_x(V_{DS}) \rangle = F_{SAT}(V_{DS}) v_{sat}$$

$$F_{SAT}(V_{DS}) \equiv \frac{V_{DS}/V_{DSAT}}{\left[1 + (V_{DS}/V_{DSAT})^\beta\right]^{1/\beta}}$$

$$V_{DS} \ll V_{DSAT} : F_{SAT}(V_{DS}) \rightarrow \frac{V_{DS}}{V_{DSAT}}$$

$$V_{DS} \gg V_{DSAT} : F_{SAT}(V_{DS}) \rightarrow 1$$

$$\langle v_x(V_{DS}) \rangle \rightarrow \frac{V_{DS}}{V_{DSAT}} v_{sat}$$

$$\langle v_x(V_{DS}) \rangle \rightarrow v_{sat} \quad \checkmark$$

$$\langle v_x(V_{DS}) \rangle \rightarrow \frac{V_{DS}}{v_{sat} L / \mu_n} v_{sat}$$

$$\langle v_x(V_{DS}) \rangle \rightarrow \mu_n \frac{V_{DS}}{L} \quad \checkmark$$

Lundstrom: 2018

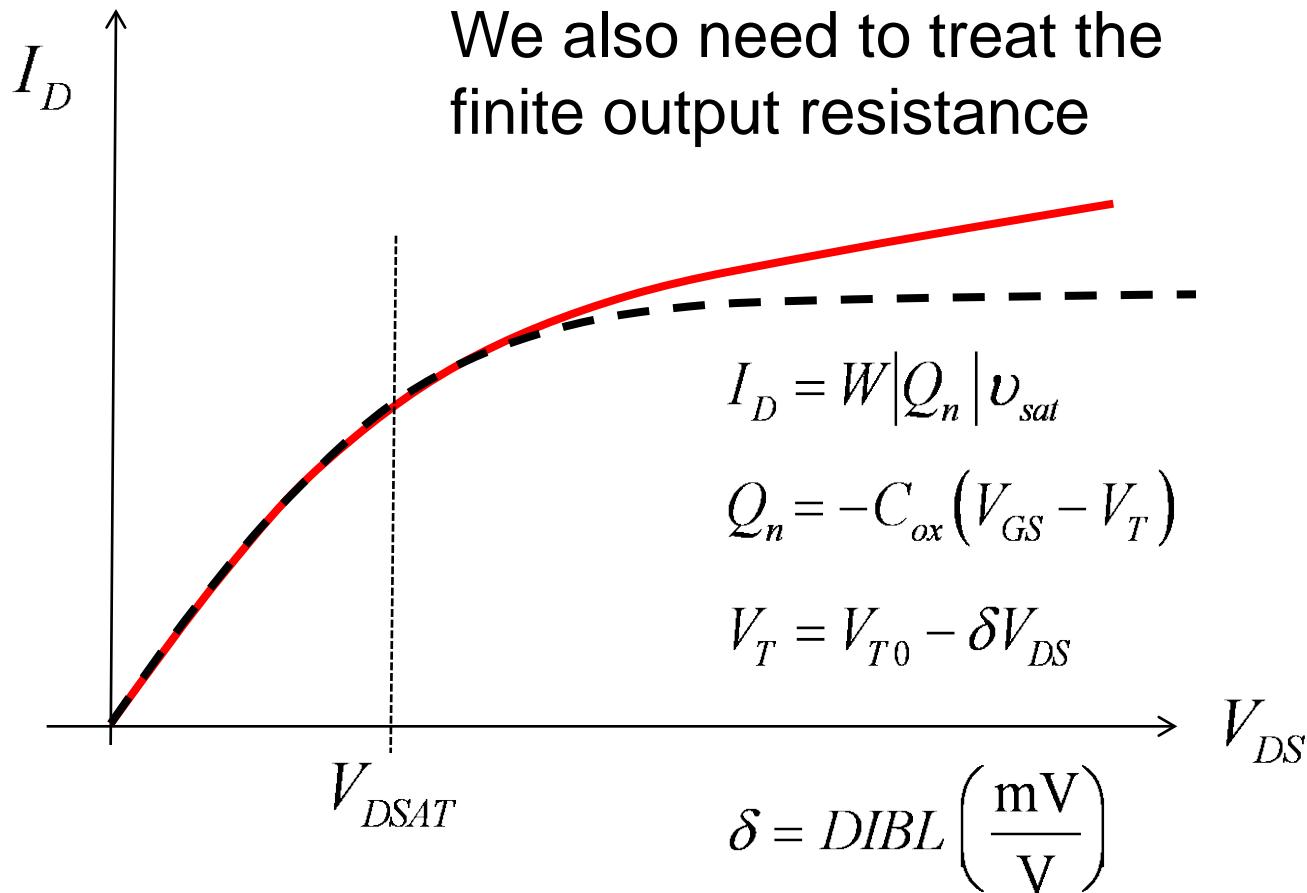
Saturation function: $F_{SAT}(V_D)$

$$\langle v_x(V_{DS}) \rangle = F_{SAT}(V_{DS}) v_{sat}$$

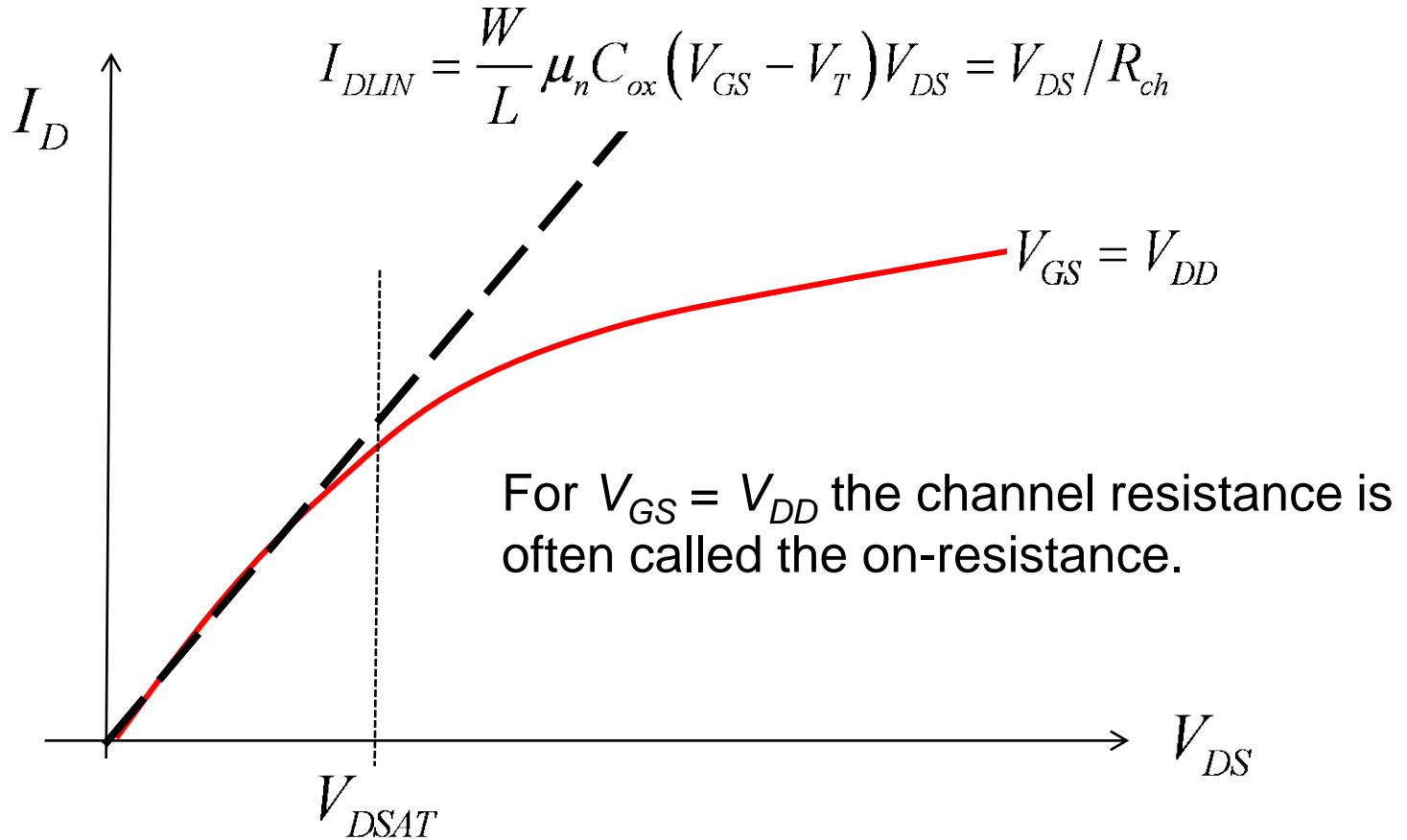
$$F_{SAT}(V_{DS}) = \frac{V_{DS}/V_{DSAT}}{\left[1 + (V_{DS}/V_{DSAT})^\beta\right]^{1/\beta}}$$

Although this is just an empirical method to produce a smooth curve that properly goes between the small and large V_D limits, it works very well in practice (and for several types of FETs), which suggests that it captures something important about MOSFETs.

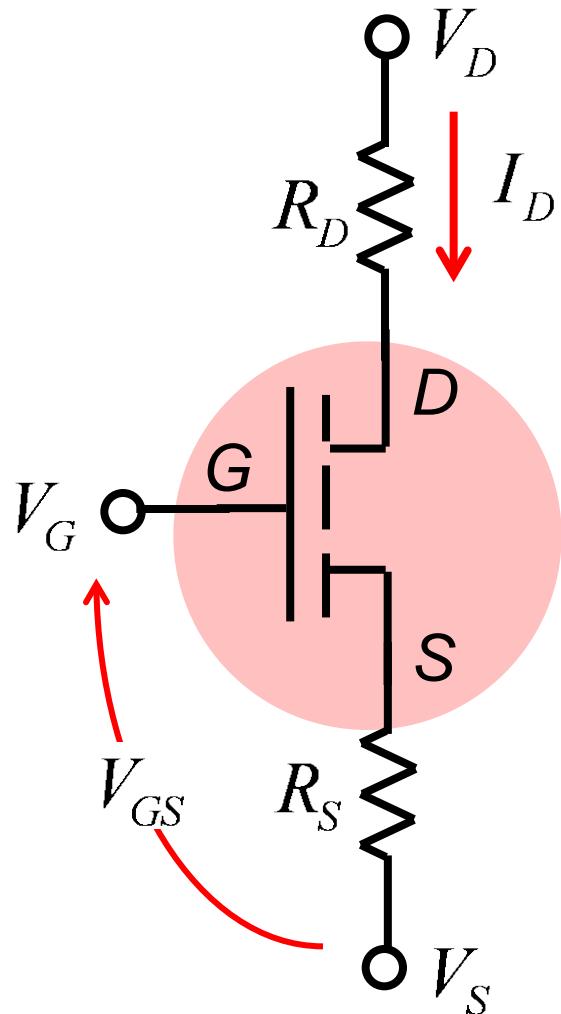
Output resistance



Channel resistance



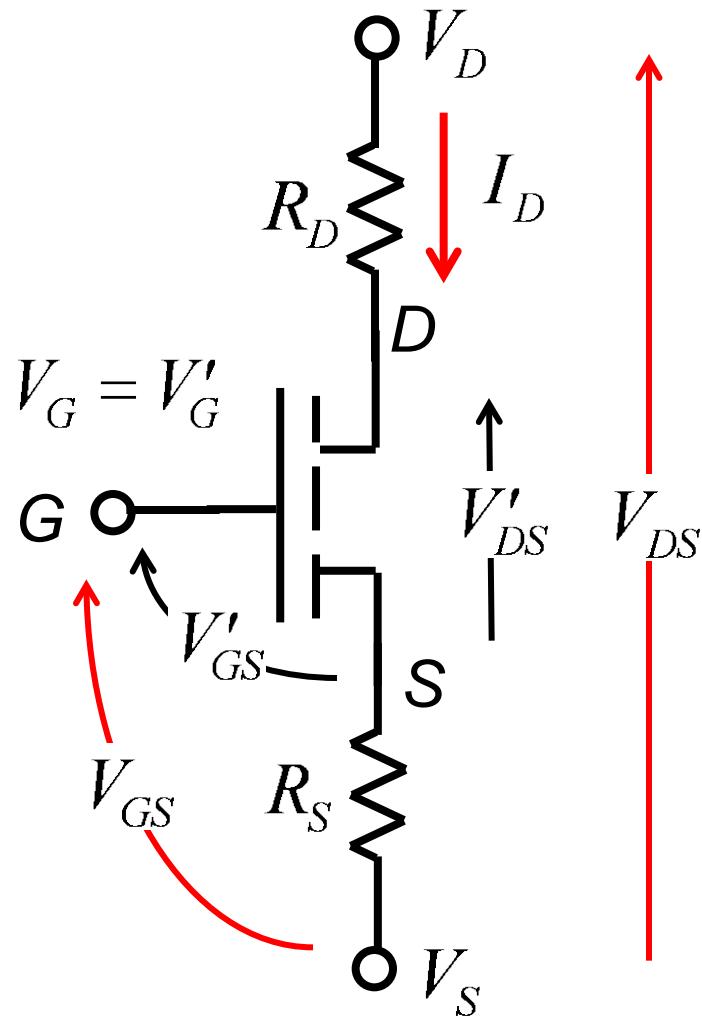
Series resistance



Parasitic resistances connect the intrinsic MOSFET to the contacts.

(There is a gate resistance too, which can be important for RF applications.)

Intrinsic vs. extrinsic voltages

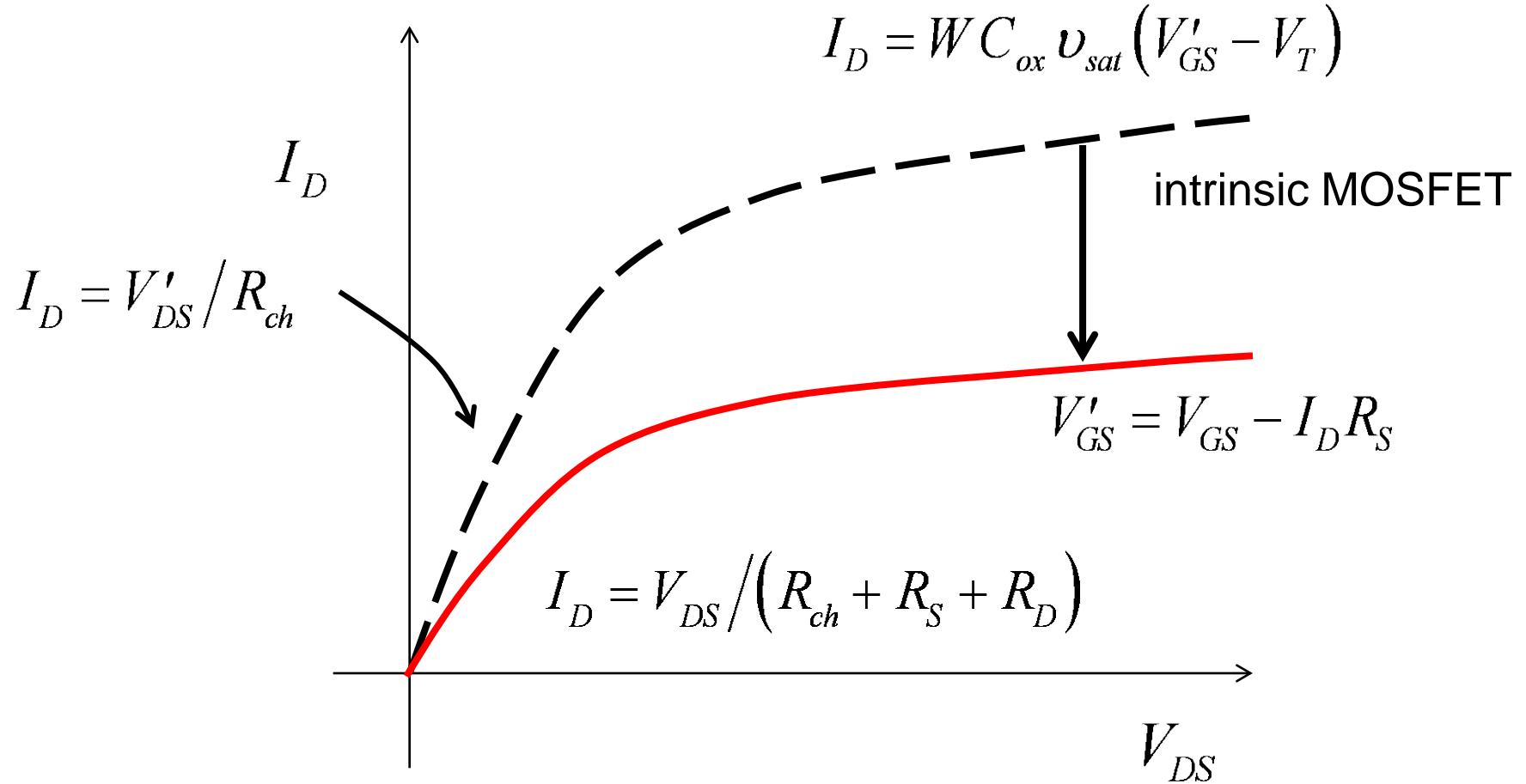


$$V'_G = V_G$$

$$V'_D = V_D - I_D(V'_G, V'_S, V'_D)R_D$$

$$V'_S = V_S + I_D(V'_G, V'_S, V'_D)R_S$$

Two effects of series resistances



Simple (level 0) VS model

$$1) \quad I_D/W = |Q_n(V'_{GS})| \langle v(V'_{DS}) \rangle$$

$$2) \quad Q_n(V'_{GS}) = -C_{ox}(V'_{GS} - V_T) \quad (V'_{GS} > V_T)$$

$$V_T = V_{T0} - \delta V'_{DS}$$

$$3) \quad \langle v(V'_{DS}) \rangle = F_{SAT}(V'_{DS}) v_{sat}$$

$$4) \quad F_{SAT}(V'_{DS}) = \frac{V'_{DS}/V_{DSAT}}{\left[1 + (V'_{DS}/V_{DSAT})^\beta\right]^{1/\beta}}$$

$$5) \quad V_{DSAT} = \frac{v_{sat} L}{\mu_n}$$

There are only 8 device-specific parameters in this model:

$$C_{ox}, V_{T0}, \delta, v_{sat}, \mu_n, L$$

$$R_{SD} = R_S + R_D$$

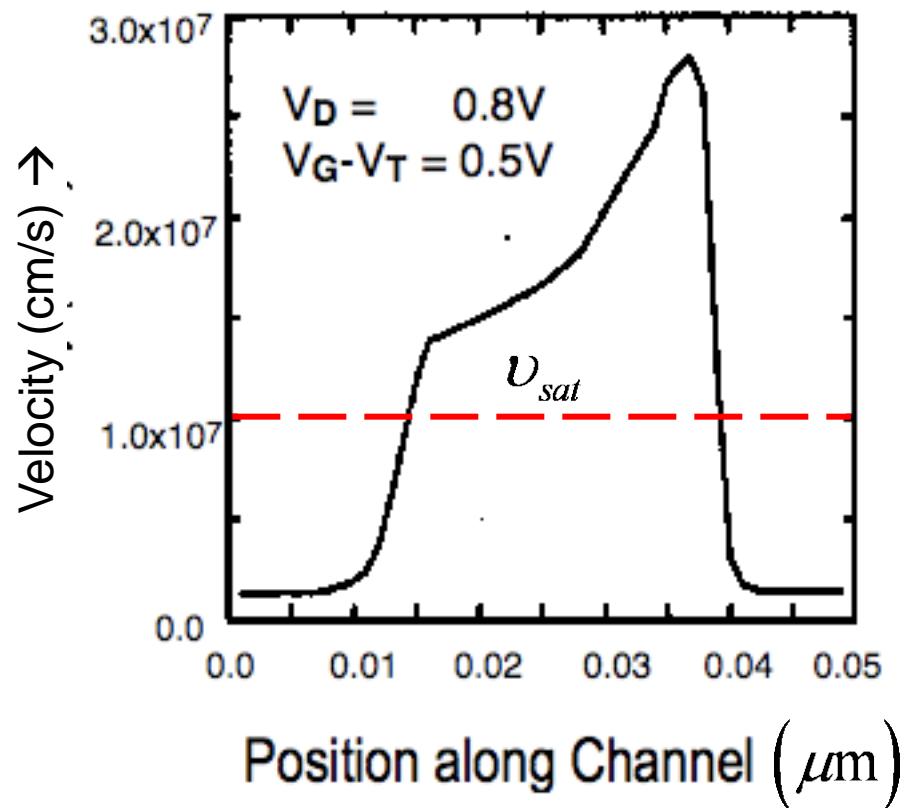
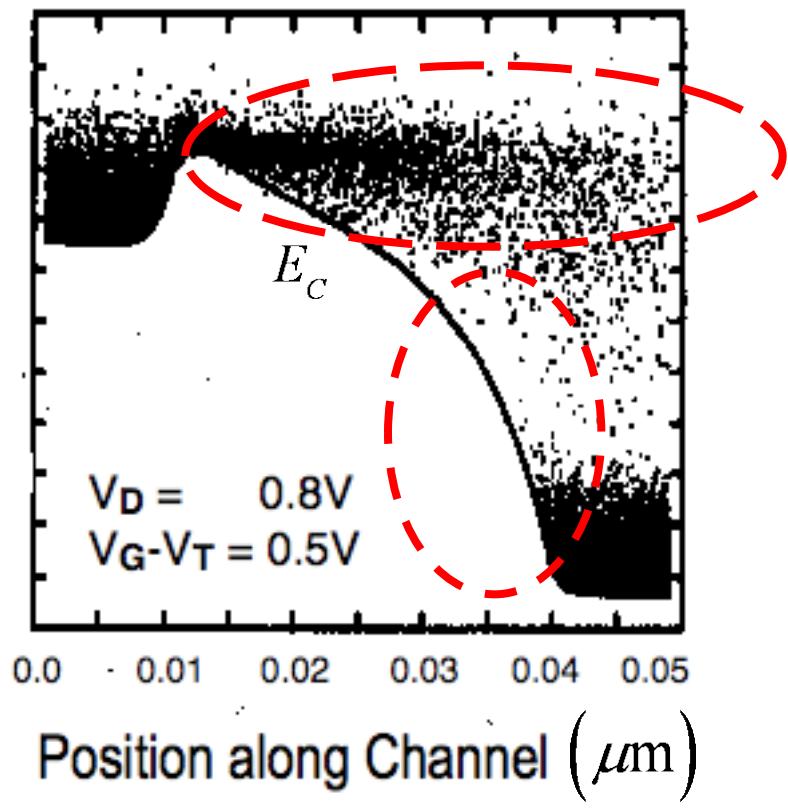
$$+ \quad \beta$$

Physics of the VS model

- 1) Diffusive (collision-dominated) transport is assumed, so that mobility is a well-defined concept.
- 2) We assume that the carrier velocity is clamped at the high-field (scattering limited) saturation velocity in the bulk.

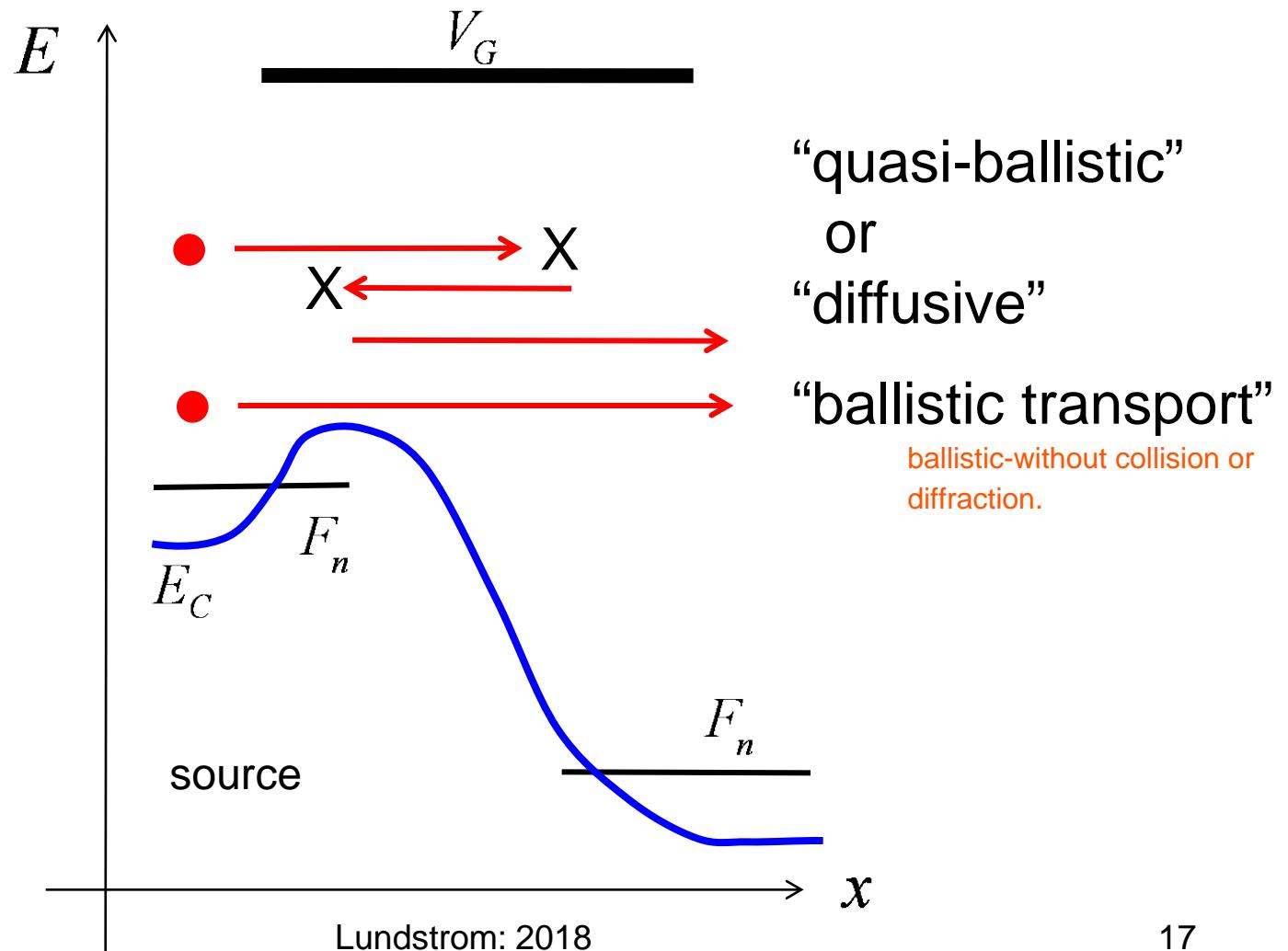
Neither of these assumptions is valid for modern, nanoscale MOSFETs.

Velocity overshoot



D. Frank, S. Laux, and M. Fischetti, Int. Electron Dev. Mtg., Dec., 1992.

Importance of transport



The MIT VS Model

In spite of these concerns, the MVS model does a remarkably good job of fitting the IV characteristics of nanoscale FETs.

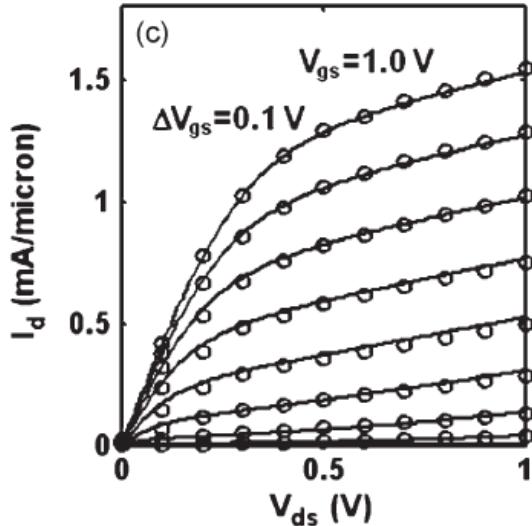
The MIT VS Model

1674

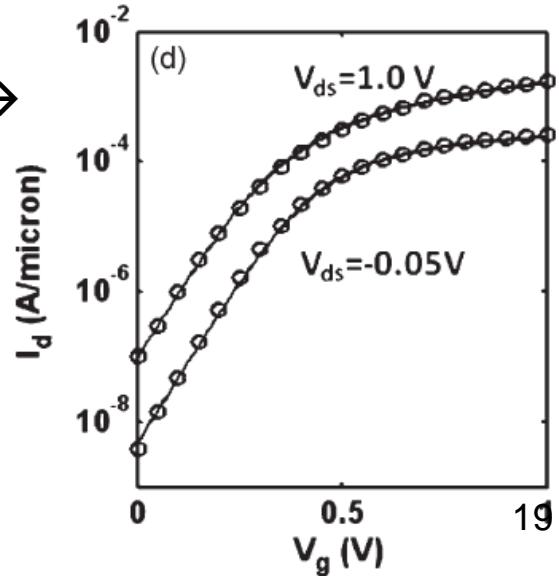
IEEE TRANSACTIONS ON ELECTRON DEVICES, VOL. 56, NO. 8, AUGUST 2009

A Simple Semiempirical Short-Channel MOSFET Current–Voltage Model Continuous Across All Regions of Operation and Employing Only Physical Parameters

Ali Khakifirooz, *Member, IEEE*, Osama M. Nayfeh, *Member, IEEE*, and Dimitri Antoniadis, *Fellow, IEEE*



← 32 nm technology →



Lundstrom: 2018

The MIT VS Model

In this course, we will show that the mobility and high-field saturation velocity should be re-interpreted:

$$\frac{1}{\mu_n} \rightarrow \frac{1}{\mu_{app}}$$

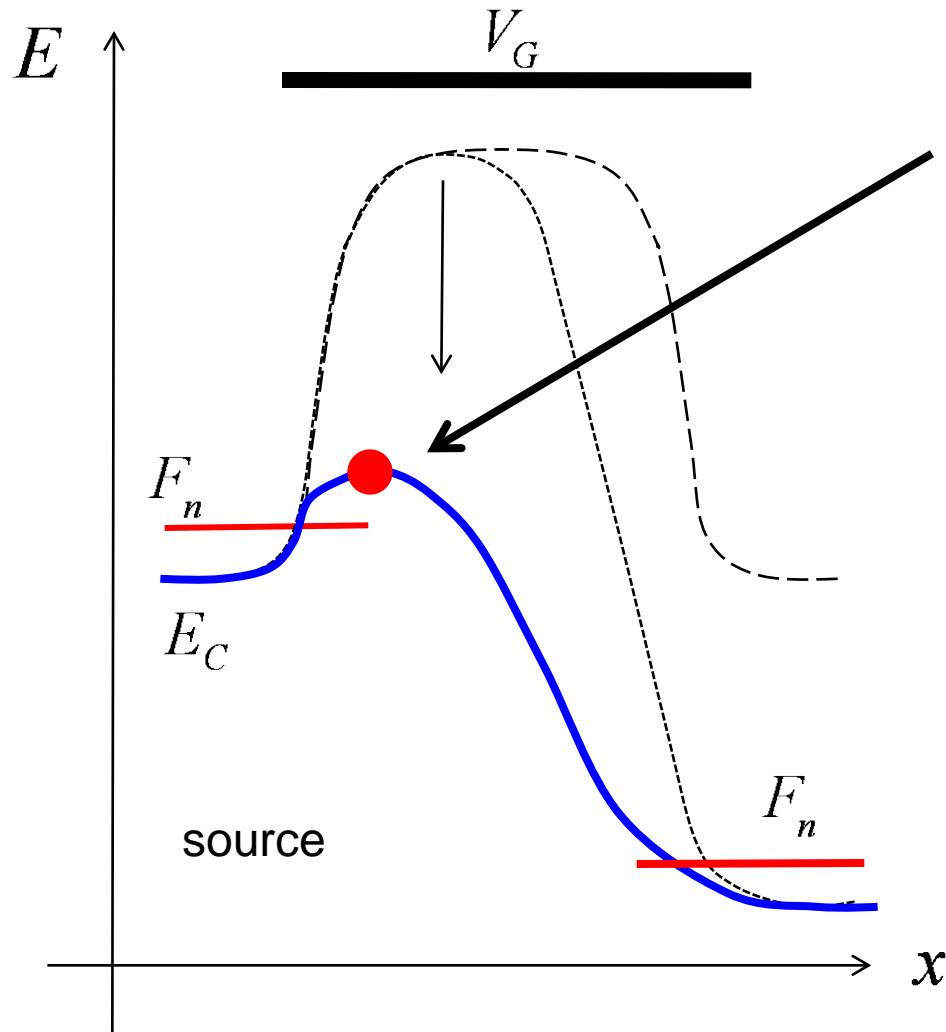
“apparent mobility”

$$v_{sat} \rightarrow v_{inj}$$

“injection velocity”

We will show that these two parameters have clear, well-defined physical interpretations.

Focus on the “virtual source”



$$Q_n = -C_{ox}(V_{GS} - V_T)$$

nearly
independent
of drain bias

This is true in an
electrostatically
well-designed
MOSFET.

Summary

$$1) \quad I_D/W = Q_n(V'_{GS}, V'_{DS}) \langle v(V'_{DS}) \rangle$$

$$2) \quad Q_n(V'_{GS}) = -C_{ox}(V'_{GS} - V_T) \quad (V'_{GS} > V_T)$$
$$V_T = V_{T0} - \delta V'_{DS}$$

$$3) \quad \langle v(V'_{DS}) \rangle = F_{SAT}(V'_{DS}) v_{sat}$$

$$4) \quad F_{SAT}(V'_{DS}) = \frac{V'_{DS}/V_{DSAT}}{\left[1 + (V'_{DS}/V_{DSAT})^\beta\right]^{1/\beta}}$$

$$5) \quad V_{DSAT} = \frac{v_{sat} L}{\mu_n}$$

There are only 8 device-specific parameters in this model:

$$C_{ox}, V_{T0}, \delta, v_{sat}, \mu_n, L$$

$$R_{SD} = R_S + R_D$$

$$+ \quad \beta$$

Essentials of MOSFETs

Unit 2: Essential Physics of the MOSFET

Lecture 2.6: Unit 2 Recap

Mark Lundstrom

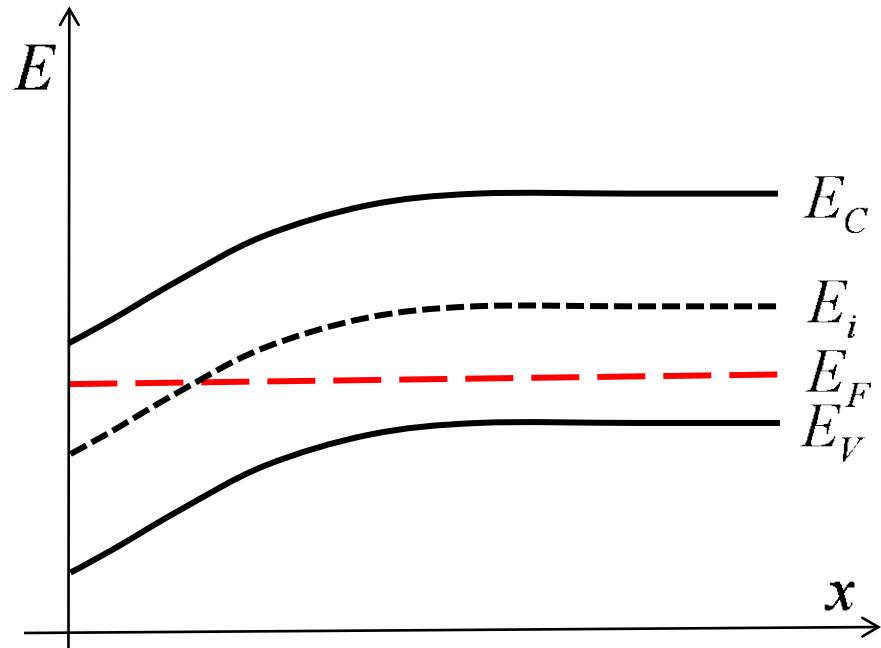
lundstro@purdue.edu
Electrical and Computer Engineering
Purdue University
West Lafayette, Indiana USA

Unit 2

- 2.1 Energy Band Review
- 2.2 Energy Band View of the MOSFET
- 2.3 MOSFET IV Theory
- 2.4 The Square Law MOSFET
- 2.5 The Virtual Source Model

Review of energy band diagrams

Draw the band diagram



$$\frac{d\mathcal{E}}{dx} = \frac{\rho(x)}{K_S \mathcal{E}_0}$$

Read the band diagram

$$V(x) \propto -E_C(x)$$

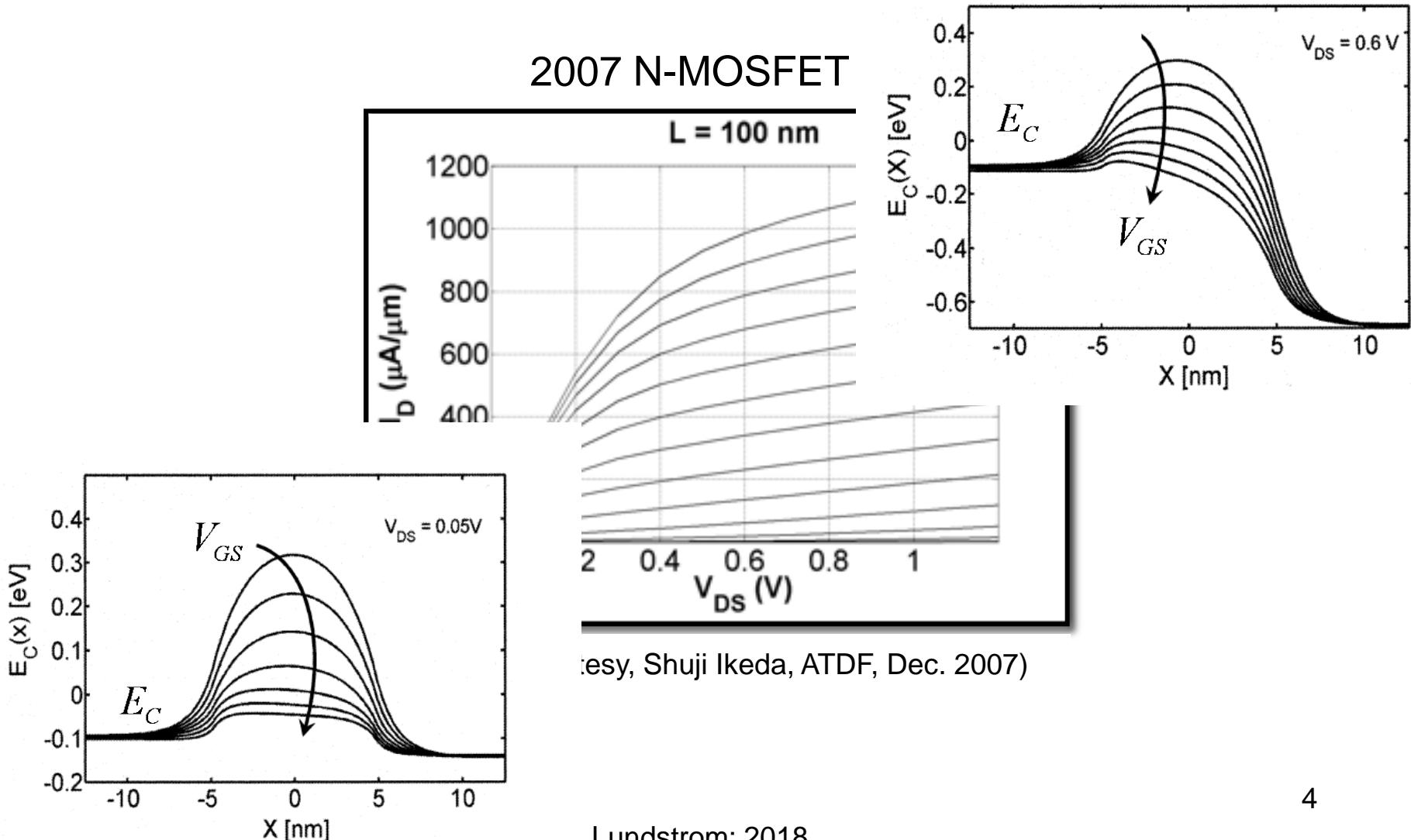
$$\mathcal{E} \propto dE_C(x)/dx$$

$$\log n(x) \propto E_F - E_i(x)$$

$$\log p(x) \propto E_i(x) - E_F$$

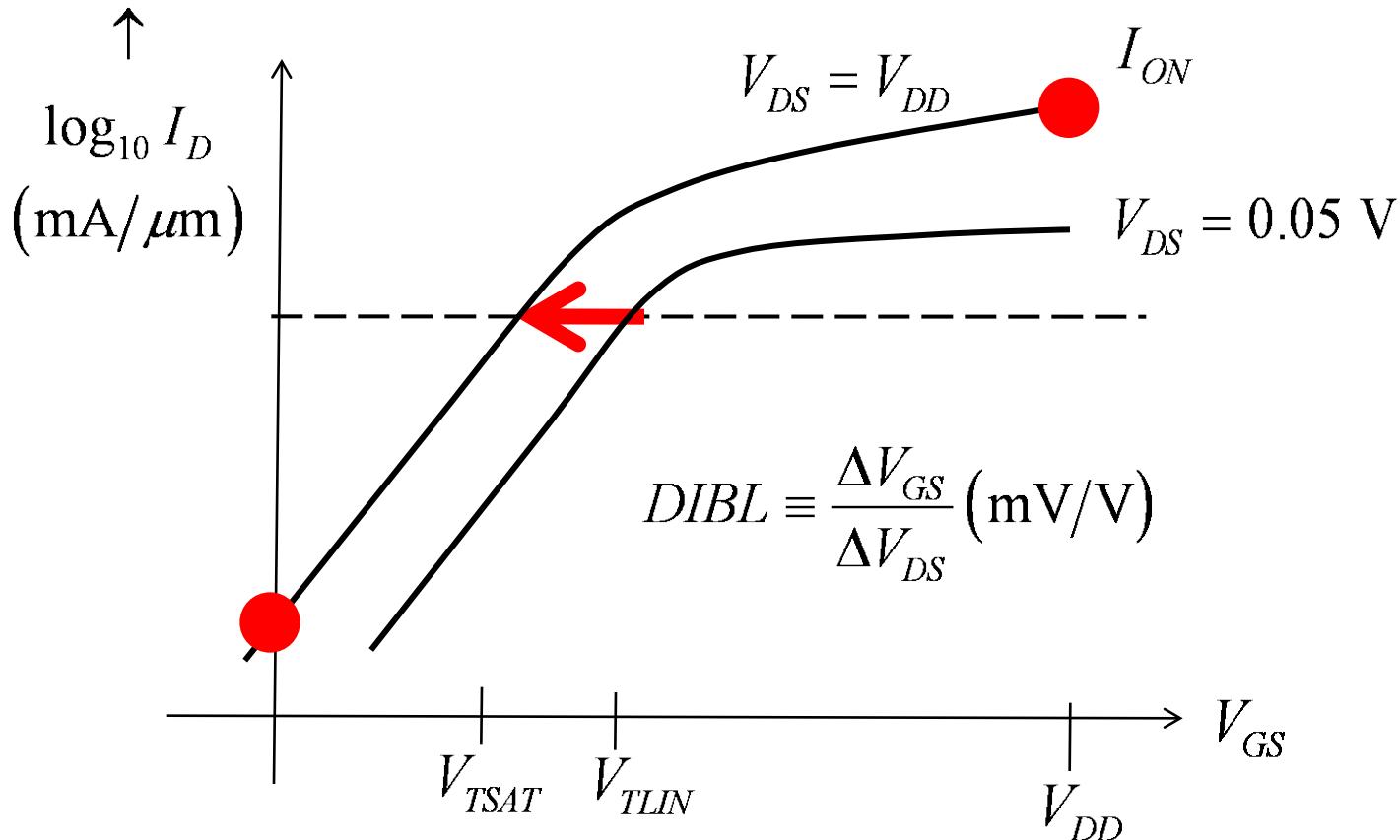
$$\rho(x) \propto (p - n + N_D - N_A)$$

How transistors work

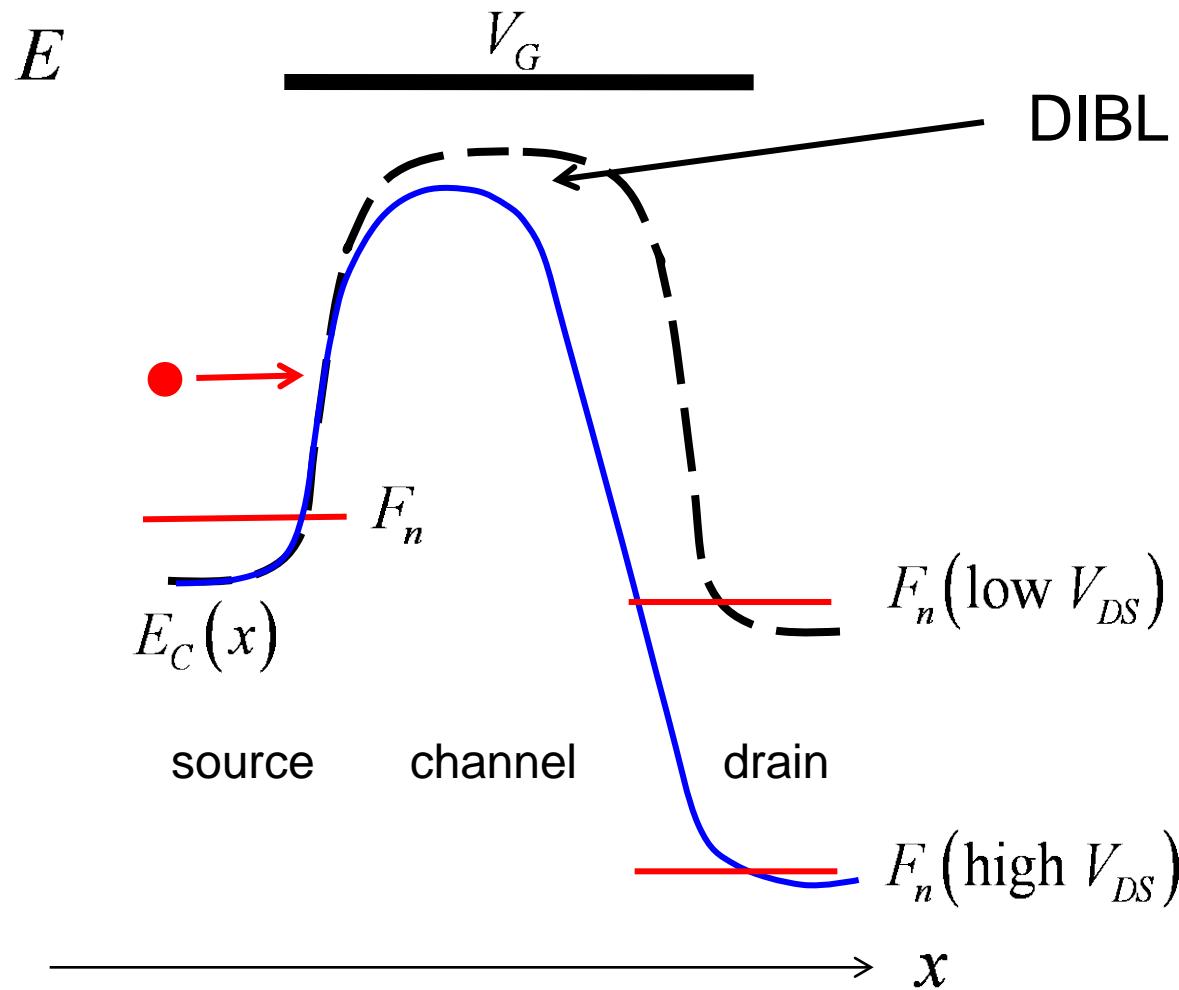


DIBL

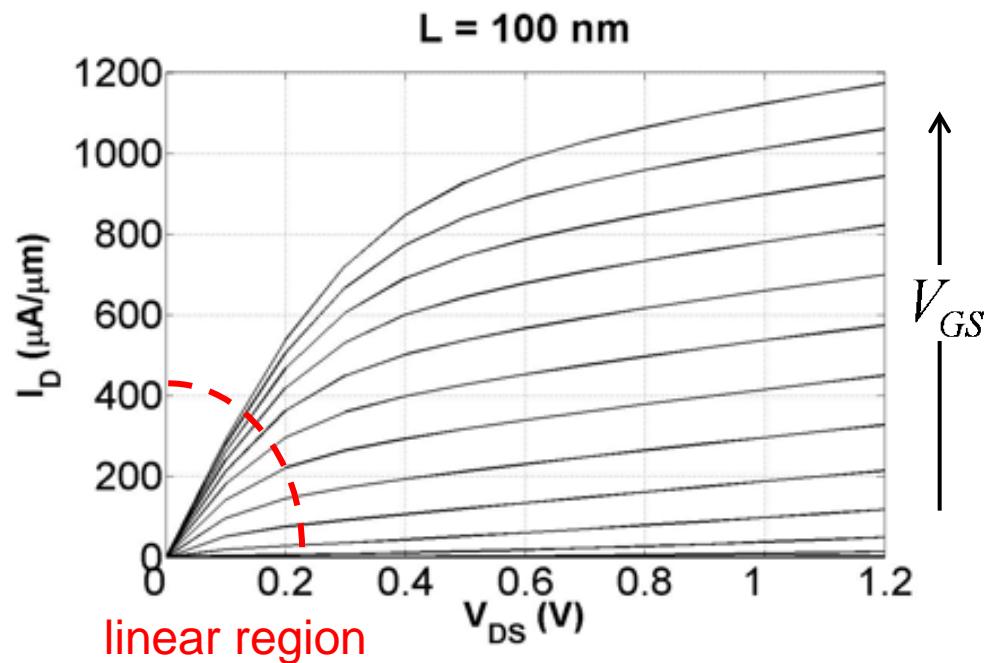
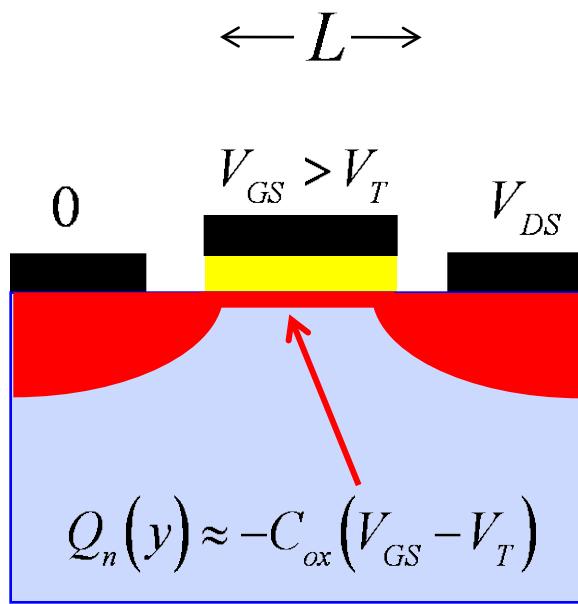
transfer characteristics:



Understanding DIBL with an e-band diagram



MOSFET IV: Low V_{DS} (linear region)



$$I_D = W |Q_n(x)| \langle v_x(x) \rangle$$

$$Q_n = -C_{ox}(V_{GS} - V_T)$$

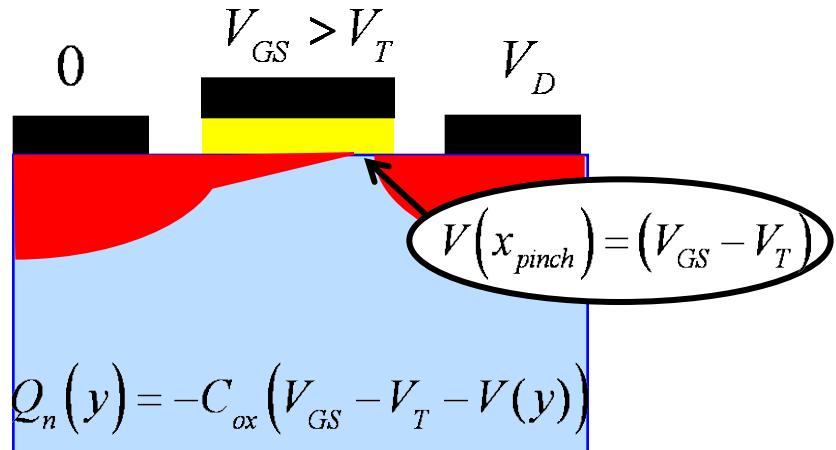
$$\langle v_x \rangle = -\mu_n \mathcal{E}_x$$

$$\mathcal{E}_x = -V_{DS}/L$$

$$I_D = \frac{W}{L} \mu_n C_{ox} (V_{GS} - V_T) V_{DS}$$



MOSFET IV: High V_{DS} (beyond pinch-off)

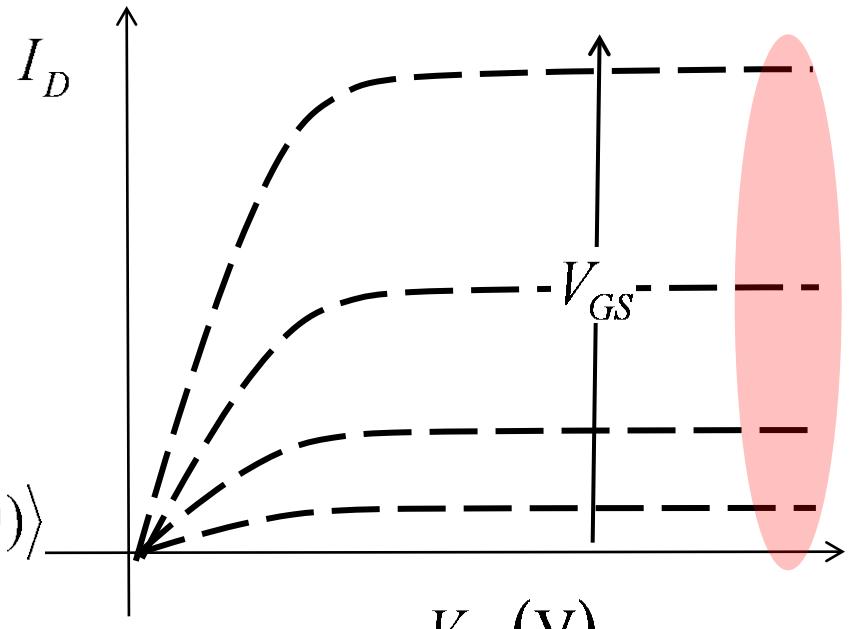


$$I_D = W |Q_n(x)| \langle v_x(x) \rangle = W |Q_n(0)| \langle v_x(0) \rangle$$

$$Q_n(0) = -C_{ox}(V_{GS} - V_T)$$

$$\langle v_x(0) \rangle = -\mu_n \mathcal{E}_x(0)$$

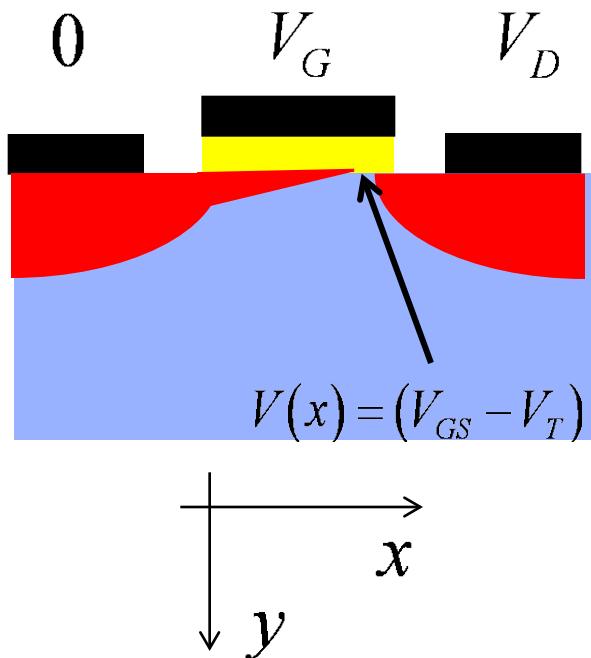
$$\mathcal{E}_x(0) \approx -V(x_{pinch})/L = -(V_{GS} - V_T)/L$$



$$I_D = \frac{W}{2L} \mu_n C_{ox} (V_{GS} - V_T)^2$$



Complete IV characteristic



$$V_{GS} > V_T$$

$$V_{DS} < V_{GS} - V_T$$

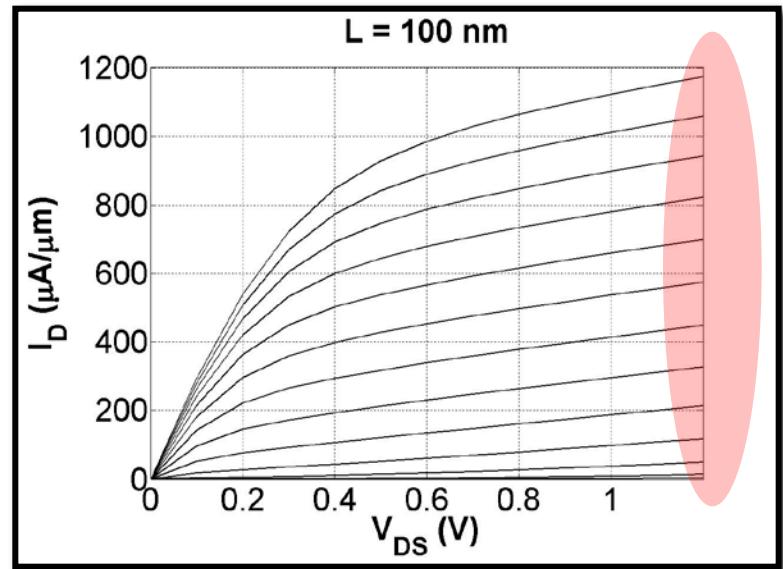
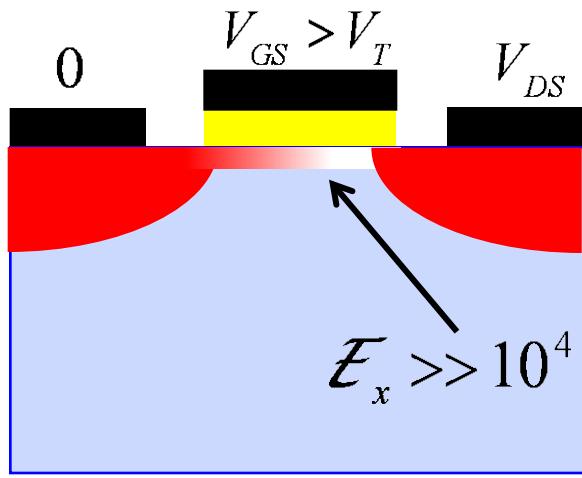
$$I_D = +\mu_n C_{ox} \frac{W}{L} \left[(V_{GS} - V_T) V_{DS} - \frac{V_{DS}^2}{2} \right]$$

$$V_{GS} > V_T$$

$$V_{DS} > V_{GS} - V_T$$

$$I_D = +\mu_n C_{ox} \frac{W}{2L} (V_{GS} - V_T)^2$$

MOSFET IV: High V_{DS} (velocity saturation)



$$I_D = W |Q_n(x)| \langle v_y(x) \rangle$$

(Courtesy, Shuji Ikeda, ATDF, Dec. 2007)

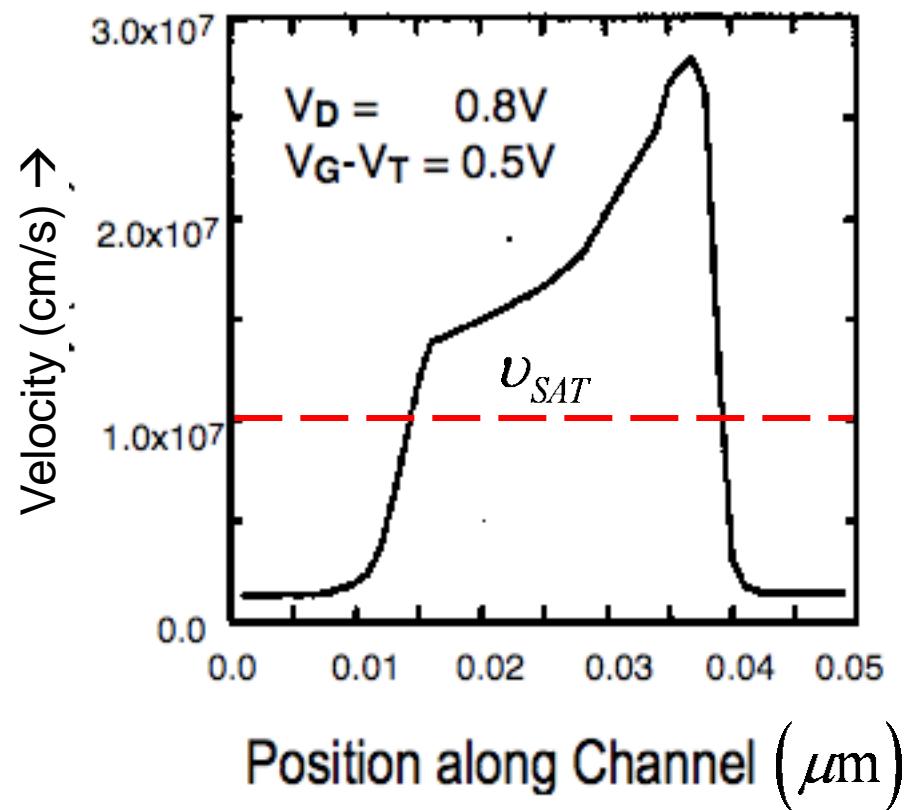
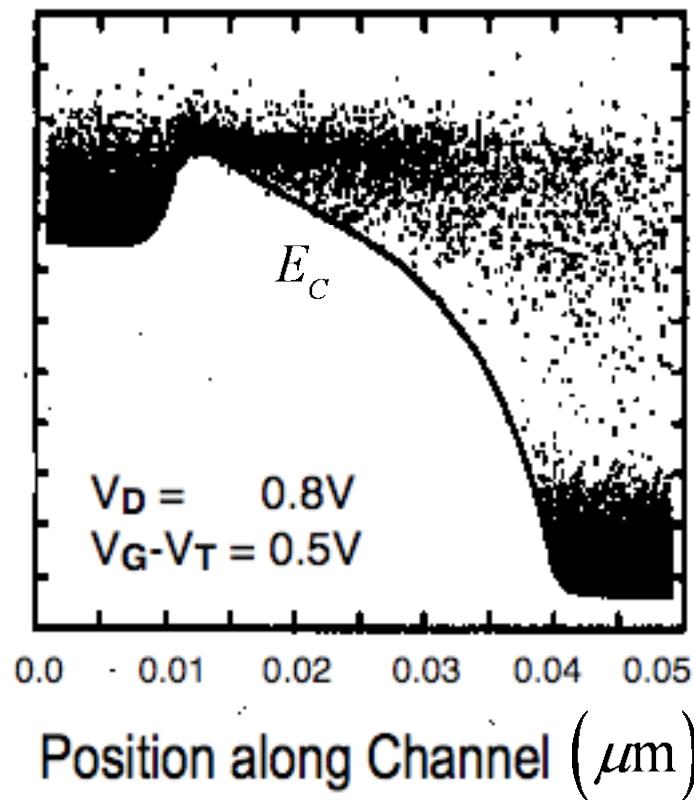
$$Q_n = -C_{ox}(V_{GS} - V_T)$$

$$\langle v_x \rangle = v_{sat}$$

$$I_D = W C_{ox} v_{sat} (V_{GS} - V_T)$$

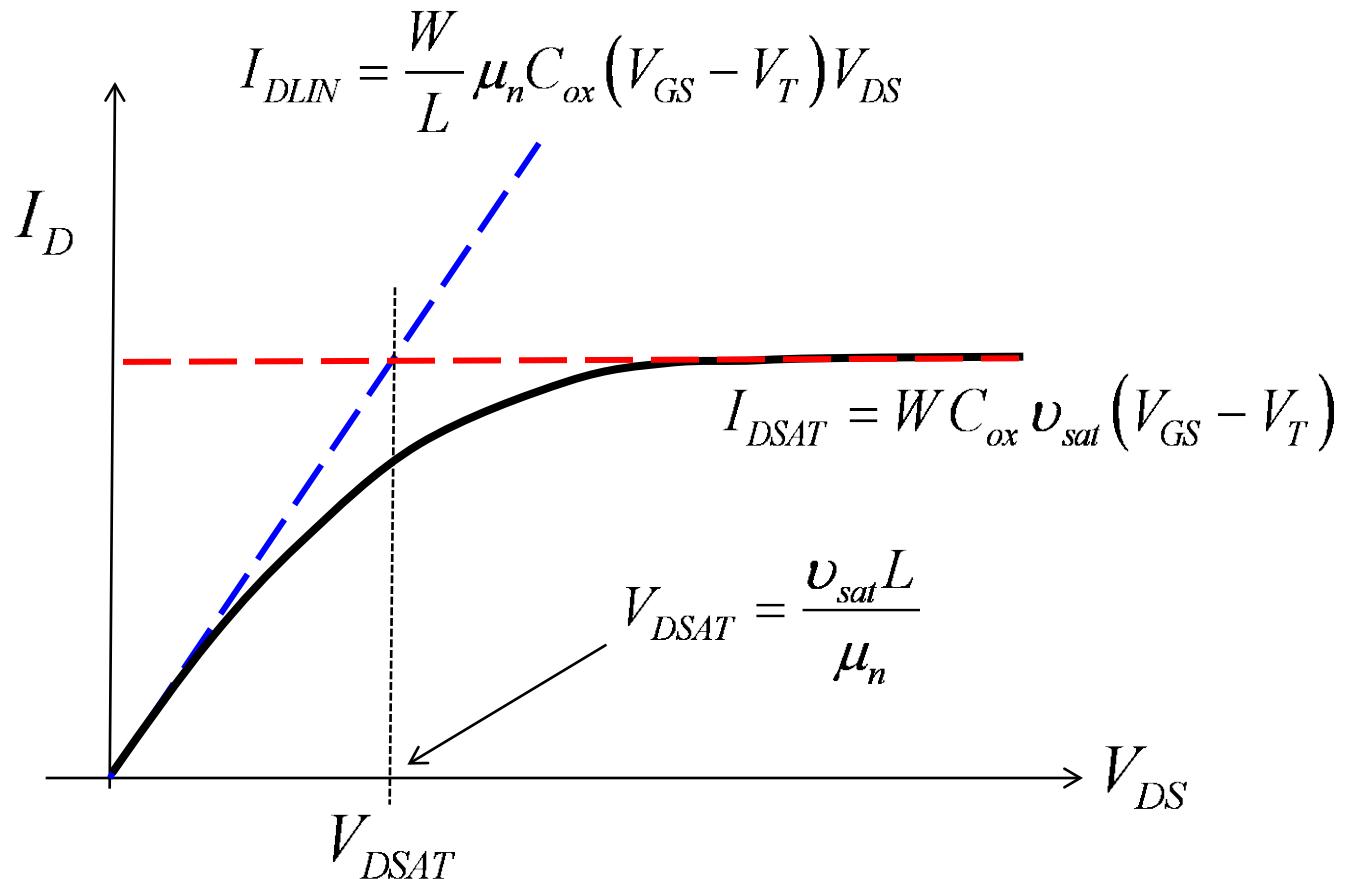


Velocity overshoot



D. Frank, S. Laux, and M. Fischetti, Int. Electron Dev. Mtg., Dec., 1992.

From two-piece to continuous model



We have developed a 2-piece approximation to the MOSFET IV characteristic.

Simple (Level 0) VS model

$$1) \quad I_D/W = |Q_n(V'_{GS})| \langle v(V'_{DS}) \rangle$$

$$2) \quad Q_n(V'_{GS}) = -C_{ox}(V'_{GS} - V_T) \quad (V'_{GS} > V_T)$$

$$V_T = V_{T0} - \delta V'_{DS}$$

$$3) \quad \langle v(V'_{DS}) \rangle = F_{SAT}(V'_{DS}) v_{sat}$$

$$4) \quad F_{SAT}(V'_{DS}) = \frac{V'_{DS}/V_{DSAT}}{\left[1 + (V'_{DS}/V_{DSAT})^\beta\right]^{1/\beta}}$$

$$5) \quad V_{DSAT} = \frac{v_{sat} L}{\mu_n}$$

There are only 8 device-specific parameters in this model:

$$C_{ox}, V_{T0}, \delta, v_{sat}, \mu_n, L$$

$$R_{SD} = R_S + R_D$$

$$+ \beta$$

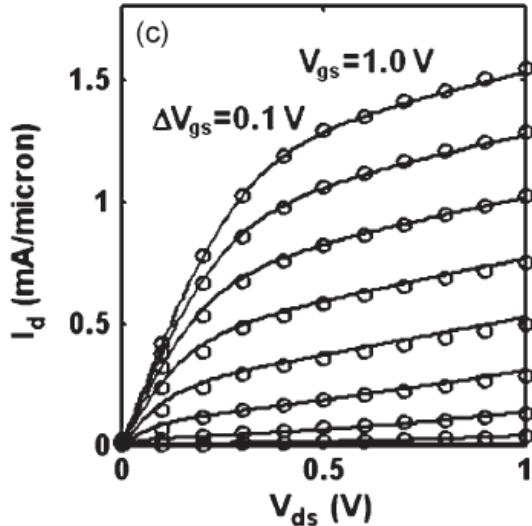
The MIT VS Model

1674

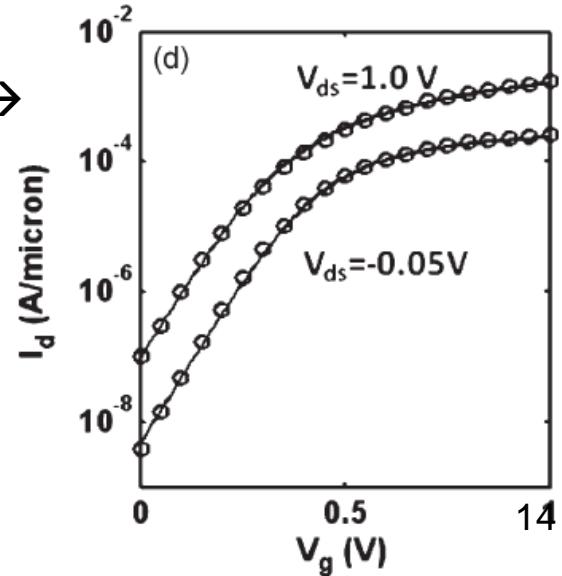
IEEE TRANSACTIONS ON ELECTRON DEVICES, VOL. 56, NO. 8, AUGUST 2009

A Simple Semiempirical Short-Channel MOSFET Current–Voltage Model Continuous Across All Regions of Operation and Employing Only Physical Parameters

Ali Khakifirooz, *Member, IEEE*, Osama M. Nayfeh, *Member, IEEE*, and Dimitri Antoniadis, *Fellow, IEEE*



← 32 nm technology →



Lundstrom: 2018

The MIT VS Model

In this course, we will show that the mobility and high-field saturation velocity should be re-interpreted:

$$\frac{1}{\mu_n} \rightarrow \frac{1}{\mu_{app}} \quad \text{"apparent mobility"}$$

$$v_{sat} \rightarrow v_{inj} \quad \text{"injection velocity"}$$

We will show that these two parameters have clear, well-defined physical interpretations.

Unit 3

Before we discuss carrier transport in nanoscale MOSFETs (e.g. mobility and saturation velocity), we will examine the important topic of **MOS electrostatics**.

MOS electrostatics describe the influence of the terminal voltages on the energy barrier between the source and drain.

Properly designed MOSFETs have “electrostatic integrity”:

$$Q_n(x=0) \approx -C_{ox}(V_{GS} - V_T) \quad \text{C/cm}^2$$

Essentials of MOSFETs

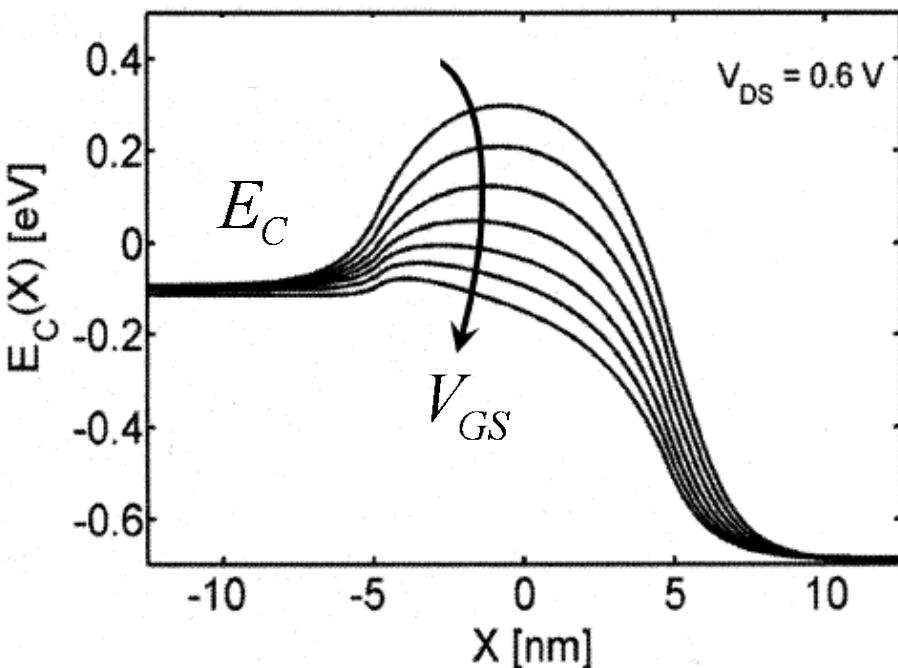
Unit 3: MOS Electrostatics

Lecture 3.1: Energy Band Diagram Approach

Mark Lundstrom

lundstro@purdue.edu
Electrical and Computer Engineering
Purdue University
West Lafayette, Indiana USA

Electrostatics



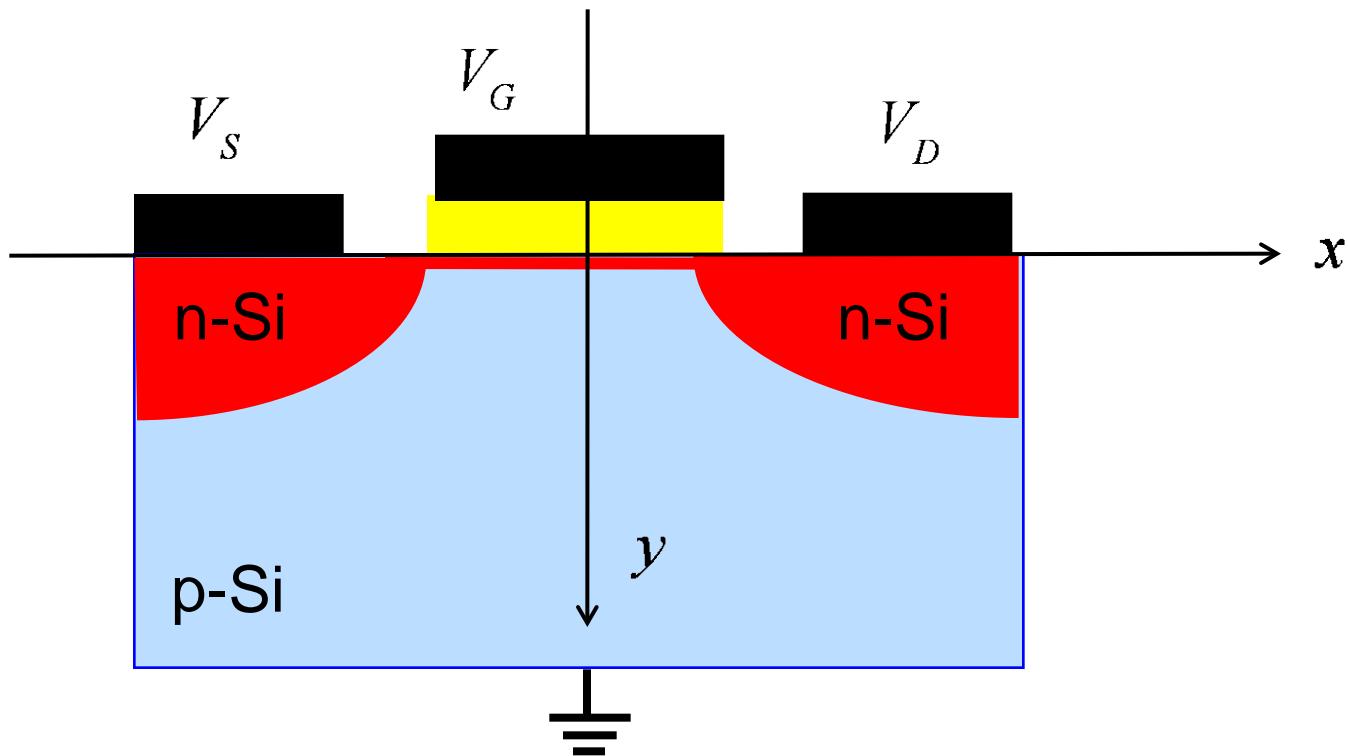
Unit 3: electrostatics

$$I_D/W = |Q_n(V_{GS}, V_{DS})| \langle v_x(V_{GS}, V_{DS}) \rangle$$

Unit 4: transport

Before developing analytical expressions, we should understand MOS electrostatics from an energy band perspective.

2D energy band diagrams



$$E_C(x, y) = E_{C0} - q\psi(x, y)$$

$$E_V(x, y) = E_{V0} - q\psi(x, y)$$

The potential, $\Psi(x, y)$, in the semiconductor is controlled by the voltages applied to the terminals.

Poisson equation

Goal: Find: $\psi(x, y)$

Solve the Poisson equation:

$$\nabla \cdot \vec{D}(x, y) = \rho(x, y)$$

$$\vec{D} = \epsilon_s \vec{\mathcal{E}} = \kappa_s \epsilon_0 \vec{\mathcal{E}}$$

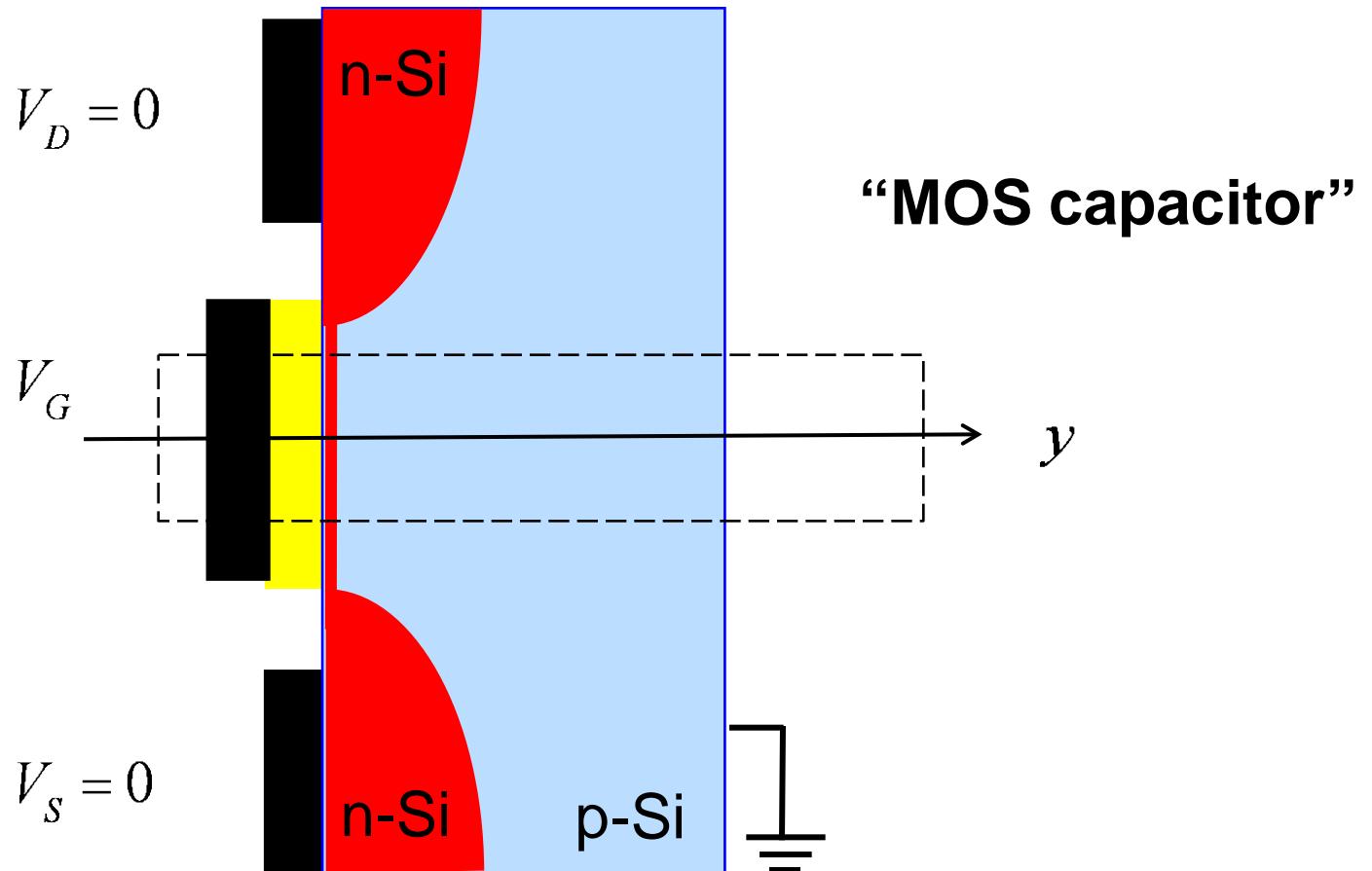
$$\vec{\mathcal{E}} = -\vec{\nabla} \psi$$

$$\nabla^2 \psi(x, y) = -\frac{\rho(x, y)}{\epsilon_s}$$

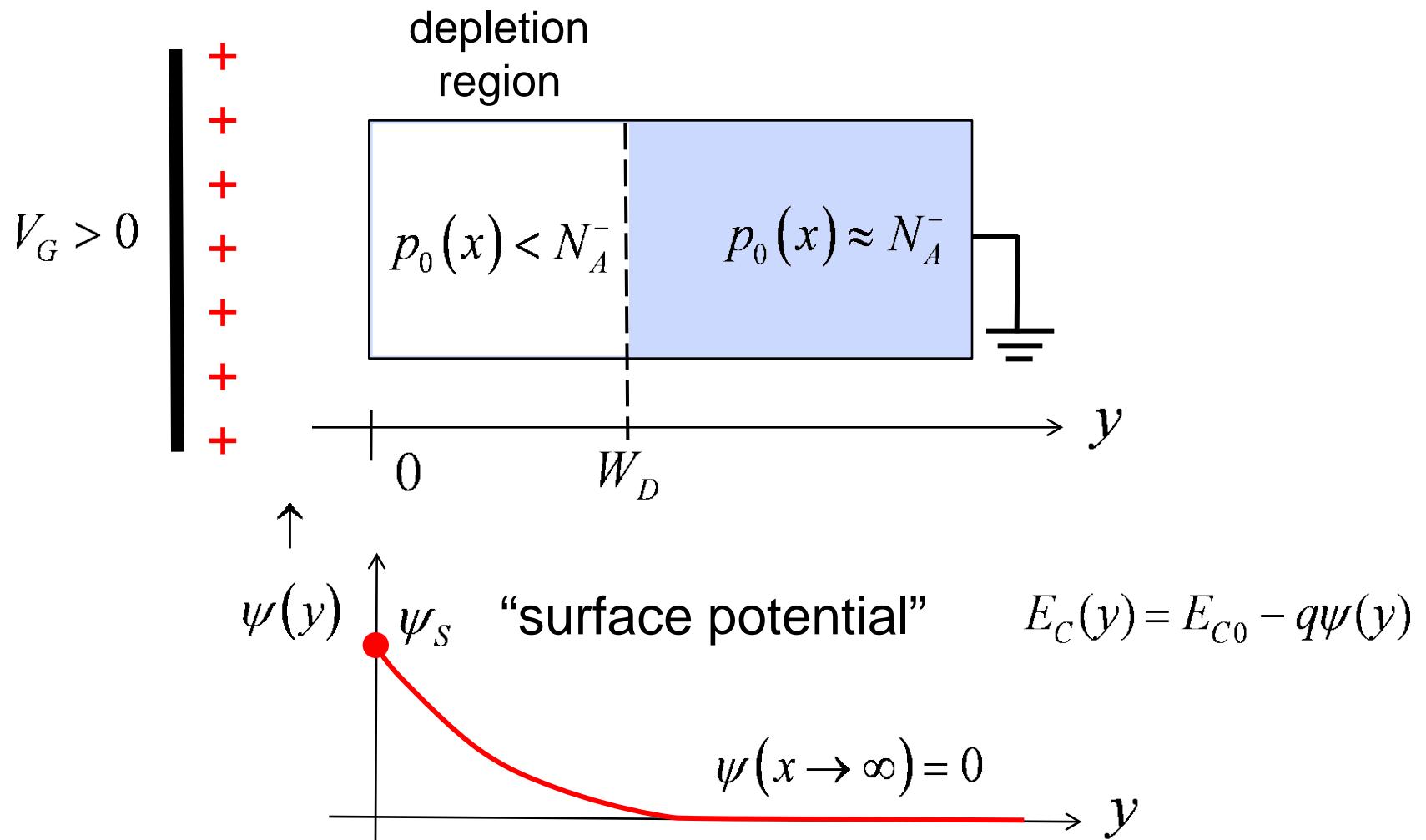
$$\rho(x, y) = q [p(x, y) - n(x, y) + N_D^+(x, y) - N_A^-(x, y)]$$

Drawing an energy band diagram provides us with a qualitative solution to the Poisson equation.

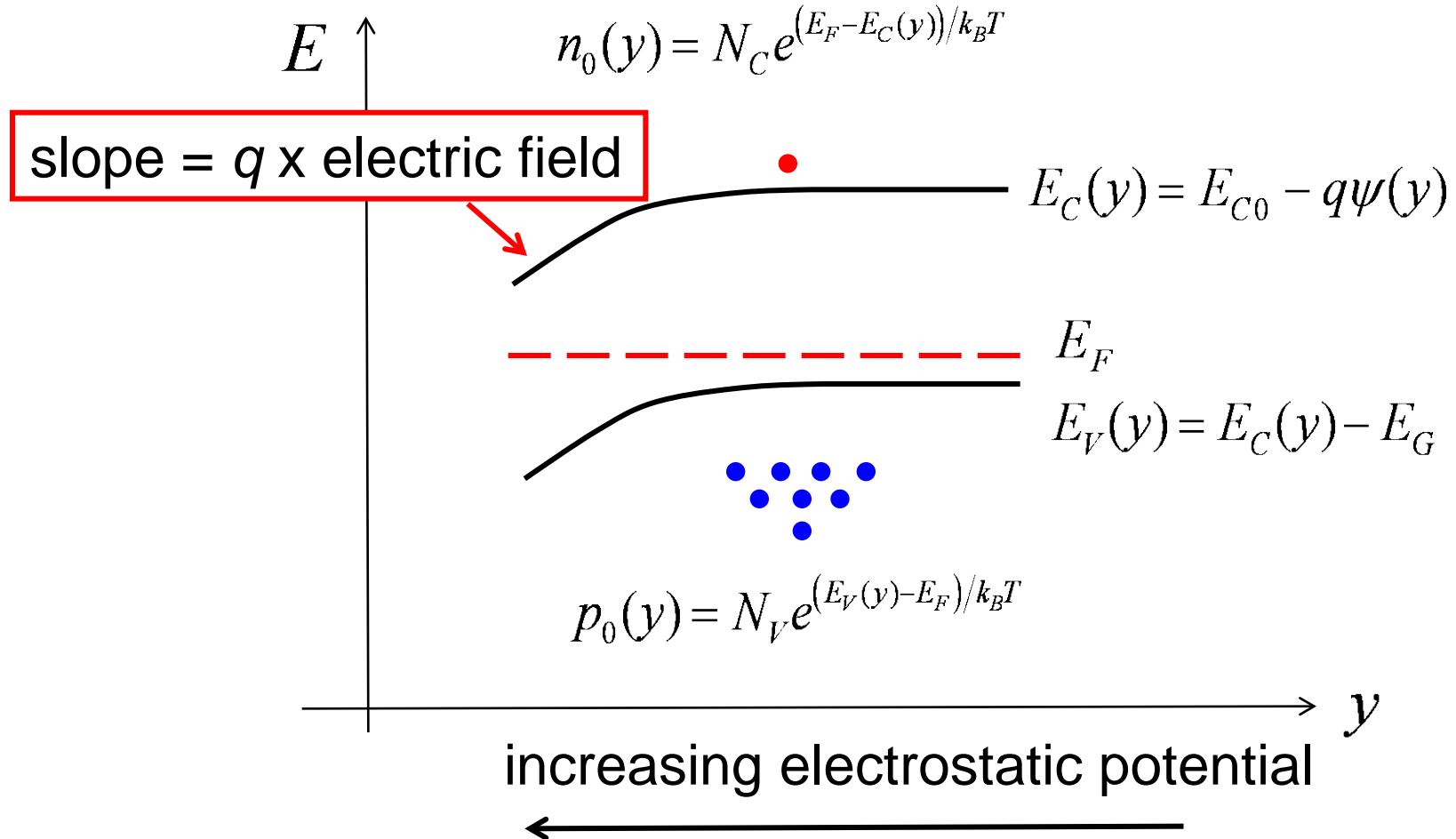
The 1D MOS Capacitor



Electrostatic potential vs. position



Band bending

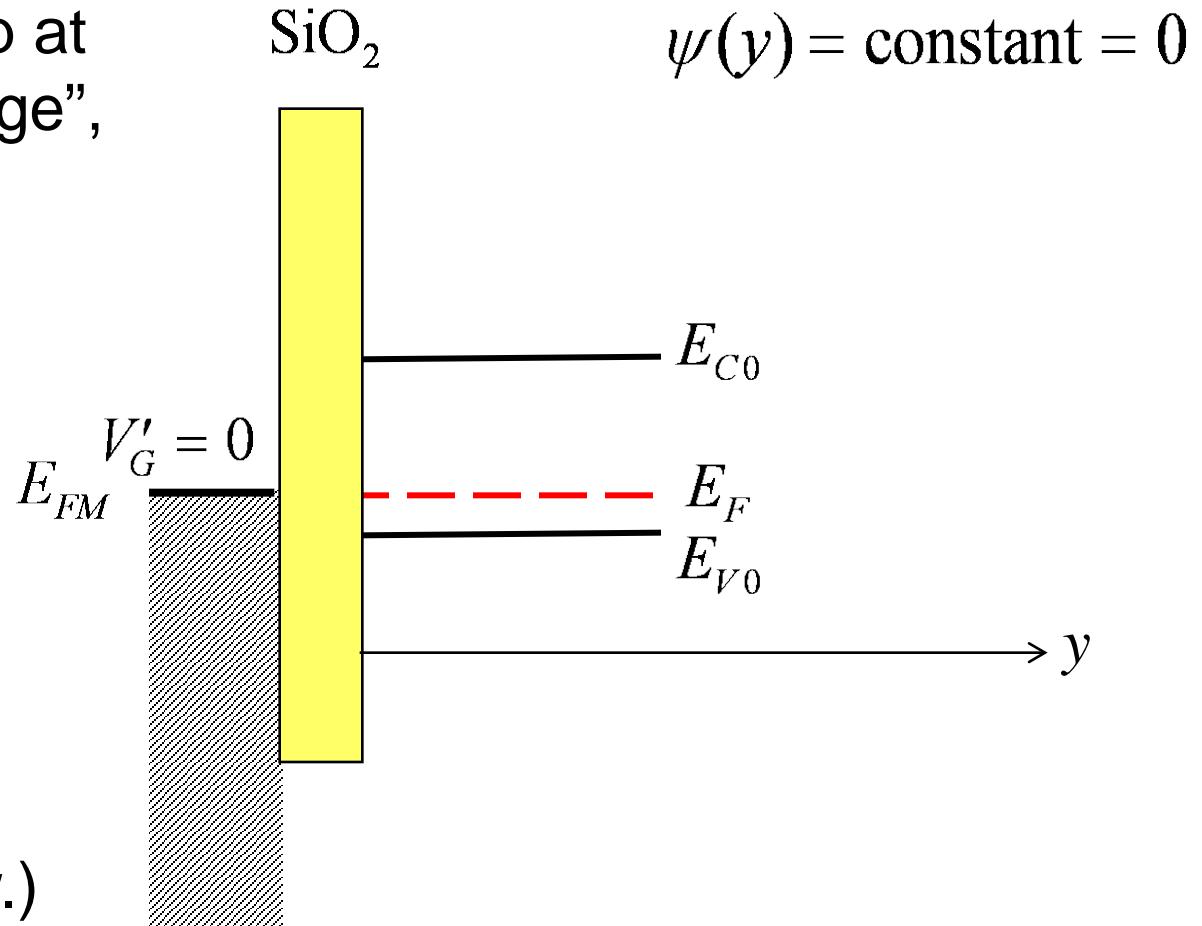


“Flat-band” conditions

Fermi levels line up at the “flat-band voltage”,

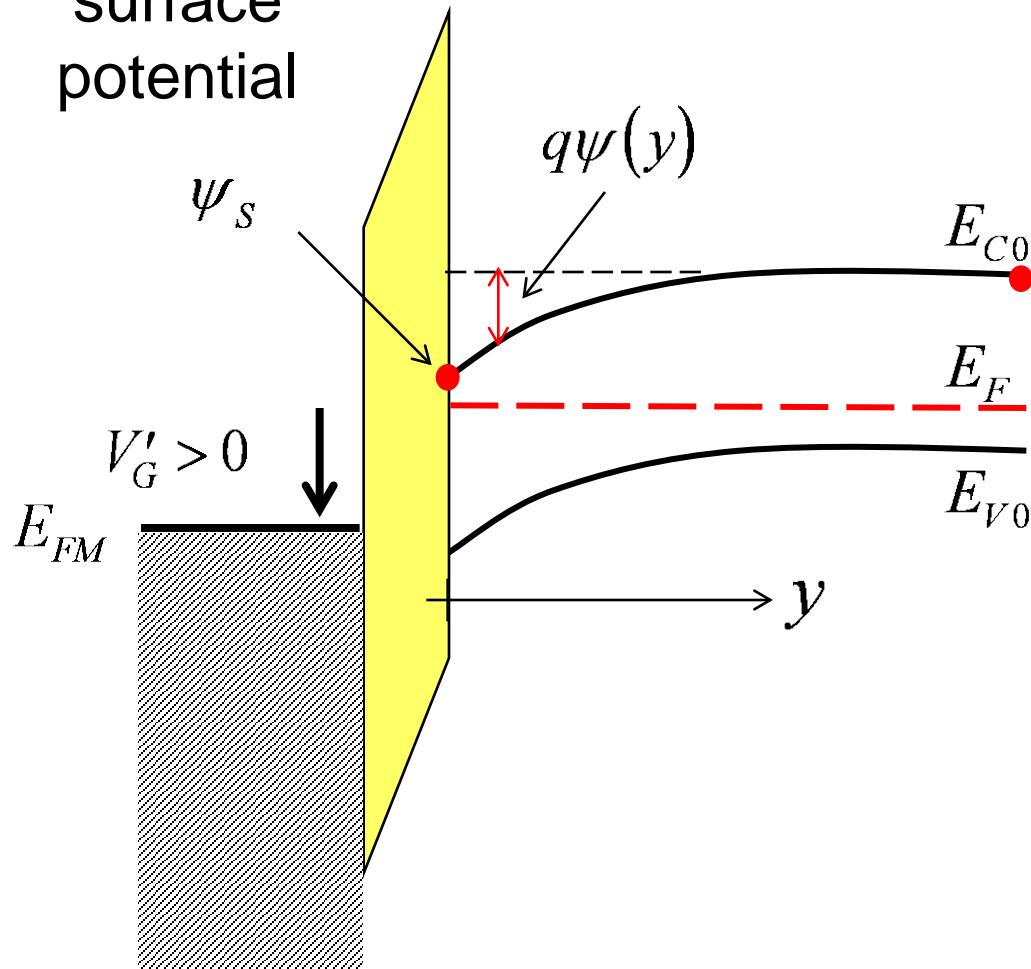
$$V'_G = 0$$

(Ignore metal-semi workfunction differences for now.)



Applied gate voltage

surface potential



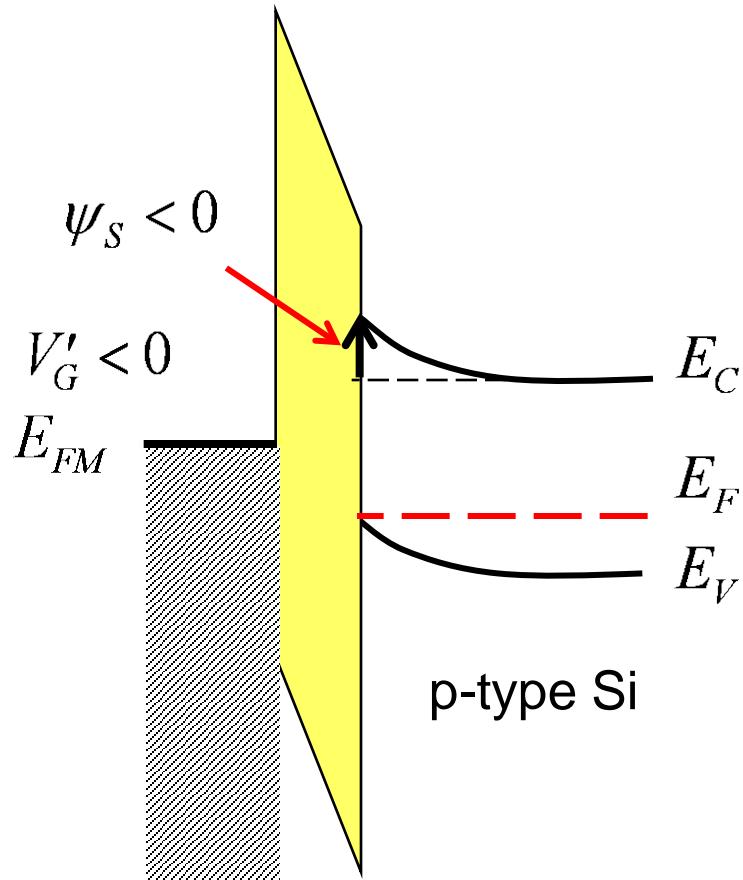
$$E_C(x \rightarrow \infty) = E_{C0}$$

$\psi(y) = 0$ (arbitrary reference for zero potential)

$$E_C(y) = E_{C0} - q\psi(y)$$

$$\psi(y) = \frac{E_{C0} - E_C(y)}{q}$$

$V'_G < 0$: “accumulation”



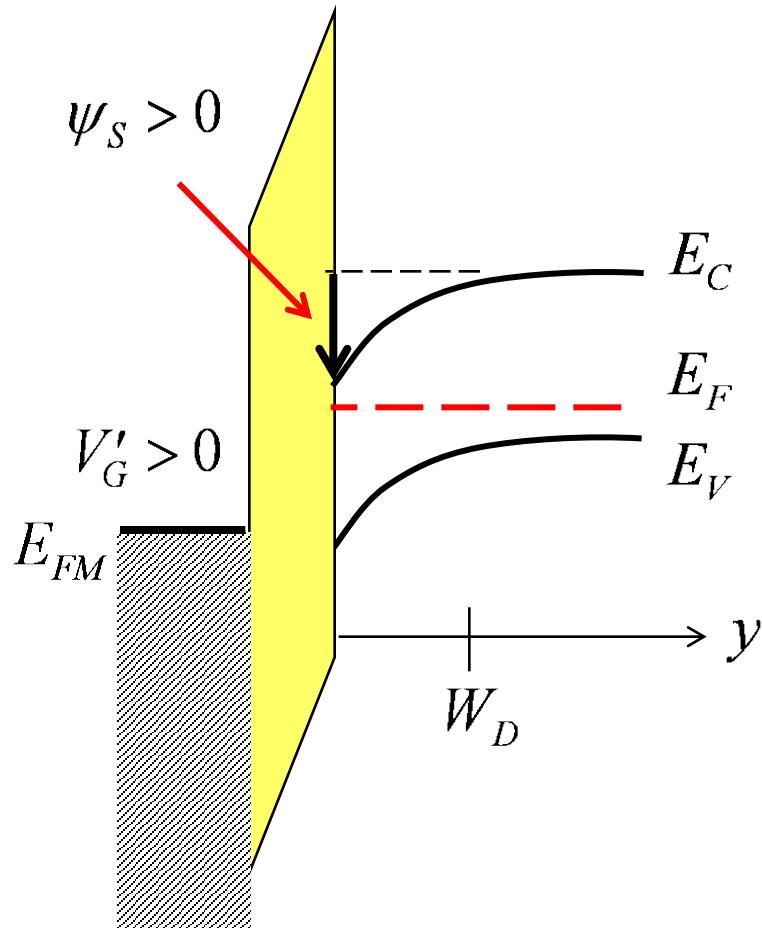
- surface potential < 0
- bands bend up
- hole density increases exponentially near the surface.

$$p_0(y) = N_V e^{(E_V(y) - E_F)/k_B T}$$

$$Q_S = +q \int_0^{\infty} (p_0(y) - N_A^-) dy \text{ C/cm}^2$$

(accumulation charge piles up very near the interface)

$V'_G > 0$: “depletion”



- surface potential > 0
- bands bend down
- space charge density $y < W_D$:

$$p_0(y) = N_V e^{(E_V(y) - E_F)/k_B T} \approx 0$$

$$n_0(y) = N_C e^{(E_F - E_C(y))/k_B T} \approx 0$$

$$\rho(y) \approx -qN_A^- \quad (y < W_D) \quad \text{C/cm}^3$$

“depletion charge”

$$\rho(y) \approx 0 \quad (y \geq W_D) \quad \text{C/cm}^3$$

$V'_G = V'_T$: onset of “inversion”

Electron concentration in the bulk:

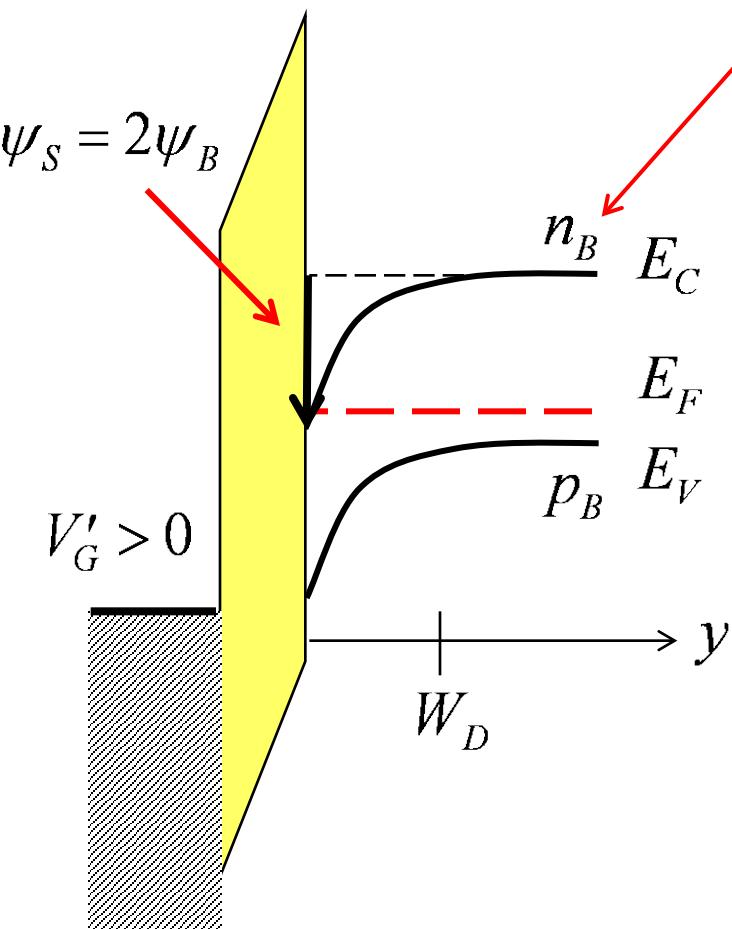
$$n_B = n_i^2 / N_A \ll p_B$$

Electron concentration at the surface:

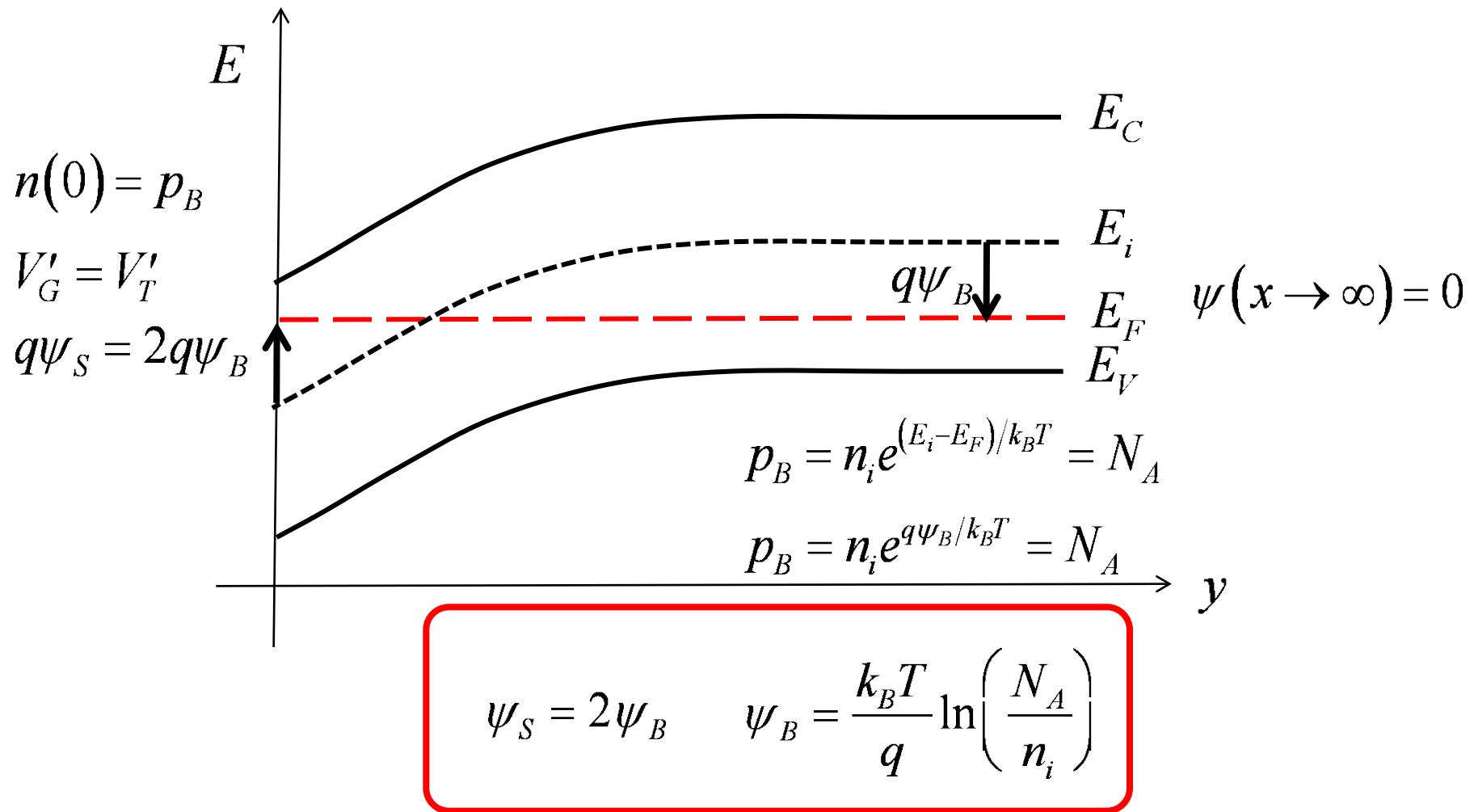
$$n_0(y=0) = N_C e^{(E_F - E_C(0))/k_B T} = n_B e^{q\psi_s/k_B T}$$

Band bending to make electron concentration at the surface = hole concentration in the bulk:

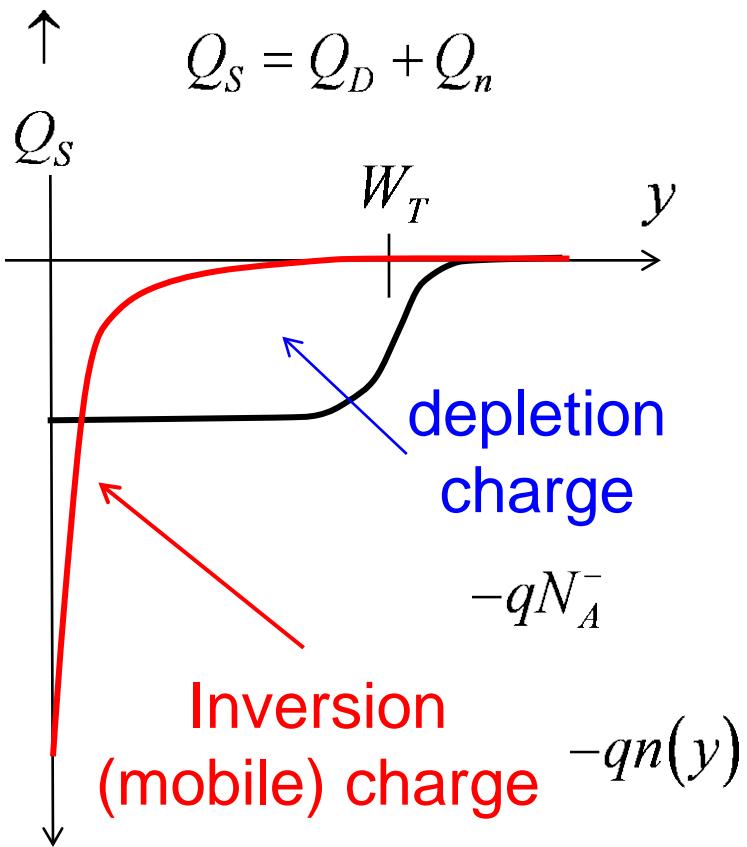
$$n_B e^{q\psi_s/k_B T} = N_A \quad \text{surface is “inverted”}$$



Onset of “inversion”



$V'_G > V_T$: “inversion”



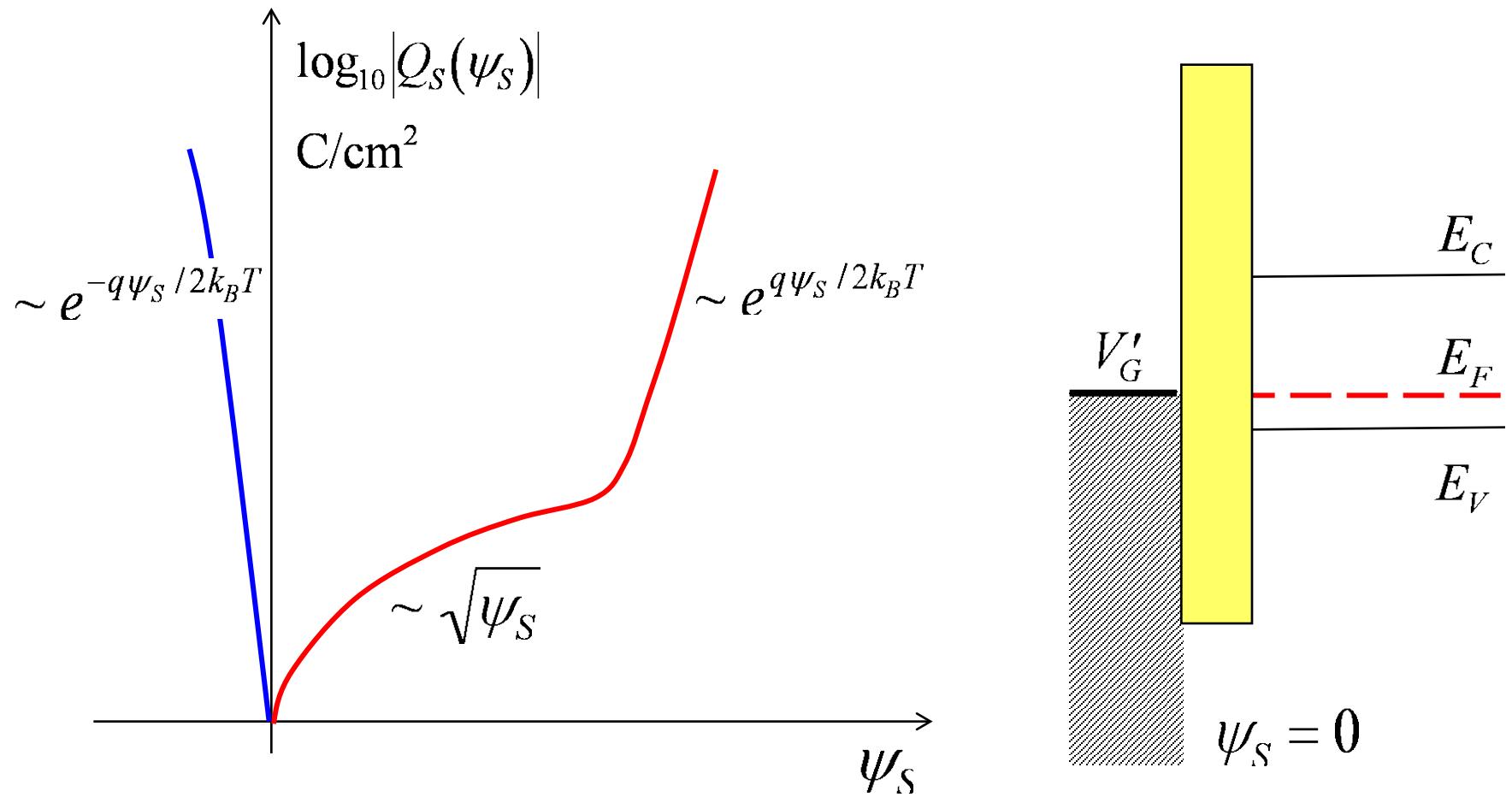
$$\psi_s \approx 2\psi_B \quad \psi_B = \frac{k_B T}{q} \ln\left(\frac{N_A}{n_i}\right)$$

Hard to bend the bands further.

$$W_T = \sqrt{2\epsilon_s(2\psi_B)/qN_A}$$

Electron charge piles up very near to the surface.

Total charge in semiconductor vs. surface potential

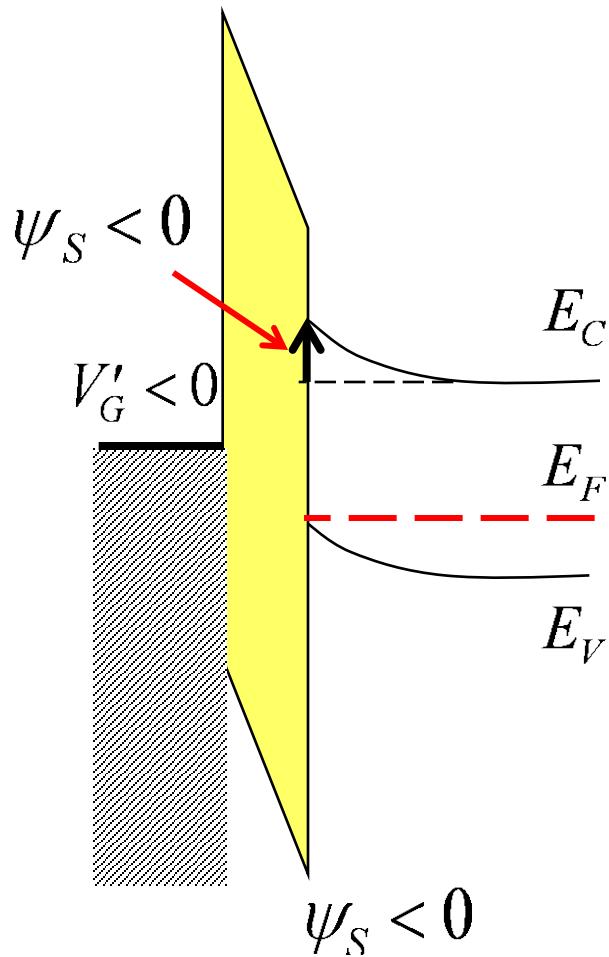


Exercise

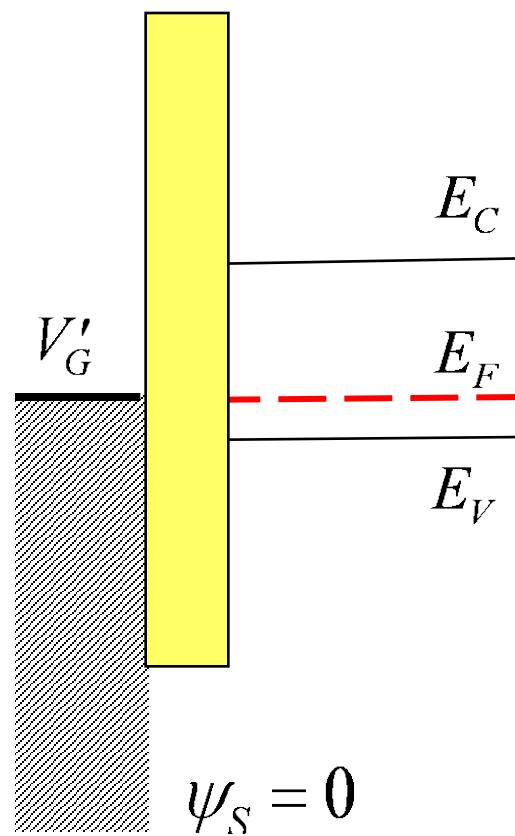
Re-do the previous two slides for an n-type semiconductor.

Summary

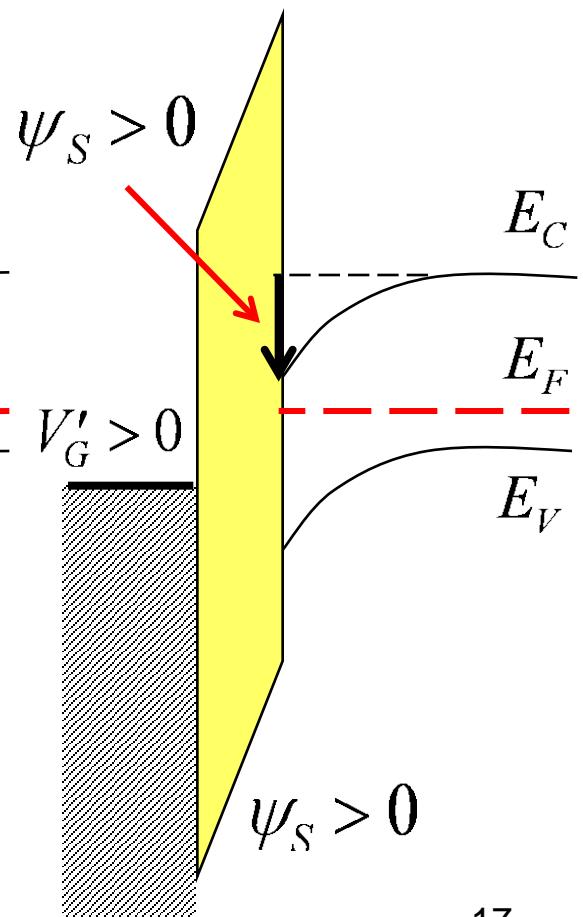
accumulation



flatband



depletion/
inversion



Next topic

Our goal is to solve the Poisson equation for $\psi(x, y)$.

In general, a numerical solution is required, but

In depletion, we can solve the problem analytically using the **depletion approximation**.

Essentials of MOSFETs

Unit 3: MOS Electrostatics

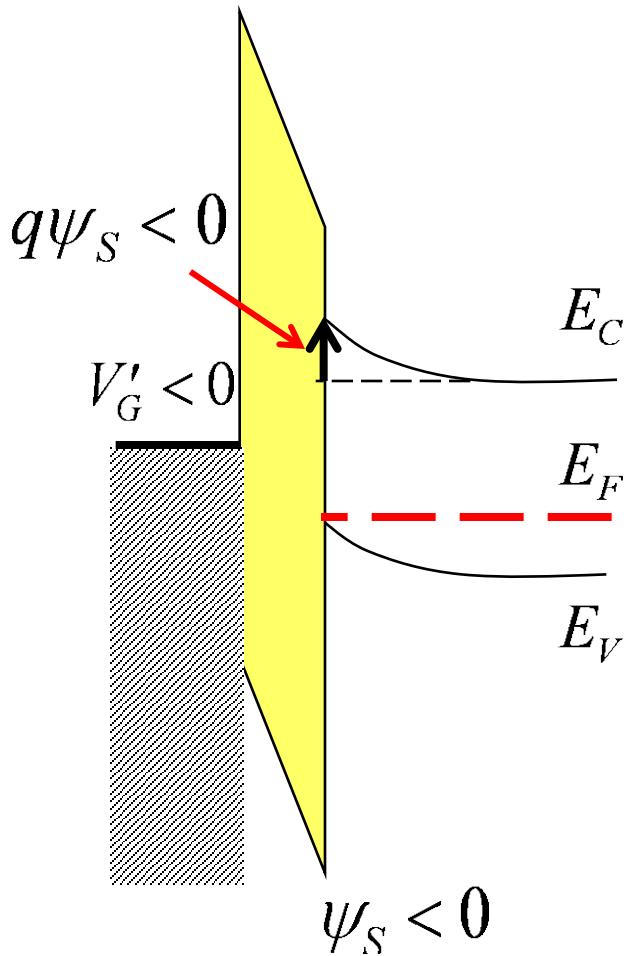
Lecture 3.2: The Depletion Approximation

Mark Lundstrom

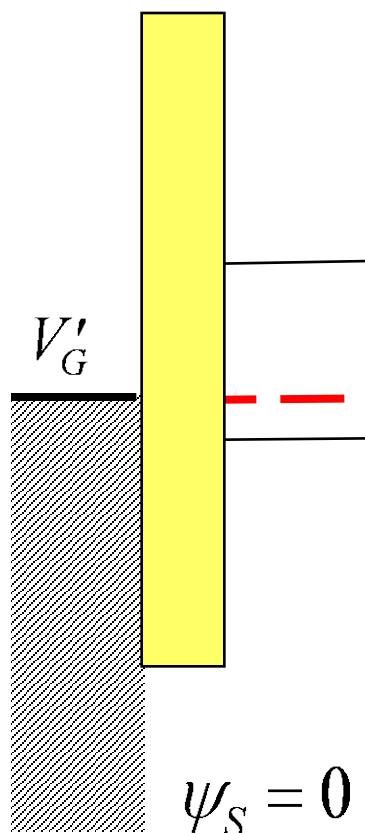
lundstro@purdue.edu
Electrical and Computer Engineering
Purdue University
West Lafayette, Indiana USA

1D MOS electrostatics

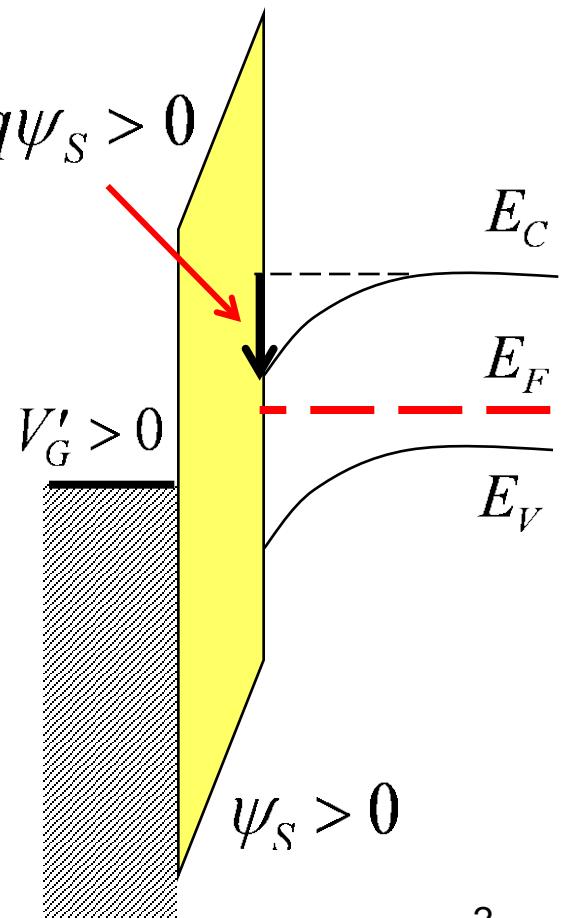
accumulation



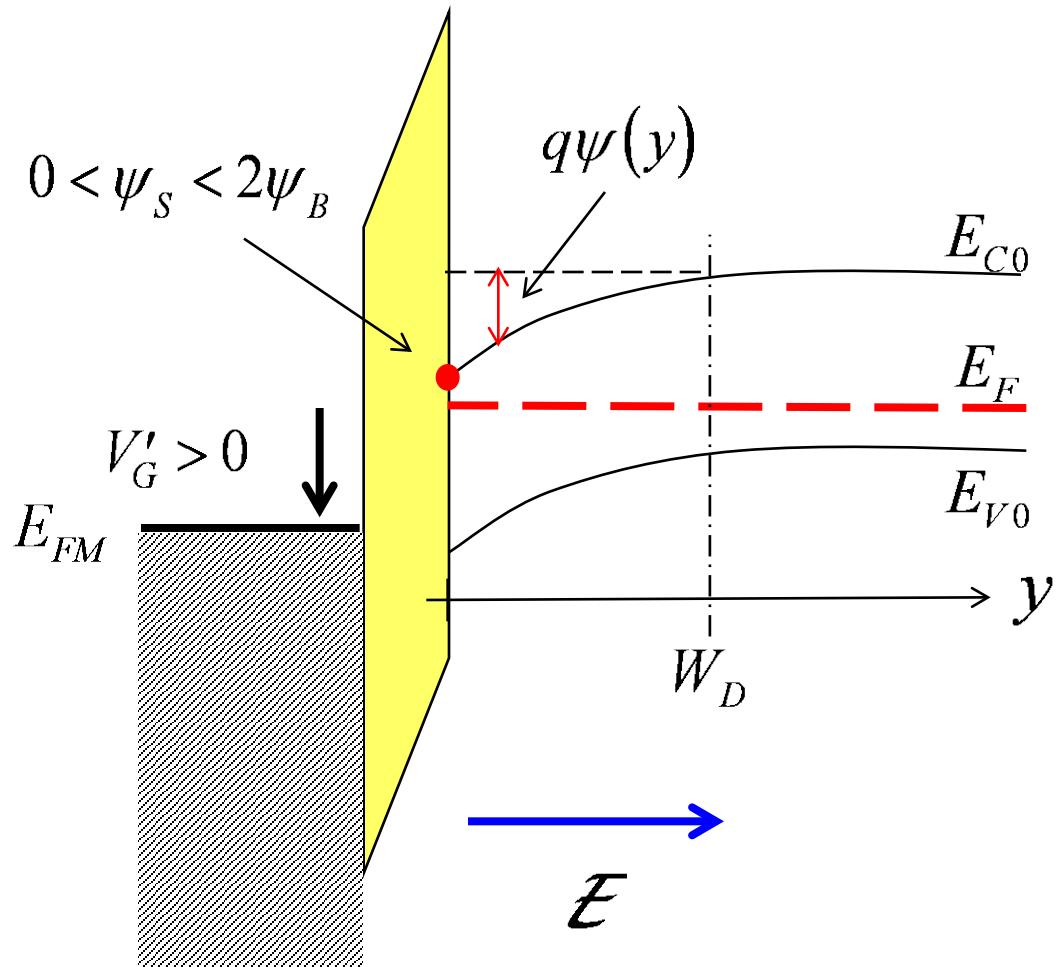
flat band



depletion/
inversion



Depletion



$$\rho(y) = q [p(y) - n(y) - N_A^-]$$

$y < W_D :$

$$p(y) \ll N_A^-$$

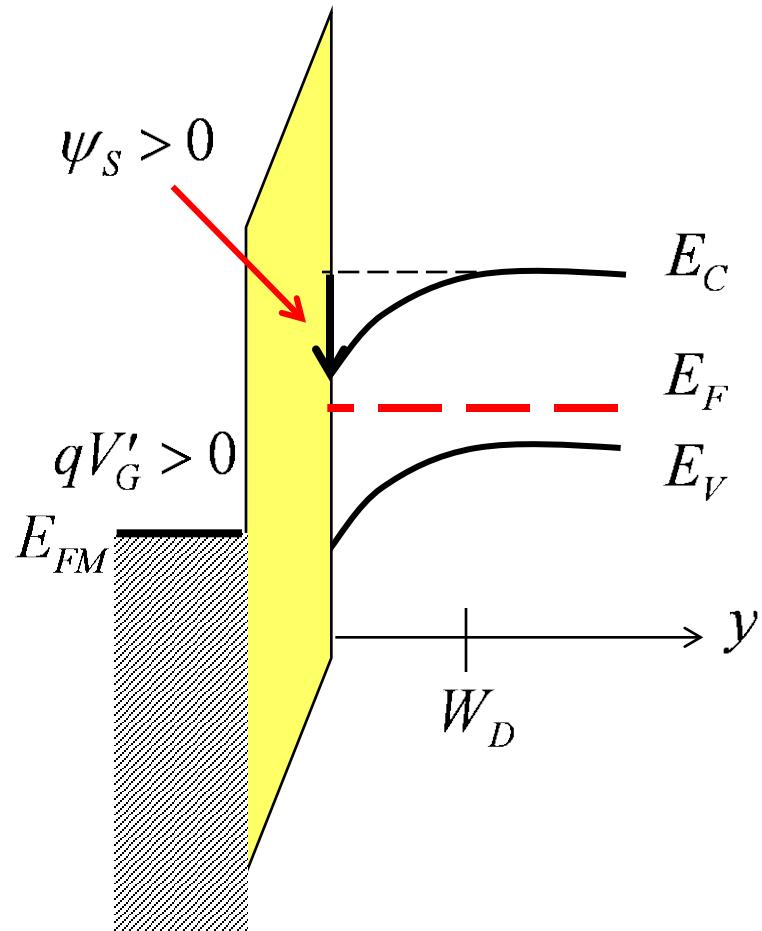
$$n(y) \ll N_A^-$$

$\rho(y) \approx -qN_A^-$

$y \geq W_D :$

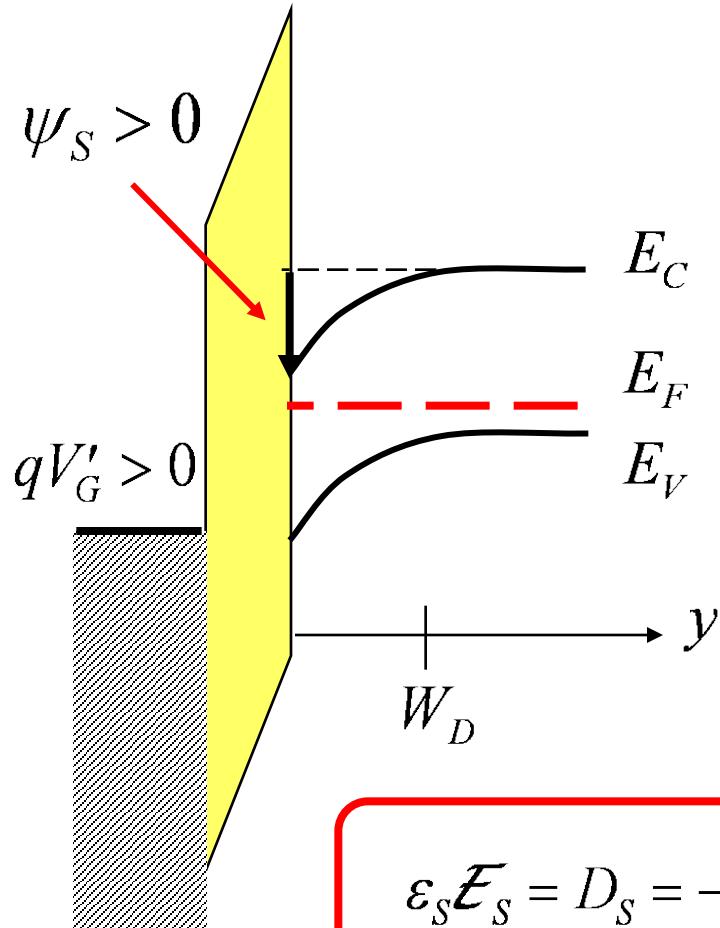
$$\rho(y) \approx 0$$

Poisson equation

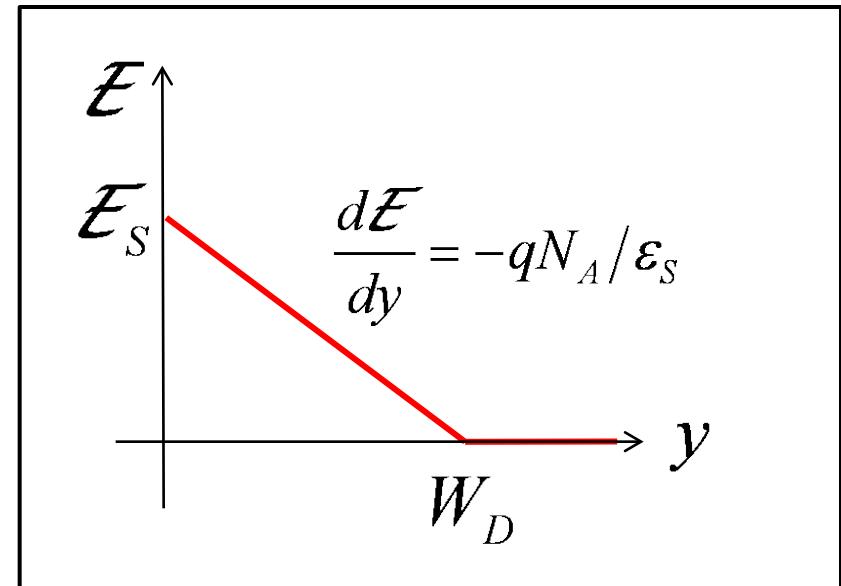


$$\left\{ \frac{dD(y)}{dy} = \rho(y) \right.$$
$$\left. \frac{d\mathcal{E}}{dy} = \frac{\rho(y)}{\epsilon_s} = -\frac{qN_A}{\epsilon_s} \right.$$
$$\epsilon_s = K_{Si}\epsilon_0$$

Electric field



$$\epsilon_s \mathcal{E}_s = D_s = -Q_s$$

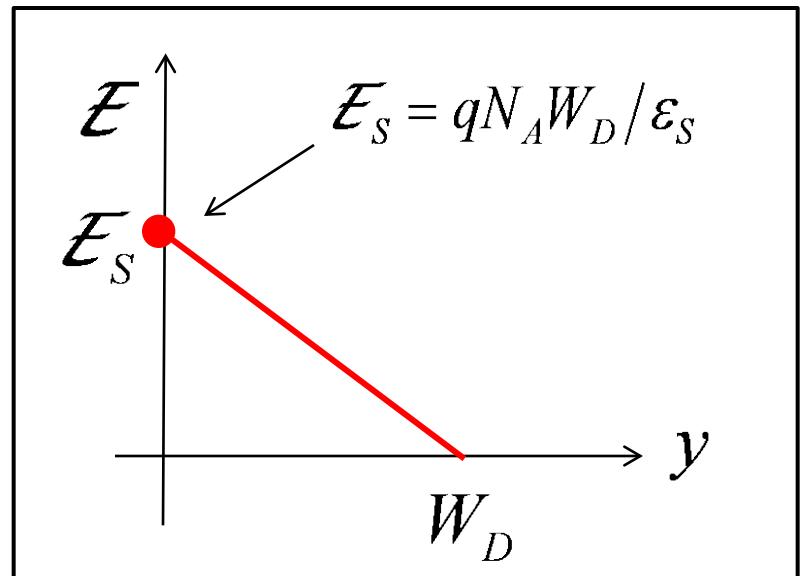
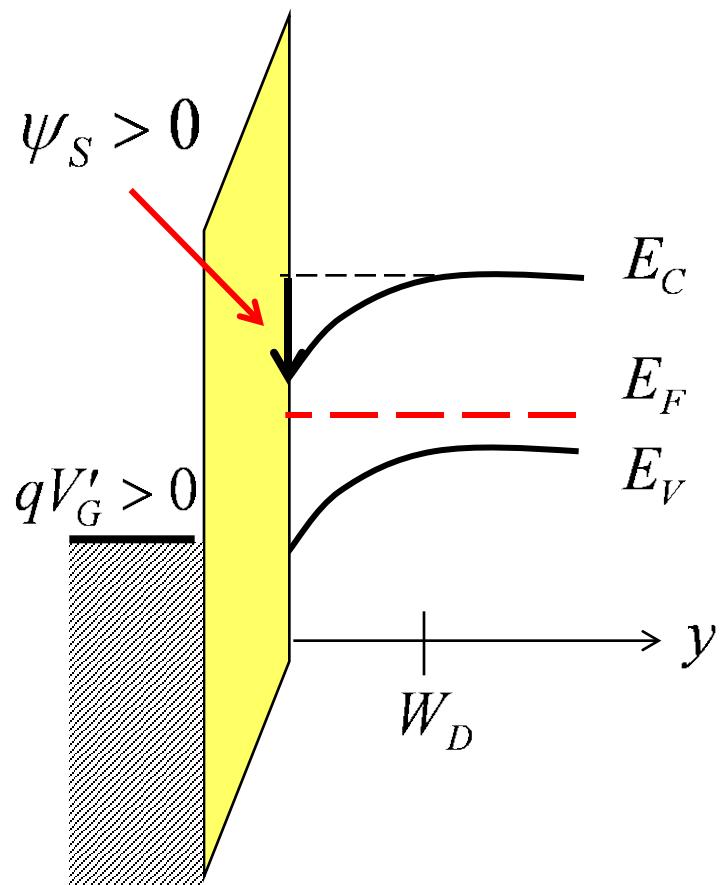


$$\mathcal{E}(y) = \frac{qN_A}{\epsilon_s} (W_D - y)$$

$$\mathcal{E}_s = \frac{qN_A W_D}{\epsilon_s} = \frac{-Q_s}{\epsilon_s}$$

Lundstrom: 2018

Electrostatic potential



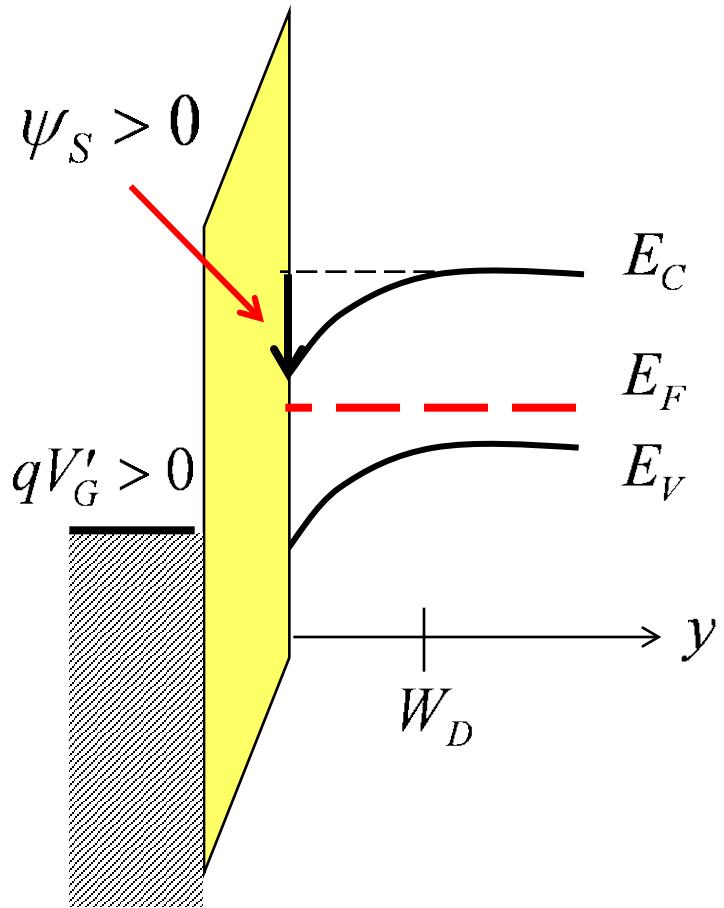
$$\mathcal{E}(y) = -d\psi(y)/dy \quad \psi(y) = -\int \mathcal{E}(y) dy$$

$$\psi_S = \frac{1}{2} \mathcal{E}_S W_D$$

Lundstrom: 2018

$$W_D = \sqrt{2\epsilon_s \psi_S / qN_A}$$

Depletion charge per cm²



$$W_D = \sqrt{2\epsilon_s \psi_s / qN_A}$$

$$Q_D = -qN_A W_D = -\sqrt{2qN_A \epsilon_s \psi_s} \text{ C/cm}^2$$

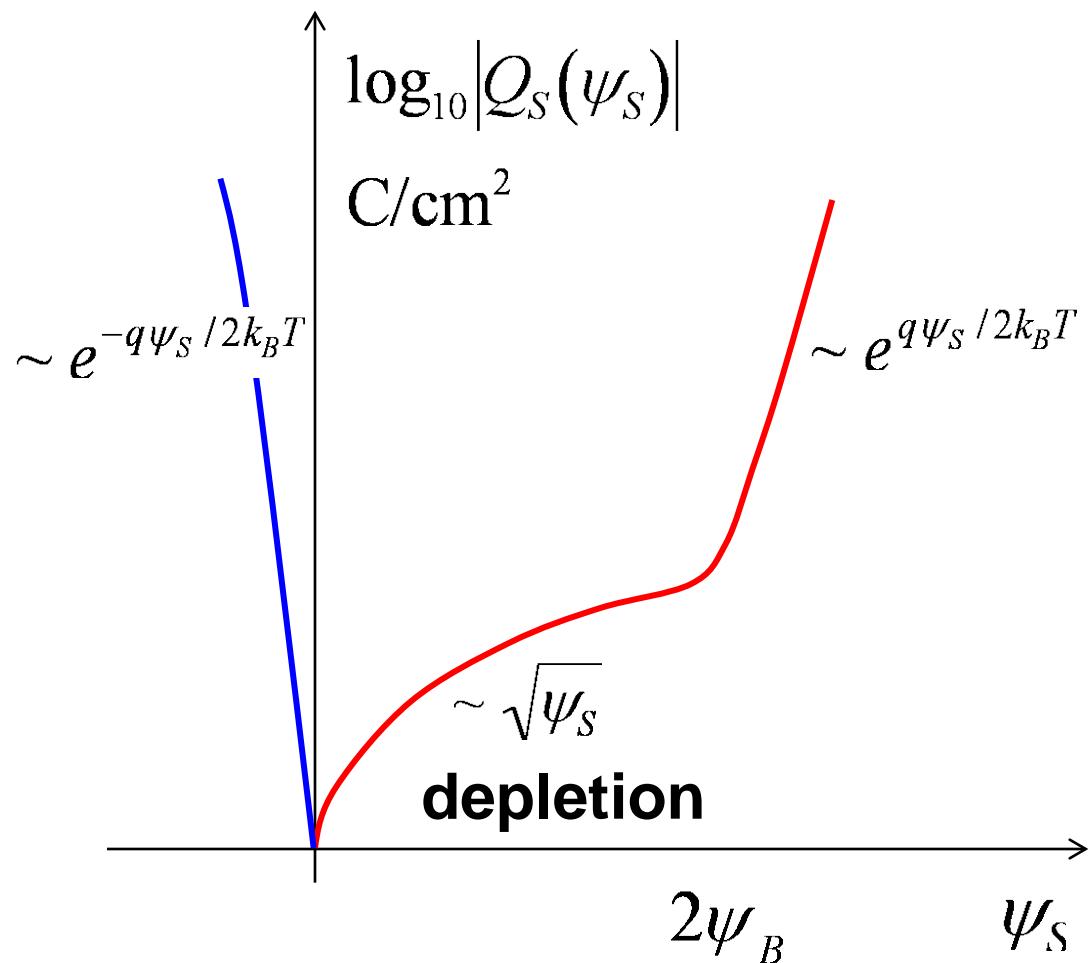
(depletion charge)

$$Q_S = Q_D + Q_n \text{ C/cm}^2$$

(total charge in semiconductor)

$$|Q_S| \approx |Q_D| \sim \sqrt{\psi_s}$$

MOS electrostatics



Example

P-type Si doped at: $N_A = 10^{18} \text{ cm}^{-3}$

$T = 300 \text{ K}$ $n_i(300 \text{ K}) = 10^{10} \text{ cm}^{-3}$ $\kappa_{Si} = 11.8$

$\psi_s = 0.5 \text{ V}$ $k_B T / q = 0.026 \text{ V}$

Find:

- i) the width of the depletion layer
- ii) the electric field at the surface

Example

P-type Si doped at: $N_A = 10^{18} \text{ cm}^{-3}$

$$T = 300 \text{ K} \quad n_i(300 \text{ K}) = 10^{10} \text{ cm}^{-3}$$

$$\psi_S = 0.5 \text{ V}$$

1) Check to see if we are in depletion or inversion.

$$\psi_S < 2\psi_B ?$$

$$\psi_B = \frac{k_B T}{q} \ln\left(\frac{N_A}{n_i}\right) = 0.026 \ln\left(\frac{10^{18}}{10^{10}}\right) = 0.48 \text{ V}$$

Lundstrom: 2018

$$\psi_S < 2\psi_B ?$$

$$0.5 < 0.96 \text{ V}$$

depletion ✓

Depletion layer thickness

P-type Si doped at: $N_A = 10^{18} \text{ cm}^{-3}$

$$T = 300 \text{ K} \quad n_i(300 \text{ K}) = 10^{10} \text{ cm}^{-3} \quad \kappa_{Si} = 11.8$$

$$\psi_s = 0.5 \text{ V} \quad k_B T / q = 0.026 \text{ V}$$

$$W_D = \sqrt{2\epsilon_s \psi_s / q N_A}$$

$$W_D = \sqrt{2(11.8)(8.854 \times 10^{-12})(0.5) / [(1.6 \times 10^{-19})10^{24}]}$$

$$W_D = 25.6 \text{ nm}$$

$$W_D = 25.6 \times 10^{-9} \text{ m}$$

Electric field at the surface

P-type Si doped at: $N_A = 10^{18} \text{ cm}^{-3}$

$$T = 300 \text{ K} \quad n_i(300 \text{ K}) = 10^{10} \text{ cm}^{-3} \quad \kappa_{Si} = 11.8$$

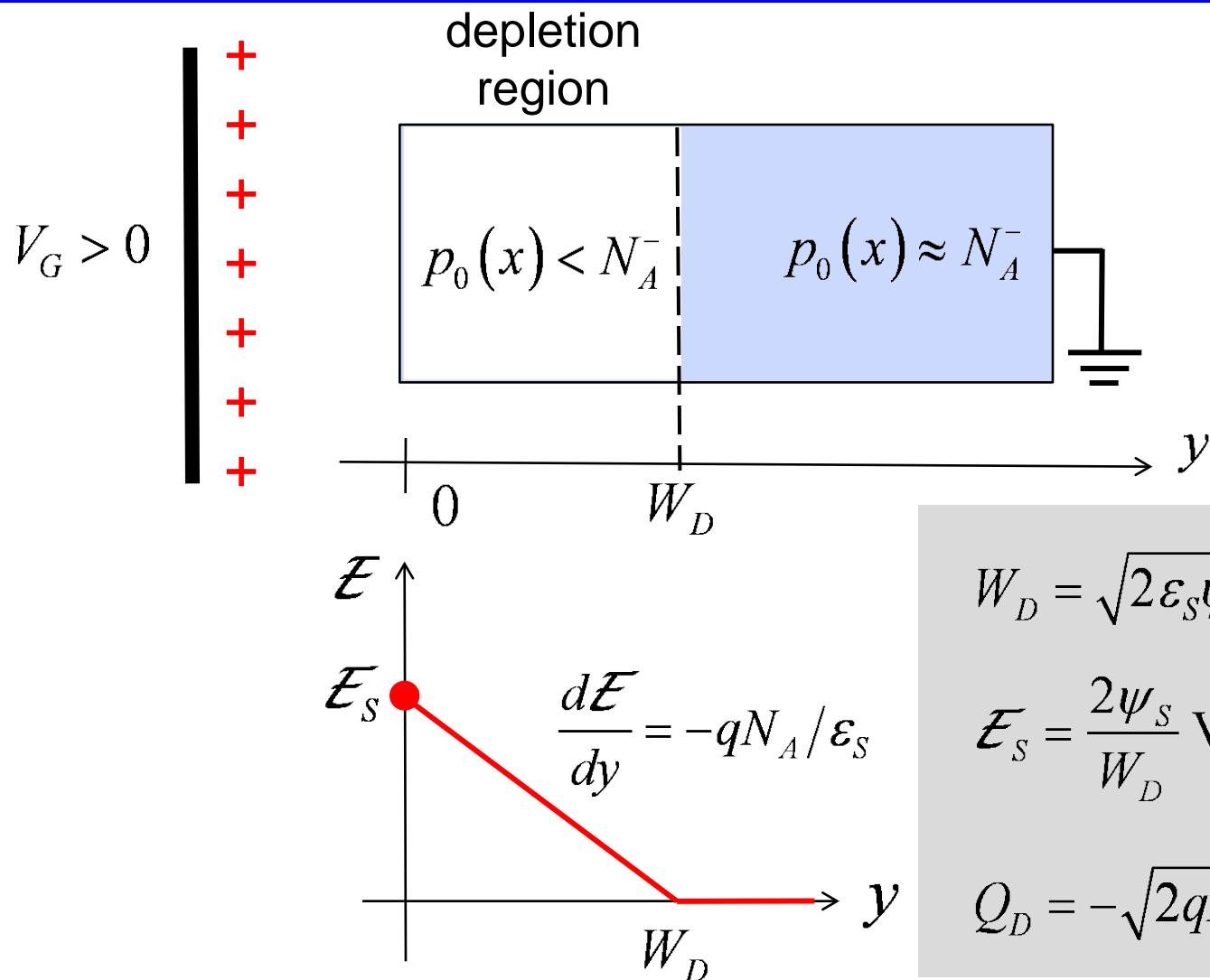
$$\psi_s = 0.5 \text{ V} \quad k_B T / q = 0.026 \text{ V}$$

$$\psi_s = \frac{1}{2} \mathcal{E}_s W_D \quad \mathcal{E}_s = \frac{2\psi_s}{W_D} = \frac{2(0.5)}{25.6 \times 10^{-9}} \quad \mathcal{E}_s = 3.9 \times 10^5 \frac{\text{V}}{\text{cm}}$$

$$\mathcal{E}_s = \frac{2\psi_s}{W_D} \quad \mathcal{E}_s = 3.9 \times 10^7 \frac{\text{V}}{\text{m}}$$

$$\mathcal{E}_s = 390 \frac{\text{kV}}{\text{cm}}$$

Summary



Next topic

Given a surface potential, we can compute the electric field and depletion layer thickness (if we are in depletion), but what gate voltage produced this surface potential?

That is the subject of the next lecture.

Essentials of MOSFETs

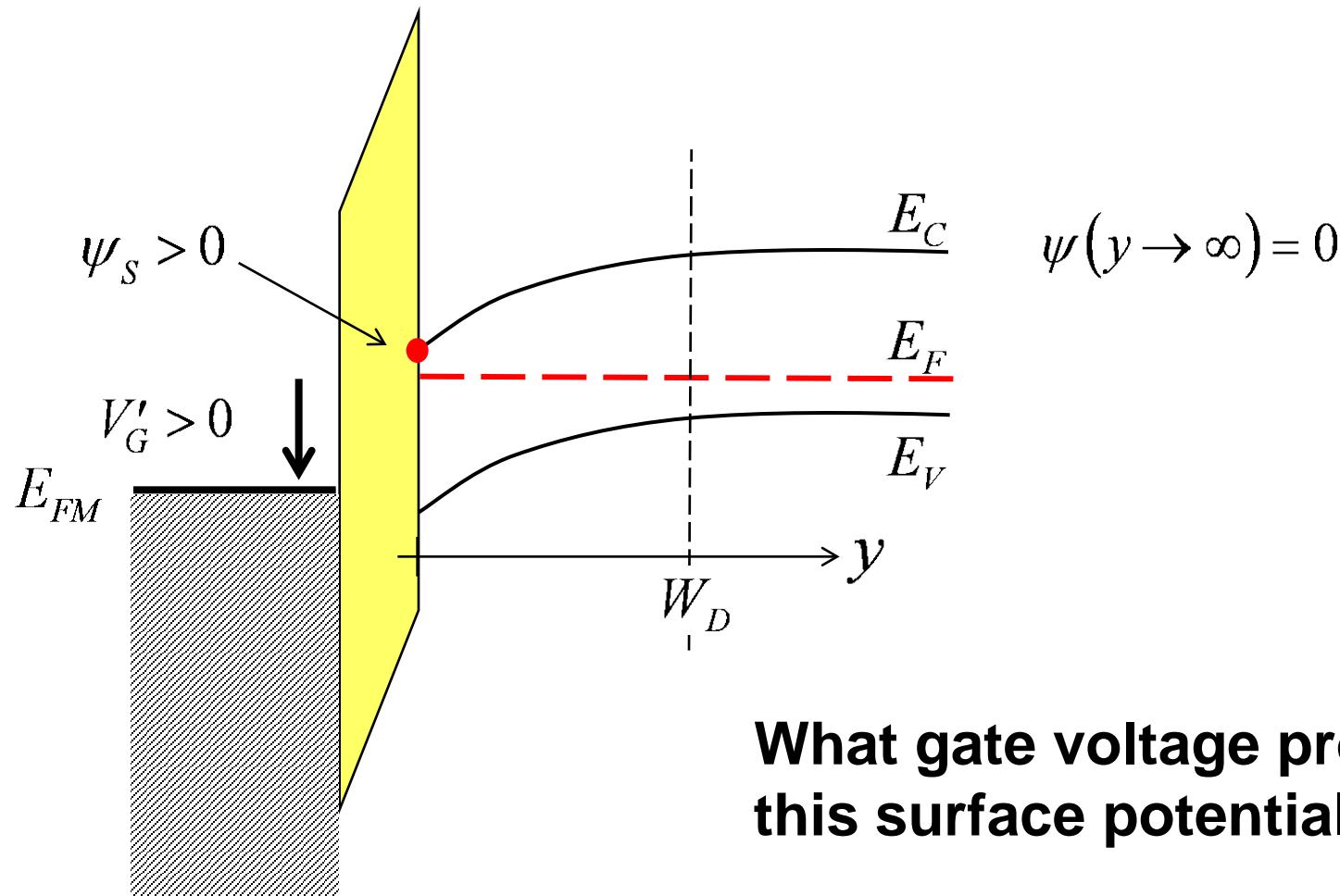
Unit 3: MOS Electrostatics

Lecture 3.3: Gate Voltage and Surface Potential

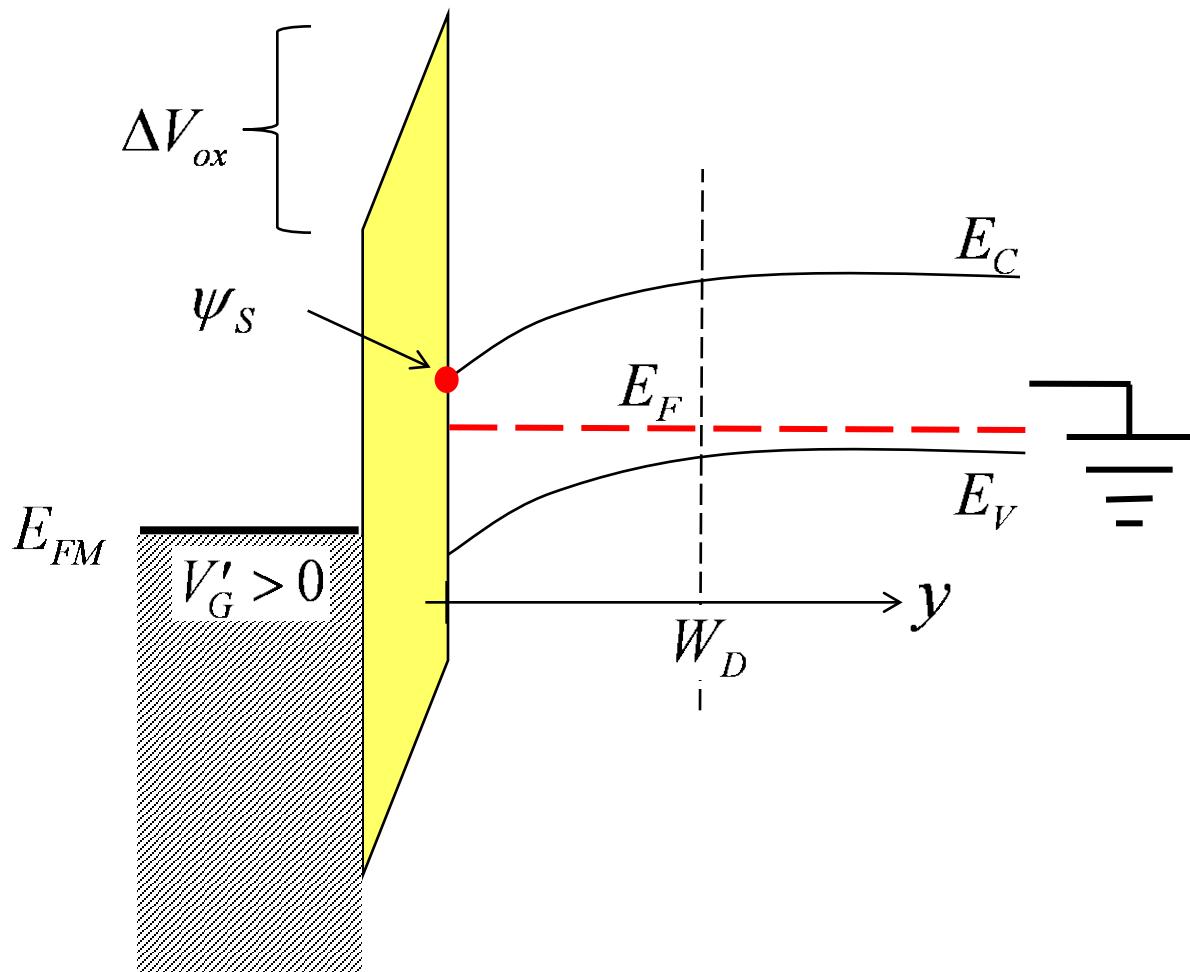
Mark Lundstrom

lundstro@purdue.edu
Electrical and Computer Engineering
Purdue University
West Lafayette, Indiana USA

Band bending depends on surface potential



Gate voltage and surface potential



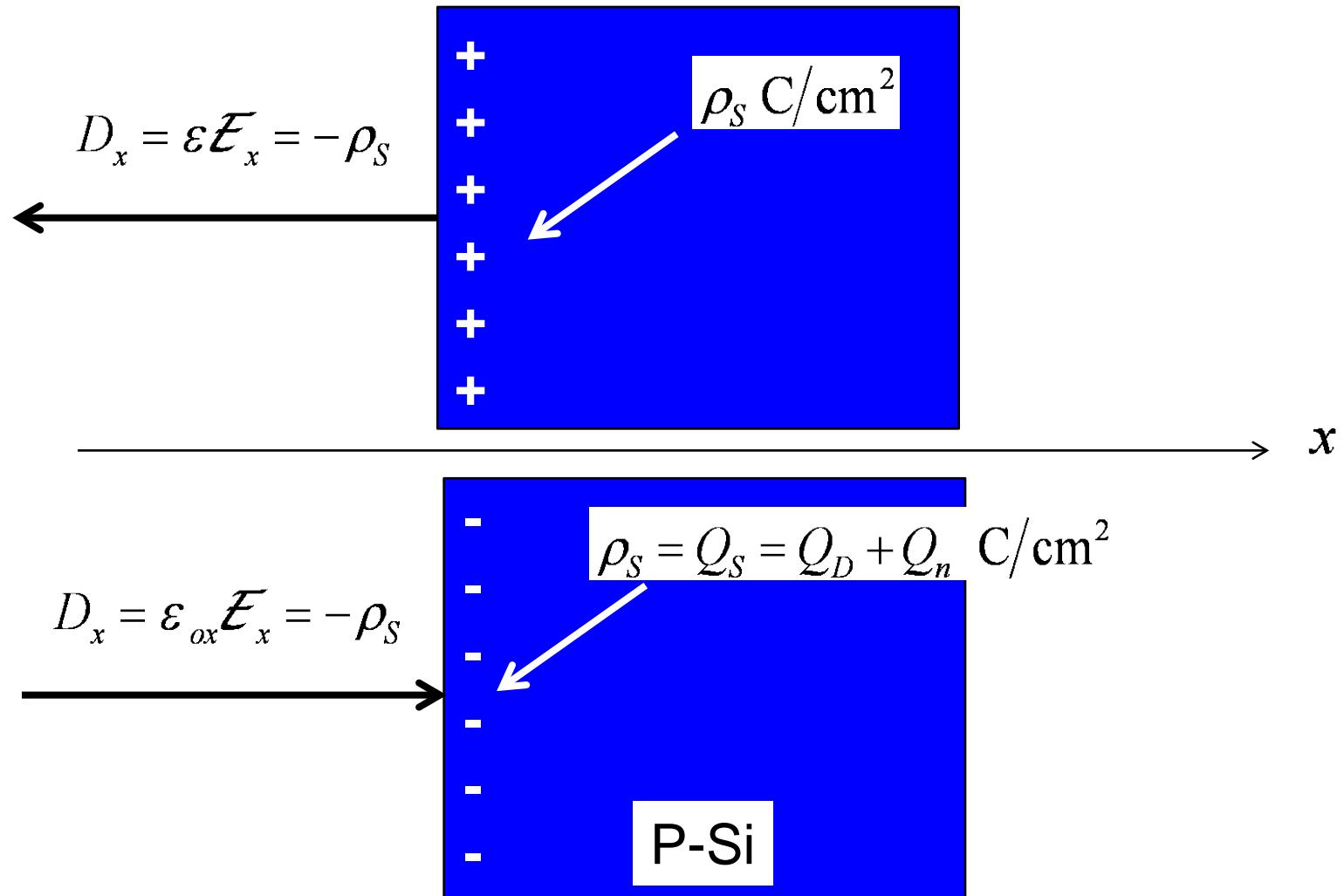
$$V_G = \Delta V_{ox} + \Delta V_{Si}$$

$$\Delta V_{Si} = \psi_s$$

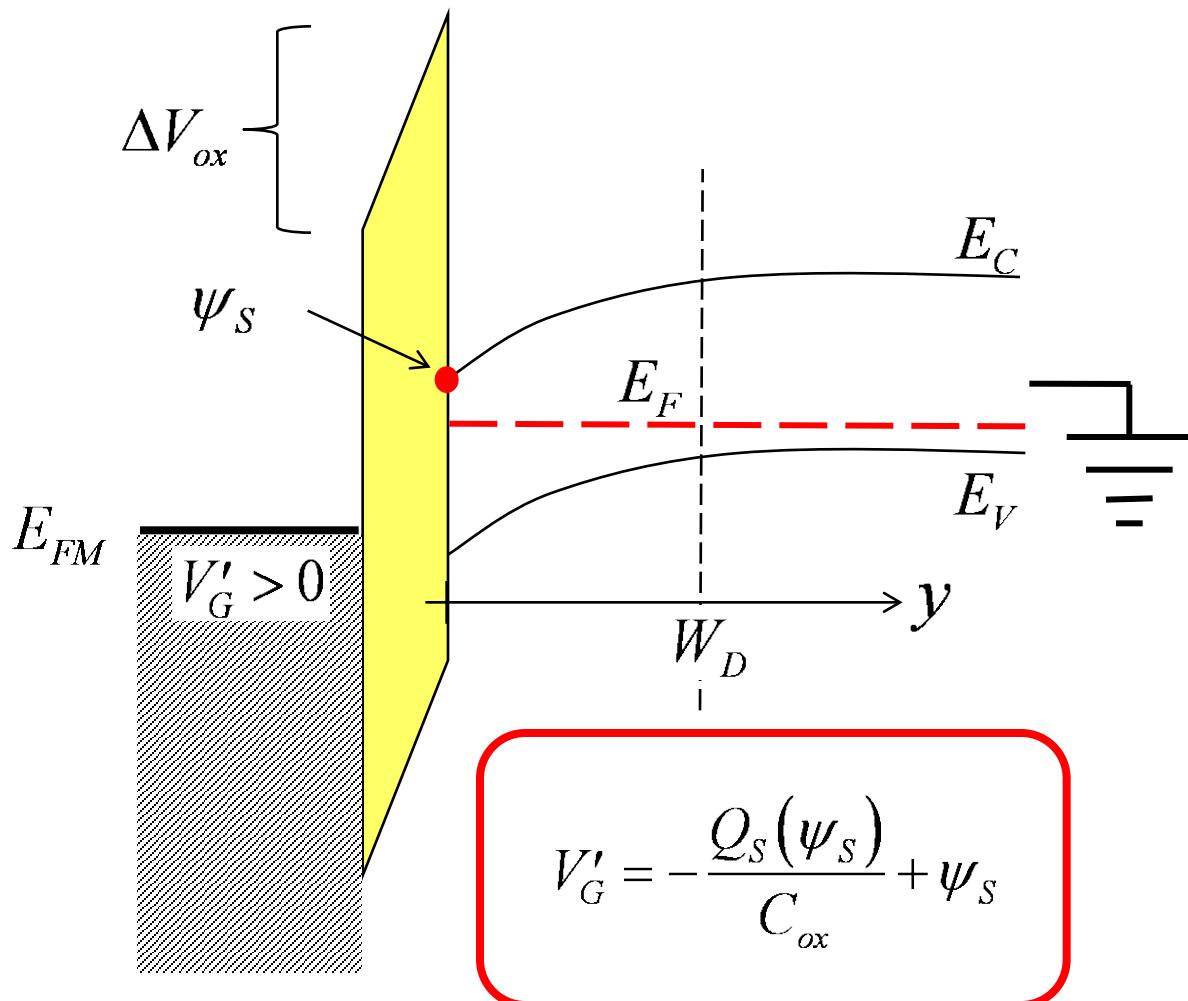
$$\Delta V_{ox} = \mathcal{E}_{ox} t_{ox}$$

$$\varepsilon_{ox} \mathcal{E}_{ox} = -Q_S(\psi_s)$$

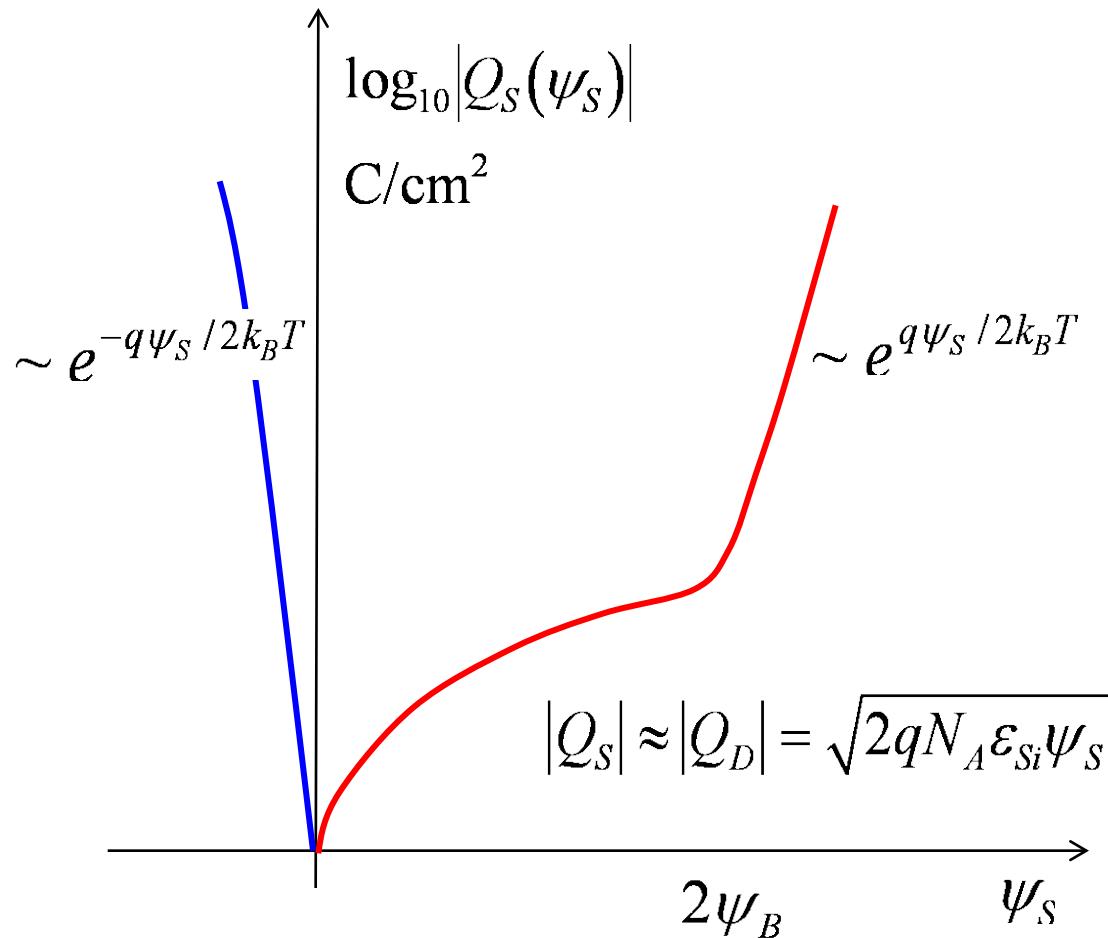
Normal D-field and sheet charge



Gate voltage and surface potential



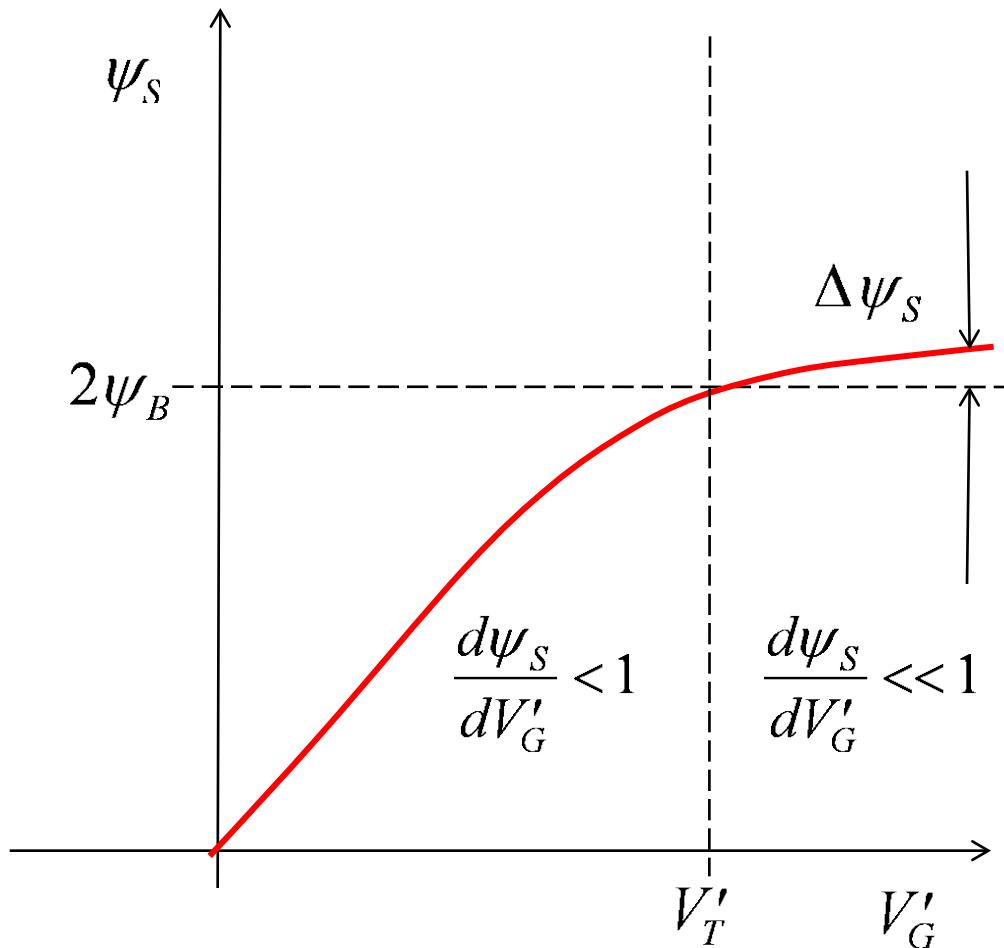
MOS electrostatics



$$V'_G = -\frac{Q_s(\psi_s)}{C_{ox}} + \psi_s$$

- given ψ_s
- determine Q_s
- find V_G

Surface potential vs. gate voltage



$$V'_G = -\frac{\mathcal{Q}_s(\psi_s)}{C_{ox}} + \psi_s$$

Threshold voltage

The gate voltage needed
to make: $\psi_s = 2\psi_b$

$$V'_G = V'_T = -\frac{Q_s(2\psi_b)}{C_{ox}} + 2\psi_b$$

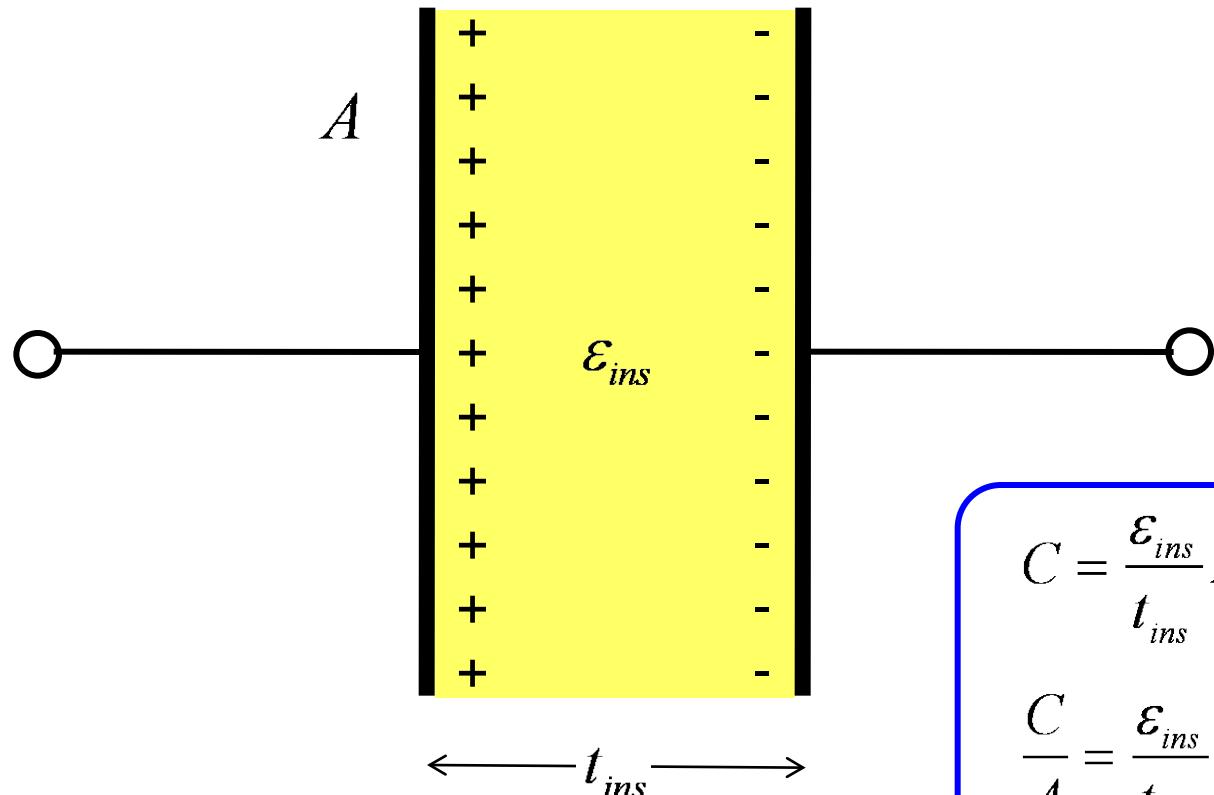
$$\begin{aligned} Q_s(2\psi_b) &= Q_d(2\psi_b) + Q_n(2\psi_b) \\ &\approx Q_d(2\psi_b) \end{aligned}$$

$$V'_G = -\frac{Q_s(\psi_s)}{C_{ox}} + \psi_s$$

$$V'_T = \frac{\sqrt{2qN_A\varepsilon_s(2\psi_b)}}{C_{ox}} + 2\psi_b$$

$$\psi_b = \frac{k_B T}{q} \ln\left(\frac{N_A}{n_i}\right)$$

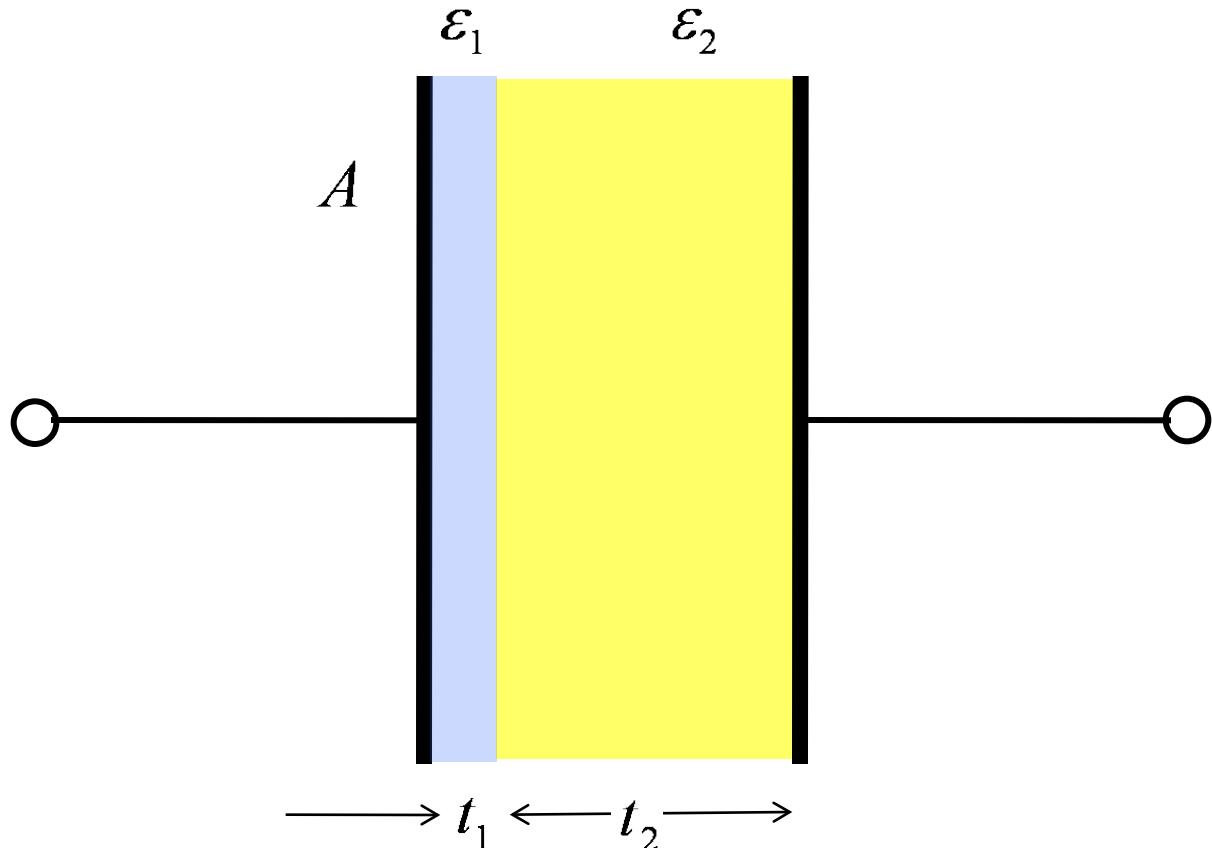
A word about capacitance



$$C = \frac{\epsilon_{ins}}{t_{ins}} A \quad \text{F}$$

$$\frac{C}{A} = \frac{\epsilon_{ins}}{t_{ins}} \quad \text{F/cm}^2$$

Capacitor with two dielectrics

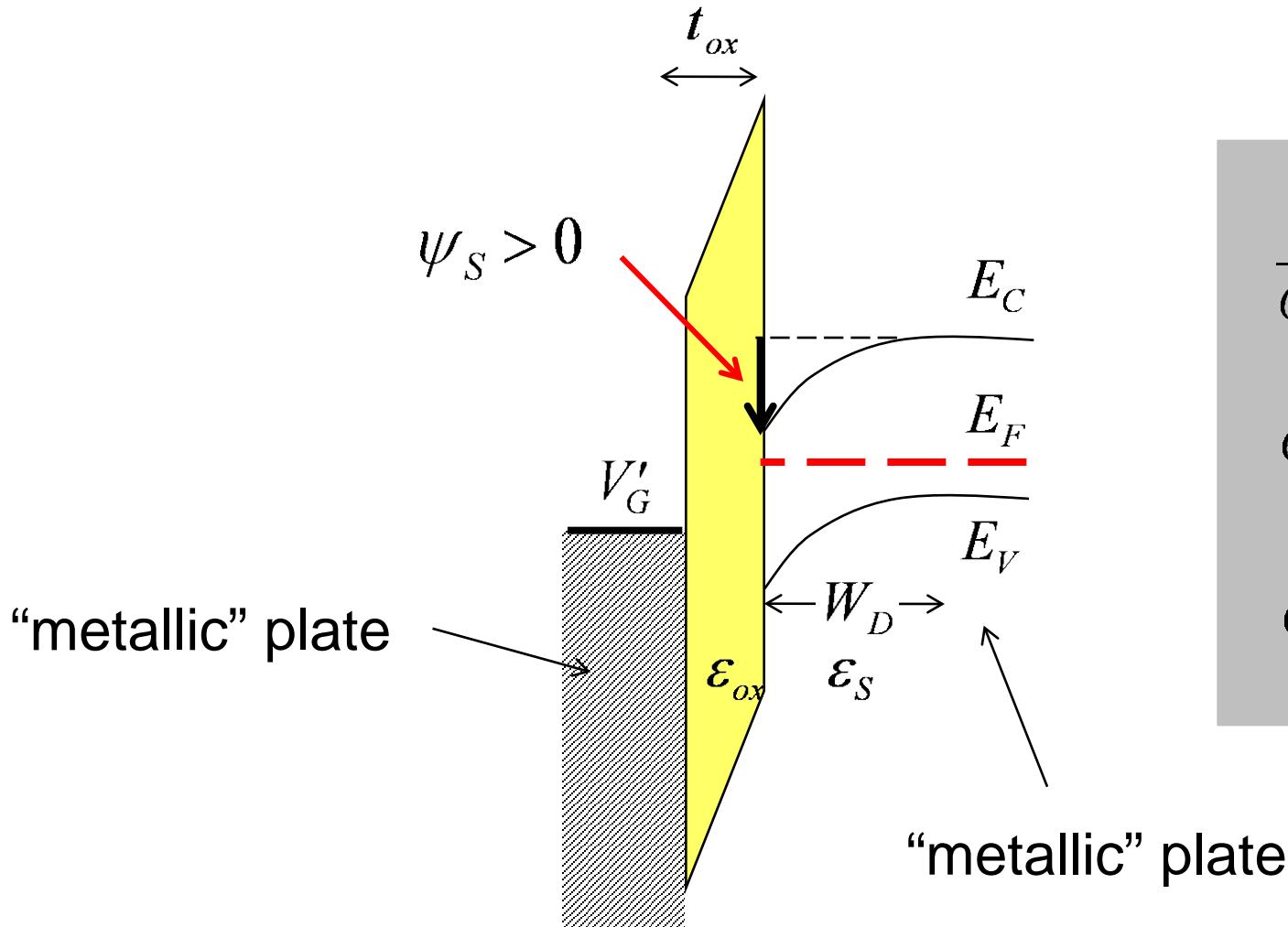


$$\frac{1}{C} = \frac{1}{C_1} + \frac{1}{C_2}$$

$$C_1 = \frac{\epsilon_1}{t_1} A$$

$$C_2 = \frac{\epsilon_2}{t_2} A$$

Depletion capacitance



$$\frac{1}{C} = \frac{1}{C_{ox}} + \frac{1}{C_D}$$

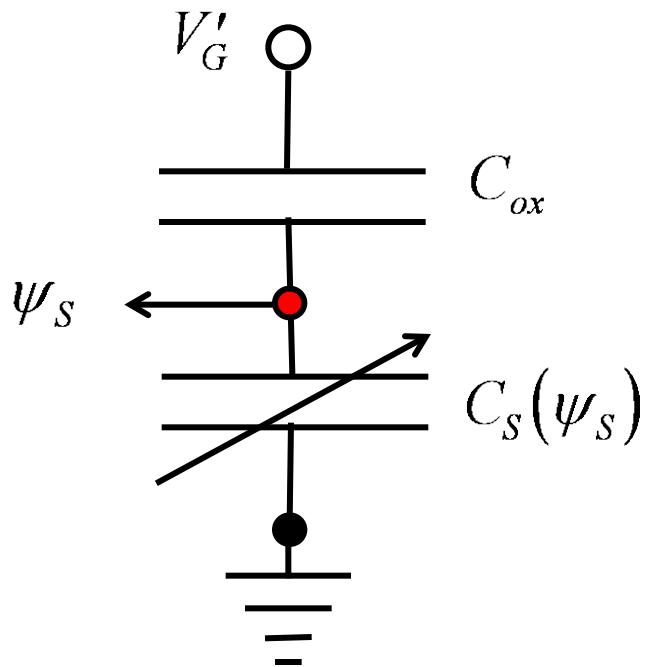
$$C_{ox} = \frac{\epsilon_{ox}}{t_{ox}}$$

$$C_D = \frac{\epsilon_s}{W_D(\psi_S)}$$

Approximate ψ_s vs. V_G relation (depletion)

$$V'_G = -\frac{Q_s(\psi_s)}{C_{ox}} + \psi_s$$

approximate solution



Below threshold:

at low voltage the semiconductor capacitance is the depletion capacitance.

$$C_s \approx C_D = \frac{\epsilon_s}{W_D(\psi_s)}$$

$$\psi_s \approx V'_G \frac{C_{ox}}{C_{ox} + C_D} = \frac{V'_G}{1 + C_D / C_{ox}}$$

$$\psi_s \approx \frac{V'_G}{m}$$

$$m = 1 + C_D / C_{ox}$$

(depletion)

Example

$$N_A = 10^{18} \text{ cm}^{-3}$$

$$W_D = 25 \text{ nm}$$

$$\kappa_{ox} = 3.9$$

$$\psi_S = 0.5 \text{ V}$$

$$t_{ox} = 2 \text{ nm}$$

$$\kappa_{Si} = 11.8$$

$$\psi_S = \frac{V'_G}{m}$$

$$m = 1 + C_D / C_{ox}$$

$$m = 1 + \frac{\epsilon_{Si}}{\epsilon_{ox}} \frac{t_{ox}}{W_D}$$

$$C_D(\psi_S) = \frac{\epsilon_s}{W_D(\psi_S)}$$

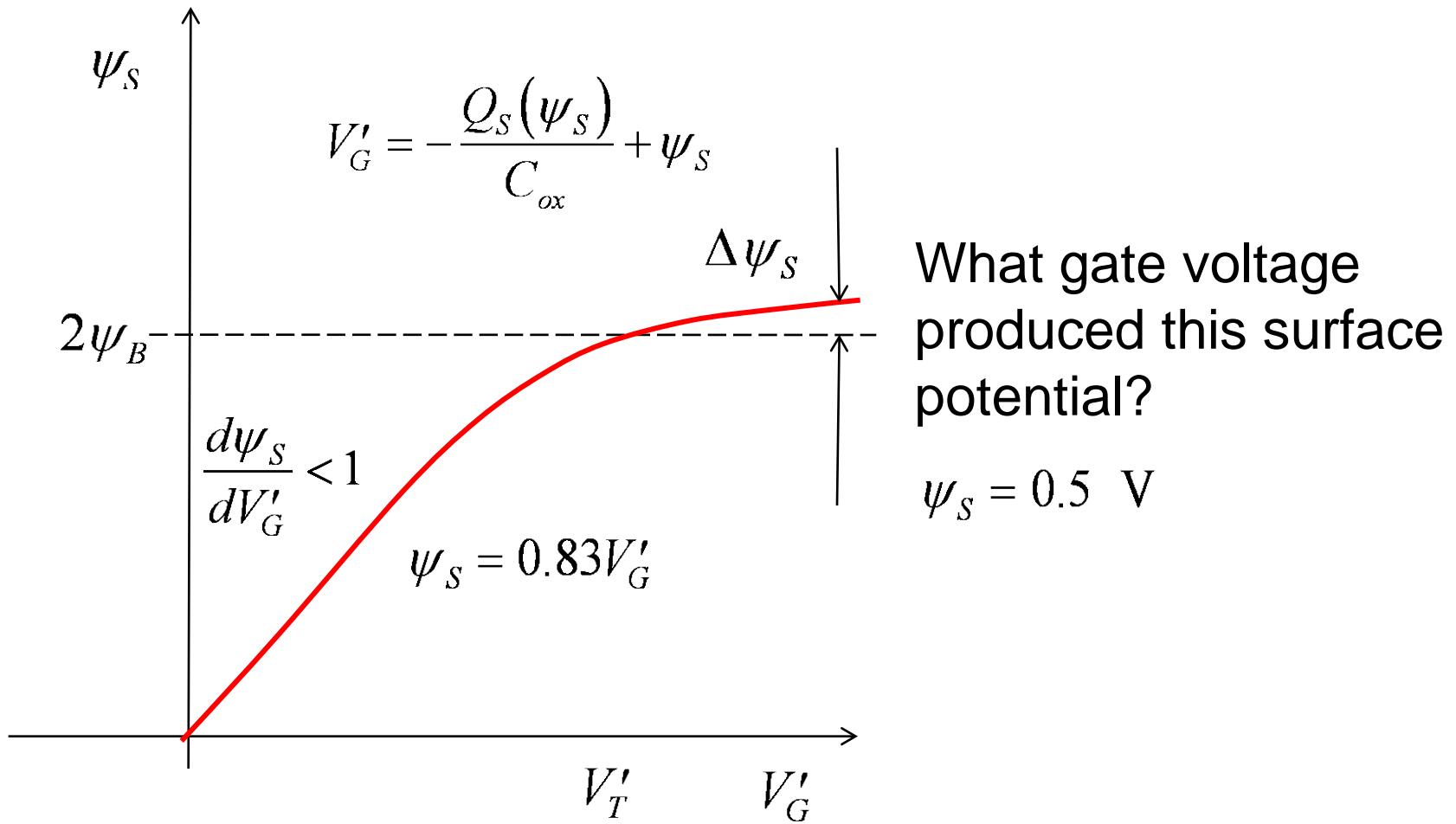
$$C_{ox} = \frac{\epsilon_{ox}}{t_{ox}}$$

$$m = 1 + \left(\frac{11.8}{3.9} \right) \left(\frac{2}{25} \right) \approx 1.2$$

$$\psi_S = \frac{V'_G}{m} = 0.83 V'_G$$

Lundstrom: 2018

Surface potential vs. gate voltage



Gate voltage

$$V'_G = -\frac{Q_D(\psi_s)}{C_{ox}} + \psi_s$$

$$\begin{aligned} V'_G &= \frac{4.1 \times 10^{-7}}{1.73 \times 10^{-6}} + 0.5 \\ &= 0.24 + 0.5 \\ &= 0.75 \text{ V} \end{aligned}$$



$$W_D(\psi_s = 0.5) = 25 \text{ nm}$$

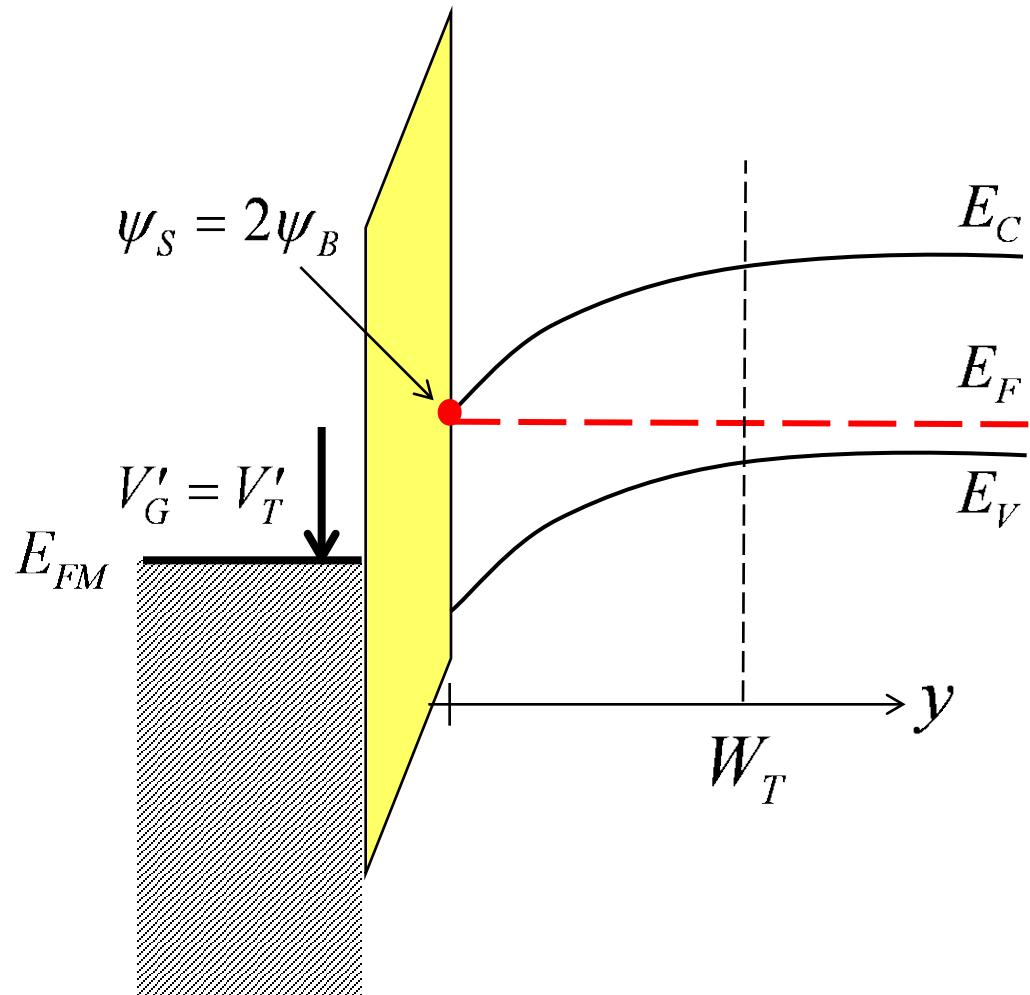
$$Q_D = -qN_A W_D \text{ C/cm}^2$$

$$Q_D = -4.1 \times 10^{-7} \text{ C/cm}^2$$

$$C_{ox} = \frac{\epsilon_{ox}}{t_{ox}}$$

$$\begin{aligned} &= \frac{(3.9)(8.854 \times 10^{-14})}{2 \times 10^{-9}} \\ &= 1.73 \times 10^{-6} \text{ F/cm}^2 \end{aligned}$$

Surface potential at the onset of inversion



threshold voltage

Lundstrom: 2018

$$\psi_S = 2\psi_B$$

$$\psi_B = \frac{k_B T}{q} \ln \left(\frac{N_A}{n_i} \right)$$

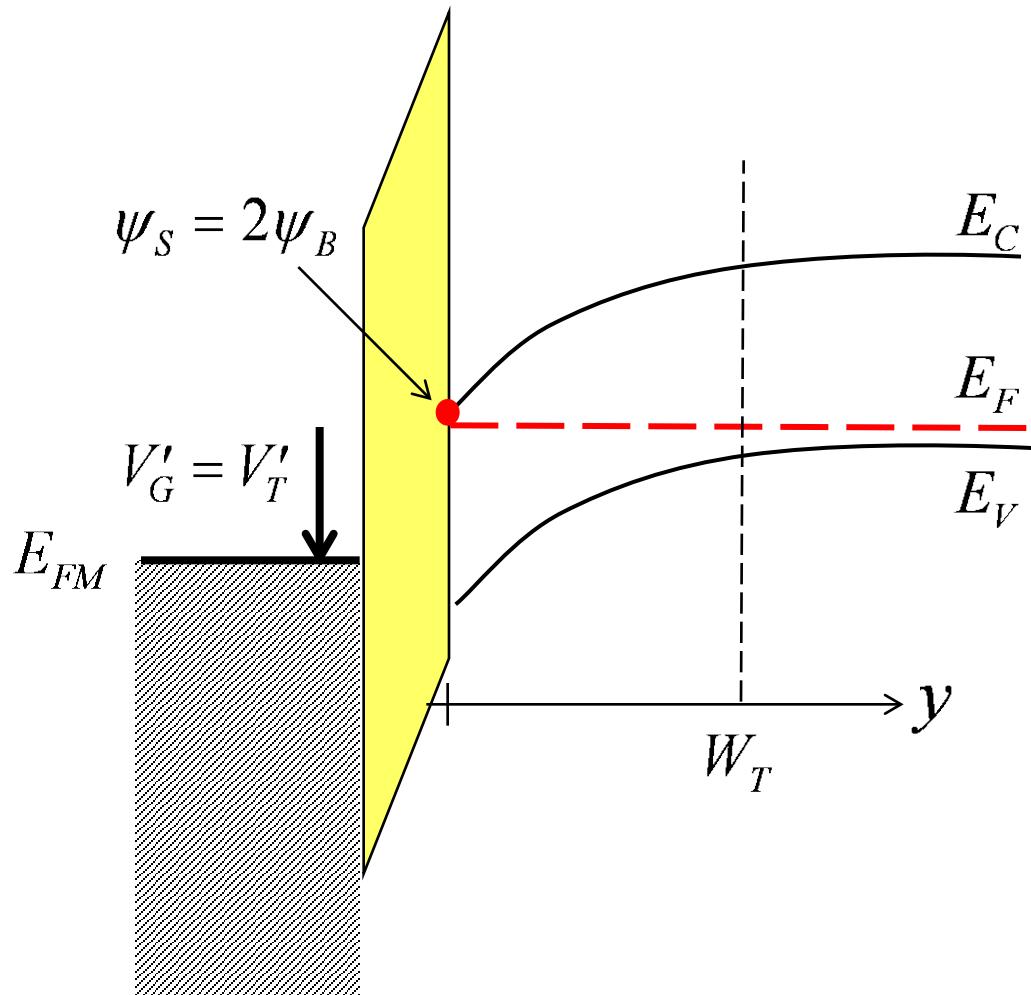
$$N_A = 10^{18} \text{ cm}^{-3}$$

$$n_i = 10^{10} \text{ cm}^{-3}$$

$$\psi_B = 0.48 \text{ V}$$

$$2\psi_B = 0.96 \text{ V}$$

Internal quantities at the onset of inversion



$$W_T = \sqrt{2\epsilon_s(2\psi_B)/qN_A}$$

$$\mathcal{E}_s = \frac{-Q_D}{\epsilon_{Si}}$$

$$W_T = 35 \text{ nm}$$

$$Q_D = -5.6 \times 10^{-7} \text{ C/cm}^2$$

$$Q_D/q = 3.5 \times 10^{12} \text{ #/cm}^2$$

$$\mathcal{E}_s = 5.4 \times 10^5 \text{ V/cm}$$

Gate voltage at the onset of inversion

$$V'_T = -\frac{Q_D(2\psi_B)}{C_{ox}} + \psi_S$$

$$\begin{aligned} V'_T &= \frac{5.6 \times 10^{-7}}{1.73 \times 10^{-6}} + 0.96 \\ &= 0.32 + 0.96 \\ &= 1.28 \text{ V} \end{aligned}$$



$$2\psi_B = 0.96 \text{ V}$$

$$W_D = 35 \text{ nm}$$

$$Q_D = -5.6 \times 10^{-7} \text{ C/cm}^2$$

$$C_{ox} = 1.73 \times 10^{-2} \text{ F/m}^2$$

Note: This is a rather large threshold voltage because we have not included the effect of the metal-semiconductor work function difference (to be discussed in the next lecture).

Summary

- 1) The gate voltage induces charge in the semiconductor by bending the bands.
- 2) There is a simple (exact) relation between the gate voltage and the surface potential, but it must be solved numerically.

$$V'_G = -\frac{Q_s(\psi_s)}{C_{ox}} + \psi_s$$

- 3) There is an approximate relation between gate voltage and surface potential that works well in depletion.

$$\psi_s = \frac{V'_G}{m} \quad m = 1 + \frac{C_d(\psi_s)}{C_{ox}}$$

Next topic

We have discussed an ideal MOS capacitor. In the next lecture we will add two important factors that affect real MOS capacitors.

Essentials of MOSFETs

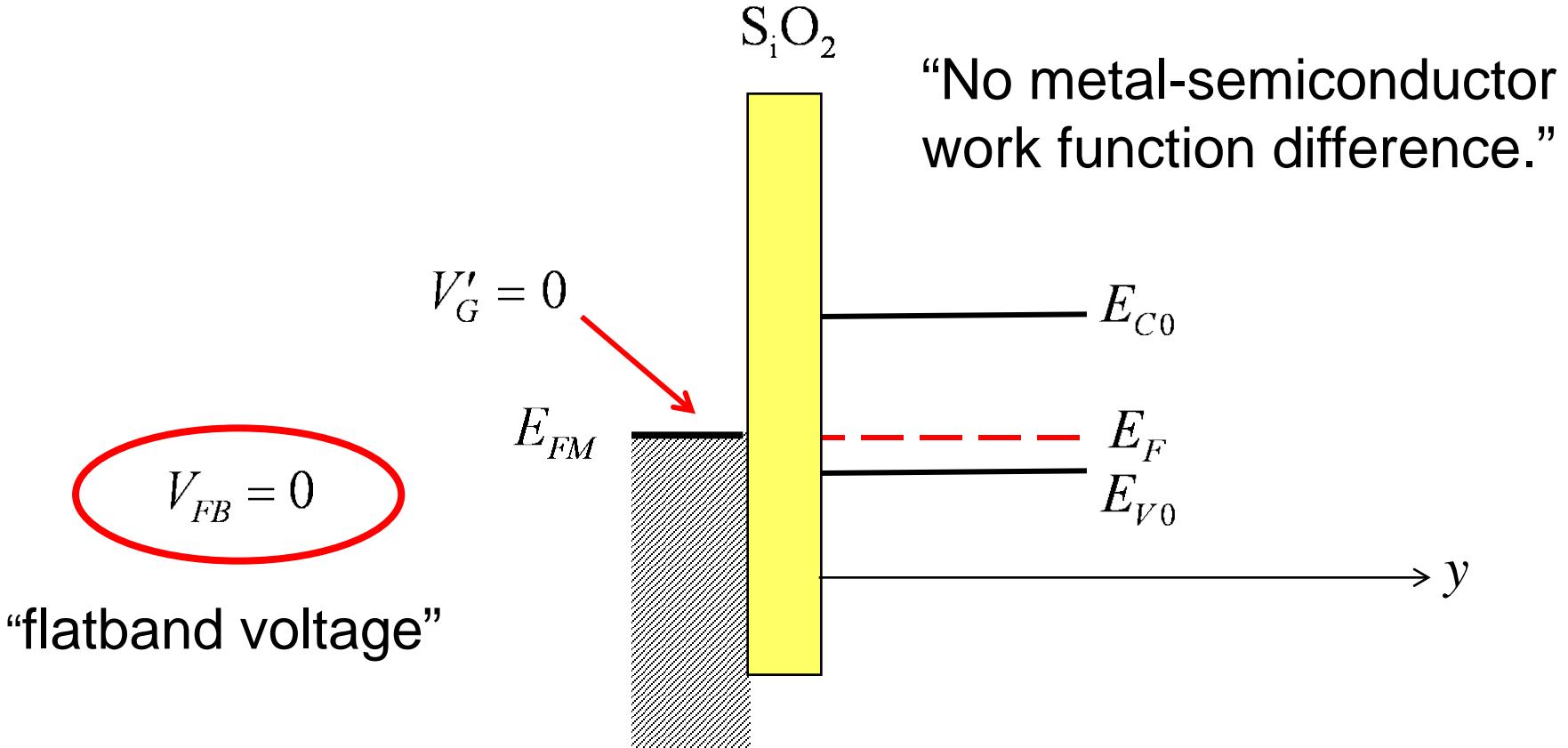
Unit 3: MOS Electrostatics

Lecture 3.4: Flat-band Voltage

Mark Lundstrom

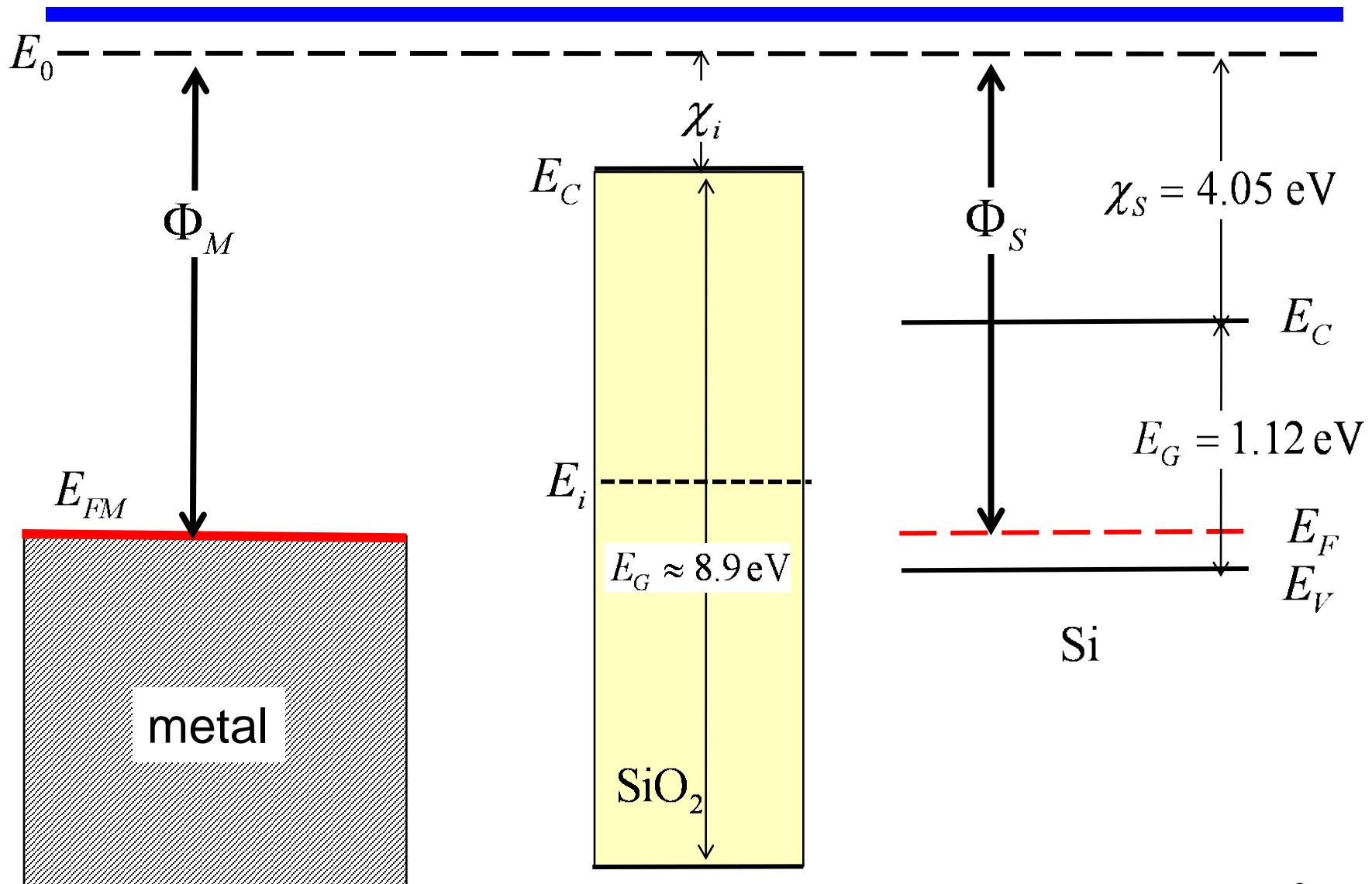
lundstro@purdue.edu
Electrical and Computer Engineering
Purdue University
West Lafayette, Indiana USA

Hypothetical, ideal MOS-C

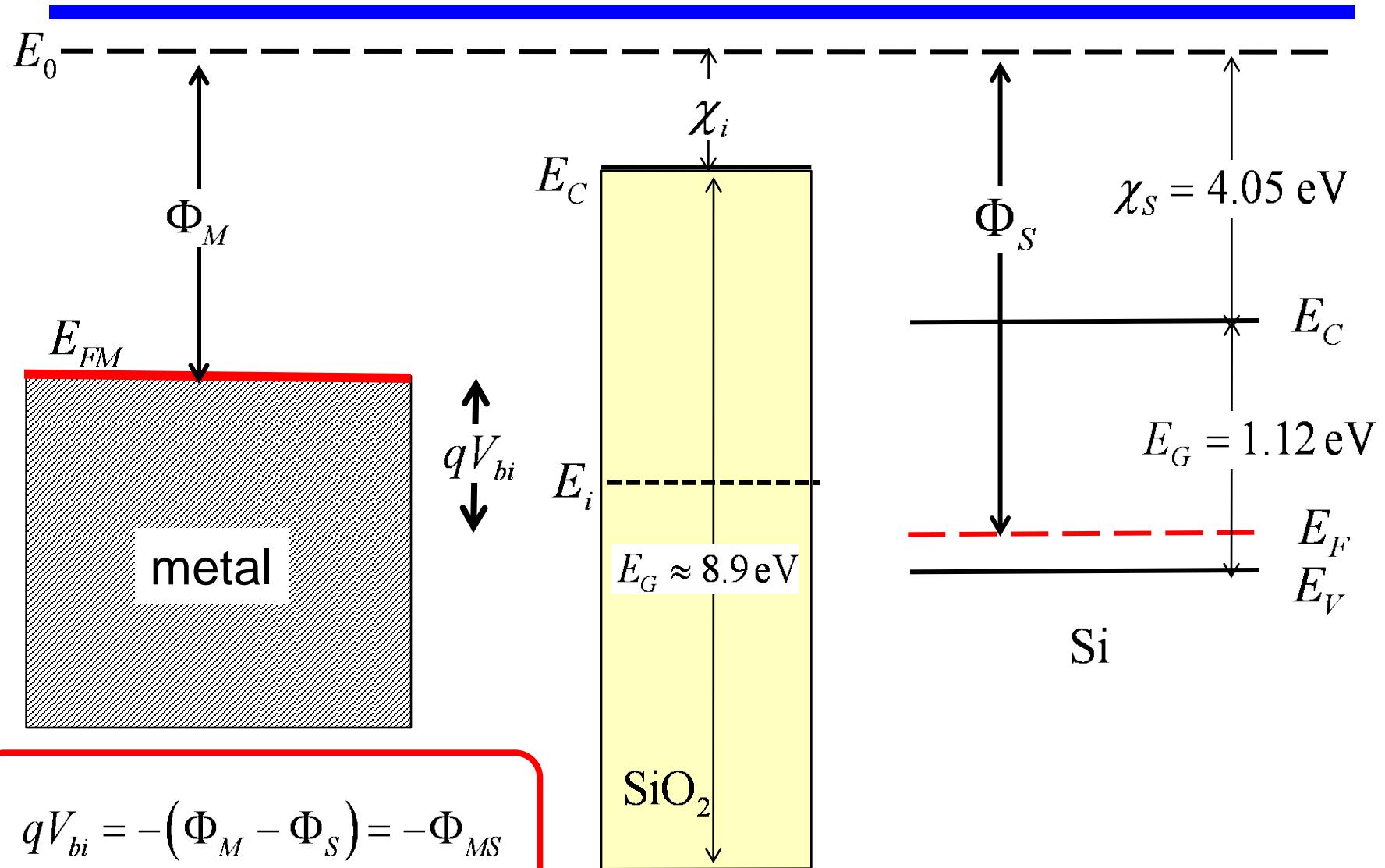


The Fermi level in this “special” metal lines up with the Fermi level in the semiconductor at zero gate voltage.)

Hypothetical, ideal MOS-C



Real MOS-C



Real MOS-C at $V_G = 0$

$$V_{bi} = -\frac{\Phi_{MS}}{q} = -\phi_{ms}$$
$$V_{FB} = -V_{bi} = \phi_{ms}$$

$$V_G = 0$$

metal

$$\psi(\text{metal}) = V_{bi}$$

$$\psi(x=0) = \psi_s \text{ surface potential}$$

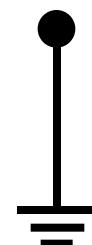
$$\psi(x) = 0 \text{ in the bulk}$$

$$E_C$$

$$E_i$$

$$E_F$$

$$E_V$$



Example

Aluminum metal and p-type Si

$$N_A = 10^{18} \text{ cm}^{-3}$$

$$p_0 = N_V e^{(E_V - E_F)/k_B T} \text{ cm}^{-3}$$

$$E_F - E_V = k_B T \ln\left(\frac{N_V}{N_A}\right)$$

$$N_V = 1.83 \times 10^{19} \text{ cm}^{-3}$$

$$\frac{E_F - E_V}{q} = 0.08 \text{ eV}$$

$$\Phi_M = 4.08 \text{ eV}$$

$$\Phi_S = \chi_S + E_G - (E_F - E_V)/q$$

$$\chi_S = 4.05 \text{ eV} \quad E_G = 1.12 \text{ eV}$$

$$\Phi_S = 5.09 \text{ eV}$$

$$\phi_{ms} = \frac{(\Phi_M - \Phi_S)}{q} = -1.01 \text{ V}$$

$$V_{FB} = \phi_{ms} = -1.01 \text{ V}$$



Gate voltage vs. surface potential

$$V'_G = -\frac{Q_s(\psi_s)}{C_{ox}} + \psi_s$$

$$V_G = V'_G + V_{FB} = V_{FB} - \frac{Q_s(\psi_s)}{C_{ox}} + \psi_s$$

$$V_G = V_{FB} - \frac{Q_s(\psi_s)}{C_{ox}} + \psi_s$$

$$V_{FB} = \phi_{ms} = \Phi_{MS}/q$$

Recall: Threshold voltage example

$$V'_G = -\frac{Q_s(\psi_s)}{C_{ox}} + \psi_s$$

$$V'_T = -\frac{Q_d(2\psi_B)}{C_{ox}} + 2\psi_B$$

$$V'_T = 1.28 \text{ V}$$

$$V_T = V_{FB} + 1.28 \text{ V} = 0.27 \text{ V} \quad \checkmark$$

$$2\psi_B = 0.96 \text{ V}$$

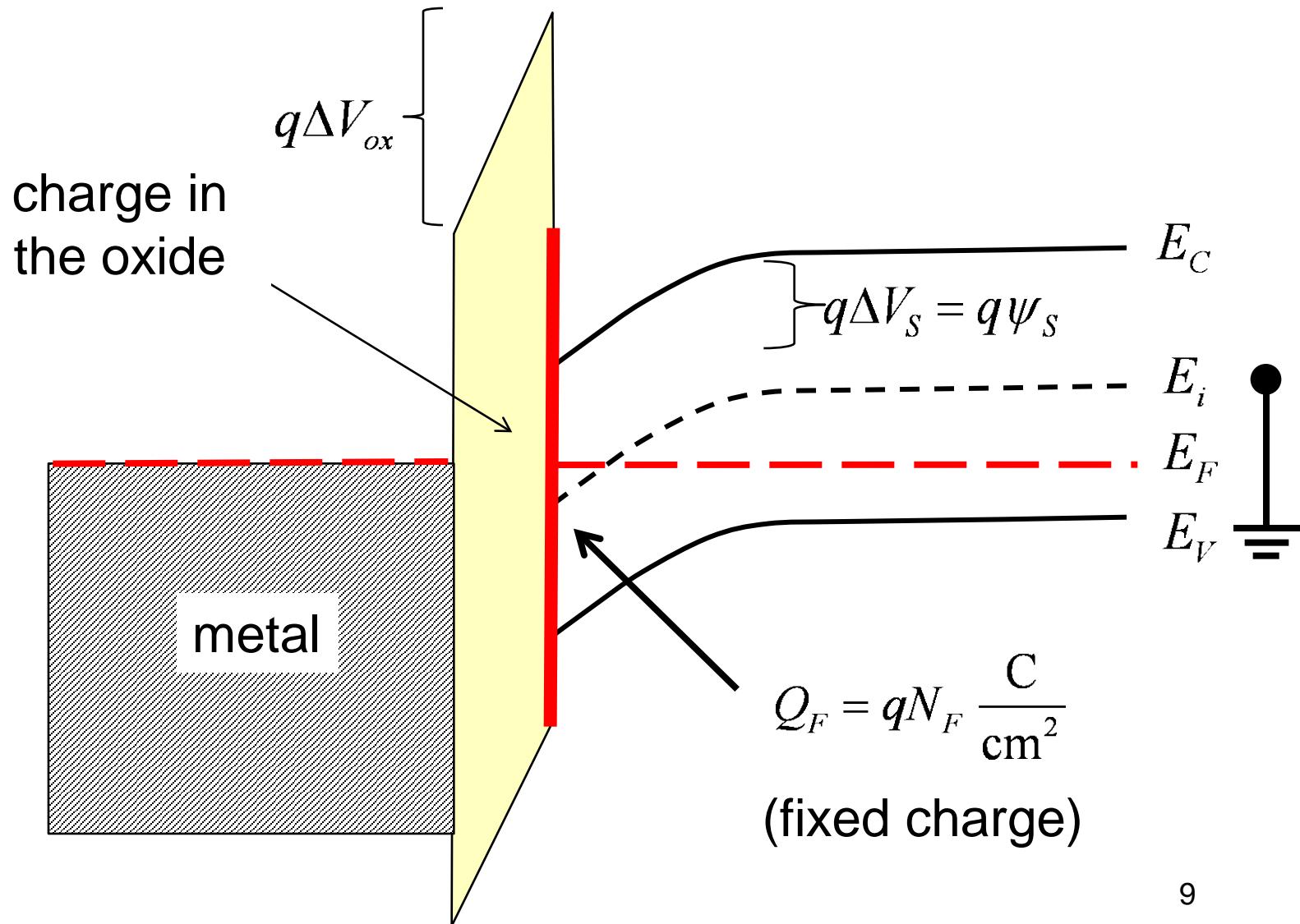
$$W_D = 35 \text{ nm}$$

$$Q_D = -5.6 \times 10^{-7} \text{ C/cm}^2$$

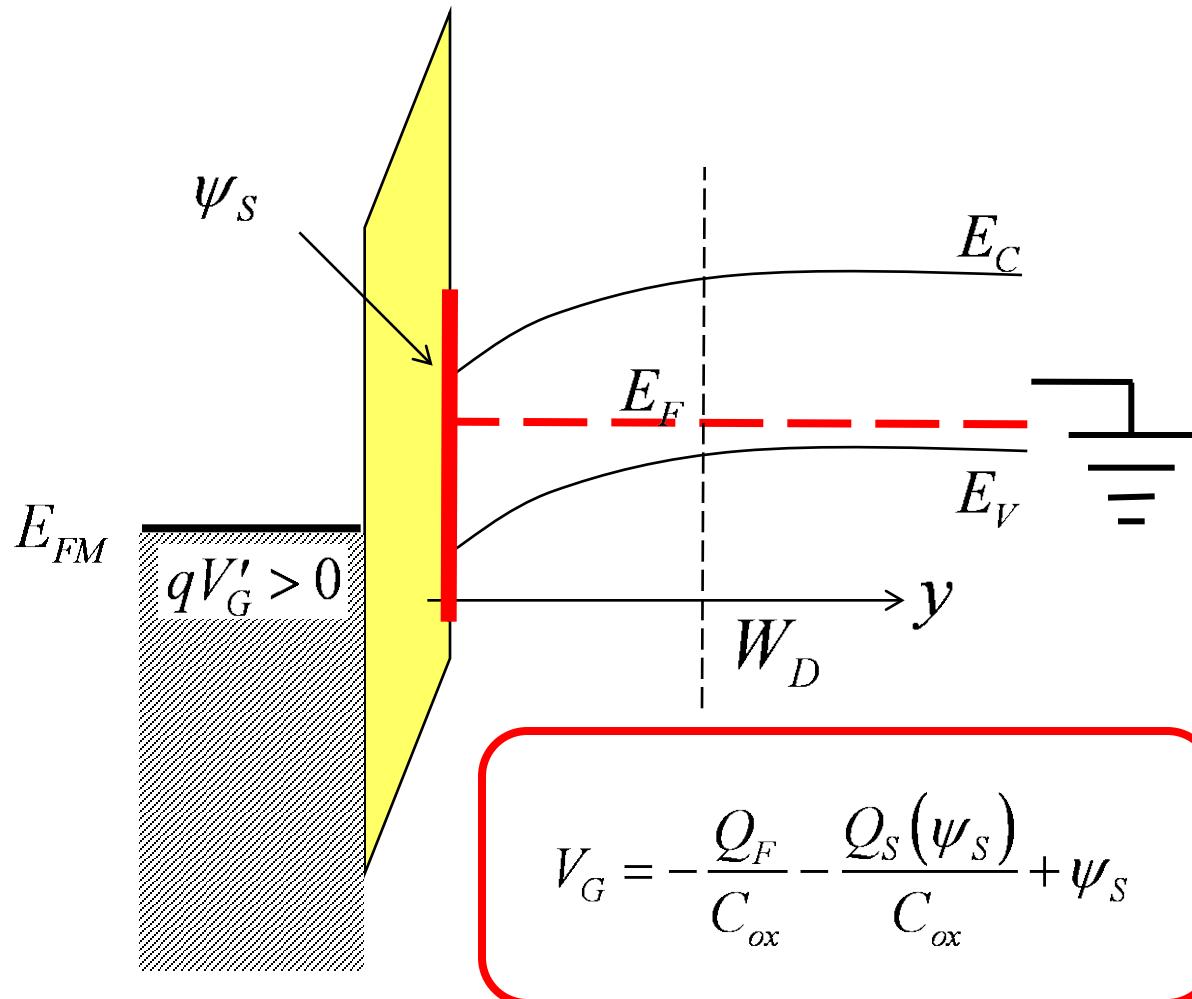
$$C_{ox} = 1.73 \times 10^{-2} \text{ F/m}^2$$

$$V_{FB} = -1.01 \text{ V}$$

Charge at the oxide-semiconductor interface



Volt drop across the oxide at fixed surface potential



$$V'_G = \Delta V_{ox} + \Delta V_{Si}$$

$$\Delta V_{ox} = \mathcal{E}_{ox} t_{ox}$$

$$\varepsilon_{ox} \mathcal{E}_{ox} = -Q_F - Q_S(\psi_S)$$

$$\Delta V_{ox} = \frac{-Q_F - Q_S(\psi_S)}{C_{ox}}$$

$$C_{ox} = \frac{\varepsilon_{ox}}{t_{ox}} \text{ F/cm}^2$$

Flat-band voltage again

$$V_G = V_{FB} - \frac{Q_s(\psi_s)}{C_{ox}} + \psi_s$$

$$V_{FB} = \phi_{ms} - \frac{Q_F}{C_{ox}}$$

Recall: Threshold voltage example

$$V_T = V_{FB} - \frac{Q_D(2\psi_B)}{C_{ox}} + 2\psi_B$$

$$V_T = 0.27 \text{ V}$$

$$Q_F = qN_F = 1.6 \times 10^{-8} \text{ C/cm}^2$$

$$\frac{Q_F}{C_{ox}} = 0.01 \text{ V}$$

$$V_{FB} = \phi_{ms} - \frac{Q_F}{C_{ox}}$$

$$V_T = 0.26 \text{ V}$$



$$2\psi_B = 0.96 \text{ V}$$

$$W_D = 35 \text{ nm}$$

$$Q_D = -5.6 \times 10^{-7} \text{ C/cm}^2$$

$$C_{ox} = 1.73 \times 10^{-6} \text{ F/cm}^2$$

$$\phi_{ms} = -1.01 \text{ V}$$

$$N_F = 10^{11} \text{ cm}^{-2} \text{ (positive charges)}$$

Summary

- 1) The flat-band voltage in a real MOS-C is non-zero.

$$V_{FB} = \phi_{ms} - \frac{Q_F}{C_{ox}}$$

- 2) The gate voltage relation is:

$$V_G = V_{FB} - \frac{Q_S(\psi_S)}{C_{ox}} + \psi_S$$

Next topic

Measuring the small signal capacitance as a function of DC bias voltage is a powerful technique for characterizing MOS structures.

Understanding MOS CV characteristics is the subject of the next lecture.

Essentials of MOSFETs

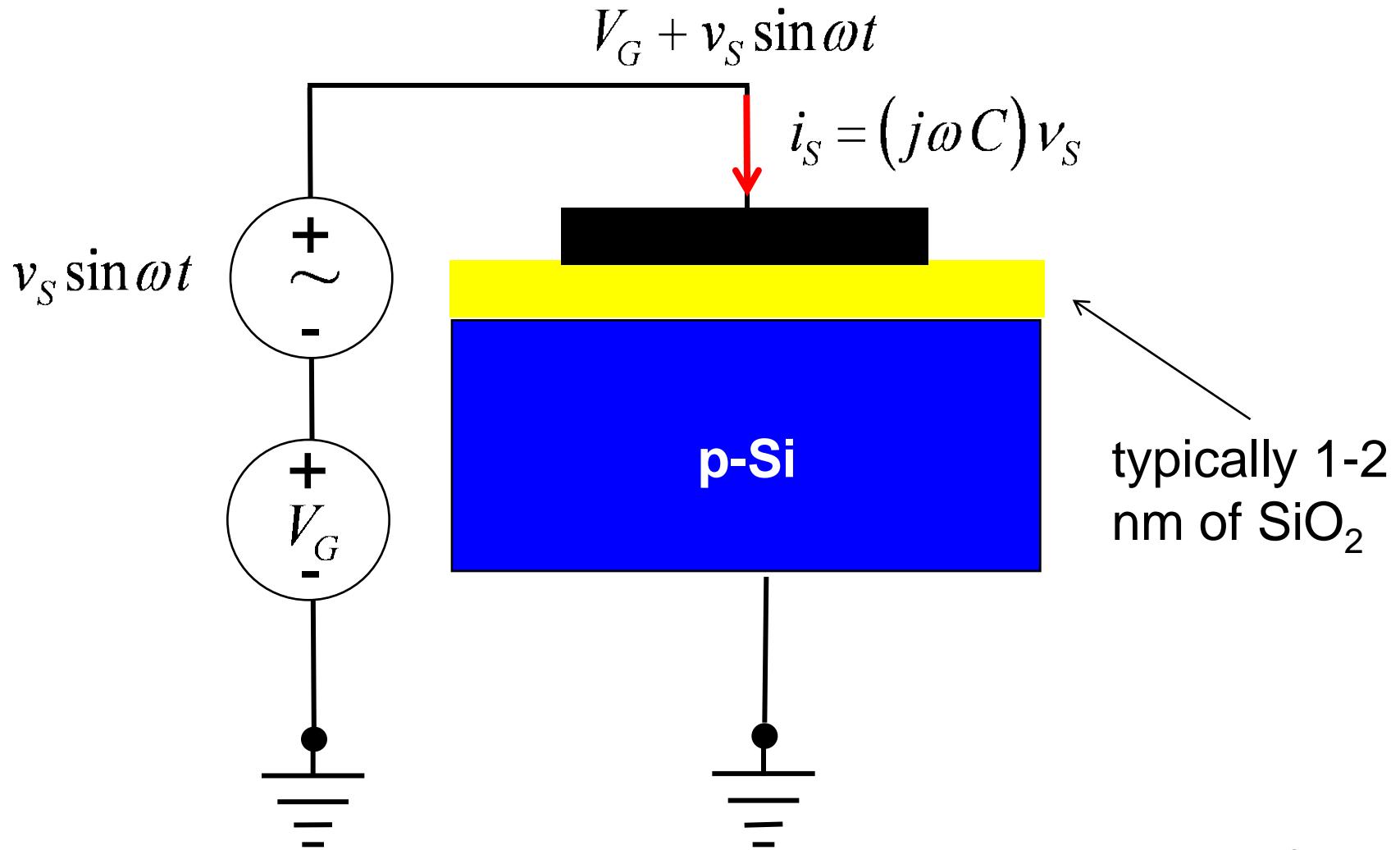
Unit 3: MOS Electrostatics

Lecture 3.5: MOS CV

Mark Lundstrom

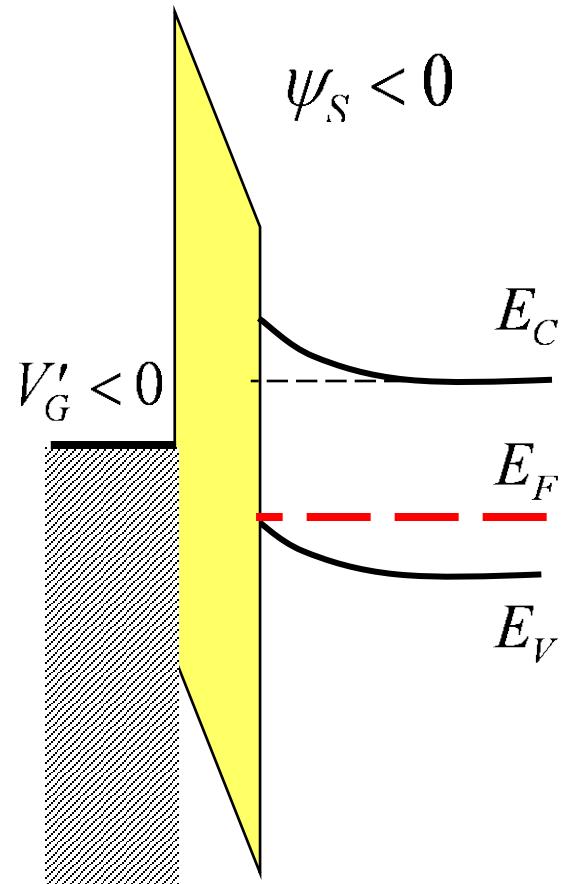
lundstro@purdue.edu
Electrical and Computer Engineering
Purdue University
West Lafayette, Indiana USA

MOS capacitor

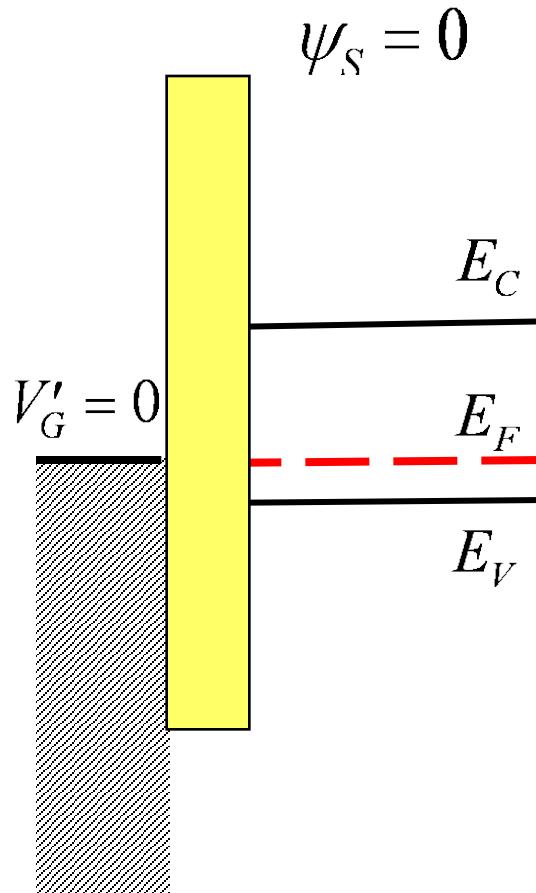


Effect of DC bias

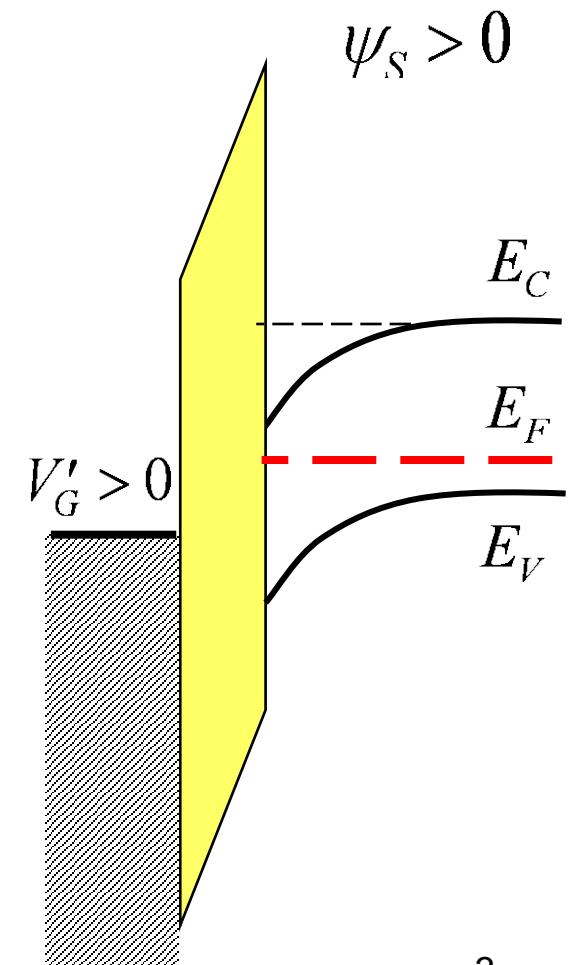
accumulation



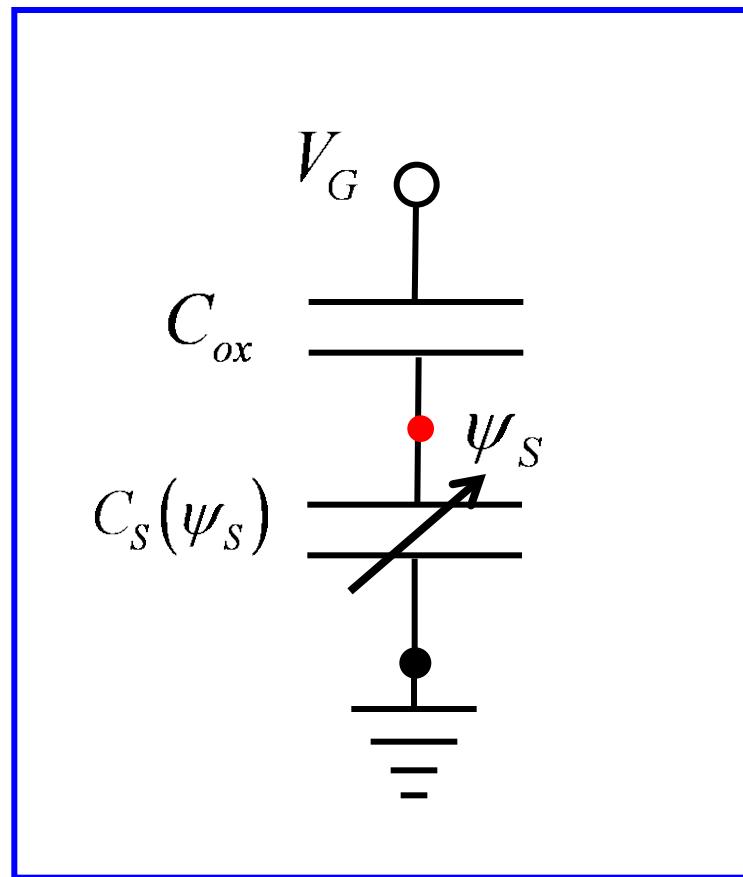
flat band



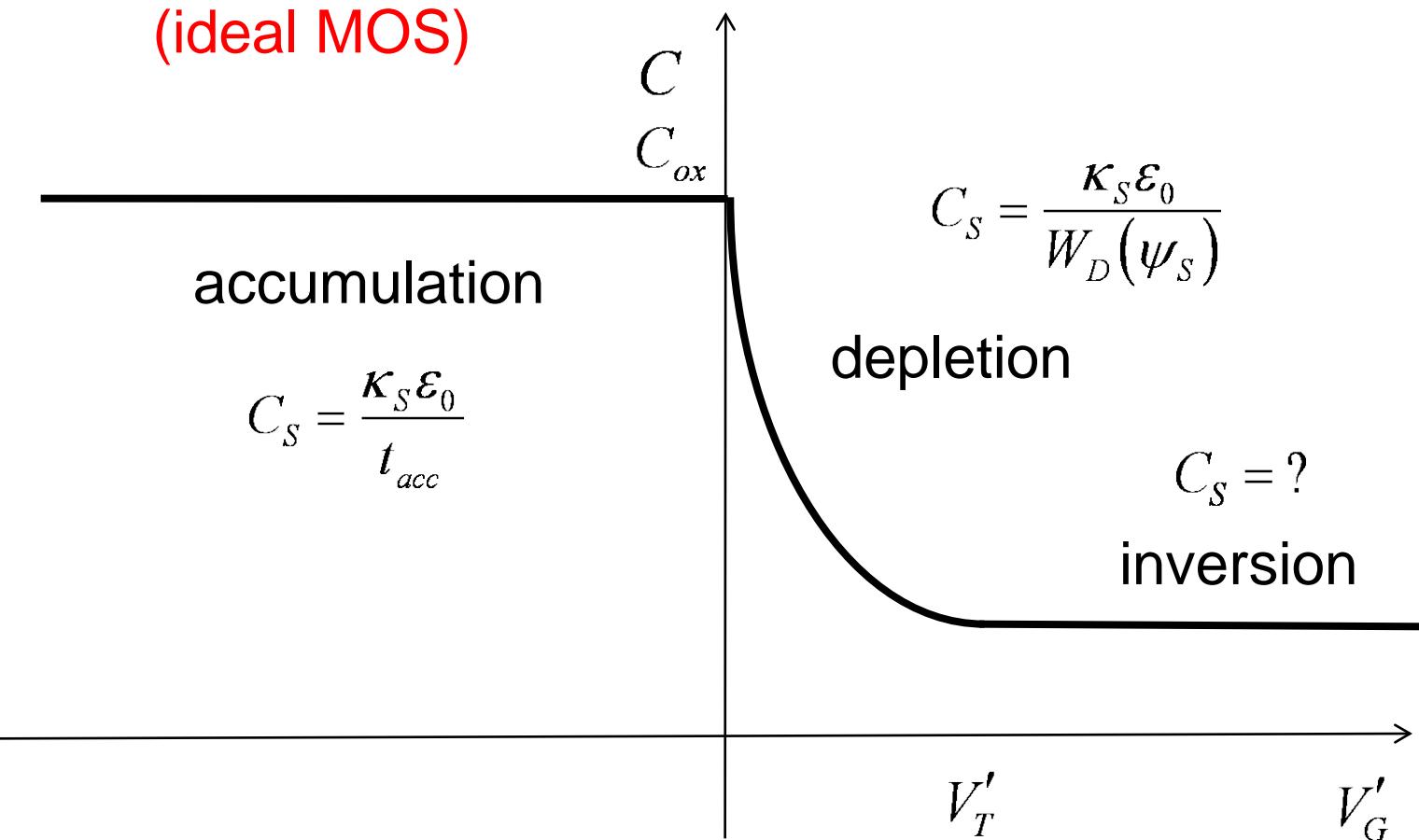
depletion/
inversion



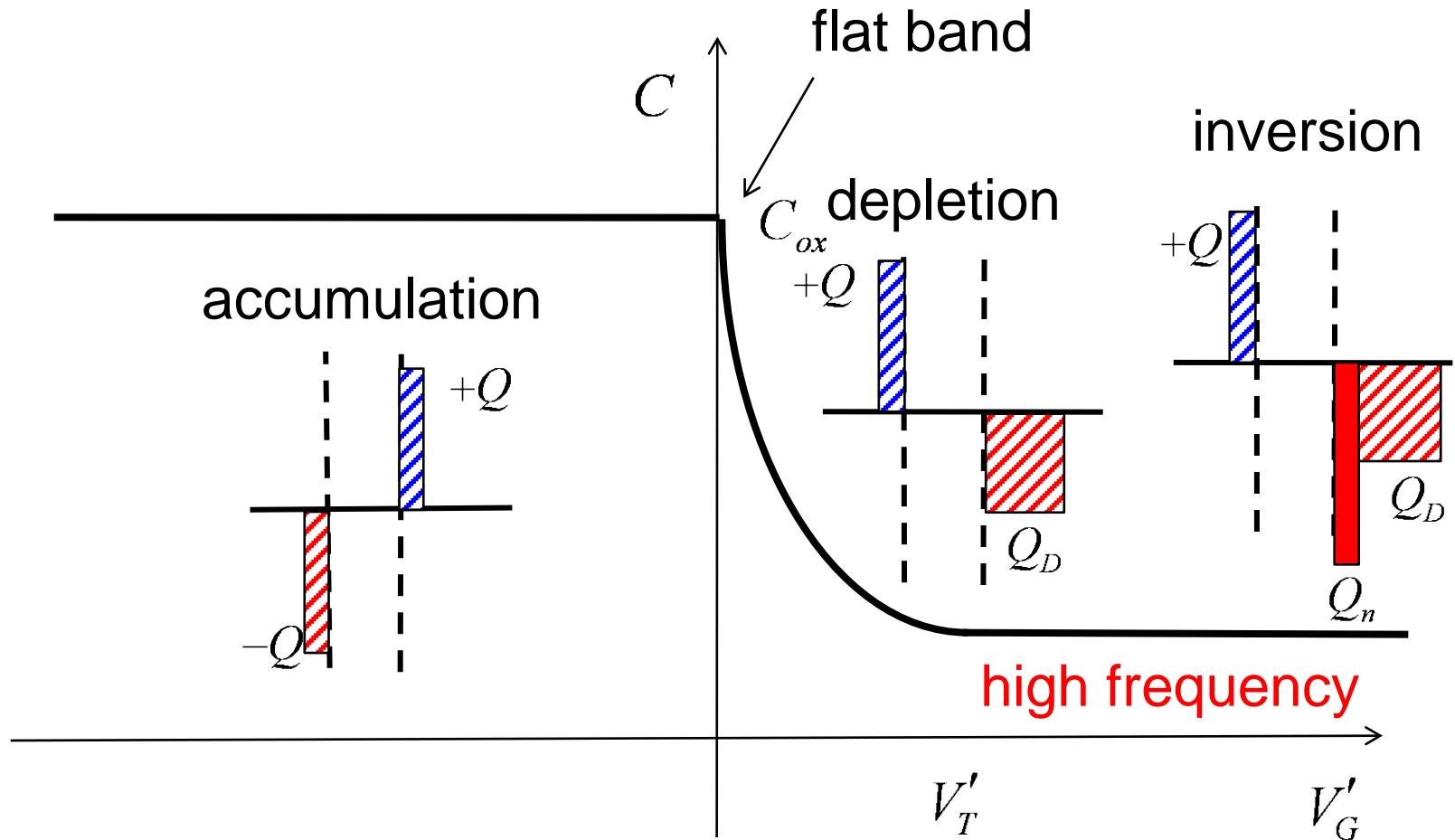
Two capacitors in series



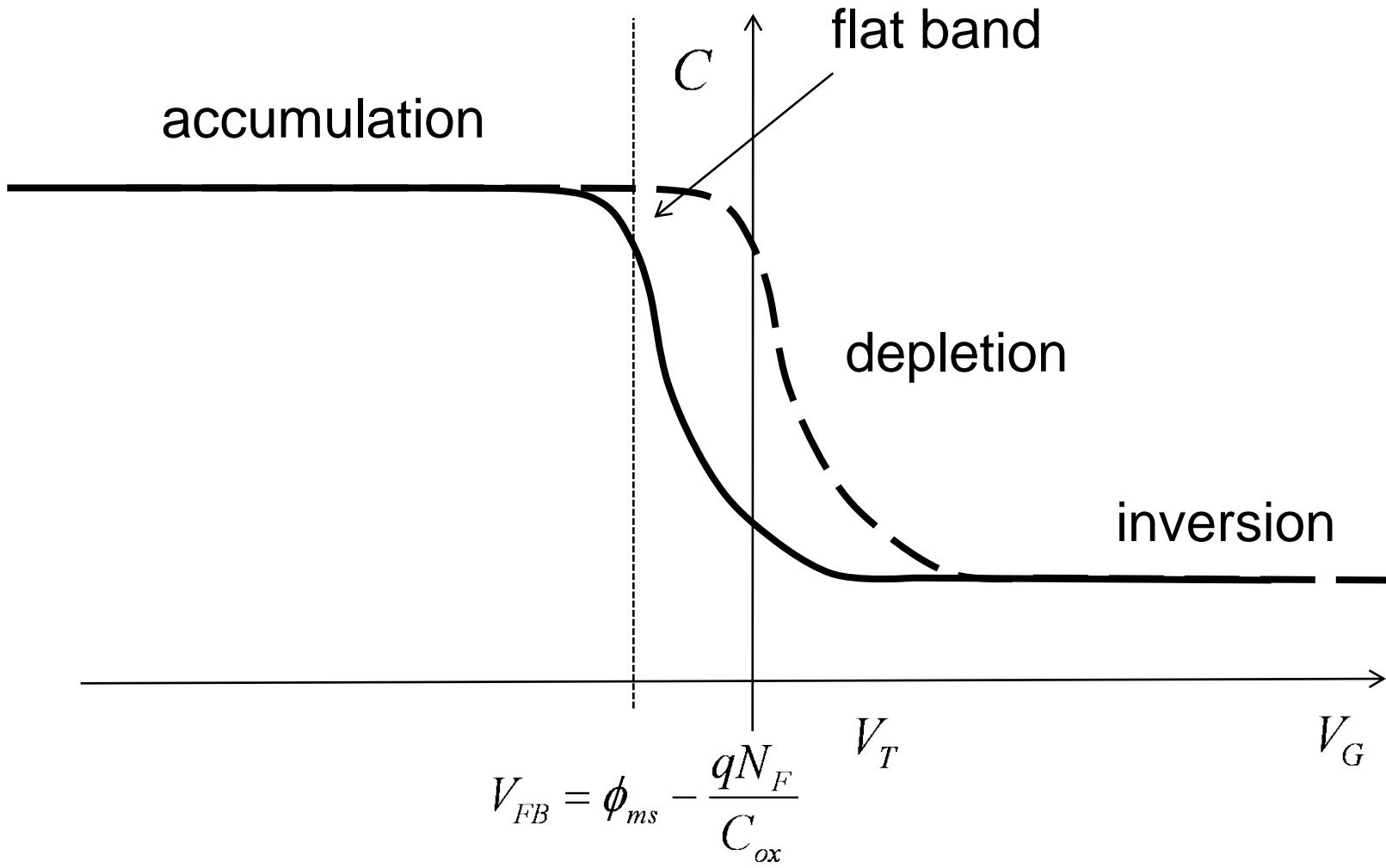
Small signal gate capacitance vs. d.c. gate bias



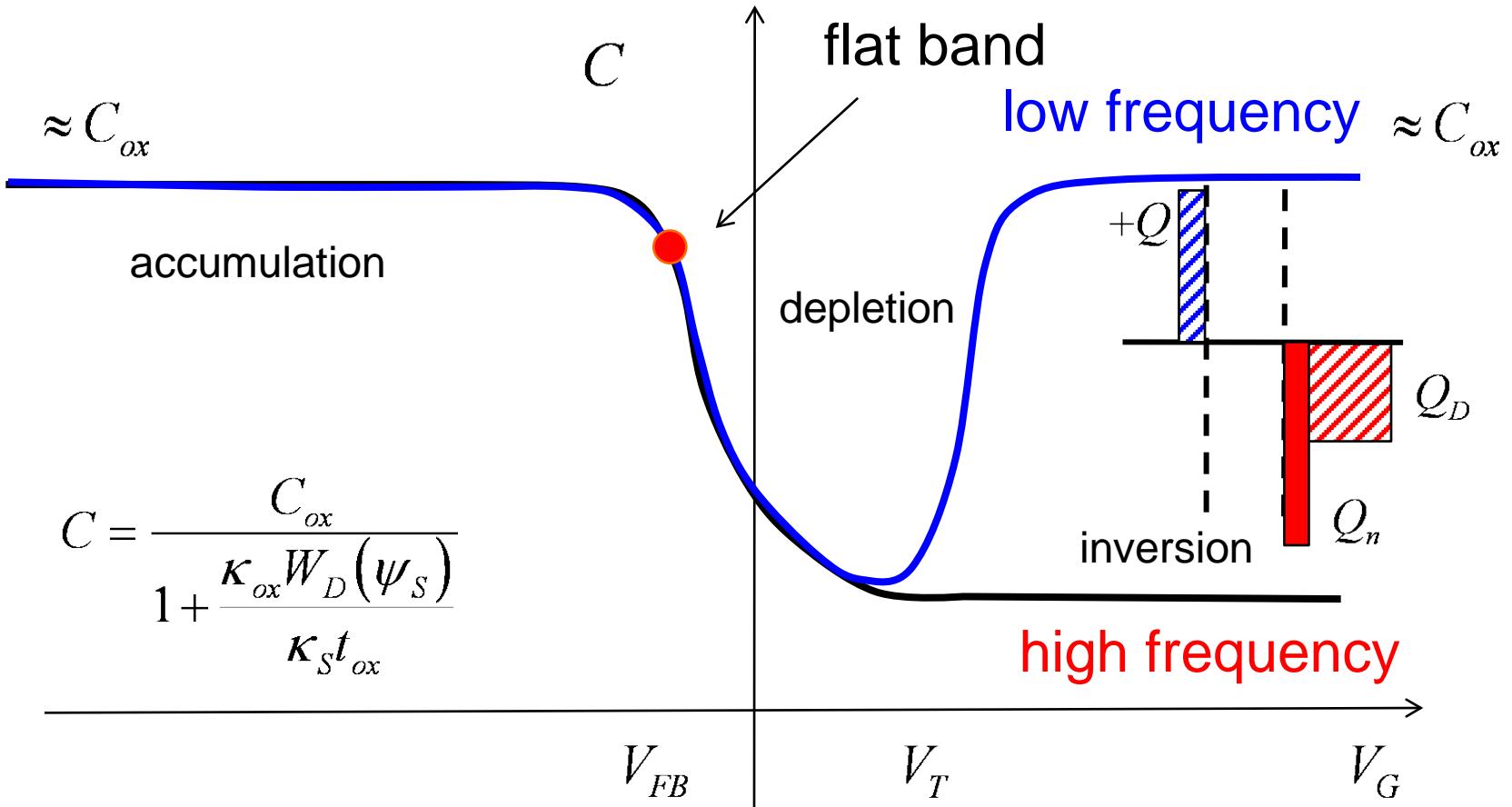
s.s. gate capacitance vs. d.c. gate bias



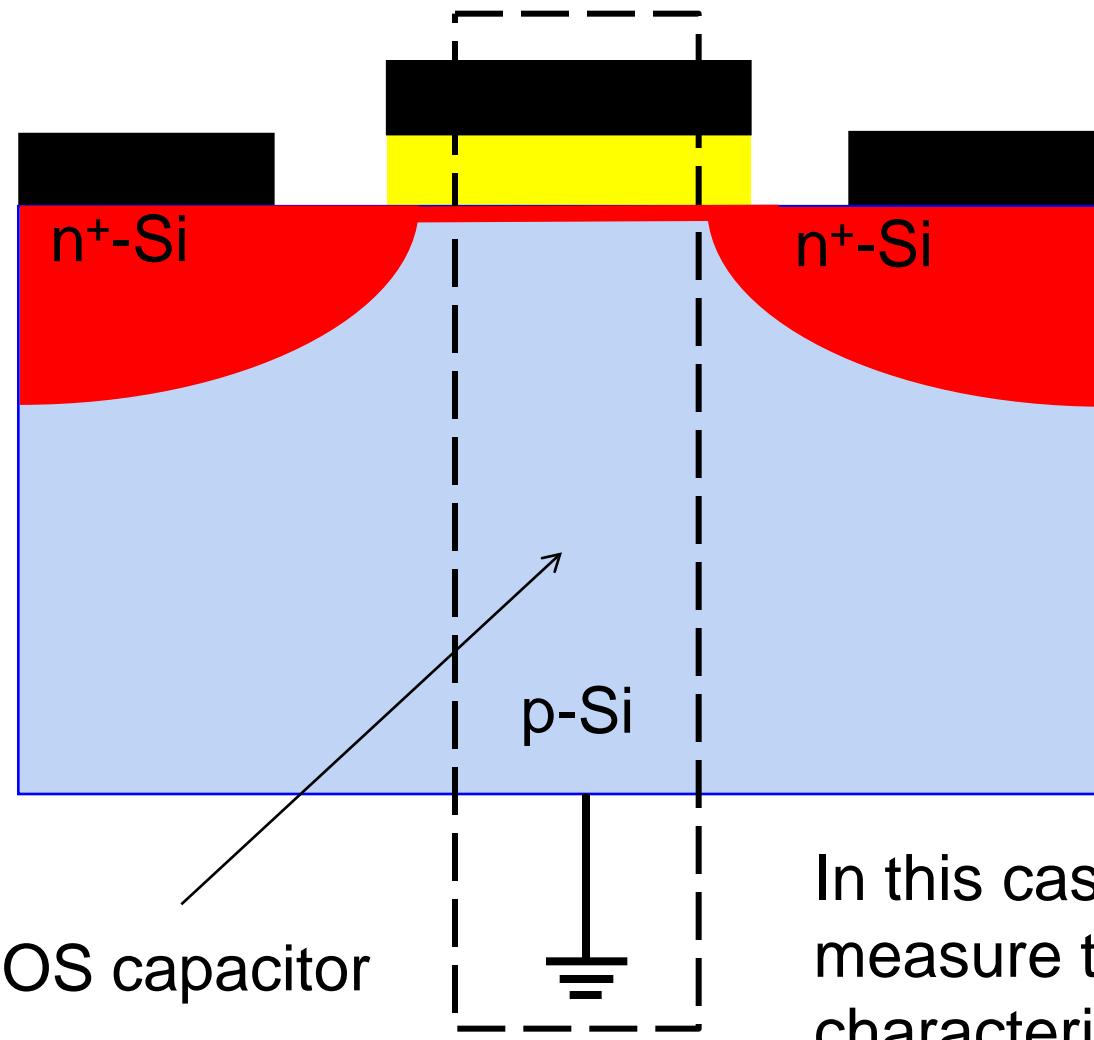
Realistic MOS CV / Critical voltages



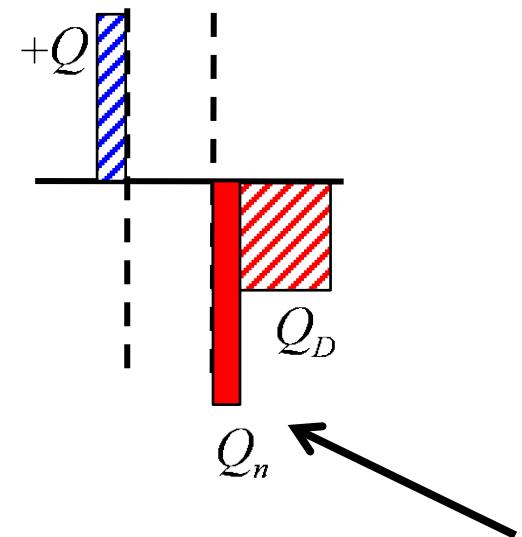
High frequency vs. low frequency CV



HF vs. LF CV



inversion



In this case, we will always measure the low frequency characteristic.

CV measurements as an analysis tool

1) accumulation

$$C = C_{ox} = \frac{K_{ox} \epsilon_0}{t_{ox}}$$

2) flat band

depletion

$$C = \frac{C_{ox}}{1 + \frac{\kappa_{ox} W_D(2\psi_B)}{\kappa_s t_{ox}}}$$

3) inversion

$$V_{FB} = \phi_{ms} - \frac{qN_F}{C_{ox}} V_T$$

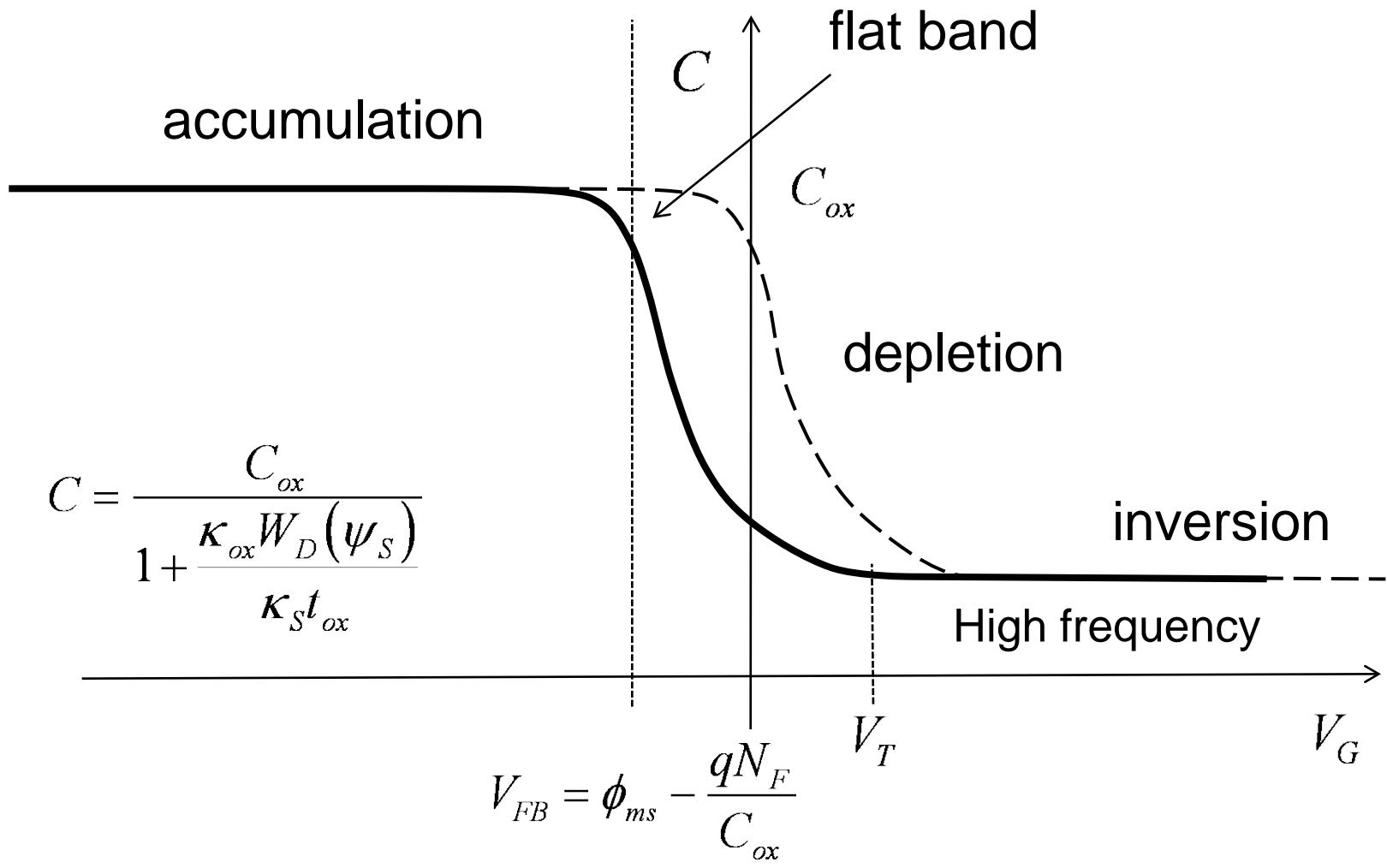
V_G

CV measurements as an analysis tool

- Oxide thickness
- Flatband voltage
- Doping density

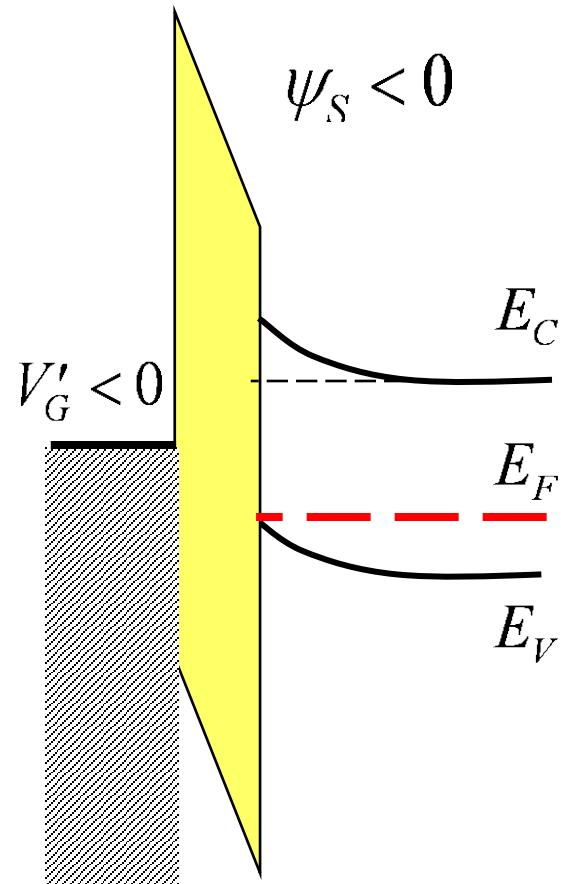
More advanced techniques can also probe various types of charges in the oxide and at the oxide/semiconductor interface.

Summary (i)

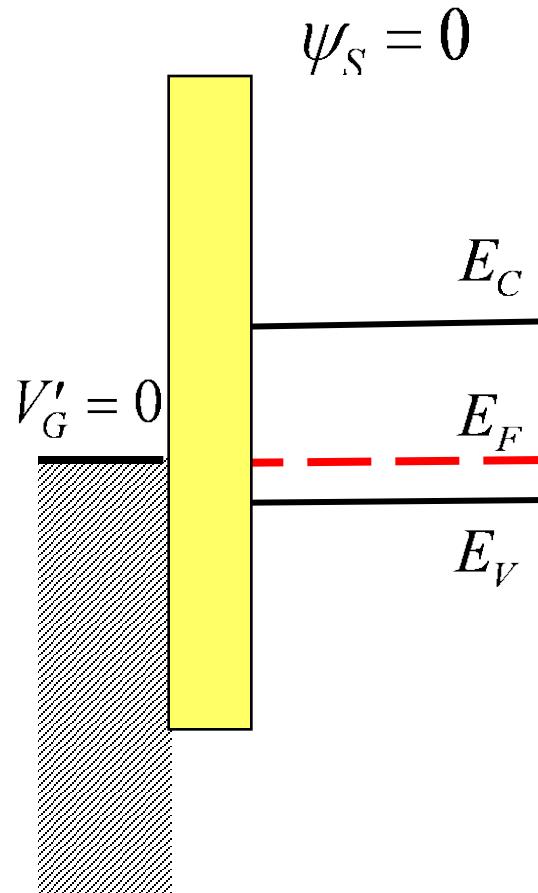


Summary (ii)

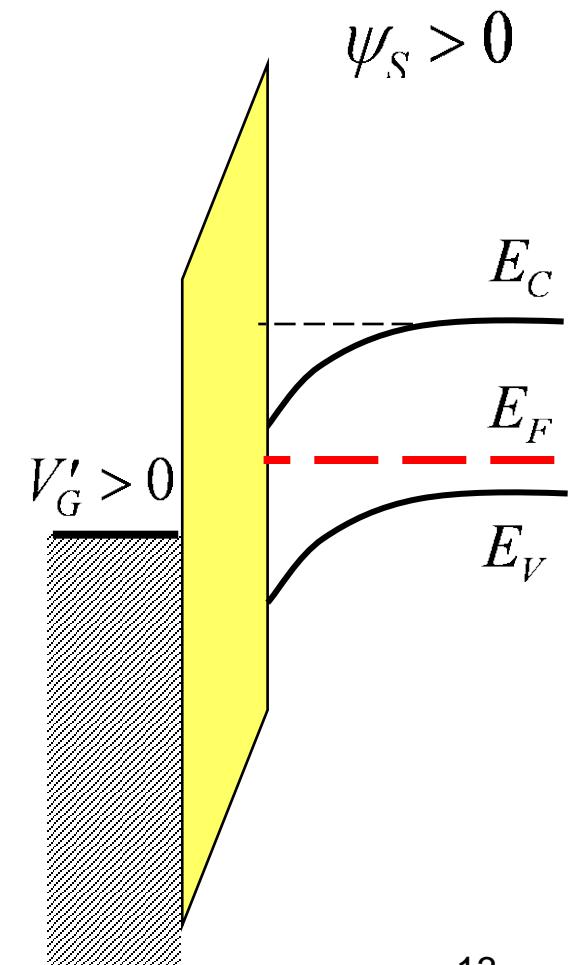
accumulation



flat band

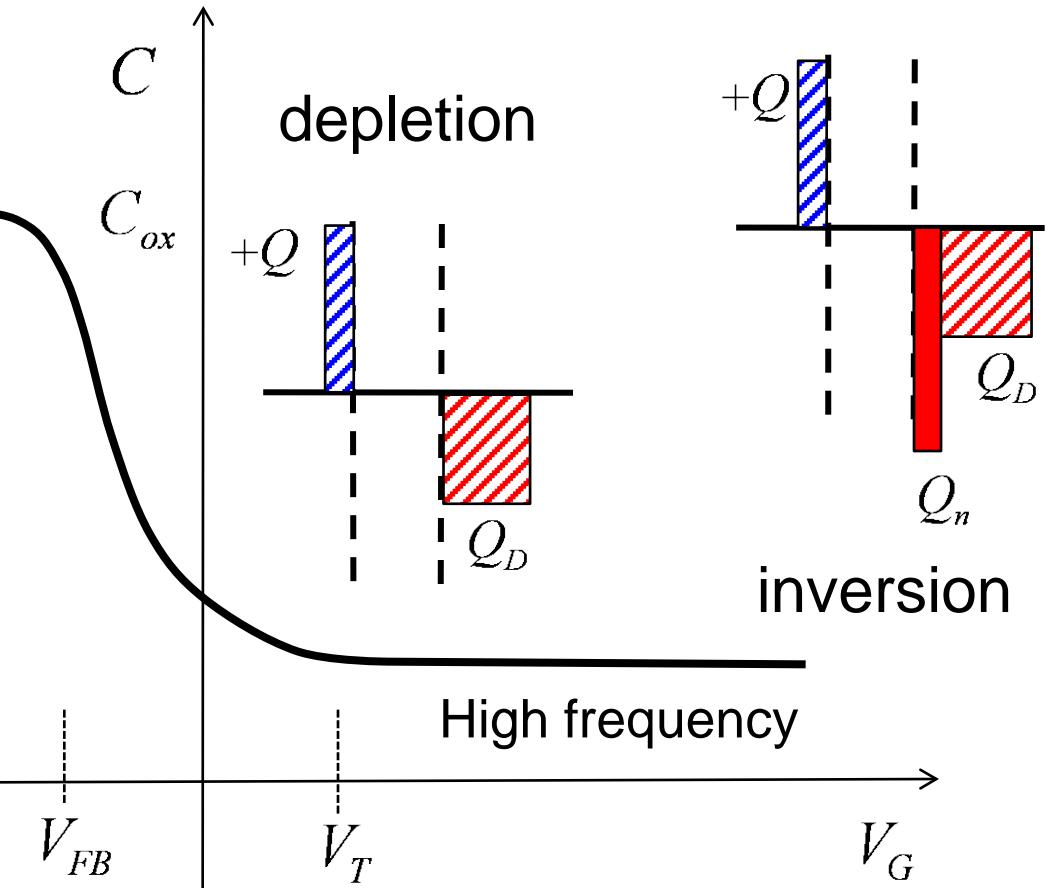
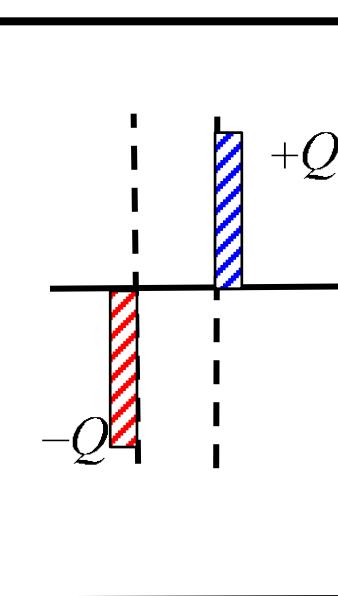


depletion/
inversion



Summary (iii)

accumulation



Next topic

The mobile charge (the electron or holes in the inversion layer) carries the current in a device.

Our goal in the next lecture is to understand how the mobile charge varies with surface potential and gate voltage.

$$Q_n(\psi_s) \qquad Q_n(V_G)$$

Essentials of MOSFETs

Unit 3: MOS Electrostatics

Lecture 3.6: The Mobile Charge vs. Surface Potential

Mark Lundstrom

lundstro@purdue.edu

Electrical and Computer Engineering
Purdue University
West Lafayette, Indiana USA

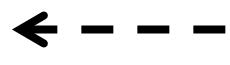
MOSFET drain current

$$I_{DS}/W = -Q_n(V_{GS}) \langle v_x(V_{DS}) \rangle$$

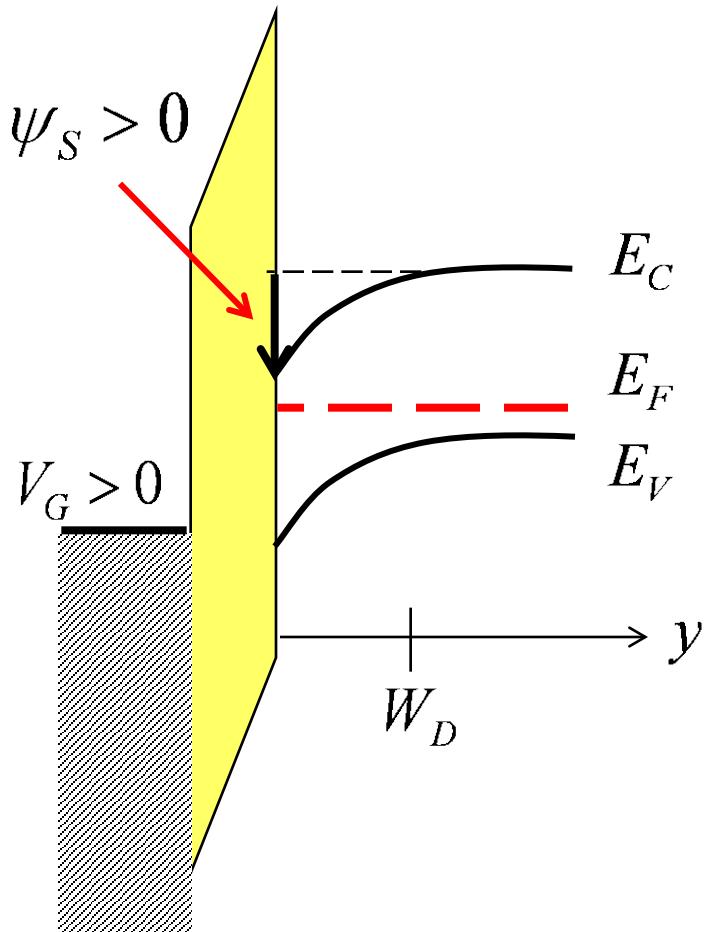
We have been discussing Q_S and Q_D , but we need \mathbf{Q}_n as a function of **surface potential** and **gate voltage**.

$$Q_S = Q_D + Q_n \text{ C/cm}^2$$

$Q_n(\psi_S)$  **this lecture**

$Q_n(V_G)$  next lecture

Mobile charge (per cm³) vs.depth

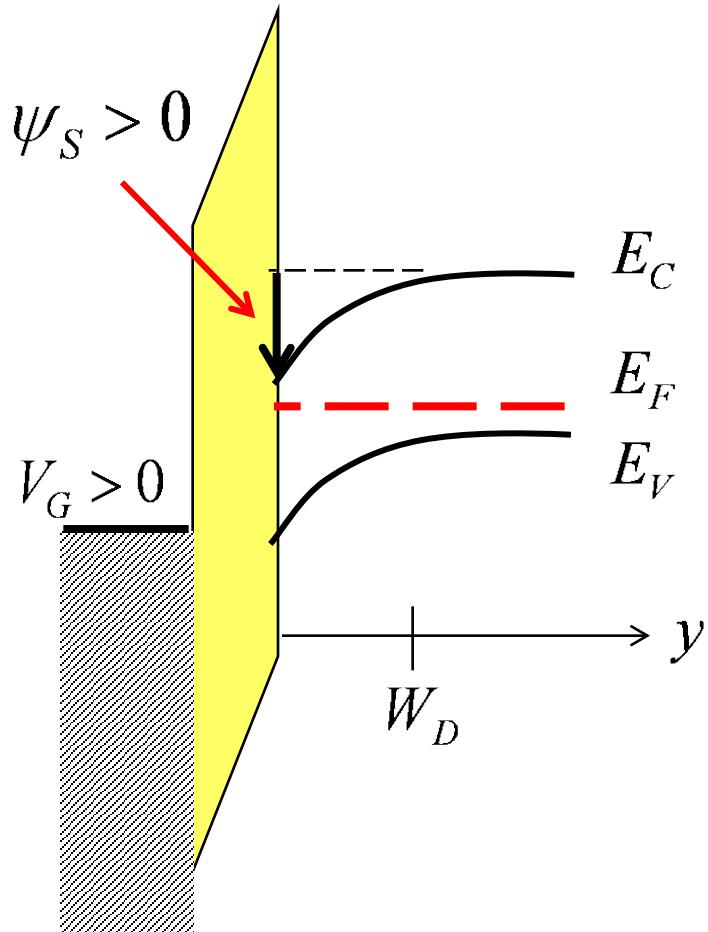


$$n_0(y) = N_C e^{(E_F - E_C(y))/k_B T} \text{ cm}^{-3}$$

$$n_0(y) = n_B \times e^{q\psi(y)/k_B T}$$

$$n_B = \frac{n_i^2}{N_A}$$

Mobile sheet charge (per cm²)

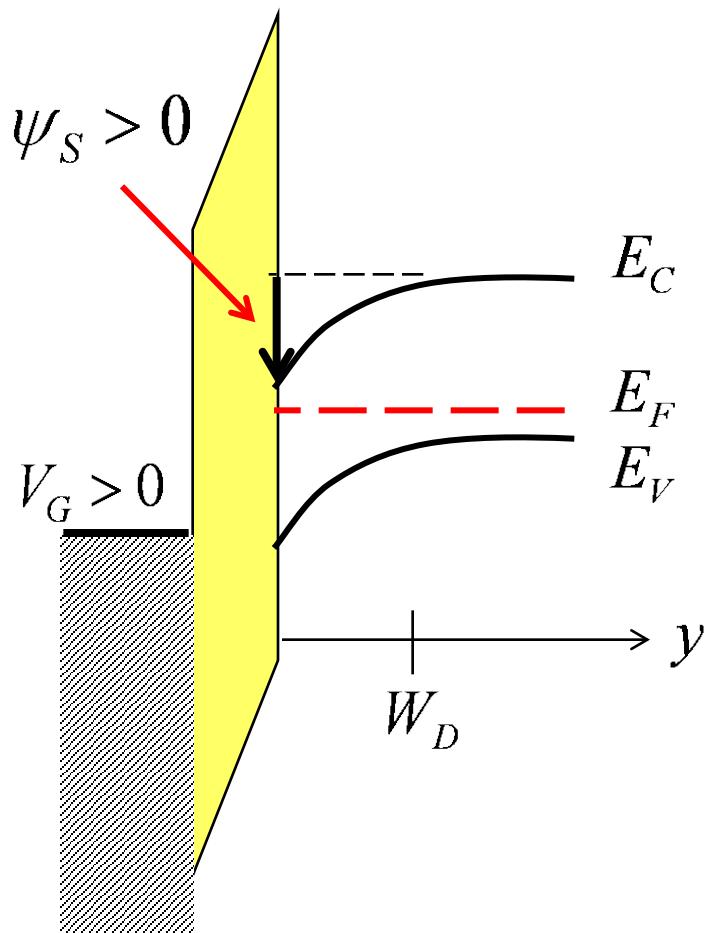


Lundstrom: 2018

$$\begin{aligned} Q_n &= -q \int_0^{\infty} n(y) dy \\ &= -q \int_0^{\infty} n_B e^{q\psi(y)/k_B T} dy \\ &= -qn_B \int_{\psi_S}^0 e^{q\psi(y)/k_B T} \frac{dy}{d\psi} d\psi \\ &\approx \frac{qn_B}{\mathcal{E}_S} \int_{\psi_S}^0 e^{q\psi(y)/k_B T} d\psi \\ Q_n &\approx -qn_B e^{q\psi_S/k_B T} \left(\frac{k_B T / q}{\mathcal{E}_S} \right) \end{aligned}$$

$$Q_n \approx -qn(0) \times t_{inv}$$

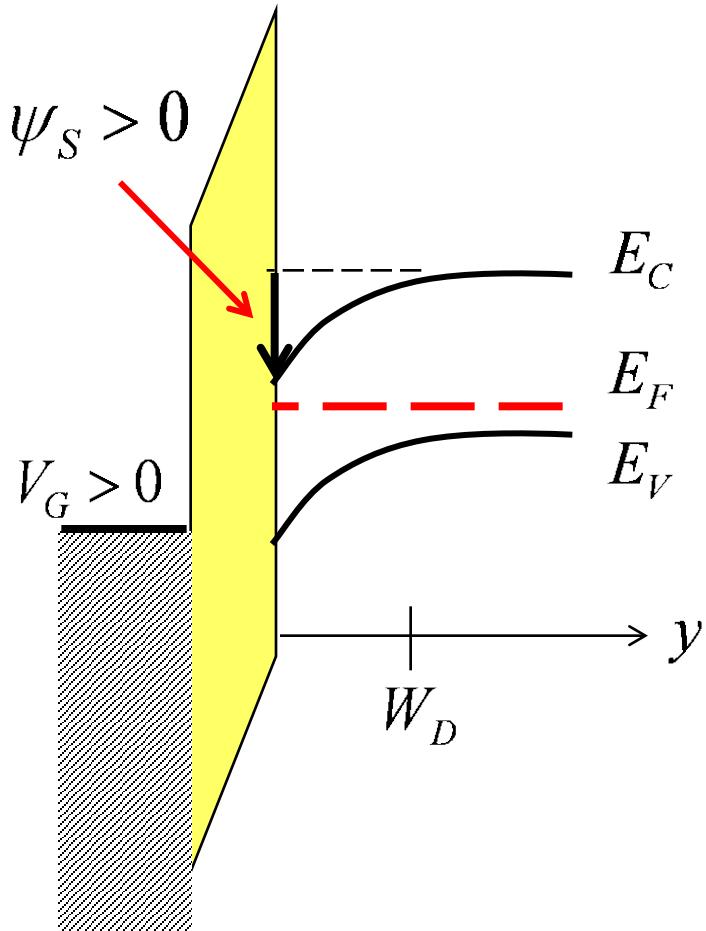
Mobile sheet charge (per cm²)



$$Q_n(\psi_s) \approx -qn_B e^{q\psi_s/k_B T} \left(\frac{k_B T/q}{\mathcal{E}_s} \right)$$

valid above **or** below threshold
(if we are careful)

Mobile sheet charge: below threshold



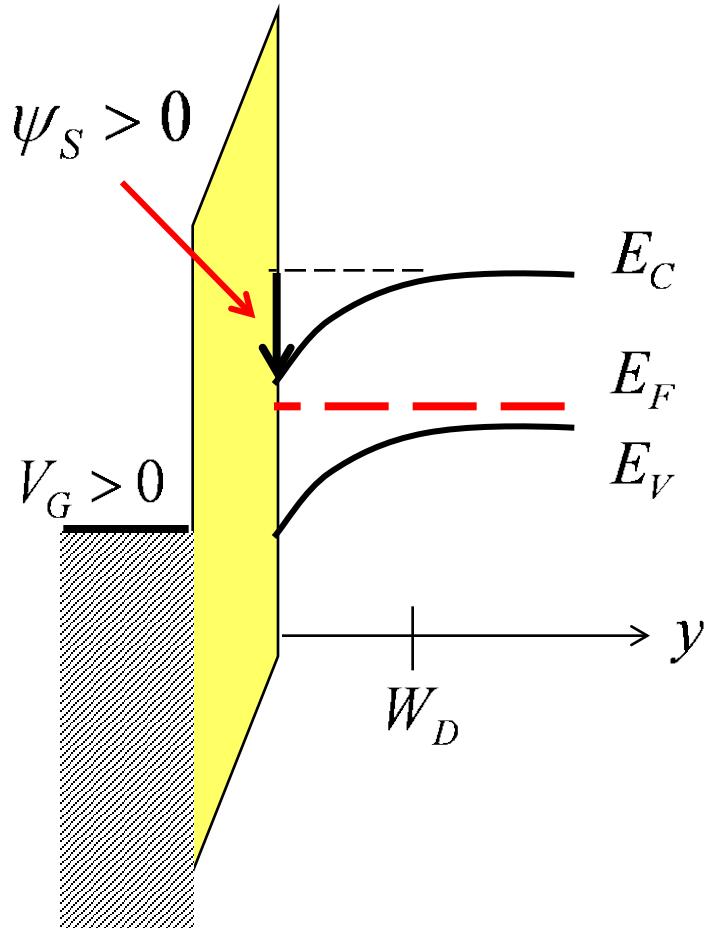
$$Q_n(\psi_S) \approx -q n_B e^{q\psi_S/k_B T} \left(\frac{k_B T / q}{\mathcal{E}_S} \right)$$

$$\mathcal{E}_S = (2qN_A\psi_S/\varepsilon_S)^{1/2}$$

$$Q_n(\psi_S) \approx -\frac{n_i^2 k_B T / N_A}{(2qN_A\psi_S/\varepsilon_S)^{1/2}} e^{q\psi_S/k_B T}$$

$$Q_n(\psi_S) \propto e^{q\psi_S/k_B T}$$

Mobile sheet charge: **above threshold**



$$Q_n(\psi_S) \approx -qn_B e^{q\psi_S/k_B T} \left(\frac{k_B T/q}{\mathcal{E}_S} \right)$$

$$\varepsilon_s \mathcal{E}_S = -Q_S(\psi_S) = -Q_D(\psi_S) - Q_n(\psi_S)$$

1) Below threshold:

$$Q_S(\psi_S) \approx Q_D(\psi_S)$$

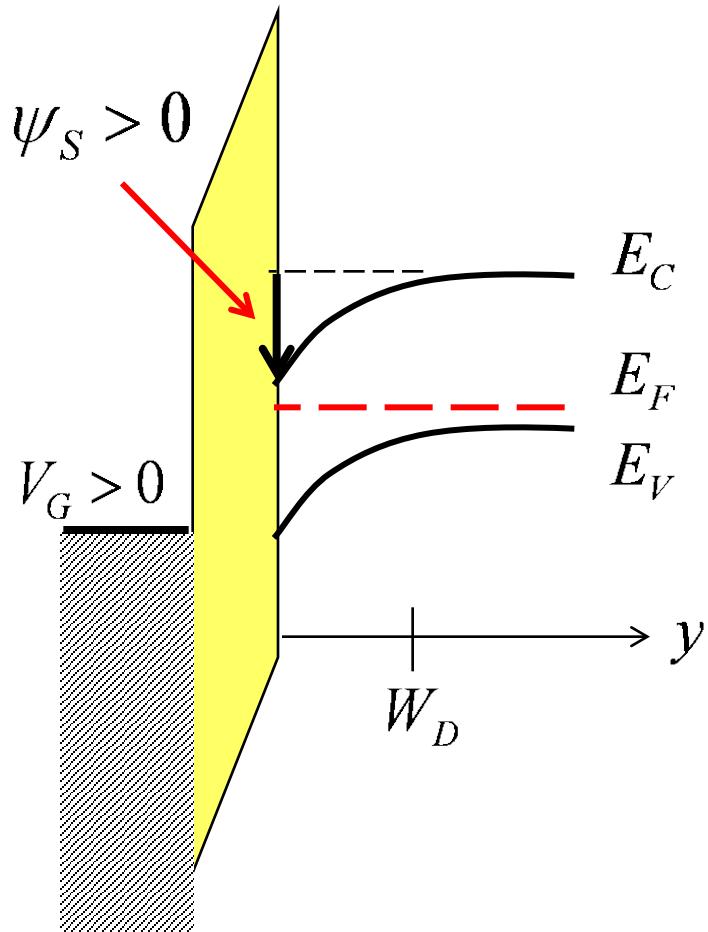
$$\mathcal{E}_S = (2qN_A\psi_S/\varepsilon_s)^{1/2}$$

2) Above threshold:

$$Q_S(\psi_S) \approx Q_n(\psi_S)$$

$$\varepsilon_s \mathcal{E}_S \approx -Q_n(\psi_S)$$

Mobile sheet charge: above threshold



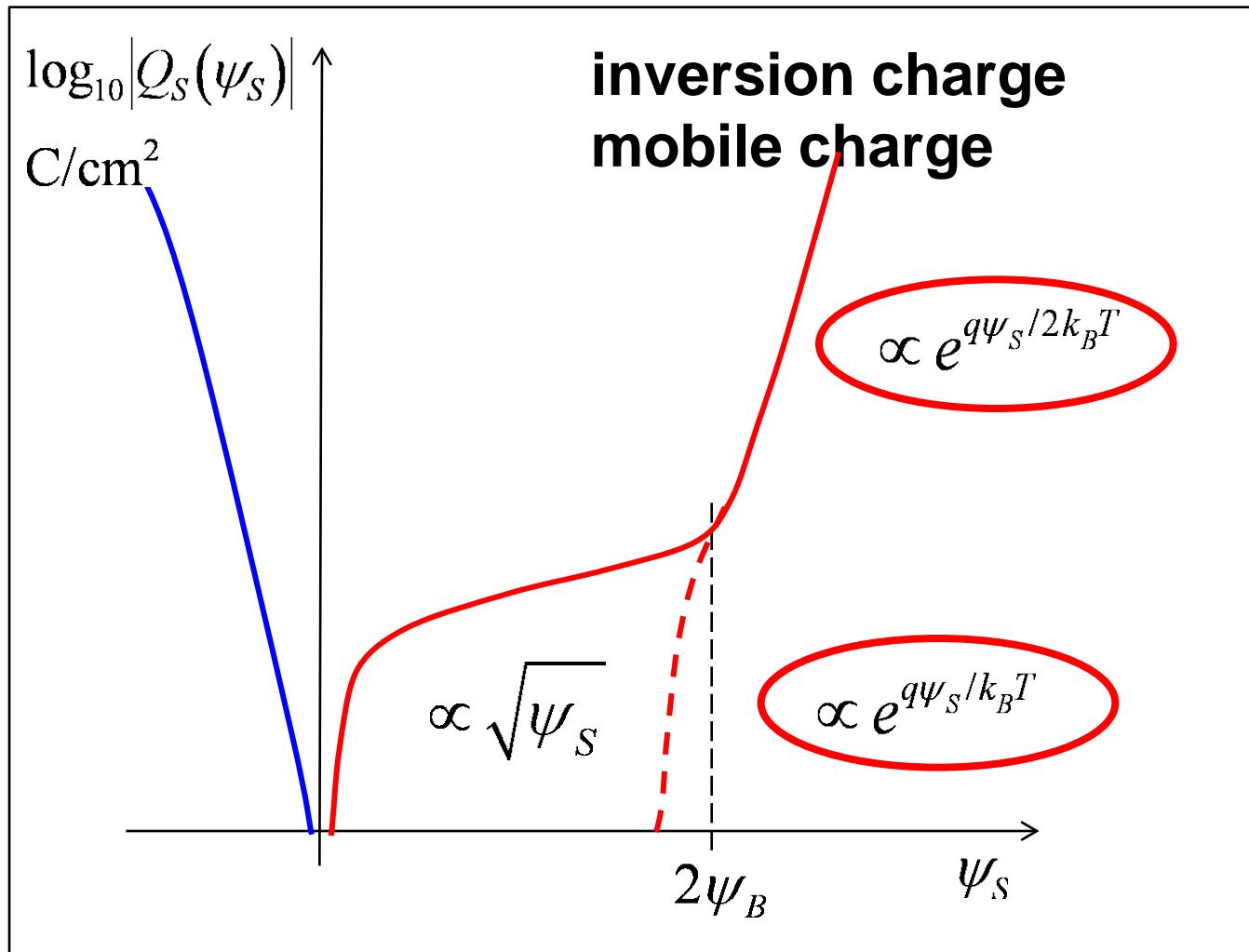
$$Q_n(\psi_S) \approx -qn_B e^{q\psi_S/k_B T} \left(\frac{k_B T / q}{\bar{\mathcal{E}}_S} \right)$$

$$\mathcal{E}_S \rightarrow \bar{\mathcal{E}}_S \quad \bar{\mathcal{E}}_S = \frac{Q_n(\psi_S)}{2\mathcal{E}_S}$$

$$Q_n(\psi_S) = -\sqrt{2\mathcal{E}_S k_B T \left(n_i^2 / N_A \right)} \times e^{q\psi_S / 2k_B T}$$

$$Q_n(\psi_S) \propto e^{q\psi_S / 2k_B T}$$

Mobile charge vs. surface potential



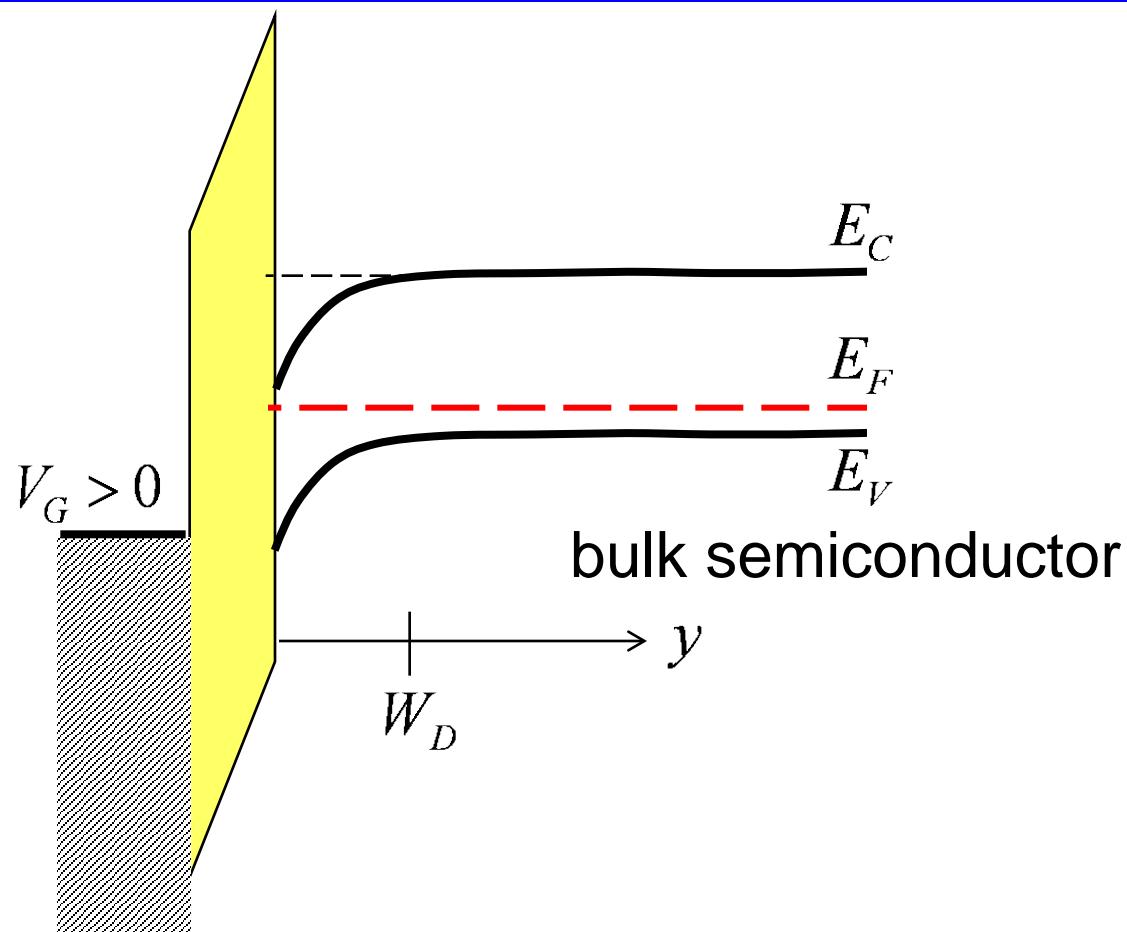
Exact solution

It is possible to solve the problem so that we go smoothly from subthreshold to above threshold.

The exact solution involves solving **the Poisson-Boltzmann equation** as discussed in these notes:

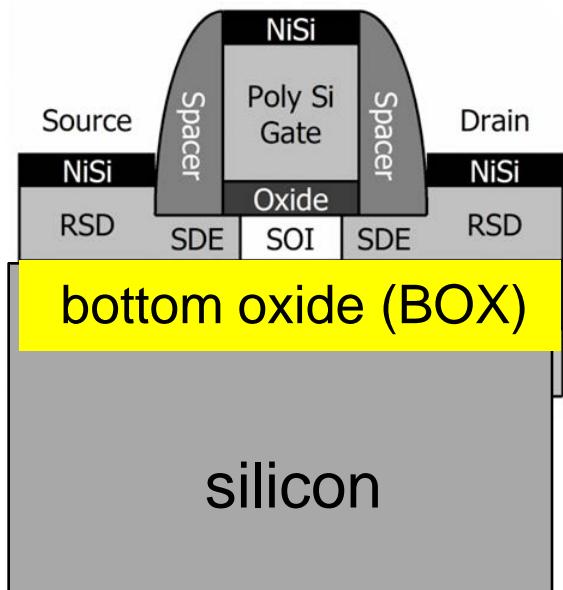
<https://nanohub.org/resources/5338>

Bulk MOS-C

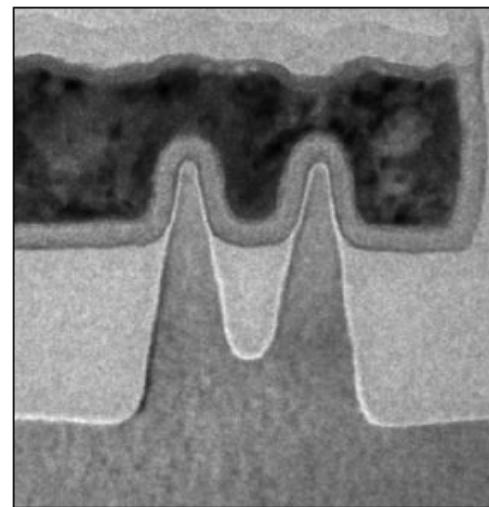


12" wafers are 775 micrometers thick

Fully depleted ultra thin body (UTB) MOS structures

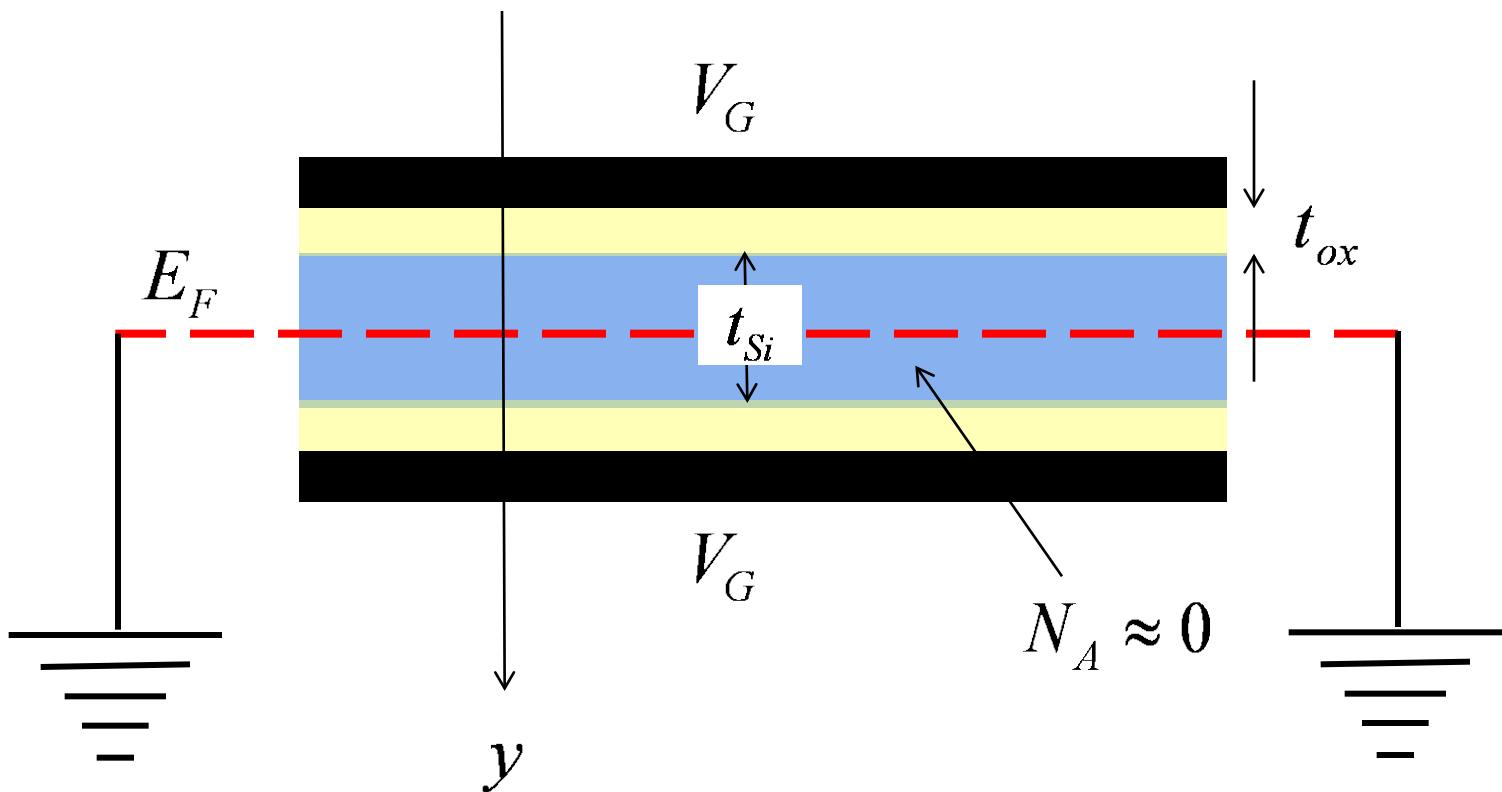


(ETSOI: Source: IBM, 2009)



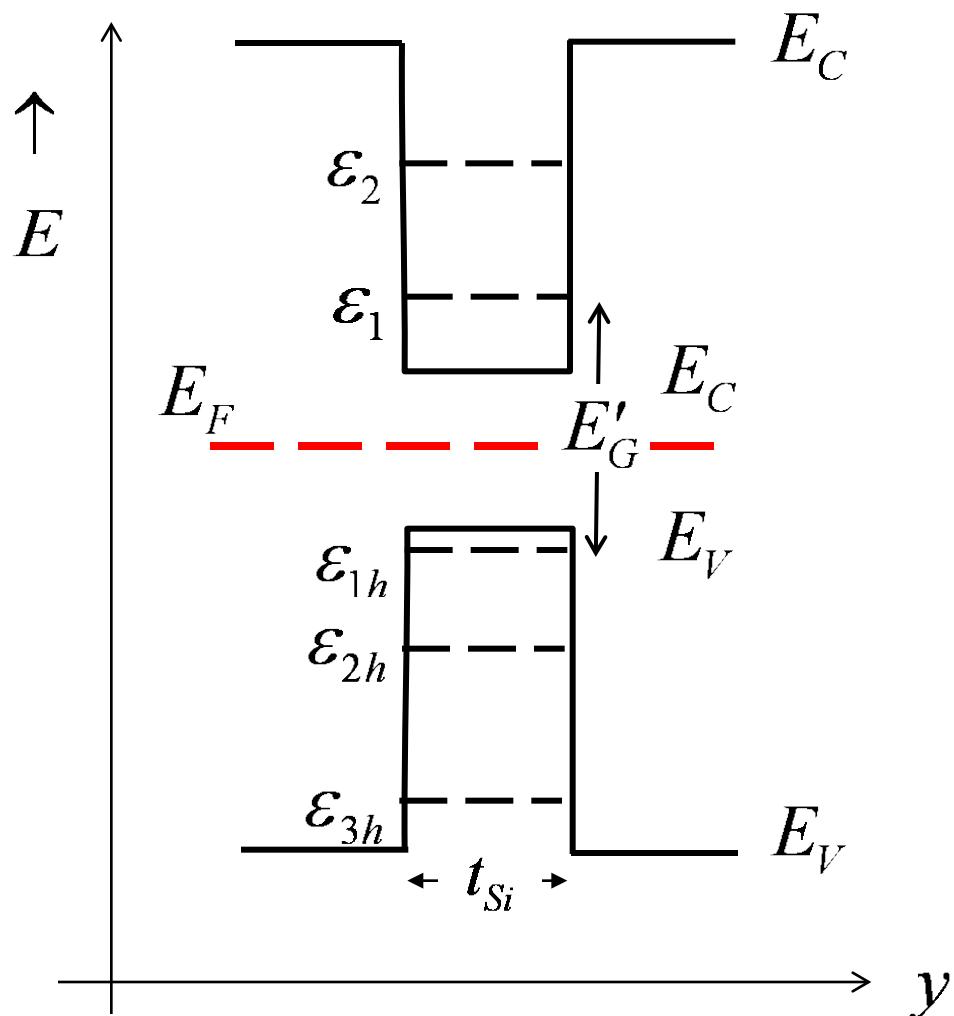
(FinFET: Source: Intel, 2015)

FD UTB double gate-C



We will assume a symmetrical, double gate geometry, which makes this discussion relevant to FinFETs as well.

FD UTB energy band diagram

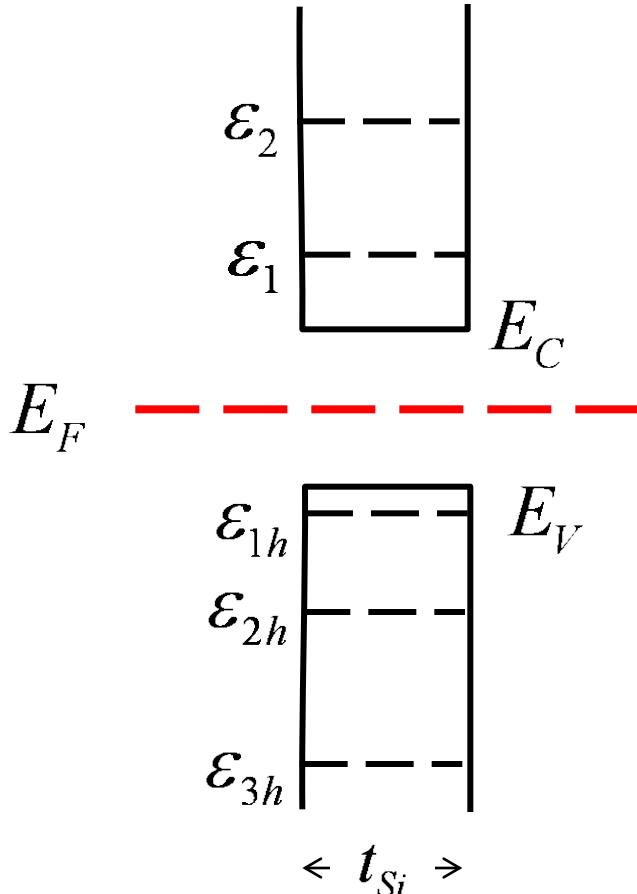


$$\varepsilon_n = \frac{\hbar^2 n^2 \pi^2}{2m^* t_{Si}^2}$$

$$E'_G = E_G + \varepsilon_1 + \varepsilon_{1h}$$

(Neglect band bending, so the potential is constant.)

2D carrier densities



$$n_{S1} = N_C^{2D} \mathcal{F}_0(\eta_{F1}) \text{ cm}^{-2}$$

$$N_C^{2D} = g_V \frac{m_n^* k_B T}{\pi \hbar^2} \text{ cm}^{-2}$$

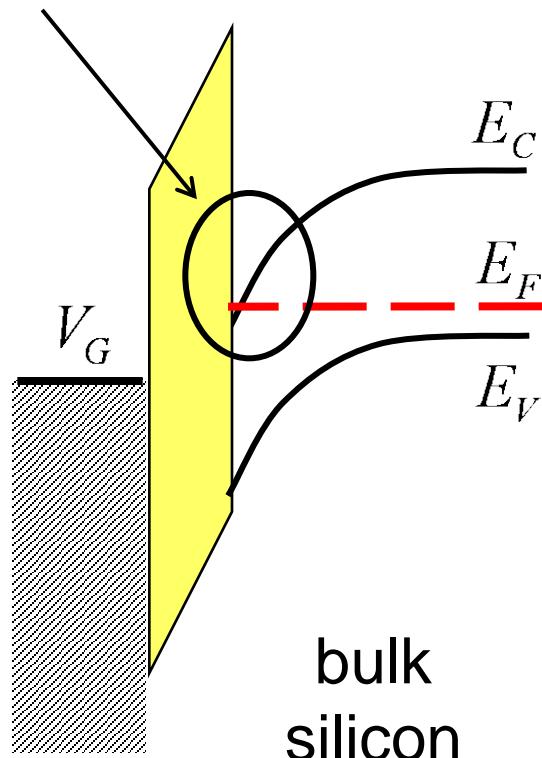
$$\eta_{F1} = \frac{(E_F - E_C - \varepsilon_1)}{k_B T}$$

Boltzmann statistics:

$$n_{S1} = N_C^{2D} e^{(E_F - E_C - \varepsilon_1)/k_B T} \text{ cm}^{-2}$$

Quantum confinement in a bulk MOS-C

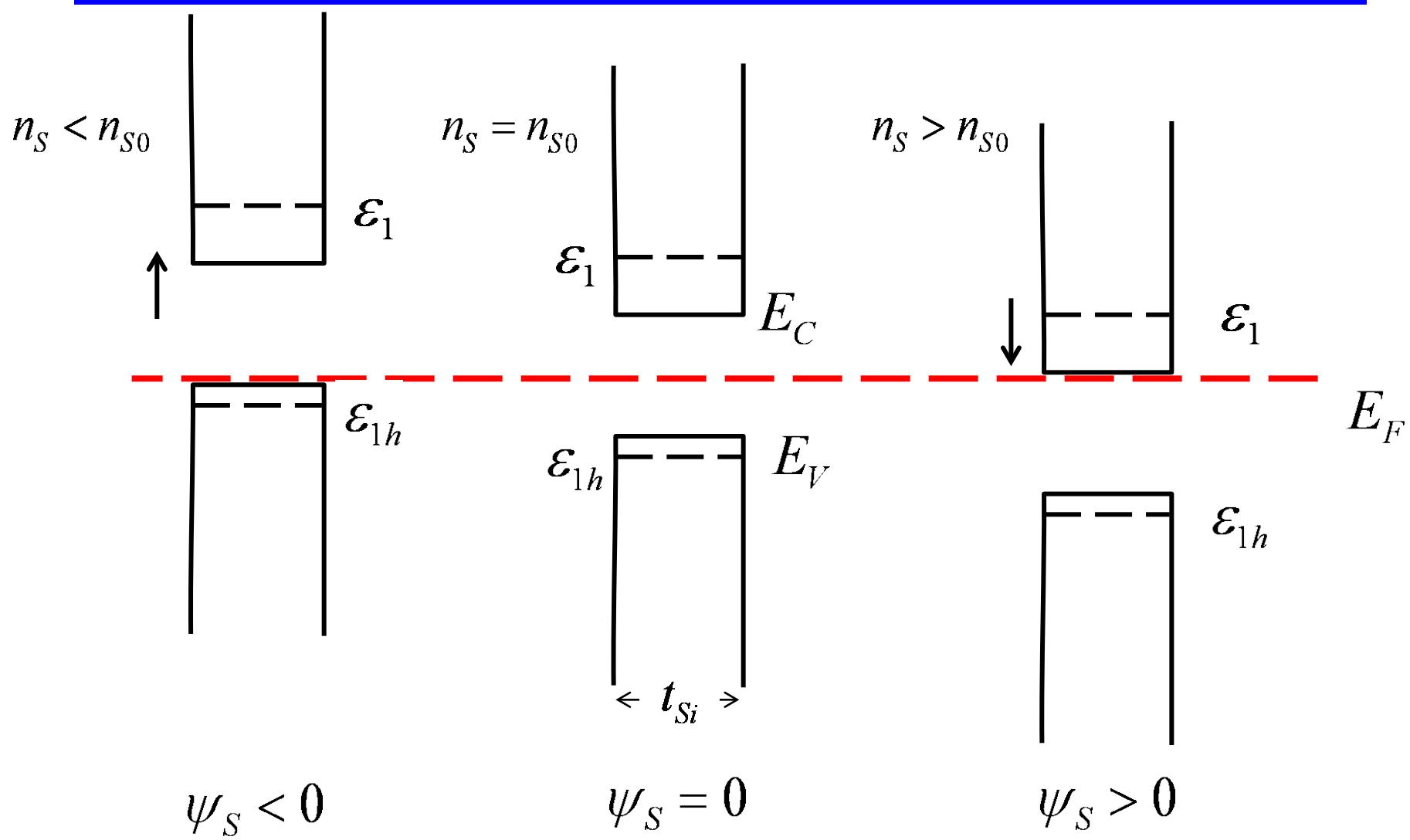
“quantum well”



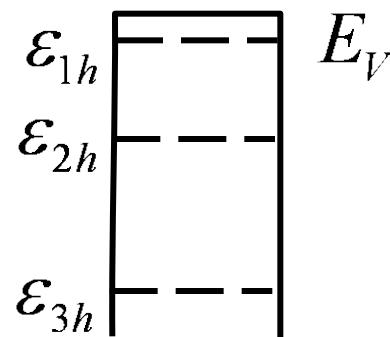
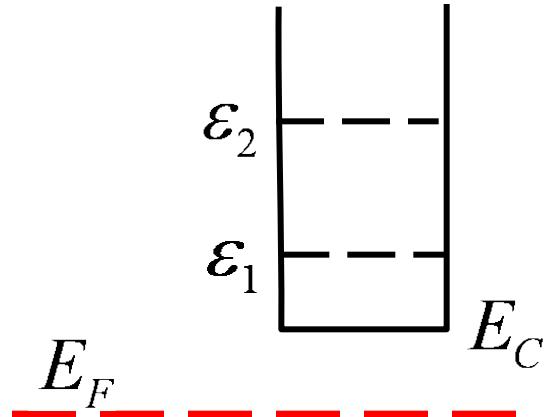
In the bulk, the confining potential is due to electrostatics.

In fully depleted ultra thin body structures, the confining potential is due to the physical structure.

FD UTB for various potentials



Carrier densities and semiconductor potential



$\leftarrow t_{Si} \rightarrow$

$$n_{S1} = N_C^{2D} e^{(E_F - E_C - \epsilon_1)/k_B T} \text{ cm}^{-2}$$

$$E_C = E_{C0} - q\psi_s$$

$$n_S = n_{S0} e^{q\psi_s/k_B T}$$

$$p_S = p_{S0} e^{-q\psi_s/k_B T}$$

(These eqns. assume that
only 1 subband is occupied.)

Mobile sheet charge (per cm²)

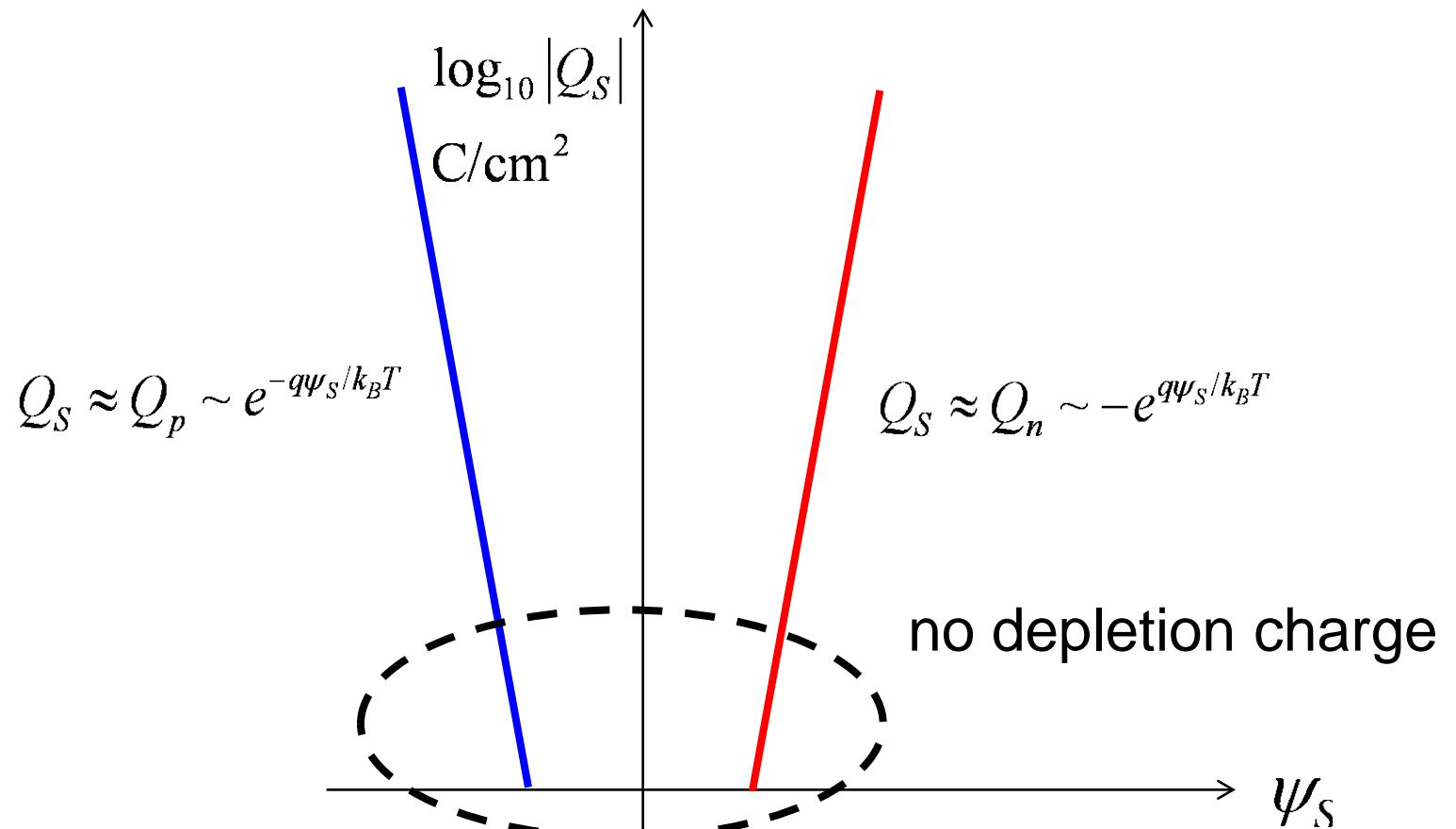
$$Q_S = q(p_s - n_s) \quad \text{C/cm}^2$$

$$Q_n(\psi_s) = -qn_{s0} e^{q\psi_s/k_B T}$$

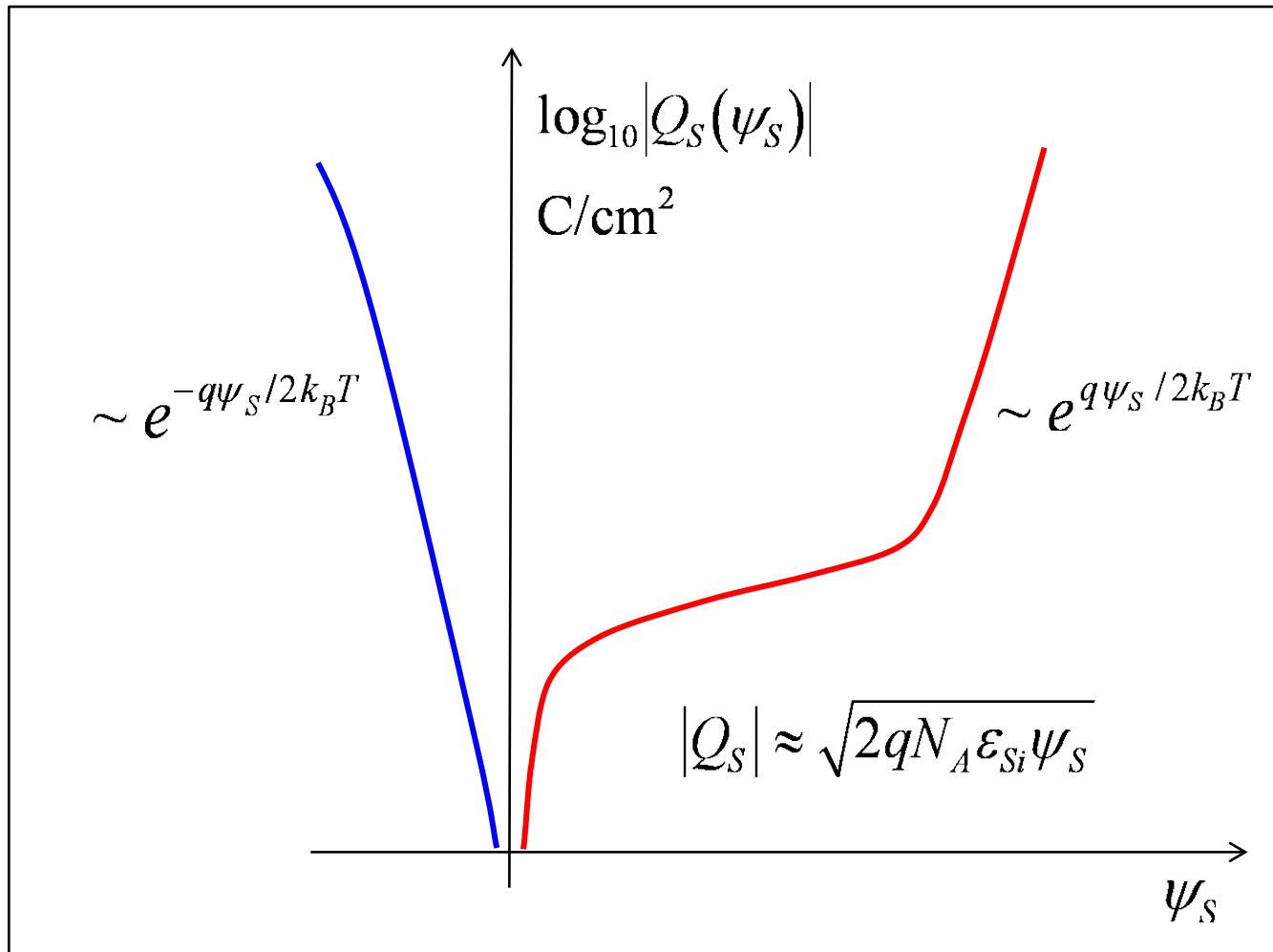
Valid above and below threshold.

Charge vs. surface potential

$$Q_s(\psi_s) = q(p_{s0} e^{-q\psi_s/k_B T} - n_{s0} e^{q\psi_s/k_B T})$$



Recall: $Q_s(\psi_s)$ for bulk MOS



Summary

Bulk semiconductor:

$$\psi_s < 2\psi_B : \quad Q_n(\psi_s) \approx - \left(\frac{n_i^2 k_B T / N_A}{\sqrt{(2qN_A \psi_s / \epsilon_s)^{1/2}}} \right) e^{q\psi_s / k_B T}$$

$$\psi_s > 2\psi_B : \quad Q_n(\psi_s) = - \sqrt{2\epsilon_s k_B T (n_i^2 / N_A)} \times e^{q\psi_s / 2k_B T}$$

Fully depleted, ultra thin body:

$$\psi_s > 0 : \quad Q_n(\psi_s) = -qn_{S0} e^{q\psi_s / k_B T}$$

Next topic

$$I_{DS}/W = -Q_n(V_{GS}) \langle v_x(V_{DS}) \rangle$$

We have been discussing Q_S and Q_D , but we need \mathbf{Q}_n as a function of **surface potential** and **gate voltage**.

$$Q_S = Q_D + Q_n \text{ C/cm}^2$$

$Q_n(\psi_S)$ ← this lecture ✓

$Q_n(V_G)$ ← -- next lecture

Essentials of MOSFETs

Unit 3: MOS Electrostatics

Lecture 3.7: The Mobile Charge vs. Gate Voltage

Mark Lundstrom

lundstro@purdue.edu

Electrical and Computer Engineering
Purdue University
West Lafayette, Indiana USA

MOSFET drain current

$$I_{DS}/W = -Q_n(V_{GS}) \langle v_x(V_{DS}) \rangle$$

We have been discussing Q_S and Q_D , but we need \mathbf{Q}_n as a function of **surface potential** and **gate voltage**.

$$Q_S = Q_D + Q_n \text{ C/cm}^2$$

$Q_n(\psi_S)$ ← --- last lecture

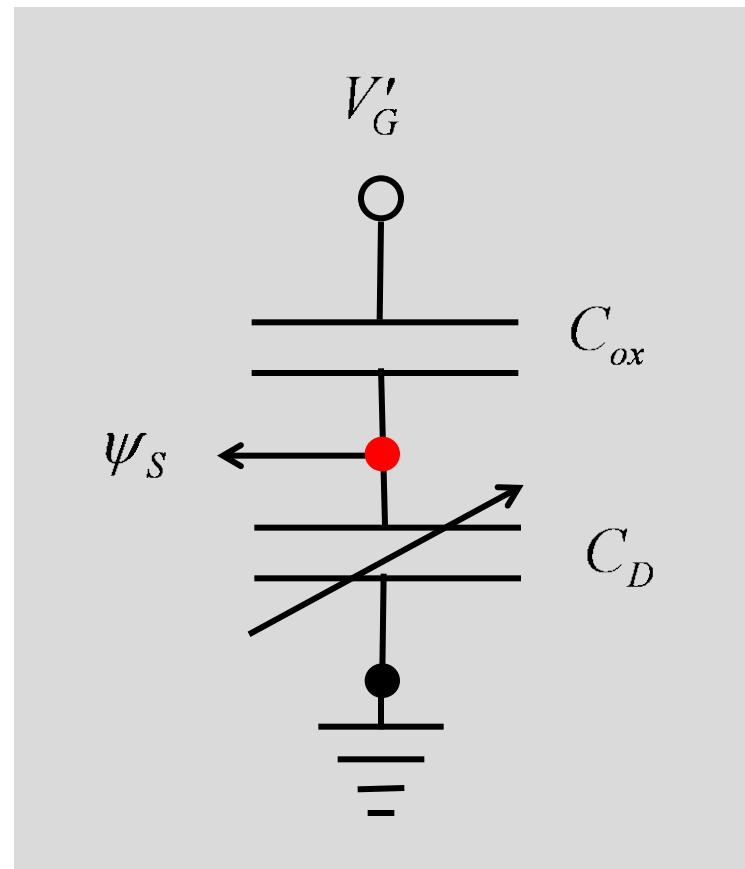
$Q_n(V_G)$  this lecture

1) Subthreshold charge vs. **gate voltage**

$$V'_G = V_G - V_{FB} = -\frac{Q_S(\psi_S)}{C_{ox}} + \psi_S$$

$$\psi_S = \frac{V'_G}{m}$$

$$m = 1 + C_D / C_{ox}$$



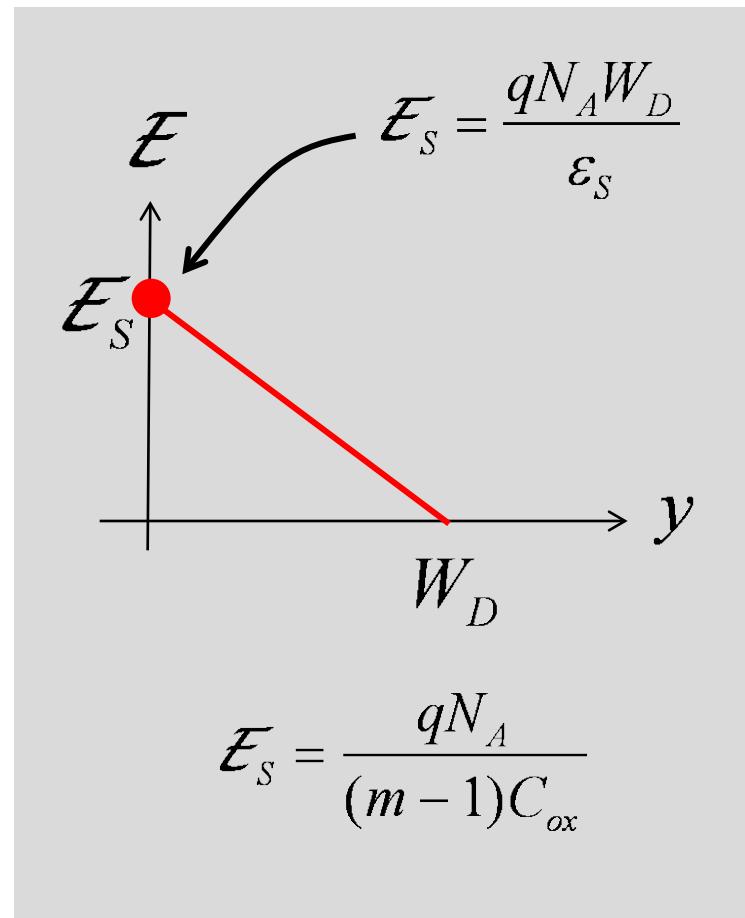
Mobile charge vs. gate voltage: Subthreshold

$$C_D = \frac{\epsilon_s}{W_D}$$

$$\mathcal{E}_s = \frac{qN_A W_D}{\epsilon_s} = \frac{qN_A}{C_D}$$

$$m = 1 + C_D / C_{ox}$$

$$C_D = (m - 1)C_{ox}$$



Subthreshold charge vs. gate voltage

$$Q_n(\psi_s) = -qn_B e^{q\psi_s/k_B T} \left(\frac{k_B T / q}{\mathcal{E}_s} \right)$$

$$n_B = n_i^2 / N_A$$

$$\mathcal{E}_s = \frac{qN_A}{(m-1)C_{ox}}$$

$$Q_n(\psi_s) = -(m-1)C_{ox} \left(\frac{k_B T}{q} \right) \left(\frac{n_i}{N_A} \right)^2 e^{q\psi_s/k_B T}$$

Lundstrom: 2018

$$\psi_s = V'_G / m$$

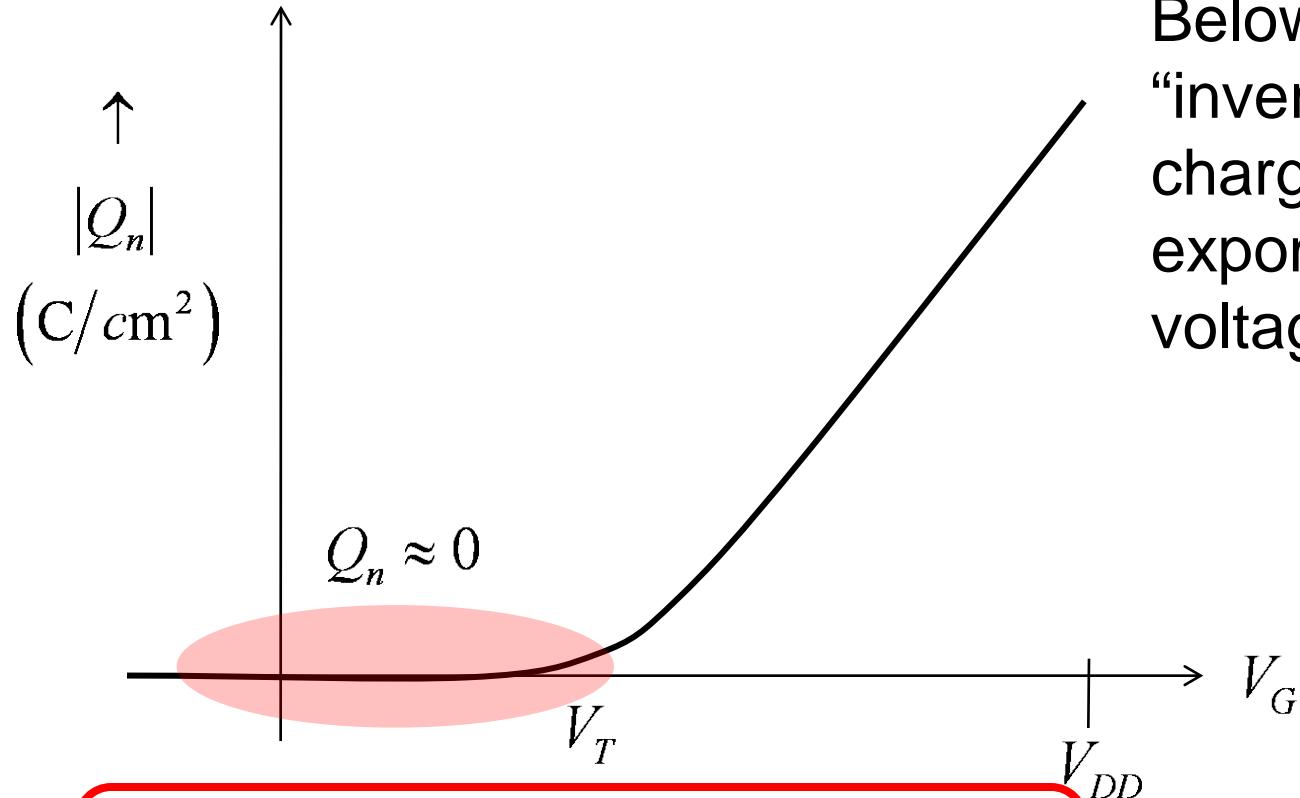
$$\psi_B = \frac{k_B T}{q} \ln \left(\frac{N_A}{n_i} \right)$$

$$\left(\frac{n_i}{N_A} \right)^2 = e^{-2q\psi_B/k_B T}$$

$$2\psi_B = V'_T / m$$

$$\left(\frac{n_i}{N_A} \right)^2 = e^{-qV_T/mk_B T}$$

Subthreshold charge vs. gate voltage



Below threshold, the “inversion” layer charge increases exponentially with gate voltage.

$$Q_n(V_G) = -(m-1)C_{ox} \left(\frac{k_B T}{q} \right) e^{q(V_G - V_T)/mk_B T}$$

2) Mobile charge vs. gate voltage: Above threshold

$$V_G = V_{FB} - \frac{Q_S(\psi_s)}{C_{ox}} + \psi_s$$

$$V_G = V_{FB} - \frac{Q_D(\psi_s)}{C_{ox}} - \frac{Q_n(\psi_s)}{C_{ox}} + \psi_s$$

◆ $V_T = V_{FB} - \frac{Q_D(2\psi_B)}{C_{ox}} + 2\psi_B$

$$V_G = V_{FB} - \frac{Q_D(2\psi_B + \delta\psi_s)}{C_{ox}} - \frac{Q_n(2\psi_B + \delta\psi_s)}{C_{ox}} + 2\psi_B + \delta\psi_s \quad \delta\psi_s \ll 2\psi_B$$

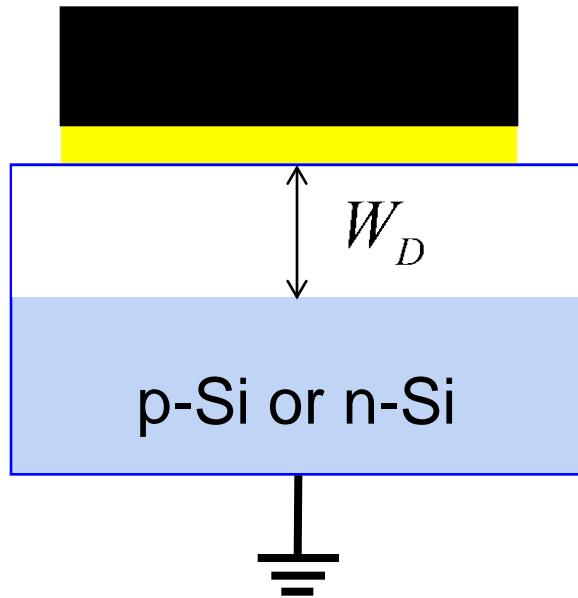
$$V_G \approx V_{FB} - \frac{Q_D(2\psi_B)}{C_{ox}} - \frac{Q_n(2\psi_B + \delta\psi_s)}{C_{ox}} + 2\psi_B$$

◆ ◆ ◆

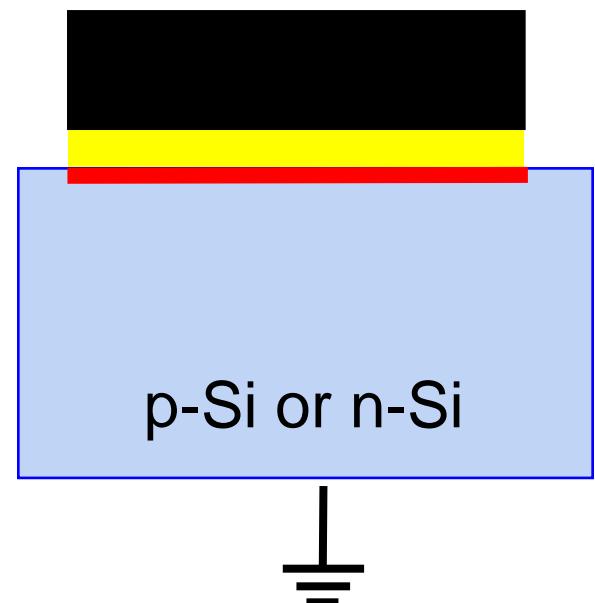
$$Q_n \approx -C_{ox}(V_G - V_T)$$

About the gate capacitance

$$V_G < V_T$$



$$V_G \gg V_T$$



$$\frac{1}{C_G} = \frac{1}{C_{ox}} + \frac{1}{C_S} \quad C_S = C_D = \frac{\epsilon_s}{W_D}$$

$$\frac{1}{C_G} = \frac{1}{C_{ox}} + \frac{1}{C_S} \quad C_S = C_S(\text{inv}) \equiv \frac{\epsilon_s}{t_{inv}}$$

$$C_G(\text{depl}) < C_{ox}$$

$$C_G(\text{inv}) < C_{ox}$$

Semiconductor capacitance

$$\frac{1}{C_G} = \frac{1}{C_{ox}} + \frac{1}{C_S}$$

$$C_S \equiv -\frac{dQ_S}{d\psi_S}$$

Depletion:

$$Q_S \approx Q_D$$

$$C_S = C_D = \frac{\epsilon_S}{W_D}$$

Strong inversion:

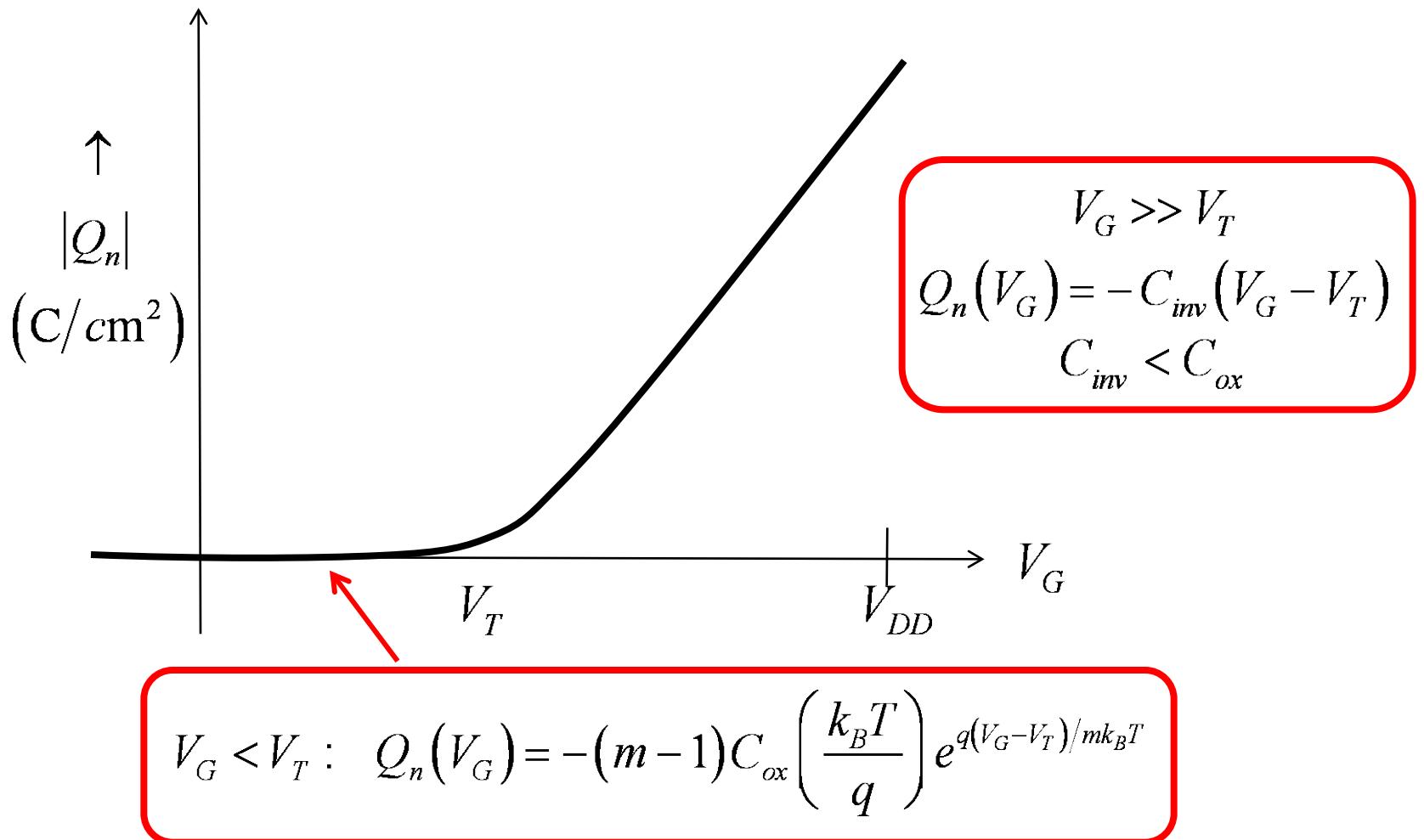
$$Q_S \approx Q_n \propto e^{q\psi_S/2k_B T}$$

$$C_S = \frac{|Q_n|}{2k_B T/q}$$

$$C_S(\text{inv}) \equiv \frac{\epsilon_S}{t_{inv}}$$

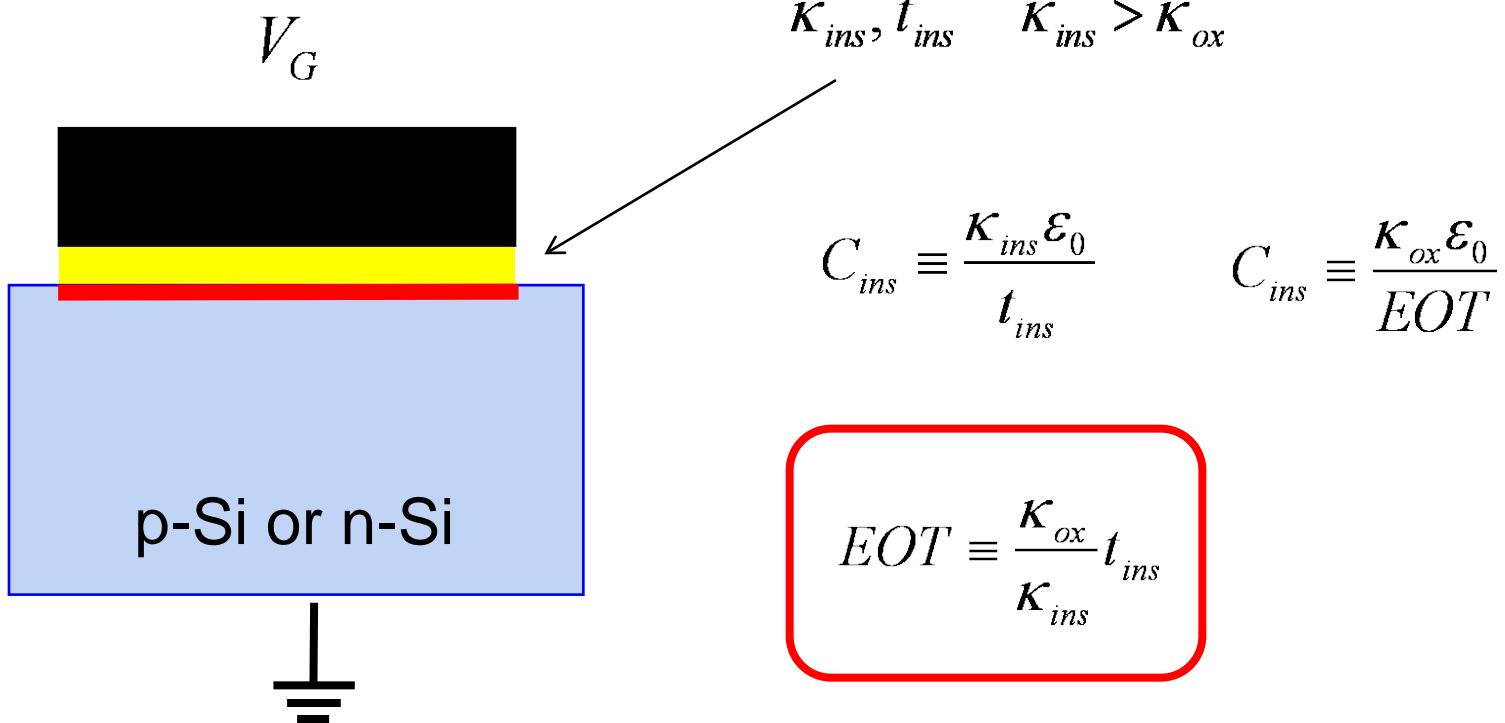
$$t_{inv} \ll W_D$$

Mobile charge vs. gate voltage



Equivalent Oxide Thickness (EOT)

high-k gate insulators



Capacitance Equivalent Thickness (CET)

$$\frac{1}{C_G(\text{inv})} = \frac{1}{C_{ox}} + \frac{1}{C_S(\text{inv})}$$
$$C_G(\text{inv}) = \frac{C_{ox} C_S(\text{inv})}{C_{ox} + C_S(\text{inv})} < C_{ox}$$

EOT

$$C_G(\text{inv}) \equiv \frac{\kappa_{ox} \epsilon_0}{CET}$$

Note: We will frequently refer to the gate capacitance in inversion, $C_G(\text{inv})$, as just C_{inv} . Remember that $C_S(\text{inv})$ is just part of the total gate capacitance in inversion..

Example

On-current conditions: $Q_n = -q \times 10^{13} \text{ C/cm}^2$

Oxide thickness: $t_{ox} = 1 \text{ nm} = 10^{-7} \text{ cm}$ $\kappa_{ox} = 3.9$

$$C_{ox} = \frac{\epsilon_{ox}}{t_{ox}} = 3.45 \times 10^{-6} \text{ F/cm}^2$$

$$C_s(\text{inv}) = \frac{-Q_n}{2(k_B T/q)} = 30.8 \times 10^{-6} \text{ F/cm}^2$$

$$\begin{aligned}\frac{1}{C_G(\text{inv})} &= \frac{1}{C_{ox}} + \frac{1}{C_s(\text{inv})} \\ &= \frac{1}{3.45 \times 10^{-6}} + \frac{1}{30.8 \times 10^{-6}}\end{aligned}$$

$$C_G(\text{inv}) = 3.10 \times 10^{-6} \text{ F/cm}^2$$

$$C_G(\text{inv}) = C_{inv} = 0.90 C_{ox}$$

$$C_G(\text{inv}) \equiv \frac{\kappa_{ox} \epsilon_0}{CET} \quad CET = 1.1 \text{ nm}$$

Why is it hard to bend the bands more than $2\psi_B$?

below threshold:

$$C_S = \frac{d(-Q_D)}{d\psi_S} = C_D < C_{ox}$$

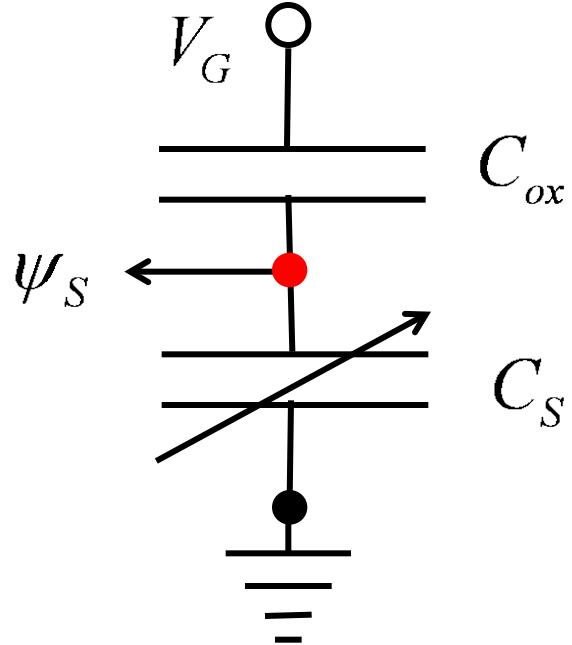
$$m = 1 + C_S / C_{ox} \approx 1 - 1.4$$

above threshold:

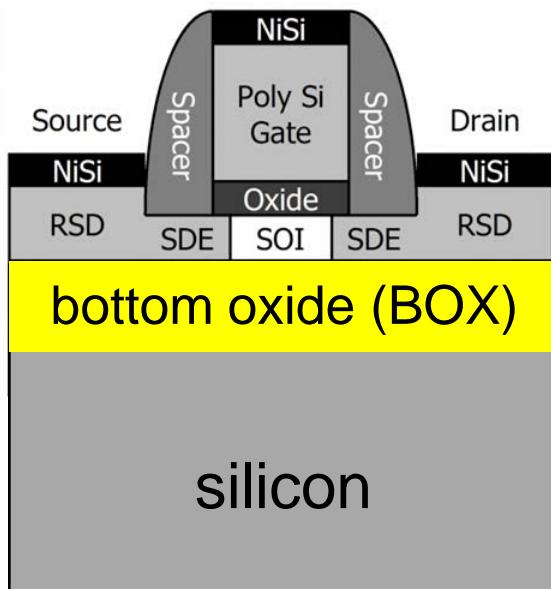
$$C_S = \frac{\epsilon_S}{t_{inv}} \gg C_{ox}$$

$$m \gg 1$$

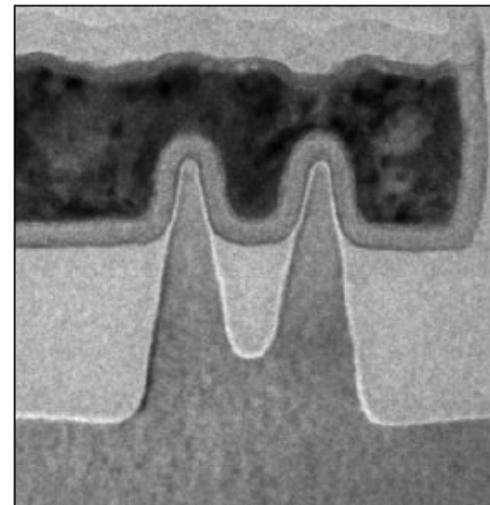
$$\Delta\psi_S = \Delta V_G \frac{C_{ox}}{C_{ox} + C_S} = \frac{\Delta V_G}{m}$$



Fully depleted ultra thin body (UTB) MOS structures



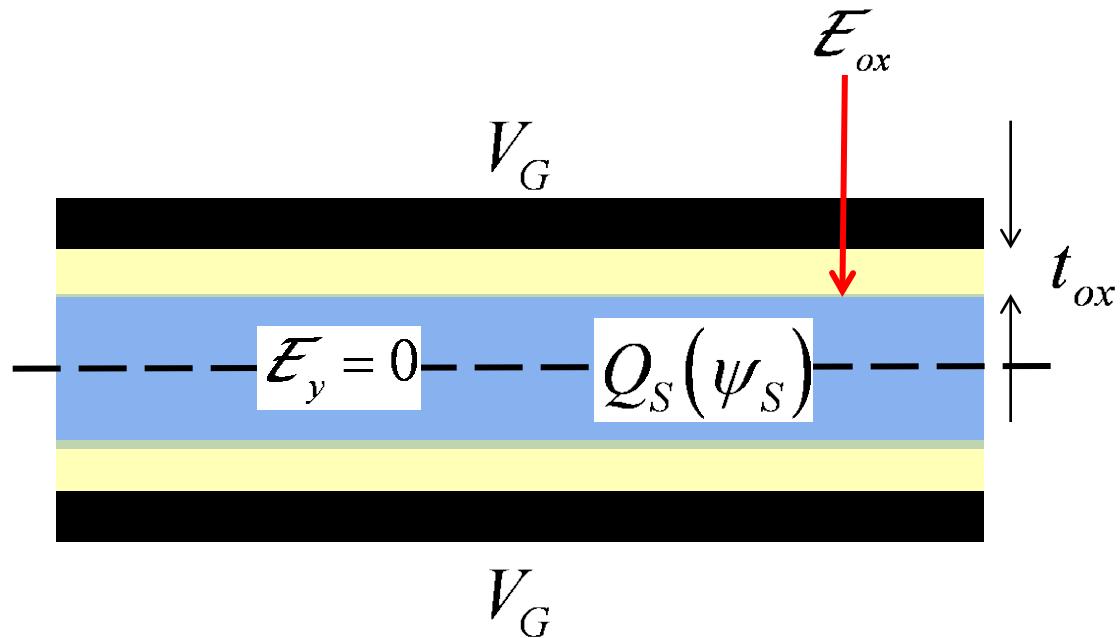
(ETSOI: Source: IBM, 2009)



(FinFET: Source: Intel, 2015)

Does anything change for these UTB MOS structures?

3) Mobile charge vs. gate voltage: Subthreshold UTB



Two gates, twice the capacitance.

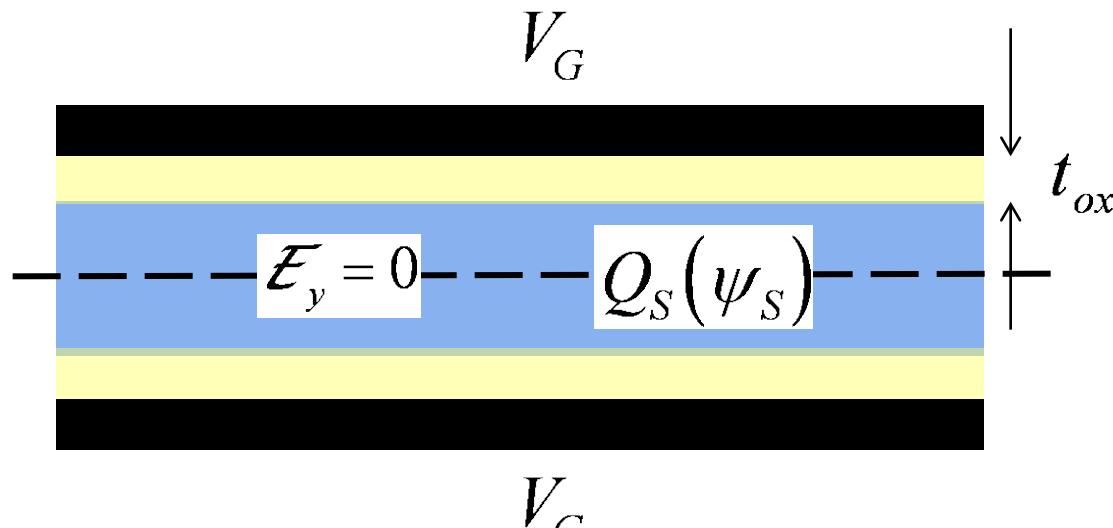
$$\varepsilon_{ox} E_{ox} = -Q_S(\psi_S)/2$$

$$\begin{aligned}\Delta V_{ox} &= E_{ox} t_{ox} \\ &= -Q_S(\psi_S)/2 C_{ox}\end{aligned}$$

$$V_G = \Delta V_{ox} + \psi_S$$

$$V_G = -\frac{Q_S(\psi_S)}{2C_{ox}} + \psi_S$$

Mobile charge vs. gate voltage: Subthreshold UTB



subthreshold

$$V_G = -\frac{Q_s(\psi_s)}{2C_{ox}} + \psi_s \quad Q_s(\psi_s) \approx 0 \rightarrow V_G \approx \psi_s$$

$$\psi_s = \frac{V_G}{m}$$

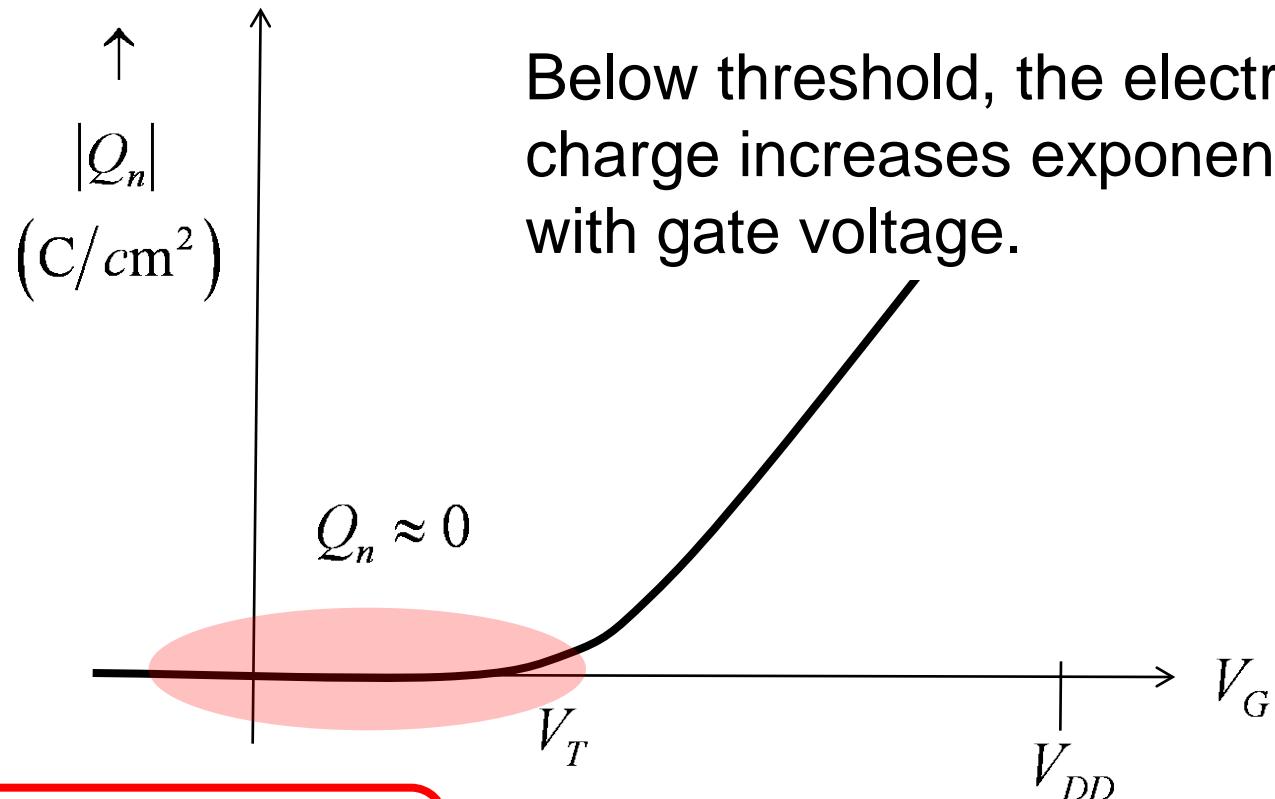
$m = 1$

Aside

$$SS = \left(\frac{\partial \log_{10} I_{DS}}{\partial V_{GS}} \Bigg|_{V_{DS}} \right)^{-1} = 2.3m \frac{k_B T}{q}$$

$m = 1$ for fully depleted structures.

Subthreshold charge vs. gate voltage



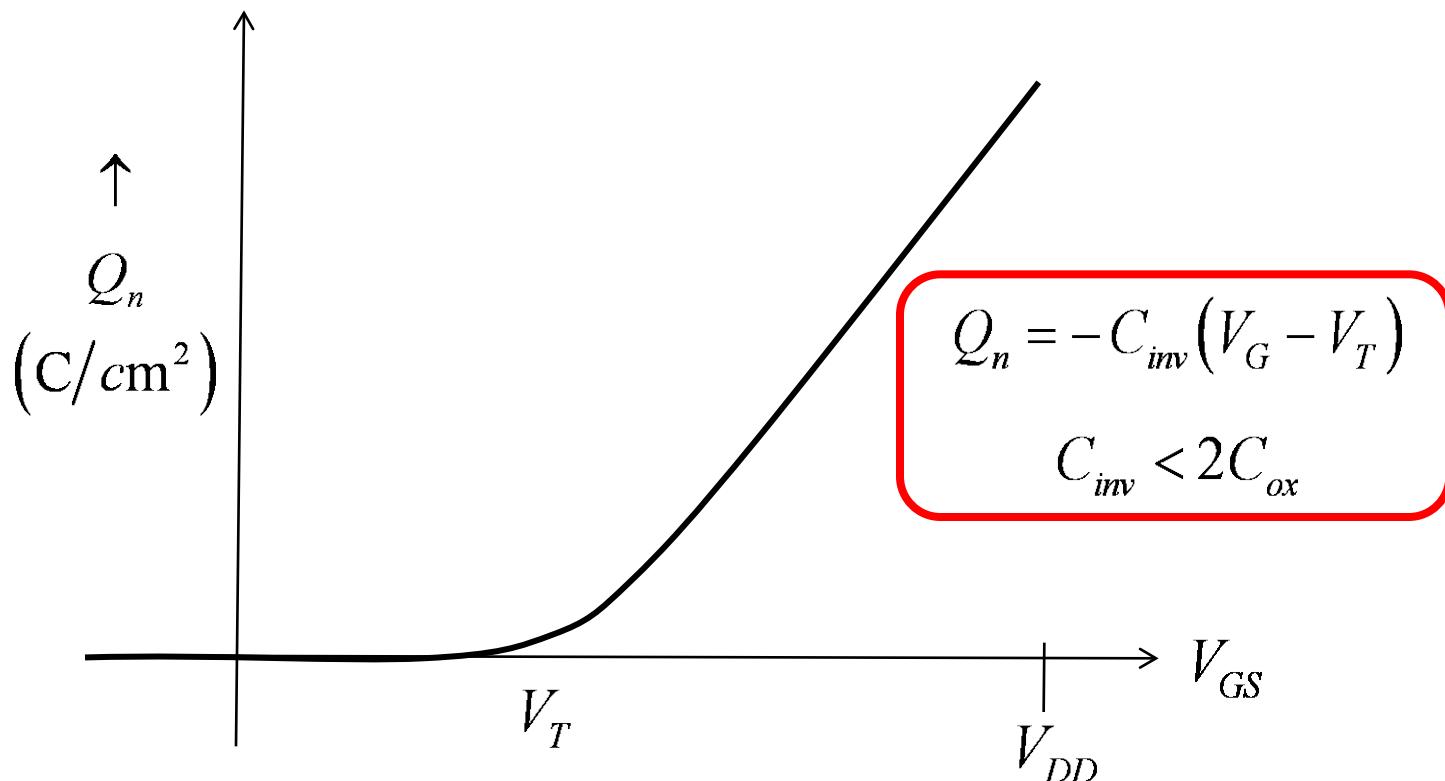
$$Q_n(V_G) \propto e^{q(V_G - V_T)/k_B T}$$

(FD UTB)

$$Q_n(V_G) \propto e^{q(V_G - V_T)/mk_B T}$$

(bulk MOS)

4) Above threshold charge vs. gate voltage



$$\frac{1}{C_{inv}} = \frac{1}{2C_{ox}} + \frac{1}{C_s(\text{inv})}$$

$$C_s = \frac{d(-Q_s)}{d\psi_s}$$

Summary

$$V_G \ll V_T : \quad Q_n(V_G) = -C_Q \left(\frac{k_B T}{q} \right) e^{q(V_G - V_T)/k_B T} \quad \text{FD UTB}$$

$$V_G \gg V_T : \quad Q_n(V_G) = -C_{inv}(V_G - V_T) \quad C_{inv} < 2C_{ox}$$

$$V_G \ll V_T : \quad Q_n(V_G) = -(m-1)C_{ox} \left(\frac{k_B T}{q} \right) e^{q(V_G - V_T)/mk_B T} \quad \text{bulk}$$

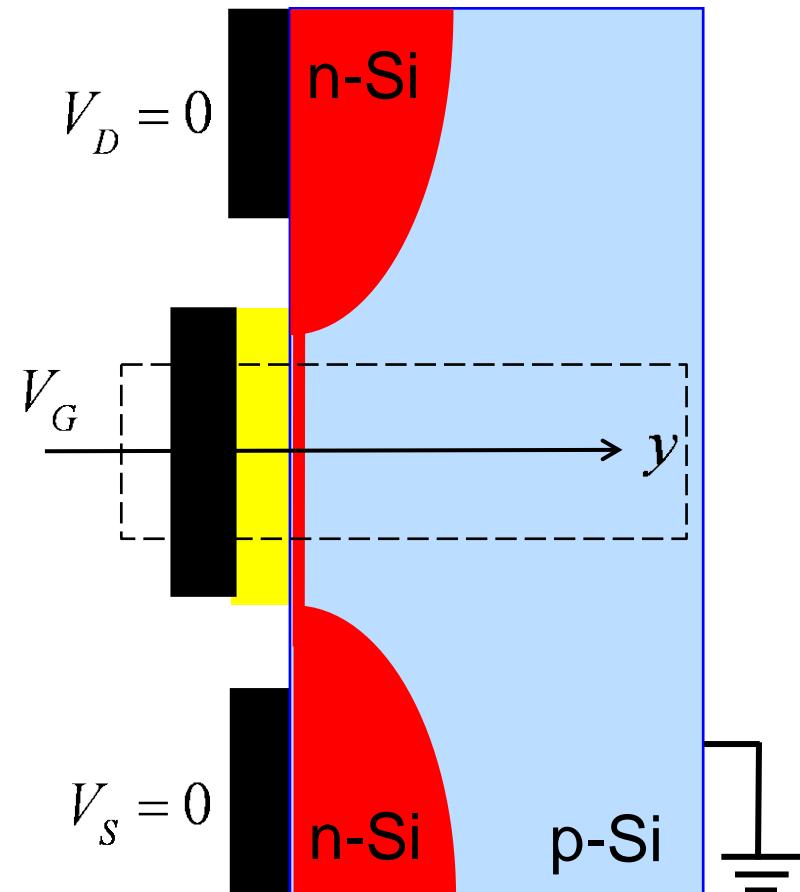
$$V_G \gg V_T : \quad Q_n = -C_{inv}(V_G - V_T) \quad C_{inv} < C_{ox}$$

Next topic

We have considered 1D MOS electrostatics – in the direction normal to the channel.

We now understand how the mobile charge varies with gate voltage above and below threshold.

Next Lecture: How does two-dimensional electrostatics affect MOSFETs?



Essentials of MOSFETs

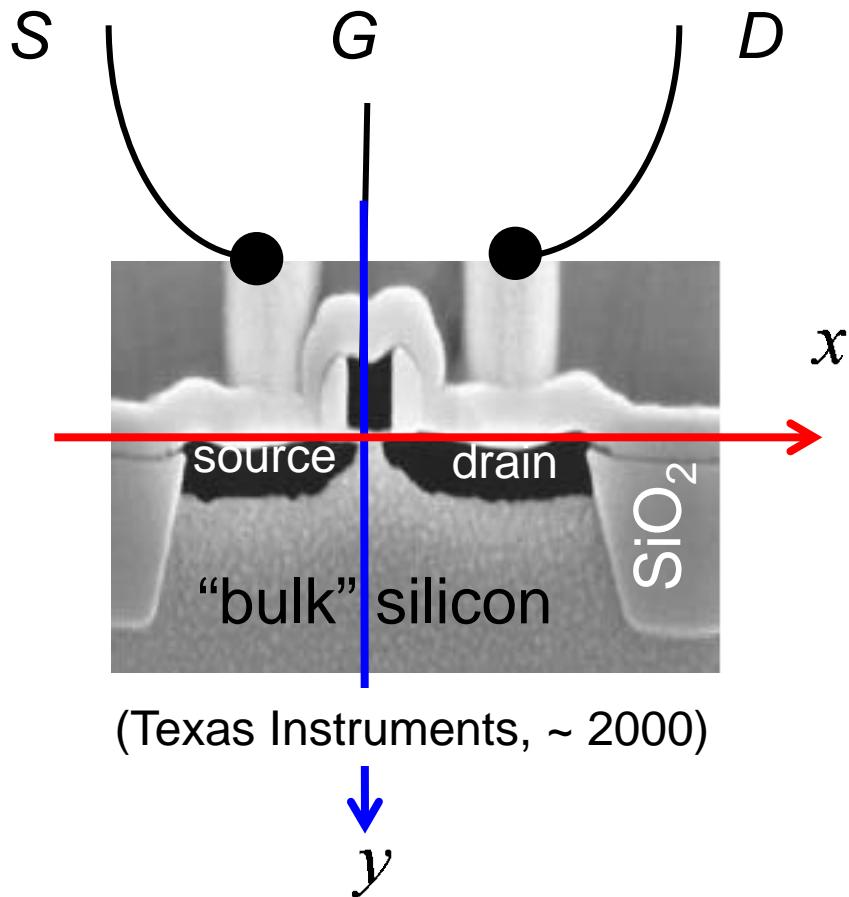
Unit 3: MOS Electrostatics

Lecture 3.8: 2D MOS Electrostatics

Mark Lundstrom

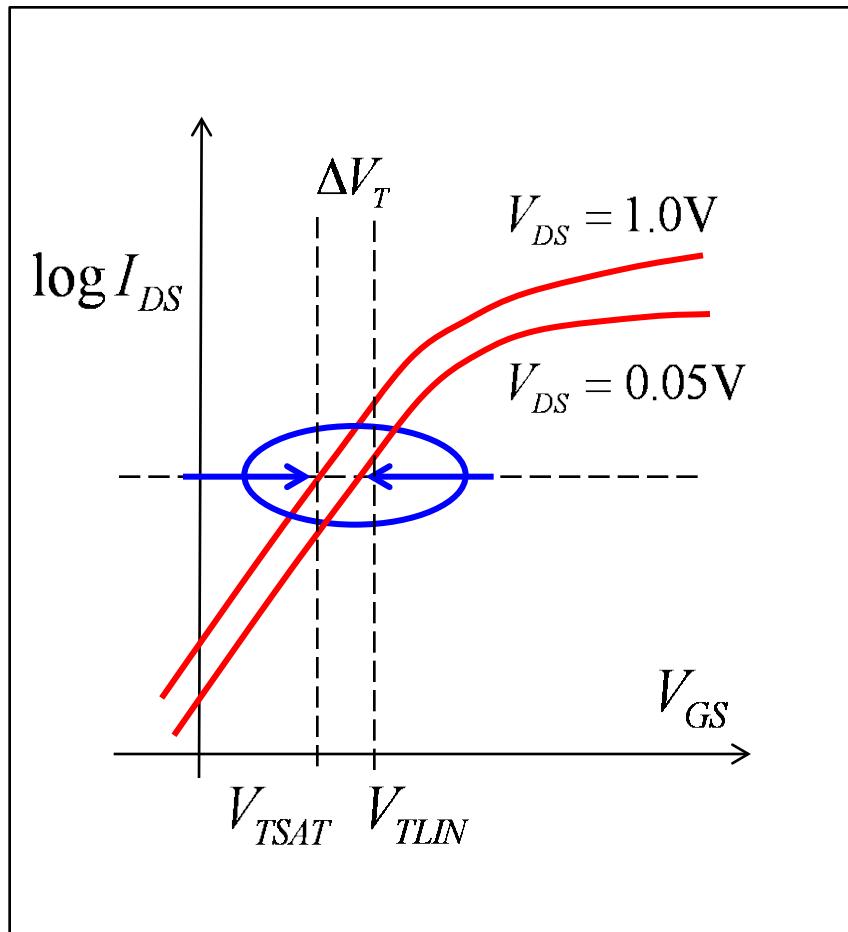
lundstro@purdue.edu
Electrical and Computer Engineering
Purdue University
West Lafayette, Indiana USA

2D electrostatics



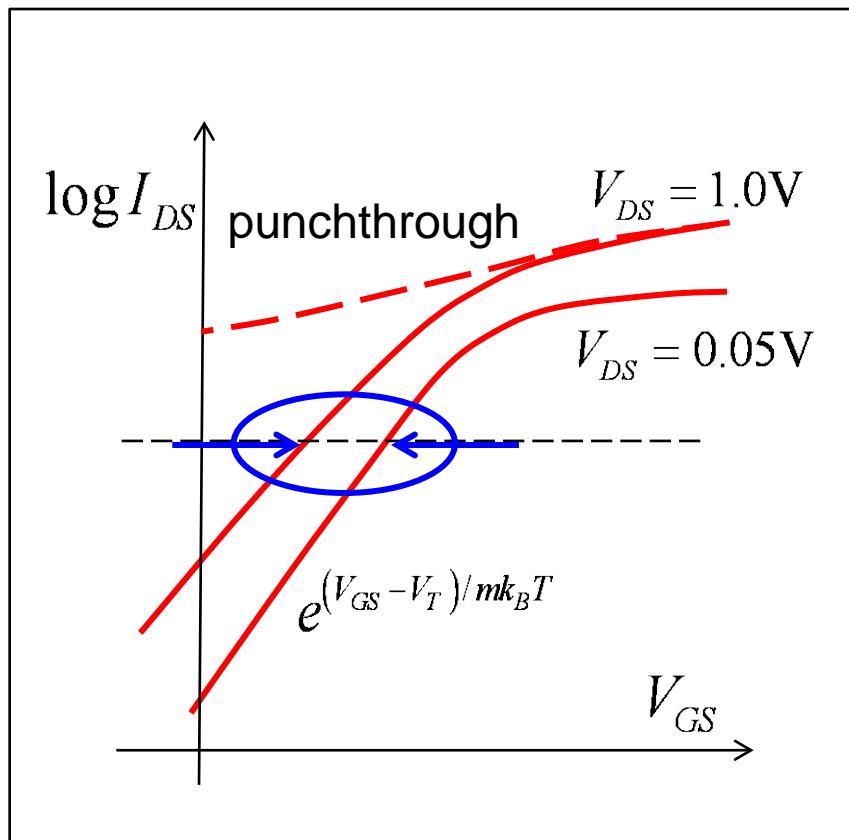
MOSFET operation is determined by the 2D (and 3D) energy bands, which vary in space according to the 2D electrostatic potential: $\psi(x, y)$

Effect of 2D electrostatics (i)



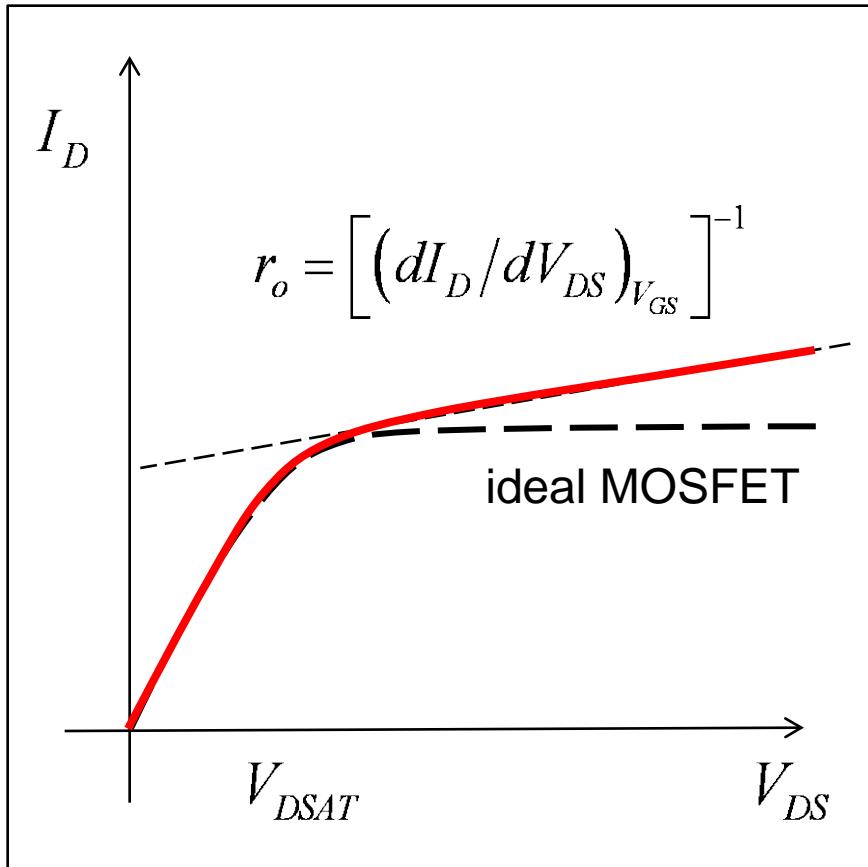
- 1) Threshold voltage decreases as the drain voltage increases
- 2) Threshold voltage decreases as the channel length decreases
- 3) DIBL increases as channel length decreases

Effect of 2D electrostatics (ii)



- 1) SS may increase as the drain voltage increases
- 2) SS may increase as the channel length decreases
- 3) In severe cases, the device may “punch through”

Effect of 2D electrostatics (iii)



- 1) Output resistance decreases as channel length decreases.

2D Poisson equation

These effects are referred to as “short channel effects.” To understand them, we should understand the solution to the 2D Poisson equation.

$$\nabla \cdot \vec{D}(x,y) = \rho(x,y)$$

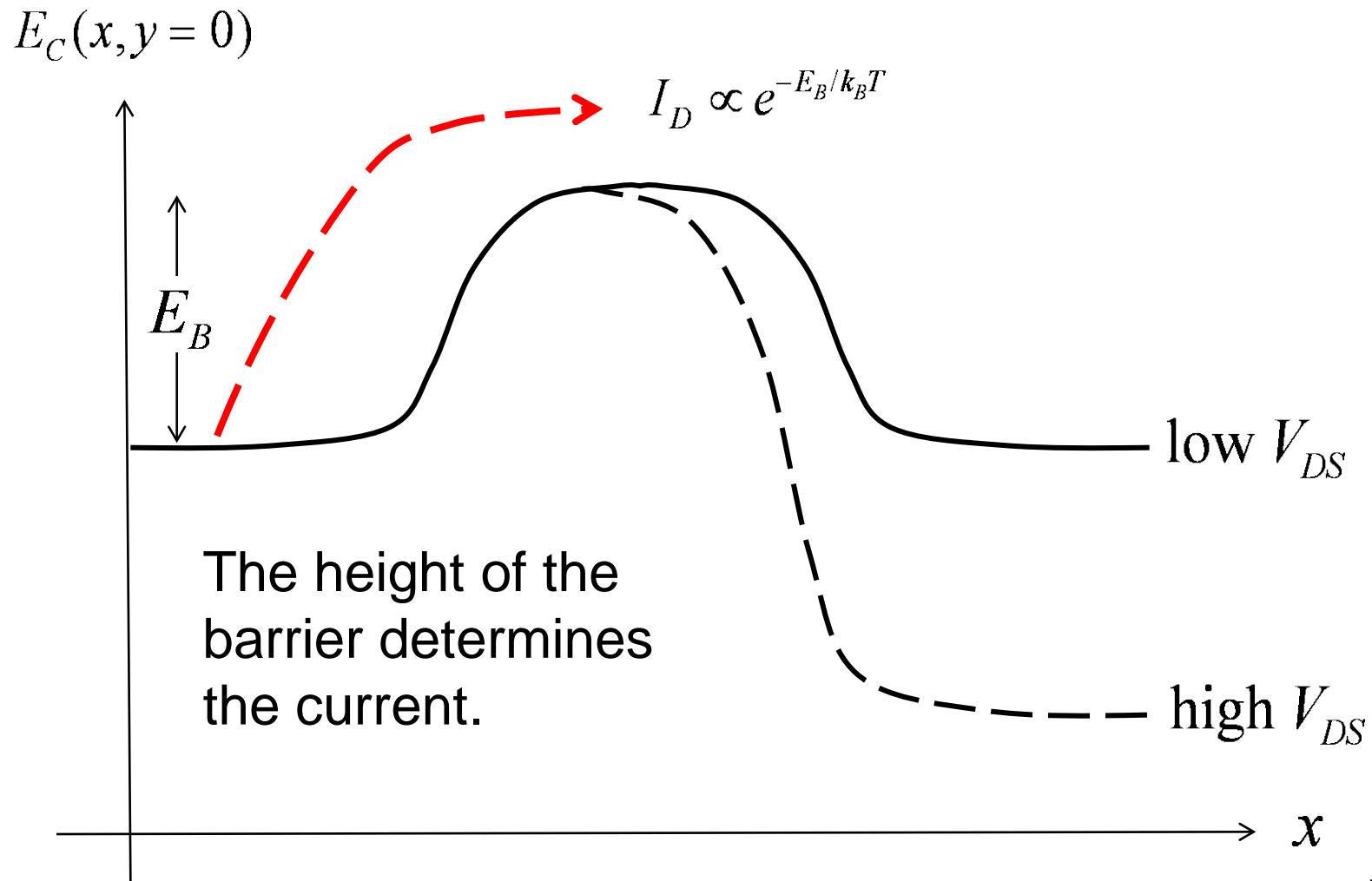
$$\vec{D}(x,y) = \epsilon_s \vec{\mathcal{E}}(x,y)$$

$$\vec{\mathcal{E}}(x,y) = -\vec{\nabla} \psi(x,y)$$

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} = -\frac{\rho(x,y)}{\epsilon_s}$$

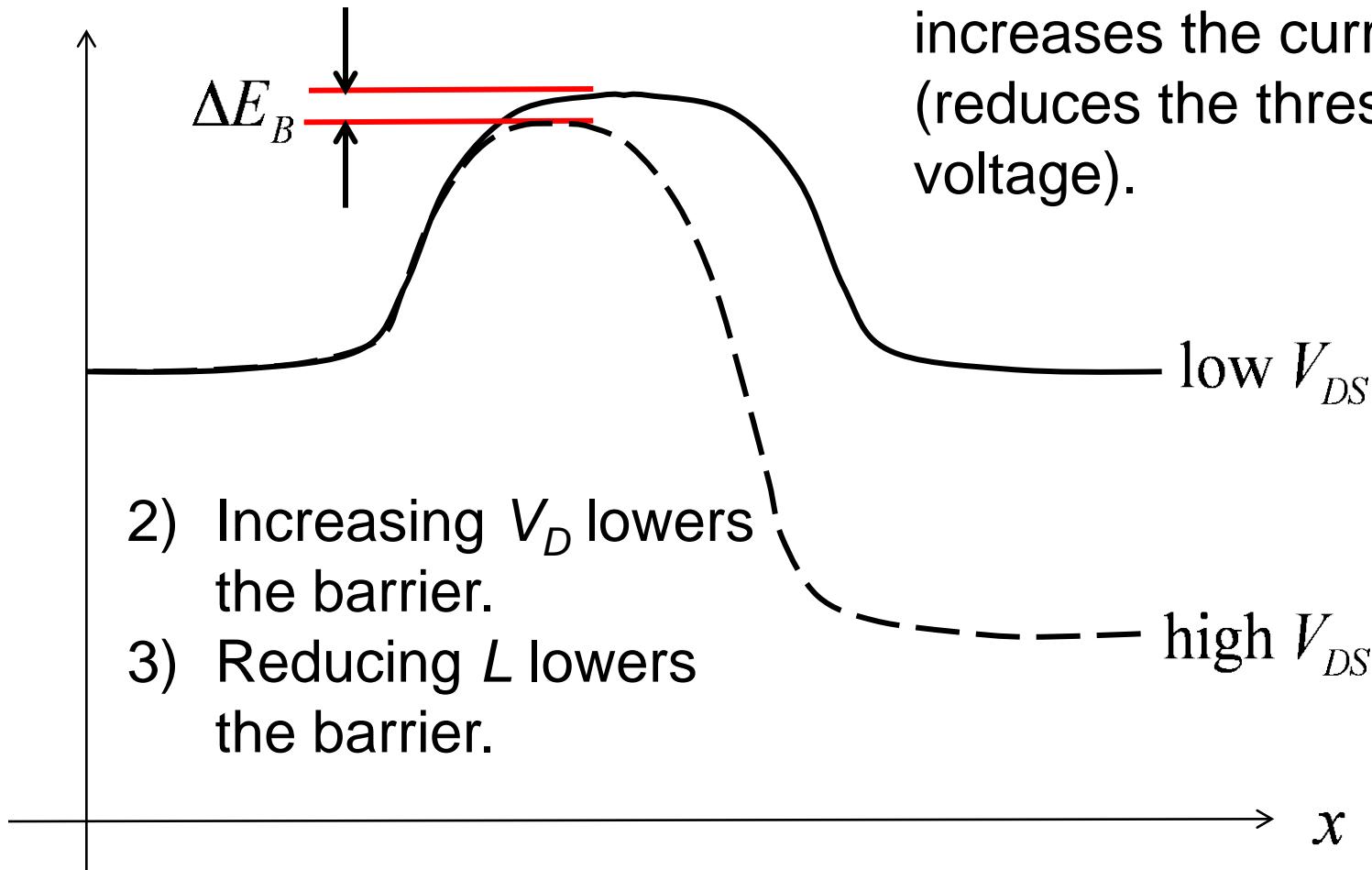
In this lecture, we will develop a **qualitative understanding** of the solutions.

V_T roll-off: The barrier lowering view

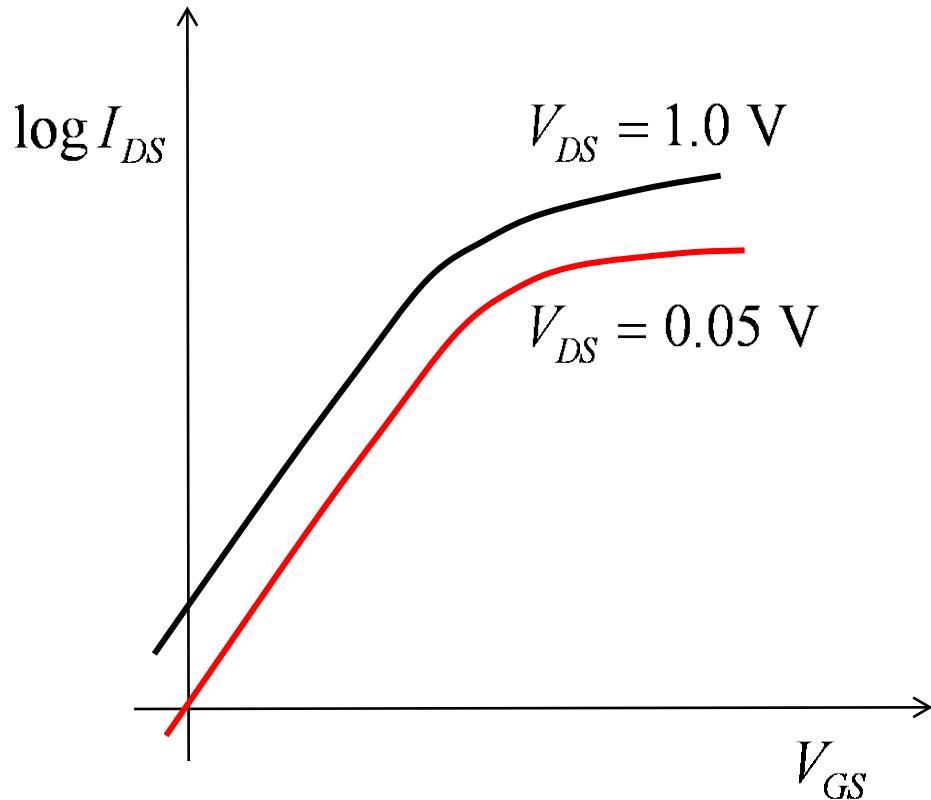


Barrier lowering

$E_C(x, y = 0)$



Effects of barrier lowering



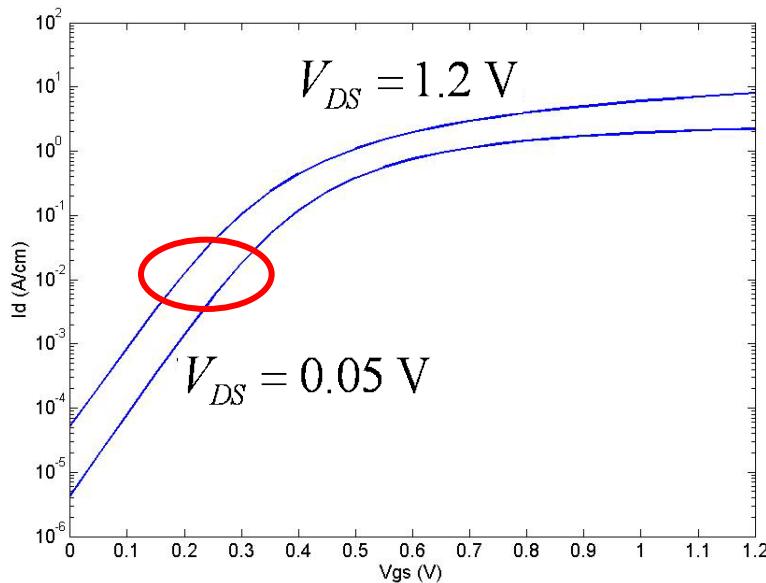
A small amount of barrier lowering reduces V_T (causes DIBL), but the SS does not change.

More barrier lowering also increases SS.

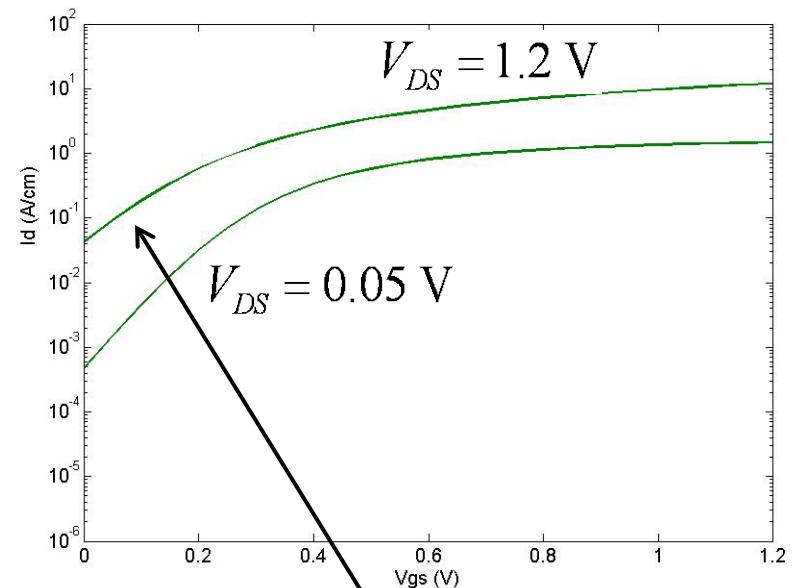
A large amount of barrier lowering is catastrophic – the gate loses control of the barrier.

Example

$L = 105 \text{ nm}$



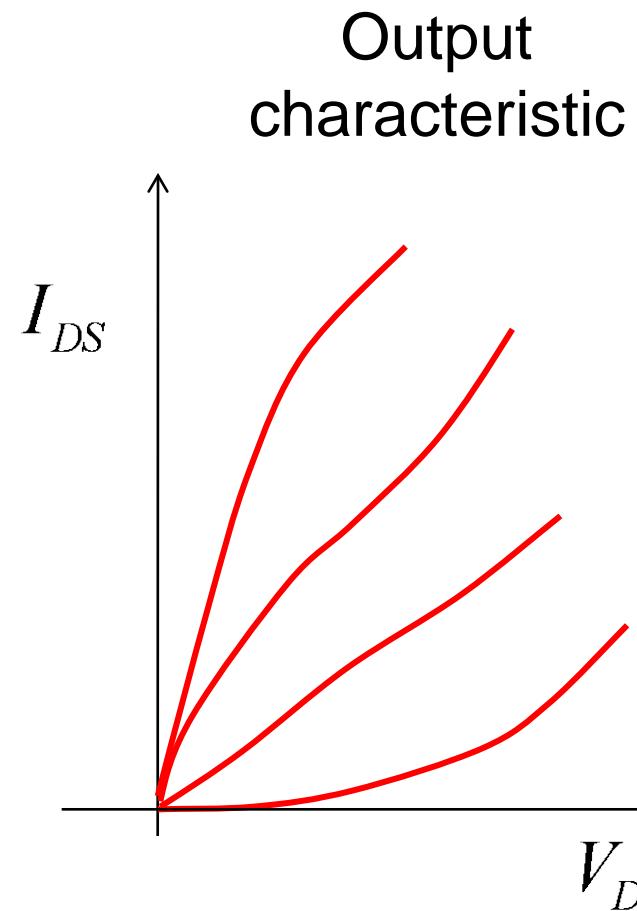
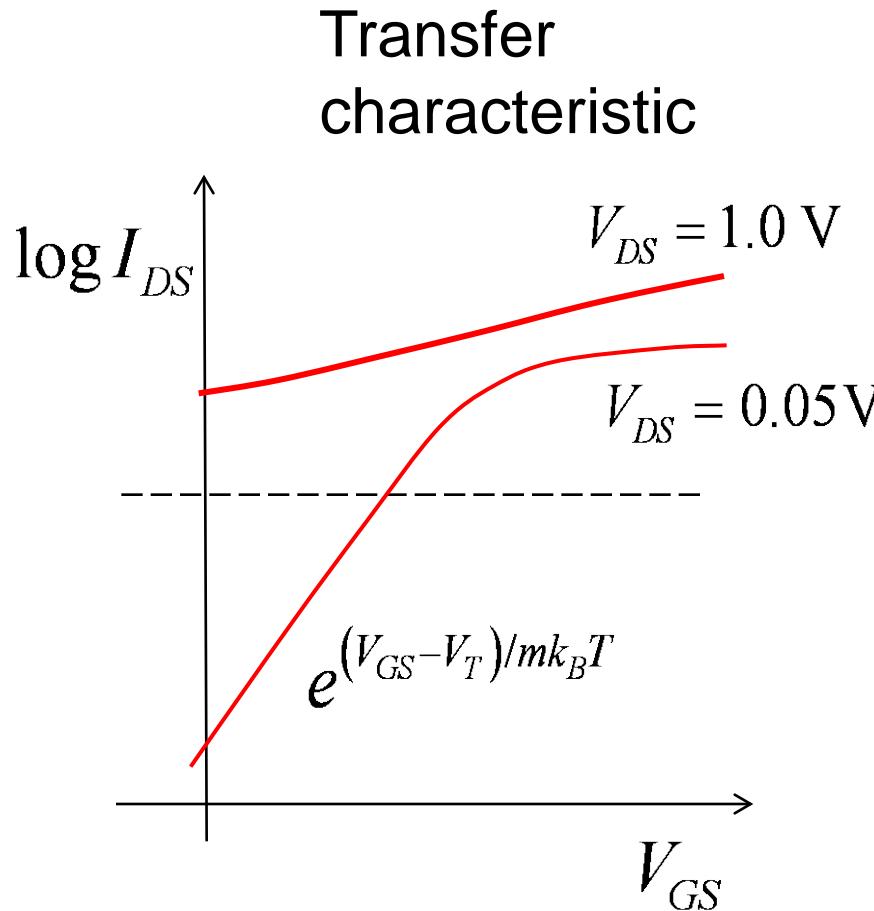
$L = 85 \text{ nm}$



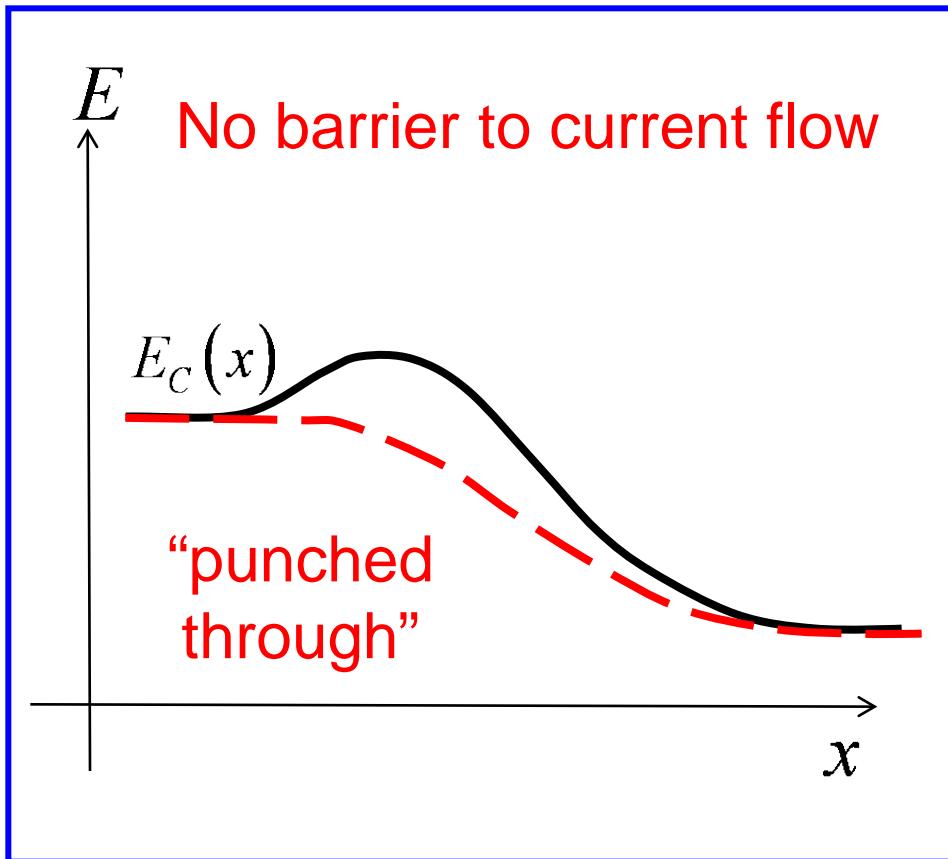
(Courtesy, Shuji Ikeda, ATDF,
Dec. 2007)

Increased DIBL and SS

Punch-through



Physics of punch-through



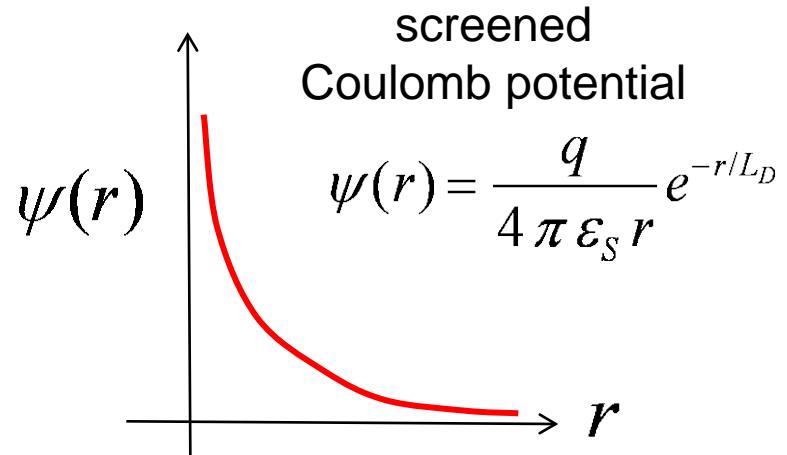
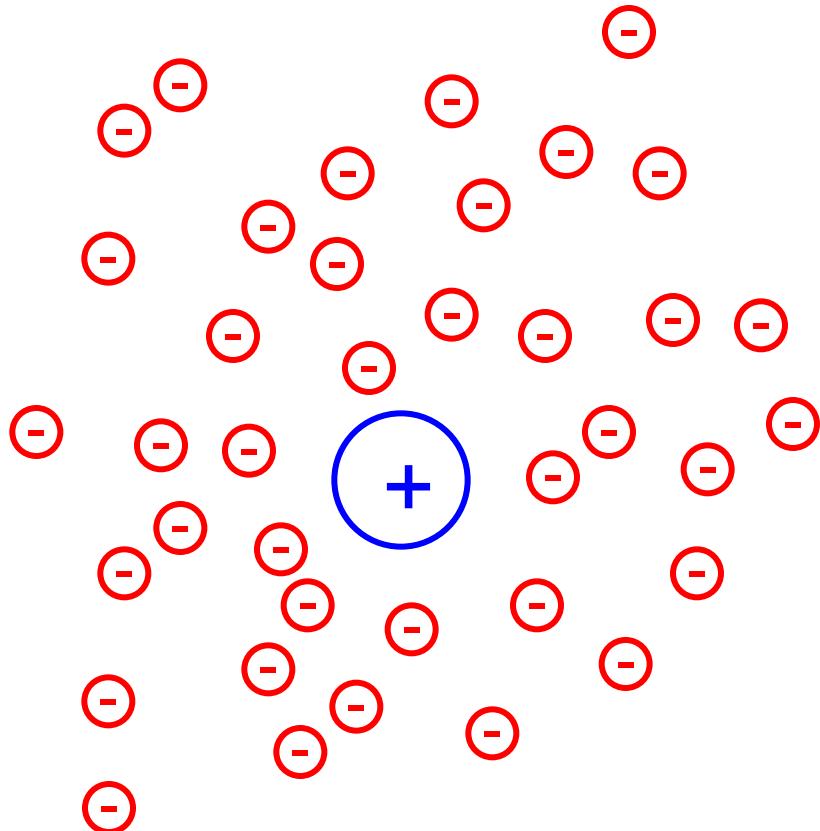
even for a small voltage larger will be the current. gate cant able to control the current.

Controlling 2D electrostatics

Question: How do we control 2D electrostatics in short channel MOSFETs?

Answer: Screen out the 2D fields.

Screening by free carriers



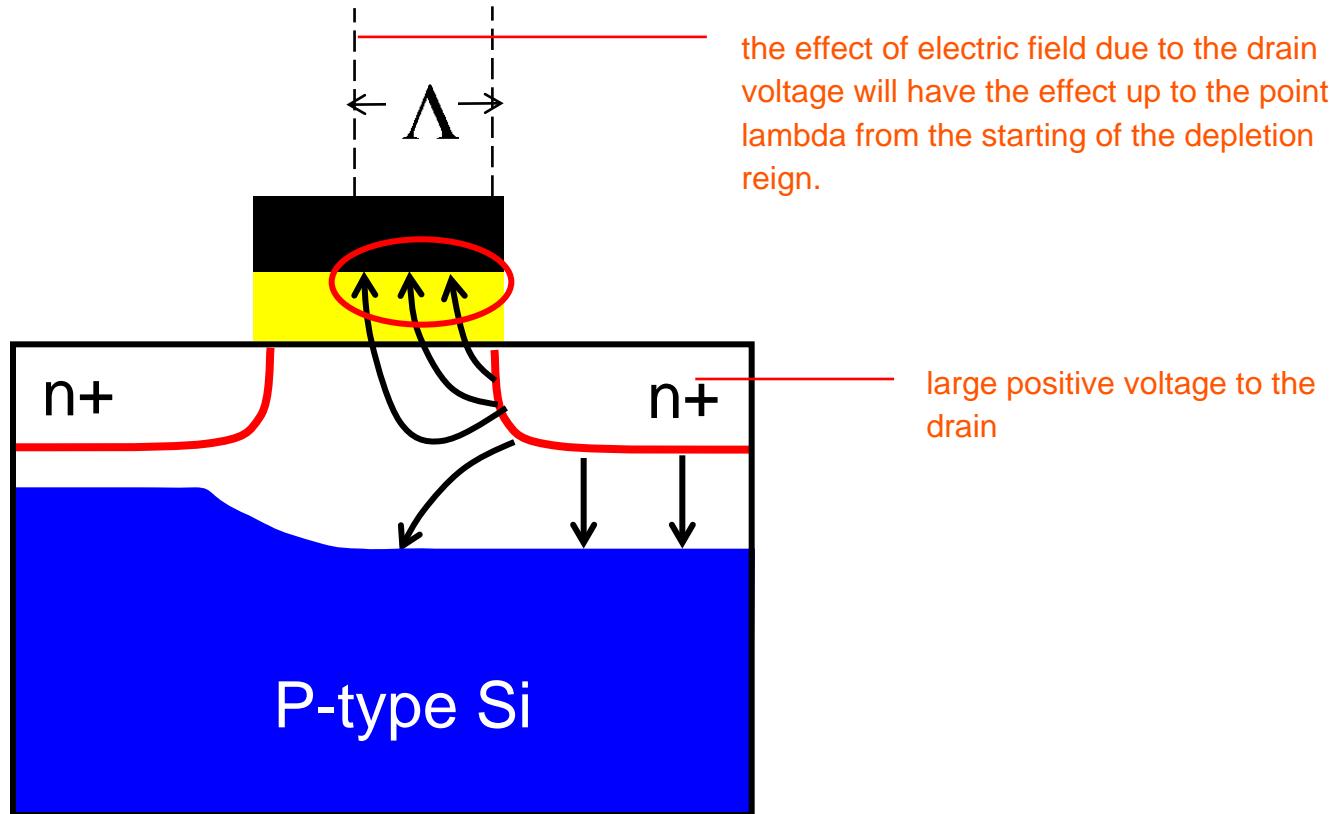
influence of the positive charge over a distance is known as screening length.

$$L_D = \sqrt{\frac{\epsilon_S k_B T}{q^2 n_0}}$$

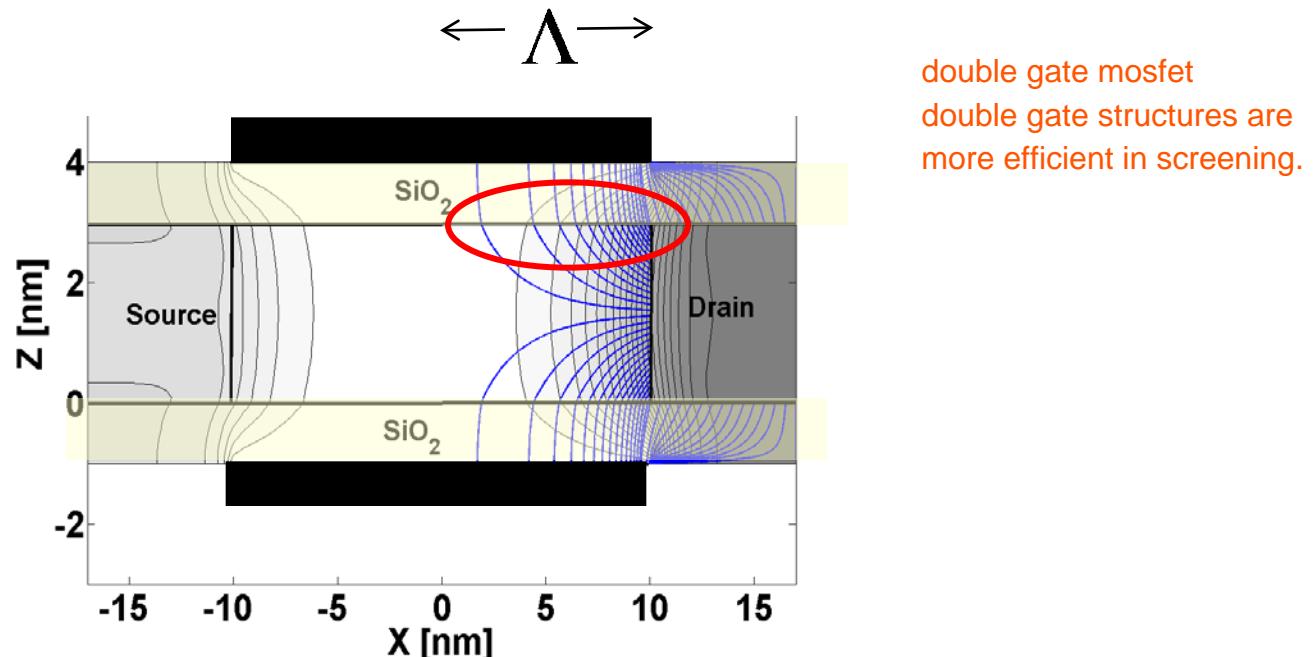
Debye length

due to the attraction of the positive charge all electrons will be attracted but the electrons will be repelled by each others charge. hence their net charge will be cancelled out.

Geometric screening length: Bulk MOSFET



Geometric screening length: DG MOSFET

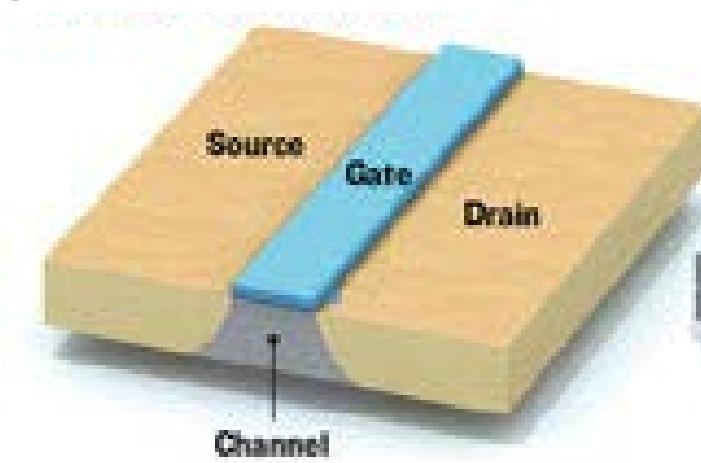


$$t_{ox} = 1 \text{ nm}$$

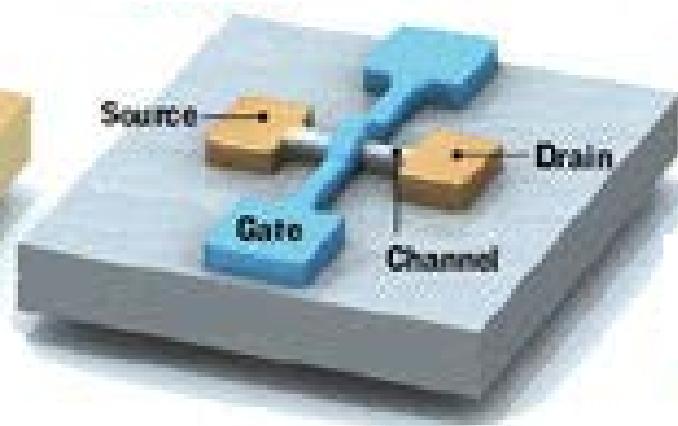
Off-state: $V_G = 0V$, $V_D = 1V$, $I_{off} = 0.1\mu\text{A}/\mu\text{m}$ (by H. Pal, Purdue, 2012)

Non-planar MOSFETs

planar transistor



FinFET



“Transistors go Vertical,” *IEEE Spectrum*, Nov. 2007.

See also: “Integrated Nanoelectronics of the Future,” Robert Chau, Brian Doyle, Suman Datta, Jack Kavalieros, and Kevin Zhang, *Nature Materials*, **6**, 2007

Computing Λ

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} = -\frac{\rho(x,y)}{\epsilon_s} = \frac{qN_A(x,y)}{\epsilon_s} \quad (\text{subthreshold})$$

$$\Lambda_{NW} < \Lambda_{DG\ SOI} < \Lambda_{BULK}$$

$$L_{\min} \approx 3\Lambda$$

D. J. Frank, Y. Taur, and H.-S. P. Wong, “Generalized scale length for two-dimensional effects in MOSFETs,” *IEEE Electron Device Lett.*, **19**, pp. 385–387, 1998.

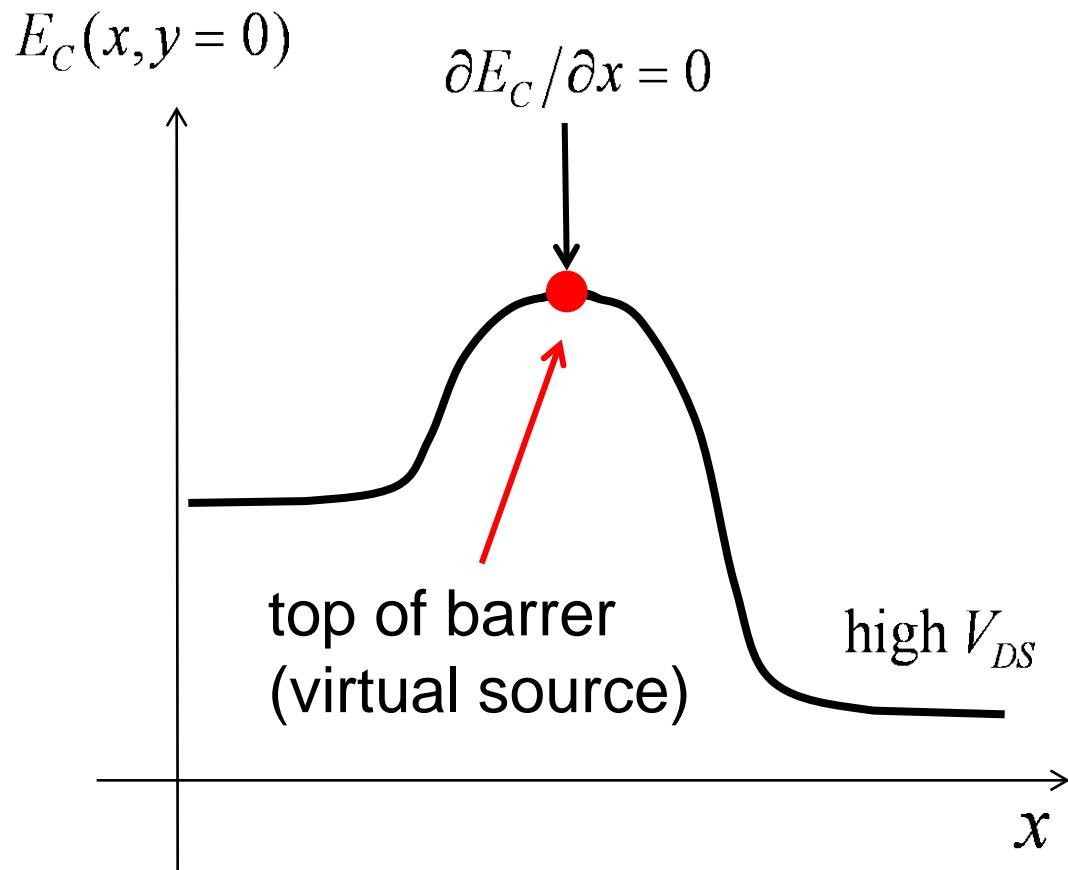
Qian Xie, Jun Xu, and Yuan Taur, “Review and Critique of Analytic Models of MOSFET Short-Channel Effects in Subthreshold,” *IEEE Trans. Electron Dev.*, **59**, pp 1569-1579, 2012.

Understanding 2D electrostatics

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} = \frac{qN_A(x,y)}{\epsilon_S}$$

- 1) Barrier lowering
- 2) Geometric screening length
- 3) Capacitor model
(Lundstrom, *Fundamentals of Nanotransistors*, Sec. 10.5)

The Virtual Source and 2D electrostatics



$$\partial E_C / \partial x = 0 \rightarrow \partial^2 \psi / \partial x^2 = 0$$

$$\partial^2 \psi / \partial y^2 \gg \partial^2 \psi / \partial x^2$$

GCA (Gradual Channel Approximation) is valid

1D Poisson eqn. (see slide 6)

$$Q_n = -C_{ox} (V_{GS} - V_T)$$

$$V_T = V_{T0} - \delta V_{DS}$$

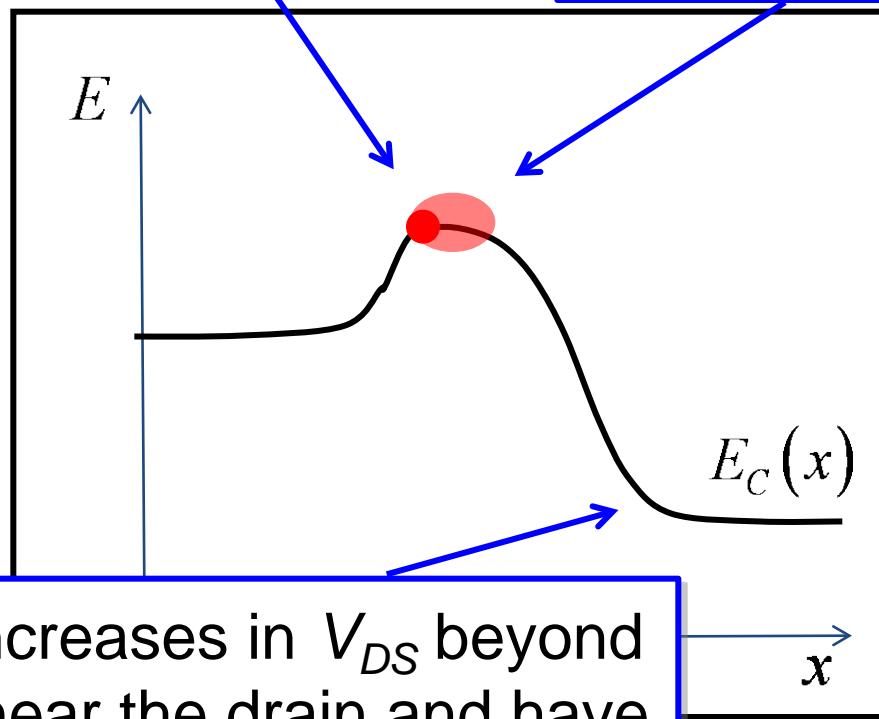
The job of the device designer is to make sure that this is true.

“Well-tempered MOSFET”

1) $Q_n(0) \approx -C_{inv}(V_{GS} - V_T)$

2) region under strong control of gate ($m \sim 1$)

$V_T = V_{T0} - \delta V_{DS}$
 $m = \text{constant}$



3) Additional increases in V_{DS} beyond V_{DSAT} drop near the drain and have a **small effect** on I_D (small DIBL)

(After D.A. Antoniadis, MIT)

Summary

- 1) 2D MOS electrostatics degrade device performance (V_T depends on L , DIBL and SS increase).
- 2) The goal of MOSFET design is to make 1D electrostatics hold at the VS – with small DIBL and SS.
- 3) To achieve this, we engineer the device such that the gate voltage controls the height of the source to channel energy barrier and the drain voltage has a small influence.

Next topic

Now that we have a better understanding of MOS electrostatics, it's time to re-visit the VS model.

Essentials of MOSFETs

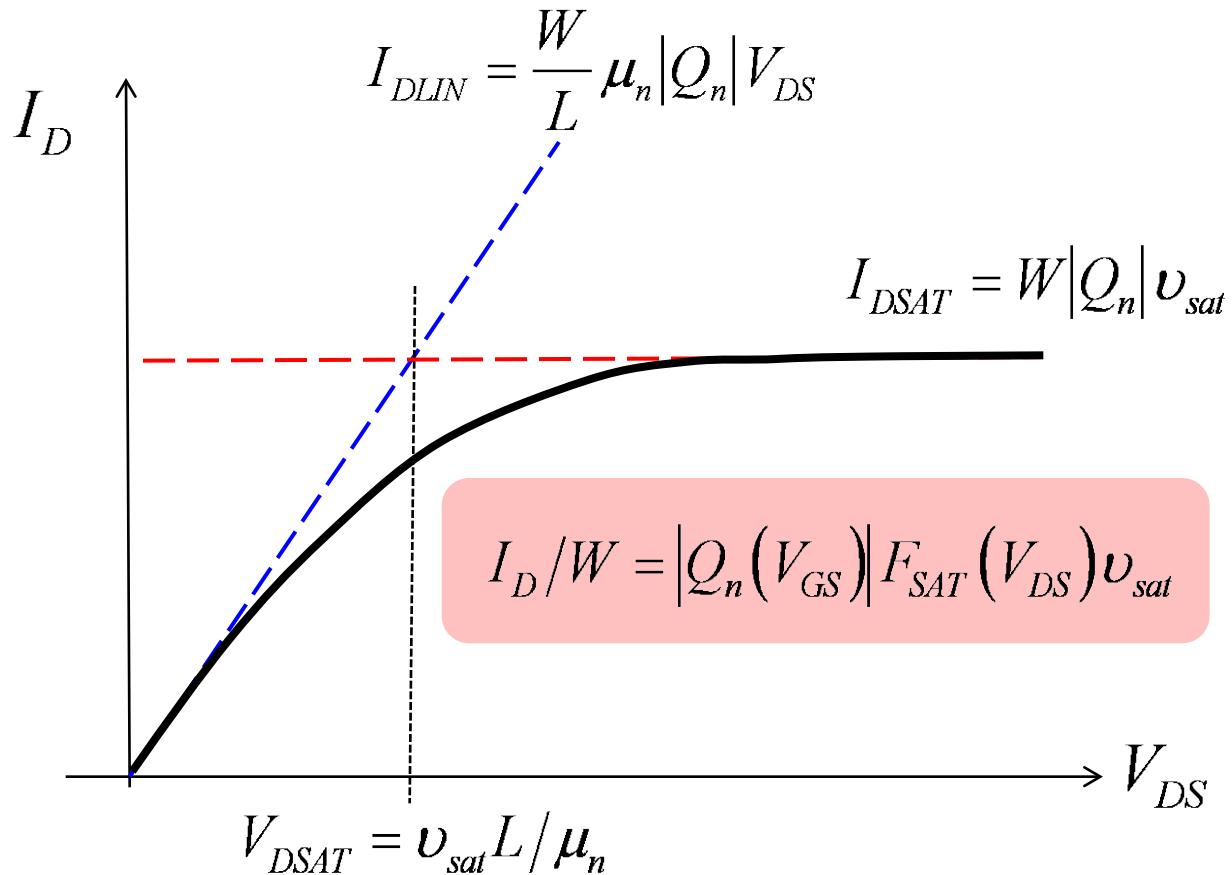
Unit 3: MOS Electrostatics

Lecture 3.9: The VS Model Revisited

Mark Lundstrom

lundstro@purdue.edu
Electrical and Computer Engineering
Purdue University
West Lafayette, Indiana USA

Level 0 VS Model



Level 0 VS model

$$1) \quad I_D/W = |Q_n(V_{GS})| \langle v_x(V_{DS}) \rangle$$

$$2) \quad Q_n(V_{GS}) = -C_{ox}(V_{GS} - V_T) \quad (V_{GS} > V_T)$$

$$V_T = V_{T0} - \delta V_{DS}$$

$$Q_n(V_{GS}) = 0 \quad (V_{GS} \leq V_T)$$

$$3) \quad \langle v(V_{DS}) \rangle = F_{SAT}(V_{DS}) v_{sat}$$

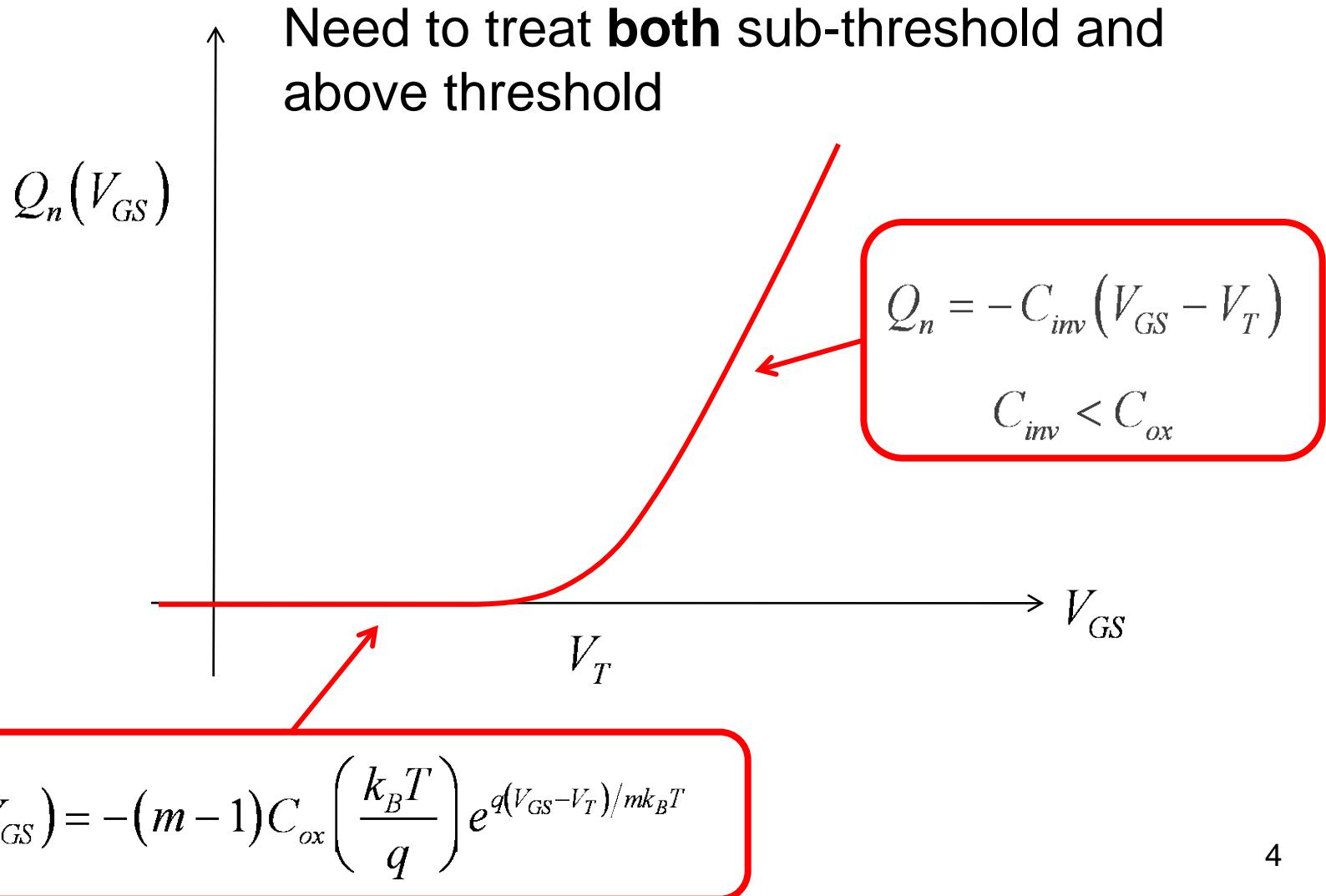
$$4) \quad F_{SAT}(V_{DS}) = \frac{V_{DS}/V_{DSAT}}{\left[1 + (V_{DS}/V_{DSAT})^\beta\right]^{1/\beta}}$$

$$5) \quad V_{DSAT} = v_{sat} L / \mu_n$$

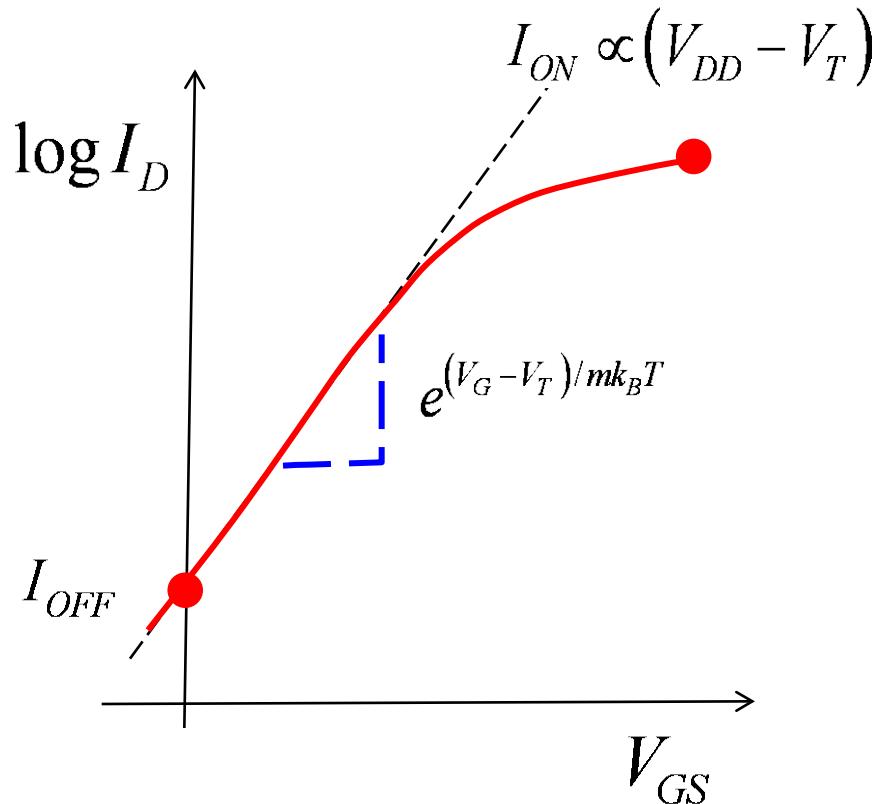
There are only 8 device-specific parameters in this model:

$$C_{ox}, V_{T0}, \delta, v_{sat}, \mu_n, L$$
$$R_{SD} = R_S + R_D, \beta$$

Improving the VS model: inversion charge



But first: the subthreshold current



- 1) subthreshold swing
- 2) off-current
- 3) on-current

$$Q_n(V_{GS}) \propto e^{q(V_{GS} - V_T)/mk_B T}$$

$$I_D \propto Q_n(V_{GS})$$

$$I_D \propto e^{q(V_{GS} - V_T)/mk_B T}$$

Subthreshold swing

$$I_D \propto e^{q\psi_S/k_B T} \quad \psi_S = V_{GS}/m \quad m \geq 1$$

$$\ln I_D = \frac{\psi_S}{(k_B T/q)} + c \quad \log_{10} I_D = \frac{\psi_S}{2.3(k_B T/q)} + \frac{c}{2.3}$$

$$\frac{\partial(\log_{10} I_D)}{\partial V_{GS}} = \frac{\partial(\log_{10} I_D)}{\partial \psi_S} \times \frac{\partial \psi_S}{\partial V_{GS}} = \frac{1}{2.3(k_B T/q)} \times \frac{1}{m} \quad \frac{\text{Decades of } I_D}{\text{Volts of } V_{GS}}$$

$$SS = \left(\frac{\partial(\log_{10} I_D)}{V_{GS}} \right)^{-1} = 2.3m(k_B T/q) \frac{\text{mV}}{\text{dec}}$$

Subthreshold swing

$$S = 2.3m \left(k_B T / q \right) \frac{\text{mV}}{\text{dec}} \quad m \geq 1$$

$m \approx 1.1 - 1.4$ typically

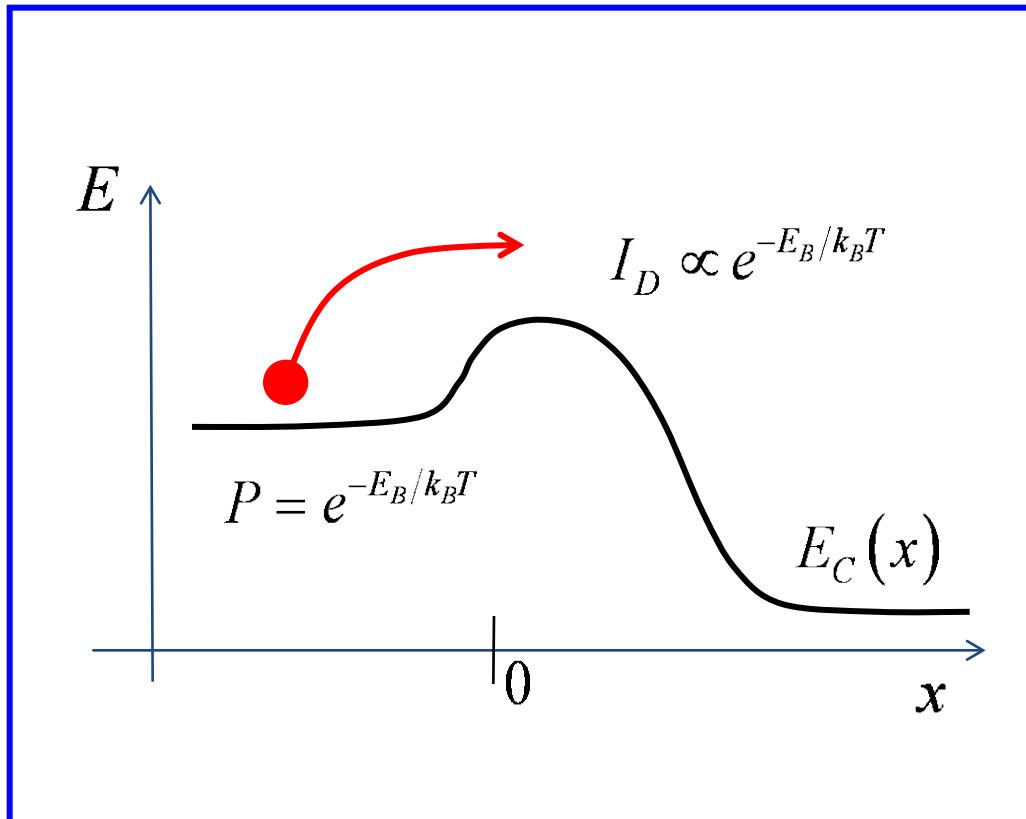
$SS > 60 \frac{\text{mV}}{\text{dec}}$ ($T = 300K$)

$SS < 100 \frac{\text{mV}}{\text{dec}}$ (typically)

Why is a small SS important?

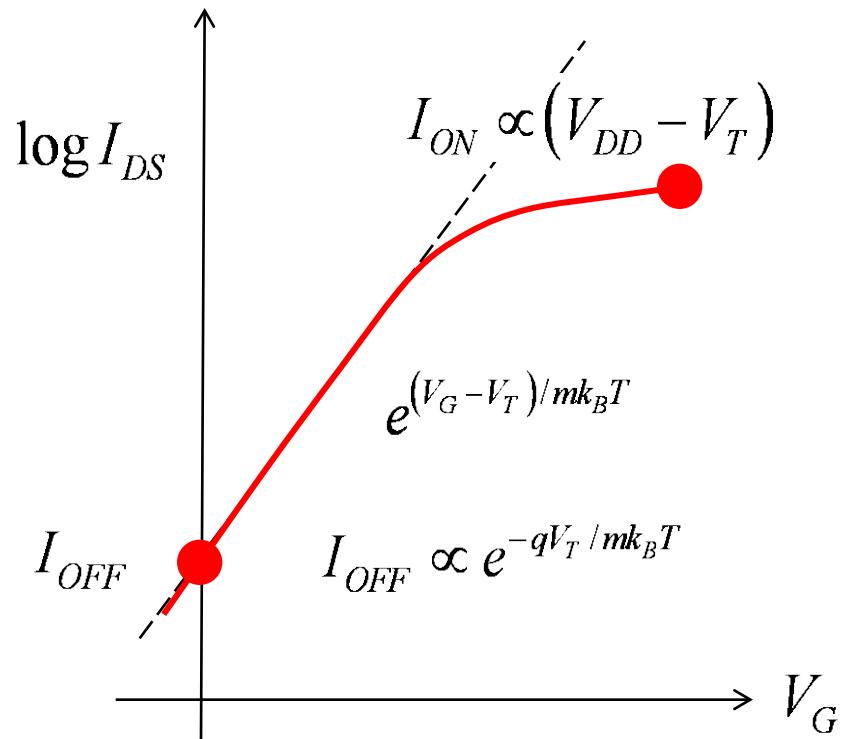
$$P_D \propto V_{DD}^2$$

Why is $S > 60 \text{ mV/decade}$?

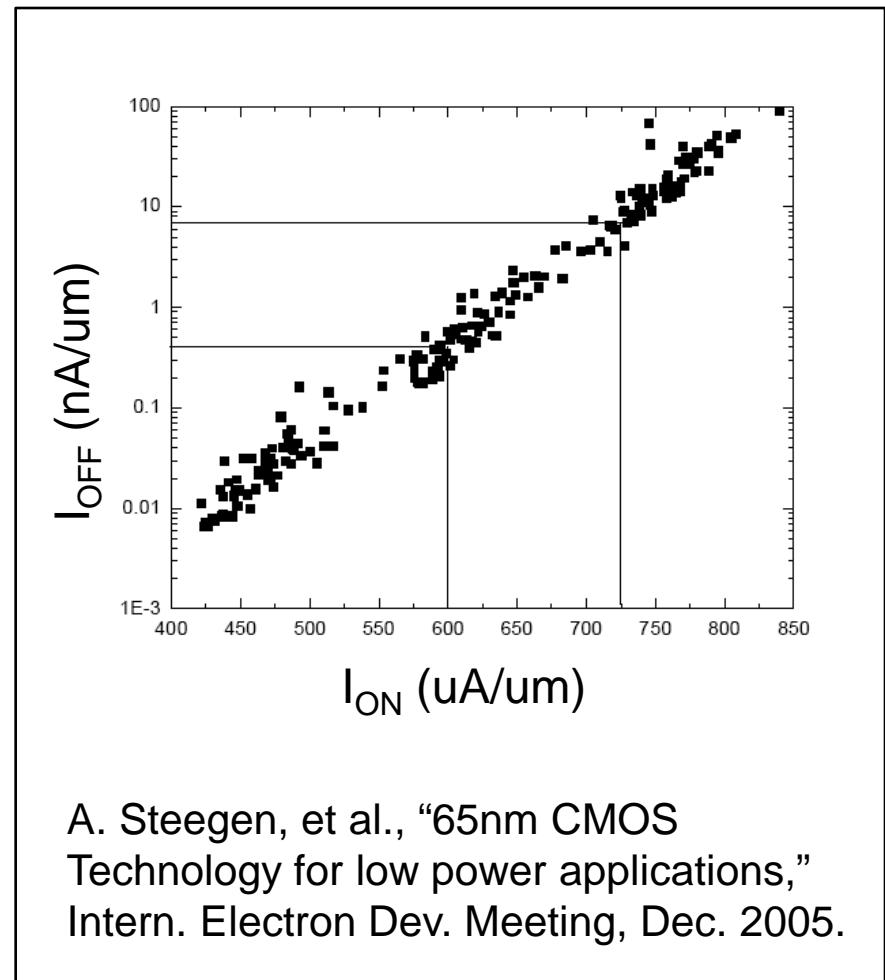


injection of thermal carriers over a barrier

Relation between I_{OFF} and I_{ON}



$$\log_{10} I_{OFF} = c_1 + c_2 I_{ON}$$



A. Steegen, et al., "65nm CMOS Technology for low power applications," Intern. Electron Dev. Meeting, Dec. 2005.

Back to the VS model

Now let us return to the question at hand:

“How do we describe $Q_n(V_{GS})$ **continuously** below and above threshold?

Empirical treatment of inversion charge

$$Q_n(V_{GS}) = -C_{inv}m(k_B T/q) \ln\left(1 + e^{q(V_{GS}-V_T)/mk_B T}\right)$$

$$Q_n(V_{GS}) = -(m-1)C_{ox}(k_B T/q)e^{q(V_{GS}-V_T)/mk_B T}$$

----- correct result -----

G. T. Wright, "Threshold modelling of MOSFETs for CAD of CMOS VLSI,"
Electron Lett., **21**, pp. 223–224, Mar. 1985.

Empirical treatment of inversion charge

$$Q_n(V_{GS}) = -C_{inv}m(k_B T/q) \ln\left(1 + e^{q(V_{GS}-V_T)/mk_B T}\right)$$

$V_{GS} > V_T$:

$$\ln(1 + x) \approx \ln(x)$$

$$Q_n(V_{GS}) \approx -C_{inv}(V_{GS} - V_T)$$

$$Q_n(V_{GS}) = -C_{inv}(V_{GS} - V_T)$$

correct

G. T. Wright, "Threshold modelling of MOSFETs for CAD of CMOS VLSI,"
Electron Lett., **21**, pp. 223–224, Mar. 1985.

Empirical treatment of inversion charge

expression used in the MIT VS Model

$$Q_n(V_{GS}, V_{DS}) = -C_{inv}m(k_B T / q) \ln\left(1 + e^{q(V_{GS} - V_T + \alpha(k_B T_L / q)F_f) / m k_B T}\right)$$

$$V_T = V_{T0} - \delta V_{DS}$$

DIBL

different V_T 's in strong and weak inversion

Ali Khakifirooz, Osama M. Nayfeh, and Dimitri Antoniadis, "A Simple Semi-empirical Short-Channel MOSFET Current–Voltage Model Continuous Across All Regions of Operation and Employing Only Physical Parameters," *IEEE Trans. Electron Devices*, **56**, pp. 1674-1680, 2009.

Level 1 VS model

$$1) \quad I_D/W = |Q_n(V_{GS}, V_{DS})| \langle v_x(V_{DS}) \rangle$$

$$2) \quad Q_n(V_{GS}, V_{DS}) = -C_{inv} m(k_B T / q) \ln \left(1 + e^{q(V_{GS} - V_T + \alpha(k_B T_L / q) F_f) / m k_B T} \right)$$

$$V_T = V_{T0} - \delta V_{DS}$$

$$3) \quad \langle v_x(V_{DS}) \rangle = F_{SAT}(V_{DS}) v_{sat}$$

$$4) \quad F_{SAT}(V_{DS}) = \frac{V_{DS}/V_{DSAT}}{\left[1 + (V_{DS}/V_{DSAT})^\beta \right]^{1/\beta}}$$

$$5) \quad V_{DSAT} = \frac{v_{sat} L}{\mu_n}$$

Only 10 device-specific parameters in this model:

$C_{inv}, V_{T0}, \delta, m, v_{sat}, \mu_n,$
 $L, R_{SD} = R_S + R_D,$
 α, β

Discussion

With this extension (subthreshold to above threshold conduction), the VS model accurately describes modern transistors providing:

- 1) The high-field saturation velocity is viewed as an empirical, fitting parameter.
- 2) The mobility of carriers in the inversion is viewed as an empirical, fitting parameter.
- 3) **But** we will see later, that these empirical parameters can be given a clear, physical interpretation.

Download the MVS model at: <https://nanohub.org/publications/15>

Summary

- 1) Using a semi-empirical expression for Q_n , we have extended the VS model to treat subthreshold to above threshold.
- 2) Excellent fits to measured transistor IV characteristics generally result.
- 3) But the physical understanding of the mobility and saturation velocity *at the nanoscale* needs to be clarified.

Essentials of MOSFETs

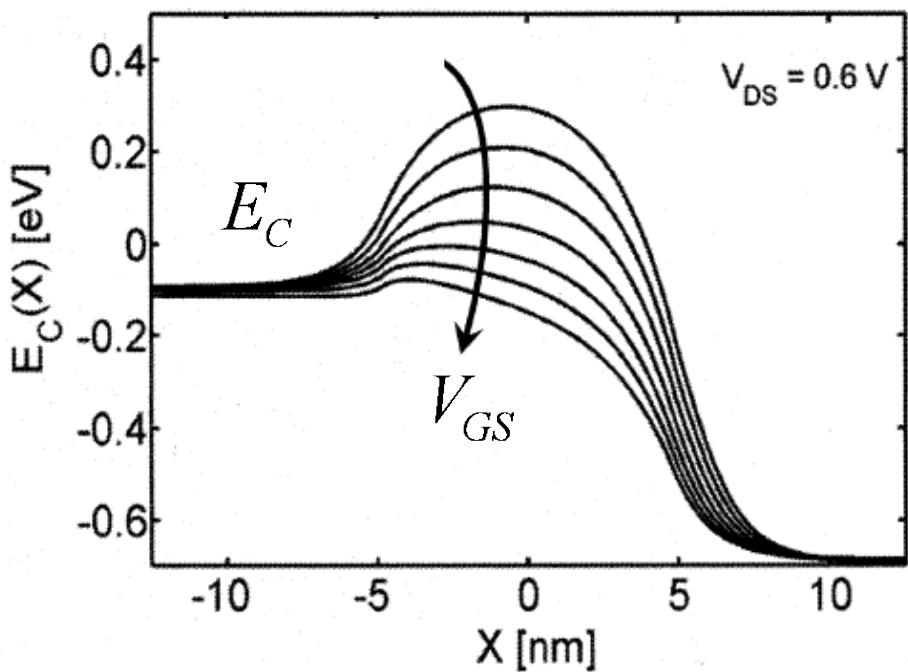
Unit 3: MOS Electrostatics

Lecture 3.10: Unit 3 Recap

Mark Lundstrom

lundstro@purdue.edu
Electrical and Computer Engineering
Purdue University
West Lafayette, Indiana USA

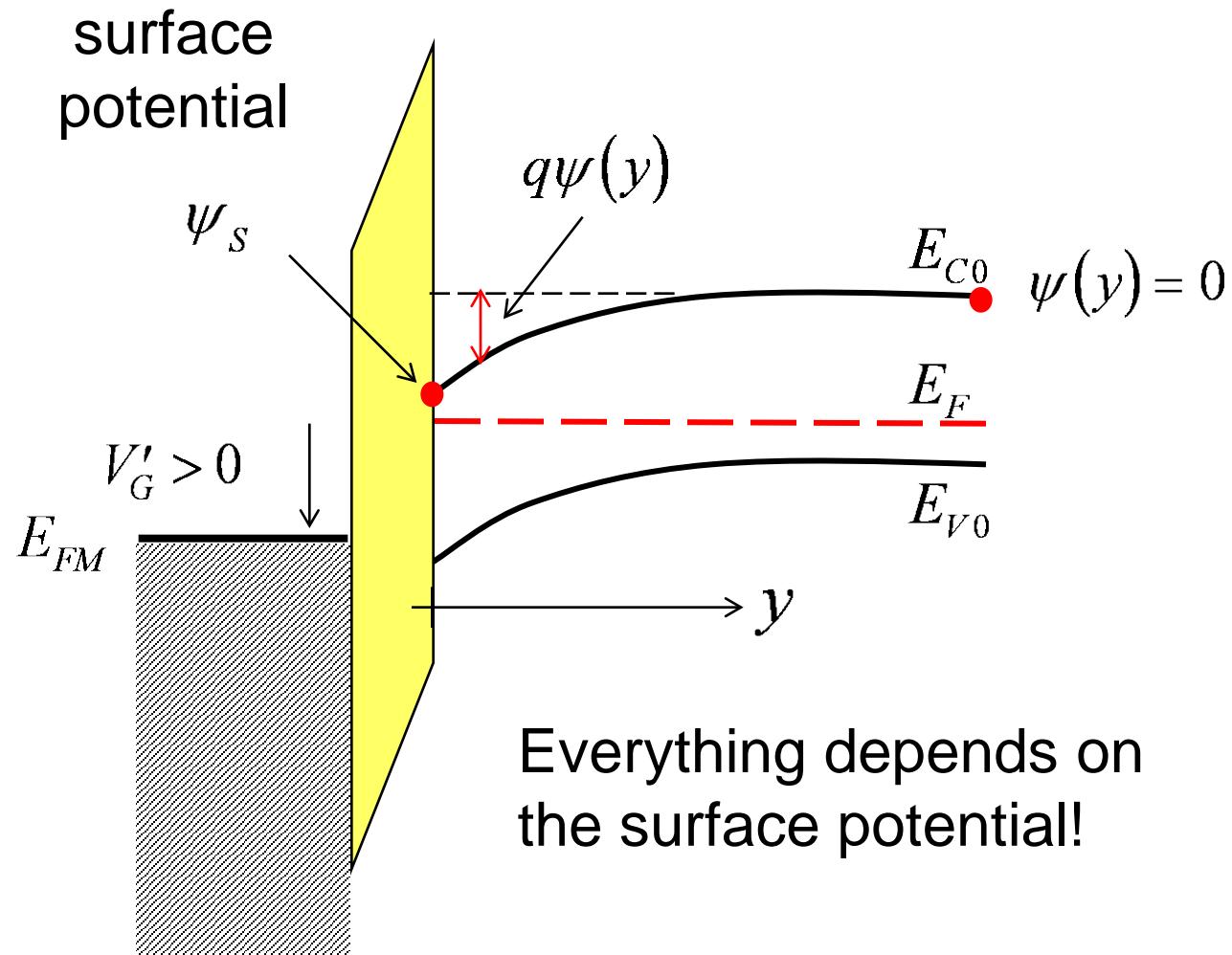
Unit 3



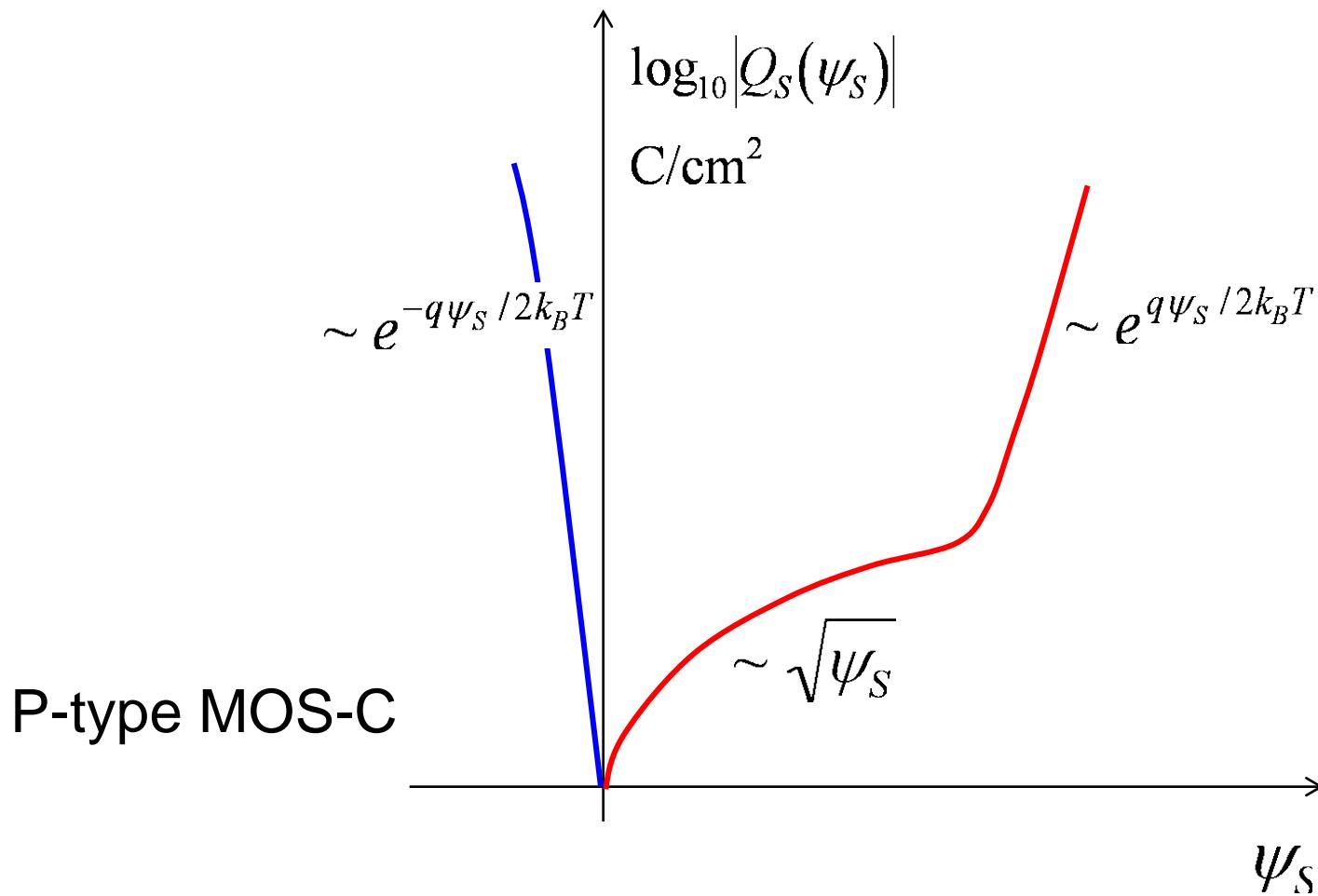
Unit 3: electrostatics

$$I_D/W = |Q_n(V_{GS}, V_{DS})| \langle v_x(V_{GS}, V_{DS}) \rangle$$

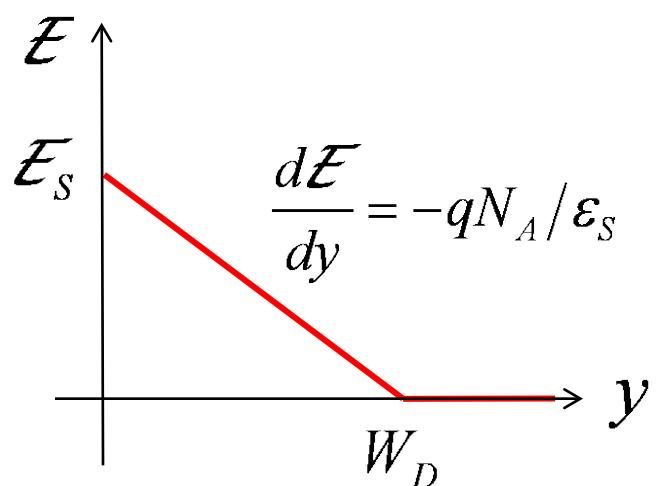
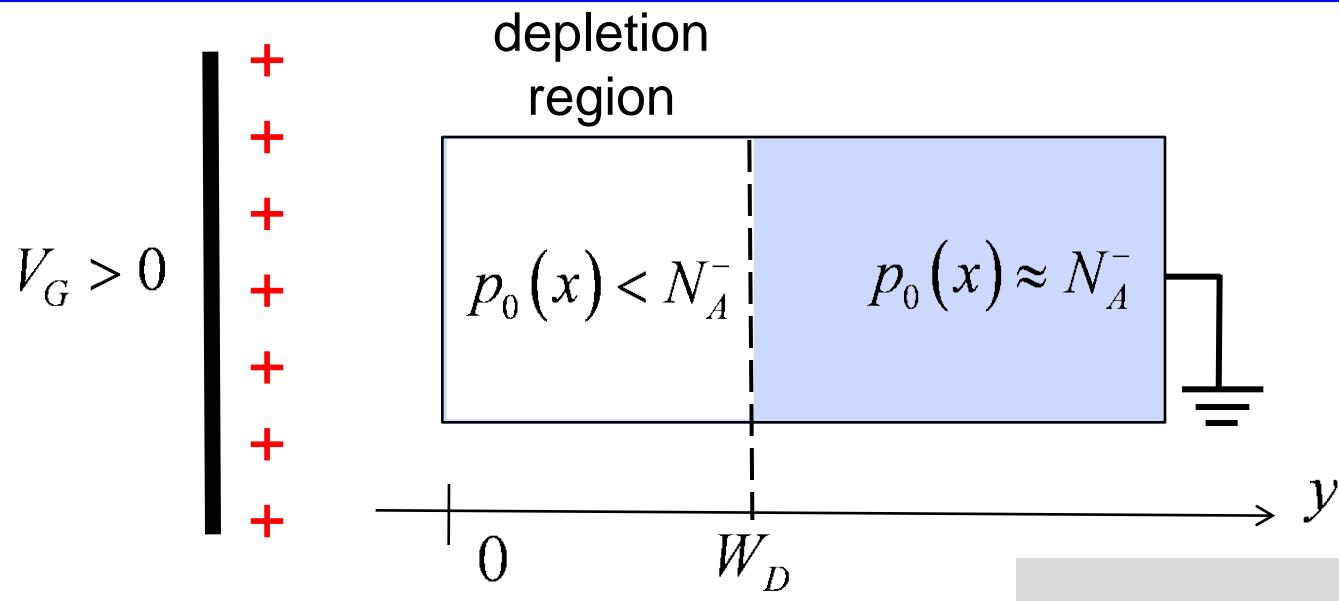
Energy band approach to MOS electrostatics



Charge vs. surface potential



Depletion approximation

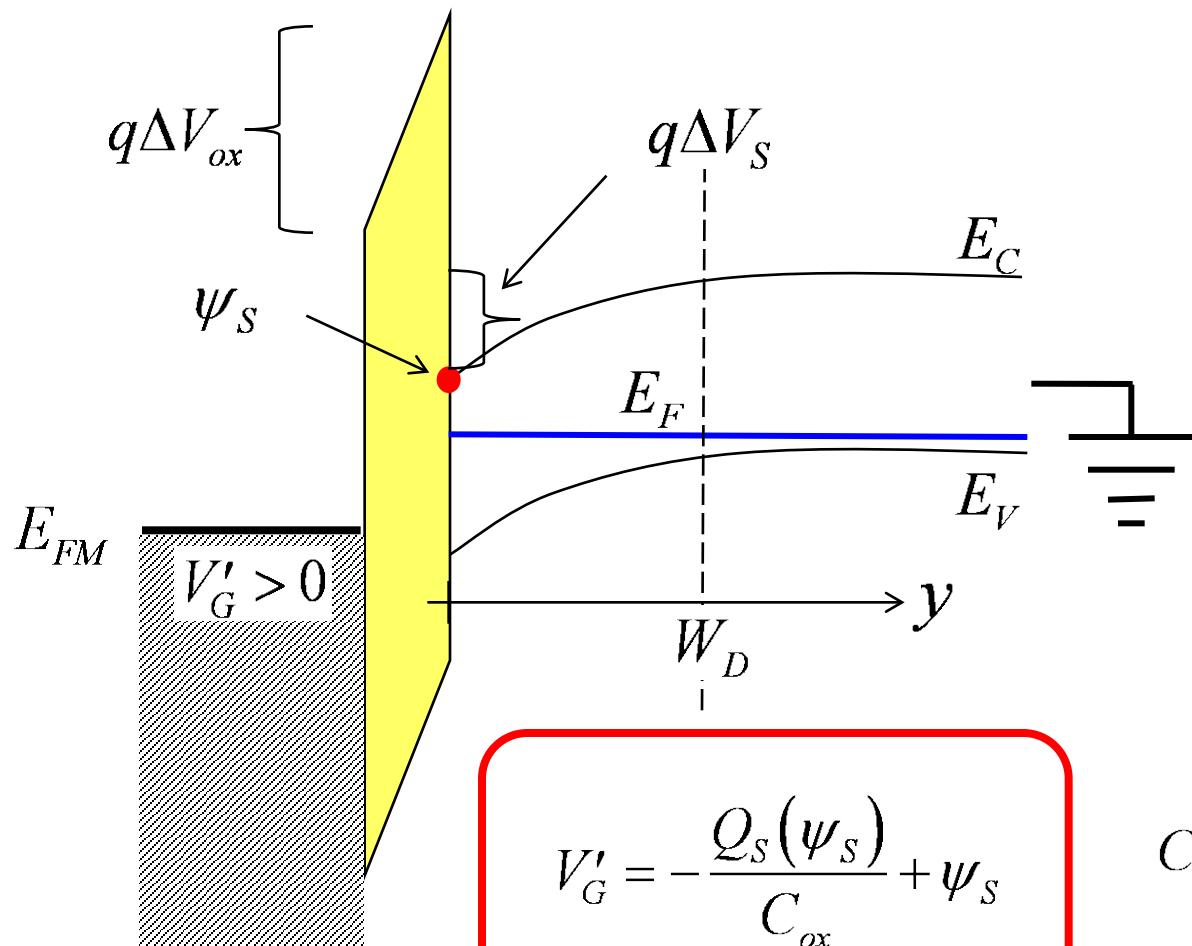


$$W_D = \sqrt{2\epsilon_S \psi_S / qN_A} \text{ m}$$

$$\mathcal{E}_S = \frac{2\psi_S}{W_D} \text{ V/m}$$

$$Q_D = -\sqrt{2qN_A \epsilon_S \psi_S} \text{ C/m}^2$$

Gate voltage and surface potential



$$V_G = \Delta V_{ox} + \psi_s$$

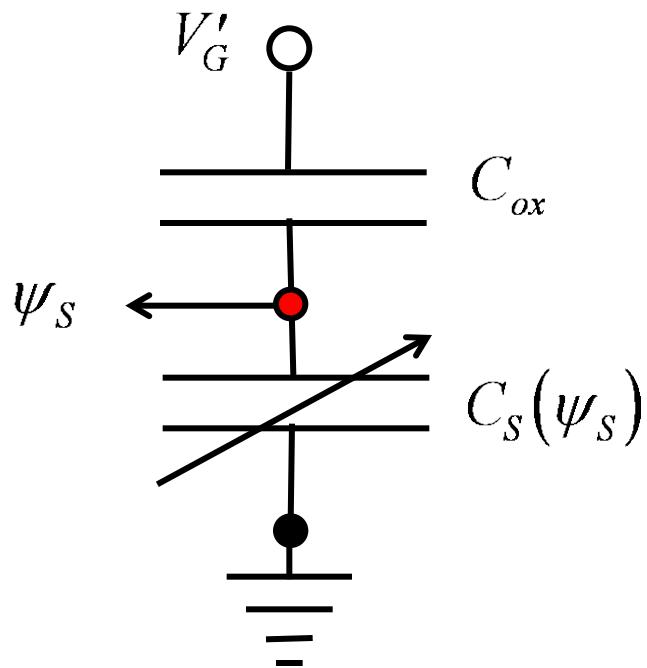
$$V'_G = -\frac{Q_s(\psi_s)}{C_{ox}} + \psi_s$$

$$C_{ox} = \frac{\epsilon_{ox}}{t_{ox}} \text{ F/cm}^2$$

Approximate gate vs. surface potential

$$V'_G = -\frac{Q_s(\psi_s)}{C_{ox}} + \psi_s \quad (\text{exact})$$

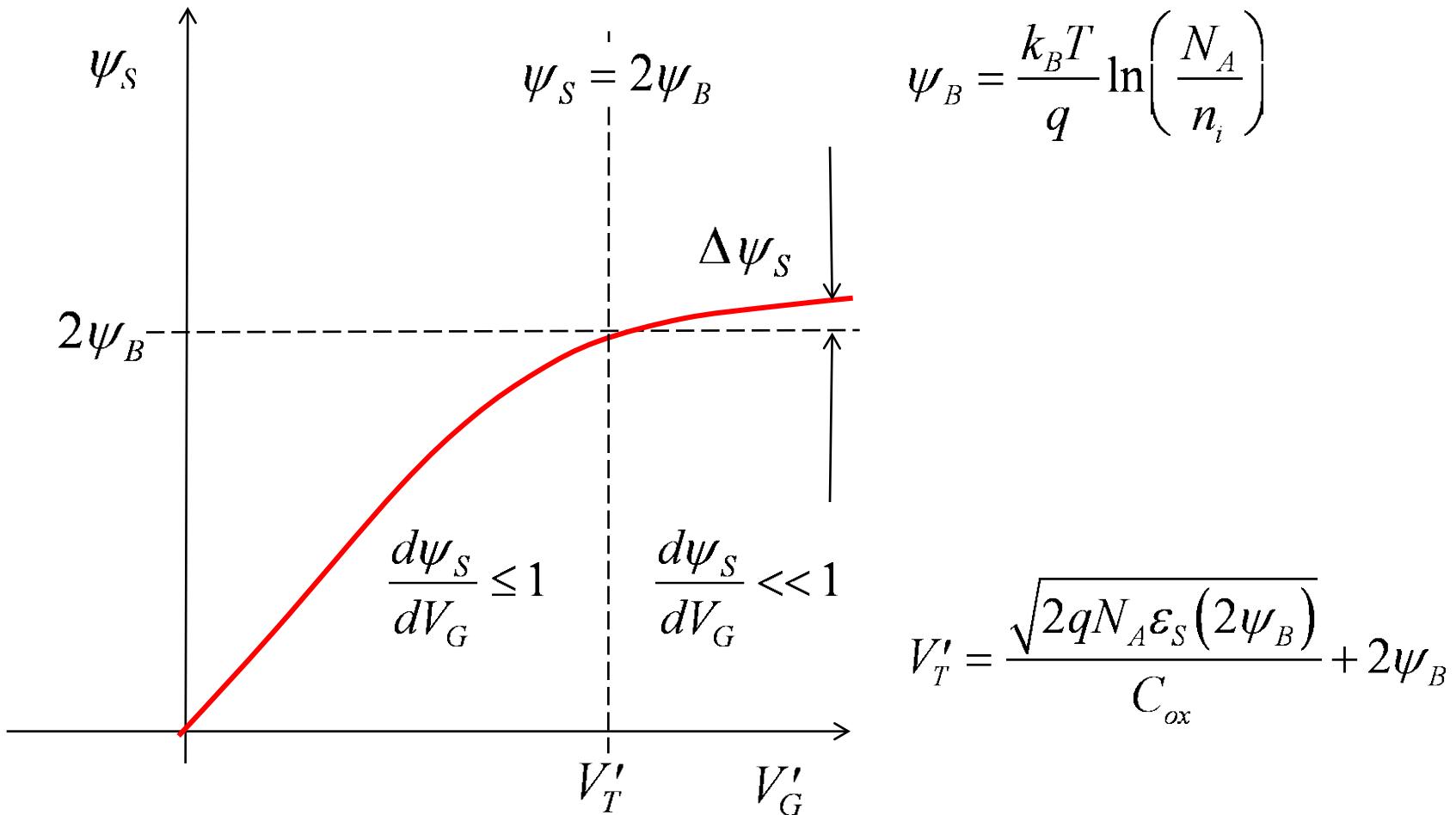
approximate solution



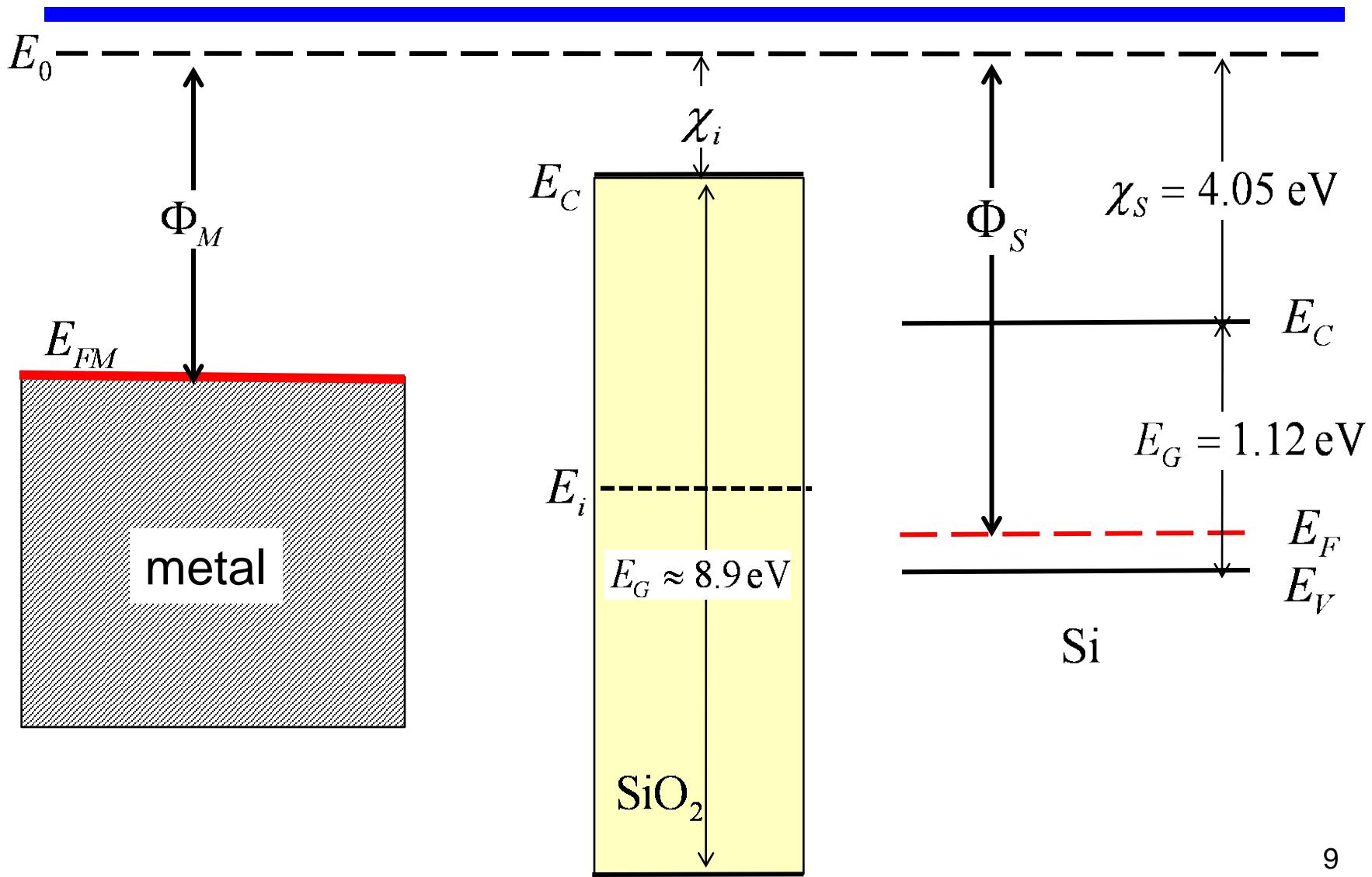
$$\psi_s \approx \frac{V'_G}{m}$$
$$m = 1 + C_D / C_{ox}$$
$$m \geq 1$$

(depletion)

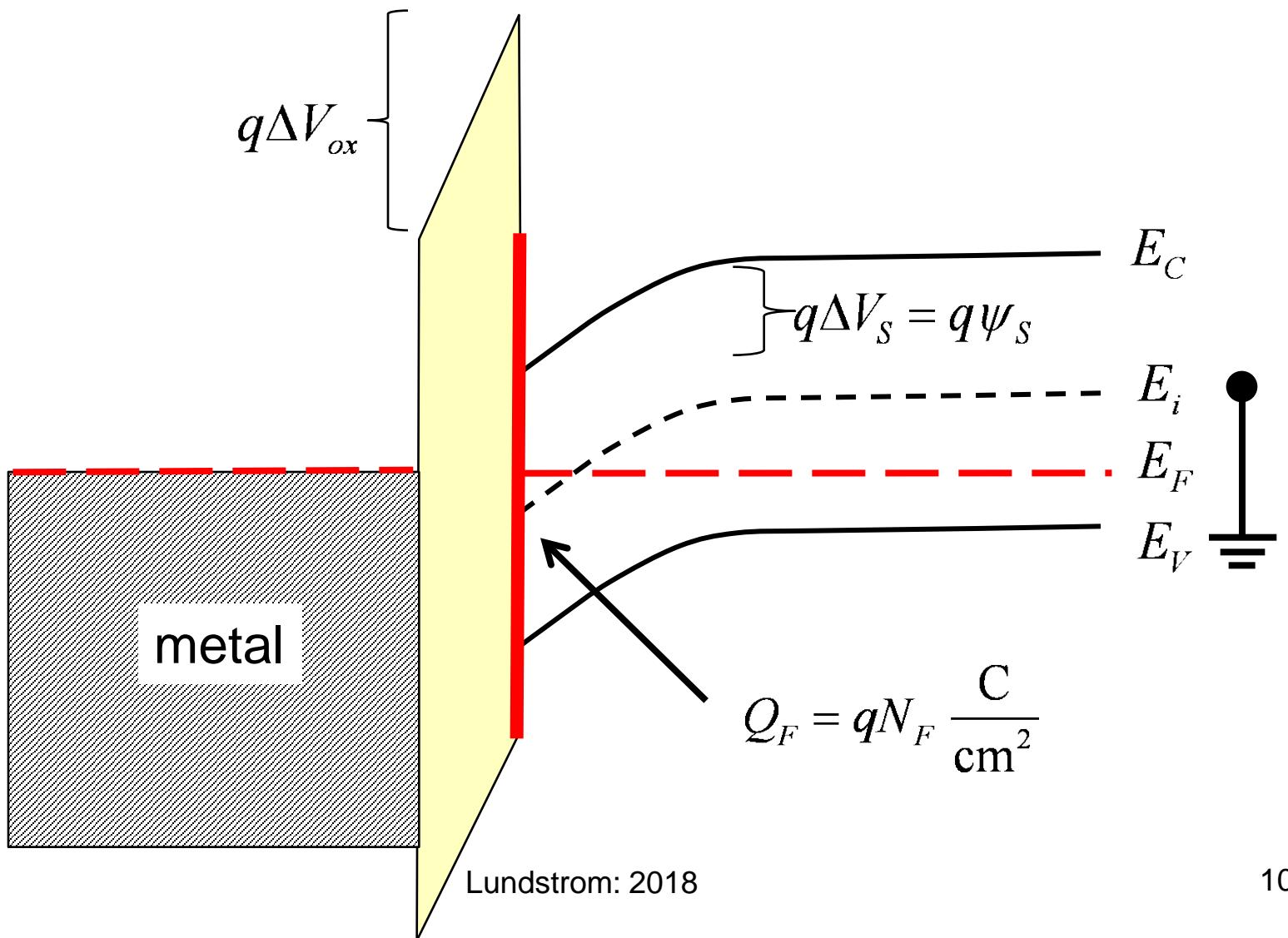
Surface potential vs. gate voltage



Workfunctions



Charge at the oxide-semiconductor interface



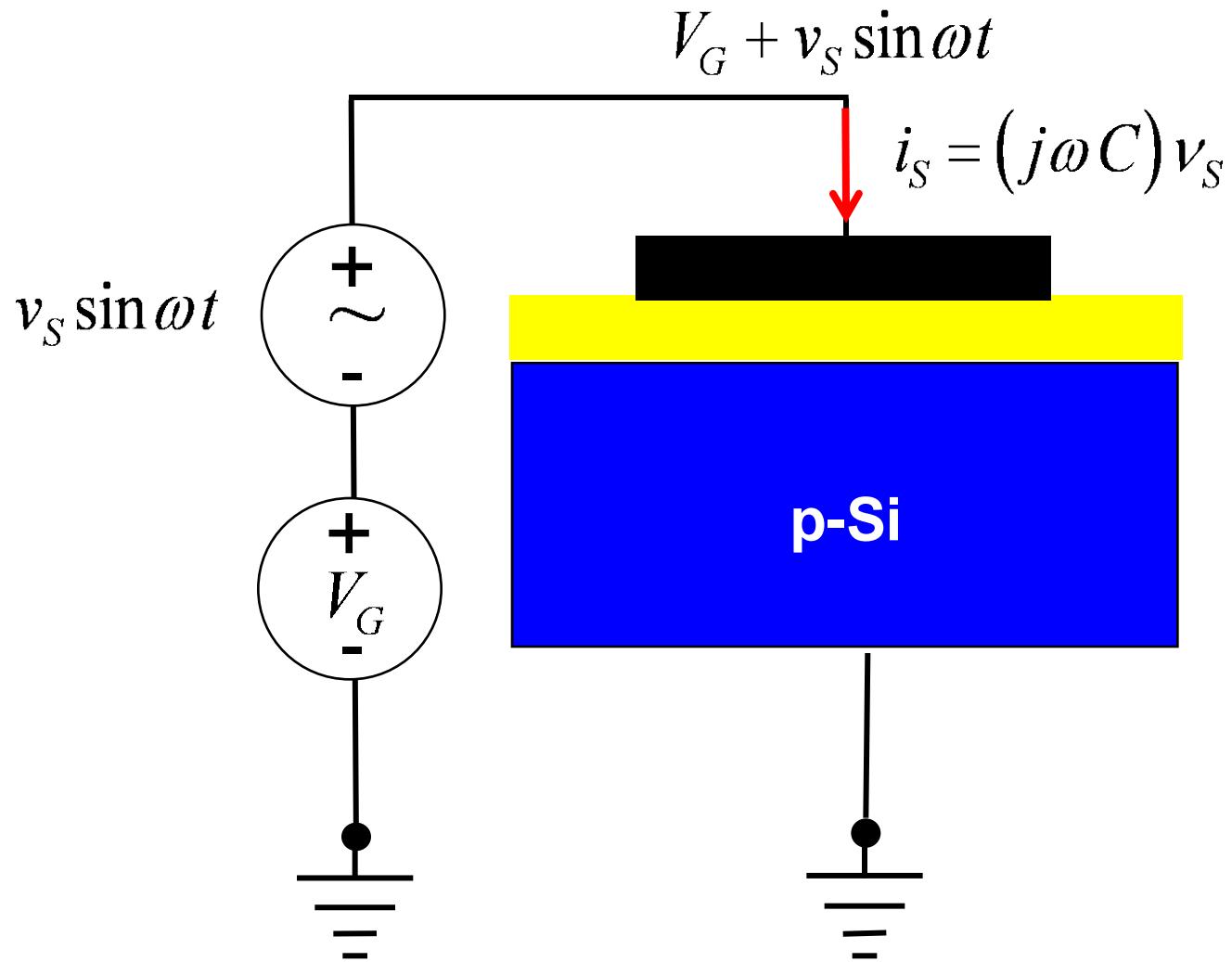
Flatband voltage

$$V_G = V_{FB} - \frac{Q_s(\psi_s)}{C_{ox}} + \psi_s$$

$$V_{FB} = \phi_{ms} - \frac{Q_F}{C_{ox}}$$

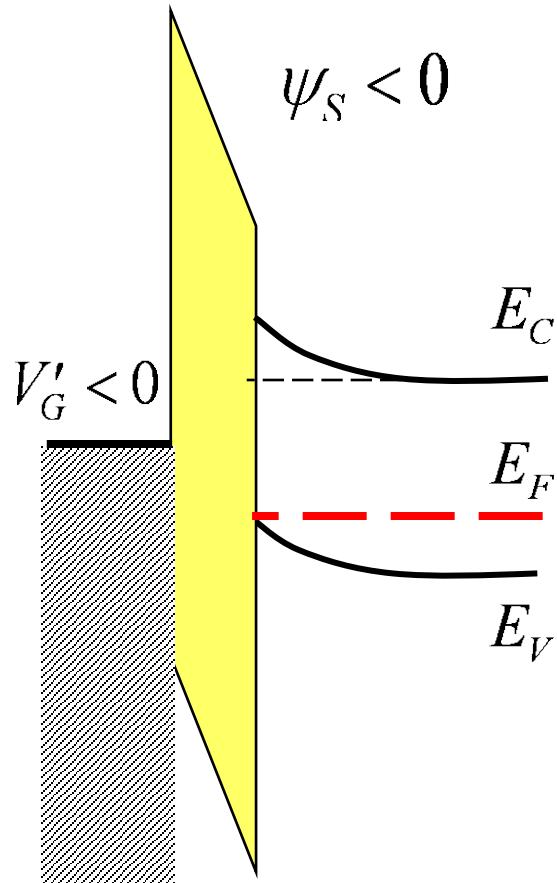
for $V_G = 0$, $\psi_s \neq 0$

MOS small signal capacitance

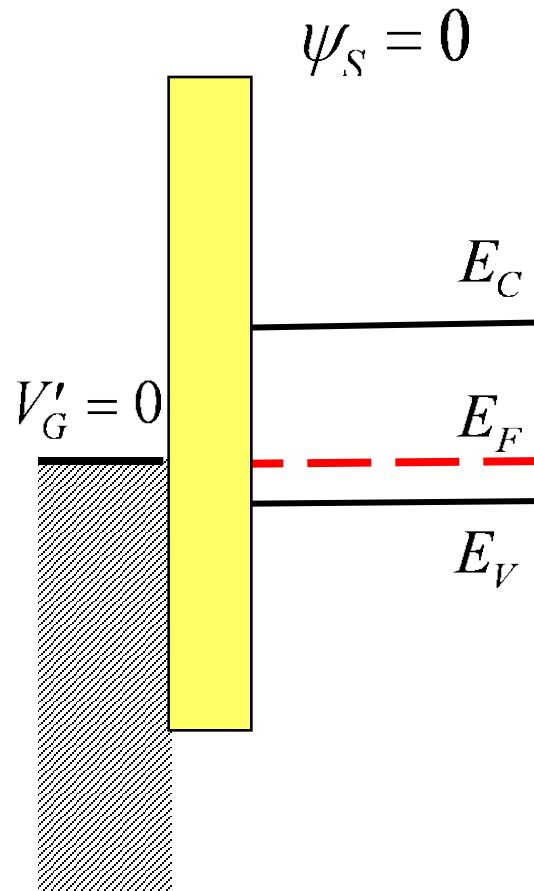


DC bias from accumulation to inversion

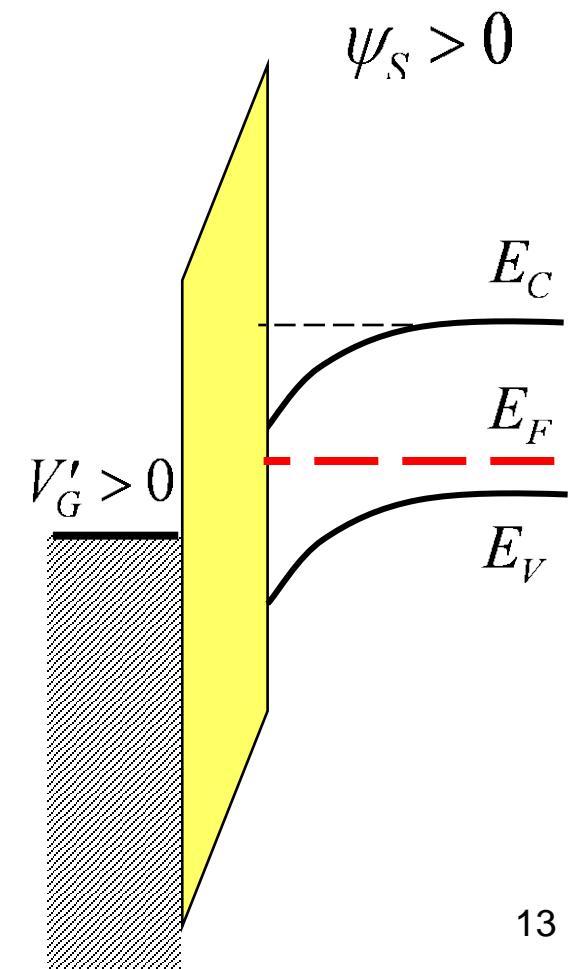
accumulation



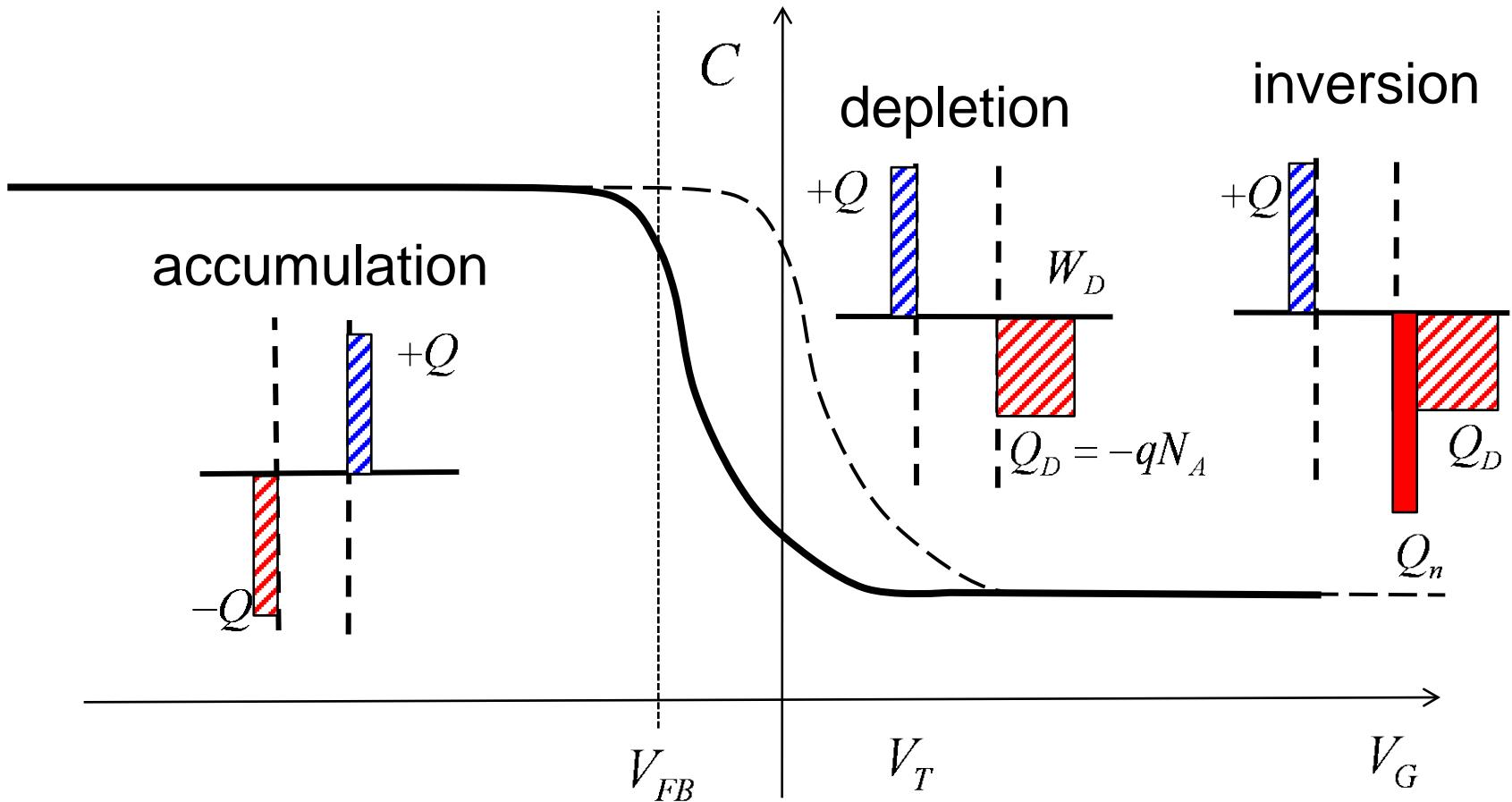
flat band



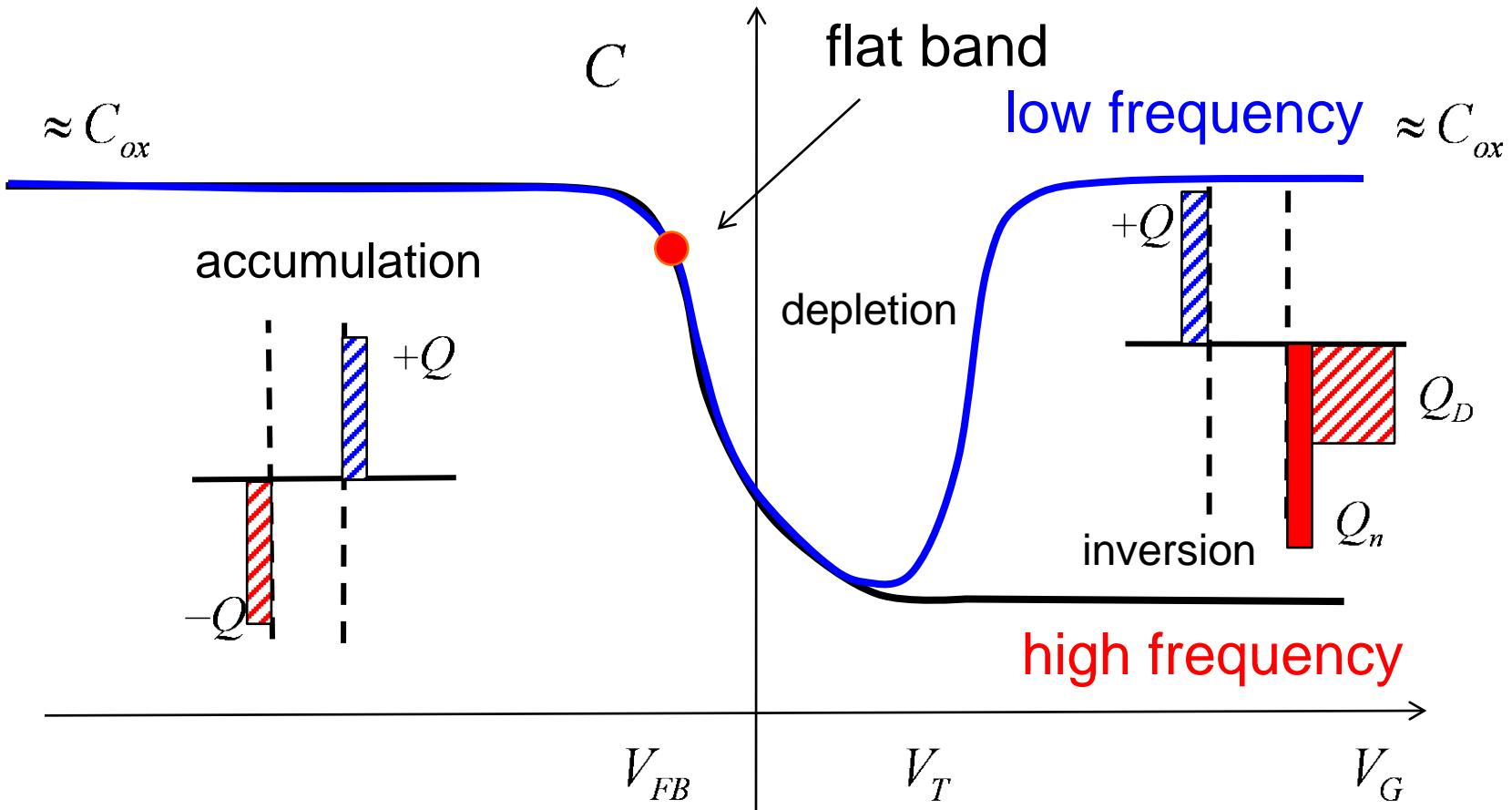
depletion/
inversion



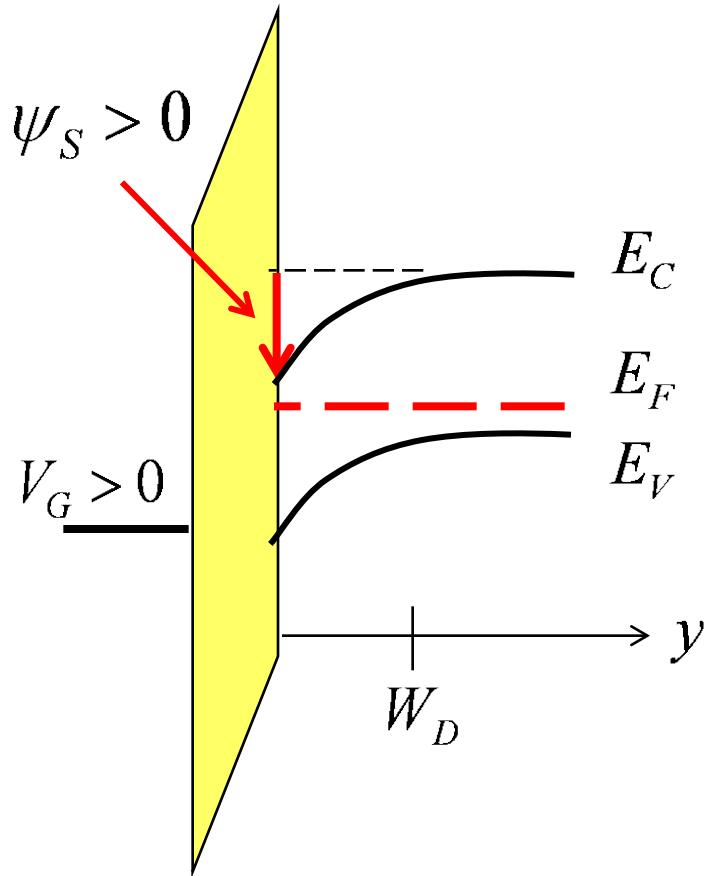
MOS high frequency CV



High frequency vs. low frequency CV



Mobile charge



The mobile charge carries the current.

$$Q_n = -q \int_0^{\infty} n(y) dy \text{ C/cm}^2$$

(electrons in a P-type semiconductor)

Expect: $Q_n \propto e^{q\psi_S/k_B T}$

Mobile charge vs. surface potential

Bulk semiconductor:

$$\psi_s < 2\psi_b : \quad Q_n(\psi_s) \approx - \left(\frac{n_i^2 k_B T / N_A}{\sqrt{(2qN_A \psi_s / \epsilon_s)^{1/2}}} \right) e^{q\psi_s / k_B T}$$

$$\psi_s > 2\psi_b : \quad Q_n(\psi_s) = - \sqrt{2\epsilon_s k_B T (n_i^2 / N_A)} \times e^{q\psi_s / 2k_B T}$$

Fully depleted, ultra thin body:

$$\psi_s > 0 : \quad Q_n(\psi_s) = -qn_{S0} e^{q\psi_s / k_B T}$$

Mobile charge vs. gate voltage

bulk

$$V_G \ll V_T : \quad Q_n(V_G) = -(m-1)C_{ox} \left(\frac{k_B T}{q} \right) e^{q(V_G - V_T)/mk_B T}$$

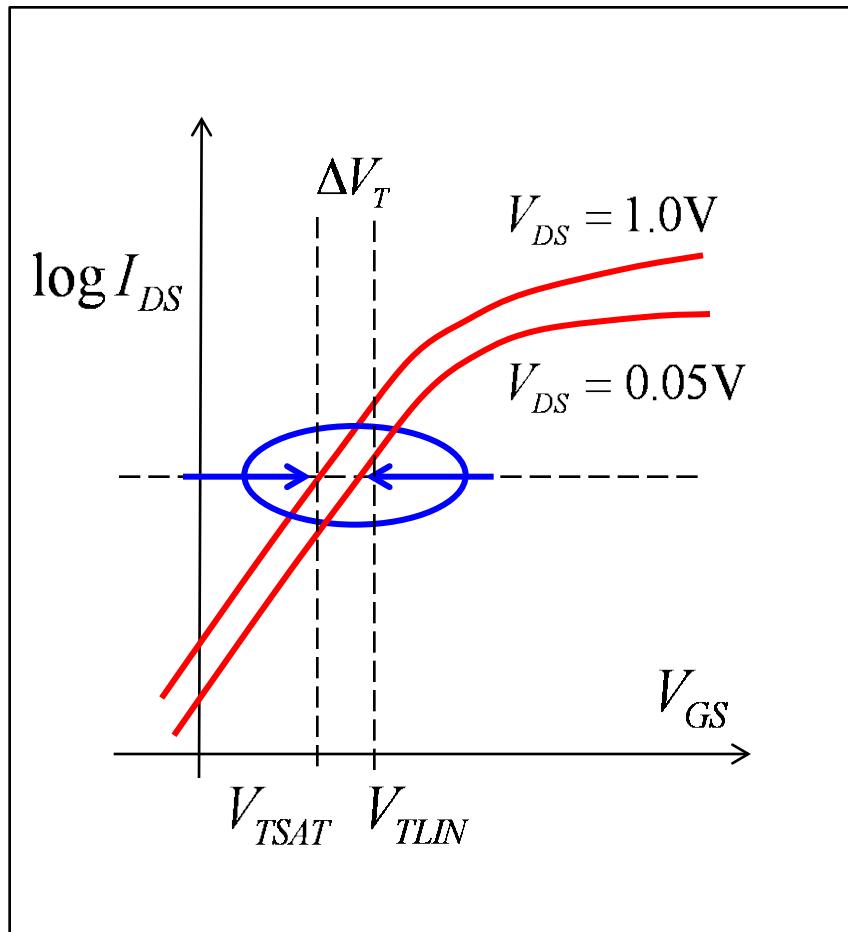
$$V_G \gg V_T : \quad Q_n = -C_{inv}(V_G - V_T) \quad C_{inv} < C_{ox}$$

FD UTB

$$V_G \ll V_T : \quad Q_n(V_G) = -C_Q \left(\frac{k_B T}{q} \right) e^{q(V_G - V_T)/k_B T} \quad m = 1$$

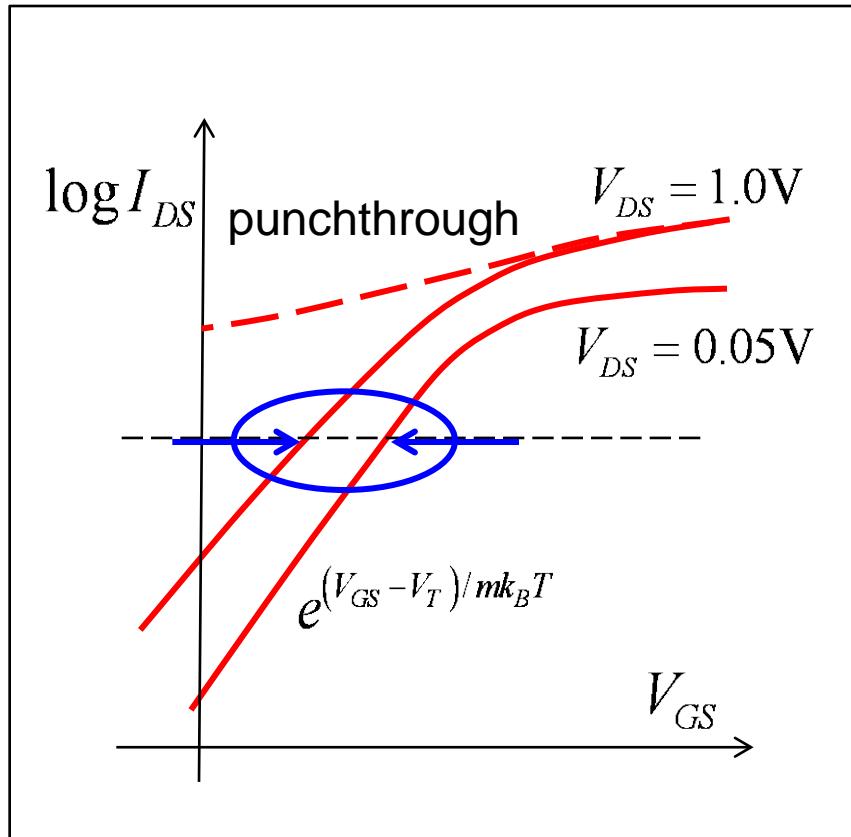
$$V_G \gg V_T : \quad Q_n(V_G) = -C_{inv}(V_G - V_T) \quad C_{inv} < 2C_{ox}$$

2D electrostatics



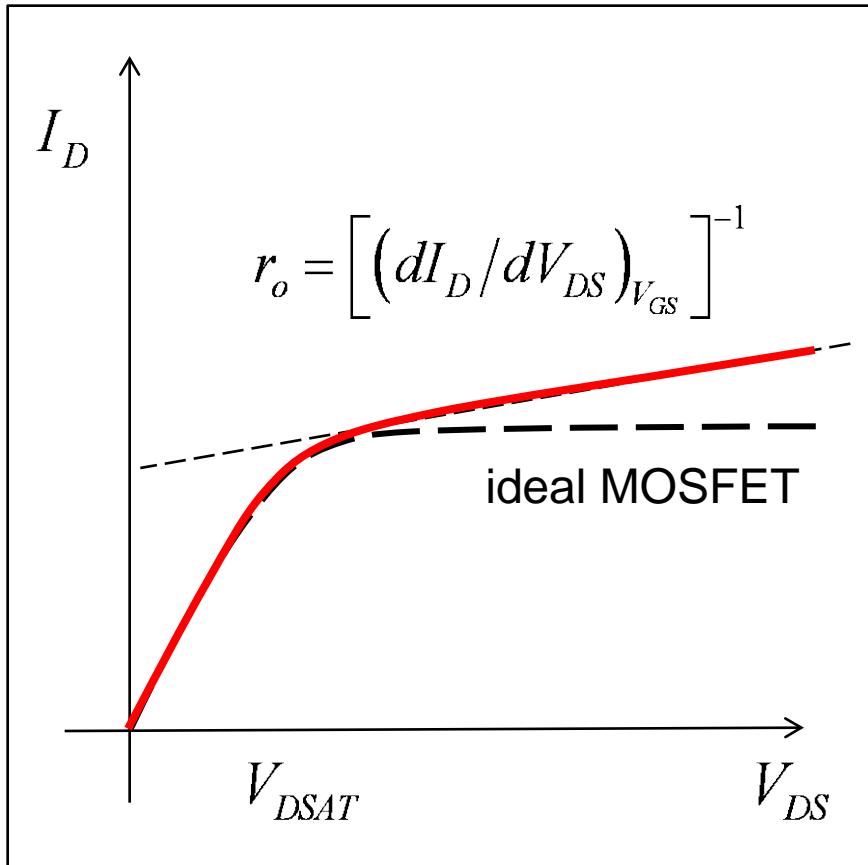
- 1) Threshold voltage decreases as the drain voltage increases
- 2) Threshold voltage decreases as the channel length decreases
- 3) DIBL increases as channel length decreases

2D electrostatics



- 1) SS may increase as the drain voltage increases
- 2) SS may increase as the channel length decreases
- 3) In severe cases, the device may “punch through”

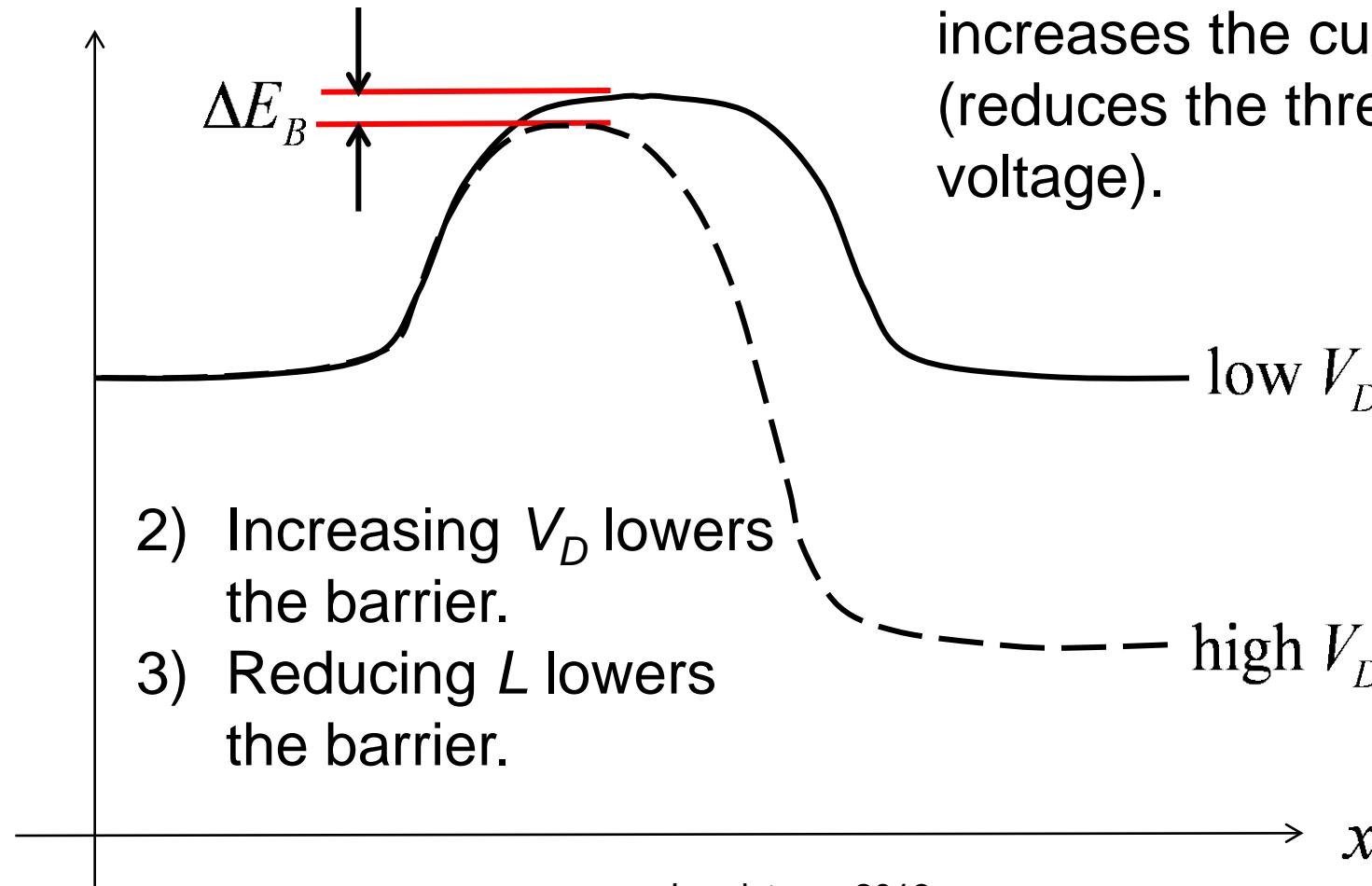
2D electrostatics



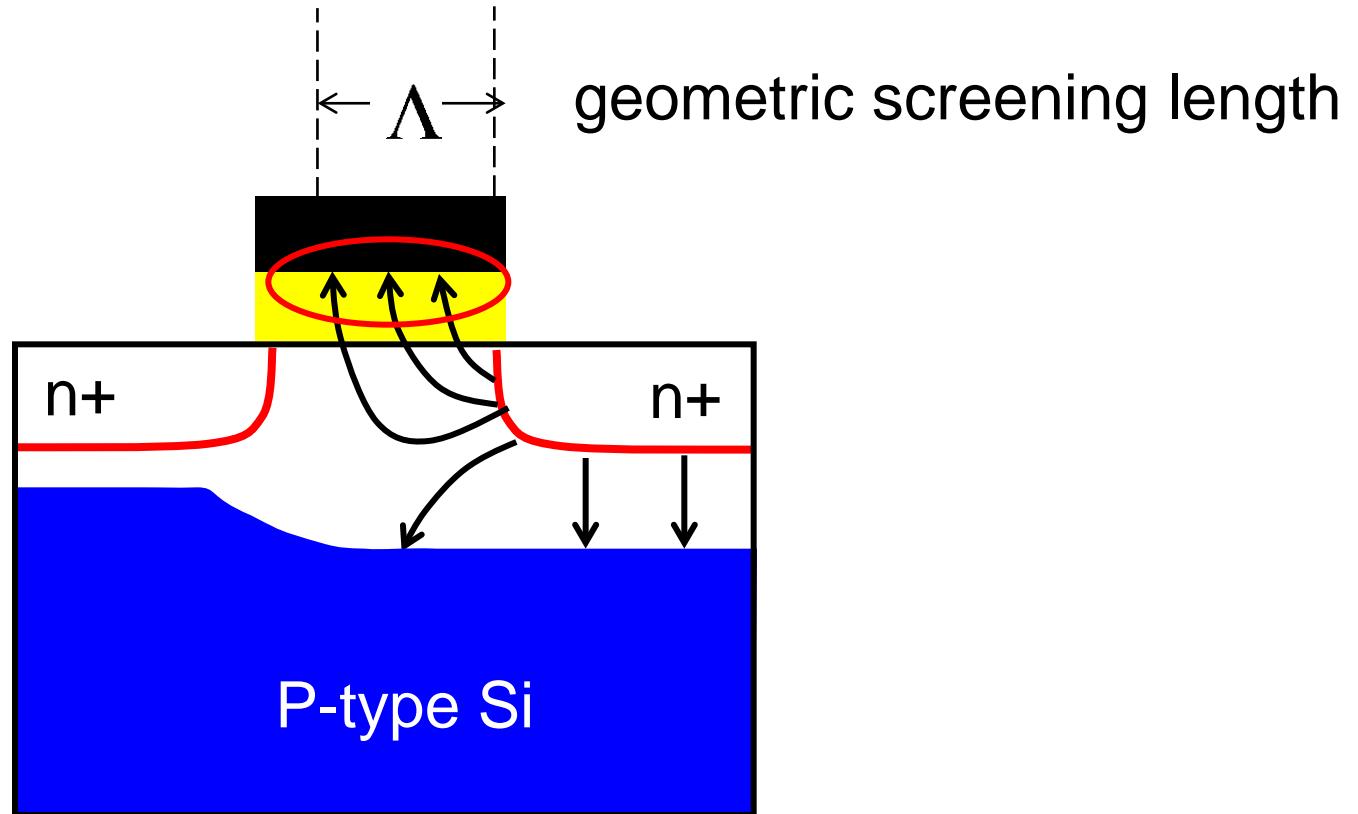
- 1) Output resistance decreases as channel length decreases.

Barrier lowering view of 2D electrostatics

$E_C(x, y = 0)$

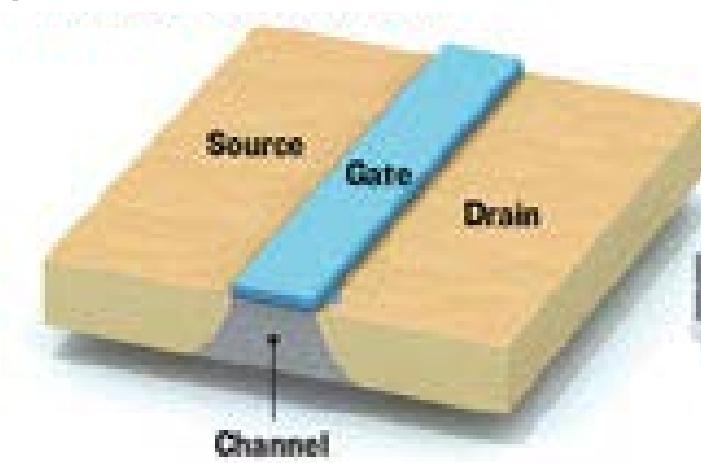


Controlling 2D electrostatics

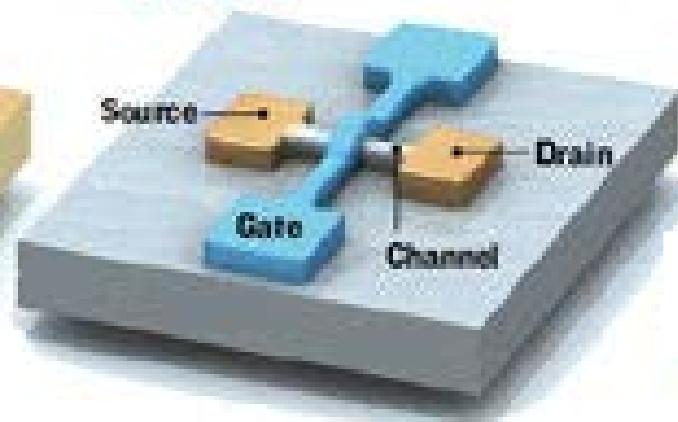


2D electrostatics

planar transistor



FinFET



“Transistors go Vertical,” *IEEE Spectrum*, Nov. 2007.

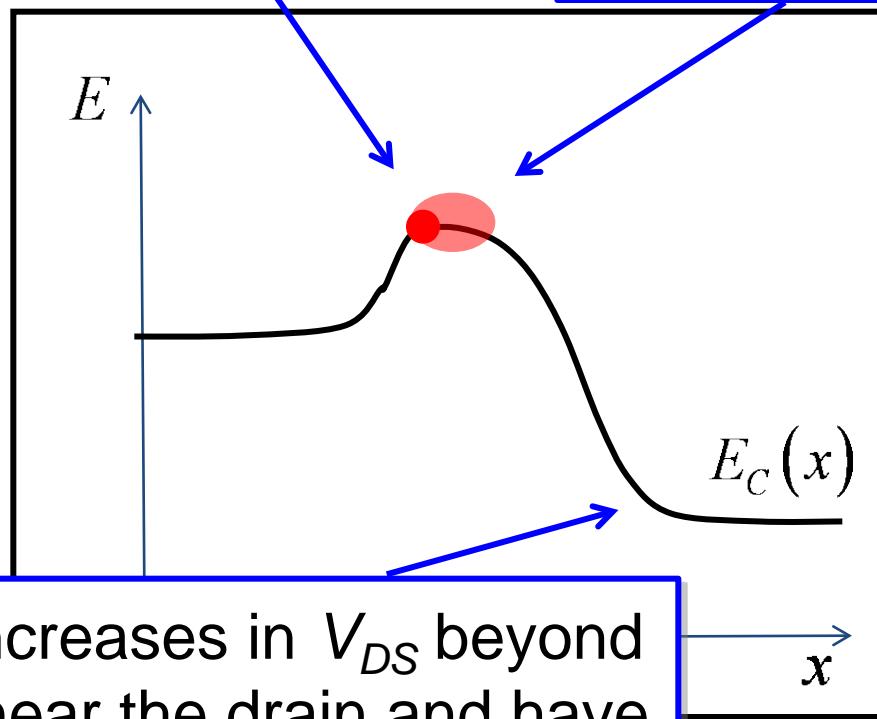
See also: “Integrated Nanoelectronics of the Future,” Robert Chau, Brian Doyle, Suman Datta, Jack Kavalieros, and Kevin Zhang, *Nature Materials*, **6**, 2007

“Well-tempered MOSFET”

1) $Q_n(0) \approx -C_{inv}(V_{GS} - V_T)$

2) region under strong control of gate ($m \sim 1$)

$V_T = V_{T0} - \delta V_{DS}$
 $m = \text{constant}$



3) Additional increases in V_{DS} beyond V_{DSAT} drop near the drain and have a **small effect** on I_D (small DIBL)

Level 0 VS model

$$1) \quad I_D/W = |Q_n(V_{GS})| \langle v_x(V_{DS}) \rangle$$

$$2) \quad Q_n(V_{GS}, V_{DS}) = -C_{ox}(V_{GS} - V_T) \quad (V_{GS} > V_T)$$

$$V_T = V_{T0} - \delta V_{DS}$$

$$Q_n(V_{GS}) = 0 \quad (V_{GS} \leq V_T)$$

$$3) \quad \langle v(V_{DS}) \rangle = F_{SAT}(V_{DS}) v_{sat}$$

$$4) \quad F_{SAT}(V_{DS}) = \frac{V_{DS}/V_{DSAT}}{\left[1 + (V_{DS}/V_{DSAT})^\beta\right]^{1/\beta}}$$

$$5) \quad V_{DSAT} = v_{sat} L / \mu_n$$

There are only 8 device-specific parameters in this model:

$$C_{ox}, V_{T0}, \delta, v_{sat}, \mu_n, L \\ R_{SD} = R_S + R_D, \beta$$

Level 1 VS Model

$$1) \quad I_D/W = |Q_n(V_{GS}, V_{DS})| \langle v_x(V_{DS}) \rangle$$

$$2) \quad Q_n(V_{GS}, V_{DS}) = -C_{inv} m(k_B T / q) \ln \left(1 + e^{q(V_{GS} - V_T + \alpha(k_B T_L / q) F_f) / m k_B T} \right)$$

$$V_T = V_{T0} - \delta V_{DS}$$

$$3) \quad \langle v_x(V_{DS}) \rangle = F_{SAT}(V_{DS}) v_{sat}$$

$$4) \quad F_{SAT}(V_{DS}) = \frac{V_{DS}/V_{DSAT}}{\left[1 + (V_{DS}/V_{DSAT})^\beta \right]^{1/\beta}}$$

$$5) \quad V_{DSAT} = \frac{v_{sat} L}{\mu_n}$$

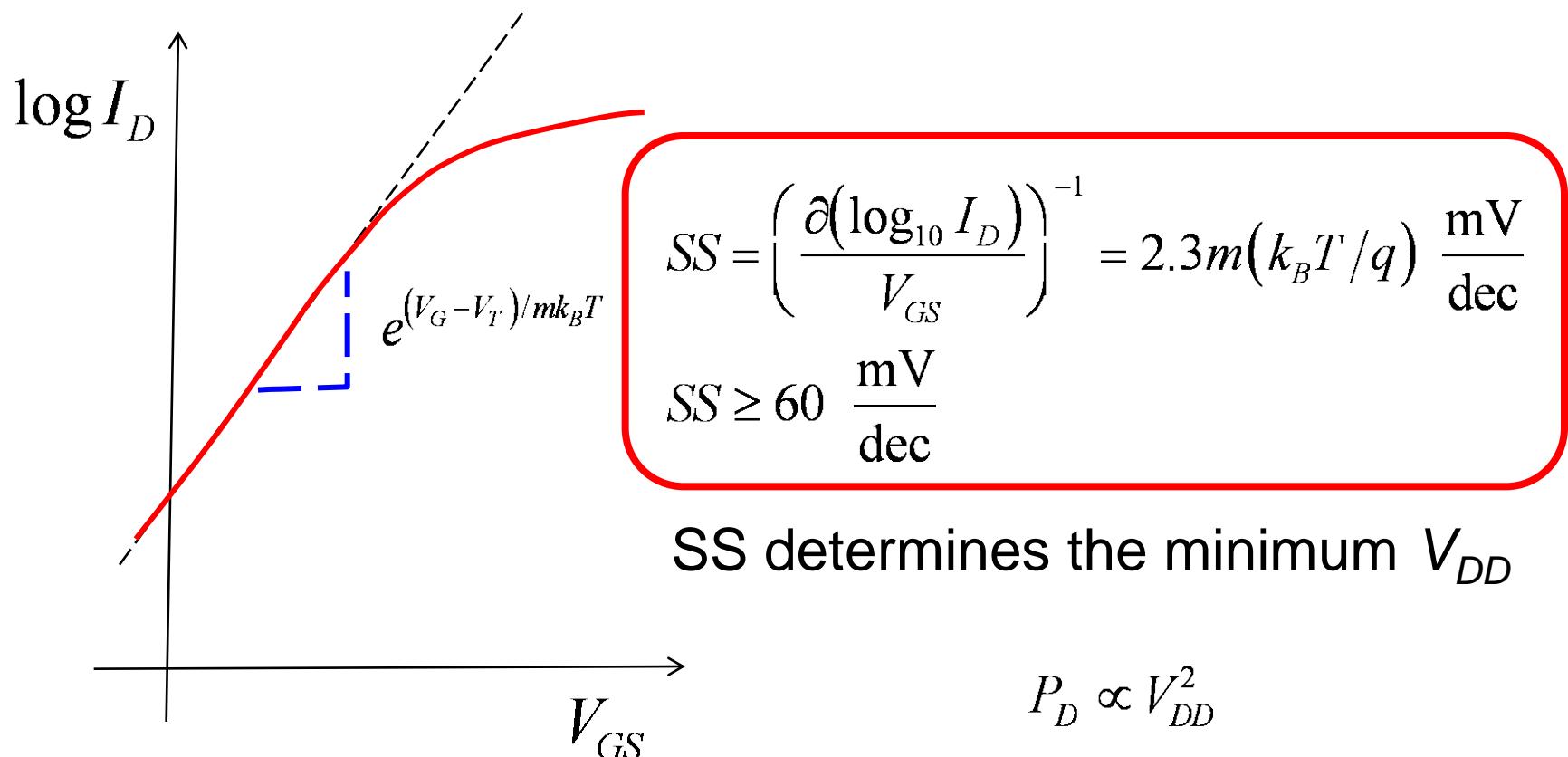
Only 10 device-specific parameters in this model:

$$C_{inv}, V_{T0}, \delta, m, v_{sat}, \mu_n,$$

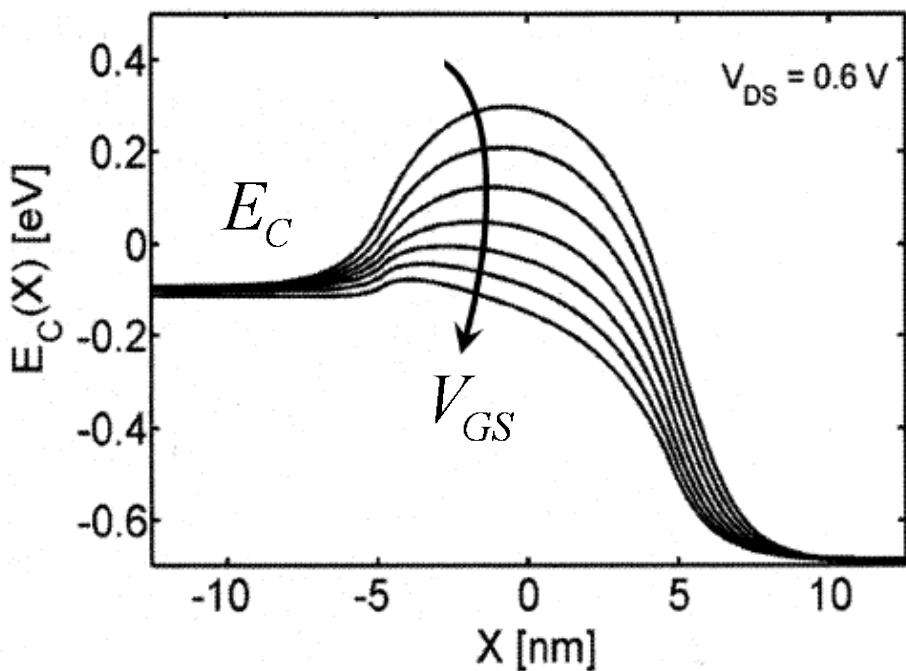
$$L, R_{SD} = R_S + R_D,$$

$$\alpha, \beta$$

Subthreshold swing



On to Unit 4



Unit 3: electrostatics

$$I_D/W = -Q_n(V_{GS}, V_{DS}) \langle v_x(V_{GS}, V_{DS}) \rangle$$

Unit 4: transport

Essentials of MOSFETs

Unit 4: Transmission Theory of the MOSFET

Lecture 4.1: The Landauer Approach

Mark Lundstrom

lundstro@purdue.edu
Electrical and Computer Engineering
Purdue University
West Lafayette, Indiana USA

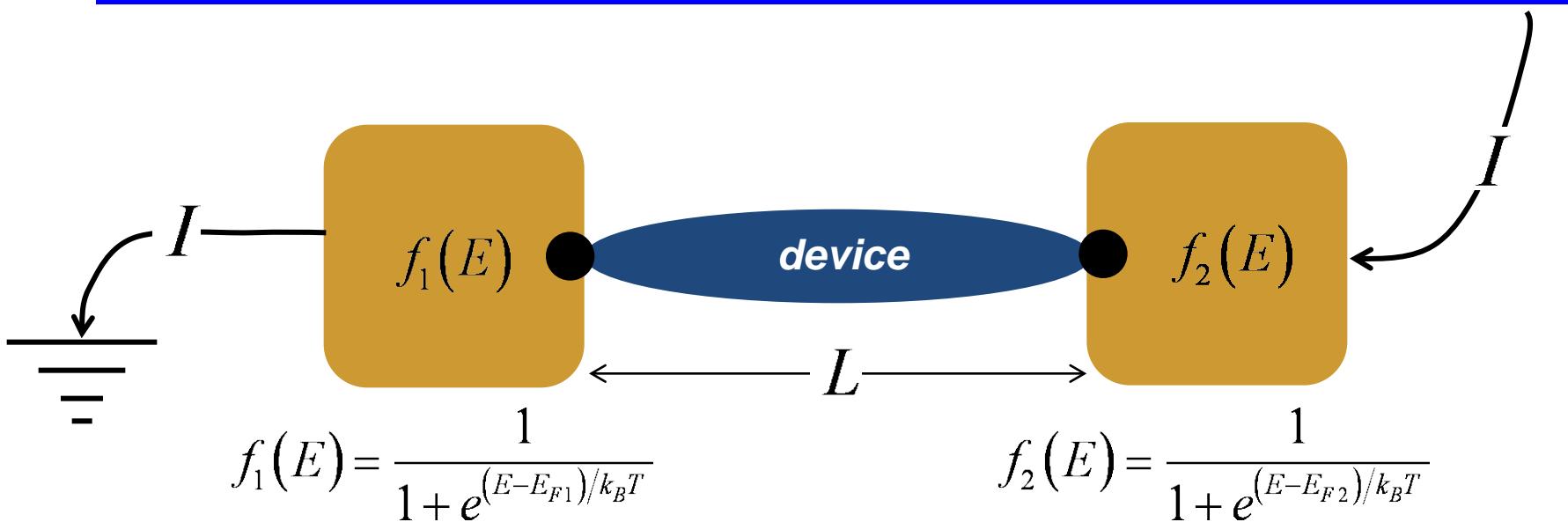
Introduction

The Landauer Approach is a simple and very physical way to describe electron transport from the ballistic to diffusive limits (i.e. from short to long channel MOSFETs).

Our description of this approach will be very simple and intuitive. Those who want a deeper understanding, should consult:

Supriyo Datta, *Lessons from Nanoelectronics*, 2nd Ed., Part A: Basic Concepts, World Scientific Publishing Co., Singapore, 2017.

Current in a nano device



How does the current that flows in contact 2, depend on the voltages on the two contacts?

Current

$$I = \frac{2q}{h} \int T(E) M(E) (f_1(E) - f_2(E)) dE$$

Fundamental
constants

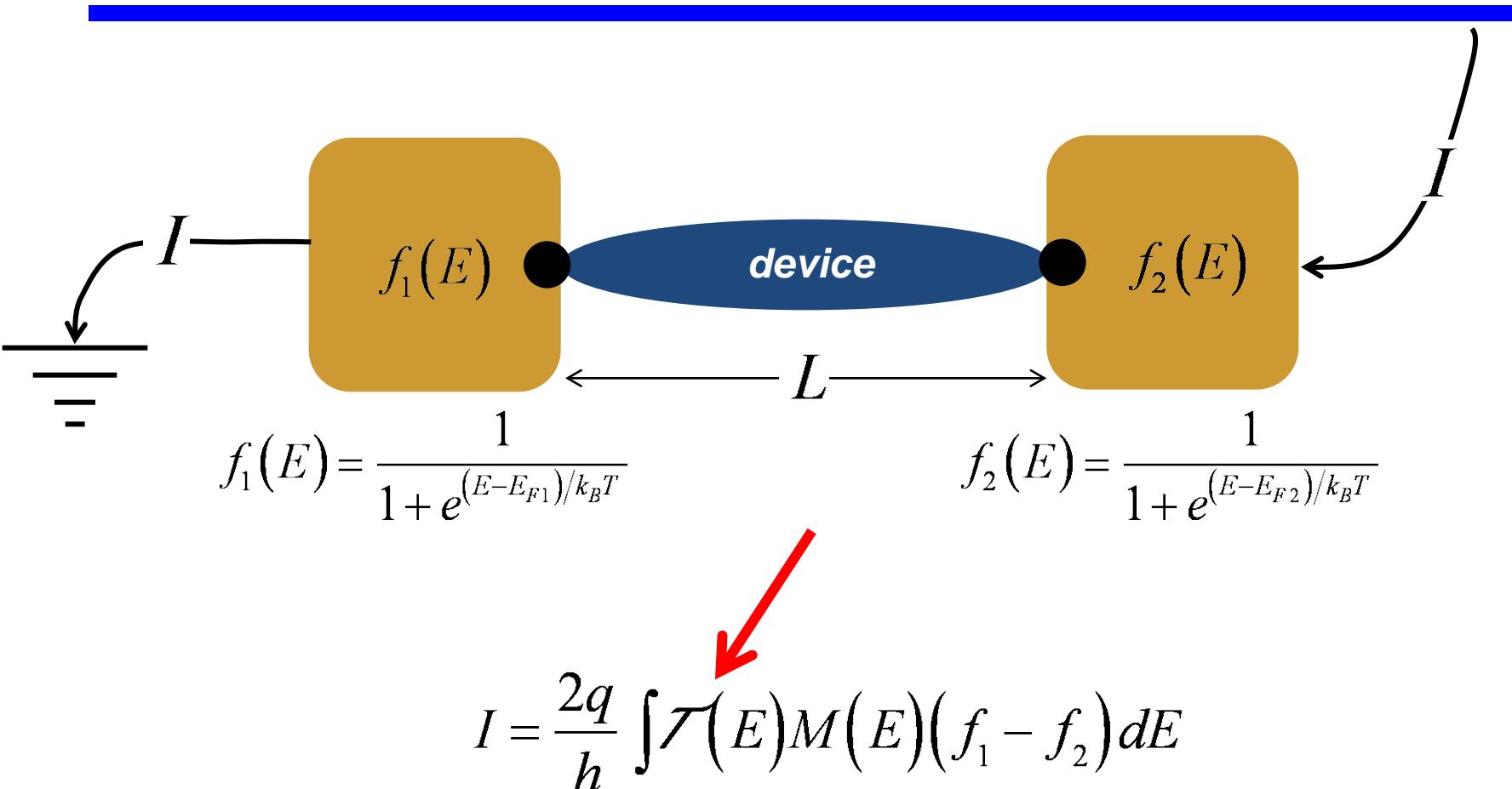
Transmission:
 $0 < T(E) \leq 1$

No. of
Channels

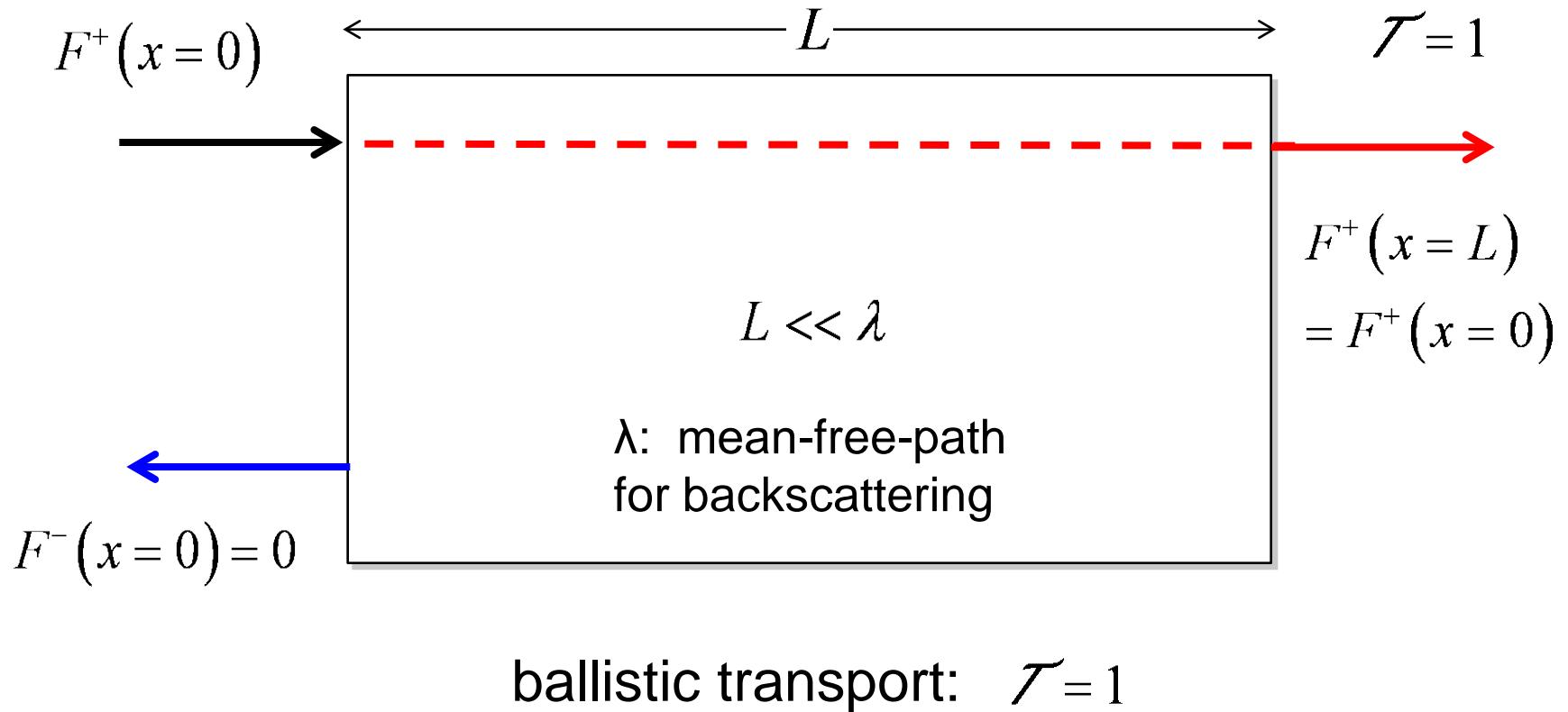
Ideal
“Landauer”
contacts

Can be derived from rigorous transport theory (the Boltzmann equation), but this expression is intuitively easy to understand.

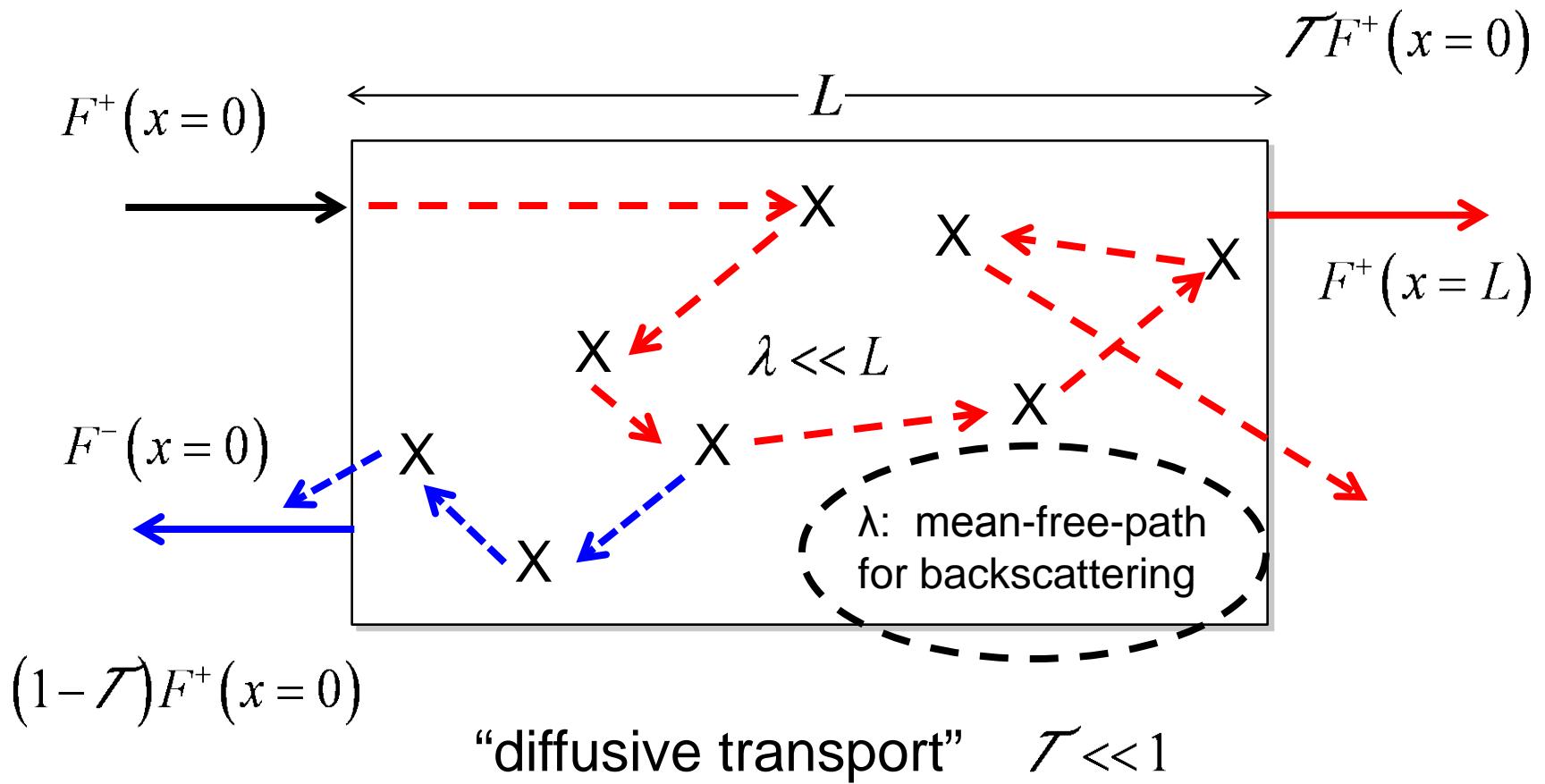
What is transmission?



Transmission (ballistic)



Transmission (diffusive)



Transmission (general)

$$\mathcal{T}(E) = \frac{\lambda(E)}{\lambda(E) + L}$$

λ is the “mean-free-path for backscattering”

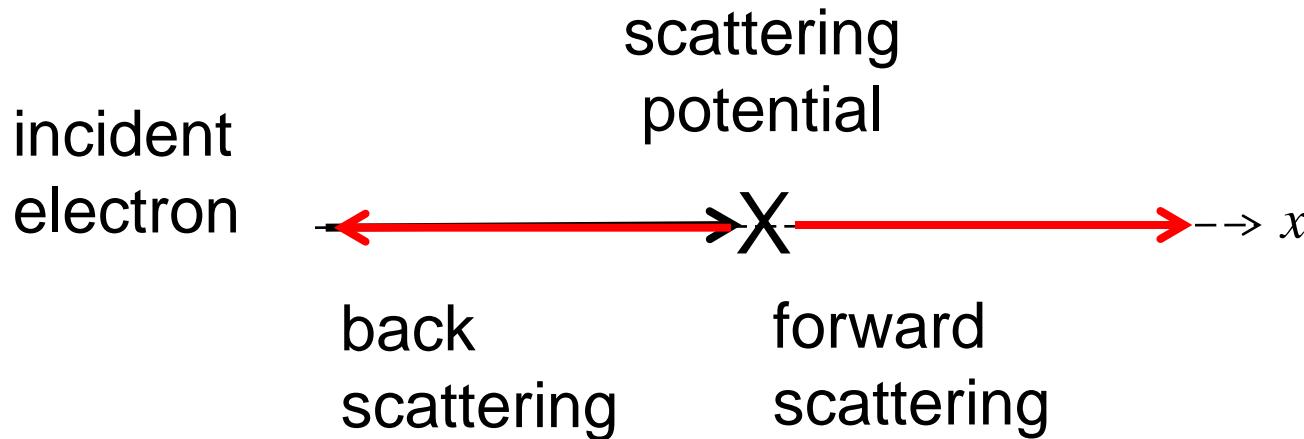
This expression can be derived with relatively few assumptions.

1) Diffusive: $L \gg \lambda$ $\mathcal{T} = \frac{\lambda}{L} \ll 1$

2) Ballistic: $L \ll \lambda$ $\mathcal{T} = 1$

$$\lambda(E) \neq v(E)\tau(E) = \Lambda$$

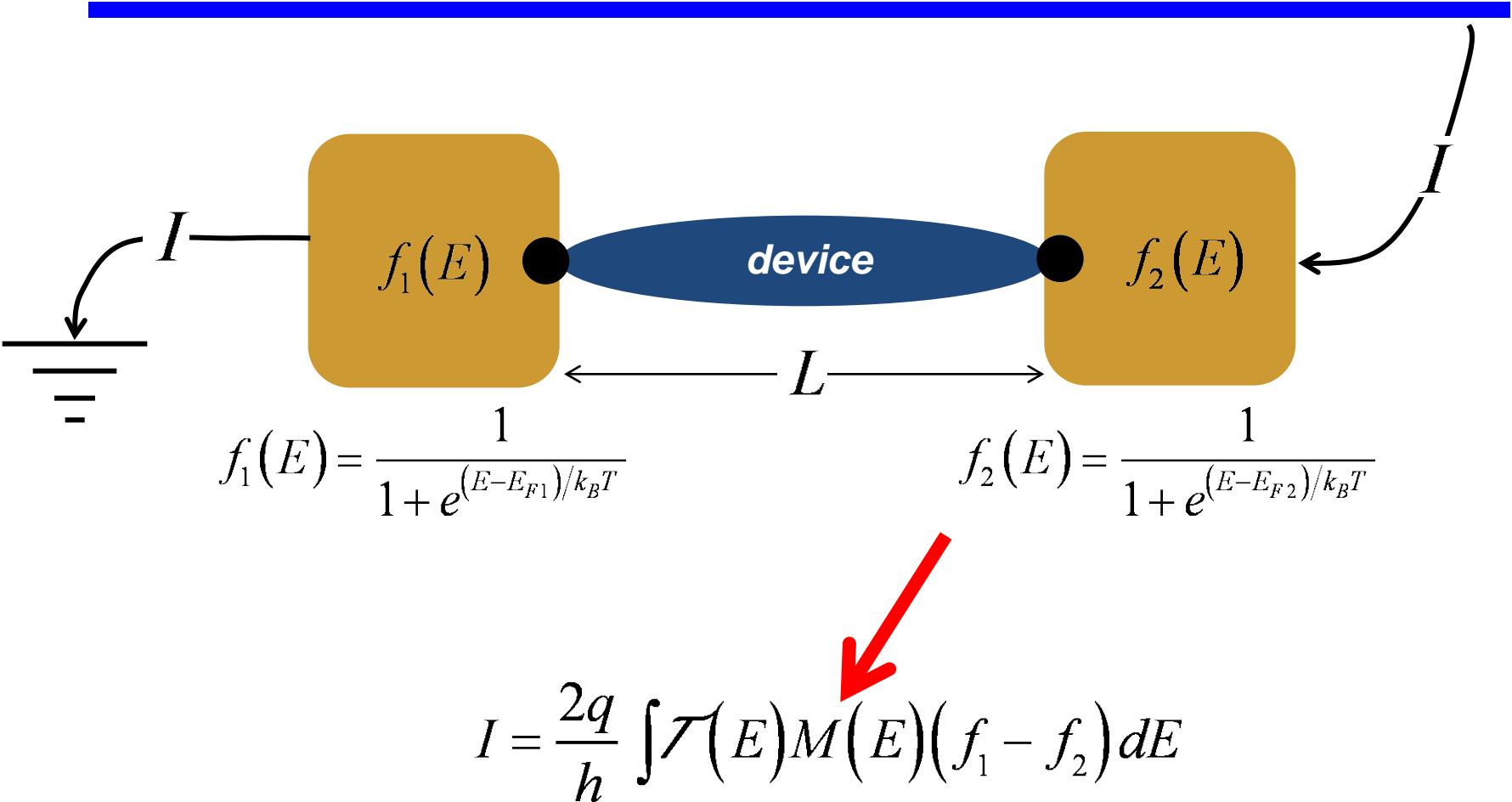
MFP for backscattering in 1D



If we assume that the scattering is ***isotropic*** (equal probability of scattering forward or back) then average time between **backscattering** events is 2τ .

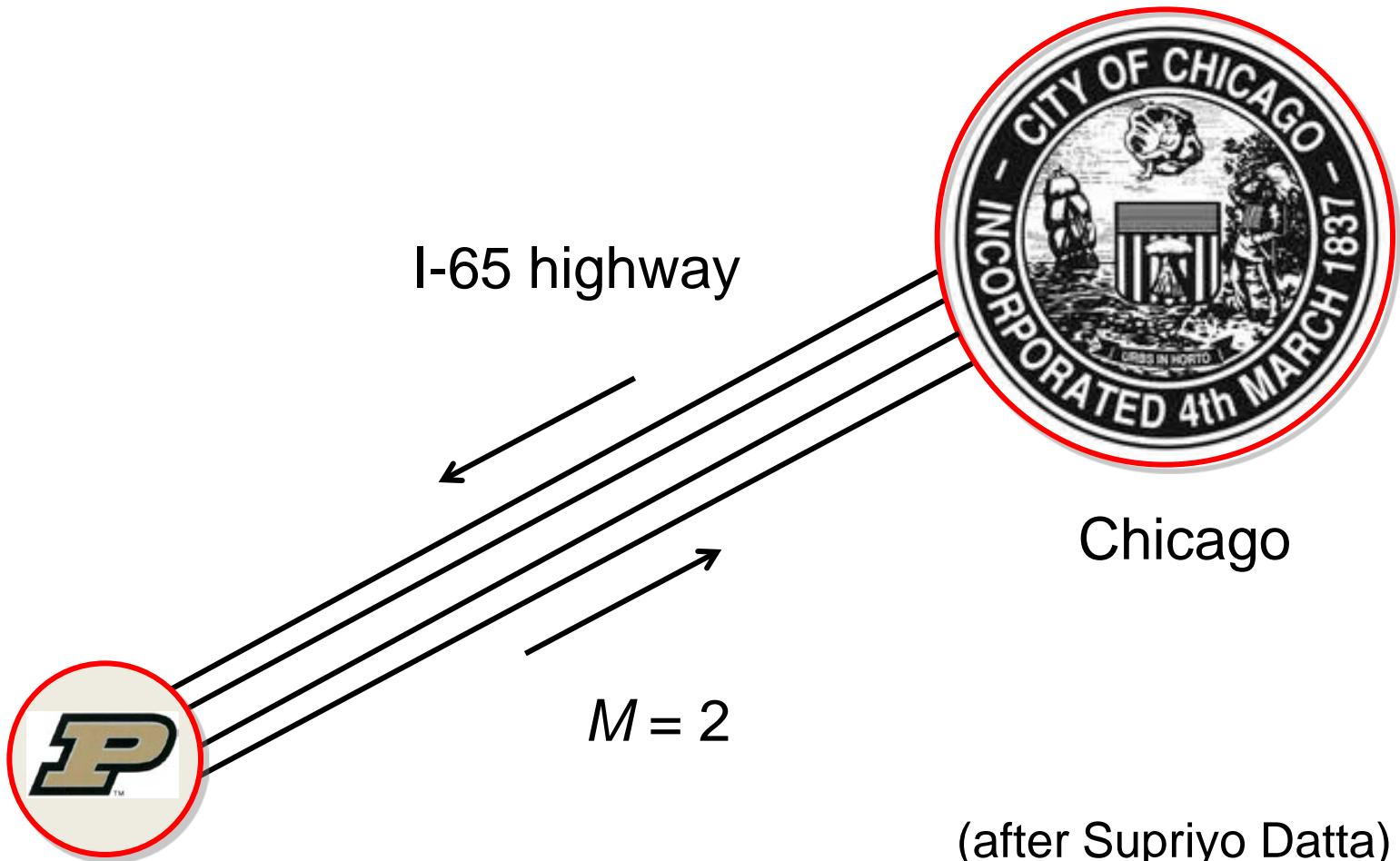
$$\lambda(E) = 2\nu(E)\tau(E) \quad \left\{ \Lambda(E) = \nu(E)\tau(E) \right\}$$

What is a channel?



(channels are also called “modes”)

Channels are like lanes on a highway



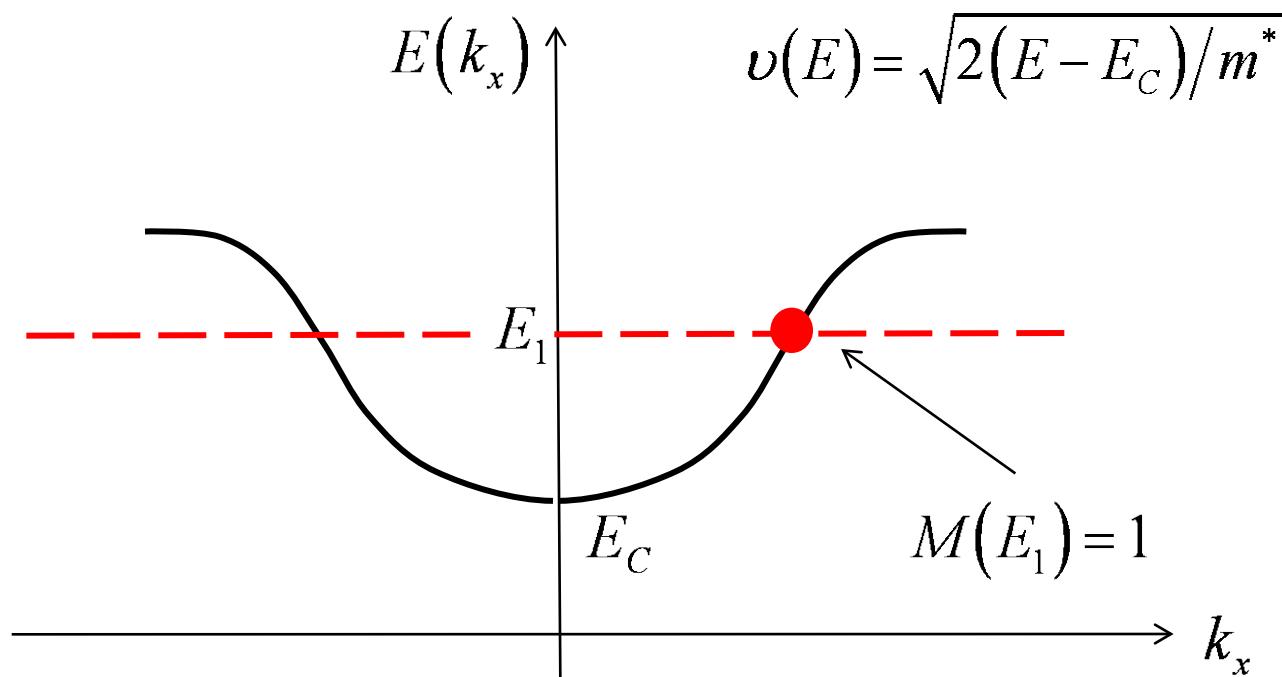
Purdue University

Lundstrom: 2018

Channels (modes) from $E(k)$

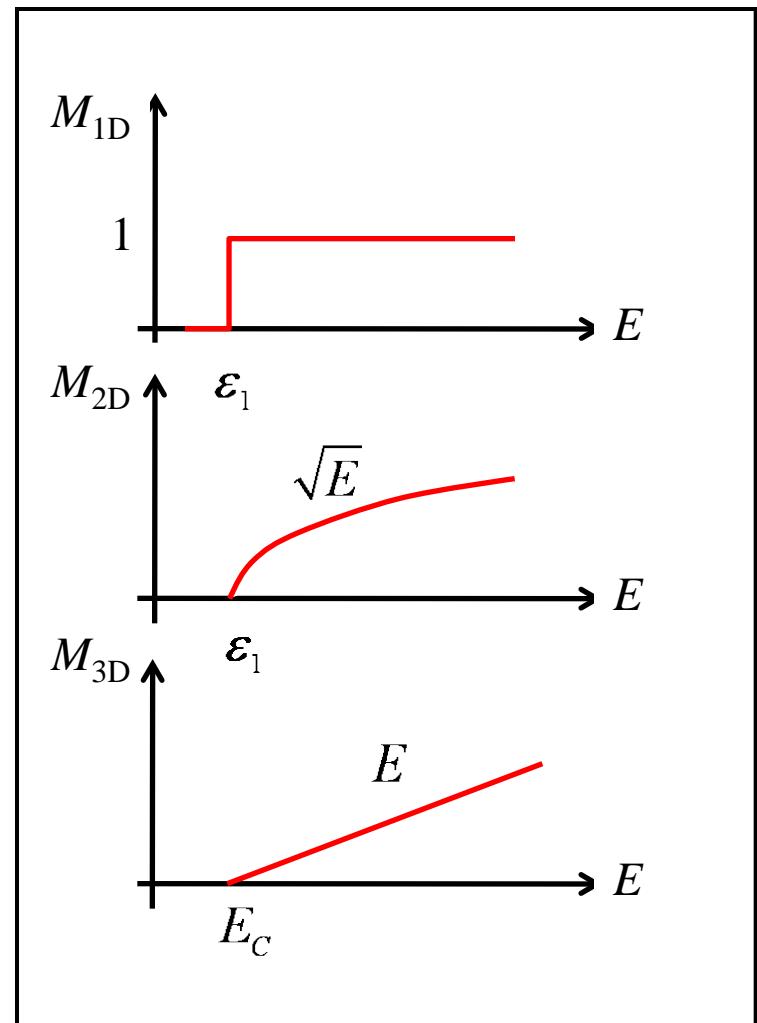
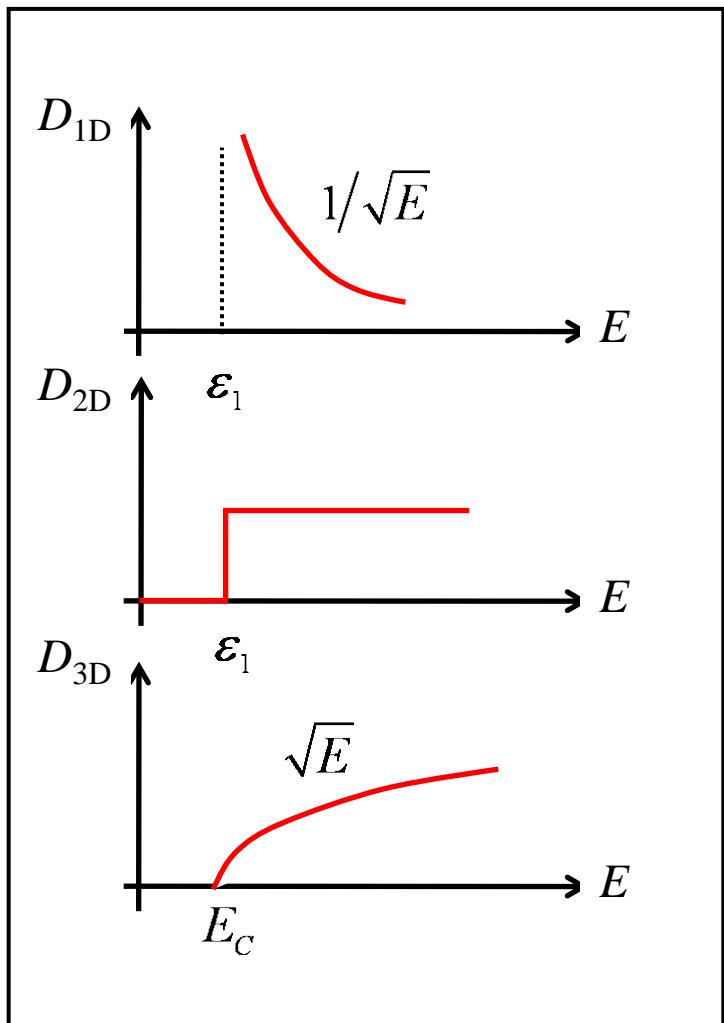
A channel is a state with a velocity.

$$M(E) = \frac{\hbar}{4} \langle v_x^+(E) \rangle D_{1D}(E)$$



(Easily generalized to arbitrary band structures in 2D and 3D.)

$DOS(E)$ vs. $M(E)$ (parabolic bands)



Transmission, channels, and MFP

$$\tau(E) = \frac{\lambda(E)}{\lambda(E) + L}$$

For two-dimensional electrons
(i.e. in the channel of a MOSFET):

$$M(E) = W \frac{\sqrt{2m^*(E - E_c)}}{\pi\hbar}$$

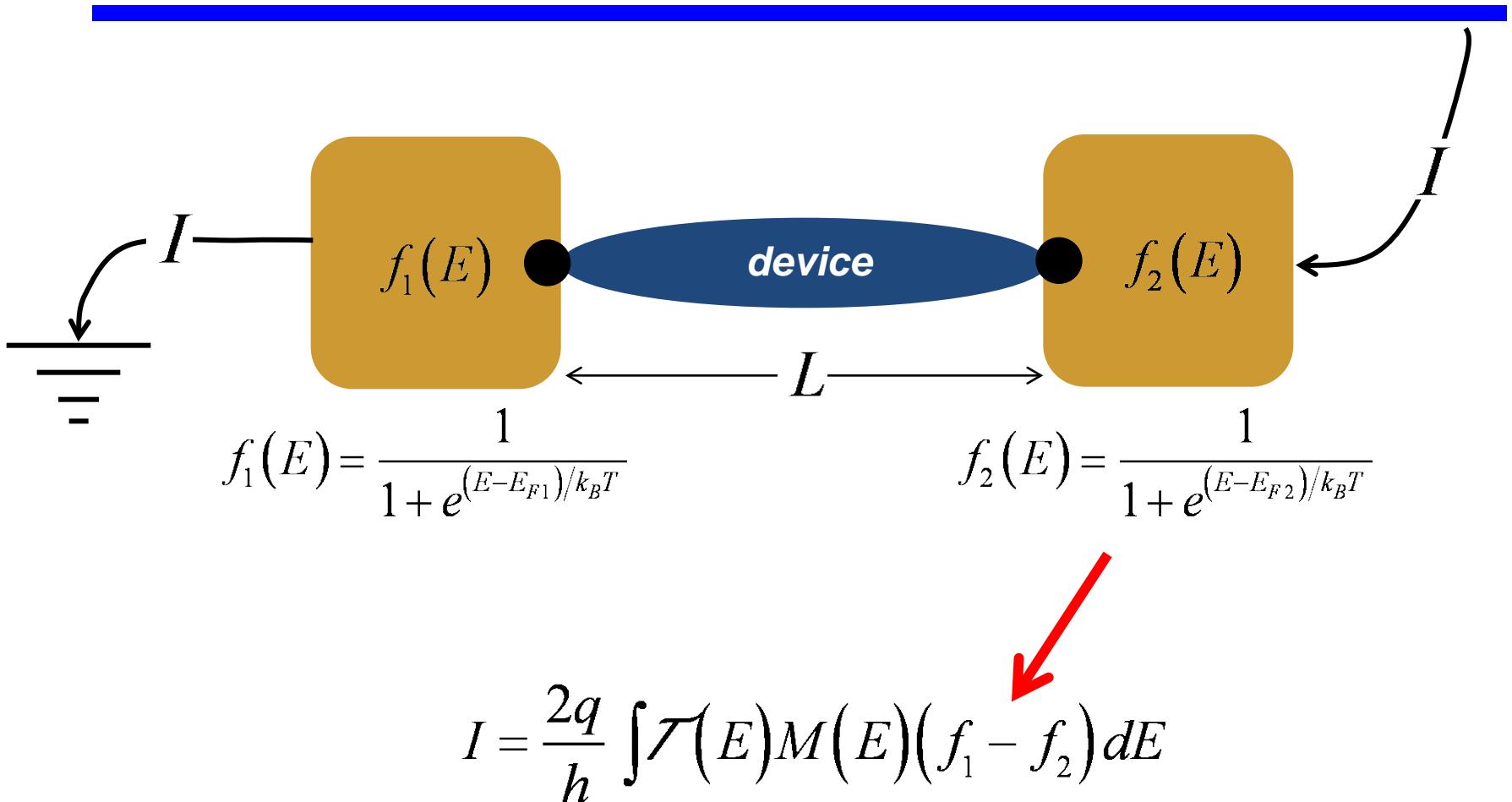
$$M_{2D}(E) = M(E)/W$$

Parabolic bands + large
structures with many channels.

$$\lambda(E) = \frac{\pi}{2} \Lambda(E)$$

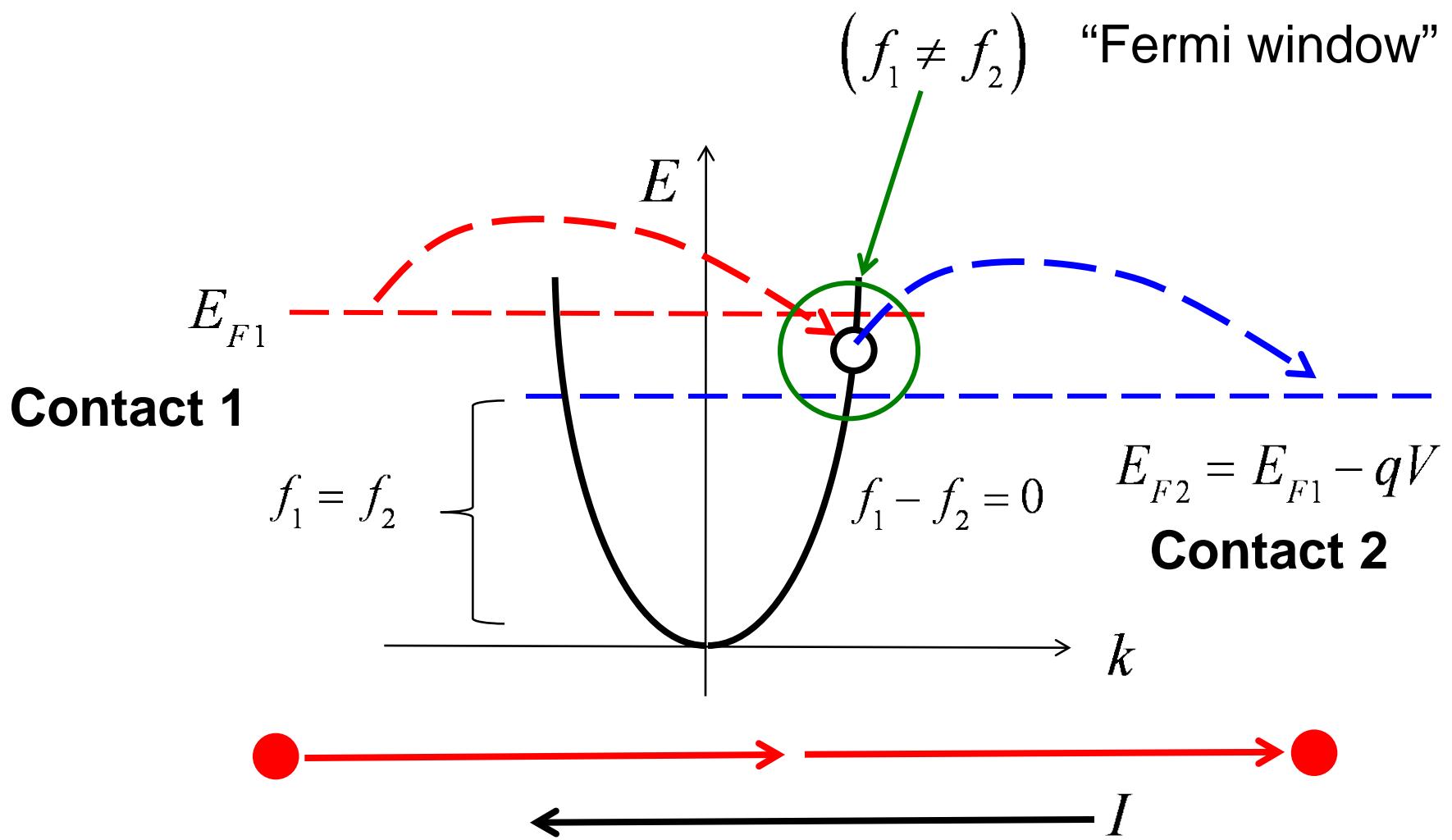
$$\Lambda(E) = v(E)\tau(E)$$

Fermi window



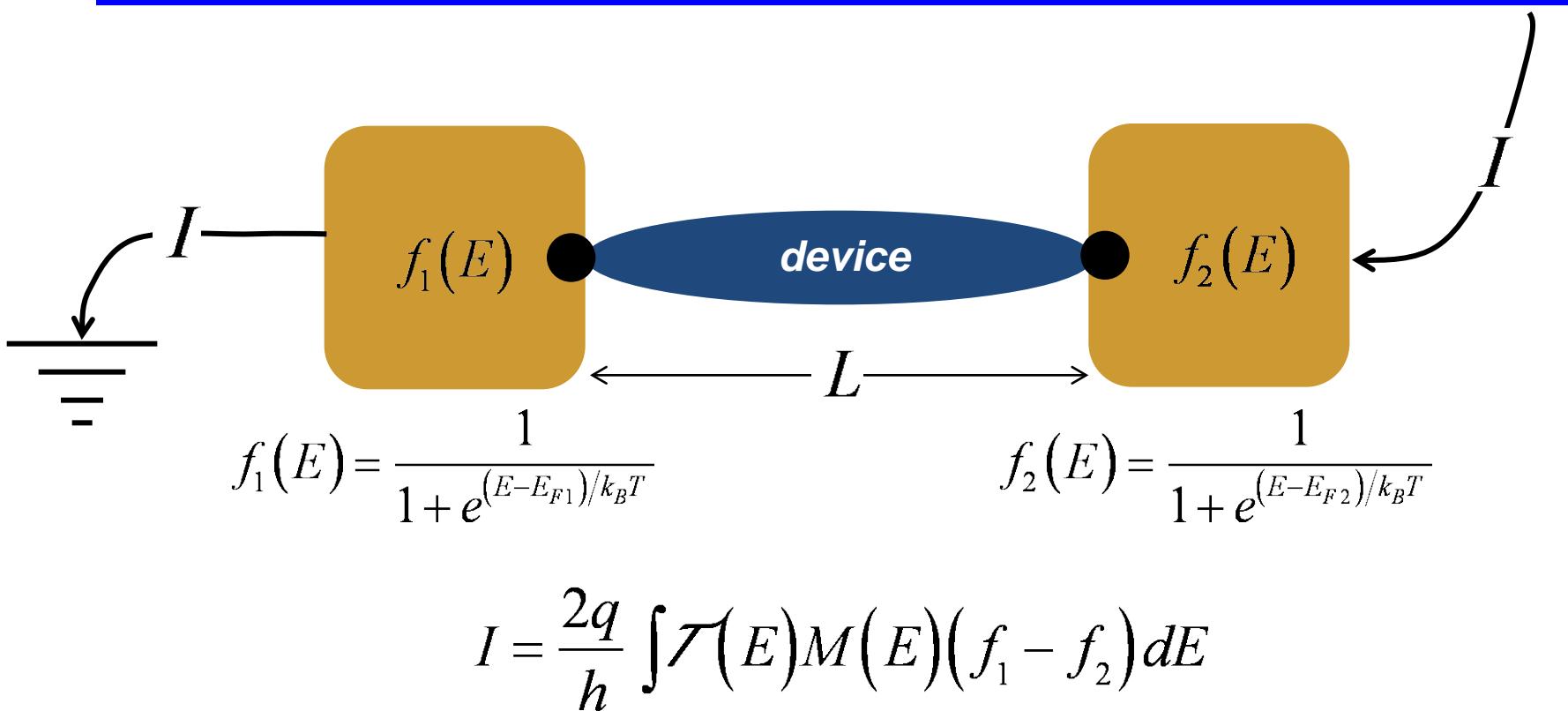
The range of energies over which $(f_1 - f_2) \neq 0$

How current flows ($T = 0$ K)



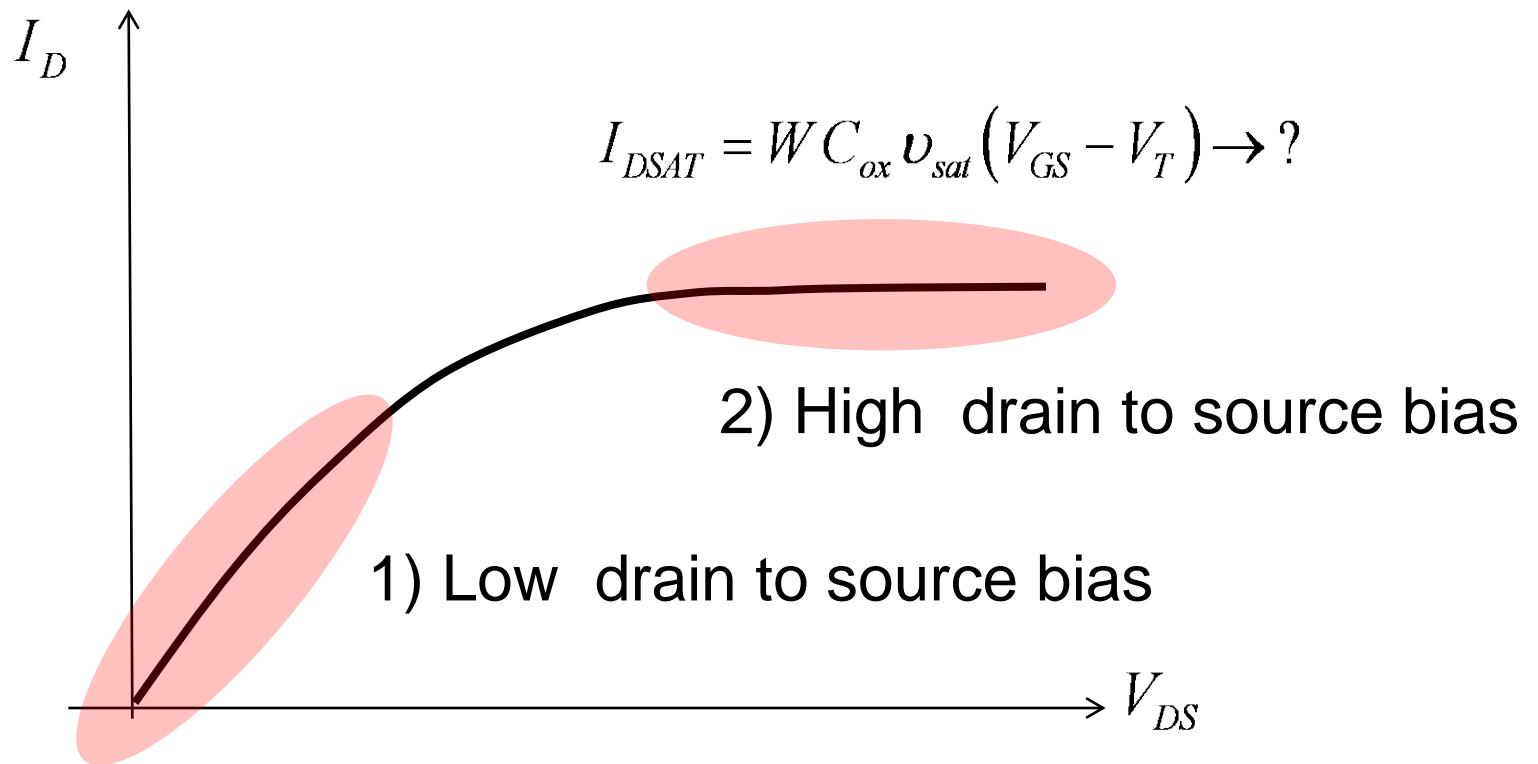
Lundstrom: 2018

Summary



Can be used to describe the current in small and large devices and in short to long devices.

Next topic



$$I_{DLN} = \frac{W}{L} \mu_n C_{ox} (V_{GS} - V_T) V_{DS} \rightarrow ?$$

Lundstrom: 2018

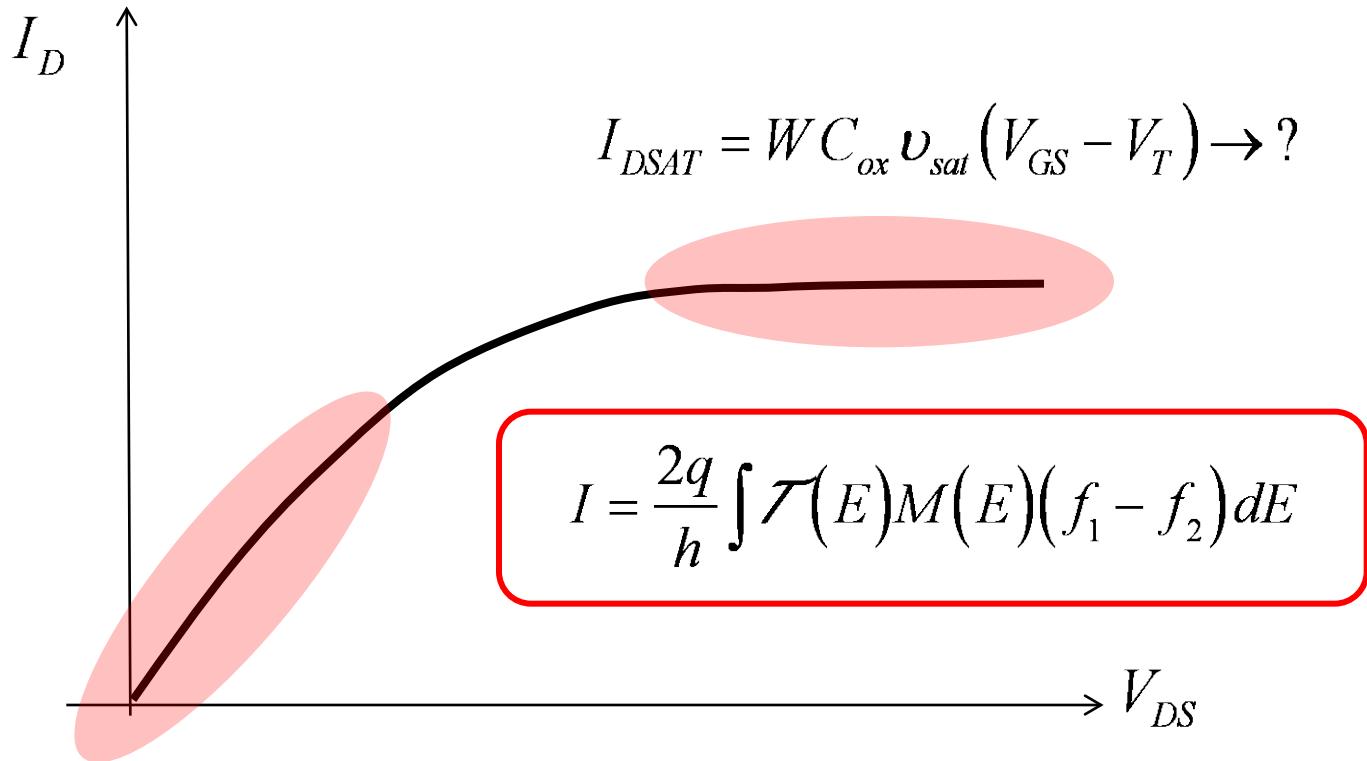
Unit 4:
Transmission Theory of the MOSFET

Lecture 4.2:
Landauer at Low and High Bias

Mark Lundstrom

lundstro@purdue.edu
Electrical and Computer Engineering
Purdue University
West Lafayette, Indiana USA

Low and high bias Landauer expressions



$$I_{DLN} = \frac{W}{L} \mu_n C_{ox} (V_{GS} - V_T) V_{DS} \rightarrow ?$$

Lundstrom: 2018

1) Low bias

$$I = \frac{2q}{h} \int \mathcal{T}(E) M(E) \left(f_1(E) - \underline{f_2(E)} \right) dE$$

Fermi window

$$f_1(E) = \frac{1}{1 + e^{(E-E_{F1})/k_B T}} = f_0(E) \quad \delta E_F = -qV$$

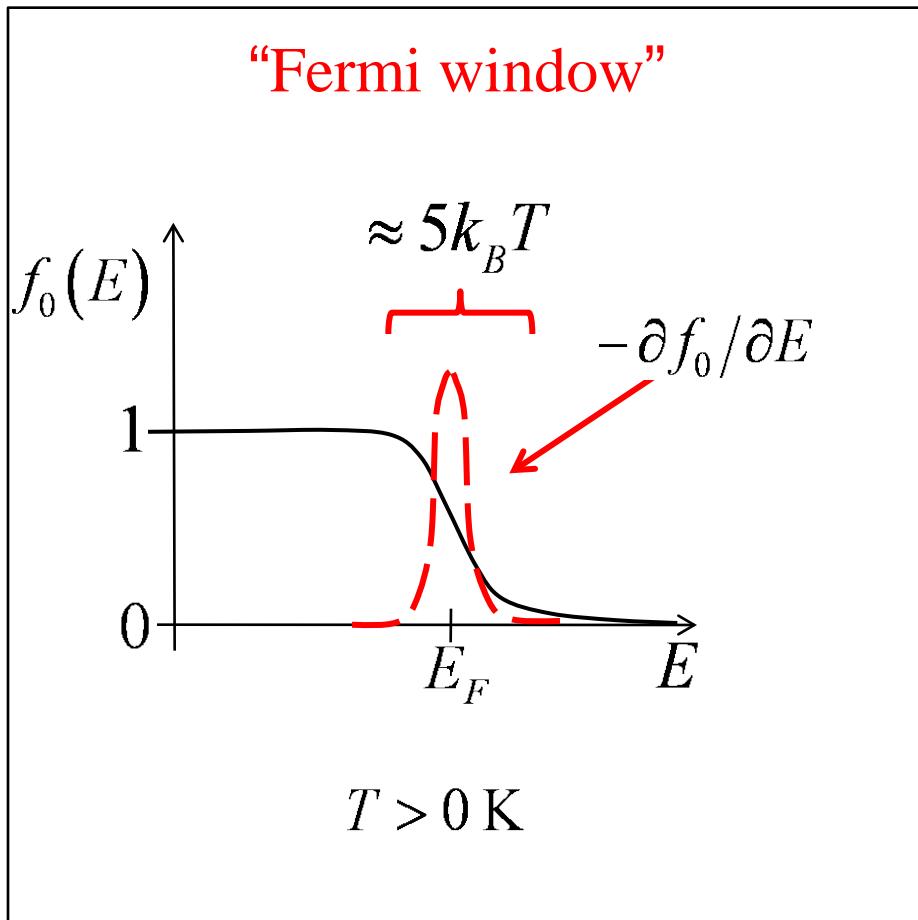
$$f_2(E) \approx f_1(E) + \frac{\partial f_1}{\partial E_F} \delta E_F$$

$$f_2(E) \approx f_1(E) + \left(-\frac{\partial f_1}{\partial E} \right) \delta E_F$$

$$f_1(E) - f_2(E) = \left(-\frac{\partial f_0}{\partial E} \right) (qV)$$

$$f_1(E) - f_2(E) \approx -\left(-\frac{\partial f_1}{\partial E} \right) \delta E_F \quad \text{Lundstrom: 2018}$$

Fermi window: Low bias

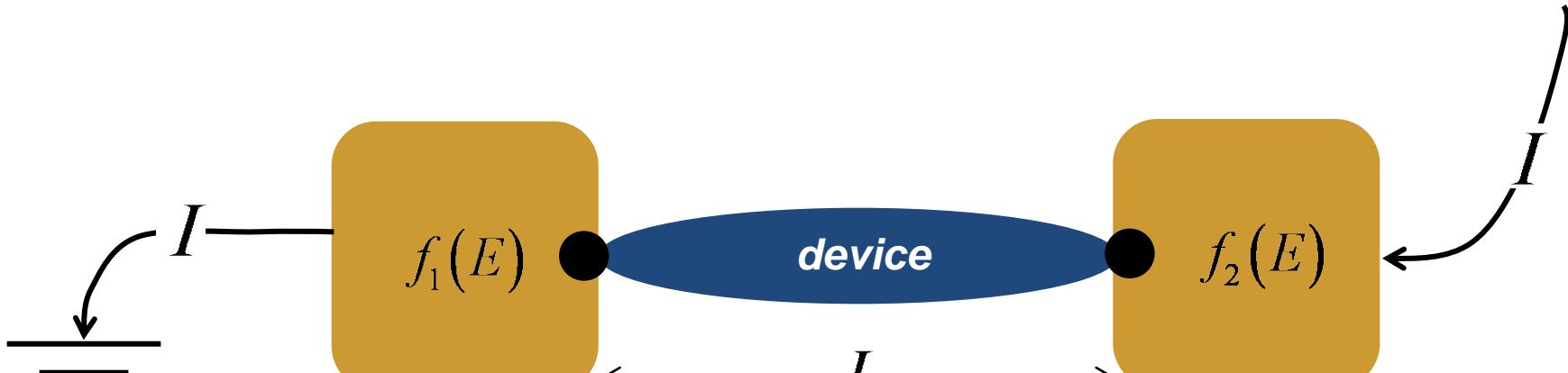


$$W_F(E) = \left(-\frac{\partial f_0}{\partial E} \right)$$

$$\int W_F(E) dE = 1$$

(window function)

Current for a small voltage difference



$$f_1(E) = \frac{1}{1 + e^{(E-E_{F1})/k_B T}}$$

$$f_2(E) = \frac{1}{1 + e^{(E-E_{F2})/k_B T}}$$

$$I = \frac{2q}{h} \int \mathcal{T}(E) M(E) (f_1 - f_2) dE$$

$$f_1(E) - f_2(E) \rightarrow \left(-\frac{\partial f_1}{\partial E} \right) (qV) \Rightarrow I = GV$$

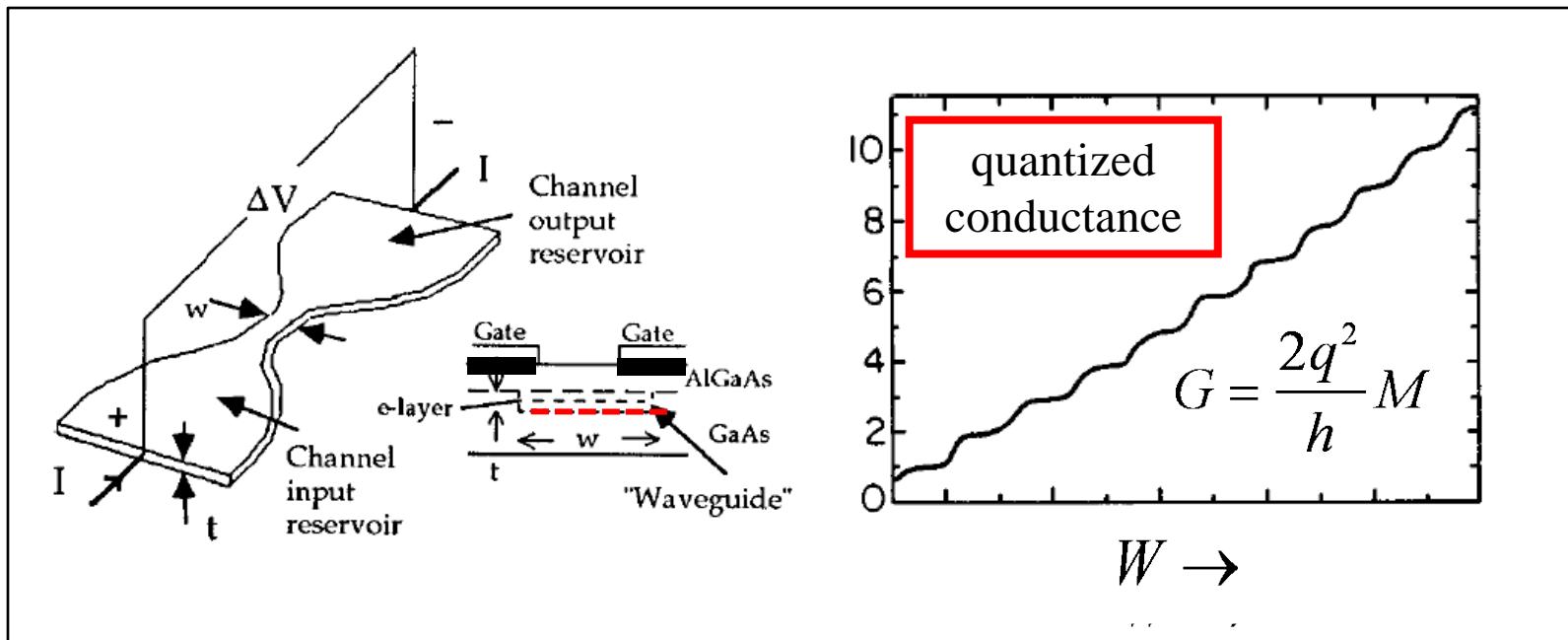
Small bias conductance

$$I = GV \quad \text{A}$$

$$G = \frac{2q^2}{\hbar} \int \mathcal{T}(E) M(E) \left(-\frac{\partial f_0}{\partial E} \right) dE \quad S$$

$$\begin{array}{ccc} \mathcal{T}(E) = 1 & \left(-\frac{\partial f_0}{\partial E} \right) = \delta(E_F) & G = \frac{2q^2}{h} M(E_F) \\ (\text{ballistic}) & (\text{T} \approx 0 \text{ K}) & \end{array}$$

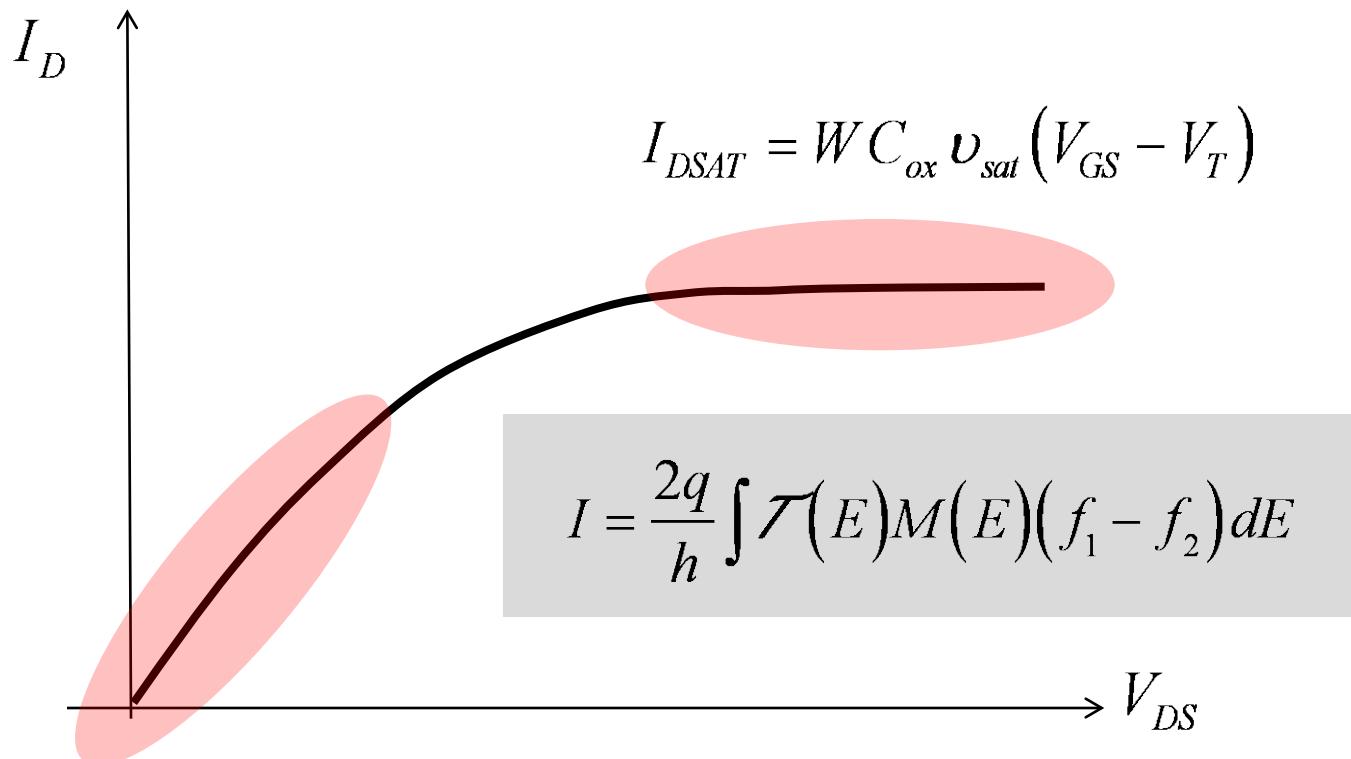
Quantized conductance



D. Holcomb, *American J. Physics*, **67**, pp. 278-297 1999.

Data from: B. J. van Wees, et al., *Phys. Rev. Lett.* **60**, 848851, 1988.

1) Linear Current in the Landauer Approach

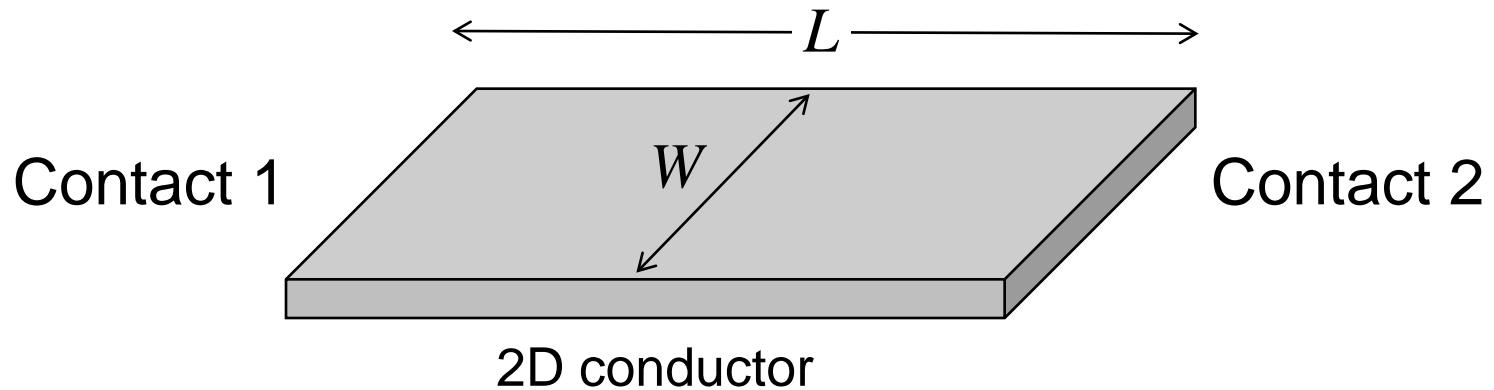


$$I_{DLIN} = \frac{W}{L} \mu_n C_{ox} (V_{GS} - V_T) V_{DS}$$

$$I_{DLIN} = \left[\frac{2q^2}{h} \int \mathcal{T}(E) M(E) \left(-\frac{\partial f_0}{\partial E} \right) dE \right] V_{DS}$$

Aside: Bulk semiconductors

Before we consider the high bias case, let's consider a bulk semiconductor (many MFP's long in both directions).



$$G = \sigma_s \frac{W}{L} \quad \sigma_s = G \frac{L}{W} \quad \Omega/\square \quad \sigma_s \equiv n_s q \mu_n$$

Conductivity (bulk)

$$G = \frac{2q^2}{h} \int \mathcal{T}(E) M(E) \left(-\frac{\partial f_0}{\partial E} \right) dE \quad \sigma_s = G \frac{L}{W}$$

$$\sigma_s = \frac{2q^2}{h} \int [\mathcal{T}(E)L] |M(E)/W| \left(-\frac{\partial f_0}{\partial E} \right) dE$$

$$\mathcal{T}(E) = \frac{\lambda(E)}{\lambda(E) + L} \rightarrow \frac{\lambda(E)}{L} \quad \text{diffusive}$$

$$M(E) = W \frac{\sqrt{2m^*(E - E_c)}}{\pi \hbar} \quad \text{2D}$$

Sheet conductivity

$$\sigma_s = \frac{2q^2}{h} \int \lambda(E) \left(M(E)/W \right) \left(-\frac{\partial f_0}{\partial E} \right) dE$$

$$\lambda(E) = \lambda_0 \quad M(E)/W = \frac{\sqrt{2m^*(E - E_C)}}{\pi\hbar} \quad f_0(E) = \frac{1}{1 + e^{(E - E_F)/k_B T}}$$

$$\sigma_s = \frac{q^2}{h} \lambda_0 \frac{\sqrt{2\pi m^* k_B T}}{\pi\hbar} e^{(E_F - E_C)/k_B T} \quad (\text{non-degenerate})$$

$$n_s = \frac{m^* k_B T}{\pi\hbar^2} e^{(E_F - E_C)/k_B T}$$

Sheet conductivity

$$\sigma_s = \frac{q^2}{h} \lambda_0 \frac{\sqrt{2\pi m^* k_B T}}{\pi \hbar} e^{(E_F - E_C)/k_B T} \equiv n_s q \mu_n \quad n_s = \frac{m^* k_B T}{\pi \hbar^2} e^{(E_F - E_C)/k_B T}$$

$$\mu_n = \frac{v_T \lambda_0}{2(k_B T / q)}$$

$$v_T = \sqrt{\frac{2k_B T}{\pi m^*}} \text{ m/s}$$

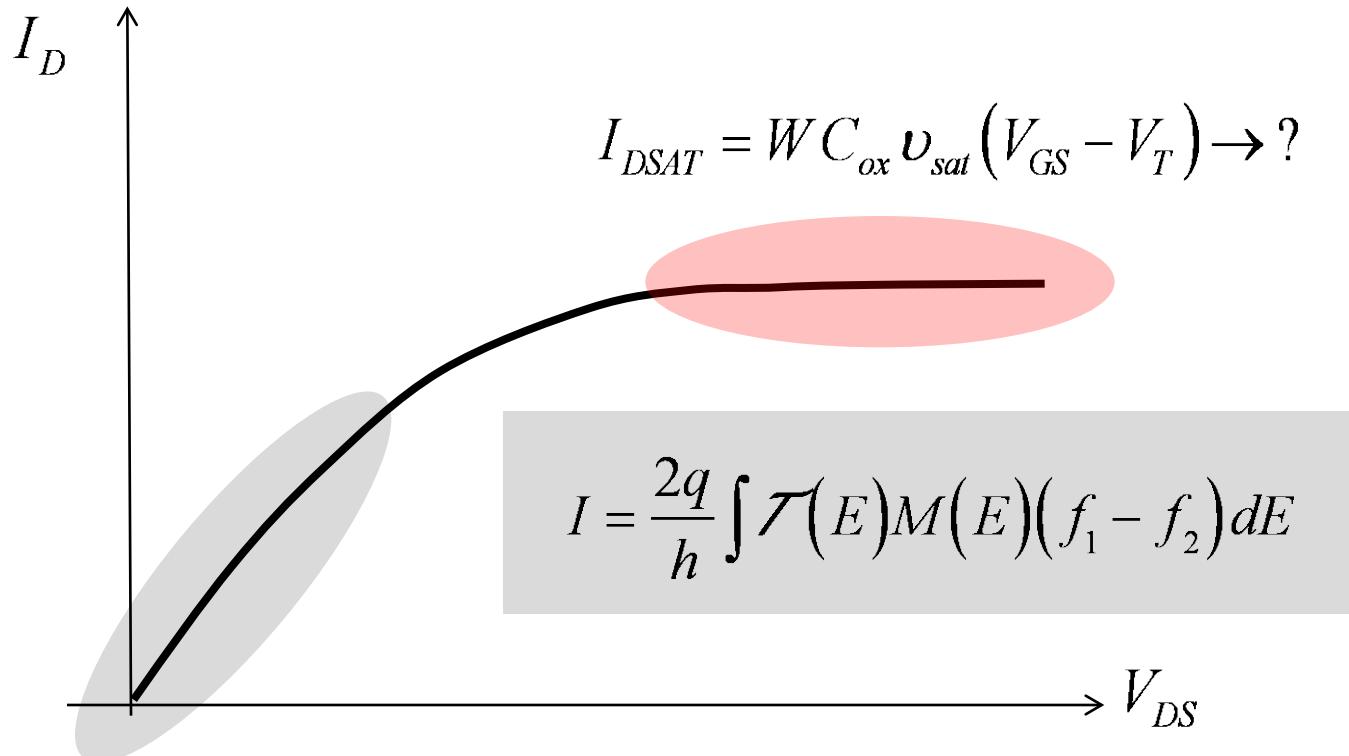
uni-directional thermal
velocity (non-degenerate)

$$\frac{D_n}{\mu_n} = \frac{k_B T}{q}$$

$$D_n = \frac{v_T \lambda_0}{2} \text{ cm}^2/\text{s}$$

(Einstein relation)

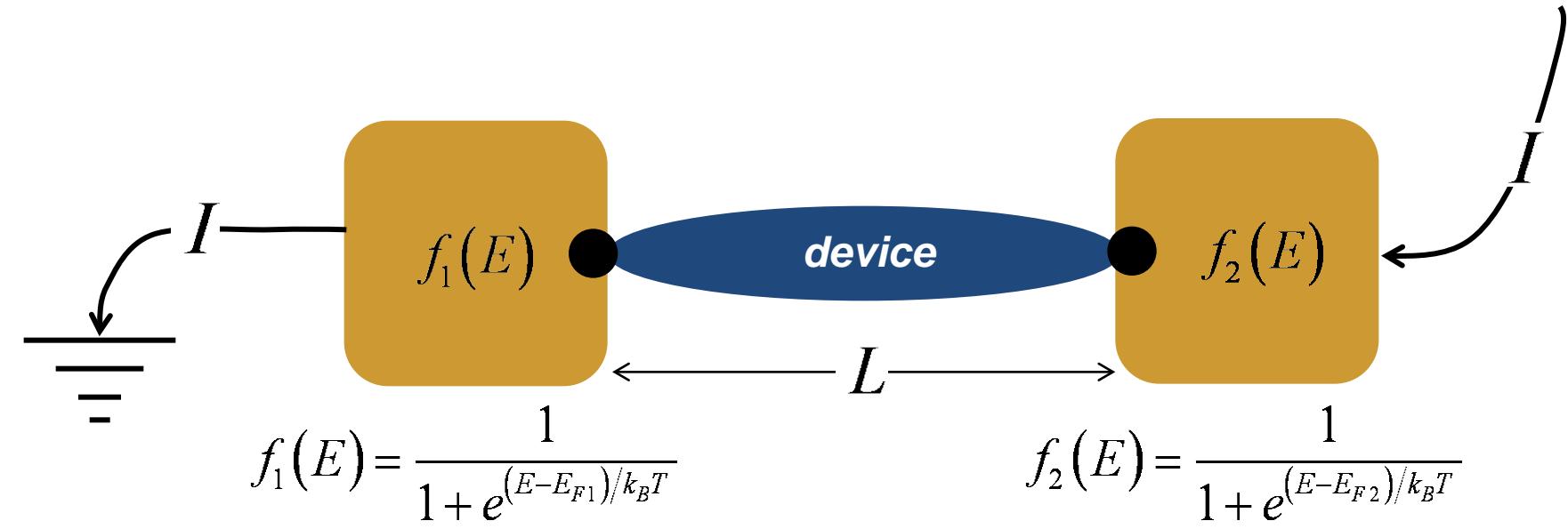
2) Saturation Current in the Landauer Approach



$$I_{DLIN} = \frac{W}{L} \mu_n C_{ox} (V_{GS} - V_T) V_{DS}$$

$$I_{DLIN} = \left[\frac{2q^2}{h} \int \mathcal{T}(E) M(E) \left(-\frac{\partial f_0}{\partial E} \right) dE \right] V_{DS}$$

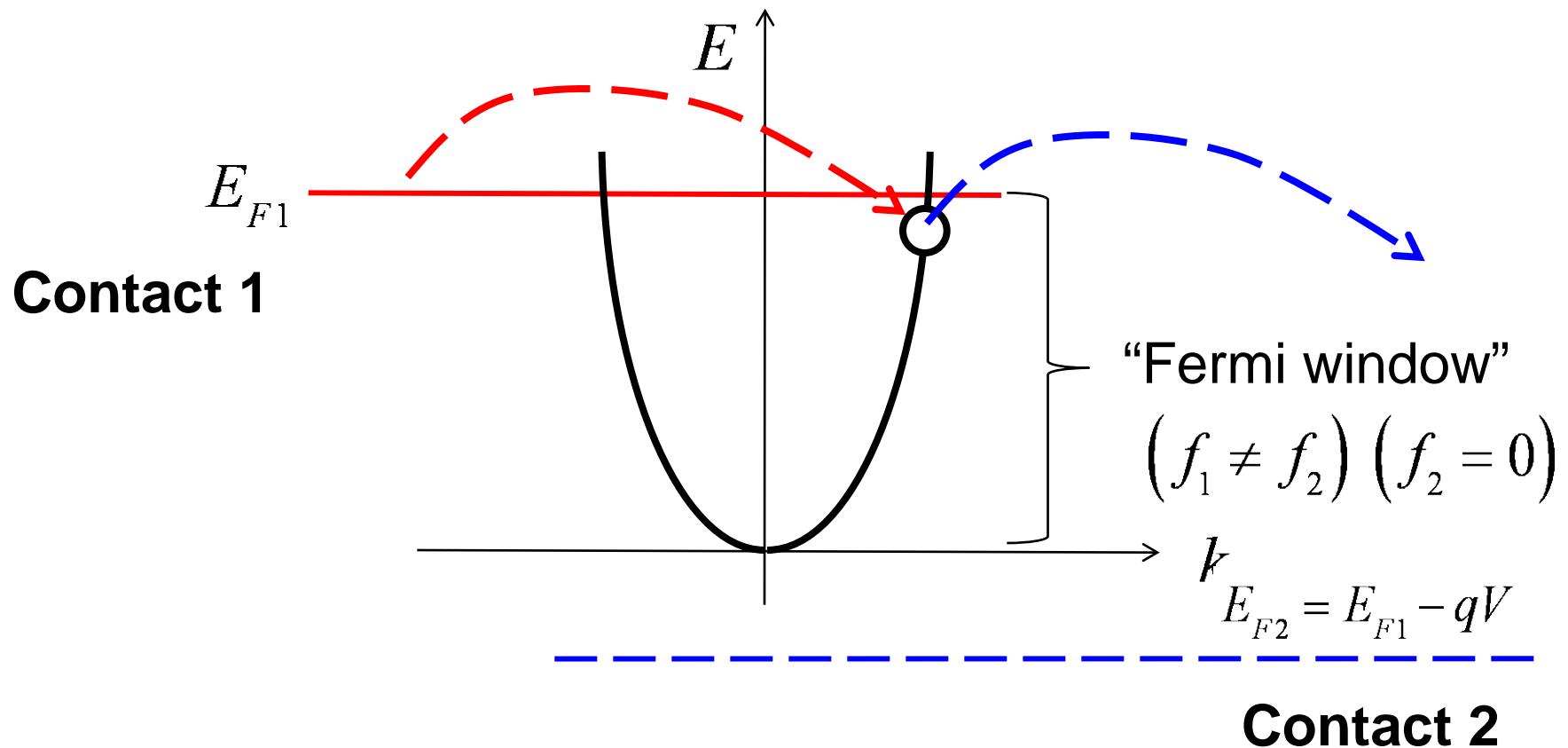
Current for a large voltage difference



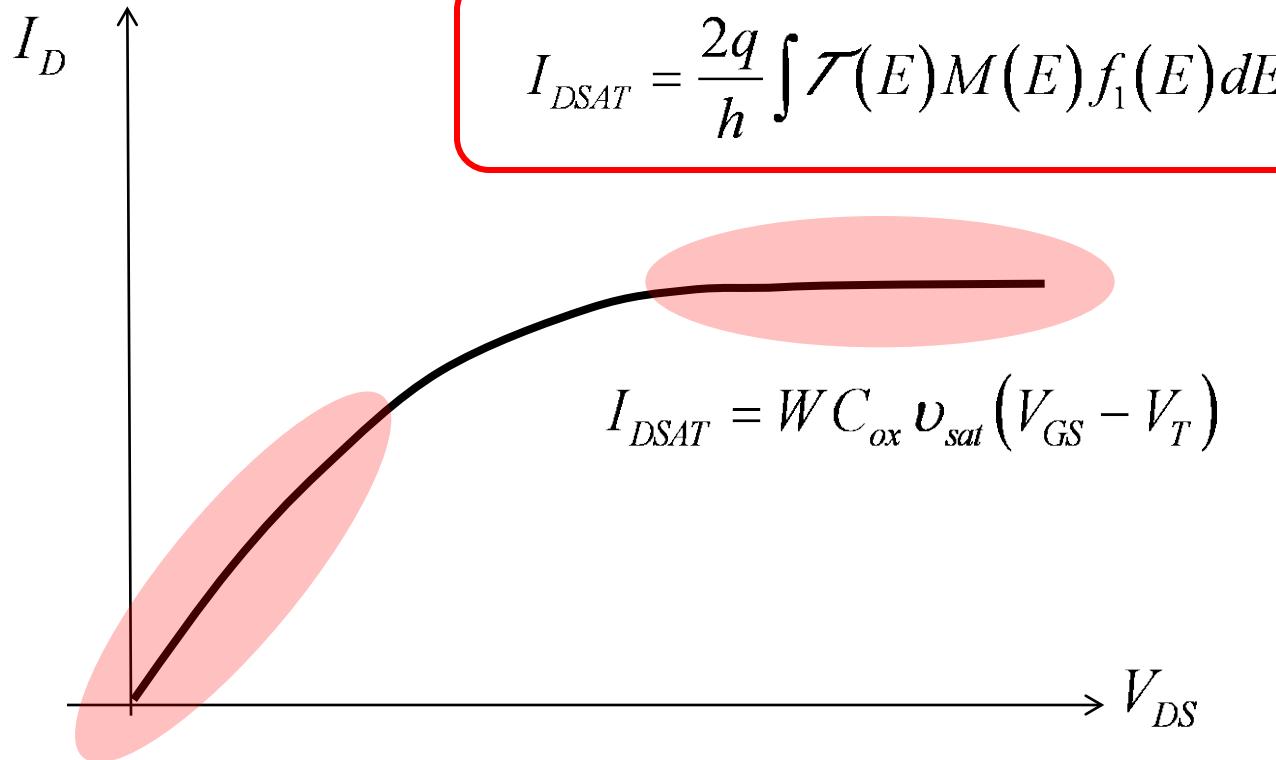
$$I = \frac{2q}{h} \int \mathcal{T}(E) M(E) (f_1 - f_2) dE$$

$$f_1 = f_0(E) = \frac{1}{1 + e^{(E - E_{F1})/k_B T}} \quad E_{F2} = E_{F1} - qV_D \quad f_2 = \frac{1}{1 + e^{(E - E_{F1} + qV_D)/k_B T}} \approx 0$$

How current flows



Current in the Landauer Approach



$$I_{DLIN} = \frac{W}{L} \mu_n C_{ox} (V_{GS} - V_T) V_{DS}$$
$$I_{DLIN} = \left[\frac{2q^2}{h} \int \mathcal{T}(E) M(E) \left(-\frac{\partial f_0}{\partial E} \right) dE \right] V_{DS}$$

Summary

$$I = \frac{2q}{h} \int \mathcal{T}(E) M(E) (f_1 - f_2) dE$$

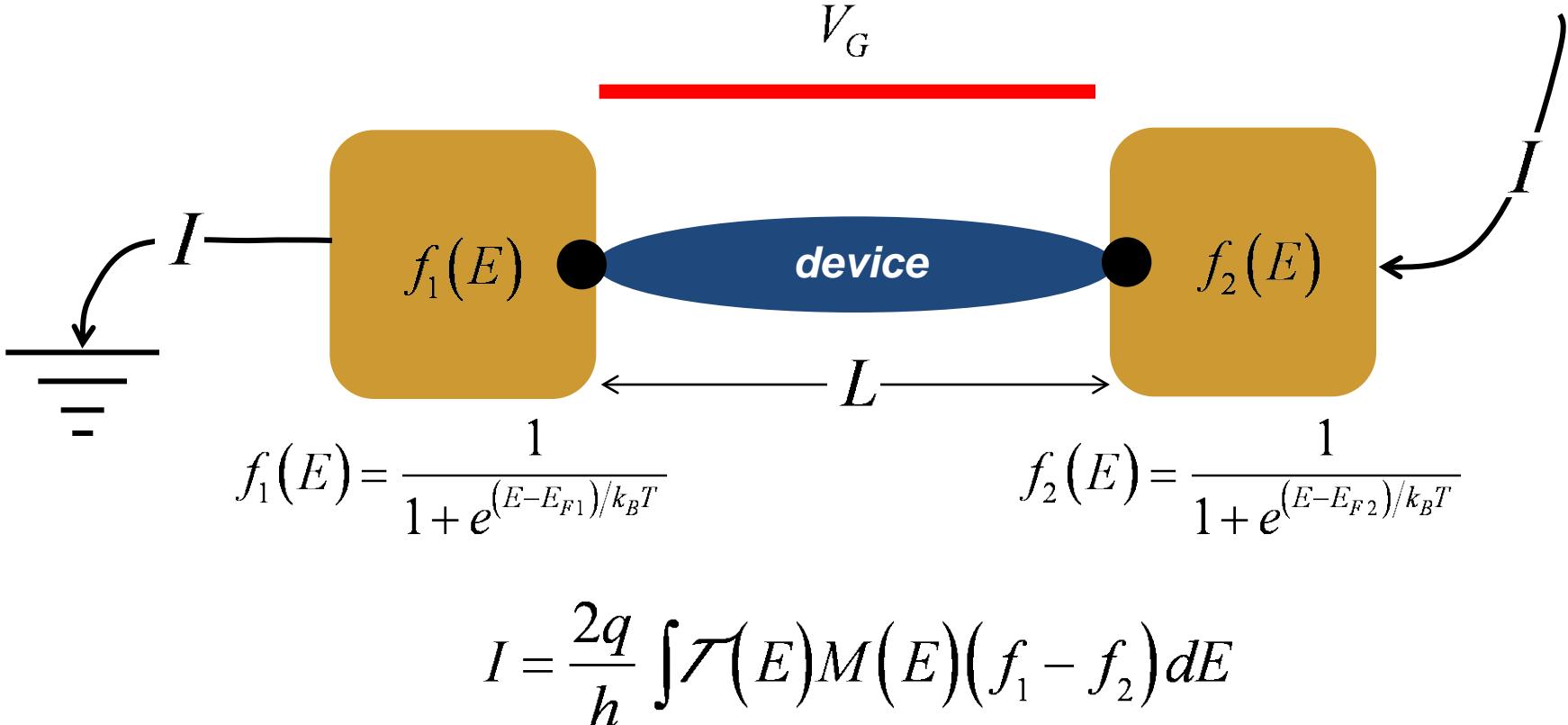
1) Linear region:

$$I_{DLIN} = \left[\frac{2q^2}{h} \int \mathcal{T}(E) M(E) \left(-\frac{\partial f_0}{\partial E} \right) dE \right] V_{DS}$$

2) Saturation region:

$$I_{DSAT} = \frac{2q}{h} \int \mathcal{T}(E) M(E) f_1(E) dE$$

Next topic



1) Ballistic MOSFET

$$\mathcal{T} = 1$$

2) MOS electrostatics

Essentials of MOSFETs

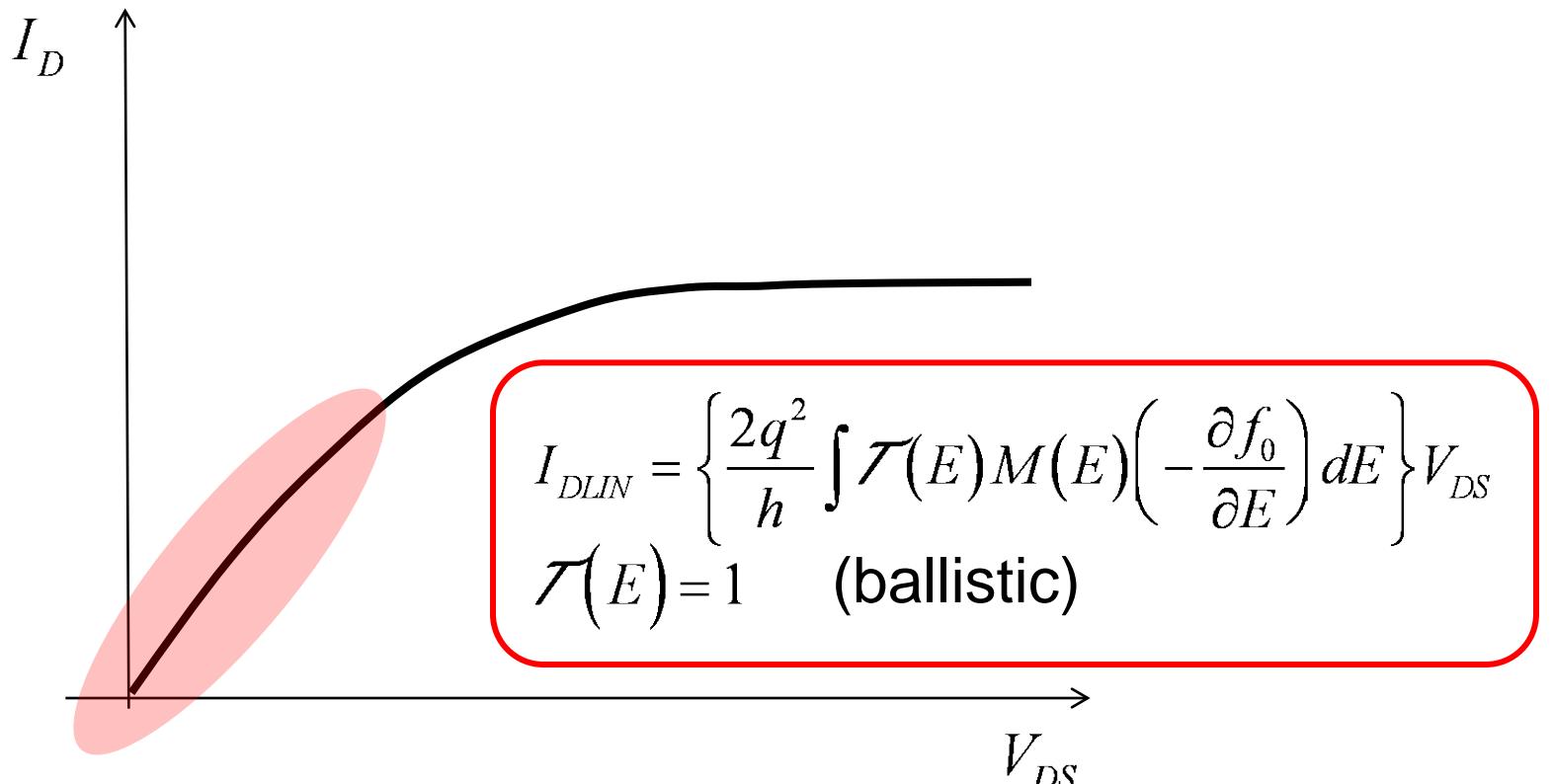
Unit 4: Transmission Theory of the MOSFET

Lecture 4.3: The Ballistic MOSFET

Mark Lundstrom

lundstro@purdue.edu
Electrical and Computer Engineering
Purdue University
West Lafayette, Indiana USA

1) Linear region



$$I_{DLIN} = \frac{W}{L} \mu_n C_{ox} (V_{GS} - V_T) V_{DS} \rightarrow ?$$

Linear region with MB statistics (i)

$$I_{DLIN} = G_{CH} V_{DS}$$

$$G_{CH} = \frac{2q^2}{h} \int_{E_C}^{\infty} \mathcal{T}(E) M(E) \left(-\frac{\partial f_0}{\partial E} \right) dE$$

(See Sections 13.3 and 15.4 of FoN lecture notes for the complete derivation.)

$$M(E) = W g_V \frac{\sqrt{2m^*(E - E_C)}}{\pi \hbar} \quad (2D)$$

$$\mathcal{T}(E) = 1$$

$$f_0(E) = \frac{1}{1 + e^{(E - E_F)/k_B T}} \approx e^{(E_F - E)/k_B T}$$

$$n_S = N_{2D} e^{(E_F - E_C)/k_B T} \quad (\text{nondegenerate})$$

$$N_{2D} = \left(g_V \frac{m^*}{\pi \hbar^2} k_B T \right)$$

Linear region with MB statistics (ii)

$$G_{CH} = \frac{2q^2}{h} \int_{E_C}^{\infty} \mathcal{P}(E) M(E) \left(-\frac{\partial f_0}{\partial E} \right) dE$$

$$G_{CH} = W(qn_s) \frac{v_T}{2(k_B T/q)} \quad v_T = \sqrt{\frac{2k_B T}{\pi m^*}}$$

$$qn_s = -Q_n = C_{inv} (V_{GS} - V_T)$$

$$G_{CH} = W C_{inv} (V_{GS} - V_T) \frac{v_T}{2(k_B T/q)}$$



$$M(E) = g_V W \frac{\sqrt{2m^*(E - E_C)}}{\pi \hbar}$$

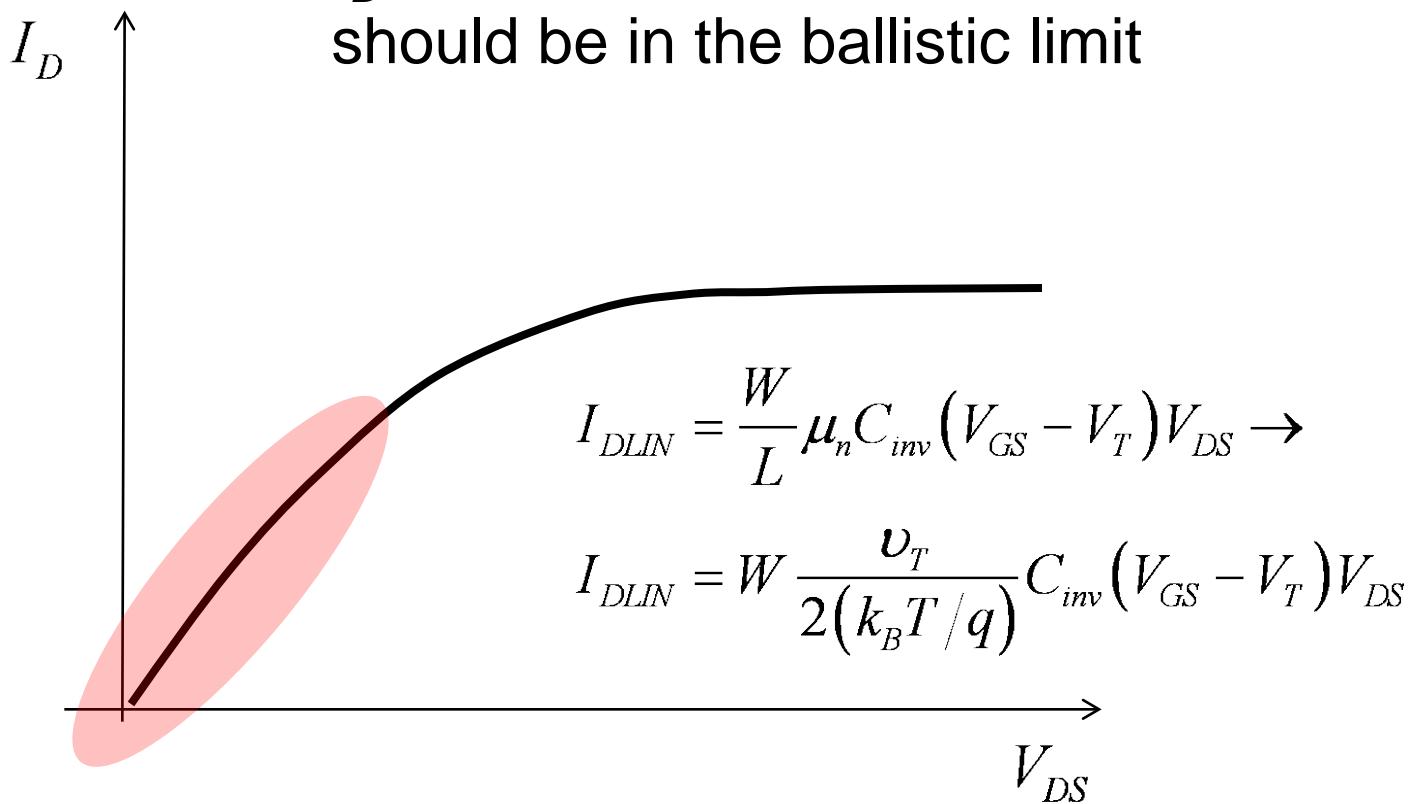
$$\mathcal{P}(E) = 1$$

$$f_0(E) = e^{(E_F - E)/k_B T}$$

$$n_s = N_{2D} e^{(E_F - E_C)/k_B T}$$

$$N_{2D} = \left(g_V \frac{m^*}{\pi \hbar^2} k_B T \right)$$

1) Linear region

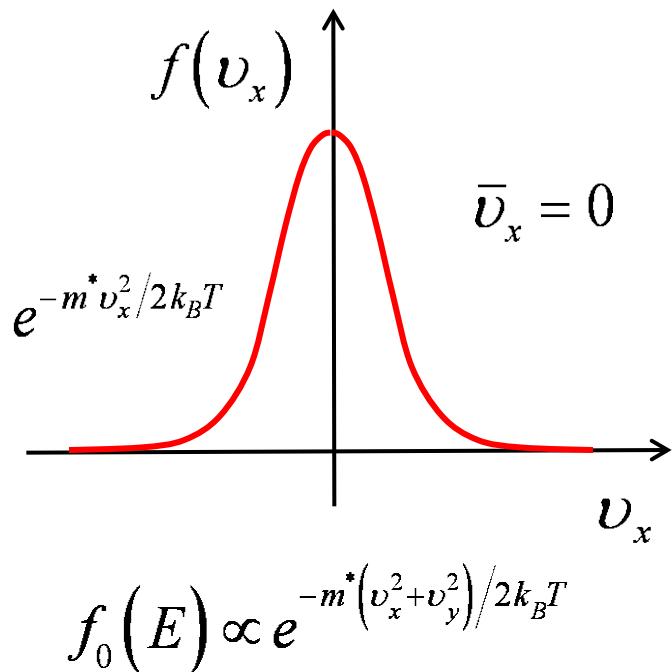


Questions

$$I_{DLIN} = W \frac{v_T}{2(k_B T / q)} C_{inv} (V_{GS} - V_T) V_{DS} \quad v_T = \sqrt{\frac{2k_B T}{\pi m^*}}$$

- 1) How do we interpret the velocity, v_T ?
- 2) Why does the traditional model, with (W/L) times mobility fit measured data for nanoscale MOSFETs so well?

Equilibrium Maxwellian velocity distribution

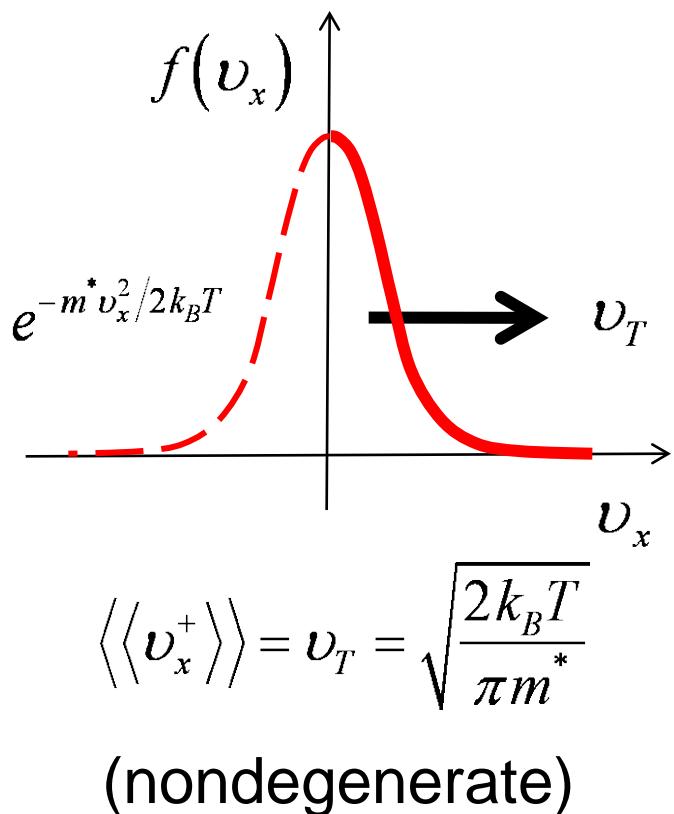


$$f_0(E) = \frac{1}{1 + e^{(E - E_F)/k_B T}} \approx e^{(E_F - E)/k_B T}$$

$$E = E_C + \frac{1}{2} m^* v^2$$

$$f_0(E) = e^{(E_F - E_C)/k_B T} \times e^{-m^* v^2/2k_B T}$$

Unidirectional thermal velocity



average over angle:

$$\langle v_x^+(E) \rangle = \frac{2}{\pi} v(E) \quad (2D)$$

average over energy:

$$\langle\langle v_x^+(E) \rangle\rangle = \frac{\int_{E_C}^{\infty} \langle v_x^+(E) \rangle D_{2D}(E) f_0(E) dE}{\int_{E_C}^{\infty} D_{2D}(E) f_0(E) dE}$$

(Exercise 12.2, p. 189, of FoN)

Ballistic mobility

$$I_{DLIN} = \frac{W}{L} \mu_n C_{inv} (V_{GS} - V_T) V_{DS}$$

traditional

$$I_{DLIN} = W \left(\frac{v_T}{2(k_B T / q)} \right) C_{inv} (V_{GS} - V_T) V_{DS}$$

Landauer

$$I_{DLIN} = \frac{W}{L} \left(\frac{v_T L}{2(k_B T / q)} \right) C_{inv} (V_{GS} - V_T) V_{DS}$$

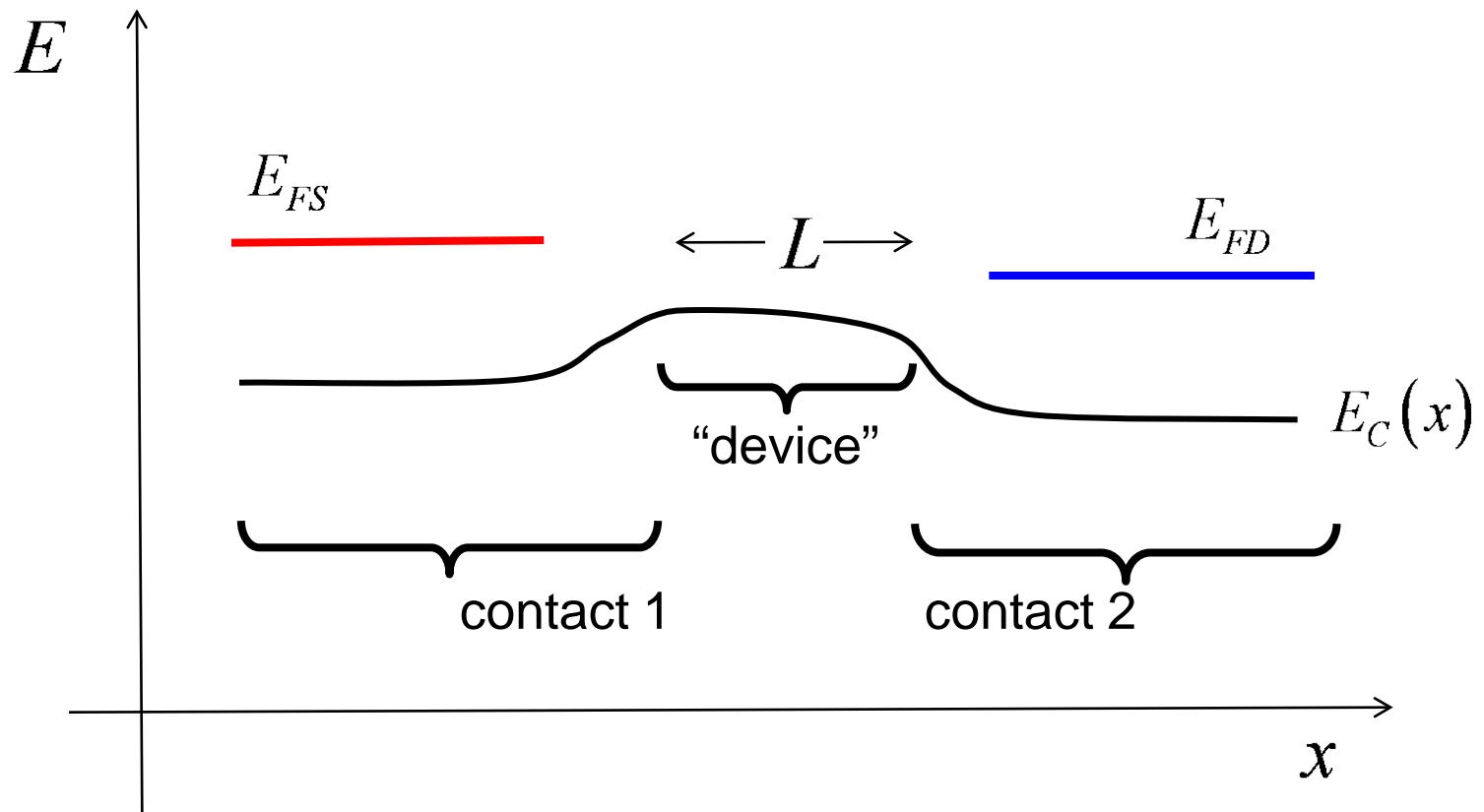
$$\mu_B \equiv \frac{v_T L}{2(k_B T / q)}$$

“ballistic mobility”

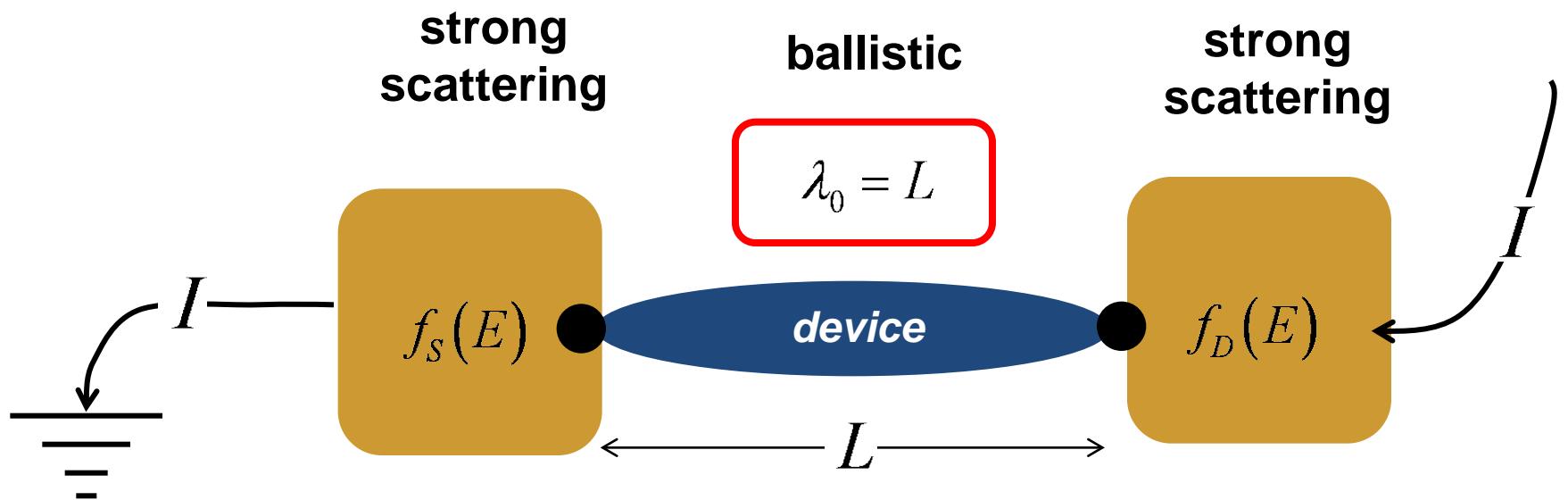
$$\mu_n = \frac{v_T \lambda_0}{2(k_B T / q)}$$

“diffusive mobility”

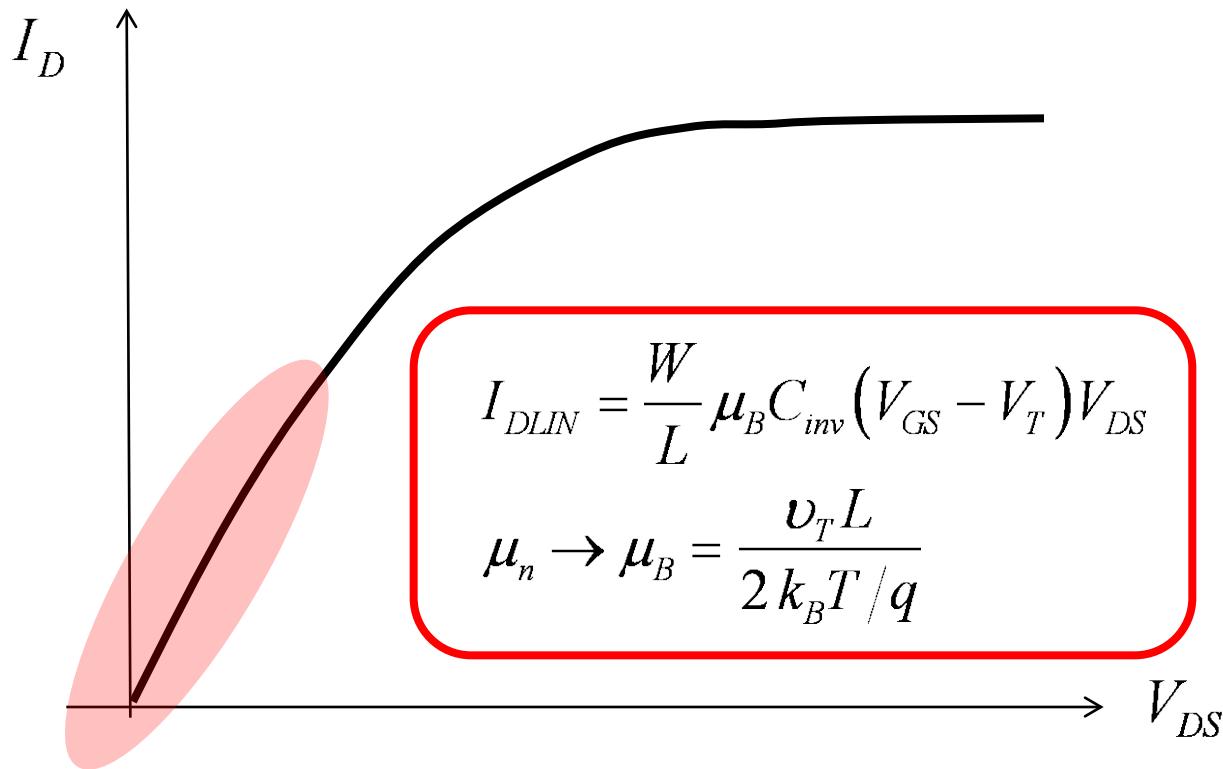
Low bias energy band diagram



MFP in a ballistic device



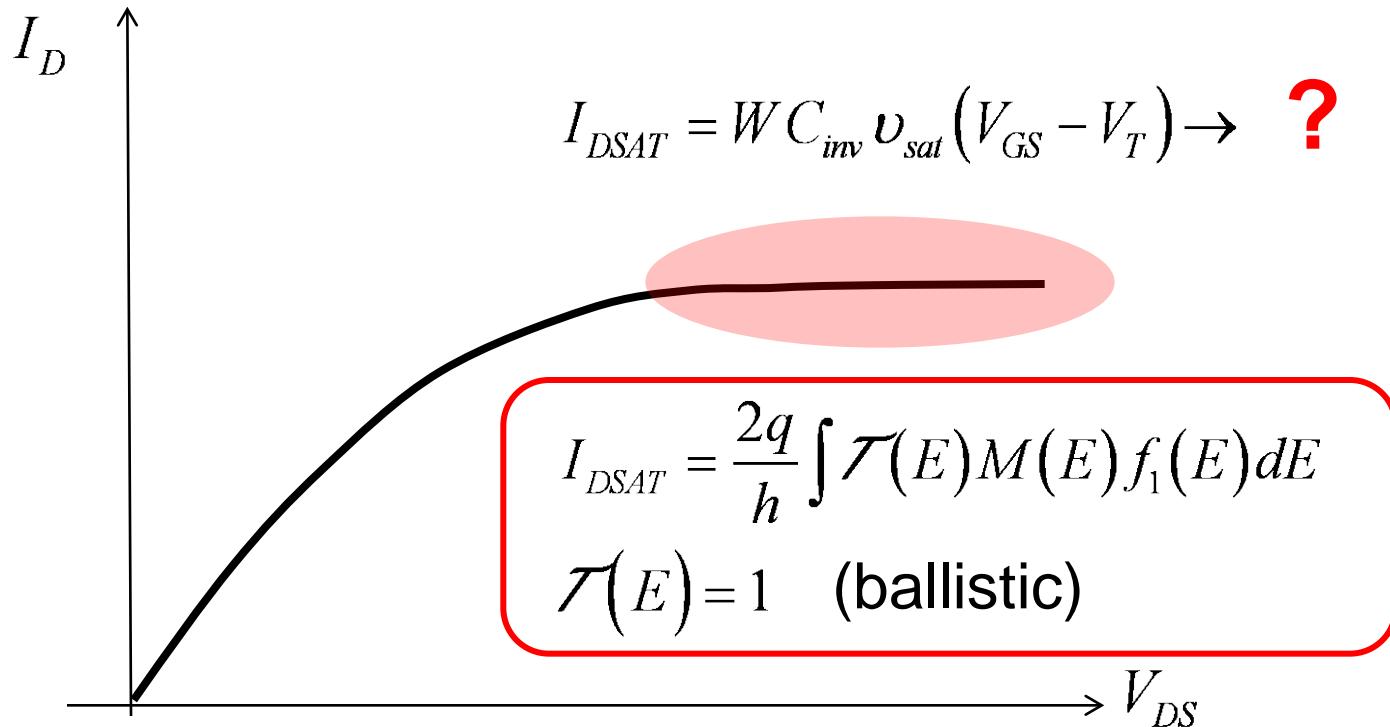
1) Linear region summary



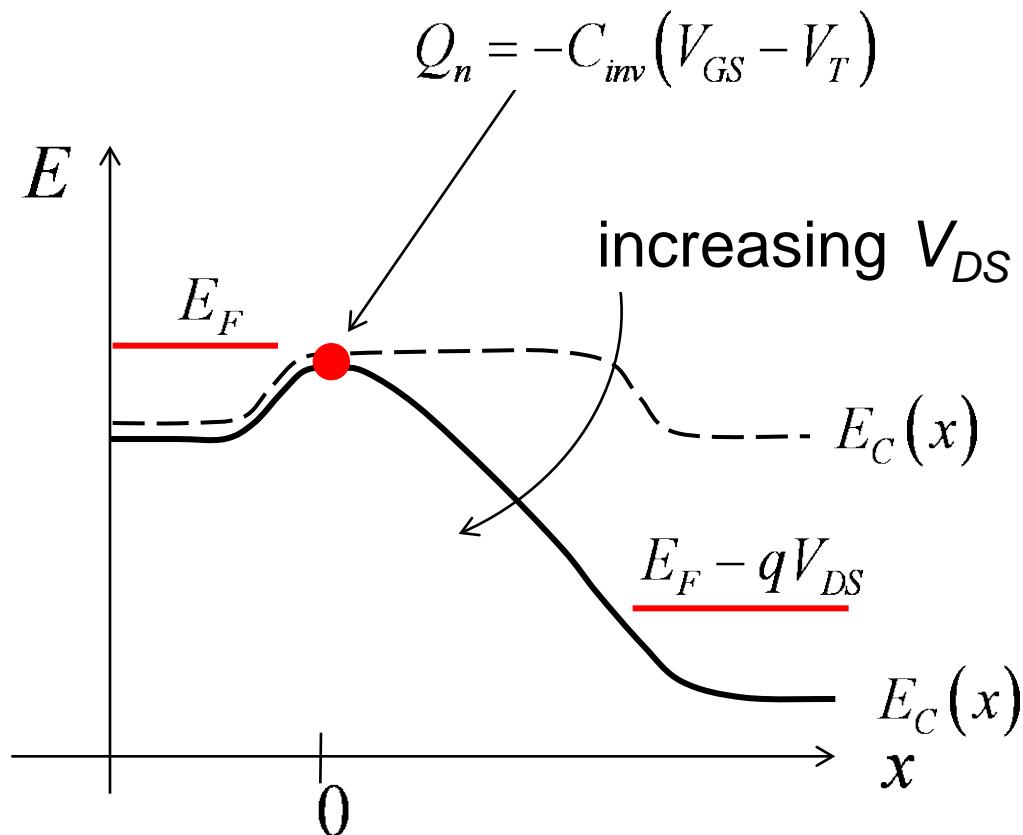
ballistic limit: μ_B

diffusive limit: μ_n

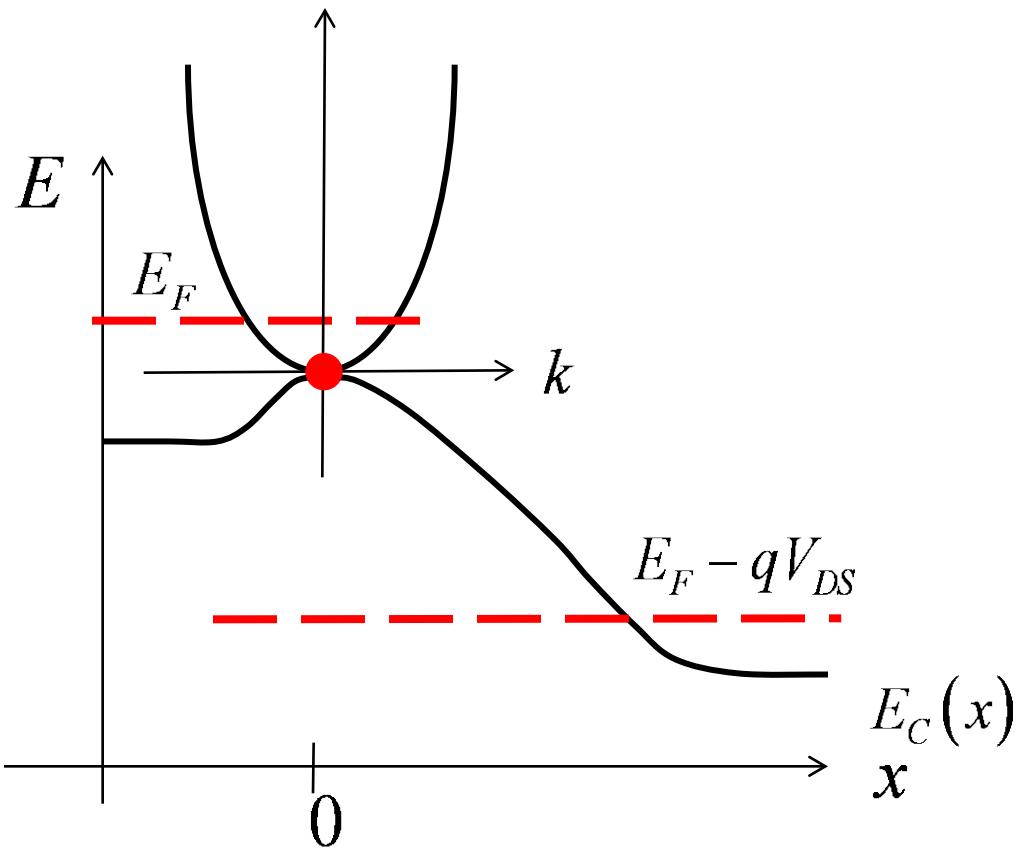
2) Saturation region



Focus on the VS (the top of the barrier)



Electron density at the VS



$$n_S^+ = \frac{N_{2D}}{2} e^{(E_F - E_C)/k_B T}$$

$$n_S^- = \frac{N_{2D}}{2} e^{(E_F - qV_{DS} - E_C)/k_B T} \approx 0$$

$$q(n_S^+ + n_S^-) \approx qn_S^+ = -Q_n$$

$$Q_n = -C_{inv}(V_{GS} - V_T)$$

$$qn_S^+ = C_{inv}(V_{GS} - V_T) = qn_S$$

Saturation region with MB statistics

$$I_{DSAT} = \frac{2q}{h} \int_{E_C}^{\infty} \mathcal{T}(E) M(E) f_1(E) dE$$

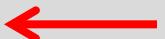
$$M(E) = g_v W \frac{\sqrt{2m^*(E - E_C)}}{\pi\hbar}$$

$$\mathcal{T}(E) = 1$$

$$f_1(E) = f_0(E) = e^{(E_F - E)/k_B T}$$

$$v_T = \sqrt{2k_B T / \pi m^*}$$

$$n_s = \frac{N_{2D}}{2} e^{(E_F - E_C)/k_B T}$$



(See Sections 13.4 and 15.4 of FoN lecture notes for the complete derivation.)

Saturation region with MB statistics

$$I_{DSAT} = \frac{2q}{h} \int_{E_C}^{\infty} \mathcal{T}(E) M(E) f_1(E) dE$$

$$I_{DSAT} = W(qn_S)v_T$$

$$qn_S = -Q_n = C_{inv}(V_{GS} - V_T)$$

$$I_{DSAT} = WC_{inv}(V_{GS} - V_T)v_T$$



$$M(E) = g_V W \frac{\sqrt{2m^*(E - E_C)}}{\pi\hbar}$$

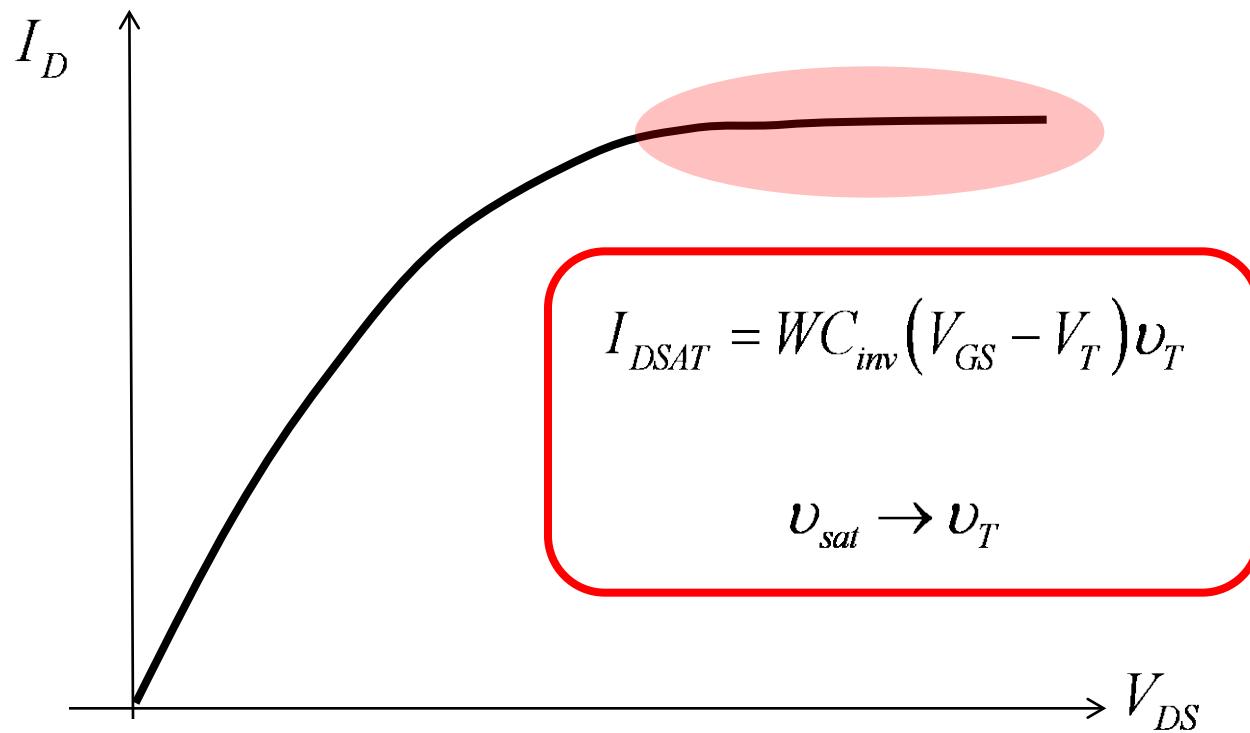
$$\mathcal{T}(E) = 1$$

$$f_1(E) = f_0(E) = e^{(E_F - E)/k_B T}$$

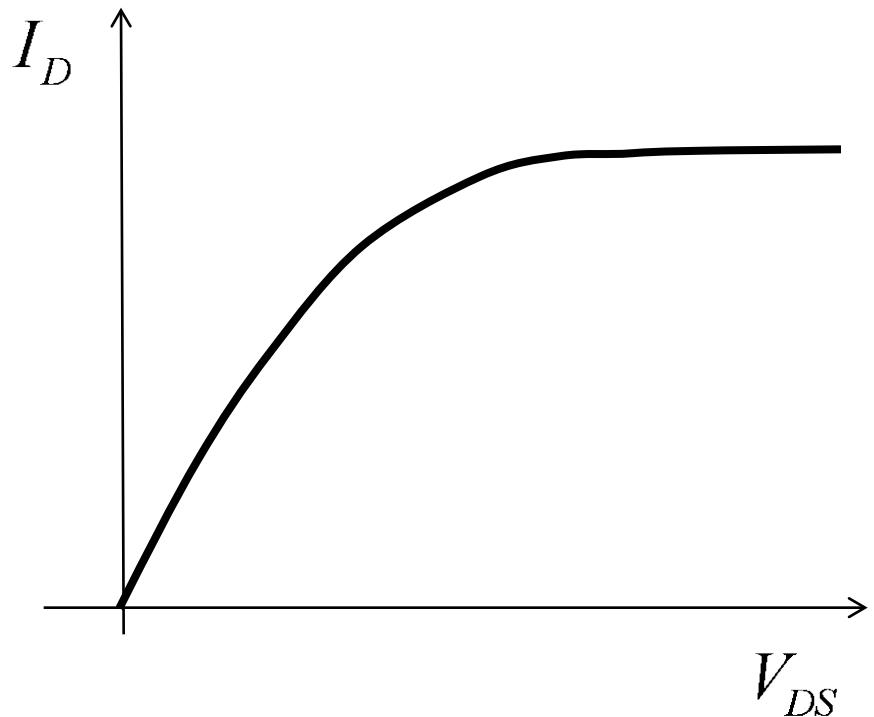
$$v_T = \sqrt{2k_B T / \pi m^*}$$

$$n_S = \frac{N_{2D}}{2} e^{(E_F - E_C)/k_B T}$$

2) Saturation region summary



3) Ballistic MOSFET: full V_{DS} range



$$I_D = \frac{2q}{h} \int \mathcal{T}(E) M(E) (f_1 - f_2) dE$$

$$I_D = I_{S \rightarrow D} - I_{D \rightarrow S}$$

$$I_{S \rightarrow D} = -W Q_n^+ v_T$$

$$I_{D \rightarrow S} = -W Q_n^- v_T$$

$$I_D = -W v_T (Q_n^+ - Q_n^-)$$

$$= -W v_T Q_n^+ \left(1 - Q_n^- / Q_n^+\right)$$

Full V_{DS} expression

$$I_D = -W v_T Q_n^+ \left(1 - Q_n^- / Q_n^+ \right)$$

$$Q_n = Q_n^+ + Q_n^- = Q_n^+ \left(1 + Q_n^- / Q_n^+ \right)$$

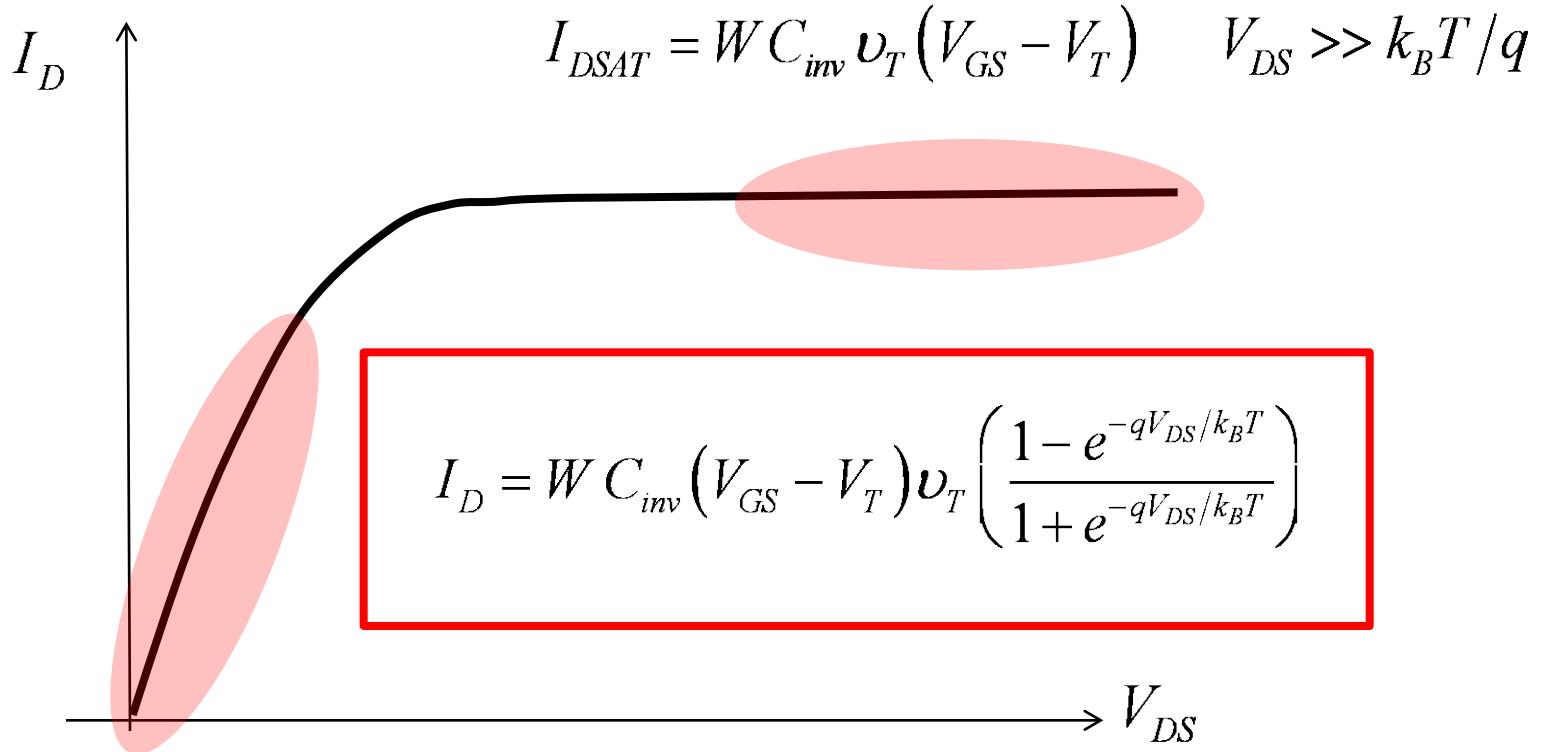
$$Q_n^+ = \frac{Q_n}{\left(1 + Q_n^- / Q_n^+ \right)}$$

$$I_D = -W v_T Q_n \frac{\left(1 - Q_n^- / Q_n^+ \right)}{\left(1 + Q_n^- / Q_n^+ \right)}$$

$$\frac{Q_n^-}{Q_n^+} = e^{-qV_{DS}/k_B T}$$

$$I_D = W C_{inv} (V_{GS} - V_T) v_T \left(\frac{1 - e^{-qV_{DS}/k_B T}}{1 + e^{-qV_{DS}/k_B T}} \right)$$

Full range ballistic model (nondegenerate)



$$I_{DLIN} = \frac{W}{L} \mu_B C_{inv} (V_{GS} - V_T) V_{DS} \quad \mu_B = \frac{v_T L}{2 k_B T / q} \quad V_{DS} \ll k_B T / q$$

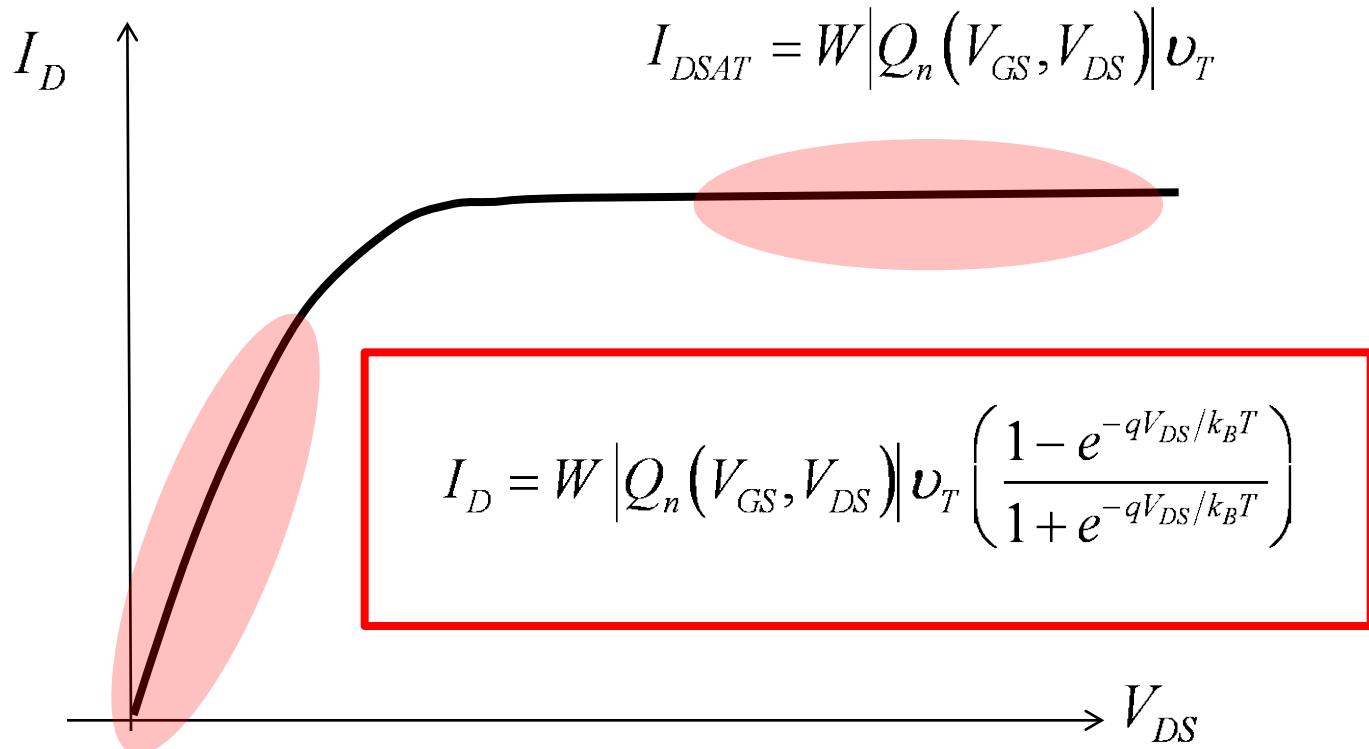
From subthreshold to above threshold

$$V_{GS} \gg V_T$$

$$C_{inv}(V_{GS} - V_T) = |Q_n(V_{GS}, V_{DS})|$$

$$V_T = V_{T0} - \delta V_{DS}$$

From subthreshold to above threshold



$$I_{DLIN} = \frac{W}{L} \mu_B |Q_n(V_{GS}, V_{DS})| V_{DS} \quad \mu_B = \frac{v_T L}{2 k_B T / q}$$

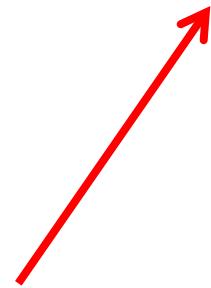
Summary

The traditional, linear region expression for I_D can be extended to the ballistic regime by replacing the mobility with the **ballistic mobility**.

The traditional, saturation region expression for I_D can be extended to the ballistic regime by replacing the high-field saturation velocity with the uni-directional thermal velocity.

Next lecture

$$I_D = W |Q_n(V_{GS}, V_{DS})| \langle v_x(V_{GS}, V_{DS}) \rangle$$



In the next lecture, we will examine the average velocity at the top of the barrier (the VS) vs. gate and drain bias.

Essentials of MOSFETs

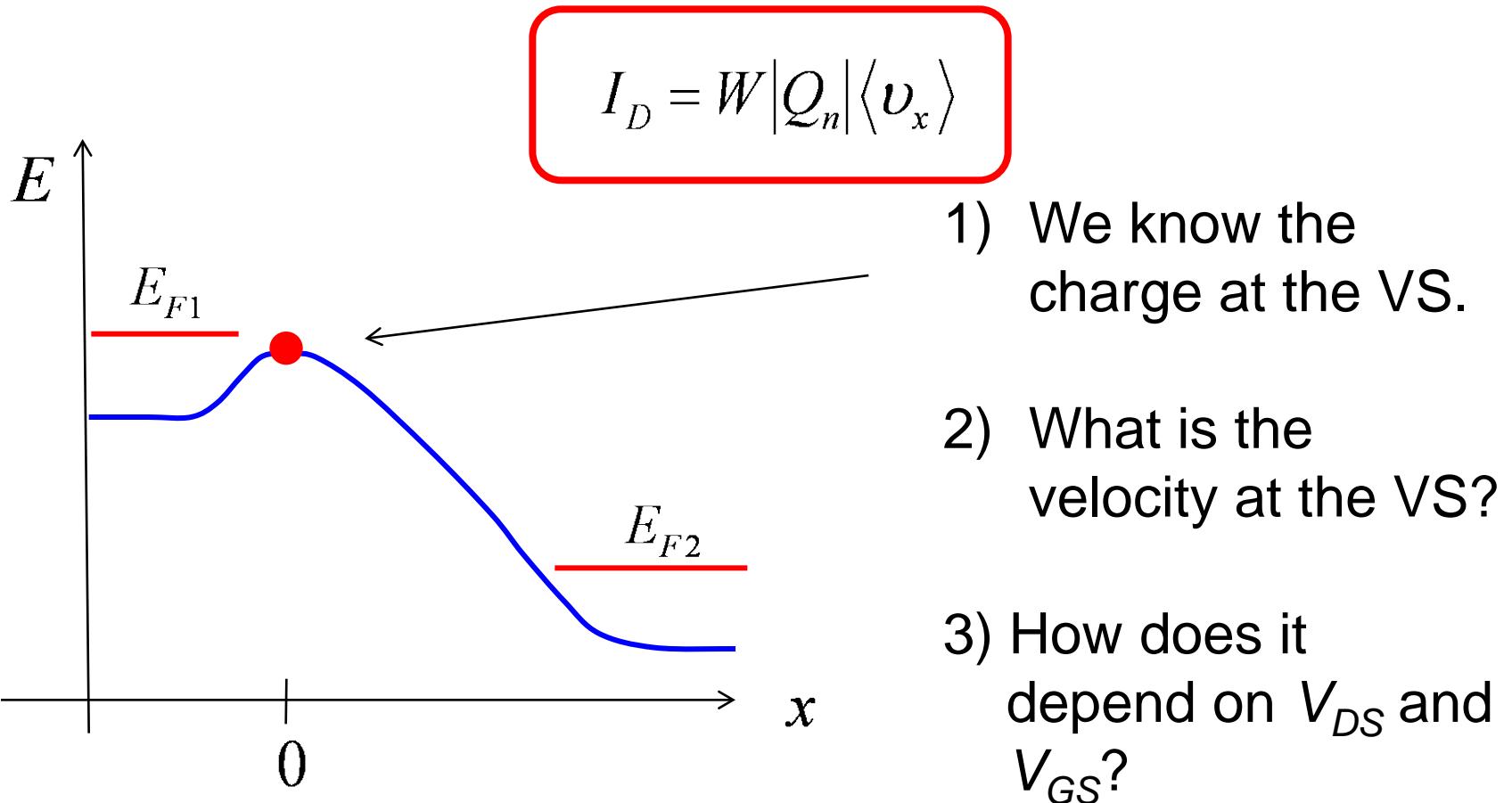
Unit 4: Transmission Theory of the MOSFET

Lecture 4.4: Velocity at the Virtual Source

Mark Lundstrom

lundstro@purdue.edu
Electrical and Computer Engineering
Purdue University
West Lafayette, Indiana USA

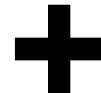
Focus on the top of the barrier (the VS)



Approach

Definition of current:

$$I_D = W |Q_n(x=0, V_{GS}, V_{DS})| \langle v_x(x=0, V_{GS}, V_{DS}) \rangle$$

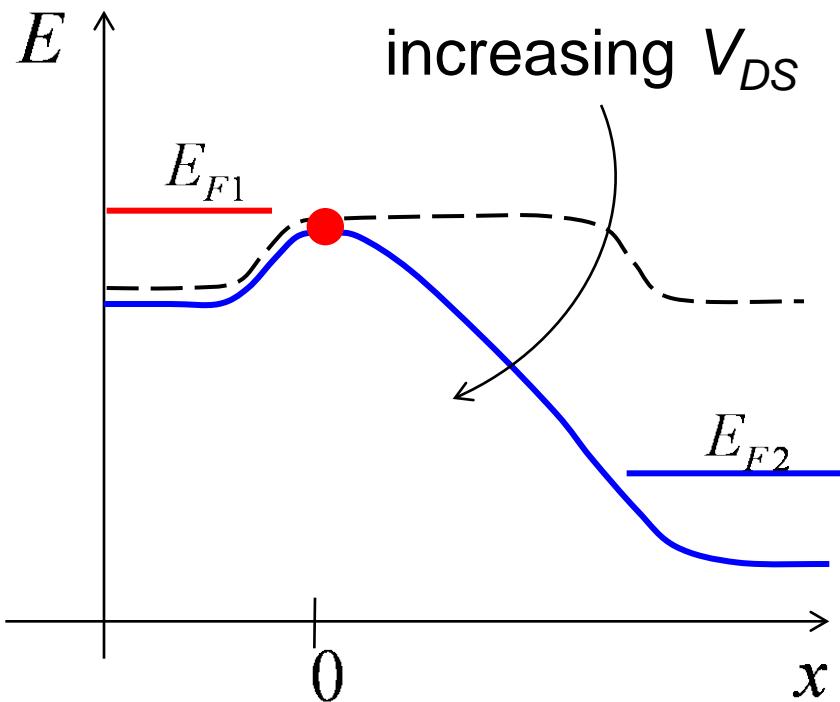


Ballistic IV:

$$I_D = W |Q_n(x=0, V_{GS}, V_{DS})| v_T \left(\frac{1 - e^{-qV_{DS}/k_B T}}{1 + e^{-qV_{DS}/k_B T}} \right)$$

(assumes nondegenerate conditions)

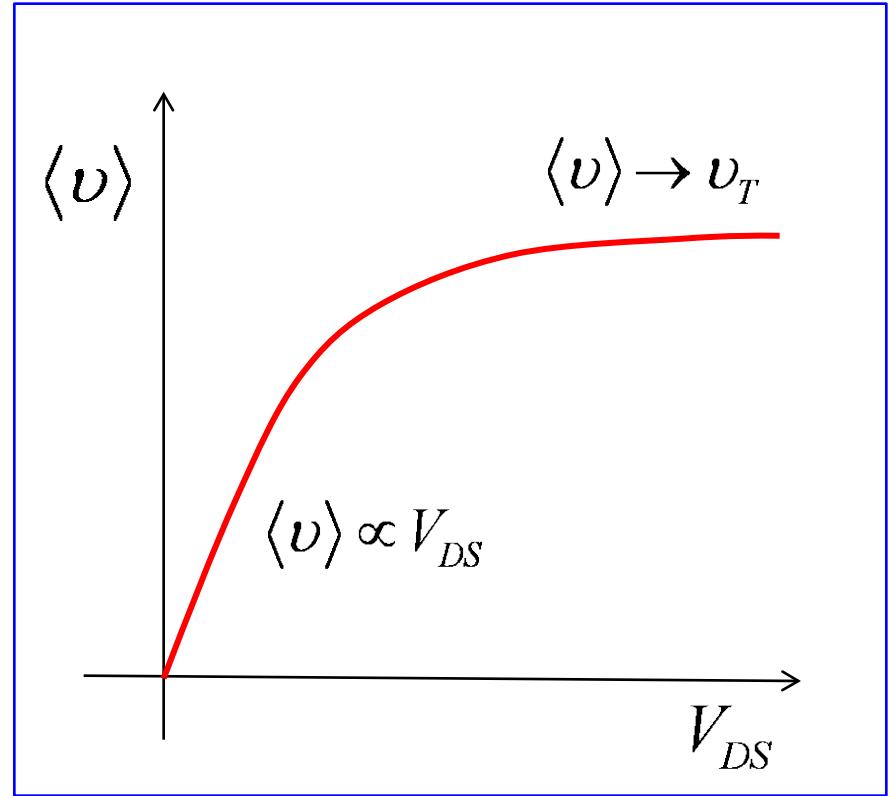
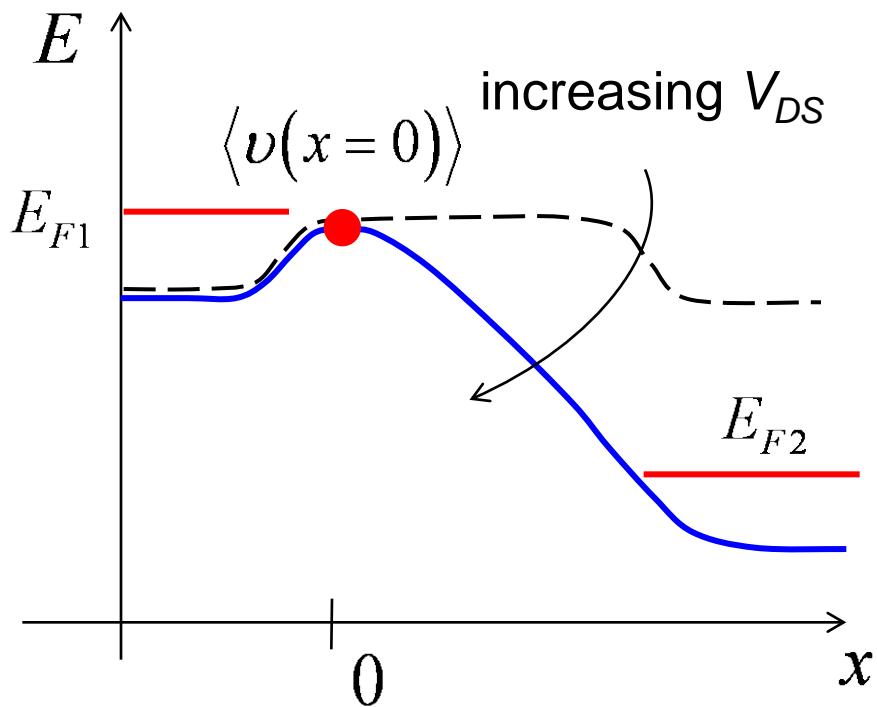
Average velocity at the VS



$$\langle v_x(x=0) \rangle = v_T \frac{(1 - e^{-qV_{DS}/k_B T})}{(1 + e^{-qV_{DS}/k_B T})}$$

(nondegenerate carrier statistics)

Velocity vs. V_{DS}



$$\langle v_x(x=0) \rangle = v_T \frac{(1 - e^{-qV_{DS}/k_B T})}{(1 + e^{-qV_{DS}/k_B T})}$$

Velocity for small V_{DS}

$$\langle v(x=0) \rangle = v_T \frac{(1 - e^{-qV_{DS}/k_B T})}{(1 + e^{-qV_{DS}/k_B T})}$$

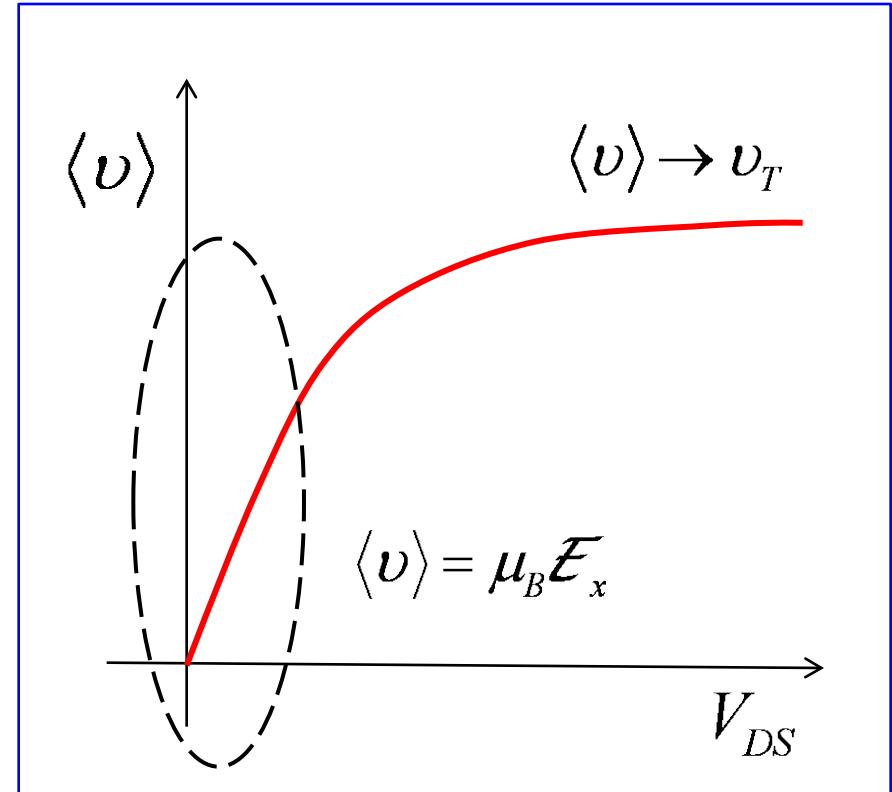
$$V_{DS} \ll k_B T / q \quad e^x \approx 1 + x$$

$$\langle v(x=0) \rangle = \frac{v_T}{2(k_B T / q)} V_{DS}$$

$$\langle v(x=0) \rangle = \left(\frac{v_T L}{2(k_B T / q)} \right) \left(\frac{V_{DS}}{L} \right)$$

$$\mu_B \equiv \frac{v_T L}{2(k_B T / q)} \text{ cm}^2/\text{V-s}$$

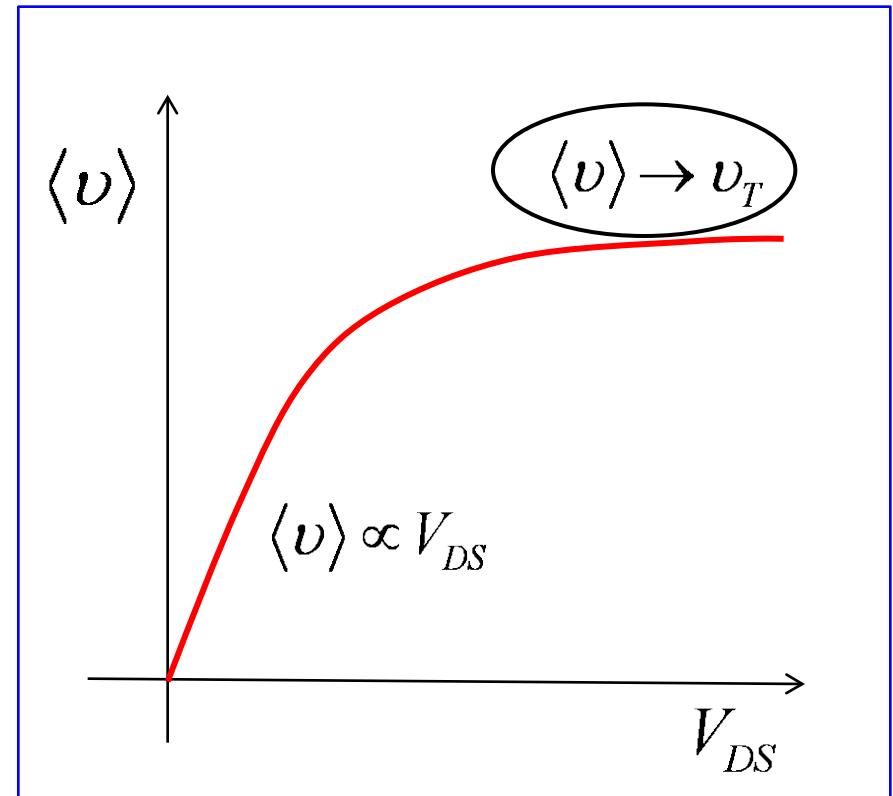
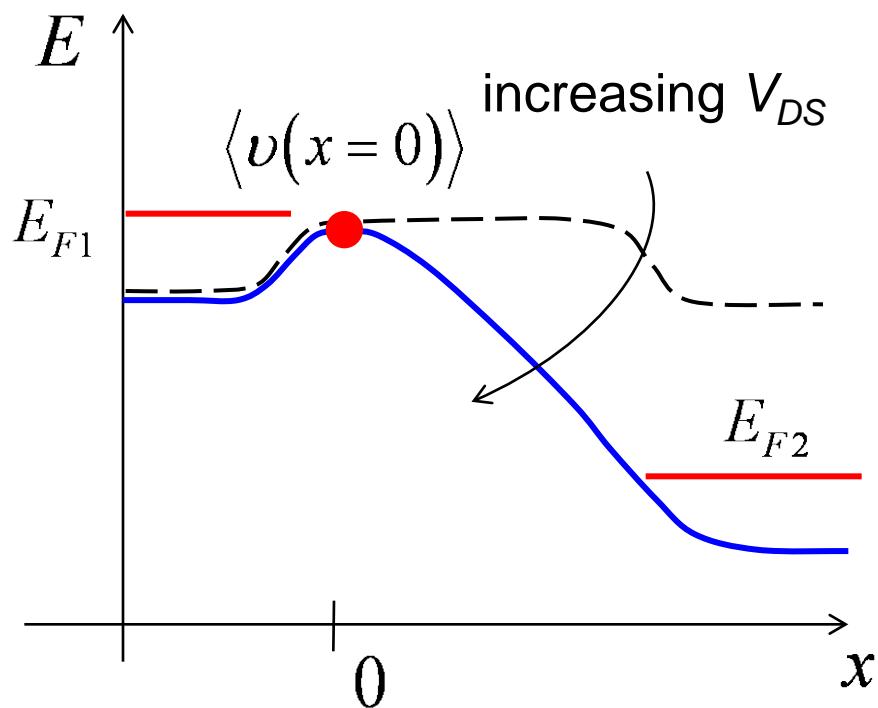
$$\langle v(x=0) \rangle = \mu_B \mathcal{E}_x$$



“ballistic mobility”

Lundstrom: 2018

Velocity for large V_{DS}

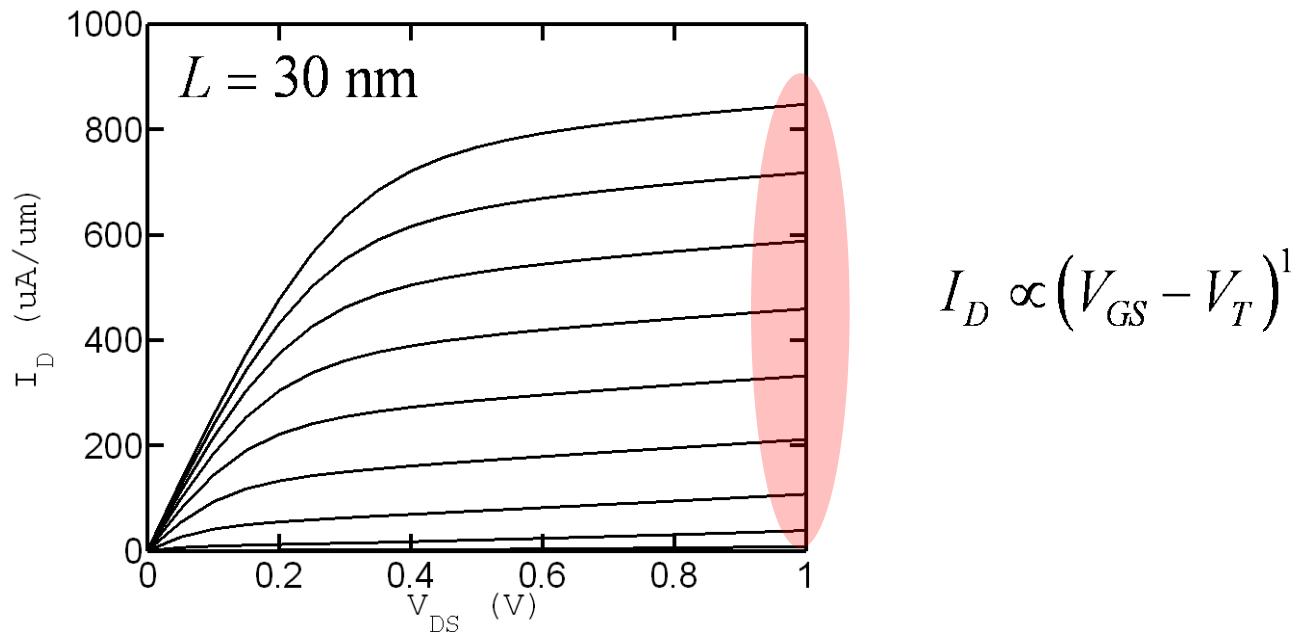


$$\langle v_x(x=0) \rangle = v_T \frac{\left(1 - e^{-qV_{DS}/k_B T}\right)}{\left(1 + e^{-qV_{DS}/k_B T}\right)}$$

$V_{DS} \gg k_B T / q$

The velocity at the VS **saturates** in a ballistic MOSFET.

The “signature” of velocity saturation in MOSFETs



ETSOI MOSFET data provided by A.
Majumdar, IBM Research, 2015.

Physics of velocity saturation

In a long channel MOSFET with a high electric field, the carrier velocity saturates at high drain bias because of **strong scattering**.

It saturates in the high-field region **near the drain**.

In a ballistic MOSFET there is no scattering, but the velocity saturates at high drain bias.

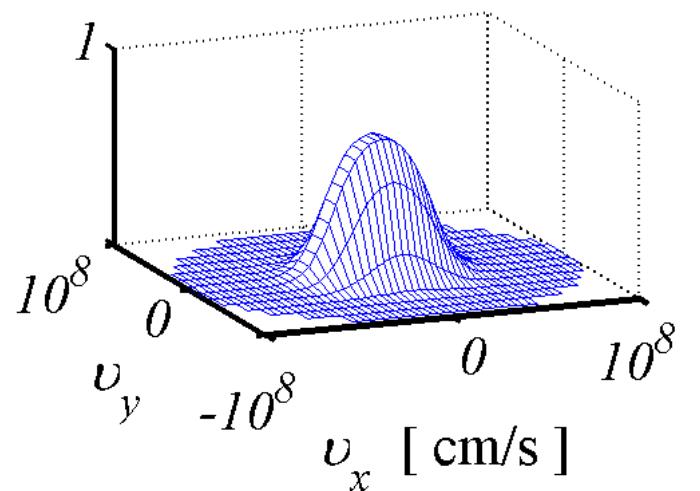
It saturates at the VS, where the E-field is zero.

What is the physics of velocity saturation in a ballistic MOSFET?

Equilibrium Maxwellian velocity distribution

$$f_0(E) = \frac{1}{1 + e^{(E - E_F)/k_B T}} \quad f_0(E) \approx e^{-(E - E_F)/k_B T} \quad E = E_C + \frac{m^* v^2}{2} \quad v^2 = v_x^2 + v_y^2$$

$$f_0(v_x, v_y) \propto e^{-m^*(v_x^2 + v_y^2)/2k_B T}$$

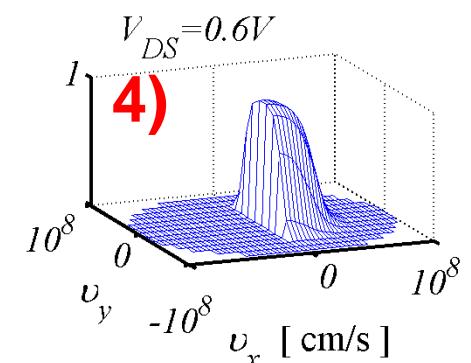
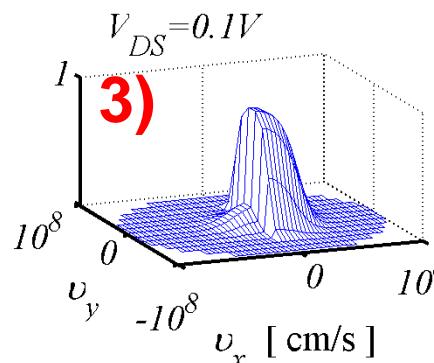
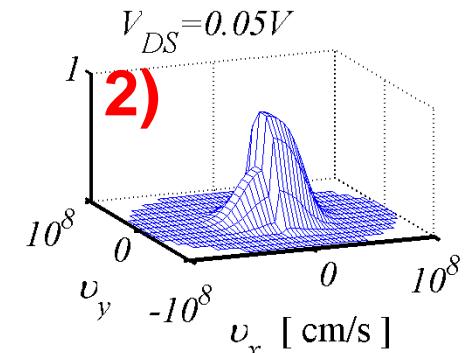
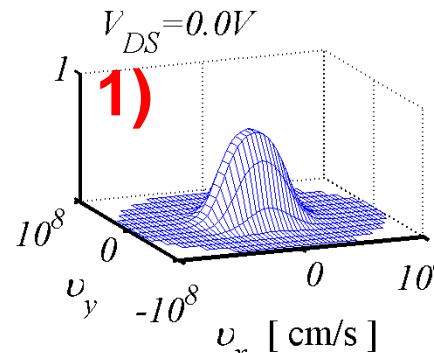
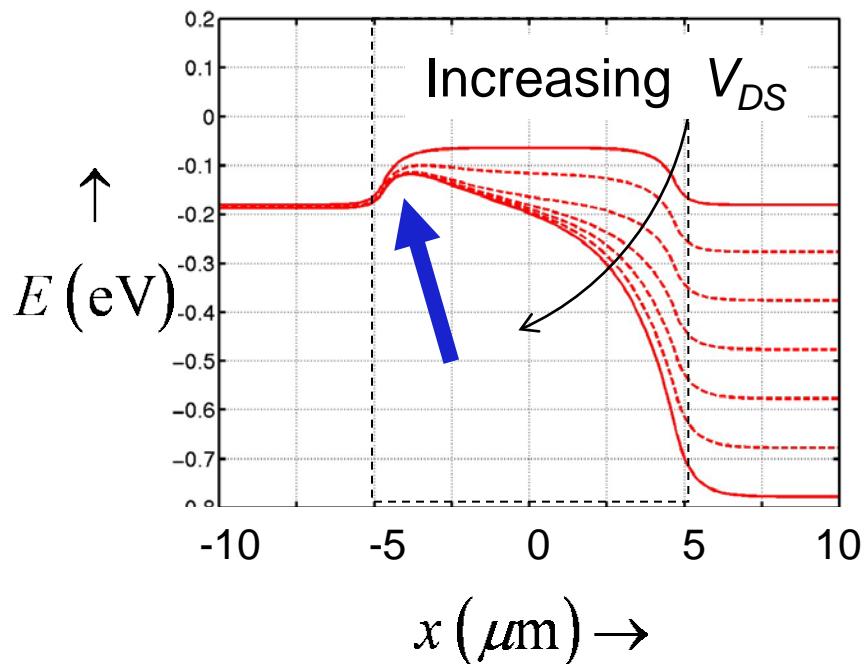


In the bulk, scattering maintains equilibrium.

Filling states at the top of the barrier

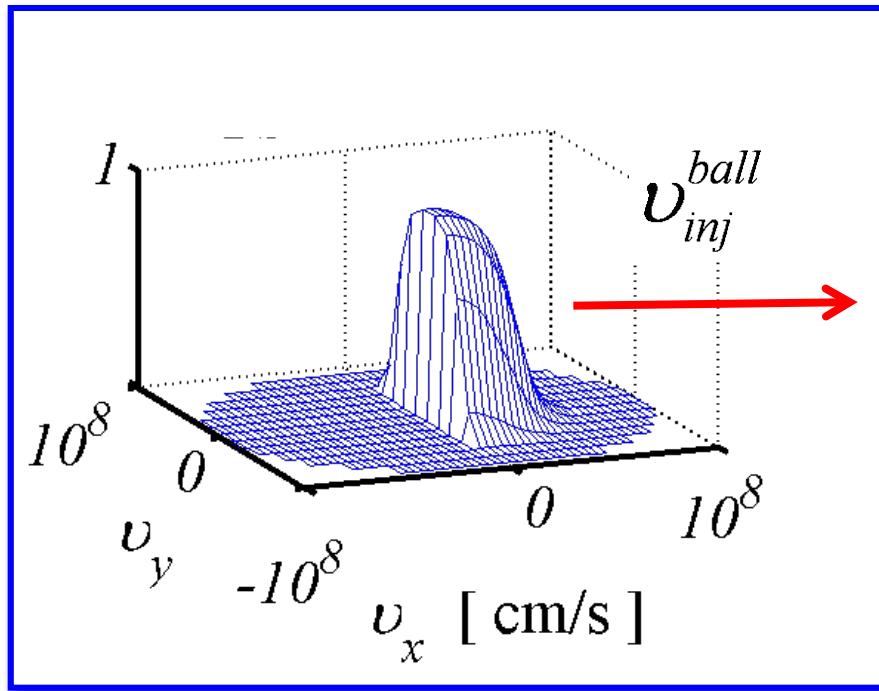
$$f(v_x, v_y)$$

E_X vs. x for $V_{GS} = 0.5V$



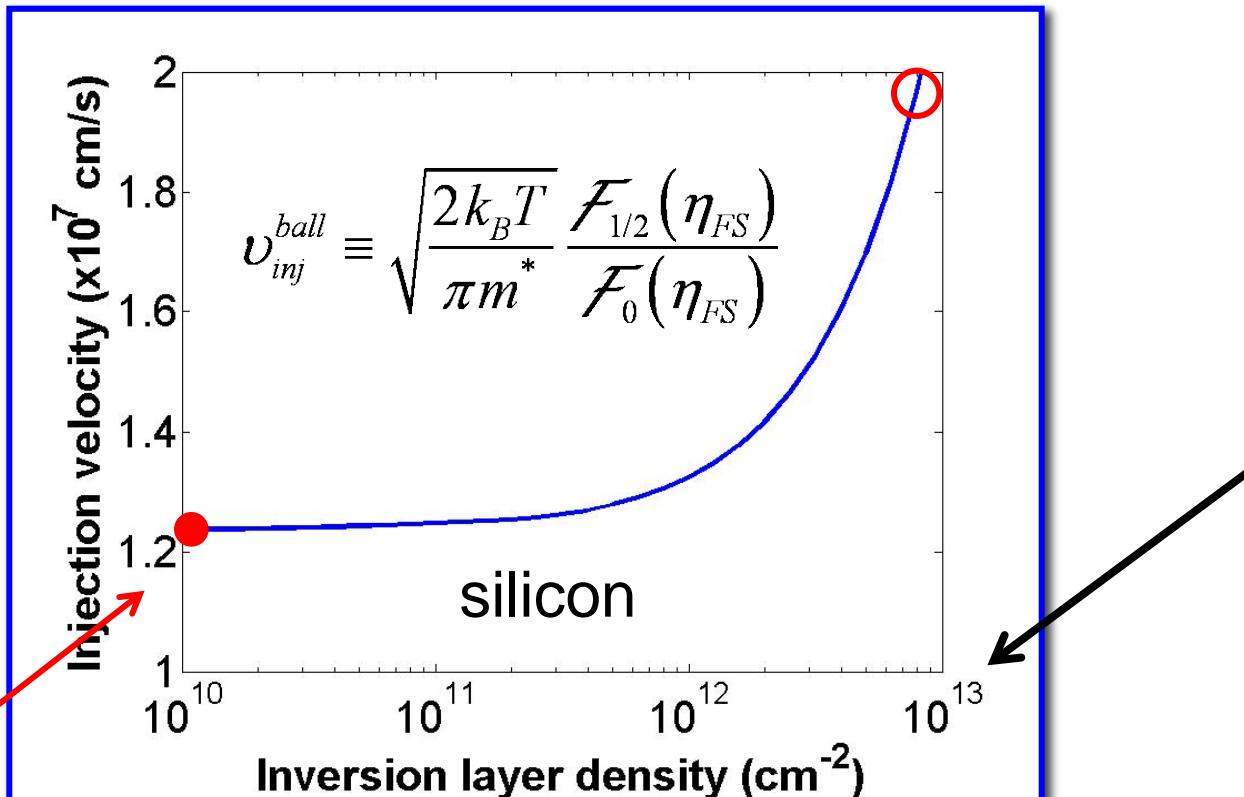
(Numerical simulations of an $L = 10$ nm double gate Si MOSFET from J.-H. Rheu and M.S. Lundstrom, *Solid-State Electron.*, **46**, 1899, 2002)

Ballistic injection velocity



$$\langle\langle v_x(0) \rangle\rangle = v_{inj}^{ball}(E_{F1}) = v_T \frac{\mathcal{F}_{1/2}(\eta_{F1})}{\mathcal{F}_0(\eta_{F1})} \quad v_{inj}^{ball} = v_T = \sqrt{\frac{2k_B T}{\pi m^*}} \quad (\text{nondegenerate})$$

Gate voltage dependent ballistic injection velocity



$$m^* = 0.19 m_0$$

For more discussion

To see how Fermi-Dirac statistics are included in the analysis, see:

Lecture 14 in:

Mark Lundstrom, *Fundamentals of Nanotransistors*,
World Scientific Publishing Co., Singapore, 2018.

Summary

In a ballistic MOSFET, the velocity saturates for high drain voltages.

But it saturates at the top of the barrier (the VS) where the electric field is zero.

The velocity saturates at the **ballistic injection velocity**, which is a key figure of merit for a transistor.

Next topic

We knew how to treat the diffusive case, where there is a lot of carrier scattering (with the traditional model).

We now know how to treat the ballistic case, where there is no scattering (with the ballistic model).

How do we treat MOSFETs between these two limits?

That is the subject of the next lecture.

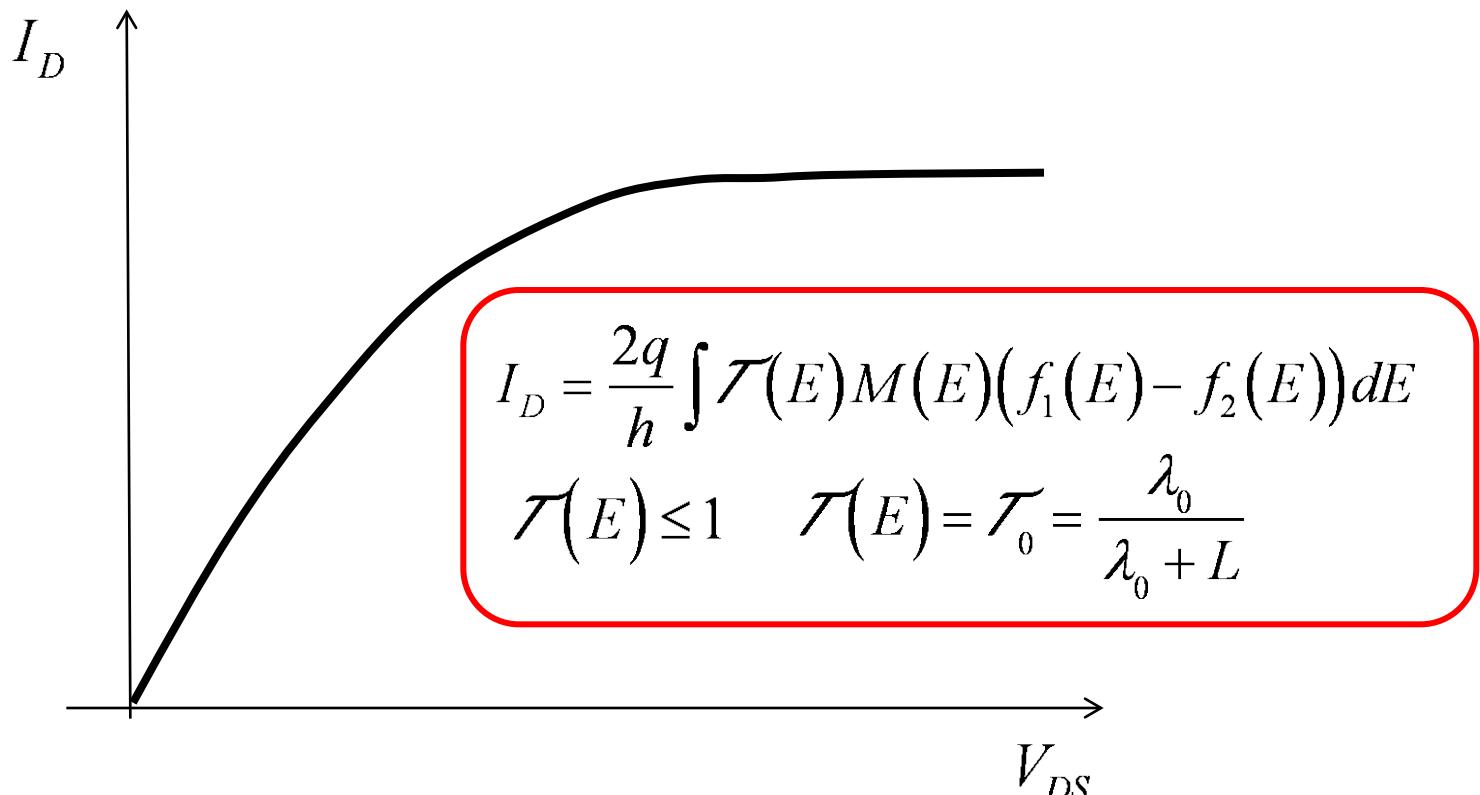
Unit 4:
Transmission Theory of the MOSFET

Lecture 4.5:
Transmission Theory
of the MOSFET

Mark Lundstrom

lundstro@purdue.edu
Electrical and Computer Engineering
Purdue University
West Lafayette, Indiana USA

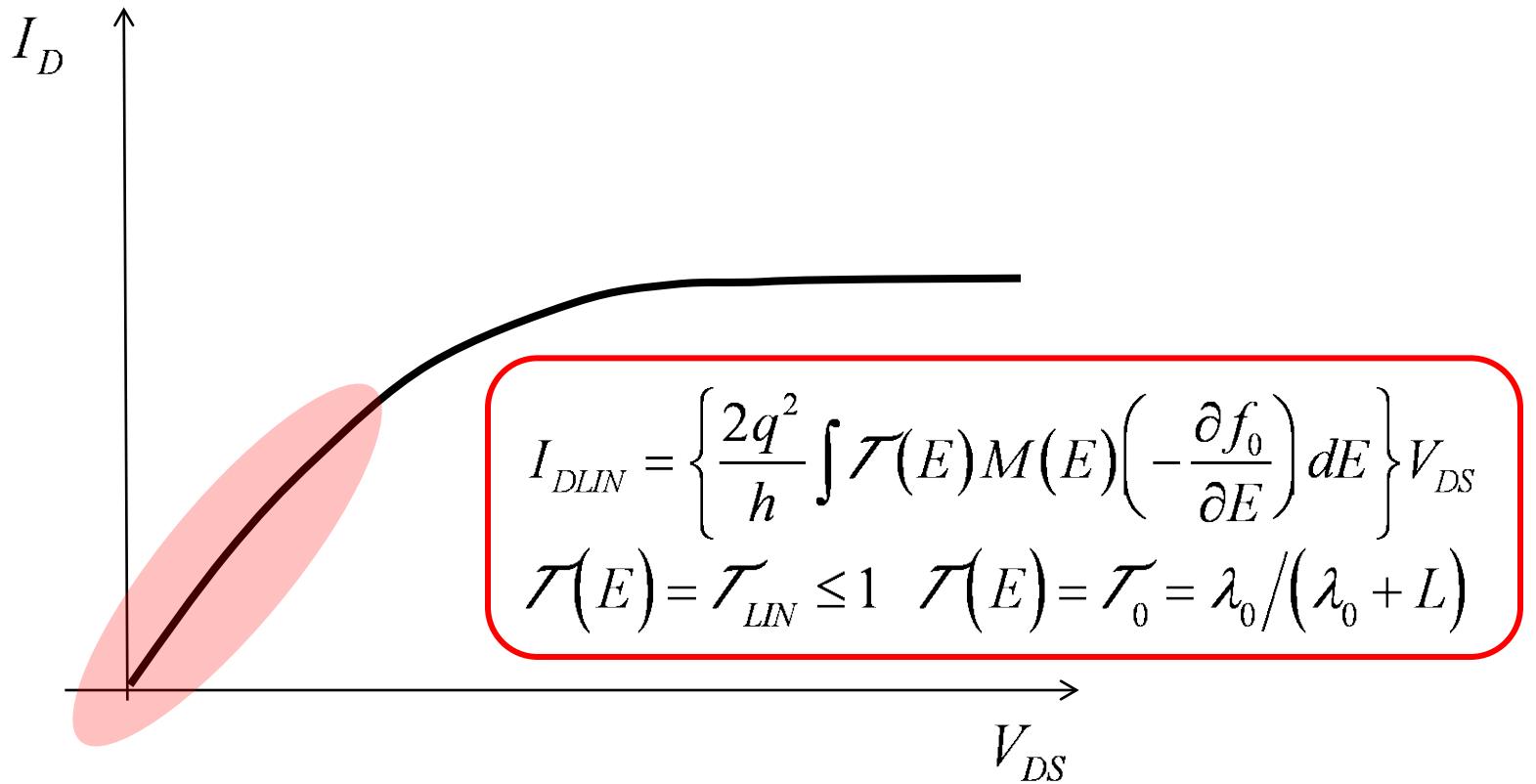
Transmission theory



1) Linear region

2) Saturation region

1) Linear region



$$I_{DLIN} = W \left(\frac{v_T}{2(k_B T/q)} \right) |Q_n| V_{DS} \rightarrow W \mathcal{T}_{LIN} \left(\frac{v_T}{2(k_B T/q)} \right) |Q_n| V_{DS} \quad ?$$

Ballistic to diffusive linear current

$$I_{DLIN} = W \mathcal{T}_{LIN} \left(\frac{v_T}{2(k_B T / q)} \right) |Q_n| V_{DS}$$

$$\mathcal{T}_{LIN} = \frac{\lambda_0}{\lambda_0 + L}$$

$$I_{DLIN} = \frac{W}{\lambda_0 + L} \left(\frac{v_T \lambda_0}{2(k_B T / q)} \right) |Q_n| V_{DS}$$

$$I_{DLIN} \propto \frac{W}{\lambda_0 + L} \text{ not } \frac{W}{L}$$



μ_n

$$I_{DLIN} = \frac{W}{L + \lambda_0} \mu_n |Q_n| V_{DS}$$

$$L \rightarrow L + \lambda_0$$

Alternative formulation

$$I_{DLIN} = \frac{W}{L + \lambda_0} \mu_n |Q_n| V_{DS} \quad \mathcal{T}_{LIN} = \frac{\lambda_0}{\lambda_0 + L}$$

$$I_{DLIN} = \frac{W}{L} \left(\frac{L}{L + \lambda_0} \right) \mu_n |Q_n| V_{DS} \quad \mu_n = \left(\frac{v_T \lambda_0}{2(k_B T / q)} \right)$$

$$I_{DLIN} = \frac{W}{L} \left(\frac{1}{1 + \lambda_0 / L} \right) \mu_n |Q_n| V_{DS} \quad \frac{\lambda_0}{L \mu_n} = \frac{1}{\mu_B}$$

$$I_{DLIN} = \frac{W}{L} \left(\frac{1}{1/\mu_n + \lambda_0 / (L \mu_n)} \right) |Q_n| V_{DS} \quad \mu_B = \left(\frac{v_T L}{2(k_B T / q)} \right)$$

Apparent mobility

$$I_{DLIN} = \frac{W}{L} \left(\frac{1}{1/\mu_n + 1/\mu_B} \right) |Q_n| V_{DS} \quad \mu_n = \left(\frac{v_T \lambda_0}{2(k_B T / q)} \right) \quad \mu_B = \left(\frac{v_T L}{2(k_B T / q)} \right)$$

$$I_{DLIN} = \frac{W}{L} \mu_{app} |Q_n| V_{DS}$$

$$\frac{1}{\mu_{app}} = \frac{1}{\mu_n} + \frac{1}{\mu_B}$$

Example

Estimate the apparent mobility for a 22 nm N-MOSFET.

$$\mu_n \approx 200 \text{ cm}^2/\text{V-s}$$

$$\mu_B = \frac{v_T L}{2 k_B T / q}$$

$$v_T = \sqrt{\frac{2k_B T}{\pi m_t^*}} = 1.2 \times 10^7 \text{ cm/s}$$

$$\mu_B \approx 500 \text{ cm}^2/\text{V-s}$$

$$\frac{1}{\mu_{app}} = \frac{1}{\mu_n} + \frac{1}{\mu_B} = \frac{1}{200} + \frac{1}{500}$$

$$\mu_{app} \approx 140 \frac{\text{cm}^2}{\text{V-s}}$$

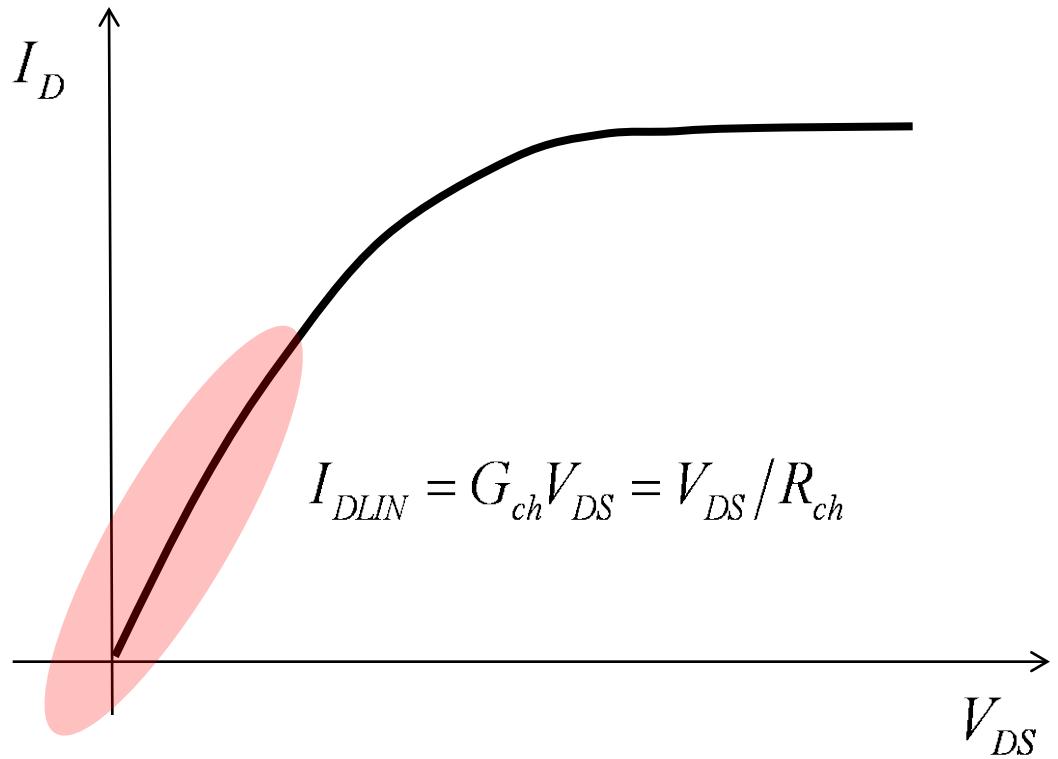
This device operates in the quasi-ballistic regime.

(Assumes confinement in the <100> direction.)

Exercise

Repeat the previous exercise for a MOSFET with a 10 nm long channel (assume the same diffusive mobility).

Channel resistance



$$I_{DLIN} = \frac{W}{L} \left(\frac{1}{1/\mu_n + 1/\mu_B} \right) |Q_n| V_{DS}$$

$$G_{ch} = \frac{W}{L} \left(\frac{1}{1/\mu_n + 1/\mu_B} \right) |Q_n|$$

$$R_{ch} = \frac{1}{G_{ch}} = \left(\frac{1/\mu_n + 1/\mu_B}{|Q_n|} \right) \frac{L}{W}$$

$$R_{ch} = R_{diff} + R_B$$

Channel resistance

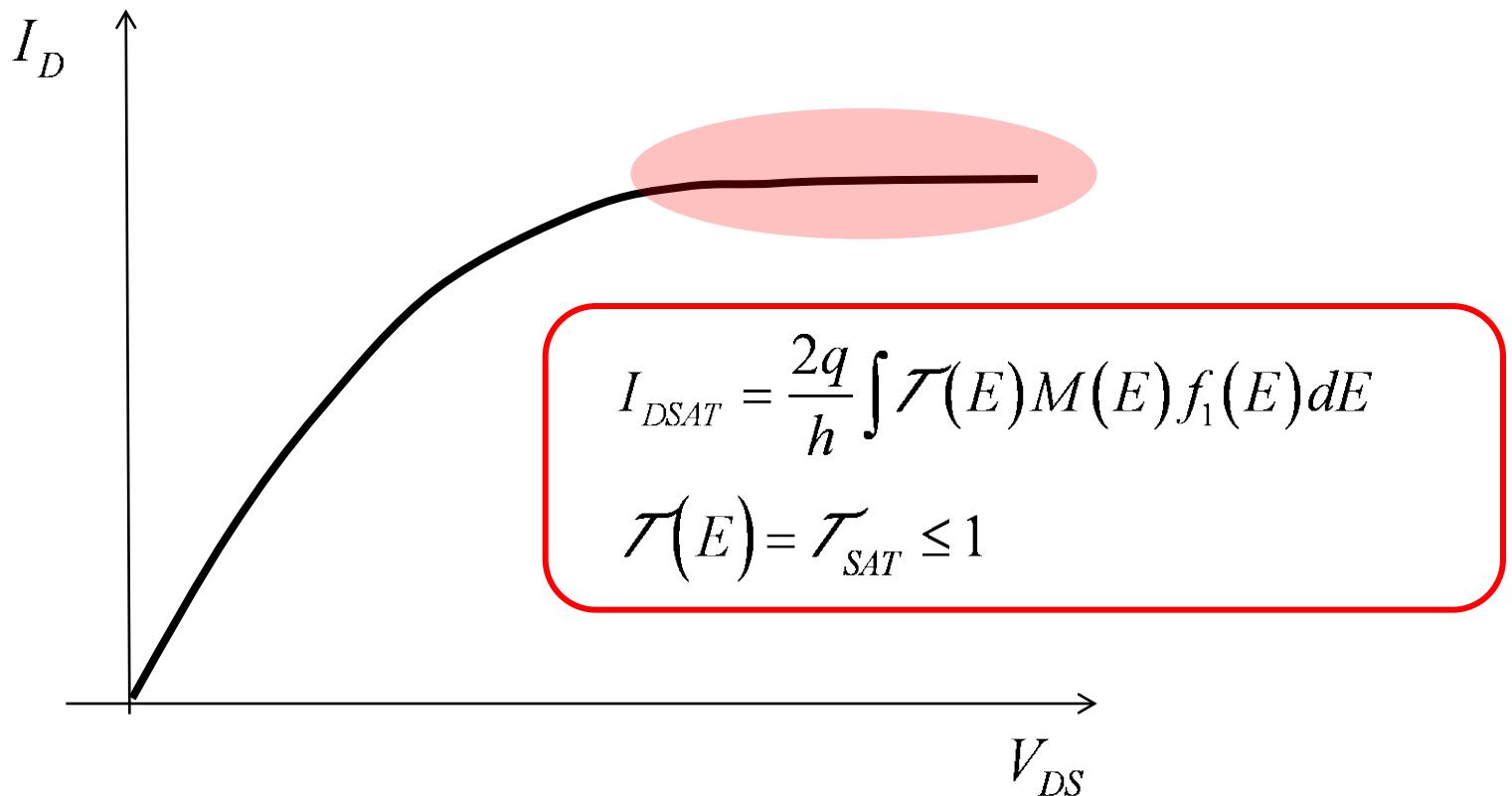
$$R_{ch} = \frac{1}{G_{ch}} = \left(\frac{1/\mu_n + 1/\mu_B}{|Q_n|} \right) \frac{L}{W} \quad \mu_n = \left(\frac{v_T \lambda_0}{2(k_B T / q)} \right) \quad \mu_B = \left(\frac{v_T L}{2(k_B T / q)} \right)$$

$$R_{ch} = R_{diff} + R_B \quad R_{diff} = \left(\frac{1}{\mu_n |Q_n|} \right) \frac{L}{W} \quad R_B = \left(\frac{1}{\mu_B |Q_n|} \right) \frac{L}{W}$$

$$R_B = \left(\frac{2 k_B T / q}{v_T |Q_n|} \right) \frac{1}{W} \quad \text{independent of channel length}$$

$$R_{ch}(L) = R_{diff}(L) + R_B$$

2) Saturation region



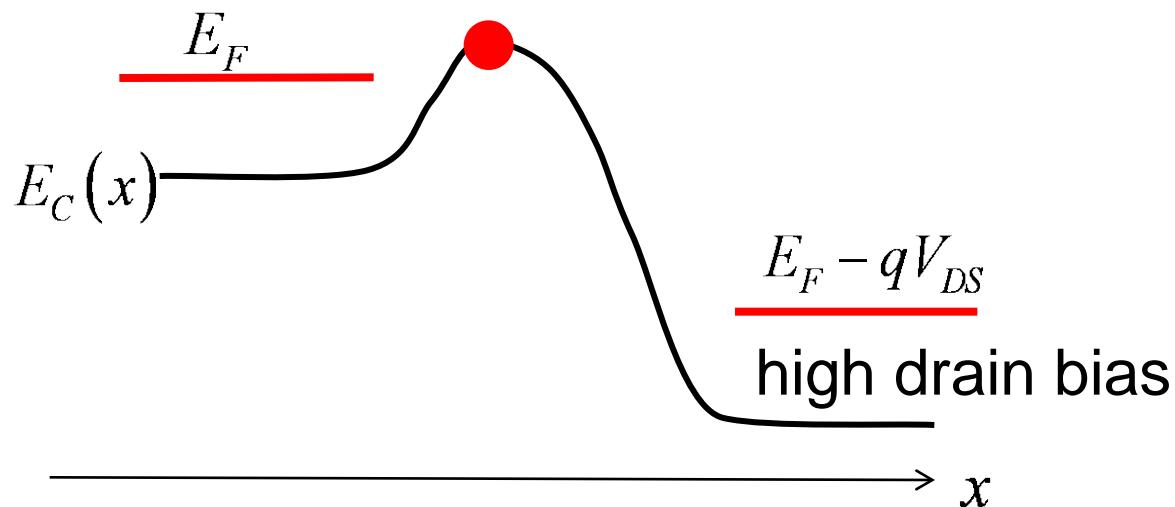
$$I_{DSAT} = W |Q_n| v_T \rightarrow \mathcal{T}_{SAT} W |Q_n| v_T \quad ?$$

This is wrong!

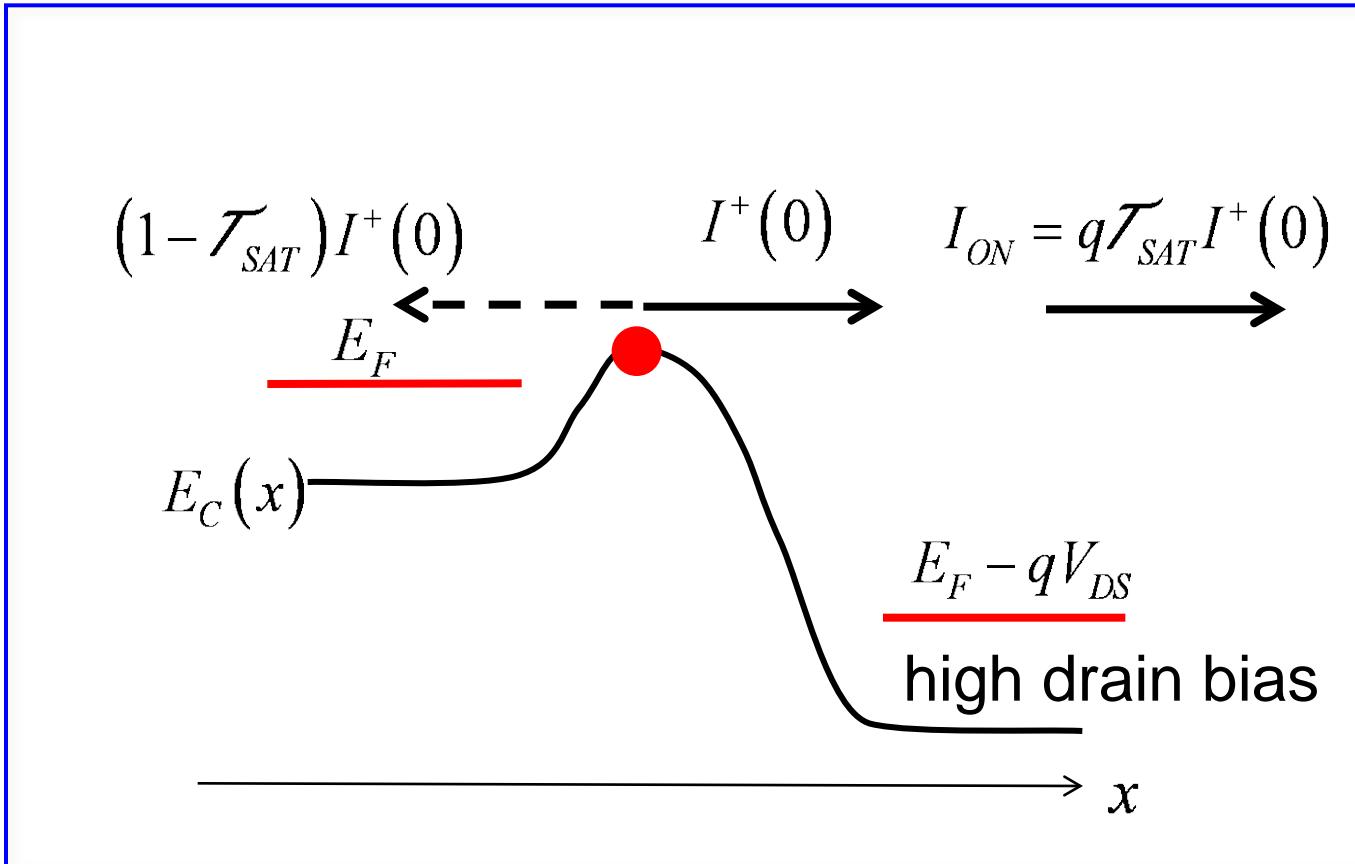
Focus on the VS

MOS electrostatics

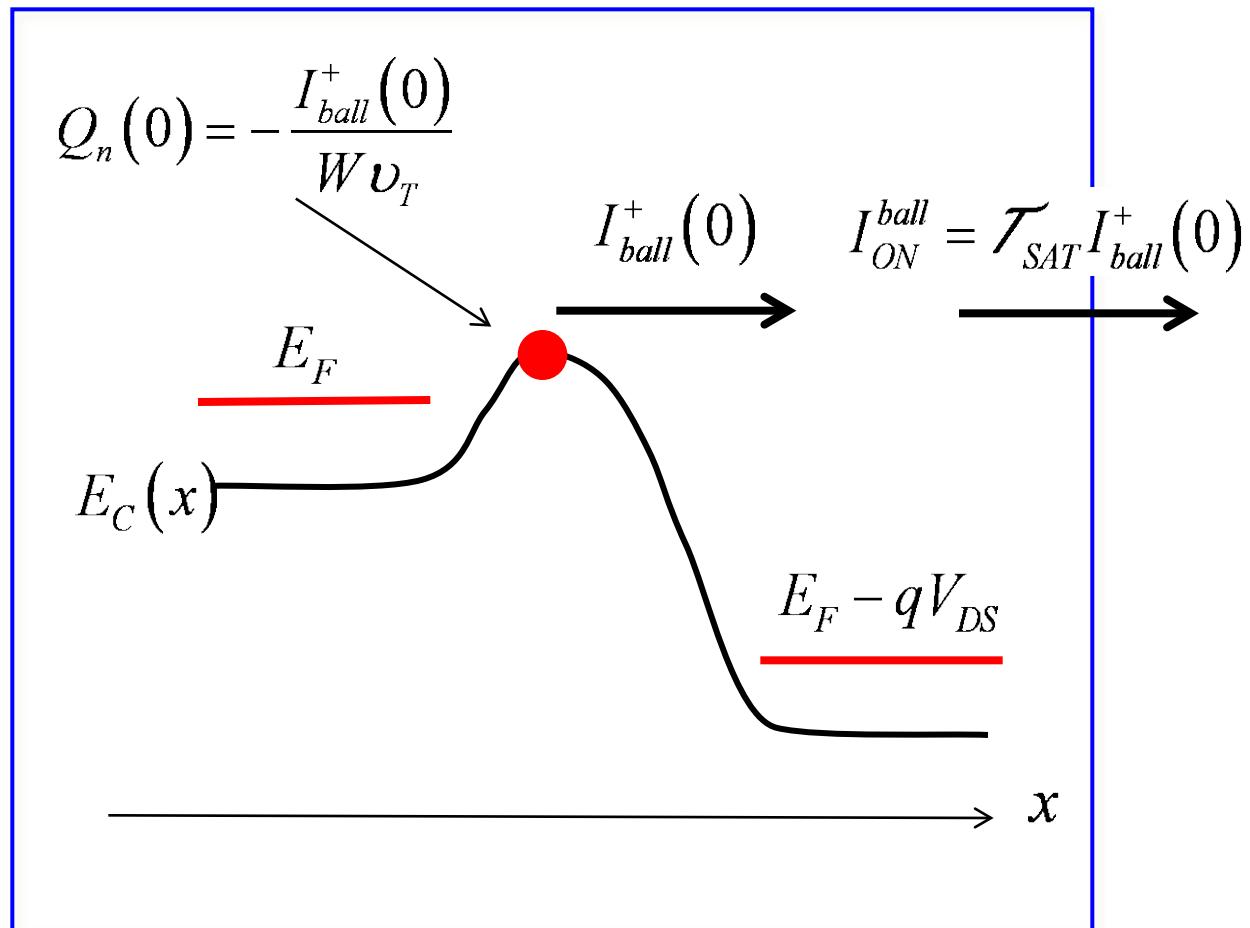
$$Q_n = -C_{ox}(V_{GS} - V_T)$$



On-current and transmission



Ballistic case first



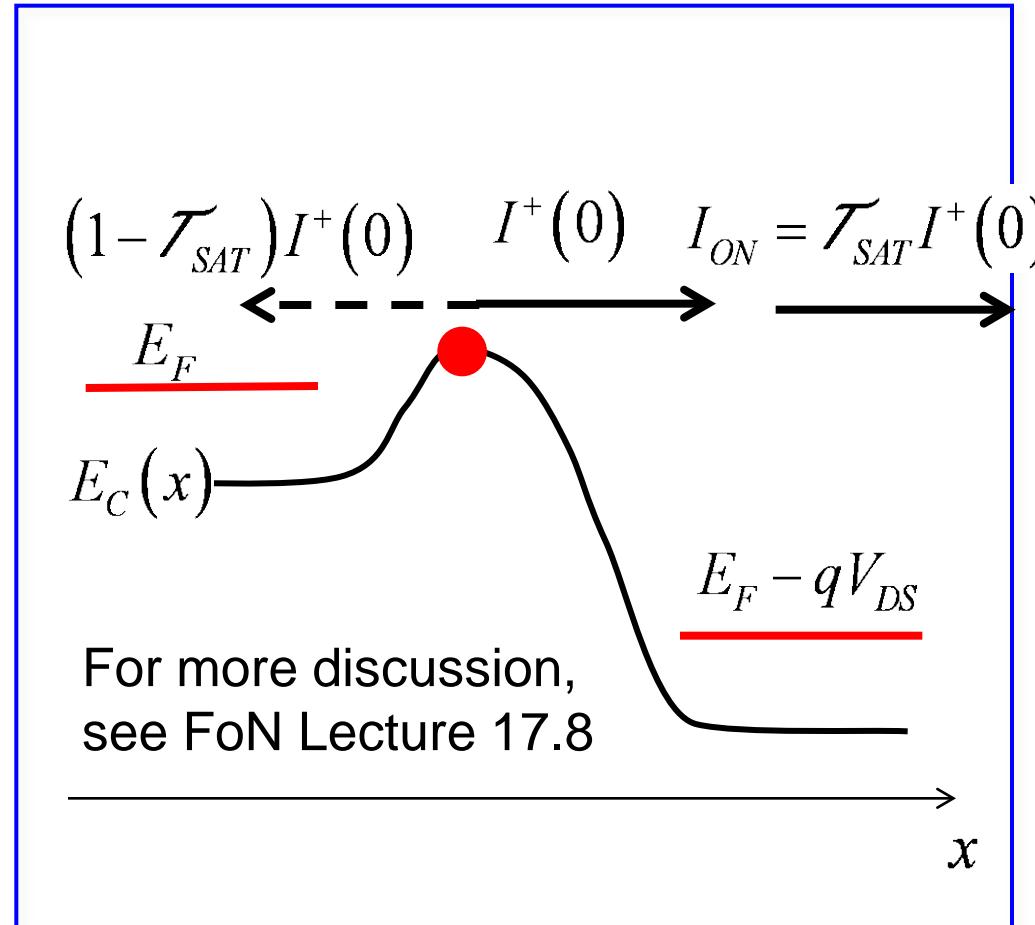
Now include scattering

$$Q_n(0) = -\frac{I_{ball}^+(0)}{Wv_T}$$

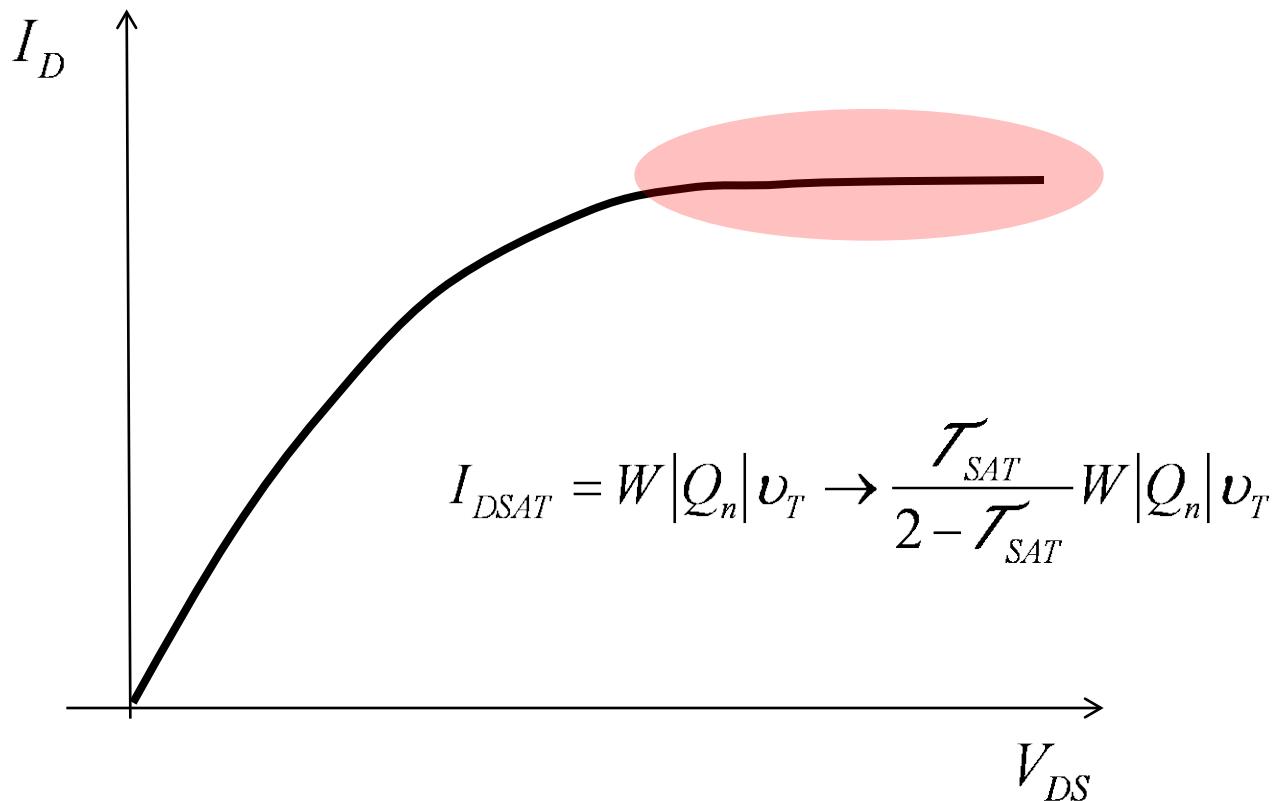
$$Q_n(0) = -\frac{I^+(0) + (1 - \mathcal{T}_{SAT})I^+(0)}{Wv_T}$$

$$I^+(0) = \frac{I_{ball}^+(0)}{(2 - \mathcal{T}_{SAT})}$$

$$I_{ON} = \left(\frac{\mathcal{T}_{SAT}}{2 - \mathcal{T}_{SAT}} \right) I_{ball}^+(0)$$

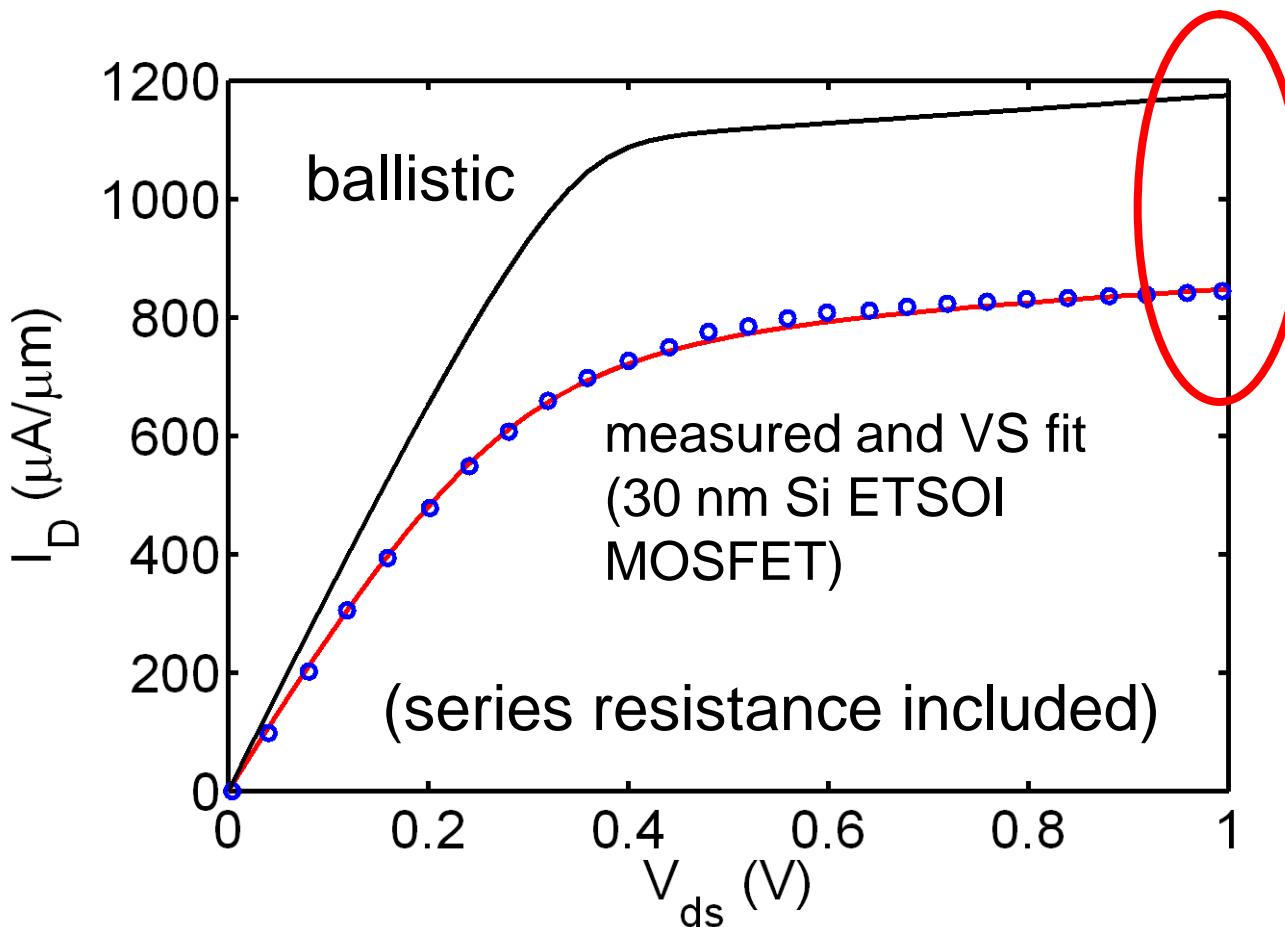


2) Saturation region



The extra factor accounts for MOS electrostatics. The charge at the VS is determined by electrostatics – not backscattering.

Transmission in saturation



$$\frac{I_{ON}}{I_{ON}^{ball}} = 0.70$$

$$\frac{I_{DSAT}}{I_{DSAT}^{ball}} = \left(\frac{\mathcal{T}_{SAT}}{2 - \mathcal{T}_{SAT}} \right)$$

$$\mathcal{T}_{SAT} = 0.82$$

Fig. 15.2 *Fundamentals of Nanotransistors*, World Scientific Lecture Notes, 2015.

The injection velocity

$$I_{DSAT} = W |Q_n| v_{inj} \quad v_{inj} = \frac{\tau_{SAT}}{2 - \tau_{SAT}} v_T$$

- 1) Traditional (velocity saturation) MOSFET model: $v_{inj} = v_{sat}$
- 2) Ballistic MOSFET model: $v_{inj} = v_T$
- 3) Transmission model: $v_{inj} = \frac{\tau_{SAT}}{2 - \tau_{SAT}} v_T < v_T$

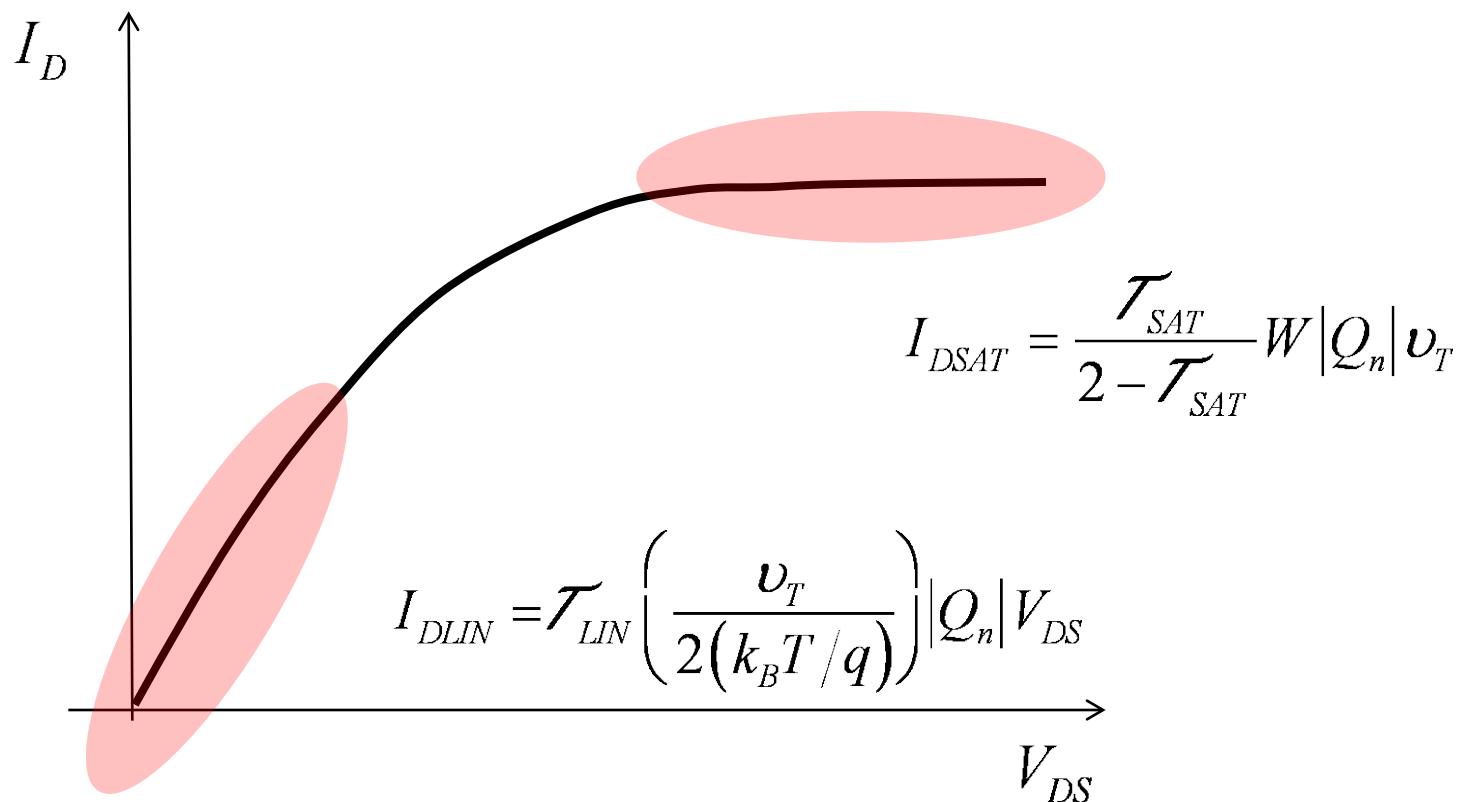
Example

What is the injection velocity of the $L = 30$ nm ETSOI MOSFET discussed earlier?

$$\mathcal{T}_{SAT} = 0.82 \quad v_T = \sqrt{\frac{2k_B T}{\pi m_t}} = 1.2 \times 10^7 \text{ cm/s}$$

$$v_{inj} = \frac{\mathcal{T}_{SAT}}{2 - \mathcal{T}_{SAT}} v_T = 0.7 v_T = 0.84 \times 10^7 \text{ cm/s}$$

Linear and saturation region transmission



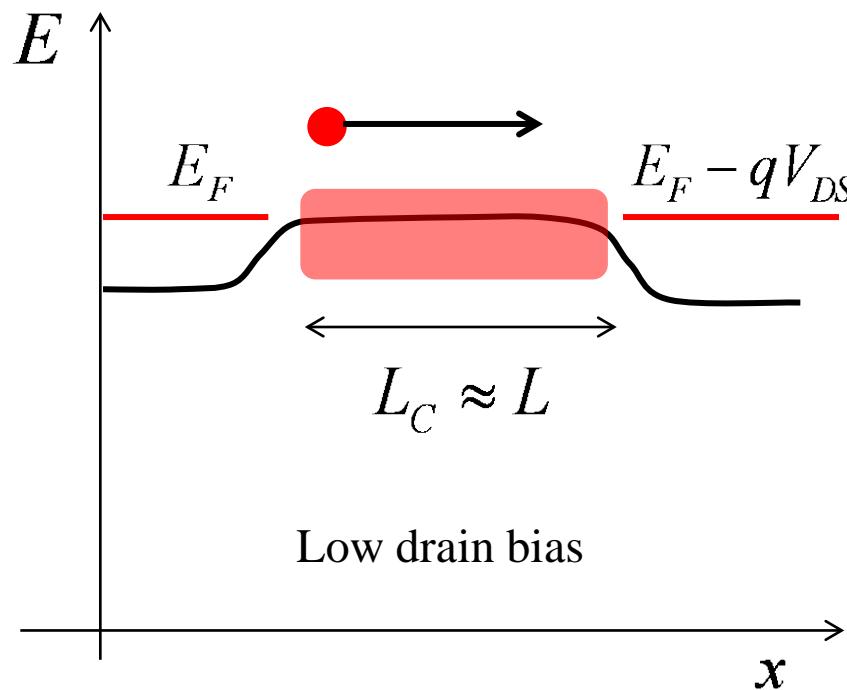
$\mathcal{T}_{LIN} < \mathcal{T}_{SAT}$ Why?

Scattering under low and high V_{DS}

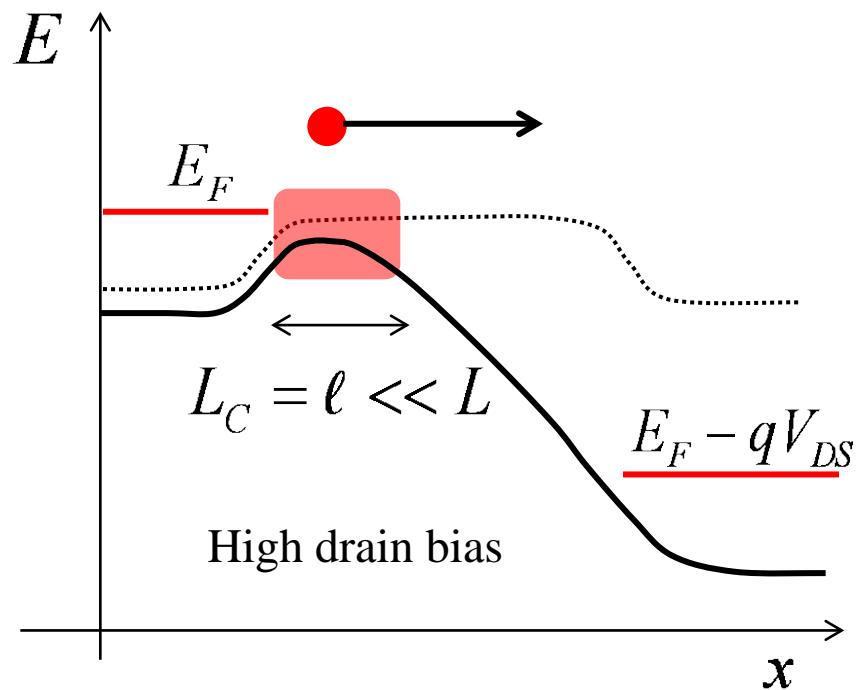
$$\mathcal{T}_{LIN} = \frac{\lambda_0}{\lambda_0 + L}$$

$$\mathcal{T}(V_{DS}) = \frac{\lambda_0}{\lambda_0 + L_C(V_{DS})}$$

$$\mathcal{T}_{SAT} = \frac{\lambda_0}{\lambda_0 + \ell}$$

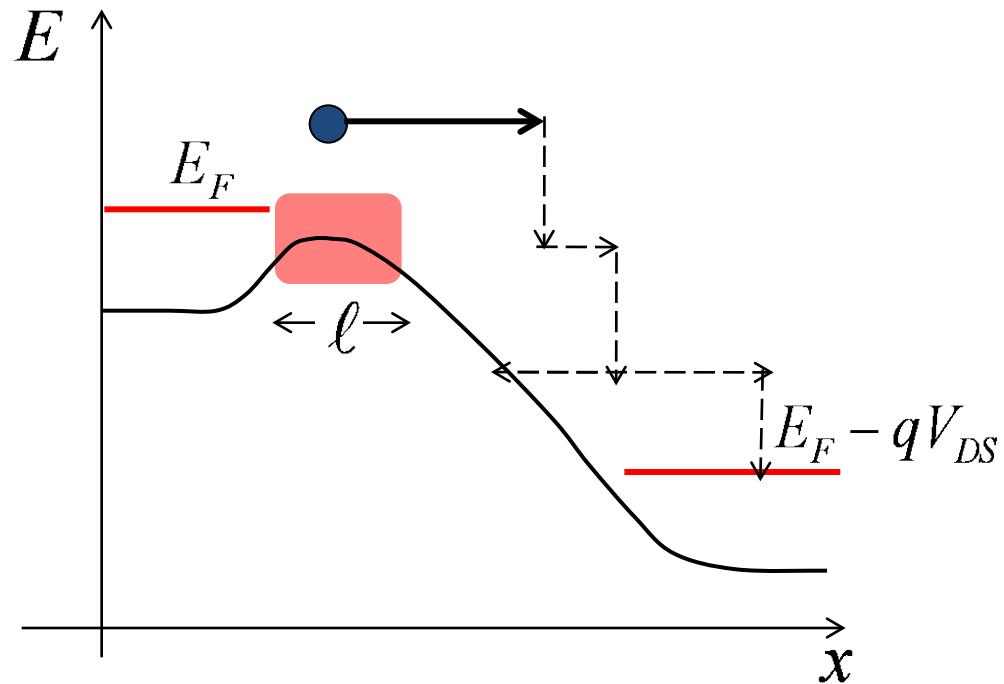


Low drain bias



High drain bias

Operation near the “ballistic limit”

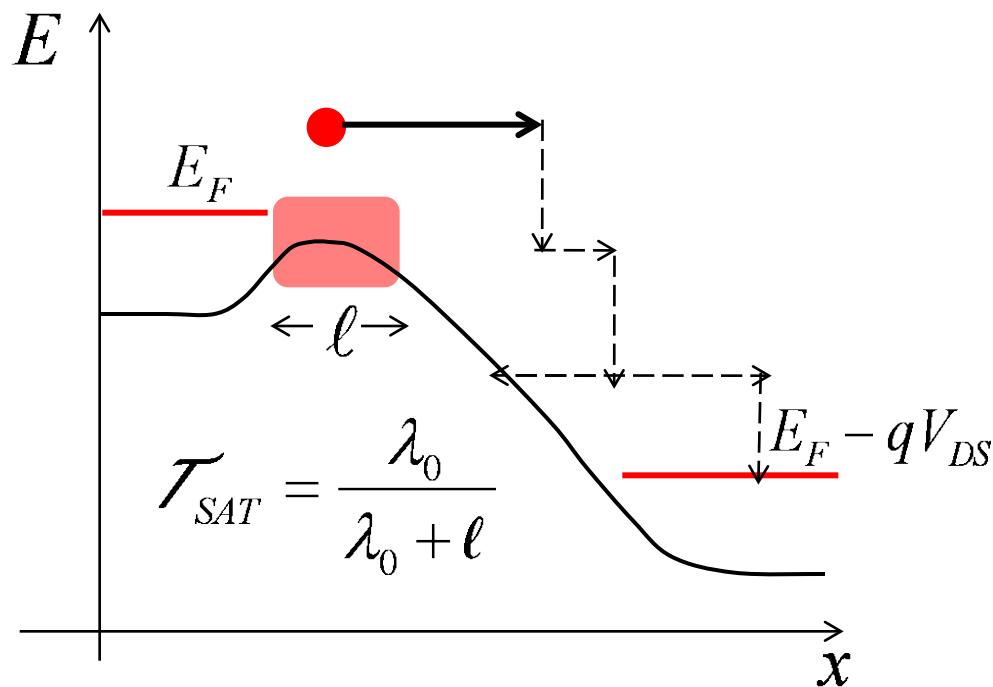


$$\mathcal{T}_{SAT} = \frac{\lambda_0}{\lambda_0 + \ell}$$

Operation near the ballistic limit current just means that $\mathcal{T}_{SAT} \rightarrow 1$,

it does not imply that there is little scattering.

Is mobility relevant at the nanoscale?



- mobility is related to the near-eq. MFP
- backscattering in the critical region is also controlled by the near-eq. MFP.
- mobility determines the on-current
- but the MFP near the drain is very short.

Summary

- Scattering lowers the drain current.
- Transmission is **higher** under high drain bias than under low drain bias.
- Under low drain bias, transmission is determined by backscattering in the **entire channel**.
- Under high drain bias, transmission is determined by backscattering in a **short, “bottleneck region”** near the top of the barrier.
- **Apparent mobility and injection velocity** allow us to use traditional MOSFET theory.

Next topic

In the next lecture, we will revisit the VS model.

Essentials of MOSFETs

Unit 4: Transmission Theory of the MOSFET

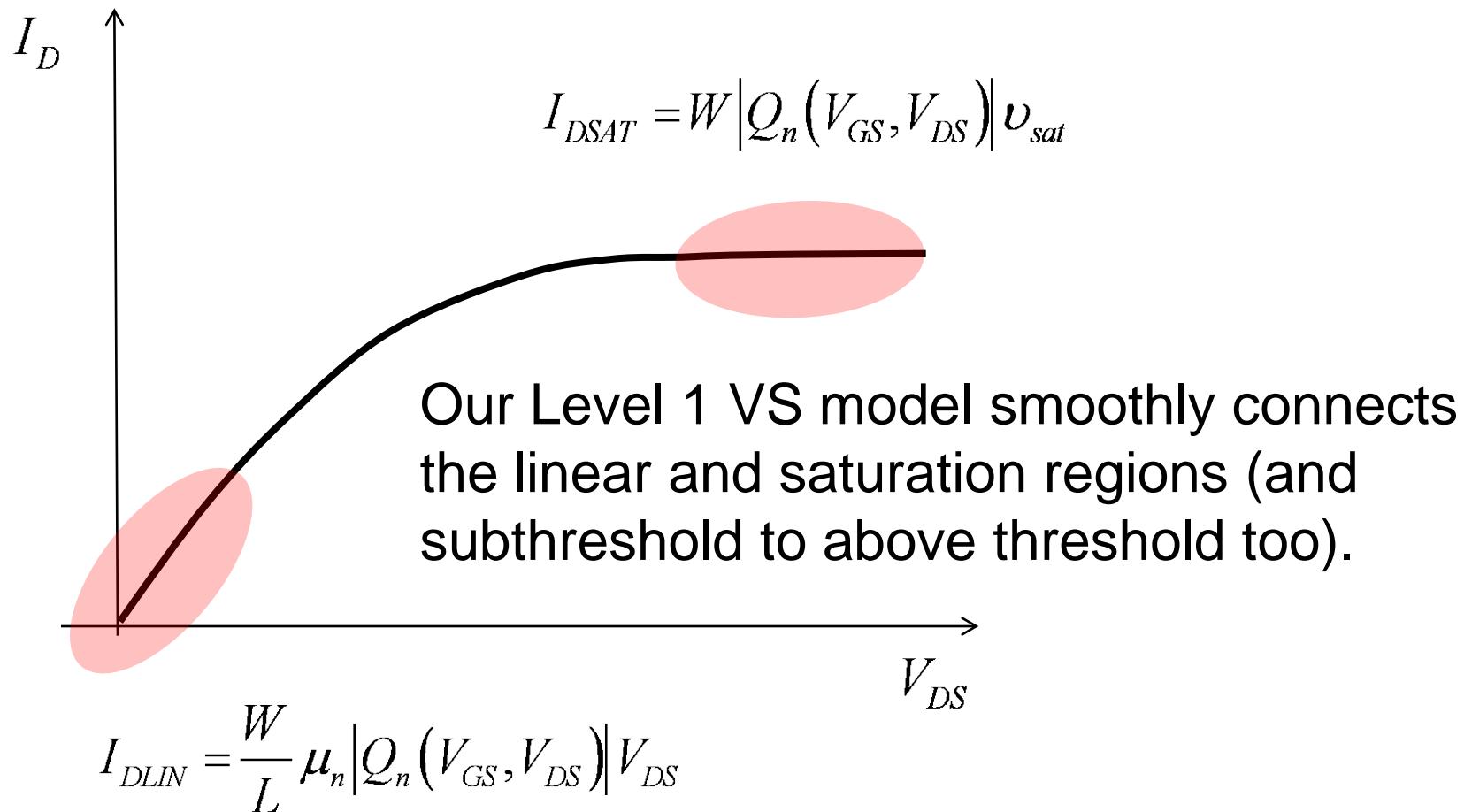
Lecture 4.6: The VS Model Revisited

Mark Lundstrom

lundstro@purdue.edu
Electrical and Computer Engineering
Purdue University
West Lafayette, Indiana USA

Lundstrom: 2018

Traditional (diffusive) model



Level 1 VS model

$$1) \quad I_D/W = |Q_n(V_{GS}, V_{DS})| \langle v_x(V_{DS}) \rangle$$

$$2) \quad Q_n(V_{GS}, V_{DS}) = -C_{inv} m(k_B T / q) \ln \left(1 + e^{q(V_{GS} - V_T + \alpha(k_B T_L / q) F_f) / m k_B T} \right)$$

$$V_T = V_{T0} - \delta V_{DS}$$

$$3) \quad \langle v_x(V_{DS}) \rangle = F_{SAT}(V_{DS}) v_{sat}$$

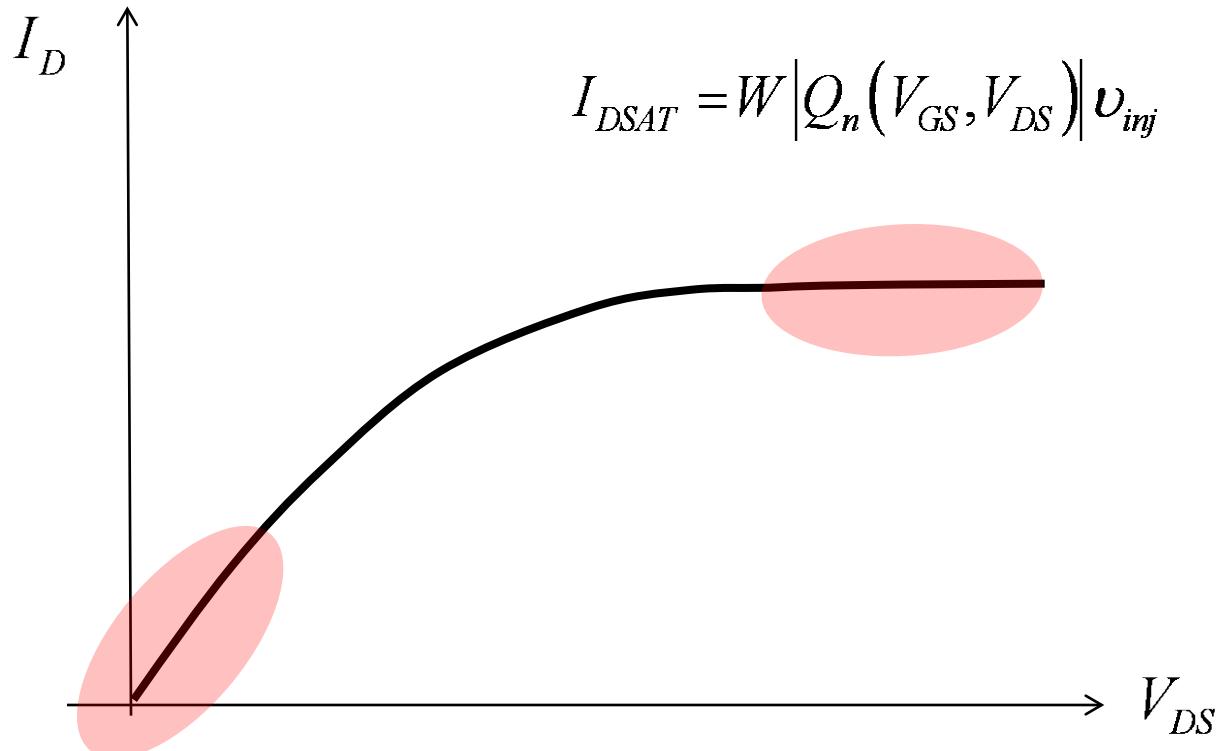
$$4) \quad F_{SAT}(V_{DS}) = \frac{V_{DS}/V_{DSAT}}{\left[1 + (V_{DS}/V_{DSAT})^\beta \right]^{1/\beta}}$$

$$5) \quad V_{DSAT} = \frac{v_{sat} L}{\mu_n}$$

Only 10 device-specific parameters in this model:

$$C_{inv}, V_{T0}, \delta, m, v_{sat}, \mu_n, \\ L, R_{SD} = R_S + R_D, \\ \alpha, \beta$$

Transmission model



$$I_{DSAT} = W |Q_n(V_{GS}, V_{DS})| v_{inj}$$

$$v_{inj} = \left(\frac{\mathcal{T}_{SAT}}{2 - \mathcal{T}_{SAT}} \right) v_T$$

$$\mathcal{T}_{SAT} = \frac{\lambda_0}{\lambda_0 + \ell}$$

$$\ell \ll L$$

$$I_{DLIN} = W \mu_{app} |Q_n(V_{GS}, V_{DS})| V_{DS} \quad \frac{1}{\mu_{app}} = \frac{1}{\mu_n} + \frac{1}{\mu_B}$$

Level 2 VS model

$$\mu_n \rightarrow \mu_{app} \quad \frac{1}{\mu_{app}} = \frac{1}{\mu_n} + \frac{1}{\mu_B}$$

$$v_{sat} \rightarrow v_{inj} \quad v_{inj} = \left(\frac{\tau_{SAT}}{2 - \tau_{SAT}} \right) v_T$$

Level 2 VS model

$$1) \quad I_D/W = |Q_n(V_{GS}, V_{DS})| \langle v_x(V_{DS}) \rangle$$

$$2) \quad Q_n(V_{GS}, V_{DS}) = -C_{inv} m(k_B T / q) \ln \left(1 + e^{q(V_{GS} - V_T + \alpha(k_B T_L / q) F_f) / m k_B T} \right)$$

$$V_T = V_{T0} - \delta V_{DS}$$

$$3) \quad \langle v_x(V_{DS}) \rangle = F_{SAT}(V_{DS}) v_{inj}$$

$$4) \quad F_{SAT}(V_{DS}) = \frac{V_{DS}/V_{DSAT}}{\left[1 + (V_{DS}/V_{DSAT})^\beta \right]^{1/\beta}}$$

$$5) \quad V_{DSAT} = \frac{v_{inj} L}{\mu_{app}}$$

Still only 10 device-specific parameters in this model:

$$C_{inv}, V_{T0}, \delta, m, v_{inj}, \mu_{app},$$

$$L, R_{SD} = R_S + R_D,$$

$$\alpha, \beta$$

Transport physics at the nanoscale

Let's take a quick, second look at how our VS model treats:

- 1) Transport in the linear region
- 2) Transport in the saturation region

Linear region: The apparent MFP

$$\frac{1}{\mu_{app}} = \frac{1}{\mu_n} + \frac{1}{\mu_B} \quad \mu_n = \frac{v_T \lambda_0}{(2k_B T / q)} \quad \mu_B = \frac{v_T L}{(2k_B T / q)} \quad \boxed{\mu_{app} = \frac{v_T \lambda_{app}}{(2k_B T / q)}}$$

“Mathiessen’s Rule”

$$\frac{1}{\lambda_{app}} = \frac{1}{\lambda_0} + \frac{1}{L}$$

The apparent MFP is the shorter of the scattering limited MFP in the bulk and the channel length.

Saturation region: Injection velocity

$$I_{DSAT} = W|Q_n|v_{inj} \quad v_{inj} = \left(\frac{\mathcal{T}_{SAT}}{2 - \mathcal{T}_{SAT}} \right) v_T \quad \mathcal{T}_{SAT} = \frac{\lambda_0}{\lambda_0 + \ell}$$

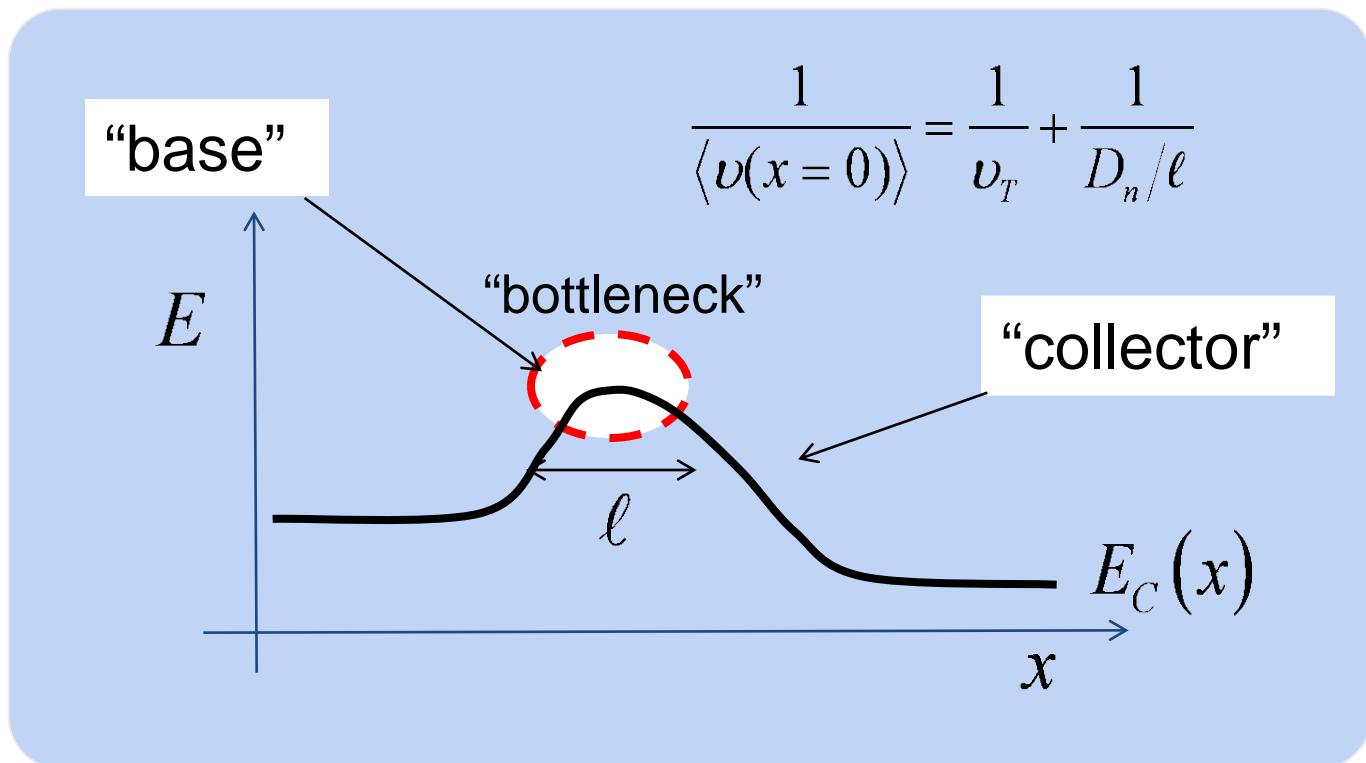
$$I_{DSAT} = W|Q_n| \left[\frac{1}{v_T} + \frac{1}{(D_n/\ell)} \right]^{-1}$$

$$D_n = \frac{v_T \lambda_0}{2}$$

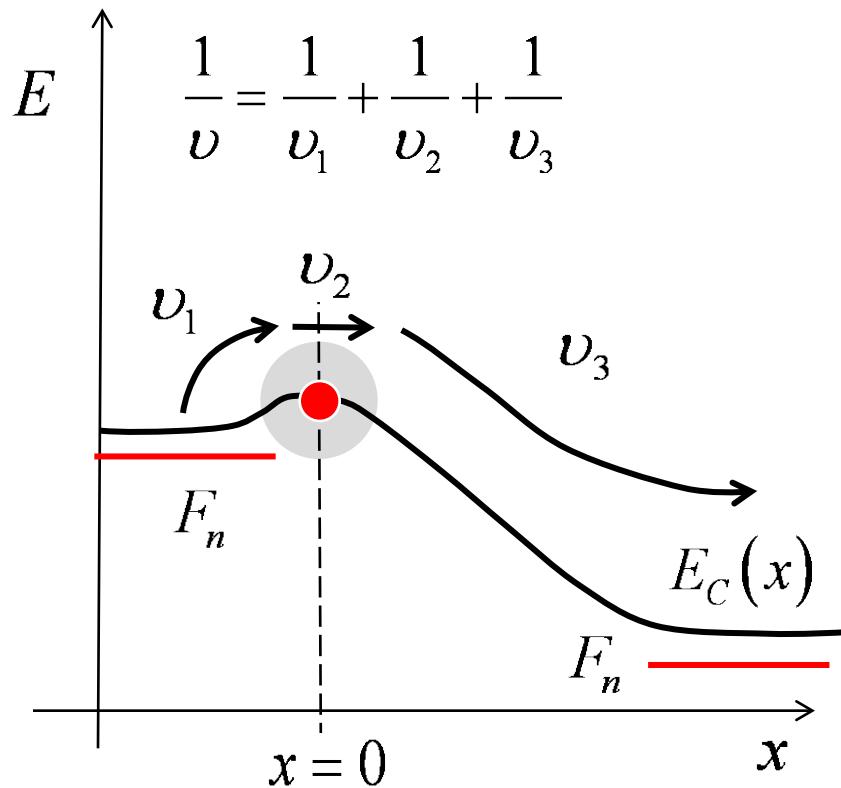
How do we interpret this result?

Saturation current in a nanoscale MOSFET

$$I_{DS} = WC_{inv} \left(V_{GS} - V_T \right) \langle v(x=0) \rangle$$



Injection velocity



$$\frac{1}{v} = \frac{1}{v_1} + \frac{1}{v_2} + \frac{1}{v_3}$$

$$\frac{1}{v} \approx \frac{1}{v_1} + \frac{1}{v_2}$$

$$\frac{1}{v_{inj}} = \frac{1}{v_T} + \frac{1}{D_n/\ell}$$

Summary

1674

IEEE TRANSACTIONS ON ELECTRON DEVICES, VOL. 56, NO. 8, AUGUST 2009

A Simple Semiempirical Short-Channel MOSFET Current–Voltage Model Continuous Across All Regions of Operation and Employing Only Physical Parameters

Ali Khakifirooz, *Member, IEEE*, Osama M. Nayfeh, *Member, IEEE*, and Dimitri Antoniadis, *Fellow, IEEE*

$$\frac{1}{\mu_n} \rightarrow \frac{1}{\mu_{app}}$$

“apparent mobility”

$$v_{sat} \rightarrow v_{inj}$$

“injection velocity”

Next lecture

How to use the VS model to characterize the electrical performance of nanotransistors.

Essentials of MOSFETs

Unit 4: Transmission Theory of the MOSFET

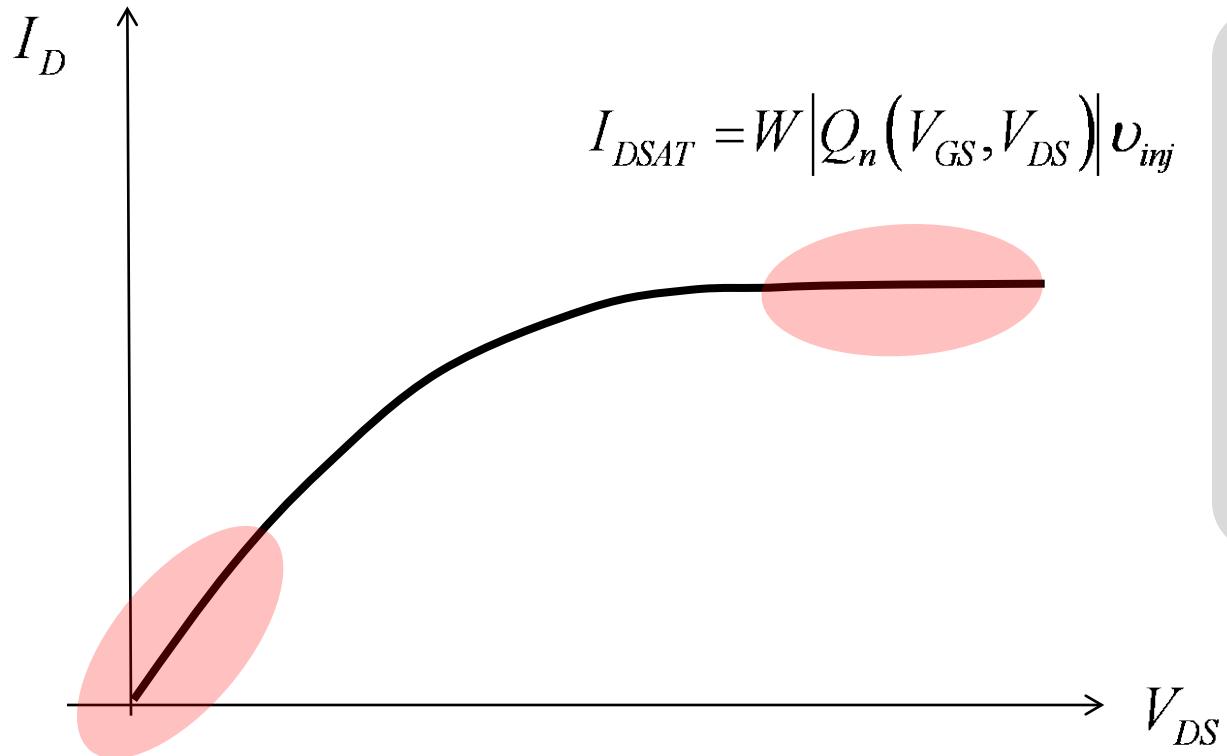
Lecture 4.7: Analysis of Experiments

Mark Lundstrom

lundstro@purdue.edu
Electrical and Computer Engineering
Purdue University
West Lafayette, Indiana USA

Lundstrom: 2018

Transmission model



$$v_{inj} = \left(\frac{\mathcal{T}_{SAT}}{2 - \mathcal{T}_{SAT}} \right) v_T$$

$$\mathcal{T}_{SAT} = \frac{\lambda_0}{\lambda_0 + \ell}$$

$$\ell \ll L$$

$$I_{DLIN} = \frac{W}{L} \mu_{app} |Q_n(V_{GS}, V_{DS})| V_{DS}$$

$$\mu_{app} = \frac{\mu_n \mu_B}{\mu_n + \mu_B} = \left(\frac{\mathcal{T}_{LIN} L v_T}{2 k_B T / q} \right) \quad \mathcal{T}_{LIN} = \frac{\lambda_0}{\lambda_0 + L}$$

MVS Model

$$1) \quad I_{DS} = W |Q_n(V_{GS}, V_{DS})| \langle v(V_{DS}) \rangle$$

$$2) \quad Q_n(V_{GS}, V_{DS}) = -C_{inv} m(k_B T / q) \ln \left(1 + e^{q(V_{GS} - V_T + \alpha(k_B T_L / q) F_f) / m k_B T} \right)$$

$$V_T = V_{T0} - \delta V_{DS}$$

$$3) \quad \langle v(V_{DS}) \rangle = F_{SAT}(V_{DS}) v_{inj}$$

$$4) \quad F_{SAT}(V_{DS}) = \frac{V_{DS}/V_{DSAT}}{\left[1 + (V_{DS}/V_{DSAT})^\beta \right]^{1/\beta}}$$

$$5) \quad V_{DSAT} = \frac{v_{inj} L}{\mu_{app}}$$

Only 10 device-specific parameters in this model:

$$C_{inv}, L,$$

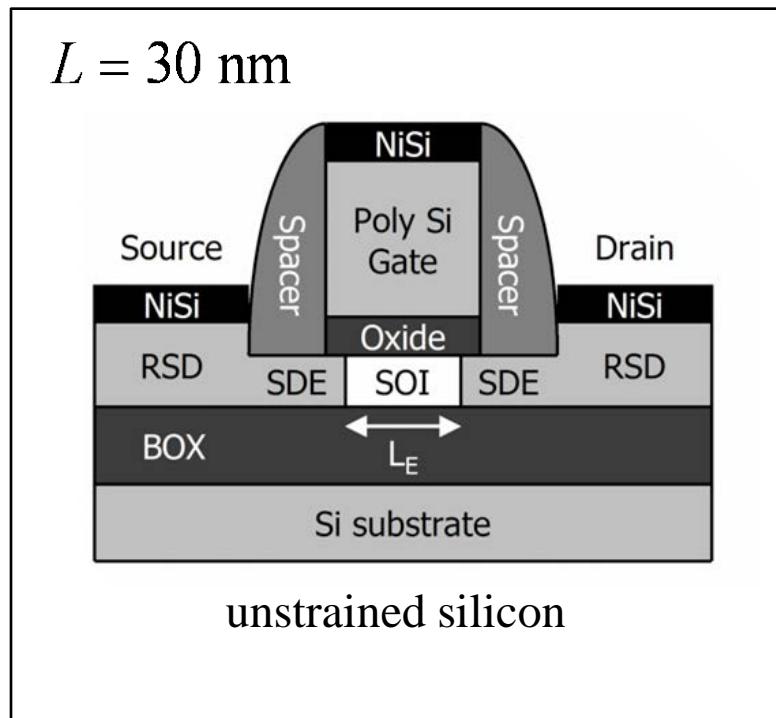
$$V_{T0}, \delta, m, v_{inj}, \mu_{app}, R_{SD} = R_S + R_D,$$

$$\alpha, \beta$$

Objective

To show how the VS model can be used to experimentally characterize Si and III-V FETS.

ETSOI N-MOSFETs



$$T_{Si} = 6.1 \pm 0.4 \text{ nm}$$

$$C_{inv} = 1.98 \mu\text{F}/\text{cm}^2$$

$$m^* = 0.22m_0$$

$$v_T = 1.14 \times 10^7 \text{ cm/s}$$

Long channel mobility:

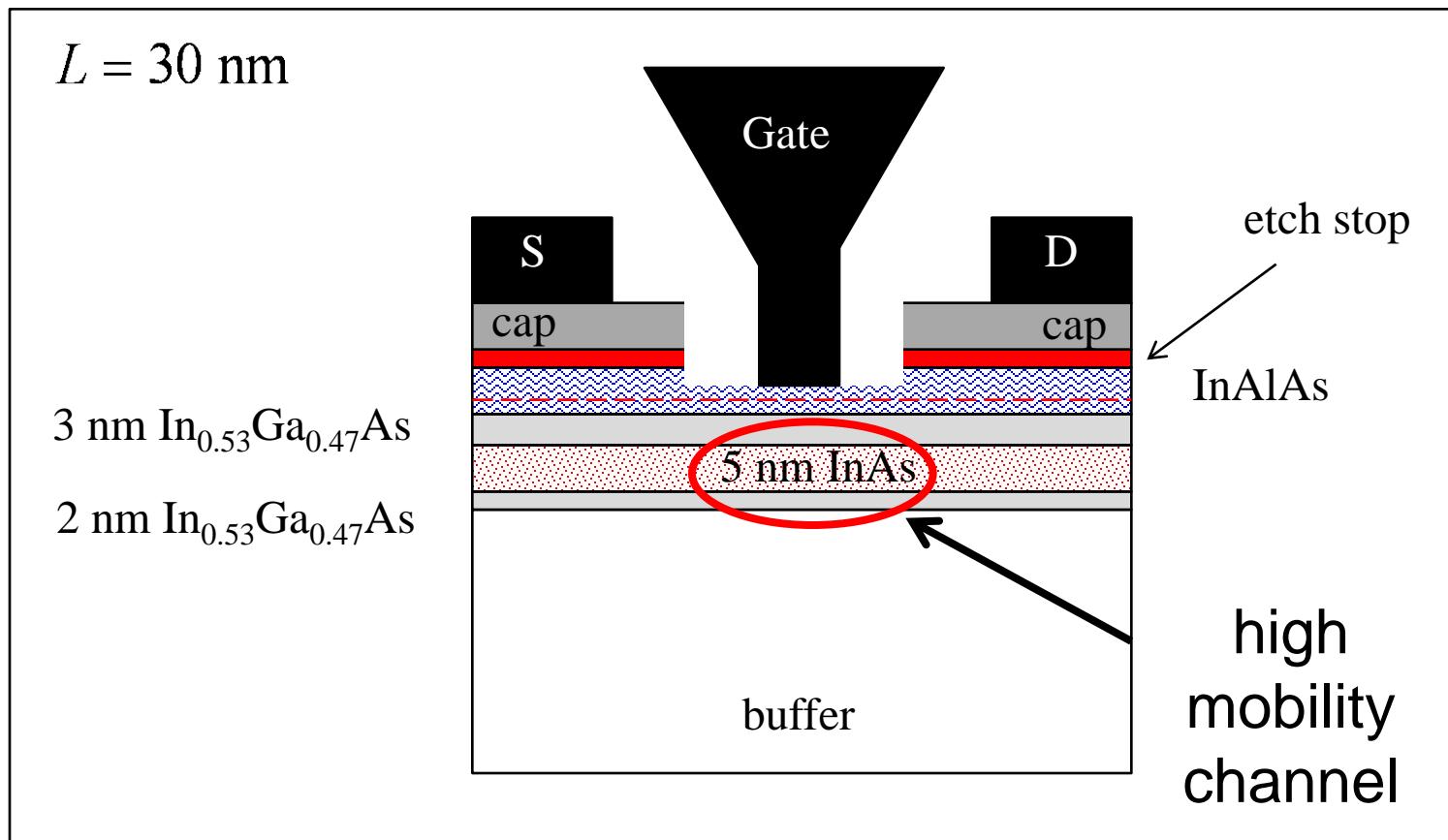
$$\mu_n = 350 \frac{\text{cm}^2}{\text{V-s}}$$

$$\text{MFP: } \lambda_0 = 15.8 \text{ nm}$$

For more details on this device see the following paper:

A. Majumdar and D.A. Antoniadis, "Analysis of Carrier Transport in Short-Channel MOSFETs," *IEEE Trans. Electron. Dev.*, **61**, pp. 351- 358, 2014.

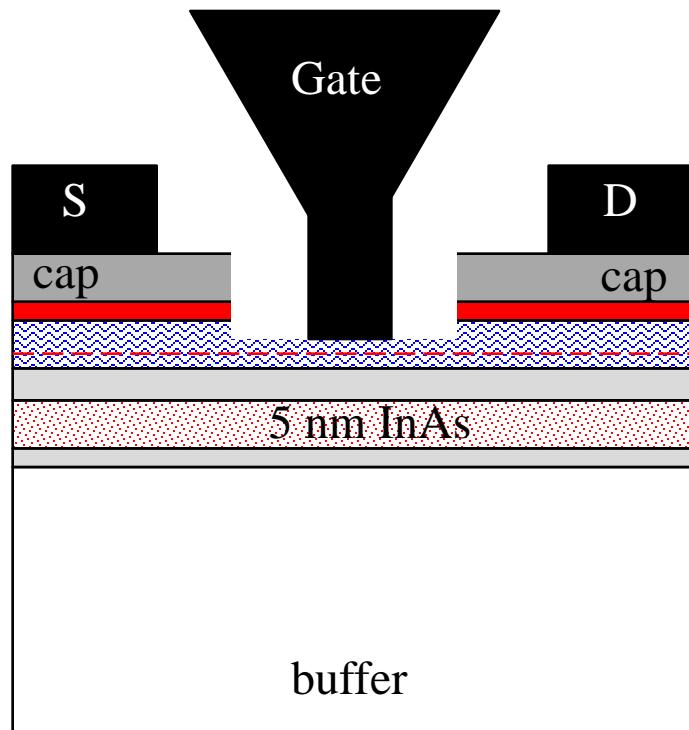
III-V HEMT



For information on this device, see: D.H. Kim and Jesus del Alamo, “30-nm InAs Pseudomorphic HEMTs on a InP Substrate With a Current-Gain Cutoff Frequency of 628 GHz,” *IEEE Electron Device Letters*, **29**, pp. 830-833, 2008.

high electron mobility transistor

III-V HEMT



$$T_{InAs} = 5 \text{ nm}$$

$$C_{inv} = 1.08 \mu\text{F}/\text{cm}^2$$

$$m^* = 0.022 m_0$$

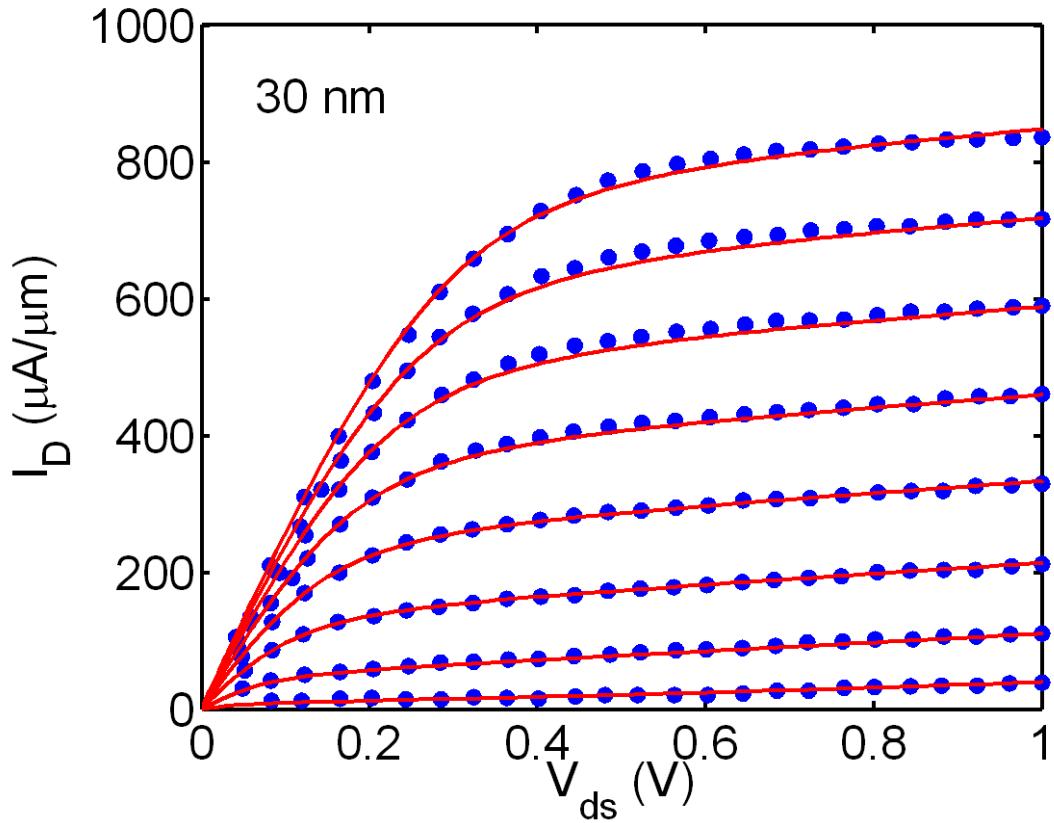
$$v_T = 3.62 \times 10^7 \text{ cm/s}$$

Long channel mobility:

$$\mu_n = 12,195 \frac{\text{cm}^2}{\text{V}\cdot\text{s}}$$

$$\text{MFP: } \lambda_0 = 171 \text{ nm}$$

MVS fit to experimental Si ETSOI data



Inputs to the model:

$$L, C_{inv}, \alpha, \beta$$

Outputs from the fitting process:

$$\mu_{app} = 220 \frac{\text{cm}^2}{\text{V}\cdot\text{s}}$$

$$v_{inj} = 0.82 \times 10^7 \text{ cm/s}$$

$$R_{S0} + R_{D0} = 130 \Omega\cdot\mu\text{m}$$

Linear region transmission

$$\mathcal{T}_{LIN} = \frac{\lambda_0}{\lambda_0 + L}$$

$$\frac{1}{\mu_{app}} = \frac{1}{\mu_n} + \frac{1}{\mu_B} \quad \mu_{app} = \frac{\mu_n \mu_B}{\mu_n + \mu_B}$$

$$\mu_n = \frac{v_T \lambda_0}{(2 k_B T / q)} \quad \mu_B = \frac{v_T L}{(2 k_B T / q)}$$

$$\frac{\mu_{app}}{\mu_B} = \frac{\lambda_0}{\lambda_0 + L} = \mathcal{T}_{LIN}$$

$$\mathcal{T}_{LIN} = \frac{\mu_{app}}{\mu_B}$$

$$\mu_{app} = 220 \frac{\text{cm}^2}{\text{V-s}}$$

$$\mu_B = 658 \frac{\text{cm}^2}{\text{V-s}}$$

$$\mathcal{T}_{LIN} = \frac{220}{658} = 0.33$$

$$\mathcal{T}_{LIN} = 0.33$$

Saturation region transmission

$$\mathcal{T}_{SAT} = \frac{\lambda_0}{\lambda_0 + \ell}$$

$$v_{inj} = \left(\frac{\mathcal{T}_{SAT}}{2 - \mathcal{T}_{SAT}} \right) v_T$$

$$\mathcal{T}_{SAT} = \left(\frac{2}{1 + v_T / v_{inj}} \right)$$

$$\mathcal{T}_{SAT} = \left(\frac{2}{1 + v_T / v_{inj}} \right)$$

$$v_T = \sqrt{\frac{2k_B T}{\pi m^*}} = 1.14 \times 10^7 \text{ cm/s}$$

$$v_{inj} = 0.82 \times 10^7 \text{ cm/s}$$

$$\mathcal{T}_{SAT} = \frac{2}{1 + 1.14 / 0.82} = 0.84$$

$$\mathcal{T}_{SAT} = 0.84$$

Linear vs. saturation region transmission

$$\mathcal{T}_{LIN} = \frac{\mu_{app}}{\mu_B}$$

$$\mathcal{T}_{SAT} = \left(\frac{2}{1 + v_T/v_{inj}} \right)$$

$$\mathcal{T}_{SAT} = 0.84 > \mathcal{T}_{LIN} = 0.33$$

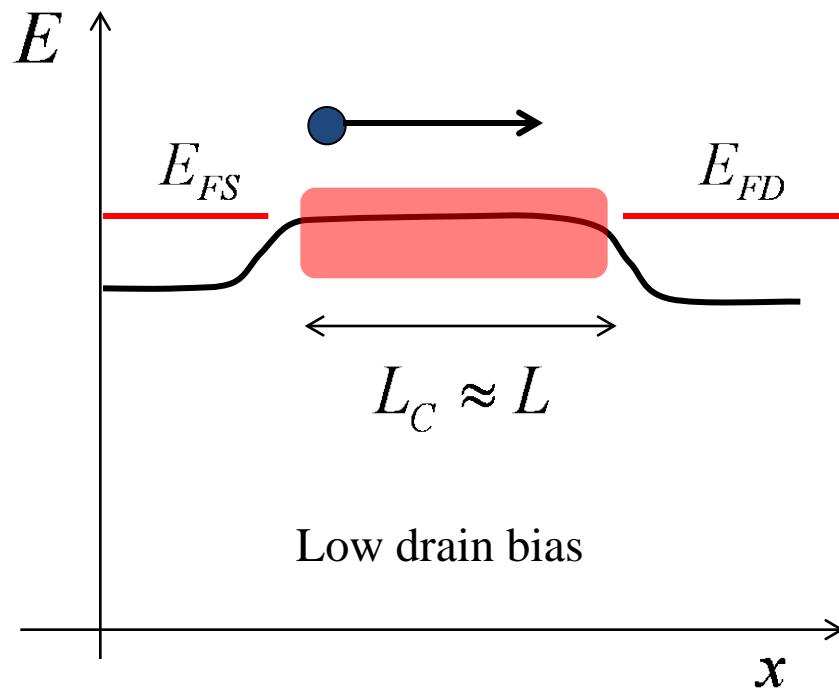


Linear vs. saturation region transmission

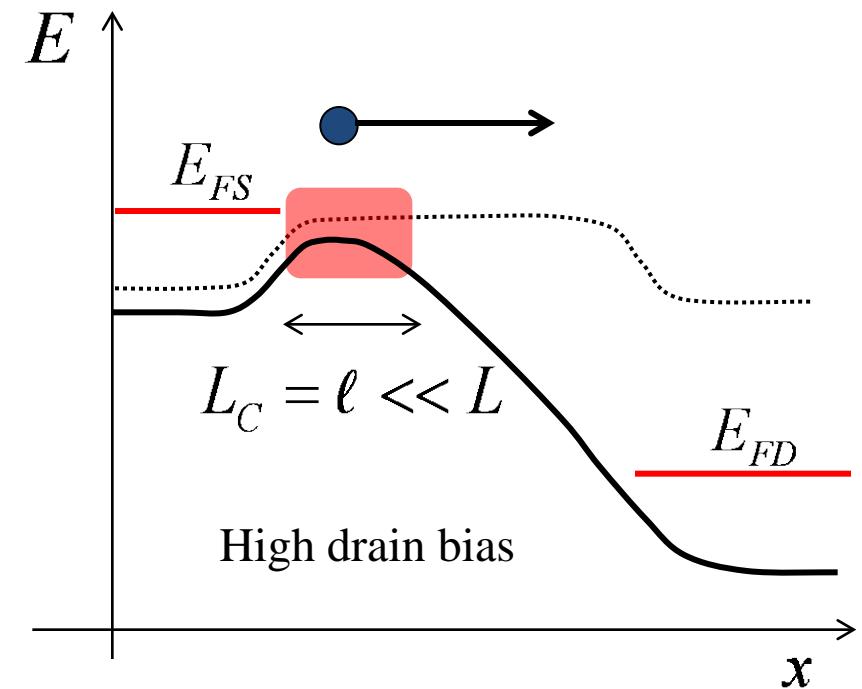
$$\mathcal{T}_{LIN} = \frac{\lambda_0}{\lambda_0 + L}$$

$$\mathcal{T}(V_{DS}) = \frac{\lambda_0}{\lambda_0 + L_C(V_{DS})}$$

$$\mathcal{T}_{SAT} = \frac{\lambda_0}{\lambda_0 + \ell}$$

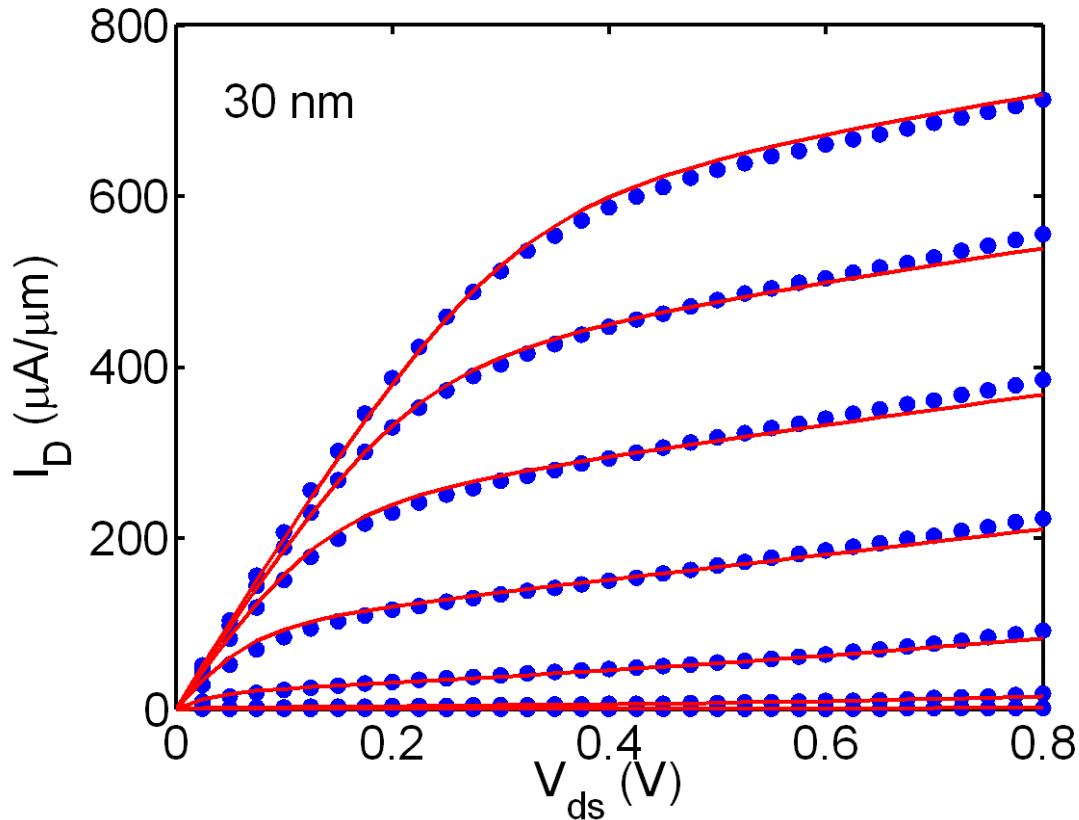


Low drain bias



High drain bias

MVS Fits to experimental III-V HEMT data



$$\mu_{app} = 1800 \frac{\text{cm}^2}{\text{V}\cdot\text{s}}$$

$$\left(\mu_n = 12,500 \frac{\text{cm}^2}{\text{V}\cdot\text{s}} \right)$$

$$v_{inj} = 3.50 \times 10^7 \text{ cm/s}$$

$$R_{S0} + R_{D0} = 400 \Omega\cdot\mu\text{m}$$

MVS fits to experimental III-V HEMT data

$$\mathcal{T}_{LIN} = \frac{\mu_{app}}{\mu_B}$$

$$\mu_B = v_T L / (2 k_B T / q) = 2088 \frac{\text{cm}^2}{\text{V}\cdot\text{s}}$$

$$\mathcal{T}_{LIN} = \frac{1800}{2088} = 0.86$$

$$\mathcal{T}_{SAT} = \left(\frac{2}{1 + v_T / v_{inj}} \right)$$

$$\mathcal{T}_{SAT} = \left(\frac{2}{1 + 3.62 / 3.50} \right) = 0.98$$

$$\mu_{app} = 1800 \frac{\text{cm}^2}{\text{V}\cdot\text{s}}$$

$$v_{inj} = 3.50 \times 10^7 \text{ cm/s}$$

$$v_T = 3.62 \times 10^7 \text{ cm/s}$$

$$R_S = R_D = 200 \Omega\cdot\mu\text{m}$$

Si vs. III-V $L = 30$ nm N-FET comparison

Si ETSOI

$$\mathcal{T}_{LIN} = 0.33$$

$$\mathcal{T}_{SAT} = 0.84$$

$$v_{inj} = 0.82 \times 10^7 \text{ cm/s}$$

Long channel mobility:

$$\mu_n = 350 \frac{\text{cm}^2}{\text{V-s}}$$

Apparent mobility:

$$\mu_{app} = 220 \frac{\text{cm}^2}{\text{V-s}}$$

III-V HEMT

$$\mathcal{T}_{LIN} = 0.86$$

$$\mathcal{T}_{SAT} = 0.98$$

$$v_{inj} = 3.50 \times 10^7 \text{ cm/s}$$

Long channel mobility:

$$\mu_n = 12,195 \frac{\text{cm}^2}{\text{V-s}}$$

Apparent mobility:

$$\mu_{app} = 1800 \frac{\text{cm}^2}{\text{V-s}}$$

Linear region analysis (i)

$$\frac{1}{\mu_{app}} = \frac{1}{\mu_n} + \frac{1}{\mu_B}$$

$$\mu_n = \frac{v_T \lambda_0}{2 k_B T / q}$$

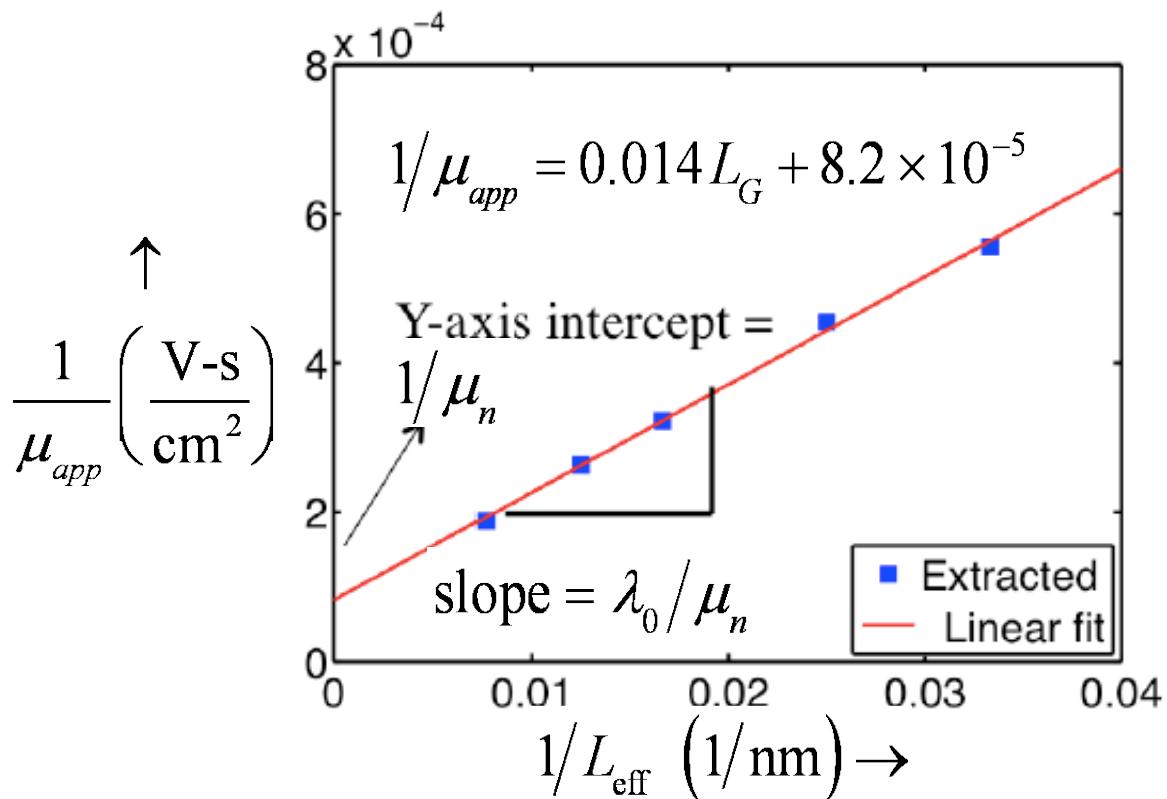
$$\mu_B = \frac{v_T L}{2 k_B T / q}$$

$$\frac{1}{\mu_{app}} = \frac{1}{\mu_n} + \left(\frac{\lambda_0}{\mu_n} \right) \frac{1}{L}$$

From the y-intercept, we find the mobility and from the slope, the MFP.

S. Rakheja, M. Lundstrom, and D. Antoniadis, “A physics-based compact model for FETs from diffusive to ballistic carrier transport regimes,” in IEDM Tech. Dig., 2014.

Linear region analysis: III-V HEMT



$$\mu_n = 12,195 \frac{\text{cm}^2}{\text{V-s}}$$

$$\lambda_0 = 171 \text{ nm}$$

Linear region analysis (ii)

$$\frac{1}{\mu_{app}} = \frac{1}{\mu_n} + \frac{1}{\mu_B}$$

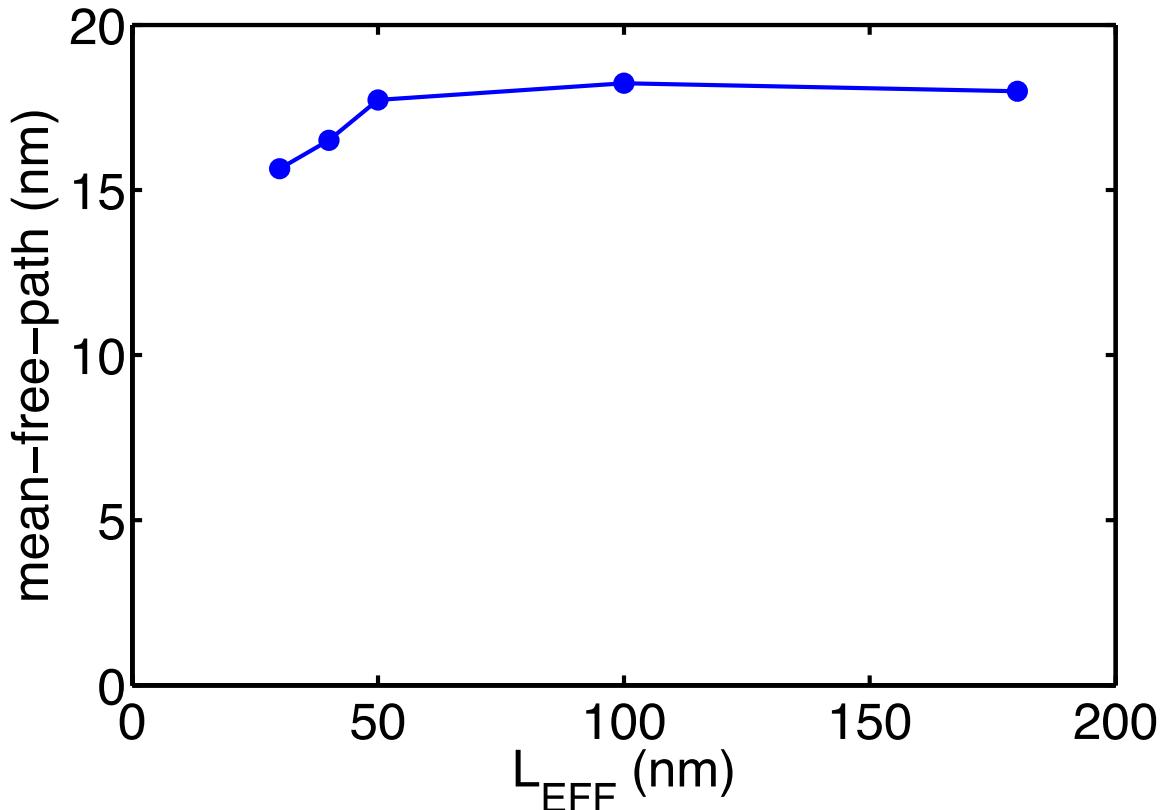
$$\mu_n = \frac{v_T \lambda_0}{2 k_B T / q}$$

$$\mu_B = \frac{v_T L}{2 k_B T / q}$$

$$\frac{1}{\lambda_0(L)} = \frac{v_T}{2(k_B T / q)} \frac{1}{\mu_{app}} - \frac{1}{L}$$

Must know the unidirectional thermal velocity.

Linear region analysis: Si ETSOI



- 1) Fundamental
(long-range
Coulomb
oscillations)?
- 2) Processing
related?
- 3) Backscattering
from the drain?

(Analysis by Xingshu Sun, Purdue University, 2015.)

Recent work on this topic

Dimitri Antoniadis, “On Apparent Mobility in Si nMOSFETs From Diffusive to Ballistic Regime,” *IEEE Trans. Electron Devices*, **63**, pp. 2650-2656, 2016.

Saturation region analysis

$$\mathcal{T}_{SAT} = \frac{\lambda_0}{\lambda_0 + \ell}$$

$$v_{inj} = \left(\frac{\mathcal{T}_{SAT}}{2 - \mathcal{T}_{SAT}} \right) v_T$$

$$\mathcal{T}_{SAT} = \left(\frac{2}{1 + v_T/v_{inj}} \right)$$

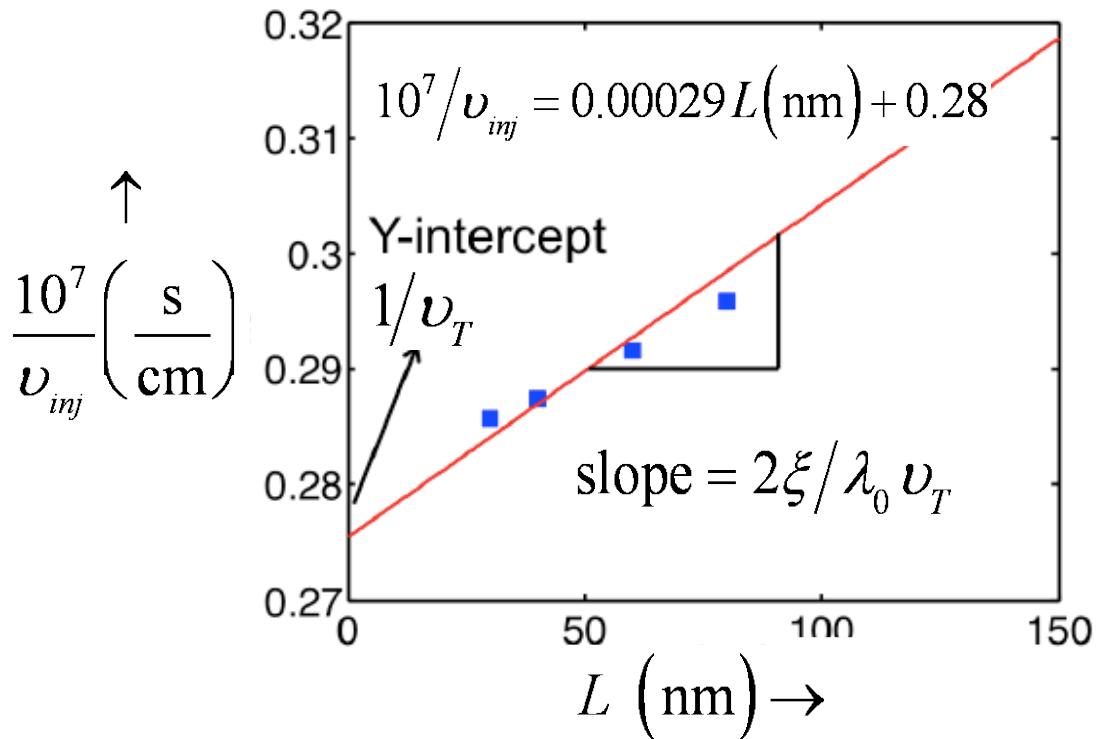
$$v_{inj} = \left(\frac{v_T \lambda_0}{\lambda_0 + 2\ell} \right)$$

Assume: $\ell = \xi L$

$$\frac{1}{v_{inj}} = \frac{1}{v_T} + \left(\frac{2\xi}{\lambda_0 v_T} \right) L$$

From the y-intercept, we find the thermal velocity and from the slope, we find the critical length.

Saturation region analysis (III-V HEMT)



$$v_T = 3.57 \times 10^7 \text{ cm/s}$$
$$\xi = 0.09$$

Summary

The MVS / transmission model provides us with a way to extract key physical parameters from **good** nanotransistors.

For more discussion about analyzing experimental data, see:

A. Majumdar and D.A. Antoniadis, “Analysis of Carrier Transport in Short-Channel MOSFETs,” *IEEE Trans. Electron. Dev.*, **61**, pp. 351- 358, 2014.

S. Rakheja, M. Lundstrom, and D. Antoniadis, “A physics-based compact model for FETs from diffusive to ballistic carrier transport regimes,” in IEDM Tech. Dig., 2014.

Next lecture

A summary of the key “take aways” for Unit 4.

Essentials of MOSFETs

Unit 4: Transmission Theory of the MOSFET

Lecture 4.8: Unit 4 Recap

Mark Lundstrom

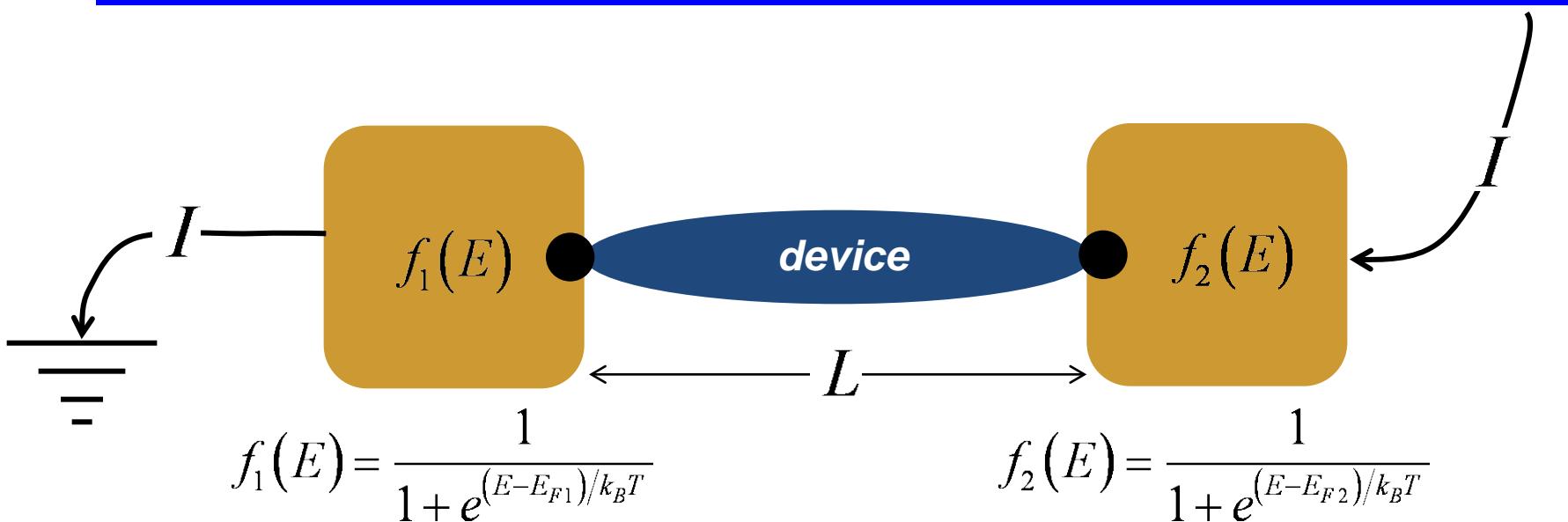
lundstro@purdue.edu
Electrical and Computer Engineering
Purdue University
West Lafayette, Indiana USA

Lundstrom: 2018

Unit 4 topics

- 4.1 Landauer Approach
- 4.2 Landauer at low and High Bias
- 4.3 The Ballistic MOSFET
- 4.4 Velocity at the Virtual Source
- 4.5 Transmission Theory of the MOSFET
- 4.6 The VS Model Revisited
- 4.7 Analysis of Experiments

Landauer Approach



$$I = \frac{2q}{h} \int \mathcal{T}(E) M(E) (f_1 - f_2) dE$$

Can be used to describe the current in small and large devices and in short to long devices.

Landauer at low and high bias

$$I = \frac{2q}{h} \int \mathcal{T}(E) M(E) (f_1 - f_2) dE$$

1) Linear region:

$$I_{DLIN} = \left[\frac{2q^2}{h} \int \mathcal{T}(E) M(E) \left(-\frac{\partial f_0}{\partial E} \right) dE \right] V_{DS}$$

2) Saturation region: $I_{DSAT} = \frac{2q}{h} \int \mathcal{T}(E) M(E) f_1(E) dE$

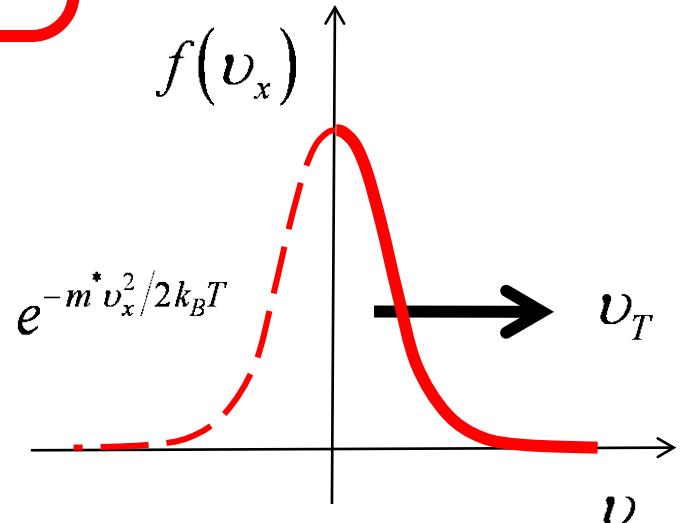
Transmission, MFP, and diffusion coefficient

$$\mathcal{T}_0 = \frac{\lambda_0}{\lambda_0 + L}$$

$$D_n = \frac{k_B T}{q} \mu_n$$

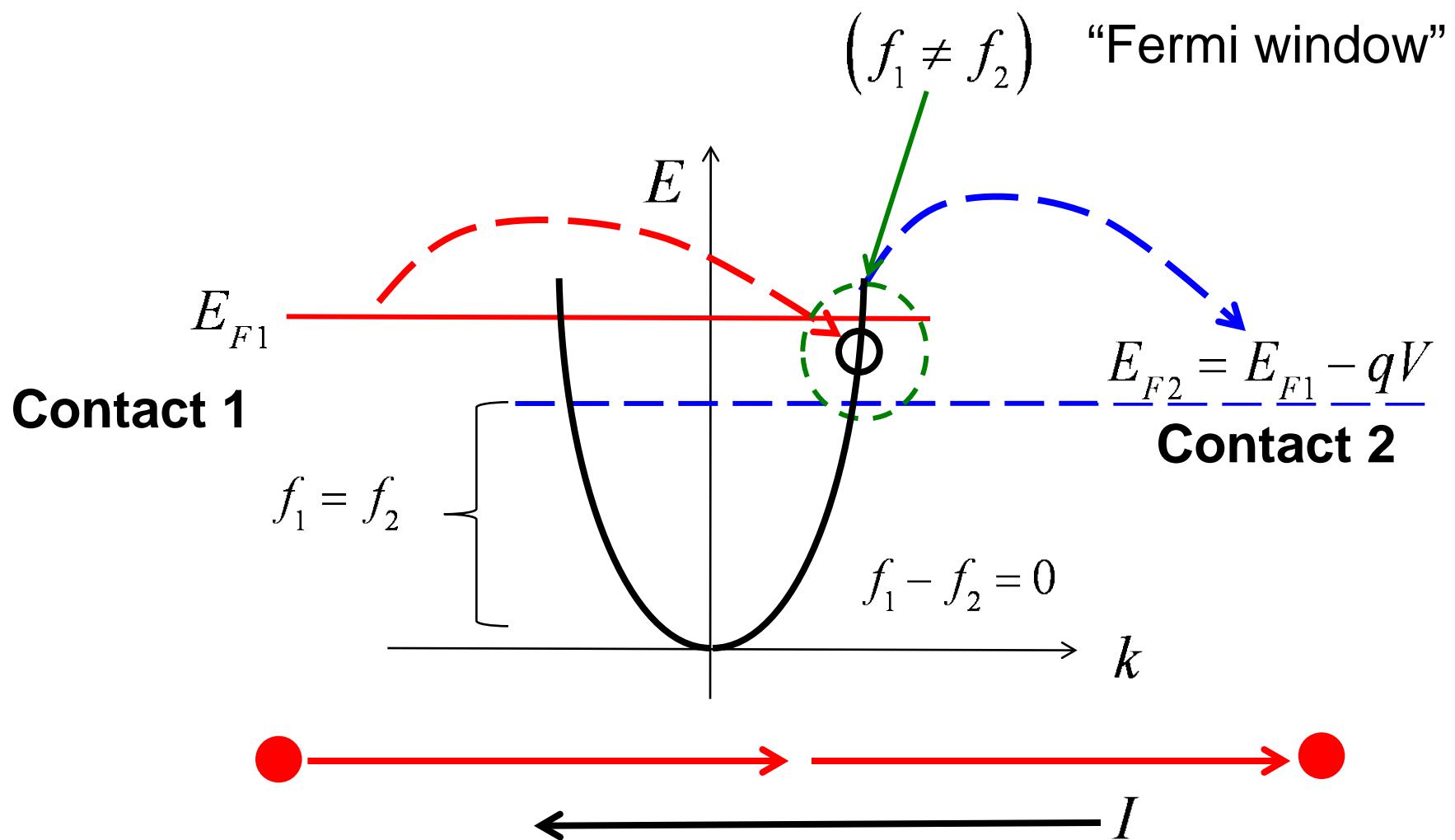
Einstein relation

$$D_n = \frac{v_T \lambda_0}{2} \text{ cm}^2/\text{s}$$

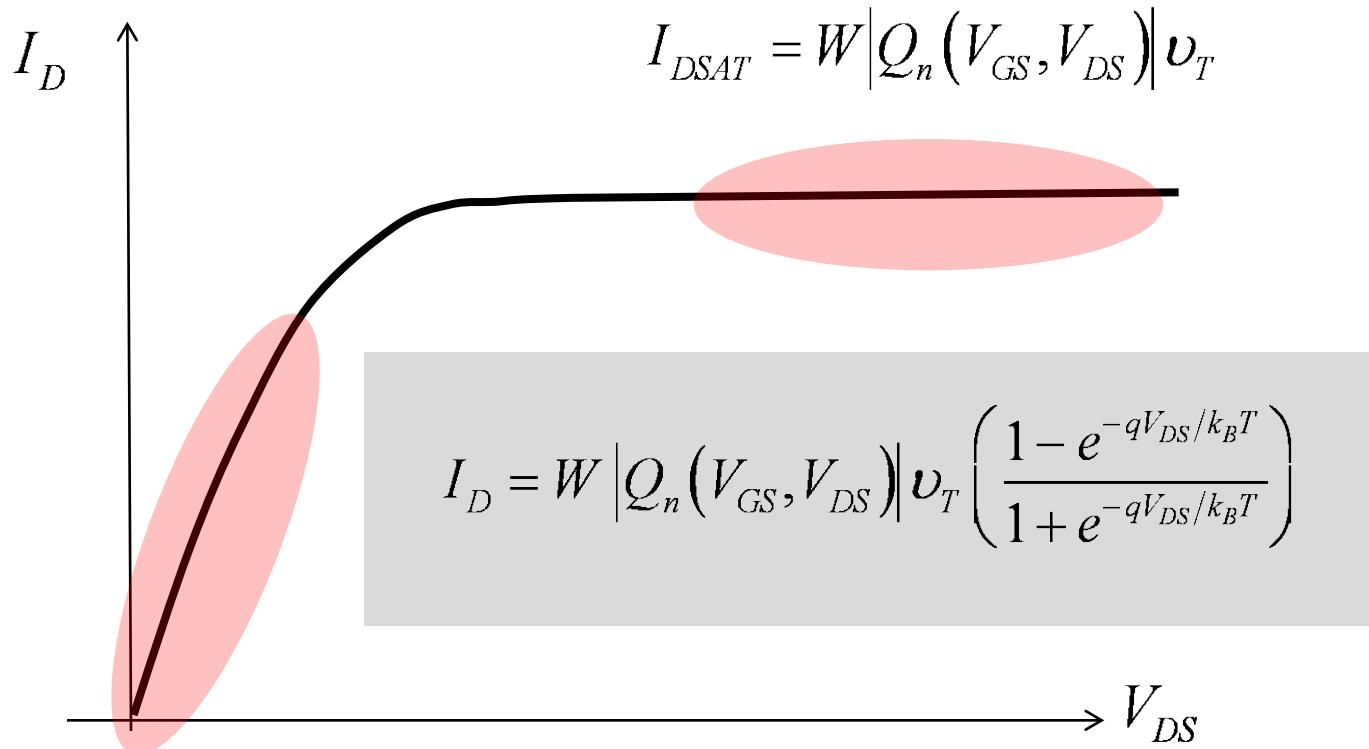


$$\langle\langle v_x^+ \rangle\rangle = v_T = \sqrt{\frac{2k_B T}{\pi m^*}}$$

Femi window and current flow



The Ballistic MOSFET



$$I_{DLIN} = \frac{W}{L} \mu_B |Q_n(V_{GS}, V_{DS})| V_{DS} \quad \mu_B = \frac{v_T L}{2 k_B T / q}$$

Ballistic vs. diffusive mobility

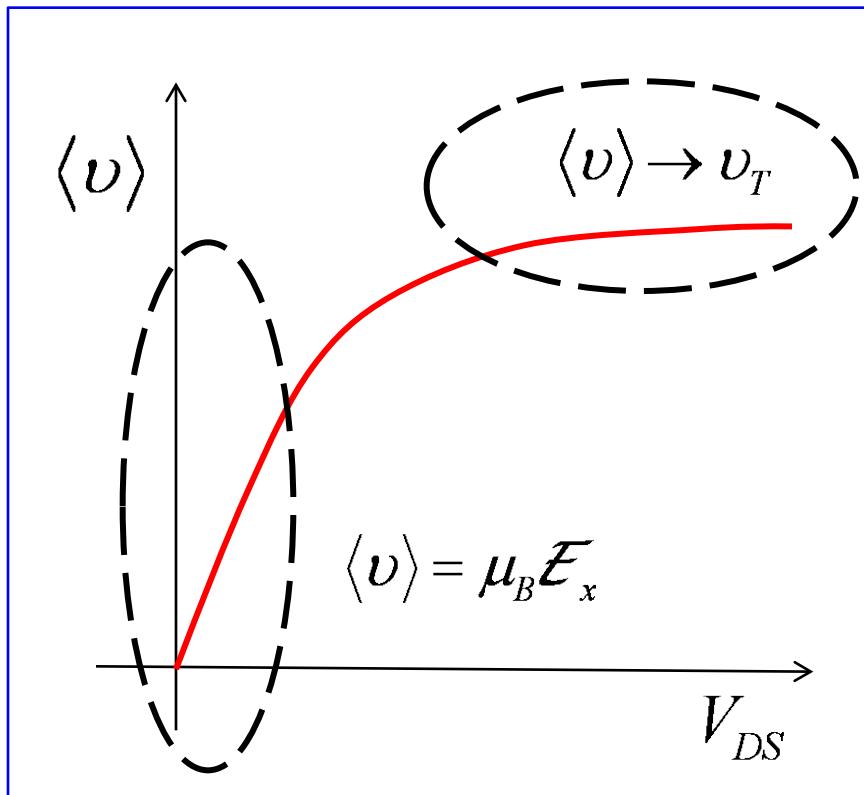
$$\mu_B \equiv \frac{v_T L}{2(k_B T / q)}$$

“ballistic mobility”

$$\mu_n = \frac{v_T \lambda_0}{2(k_B T / q)}$$

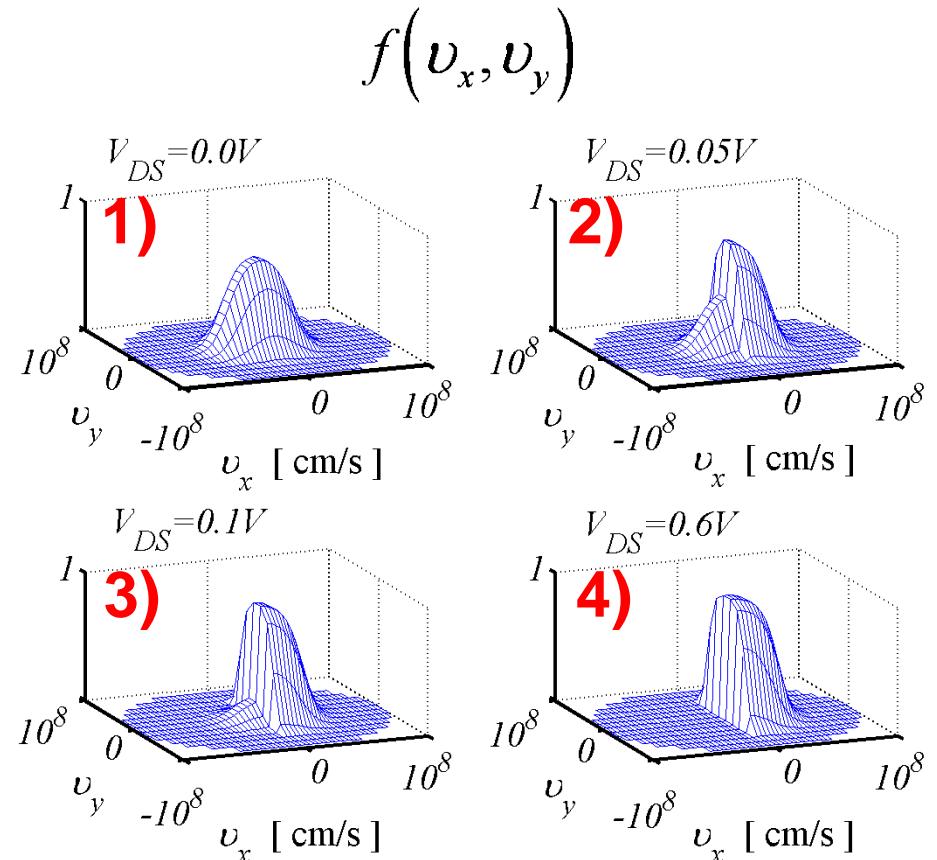
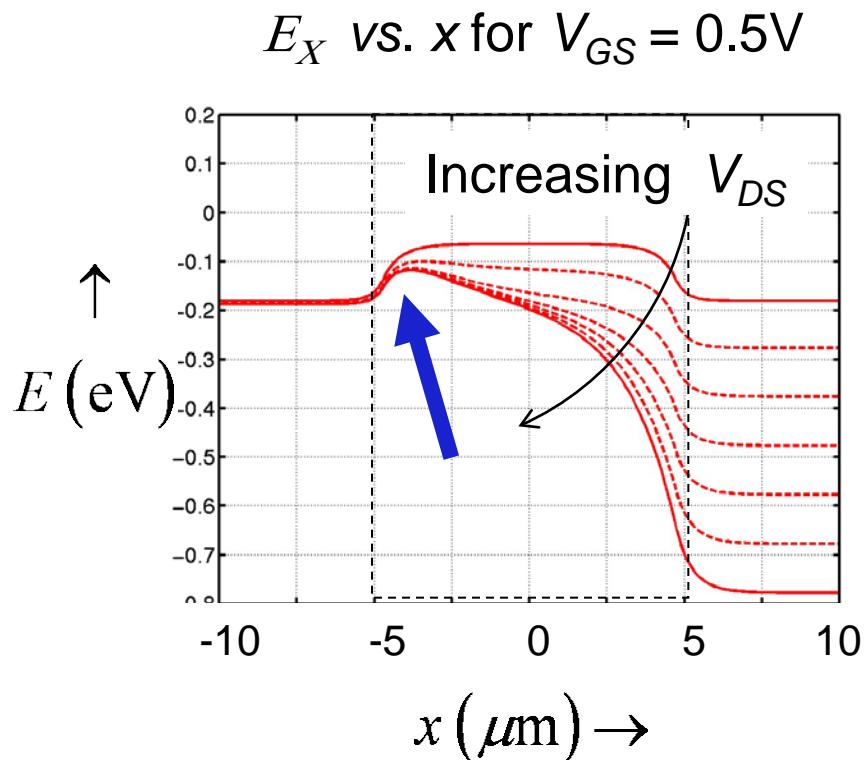
“diffusive mobility”

Ballistic velocity vs. V_{DS} at the VS



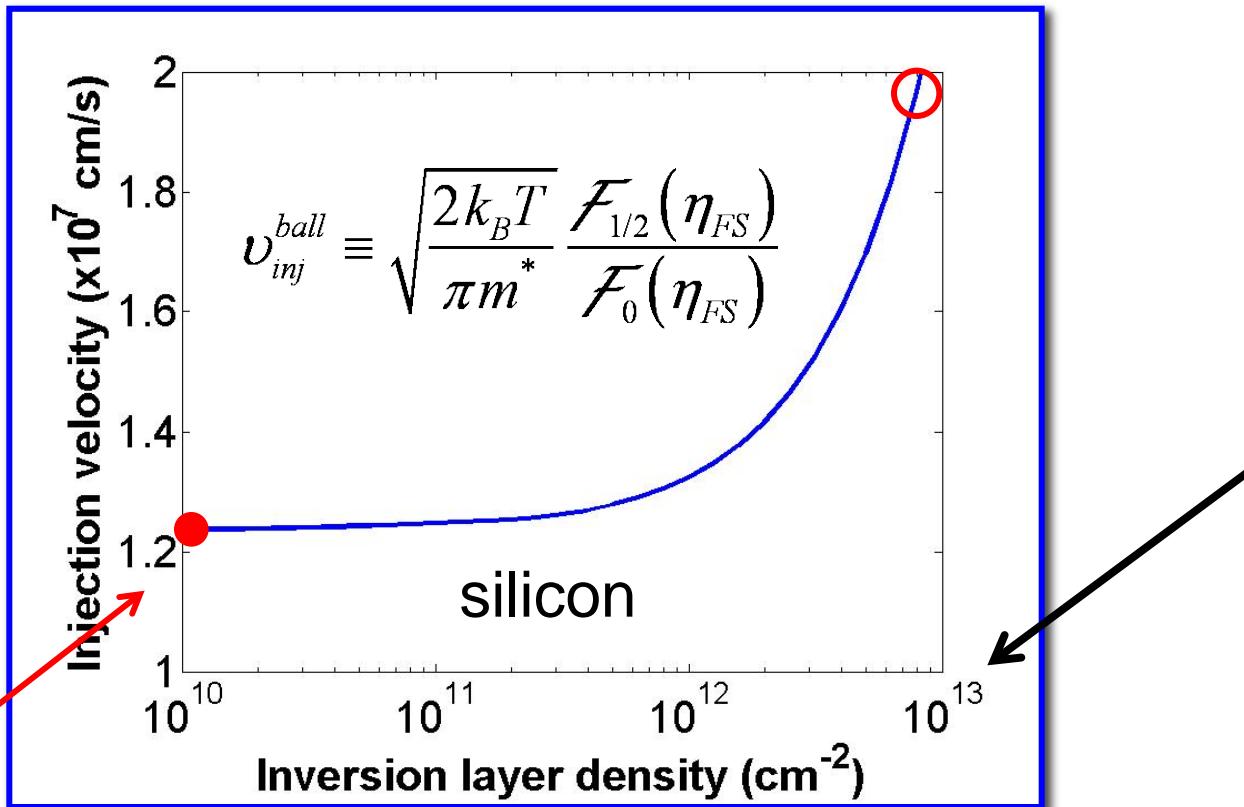
Velocity
saturation
with no
scattering!

Physics of velocity saturation in ballistic MOSFETs



(Numerical simulations of an $L = 10$ nm double gate Si MOSFET from J.-H. Rheu and M.S. Lundstrom, *Solid-State Electron.*, **46**, 1899, 2002)

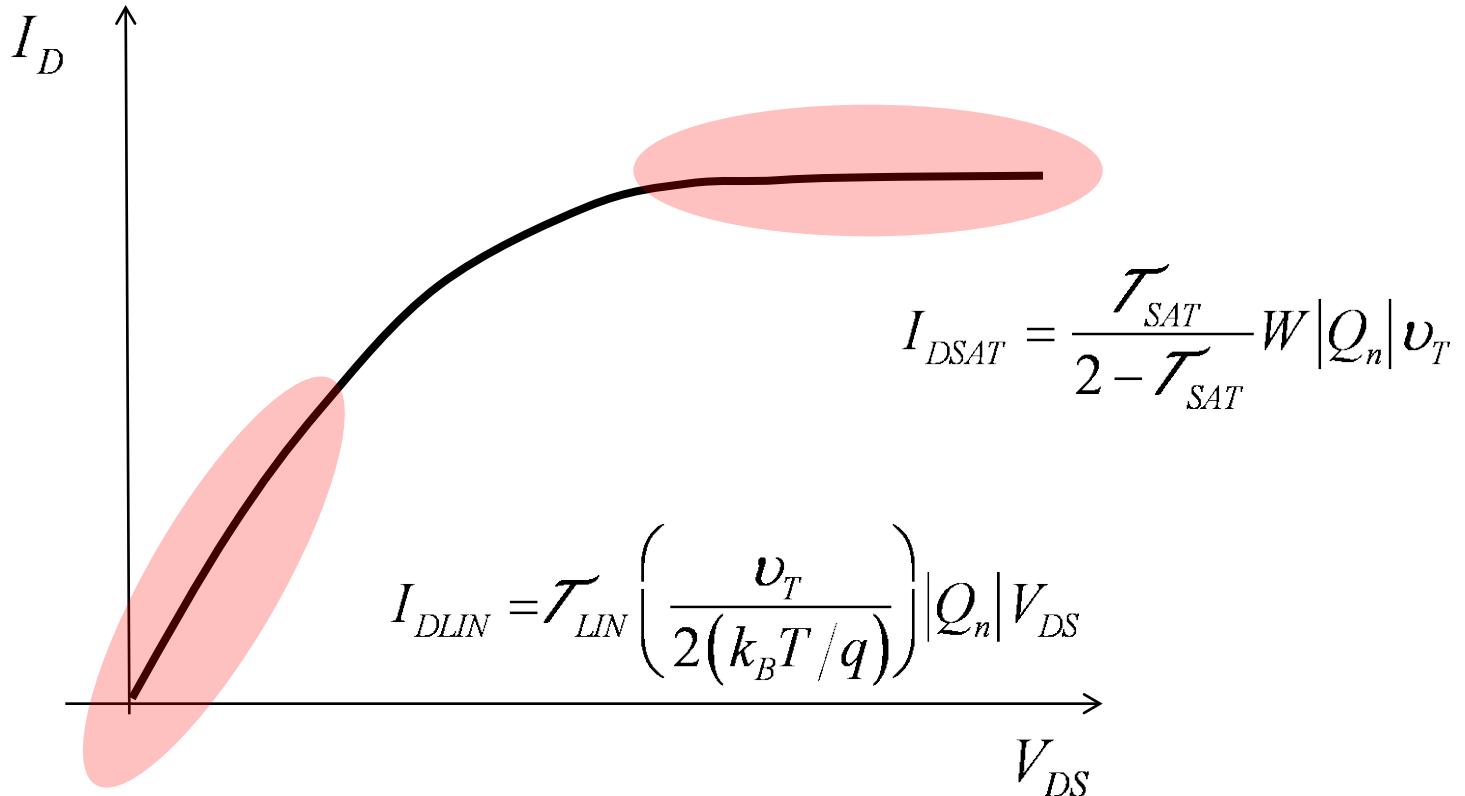
Injection velocity vs. gate voltage



$$v_T = \sqrt{\frac{2k_B T}{\pi m^*}}$$

non-degenerate

Transmission theory of the MOSFET



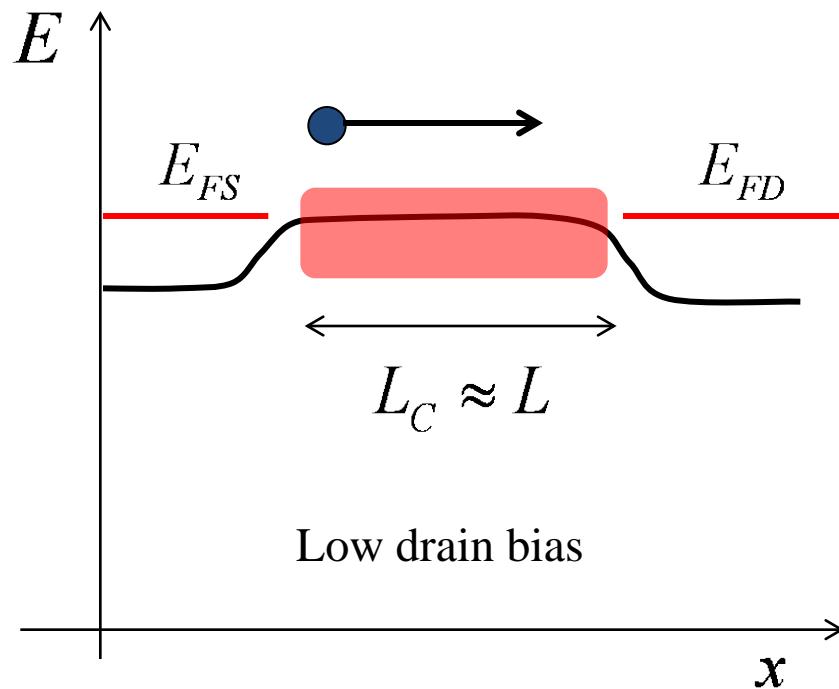
$$\mathcal{T}_{LIN} < \mathcal{T}_{SAT}$$

Linear vs. saturation region transmission

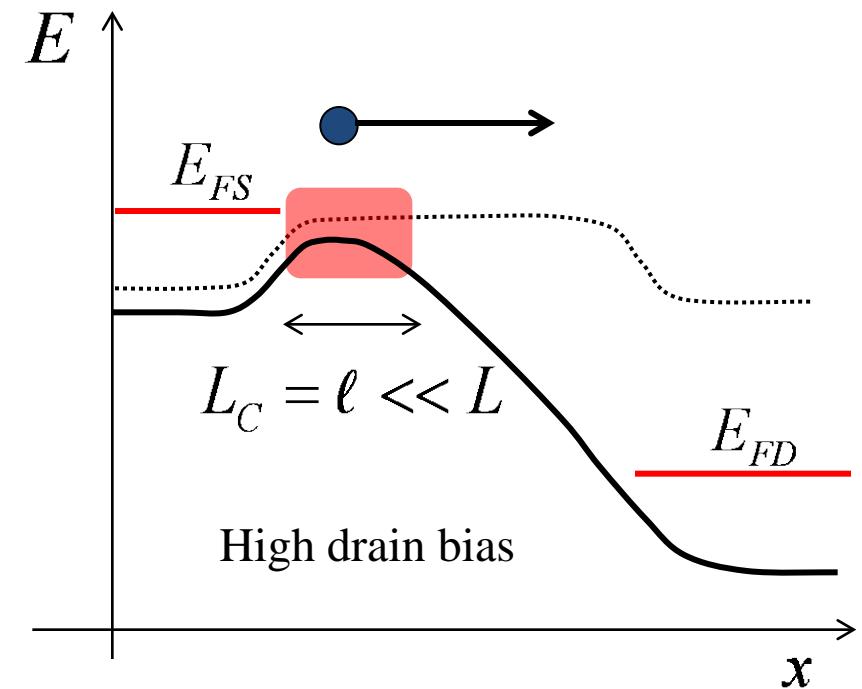
$$\mathcal{T}_{LIN} = \frac{\lambda_0}{\lambda_0 + L}$$

$$\mathcal{T}(V_{DS}) = \frac{\lambda_0}{\lambda_0 + L_C(V_{DS})}$$

$$\mathcal{T}_{SAT} = \frac{\lambda_0}{\lambda_0 + \ell}$$

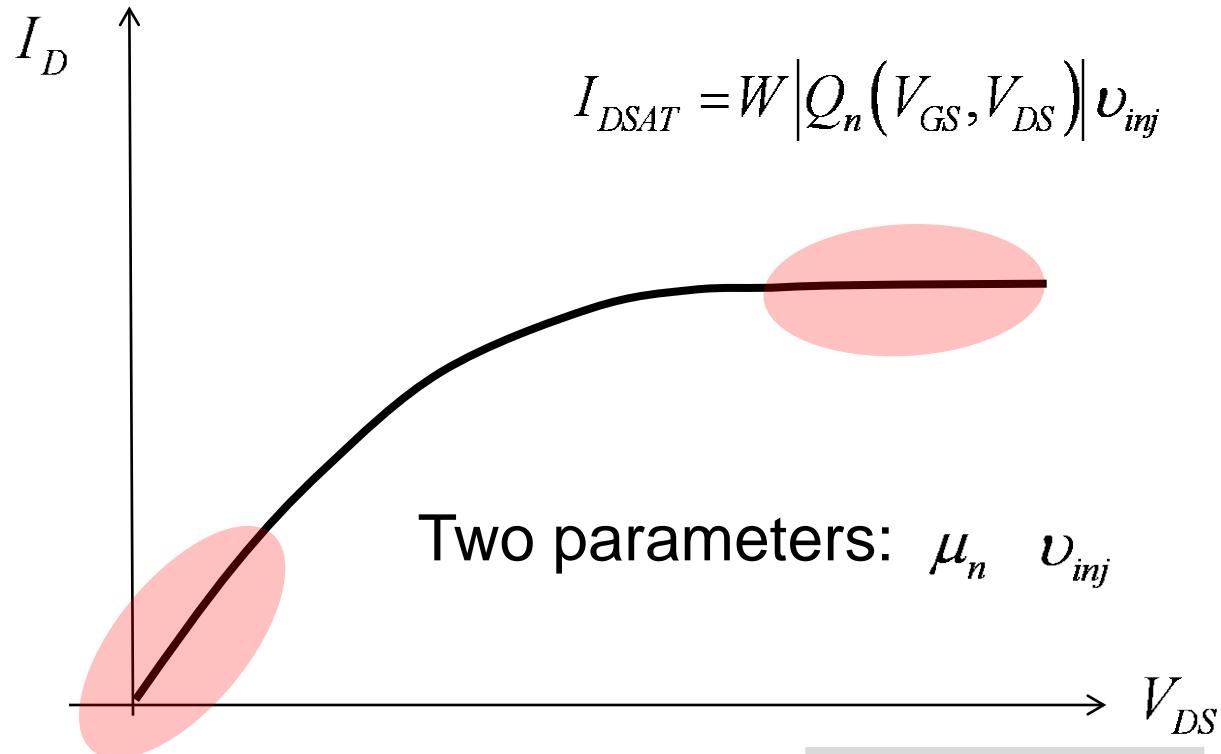


Low drain bias



High drain bias

Alternative formulation



$$I_{DLIN} = W \mu_{app} |Q_n(V_{GS}, V_{DS})| V_{DS}$$

$$\frac{1}{\mu_{app}} = \frac{1}{\mu_n} + \frac{1}{\mu_B}$$

$$v_{inj} = \left(\frac{\mathcal{T}_{SAT}}{2 - \mathcal{T}_{SAT}} \right) v_T$$

$$\mathcal{T}_{SAT} = \frac{\lambda_0}{\lambda_0 + \ell}$$

$$\ell \ll L$$

Level 2 VS model

$$1) \quad I_D/W = |Q_n(V_{GS}, V_{DS})| \langle v_x(V_{DS}) \rangle$$

$$2) \quad Q_n(V_{GS}, V_{DS}) = -C_{inv} m(k_B T / q) \ln \left(1 + e^{q(V_{GS} - V_T + \alpha(k_B T_L / q) F_f) / m k_B T} \right)$$

$$V_T = V_{T0} - \delta V_{DS}$$

$$3) \quad \langle v_x(V_{DS}) \rangle = F_{SAT}(V_{DS}) v_{inj}$$

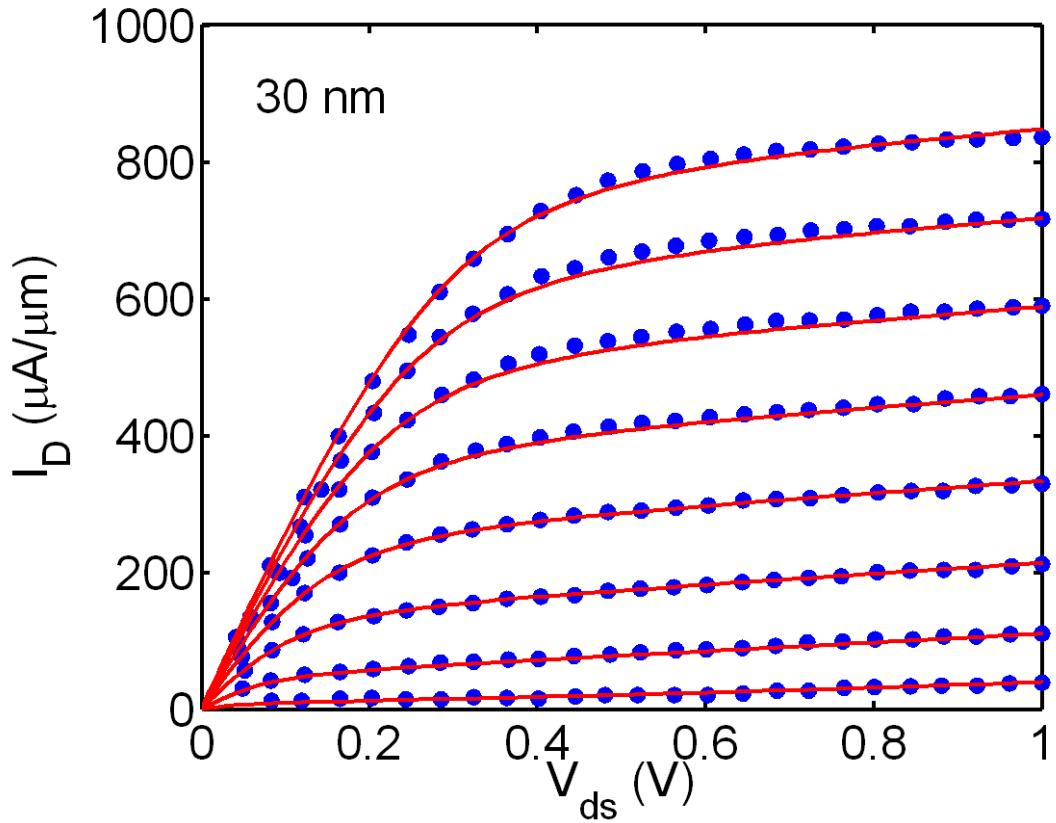
$$4) \quad F_{SAT}(V_{DS}) = \frac{V_{DS}/V_{DSAT}}{\left[1 + (V_{DS}/V_{DSAT})^\beta \right]^{1/\beta}}$$

$$5) \quad V_{DSAT} = \frac{v_{inj} L}{\mu_{app}}$$

Only 10 device-specific parameters in this model:

$$C_{inv}, V_{T0}, \delta, m, v_{inj}, \mu_{app}, L, R_{SD} = R_S + R_D, \alpha, \beta$$

MVS Fits to experimental Si ETSOI data



$$\mu_{app} = 220 \frac{\text{cm}^2}{\text{V}\cdot\text{s}}$$

$$v_{inj} = 0.82 \times 10^7 \text{ cm/s}$$

$$R_{S0} + R_{D0} = 130 \Omega\cdot\mu\text{m}$$

A. Majumdar and D.A. Antoniadis, "Analysis of Carrier Transport in Short-Channel MOSFETs," *IEEE Trans. Electron. Dev.*, **61**, pp. 351- 358, 2014.

MVS analysis of well-tempered MOSFETs

$$\mathcal{T}_{LIN} = \frac{\mu_{app}}{\mu_B} \quad \mathcal{T}_{SAT} = \left(\frac{2}{1 + v_T/v_{inj}} \right)$$

Unit 4

- 4.1 Landauer Approach
- 4.2 Landauer at low and High Bias
- 4.3 The Ballistic MOSFET
- 4.4 Velocity at the Virtual Source
- 4.5 Transmission Theory of the MOSFET
- 4.6 The VS Model Revisited
- 4.7 Analysis of Experiments

Essentials of MOSFETs

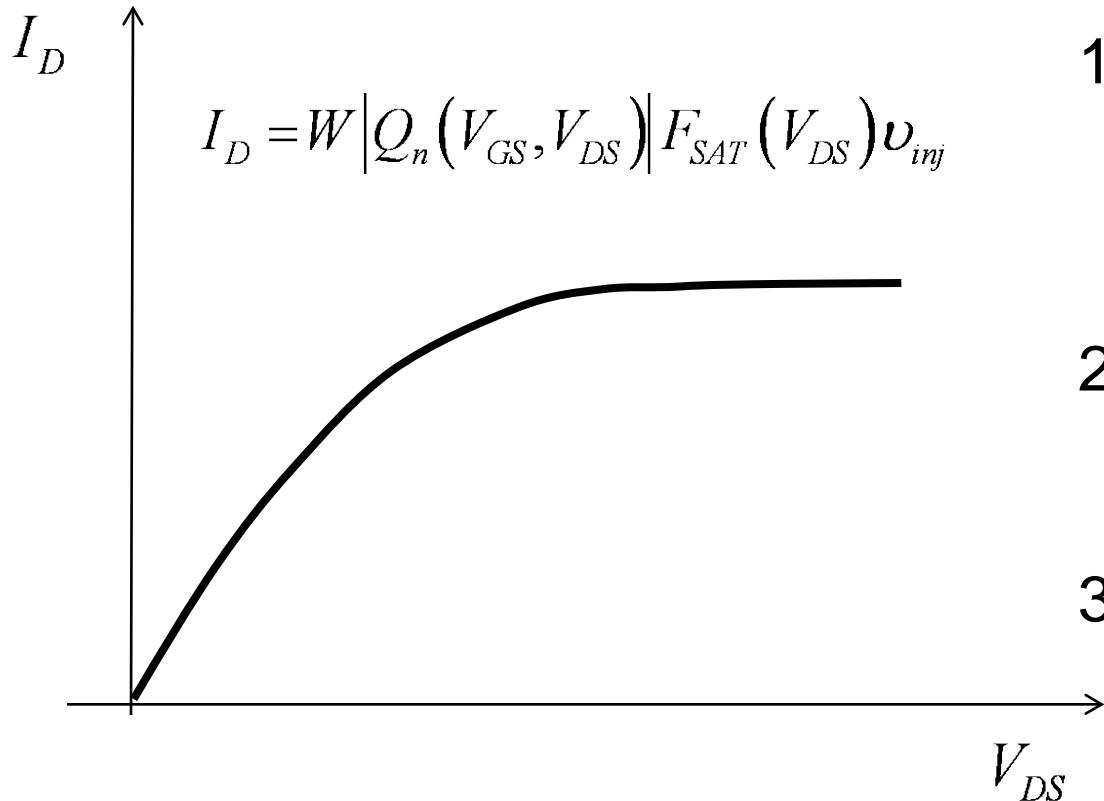
Unit 5: Additional Topics

Lecture 5.1: Limits of MOSFETs

Mark Lundstrom

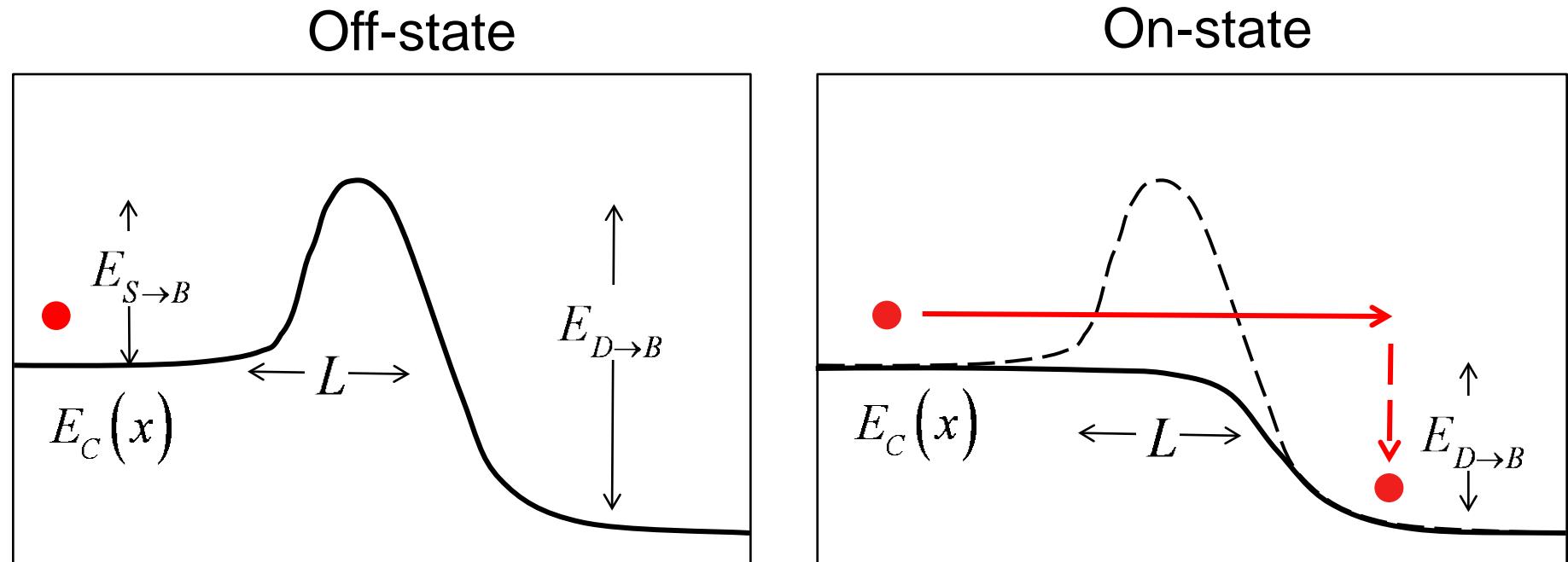
lundstro@purdue.edu
Electrical and Computer Engineering
Purdue University
West Lafayette, Indiana USA

Questions (for digital logic)



- 1) What are the fundamental limits of MOSFETs?
- 2) How close are we to those limits today?
- 3) What sets the limits in practice?

MOSFETs are barrier controlled devices



(Recall Lecture 2.2)

Fundamental limits of MOSFETs

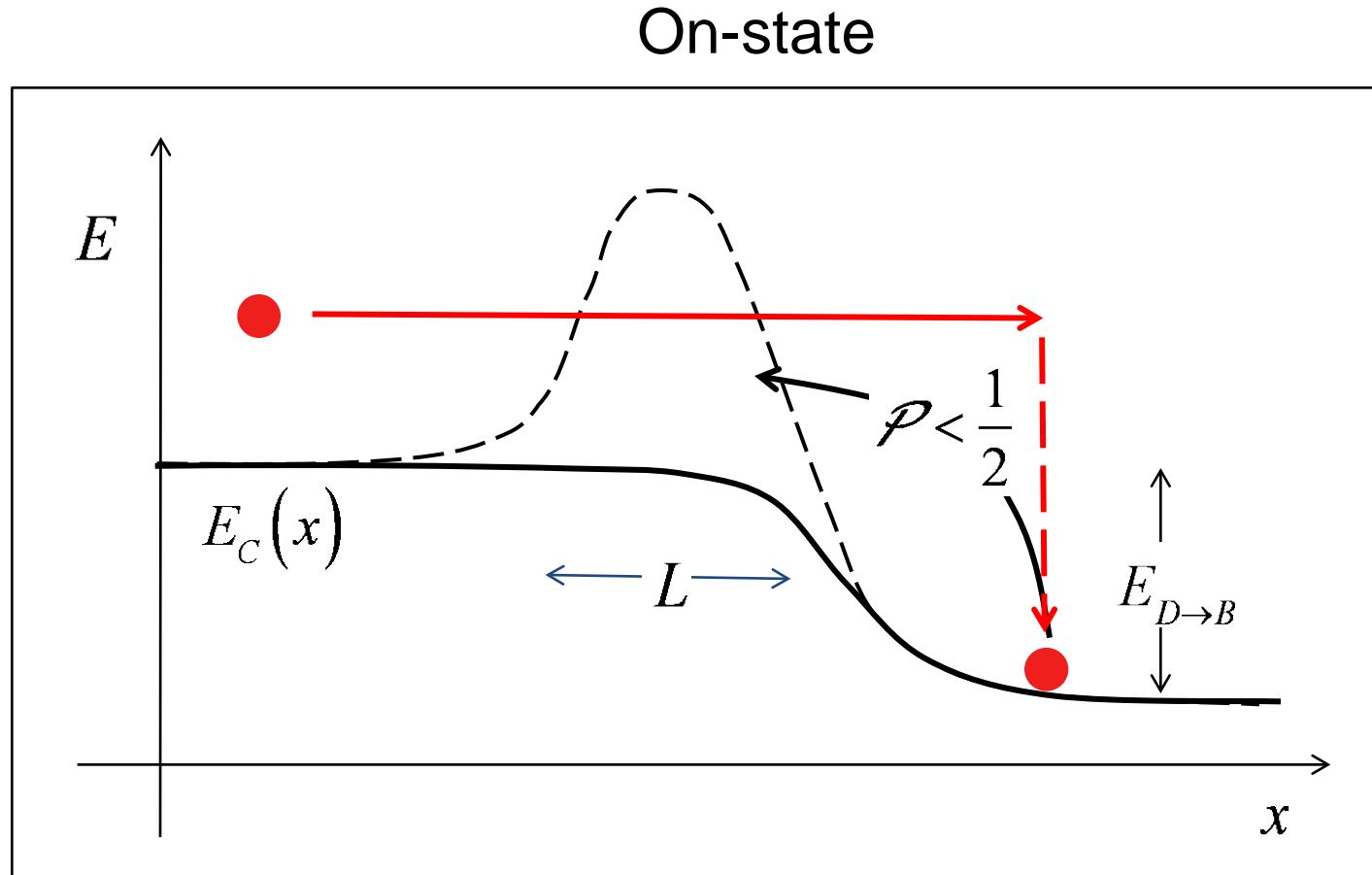
We will use some very simple arguments to estimate some ultimate limits for transistors. Our approach is similar to (but not quite the same) as the approach of Zhirnov, *et al.*

V. V. Zhirnov, R.K. Cavin III, J.A. Hutchby, and G.I. Bourianoff,
"Limits to Binary Logic Switch Scaling – A Gedanken Model,"
Proc. IEEE, **91**, pp. 1934 - 1939 , 2003.

Three questions

- 1) What is the minimum switching energy for a MOSFET?
- 2) What is the minimum channel length?
- 3) What is the minimum switching time?

1) Minimum switching energy



Minimum switching energy

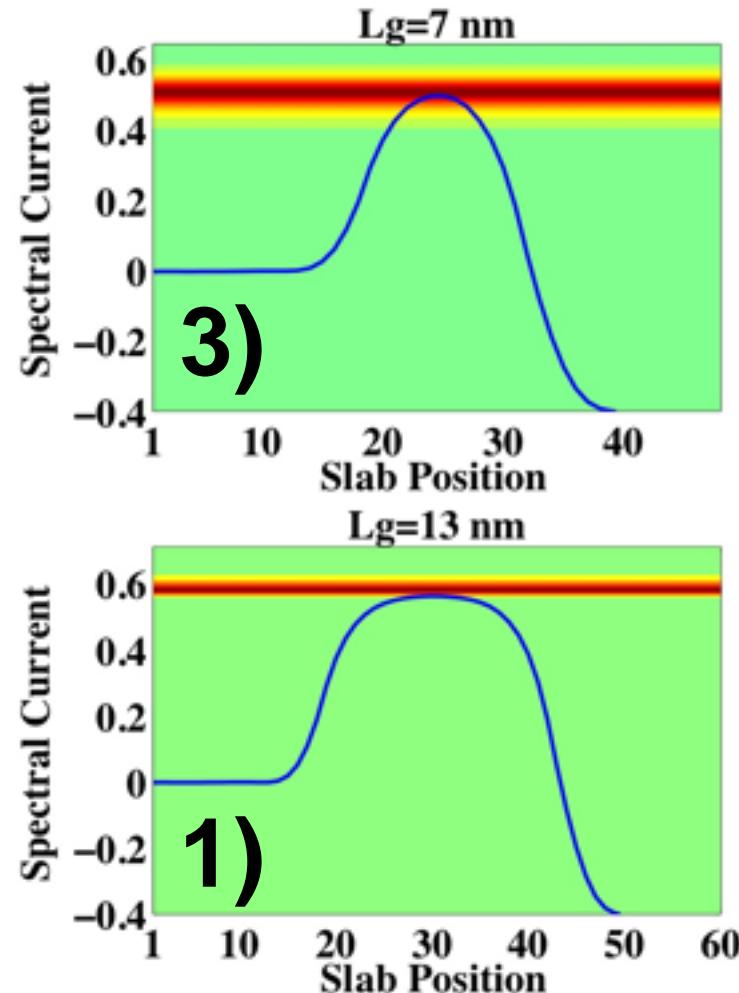
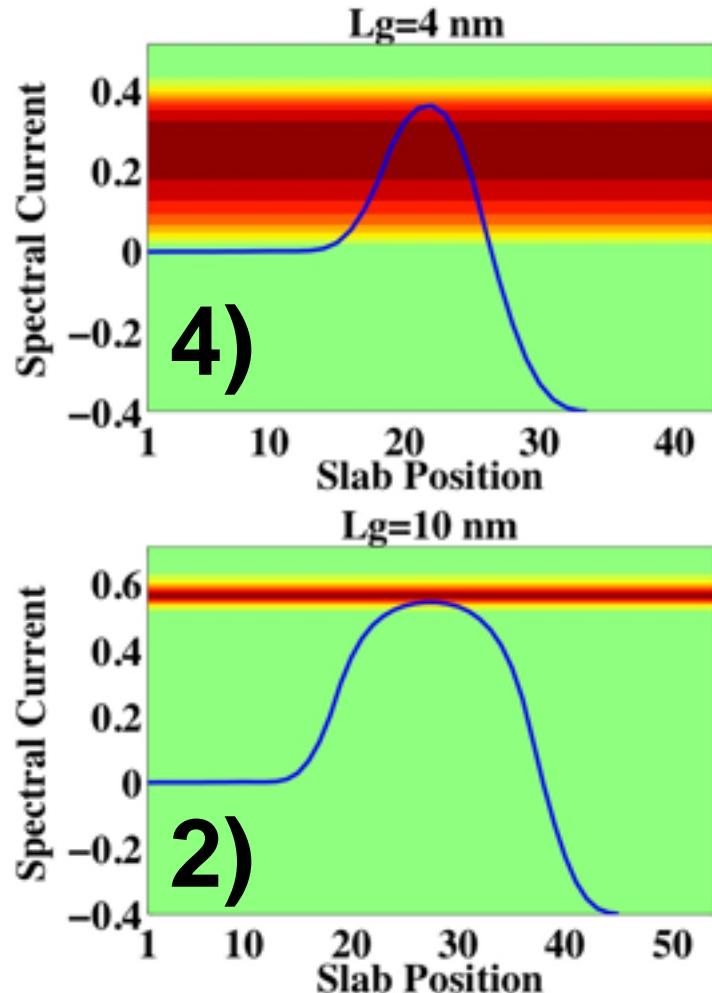
To have a switching event, the electron must stay in the drain and not be thermionically re-emitted back to the source. We require:

$$\mathcal{P} = e^{-E_{D \rightarrow B}/k_B T} < \frac{1}{2}$$

$$e^{-E_{\min}/k_B T} = \frac{1}{2}$$

$$E_{\min} = k_B T \ln 2$$

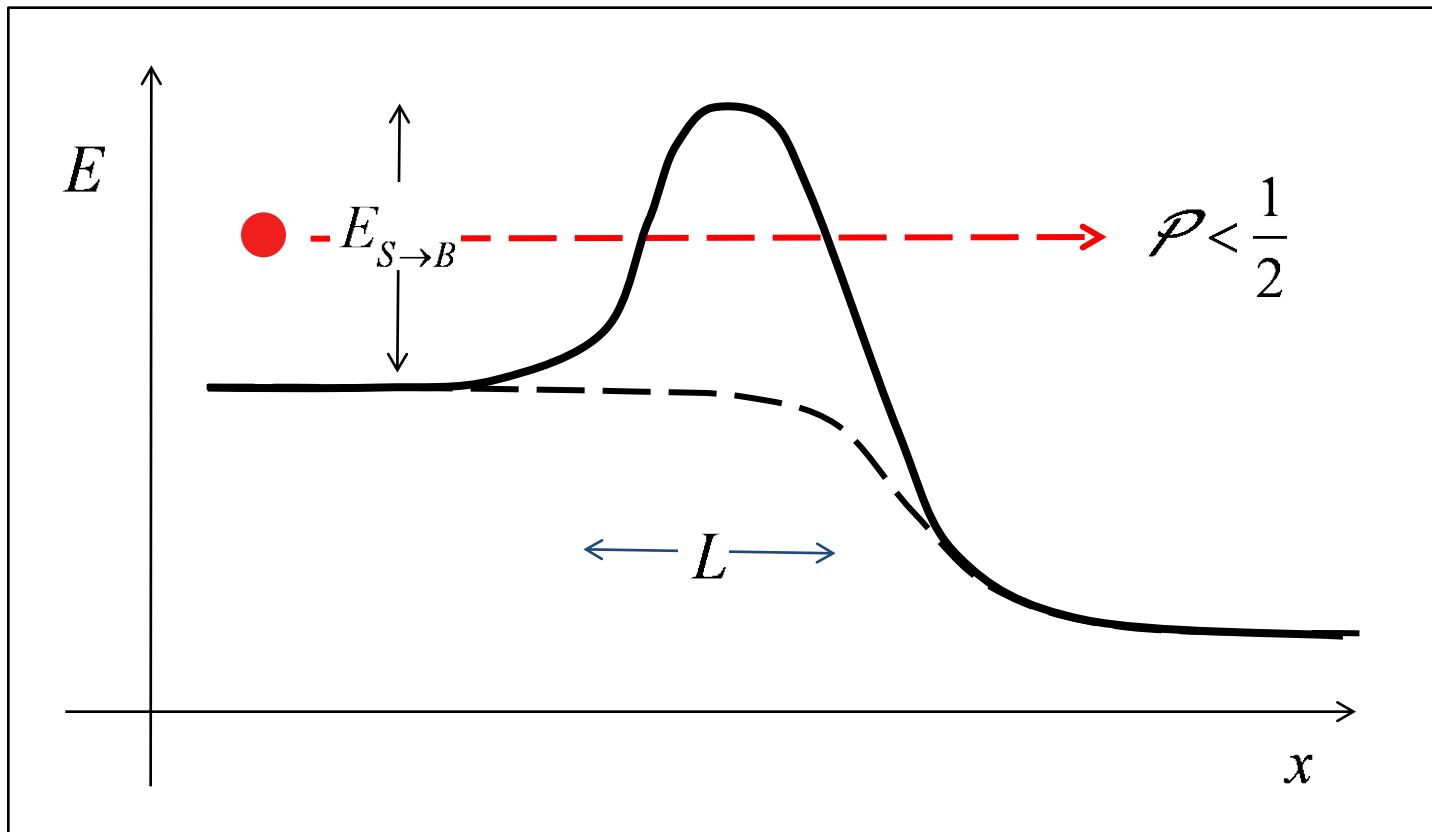
2) Minimum channel length and QM tunneling



(Provided by M. Luisier, ETH Zurich)

Minimum channel length

Off-state



Minimum channel length

$$\mathcal{P} = e^{-2\sqrt{2m^*E_{S \rightarrow B}}L/\hbar} \quad (\text{WKB approximation})$$

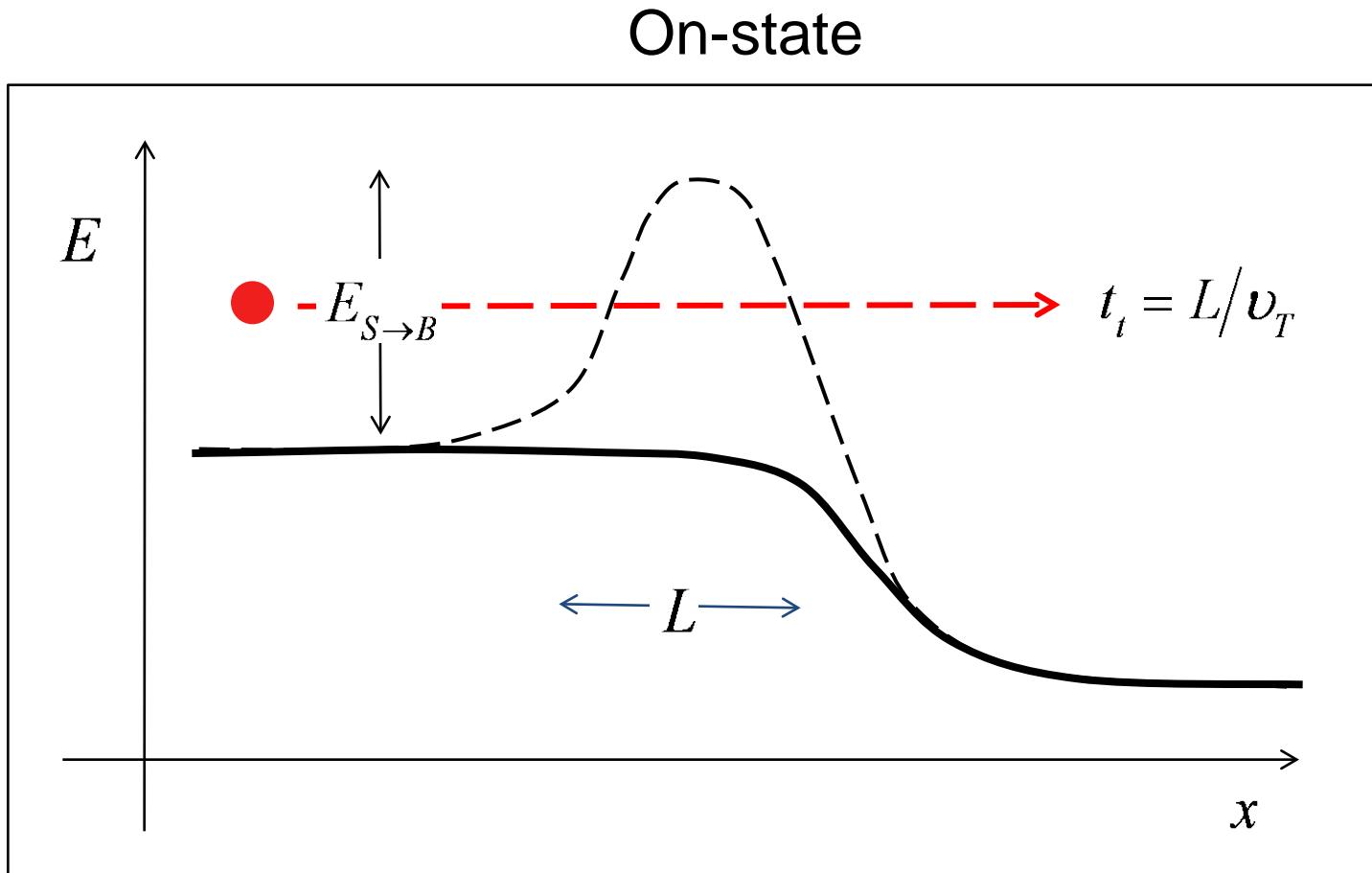
$$\mathcal{P} < \frac{1}{2} \quad (\text{To say that the device is off})$$

$$L > \frac{\hbar}{\sqrt{2m^*E_{S \rightarrow B}}}$$

$$E_{S \rightarrow B} = E_{\min}$$

$$L_{\min} = \frac{\hbar}{\sqrt{2m^*E_{\min}}}$$

3) Minimum switching time



Minimum switching time

The minimum switching time is the **transit time** – the time it takes for an electron to cross the channel.

$$t_t = \frac{L}{v_T}$$

$$\tau_{\min} = \frac{L_{\min}}{v_T} \quad v_T = \sqrt{\frac{2k_B T}{\pi m^*}} \quad E_{\min} = k_B T \ln 2$$

(Discarding some constants on the order of unity...)

$$\tau_{\min} = \frac{\hbar}{E_{\min}}$$

“Fundamental limits”

$$E_{\min} = k_B T \ln 2 = 0.017 \text{ eV}$$

$$L_{\min} = \frac{\hbar}{\sqrt{2m^* E_{\min}}} = 1.5 \text{ nm}$$

$$(m^* = m_0)$$

$$\tau_{\min} = \frac{\hbar}{E_{\min}} = 40 \text{ fs}$$

Discussion

The minimum switching energy, $E_{\min} = k_B T \ln 2$, is usually derived using thermodynamic arguments.

The minimum channel length and switching times can also be derived from the two quantum mechanical uncertainty relations:

$$\Delta E \Delta t \geq \frac{\hbar}{2}$$

$$\Delta p \Delta x \geq \frac{\hbar}{2}$$

Discussion

The fact that the fundamental limits for MOSFETs can also be derived with some very general thermodynamic and quantum mechanic arguments, suggests that they apply to any switching device.

22 nm technology

$$V_{DD} = 0.7 \text{ V}$$

$$C_{inv} \approx 2.9 \times 10^{-6} \text{ F/cm}^2$$

$$C_G = C_{inv}(WL) \approx 24 \times 10^{-18} \text{ F}$$

$$W = L = 22 \text{ nm}$$

$$I_{ON} \approx 1 \times 10^{-3} \text{ A}/\mu\text{m}$$

$$\frac{1}{2} C_G V_{DD}^2 \approx 1200 \times E_{\min}$$

$$L = 22 \text{ nm} = 15 \times L_{\min}$$

$$\tau = \frac{C_G V_{DD}}{I_{ON}} \approx 11 \times \tau_{\min}$$

Channel length and switching time are within about a factor of 10 above fundamental limits. Energy (and power) are orders of magnitude larger than the fundamental limits.

What sets the limits in practice?

1) Minimum switching energy: $E_{\min} = k_B T \ln(2)$

In practice, the switching energy is: $E_{swn} = \frac{1}{2} C_{sw} V_{DD}^2$

The switching capacitance includes all the parasitic capacitances and the capacitance of the wiring. It can be quite large ($\sim 1 \text{ fF/node}$).

The power supply voltage is far above the minimum because low error computation requires an on-off ratio of $\sim 10^4$. The MOSFET IV characteristic then dictates that $V_{DD} \sim 1 \text{ V}$.

What sets the limits in practice?

2) Minimum channel length: $L_{\min} = \frac{\hbar}{\sqrt{2m^* E_{\min}}}$

The fundamental limit is set by quantum mechanical tunneling assuming an on-off ratio of 1.

In practice, 2D electrostatics is the main concern (leading to the replacement of planar FETs by FinFETs).

The increasing cost of lithography is also a concern.

What sets the limits in practice?

3) Minimum switching speed:

$$\tau_{\min} = \frac{\hbar}{E_{\min}}$$

In practice, the switching speed is:

$$\tau_{dev} = \frac{C_{dev} V_{DD}}{I_{ON}}$$

The device capacitance includes a parasitic capacitance.

$$C_{dev} = WLC_{inv} + C_{par}$$

The parasitic gate-drain capacitance is a larger and larger fraction of the total gate capacitance.

Summary

- 1) Transistors are approaching some very fundamental limits.
- 2) Practical, technology considerations such as series resistance, parasitic capacitance, BTBT leakage currents, etc. are likely to set the practical limits.
- 3) It is unlikely that any digital switching device that is fundamentally better than a MOSFET exists.

Next topic

We have focused on digital and analog applications of MOSFETs, but there are other applications.

For example, MOSFETs can be used as switching devices in power electronic systems.

Essentials of MOSFETs

Unit 5: Additional Topics

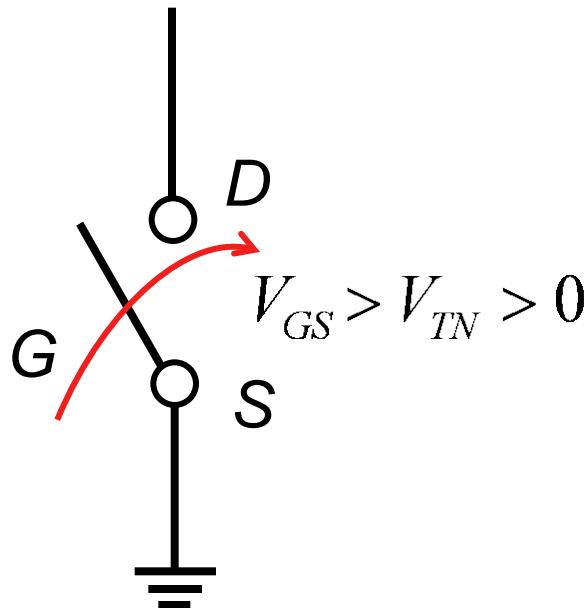
Lecture 5.2: Power MOSFETs

Mark Lundstrom

lundstro@purdue.edu
Electrical and Computer Engineering
Purdue University
West Lafayette, Indiana USA

Power MOSFETs are used as switches

$$V_{DS} > 0$$

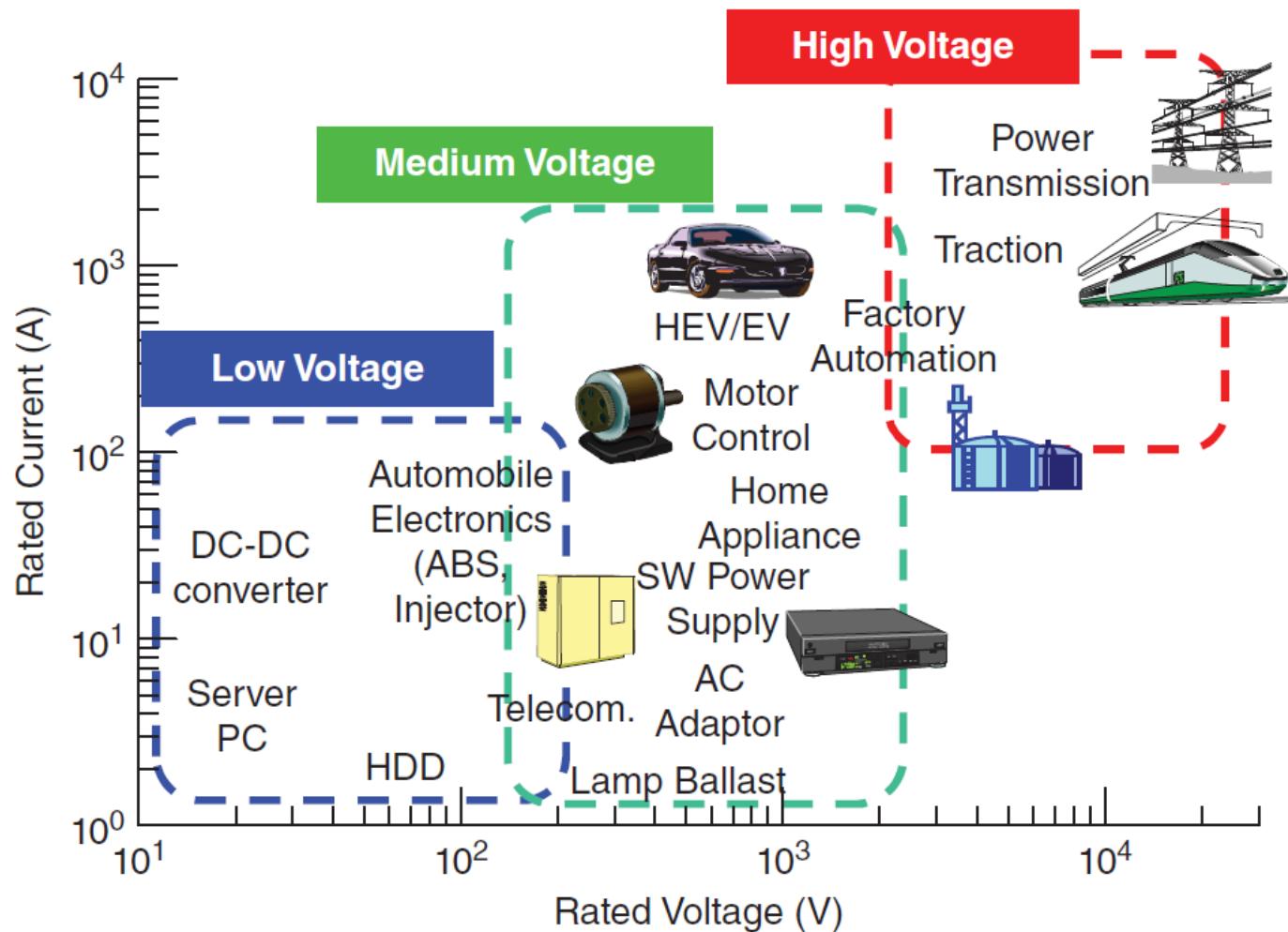


When **on**, the resistance of the switch should be as low as possible.

When **off**, the device should block current to as high a voltage as possible.

- 1) Low on resistance**
- 2) High breakdown voltage**

Applications



T. Kimoto and J. A. Cooper, *Fundamentals of Silicon Carbide Technology*, Wiley (2014).

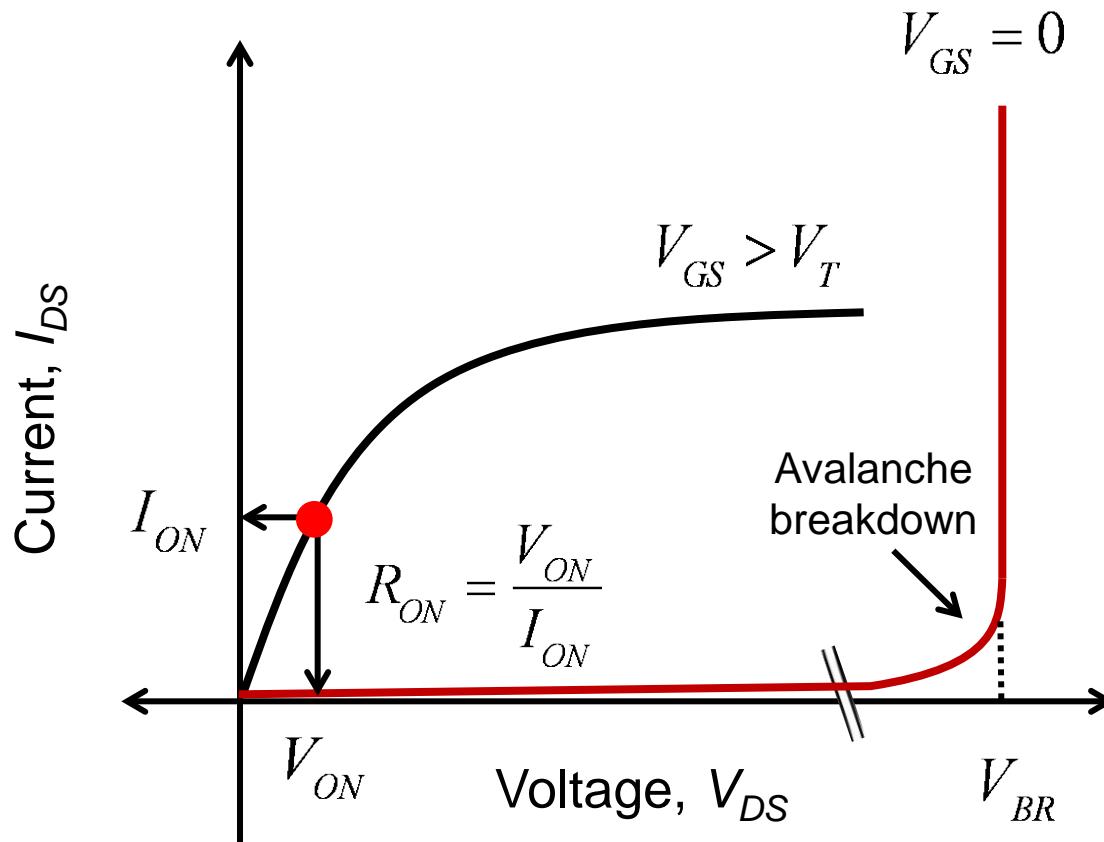
Power semiconductor devices

There are several different types of three-terminal, power semiconductor switching devices – e.g.

- Vertical Diffused MOS transistors (DMOSFET)
- V-groove trench MOSFETs
- Trench MOSFETs
- Insulated Gate Bipolar Transistors (IGBT)
- Superjunction devices

To illustrate some of the considerations, we will only discuss the DMOSFET transistor.

On and off-states



Device design goals

On-state: Maximize I_{ON} (minimize R_{ON})

Off-state: Maximize V_{BR}

Minimize area

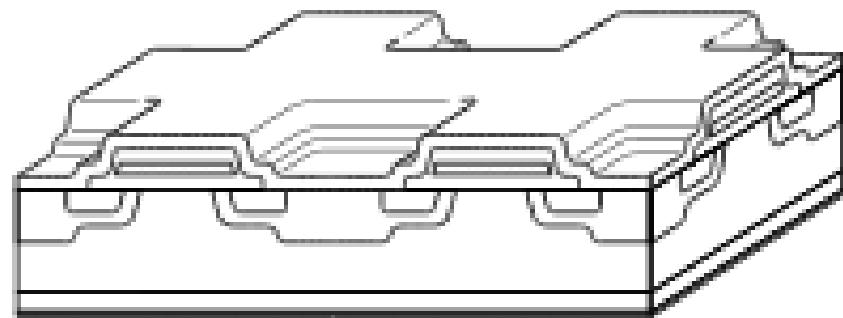
Specific on-resistance: $R_{ON,SP} = R_{ON} \times A$

Device design goals

High currents require large W (MOSFET width)

High breakdown voltage requires that we spread the voltage drop out (long channels) to minimize the electric field.

To achieve these goals, power MOSFETs use vertical current flow and cellular structures.



https://en.wikipedia.org/wiki/Power_MOSFET

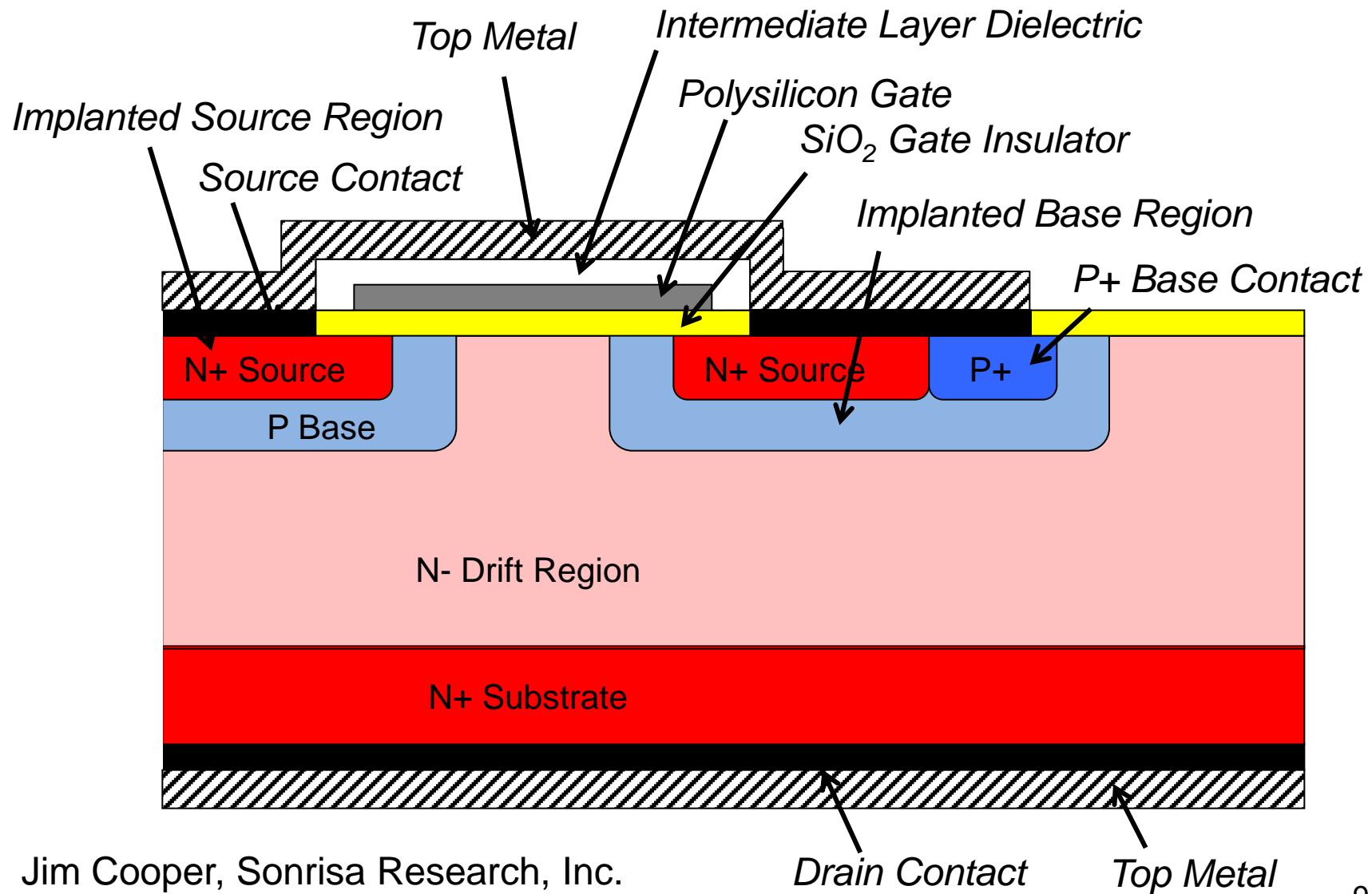
Area must be minimized

$$R = \rho \frac{L}{A} \quad R \times A = R_{SP} = \rho L \Omega \text{-cm}^2 \quad \text{specific resistance}$$

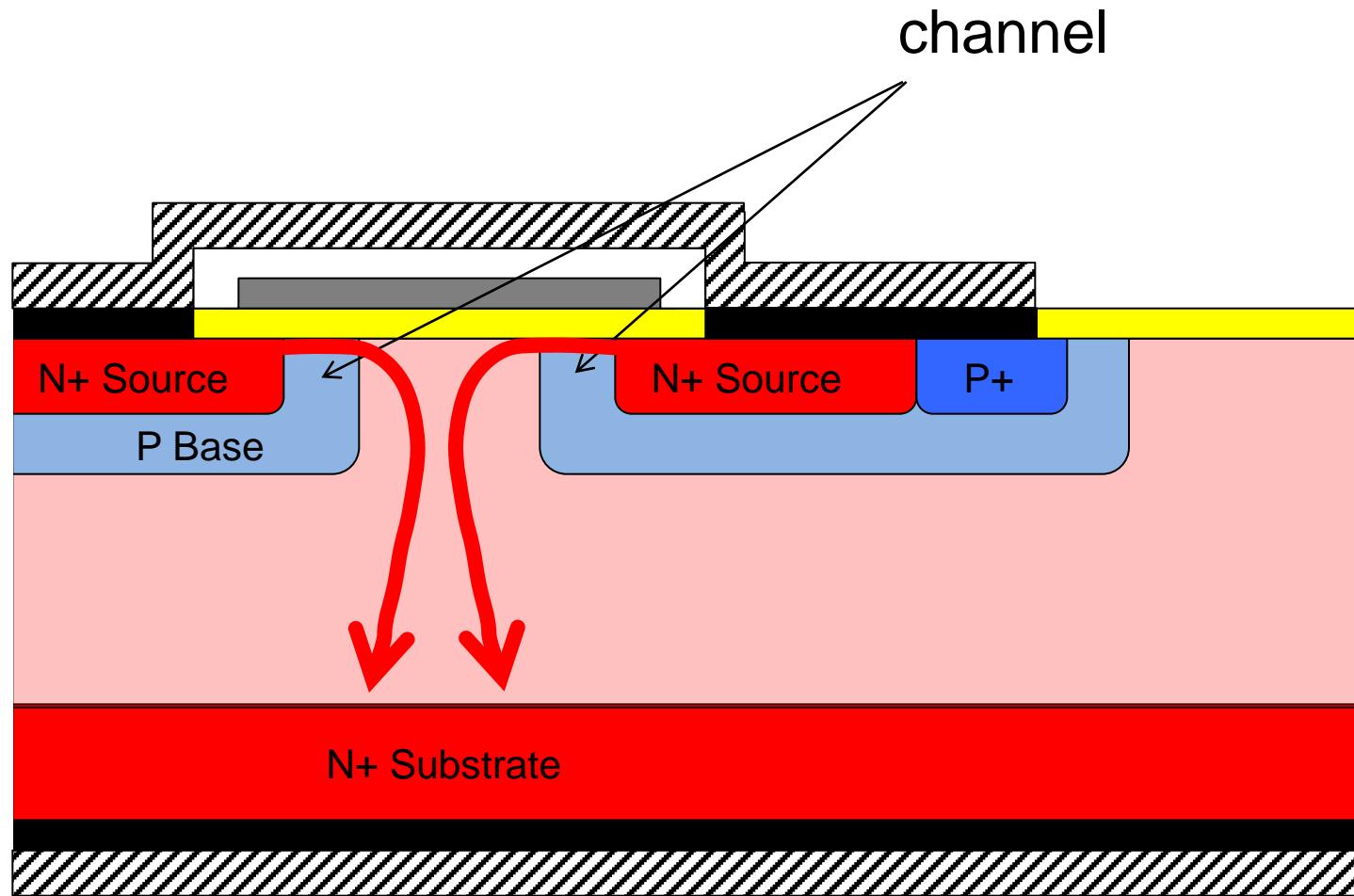
$$A_{DIE} (\text{cm}^2) = \frac{R_{ON,SP} (\Omega - \text{cm}^2)}{R_{ON} (\Omega)} \quad \longrightarrow \quad \text{Die cost}$$



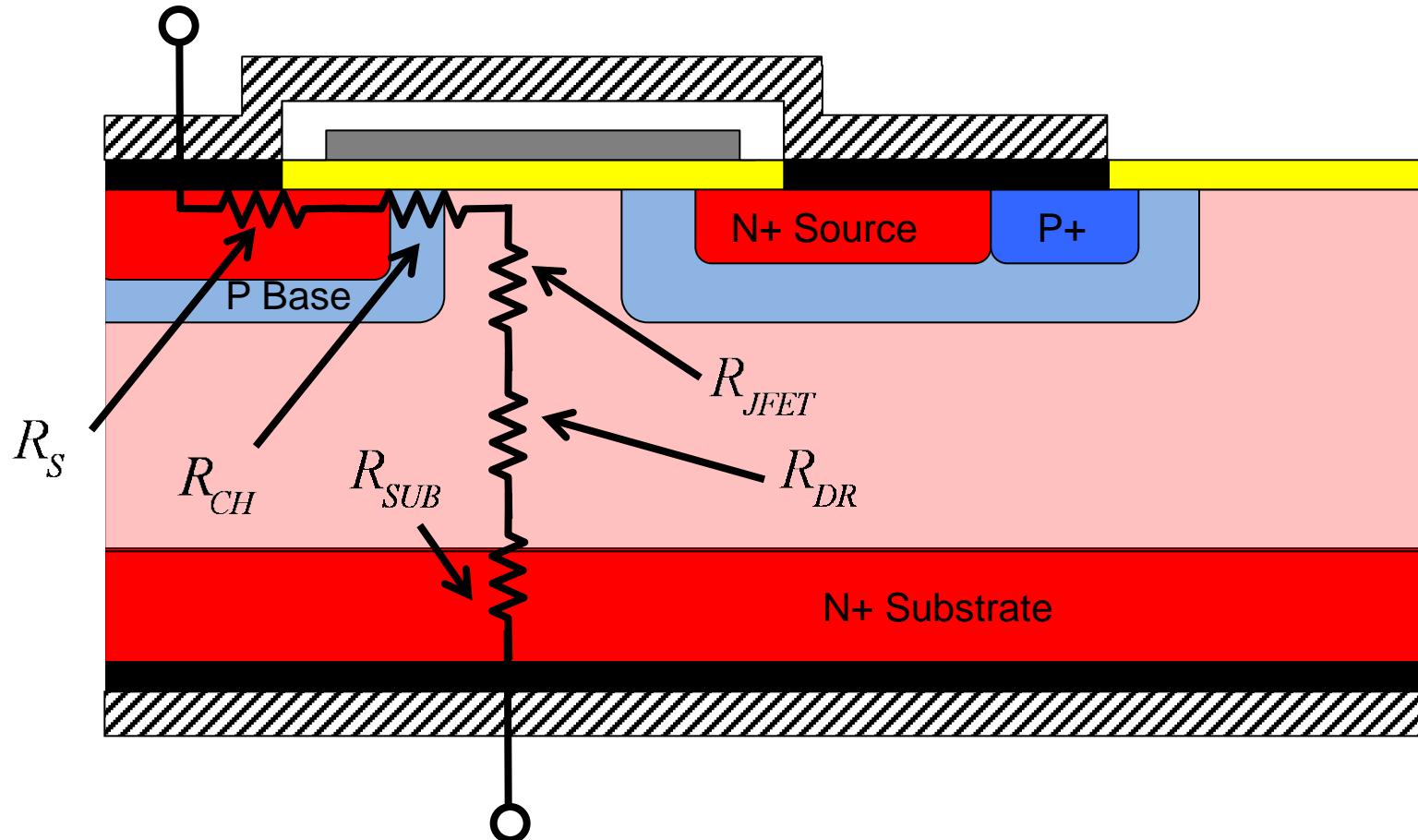
DMOSFET



DMOSFET: On-state



DMOSFET: On-state resistances



On resistance

$$R_{ON} = R_S + R_{CH} + R_{JFET} + R_{DR} + R_{SUB}$$

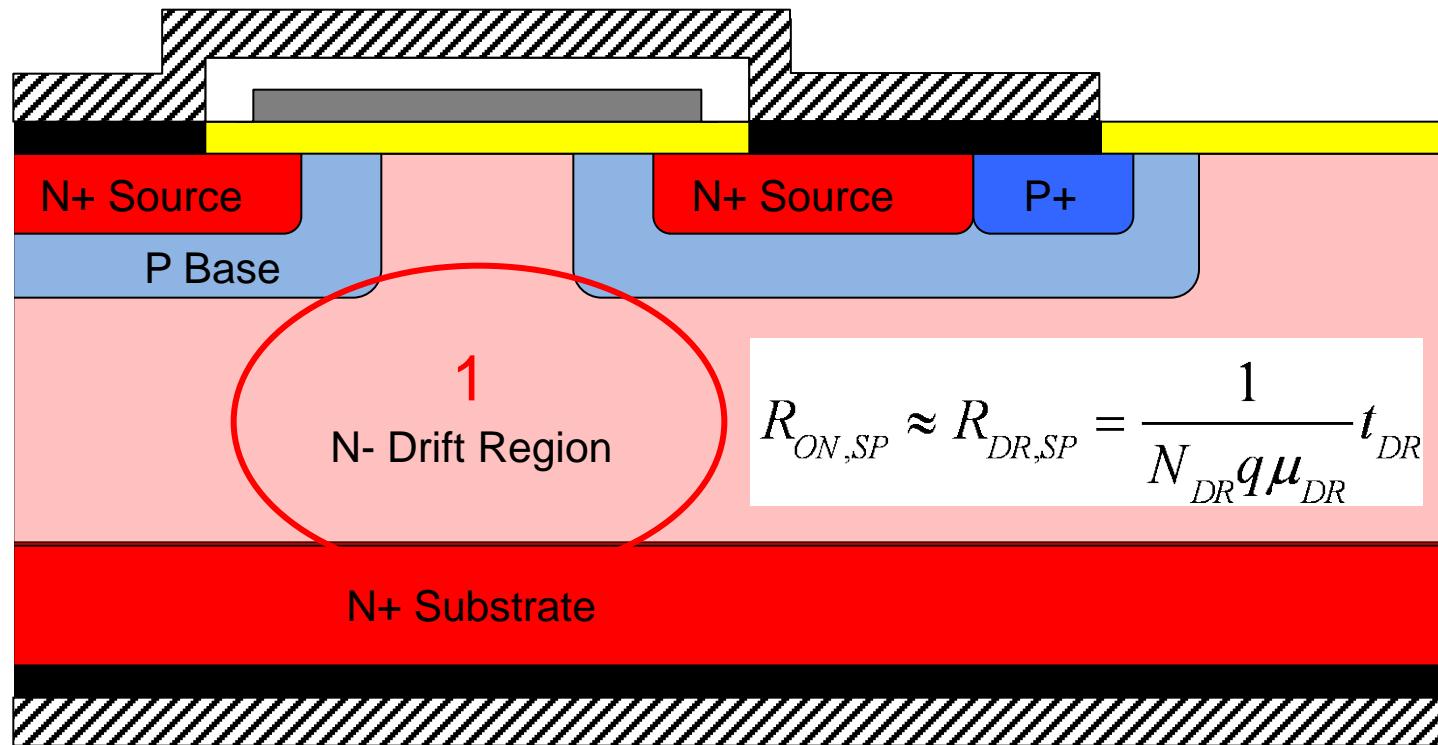
For high voltage MOSFETs, the N⁻ region is thick, so R_{DR} dominates.

$$R = \rho \frac{t_{DR}}{A}$$

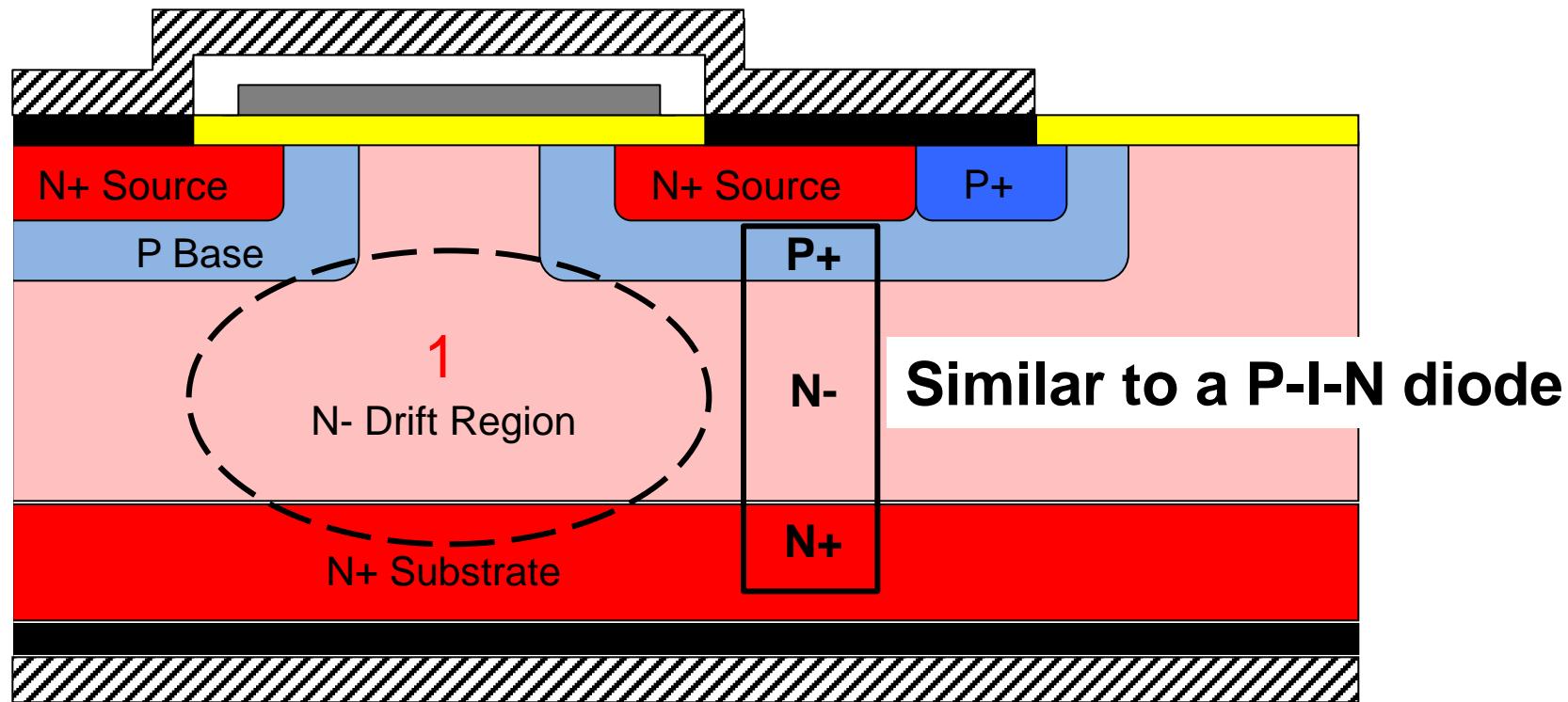
$$R_{DR,SP} = \frac{1}{N_{DR} q \mu_{DR}} t_{DR} \left(\Omega \cdot \text{cm}^2 \right)$$

$$R \times A = R_{SP} = \rho t_{DR} \Omega \cdot \text{cm}^2$$

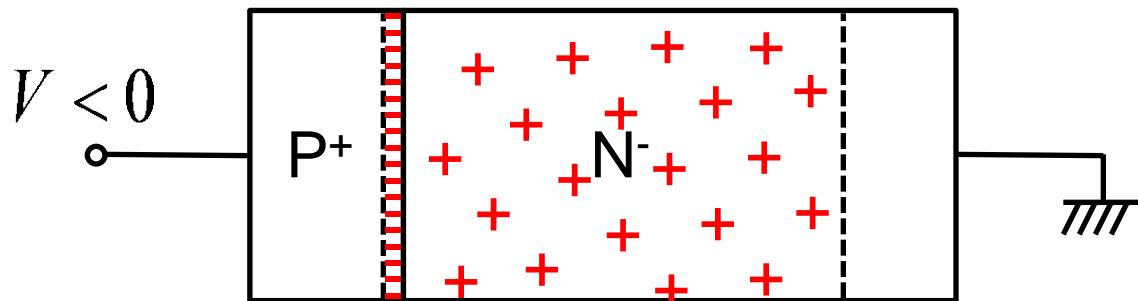
DMOSFET: On resistance



DMOSFET: Breakdown voltage

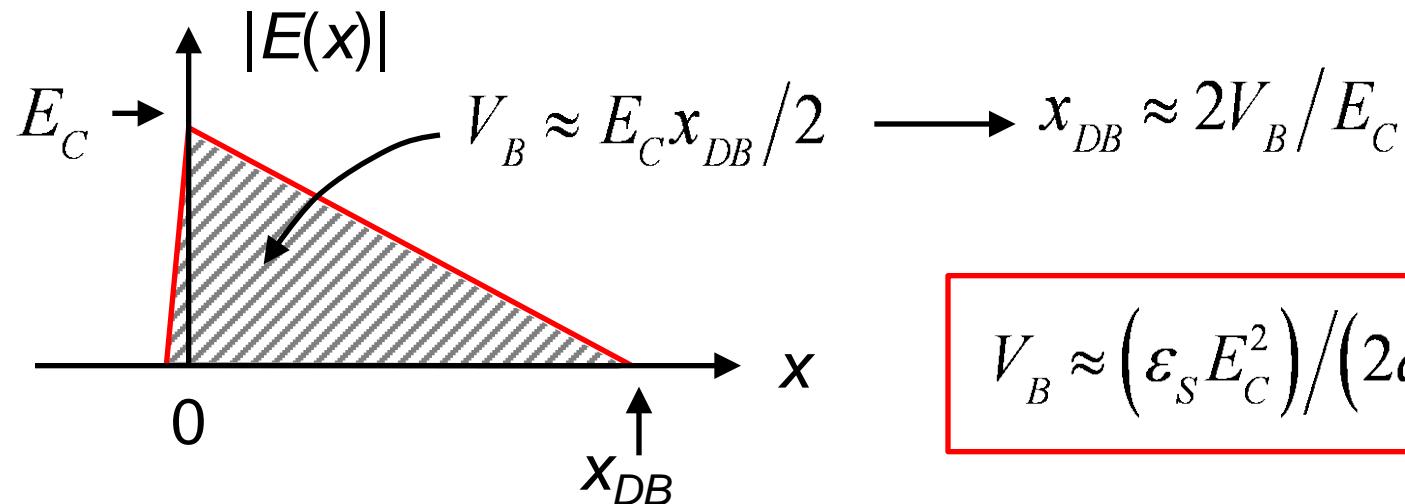


P+/N- diode at reverse breakdown



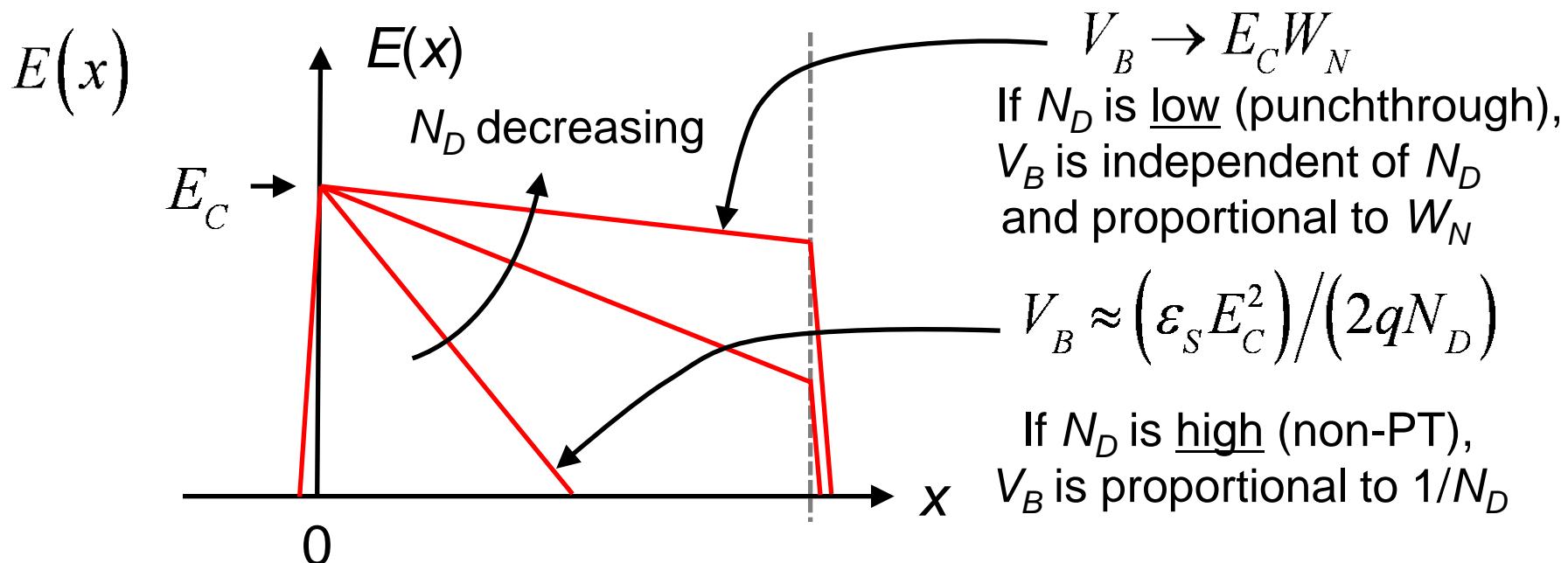
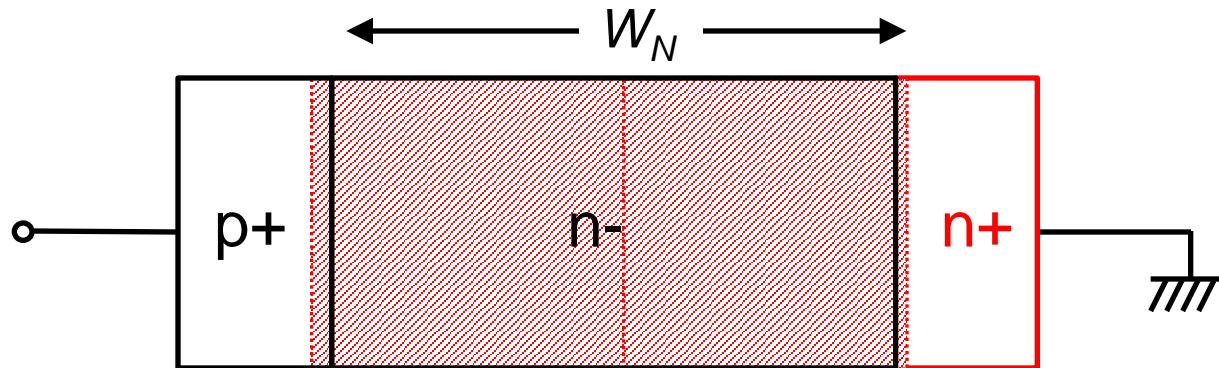
For a one-sided step junction,

$$x_{DB} \approx \sqrt{2\epsilon_s V_B / (qN_D)}$$

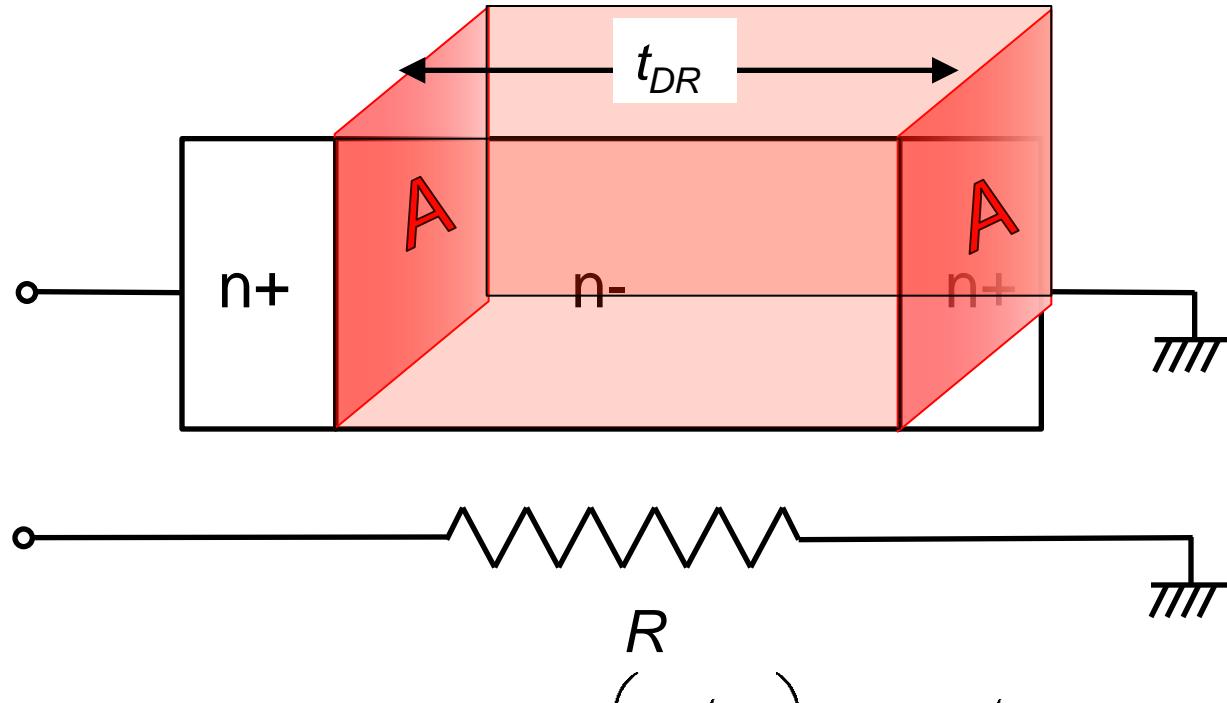


$$V_B \approx (\epsilon_s E_C^2) / (2qN_D)$$

P+/N- diode at reverse breakdown



Specific on-resistance



$$R_{ON,SP} = R \times A = \left(\rho \frac{t_{DR}}{A} \right) A = \frac{t_{DR}}{q\mu_n N_D}$$

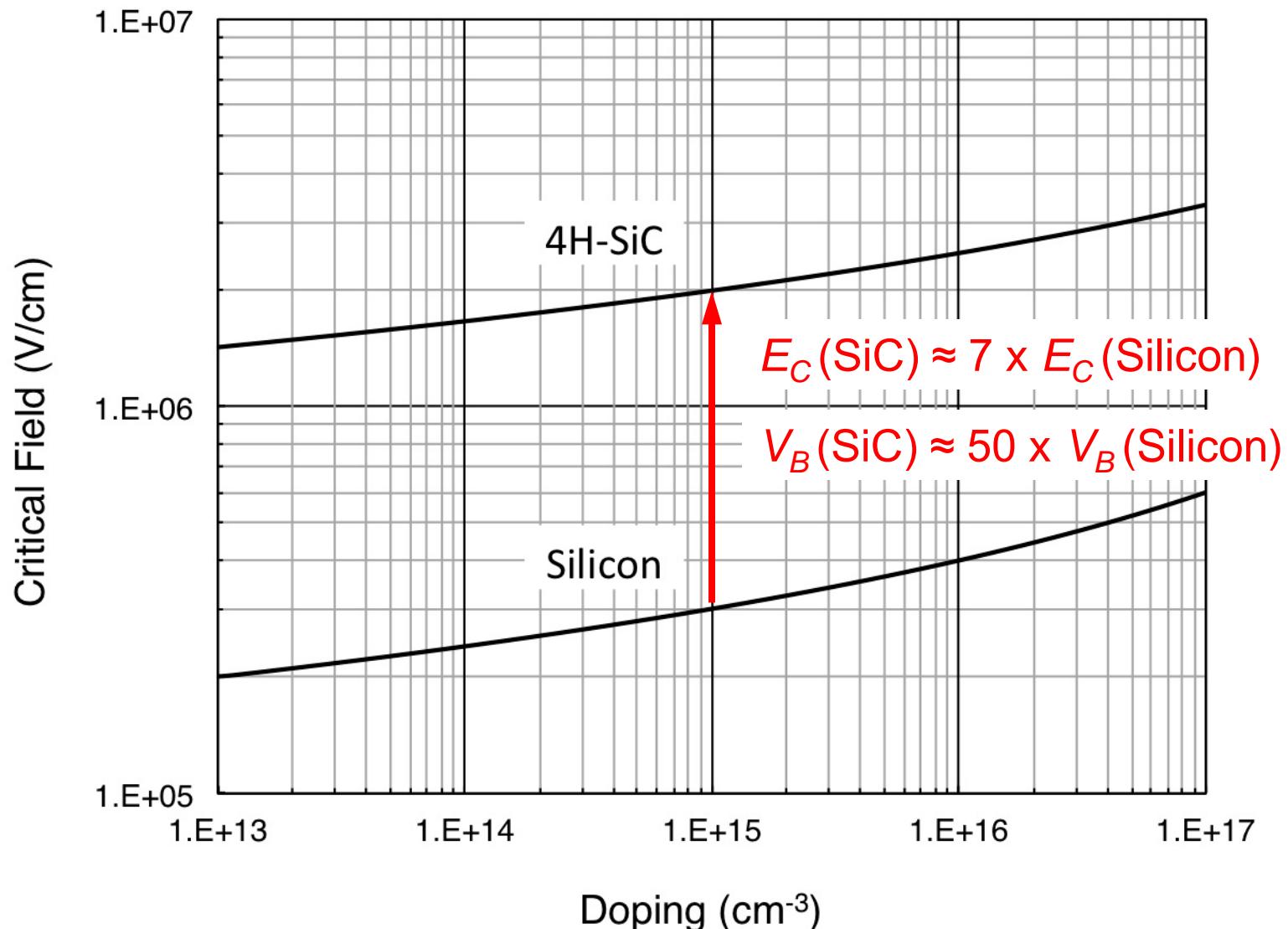
Resistance – breakdown voltage trade-off

$$V_B \approx (\epsilon_s E_C^2) / (2qN_D) \quad \Rightarrow \quad N_D = (\epsilon_s E_C^2) / (2qV_B)$$

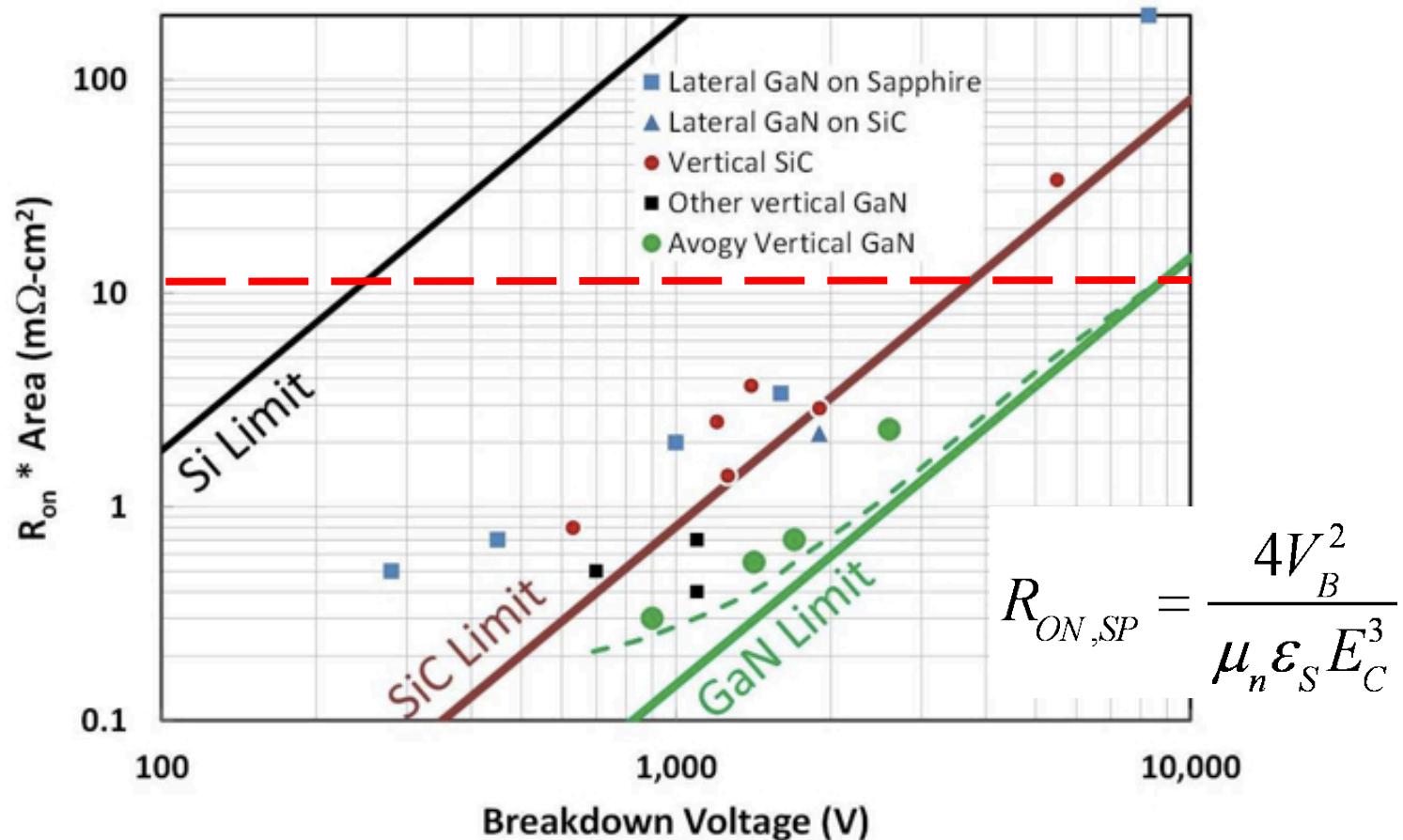
$$V_B = E_C x_{DB} / 2 \quad \Rightarrow \quad x_{DB} = 2V_B / E_C = t_{DR}$$

$$R_{ON,SP} = \frac{t_{DR}}{q\mu_n N_D} \quad \Rightarrow \quad R_{ON,SP} = \frac{4V_B^2}{\mu_n \epsilon_s E_C^3}$$

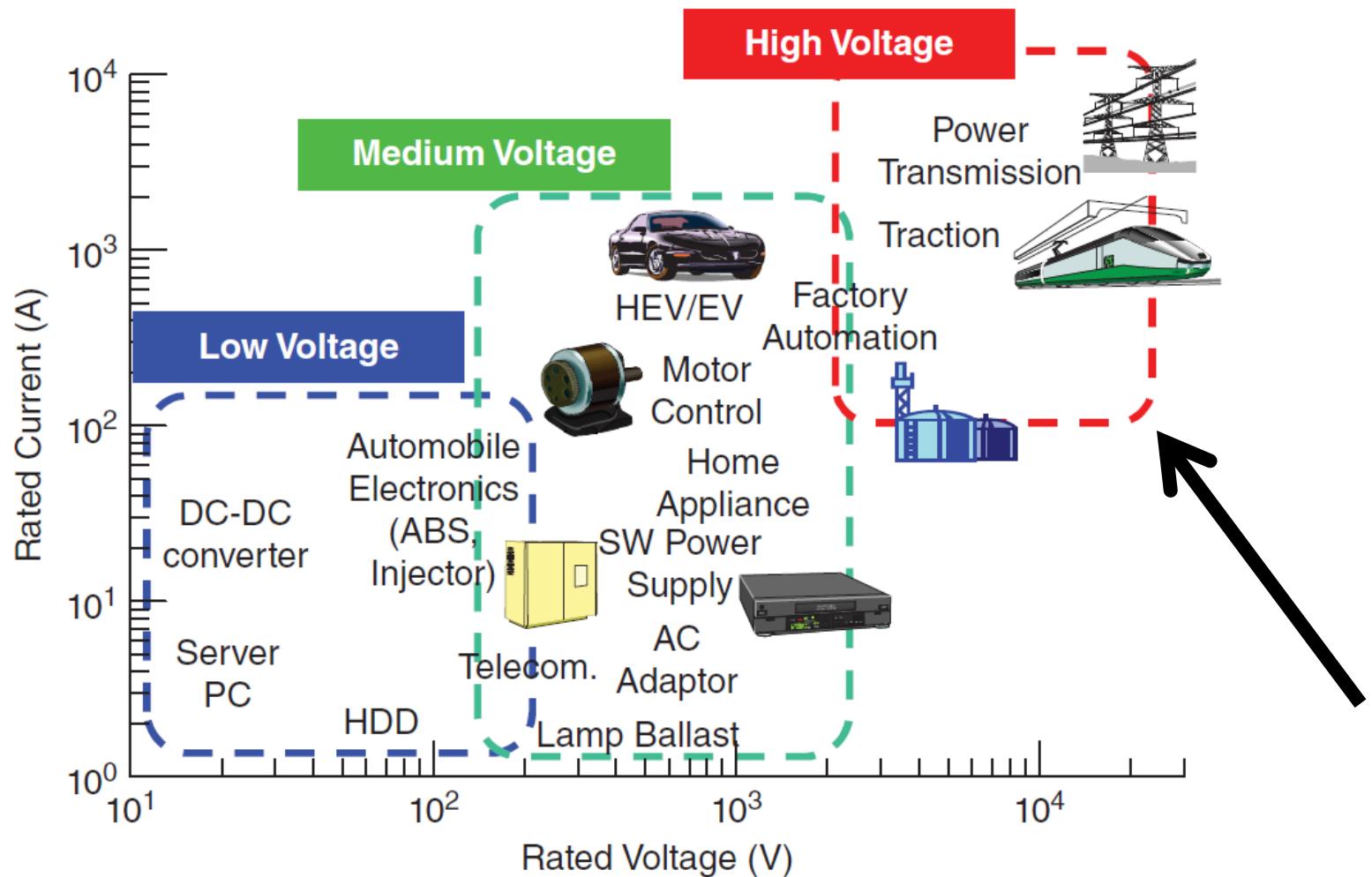
Critical field for breakdown



On resistance vs. blocking voltage

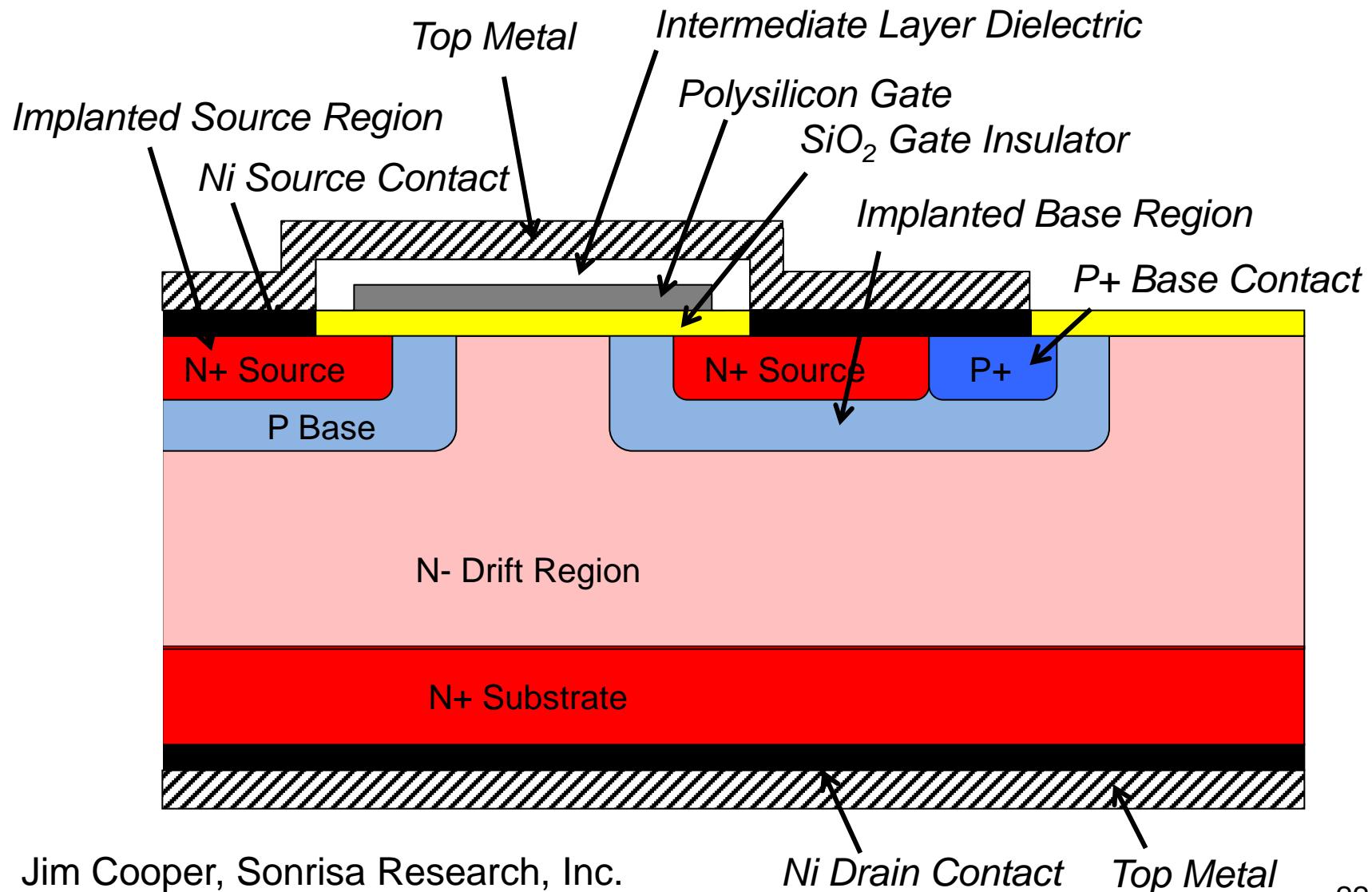


Applications

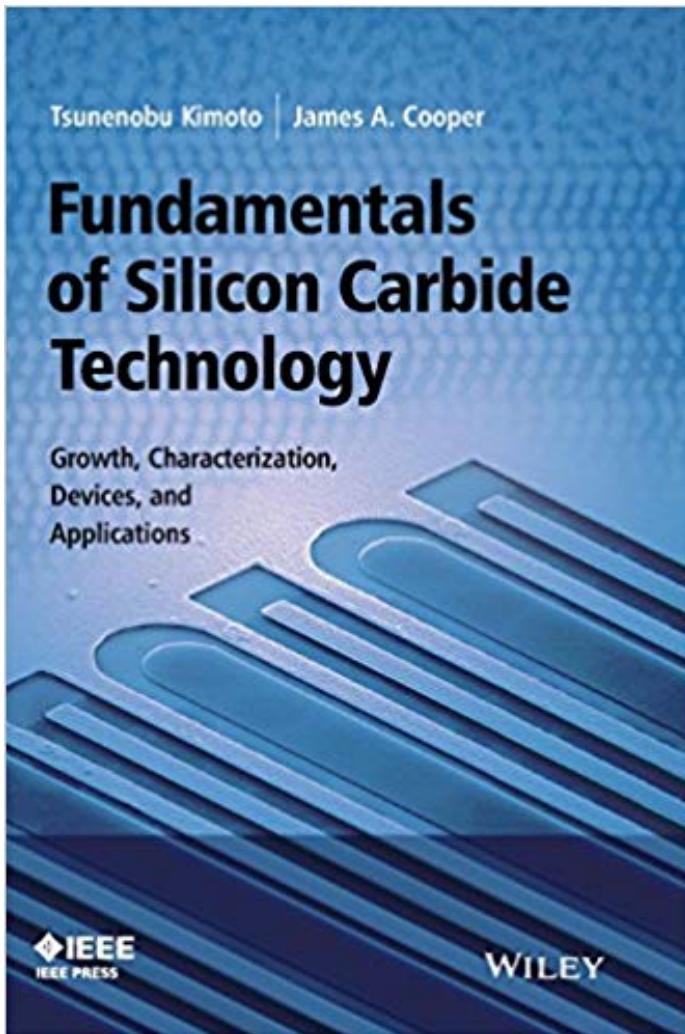


T. Kimoto and J. A. Cooper, *Fundamentals of Silicon Carbide Technology*, Wiley (2014).

DMOSFET



For more information



Tsunenobu Kimoto and
James A. Cooper,
*Fundamentals of Silicon
Carbide Technology*,
Wiley (2014).

Summary

- 1) MOSFETs have important applications beyond digital and RF/analog electronics.
- 2) Power MOSFET device geometries and design considerations are much different.
- 3) Power electronics is an important and rapidly advancing field.

Next topic

MOSFETs are not the only type of transistor. In the next lecture, we'll discuss the HEMT – an important RF device.

Essentials of MOSFETs

Unit 5: Additional Topics

Lecture 5.3: **High Electron Mobility Transistors (HEMTs)**

Mark Lundstrom

lundstro@purdue.edu

Electrical and Computer Engineering
Purdue University
West Lafayette, Indiana USA

Lundstrom: 2018

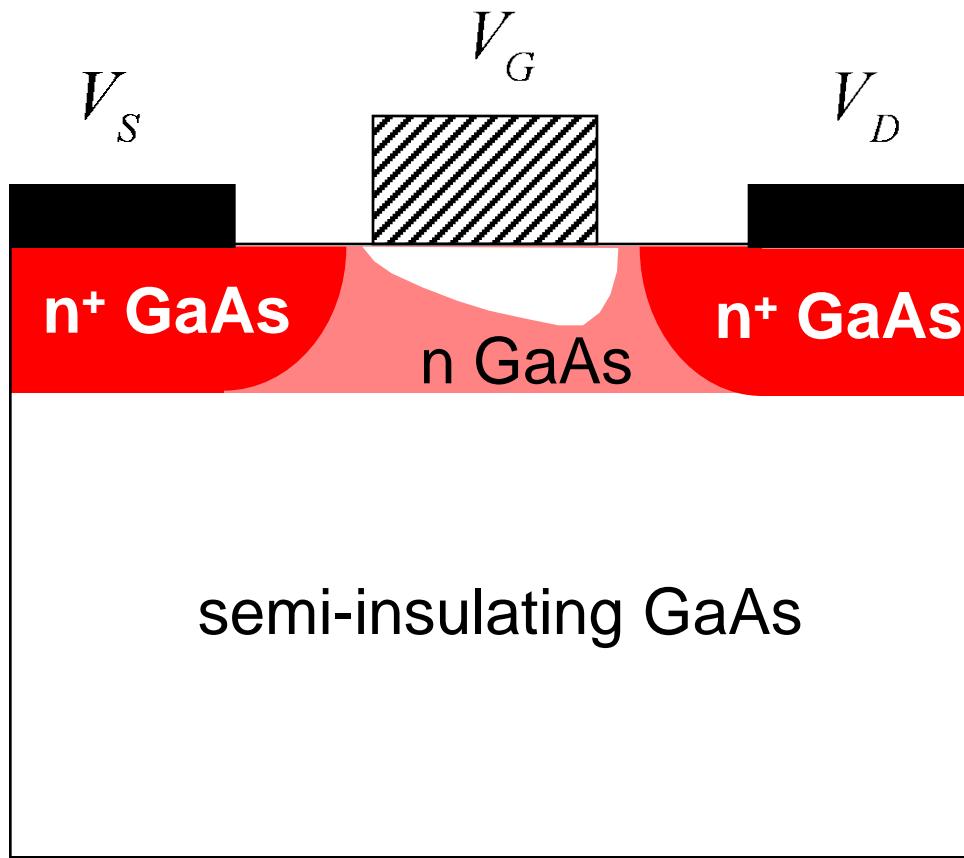
Transistors

MOSFET	HFET	RBT
MOST	DHFET	RHET
IGFET	HIGFET	QWBRTT
DMOS	SISFET	TETRAN
HEXFET	PBT	SIT
VMOS	LRTFET	NWFET
TFT	VMT	CNT FET
MISFET	BJT	SB FET
JFET	HBT	BTBT FET
VFET	DGBT	induced base transistor
MESFET	THETA	planar doped barrier transistor
MOSFET	RST	metal base transistor
HEMT	BICFET	Stark-effect transistor
TEGFET	RTBT	delta-doped channel heterojunction FET

***Most of these are
barrier-controlled
transistors.***

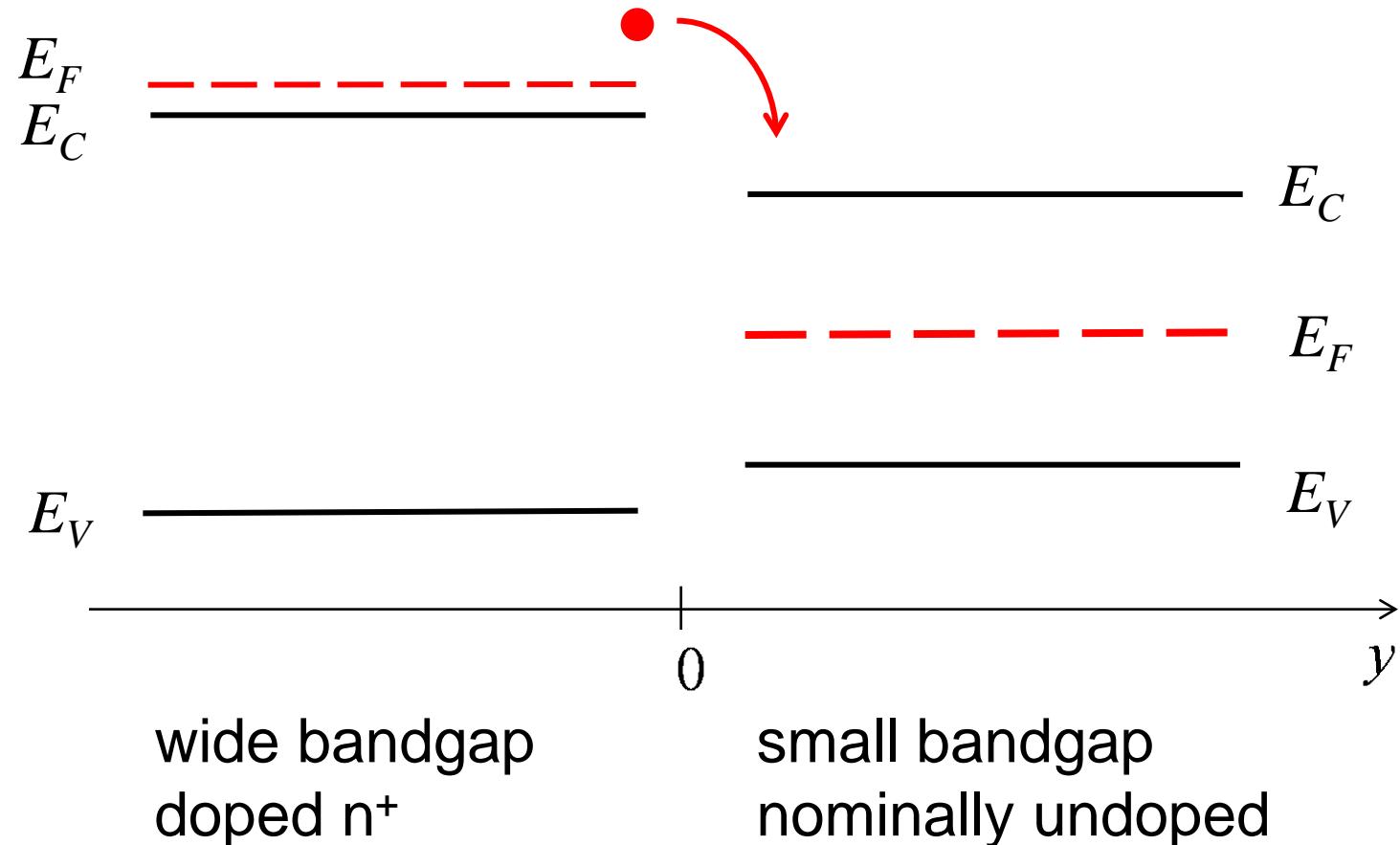
GaAs MESFET

NO INSULATING LAYER BECAUSE VERY VERY TOUGH TO BUILD HIGH QUALITY INSULATING LAYERS.



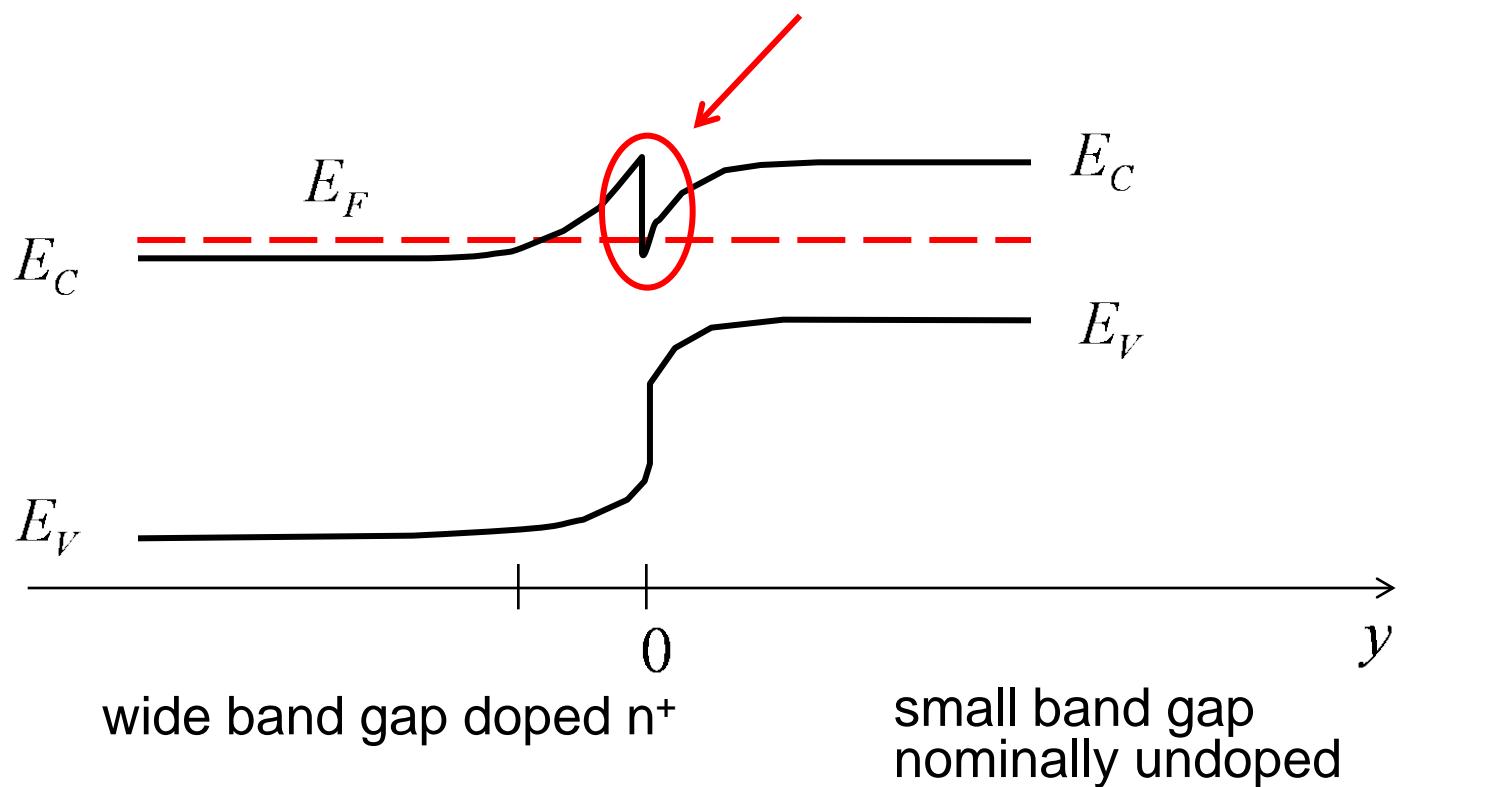
- high mobility
 $\mu_n(10^{14}) \sim 8500 \text{ cm}^2/\text{V}\cdot\text{s}$
- mobility and doping
 $\mu_n(10^{17}) \sim 4700 \text{ cm}^2/\text{V}\cdot\text{s}$
 $\mu_n(10^{18}) \sim 2800 \text{ cm}^2/\text{V}\cdot\text{s}$
- for high g_m , need both charge **and** velocity
- SB gate limits V_G

“Modulation doping”

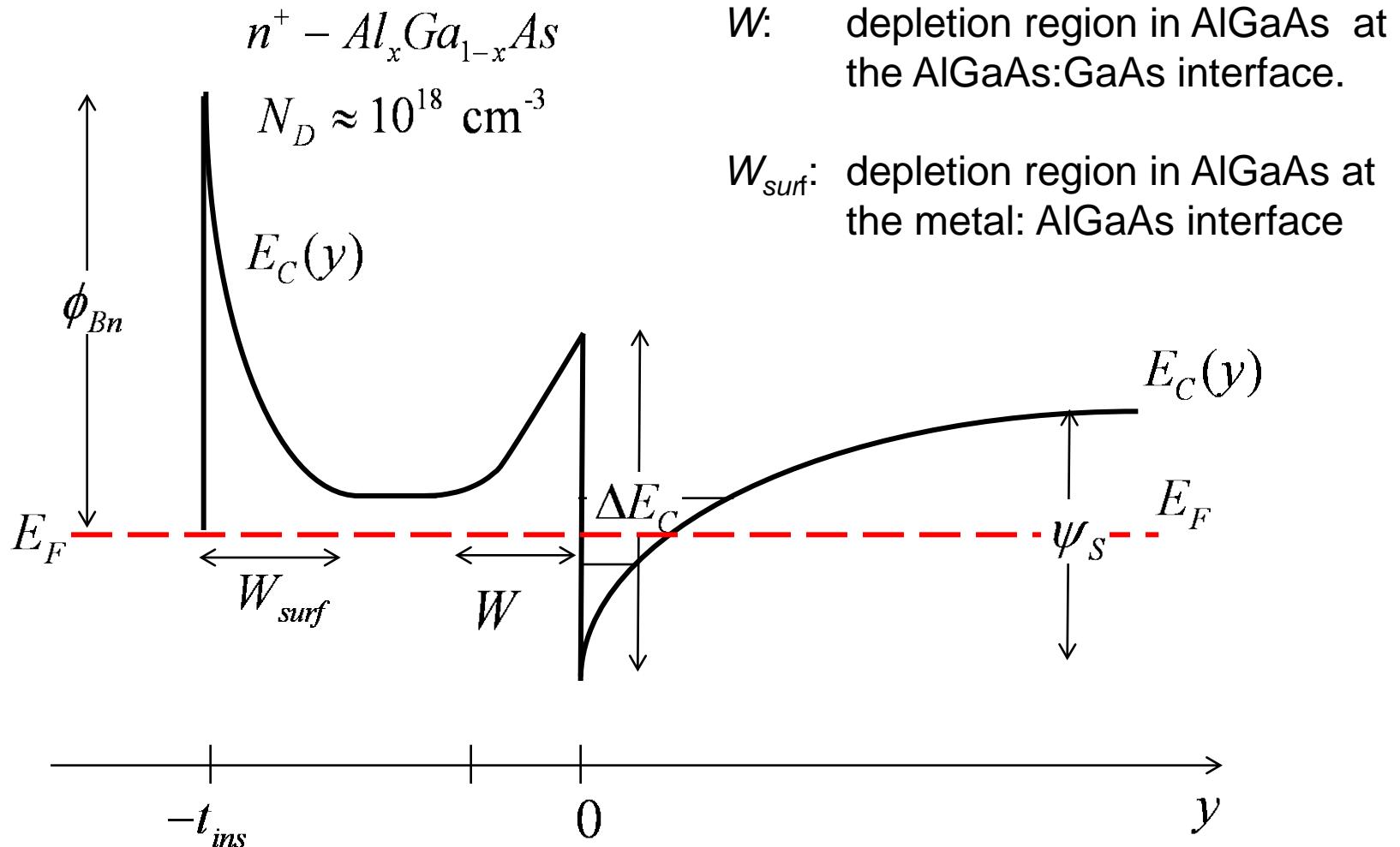


Modulation doping

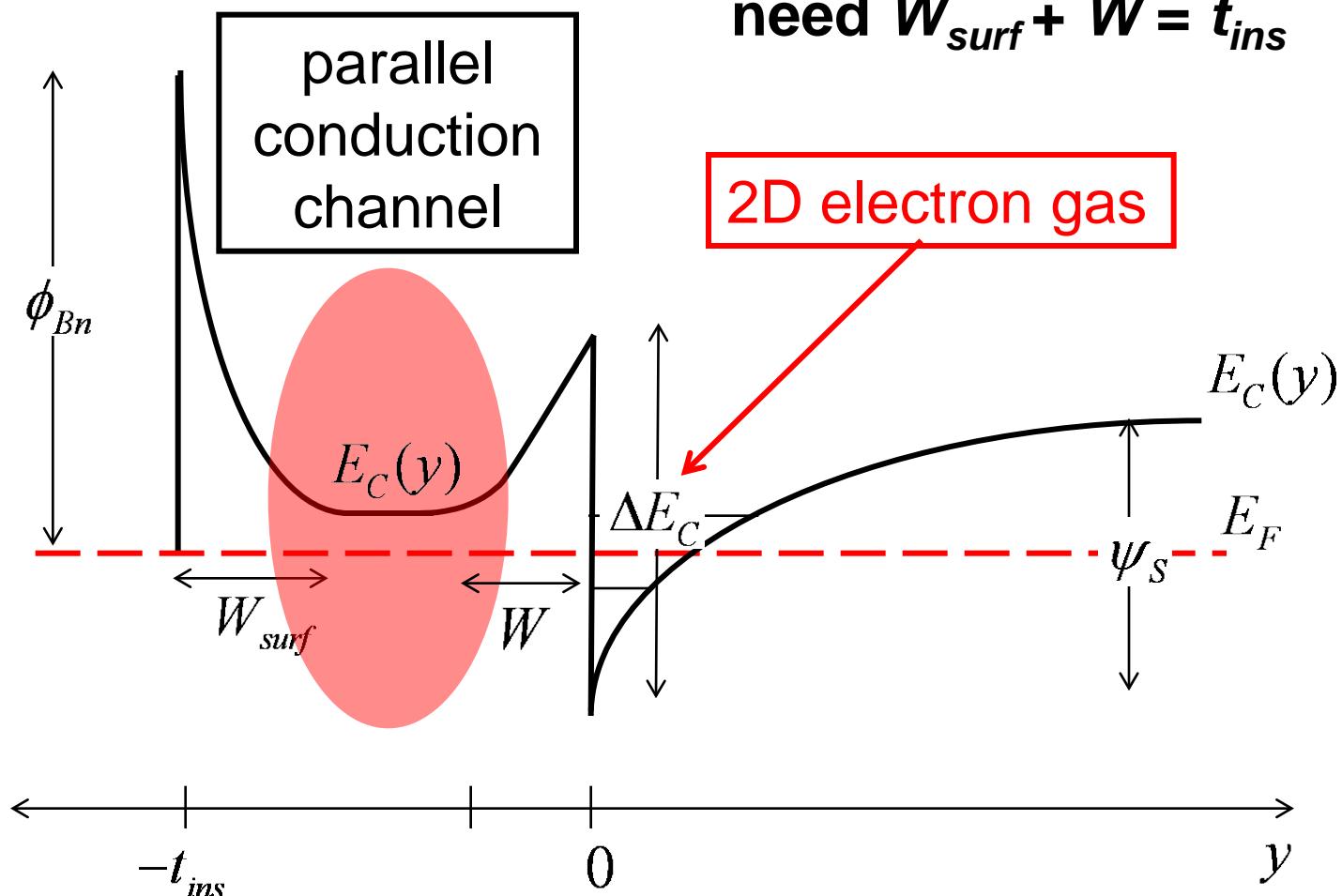
**“2D electron gas”
(high carrier density *and* high μ_n)**



Equilibrium energy band diagram



Parallel conduction should be avoided

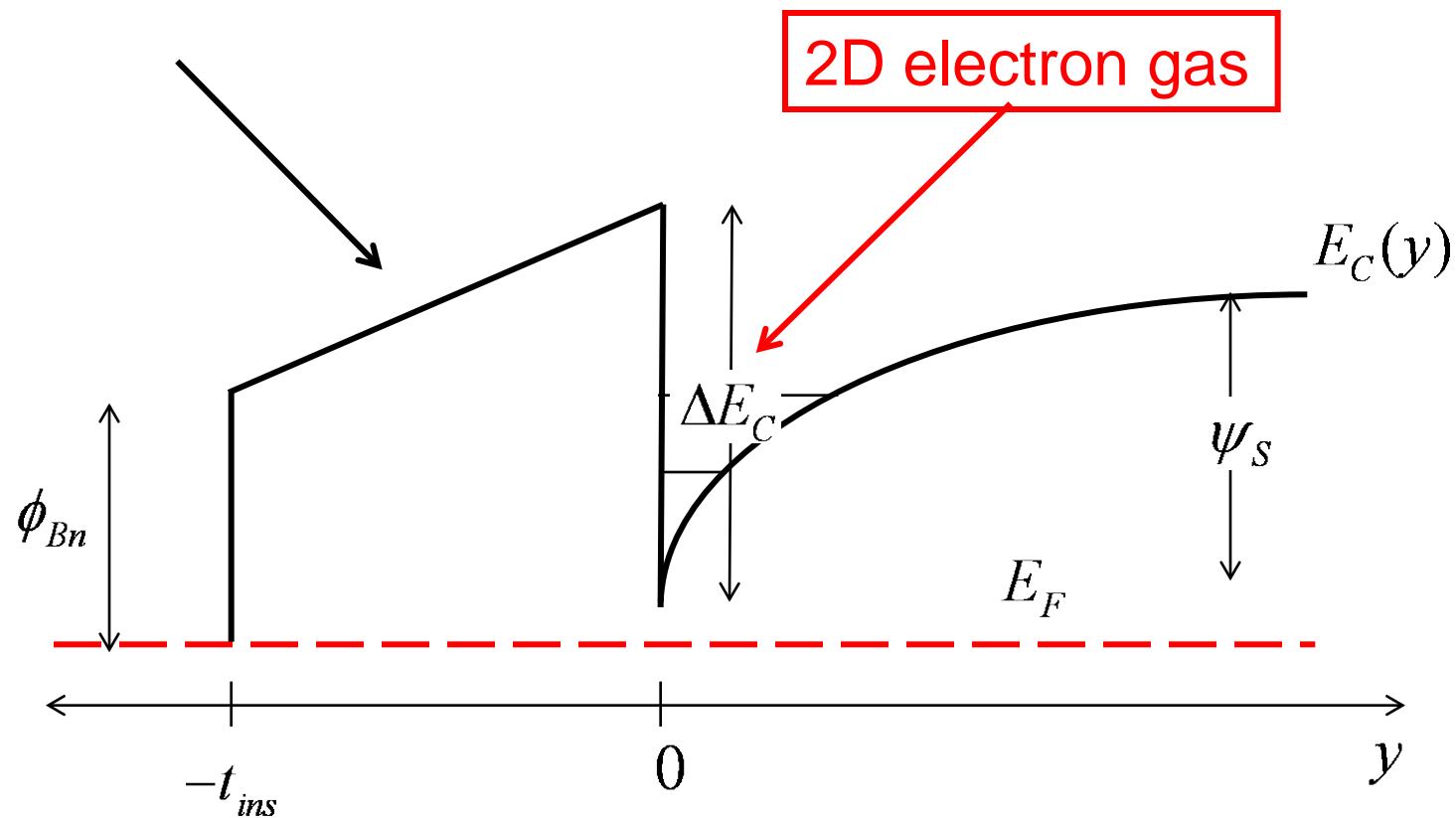


IF WE ARE NOT CAREFUL THERE WILL UNDEPLETED LAYER IN THE RED ROUNDER REIGN WHICH IS NOT DESIRABLE

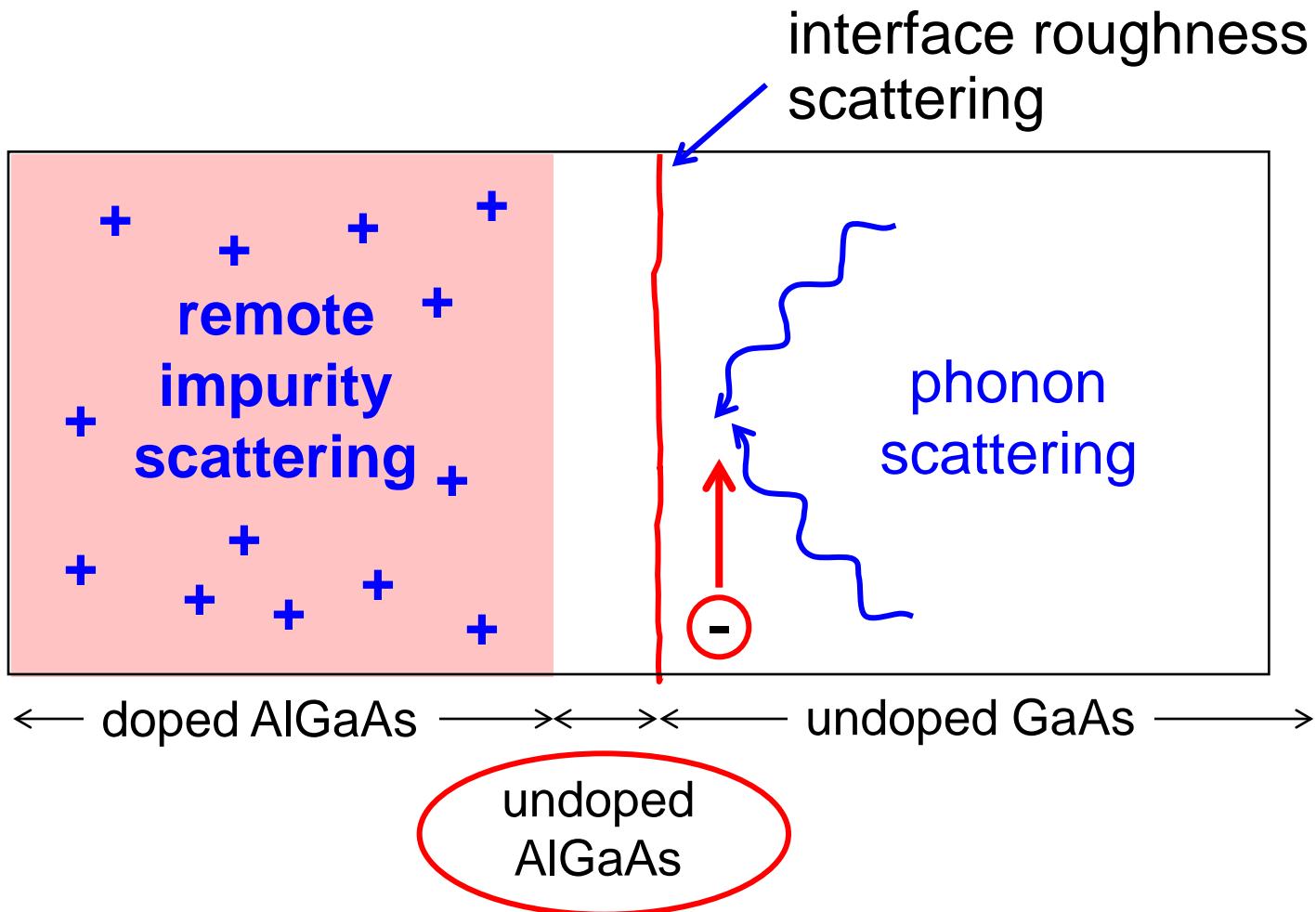
Lundstrom: 2018 SO THERE MUST NOT BE UNDEPLETED PARALLEL CONDUCTION PATH

Why dope the wide bandgap layer?

“like a gate insulator”

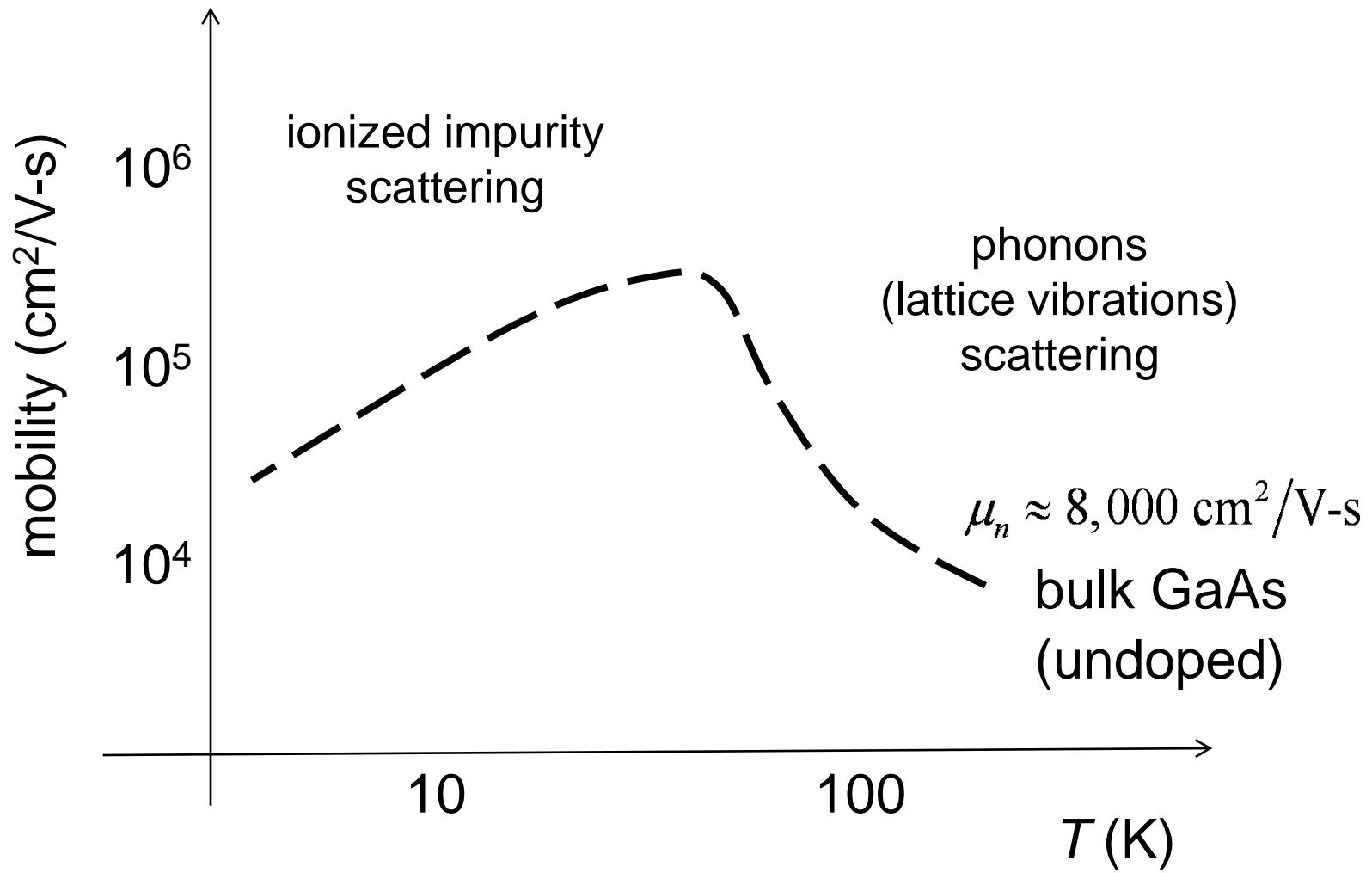


Scattering mechanisms (mobility)



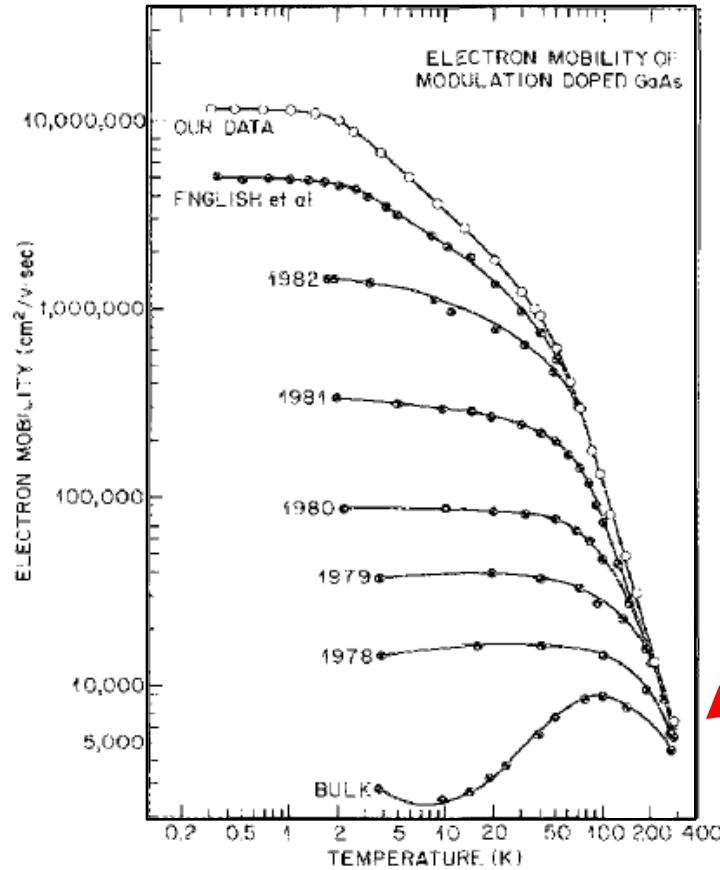
(after Solomon and Morkoc)

Mobility vs. temperature



Mobility vs. temperature (modulation doped)

$\mu_n > 10^7 \text{ cm}^2/\text{V-s} !!$

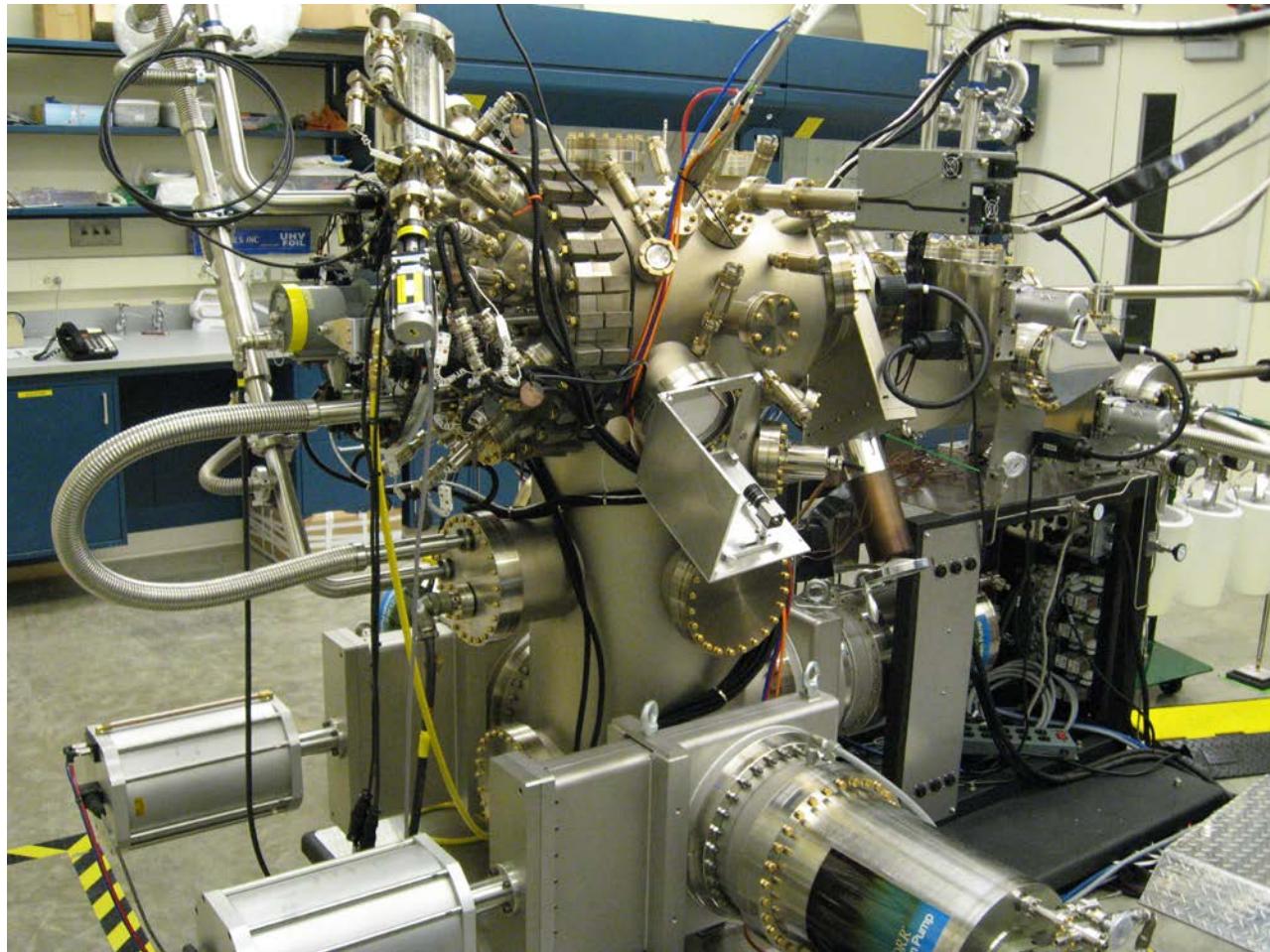


$\mu_n \approx 10^4 \text{ cm}^2/\text{V-s}$
300 K



L. Pfeiffer, K.W. West, H.L. Stormer, and K.W. Baldwin, "Electron mobilities exceeding $10^7 \text{ cm}^2/\text{V s}$ in modulation-doped GaAs," *Appl. Phys. Lett.*, **55**, 1888, 1989. 11

Molecular beam epitaxy



Michael Manfra Lab, Birck Nanotechnology Center, Purdue University

From physics to technology

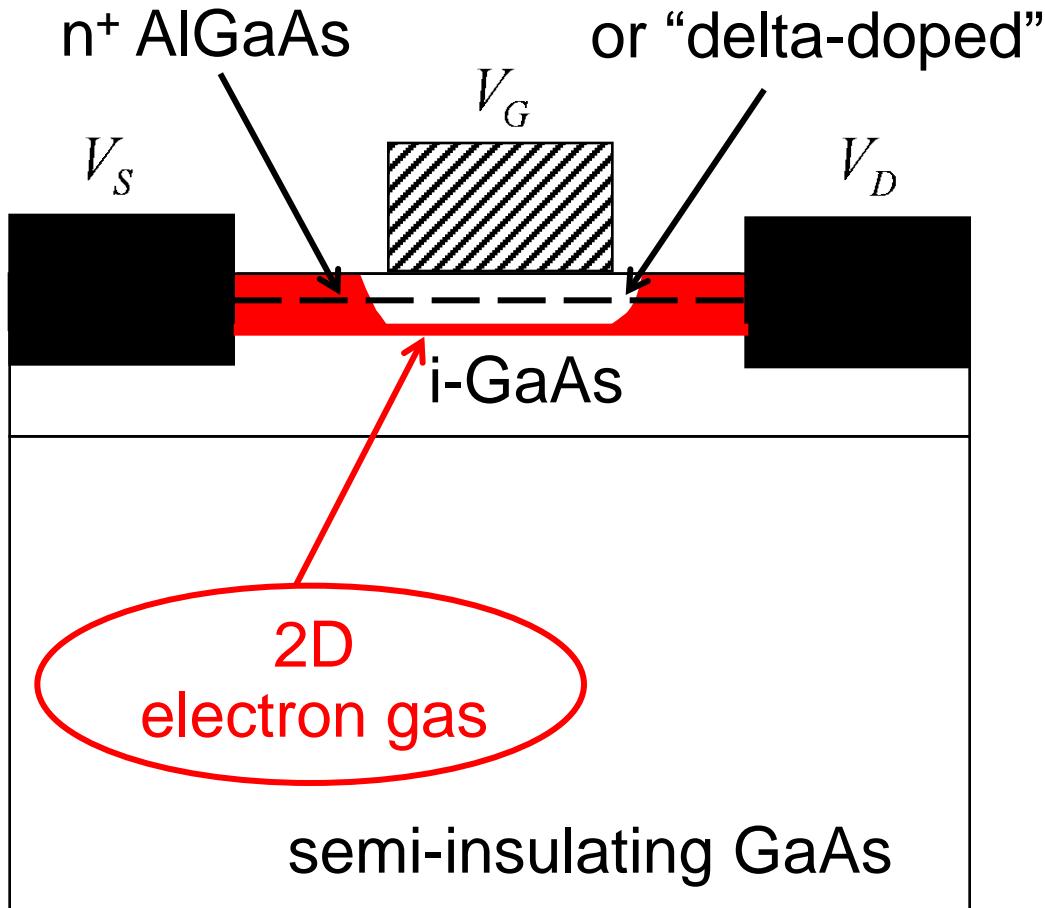
Modulation doping was discovered in 1978 by Dingle. Two years later, the High Electron Mobility Transistor (HEMT) was invented.

T. Mimura et al., "A New Field-Effect Transistor with Selectively Doped GaAs/n-Al_xGa_{1-x}As Heterojunctions," *Jpn. J. Appl. Phys.* **19**, L225 (1980)

For an interesting perspective on this important device, see:

s A. del Alamo, "The High Electron Mobility Transistor at 30: Impressive Accomplishments and Exciting Prospects," Proc. of the 2011 Int. Conf. on Compound Semiconductor Manufacturing Technology (2011) <http://hdl.handle.net/1721.1/87102>

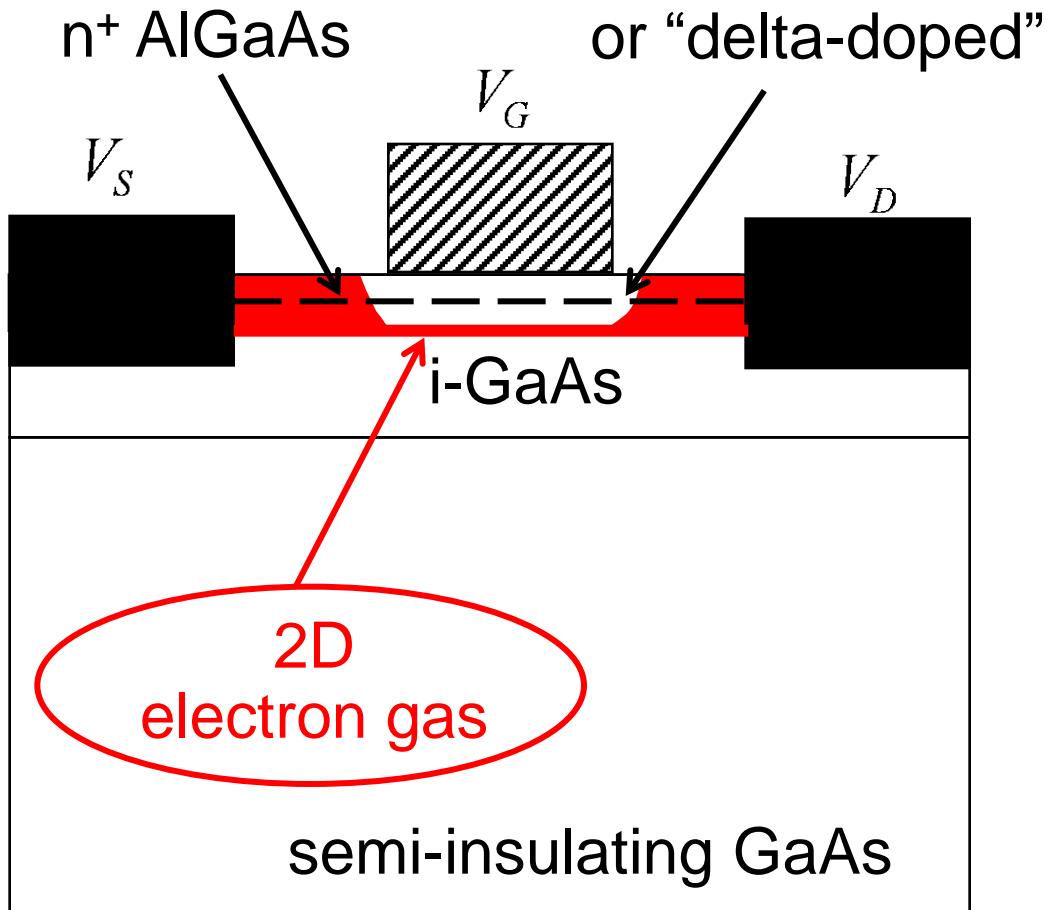
Heterostructure FET



Note:

InP / InAlAs more common now

Why delta doping?



- 1) Higher channel charge results
- 2) Higher gate breakdown voltage results
- 3) Gate electrode is closer to channel:
 - suppresses 2D effects
 - higher g_m

Names

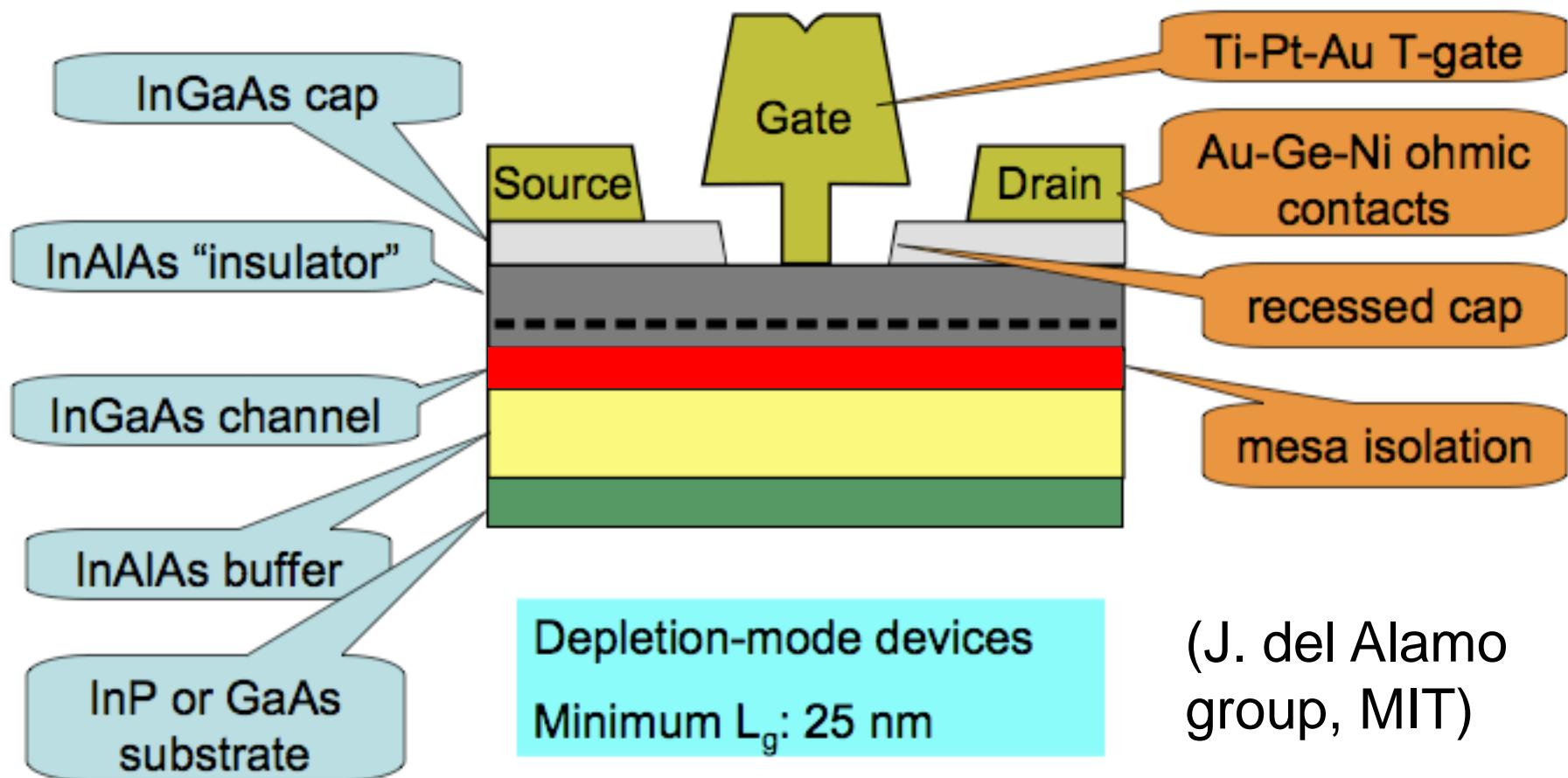
HEMT: “**H**igh **E**lectron **M**obility **T**ransistor”

MODFET: “**M**odulation-**D**oped **F**ield-**E**ffect **T**ransistor”

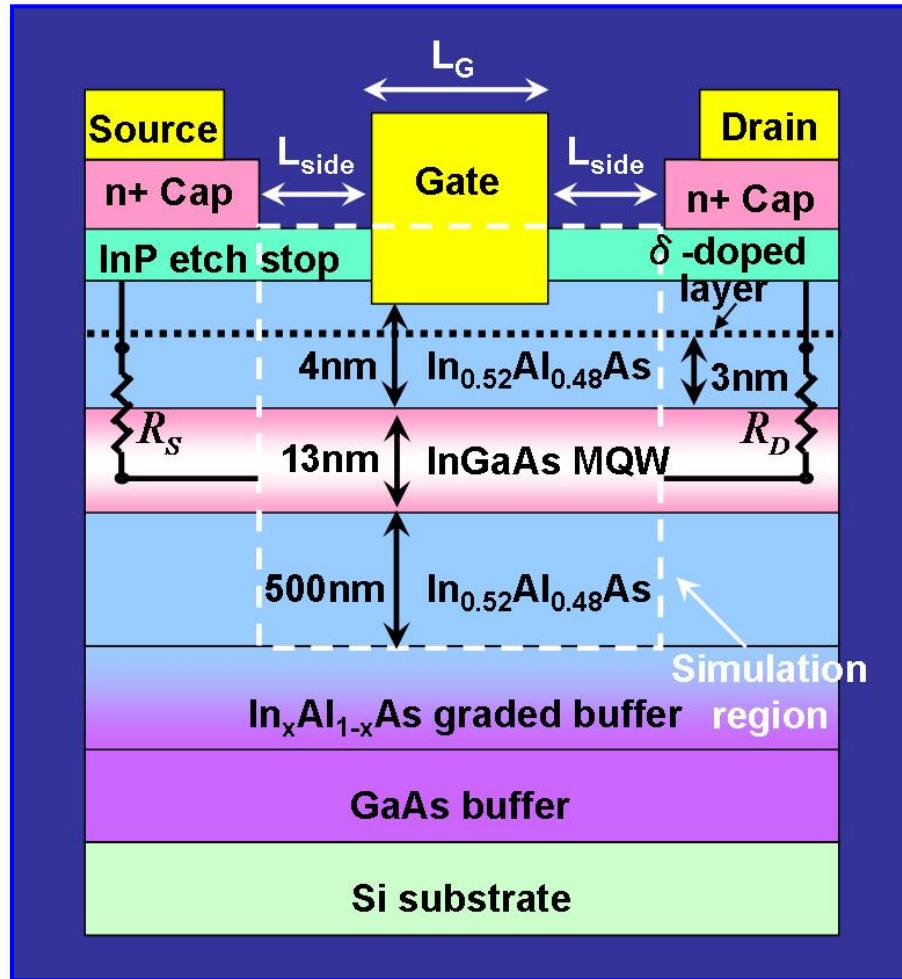
SDHT: “**S**electively-**D**oped **H**eterostructure **T**ransistor”

TEGFET: “**T**wo-dimensional **E**lectron **G**as **F**ield-**E**ffect
Transistor”

InGaAs HEMT



Layer structure

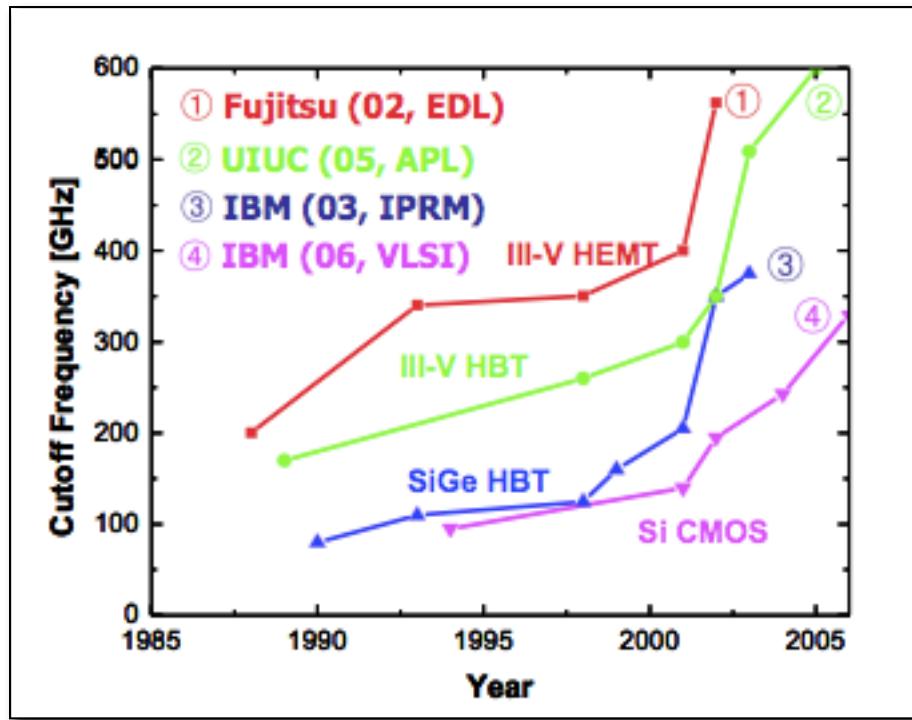


Jesus del Alamo group (MIT)

Applications

- 1) Initially driven by high-speed logic
- 2) **Low noise amplifiers (micro/millimeter waves)**
 - satellite communication, radio astronomy, electronic warfare
 - cell phones
- 3) **Millimeter-wave power amplifiers**

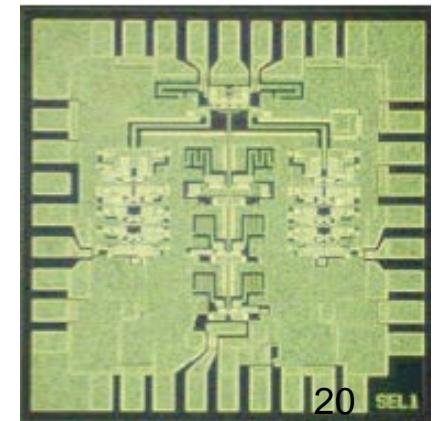
InGaAs HEMT technology



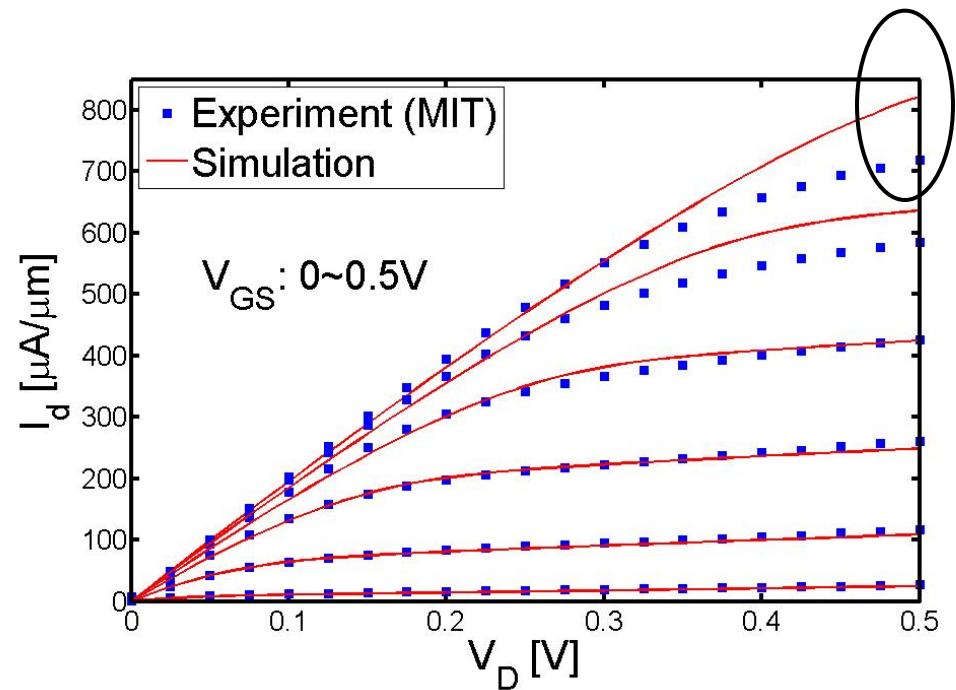
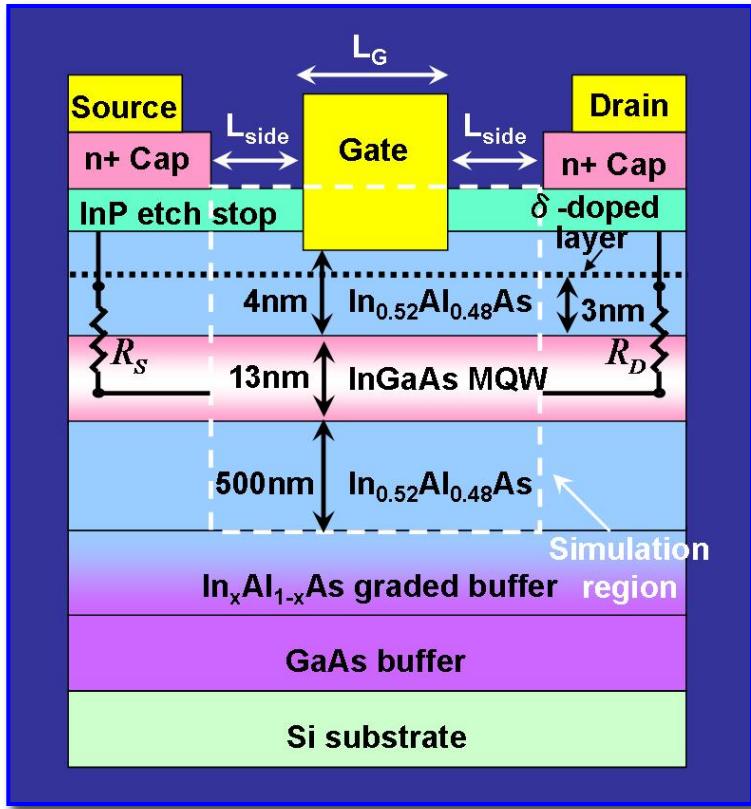
Fairly mature technology at SSI level:

- 120 Gb/s MUX (NEC, 2004)
- 110 Gb/s DEMUX (NEC, 2004)
- 140 Gb/s Selector (Fujitsu, 2004)
- 160-215 GHz Amp (TRW, 2002)
- Space qualified

100Gb/s selector IC
(NTT 2003)



Comparison with experiment: InGaAs HEMTs



Jesus del Alamo group (MIT)

near-ballistic operation

III-V MOSFETs

There has recently been significant progress in III-V MOSFETs. For a review of the current state of the field, see:

J. A. del Alamo, D. A. Antoniadis, J. Lin, W. Lu,
A. Vardi, and X. Zhao, “Nanometer-Scale III-V MOSFETs,”
J. Electron Devices Society, vol. 4, pp. 205-214, 2016.

Summary

- 1) III-V FETs are an important technology for high-frequency RF applications.
- 2) Both HEMTs and HBTs have achieved THz speeds.
- 3) HEMTs operate in exactly the same “barrier controlled mode” as Si MOSFETs, so the VS model describes them well.
- 4) III-V HEMTs operate near the ballistic limit.

Thanks to Profs. J. del Alamo (MIT), Mark Rodwell (UCSB), Peide Ye and Mike Manfra (Purdue) for their help in putting together this lecture.

Next topic

The very first transistors (Bell Labs, 1947), were bipolar transistors.

Bipolar transistors are also barrier controlled transistors.

Modern bipolar transistors (heterostructure bipolar transistors) also have important applications in RF.

Essentials of MOSFETs

Unit 5: Additional Topics

Lecture 5.4: Review of PN Junctions

Mark Lundstrom

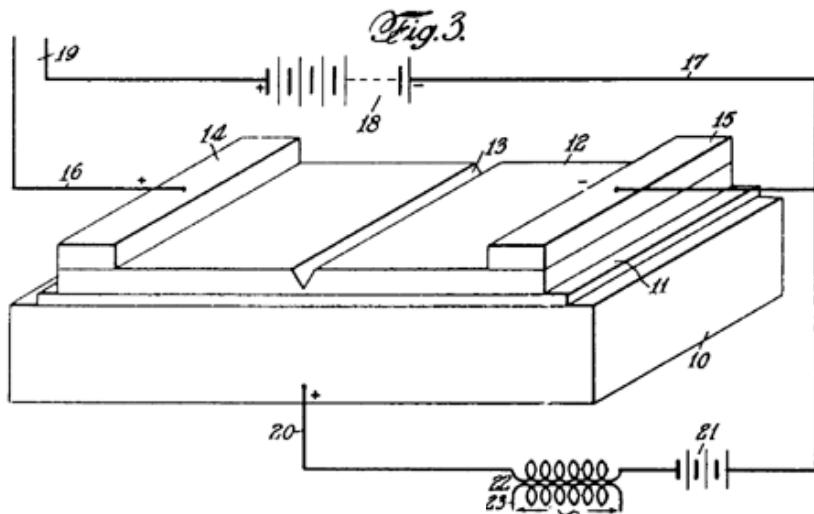
lundstro@purdue.edu
Electrical and Computer Engineering
Purdue University
West Lafayette, Indiana USA

Transistors

MOSFET	HFET	RBT
MOST	DHFET	RHET
IGFET	HIGFET	QWBRTT
DMOS	SISFET	TETRAN
HEXFET	PBT	SIT
VMOS	LRTFET	NWFET
TFT	VMT	CNT FET
MISFET	BJT	SB FET
JFET	HBT	BTBT FET
VFET	DHBT	induced base transistor
MESFET	THETA	planar doped barrier transistor
MOSFET	RST	metal base transistor
HEMT	BICFET	Stark-effect transistor
TEGFET	RTBT	delta-doped channel heterojunction FET

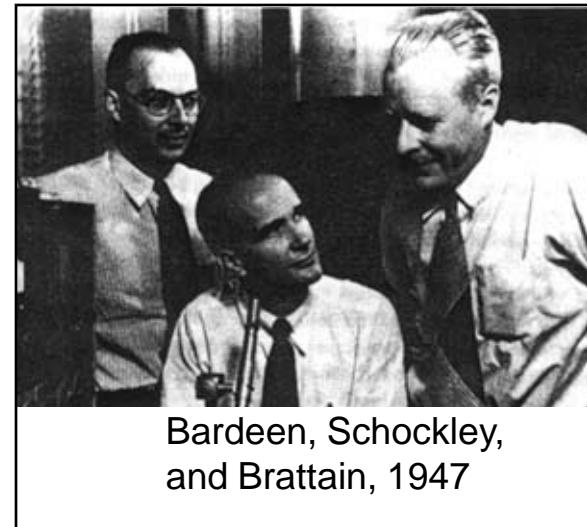
***Bipolar
transistors are
barrier-
controlled
transistors too.***

Invention (and discovery) of the transistor

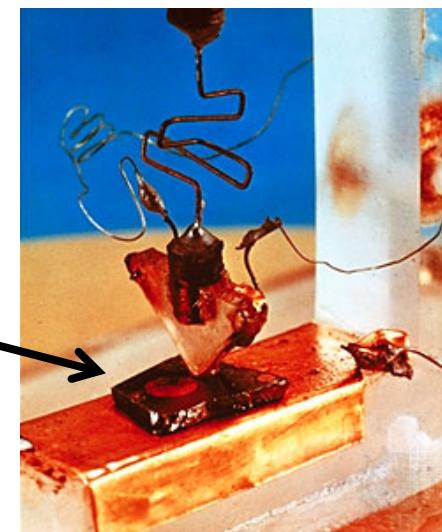


Field-Effect Transistor
Lillienfield, 1925
Heil, 1935

“base”



Bardeen, Shockley,
and Brattain, 1947

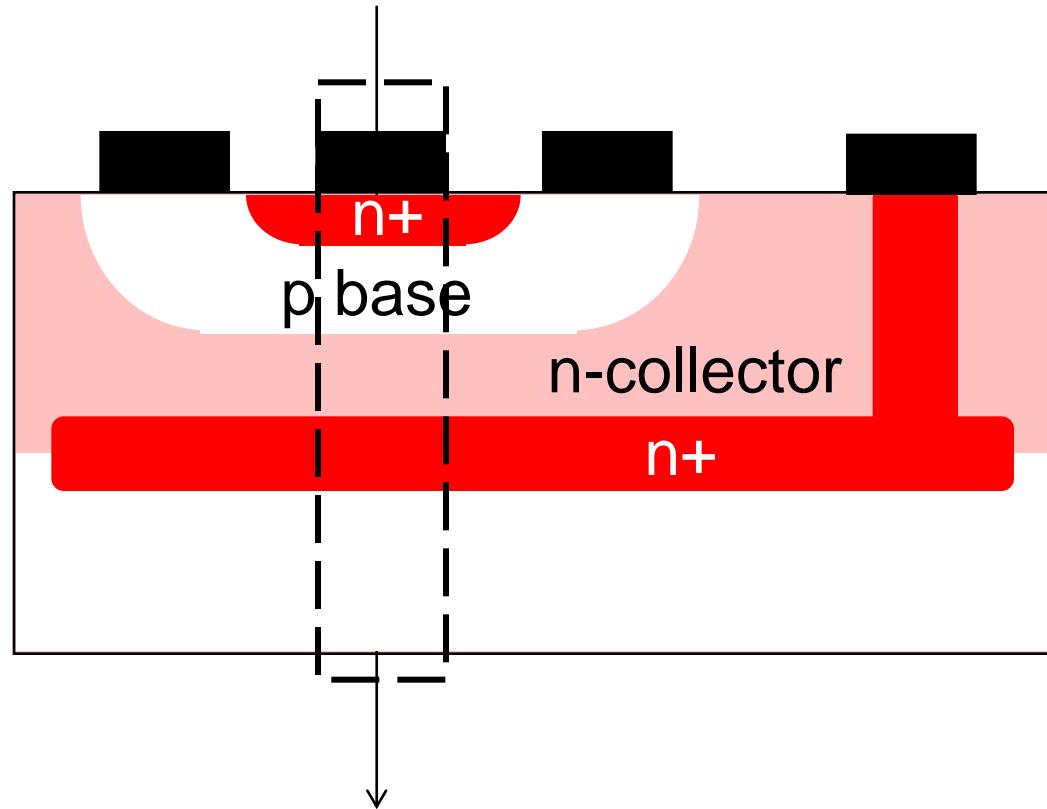


Lundstrom: 2018

Lucent / Bell Labs

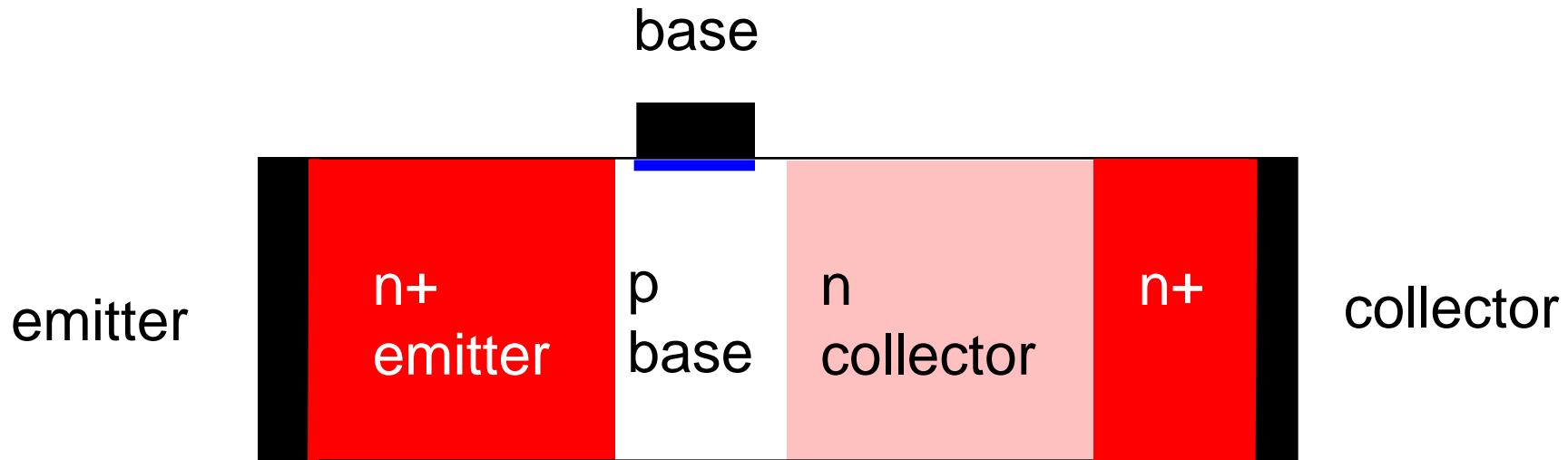
Bipolar transistors

Double diffused Bipolar Junction Transistor (BJT)



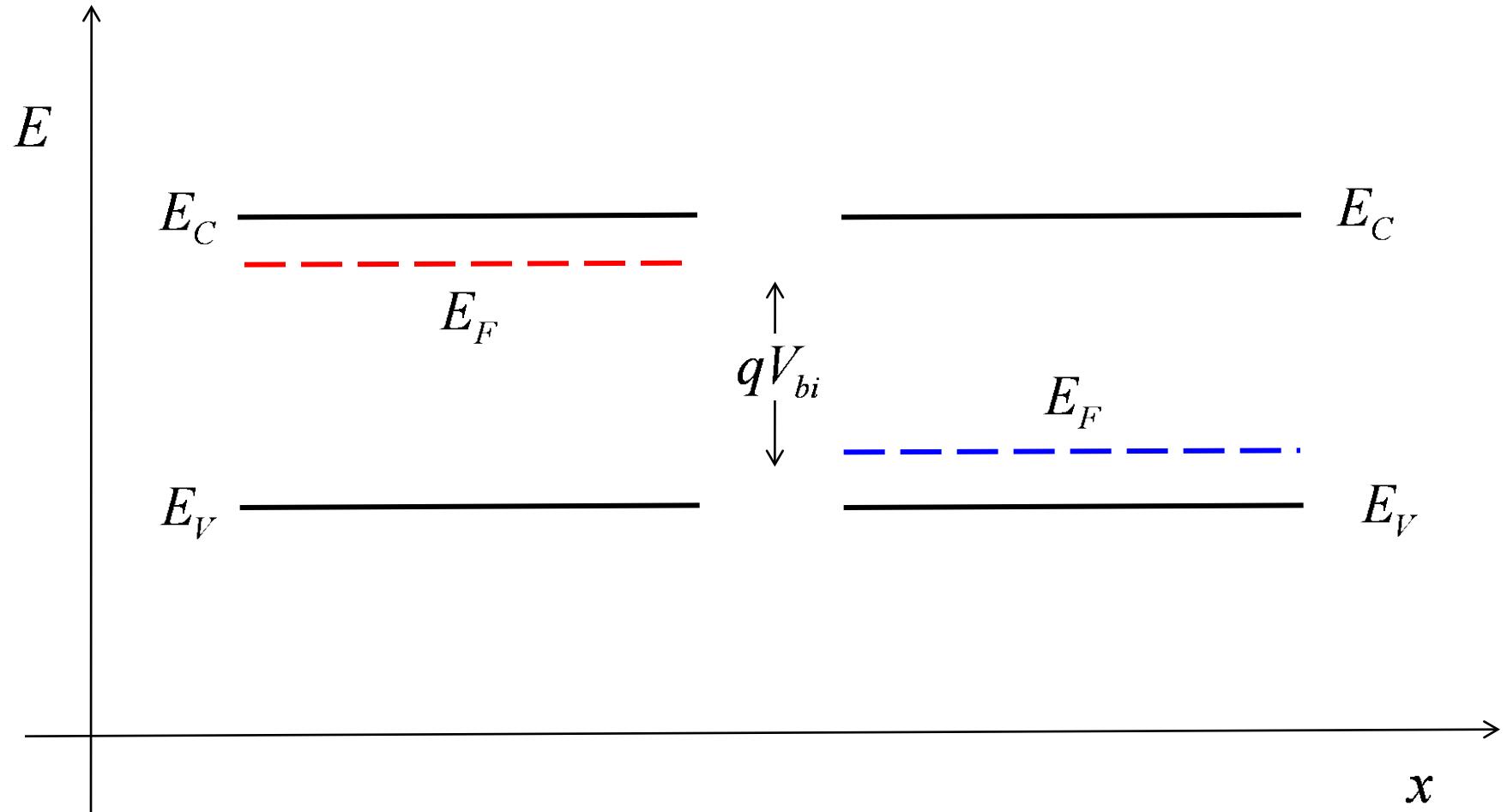
“ideal” 1D BJT

Bipolar transistors

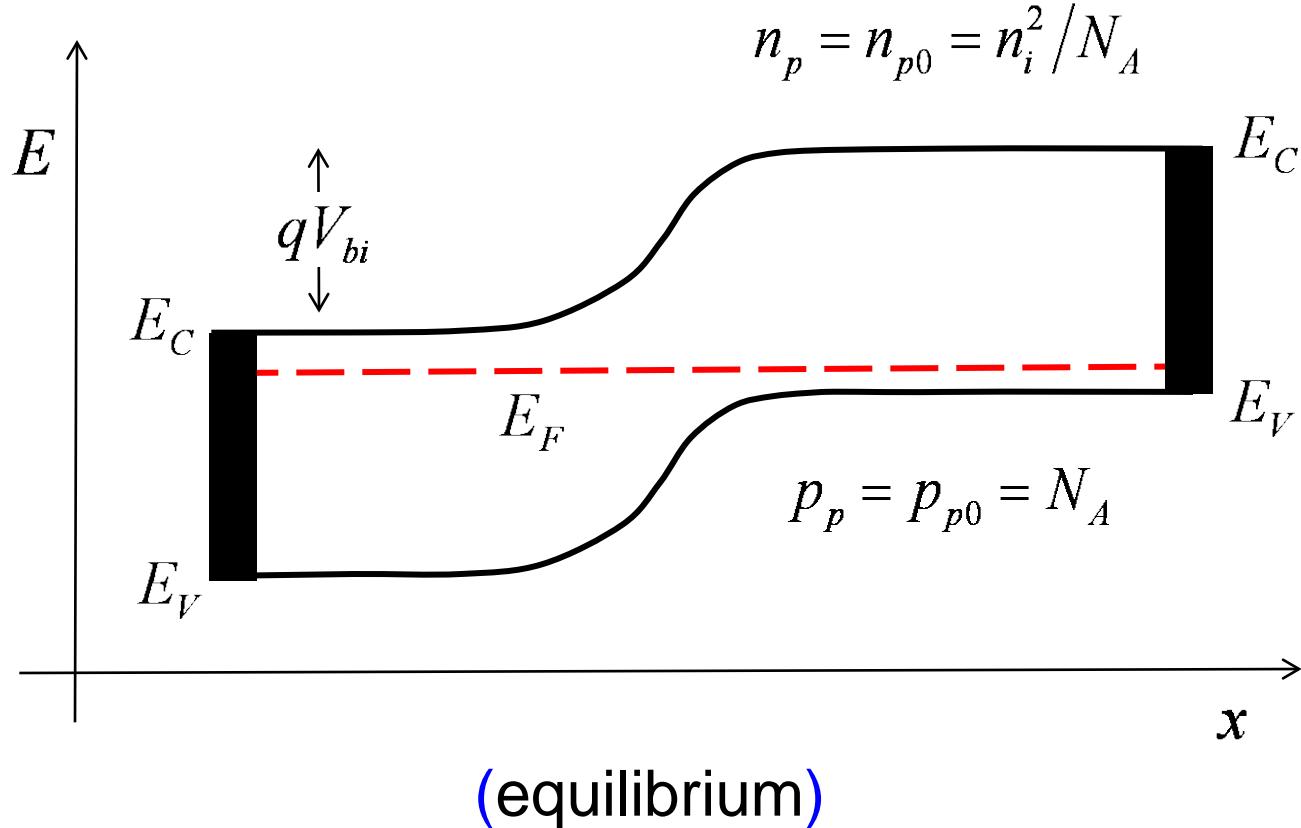


Just as the MOS-C is the heart of a MOSFET, BJT's are made up of a basic building block: The PN junction.

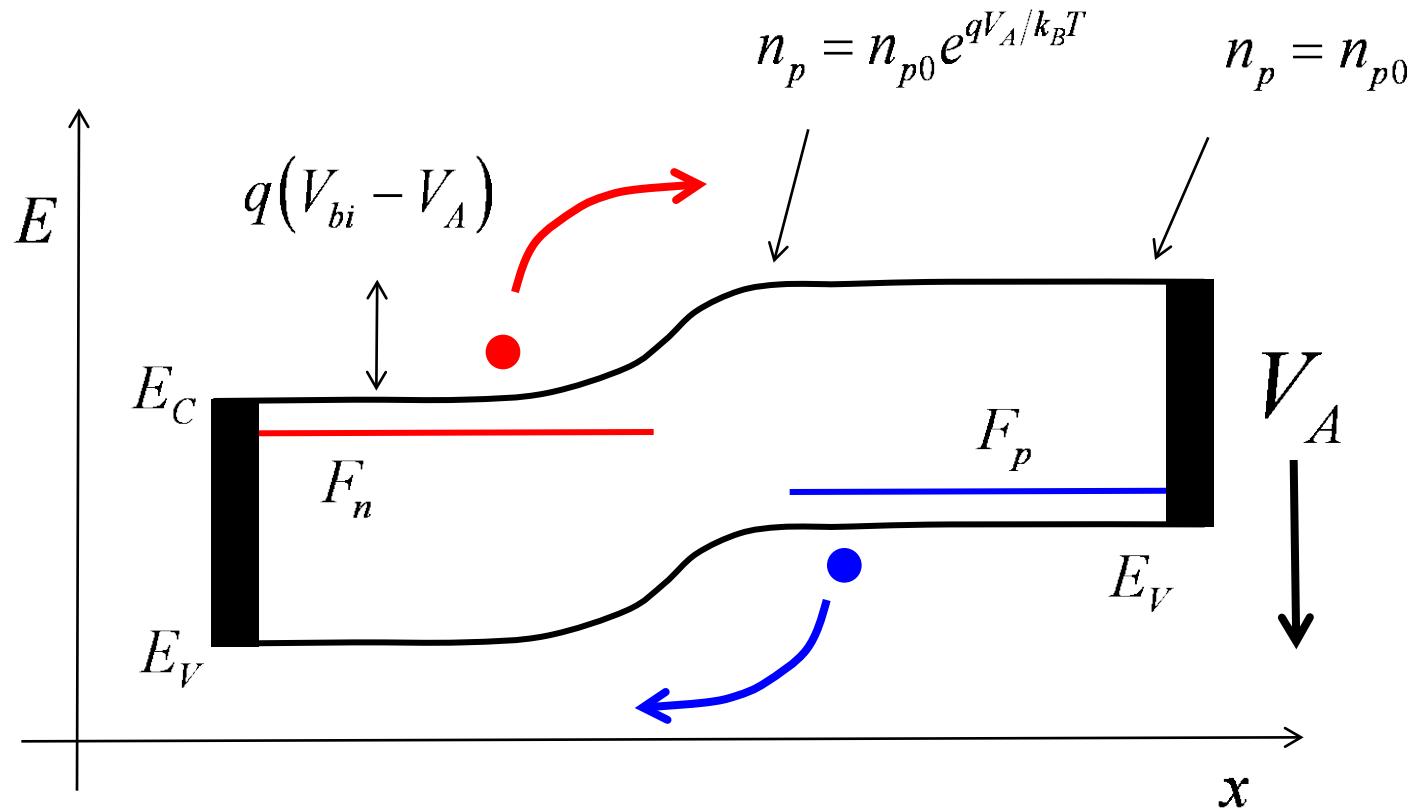
Primer on PN junctions



Equilibrium energy band diagram

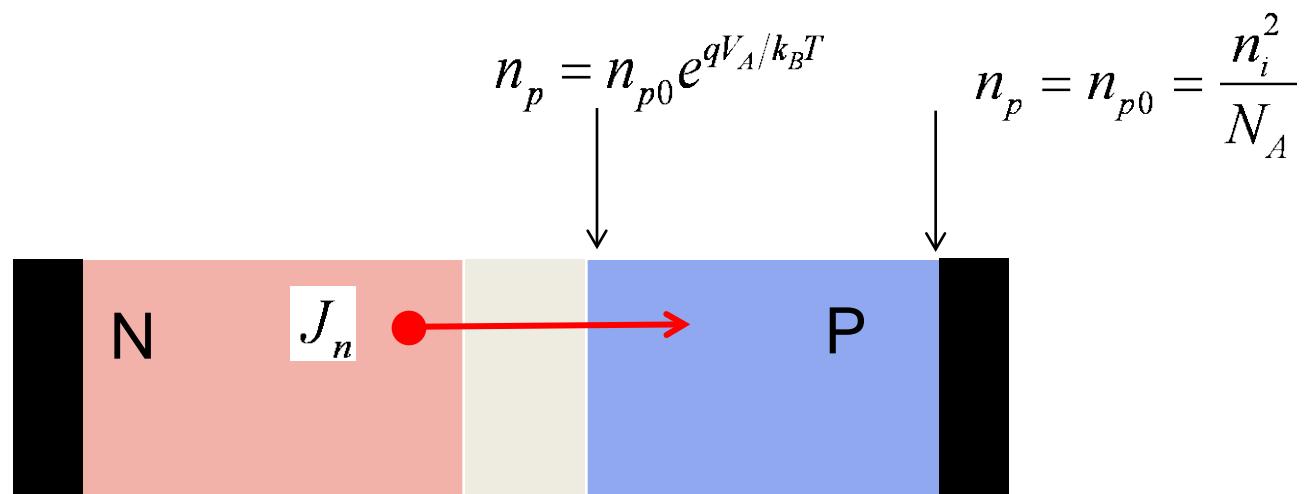


Forward bias



(forward bias)

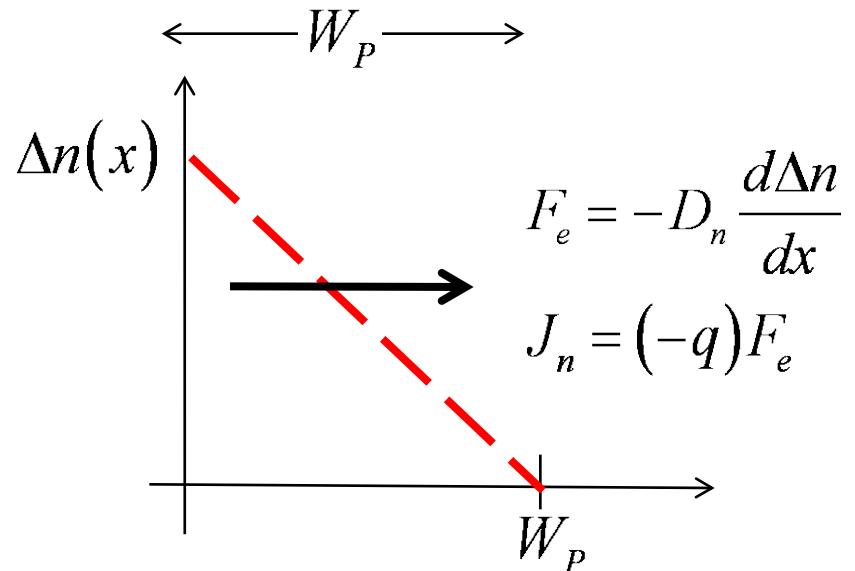
Electron current in a FB PN junction



$$\Delta n_p(0) = n_{p0} \left(e^{qV_A/k_B T} - 1 \right)$$

$$J_n = q n_{p0} \frac{D_n}{W_P} \left(e^{qV_A/k_B T} - 1 \right)$$

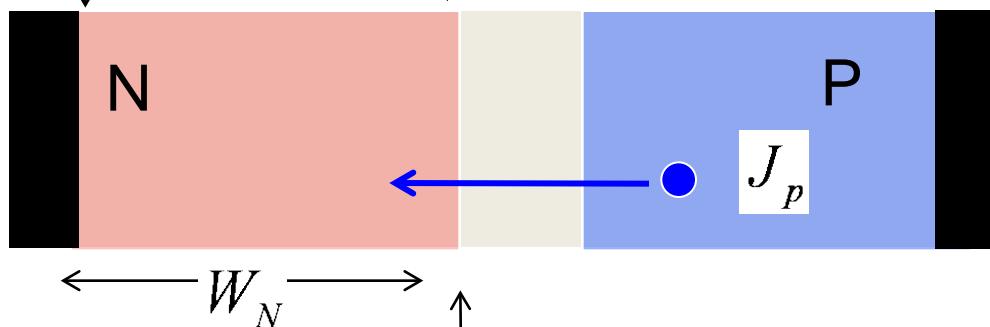
(magnitude of current)



Hole current in a FB PN junction

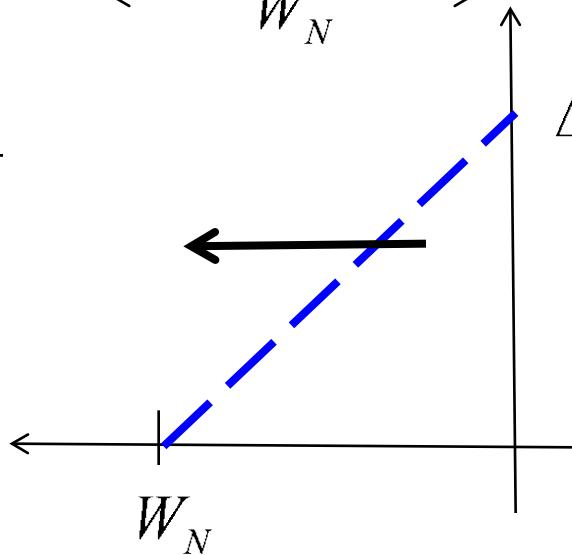
$$p_n = p_{n0} = \frac{n_i^2}{N_D}$$

$$p_n = p_{n0} e^{qV_A/k_B T}$$



$$F_h = -D_p \frac{d\Delta p}{dx}$$

$$J_p = (+q) F_h$$

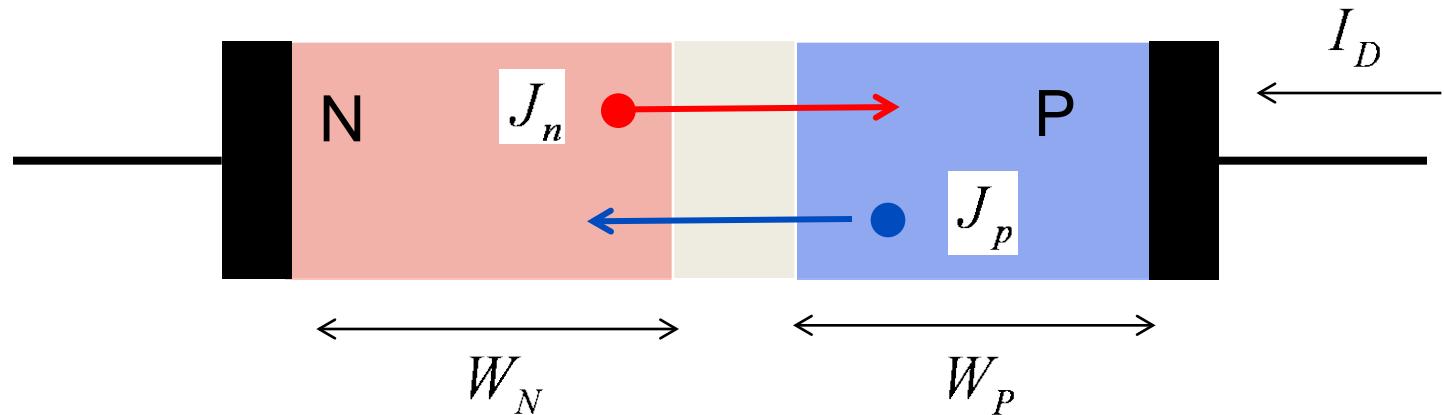


$$\Delta p_n(0) = p_{n0} (e^{qV_A/k_B T} - 1)$$

$$J_p = q p_{n0} \frac{D_p}{W_N} (e^{qV_A/k_B T} - 1)$$

(magnitude of current) 10

Current in a FB PN junction

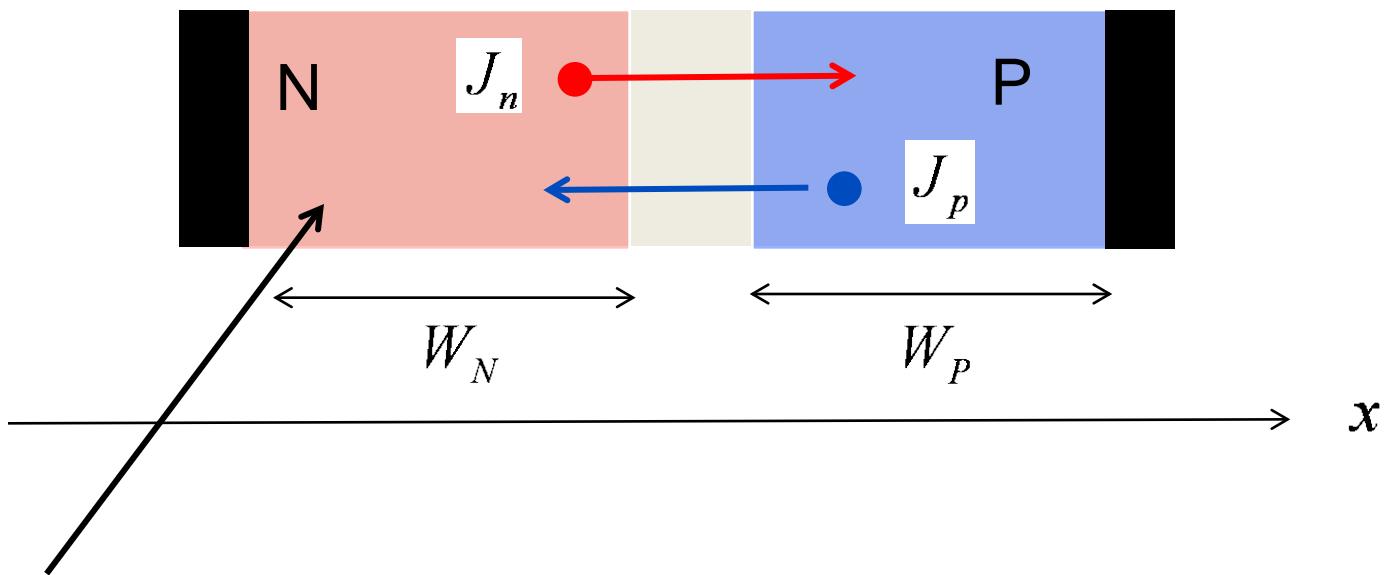


$$J_n = q \frac{D_n}{W_P} \frac{n_i^2}{N_A} \left(e^{qV_A/k_B T} - 1 \right)$$

$$J_p = q \frac{D_p}{W_N} \frac{n_i^2}{N_D} \left(e^{qV_A/k_B T} - 1 \right)$$

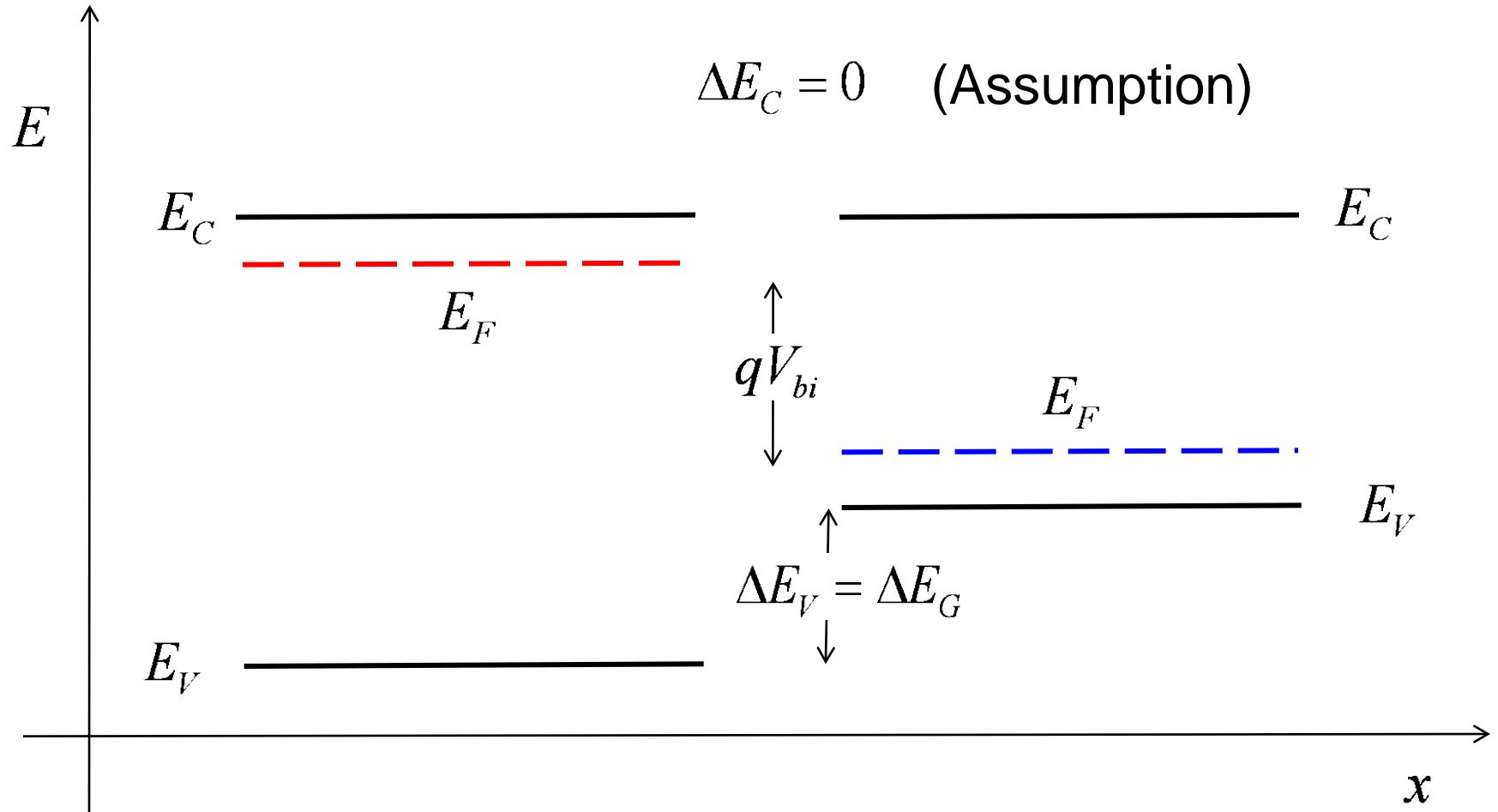
$$I_D = A_D (J_n + J_p)$$

Heterojunction diodes

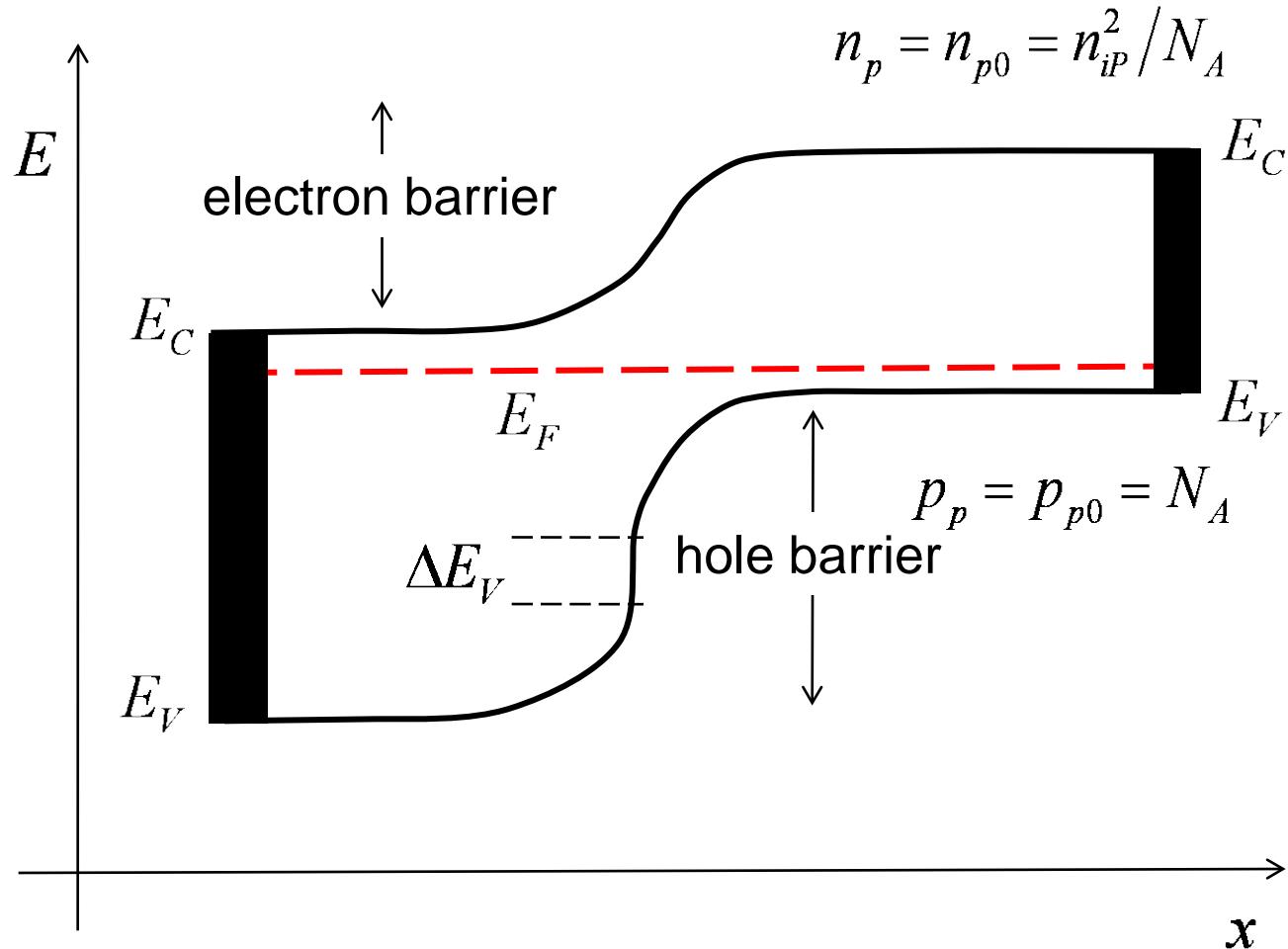


$$E_{GN} > E_{GP}$$

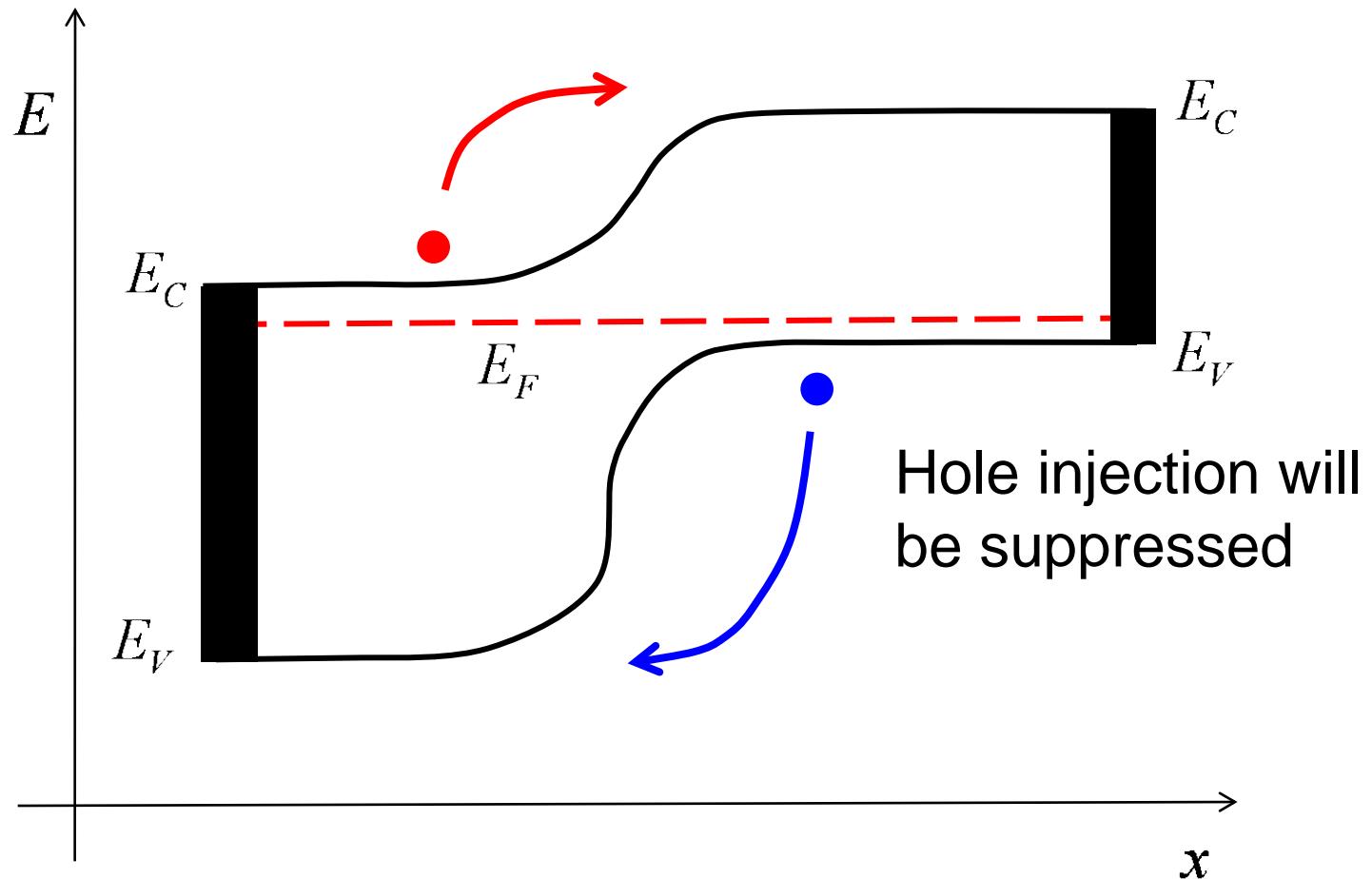
Np heterojunction



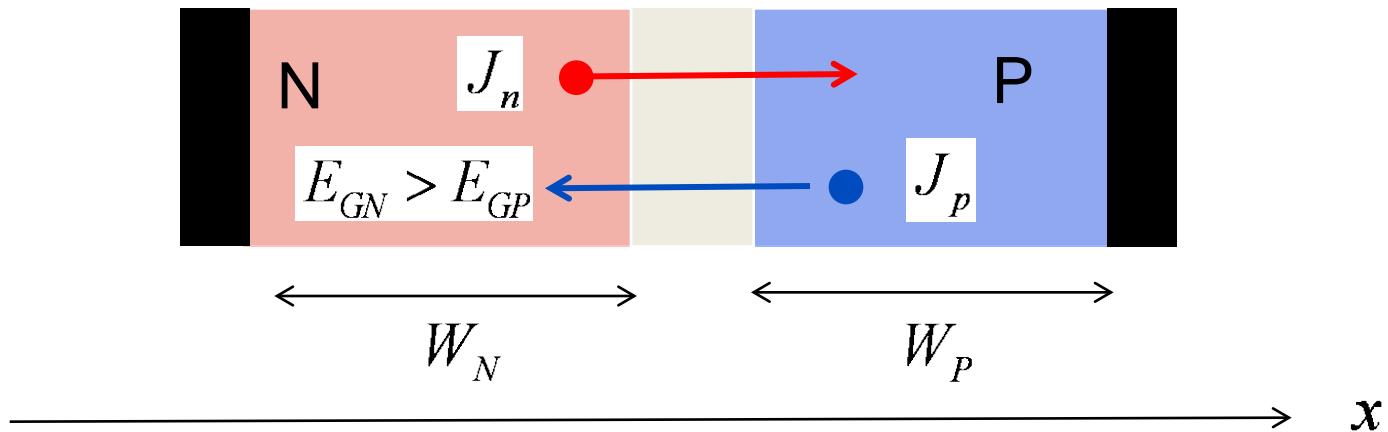
Equilibrium energy band diagram



Equilibrium energy band diagram



Current in a FB Np heterojunction



$$J_n = q \frac{D_n}{W_P} \frac{n_{iP}^2}{N_A} \left(e^{qV_A/k_B T} - 1 \right)$$

$$J_p = q \frac{D_p}{W_N} \frac{n_{iN}^2}{N_D} \left(e^{qV_A/k_B T} - 1 \right)$$

$$n_i^2 \propto e^{-E_G/k_B T}$$

$$E_{GN} > E_{GP}$$

$$n_{iN}^2 \ll n_{iP}^2$$

$$J_p \ll J_n$$

Summary

- 1) PN diodes are bipolar devices – current is due to electrons and holes.
- 2) A forward bias lowers the barrier for electron and hole injection (i.e. it is a barrier-controlled device).
- 3) The current increases exponentially with forward bias (a 60 mV increase in FB increases I_D by 10X).
- 4) It is possible to suppress one of the two current components by using heterojunctions.

Next topic

An NPN bipolar transistor consists of two PN junctions with a shared P region.

Since we understand PN junctions, we can understand BJTs.

That is our topic for the next lecture.

Essentials of MOSFETs

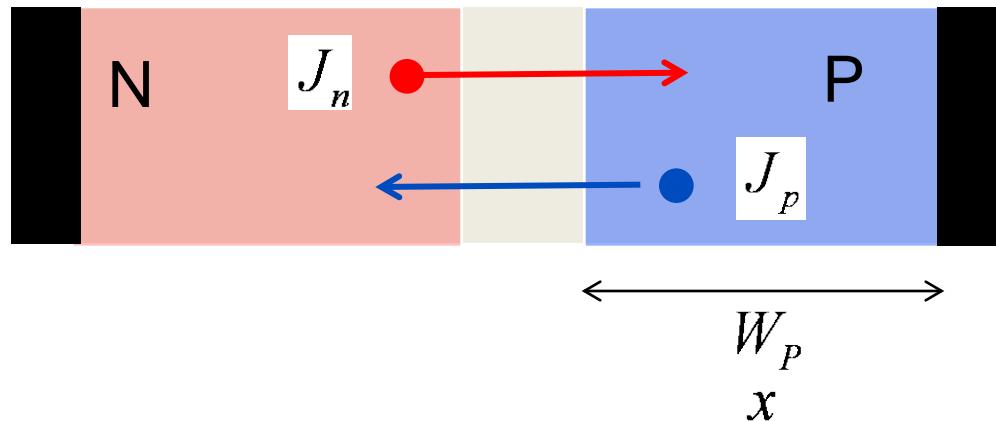
**Unit 5:
Additional Topics
Lecture 5.5:
Heterostructure
Bipolar Transistors (HBTs)**

Mark Lundstrom

lundstro@purdue.edu
Electrical and Computer Engineering
Purdue University
West Lafayette, Indiana USA

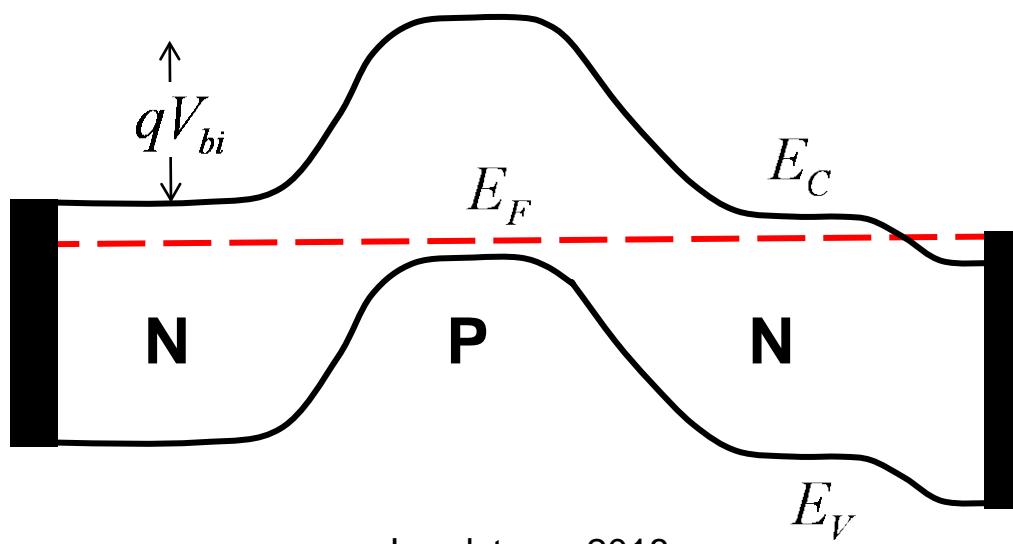
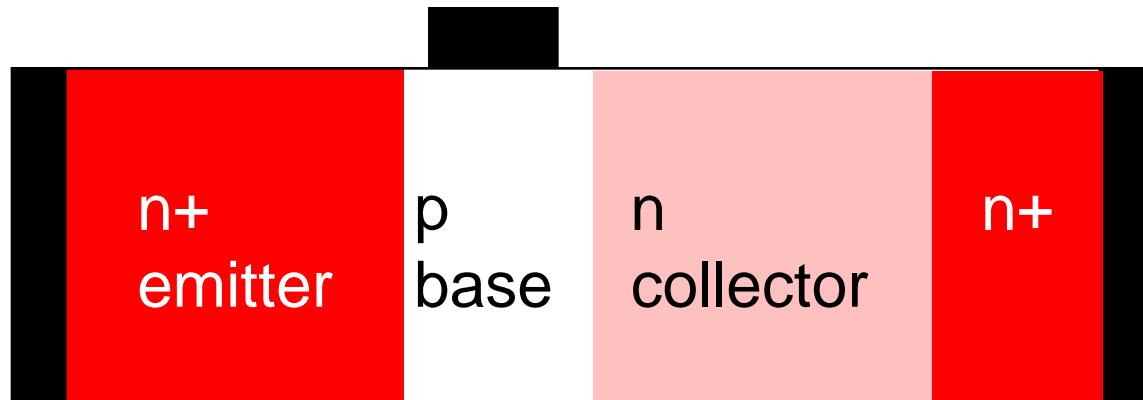
Lundstrom: 2018

Currents in a PN junction



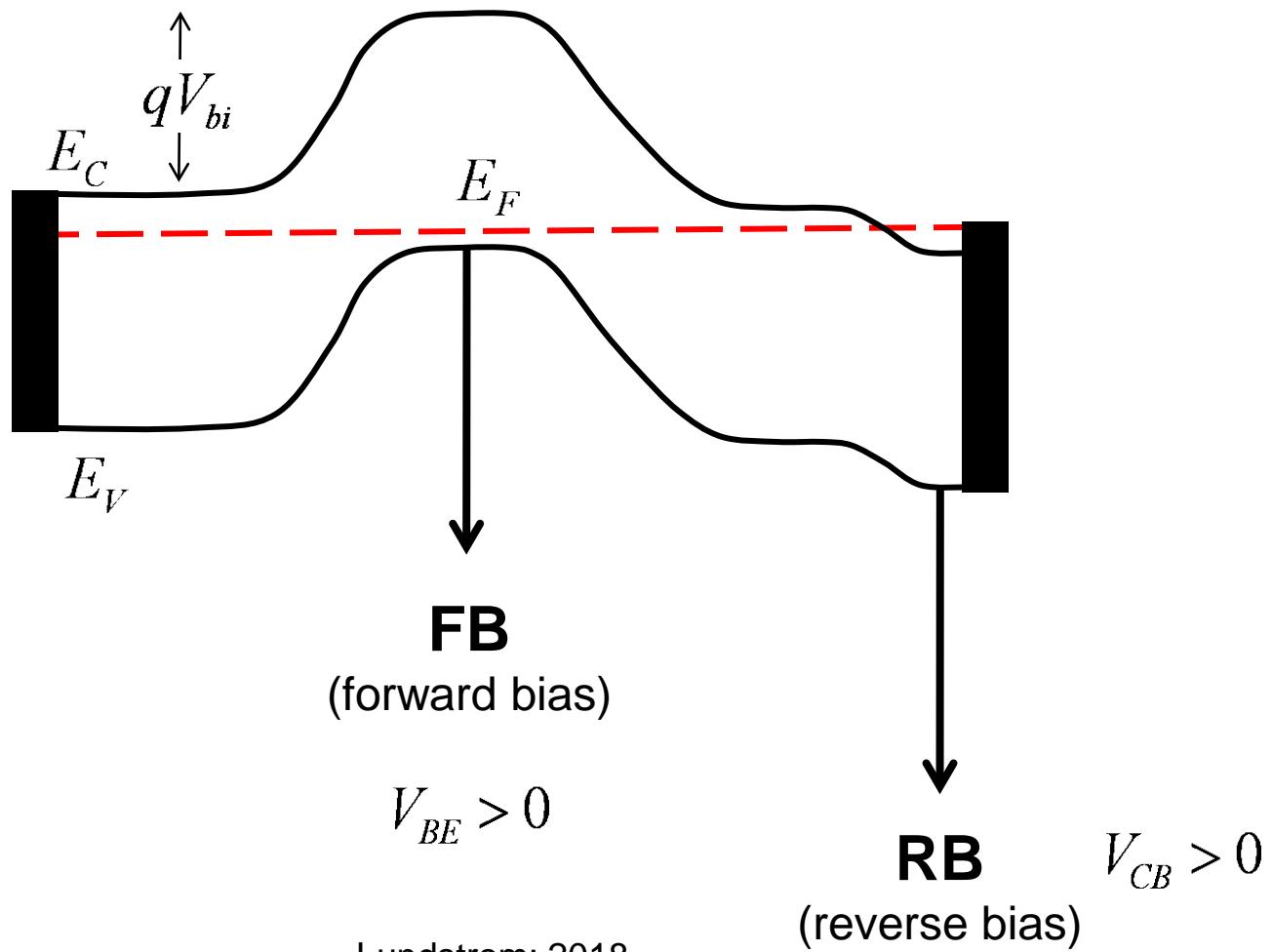
$$J_n = q \frac{D_n}{W_P} \frac{n_{iP}^2}{N_A} \left(e^{qV_A/k_B T} - 1 \right)$$
$$J_p = q \frac{D_p}{W_N} \frac{n_{iN}^2}{N_D} \left(e^{qV_A/k_B T} - 1 \right)$$

Equilibrium E-band diagram: bipolar transistor

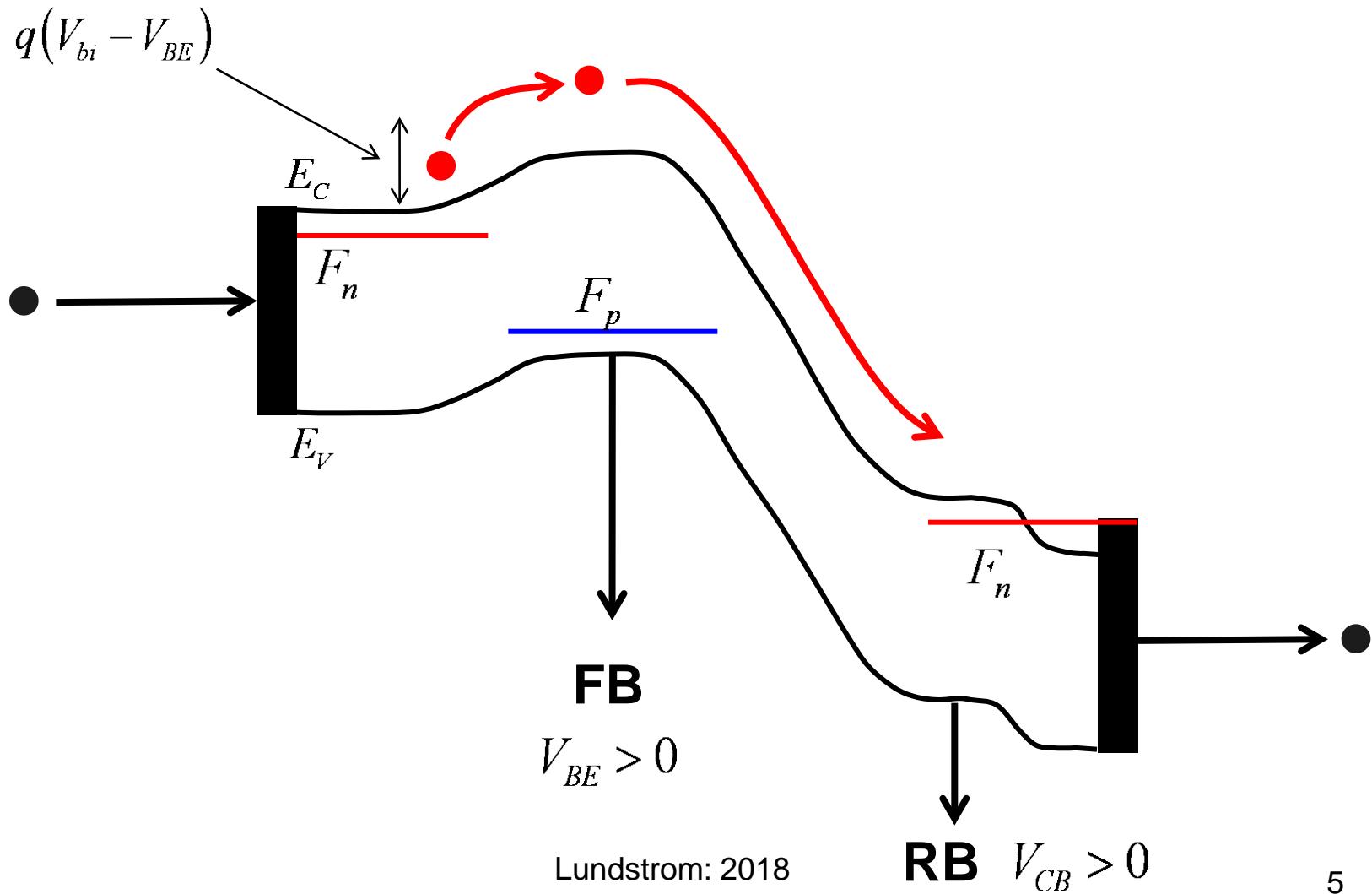


Lundstrom: 2018

The BJT: a barrier controlled device



The BJT: a barrier controlled device

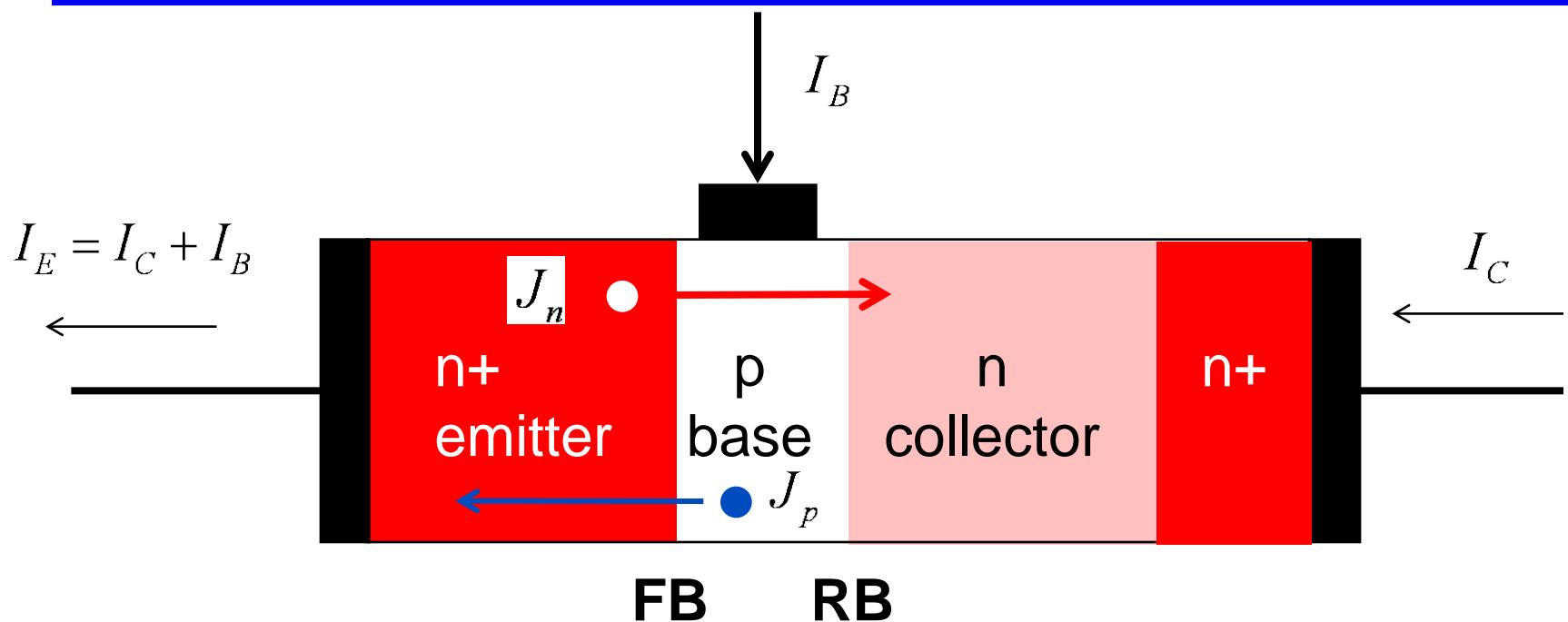


MOSFETs and BJTs

“The Insulated Gate Field Effect Transistor – A Bipolar Transistor in Disguise,” *RCA Review*, vol. 34, pp. 80-94, 1973.

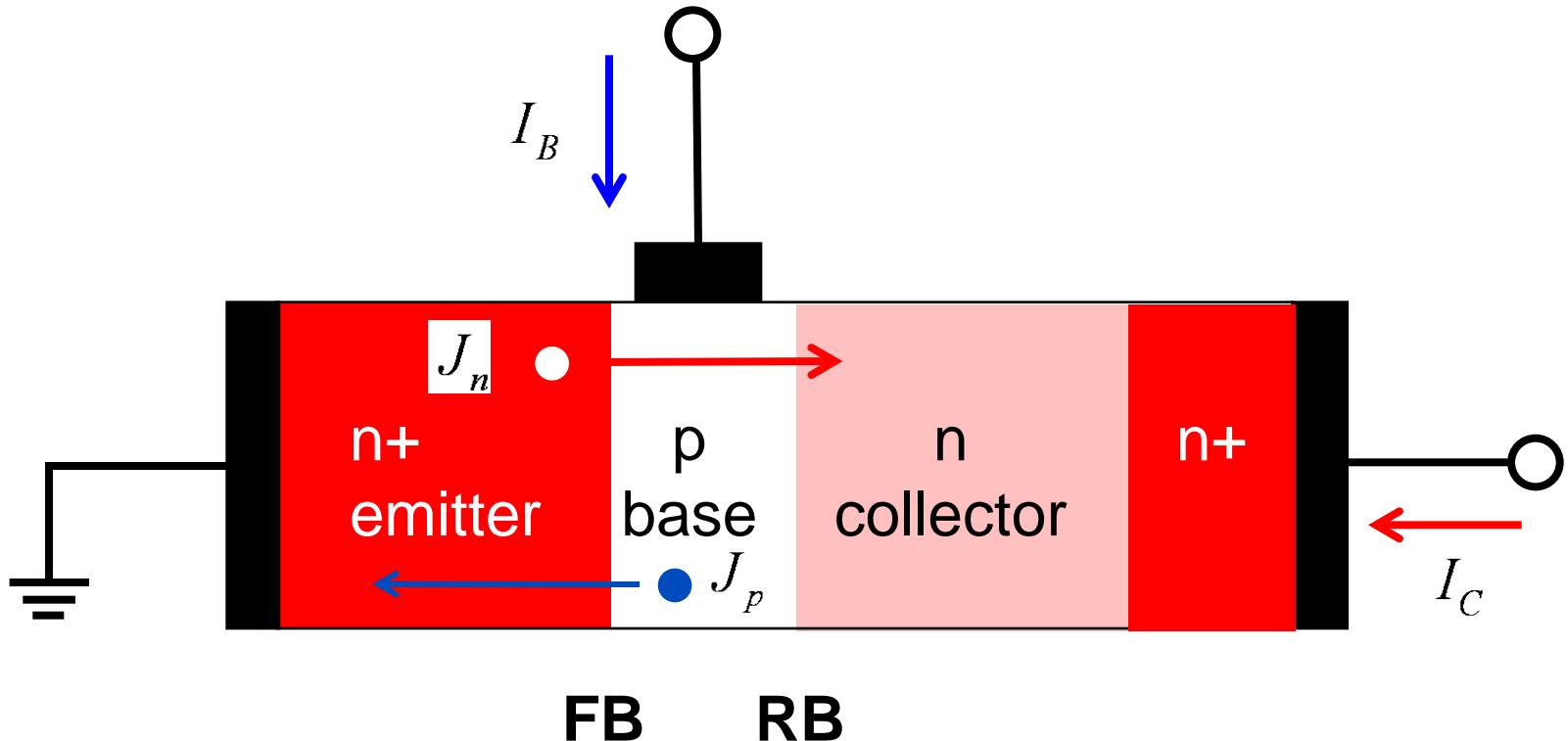
“Although both the Insulated Gate Field-Effect transistor (IGFET) and the bipolar transistor both operate on charge control principles, these device are widely believed to be intrinsically different in the details of their operation.”

Forward active region of operation



$$I_C = A_E J_n = q A_E \frac{D_n}{W_B} \frac{n_{iB}^2}{N_{AB}} \left(e^{qV_{BE}/k_B T} - 1 \right) \quad I_B = A_E J_p = q A_E \frac{D_p}{W_E} \frac{n_{iE}^2}{N_{DE}} \left(e^{qV_{BE}/k_B T} - 1 \right)$$

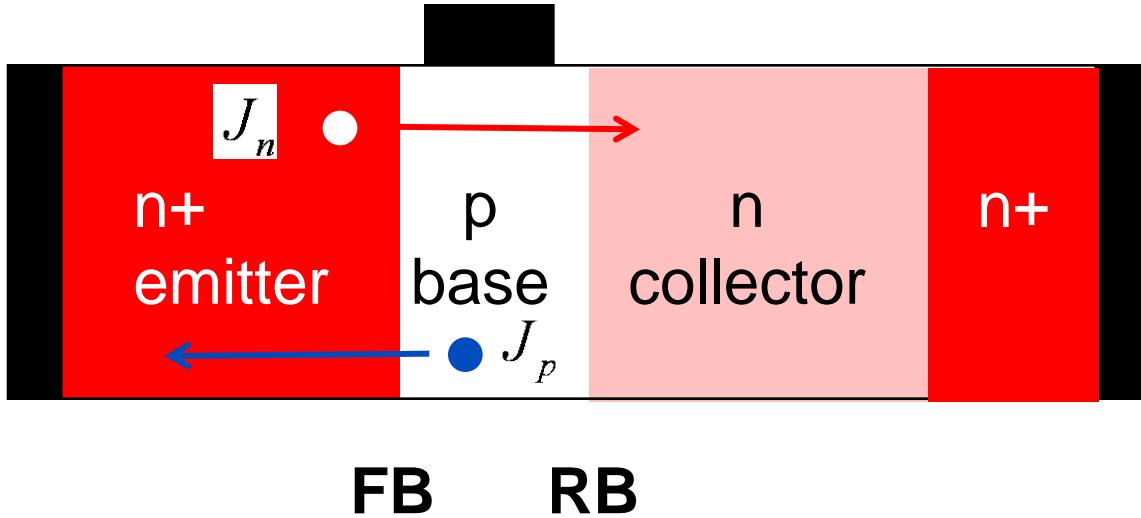
A current controlled device?



$$I_B = qA_E \frac{D_p}{W_E} \frac{n_{iE}^2}{N_{DE}} \left(e^{qV_{BE}/k_B T} - 1 \right)$$

$$V_{BE} \propto \ln(I_B)$$

Current gain



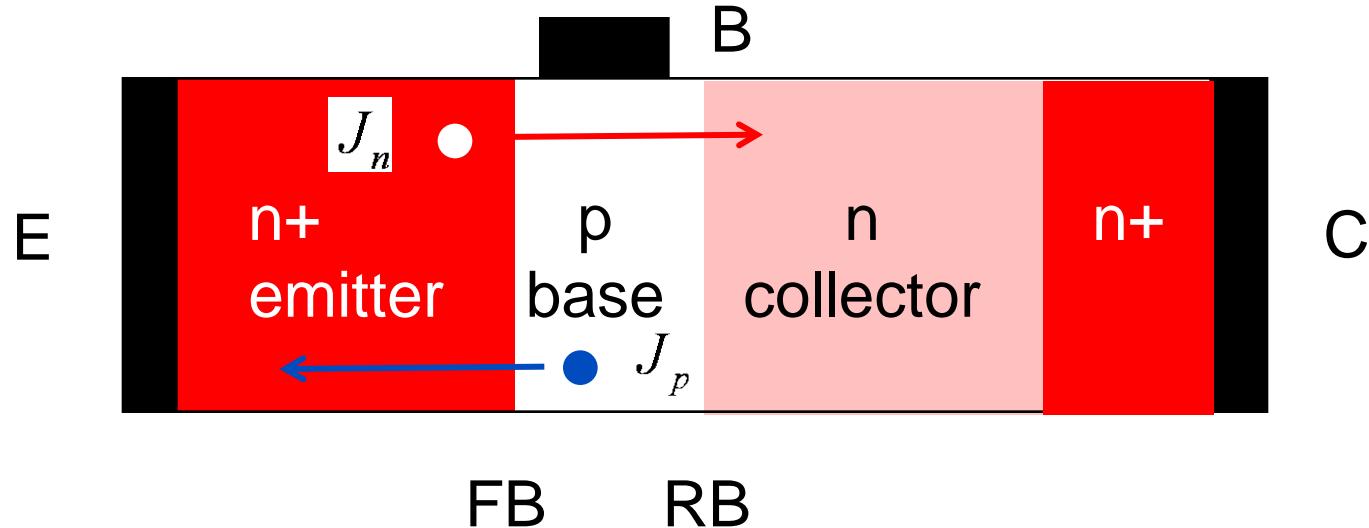
$$I_C = qA_E \frac{D_n}{W_B} \frac{n_{iB}^2}{N_{AB}} \left(e^{qV_{BE}/k_B T} - 1 \right)$$

$$I_B = qA_E \frac{D_p}{W_E} \frac{n_{iE}^2}{N_{DE}} \left(e^{qV_{BE}/k_B T} - 1 \right)$$

$$\frac{I_C}{I_B} = \left[\frac{D_n}{D_p} \frac{W_E}{W_B} \frac{N_{DE}}{N_{AB}} \frac{n_{iB}^2}{n_{iE}^2} \right] = \beta$$

$$I_C = \beta I_B$$

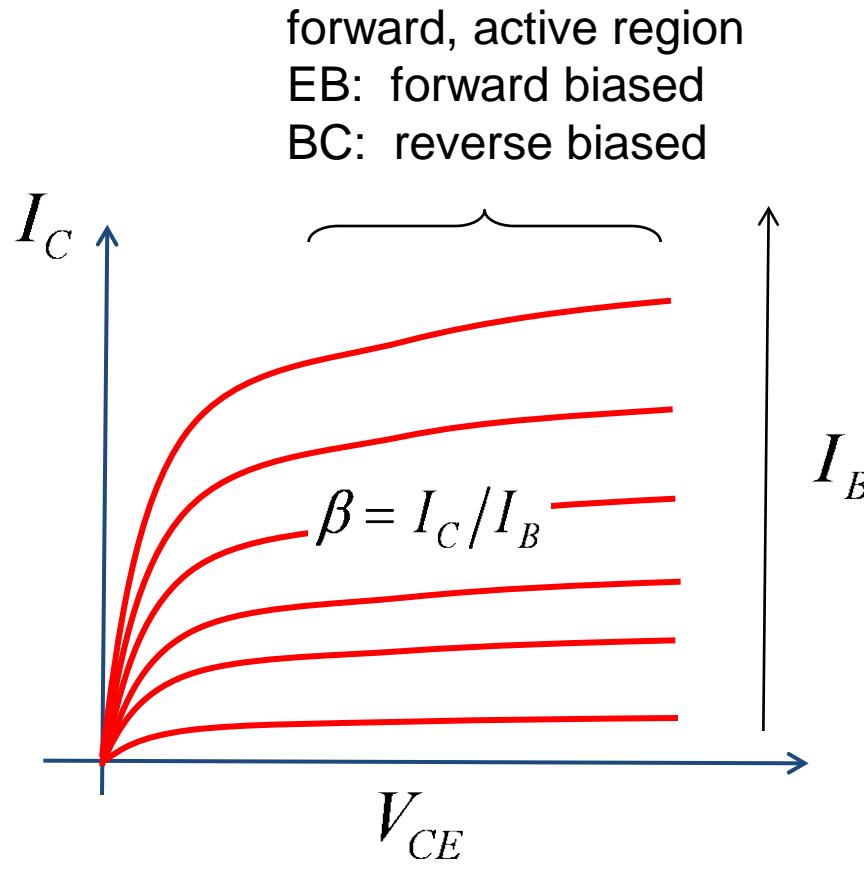
Forward active region



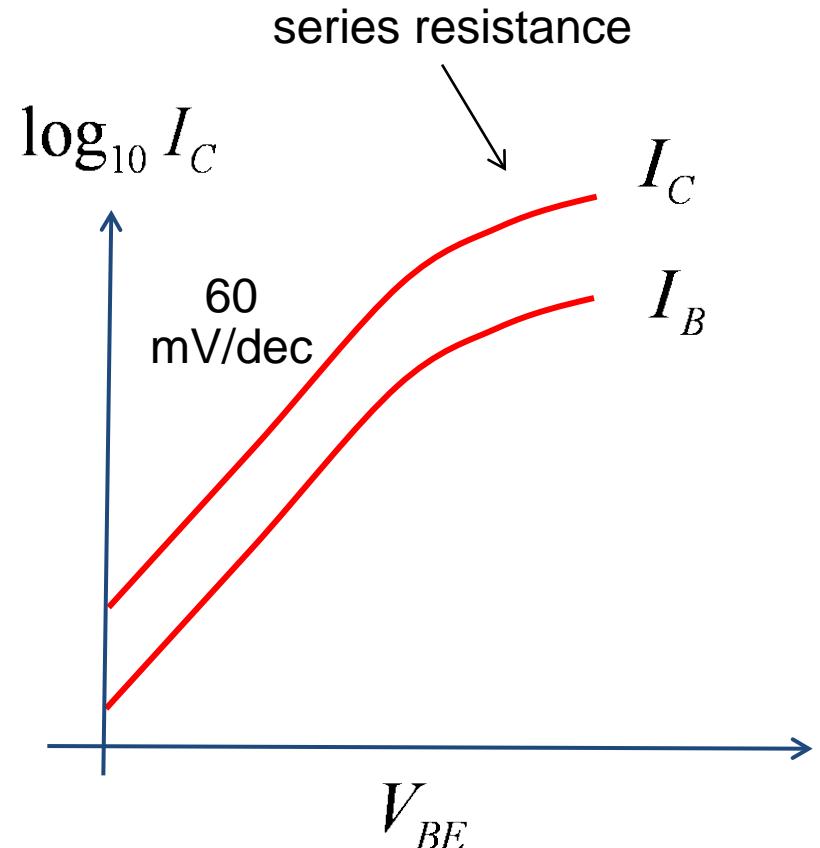
$$\beta = \frac{I_C}{I_B} = \frac{N_{DE}}{N_{AB}} \frac{D_n}{D_p} \frac{W_E}{W_B} \frac{n_{iB}^2}{n_{iE}^2}$$

(In the forward active region of operation)

IV characteristics



output characteristic



transfer characteristic

MOSFETs vs. BJTs

For analog / RF applications, two important figures of merit are:

Transconductance

$$g_m = \frac{\partial I_C}{V_{BE}} \Big|_{V_{CE}}$$

$$g_m = \frac{\partial I_D}{V_{GS}} \Big|_{V_{DS}}$$

Gain-bandwidth product

$$f_T = \frac{g_m}{2\pi C_{tot}}$$

Transconductance

MOSFET

Saturation region:

$$I_D = k_n(V_{GS} - V_T)$$

$$g_m = \partial I_D / \partial V_{GS} \Big|_{V_{DS}}$$

$$g_m = k_n$$

$$g_m = \frac{I_D}{(V_{GS} - V_T)}$$

BJT

Forward active region:

$$I_C = I_{C0} e^{qV_{BE}/k_B T}$$

$$g_m = \partial I_C / \partial V_{BE} \Big|_{V_{CE}}$$

$$g_m = \frac{I_{C0}}{(k_B T / q)} e^{qV_{BE}/k_B T}$$

$$g_m = \frac{I_C}{k_B T / q}$$

MOSFET vs. BJT transconductance

MOSFET	BJT
$g_m = \frac{I_D}{(V_{GS} - V_T)}$	$g_m = \frac{I_C}{k_B T/q}$
$I_D = 1 \text{ mA}$	$I_C = 1 \text{ mA}$
$V_{GS} = 1.0 \text{ V}$	$k_B T/q = 0.026 \text{ V}$
$g_m = 2.5 \text{ mS}$	$g_m = 40 \frac{\text{mA}}{\text{V}} = 40 \text{ mS}$

Importance of series resistance

$$f_T = \frac{g_m}{2\pi C_{tot}}$$

$$2\pi f_T = \frac{g_m}{C_{tot}} = \frac{1}{\tau}$$

$$\tau = t_t + \frac{C_{par}}{g_m}$$

$$\tau = \frac{C_{tot}}{g_m} = \frac{C_{eb}}{g_m} + \frac{C_{par}}{g_m}$$

$$\frac{C_{eb}}{g_m} = \frac{\partial Q_n / \partial V_{BE}}{\partial I_C / \partial V_{BE}} = t_t$$

$$\tau = t_t + \frac{C_{par} k_B T / q}{I_C}$$

(Need high drive currents and low parasitic resistance.)

MOSFETs vs. BJTs

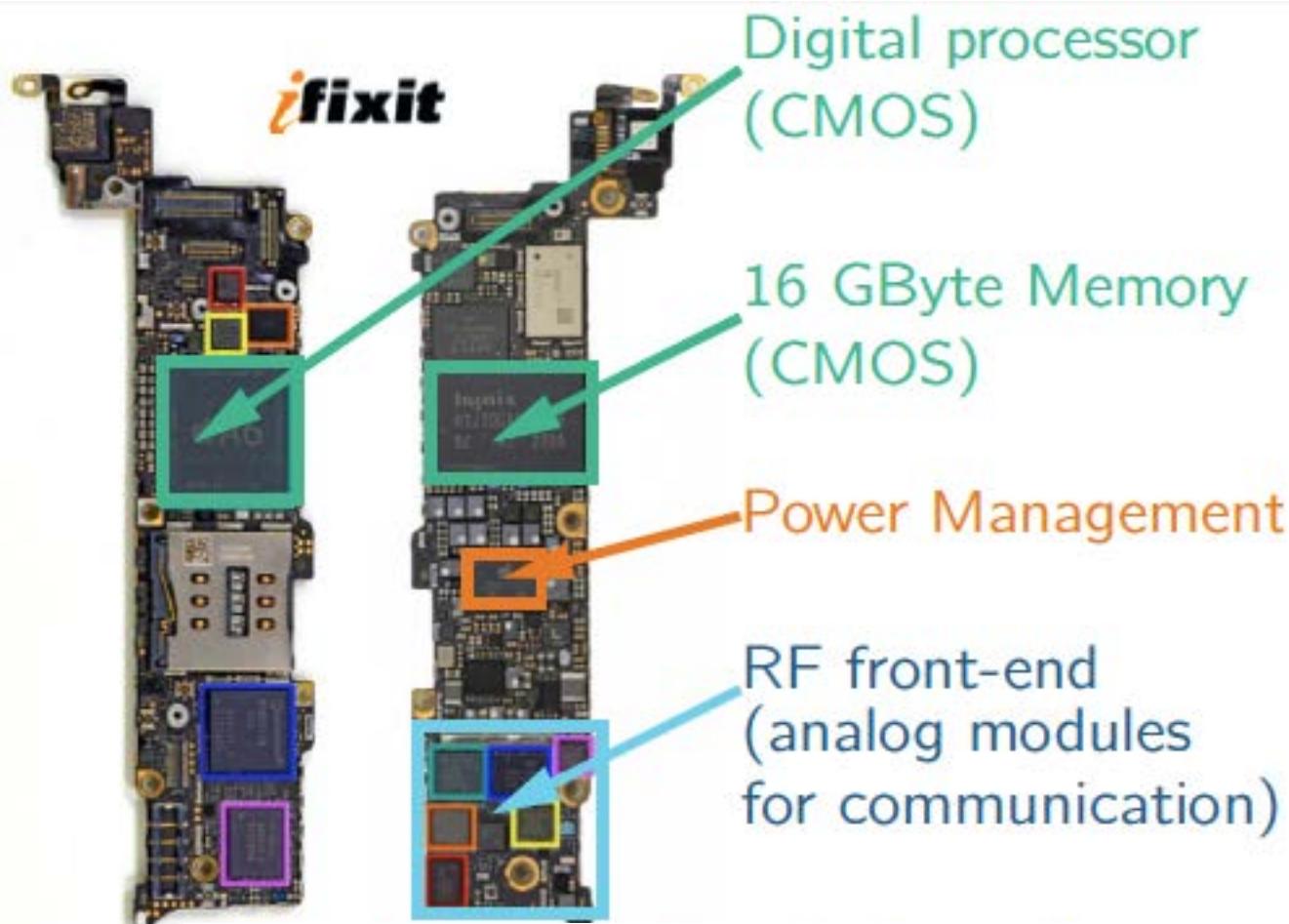
BJTs are superior RF/analog devices. A typical BJT has more than 10X the transconductance of a typical MOSFET.

Both devices can achieve high f_T
$$f_T = \frac{g_m}{2\pi C_{tot}}$$

BJTs offer higher drive currents (higher g_m)

→ Well-suited for RF power amplifiers

SiGe HBTs



Circuit board of an iPhone 5

Heterojunction bipolar transistors

Modern bipolar transistors are Heterojunction Bipolar Transistors (HBTs).

- 1) What is an HBT?
- 2) Why are they useful?

Design trade-offs for homojunction BJTs

$$\beta = \frac{N_{DE}}{N_{AB}} \frac{D_n}{D_p} \frac{W_E}{W_B} \frac{n_{iB}^2}{n_{iE}^2}$$

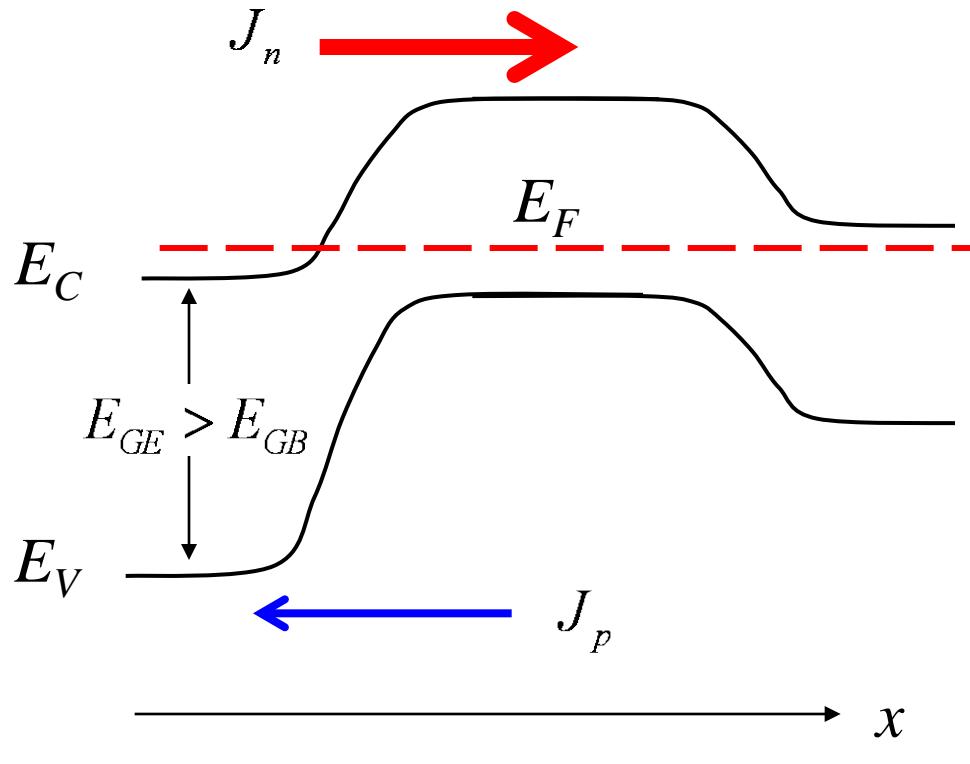
For homojunction BJT's: $n_{iE} = n_{iB}$

For high current gain: $N_{DE} \gg N_{AB}$

Would prefer a thin, heavily doped base.

- thin for speed
- heavily doped for base resistance

Heterojunction bipolar transistors (HBTs)



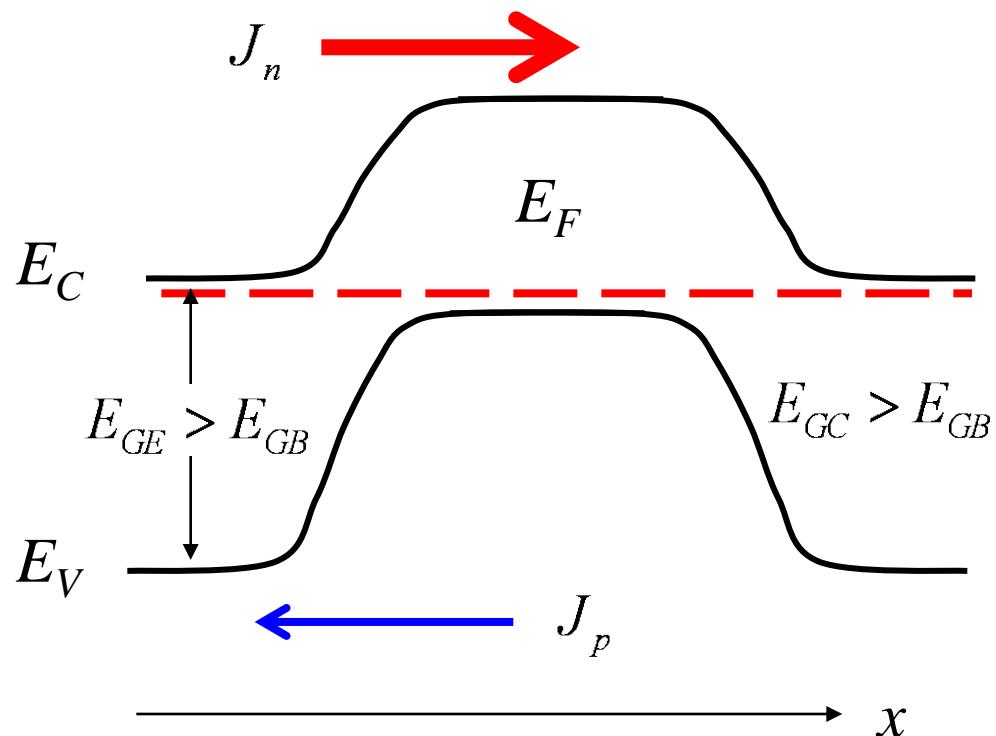
wide bandgap emitter

$$\beta = \frac{N_{DE}}{N_{AB}} \frac{D_n}{D_p} \frac{W_E}{W_B} \frac{n_{iB}^2}{n_{iE}^2}$$

$$n_i^2 = N_C N_V e^{-E_G/k_B T}$$

$$\beta \approx \frac{N_{DE}}{N_{AB}} \frac{D_n}{D_p} \frac{W_E}{W_B} e^{\Delta E_G/k_B T}$$

Double HBT (DHBT)



- symmetrical operation
- reduced collector offset voltage
- higher collector breakdown voltage

Modern III-V HBTs

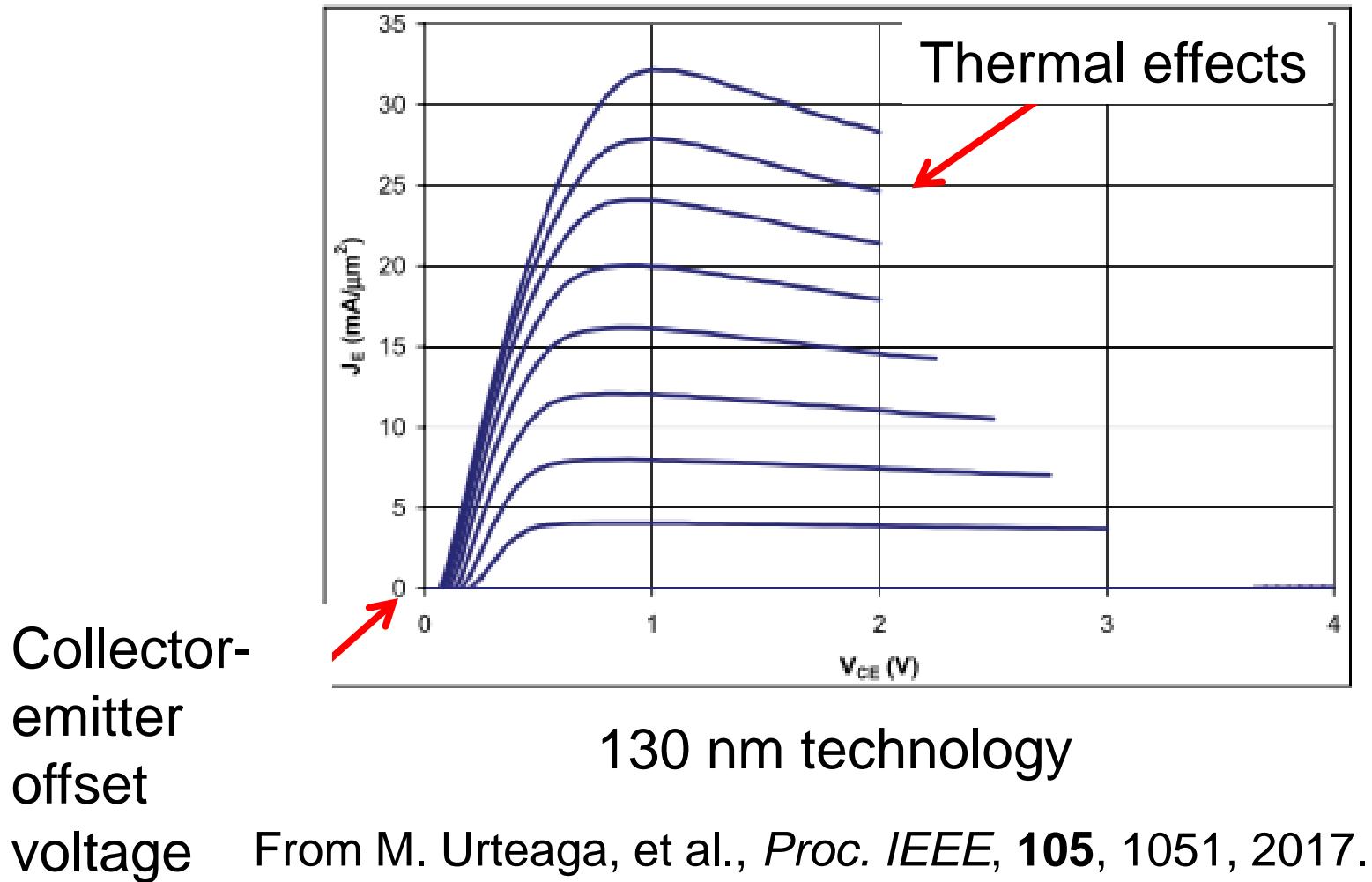
For a good review of the current state-of-the-art, see:

M. Urteaga, Z. Griffith, M. Seo, J. Hacker, and M. Rodwell, “InP HBT Technologies for THz Integrated Circuits,” *Proc. IEEE*, Vol. 105, pp.1051-1067, 2017.

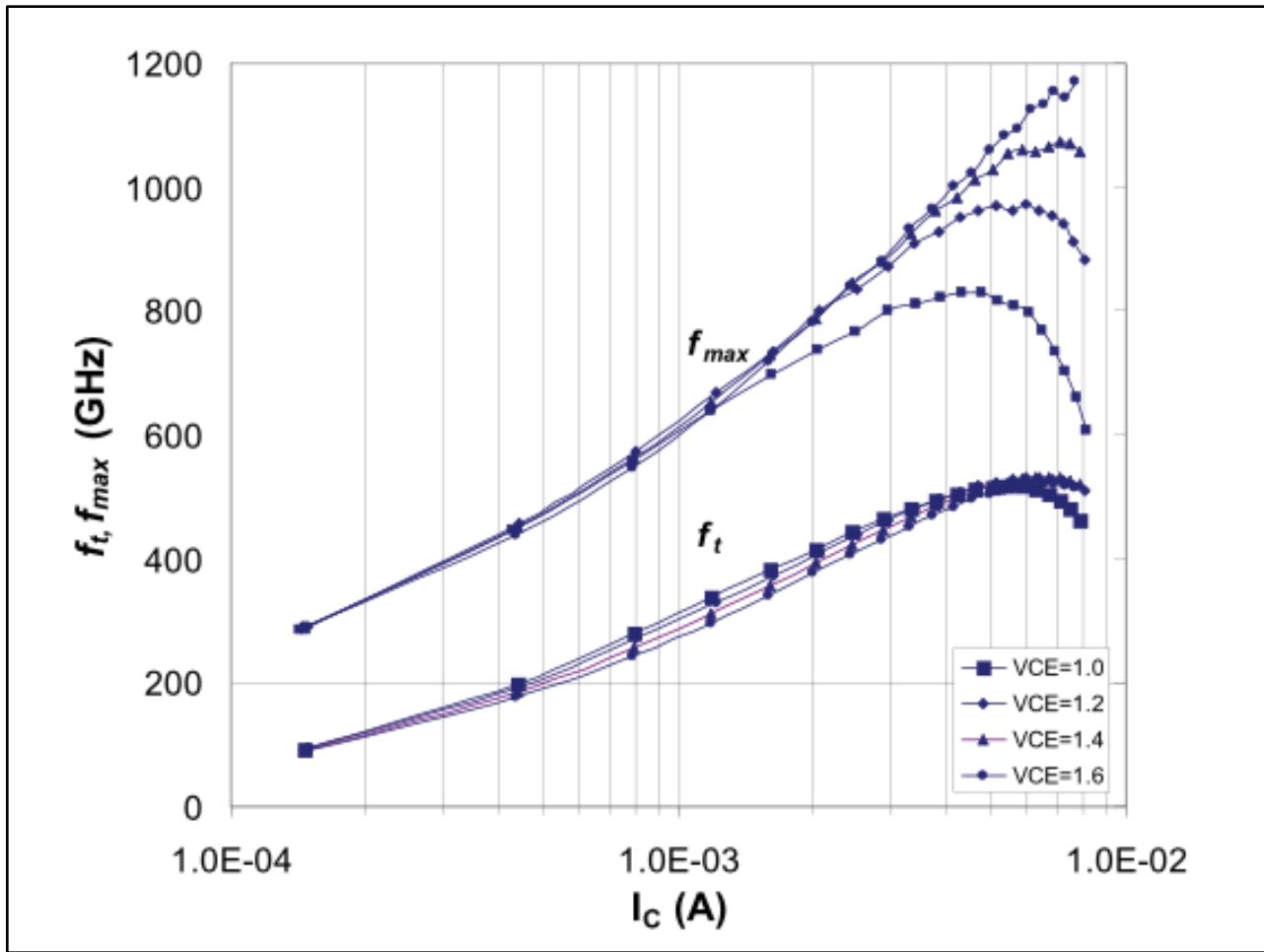
For a discussion about the critical role of parasitics, see:

J. C. Rode, H.-W. Chiang, P. Choudhary, V. Jain, B. J. Thibeault, W. J. Mitchell, M. J. W. Rodwell, M. Urteaga, D. Loubichev, A. Snyder, Y. Wu, J. M. Fastenau, and A. W. K. Liu, “Indium Phosphide Heterobipolar Transistor Technology Beyond 1-THz Bandwidth,” *IEEE Trans. Electron Dev.*, Vol. 62, pp. 2779-2785, 2015.

Output characteristics

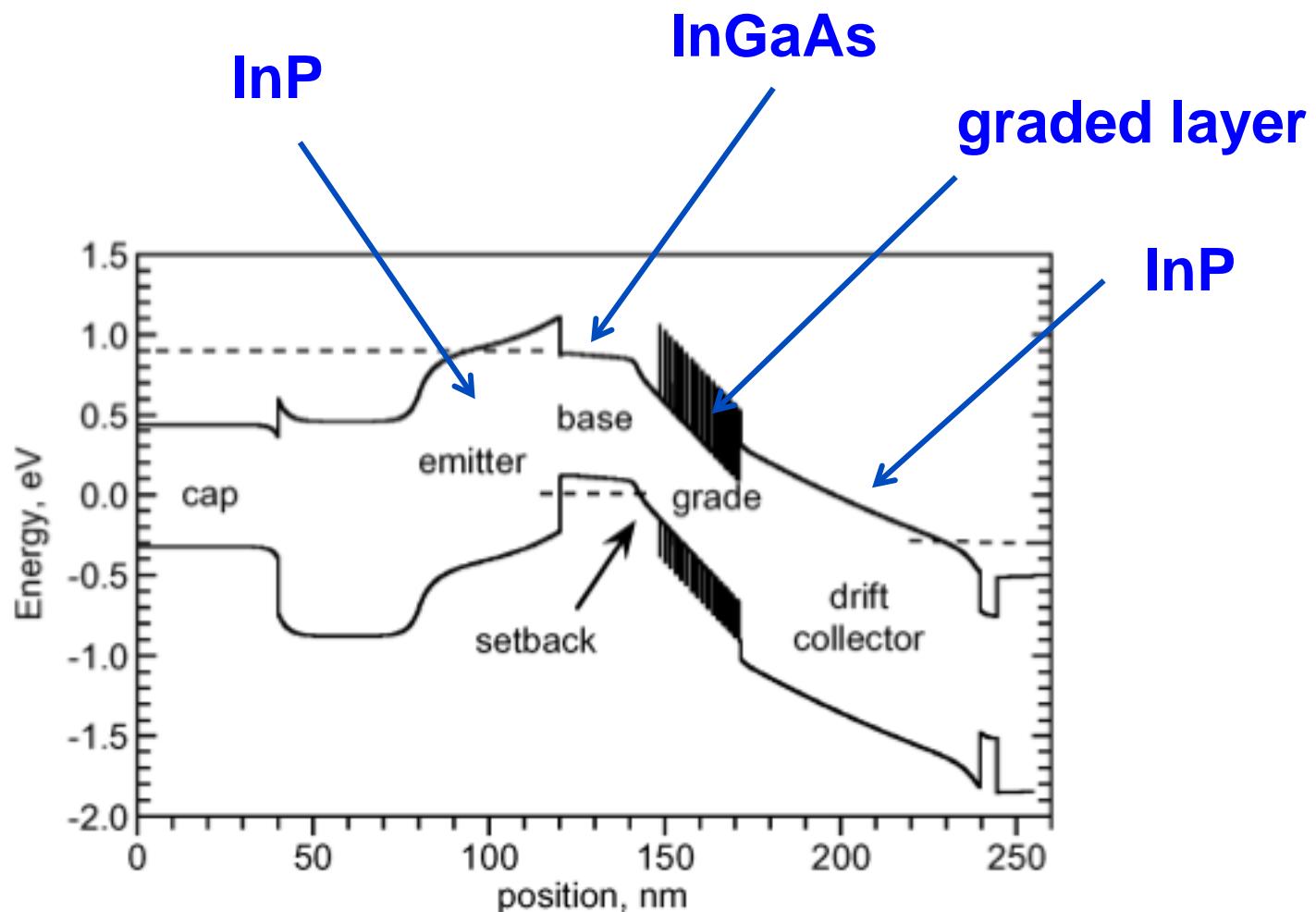


f_T and f_{max} vs. I_C



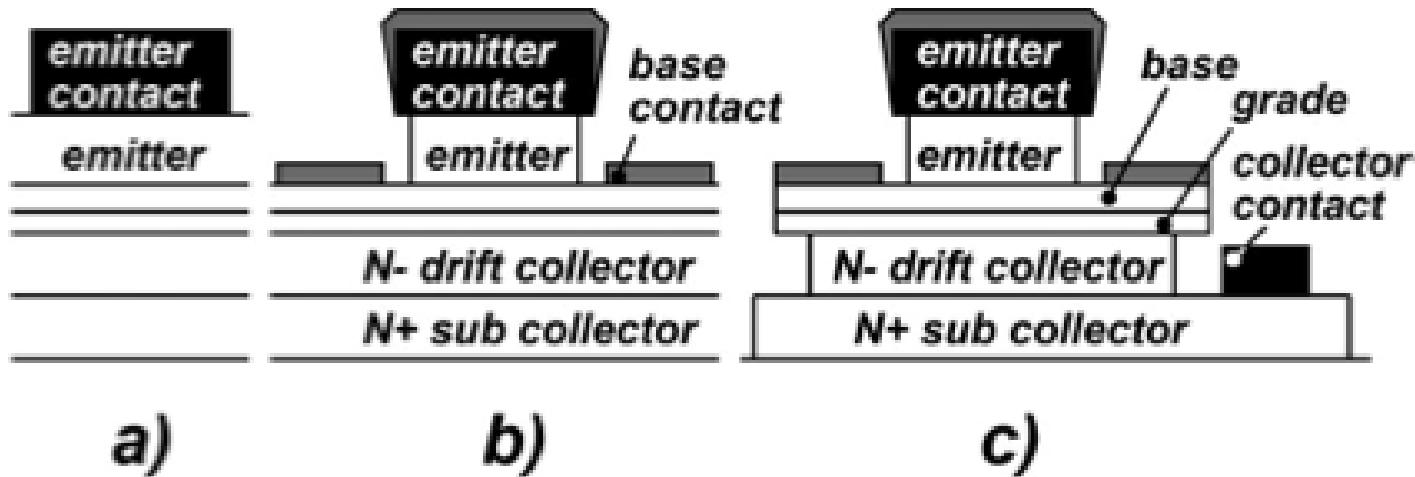
From M. Urteaga, et al., Proc. IEEE, 105, 1051, 2017.

DHBT: InP: InGaAs: InP



From M. Rodwell, et al., *Proc. IEEE*, **96**, 271, 2008.

Mesa fabrication process



- a) Emitter metal deposition
- b) Emitter etch and self-aligned base contact metal deposition
- c) Base mesa etch, collector mesa etch, collector contact deposition

InP HBTs

Key scaling challenges:

- emitter & base contact resistivity
- current density → device heating
- parasitic capacitances
- current gain (surface recombination)

Summary

- 1) III-V HBTs are an important technology for high-frequency RF power applications.
- 2) SiGe HBTs are a critical technology for wireless electronics
- 3) HBTs offer speeds close to that of HEMTs ($f_T = 600$ GHz, $f_{max} = 1200$ GHz).
- 4) Compared to HEMTs, HBTs deliver higher power and integration density.
- 5) HBTs operate in the same “barrier controlled mode” as Si MOSFETs.

Thanks to Mark Rodwell (UCSB), for his help in putting together this lecture.

Next topic

Whatever the type of transistor, a good compact circuit model is needed to design circuits.

We briefly discussed compact circuit models in Unit 1. In the next lecture, we'll say a little more about compact models for circuit design.

Essentials of MOSFETs

Unit 5: Additional Topics

Lecture 5.6: A Second Look at Compact Circuit Models

Mark Lundstrom

lundstro@purdue.edu

Electrical and Computer Engineering
Purdue University
West Lafayette, Indiana USA

Lundstrom: 2018

Type 1 Compact Models

In Lecture 1.5, we discussed two types of “compact models.”

1) Compact physical models

These models aim to describe a device in terms of a few parameters with strong physical significance. These kinds of models are useful for device characterization, process monitoring, and for the conceptual understanding that guides device research.

Our focus in this course has been on this type of compact model.

Type 2 Compact Models

2) Compact device models for circuit simulation

These models accurately describe the electrical behavior of a device **in a form suitable for use in numerical circuit simulation programs**. To describe everything relevant to a circuit, these models are more complex and semi-empirical, but the core of the model is usually a compact physics model.

These kinds of compact models interface semiconductor technology to circuit design.

Compact models for circuit simulation

1) **Table based**

Tables of measured or simulated data. The model then interpolates (smoothly) to describe a device across a range of biases.

2) **Physics based**

Analytical or simple numerical expressions that describe a device across a range of biases, layout geometries, and temperatures.

Why physics-based?

- + Physics-based model relate the needs of designers to the manufacturing process.
- + Generally results in the fewest number of model parameters and simplifies model calibration.
- + These models provide the best basis for statistical and mismatch modeling.
- + As manufacturing processes evolve, these models can quickly adapt.
- Development time

MVS for circuit simulation

Versions of the MVS model for circuit simulation exist and are useful for technology development.

But the MVS model is not an industry-strength model for digital circuit simulation because it:

- doesn't smoothly go from short to long channels
- doesn't treat a number of second and third order effects
- doesn't treat noise
- doesn't treat layout dependencies
- etc.

Source referenced models

Square Law Model

$$\begin{array}{l} V_{GS} > V_T \\ V_{DS} < V_{GS} - V_T \end{array} \quad I_D = \mu_n C_{ox} \frac{W}{L} \left[(V_{GS} - V_T) V_{DS} - \frac{V_{DS}^2}{2} \right]$$

$$\begin{array}{l} V_{GS} > V_T \\ V_{DS} > V_{GS} - V_T \end{array} \quad I_D = \mu_n C_{ox} \frac{W}{2L} (V_{GS} - V_T)^2$$

Problems occur when V_{DS} changes sign.

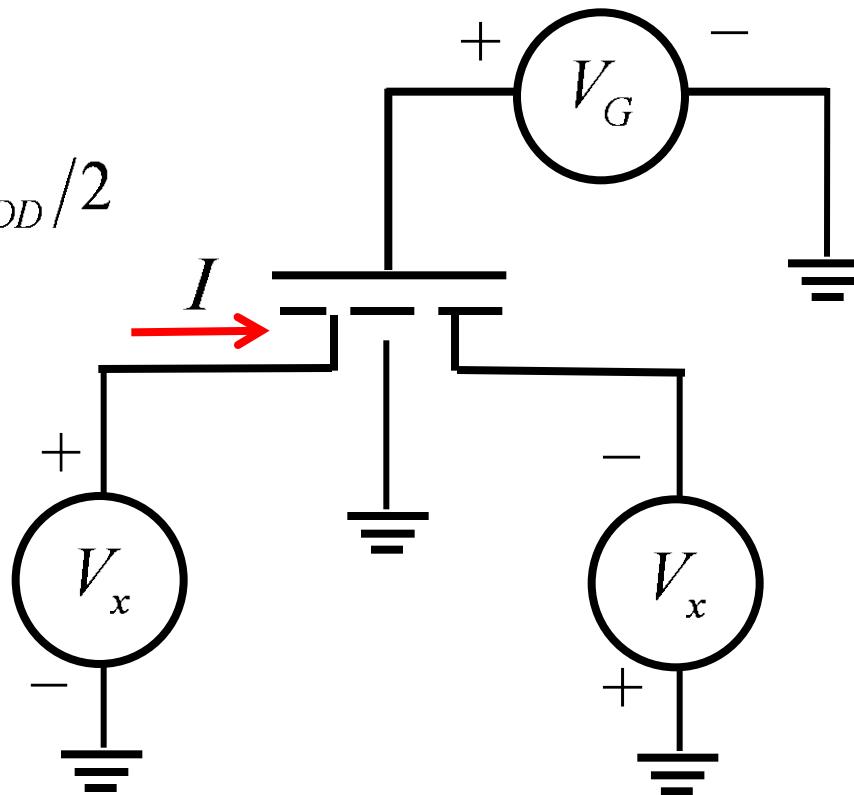
For an early example that is not source referenced, see:

H. Shichman and D.A. Hodges, “Modeling and simulation of insulated gate field-effect transistor switching circuits,” *IEEE J. Solid State Circuits*, **SC-3**, 285-289, 1968.

“Gummel symmetry test”

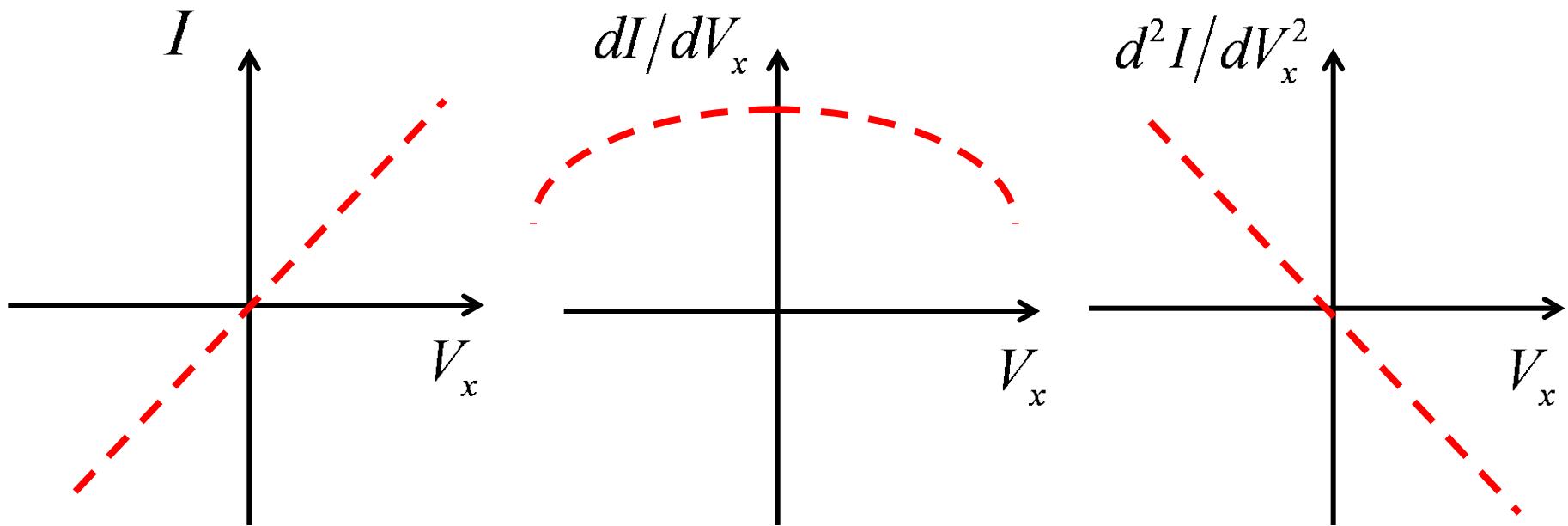
$$-V_{DD}/2 \leq V_x \leq +V_{DD}/2$$

$$V_{DS} = 2V_x$$



Many models had problems near the origin where $V_{DS} = 0$.

Gummel Symmetry Test



Designers need robust, well-behaved models, but for early stage R&D, simpler models are still useful.

Device research

“The principal applications of any sufficiently new and innovative technology always have been – and will continue to be – applications created by that technology.” – Herbert Kroemer, 2001 Noble Lecture

But today, electronics is mature, many devices are available, and materials and device research is more and more driven by applications.

Compact circuit models allow materials and device researchers to explore impact on applications at an early stage.

Finding compact models

Device researchers and developers may be able to locate and suitably modify a model for their purposes. (Some open source models can be found at the **NEEDS** site: needs.nanoHUB.org.)

Designers need industrial strength compact models. The **Compact Model Coalition** provides members with industry-standard models (<http://www.si2.org/cmc/>)

BSIM: An industry standard MOSFET model

BSIM (Berkeley Short Channel IGFET Model) was first introduced in 1987. Since that time, BSIM has been continually enhanced, updated, and extended and now refers to a family of models. (<https://en.wikipedia.org/wiki/BSIM>)

BSIM is one of the models selected by the Compact Model Coalition as an industry standard.
(<http://www.si2.org/cmc/>)

BSIM is supported by all commercial circuit simulation platforms.

Writing your own compact model

Verilog-A is an industry standard modeling language for analog systems. It can also be used to write compact circuit models that can be used by all commercial circuit simulators and by some open source simulators.

A good starting point is:

Geoffrey Coram, “Writing Your First Verilog-A Compact Model,” <https://nanohub.org/resources/20579>

Additional tutorial information is available at:
www.needs.nanoHUB.org

Writing your own compact models

MAPP (Model and Algorithm Prototyping Platform) is a MATLAB-based, open-source, environment for developing, testing, experimentally validating compact models and for inserting them in open SPICE-compatible simulation platforms. MAPP is also useful in developing and debugging high-quality Verilog-A compact models.

Additional information and a pointer to the github repository are available at: www.needs.nanoUB.org

Before writing your own compact model

Familiarize yourself with the many lessons learned over the decades in developing compact circuit models for MOSFETs.

Start by searching “Colin C. McAndrew”

Summary

- 1) Compact circuit models link semiconductor manufacturing and circuit design.
- 2) Compact circuit models also link device and materials R&D to circuit design and applications.
- 3) Developing compact circuit models requires a good understanding of the device physics, what goes on inside the circuit simulator, and the intended application.
- 4) Many of the lessons learned in developing compact circuit models for MOSFETs can be transferred to other devices.

Next topic

We've covered several different topics in Unit 5. In the next lecture, I'll review the main points.

Essentials of MOSFETs

Unit 5: Additional Topics

Lecture 5.7: Unit 5 Recap

Mark Lundstrom

lundstro@purdue.edu
Electrical and Computer Engineering
Purdue University
West Lafayette, Indiana USA

Lundstrom: 2018

Unit 5 Topics

L5.1: Limits of MOSFETs

L5.2: Power MOSFETs

L5.3: High Electron Mobility Transistors (HEMTs)

L5.4: Review of PN Junctions

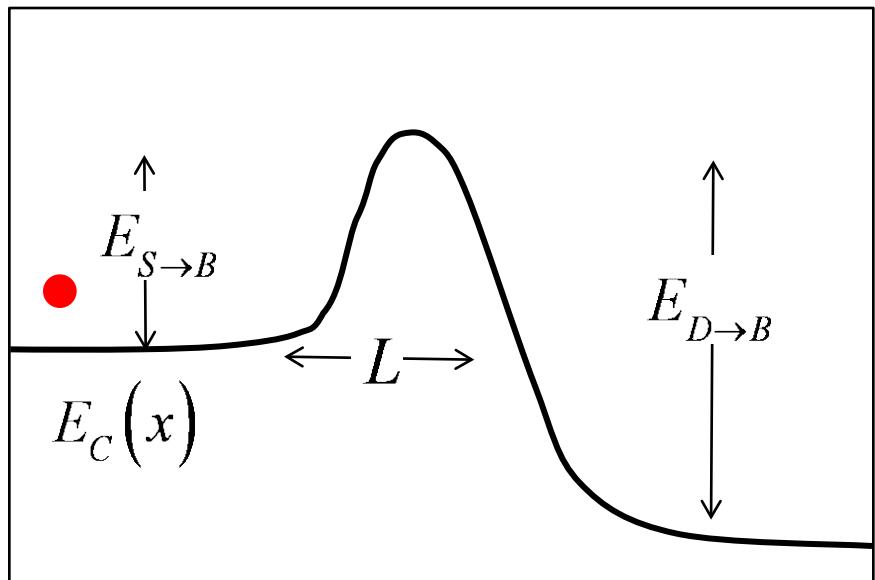
L5.5: Heterostructure Bipolar Transistors (HBTs)

L5.6: A Second Look at Compact models

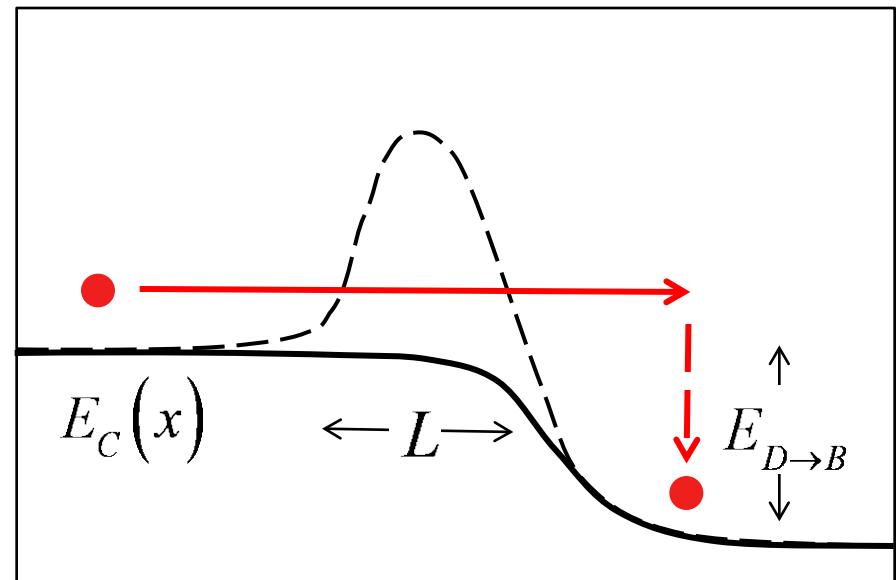
Because we covered so much ground and so many topics, I'll simply mention a few key things that you should take away from Unit 5.

Limits of MOSFETs

Off-state



On-state



Channel length, speed, energy dissipation

Limits

$$E_{\min} = k_B T \ln 2 = 0.017 \text{ eV}$$

$$L_{\min} = \frac{\hbar}{\sqrt{2m^* E_{\min}}} = 1.5 \text{ nm}$$

$$(m^* = m_0)$$

$$\tau_{\min} = \frac{\hbar}{E_{\min}} = 40 \text{ fs}$$

CMOS today

$$\frac{1}{2}C_G V_{DD}^2 \approx 57,000 \times E_{\min}$$

$$L = 20 \text{ nm} = 13 \times L_{\min}$$

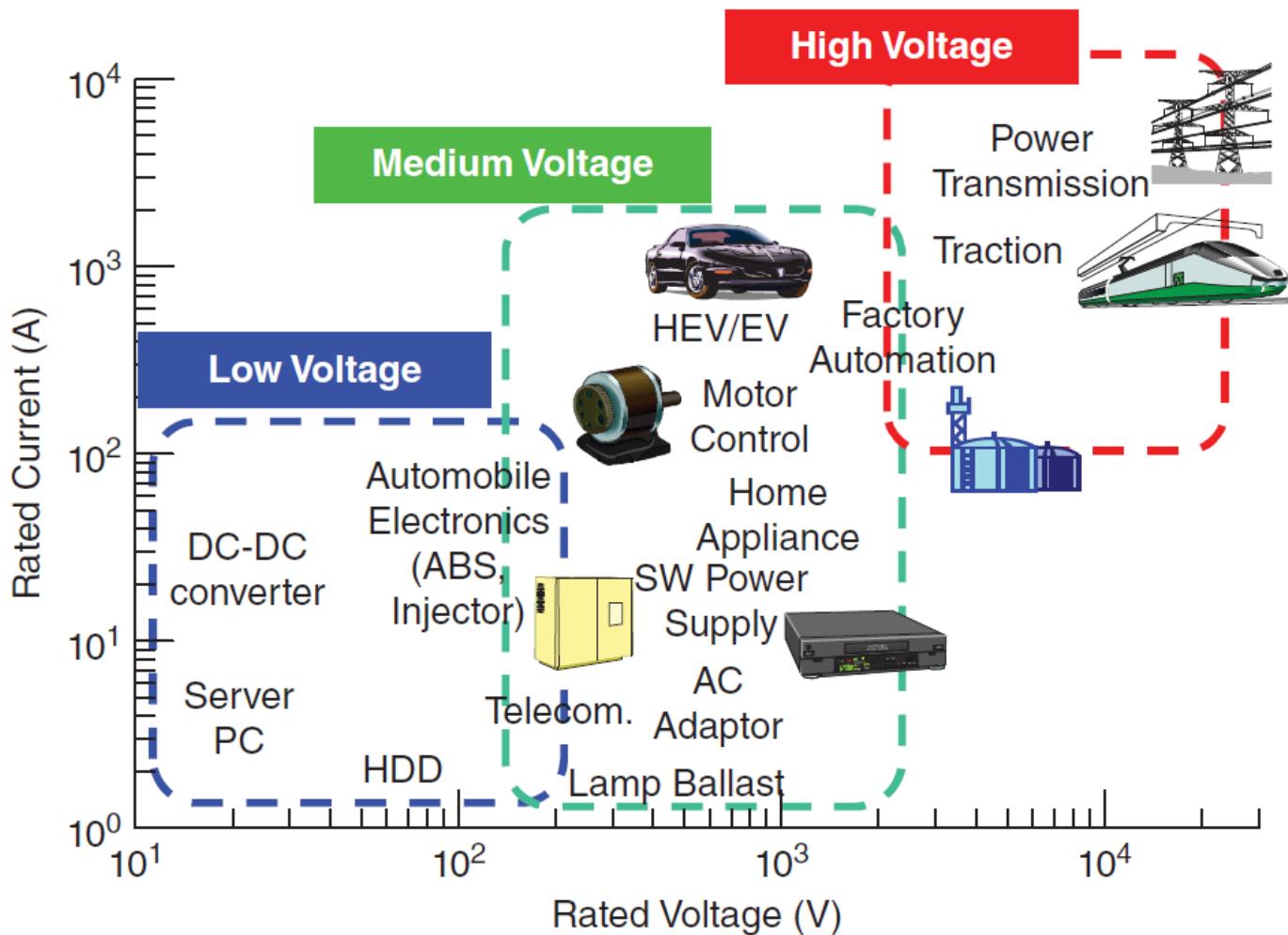
$$\tau = \frac{C_G V_{DD}}{I_{ON}} \approx 11 \times \tau_{\min}$$

22 nm technology

Summary of limits

- 1) Transistors are approaching some very fundamental limits.
- 2) Practical, technology considerations such as series resistance, parasitic capacitance, BTBT leakage currents, etc. are likely to set the practical limits.
- 3) It is unlikely that a digital switching device that is fundamentally better than a MOSFET exists.

Power semiconductor devices



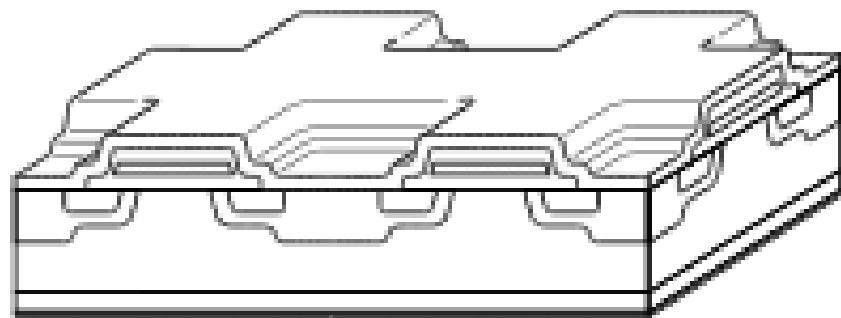
T. Kimoto and J. A. Cooper, *Fundamentals of Silicon Carbide Technology*, Wiley (2014).

Power MOSFET design

High currents require large W (MOSFET width)

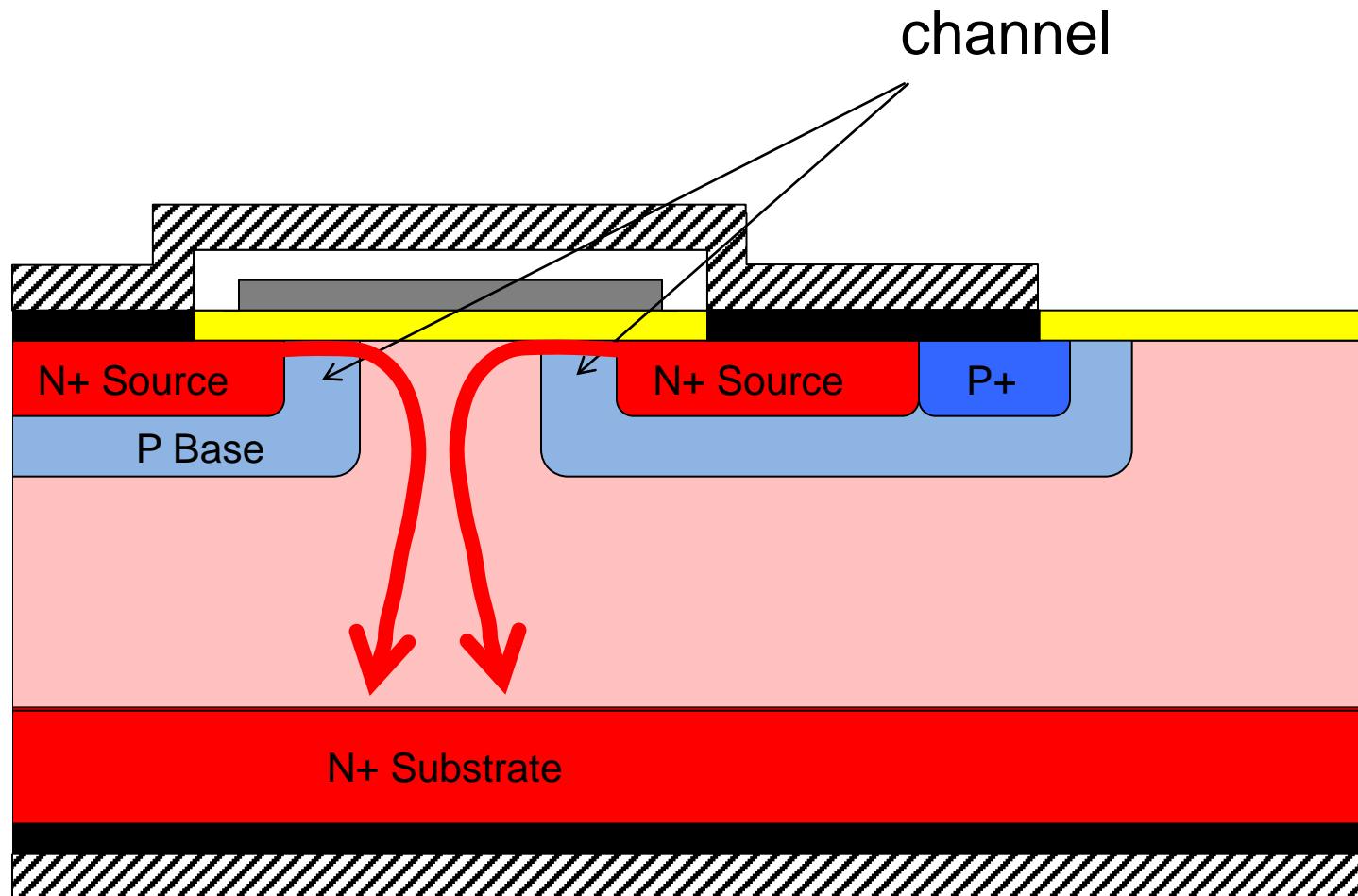
High breakdown voltage requires that we spread the voltage drop out (long channels) to minimize the electric field.

To achieve these goals, power MOSFETs use vertical current flow and cellular structures.

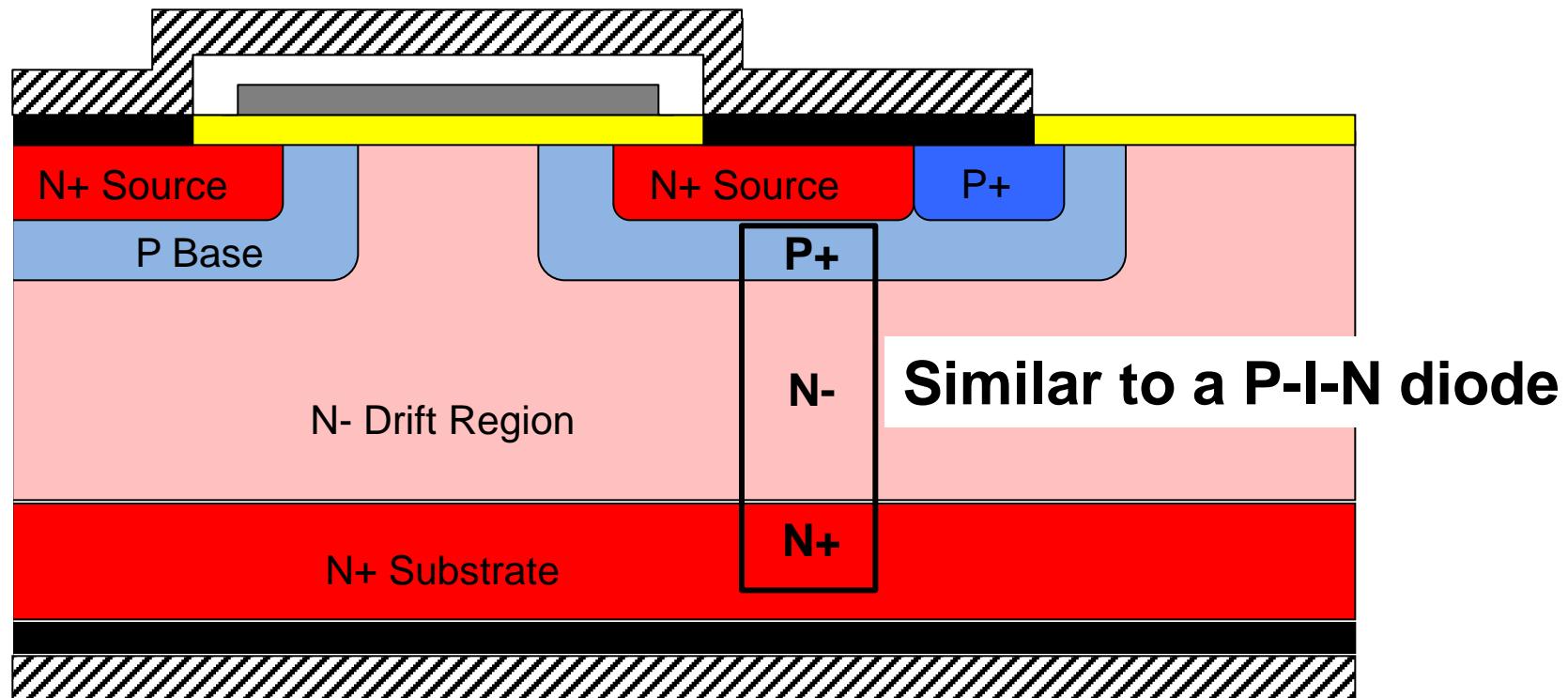


https://en.wikipedia.org/wiki/Power_MOSFET

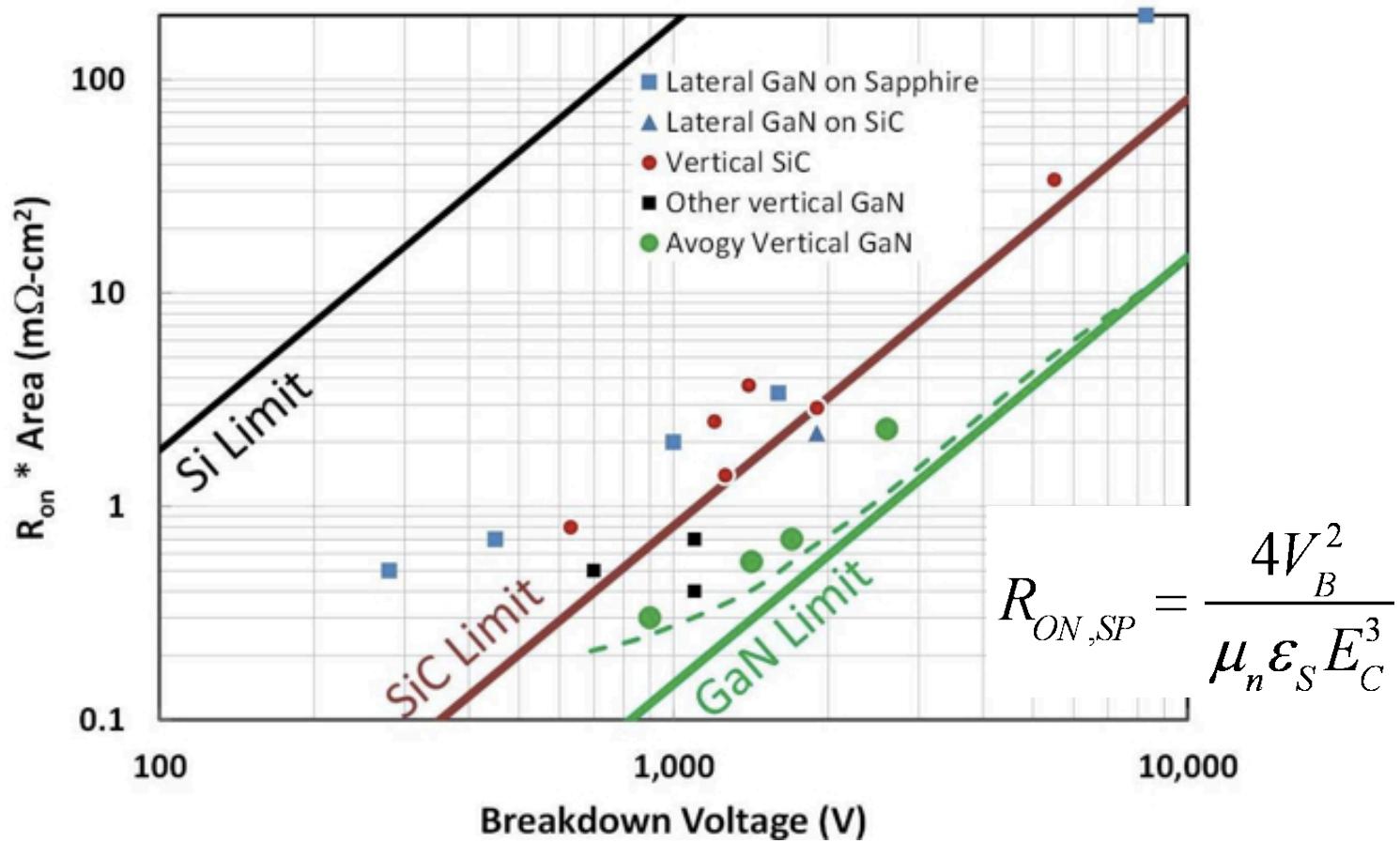
Diffused MOSFET DMOSFET (ON)



DMOSFET (OFF)

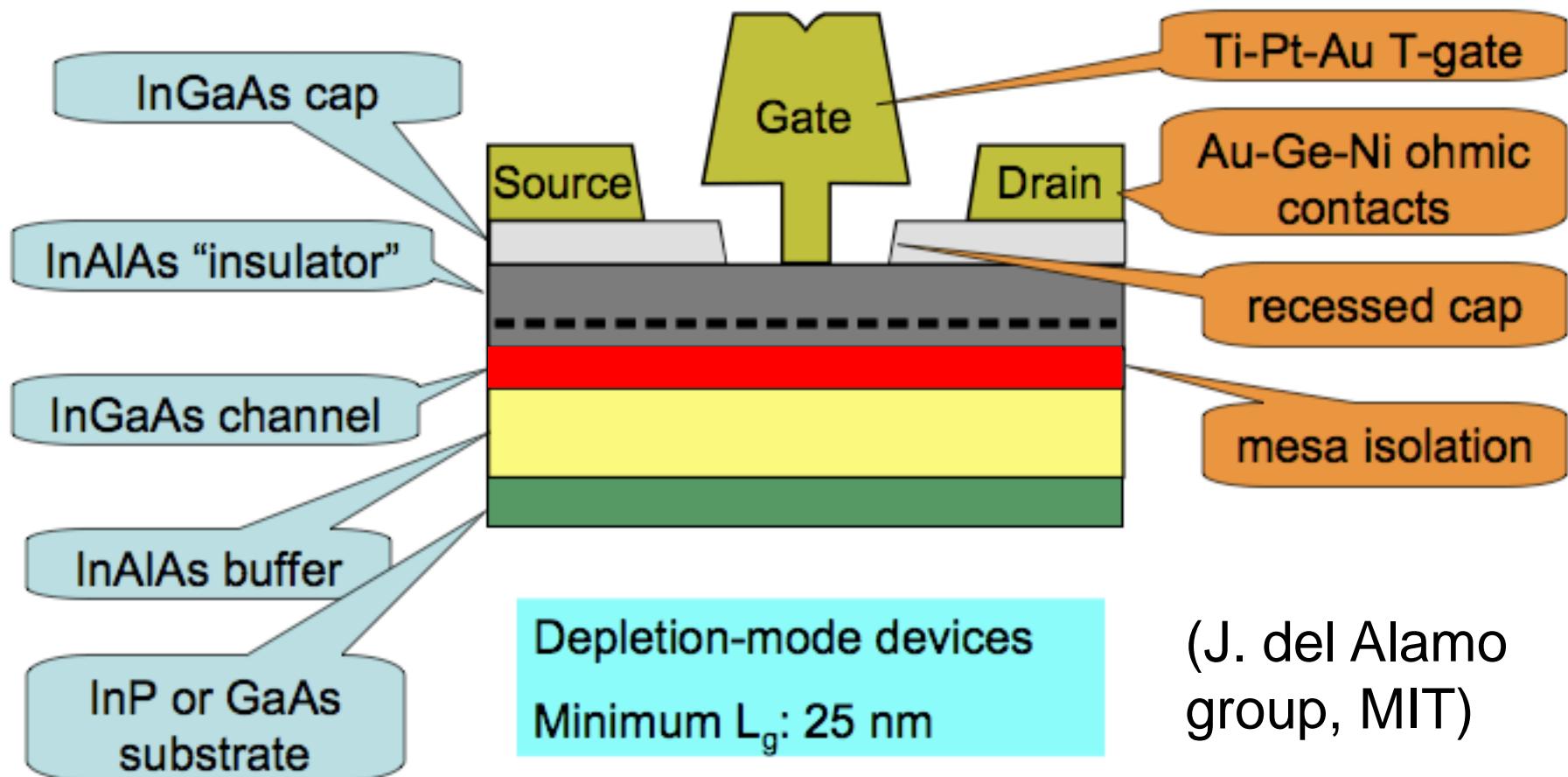


On resistance vs. blocking voltage

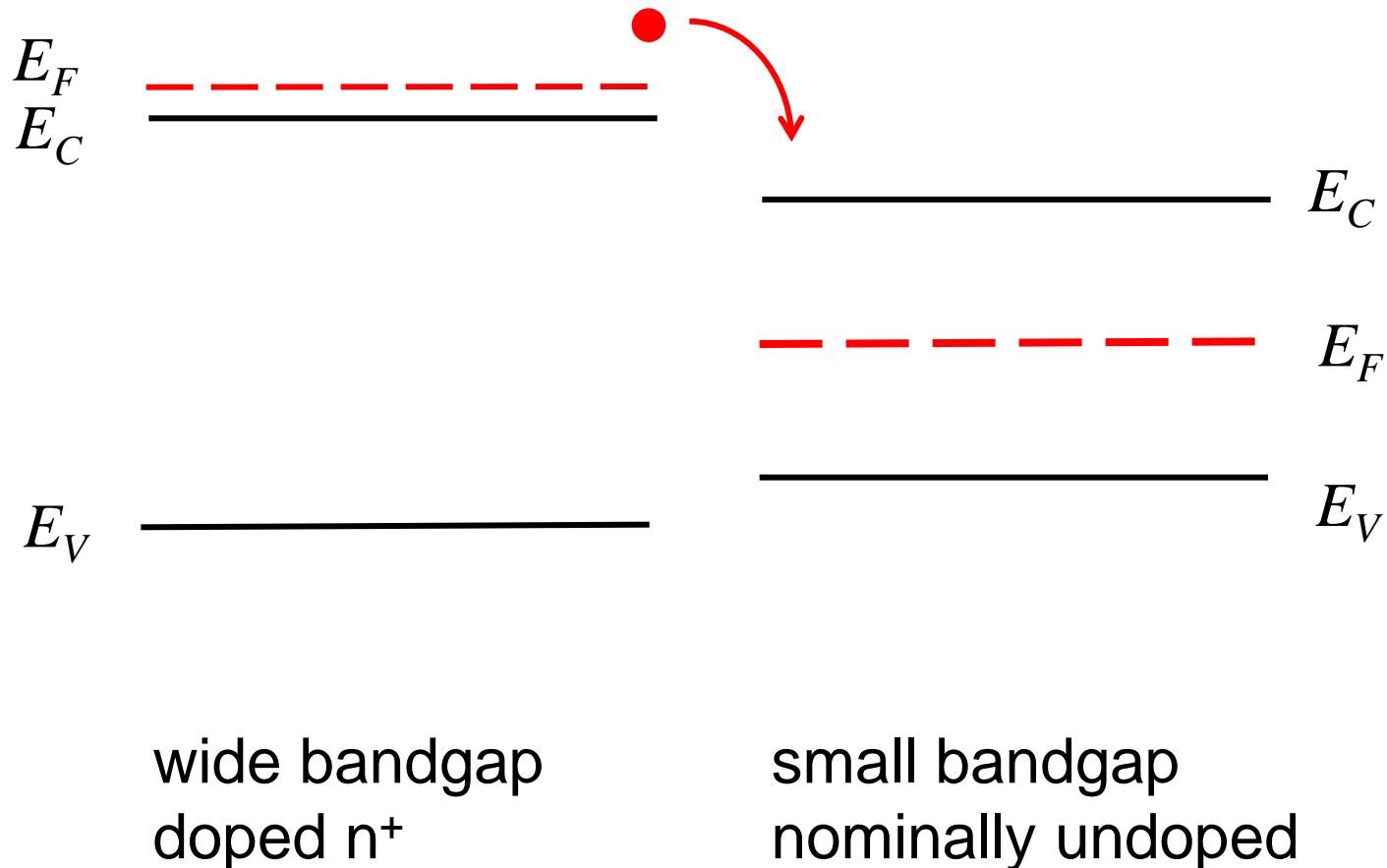


Lundstrom: 2018

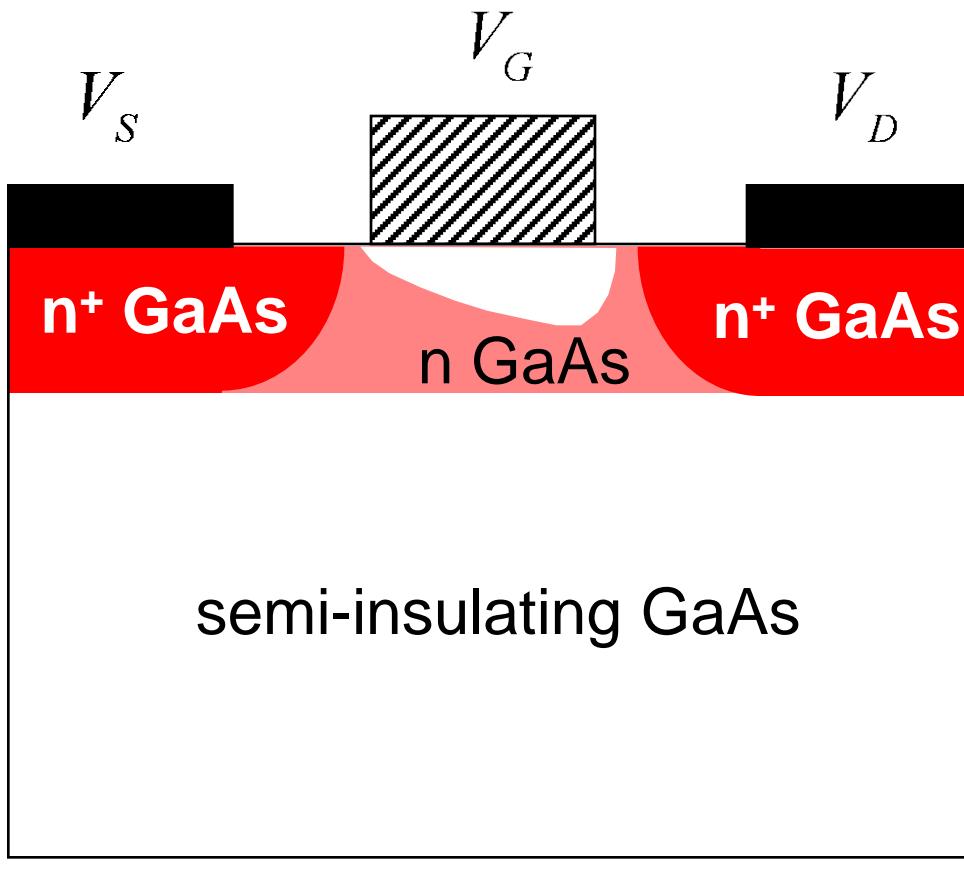
InGaAs HEMT



“Modulation doping”



GaAs MESFET

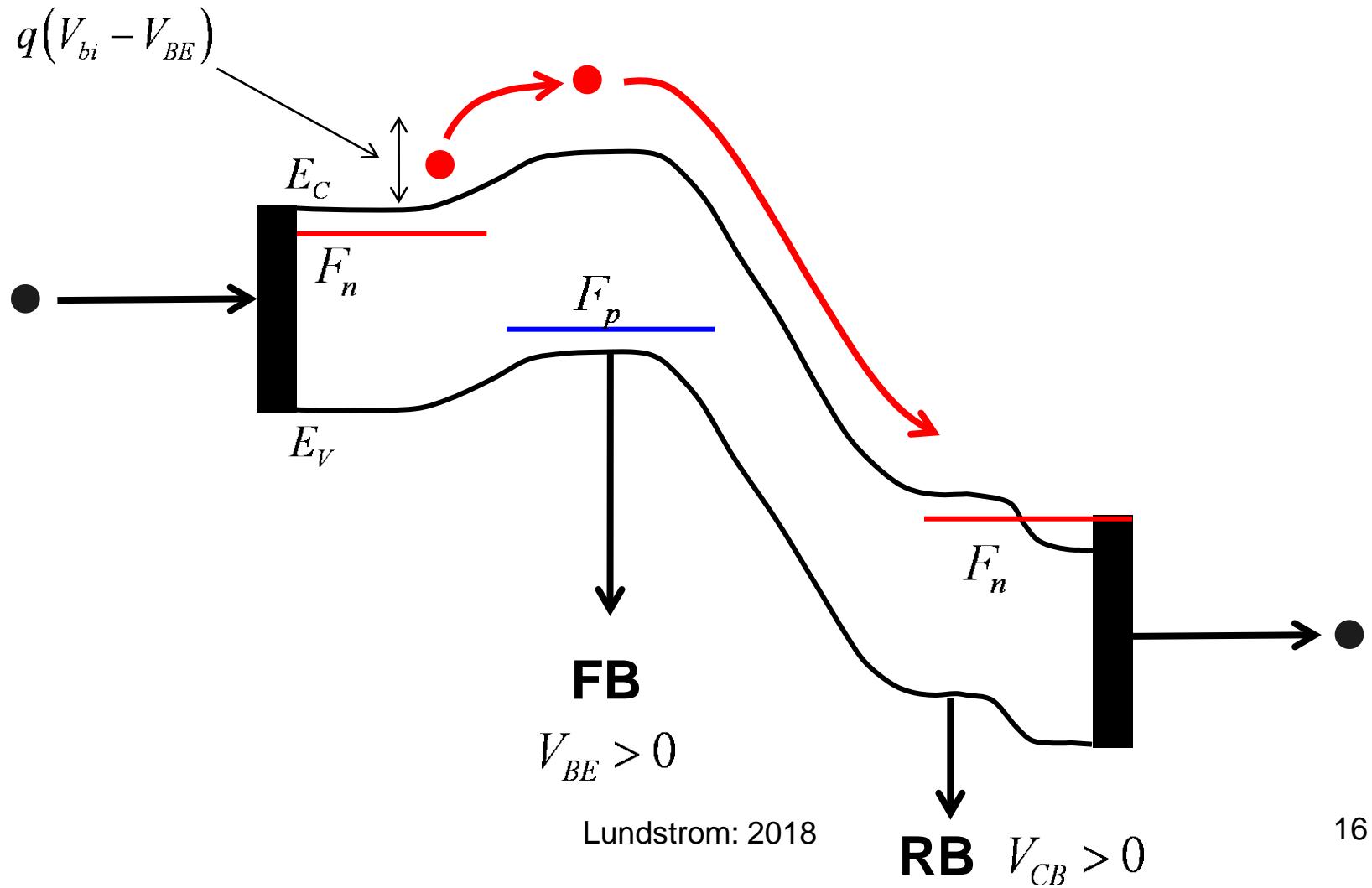


- high mobility
 $\mu_n(10^{14}) \sim 8500 \text{ cm}^2/\text{V}\cdot\text{s}$
- mobility and doping
 $\mu_n(10^{17}) \sim 4700 \text{ cm}^2/\text{V}\cdot\text{s}$
 $\mu_n(10^{18}) \sim 2800 \text{ cm}^2/\text{V}\cdot\text{s}$
- for high g_m , need both charge **and** velocity
- SB gate limits V_G

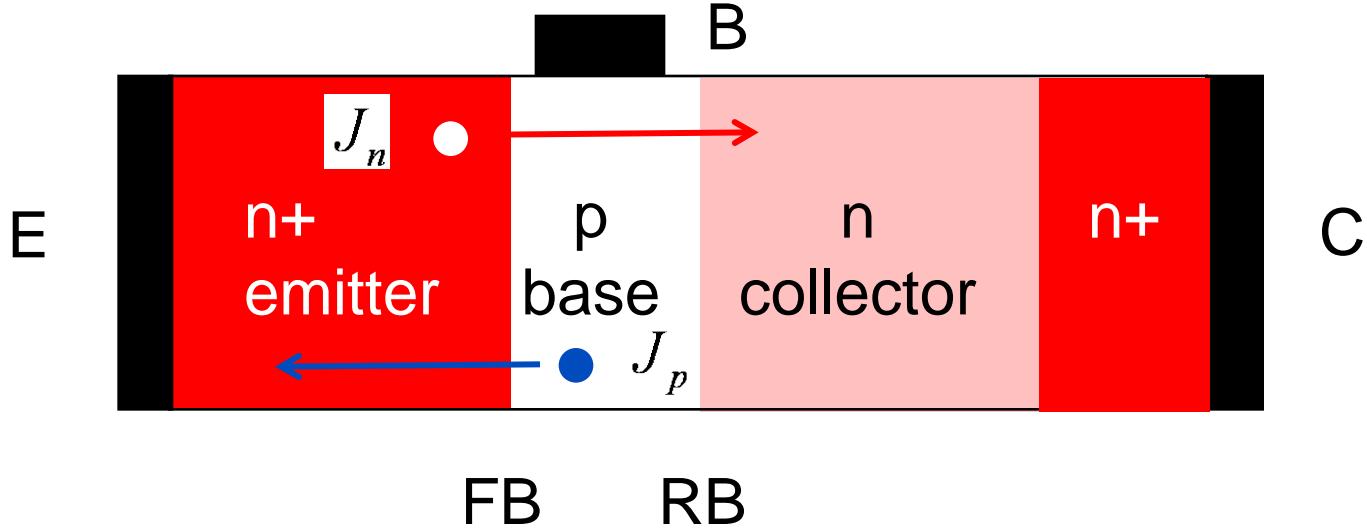
HEMT summary

- 1) III-V FETs are an important technology for high-frequency RF applications.
- 2) Both HEMTs and HBTs have achieved THz speeds.
- 3) HEMTs operate in exactly the same “barrier controlled mode” as Si MOSFETs, so the VS model describes them well.
- 4) III-V HEMTs operate near the ballistic limit.

BJTs are barrier controlled transistors



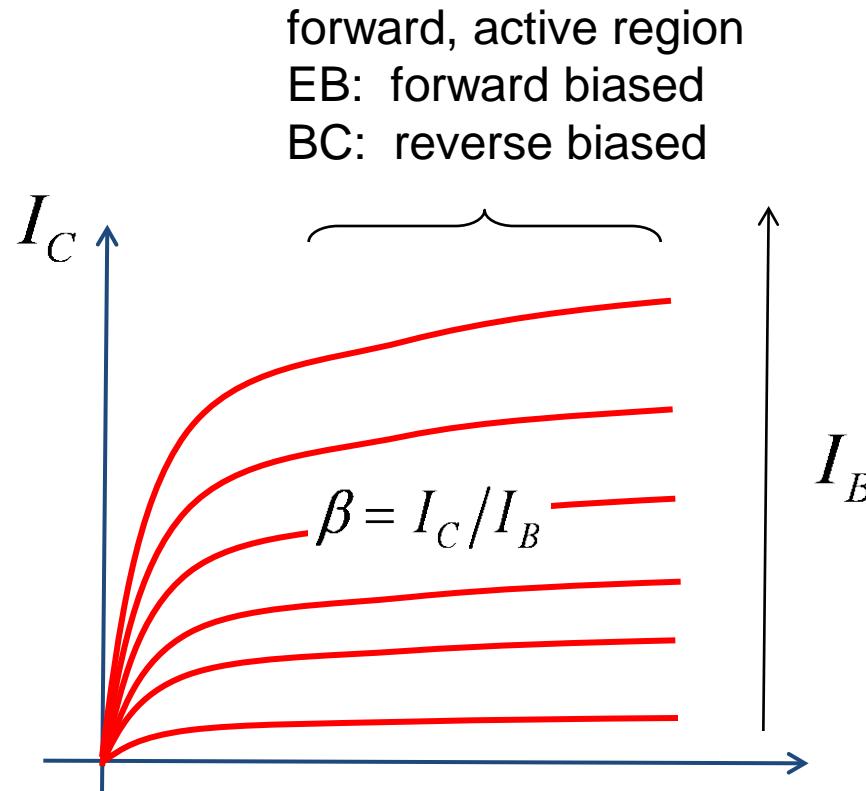
Common emitter current gain



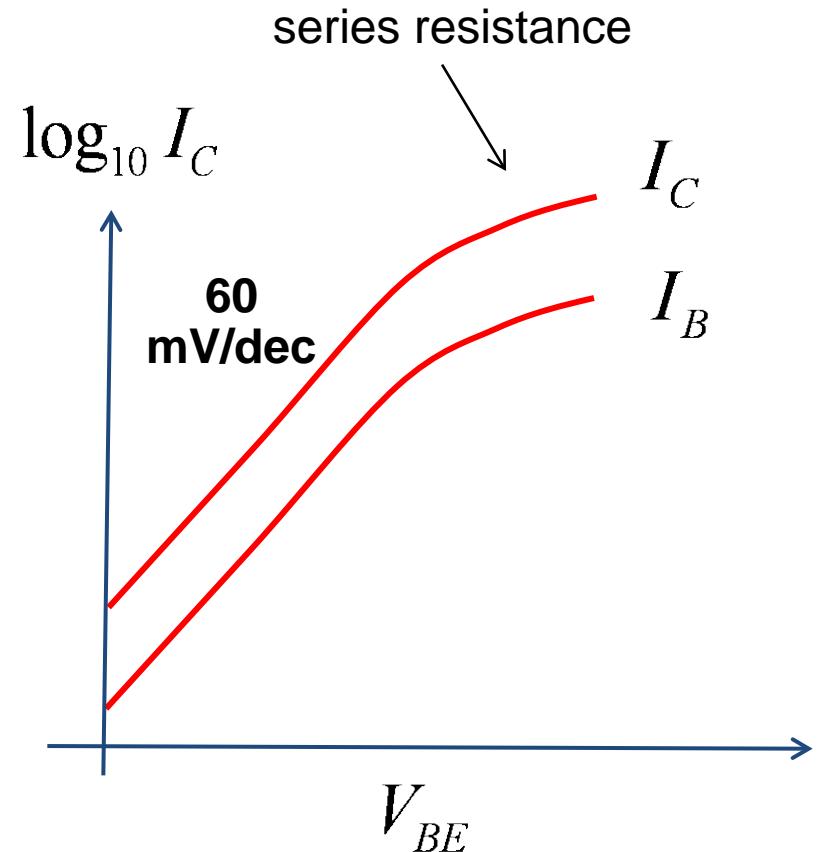
$$\beta = \frac{I_C}{I_B} = \frac{N_{DE}}{N_{AB}} \frac{D_n}{D_p} \frac{W_E}{W_B} \frac{n_{iB}^2}{n_{iE}^2}$$

(In the forward active region of operation)

IV characteristics



output characteristic



Lundstrom: 2018

transfer characteristic₁₈

MOSFET vs. BJT (transconductance)

MOSFET

$$g_m = \frac{I_D}{(V_{GS} - V_{th})}$$

$$I_D = 1 \text{ mA}$$

$$V_{GS} = 1.0 \text{ V} \quad V_{th} = 0.2 \text{ V}$$

$$g_m = 2.5 \text{ mS}$$

BJT

$$g_m = \frac{I_C}{V_T}$$

$$I_C = 1 \text{ mA}$$

$$V_T = 0.026 \text{ V}$$

$$g_m = 40 \frac{\text{mA}}{\text{V}} = 40 \text{ mS}$$

MOSFET vs. BJT (high frequency)

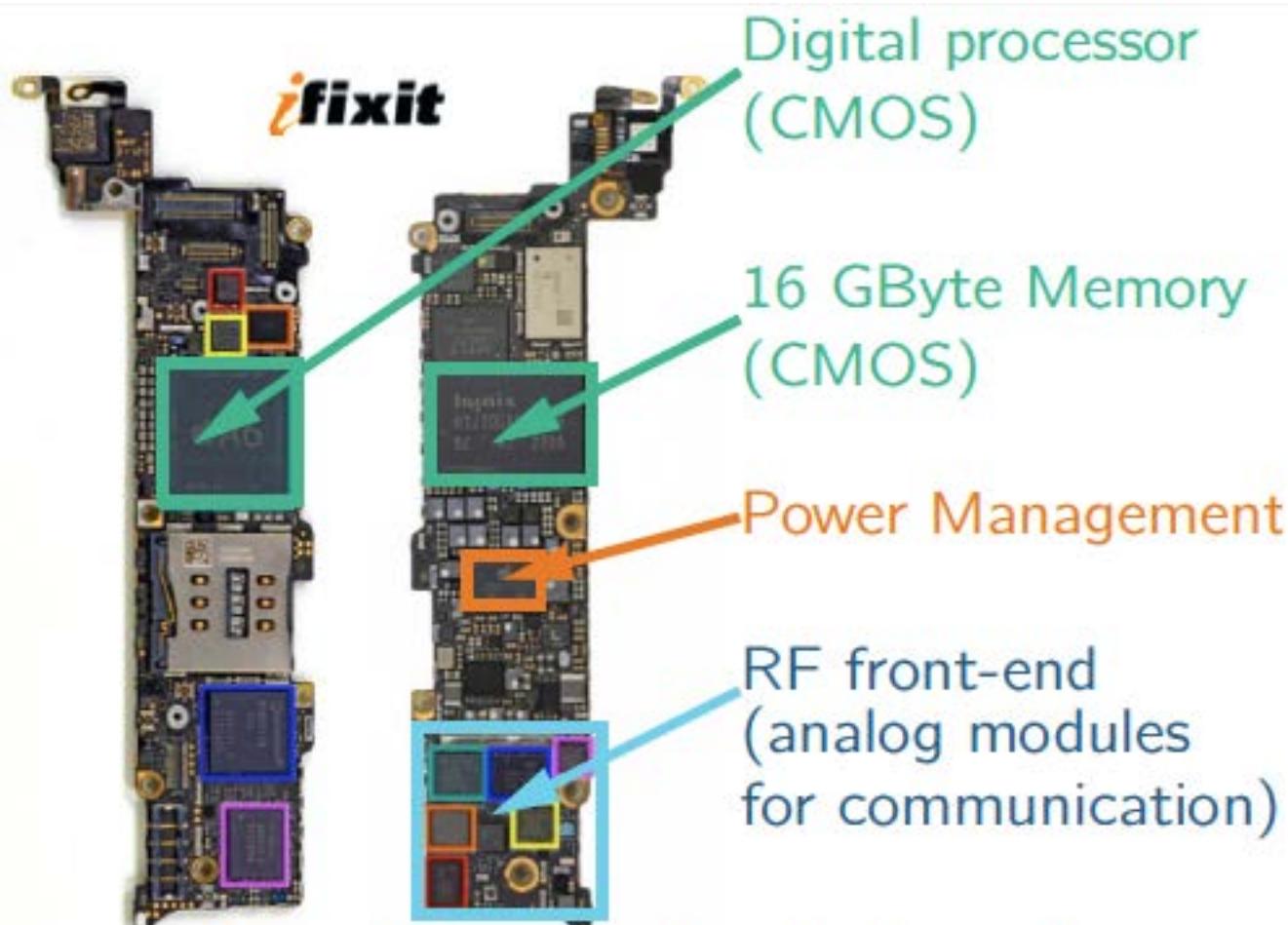
BJTs are superior RF/analog devices. A typical BJT has more than 10X the transconductance of a typical MOSFET.

Both devices can achieve high f_T .
$$f_T = \frac{g_m}{2\pi C_{tot}}$$

BJTs offer higher drive currents (higher g_m)

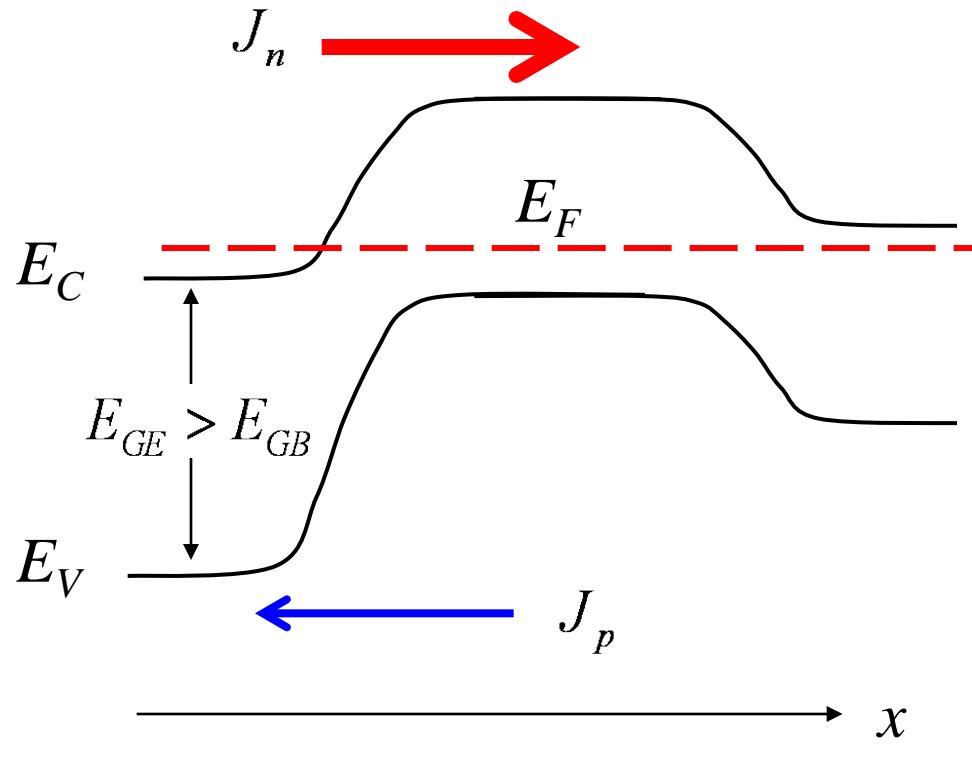
→ Well-suited for RF power amplifiers

SiGe HBTs



Circuit board of an Iphone 5

HBTs



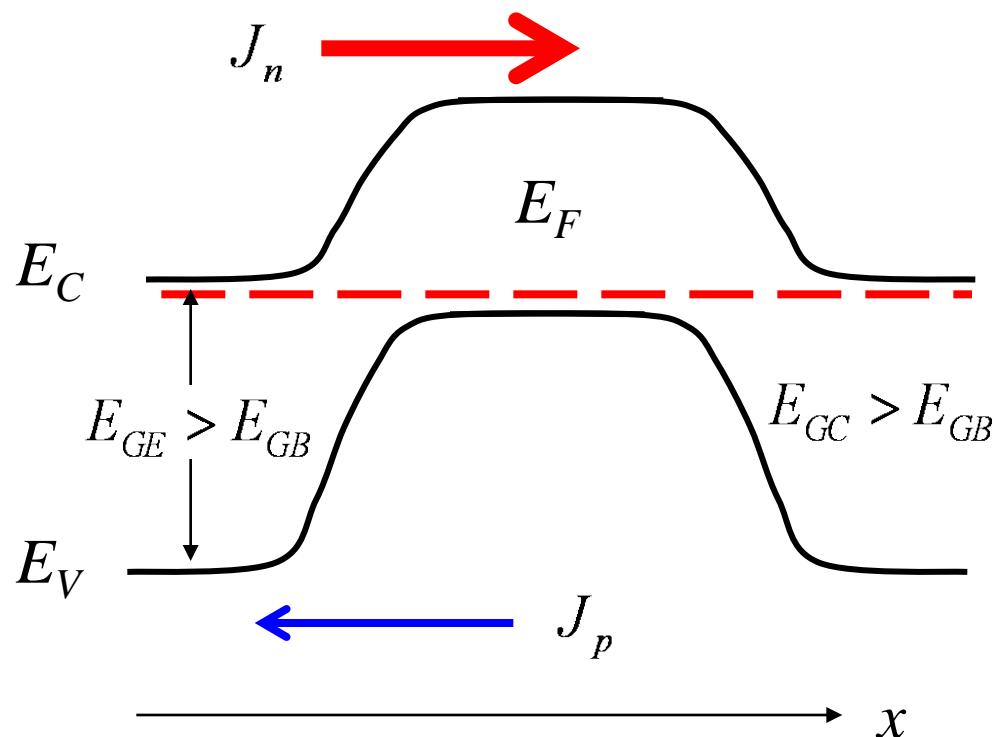
wide bandgap emitter

$$\beta = \frac{N_{DE}}{N_{AB}} \frac{D_n}{D_p} \frac{W_E}{W_B} \frac{n_{iB}^2}{n_{iE}^2}$$

$$n_i^2 = N_C N_V e^{-E_G/k_B T}$$

$$\beta \approx \frac{N_{DE}}{N_{AB}} \frac{D_n}{D_p} \frac{W_E}{W_B} e^{\Delta E_G/k_B T}$$

DHBTs



- symmetrical operation
- reduced collector offset voltage
- higher collector breakdown voltage

HBT summary

- 1) III-V HBTs are an important technology for high-frequency RF power applications.
- 2) SiGe HBTs are a critical technology for wireless electronics
- 3) HBTs offer speeds close to that of HEMTs ($f_T = 600$ GHz, $f_{max} = 1200$ GHz).
- 4) Compared to HEMTs, HBTs deliver higher power and integration density.
- 5) HBTs operate in the same “barrier controlled mode” as Si MOSFETs.

Compact circuit models

- 1) Compact circuit models link semiconductor manufacturing and circuit design.
- 2) Compact circuit models also link device and materials R&D to circuit design and applications.
- 3) Developing compact circuit models requires a good understanding of the device physics, what goes on inside the circuit simulator, and the intended application.
- 4) Many of the lessons learned in developing compact circuit models for MOSFETs can be transferred to other devices.

Course Summary

Unit 1: Transistors, compact models, and circuits

Unit 2: Essential physics of the MOSFET

Unit 3: MOS Electrostatics

Unit 4: Transmission theory of the MOSFET

Unit 5: Additional topics