Virtual Device Simulation

Virtual Process Integration

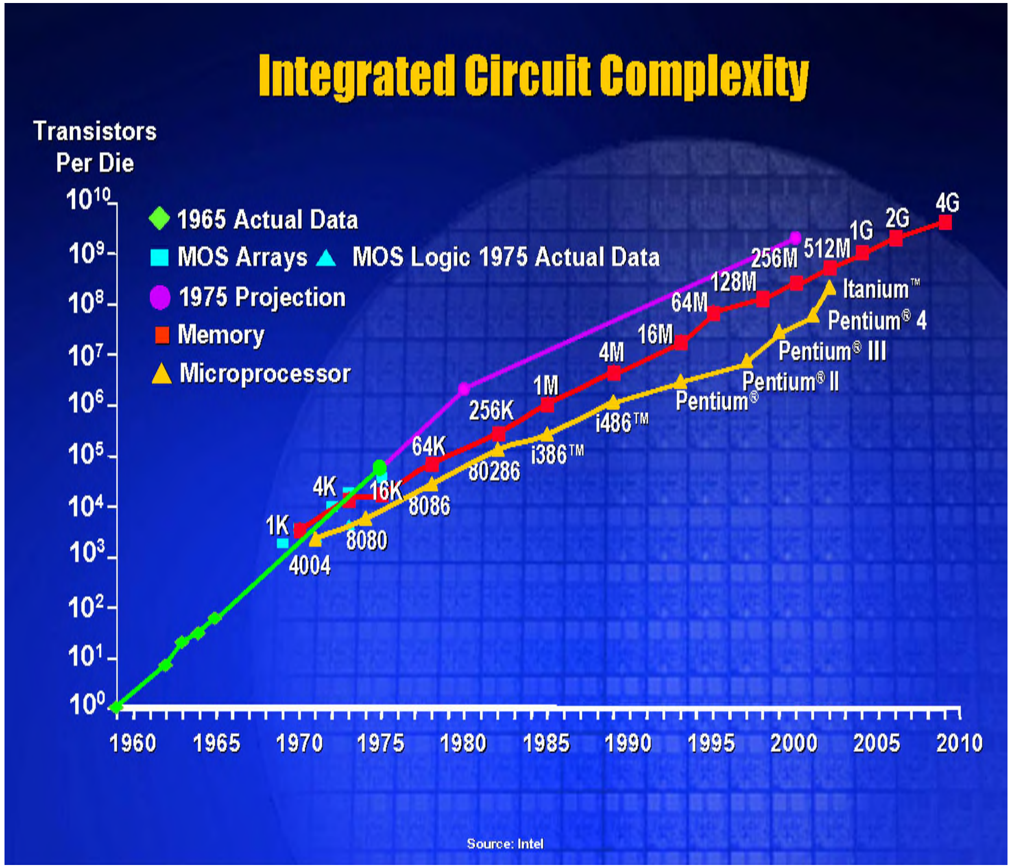
Dr Zhou Xing

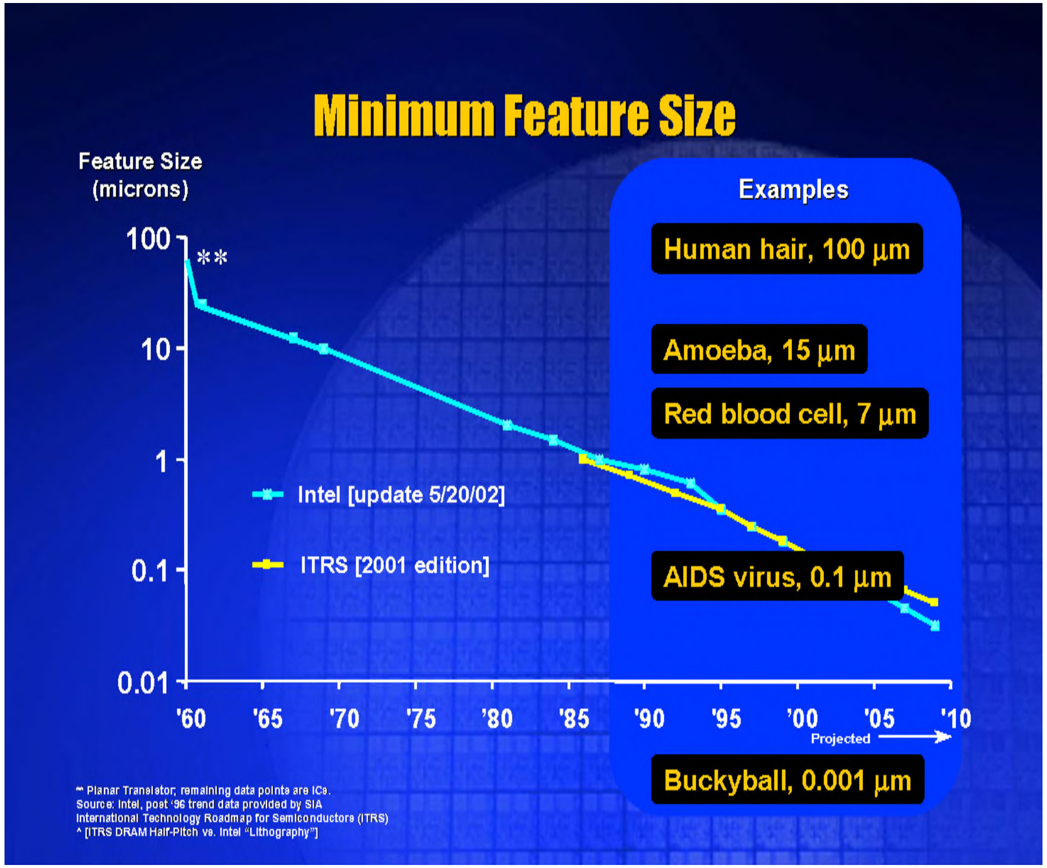
Office: S1-B1c-95
Phone: 6790-4532
Email: exzhou@ntu.edu.sg

Web: https://www3.ntu.edu.sg/home/exzhou/Teaching/EE4613/

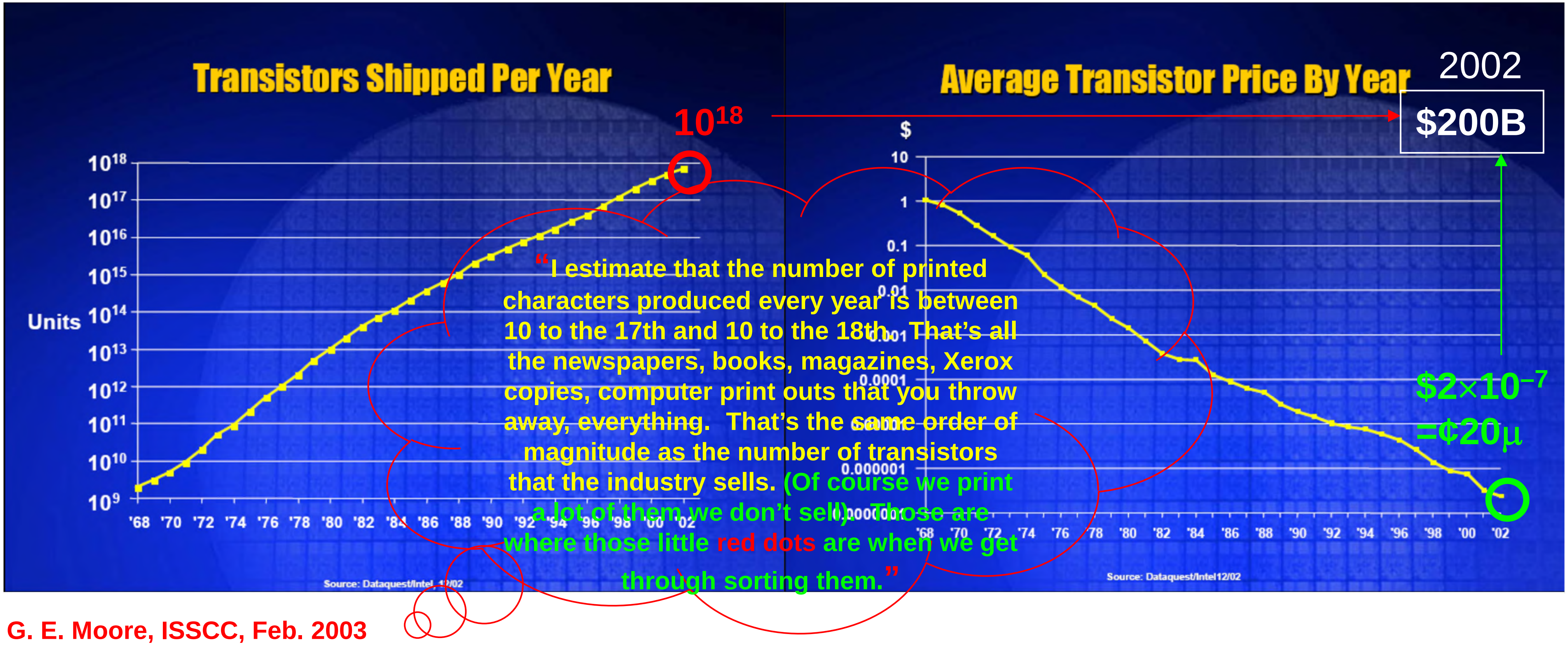
Technology Scaling

Gordon Moore (2003): *"No exponential is forever. But we can delay 'forever'."*





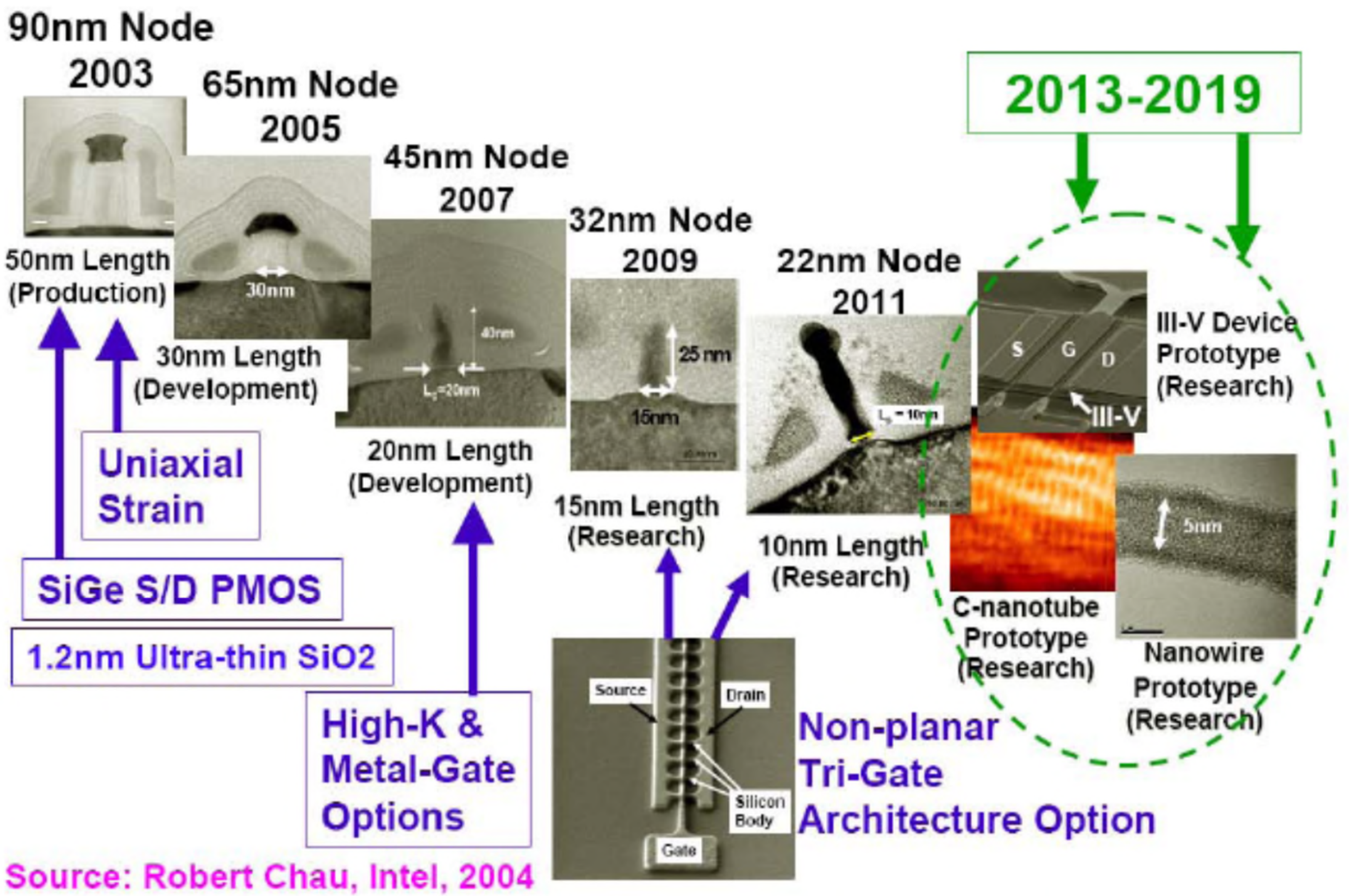
Moore’s Law: Essence of Technology Scaling



G. E. Moore, ISSCC, Feb. 2003

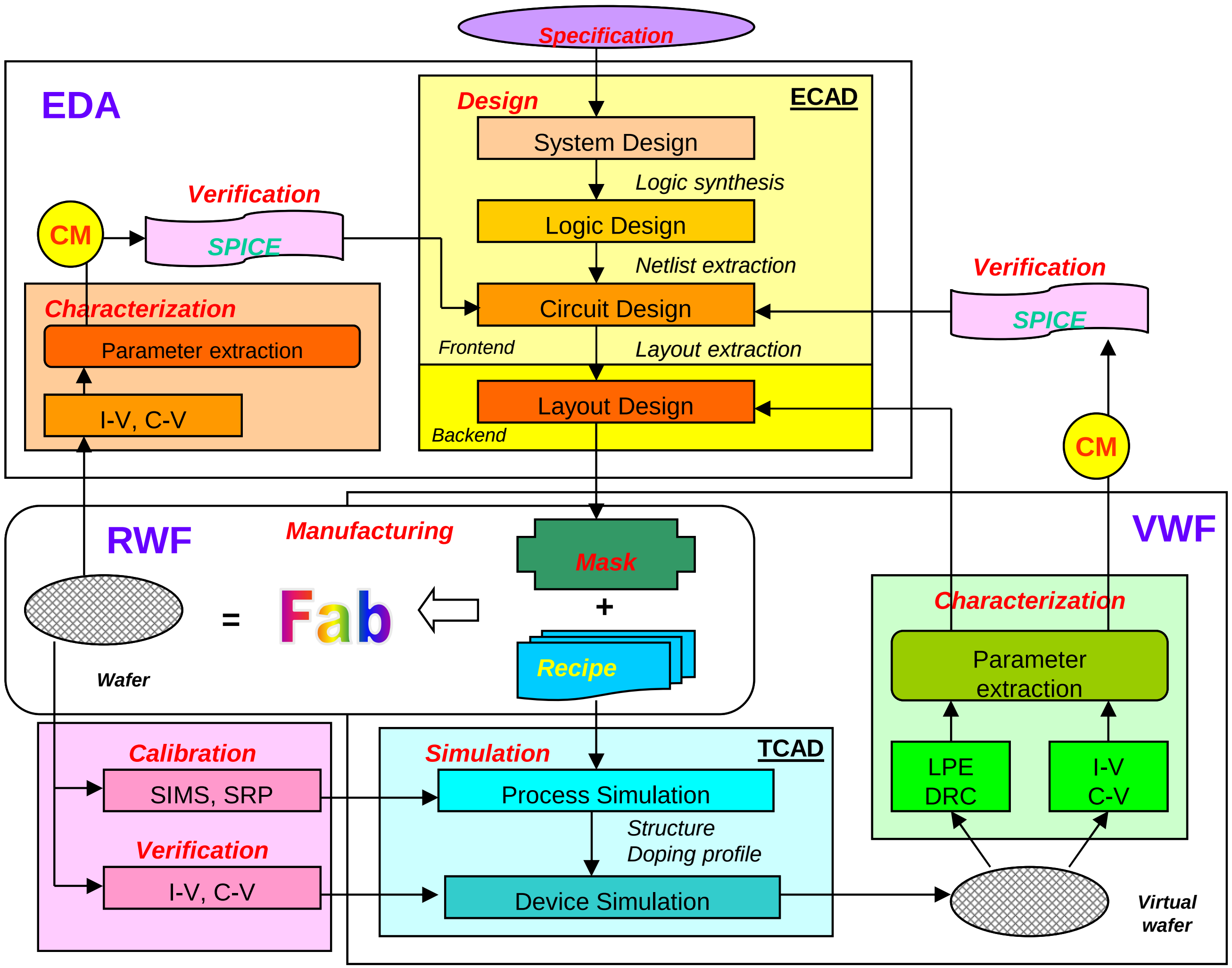
What really dictates technology scaling is the Economics (as “governed” by Moore’s law), and the essence of which is “yield”, which is mainly determined by **variability**/**reliability**.

CMOS Technology Generations

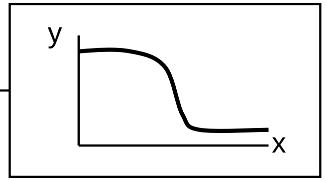
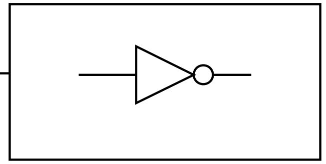
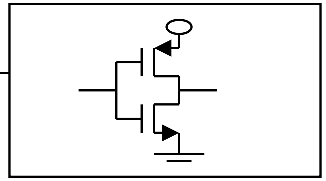
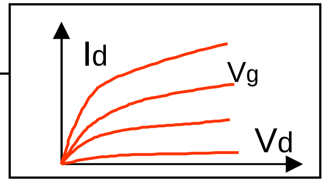
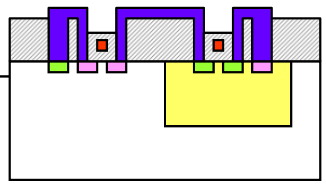


Overall Picture: Chip Design and Wafer Fabrication

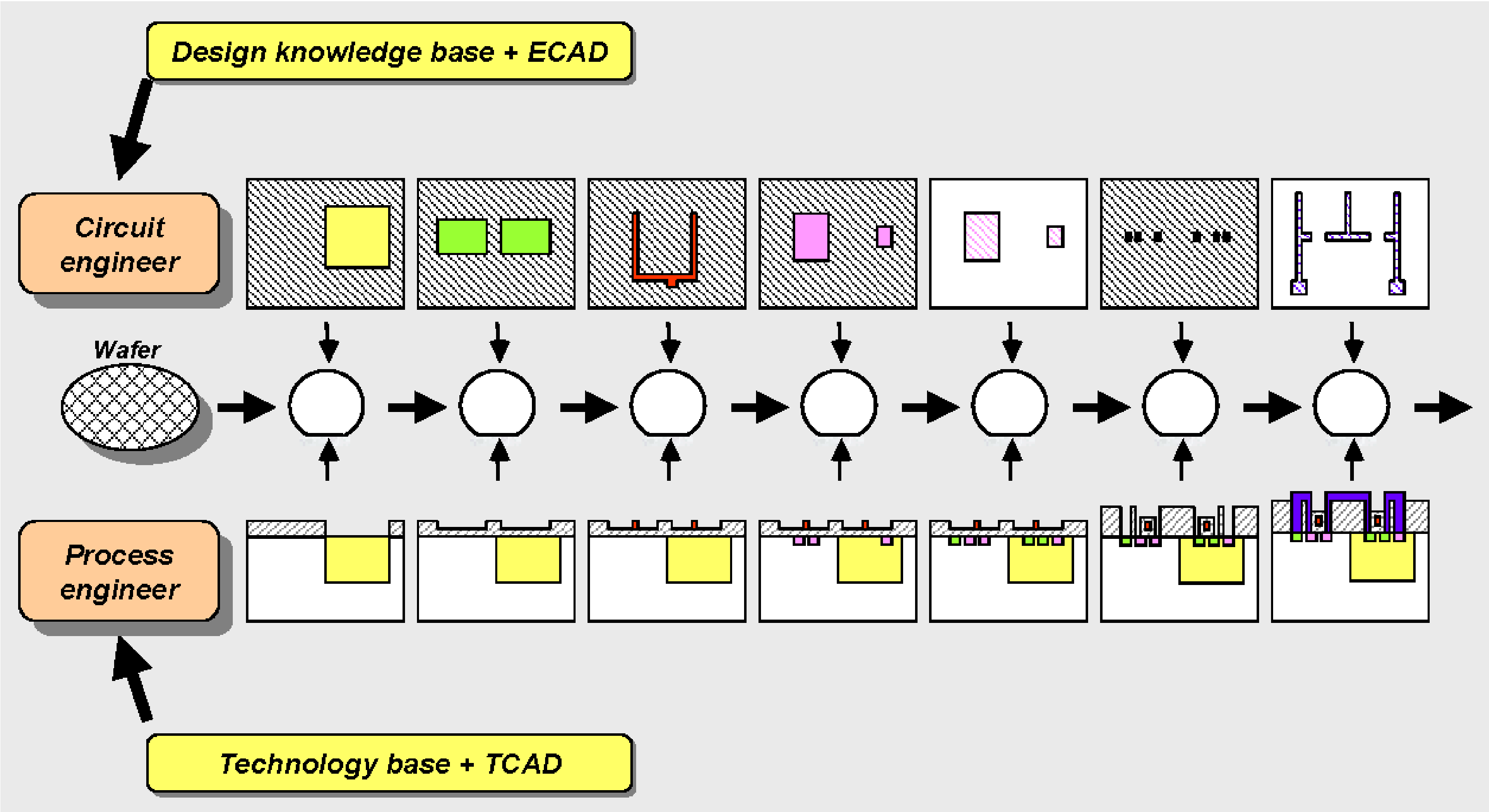
Design–Manufacturing–Characterization–Simulation–Verification



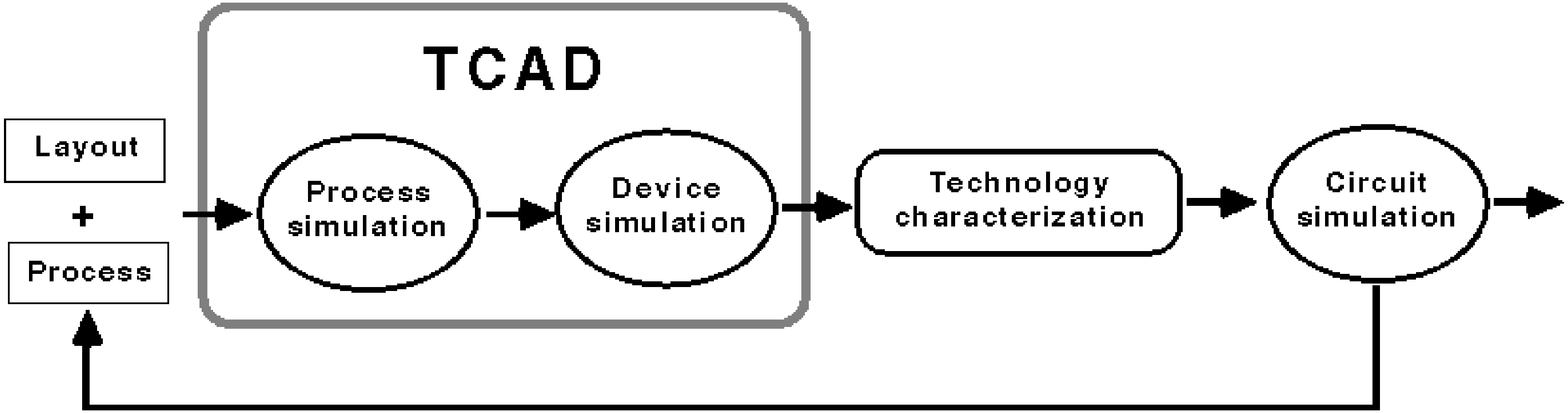
Multi-Level Representation

Level	Representation	Function	Graphical
System	Functional	y = f(x)	
Logic	Boolean	x y 0 1 1 0	
Circuit	Transistor	[I] = [Y][V]	
Device	Analytical	Ids = f(Vgs, Vds, Vbs, W, L, T)	
Technology	Device	$\nabla^2 \psi = -\rho/\varepsilon$ $\frac{d\rho}{dt} + \nabla\cdot \mathbf{J} = G - U$ $\mathbf{J} = q\mu n \nabla\psi + qD\nabla n$	

Layout + Process = Chip

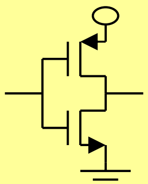
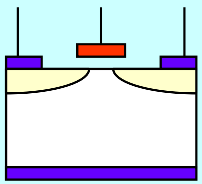
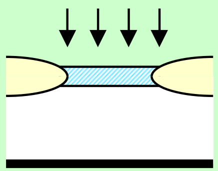


TCAD — Physical Simulation



- <u>Goal</u>: Emulate physical phenomena — *"virtual" wafer fabrication*
  - Semiconductor processing
  - Device operation and electrical characterization
  - Parasitic electrical effects
  - Circuit performance

Target–Variable Relationship

Level	**Variables**	**Targets**
Circuit 	Spice: Model parameters, ... Geometrical: Channel length, width, ... Electrical: Supply voltage, substrate bias, ...	Digital: Delay, rise/fall time, drivability, off-state current, noise margin, ... Analog: Voltage gain, cutoff frequency, slew rate, gain-bandwidth, ...
Device 	Structural: Oxide thickness, junction depth, sheet resistance, ... Doping: Peak/surface concentration, ... Electrical: Supply voltage, substrate bias, ...	Electrical: Threshold, transconductance, subthreshold swing, saturation current, punchthrough current, junction capacitance, lifetime, ... Physical: Potential, field, charge, current, carriers, velocity, ...
Process 	Oxidation: Temperature, time, ambient, ... Implantation: Dose, energy, tilt, damage, ... Diffusion: Defect, stress, OED, TED, ...	Layer: Oxide thickness, junction depth, sheet resistance, ... Profile: Peak/surface concentration, projected range/straggle, ...

Three Ways of Obtaining Device Characteristics

Experimental

- Wafer
- SMU, oscilloscope

Numerical

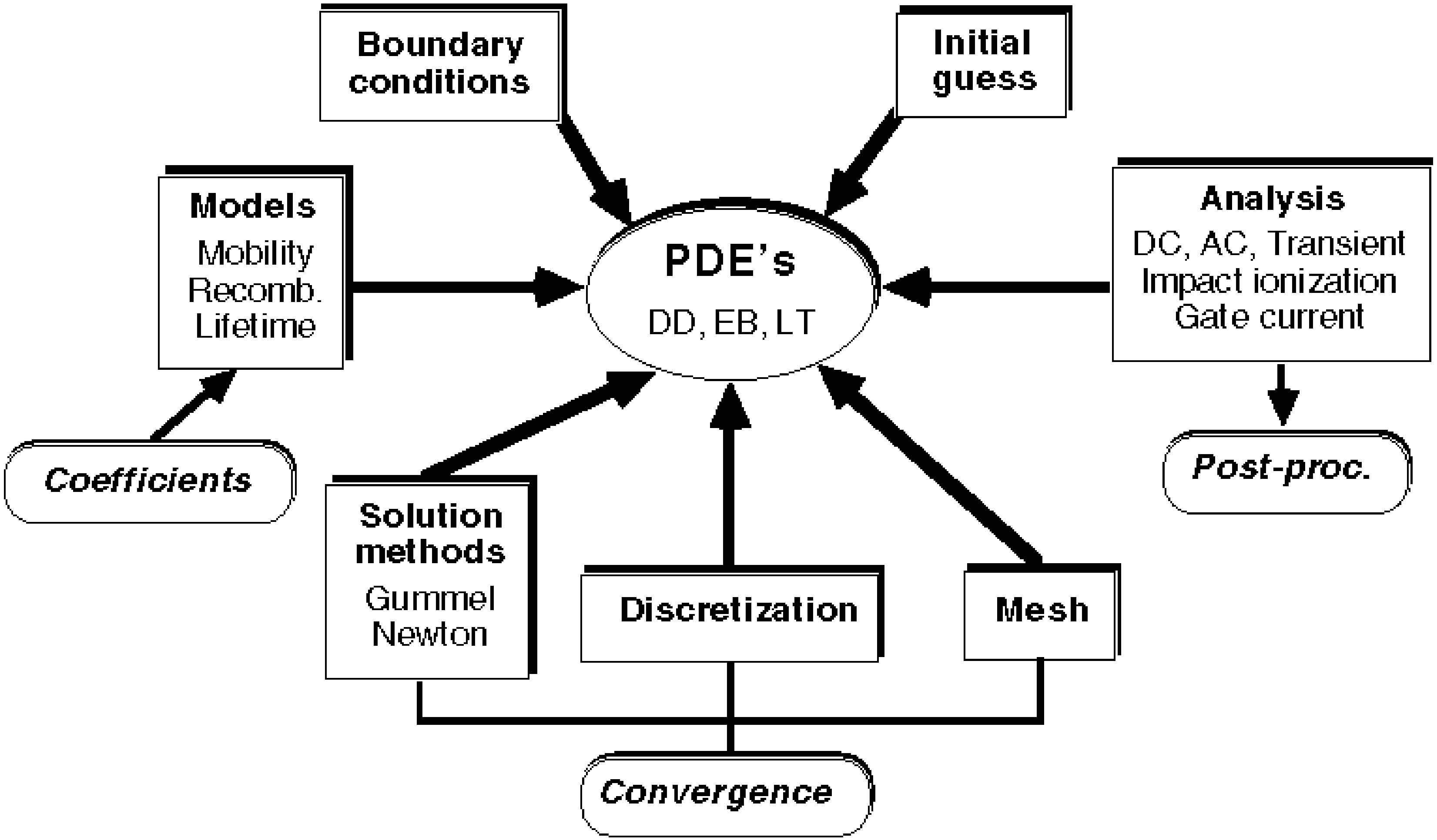
- Partial differential equations + B.C.'s
- 2D/3D grid, finite-element
- MEDICI

Analytical

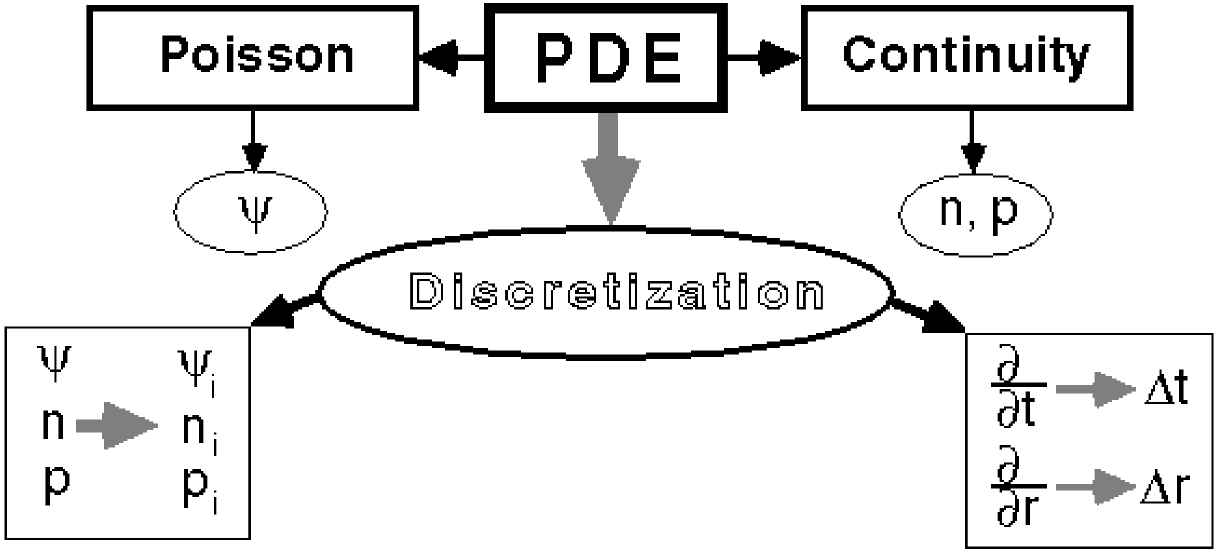
- Closed-form equations
- Matrix, iteration
- SPICE

- ***What are the advantages of the numerical approach to device characterization?***

Basic Ingredients in Device Simulation



Discretization and Solution Methods



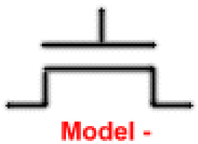
- The PDE's describes the bulk behavior of semiconductor devices
- The continuous functions (ψ, n, p) are represented by vectors of function values at the nodes
- The differential operators are replaced by suitable difference operators
- Solving 3 unknown functions becomes solving for 3*N* unknown real numbers

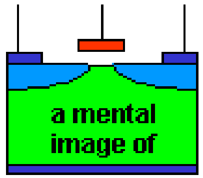
What Is a Model, and Modeling?

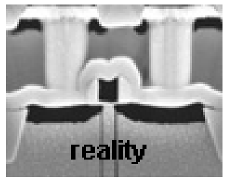
John von Neumann

"The sciences do not try to explain, they hardly even try to interpret, they mainly make models. By a model is meant a mathematical construct which, with the addition of certain verbal interpretations, describes observed phenomena. The justification of such a mathematical construct is solely and precisely that it is expected to work."

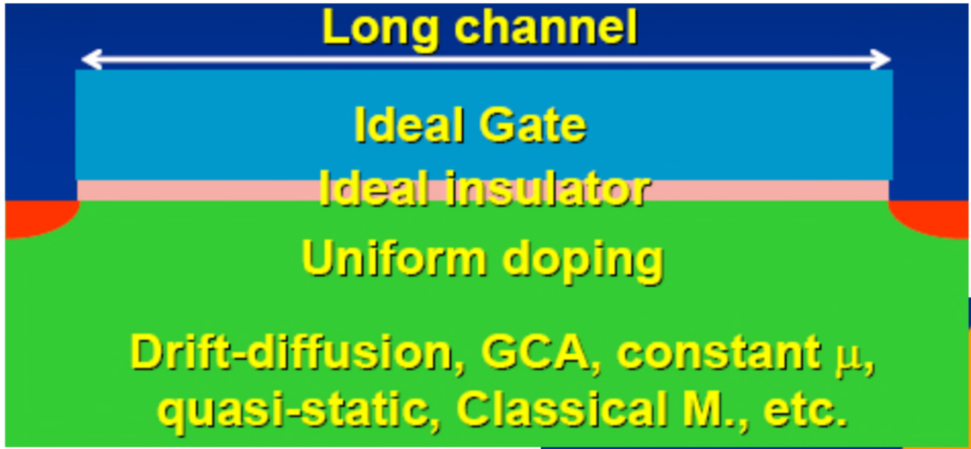
A model is a mental image of reality

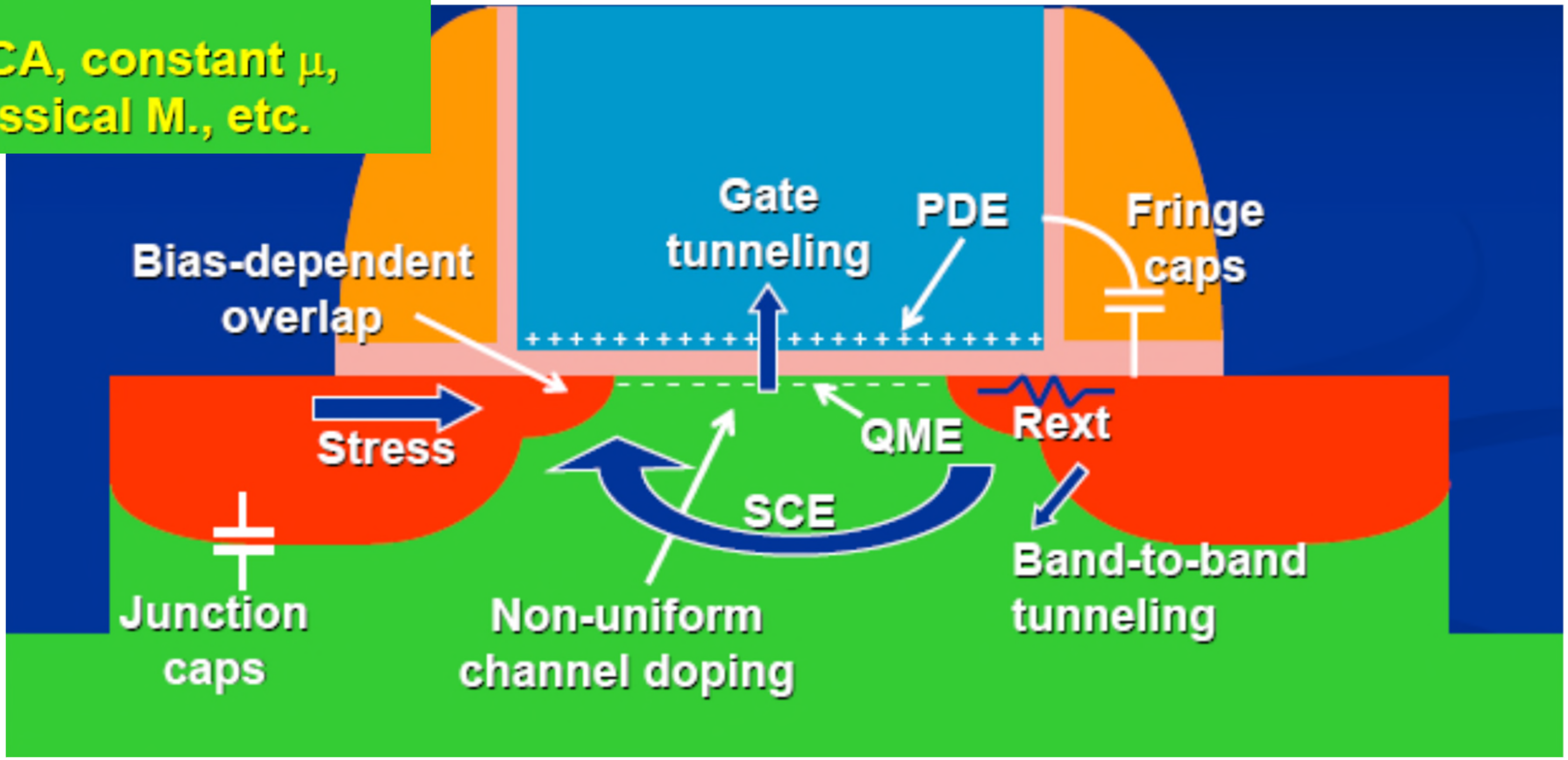






Ideal vs Real MOSFET

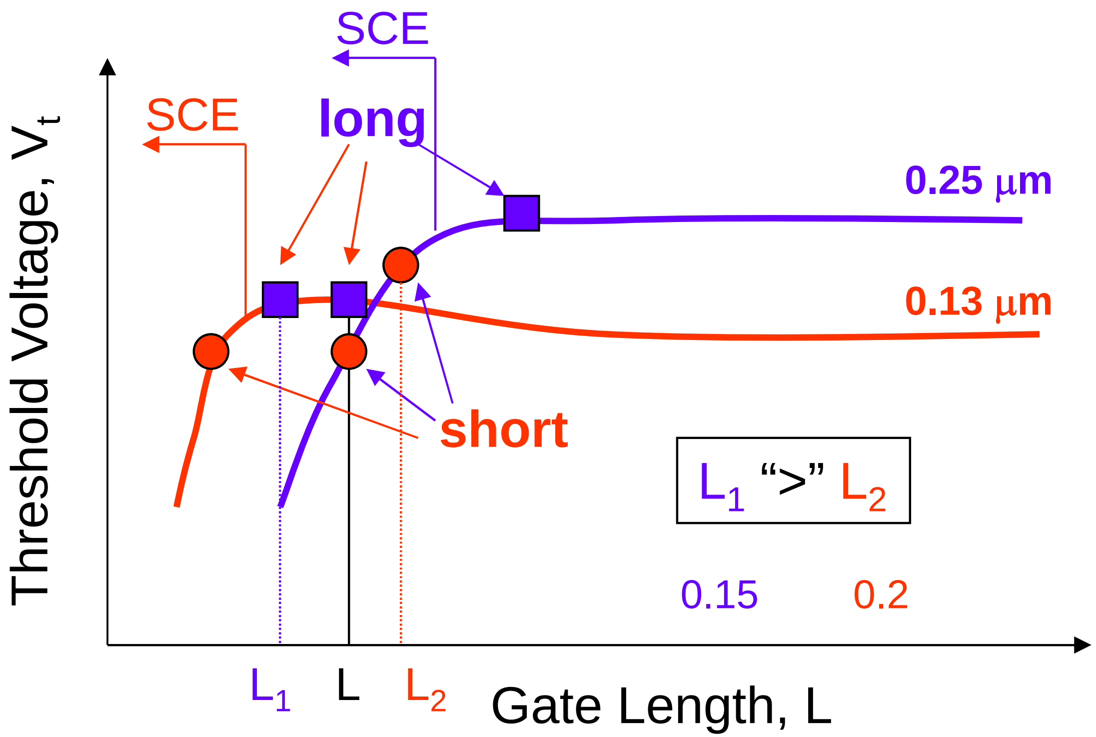




Long-Channel or Short-Channel?

- **Short-channel effect (SCE)** — technology dependent (depends on where the device "sits" on the V_t – L curve, not the actual dimension)

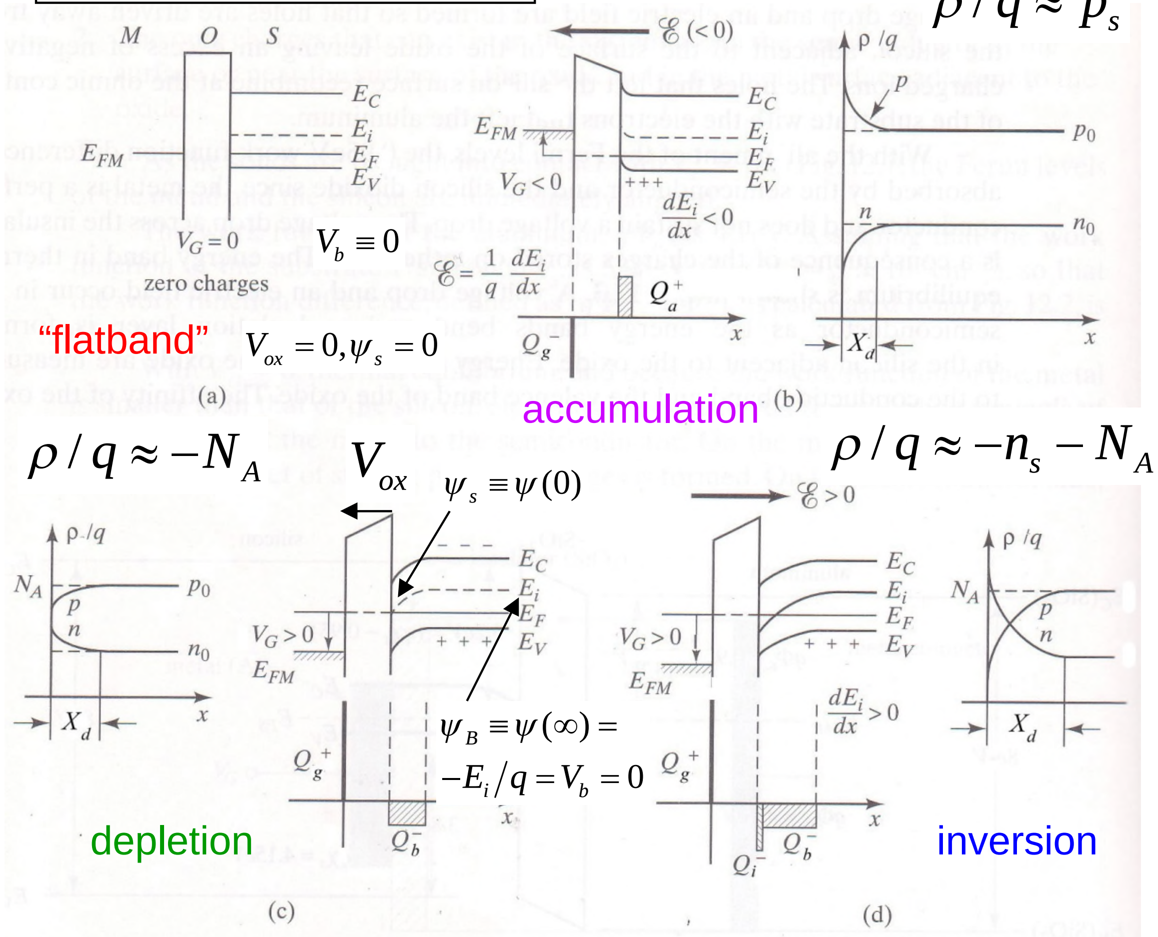
- **Challenge in modeling** — geometry dependence in the SCE regime



Ideal MOS Capacitor and Operation

- **Ideal MOS**: No work function difference between metal and Si, and no charges in SiO_2; thus, all bands are "flat" (V_{ox}=0, ψ_s = 0) at zero gate-bulk bias (V_{gb} = 0).
- In an MOS capacitor, since there's **no (DC) current**, the system is <u>always</u> at **thermal equilibrium**, even <u>with bias</u>.
- **Three regions of operation**
  - ***Accumulation*** ($V_{gb} < 0$; $\psi_s < 0$)
  - ***Depletion*** ($0 < V_{gb} < V_t$; $0 < \psi_s < 2\phi_F$)
  - ***Strong-inversion*** ($V_{gb} > V_t$; $\psi_s > 2\phi_F$)
- **Basic governing equations**
  - ***Voltage balance*** (KVL): $V_{gb} = V_{ox} + \psi_s$
  - ***Charge balance*** (neutrality): $Q_g = -Q_{sc}$
  - ***Gauss law***: $\varepsilon_{ox}\mathcal{E}_{ox} = Q_g$; $\varepsilon_{Si}\mathcal{E}_s = -Q_{sc}$

$$\rho = q(p - n + N_D - N_A)$$



(a) Ideal MOS (p-bulk) at zero bias (flatband); (b) accumulation at $-V_{gb}$; (c) depletion at $+V_{gb}$; (d) strong-inversion at large $+V_{gb}$.

Real MOS: Gate–Bulk Work Function Difference

In real MOS, metal work function and Si work function (depends on doping) won't be the same, given by

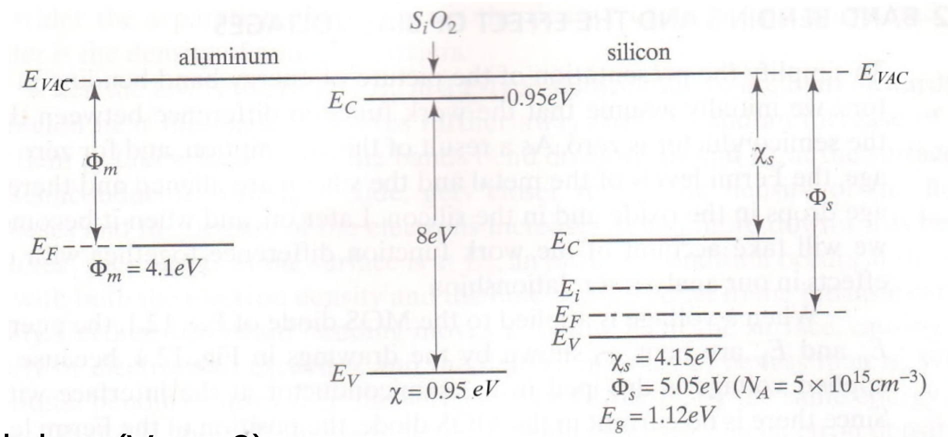
$$\phi_{MS} = \Phi_M - \Phi_S = -(E_{Fm} - E_F)/q$$

After connecting the metal gate to the Si bulk with a wire (V_{gb} = 0), the two Fermi levels will line up. Band bending occurs such that the total potential drop in MOS will balance the initial Fermi level difference:

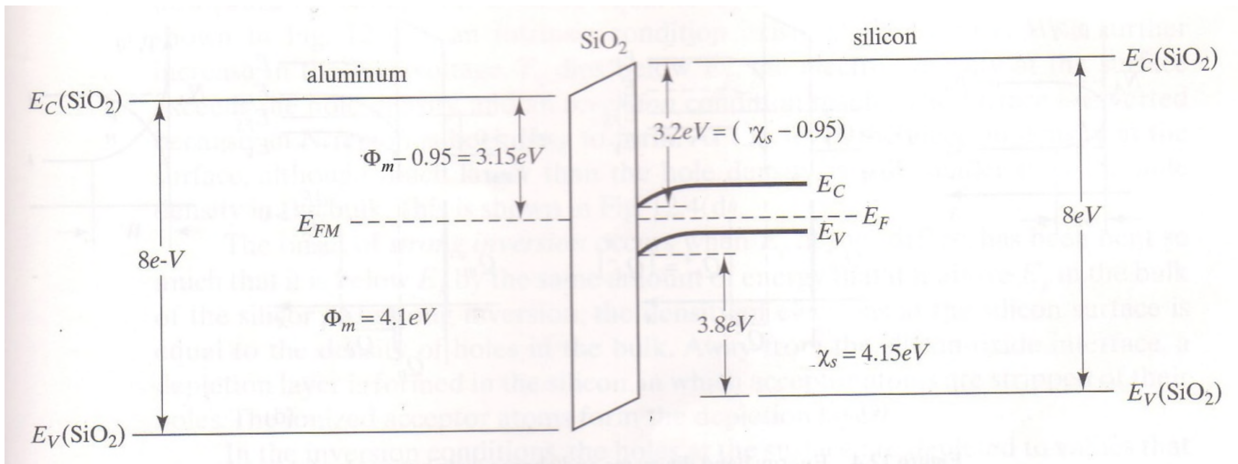
$$V_{ox} + \psi_s = -\phi_{MS} \text{ or } V_{ox} + \psi_s + \phi_{MS} = V_{gb} = 0$$

This is similar to pn junction built-in voltage (V_{bi}) except MOS is always at *equilibrium*, so np=n_i^2 everywhere; while for pn junction, np≠n_i^2 in the depletion region (drift = diffusion).

Before contact:



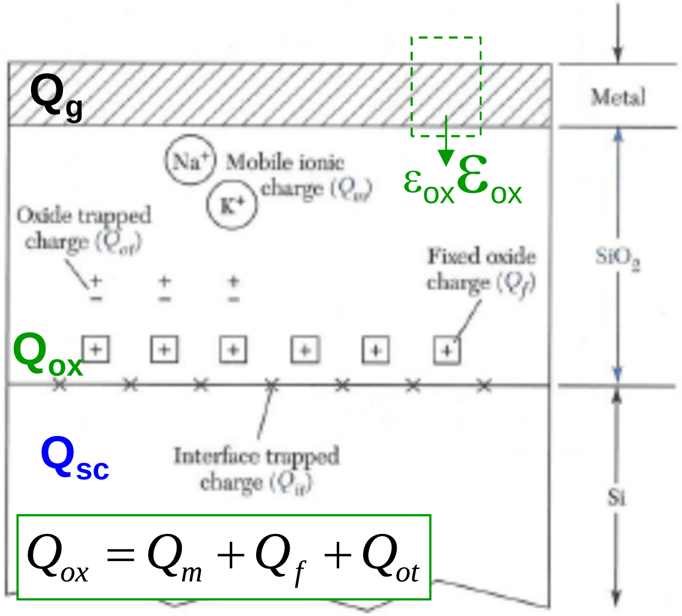
After contact:
at thermal equilibrium (V_{gb} = 0)



Note: Work function, affinity, and band gap are all material properties (function of temperature) and do not change with applied bias.

Real MOS: Mobile/Fixed/Trapped Charges in the Oxide

In real MOS, there are various charges distributed in the oxide during wafer processing, including: mobile ions (Q_m), fixed oxide charge (Q_f), and oxide trapped charge (Q_{ot}), as well as interface trapped charge (Q_{it}). To simplify analysis, we assume all these charges can be represented by an equivalent sheet charge $\mathbf{Q_{ox}}$ (per unit area) at the SiO_2/Si interface (ignoring Q_{it}):



$$Q_{ox} = Q_m + Q_f + Q_{ot}$$

From charge balance, we have $Q_g + Q_{ox} + Q_{sc} = 0$

From potential balance, we have $V_{gb} = V_{ox} + \psi_s + \phi_{MS}$

From Gauss law, we have $\varepsilon_{ox}\mathcal{E}_{ox} = Q_g$, with $\mathcal{E}_{ox} = V_{ox}/T_{ox}$ and $C_{ox} = \varepsilon_{ox}/T_{ox}$ $\longrightarrow$

$$V_{ox} = T_{ox}\mathcal{E}_{ox} = T_{ox}\left(Q_g/\varepsilon_{ox}\right) = Q_g/C_{ox} = -\left(Q_{ox} + Q_{sc}\right)/C_{ox}$$

Define **Flatband voltage**: $V_{FB} = \phi_{MS} - Q_{ox}/C_{ox} = \phi_{MS} - \left(Q_m + Q_f + Q_{ot}\right)/C_{ox}$

$$\therefore\ V_{gb} = -\left(Q_{ox} + Q_{sc}\right)/C_{ox} + \psi_s + \phi_{MS} = V_{FB} - Q_{sc}/C_{ox} + \psi_s$$

So, non-idealities due to work function difference and oxide charge can be “absorbed” into flatband, if we ***define*** $V_{gf} \equiv V_{gb} - V_{FB} = V_{ox} + \psi_s$. When $V_{gf} = 0$ ($V_{gb} = V_{FB}$), we have $V_{ox} = 0, \psi_s = 0$, i.e., bands are flat. This is similar to ideal MOS ($V_{FB} = 0$): when $V_{gb} = 0$, we have $V_{ox} = 0, \psi_s = 0$.

Field and Potential Distribution in Depletion Region

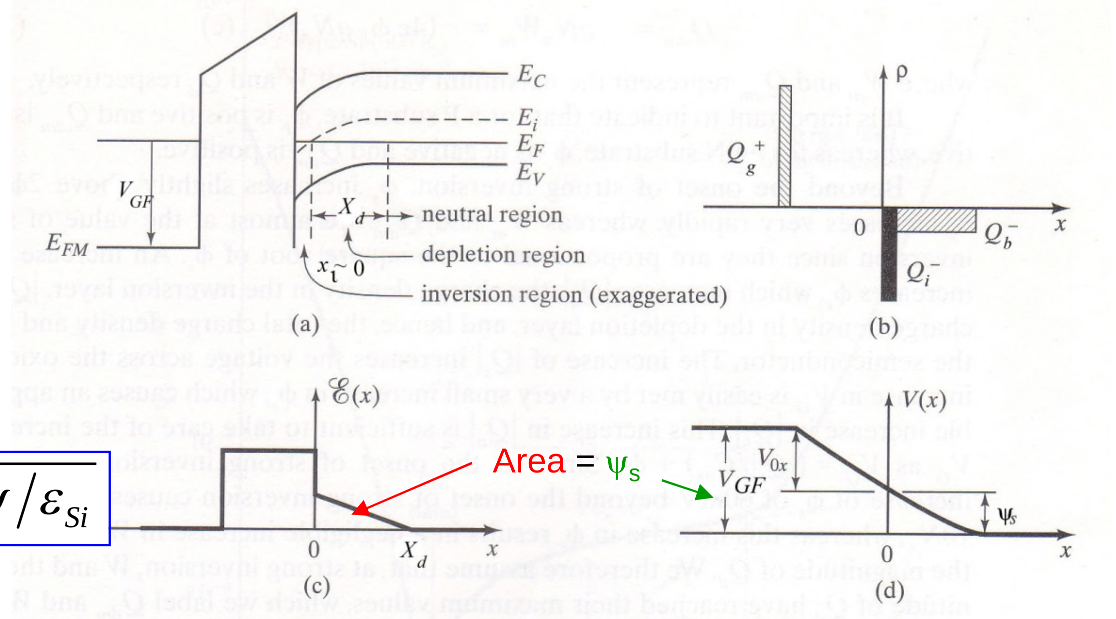
Full-depletion approximation

In depletion region, both electrons and holes can be ignored compared to N_A, so space charge (per unit area) is: $Q_{sc} = -qN_A X_d$

Poisson equation is given by

$$d^2\psi/dx^2 = -\rho/\varepsilon_{Si} = qN_A/\varepsilon_{Si} \qquad \frac{d^2\psi}{dx^2} = \mathcal{E}_x \frac{d\mathcal{E}_x}{d\psi}$$

$$\frac{\mathcal{E}_x^2}{2} = \int_0^{\mathcal{E}_x} \mathcal{E}_x d\mathcal{E}_x = \int_0^{\psi} \frac{d^2\psi}{dx^2} d\psi = \frac{qN_A\psi}{\varepsilon_{Si}} \rightarrow \boxed{\mathcal{E}_x(x) = \sqrt{2qN_A\psi/\varepsilon_{Si}}}$$



So, surface field: $\mathcal{E}_s \equiv \mathcal{E}_x(0) = \sqrt{2qN_A\psi_s/\varepsilon_{Si}}$

From **Gauss law** (at Si surface): $-\varepsilon_{Si}\mathcal{E}_s = Q_{sc}$

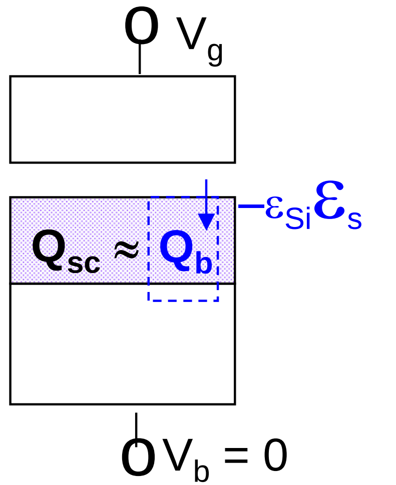
$$\rightarrow \mathcal{E}_s = \sqrt{2qN_A\psi_s/\varepsilon_{Si}} = -Q_{sc}/\varepsilon_{Si} = qN_A X_d/\varepsilon_{Si} \rightarrow$$

So, **depletion width** (function of V_{gb}):

$$\boxed{X_d = \sqrt{2\varepsilon_{Si}\psi_s/qN_A}}$$

$$\downarrow$$

$$\psi_s = (qN_A/2\varepsilon_{Si})X_d^2$$



Potential distribution

$$\mathcal{E}_x = -\frac{d\psi}{dx} \qquad \left(1/\sqrt{\psi}\right)d\psi = -\sqrt{2qN_A/\varepsilon_{Si}}\,dx$$

From 2nd integral of Poisson:

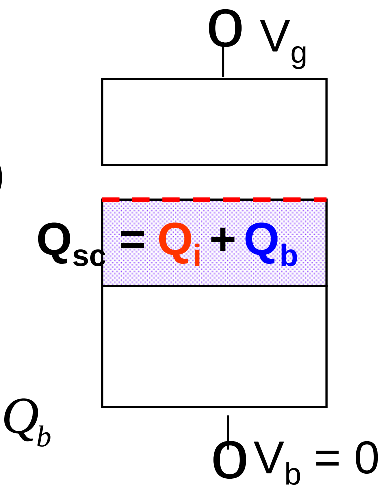
$$\int_0^{\psi}\left(1/\sqrt{\psi}\right)d\psi = -\int_{X_d}^{x}\sqrt{2qN_A/\varepsilon_{Si}}\,dx \rightarrow \boxed{\psi(x) = (qN_A/2\varepsilon_{Si})(X_d - x)^2}$$

$x = 0$

Charge-Sheet Approximation and Threshold Voltage

Charge-sheet approximation (CSA) in strong inversion

In strong inversion, electrons (n) are comparable to N_A, so the space charge (per unit area) Q_{sc} includes both inversion charge $Q_i = -qn*x_i$ (in a layer of x_i) and depletion charge $Q_b = -qN_A*X_d$, and they are in principle not separable. In the **charge-sheet approximation**, we assume Q_i is a sheet of charge with negligible thickness ($x_i \approx 0$); and under *full-depletion approximation*, in the depletion region of thickness X_d, there is no free carriers. Then: $Q_{sc} = Q_i + Q_b$



Threshold voltage definition and maximum depletion width

We define the **threshold voltage** (V_t) to be <u>the gate-to-bulk voltage</u> (V_{gb}) <u>at which surface potential is equal to twice of the bulk Fermi potential</u> ($2\phi_F$). For $V_{gb} > V_t$, we assume surface potential is "pinned" to $2\phi_F$, and depletion layer reaches its maximum value X_{dm}. This is due to sufficient electrons to "screen" the gate electric field after the onset of strong inversion.

For $V_{gb} \geq V_t$, **maximum depletion width**:

$$X_{dm} = \sqrt{\frac{4\varepsilon_{Si}\phi_F}{qN_A}}$$

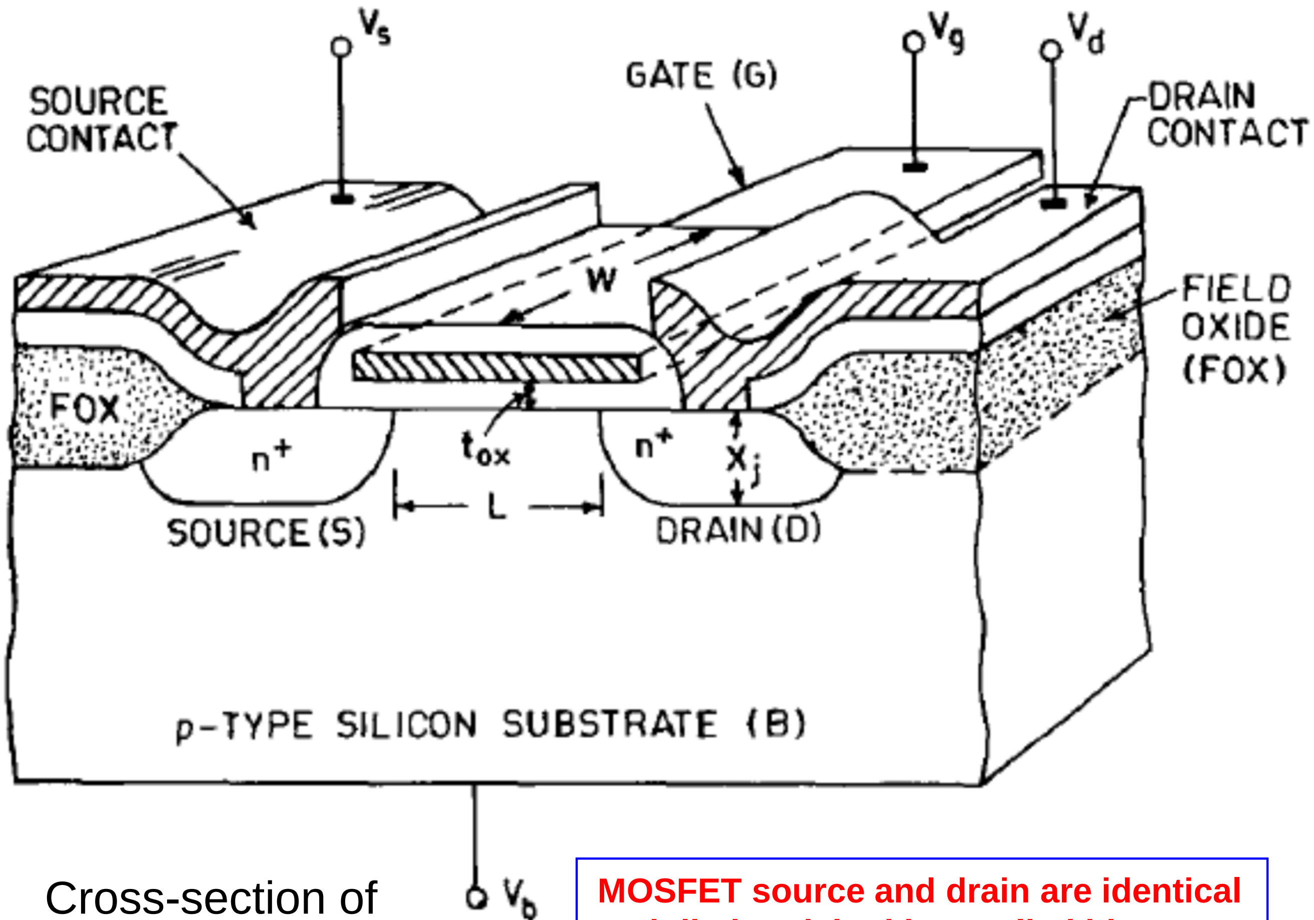
Recall potential balance: $V_{gf} \equiv V_{gb} - V_{FB} = V_{ox} + \psi_s = -Q_{sc}/C_{ox} + \psi_s$ and $Q_{sc} \approx Q_b = -qN_A X_d = -\sqrt{2q\varepsilon_{Si}N_A\psi_s}$ (CSA)

$$\therefore \quad V_t \equiv V_{gb}\Big|_{\psi_s = 2\phi_F} = V_{FB} - Q_b\left(2\phi_F\right)/C_{ox} + 2\phi_F = V_{FB} + \Upsilon\sqrt{2\phi_F} + 2\phi_F$$

where $\Upsilon = \sqrt{2q\varepsilon_{Si}N_A}/C_{ox}$ is the **body factor**.

MOSFET Structure and Naming

- **Four terminals**
  - ***Gate***: control current flow in an 'inversion' ***channel***
  - ***Source***: supply carriers
  - ***Drain***: collect carriers
  - ***Bulk***: silicon body
- **Intrinsic device parameters**
  - ***W/L***: channel width/length
  - T_{ox}: oxide thickness
  - N_A: substrate (bulk/body) doping
- **2D problem – two 1D solutions**
  - ***Electrostatics by vertical field*** ($\sim V_{gb}/T_{ox}$)
  - ***Current transport by lateral field*** ($\sim V_{ds}/L$)
  - ***Key***: Inversion-carrier "imref split" $V_{cb}(y) = [E_{Fn}(y) - E_F]/q \quad (E_{Fp} = E_F)$

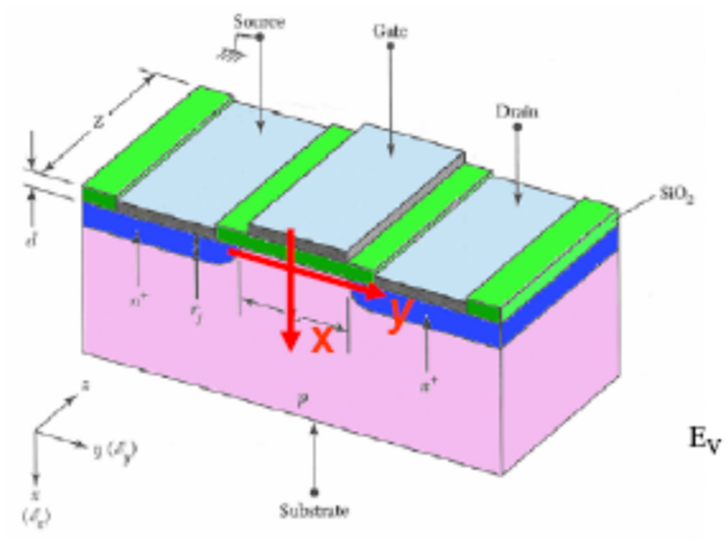


Cross-section of an NMOS with p-type Si substrate.

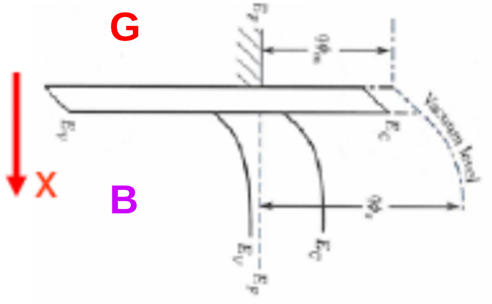
MOSFET source and drain are identical and distinguished by applied bias.

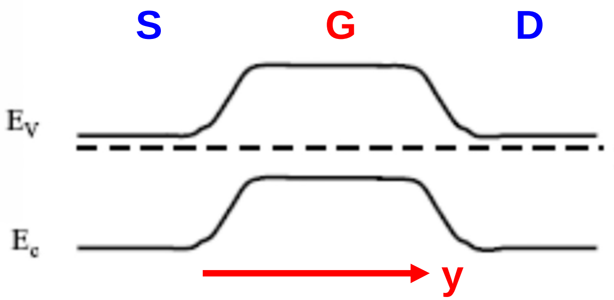
$$I_{ds} = f(V_{gs}, V_{ds}, V_{bs})$$

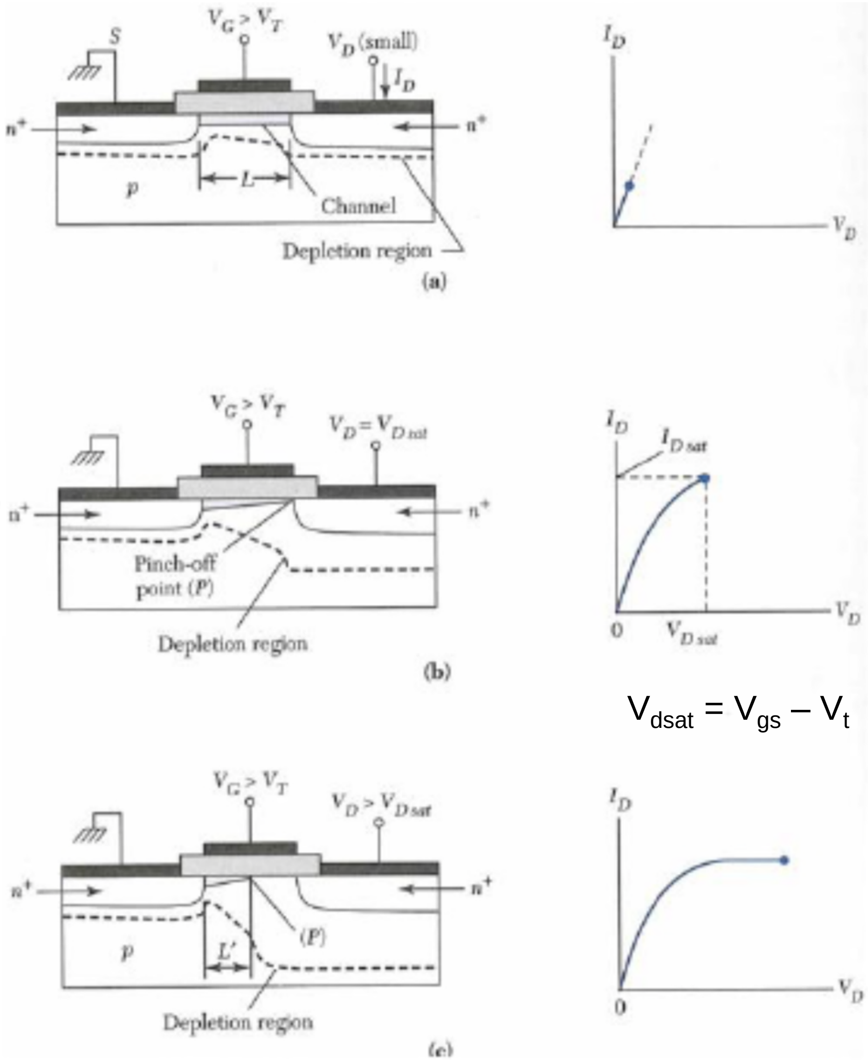
MOSFET in Equilibrium and Regions of Operation



MOSFET & equilibrium energy band.







$V_{dsat} = V_{gs} - V_t$

(a) Linear (triode), (b) "pinch-off", (c) saturation.

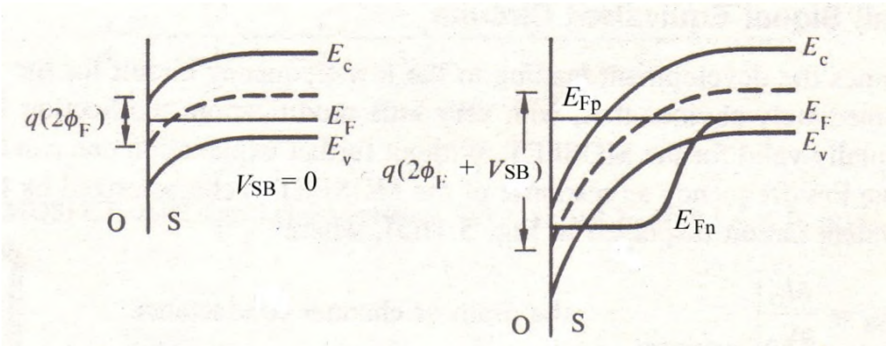
- **V_{gs}**: Vertical field controls the conductance (carriers)
- **V_{ds}**: Lateral field controls current transport (flow)

- Cut-off: $V_{gs} < V_t$ (no free carriers → no current*)
- Turn-on: $V_{gs} \geq V_t$ ($V_{ds} = 0$: $I_{ds} = 0$; $V_{ds} > 0$: $I_{ds} > 0$)
- Linear: $V_{gs} \geq V_t$ & $V_{gd} \geq V_t$ (full inversion channel)
- Saturation: $V_{gs} \geq V_t$ & $V_{gd} < V_t$ (drain "pinched-off")

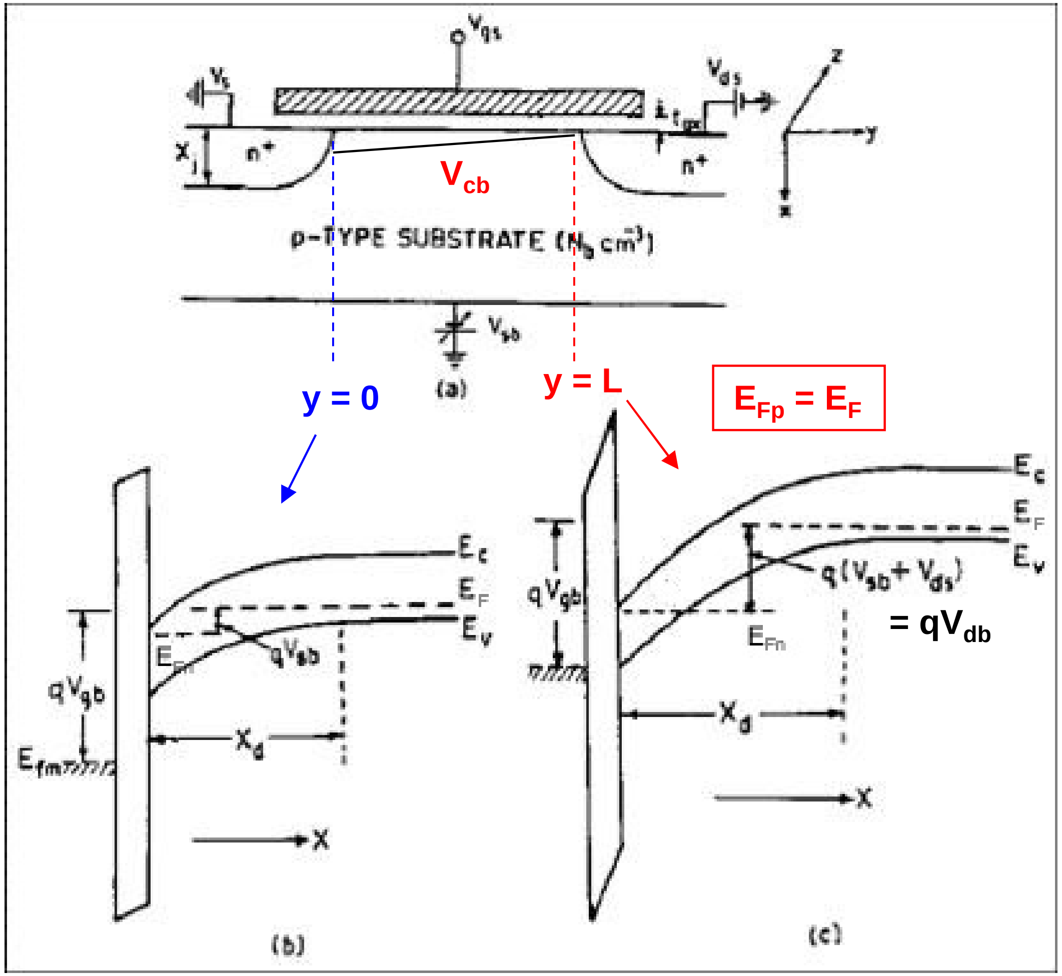
(*There is still small diffusion current, like a thick-base [long-channel] BJT.)

MOSFET Operation: Due to Inversion Carrier Imref-Split

- MOSFET at equilibrium ($V_{ds} = 0$): no current flow even if channel is created at $V_{gs} = V_t$ ($\psi_s = 2\phi_F$)
- When $V_{sb} \neq 0$ ($V_{sb} > 0$ in NMOS), electron imref will "split" from hole imref with $qV_{sb} = E_{Fn} - E_{Fp}$, so $\psi_s = 2\phi_F + V_{sb}$.

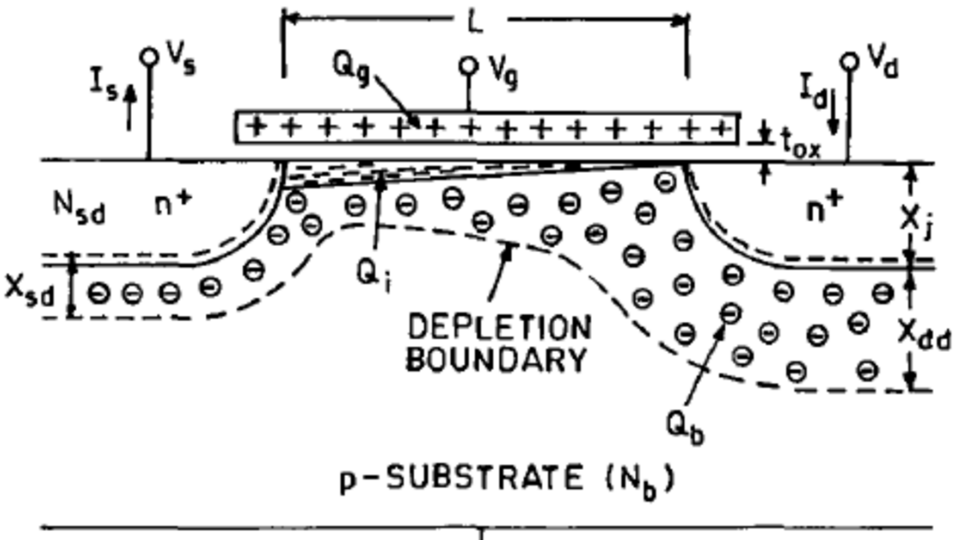


- When $V_{ds} \neq 0$ ($V_{ds} > 0$ in NMOS), holes are still at quasi-equilibrium (since no 'source' nor 'drain'), so we can assume $E_{Fp} = E_F$. However, electron imref will change from V_{sb} at source end to $V_{db} = V_{sb} + V_{ds}$ at drain end relative to E_F, and varying along the channel as $V_{cb}(y)$ ['c' stands for '<u>c</u>hannel'].
- It is the <u>gradient</u> of $V_{cb}(y)$ that drives electrons *drifting/diffusing* from source to drain along y.



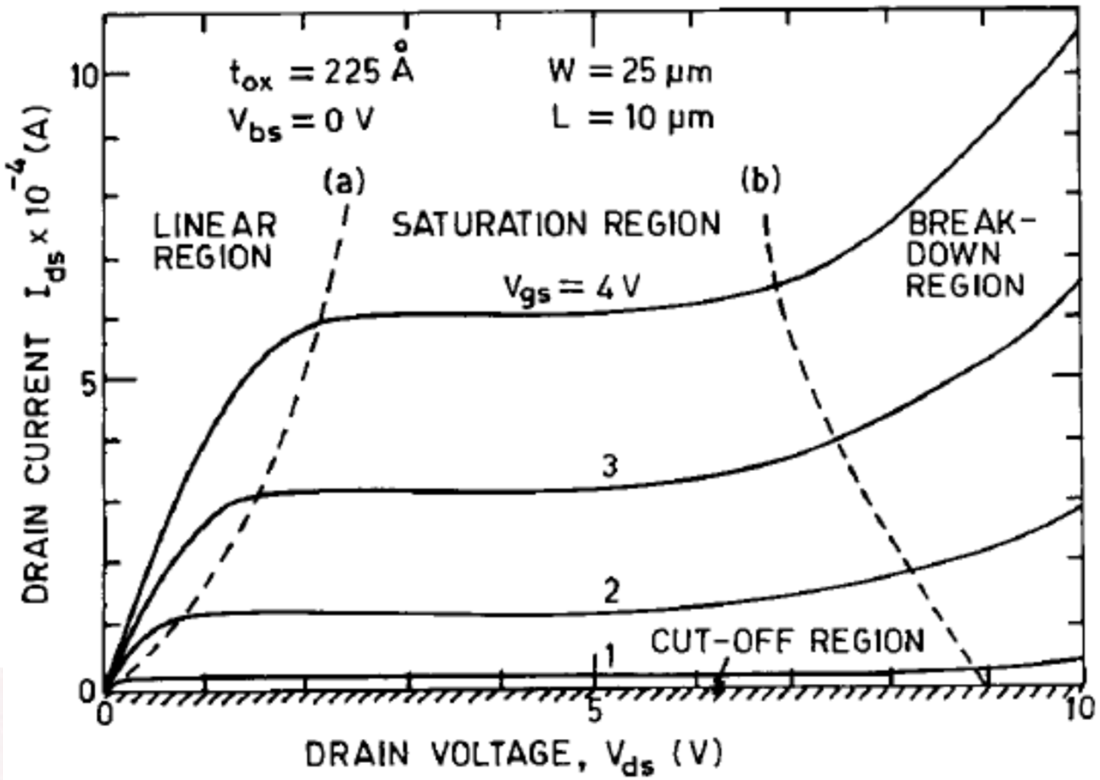
Key to understanding MOSFET operation: Band diagram in the x direction along a cutline at (b) source-end (y = 0) and (c) drain-end (y = L).

Regions of Operation: Output Characteristics



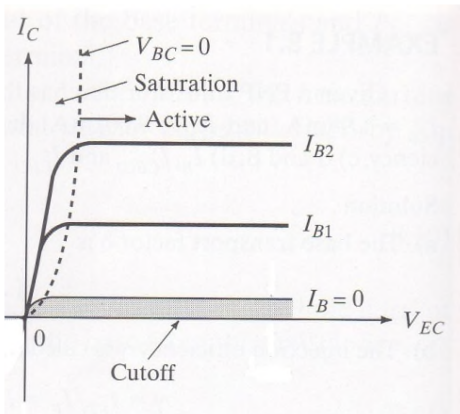
Vb

Cross-sectional view of a long-channel NMOS in strong inversion.



<u>Compare with BJT (CE):</u>

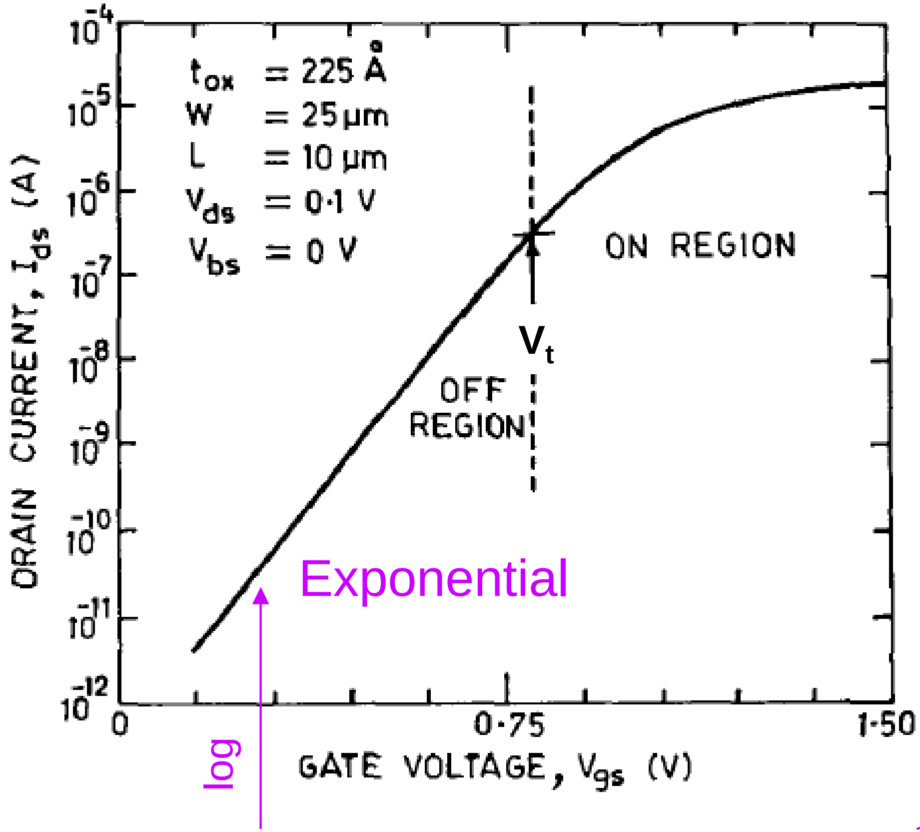
<u>BJT</u>:	<u>MOS</u>:
E–C	S–D
Saturation	Linear
Active	Saturation
$I_c \propto I_b$	$I_d \propto V_{gs} - V_t$

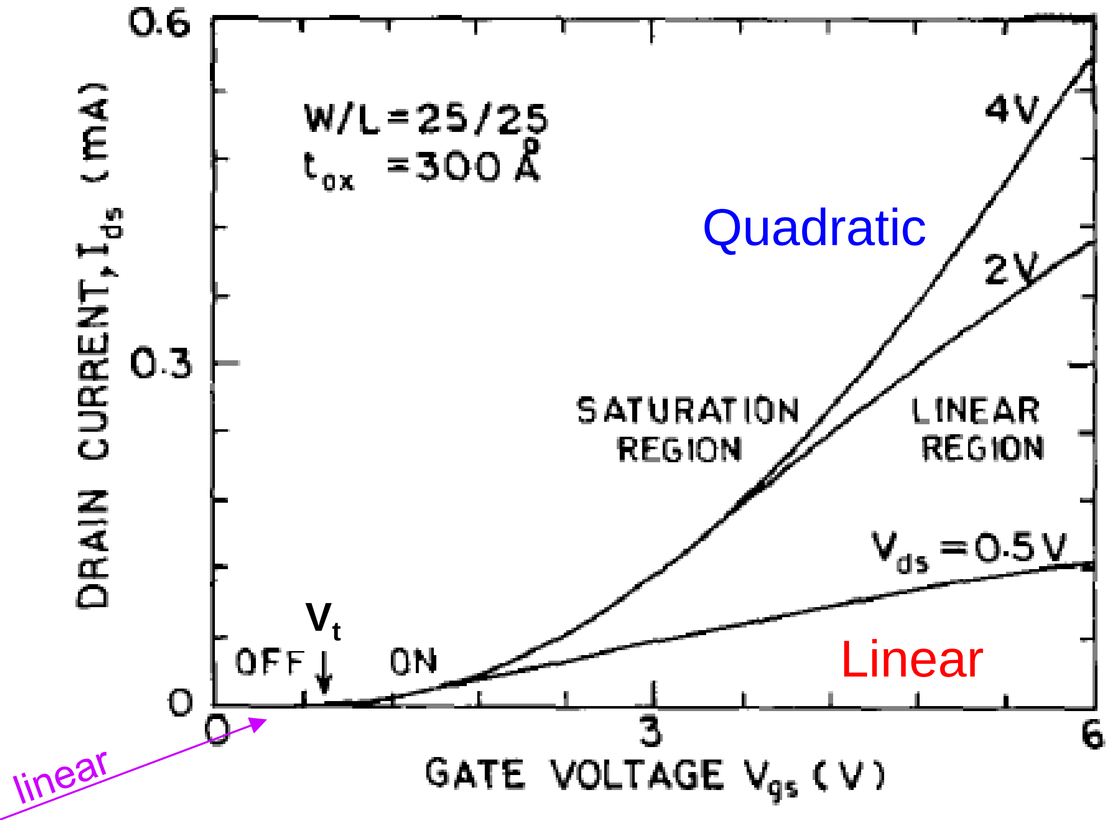


Three regions of operation (excluding break-down):
Cut-off: $V_{gs} < V_t$; Linear $V_{gd} > V_t$ (or $V_{ds} < V_{gs} - V_t$);
Saturation: $V_{gd} < V_t$ (or $V_{ds} > V_{gs} - V_t = V_{dsat}$).

(<u>Note</u>: $V_{dsat} = V_{gs} - V_t$; $V_{gd} = V_{gs} - V_{ds}$)

Regions of Operation: Transfer Characteristics





Low gate–source bias ($V_{gs} < V_t$): No inversion layer; diffusion dominant. MOS behaves like a wide-base (long-channel) BJT with $I_{ds} \propto \exp[(V_{gs} - V_t)/v_{th}]$.

High drain–source bias ($V_{ds} > V_{dsat}$): Drain side "pinched-off". MOS behaves like a current source.

Low drain–source bias ($V_{ds} < V_{dsat}$): Full channel. MOS behaves like a voltage-controlled resistor.

MOSFET Source-Referenced Threshold Voltage

MOSFET threshold voltage definition (source-referenced)

We define the **threshold voltage** (V_t) to be the gate-to-*source* voltage (V_{gs}) at which source-end surface potential is equal to twice of the bulk Fermi potential ($2\phi_F$) with reference to source–bulk voltage V_{sb}.

Potential balance — Gauss law — Charge balance — Charge-sheet approximation

$$\left(V_{gs}+V_{sb}\right)-V_{FB}=V_{gb}-V_{FB}\equiv V_{gf}=V_{ox}+\psi_s=Q_g/C_{ox}+\psi_s=-Q_{sc}/C_{ox}+\psi_s\approx -Q_b/C_{ox}+\psi_s$$

$$=-\left(-qN_AX_d\right)/C_{ox}+\psi_s=+\sqrt{2q\varepsilon_{Si}N_A\psi_s}/C_{ox}+\psi_s$$

Full-depletion approximation $X_d=\sqrt{2\varepsilon_{Si}\psi_s/qN_A}$

$$V_{gs}=V_{FB}-V_{sb}-Q_b/C_{ox}+\psi_s$$

$$V_t\equiv V_{gs}\Big|_{\psi_s=2\phi_F+V_{sb}}=V_{FB}-V_{sb}+\left[-Q_b\left(\psi_s\right)/C_{ox}+\psi_s\right]\Big|_{\psi_s=2\phi_F+V_{sb}}=V_{FB}-V_{sb}-Q_b\left(\psi_s=2\phi_F+V_{sb}\right)/C_{ox}+\left(2\phi_F+V_{sb}\right)$$

$$\therefore\quad V_t\equiv V_{gs}\Big|_{\psi_s=2\phi_F+V_{sb}}=V_{FB}+\Upsilon\sqrt{2\phi_F+V_{sb}}+2\phi_F$$

where $\Upsilon=\sqrt{2q\varepsilon_{Si}N_A}/C_{ox}$ is the **body factor**.

Body effect — Threshold-voltage shift due to non-zero V_{sb}

For NMOS, V_{sb} > 0 so that source/drain-to bulk diodes always reverse biased.

$$V_t\left(V_{sb}\right)=V_{t0}+\Upsilon\left(\sqrt{2\phi_F+V_{sb}}-\sqrt{2\phi_F}\right)$$

$$V_{t0}\equiv V_t\Big|_{V_{sb}=0}=V_{FB}+\Upsilon\sqrt{2\phi_F}+2\phi_F$$

Current–Voltage in Linear (Triode) Region

- **MOSFET analysis** — major assumptions (NMOS as example)
  - ***"GCA" – Gradual Channel Approximation***: $dE_y/dy << dE_x/dx$
  - ***"Unipolar" – hole current can be neglected*** ($E_{Fp} \approx E_F$) in normal region (excluding breakdown)
  - ***Built-in voltages for the source/drain diodes can be ignored*** *(long channel)*
  - ***No recombination/generation*** *and* ***constant mobility***
  - ***Current flows in the y direction only***
- **First-order equation derivation**
  - ***Charge-sheet approximation (CSA)***
  - ***"Pinned" surface potential*** *at strong inversion* ($2\phi_F$)
  - ***Constant bulk charge along channel***
  - ***Drift-current only in linear region***

$$\boxed{\psi_s(y) = \psi_s(0) + V_{cb}(y) = 2\phi_F + V_{sb} + V(y)} \quad (0 \le V \le V_{ds})$$

$$V_{gb} - V_{FB} = V_{ox} + \psi_s = Q_g / C_{ox} + \psi_s = -(Q_b + Q_i)/C_{ox} + \psi_s$$

$$Q_i = -C_{ox}\left(V_{gb} - V_{FB} - \psi_s\right) - Q_b \qquad \boxed{Q_b \approx -\Upsilon C_{ox}\sqrt{2\phi_F + V_{sb}}}$$

$$= -C_{ox}\left[V_{gb} - V_{FB} - 2\phi_F - V_{sb} - V(y) - \Upsilon\sqrt{2\phi_F + V_{sb}}\right]$$

$$= -C_{ox}\left[V_{gs} - V_t - V(y)\right] \qquad \boxed{V_t \equiv V_{FB} + \Upsilon\sqrt{2\phi_F + V_{sb}} + 2\phi_F}$$

$$I_{ds}(y) \approx W\int_0^\infty J_{n,drift}(y)\,dx = W\int_0^\infty qn(x,y)\,\mu_n\left(-d\psi_s/dy\right)dx$$

$$= -W\mu_n Q_i(y)\,dV/dy \qquad \left[Q_i(y) \equiv \int_0^\infty qn(x,y)\,dx\right]$$

$$I_{ds} = \frac{W}{L}\mu_n\int_0^{V_{ds}} -Q_i(y)\,dV \qquad \left(\int_0^L dy \sim \int_{\psi_s(0)}^{\psi_s(L)} d\psi_s = \int_{V_{sb}}^{V_{db}} dV_{cb} = \int_0^{V_{ds}} dV\right)$$

Linear law (I_{ds} is a *linear* function of V_{gs}) ["Sah equation"]:

$$\boxed{I_{ds} = \mu_n C_{ox}\frac{W}{L}\left(V_{gs} - V_t - \frac{1}{2}V_{ds}\right)V_{ds}} \qquad (V_{gs} > V_t, V_{gd} > V_t)$$

Current–Voltage in Saturation (Pinch-off) Region

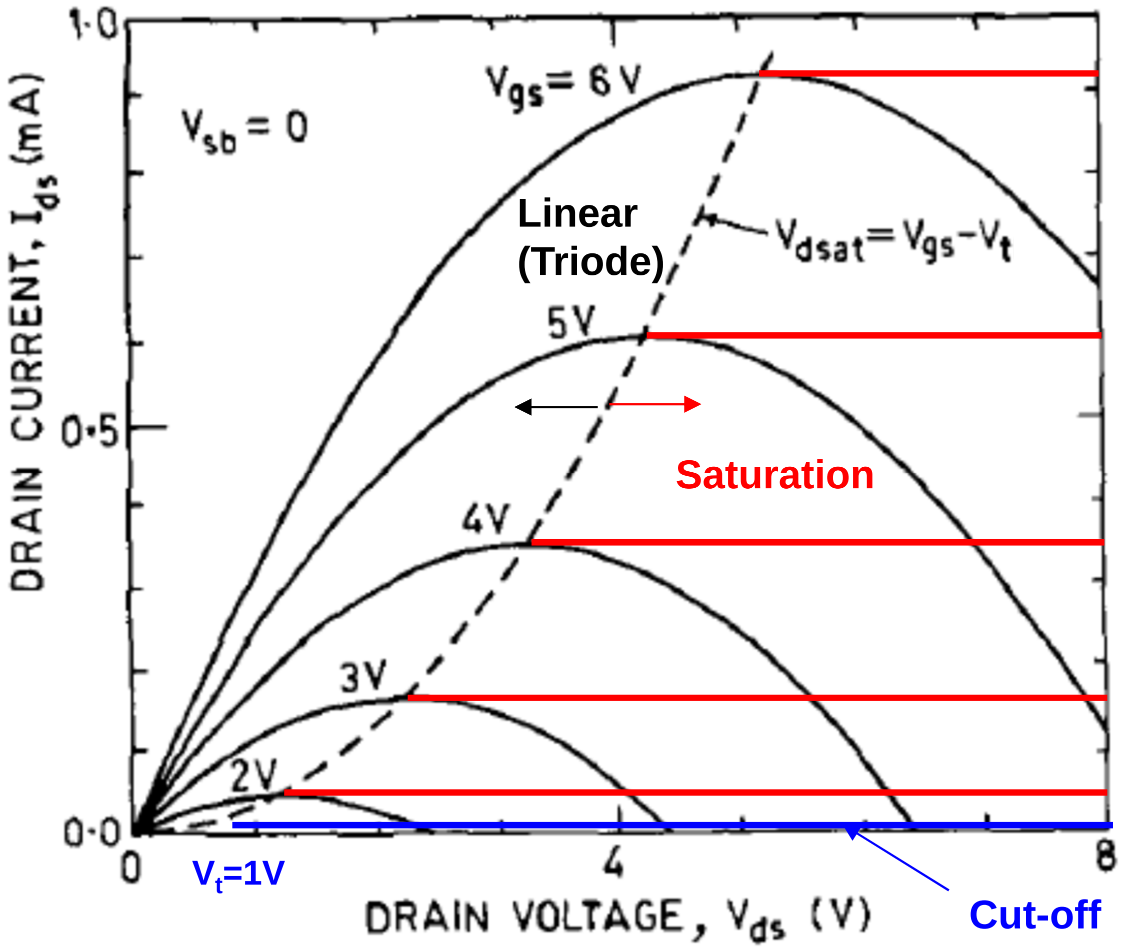
The peak of the linear current is reached when $dI_{ds}/dV_{ds} = \mu_n C_{ox} W/L\left(V_{gs} - V_t - V_{ds}\right) = 0$

For $V_{ds} \geq V_{gs} - V_t \equiv V_{dsat}$, GCA is not valid. Also, $Q_i\left(V = V_{dsat}\right) \approx 0$, channel is said to be "pinched-off." V_{dsat} is called *saturation* or *pinch-off voltage*, and the corresponding current is the *saturation current*:

Square law (I_{ds} is a *quadratic* function of V_{gs}):

$$I_{ds} = \frac{1}{2}\mu_n C_{ox} \frac{W}{L}\left(V_{gs} - V_t\right)^2 = I_{dsat} \qquad \begin{pmatrix} V_{gs} > V_t, \\ V_{gd} < V_t \end{pmatrix}$$

The "pinch-off" picture (Q_i = 0 assumption) is not physically correct since it requires the field to be infinite $E_y(y) = J_{ds}(y)/\mu_n Q_i(y)$ at pinch-off and carriers travel with infinite drift velocity. A more correct picture is that Q_i at pinch-off is very small but finite, with carriers drift under the large field in the pinch-off region at a saturated velocity.



MOSFET first-order piece-wise linear/square-law model.

Velocity Saturation and Saturation Current

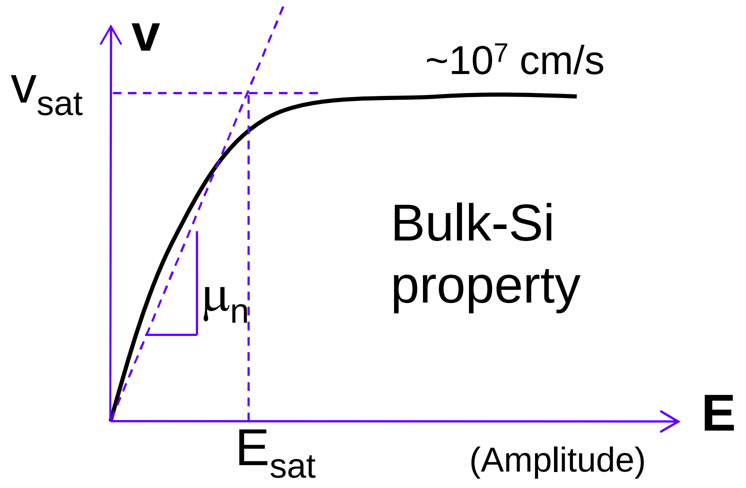
Velocity–field relation — piecewise model

$$v = \begin{cases} \dfrac{\mu_n E}{1+E/E_{sat}} & E < E_{sat} \\ v_{sat} & E \geq E_{sat} \end{cases}$$

$$I_{ds}(y) \approx -WQ_i(y)\frac{\mu_n E}{1+E/E_{sat}}$$

$$I_{ds}(y)\left(1+\frac{1}{E_{sat}}\frac{dV}{dy}\right) = -WQ_i(y)\mu_n\frac{dV}{dy}$$

$$Q_i = -C_{ox}\left(V_{gb} - V_{FB} - (2\phi_F + V_{sb} + V)\right) - Q_b$$



- **Saturation field**

$$v_{sat} = \frac{\mu_n E_{sat}}{1+E_{sat}/E_{sat}} \rightarrow E_{sat} = \frac{2v_{sat}}{\mu_n}$$

- **Lateral-field mobility**

$$Q_b \approx -\Upsilon C_{ox}\sqrt{2\phi_F + V_{sb} + V}$$

$$\mu_{eff} = \frac{\mu_n}{1+V_{ds}/(E_{sat}L_{eff})}$$

$$A_b = 1 + \frac{\Upsilon}{2\sqrt{2\phi_F + V_{sb}}}$$

$$I_{ds} = \mu_{eff}C_{ox}\frac{W}{L}\left(V_{gs} - V_t - \frac{1}{2}A_b V_{ds}\right)V_{ds} \quad (1)$$

Saturation current

$$I_{dsat} = -Wv_{sat}Q_{sat} = Wv_{sat}C_{ox}\left(V_{gs} - V_t - A_b V_{dsat}\right) \quad (2)$$

(1)(V_{dsat}) = (2): $V_{dsat} = \dfrac{E_{sat}L_{eff}(V_{gs}-V_t)}{V_{gs} - V_t + A_b E_{sat} L_{eff}}$

$$I_{dsat} = Wv_{sat}C_{ox}\frac{(V_{gs}-V_t)^2}{V_{gs} - V_t + A_b E_{sat} L_{eff}} \xrightarrow{L \to 0} \propto (V_{gs} - V_t)$$

Linear!

Charge-Sharing Model: V_t "Roll-Off"

- **Charge-sharing model**

$$L_{eff} = L_g - 2\sigma X_j \quad (\sigma \sim 0.7)$$

- ***Without charge-sharing***

$$V_t = V_{FB} - Q_{bm}/C_{ox} + 2\phi_F$$

$$C_{ox} = \varepsilon_{ox}/T_{ox} \qquad Q_{bm} = -qN_A X_{dm}$$

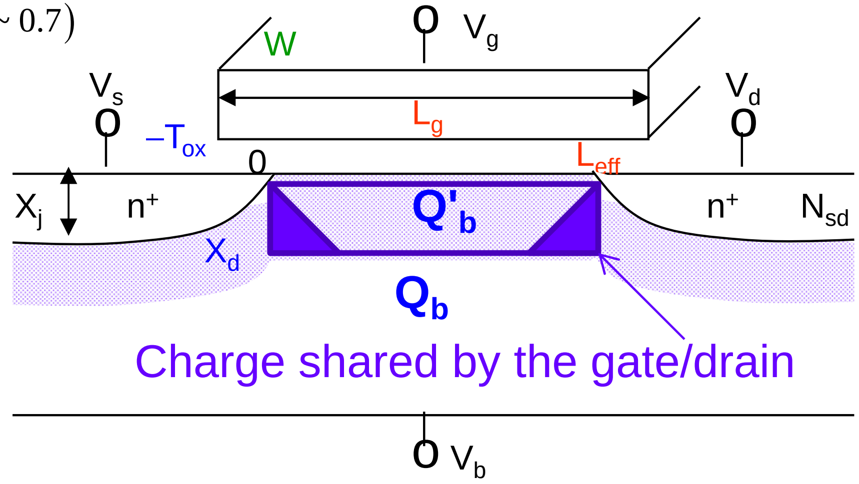
- ***With charge-charing***

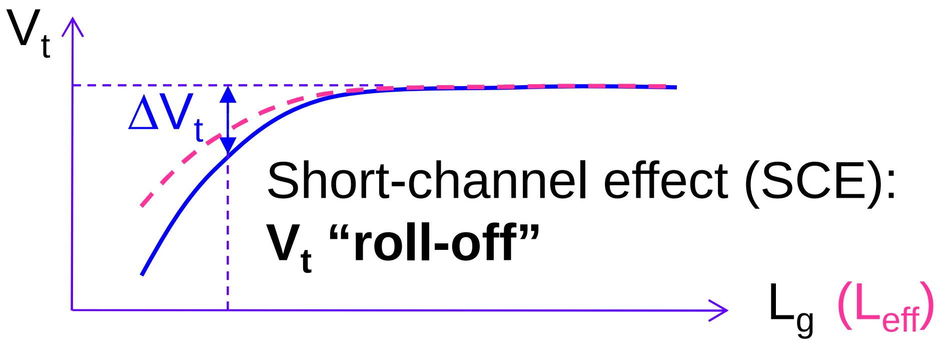
$$X_{dm} = \sqrt{2\varepsilon_{Si}\left(2\phi_F + V_{sb}\right)/qN_A}$$

$$V_t' = V_{FB} - Q_{bm}'/C_{ox} + 2\phi_F$$

$$\Delta V_t \equiv V_t - V_t' = -\frac{Q_{bm}}{C_{ox}}\left(1 - \frac{Q_{bm}'}{Q_{bm}}\right) = -\frac{Q_{bm}}{C_{ox}}\frac{X_{dm}}{L_{eff}}$$

$$= \frac{qN_A X_{dm}}{\varepsilon_{ox}/T_{ox}}\frac{X_{dm}}{L_{eff}} = \frac{4\varepsilon_{Si}\phi_F}{\varepsilon_{ox}}\frac{T_{ox}}{L_g - 2\sigma X_j}$$





Simple "Triangle" Model

Total bulk charge:

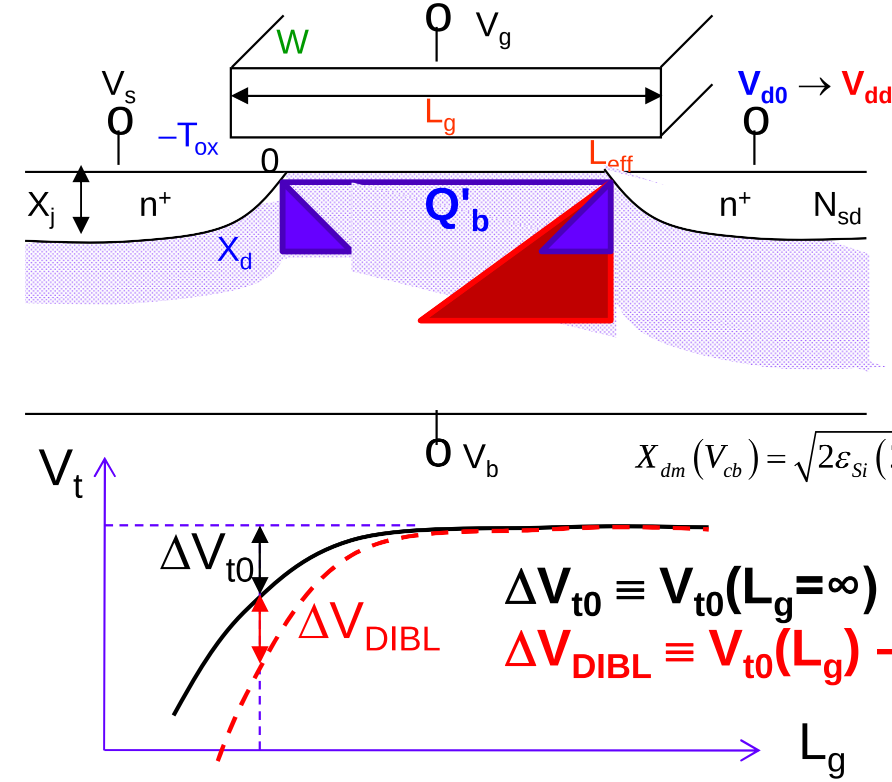
$$Q_B' = -qN_A W\left(L_{eff} X_d - X_d^2\right)$$

$$Q_B = -qN_A W\left(L_{eff} X_d\right)$$

Bulk charge per unit area:

$$\therefore \frac{Q_b'}{Q_b} = \frac{Q_B'}{Q_B} = 1 - \frac{X_d}{L_{eff}}$$

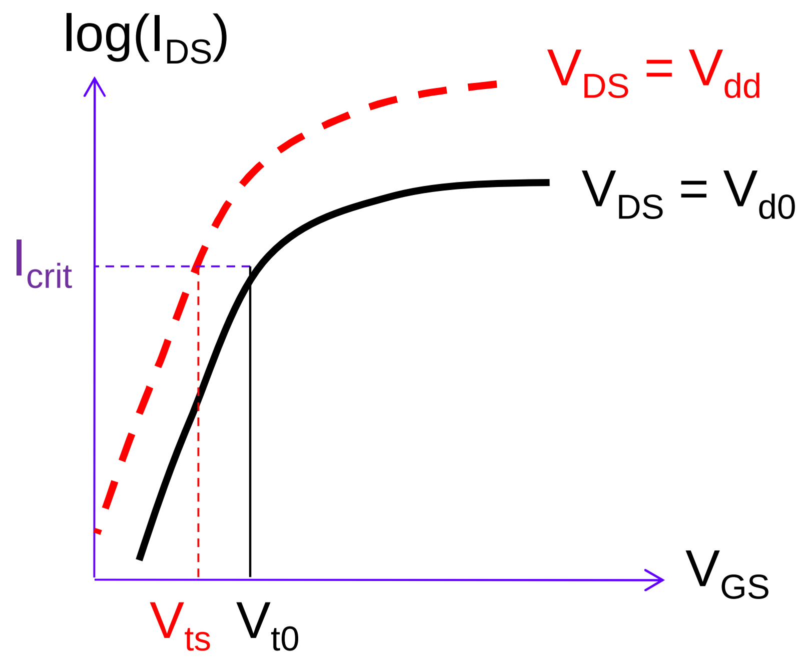
DIBL: Drain-Induced Barrier Lowering

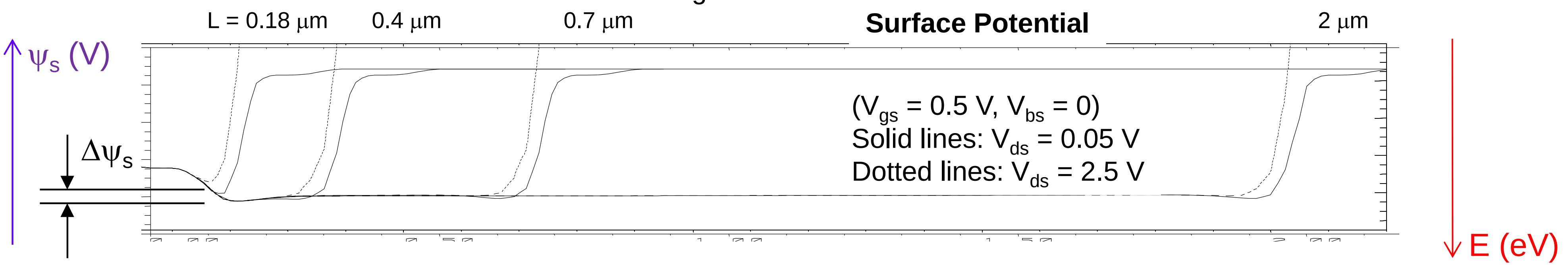


$$X_{dm}(V_{cb}) = \sqrt{2\varepsilon_{Si}\left(2\phi_F + V_{cb}\right)/qN_A}$$

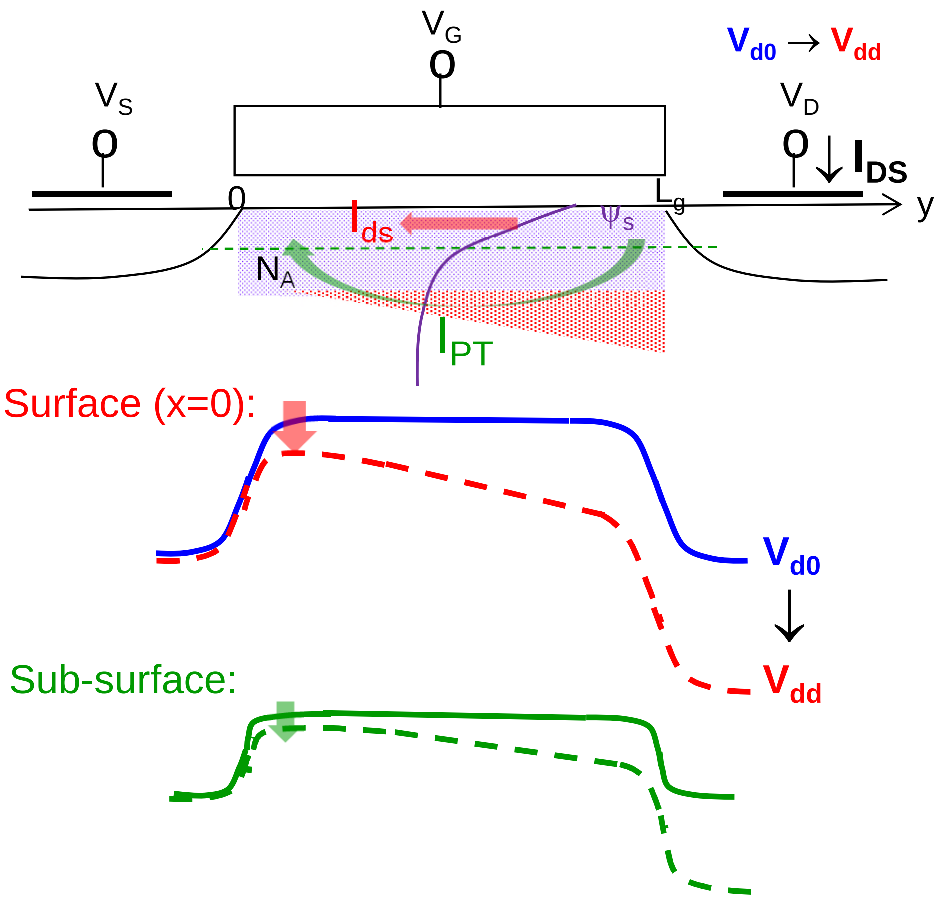
$$\Delta V_{t0} \equiv V_{t0}(L_g = \infty) - V_{t0}(L_g)$$

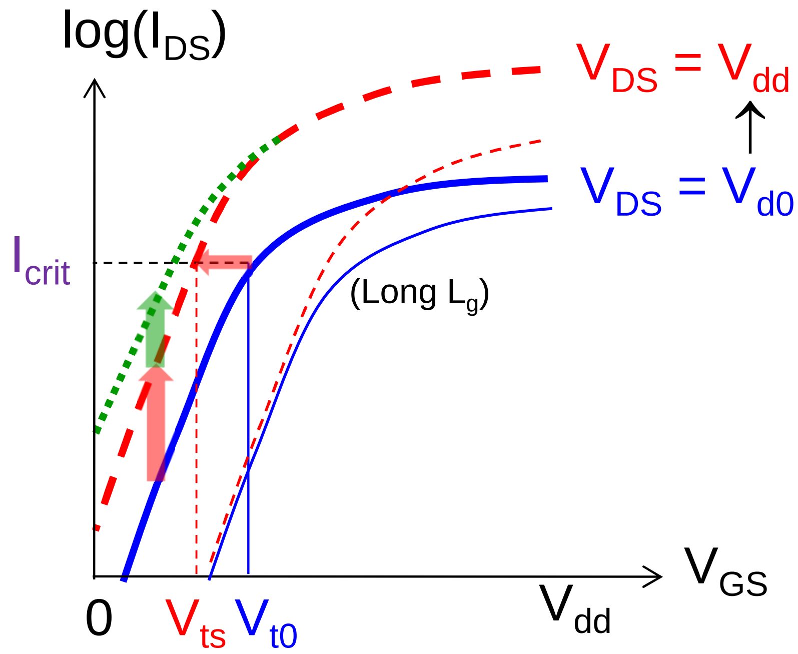
$$\Delta V_{DIBL} \equiv V_{t0}(L_g) - V_{ts}(L_g)$$

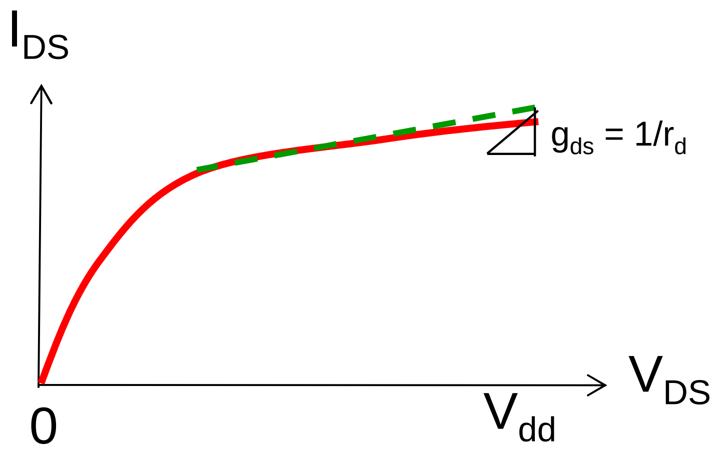




DIBL: Geometry/Doping/Bias-Dependent Effects







- ❑ **DIBL: Drain-Induced Barrier Lowering**
- ❑ **Punch-through (PT) – "sub-surface DIBL"**

Summary of Important Equations

❑ Threshold voltage

➢ ***Long-channel (1D theoretical model)***

$$V_t \equiv V_{gs}\Big|_{\psi_s=2\phi_F+V_{sb}} = V_{FB} + \Upsilon\sqrt{2\phi_F + V_{sb}} + 2\phi_F$$

$$\phi_F = \frac{kT}{q}\ln\left(\frac{N_A}{n_i}\right) \qquad \Upsilon = \sqrt{2q\varepsilon_{Si}N_A}/C_{ox} \qquad C_{ox} = \varepsilon_{ox}/T_{ox}$$

$$V_{FB} \equiv \phi_{MS} - Q_{ox}/C_{ox} = \Phi_M - \left(\chi + E_g/2 + \phi_F\right) - Q_{ox}/C_{ox}$$

➢ ***Short-channel (triangle charge-sharing model)***

$$V_{t0}\left(L_g\right) \equiv V_{t0\_long} - \Delta V_{t0} = V_{t0\_long} - \frac{4\varepsilon_{Si}\phi_F}{\varepsilon_{ox}}\frac{T_{ox}}{L_g - 2\sigma X_j}$$

➢ ***Short-channel DIBL***

$$\Delta V_{DIBL}\left(L_g\right) \equiv V_{t0}\left(L_g\right) - V_{ts}\left(L_g\right)$$

❑ Drain current

➢ ***Linear***

$$I_{ds} = \mu_{eff}C_{ox}\frac{W}{L}\left(V_{gs} - V_t - \frac{1}{2}A_bV_{ds}\right)V_{ds}$$

➢ ***Subthreshold***

$$I_{ds} = \mu_0 C_d v_{th}^2 \frac{W}{L} e^{\left(V_{gs}-V_t\right)/\left(nv_{th}\right)}\left(1 - e^{-V_{ds}/v_{th}}\right)$$

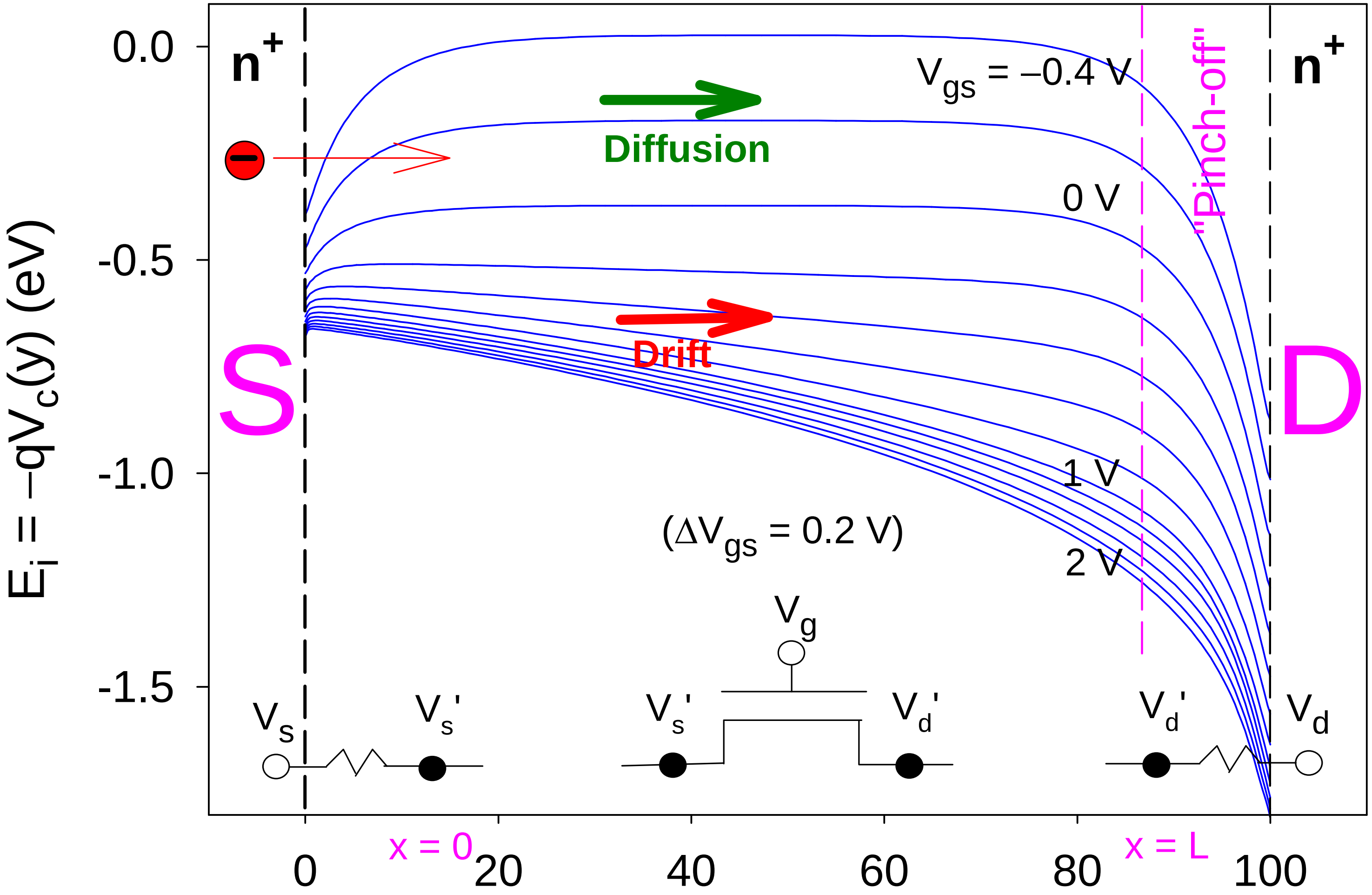
$$n = 1 + C_d/C_{ox}$$

$$C_d = \varepsilon_{Si}/X_{dm} = \frac{\Upsilon C_{ox}}{2\sqrt{2\phi_F + V_{sb}}}$$

➢ ***Saturation***

$$I_{dsat} = Wv_{sat}C_{ox}\frac{\left(V_{gs} - V_t\right)^2}{V_{gs} - V_t + A_bE_{sat}L_{eff}} \Rightarrow \propto \begin{cases}\left(V_{gs} - V_t\right)^2 & \left(L_{eff} \to \infty;\ \text{long-channel: quadratic}\right)\\ \left(V_{gs} - V_t\right) & \left(L_{eff} \to 0;\ \text{short-channel: linear}\right)\end{cases}$$

Gate-Controlled Drift ("ON") and Diffusion ("OFF")



Threshold Voltage Definition (I_{crit}@V_{t0})