MOSFET Principles

Formulas from Chapter 2

- Oxide capacitance: $C_{ox} = \frac{\epsilon_{ox}}{t_{ox}}$
- Work function of semiconductor: $q\phi_s = q\chi_s + \frac{E_g}{2} + q\phi_F$
- Fermi potential (definition): $q\phi_F = E_i - E_F$
- Fermi potential (p-type): $\phi_F = +\frac{kT}{q} ln\left(\frac{N_A}{n_i}\right)$
- Fermi potential (n-type): $\phi_F = -\frac{kT}{q} ln\left(\frac{N_D}{n_i}\right)$
- Work-function difference: $q\phi_{ms} = q\phi_m - q\left(\chi_s + \frac{E_g}{2q} + \phi_F\right)$
- Flat-band voltage: $V_{FB} = \phi_{ms} - \frac{qN_{oc}}{C_{ox}}$ where qN_{oc} is the density of charge.
- For strong inversion: $\varphi_S = 2\phi_F$

- Inversion-layer charge: $Q_I = (V_{GS} - V_T)C_{ox}$
- Voltage drop across gate oxide: $V_{ox}C_{ox} = Q_d + Q_I$
- Voltage drop across gate oxide: $V_{ox} = (V_{GS} - V_{FB}) - \varphi_s$
- Threshold voltage: $V_T = V_{FB} + 2 \mid \phi_F \mid + \frac{Q_d}{C_{ox}}$
- Depletion-region charge density: $Q_d = qN_A w_d = C_{ox}\gamma\sqrt{\varphi_s}$
- Depletion-layer width: $w_d = \sqrt{\frac{2\epsilon_s \varphi_s}{qN_A}}$
- Threshold voltage: $V_T = V_{FB} + 2 \mid \phi_F \mid + \gamma\sqrt{2 \mid \phi_F \mid}$
- Body factor: $\frac{\sqrt{2\epsilon_s qN_D}}{C_{ox}}$ where N_D is the dopant concentration.

MOSFET Structure

See slide 2.

Threshold Voltage and Strong Inversion

- Gate voltage determines the potential at the silicon surface φ_S.

- Due to the work-function difference between the metal and semiconductor, a voltage V_{FB} must be applied between gate and channel to flatten the bands. V_{FB} is called the *flat-band voltage*.
- See slide 3 and overview MOSFET operation.
- See slide 4.
  - This is the flat band condition.
  - Observe that the Fermi level is constant in the y direction.
  - This creates an energy barrier that prevents electrons from flowing between source and drain.
  - As long as there is an energy barrier between the source and the drain there is no flow of electron.
  - MOSFET is in *cutoff*.
- To turn the MOSFET *ON*, a positive voltage that reduces the

potential barrier between source and drain should be applied to the gate.

- The energy barrier at the surface is reduced, as seen in slide 5.
- Bands bent at the silicon surface, moving the conduction band toward the Fermi level.
- Once the Fermi level is closer to the conduction band than to the valence band, the occupancy of the conduction band states becomes more likely that the occupancy of the valence band states.
- The electron concentration becomes larger than the hole concentration.
- The inversion layer as been created.
- When $\varphi_S = 2\phi_F$[a], strong inversion is reached.

[a] ϕ_F is the Fermi level and the middle of the energy gap in the bulk

- Further increases in the gate voltage cause a quick increase in the electron concentration in the inversion layer. The additional gate voltage appears across the oxide and φ_S remains constant.

- See slide 6.

- The gate voltage required for $\varphi_S = 2\phi_F$ is the *threshold voltage*, V_T.

- When a positive effective voltage $V_{GS} - V_{FB}$ is applied to the gate, holes are rejected from the region under the gate. The charge due to the uncompensated acceptor atom ions left behind is called the *depletion region charge*. The associated charge density, Q_d, can be expressed in C/m^2.

- The following approximation is used:
  - the inversion-layer charge is neglected if the gate voltage is below V_T; under this conditions, the only uncompensated

charge in the silicon substrate is Q_d.

– the surface potential is pinned at

$$\varphi_S \approx 2\phi_F$$

for $V_{GS} > V_T$. Increases beyond V_T produce the inversion layer charge Q_I,

$$Q_I = (V_{GS} - V_T)\, C_{ox}$$

- Situation is similar to two caps in series. One fraction of the effective applied voltage appears across the oxide capacitor, the other across the depletion region.
- The voltage drop across the oxide is given as the difference between the effective gate potential on one side of the oxide and the surface potential. Thus

$$V_{GS} - V_{FB} - \varphi_S = \frac{Q_d + Q_I}{C_{ox}}$$

- At the onset of strong inversion, the surface potential is $\varphi_S = 2\phi_F$, $V_{GS} = V_T$, and $Q_I = 0$; thus

$$V_T - V_{FB} - 2\phi_F \mid= \frac{Q_d}{C_{ox}}$$

and

$$V_T = V_{FB} + 2\phi_F + \frac{Q_d}{C_{ox}}$$

- In chapter 2, we found that

$$Q_d = qN_A w_m = \sqrt{2\epsilon_s q N_A (2\phi_F)}$$

- In terms of the *body factor*, $\gamma = \frac{\sqrt{2\epsilon_s q N_A}}{C_{ox}}$, the threshold voltage can be expressed as

$$V_T = V_{FB} + 2\phi_F + \gamma\sqrt{2\phi_F}$$

This expression is valid if the substrate and the source are at the same potential.

- In slide 5,
  - The situation for $V_{DS} = 0$ is shown.
  - There is no drain current and the quiescent point remains at the origin.
- If a positive voltage is applied to the drain, the conduction band at the drain is lowered with respect to the source and electrons flow from source to drain. This is shown in slide 7.
- The drain current magnitude is determined by the amount of voltage applied to the drain and by the density of charge in the inversion layer.
- Since

$$Q_I = (V_{GS} - V_T)\,C_{ox}$$

the drain current can be expressed

$$I_D = \beta(V_{GS} - V_T)V_{DS}$$

where β is a proportionality factor.

Body Effect

- If a negative source-to-substrate voltage is applied, the energy of the bands in the substrate is increased.
- The barrier between source/drain and substrate is increased by qV_{SB}. See slide 8 and compare (a) with slide 5.
- As a consequence, the potential needs to be increased to $\varphi_S = 2\phi_F + V_{SB}$. In the other hand, the gate-to-bulk voltage is increased to $V_{GS} + V_{SB} - V_{FB}$ and the effect of V_{SB} cancels.
- The bulk voltage affects Q_d through its dependence on the surface potential.
  - when $V_{SB} = 0$ under strong inversion

$$Q_d/C_{ox} = \gamma\sqrt{\varphi_S} = \gamma\sqrt{2\phi_F}$$

– When V_{SB} is applied, $\varphi_S = 2\phi_F + V_{SB}$ and

$$Q_d/C_{ox} = \gamma\sqrt{2\phi_F + V_{SB}}$$

– The threshold voltage expression is also modified

$$V_T = V_{FB} + 2\phi_F + \gamma\sqrt{2\phi_F + V_{SB}}$$

– The change in threshold voltage is

$$\begin{aligned} \Delta V_T &= V_T(V_{SB}) - V_T(0) \\ &= \gamma(\sqrt{2\phi_F + V_{SB}} - \sqrt{2\phi_F}) \end{aligned}$$

Saturation

- Previous discussion assumes inversion region exists along the bulk at both source and drain sides.
- If a positive voltage is applied to the drain, electrons flow from source to drain.

- The positive drain voltage reduces the difference between the gate and drain voltages.
- There is a point at which $V_{GD} < V_T$ and the inversion region disappears at the drain side.
- This is called *channel pinch-off*.
- See slide 9
- A depletion region is formed at the drain side.
  - This depletion region, however, has little effect on the drain current
  - The current is limited by the electron concentration at the pinch-off point.
  - The electron concentration is controlled by the gate-to-source voltage.
  - The drain current is controlled by the gate voltage.

- The MOSFET works like a voltage-controlled current source.
- This operating regime is called *saturation region*.
- Very small devices are *short-channel* MOSFETs.
  - In these devices the drain current saturates at smaller drain-to-source voltages.
  - The small size of the device makes the electric field in the channel very large.
  - Carrier mobility is reduced due to the large electric field.
  - This can happen before channel pinch-off.

Types of MOSFET: see slide 10.

Note: For p-channel MOSFETs use:

$$V_T = V_{FB} - 2 \mid \phi_F \mid - \gamma \sqrt{2 \mid \phi_F \mid + \mid V_{SB} \mid}$$

Example 5.1

Data: $t_{ox} = 10$ nm; $N_A = 5 \times 10^{16} cm^{-3}$; $N_{OC} = 5 \times 10^{10} cm^{-2}$.

Constants: $V_T = 0.026V$; $\epsilon_{ox} = 3.9 \times 8.85 \times 10^{-12} F/m$;
$\epsilon_s = 11.8 \times 8.85 \times 10^{-12} F/m$; $n_i = 1.02 \times 10^{10} cm^{-3}$; $E_g = 1.12eV$

Problem: Find V_T if bulk is biased at 0 and $-5V$, respectively.

MOSFET Fabrication

Ion implantation: see slides 11 and 12.

NMOS

This is only a summary. Check book for full explanation.

- Cross section & top view of inverter are shown in slide 13

  Thick oxide is used to isolate devices. Notice that metal connections with next device forms a MOSFET gate; thick oxide increases V_T thus preventing this parasitic MOSFET from operating. See formulas from chapter 2 to verify this.

  Also, observe that n+ of EMOSFET drain and of DMOSFET source are merged into a single n+ region. Layer merging is employed whenever possible to reduce parasitic effects and circuit size.

- Fabrication steps are shown in Slide 14 (should be animated).

1. Create isolation p+ and oxide regions around active area.
   (a) Grow SiO_2 buffer.
   (b) CVD of Si_3N_4 layer to protect active area.
   (c) photo 1 to open windows for isolation
   (d) etch Si_3N_4 over isolation regions
   (e) ion implant p+
   (f) thermal oxidation to create field oxide
2. etch Si_3N_4 and buffer oxide.
3. photoresist and photo 2 to define DMOSFET channel region
4. implant channel
5. remove photoresist
6. grow gate oxide
7. Photo 3: to etch source-drain and gate contact hole
8. CVD poly to form drain of EMOSFET and source/gate of DMOSFET

9. Photo 4: to pattern poly and etch gate oxide
10. CVD of phosphorous doped oxide
11. phosphorous diffusion into source/drain regions; poly prevents diffusion into gates (self-alignment).
12. Photo 5: contact hole etching for drain/source region
13. Al deposition and Photo 6 for metalization.

CMOS

- Basic idea: slide 15.
- Layout of inverter: slide 16.

  Substrate of DMOSFET is an n-well deposited on p-type substrate.

  Fabrication steps are shown in Slide 17 (should be animated).

MOSFET Modeling

LEVEL 2 Spice model

- Differential form of Ohm's Law:

$$j = q\mu_0 n E$$

- Take averages of quantities in the above equation.
  - channel cross section: $x_{ch}W$
  - Average electric field: V_{DS}/L_{eff}
- Ohm's law using average field:

$$\begin{aligned} I_D(A) &= j x_{ch} W \\ &= \mu_0 W q n x_{ch} V_{DS}/L_{eff} \\ &= \frac{\mu_0 W}{L_{eff}} \bar{Q}_I V_{DS} \end{aligned}$$

where $\bar{Q}_I = qnx_{ch}$ is the average value of the inversion-layer charge density.

- Using the model previously introduced $\bar{Q}_I = 0$ for $V_{GS} < V_T$ and $\bar{Q}_I = (V_{GS} - V_T)C_{ox}$ for $V_{GS} \geq V_T$. Thus, in the triode region

$$I_D = \beta(V_{GS} - V_T)V_{DS}$$

$$\beta = \frac{\mu_0 W C_{ox}}{L_{eff}}$$

- The above expression neglects non-uniformities on Q_I created by V_{DS}. See slide 18.
  - The reduction in Q_I by the drain end of the channel can be modeled through the increase in the threshold voltage caused by the body effect.

$$v_T(\varphi_S) = V_{FB} - V_{SB} + \varphi_S + \gamma\sqrt{\varphi_S}$$

– At the source end of the channel,

$$\varphi_S = 2\phi_F + V_{SB} = \varphi_{src}$$

– At the drain end,

$$\varphi_S = 2\phi_F + V_{SB} + V_{DS} = \varphi_{drn}$$

– The average Q_I can be found from

$$\begin{aligned}\bar{Q_I} &= \frac{\int_{\varphi_{src}}^{\varphi_{drn}} Q_I(\varphi_S) d\varphi_S}{(2\phi_F + V_{SB} + V_{DS}) - (2\phi_F + V_{SB})} \\ &= \frac{\int_{\varphi_{src}}^{\varphi_{drn}} (V_{GS} - v_T)\, C_{ox} d\varphi_S}{V_{DS}} \\ &= \frac{\int_{\varphi_{src}}^{\varphi_{drn}} \left(V_{GS} - V_{FB} + V_{SB} - \varphi_S - \gamma\sqrt{\varphi_S}\right) C_{ox} d\varphi_S}{V_{DS}}\end{aligned}$$

$$
\begin{aligned}
&= \frac{C_{ox}}{V_{DS}}\Bigg((V_{GS} - V_{FB} + V_{SB})\varphi_S - \frac{1}{2}\varphi_S^2 \\
&\quad -\frac{2}{3}\gamma\varphi_S^{\frac{3}{2}}\Bigg) \\
&= \frac{C_{ox}}{V_{DS}}\{\left(V_{GS} - V_{FB} - 2\phi_F - \frac{V_{DS}}{2}\right)V_{DS} \\
&\quad -\frac{2}{3}\gamma\left(\sqrt{(2\phi_F + V_{SB} + V_{DS})^3} - \sqrt{(2\phi_F + V_{SB})^3}\right)\}
\end{aligned}
$$

– SPICE LEVEL 2 model for the drain current in the triode region becomes

$$
\begin{aligned}
I_D &= \beta\{\left(V_{GS} - V_{FB} - 2\phi_F - \frac{V_{DS}}{2}\right)V_{DS} \\
&\quad -\frac{2}{3}\gamma\left(\sqrt{(2\phi_F + V_{SB} + V_{DS})^3} - \sqrt{(2\phi_F + V_{SB})^3}\}\right)\}
\end{aligned}
$$

where

$$\beta = \frac{\mu_0 W C_{ox}}{L_{eff}} = KP\frac{W}{L_{eff}}$$

– See equation plot in slide 19.

– The voltage at which I_D saturates can be found by setting $\frac{\partial I_D}{\partial V_{DS}} = 0$ in the above equation and solving for V_{DS}. The result is

$$\begin{aligned} V_{DS_{sat}} &= V_{GS} - V_{FB} - 2\phi_F \\ &\quad -\frac{\gamma^2}{2}\left(\sqrt{1 + \frac{4}{\gamma^2}(V_{GS} - V_{FB} + V_{SB})} - 1\right) \end{aligned}$$

– This model is difficult to use because V_T is not a parameter.

– Calculating the $3/2$ power is computationally intensive.

LEVEL 3 Spice Model

– Obtained by approximating the $3/2$-power term in the LEVEL 2 model by the first three terms in its Taylor series.

– The following equation is obtained for the triode region $(0 \leq V_{DS} < V_{DS_{sat}})$:

$$I_D = \beta \left(V_{GS} - V_T - \frac{1 + F_B}{2} V_{DS} \right) V_{DS}$$

where

$$F_B = \frac{\gamma}{2\sqrt{2\phi_F + V_{SB}}}$$

– For the saturation region $(V_{DS} \geq V_{DS_{sat}})$

$$I_D = \frac{\beta}{2(1 + F_B)} (V_{GS} - V_T)^2$$

– The drain-to-source voltage at the onset of saturation is

$$V_{DS_{sat}} = \frac{V_{GS} - V_T}{1 + F_B}$$

LEVEL 1 Spice model

– Obtained by neglecting F_B in LEVEL 2 model.

– Triode region

$$I_D = \beta(V_{GS} - V_T)V_{DS} - \frac{1}{2}V_{DS}^2$$

– Saturation region:

$$I_D = \frac{\beta}{2}(V_{GS} - V_T)^2$$

– The drain-to-source voltage at the onset of saturation is

$$V_{DS_{sat}} = V_{GS} - V_T$$

Second Order Effects Incorporated into LEVEL 3 Spice Model

Mobility Reduction with Gate Voltage

- Mobility is reduced by increased scattering.
- Scattering depends on channel thickness, and thus on gate voltage.

- See slide 21.
- The following semi-empirical equation is used in Spice:

$$\mu_s = \frac{\mu_0}{1 + \theta(V_{GS} - V_T)}$$

- The surface mobility μ_s is used instead of the low-field mobility μ_0.
- The parameter θ is called the *mobility modulation constant*.

Drift Velocity Saturation

- High drain-to-source electric fields produce a reduction in the mobility of channel carriers.
- The *effective mobility* can be expressed in terms of V_{DS} and L_{eff}:

$$\mu_{eff} = \frac{\mu_s}{1 + \frac{\mu_s}{v_{max}} \frac{V_{DS}}{L_{eff}}}$$

- The strength of the effect is controlled by the *maximum drift velocity*, v_{max}. Setting $v_{max} = \infty$ eliminates the effect.
- Velocity saturation also affects $V_{DS_{sat}}$.

Finite Output Resistance

- Due to two effects: *channel-length modulation* and *drain-induced barrier-lowering* (DIBL).
- First is dominant for large devices. It is modeled by varying the channel length, setting it to $L_{eff} - L_{pitch}$, where L_{pitch} depends on V_{DS}.
- See table 5.4.
- Channel-length modulation model is controlled by Spice's parameter κ.
- DIBL is more important in short-channel devices. It is modeled

by modifying the expression for V_T,

$$V_T = V_{FB} + 2\phi_F + \gamma\sqrt{2\phi_F + V_{SB}} - \sigma_D V_{DS}$$

- The coefficient σ_D is calculated by Spice from an equation that involves L_{eff}, C_{ox} and Spice parameter η, known as the *static feedback coefficient*. See table 5.5.
- See slide 22.

Short-channel effects on V_T

- See slide 23.
- The charge in the trapezoidal section dominates in short-channel MOSFETs.
- *Charge-sharing factor* (F_s): ratio between trapezoidal and rectangular areas.

- Threshold voltage is modified as in:

$$V_T = V_{FB} + 2\phi_F + F_s\gamma\sqrt{2\phi_F + V_{SB}} - \sigma_D V_{DS}$$

- Important in devices with channels shorter than $2\mu m$.

Narrow-channel effects on V_T

- See slide 24.
- Similar to short channel.
- The expression for the threshold voltage is modified to:

$$V_T = V_{FB} + 2\phi_F + F_s\gamma\sqrt{2\phi_F + V_{SB}} - \sigma_D V_{DS} + F_n(2\phi_F + V_{SB})$$

- F_n is calculated from the channel width W and a parameter δ that modulates the strength of this effect.

Sub-threshold Current

- There is a small channel current for $V_{GS} < V_T$.

- The inversion layer charge is very small for $V_{GS} < V_T$, and even moderate V_{DS} causes concentration gradients.
- Current is due to diffusion and is modeled with the equation:

$$I_{D_{subth}} = I_{D0} e^{V_{DS}/n_s kT}$$

- I_{D0} is selected to obtain a smooth transition from sub-threshold to above threshold drain current.
- See table 5.5 for a formula for n_s.

Parameter Measurement

V_{T0} and KP

- Small V_{DS}: linear part of I_D versus V_{DS}.
- Small $V_{GS} - V_T$ so that $\mu_{eff} \approx \mu_0$.
- Conditions can also be expressed as

$$V_{DS}(1 + F_B) \ll V_{GS} - V_T \ll 1/\theta$$

- Equation is:

$$I_D \approx \beta_0(V_{GS} - V_T)V_{DS}$$

- Slope gives β_0 and x-intercept yields V_T.
- See slide 26.
- $\beta_0 \equiv \mu_0 C_{ox} \frac{W}{L_{eff}}$.
- From the measured β_0 you can find

$$KP = \frac{\beta_0 W}{L_{eff}}$$

θ

- θ is the *mobility modulation constant*.
- I_D versus V_{GS} stops being linear due to the mobility reduction

effect.

- Do the measurement at higher V_{GS} but still small V_{DS}:

$$V_{DS}(1 + F_B) \ll V_{GS} - V_T$$

- Measure I_D versus V_{GS} See Slide 28.
- Use

$$I_D \approx \beta V_{DS}(V_{GS} - V_T)$$

to find β for different values of V_{GS}.

- From

$$\beta = \frac{\beta_0}{1 + \theta(V_{GS} - V_T)}$$

- Plot $\frac{\beta_0}{\beta} - 1$ versus $V_{GS} - V_T$ and find θ from the slope.
- Requires previous estimation of V_T and β_0.