MOS transistor modeling and design

CMOS IC design flow

1. Understand the application
2. Requirement specification and simulation test bench
3. Step-by-step system and circuit design
4. Design verification: corner simulations with test bench
5. Design adjustments
6. Lay-out
7. Extraction parasitics
8. Design verification: corner simulations with test bench
9. Minor design adjustments

increasing costs of recalling design decisions

The big fear
A seemingly unimportant detail overseen at an early stage of the design, turns out to be a show-stopper later.

top-down design ↓ ↑ bottom-up show stopper

The devil is in the details

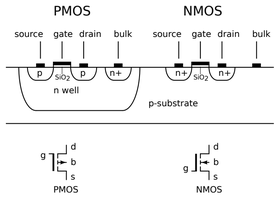
Structured analog design

1. Use relative simple models that model the dominant effects at early stages of the design and increase model complexity at later stages
2. Be aware of simplifications and include error budgets for them
3. Use a design approach that maximally exploits the intended dominant effects

Required knowledge

1. Physical operation of semiconductor devices
   - Basic operating principle and operating regions
   - Secondary effects
2. Simulation models
   - Models used in computer simulations
   - Models for setting-up design equations:
     Relation between:
     * Performance aspects of interest
     * Technology parameters
     * Device geometry
     * Operating conditions
   - How to derive simple models from the complex ones
     1. Use simplified model equations
     2. Extract model parameters for simple models from single-device simulations with complex models
     3. Use graphs and apply scaling

CMOS technology



Multiple-finger and/or Multiple-device Lay-out

- << drain/source capacitance
- << gate series resistance
- >> device matching

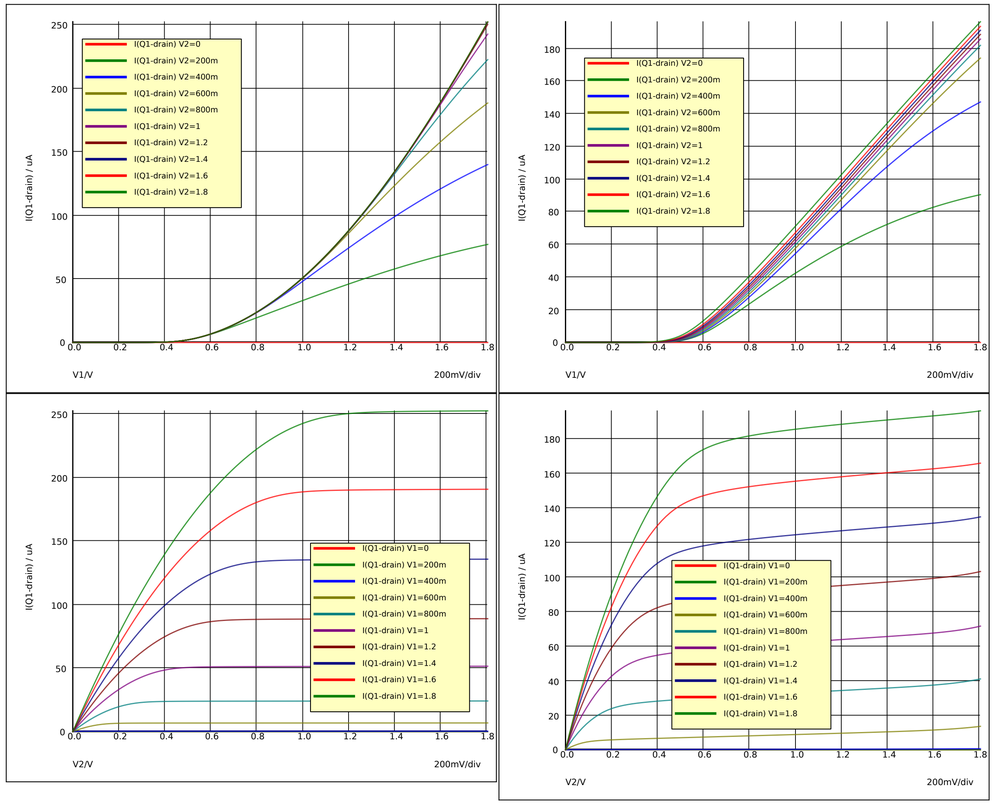
s	s
g	g
d	d
g	g
s	s
g	g
d	d
g	g
s	s

NMOS basic operation

1. V_B=0, V_S=0, V_G=0, V_D>0
   - Surface potential at interface oxide and substrate
   - No current flow from drain to source
2. V_G increases
   - Surface potential increases due to capacitive coupling with the gate
   - Source injects electrons into p-region
   - Diffusion current from drain to source (Lateral NPN)
   - Weak inversion operation, exponential relationship $I_{DS}(V_{GS})$
3. V_G increases further
   - Channel extends and reduces capacitive coupling
   - Quadratic relationship between I_{DS} and V_{GS}: strong inversion
   - V_T is voltage at transistion from weak to strong inversion
   - If V_G-V_D<V_T the device operates in 'pinch-off' or saturation
4. V_G increases further V_G-V_D>=V_T
   - The channel extends to the drain: linear region (voltage-controlled resistor)
5. Increase of VDS results in a wider drain depletion layer (CLM). This effectively results in a shorter channel and an increase of I_{DS}.

Small-geometry effects

1. Thin oxide: high vertical field strength reduces carrier mobility: VFMR
2. Sort channel: high lateral field strength causes velocity saturation
3. Short channel: drain depletion layer extends under the channel. This increases the coupling of the surface potential with the gate voltage and lowers V_T. Result: increase of I_{DS} (DIBL) with V_{DS}.



CMOS18 NMOS, L=18um W=22um
No short-channel effects

CMOS18 NMOS, L=180nm W=220nm
Short channel effects: (velocity saturation and DIBL)

MOS level-1 model

1. Often used for hand calculations
2. Weak inversion not modeled
3. Abrubts transition from linear to saturation region
4. Small-geometry effects not modeled

MOS BSIM models BSIM3 and higer

1. Optimized for numeric accuracy, speed and stability
2. No focus on physical basis
3. Often used for modern IC processes
4. Small-signal uses Ward-Dutton capacitance matrix
5. Different expressions for different operating regions

MOS EKV model

1. Physical basis
2. One expression for all operating regions
3. Takes bulk as reference electrode rather than source
4. Provides inversion coefficient as design parameter
5. Implemented in SLiCAP

NMOS level-1 equations

Strong inversion, no small geometry effects:

1. Normal mode (forward) cut-off region V_{DS}>0, V_{GS}<V_{T0}

$$I_{DS} = 0$$

2. Normal mode (forward) linear region V_{DS}>0, V_{GS}>V_{DS}+V_{T0}

$$I_{DS} = \frac{W}{L}\beta(V_{DS} - V_{t0})(1 + \lambda V_{DS})$$

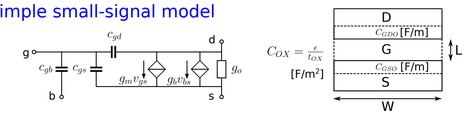
3. Normal mode (forward) saturation region

$$I_{DS} = \frac{W}{L}\frac{\beta}{2}(V_{DS} - V_{t0})^2(1 + \lambda V_{DS})$$

4. Reverse mode V_{DS}<0, swap drain and source
5. V_{T0} changes with bulk voltage (back-gate effect)

$$V_{t0} = \gamma(\sqrt{\phi - V_{BS}} - \sqrt{\phi}) \qquad \phi = \text{surface potential}$$

Simple small-signal model



$C_{OX} = \frac{\epsilon}{t_{OX}}$ [F/m²]

1. Cut-off region

$$c_{gs} = W \cdot C_{GSO} \quad c_{gd} = W \cdot C_{GDO} \quad c_{gb} = WL \cdot C_{OX} \quad g_m = 0 \quad g_o = 0$$

2. Linear region

$$c_{gs} = W \cdot C_{GSO} + \frac{WL}{2}C_{OX} \quad c_{ds} = W \cdot C_{GDO} + \frac{WL}{2}C_{OX} \quad c_{gb} = 0$$

$$g_m = \frac{W}{L}\beta V_{DS} \quad g_o = \frac{W}{L}\beta(V_{GS} - V_{t0})$$

3. Saturation region

$$c_{gs} = W \cdot C_{GSO} + WL\frac{2}{3}C_{OX} \quad c_{ds} = W \cdot C_{GDO} \quad c_{gb} = 0$$

$$g_m = \sqrt{\frac{W}{L}2\beta I_{DS}} \quad g_o = \frac{W}{L}\lambda I_{DS}$$

- small-signal parameters often listed by simulator: (operating point information)
- can be determined with simulation test bench: (see section 4.4.5)
- Drain-bulk and source-bulk junction capacitances can be added at a later stage
- Dominant elements should be g_m and c_{gs}.

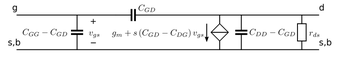
Ward-Dutton capacitance matrix

How to translate the spacial charge distribution in the physical device into a lumped-element model? (see section 4.4.6)

Ward-Dutton: 4-terminal network with capacitive terminal currents:

$$\begin{pmatrix} i_g \\ i_d \\ i_s \\ i_b \end{pmatrix} = s \begin{pmatrix} C_{GG} & -C_{GD} & -C_{GS} & -C_{GB} \\ -C_{DG} & C_{DD} & -C_{DS} & -C_{DB} \\ -C_{SG} & -C_{SD} & C_{SS} & -C_{SB} \\ -C_{BG} & -C_{BD} & -C_{BS} & C_{BB} \end{pmatrix} \begin{pmatrix} v_g \\ v_d \\ v_s \\ v_b \end{pmatrix}$$

translation to hybrid-pi equivalent model:

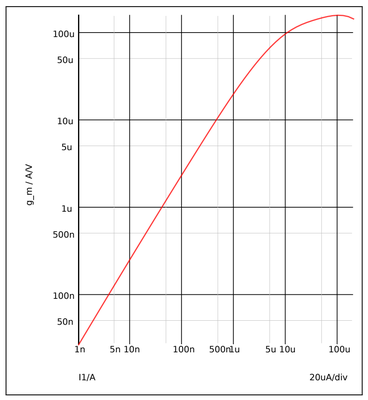


Stationary noise model

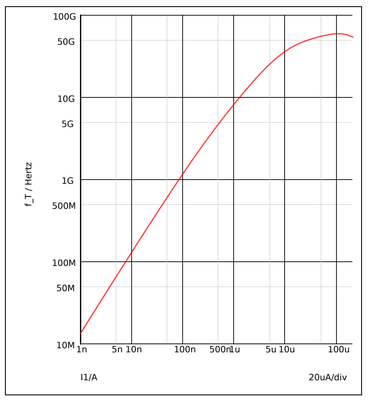
1. Current noise associated with channel current I_{DS}
   - Saturation region:
     $S_{id} = 4kT\frac{2}{3}g_m\left(1 + \frac{f_L}{f}\right)$ [A²/Hz]
   - Linear region:
     $S_{id} = 4kTg_o$ [A²/Hz]
2. 1/f noise due to imperfect atomic structure at gate-channel interface
   - low cross-over frequency f_L: use large device at low current.

Transconductance and cut-off frequency

The cut-off frequency f_T is the frequency at which the current gain equals unity:
$f_T = \frac{g_m}{c_{iss}}$ important figure of merit for an IC process; c_{iss} = input capacitance with shorted output.



CMOS18 NMOS L=180nm W=220nm $g_m(I_{DS})$



CMOS18 NMOS L=180nm W=220nm $f_T(I_{DS})$

Channel current dependency of f_T dominated by that of g_m! Input capacitance c_{iss} only weekly depends on I_{DS}.

Operating point, transconductance efficiency and device scaling

The transconductance efficiency g_m/I_{DS} is a measure for the inversion level at which the device operates:
- Weak inversion g_m/I_{DS} > 30
- Moderate inversion 10< g_m/I_{DS} < 30
- Strong inversion: g_m/I_{DS} <10 .

Increase W:
- Larger current-drive capability
- Increase of max g_m and f_T

Increase L:
- Decrease of CLM and output conductance
- Decrease of max g_m and f_T

Increase inversion level:
- Increase of gm and f_T
- Increase of V_{GS}.

Decrease inversion level:
- Decrease of g_m and f_T
- Improve matching and flicker noise.

Conclusion

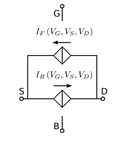
At an early stage of the design, we focus on noise, (load) drive capability and bandwidth.
- For noise and dynamic behavior we use linear models.
- For design of clipping levels we use device scaling.
- For analysis of dynamic nonlinearity we use quasi dynamic eigenvalues (stepping of the small-signal parameters with the operating point. SLiCAP supports this with: stepMethod('array').

We are able to obtain the small-signal parameters from simulation with a single device.
We use device scaling and selection of the inversion level for performance optimization.

Alternative method

Instead of obtaining the small-signal parameters from operating point simulations we can also use expressions that relate them to the technology, the geometry and the operating conditions.
A single expression for all operating ranges enables parameter stepping without changing of expressions.

EKV model



Static model
I_{DS}=I_F-I_R

Oxide capacitance per unit of area: $C'_{OX} = \frac{\varepsilon_{ox}\varepsilon_r}{t_{ox}}$ substrate factor: $n \approx 1 + \frac{C'_{DEP}}{C_{OX}}$

Transconductance factor: $\beta_{sq} = \mu_0 C'_{OX}$ Technology current: $I_0 = 2n\mu_0 C'_{OX} U_T^2 = 2n\beta_{sq} U_T^2$

Weak inversion: $I_{F,R} = I_0 \frac{W}{L}\exp\left(\frac{V_G - V_{T0} - V_{S,D}}{nU_T}\right)$ Strong inversion: $I_{F,R} = \frac{W}{L}\frac{\beta_{sq}}{2n}(V_G - V_{T0} - nV_{S,D})^2$

Transition from weak to strong inversion: $F(x) = \left(\ln\left(1 + \exp\left(\frac{x}{2}\right)\right)\right)^2$ $\exp(x)$ if $x \ll 0$ $\left(\frac{x}{2}\right)^2$ if $x \gg 0$

Function returns forward and reverse inversion coefficient: $IC_{F,R} = F\left(\frac{V_G - V_{T0} - nV_{S,D}}{nU_T}\right)$

Forward and reverse current can be written as a function of the inversion coefficient:
$I_{F,R} = I_0 \frac{W}{L} IC_{F,R} = 2n\beta_{sq} U_T^2 \frac{W}{L} IC_{F,R}$ $I_{DS} = I_F - I_R$

VFMR: $\beta'_{sq} = \frac{\beta_{sq}}{1 + 2\theta U_T\sqrt{IC_F}}$ CLM: $I_{DS} = (I_F - I_R)\left(1 + \frac{V_D - V_S}{V_{AL}L}\right)$ note: only forward operation.

Velocity saturation as mobility reduction: $\beta'_{sq} = \frac{\beta_{sq}}{1 + \left(\theta + \frac{1}{E_{CRIT}L}\right)2U_T\sqrt{IC_F}}$ $IC_{CRIT} \approx \frac{1}{\left(4nU_T\left(\theta + \frac{1}{L \cdot E_C}\right)\right)^2}$

Changes in capacitances from saturation to linear region are described with the aid of inversion coefficients.
Transconductance is also modeled with the aid of the inversion coefficients.

Obtain SLiCAP EKV model parameters from BSIM model simulations

Plot device characteristics in both simulators and adjust parameters.
Sufficently accurate for taking early-stage design parameters.
SLiCAP demo: section 4.5.7!

Inversion coefficient

Forward operation IC_F, saturation region:
- weak inversion: < 0.1
- strong inversion: >10
- moderate inversion: 0.1 .. 10