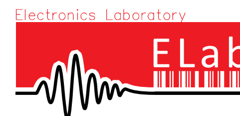
Design analogique en technologies fines

Marc Pastre
Ecole Polytechnique Fédérale de Lausanne, Suisse
marc.pastre@epfl.ch



Outline

- Deep submicron technologies
- EKV model
- Inversion factor based design methodology
- Circuit design
- Digital calibration

Introduction

- Deep-submicrometer CMOS design becomes challenging for the designer:

  CMOS evolution has come to a point where new phenomena need to be taken into account, in particular as far as analog circuits are concerned.

- In order to predict the performances of the CMOS circuits, accurate and efficient MOSFET models should account for most important physical effects in advanced technology.

- Scaling of modern deep submicron CMOS technologies also imposes that the supply voltage is continuously decreased. However, the threshold voltage of the MOSFETs remains almost the same.

  Therefore, modeling should also address weak and moderate inversion.

Deep sub-micron technologies

Acknowledgments

- J-M. Sallese, EPFL
- C. Enz, CSEM & EPFL
- F. Krummenacher, EPFL
- M. Bucher, TUC
- W. Grabinski, GMC Suisse

MOSFET Modeling

Non-ideal characteristics in advanced CMOS transistors.

- *Mobility reduction & Velocity saturation*
- *Drain Induced Barrier Lowering, Channel length modulation, Charge sharing*
- *Reverse short channel effect*

Modeling versus analog parameters.

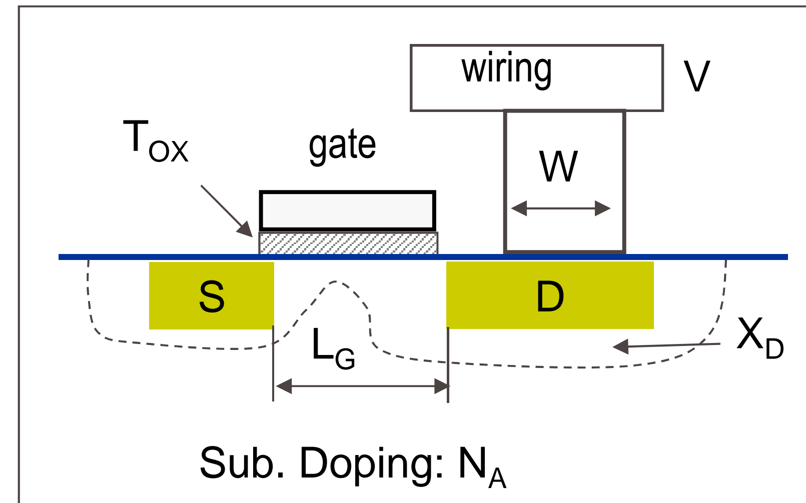
- *Inversion charge linearization.*
- *Pinch-off voltage.*
- *Inversion factor coefficient*
- g_m/I_D *invariant function.*
- *Intrinsic voltage gain and transition frequency*

Introduction
Constant Field Scaling Principles

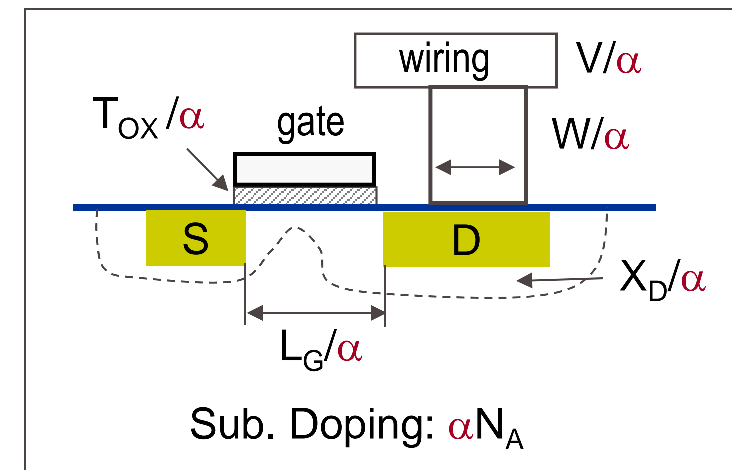
- Scaling consists in changing the physical parameters of a MOSFET so that the scaled device will have similar behavior.

  - A large FET is scaled down by a factor α (> 1) leading to a smaller FET that is expected to have similar behavior.
  - Reducing all voltages and dimensions by the scaling factor and increasing the doping and charge densities by the same factor leads to the same electric field distribution inside the FET:

    constant field scaling
  - Time delay (CV/I) decreases in proportion to $1/\alpha$ and density in proportion to α^2



If Constant field scaling



Introduction – Constant Field Scaling Principles

- Scaling (constant field scaling – $\alpha > 1$)

Quantity	Scaling
Dimensions, L_G, W, T_{OX}	$1/\alpha$
Area	$1/\alpha^2$
Capacitances	$1/\alpha$
Capacitances per unit area	α
Devices per unit of chip area	α^2
Charges	$1/\alpha^2$
Doping concentrations	α
Voltages … and ideally also $\mathbf{V_T}$	$1/\alpha$
Bias currents	$1/\alpha$
Transistor transit time	$1/\alpha^2$
Gate delay	$1/\alpha$
Power dissipation	$1/\alpha^2$

Introduction - Limits to Scaling

- Limitations are mainly due to:
  - Nonscaling of the built-in potential (& junctions).
  - Non scaling of the subthreshold slope.
  - Non scaling of the threshold voltage.

  } Degradation of the OFF state current.

  - Quantum mechanical tunneling currents (gate to channel and source-dain).
  - Discrete nature of dopants in nm scaled devices: matching issues.

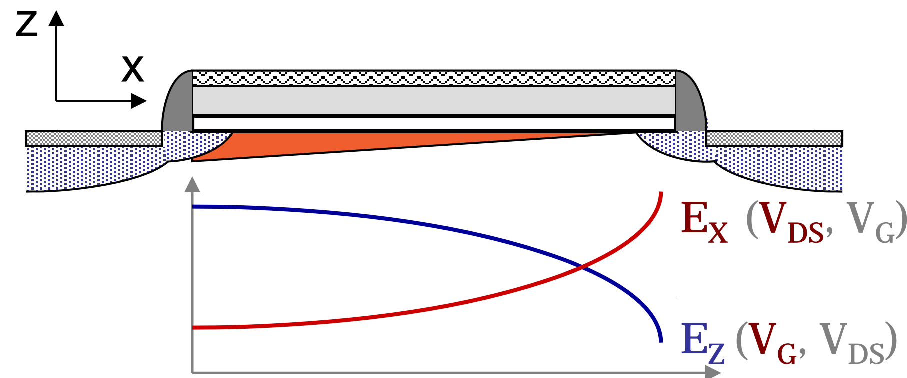
- Lowering of the nominal voltage down to 1 volt will also reduce the available dynamic range, pushing the devices to operate in weak-moderate inversion.

- Then, while scaling of CMOS technology improves digital applications, this evolution is rather detrimental for analog design since it introduces non-ideal characteristics.

Mobility Reduction

- Mobility is the key parameter that impacts the current density in a MOSFET. But moving to more advanced technologies may also degrade it:

  The normal and and lateral electric fields alter the mobility in the channel.



Mobility is then voltage and position dependent

- Causes to mobility reduction due to the normal field E_z:
  - Coulomb scattering μ_c: interaction with ionized impurities (at low field, high doping)
  - Phonon scattering μ_{ph} : interaction with lattice vibrations (at medium field)
  - Surface roughness μ_{rs} : roughness of the Si-SiO_2 interface (at high field)

  With downscaling of MOSFETs, these effects become more pronounced.

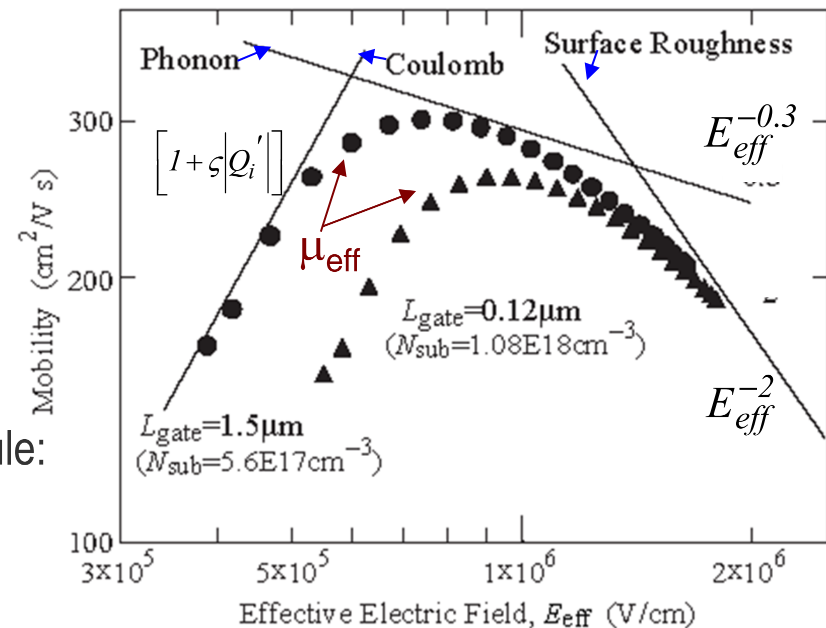
Mobility Reduction

- Along the channel, the normal electric field affects the mobility:
  - This is modelled through an effective field E_{eff} that accounts for the spatial distribution of inversion and depletion charges.

$$E_{eff} = \left| \frac{\eta \cdot Q_I' + Q_B'}{\varepsilon_{SC}} \right| \qquad \begin{cases} \eta_n = 1/2 \\ \eta_p = 1/3 \end{cases}$$

- Low normal field (E_Z) mobility is described by 3 contributions:

$$\begin{cases} \text{Coulomb scattering} & \mu_c \propto \left[1+\varsigma\left|Q_i'\right|\right] \\ \text{Surface scattering} & \mu_{sr} \propto \left[E_{eff}\right]^{-2} \\ \text{Phonon scattering} & \mu_{ph} \propto \left[E_{eff}\right]^{-0.3} \end{cases}$$



- These are combined through Mathiessen's rule:

$$\frac{1}{\mu_{eff}} = \frac{1}{\mu_c} + \frac{1}{\mu_{sr}} + \frac{1}{\mu_{ph}}$$

Mobility Reduction

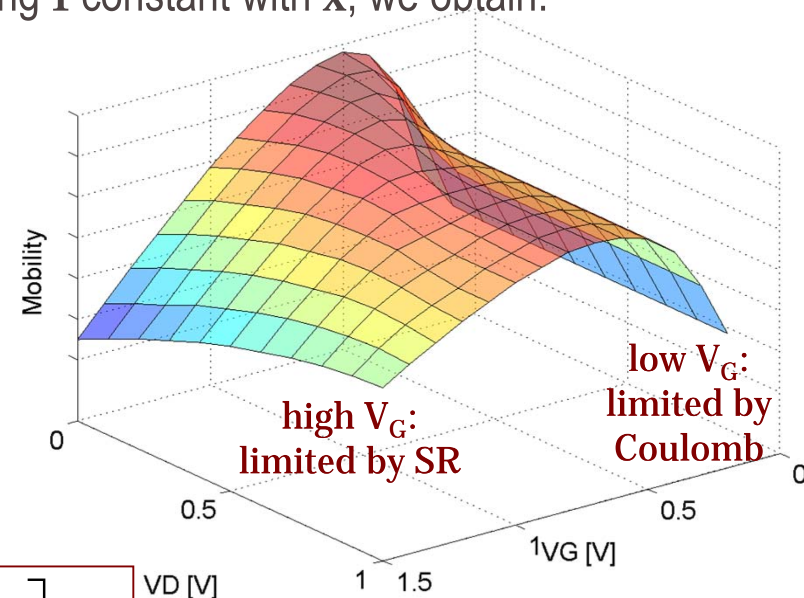
- Integrating the expression of the drain current along the channel, we obtain an equivalent global mobility that is now position independent:
  - The general expression for the drain current is given by $I = W \cdot \mu(x) \cdot F(Q_i)$
    Integrating along the channel, and assuming $\mathbf{I}$ constant with $\mathbf{x}$, we obtain:

$$\int_S^D \frac{I}{\mu(x)} \cdot dx = W \cdot \int_S^D F(Q_i(x)) \cdot dx$$

$$I = \frac{W}{L} \cdot L \cdot \left(\int \frac{1}{\mu(x)} \right)^{-1} \cdot \int_S^D F(Q_i(x)) \cdot dx$$

This leads to an effective mobility:

$$\frac{1}{\mu_{eff}} = \frac{1}{L} \int_S^D \left[\frac{1}{\mu_C(Q(x))} + \frac{1}{\mu_{ph}(Q(x))} + \frac{1}{\mu_{sr}(Q(x))} \right] \cdot dx$$



Bucher M. in 'Trans. Level Modeling for Anal. RF IC Design', Springer

Velocity Saturation

- Near the drain, when V_D exceeds a saturation voltage, high lateral electric fields E_x causes saturation of the carrier drift velocity.
- This effect plays an important role in short-channel transistors, where electric field near the drain has a non negligible extension and gives rise to impact ionization.
  - A semi empirical approach to relate velocity saturation to mobility is given by:

$$v_{drift} = v_{sat} \frac{E_x / E_C}{\left(1 + (E_x / E_C)^\alpha\right)^{1/\alpha}} \quad \Rightarrow \quad \mu \approx \frac{\mu_z}{\left[1 + (E_x / E_C)^\alpha\right]^{1/\alpha}}$$

α : 2 for electrons; 1 for holes

where $E_C = v_{SAT}/\mu_z$ is the critical field at which the carrier velocity starts to saturate:

- Electrons: $v_{sat} \cong 10^5\ m/s$ $\quad E_c \cong 1\ V/\mu m$
- Holes: $v_{sat} \cong 8.10^4\ m/s$ $\quad E_c \cong 3\ V/\mu m$

- μ_z should also include mobility reduction due to the vertical field E_z.

CAUGHEY, D. M.et. al,.Proc. Inst. Elect. Electron.Engrs., 1967, vol. 55, p. 2192

Drain Induced Barrier Lowering (DIBL)

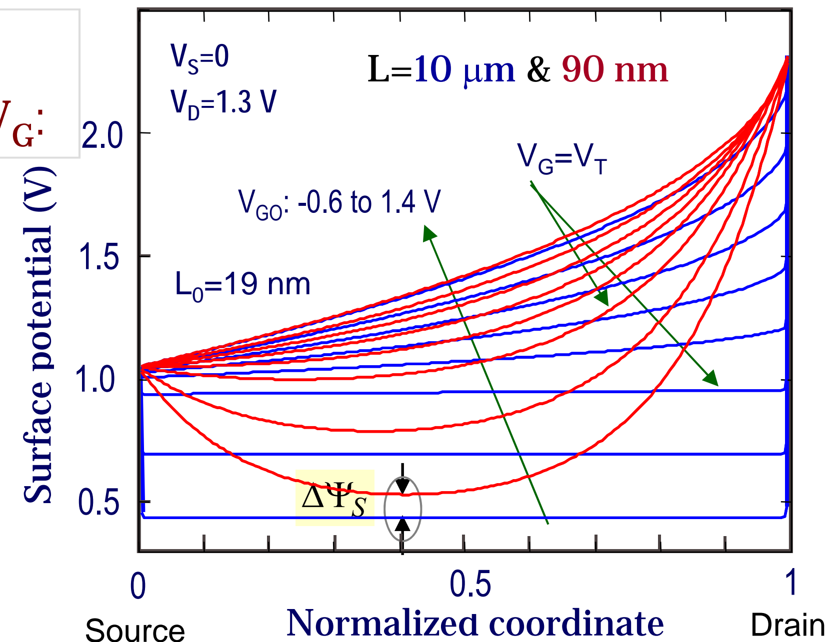
- In short channel devices, the lateral electric field ($\sim V_{DS}$) cannot be neglected with respect to the normal component ($\sim V_G$), *particularly in weak inversion*.
- Electrostatic solution relies upon a quasi 2D description of the channel in WI ($Q_{mob} \ll Q_{dep}$).

DIBL lowers the surface potential in WI by $\Delta\psi_S$.
This can be seen as a decrease in $\mathbf{V_{T0}}$ with V_{DS} & $\mathbf{V_G}$:

$$-\Delta V_{T0} \overset{L_G > 5L_0}{\sim} \left(f(V_G) + V_D\right) \cdot \exp\left(\frac{-L_G}{2 \cdot L_0}\right)$$

where $\mathbf{L_0}$ is a characteristic length:

$$L_0 = \eta(\approx 1) \cdot \sqrt{\frac{\varepsilon_{si}}{C_{ox}} h_{depl}} \sim \frac{1}{C_{ox}^{1/2} \cdot (N_{sub})^{1/4}}$$



- In weak inversion, short channel transistors are significantly affected by a high drain voltage. Reduction of DIBL requires thinner gate oxides and higher substrate doping.

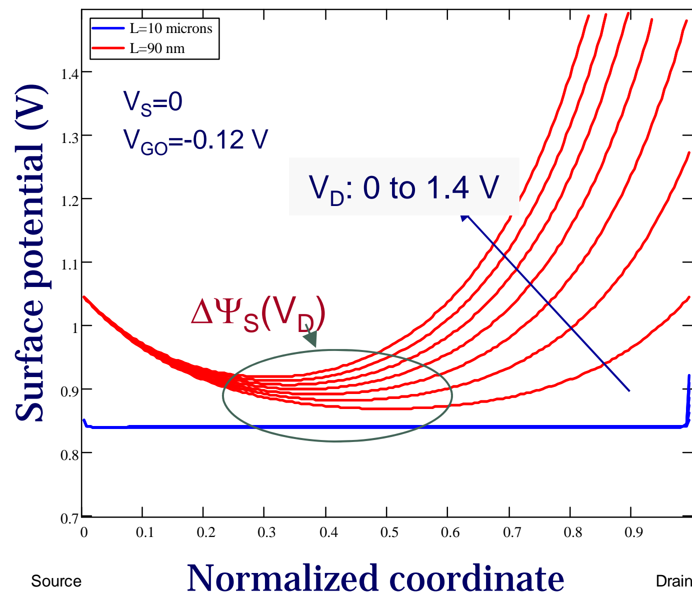
Zhi-Hong Liu et. al., Trans.Elec. Dev. vol. 40, n. 1, p. 86 - 1993

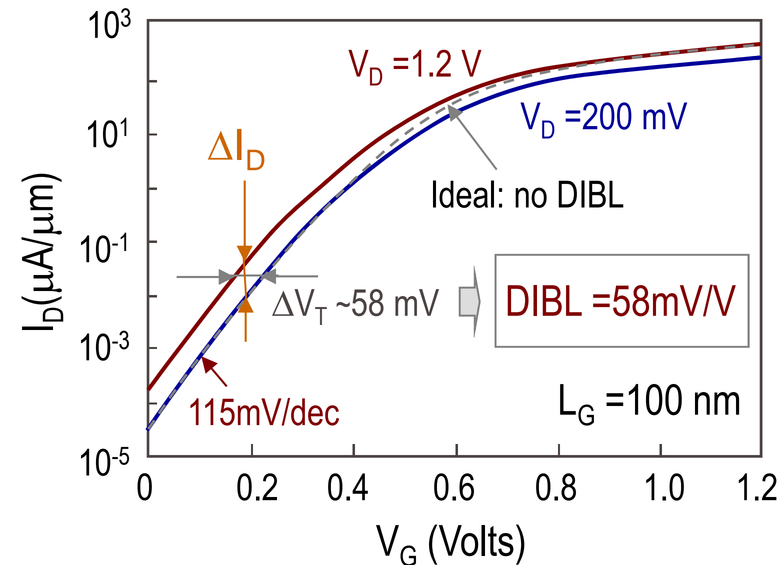
Drain Induced Barrier Lowering (DIBL)

- Lowering of the surface potential with V_D (in WI) will increase the current for short channel devices, eventhough the MOSFET is saturated.
  - *Degradation of the output conductance g_{ds}, maximum voltage gain g_m/g_{ds} and I_{on}/I_{off}.*

$\Delta\Psi_S(V_D) \sim V_D \rightarrow \Delta I_D \sim \exp(\alpha V_D) \rightarrow Ln(\Delta I_D) \sim \alpha V_D$

$\Delta\Psi_S(V_D) \sim f(V_G) \rightarrow Ln(\Delta I_D) \sim f(V_G) \rightarrow$ Degradation of WI Slope

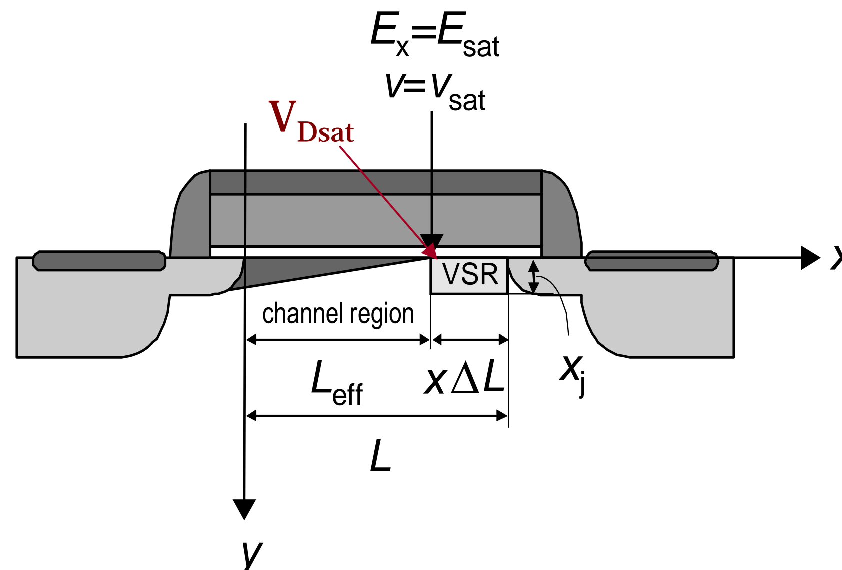




- DIBL is one of the most important short channel effect in weak inversion, together with charge sharing and reverse short-channel effect (RSCE).

Channel Length Modulation

- In Strong Inversion, when $E_x >> E_C$ (saturation),.
- Increasing V_D above V_{DSAT} creates a velocity saturated region at the drain and shifts the pinch-off point toward the source by ΔL.
- In order to maintain the drain current, the inversion charge density at the drain has to reach an asymptotic value depending on I_D.



In addition to CLM, the velocity saturated region is responsible for hot electron generation & substrate current.

H. Wong et. al., Trans.Elec. Dev. vol. 44, n. 11, p. 2033 - 1997

Channel Length Modulation

- From a 2D analysis, the saturated region extension is given by

$$\Delta L \cong L_C \cdot \ln\left(\left[\frac{V_{DS}-V_{DSsat}}{L_C \cdot E_C}\right] + \sqrt{1+\left[\frac{V_{DS}-V_{DSsat}}{L_C \cdot E_C}\right]^2}\right)$$

$$L_C = \sqrt{\frac{\varepsilon_{si} \cdot X_J}{C_{OX}'}}$$

junction depth

The smaller the junction depth and oxide thickness, the smaller CLM effect.

- Definition of the channel length modulation voltage V_M (corresponding to the Early voltage in a bipolar) :

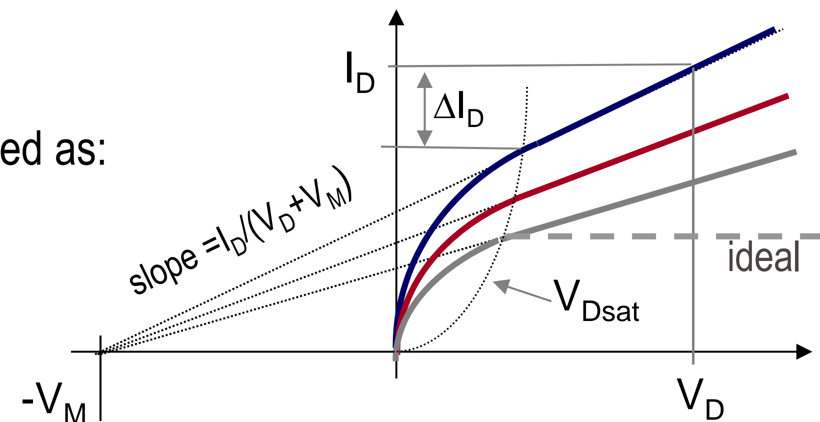
$$g_{ds} = \frac{I_D}{V_M + V_D} = \frac{\Delta I_D}{V_D - V_{Dsat}}$$

- But the output conductance can also be defined as:

$$\frac{\Delta I_D}{I_D} = \frac{\Delta L}{L} \approx \frac{L_C}{L} \cdot \ln\left(1 + \frac{V_{DS}-V_{DSsat}}{L_C \cdot E_C}\right)$$

- Equating both terms gives :

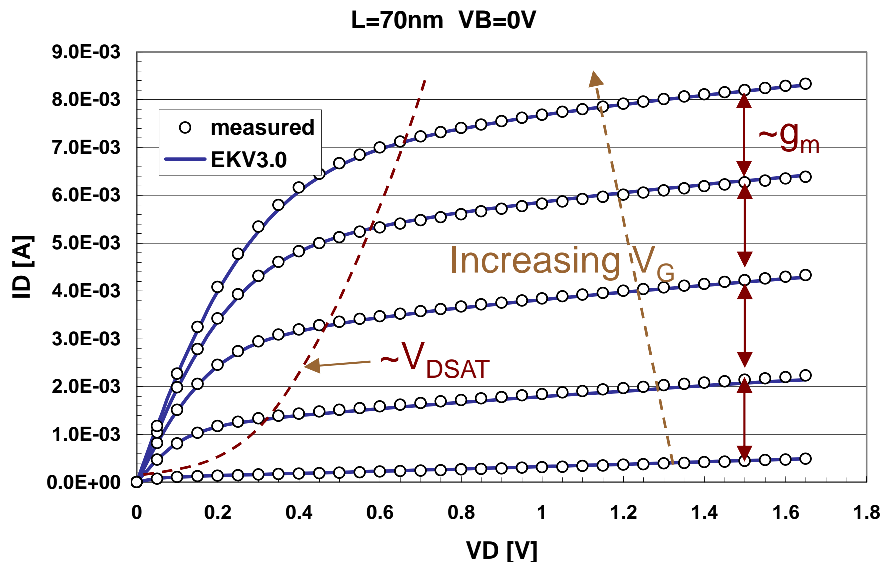
$$V_D + V_M \approx V_M \approx L_G \cdot E_C \quad (V_D - V_{Dsat} << L_C E_C)$$

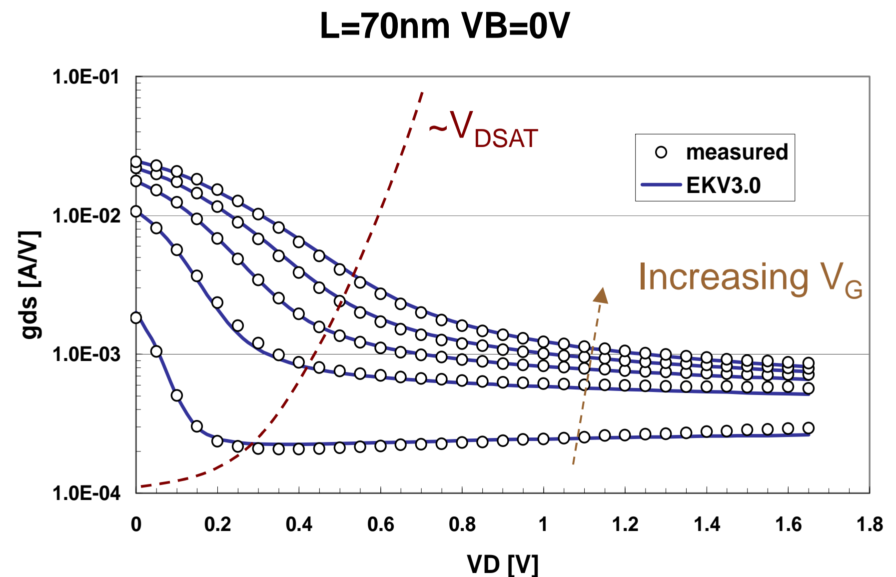


Wong H., TED Vol. 44 (11), p. 2033, 1997

Channel Length Modulation

- Drain current and output conductance in 70 nm effective channel length.
  - The current becomes a linear function of the charge and therefore of the overdrive voltage. $I_D \overset{vel.sat.}{\approx} (V_{GO} - V_{Dsat} - 2\phi_F) \cdot C_{OX} \cdot v_{sat} \cdot W$ ⇒ $g_m \overset{vel.sat.}{\approx} C_{OX} \cdot v_{sat} \cdot W$ indp. of L !
  - g_{ds} (in saturation) is dominated by channel length modulation.





Bucher M. in ‘Trans. Level Modeling for Anal. RF IC Design’, Springer

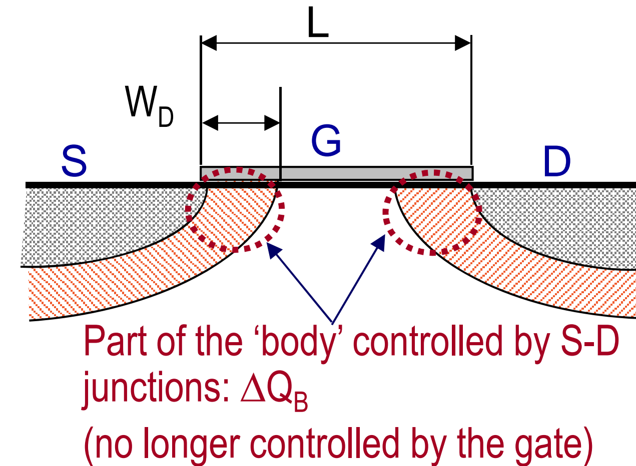
Charge Sharing

- Charge sharing affect short channel devices:

  Depletion from source and drain junctions overlap with the channel region and modify the average doping of the bulk, and hence reduce V_T.

  – These extensions are given by:

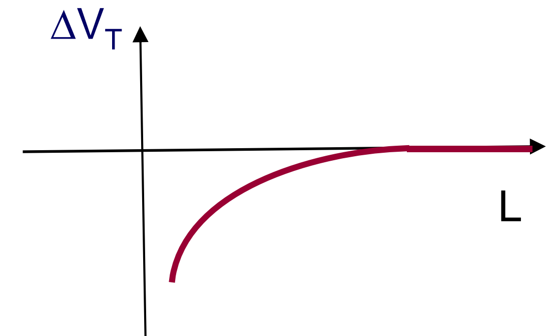
$$W_{S,D} \approx \sqrt{\frac{2 \cdot \varepsilon_{sc}}{q \cdot N_{sub}} \cdot \left(V_{bi} + V_{S,D}\right)}$$



- Rough estimation of the threshold voltage variation:

$$\approx -C_{ox} \cdot \gamma \cdot \sqrt{2\phi_F} \quad \text{and} \quad 2\phi_F \approx 1V$$

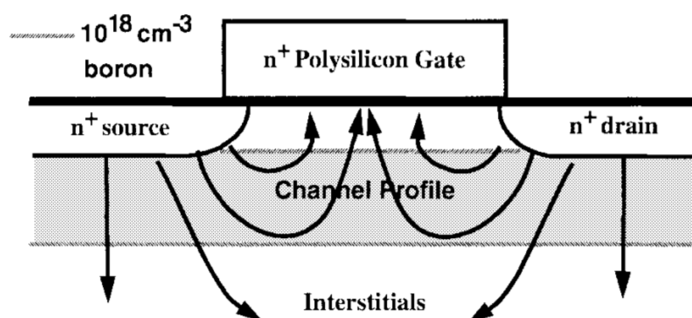
$$\Delta V_T \approx \frac{-\Delta Q_B}{C'_{ox}} \approx \frac{1}{C'_{ox}} \cdot \left[Q_{B@V_T} \cdot \left(W_D + W_S\right)\right] \cdot \frac{1}{L} \approx \gamma \cdot \left(\frac{W_D + W_S}{L}\right)$$

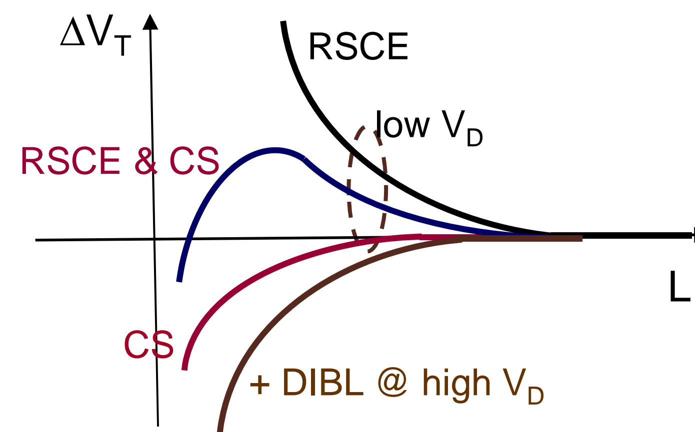


  – *The same approach applies to narrow channels, but then V_T increases instead of decreasing.*

Reverse Short Channel Effect.

- Whereas the threshold voltage is expected to decrease when decreasing the gate length, non-uniform lateral doping profile and dopant diffusion at the Si/SiO_2 interface may result in an increase in V_T.
- The behavior of the threshold voltage rollup is strongly influenced by S/D implant doses and anneals.
- ΔV_T can reach few 100 mV in more advanced processes, and $\Delta V_T \sim T_{OX}$.

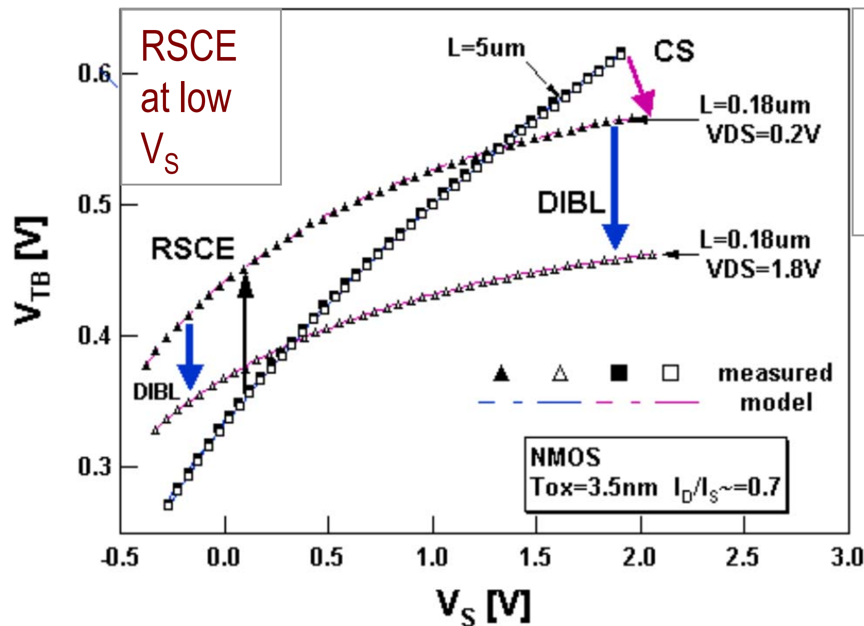




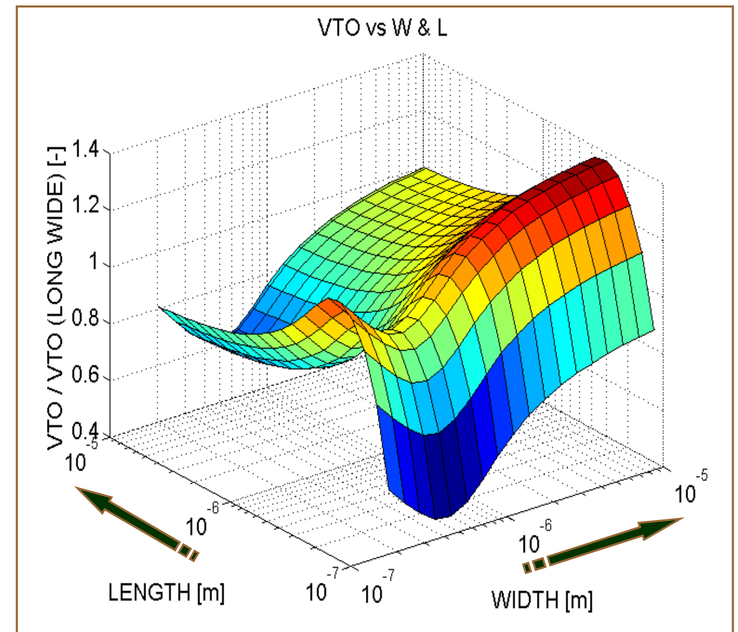
Mazuré, EDL vol. 10, n.12, 1989 – Rousseau EDL vol 18, n.2 1997

Combination of DIBL, Charge Sharing, RSCE

- Combination of all effects:
  - V_T roll-up: Non-uniform doping/oxide charge (RSCE)
  - V_T roll-off: Charge-sharing and DIBL
    - Charge Sharing (CS) reduces substrate effect
    - DIBL reduces threshold voltage @ short L, high V_D



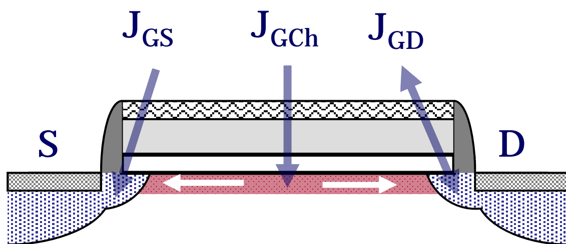
V_T versus length and width including DIBL, CS, vel. Sat. and CLM effects @ L_G = 70nm.



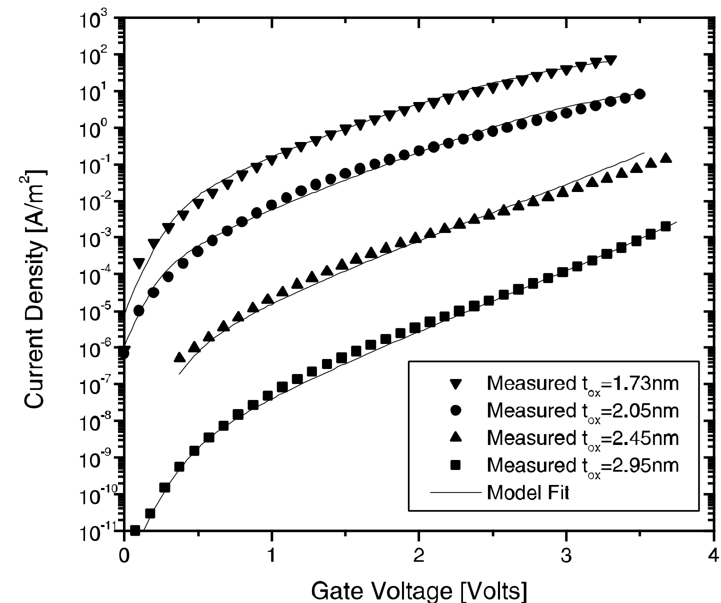
Bucher M., WCM 2004.

Gate Tunelling Current

- With the downscaling of MOS devices, gate equivalent oxide thicknesses have been reduced down to 1.2 nm.
  - T_{ox} is the main technological factor for the gate tunneling current .
- Depending on the biasing voltages, the gate current can become relatively large and affects source and drain currents. It has distinct components:
  - Gate to 'channel' tunneling current (intrinsic).
  - Gate to source/drain overlap tunneling current (extrinsic), enhanced in scaled MOSFETs.



Moving from 3 to 1.7nm SiO_2 thickness increases the gate leakage current by more than 10^6



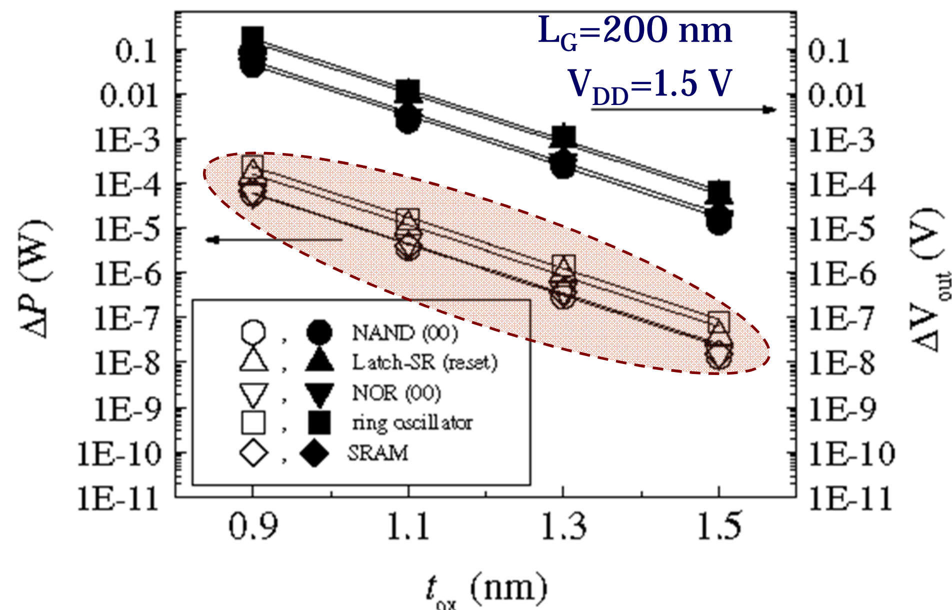
Ranuarez et al. Microelectronics Reliability vol. 46, p. 1939 (2006)

Clerc R. et.al. SSE vol. 43, p. 1705 (2001)

Gate Tunneling Current

- At 0.9 nm, direct tunneling currents are comparable with drain currents.
  - Tends to become the main contribution to sub-threshold current in a thin-oxide MOSFET.

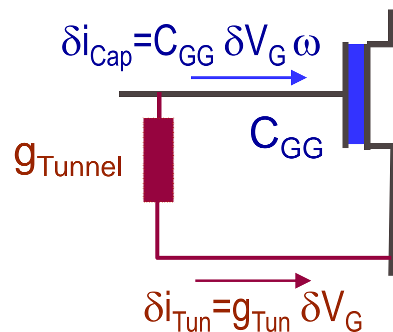
A consequence of the tunneling current is an exponential increase in static power dissipation with respect to T_{OX}



A. Marras et al. , Microelectronics Reliability vol. 45, p. 499 (2005)

Gate Tunneling Current

- From the gate impedance point of view, the gate current can be seen as a tunnel conductance in parallel to the conventional input capacitance.

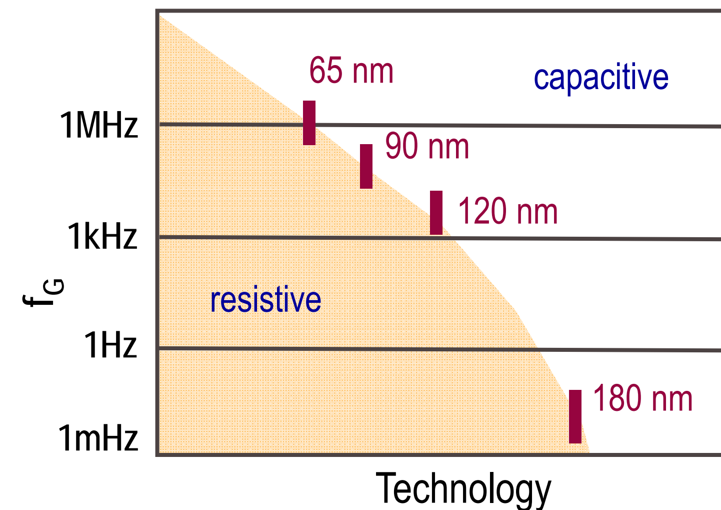


$$f_G = \frac{g_{Tun}}{2\pi C_{GG}}$$

(~ independent of WL for wide/long devices dominated by intrinsic leakage)

- $f > f_G$ ⟶ Capacitive input impedance
- $f < f_G$ ⟶ Resistive input Impedance

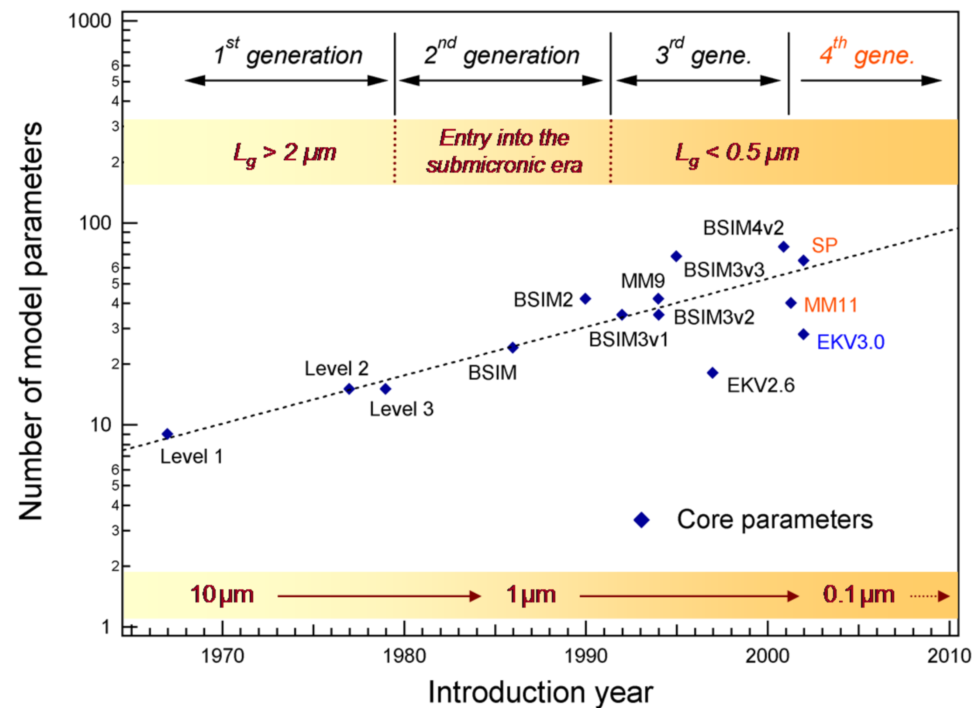
Input impedance appears resistive below 0.1 Hz in 180-nm technologies and below 1 MHz in 65-nm CMOS.



Since MOS gates are no longer ideally insulated, storage nodes retention time is decreased, thus increasing the minimum operating frequency (needed to read that node).

Compact Models

- Increasing importance of short channel effects, including quantum mechanical and non static effects, leads to an increase in the model complexity and number of model parameters that requires the development of a new class of surface-potential or charge based models model.
  - Ideally, physics based compact models should also be predictive with respect to some basic technological parameters.



Compact Models

- BSIM 3v3, BSIM 4: threshold-voltage-based models.
  - Regional approximations
  - Smoothing functions used as *a model* (e.g. drain current)
  - Increasing complexity, dramatic number of parameters

3rd generation

- (MM11,SP), PSP: surface-potential-based models
  - Models close to physics
  - Explicit formulation of the surface potential
  - Symmetric model structure (idem EKV)
- EKV 3.0: charge linearization model
  - Bulk used as a reference: symmetric model structure
  - Alternative to surface-potential-based models.
  - EKV is a design oriented model.

4th generation

Example of EKV3.0 model parameter list…

Visit http://legwww.epfl.ch/ekv/

- **Setup pars.**
  - SIGN 1 (nmos), -1 (pmos)
  - TG -1 (enhancement)
  - SCALE scaling factor L, W
  - QOFF charge model off
- **Oxide, Substrate and Gate Doping related pars. (7)**
  - COX oxide capacitance
  - XJ junction depth
  - VTO threshold voltage
  - PHIF fermi-bulk voltage
  - GAMMA body factor
  - GAMMAG gate factor
  - N0 long channel slope
- **Quantum Mechanical effect (3)**
  - AQMA QME accumulation
  - AQMI QME inversion
  - ETAQM QME coefficient
- **Vertical Field Mobility effect (6)**
  - K transconductance fact.
  - E0 1st order coefficient
  - E1 2nd order coefficient
  - ETA QB and QI balance
  - ZC Coulomb sc. Par. 1
  - THC Coulomb sc. Par. 2
- **Mobility geometrical pars. (4)**
  - LA char. mobility length A
  - LB char. mobility length B
  - KA char. mobility factor A
  - KB char. mobility factor B
- **Velocity Saturation & CLM (4)**
  - UCRIT critical long. field
  - DELTA order of vsat model
  - LAMBDA CLM effect
  - ACLM pocket implant factor

- **Long-channel VT & RSCE (5)**
  - LVT VTO corr. char. length
  - AVT VTO corr. factor
  - LR RSCE char. length
  - QLR RSCE factor charge
  - NLR RSCE factor doping
- **INWE (3)**
  - WR INWE char. length
  - QWR INWE factor charge
  - NWR INWE factor doping
  - Charge Sharing effect (5)
  - LETA0 Long-ch. CS factor
  - LETA 1st order CS factor
  - LETA2 2nd order CS factor
  - NCS CS slope factor degr.
  - WET Narrow-ch. CS factor
- **DIBL effect (2)**
  - ETAD char. length factor DIBL
  - SIGMAD bias factor DIBL
- **Halo-related gds degradation (5)**
  - FPROUT
  - PDITS
  - PDITSL
  - PDITSD
  - DDITS
- **Gate current pars. (3)**
  - XB crit. difference potential
  - EB crit. electrical field
  - KG transc. factor Igate
- **Impact ionization (3)**
  - IBA II current factor A
  - IBB II current factor B
  - IBN II current coefficient

- **Geometrical pars. (10)**
  - DL gate length offset
  - DLC gate length CV offset
  - DW gate width offset
  - DWC gate width CV offset
  - LDW short-ch. DW correct.
  - WDL narrow-ch. DL correct.
  - LL hyperbolic length fact.
  - LLN exp. Length fact.
  - XL
  - XW
  - LIBB length scaling IBB
- **Width Scaling pars. (9)**
  - WE0 width scaling E0
  - WE1 width scaling E1
  - WUCRIT width scaling UCRIT
  - WLAMBDA width scaling LAMBDA
  - WETAD width scaling ETAD
  - WQLR width scaling QLR
  - WNLR width scaling NLR
  - WLR width scaling LR
  - WIBB width scaling IBB
- **Overlap & fringing capacitance (6)**
  - GAMMAOV overlap body factor
  - VFBOV overlap flat-band voltage
  - LOV overlap length
  - VOV overlap bias factor
  - KJF inner fringing cap. par.
  - CJF inner fringing cap. Factor
- ….. + 12 Temp- Params

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Electronics Laboratory ELab

The BSIM-EKV Modeling Partnership Announcement

BSIM and EKV groups have agreed to collaborate on the long-term development and support of BSIM6 as a world-class open-source MOSFET SPICE model.

EKV model

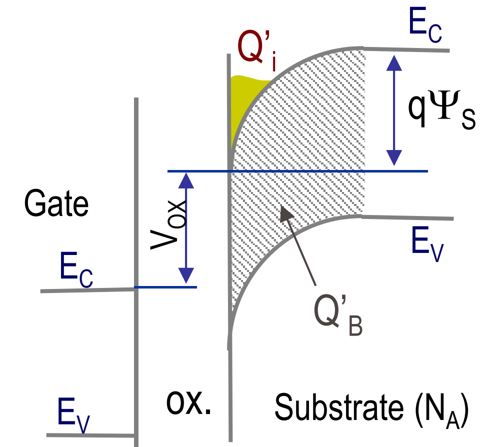
Charge Linearization Concept.

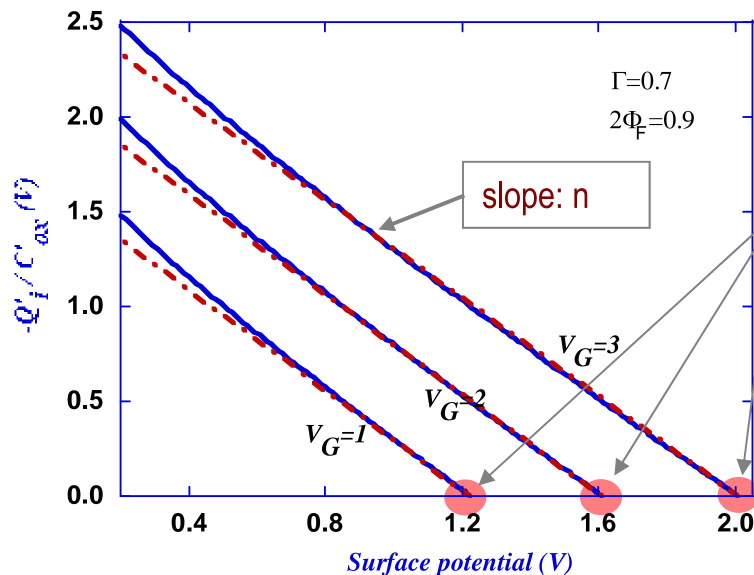
- Potential drop and charge linearization: (the bulk is the reference)

$$V_G = V_{FB} + \psi_S - \frac{Q_B'}{C_{OX}'} - \frac{Q_I'}{C_{OX}'} \quad (V_{OX})$$

$$\frac{-Q_B'}{C_{ox}'} \approx \frac{\sqrt{2 \cdot q \cdot \varepsilon \cdot N_A}}{C_{ox}'} \cdot \sqrt{\psi_S}$$

$$V_G = V_{FB} + \psi_S + \Gamma \cdot \sqrt{\Psi_S} - \frac{Q_I'}{C_{OX}'}$$





pinch-off surface potential : $\Psi_P(V_G) = \Psi_S(Q_i=0)$

$$\frac{Q_i'}{C_{ox}'} = n \cdot (\Psi_S - \Psi_P)$$

The pinch-off surface potential will then depend on V_G

Slope factor:

$$n \approx 1 + \frac{\Gamma}{2 \cdot \sqrt{\Psi_P}}$$

$\Phi_F = U_T \, Ln(N_A/n_i)$ ~0.4V for 10^{17} cm^{-3}

n will then also depend on V_G

Sallese J.M. et. al. SSE vol.47, p.677 (2003)

The Pinch-Off Voltage: link between gate and channel.

- We notice that in strong inversion, the source and drain potentials are linked to the surface potential through:

$$V_{S,D} \overset{SI}{\approx} \psi_s^{S,D} - 2 \cdot \phi_F$$

- Therefore, in strong inversion, the charge density at source and drain can as well be written in terms of the source and drain potentials.

$$\frac{Q'_{S,D}}{n \cdot C'_{ox}} = \left[\Psi_s^{S,D} - 2\phi_F\right] - \left[\Psi_P - 2\phi_F\right] \overset{SI,Lin}{\sim} V_{S,D} - \left[\Psi_P - 2\phi_F\right]$$

- From this relation, we propose to define a 'pinch-off' voltage V_P as: $V_P = \Psi_P - 2\phi_F$

  As for ψ_P, V_P will only depend on the gate voltage.
  V_P is the potential of the gate that as 'seen' from the channel.

- In strong inversion, the relation between charges and potentials can then be written as:

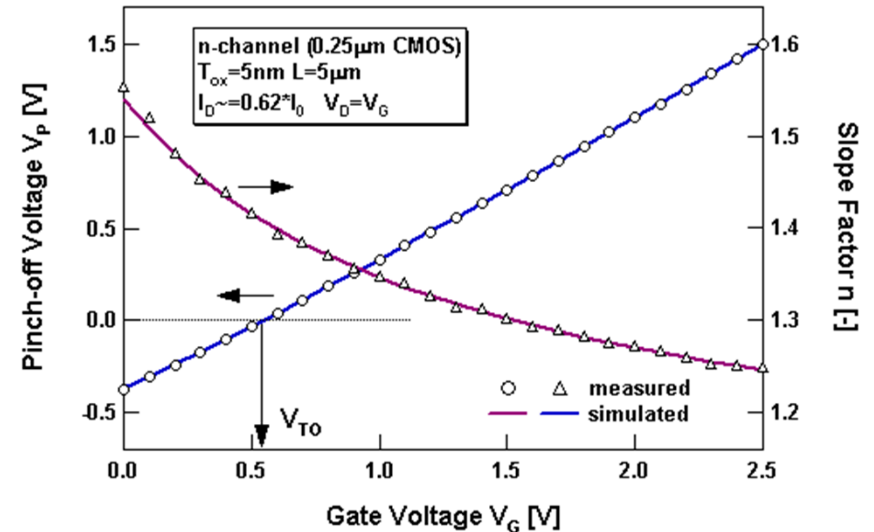
$$-Q'_{i@D,S} \overset{SI}{\approx} n \cdot C_{ox} \cdot \left(V_p - V_{S,D}\right)$$

The Pinch-Off Voltage: approximate expression.

- An interesting approximation of the pinch-off is given by:

$$V_p \approx \frac{V_G - V_{T0}}{n}$$

Where n is still the slope factor.



- More intuitively, the slope factor can also be obtained from a capacitance representation of the device:

$$\frac{dV_G}{dQ_G} = \frac{d\psi_S}{dQ_G} + \frac{dV_{ox}}{dQ_G} = \left(\frac{-dQ_I}{d\psi_S} - \frac{dQ_B}{d\psi_S} \right)^{-1} + \frac{1}{C_{ox}} = \frac{1}{C_D - n \cdot C_{ox}} + \frac{1}{C_{ox}}$$

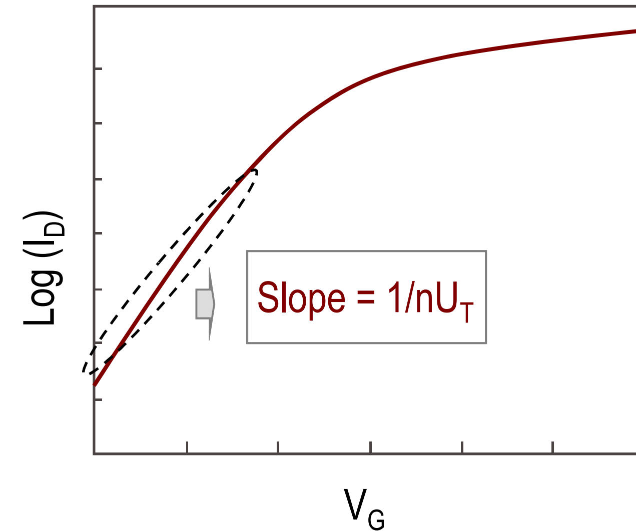
C_D: depletion capacitance which depends on the doping of the substrate.

- Keeping V_G fixed, we obtain:

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

What about weak inversion ?

- *n* is also the slope of the Log (I_D) vs V_G characteristic in weak inversion that equals $1/nU_T$
  - *n* varies between 1.1 (strong inversion) up to 1.6 (weak inversion).



- As for strong inversion, we can show that in weak inversion, charge densities at source and drain are still function of the difference between V_P and V_S or V_D:

$$Q_{S,D} = -2 \cdot n \cdot C_{ox} \cdot U_T \cdot e^{\frac{V_P - V_{S,D}}{U_T}}$$

- Without demonstration, the general relation between charges and potentials is :

$$Ln\left(\frac{-Q_{S,D}}{2 \cdot n \cdot C_{ox} \cdot U_T}\right) - 2 \cdot \frac{Q_{S,D}}{2 \cdot n \cdot C_{ox} \cdot U_T} = \frac{V_P - V_{S,D}}{U_T}$$

Drain Current & normalization.

- Charge linearization can also be used in the current expression.
  - Adopting the drift-diffusion transport model, we have:

$$I_D = \mu W(-Q_i' \frac{d\psi_S}{dx} + U_T \frac{dQ_i'}{dx}) \Longrightarrow I_D = \mu W(-Q_i' \frac{dQ_i'}{dx} \cdot \frac{1}{n \cdot C_{ox}} + U_T \frac{dQ_i'}{dx})$$

- Since I_d is constant along the channel, integration from S to D gives:

$$I_D \cdot L = \mu W \int_S^D \left(-\frac{Q_i}{nC_{Ox}} + U_T \right) \cdot dQ_i \Longrightarrow I_D = 2n\mu C_{OX} U_T^2 \frac{W}{L} \left(\left[\frac{Q_i}{2nC_{OX}U_T} \right] - \left[\frac{Q_i}{2nC_{OX}U_T} \right]^2 \right) \Bigg|_S^D$$

(assuming a constant mobility μ)

- We define a specific current I_{SP} and specific charge density Q_{SP} that depend only on the technological parameters:

$$I_{SP} = 2n\mu C_{OX} U_T^2 \frac{W}{L} = 2nK_P \frac{W}{L} U_T^2 \quad \text{with} \quad K_P = \mu C_{OX}$$

$$Q_{SP} = 2nC_{OX} U_T$$

- The current can be written:

$$\frac{I_D}{I_{SP}} = \left[\left(\frac{Q_S}{Q_{SP}} \right)^2 - \left(\frac{Q_S}{Q_{SP}} \right) \right] - \left[\left(\frac{Q_D}{Q_{SP}} \right)^2 - \left(\frac{Q_D}{Q_{SP}} \right) \right]$$

Sallese J.M. et. al. SSE vol.47, p.677 (2003)

Drain Current & transconductances.

- Since the inversion charge densities at source and drain are always given by a general form involving the difference between V_P and the source and drain potentials V_S and V_D:

  $Q_D = F(V_P - V_D)$ and $Q_S = F(V_P - V_S)$

  The current takes a simple form:

$$I_D = H(V_P - V_S) - H(V_P - V_D)$$

- This will have important implications in small signal analysis (even under NQS):
  - A variation δV of V_P is equivalent to a **simultaneous** variation $-\delta V$ of V_D and V_S

$$\frac{\partial I_D}{\partial V_G} = \frac{\partial I_D}{\partial V_P} \cdot \frac{\partial V_P}{\partial V_G} = \left(\frac{\partial I_D}{\partial(-V_S)} - \frac{\partial I_D}{\partial(-V_D)} \right) \cdot \frac{1}{n} \Longrightarrow g_m = \frac{g_{ms} - g_{md}}{n}$$

- In saturation, g_{md}~0 and we obtain: $g_m \overset{sat}{\approx} \dfrac{g_{ms}}{n}$

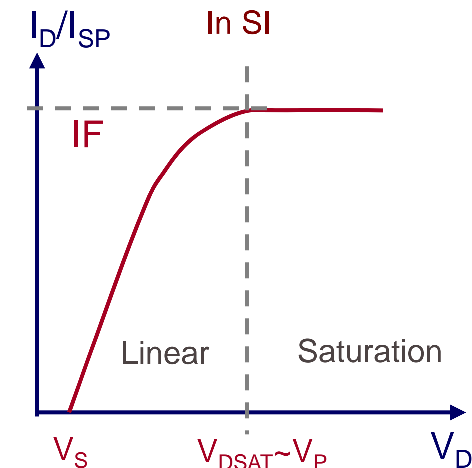
The Inversion Factor IF.

- In saturation, the drain current doesn't change any more when increasing the drain voltage above V_{DSAT}, which means a negligible mobile charge at the drain:
  - Therefore, in strong inversion, V_{DSAT} is almost V_P.

$$-Q'_{i@D} \overset{SI,sat}{\approx} n \cdot C_{ox} \cdot \left(V_p - V_{Dsat}\right) \approx 0$$

$$\frac{I}{I_{SP}} \overset{sat}{\approx} \left[\left(\frac{Q_S}{Q_{SP}}\right)^2 - \left(\frac{Q_S}{Q_{SP}}\right)\right]$$

    - *Note that $V_D = V_G$ ensures saturation since $V_D(=V_G) > V_P$.*



- By definition, the inversion factor IF is the normalized current of the device operating in saturation:

$$IF \overset{SAT}{=} \frac{I}{I_{SP}}$$

$$-Q'_{i@S} \overset{SI,sat}{\approx} n \cdot C_{ox} \cdot \left(V_p - V_S\right)$$

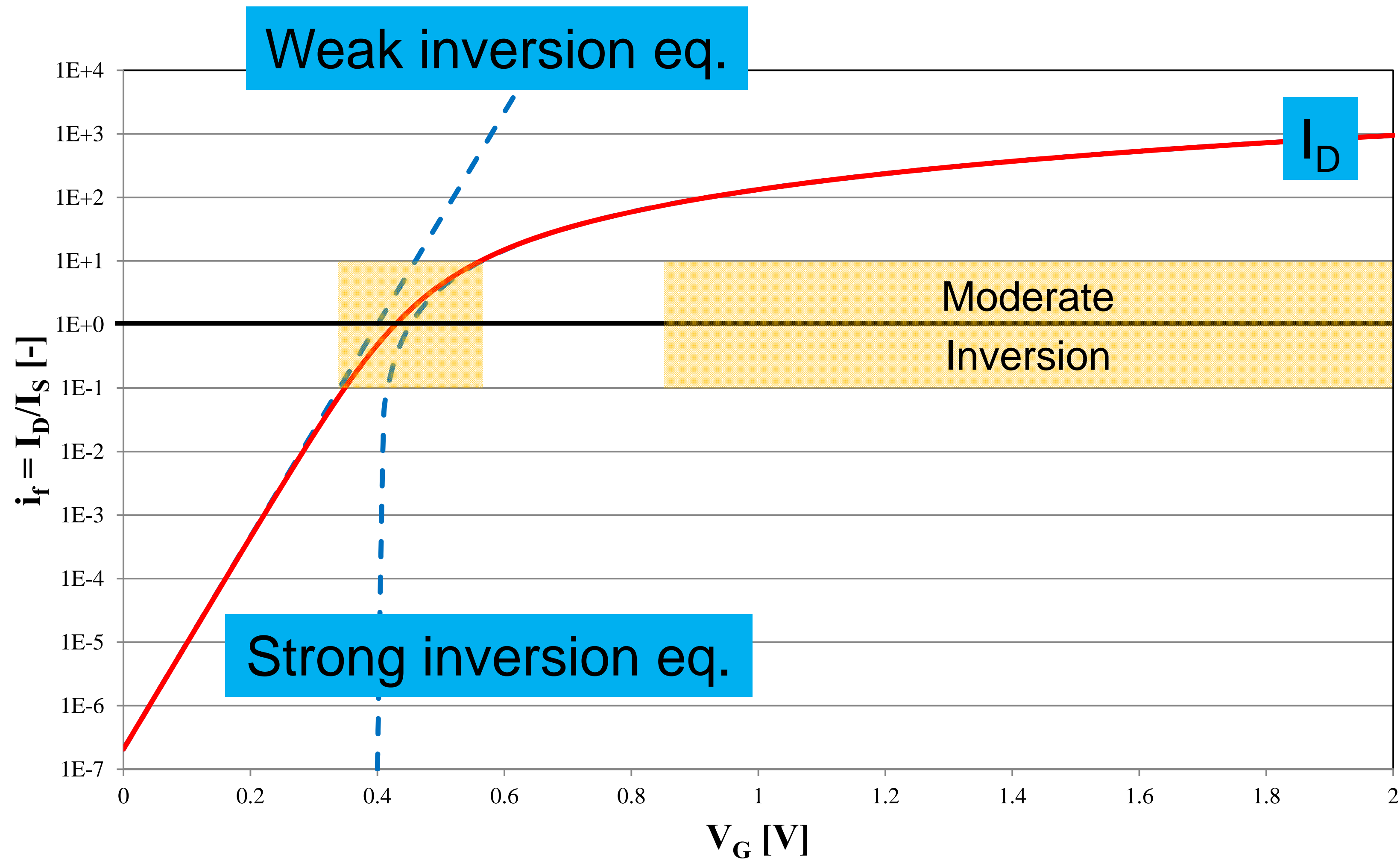
$$\frac{I}{I_{SP}} \overset{SI,sat}{\approx} \left(\frac{Q_S}{Q_{SP}}\right)^2$$

in strong inv. ⇒

$$IF \overset{SI,SAT}{\approx} \left(\frac{V_P}{2U_T}\right)^2$$

'IF' will now 'replace' V_G'

V_G vs IF

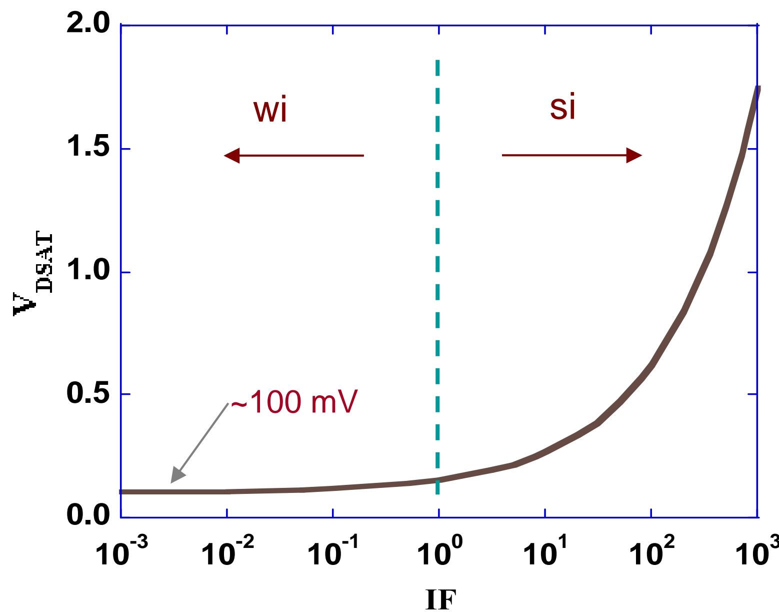


The Saturation Voltage.

- In strong inversion, saturation is then given by : $V_{DSat} \overset{SI}{\approx} V_P \approx 2U_T \cdot \sqrt{IF}$

- In weak inversion, the saturation voltage can be approximated to : $V_{DSat} \overset{WI}{\approx} 4U_T$

- Therefore, we can define a ***semi-empirical relation*** for the saturation voltage as a function of the Inversion Factor IF that covers all modes of operation from weak to strong inversion:



$$V_{D_SAT} \approx U_T \cdot \left(4 + 2\sqrt{IF}\right)$$

Extraction of the I_{SP} Parameter

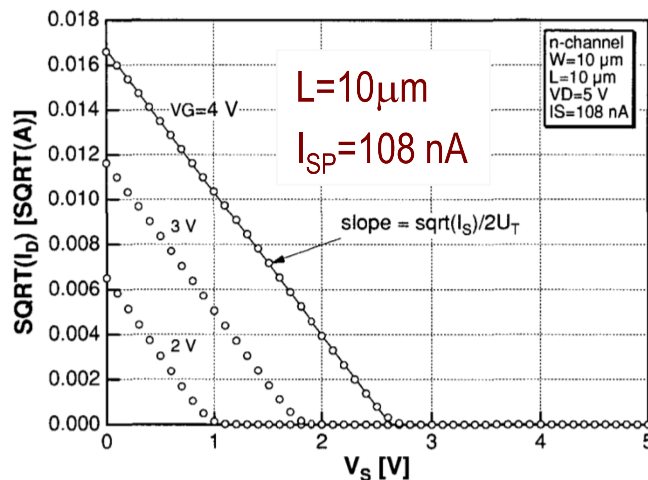
Extraction of the Specific Current I_{SP}

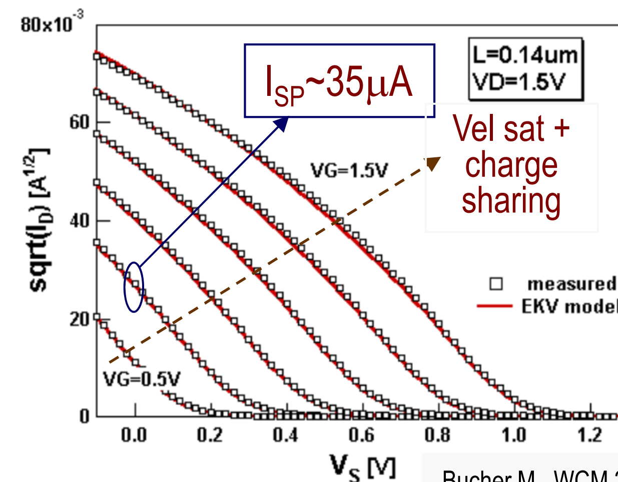
- Assuming strong inversion and saturation, the drain current can be written as: (neglecting diffusion)

$$I_D \overset{SI,SAT}{\approx} I_{SP}\left[\frac{Q_S}{2nC_{OX}U_T}\right]^2 \approx I_S\left[\frac{V_P - V_S}{2U_T}\right]^2$$

- The specific current is then obtained from the slope of $I_D^{1/2}$ versus V_S:

$$\sqrt{I_{SP}} \overset{SI,SAT}{\approx} 2U_T \cdot \frac{-\partial\sqrt{I_D}}{\partial V_S}$$





Bucher M., WCM 2004.

Noise versus Inversion Factor

- Under quasi-static operation, the drain current is constant along the channel.
  - Integrating from source ($x=0$) to x gives $Q_i(x)$ as solution of :

$$\frac{x}{L}\cdot\frac{I_D}{I_{SP}}\overset{Sat}{=}\frac{x}{L}\cdot IF=\left(\left[\frac{Q_i}{Q_{SP}}\right]-\left[\frac{Q_i}{Q_{SP}}\right]^2\right)\Bigg|_S^{Q_i(x)}$$

$$\Rightarrow \quad Q_i\left(\frac{x}{L}\right) \quad versus \quad IF$$

- In saturation, the drain current noise spectral density becomes:

$$S_{\delta I_{nd}^2}=\frac{4\cdot k_B\cdot T\cdot\mu}{L_{elec}^2}\cdot\left[W\cdot\int|Q_i|dx\right]\overset{Sat}{=}4\cdot e\cdot I_{SP}\cdot\left[\frac{1}{12\cdot IF}\cdot\left(\left(1+4IF\right)^{3/2}-1\right)-\frac{1}{2}\right]$$

- Approximation in strong inversion:

$$S_{\delta I_{nd}^2}\overset{SI-Sat}{\approx}4\cdot e\cdot I_{SP}\cdot\frac{2}{3}IF^{1/2}$$

- Approximation in weak inversion:

  ~ Shot noise in junctions .

$$S_{\delta I_{nd}^2}\overset{WI-Sat}{\approx}2\cdot e\cdot I_{SP}\cdot IF$$

The g_{ms}/I_D Invariant.

- An equivalent formulation of the current is given by:

$$I_D = \mu W(-Q_i')\frac{dV_{ch}}{dx} \quad \xrightarrow{\text{Integration along the channel}} \quad I_D = \int_{V_S}^{V_D} \beta \cdot \frac{-Q_i}{C_{ox}} \cdot dV$$

- Then, from the definition of the source transconductance, we obtain:

$$g_{ms} = -\frac{\partial I_D}{\partial V_S} = \beta \frac{-Q_S}{C_{ox}} = 2n\mu C_{ox}\frac{W}{L}U_T \cdot \left(\frac{-Q_S}{2nC_{ox}U_T}\right) = \frac{I_S}{U_T} \cdot \left(\frac{-Q_S}{2nC_{ox}U_T}\right)$$

- In addition, in saturation, Q_D~0 and Q_S is related to the Inversion Factor:

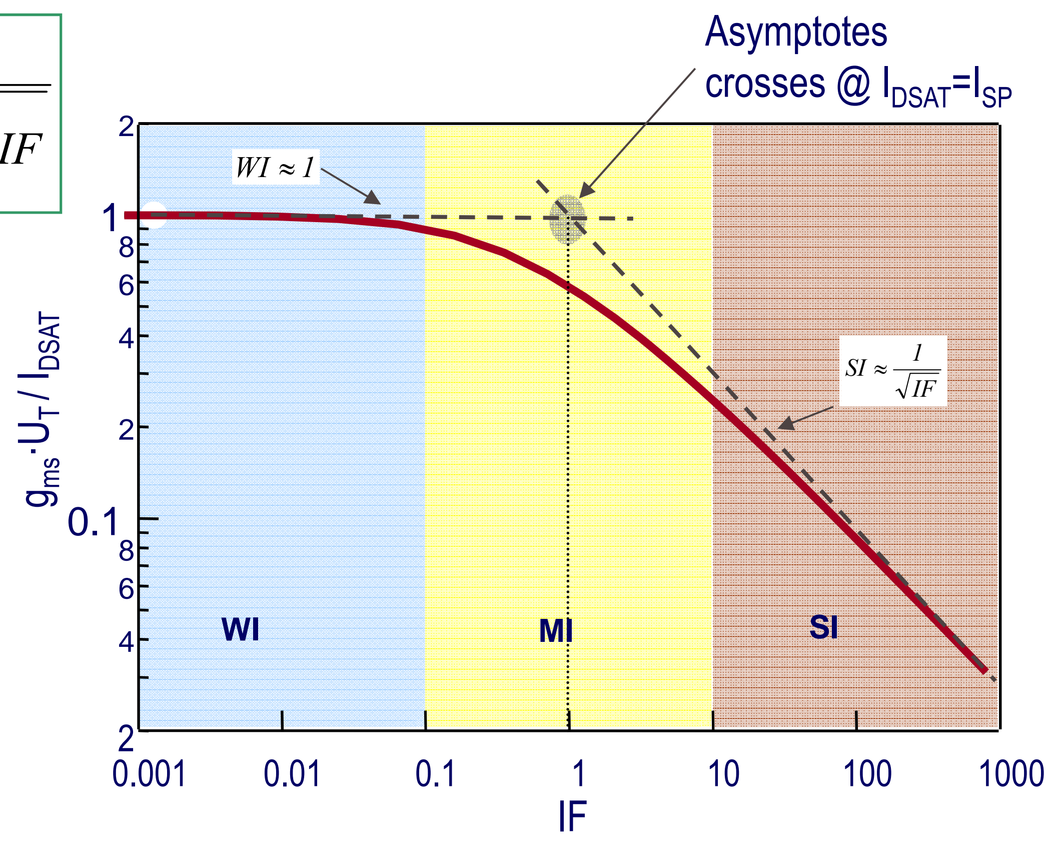
$$I_D^{SAT} = I_{SP} \cdot \left(\left[\frac{Q_S}{2nC_{OX}U_T}\right]^2 - \left[\frac{Q_S}{2nC_{OX}U_T}\right]\right) \Longrightarrow -\left[\frac{Q_S}{2nC_{OX}U_T}\right]^{SAT} = \frac{\sqrt{1+4 \cdot IF} - 1}{2}$$

The g_{ms}/I_D Invariant.

- In saturation, the source transconductance-to-current ration is only dependent on the inversion factor *IF*, *and not on the device parameters*:

$$\frac{g_{ms}}{I_{DSat}} \cdot U_T = n \frac{g_m}{I_{DSat}} \cdot U_T = \frac{1}{\frac{1}{2} + \sqrt{\frac{1}{4} + IF}}$$

The definition of I_{SP} (the current normalization factor) is consistent with the transition from weak to strong inversion, when the MOSFET operates in saturation.

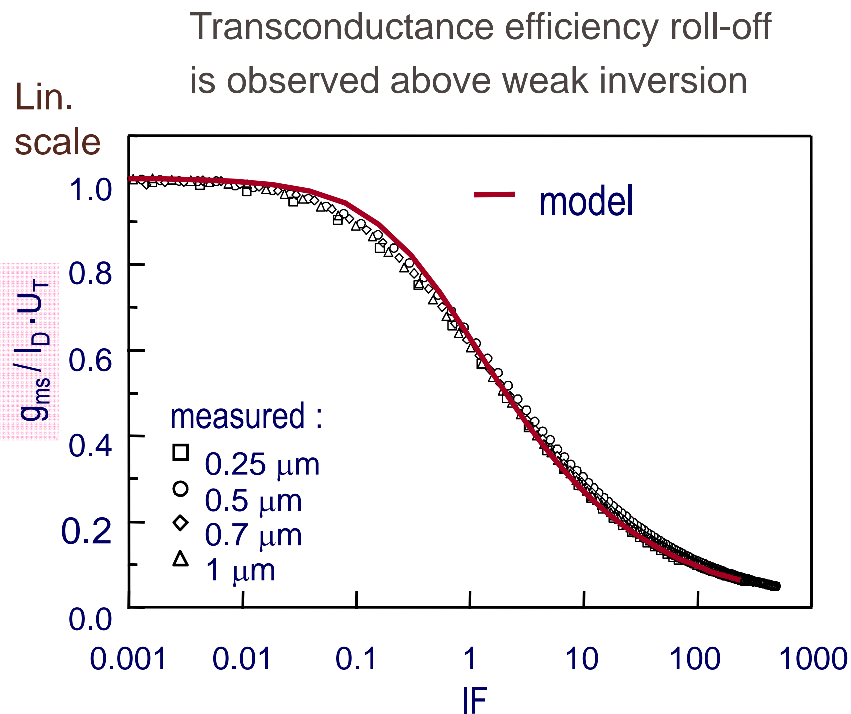


The g_{ms}/I_D Invariant.

- Measured g_{ms}/I_D in 'long' channels.
  - *Agreement between the model derived from charge linearization and drift-diffusion transport with constant mobility seems accurate for g_{ms}/I_D and g_m/I_D characteristics.*

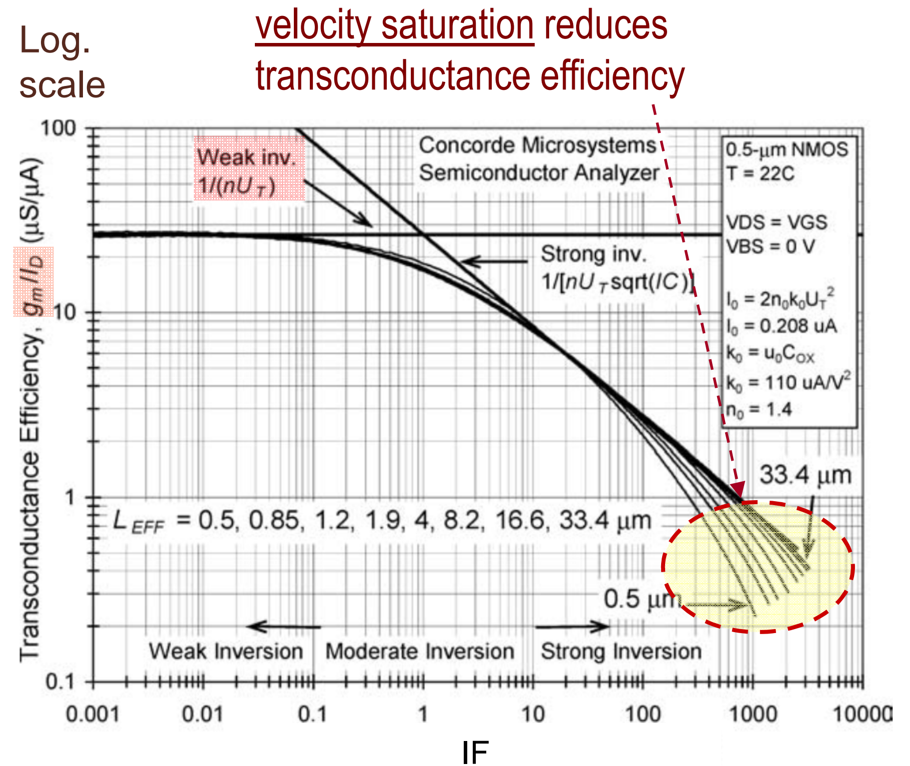
Transconductance efficiency roll-off is observed above weak inversion

Lin. scale



Log. scale

velocity saturation reduces transconductance efficiency



Impact of Velocity Saturation on g_m/I_D

- Assuming strong inversion and saturation, the drain current is given by:

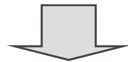
$$I_D \overset{SI,SAT}{\approx} \frac{\beta}{2n} V_{GTO}^2 \Rightarrow g_{m.} \overset{SI,SAT}{\approx} \sqrt{2\frac{\beta}{n} I_D} \Rightarrow \frac{g_m}{I_D} \overset{SI,SAT}{\approx} \frac{1}{nU_T} \cdot \frac{1}{\sqrt{IF}}$$

- At high V_D, velocity saturation will limit the current:

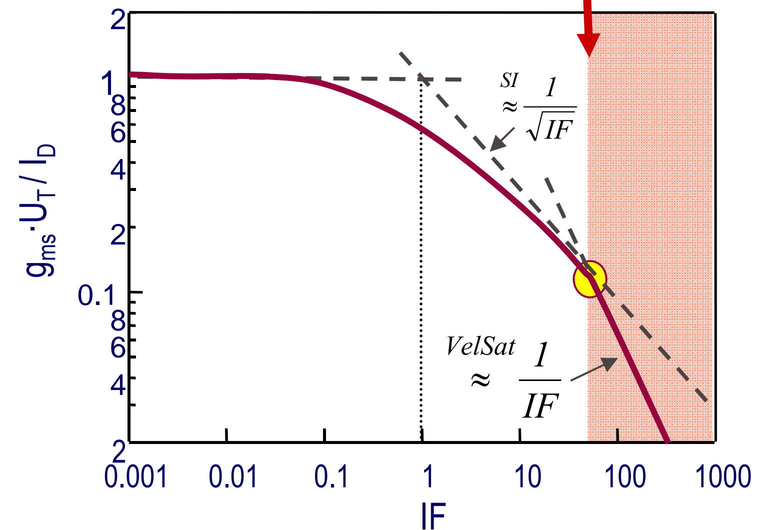
$$I_D \overset{VSat}{\approx} W \cdot C_{ox} \cdot (V_{GO} - V_{Dsat} - 2\phi_F) \cdot V_{sat} \Rightarrow g_m \overset{VSat}{\approx} WC_{ox}V_{sat} \Rightarrow \frac{g_m}{I_D} \overset{SI,SAT}{\approx} \frac{V_{Sat}WC_{OX}}{I_{SP}} \cdot \frac{1}{IF}$$

$$IF_{crit} \approx (V_{Sat}L)^2 / (\mu U_T)^2$$

- If $V_{sat}=10^5\ ms^{-1}$ & $\mu=0.05\ m^2V^{-1}s^{-1}$



$IF_{crit}(L=1\mu m) \sim 64000$
$IF_{crit}(L=0.1\mu m) \sim 64$



Velocity saturation deteriorates g_{ms}/I_D characteristics.

Intrinsic Voltage Gain

- The intrinsic voltage gain is the product of transconductance efficiency and Early voltage:

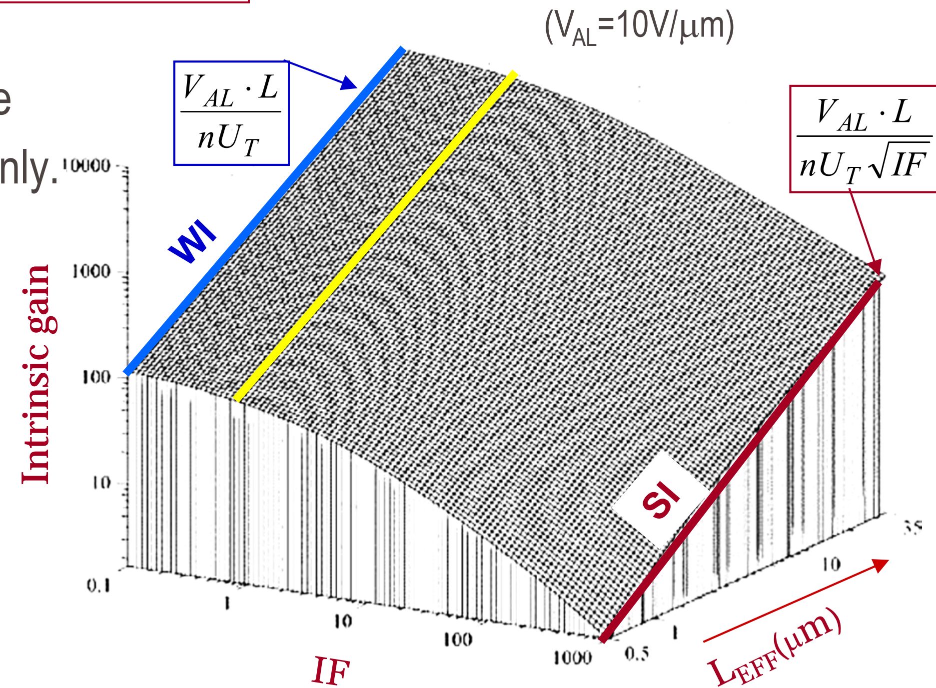
$$A_V = \frac{g_m}{g_{ds}} = \frac{g_m}{I_D} \cdot \frac{I_D}{g_{ds}} = \left[\frac{g_m}{I_D}\right] \cdot V_{Early} \overset{SAT}{=} \frac{L_G}{\frac{1}{2} + \sqrt{\frac{1}{4} + IF}} \cdot \frac{V_{AL}}{nU_T}$$

Early voltage normalized to 1 µm.(tech. dependent)

– In saturation, A_V should depend on the Inversion Factor and channel length only.

Gain is maximum in weak inversion for long channel devices and minimum in strong inversion for short channel devices.

There is a large design range available in analog MOS design.



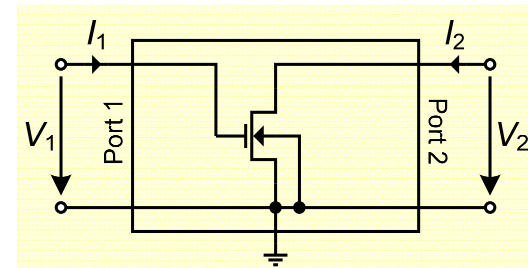
D. M. Binkley et-al Analog Integrated Circuits and Signal Processing, 47, 137–163, 2006

Transition Frequency

- The transit frequency f_t is simply controlled by the transconductance and gate capacitance. It represents the 'intrinsic' gate speed (free from junction capacitances) and is related to the current gain (H_{21}):
  - f_t is defined as the frequency for which the current gain of the two-port shown below is equal to unity

not negligible in short channel devices !

$$h_{21} \equiv \left.\frac{I_2}{I_1}\right|_{V_2=0} = \frac{Y_{21}}{Y_{11}} \approx \frac{g_m - j\omega C_{GD}}{j\omega C_{GG}} \cong \frac{g_m}{j\omega C_{GG}} = \frac{\omega_t}{j \cdot \omega}$$



$$\omega_t = \frac{g_m}{C_{GG}} = \frac{g_m}{C_{GGi} + C_{Go}}$$

Layout dependent (overlap)

In strong inversion and saturation $C_{GSi} \approx 2/3 C_{ox}$ and $C_{GDi} \approx 0$:

$$\omega_t = \frac{g_m}{C_G} \approx 3 \cdot \frac{\mu}{L^2} \cdot U_T \sqrt{IF} \approx \frac{3}{\tau_D} \sqrt{IF}$$

Ignoring extrinsic capacitances

It follows that to first order, the unity gain frequency of transistors depends only on effective gate-overdrive voltage and on channel length.

Transition Frequency

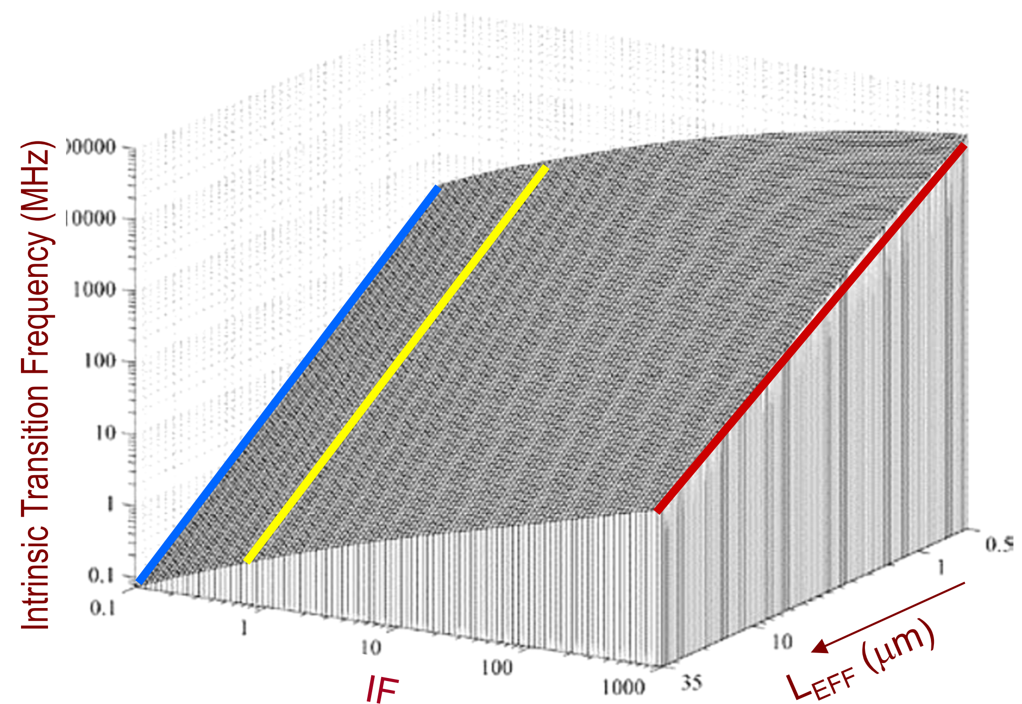
- NMOS intrinsic TF as a function of inversion factor and channel length

Ideally, transition frequency increases linearly with IF in weak inversion and with √IF in strong inversion.

For a given IF, f_T decreases as $1/L^2$ (1/L if velocity saturation) with increasing channel length for all regions of operation

This is where intrinsic voltage gain is maximized and flicker noise minimized.

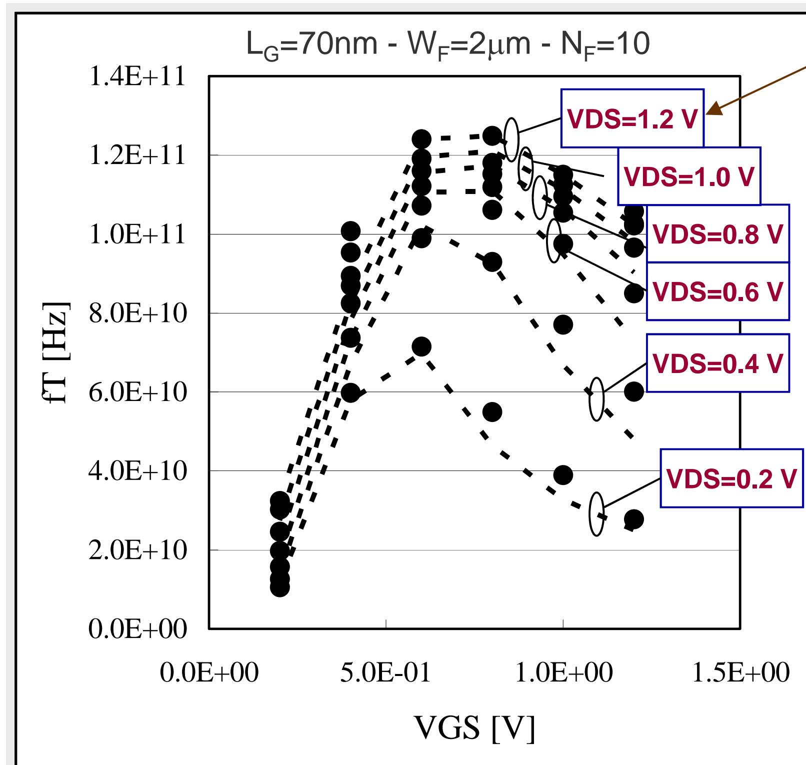
Extrinsic overlap & junction capacitances will reduce operating bandwidth (short-channel devices).



D. M. Binkley et-al Analog Integrated Circuits and Signal Processing, 47, 137–163, 2006

Transition Frequency

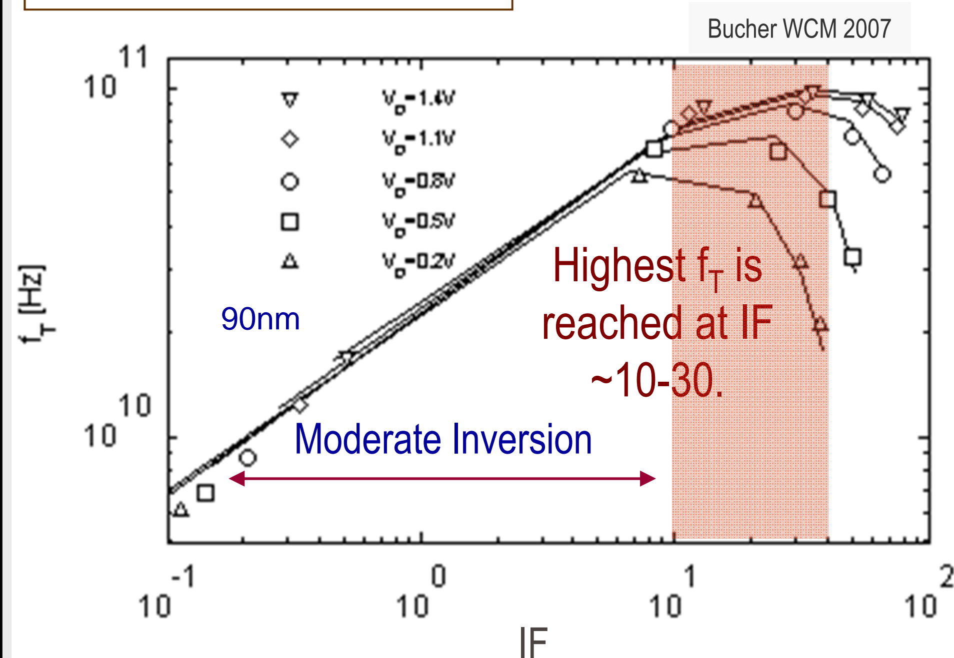
- ...but at high inversion levels, short channel degrades the transconductance efficiency and reduces the increase in bandwidth: the optimum is located close to the maximum in g_m (almost independent of V_{DS}).



Yoshitomi S. MIXDES 2007

Increasing V_{DS} also increases f_T (increase in g_m)

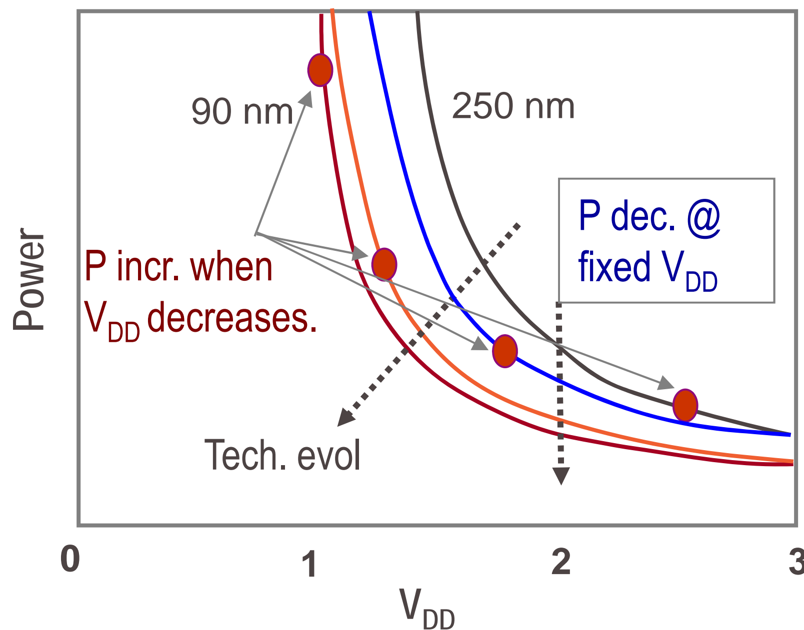
Bucher WCM 2007



Moderate Inversion is then a good trade-off (less µ reduction)

Example: Power Scaling from the g_m/I Invariant

- Power consumption of analog circuits is proportional to the signal-to-noise ratio and to the signal frequency.
  - For analog circuits more performance requires higher power consumption.
- For a given power, the SNR performance drops when moving to newer technologies, simply because of reduced maximum signal swing due to lower supply voltage.



Example of minimum power consumption for simple unity gain voltage buffer.

If the supply voltage could be maintained constant, new technologies would indeed tend to lower the power consumption.

…but this is not true if the supply voltage has also to be scaled down.

Annema A.J. et. al, JSSC vol.. 40, n.1, p.132 (2005)

Example: Power Scaling from the g_m/I Invariant

- How SNR evolves assuming constant power consumption ?
  - *We assume the same scaling factor α for the technology and for the supply voltage.*

Scaling α (<1)

- Supply voltage V_{DD} $\quad V_{DD} \longrightarrow V'_{DD} = \alpha V_{DD}$
- (Large) signal amplitude V_{PP} $\quad V_{PP} \longrightarrow V'_{PP} = \alpha V_{PP}$
- DC bias current I_D $\quad I_D \longrightarrow I'_D = I_D/\alpha$ (cst power)

If we maintain the same gate overdrive voltage, in saturation, IF will be unchanged, and so for g_m/I_D $\Longrightarrow$ $S_{\delta V_G} \div 1/g_m \div 1/I_D$

- Input noise signal $\sim V^2_{noise}$ $\quad V^2_{noise} \overset{cst.IF}{\div} 1/I_D \longrightarrow V'^2_{noise} \overset{cst.IF}{\div} \alpha/I_D = \alpha V^2_{noise}$
- Signal to noise ratio $\quad \frac{Signal}{Noise} \div \frac{V^2_{PP}}{V^2_{noise}} \longrightarrow \left.\frac{Signal'}{Noise'}\right|_{scaled} \div \frac{V'^2_{PP}}{V'^2_{noise}} = \alpha \frac{Signal}{Noise}$

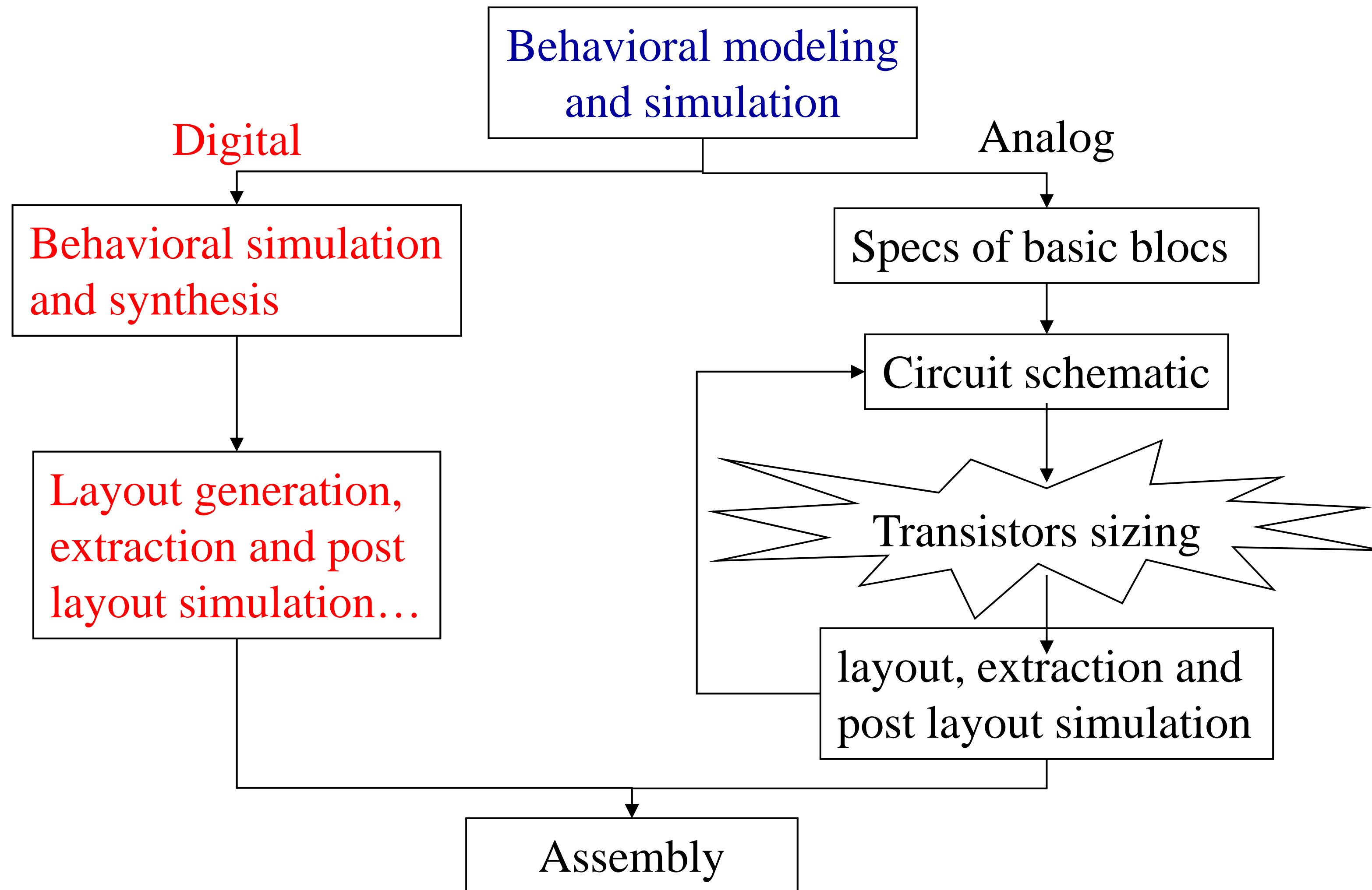
- SNR is degraded upon scaling assuming constant power, when large signals are considered.
- However, if V_{PP} doesn't need to be scaled (small signal), SNR will increase by 1/α upon scaling

Inversion factor based design methodology

Acknowledgments

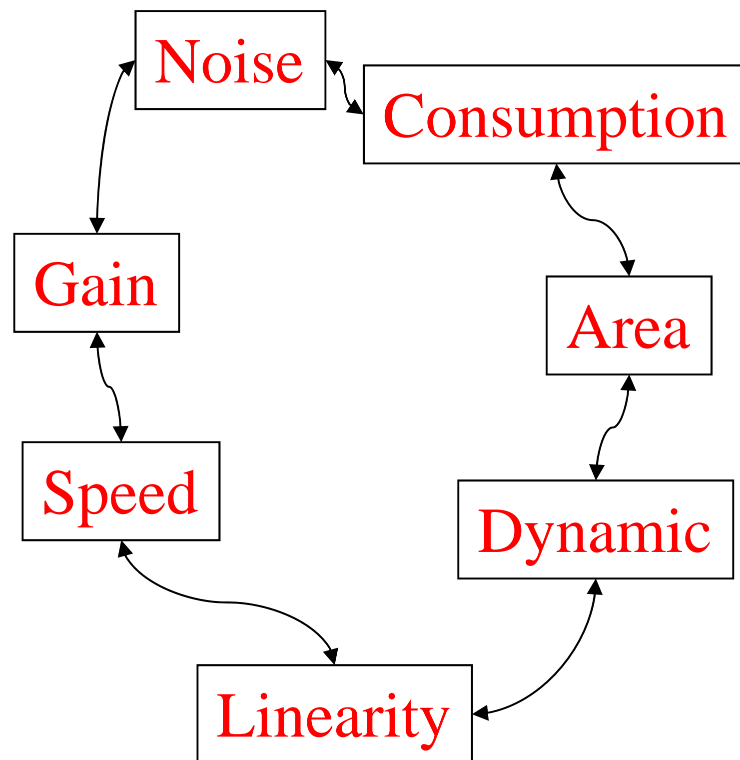
– M. Kayal, EPFL

Top Down System Design

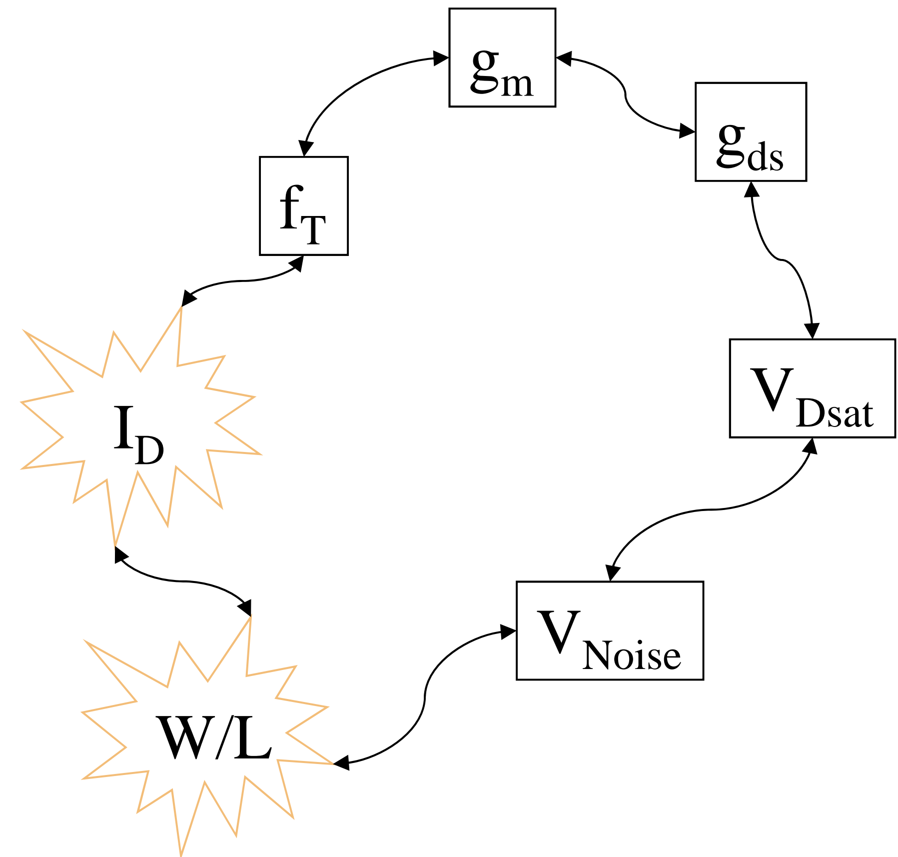


Circuits Parameters / Transistors Parameters

Example of circuit level parameters



Example of transistor level parameters

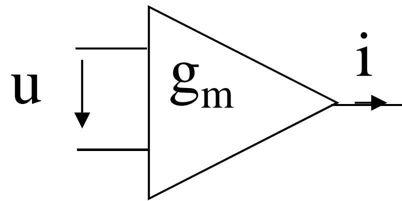


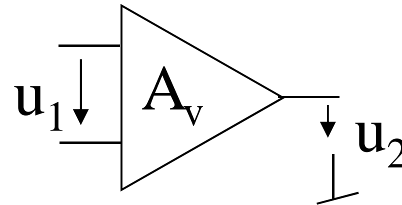
Main Questions

HOW TO:

- Deal with increasing circuit **complexity** ?
- Deal with very **demanding specifications** sets ?
- Estimate **technology limits** ?
- Bridge the gap between **hand-calculations** and simulations ?
- Use the **device physics understanding** for analog design ?
- Optimize analog circuits and find the best **trade-offs** ?
- Using design flow tools for **analog design assistance** ?

Basic Analog Cells



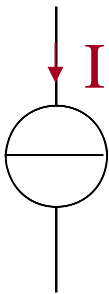


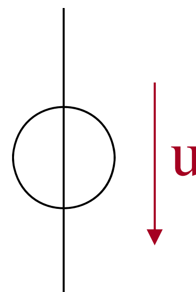
OTA

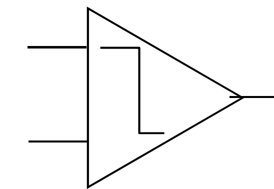
$R_{out} = \infty$

OP-Amp

$R_{out} = 0$







Comparator

Current Source

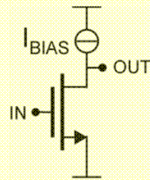
$R_{out} = \infty$

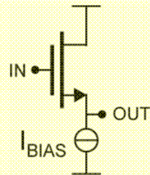
Voltage Source

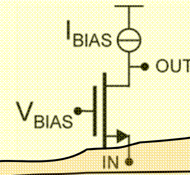
Rout= 0

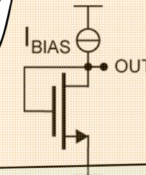
Basic analog structures

Common Source Common Drain Common Gate Diode



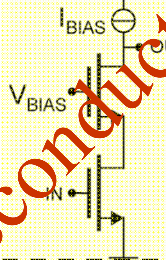


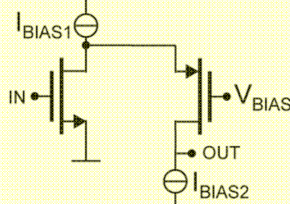


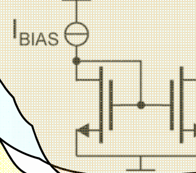


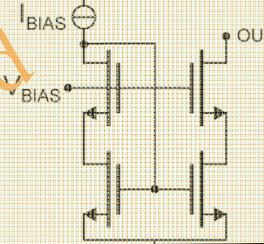
Bias

CS CS Current Mirror Current Mirror





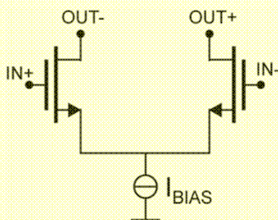


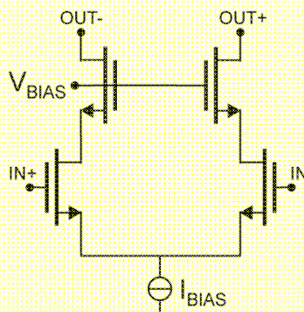


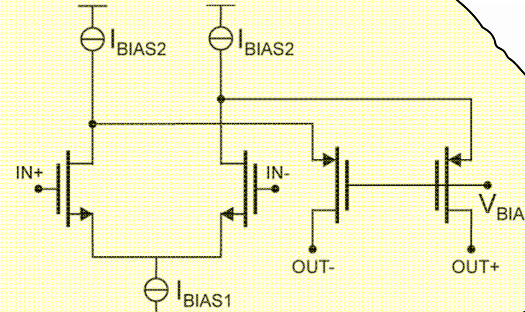
Transconductances

Load

Diff Pair Diff Pair Diff Pair







Observations

- MOS Device behavior depends strongly on the inversion level of the device.
- Sizing of geometrical dimensions W and L depends, therefore, on the inversion level.
- Design parameters are well controlled when the device is in very strong or very weak inversion.

MOS model dedicated to analog design:

- **Small number of parameters** with physical meaning
- Model equations **approximations** without a great loss of accuracy
- Accurate modeling of **weak and moderate inversion** behavior
- **Continuous expressions** of current derivatives

Inversion Factor

- Inversion Factor I_F is a normalized value that describes the inversion level

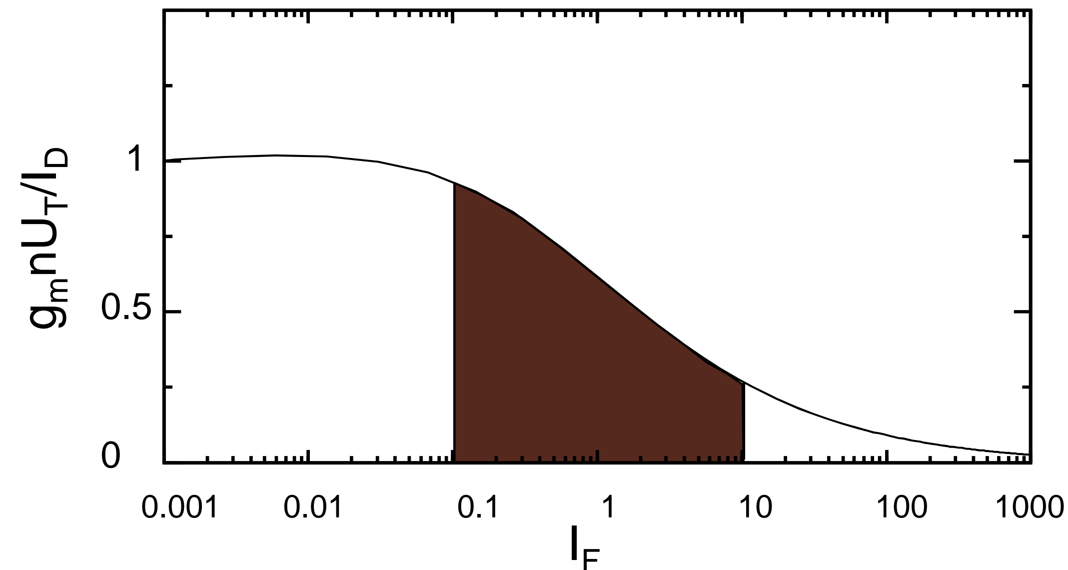
$$I_F = \frac{I_D}{2n\beta U_T^2} = \frac{I_D}{2n\mu C_{ox}\frac{W}{L}U_T^2} = \frac{I_D}{2nK_p\frac{W}{L}U_T^2} = \frac{I_D}{I_S}$$

- Strong inversion $I_F > 10$
- Weak inversion $I_F < 0.1$
- Moderate Inversion $0.1 \leq I_F \leq 10$

$$\frac{g_m}{I_D} = \frac{1}{nU_T} \cdot \frac{1}{\frac{1}{2} + \sqrt{\frac{1}{4} + I_F}}$$

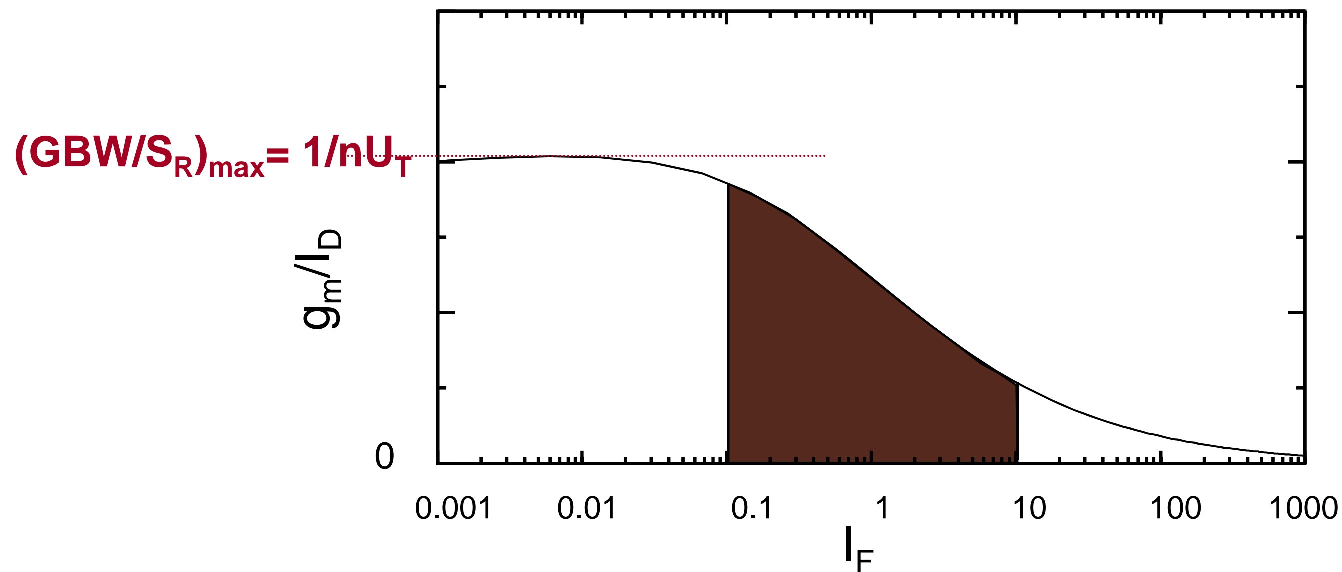
The Normalized g_m/I_D Characteristic

- The choice of g_m/I_D ratio translates current into transconductance (g_m efficiency and consumption efficiency).
- It gives an indication of the device operating region (strong, moderate or weak).
- The greater the g_m/I_D value, the greater the transconductance we obtain at a constant current value.
- The normalized g_m/I_D is almost invariant to process parameters and scaling



Maximum of g_m/I_D

$$\frac{g_m}{I_D} = \frac{1}{nU_T} \cdot \frac{1}{\frac{1}{2} + \sqrt{\frac{1}{4} + I_F}} \leq 29.5\,(n = 1.3)$$



Transistor level design

Design parameters	**Design variables**
saturation voltage V_{DSsat}	saturation current I_{Dsat}
transconductance g_m	inversion factor I_F
output conductance g_{DS}	transistor width W
parasitic capacitances	transistor length L
intrinsic gain A_i	ratio W/L
transition frequency f_t	area WL
equivalent noise	

?

Modeling versus analog parameters

$$V_{DSsat} = V_t \cdot (2\sqrt{IF} + 4)$$

$$\frac{g_m}{I_{Dsat}} = \frac{1}{nV_t} \cdot \frac{1}{\frac{1}{2} + \sqrt{IF + \frac{1}{4}}}$$

$$g_{DS} = \frac{I_{Dsat}}{L \cdot V_a}$$

$$A_i = \left(\frac{g_m}{I_{Dsat}}\right) \cdot L \cdot V_a$$

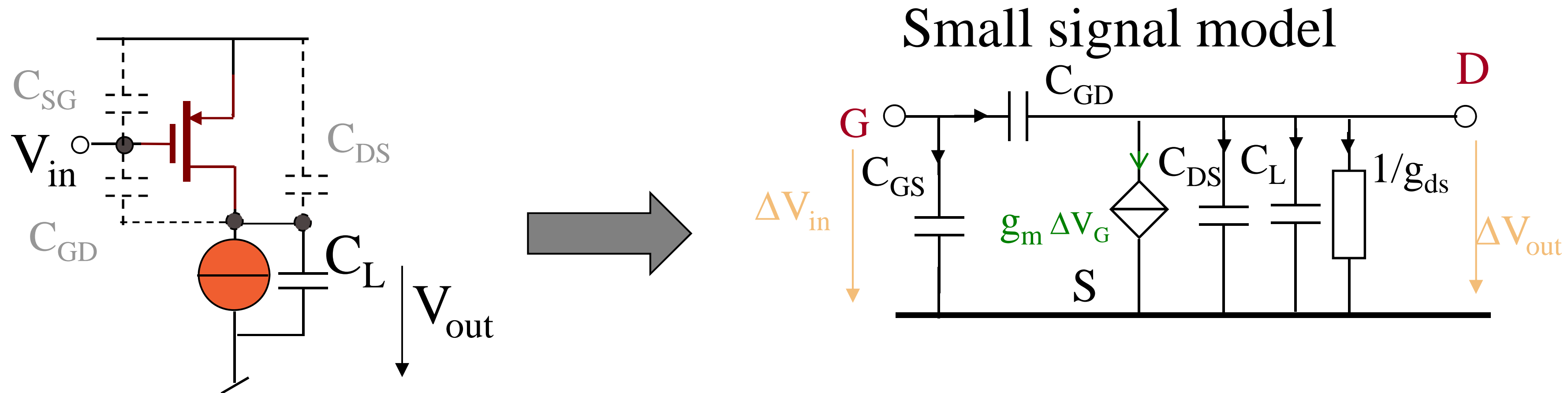
$$(\Sigma C_i)_{MAX} = C_{ox} \cdot \frac{I_{Dsat}}{2nKPV_t^2 \cdot IF} \cdot L^2$$

$$f_t = \frac{g_m}{2\pi(\Sigma C_i)_{MAX}} = \left(\frac{g_m}{I_{Dsat}}\right) \cdot \frac{nKPV_t^2 \cdot IF}{\pi C_{ox} \cdot L^2}$$

$$v_{n,th}^2 = \frac{4KT \cdot n \cdot \frac{1}{1+IF} \cdot \left(\frac{1}{2} + \frac{2}{3}IF\right)}{\left(\frac{g_m}{I_{Dsat}}\right) \cdot I_{Dsat}}$$

$$v_{n,fl}^2 = \frac{2nKPV_t^2 \cdot KF \cdot IF}{L^2 C_{ox} f^{AF} \cdot I_{Dsat}}$$

Common Source Design



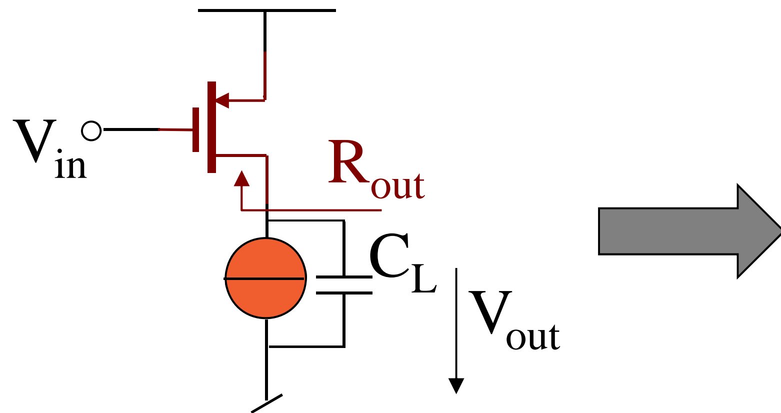
$$(\Delta V_{in} - \Delta V_{out})\, j\omega C_{GD} = g_m \Delta V_{in} + \Delta V_{out}\, g_{ds} + \Delta V_{out}\, j\omega C_L$$

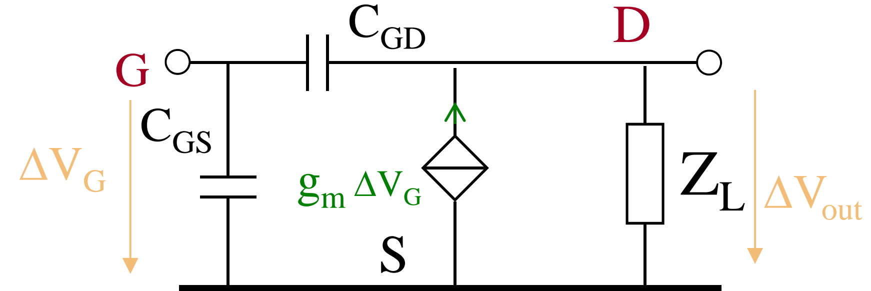
$$\Delta V_{in}\,(j\omega C_{GD} - g_m) = \Delta V_{out}\,(j\omega C_{GD} + j\omega C_L + g_{ds})$$

$$\frac{\Delta V_{out}}{\Delta V_{in}} = -\frac{g_m}{g_{ds}} \frac{1 - j\omega C_{GD}/g_m}{1 + j\omega (C_L + C_{GD})/g_{ds}}$$

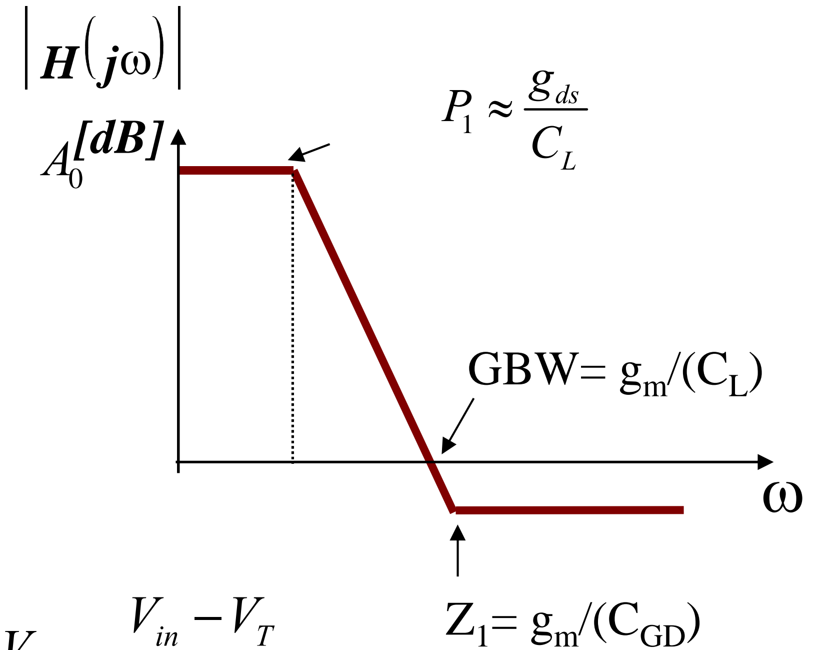
A_0 — Zero — Pole

Common Source (1/2)





$$\frac{\Delta V_{out}}{\Delta V_{in}} = -\frac{g_m}{g_{ds}} \frac{1 - j\omega C_{GD} / g_m}{1 + j\omega (C_L + C_{GD}) / g_{ds}}$$

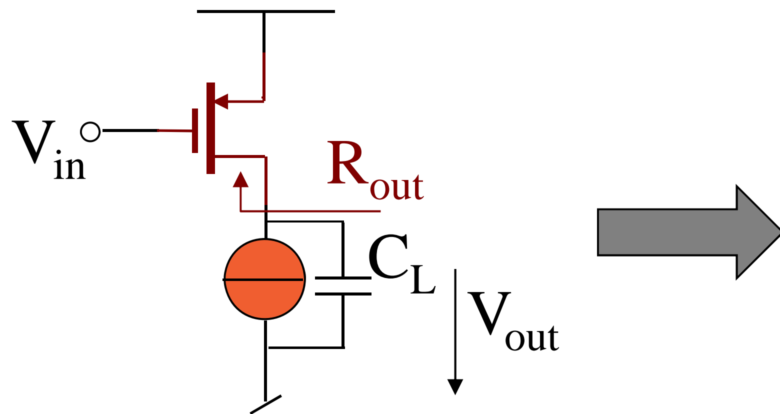


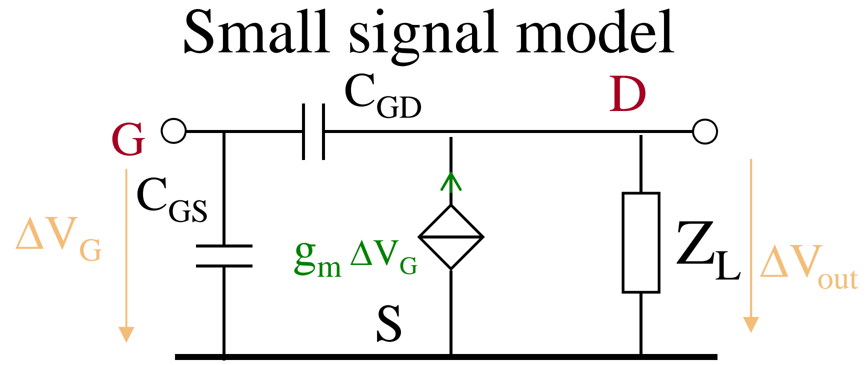
$$A_0 = -\frac{g_m}{g_{ds}} \text{ (Intrinsic gain)}$$

$$GBW = \frac{g_m}{C_L}$$

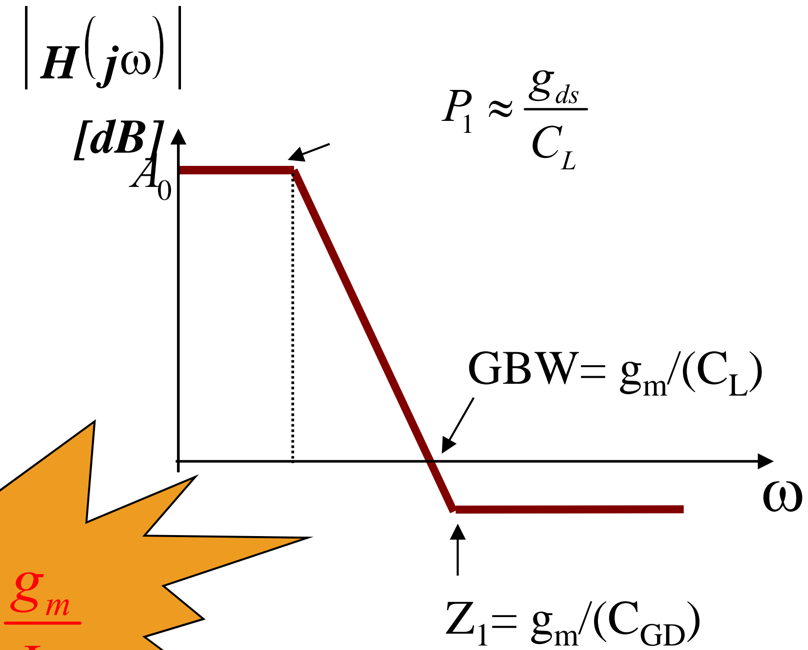
$$S_R = \frac{I_D}{C_L} \qquad V_{out,\max} = V_{DD} - V_{Dsat} = V_{DD} - \frac{V_{in} - V_T}{n}$$

Common Source (2/2)





$$\frac{\Delta V_{out}}{\Delta V_{in}} = -\frac{g_m}{g_{ds}} \frac{1 - j\omega C_{GD}/g_m}{1 + j\omega (C_L + C_{GD})/g_{ds}}$$



$$A_0 = -\frac{g_m}{g_{ds}} = \frac{g_m}{I_D}.U_a.L \text{ (Intrinsic gain)}$$

$$GBW = \frac{g_m}{C_L} = \frac{g_m}{I_D} \cdot \frac{I_D}{C_L}$$

$$S_R = \frac{I_D}{C_L}$$

$$\frac{GBW}{S_R} = \frac{g_m}{I_D}$$

Exercise: Transistor Design

- Electrical specifications:
  - $GBW = 2\pi f_T \geq 2.\ \pi.10$Mhz
  - $A_0 \geq 40$ dB
  - $SR \geq 4.10^6$ V/s
  - $C_L = 1$pF
- Technology Data:
  - Early Voltage, $U_{a,P} = 7$ V/ μm
  - $I_{S,P} = 0.18\ \mu A$ ($n = 1.3$ & $K_P = 100\ \mu A/\ V^2$)

Structured Transistor Sizing

$$\frac{g_m}{I_D} = \frac{2\pi f_T}{S_R} = \frac{2.\pi.10\,MHz}{4.10^6 V/s} \approx 15.7$$

$$\frac{g_m}{I_D} = \frac{1}{nU_T} \cdot \frac{1}{\frac{1}{2} + \sqrt{\frac{1}{4} + I_F}} \rightarrow I_F = 1.6$$

$$A_0 = -\frac{g_m}{g_{ds}} = \frac{g_m}{I_D} . U_a . L \rightarrow L \geq 0.8\mu m$$

$$S_R \,\&\, C_L \rightarrow I_D = 4\mu A$$

$$I_F = \frac{I_D}{2nK_p \frac{W}{L} U_T^2} \rightarrow W = \frac{(L=1).I_D}{I_F 2nK_p U_T^2} \approx 14\mu m$$

$$\frac{W}{L} \approx \frac{14\,\mu m}{1\,\mu m}$$

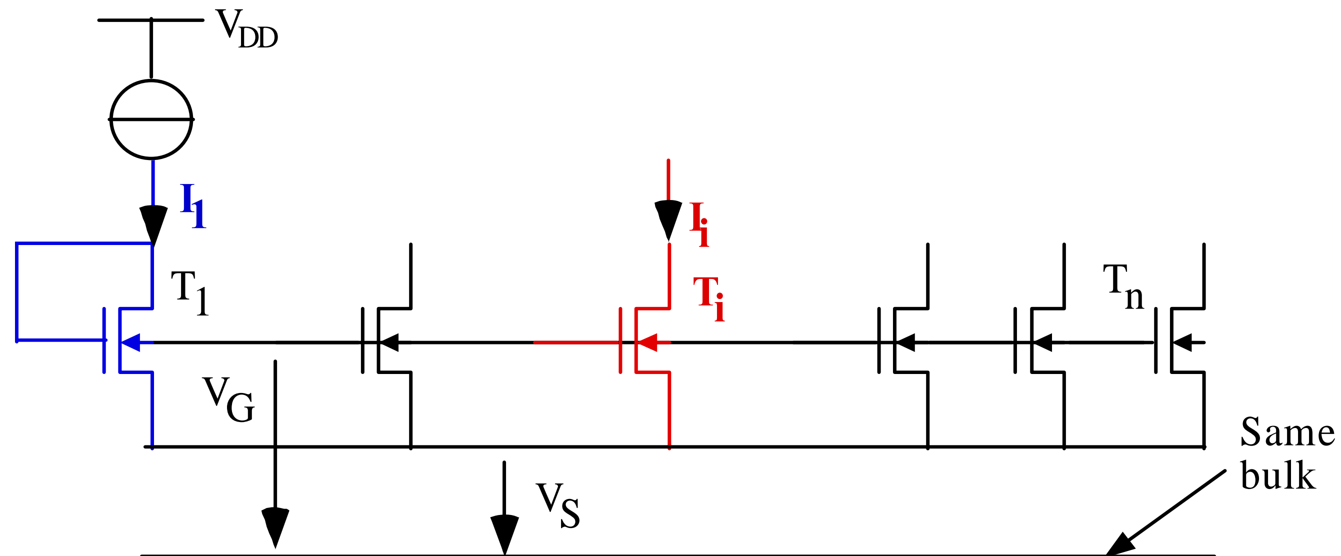
Hand Calculation of Transistor Using inversion factor

$\frac{g_m}{I_D}$	$\frac{1}{nU_T} \cdot \frac{1}{\frac{1}{2}+\sqrt{\frac{1}{4}+I_F}}$ With $I_F = \frac{I_D}{2n\mu C_{ox}\frac{W}{L}U_T^2}$
$\frac{GBW}{S_R}$	$\frac{g_m}{I_D}$ (maximum is 29.5)
Saturation Voltage	$V_{Dsat} = U_T(2\sqrt{I_F}+4)$, minimum 100 mV
DC Gain	$A_0 = \frac{g_m}{g_{ds}} = \frac{g_m}{I_D}.U_a.L = \frac{GBW}{S_R}.U_a.L$ (U_a is Early Voltage)
Thermal Noise	$f(\frac{1}{g_m})$
Flicker Noise	$f(\frac{1}{W.L})$
Mismatch	$\frac{\Delta I_D}{I_D} = \frac{\Delta\beta}{\beta} - \frac{g_m}{I_D}\Delta V_{T0}$ $\Delta V_G = \Delta V_{T0} - \frac{I_D}{g_m}\frac{\Delta\beta}{\beta}$

I_F & L

Current Mirror mismatch and sizing

Multiple Current Mirror



$I_i/I_1 = \beta i/\beta_1.$

$$With\ \beta = k_P \frac{W}{L} \ \&\ k_P = \mu Cox$$

Mismatch Effect on Current

$$I_D = \frac{\beta}{2n}(V_G - V_{T0})^2$$

$$\Delta I_D = \frac{\partial I_D}{\partial \beta}\Delta\beta + \frac{\partial I_D}{\partial (V_G - V_{T0})}\Delta(V_G - V_{T0}) = \frac{1}{2n}(V_G - V_{T0})^2 \Delta\beta - g_m \Delta V_{T0}$$

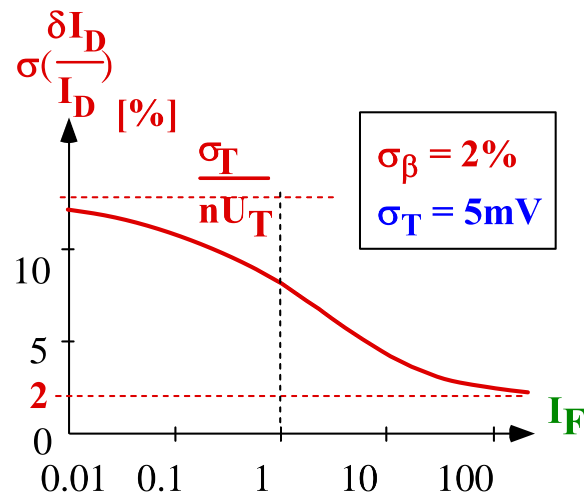
$$\frac{\Delta I_D}{I_D} = \frac{\Delta\beta}{\beta} - \frac{g_m}{I_D}\Delta V_{T0}$$

Matching of Currents as a Function of g_m/I_D

$$\frac{\Delta I_D}{I_D} = \frac{\Delta\beta}{\beta} - \frac{g_m}{I_D}\Delta V_{T0} \qquad \sigma\left(\frac{\Delta I_D}{I_D}\right) = \sqrt{\sigma_\beta^2 + \left(\frac{g_m}{I_D}\sigma_T\right)^2}$$

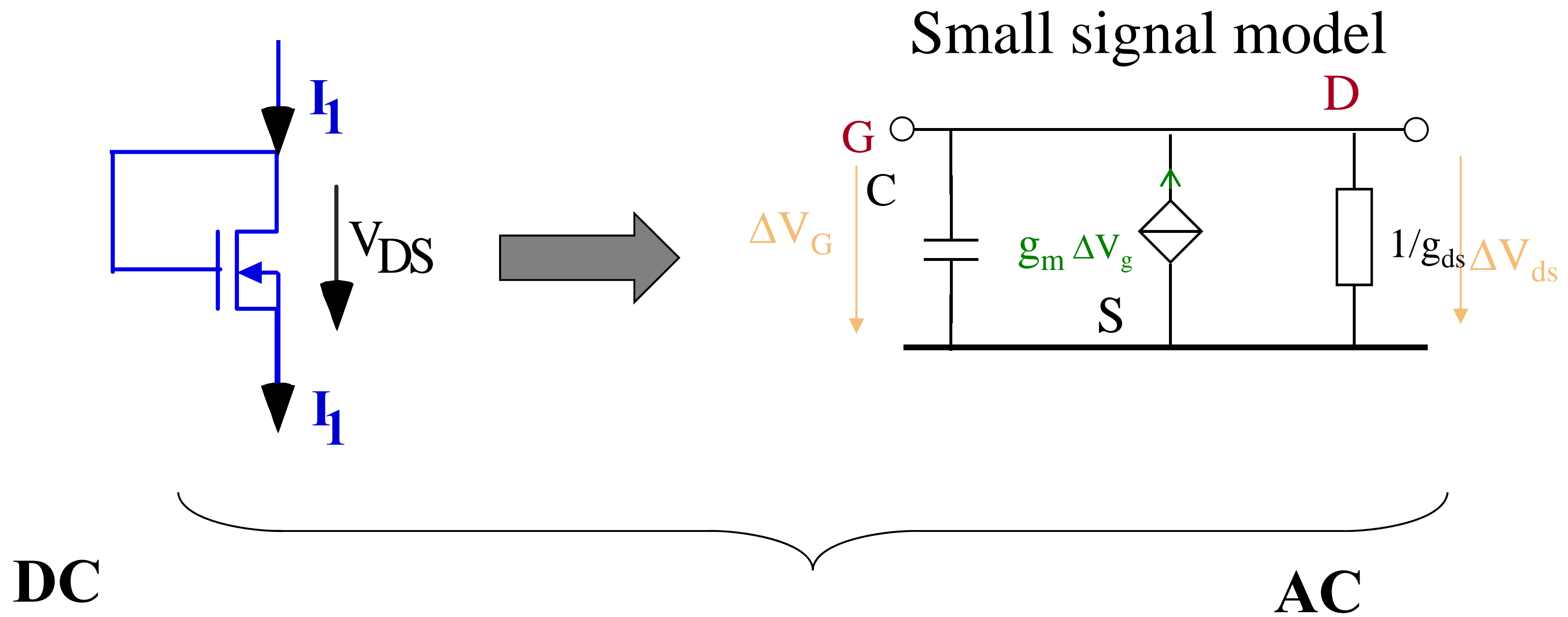
Techno: $A_\beta = \sigma_\beta \sqrt{W.L}$ et $A_T = \sigma_T \sqrt{W.L}$

Example (with $nU_T = 40mV$):



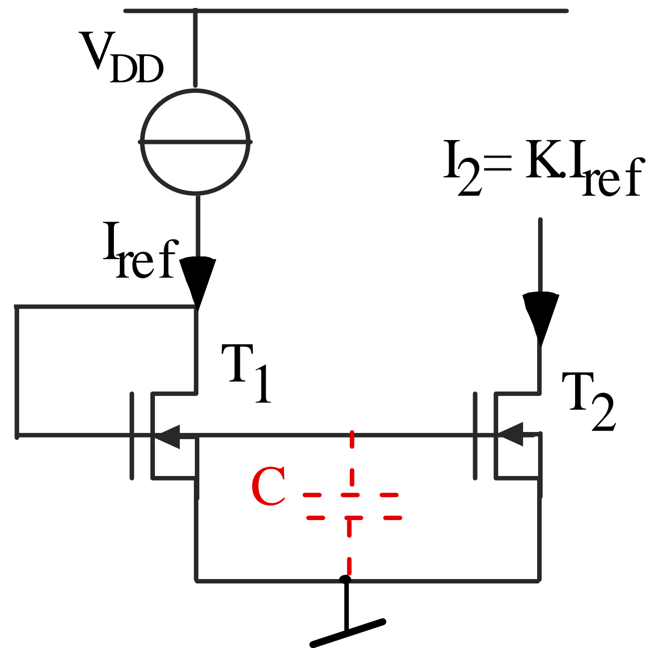
- Bad precision of currents in weak inv.

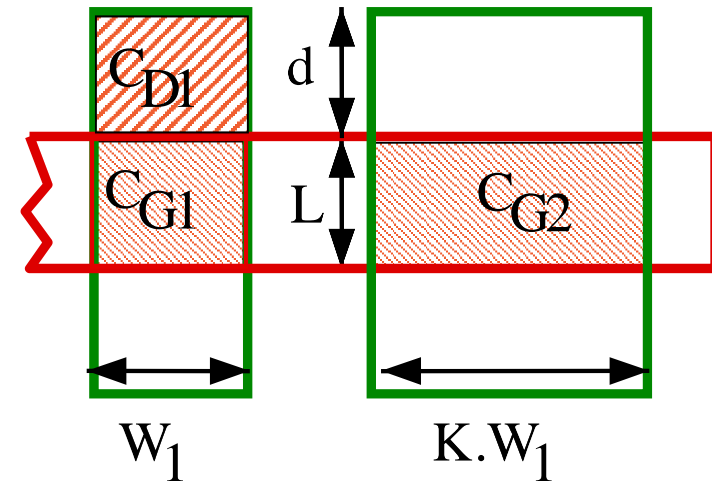
Diode Connected Transistor



$$R = \frac{\Delta V_{DS}}{\Delta I_1} = 1/g_m$$

Dynamic Behavior





$C = C_{G1} + C_{G2} + C_{D1}$

$C_G = W.L.C_{ox}$ in strong inv.

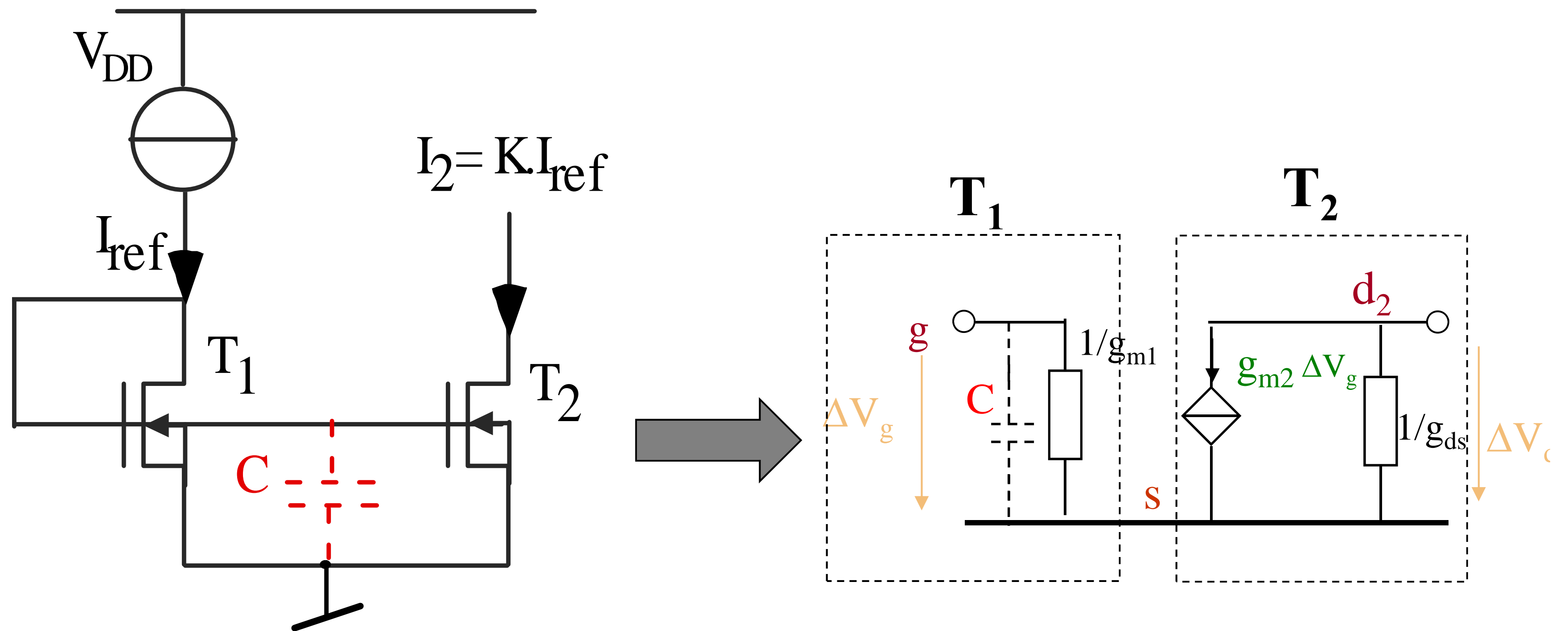
$C_G = W.L.C_{ox.}(1\text{-}1/n)$ in weak inv.

$C_D = C_J.W.d$

C_{OX} ≈ 3.4 fF/µm² - tech : 0.5 µm
≈ 8.5 fF/µm² - tech : 0.18 µm
≈ 14 fF/µm² - tech : 0.09 µm

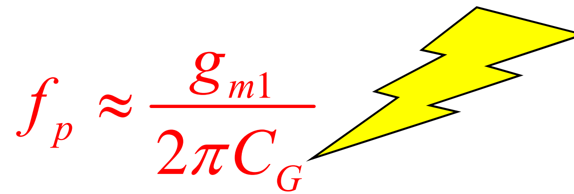
C_j ≈ 0.74 fF/µm²

Small Signal Model



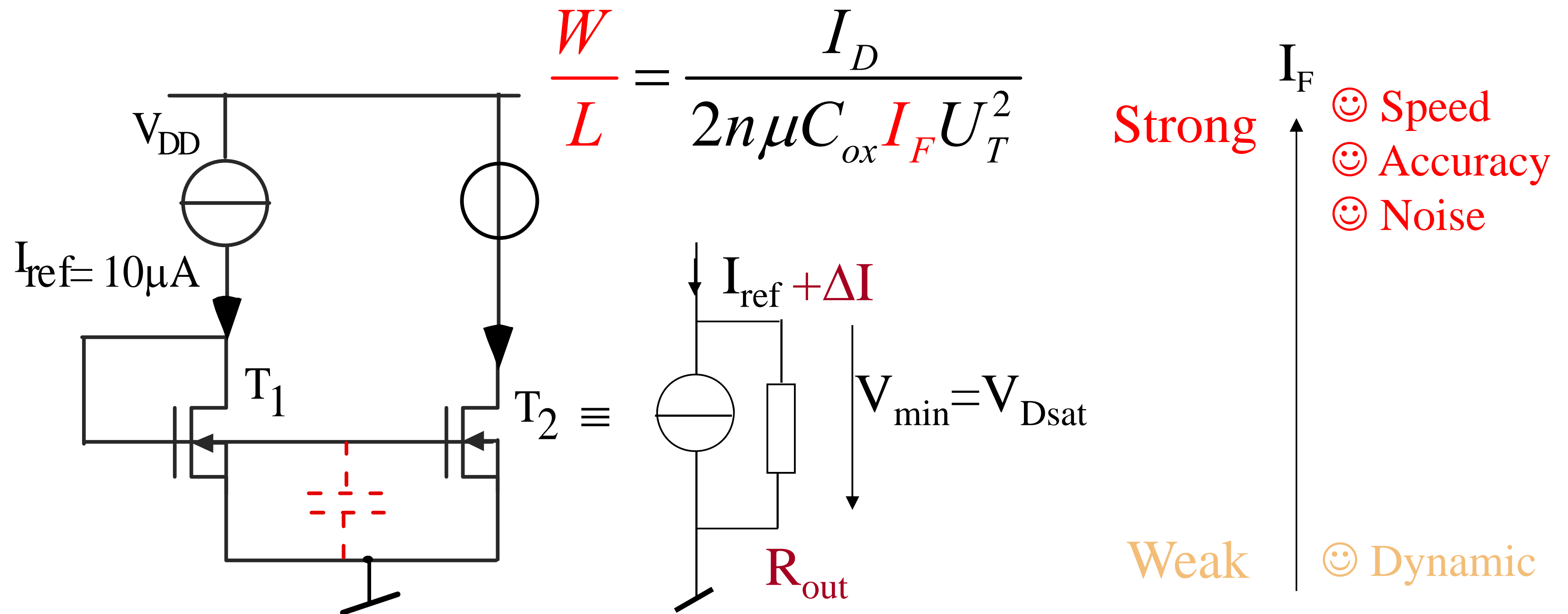
$$\bar{i_2} = \frac{\bar{i_1}}{1 + j\frac{\omega}{\omega_p}}$$

$$f_p \approx \frac{g_{m1}}{2\pi C_G}$$



$$f_p \approx \frac{g_{m1}}{2\pi K.W.L.C_{ox}}$$

Current Mirrors: Trade-off



$$\sigma(\frac{\Delta I_D}{I_D}) = \sqrt{\sigma_\beta^2 + (\frac{g_m}{I_D}\sigma_T)^2}$$

$$R_{out} = \frac{1}{g_{ds}} = \frac{U_A}{I_D} = \frac{U_a.L}{I_D}$$

$$V_{Dsat} = U_T(2\sqrt{I_F}+4)$$

$$\omega_p \approx f(\frac{g_m}{W.L})!!!!$$

Hand Calculation of Current Mirror Using inversion factor

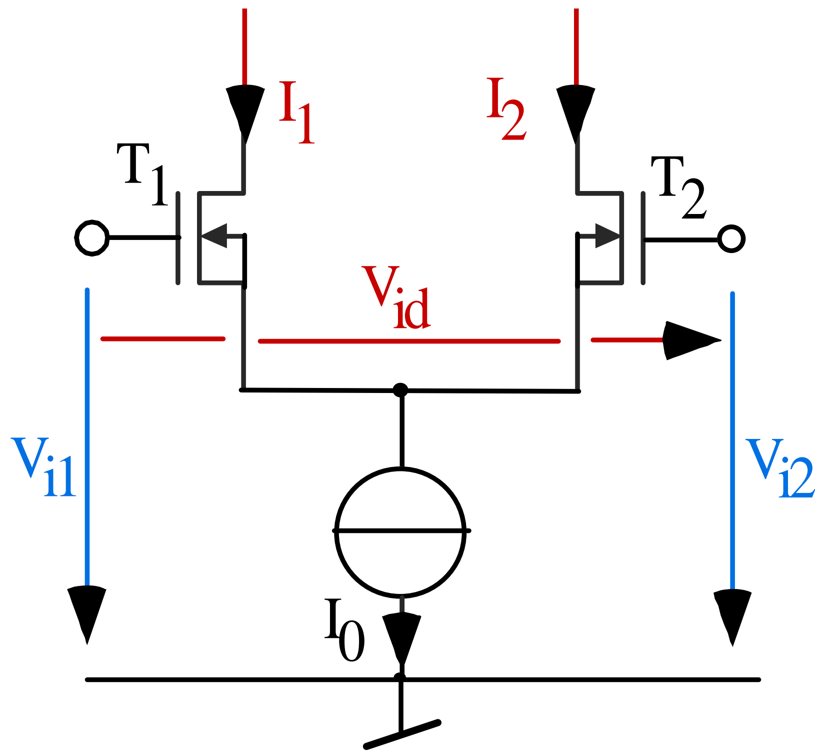
$\frac{g_m}{I_D}$	$\frac{1}{nU_T} \cdot \frac{1}{\frac{1}{2}+\sqrt{\frac{1}{4}+I_F}}$	$With\ \ I_F = \frac{I_D}{2n\mu C_{ox}\frac{W}{L}U_T^2}$
Matching	$\sigma(\frac{\Delta I_D}{I_D}) = \sqrt{\sigma_\beta^2 + (\frac{g_m}{I_D}\sigma_T)^2}$	☺ Strong inversion
Saturation Voltage	$V_{Dsat} = U_T(2\sqrt{I_F}+4)$	☺ Weak inversion
Output Resistance	$R_{out} = \frac{1}{g_{ds}} = \frac{U_A}{I_D} = \frac{U_a.L}{I_D}$!!!	
Parasitical pole	$\omega_p \approx f(\frac{g_m}{W.L})$	☹

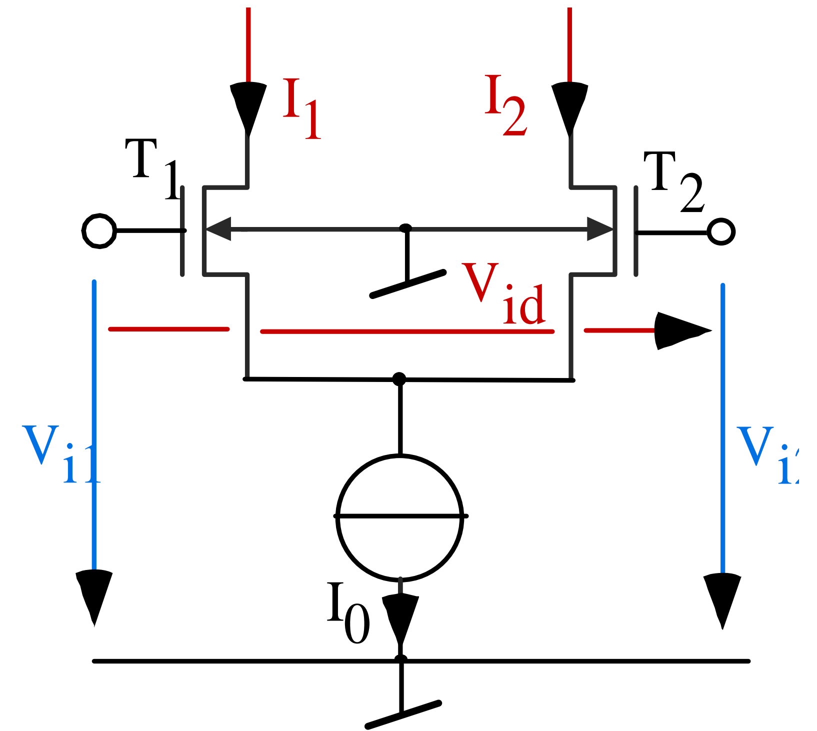
ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Electronics Laboratory ELab

Differential Pair mismatch and sizing

Differential Pair



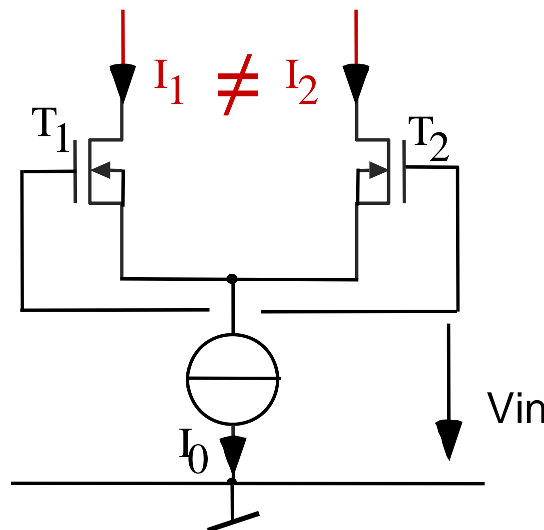


Special well connected to sources

✓ **$\Delta V_s = 0$ -> highest common mode**

Substrate connected to Vss (P-subs.)

DC-Offset

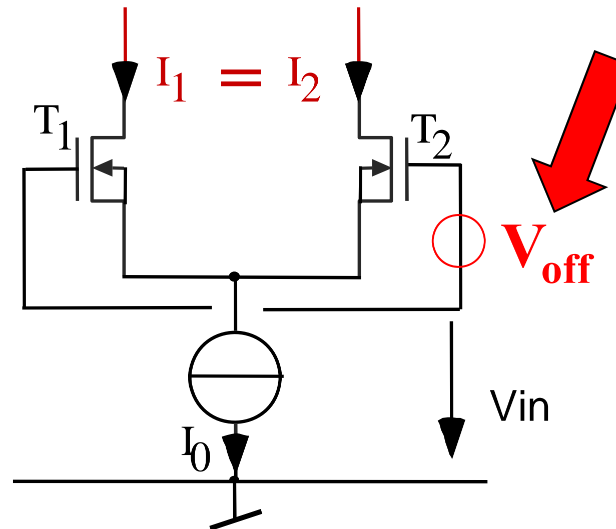


Mismatch effect on diff. Pair

$$\sigma_\beta = \sigma\,(\delta\beta/\beta)\ \&\ \sigma_T = \sigma\,(\delta V_{T0})$$

Techno: $A_\beta = \sigma_\beta \sqrt{W.L}$ et $A_T = \sigma_T \sqrt{W.L}$

DC-Offset



Mismatch effect
modeling

$$\Delta V_G = \Delta V_{T0} - \frac{I_D}{g_m}\frac{\Delta\beta}{\beta}$$

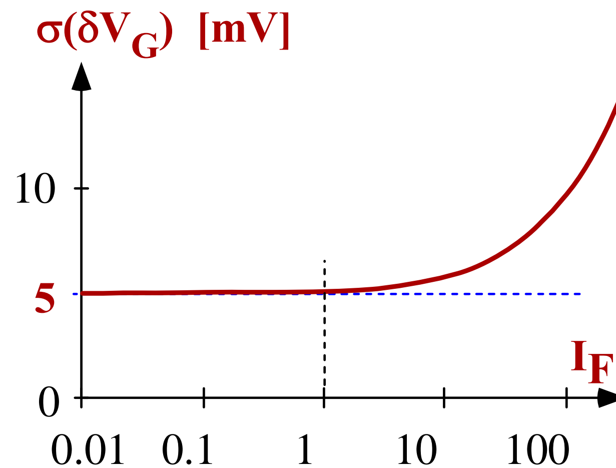
$$\sigma(\Delta V_G) = \sqrt{\sigma_T^2 + (\frac{I_D}{g_m}\sigma_\beta)^2}$$

☺ ***Offset is minimum when gm is maximum -> weak inv.***

Example

$$\sigma(\Delta V_G) = \sqrt{\sigma_T^2 + (\frac{I_0}{2g_m}\sigma_\beta)^2}$$

- With $\sigma_\beta = 2\%$ and $\sigma_T = 5mV$

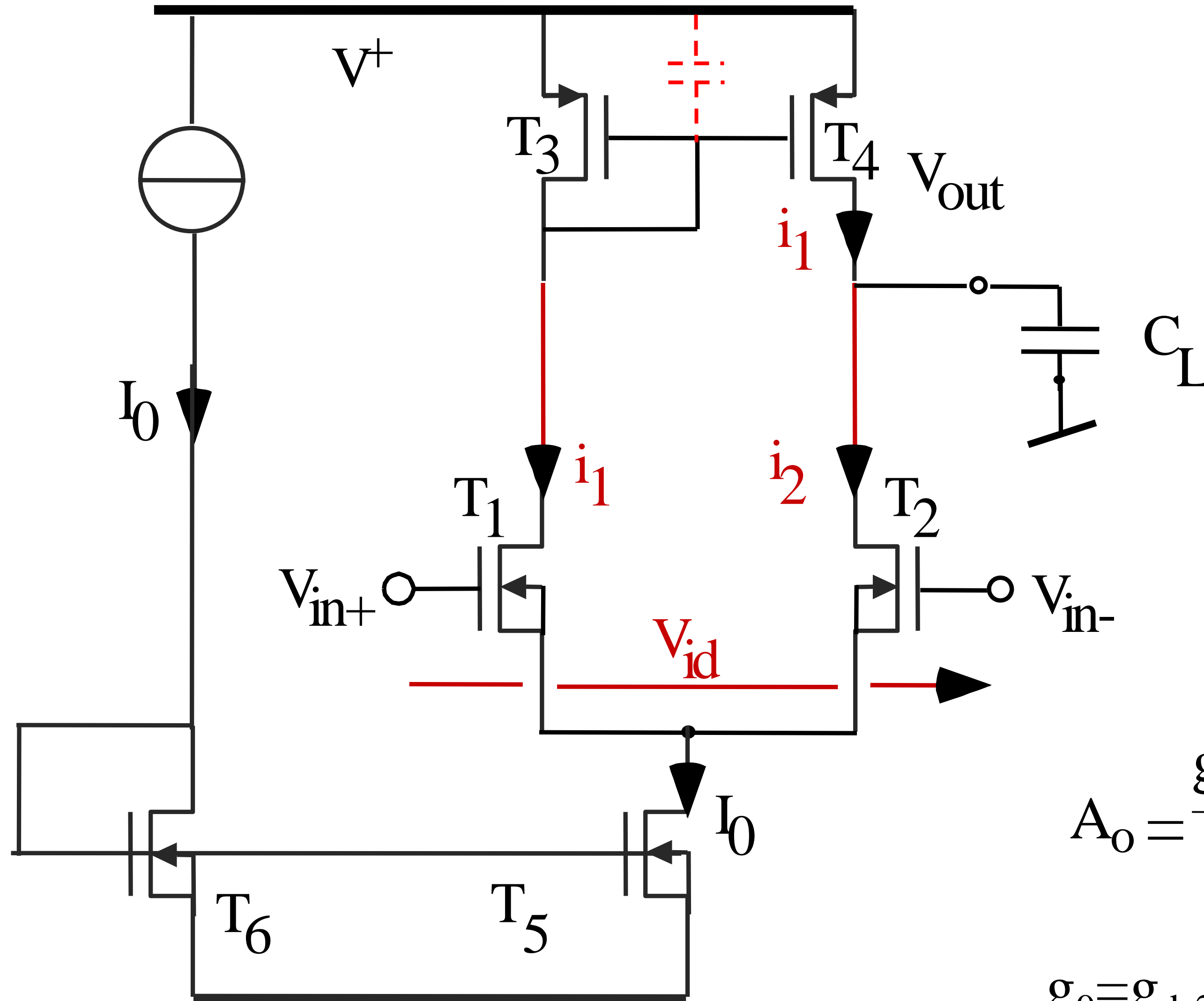


- Low offset Voltage in weak inv.

OTA Cell Specifications

- Gain- A_0 [dB]
- Gain Band width product – GBW [rad/s]
- Slew Rate- S_R [V/μs]
- Common-mode input range- CMR [V]
- Common-mode rejection ratio- CMRR [dB]
- Power-supply rejection ratio-PSRR [dB]
- Output-voltage swing-ΔV_{out} [V]
- Offset- [mV]
- Noise – nv / sqrt (Hz)

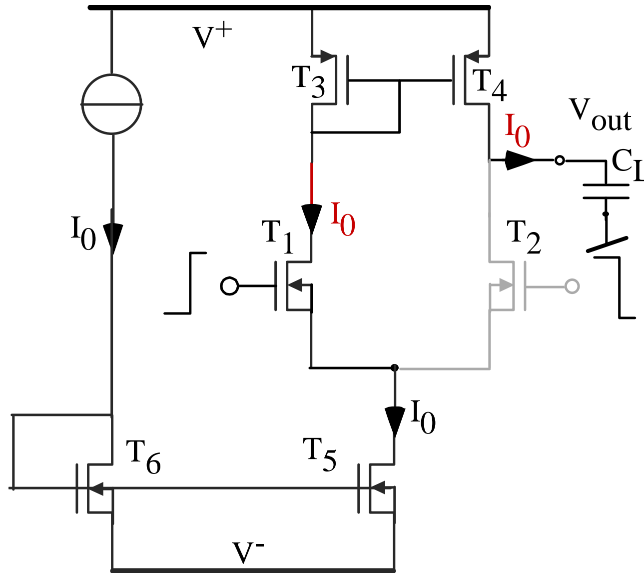
OTA

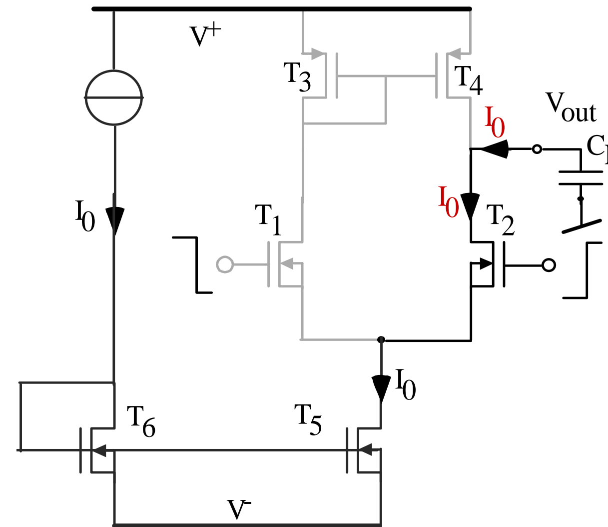


$$A_o = \frac{g_{m1,2}}{2 g_o}$$

$$g_0 = g_{ds2} + g_{ds4}$$

OTA-Slew Rate





$$S_R = \left|\frac{dV_{out}}{dt}\right|_{max} = \left|\frac{dI_{out}}{C_L}\right|_{max} = \frac{I_0}{C_L}$$

OTA Basic Structure Design

- ***Small Signal Voltage Gain:***

$$Av = g_{m2}.R_{o,OTA} = \frac{g_{m2}}{g_{ds2} + g_{ds4}} = \frac{g_{m2}}{I_{D2}} \frac{L}{(1/U_{a2} + 1/U_{a4})}$$

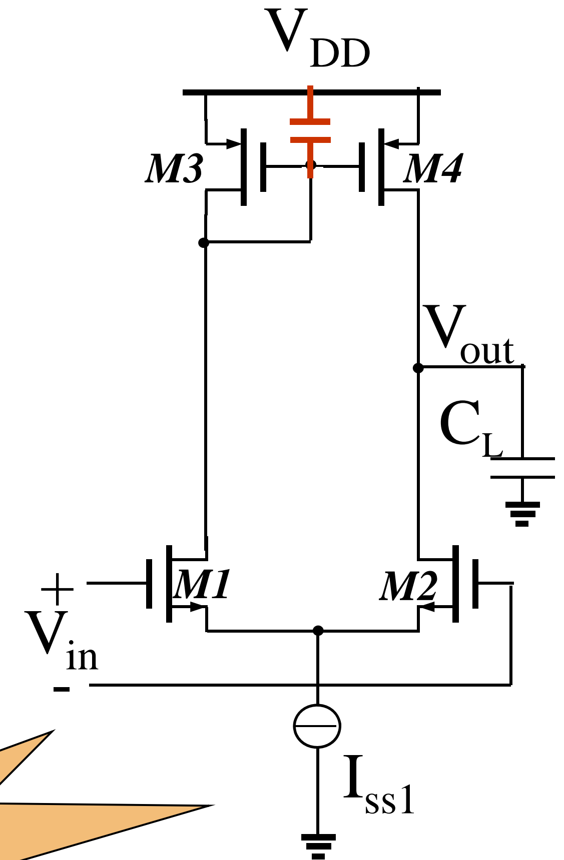
$$Gain\ Band\ Width\ product, GBW \approx \frac{g_{m2}}{I_{D2}} . \frac{I_{ss1}}{2C_L}$$

$$P_2 \approx \frac{g_{m3}}{C_{gs3} \cdot \left(1 + C_{gs4}/C_{gs3} + \sum C^B_{parasitic}/C_{gs3}\right)}$$

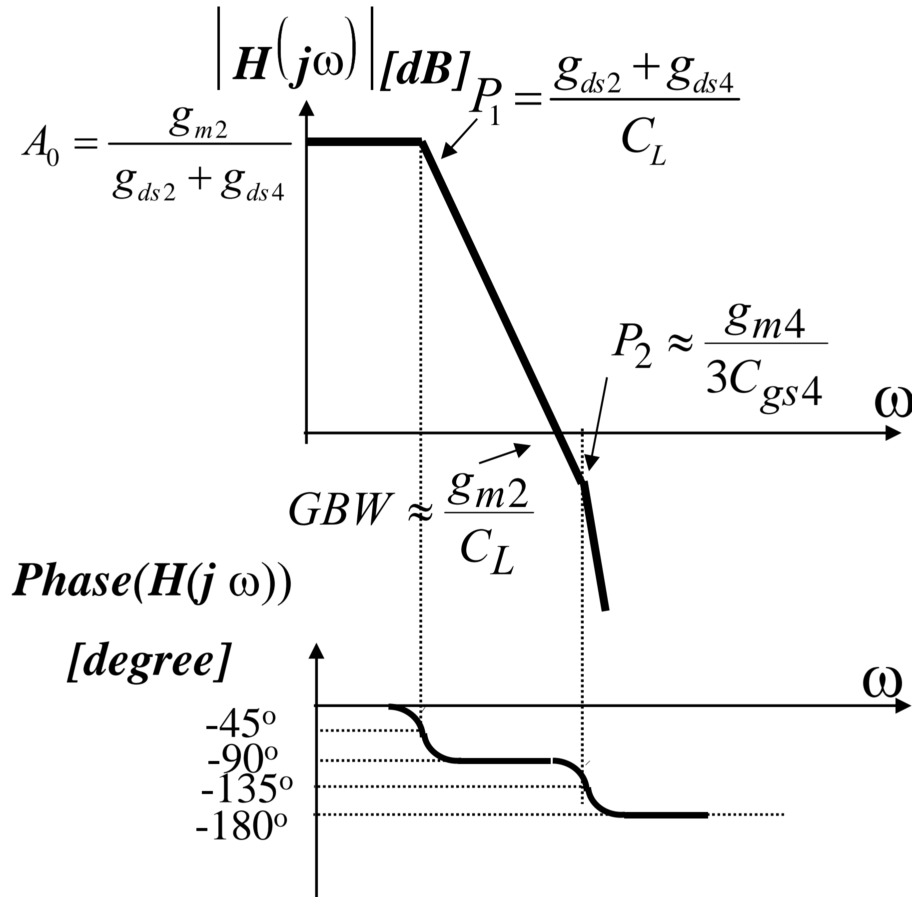
$$P_2 \approx \frac{g_{m3}}{3C_{gs3}} = \frac{g_{m3}}{I_{D3}} . \frac{I_{D3}}{3C_{gs3}}$$

$$SlewRate: \ S_R = \frac{I_{SS1}}{C_L} = \frac{2I_{D1,3}}{C_L}$$

$$\frac{GBW}{S_R} \approx \frac{g_{m2}}{2I_{D2}} = \frac{g_{m2}}{I_{ss1}}$$



OTA Basic Structure: Bode Plot



$P_2 \approx 3$ GBW!!!

$$\frac{g_{m4}}{I_{D4}} > \frac{g_{m2}}{I_{D1}} \frac{9C_{GS4}}{C_L}$$

PM= 180°- arctg (ω_{GBW}/P_1)- arctg (ω_{GBW}/P_2) °

OTA Basic Structure: Noise

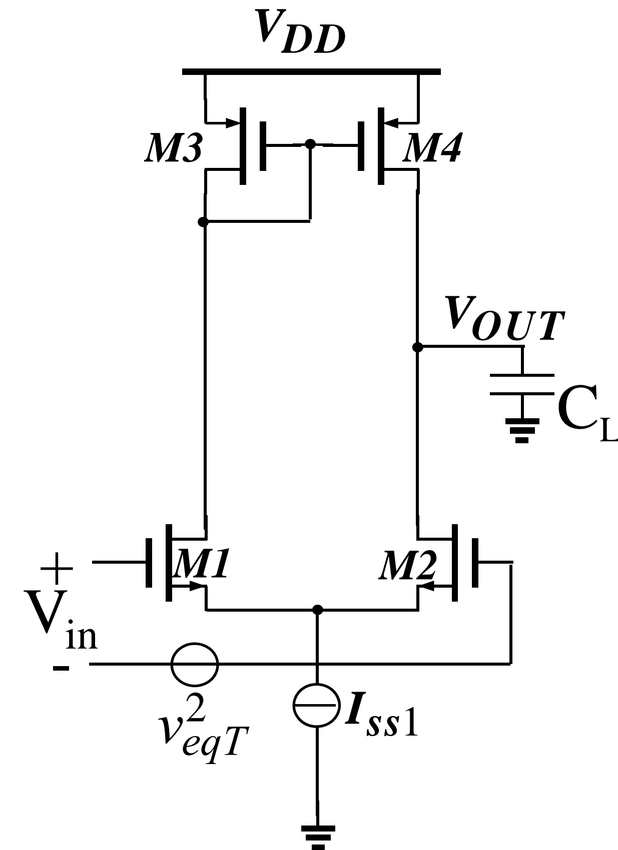
Total equivalent input noise : v_{eqT}^2:

$$v_{eqT}^2 = v_{eq1}^2 + v_{eq2}^2 + \left(\frac{g_{m4}}{g_{m2}}\right)^2 \cdot \left(v_{eq3}^2 + v_{eq4}^2\right)$$

Noise of active load is scaled by $\left(\frac{g_{m4}}{g_{m2}}\right)^2$

Minimize effect of device noise **:**

$$\frac{g_{m_4}}{I_{D_4}} << \frac{g_{m_2}}{I_{D_1}}$$



OTA - Input Offset Voltage(V_{off})

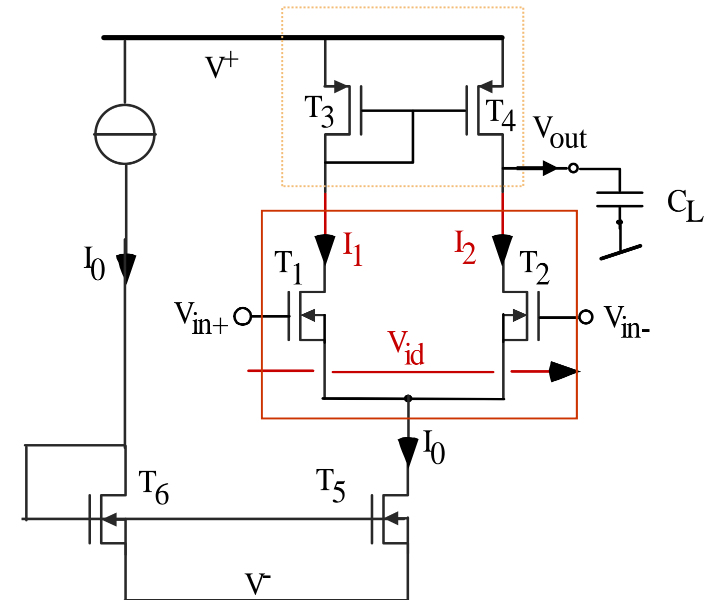
$$\Delta I_{out} = I_4 - I_2 = I_1(\varepsilon_m - \varepsilon_p)$$

$$\varepsilon_m = \frac{\Delta I_m}{I_m} = \frac{\Delta\beta_m}{\beta_3} - \frac{g_{m3}}{I_0/2}\Delta V_{T0,m}$$

$$\varepsilon_p = \frac{\Delta I_p}{I_p} = \frac{\Delta\beta_p}{\beta_1} - \frac{g_{m1}}{I_0/2}\Delta V_{T0,p}$$

$$I_{off} = \Delta I_{out} = \frac{I_0}{2}\left[\frac{\Delta\beta_m}{\beta_3} - \frac{\Delta\beta_p}{\beta_1}\right] - g_{m3}\Delta V_{T0m} + g_{m1}\Delta V_{T0p}$$

$$V_{off} = \frac{\Delta I_{out}}{g_{m1}} = \frac{I_0}{2g_{m1}}\left[\frac{\Delta\beta_m}{\beta_3} - \frac{\Delta\beta_p}{\beta_1}\right] - \frac{g_{m3}}{g_{m1}}\Delta V_{T0m} + \Delta V_{T0p}$$



$$\sigma^2(V_{off}) = \frac{1}{W_1.L_1}\left[A_T^2 + \left(\frac{I_0}{2g_{m1}}\right)^2 A_B^2\right] + \frac{1}{W_4.L_4}\left[A_B^2 + \left(\frac{2g_{m3}}{I_0 g_{m1}}\right)^2 A_T^2\right]$$

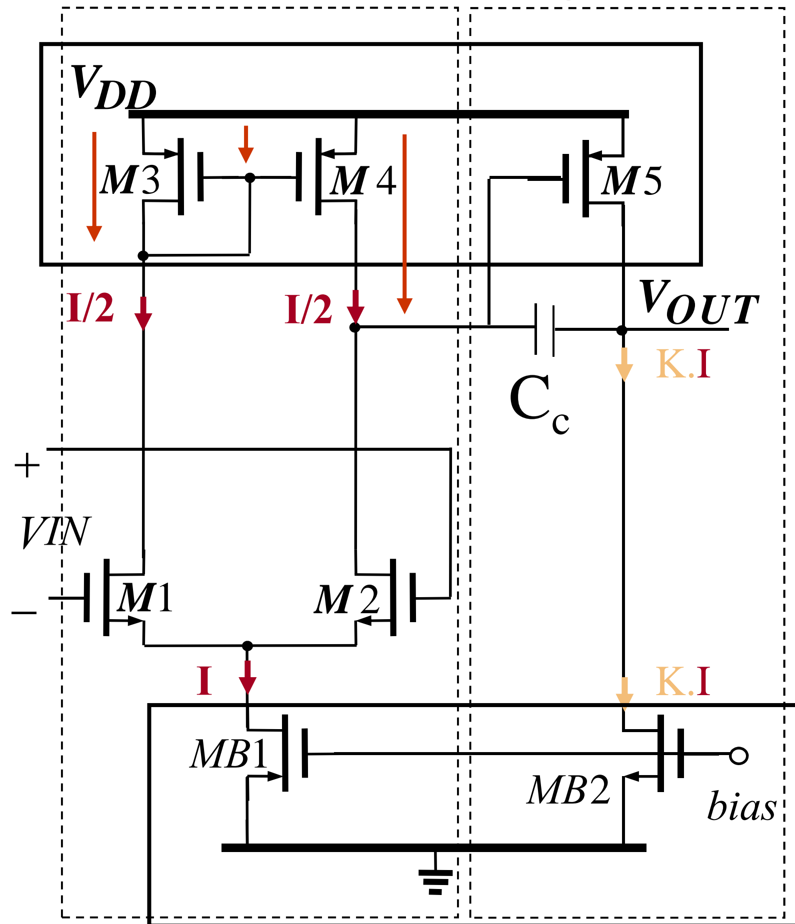
☺ $g_{m1,2} >> g_{m3,4}$

Circuit Level Tradeoffs

Speed	$\frac{GBW}{S_R} \approx \frac{g_{m2}}{2I_{D2}}$
Stability	$\frac{g_{m3}}{I_{D3}} > \frac{g_{m2}}{I_{D2}} \frac{9C_{GS3}}{C_L}$
Noise & Offset	$\frac{g_{m3}}{I_{D3}} << \frac{g_{m2}}{I_{D2}}$
DC Gain	$\frac{g_{m2}}{I_{D2}} \frac{L}{\left(1/U_{a2} + 1/U_{a4} \right)} \approx \frac{GBW.L}{S_R} \cdot \frac{U_a}{2}$
Consumption	$I_{SS1}V_{DD}=S_R.C_L.\ V_{DD}$
….	….

Miller Op-Amp

Op. Amp.: Two stages



- Common source output stage
- Open Loop Gain:

$$Av = = \frac{g_{m1}}{g_{ds2} + g_{ds4}} \cdot \frac{g_{m5}}{g_{ds5} + g_{ds_{MB2}}} =$$

$$\frac{g_{m2}}{I_{D2}} \frac{L}{\left(\frac{1}{U_{an}} + \frac{1}{U_{ap}} \right)} \frac{g_{m5}}{I_{D5}} \frac{L}{\left(\frac{1}{U_{an}} + \frac{1}{U_{ap}} \right)}$$

- Low systematic offset

$$\frac{\left(\frac{W}{L} \right)_{MB1}}{K \left(\frac{W}{L} \right)_{MB2}} = \frac{2 \cdot \left(\frac{W}{L} \right)_{M4}}{K \cdot \left(\frac{W}{L} \right)_{M5}}$$

Output Swing > OTA (Rail to Rail.)

Op. Amp.-2 stages: Compensation-2

$$GBW \approx \frac{g_{m1}}{C_c} \qquad P_2 \approx \frac{g_{m5}}{C_c + C_2} \qquad Z \approx \frac{g_{m5}}{C_c}$$

(60 degree phase margin)

$$\frac{P_2}{GBW} > 2.2 \longrightarrow \frac{g_{m5}}{g_{m1}} > \frac{2.2(C_c+C_L)}{C_c} \Rightarrow \frac{g_{m5}}{I_5} > \frac{1}{2K}\frac{2.2(C_c+C_L)}{C_c}\frac{g_{m1}}{I_1}$$

$$\frac{Z}{GBW} > 10 \longrightarrow \frac{g_{m5}}{g_{m1}} > 10 \Rightarrow \frac{g_{m5}}{I_5} > \frac{5}{K}\frac{g_{m1}}{I_1}$$

$$\frac{g_{m5}}{I_5} > \frac{5}{K}\frac{GBW}{S_R}$$

Circuit Level Tradeoffs

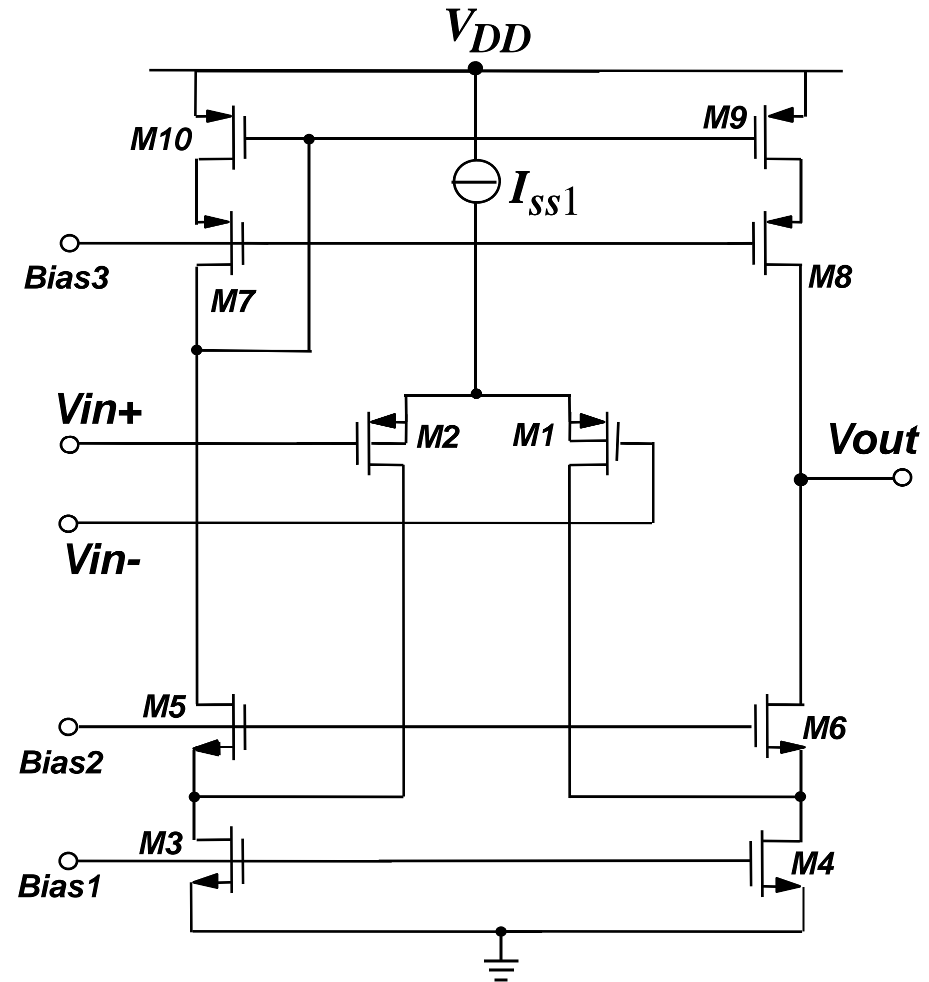
Speed	$\frac{GBW}{S_R} \approx \frac{g_{m2}}{2I_{D2}}$
Stability	$\frac{g_{m5}}{I_{D5}} > \frac{5}{K}\frac{GBW}{S_R} = \frac{5}{K}\frac{g_{m2}}{2I_{D2}}$ $\frac{g_{m5}}{I_5} > \frac{1}{2K}\frac{2.2(C_c + C_L)}{C_c}\frac{g_{m1}}{I_1}$
Noise & Offset	$\frac{g_{m3}}{I_{D3}} << \frac{g_{m1}}{I_{D1}}$
DC Gain	$\frac{g_{m2}}{I_{D2}}\frac{g_{m5}}{I_{D5}}\frac{L^2}{(1/U_{an} + 1/U_{ap})}$
Consumption	$(K+1)I_{SS1}V_{DD}$
….	….

$With\ \frac{I_{D5}}{I_{D1,2}} = 2K$

Cascoded OTAs

OTA/ Folded-Cascode

- Configuration:
  - Common source + common Gate
- Current in the input devices can be set to be different from the active load current



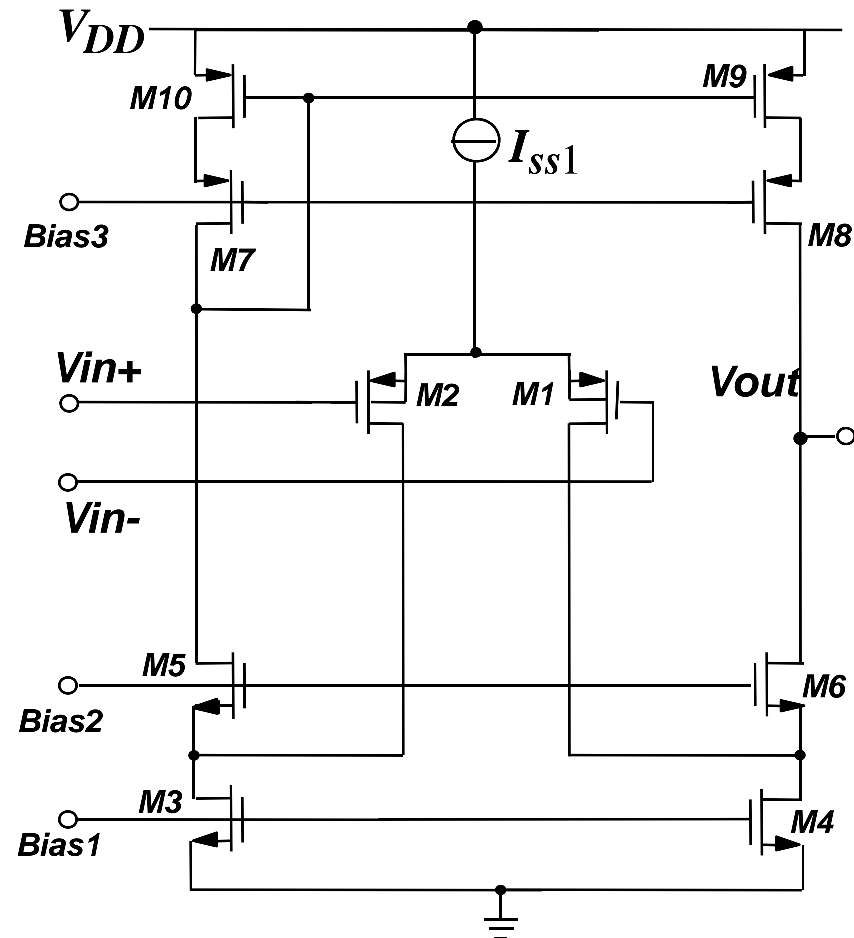
Folded-Cascode: Output Swing

- Output swing:

$$V_{out}^{Max} = V_{DD} - \left(V_{dsat}^{M8} + V_{dsat}^{M6} + V_{ds}^{M4} + V_{ds}^{M9}\right)$$

- Maximize output swing by:
  - –Minimizing the V_{dsat} of cascode
    - ☹ Degrades speed

✓ *Bias M5,6 & M7,8 to get M3,4 & M9,10 operate close to edge of saturation*



Folded-Cascode: Open-loop Gain

- Gain

$$Av = g_{m1}.R_{o,OTA} = \frac{g_{m1}}{g_{ox} + g_{oy}}$$

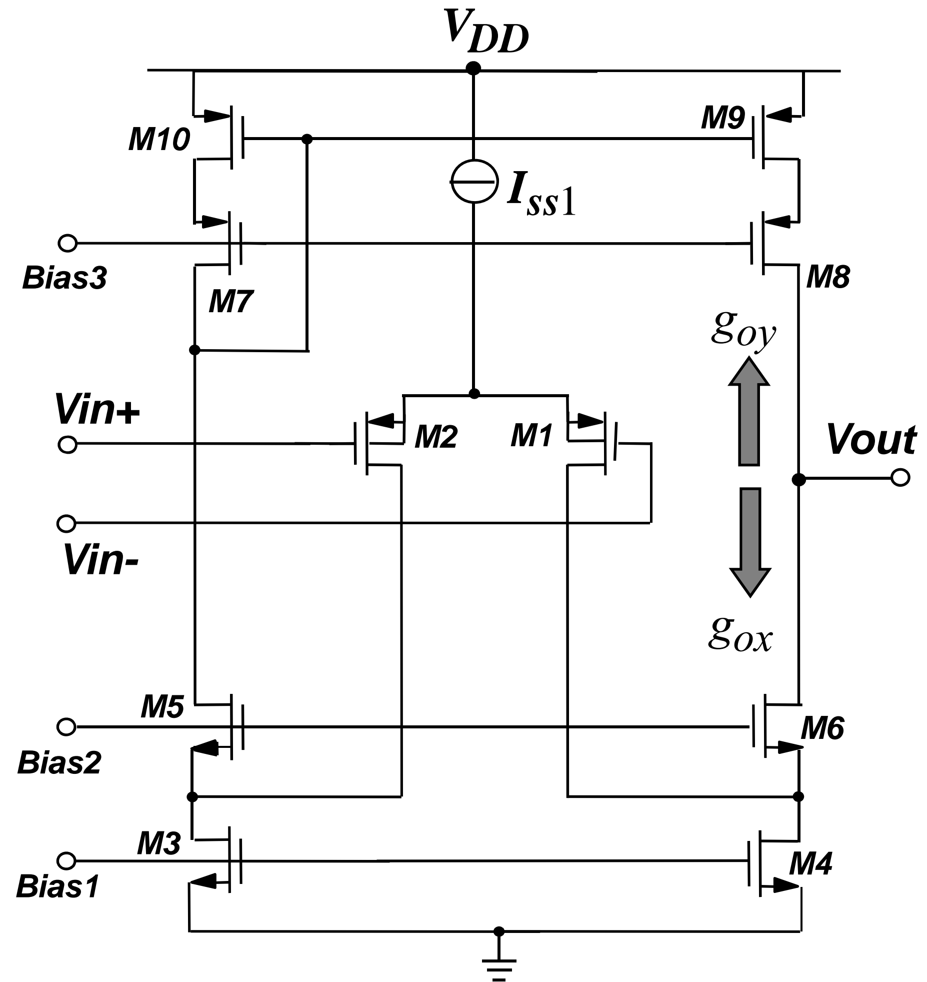
$$g_{ox} \approx \frac{g_{ds6} \cdot (g_{ds1} + g_{ds4})}{g_{m6}}$$

$$g_{oy} \approx \frac{g_{ds8} \cdot g_{ds9}}{g_{m8}}$$

$$Av \approx \frac{g_{m1} \cdot g_{m6}}{2.g_{ds6} \cdot (g_{ds1} + g_{ds4})}$$

- Output resistance

$$R_{o,OTA} \approx \frac{g_{m6}}{2.g_{ds6} \cdot (g_{ds1} + g_{ds4})}$$



Folded-Cascode: Dynamic Characteristics

- Dominant pole @ output node

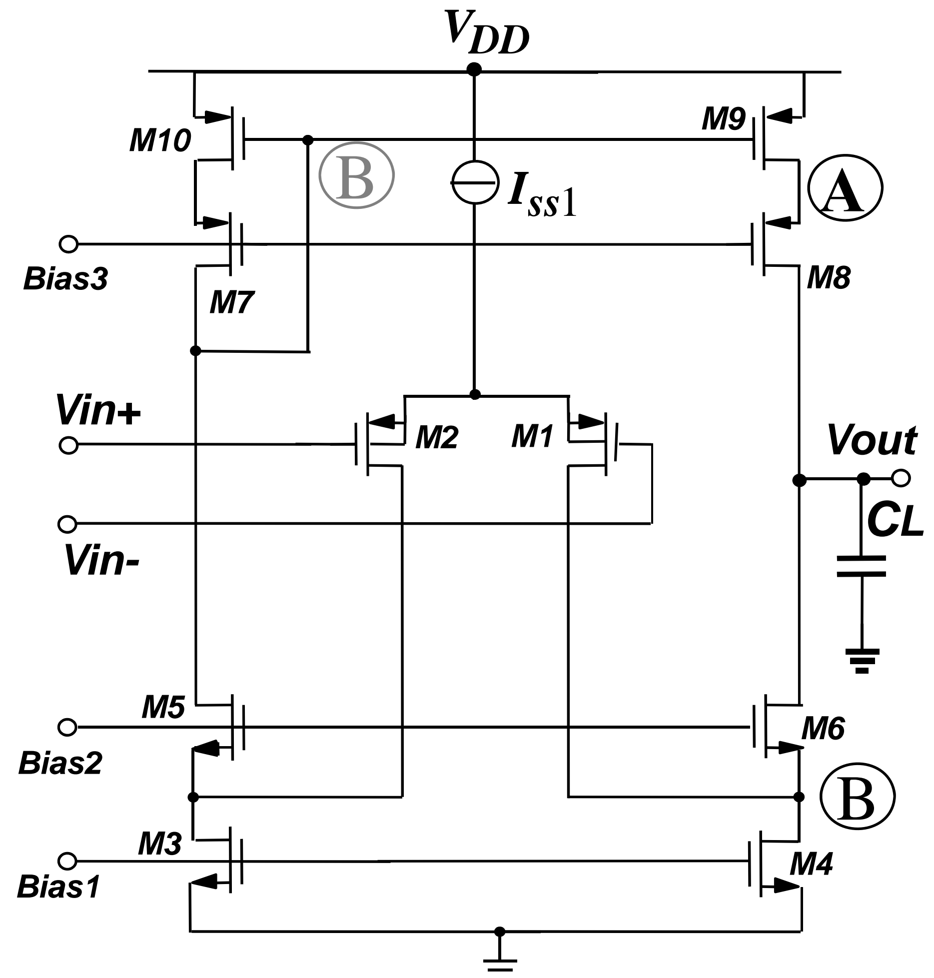
$$GBW \approx \frac{g_{m1}}{C_L}$$

- Non-dominant pole @ nodes A & B

  – Pole at node B is lower in frequency due to more parasitic:

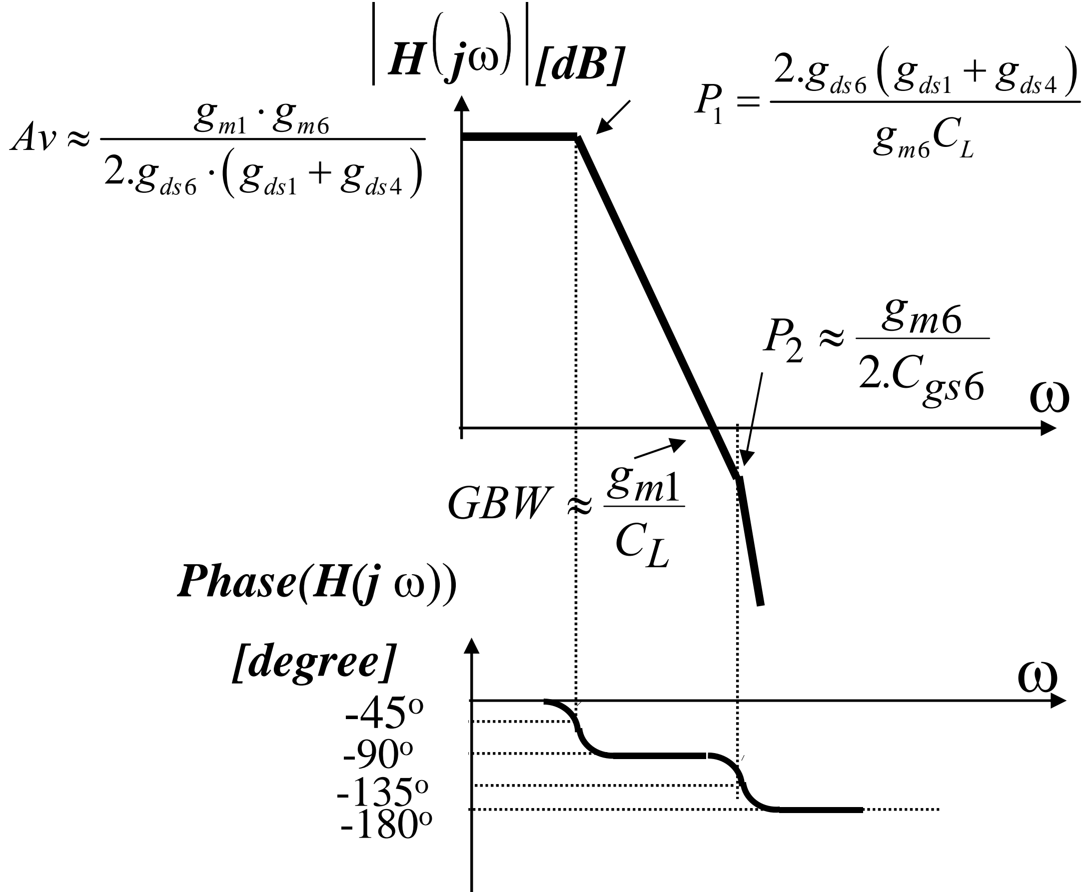
$$P_2 \approx \frac{g_{m6}}{C_{gs6} \cdot \left(1 + \sum C^B_{parasitic} \Big/ C_{gs6}\right)}$$

$$P_2 \approx \frac{2\pi \cdot f_T^{M6}}{2} = \frac{g_{m6}}{2\, C_{gs6}} \quad \text{(Estimation)}$$



✓ *Speed in the range of basic OTA structure while gain is very high*

Folded-Cascode: Bode Plot



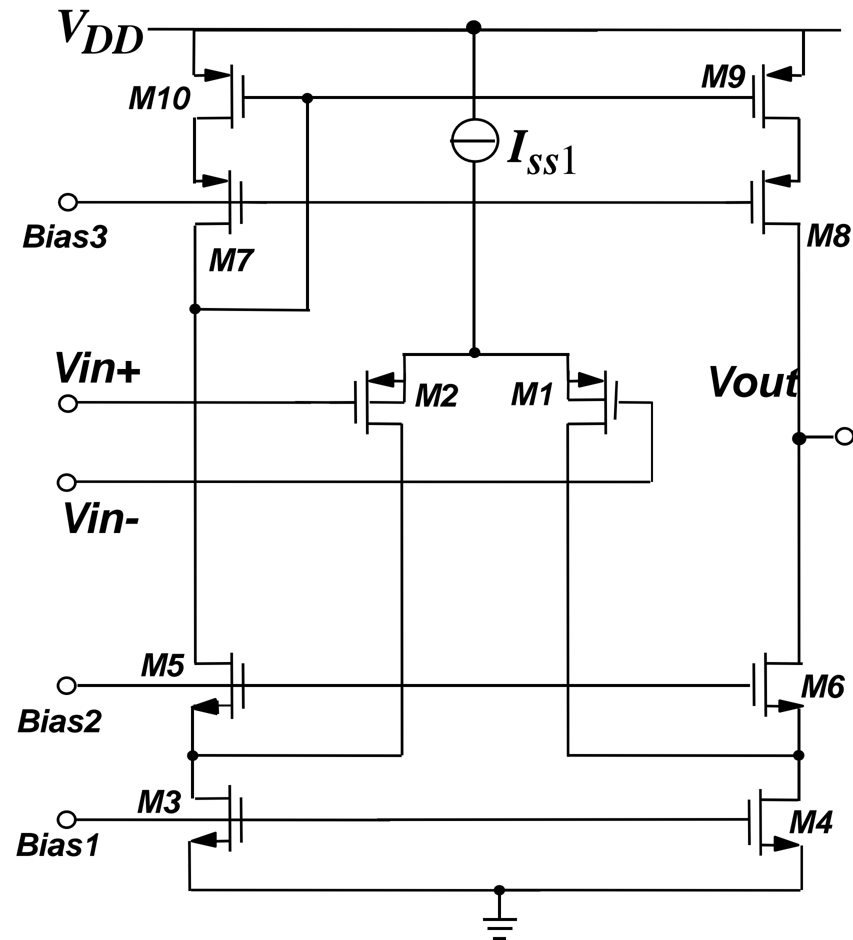
$$PM = 180° - \operatorname{arctg}(\omega_{GBW}/P_1) - \operatorname{arctg}(\omega_{GBW}/P_2)\ °$$

Folded-Cascode: Noise

Total equivalent input noise: v_{eqT}^2:

$$v_{eqT}^2 = v_{eq1}^2 + v_{eq2}^2 + \left(\frac{g_{m4}}{g_{m1}}\right)^2 \cdot \left(v_{eq3}^2 + v_{eq4}^2\right) + \left(\frac{g_{m9}}{g_{m1}}\right)^2 \cdot \left(v_{eq9}^2 + v_{eq10}^2\right)$$

No noise contribution by cascode devices!



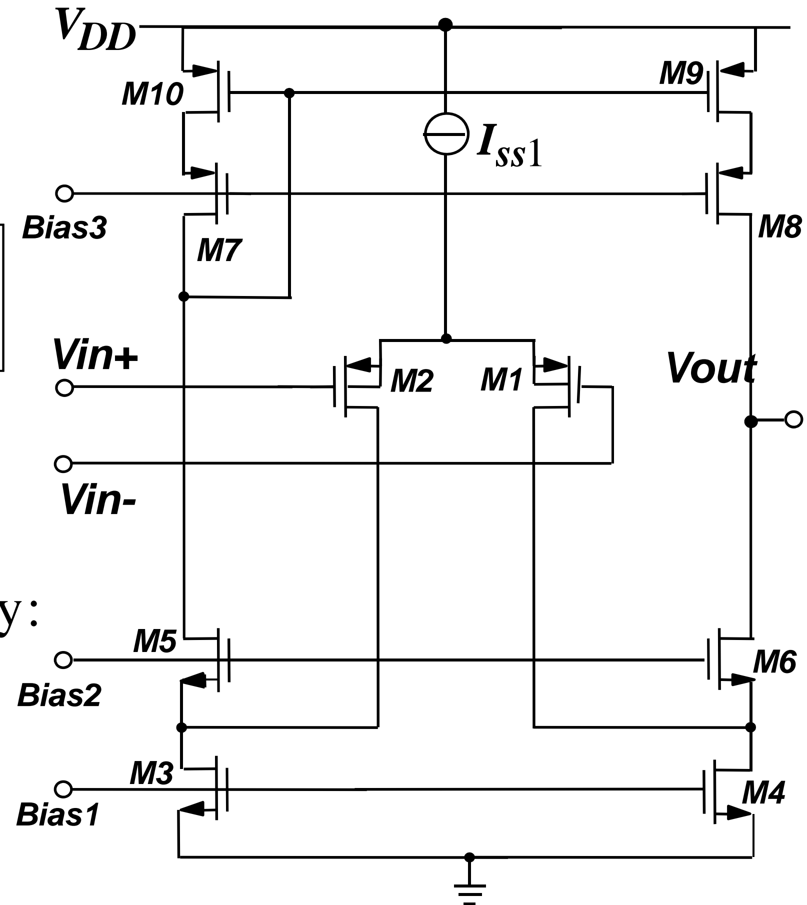
→ Minimize effect of load device noise by:

$g_{m4} << g_{m1}$ and $g_{m9} << g_{m1}$

→ Same as the case for basic OTA structure:

$L_4, L_9 > L_1$ both thermal & flicker noise contribution of load reduced

Folded-Cascode Offset



$$\sigma^2(V_{os}) = \sigma(\partial V_{G1})^2 + (\frac{g_{m3}}{g_{m1}})^2 \sigma(\partial V_{G3})^2 + (\frac{g_{m9}}{g_{m1}})^2 \sigma(\partial V_9)^2$$

$\rightarrow$ Minimize effect of load device contribution by:

$g_{m4} << g_{m1}$ and $g_{m9} << g_{m1}$

$\rightarrow$ Same as the case for basic OTA structure:

$L_4, L_9 > L_1$ both thermal & flicker noise contribution of load reduced

Circuit Level Tradeoffs

Speed	$\frac{GBW}{S_R} \approx \frac{g_{m2}}{I_{D2}}$
Stability	$\frac{g_{m6}}{I_{D6}} > 3(2K-1)\frac{g_{m2}}{I_{D2}}\frac{\Sigma C_B}{C_C}$
Noise & Offset	$\frac{g_{m4}}{I_{D4}} << 2K\frac{g_{m2}}{I_{D2}}$ & $\frac{g_{m9}}{I_{D9}} << (2K-1)\frac{g_{m2}}{I_{D2}}$
DC Gain	$\frac{g_{m2}}{I_{D2}}\frac{g_{m6}}{I_{D6}}L^2(U_{an}+U_{ap})$
Consumption	$2K.I_{SS1}V_{DD}$
….	….

$$With\ \frac{I_{D4}}{I_{D2}} = 2K\ \&$$

$$\frac{I_{D6}}{I_{D2}} = 2K-1$$

Fully Differential OTA

- Advantages:
  - Double the voltage swing compared to single-ended
  - Immunity to noise injected from substrate & supply.
- Disadvantages:
  - Requires extra circuitry to establish the common-mode output voltage. (Keep *Vx* constant)

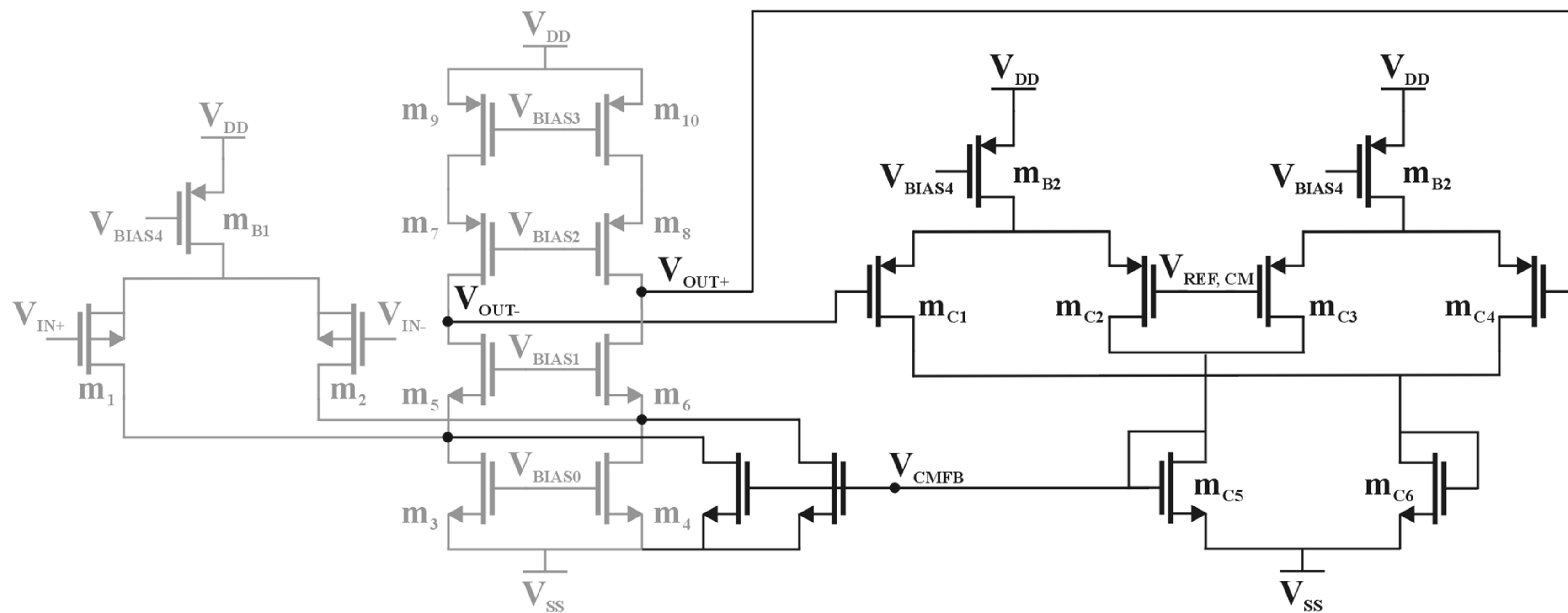
$$V_x = \frac{V_{out+} + V_{out-}}{2}$$

Stability of the C.M.F.B. loop has to be checked and insured. Inadequate phase margin causes common-mode oscillations

CMFB topologies

- amplifier with two differential pairs:

Example: Fully differential folded cascode OTA with CMFB circuit using two differential pairs



Digital calibration

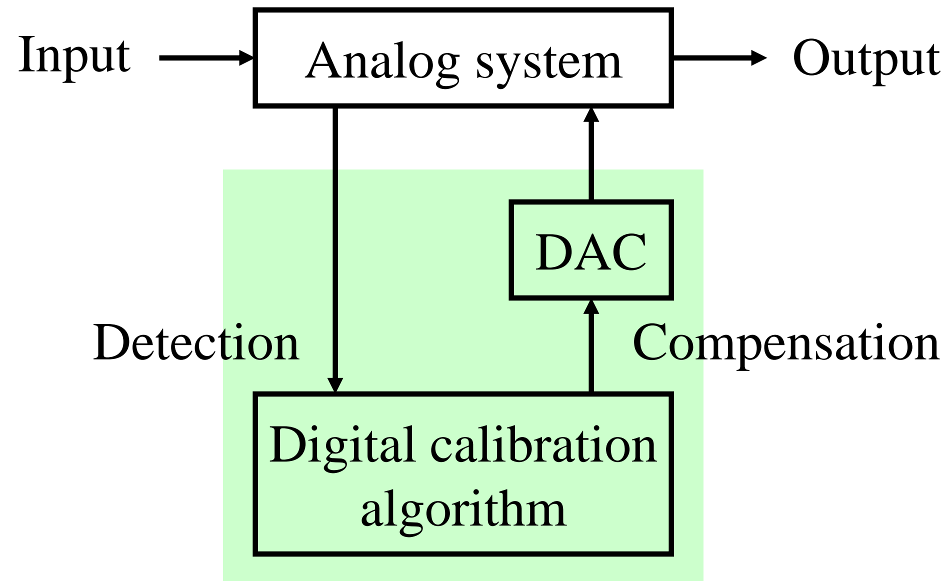
Outline

- Introduction
- Digital compensation of analog circuits
- Successive approximations:
  - Algorithm
  - Working condition
- Sub-binary DACs for successive approximations:
  - Resolution
  - Radix
  - Tolerance to component mismatch
  - Architectures
  - Design
- Conclusion

Digital calibration

- Analog design is complex
- Evolution of manufacturing process towards digital makes it even more difficult
  - Low supply voltage
  - Low-precision devices
- But digital can also help analog
  - Performance improvement by calibration
  - Relaxation of design constraints

Analog signal processing enhanced by digital calibration



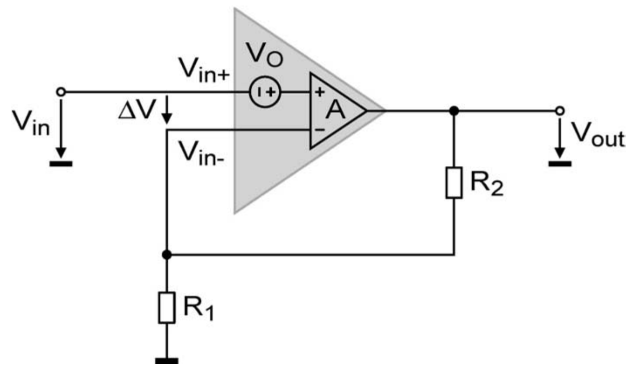
- High-precision calibration of low-precision circuits
- Alternative to intrinsically precise circuits (matching & high area)

Compensation methodology

- Detection configuration
  - Continuous: normal operation configuration
  - Interrupted: special configuration
- Detection node(s)
  - Imperfection sensing
  - Usually voltage-mode
- Compensation node(s)
  - Imperfection correction
  - Current-mode

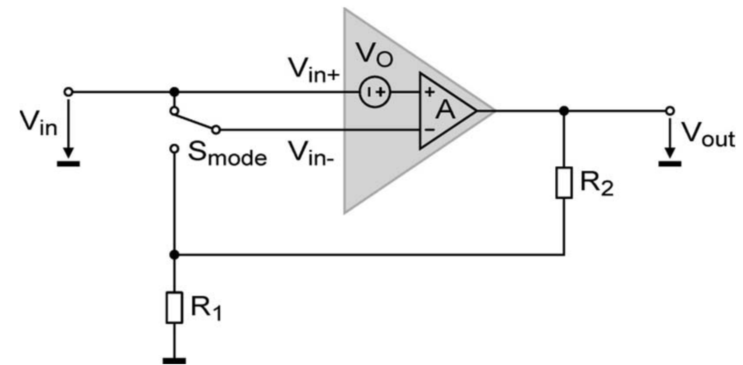
Offset cancellation in OAs

Closed-loop



$$\Delta V \cong -V_O$$

Open-loop

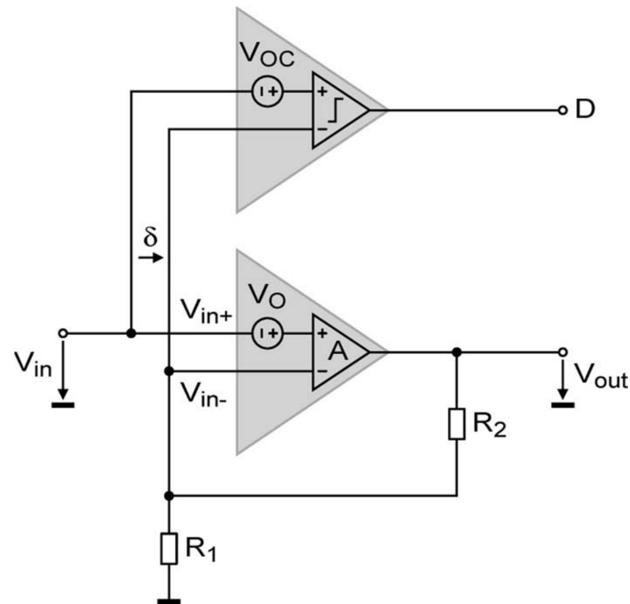


$$V_{out} = AV_O$$

- Closed-loop: calibration during operation possible
- Open-loop: higher detection level

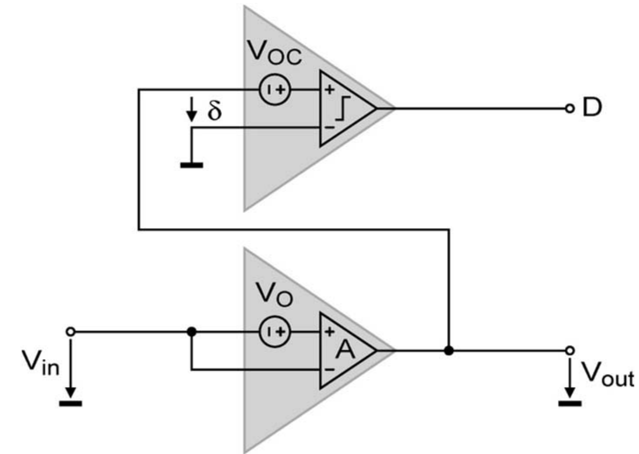
Offset detection in OAs

Closed-loop



$$V_{O;Compensated} = V_{OC}$$

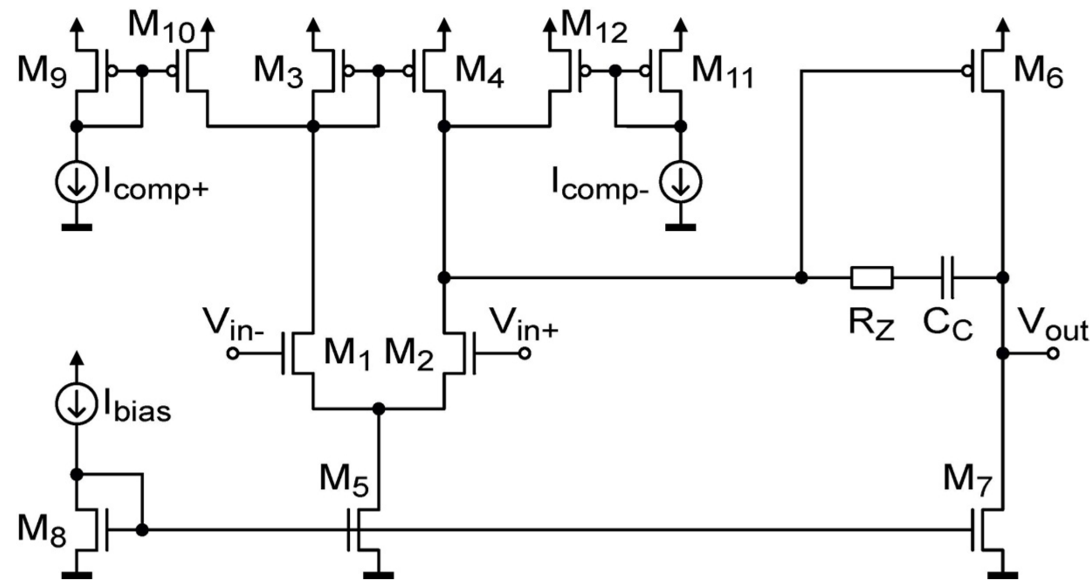
Open-loop



$$V_{O;Compensated} = -V_{OC} / A$$

- Open-loop configuration less sensitive to imperfections

Offset compensation in OAs

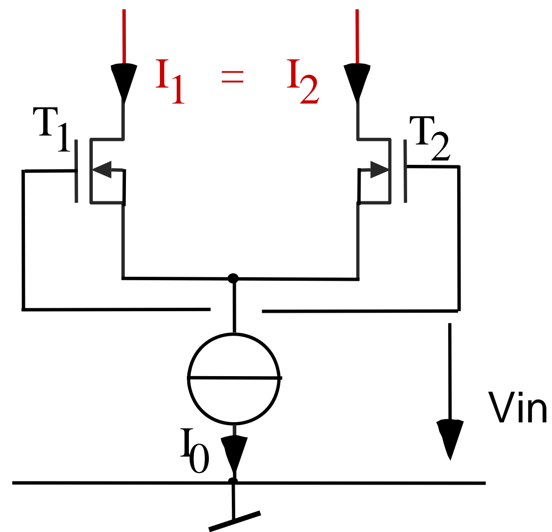


- Compensation by current injection
- Unilateral/bilateral
- Compensation current sources: M/2+M DACs

Choice of compensation node(s)

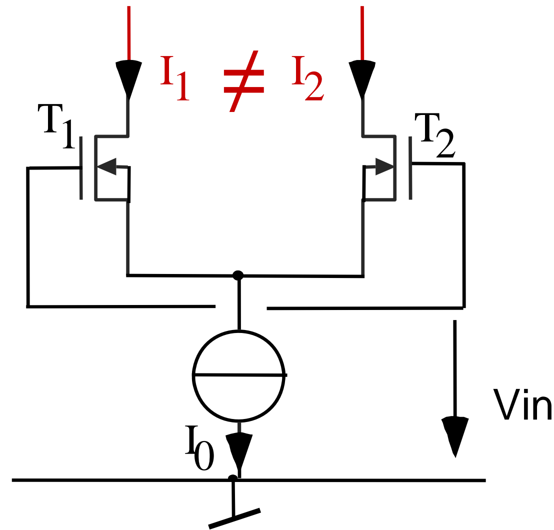
- Compensation current corrects imperfection only
- Current injected by a small current mirror, taking into account:
  - Channel length modulation
  - Saturation voltage
- Connection of the current mirror does not affect the compensation node characteristics:
  - Impedance
  - Parasitic capacitance
  - System parameters linked to parasitics

Ideal differential pair



- $I_1 = I_2 \Rightarrow$ No offset

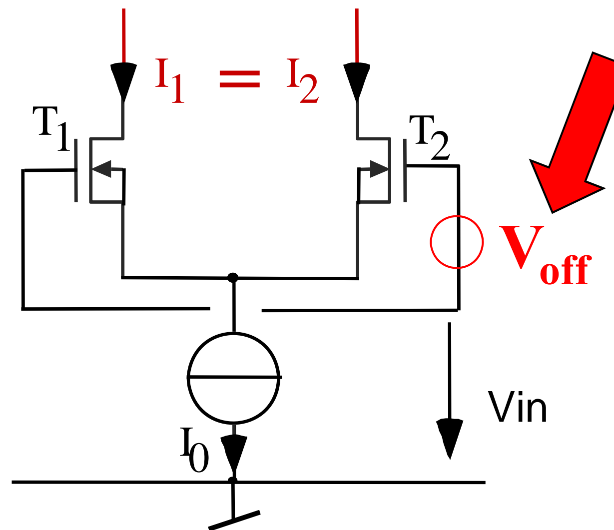
Differential pair: Offset



$$\sigma_\beta = \sigma\,(\delta\beta/\beta)\ \&\ \sigma_T = \sigma\,(\delta V_{T0})$$

$$A_\beta = \sigma_\beta \sqrt{W.L} \qquad A_T = \sigma_T \sqrt{W.L}$$

Differential pair: Offset modeling



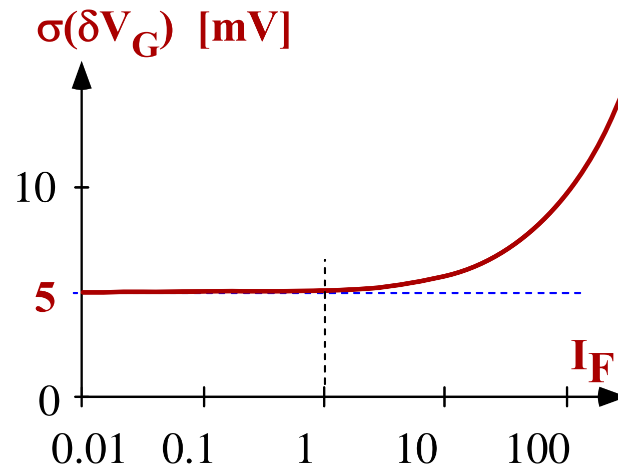
$$\Delta V_G = \Delta V_{T0} - \frac{I_D}{g_m}\frac{\Delta\beta}{\beta} \qquad \sigma(\Delta V_G) = \sqrt{\sigma_T^2 + (\frac{I_D}{g_m}\sigma_\beta)^2}$$

☺ ***Offset is minimum when g_m is maximum -> weak inv.***

Offset: Example

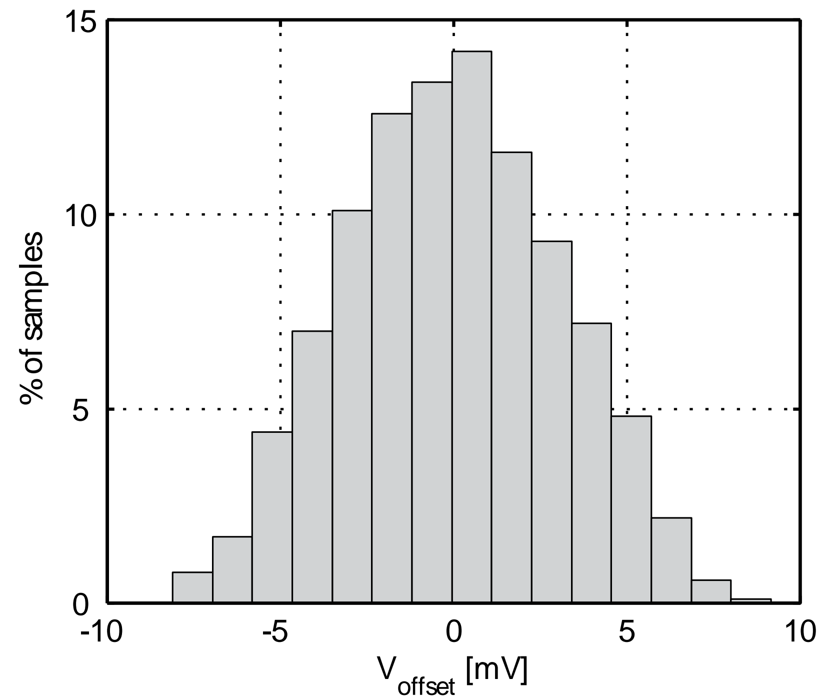
$$\sigma(\Delta V_G) = \sqrt{\sigma_T^2 + (\frac{I_0}{2g_m}\sigma_\beta)^2}$$

- With $\sigma_\beta = 2\%$ and $\sigma_T = 5mV$



- Low offset Voltage in weak inv.

Offset distribution before compensation

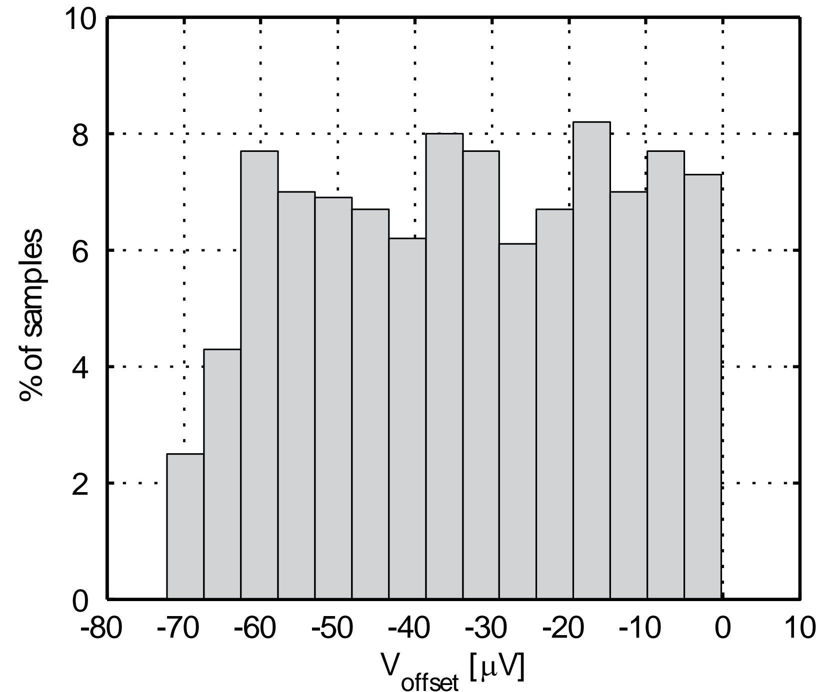


- Gaussian distribution
- Depends on component matching

Offset reduction

- Offset can be reduced by:
  - Matching $\Rightarrow$ Increase area
  - Digital calibration
- Digital calibration circuits can be made very small
- In deep sub-micron technologies:
  - Design analog circuits with reasonable performance
  - Enhance critical parameters by digital calibration
- Mixed-signal solution is optimal in terms of global circuit area

Offset distribution after digital compensation



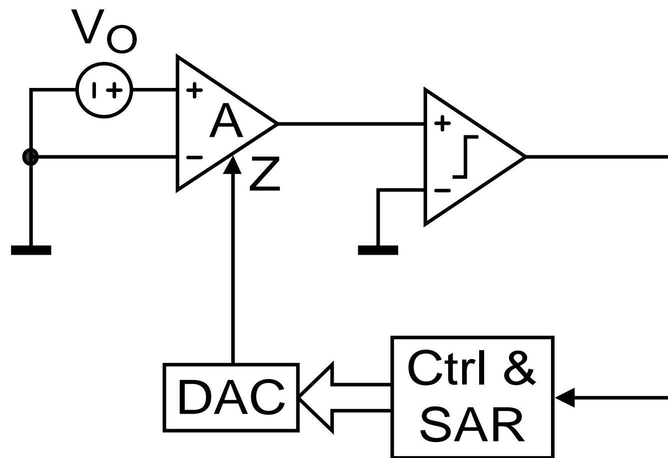
- Uniform distribution (in a 1 LSB interval)
- Residual offset depends on DAC resolution

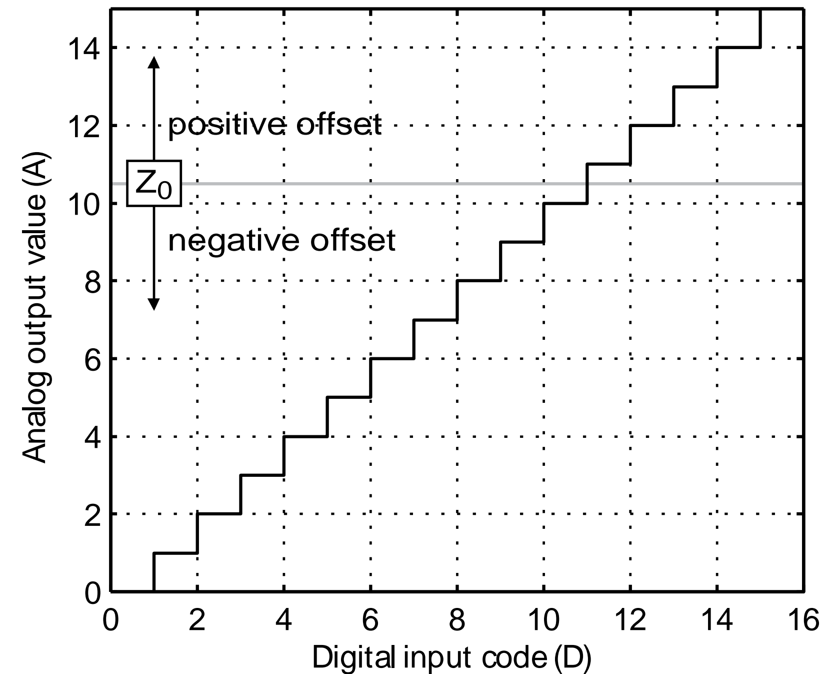
Successive approximations

```
reset all d_i = 0
for i = n downto 1
   set d_i = 1
   if C_out > 0
      reset d_i = 0
   end if
end for
```

- The algorithm decides on the basis of comparisons
- A comparator senses the sign of the imperfection
- Working condition: $b_i \leq b_1 + \sum_{j=1}^{i-1} b_j$ (i $\in$ [2, n])

Offset compensation: Target

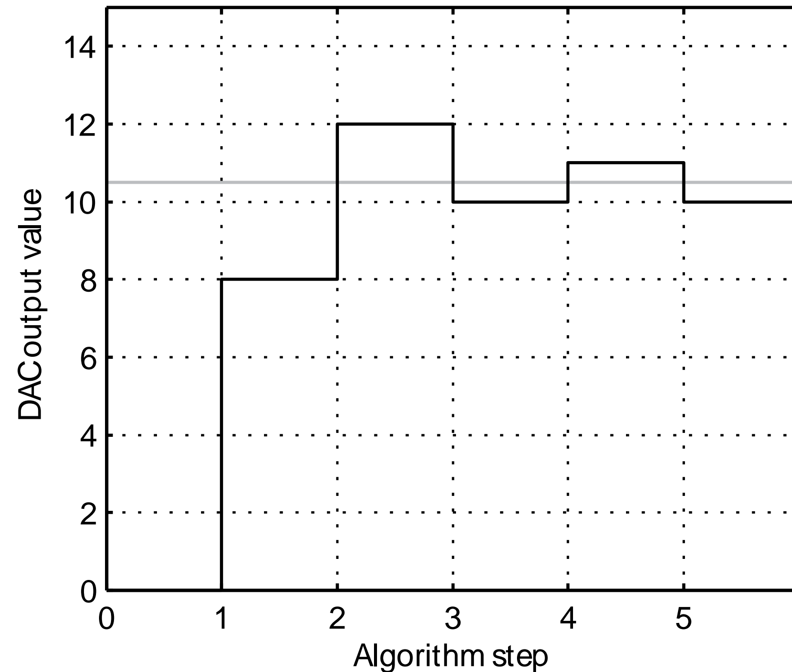




- Z_0: Correction value that perfectly cancels the offset
- $A < Z_0$: Resulting offset negative
- $A > Z_0$: Resulting offset positive

Offset compensation: Algorithm execution

```
reset all d_i = 0
for i = n downto 1
   set d_i = 1
   if C_out > 0
      reset d_i = 0
   end if
end for
```



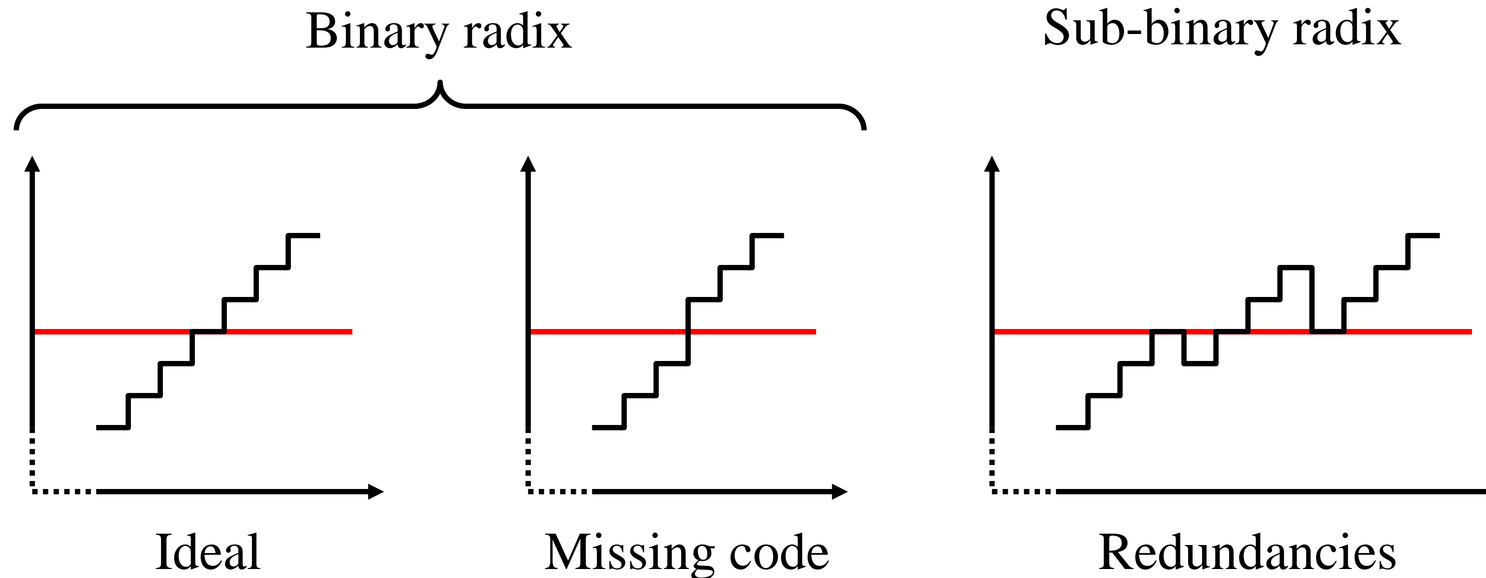
- From MSB to LSB
- Bit kept if compensation value insufficient

DAC Resolution

$$Resolution = \frac{FullScale}{LSB} \leq \frac{V_{offset;uncompensated;\max}}{V_{offset;compensated;\max}}$$

- Full scale chosen to cover whole uncompensated offset range
- Resolution corresponds to residual offset achieved after compensation

DACs for successive approximations

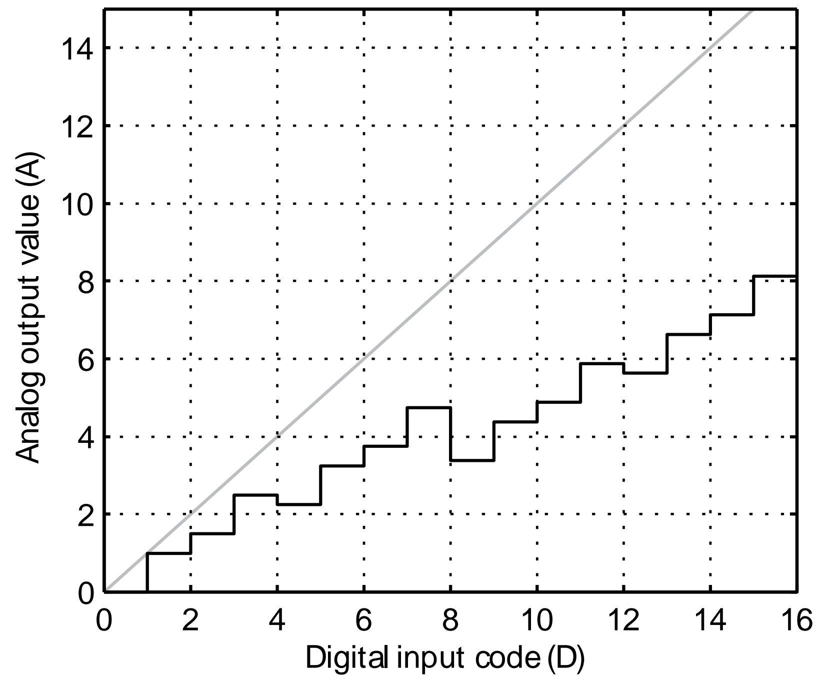


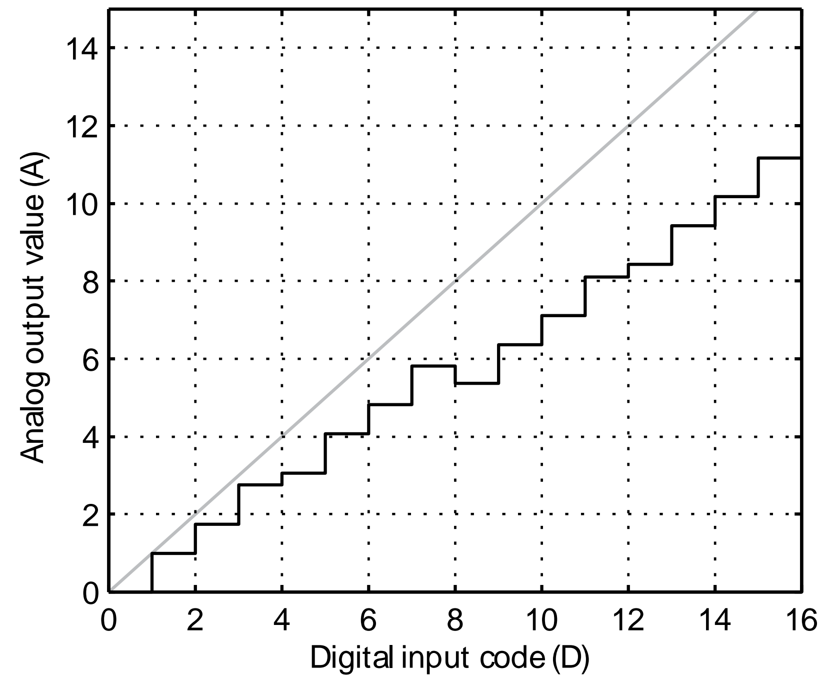
- Imperfections in DACs for compensation
  - Missing codes are problematic
  - Redundancies are acceptable

Sub-binary radix DACs

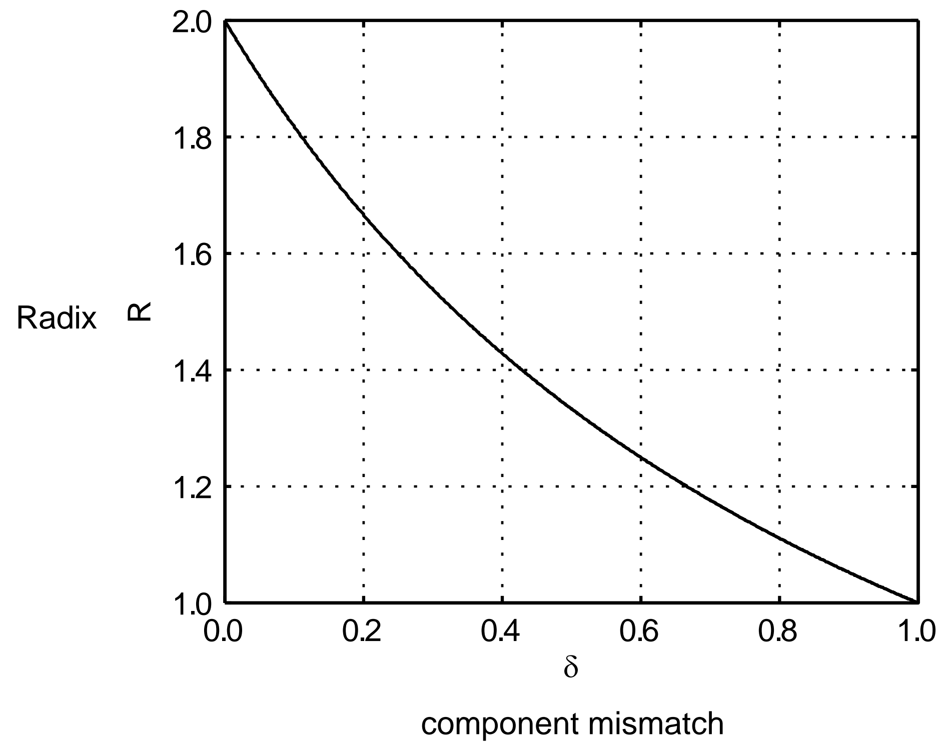
- Code redundancies are voluntarily introduced to:
  - Account for variations of component values
  - Avoid missing codes
- Arbitrarily high resolutions can be achieved without exponential increase of area
- For successive approximations:
  - Precision is not important
  - Resolution is the objective
- **Sub-binary DACs are ideal in conjunction with successive approximations**
  - **Very low area**

Sub-binary DACs: Radix



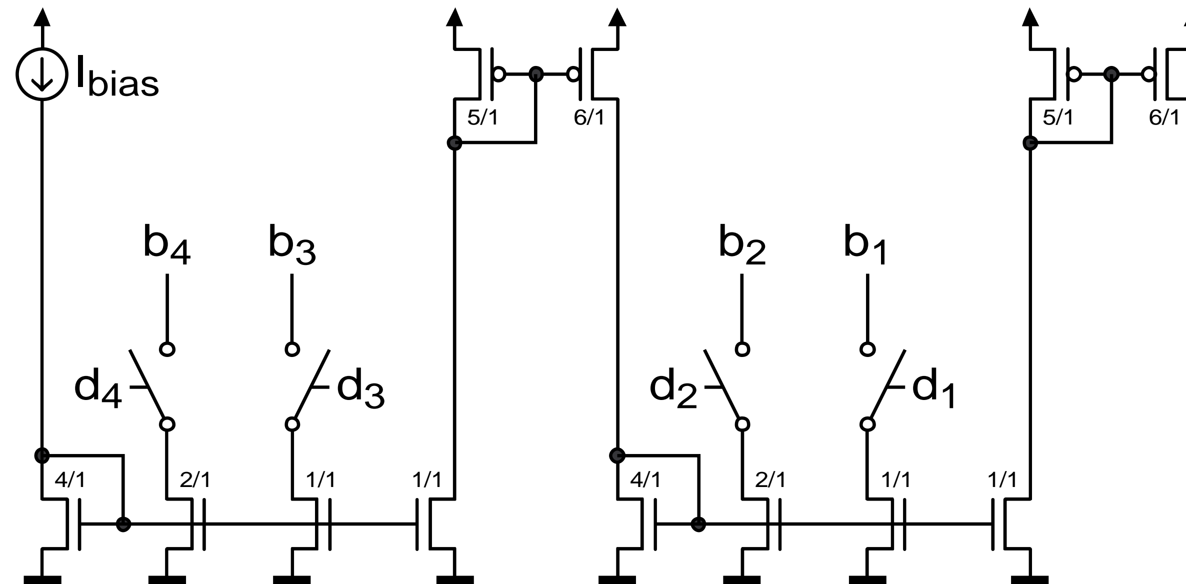


Radix: Tolerance to component mismatch



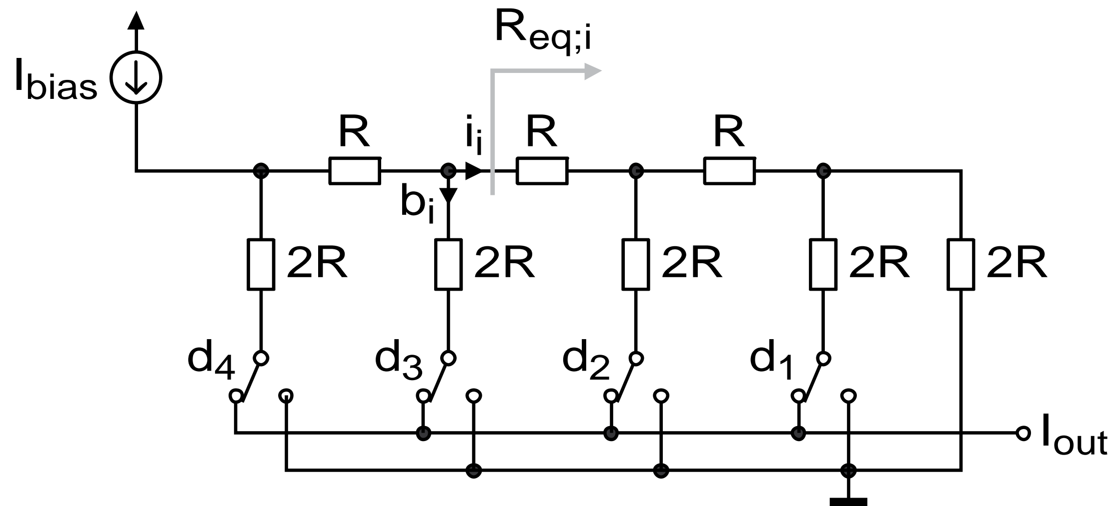
- Radix-2 tolerates no mismatch!
- 100 % mismatch $\Rightarrow$ thermometric DAC (radix-1)

Implementation: Multi-stage current mirrors



- Binary radix in each stage (NMOS mirrors)
- Sub-binary radix inter-stage (PMOS mirrors)

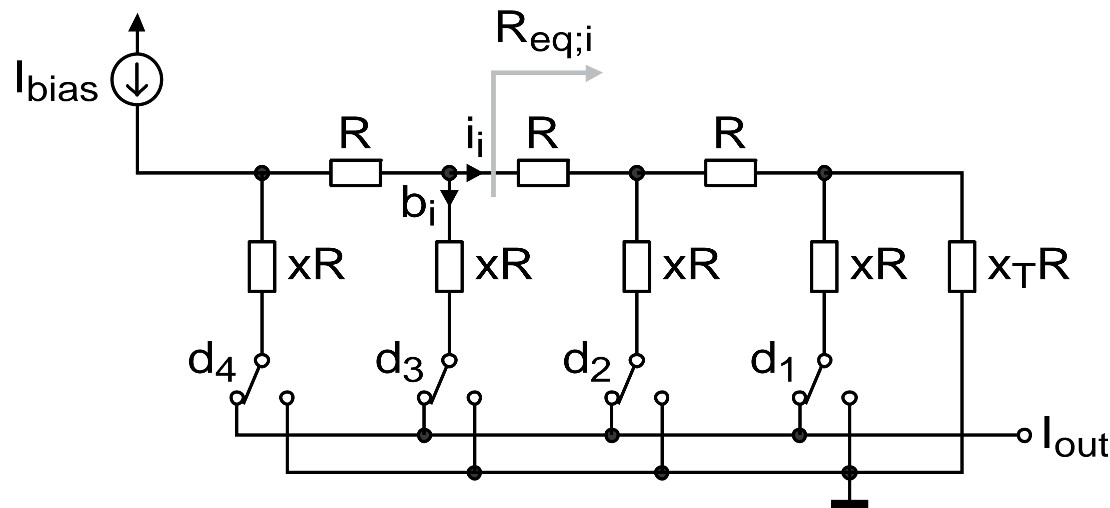
Implementation: Current-mode R/2R converters



$$b_i = b_1 + \sum_{j=1}^{i-1} b_j$$

- $R_{eq;i} = 2R$ ($\forall$ i)
- Current divided equally in each branch ($i_i = b_i$)
- Component imperfection $\Rightarrow$ current imbalance $\Rightarrow$ missing code (or redundancy)

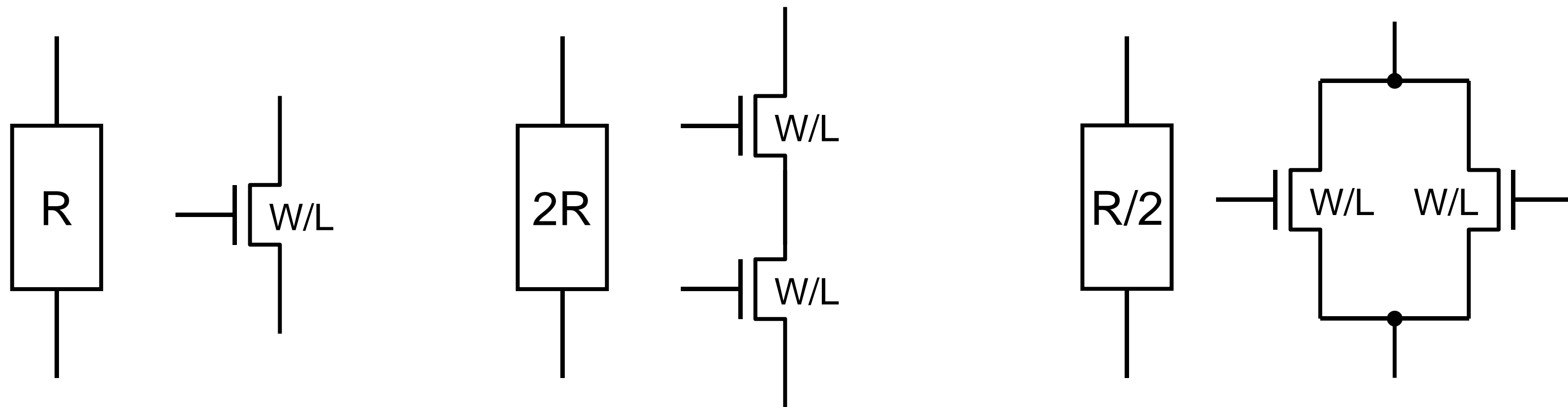
R/2⁺R converters



$$b_i < b_1 + \sum_{j=1}^{i-1} b_j$$

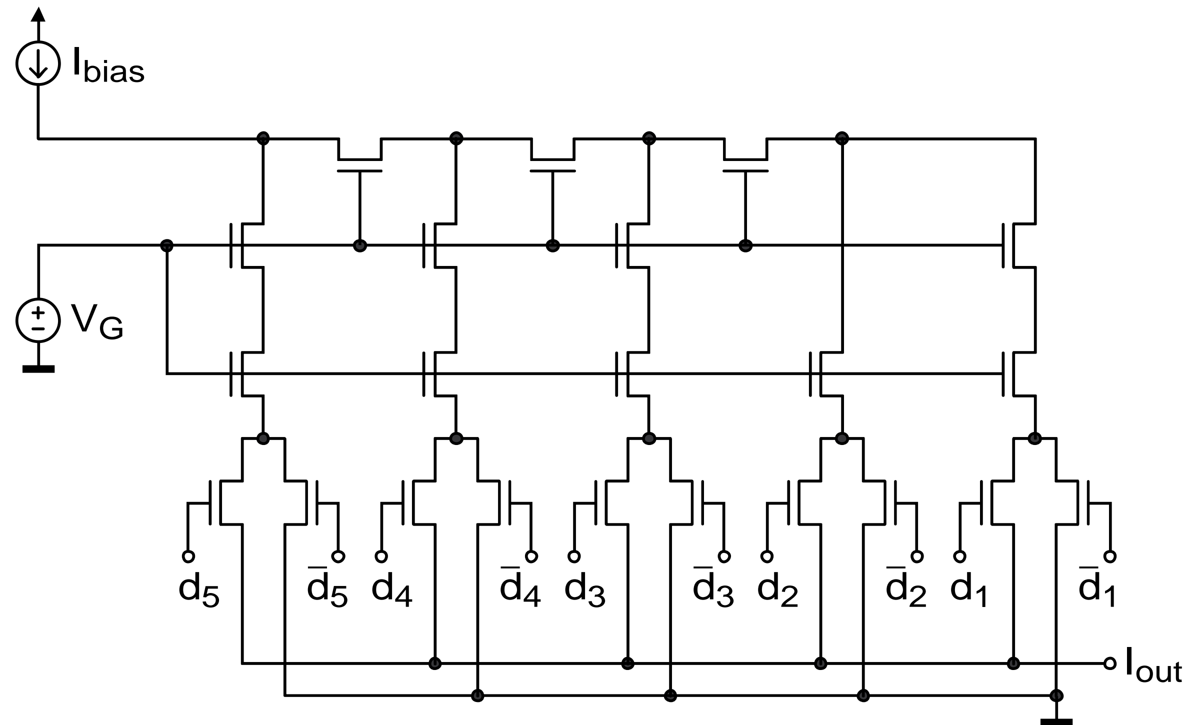
- $x > 2$; $R_{eq;i} < xR$; $i_i > b_i$
- Current division voluntarily unbalanced
  - Radix < 2 (sub-binary)
  - Code redundancies

R/2+R converters: Pseudo-MOS implementation



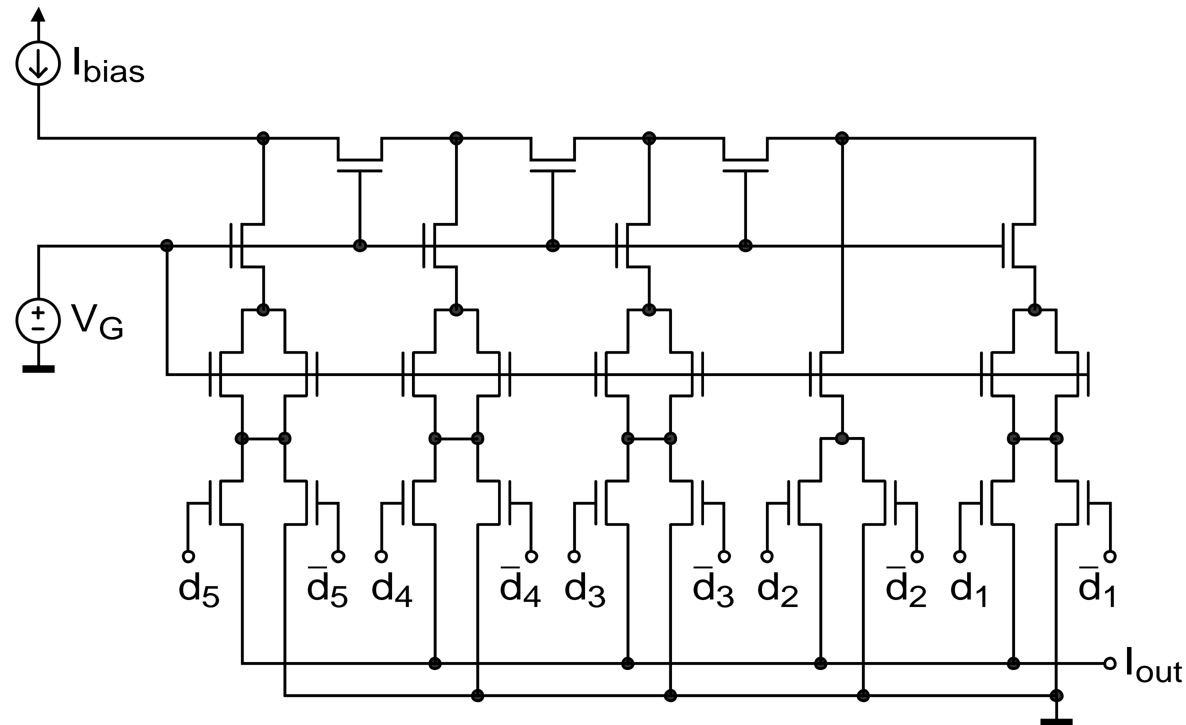
- Resistors can advantageously be replaced by transistors to implement the current division
- Unit-size device with fixed W/L implements R
- Unit-size devices are put in series (2R) or in parallel (R/2)
- Unit-size transistors are kept very small
- Condition: V_G identical for all transistors

M/3M converters



- Radix 1.77 ; maximum mismatch 13 %
- $V_G = V_{DD}$ allows driving d_i directly with logic

M/2.5M converters

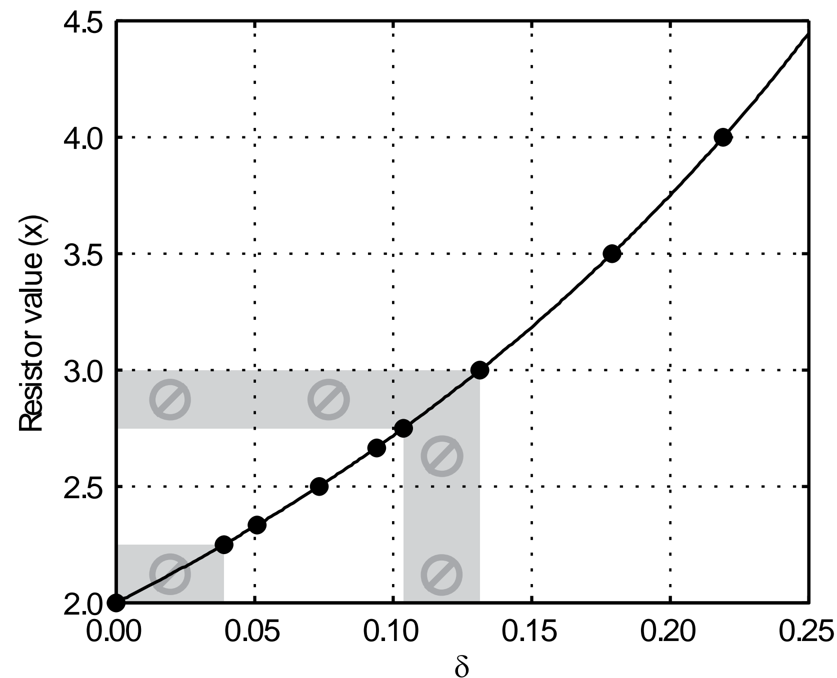


- Radix 1.86 ; maximum mismatch 7.3 %
- $V_G = V_{DD}$ allows driving d_i directly with logic

Current collectors & Output stage

- Current mirrors simple to implement
- ΔV is not problematic

Possible resistor/mismatch choices



- A large range of resistor values can be implemented easily
- $\varnothing \Rightarrow$ More than 6 transistors required to implement xR

Conclusion

Conclusion

- Deep submicron scaling of MOSFETs presents many advantages with respect to speed and circuit density, as required for example in RF applications.
- However, for analog applications, short channel effects, gate leakage current and matching will degrade the device performances. Moreover, it seems that power consumption will also increase upon lowering supply voltages.
- Shift towards lower gate overdrive due to the low-voltage operation required by deep-submicron processes will need an accurate modeling of moderate-to-weak inversion behavior, even for digital applications.

  This can be addressed through the g_m/I_D based design methodology.
- Calibration is necessary in deep sub-micron technologies to reach high performance

  Digital calibration circuits can be made very small:
  - Successive approximations
  - M/2⁺M sub-binary DAC