Industrial Challenges of Advanced Analog Design in Deep Submicron Process Technologies

July 2012

Ahmed Helmy, Ph.D.

Sr. Staff Design Manager, Intel Corp., Chandler AZ

Associate Professor, Arizona State University

Introduction

- Designing analog circuits on deep sub-micron process technologies (32, 22, 14, 10, 7, 5 nm) has its own set of challenges.

- With supply voltage scaling down, threshold voltages holding its values to maintain moderate leakage currents as demanded by low power designs for modern and mobile gadgets, the headroom is running out.

- Coupled with degraded rout and denser metal stack, stringent density rules, higher BE parasitics make analog design even tougher, especially for high yield products that are required to ship hundreds of millions of units with very low wafer cost.

- This focused short course is intended to address these challenges from an industrial view point, starting from the basics, building its way up to designing advanced analog blocks in deep submicron process nodes, while keeping yield as the number one challenge.

Who is an analog designer ?

- Studied analog design in collage
- Got a degree in analog design, BSC, MSC, PHD
- Taught analog design at the university
- Sat among people discussing analog design
- Talked about analog design
- Worked in an analog design company
- Designed many analog design circuits
- Laid out many analog design circuit
- Taped out many analog designs
- Sat in the lab to measure and debug analog designs
- Told his family that he is a designer
- Posted on his FB page that he is a designer

Who is a brain surgeon?

Qualities of an Analog Designer

- Never say this is what the simulator gave
- Meticulous
- Pessimistic
  - Always consider worst cases.
- Miser (area and power)
- Says this in his/her sleep
  - Variation
  - Random, Systematic, bias, PVT

Analog Designer which one? ☺

- **High-context cultures** (including much of the Middle East, Asia, Africa, and South America) are relational, collectivist, intuitive, and contemplative. This means that people in these cultures emphasize interpersonal relationships. Developing trust is an important first step to any business transaction. These cultures are collectivist, preferring group harmony and consensus to individual achievement. **People in these cultures are less governed by reason than by intuition or feelings**. Words are not so important as context, which might include the speaker's tone of voice, facial expression, gestures, posture—and even the person's family history and status. High-context communication tends to be more indirect and more formal. Flowery language, humility, and elaborate apologies are typical.

- **Low-context cultures** (including North America and much of Western Europe) are logical, linear, individualistic, and action-oriented. People from low-context cultures value logic, facts, and directness. **Solving a problem means lining up the facts and evaluating one after another**. Decisions are based on fact rather than intuition. Discussions end with actions. And communicators are expected to be straightforward, concise, and efficient in telling what action is expected. To be absolutely clear, they strive to use precise words and intend them to be taken literally. Explicit contracts conclude negotiations. This is very different from communicators in high-context cultures who depend less on language precision and legal documents. High-context business people may even distrust contracts and be offended by the lack of trust they suggest.

Day 1: July 17

Morning Session:
9am → 12pm

- Setting the stage:
  - Introduction and basic transistor operation
  - What to expect out of the transistor in 22nm and beyond.
  - Device matching considerations….in 22nm

Afternoon Session
1pm →5pm

- Building Block design in deep sub-micron
  - Bias, can it still be global?
  - Different topologies of bias circuitry
  - Bias stability and matching considerations…

Day 2: July 18

Morning Session:
9am → 12pm

- Diff amp, …
- Opamps….
- VCOs

Afternoon Session
1pm →5pm

- Noise…
- Analog layout…
- Yield considerations
  - PVT
  - Bias
  - How many sigma's?

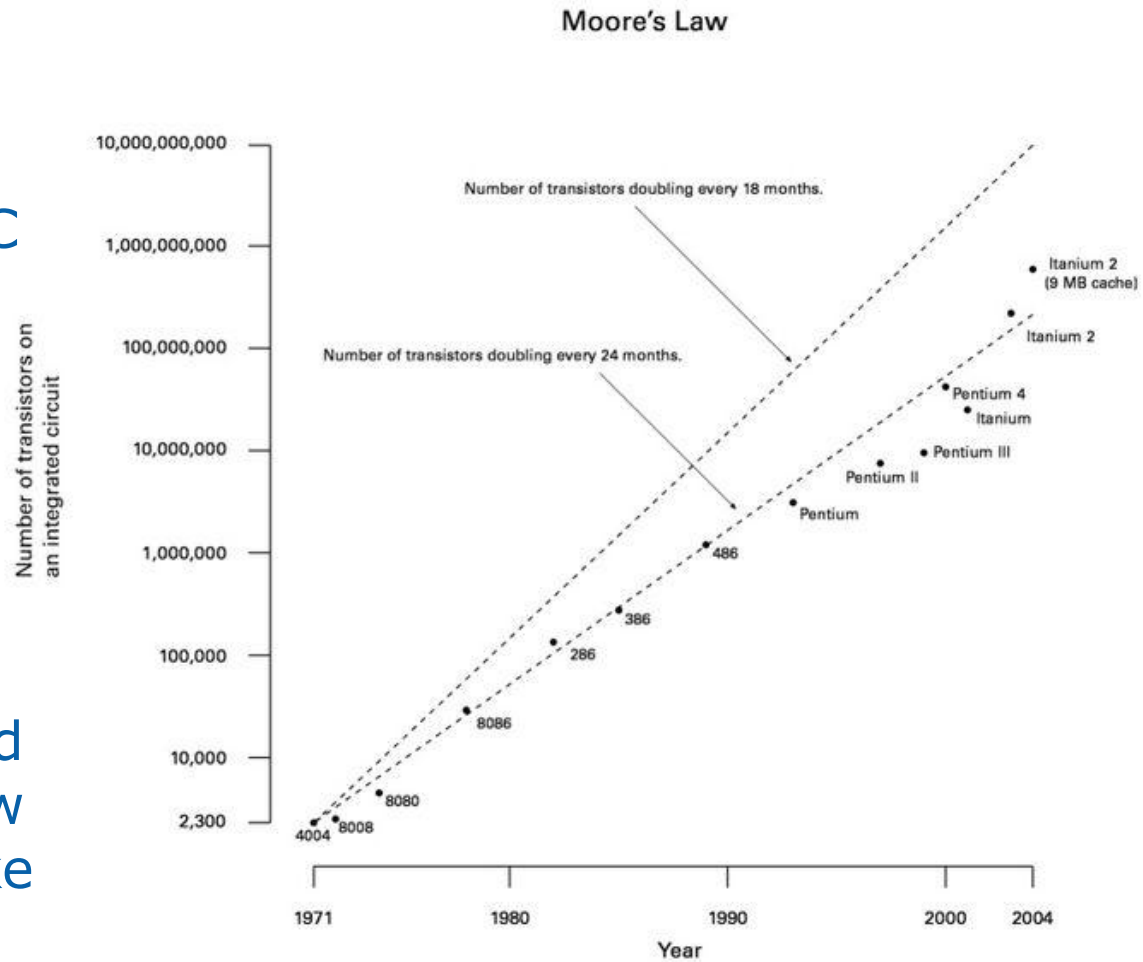
Technology Trend

- From Sand to Si
- http://www.youtube.com/watch?v=Q5paWn7bFg4&feature=related
- Area → form factor, mobility, hand held
- Power → cost of electric bill, battery life
- Speed, speed, speed
- Capacity
- User experience → apple
- Wave length
- ROI/wafer cost
- Idsat (Speed), Cox, gate oxide thickness, Supply voltage, threshold voltage, leakage

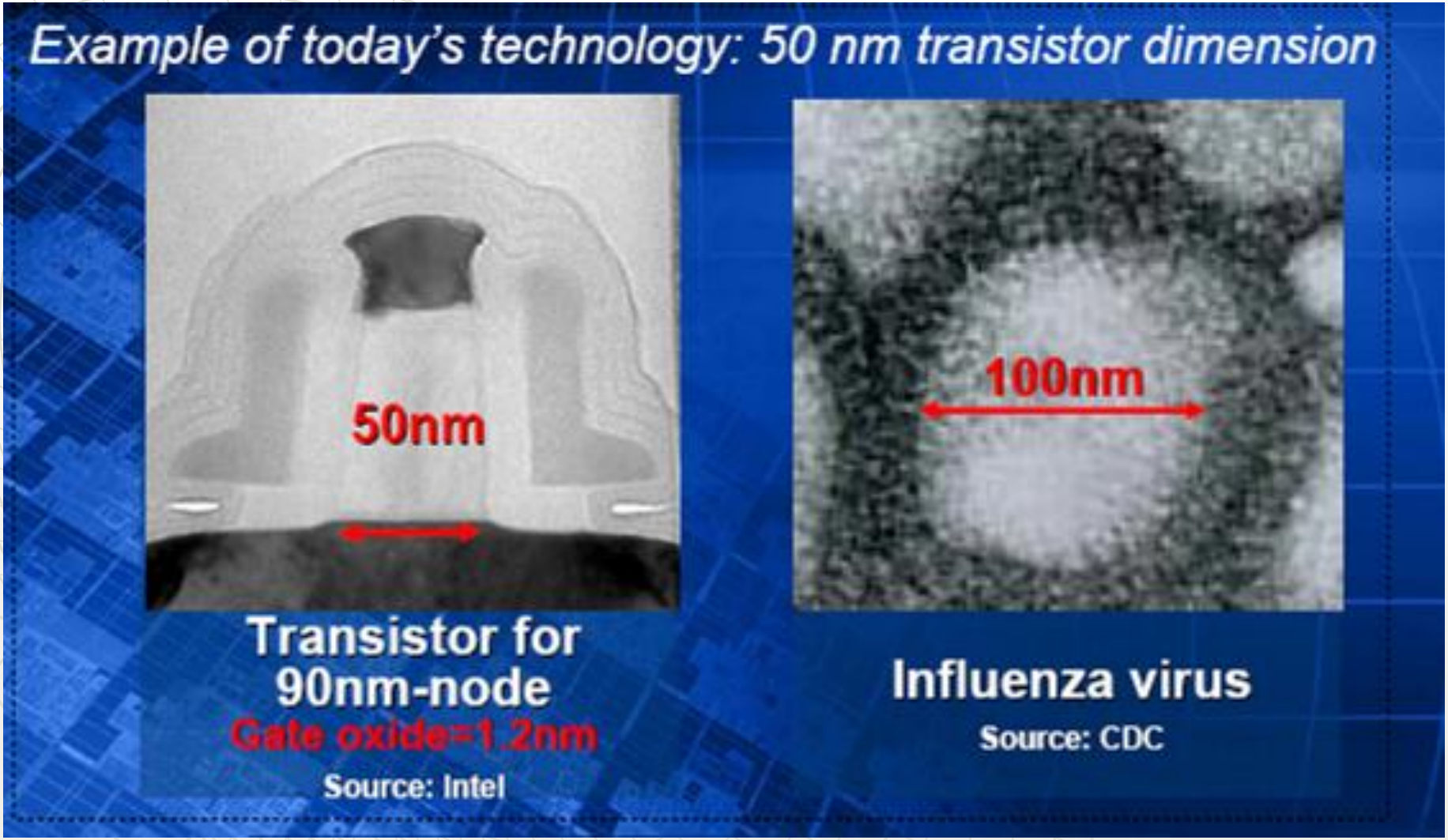
As technology feature size is reduced, the doping of the diffused regions is increased to avoid total pinch off effect. Higher doping leads to higher depletion capacitance therefore lower supply is needed to avoid junction break down

Technology Trend

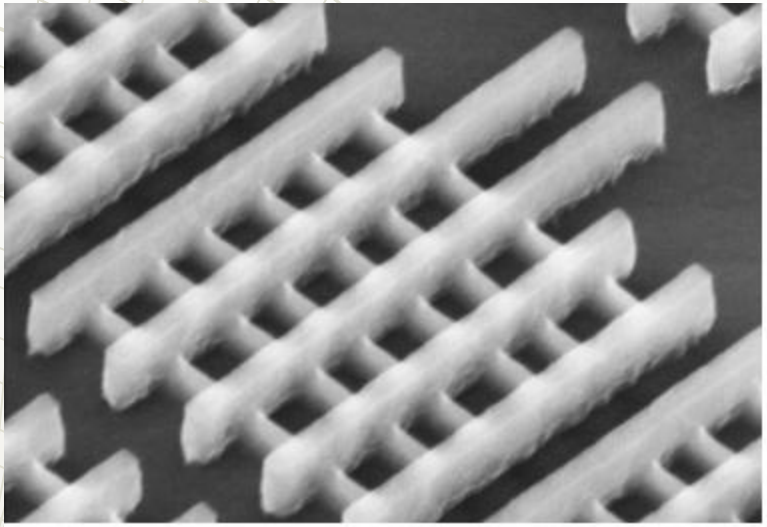
- Si development is still obeying Moore's law
- # of xtors/area on an IC doubles every 2 years
- If Moore's Law were applicable to the airline industry, a flight from New York to Paris in 1978 that cost $900 and took 7 hours, would now cost about $0.1 and take ~ 2 seconds.



The MOS

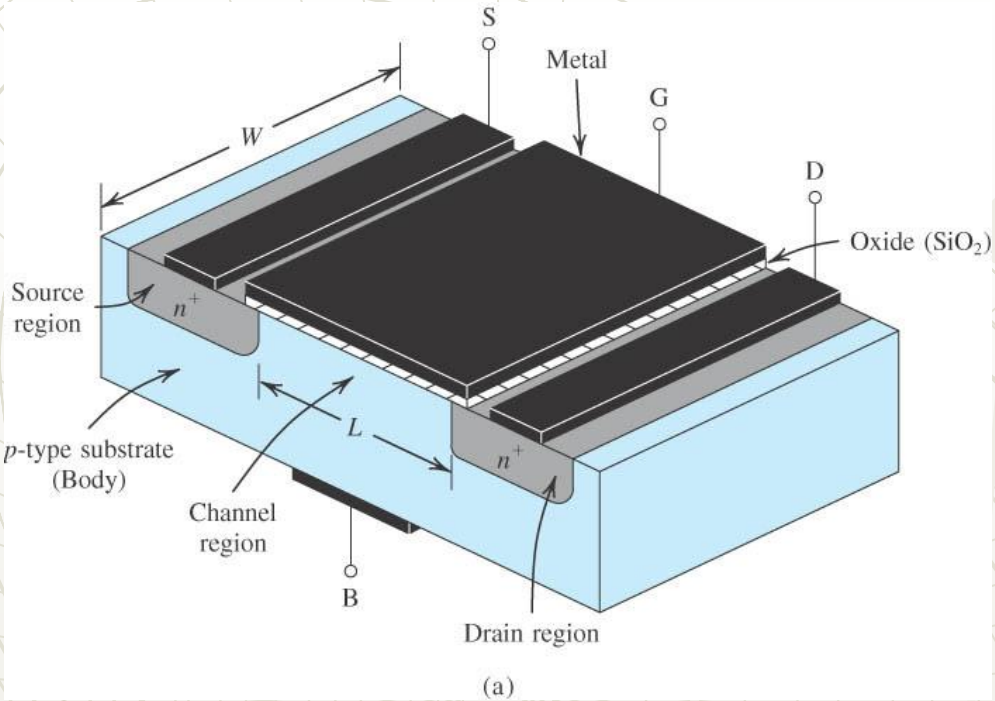

Example of today's technology: 50 nm transistor dimension
50nm
Transistor for
90nm-node
Gate oxide=1.2nm
Source: Intel
100nm
Influenza virus
Source: CDC

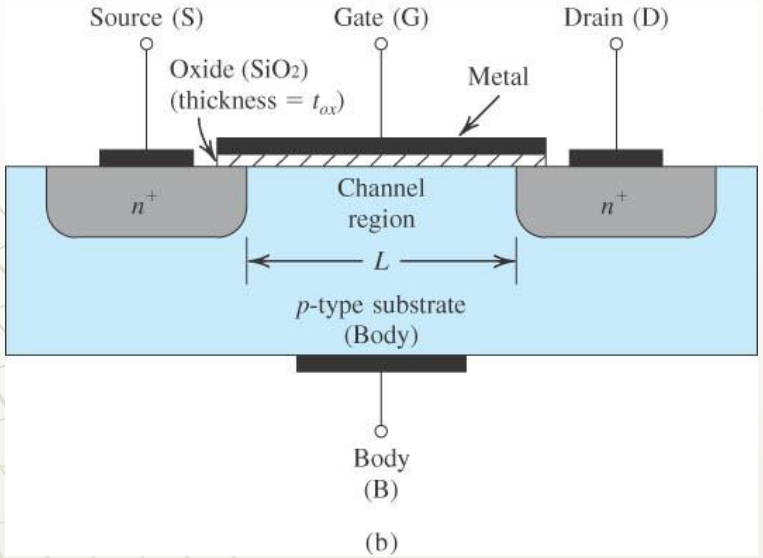
3D Transistors, why? What does it mean to consumers 22nm and beyond



A. Helmy, PhD

Back to Basics

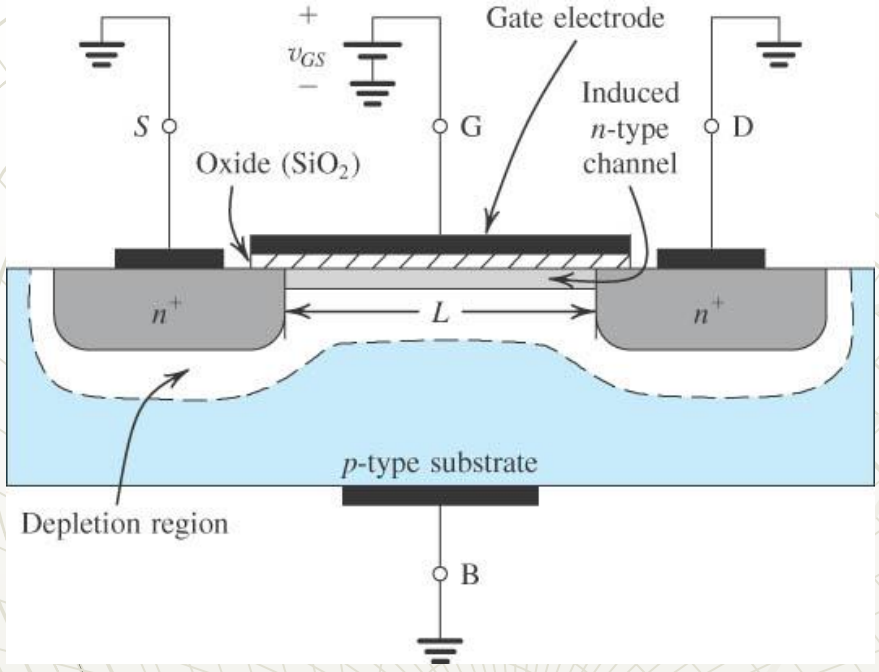


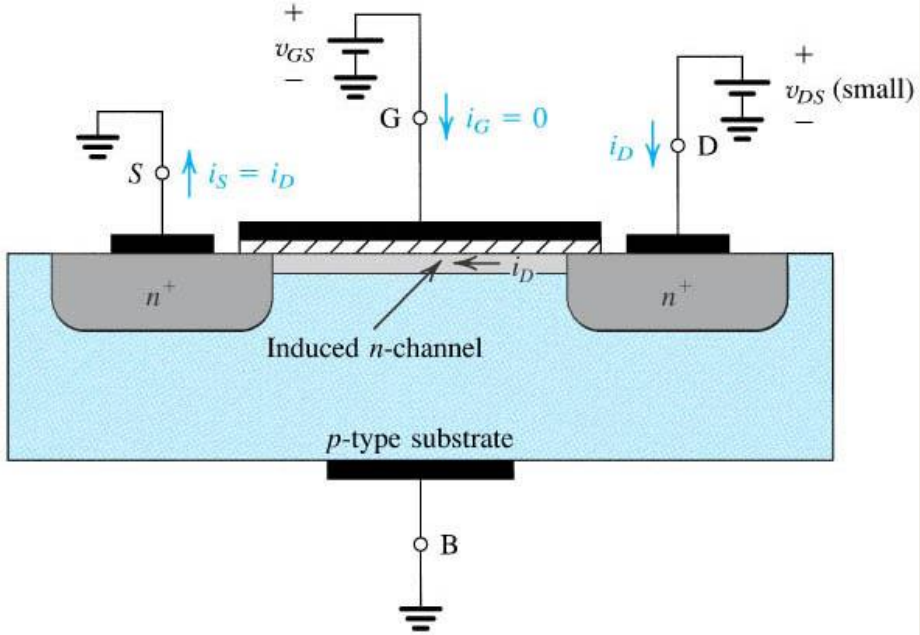


What can you change as a designer?

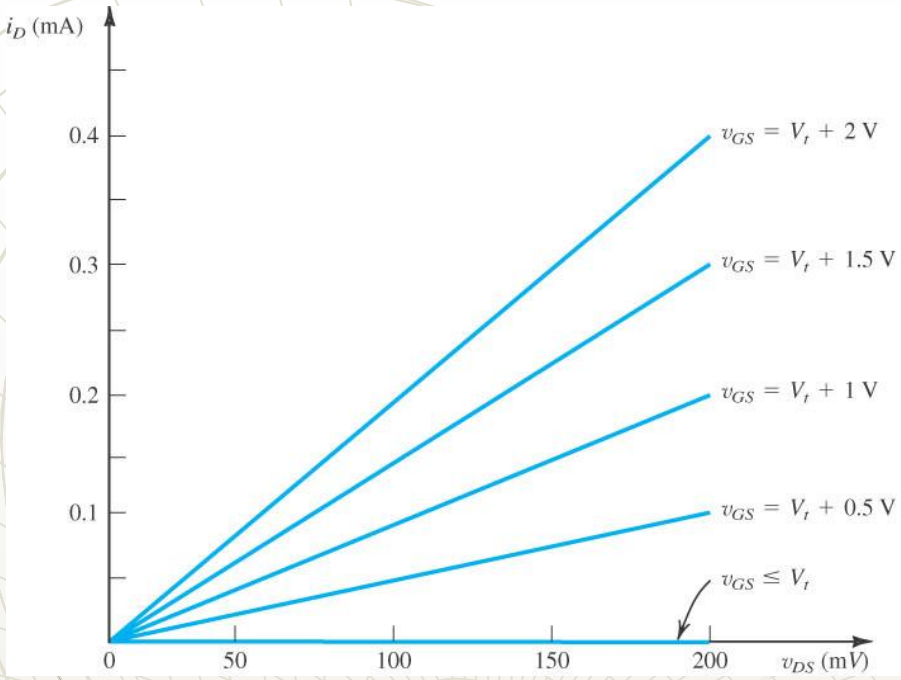
W (yes), L (?)

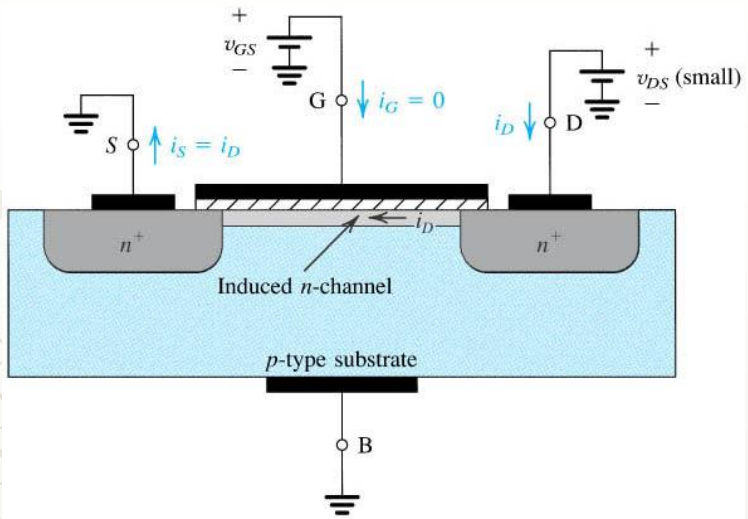
Channel Inversion





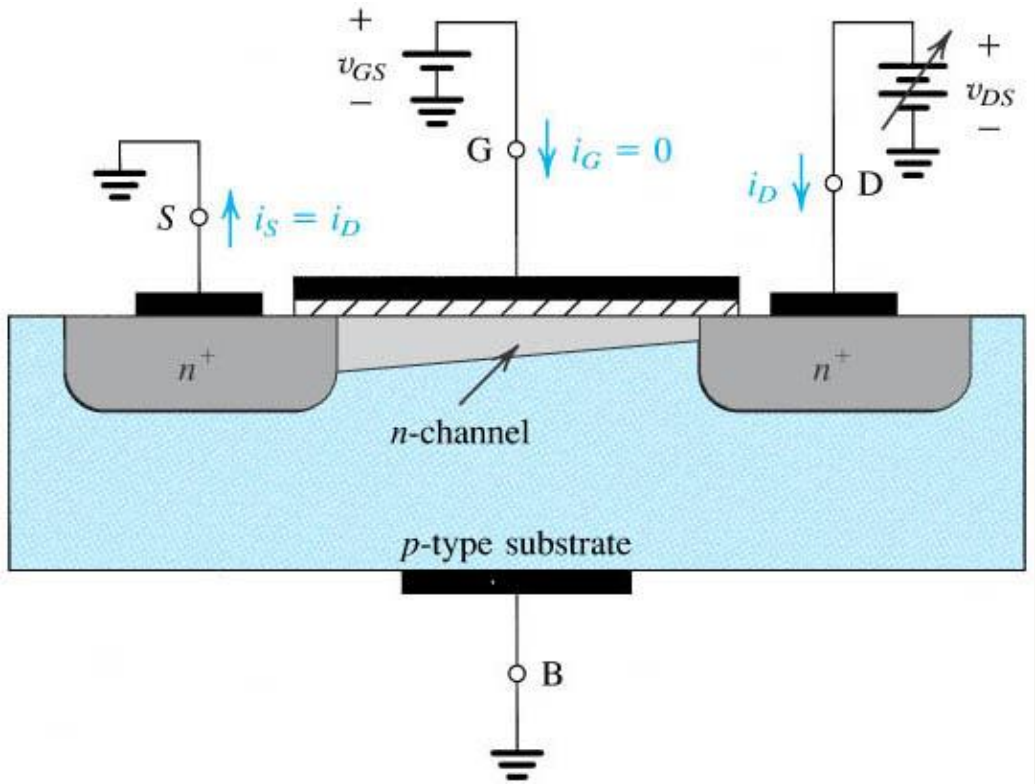
Channel Resistance





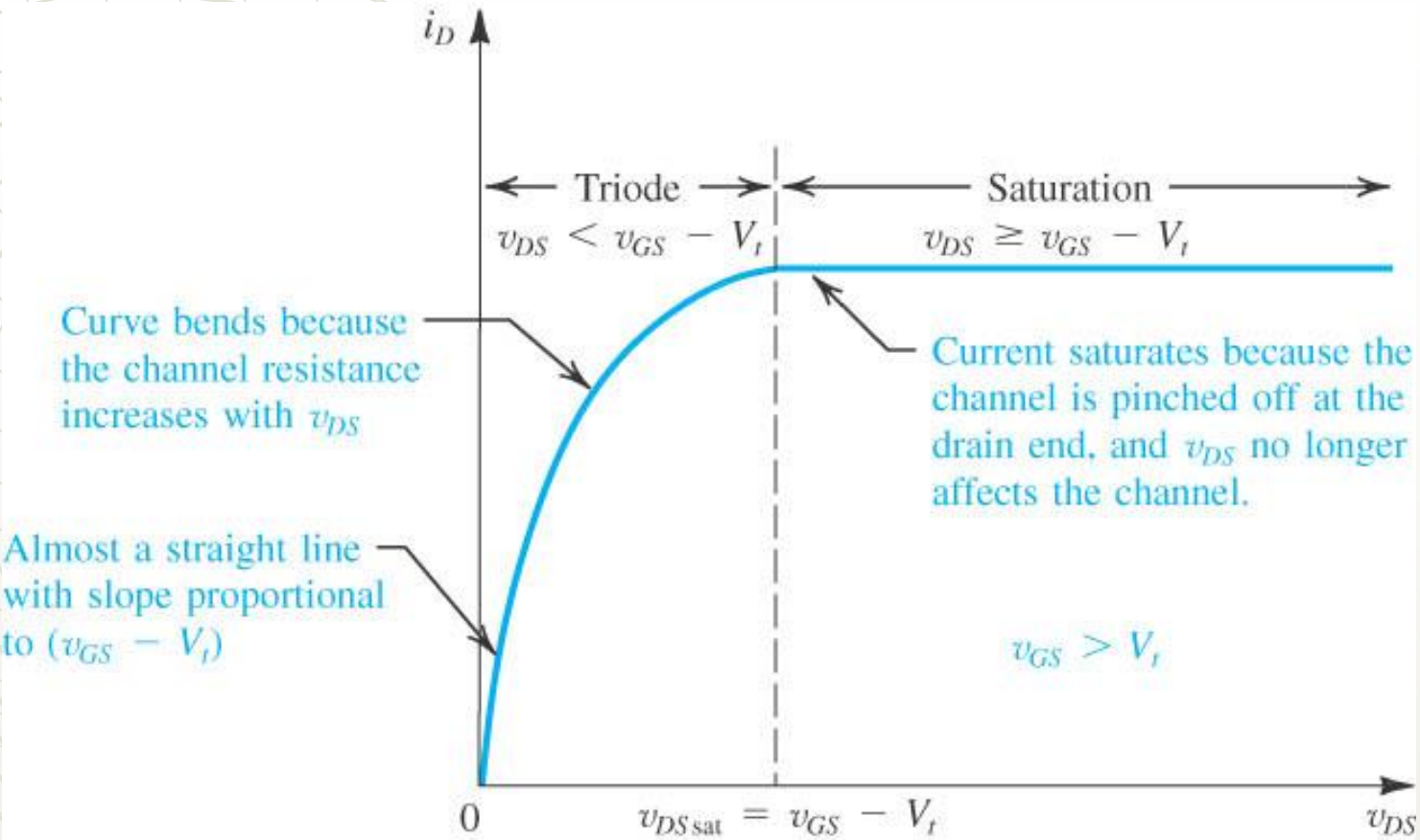
- What is V_{th}?
- $V_{gs} > V_{th}$ for channel to be inverted.

High Vds

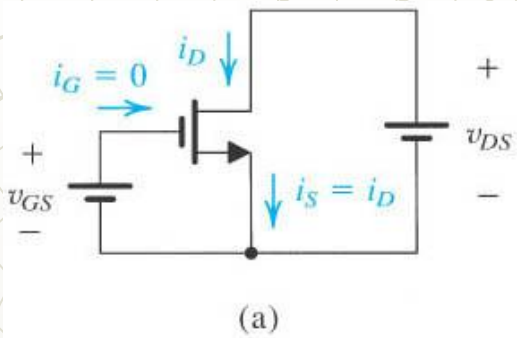


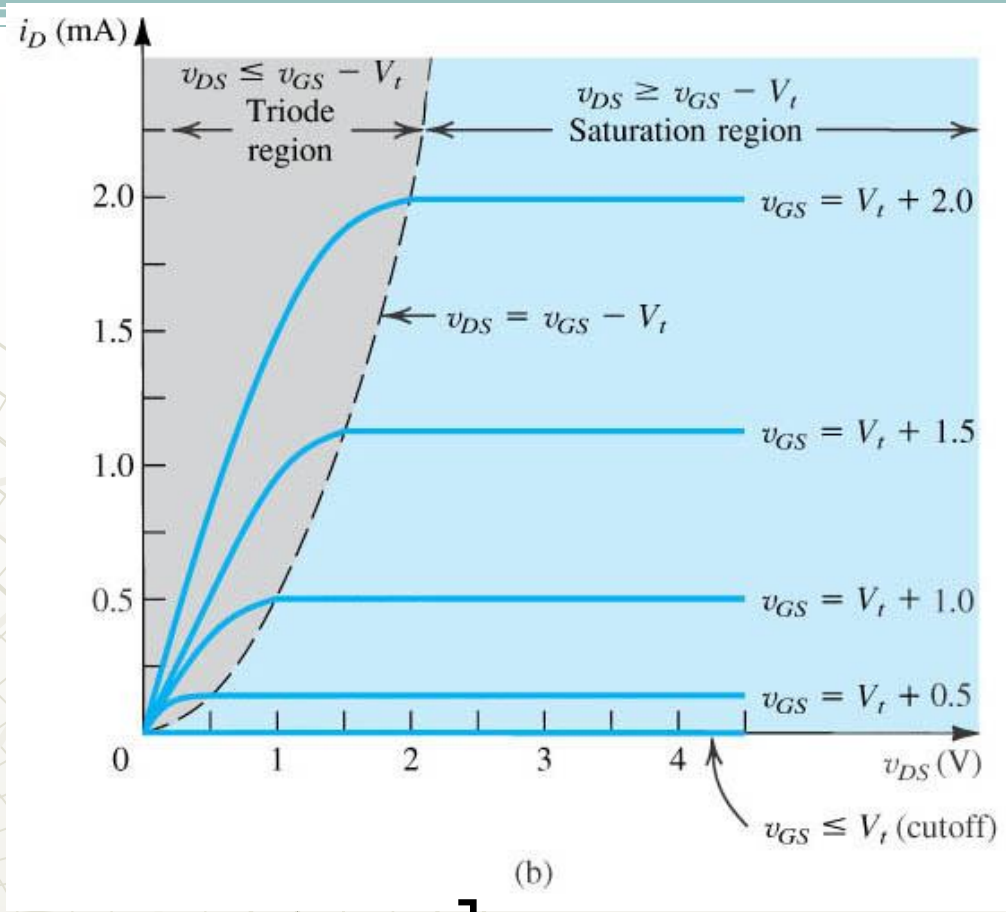
- As V_{ds} goes up, what happens to V_{gd}?
- What happens to the channel resistance?

Impact of High Vds and Pinch-off



NMOS I-V C/C





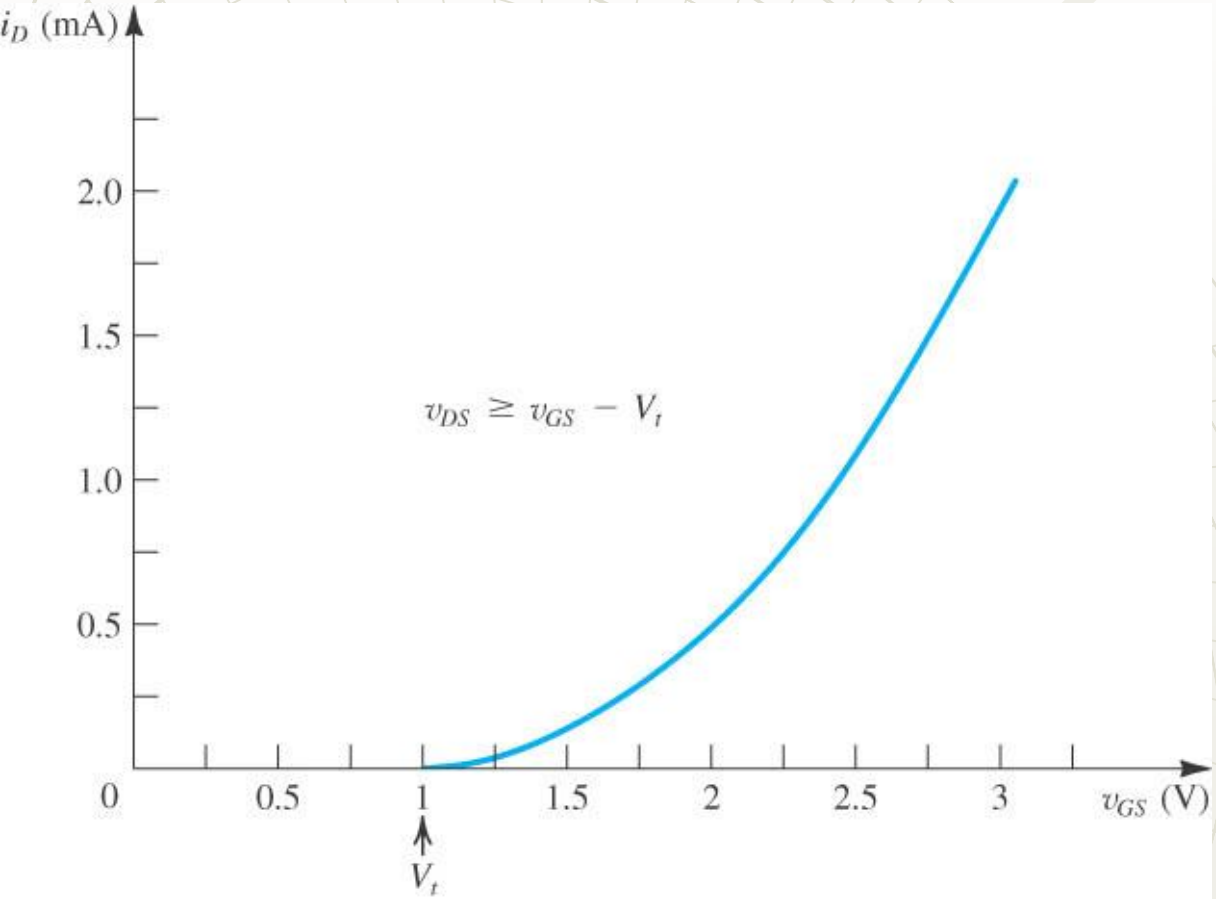
$$i_D = (\mu_n C_{ox})\left(\frac{W}{L}\right)\left[(v_{GS} - V_t)v_{DS} - \tfrac{1}{2}v_{DS}^2\right]$$ **Triode Region**

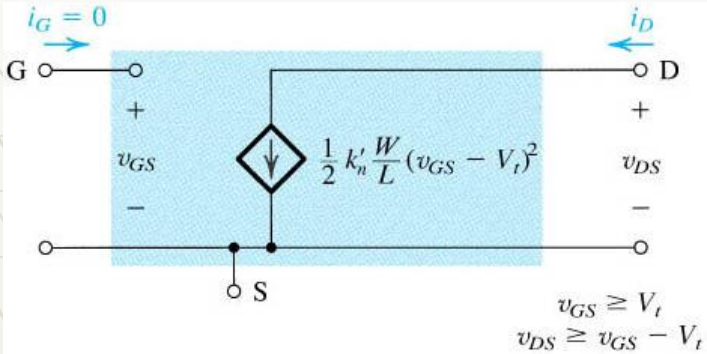
$$i_D = \tfrac{1}{2}(\mu_n C_{ox})\left(\frac{W}{L}\right)(v_{GS} - V_t)^2$$ **Saturation Region**

$i_D - v_{GS}$ Relation

$$i_D = \frac{1}{2}\left(\mu_n C_{ox}\right)\left(\frac{W}{L}\right)\left(v_{GS} - V_t\right)^2$$

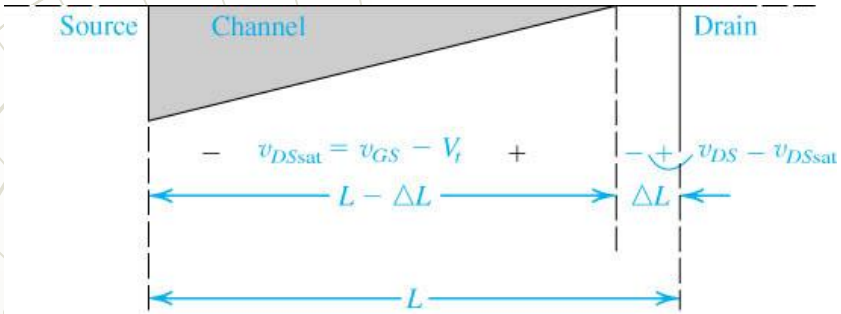
Saturation Region

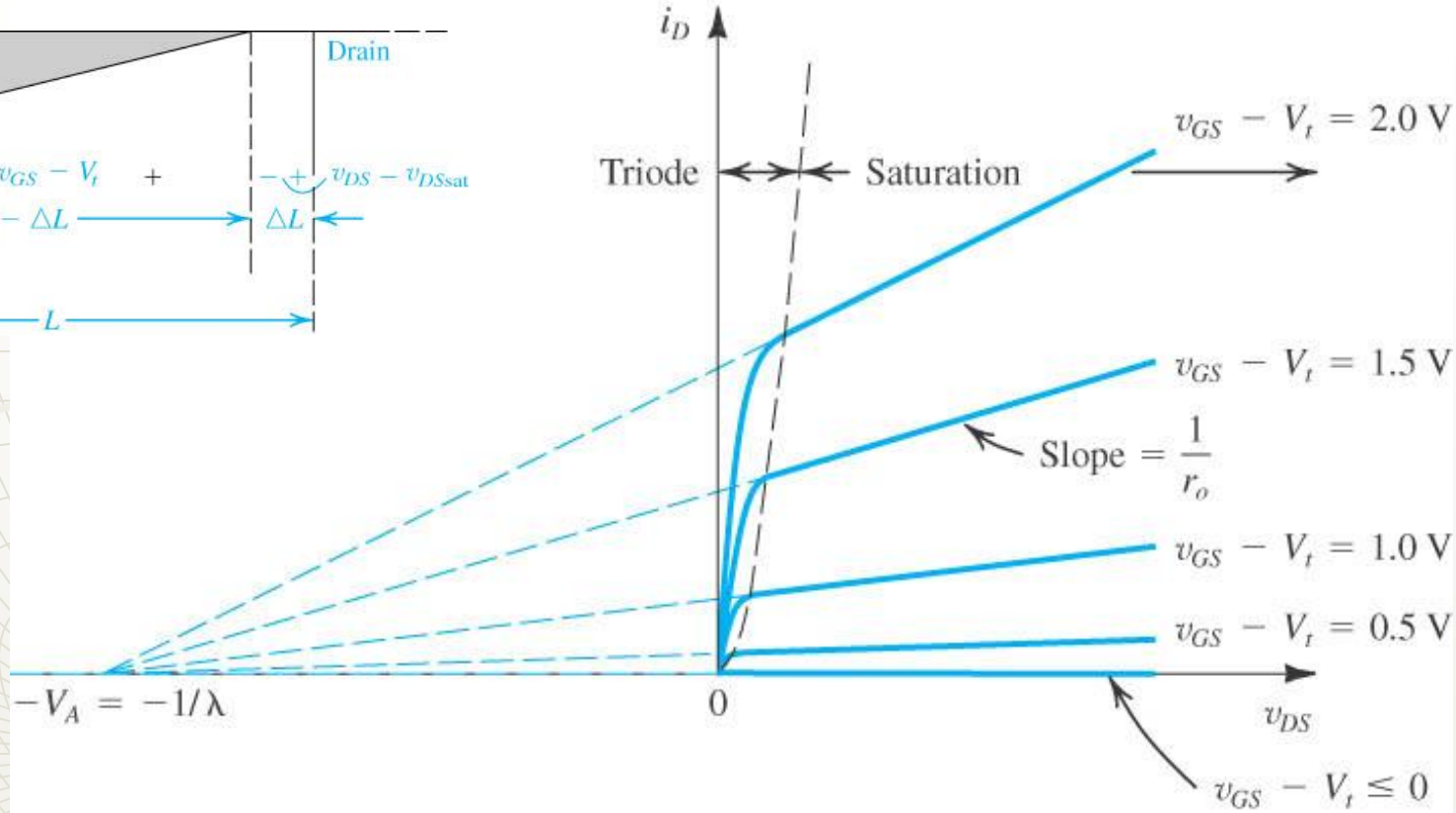




Output Resistance

Take a mental picture





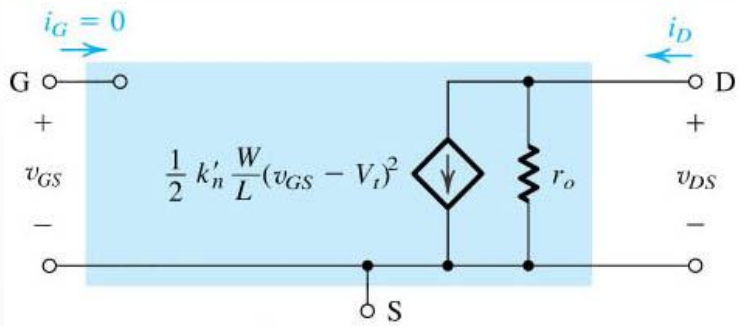
- Gate Length impact?
- Channel Length Modulation-Effect

Model with finite O/P Resistance

$$\lambda = \frac{\lambda'}{L}$$

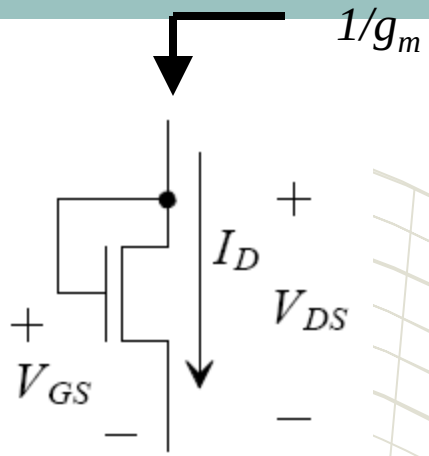
$$V_A = 1/\lambda$$

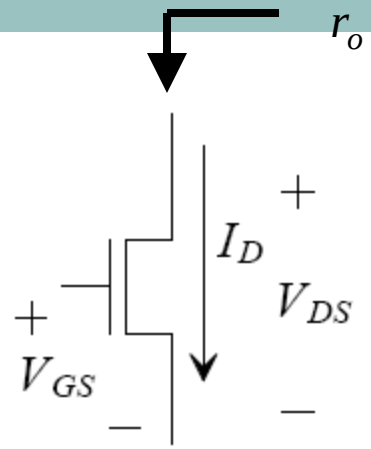
$$r_o = \frac{V_A}{I_D}$$

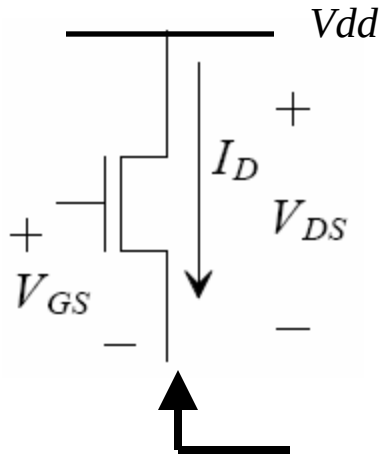


$$i_D = \tfrac{1}{2}(\mu_n C_{ox})\left(\frac{W}{L}\right)(v_{GS} - V_t)^2(1 + \lambda v_{DS})$$

1. Triode
2. Rout
3. Vds mismatch







$$V_{DS} > (V_{GS} - V_{THN}) \,\&\, V_{GS} = V_{DS}$$

$$0 > -V_{THN}$$

Always true, always in sat if I > 0

$$I_{DS} = g_m V_{GS} = gm\, V_{DS}$$

$$V_{DS}/I_{DS} = 1/g_m$$

$$1/g_m \,||\, 1/g_{mb} \,||\, r_o \approx 1/g_m$$

<u>Design Practices for Device Sizing and Biasing:</u>

For analog design, we use *L* of at least 2x minimum gate length (from 2x to 5x). This is required for good output resistance (gain) and matching ($\Delta V_{THN}\ \alpha\ Area^{-1}$). ***Trade off between speed and gain. (gain α r_o), (L α r_o)***

For $V_{DS,\,SAT}$ *the common practice is to choose around 5-10 % of VDD. For a 2.5V design,* $V_{DS,\,SAT}$ *~ 125mV* ***<u>gate overdrive</u>***

3. What happens to device parameters with temperature?

The parameters that change with temperature are threshold voltage and mobility

$$KP_n = \mu_n . C_{ox} = \mu_n . \frac{\varepsilon_{ox}}{t_{ox}} \text{ where } \mu_n \text{ is the mobility}, \varepsilon_{ox} \text{ is the oxide permitivitty and } t_{ox} \text{ is the oxide thickness}$$

threshold voltage vs. temperature

$$V_{THN} = -V_{ms} - 2V_{fb} + \frac{Q'_{b0} - Q'_{ss}}{C'_{ox}}, Q'_{ss} \text{ is a constant}$$

$$Q'_{b0} = \sqrt{2qN_A\varepsilon_{si}\left|-2V_{fp}\right|}, V_{fb} = -\frac{kT}{q}\ln\frac{N_A}{n_i}$$

$$V_{ms} = V_G - V_{fb} = \frac{kT}{q}\ln\frac{N_{D,poly}}{n_i} - V_{fb}$$

$$V_{THN} = -\frac{kT}{q}\ln\frac{N_{D,poly}}{n_i} + \frac{kT}{q}\ln\frac{N_A}{n_i} + \frac{\sqrt{2qN_A\varepsilon_{si}\left|-2V_{fp}\right|} - Q'_{ss}}{C'_{ox}}$$

$$\frac{\partial V_{THN}}{\partial T} = -\frac{k}{q}\ln\frac{N_{D,poly}}{N_A} + \frac{Q'_{b0}}{C'_{ox}.2T} \approx -\frac{k}{q}\ln\frac{N_{D,poly}}{N_A}$$

the term k/q is the change in the thermal voltage V_t with temperature

$$\frac{\partial V_t}{\partial T} = \frac{\partial}{\partial T}\left(\frac{kT}{q}\right) = \frac{k}{q} = 0.085mV/C^o$$

if $N_{D,poly} = 10^{20}$ and $N_A = 10^{15}$ an approx values for long-channel device

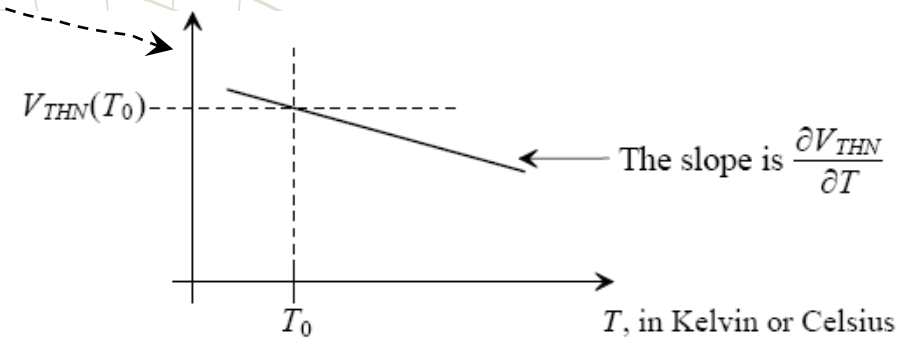
$\frac{\partial V_{THN}}{\partial T} = -1mV/C^o$ so the threshold voltage V_{THN} decreases with temp (more leakage)

if $N_{D,poly} = 10^{20}$ and $N_A = 10^{17}$ an approx values for short-channel device

$\frac{\partial V_{THN}}{\partial T} = -0.6mV/C^o$ why do we need to increase doping for short-channel?

$$TCV_{THN} = \frac{1}{V_{THN}} . \frac{\partial V_{THN}}{\partial T}$$, temperature coefficient of the threshold voltage

$$V_{THN}(T) = V_{THN}(T_o).(1 + TCV_{THN}.(T - T_o))$$



for short channel device $V_{THN} = 0.28V, TCV_{THN} = -2143\, ppm / C^o$

mobility vs. temperature

The reduction in mobility with increasing temperature is modeled as (not including non-linear variations in mobility)

$$\mu(T) = \mu(T_o).\left(\frac{T_o}{T}\right)^{1.5}$$, mobility

$$KP(T) = KP(T_0).\left(\frac{T_o}{T}\right)^{1.5}$$, transconductance parameter

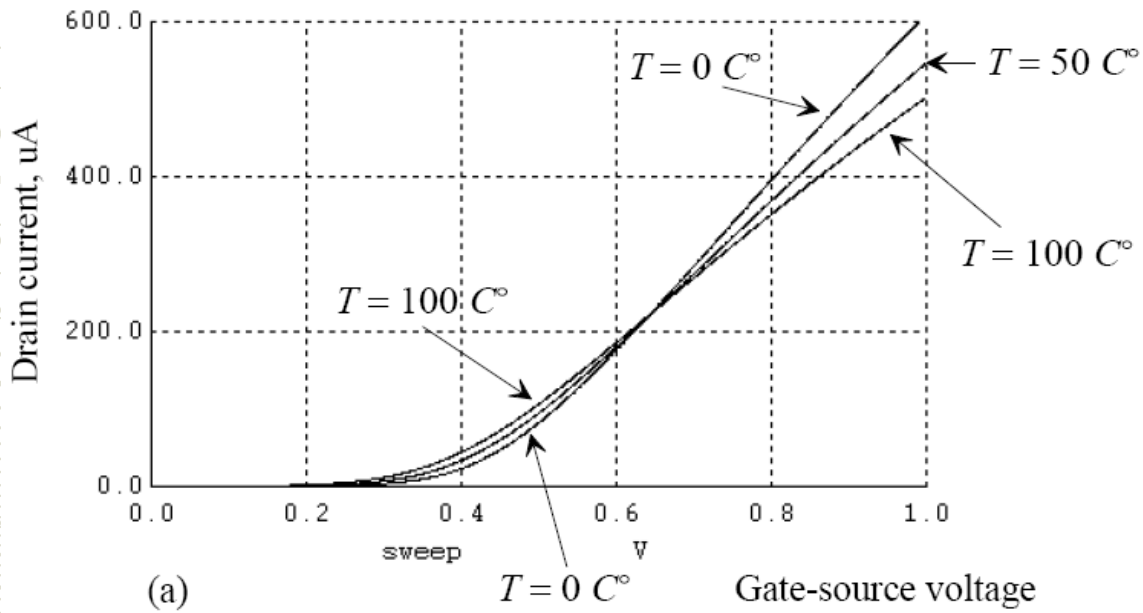
$$\frac{\partial KP(T)}{\partial T} = KP(T_0).(-1.5)\left(\frac{T_o}{T}\right)^{2.5} . \frac{1}{T_o}$$

$$TCKP = \frac{1}{KP(T)} . \frac{\partial KP(T)}{\partial T} = \frac{-1.5}{T}$$, temperature coefficient of the transconductance parameter

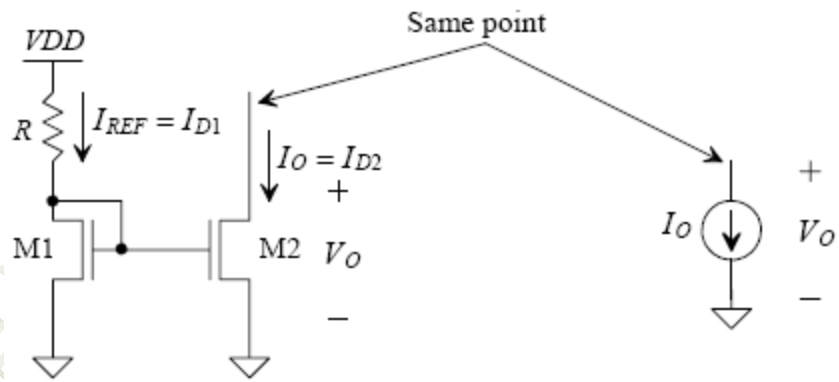
$$KP(T) = KP(T_0).(1 - 1.5.\frac{T - T_o}{T})$$

$$as\ T \uparrow \mu \downarrow I_D \downarrow$$ *Dominant at high* V_{GS}

$$as\ T \uparrow V_{TH} \downarrow I_D \uparrow$$ *Dominant at low* V_{GS}



4. Current Mirrors in CMOS Processes and Mismatch Mechanisms



$$I_{REF} = \frac{KP_n}{2}\left(\frac{W}{L}\right)_1 (V_{GS1} - V_{THN})^2 (1 + \lambda(V_{DS1} - V_{DS1,SAT})$$

$$V_{DS1} = V_{GS1} = V_{GS2} \text{ and } V_{DS1,SAT} = V_{DS2,SAT} = V_{GS1} - V_{THN}$$

$$I_o = \frac{KP_n}{2}\left(\frac{W}{L}\right)_2 (V_{GS1} - V_{THN})^2 (1 + \lambda(V_o - V_{DS1,SAT})$$

$$\frac{I_o}{I_{REF}} = \frac{(W/L)_2}{(W/L)_1} \cdot \frac{(1 + \lambda(V_o - V_{DS1,SAT})}{(1 + \lambda(V_{DS1} - V_{DS1,SAT})}$$

$$if\ \ L_1 = L_2 \text{ and } \lambda = 0 \text{ (long channel)}$$

$$\frac{I_o}{I_{REF}} = \frac{W_2}{W_1}$$

Mismatch Mechanisms in Current Mirrors:

1- V_{THN} mismatch

$$V_{THN1} = V_{THN} - \frac{\Delta V_{TH}}{2}$$

$$V_{THN2} = V_{THN} + \frac{\Delta V_{TH}}{2}$$

$$\frac{I_o}{I_{REF}} \approx 1 - \frac{2\Delta V_{TH}}{V_{DS,SAT}}$$

For better matching higher $V_{DS,SAT}$ is needed, i.e. higher V_{gs}-V_t

But, higher VDS,SAT causes the current mirror to enter triode early, current mirror range VDD - VDS,SAT

2- KP_n mismatch

$$KP_{n1} = KP_n - \frac{\Delta KP_n}{2}$$

$$KP_{n2} = KP_n + \frac{\Delta KP_n}{2}$$

$$\frac{I_o}{I_{REF}} \approx 1 + \frac{\Delta KP_n}{KP_n}$$

As geometry feature ↓

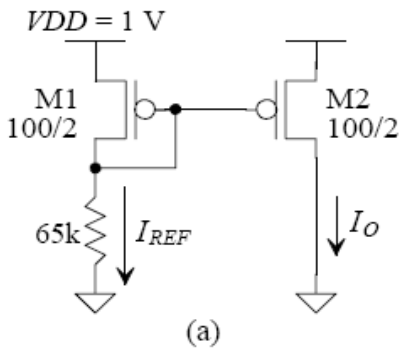
Cox ↑ and it is better controlled, 3 is the most dominant mismatch factor

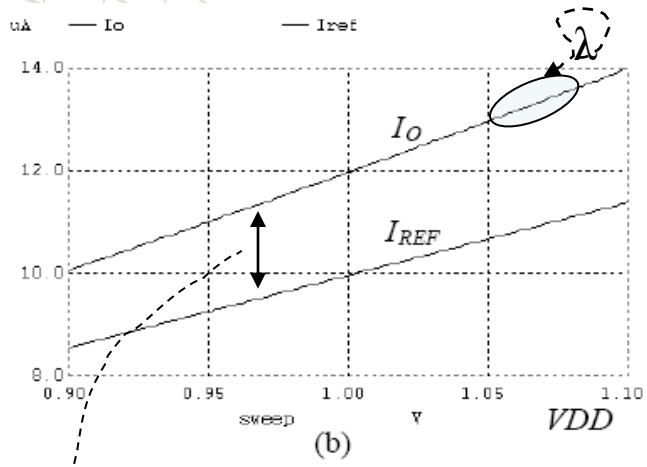
3- V_{DS} mismatch

$$\frac{I_o}{I_{REF}} \approx \frac{1 + \lambda_2 . V_o}{1 + \lambda_1 . V_{DS}}$$

Biasing Current Mirrors

Goal: Want a reference current that is independent on supply voltage and ground noise.





$$I_{REF} = \frac{V_{DD} - V_{SG}}{R}$$

$$\frac{\partial I_{REF}}{\partial V_{DD}} = \frac{1}{R}, \text{ assuming } V_{SG} \text{ is const}$$

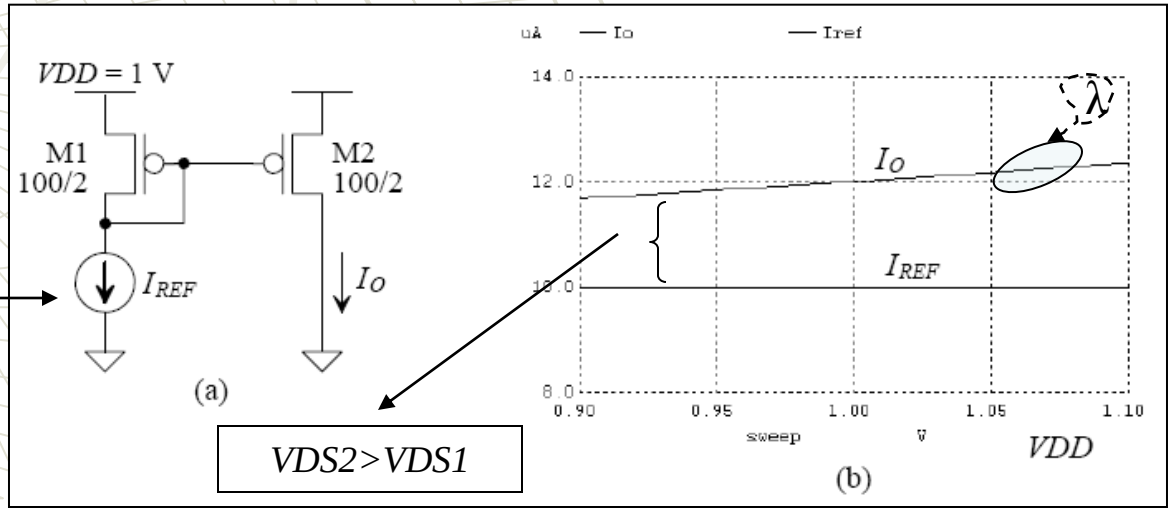
$$\frac{\partial I_{REF}}{\partial V_{DD}} = \frac{1}{65k} = 15.4\mu A/V$$

BAD

$$I_o = f(I_{REF}, V_{DS2}, \lambda)$$

Different V_{DS}

Replacing R in the drain side with an ideal current source (ideal MOS)



Using a MOSFET-only reference circuit

$$V_{DD} = V_{SG3} + V_{GS1}$$

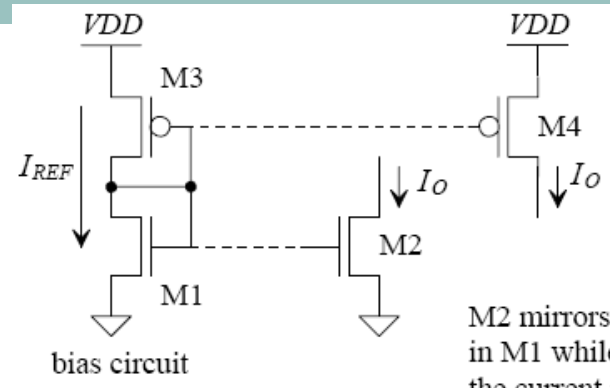
$$V_{DD} = \sqrt{\frac{2I_{REF}}{KP_p \frac{W_3}{L_3}}} + V_{THP} + \sqrt{\frac{2I_{REF}}{KP_n \frac{W_1}{L_1}}} + V_{THN}$$

Assume long channel $\lambda=0$

$$(V_{DD} - V_{THP} - V_{THN})^2 = I_{REF} \cdot \left[\sqrt{\frac{2L_3}{KP_p.W_3}} + \sqrt{\frac{2L_1}{KP_n.W_1}}\right]^2 = I_{REF}.const$$

$$\frac{\partial I_{REF}}{\partial V_{DD}} = 2(V_{DD} - V_{THP} - V_{THN})/const = 12\mu A/V$$

Still BAD



M2 mirrors the current in M1 while M4 mirrors the current in M3.

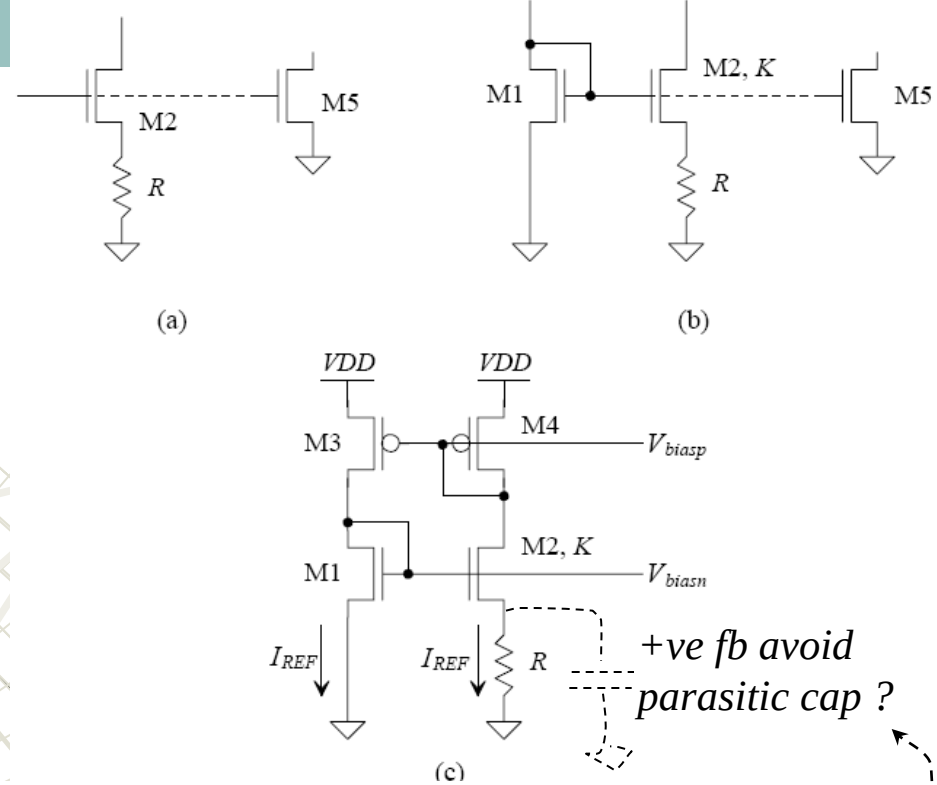
Replace the resistor bias with a MOSFET. Both M1 and M3 can be considered as an R but both are in the drain side of the other transistor

Next we will study the supply independent biasing and ***R connected to the source side of the current mirror instead of the drain side***

(a) Don't know the value of the mirrored current because VGS2, VGS5 are not obviously related

(b) M1 and M5 are mirrors, but how to force the same current in M1 and M2

(c) Addition of a PMOS current mirror to force same IREF in M1 and M2



+ve fb avoid parasitic cap ?

$$V_{GS1} = V_{GS2} + I_{REF}.R$$

V_{GS1} must be > V_{GS2}, so we make $(W/L)_2 > (W/L)_1$ so that the same current I_{REF} can pass at a lower V_{GS2}

ß multiplier ref circuit

$$\beta = KP_n.\frac{W}{L}$$

$$V_{GS} = \sqrt{\frac{2I_D}{\beta}} + V_{THN}$$

$$\beta_2 = K.\beta_1, \text{ or } W_2 = K.W_1$$

ß multiplier ref circuit

$$I_{REF} = I_2 = \frac{V_{GS1} - V_{GS2}}{R} = \frac{1}{R}\left(\sqrt{\frac{2I_{REF}}{KP_n\frac{W_1}{L_1}}} - \sqrt{\frac{2I_{REF}}{K.KP_n\frac{W_1}{L_1}}}\right)^2$$

$$I_{REF} = \frac{2}{R^2 KP_n\frac{W_1}{L_1}}\left(1 - \frac{1}{\sqrt{K}}\right)^2 \neq f(V_{DD})$$

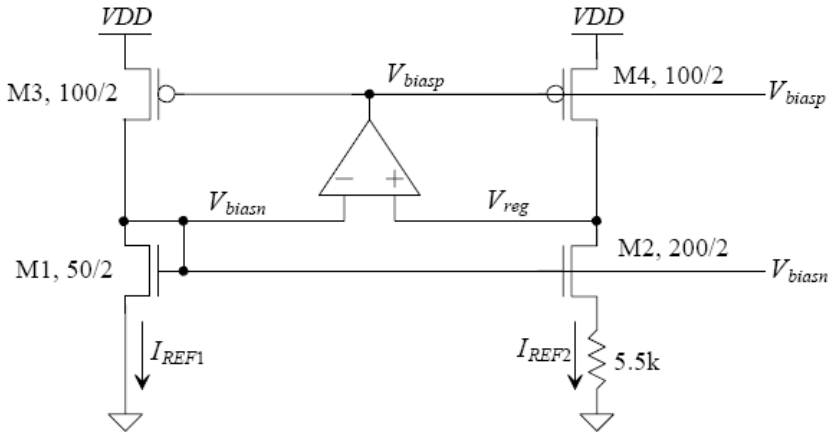
$$V_{DS,SAT} = V_{GS} - V_{THN} = \sqrt{\frac{2I_{REF}}{KP_n\frac{W_1}{L_1}}} = \frac{2}{R.KP_n.\frac{W_1}{L_1}}\left(1 - \frac{1}{\sqrt{K}}\right)$$

for $K = 4$

$$g_m = KP_n.\frac{W}{L}.V_{DS,SAT} = \sqrt{2KP_n.\frac{W}{L}.I_{REF}} = \frac{1}{R}$$

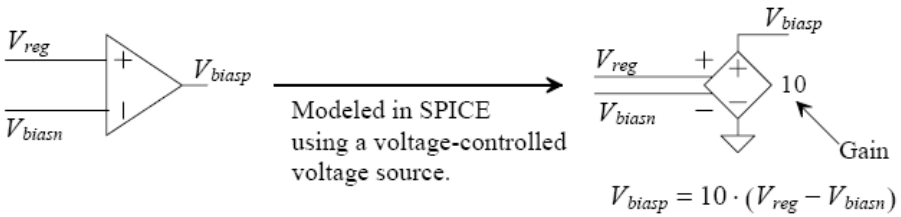
Constant gm, if R is external, gm becomes process independent but watch for instability

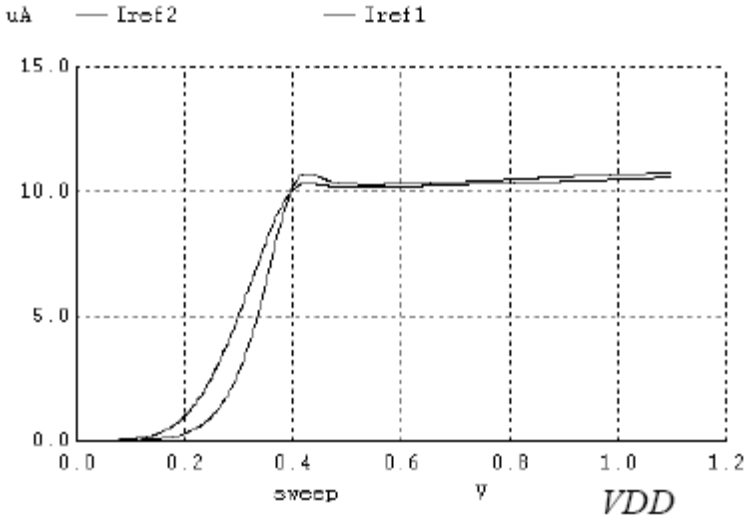
A. Helmy, PhD



M4 is not diode connected any more so its drain voltage can follow M3 drain voltage.

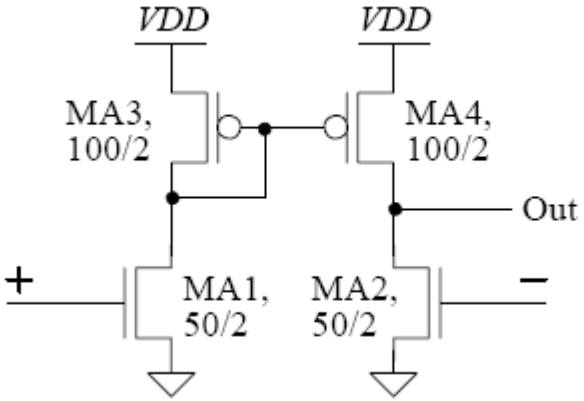
If Vreg is high, Vbiasp is high and I in M4 goes down and hence lowering Vreg, also M3 current goes low and Vbiasn goes low to follow Vreg





The amplifier tries to equate V_{reg} and V_{biasn} *making it independent of* V_{DD}

Typical implementation of regulation op-amp



Cascoding current mirror

- *Vds variation is the biggest cause of current mismatch*
- *To have a constant current Vds must be constant, but this will force a constant Vo*
- *If we increase the output resistance this will make the current more constant with VDS specially in short channel process nodes, solution is (cascoding) but* ***headroom??!!***

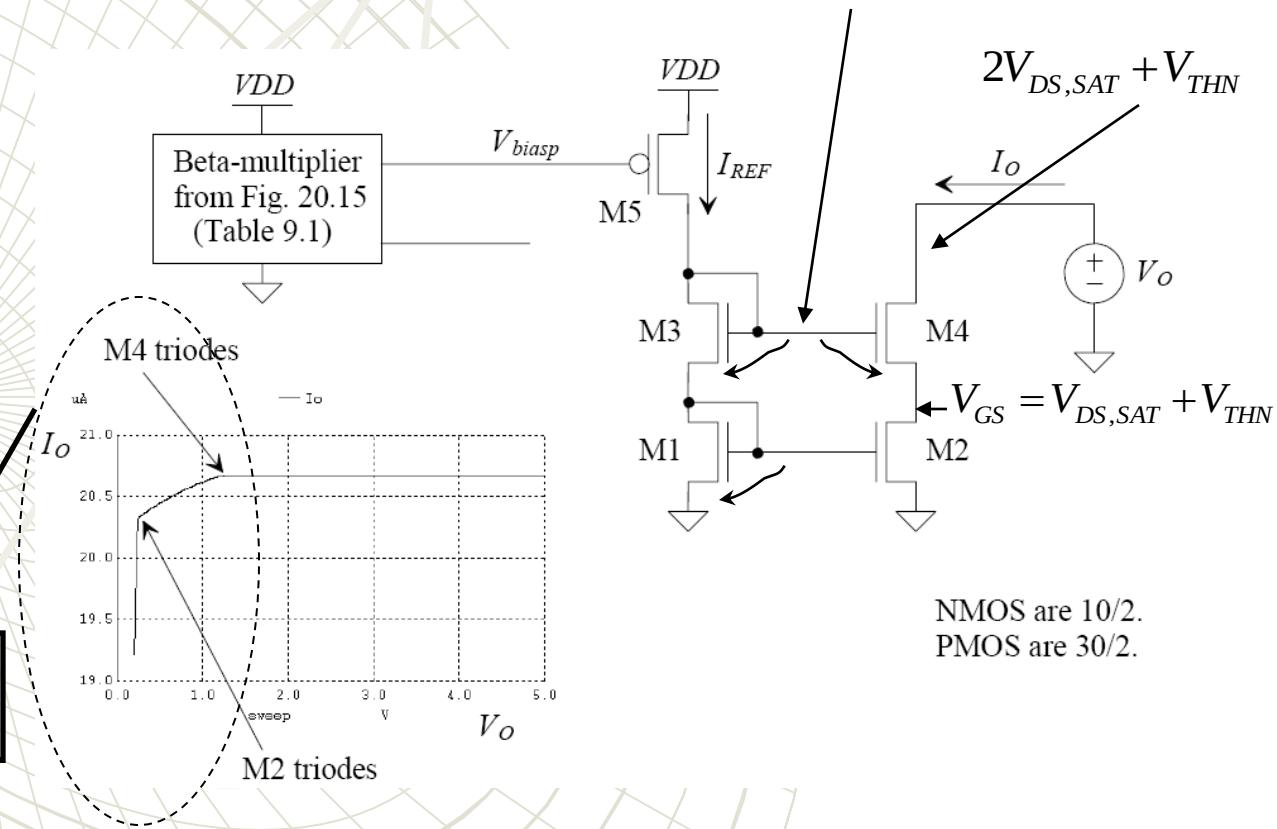
I.Q

$$V_{DS,SAT} = V_{GS} - V_{THN}$$

$$2V_{GS} = 2(V_{DS,SAT} + V_{THN})$$

Vo,min is M4 drain voltage below which M4,M2 triodes

*This is not the minimum Vo that can be achieved because (absolute min happens when there is 1*$V_{DS,SAT}$ on each of M2 and M4, i.e 2 $V_{DS,SAT}$*

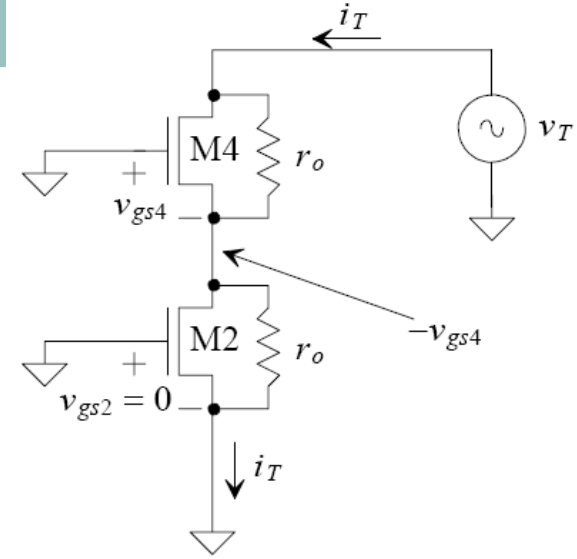


Notice how cascode eats from your voltage span "headroom"

Lets check the cascode output ss resistance

Cascode XTR

Rail XTR



$$R_o = \frac{v_T}{i_T}$$

$$-v_{gs4} = i_T . r_{o2}$$

$$i_T = g_{m4} v_{gs4} + \frac{v_T - (-v_{gs4})}{r_{o4}}$$

Assume $r_{o2} = r_{o4}$

$$i_T = -g_{m4} i_T . r_{o2} + \frac{v_T}{r_{o4}} - i_T$$

imp

$$R_o = \frac{v_T}{i_T} = (2 + g_{m4} r_{o2}) r_{o4} \approx g_{m4} r_{o4} r_{o2} \approx g_m r_o^2 = A_v . r_o$$

Self gain of M_4, $g_{m4} r_{o4}$, *multiplies* r_{o2} *of* M_2

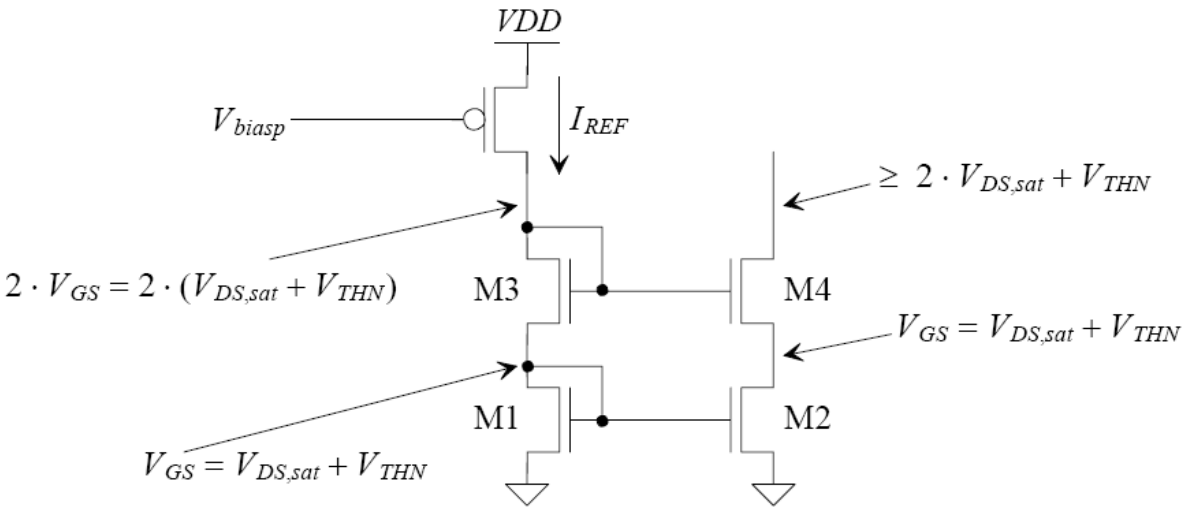
For small signal put all DC gate voltages =0

$$g_m r_o^2 >> r_o$$

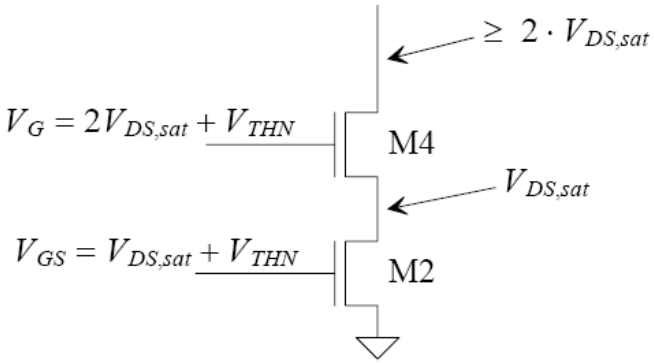
Cascoding increases output resistance compared to a single transistor but is it a good solution for short channel devices in modern process nodes where Vdd is scaled down and your headroom is limited??!!

Need a wide swing cascode

Wide swing cascode current mirror



(a) Regular cascode structure



(b) Low-voltage (aka wide-swing) structure

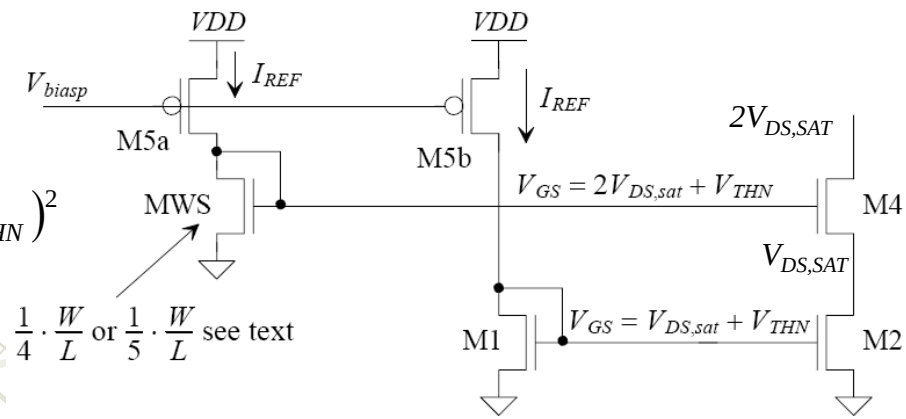
Wide swing cascode current mirror

$$I_{REF} = \frac{KP_n}{2} \cdot \frac{W_1}{L_1} (V_{GS} - V_{THN})^2$$

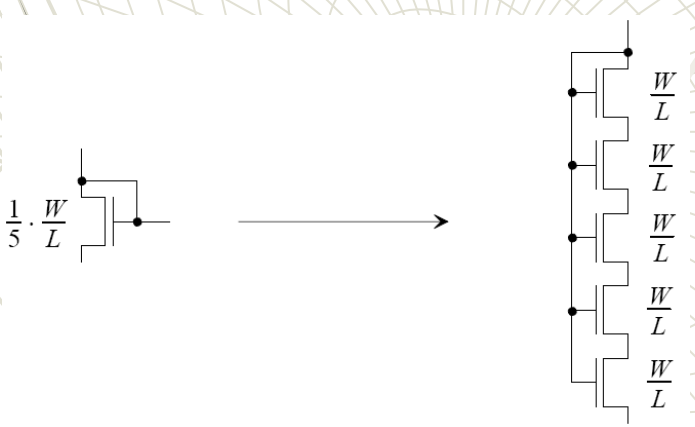
$$I_{REF} = \frac{KP_n}{2} \cdot \frac{W_{MWS}}{L_{MWS}} (2V_{DS,SAT})^2 = \frac{KP_n}{2} \cdot \frac{W_{MWS}}{L_{MWS}} .4.(V_{GS} - V_{THN})^2$$

$$\frac{W_1}{L_1} = \frac{W_{MWS}}{L_{MWS}} .4 \quad \text{if } W_1 = W_{MWS} \text{ then}$$

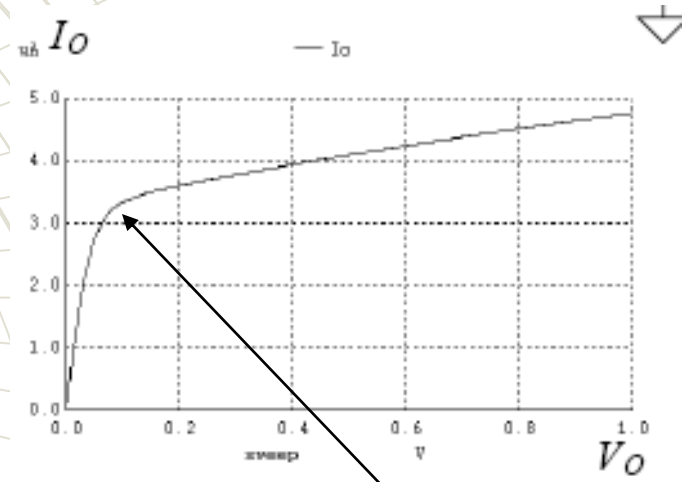
$$4L_1 = L_{MWS}$$



If we increase LMWS to $5L_1$, min voltage will be > 2VDS,SAT



I.Q

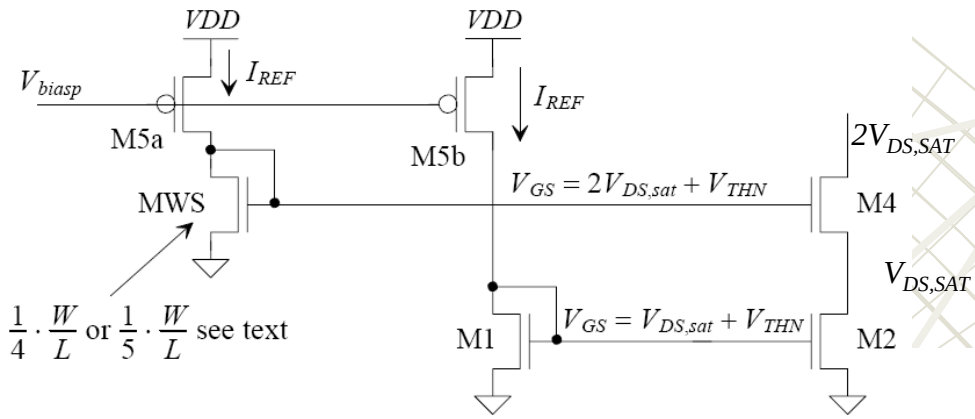


2VDS,sat

Better matching as Vt = f(L)

However, transistors M1 and M2 still do not have a perfectly matched V_{DS}

M1 "sees" M5b on its drain, M2 "sees" M4 source. A better design practice is to match the rail devices better.



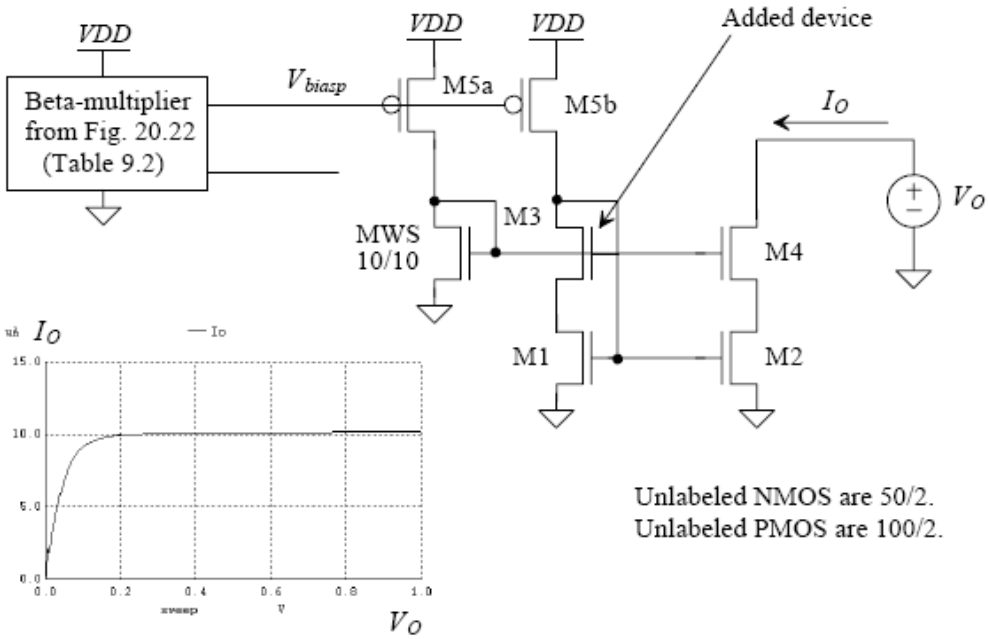


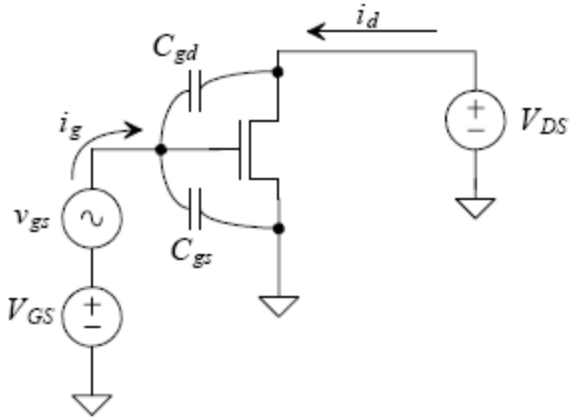
Figure 20.38 Wide-swing cascode current source in the short-channel process (good). Notice the drain-to-source voltages of M1 and M2 are the same.

How fast a MOS is? transition frequency f_T

for small signal analysis $V_{DS} = 0$, C_{gs} *and* C_{gd} *appear in //*

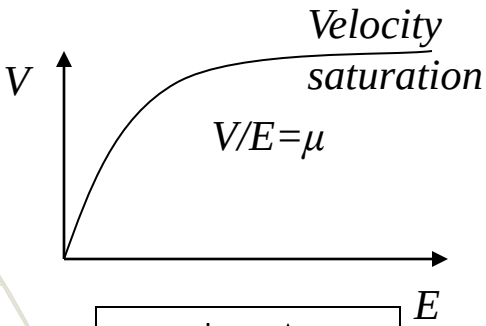
$$v_{gs} = \frac{i_g}{j.w.(C_{gs} + C_{gd})}, \; i_d = g_m.v_{gs}$$

$$\left|\frac{i_d}{i_g}\right| = \frac{g_m}{2\pi f(C_{gs} + C_{gd})}$$



MOS transition frequency f_T *is the frequency at which its current gain is unity* $H_{21}=1$

Given that $C_{gs} = \frac{2}{3}.W.L.C_{ox} >> C_{gd}$ *in saturation* *ch9*



$$f_T = \frac{g_m}{2\pi C_{gs}} = \frac{3KP_n}{4\pi L^2 C_{ox}}(V_{GS} - V_{THN}) = \frac{3\mu_n}{4\pi}.\frac{V_{DS,SAT}}{L^2} \; \alpha \; \sqrt{I_D}$$

for short channel devices, due to velocity saturation, with decreasing length μ/L is relatively constant; yielding:

$$f_T \; \alpha \; \frac{V_{DS,SAT}}{L}$$

As $L\downarrow$ $E\uparrow$

Hence $\mu\downarrow$ *and*

$\mu/L \sim const$

$$g_m = KP_n.\frac{W}{L}V_{DS,SAT}$$

$$g_m = \sqrt{2.KP_n.\frac{W}{L}.I_D}$$

A Figure of merit (FOM) for comparing processes:

Used to compare transistors with different (W/L)s "the gain-f_T (GFT) product"

Transistor self DC gain

$$GFT = g_m . r_o . f_T$$

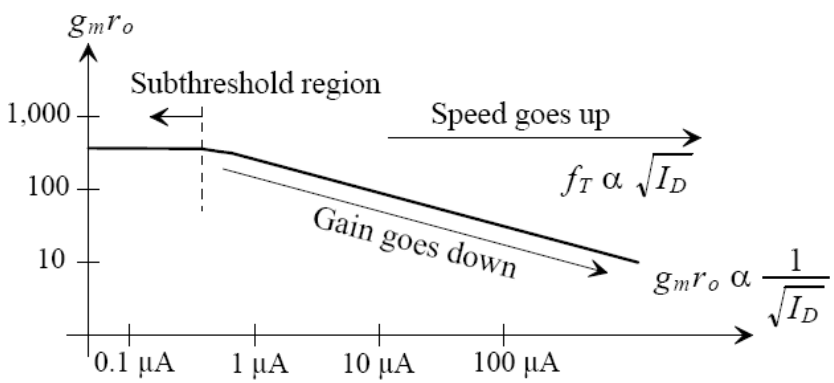
$$GFT = g_m . r_o . f_T = \frac{g_m^2}{2\pi C_{gs}} . \frac{1}{\lambda I_D} = \frac{3\mu_n}{2\pi\lambda L^2} \neq f(I_D)$$

Constant for same L independent of bias nor current

MOS self gain

$$g_m . r_o = \frac{\sqrt{2KP_n \frac{W}{L} . I_D}}{\lambda I_D} = \frac{\sqrt{2KP_n \frac{W}{L}}}{\lambda} . \frac{1}{\sqrt{I_D}} = \frac{2}{\lambda(V_{GS} - V_{THN})} = \frac{2}{\lambda V_{DS,SAT}}$$

This shows that the DC gain of a MOSFET decreases with I_D while speed increases with I_D



For high f_T, use min L and high $V_{DS,SAT}$
But low L has low r_o and low gain thus GBW ≈ const,
Also at large $V_{DS,SAT}$ the MOS triodes
early hence produces low swing
$f_{Tn} > f_{Tp}$ as $\mu_n > \mu_p$

I.Q

6 sigma Understanding Variation

If everything were predictable, human beings, who have pretty good brains, will undoubtedly learn to predict everything,

→ and they will thereupon have no motive to do anything at all,

because they will recognize that the future is totally determined and cannot be influenced by any human action.

On the other hand, if everything were unpredictable,

they will gradually discover that there is no rational basis for any decision whatsoever and, as in the first case, they will thereupon have no motive to do anything at all.

Neither scheme would make sense. The world is a mixture of the two.

Some things are predictable and others be unpredictable.

They will then, amongst many other things, have the very important task of finding out which is which.

Variation is inherent to a process……

Understanding Variation

	X	$X-\bar{X}$	$(X-\bar{X})^2$
1	130	–20	400
2	166	16	256
3	178	28	784
4	131	–19	361
5	140	–10	100
6	125	–25	625
7	127	–23	529
8	145	–5	25
9	110	–40	1600
10	184	34	1156
11	161	11	121
12	194	44	1936
13	171	21	441
14	125	–25	625
15	163	13	169
Sum	2250		9128
Mean	150		
S^2			652
S^1			25.53

$$\sum^{n}(X_i - \bar{X})^2$$

$$\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}$$

$$\sqrt{\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}}$$

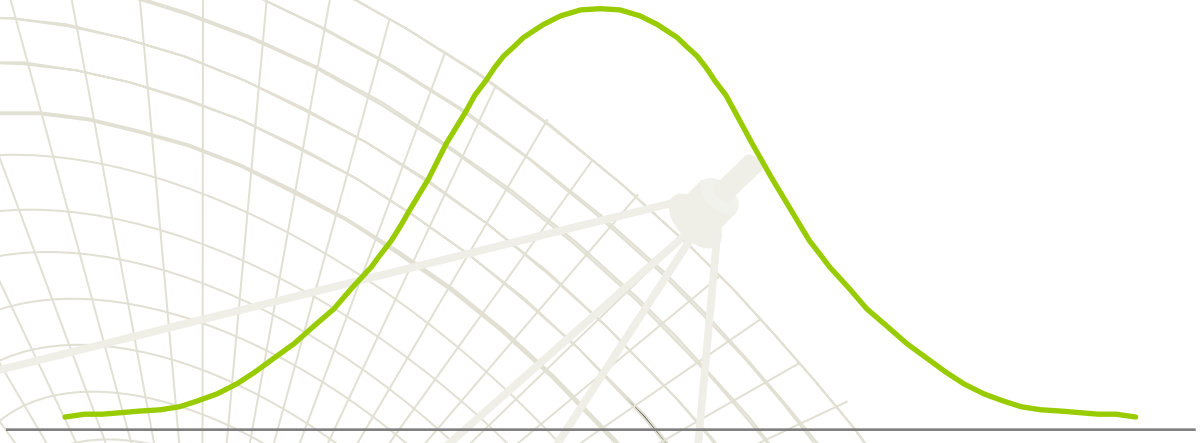
Sample Mean

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

Sample Standard Deviation

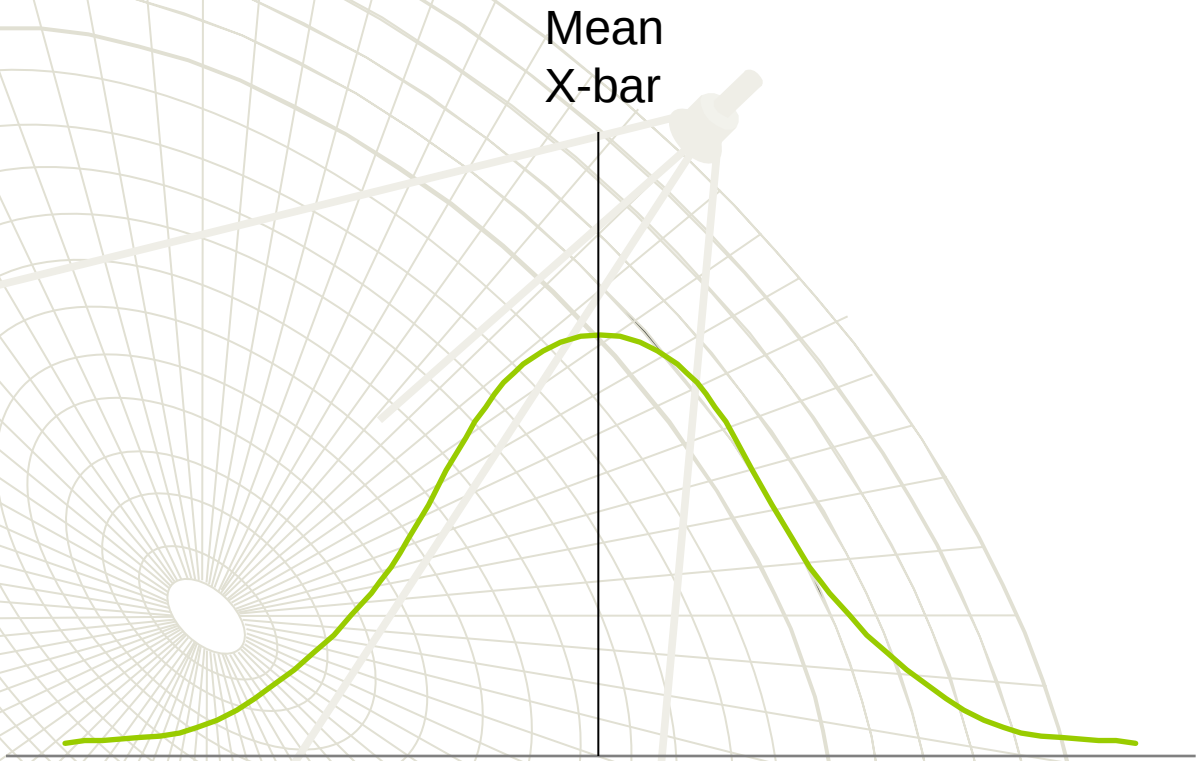
$$s = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}}$$

Understanding Variation

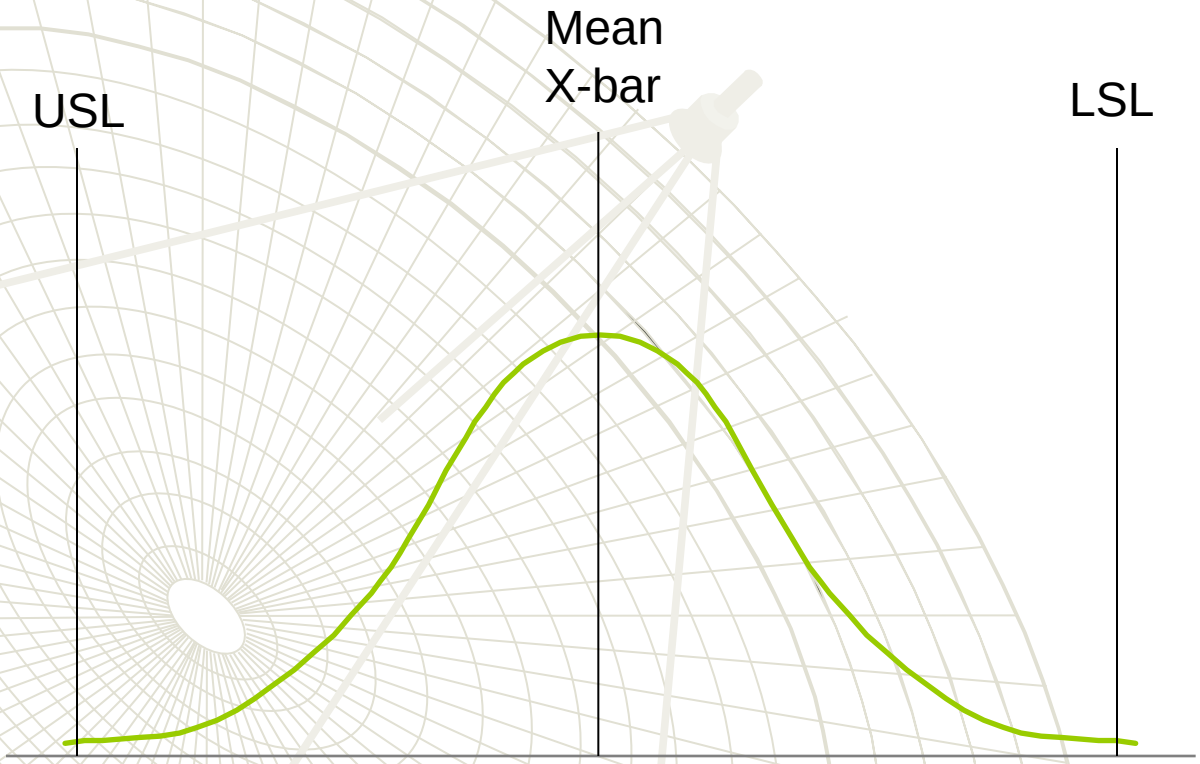


If we take the height of all the people of India and draw a distribution
Of frequencies it will tend to follow a normal distribution

Understanding Variation

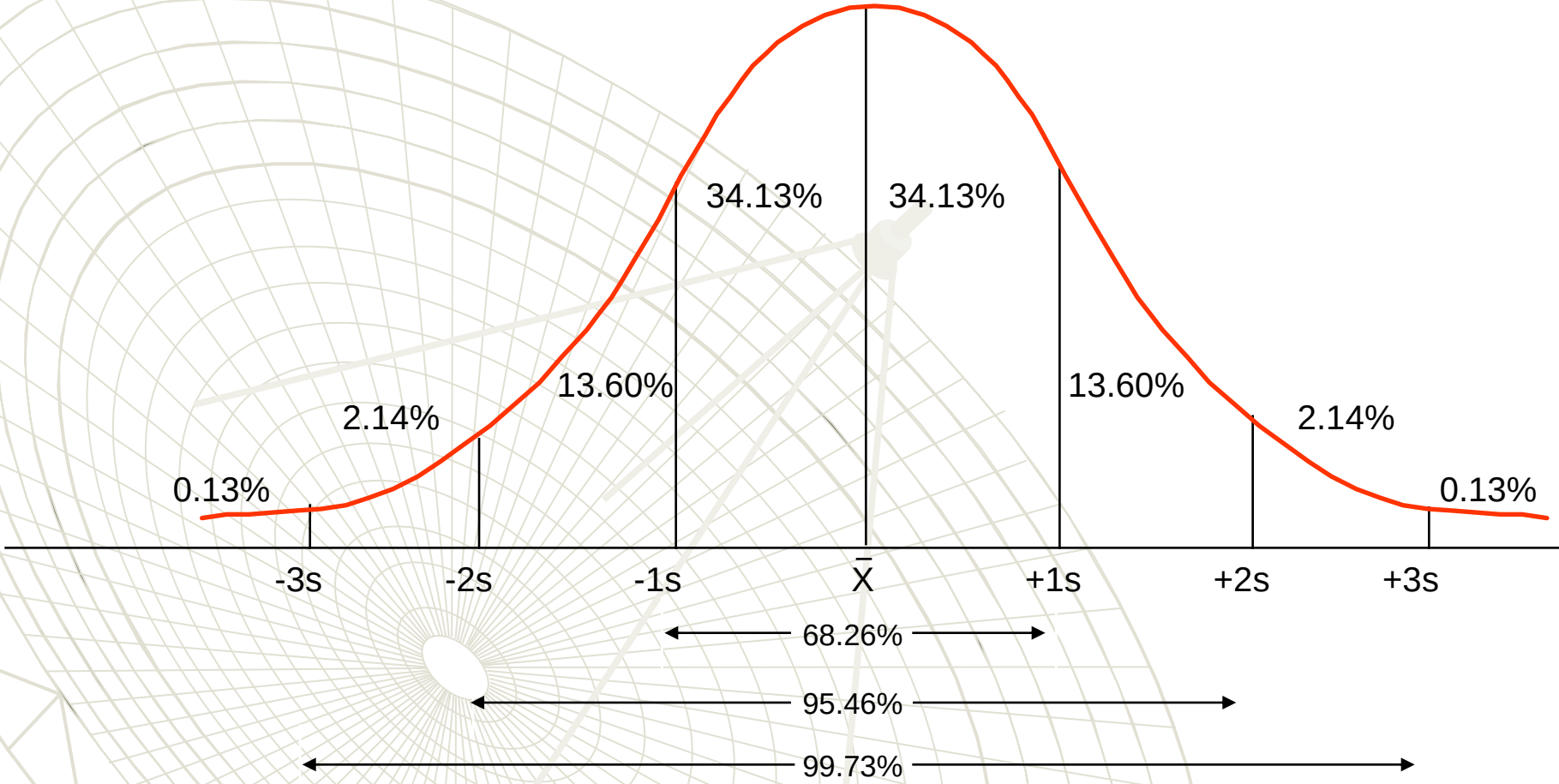


Understanding Variation



LSL and USL are those specification limits beyond which your product doesn't have a salable value in the market

Understanding Variation



σ = SIGMA

68.26% Fall Within +\- 1 Sigma

95.46% Fall Within +\- 2 Sigma

99.73% Fall Within +\- 3 Sigma

A. Helmy, PhD

Common vs. Special Cause

Type of Variation	Characteristics
Common Cause	**Always Present** Expected Predictable Normal
Special Cause	**Not Always Present** Unexpected Unpredictable Not Normal Has a surprise element

Understanding Variation

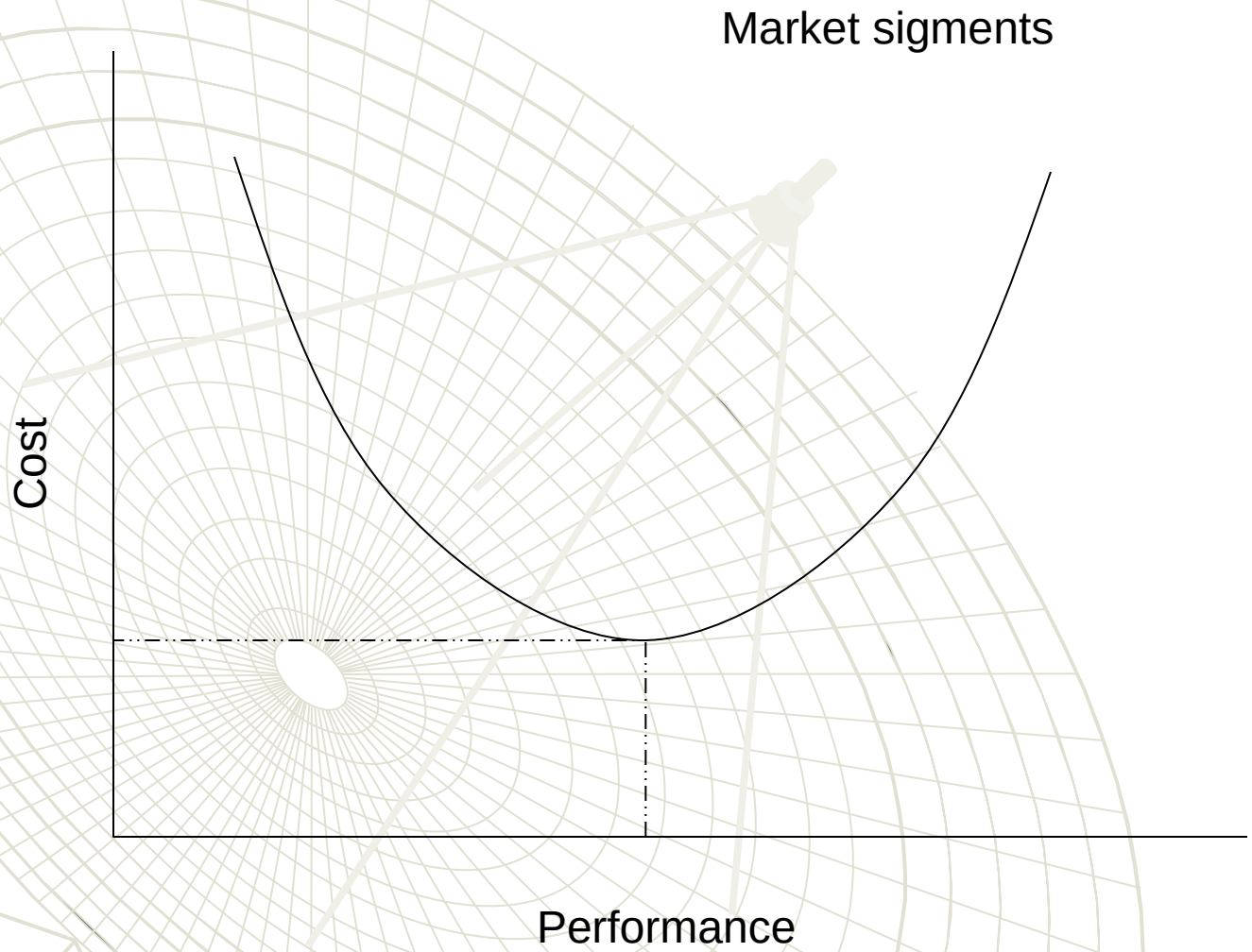
- Unsafe drinking water almost 15 minutes each day
- No electricity for almost 7 hours each month
- 20000 lost articles of mail per hour
- 5000 incorrect surgical operations per week
- 50 Newborn Babies Dropped At Birth By Doctors Each Day
- 200000 wrong drug prescriptions each year

Can you believe, all above correspond to a performance level of 99%!

If We Accepted The Goal Of 99.9% Quality

- **Then We Would Have…**
- One Hour Of Unsafe Drinking Water Every Month
- Two Unsafe Plane Landings Per Day At O'Hare International Airport In Chicago
- 16,000 Pieces Of Mail Lost By The U.S. Postal Service Every Hour
- 500 Incorrect Surgical Operations Each Week
- 22,000 Checks Deducted From The Wrong Bank Accounts Each Hour
- 32,000 Missed Heartbeats Per Person, Per Year

Taguchi Model

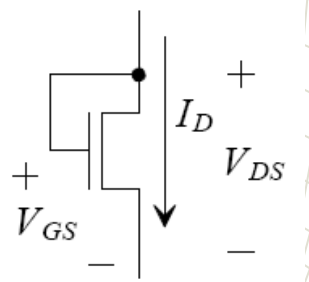


Common source amplifier (CS)

Why active load *has high r, small area, high r gives higher gain. Less power*

A diode-connected MOSFET can be thought of as a small signal resistor with a value of $1/g_m$.

$\approx 1/g_m$



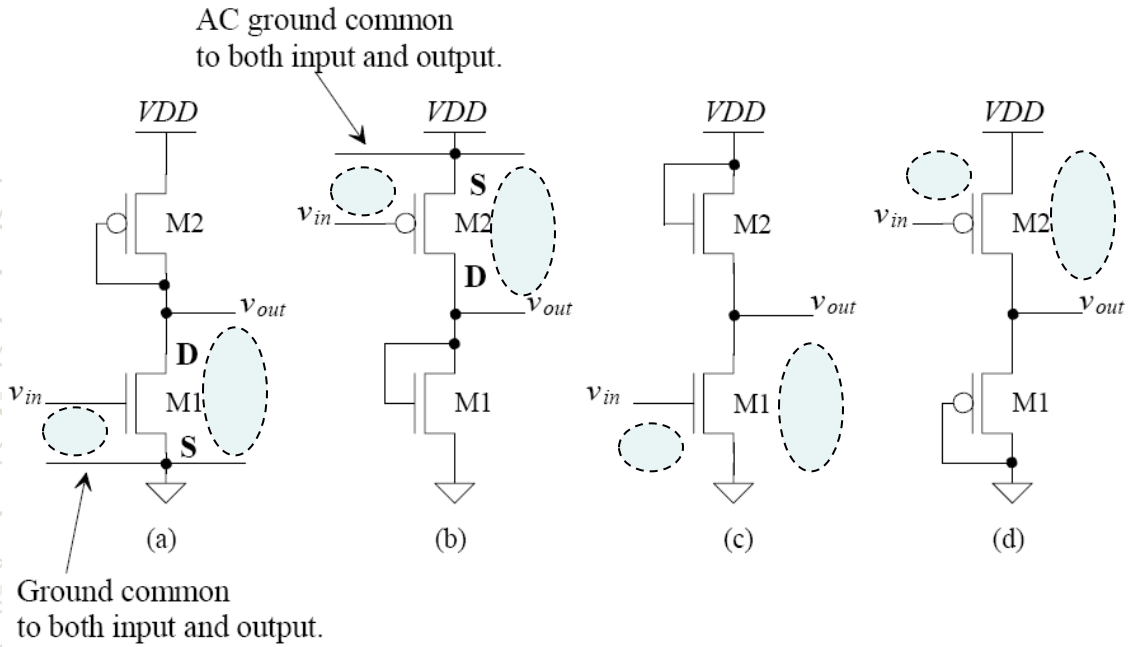
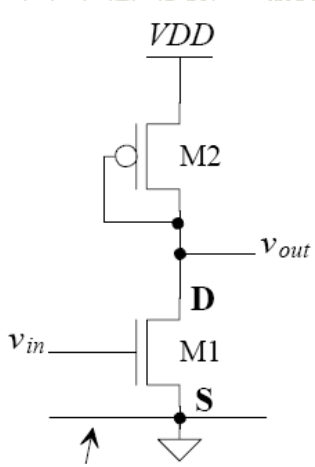
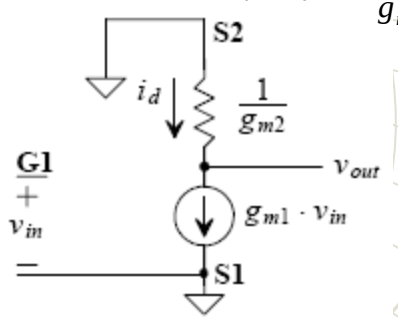


Figure 21.1 Four possible configurations of common-source amplifiers with gate-drain loads.



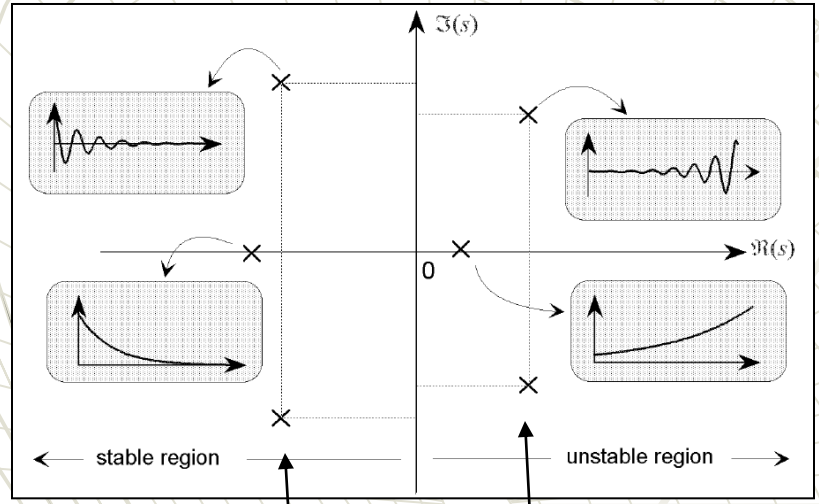
$$r_{o1} // r_{o2} >> \frac{1}{g_{m2}}$$

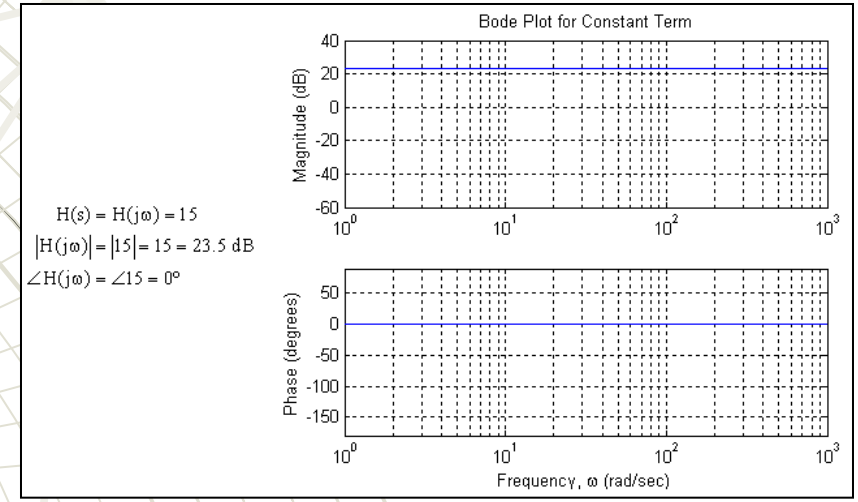


$$\frac{v_{out}}{v_{in}} = \frac{-i_d \cdot \frac{1}{g_{m2}}}{i_d \cdot \frac{1}{g_{m1}}} = \frac{-1/g_{m2}}{1/g_{m1}} = \frac{-g_{m1}}{g_{m2}}$$

Poles and Zeros

$$H(s) = \frac{N(s)}{D(s)} = K\frac{(s-z_1)(s-z_2)\ldots(s-z_{m-1})(s-z_m)}{(s-p_1)(s-p_2)\ldots(s-p_{n-1})(s-p_n)},$$



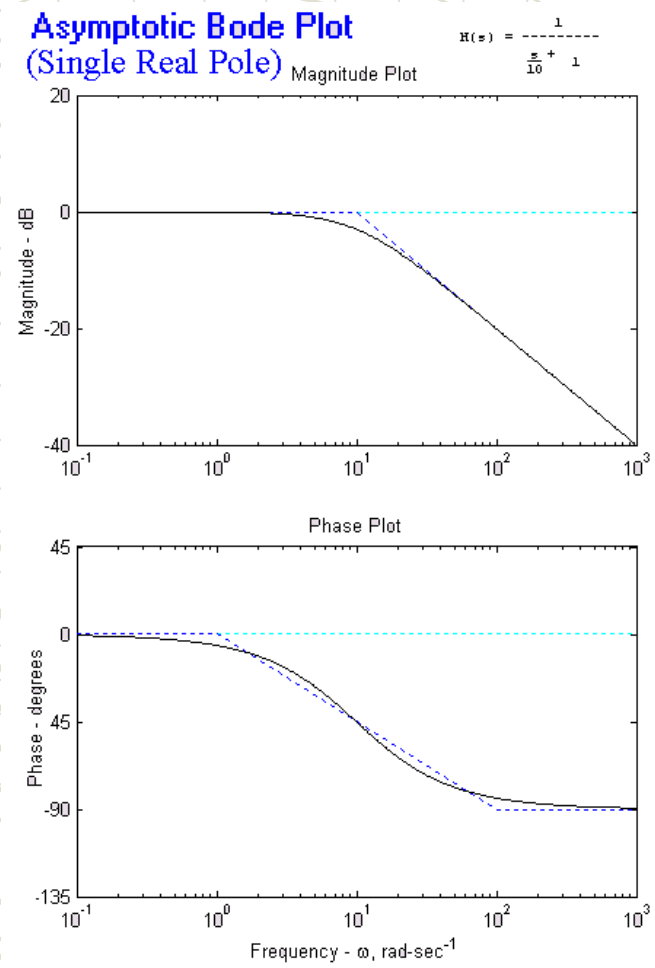


The phase is 180 for a –ve const (inverted amp "CS")

Must occur in conj pairs

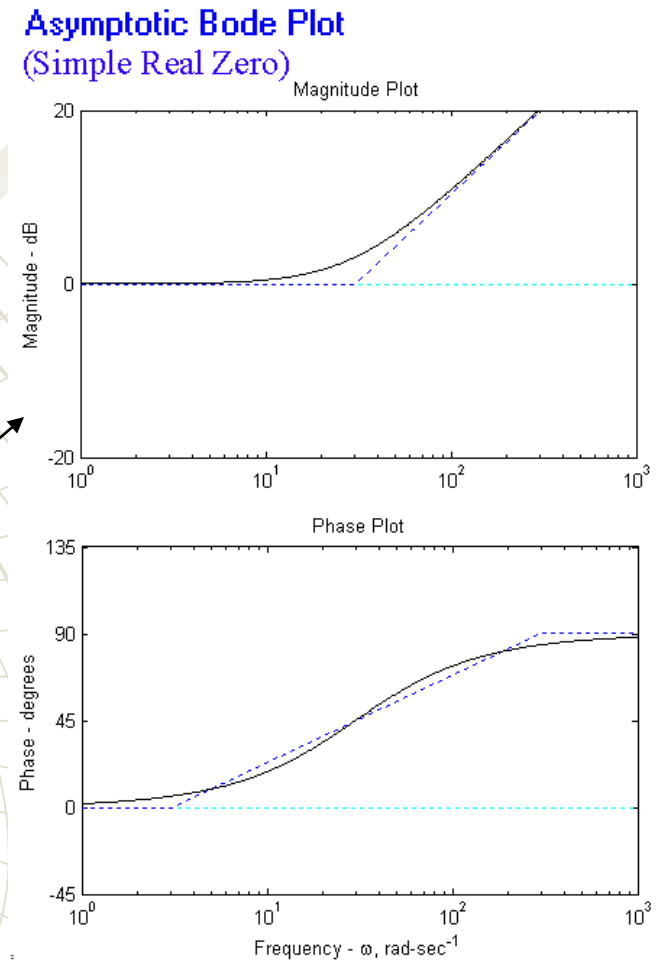
LHP Pole

LHP Zero



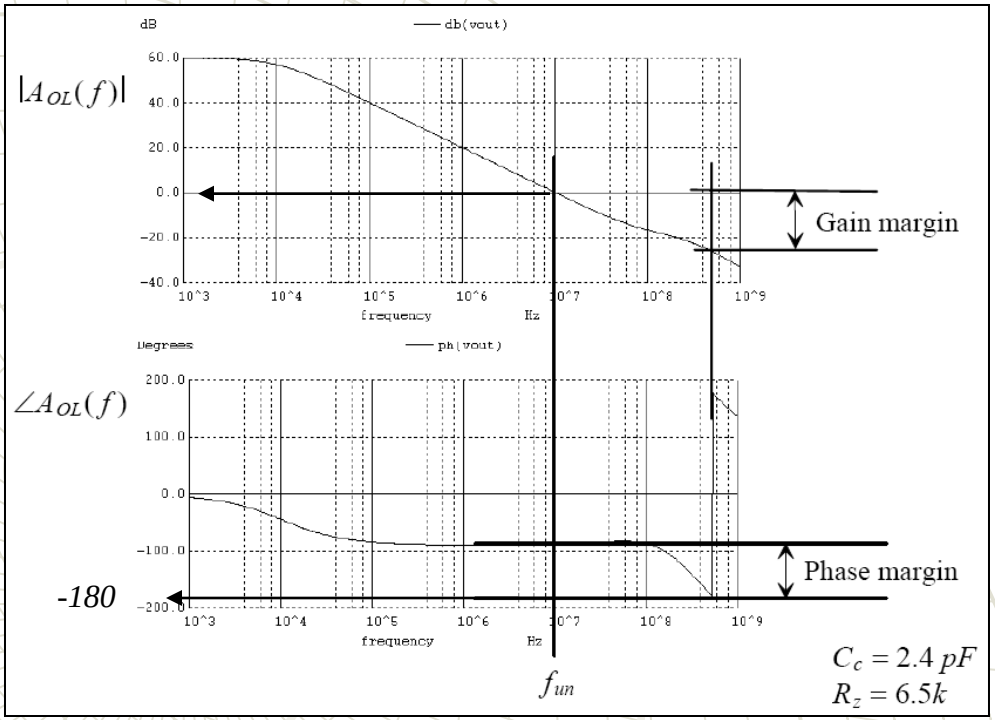
RHP Zero

$$H(s) = 1 + \frac{s}{30}$$



Stability Criteria, more on this topic later

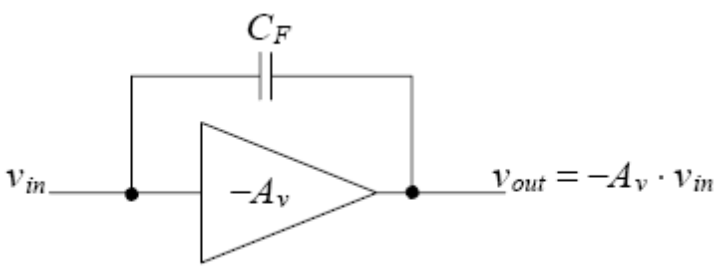
Open loop gain



The frequency response of CS amp

Overview: Miller's Theorem

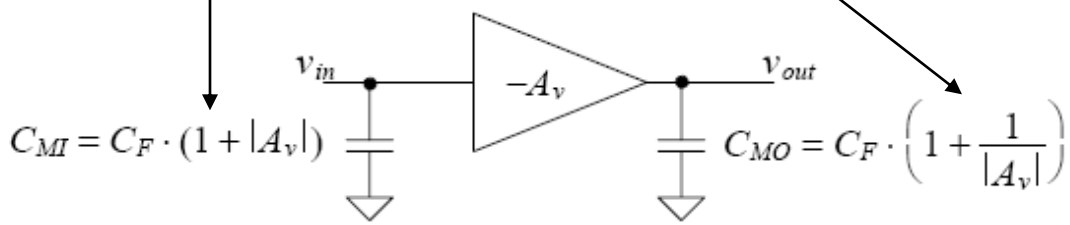
Consider an inverting amplifier as shown below:



The current applied by v_{in} to the capacitor is:

$$i = \frac{v_{in} - v_{out}}{Z_F} = jwC_F.(1+|A_v|).v_{in} = jwC_F.(1+\frac{1}{|Av|}).v_{out}$$

This looks like a single cap load at the input with a value of: and a cap at the output with a value of



(Effective input capacitance gets multiplied), also effective input capacitance is multiplied by the gain, this gives large Cin and hence slow circuit. Note that ***Miller theorem is an approximation*** *since direct connection between in and out through "C_F" is missing "shorting input and output at high frequency is not represented in Miller theory",*

Frequency response

DC gain is $-g_{m1}/g_{m2}$

Two RC time constants exist:

$$\tau_{in} = R_s(C_{mi} + C_{gs1})$$

$$C_{mi} = C_{gd1}.(1 + \frac{g_{m1}}{g_{m2}})$$

and

$$\tau_{out} = \frac{1}{g_{m2}}.(C_{sg2} + C_{mo})$$

$$C_{mo} = C_{gd1}.\left(1 + \frac{g_{m2}}{g_{m1}}\right)$$

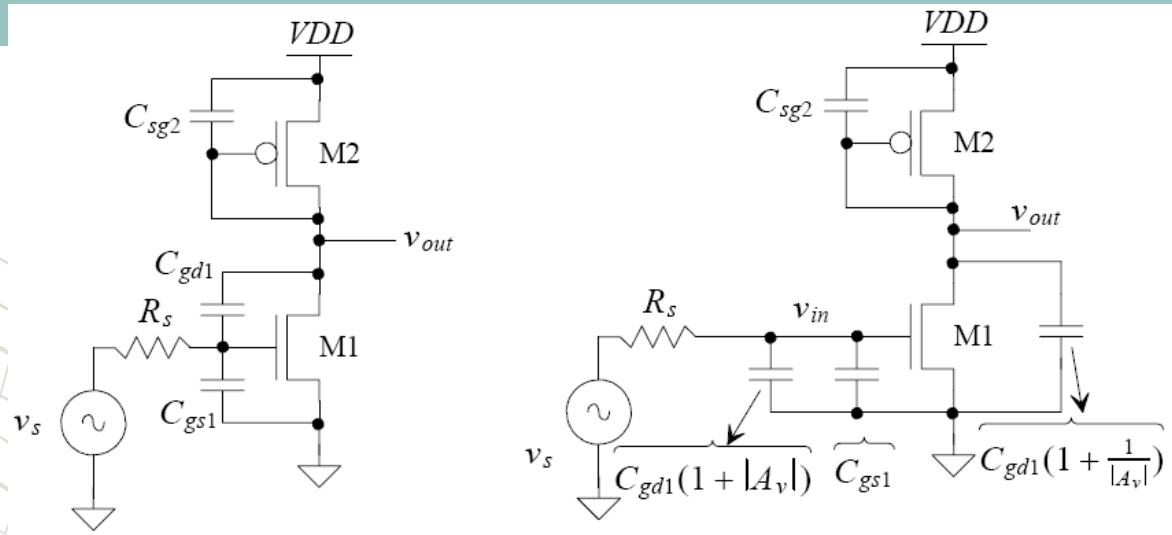
We can now define the frequency response as

$$A_v(f) = \frac{-g_{m1}/g_{m2}}{(1 + j\frac{f}{f_{in}}).(1 + j\frac{f}{f_{out}})}$$

where

$$f_{in} = \frac{1}{2\pi\tau_{in}}$$

$$f_{out} = \frac{1}{2\pi\tau_{out}}$$



(a)

(b) After applying Miller's theorem

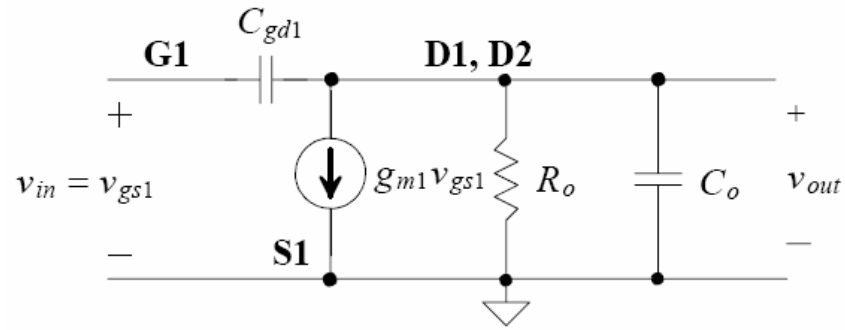
Right Half Plane (RHP) Zero

One important issue Miller's analysis ignores is the 'zero' due to capacitance between input and output, which will be effective at high frequency.

The i/p feeding directly to the o/p because of the RHP zero without the inversion can result in instability when the amp is used w/ f/b

Lets find the RHP zero

Consider the small signal model, where Ro models both 'load' and input transistor output impedance.



$$R_o = \frac{1}{g_{m2}} // r_{o2} // r_{o1} \approx \frac{1}{g_{m2}}$$

$$C_o = C_{sg2}$$

Summing currents at the output:

$$\frac{v_{out} - v_{in}}{1/jwC_{gd1}} + \frac{v_{out}}{R_o \,||\, (1/jwC_o)} + g_{m1}.v_{in} = 0$$

$$\frac{v_{out}}{v_{in}} = -g_{m1}.R_o.\frac{1 - jw\dfrac{C_{gd1}}{g_{m1}}}{1 + jw(C_{gd1} + C_o).R_o}$$

$$\tau_{out} = \frac{1}{g_{m2}}.(C_{sg2} + C_{mo})$$

$$C_{mo} = C_{gd1}.\left(1 + \frac{g_{m2}}{g_{m1}}\right)$$

$$A_v(f) = -g_{m1}.R_o.\frac{\left(1 - j\dfrac{f}{f_z}\right)}{\left(1 + j\dfrac{f}{f_{out}}\right)}$$

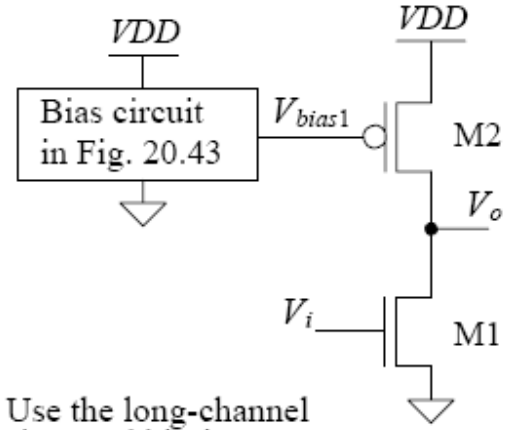
The output pole is similar to what was derived before if Av=-g_{m1}/g_{m2} is small

$$f_Z = \frac{g_{m1}}{2\pi C_{gd1}}$$

A zero in the right half plane (RHP) has the same magnitude response as a zero in the left hand plane (LHP). However RHP zero has the same phase impact as a pole in the LHP. Can cause instability in feedback configurations. Phase margin & +ve fb.

Amplifiers with Current Source Loads why??

This provides the largest possible load resistance and thus the largest gain

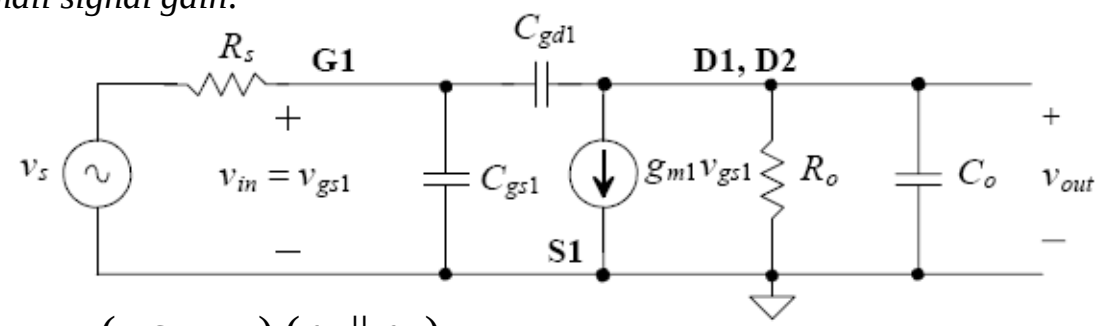


Use the long-channel sizes and biasing seen in Table 9.1.

Common Source Amplifier

Also called Class-A amplifier. Both M1 and M2 conduct current (in saturation) for proper operation.

Small signal gain:



$$v_o = (-g_{m1}.v_{in}).(r_{o1} \parallel r_{o2})$$

$$\frac{v_o}{v_{in}} = -g_{m1}.(r_{o1} \parallel r_{o2}) = \frac{-g_{m1}}{g_{o1} + g_{o2}}$$ DC Gain

Frequency Response of CS Amplifier:

*The o/p node at the drain is a high Z node, with only D connection $Z_h = r_{o1} // r_{o2}$. Source or drain connected is a low Z node $Z_l = 1/g_m$. High z nodes have low freq pole that limit the speed/BW (but the gain is high) again it is a gain*BW ≈ const*

Using Miller approximation

$$f_{in} \approx \frac{1}{2\pi\left(C_{gs1} + C_{gd1}.\left(1+|A_v|\right)\right).R_s}$$

$$f_{out} \approx \frac{1}{2\pi\left(C_{dg2} + C_{gd1}.\left(1+\frac{1}{|A_v|}\right)\right).(r_{o1} // r_{o2})}$$

$$f_z = \frac{g_{m1}}{2\pi.C_{gd1}}$$

$$A_v = -g_{m1}.(r_{o1} // r_{o2})$$

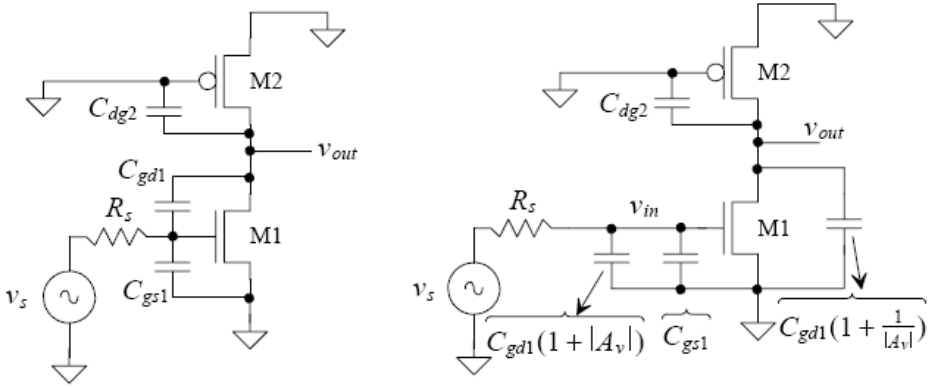
$$A_v(f) = -g_{m1}.(r_{o1} \| r_{o2}).\frac{\left(1 - j\frac{f}{f_z}\right)}{\left(1 + j\frac{f}{f_{in}}\right)\left(1 + j\frac{f}{f_{out}}\right)}$$

Assume

$$g_{m1} = 150\mu A/V$$
$$r_{o1} = 5M\Omega$$
$$r_{o1} = 4M\Omega$$
$$A_v = -150\mu A/V.(5M\Omega \| 4M\Omega) = 333V/V = 50dB$$
$$C_{in} = (C_{gs1} + C_{gd1}.(A_v)) \approx 690\,fF$$
$$C_{out} = (C_{gd2} + C_{gd1}.(1+\frac{1}{A_v})) \approx 6 + 2\,fF = 8\,fF$$
$$f_{in} = 2.3MHz,\ f_{out} = 9MHz,\ f_z = 11.9\,GHz$$

Compare poles and zero freq to previous lecture

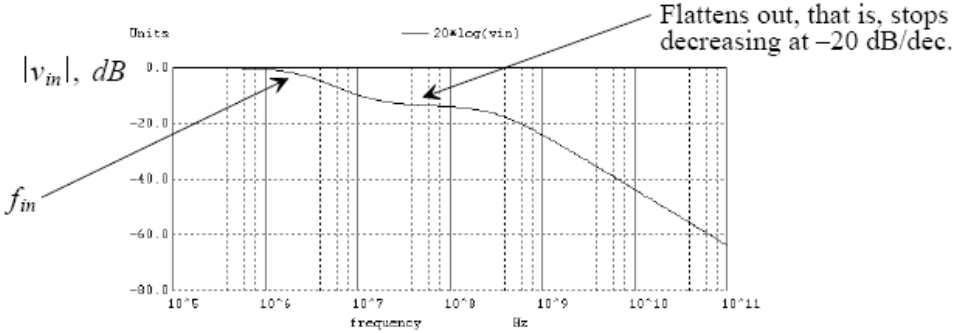
A. Helmy, PhD



(a) (b) After applying Miller's theorem

In reality f_{out} is at a much higher frequency, around 200MHz. Why ?! Because of pole-splitting !! As A_v starts dropping at 20dB/dec beyond f_{in}, v_{in} sees less and less capacitance $C_{gd1}(1+A_v)$, i.e. higher Z, So it ceases to decrease "flattens out" and f_{out} appears at a much higher frequency. This happens for high A_v. If R_s is zero, there won't be pole splitting nor f_{in} and f_{out} equation in this case will be the only pole and its expression will be accurate.

End effect is that the second pole 'splits' and starts moving away. Why didn't we have this in the previous lecture?? Let us analyze pole splitting in detail.



No Miller Approx.

$$R_o = r_{o1} \parallel r_{o2}$$

$$C_o = C_{gd2}$$

Note that this is not $\frac{v_{out}}{v_s}$

$$\frac{v_{out}}{v_{in}} = -g_{m1}.R_o.\frac{1 - jw\frac{C_{gd1}}{g_{m1}}}{1 + jw.(C_{gd1} + C_o).R_o}$$

Similar to Previous lecture summing I's at node D1, D2

At node G1, KCL gives:

$$\frac{v_{in} - v_s}{R_s} + \frac{v_{in}}{1/jwC_{gs1}} + \frac{v_{in} - v_{out}}{1/jwC_{gd1}} = 0$$

$$v_{in} = \frac{\frac{v_s}{R_s} + v_{out}.j\omega C_{gd1}}{\frac{1}{R_s} + j\omega C_{gs1} + j\omega C_{gd1}}$$

$$v_{out} = -g_{m1}R_o.\frac{1 - j\omega\frac{C_{gd1}}{g_{m1}}}{1 + j\omega(C_{gd1} + C_o).R_o}.\frac{\frac{v_s}{R_s} + v_{out}.j\omega C_{gd1}}{\frac{1}{R_s} + j\omega C_{gs1} + j\omega C_{gd1}}$$

$$f_z = \frac{g_{m1}}{2\Pi C_{gd1}}$$

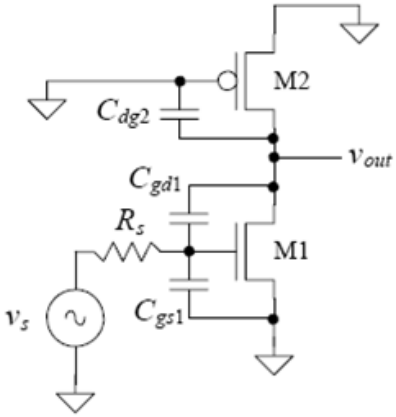
$$\frac{v_{out}}{v_s} = \frac{-g_{m1}.R_o.(1 - s\frac{C_{gd1}}{g_{m1}})}{A.s^2 + B.s + 1}$$

$$A = R_o.R_s.(C_{gd1}.C_{gs1} + C_o.C_{gs} + C_o.C_{gd1})$$

$$B = [(C_{gd1} + C_o).R_o + (C_{gs1} + C_{gd1})R_s + C_{gd1}.g_{m1}.R_o.R_s]$$

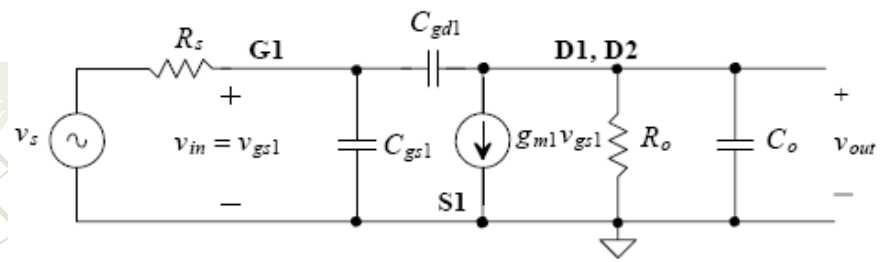
$$B = (C_{gd1} + C_{dg2}).r_{o1} \parallel r_{o2} + (C_{gs1} + C_{gd1}(1 + A_v)).R_s$$

A. Helmy, PhD



$$f_{in} \approx \frac{1}{2\pi(C_{gs1} + C_{gd1}.(1 + |A_v|)).R_s}$$

$$f_{out} \approx \frac{1}{2\pi\left(C_{dg2} + C_{gd1}.\left(1 + \frac{1}{|A_v|}\right)\right).(r_{o1} // r_{o2})}$$



At low frequencies, term B dominates since "s^2 or w^2 is small", yielding the low frequency denominator

$1 + jw.B$ *where* $A_v = g_{m1}(r_{o1} \parallel r_{o2}) = g_{m1}R_o, \quad C_o = C_{dg2}$

low freq pole is located at

$$f_1 \approx \frac{1}{2\pi[(C_{gd1} + C_{dg2}).r_{o1} \parallel r_{o2} + (C_{gs1} + C_{gd1}(1 + A_v)).R_s]}$$

if $C_{gd1}(1 + A_v) >> $ *other caps*

$$f_1 \approx \frac{1}{2\pi C_{gd1}(1 + A_v)R_s}$$ similar to f_{in} using miller approx

if $R_s = 0$, $A = 0$ and

$$f_1 \approx \frac{1}{2\pi(C_{gd1} + C_{dg2}).r_{o1} \parallel r_{o2}}$$ similar to f_{out} using miller approx

To determine f2 we factor f1 out of the denominator, we can put the denominator in the form: $(1+j\frac{f}{f_1}).(1+j\frac{f}{f_2})$

$$\left[1+s\left[\left(C_{gd1}+C_o\right)R_o+\left(C_{gs1}+C_{gd1}\right)R_s+C_{gd1}g_{m1}R_oR_s\right]\right].$$
$$\left(1+\frac{s^2\left[R_oR_s\left(C_{gd1}C_{gs1}+C_oC_{gs1}+C_oC_{gd1}\right)\right]}{s\left[\left(C_{gd1}+C_o\right)R_o+\left(C_{gs1}+C_{gd1}\right)R_s+C_{gd1}g_{m1}R_oR_s\right]+1}\right)$$

$$As^2+Bs+1=$$
$$(1+Bs)(1+As^2/(1+Bs))$$

If Rs=0 second term goes to unity i.e f2 is not present, amp tr function has a single pole associated with the output, else Rs needs to be large. Looking at the second term and dividing it by sRoRs gives

$$\left(1+j\frac{2\pi f.\left[\left(C_{gd1}C_{gs1}+C_oC_{gs1}+C_oC_{gd1}\right)\right]}{\left[\left(C_{gd1}+C_o\right)/R_s+\left(C_{gs1}+C_{gd1}\right)/R_o+C_{gd1}g_{m1}\right]+1/sR_oR_s}\right),\quad \text{Rs is large}$$

Also note that $g_m >> g_o \approx \frac{1}{r_o}$ *Solving for f2, we have:* $f_2=\frac{g_{m1}.C_{gd1}}{2\Pi.(C_{gd1}.C_{gs1}+C_o.C_{gs1}+C_o.C_{gd1})}$

This is quite different from the 'output pole' approximated with Millers Theorem. Substituting for g_m *and C's wee have f2=240MHz. Why do we use the term 'pole splitting'? If we place an intentional capacitor Cc in parallel to* Cgd1 *f1 decreases while f2 pushes out.*

$$A=R_o.R_s.(C_{gd1}.C_{gs1}+C_o.C_{gs}+C_o.C_{gd1})$$
$$B=[(C_{gd1}+C_o).R_o+(C_{gs1}+C_{gd1})R_s+C_{gd1}.g_{m1}.R_o.R_s]$$

Generally for any 2 stage amplifier that can be put in this form

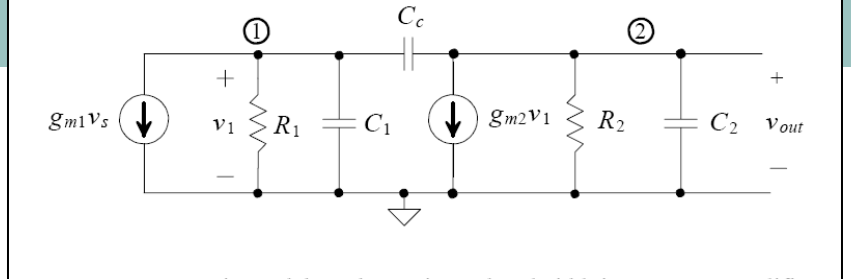
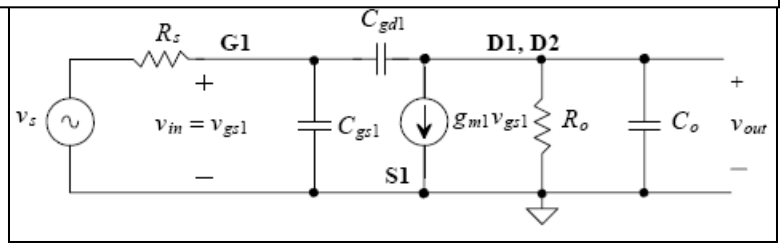


Figure 21.25 Generic model used to estimate bandwidth in a CMOS amplifier.



$$f_z = \frac{g_{m2}}{2\pi.C_c}$$

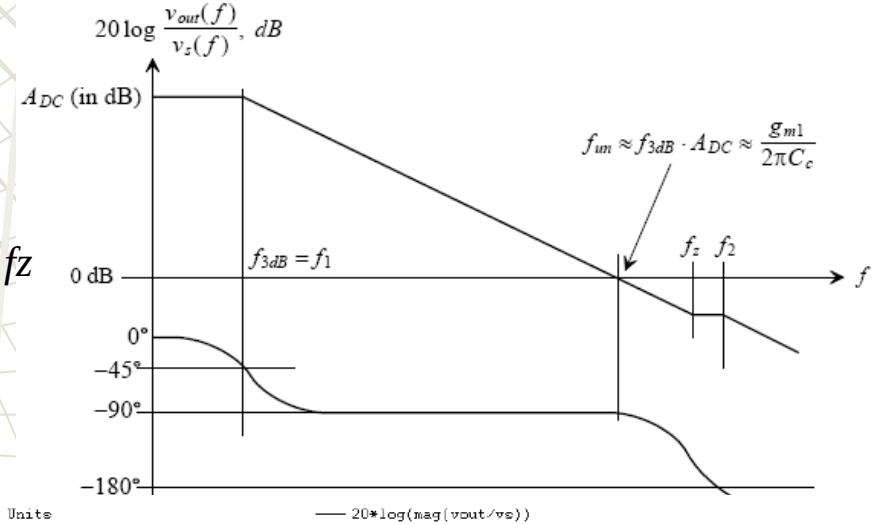
$$f_1 \approx \frac{1}{2\pi\left[(C_c + C_2).R_2 + (C_1 + C_c(1 + g_{m2}R_2)).R_1\right]}, \text{ for large gain } g_{m2}R_2$$

$$f_1 \approx \frac{1}{2\pi\left[(C_c(g_{m2}R_2)).R_1\right]} = f_{3dB}$$

$$f_2 \approx \frac{g_{m2}C_c}{2\pi.(C_cC_1 + C_1C_2 + C_cC_2)}$$

As C2 α CL increases, f2 decreases, hurt stability

$$A_v(f) = \frac{v_{out}(f)}{v_s(f)} = g_{m1}R_1g_{m2}R_2.\frac{\left(1 - j\frac{f}{f_z}\right)}{\left(1 + j\frac{f}{f_1}\right)\left(1 + j\frac{f}{f_2}\right)}$$



If Cc is made large we will get pole splitting and f1 << f2, fz i.e f1 is called the dominant pole

$$A_v(f) = \frac{v_{out}(f)}{v_s(f)} = \frac{g_{m1}R_1g_{m2}R_2}{\left(1 + j\frac{f}{f_1}\right)} = \frac{g_{m1}R_1g_{m2}R_2}{(1 + j.2\pi f.g_{m2}R_2R_1C_c)}$$

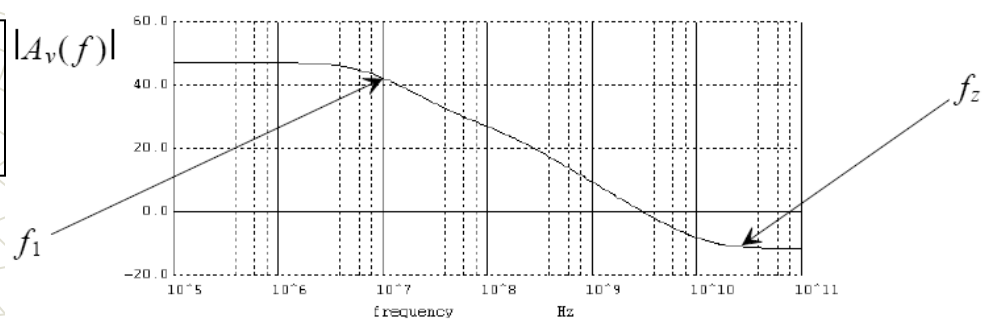
$f / f_1 >> 1$

$$A_v(f) = \left|\frac{v_{out}(f)}{v_s(f)}\right| = \frac{g_{m1}}{(2\pi f.C_c)}$$

Increasing Cc will cause fz and f1 to decrease need zero cancellation

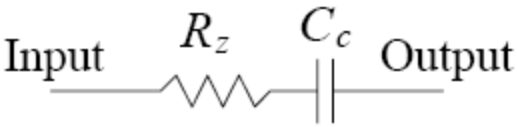
$$f_{un} = A_v.f_{3dB} = \frac{g_{m1}}{(2\pi.C_c)}$$

Unity gain frequency



As discussed in the previous lecture, the amp i/p feeds directly through Cc to the amp o/p w/o phase inversion. For high gain amp the lack of inversion when the amp uses fb can result in an unstable amp. To avoid this situation we add a resistor in series with Cc to attenuate the high freq signals (where the zero occur) and push the zero to a higher freq

$$f_z = \frac{g_{m2}}{2\pi.C_c} = \frac{1}{2\pi.C_c.\frac{1}{g_{m2}}}$$



With a resistor

$$f_z = \frac{1}{2\pi.C_c.\left(\frac{1}{g_{m2}} - R_Z\right)}$$

Can you work this out ??

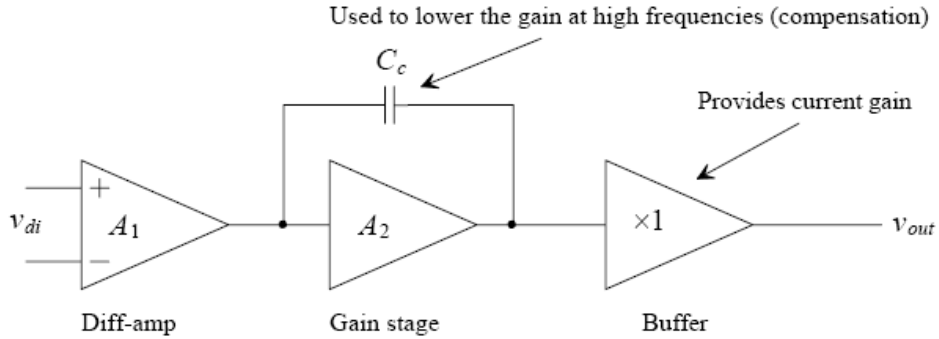
If Rz=1/gm$_2$ the zero disappears (pushed to inf freq). if Rz>1/gm2 zero is pushed to LHP to provide 90 degrees phase boast that aids stability (lead compensation), any practical design that uses pole splitting should include zero-nulling resistor

Very imp notes and summary

- High imp nodes has low freq poles
- High o/p res amp can not drive high C loads “f2 decreases hurt stability”
- High o/p res amp can not drive low res loads “its gain will drop due to loading, needs a buffer with low o/p resistance”
- Increasing Cc will cause fz to decrease need zero cancellation else stability will be hurt and f1 will decrease as well
- Next lecture, OTA and folded cascodes has all nodes at low imp 1/gm, all of its poles are at high freq, except output node that has a high imp “dominant pole”, thus compensation can be done via load cap. Opposite to two stage amps

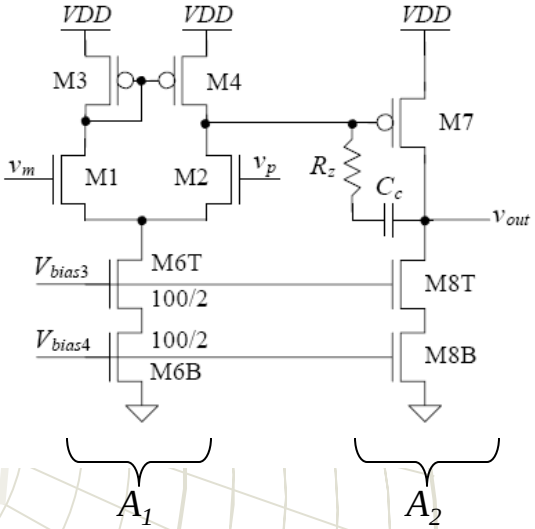
Specifications:

AOL=open loop gain
CMR=common mode range
CMRR=common mode rejection ratio
PSRR=power supply rejection ratio
Output voltage range
Current sourcing/sinking capability
Power dissipation
Total Harmonic Distortion

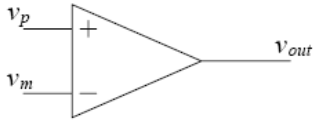


Basic 2-Stage OP-AMP Design

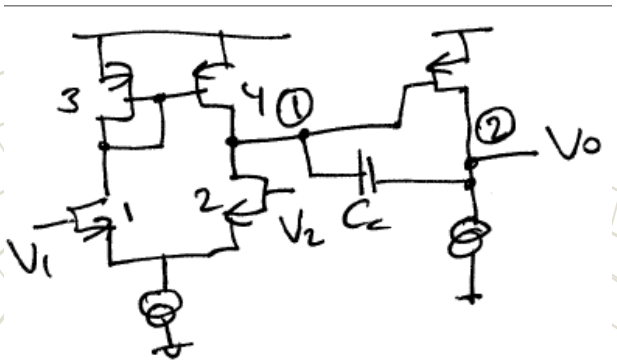
$$DC\ Gain = A_{OL,DC} = A_1 . A_2 = g_{mn} . (r_{on} || r_{op}) . g_{mp} . r_{o7}$$

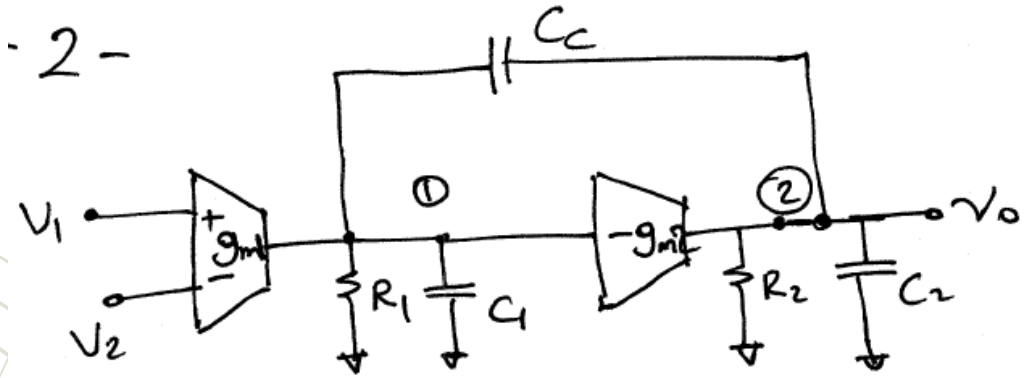


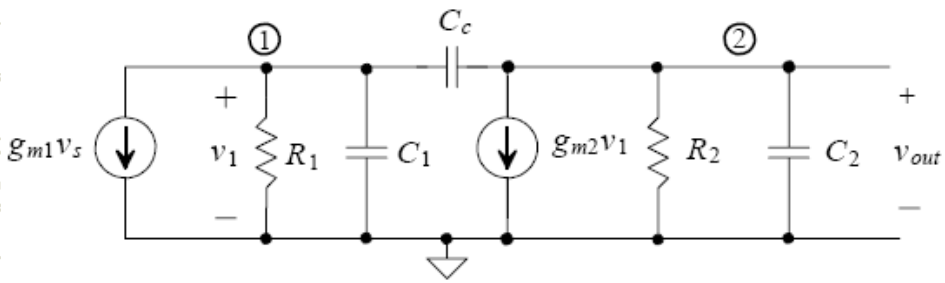
Parameters from Table 9.2 with biasing circuit from Fig. 20.47. Unlabeled NMOS are 50/2 and PMOS are 100/2. Scale factor is 50 nm.



Op-amp schematic symbol







As shown in lecture 4 for a generic amplifier a RHP zero is located at

$$f_z = \frac{g_{m2}}{2\pi . C_c}$$

$$f_1 \approx \frac{1}{2\pi\left[(C_c + C_2).R_2 + \left(C_1 + C_c\left(1 + g_{m2}R_2\right)\right).R_1\right]}, \text{ for large gain } g_{m2}R_2$$

$$f_1 \approx \frac{1}{2\pi\left[\left(C_c\left(g_{m2}R_2\right)\right).R_1\right]} = f_{3dB}$$

$$f_2 \approx \frac{g_{m2}C_c}{2\pi.\left(C_cC_1 + C_1C_2 + C_cC_2\right)}$$

2nd pole

$$A_v(f) = \frac{v_{out}(f)}{v_s(f)} = g_{m1}R_1 g_{m2}R_2 \frac{\left(1 - j\frac{f}{f_z}\right)}{\left(1 + j\frac{f}{f_1}\right)\left(1 + j\frac{f}{f_2}\right)}$$

DC gain

- *Dominant pole is set by Cc multiplied by the low freq gain of second stage*
- *As Cc increases f1 decreases and f2 increases causing pole splitting*

As C2 α CL increases, f2 decreases, hurt stability

If Cc is made large will get pole splitting and f1 << f2, fz i
f1 is called the dominant pole. The main motive is to make
2-pole, 1-zero system looks like a single pole integrator

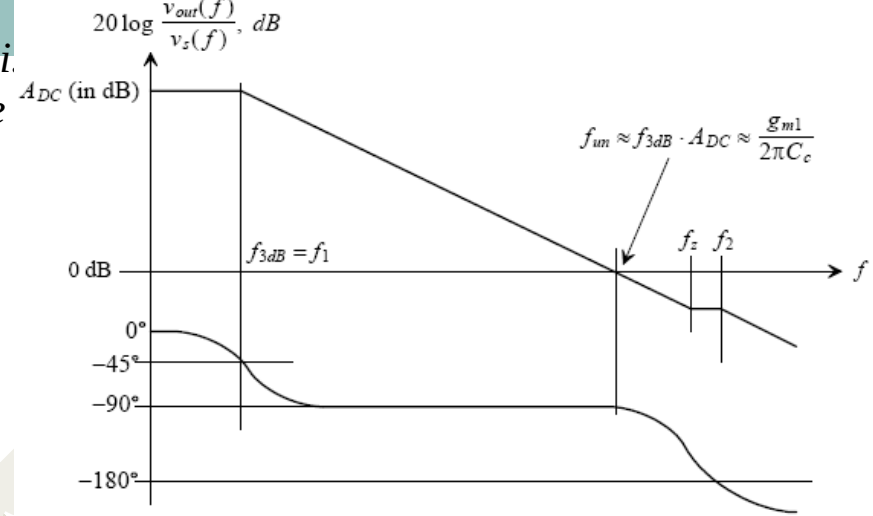


Figure 21.26 Magnitude and phase responses of a generic CMOS amplifier.

$$A_v(f) = \frac{v_{out}(f)}{v_s(f)} = \frac{g_{m1}R_1 g_{m2}R_2}{\left(1 + j\frac{f}{f_1}\right)} = \frac{g_{m1}R_1 g_{m2}R_2}{(1 + j.2\pi f.g_{m2}R_2R_1C_c)}$$

At frequencies above f1 we have $f / f_1 >> 1$

$$A_v(f) = \left|\frac{v_{out}(f)}{v_s(f)}\right| = \frac{g_{m1}}{(2\pi f.C_c)}$$

$$f_{un} = \frac{g_{m1}}{(2\pi.C_c)}$$ *Unity gain frequency*

Increasing Cc will cause fz and f1 to decrease need zero cancellation

$$BW = f_1 = f_{3dB} = \frac{1}{(2\pi.g_{m2}R_2R_1C_c)}$$

$$f_{un} = \frac{g_{m1}}{(2\pi.C_c)} = g_{m1}R_1 g_{m2}R_2 \frac{1}{(2\pi.g_{m2}R_2R_1C_c)}$$

Unity gain frequency or gain band width product is f_{un}

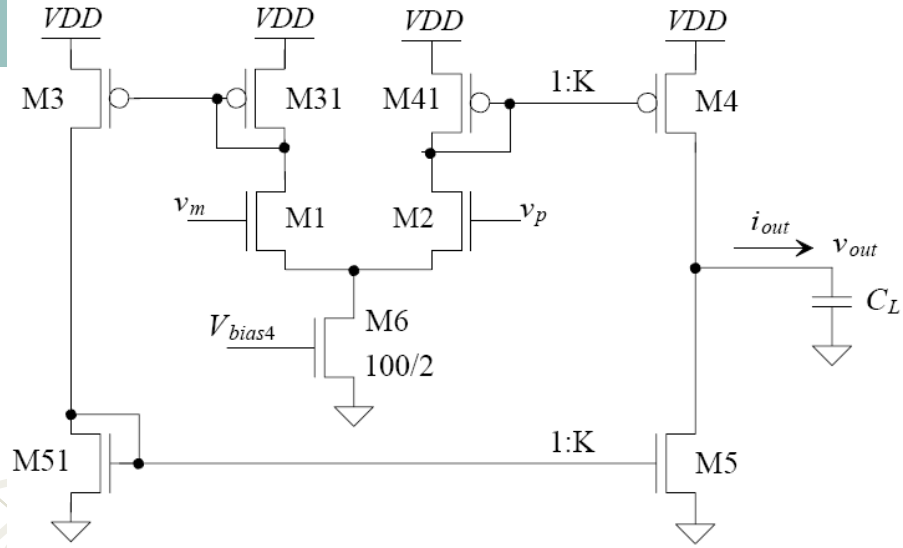
As f1 and f2 split fz is still present as a RHP zero, impacting phase margin. Also note that as gm1=gm2 unity gain and zero frequencies may overlap => almost guaranteed low phase margin !!

Hence we need to cancel it out.

Operational Trans-conductance Amplifier (OTA)

A voltage to current converter that can only drive capacitive loads (no resistive loads). A typical one is a gm/gm stage

All nodes are low impedance (sees 1/gm as impedance, not ro), except output node



o/p compensated opamp

$$\beta_4 = K.\beta_{41} = K.\beta_{31} = K.\beta_3$$

$$K.\beta_{51} = \beta_5$$

$$i_3 = i_{31}$$

$$i_4 = K.i_{41}$$

$$v_{gs1} = -v_{gs2}$$

$$v_p - v_m = v_d = v_{gs1} - v_{gs2} = 2v_{gs1}$$

$$-i_{31} = i_{41} = -i_3 = i_d = g_{mn} v_{gs1} = \frac{g_{mn}}{2}(v_p - v_m)$$

$$i_3 = -i_{41} = -i_4 / K$$

$$i_5 = K.i_{51}$$

$$i_3 = i_{51} = -i_4 / K = i_5 / K$$

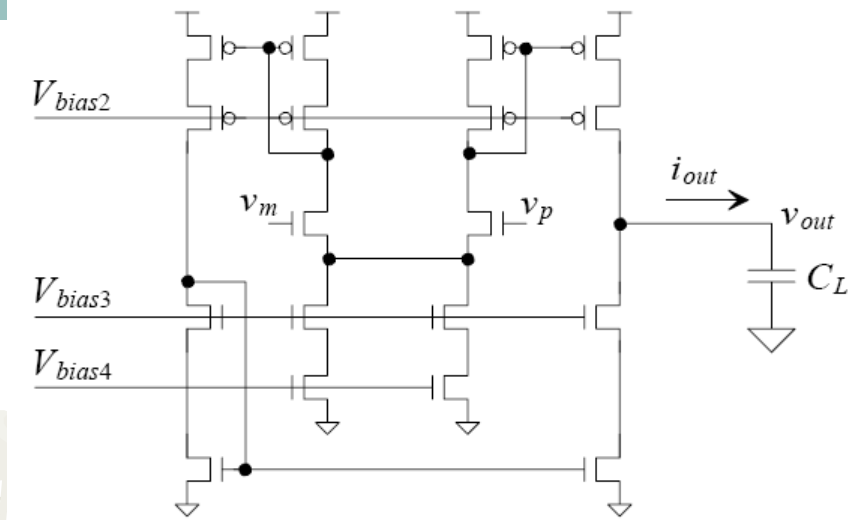
$$i_4 = -i_5 = -K.i_3 = K.i_d$$

$$v_{out} = i_{out}.(r_{o4} \parallel r_{o5}) = (i_4 - i_5).(r_{o4} \parallel r_{o5}) = 2K.i_d.(r_{o4} \parallel r_{o5}) = 2K\frac{g_{mn}}{2}(v_p - v_m)(r_{o4} \parallel r_{o5})$$

$$A_v = \frac{v_{out}}{v_p - v_m} = K.g_{mn}.(r_{o4} \parallel r_{o5})$$

$$g_{mOTA} = \frac{i_{out}}{v_p - v_m} = \frac{2K.i_d}{2.i_d / g_{mn}} = K.g_{mn}$$

The ideal OTA has infinite output resistance. If we cascode the current mirrors, we will be increasing the OTA output resistance without the need for big devices.



$$A_v = g_{mn}.\left(R_{ocasn} \parallel R_{ocasp}\right)$$

$$f_{3dB} = \frac{1}{2\pi\left(R_{ocasn} \parallel R_{ocasp}\right).C_L}$$

$$GBW = f_{un} = A_v.f_{3dB} = \frac{g_{mn}}{2\pi.C_L}$$ *As before*

Single stage o/p compensated 1 pole
no need for CC, high PSRR

Increasing the load capacitance decreases the unity gain freq making the OTA more stable "this is an output compensated amp". Unlike the two stage op-amp where the op-amp can become unstable with large load cap (as f2 will decrease). The OTA only becomes more stable with large load caps. Adding a second stage to the OTA make it like what we discussed before and it may become unstable

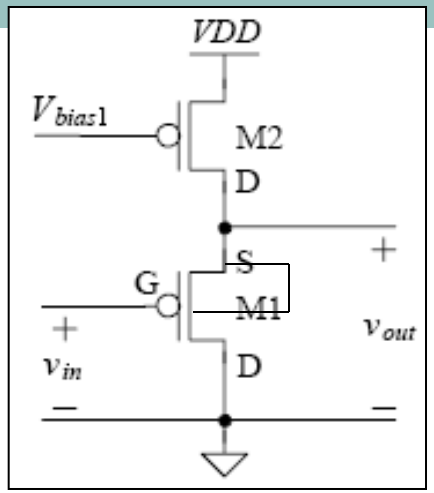
CC ↑ f1 ← f2 →

CC↑ fz ← zero nulling is needed

CL ↑ f2 ←

For 2 stage opamp
Increase CC to counter balance high CL, but PSRR will degrade

The Source Follower (SF Common-drain Amplifier)



$$\frac{v_{out}}{v_{in}} = g_{m1}\left(R_{S_of_M1} \parallel R_{D_of_M2}\right) \text{no body effect}$$

$$\frac{v_{out}}{v_{in}} = g_{m1}\left(\frac{1}{g_{m1}} \parallel r_o\right) \approx g_{m1}\left(\frac{r_o}{g_{m1}r_o+1}\right) = \frac{g_{m1}r_o}{g_{m1}r_o+1} \approx 1 \text{(specially if M2 is cascode)}$$

Used as a level shifter $v_{out}=v_{sg1}+v_{in}$

Vin min

Vout > vsg-vtp, vout > vout-vin-vtp, vin > -vtp

$V_{in\,max} < V_{dd} - v_{ds,sat} - v_{sg}$ ⇒ $\boxed{-v_{tp} < V_{in} < V_{dd} - v_{ds,sat} - v_{sg}}$

Input capacitance of SF

Consider the charge entering the gate capacitance:

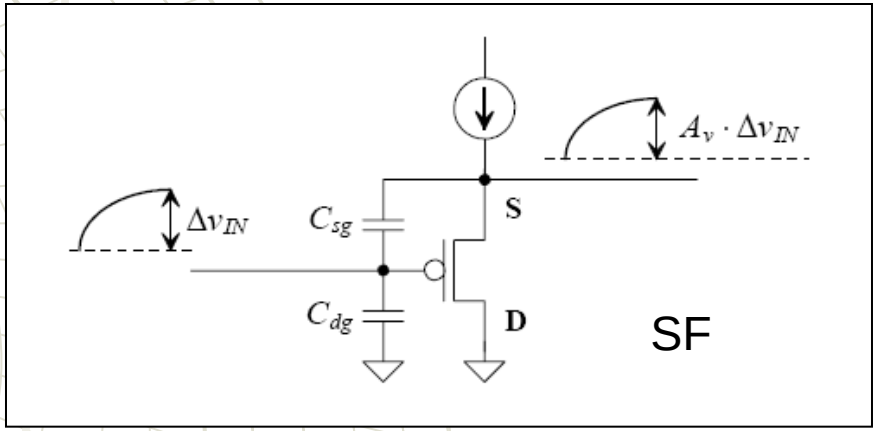
$$Q_{IN} = \Delta v_{in}.C_{dg} + (\Delta v_{in} - A_v \bullet \Delta v_{in}).C_{sg}$$

$\boxed{Av \le 1}$

$$Q_{IN} = \Delta v_{in}.C_{dg} + (1 - A_v).\Delta v_{in} C_{sg}$$

$$Q_{IN} = \Delta v_{in}.C_{dg} \quad \text{as } Av \approx 1$$

$$C_{IN} = \frac{Q_{IN}}{\Delta v_{in}} = C_{dg} + C_{sg}(1 - A_v) \approx C_{dg}$$



The source-gate capacitance does not affect the input capacitance (unlike the CS that can have a very large input capacitance due to Miller effect). Can be used in amps that need low input cap.
Cons noise performance is poorer than CS, CD due to its low gain

Frequency behavior of SF

If we drive the SF with a source whose resistance is Rs, then the input pole will be

$$f_{in} = \frac{1}{2\pi R_s C_{in}} = \frac{1}{2\pi R_s C_{dg}}$$

A very high frequency pole due to low Cdg

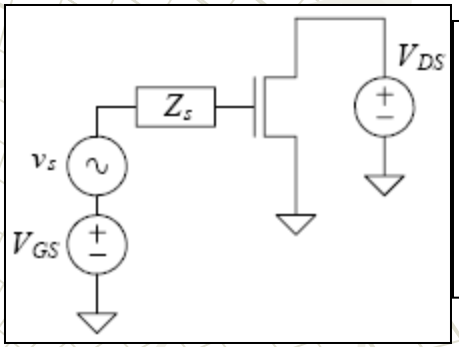
On the output

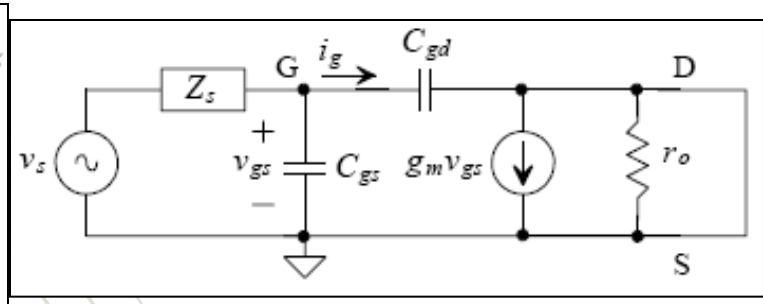
$$f_{out} = \frac{1}{2\pi R_{oSF} C_{load}} = \frac{g_m}{2\pi C_{load}}$$

$R_{oSF}=1/g_m$, *but* g_m *is freq dependent*

A high freq pole due to small Rout

To get how gm depends on freq





$$v_{gs}(f) = v_s . \frac{1/\, j2\Pi f(C_{gs}+C_{gd})}{\frac{1}{j2\Pi f(C_{gs}+C_{gd})} + Z_s}$$

$$= \frac{v_s}{1 + jZ_s.2\Pi f.(C_{gs}+C_{gd})}$$

$$i_d(f) = g_m.v_{gs}(f) = \frac{g_m.v_s}{1 + Z_s.j2\Pi f(C_{gs}+C_{gd})}$$

$$g_m(f) = \frac{g_m}{1 + Z_s.j2\Pi f(C_{gs}+C_{gd})}$$

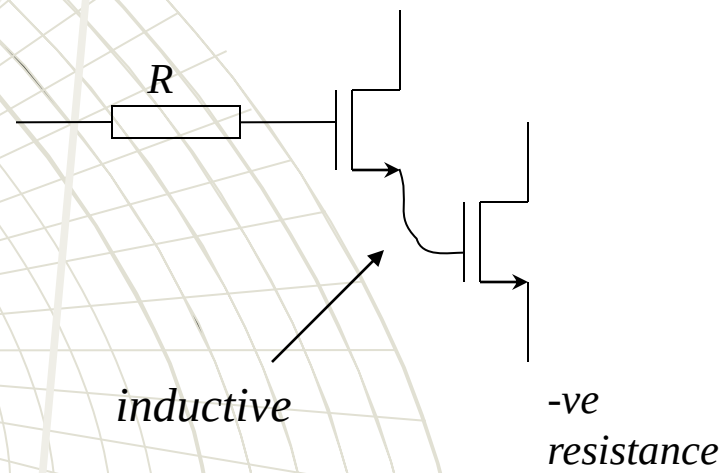
$$R_{out} = \frac{1}{g_m(f)} = \frac{1}{g_m}.(1 + j2\Pi f.Z_s(C_{gs}+C_{gd})$$

$$L_s = \frac{R_s(C_{gs}+C_{gd})}{g_m} \; if\; Z_s = R_s$$

Rout or R into Source is resistive at low freq and an R in series with an inductive load at high freq, may induce ringing while driving capacitance loads due to RLC

$$R_{out} = \frac{1}{g_m}\left(1 - \omega^2 .L.\left(C_{gs} + C_{gd}\right)\right) \quad if\, Z_s = j\omega L$$

Happens if we cascade two SF, Rout (inductive) of the first stage will be Zs of the second stage, giving –ve Rout of the second stage. Used in RF oscillator designs inductor placed in the gate of the MOS to AC ground to create –ve resistance



SF are used as output buffers, due to their low output resistance, and small input capacitance, also SF won't have much effect on the overall freq response of the amplifier that they buffer "high fin and fout"

CS with source degeneration resistor

$$A_v = \frac{v_{out}}{v_{in}} = \frac{i_d}{v_{in}} \cdot \frac{v_{out}}{i_d} = G_m . R_{out}$$

$$v_{in} = v_{gs} + v_s$$

$$i_d = g_{m1} v_{gs} - g_{mb1} v_s$$

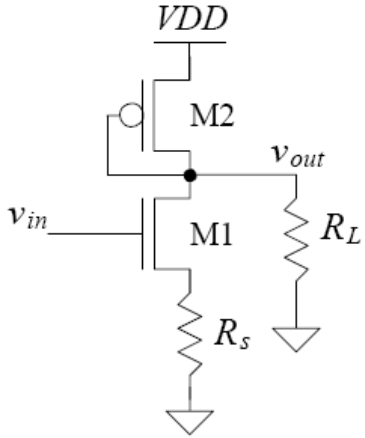
$$v_s = i_d . R_s$$

$$G_m = \frac{i_d}{v_{in}} = \frac{g_{m1}}{1 + (g_{m1} + g_{mb1}) R_s}$$

$$R_{out} \approx g_{m1} r_{o1} R_s \,//\, \frac{1}{g_{m2}} \,//\, r_{o2} \,//\, R_L \approx \frac{1}{g_{m2}}$$ → *M1 o/p resistance is gm1*Rs larger than its ro1, similar to the cascode action.*

$$A_v = G_m R_{out} = \frac{-g_{m1}}{1 + (g_{m1} + g_{mb1}) R_s} \cdot \frac{1}{g_{m2}}$$

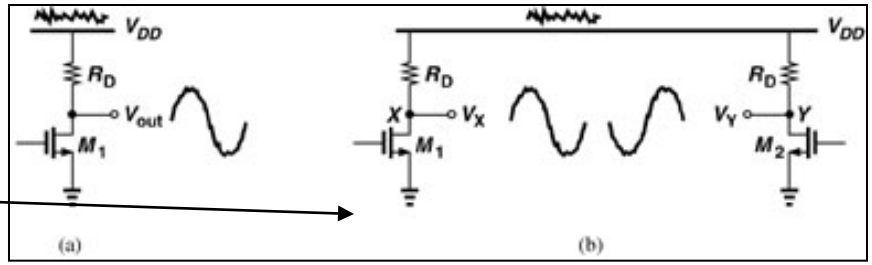
Rs decreased the gain and increased Rout of M1

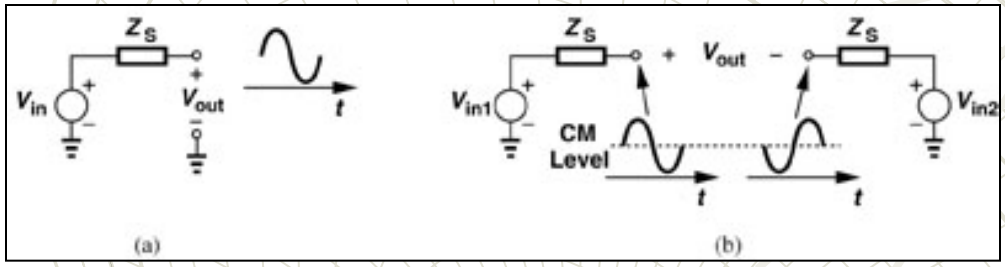


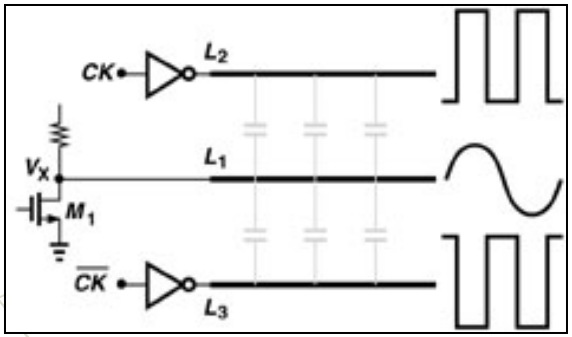
1) *What will be the gain if M2 is a current source load, not diode connected*
2) *What will be the gain if M2 is a current source load with infinite output resistance "ideal current source" and RL is also infinity. Comment on the results.*

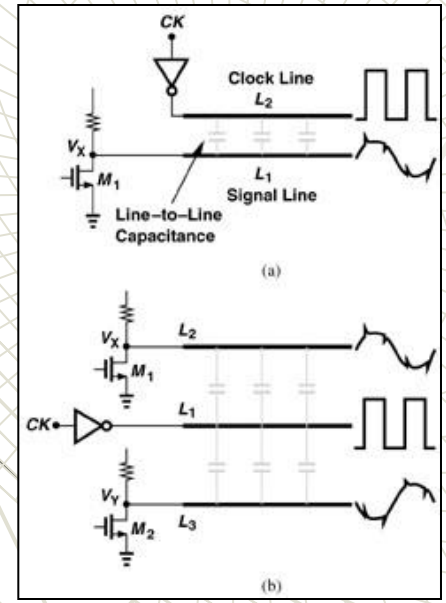
Differential vs. single ended

1) *Cmrr*
2) *Psrr*
3) *Voltage swing (vout=Vx-Vy)=2(Vdd-Vds,sat)*
4) *Cons, possible area increase, but it is worth it*
5) *Noisy diff, has an impact that cancels out on SE*









Why differential?

- Noise immunity {CMRR/PSRR}
- Larger output swing (Vx-Vy)
- Simpler biasing
- Higher linearity (remove even harmonics)
- Cons : area increase → worth it?

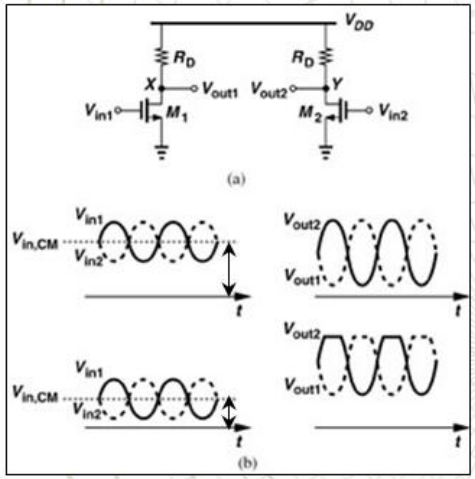
1) *why do we use tail bias current, not to impact output CM dependence on input common mode.*

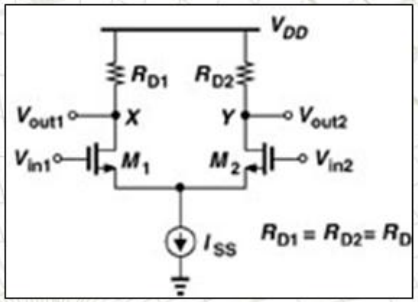
2) *Vdd, Vdd-iss*RD, indep on i/p cm*

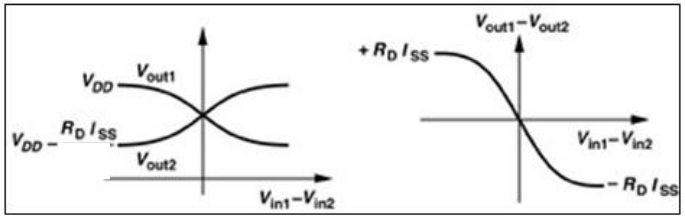
3) *Max linearity at vin1=vin2*

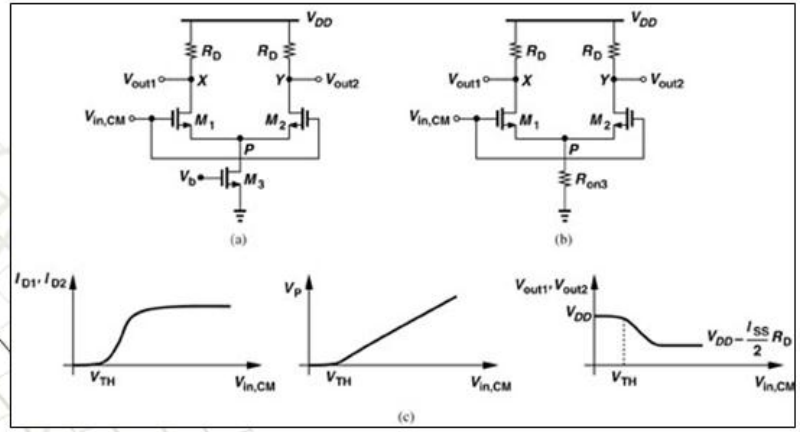
4) *The higher the i/p CM the lower the o/p swing, low i/p CM is desirable*

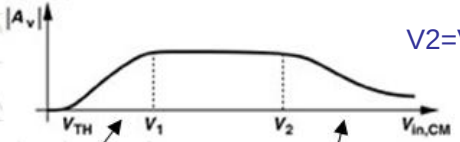
5) *Vin CM max and diff gain trade off*









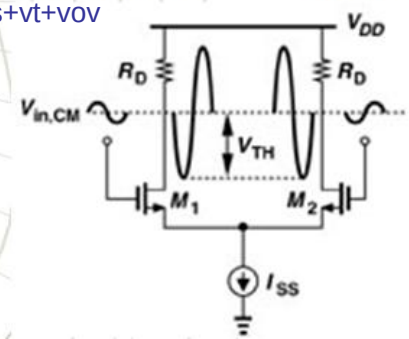


V2=Vt+Vdd-0.5*Rd

Tail current xtor triodes

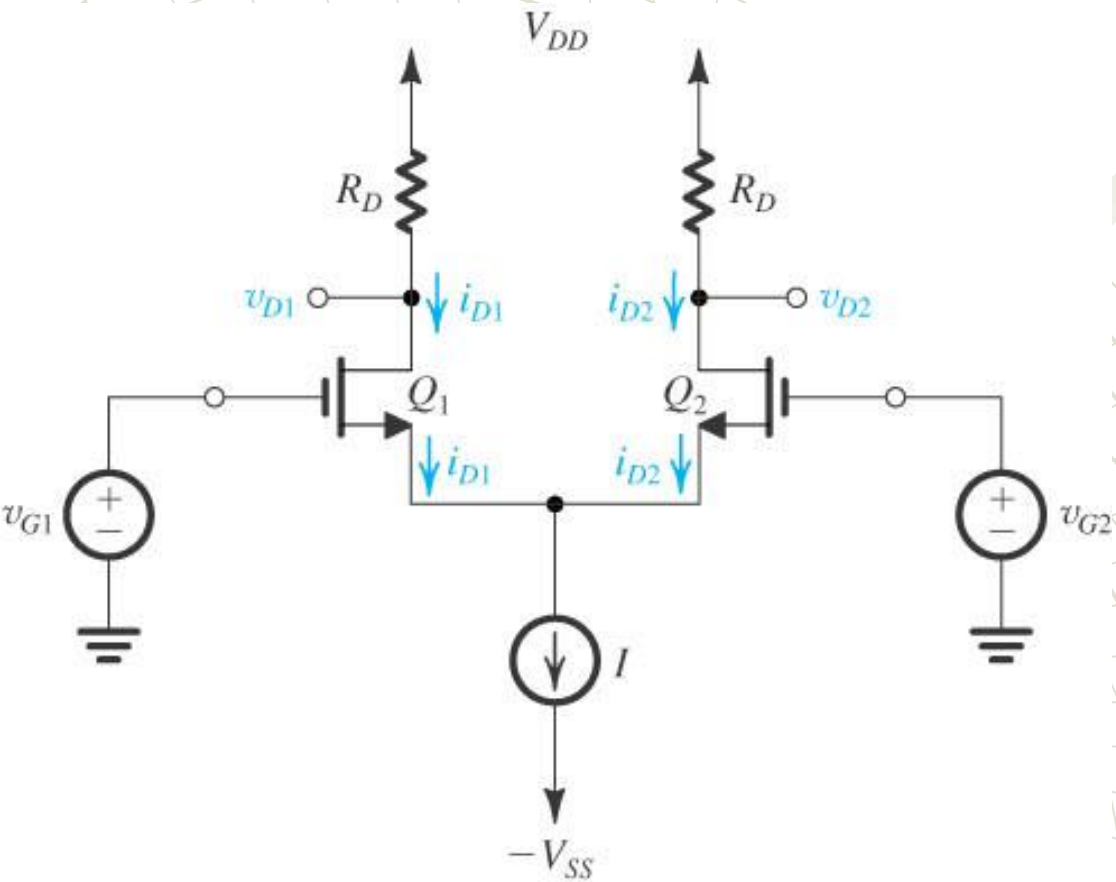
V1=-vss+Vcs+vt+vov

i/p xtor triodes



A. Helmy, PhD

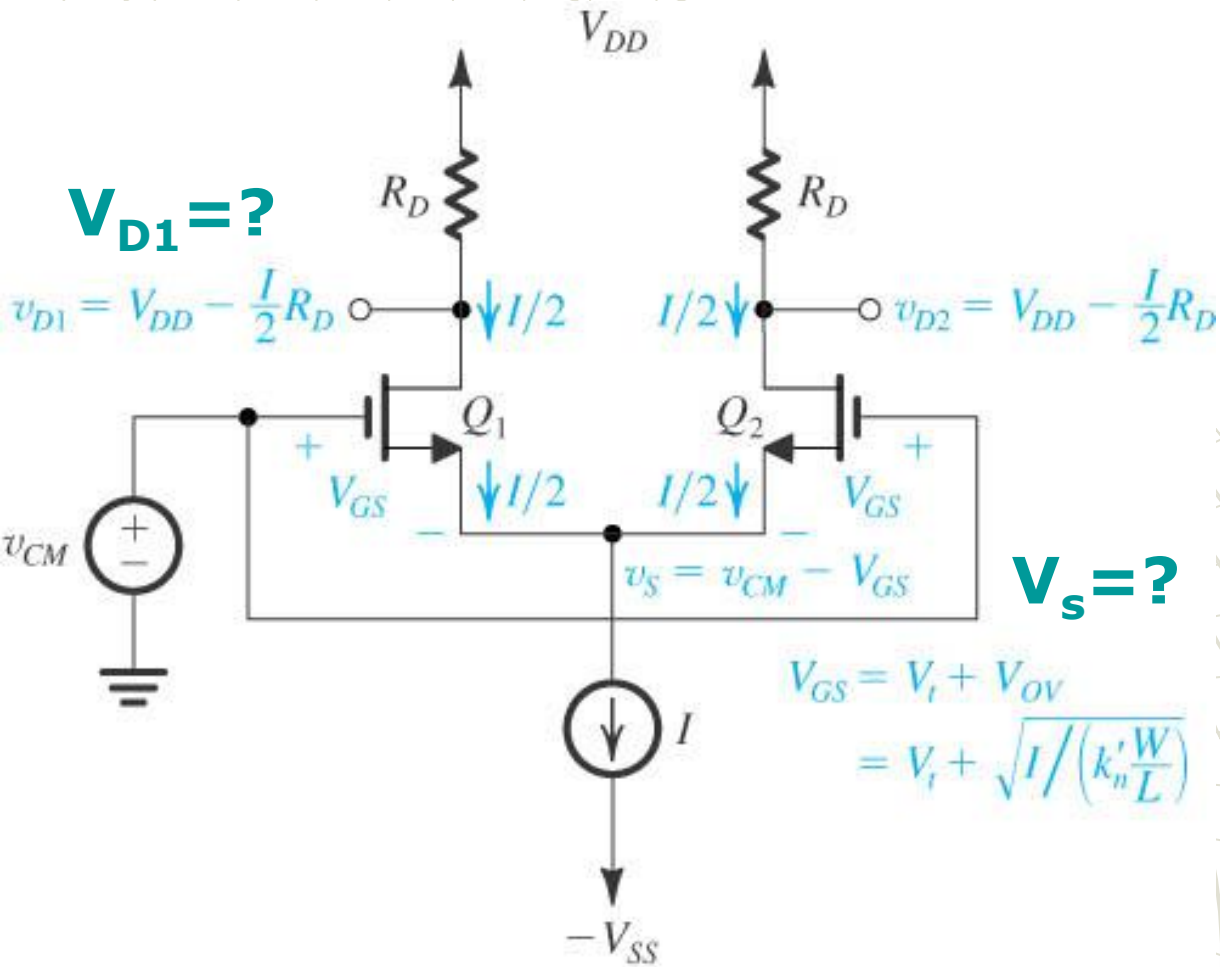
Large Signal Operation



- Diff IN
- Diff Out

- $V_{G1}=V_{CM}+V_{id}/2$
- $V_{G2}=V_{CM}-V_{id}/2$
- Vid=?
- VCM=?

Common mode Bias (Same I/P to both sides)



V_{D1}=?

V_s=?

$$\frac{I}{2} = \frac{1}{2} k_n' \frac{W}{L} (V_{GS} - V_t)^2$$

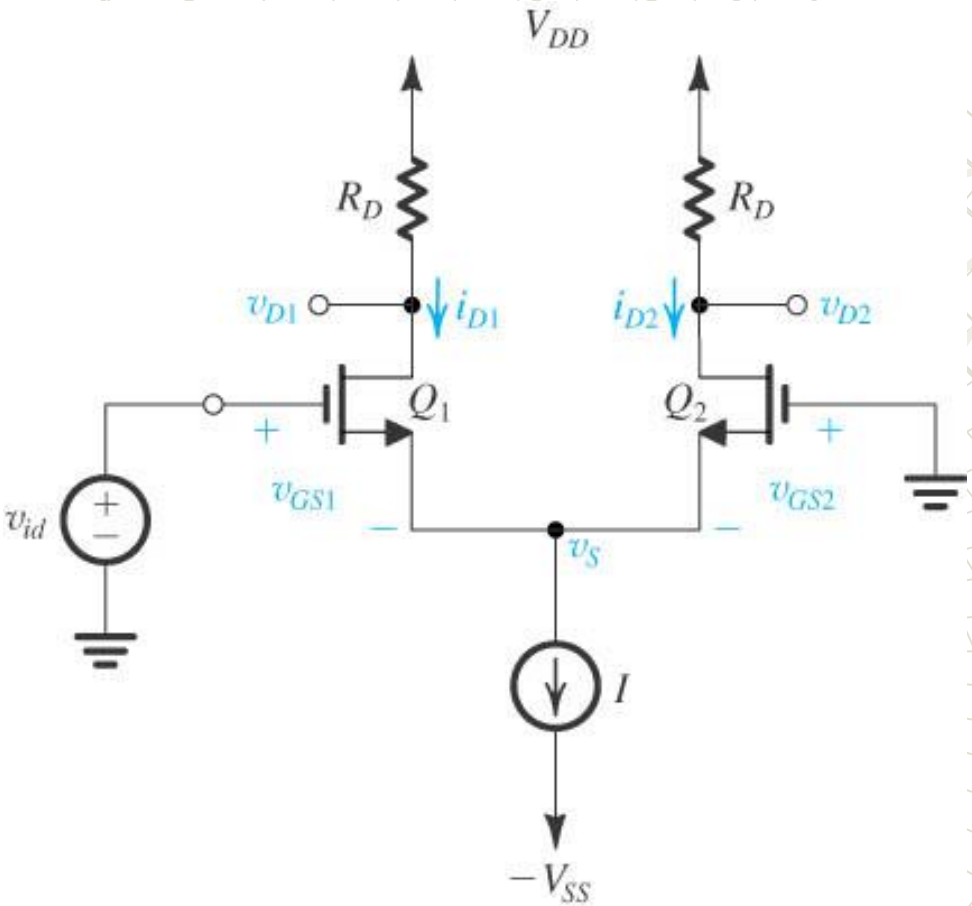
$$\frac{I}{2} = \frac{1}{2} k_n' \frac{W}{L} V_{OV}^{\ 2}$$

$$V_{OV} = \sqrt{I / k_n' (W/L)}$$

V_{CMmax}=?

V_{CMmin}=?

Differential Input Signal $V_{OV} = \sqrt{I/k_n^{'}(W/L)}$



- $V_{id}=V_{GS1}-V_{GS2}$
- If $i_{D1}=I$: Q_2 off
- Vs=-Vt
- I=1/2 KW/L (vgs1-vt)^2

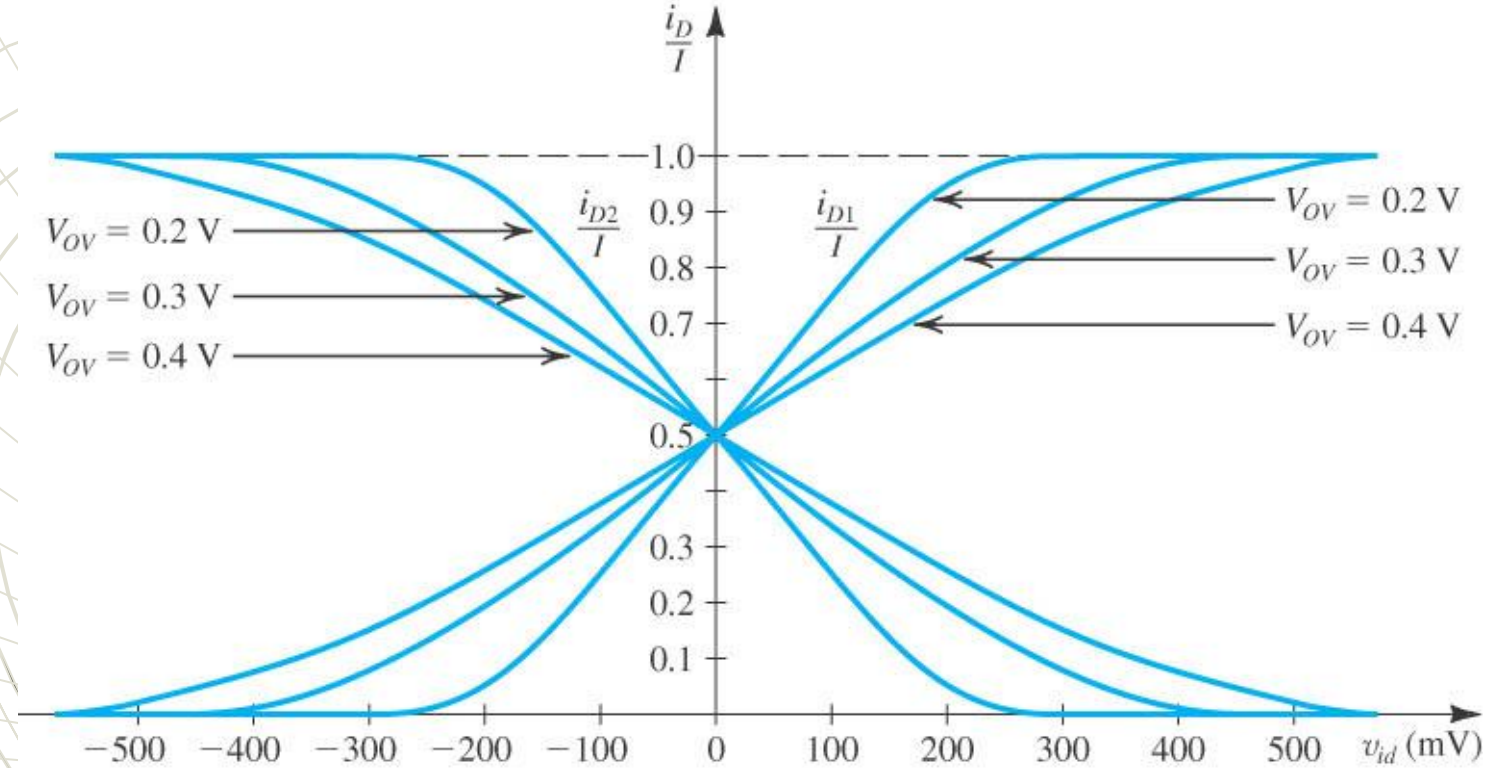
$$v_{gs1} = Vt + \sqrt{2}V_{OV}$$

$$v_{id\max} = vgs1 + vs$$

$$v_{id\max} = \sqrt{2}V_{OV}$$

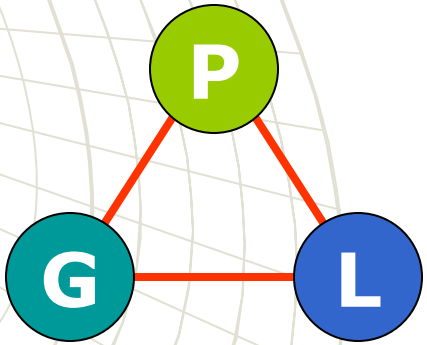
$$-\sqrt{2}V_{OV_{CM}} \le v_{id} \le \sqrt{2}V_{OV_{CM}}$$

Impact of Overdrive



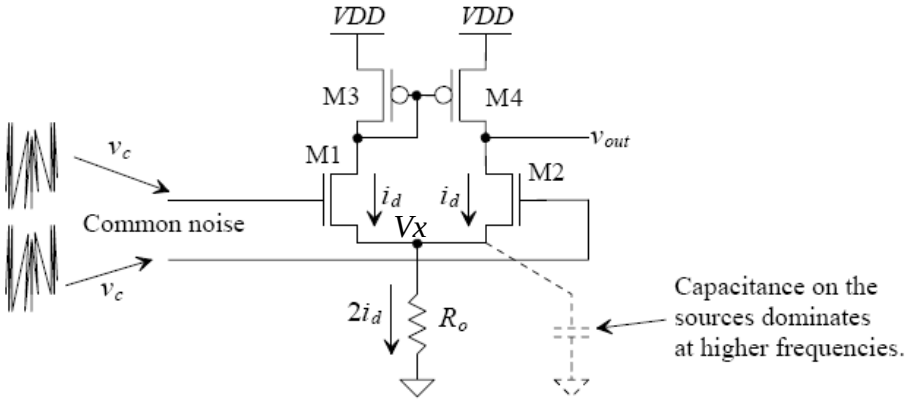
High V_{ov} -> High Linearity

How to get high V_{ov}?



Common mode rejection ratio CMRR

An imp aspect of the diff amplifier is its ability to reject common signal applied to both inputs. Often, in analog systems noise is coupled as common mode (due to its random nature) and it is desirable to reject this noise



Both id1 and id2 flow opposite to each other, if Ro=∞ they should cancel each other, CMRR= ∞, for Ro ≠ ∞ it will modulate Vx node

Very imp to keep this C low

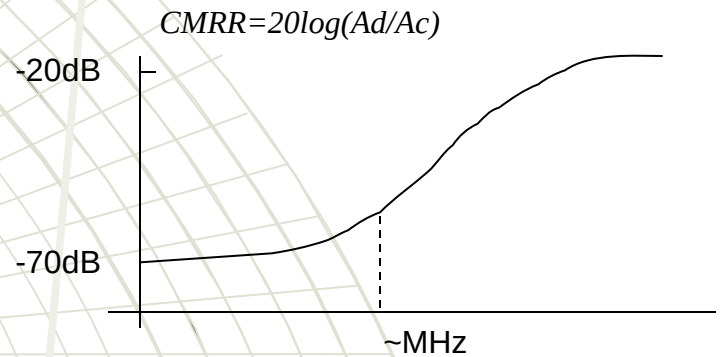
$$v_c = v_{gs1,2} + 2i_d R_o$$

$$v_c = \frac{i_d}{g_m} + 2i_d R_o = i_d\left(\frac{1}{g_m} + 2R_o\right) \approx i_d 2R_o$$

$$v_{out} = -i_d \frac{1}{g_{m3}} = -i_d \frac{1}{g_{m4}}$$

$$A_c = \frac{v_{out}}{v_c} = \frac{-i_d}{g_{m3,4} i_d 2R_o} = -\frac{1}{2g_{m3,4}R_o}$$

$$CMRR = 20.\log\left|\frac{A_d}{A_c}\right| = 20.\log\left[g_{m1,2}(r_{o2} \parallel r_{o4}).2g_{m3,4}R_o\right]$$

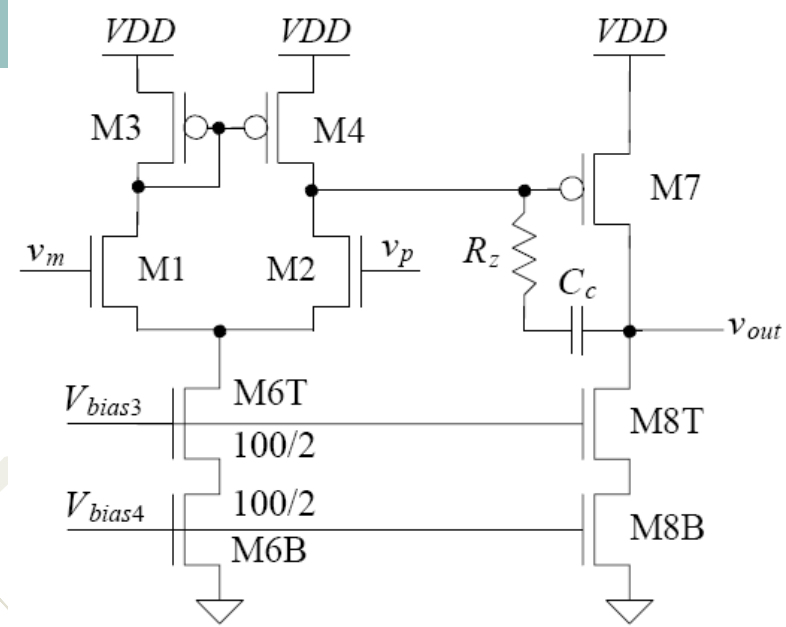


The higher the tail current source resistance, the lower the common mode gain gets – it is good for op-amps. As common mode frequency increases, Co becomes dominant. Cascoding helps?

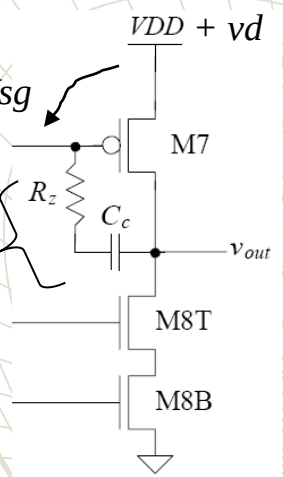
How about high frequency noise?

Consider the same two stage amp:

At low frequencies mostly bias and asymmetries cause PSRR. At high frequencies, output stage dominates



Low imp at high freq



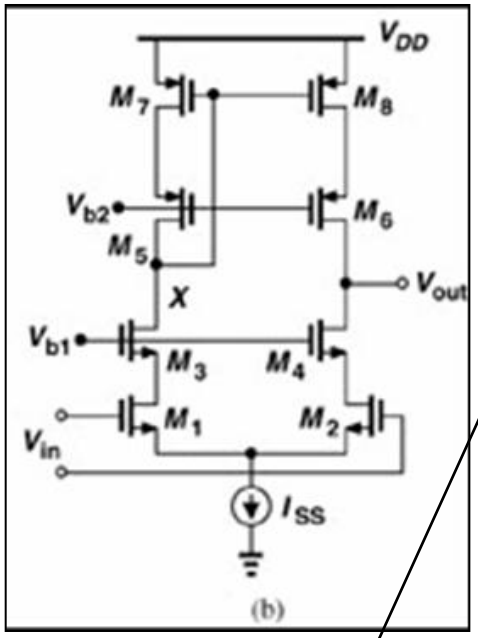
Δvout=vd+ Δvgs

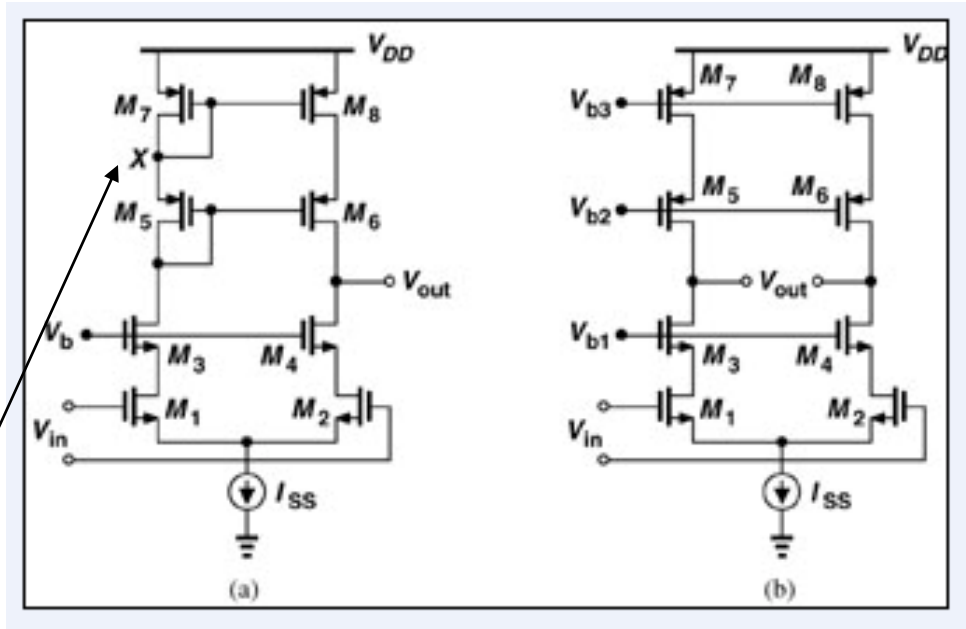
Assuming I is fixed

Δvout≈vd, i.e Ap=1

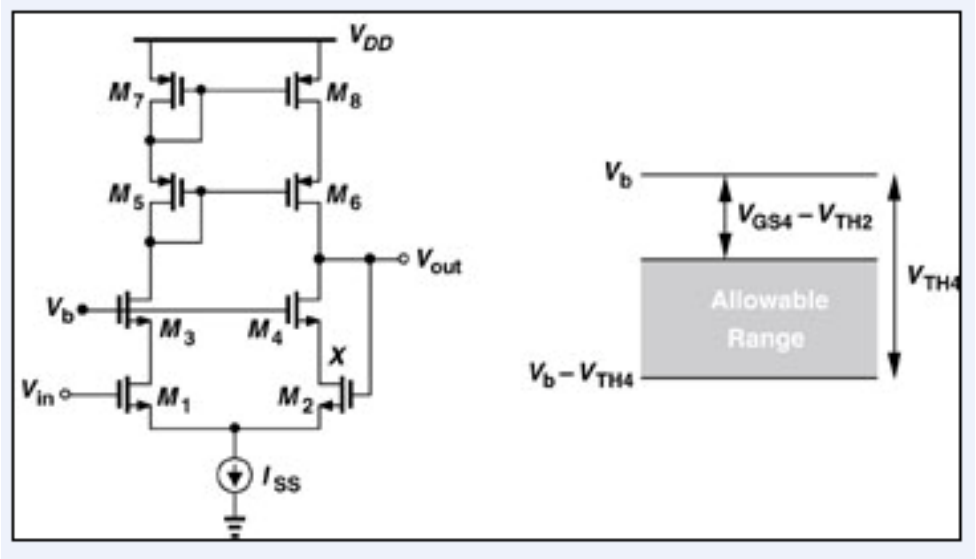
As open loop gain A_D reduces, assuming Ap=1, PSRR reduces to 0dB at high frequency

Draw back of telescopic cascode amplifiers





- *Limited output swing*
  - *Single ended worse than diff for 2 reasons (se vs. diff, diode connected vs. current mirror)*
- *Mirror pole in single ended node X*
- *Cannot be used as a unity gain buffer due to limited swing and high delta in i/p vs. o/p CM*
- *Vout<vx+vth2 and vout>vb-vth4*
- *Since vx=vb-vgs4*
- *Vb-vth4<vout<vb-vgs4+vth2*
- *Vmax-vmin = vth4-(vgs4-vth2)*



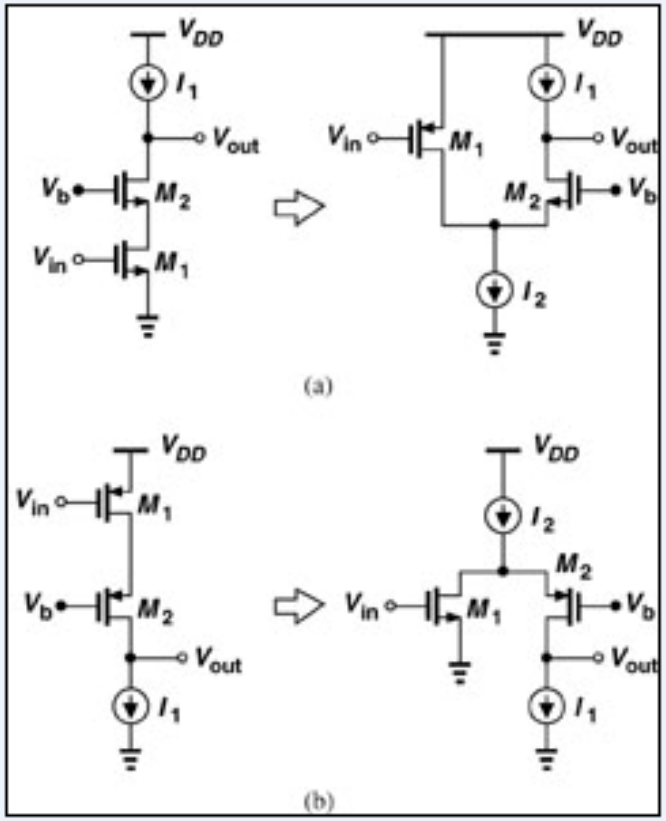
Folded cascode amplifier

To overcome the drawbacks of telescopic cascode (limited o/p swing, and difficulty in implementing unity gain buffers)

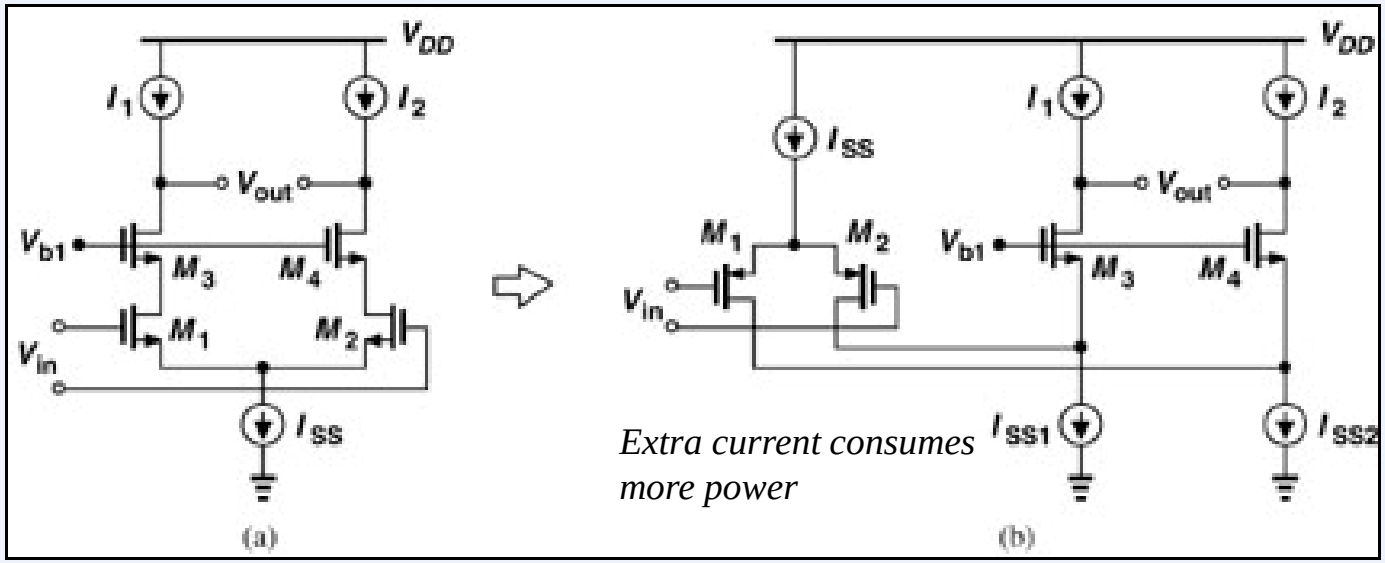
*i/p device is replaced by the opposite type without stacking. Very versatile **single stage amplifier**. Since it is single stage, it requires only output compensation, no CC and hence good PSRR (supply and gnd noise shielded by the cascode)*

(a)i/p CM cannot exceed vb1-vgs3+vth1 or M1 will triode

(b)i/p CM cannot be less than vb1-vgs3-vth1, thus we can short i/p and o/p with no swing limitation



Refreshing 'common-gate' and 'transimpedance' topologies:



Extra current consumes more power

Folded cascode

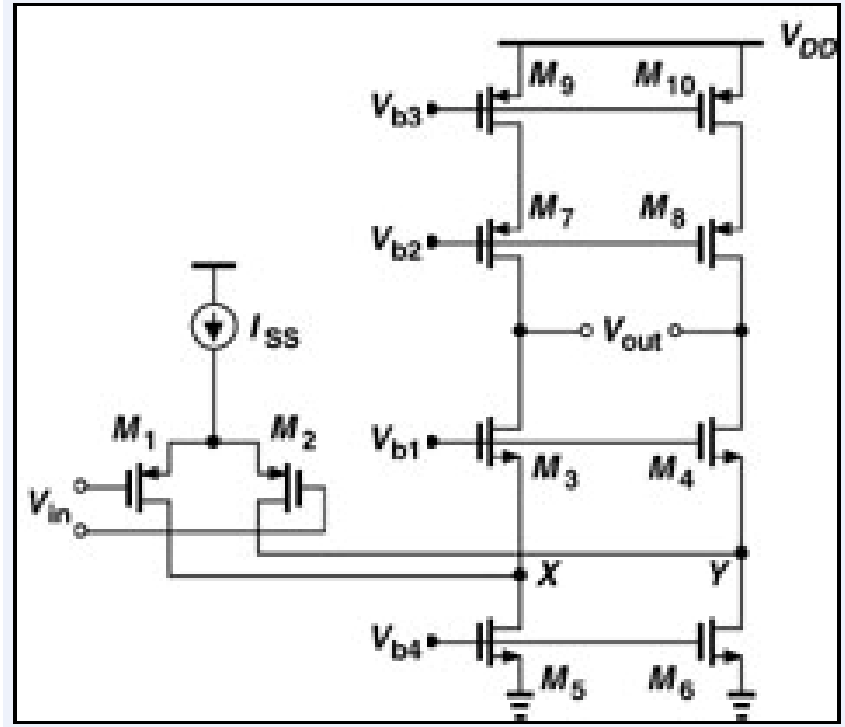
Vout min = vod3+vod5

Vout max= vdd-(vod7+vod9)

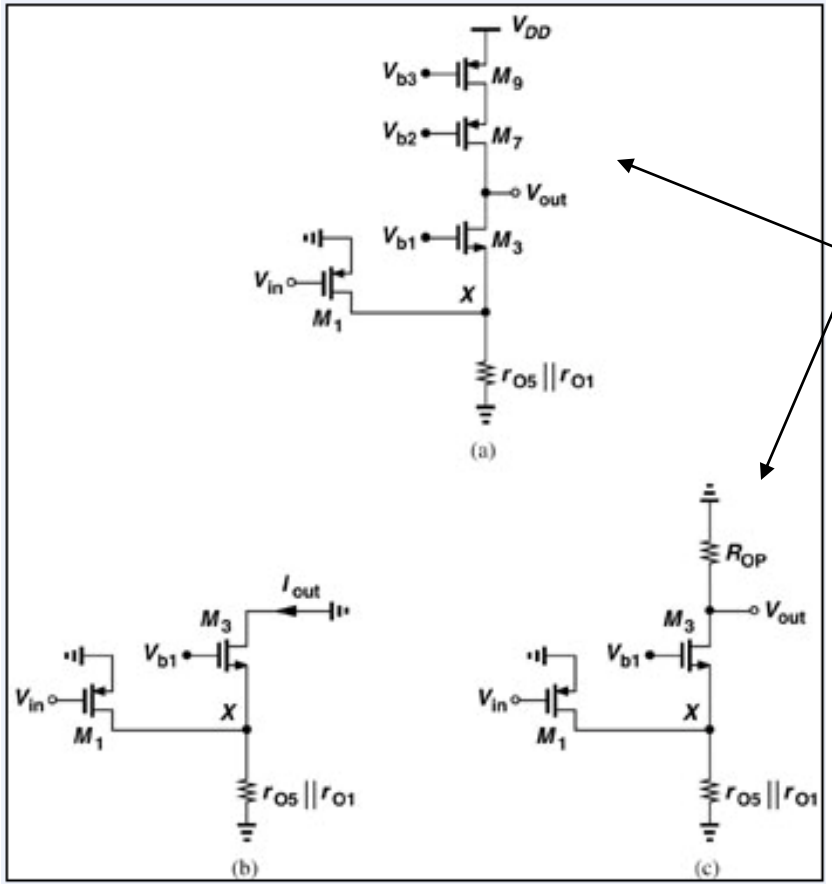
p-p swing is vdd- (vod3+vod5+vod7+vod9)

Telescopic cascode was less by the overdrive of the tail current transistor

But M5 and M6 will carry high current and their over drive will be higher if their cap contributions at nodes X and Y is to be minimized, thus swing of folded cascode is ***<u>slightly</u>*** *> telescopic*



1



Av=Gm Rout,

Gm =gm1 since iout is I1 and 1/gm3<<ro1||ro5

Rop=gm7ro7ro9

Rout=Rop||gm3ro3(ro1||ro5)

Av=gm1{gm7ro7ro9|| gm3ro3(ro1||ro5)}

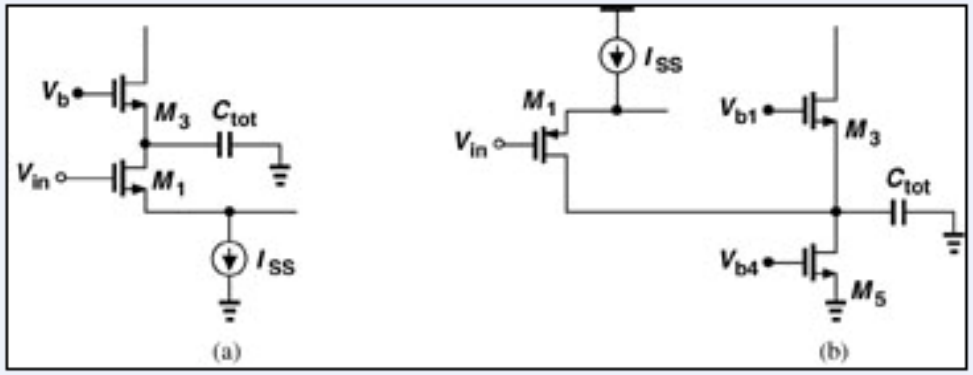
ro1 is in || with ro5 hence

Av folded < Av telescopic by 2-3X

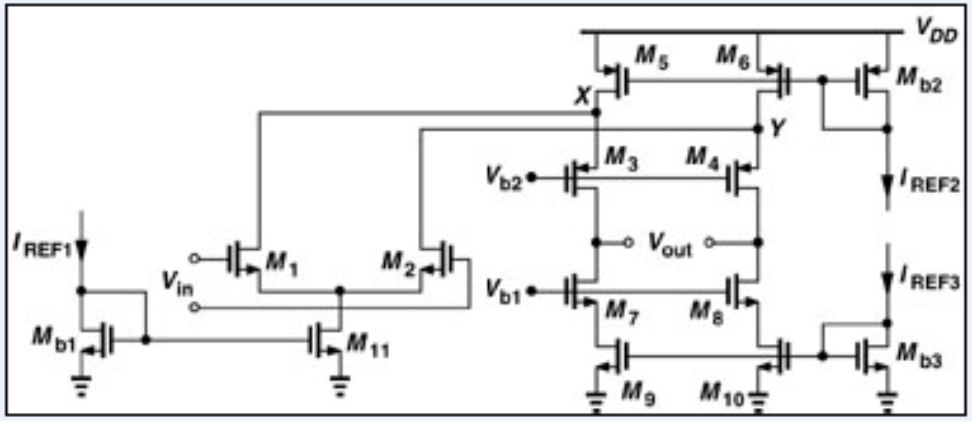
2

Pole at folding point is more dominant than the corresponding pole of the telescopic cascode. (note that M5 is a large device carrying high current). Folded cascode is a bit slower than telescopic cascode when unloaded with Cl. When loaded with Cl the dominant pole that limits the BW is that of the output node.

3



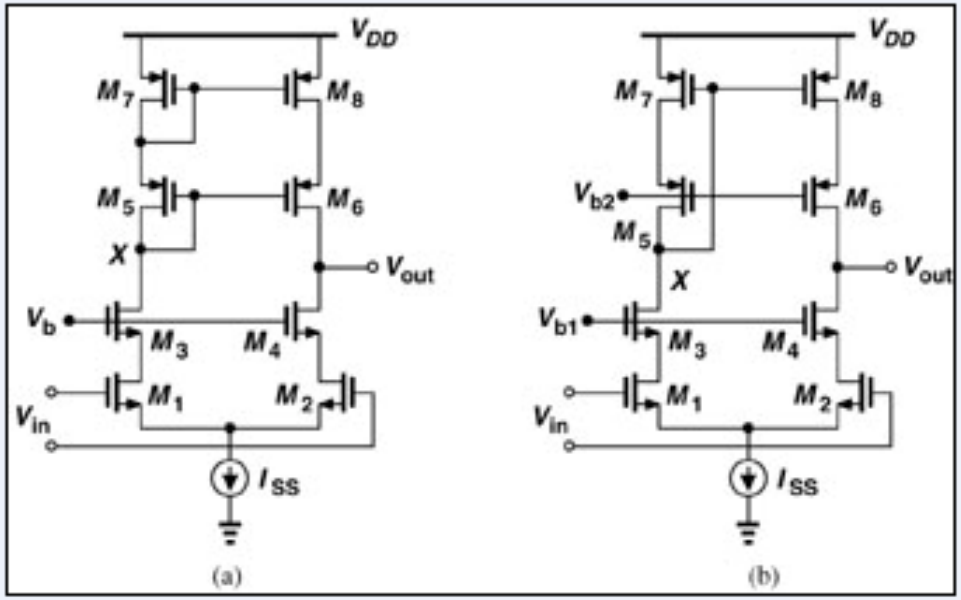
NMOS i/p with PMOS cascode will have higher gain than a PMOS i/p with NMOS cascode because of higher mobility of NMOS (higher gm) but this comes at the expense of lower BW (why?) because mirror pole of node X and Y is lower than the PMOS i/p with NMOS cascode case due to higher R and C of node X&Y (low gm3 higher 1/gm3, and bigger M5 high C (wider to carry the same current as its NMOS counterpart due to its low mobility)



SE o/p cascode.

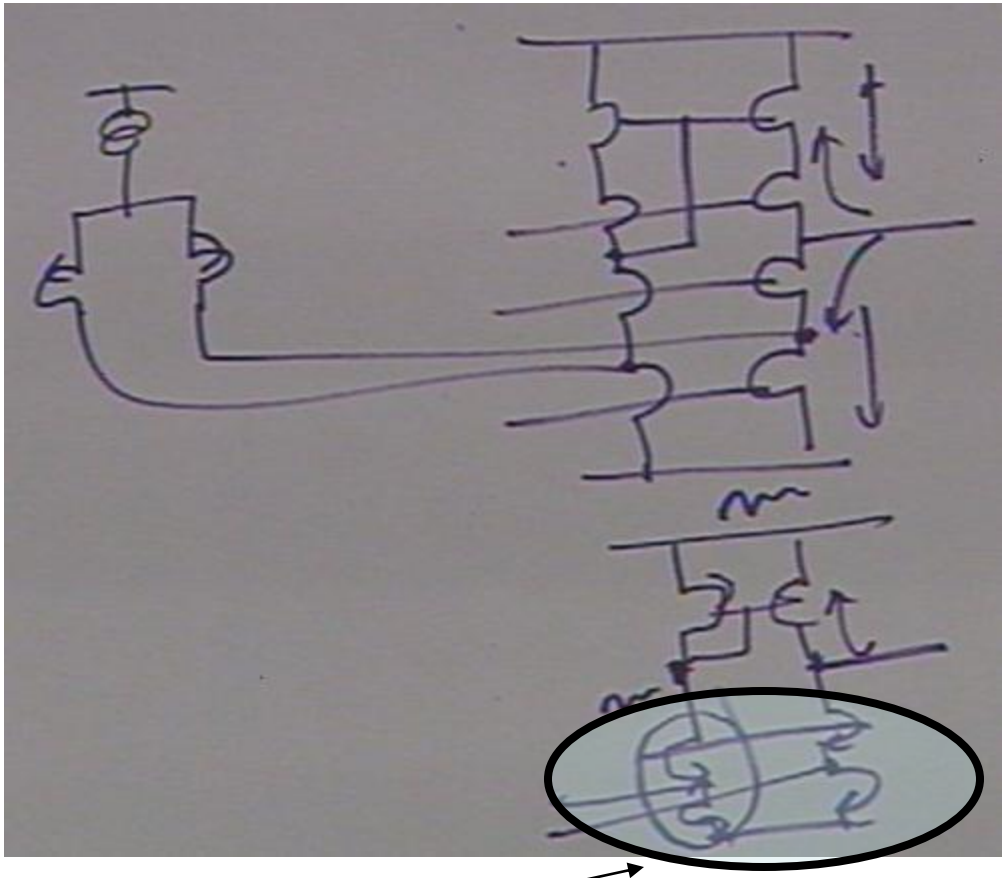
Lower BW due to lower mirror pole at X much higher C than fully differential case

Half the swing of the fully differential



Folded cascode
with cascode load

Folded cascode
with single load,
better swing, but
lower gain



Folding must have a cascode to fold to,
i.e ***this*** must be a cascode

OTA, telescopic cascode, folded cascode are all suitable for driving C loads but not R load, we need an o/p buffer

Critical Specifications For Amplifiers

Distortion:

Small signal (AC) analysis assumes signal swings (voltages and currents) are much smaller than DC operating points. When signal swings increase to a point where voltage swings are rail-to-rail, small signal assumption may not hold true.

Consider a common source amplifier;

$$Av = gm1.(ro1 \| ro2)$$

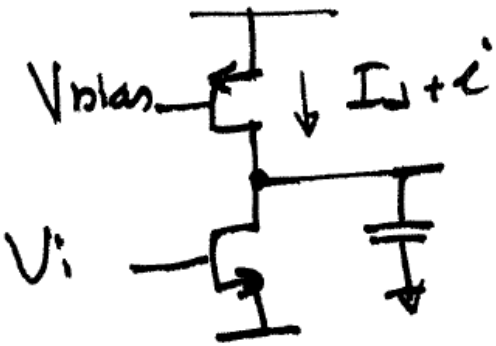
$$= \frac{\sqrt{2.k'.(W/L).(Id+i)}}{gds1+gds2}$$

$$= \frac{\sqrt{2.k'.(W/L).(Id+i)}}{2.(\lambda n+\lambda p)(Id+i)}$$

$$= \frac{\sqrt{2.k'.(W/L)}}{2.(\lambda n+\lambda p)\sqrt{(Id+i)}}$$

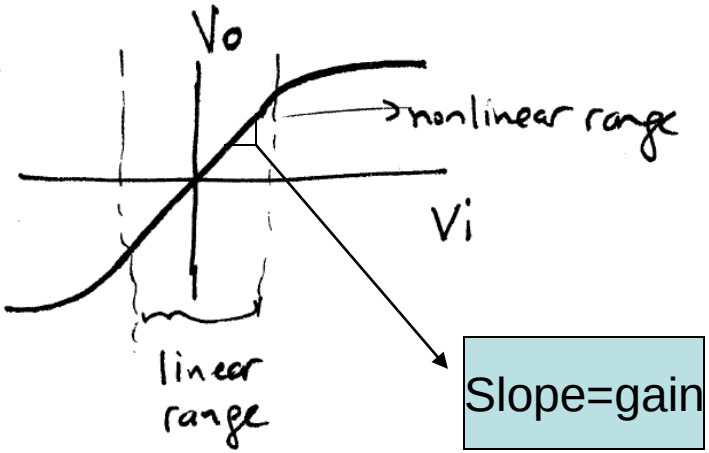
More current Id means more linear performance. Power vs. linearity trade off

Can we check for distortion using ac analysis in spectre?



If Id and i are comparable, $\frac{1}{\sqrt{(Id+i)}}$ *component becomes highly non linear. Distortion is characterized by voltage gain becoming a function of input signal Vi or i*

When gain decreases, distortion increases, distortion increases with freq because gain drops with frequency



Frequency Domain Characterization Of Distortion

Apply a spectrally pure sine way to the input.

$$Vin(t) = Vp.\sin(2.\pi.f.t)$$

The amplifier will generate a series of tones due to the $\sqrt{}$ *term.*

$$Vout(t) = a1.Vp.\sin(2\pi ft) + a2.Vp.\sin(2.2\pi ft) + a3.Vp.\sin(3.2\pi ft) + ... + an.Vp.\sin(n.$$

$ai.Vp.\sin(i.2\pi ft)$ *is the ith harmonic of the input.*

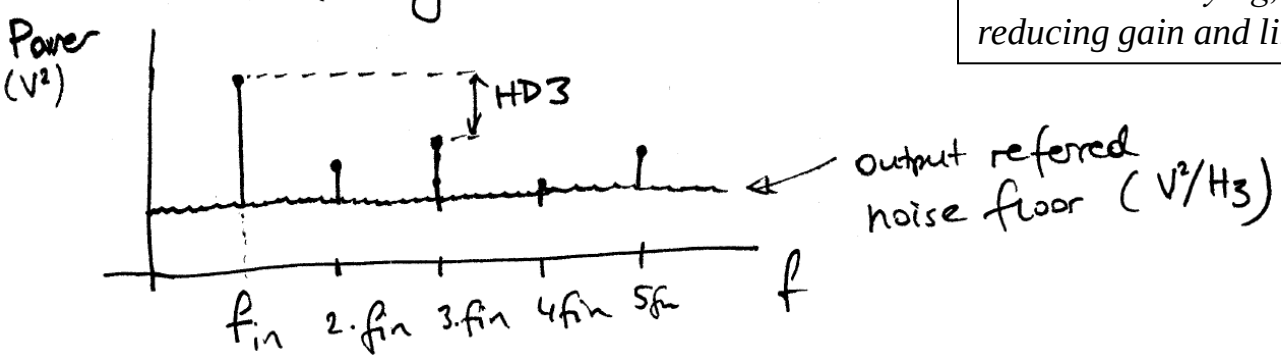
$a1.Vp.\sin(2\pi ft)$ *is the fundamental of the input.*

The nth term harmonic distortion is given by HDn=an/a1, for n>1

Total harmonic distortion is: $THD = \sqrt{\dfrac{a_2^{\,2} + a_3^{\,2} + ...a_n^{\,2}}{a_1^{\,2}}}$

o/p buffers, amps used to drive a large load capacitance or low resistance, are examples of amps where low THD is imp. If the o/p amp is biased such that DC is >> time varying, too much power is lost. If Bias ≈ time varying, fb is used to reduce distortion by reducing gain and linearizing the amp

In frequency domain:



Gain Enhanced (Gain boosted Amps)

Regulated Drain Mirrors Revisited (Chap 20)
Remember the wide swing cascade.
Now let us stabilize the drains of current sources with op-amps.

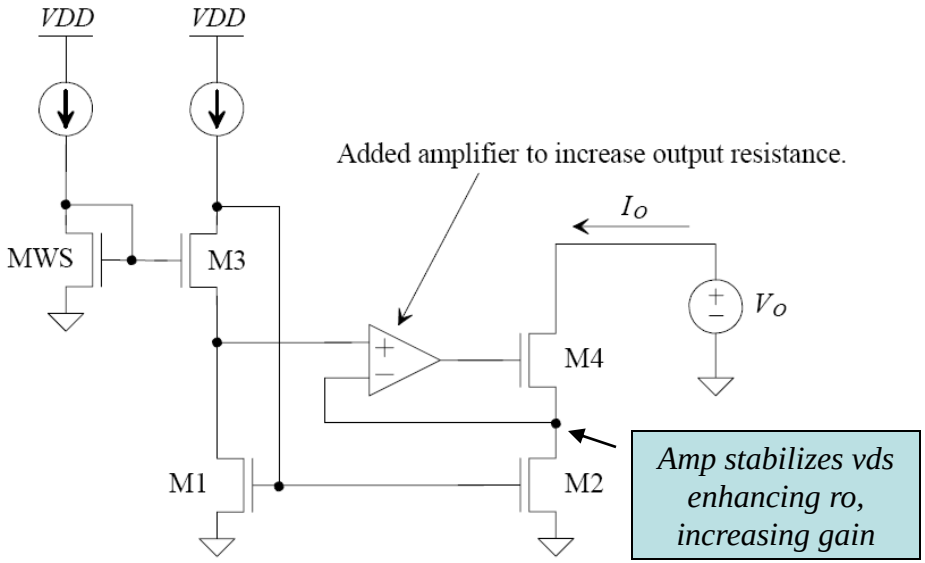


Figure 20.40 Regulating the drain of M2 using an amplifier.

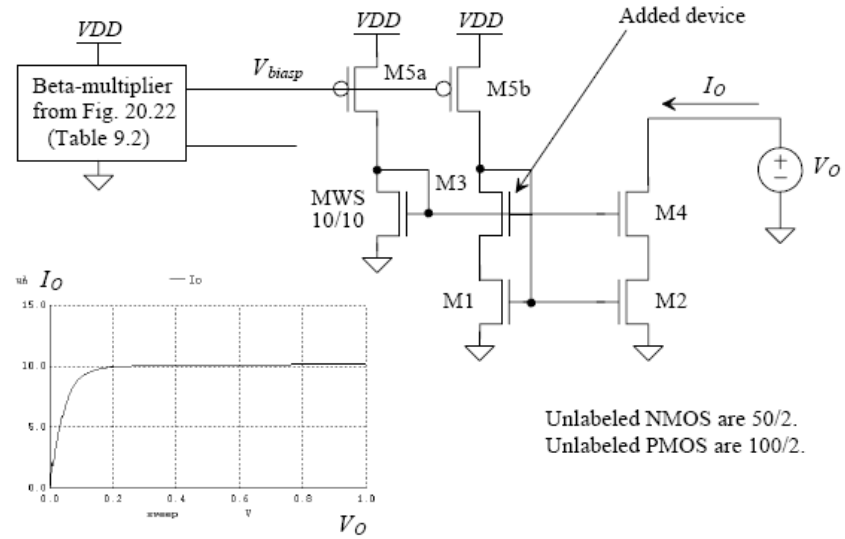
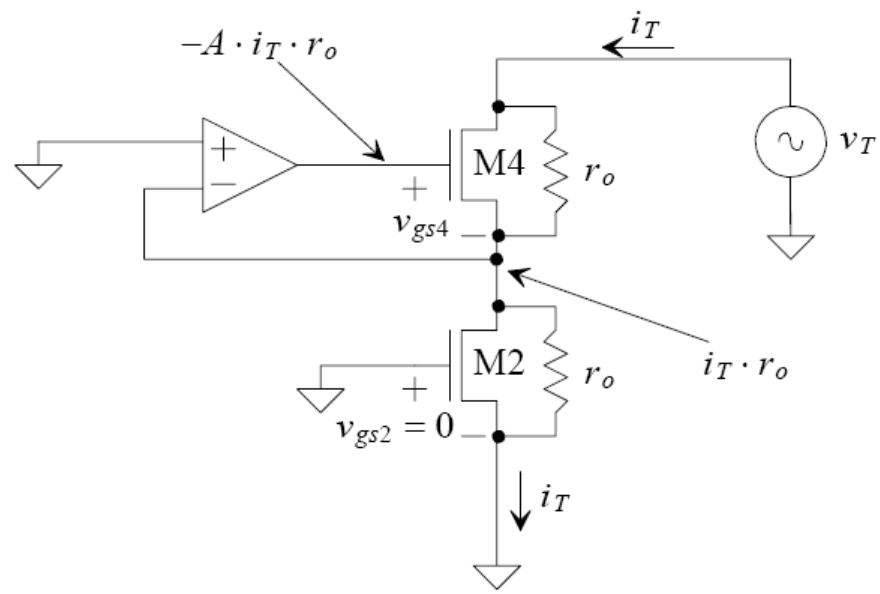


Figure 20.38 Wide-swing cascode current source in the short-channel process (good). Notice the drain-to-source voltages of M1 and M2 are the same.



$$v_{gs4} = -i_t \cdot r_o \cdot (A+1)$$

$$i_T = g_m \cdot \left(-i_t \cdot r_o \cdot (A+1)\right) + \frac{v_t - i_t \cdot r_o}{r_o}$$

yielding

$$R_o = \frac{v_t}{i_t} = 2r_o + g_m r_o^2 \cdot (A+1) \approx \underbrace{g_m \cdot r_o \cdot A}_{A \times \text{cascode self gain}} \cdot r_o$$

Gain of the amplifier A, multiplies the cascade Rout.

A simple implementation is to use a CS amp for the amplifier.

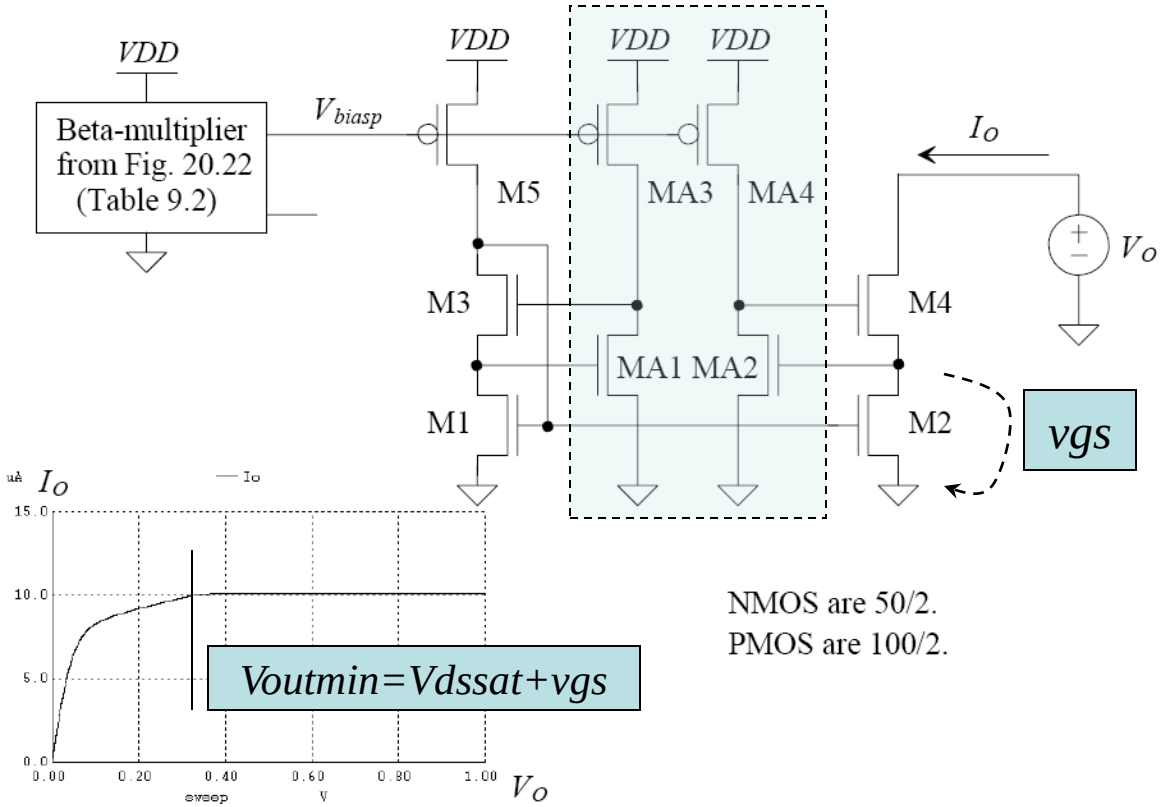


Figure 20.42 Using amplifiers to regulate the drain potentials in a current mirror.

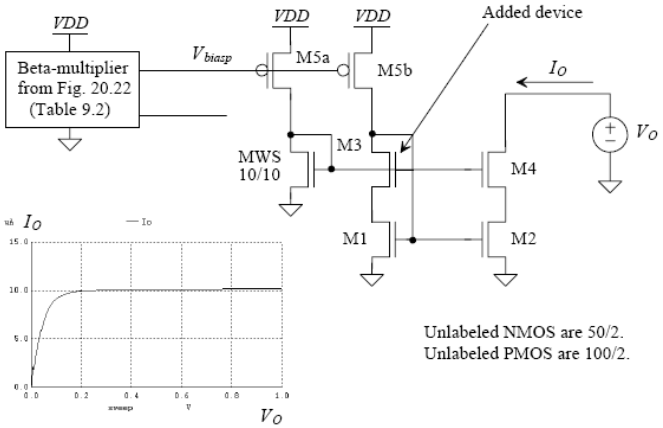


Figure 20.38 Wide-swing cascode current source in the short-channel process (good). Notice the drain-to-source voltages of M1 and M2 are the same.

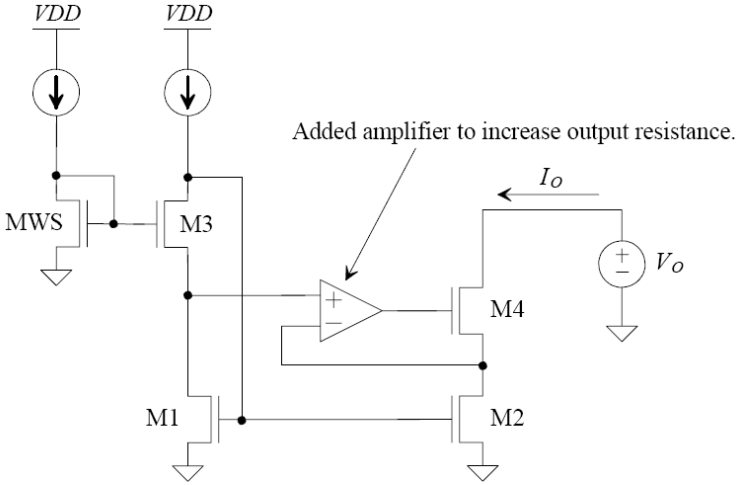


Figure 20.40 Regulating the drain of M2 using an amplifier.

$A=gm.ro/2$

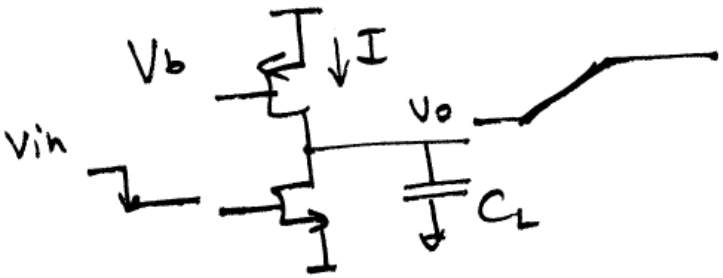
$Rout=gm.ro/2*gm.ro^2$

$Rout=gm^2.ro^3/2$

Enhancing the gain of a single stage and increasing its rout make it more stable, pole more dominant, bw of the inserted amp > f3db of the main amp, cl>> cjun

Slew Rate Limiting In Amplifiers

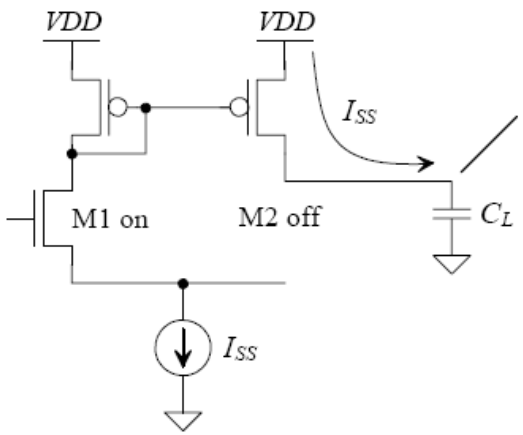
Slew rate is a "large signal" problem. It is associated with load capacitance.
Consider a CS stage. When input transitions from high to low, the fastest rate output which makes a low to high transition is the slew rate.

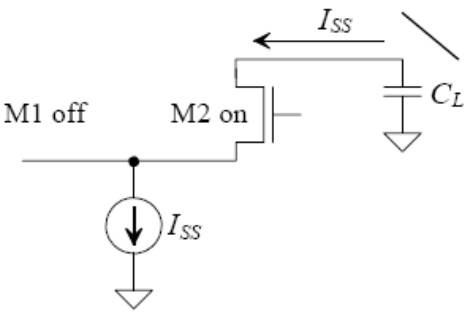


$$I = C_L . \frac{dVout}{dt}$$

$$\frac{dVout}{dt} = SR = \frac{I}{C_L}$$

This is not a step response, where input is a small step. This is a large signal step input.





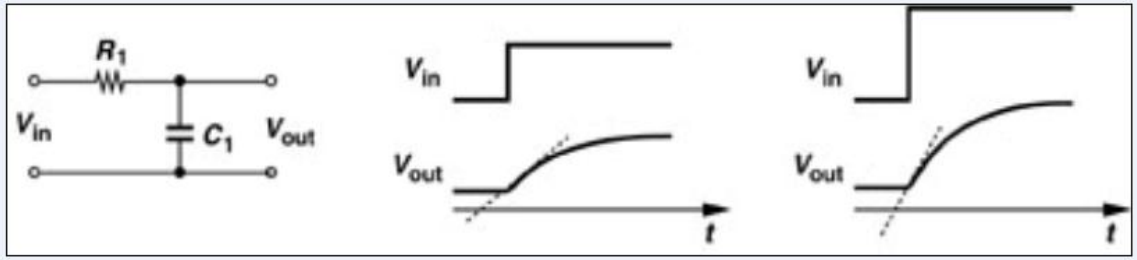
$$\frac{dVout}{dt} = \frac{I_{SS}}{C_L}$$

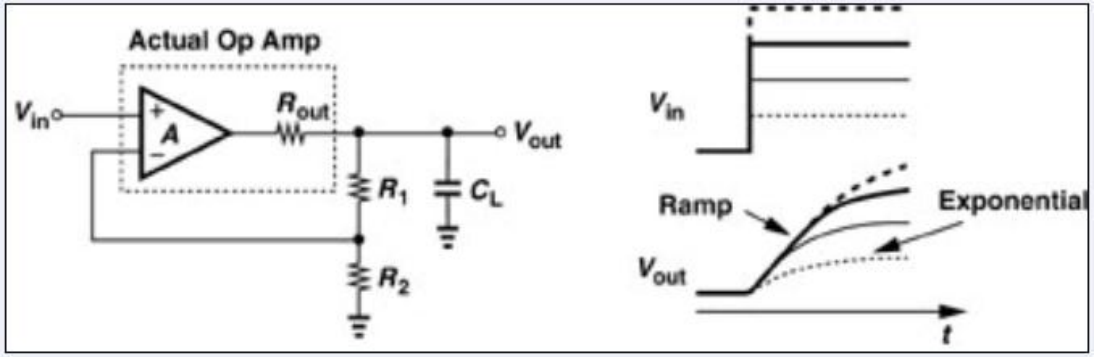
Indep on i/p step size

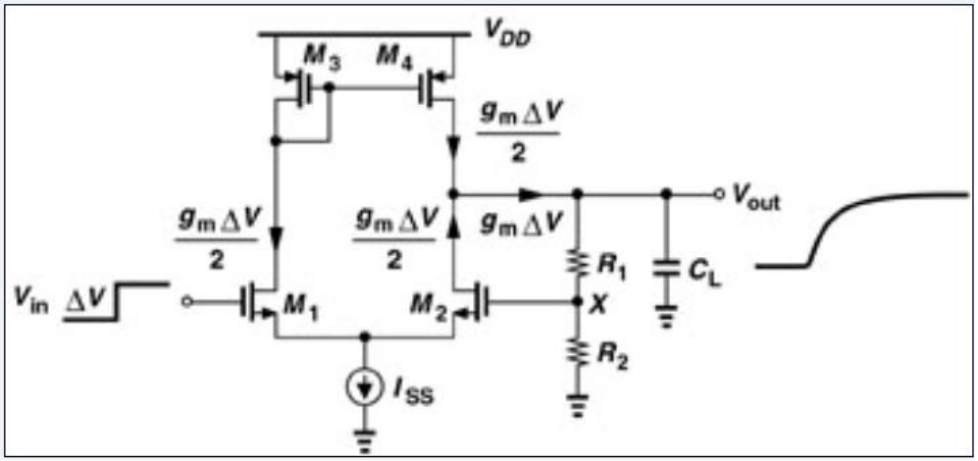
$$V_{out} = V_o\left(1 - e^{-t/\tau}\right)$$

$$\frac{dV_{out}}{dt} = \frac{V_o}{\tau}.e^{-t/\tau}$$

$$\tau \alpha \frac{1}{f_{un}} \alpha \frac{C_L}{g_m}$$







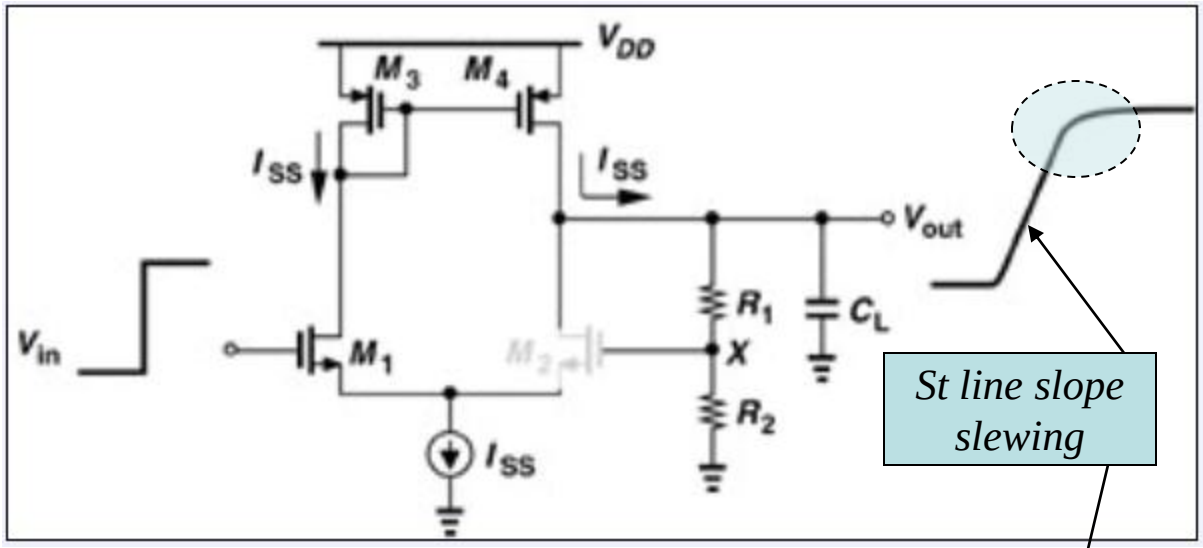
Note that the output slope is time dependent and that it is proportional to input step value Vo, unlike SR which is an output slope that is independent on the input step value, rather it depends on Iss and CL.

Linear settling τ is related to the small signal (linear) BW of the amp

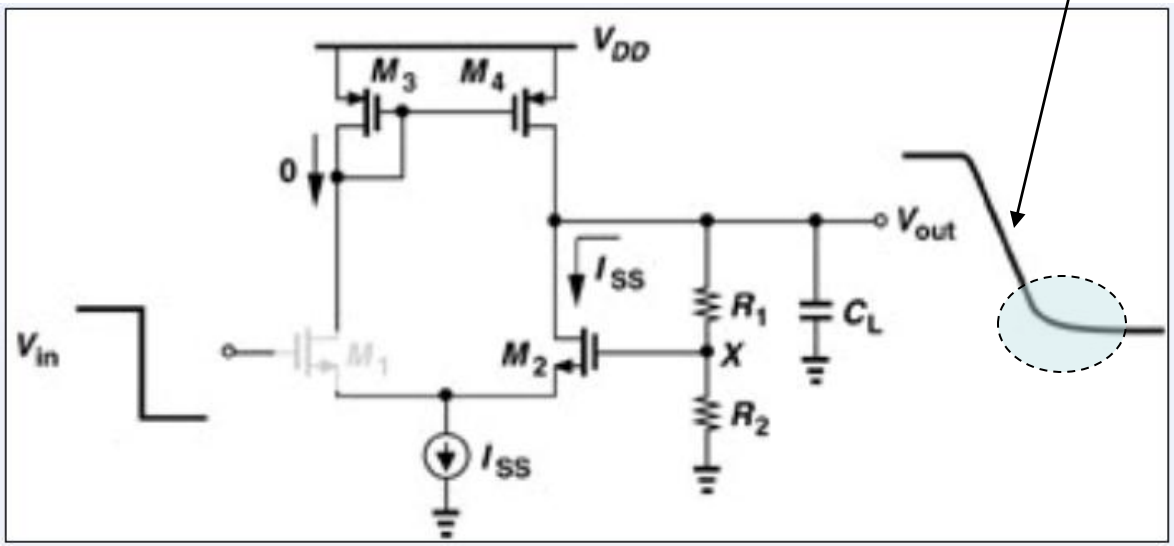
On the other hand SR is a non linear phenomena hence it causes distortion

Linear settling time limited due to small value of the input step. Note that diff pair is in saturation

SR limited due to large value of the input step, note that diff pair is in (triode M1, and off M2) till vout increases by slewing and Vx rises opening M2 and entering in the linear setting zone where the slope is limited by settling time τ



St line slope slewing

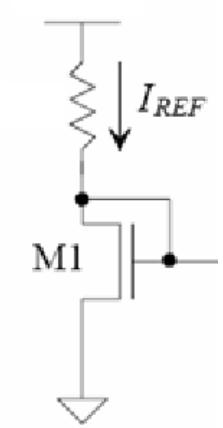


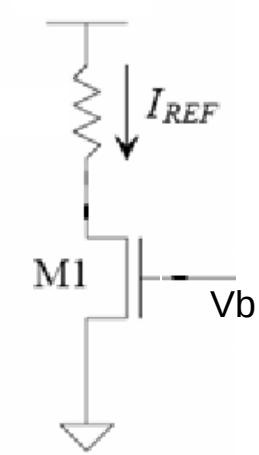
Example i/p sin wave to an amp with gain A SR must be > A.Vm.ωo else distortion will occur

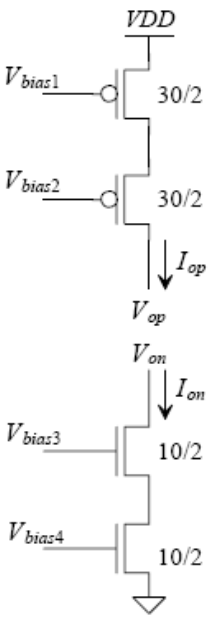
Differential amplifiers

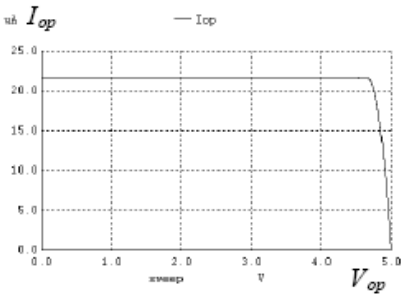
Passive load	**SE output**	**Diff output**
Active load	SE output	Diff output ??? CMFB (covered only in eee523)

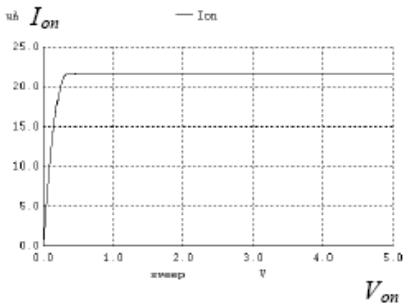
Setting DC bias: Diode connection vs. current source transistors

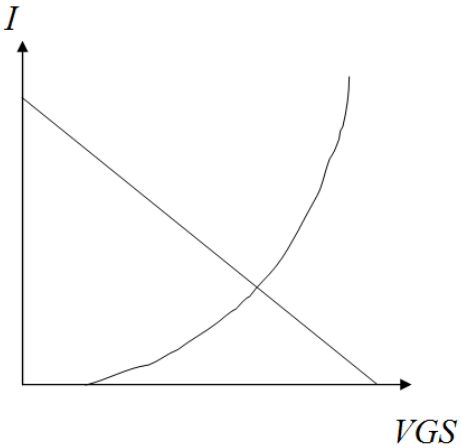


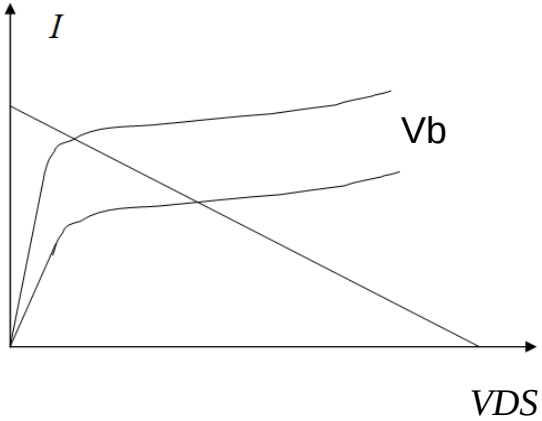






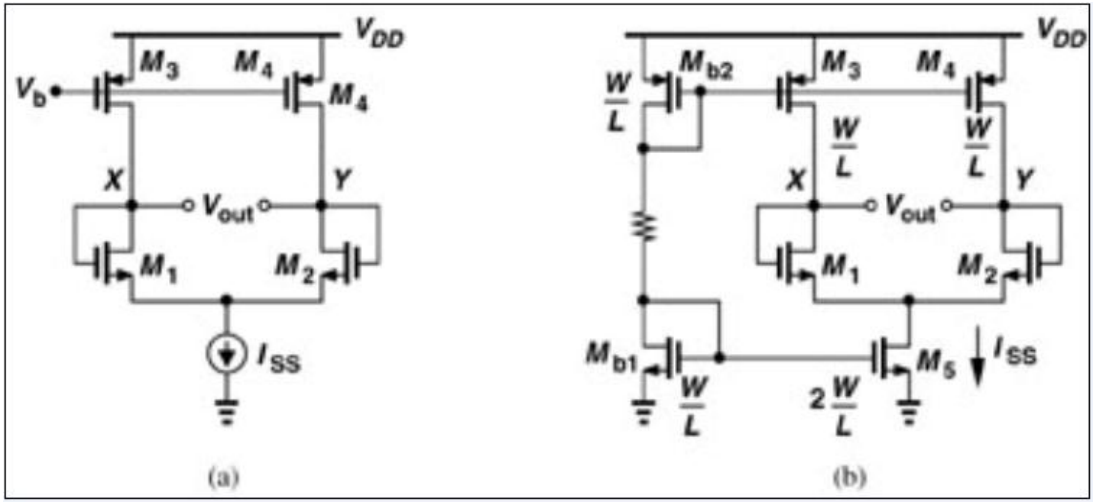


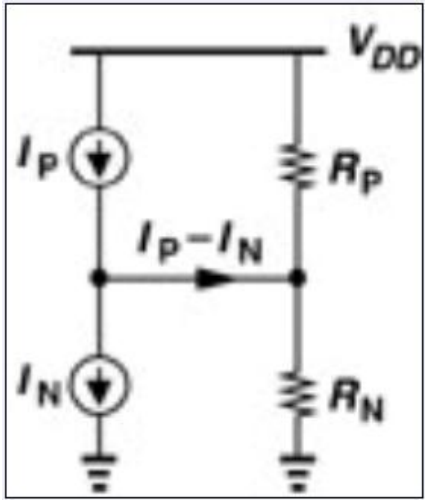




No guarantee that pmos current = nmos current, to satisfy KCL one of the branches will triode and Vout CM will be either very close to VDD or gnd. This is called a CM drift, has nothing to do with differential drift "offset" which is due to mismatch and asymmetrical diff branches, while CM drift is due to PMOS/NMOS mismtach in one branch during CM (i/p shorted) operation

Setting DC bias: Diode connection vs. current source transistors





If I3,4 in sat > Iss/2, M3,4 will triode so that their current falls to Iss/2 and satisfy KCL, Vx, Vy goes to near Vdd

*If I3,4 in sat < Iss/2 , both Vx, Vy will drop and M5 triodes to produce only 2*I3,4 to satisfy KCL*

Delta current must flow in rop and ron, creating a voltage error =

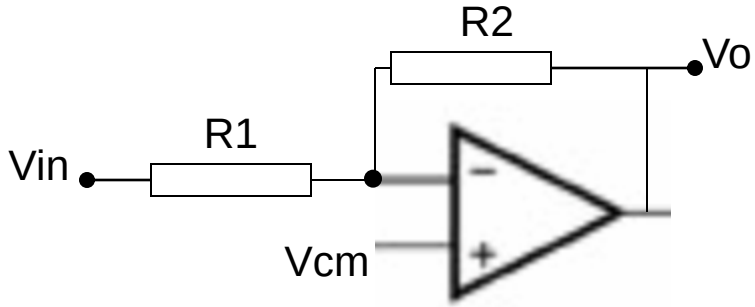
(Ip-In)(rop//ron) very high due to current sourece configuration causing transistors to triode and o/p CM to be either vdd or gnd

In high gain amplifiers (no diode connected loads), o/p CM level is sensitive to device properties and mismatch and it can't be stabilized by means of differential feedback. Diff feedback cannot define the cm level. CMFB is needed to set o/p CM and ensure proper amp operation.

Benefits of diff output vs. SE output

1) *Better CMRR (common mode noise canceled)*
2) *Better PSRR*
3) *Voltage swing 2X SE, may be higher than Vdd*
4) *Even harmonic distortion canceled*
5) *No mirror pole, higher closed loop speed*
6) *Matching is more critical in diff o/p opamps (matching to lower the offset and matching to equate o/p CM), while in SE only matching to lower offset is needed.*

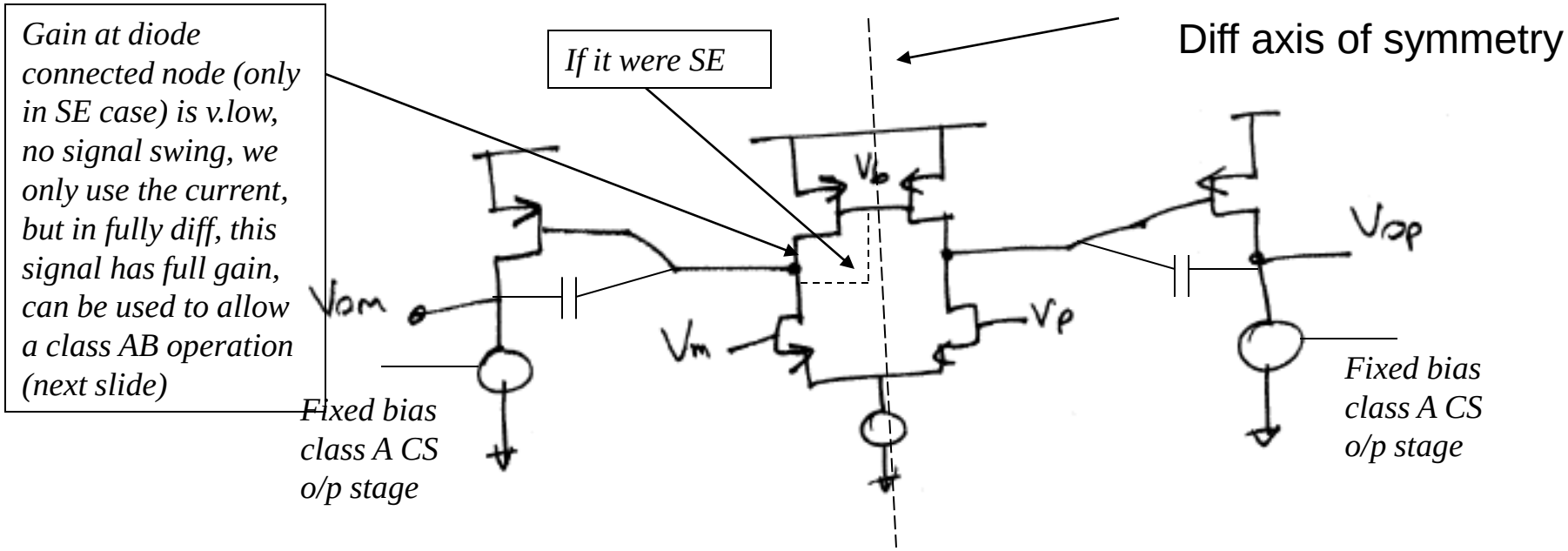
Diff output amps are fully symmetric if compared to SE output amps where one of the branches has diode connected transistors that set the o/p DC CM voltage. In previous lectures we discussed fully differential telescopic cascodes, fully differential folded cascodes. Down side of diff o/p amps is that due to using current source and sink transistors, DC level is hard to set. i/p CM is set by i/p sources, while o/p CM needs CMFB circuit, no global fb can help. While in SE amps, global fb is used to set the o/p CM level.



Ideal opamp A=inf

Non-ideal opamp A≠inf

How is Vcm copied to the output



At DC Vp=Vm, a SE diode connected copy Vsg to set the o/p CM, in a diff amp this is not controlled

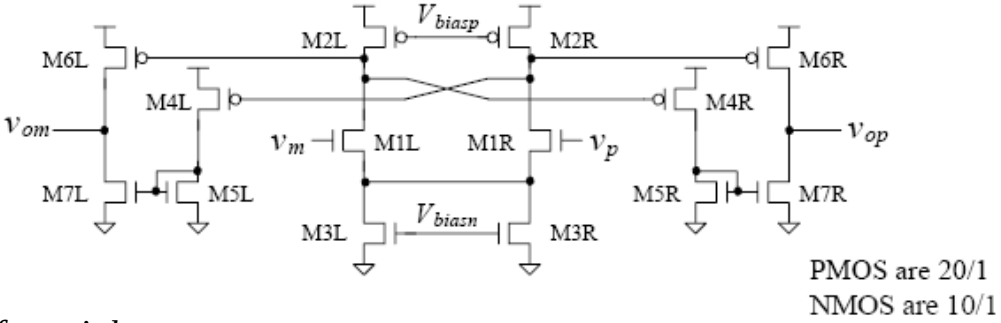
<u>CM definition</u>

At zero diff i/p, o/p CM should be equal (i.e Vop, cm=Vom, cm) if not this is considered a diff output mismatch that will cause an offset, this can be fixed by an i/p referred offset.

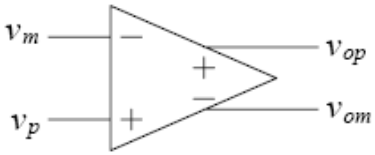
How is Vcm copied to the output (Vgs1=Vgs2) at DC

Introduction to Fully Differential OpAmps

$$v_o = v_{op} - v_{om} = A_{OL}.(v_p - v_m)$$

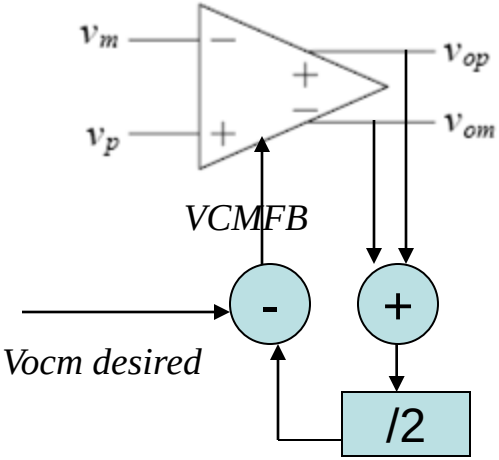


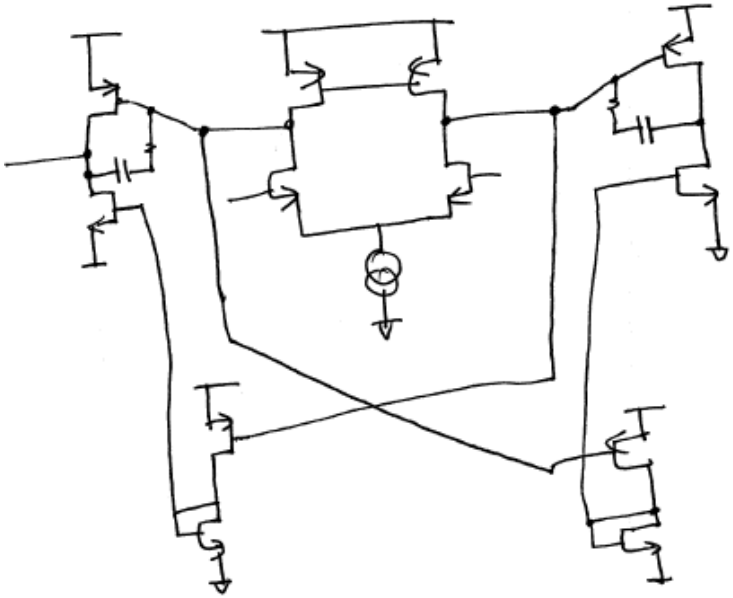
In fully differential circuits, both input and output are differential. 'Common-mode' signals is the average of input/output signals. Vcm=VDD/2 or any convenient DC level.



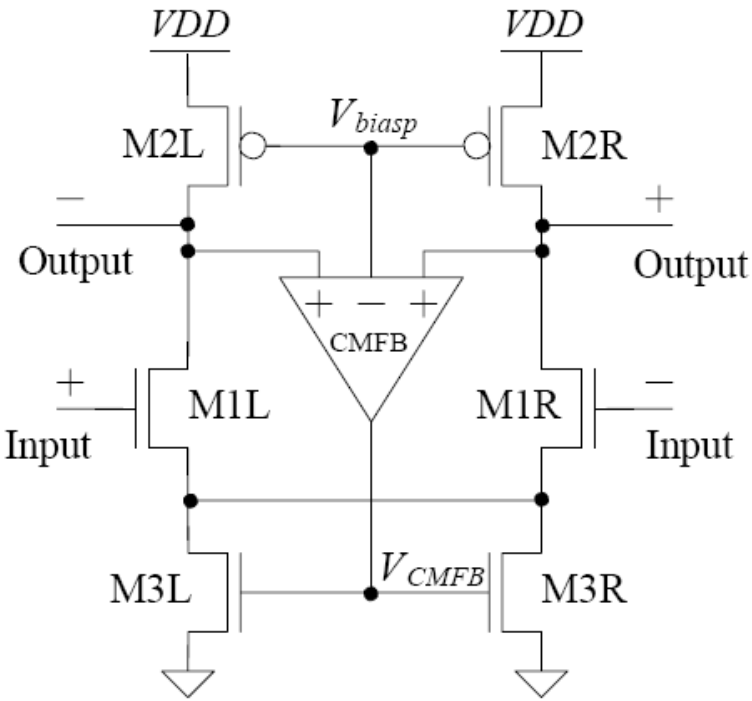
However, 'common-mode' signals need to be controlled and actively tracked. The circuits that achieve this are called 'common-mode feedback' (CMFB) circuits.

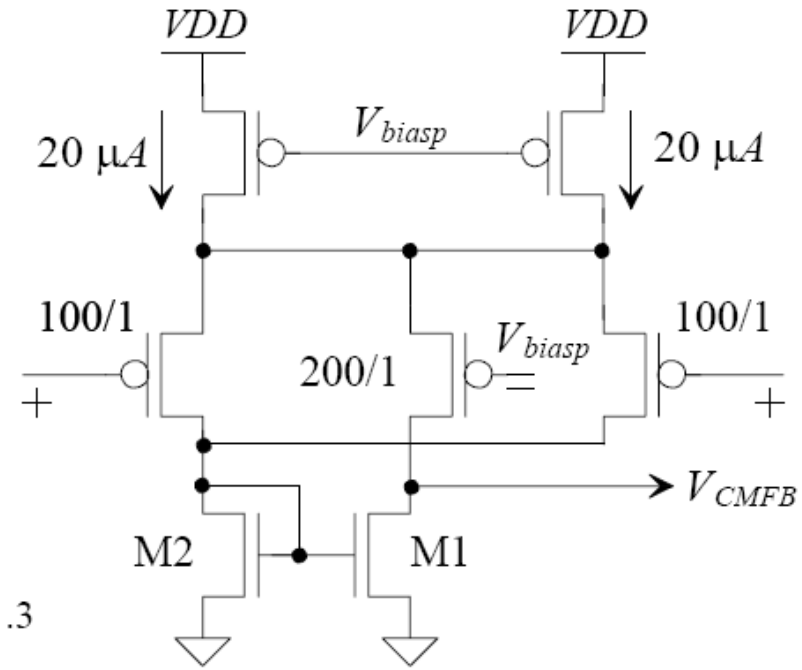
Class AB o/p stage





Where should VCMFB go? i.e. what should it be controlling ? Ans. A node that does not have an ac signal, so as not to interfere with the VCMFB error signal. Can control tail current sink or top current source

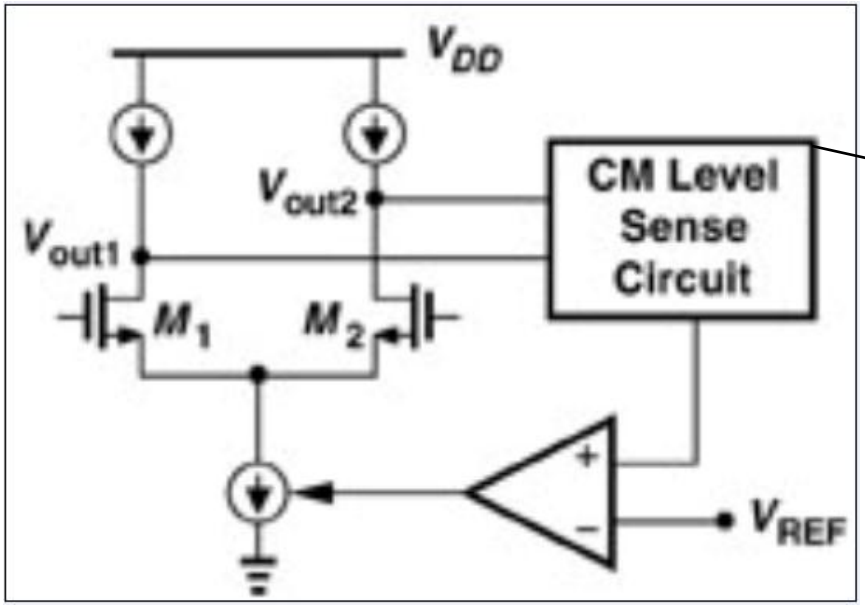


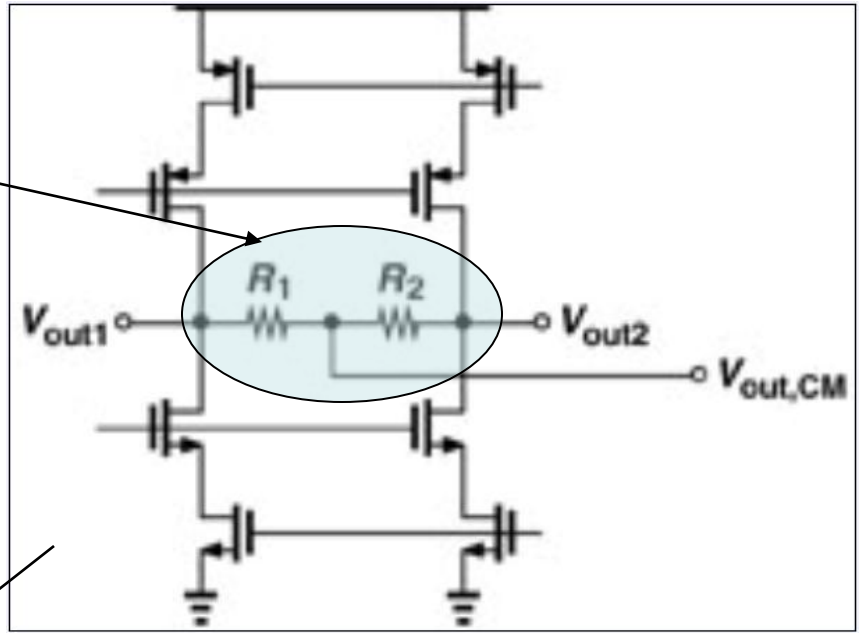


.3

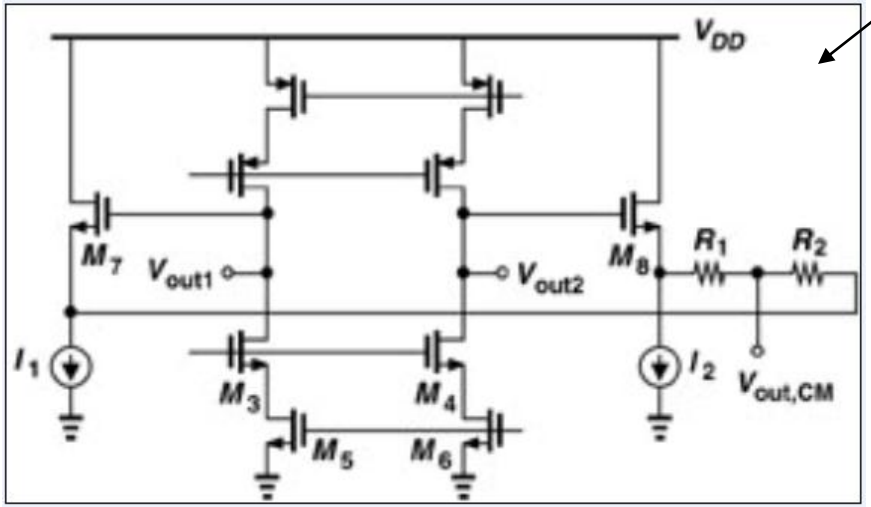
If Vocm increase, middle pmos tends to shut off, its current decrease, diode connected current decrease, mirror current M1 decrease, VCMFB increase, tail curent M3L,R increase, bringing Vocm down (CMFB)

A cmfb network must be added to sense the cm level of the two diff outputs and accordingly adjust one of the bias currents in the amp

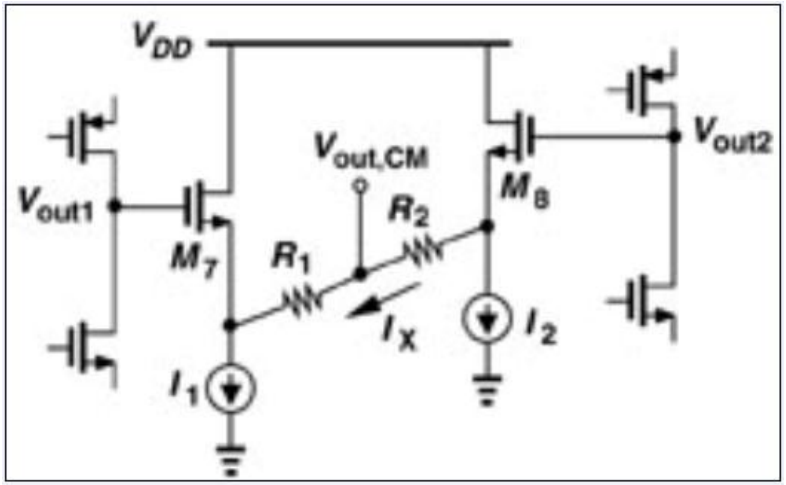




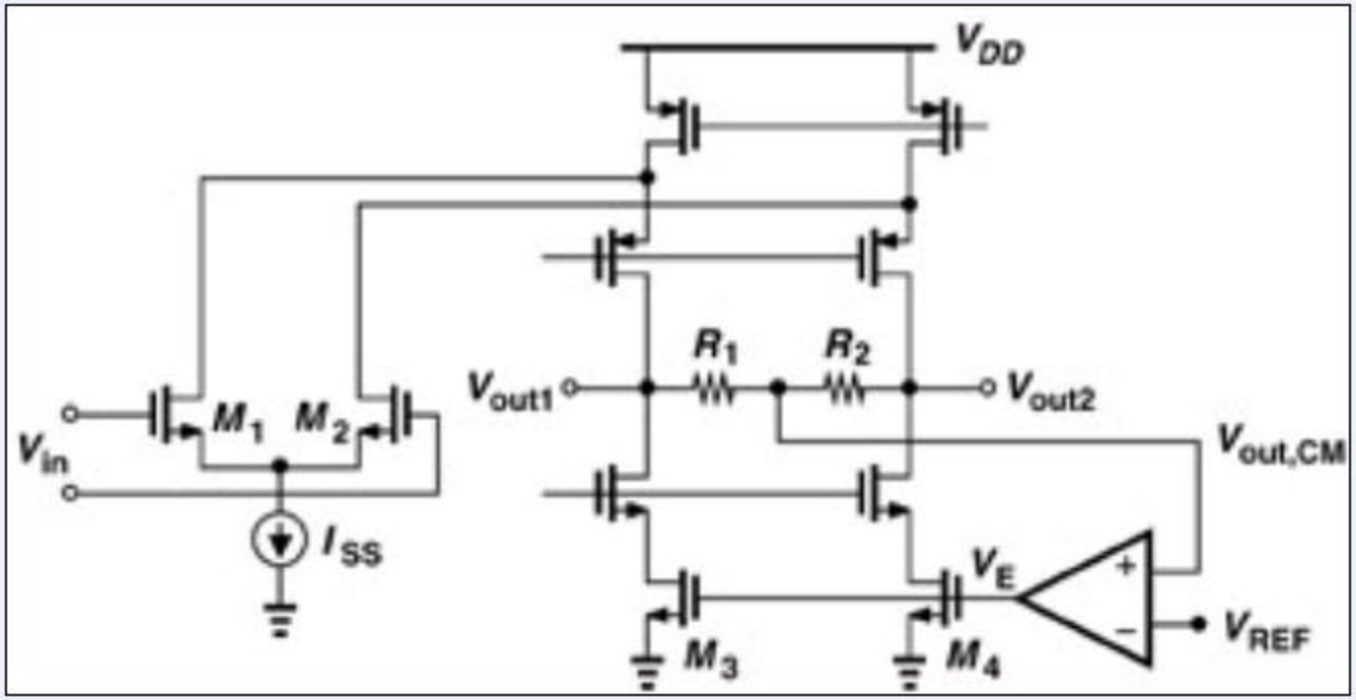
Vout, cm=vout1+vout2/2, if R1=R2



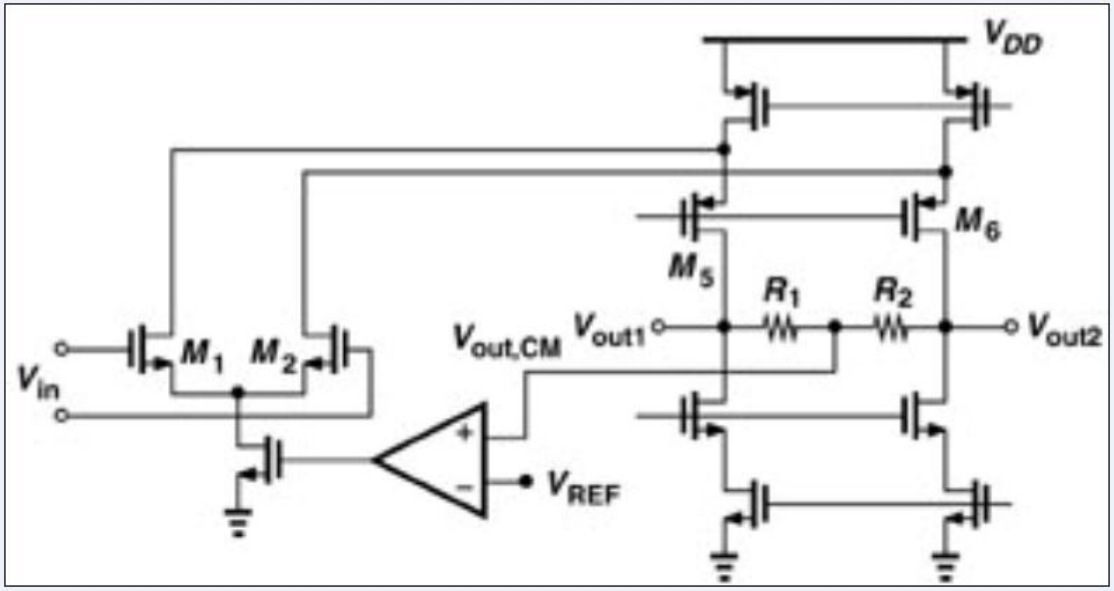
Sf buffers to avoid loading but Wasted swing due to CM sensing



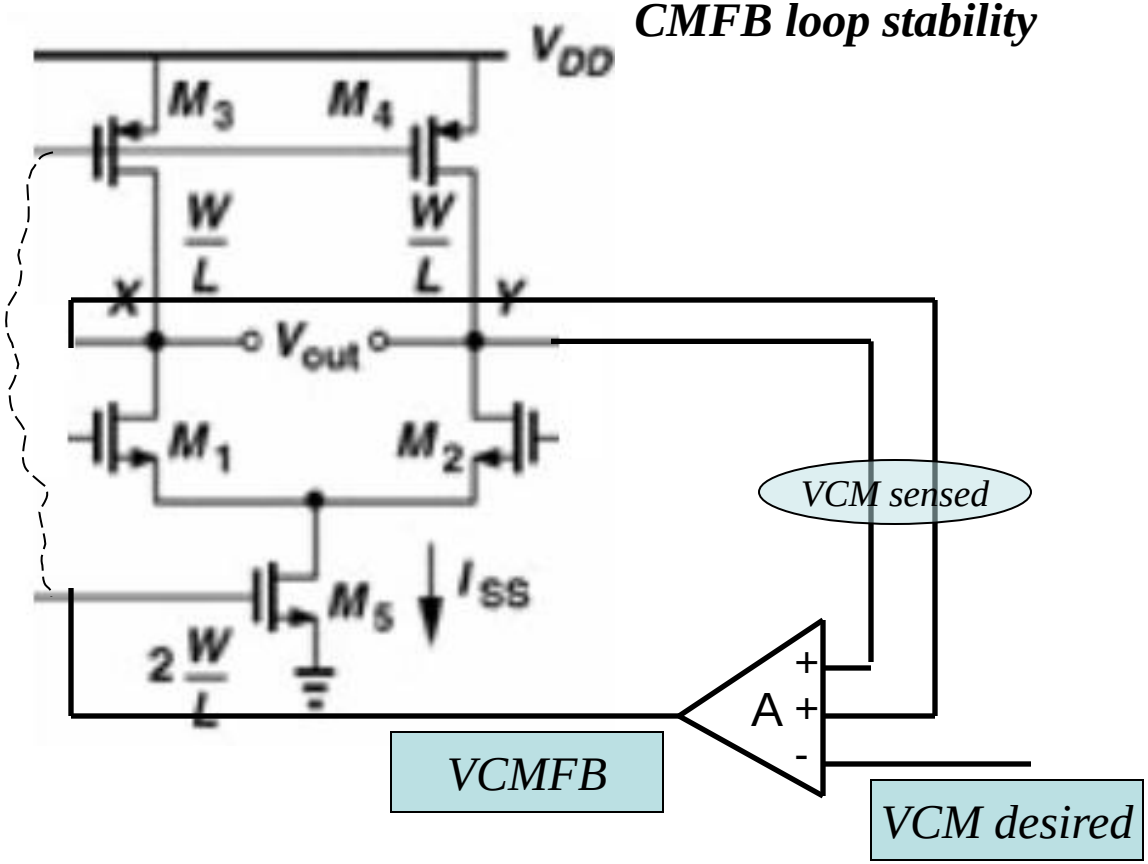
Avoid current starving of M7.8 (I1,2 R1,2 must be large)

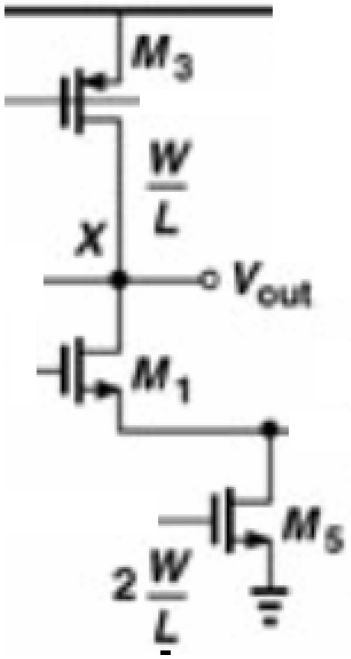


Loop gain if high will equate vout, cm to vref



CMFB loop stability





gm5.ro3.A=loop gain

VCMFB=loop gain * (VCM sensed –VCM desired)

2 amps in a fb loop, stability 2 amps + stabilizing the CMFB loop is v.imp

Simulation and Test Benches Quality Guide Lines

Purpose of this presentation

- Setting common guide lines to building accurate test benches that ensures good Silicon to Simulation correlation
  - *"What should I account for when I am building my test bench"*
- This is not project specific.
- This is based on many projects and Si debug to reproduce Si behavior in simulation.
- **All DE's are highly recommended to follow these guide lines and show in their design reviews how these guide lines are being followed, and document the validation needed to meet them.**

Introduction

- Si is not forgiving, i.e. do not expect things to work on their own, chances are; if you don't check and fix it , it will be broken.
- Our process is even more unforgiving vs. any other foundry due to our design rules and yield requirements.
  - These are present to reduce the wafer cost and increase our profit margins.
- 14nm will even be more and more unforgiving
- Even if your design works this does not mean you succeeded!!
  - Tens/hundreds of millions of units of your design must work and meet all specs to be successful.
- The parts outside our yield envelope is equal to the parts shipped by other companies.

  http://www.youtube.com/watch?v=981OA_BZ42Q
  - Our analog designers are not as your analog designers, we are ……

Pre lay schematic simulation

- Test benches must include
  - DUT (design under test, the block that you are designing)..obviously ☺
  - **Parasitics of critical nets inside your DUT**
    - Internal nets to the design
    - Parastics of I/O pins of the DUT (since they will have relatively long routes)
    - **Parasitics of High Speed nets**
    - **Parasitics of high current nets**
    - **Parasitics of bias nets**
    - The above nets will need lots of metal tracks to pass RV and hence it will be associated with high parasitic r's and c's, it is better to estimate it and attach it to test benches in the pre-lay phase to avoid unpleasant surprises at the POLO phase.
    - Many tools are available to estimate these net parasitic.
      - Parest **+** the follwoing
      - Short CDC methodology and calculator
      - RC cal
      - Upf data for sheet rho
      - A rough rule of thumb is 0.2fF/um for the cap of a route (back of the envelope)
      - **Ftrc**
        - » **Make sure to include ALL the info in the ftrc GUI, mainly spr, spl, MCF, else ftrc will be too optumistic.**

Pre lay schematic simulation

- **Loading** to your DUT
  - Block loading your DUT must be in your test bench
  - include the estimated parasitics of the load cap and res of the loading block input stage and second stage (to account for Miller cap).
    - i.e. → If a divider (DUT) is driving a cml buffer (load), put the load and its input estimated cap and res as the load to your DUT, put also the output Cload of the cml buffer to account for Miller cap effect.
- **Interface** spec
  - **Do not assume** your input signal, talk to the DE of the block driving you, agree on a spec, *document it in a ticket*, put safety margins to it, use the WC spec as your input across PVT.
  - **Or better and when available** include the driving block of your DUT to your TB to get the input signal values as it varies across PVT, hence you will not use a constant WC single value across all PVTs
  - Include the estimated res and cap to the output signal of the block driving your DUT

Pre lay schematic simulation

- **Overlap blocks in your test benches when you hand over specs across interfaces or when you break big designs to stages (due to large simulation time)**
  - See next slide for explanation
- **Bias**
  - Use bias spec with its min-typ-max, add to it its 5 sigma number
  - Always include the bias random variation in your bias values
  - **Or when available** include the bias block in your TB to get its value across PVT
  - Include the bias in your Monte Carlo runs
  - Never use ideal current sources for production checks, these can give fake EOS violations (an ideal current source into an off transistor, can give 1k Volt !!!)
    - Why would I put a current source into an OFF transistor to start with → I am smarter than that
    - Well to check your DUT when it is ON and when it is OFF, and the bias block and your DUT can be enabled and disabled independently, i.e. **leave no stone unturned**

Pre lay schematic simulation

- **Supply noise**
  - Do not assume clean supply **it is not**, apply supply noise (amplitude and frequency ranges to your VCC)
  - Use min and max VCC, add to it supply noise, (30mV p2p 1MHz to 200MHz)
  - Note that VCCmin is 0.94V and VCCmax is 1.1V (do not let noise take your VCC below the min nor above the max)
- CM noise
  - Do not assume clean CM, add CM noise and check your DUT sensitivity to it.
- Duty Cycle variation.
  - Do not assume a perfect clock coming to you
  - Apply Duty Cycle error, rise/fall time to your input clocks, **account for DC random variation up to 5 sigma**

PVT

- Account for MFC and CPR variation with the corners list
- Meet performance at
  - tttt
  - rsss
  - rfff
  - rsfs
  - rfsf
  - rssf
  - rffs
  - All the above blended with
    - Min - Max Temp -40C and 110C
    - Min - Max Vcc 0.94V and 1.1V

Sensitivity

- Know your design weak points
  - Sensitivity to bias current and/or voltage on both high and low ends of the bias values, load, routing
  - Sensitivity to i/p swing, i/p CM, CM/supply noise, duty cycle
  - Sensitivity to freq (your BW)
  - Sensitivity to Vt, Idsat, rout
  - If you want to check your Vt sensitivity between rsss and ssss, use rsss and add a voltage source to increase your Vt across your GS junction.
  - FTRC, TL
  - You should be able to say this in your sleep
    - My WC corner is …….

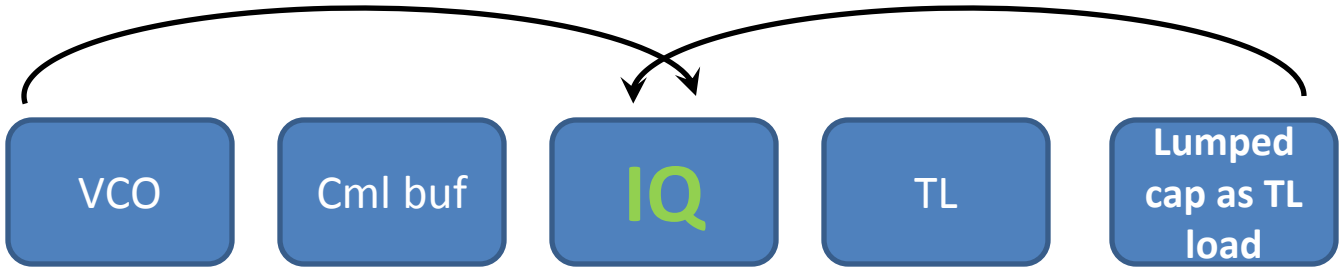
Monte Carlo

- rsss minV cold, corresponding bias, corresponding stimulus, make sure your spec is met with 5 sigma. For example
  - Mean gain at rsss minV cold is 500V/V
  - Sigma is 15
  - Spec should be greater than 500 -5*15
  - Gain spec is > 425V/V
- Repeat the above at your **WC corner**, if rsss minv cold is not your WC corner.

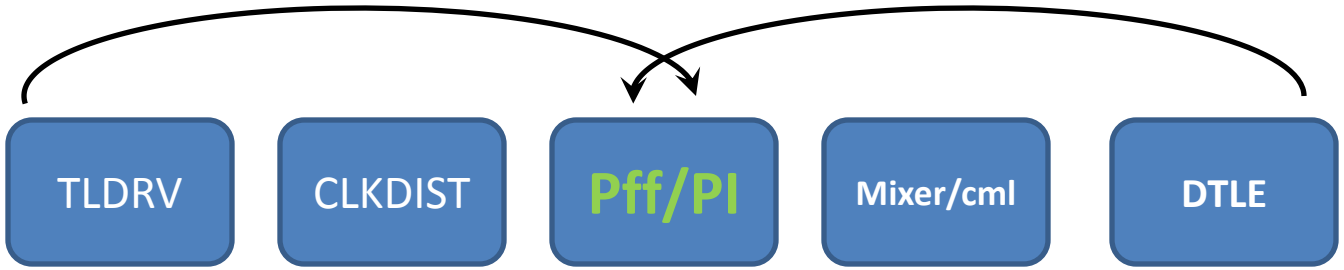
POST lay

- POLO simulation
  - Previous slides still applies → big time
  - POLO for the DUT/BIAS/LOAD/driving block
  - Overlap
  - Make sure the parasitics are loaded into the simulator, not dropped, stitched (if applicable) and intact

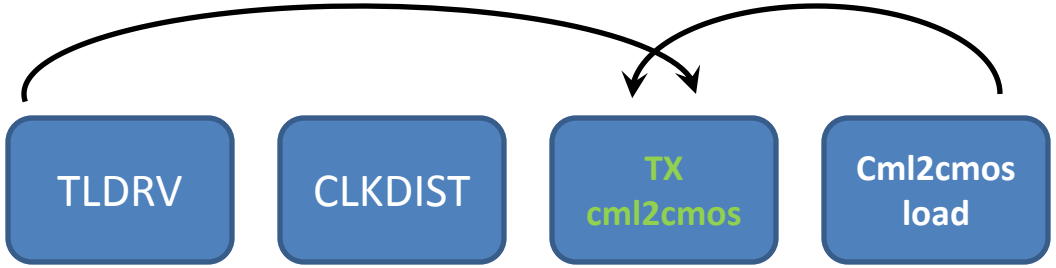
Overlap Need to get the i/p and o/p specs of the IQ for example



Overlap Need to get the i/p and o/p specs of the PI for example



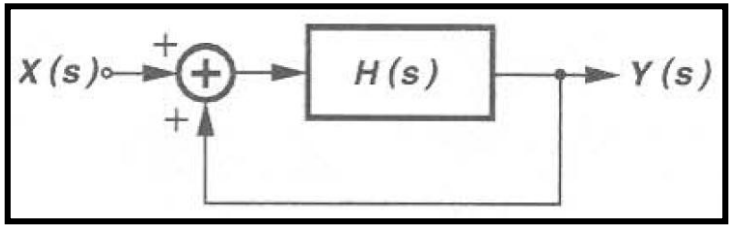
Overlap Need to get the i/p and o/p specs of the TX input stage for example



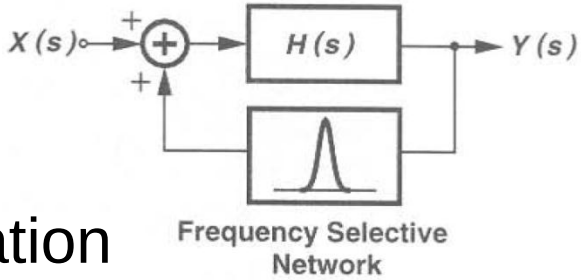
VCO Outline

- LC-VCO Quick Overview
- Basic Concepts and tradeoffs
- Architecture Options
- Practical Considerations in 65nm and beyond
  - Deep Submicron Challenges
  - Inductor Design
  - Varactor Design
  - Oscillator design
  - Sample Data and what to look for.

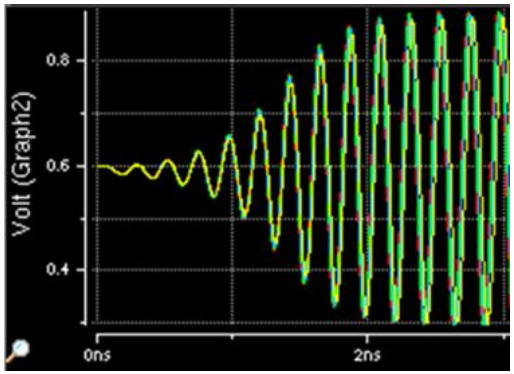
LC-VCO Quick Overview

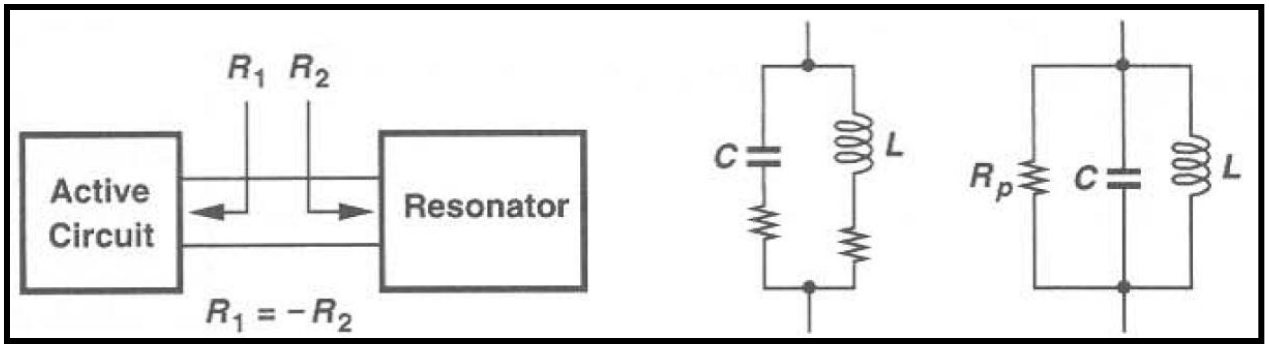


$$\frac{Y(s)}{X(s)} = \frac{H(s)}{1 - H(s)}.$$



- Oscillators are +ve feedback systems with
  - Loop gain=1 (at sustainable oscillations)
  - Total phase shift around the feedback loop = 0
- LC is used in the loop to stabilize the oscillation
  - Sharp band pass filter to select o/p freq and ensure low PN + sine o/p
- Frequency of oscillation is set by the tank resonance freq
- Loop gain should be > 1 to start oscillation (amplify noise)
- As amp increases, loop gain decreases, maintaining steady oscillation

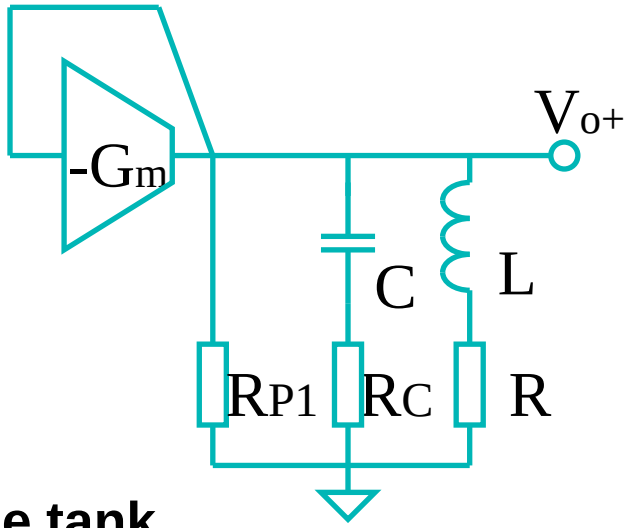




Required Condition of Oscillation

Loop gain = $G_m R_p > 1$ $\quad Q_c >> Q_L$

$$\begin{cases} Q = \dfrac{\omega L}{R} \\ Rp = R \cdot \left(1 + Q^2\right) \end{cases} \rightarrow Gm \cdot \underbrace{R \cdot (1 + Q^2)}_{Rp} > 1$$



- **Rp is the effective parallel resistance of the tank**
- **Q is the loaded quality factor of the tank α Rp**
- **Loop gain α Gm**
- **Loop gain α Q**
- **Worst case is low Q ind + low Q tank (all Cbank switched in the design "1111" lowest freq of osc)**
- **Worst case is lowest Gm at "ssss" and lowest tail bias current "1111" for the sel bits.**
- **Lowest amplitude of oscillation, <u>must check for these conditions</u>**

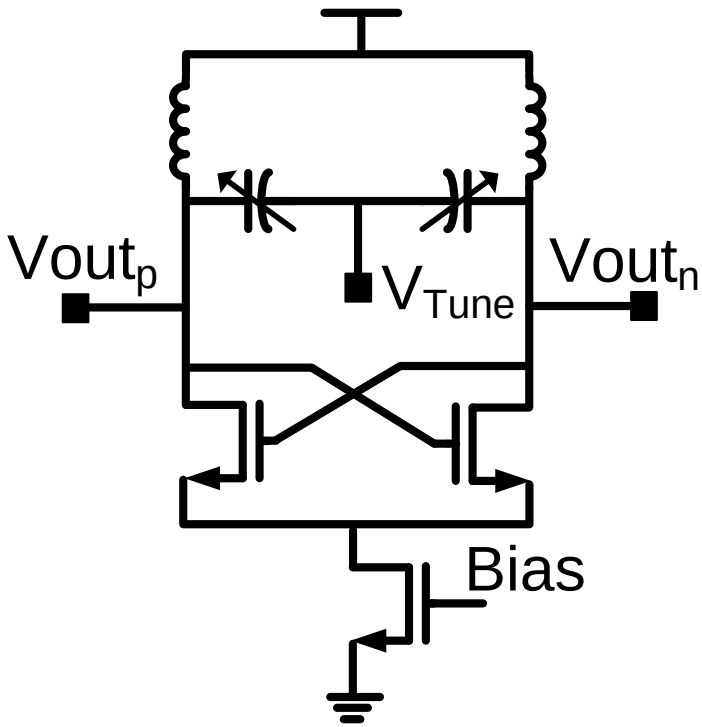
Basic Concepts

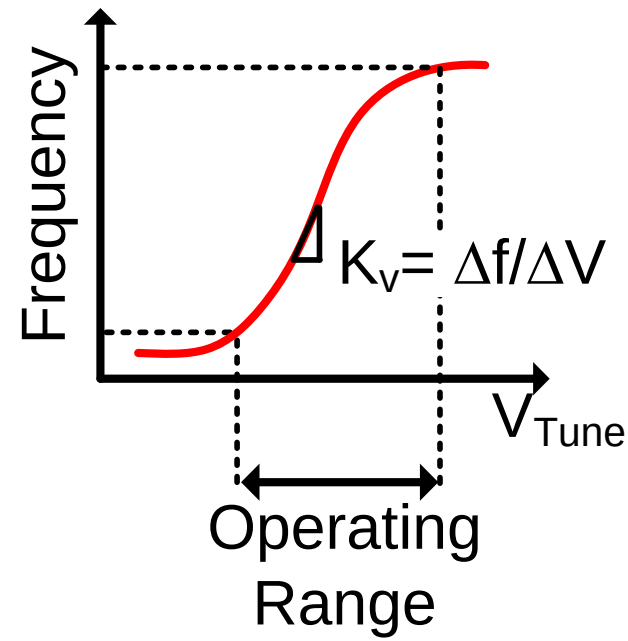
- High Q of the tank is needed (ind Q) for
  - high loop gain,
  - low PN (sharper resonance has low PN skirt),
  - high swing for a certain bias current + Low power dissipation.
    - Rp ~ Q, Vswing ~ I.Q, higher Q yields lower I for same Vswing, hence lower power
  - Active device noise injection decreases
    - Lower I implies lower active device noise
- Low freq noise (1/f) on vtune impacts PN, tradeoff between Kv (tuning) and PN is needed, more on this later

Deep Submicron trends

- As technology feature size is reduced,
  - the doping concentration of the diffused regions is increased to avoid total pinch off effect.
  - Higher doping concentration leads to higher depletion capacitance
  - therefore lower supply is needed to avoid junction break down
- Supply voltage scales down,
  - Voltage swing decreases,
  - higher relative PN,
  - rev bias of the varact is less,
  - limit tuning range, specially at high freq, as Cpex is high, varac can change only a small frac of the tank C
- Low KVco is better for PN worse for tuning range.
- Coarse tuning is needed using digitally controlled Cbank to cover frequencies needed.

Voltage Controlled Oscillator (VCO)



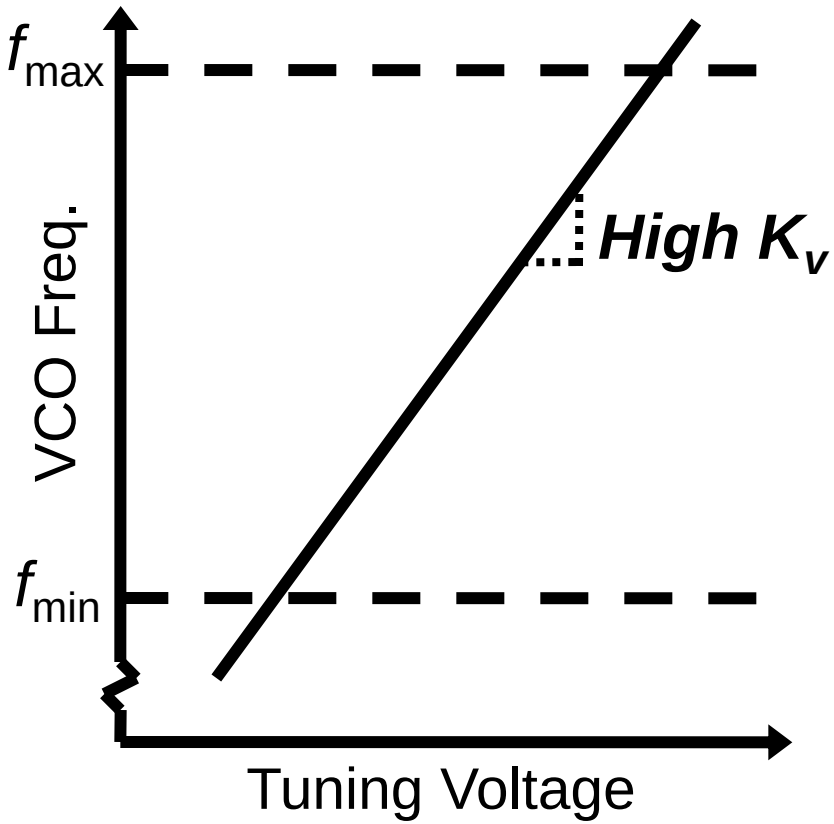


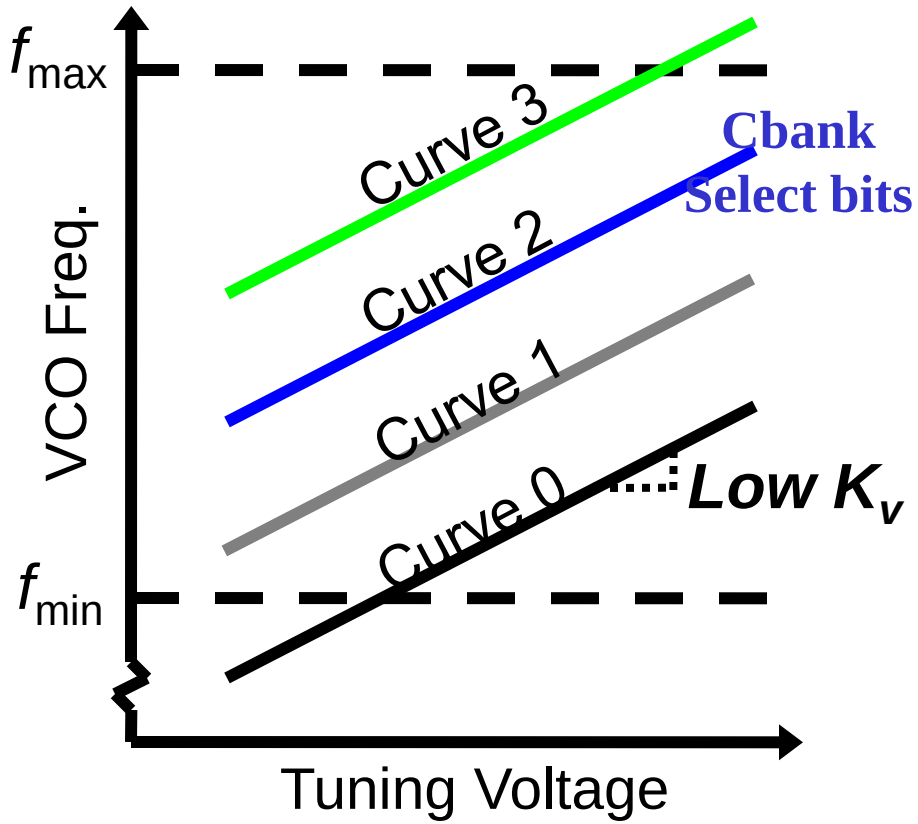
$$f(V_{Tune}) = \frac{1}{2\pi\sqrt{L \cdot (C_p + C_V(V_{Tune}))}}$$

- **Implementation Issues :**
  - **Phase noise**
  - **Tuning range**
  - **Gain (K_v) variation over operating range**

$$K_{VCO} = \frac{df(V_{Tune})}{dV_{Tune}}$$

VCO Design Tradeoffs





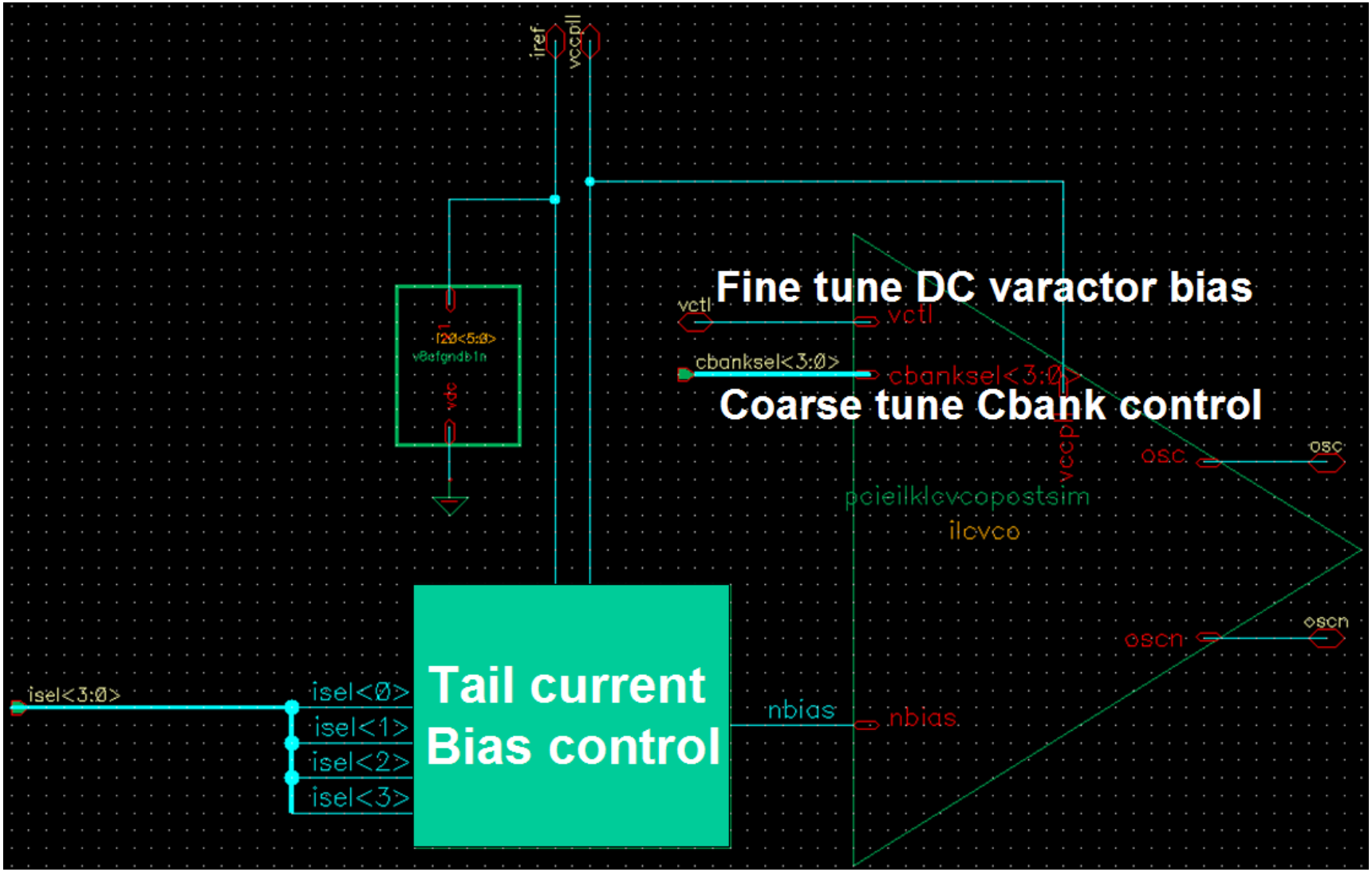
Typical VCO

- **No tuning required**
- **high phase noise**

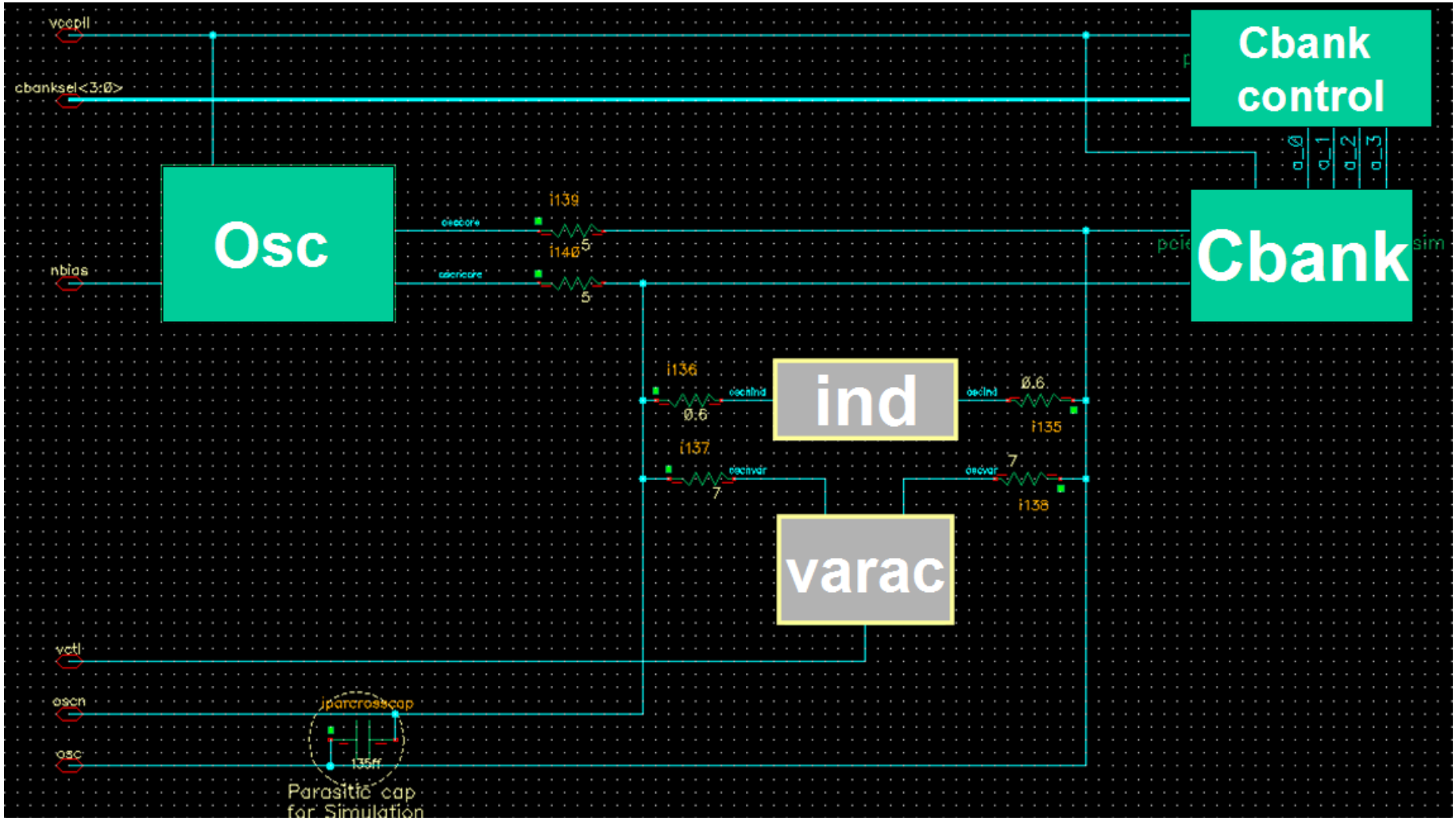
Coarse-Tune VCO

- **low phase noise**
- **Curve selection req'd**

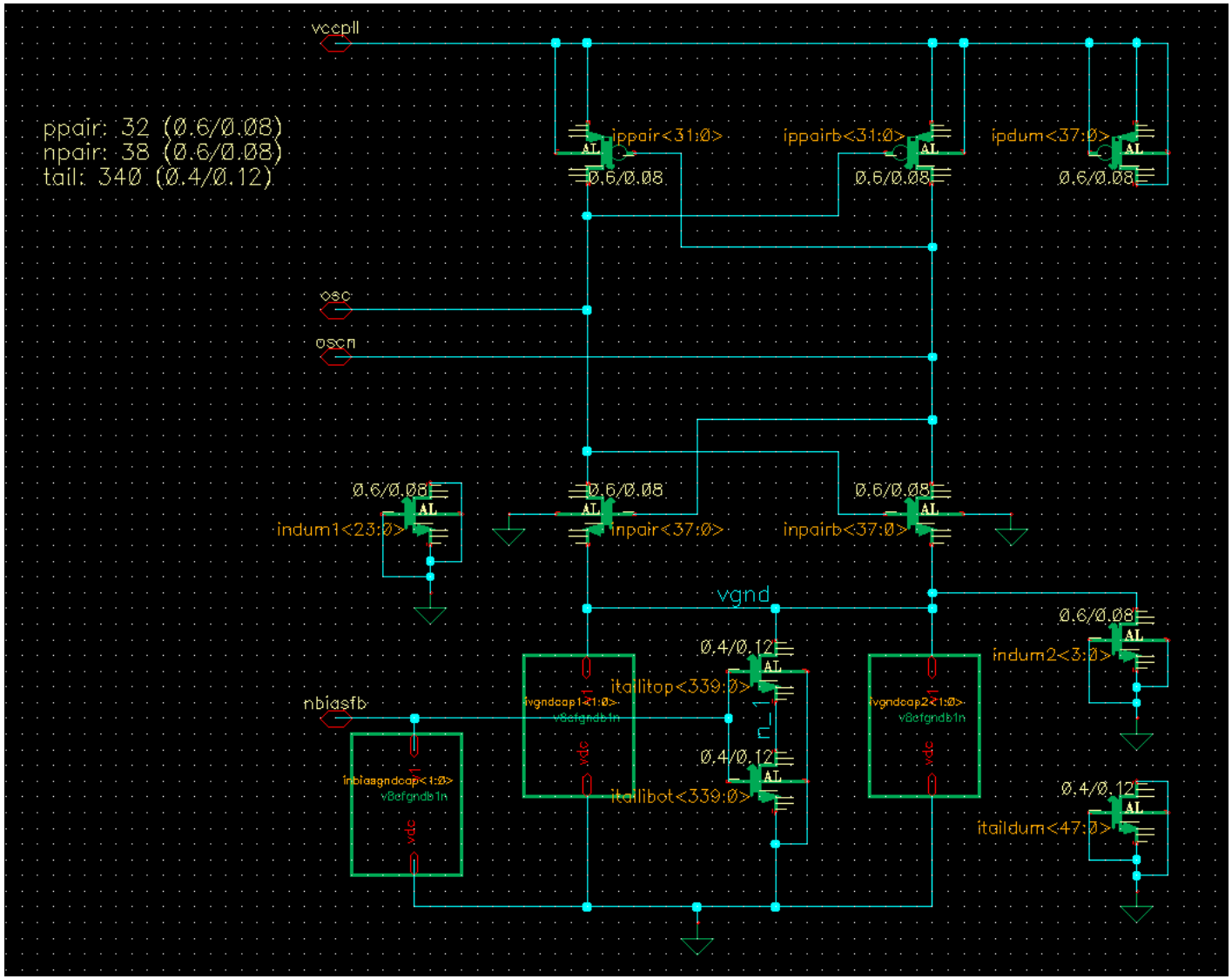
Voltage Controlled Oscillator (VCO)



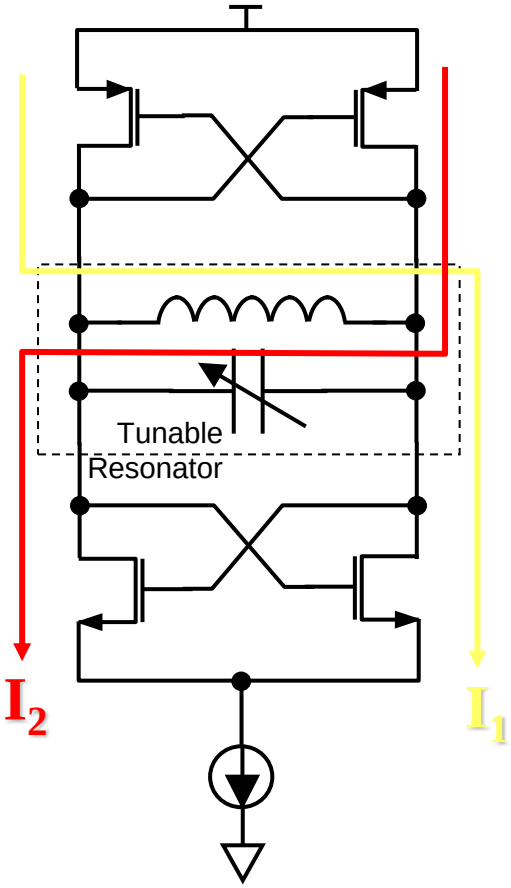
Voltage Controlled Oscillator (VCO)



Voltage Controlled Oscillator (VCO)



LC VCO Architectures



NPmos VCO

Pros:

- Double output amplitude for similar bias current.
- lower power for similar output swing.
- Larger negative resistance for similar current.
- more symmetric oscillation waveform which reduces up conversion of flicker noise therefore resulting in lower phase noise.

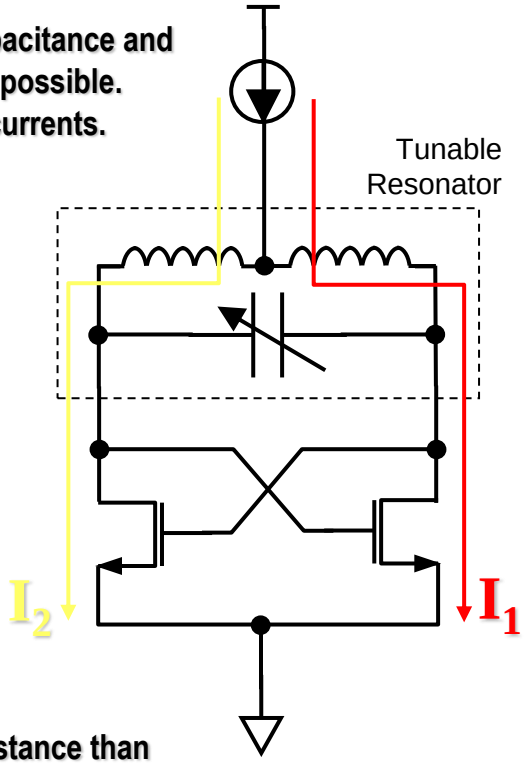
Cons:

- Increased parasitic capacitance.

Nmos VCO

Pros:

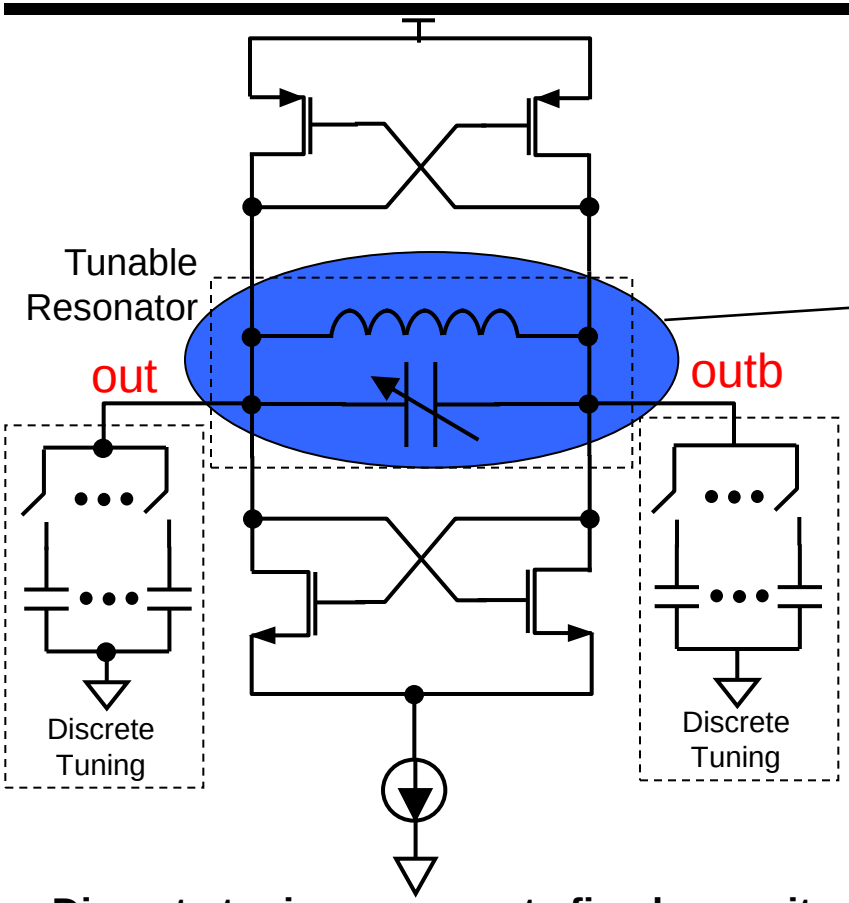
- Lower parasitic capacitance and wider tuning range possible.
- Can handle larger currents.

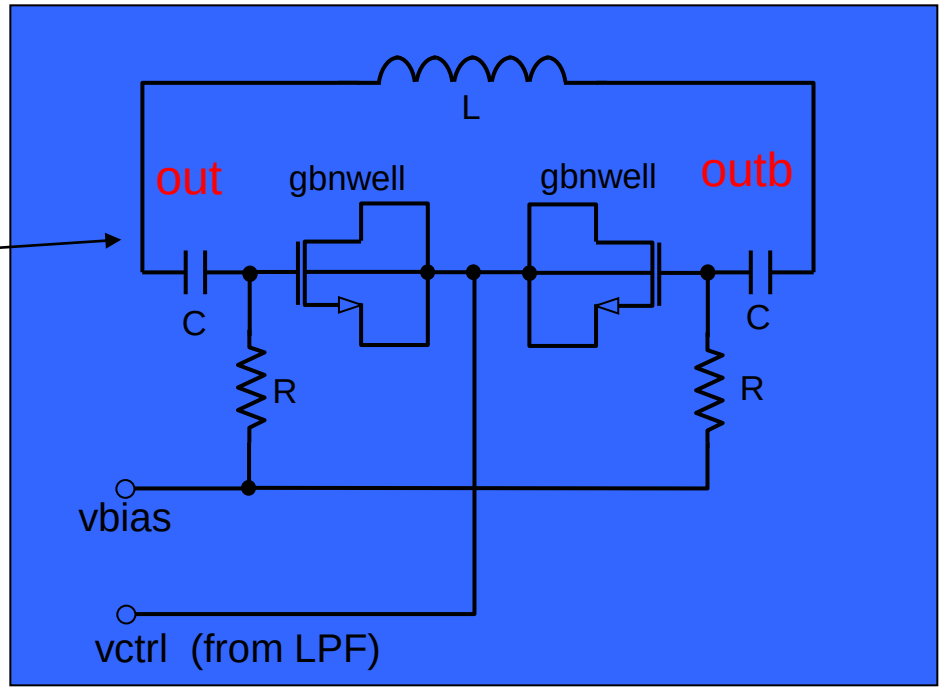


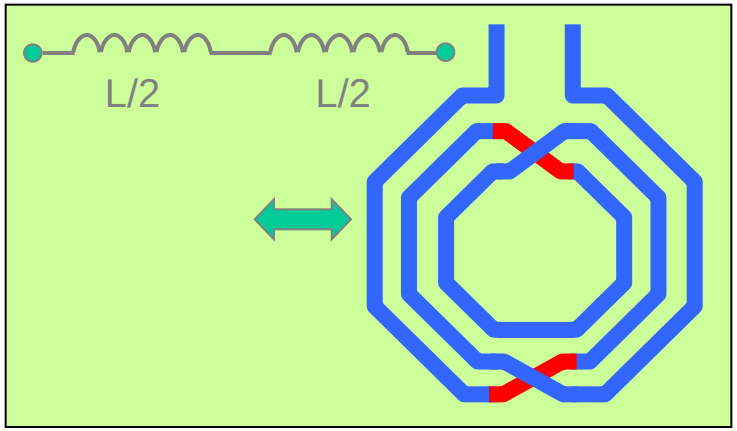
Cons:

- lower negative resistance than that of the NPmos structure.
- Higher phase noise due to asymmetry in output signals.
- higher power consumption.

LC VCO Architecture







- **Discrete tuning represents fixed capacitance contributed by the active transistors and digital tuning capacitors.**
- **<u>Increasing active transistor width degrades the fine tuning range.</u>**
- **As we switch in the digital capacitors, the tuning range decreases.**
- **<u>Tradeoff between digital coarse tuning range and fine tuning range.</u>**

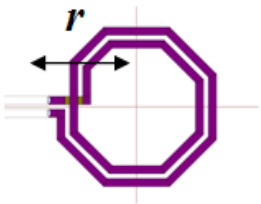
Deep Submicron Challenges

- Deep submicron implies
  - Higher freq
  - Higher integration levels
  - Low supply voltage
  - SOC
- Discrete L and W values for cost efficient litho.
- Stringent design rules for cost efficient litho.
  - Must be very **layout aware**.
- Very high 1/f corner specially for the shortest channel devices
- Higher substrate noise coupling in highly integrated circuits
- Highly packed metal stack has a higher parasitic impact

Inductor Basics

- **Important design factors for inductors are: L, Q, SRF, and Cost (which means Si area)**

$$L \approx \mu n^2 r$$



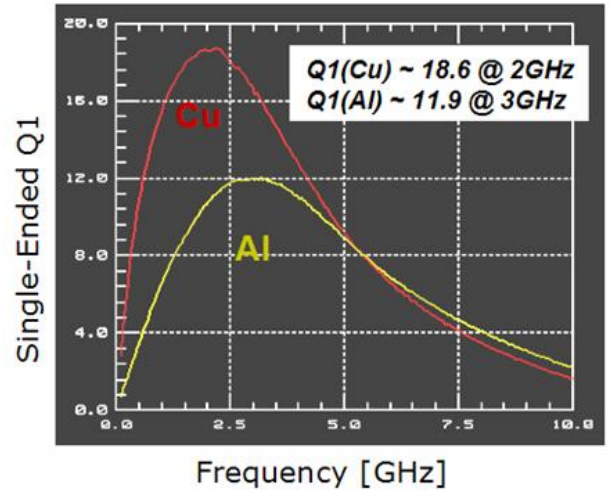
- **Quality factor is the ratio of peak energy stored to energy dissipated in the inductor**

$$Q \equiv \frac{2\pi E_{peak}}{E_{diss}}$$

- **The self resonance frequency is given by**

$$SRF = \frac{1}{2\pi\sqrt{LC}}$$

- **This shows that inductors do not scale with lithography. More accurately they scale inversely, so they are more expensive in each new generation**

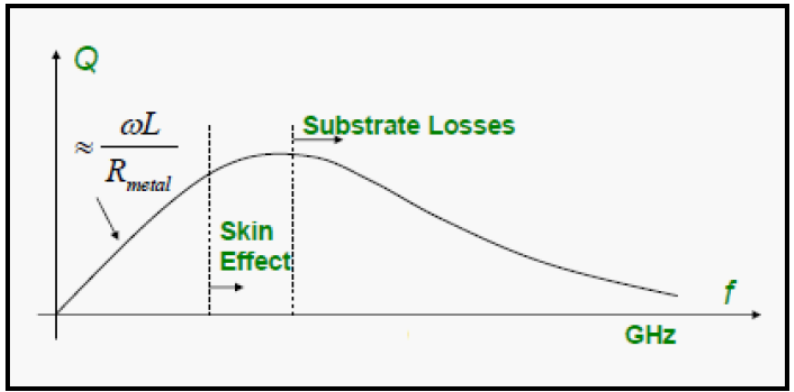


Scaling hurts Q

$$R_s \approx \frac{l}{w \cdot \sigma \cdot \delta\left(1 - e^{-t/\delta}\right)}$$

$$\delta = \sqrt{\frac{2}{\omega \mu_0 \sigma}}$$

Cu much better than Al

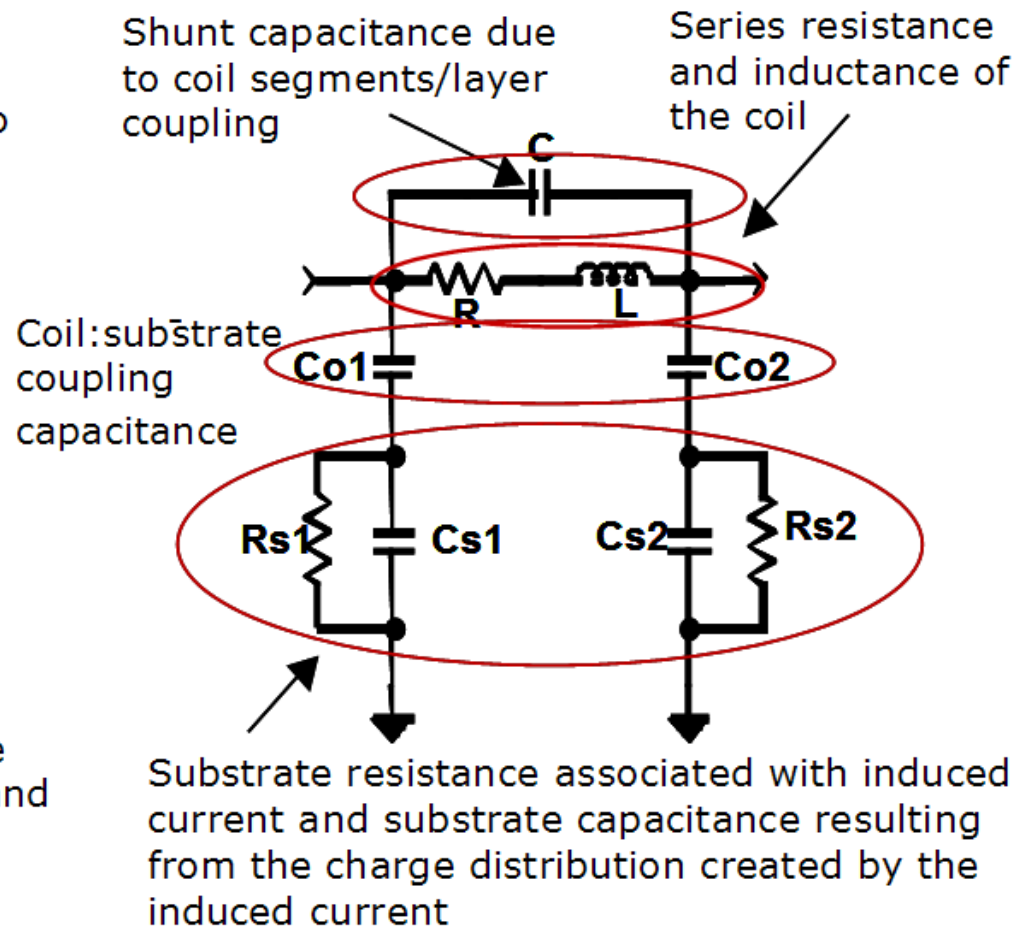


Inductor Losses: Substrate

- **The substrate loss has two components.**
  - Loss due to current flowing into the substrate through Co1
  - Image currents induced by current flowing in the spiral above
  - The power lost is given by i^2R

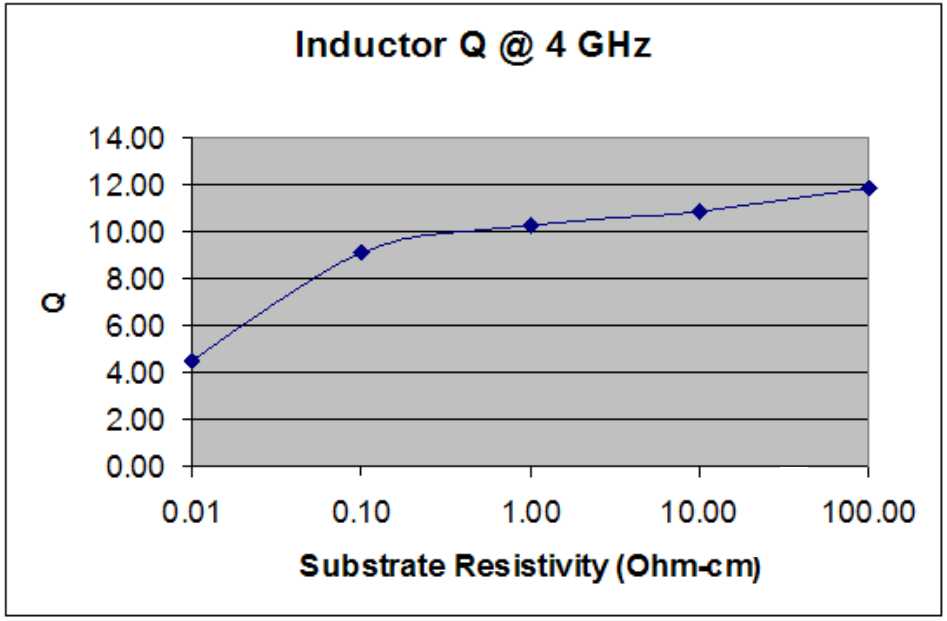
$$R_s \approx \frac{2}{w \cdot l \cdot G_{sub}}$$

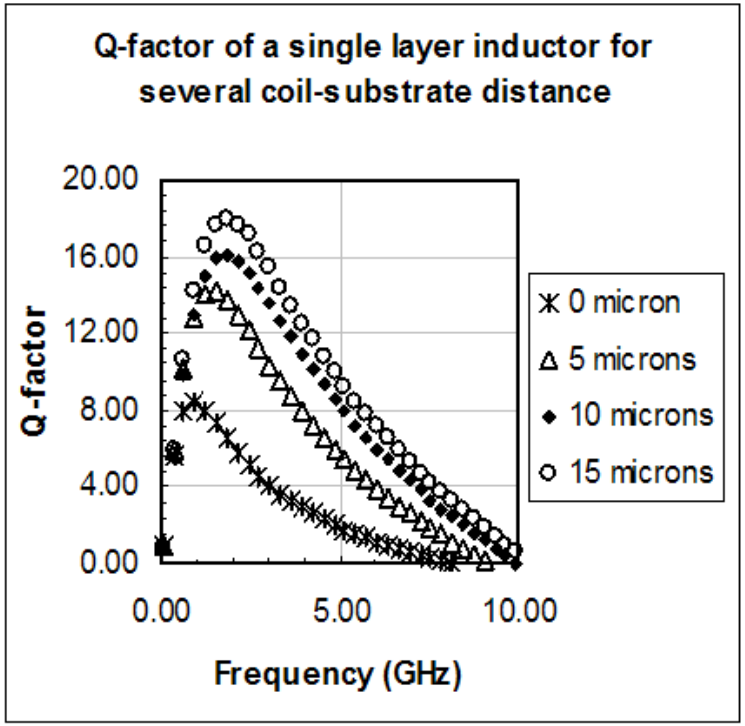
G_{sub} is a fitting parameter that captures the substrate conductance and the distance between the coil and substrate



Higher res substrates or greater distance from the substrate reduces the loss and improves Q

Substrate Resistivity & Inductor Q





General trend in inductor Q for a 3-turn, ~2nH inductor.
The metal stack is constant, only substrate conductivity is changed

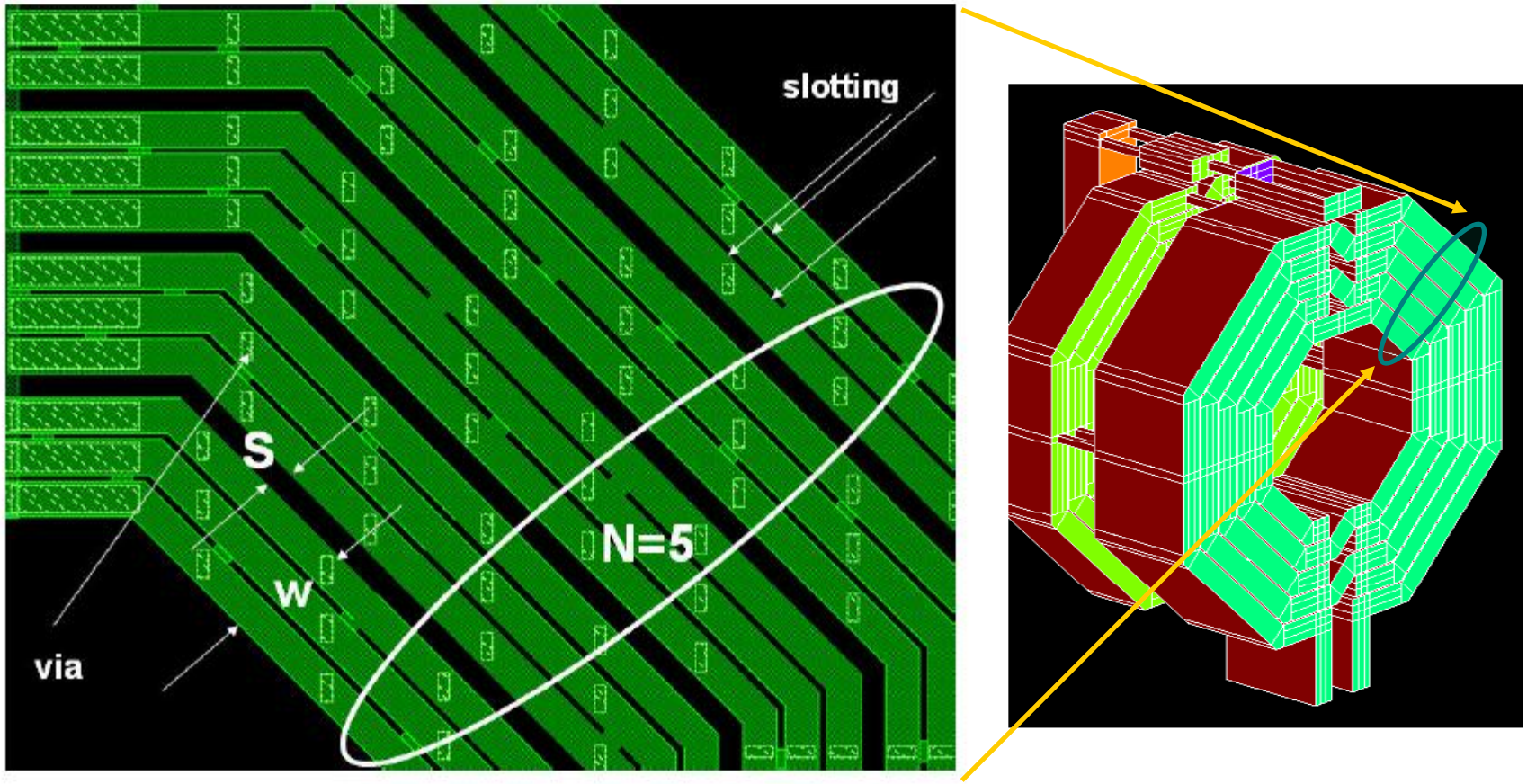
Increasing the substrate resistivity or the distance between the coil and the substrate significantly improves Q

Practical Considerations during ind design

Inductance L
Desired minimum quality factor Q at the operating frequency fo
The operating frequency fo
Minimum self resonance frequency fr
Maximum DC current
Maximum rms current
Bump pitch for flip chip
Maximum metal density rules
Metal design rules, maximum width, minimum width, minimum spacing

set of parameters that should be considered during designing the inductor

Practical Considerations



Inductor layout showing design parameters and some DFM rules

Inductor behavior vs. Freq

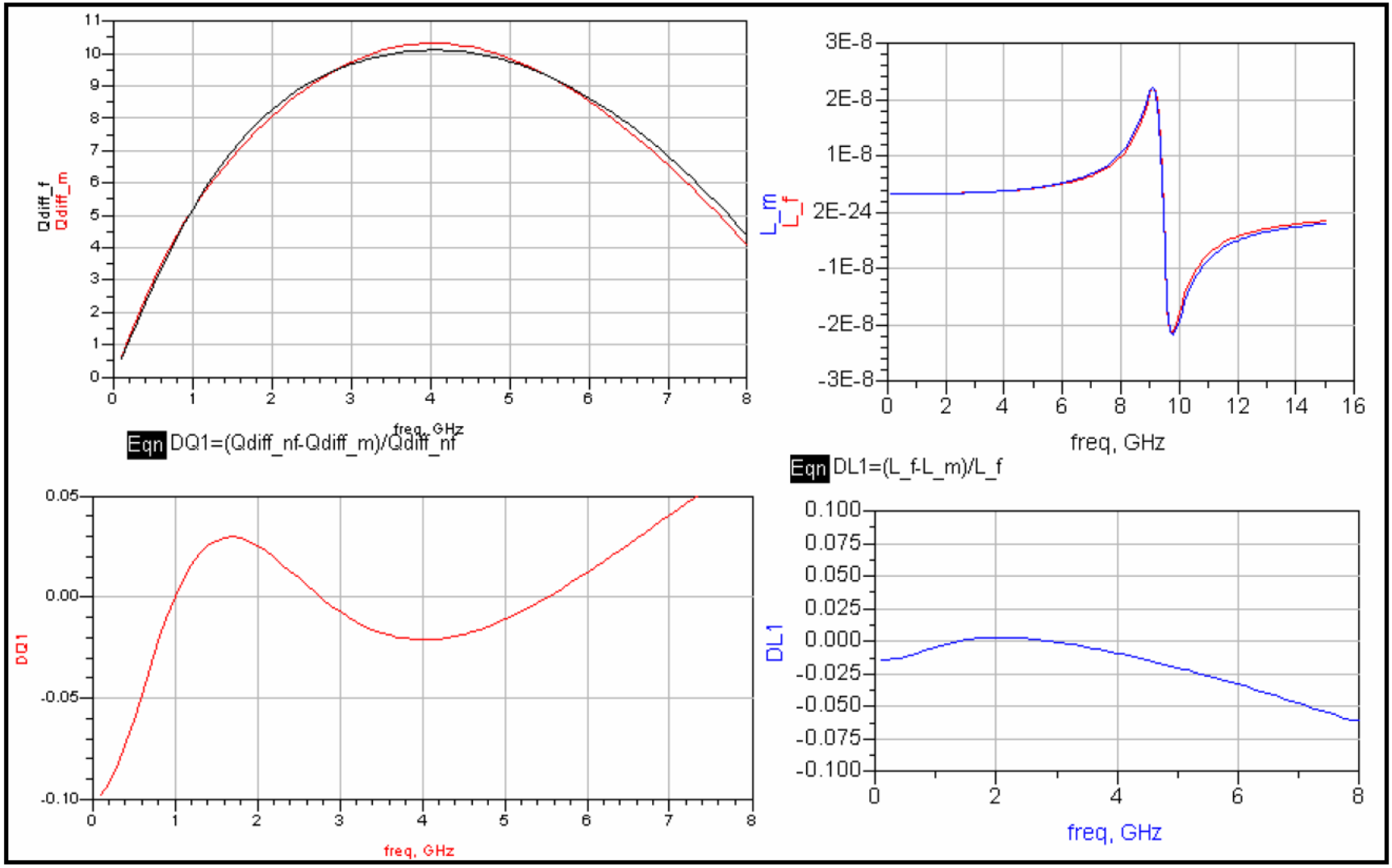
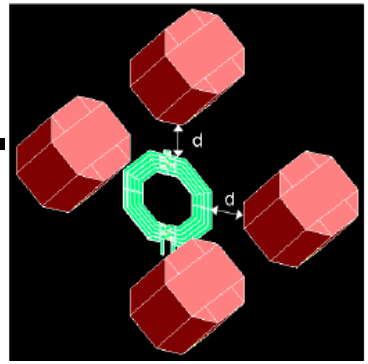
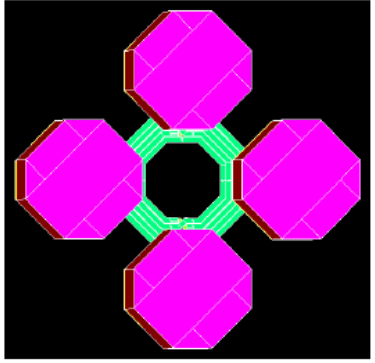
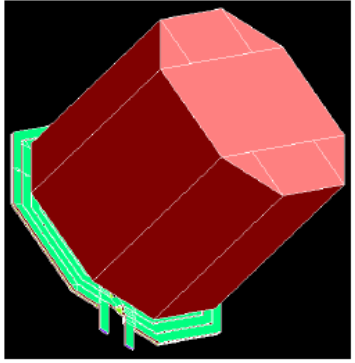


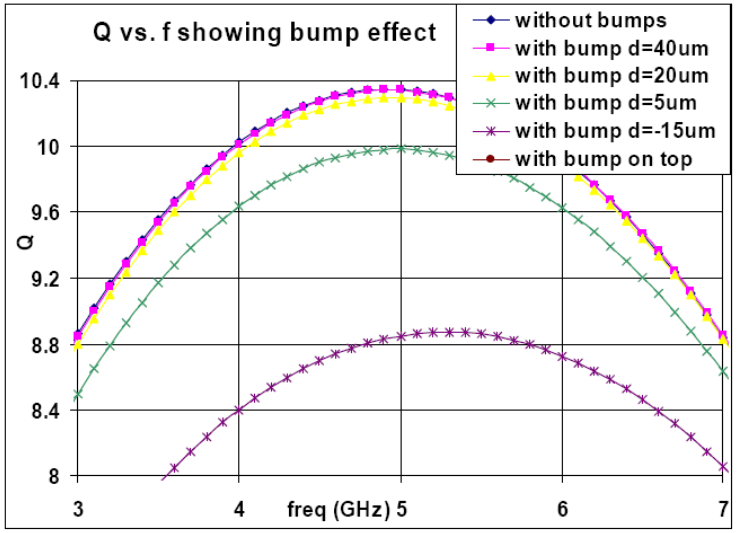
Figure 5.11 Q and L macro model fit vs. sp file and percentage error

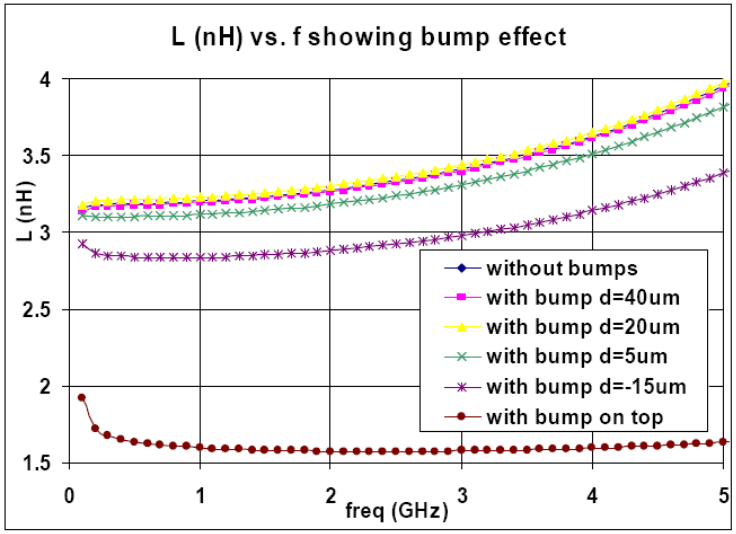
Practical Considerations Bump Impact

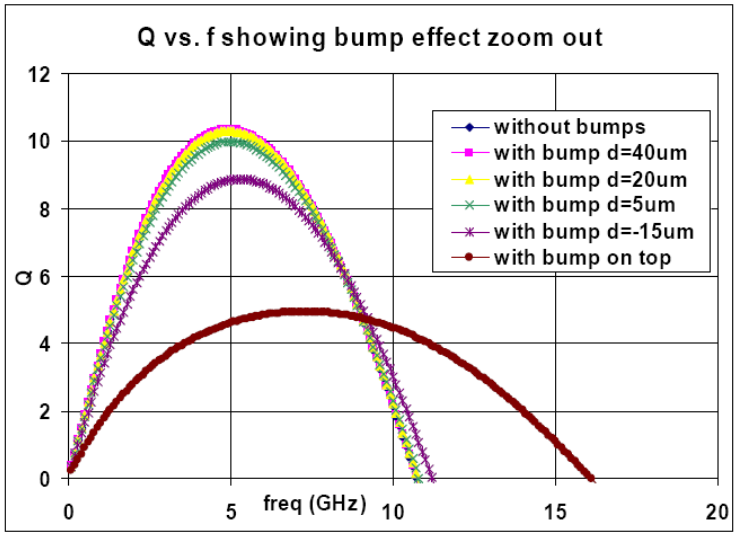


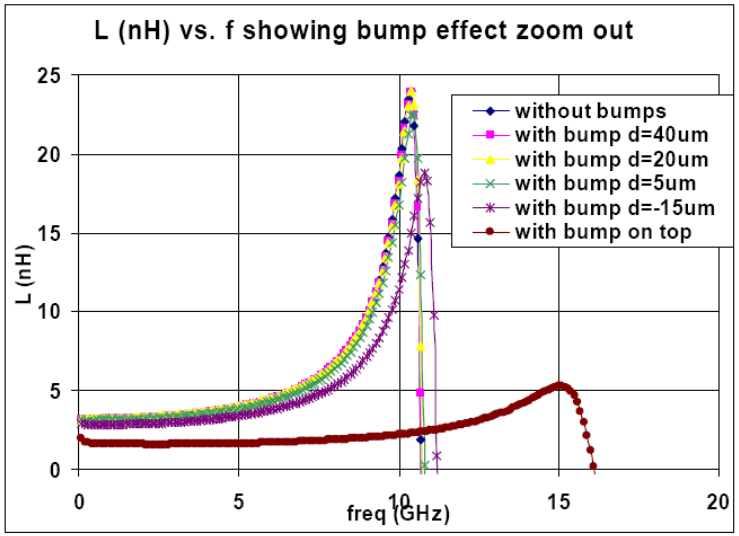




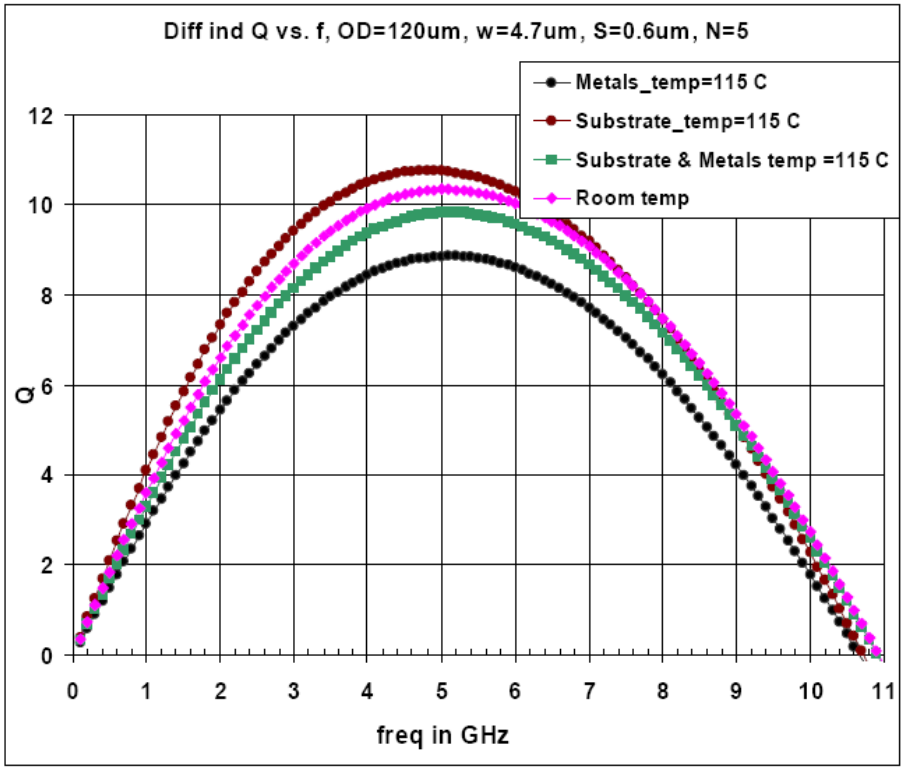


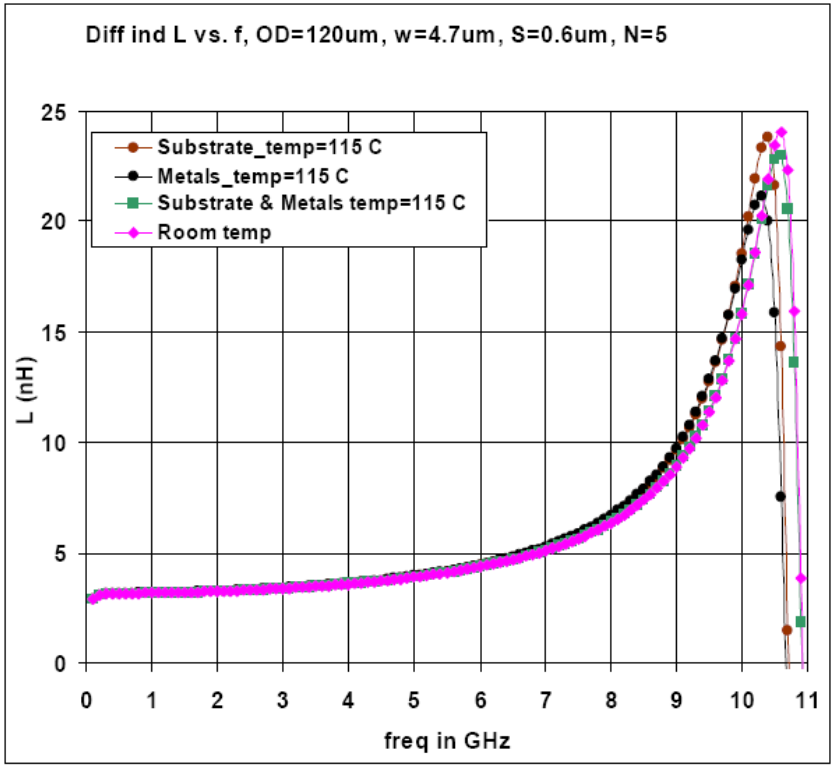




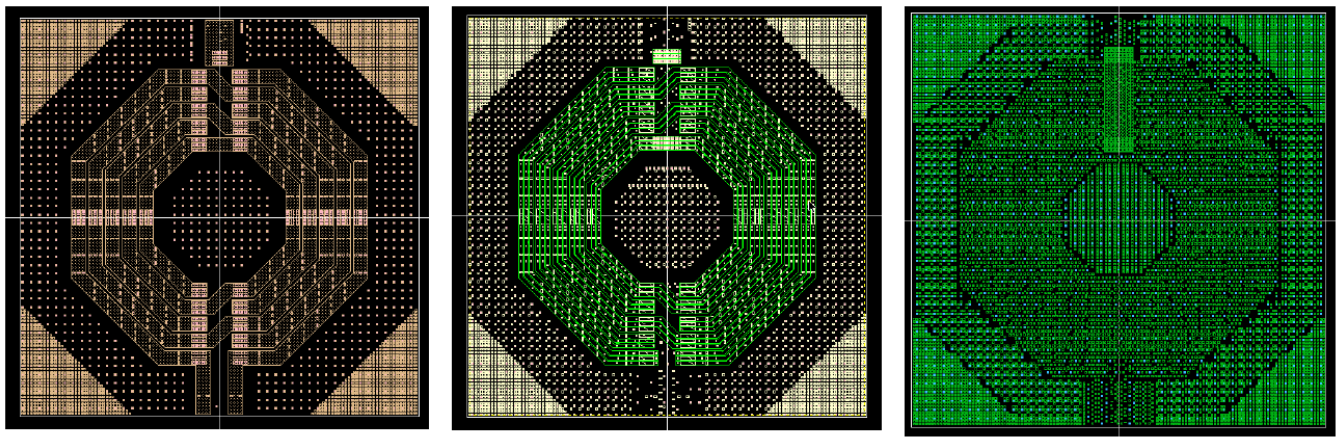


Practical Considerations - Temp impact





Practical Considerations – Dummification



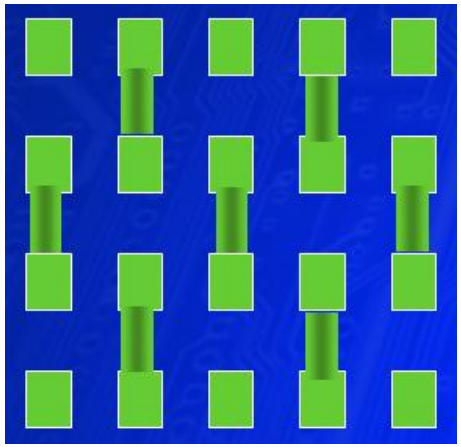
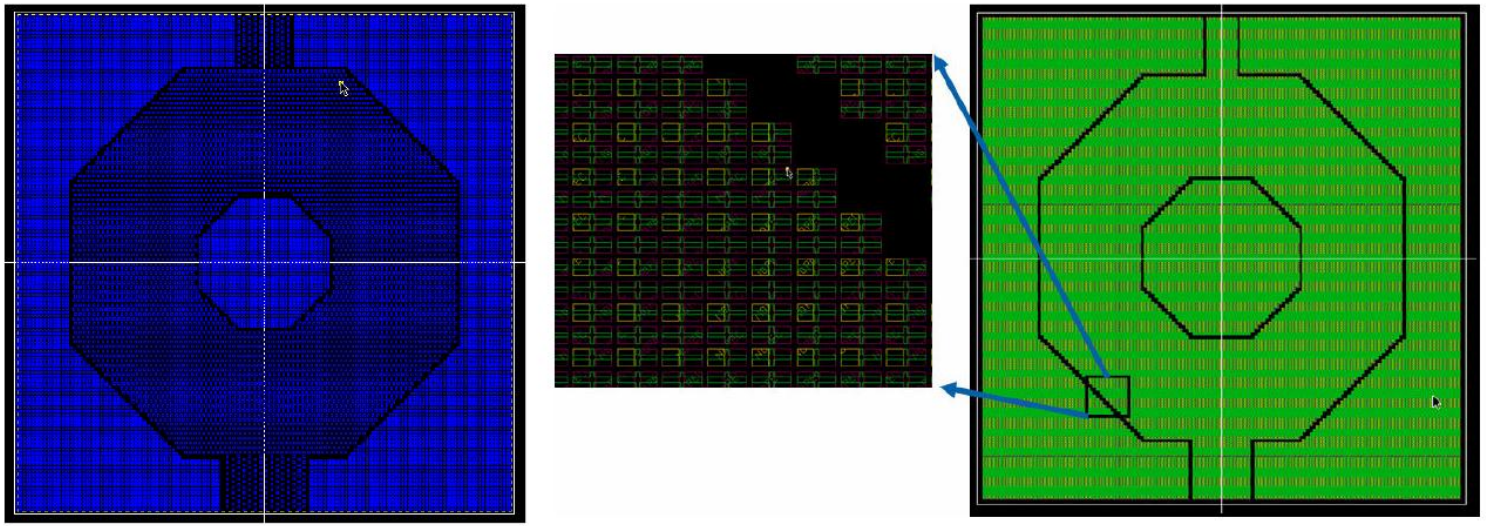


Figure 5.19 inductor with metal fill patterns, top metal (left), n-1 metal (middle), n-2 metal (right)

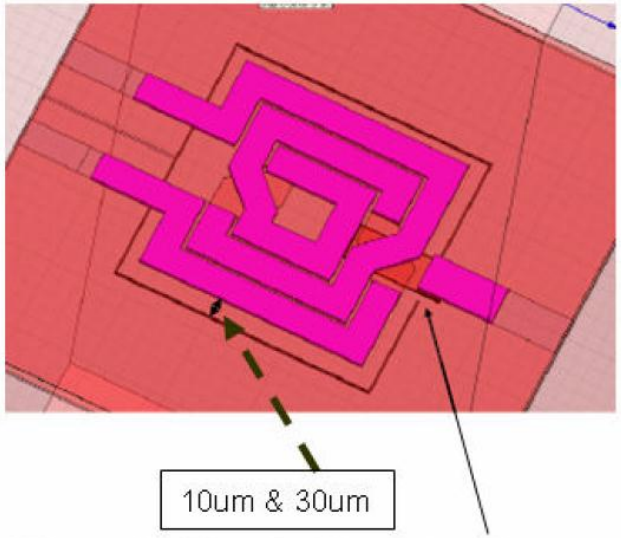


Non stacked dummification is required and yet a 5-10% Q degradation at Qmax is to be expected due to dummies

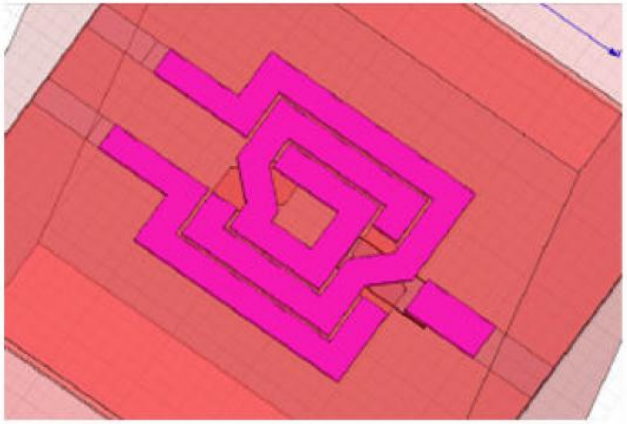
Figure 5.20 inductor with metal fill patterns, metal 1, (left), ploy, ndiff, pdiff (right)

Practical Considerations GR Impact

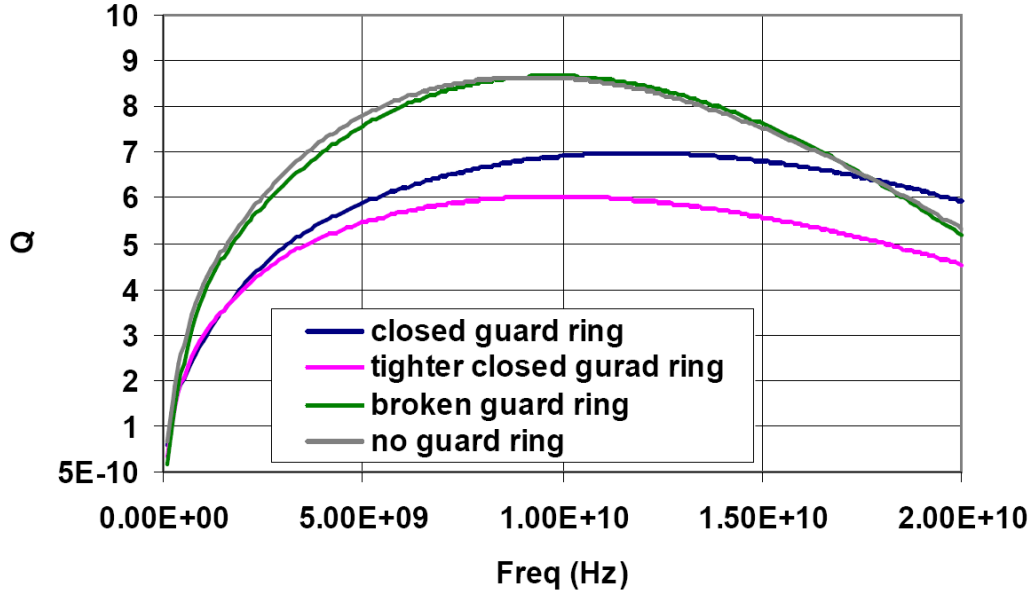
With broken loop M1 tie ring (to accommodate tab exit)



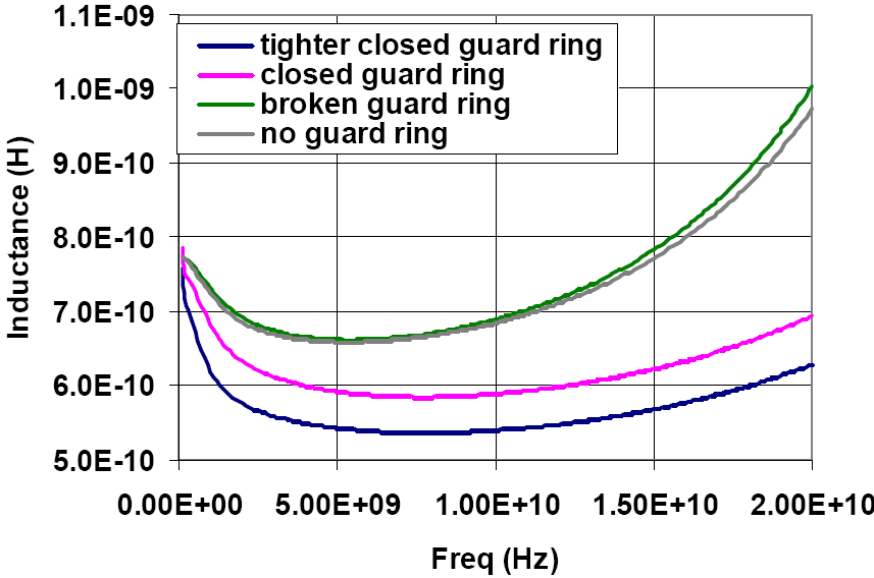
Without tie ring



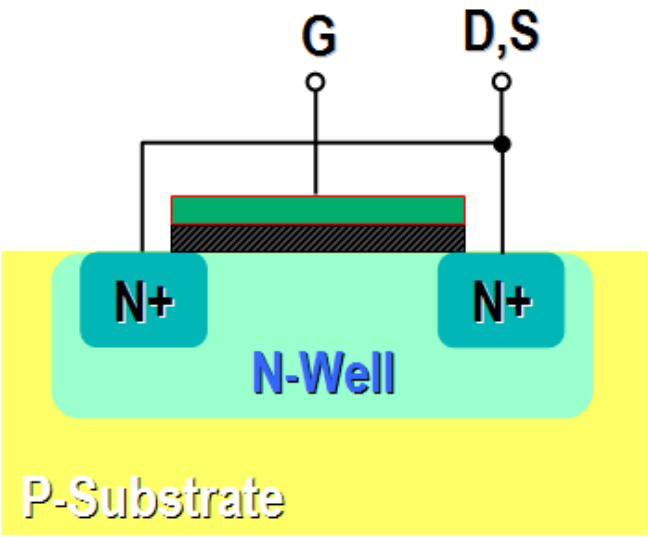
additional simulation to show effect of closed tie ring

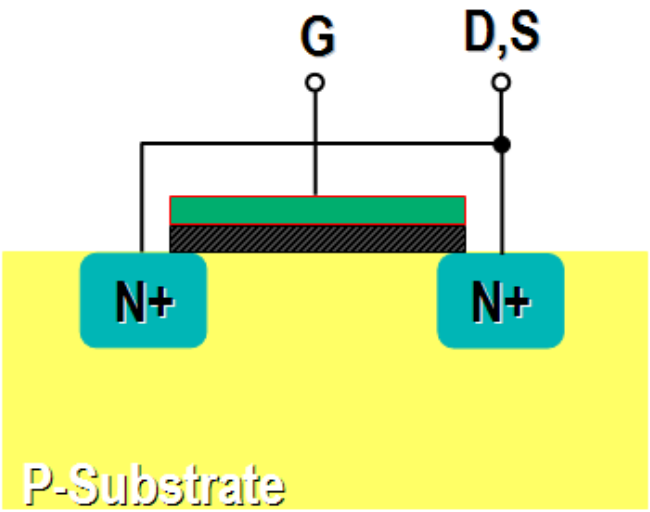


Differential inductance vs. freq

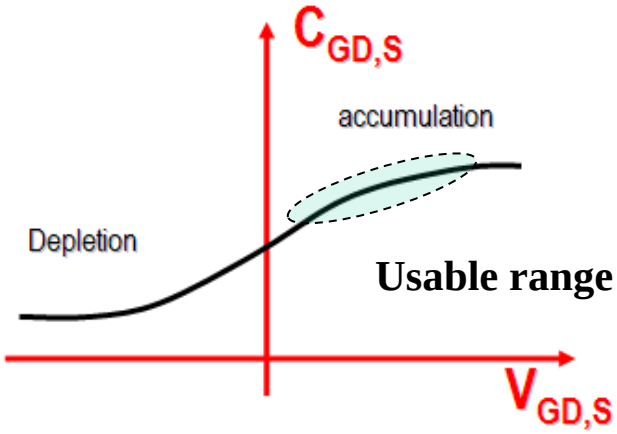


LC-VCO – CMOS Varactor

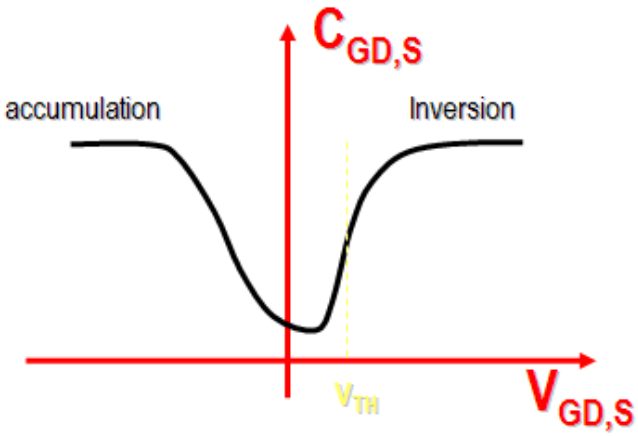




Accumulation Mode



Inversion Mode



Parameters to be considered in a varactor

Quality factor *Q*: must be as large as possible to minimize losses and hence save power and minimize phase noise.

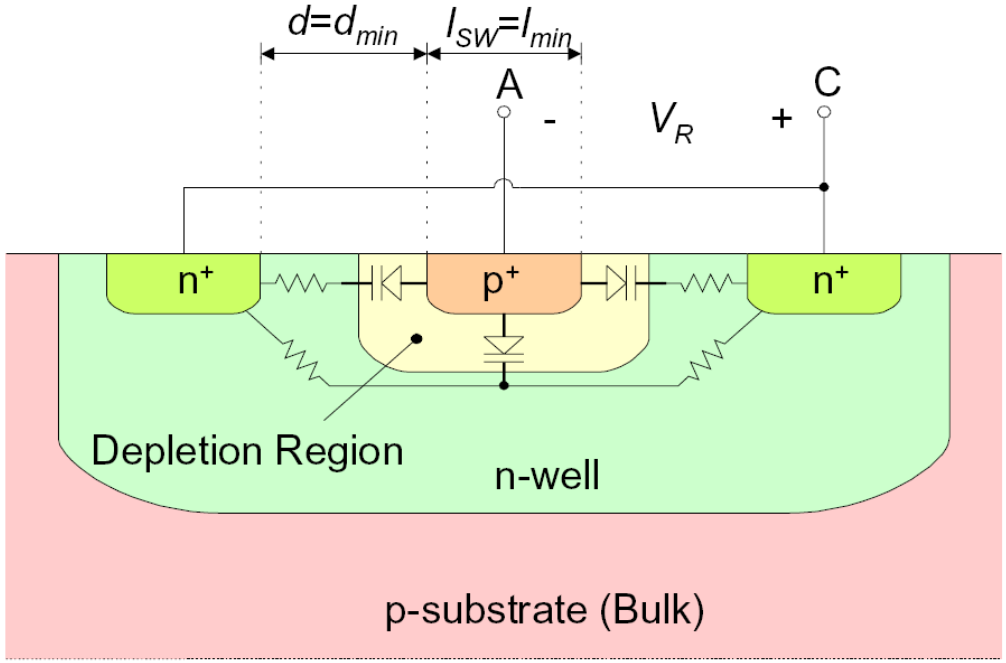
C_{max}/C_{min}: must be as large as possible to maximize the tuning range.

$K_{VCO,max}/K_{VCO,min}$: must be as close to one as possible to give linear frequency tuning.

LC-VCO p-n varactor

P+–to–n-well Junction - Layout Considerations

Maximizing *Q*

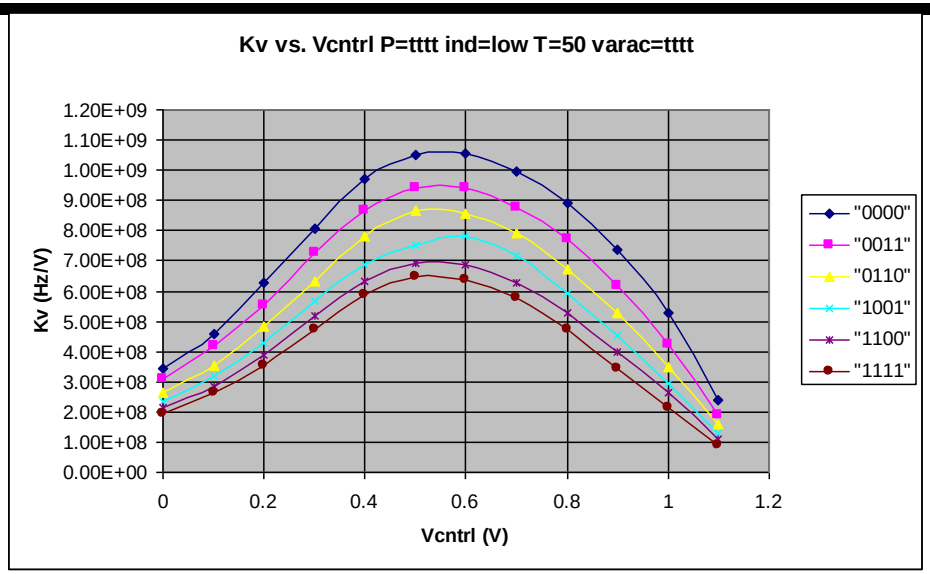


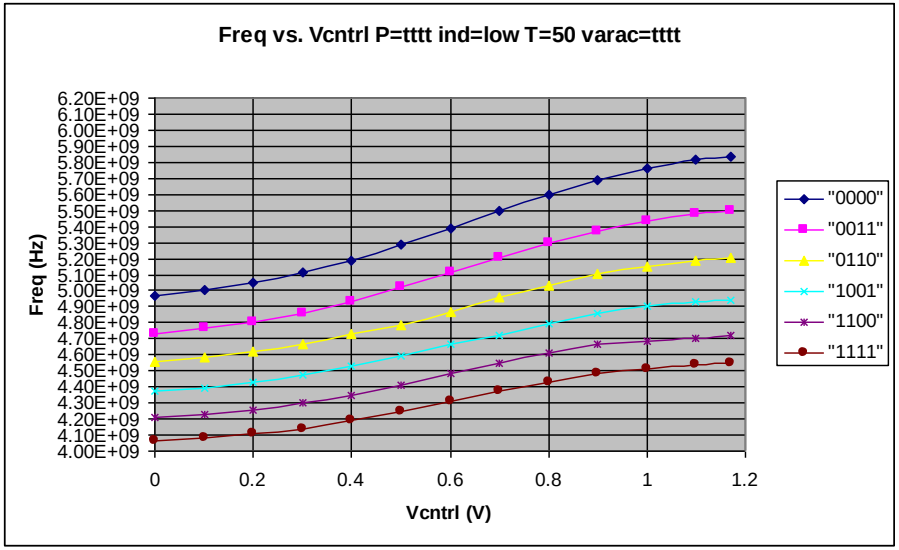
Using minimum dimensions results in:

- maximum Q
- but also minimum C_{max}/C_{min}, due to denser layout and hence larger stray capacitances.

Consequently, layout parameters must be chosen as a compromise between Q and C_{max}/C_{min}.

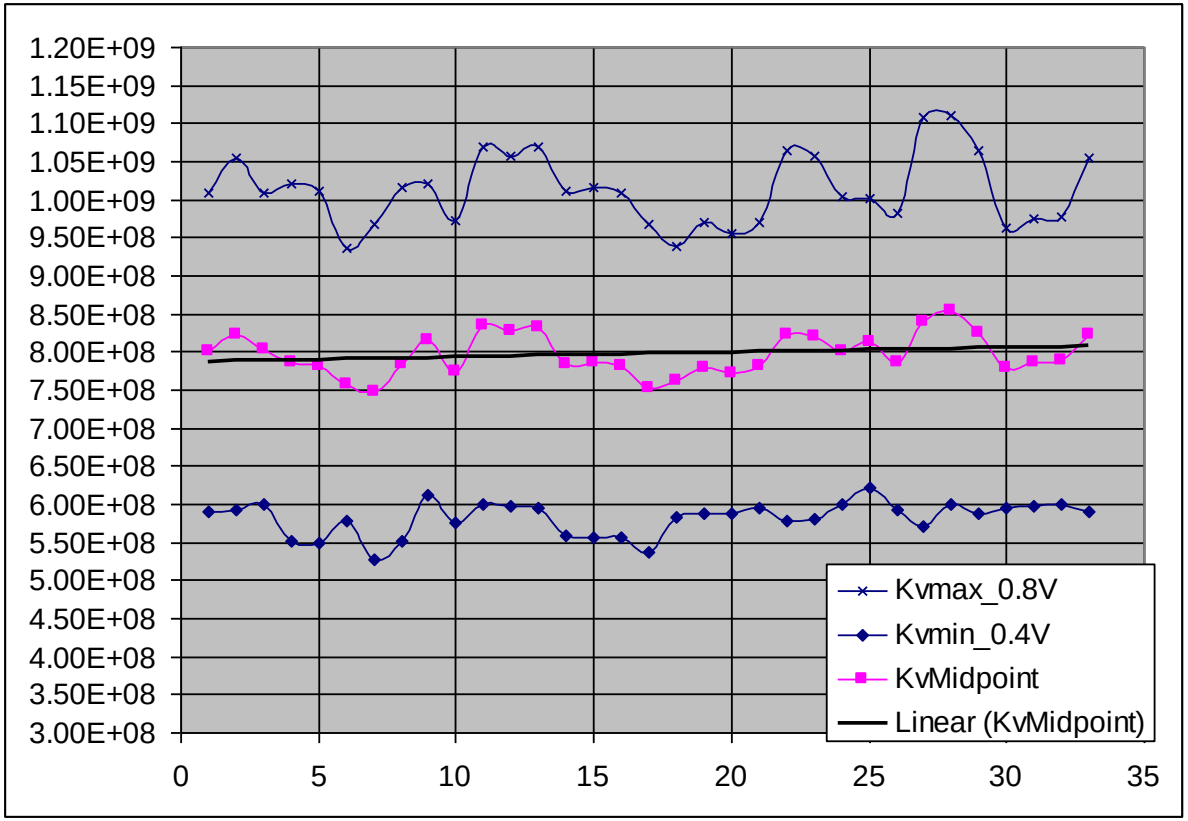
VCO frequency and gain simulations



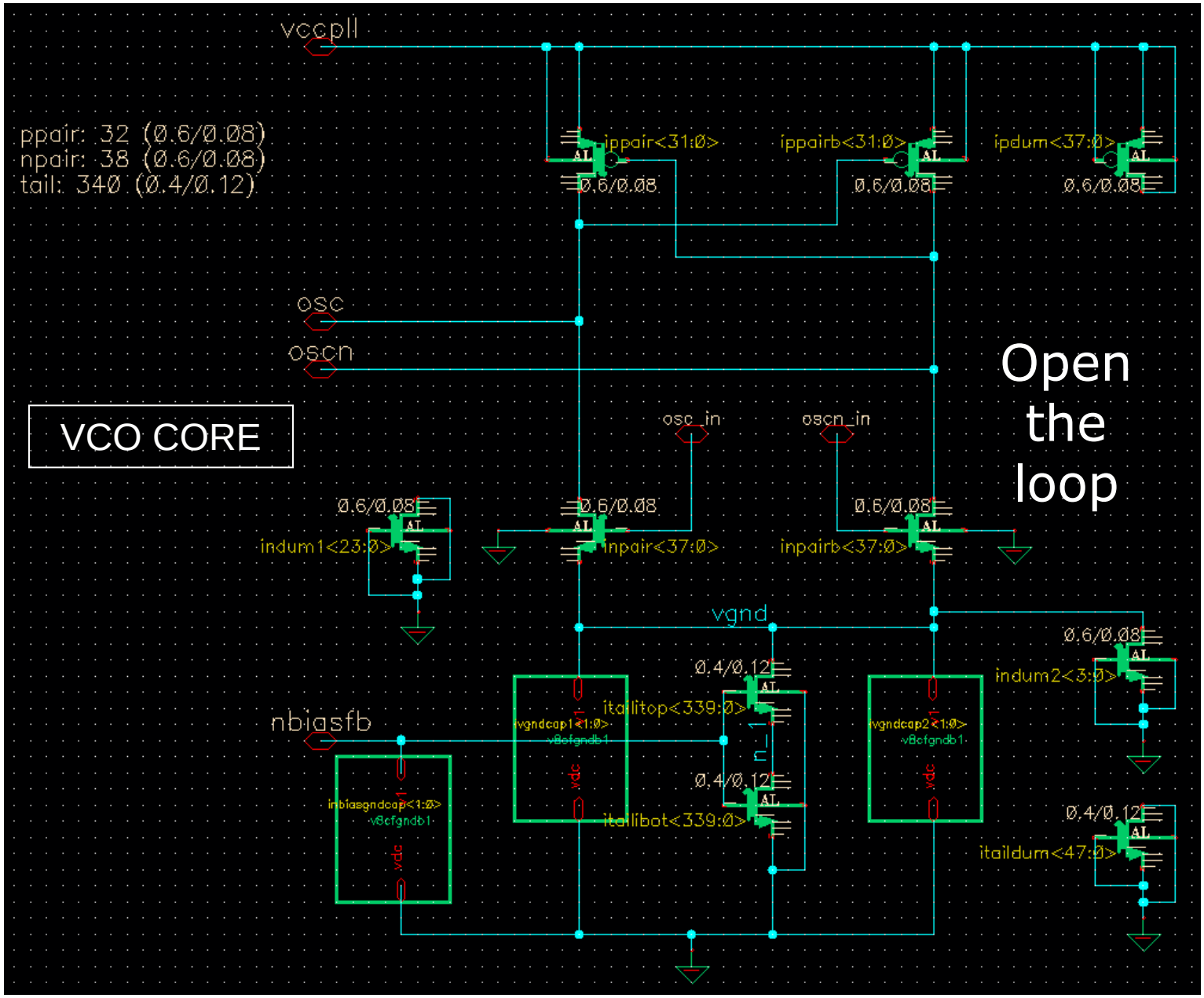


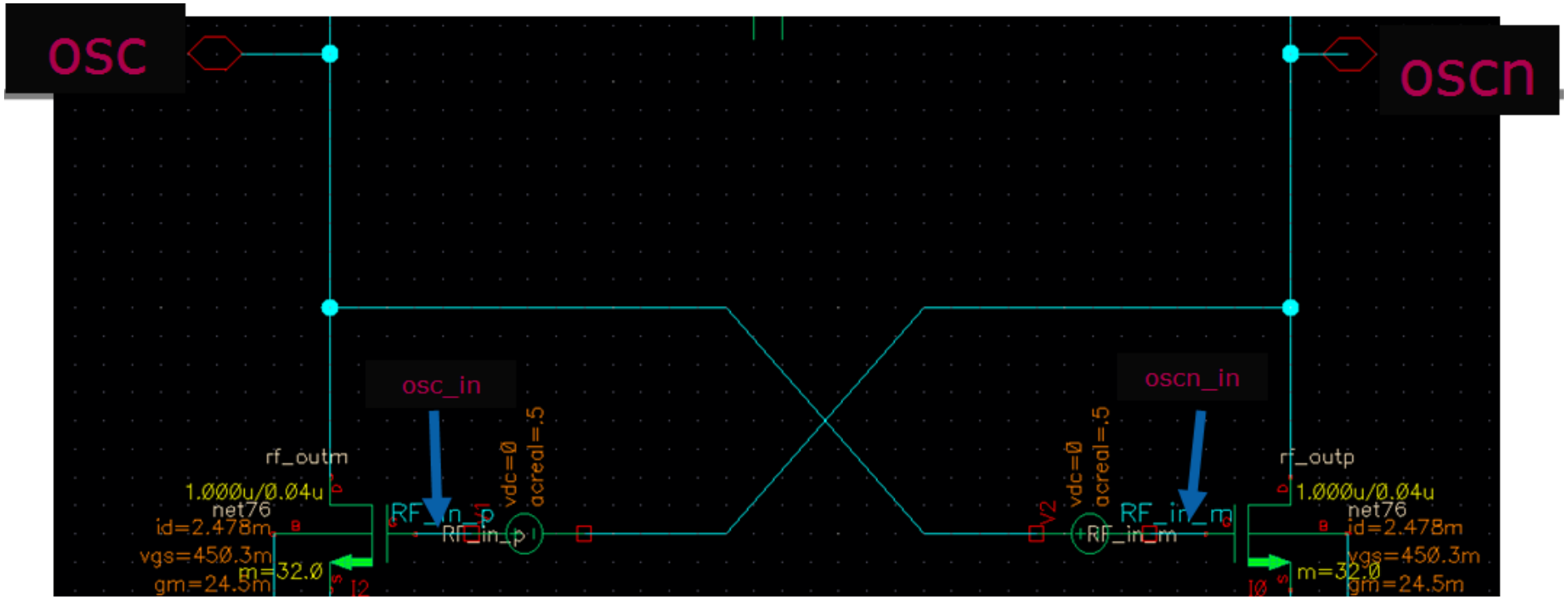
- Right plot shows VCO freq vs. Vctl and Cbank control settings
  - Cbank control 4 bits
  - Vctl range considered 0.4-0.8V
  - Curve spacing ~ 60MHz (averaged over useful control range)
  - For fixed Cbank control frequency varies ~ +/-150MHz over PVT
    - Provided the above locking range prediction - roughly between curves (0001-1001), i.e. from light blue to dark blue.
- On the left plot showed VCO gain change over Vctl
  - Maximum Kv is ~ 1000 MHz/V, min Kv is ~ 600MHz/V (Vcntrl=0.4V)

VCO Gain range

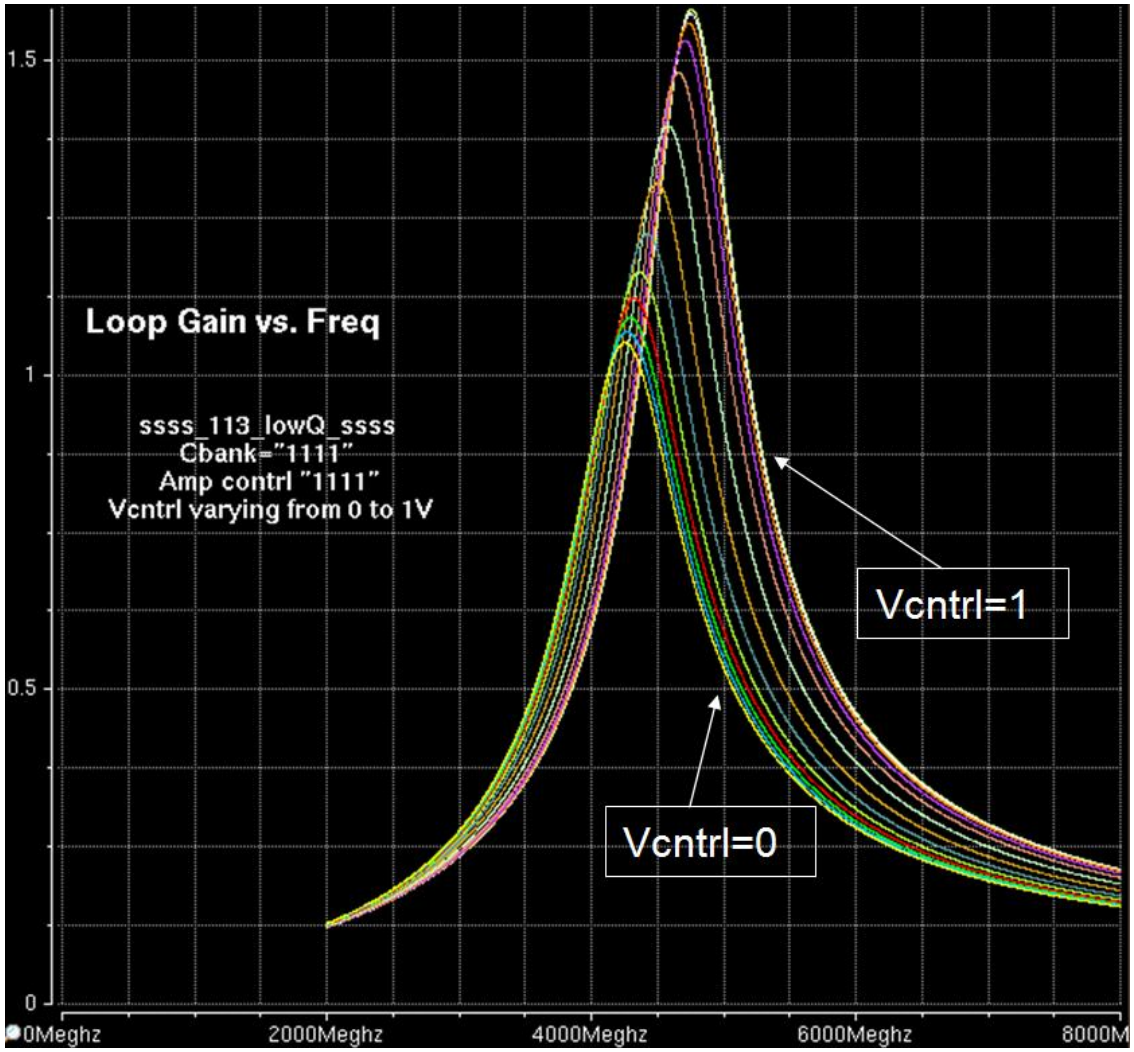


- VCO gain range has been simulated over PVT in assumption:
  - Vctl range 0.4 to 0.8V
  - Kv variation is 600-1000 MHz/V with a median ~ 800MHz/v


vccpll
ppair: 32 (0.6/0.08)
npair: 38 (0.6/0.08)
tail: 340 (0.4/0.12)
ippair<31:0>
ippairb<31:0>
ipdum<37:0>
0.6/0.08
osc
oscn
VCO CORE
osc_in
oscn_in
Open the loop
indum1<23:0>
inpair<37:0>
inpairb<37:0>
vgnd
indum2<3:0>
0.4/0.12
itailtop<339:0>
itailbot<339:0>
itaildum<47:0>
nbiasfb
vgndcap1<1:0>
vgndcap2<1:0>
inbiasgndcap<1:0>
vdc



Loop gain = (osc –oscn)/(osc_in-oscn_in)


Loop Gain vs. Freq
ssss_113_lowQ_ssss
Cbank="1111"
Amp contrl "1111"
Vcntrl varying from 0 to 1V
Vcntrl=1
Vcntrl=0
1.5
1
0.5
0
0Meghz
2000Meghz
4000Meghz
6000Meghz
8000M

	x6clcvco	x6clcvco	x6clcvco	x6clcvco	x6clcvco	x6clcvco	x6clcvco	x6clcvco	x6clcvco	x6clcvco	x6clcvco	x6clcvco	x6clcvcotoppostsim_
frequency0<MegaHz>	NONE	NONE	NONE	NONE	NONE	NONE	NONE	NONE	NONE	4431.11	4491.43	4515.73	4525.67
max0<V>	0.661023	0.661173	0.661375	0.661656	0.662091	0.662922	0.665121	0.678849	0.716951	0.745287	0.755869	0.760659	0.762565
min0<V>	0.599	0.599	0.599	0.599	0.599	0.599	0.599	0.599	0.599	0.584135	0.576198	0.573025	0.571842

Reference VCC: 1.17

Wave Selector

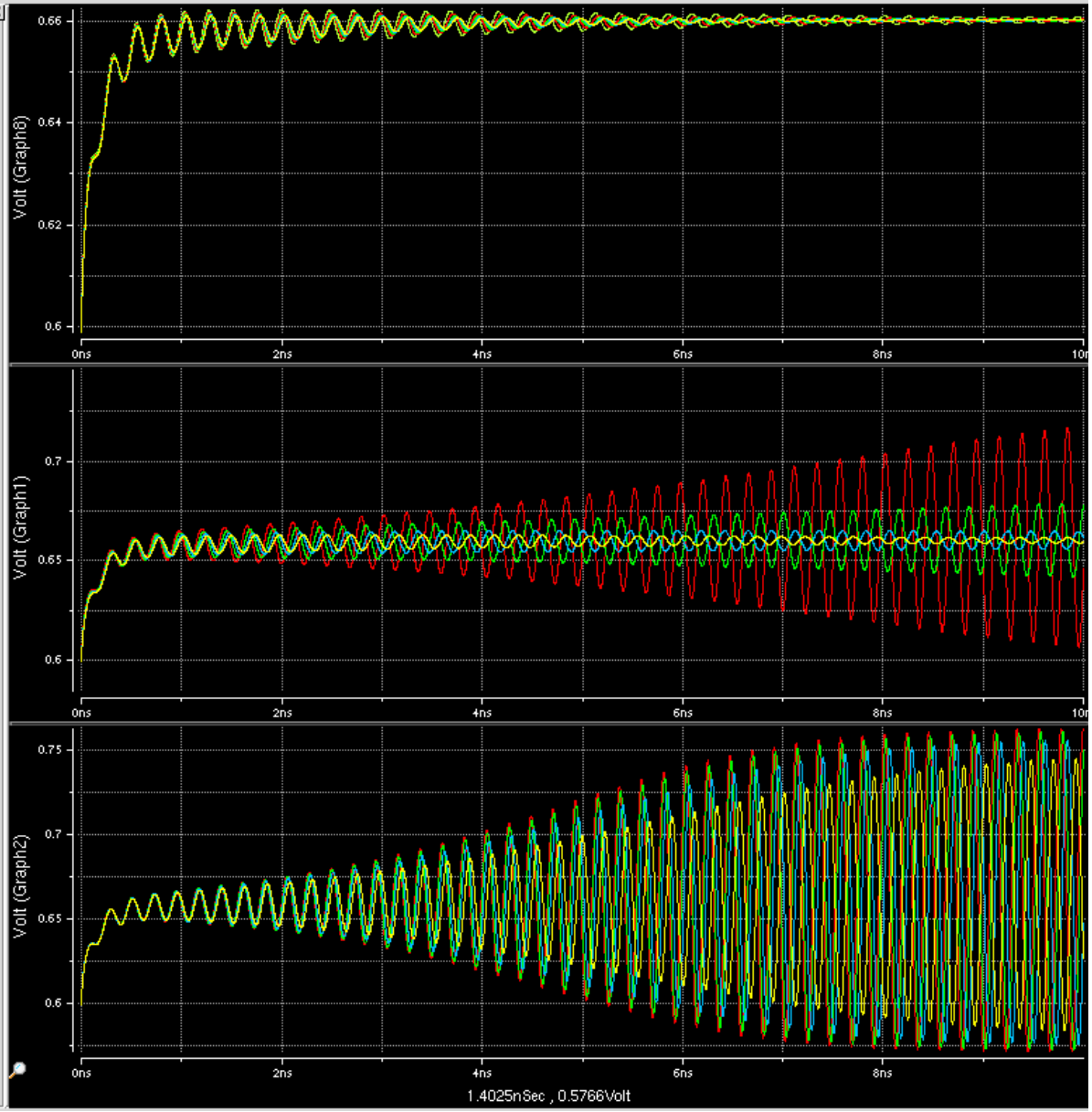
Run: x6clcvcotoppostsim_exp_post_tail

Find:

x6clcvcotoppostsim_exp_post_tail_1
x6clcvcotoppostsim_exp_post_tail_2
x6clcvcotoppostsim_exp_post_tail_3
x6clcvcotoppostsim_exp_post_tail_4
x6clcvcotoppostsim_exp_post_tail_5
x6clcvcotoppostsim_exp_post_tail_6
x6clcvcotoppostsim_exp_post_tail_7
x6clcvcotoppostsim_exp_post_tail_8
x6clcvcotoppostsim_exp_post_tail_9
x6clcvcotoppostsim_exp_post_tail_10
x6clcvcotoppostsim_exp_post_tail_11
x6clcvcotoppostsim_exp_post_tail_12
x6clcvcotoppostsim_exp_post_tail_13

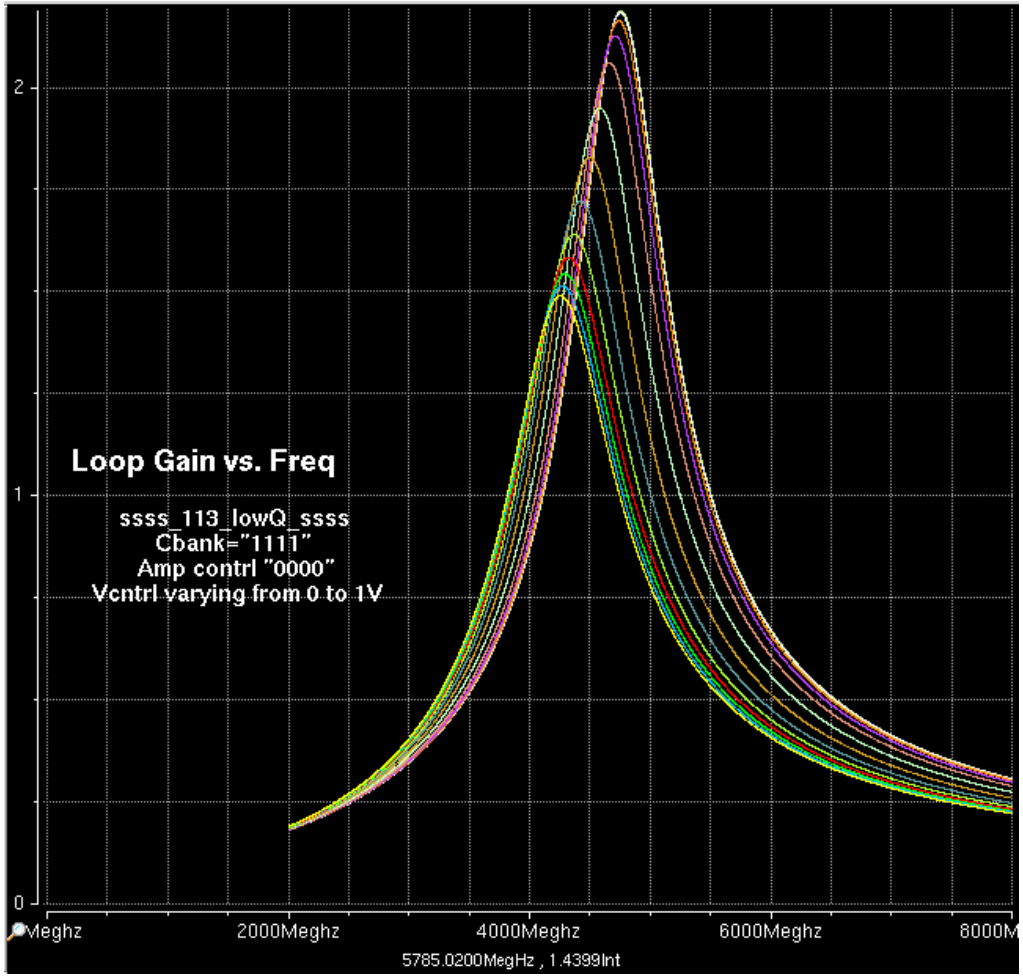
Find:

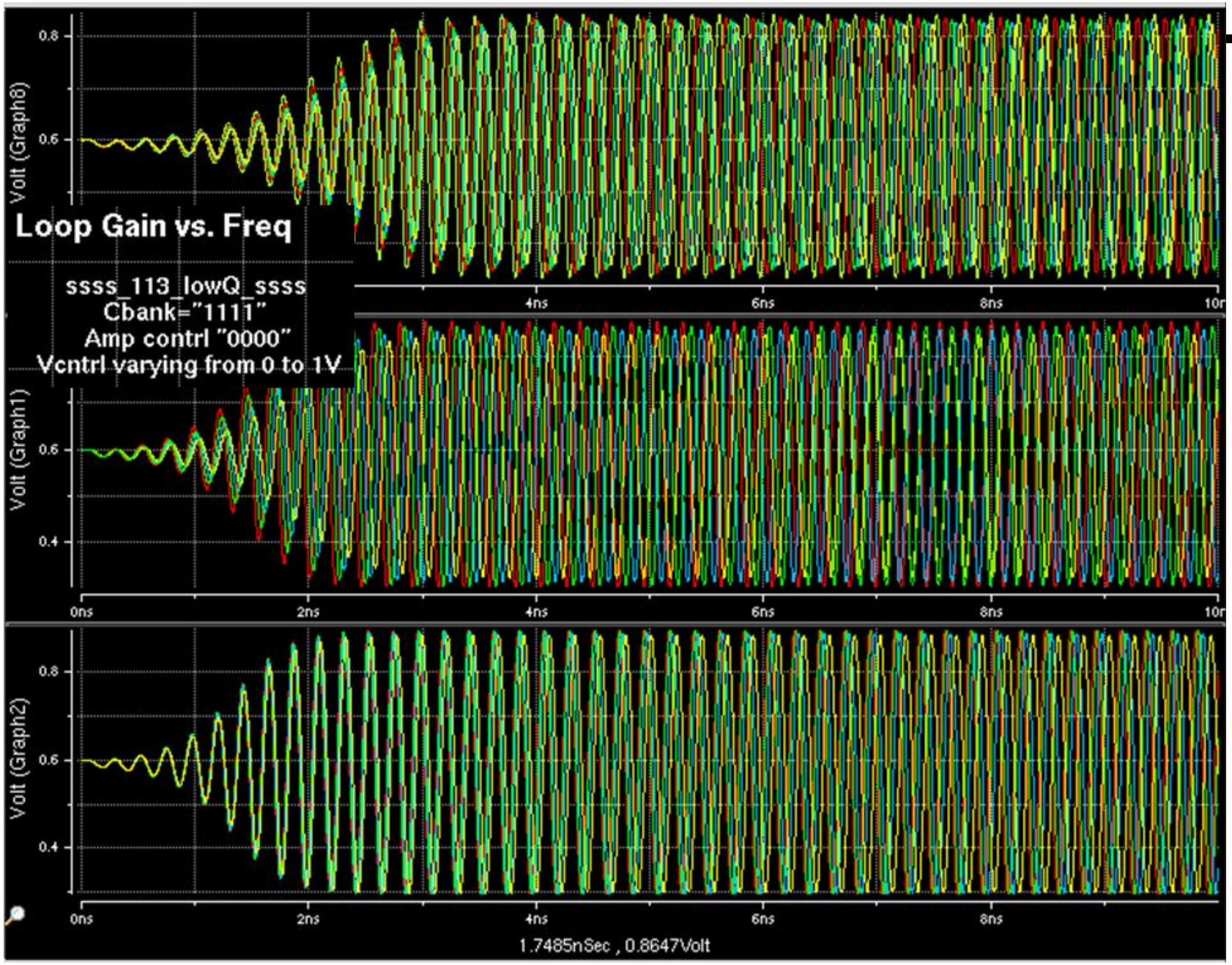
V x6clcvcotoppostsim/llcvco/idiffcore:vgnd
VP x6clcvcotoppostsim%osc
VP x6clcvcotoppostsim%oscn
VP x6clcvcotoppostsim%vctl
VP x6clcvcotoppostsim/llcvco/idiffcore%nbiasfb



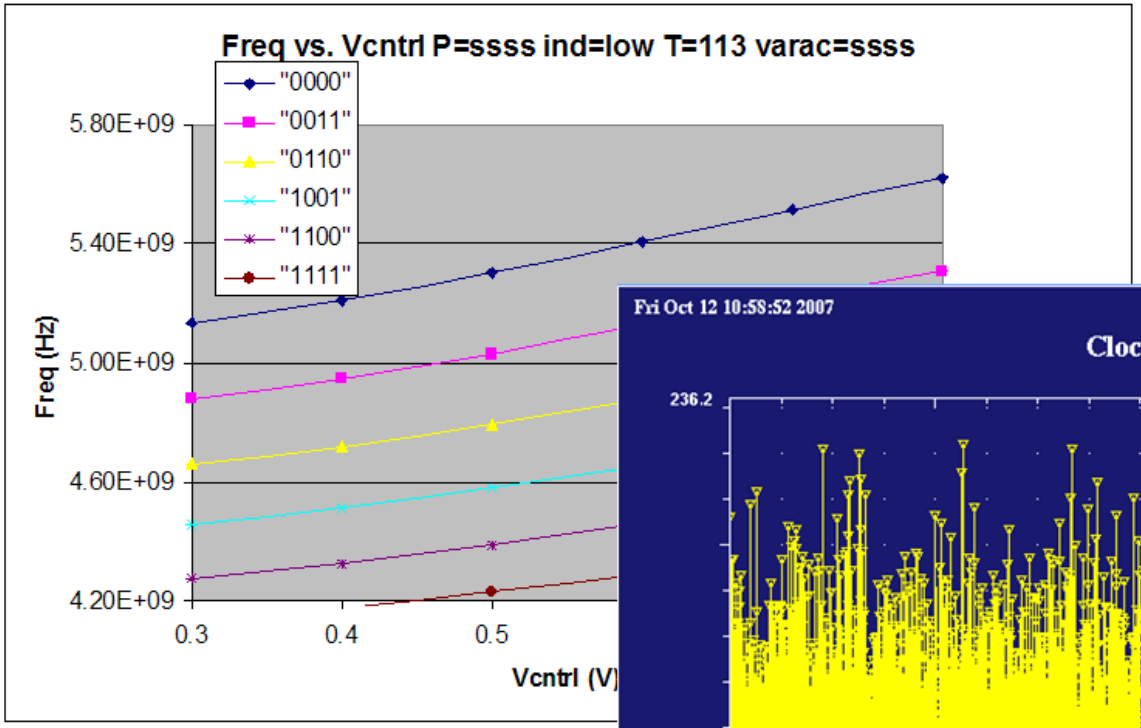
Loop Gain vs. Freq

ssss_113_lowQ_ssss
Cbank="1111"
Amp contrl "1111"
Vcntrl varying from 0 to 1V

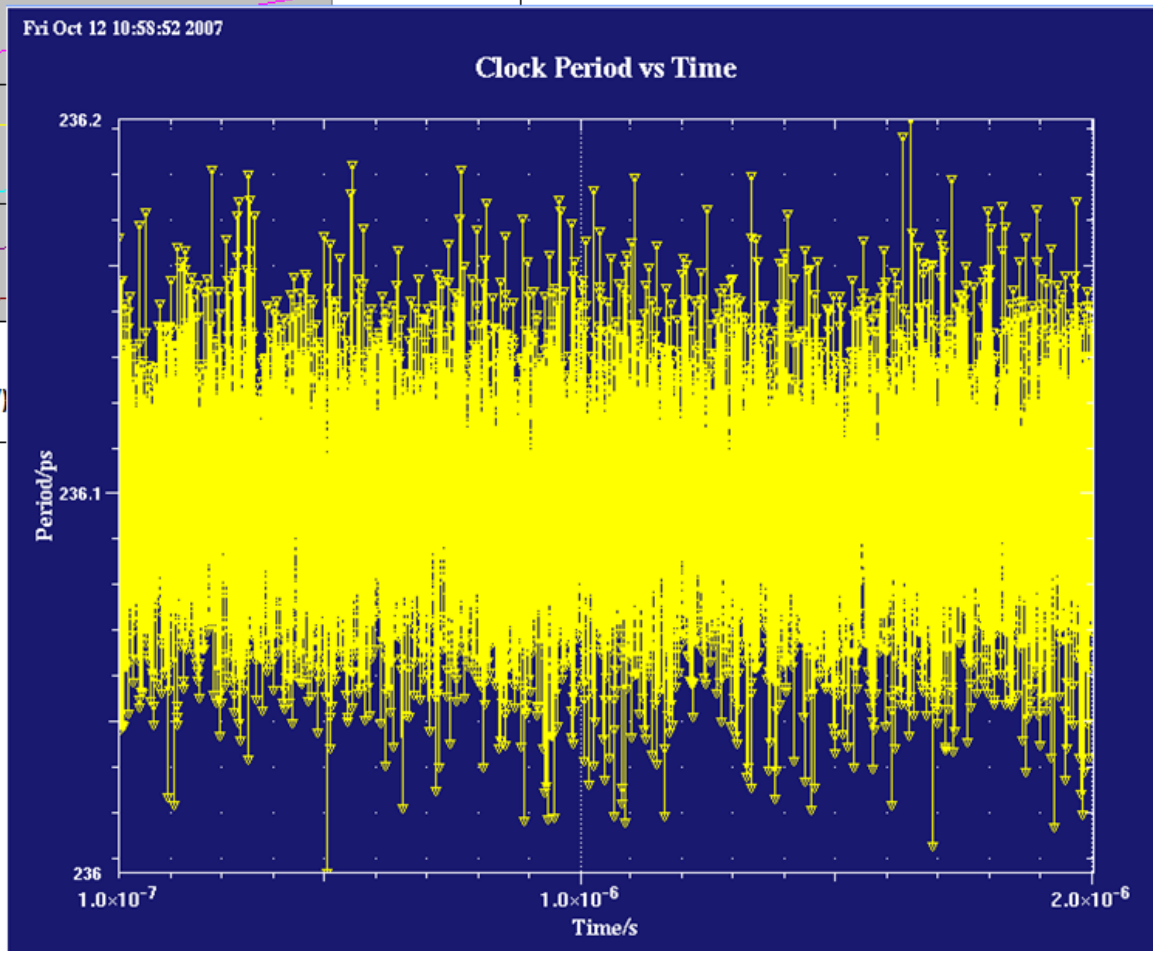

Loop Gain vs. Freq
ssss_113_lowQ_ssss
Cbank="1111"
Amp contrl "0000"
Vcntrl varying from 0 to 1V
2
1
0
Meghz
2000Meghz
4000Meghz
6000Meghz
8000M
5785.0200MegHz , 1.4399Int


Loop Gain vs. Freq
ssss_113_lowQ_ssss
Cbank="1111"
Amp contrl "0000"
Vcntrl varying from 0 to 1V
Volt (Graph8)
Volt (Graph1)
Volt (Graph2)
0.8
0.6
0.4
0ns
2ns
4ns
6ns
8ns
1.7485nSec , 0.8647Volt

split_name	vctl	a0	a1	a2	a3	myskew	mytemp	4n_2port_indu	n_diff_60f_varactor_skew
kappa_ahmed_all_8	0.5	1.17	1.17	1.17	1.17	ssss	113	lowQ	ssss

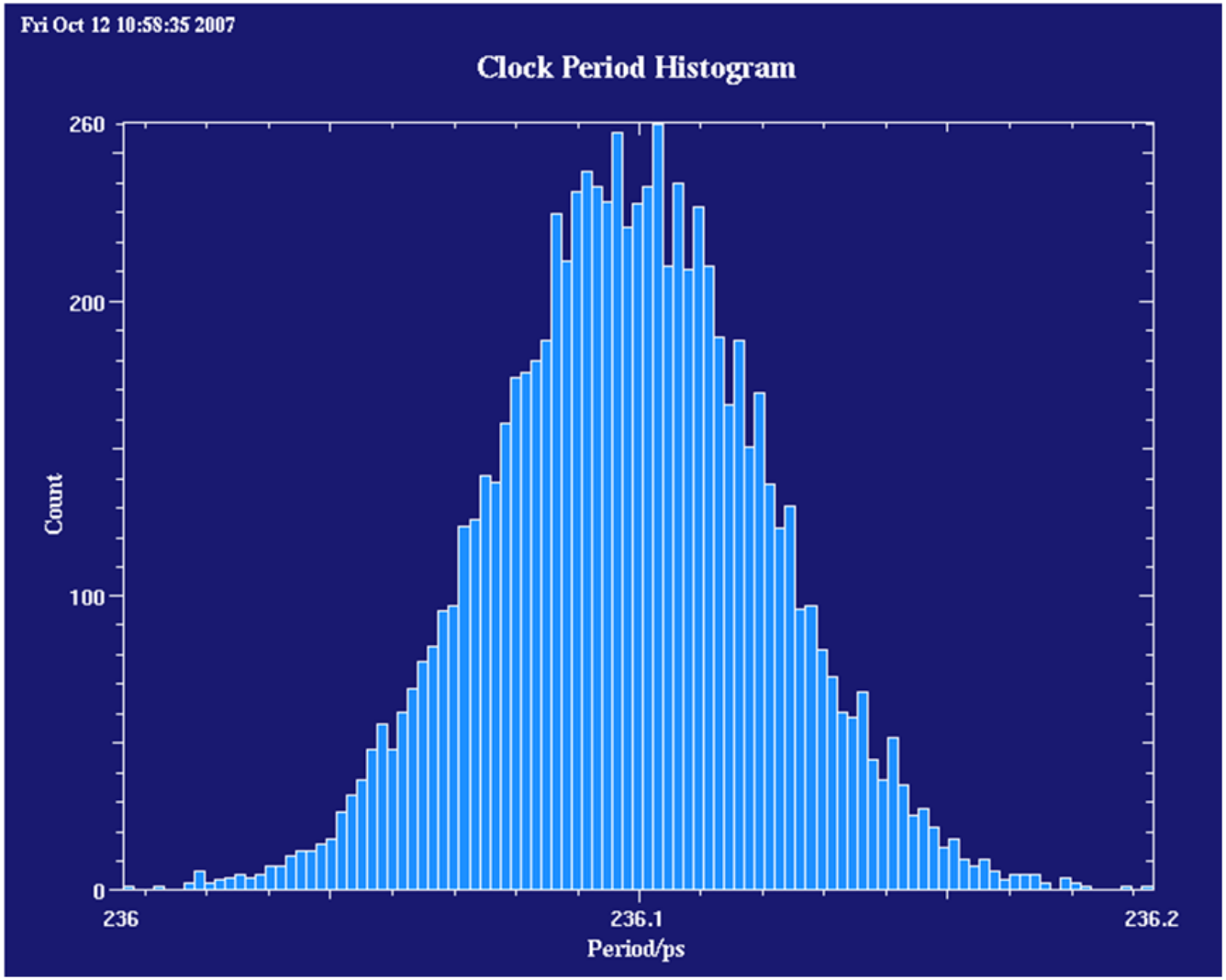


4.23 GHz

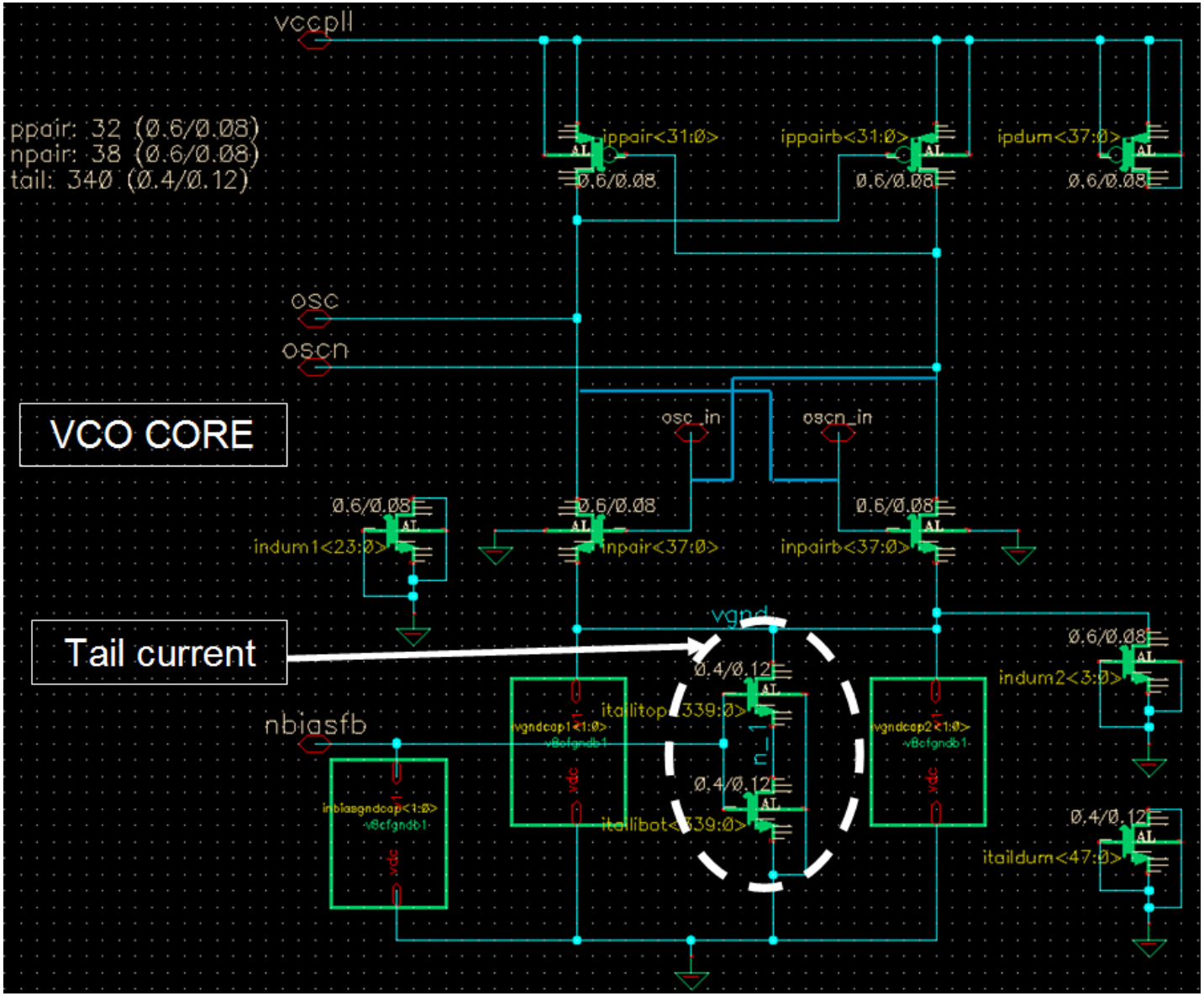


Corner #6

split_name	vctl	a0	a1	a2	a3	myskew	mytemp	4n_2port_indun	_diff_60f_varactor_skew
kappa_ahmed_all_8	0.5	1.17	1.17	1.17	1.17	ssss	113	lowQ	ssss



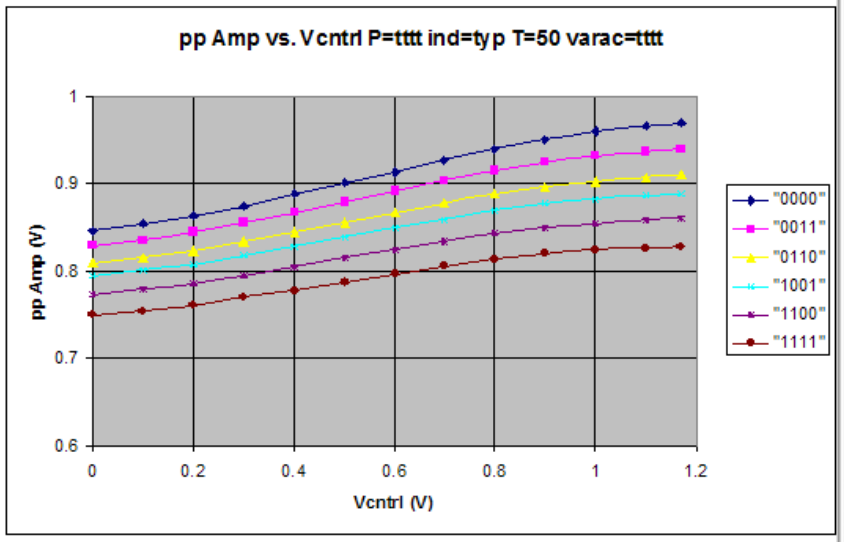
4.23 GHz


vccpll
ppair: 32 (0.6/0.08)
npair: 38 (0.6/0.08)
tail: 340 (0.4/0.12)
ippair<31:0>
ippairb<31:0>
ipdum<37:0>
0.6/0.08
osc
oscn
VCO CORE
osc_in
oscn_in
indum1<23:0>
inpair<37:0>
inpairb<37:0>
vgnd
Tail current
0.4/0.12
indum2<3:0>
nbiasfb
vgndcap1<1:0>
vgndcap2<1:0>
inbiasgndcap<1:0>
itailitop<339:0>
itailibot<339:0>
itaildum<47:0>

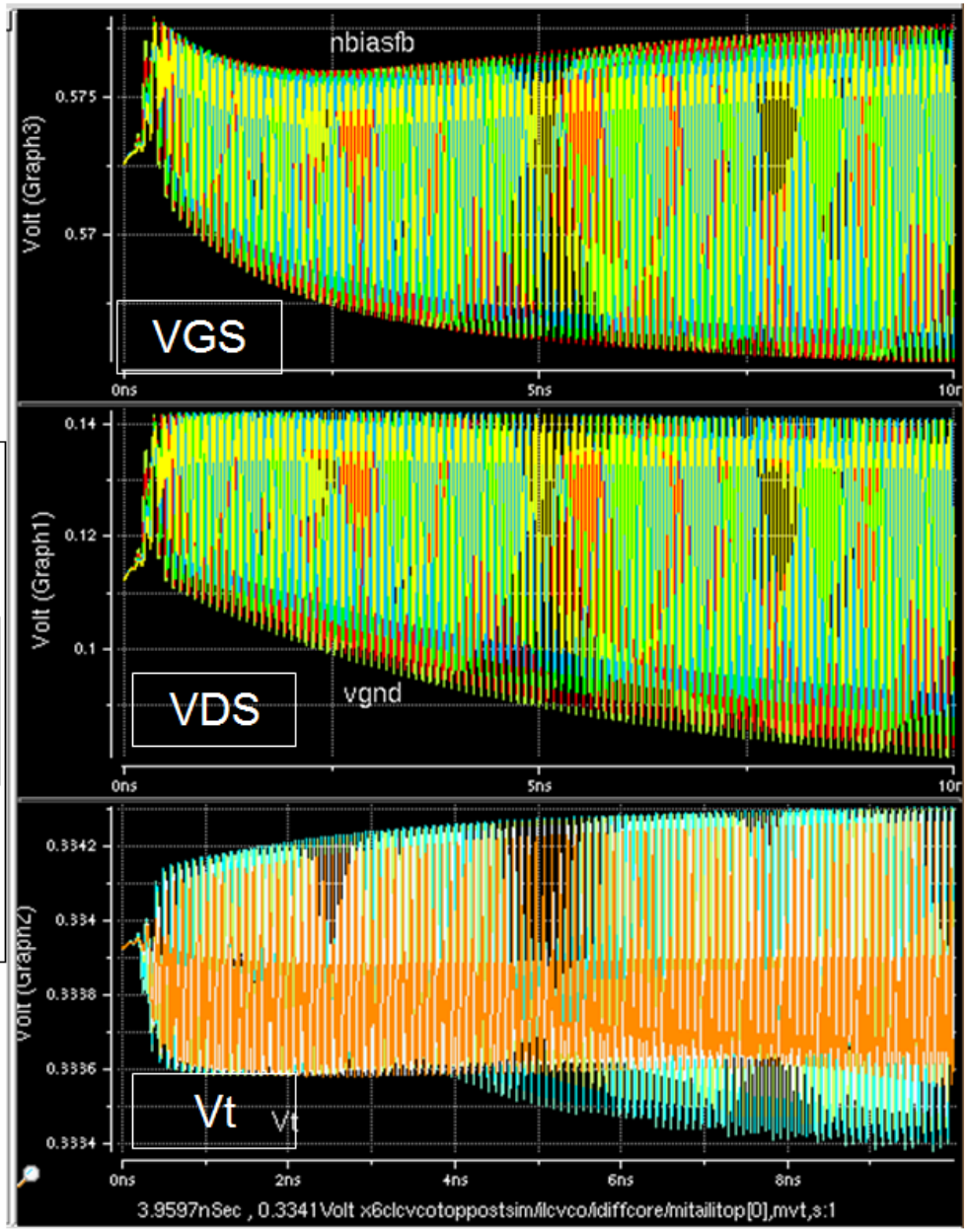
	Parameter	Discrete
1	a0	=vss()
2	a1	=vss()
3	a2	=vss()
4	a3	=vss()
5	hertz	
6	myskew	tttt
7	mytemp	50.0
8	sel0	=vss()
9	sel1	=vss()
10	sel2	=vss()
11	sel3	=vss()
12	1p4n_2port_inductor_skew	typ0
13	n_diff_60f_varactor_skew	tttt
14	vctl	7 0.8 0.9 1.0 1.1 1.17

Max Freq

Max Amp



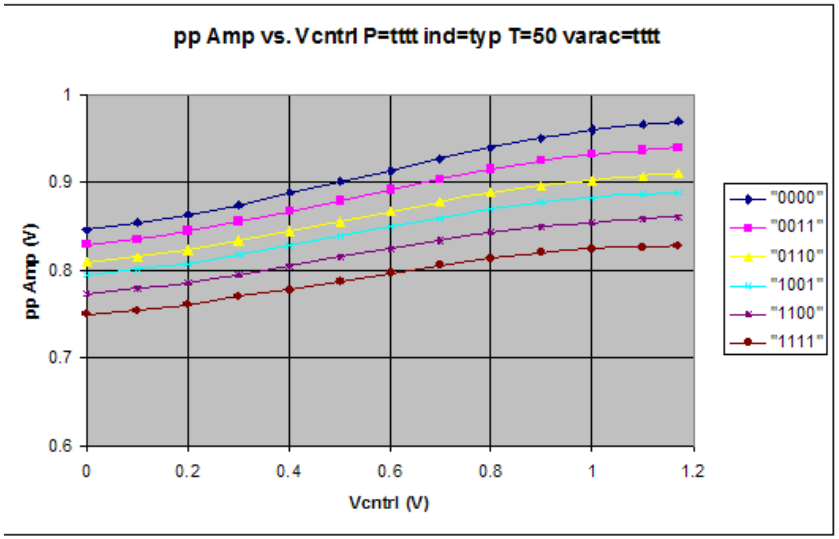
DEEP TRIODE

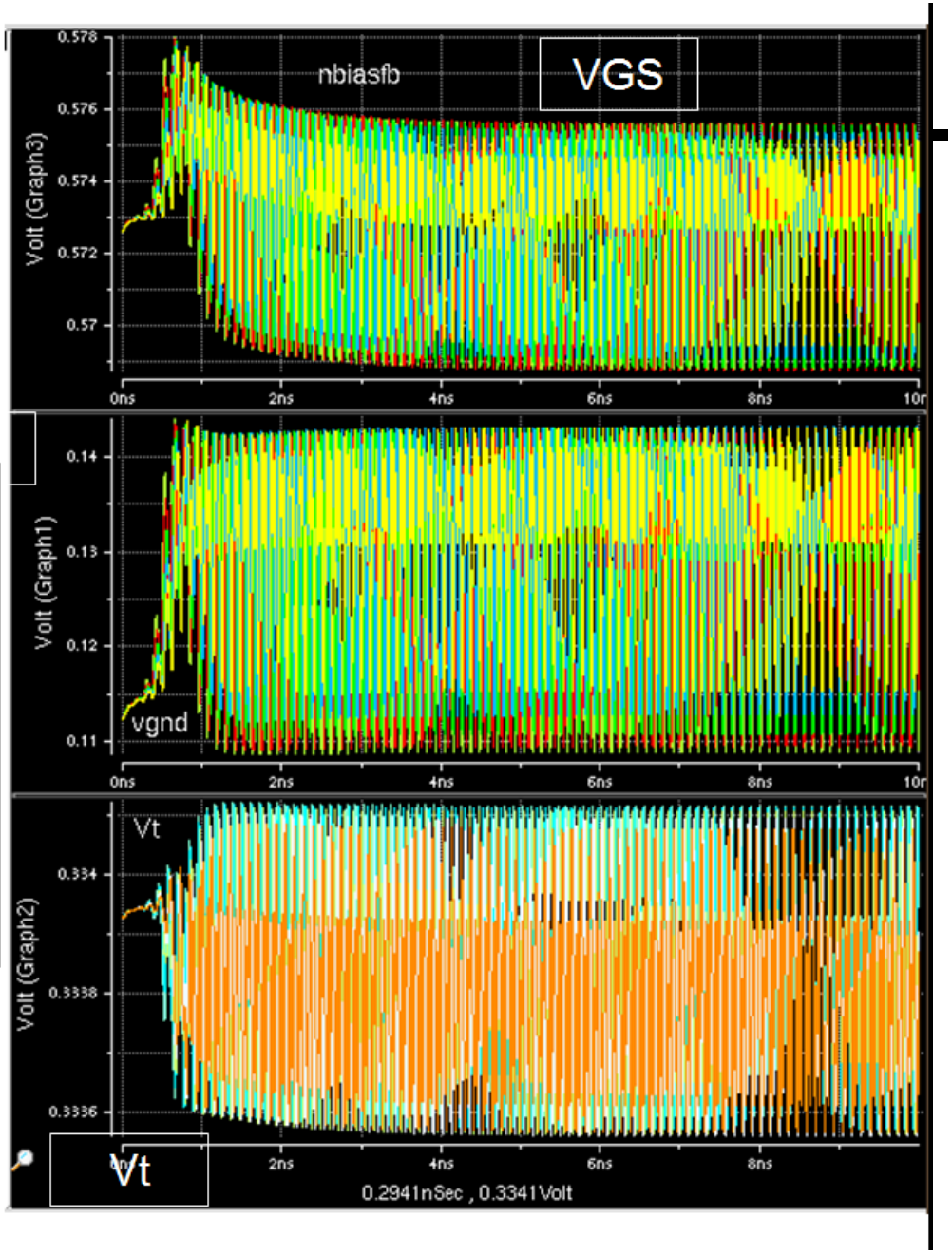


	Parameter	Discrete
1	a0	=vcc()
2	a1	=vcc()
3	a2	=vcc()
4	a3	=vcc()
5	hertz	
6	myskew	tttt
7	mytemp	50.0
8	sel0	=vss()
9	sel1	=vss()
10	sel2	=vss()
11	sel3	=vss()
12	1p4n_2port_inductor_skew	typ0
13	n_diff_60f_varactor_skew	tttt
14	vctl	0.4 0.5 0.6 0.7 0.8

Min Freq

Max Amp





TRIODE

- Tail current transistor goes into triode for high amp settings
- This will impact PN at such settings
- This should be prevented by the aac

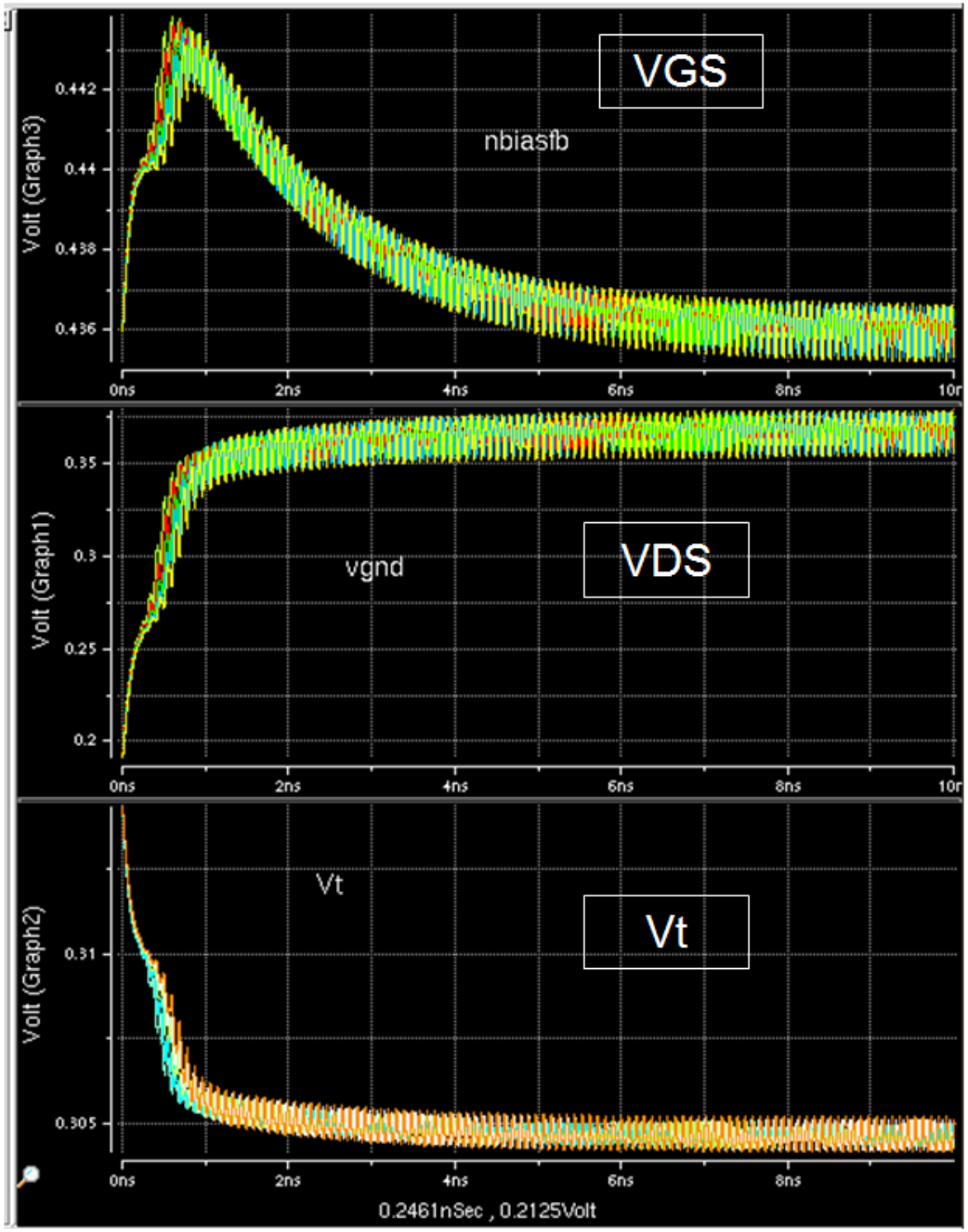
	Parameter	Discrete
1	a0	=vss()
2	a1	=vss()
3	a2	=vss()
4	a3	=vss()
5	hertz	
6	myskew	tttt
7	mytemp	50.0
8	sel0	=vcc()
9	sel1	=vcc()
10	sel2	=vcc()
11	sel3	=vcc()
12	1p4n_2port_inductor_skew	typQ
13	n_diff_60f_varactor_skew	tttt
14	vctl	0.4 0.5 0.6 0.7 0.8

Max Freq

Min Amp

	x6clcvco	x6clcvco	x6clcvco	x6clcvco	x6clcvcotop
frequency0<MegaHz>	5361.13	5441.29	5548.0	5675.16	5796.62
max0<V>	1.01151	1.01724	1.02375	1.03164	1.04137
min0<V>	0.449417	0.437527	0.423509	0.410887	0.401244

Saturation



Conclusion

- Deep submicron LC-VCO design has its own challenges
- Designing high Q inductors specially in CMOS technologies is a key challenge
- Picking L, C value is imp to trade off, PN, Swing, freq of osc, and power
- Design rules are making it even worse
- Use CMOS architecture
- Design for low PN
- Low Kv and Cbank selection is a key for optimum tuning range, PN tradeoff
- Make sure that oscillation can start at worst case

Analog designers think a lot about matching

- *Basic operation of some circuits depends directly on matching*

e.g. Diff amps

- *Common-mode rejection limited by matching*

- *Supply noise rejection limited by mismatch in fully-differential circuits*

In fully-differential circuits supply noise appears like a common-mode signal applied to the supply

- *Amplifier offsets greatly degrade the performance of bandgap references*

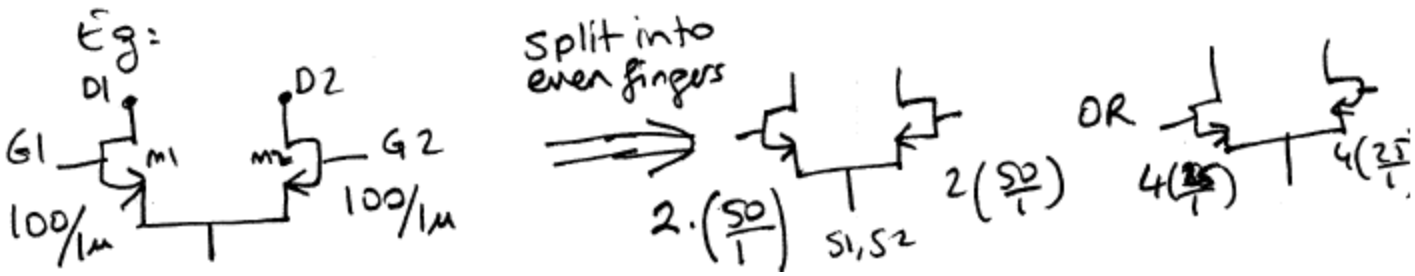
Overview of good layout practices for amplifiers.

The Art of Analog Layout by [Alan Hastings]

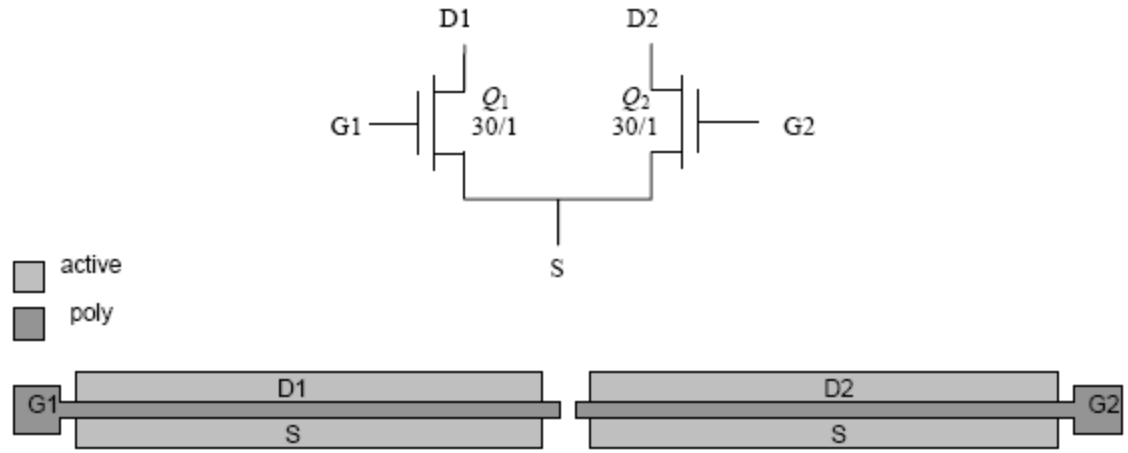
Adobe Acrobat Document

1. ***Diff pairs***

Always use common centroid technique

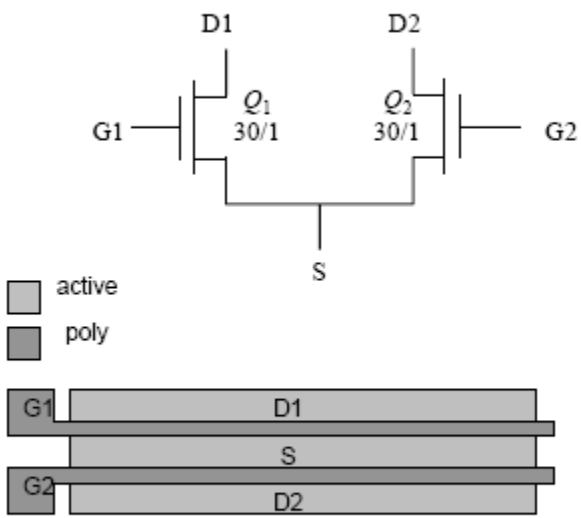


Basic Differential Pair Layout



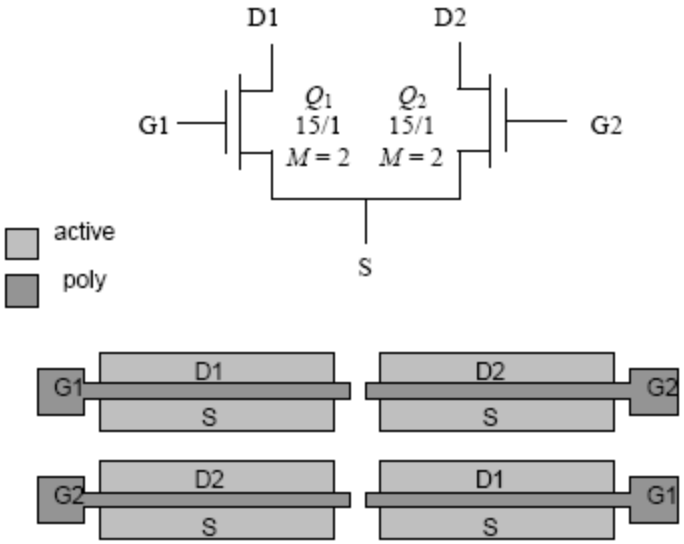
Good matching in the absence of cross-chip gradients; both drain currents flow in same direction.

Alternative Differential Pair Layout



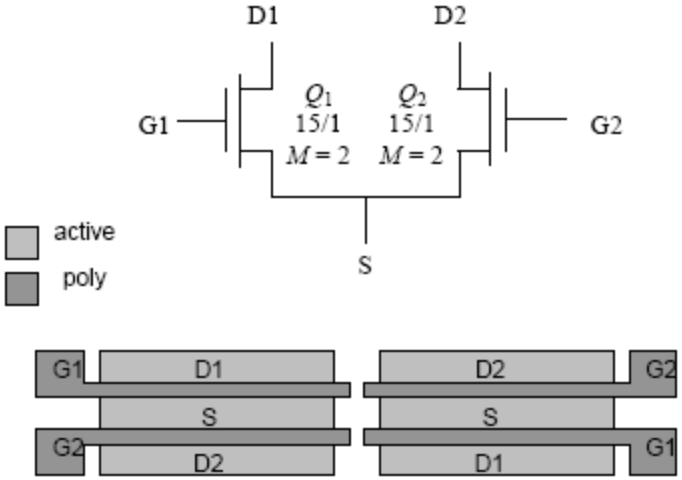
More compact, but worse matching than previous case; drain currents flow in opposite directions.

Common Centroid Layout



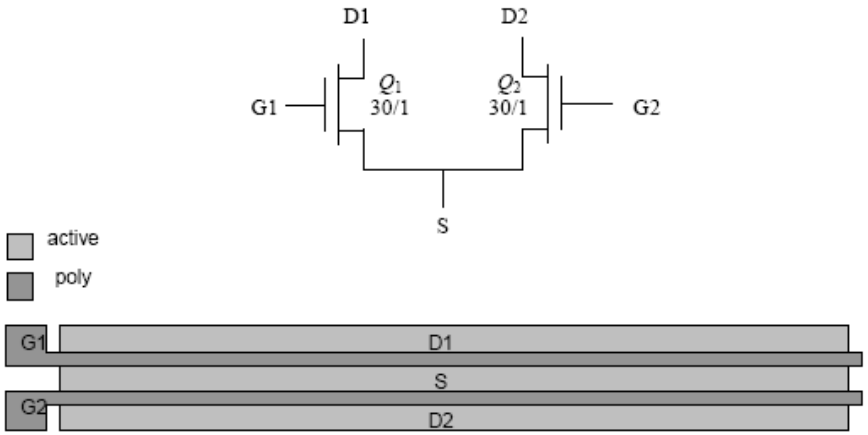
Q_1 and Q_2 have a "common centroid", which makes them immune from cross-chip gradients. Best matching performance possible.

Alternative Common-Centroid Layout



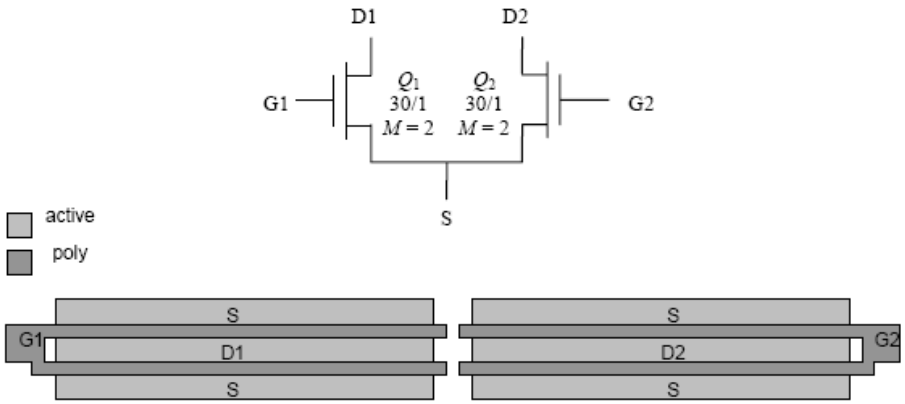
Immune from cross-chip gradients like previous case, but area is saved by sharing sources.

Differential Pair with Very Wide Transistors



Very wide transistors can lead to awkward layout and significant series gate resistance.

Multi-Finger Transistors

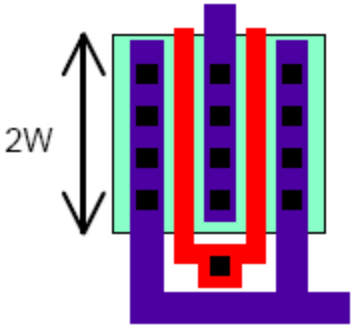


Multi-fingered gates save space and reduce series resistance in gate. Notice that drains are selected to minimize parasitic capacitance to bulk.

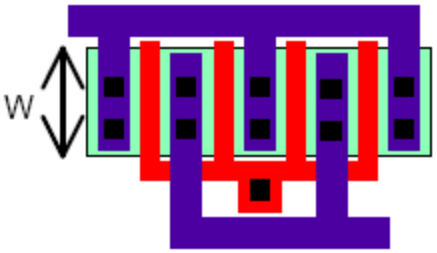
Multifinger transistors

For transistors that requires very large W/L ratio, a " folded " layout like figure (a) maybe not enough to reduce high gate resistance.

A layout like figure (b) is an improved version of figure (a) using multifinger transistor. The gate resistance is reduced by a factor of 4. While multifinger transistor reduces gate resistance, it raises source and drain capacitance, which introduces an trade-off.

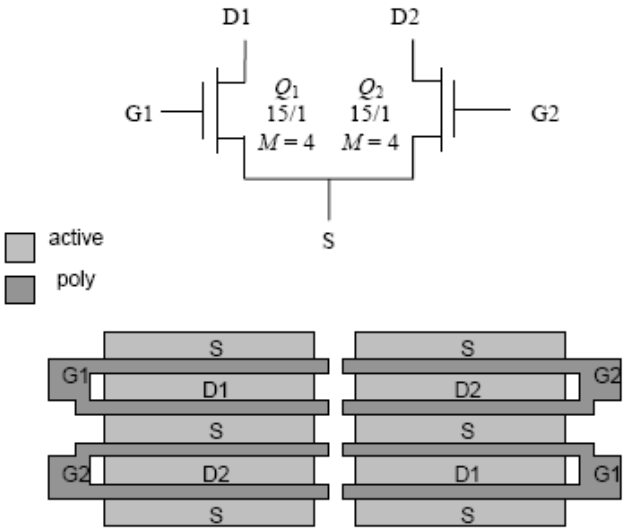


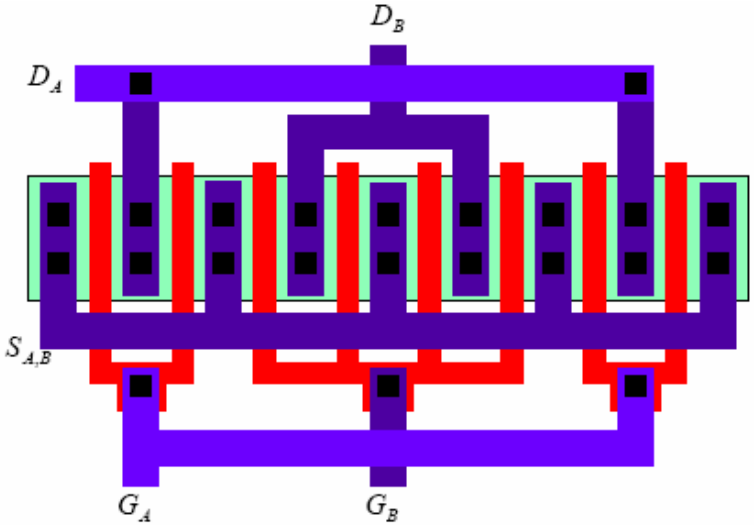
(a)



(b)

Common Centroid Layout with Multi-Fingered Gates





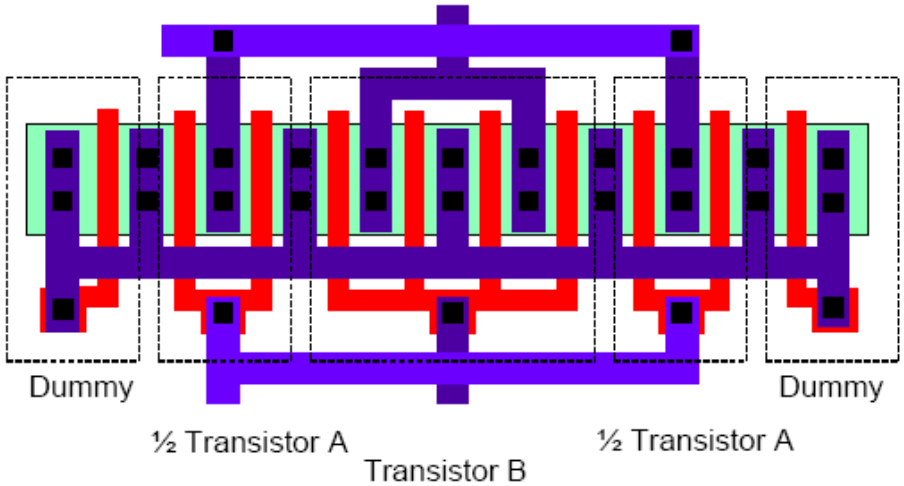
1D common centroid

Drain-to-bulk parasitic capacitance is minimized; sources are partially shared to save area.

2D common centroid

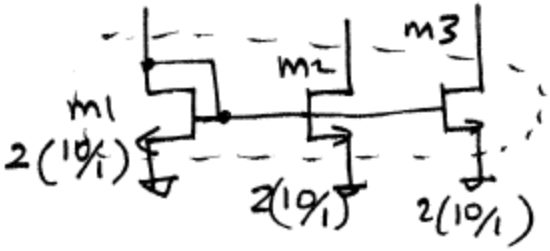
This is an improved version of the layout above. Two dummy transistors are added to the left and right sides. On the layout above, transistor A sees different ambient environment to the left and to the right. Dummy transistors improves the matching between transistor A and B by providing similar environment to the circuit that is on the boundary of a layout.

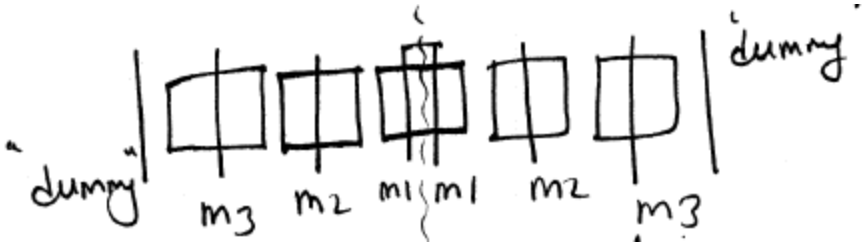
1D common centroid with dummies



2. ***Current mirrors***

matching is less critical than diff pair. If you have the area always use common centroid technique, if not use inter-digitated muli-finger devices





Matched parasitics

Not values (in case of gradient)

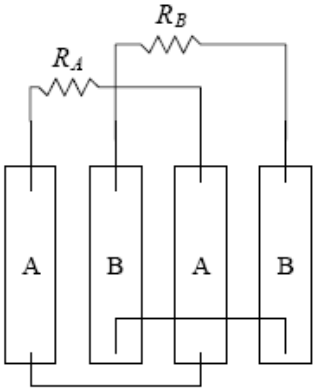
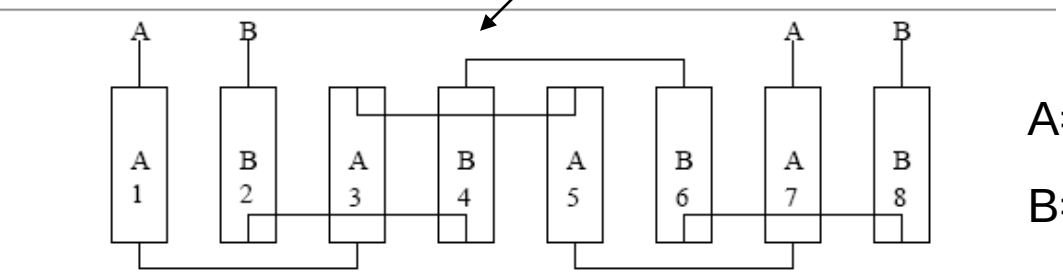


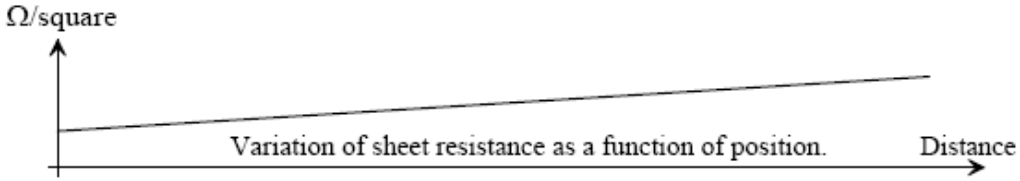
Figure 5.6 Layout of interdigitated resistors.

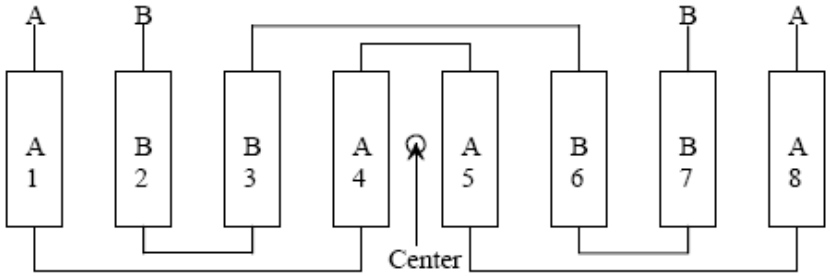


A=16

B=20

(a) Interdigitated layout





Matched values

Not parasitics (in case of gradient)

A=18

B=18

(b) Common-centroid layout

Figure 5.7 (a) Interdigitated layout and (b) common-centroid layout.

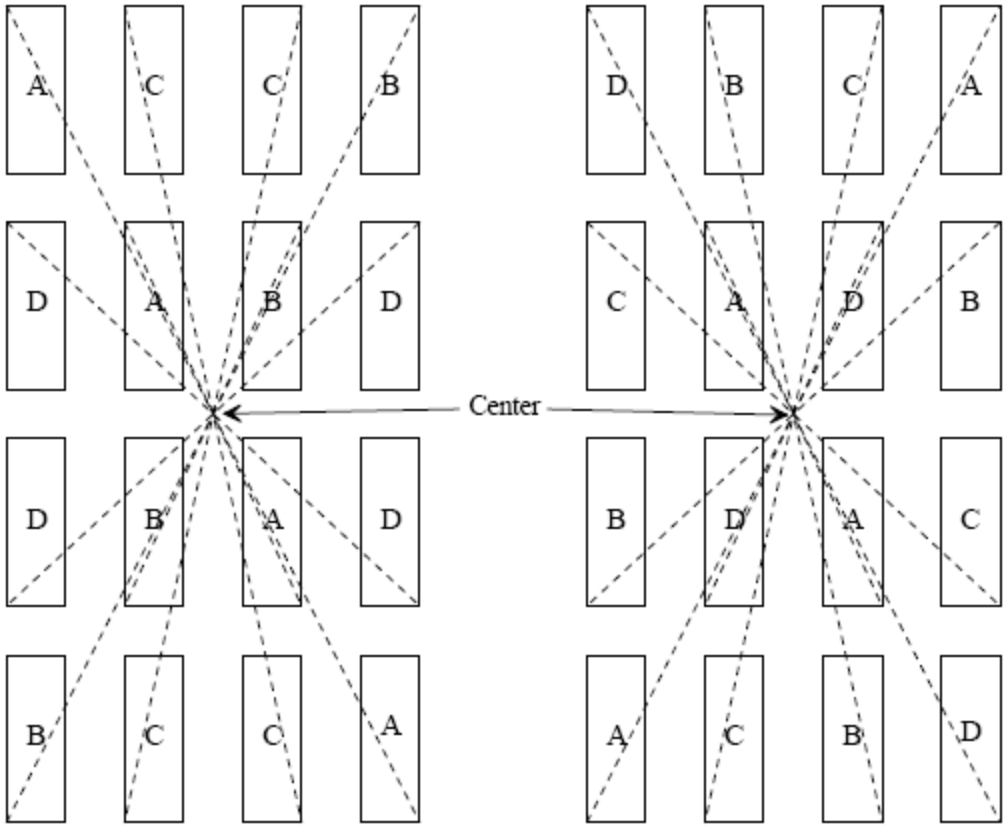
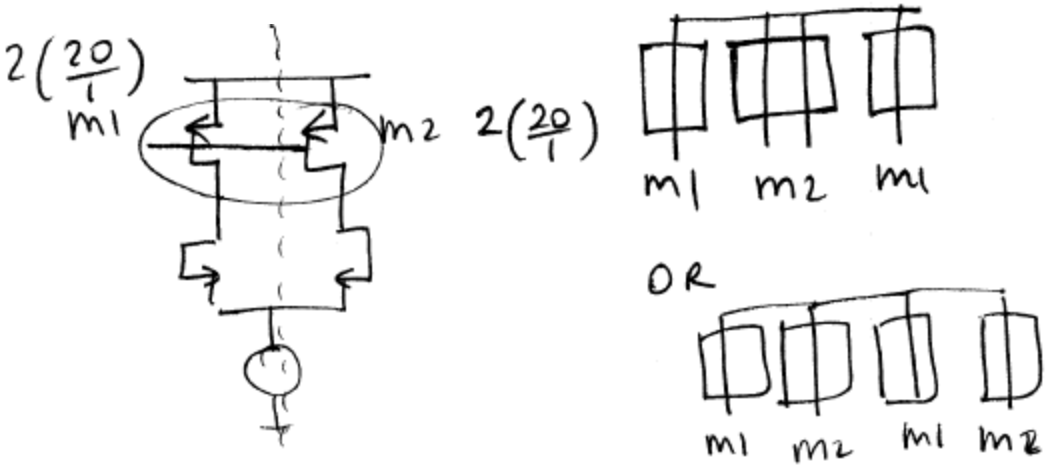


Figure 5.8 Common-centroid layout of four matched resistors (or elements).

Can be used for R's, C's, and transistors after dividing the elements into smaller parts

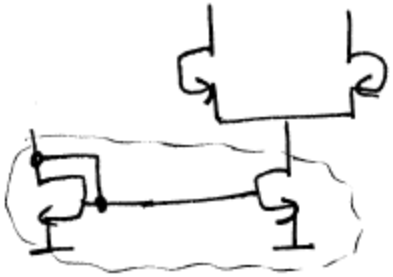
3. ***Load Devices***

If you have the area always use common centroid technique, if not use inter-digitated muli-finger devices



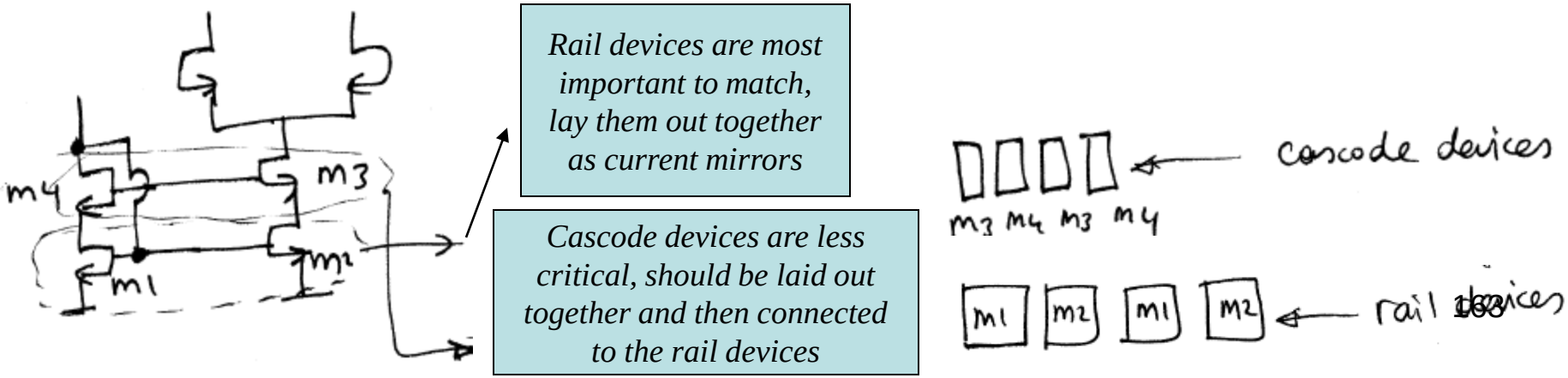
4. ***Tail current sources***

Same as current mirrors



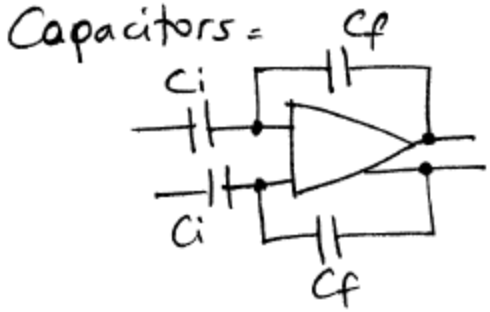
5. ***Cascode Stages***

Same as current mirrors



6. ***Capacitors (MIM cap, MOS cap, etc)***

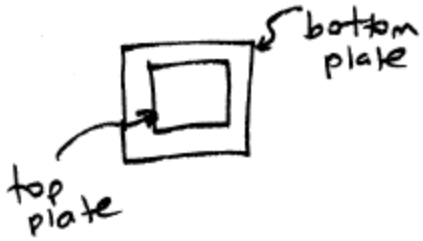
Divide into smaller units and use common centroid

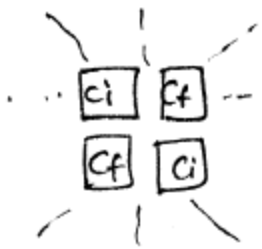


$C_i = 0.5pF$

$C_f = 0.5pF$

Divide C_i & C_f into much smaller caps. $C_{unit} \approx 50fF - 100fF$





General layout guidelines

1. *Try to Minimize (parasitics) and equalize (routing) in layout*
2. *Use one unit and use it repeatedly (Use the same W and L and vary M (takes out ΔW and ΔL effects)*
3. *Draw one half and mirror it*
4. *Use wide metals lines for low R, but this will have larger C*
5. *Slotting, stacking (strapping) to get high R without violating max W design rule*
6. *Use nwell shielding to minimize C to substrate for high speed signals*
7. *Dummification (avoid stacking to minimize parasitic capacitance to substrate)*
8. *Route currents a long way, not voltages - IR drops can cause big mismatches*