وَمَا أُوتِيتُمْ مِنَ الْعِلْمِ إِلَّا قَلِيلًا

Analog IC Design

Lecture 16
OTA Frequency Compensation

Dr. Hesham A. Omran

Integrated Circuits Laboratory (ICL)
Electronics and Communications Eng. Dept.
Faculty of Engineering
Ain Shams University

Outline

- Recapping previous key results
- Frequency compensation
- Compensation of single-stage OTAs
- Compensation of two-stage OTA
  - Miller compensation
  - The feedforward zero

Outline

- ❑ Recapping previous key results
- ❑ Frequency compensation
- ❑ Compensation of single-stage OTAs
- ❑ Compensation of two-stage OTA
  - Miller compensation
  - The feedforward zero

MOSFET in Saturation

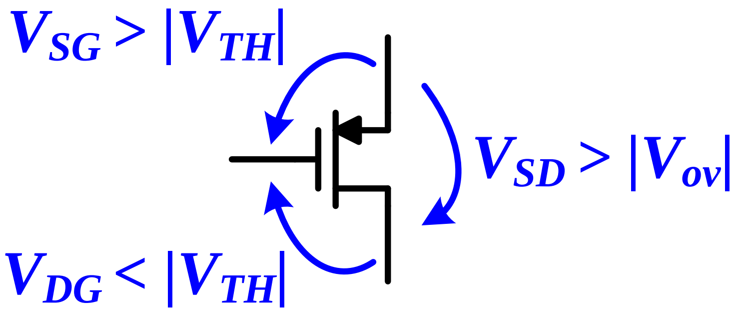
- The channel is pinched off if the difference between the gate and drain voltages is not sufficient to create an inversion layer

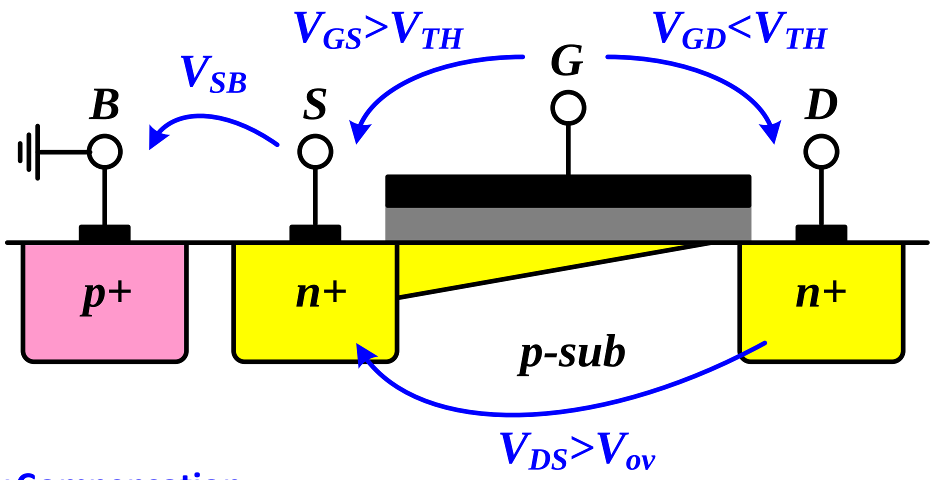
$$V_{GD} \leq V_{TH} \quad or \quad V_{DS} \geq V_{ov}$$

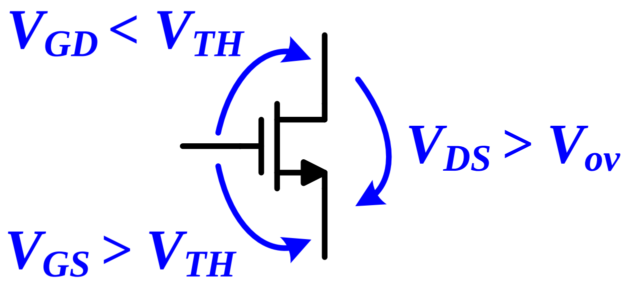
- Square-law (long channel MOS)

$$I_D = \frac{\mu_n C_{ox}}{2} \frac{W}{L} \cdot V_{ov}^2 (1 + \lambda V_{DS})$$

$$V_{SB} \uparrow \quad \Rightarrow \quad V_{TH} \uparrow$$







Regions of Operation Summary

OFF (Subthreshold)

$V_{GS} < V_{TH}$

ON

$V_{GS} > V_{TH}$

Triode

$V_{DS} < V_{ov}$

Or

$V_{GD} > V_{TH}$

$$I_D = \mu C_{ox}\frac{W}{L}\left(V_{ov}V_{DS} - \frac{V_{DS}^2}{2}\right)$$

Pinch-Off (Saturation)

$V_{DS} \geq V_{ov}$

Or

$V_{GD} \leq V_{TH}$

$$I_D = \frac{\mu C_{ox}}{2}\frac{W}{L}V_{ov}^2(1 + \lambda V_{DS})$$

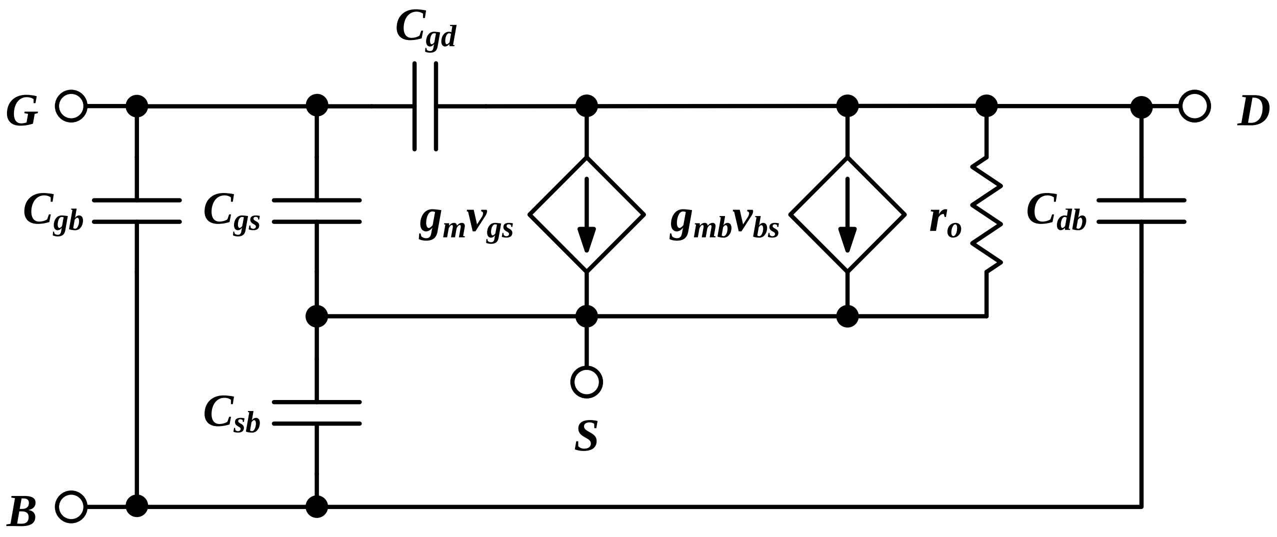
High Frequency Small Signal Model

$$g_m = \frac{\partial I_D}{\partial V_{GS}} = \mu C_{ox}\frac{W}{L}V_{ov} = \sqrt{\mu C_{ox}\frac{W}{L}\cdot 2I_D} = \frac{2I_D}{V_{ov}}$$

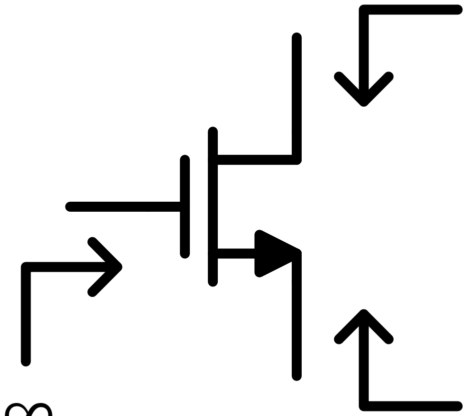
$$g_{mb} = \eta g_m \qquad \eta \approx 0.1 - 0.25$$

$$r_o = \frac{1}{\partial I_D/\partial V_{DS}} = \frac{V_A}{I_D} = \frac{1}{\lambda I_D} \qquad V_A \propto L \leftrightarrow \lambda \propto \frac{1}{L} \qquad V_{DS}\uparrow V_A\uparrow$$

$$C_{gb} \approx 0 \qquad C_{gs} \gg C_{gd} \qquad C_{sb} > C_{db}$$



Rin/out Shortcuts Summary



$$r_o[1 + (g_m + g_{mb})R_S]$$

H.I.N.

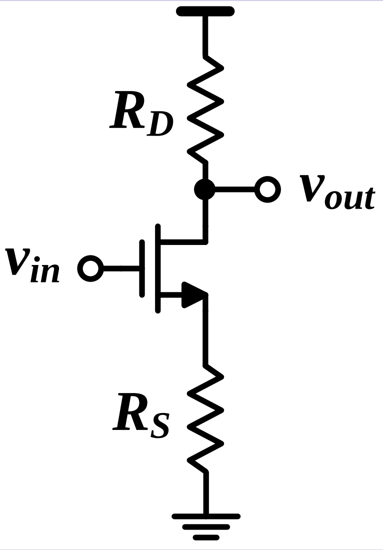
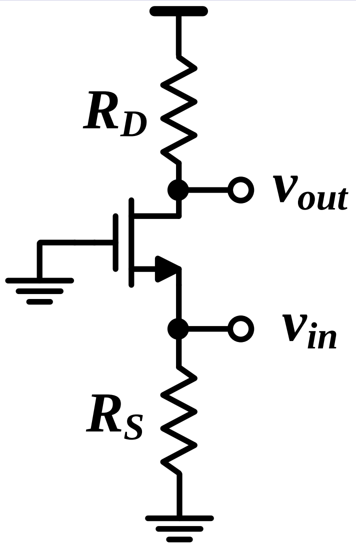
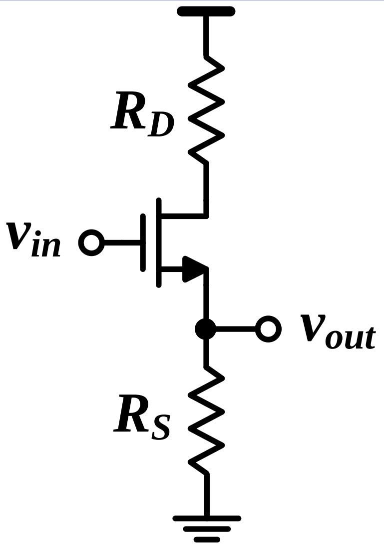
$$\frac{1}{g_m + g_{mb}}\left(1 + \frac{R_D}{r_o}\right)$$

L.I.N.

∞

At low frequencies **ONLY**

Summary of Basic Topologies

	CS	CG	CD (SF)
			
	Voltage & current amplifier	Voltage amplifier Current buffer	Voltage buffer Current amplifier
Rin	∞	$R_S \parallel \frac{1}{g_m + g_{mb}}\left(1 + \frac{R_D}{r_o}\right)$	∞
Rout	$R_D \parallel r_o[1 + (g_m + g_{mb})R_S]$	$R_D \parallel r_o$	$R_S \parallel \frac{1}{g_m + g_{mb}}\left(1 + \frac{R_D}{r_o}\right)$
Gm	$\frac{-g_m}{1 + (g_m + g_{mb})R_S}$	$g_m + g_{mb}$	$\frac{g_m}{1 + R_D/r_o}$

Differential Amplifier

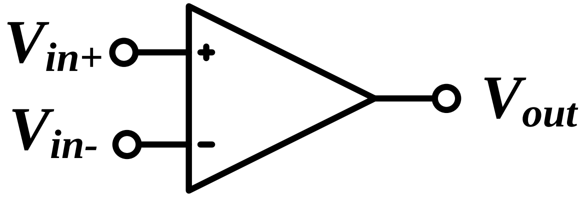
	Pseudo Diff Amp	Diff Pair (w/ ideal CS)	Diff Pair (w/ R_{SS})
A_{vd}	$-g_m R_D$	$-g_m R_D$	$-g_m R_D$
A_{vCM}	$-g_m R_D$	0	$\frac{-g_m R_D}{1+2(g_m+g_{mb})R_{SS}}$
A_{vd}/A_{vCM}	1	∞	$2(g_m+g_{mb})R_{SS} \gg 1$

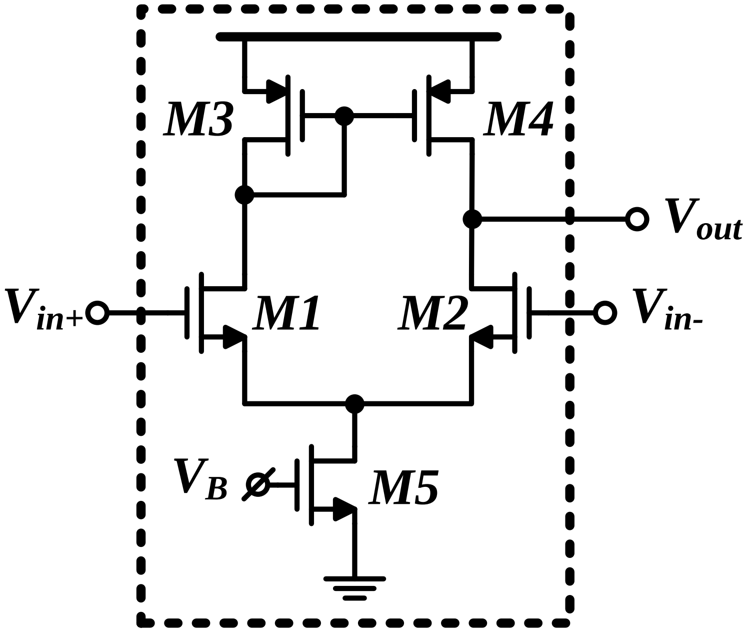
$$A_{vCM2d} = \frac{v_{od}}{v_{iCM}} \approx \frac{\Delta R_D}{2R_{SS}} + \frac{\Delta g_m R_D}{2g_{m1,2}R_{SS}}$$

$$CMRR = \frac{A_{vd}}{A_{vCM2d}}$$

Op-Amp

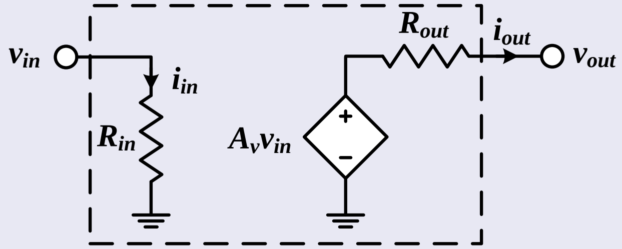
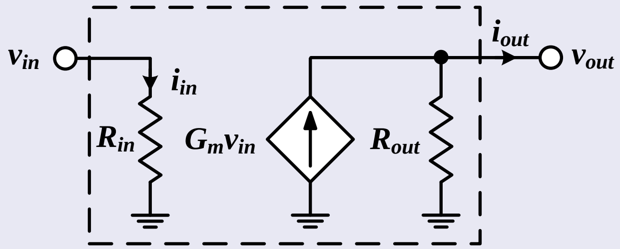
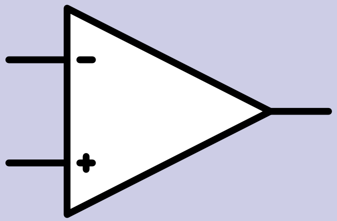
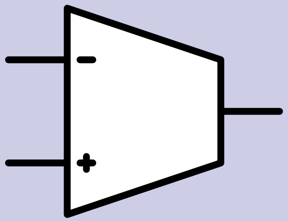
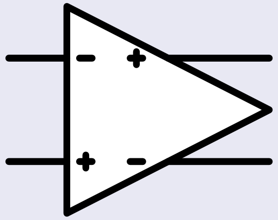
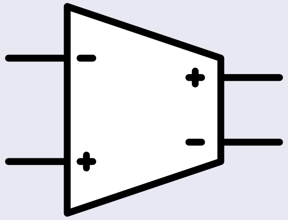
- An op-amp is simply a high gain differential amplifier
  - The gain can be increased by using cascodes and multi-stage amplification
- The diff amp is a key block in many analog and RF circuits
  - DEEP understanding of diff amp is ESSENTIAL





Op-Amp vs OTA

- In short, an OTA is an op-amp without an output stage (buffer)
- Some designers just use op-amp name and symbol for both

	Op-amp	OTA
Rout	**LOW**	**HIGH**
Model		
Diff input, SE output		
Fully diff		

V-star (V^*)

- V-star (V^*) is inspired by V_{ov} but calculated from actual simulation data

$$g_m = \frac{2I_D}{V^*} \leftrightarrow V^* = \frac{2I_D}{g_m} = \frac{2}{g_m/I_D}$$

- Figures-of-merit in terms of V^*

$$g_m r_o = \frac{2I_D}{V^*} \cdot \frac{1}{\lambda I_D} = \frac{2}{\lambda V^*}$$

$$f_T = \frac{g_m}{2\pi C_{gg}} = \frac{1}{2\pi} \cdot \frac{2I_D}{V^*} \cdot \frac{1}{C_{gg}}$$

$$\frac{g_m}{I_D} = \frac{2}{V^*}$$

- The boundary between weak and strong inversion ($n = 1.2 \rightarrow 1.5$)

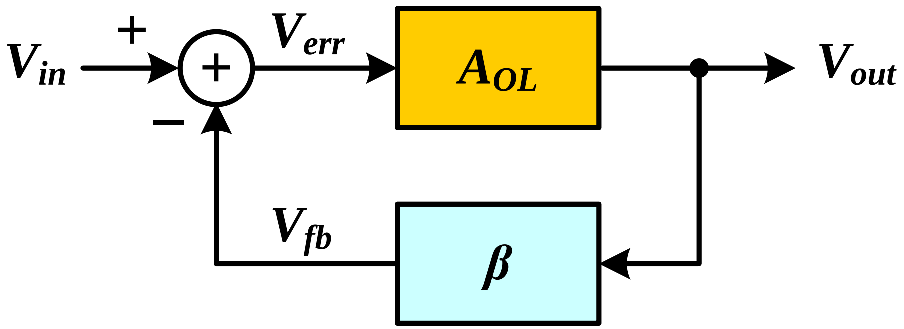
$$V_{ov}(SI) = V^*(WI) = 2nV_T \approx 60 \rightarrow 80mV$$

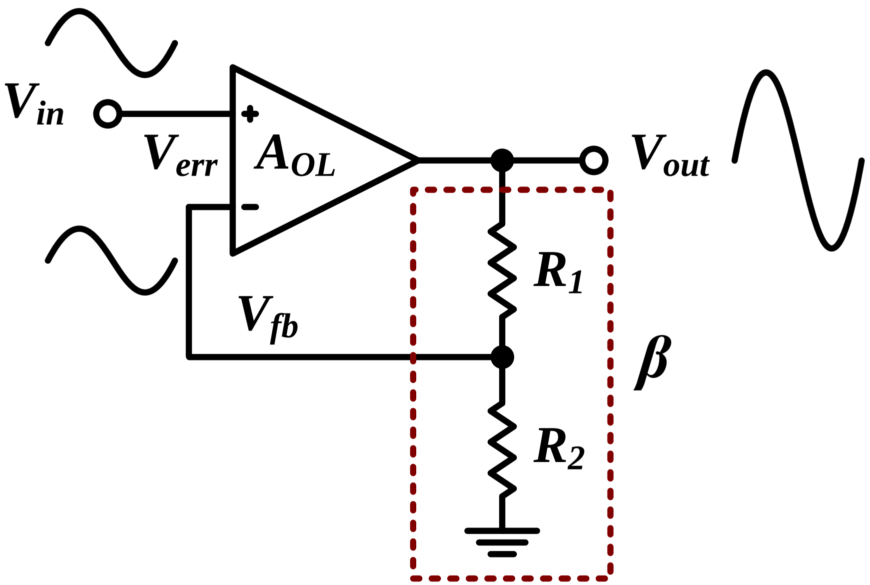
Negative Feedback

$$\beta = \frac{R_2}{(R_1 + R_2)}$$

$$A_{CL} = \frac{V_{out}}{V_{in}} = \frac{A_{OL}}{1 + \beta A_{OL}} = \frac{A_{OL}}{1 + \beta A_{OL}} \approx \frac{1}{\beta} = \frac{R_1 + R_2}{R_2} = 1 + \frac{R_1}{R_2}$$

$$\omega_{p,CL} = (1 + \beta A_{OLo})\omega_{P,OL}$$





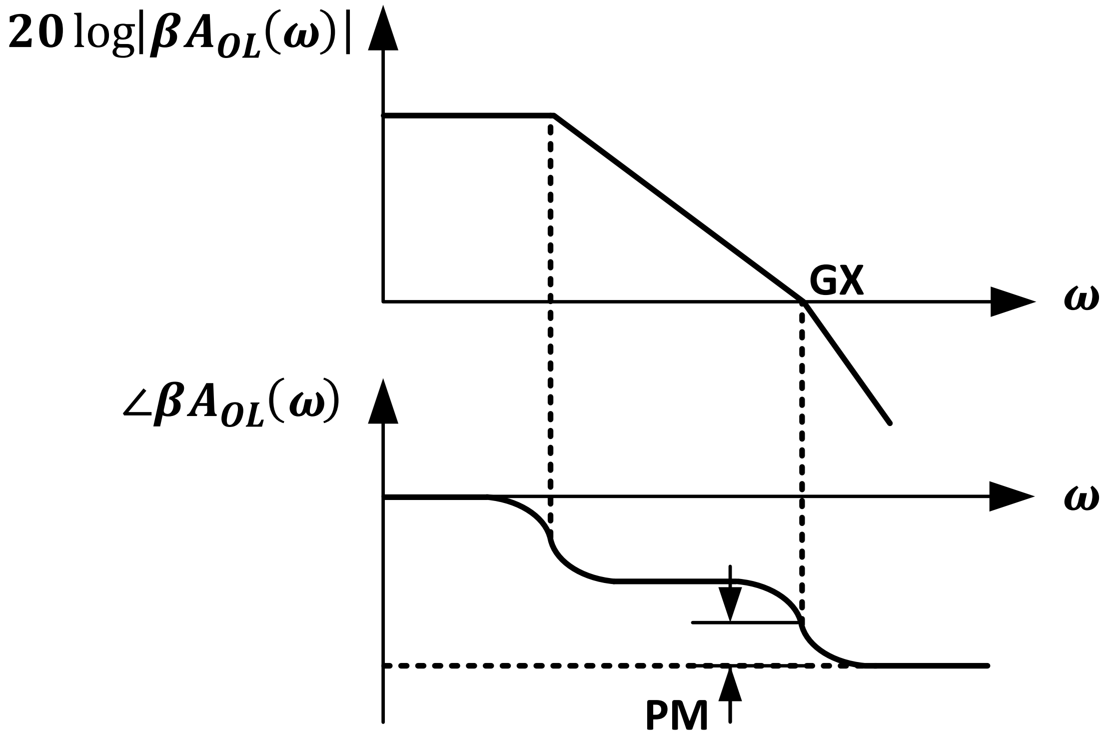
Outline

- ❑ Recapping previous key results
- ❑ Frequency compensation
- ❑ Compensation of single-stage OTAs
- ❑ Compensation of two-stage OTA
  - Miller compensation
  - The feedforward zero

Phase Margin and the Ultimate GBW

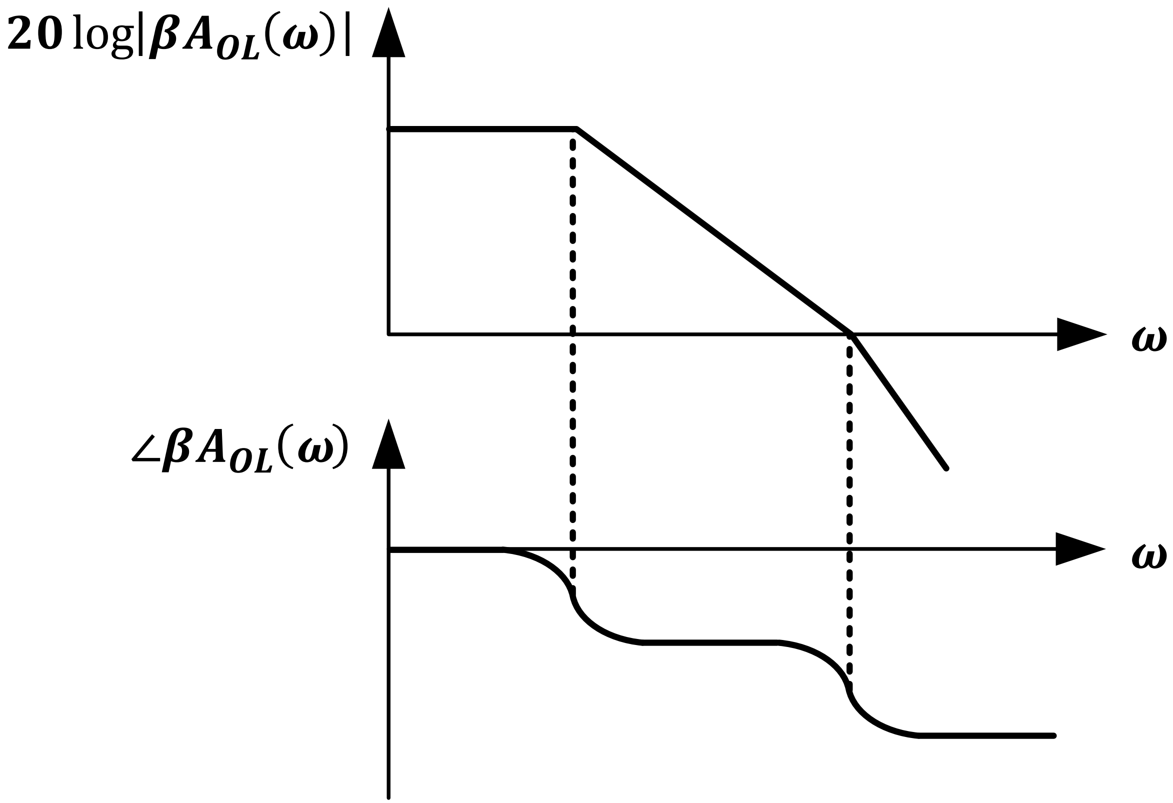
- If $\omega_{p2} = \omega_u$: PM = 45°
  - Typically inadequate (peaking/ringing)
- Thus ω_{p2} should be $> \omega_u$ → $\omega_{p1} \ll \omega_u < \omega_{p2}$
  - ω_{p1} defines OL BW and ω_{p2} defines ultimate GBW (max CL BW)

Frequency domain peaking
→ noise amplification
Time domain ringing
→ poor settling time



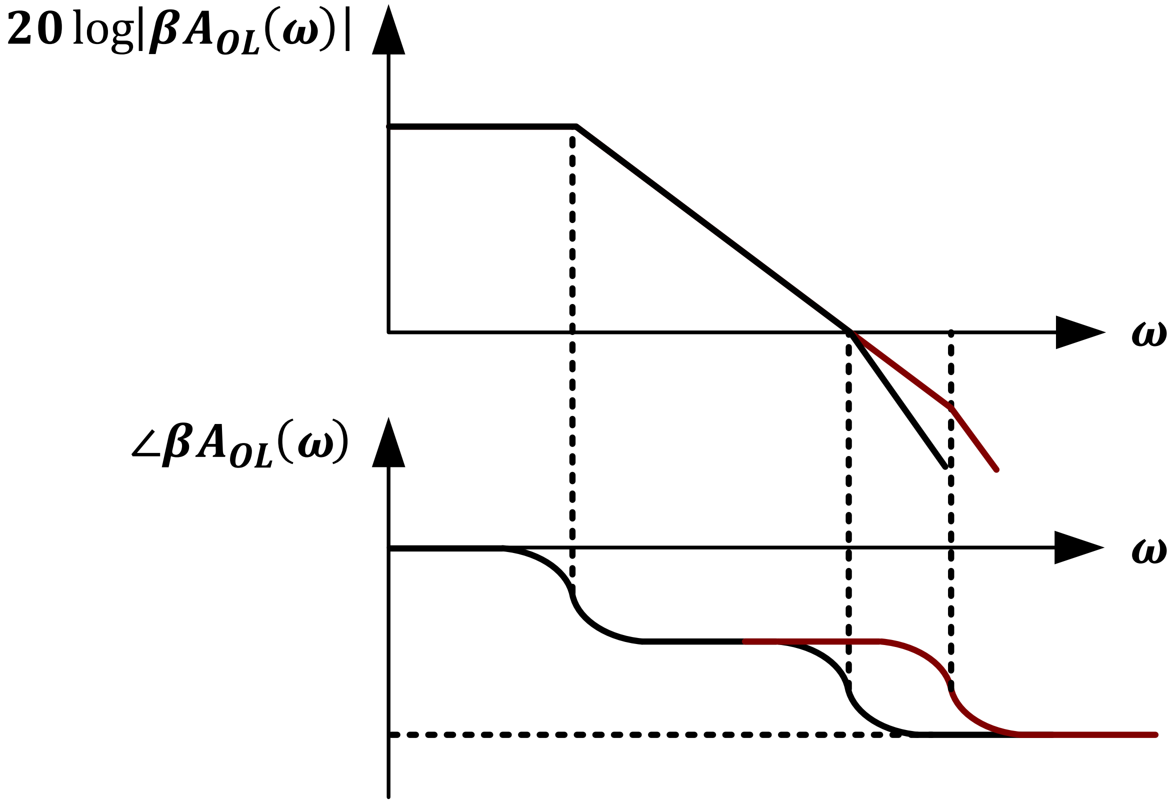
Frequency Compensation

- Frequency compensation: Modify the system to achieve a specific PM to control frequency domain peaking and time domain ringing
- GX < PX is not enough: We actually need $\omega_u < \omega_{p2}$



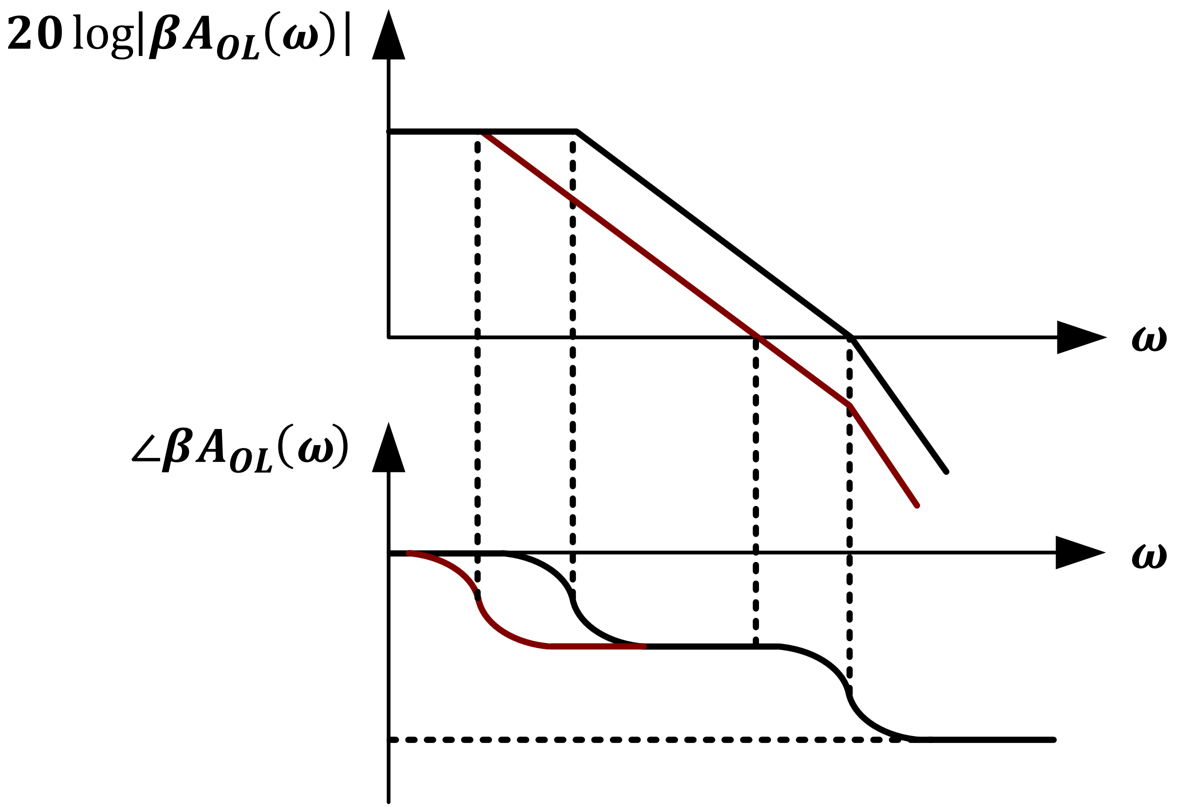
Frequency Compensation

- We need GX < PX (actually $\omega_u < \omega_{p2}$)
  - Push ω_{p2} **outwards**: lower resistance/capacitance
    - Not always feasible for free



Frequency Compensation

- We need GX < PX (actually $\omega_u < \omega_{p2}$)
  - Push GX / ω_u / ω_{p1} **inwards**: lower GBW



Compensation of Popular OTA Topologies

- Single-stage OTAs
  - 5T OTA
  - Telescopic cascode OTA
  - Folded cascode OTA
- Two-stage OTA
- Three-stage OTA
- Gain boosted OTA

Outline

- Recapping previous key results
- Frequency compensation
- Compensation of single-stage OTAs
- Compensation of two-stage OTA
    - Miller compensation
    - The feedforward zero

5T OTA

- $\omega_{p1} \ll \omega_u < \omega_{p2}$
- The H.I.N. sets BW_{OL}

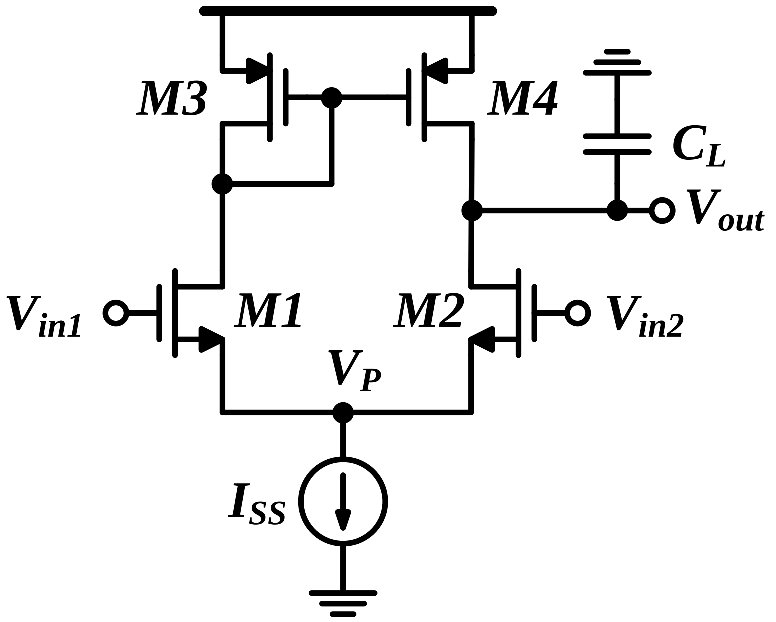
$$\omega_{p1} = \omega_{pout} \approx \frac{1}{R_{out}C_{out}}$$

- The first non-dominant pole (mirror node) sets the ultimate GBW
  - Ultimate CL bandwidth (buffer)

$$\omega_{p2} = \omega_{pM} \approx \frac{g_{m3}}{C_M} \sim \frac{\omega_T}{3}$$

- LHP zero:

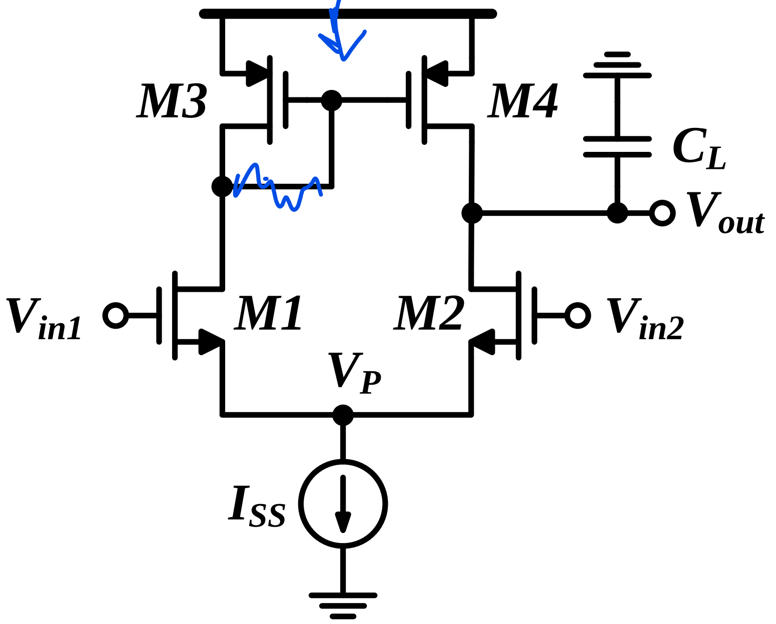
$$\omega_z = 2\omega_{p2}$$



5T OTA

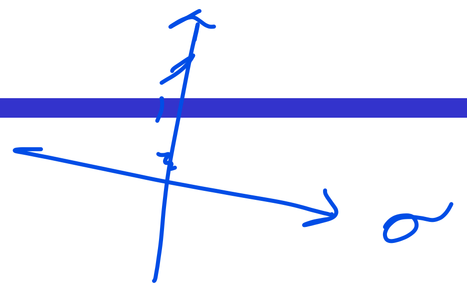
- ❑ Fully differential?

floating voltage,



5T OTA

s-plane



- Fully differential with mismatch?

$$H(s) = \frac{A_o/2}{1+\frac{s}{\omega_{p1}}} + \frac{A_o/2}{1+\frac{s}{\omega_{p1}+\Delta\omega}}$$

multiple paths

two poles

one zero

↓

contributes

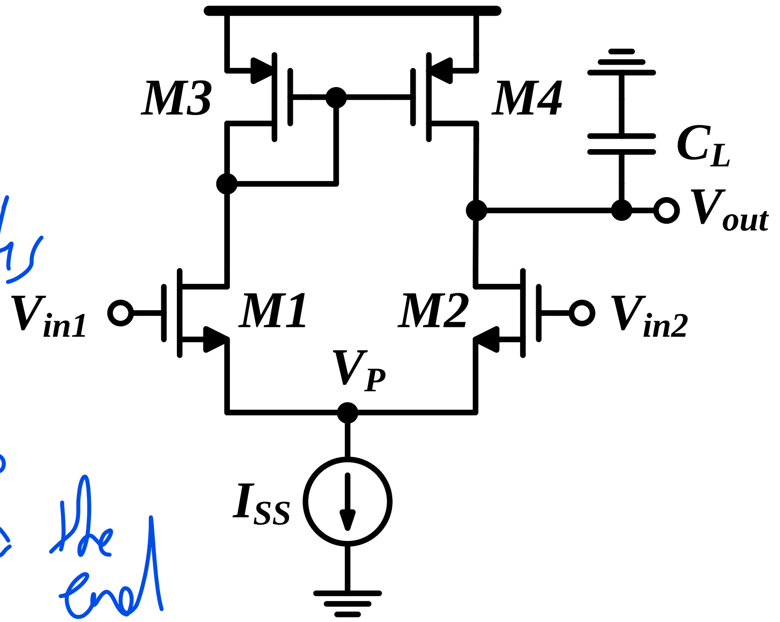
to

one pole

@ the end

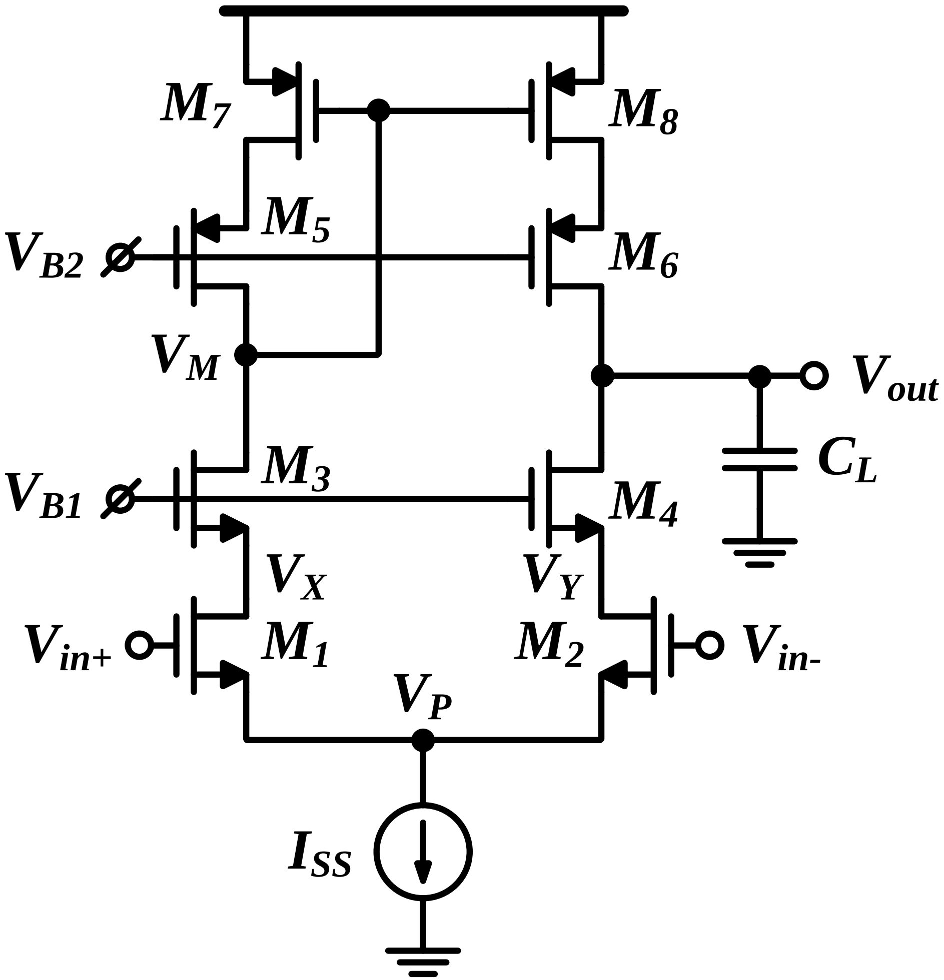
$$= \frac{A_o\left\{1+\frac{s}{2}\left(\frac{1}{\omega_{p1}}+\frac{1}{\omega_{p1}+\Delta\omega}\right)\right\}}{\left(1+\frac{s}{\omega_{p1}}\right)\left(1+\frac{s}{\omega_{p1}+\Delta\omega}\right)}$$

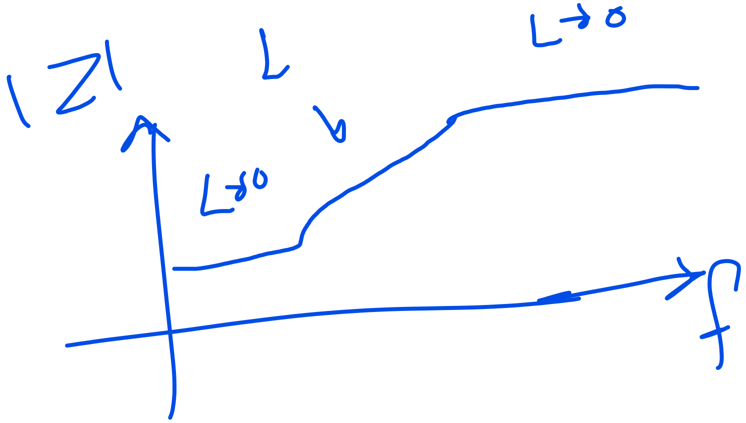
$$= \frac{A_o\left(1+\frac{s}{\omega_z}\right)}{\left(1+\frac{s}{\omega_{p1}}\right)\left(1+\frac{s}{\omega_{p1}+\Delta\omega}\right)}$$



Telescopic Cascode

- Higher DC gain, but limited swing and additional poles





$Z_L = j\omega L = Z_x$

$L = \frac{Z_x}{j\omega}$

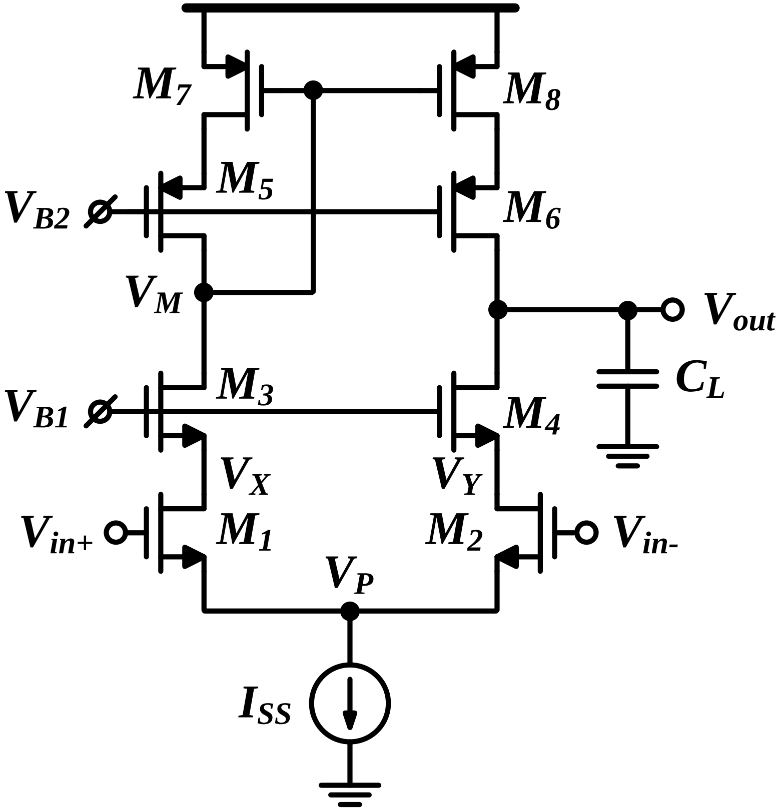
Telescopic Cascode: Poles

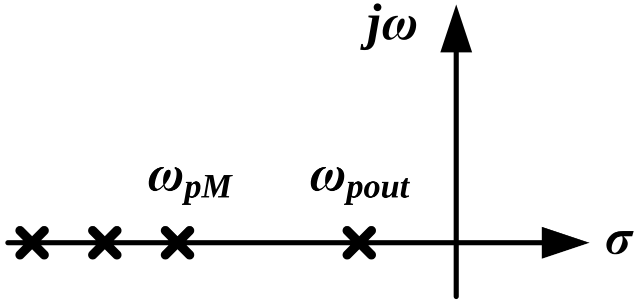
- $\omega_{p1} \ll \omega_u < \omega_{p2}$
- The H.I.N. sets BW_{OL}

$$\omega_{p1} = \omega_{pout} = \frac{1}{R_{out} C_{out}}$$

- V_X and V_Y contribute a single pole
- Non-dominant poles considerations:
  - C_{gs} is larger than other caps
  - PMOS contributes larger capacitances (lower ID/W)
- The first non-dominant pole (mirror node) sets the ultimate GBW

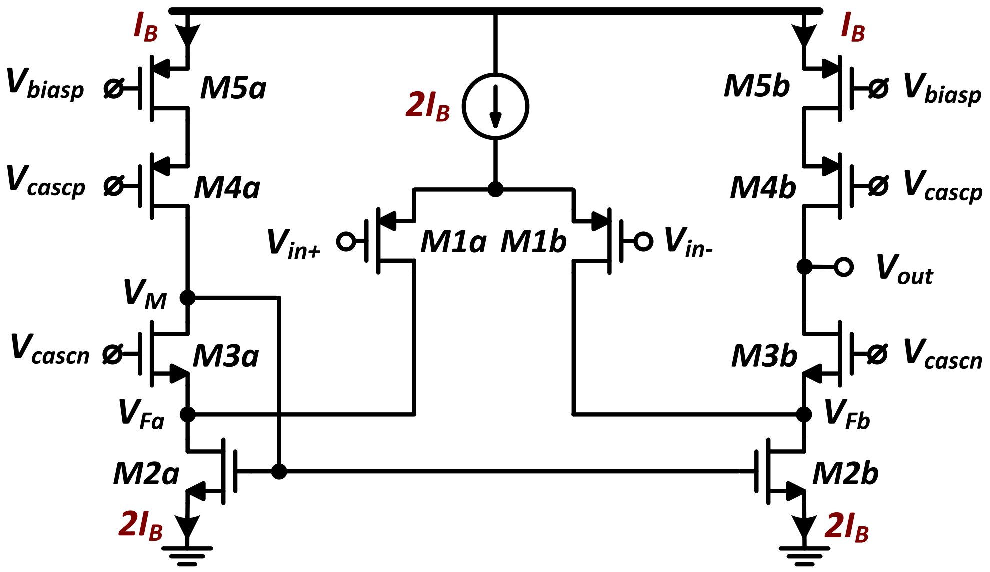
$$\omega_{p2} = \omega_{pM} \approx \frac{g_{m7}}{C_M} \sim \frac{\omega_T}{4}$$

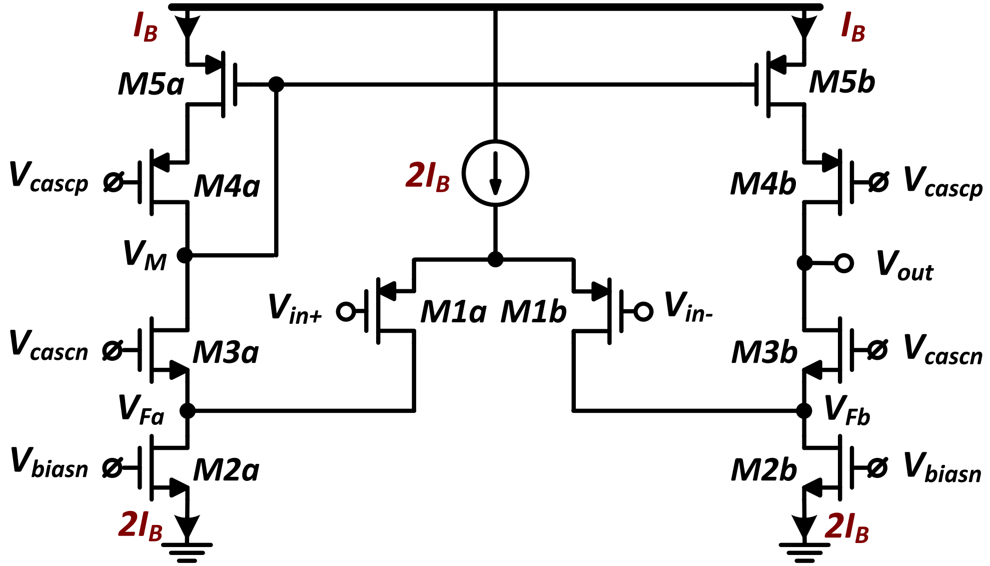




Folded Cascode

- Two possible implementations for SE output
- Compared to telescopic cascode:
  - More power (~2x)
  - Lower gain ($r_{o1}||r_{o2}$)
  - More nodes/poles
  - More complex
  - But input and output ranges decoupled





Folded Cascode: Poles

- ❑ $\omega_{p1} \ll \omega_u < \omega_{p2}$
- ❑ The H.I.N. sets BW_{OL}

$$\omega_{p1} = \omega_{pout} = \frac{1}{R_{out}C_{out}}$$

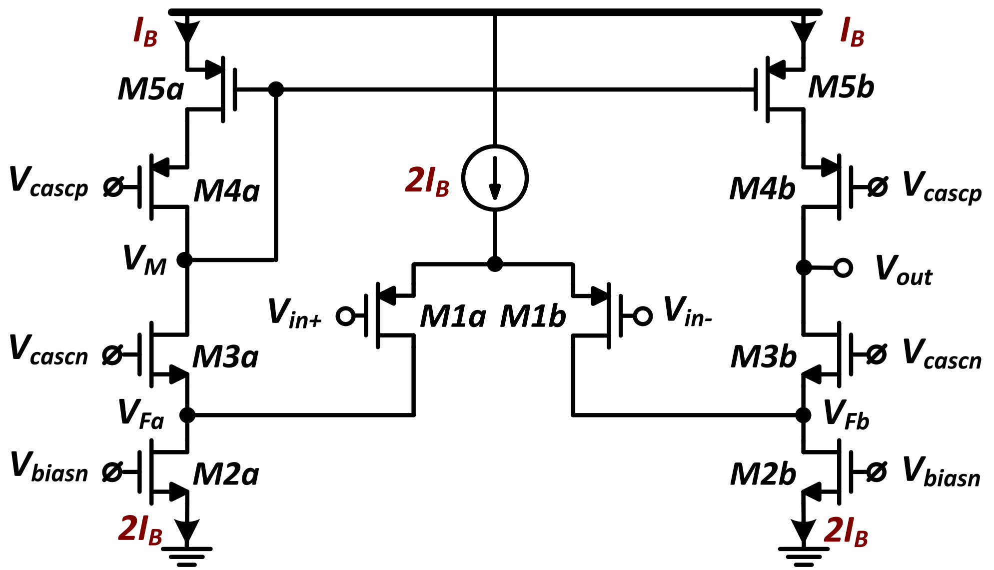
- ❑ The first non-dominant pole sets the ultimate GBW
  - Mirror node (V_M)

$$\omega_{p2} = \omega_{pM} \approx \frac{g_{m5}}{C_M}$$

  - **Or** folding node (V_F)

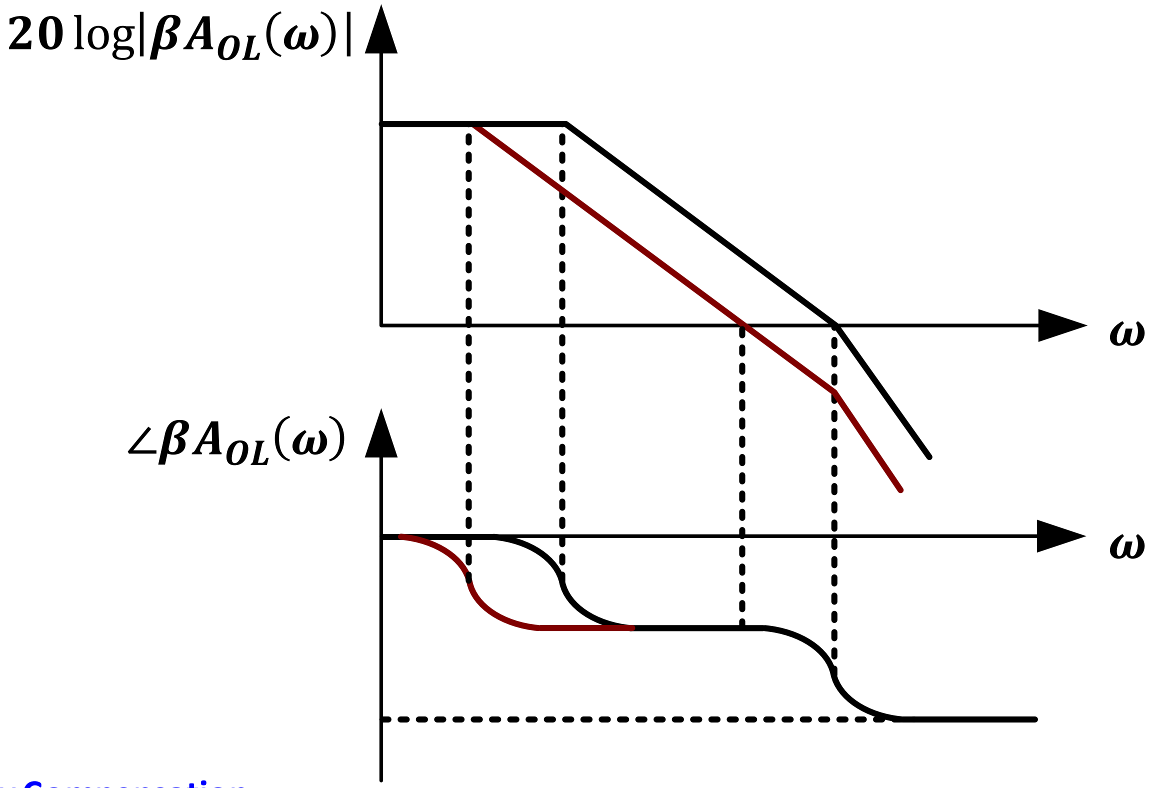
$$\omega_{p2} = \omega_{pF} \approx \frac{g_{m3}}{C_F}$$

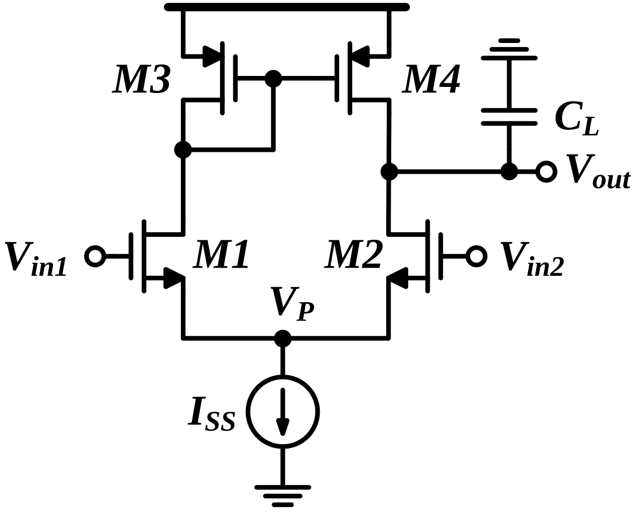
- ❑ V_{Fa} and V_{Fb} contribute a single pole
- ❑ Considerations:
  - C_{gs} larger than other caps
  - PMOS larger than NMOS
  - M2 has large capacitance (double the current)



Single-Stage OTAs: Compensation

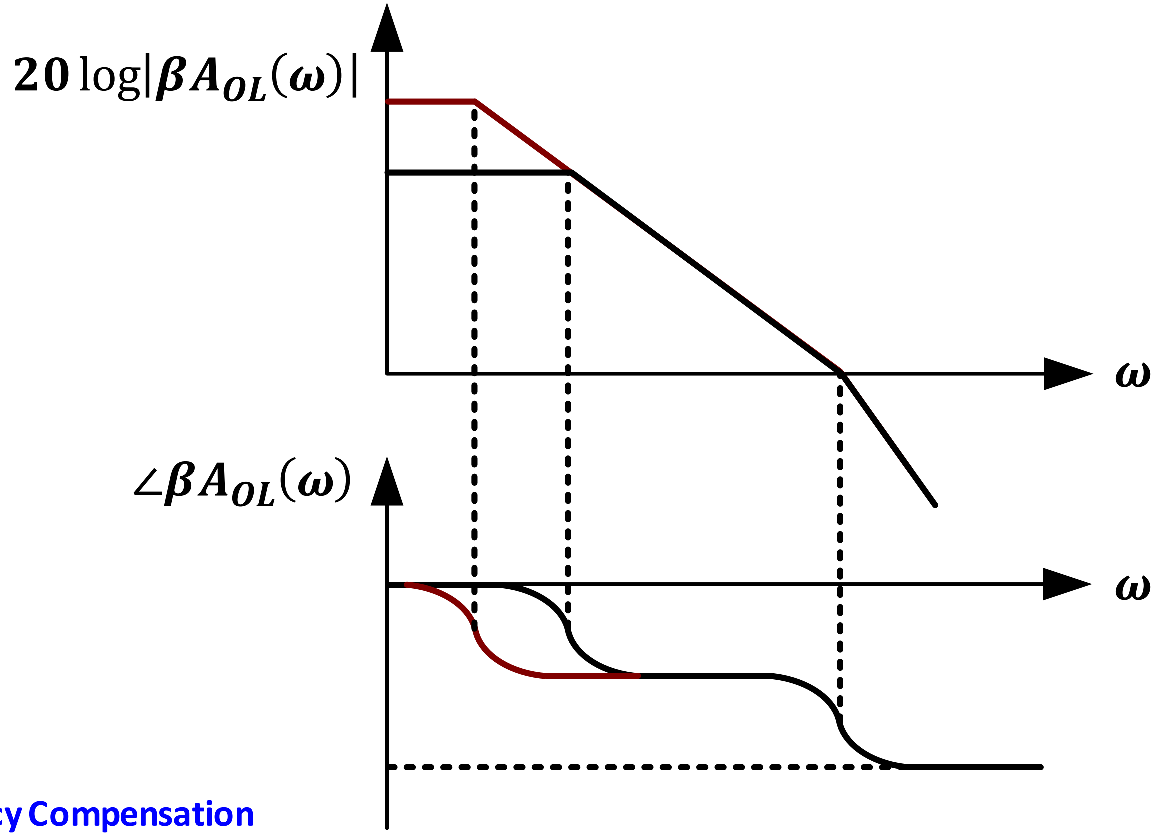
- Push GX / ω_u / ω_{p1} **inwards**: lower GBW
  - Increase C_L
  - **Single-stage OTAs are compensated by large load capacitance**

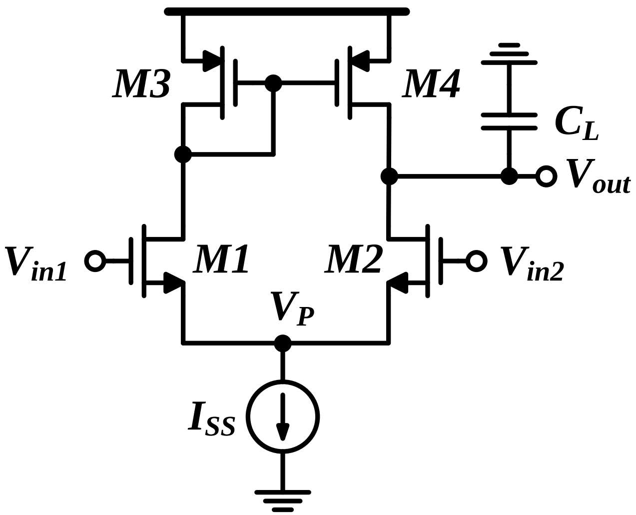




Single-Stage OTAs: Compensation

- Push GX / ω_u / ω_{p1} **inwards**: lower GBW
  - Increase C_L
  - **Single-stage OTAs are compensated by large load capacitance**
- Increasing R_{out} does not affect PM



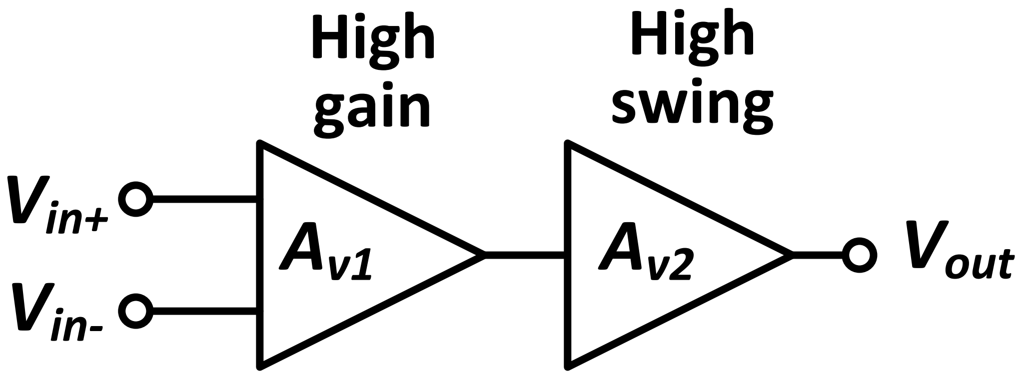


Outline

- ❑ Recapping previous key results
- ❑ Frequency compensation
- ❑ Compensation of single-stage OTAs
- ❑ Compensation of two-stage OTA
  - Miller compensation
  - The feedforward zero

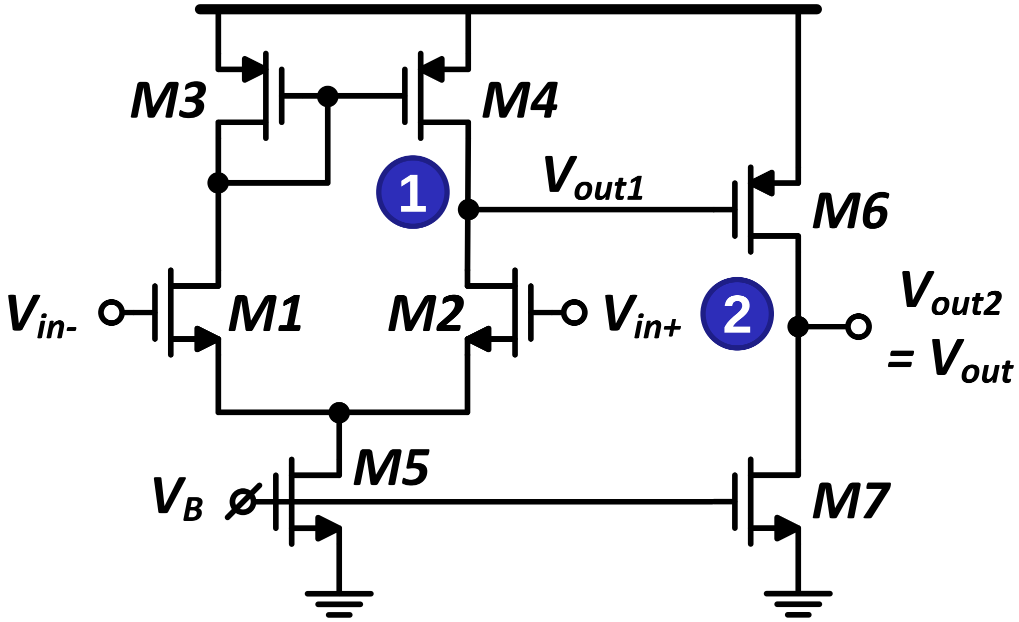
Two-Stage OTA

- Isolates the gain and swing requirements
  - But much more power consumption
  - And complicates stability requirements
- Three-stage OTA exists, but quite difficult to stabilize
- First stage can be 5T-OTA or cascode
- Second stage is typically a simple common-source
  - Allows maximum output swing



Two-Stage OTA: Poles

- Two gain stages → Two H.I.N.s
  - Two dominant poles!
- Internal pole: $\omega_{p1} = \frac{1}{R_{out1}C_1}$
- Output pole: $\omega_{p2} = \frac{1}{R_{out2}C_2}$

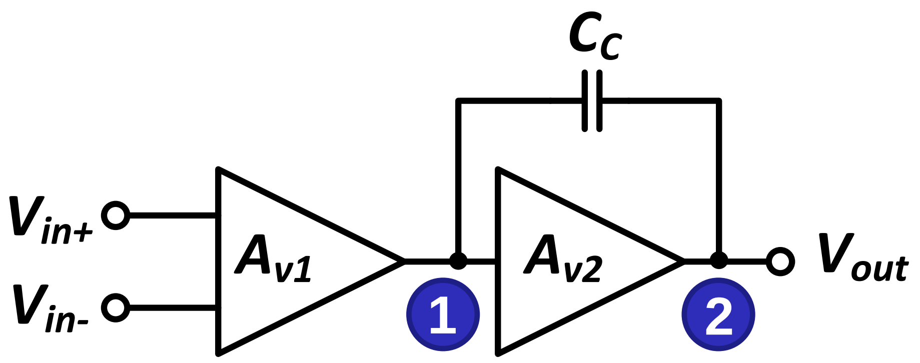


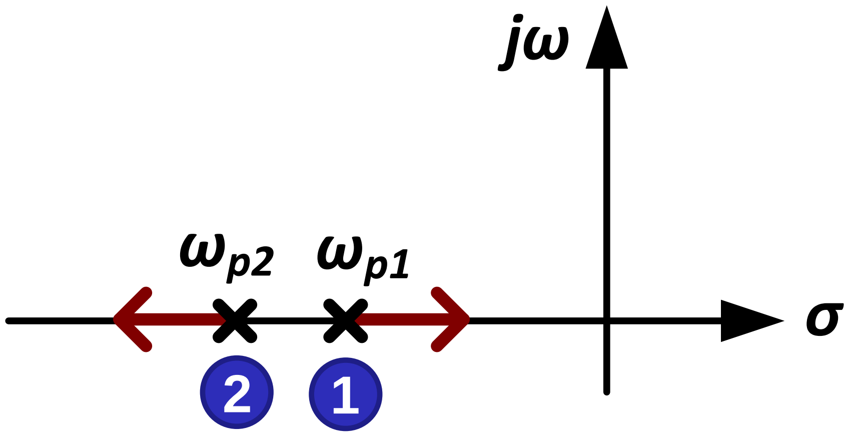
Outline

- Recapping previous key results
- Frequency compensation
- Compensation of single-stage OTAs
- Compensation of two-stage OTA
  - Miller compensation
  - The feedforward zero

Two-Stage OTA: Miller Compensation

- Exploit Miller capacitance multiplication
- Pole splitting
  - Push ω_{p1} inwards → push GX inwards
  - Push ω_{p2} outwards → push PX outwards





CS HFR: Reminder

- Surprisingly, exact analysis gives a quite complex expression
  - See [Johns & Martin 2012] or [Razavi 2017]
- If dominant pole approximation is applied

$$\omega_{pd} \approx \frac{1}{b_1} = \frac{1}{R'_{sig}\left[C_{gs} + C_{gd}(1 + g_m R_{out})\right] + R_{out}\left(C_{out} + C_{gd}\right)}$$

  - Same result as OCTC (both based on same approximation)
- Additionally, dominant pole approx gives an expression for ω_{pnd}

$$\omega_{pnd} \approx \frac{1}{b_2 \omega_{p1}} = \frac{b_1}{b_2} = \frac{g_m C_{gd}}{C_{gd}\left(C_{gs} + C_{out}\right) + C_{gs} C_{out}}$$

  - If a large cap is connected parallel to C_{gd}: $\omega_{pnd} \approx \frac{g_m}{C_{gs}+C_{out}}$
    - Can be derived intuitively without analysis (how?)

Two-Stage OTA: Miller Compensation

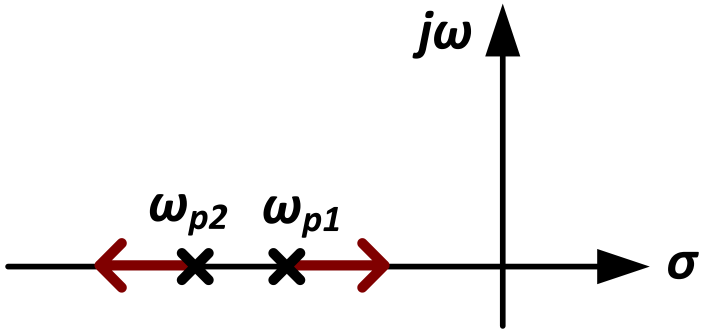
- Before compensation

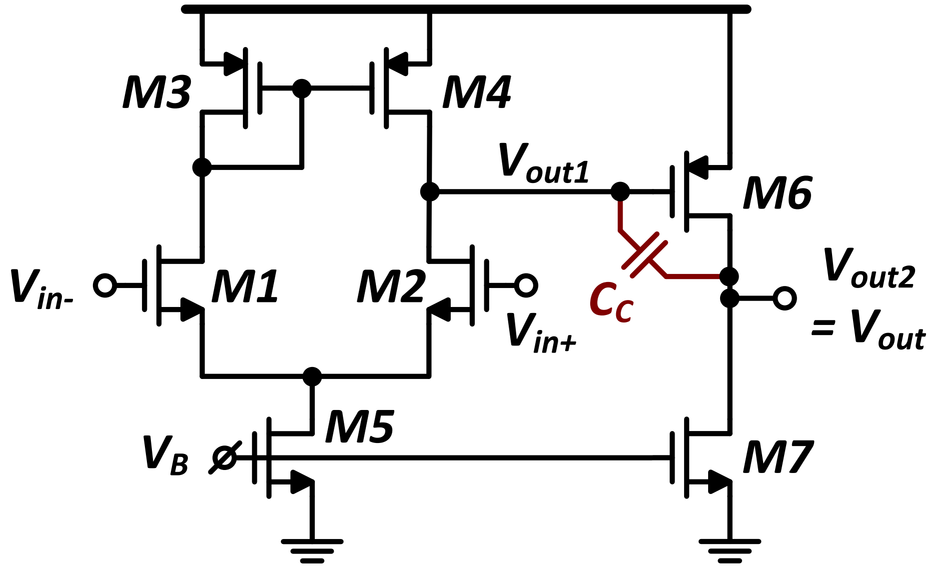
$$\omega_{p1} = \frac{1}{R_{out1}C_1} \quad \& \quad \omega_{p2} = \frac{1}{R_{out2}C_2}$$

- After compensation: Pole splitting

$$\omega_{p1} \approx \frac{1}{R_{out1}[(G_{m2}R_{out2})C_C + C_1] + R_{out2}(C_2 + C_C)} \approx \frac{1}{R_{out1}(G_{m2}R_{out2})C_C}$$

$$\omega_{p2} \approx \frac{G_{m2}C_C}{C_C(C_1 + C_2) + C_1C_2} \approx \frac{G_{m2}}{C_1 + C_2}$$





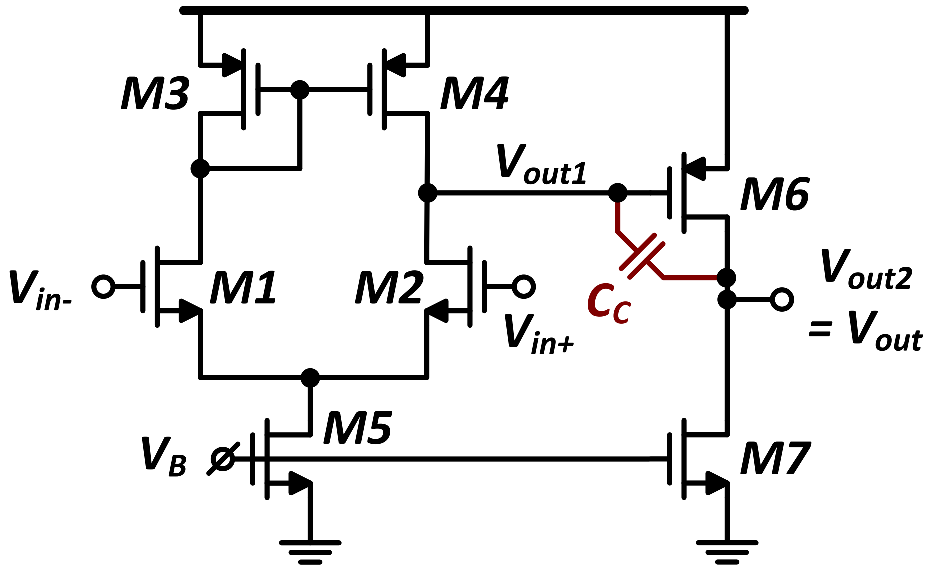
Stability Requirement: $\boldsymbol{\omega_u}$ and $\boldsymbol{\omega_{p2}}$

$$\omega_{p1} \approx \frac{1}{R_{out1}(G_{m2}R_{out2})C_C}$$

$$\omega_{p2} \approx \frac{G_{m2}}{C_1 + C_2} \approx \frac{G_{m2}}{C_L}$$

$$GBW \approx G_{m1}R_{out1}G_{m2}R_{out2} \cdot \frac{1}{R_{out1}(G_{m2}R_{out2})C_C}$$

$$GBW = \omega_u \approx \frac{G_{m1}}{C_C}$$



- **For critical damped response: $\boldsymbol{\zeta = 1}$, $\boldsymbol{Q = 0.5}$, and $\boldsymbol{PM \approx 76^o}$**

$$\omega_{p2} \approx 4\omega_u \rightarrow \frac{G_{m2}}{C_L} \approx \frac{4G_{m1}}{C_C}$$

Bias Current in 1st and 2nd Stages

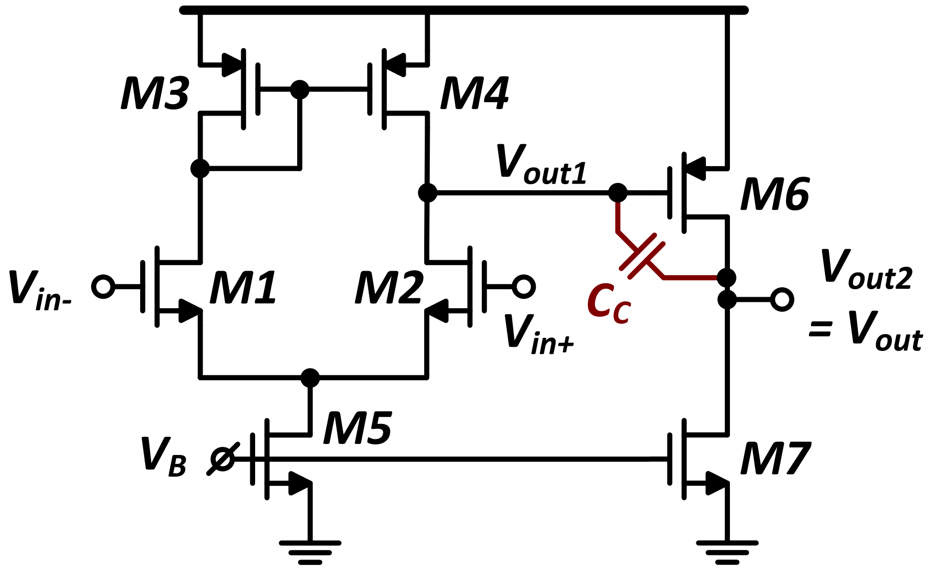
$$\frac{G_{m2}}{C_L} \approx \frac{4G_{m1}}{C_C}$$

- Assume $C_L = 4pF \; and \; C_C = C_L/2 = 2pF$

$$\frac{G_{m2}}{G_{m1}} = 4 \times \frac{4}{2} = 8$$

- If both stages use the same gm/ID
  - Note that $I_{B1} = 2 \times I_{D1,2}$

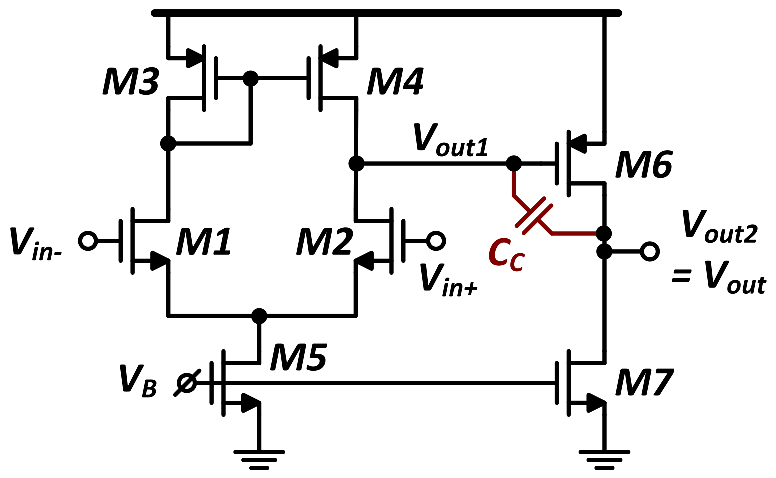
$$\frac{I_{B2}}{I_{B1}} = 4$$

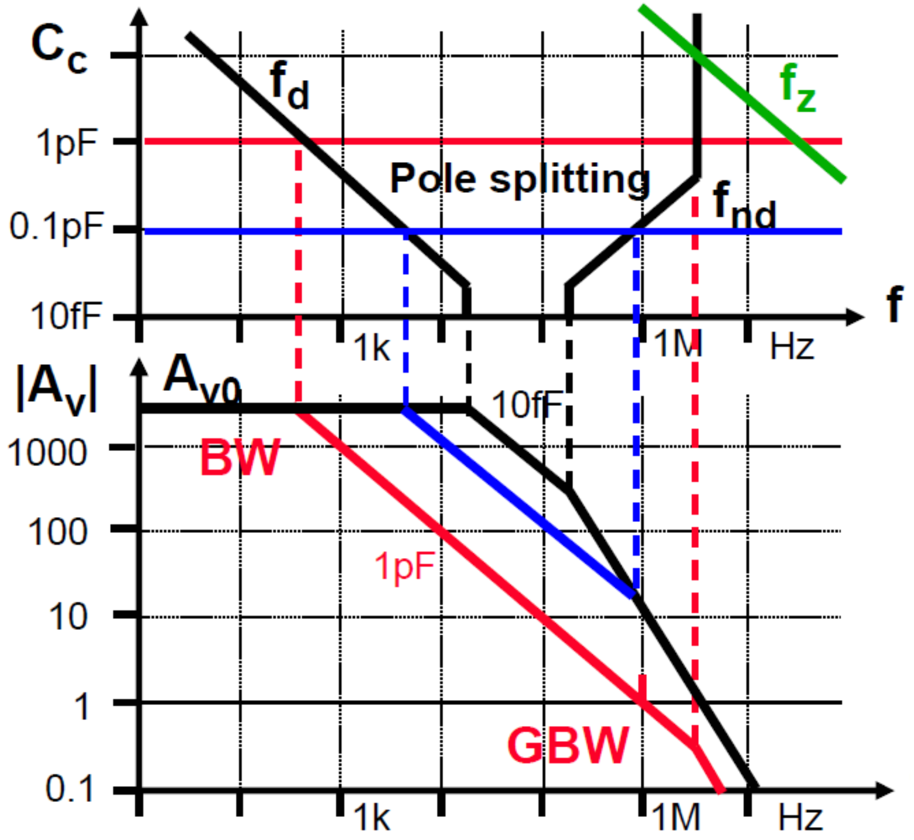


- 80% of the power is consumed in the second stage to achieve stability
  - 80% of bias current do not contribute (directly) to GBW
  - Miller OTA is very energy inefficient!

Miller OTA: Pole Splitting with C_C

- Too large C_C does not give more pole splitting: just smaller GBW
  - Usually we choose $C_1 < C_C < C_L$
  - Reasonable starting point: $C_C \approx (0.3 \rightarrow 0.5) \times C_L$

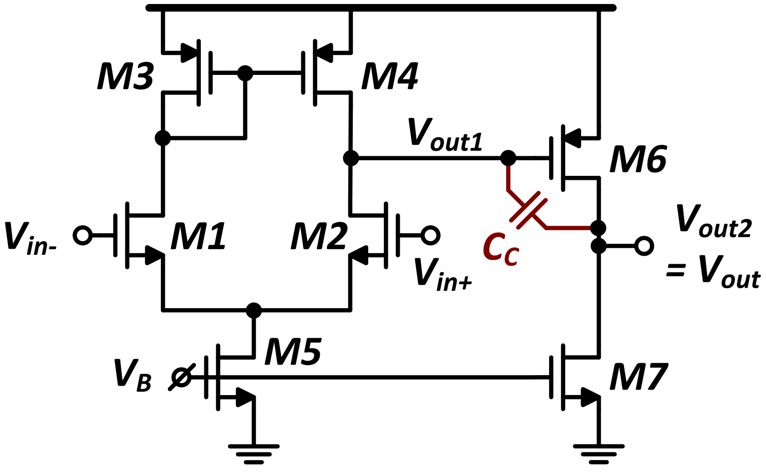


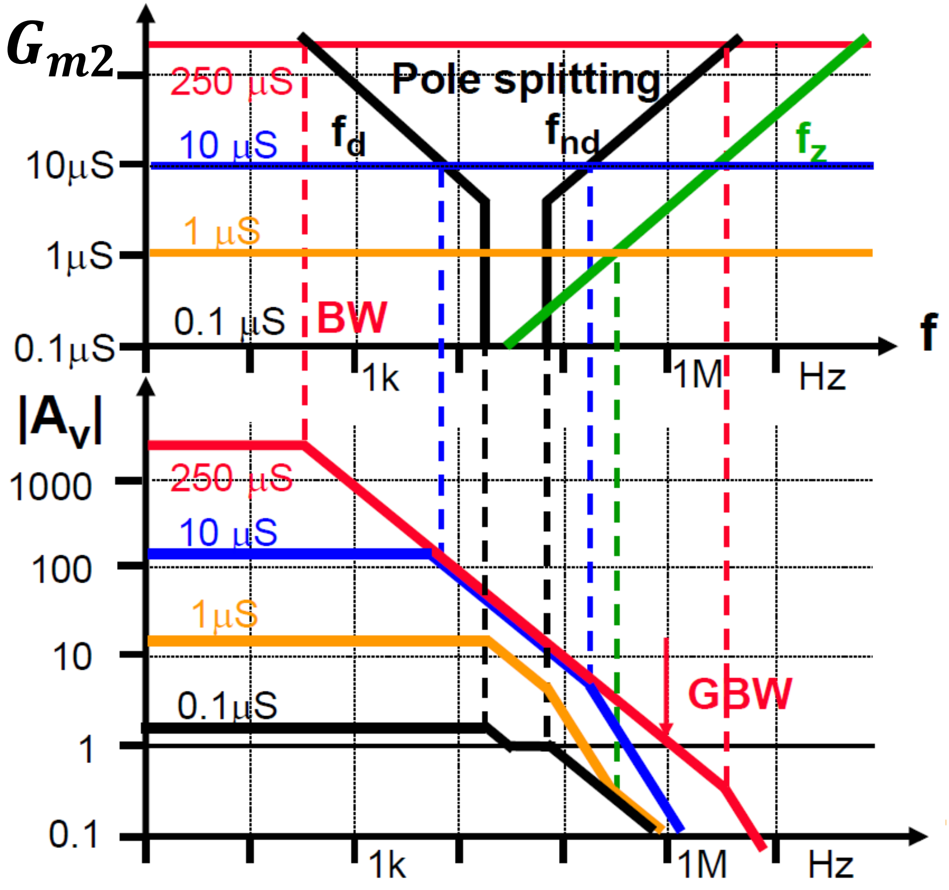


[Sansen, Essentials, 2006]

Miller OTA: Pole Splitting with G_{m2}

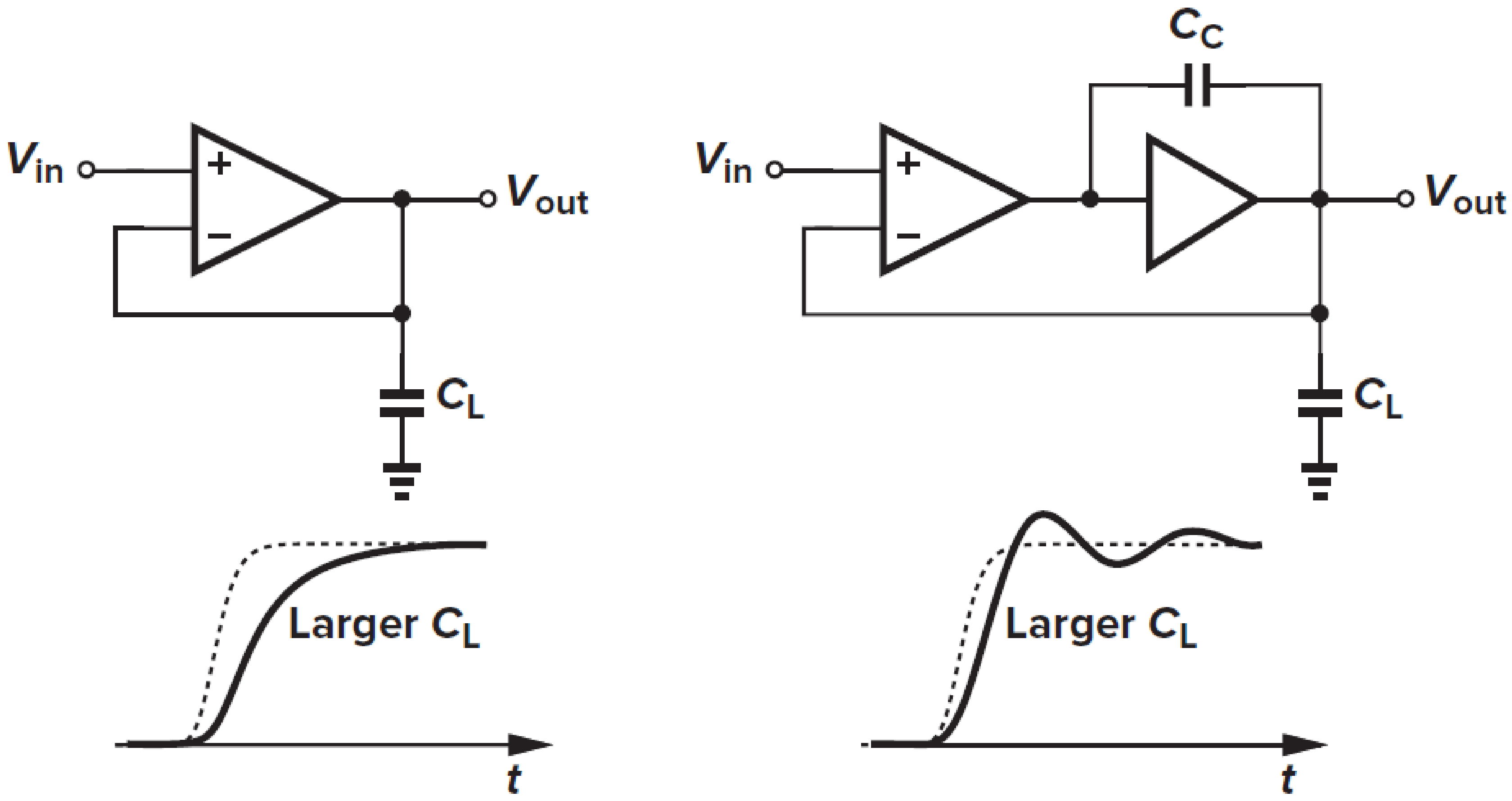
- Increasing G_{m2} works even better than increasing C_C
  - But more power consumption in the 2nd stage





[Sansen, Essentials, 2006]

Single vs Two-Stage OTA: Sensitivity to C_L



[Razavi, 2017]

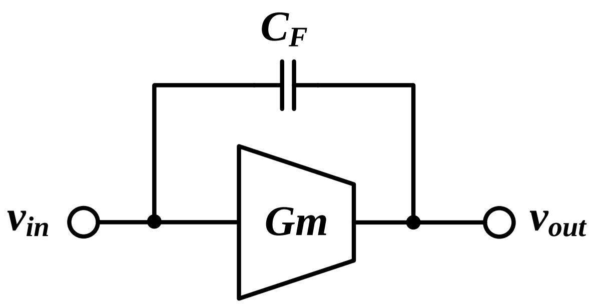
Outline

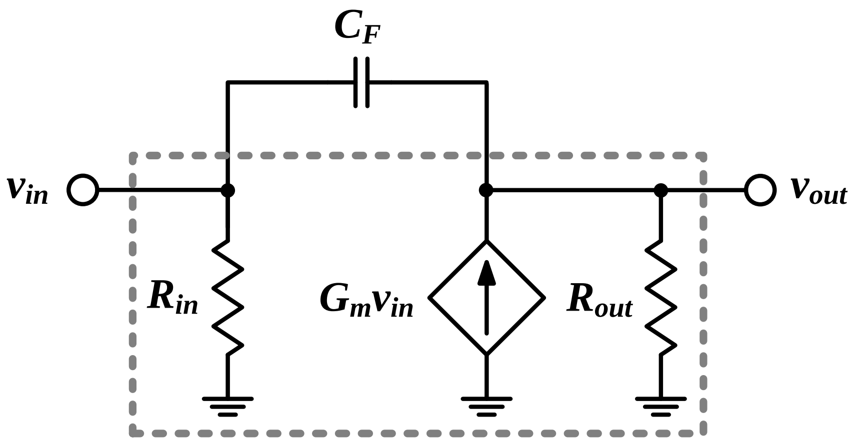
- Recapping previous key results
- Frequency compensation
- Compensation of single-stage OTAs
- Compensation of two-stage OTA
  - Miller compensation
  - The feedforward zero

The Feedforward Zero: Reminder

- $v_{out} = 0 \;\rightarrow i_{out} = 0$
- $v_{in}sC_F = -G_m v_{in}$

$$\boxed{s_z = -\frac{G_m}{C_F}}$$



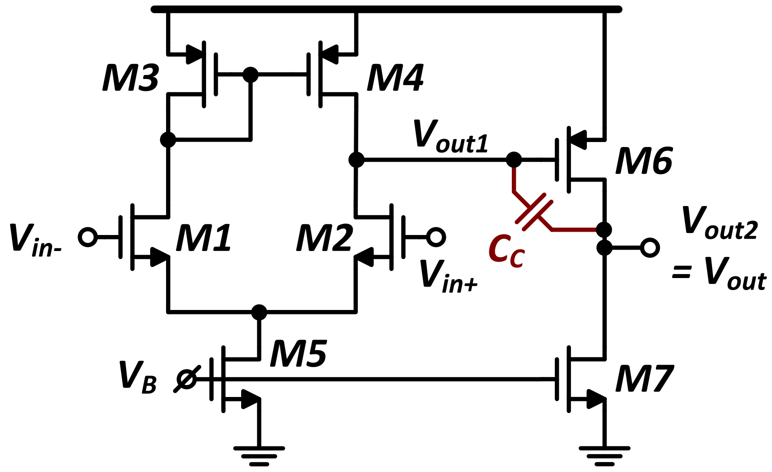


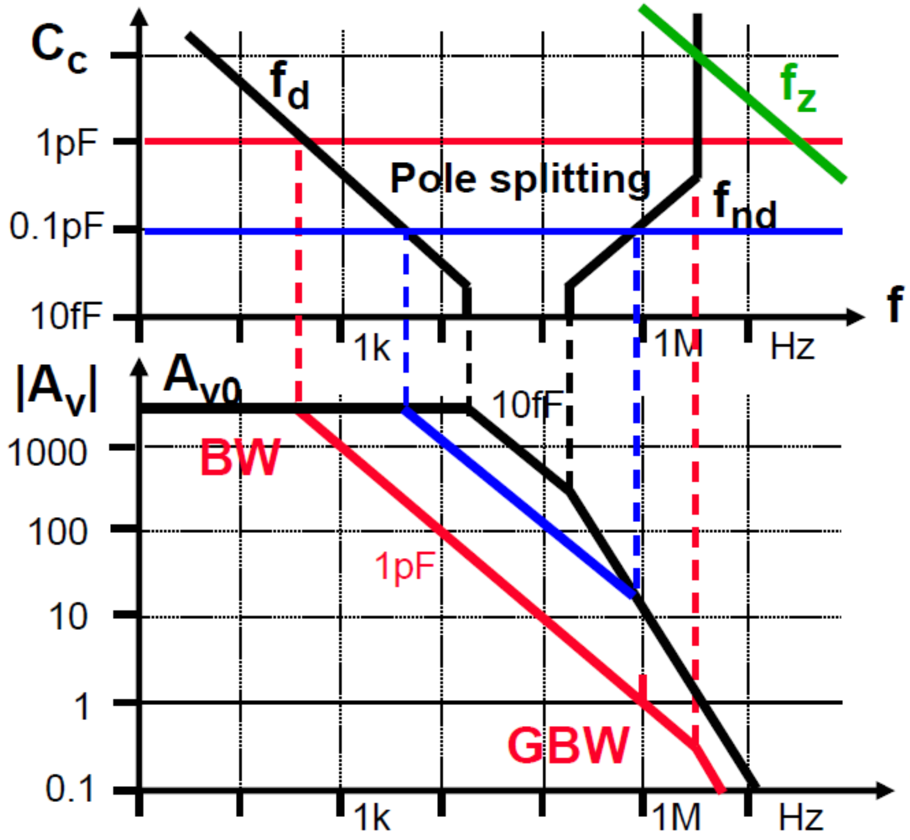
- LHP zero if G_m is +ve (e.g. CD)
- **RHP zero if G_m is –ve (e.g. CS)**
  - Mag inc and phase drops
  - Very bad for FB loop stability

Miller OTA: The RHP Zero

- RHP zero is bad for both magnitude and phase
  - Pushes GX outwards and pushes PX inwards
  - Increasing C_C may hurt stability!

$$\omega_z = \frac{g_{m6}}{C_C + C_{gd}} \approx \frac{G_{m2}}{C_C}$$





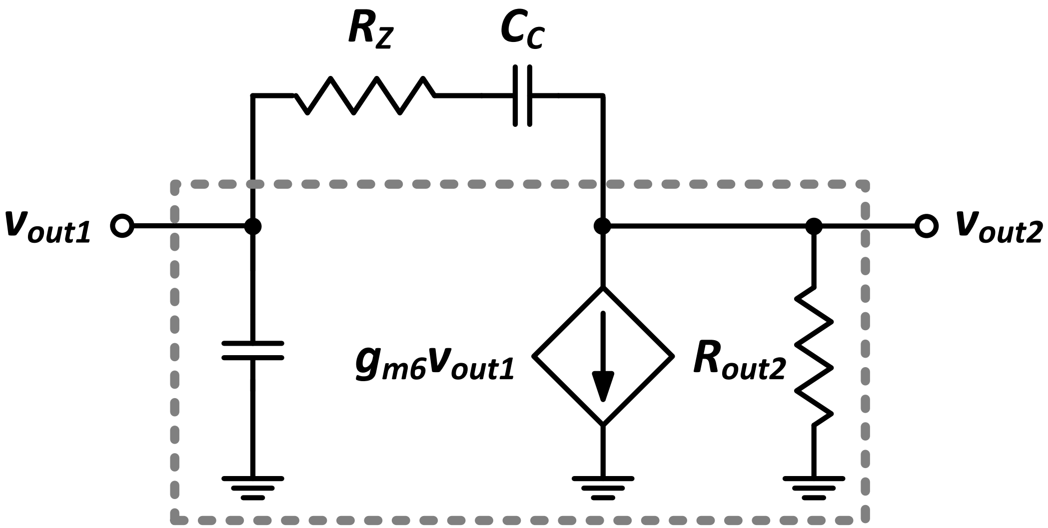
[Sansen, Essentials, 2006]

Handling the RHP Zero

- Add a resistance to control the value of the zero (other tricks exist)

$$\frac{v_{out1}}{R_Z + \frac{1}{s_Z C_C}} = g_{m6} v_{out1}$$

$$s_Z = \frac{1}{C_C\left(\frac{1}{g_{m6}} - R_Z\right)}$$



Miller Zero Placement

- Place the zero at ∞

$$\omega_Z = \frac{1}{C_C\left(\frac{1}{G_{m2}} - R_Z\right)} \rightarrow R_Z = \frac{1}{G_{m2}}$$

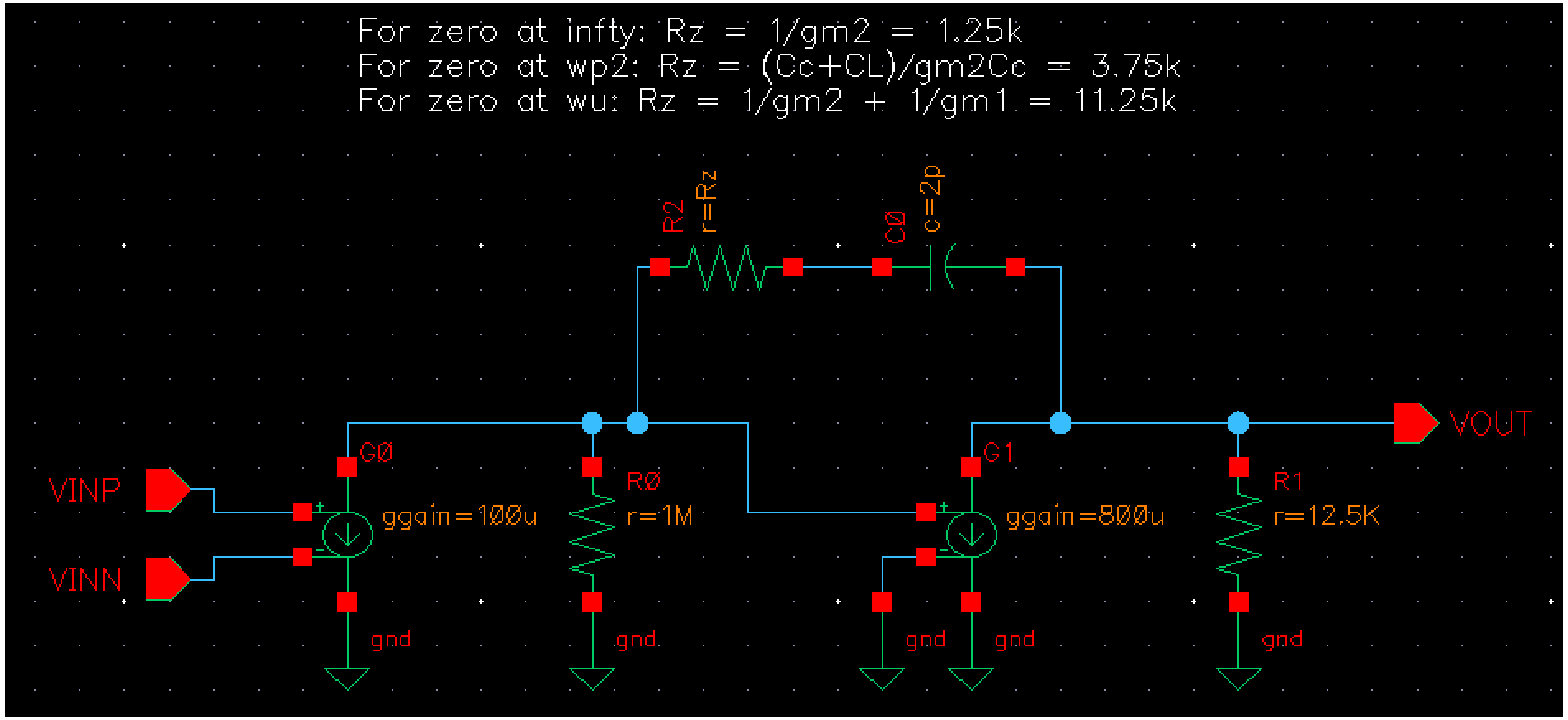
- Some designers try to move the zero to the LHP to cancel the first non-dominant pole and/or improve the PM

$$\frac{1}{C_C\left(R_Z - \frac{1}{G_{m2}}\right)} = \frac{G_{m2}}{C_L} \rightarrow R_z = \frac{C_L + C_C}{G_{m2} C_C}$$

  - Practically never achieved due to variations
  - Actually does not lead to faster response (why?)
- Pushing the LHP zero to lower frequencies is even worse
  - Very poor settling time (a.k.a. pole-zero doublet)
  - Noise amplification

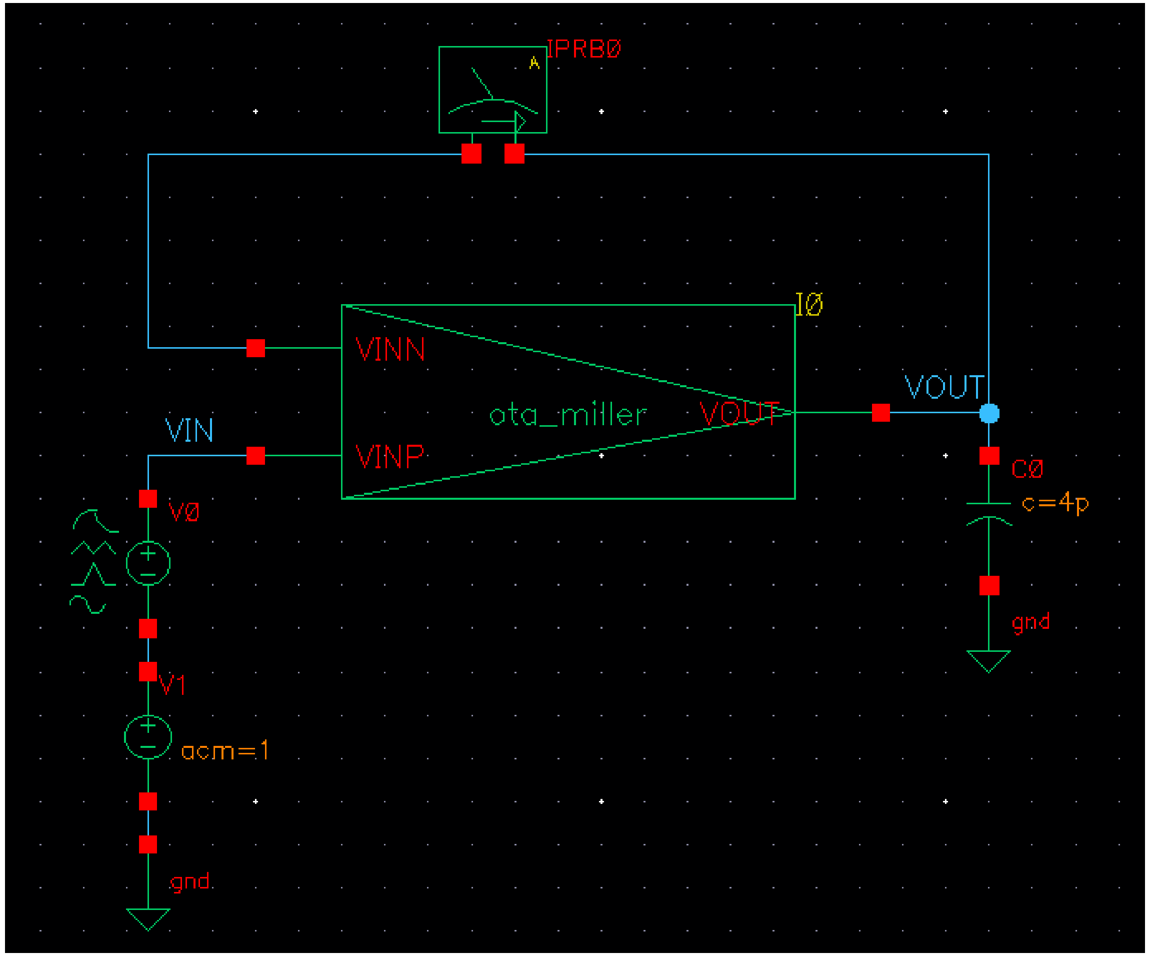
Miller OTA Behavioral Model

- Try three different strategies for Miller zero placement
  - ∞
  - Cancel ω_{p2}
  - At ω_u (this is a designer mistake rather than a valid strategy!)

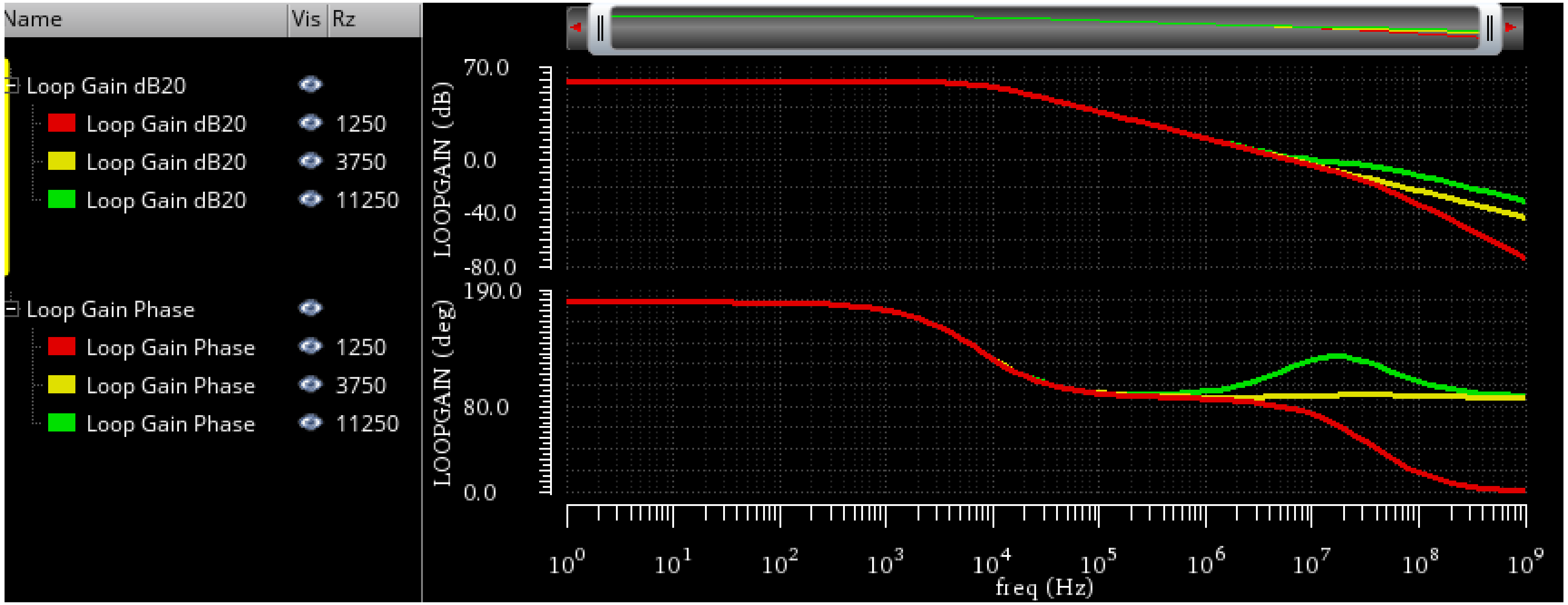


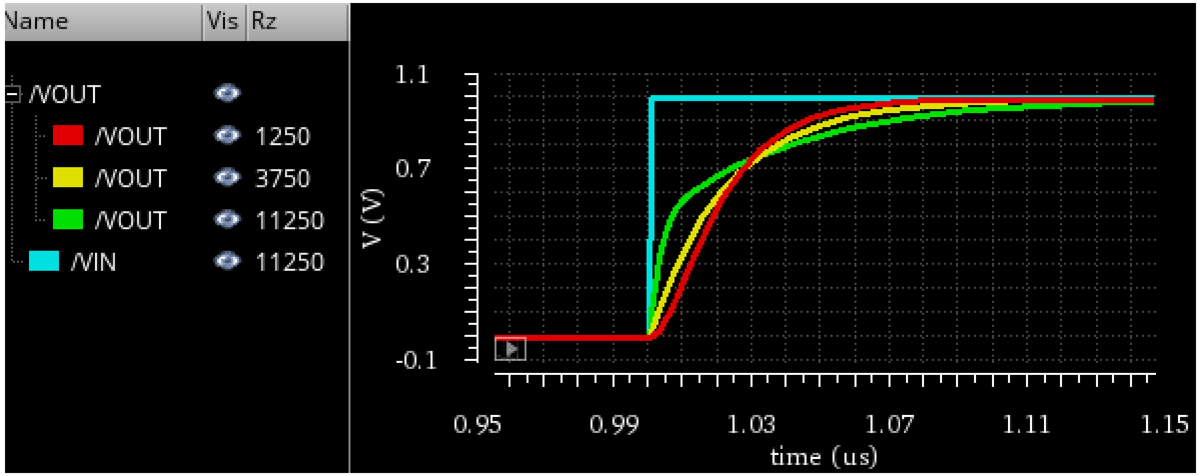
TB for LG and Step Response

- ❑ Verify stability, ac response, and transient response in unity-gain buffer configuration.

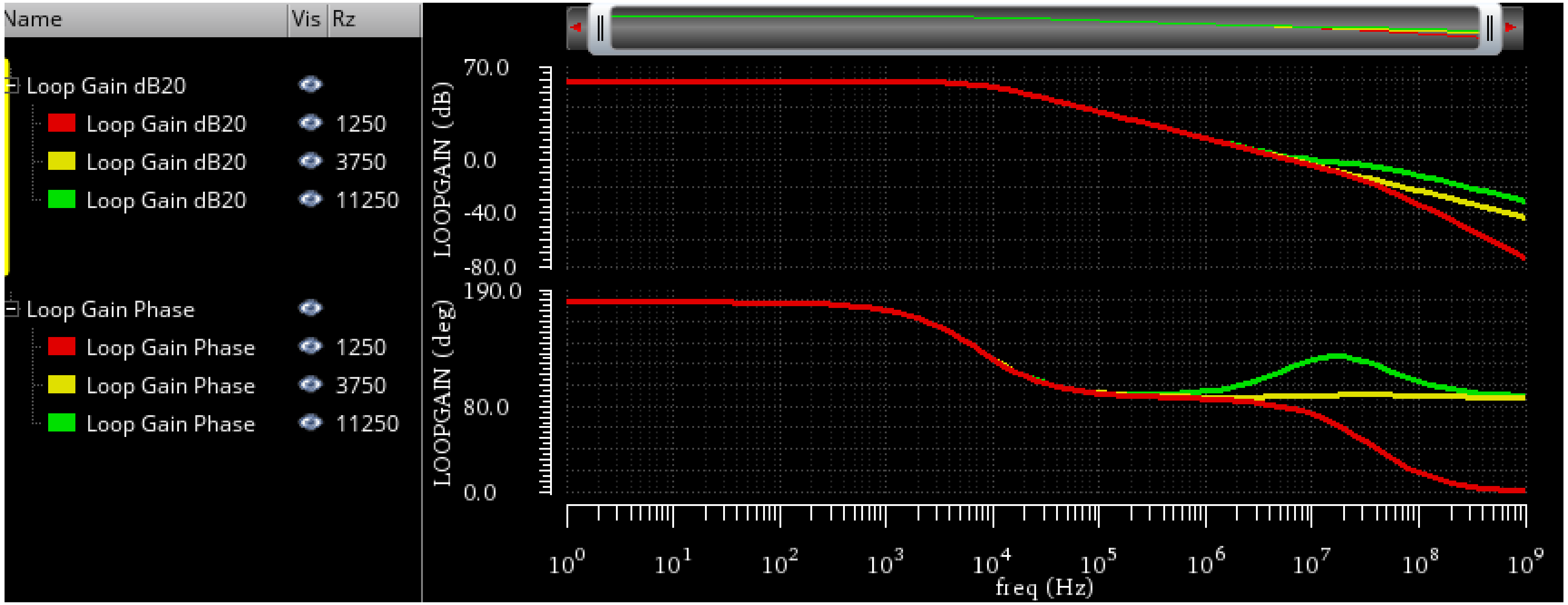


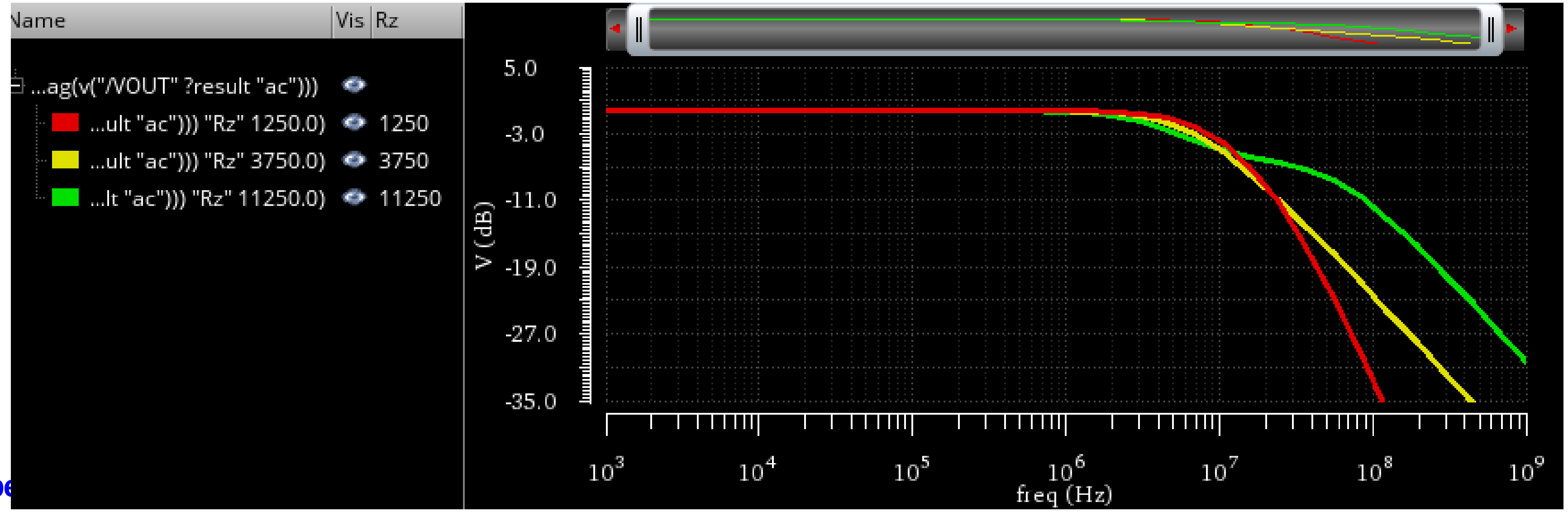
LG and Step Response





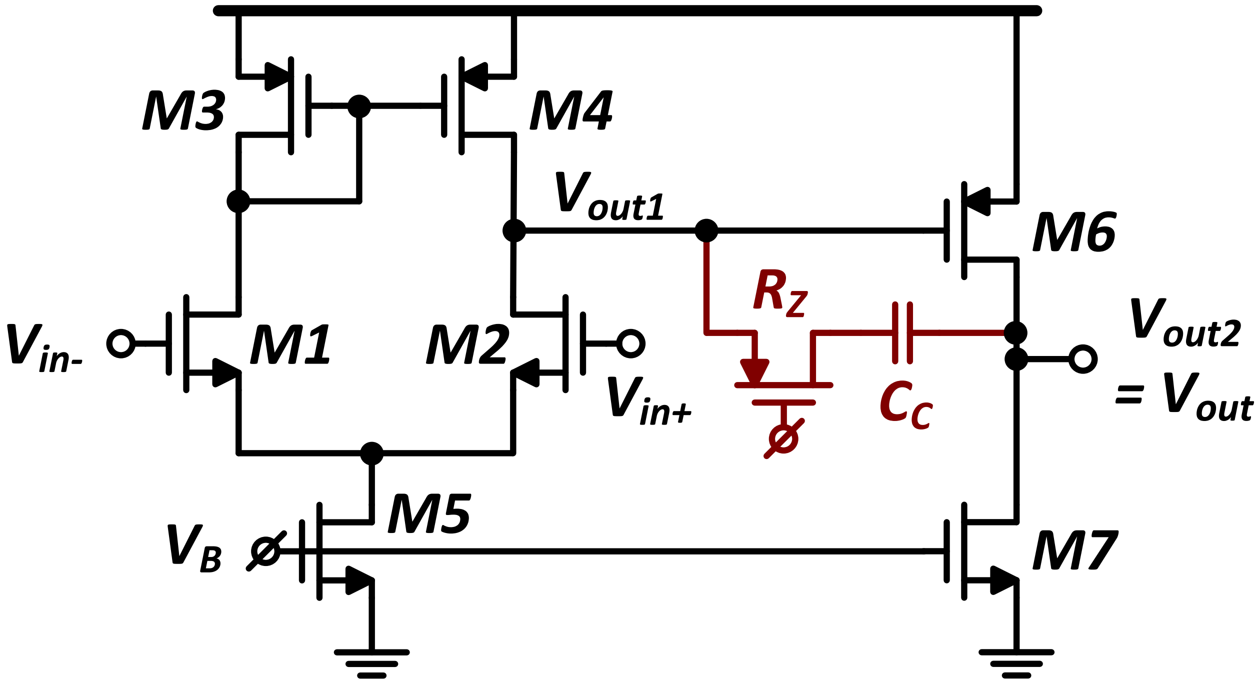
LG and CL Gain





R_Z Implementation

- R_Z can be implemented as a simple resistor or using a transistor in the linear region
  - Clever biasing can make it track variations (maintain $R_Z \approx \frac{1}{g_{m6}}$)
  - Why R_Z is implemented as PMOS?



Thank you!

References

- A. Sedra and K. Smith, “Microelectronic Circuits,” Oxford University Press, 7th ed., 2015
- B. Razavi, “Design of Analog CMOS Integrated Circuits,” McGraw-Hill, 2nd ed., 2017.
- W. Sansen, “Analog design essentials,” Springer, 2006.
- T. C. Carusone, D. Johns, and K. W. Martin, “Analog Integrated Circuit Design,” 2nd ed., Wiley, 2012.
- B. Murmann, EE214 Course Reader, Stanford University.