وَمَا أُوتِيتُمْ مِنَ الْعِلْمِ إِلَّا قَلِيلًا

# Analog IC Design

# Lecture 16
# OTA Stability and Compensation

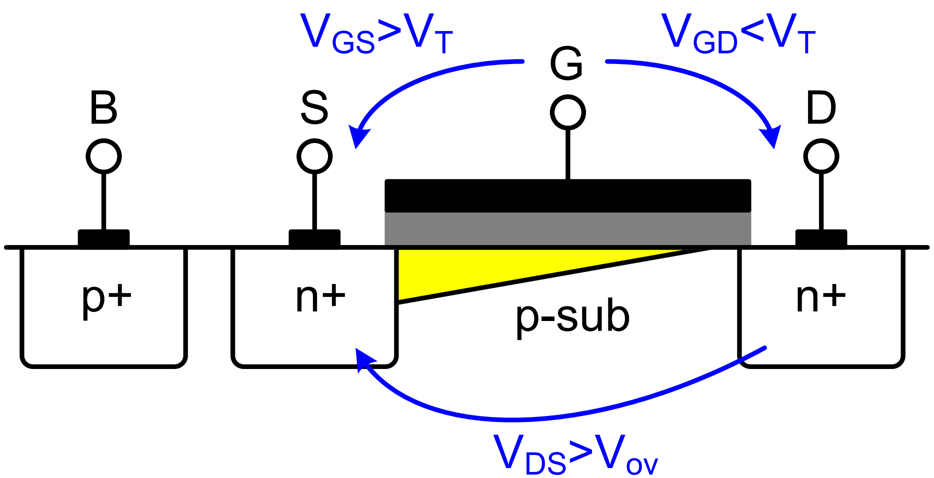

Dr. Hesham A. Omran

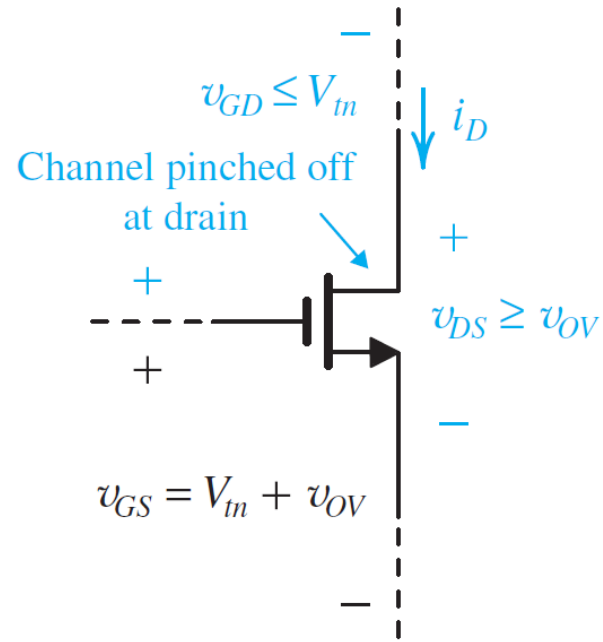
Integrated Circuits Lab (ICL)
Electronics and Communications Eng. Dept.
Faculty of Engineering
Ain Shams University

# 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

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




 [Sedra/Smith, 2015] 

# Regions of Operation Summary

**OFF (Subthreshold)**

$V_{GS} < V_T$

**ON**

$V_{GS} > V_T$

**Triode**

$V_{DS} < V_{ov}$

Or

$V_{GD} > V_T$

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

**Pinch-Off (Saturation)**

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

Or

$V_{GD} \leq V_T$

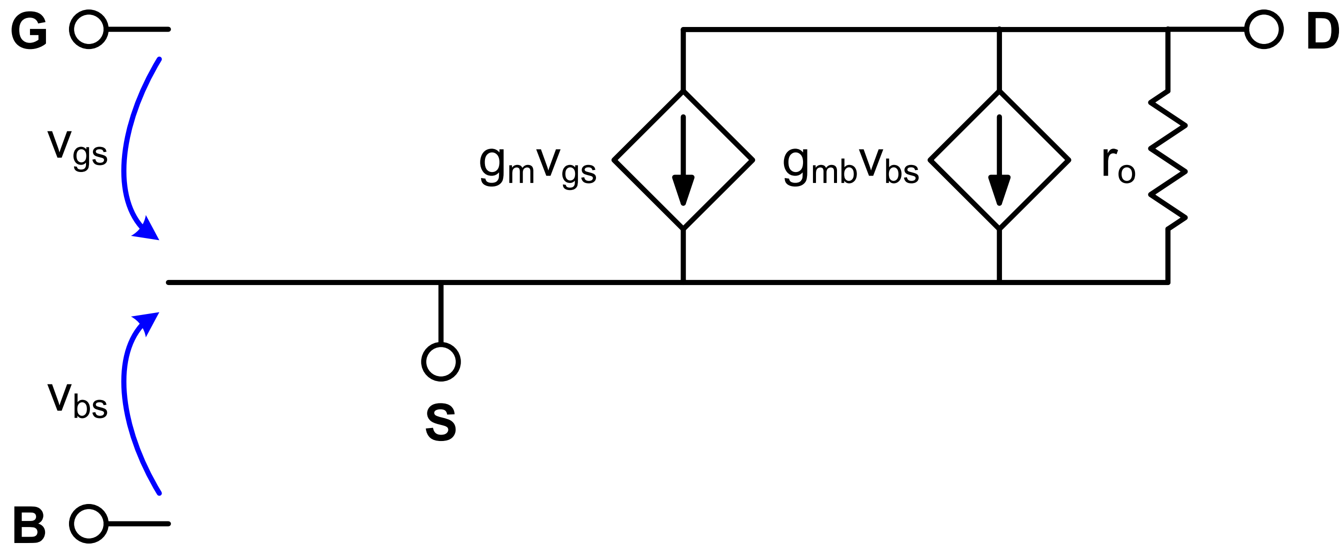
$$I_D = \frac{\mu C_{ox}}{2} \frac{W}{L} V_{ov}^2 (1 + \lambda V_{DS})$$

# Low-Frequency Small-Signal Model

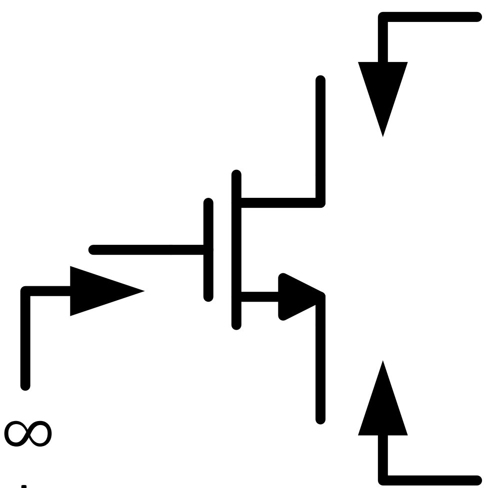
$$g_m = \frac{\partial I_D}{\partial V_{GS}} = \mu C_{ox}\frac{W}{L}V_{ov} = \sqrt{\mu C_{ox}\frac{W}{L}\cdot 2I_D} = \frac{2I_D}{V_{ov}}$$

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

$$r_o = \frac{1}{\frac{\partial I_D}{\partial V_{DS}}} = \frac{1}{\lambda I_D}, \qquad \lambda \propto \frac{1}{L}$$



# Rin/out Shortcuts Summary



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

H.I.N.

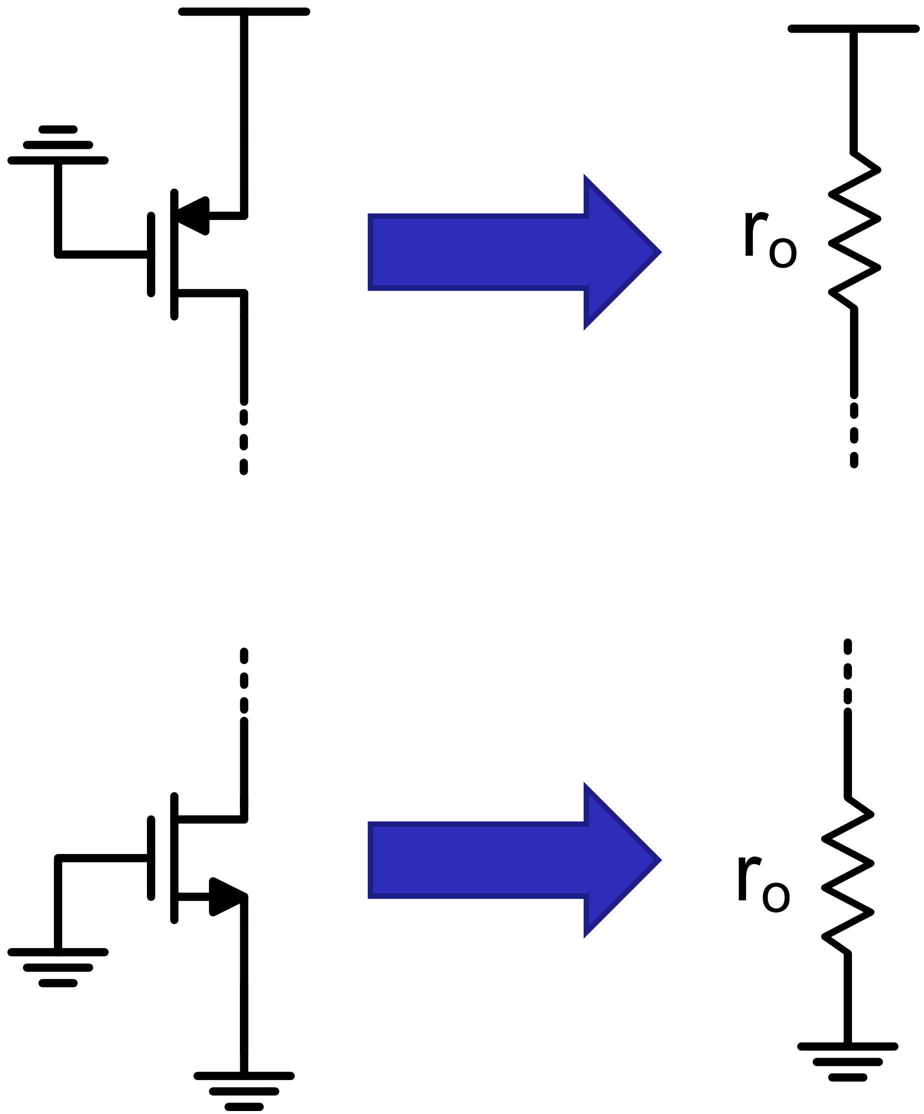
$\infty$

At low frequencies **ONLY**

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

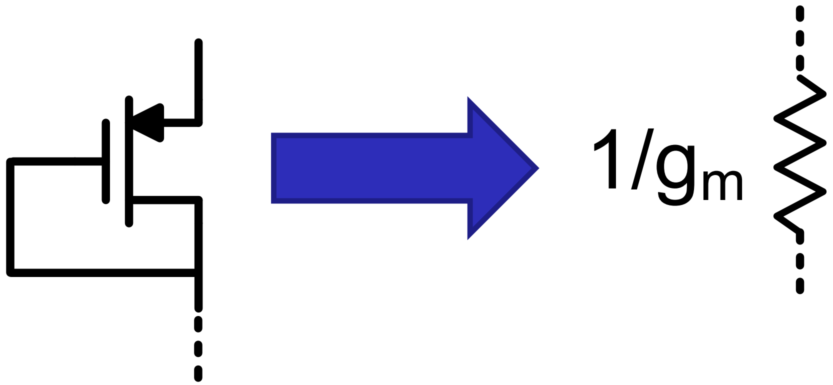
L.I.N.

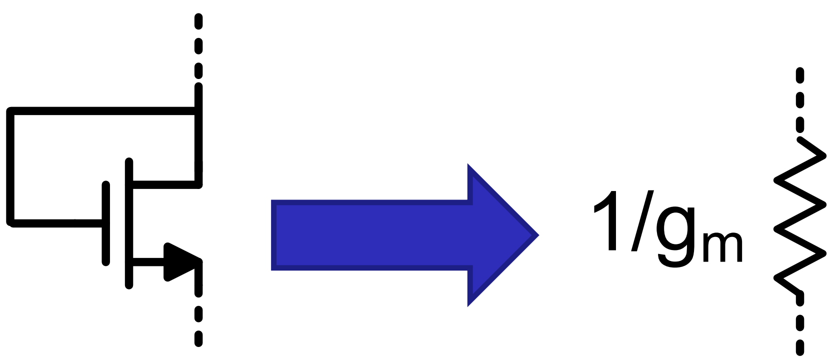
# Active Load (Source OFF)



# Diode Connected (Source Absorption)

- Always in saturation
- Bulk effect: $g_m \rightarrow g_m + g_{mb}$



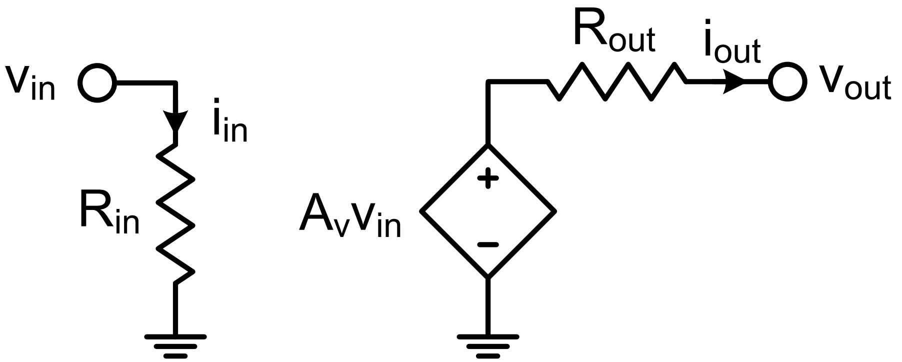
# Why GmRout?

$$R_{out} = \frac{v_x}{i_x} \text{ @ } v_{in} = 0$$

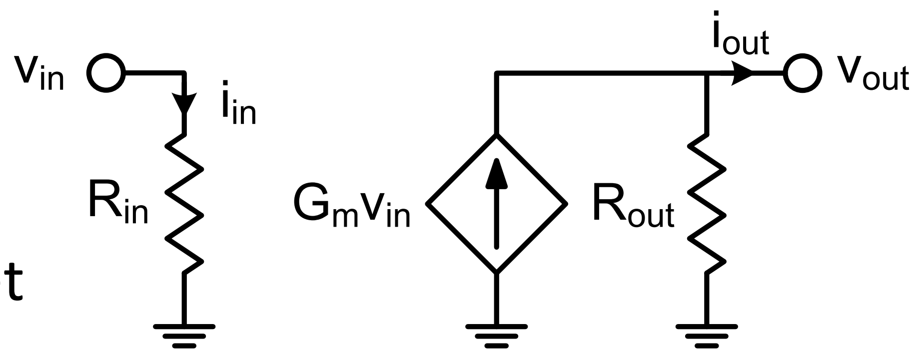
$$G_m = \frac{i_{out,sc}}{v_{in}}$$

$$A_v = G_m R_{out}$$

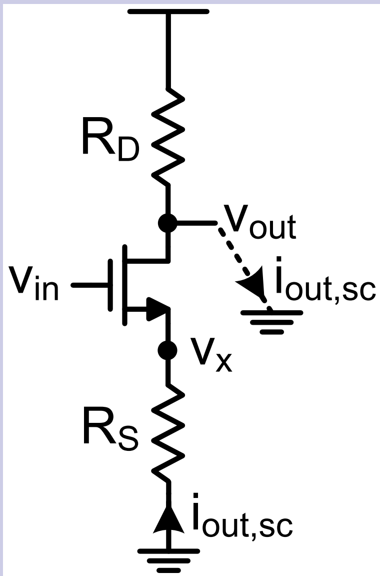
$$A_i = G_m R_{in}$$

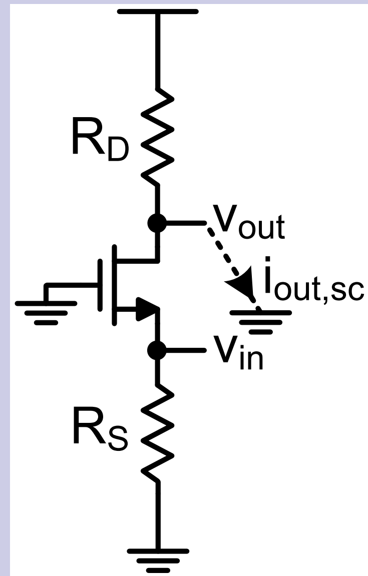


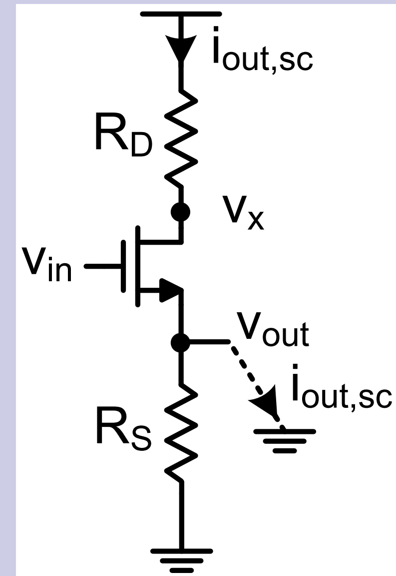

- Divide and conquer
  - Rout simplified: vin=0
  - Gm simplified: vout=0
  - We already need Rin/out
  - We can quickly and easily get Rin/out from the shortcuts



# Summary of Basic Topologies



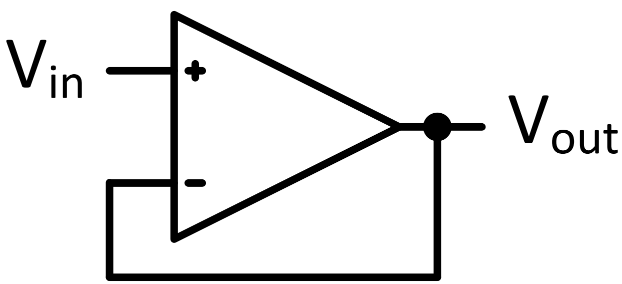

| | CS | CG | CD (SF) |
|---|---|---|---|
| | Voltage & current amplifier | Current buffer | Voltage buffer |
| **Rin** | $\infty$ | $R_S // \frac{1}{g_m + g_{mb}}\left(1 + \frac{R_D}{r_o}\right)$ | $\infty$ |
| **Rout** | $R_D // r_o[1 + (g_m + g_{mb})R_S]$ | $R_D // r_o$ | $R_S // \frac{1}{g_m + g_{mb}}\left(1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

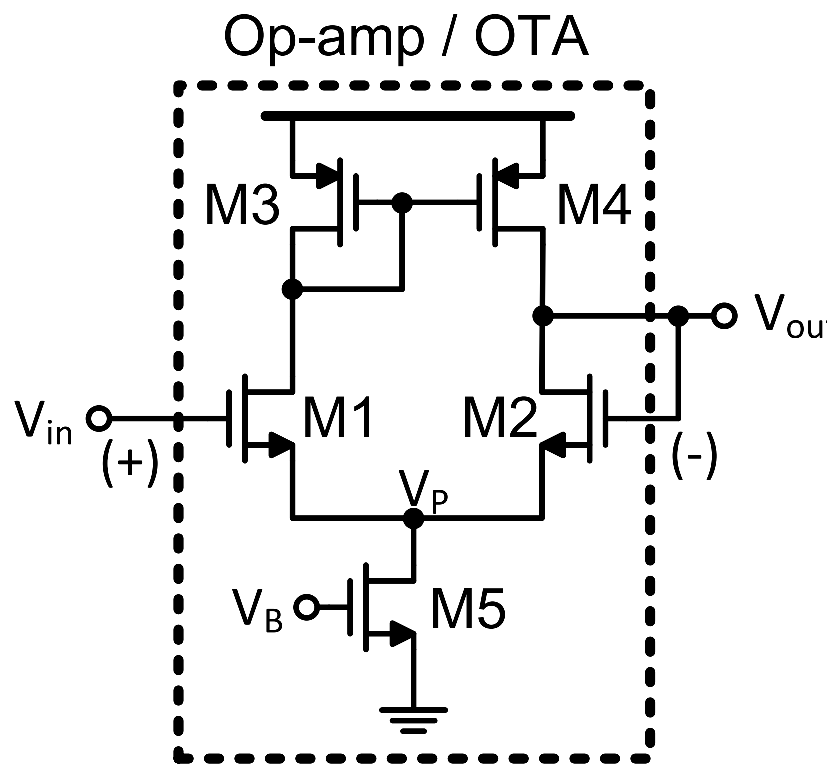
# Differential Amplifier

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

# What is an OTA / Op-Amp?

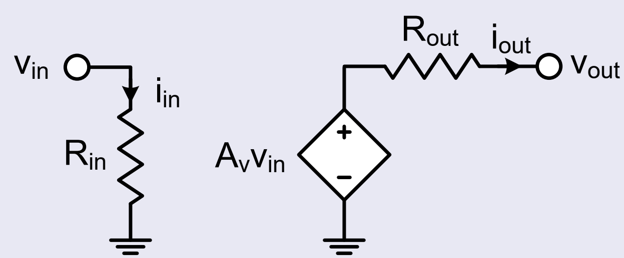
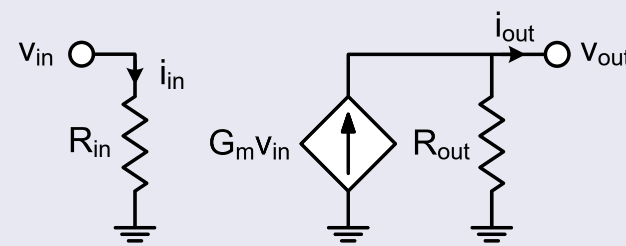
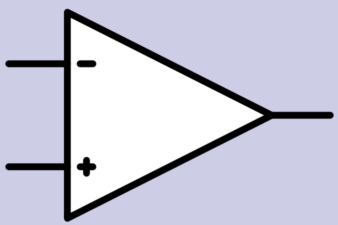
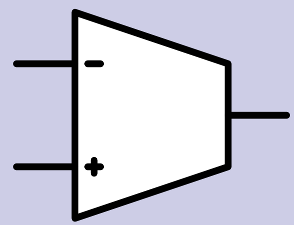
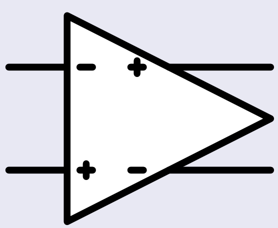
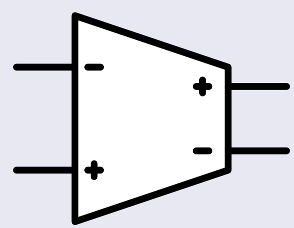
- An op-amp is simply a high gain differential amplifier
- The gain can be increased by using cascodes and multi-stage amplifiers



# Op-Amp vs OTA

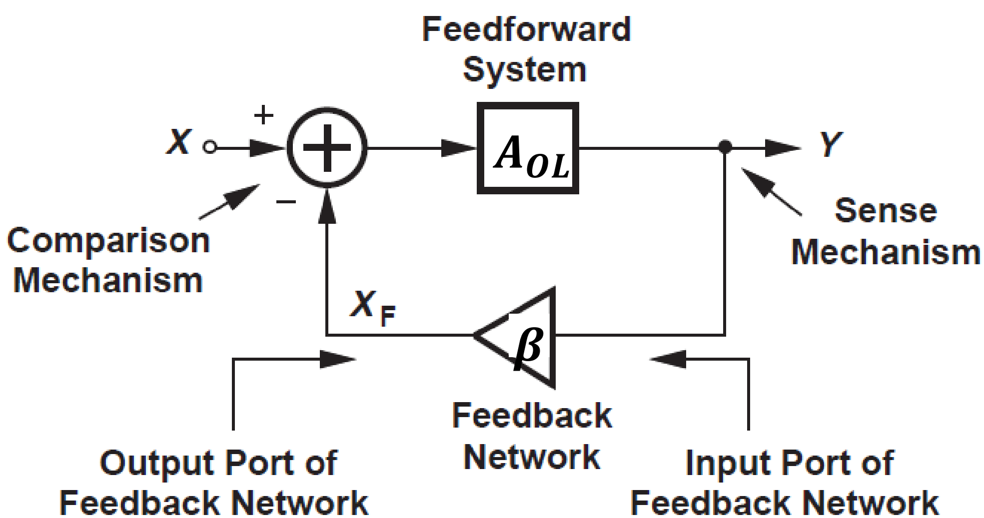
- An OTA is an op-amp without an output stage (buffer)
- Some designers just use op-amp name and symbol for both

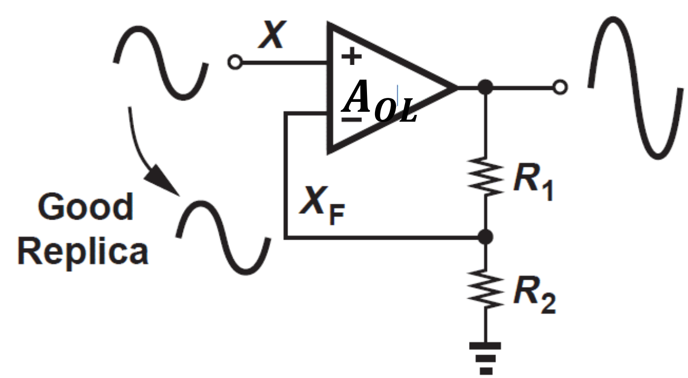
| | Op-amp | OTA |
|---|---|---|
| **Rout** | **LOW** | **HIGH** |
| **Model** |  |  |
| **Diff input, SE output** |  |  |
| **Fully diff** |  |  |

# Negative Feedback

- $A_{OL}$ = Open loop (OL) gain $\gg 1$
- $A_{CL} = \frac{Y}{X}$ = Closed loop (CL) gain
- Error signal = $X - X_F$

$$Y = A_{OL}(X - X_F) = A_{OL}(X - \beta Y)$$
$$A_{CL} = \frac{Y}{X} = \frac{A_{OL}}{1 + \beta \cdot A_{OL}} \approx \frac{1}{\beta}$$

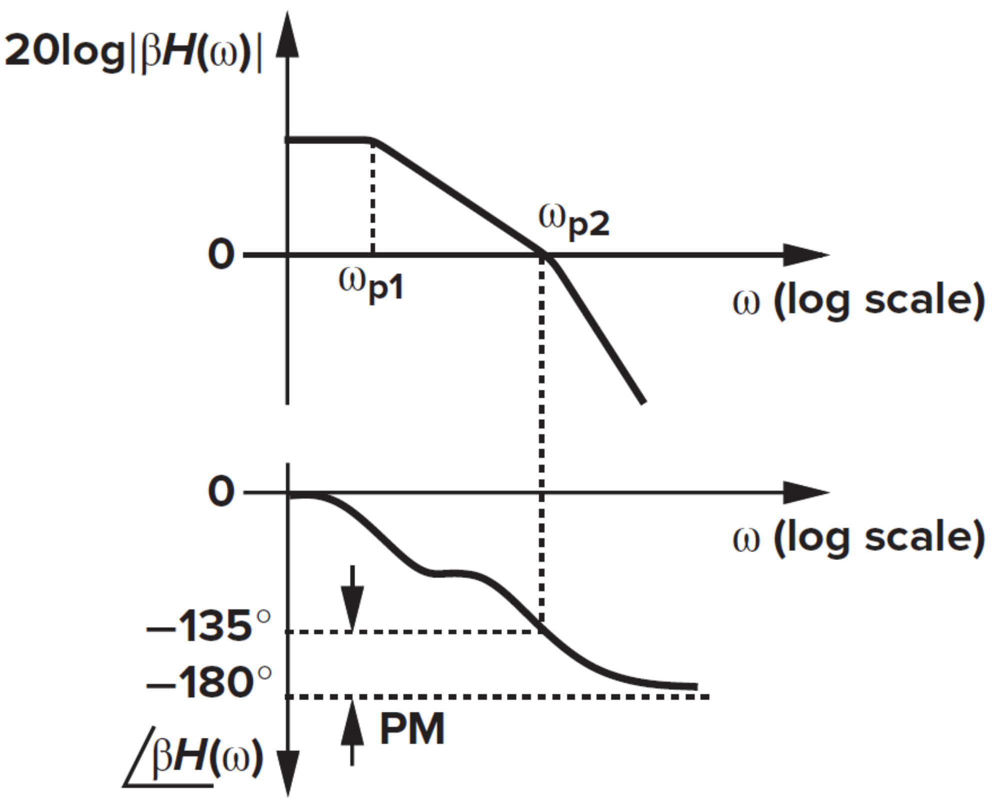


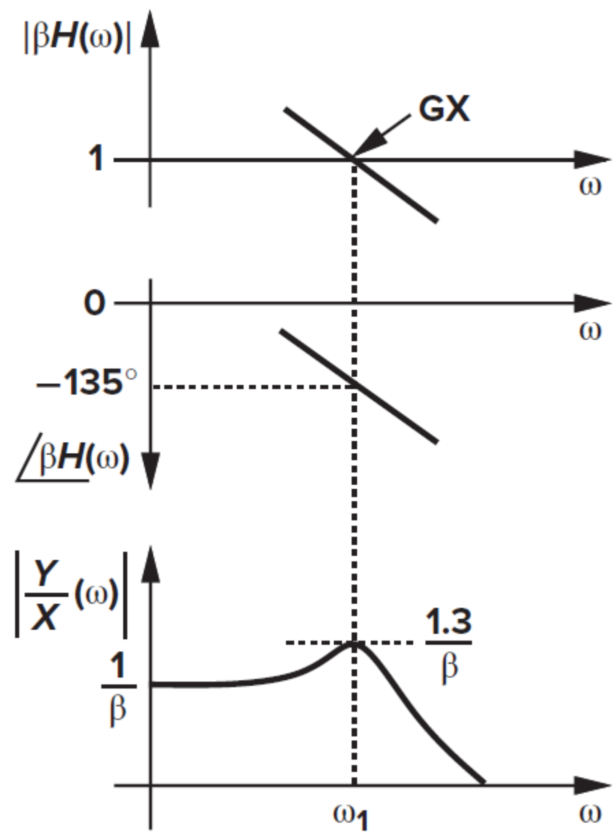




[Razavi, 2014]

# Stability: Phase Margin

- If $\omega_{p2} = \omega_u$: PM = 45° → typically inadequate (peaking/ringing)
- The ultimate $\omega_u$ cannot exceed $\omega_{p2}$ → $\omega_{p1} < \omega_u < \omega_{p2}$
  - For $\omega < \omega_u$ the Bode plot is similar to a 1st order system





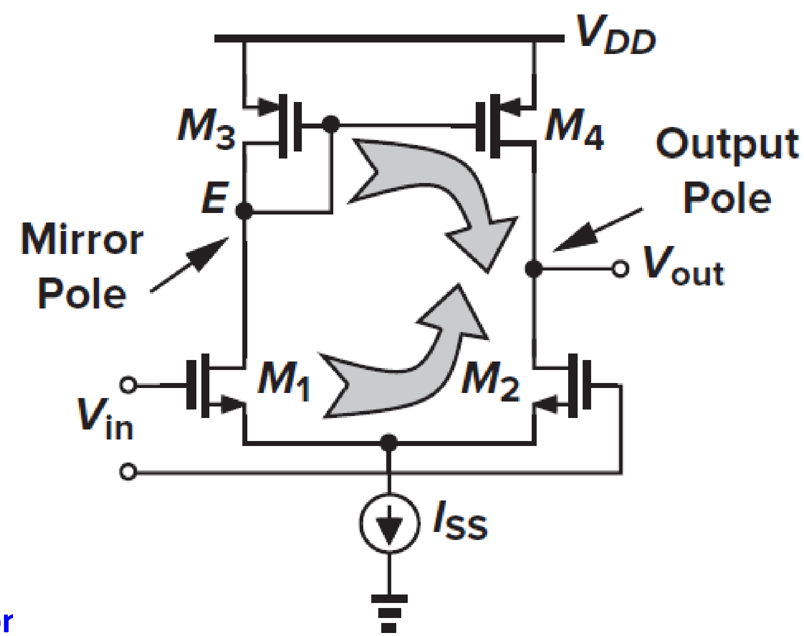

[Razavi, 2017]

# Famous OTA Topologies

1. Simple single-stage OTA
2. Telescopic cascode OTA
3. Folded cascode OTA
4. Two-stage OTA
5. Gain boosted OTA

# Simple Single-Stage OTA

- Simple, but limited gain
- $\omega_{p1} < \omega_u < \omega_{p2}$
- The H.I.N. sets the dominant pole
  - OL bandwidth
- The first non-dominant pole (mirror pole) sets the max GBW
  - Max CL bandwidth (buffer)





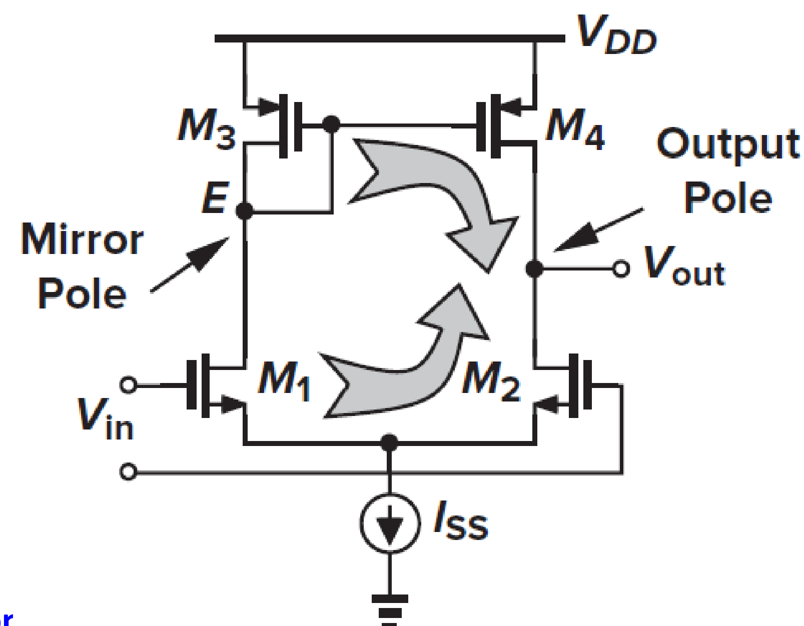

[Razavi, 2017]

# Simple Single-Stage OTA: Poles

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

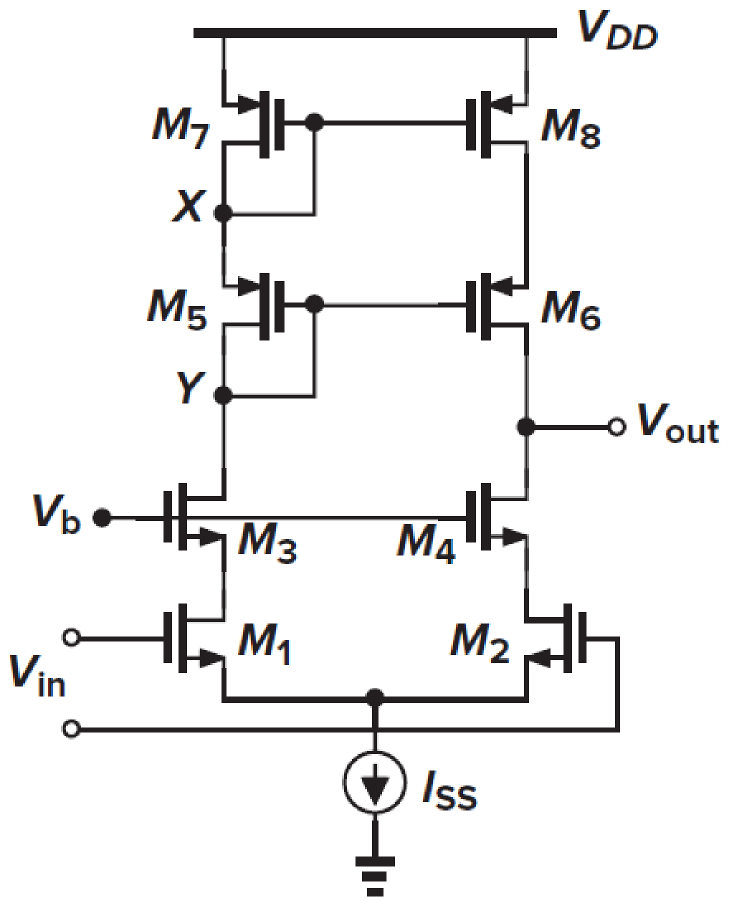
$$\omega_{p2} \approx \frac{g_{m3}}{C_E}$$

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



# Telescopic Cascode

- Higher DC gain, but limited swing and additional poles

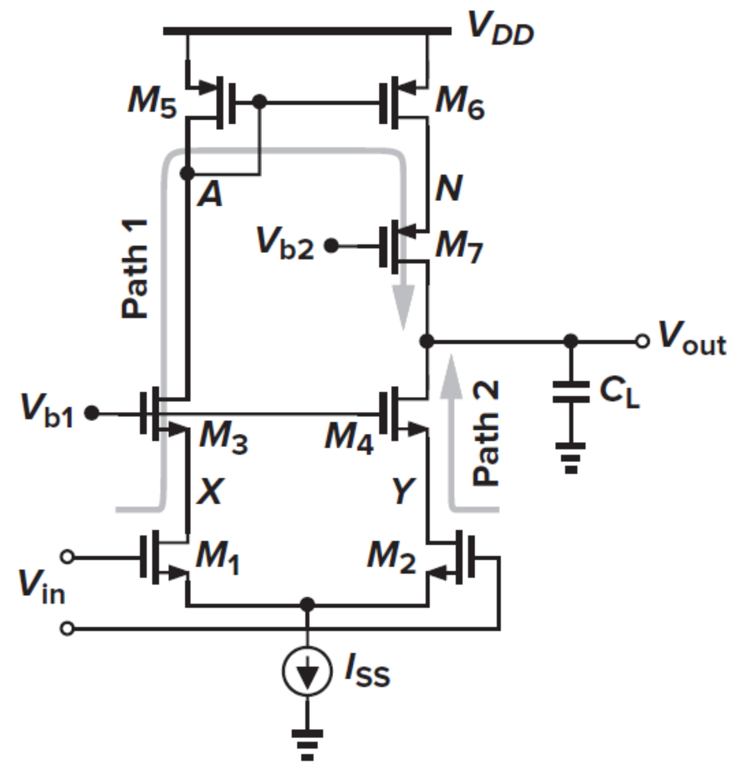


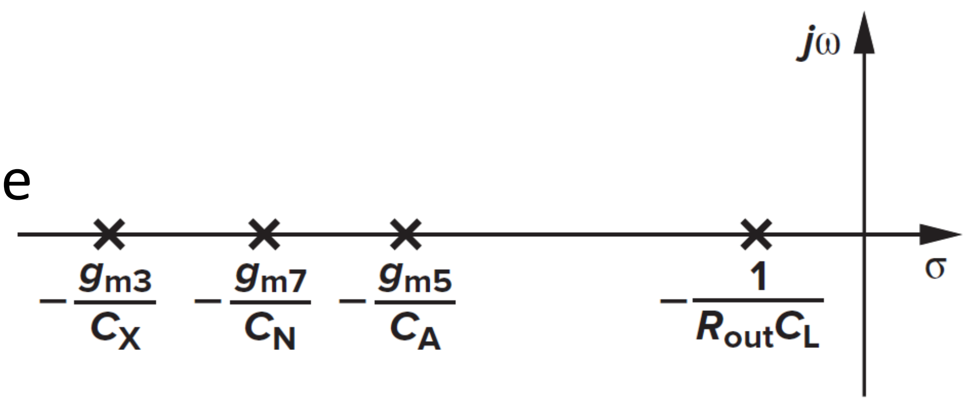



[Razavi, 2017]

# SE Output Telescopic Cascode: Poles

- ❑ $\omega_{p1} < \omega_u < \omega_{p2}$
- ❑ The H.I.N. sets the dominant pole
  - OL bandwidth
- ❑ The first non-dominant pole (mirror pole) sets the ultimate max GBW
  - Max CL bandwidth (buffer)
- ❑ $C_{gs}$ is larger than other caps
- ❑ PMOS contributes larger capacitances (low ID/W)
- ❑ X and Y contribute a single pole

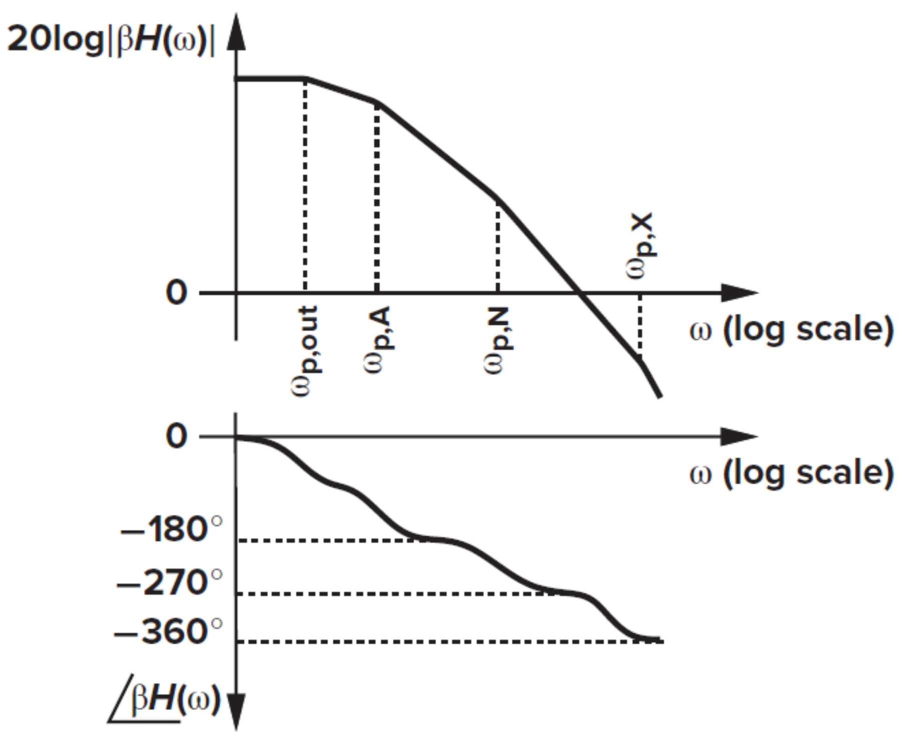


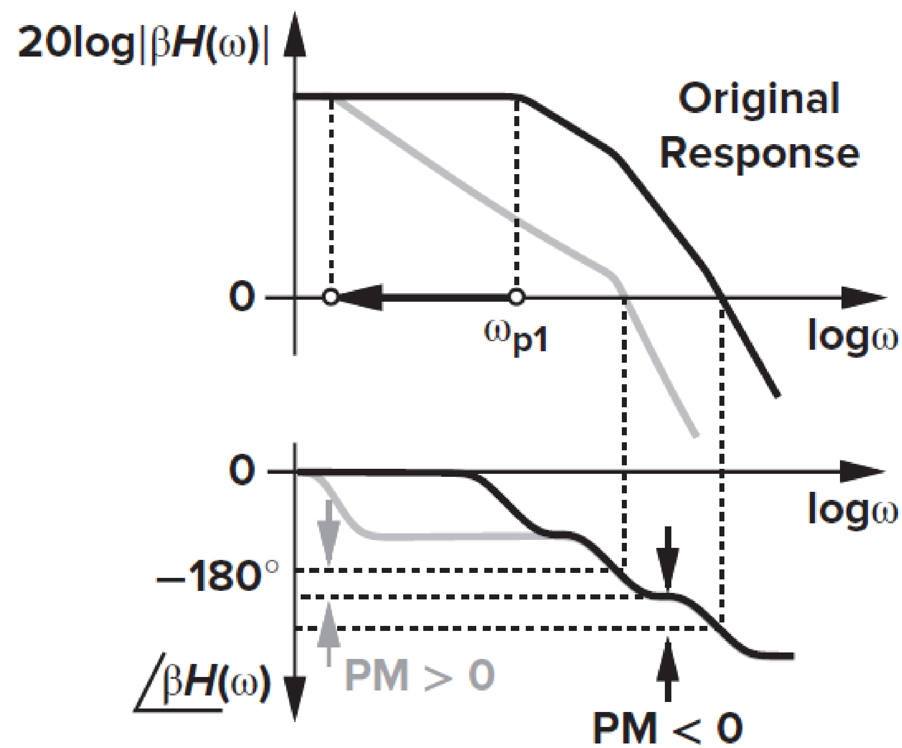




[Razavi, 2017]

# SE Telescopic Cascode: Compensation

- Push GX in: lower GBW
  - Increase $C_L$
  - **Single-stage OTAs are compensated by large load capacitance**

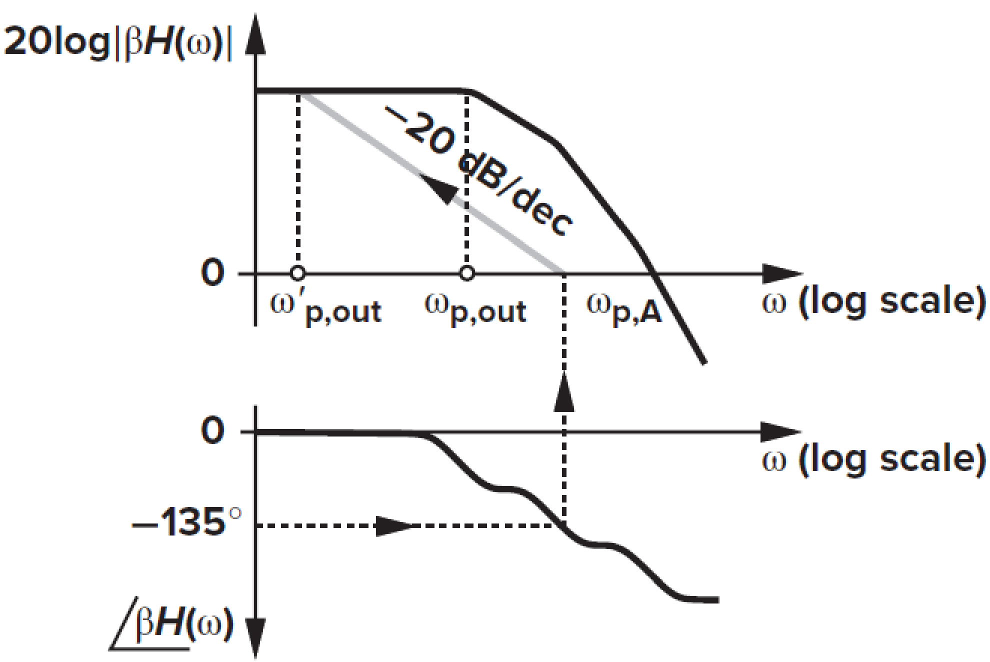


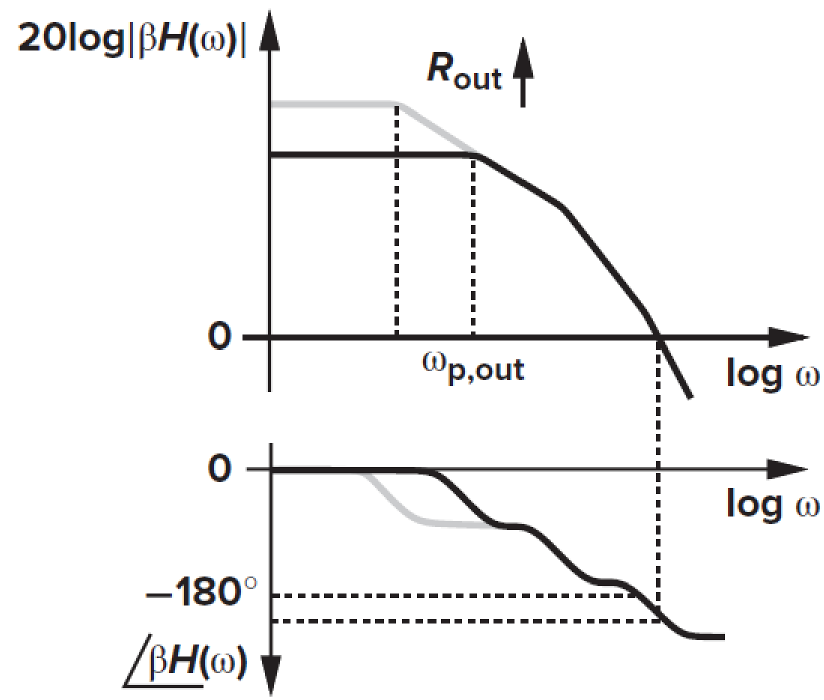




[Razavi, 2017]

# SE Telescopic Cascode: Compensation

- ❑ Push GX in: lower GBW
  - Increase $C_L$
  - **Single-stage OTAs are compensated by large load capacitance**
- ❑ Increasing $R_{out}$ does not affect PM

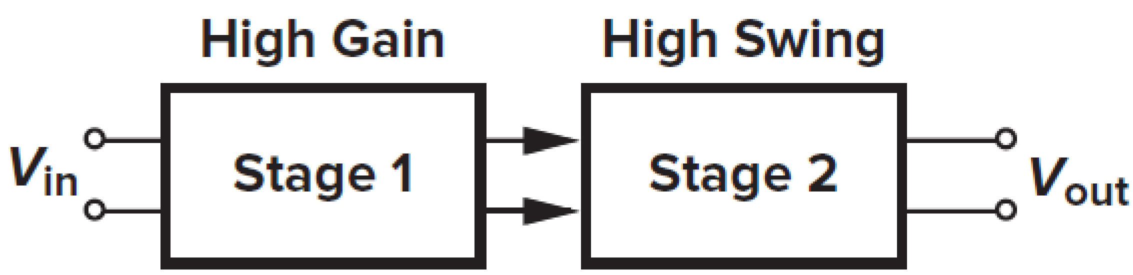



[Razavi, 2017]

# Two-Stage OTA

- Isolates the gain and swing requirements
- But more power consumption
- And complicates stability requirements
  - More than two stages exist, but quite difficult to stabilize
- Second stage is typically configured as a simple common-source stage so as to allow maximum output swings

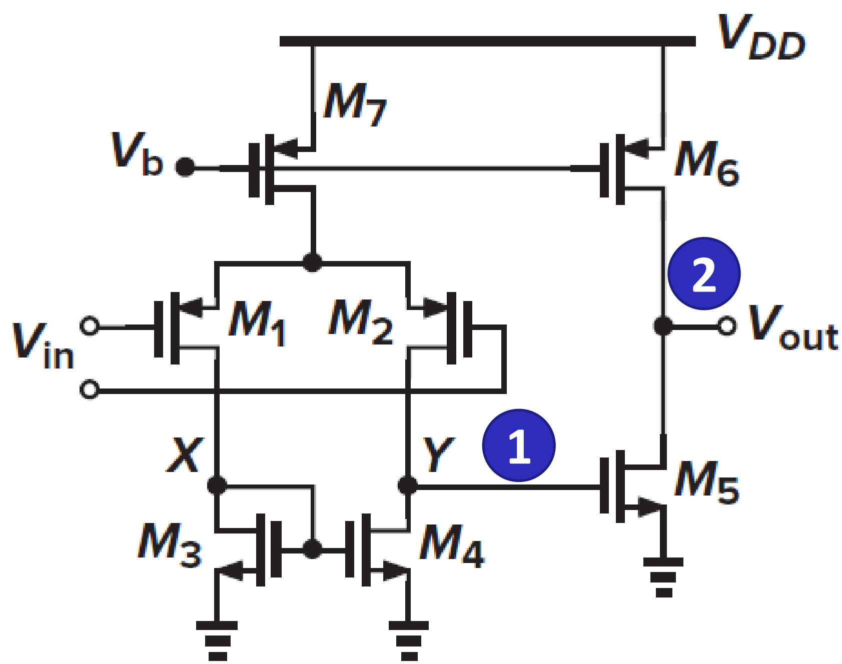



[Razavi, 2017]

# Two-Stage OTA: Poles

- Two H.I.N.s at 1 and 2
  - Two dominant poles

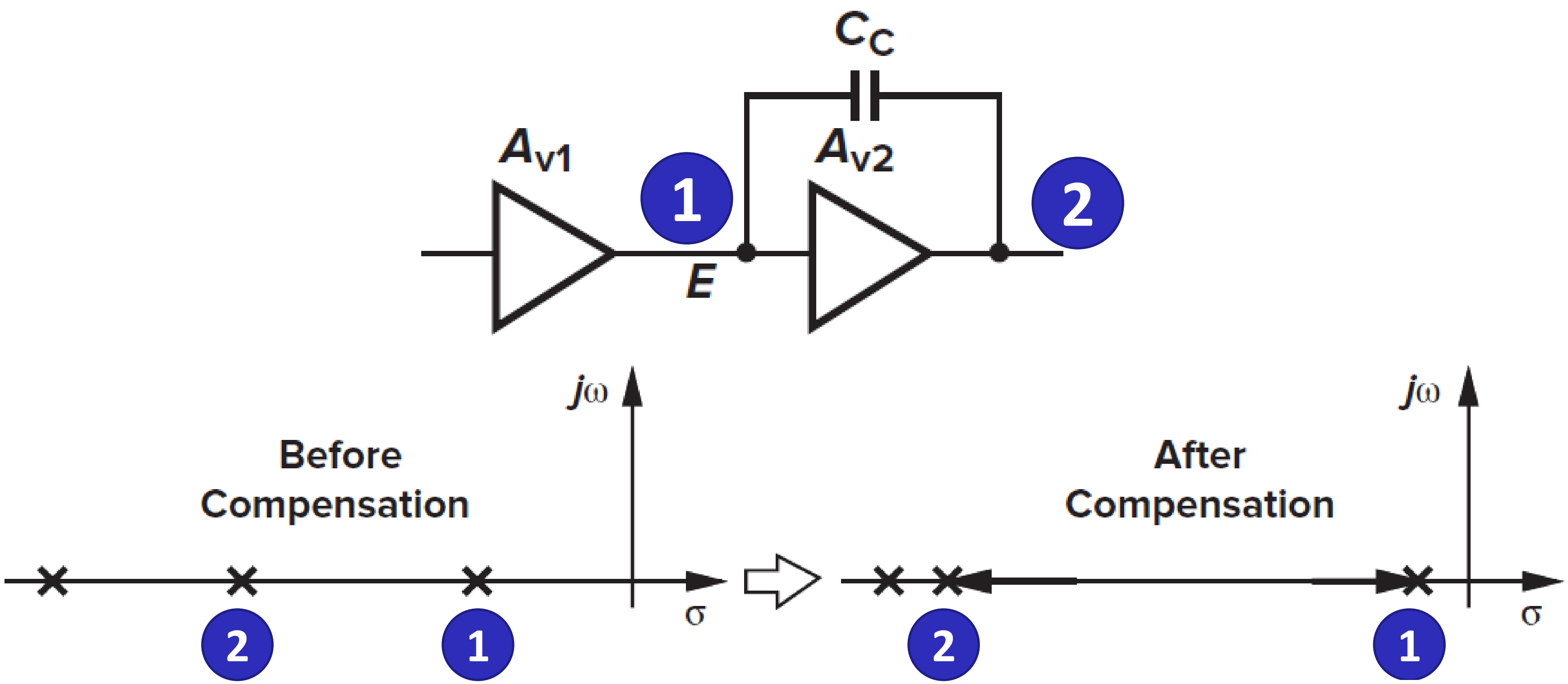
$$\omega_{p1} = \frac{1}{R_{out1} C_1} \quad \& \quad \omega_{p2} = \frac{1}{R_{out2} C_2}$$




[Razavi, 2017]

# Two-Stage OTA: Miller Compensation

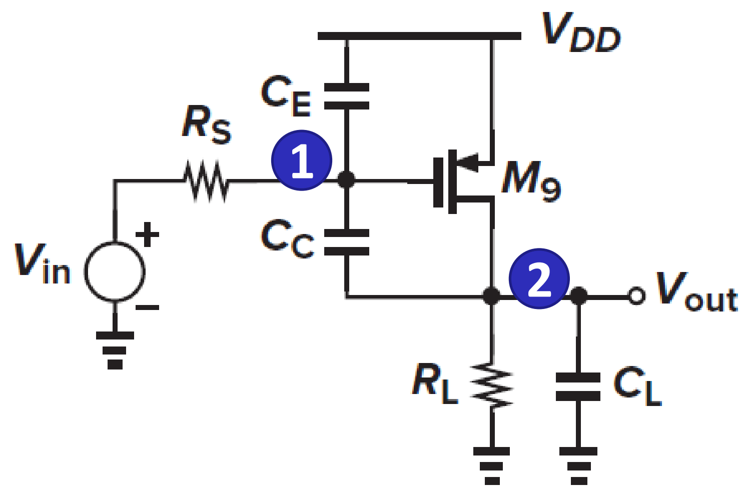
- Exploit Miller capacitance multiplication
- Pole splitting
  - Push the pole @ 1 inwards
  - Push the pole @ 2 outwards

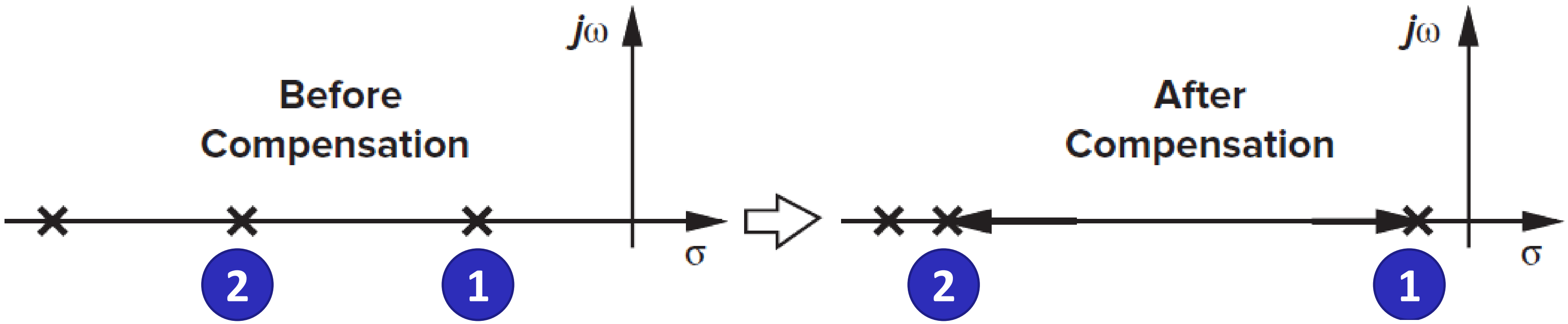



[Razavi, 2017]

# Two-Stage OTA: Miller Compensation

- Exploit Miller capacitance multiplication
- Pole splitting




[Razavi, 2017]

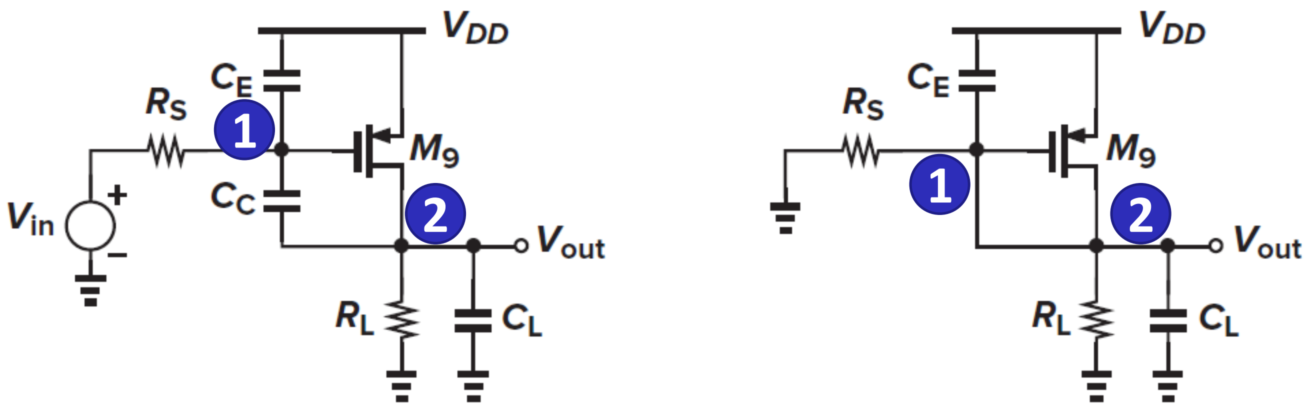
# Two-Stage OTA: Miller Compensation

- Before compensation

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

- After compensation

$$\omega_{p1} \approx \frac{1}{R_{out1}[(G_{m2}R_{out2})C_C + C_1]} \quad \& \quad \omega_{p2} \approx \frac{G_{m2}}{C_1 + C_2}$$



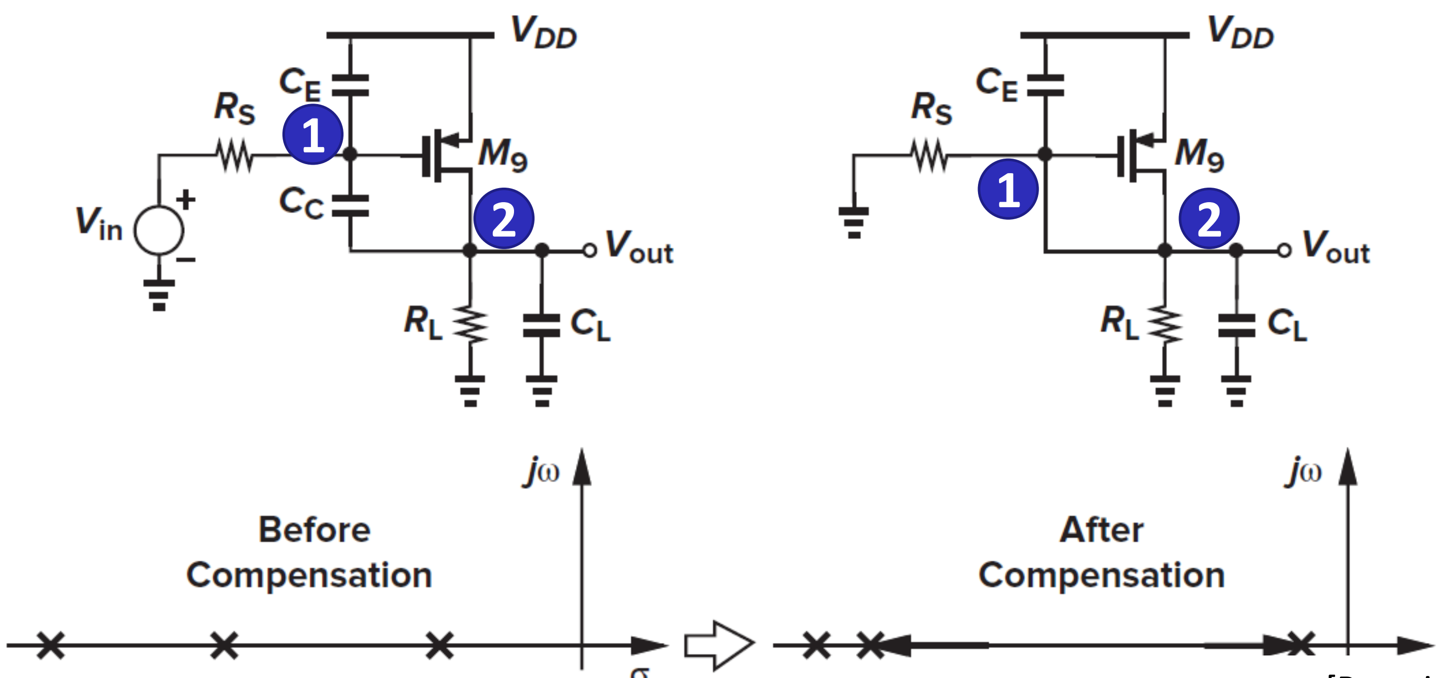

[Razavi, 2017]

# Two-Stage OTA: Miller Compensation

- After compensation: more accurate expressions

$$\omega'_{p1} \approx \frac{1}{R_S[(1+g_{m9}R_L)(C_C+C_{GD9})+C_E]+R_L(C_C+C_{GD9}+C_L)}$$

$$\omega'_{p2} \approx \frac{R_S[(1+g_{m9}R_L)(C_C+C_{GD9})+C_E]+R_L(C_C+C_{GD9}+C_L)}{R_S R_L[(C_C+C_{GD9})C_E+(C_C+C_{GD9})C_L+C_E C_L)]}$$




[Razavi, 2017]

# Two-Stage Miller OTA

- After compensation

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

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

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

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

# Two-Stage Miller OTA

$$\omega_{p2} \approx \frac{G_{m2}}{C_L}$$

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

- For $PM \approx 70^o$

$$\omega_{p2} \approx 3\omega_u$$

- Take additional margin to account for parasitic capacitors

$$\omega_{p2} \approx 4\omega_u$$

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

# Two-Stage Miller OTA

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

- Assume $C_L = 5pF \; and \; C_{Cmax} = 2pF$

$$\frac{G_{m2}}{G_{m1}} = 4 \times \frac{5}{2} = 10$$

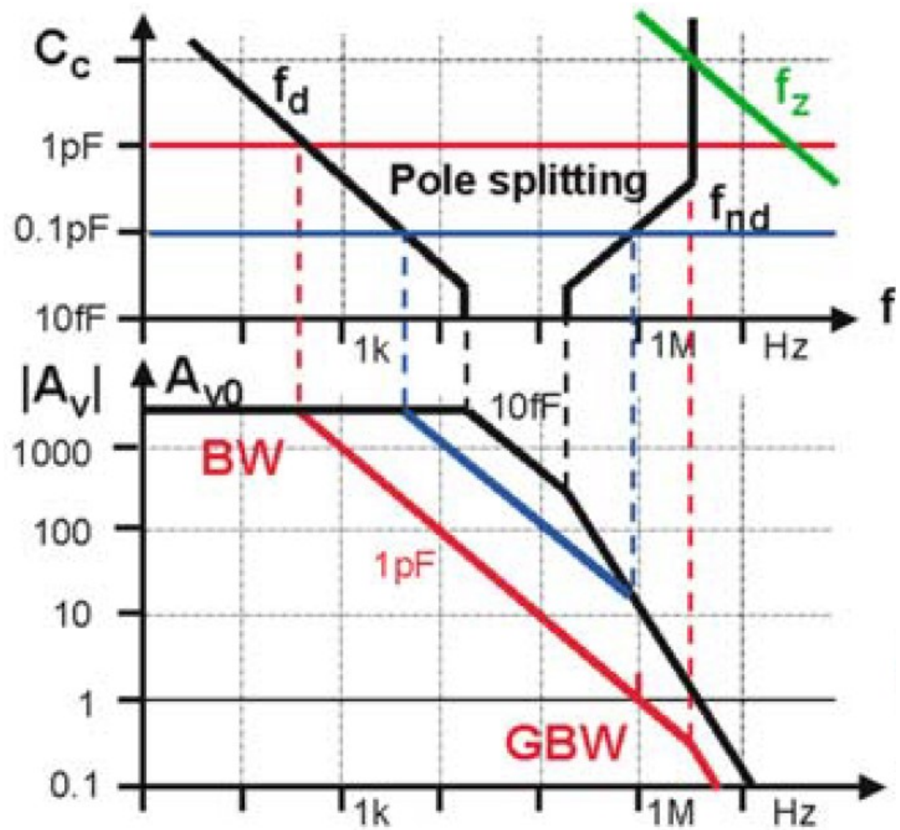
- If both stages use the same gm/ID

$$\frac{I_{D2}}{I_{D1}} = 10$$

- More than 80% of the power is consumed in the second stage to achieve stability
  - Miller OTA is very energy inefficient!

# Miller OTA: Pole Splitting with $C_C$

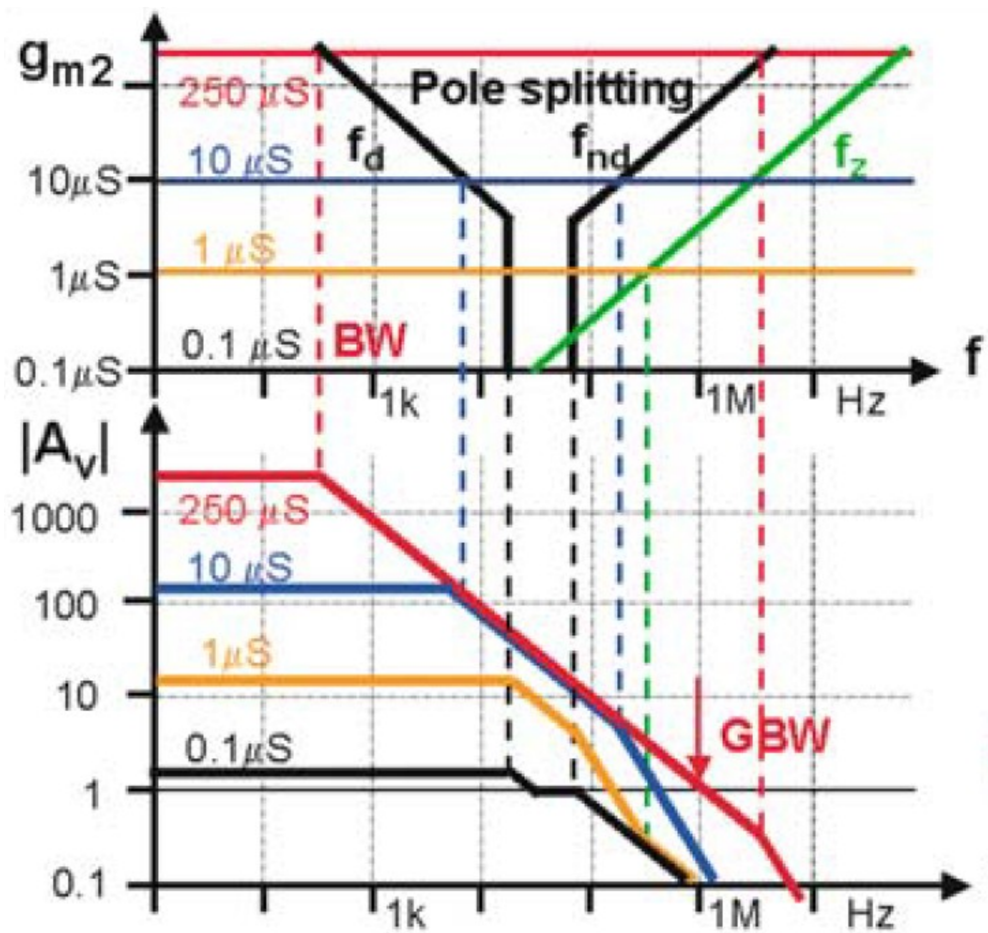
- Too large $C_C$ does not give more pole splitting: just smaller GBW
  - Usually we choose $C_1 < C_C < C_L$
  - Reasonable starting point: $C_C = (0.3 - 0.5) \times C_L$




[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



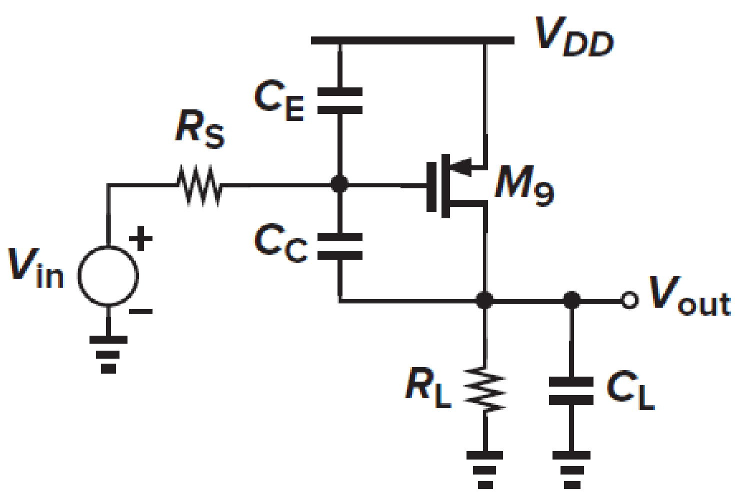


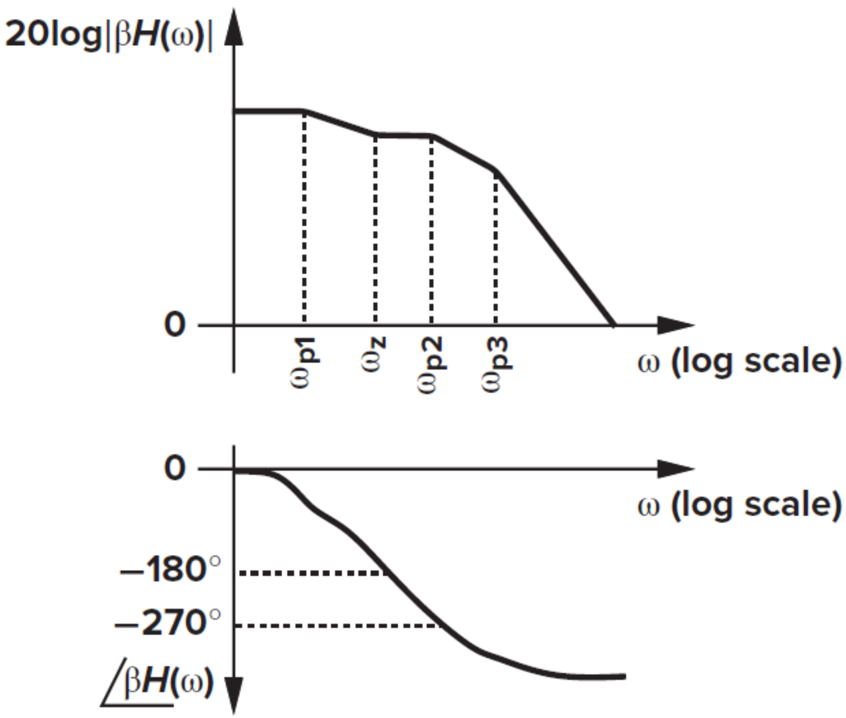

[Sansen, Essentials, 2006]

# Two-Stage Miller OTA: RHP Zero

- RHP zero is bad for both magnitude and phase
  - Increases GX and decreases PX

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



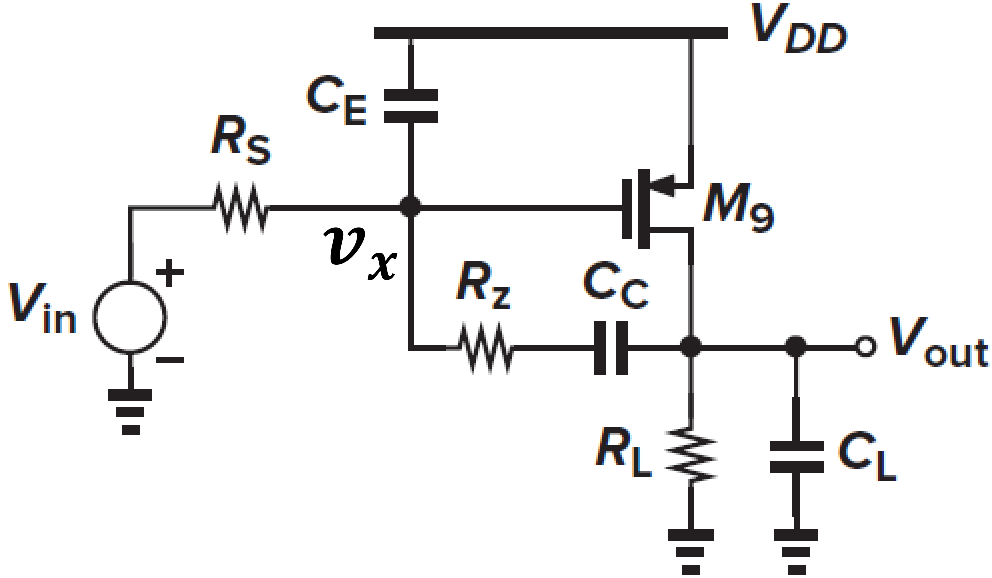

[Razavi, 2017]

# Two-Stage Miller OTA: RHP Zero

- Add a resistance to control the value of the zero → Move it to LHP

$$\frac{v_x}{R_Z + \frac{1}{s_z C_C}} = g_m v_x$$

$$s_z = \frac{1}{C_C\left(\frac{1}{g_m} - R_Z\right)}$$






[Razavi, 2017]

# Two-Stage Miller OTA: LHP Zero Placement

- Can we cancel the first non-dominant pole with the LHP zero?
  - Theoretically yes

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

  - But practically impossible due to variations
  - The case of $\omega_z$ in the vicinity of $\omega_u$ will cause several disadvantages
    - GX at high frequency in the vicinity of many non-dominant poles
    - Noise amplification
    - **Do NOT do it**
  - Warning: Problems 10.19 and 10.20 in [Razavi, 2017] assumes that $A_o\omega_{p1}=\omega_u \gg \omega_{p2}$, which is incorrect

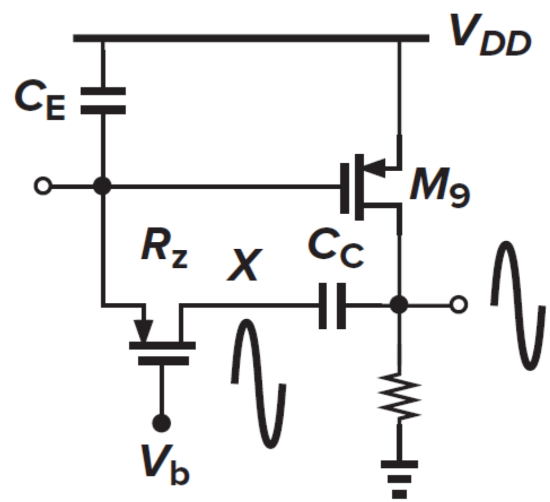
# Two-Stage Miller OTA: LHP Zero Placement

- Of course we cannot cancel the dominant pole as well
  - Will cause poor settling behavior (pole-zero doublet)
  - A component of the output proportional to the mismatch between $\omega_{p1}$ and $\omega_z$ will settle slowly with $\tau = 1/\omega_{p1}$ instead of $\tau = 1/\omega_u$
  - See Problem 10.19 in [Razavi, 2017] (note that the assumption that $A_o\omega_{p1} = \omega_u \gg \omega_{p2}$ is incorrect).

- **Conclusion: The LHP zero should be placed between the 1st and 2nd non-dominant poles**
  - **Make sure $\omega_z > 2\omega_u$ under all conditions**

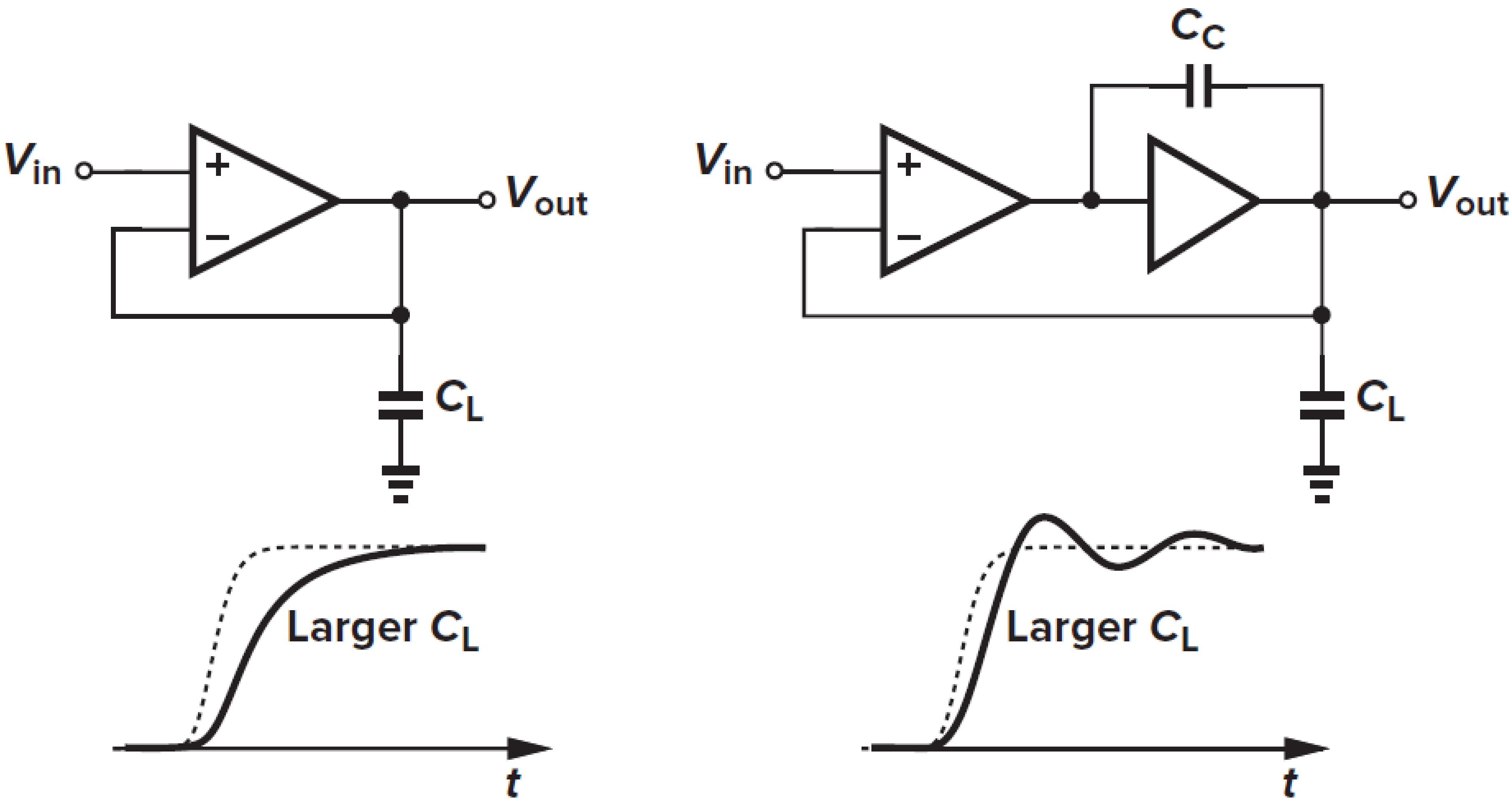
# Two-Stage Miller OTA: $R_Z$ Implementation

- $R_Z$ can be implemented using a transistor in triode




[Razavi, 2017]

# Single vs Two-Stage OTA: Sensitivity to $C_L$

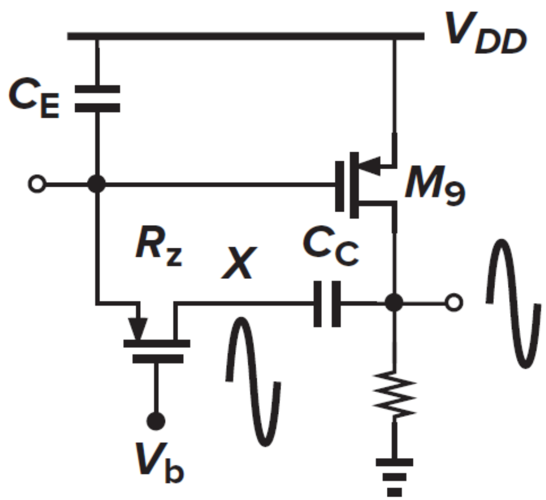


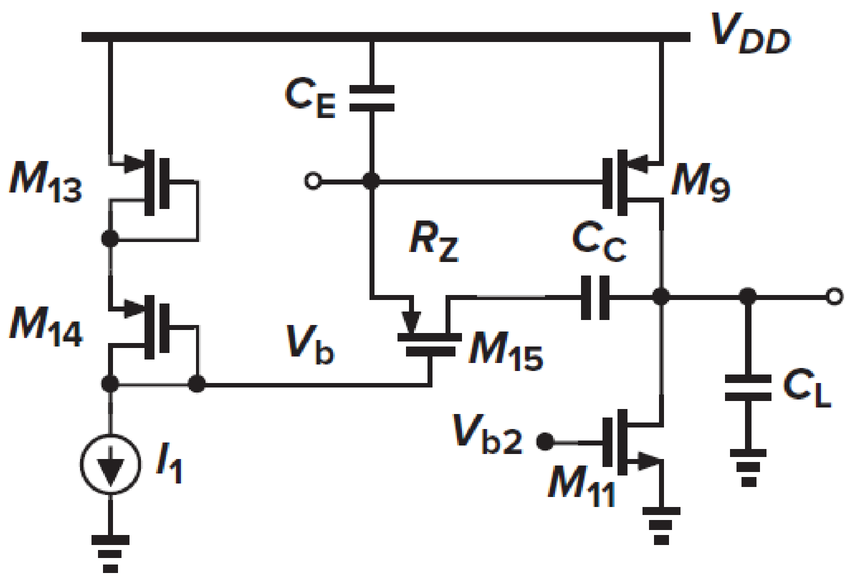

[Razavi, 2017]

# Thank you!

# Two-Stage Miller OTA: $R_Z$ Implementation

- $R_Z$ can be implemented using a transistor in triode
- Generation of Vb for proper temperature and process tracking.
  - The principal drawback of this method is that it assumes square-law characteristics for all of the transistors.



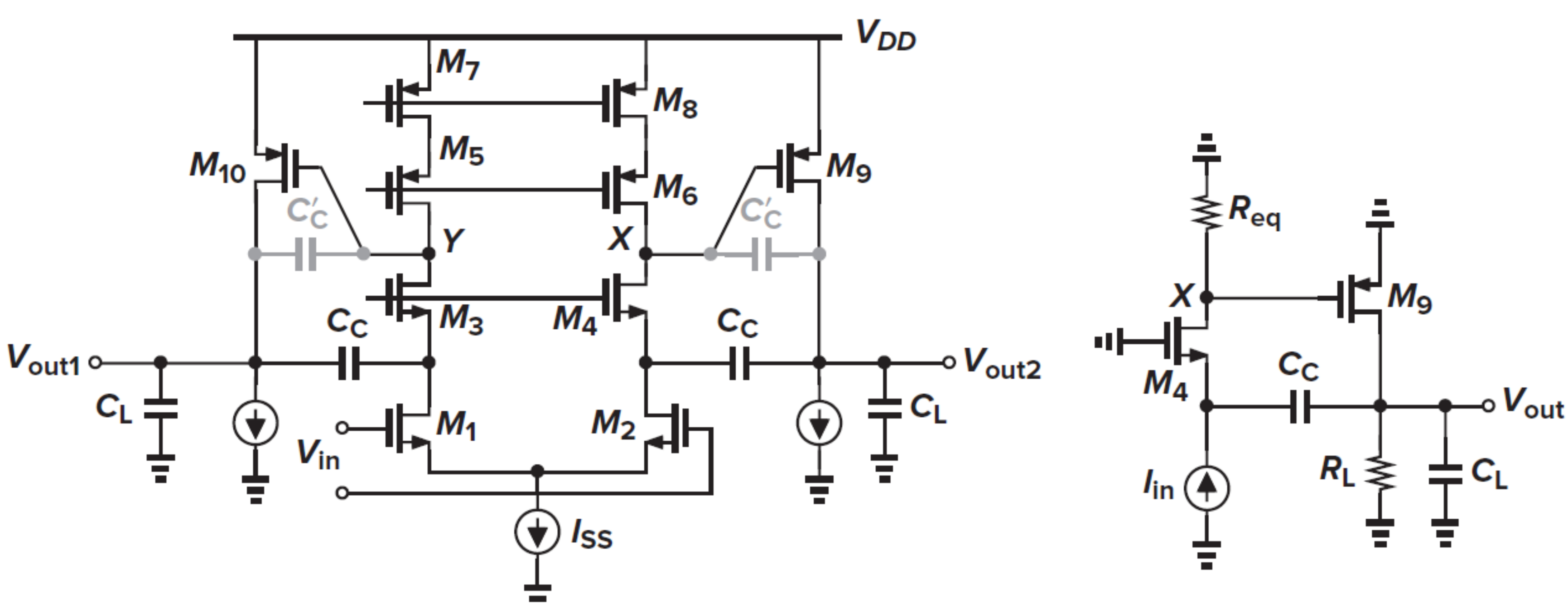






[Razavi, 2017]

# Two-Stage Miller OTA: Compensation

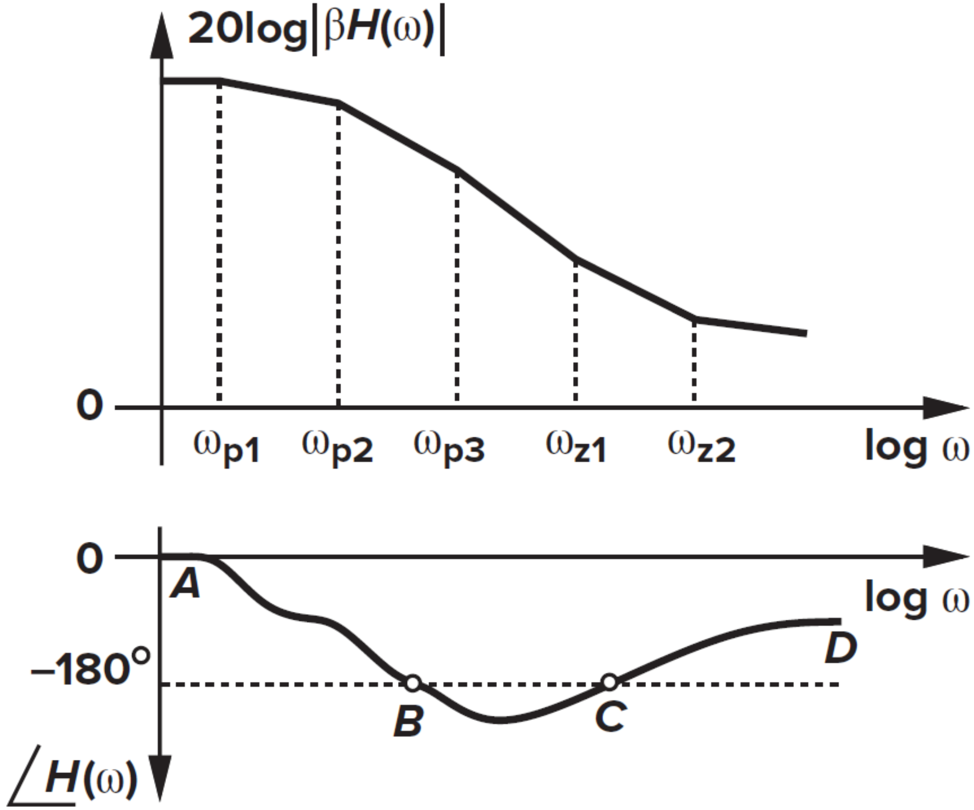
- Other compensation techniques exist to avoid the RHP zero
  - The common idea is to cancel the feedforward path due to $C_C$
  - Example 1: CC/CG stage inserted in series with $C_C$
  - Example 2: $C_C$ placed between the source of the cascode devices and the output nodes




[Razavi, 2017]

# Systems with Multiple 180° Crossings

- If $\angle\beta H$ crosses 180° an even (odd) number of times while $|\beta H| > 1$, then the system is stable (unstable)




[Razavi, 2017]