وَمَا أُوتِيتُمْ مِنَ الْعِلْمِ إِلَّا قَلِيلًا

Analog IC Design

Lecture 15
OTA / Op-Amp Topologies

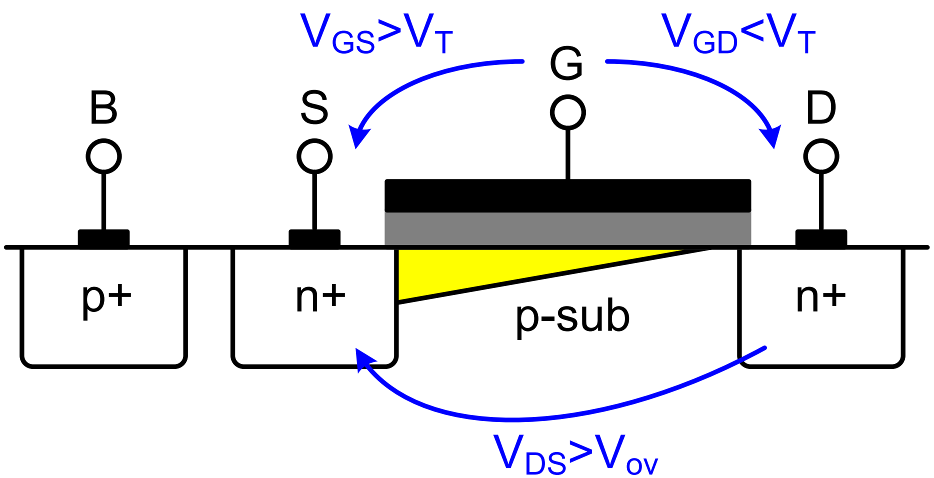
Dr. Hesham A. Omran

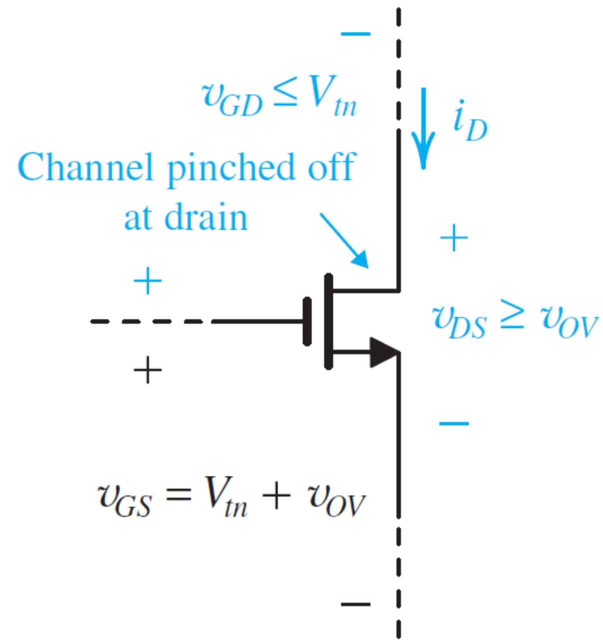
Integrated Circuits Lab (ICL)
Electronics and Communications Eng. Dept.
Faculty of Engineering
Ain Shams University

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

$V_{GD} \leq V_T$

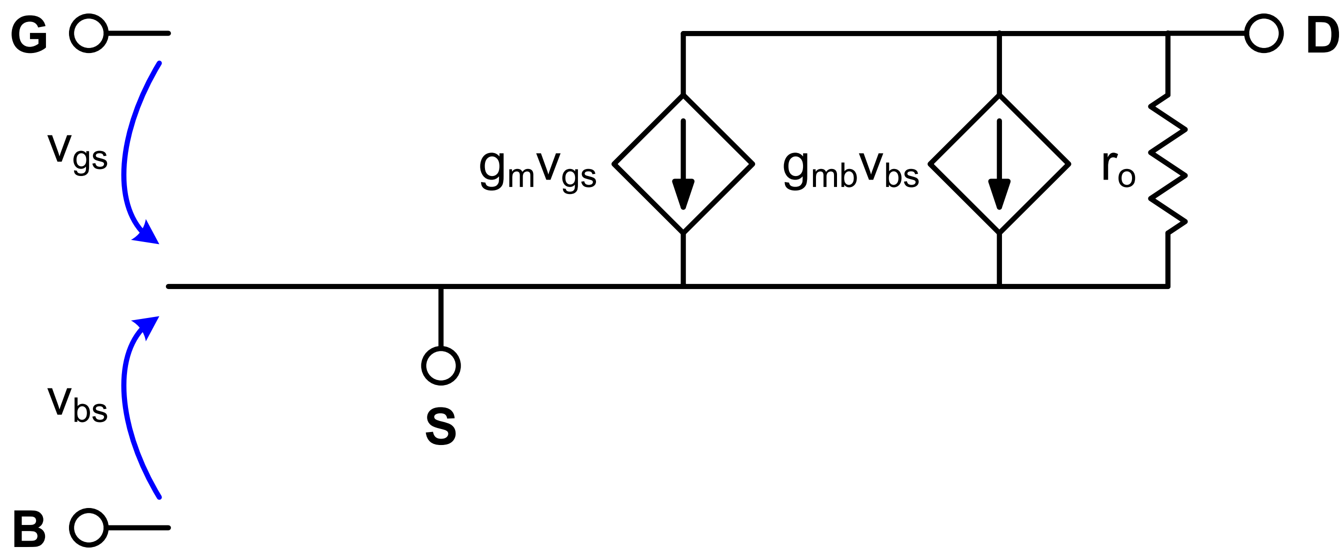
$$I_D = \frac{\mu C_{ox}}{2}\frac{W}{L}V_{ov}^2(1 + \lambda V_{DS})$$

Low-Frequency Small-Signal Model

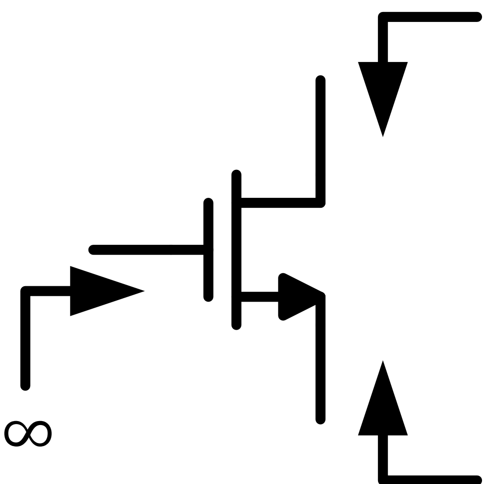
$$g_m = \frac{\partial I_D}{\partial V_{GS}} = \mu C_{ox}\frac{W}{L}V_{ov} = \sqrt{\mu C_{ox}\frac{W}{L}\cdot 2I_D} = \frac{2I_D}{V_{ov}}$$

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

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



Rin/out Shortcuts Summary



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

H.I.N.

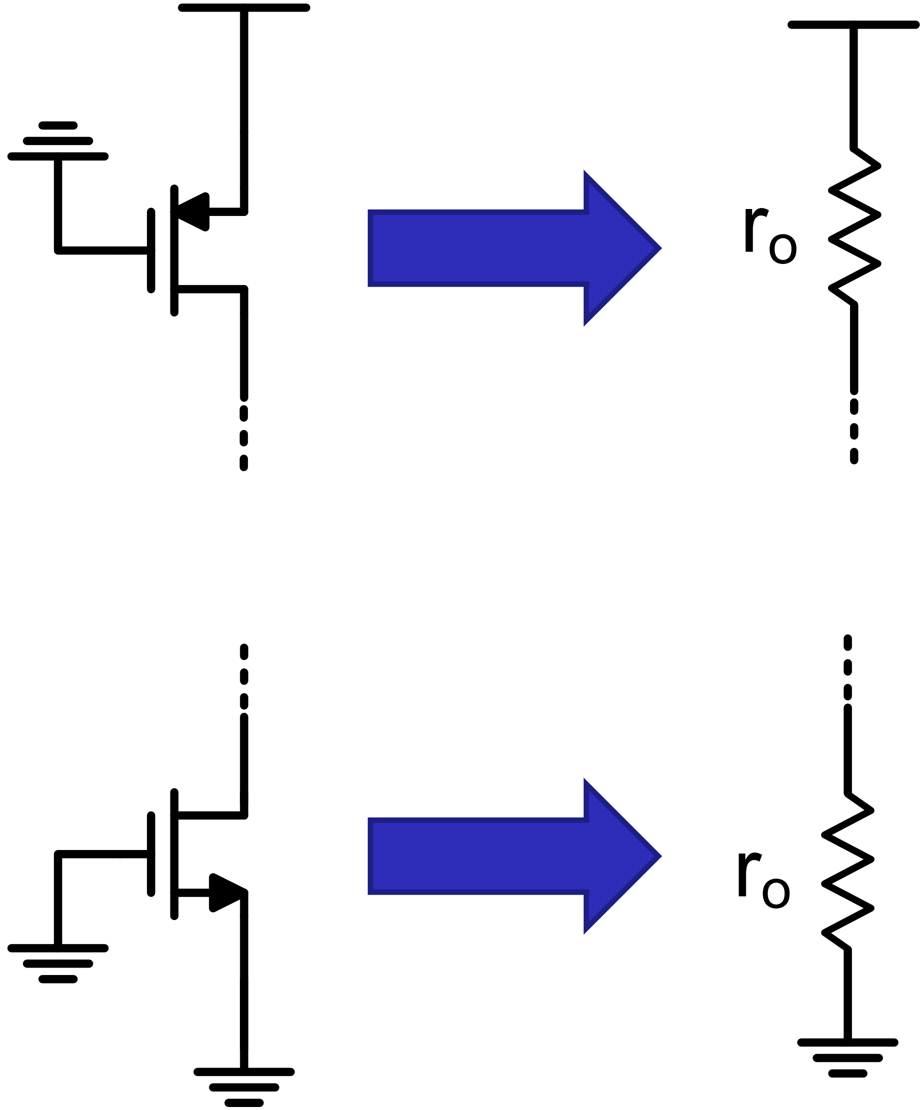
∞

At low frequencies **ONLY**

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

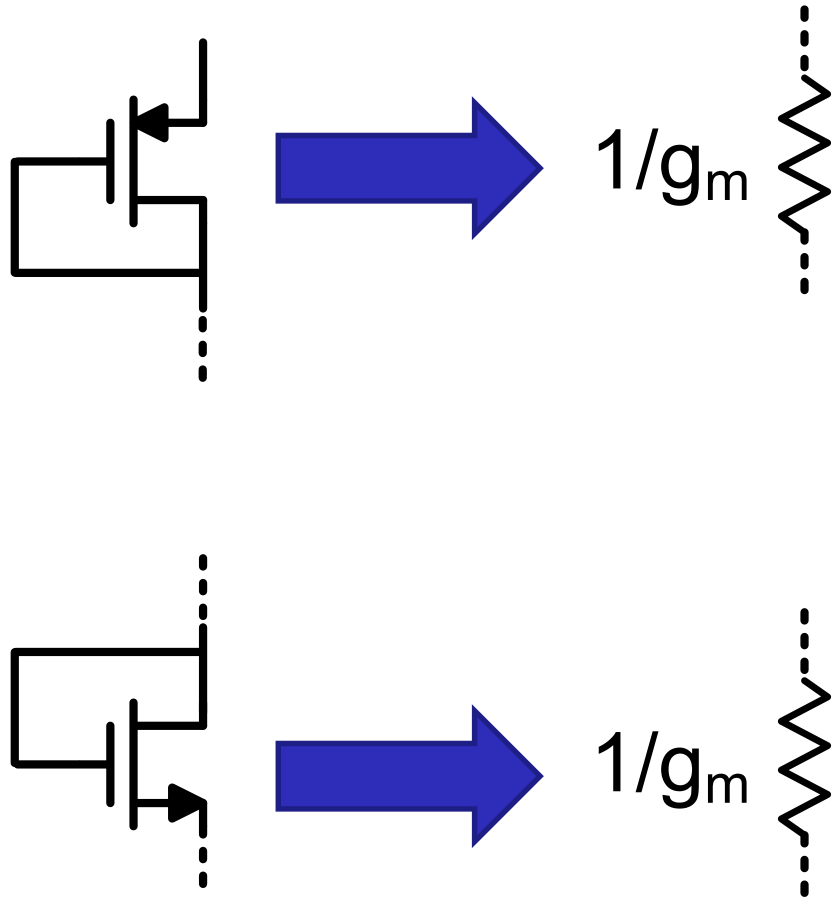
L.I.N.

Active Load (Source OFF)



Diode Connected (Source Absorption)

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



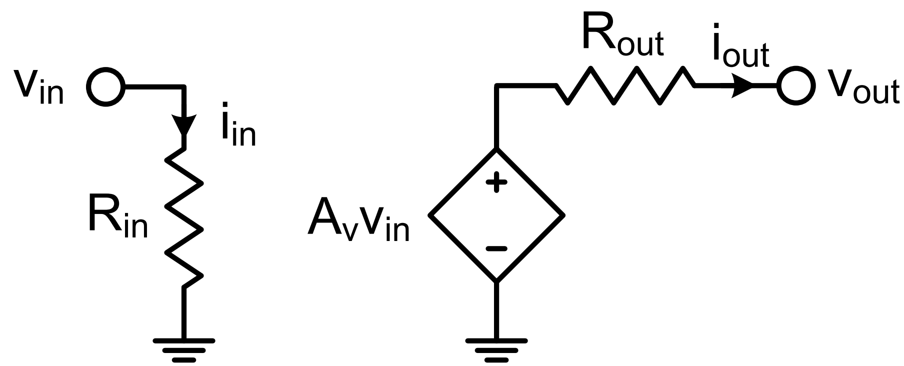
Why GmRout?

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

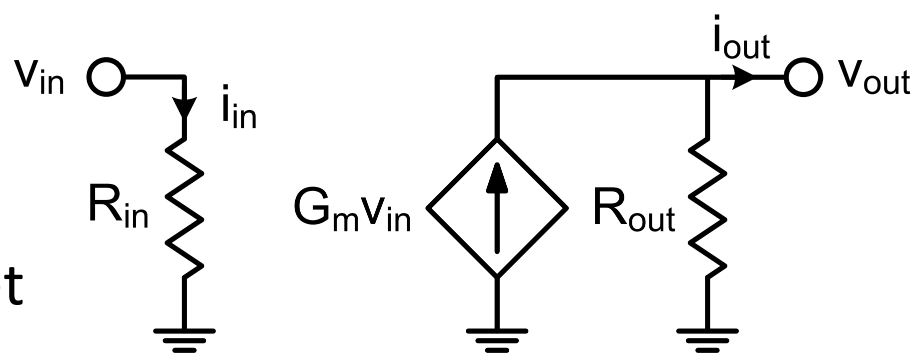
$$G_m = \frac{i_{out,sc}}{v_{in}}$$

$$A_v = G_m R_{out}$$

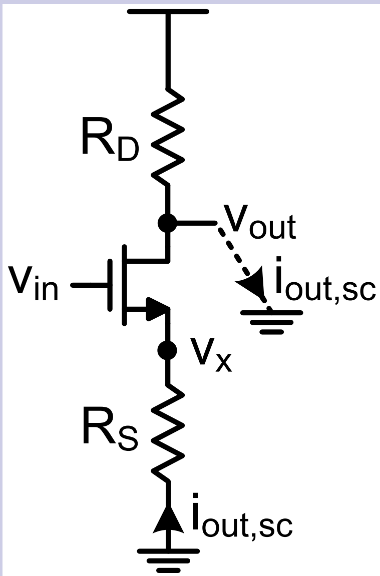
$$A_i = G_m R_{in}$$

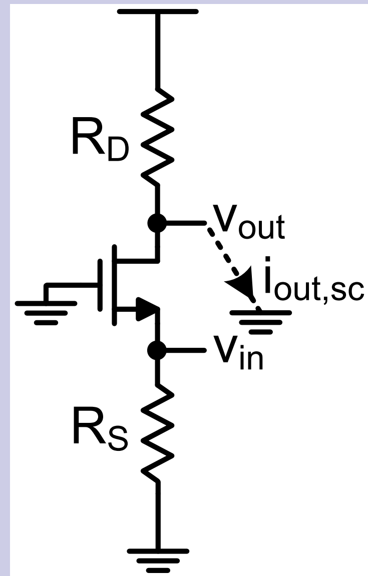


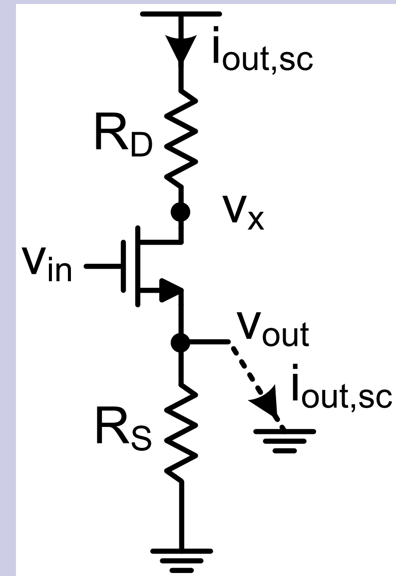
- Divide and conquer
  - Rout simplified: vin=0
  - Gm simplified: vout=0
  - We already need Rin/out
  - We can quickly and easily get Rin/out from the shortcuts



Summary of Basic Topologies







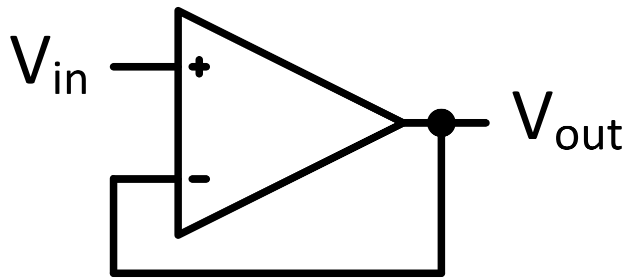
	CS	CG	CD (SF)
	Voltage & current amplifier	Current buffer	Voltage buffer
Rin	∞	$R_S // \frac{1}{g_m + g_{mb}}\left(1 + \frac{R_D}{r_o}\right)$	∞
Rout	$R_D // r_o[1 + (g_m + g_{mb})R_S]$	$R_D // r_o$	$R_S // \frac{1}{g_m + g_{mb}}\left(1 + \frac{R_D}{r_o}\right)$
Gm	$\frac{-g_m}{1 + (g_m + g_{mb})R_S}$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

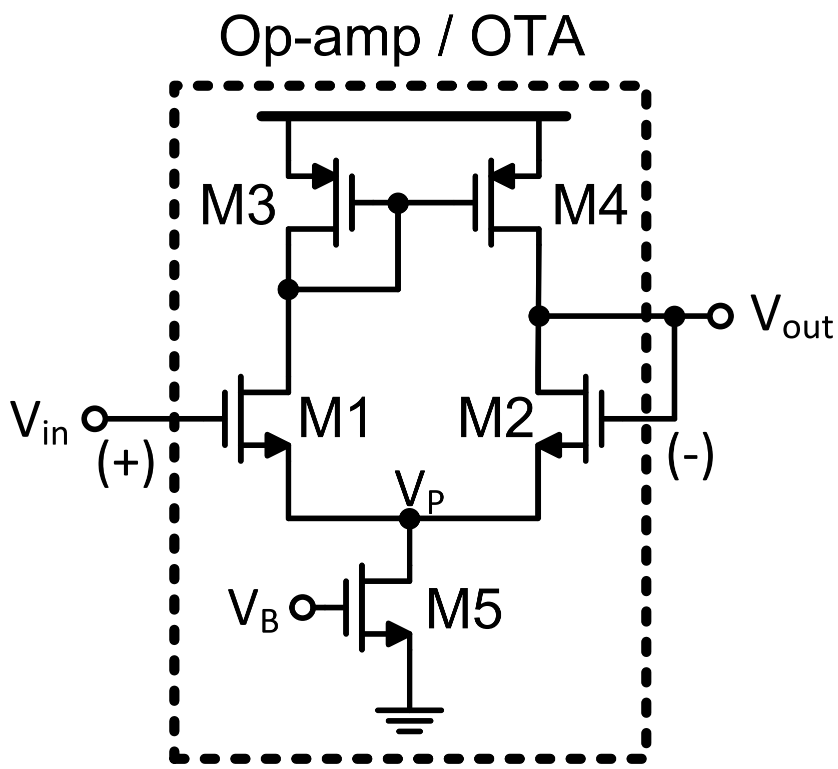
Differential Amplifier

	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	$\dfrac{-g_m R_D}{1 + 2(g_m + g_{mb})R_{SS}}$
A_{vd}/A_{vCM}	1	∞	$2(g_m + g_{mb})R_{SS} \gg 1$

What is an OTA / Op-Amp?

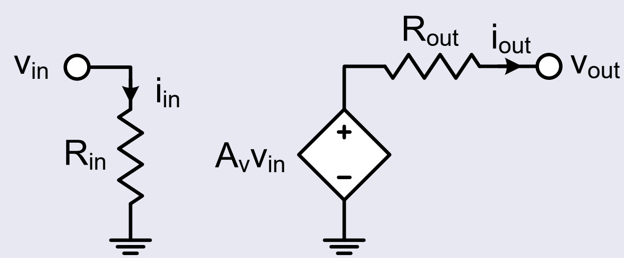
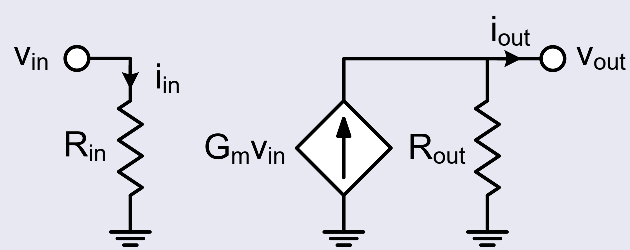
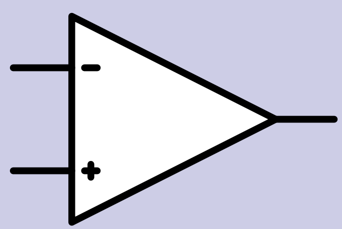
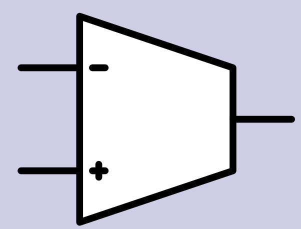
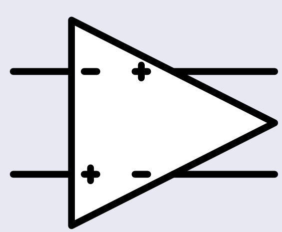
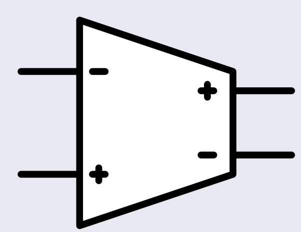
- An op-amp is simply a high gain differential amplifier
- The gain can be increased by using cascodes and multi-stage amplifiers





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

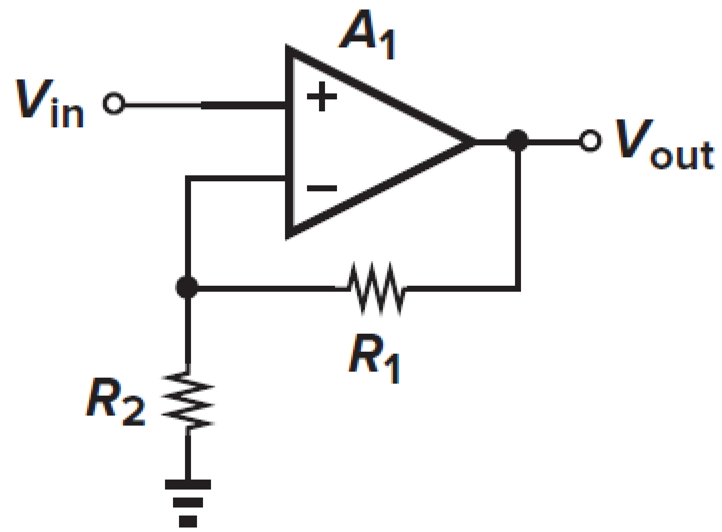
	Op-amp	OTA
Rout	**LOW**	**HIGH**
Model		
Diff input, SE output		
Fully diff		

OTA / Op-Amp

- Integral part of many analog and mixed-signal systems
  - DC bias generation
  - Amplification
  - Filtering
- Challenges
  - Supply voltage and channel length scaling
  - Energy efficiency

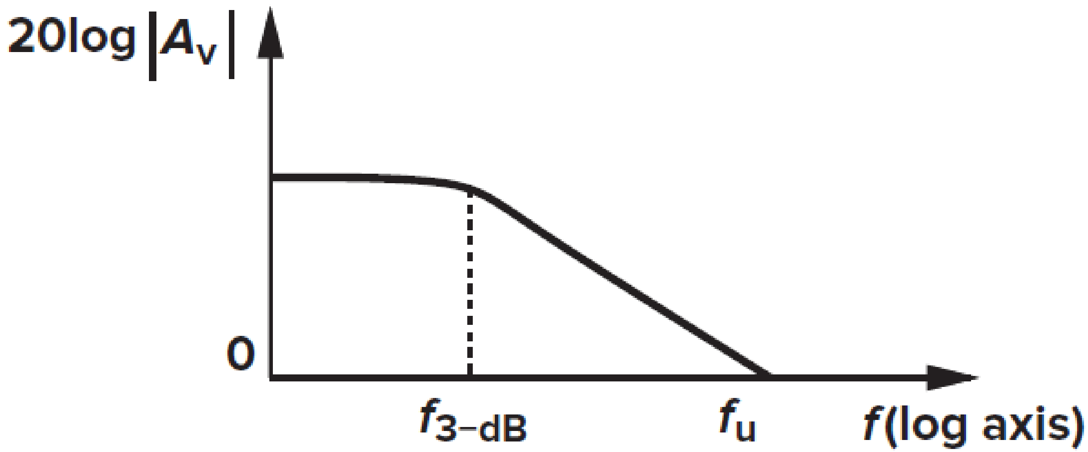
Finite Gain

- Gain determined by ratio of matched components
  - Very low sensitivity to PVT variations
- Example: A1 > 60 dB for error < 1%

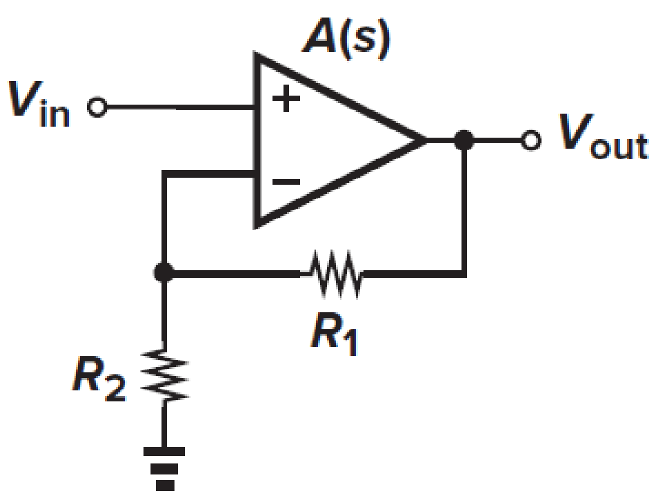


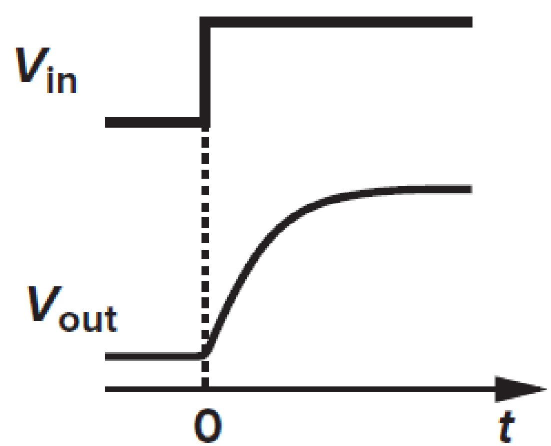
$$\frac{V_{out}}{V_{in}} \approx \left(1 + \frac{R_1}{R_2}\right)\left(1 - \frac{R_1 + R_2}{R_2}\frac{1}{A_1}\right)$$

Finite Small-Signal Bandwidth



$$\tau \approx \left(1 + \frac{R_1}{R_2}\right)\frac{1}{A_0\omega_0}$$



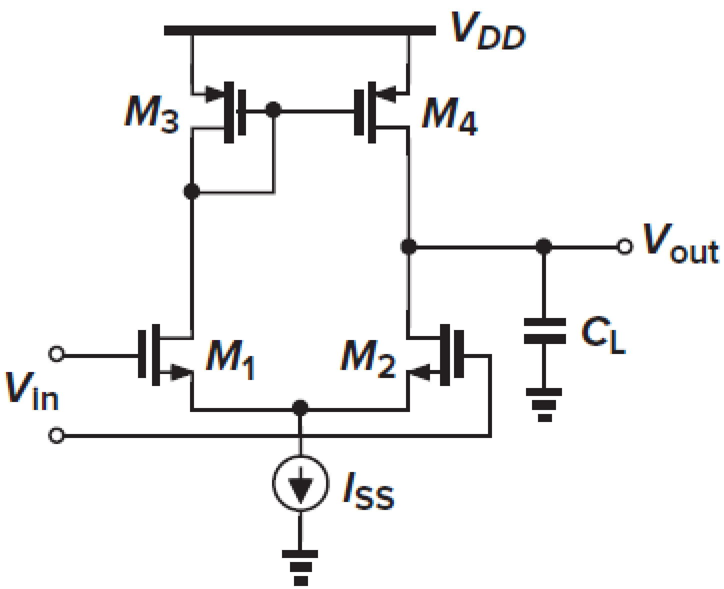


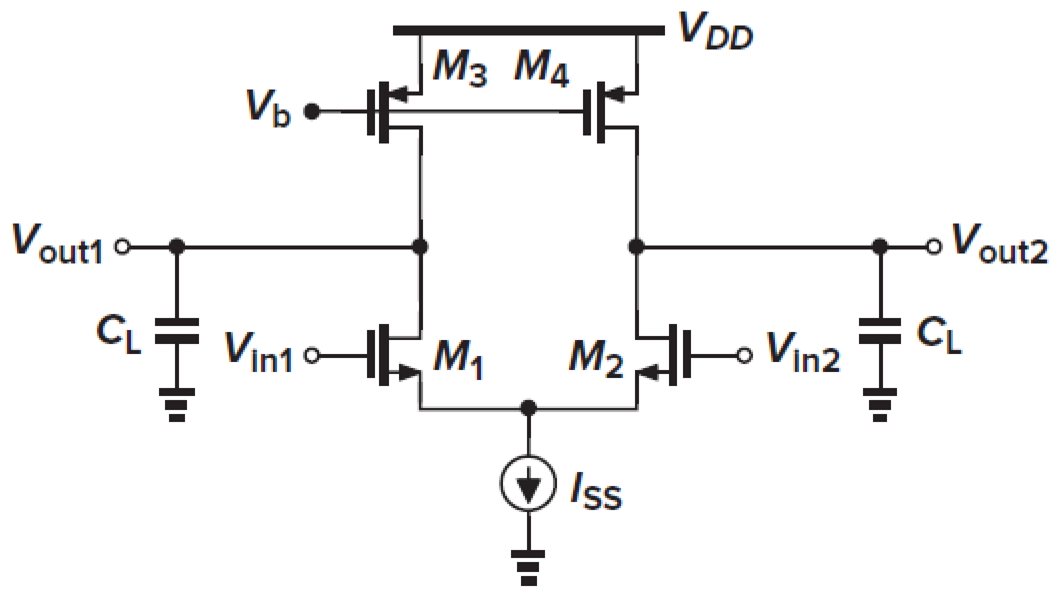
OTA Topologies

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

Simple Single-Stage OTAs

- Simple, but limited gain



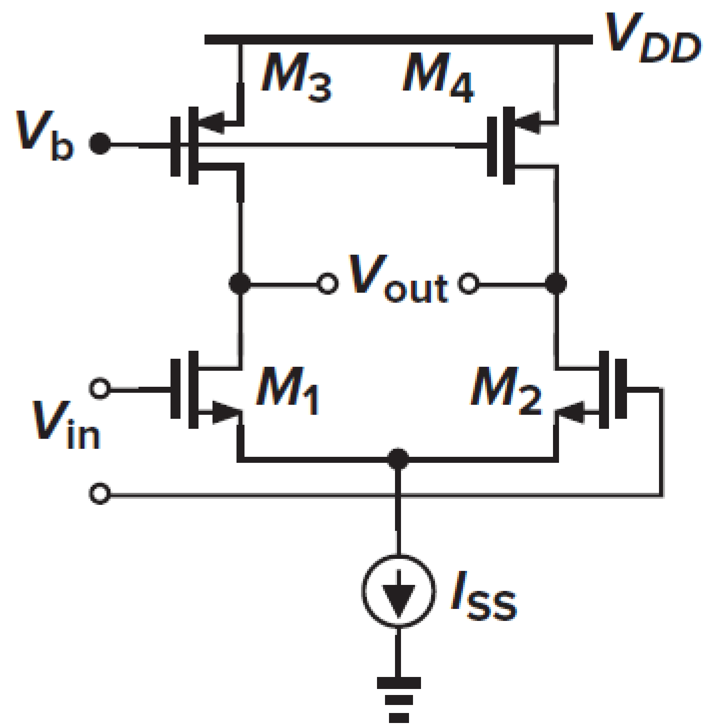


Simple Single-Stage OTA CMIR (OL)

- $V_{in,CM}$ Range

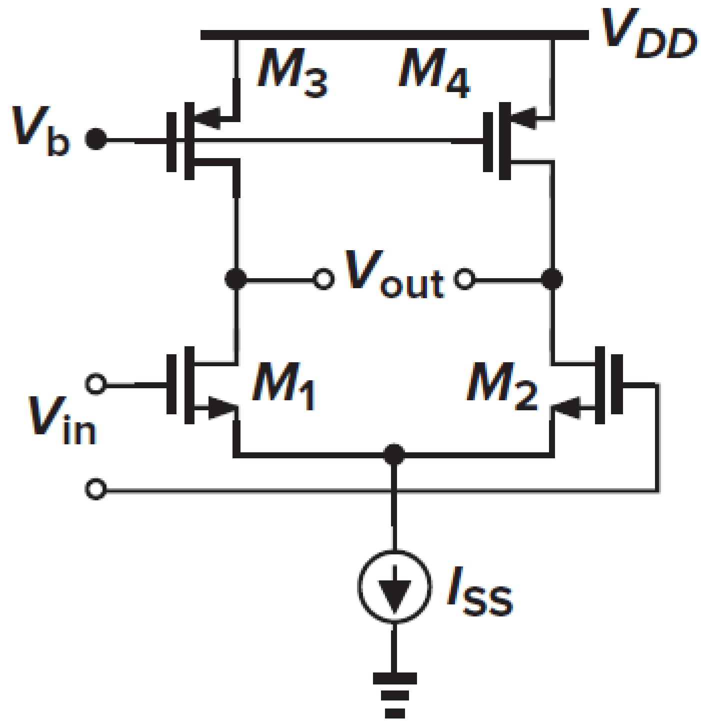
- V_{out} Range

- $V_{in,CM,max}$ and $V_{out,min}$ are coupled
- If max output swing is desired
  - M1 at edge of sat
  - No range for $V_{in,CM}$



Simple Single-Stage OTA Output Swing

- $Max\ Diff\ Swing \approx 2(V_{DD} - 3V_{ov})$

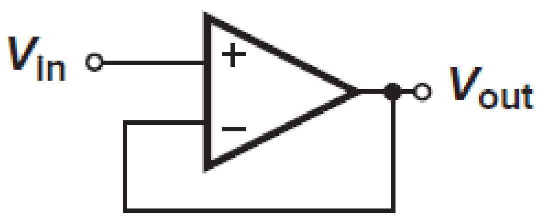


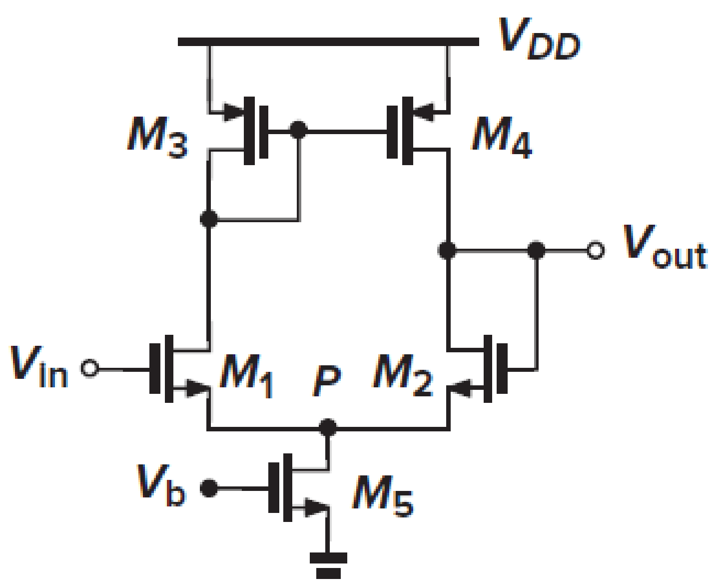
5T OTA as a Buffer

- V_{in} Range (OL, without FB)

- V_{out} Range (OL, without FB)

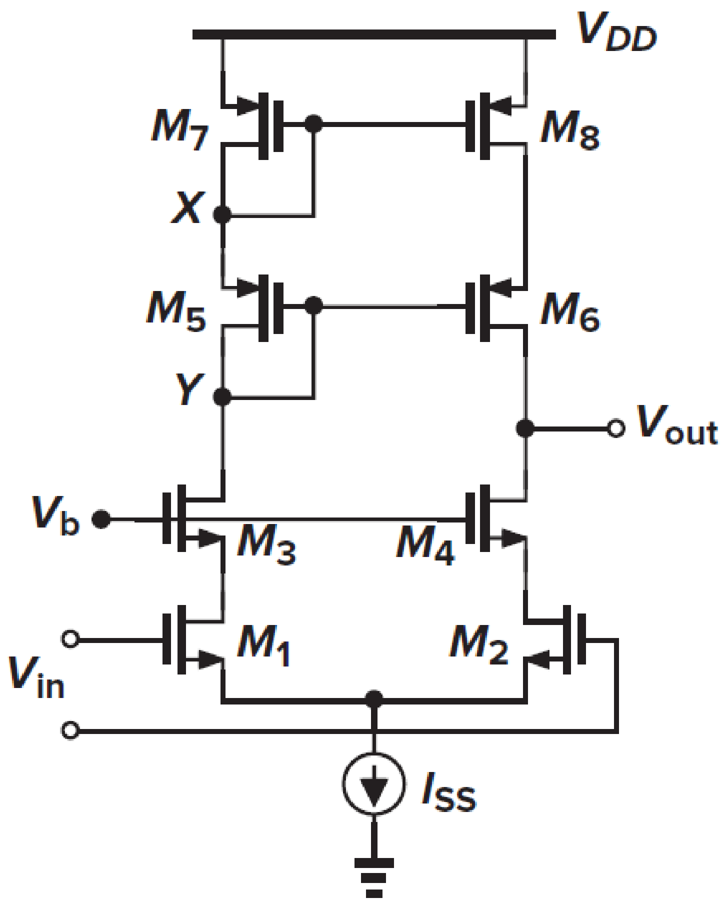
- Example: VDD = 1.2 V, VTH = 0.3 V, and Vov = 0.1 V
  - Vin (Vout) = 0.5 – 1.1 V → Max swing = 0.6 V

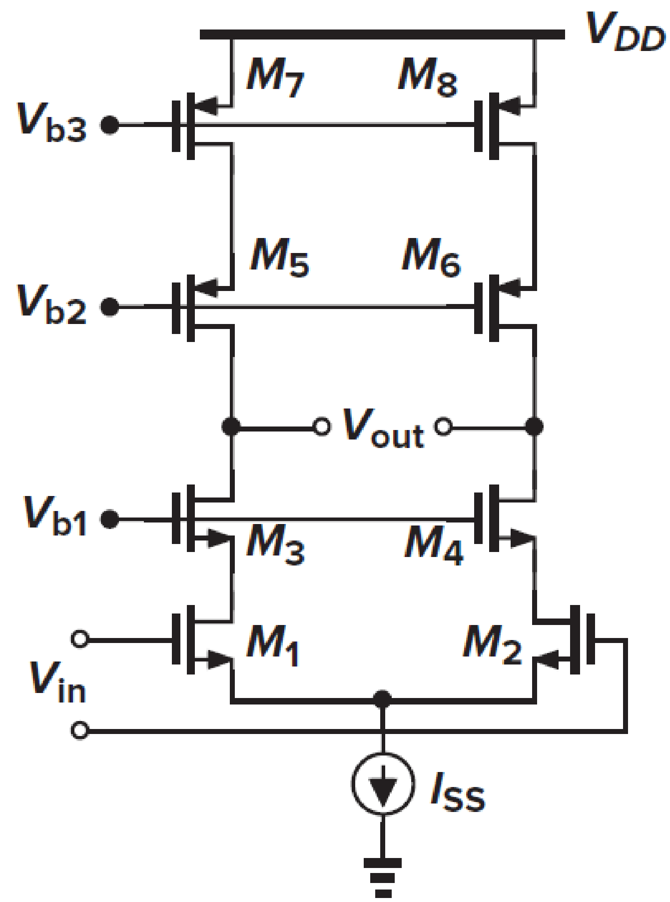




Telescopic Cascode

- Higher DC gain, but limited swing and additional poles



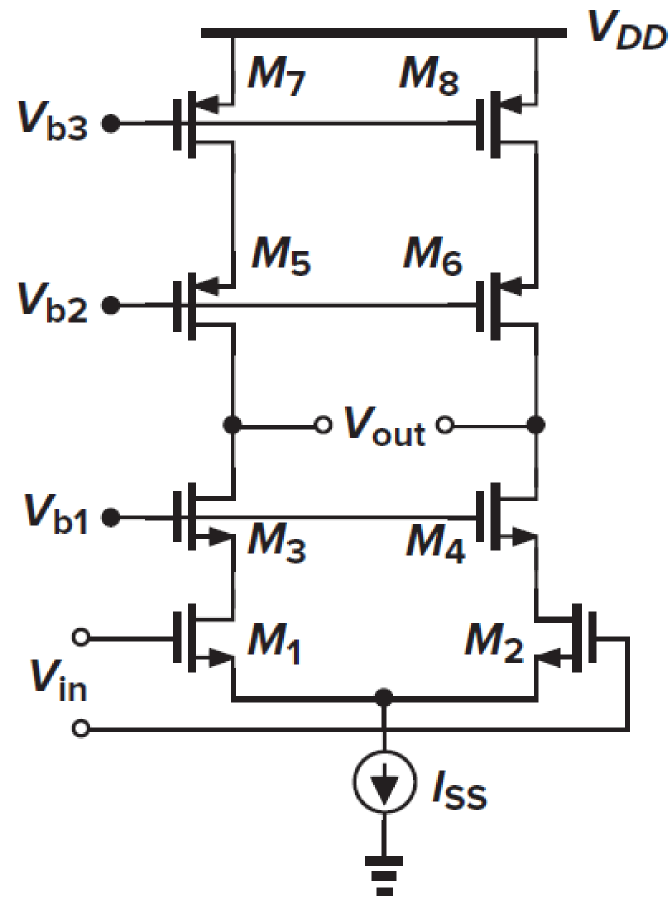


Telescopic Cascode CMIR (OL)

- $V_{in,CM}$ Range

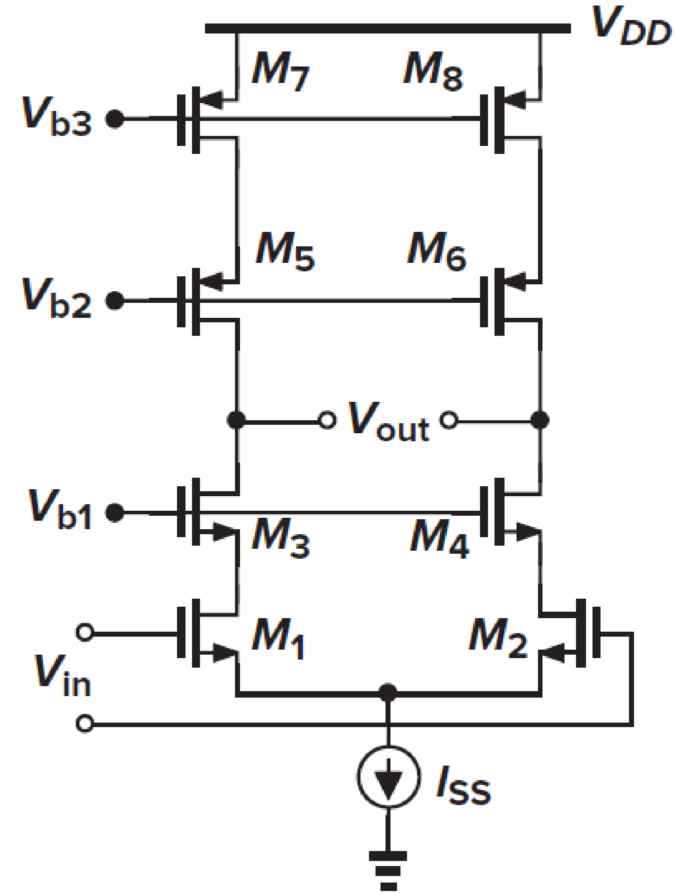
- V_{out} Range

- $V_{in,CM,max}$ and $V_{out,min}$ are coupled
- If max output swing is desired
  - M1 at edge of sat
  - No range for $V_{in,CM}$



Telescopic Cascode Output Swing (OL)

- $Max\ Diff\ Swing \approx 2(V_{DD} - 5V_{ov})$
- The choice of bias voltages is critical to maintain output swing
- $V_{b1} \geq V_{THN} + V_{ov3} + V_{ov1} + V_{ISS}$
- $V_{b2} \leq V_{DD} - V_{THP} - V_{ov5} - V_{ov7}$
- Place M1 and M7 just at the edge of sat for max output swing

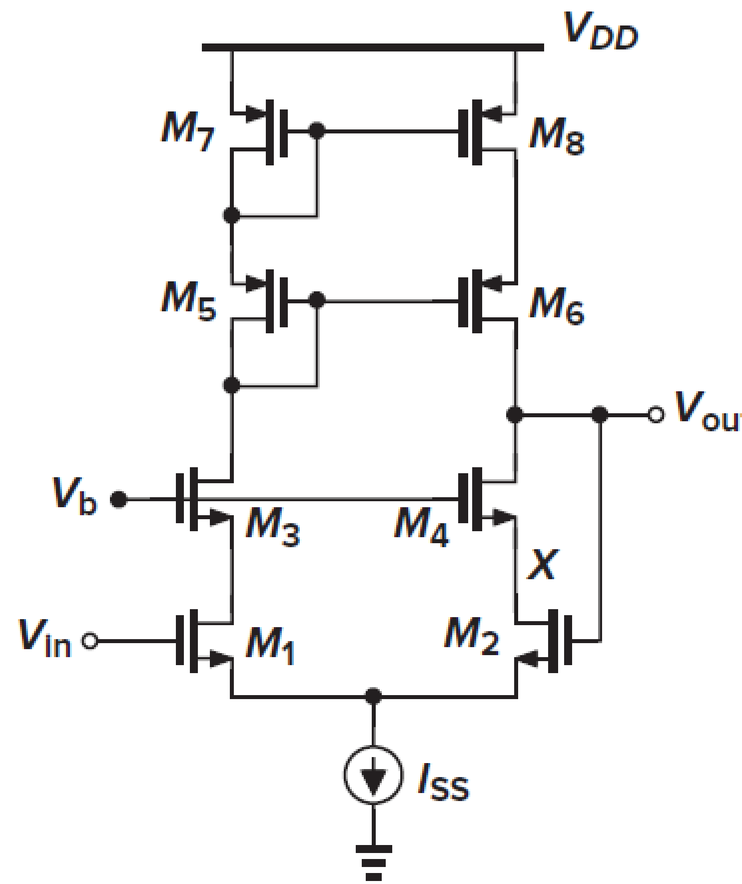


Telescopic Cascode as a Buffer

- ❑ V_{in} Range (OL, without FB)

- ❑ V_{out} Range (OL, without FB)

- ❑ $Max\ Swing = V_{THN} - V_{ov3}$
- ❑ Example: VDD = 1.2 V, VTH = 0.3 V, and Vov = 0.1 V
  - Vin (Vout) = (Vb – 0.3) – (Vb – 0.1) V
  - Max swing = 0.2 V
  - **Independent of VDD!**

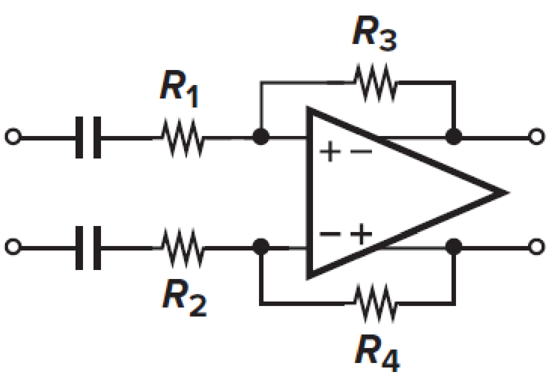


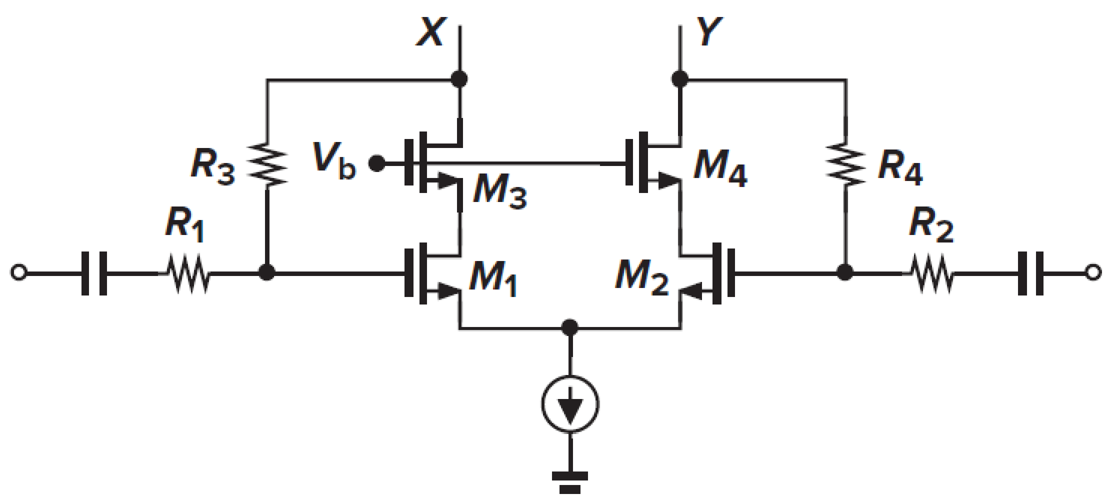
Telescopic Cascode CM Range (CL)

- Input and output CM levels are equal (why?) → similar to buffer
- $V_{in,CM}$ Range

- $V_{out,CM}$ Range

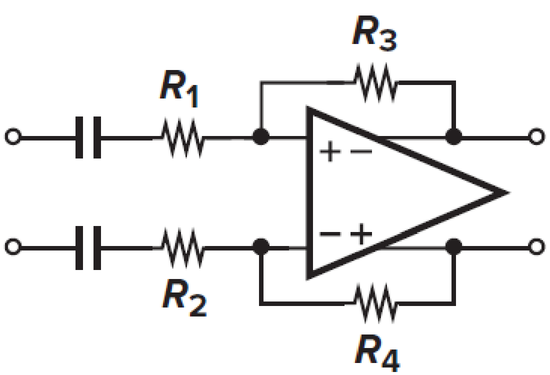
- $Max\ CM\ Range = V_{THN} - V_{ov3}$
- Example: VDD = 1.2 V, VTH = 0.3 V, and Vov = 0.1 V
  - VCM = (Vb – 0.3) – (Vb – 0.1) V → Max CM range = 0.2 V

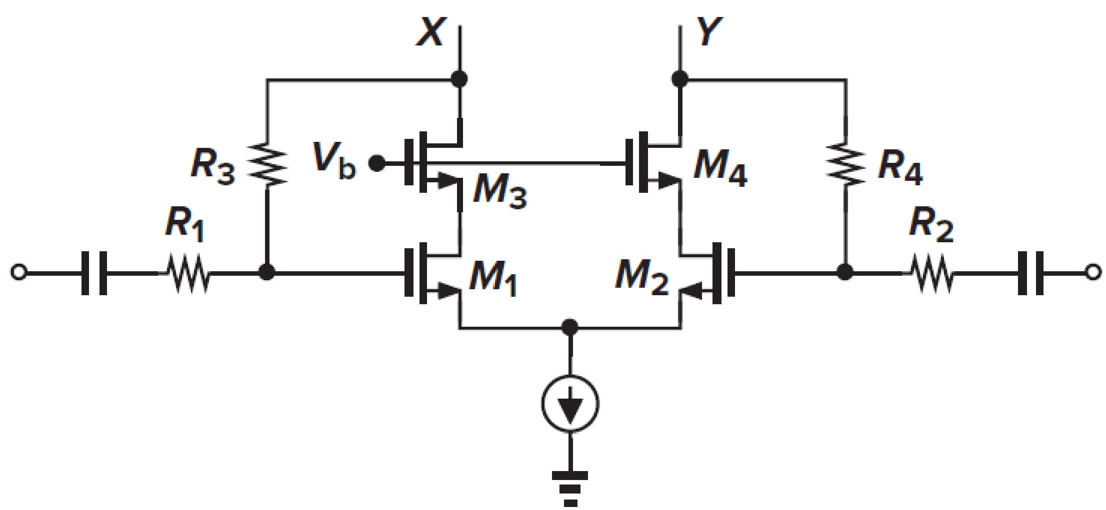




Telescopic Cascode Output Swing (CL)

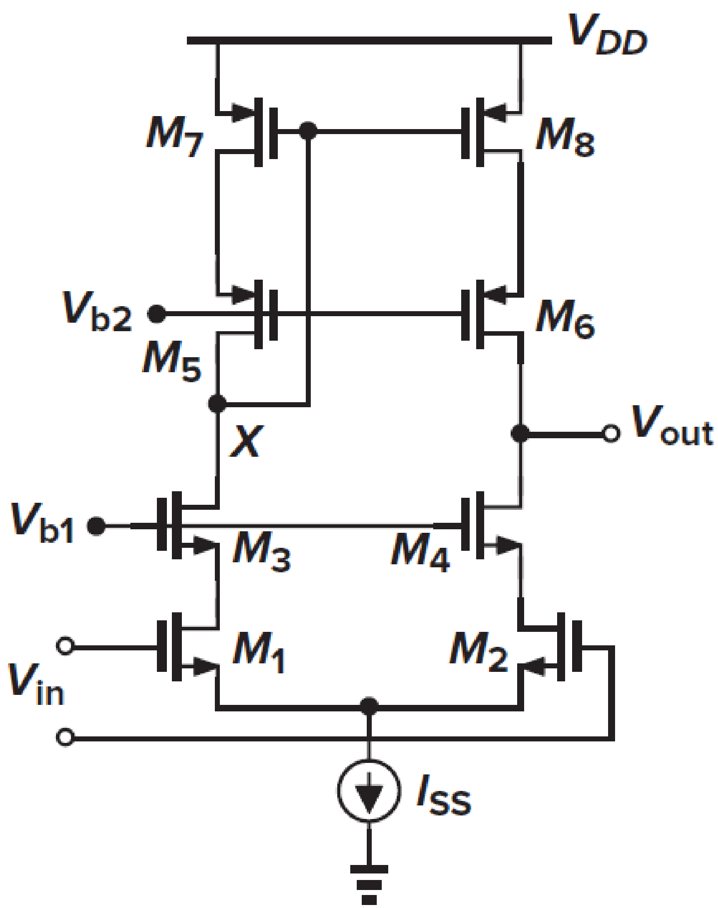
- Assume swing at input is negligible (high open-loop gain)
  - We care more about keeping M3 and M4 in sat
    - $V_{out,min} \geq V_b - V_{THN}$
  - We can place M1 and M2 at the edge
    - Set CM level at its max value: $V_{CM} = V_b - V_{ov3}$
- $Max\ Diff\ Swing = 2 \times 2 \times (V_{THN} - V_{ov3})$
- Ex: VTH = 0.3 V and Vov = 0.1 V → Max swing= 2*2*0.2 V = 0.8 V



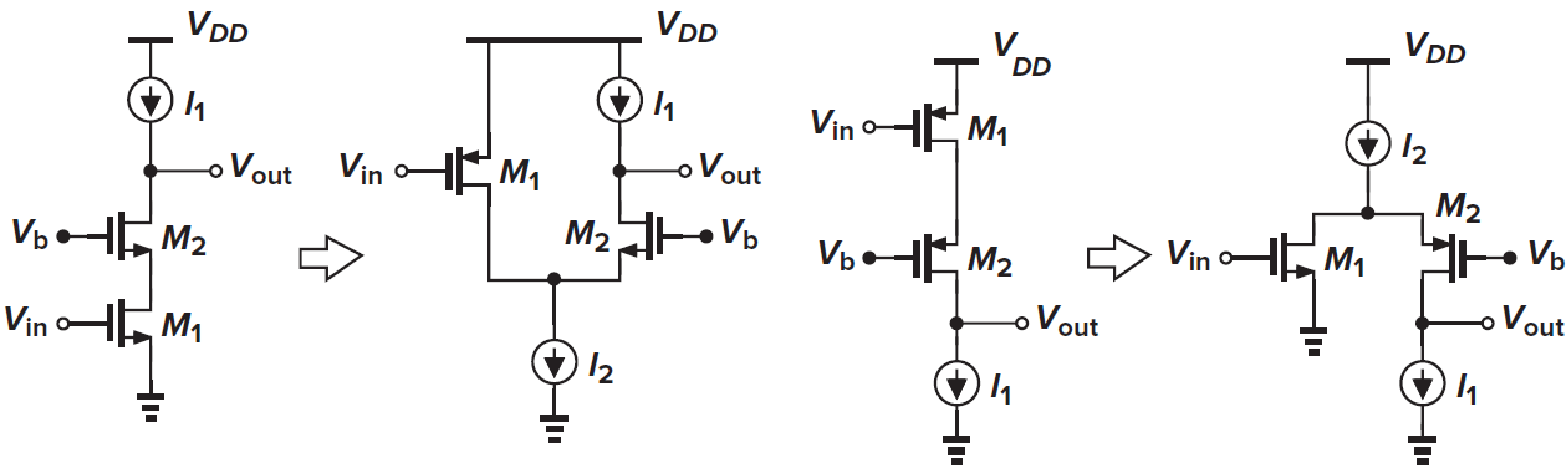


Telescopic Cascode with SE Output

- Low compliance (wide swing) current mirror load
- Note that for buffer connection, the swing is limited by M2 and M4

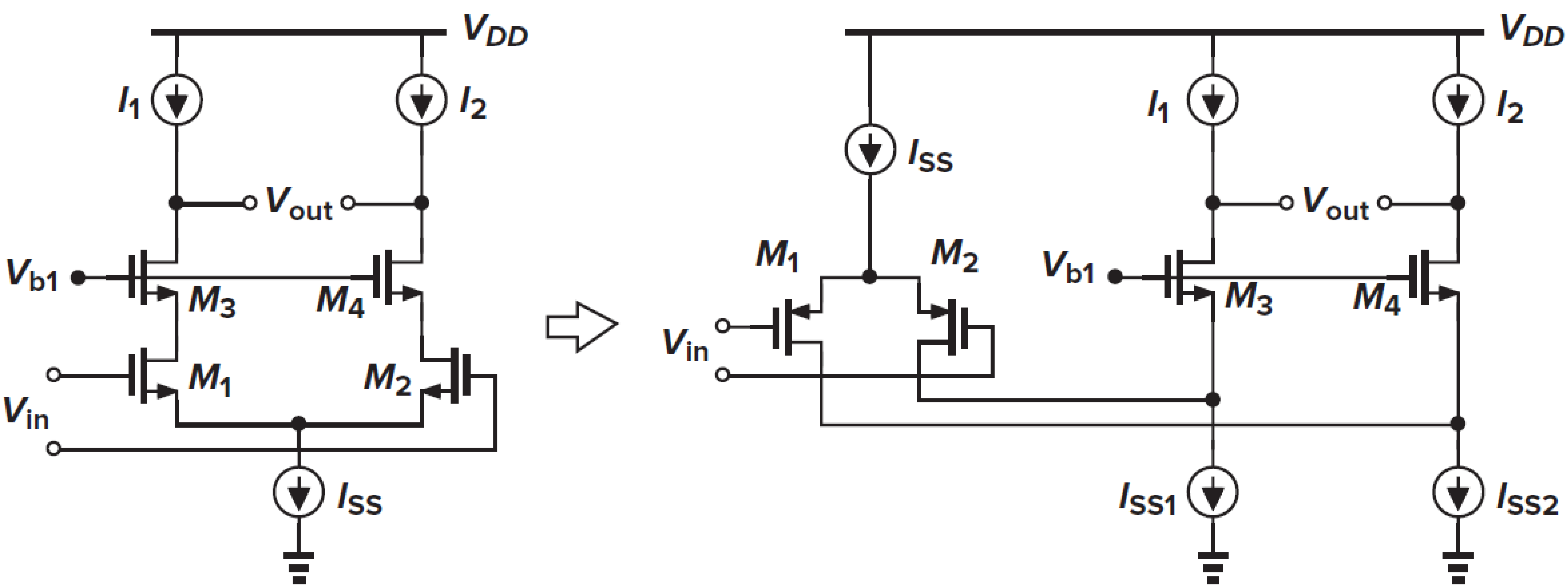


Folded Cascode



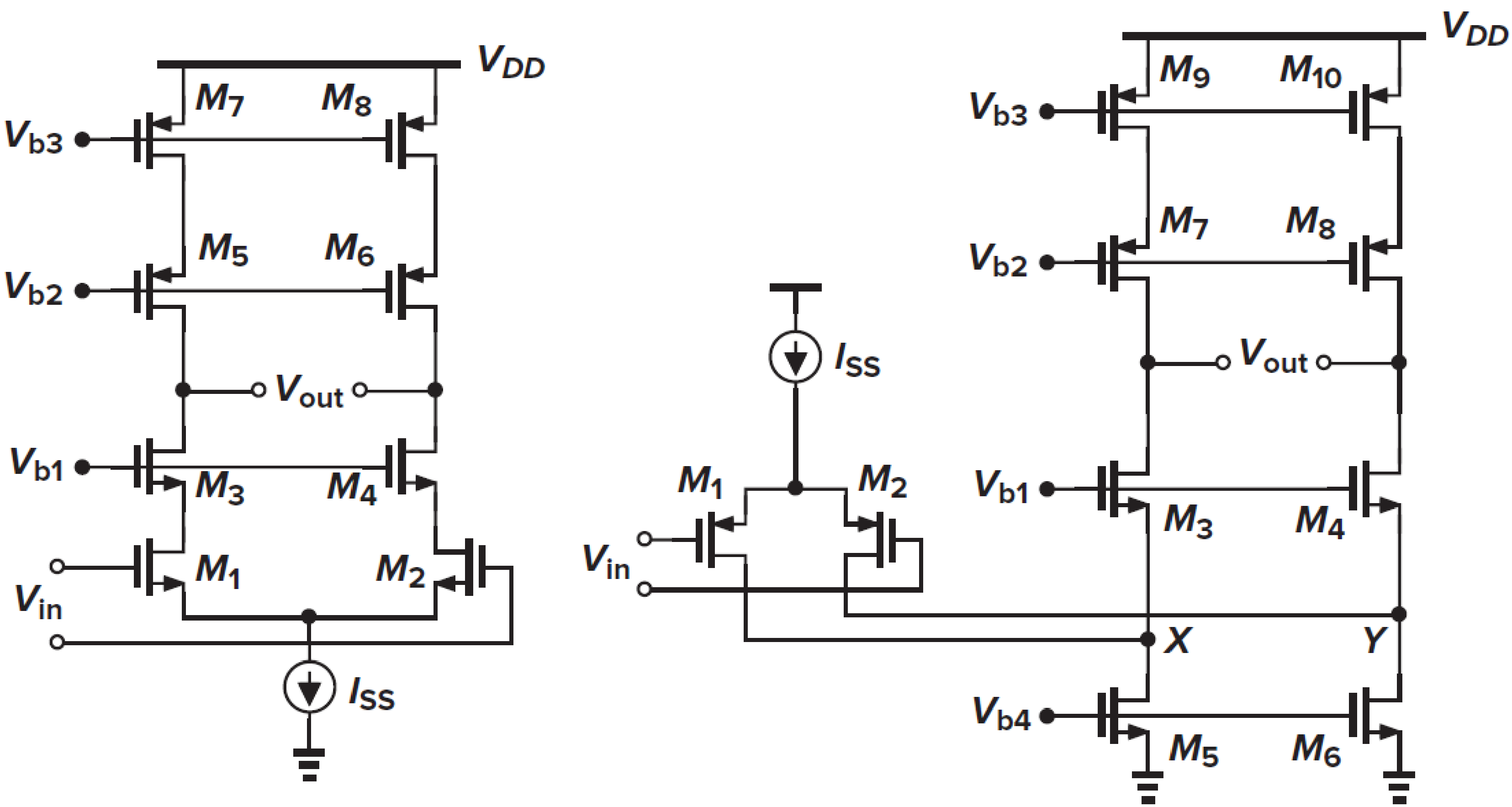
Telescopic vs Folded Cascode

- Higher power consumption: $I_{SS1,2} = \frac{I_{SS}}{2} + I_{1,2}$
  - M3,4 must be remain ON when I_{SS} is fully steered on one side
  - $I_{SS1,2} > I_{SS} \rightarrow I_{1,2} > \frac{I_{SS}}{2}$
  - Ex: $I_{SS1,2} = 1.2I_{SS} \rightarrow I_{1,2} = 0.7I_{SS} \rightarrow$ Total current $= 2.4I_{SS}$



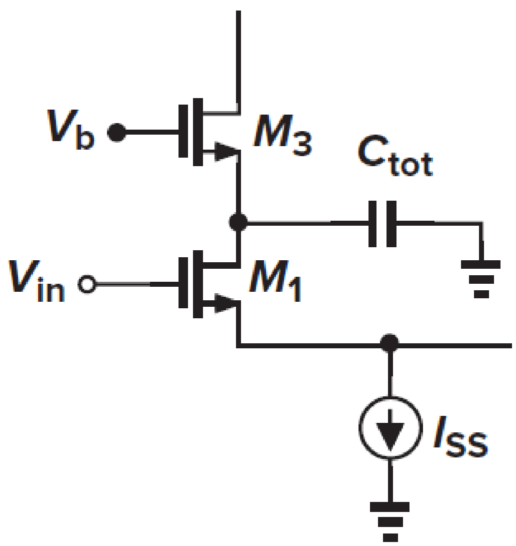
Telescopic vs Folded Cascode

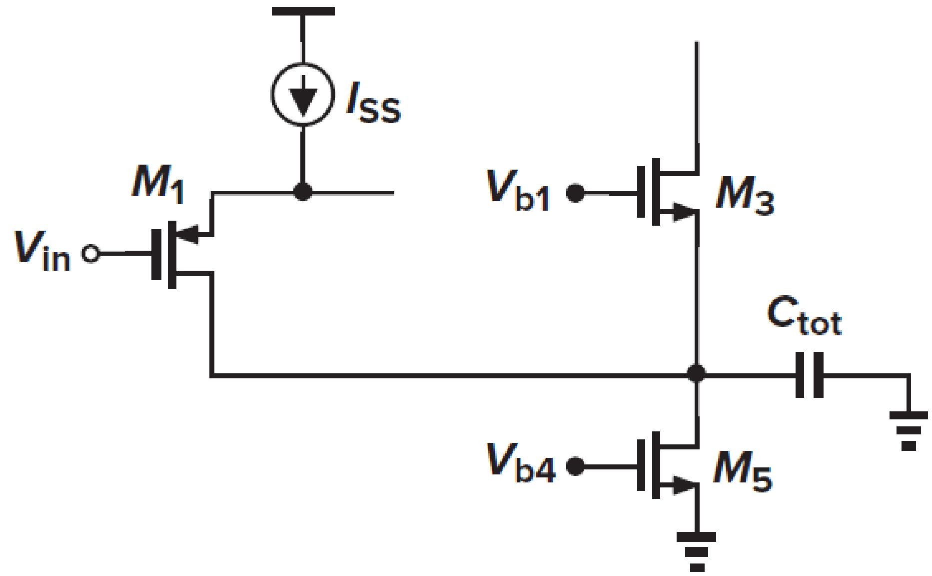
- Gain lower than telescopic cascode ($r_{o1}//r_{o5}$)
  - Around two times less



Telescopic vs Folded Cascode

- Non-dominant pole (at folding point) has higher capacitance
- Telescopic: $C_{tot} = C_{db1} + f(C_{gd1}) + C_{gs3} + C_{sb3}$
- Folded: $C_{tot} = C_{db1} + f(C_{gd1}) + C_{gs3} + C_{sb3} + \boldsymbol{C_{db5}} + \boldsymbol{C_{gd5}}$
- Note that M5 is large as it carries large current (large parasitics)





Folded Cascode CMIR (OL)

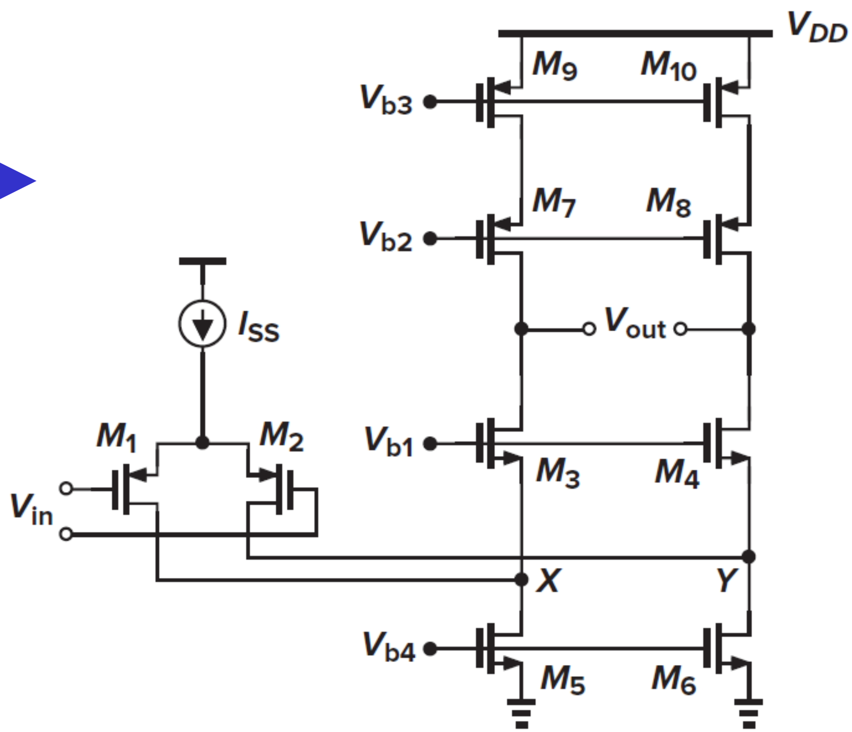
- ❑ $V_{in,CM}$ Range

- ❑ V_{out} Range

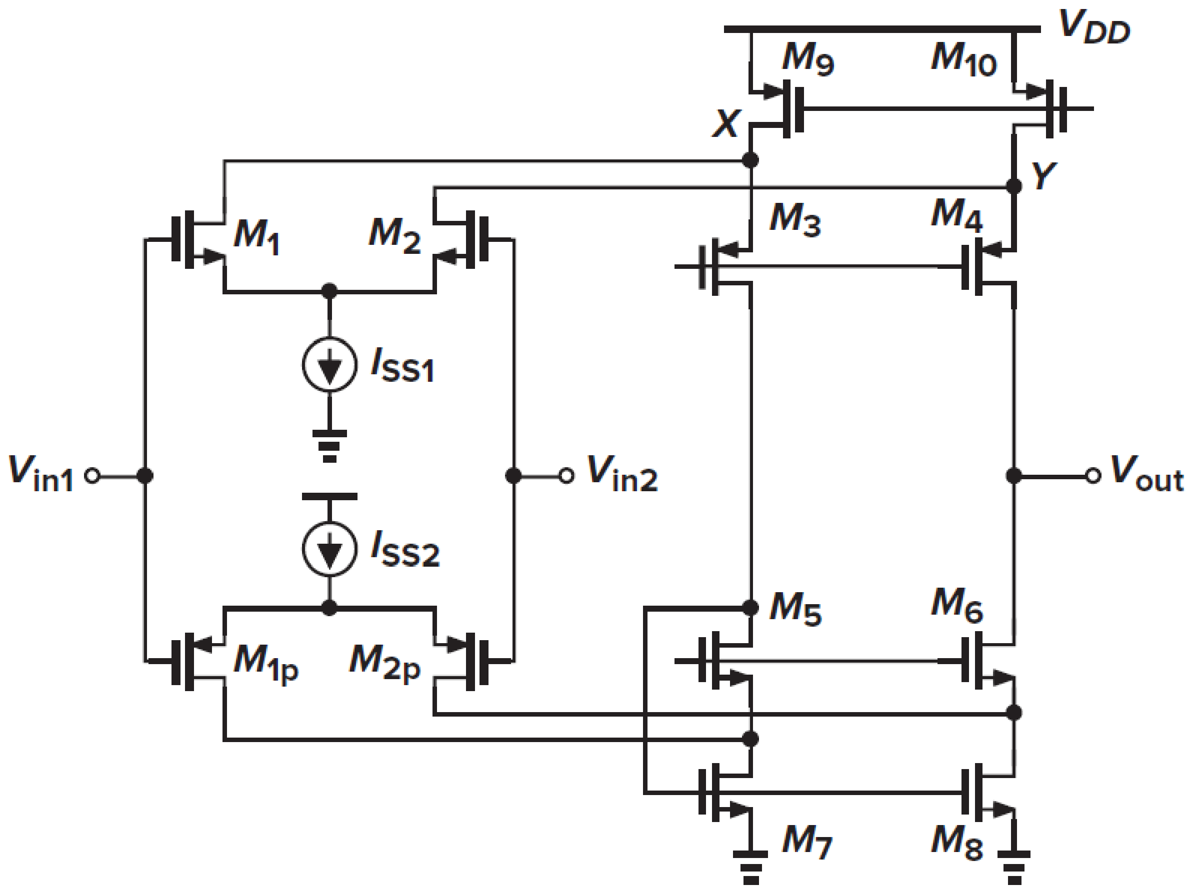
- ❑ **$V_{in,CM}$ and V_{out} are NOT coupled**
- ❑ **Main advantage over telescopic**
- ❑ Example: VDD = 1.2 V,

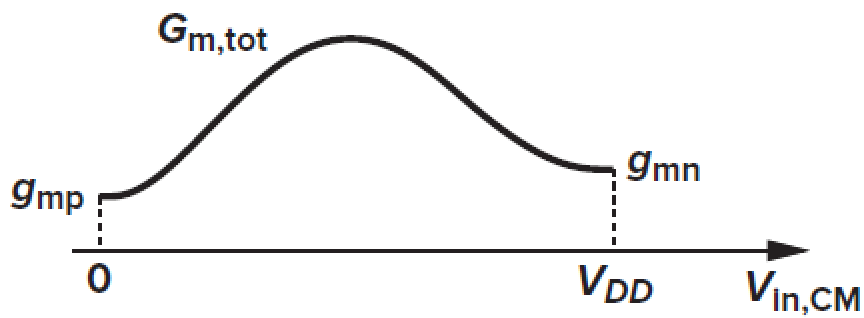
  VTH = 0.3 V, and Vov = 0.1 V

  - VCM = -0.2 V – 0.7 V
  - Max CM range = 0.9 V



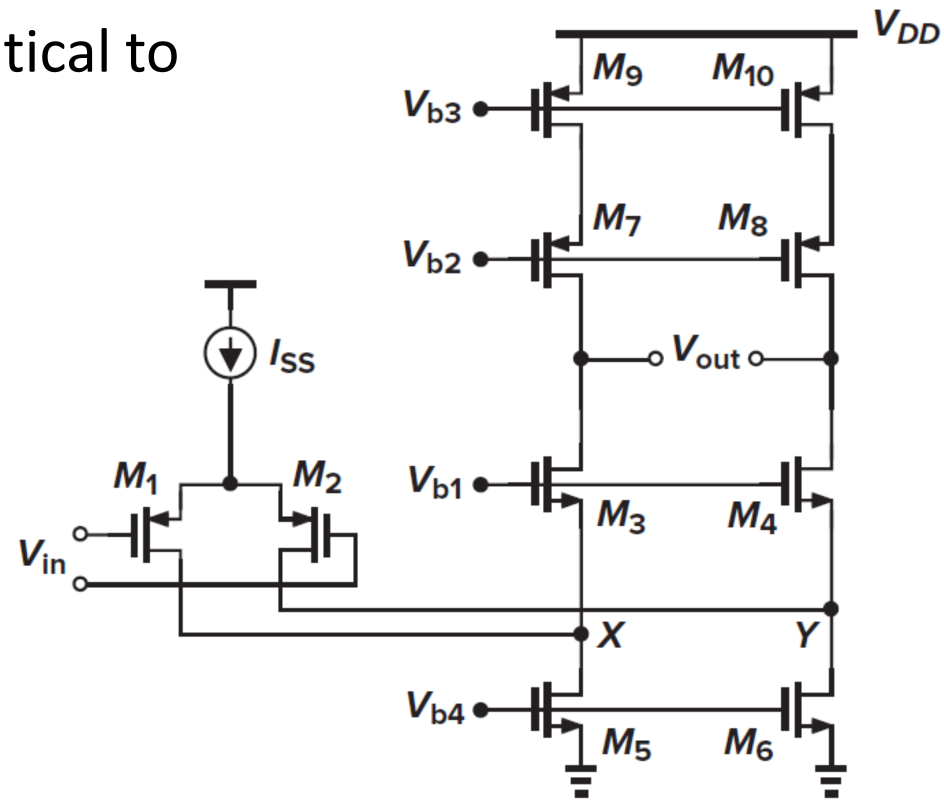
Folded Cascode with Rail-to-Rail CMIR





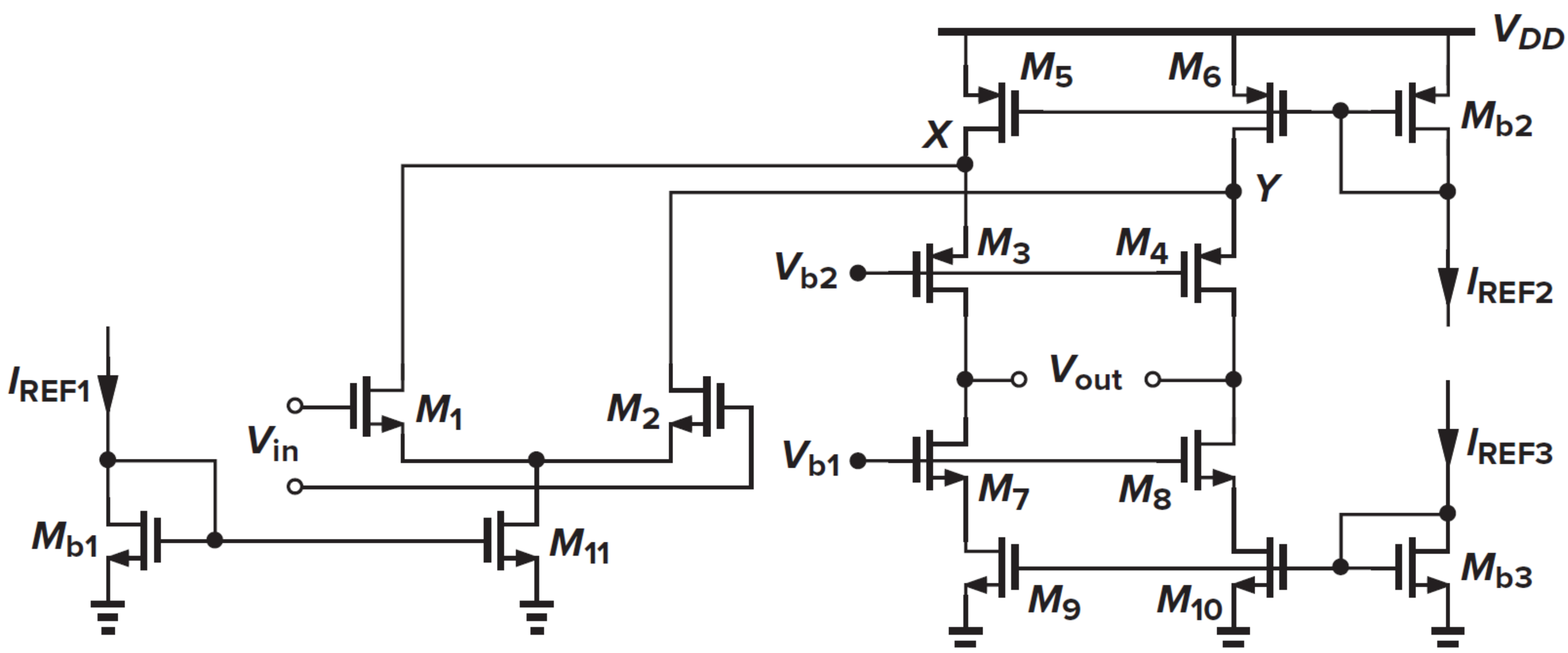
Folded Cascode Output Swing (OL)

- $Max\ Diff\ Swing \approx 2(V_{DD} - 4V_{ov})$
- Slightly better than telescopic cascode
- The choice of bias voltages is critical to maintain output swing



Folded Cascode

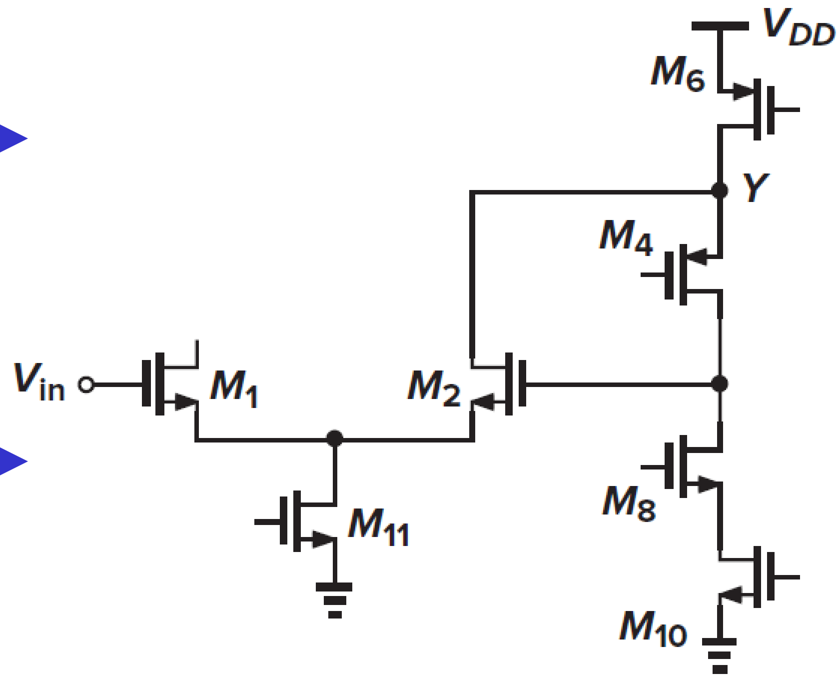
- NMOS i/p stage has higher gm for same dimensions
  - Or lower parasitics for same gm
- But the non-dominant pole at folding point is worse
- Also body effect and flicker noise may favor PMOS i/p pair



Folded Cascode as a Buffer

- V_{in} Range (OL, without FB)

- V_{out} Range (OL, without FB)



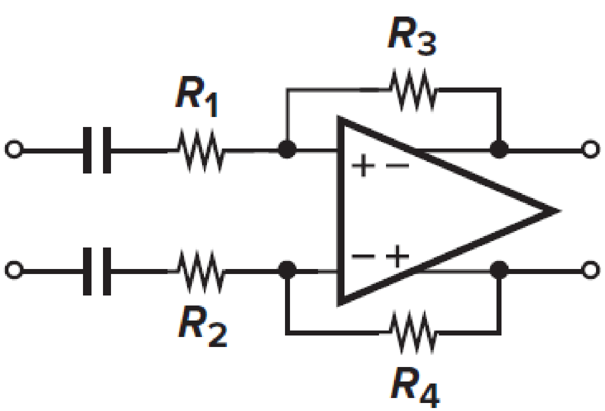
- $Max\ Swing = V_{DD} - V_{THN} - 4V_{ov}$
- Example: VDD = 1.2 V, VTH = 0.3 V, and Vov = 0.1 V
  - Max swing = 0.5 V
  - **Function of VDD**

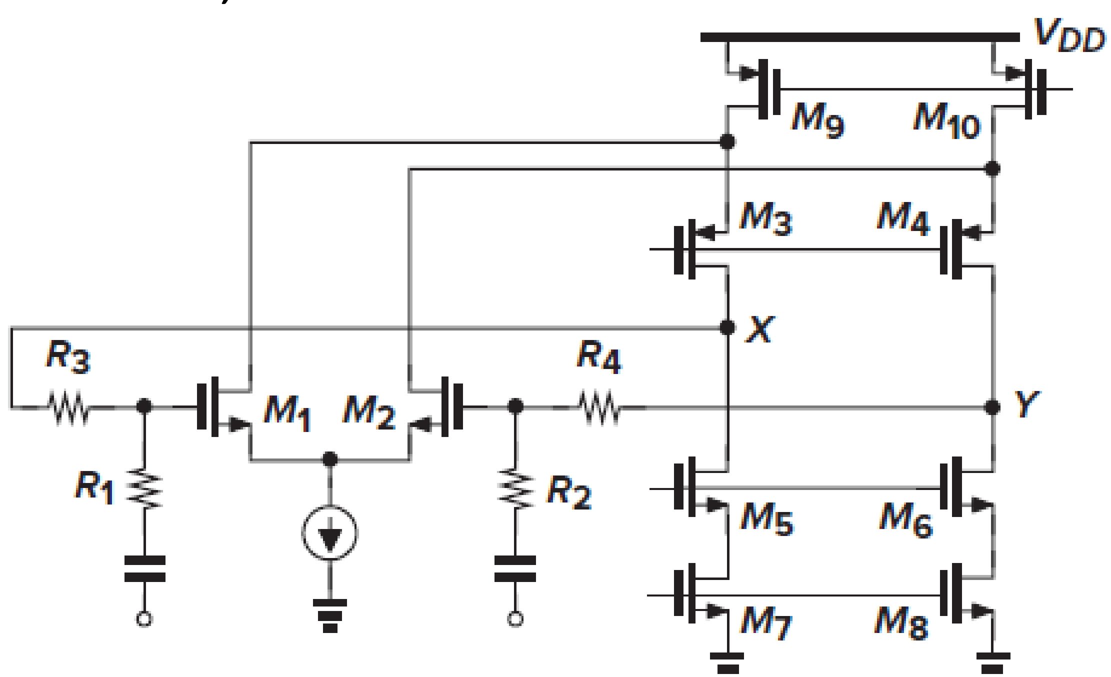
Folded Cascode CM Range (CL)

- Input and output CM levels are equal (why?) → similar to buffer
- $V_{in,CM}$ Range

- $V_{out,CM}$ Range

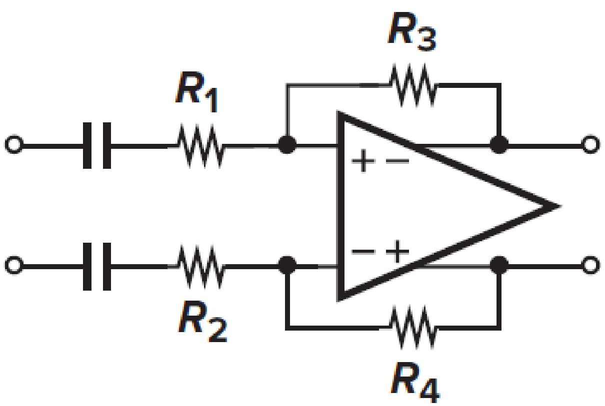
- $Max\ CM\ Range = V_{DD} - V_{THN} - 4V_{ov}$
- Example: VDD = 1.2 V, VTH = 0.3 V, and Vov = 0.1 V
  - Max CM range = 0.5 V

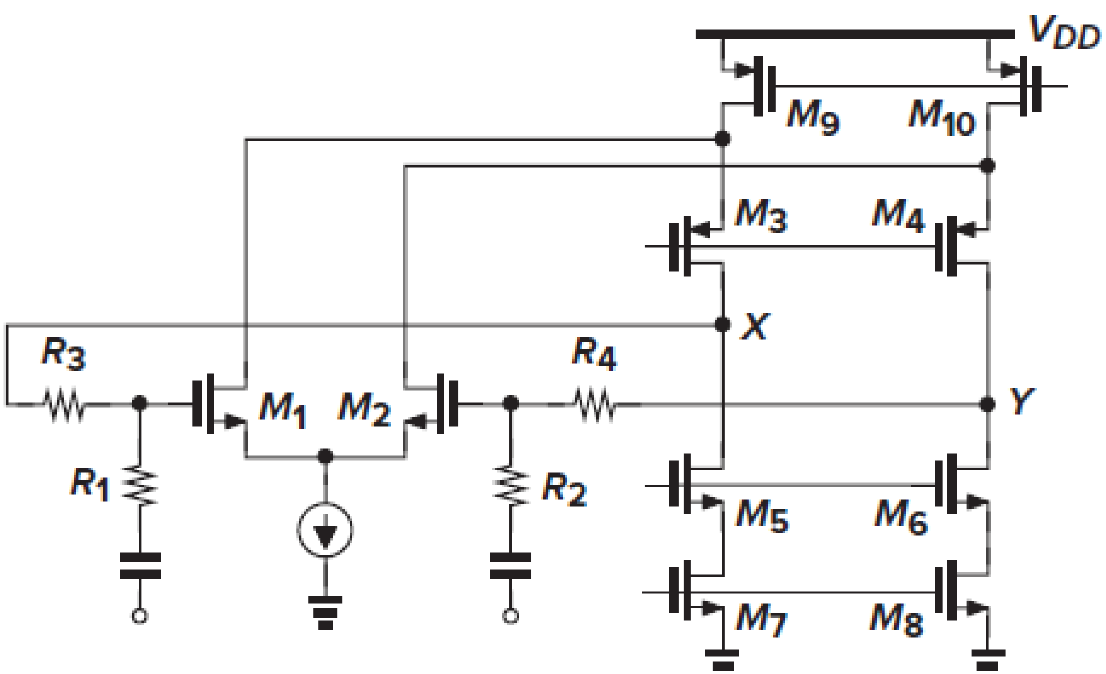




Folded Cascode Output Swing (CL)

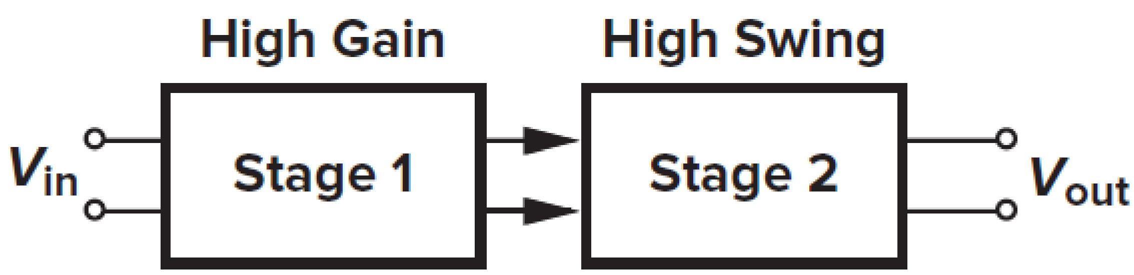
- ❑ Assume swing at input is negligible (high open-loop gain)
- ❑ $V_{in,CM}$ and V_{out} are NOT coupled
- ❑ $Max\ Diff\ Swing \approx 2(V_{DD} - 4V_{ov})$
- ❑ Example: VDD = 1.2 V, VTH = 0.3 V, and Vov = 0.1 V
  - Max diff swing = 1.6 V





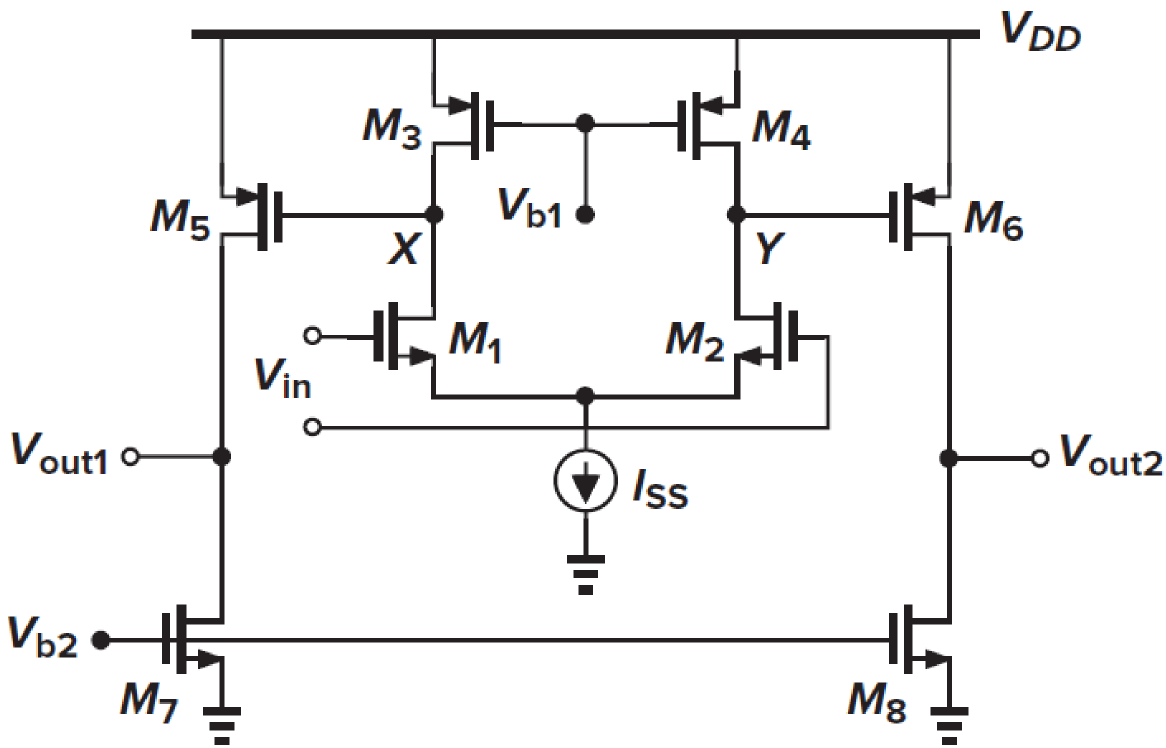
Two-Stage OTA

- 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



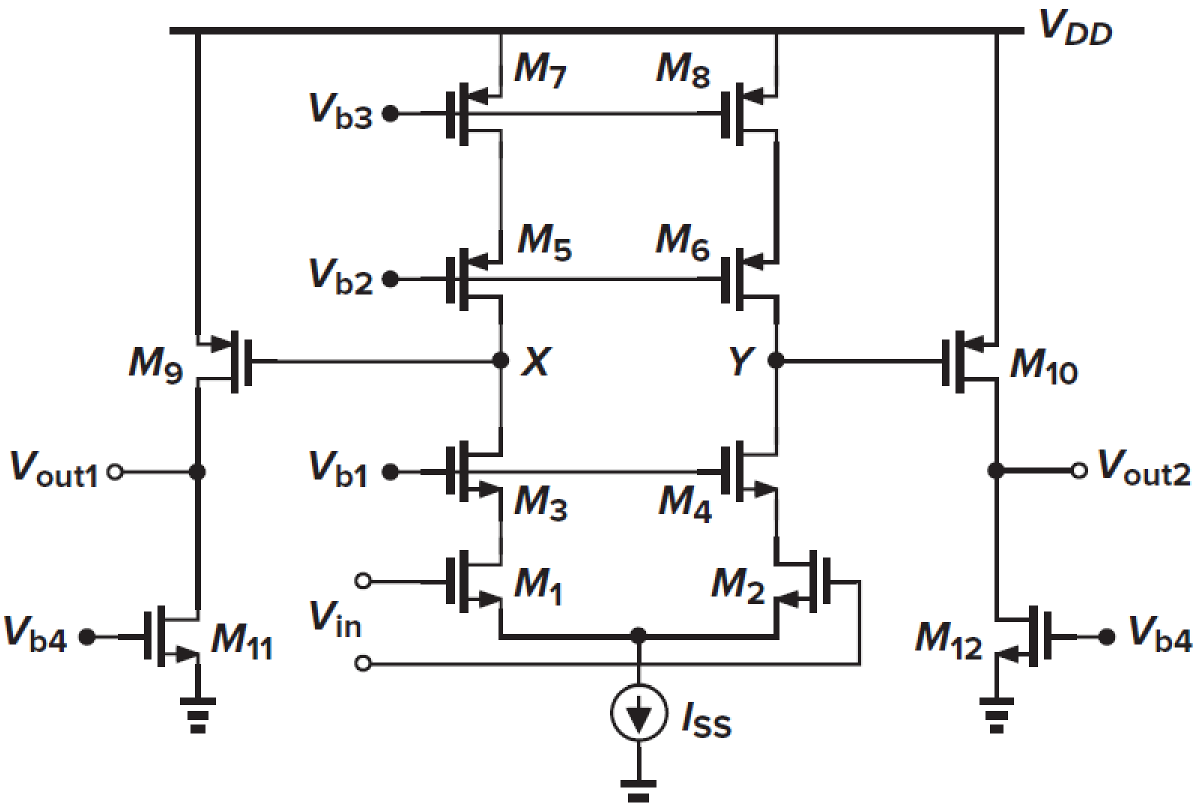
Two-Stage OTA

- $Max\ Diff\ Swing \approx 2(V_{DD} - 2V_{ov})$

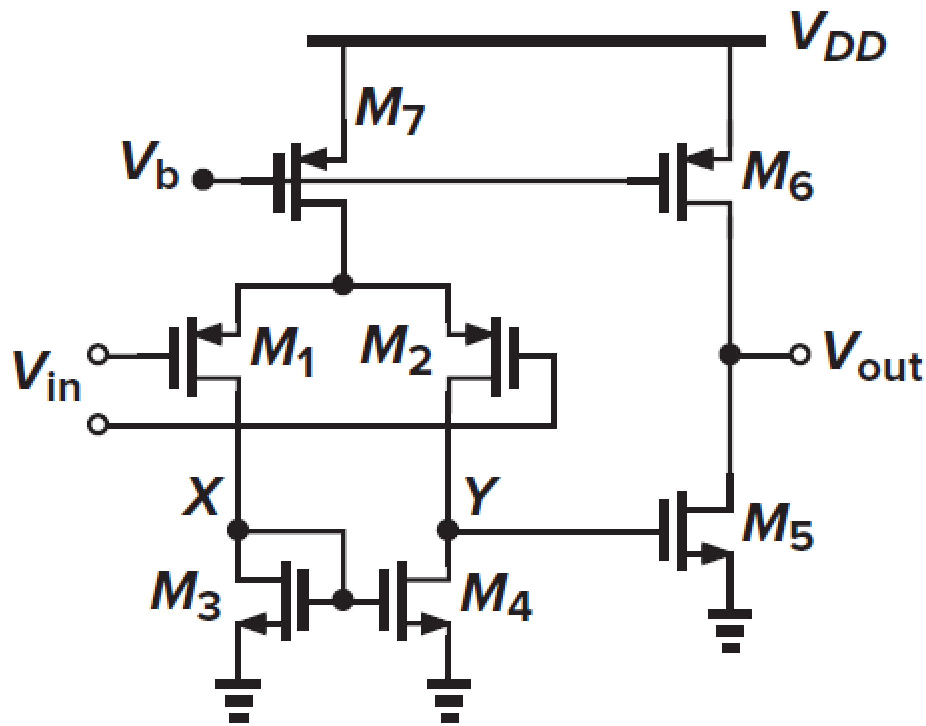


Two-Stage OTA

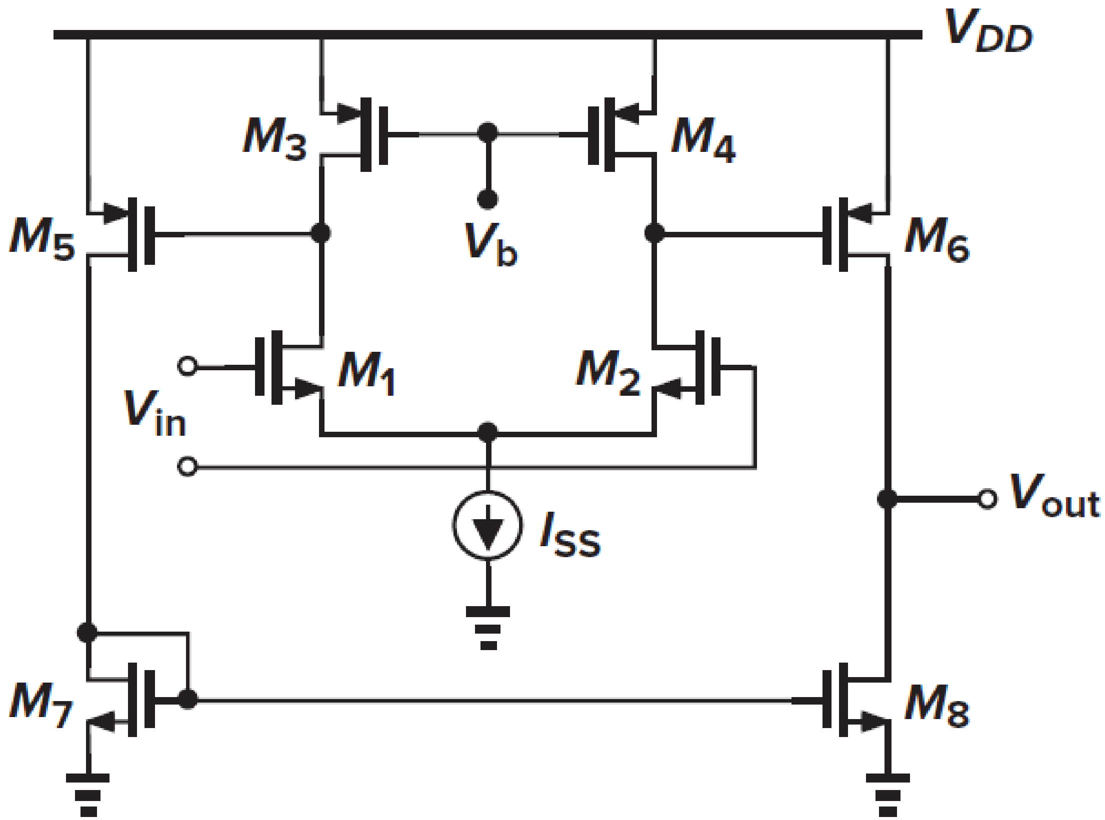
- $Max\ Diff\ Swing \approx 2(V_{DD} - 2V_{ov})$
- Voltage swing at X and Y is negligible



Two-Stage OTA with SE Output

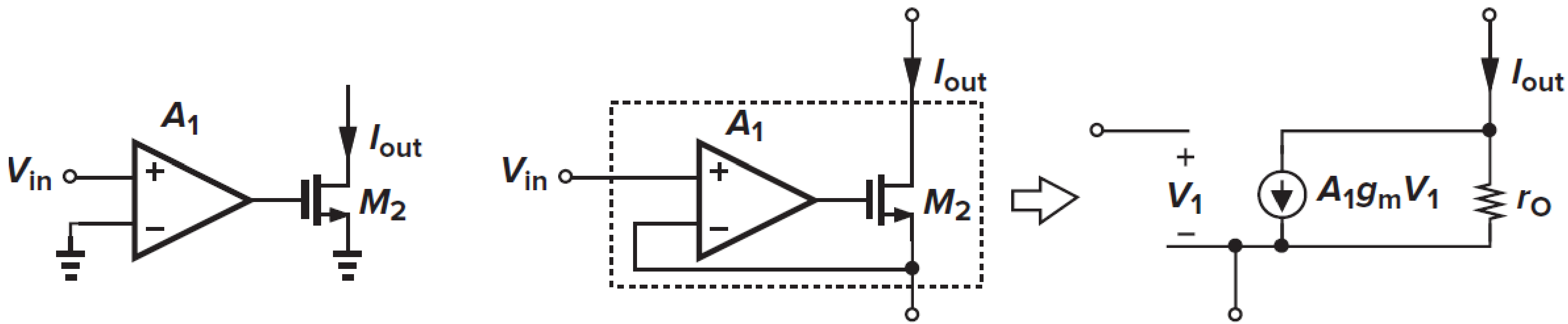


Two-Stage OTA with SE Output



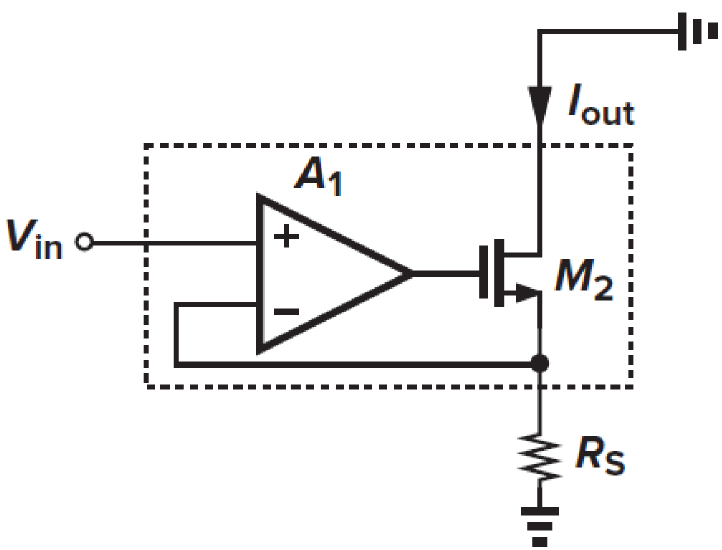
Gain Boosting: Super Transistor Perspective

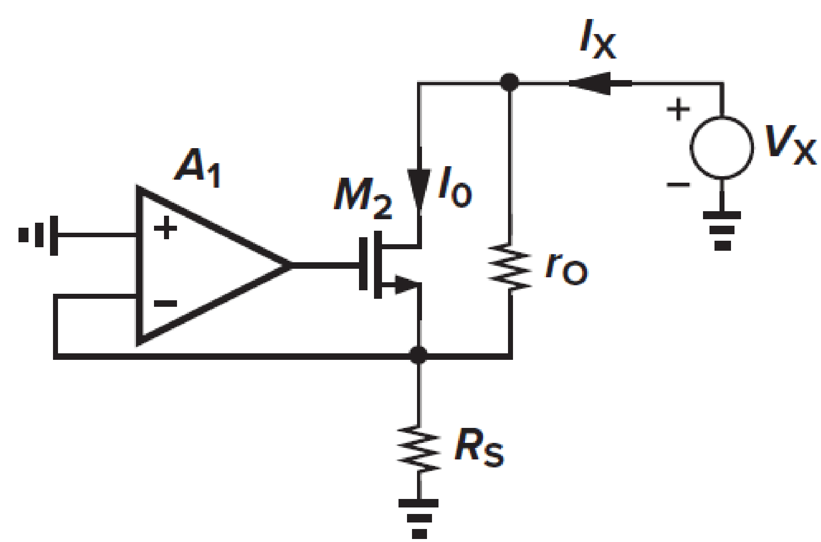
- Assume $A_1 \gg 1$
- $g_{m,super} = A_1 g_m$
- $r_{o,super} = r_o$



Gain Boosting: Super Transistor Perspective

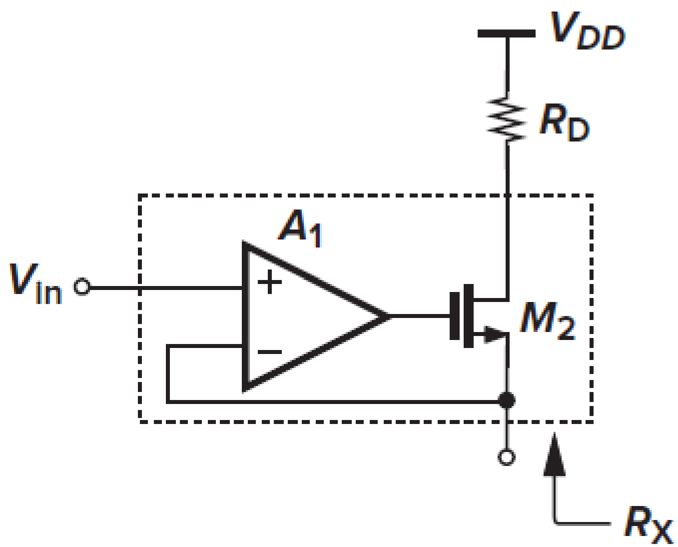
- $G_m \approx \frac{g_{m,super}}{1+g_{m,super}R_s} \approx \frac{A_1 g_m}{1+A_1 g_m R_s} \approx \frac{1}{R_S}$
- $R_{out} = r_o\left(1 + g_{m,super}R_S\right) = r_o(1 + A_1 g_m R_S)$

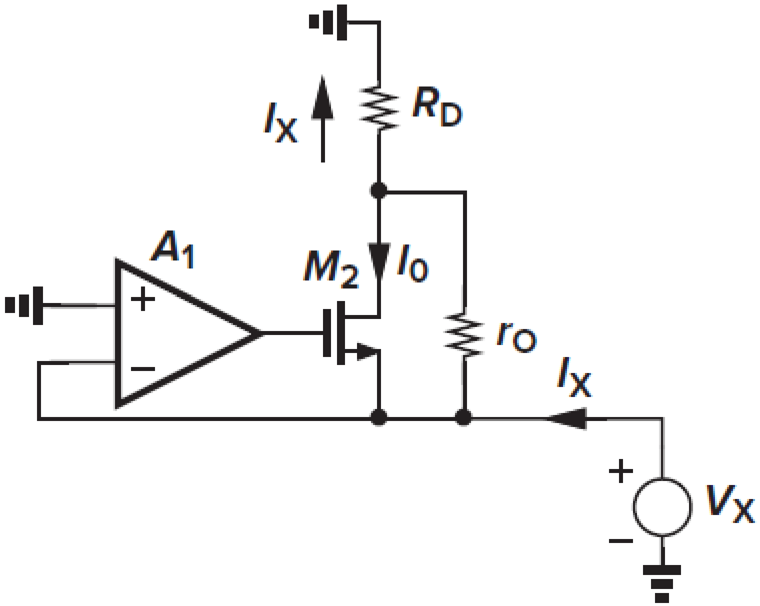




Gain Boosting: Super Transistor Perspective

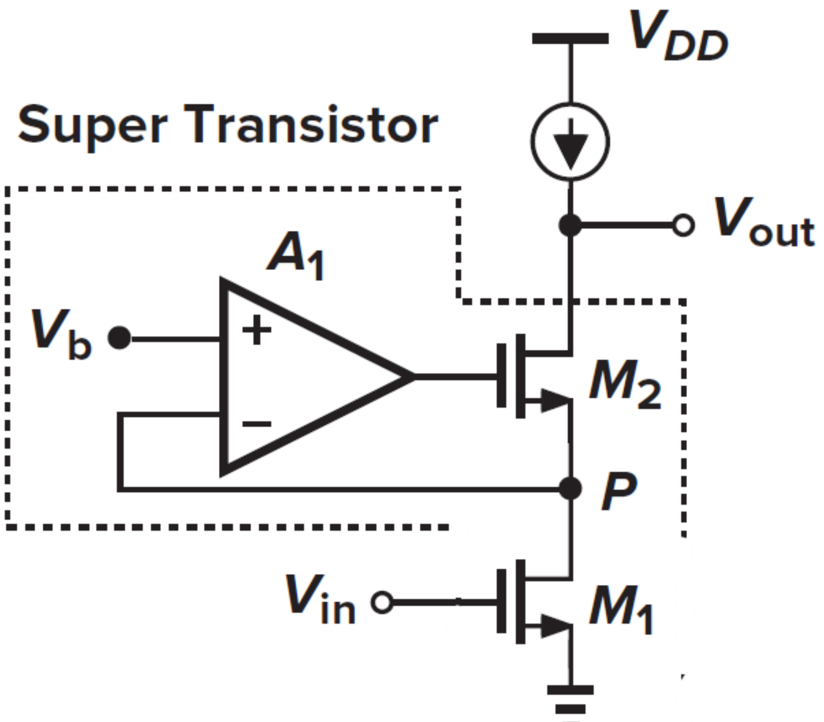
- $R_X \approx \dfrac{1}{g_{m,super}}\left(1 + \dfrac{R_D}{r_o}\right) \approx \dfrac{1}{A_1 g_m}\left(1 + \dfrac{R_D}{r_o}\right)$





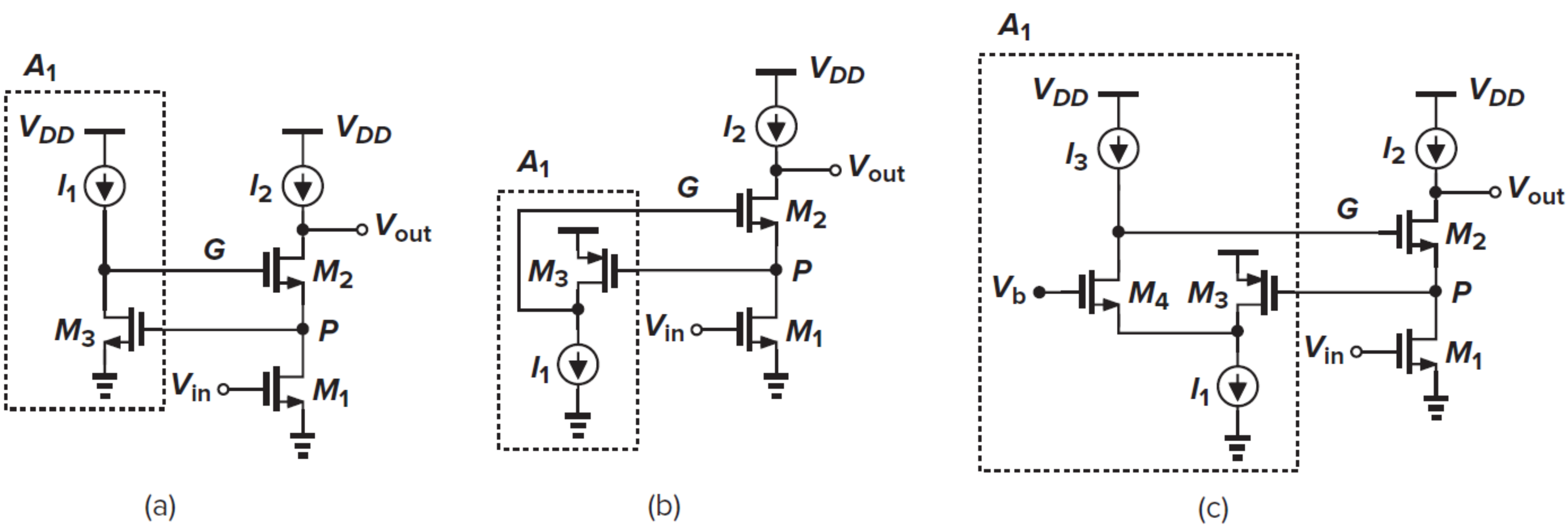
Gain Boosted (Regulated) Cascode

- $G_m \approx g_{m1}$
- $R_{out} = r_{o2}\left(1 + g_{m2,super} r_{o1}\right) = r_{o2}(1 + A_1 g_{m2} r_{o1})$
- $A_v \approx A_1(g_{m1} r_{o1})(g_{m2} r_{o2})$
- Gain is boosted while preserving headroom



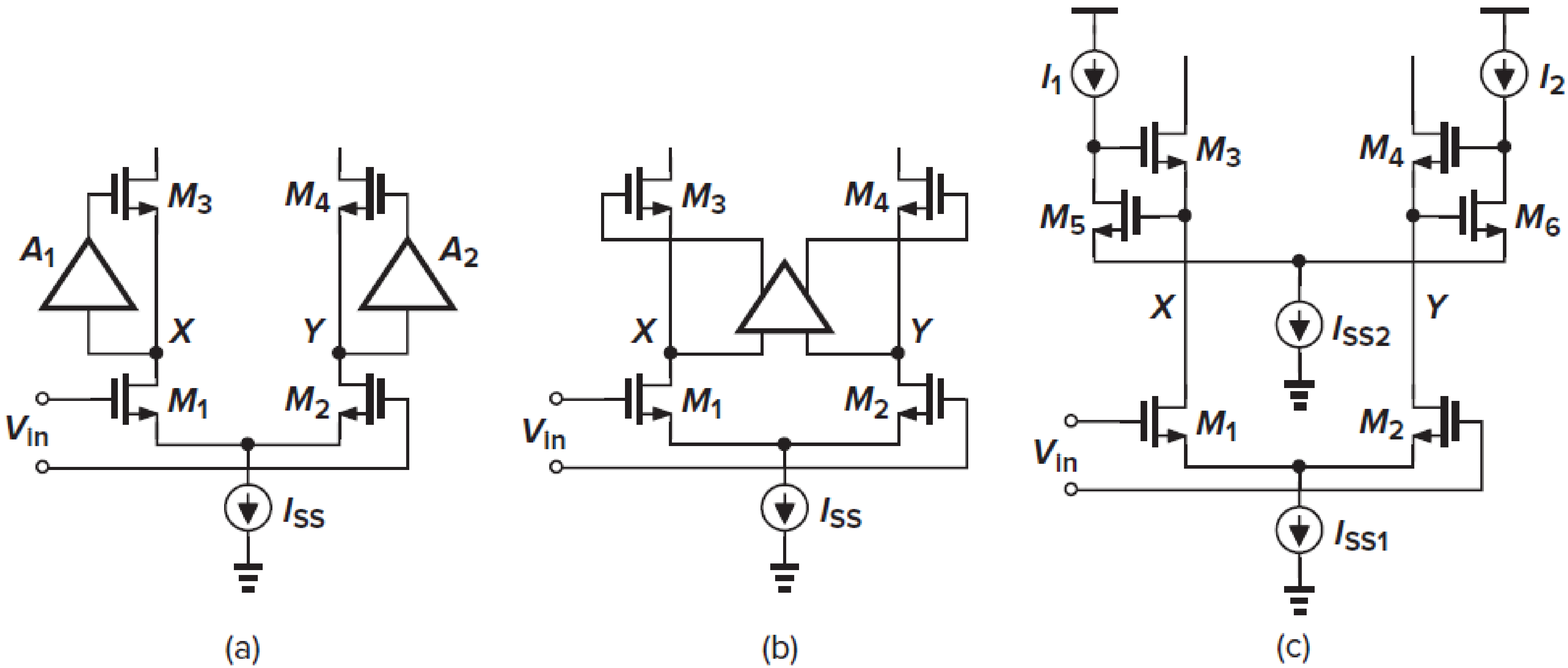
Gain Boosting Implementation

- NMOS CS (a): headroom limitation
  - $V_P = V_{TH} + V_{ov3}$ instead of V_{ov1}
- PMOS CS (b): M3 will be in triode
  - $V_G - V_P > V_{TH}$
- Folded cascode (c): M4 provide level shift



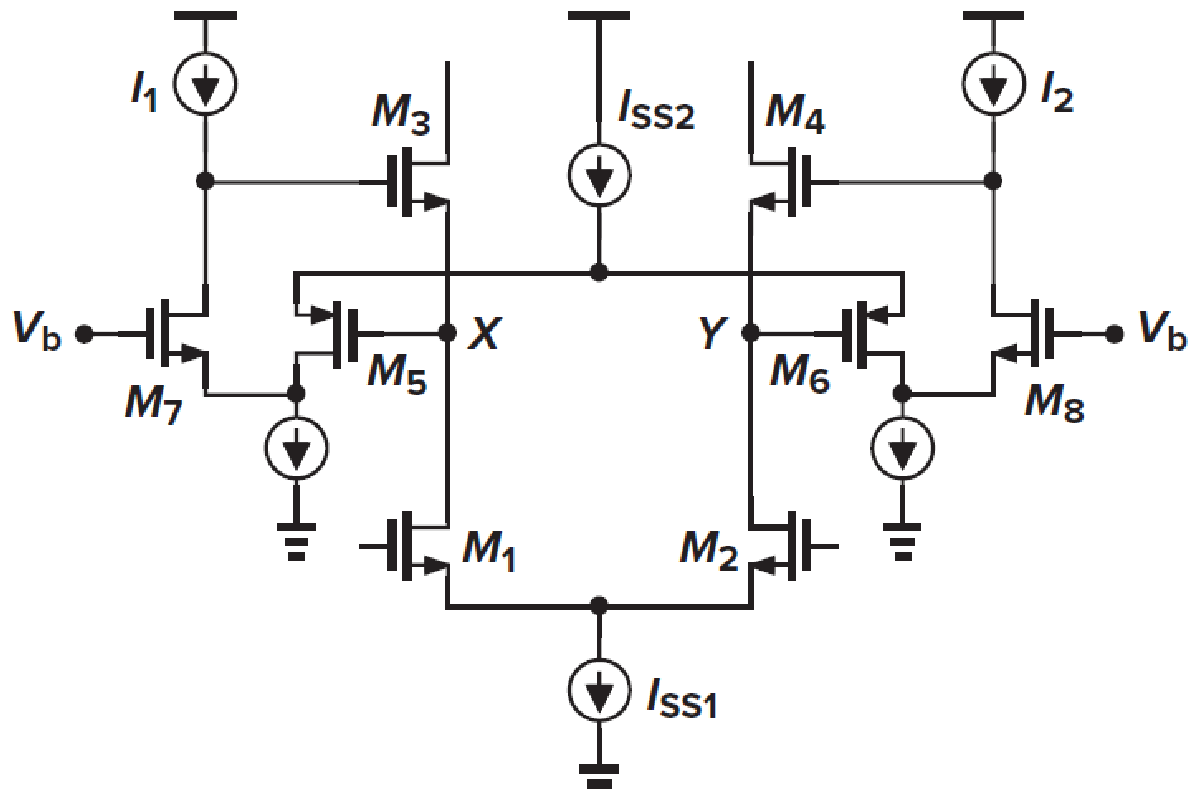
Gain Boosted OTA

- NMOS CS implementation replaced by a diff pair
- Headroom limitation
  - $V_{X,Y} = V_{TH} + V_{ov5,6} + V_{ISS2}$ instead of $V_{ov1} + V_{ISS1}$



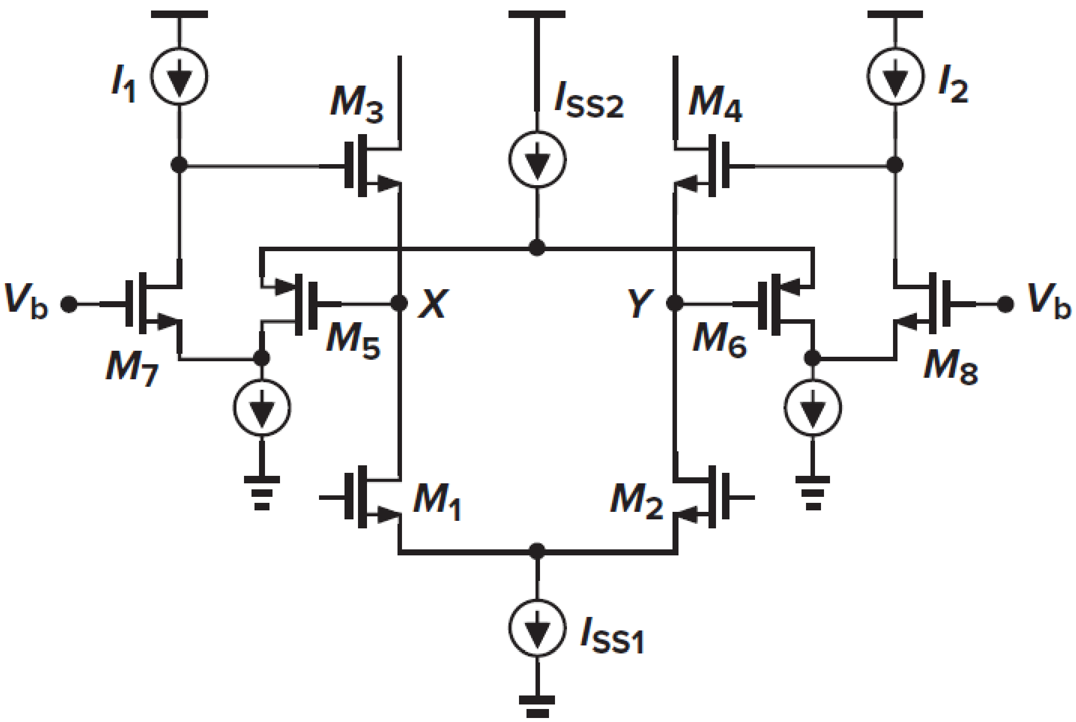
Gain Boosted OTA

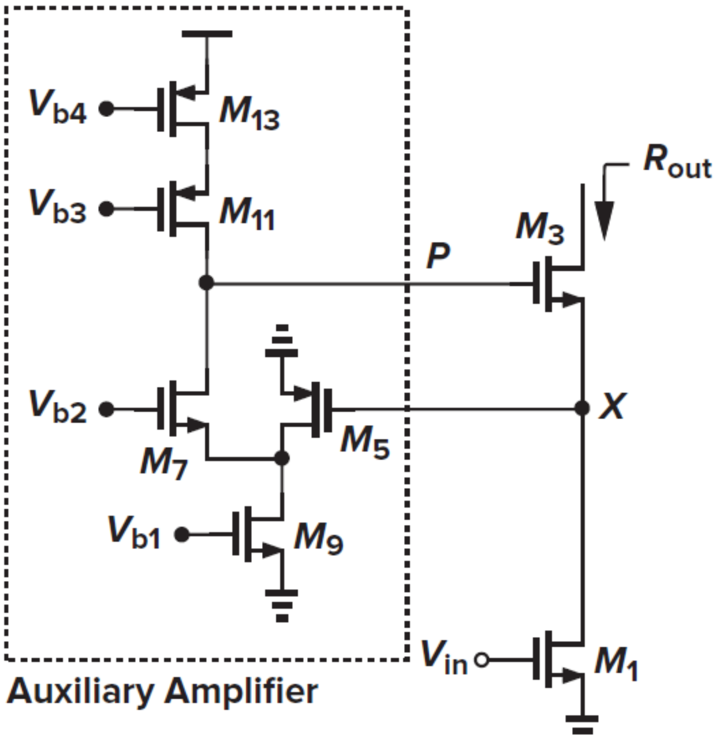
- Folded-cascode used as auxiliary amplifier
- No headroom limitation
  - $V_{X,Y} = V_{ov1} + V_{ISS1}$



Bonus Question

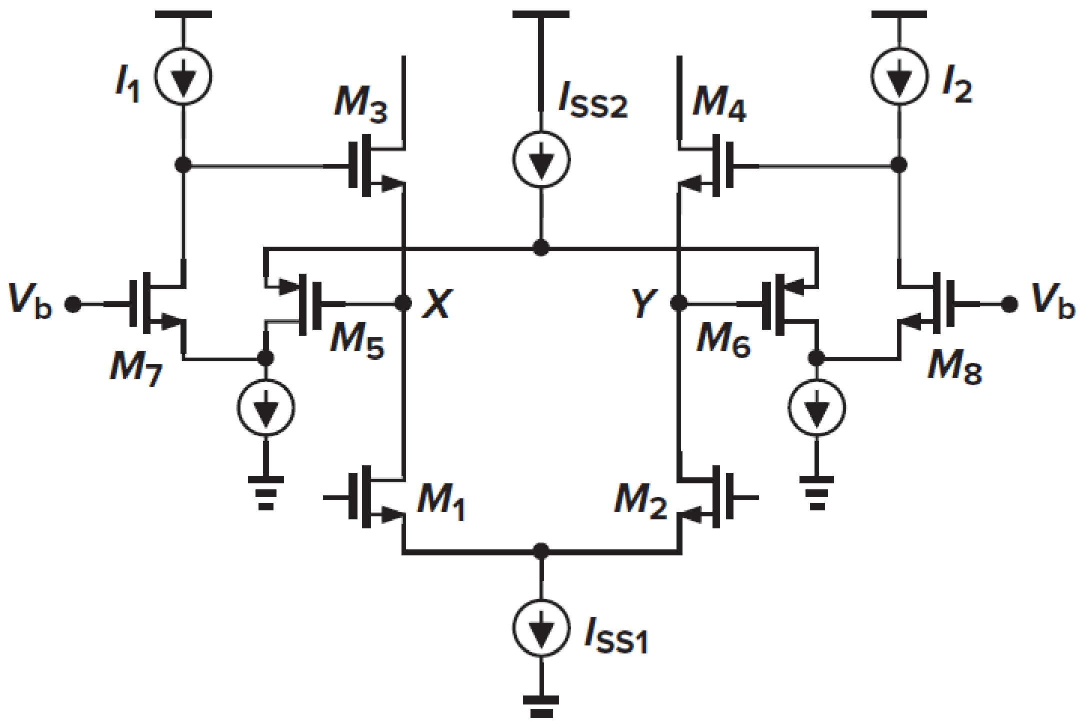
- Calculate the voltage gain. Assume all transistors have the same gm and ro. Assume the load is ideal CS (not drawn).

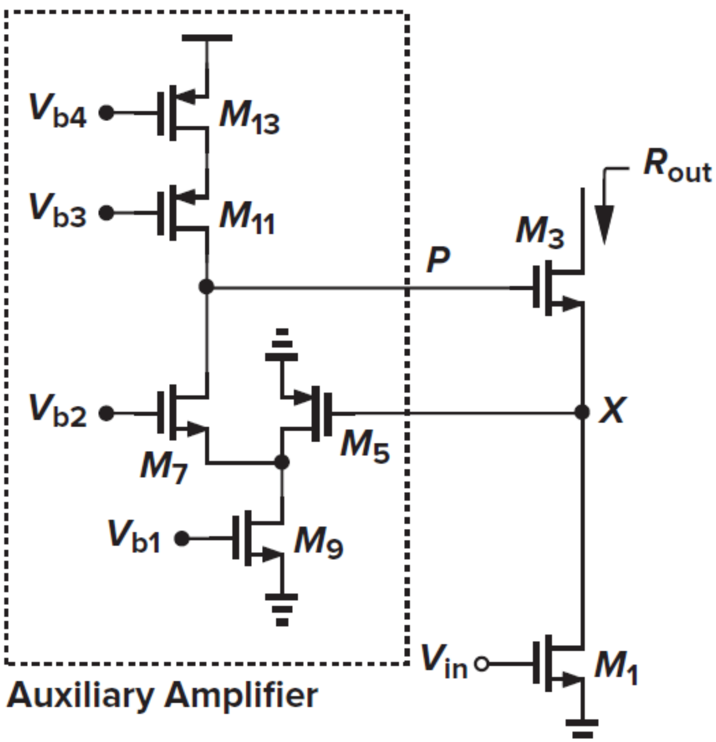




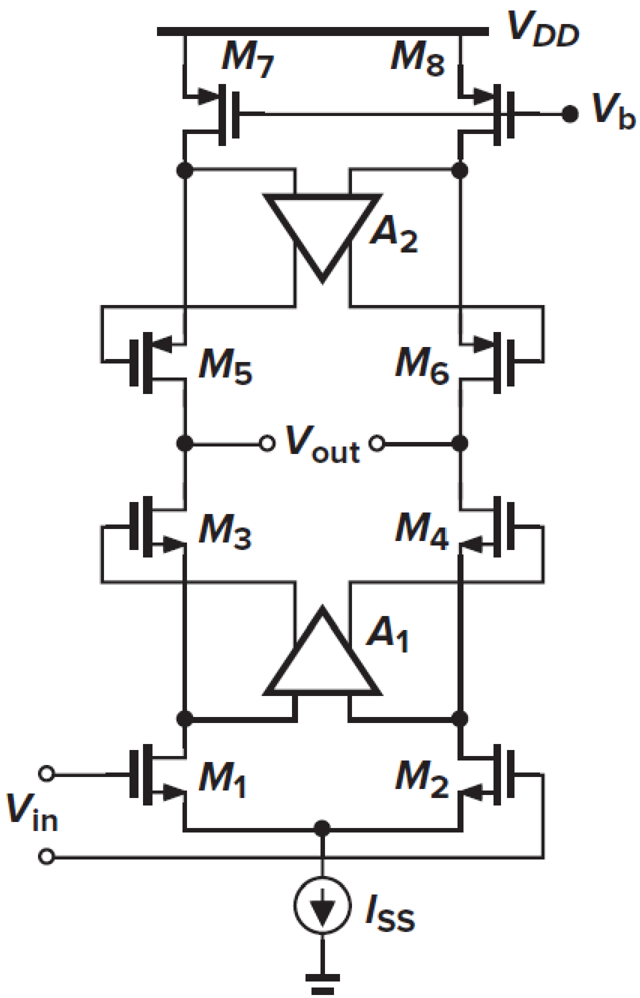
Gain Boosted OTA

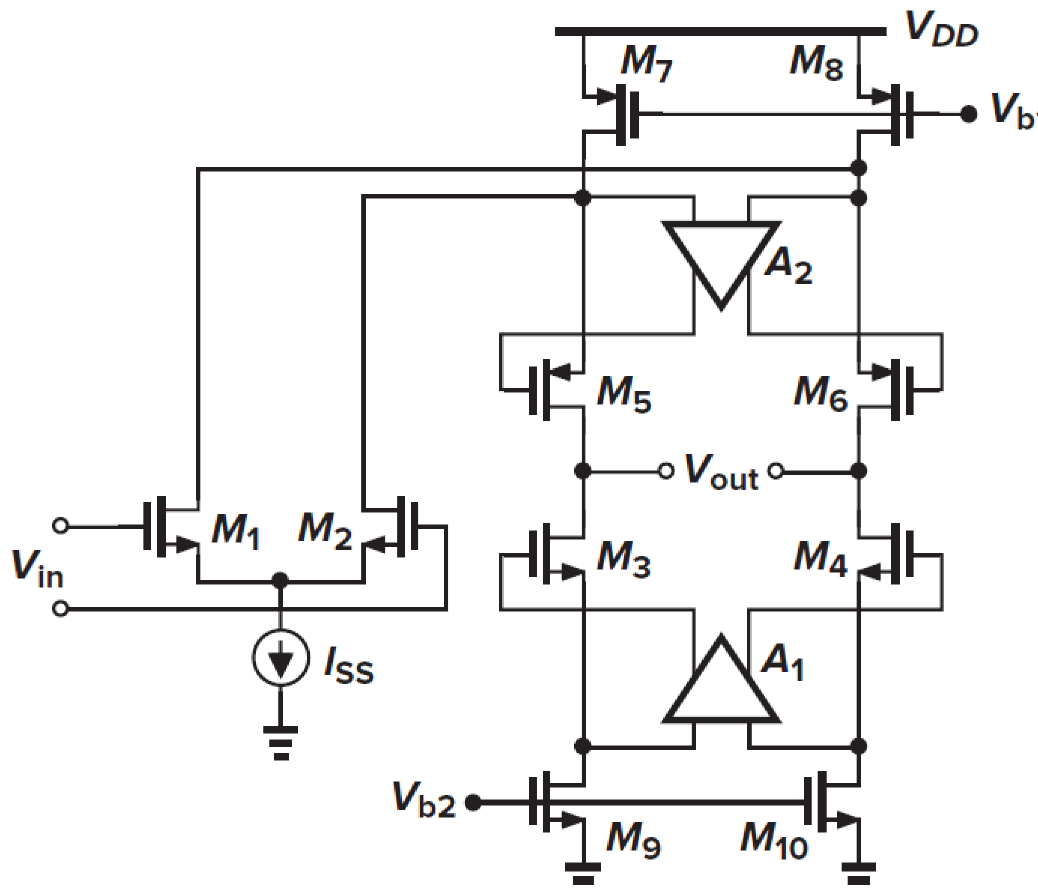
- $G_m \approx g_{m1}$
- $R_{out} \approx r_{o3}\left(1 + g_{m3,super} r_{o1}\right) \approx r_{o3}(A_1 g_{m3} r_{o1})$
- $A_1 \approx g_{m5}[g_{m11} r_{o11} r_{o13} // g_{m7} r_{o7} (r_{o5} // r_{o9})]$
- $A_v \approx \frac{1}{3}(g_m r_o)^4$ → quadruple cascode





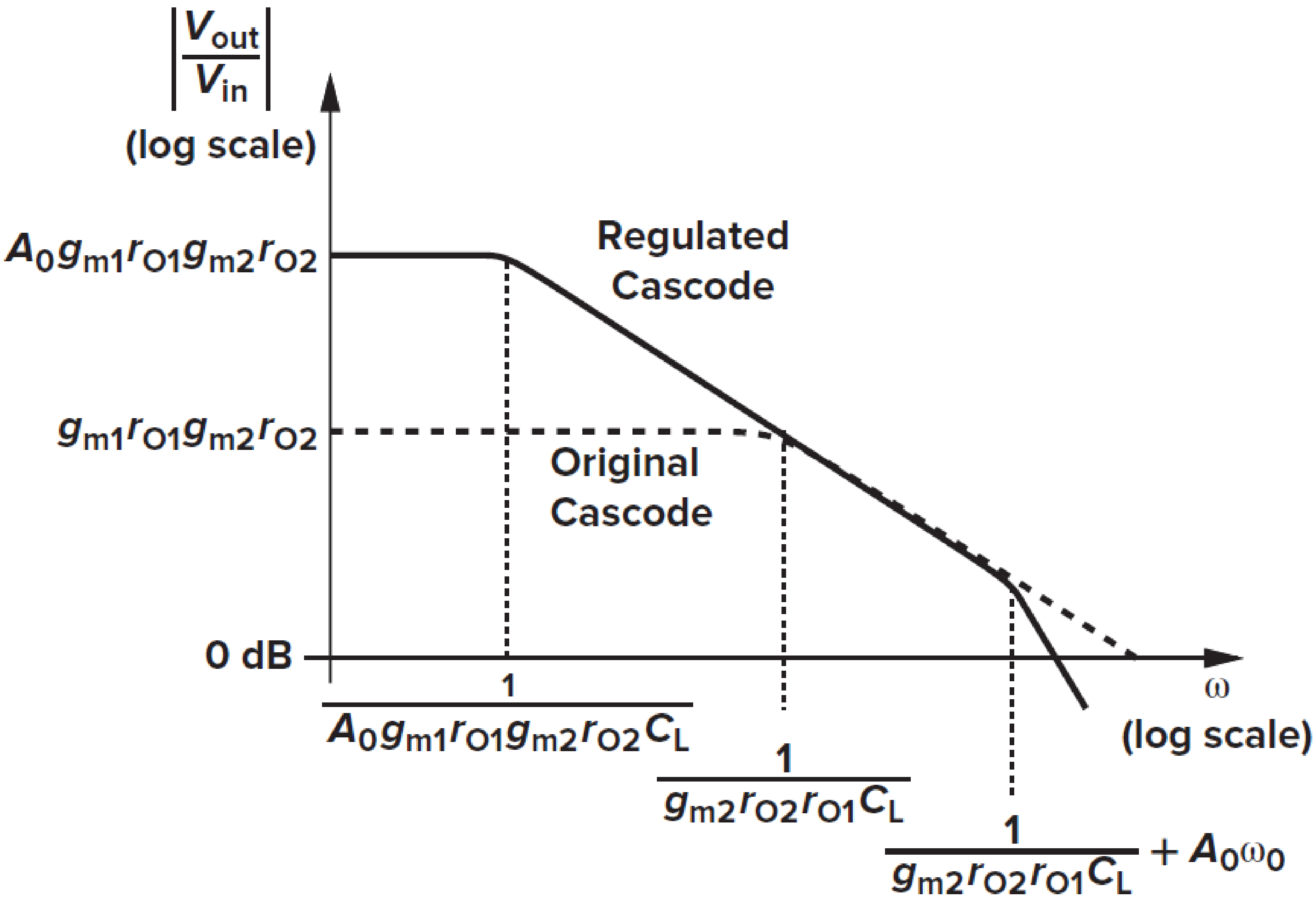
Gain Boosted OTA





Gain Boosted OTA Frequency Response

- See [Razavi, 2017] Section 9.4.3



Comparison

	Gain	Output Swing	Speed	Power Dissipation	Noise
Telescopic	Medium	Medium	Highest	Low	Low
Folded−Cascode	Medium	Medium	High	Medium	Medium
Two−Stage	High	Highest	Low	Medium	Low
Gain−Boosted	High	Medium	Medium	High	Medium

Thank you!