

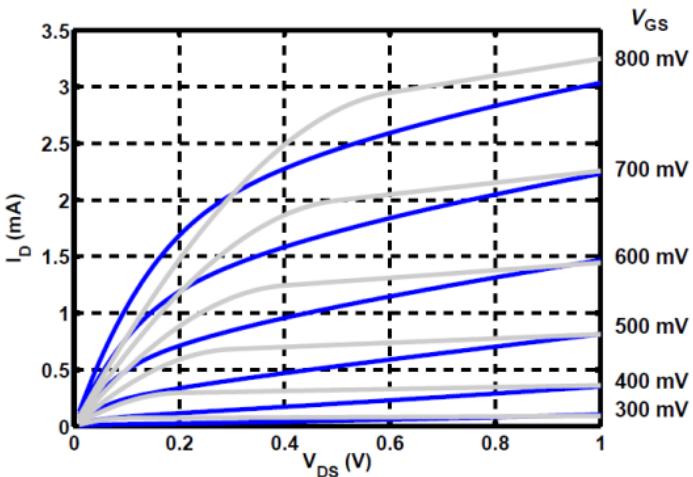
Chapter 11: Nanometer Design Studies

- 11.1 Transistor Design Considerations**
- 11.2 Deep-Submicron Effects**
- 11.3 Transconductance Scaling**
- 11.4 Transistor Design**
- 11.5 Op Amp Design Examples**
- 11.6 High-Speed Amplifier**
- 11.7 Summary**

Transistor Design Considerations

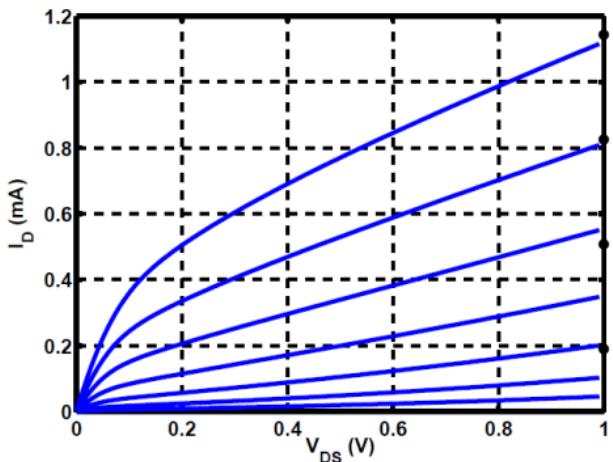
- Large signal model is necessary in two cases:
 - When transistor experiences large voltage (or current changes) due to input or output signals, disobeying the small-signal model
 - When transistor must be biased, requiring certain terminal voltages so as to carry a specified current
- Former case occurs occasionally while latter almost always occurs
- Large-signal behavior of nanometer MOSFETs departs from “long-channel” model developed
- Due to technology scaling, several effects manifest altering the I/V characteristics

Transistor Design Considerations



- Above figure plots the actual ID - V_{DS} characteristics of an NFET with $W/L = 5 \mu\text{m}/40 \text{ nm}$ and $V_{TH} \approx 300 \text{ mV}$ (using a BSIM4 model) as against a “best-fit” long-channel square-law approximation
- The two models diverge considerably
 - Cannot perform bias calculations using square-law model

Transistor Design Considerations

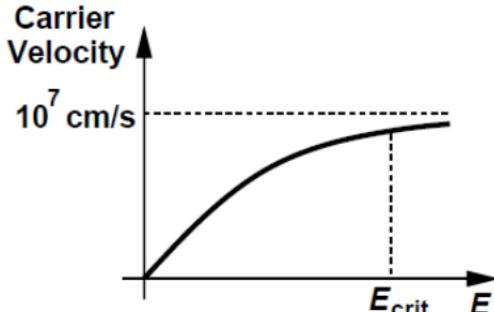


MOS small-signal model still holds for short-channel devices
Expressions relating gm and rO to bias conditions need revision
Difficult to distinguish between triode and saturation regions
Can associate a “knee” point to each curve as a rough boundary

- Above figure plots actual 40-nm device characteristics for $VGS - VTH = 50 \text{ mV}, 100 \text{ mV}, \dots, 350 \text{ mV}$
- Knee points are observed below $VDS = 0.2 \text{ V}$ ($W = 5\mu\text{m}$ and $VTH \approx 200 \text{ mV}$)

Deep-Submicron Effects

- Velocity Saturation:
- In a MOSFET, as V_{DS} and hence the electric field along the source-drain path increase, 'v' does not increase proportionally
- Carriers experience “velocity saturation”, i.e., mobility falls
- Effect arises because length of MOSFETs has shrunk from 1 μm to 40 nm while the allowable drain-source voltage has decreased from 5 V to about 1 V
- Lateral electric field has exceeded E_{crit} in fig. below



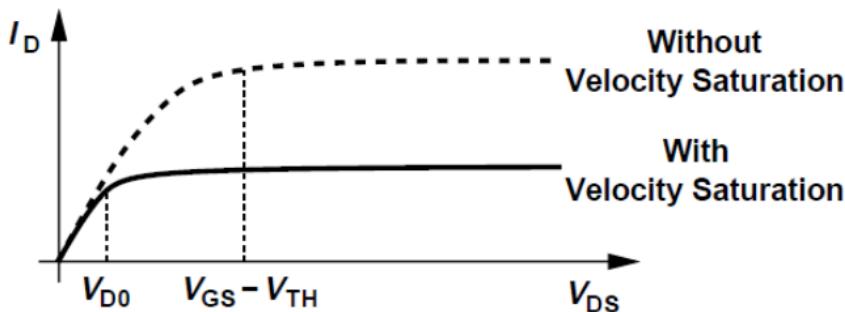
Deep-Submicron Effects

- Velocity Saturation:
- Suppose charge carriers reach saturated velocity, v_{sat} as soon as they depart from the source
- Since $I = Qd \cdot v$ where Qd is the charge density per unit length and given by $WC_{ox}(VGS - VTH)$, we have
- $$ID = WC_{ox}(V_{GS} - V_{TH})v_{sat}$$
 - Extreme velocity saturation creates three departures from square-law behavior
 - ID is linearly proportional to overdrive and independent of channel length
 - ID reaches saturation even for $VDS < VGS - VTH$
 - gm of a fully velocity-saturated MOSFET emerges to be relatively constant

$$\begin{aligned} g_m &= \frac{\partial I_D}{\partial V_{GS}} \Big|_{V_{DSconst}} \\ &= WC_{ox}v_{sat}, \end{aligned}$$

Deep-Submicron Effects

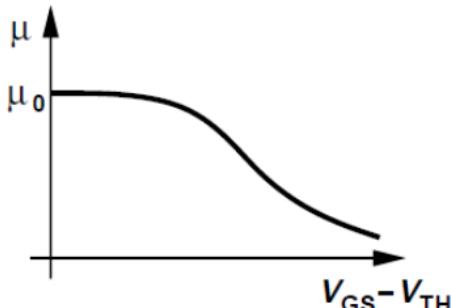
- Velocity Saturation:



- Above figure shows that I_D saturates even for $V_{DS} < V_{GS} - V_{TH}$ due to velocity saturation
- Knee points occur at relatively small V_{DS} 's even for moderate overdrives

Deep-Submicron Effects

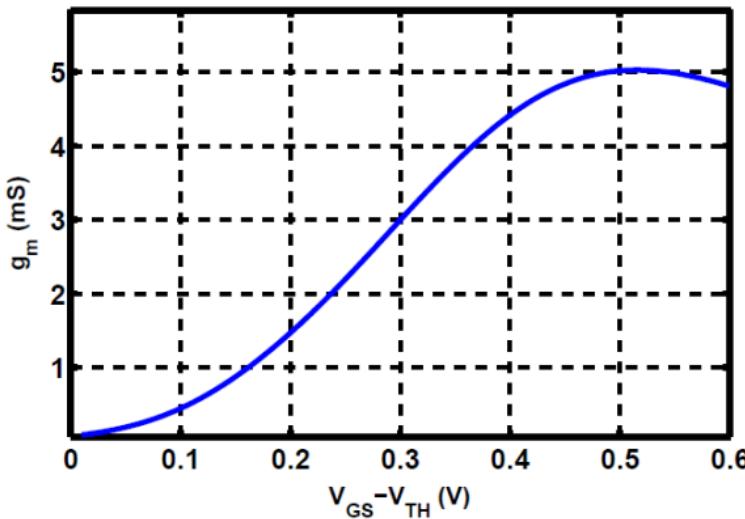
- Mobility Degradation with Electric Field:
- Mobility of charge carriers in the channel also declines as the gate-source voltage and the vertical field increase



- We intuitively expect that gm no longer follows the linear relationship $gm = \mu C_{ox}(W/L)(V_{GS} - V_{TH})$, with the overdrive voltage

Deep-Submicron Effects

- Mobility Degradation with Electric Field:



- Above figure shows the nonlinear relationship between g_m and $V_{GS} - V_{TH}$ for the 5 $\mu\text{m}/40$ nm NFET device

Transconductance Scaling

- Suppose a transistor operates in the saturation region but does not provide the required transconductance, from gm equations,

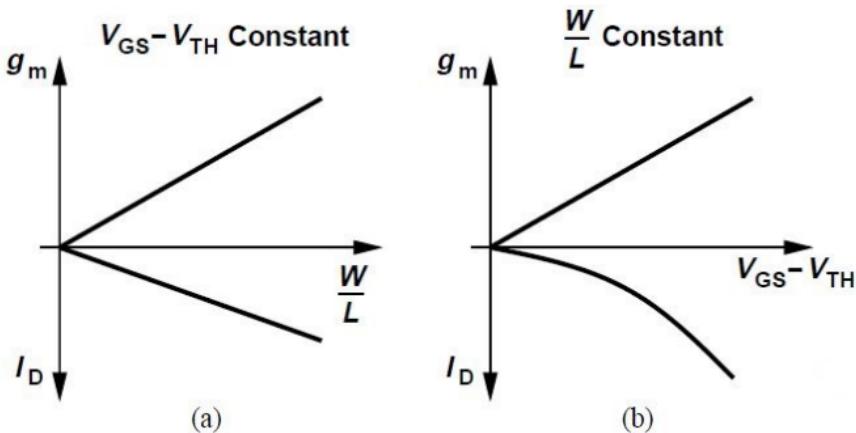
$$\begin{aligned} g_m &= \mu_n C_{ox} \frac{W}{L} (V_{GS} - V_{TH}) \\ &= \sqrt{2\mu_n C_{ox} \frac{W}{L} I_D} \\ &= \frac{2I_D}{V_{GS} - V_{TH}}, \end{aligned}$$

- Adjustment in three parameters, namely, W/L , $V_{GS} - V_{TH}$, or I_D can scale gm
- Assume a long-channel device and hence

$$I_D \approx (1/2)\mu_n C_{ox} (W/L) (V_{GS} - V_{TH})^2$$

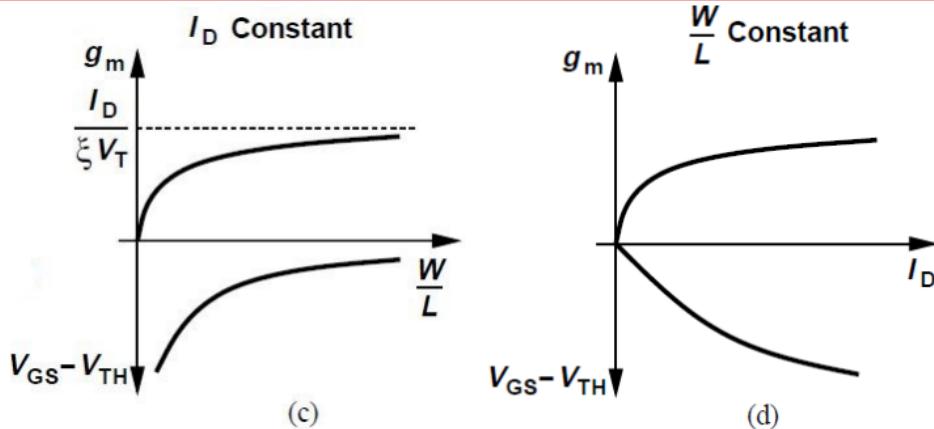
$$V_{GS} - V_{TH} \approx \sqrt{2I_D / (\mu_n C_{ox} W/L)}$$

Transconductance Scaling



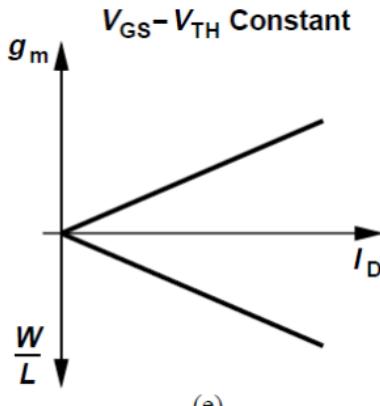
- If W/L is increased while keeping $V_{GS} - V_{TH}$ constant, both gm and ID linearly scale with W/L and so does the power consumption [Fig. (a)]
- If $V_{GS} - V_{TH}$ is increased but keep W/L constant, thus requiring a higher drain current [Fig. (b)]
- In the former case, device capacitances rise whereas in the latter, $V_{DS,min}$ increases

Transconductance Scaling

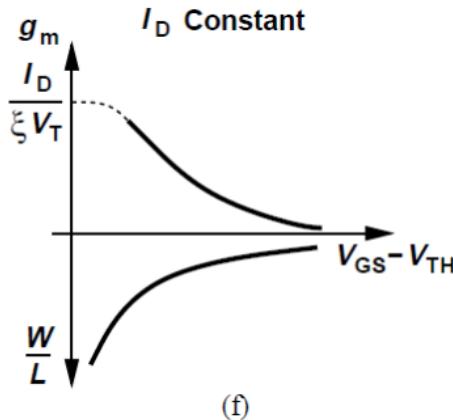


- If W/L is increased with ID constant, $V_{GS} - V_{TH}$ is lowered [Fig. (c)]
- gm does not climb indefinitely however, due to subthreshold conduction
- If W/L is held constant and ID is increased, then $V_{GS} - V_{TH}$ and hence $V_{DS,min}$ rises

Transconductance Scaling



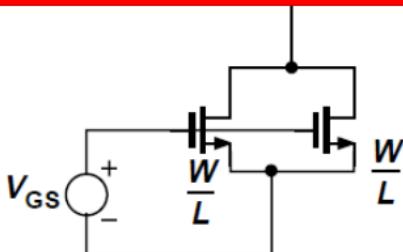
(e)



(f)

- If I_D is increased while $V_{GS} - V_{TH}$ is constant, W/L is needed to increase [Fig. (e)]
- Alternatively, if $V_{GS} - V_{TH}$ is lowered with I_D kept constant, W/L must increase [Fig. (f)]
- For $V_{GS} - V_{TH} \approx 0$, device enters subthreshold region and $g_m \approx I_D/(\xi V_T)$
- Device capacitances increase in both cases

Transconductance Scaling - Example



- Linear scaling of gm and ID with W/L holds regardless of transistor characteristics
- Consider two identical transistors are connected in parallel, each with transconductance gm
- If V_{GS} changes by ΔV , then drain current of each device changes by $gm\Delta V$, hence current of composite device changes by $2gm\Delta V$; parallel combination exhibits a transconductance $2gm$
- Scaling preserves device “current density” (ID/W)
- Bias overdrive voltage and gm/ID ratio remains constant

Transistor Design

- A given transistor in a circuit is characterized by many parameters
- Assume transistor operates in saturation
- Interested in:
 - Bias quantities: ID and $VGS - VTH (=VDS,min)$
 - Small-signal parameter: gm
 - Physical parameter: W/L
- Typical transistor design problem:
 - Given two of gm , ID and $VGS - VTH$, determine the other two parameters

Transistor Design

	Case I	Case II	Case III
Given	$I_D, V_{DS,min}$	g_m, I_D	$g_m, V_{DS,min}$
To Be Determined	$\frac{W}{L}, g_m$	$\frac{W}{L}, V_{DS,min}$	$\frac{W}{L}, I_D$
Design Revision	g_m insufficient; Raise I_D and $\frac{W}{L}$	$V_{DS,min}$ too large; Raise $\frac{W}{L}$	I_D too large; Raise $\frac{W}{L}$; Lower $V_{GS} - V_{TH}$

- Design problems shown above are “overconstrained”
- Two given parameters lead to certain values for the other two, even though results may not always be desirable
- Design revisions are necessary in such a case

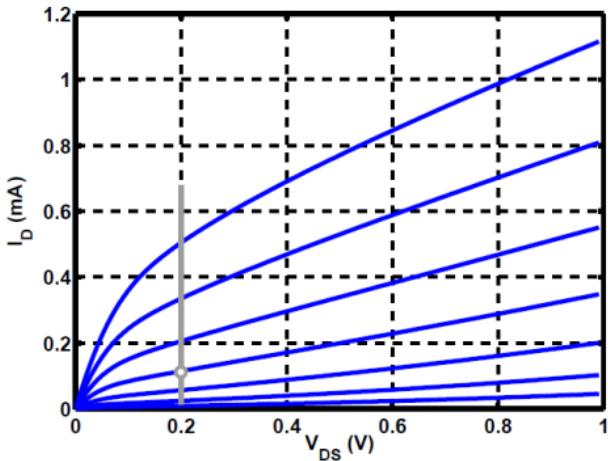
Design for given ID and VDS,min

- Suppose for a given transistor, we have chosen:
 - A bias current (according to a power budget)
 - Minimum VDS (according to voltage headroom restrictions imposed by supply voltage and output swing requirements)
- We wish to determine the dimensions and transconductance of the device, recognizing that square-law models are inaccurate
- Three-step approach
- Consider $ID = 0.5 \text{ mA}$ and $VDS,min = 200 \text{ mV}$ as an example

Design for given ID and VDS,min

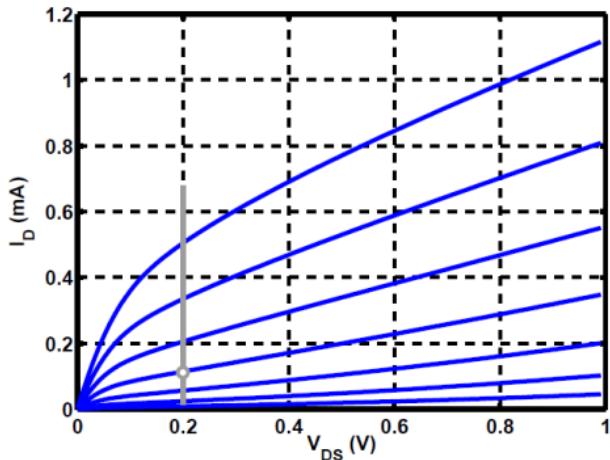
- **Step 1:** Select a reference transistor, with a width $WREF$ and length equal to minimum allowable value, L_{min} (e.g., $L_{min} = 40$ nm). Choose $WREF = 5 \mu\text{m}$ as an example
- **Step 2:** Using actual device models and a circuit simulator, plot $ID-VDS$ characteristics of the reference transistor for different values of $VGS - VTH$
 - Typically $VGS - VTH$ ranges from 50 mV to 600 mV
 - Can construct the characteristics with the overdrive incrementing in steps of 50 mV

Design for given ID and VDS,min



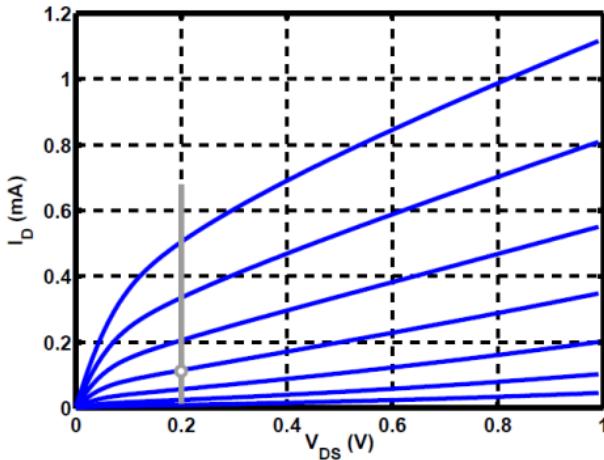
- Above figure shows results for $WREF/Lmin = 5 \mu\text{m}/40 \text{ nm}$ and overdrive ranging from 50 mV to 350 mV
- Step 3: Since in our example, $ID = 0.5 \text{ mA}$ and $VDS,min = 200 \text{ mV}$, we draw a vertical line at $VDS = 200 \text{ mV}$ and find its intersection with the plots
- If device obeyed square law, we would choose the plot for $VGS - VTH = VDS,min = 200 \text{ mV}$

Design for given ID and VDS,min



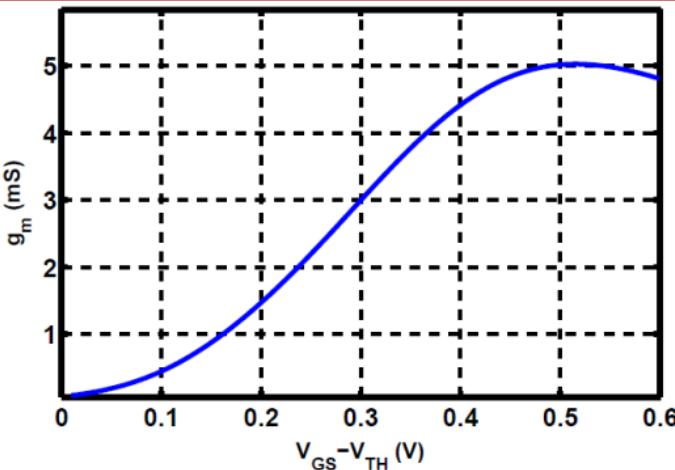
- **Step 3 (contd.):** Short-channel device remains in saturation even for $VGS - VTH = 350$ mV and $VDS = 200$ mV
 - Situation is more complex, proceed with $VGS - VTH = 200$ mV for now
- This procedure yields one operating point for the reference transistor satisfying VDS requirement

Design for given ID and $V_{DS,min}$



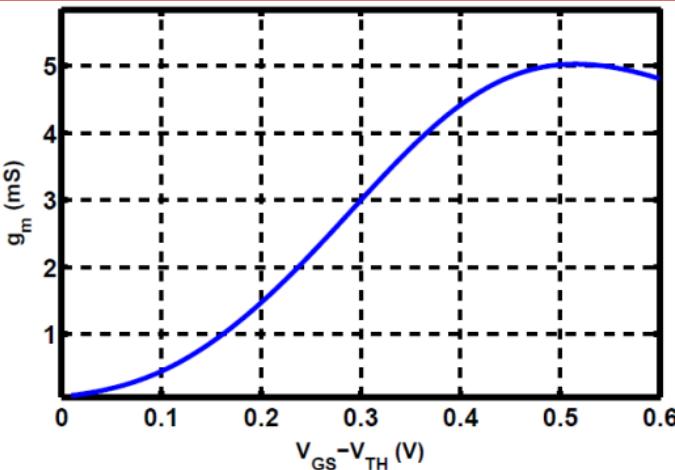
- Step 4: The drain current, ID,REF is not close to the necessary value, 0.5 mA, in our example
- We must scale the width and hence the drain current of the transistor
- Since $ID,REF \approx 100 \mu\text{A}$ above, we choose a transistor width of $(500 \mu\text{A} / 100 \mu\text{A}) \times WREF = 5WREF = 25 \mu\text{m}$

Design for given ID and VDS,min



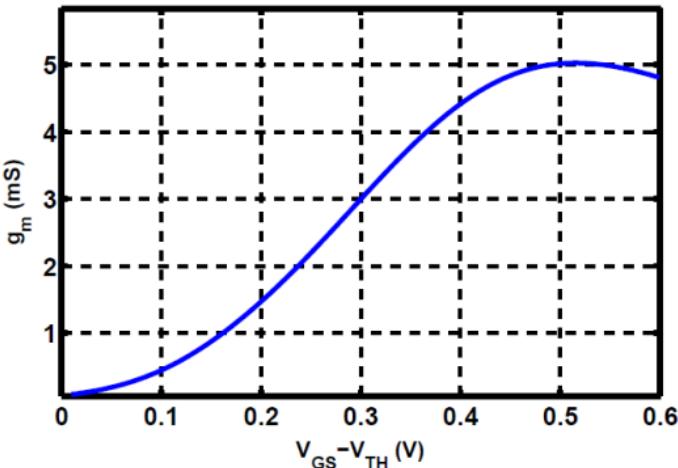
- Step 4 (contd.): From ID - VDS plots, gm of the reference transistor can be approximated from $gm = \Delta ID / \Delta VGS \approx 2 \text{ mS}$
- To get a more accurate value of gm , we plot gm versus $VGS - VTH$ for $VDS = 200 \text{ mV}$ using simulations as shown above
- This predicts $gm = 1.5 \text{ mS}$ for $VGS - VTH = 200 \text{ mV}$

Design for given ID and $V_{DS,min}$



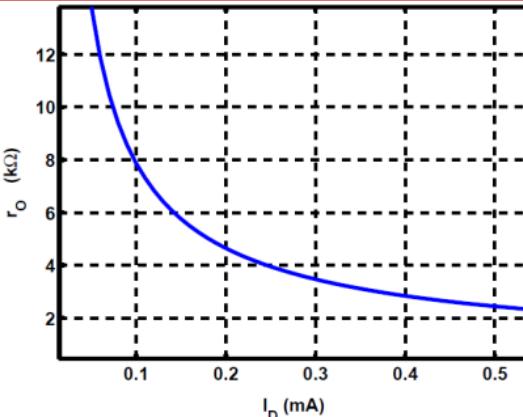
- **Step 4 (contd.):** If both width and drain current are scaled up by a factor of 5, gm increases to 7.5 mS
- If this is insufficient, W/L must be increased further
- We can perform scaling to determine width and transconductance of other transistors in the circuit, using ID and gm plots for the reference device

Design for given ID and $V_{DS,min}$



- **Step 4 (contd.):** We typically choose $V_{GS} - V_{TH} \approx V_{DS,min}$ to obtain a higher g_m even though it translates to a wider transistor

Design for given ID and $V_{DS,min}$



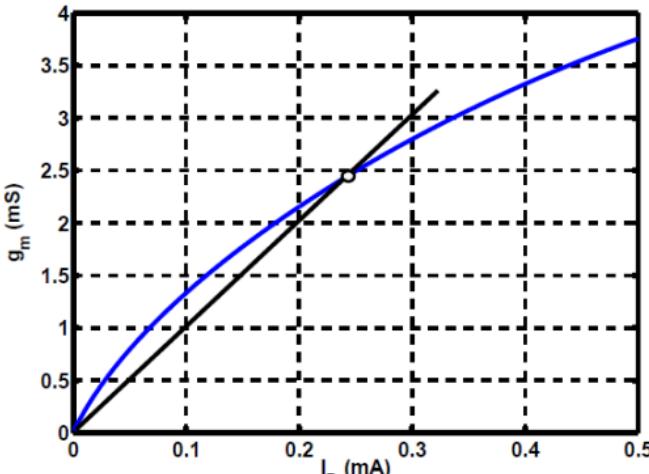
- Output impedance r_O of short-channel devices cannot be expressed as $1/(\lambda ID)$
- We use simulations to plot r_O for the reference device as a function of ID as shown above
- Reference transistor carries a current of $100 \mu\text{A}$, exhibiting $r_O = 8 \text{ k}\Omega$
- When width and drain current are scaled up by factor of 10, r_O falls by the same factor to 800Ω

Design for given gm and ID

- In many analog circuits, a given transistor must provide sufficient transconductance while consuming minimal power
- Suppose we are given a specified transconductance gm_1 and an upper limit for the drain bias current ID_1 , seeking the corresponding values of W/L and $VGS - VTH$
- Assume $gm_1 = 10 \text{ mS}$ and $ID_1 = 1 \text{ mA}$
- To determine whether gm_1 can be obtained with $ID \leq ID_1$, we note that maximum gm occurs in the subthreshold region and is given by $ID/(\xi VT)$ where $\xi = 1.5$
- For $ID = 1 \text{ mA}$, gm cannot exceed 26 mS at room temperature
- Since $gm_1 < ID_1/(\xi VT)$, this design is feasible

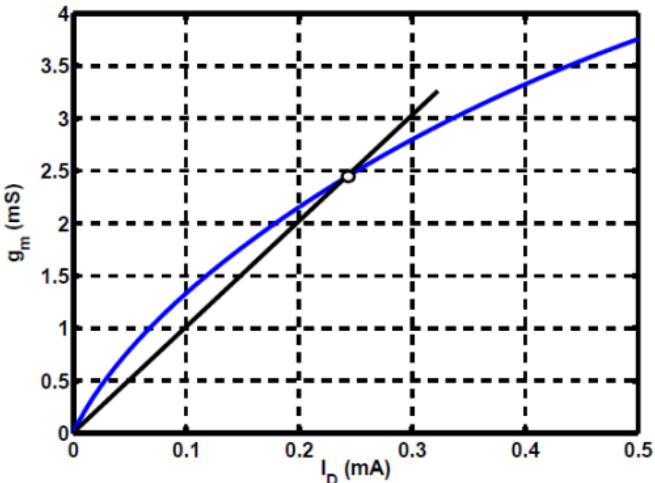
Design for given gm and ID

- Step 1: Using simulations, we plot gm as a function of ID for a reference transistor, e.g., $WREF/Lmin = 5 \mu\text{m}/40 \text{ nm}$



- Step 2: We , , , , the gm - ID plane and draw a line through the origin and this point, obtaining the intersection at $(ID,REF,gm,REF) = (240 \mu\text{A}, 2.4 \text{ mS})$ and a corresponding overdrive

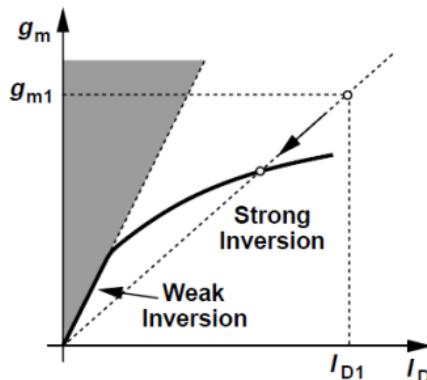
Design for given gm and ID



- Step 3: We multiply $WREF$ by $gm1/gm, REF = 4.2$ so as to travel on the straight line to point $(ID1, gm1)$ while maintaining the same overdrive
- This completes the transistor design

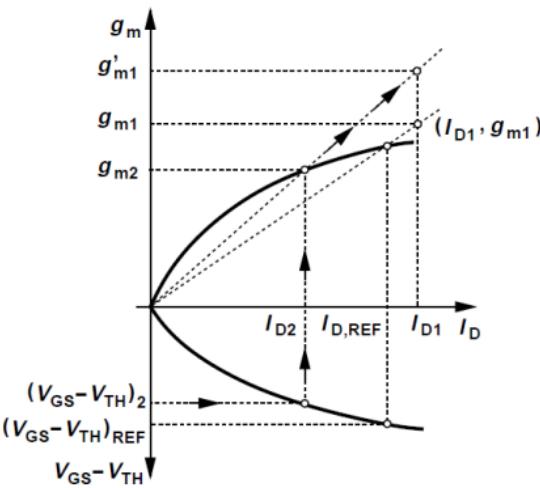
Design for given gm and ID

- Above procedure elicits two questions
- Question 1: Does the straight line passing through the origin and (ID_1, gm_1) always intersect the $gm-ID$ plot?
- For a square-law device in strong inversion,
 $gm = \sqrt{2\mu_n C_{ox} (W/L) ID}$ has a slope of infinity at the origin, guaranteeing an intersection point
- In the subthreshold region, on the other hand,
 ~~$gm \propto ID$~~ which means the (ID, gm) combinations in the gray region are not achievable



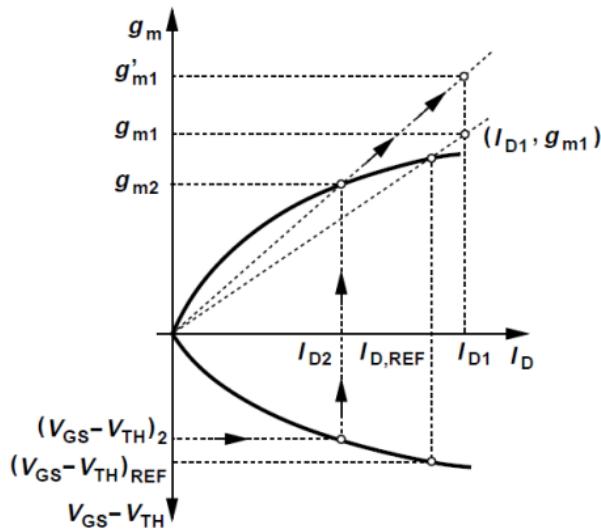
Design for given gm and ID

- **Question 2:** If $(VGS - VTH)REF$ is excessively large, then by factor must W be increased?



- In above figure, $(VGS - VTH)2 < (VGS - VTH)REF$ is desired
- We then find the corresponding current $ID2$ and transconductance $gm2$ on the $gm-ID$ plane

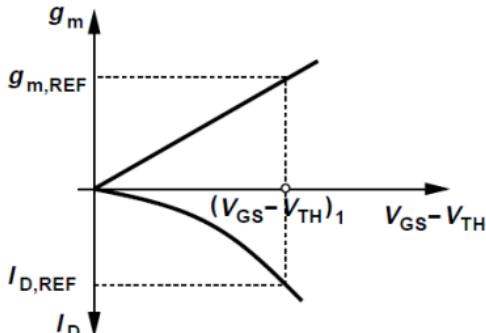
Design for given gm and ID



- Next, we draw a line through the origin and the point (ID_2, gm_2) continue to $ID = ID_1$, i.e., we multiply W_{REF} by ID_1/ID_2
- The resulting width guarantees an overdrive of $(V_{GS} - V_{TH})_2$ at a drain current of ID_1 and provides a transconductance of at least gm_1

Design for given gm and $V_{DS,min}$

- In some designs, the transconductance is dictated by some performance requirements (voltage gain, noise, etc.) and the minimum V_{DS} by the voltage headroom, without an explicit specification of ID
- Step 1: We use simulations to plot gm as a function of $V_{GS} - V_{TH}$ for the reference transistor
- Now we select $(V_{GS} - V_{TH})_1 = V_{DS,min}$ and obtain the corresponding transconductance gm,REF
- It is helpful to plot ID on the same plane and find ID,REF at $(V_{GS} - V_{TH})_1$



Design for given gm and VDS,min

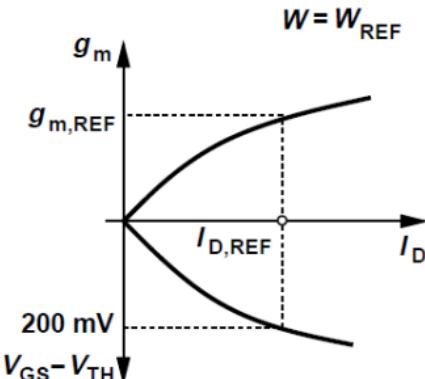
- **Step 2:** To reach the required transconductance gm_1 , we scale the transistor width by a factor of $gm_1/gm,REF$
 - ID scales by the same factor
- If the resulting ID is too large, we redesign for a given gm and ID
- Device is now smaller and has a smaller transconductance

Design for given gm

- Suppose a design problem only specifies the transconductance, and we wish to compute the remaining parameters
- Two scenarios may be envisaged
- We select a certain W/L and raise ID until the desired transconductance, gm_1 is reached
 - Required ID and hence power consumption may be excessive
 - Overdrive voltage may be excessively large
- We select a reasonable value for ID (perhaps according to a power budget) and increase W/L to obtain gm_1
- Increasing W/L (hence decreasing VGS) eventually drives device into subthreshold region, where gm cannot exceed $ID/(\xi VT)$

Design for given gm

- First, we construct plots representing gm and $VGS - VTH$ as a function of ID for the reference device, using simulations
 - Here VDS is kept constant and approximately equal to $VDD/2$
- We select a reasonable value for $VGS - VTH$, e.g., 200 mV, which points to ID,REF and gm,REF
- Next, we scale the width and drain current by a factor of $gm1/gm,REF$



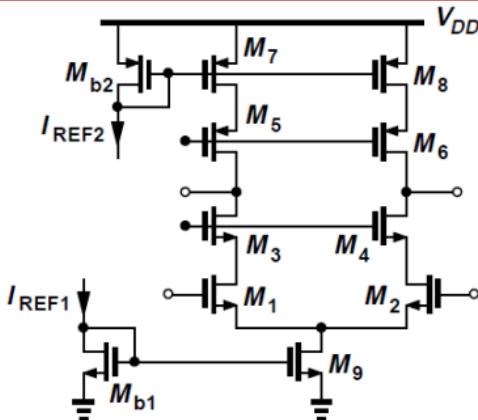
Choice of Channel Length

- If the selection of ID , $VGS - VTH$ and gm does not yield a sufficiently high rO , we must increase the transistor length
- To maintain the same drain current, overdrive and gm , the width must also be scaled proportionally
- Such scaling is not straightforward
 - If drawn length increases from $Lmin$ to $2Lmin$, effective length increases from $Lmin - 2LD$ to $2Lmin - 2LD$
- We must use simulations to construct the $ID-VDS$, gm and rO characteristics for several channel lengths, e.g., 60 nm, 80 nm, 100 nm (drawn values)

Op Amp Design Examples

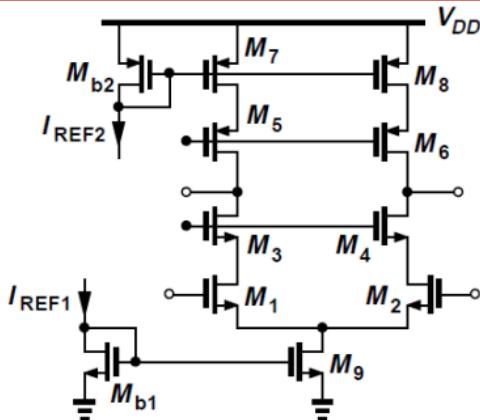
- Consider an op-amp design example in 40-nm technology
- We target the following specifications
 - Differential output voltage swing = 1 Vpp
 - Power consumption = 2mW
 - Voltage Gain = 500
 - Supply Voltage = 1 V
- Single-ended output swing of 0.5 Vpp is small enough to make telescopic or folded-cascode op-amps a plausible choice
- Begin with minimum allowable width and length devices unless otherwise dictated by current, transconductance, VD_{sat} , output resistance, etc

Design 1: Telescopic Op-Amp



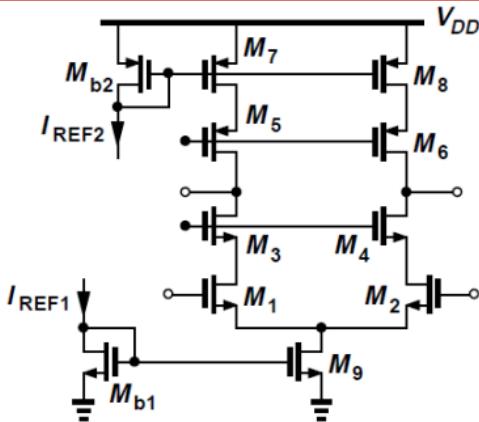
- Consider the telescopic op-amp topology shown above to meet previously mentioned requirements
- Of the total supply current of 2 mA, we allocate 50 μ A each to I_{REF1} and I_{REF2} , and 0.95 mA for each branch of the differential pair
- To accommodate a single-ended peak-peak swing of 0.5 V, we must distribute the remaining 0.5 V over $M9$, $M1,2$, $M3,4$, $M5,6$ and $M7,8$, allowing 100 mV for each

Design 1: Telescopic Op-Amp



- With the bias currents and overdrives known, we can determine W/L's by examining I/V characteristics
- For L = 40 nm, the intrinsic gain $gmrO$ of NMOS devices is around 7-10 and for PMOS devices is 5-7
- It is difficult to raise $gmrO$ beyond 10 for PFETs
- If we approximate gm as $2ID/(VGS - VTH) = 2 \times 0.95$ mA/100 mV = 19 mS, we estimate $rO \approx 530 \Omega$ from $gmrO \approx 10$

Design 1: Telescopic Op-Amp

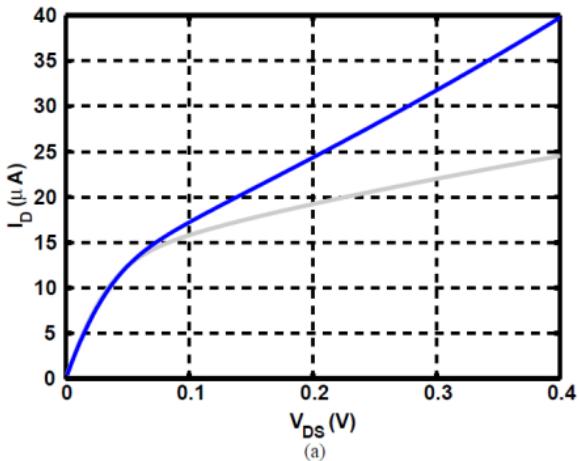


- If $gm_{1,2} \approx 19 mS$, then for the gain GmR_{out} to reach 500, the op-amp output impedance must exceed $26 k\Omega$
- However, with $gm_{3,4}rO_{3,4} \approx 10$ and $rO_{7,8} \approx 530 \Omega$, we have $(gm_{3,4}rO_{3,4})rO_{1,2} \approx 5.3 k\Omega$, obtaining a voltage gain of only 100 even if the PMOS devices have $\lambda = 0$
- Telescopic arrangement is impractical for a gain of 500

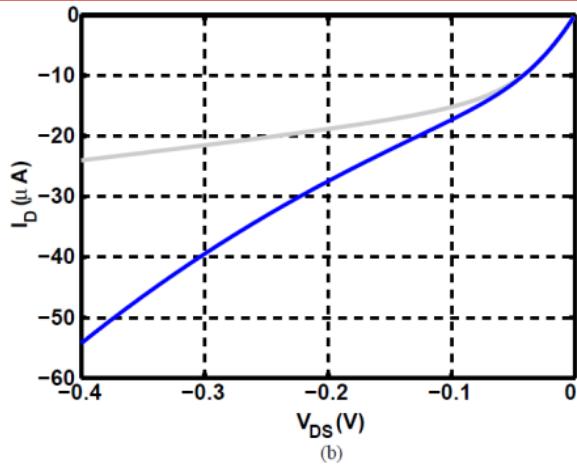
Design 1: Telescopic Op-Amp

- We still continue with the design out of curiosity
- Using simulations, we construct the I/V characteristics of NMOS and PMOS devices with $L = 40$ nm and 80 nm, since minimum length may exhibit unacceptably low rO and $gmrO$
- Simulation parameters must ensure that devices remain in saturation for $|VDS| \geq 100$ mV
- Since threshold and overdrive elude clear definitions in nanometer technologies, we must adjust VGS in simulations to ensure saturation

Design 1: Telescopic Op-Amp



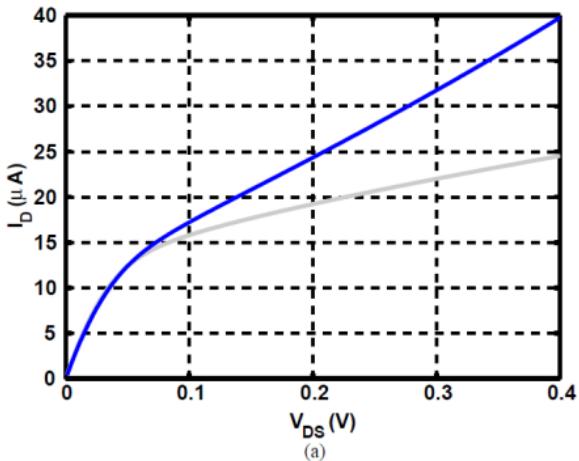
(a)



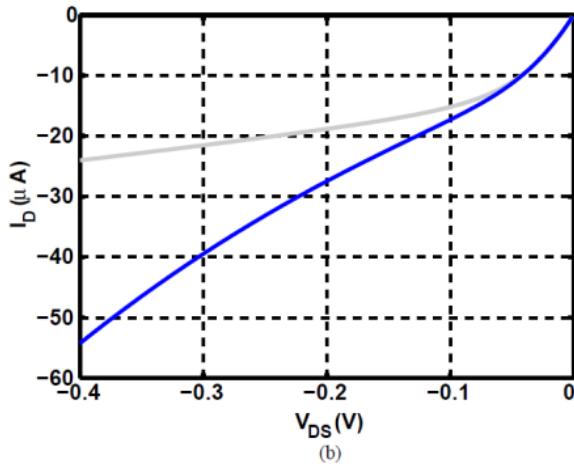
(b)

- Fig. (a) plots ID - V_{DS} curves for $(W/L)N = 5 \mu\text{m}/40 \text{ nm}$ (blue) and $10 \mu\text{m}/80 \text{ nm}$ (gray) with $V_{GS} = 300 \text{ mV}$
- Fig. (b) plots ID - V_{DS} curves for $(W/L)P = 5 \mu\text{m}/40 \text{ nm}$ (blue) and $5 \mu\text{m}/80 \text{ nm}$ (gray) with $V_{GS} = -400 \text{ mV}$
- Difficult to distinguish between triode and saturation regions, especially for PFETs

Design 1: Telescopic Op-Amp



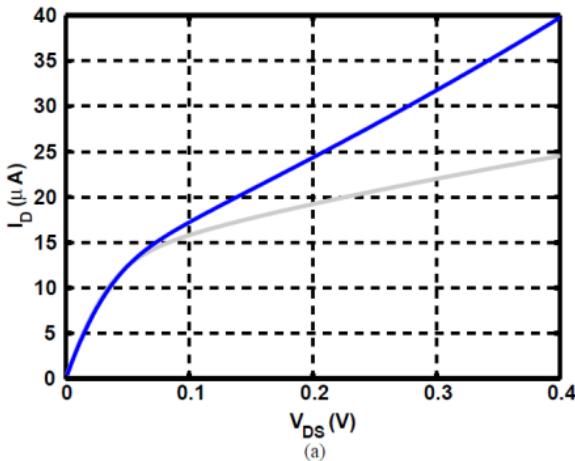
(a)



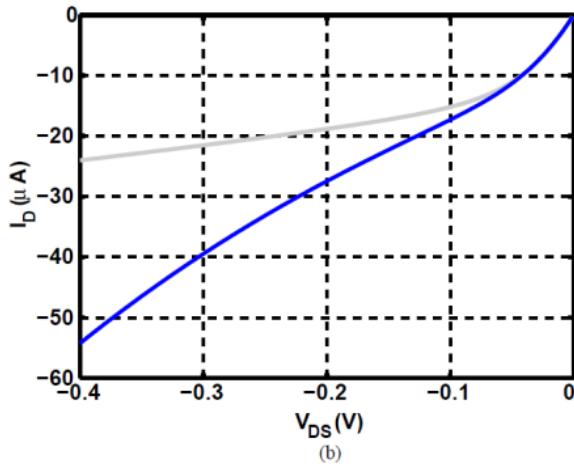
(b)

- 40-nm PMOS device displays a decreasing output impedance as $|V_{DS}|$ reaches 400 mV
- For other three characteristics, we can roughly identify a “knee” point beyond which slope falls considerably
 - VGS chosen to place this point below $|V_{DS}| = 100$ mV

Design 1: Telescopic Op-Amp



(a)

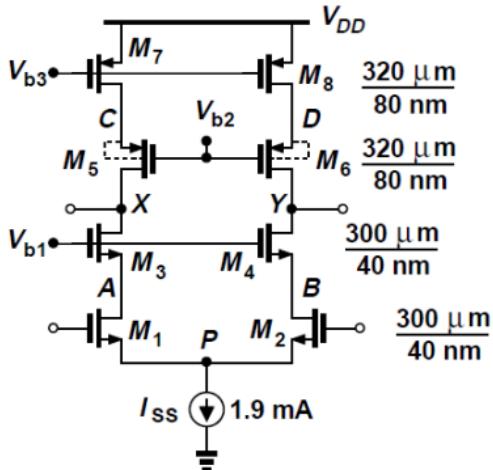


(b)

- At $V_{DS} = 100 \text{ mV}$, the $10 \mu\text{m}/80 \text{ nm}$ NMOS provides r_O of $22.8 \text{ k}\Omega$ and ID of $16 \mu\text{A}$
- If scaled up to carry $950 \mu\text{A}$, r_O drops to 385Ω
- Similarly, the $5 \mu\text{m}/80 \text{ nm}$ PMOS has an r_O of $18.45 \text{ k}\Omega$ at $V_{DS} = -100 \text{ mV}$ with $ID = 15 \mu\text{A}$, thus offering $r_O = 290 \Omega$ if scaled up to carry $950 \mu\text{A}$
- These low values of r_O are discouraging

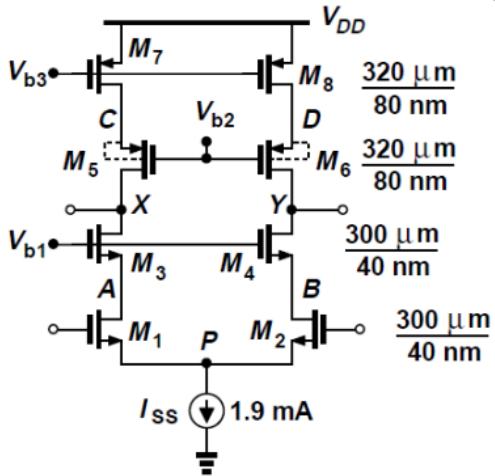
Design 1: Telescopic Op-Amp

- We scale the NMOS and PMOS device widths to accommodate a drain current of $950 \mu\text{A}$ with $V_{GS,N} = 300 \text{ mV}$, $V_{GS,P} = -400 \text{ mV}$ and $|V_{DS}| = 100 \text{ mV}$, resulting in the design shown below



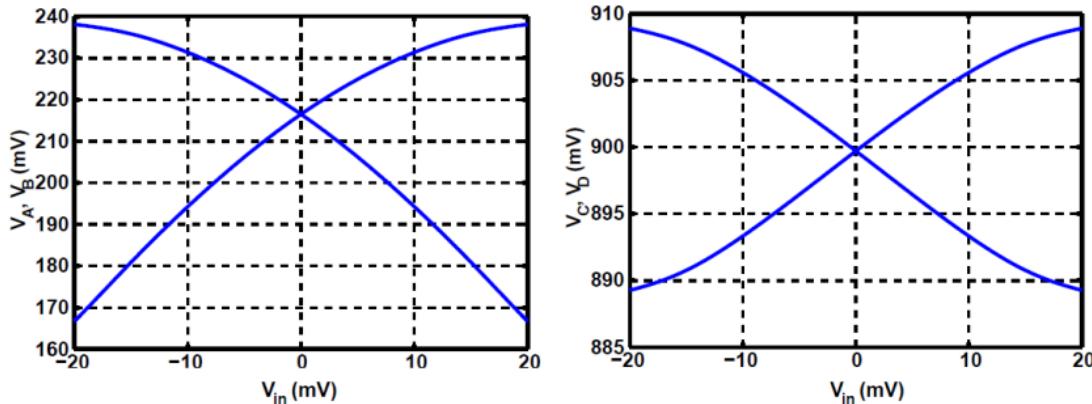
- Minimum-length signal path to maximize speed and minimize capacitances

Design 1: Telescopic Op-Amp



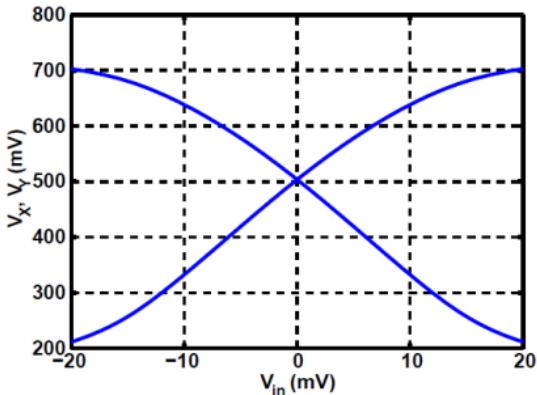
- Bias voltages are tentatively chosen as follows:
- Input common-mode level $V_{CM,in} = 100 \text{ mV}$ for the tail current source plus $V_{GS1,2}$ ($= 300 \text{ mV}$)
- $V_{b1} = V_{D1,2} (= 200 \text{ mV}) + V_{GS3,4} (= 300 \text{ mV})$
- $V_{b2} = V_{DD} - |V_{DS7,8}| - |V_{GS5,6}|$
- $V_{b3} = V_{DD} - |V_{GS7,8}|$

Design 1: Telescopic Op-Amp



- We perform a dc sweep of differential input voltage V_{in}
- Drain voltages of $M1$ and $M2$ (V_A, V_B) are around 220 mV in the middle of the range
- Drain voltages of $M7$ and $M8$ (V_C, V_D) are close to targeted value

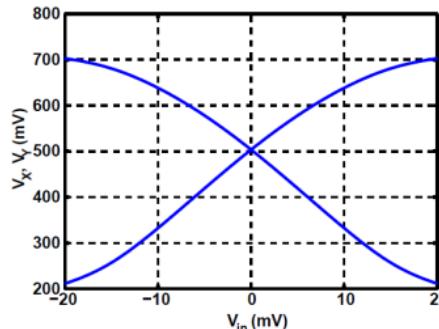
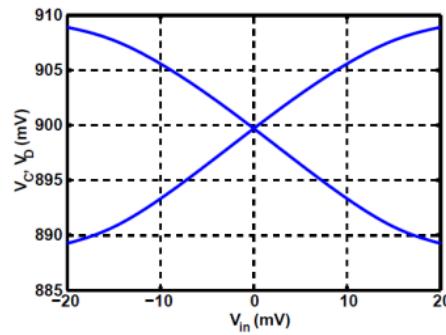
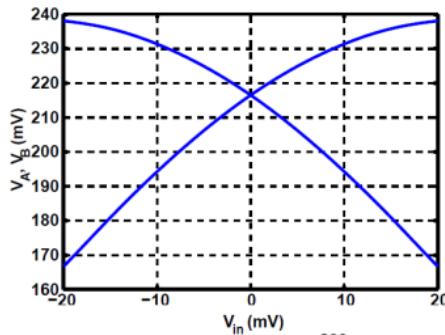
Design 1: Telescopic Op-Amp



- Output behavior shown above (V_X, V_Y)
- Slope of each single-ended output is approximately 15 near $V_{in} = 0$, yielding a differential gain of 30
- Characteristic becomes nonlinear as each output rises towards 0.7 V

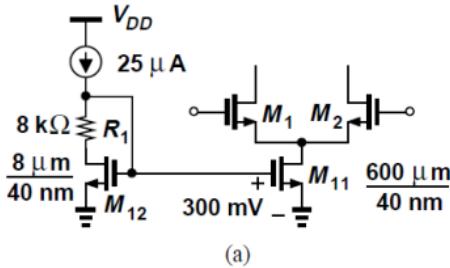
Design 1: Telescopic Op-Amp

- Gain can be raised by increasing $(W/L)_{3,4}$ to 600 $\mu\text{m}/80 \text{ nm}$, characteristics plotted below
- Characteristics exhibit a gain of about 54 but a limited output swing

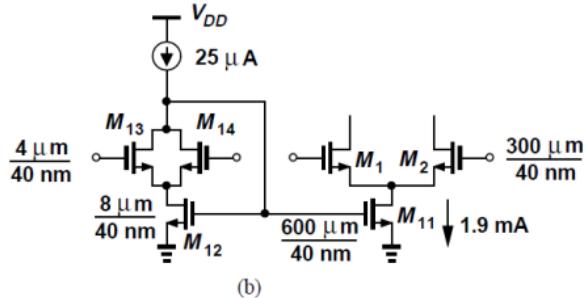
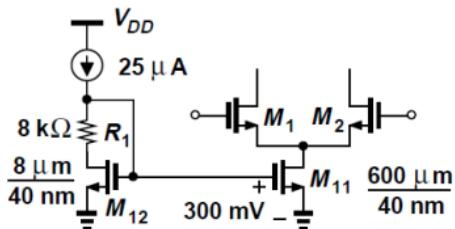


Design 1: Telescopic Op-Amp- Bias Circuit

- Op-amp circuit relies on proper choice of bias quantities ISS , V_{b1} , V_{b2} and V_{b3}
- ISS and V_{b3} must be established by current mirror action and V_{b2} by low-voltage cascode
- V_{b1} requires a different approach
- We begin with $ISS = 1.9 \text{ mA}$, choosing a channel length of 40 nm and scaling from reference device curves, a width of 600 μm for $V_{DS} = 100 \text{ mV}$
- With reference current budget of $25 \mu\text{A}$, we arrive at the arrangement of Fig. (a), with W_{12} scaled down from W_{11} by a factor of $1.9 \text{ mA}/25 \mu\text{A}$



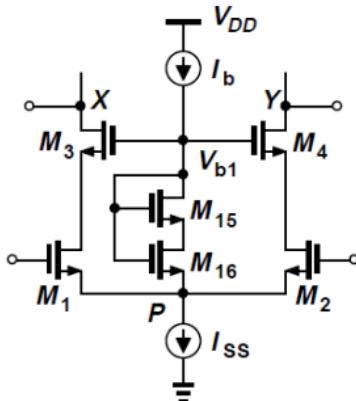
Design 1: Telescopic Op-Amp- Bias Circuit



- Since M_{11} operates with a V_{DS} of 100 mV, R_1 is inserted in series with drain of M_{12} such that $V_{DS12} = V_{GS12} - VR_1 = 100 \text{ mV}$
- Design of Fig. (a) is sensitive to CM level since $V_{DS11} = V_{CM,in} - V_{GS1,2}$ whereas $V_{DS12} = V_{GS1,2} - VR_1$
- Drain voltage of M_{12} must track $V_{CM,in}$
- This can be done as in Fig. (b) where R_1 is replaced by a differential pair driven by V_{in1} and V_{in2}
- With proper width scaling, we have $V_{GS13,14} = V_{GS1,2}$ and hence $V_{DS12} = V_{DS11}$

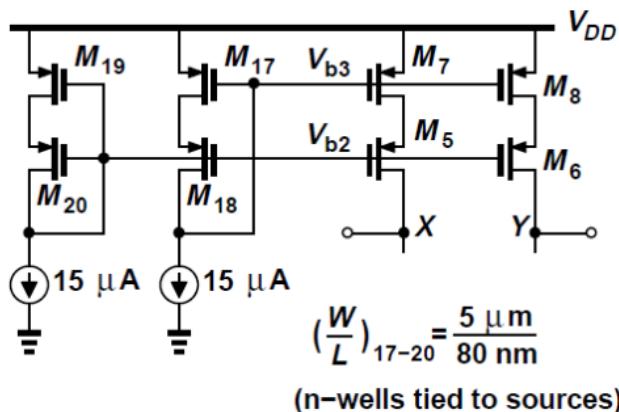
Design 1: Telescopic Op-Amp- Bias Circuit

- In op-amp circuit designed, V_{b1} must equal $V_{GS3,4} + V_{DS1,2} + V_P$, where $V_{DS1,2} = 100$ mV
- Diode-connected device in series with a drain-source voltage added to V_P can produce V_{b1}
- I_b must be much less than I_{SS}
- We select $I_b = 15 \mu\text{A}$ and hence $(W/L)_{15,16} = 10 \mu\text{m}/80 \text{ nm}$
- If $V_{CM,in}$ rises, so do V_P and hence V_{b1} , keeping $V_{DS1,2}$ constant



Design 1: Telescopic Op-Amp- Bias Circuit

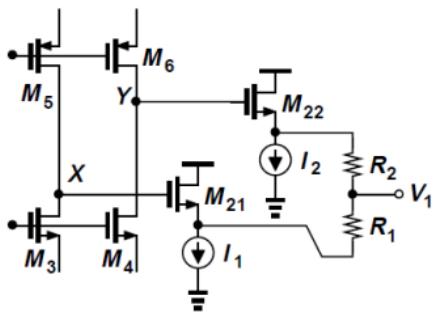
- In order to generate V_{b3} and V_{b2} , a low-voltage cascode is constructed as shown below
- $M17$ and $M18$ are scaled down respectively down from $M7,8$ and $M5,6$, ensuring that $V_{DS17} = V_{DS7,8}$
- To create $V_{b2} = VDD - |V_{DS7,8}| - |V_{GS5,6}|$, a diode-connected device $M20$ is connected in series with V_{DS} produced by $M19$



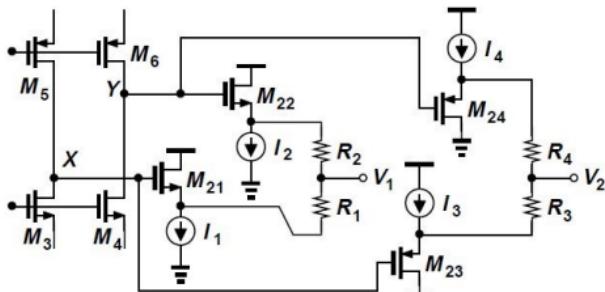
Design 1: Telescopic Op-Amp- Common-Mode Feedback

- Due to various mismatches, PMOS currents are not exactly $ISS/2$, forcing the output CM level to VDD or ground and hence requiring CMFB
- Output CM level, VCM must be sensed and fed to the NMOS or PMOS current sources
- CM level can be sensed by resistors, triode transistors or source followers
- High output impedance of op-amps dictates very large resistors and tight voltage margins demand precise CM control and preclude triode devices
- Source followers cannot measure CM level across a wide output swing

Design 1: Telescopic Op-Amp- Common-Mode Feedback



(a)



(b)

- In Fig. (a), if V_X (or V_Y) falls, I_1 (or I_2) eventually collapses, disabling the source follower
- Complement NMOS followers by PMOS counterparts
- In Fig. (b), PMOS followers M_{23} and M_{24} also sense the output CM level and drive R_3 , R_4 respectively
- Here

$$V_1 = V_{CM} - V_{GS21,22}$$

$$V_2 = V_{CM} + |V_{GS23,24}|$$

Design 1: Telescopic Op-Amp- Common-Mode Feedback

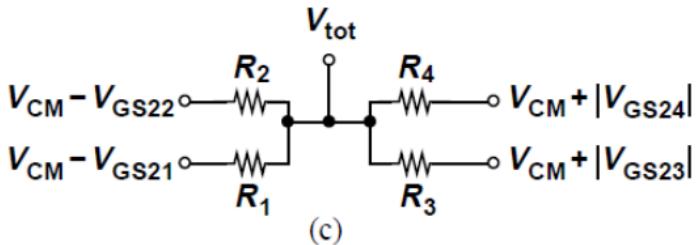
- A linear combination of V_1 and V_2 can remove V_{GS} terms, yielding a value proportional to V_{CM}
- If $\alpha V_1 + \beta V_2 = (\alpha + \beta)V_{CM} - \alpha V_{GS21,22} + \beta |V_{GS23,24}|$, then we must choose $\alpha V_{GS21,22} = \beta |V_{GS23,24}|$, obtaining $\alpha V_1 + \beta V_2 = (\alpha + \beta)V_{CM}$
- We can choose $\alpha + \beta = 1$ so that reconstructed value is equal to actual op-amp output CM level
- α and β can be implemented by $R1-R4$
- If V_1 and V_2 are shorted, the weighted sum of V_1 and V_2 is produced

Design 1: Telescopic Op-Amp- Common-Mode Feedback

- Using Fig. (c) below, it can be shown that

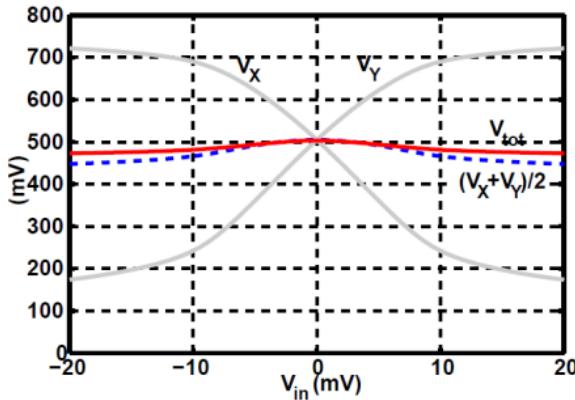
$$V_{tot} = V_{CM} + \frac{R_N |V_{GS23,24}| - R_P V_{GS21,22}}{R_N + R_P}$$

- Where $R_N = R_1 = R_2$ and $R_P = R_3 = R_4$
- We therefore choose $R_N/R_P = V_{GS21,22}/|V_{GS23,24}|$



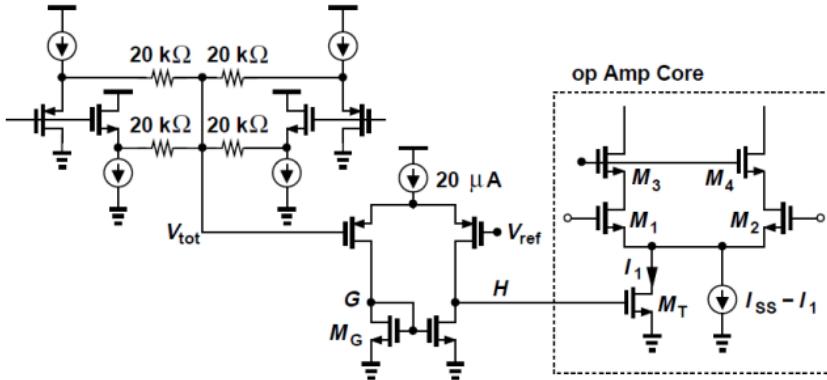
Design 1: Telescopic Op-Amp- Common-Mode Feedback

- To evaluate feasibility of this idea, we run a dc sweep and observe the behavior of V_{tot}
- We select a bias current of $10 \mu\text{A}$ and $W/L = 10 \mu\text{m}/40 \text{ nm}$ for all source followers and $R_N = R_P = 20 \text{ k}\Omega$
- The $10\text{-}\mu\text{A}$ current sources are also implemented as $10 \mu\text{m}/40 \text{ nm}$ transistors
- Figure below plots the actual CM level, defined as $(V_X + V_Y)/2$ and the reconstructed counterpart, V_{tot}



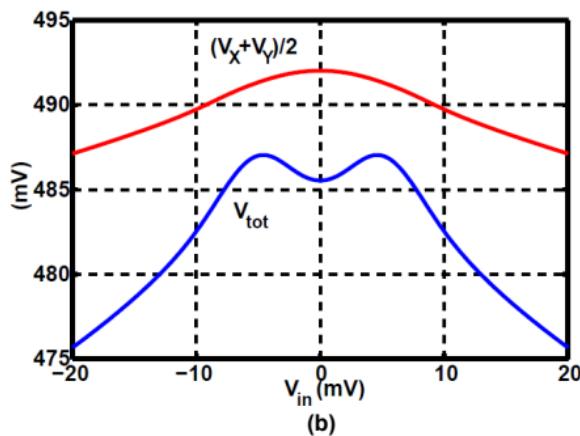
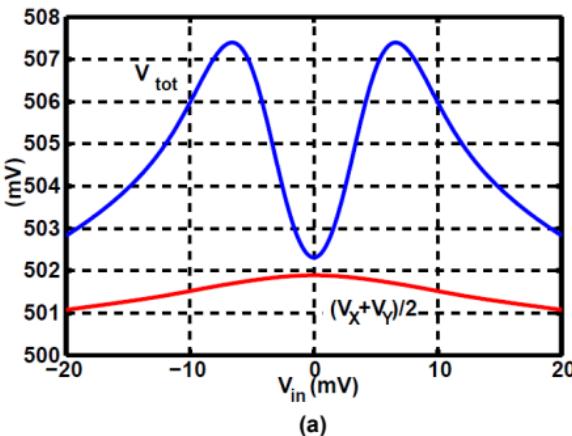
Design 1: Telescopic Op-Amp- Common-Mode Feedback

- In the next test, the CMFB loop is closed: we compare V_{tot} to a reference, amplify the error and return the result to control I_{SS}
- Error amplifier is designed as a five-transistor OTA with $W/L = 5 \mu\text{m}/80 \text{ nm}$ for all transistors, tail current of $20 \mu\text{A}$ and voltage gain of 10
- Amplifier output controls a fraction of main tail current I_1



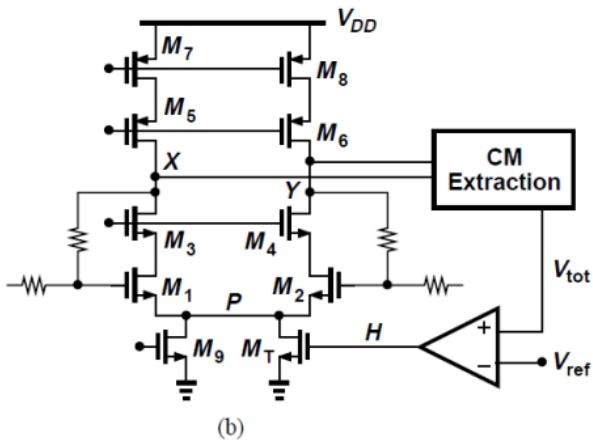
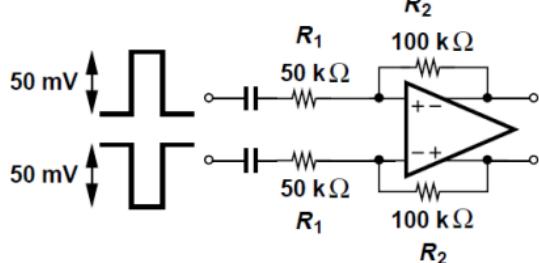
Design 1: Telescopic Op-Amp- Common-Mode Feedback

- Figure below shows closed-loop dc sweep results with $V_{ref} = 0.5 V$
 - By virtue of feedback, the CM variation is greatly reduced as V_X and V_Y reach high or low values
 - With a 10% mismatch between PMOS current sources, dc sweep is repeated [Fig. (b)]
- CMFB suppresses mismatch by adjusting I_1



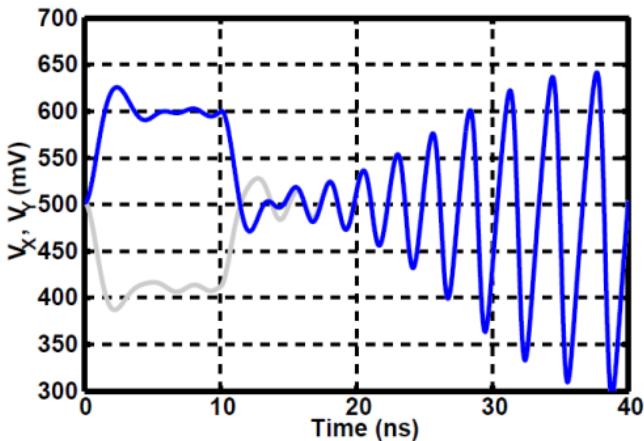
Design 1: Telescopic Op-Amp- CMFB Stability

- To investigate stability of CM loop, the overall op amp is placed its intended feedback system, applying differential pulses at input, and examining differential and common-mode behavior at output
- Fig. (a) shows a feedback topology for a nominal closed-loop gain of 2
- Fig. (b) shows a more detailed diagram highlighting the CM feedback loop



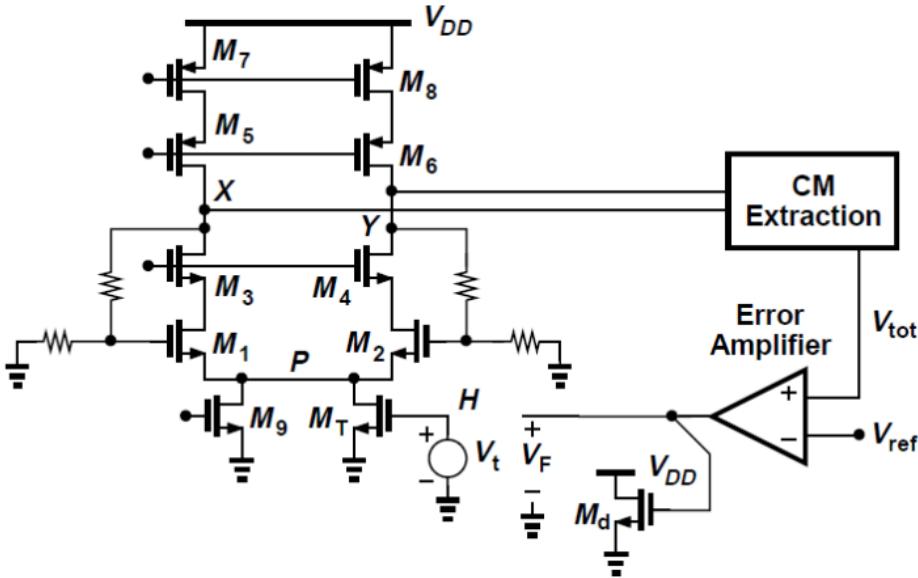
Design 1: Telescopic Op-Amp- CMFB Stability

- Output waveforms in response to an input step are plotted below, revealing common-mode instability
- CM loop contains a pole at input of error amplifier, one at node H, one at node P, one at the sources of the NMOS cascode devices and one at the main outputs [Refer Fig. (b) on previous slide]
- Loop demands compensation



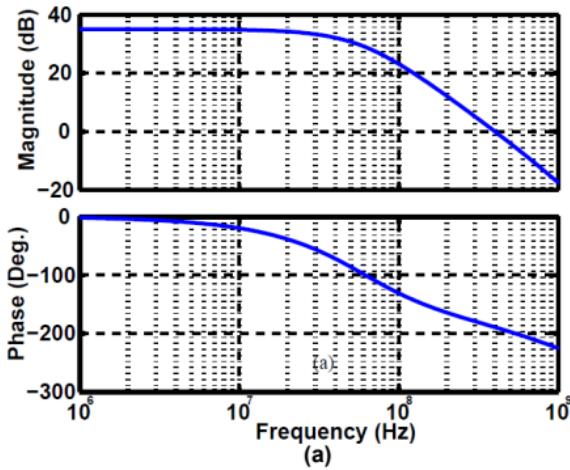
Design 1: Telescopic Op-Amp- CMFB Stability

- Break the CM loop at node H as shown below
- Error amplifier drives a dummy device M_d identical to M_T to see loading effect of the latter



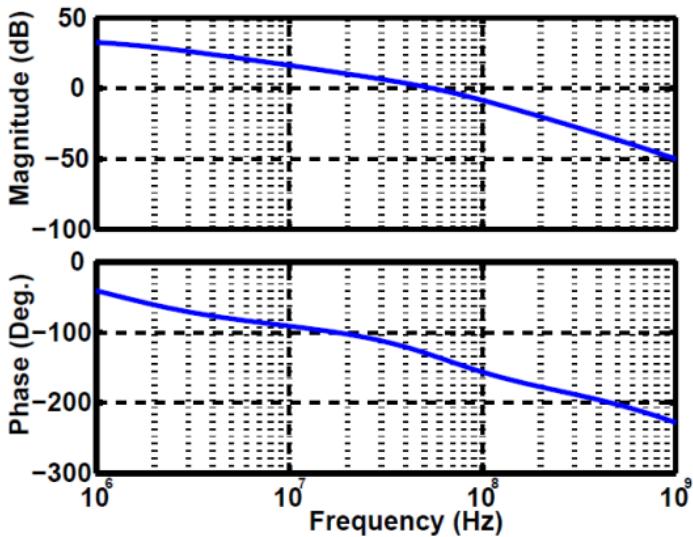
Design 1: Telescopic Op-Amp- CMFB Stability

- Fig. (a) plots the magnitude and phase of $-VF/Vt$ as a function of frequency, revealing a phase of -190° at the unity-gain frequency
- We seek a convenient node for compensation
- Error amplifier does not provide signal inversion from V_{tot} to H and hence cannot employ Miller compensation



Design 1: Telescopic Op-Amp- CMFB Stability

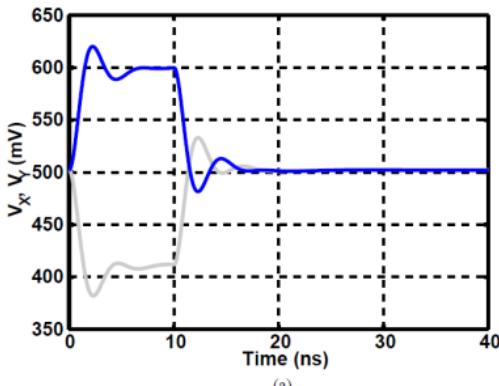
- Adding capacitance from high-impedance nodes X and Y to ground affects differential response
- A 3-pF capacitor is tied from error amplifier output to ground, obtaining the response shown in Fig. (b), with a phase margin of 50°



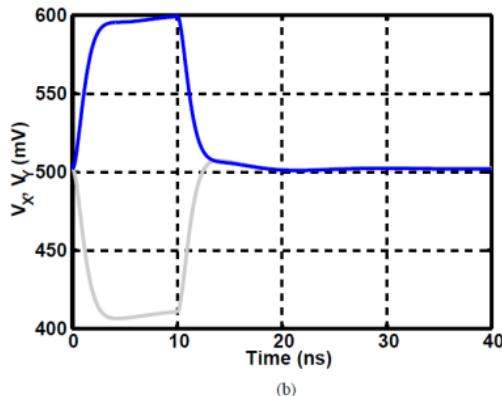
(b)

Design 1: Telescopic Op-Amp- CMFB Stability

- Closed-loop response of Fig. (a) implies that CMFB loop is properly compensated and CM level incurs differential ringing
- Pole formed by large feedback resistors and input capacitance of op-amp is located at a low frequency, degrading phase margin of differential feedback
- To compensate differential signal path, two 7-pF capacitors are connected from outputs of op-amp to its inputs to create Miller multiplication [Fig. (b)]



(a)



(b)

Design 1: Telescopic Op-Amp-Design Summary

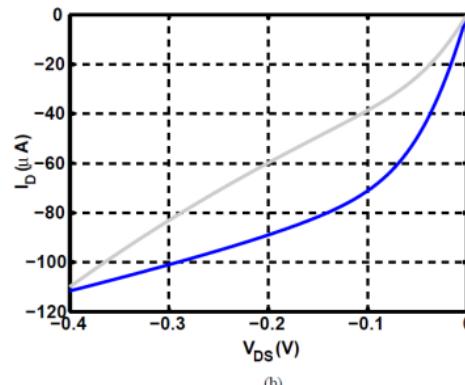
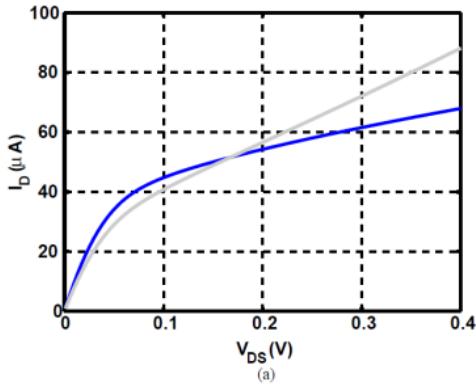
- Telescopic cascode op-amp for a voltage gain of 500 and a differential output swing of 1 V_{pp} is not achievable with a 1-V supply
- General steps followed:
 - Allocation of V_{DS} and ID to transistors according to required swings and power dissipation
 - Characterization and scaling of MOSFETs for allowable V_{DS} and desired currents
 - Quick estimate of achievable voltage gain
 - Dc sweep to study bias conditions and nonlinearity
 - Design of bias circuitry using current mirrors and low-voltage cascodes
 - CMFB design and compensation
 - Closed-loop transient analysis to study CM and differential stability

Design 2: Two-Stage Op-Amp

- Relatively high voltage gain and $1-V_{pp}$ swing point to a two-stage configuration as a feasible candidate
- Gain of 500 dictates use of cascading in the first stage
- With a gain of about 50 in the first stage, gain of second stage can be around 10
- Single-ended peak-to-peak swing at output of first stage can be as small as 50 mV, allowing redesign of cascode for greater V_{DS} 's and more robust operation
- Partitioning of power budget between two stages requires speed and/or noise specifications
- We split the power equally here, with further optimization after one round of complete design
- With about $100 \mu A$ reserved for bias network, we allocate $1.9 \text{ mA}/4 = 475 \mu A$ to each transistor branch in the first and second stages

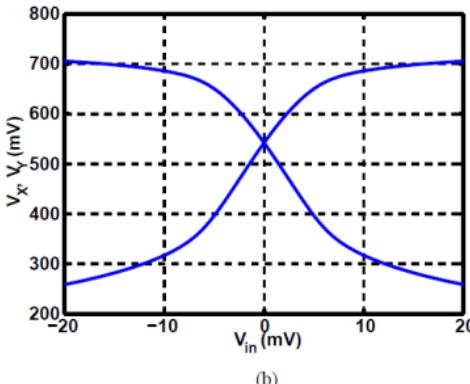
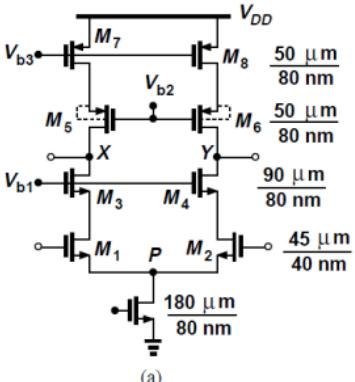
Design 2: Two-Stage Op-Amp- First Stage Design

- Telescopic-cascode configuration must accommodate a single-ended swing of 50 mVpp, allowing 0.95 V for the sum of five V_{DS} 's
- We choose $V_{DS,N} = 150 \text{ mV}$ and $V_{DS,P} = 200 \text{ mV}$ and simulate reference transistors [$W/L = 5 \mu\text{m}/40 \text{ nm}$ (gray) and $10 \mu\text{m}/80 \text{ nm}$ (blue)], seeking acceptable knee voltages
- Fig. (a) shows curves for $V_{GS,N} = 350 \text{ mV}$ while Fig. (b) shows curves for $V_{GS,P} = -450 \text{ mV}$



Design 2: Two-Stage Op-Amp- First Stage Design

- Width of NMOS transistors in the signal path must be scaled by a factor of $450 \mu\text{A}/50 \mu\text{A}$ for either $L = 40 \text{ nm}$ or $L = 80 \text{ nm}$
- Width of PMOS device must be scaled by a factor of $450 \mu\text{A}/50 \mu\text{A}$ for either $L = 40 \text{ nm}$ or $L = 80 \text{ nm}$
- For tail current device, we choose $W = (900 \mu\text{A}/50 \mu\text{A}) \times 10 \mu\text{m}$ and $L = 80 \text{ nm}$
- First-stage design is shown in Fig. (a) and simulated behavior in Fig. (b), revealing a gain of about 50

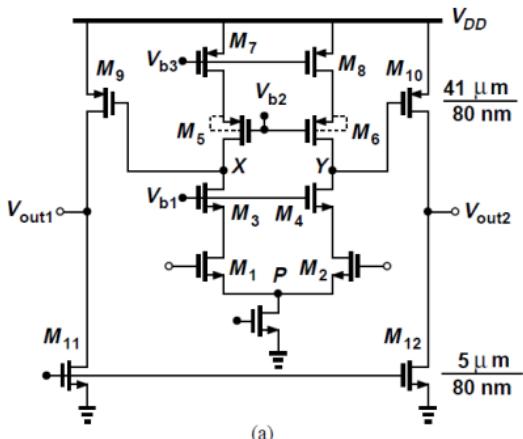


Design 2: Two-Stage Op-Amp- Second Stage Design

- Second stage must provide a voltage gain of about 10, dictating channel lengths greater than 40 nm for both NMOS and PMOS devices
- Output CM level of first stage is around 0.55 V
- Consider a transistor having $W/L = 10 \mu\text{m} / 80 \text{ nm}$ and determine its $gmrO$ if it is an NFET with $VGS \approx 0.55 \text{ V}$ or a PFET with $|VGS| \approx 0.45 \text{ V}$
- Using simulations, we get $(gmrO)N = 12.8$ and $rON = 1.86 \text{ k}\Omega$ at $VDS = 0.5 \text{ V}$ and $ID = 900 \mu\text{A}$, and $(gmrO)P = 17.5$ and $rOP = 9.75 \text{ k}\Omega$ at $|VDS| = 0.5 \text{ V}$ and $|ID| = 110 \mu\text{A}$
- We therefore select the PFET as input of second stage and scale its width to $(450 \mu\text{A}/110 \mu\text{A}) \times 10 \mu\text{m} \approx 41 \mu\text{m}$, exhibiting an output resistance of $2.38 \text{ k}\Omega$
- Drain of the PFET is tied to an NMOS current source

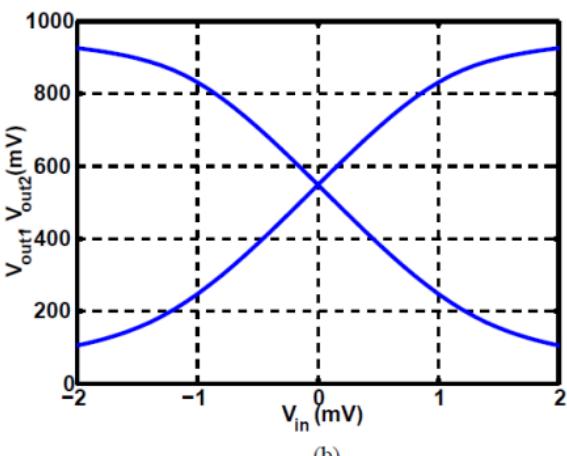
Design 2: Two-Stage Op-Amp- Second Stage Design

- The NMOS current source must not lower the gain of the second stage, $|Av_2|$ below 10
- Since $|Av_2| = g_{mP}(r_{OP})|r_{ON}| \geq 10$, we have $r_{ON} \geq 1.33r_{OP} = 3.0\text{ k}\Omega$ at $ID = 475\text{ }\mu\text{A}$
- If the 10 $\mu\text{m}/80\text{ nm}$ NFET considered before with $r_O = 1.86\text{ k}\Omega$ and $ID = 900\text{ }\mu\text{A}$ is scaled down by a factor of 2, it yields $r_{ON} = 3.72\text{ k}\Omega$, close to the desired value
- Fig. (a) shows op-amp developed so far

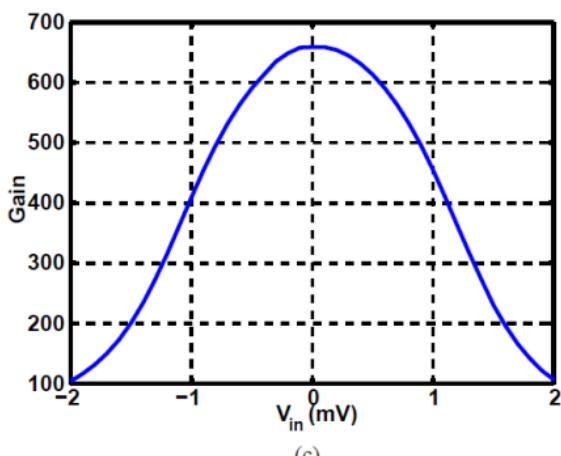


Design 2: Two-Stage Op-Amp- Second Stage Design

- Fig. (b) plots input-output characteristics
- To determine the maximum output swing that the op-amp can handle, we plot the slope of the differential characteristic in Fig. (c), noting that the differential output cannot exceed 450 mV if the gain must not drop below 500



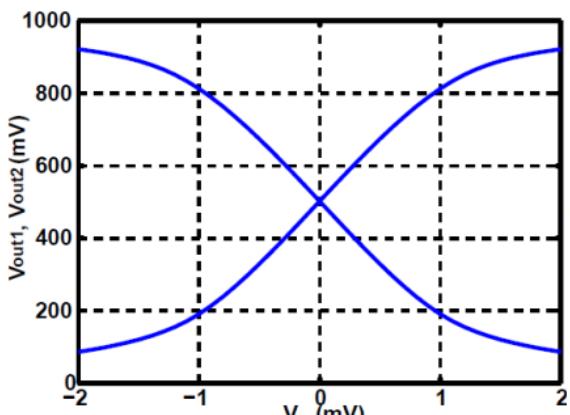
(b)



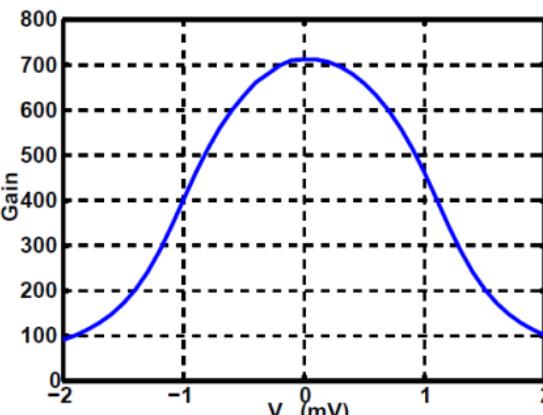
(c)

Design 2: Two-Stage Op-Amp- Second Stage Design

- To resolve this issue, we double the width and length of the output NFETs, raising the gain and arriving at the results shown below
- Now, single-ended swing reaches 530 mV for a minimum gain of 500



(a)



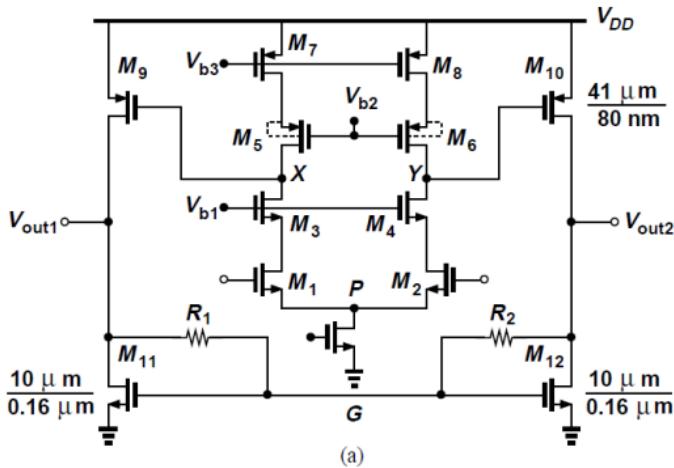
(b)

Design 2: Two-Stage Op-Amp- Common-Mode Feedback

- Two-stage op-amps generally require CMFB for both stages
- For first stage, the CMFB scheme described previously can be used
- For second stage, the lower output impedance allows the use of resistors for direct sensing of the CM level

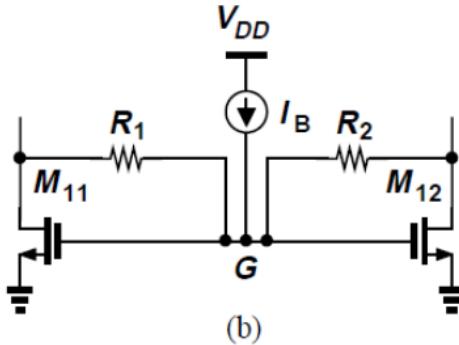
Design 2: Two-Stage Op-Amp- Common-Mode Feedback

- In topology of Fig. (a), R_1 and R_2 ($\approx 30 \text{ k}\Omega$) reconstruct the CM level at node G , applying the result to the gates of M_{11} and M_{12}
- Under equilibrium, resistors draw no current, establishing an output CM level equal to $V_{GS11,12}$
- This voltage varies by about 50 mV with PVT
- This CMFB loop is stable



Design 2: Two-Stage Op-Amp- Common-Mode Feedback

- If $V_{GS11,12}$ is not close to the desired output CM level, we inject a current I_B into node G, and the output CM level is shifted by $I_B R_1/2$ ($=I_B R_2/2$), as shown in Fig. (b)
- For example, a shift of 100 mV requires a current of $(100 \text{ mV}/30 \text{ k}\Omega) \times 2 = 6.7 \mu\text{A}$
- A positive I_B shifts the CM level down and vice versa

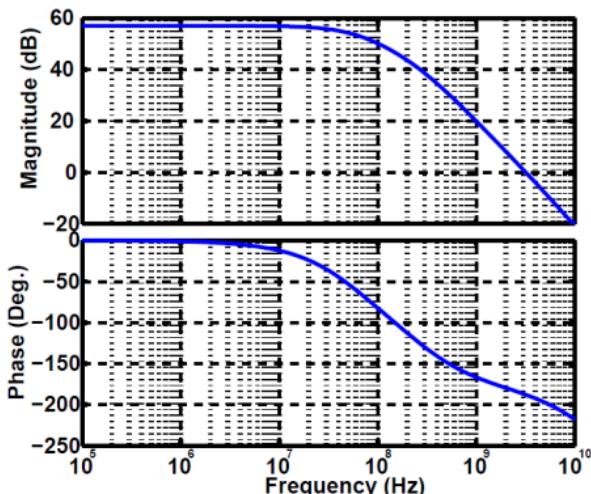


Design 2: Two-Stage Op-Amp- Frequency Compensation

- Two-stage op-amp contains several poles and demands compensation
- First non-dominant pole of two-stage op-amps typically arises from the output node and depends on the load capacitance, CL
- Choose a single-ended load capacitance of 1 pF here, obtaining an output pole frequency of around 90 MHz
- We begin with the open-loop op-amp, realizing that the feedback network may add its own effects and require changes

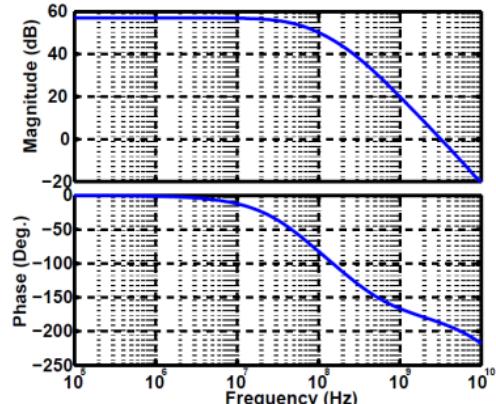
Design 2: Two-Stage Op-Amp- Frequency Compensation

- Open-loop (differential) magnitude and phase response of the circuit is plotted below
- Low-frequency gain is 57 dB (≈ 700) and unity-gain frequency is 3.2 GHz, with a phase margin of 8°
- Phase reaches -120° at 240 MHz, therefore after compensation for 60° phase margin, the unity-gain bandwidth can drop by a factor of 13



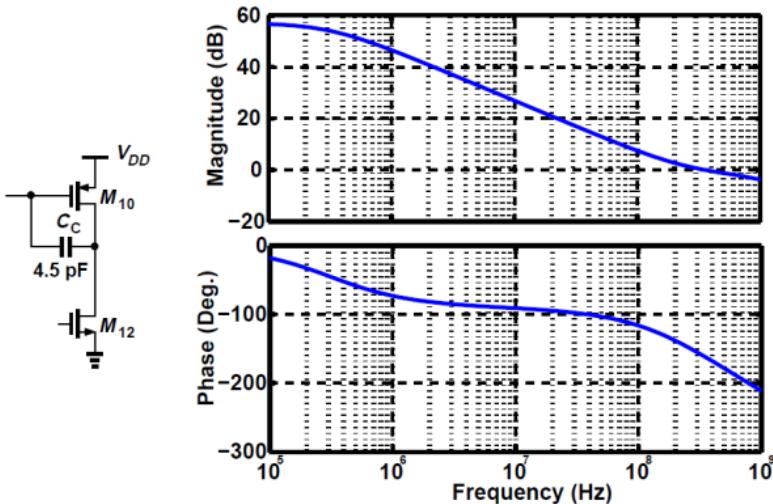
Design 2: Two-Stage Op-Amp- Frequency Compensation

- To compensate the op-amp, we begin at 240 MHz and 0 dB on the magnitude plot and draw a straight line toward the y-axis with a slope of -20 dB/dec
- The frequency at which this line intersects the magnitude plot is roughly $240 \text{ MHz}/700 = 344 \text{ kHz}$, yielding the desired value for the dominant pole
- Node X is preferred over the output node to produce the dominant pole due to Miller multiplication and pole splitting



Design 2: Two-Stage Op-Amp- Frequency Compensation

- With the output resistance of $8\text{ k}\Omega$ seen at node X and a voltage gain of about 12 provided by the output stage, we choose a Miller compensation capacitor, C_C equal to 4.5 pF so as to create a 344-kHz pole at this node
- Resulting open-loop frequency response is shown below

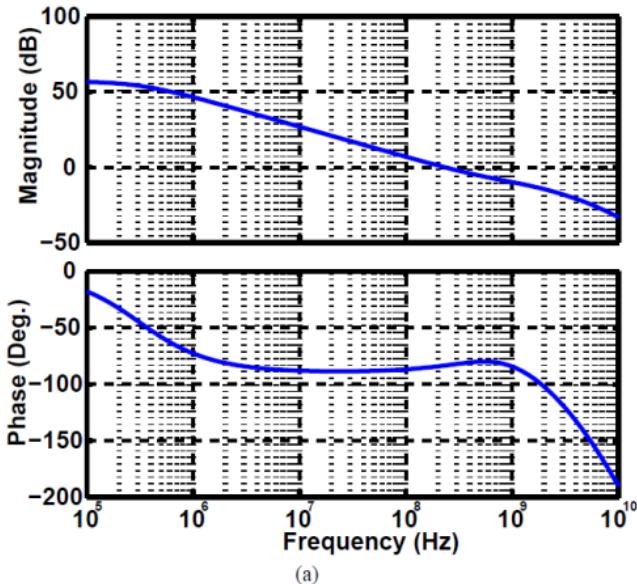
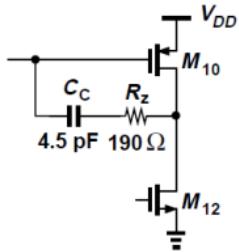


Design 2: Two-Stage Op-Amp- Frequency Compensation

- Dominant pole is now located around 340 kHz
- Gain crossover occurs at 350 MHz and the phase margin is only 18° because the zero introduced by CC, $\omega_z = gm10/CC$, is as low as 250 MHz
- We can insert a resistor R_z , in series with CC, so as to move the zero to the second pole, ω_{p2} ,
- Second pole can be roughly estimated as the frequency at which $\angle H$ reaches -135° and is equal to 185 MHz

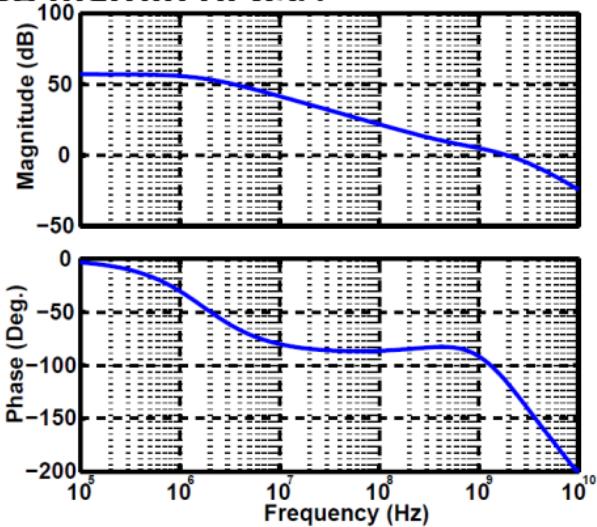
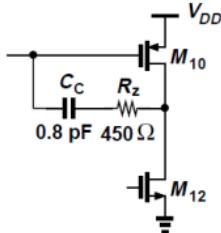
Design 2: Two-Stage Op-Amp- Frequency Compensation

- Selecting R_z according to $(\omega_p 2CC) - 1 = 190 \Omega$, the response observed is shown in Fig. (a)
- Phase margin rises to 96° because of pole-zero cancellation



Design 2: Two-Stage Op-Amp- Frequency Compensation

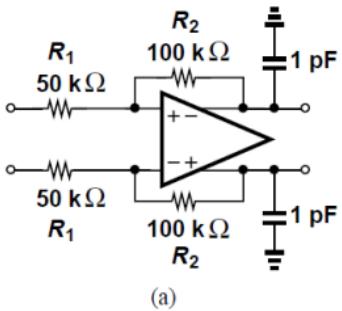
- CC can be smaller and the unity-gain bandwidth larger
- By some iteration, we choose $CC = 0.8 \text{ pF}$ and $R_z = 450 \Omega$, arriving at the response of Fig. (b)
- The op-amp now exhibits a unity-gain bandwidth of 1.9 GHz with a phase margin of 65°



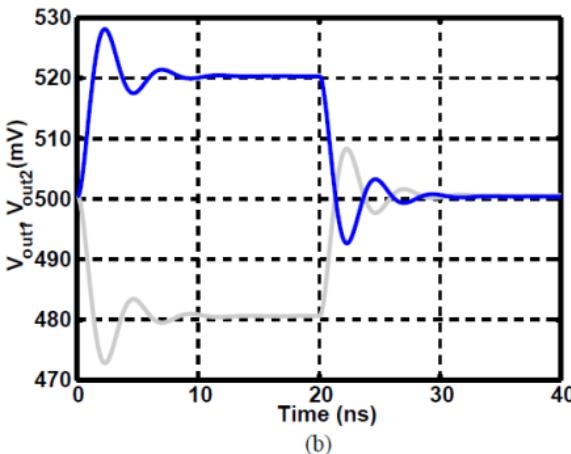
(b)

Design 2: Two-Stage Op-Amp- Closed-Loop Behavior

- Configure the op-amp as a closed-loop amplifier with a nominal gain of 2 and a load capacitance of 1 pF [Fig. (a)]
- Small-signal transient response shows significant ringing due to large resistance values used in feedback network [Fig. (b)]



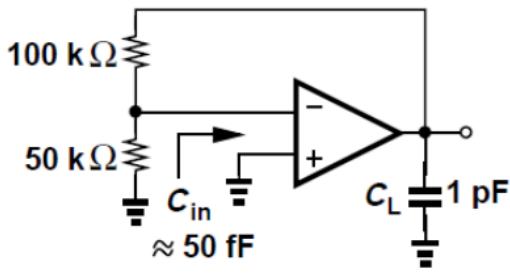
(a)



(b)

Design 2: Two-Stage Op-Amp- Closed-Loop Behavior

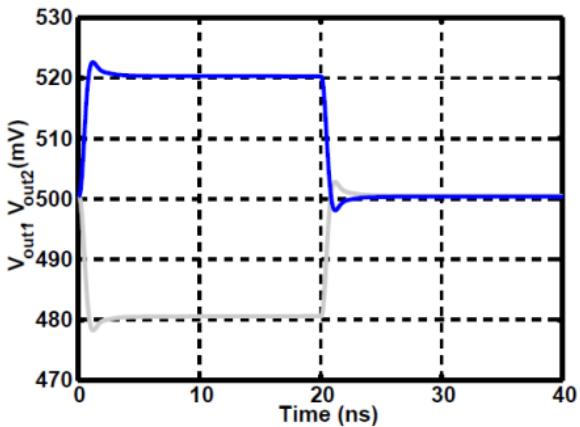
- Draw the single-ended equivalent circuit as in Fig. (a) below for loop transmission calculation
- We observe that an open-loop pole around $[2/(100 k\Omega|50 k\Omega)C_{in}]^{-1} \approx 95 \text{ MHz}$ is formed at the input of the op-amp



(a)

Design 2: Two-Stage Op-Amp- Closed-Loop Behavior

- To improve closed-loop stability, we can reduce $R1$ and $R2$ to $25\text{ k}\Omega$ and $50\text{ k}\Omega$, respectively before the open-loop gain falls appreciably at the cost of double input pole frequency
- Alternatively, we can increase the resistance in series with the compensation capacitors from $450\text{ }\Omega$ to $1500\text{ }\Omega$, arriving at the response in Fig. (b) below

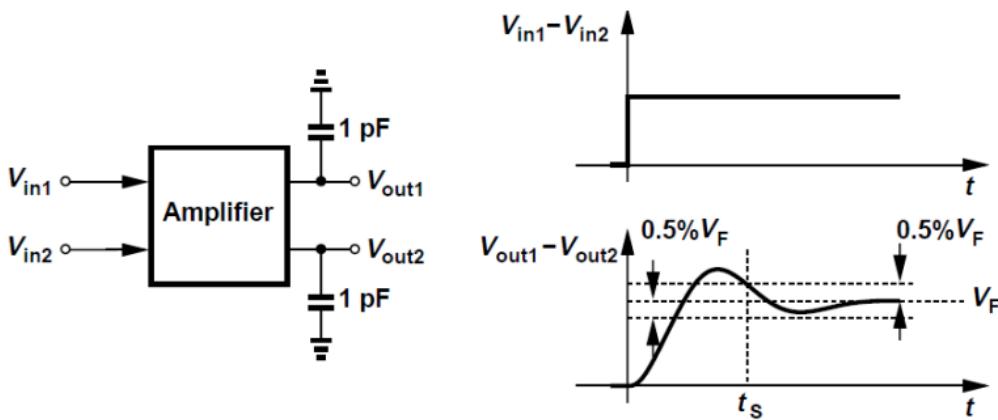


(b)

High-Speed Amplifier

- Some applications require an amplifier with fast settling and accurate gain
- We design a differential amplifier according to the following specifications:
 - **Voltage gain = 4**
 - **Gain Error $\leq 1\%$**
 - **Differential output swing = 1 V_{pp}**
 - **Load capacitance = 1 pF**
 - **Step response settling time to 0.5% accuracy = 5 ns**
 - **V_{DD} = 1 V**

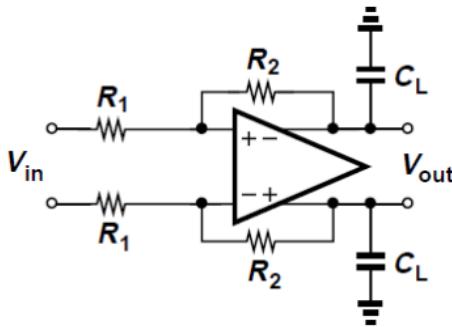
High-Speed Amplifier



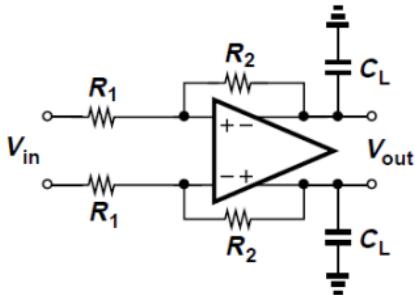
- As shown above, the settling time t_s , is defined as the time necessary for the output to reach within 0.5% of its final value
- Objective is to minimize the power consumption of the circuit

High-Speed Amplifier: Precision Issues

- Maximum tolerable gain error of 1% indicates a closed-loop configuration to make gain relatively independent of PVT
- Must design an amplifier with open-loop gain high enough to yield closed-loop gain error of less than 1%
- Also, output swing of 1 Vpp calls for a two-stage op-amp
- We arrive at the feedback arrangement shown below



High-Speed Amplifier: Precision Issues



- For the above feedback configuration, closed-loop gain is given by

$$\begin{aligned}\frac{V_{out}}{V_{in}} &= -\frac{R_2}{R_1} \frac{1}{1 + (1 + \frac{R_2}{R_1}) \frac{1}{A_0}} \\ &\approx -\frac{R_2}{R_1} \left[1 - \left(1 + \frac{R_2}{R_1} \right) \frac{1}{A_0} \right]\end{aligned}$$

- We choose $R_2/R_1 = 4$, and ensure that the gain error falls below 1%, obtaining $A_0 \geq 500$

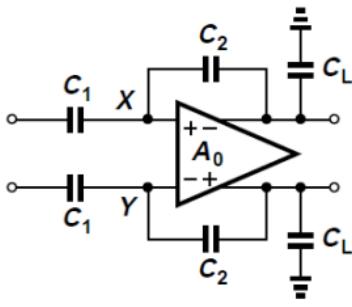
$$\left(1 + \frac{R_2}{R_1} \right) \frac{1}{A_0} \leq 0.01$$

High-Speed Amplifier: Precision Issues

- If R_1 and R_2 are large enough not to reduce the open-loop gain of the op-amp, they form a significant pole with the input capacitance, degrading the phase margin
- Consider capacitive feedback as shown in Fig. (a)
- Closed-loop gain is now C_1/C_2 , or more accurately,

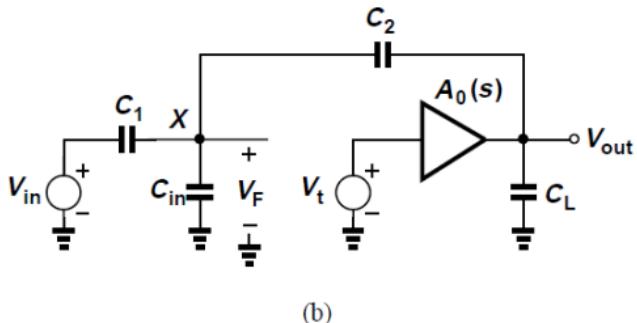
$$\frac{V_{out}}{V_{in}} \approx -\frac{C_1}{C_2} \left(1 - \frac{C_1 + C_2 + C_{in}}{C_2} \frac{1}{A_0} \right)$$

- C_{in} denotes the (single-ended) input capacitance of the op-amp



(a)

High-Speed Amplifier: Precision Issues

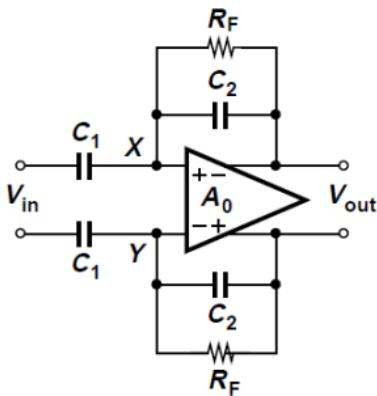


(b)

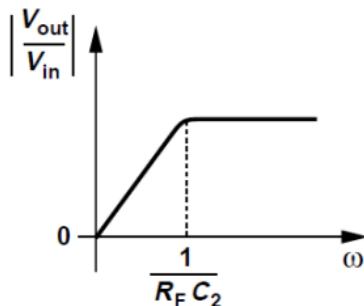
- In the single-ended counterpart of Fig. (b) for loop transmission calculation, we observe that C_1 and C_2 do not contribute additional poles because $(C_1+C_{in})C_2/(C_1+C_2+C_{in})$ simply appears in parallel with C_L

High-Speed Amplifier: Precision Issues

- To provide bias for the op-amp inputs, we add two feedback resistors so that the input and output dc levels become equal [Fig. (a)]
- Finite time constant at X and Y leads to a high-pass response



(a)



(b)

High-Speed Amplifier: Precision Issues

- If $A_0 = \infty$, then

$$\begin{aligned}\frac{V_{out}}{V_{in}}(s) &= -\frac{R_F || \frac{1}{C_2 s}}{\frac{1}{C_1 s}} \\ &= -\frac{R_F C_1 s}{R_F C_2 s + 1}\end{aligned}$$

- The corner frequency, $1/(2\sqrt{RFC2})$, must be chosen less than the minimum frequency of interest
- We proceed assuming $RFC2$ is sufficiently large
- Capacitive-feedback amplifier's gain also depends on C_{in}
- For example, if $C_{in} \approx (C_1 + C_2)/5$ then A_0 must be 20% higher than that predicted previously
- We can choose $C_1 + C_2 \gg C_{in}$ but at the cost of settling speed

High-Speed Amplifier: Speed Issues

- The amplifier must settle to 0.5% accuracy in 5 ns
- Assuming a linear, first-order circuit, we can write the step response as

$$V_{out}(t) = V_0 \left(1 - \exp \frac{-t}{\tau} \right) u(t)$$

- The time necessary for V_{out} to reach 0.995 V_0 is
 $ts = -\tau * \ln 0.005 = 5.3\tau$, i.e., τ must be no more than 0.94 ns
- Closed-loop amplifier must achieve a -3-dB bandwidth of at least $1/(2\tau) \approx 0.94$ ns ≈ 170 MHz

High-Speed Amplifier: Speed Issues

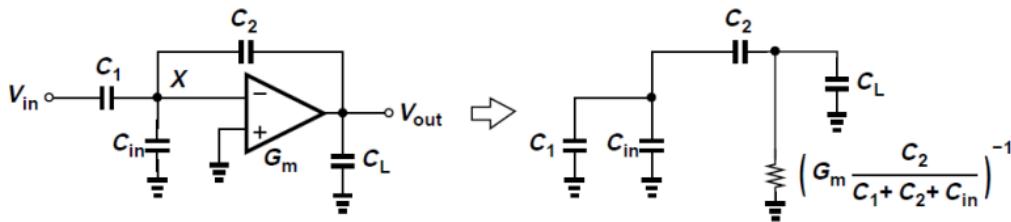
- If the op-amp is modeled by a dependent current source, $GmVin$, and an output resistance, $Rout$, then the closed-loop time constant is given by

$$\tau = \frac{C_L(C_1 + C_{in}) + C_L C_2 + C_2(C_1 + C_{in})}{G_m C_2}$$

- $GmRout$ is assumed much greater than unity; above expression can be rewritten as

$$\tau = \left(\frac{C_1 + C_2 + C_{in}}{C_2} \right) \frac{C_L + \frac{C_2(C_{in} + C_1)}{C_2 + C_{in} + C_1}}{G_m}$$

- Op-amp sets the series combination of C_2 and $C_1 + C_{in}$ in parallel with C_L



High-Speed Amplifier: Speed Issues

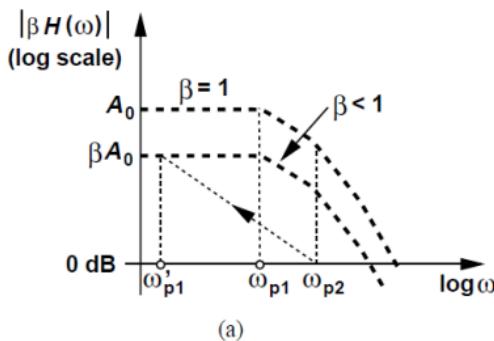
- The foregoing model is not accurate for a two-stage op-amp since the internal pole also affects the response
- Consider a frequency-compensated two-stage op-amp
- If the loop gain falls to 1 at the second pole ω_{p2} , the phase margin is about 45° for unity-gain feedback
- To compensate the op-amp for a closed-loop gain of 4, $|\beta H|$ must fall to 0 dB at ω_{p2} (i.e., circuit is not compensated for unity-gain feedback)

High-Speed Amplifier: Speed Issues

- As shown in Fig. (a), we begin at $\omega = \omega_{p2}$ and draw a line with slope -20 dB/decade toward the y-axis, seeking its intercept with the plot of $|\beta H|$ [Fig. (a)]
- Between the compensated dominant pole ω'_{p1} , and ω_{p2} , we can approximate the compensated $\beta H(s)$ as $\beta A_0 / (1 + s/\omega'_{p1})$; we set its magnitude to 1 at ω_{p2}

$$|\beta A_0 / (1 + j\omega_{p2}/\omega'_{p1})| = 1$$

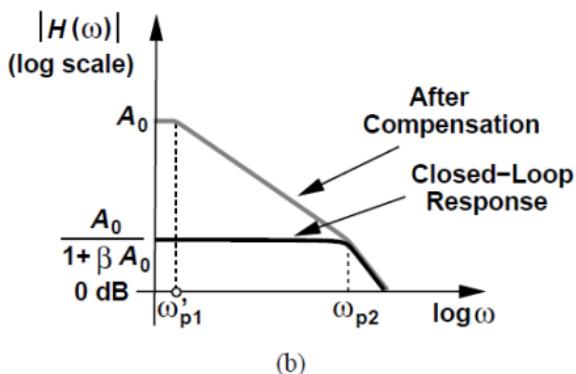
$$\square \quad \omega'_{p1} \approx \sqrt{\frac{\omega_{p2}^2}{\beta^2 A_0^2 - 1}} \quad \square \quad \omega'_{p1} \approx \frac{\omega_{p2}}{\beta A_0}$$



(a)

High-Speed Amplifier: Speed Issues

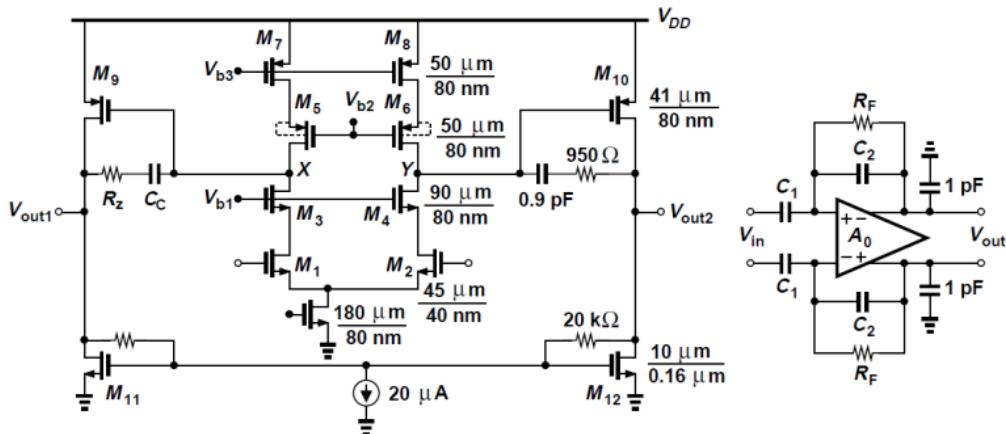
- To construct the closed-loop frequency response with the value of $\omega'p1$ calculated before, we plot the magnitude of the loop transmission $|\beta H|$, for $\beta = 1$, and after compensation [Fig. (b)]
- The closed-loop response begins at $A_0/(1 + \beta A_0)$ at low frequencies and begins to roll off at $\omega \square \omega_{p2}$
- The two responses intersect at $\omega \square \beta A_0 \omega' p_1 \square \omega_{p2}$
- We choose this bandwidth equal to $2\square(170 \text{ MHz})/125 = 2\square(1.36 \text{ MHz})$



High-Speed Amplifier: Op-Amp Design

- Based on foregoing calculations, we seek a two-stage op-amp with an open-loop gain of 500, a dominant pole at 1.36 MHz and a second pole at 170 MHz and a differential output swing of 1 V_{pp}
- Prototype designed earlier can serve our purpose
- Since the compensation can be relaxed to suit a feedback factor of 1/5, the dominant pole of the op-amp need not be as low as 344 kHz
- If the feedback factor is reduced from 1 to β , then the dominant pole can increase by roughly a factor of $1/\beta$
- Compensation capacitor can be lowered from 4.5 pF to 0.9 pF, raising the dominant pole frequency from 340 kHz to 1.7 MHz
- For the feedforward zero to cancel the second pole, R_2 must rise by the same factor, reaching 950 Ω

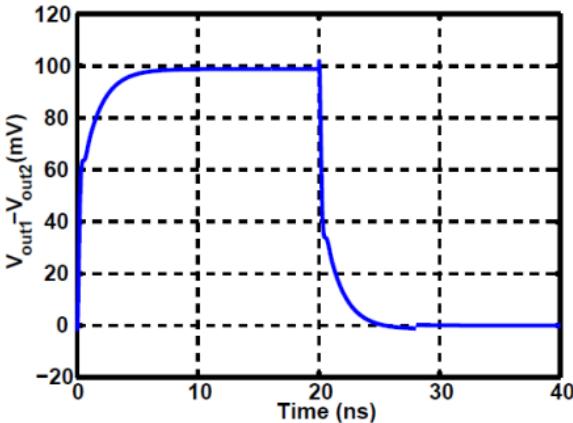
High-Speed Amplifier: Closed-Loop Small-Signal Performance



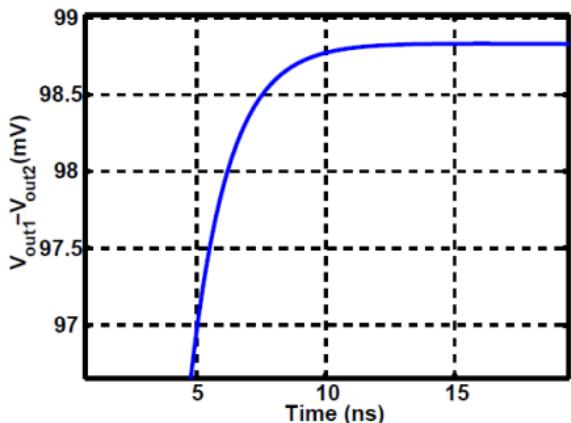
- Above figure shows overall op-amp and its closed-loop environment
- For a nominal gain of 4, we choose $C_1 = 1\text{ pF}$ and $C_2 = 0.25\text{ pF}$
- With $C_{in} \square 50\text{ fF}$, gain equation predicts a gain error less than 1% if $A_0 > 520$
- For a settling time of 5 ns, we select $R_{FC2} > 10\text{ }\mu\text{s}$, i.e., $RF = 40\text{ M}\Omega$

High-Speed Amplifier: Closed-Loop Small-Signal Performance

- Applying a differential input step of 25 mV, we expect an output around 99 mV (for 1% gain error)
- Fig. (a) is the differential output waveform and Fig. (b) is a close-up showing the fine settling
- Final value is equal to 98.80 mV, a result of insufficient open-loop gain



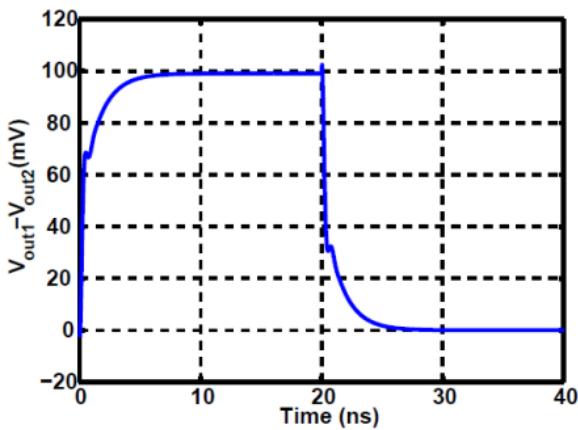
(a)



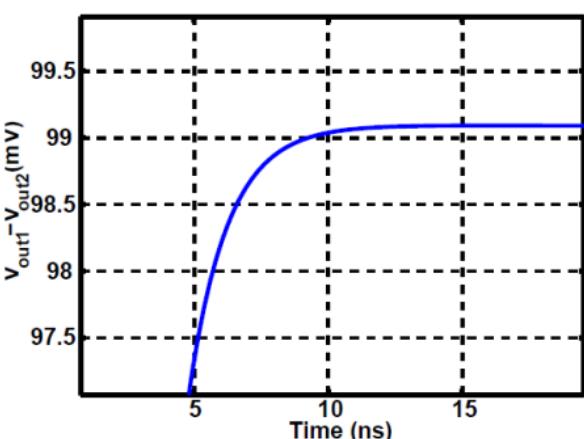
(b)

High-Speed Amplifier: Closed-Loop Small-Signal Performance

- In order to increase gain, if we raise the length (and hence width) of the first-stage input transistors, C_{in} also increases, counteracting A_0
- Instead, we double the (drawn) width and length of the NMOS cascode transistors, obtaining the output shown below
- Now the gain error is below 1%



(a)

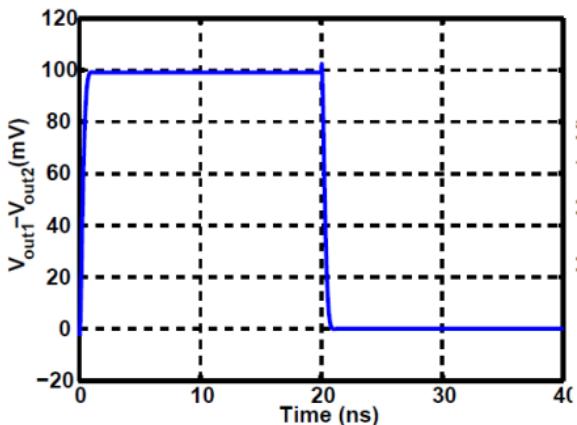


(b)

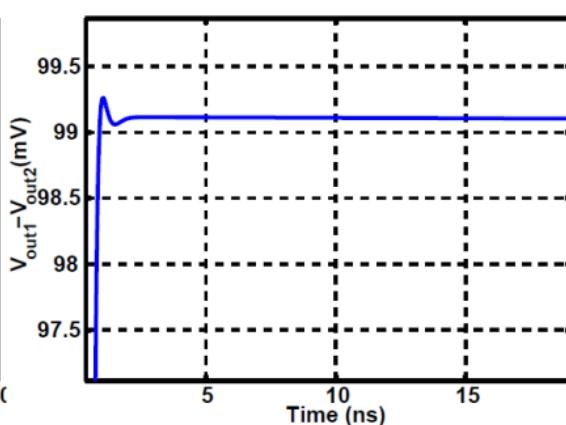
- If the output reaches 99.1 mV at $t = \square$, the settling time ts to 1% precision must be defined as follows:
 - Find the time at which $V_{out} = 99.1 \text{ mV} \square 0.01 \square 99.1 \text{ mV} \square 99.1 \text{ mV} \square 1 \text{ mV}$
 - From Fig. (b) on previous slide, we obtain $ts \square 5.8 \text{ ns}$
- From Fig. (a) on slide 104, output appears “overcompensated”, thus speed can be improved

High-Speed Amplifier: Closed-Loop Small-Signal Performance

- With $CC = 0.3 \text{ pF}$ and $Rz = 700 \Omega$, we observe the settling behavior shown below
- Settling time drops to 800 ps, a remarkable improvement
- Rz is reduced in this case, moving the zero to higher frequencies



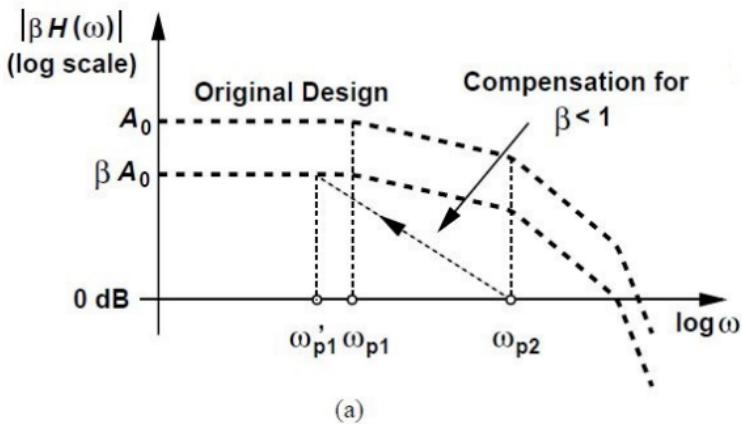
(a)



(b)

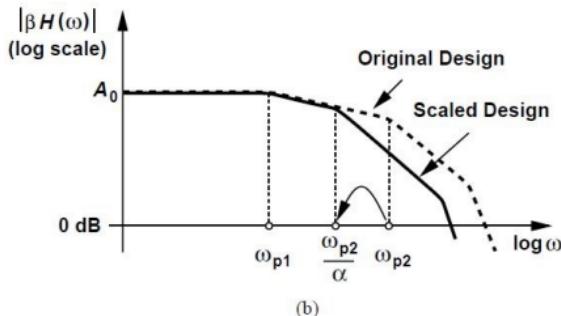
High-Speed Amplifier: Op-Amp Scaling

- If the settling time is much shorter than the required value, we can trade speed for power dissipation by “linear scaling”
- Beginning with the response in Fig. (a), we scale down all transistor widths and bias currents by a factor of β , thereby reducing the power by the same factor while retaining the voltage gain and the headroom

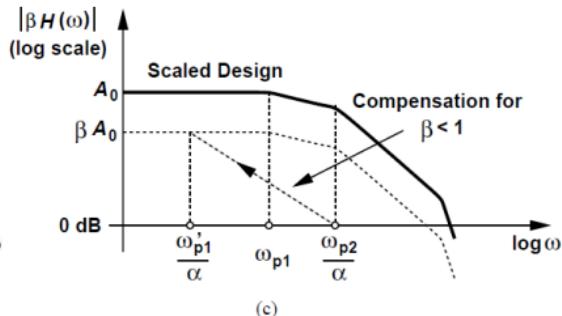


High-Speed Amplifier: Op-Amp Scaling

- With the load capacitance fixed, the output pole (before compensation) is scaled down by a factor of α [Fig. (b)]
- To maintain the same phase margin, the dominant pole after compensation must also be scaled down by this factor [Fig. (c)]
- Since the output impedance of first stage is multiplied by α , CC should remain unchanged
- To place the zero introduced by R_z atop ω_{p2}/α , we multiply R_z by α



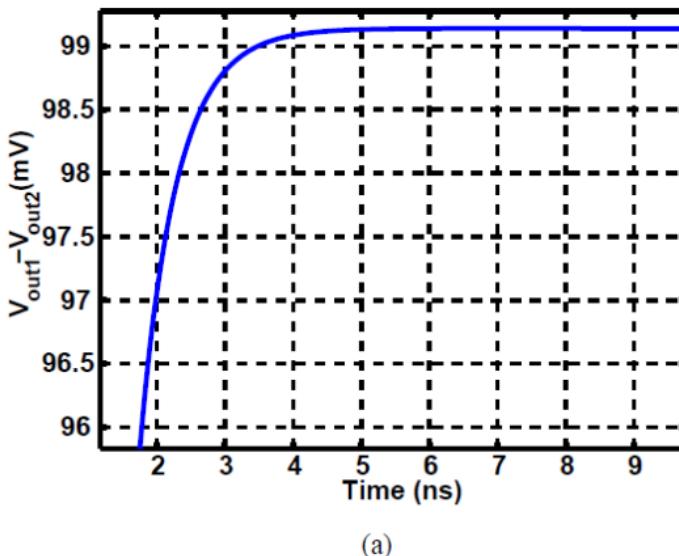
(b)



(c)

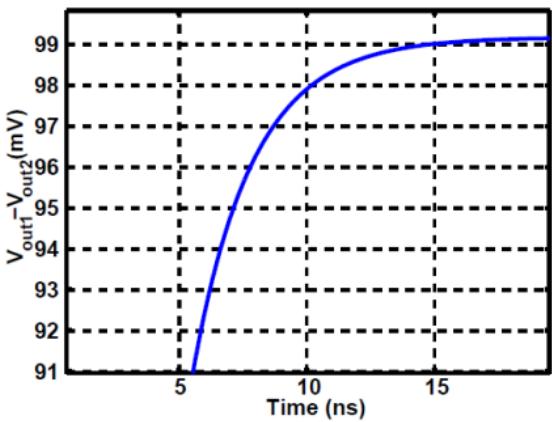
High-Speed Amplifier: Op-Amp Scaling

- Let us try $\beta = 2$
- Fig. (a) plots the output waveform, revealing the same final values as before and an overdamped response with $t_s \approx 2.5 \text{ ns}$

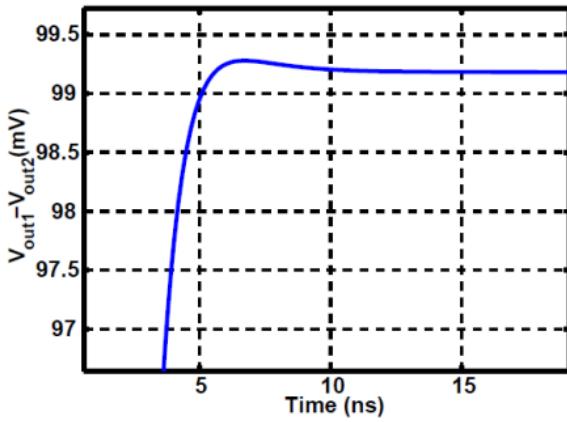


High-Speed Amplifier: Op-Amp Scaling

- We can try scaling by another factor of 4 (i.e., $\square = 8$ with respect to the original design), observing the heavily overdamped response in Fig. (b)
- Now, we adjust CC and Rz manually to optimize the speed
- With $CC = 0.15 \text{ pF}$ and $Rz = 9 \text{ k}\Omega$, the step response appears as in Fig. (c), exhibiting $ts \square 4.5 \text{ ns}$



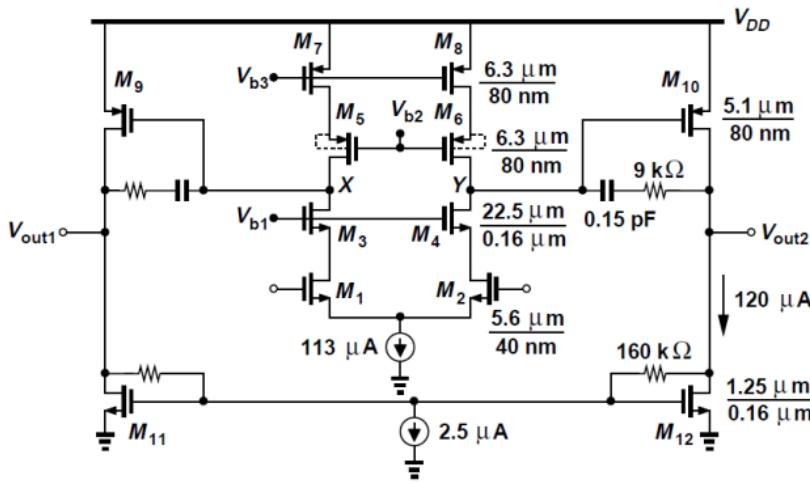
(b)



(c)

High-Speed Amplifier: Op-Amp Scaling

- Linear scaling along with some adjustment of CC and R_z affords an eightfold reduction of power and area
- Minimal design effort needed because it does not alter circuit's gain and swing values
- Scaling gives rise to longer settling and higher noise (and offset)

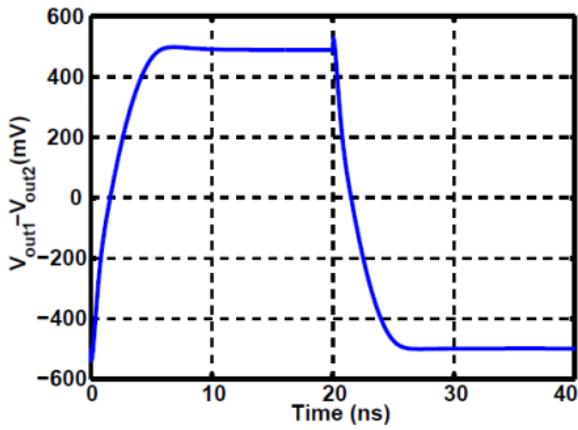


High-Speed Amplifier: Large-Signal Behavior

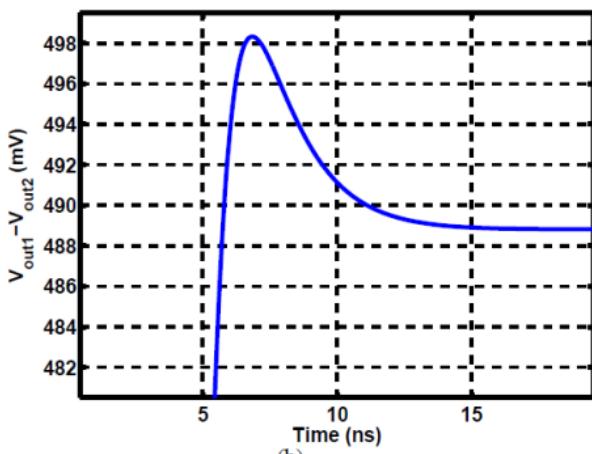
- The amplifier's ultimate test is with large signal swings (1 V_{pp,diff})
- Open-loop gain may drop as some transistors sustain less V_{DS} , and speed may suffer due to slewing
- In previous simulations, the differential output begins from zero, jumps to some value, and returns to zero
- For large-signal tests, V_{out} must swing from -0.5 V to +0.5 V, which can be accomplished by setting the initial differential conditions at the op-amp inputs such that $V_{out} = -0.5$ V at $t = 0$

High-Speed Amplifier: Large-Signal Behavior

- The large-signal response is shown in Fig. (a)
- Total change in V_{out} from $t = 20 \text{ ns}$ to $t = 40 \text{ ns}$ is equal to 987.4 mV, about 2.6 mV less than the allowed value for 1% gain error
- Settling to 1% from final value is about 6 ns



(a)



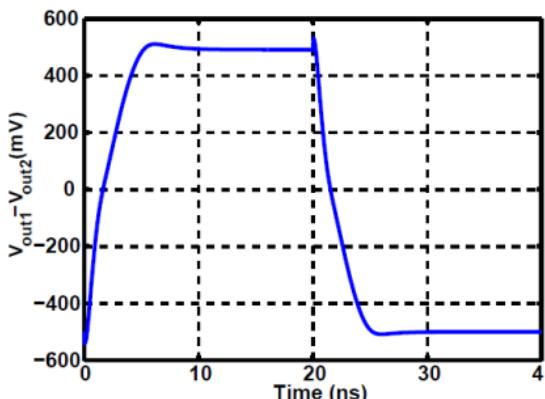
(b)

High-Speed Amplifier: Large-Signal Behavior

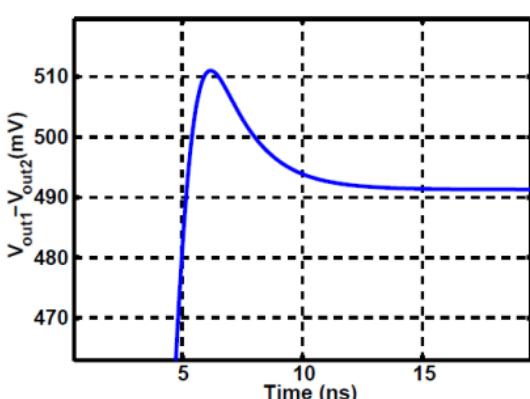
- Let us first deal with insufficient gain
- We can measure voltage gain of each stage under these conditions by dividing its differential output swing by its differential input swing (after the voltages have settled)
- We obtain $A_v = 39.5$ and 10.2 for first and second stages respectively (in small-signal operation, these values are equal to 46.3 and 11.2)
- Thus, open-loop gain has dropped from 518 to 403

High-Speed Amplifier: Large-Signal Behavior

- To raise the gain, we double W and L of the NMOS cascode transistors in the first stage and the NMOS current sources in the second, arriving at the output shown below
- Gain error is now less than 1%, but settling has become longer because the pole associated with the source of the NMOS cascode transistors degrades the phase margin



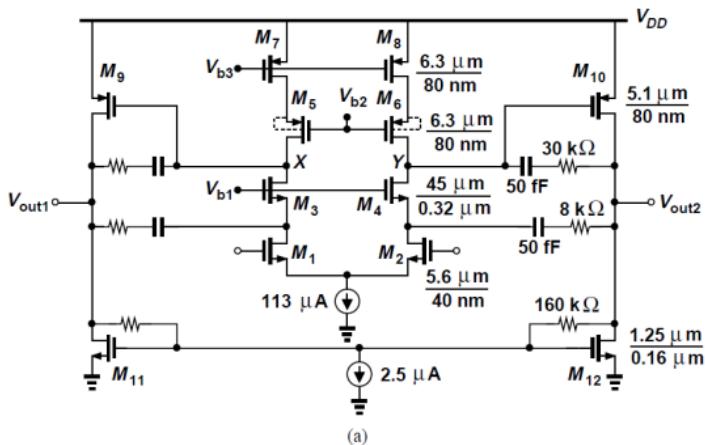
(a)



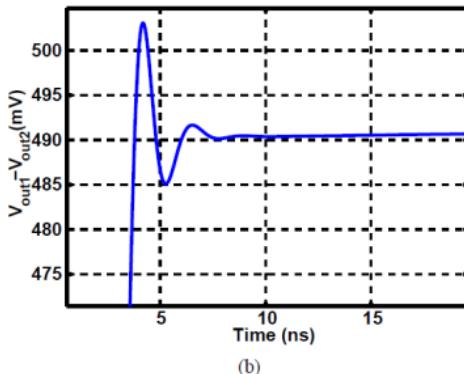
(b)

High-Speed Amplifier: Large-Signal Behavior

- To address the settling issue, we consider cascode compensation
- With some iteration, we reach the design of Fig. (a)
- As shown in Fig. (b), the settling to 1% takes less than 5 ns
- This performance is achieved with a power of $370 \mu\text{W}$



(a)



(b)

Summary

- Three steps in good analog design:
 - Closely examine the circuit's behavior and understand the root cause of undesired phenomena
 - Adjust only the circuit parameters that relate to the root cause – do not play blindly with any random device
 - Continue to explore various techniques and new ideas, many a time improving performance