MIT OPENCOURSEWARE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

6.776
High Speed Communication Circuits and Systems
Lecture 8
Broadband Amplifiers, Continued

Massachusetts Institute of Technology

March 1, 2005

Copyright © 2005 by Hae-Seung Lee and Michael H. Perrott

The Issue of Velocity Saturation

- We often assume that MOS current is a quadratic function of V_{gs}:

$$I_D = \frac{\mu_n C_{ox}}{2}\frac{W}{L}(V_{gs} - V_T)^2$$

- It can be shown, more generally

$$I_D = \frac{\mu_n C_{ox}}{2}\frac{W}{L}(V_{gs} - V_T)V_{dsat,l}$$

  - $V_{dsat,l}$ corresponds to the saturation voltage at a given length, which we often refer to as ΔV
  - In strong inversion below velocity saturation:

    $$V_{dsat,l} \approx V_{gs} - V_T$$

    which gives the quardatic equation above.

Velocity Saturation Continued

- **It can be shown that**

$$V_{dsat,l} \approx \frac{(V_{gs} - V_T)(LE_{sat})}{(V_{gs} - V_T) + (LE_{sat})} = (V_{gs} - V_T)||(LE_{sat})$$

  - **E_{sat}: electric field (lateral) at which velocity saturation occurs**
  - **If** $\frac{V_{gs} - V_T}{L} << E_{sat}$ **then** $V_{dsat,l} \approx V_{gs} - V_T$
  - **If $(V_{gs}$-$V_T)$/L approaches E_{sat} in value, then the quadratic equation is no longer valid**
  - **If** $\frac{V_{gs} - V_T}{L} >> E_{sat}$ **then** $V_{dsat,l} \approx LE_{sat} = constant$ **and the I-V characteristic becomes linear**

Analytical Device Modeling in Velocity Saturation

- **If L small (as in modern devices), than velocity saturation will impact us for even moderate values of V_{gs}-V_T**

$$I_D = \frac{\mu_n C_{ox}}{2}\frac{W}{L}(V_{gs} - V_T)[(V_{gs} - V_T)||(LE_{sat})]$$

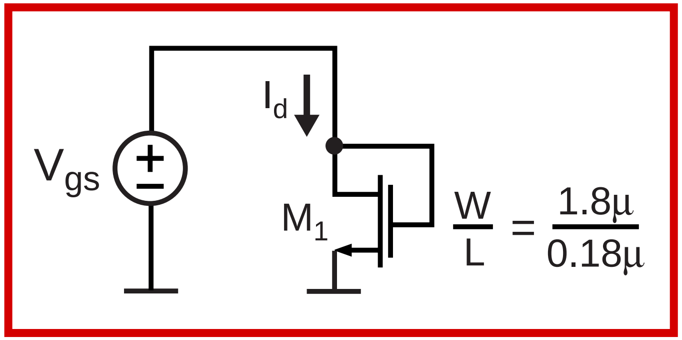
$$\Rightarrow \; I_D \approx \frac{\mu_n C_{ox}}{2} W (V_{gs} - V_T) E_{sat}$$

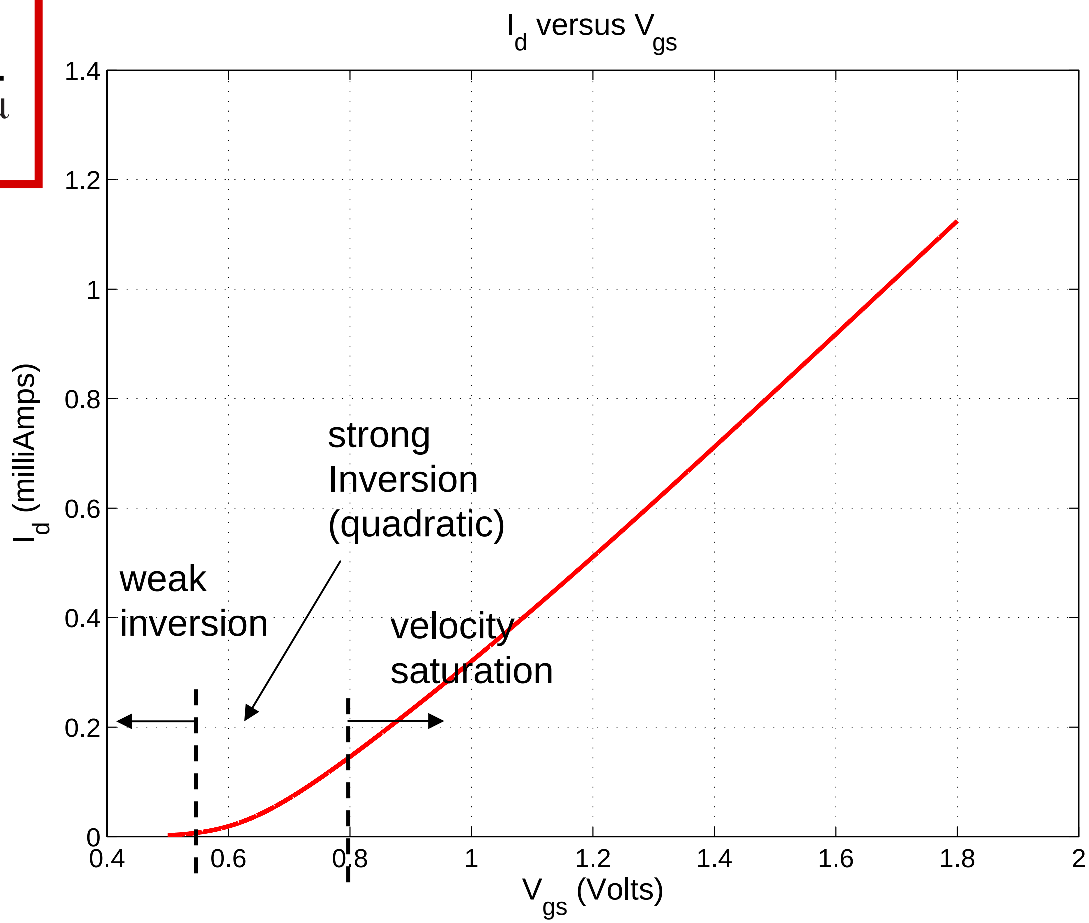
  - **Current increases linearly with V_{gs}-V_T!**
- **Transconductance in velocity saturation:**

$$g_m = \frac{dI_d}{dV_{gs}} \;\Rightarrow\; g_m = \frac{\mu_n C_{ox}}{2} W E_{sat}$$

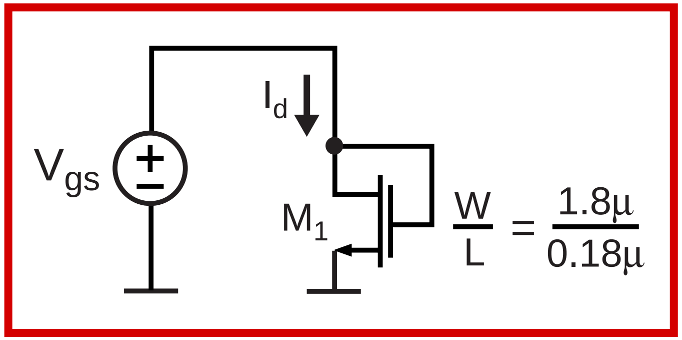
  - **No longer a function of V_{gs}- higer V_{gs} increases I_d, but little increase in g_m: wasted power**

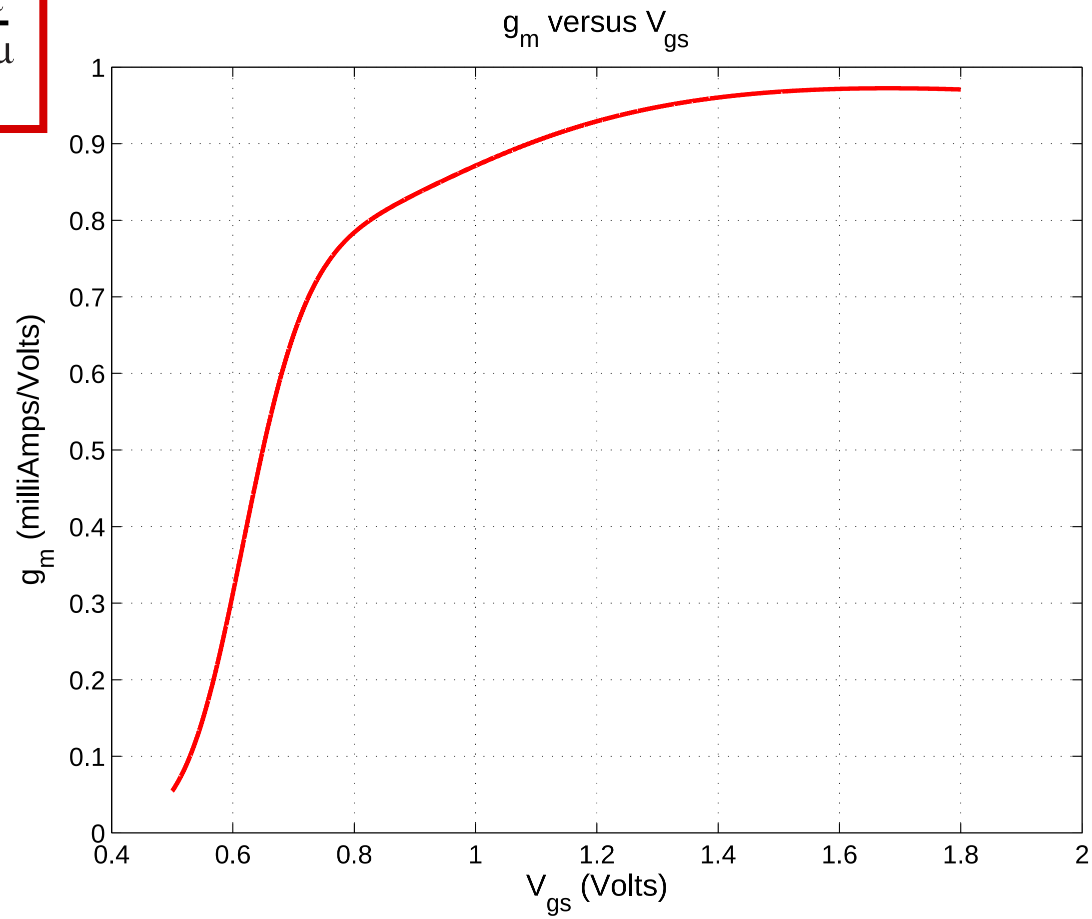
Example: Current Versus Voltage for 0.18μ Device



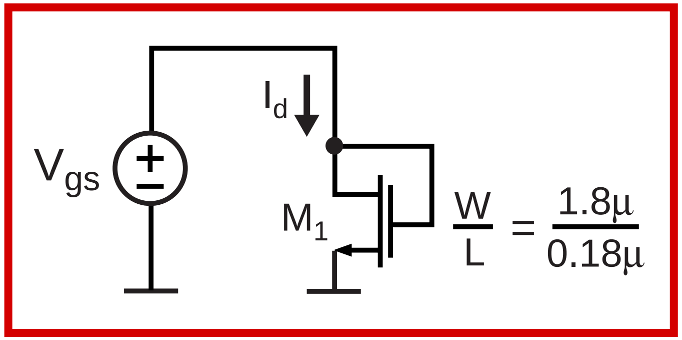


Example: G_m Versus Voltage for 0.18μ Device

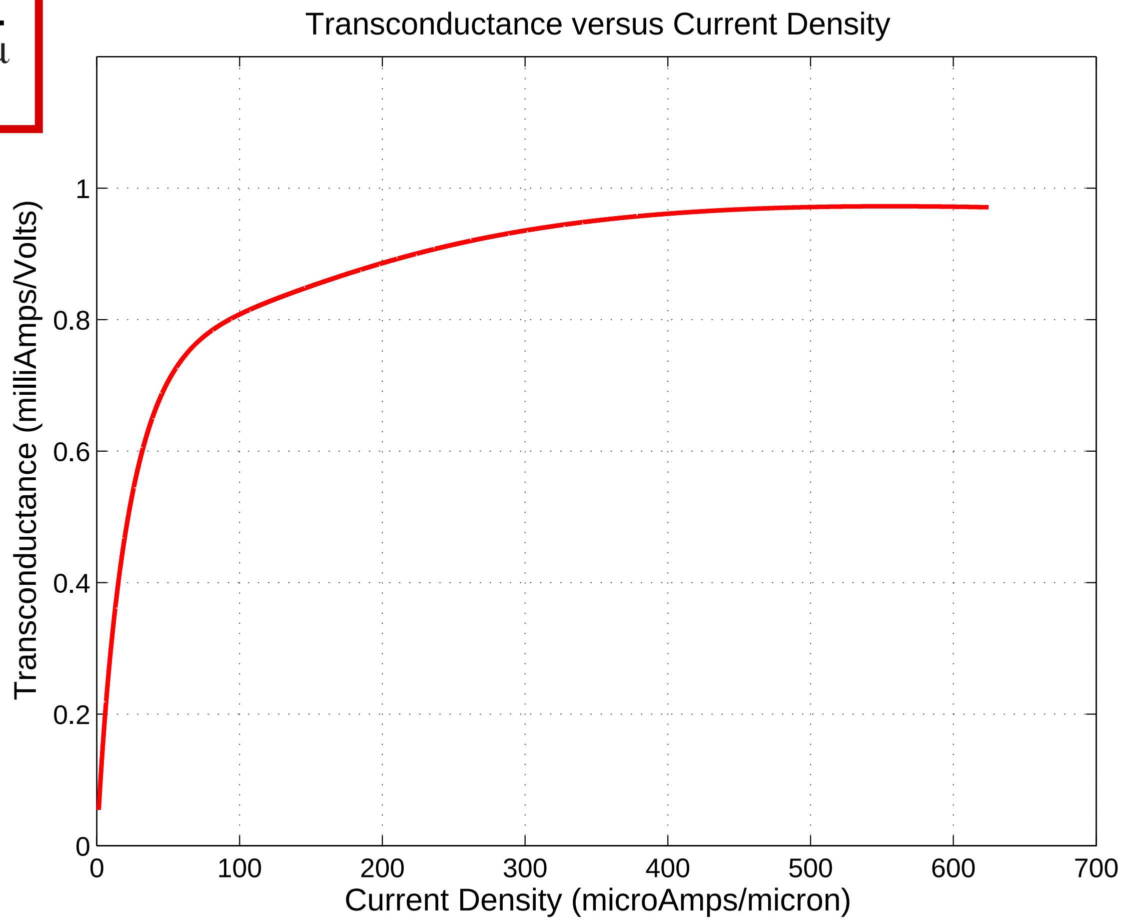




Example: G_m Versus Current Density for 0.18μ Device



Note: $I_{den} = \frac{I_d}{W} = \frac{I_d}{1.8\mu}$



How Do We Design the Amplifier?

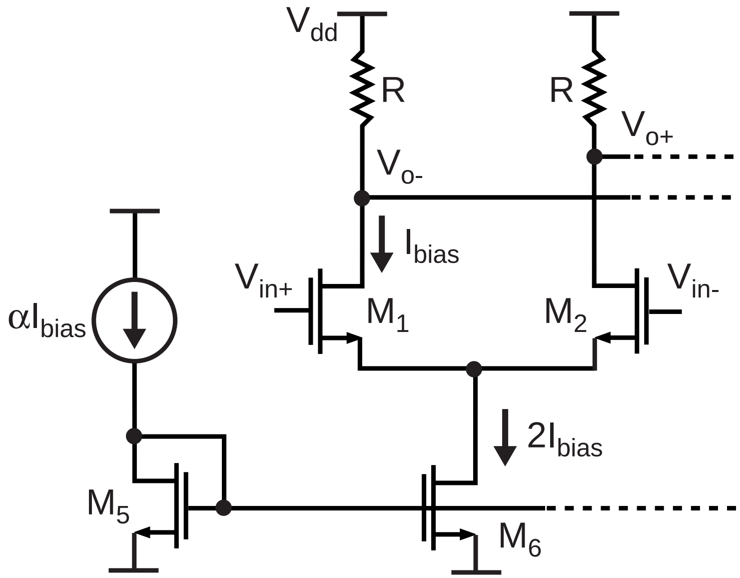
- **Highly inaccurate to assume square law behavior**
- **We will now introduce a numerical procedure based on the simulated g_m curve of a transistor**
  - **A look at transconductance:**

$$g_m = \frac{dI_d}{dV_{gs}}|_{I_d} = \frac{dWI_{den}}{dV_{gs}} = W\frac{dI_{den}}{dV_{gs}}|_{I_{den}}$$

  - **Observe that if we keep the current density ($I_{den}=I_d/W$) constant, then g_m scales directly with W**
    - This is independent of bias regime
  - **We can therefore relate g_m of devices with different widths given that they have the same current density**

$$g_m(W, I_{den}) = \frac{W}{W_o} g_m(W_o, I_{den})$$

A Numerical Design Procedure for Resistor Amp – Step 1



- **Two key equations**
  - **Set gain and swing (single-ended)**

$$(1)\quad g_m(W, I_{bias}/W)R = A$$

$$(2)\quad V_{sw} = 2I_{bias}R$$

- **Equate (1) and (2) through R**

$$\frac{A}{g_m(W, I_{bias}/W)} = \frac{V_{sw}}{2I_{bias}}$$

$$\Rightarrow\quad g_m(W, I_{bias}/W) = 2\frac{A}{V_{sw}}W\left(\frac{I_{bias}}{W}\right)$$

Can we relate this formula to a g_m curve taken from a device of width W_o?

A Numerical Design Procedure for Resistor Amp – Step 2

- **We now know:**

$$(1)\quad g_m(W, I_{bias}/W) = 2\frac{A}{V_{sw}}W\left(\frac{I_{bias}}{W}\right)$$

$$(2)\quad g_m(W, I_{den}) = \frac{W}{W_o}g_m(W_o, I_{den})$$

- **Substitute (2) into (1)**

$$\frac{W}{W_o}g_m(W_o, I_{bias}/W) = 2\frac{A}{V_{sw}}W\left(\frac{I_{bias}}{W}\right)$$

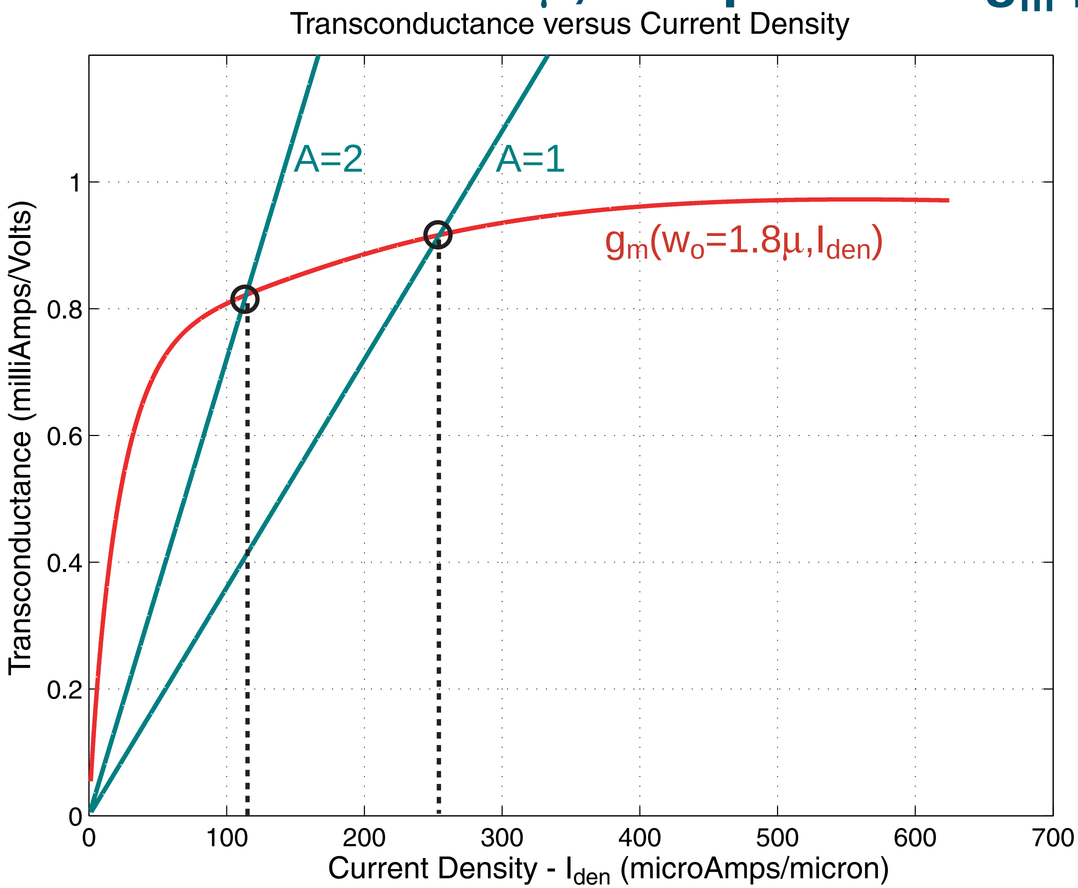
$$\Rightarrow \quad g_m(W_o, I_{den}) = 2W_o\frac{A}{V_{sw}}I_{den}$$

- **The above expression allows us to design the resistor loaded amp based on the g_m curve of a representative transistor of width W_o!**

Example: Design for Swing of 1 V, Gain of 1 and 2

$$g_m(W_o, I_{den}) = 2W_o \frac{A}{V_{sw}} I_{den}$$

- **Assume L=0.18μ, use previous g_m plot (W_o=1.8μ)**



- **For gain of 1, current density = 250 μA/μm**
- **For gain of 2, current density = 115 μA/μm**
- **Note that current density reduced as gain increases!**
  - **f_t effectively decreased**

Example (Continued)

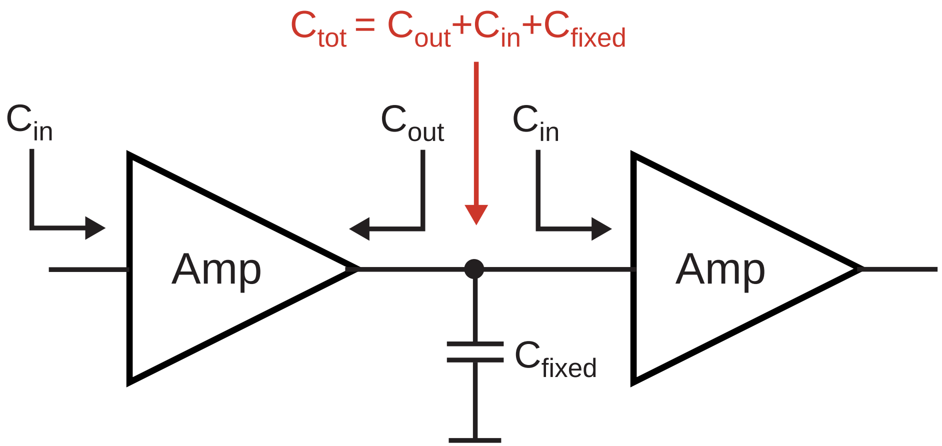
- **Knowledge of the current density allows us to design the amplifier**
  - **Recall** $V_{sw} = 2I_{bias}R$
  - **Free parameters are W, I_{bias}, and R (L assumed to be fixed)**
- **Given I_{den} = 115 μA/μm (Swing = 1V, Gain = 2)**
  - **If we choose I_{bias} = 300 μA**

$$I_{den} = \frac{I_{bias}}{W} \quad \Rightarrow \quad W = \frac{300}{115} = 2.6\mu m$$

$$V_{sw} = 2I_{bias}R \quad \Rightarrow \quad R = \frac{1}{2 \cdot 300 \times 10^{-6}} = 1.67k\Omega$$

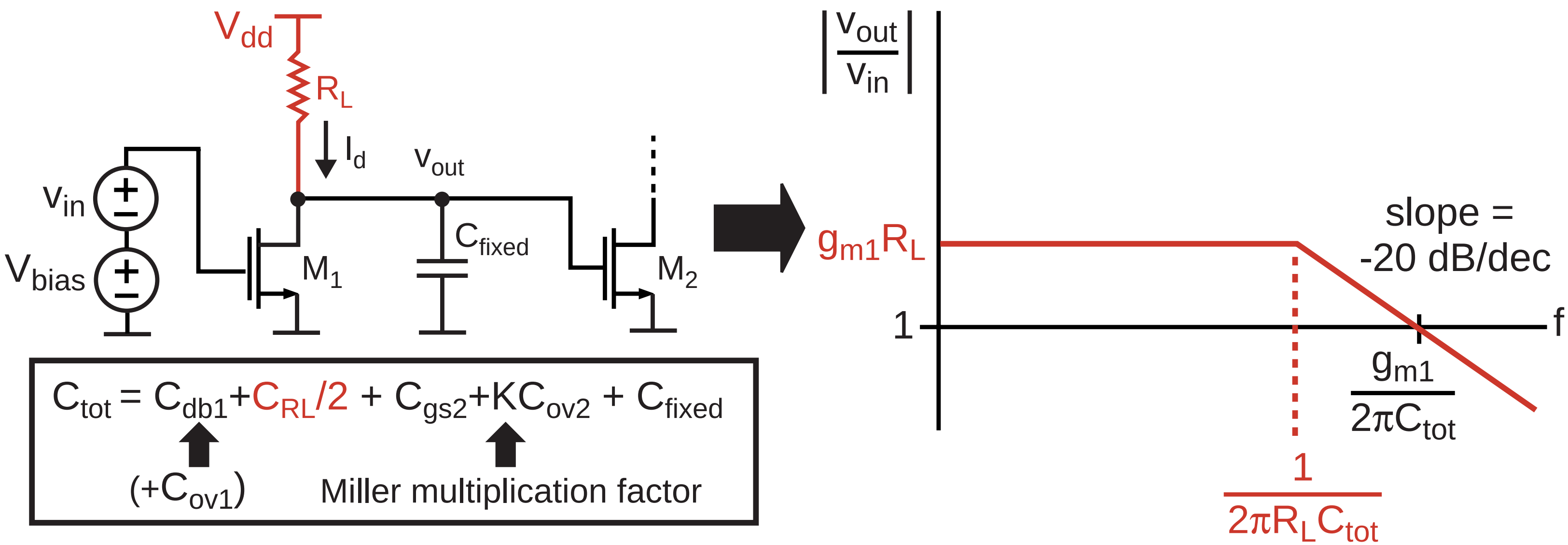
- **Note that we could instead choose W or R, and then calculate the other parameters**

How Do We Choose I_{bias} For High Bandwidth?



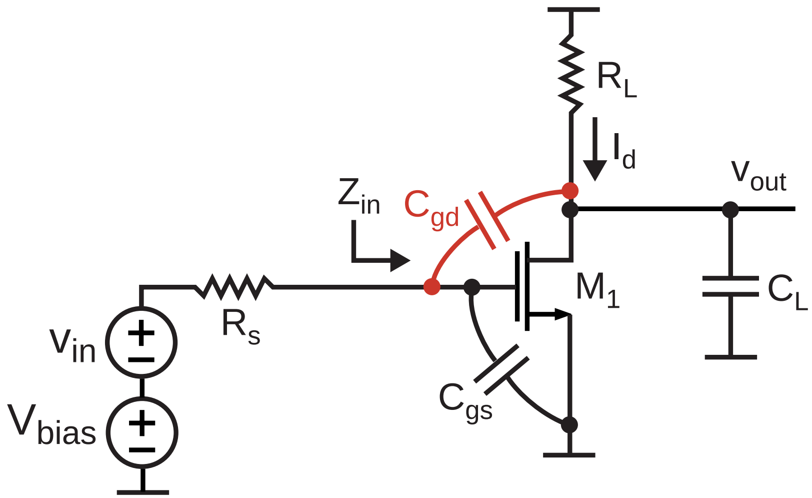
- **Pick current density just below velocity saturation**
- **As you increase I_{bias}, the size of transistors also increases to keep a constant current density**
  - **The size of C_{in} and C_{out} increases relative to C_{fixed}**
- **To achieve *the highest bandwidth*, size the devices (i.e., choose the value for I_{bias}), such that**
  - **$C_{in}+C_{out}$ dominates over C_{fixed}**
- **However, $C_{in}+C_{out}=C_{fixed}$ is roughly the point of diminishing return because the bandwidth improvement becomes marginal while power and area continue to grow proportionally**
- **Thus, $C_{in}+C_{out}=C_{fixed}$ is the most *efficient* point**

Resistor Loaded Amplifier (Unsilicided Poly)



- **We decided this was the fastest non-enhanced amplifier**
  - **Can we go faster? (i.e., can we enhance its bandwidth?)**
- **We will look at the following**
  - **Reduction of Miller effect on C_{gd}**
  - **Shunt, series, and zero peaking**
  - **Distributed amplification**

Miller Effect on C_{gd} Is Significant



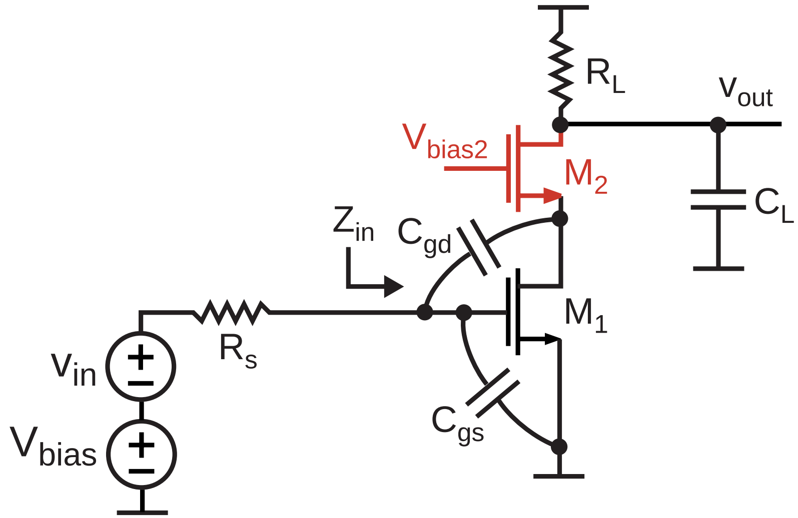
- C_{gd} is quite significant compared to C_{gs}
  - In 0.18μ CMOS, C_{gd} is about 45% the value of C_{gs}
- Input capacitance calculation

$$Z_{in} \approx \frac{1}{s(C_{gs} + C_{gd}(1 - A_v))} = \frac{1}{sC_{gs}(1 + \frac{C_{gd}}{C_{gs}}(1 + g_m R_L))}$$

  - For 0.18μ CMOS, gain of 3, input cap is almost tripled over C_{gs}!

$$Z_{in} \approx \frac{1}{sC_{gs}(1 + 0.45(4))} = \frac{1}{sC_{gs}2.8}$$

Reduction of C_{gd} Impact Using a Cascode Device



- **The cascode device lowers the gain seen by C_{gd} of M_1(the *total* gain is the same as non-cascoded amp)**

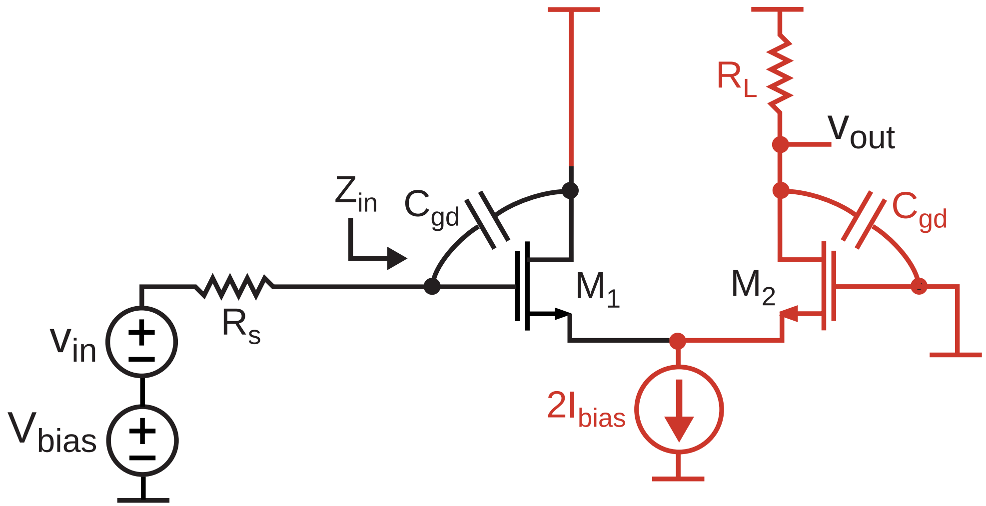
$$A_v \rightarrow g_{m1}\frac{1}{g_{m2}} \approx 1 \quad \Rightarrow \quad Z_{in} \approx \frac{1}{sC_{gs}(1+\frac{2C_{gd}}{C_{gs}})}$$

  - **For 0.18m CMOS and total gain of 3, impact of C_{gd} is reduced by 50%:**

$$Z_{in} \approx \frac{1}{sC_{gs}1.9}$$

- **Issue: cascoding lowers achievable voltage swing**

Source-Coupled Amplifier (Unilateralization)

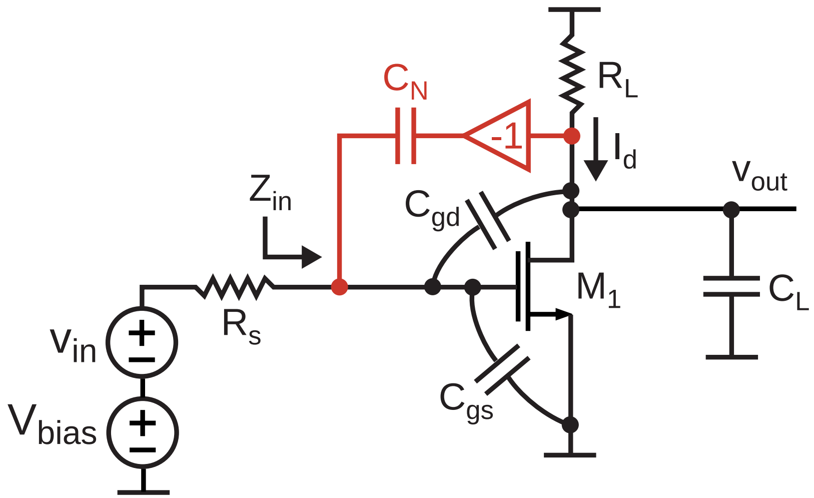


- Remove impact of Miller effect by sending signal through source node rather than drain node
  - C_{gd} not Miller multiplied AND impact of C_{gs} cut in half!

$$Z_{in} \approx \frac{1}{s(C_{gs}/2 + C_{gd})} \quad \Rightarrow \quad Z_{in} \approx \frac{1}{sC_{gs}0.95} \quad (0.18\mu \text{ CMOS})$$

- The bad news
  - Signal has to go through source node (C_{sb} significant)
  - Power consumption doubled

Neutralization



- **Consider canceling the effect of C_{gd}**
  - **Choose $C_N = C_{gd}$**
  - **Charging of C_{gd} now provided by C_N**
- **Benefit: Impact of C_{gd} reduced:**

$$C_{in} \approx C_{gs} + (1 + |A_v|)C_{gd} + (1 - |A_v|)C_{gd} = C_{gs} + 2C_{gd}$$

:same as cascode

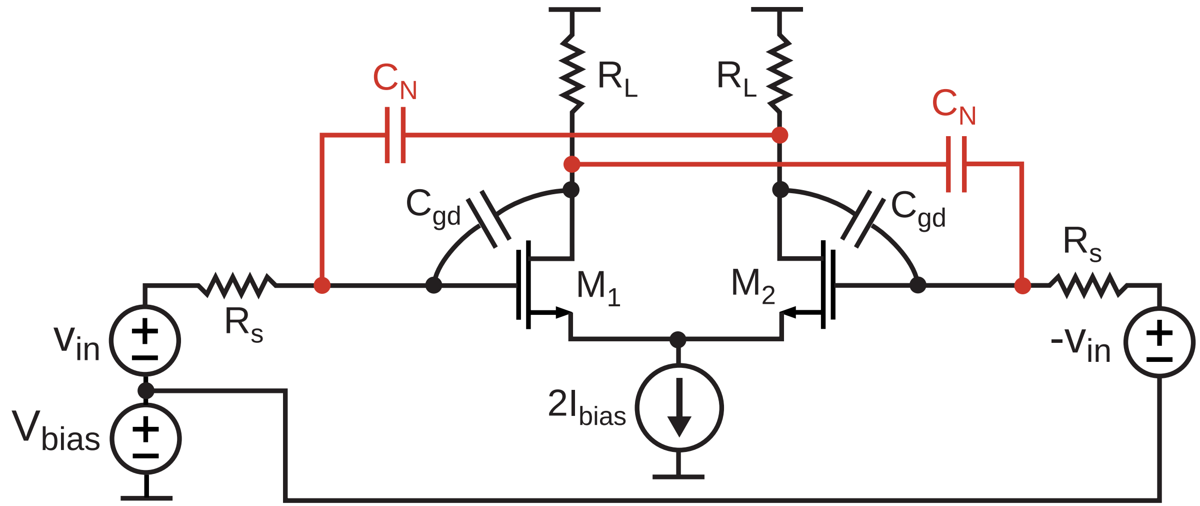
Neutralization, cont'd

- **Issues:**
  - **What happens if C_N is not precisely matched to C_{gd}?**

$$C_{in} \approx C_{gs} + (1 + |A_v|)C_{gd} + (1 - |A_v|)C_N$$

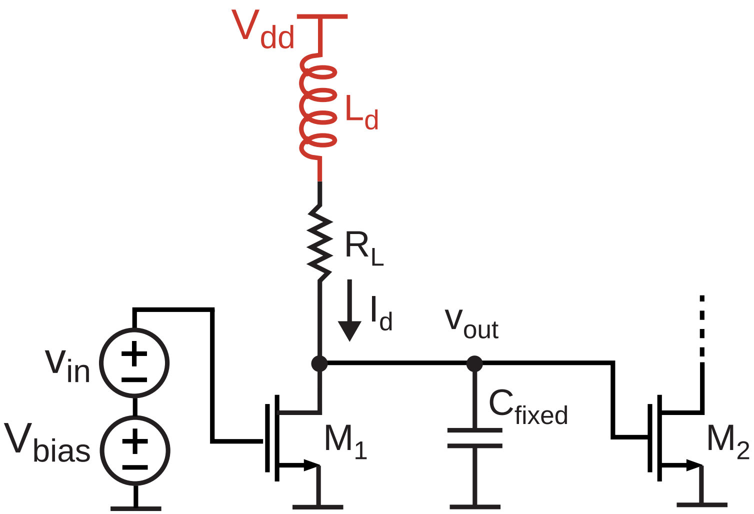
  - **Since the neutralization does not completely remove the effect of C_{gd}, we can make C_N slightly larger than C_{gd} to 'over neutralize'**
  - **Over neutralization can reduce the effect of C_{gs}, but if C_N is too large, the input capacitance is negative and can compromise stability.**
  - **At high frequencies, this can lead to inductive input impedance**
  - **How do we create the inverting amplifier?**

Practical Implementation of Neutralization



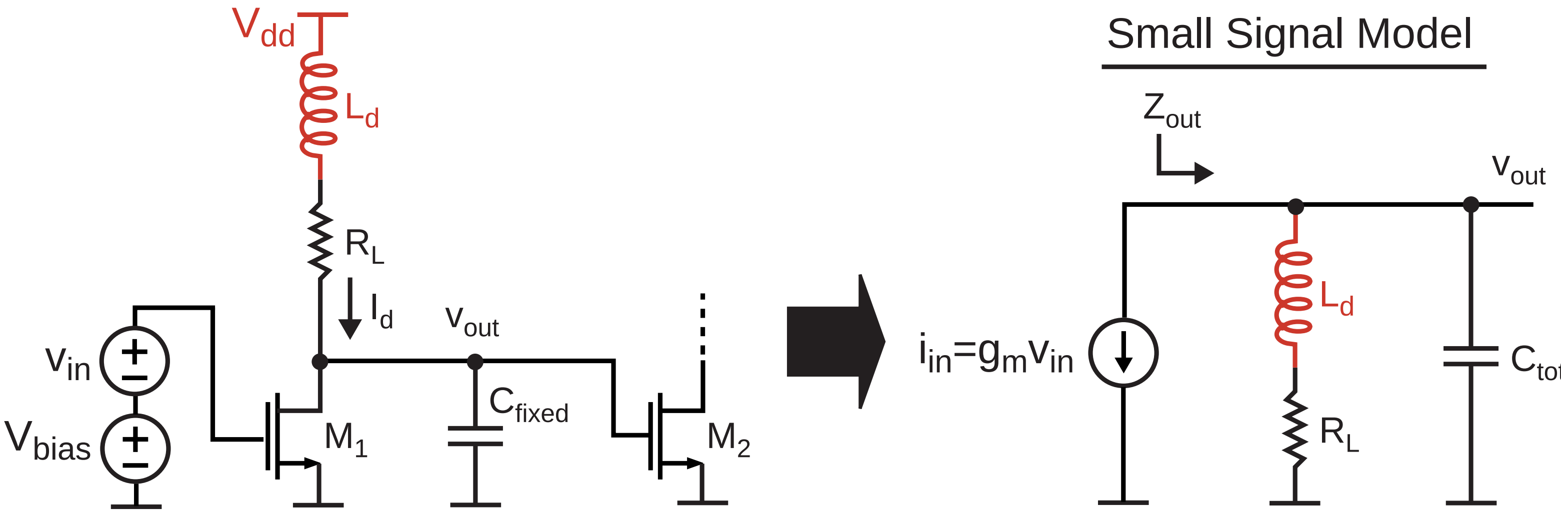
- **Leverage differential signaling to create an inverted signal**
- **Only issue left is matching C_N to C_{gd}**
  - **Often use lateral metal caps for C_N (or CMOS transistor)**
  - **If C_N too low, residual influence of C_{gd}**
  - **If C_N too high, input impedance has inductive component**
    - Causes peaking in frequency response
    - Often evaluate acceptable level of peaking using eye diagrams

Shunt-peaked Amplifier



- **Use inductor in load to extend bandwidth**
  - **Often implemented as a spiral inductor**
- **We can view impact of inductor in both time and frequency**
  - **In frequency: peaking of frequency response**
  - **In time: delay of changing current in R_L**
- **Issue – can we extend bandwidth without significant peaking?**

Shunt-peaked Amplifier - Analysis



- Expression for gain

$$A_v = g_m Z_{out} = g_m[(sL_d + R_L)||1/(sC_{tot})]$$

$$= g_m R_L \frac{s(L_d/R_L) + 1}{s^2 L_d C_{tot} + sR_L C_{tot} + 1}$$

- Parameterize with

$$m = \frac{R_L C_{tot}}{\tau}, \quad \text{where } \tau = \frac{L_d}{R_L}$$

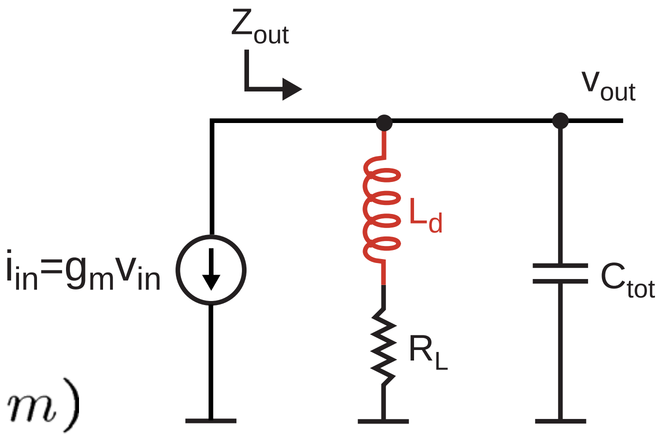
  - Corresponds to ratio of RC to LR time constants

The Impact of Choosing Different Values of m – Part 1

- **Parameterized gain expression**

$$A_v = g_m R_L \frac{\tau s + 1}{s^2\tau^2 m + s\tau m + 1}$$

- **Comparison of new and old 3 dB frequencies**



set: $s = jw, w_1 = \frac{1}{RC}, \tau = 1/(w_1 m)$

$$|A_v| = g_m R_L \left| \frac{jw/(w_1 m) + 1}{-(w/(w_1 m))^2 m + jw/(w_1 m)m + 1} \right|$$

define w_2 as new 3 dB frequency, note that w_1 is old one

$$\Rightarrow \left| \frac{jw_2/(w_1 m) + 1}{-(w_2/(w_1 m))^2 m + jw_2/w_1 + 1} \right| = \frac{1}{\sqrt{2}}$$

- **Want to solve for ω_2/ω_1**

The Impact of Choosing Different Values of m – Part 2

- **From previous slide, we have**

$$\left|\frac{jw_2/(w_1 m)+1}{-(w_2/(w_1 m))^2 m + jw_2/w_1 + 1}\right| = \frac{1}{\sqrt{2}}$$

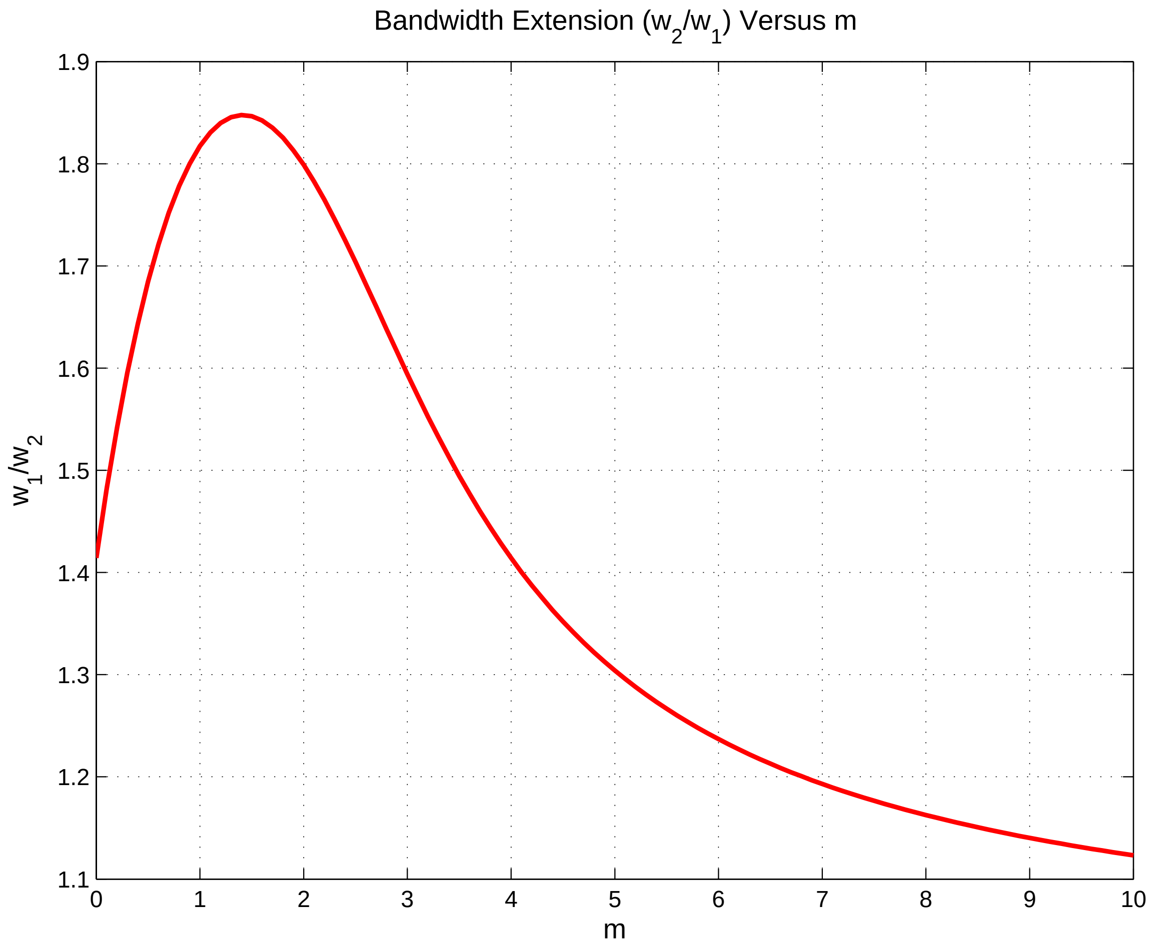
- **After much algebra**

$$\frac{w_2}{w_1} = \sqrt{\left(-\frac{m^2}{2}+m+1\right)+\sqrt{(-\left(\frac{m^2}{2}+m+1\right)^2+m^2}}$$

- **We see that *m* directly sets the amount of bandwidth extension!**
  - **Once m is chosen, inductor value is**

$$L_d = \frac{R_L^2 C_{tot}}{m}$$

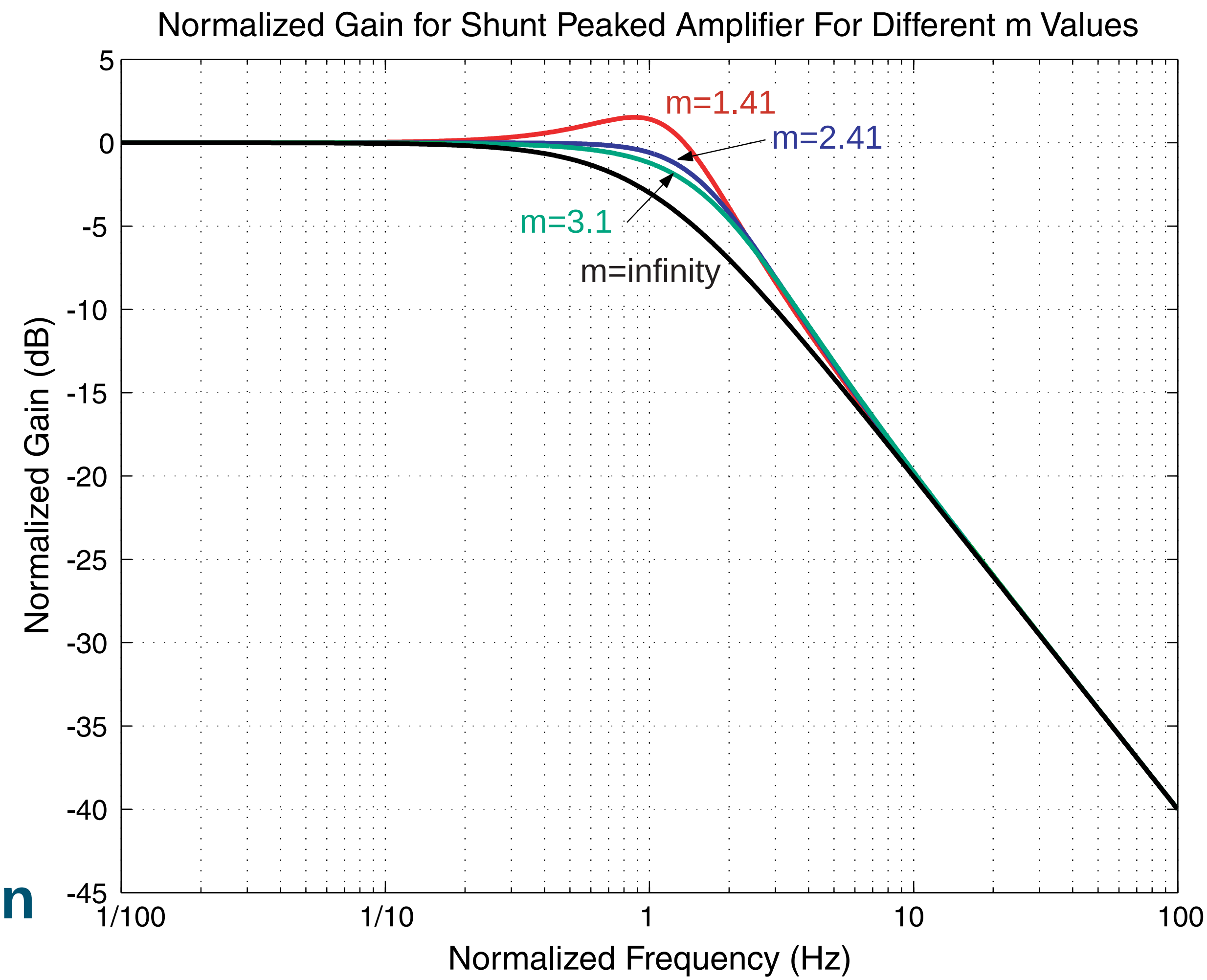
Plot of Bandwidth Extension Versus m



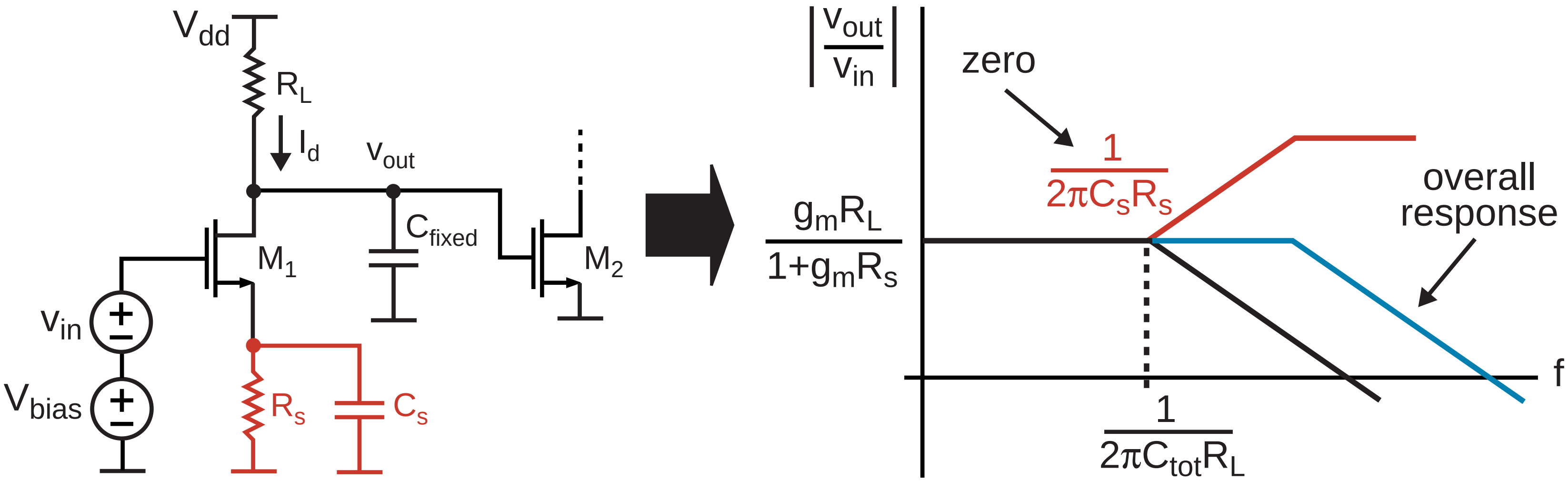
- **Highest extension: ω_2/ω_1 = 1.85 at m $\approx$ 1.41**
  - **However, peaking occurs!**

Plot of Transfer Function Versus m

- **Maximum bandwidth: m = 1.41 (extension = 1.85)**
- **Maximally flat response: m = 2.41 (extension = 1.72)**
- **Best phase response: m = 3.1 (extension = 1.6)**
- **No inductor: m = infinity**
- **Eye diagrams often used to evaluate best m**

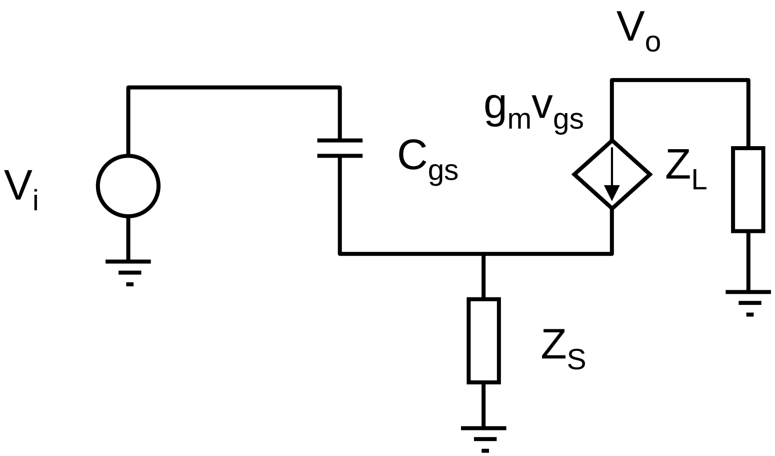


Zero-peaked Common Source Amplifier



- **Inductors are expensive with respect to die area**
- **Can we instead achieve bandwidth extension with capacitor?**
  - **Idea: degenerate gain at low frequencies, remove degeneration at higher frequencies (i.e., create a zero)**
- **Issues:**
  - **Must increase R_L to keep same gain (lowers pole)**
  - **Lowers achievable gate voltage bias (lowers device f_t)**

Zero-peaked Common Source Amplifier Analysis



$$Z_L = R_L || \frac{1}{sC_{tot}}$$

$$Z_s = R_s || \frac{1}{sC_s}$$

- **Add C_{gd} to C_{tot} (as we did previously)**
- **Ignore the feed-forward effect of C_{gd} (It contributes high frequency zero of little consequence)**
- **Analysis shows**

$$\frac{V_o}{V_i} = -\frac{g_m Z_L}{1 + g_m Z_s + sC_{gs} Z_s}$$

Zero-peaked Amplifier Analysis Continued

Assuming $\omega << \omega_T$ $\quad g_m >> sC_{gs}$

Transfer function can now be simplified to

$$\frac{V_o}{V_i} \approx -\frac{g_m R_L (1 + sR_s C_s)}{(1 + sR_L C_{tot})(1 + sR_s C_s + g_m R_s)}$$

Adds a zero at $1/R_s C_s$, but introduces a 2^{nd} pole at $\frac{1 + g_m R_s}{R_s C_s}$

Reduces low freq. gain to $\frac{g_m R_L}{1 + g_m R_s}$

The 1^{st} pole at $1/R_L C_{tot}$ can be cancelled by making $R_s C_s = R_L C_{tot}$, then the bandwidth is extended to the 2^{nd} pole

Zero-peaked Amplifier Continued

- **Pole-zero cancellation:**

$$R_s C_s = R_L C_{tot}$$

$$\frac{V_o}{V_i} = -\frac{g_m R_L}{1 + sR_sC_s + g_mR_s}$$

$$A_{vdc} = \frac{g_m R_L}{1 + g_m R_s} \quad \omega_h = \frac{1 + g_m R_s}{R_s C_s} = \frac{1 + g_m R_s}{R_L C_{tot}}$$

- **Does it really help the bandwidth?**

  If we designed the simple CS amplifier for the same gain, what would be the bandwidth? We need to first reduce R_L to

$$R_L' = \frac{R_L}{1 + g_m R_s}$$

The bandwidth is then $\omega_h = \frac{1}{R_L' C_{tot}} = \frac{1 + g_m R_s}{R_L C_{tot}}$

Same as zero-peaked amplifier!

Zero-peaked Amplifier Input impedance

Input Impedance (ignoring Miller effect for now)

$$Z_{in} = \frac{1}{sC_{gs}} + (1 + \frac{g_m}{sC_{gs}})Z_s$$

Again, $\omega << \omega_T$ $\quad g_m >> sC_{gs}$

Also, near the upper 3dB bandwidth, $sR_sC_s >> 1$

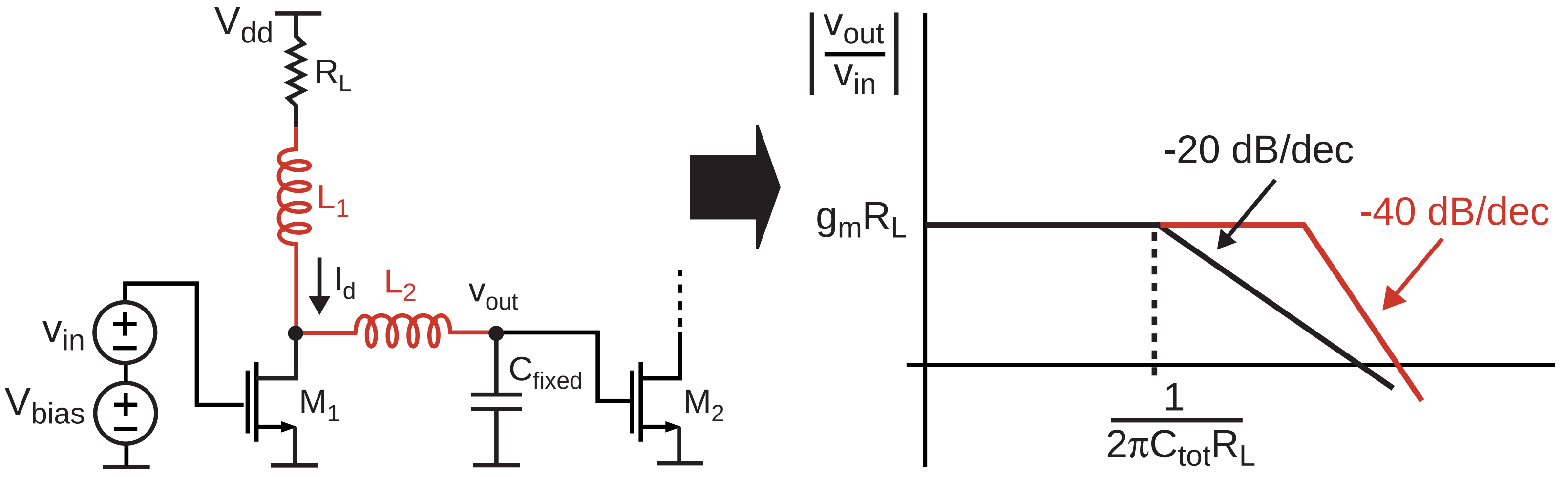
$$Z_{in} \approx \frac{1}{sC_{gs}}(1 + \frac{g_m}{sC_s})$$

$$Z_{in}(jw) \approx \frac{1}{jwC_{gs}} - \frac{g_m}{w^2C_gC_s}$$

Negative resistance!

The negative input resistance component can cause parasitic oscillation. The actual input impedance $Z_{in,tot}$ is the parallel connection between Z_{in} and KC_{gd}.

Back to Inductors – Shunt and Series Peaking



- **Combine shunt peaking with a series inductor**
  - **Bandwidth extension by converting to a second order filter response**
    - Can be designed for proper peaking
- **Increases delay of amplifier**

Refer to Tom Lee's book pp. 279-280 (2nd ed.) or 187-189 (1st ed.)