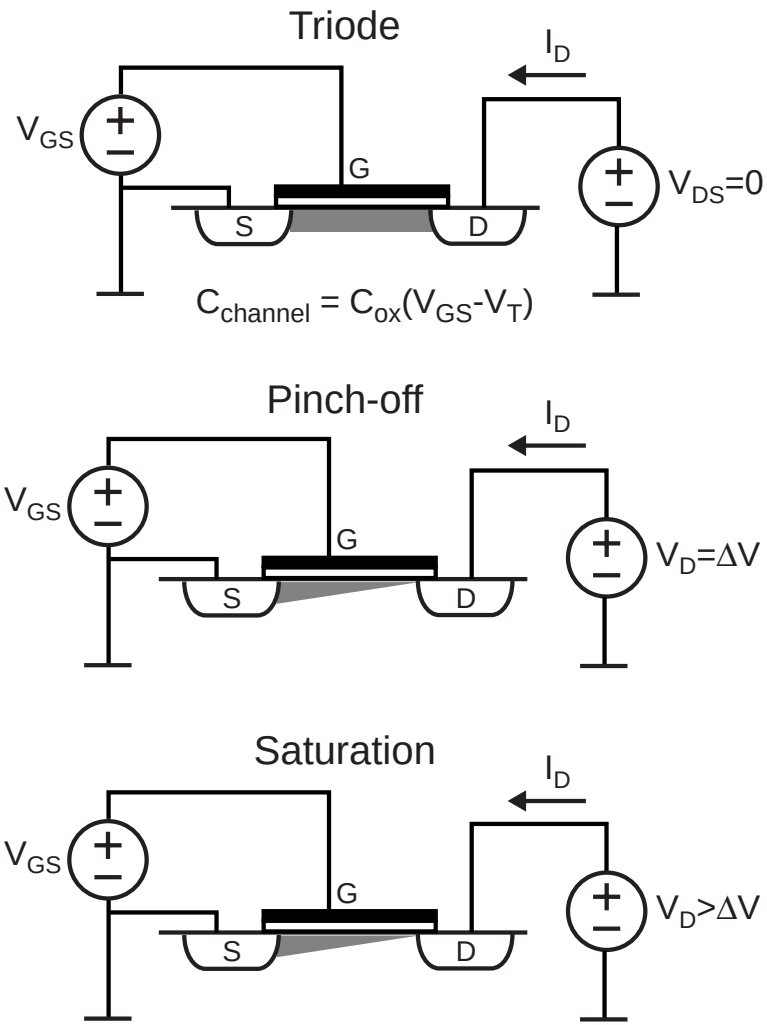
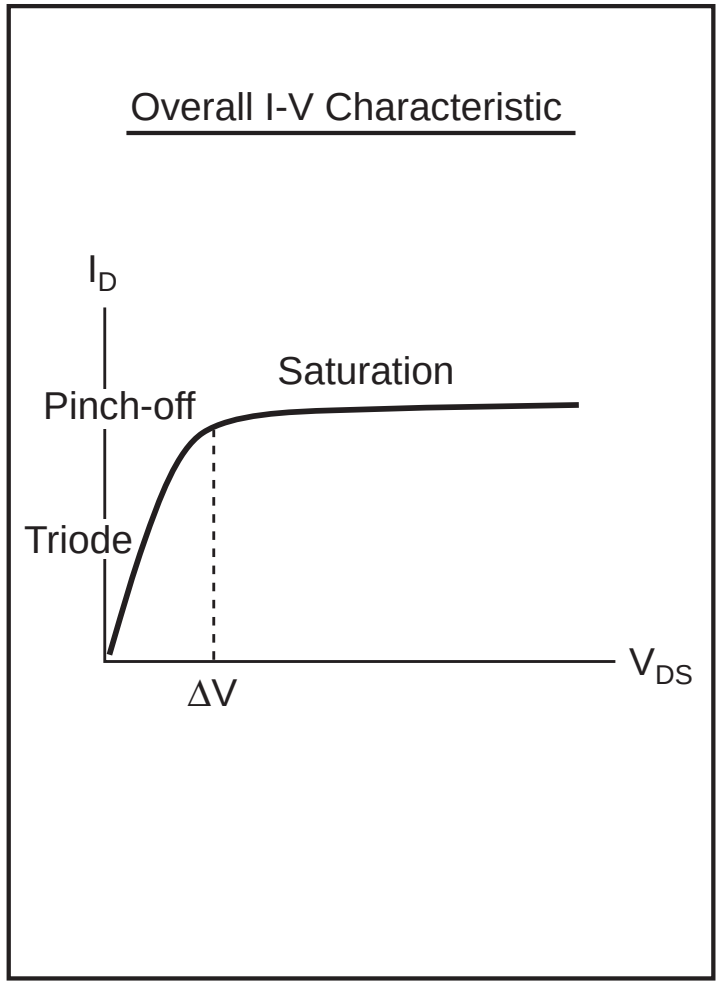
6.776
High Speed Communication Circuits
Lecture 6
MOS Transistors, Passive Components, Gain-Bandwidth Issue for Broadband Amplifiers

Massachusetts Institute of Technology

February 17, 2005

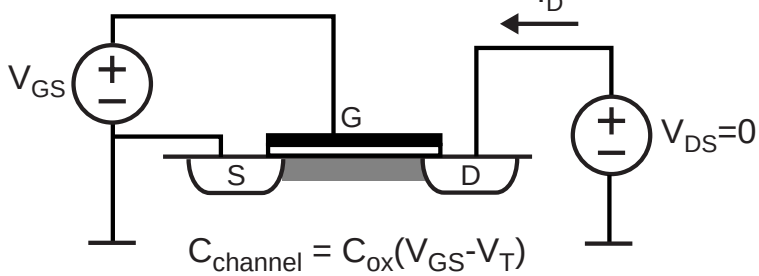
Basics of MOS Large Signal Behavior (Qualitative)





Basics of MOS Large Signal Behavior (Quantitative)

Triode

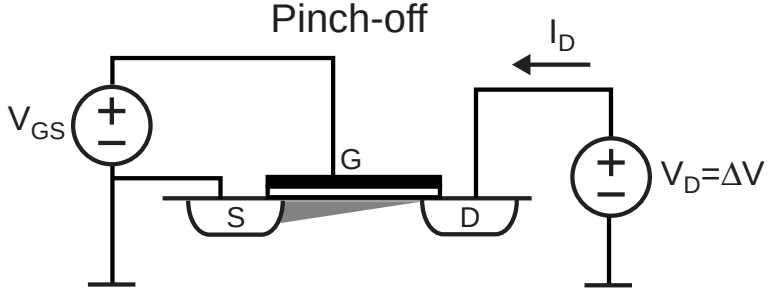


$$I_D = \mu_n C_{ox}\frac{W}{L}(V_{GS} - V_T - V_{DS}/2)V_{DS}$$

$$\text{for } V_{DS} << V_{GS} - V_T$$

$$I_D \approx \mu_n C_{ox}\frac{W}{L}(V_{GS} - V_T)V_{DS}$$

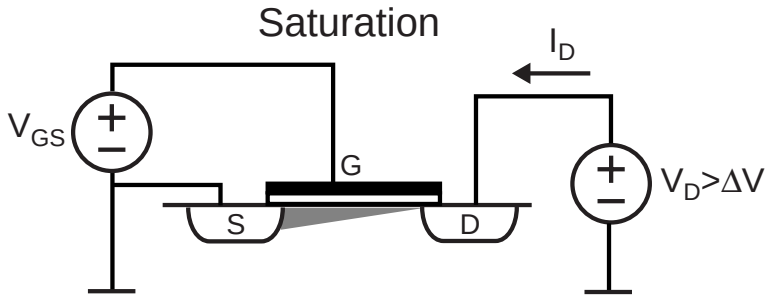
Pinch-off



$$\Delta V = V_{GS}-V_T$$

$$\Delta V = \sqrt{\frac{2I_D L}{\mu_n C_{ox} W}}$$

Saturation



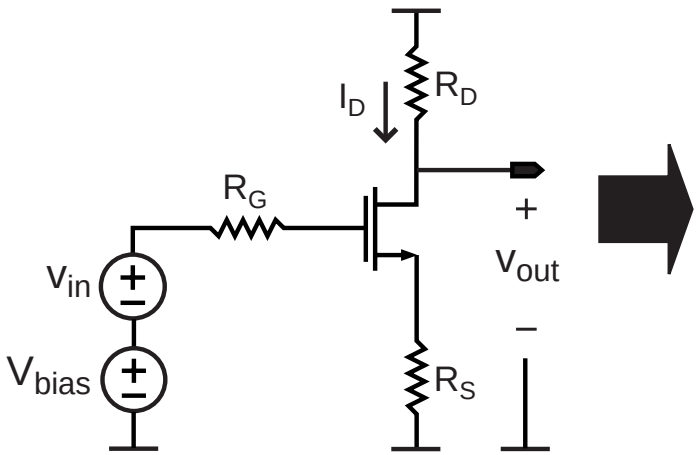
$$I_D = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}(V_{GS}-V_T)^2(1+\lambda V_{DS})$$

(where λ corresponds to channel length modulation)

Analysis of Amplifier Behavior

- **Typically focus on small signal behavior**
  - **Work with a linearized model such as hybrid-π**

- **To do small signal analysis:**

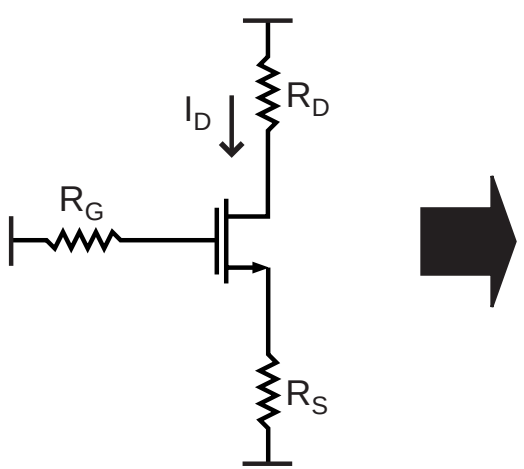


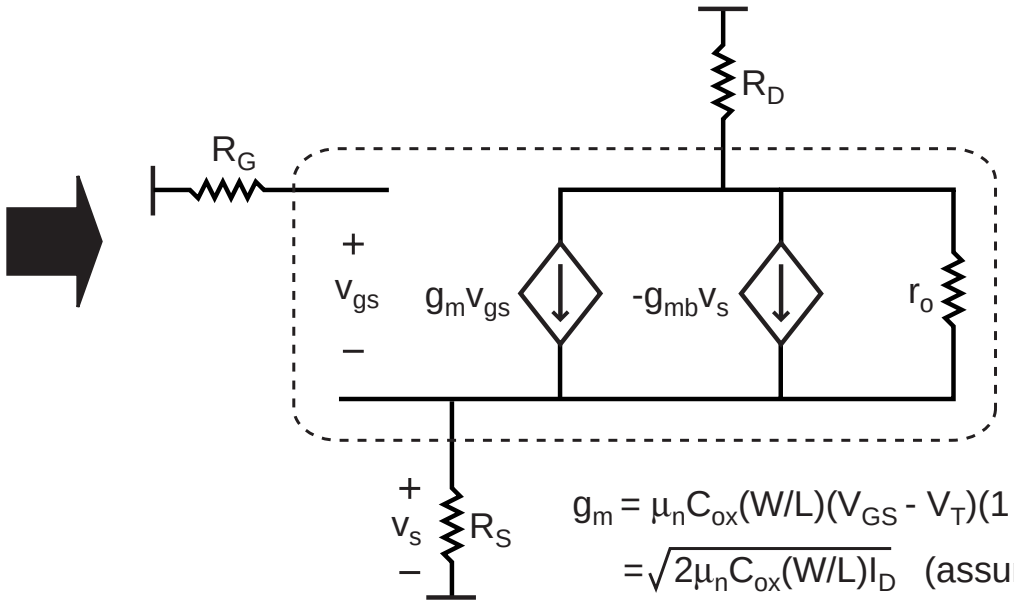
Small Signal Analysis Steps

1) Solve for bias current I_d
2) Calculate small signal parameters (such as g_m, r_o)
3) Solve for small signal response using transistor hybrid-π small signal model

MOS DC Small Signal Model

- **Assume transistor in saturation:**





$$g_m = \mu_n C_{ox}(W/L)(V_{GS} - V_T)(1 + \lambda V_{DS})$$

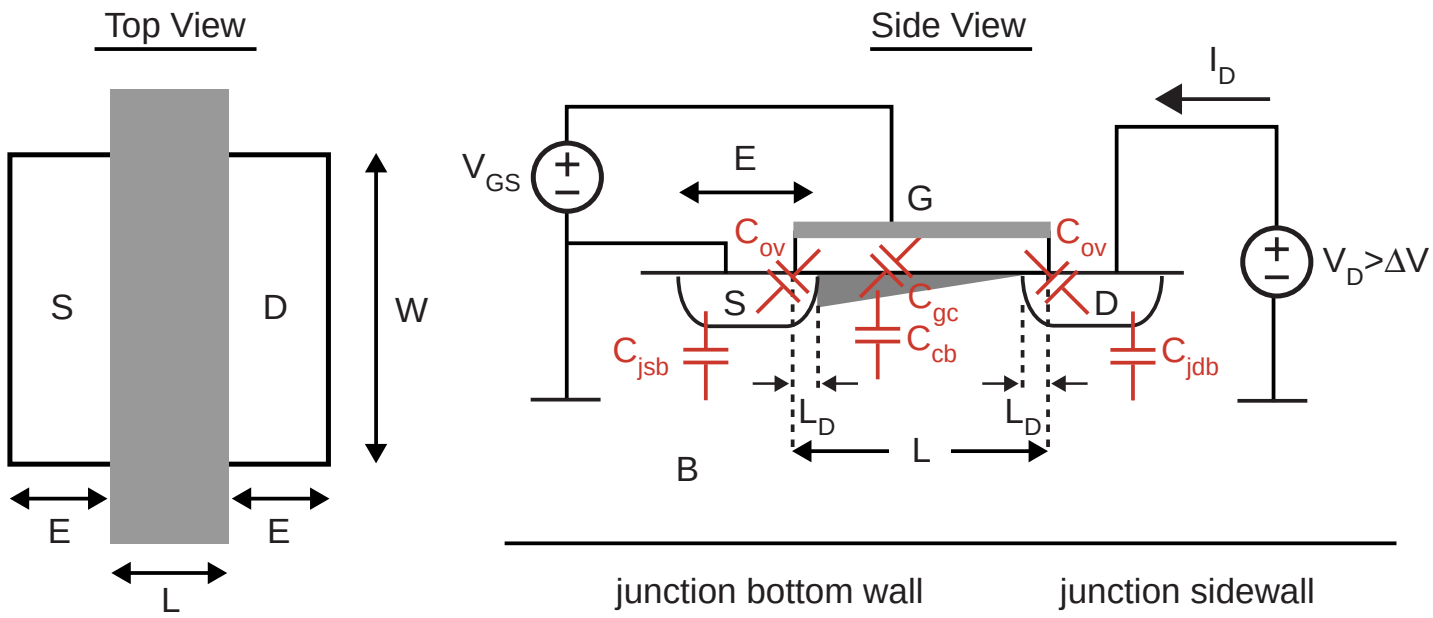
$$= \sqrt{2\mu_n C_{ox}(W/L) I_D} \quad \text{(assuming } \lambda V_{DS} << 1\text{)}$$

$$g_{mb} = \frac{\gamma g_m}{2\sqrt{2|\Phi_p| + V_{SB}}} \quad \text{where } \gamma = \frac{\sqrt{2q\varepsilon_s N_A}}{C_{ox}}$$

In practice: $g_{mb} = g_m/5$ to $g_m/3$

$$r_o = \frac{1}{\lambda I_D}$$

Capacitors For MOS Device In Saturation



junction bottom wall cap (per area)

junction sidewall cap (per length)

source to bulk cap: $C_{jsb} = \frac{C_j(0)}{\sqrt{1 + V_{SB}/|\Phi_B|}} WE + \frac{C_{jsw}(0)}{\sqrt{1 + V_{SB}/|\Phi_B|}} (W + 2E)$

drain to bulk cap: $C_{jdb} = \frac{C_j(0)}{\sqrt{1 + V_{DB}/|\Phi_B|}} WE + \frac{C_{jsw}(0)}{\sqrt{1 + V_{DB}/|\Phi_B|}} (W + 2E)$

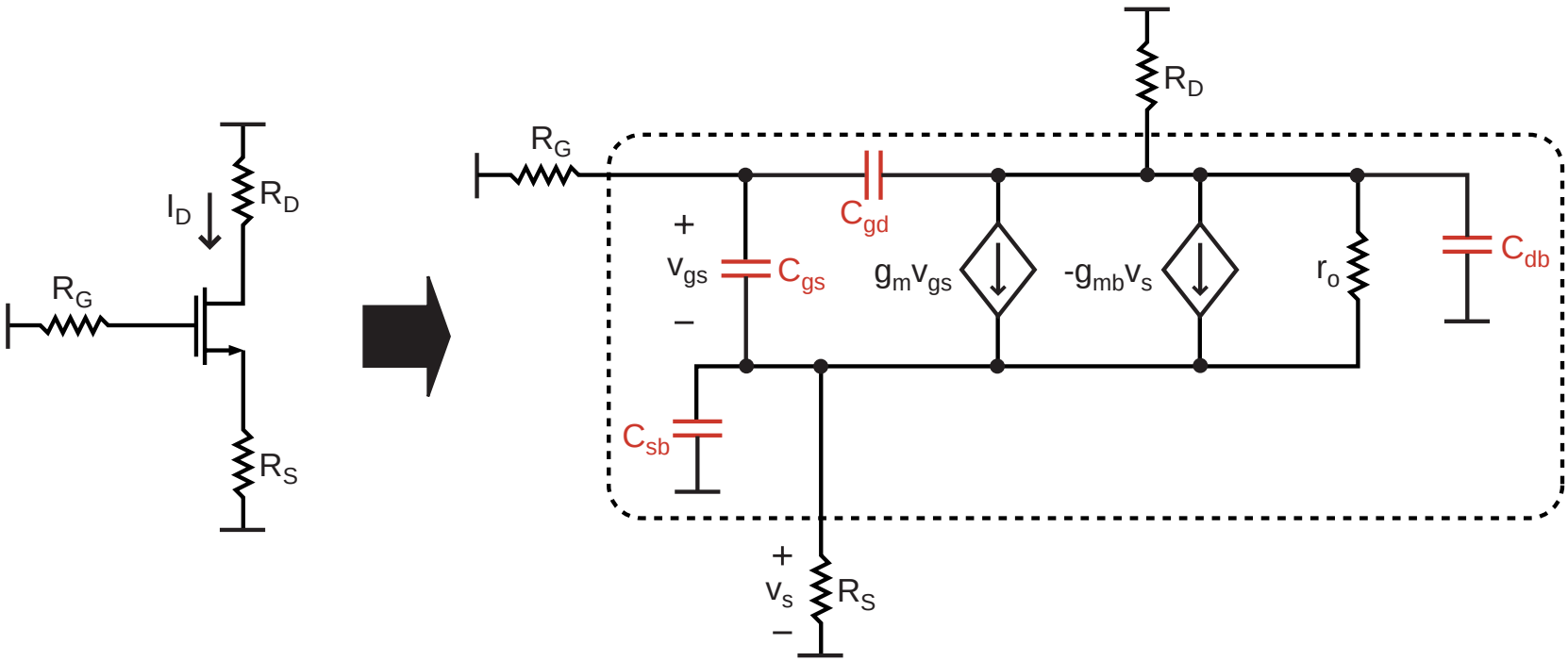
(make 2W for "4 sided" perimeter in some cases)

overlap cap: $C_{ov} = WL_DC_{ox} + WC_{fringe}$

gate to channel cap: $C_{gc} = \frac{2}{3} C_{ox}W(L-2L_D)$

channel to bulk cap: C_{cb} - ignore in this class

MOS AC Small Signal Model (Device in Saturation)



$$C_{gs} = C_{gc} + C_{ov} = \frac{2}{3} C_{ox} W(L\text{-}2L_D) + C_{ov}$$

$$C_{gd} = C_{ov}$$

$C_{sb} = C_{jsb}$ (area + perimeter junction capacitance)

$C_{db} = C_{jdb}$ (area + perimeter junction capacitance)

Wiring Parasitics

- Capacitance
  - Gate: cap from poly to substrate and metal layers
  - Drain and source: cap from metal routing path to substrate and other metal layers
- Resistance
  - Gate: poly gate has resistance (reduce by silicide) long metal lines can add resistance
  - Drain and source: some resistance in diffusion region (reduce by silicide), and from routing long metal lines
- Inductance
  - Gate: poly gate has negligible inductance, but long wires can add inductance
  - Drain and source: becomes an issue for long wires

Extract these parasitics from circuit layout

Frequency Performance of a CMOS Device

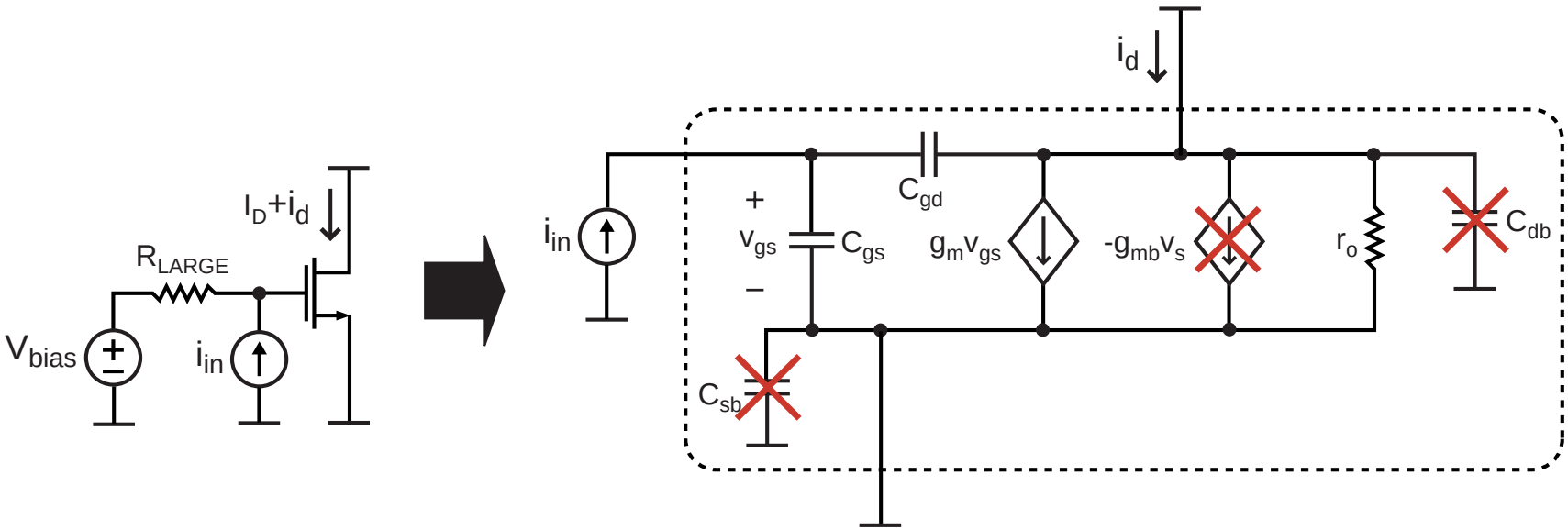
- **Two figures of merit in common use**
  - **f_t : frequency for which current gain is unity**
  - **f_{max} : frequency for which power gain is unity**

- **Common intuition about f_t**
  - **Gain, bandwidth product is conserved**

$$\Rightarrow \text{Gain} \cdot \text{Bandwidth} = f_t$$

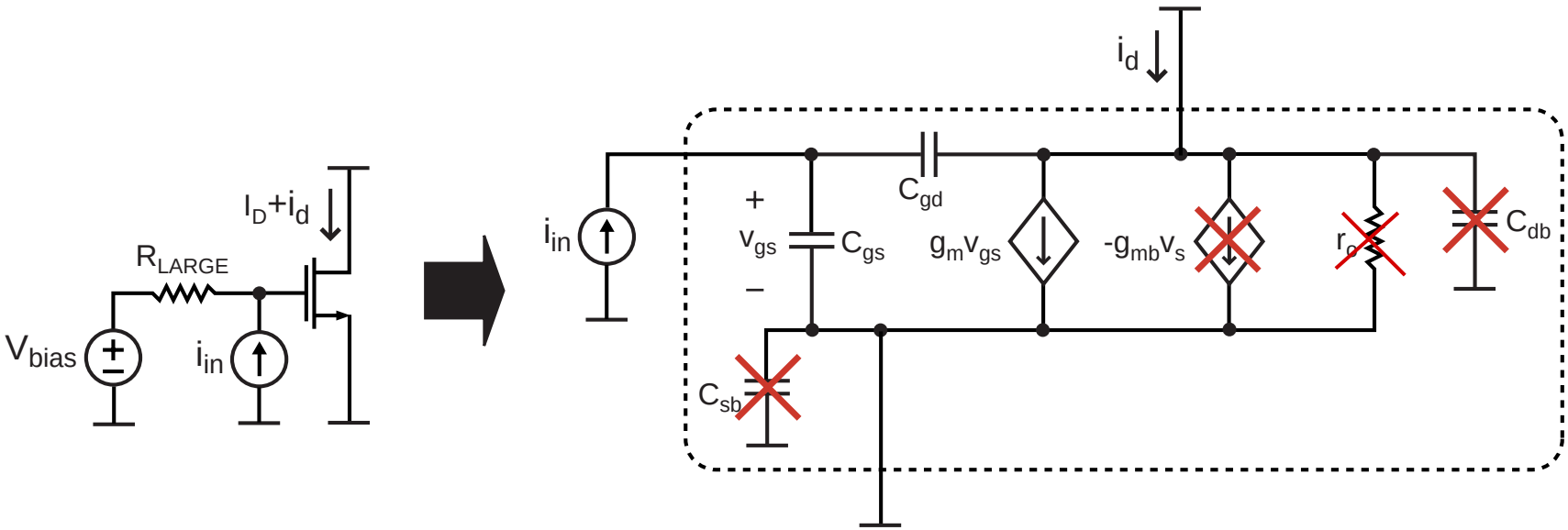
  - **We will see that MOS devices have an f_t that is a function of bias**
    - This effect strongly impacts high frequency amplifier topology selection

Derivation of f_t for MOS Device in Saturation



- **Assumption is that input is current, output of device is short circuited to a supply voltage**
  - **Note that voltage bias is required at gate**
    - The calculated value of f_t is a function of this bias voltage

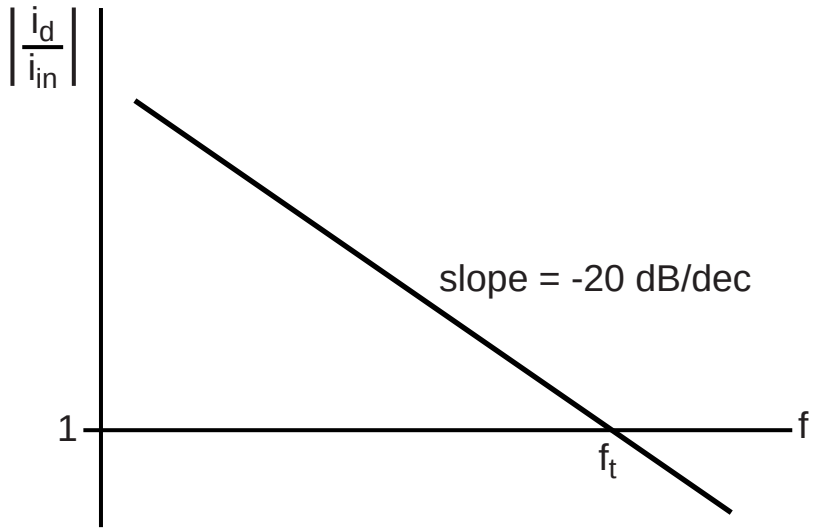
Derivation of f_t for MOS Device in Saturation



$$i_d = g_m v_{gs} = g_m \left(\frac{1}{s(C_{gs} + C_{gd})} \right) i_{in}$$

$$\Rightarrow \frac{i_d}{i_{in}} = \frac{g_m}{j\omega(C_{gs} + C_{gd})}$$

Derivation of f_t for MOS Device in Saturation



$$\omega_t = \frac{g_m}{C_{gs} + C_{gd}}$$

$$\Rightarrow \quad f_t = \frac{g_m}{2\pi(C_{gs} + C_{gd})}$$

$$i_d = g_m v_{gs} = g_m \left(\frac{1}{s(C_{gs} + C_{gd})} \right) i_{in}$$

$$\Rightarrow \quad \frac{i_d}{i_{in}} = \frac{g_m}{j\omega(C_{gs} + C_{gd})}$$

Why is f_t a Function of Voltage Bias?

$$f_t = \frac{g_m}{2\pi(C_{gs} + C_{gd})}$$

- f_t is a ratio of g_m to gate capacitance
  - g_m is a function of gate bias, while gate cap is not (in strong inversion)
- First order relationship between g_m and gate bias:

$$g_m \approx \mu_n C_{ox} \frac{W}{L}(V_{gs} - V_t)$$

  - The larger the gate bias, the higher the value for f_t
- Alternately, f_t is a function of current density

$$\frac{g_m}{C_{gs} + C_{gd}} \approx \frac{\sqrt{2\mu_n C_{ox}(W/L)I_d}}{(2/3)WLC_{ox} + W(C_{ov}/W)} \propto \sqrt{\frac{I_d}{W}} \frac{1}{L^{\frac{3}{2}}}$$

  - So f_t maximized at max current density (and minimum L)

Unity Power Gain Frequency f_{max}

From pages 176-178 (2nd ed.) 70-72 (1st ed.) of text book for derivation on f_{max}

$$\omega_{max} \approx \frac{1}{2}\sqrt{\frac{\omega_T}{r_g C_{gd}}}$$

- **r_g is the series parasitic gate resistance**
- **f_{max} can be much higher than f_T: make gate resistance small (by careful layout)**
- **Output capacitance has no effect (can be tuned out by inductor)**

Speed of NMOS Versus PMOS Devices

$$f_t = \frac{g_m}{2\pi(C_{gs} + C_{gd})}$$

- **NMOS devices have much higher mobility than PMOS devices (in typical bulk CMOS processes)**

$$\mu_n \approx 2.5\mu_p \quad \text{for many processes}$$

$$\Rightarrow \; f_t \text{ of NMOS} \approx 2.5 \times f_t \text{ of PMOS}$$

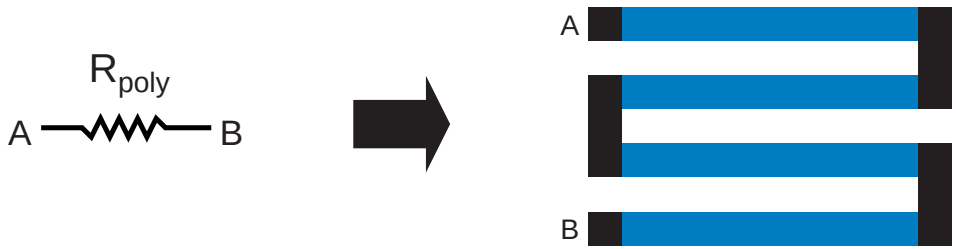
  - **NMOS devices provide approximately 2.5 x g_m for a given amount of capacitance and gate bias voltage**
  - **Also, NMOS devices provide approximately 2.5 x I_d for a given amount of capacitance and gate bias voltage**

Integrated Passive Components for RF Circuits

- **We will only consider passive components appropriate for RF use**
- **High Q, low parasitics, and good linearity are generally desired (bias circuit is an exception)**
- **Well resistors, diffused resistors, poly-n+ capacitors even poly-poly capacitors do not perform very well in these regards**

Polysilicon Resistors

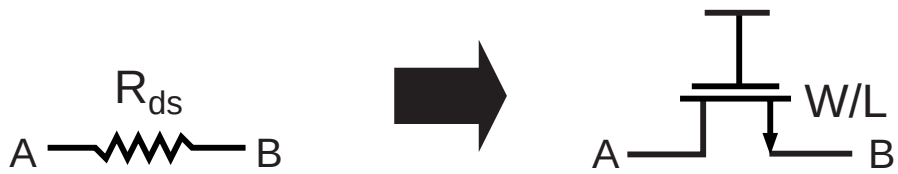
- **Use unsilicided polysilicon to create resistor**



- **Key parameters**
  - **Resistance (usually 100- 200 Ohms per square)**
  - **Parasitic capacitance (usually small)**
    - Appropriate for high speed amplifiers
  - **Linearity (excellent)**
  - **Accuracy (usually can be set within ± 15%)**

MOS Resistors

- Bias a MOS device in its triode region

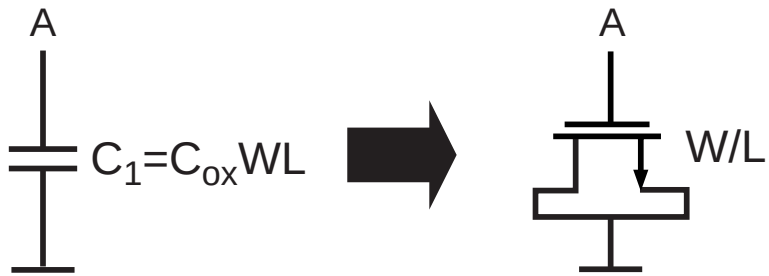


$$R_{ds} \approx \frac{1}{\mu C_{ox} W/L((V_{gs} - V_T) - V_{DS})}$$

- High resistance values can be achieved in a small area (MegaOhms within tens of square microns)
- Parasitic capacitance is large (gate capacitance!)
- Resistance is quite nonlinear
  - Appropriate for small swing circuits or DC (bias) circuits

High Density Capacitors (Biasing, Decoupling)

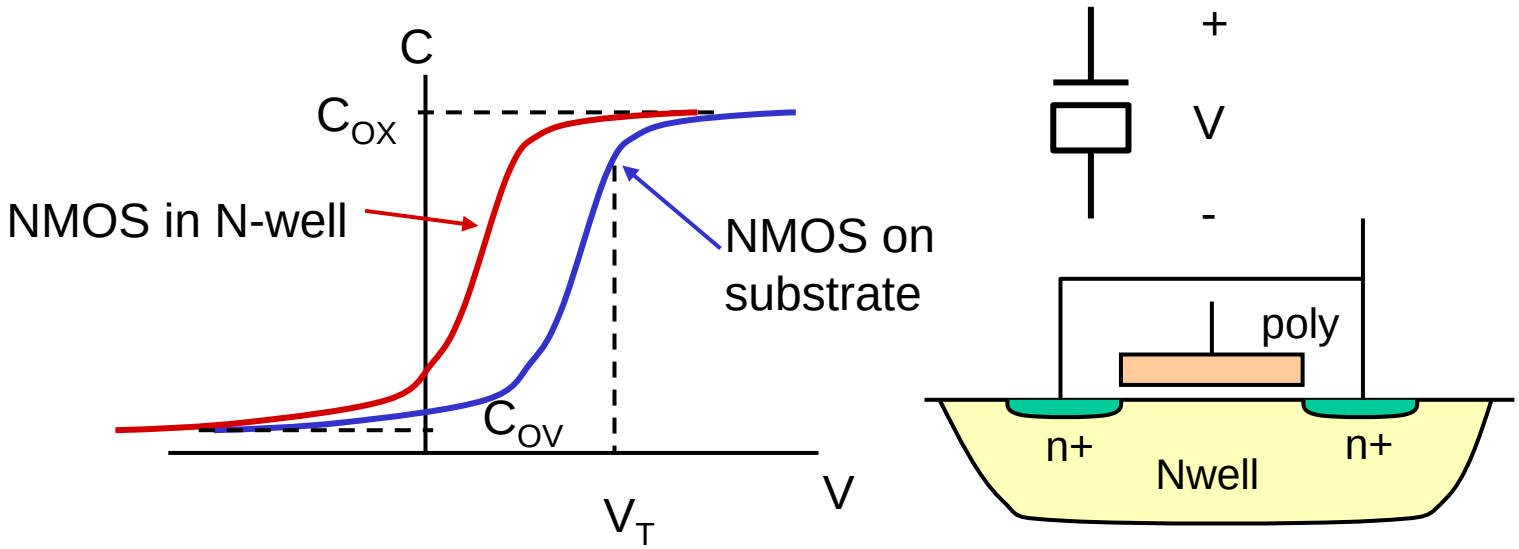
- **MOS devices offer the highest capacitance per unit area**
  - **Voltage must be high enough to invert the channel**



- **Key parameters**
  - **Capacitance value**
    - Raw cap value from MOS device is about 8-8.5 fF/μ^2 for 0.18u CMOS
  - **Q (i.e., amount of series resistance)**
    - Maximized with minimum L (tradeoff with area efficiency)

MOS Capacitors, Cnt'd

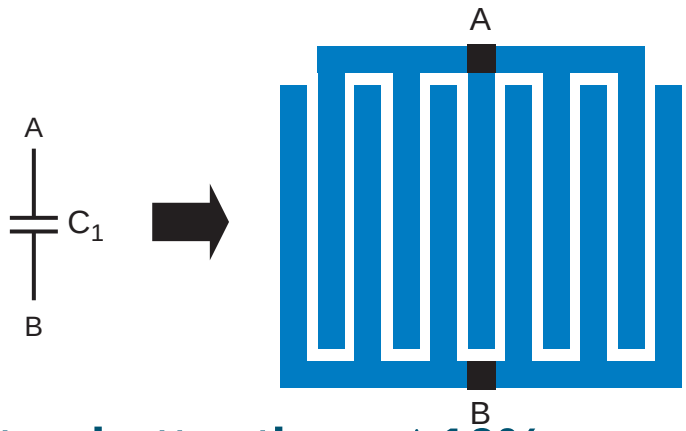
- **Putting NMOS capacitor in NWell allows operation at lower voltage**



- **The non-linearity is often exploited in VCO designs as varactors**

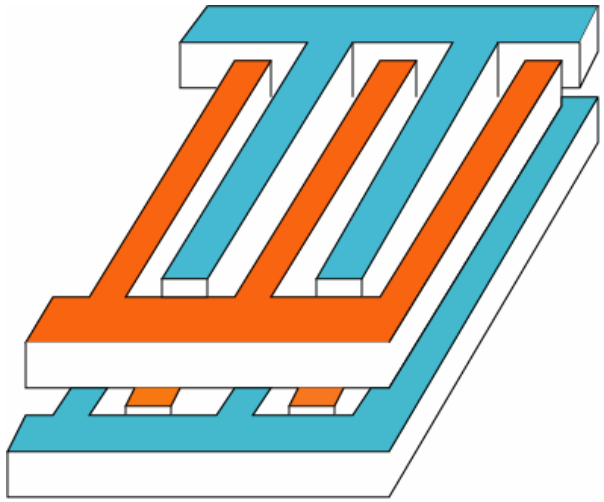
High Q Capacitors (Signal Path)

- **Lateral metal capacitors offer high Q and reasonably large capacitance per unit area**
  - **Stack many levels of metal on top of each other (best layers are the top ones), via them at maximum density**



  - **Accuracy often better than ±10%**
  - **Parasitic cap is symmetric, typically less than 10% of cap value**

Stacked Lateral Flux Capacitor



- **Example: C_T = 1.5 fF/μm^2 for 0.24μm process with 7 metals, $L_{min} = W_{min}$ = 0.24μm, t_{metal} = 0.53μm**
  - See “Capacity Limits and Matching Properties of Integrated Capacitors”, Aparicio et. al., JSSC, Mar 2002

Fractal Capacitor

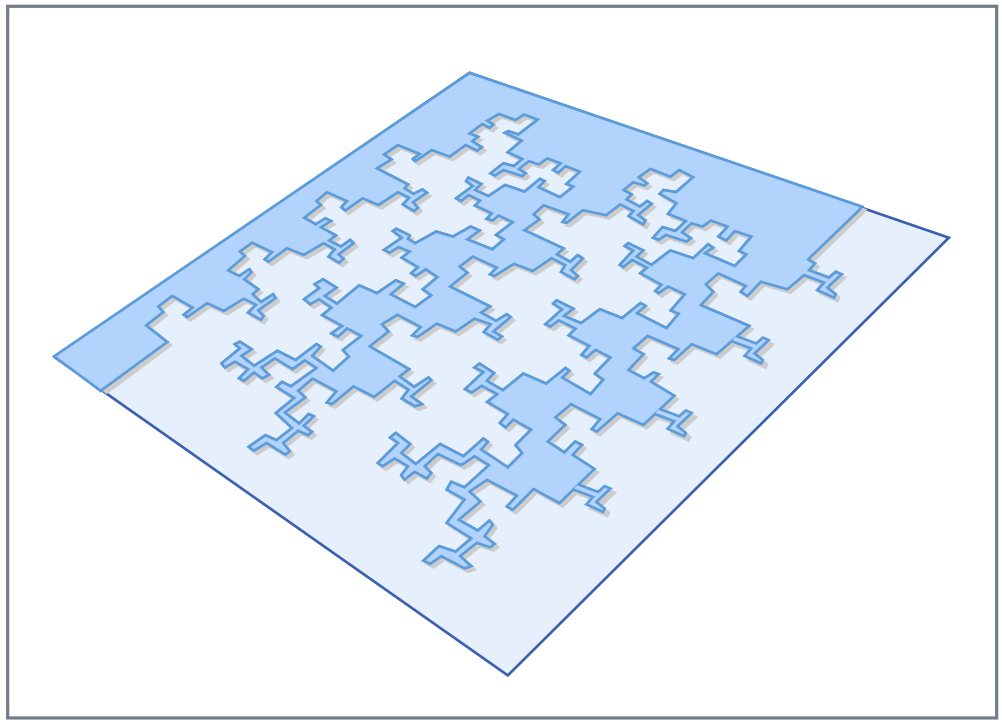


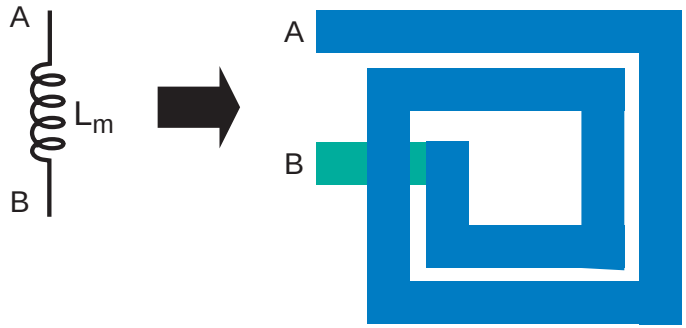
Figure by MIT OCW.

- Maximizes perimeter area: up to 10x increase in unit capacitance
- Limited by lithography

See A. Shanhani et. al., "A 12 mW, Wdie Dynamic Range CMOS Fron-End Circuit for Portable GPS Receiver," Digest of Technical Papers, ISSCC 1997

Spiral Inductors

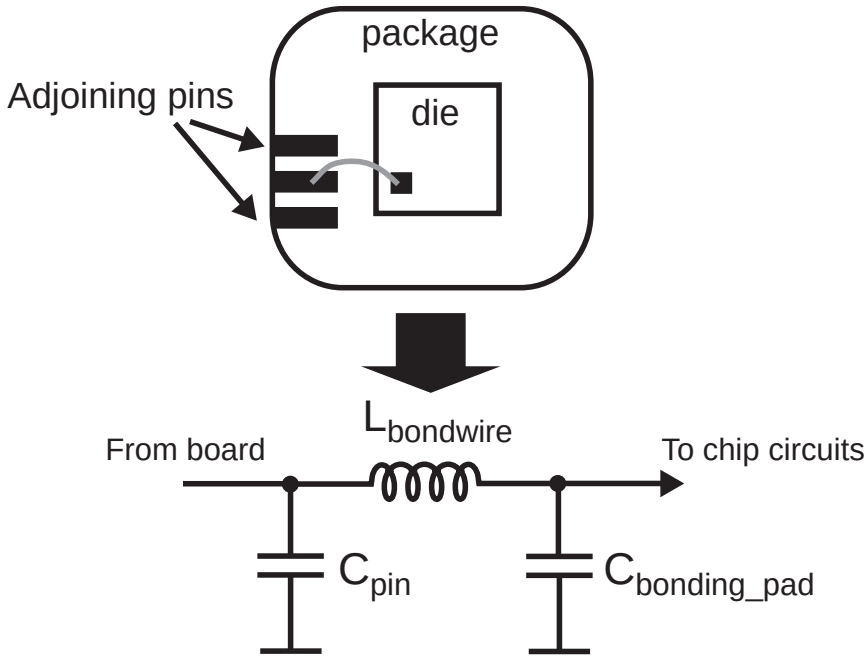
- Create integrated inductor using spiral shape on top level metals (may also want a patterned ground shield)



  - Key parameters are Q (< 10), L (1-10 nH), self resonant freq.
  - Usually implemented in top metal layers to minimize series resistance, coupling to substrate
  - See using Mohan et. al, “Simple, Accurate Expressions for Planar Spiral Inductances, JSSC, Oct, 1999, pp 1419-1424
  - Verify inductor parameters (L, Q, etc.) using ASITIC
    http://formosa.eecs.berkeley.edu/~niknejad/asitic.html

Bondwire Inductors

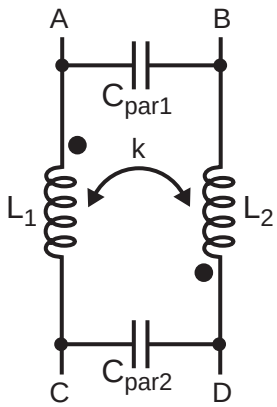
- Used to bond from the package to die
  - Can be used to advantage

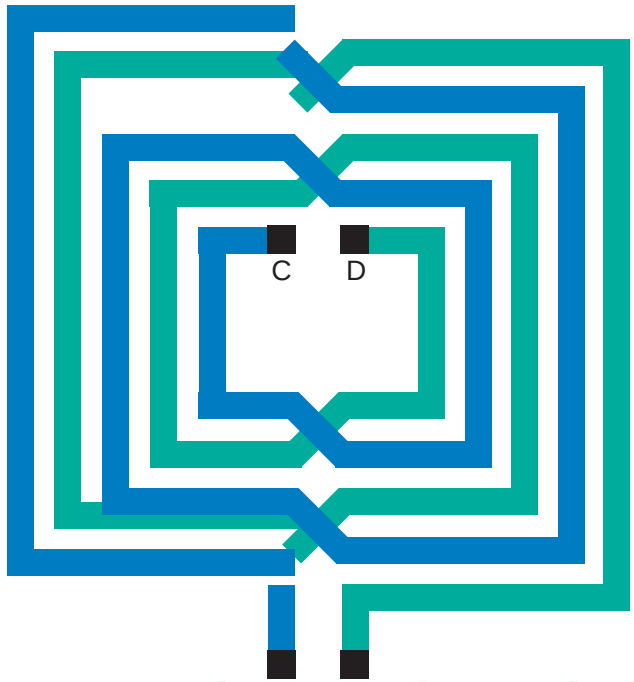


- Properties
  - Inductance ($\approx$ 1 nH/mm – usually achieve 1-5 nH)
  - Inductance value is difficult to control (chip-package alignment, bondwire height, etc.)
  - Q (much higher than spiral inductors – typically > 40)

Integrated Transformers

- Utilize magnetic coupling between adjoining wires





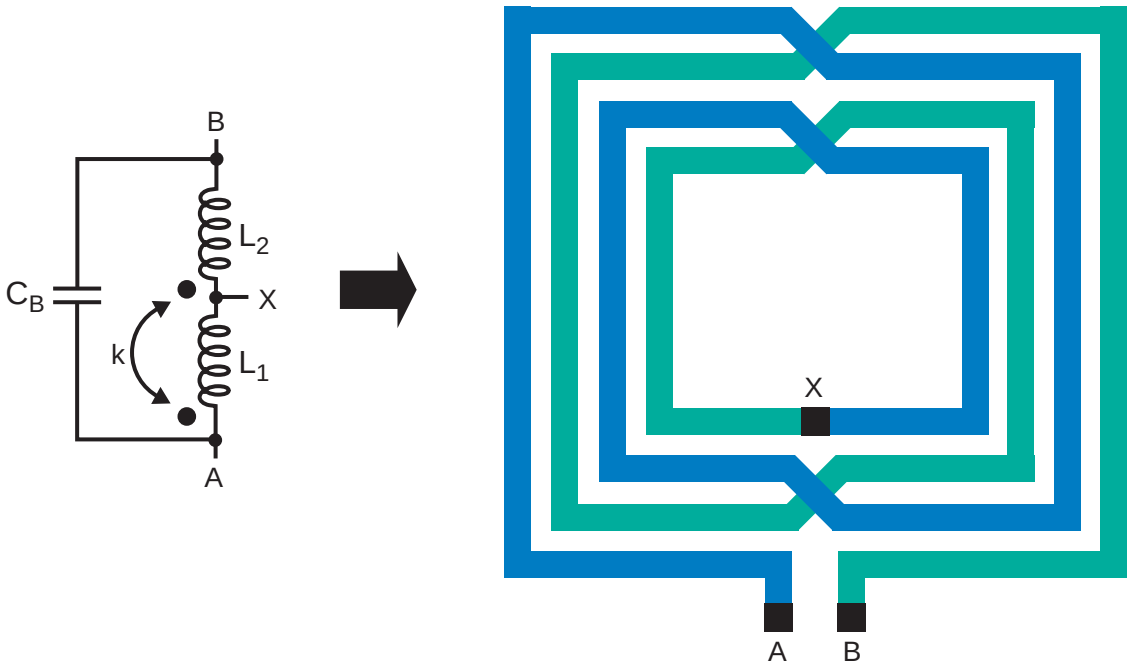
- Key parameters
  - L (self inductance for primary and secondary windings)
  - k (coupling coefficient between primary and secondary)

Note: $k = \dfrac{M}{\sqrt{L_1 L_2}}$ where M = mutual inductance

- Design – ASITIC, other CAD packages

High Speed Transformer Example – A T-Coil Network

- A T-coil consists of a center-tapped inductor with mutual coupling between each inductor half

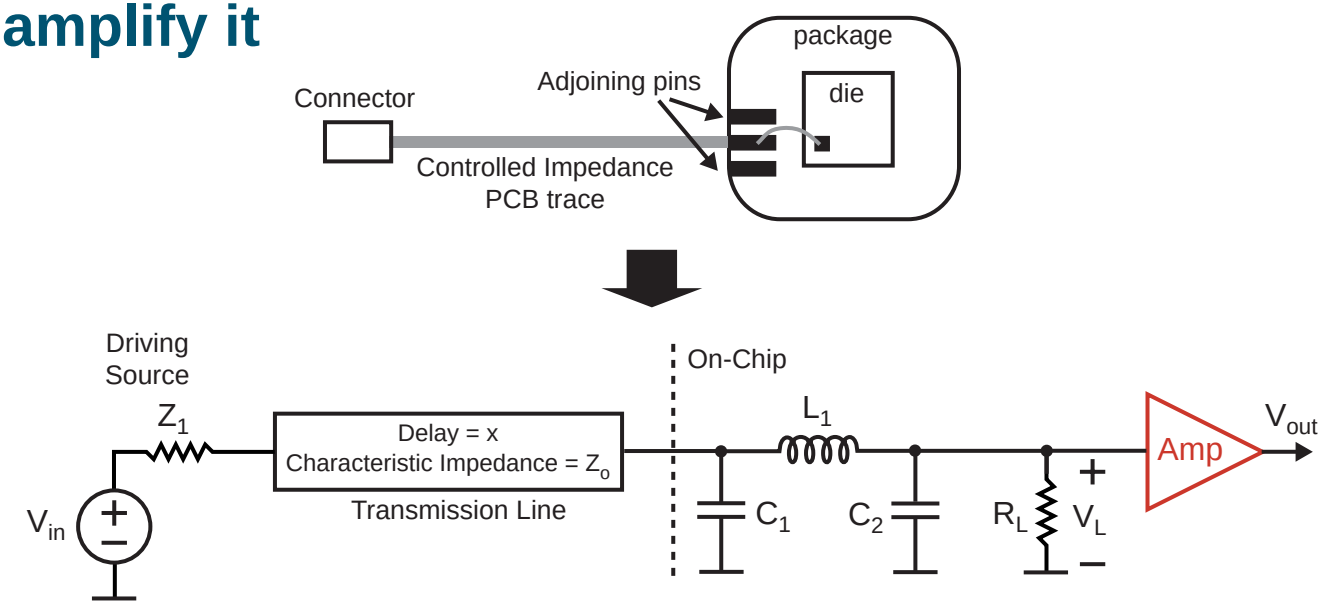


- Used for bandwidth enhancement
  - See S. Galal, B. Ravazi, “10 Gb/s Limiting Amplifier and Laser/Modulator Driver in 0.18u CMOS”, ISSCC 2003, pp 188-189 and “Broadband ESD Protection …”, pp. 182-183

Broadband Amplifiers

High Frequency, Broadband Amplifiers

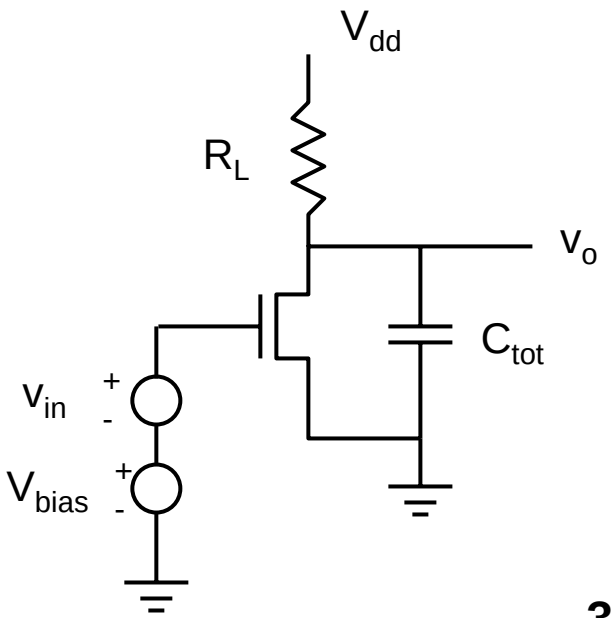
- **The first thing that you typically do to the input signal is amplify it**



- **Function**
  - **Boosts signal levels to acceptable values**
  - **Provides reverse isolation**
- **Key performance parameters**
  - **Gain, bandwidth, noise, linearity**

Gain-bandwidth Trade-off

- **Common-source amplifier example**



C_{tot}: total capacitance at output node

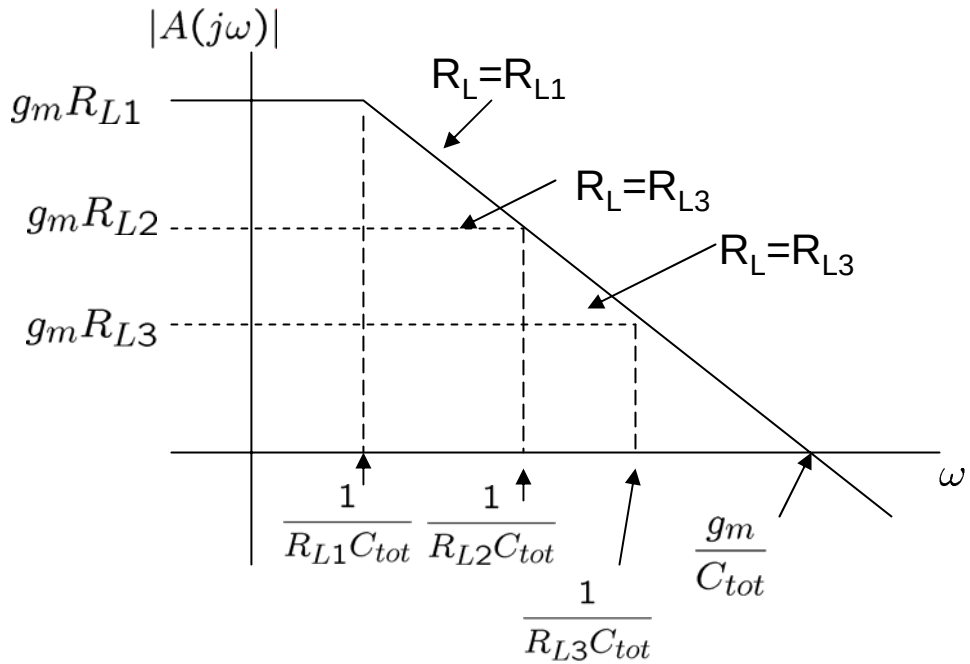
DC gain $A = g_m R_L$

3 dBbandwidth $\omega_h = \dfrac{1}{R_L C_{tot}}$

Gain-bandwidth $GB = \dfrac{g_m}{C_{tot}}$

Gain-bandwidth Trade-off

- **Common-source amplifier example**



- **Given the 'origin pole' g_m/C_{tot}, higher bandwidth is achieved only at the expense of gain**
- **The 'origin pole' g_m/C_{tot} must be improved for better GB**

Gain-bandwidth Improvement

- How do we improve g_m/C_{tot}?
- Assume that amplifier is loaded by an identical amplifier and fixed wiring capacitance is negligible
- Since $g_m = \mu_n C_{ox} \frac{W}{L}(V_{GS} - V_T)$ and $C_{tot} \propto W$

$$\frac{g_m}{C_{tot}} \propto \frac{V_{GS} - V_T}{L}$$

- To achieve maximum GB in a given technology, use minimum gate length, bias the transistor at maximum $V_{GS} - V_T$
- When velocity saturation is reached, higher $V_{GS} - V_T$ does not yield higher g_m
- In case fixed wiring capacitance is large, power consumption must be also considered

Gain-bandwidth Observations

- **Constant gain-bandwidth is simply the result of single-pole role off – it's *not* fundamental!**
- **It implies a single-pole frequency response may not be the best for obtaining gain and bandwidth simultaneously**
- **Single-pole role off is necessary for some circuits, e.g. for stability, but not for broad-band amplifiers**