High-speed CMOS integrated circuit design

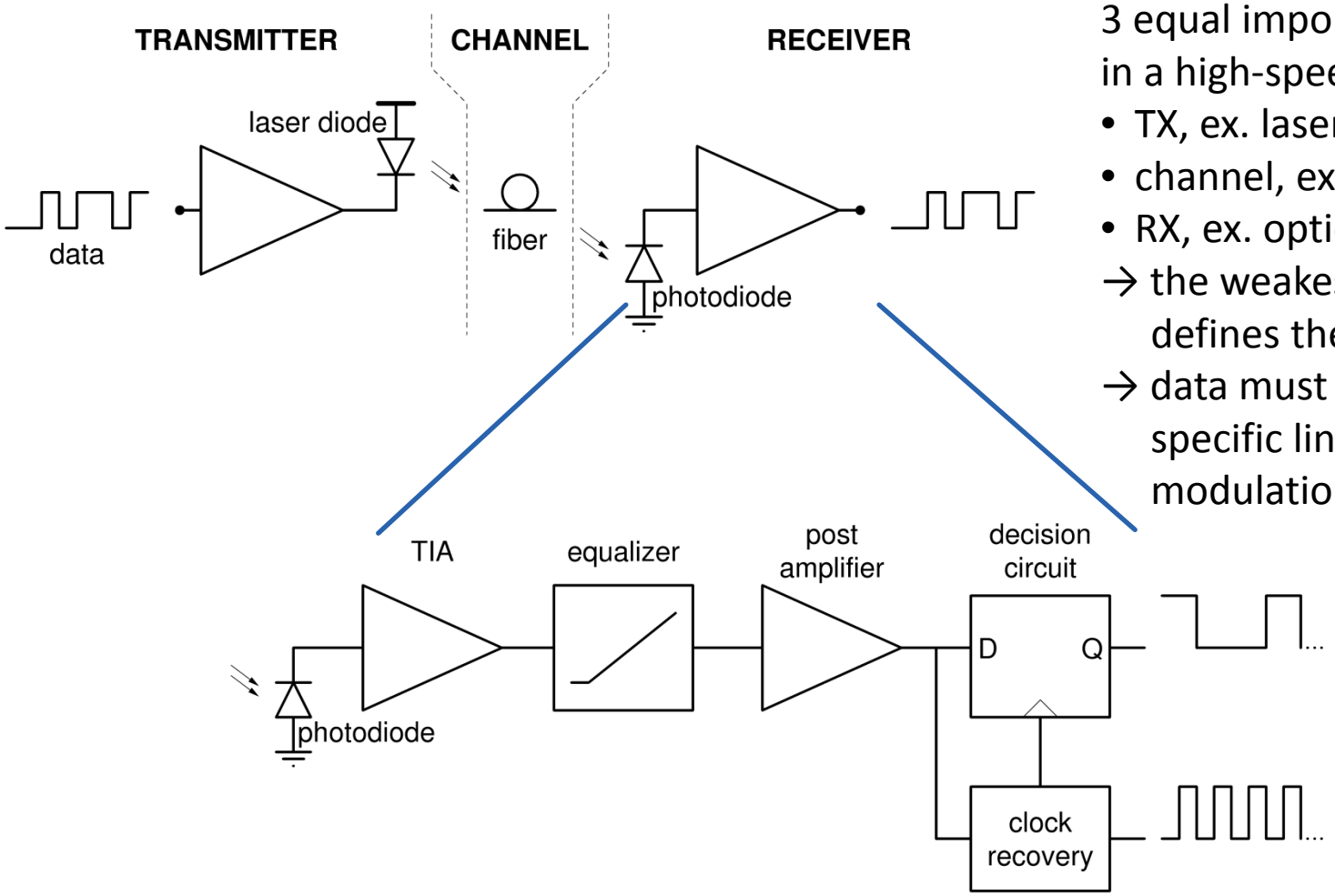
Filip Tavernier

CERN

Outline

- Basic concepts
  - Data properties
  - Circuit noise
  - Bandwidth
  - BER
  - Jitter
  - FEC
- *Circuit concepts*
  - *Common-source amplifier*
  - *Optimum current density*
  - *Bandwidth extension*
  - *Multistage amplifier*
  - *Negative impedances*

High-speed (optical) data communication



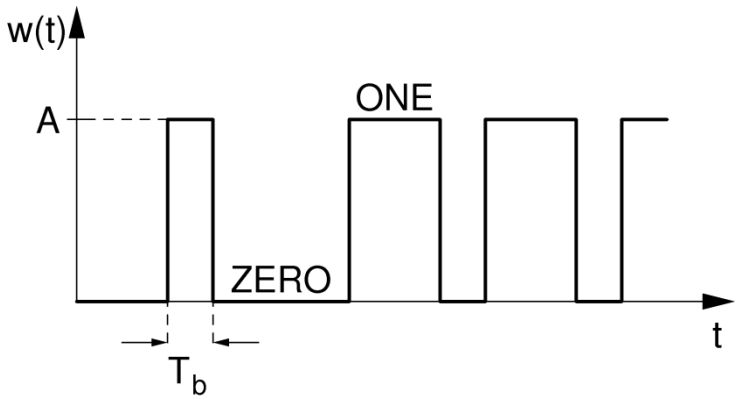
3 equal important components in a high-speed data link:

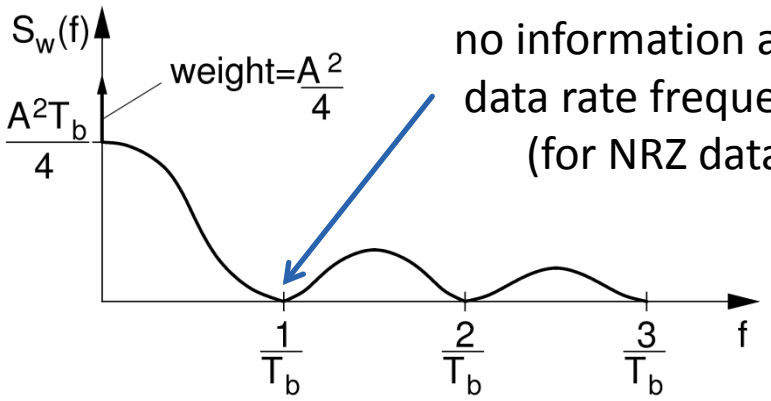
- TX, ex. laser driver
- channel, ex. optical fiber
- RX, ex. optical receiver

→ the weakest component defines the overall performance!

→ data must be conditioned for the specific link (coding, scrambling, modulation, …)

Random digital data





no information at the data rate frequency! (for NRZ data)

$$w(t) = \sum_i b_i \, p(t - iT_b)$$

$$p(t) = \begin{cases} A & \text{if } 0 \le t \le T_b \\ 0 & \text{otherwise} \end{cases}$$

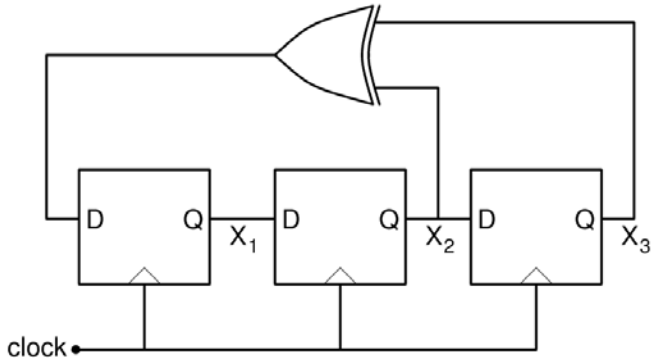
$$S_w(f) = \frac{|P(f)|^2}{4T_b}\left(1 + \frac{\delta(f)}{T_b}\right)$$

$$P(f) = AT_b \frac{\sin \pi f T_b}{\pi f T_b} e^{-j\pi f T_b}$$

$$S_w(f) = \frac{A^2 T_b}{4}\left(\frac{\sin \pi f T_b}{\pi f T_b}\right)^2\left(1 + \frac{\delta(f)}{T_b}\right)$$

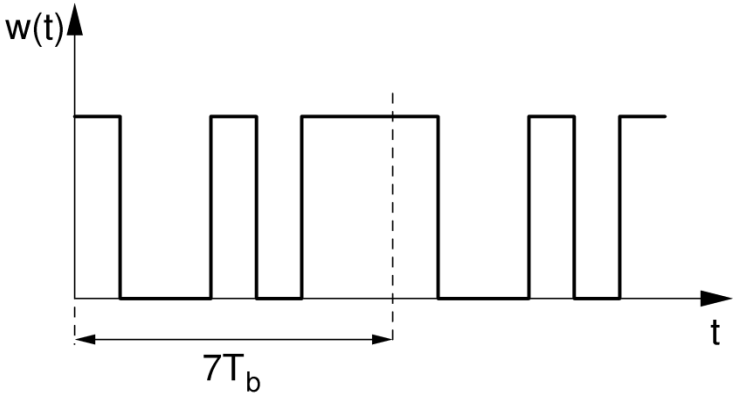
impuls at DC (DC offset)

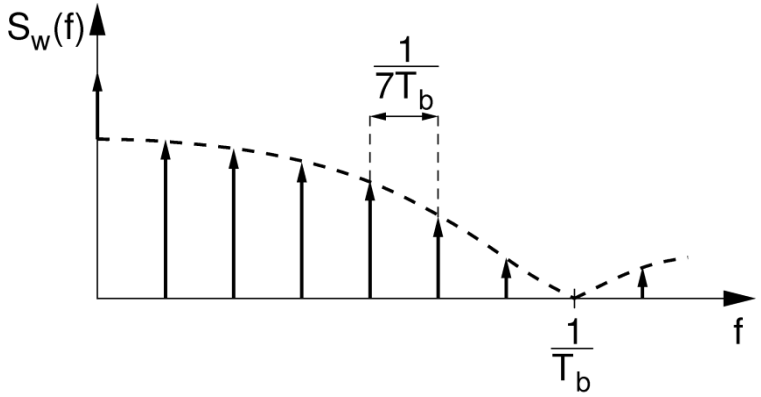
Pseudo-random digital data



state	X_1	X_2	X_3
1	1	1	1
2	0	1	1
3	0	0	1
4	1	0	0
5	0	1	0
6	1	0	1
7	1	1	0

return to state 1





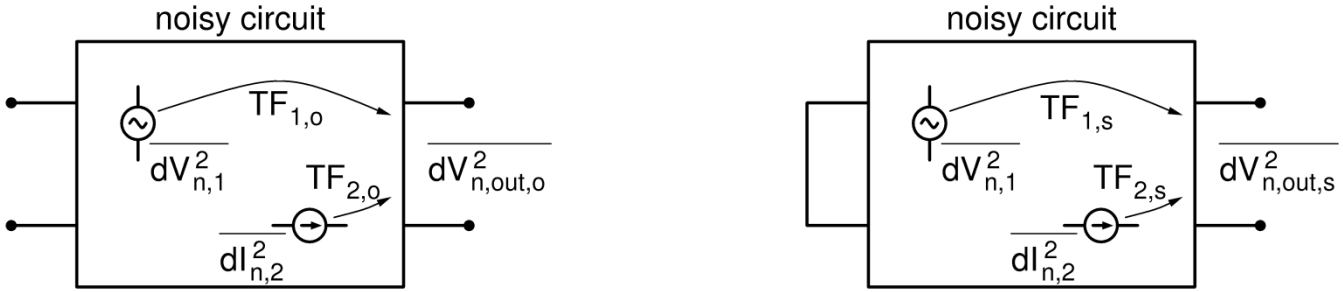
- same pattern is repeated; pattern length = 2^n-1 (n = # memory elements)
- PRBS used frequently to test high-speed data links (n typically between 7 and 31)
- continuous spectrum of truly random data becomes line spectrum because of the repetition of the same pattern every 2^n-1 bits
- distance between spectral lines inversely proportional to pattern length

Circuit noise (1)

- when working with noise, always use the **spectral noise densities**

$$\begin{cases} \overline{dV_{n,R}^2} = 4kTR \cdot df + V_R^2 \dfrac{K_R R_{sq}}{A_R} \cdot \dfrac{df}{f} & \text{for resistors} \\ \overline{dV_{n,T}^2} = 4kTR_{eff} \cdot df + \dfrac{K_T}{WLC_{ox}^2} \cdot \dfrac{df}{f} & \text{for MOS transistors} \\ \overline{dI_{n,D}^2} = 2qI_D \cdot df + I_D \dfrac{K_D}{A_D} \cdot \dfrac{df}{f} & \text{for diodes} \end{cases}$$

- every noise source in a circuit has its own transfer function towards the output
- output noise spectral density depends on the input impedance!

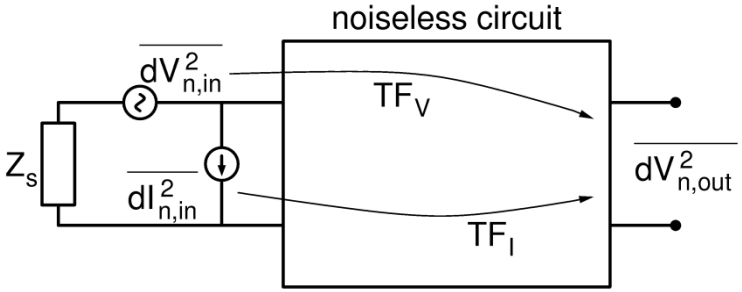


$$\overline{dV_{n,out,o}^2} = \overline{dV_{n,1}^2} \cdot \left|TF_{1,o}\right|^2 + \overline{dI_{n,2}^2} \cdot \left|TF_{2,o}\right|^2 + \ldots$$

$$\overline{dV_{n,out,s}^2} = \overline{dV_{n,1}^2} \cdot \left|TF_{1,s}\right|^2 + \overline{dI_{n,2}^2} \cdot \left|TF_{2,s}\right|^2 + \ldots$$

Circuit noise (2)

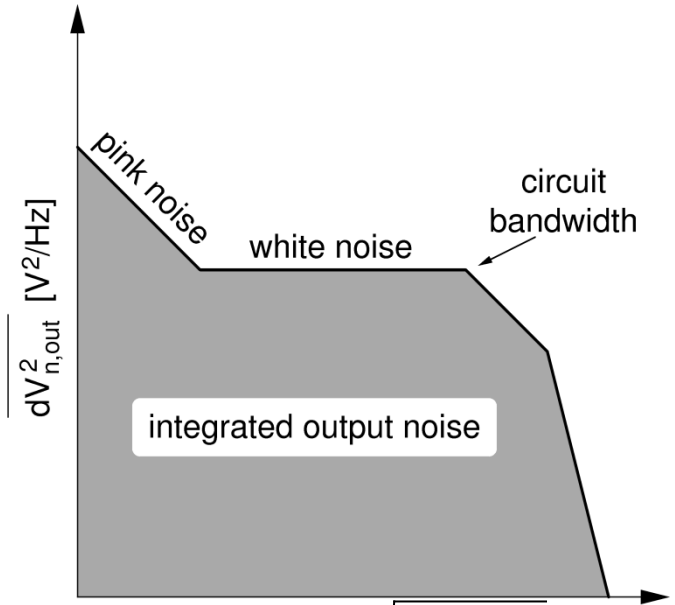
- equivalent input noise is a purely mathematical quantity (can go infinitely high if the gain of the circuit goes to zero!)
- both a voltage and a current noise source are required to characterize the circuit noise
- for an open input ($Z_s = \infty$), the output noise is only determined by the current source
- for a shorted input ($Z_s = 0$), the output noise in only determined by the voltage source
- for a finite Z_s, the output noise is determined by both the voltage and the current source



$$\overline{dV^2_{n,in}} = \frac{\overline{dV^2_{n,out,s}}}{|TF_V|^2}$$

$$\overline{dI^2_{n,in}} = \frac{\overline{dV^2_{n,out,o}}}{|TF_I|^2}$$

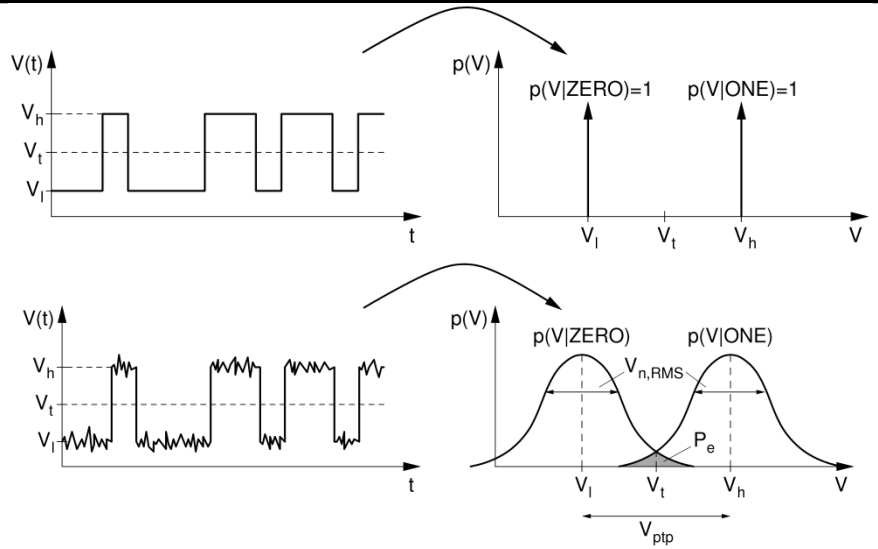
for a certain Z_s



$$V_{n,RMS} = \sqrt{\int_0^\infty \overline{dV^2_{n,out}}}$$

Circuit noise and BER

- without noise, there is no chance of making a wrong decision provided that the decision threshold is between the low and high signal values
- with noise, there is a chance of making an error which depends on the choice of the decision threshold



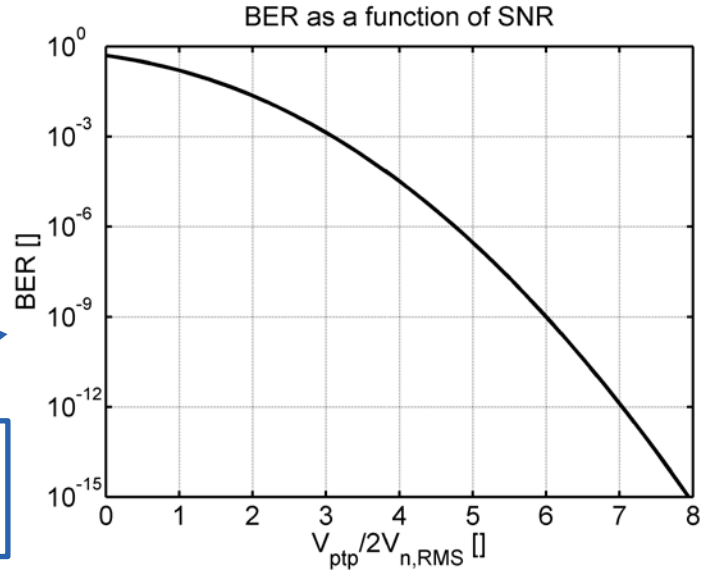
$$P_e = \frac{1}{2}P_{e,ZERO} + \frac{1}{2}P_{e,ONE}$$

$$= \frac{1}{2}\int_{V_t}^{\infty} p(V|ZERO)dV + \frac{1}{2}\int_{-\infty}^{V_t} p(V|ONE)dV$$

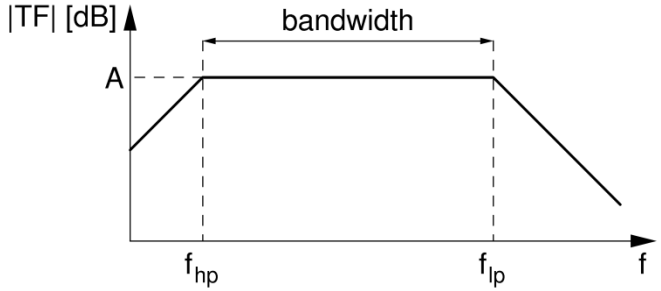
$$= \frac{1}{2}\int_{V_t}^{\infty} \frac{1}{2\pi V_{n,RMS}} e^{\frac{-(V-V_l)^2}{2V_{n,RMS}^2}} dV + \frac{1}{2}\int_{-\infty}^{V_t} \frac{1}{2\pi V_{n,RMS}} e^{\frac{-(V-V_h)^2}{2V_{n,RMS}^2}} dV$$

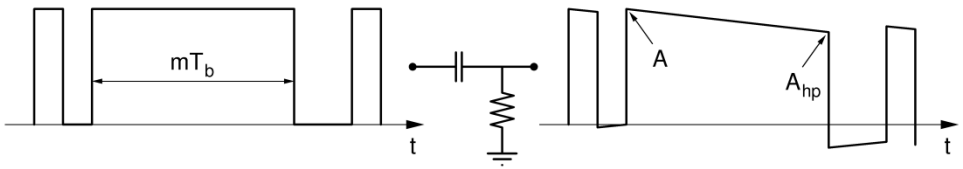
$$= Q\left(\frac{V_{ptp}}{2V_{n,RMS}}\right)$$

$V_{ptp} > 14\ V_{n,RMS}$ for BER < 10^{-12}
→ **SNR > 17 dB**

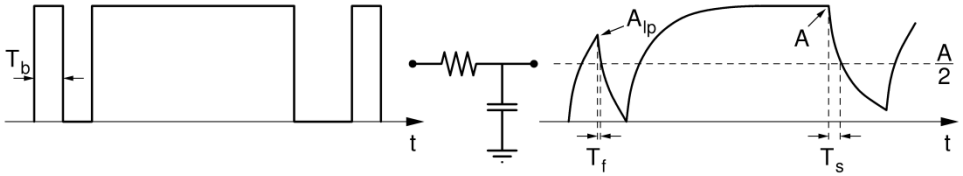


Bandwidth





$$\frac{A_{hp}}{A} = e^{\frac{-mT_b}{\tau_{hp}}}$$



$$\frac{A_{lp}}{A} = 1 - e^{\frac{-T_b}{\tau_{lp}}}$$

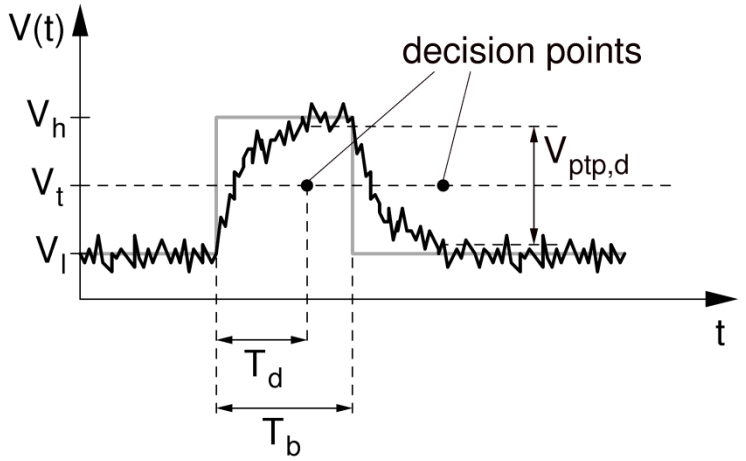
high-pass filtering

- originates from AC-coupling, offset compensation, …
- leads to DC wander with a long succession of identical bits
- scrambling data to reduce the low frequency content

low-pass filtering

- originates from limited performance of circuits at high frequencies
- leads to incomplete settling of the signal within the bit interval = ISI
- leads to jitter

Bandwidth, circuit noise and BER



finite receiver BW → signal not settled at the decision point → lower effective SNR

$$V_{ptp,d} = V_{ptp} \cdot \left(1 - 2e^{\frac{-T_d}{\tau_{lp}}}\right)$$

$$P_e = Q\left(\frac{V_{ptp,d}}{2V_{n,RMS}}\right)$$

Example:

bit rate = 1 Gbit/s (T_b = 1 ns) and f_{lp} = 500 MHz (τ_{lp} = 318 ps)

→ $V_{ptp,d}$ = 0.585 V_{ptp} for a decision point in the middle of the bit interval

→ BER ≈ 1e-5 if V_{ptp} = 14 $V_{n,RMS}$ (compared to 1e-12!)

Question:

Is it a good idea to postpone the decision point towards the end of the bit interval to profit from the better settling?

Bandwidth-noise trade-off

GOAL: minimize BER of a receiver

$$\frac{A_{lp}}{A} = 1 - e^{\frac{-T_b}{\tau_{lp}}}$$

BW perspective

minimize signal distortion

↓

minimize rise and fall times

↓

maximize bandwidth

noise perspective

$$V_{n,RMS} = \sqrt{\int_0^{\infty} dV_{n,out}^2}$$

minimize circuit noise

↓

minimize integrated noise

↓

minimize bandwidth

Sweet spot: BW is around 70 % of the bit rate (for example 700 MHz for 1 Gbit/s data)

Jitter

voltage axis

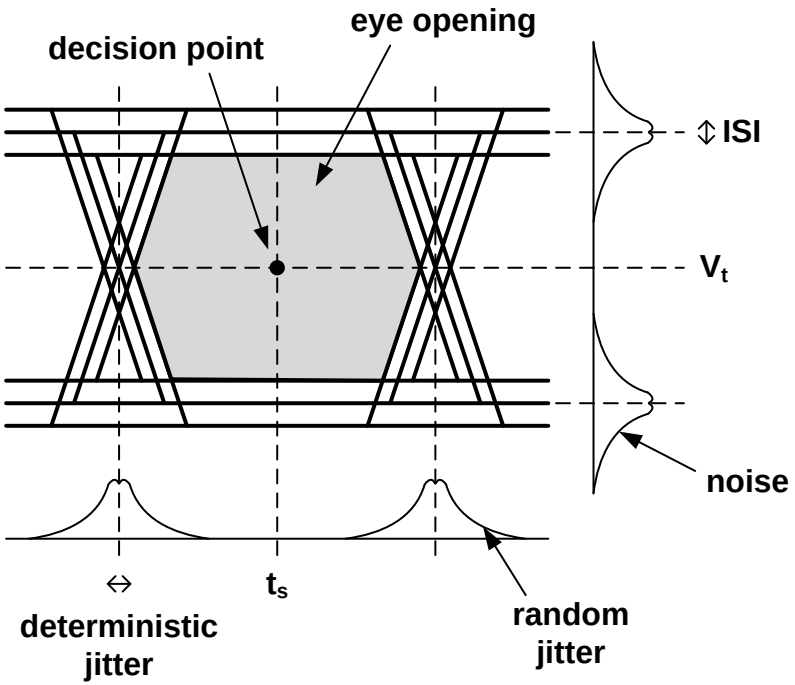
- noise leads to a finite BER
- BW limitation leads to ISI

time axis

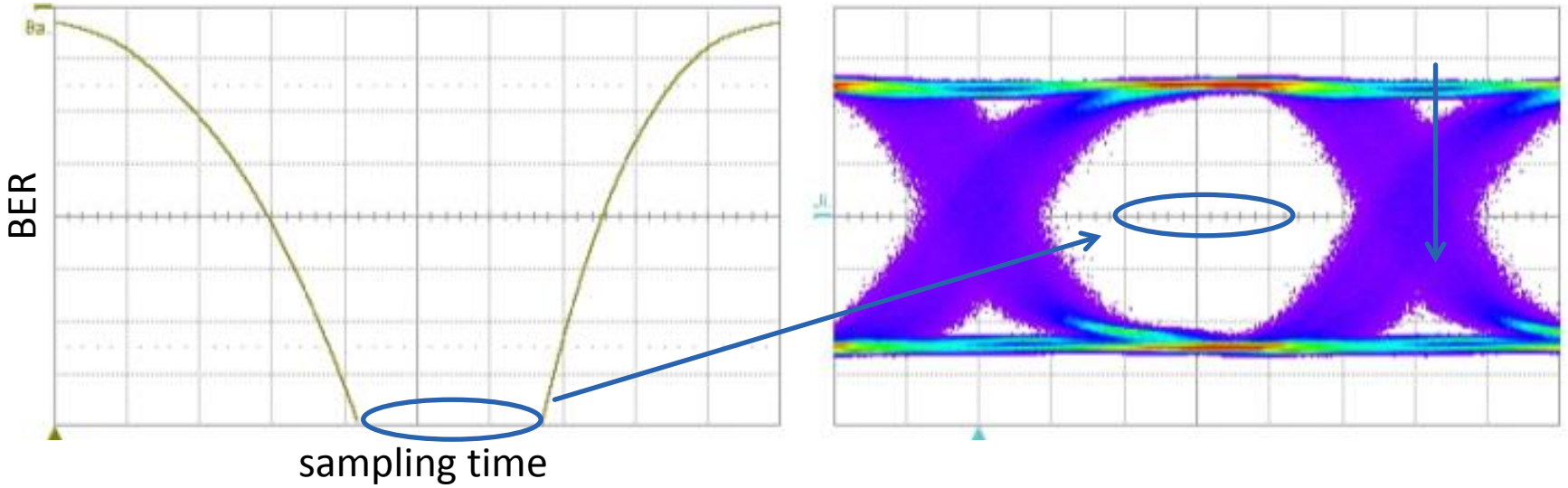
- noise leads to random jitter
- BW limitation leads to deterministic jitter

BER is determined by the sampling time as well as the threshold voltage

ideal decision point is probably in the exact center of the eye opening



Jitter and BER



optimum sampling time range is usually very limited for the minimal BER

↓

tight control of sampling time is critical

↓

difficult due to static clock phase errors, jitter in the recovered clock, …

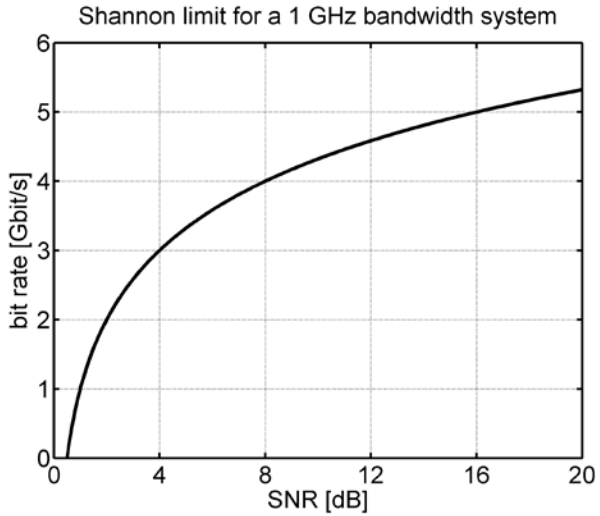
↓

ideal CDR would ‘track’ the jitter of the input data to always sample in the center of the eye

Forward error correction

- for a BER of 1e-12, an SNR of at least 17 dB is needed, assuming no bandwidth limitation
- **Shannon's channel capacity theorem** states that:
  error-free transmission over a channel with additive white Gaussian noise is possible if

$$\text{bit rate} \leq BW \cdot \left(1 + \log_2 SNR\right)$$

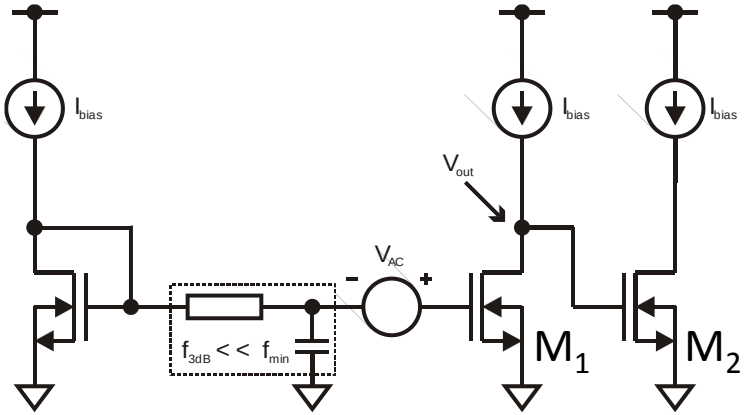


- with 17 dB SNR, error-free operation at 5 Gbit/s for a 1 GHz bandwidth!
- how? → forward error correction coding (FEC)
  FEC = add redundancy bits in the TX to correct for transmission errors
  examples: add parity bits, Reed-Solomon code, …

Outline

- *Basic concepts*
  - *Data properties*
  - *Circuit noise*
  - *Bandwidth*
  - *BER*
  - *Jitter*
  - *FEC*
- Circuit concepts
  - Common-source amplifier
  - Optimum current density
  - Bandwidth extension
  - Multistage amplifier
  - Negative impedances

Common-source amplifier



$$A_{cs} = \frac{-g_{m,M_1} r_{ds,M_1}\left(1 - \frac{s \cdot C_{gd,M_1}}{g_{m,M_1}}\right)}{1 + s \cdot r_{ds,M_1} C_l}$$

$$\left|A_{DC}\right| = g_{m,M_1} r_{ds,M_1}$$

$$f_p = \frac{1}{2\pi \cdot r_{ds,M_1} C_l}$$

$$f_z = \frac{g_{m,M_1}}{2\pi \cdot C_{gd,M_1}}$$

$$GBW = \frac{g_{m,M_1}}{2\pi \cdot C_l}$$

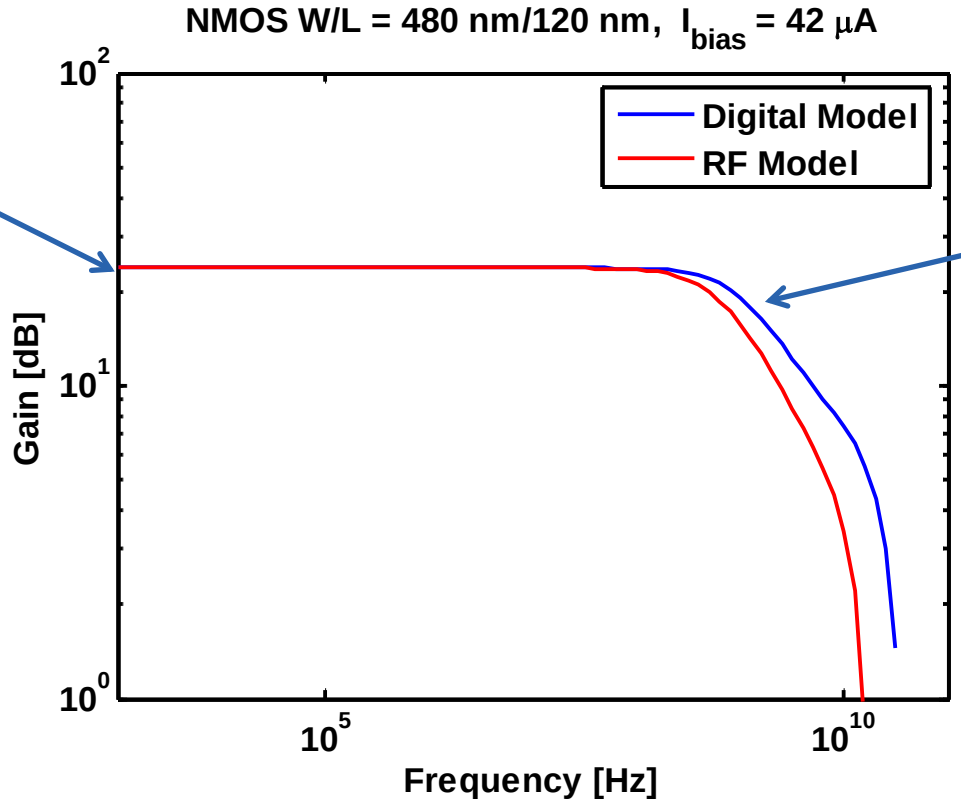
$$C_l = C_{db,M_1} + C_{gd,M_1} + C_{gs,M_2} + (1 + A_{cs}) \cdot C_{gd,M_2}$$

- Miller effect on C_{gd,M_2}
- frequency dependent C_l

- gain-bandwidth (GBW) trade-off through r_{ds,M_1}
- negative pole, reduces phase with 90°
- positive zero, increases phase with 90°
- zero at very high frequency!

Frequency response of CS amplifier

$$A_{DC} = -g_{m,M_1} r_{ds,M_1}$$



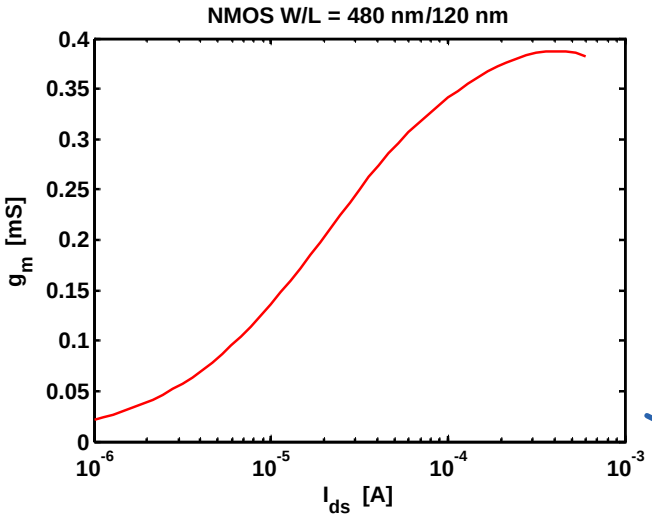
$$f_p = \frac{1}{2\pi \cdot r_{ds,M_1} C_l}$$

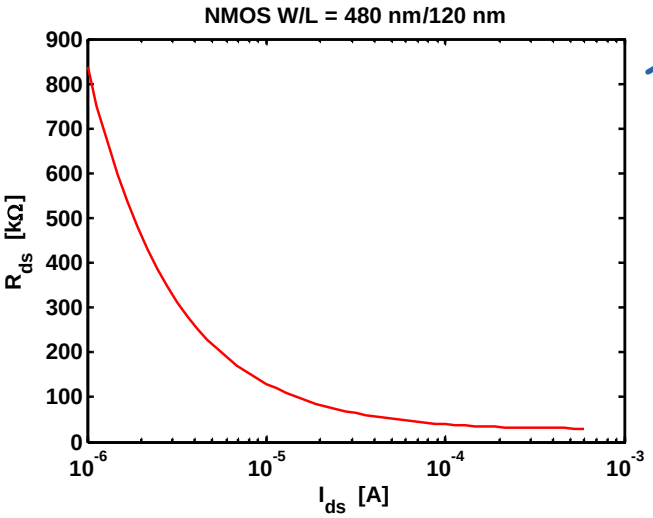
bandwidth depends strongly on parasitics!

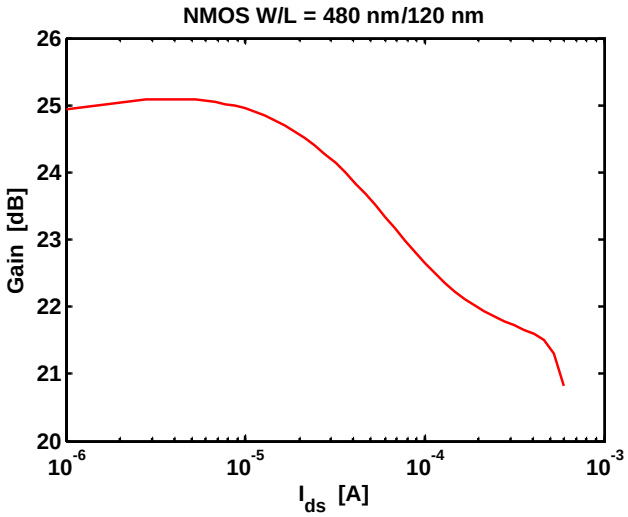
→ RF transistor model covers some of the layout parasitics

→ extracted schematics include even more of them

Transistor biasing for gain







$$r_{ds} \approx \frac{1}{\lambda I_{ds}}$$

high gain?

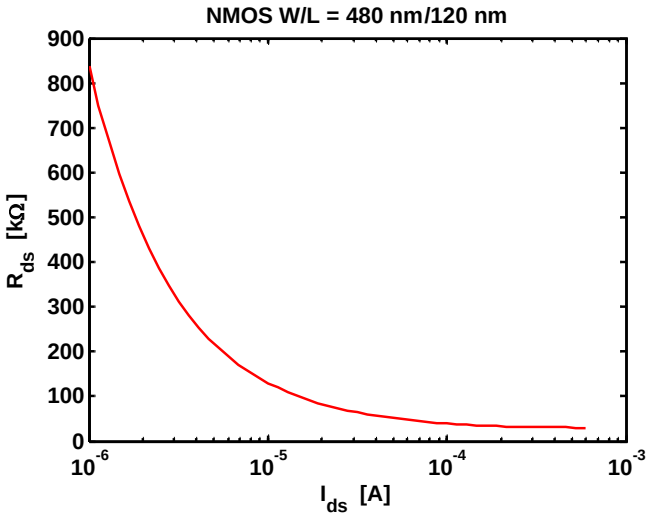
→ low current density

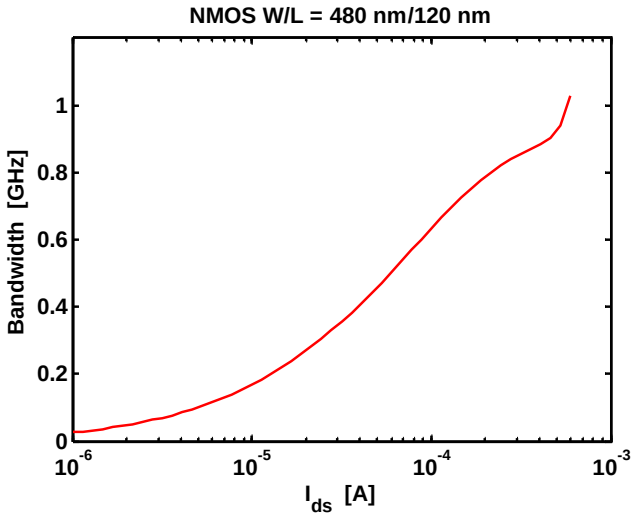
= small V_{gs}-V_{th} (weak inversion)

→ long transistor $\lambda \propto \frac{1}{L}$

Transistor biasing for bandwidth

$$C_l = C_{db,M_1} + C_{gd,M_1} + (1 + A_{cs}) \cdot C_{gd,M_2}$$ → decreases slightly with current density



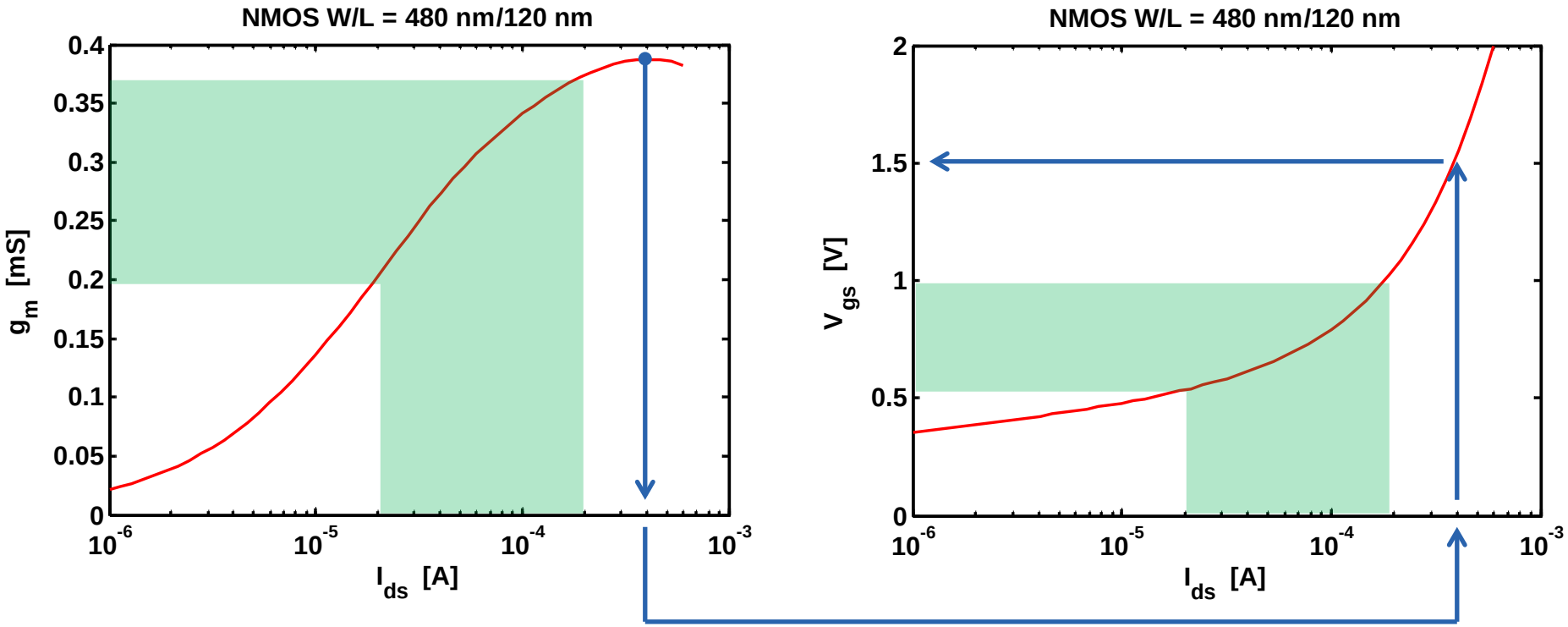


	high gain	high speed
V_{gs}-V_{th}	low	high
L	large	small

high bandwidth?
→ high current density
→ large V_{gs}-V_{th} (strong inversion)
→ short transistor

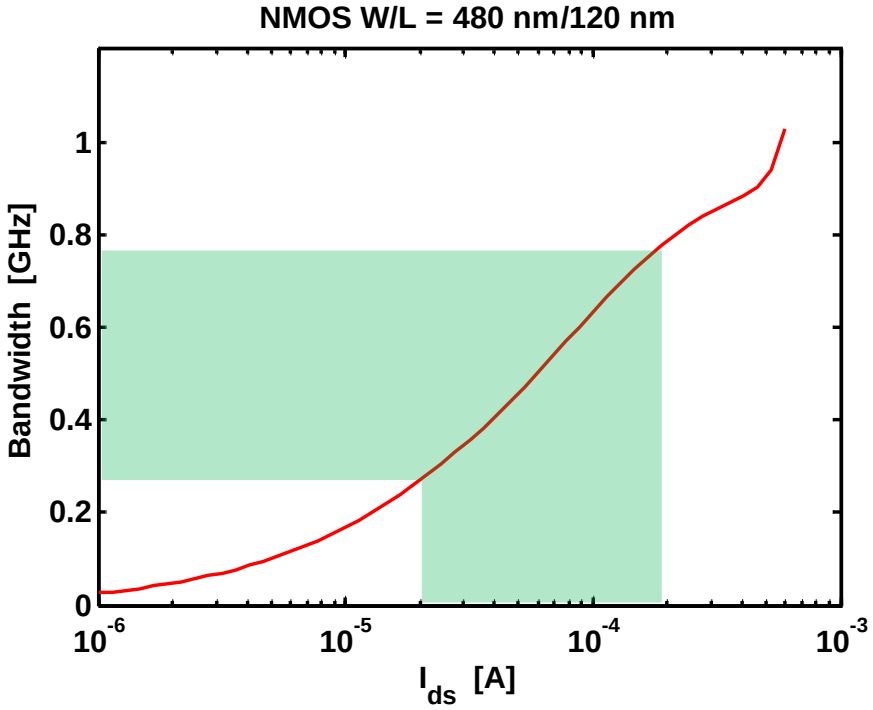
$$\lambda \propto \frac{1}{L}$$

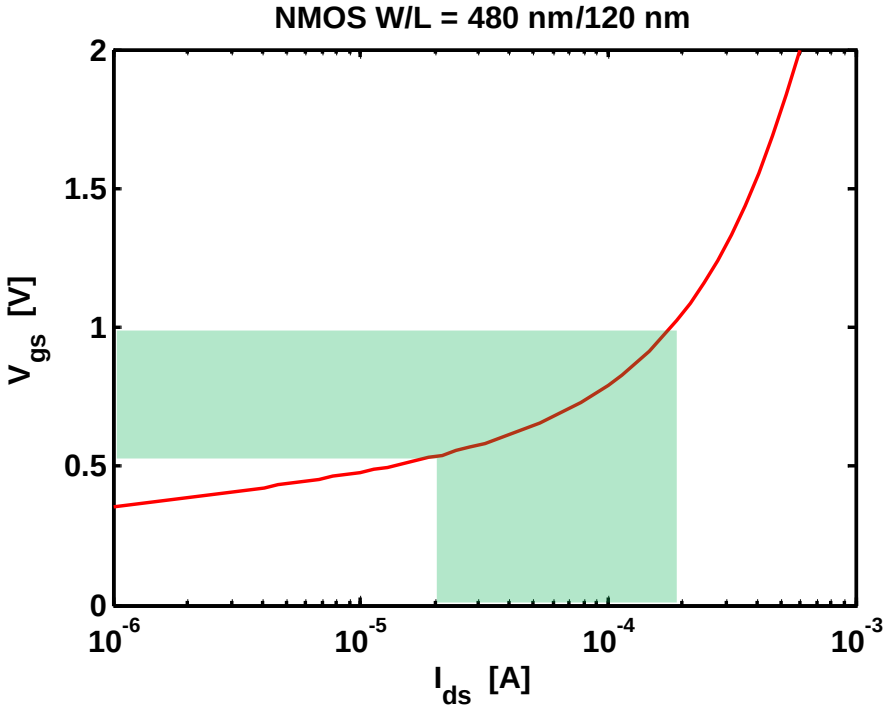
Biasing for maximum g_m?



- high-speed circuits? → maximize g_m
- maximum g_m is obtained for a large current density and large overdrive voltage
- no signal headroom for the maximum g_m (V_{dd} = 1.5 V)
- in practice: 0.5 V < V_{gs} < 1 V → 70-80 % of maximum g_m

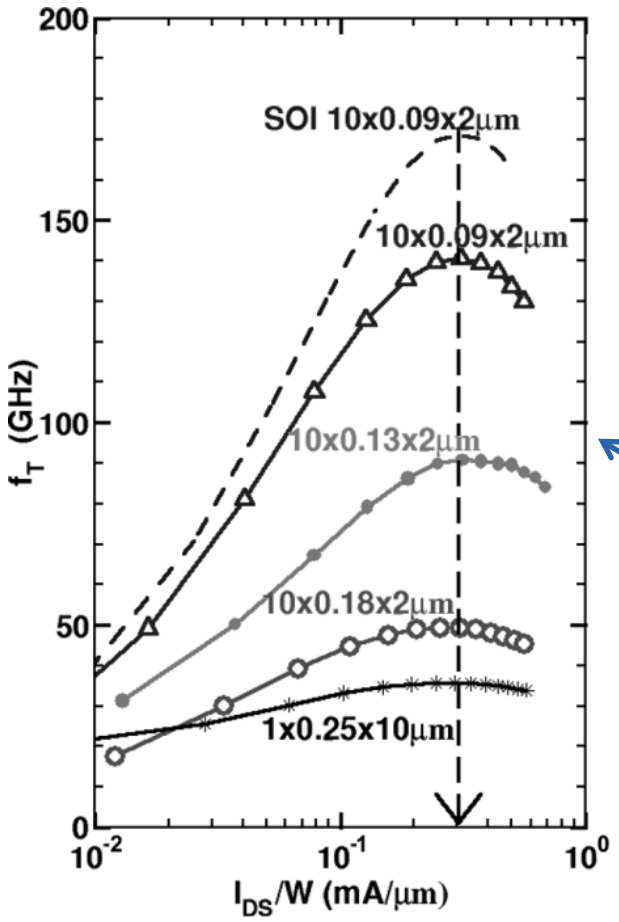
Biasing for maximum bandwidth?





- maximum bandwidth is obtained for a large current density and large overdrive voltage
- no signal headroom for the maximum bandwidth (V_{dd} = 1.5 V)
- in practice: 0.5 V < V_{gs} < 1 V → 25-75 % of maximum bandwidth

Invariance of optimum current density (1)



(T. O. Dickson et al. - JSSC vol. 41, no. 8, p. 1830, August 2006)

- intrinsic speed of transistor → f_t

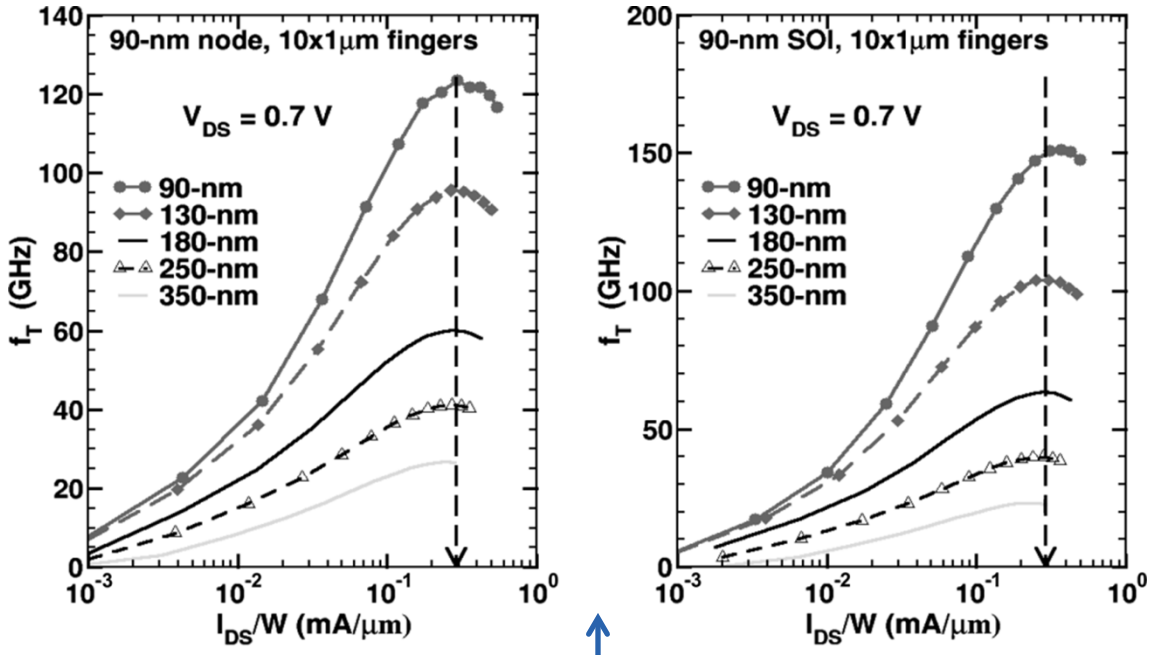
$$f_t = \frac{g_m}{2\pi \cdot C_{gs}}$$

- maximum f_t can be found at a current density of J = ± 0.3 mA/µm

$$J = \frac{I_{ds}}{W}[\mathrm{mA}/\mu\mathrm{m}]$$

- relatively broad optimum
- **independent on foundry**
- independent on technology node (due to constant field scaling)
- independent on transistor length
- valid for nMOS and pMOS transistors (for pMOS, the optimum is 0.15 mA/µm)

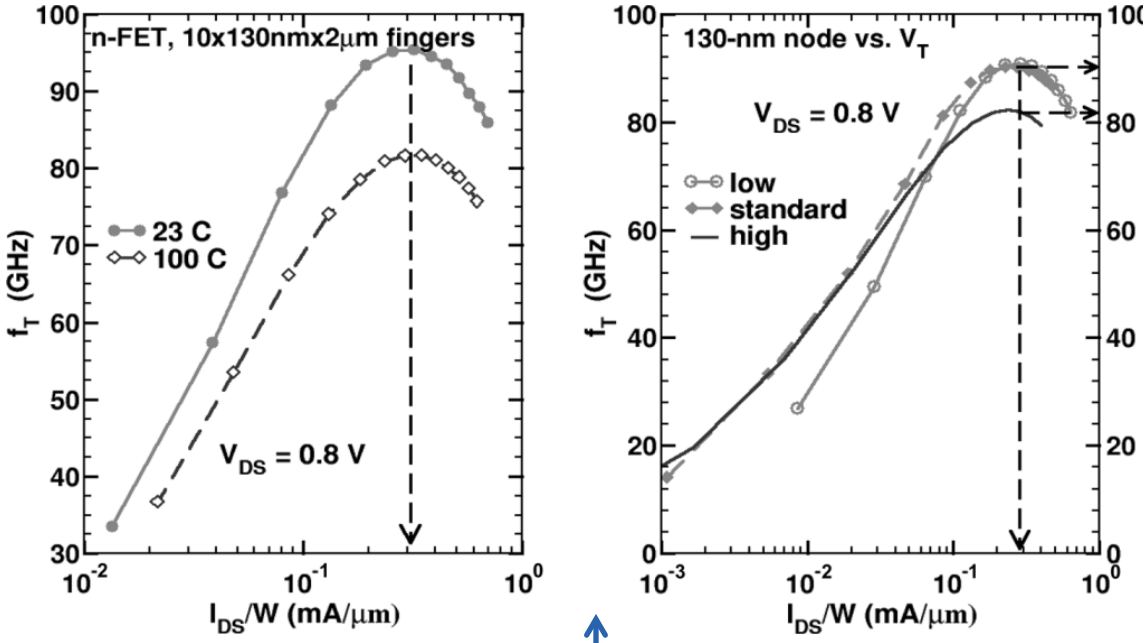
Invariance of optimum current density (2)



optimum current density

- independent on transistor length
- identical for bulk or SOI processes

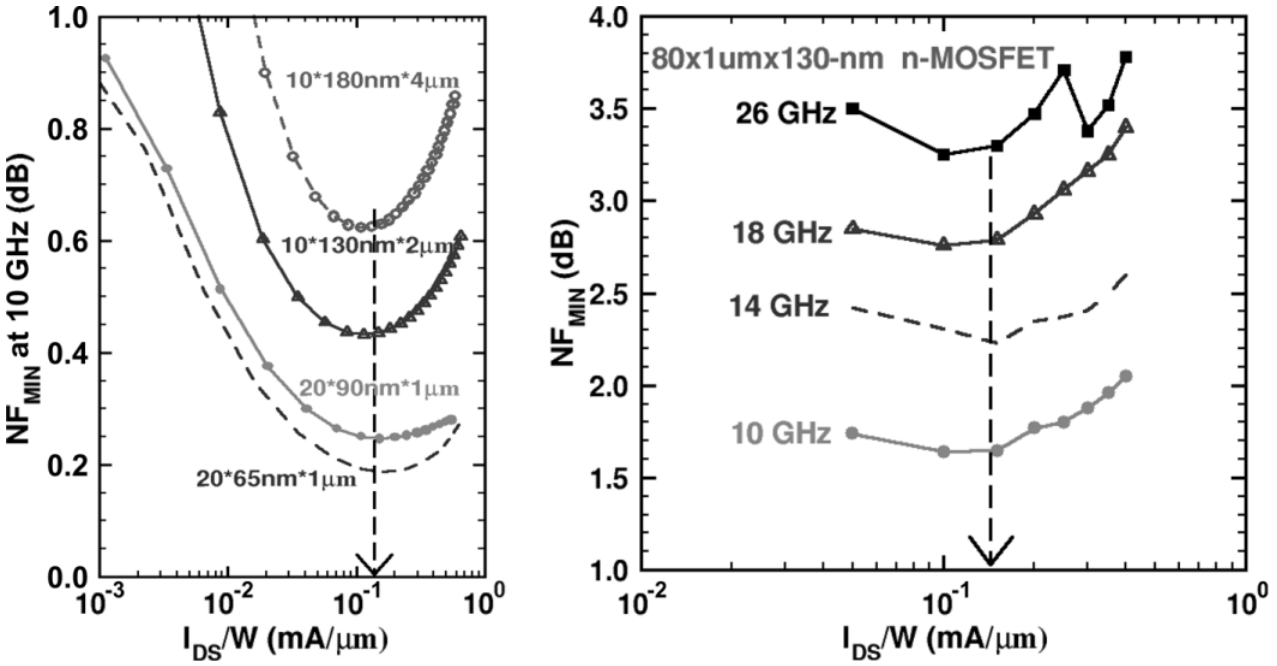
Invariance of optimum current density (3)



optimum current density

- independent on temperature
- Independent on threshold voltage

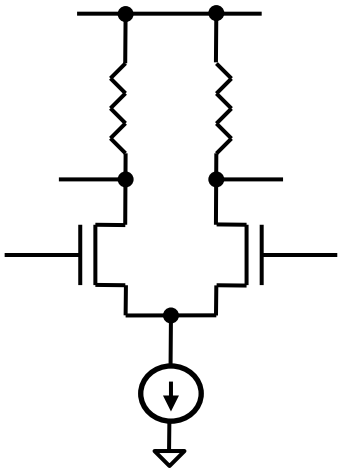
Invariance of optimum current density (4)

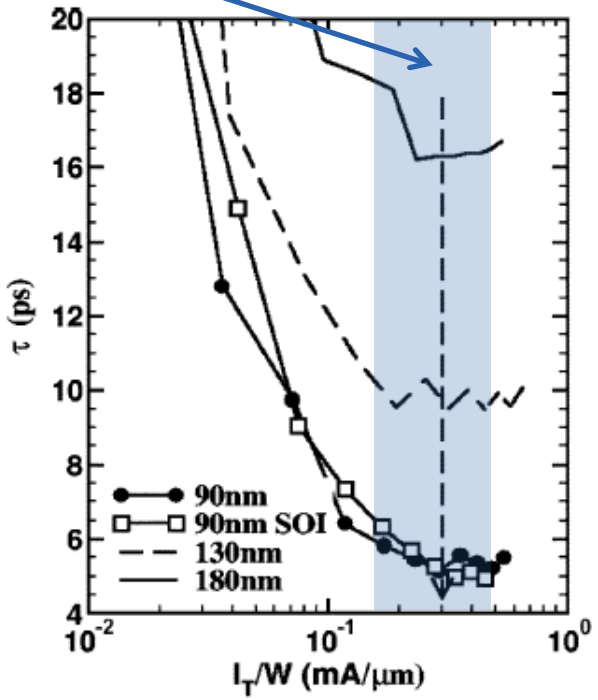


a comparable optimum can be found when considering noise (J = ± 0.15 mA/µm)
→ wider transistors and smaller $V_{gs} - V_{th}$ if noise is more important than bandwidth

CML stage and optimum current density

delay changes less than 10 % for a current density in the range 0.15 mA/µm – 0.5 mA/µm

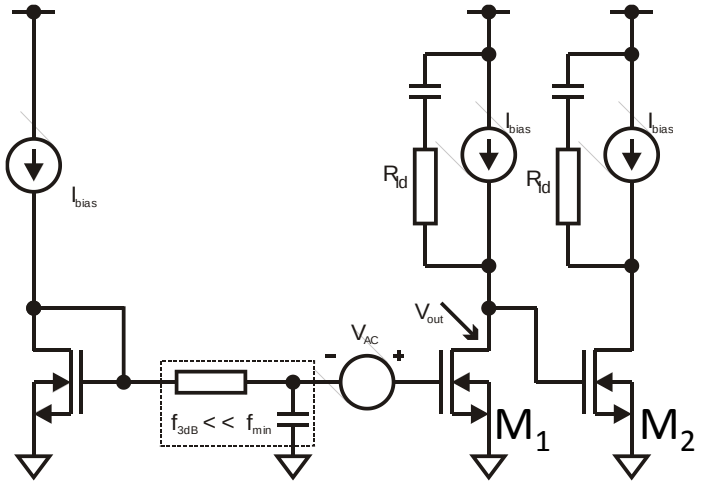


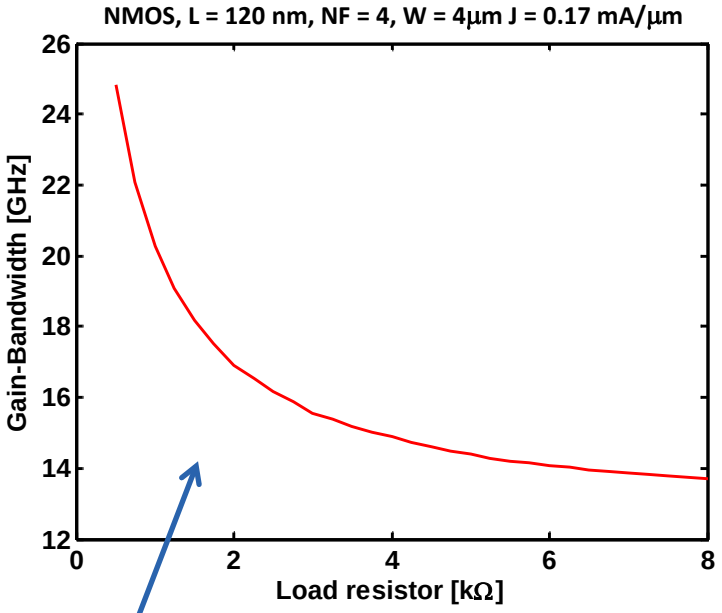


$$\tau = \frac{\Delta V}{I_T}\left[C_{gd} + C_{db} + \left(k + \frac{R_g}{R_L}\right)\left[C_{gs} + (1 + g_m R_L) C_{gd}\right]\right]$$

→ time constant at the output of the CML stage when fully switching the tail current

Common-source amplifier with resistive load





$$\left|A_{DC}\right| = g_{m,M_1}\left(r_{ds,M_1} \,//\, R_{ld}\right)$$

$$f_p = \frac{1}{2\pi \cdot \left(r_{ds,M_1} \,//\, R_{ld}\right) \cdot C_l}$$

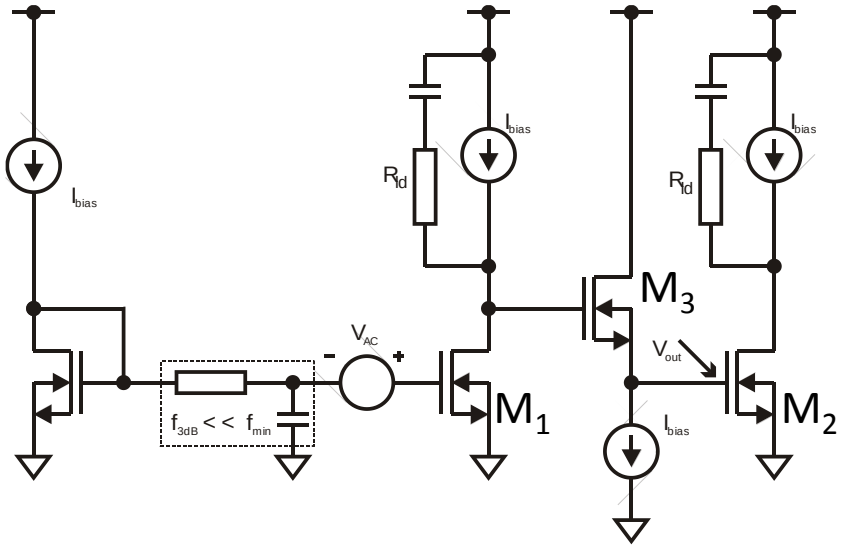
$$GBW = \frac{g_{m,M_1}}{2\pi \cdot C_l}$$

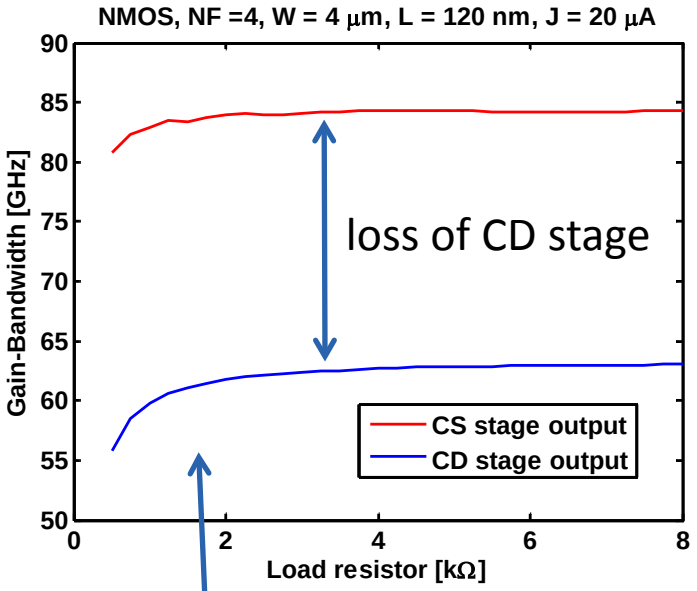
same as for simple CS amplifier?

$$C_l = C_{db,M_1} + C_{gd,M_1} + C_{gs,M_2} + \left(1 + A_{cs}\right) \cdot C_{gd,M_2}$$

No! lower DC gain
→ less pronounced Miller effect
→ smaller C_l
→ bandwidth increases faster then that the gain is reduced

Common-source amplifier with common-drain stage





$$|A_{DC}| = g_{m,M_1}\left(r_{ds,M_1} // R_{ld}\right).\frac{g_{m,M_3}}{g_{m,M_3}+g_{ds,M_3}}$$

$$\approx g_{m,M_1}\left(r_{ds,M_1} // R_{ld}\right) \rightarrow GBW = \frac{g_{m,M_1}}{2\pi \cdot C_l}$$

same as before?

$$f_p = \frac{1}{2\pi \cdot \left(r_{ds,M_1} // R_{ld}\right) \cdot C_l}$$

→ assume pole of CD stage is very large

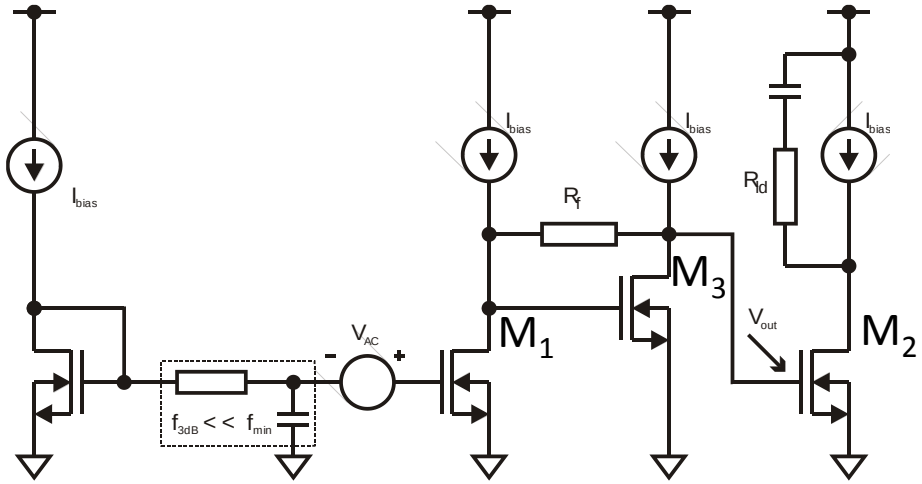
GBW does not depend on DC gain anymore

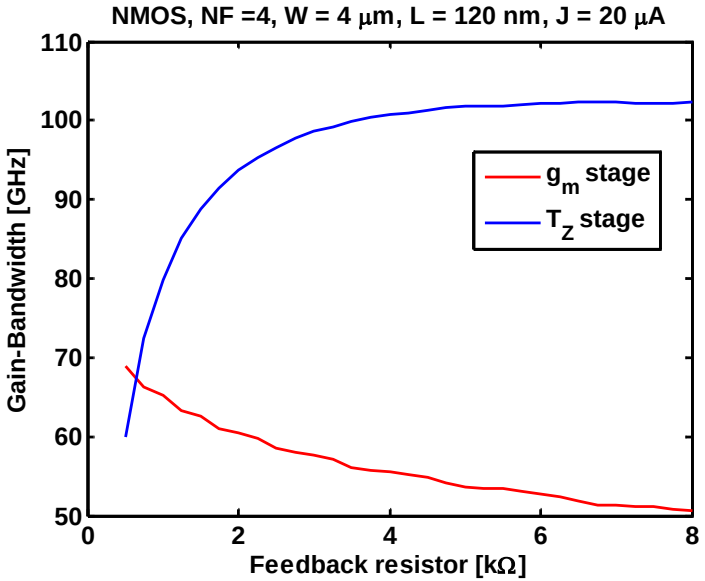
$$C_l = C_{db,M_1} + C_{gd,M_1} + C_{gd,M_3}$$

No! no parasitics from M_2

Cherry-Hooper amplifier

principle: no high impedance nodes
→ alternate transconductance and transimpedance stages





$$|A_{DC}| = g_{m,M_1}\left(r_{ds,M_1} // R_f\right)$$

$$f_{p1} = \frac{1}{2\pi \cdot \left(r_{ds,M_1} // \frac{R_f}{A_{cs}}\right) \cdot C_{l1}} \approx \frac{A_{cs}}{2\pi \cdot R_f C_{l1}}$$

$$f_{p2} \approx \frac{g_{m,M_3}}{2\pi \cdot C_{l2}}$$

2 low-impedance poles → peaking?

Inductive peaking

problem: output capacitance reduces output impedance at higher frequencies

→ lower output voltage since output current sees lower impedance

→ pole at relatively low frequency

solution: add an inductor in series with the load resistor = shunt peaking

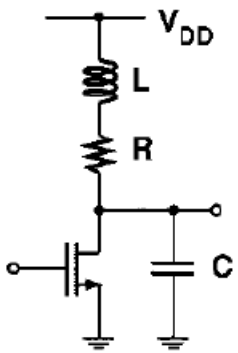
→ impedance of inductor increases with frequency

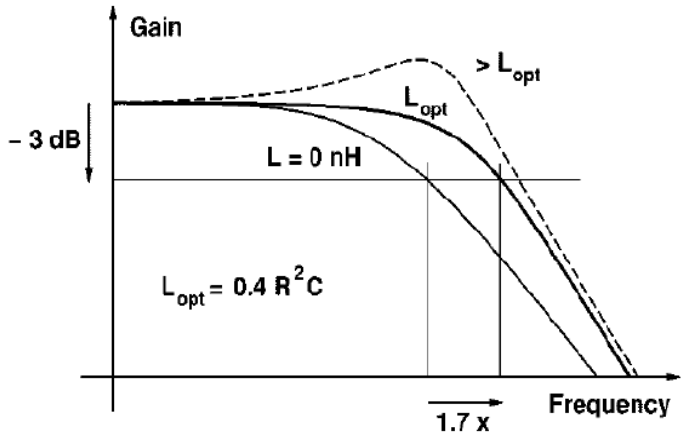
→ this increased inductor impedance can balance the reduced capacitance impedance

→ voltage gain can be maintained over wider frequency range

→ pole at a higher frequency

$$L = mR^2C$$





m	BW increase	response
0	1	no peaking
0.32	1.6	optimum group delay
0.41	1.72	maximally flat
0.71	1.85	maximum bandwidth

1st order system → 2nd order system → peaking

Active inductive peaking

problem: on-chip inductors require a lot of area

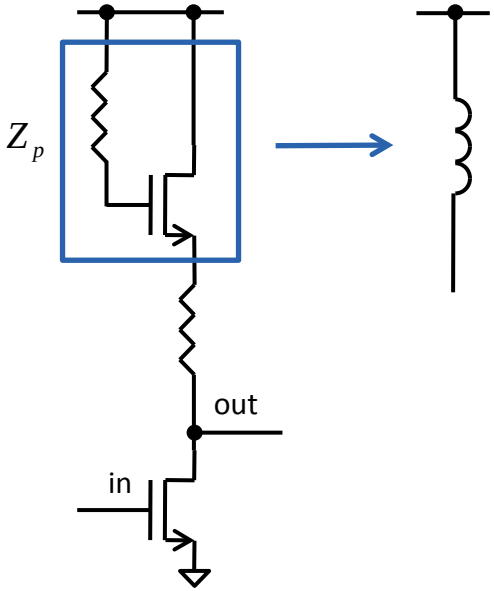
solution: emulate an inductor by means of a resistor, a capacitor and a transistor

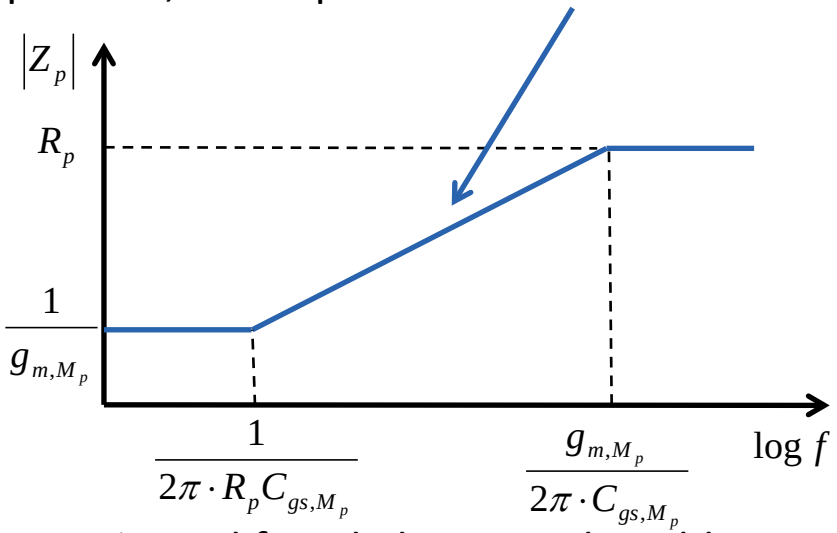
→ at low frequencies, the g_m generates a low impedance

→ at very high frequencies, gate and source of the transistor are shorted through the gate-source capacitance

→ at very high frequencies, the impedance is high as it is only determined by R_p (and by the output resistance of the transistor)

→ at intermediate frequencies, the impedance is inductive!

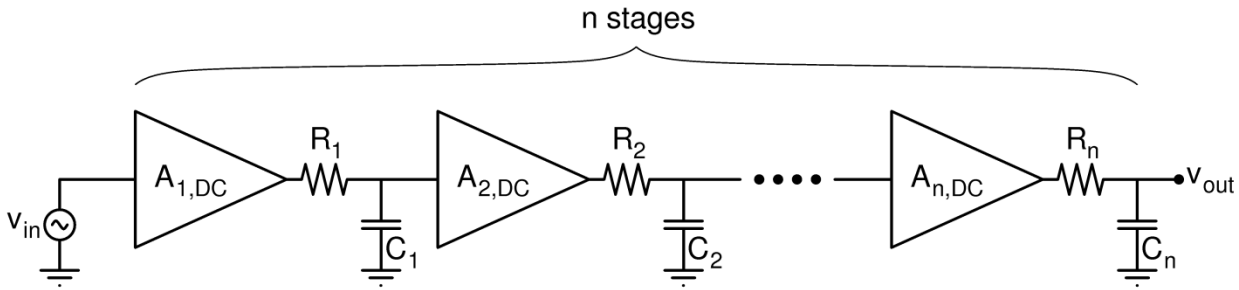




→ modify pole locations by adding a capacitor in parallel with C_{gs}

Multistage amplifier (1)

sometimes, gain **AND** bandwidth are required → multistage amplifier

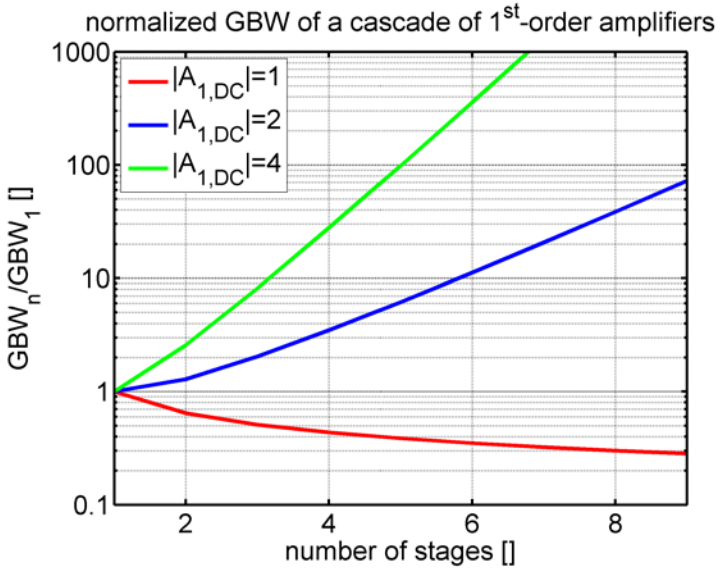


$$\left|A_{n,DC}\right| = \left|A_{1,DC}\right|^n$$

$$BW_n = \sqrt{\sqrt[n]{2}-1}\cdot BW_1$$

$$GBW_n = \left|A_{1,DC}\right|^{n-1}\cdot\sqrt{\sqrt[n]{2}-1}\cdot GBW_1$$

- GBW increases faster for higher single-stage gains
- power consumption increased by the number of stages

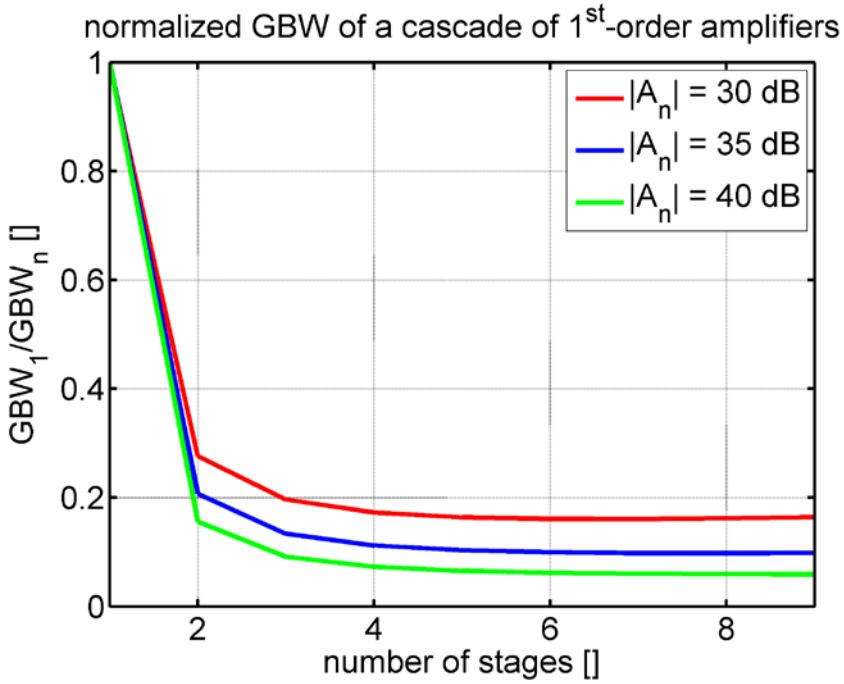


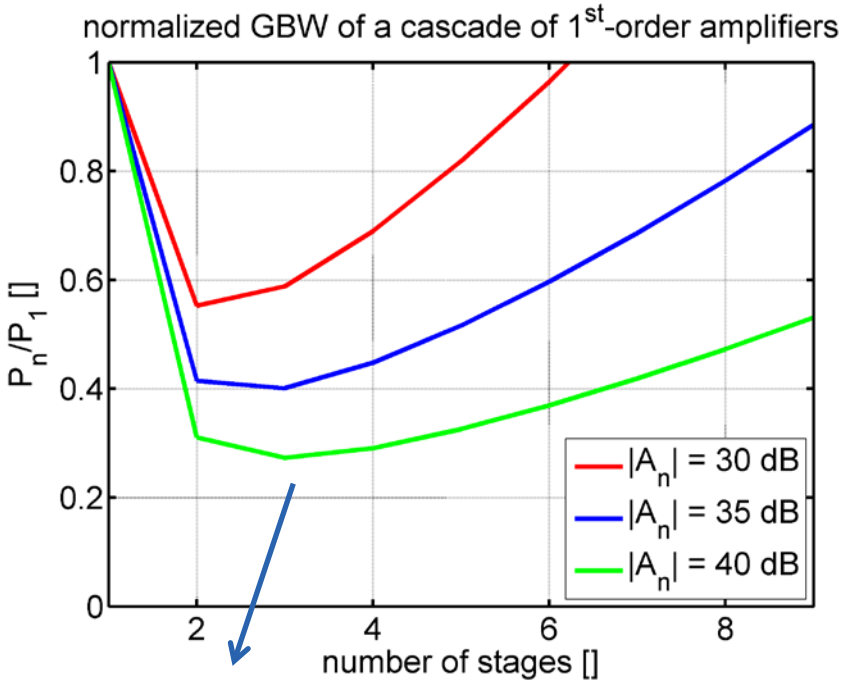
Multistage amplifier (2)

What if we want to design a multistage amplifier with a certain gain and bandwidth?
How many stages do we need?

$$\frac{GBW_1}{GBW_n} = \frac{A_n^{\frac{1}{n}-1}}{\sqrt{\sqrt[n]{2}-1}}$$

$$\frac{P_n}{P_1} = n\frac{A_n^{\frac{1}{n}-1}}{\sqrt{\sqrt[n]{2}-1}}$$





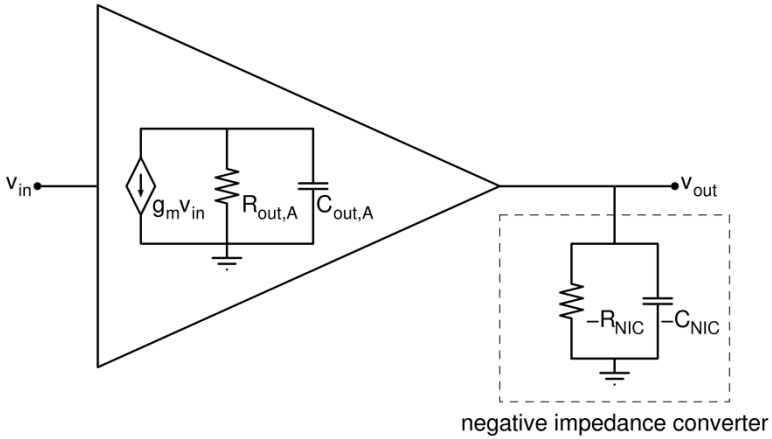
assuming GBW_n can be realized in a single stage!

$$P \propto g_m \propto GBW$$

Negative impedances?

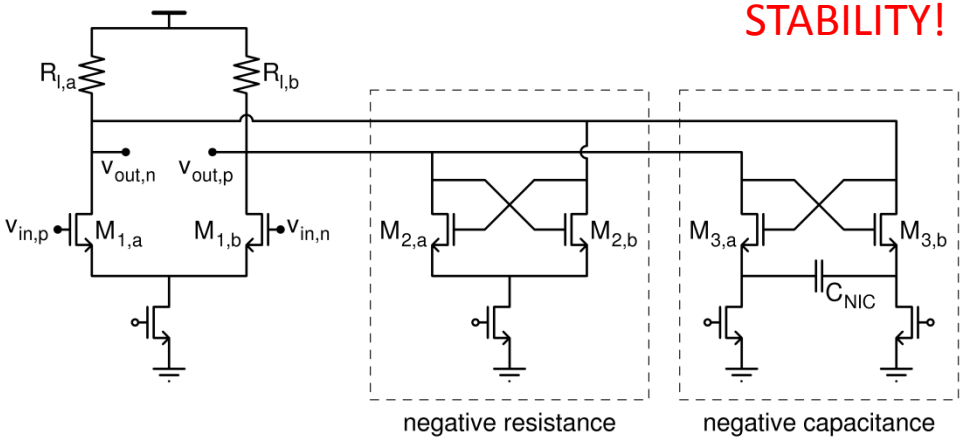
Can we increase gain and bandwidth simultaneously?

- negative resistance to increase the overall output resistance
  → gain is increased
  → bandwidth is reduced
- negative capacitance to decrease the overall output capacitance
  → bandwidth is increased



negative impedance converter

STABILITY!



negative resistance

negative capacitance

$$R_{out} = \frac{R_{NIC} R_{out,A}}{R_{NIC} - R_{out,A}}$$

$$C_{out} = C_{out,A} - C_{NIC}$$

$$Z_{NIC,res} = -\frac{2}{g_{m,M_2}}$$

$$Z_{NIC,cap} = -\frac{1}{s \cdot C_{NIC}} \left(1 + \frac{s \cdot 2C_{NIC}}{g_{m,M_3}} \right)$$

Acknowledgements

- Paulo Moreira