HIGH DATA RATE LINKS FOR HIGH ENERGY PHYSICS

Paulo Moreira
CERN, Switzerland

ECOLE IN2P3 DE MICROÉLECTRONIQUE
23 – 27 JUNE 2013,
PORQUEROLLES FRANCE

Credits

This slides were prepared with basis on the work of several people. Among them the most important contributions are from:

- Ozgur Cobanoglu – ISKO, Turkey
- Rui Francisco – CERN, Switzerland
- Ping Gui – SMU, USA
- Gianni Mazza – INFN, Italy
- Mohsine Menouni – CPPM, France
- Lukas Perktold – CERN, Switzerland
- David Porret – CERN, Switzerland
- Filip Tavernier – CERN, Switzerland

Outline

- A quick “flight” over digital communications
- Multi-Gb/s CMOS Design
- Power Dissipation In CMOS Circuits
- Optoelectronics Data Transmission Links
- All that you don’t (should) care about:
  - ESD Protection Circuits
    - Introduction
    - Input Matching using Coupled T-Coils
    - Input Matching using Inductive Compensation
  - Connectors, PCB & PACKAGE
- Laser Driver Design
- PIN – Receiver Design
- Assembling ASICS
- “Analogue-Verilog”
- Handling SEUs

Paulo.Moreira@cern.ch

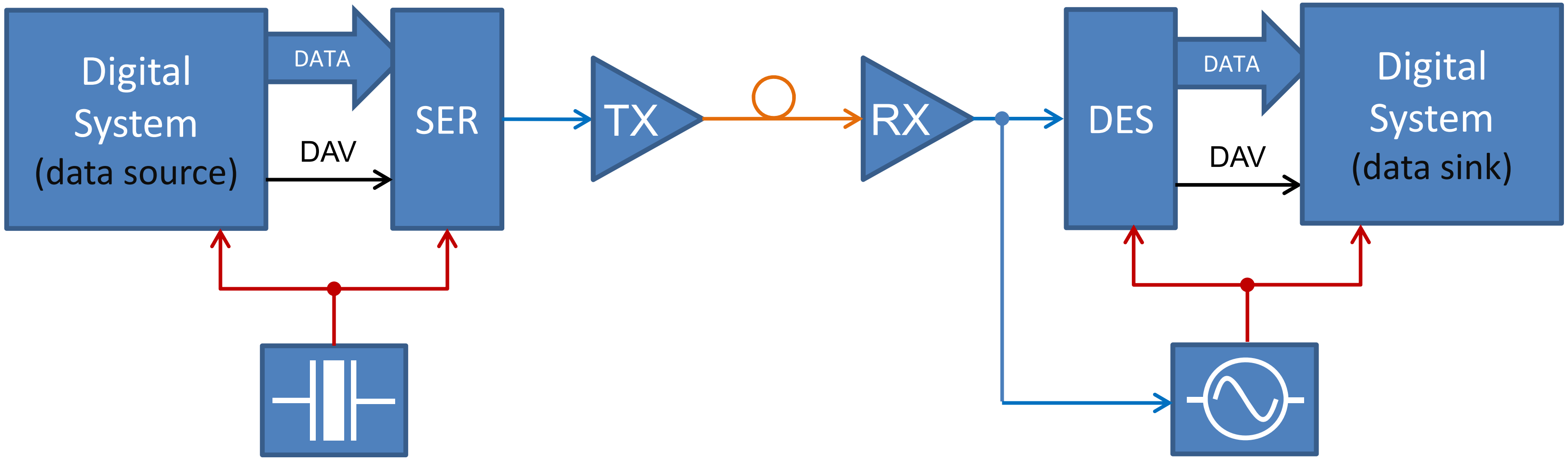
A QUICK “FLIGHT” OVER DIGITAL COMMUNICATIONS

Digital Communications

- Digital communication systems are widely used for data acquisition, trigger and control links in HEP:
  - A few exceptions: e.g. the CMS tracker data links
- Advantages:
  - Digital communications can be done virtually error free
  - When appropriate codes are used, transmission errors can be:
    - Detected
    - Corrected
  - They are "natural" for digital systems
  - Easy handling of multiple data sources and destinations
  - Easy re-routing between multiple sources and destination

Paulo.Moreira@cern.ch

Typical Architecture



Digital System (data source)

DATA

DAV

SER

TX

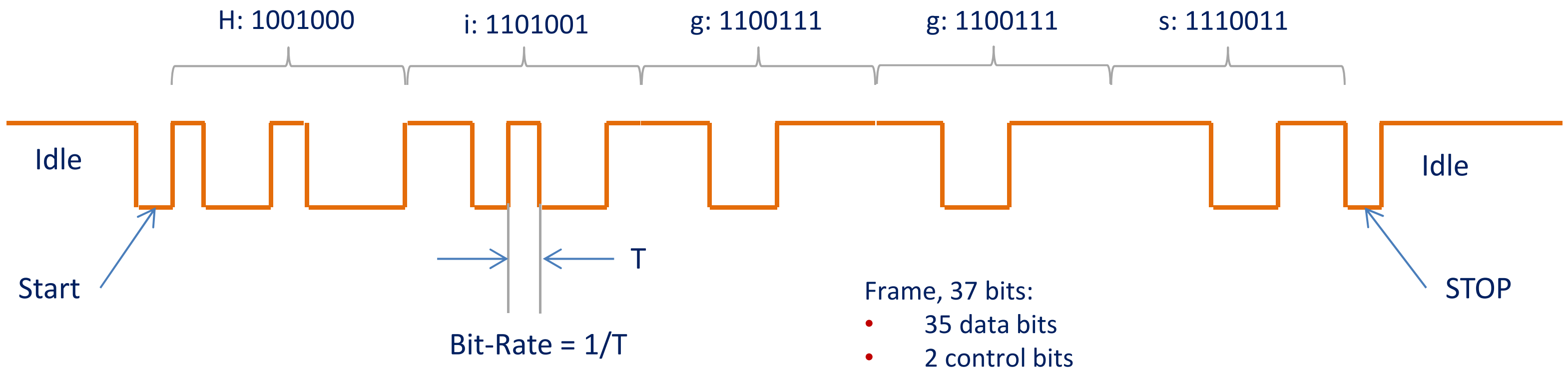
RX

DES

DATA

DAV

Digital System (data sink)



H: 1001000

i: 1101001

g: 1100111

g: 1100111

s: 1110011

Idle

Idle

Start

STOP

T

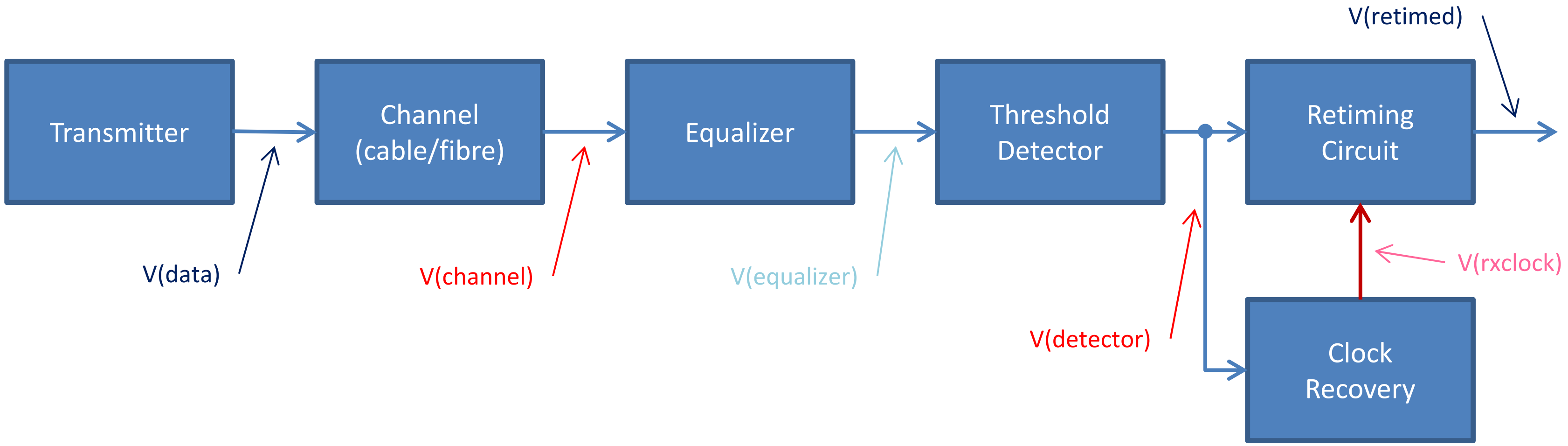
Bit-Rate = 1/T

Frame, 37 bits:

- 35 data bits
- 2 control bits
- "Code efficiency" : 35/37 = 94%

Paulo.Moreira@cern.ch

BW – Limited Channel (1/2)



Transmitter

Channel (cable/fibre)

Equalizer

Threshold Detector

Retiming Circuit

Clock Recovery

V(data)

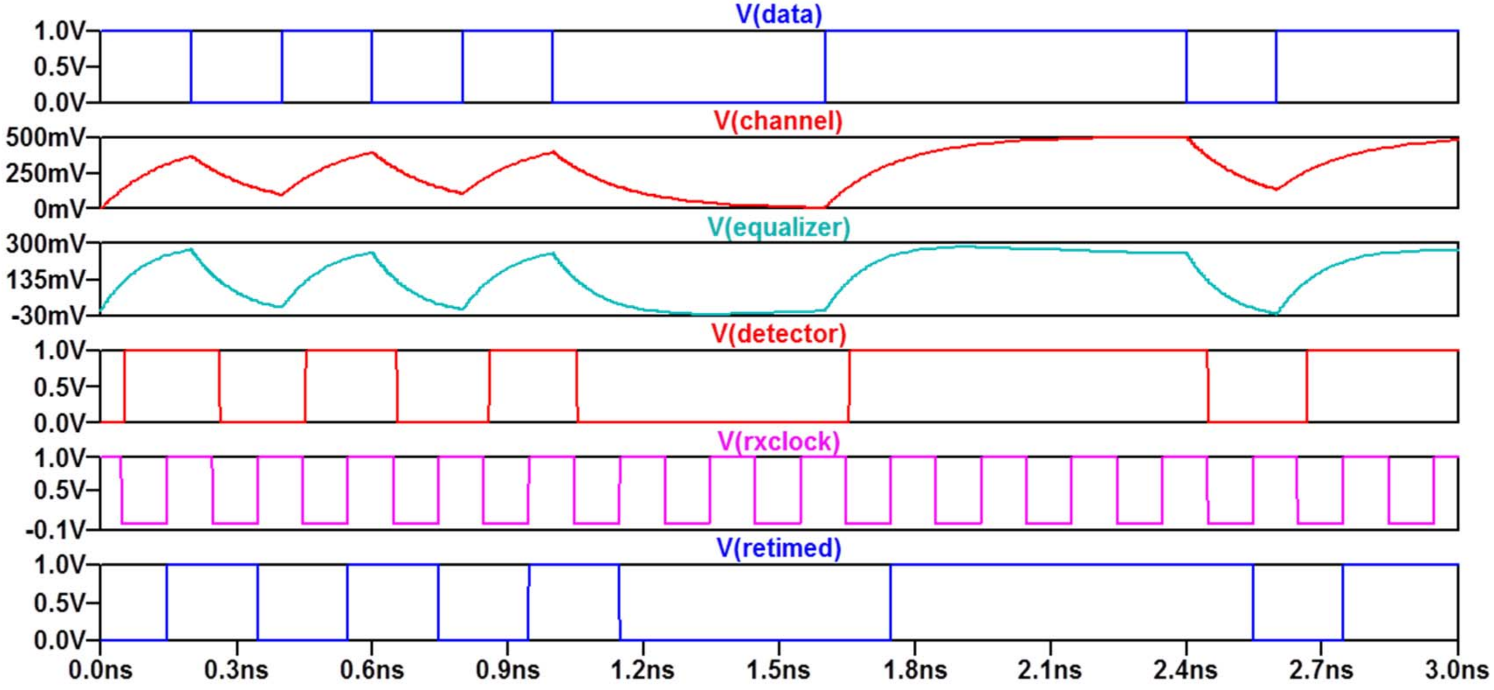
V(channel)

V(equalizer)

V(detector)

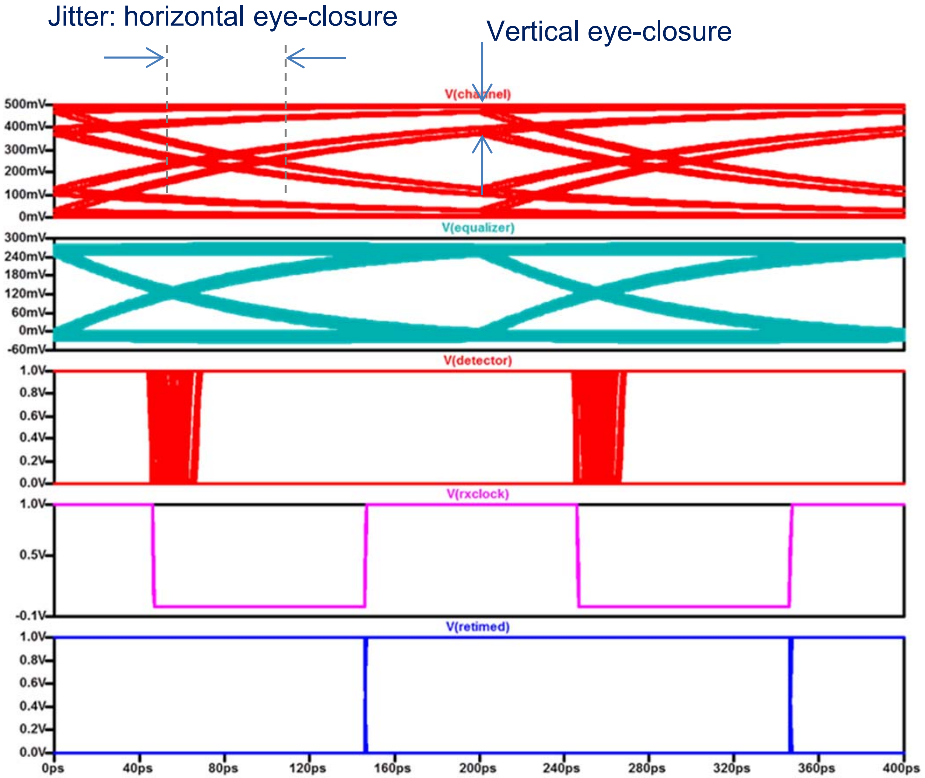
V(rxclock)

V(retimed)



Paulo.Moreira@cern.ch

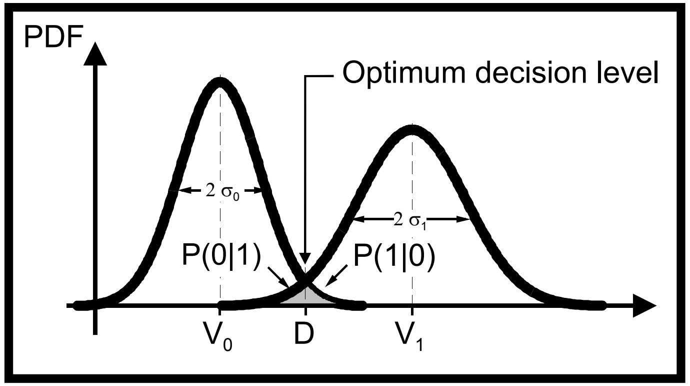
BW – Limited Channel (2/2)

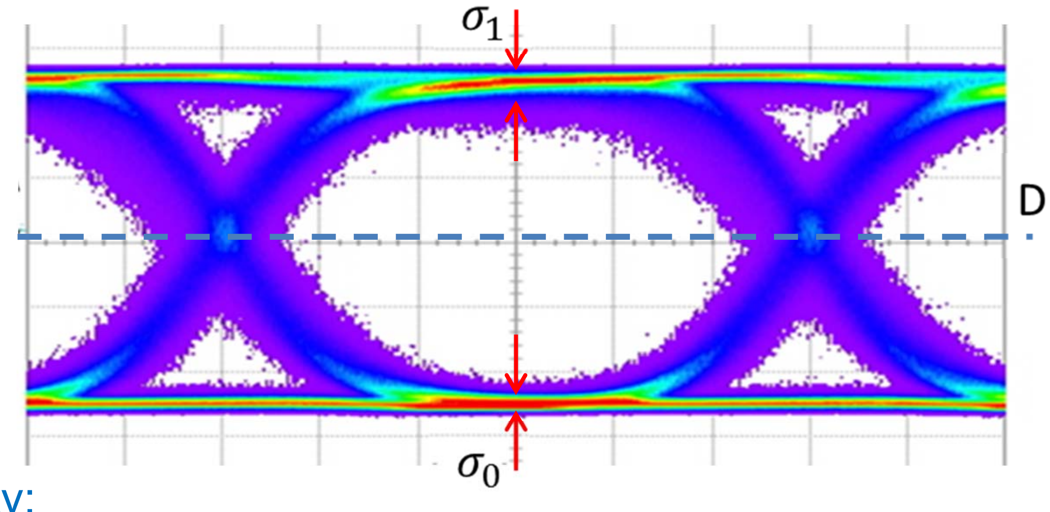


- Due to the limited channel bandwidth the transmitted pulses are broadened in time.
- If the broadening is significant, the signal corresponding to one symbol (bit) will overlap in time the signal of the next symbol!
  - This is called **Inter Symbol Interference** (ISI)
- ISI is seen in an eye-diagram as:
  - Vertical eye-closure:
    - Reduction of the vertical eye-opening
  - Horizontal eye-closure:
    - Random positions of the threshold level crossings (Jitter or Phase Noise)
- Equalization (high-pass filtering in this case) can be used to “restore” the signal high frequency content:
  - Reduces ISI
  - Improves Jitter
  - Can never fully compensate the channel
- Two additional steps are needed to restore the signal to their “original” shape:
  - **Threshold detection**, restores the symbol levels by comparing a signal with a “ pre-defined” level
  - The **signal is retimed** using the recovered clock, eliminating the jitter
- Two additional phenomena impair error free data transmission:
  - **Attenuation**
  - **Noise**

Paulo.Moreira@cern.ch

Noise





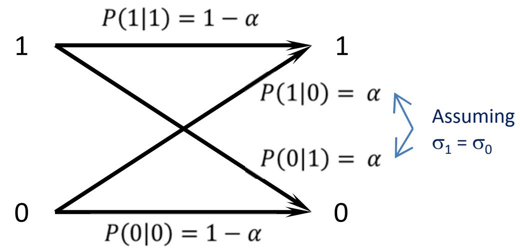
- The average bit error probability:

$$P_e = P(0|1) \cdot P(1) + P(1|0) \cdot P(0)$$

- It is a strong function of the signal to noise ratio:

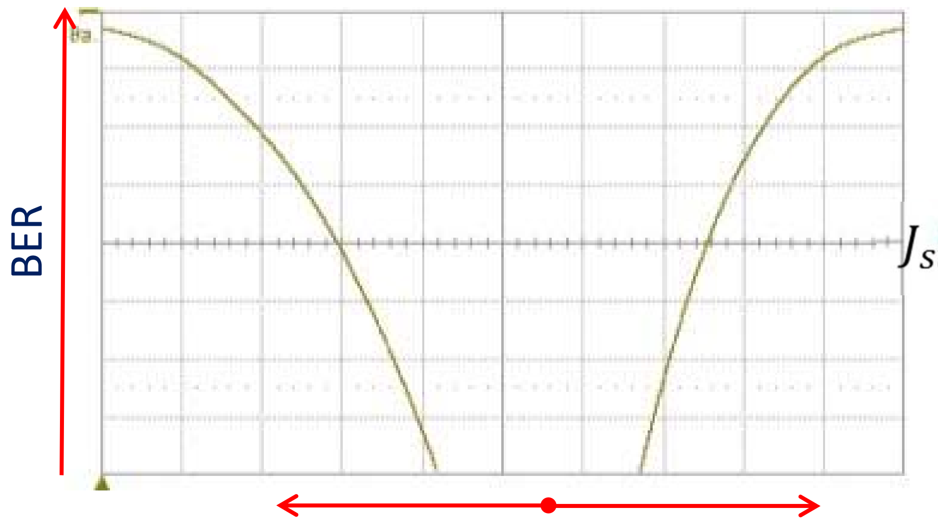
$$SNR_1 = \frac{V_1 - D}{\sigma_1} \qquad SNR_0 = \frac{D - V_0}{\sigma_0}$$

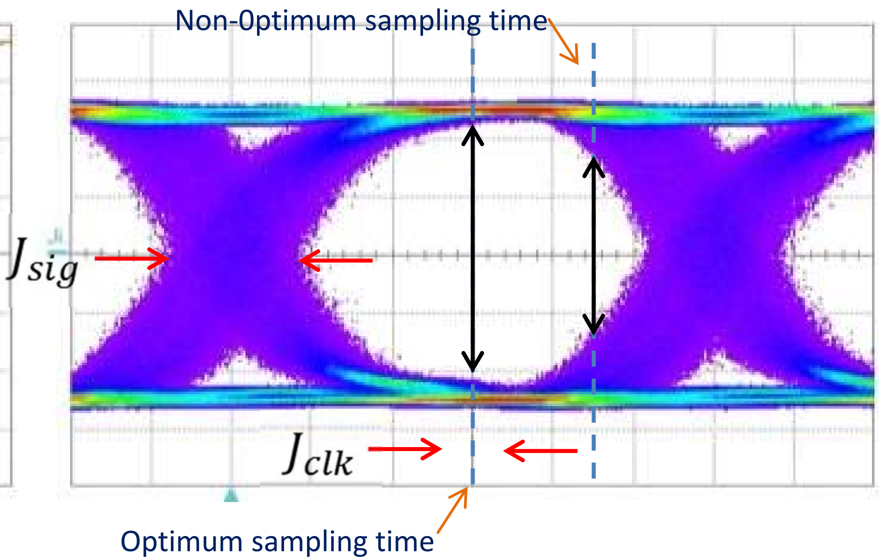
- In the limit the measured Bit Error Rate **BER** is equal to **P_e**
  - SNR = 6 → BER = 10^{-9}
    - Testing time to “count” 100 errors: 20 s @ 5 Gb/s
  - SNR = 7.9 → BER = 10^{-15}
    - Testing time to “count” 100 errors: 231 days @ 5 Gb/s



Paulo.Moreira@cern.ch

Jitter





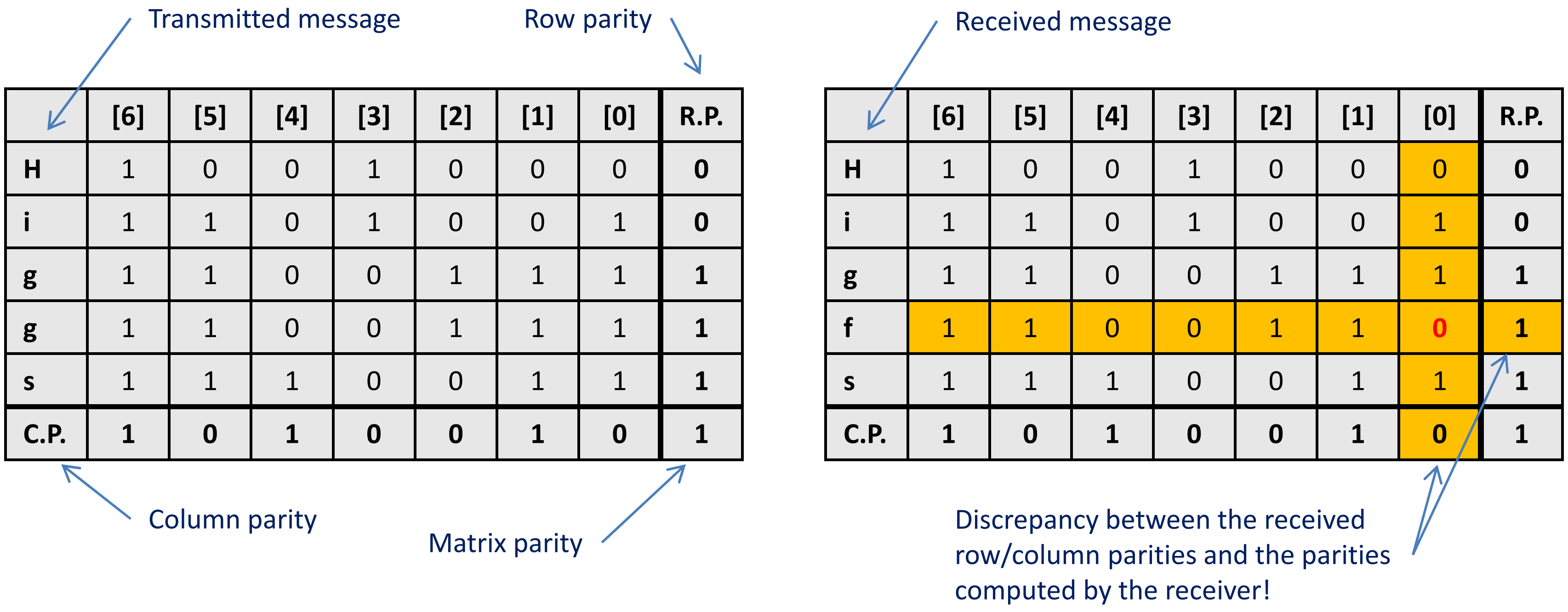
- Moving away from the optimum sampling instant drastically increases the BER because the SNR decreases due to:
  - Signal jitter
  - Signal finite rise and fall times
    - The signal magnitude |V–D| decreases because of the finite raising/fall times of the signal
- Two main causes for non-optimum sampling:
  - Retiming clock static phase error
    - E.g. due an unbalance in the CDR PLL charge-pump
  - Retiming clock jitter
    - E.g. due to the CDR tracking behaviour or VCO noise

Paulo.Moreira@cern.ch

Error Control Coding

- Due to Noise and ISI the received message might differ from the transmitted
  - Some of the transmitted bits will be wrongly detected
- Error control coding introduces extra bits in the transmitted message:
- These allow to:
  - Detect the presence of errors
  - Correct detected errors
- Error control is done at the expense of bandwidth

Even parity: Parity bits are computed such that the number of "1s" in each row and column is even



Transmitted message

Row parity

	[6]	**[5]**	**[4]**	**[3]**	**[2]**	**[1]**	**[0]**	**R.P.**
H	1	0	0	1	0	0	0	**0**
i	1	1	0	1	0	0	1	**0**
g	1	1	0	0	1	1	1	**1**
g	1	1	0	0	1	1	1	**1**
s	1	1	1	0	0	1	1	**1**
C.P.	**1**	**0**	**1**	**0**	**0**	**1**	**0**	**1**

Column parity

Matrix parity

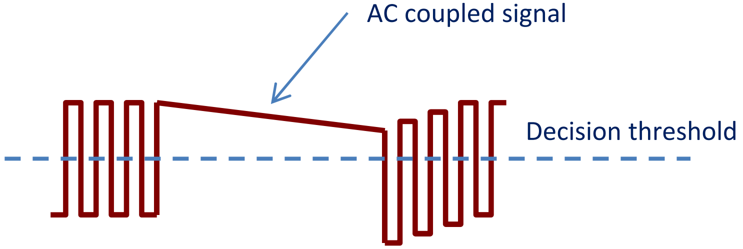
Received message

	[6]	**[5]**	**[4]**	**[3]**	**[2]**	**[1]**	**[0]**	**R.P.**
H	1	0	0	1	0	0	0	**0**
i	1	1	0	1	0	0	1	**0**
g	1	1	0	0	1	1	1	**1**
f	1	1	0	0	1	1	0	**1**
s	1	1	1	0	0	1	1	**1**
C.P.	**1**	**0**	**1**	**0**	**0**	**1**	**0**	**1**

Discrepancy between the received row/column parities and the parities computed by the receiver!

Line Coding

- Apart from detecting and correcting errors, coding is also applied to condition the signal to the transmission medium and/or the transmitter/receiver architecture. This is called line coding.
- Most common addressed problems are:
  - DC wander caused by AC coupling ("high-pass" type response), see figure below:
    - DC blocking between circuits
    - Laser-driver mean power control
    - Offset cancelation circuits in pin-receivers
  - Lack of signal transitions to keep the RX clock locked on the data
- Line codes typically have the following properties:
  - limit the low frequency content in the signal spectrum
  - Guarantee a high density of symbol transitions
- Some popular examples:
  - 8B/10B
  - Scrambling

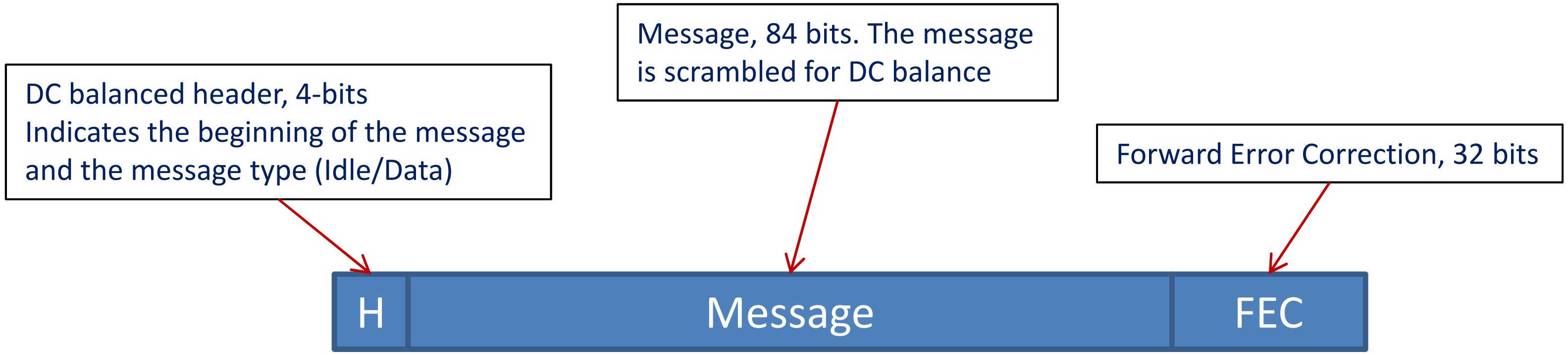


3B/4B

Input	Output		
Disparity	+	0	-
000	1101		0010
001		1001	
010		1010	
011		0011	
100		1100	
101		0101	
110		0110	
111	1011		0100

Paulo.Moreira@cern.ch

Example: The GBT Frame (1/3)



DC balanced header, 4-bits
Indicates the beginning of the message and the message type (Idle/Data)

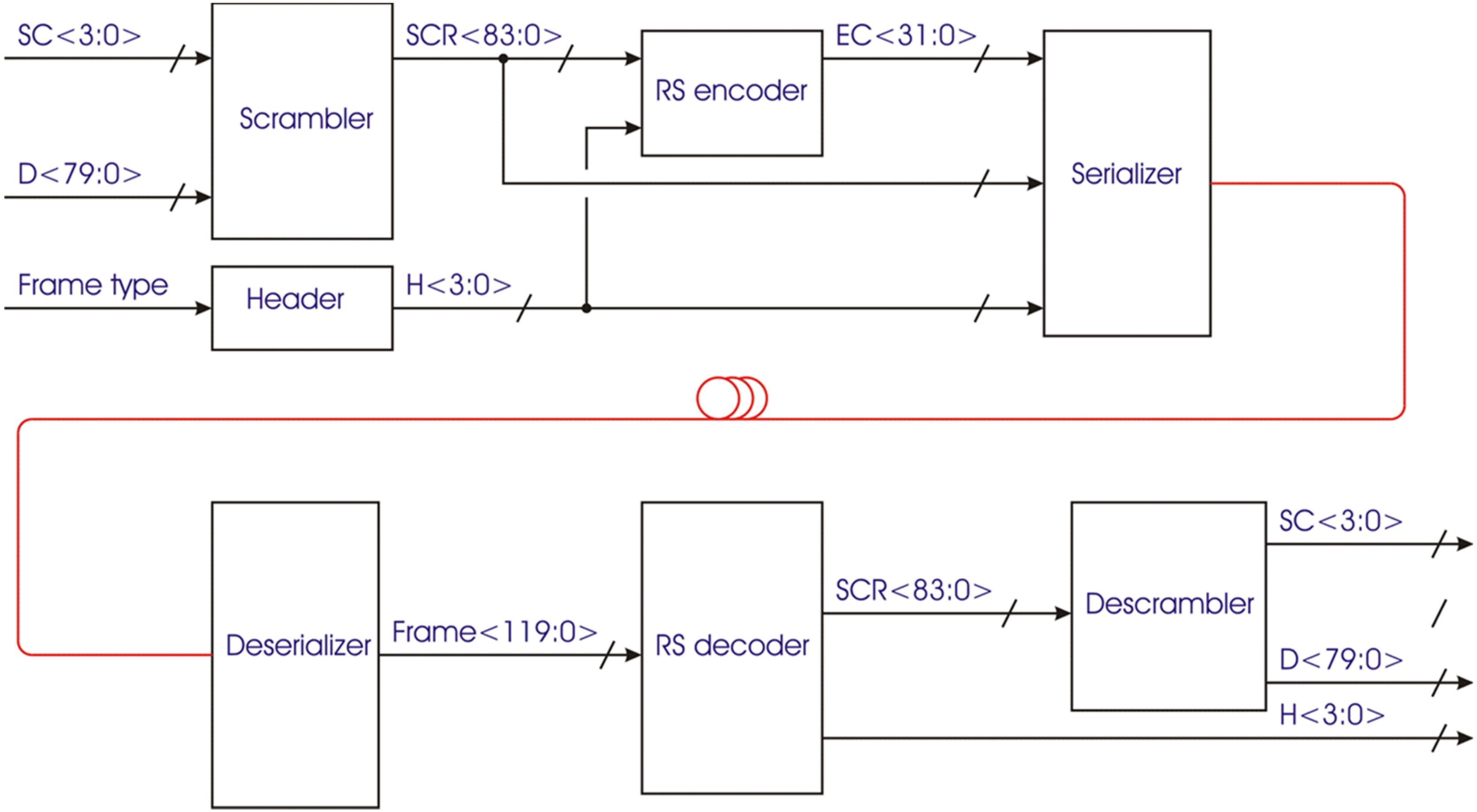
Message, 84 bits. The message is scrambled for DC balance

Forward Error Correction, 32 bits

- Link bandwidth: 4.8 Gb/s
- Payload: 3.36 Gb/s
- DC balance:
  - Scrambler
  - No bandwidth penalty
- Forward Error Correction:
  - Two encoders
    - Each encoder receives 44 bits and computes a 16 bit FEC code
  - Code: Reed-Solomon double error correction
  - 4-bit symbols (RS(15,11))
  - Interleaving: 2
  - Error correction capability:
    - **$2_{\text{Interleaving}} \times 2_{\text{RS}}$ = 4 symbols = 16-bits**
- GBT frame efficiency: 70%
  - A line code is always required for DC balance and synchronization
  - For comparison, the Gigabit Ethernet frame efficiency is 80% (at the physical level – 8B10B coding)
  - At a small penalty (10%, when compared with the Gigabit Ethernet) the GBT protocol will offer the benefits of Error Detection and Correction

Paulo.Moreira@cern.ch

Example: GBT Coding Architecture (2/3)



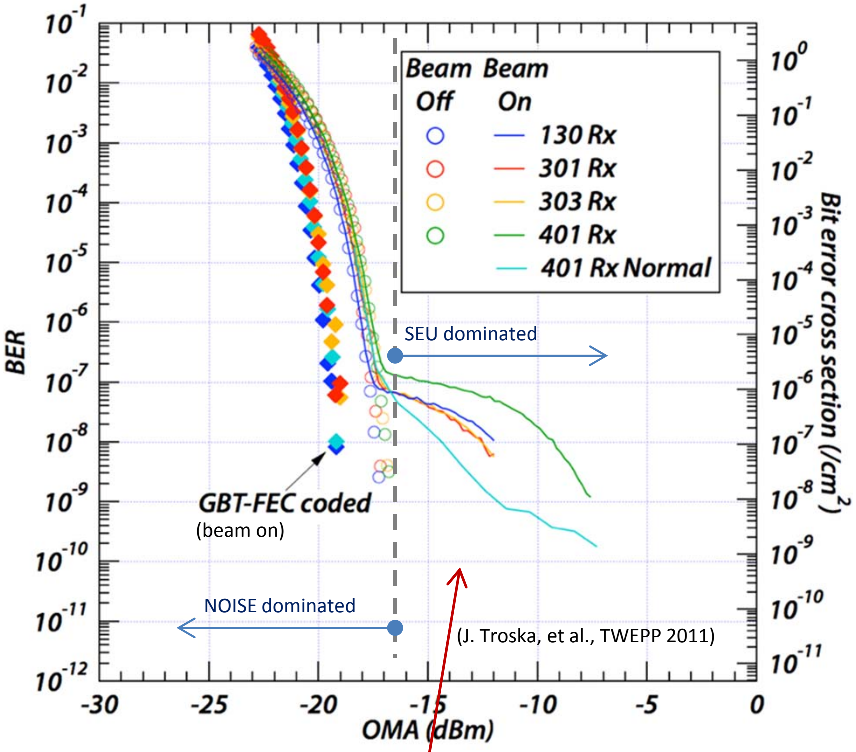
Paulo.Moreira@cern.ch

Example: GBT Forward Error Correction (FEC) (3/3)

- Radiation levels at 20 cm radius from the beam:
  - 2×10^{15} neutrons/cm^2
  - 1×10^{15} hadrons/cm^2
  - 500 kGy total dose
- High rates of Single Event Upsets (SEU) are expected for SLHC links:
  - Particle "detection" by Photodiodes used in optical receivers
  - SEUs on PIN-receivers,
  - SEUs on Laser-drivers
  - SEUs on SERDES circuits
- Experimental results confirmed that:
  - Error correction is mandatory to achieve error-rates ≤ 10^{-12}
  - Upsets lasting for multiple bit periods will occur on PIN detectors!
  - Upset lasting for multiple frames can occur in commercial TIAs

Test conditions:

- 70 MeV proton beam
- Flux of 2×10^8 p/cm^2/s
- ≈ 2 × HL-LHC trackers flux



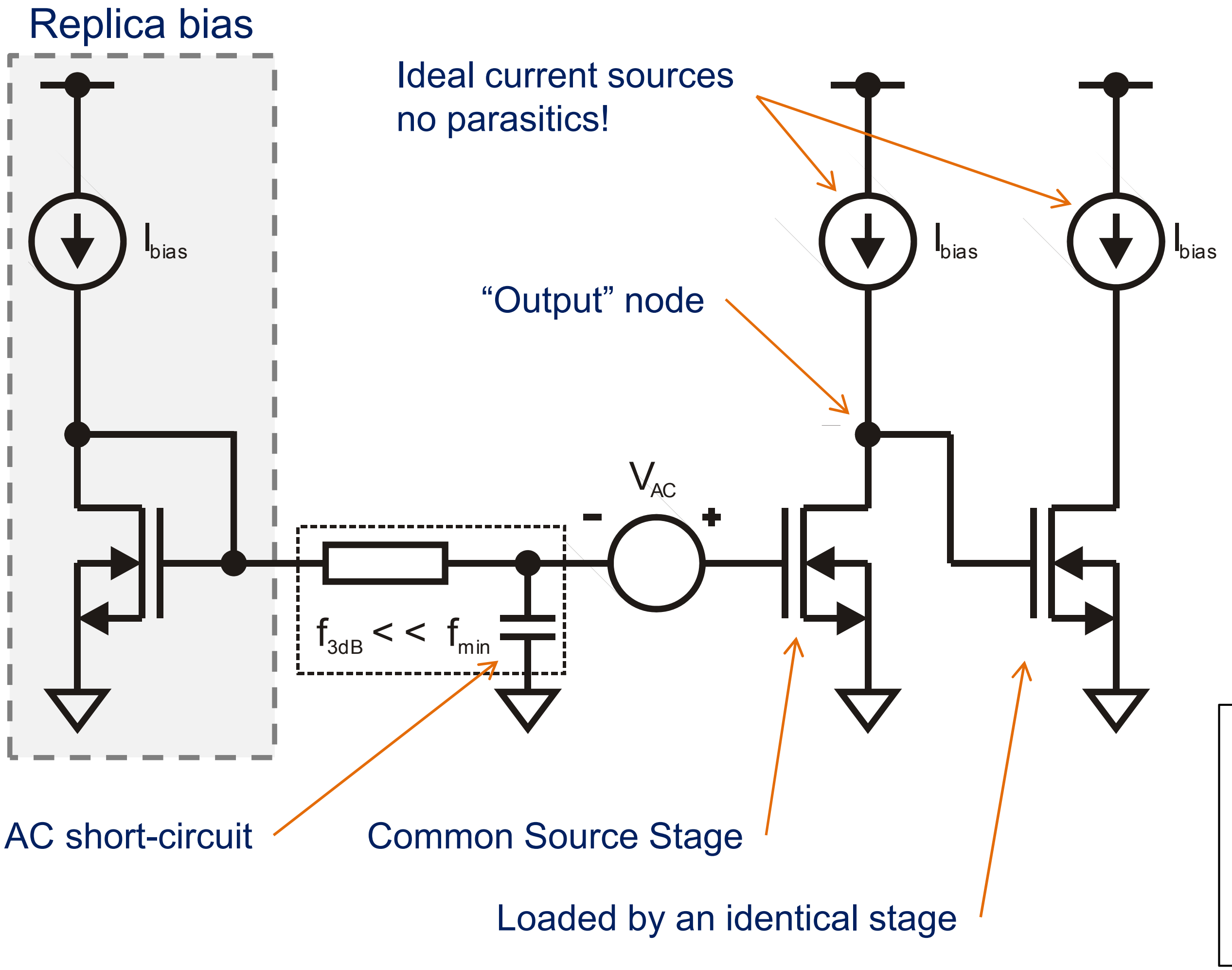
(J. Troska, et al., TWEPP 2011)

No errors were observed in the FEC-coded data-stream in the SEU dominated region with beam on!

Paulo.Moreira@cern.ch

MULTI-GB/S CMOS DESIGN

Basic CS-Amplifier



Replica bias

Ideal current sources
no parasitics!

“Output” node

AC short-circuit

Common Source Stage

Loaded by an identical stage

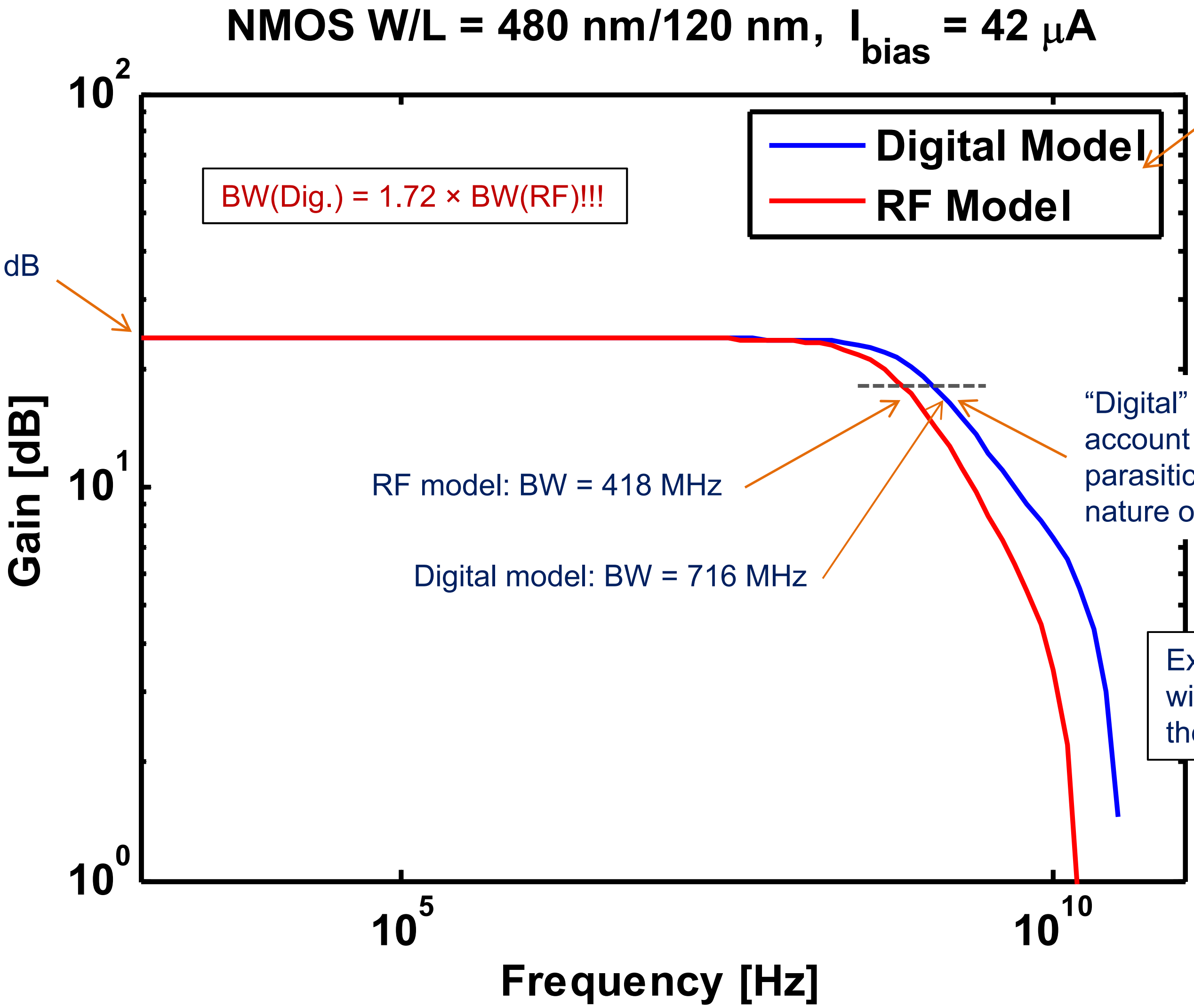
Frequency response is dominated by:

- C_{gs},
  C_{db},
  C_{dg} (“Miller multiplied”)
- g_{ds}

Note: Single ended circuits are used here for simplicity but the results can be easily extended to differential circuits.
(In most circumstances engineers are not expected to design anything else other than differential circuits!)

Paulo.Moreira@cern.ch

Frequency Response: CS Amplifier



BW(Dig.) = 1.72 × BW(RF)!!!

Gain = 21.8 dB
= 12.3

Foundry **might** provide two sets of models!

RF model: BW = 418 MHz

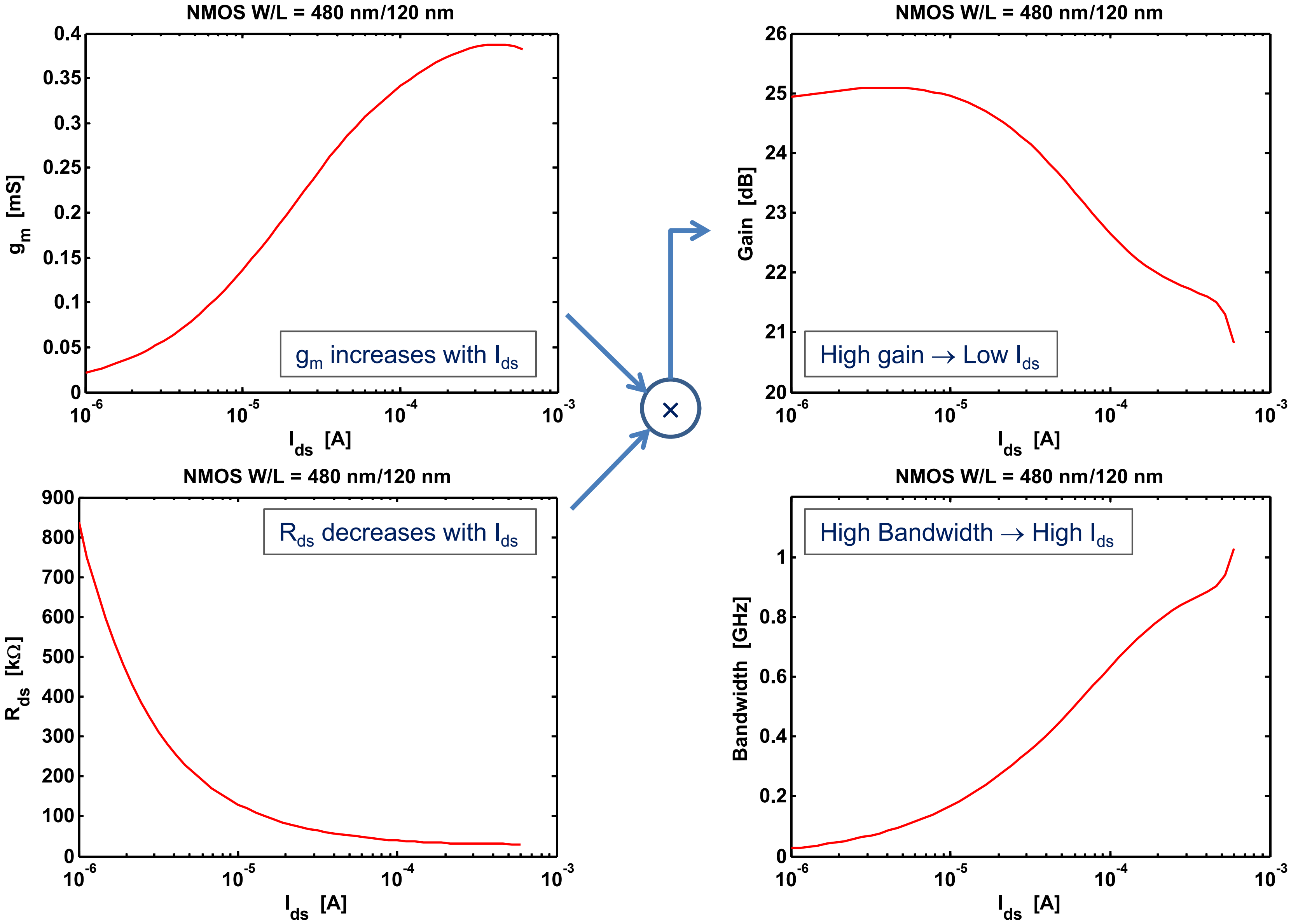
Digital model: BW = 716 MHz

"Digital" models do not account correctly for parasitics and the distributed nature of the gate!

Extracted layouts will always yield the "correct" results!

Paulo.Moreira@cern.ch

Performance (Gain & BW) Depends on I_{ds}



g_m increases with I_{ds}

High gain → Low I_{ds}

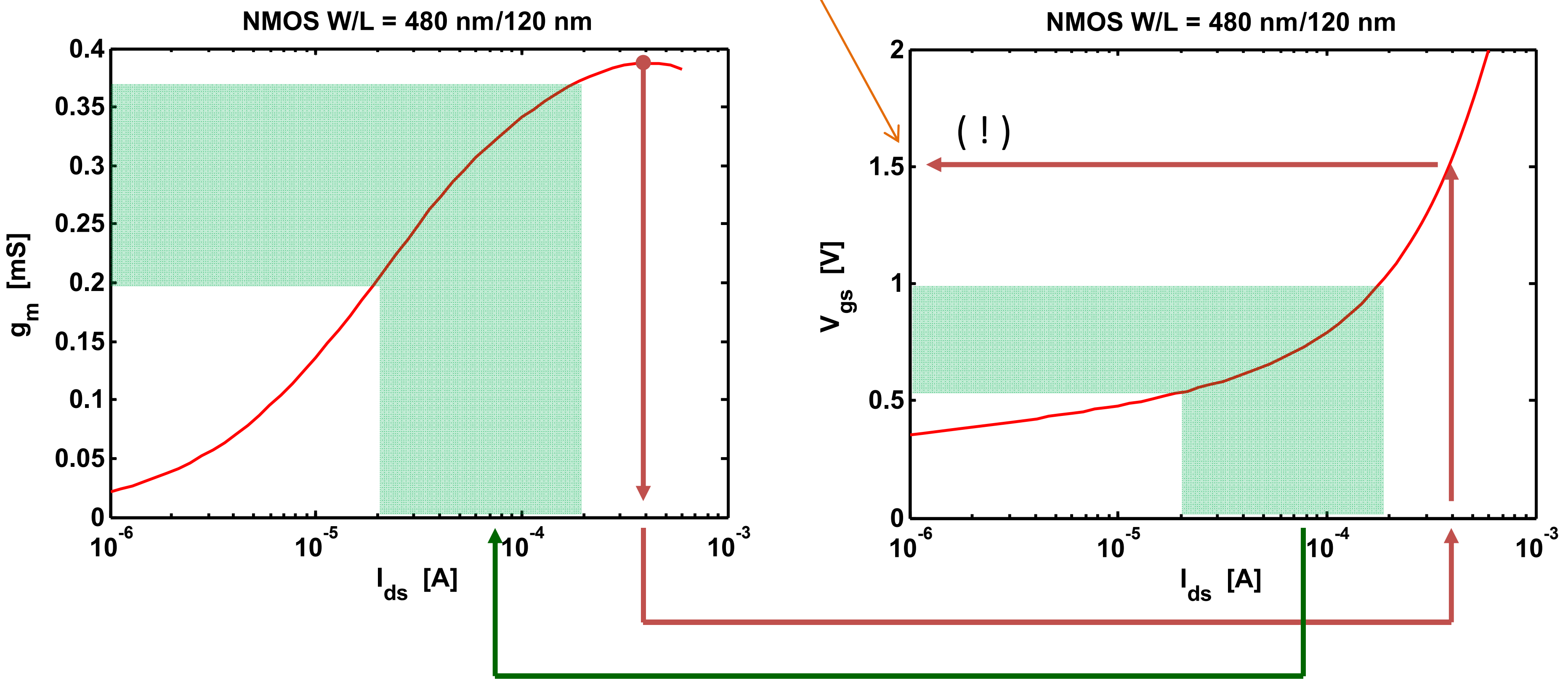
R_{ds} decreases with I_{ds}

High Bandwidth → High I_{ds}

Paulo.Moreira@cern.ch

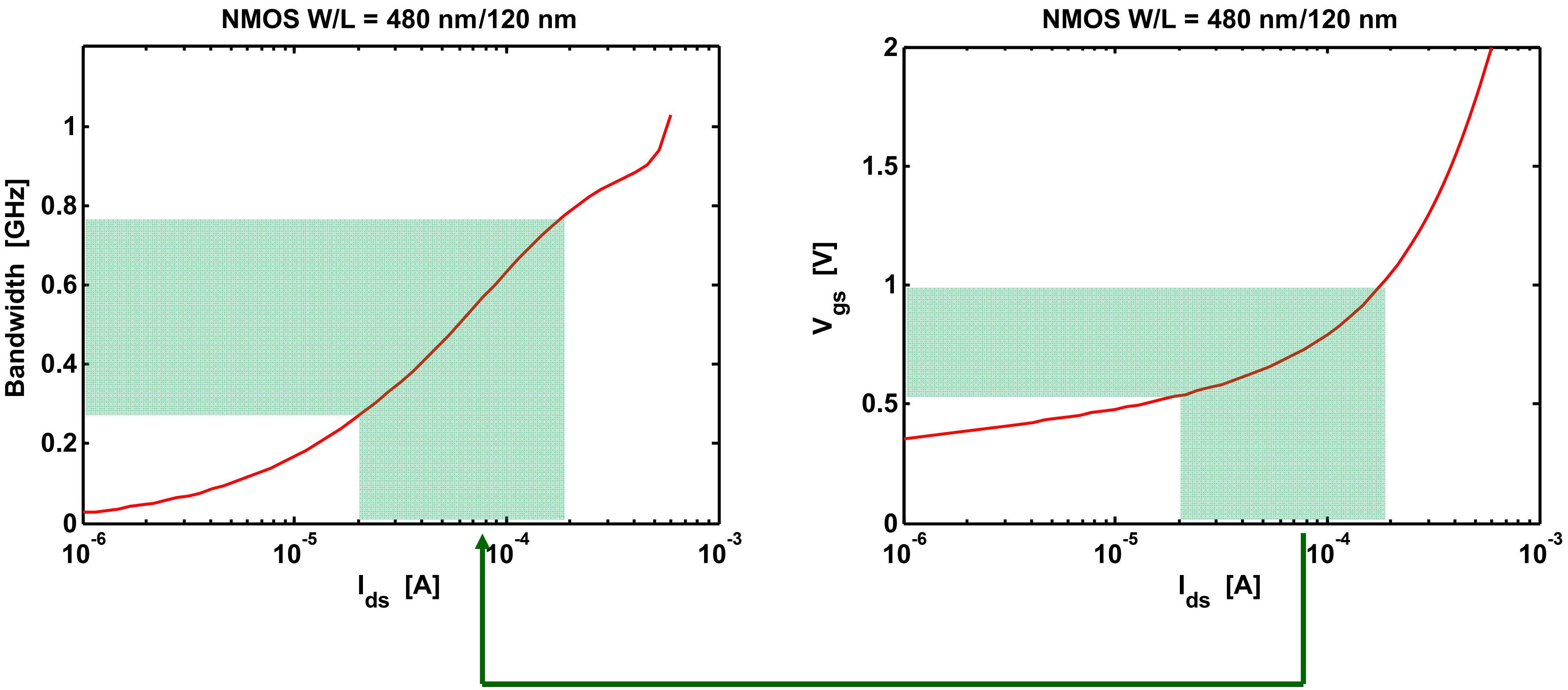
Biasing for Maximum g_m

Biasing for maximum g_m would leave
no headroom to process signals!!!



In practice: $0.5\ V < V_{gs} < 1\ V$
Practically: 50% to 80% of g_m(max)

“Available” Bandwidth: CS Amplifier

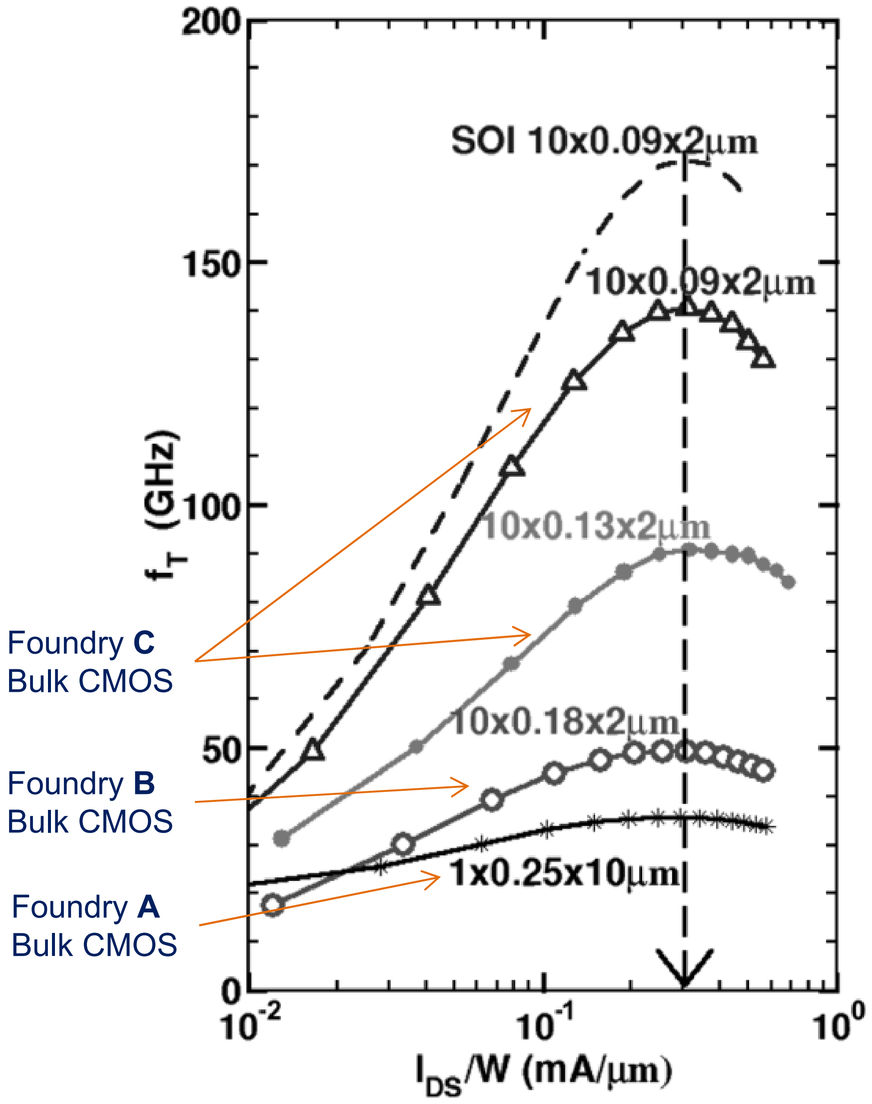


In practice: 0.5 V < V_{gs} < 1 V
Practically: 25% to 75% of BW(max)

Invariance of Characteristic Current Densities (1/6)

- As a result of constant-field scaling the optimum current densities remain largely invariant over technology nodes:
  - $J_{opt}(f_T) \approx 0.3$ mA/μm
  - $J_{opt}(f_{MAX}) \approx 0.2$ mA/μm
  - $J_{opt}(NF_{MIN}) \approx 0.15$ mA/μm
- The optimum current densities are basically independent on:
  - Technology node
  - Foundry
  - Gate length
  - Threshold voltage
- It is valid for NMOS as well as for PMOS
  - PMOS optimum current densities are 40% to 45% of the NMOS values

$$J = \frac{I_D}{W} \quad [in\ mA/\mu m\,]$$



Foundry **C**
Bulk CMOS

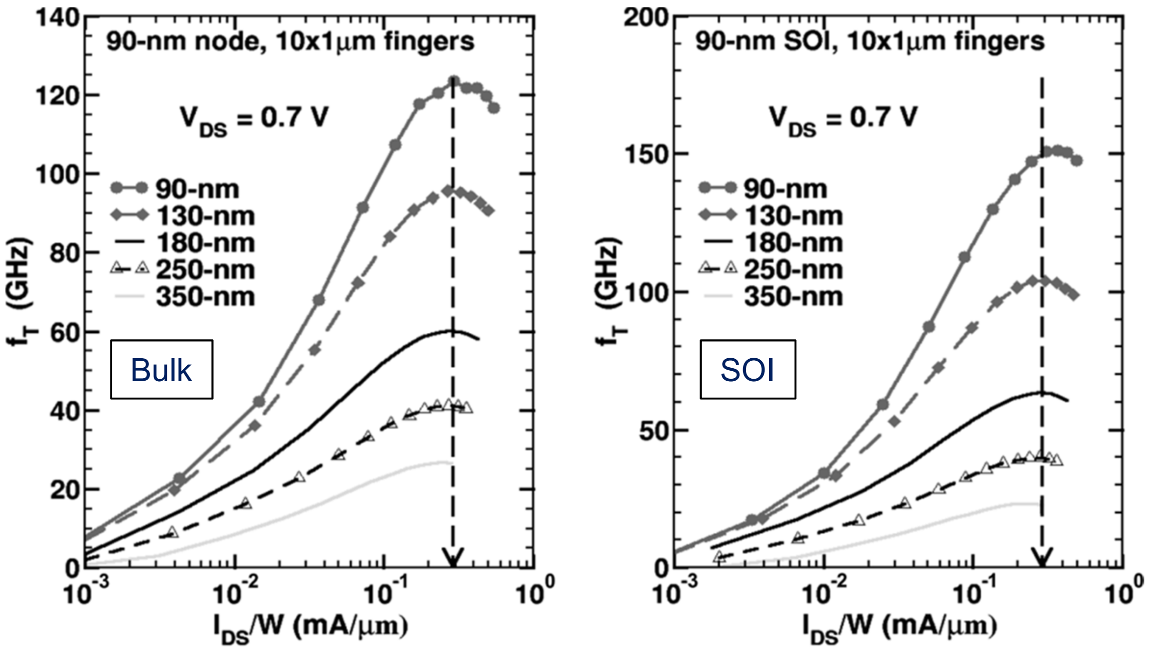
Foundry **B**
Bulk CMOS

Foundry **A**
Bulk CMOS

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

Invariance of Characteristic Current Densities (2/6)

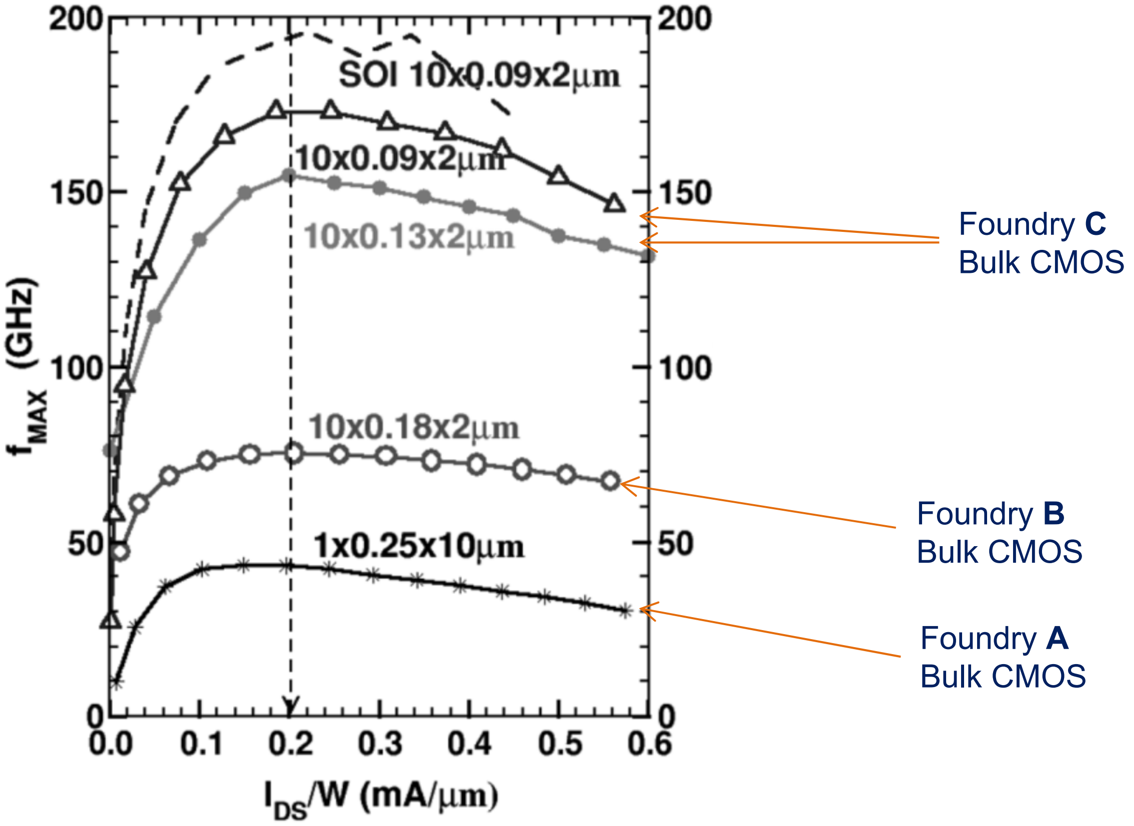
Transistors with various gate lengths in the 90-nm node.



Peak f_T current density:

- $J_{opt}(f_T) \approx 0.3$ mA/µm
- Independent of the gate length within a technology node
- Identical for Bulk and SOI

Invariance of Characteristic Current Densities (3/6)



Foundry **C**
Bulk CMOS

Foundry **B**
Bulk CMOS

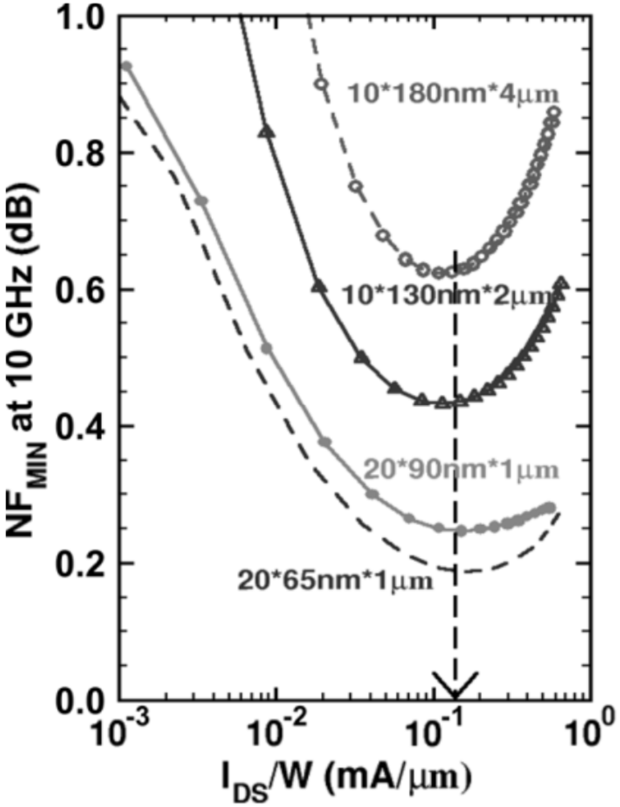
Foundry **A**
Bulk CMOS

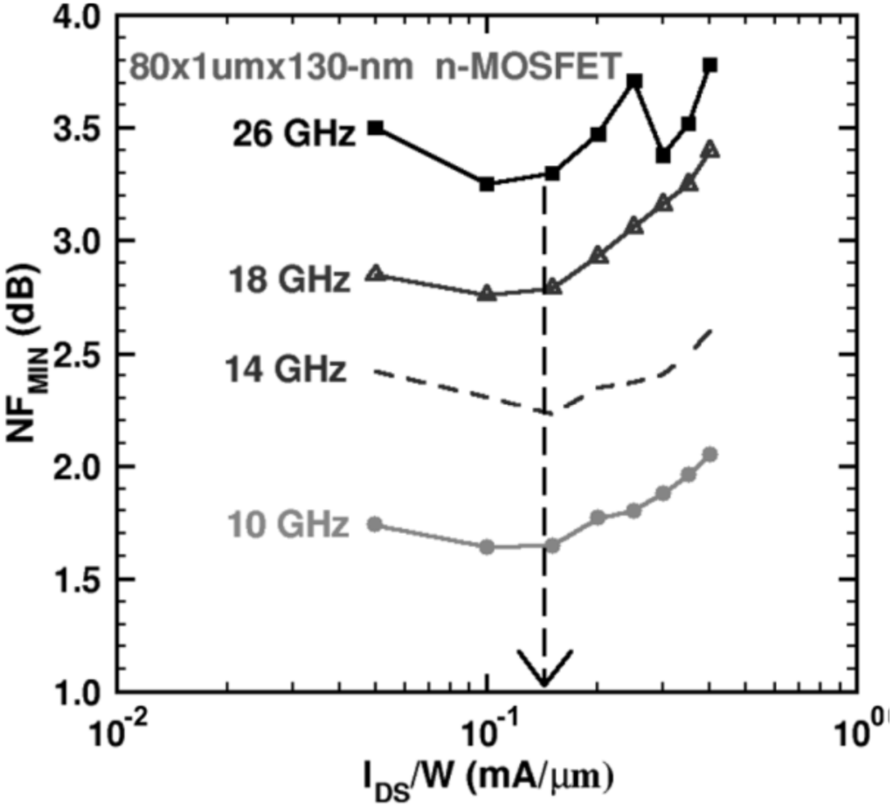
Peak f_{MAX} current density:

- $J_{opt}(f_{MAX}) \approx 0.2$ mA/µm
- Independent of the technology node
- Identical for Bulk and SOI

Paulo.Moreira@cern.ch

Invariance of Characteristic Current Densities (4/6)



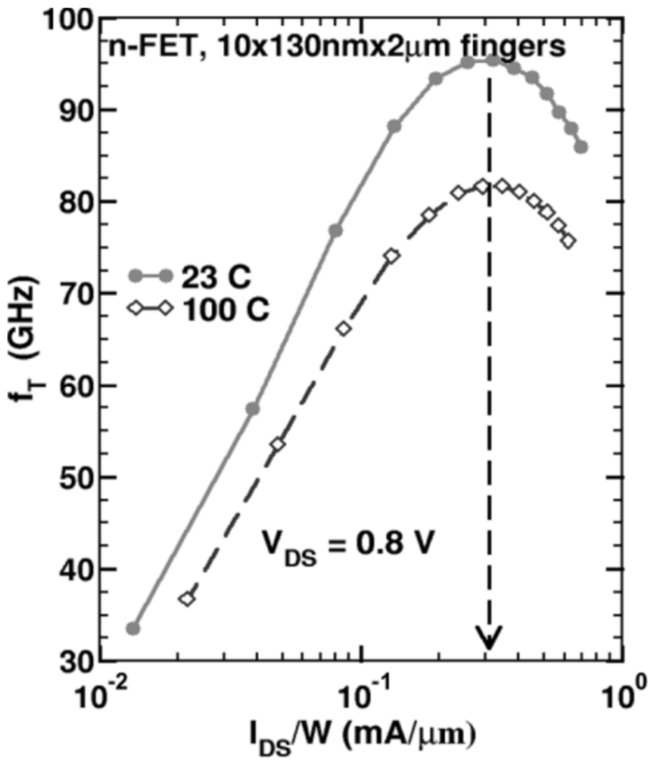


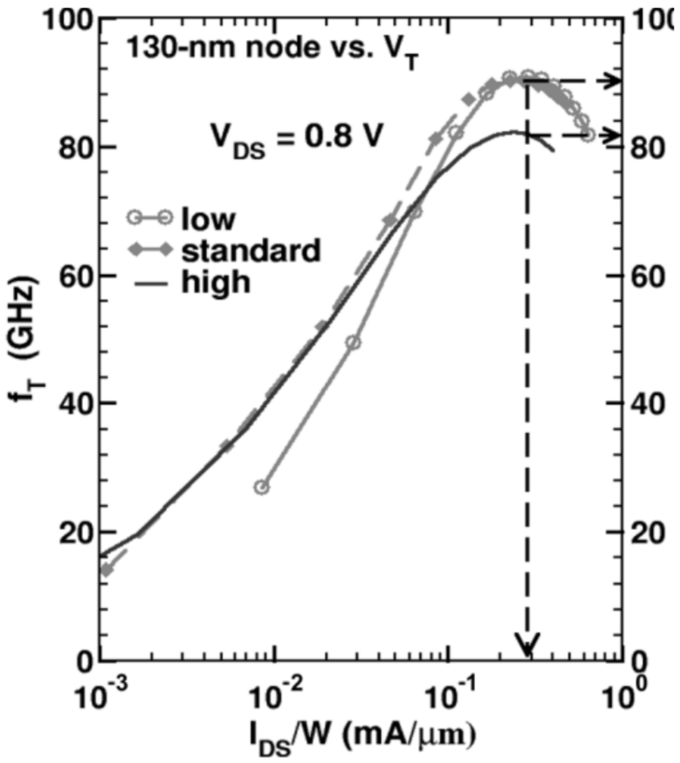
Optimum NF_{MIN} current density:

- $J_{opt}(NF_{MIN}) \approx 0.15$ mA/μm
- Independent of the technology node
- Independent of the frequency

Paulo.Moreira@cern.ch

Invariance of Characteristic Current Densities (5/6)





Optimum current density remain unchanged over:

- Temperature
- Threshold voltages

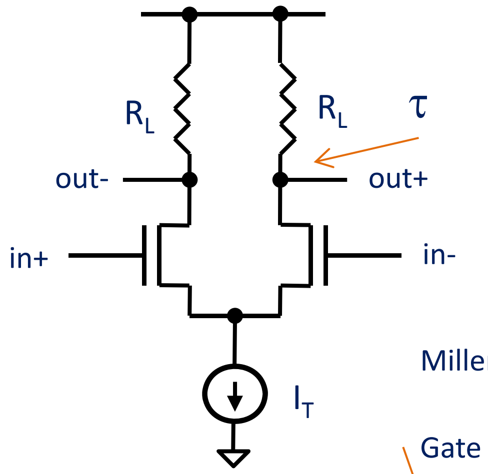
Paulo.Moreira@cern.ch

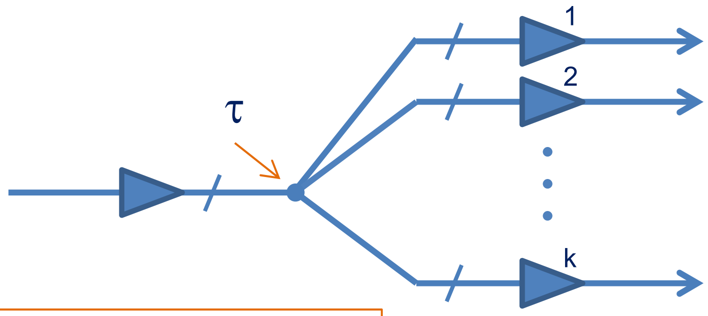
Invariance of Characteristic Current Densities (6/6)

Rule of thumb for high-frequency design:

- All rules of thumb must be critical accessed for the case at hand!
- Employ constant density biasing techniques
- Bias the NMOS transistors in between:
  - The peak f_T and peak f_{MAX} for high bandwidth:
    - $J_{opt} \in [0.2\ mA/\mu m, 0.3\ mA/\mu m]$
  - Or optimum NF_{MIN} for low noise:
    - $J_{opt} \approx 0.15\ mA/\mu m$
- For PMOS transistors use current densities are 40% to 45% of the NMOS values
- This should result in a design that is robust against process variations (1st order assumption):
- Signal headroom might dictate that lower current densities need to be used
  - The optimum values for f_T , f_{MAX} and NF_{MIN} for are relatively broad
  - Using $0.5 \times J_{opt} < J < J_{opt}$ has little influence on performance and can significantly reduce power consumption and/or increase signal handling ability

Application to NMOS – CML Gates





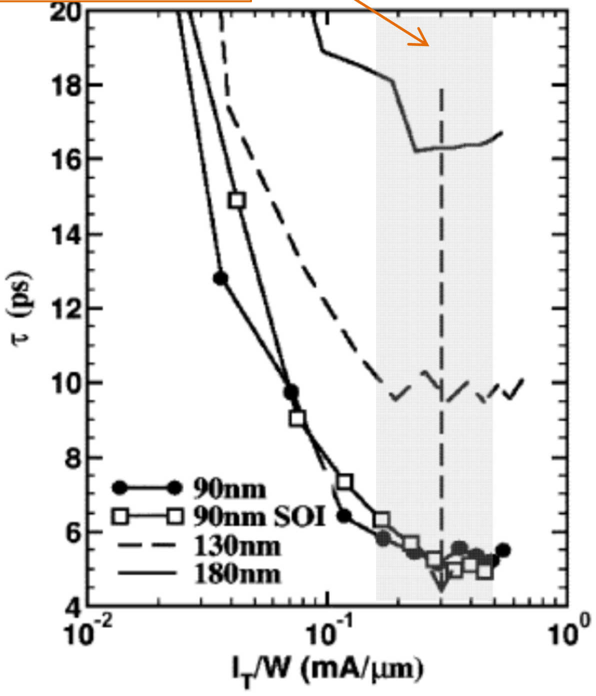
Miller multiplied C_{gd}

Gate resistance

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

$$\frac{\Delta V}{I_T} = R_L$$

Delay changes less than 10% when the current density changes from 0.15 mA/μm to 0.5 mA/μm



Gain

Make sure the “gate” has gain: $\Delta V \geq 1.5 \times \Delta V_{MIN}$

Minimum differential voltage required to fully switch the differential pair!

Transistor Layout: Capacitances

- The gate capacitance is typically dominated by the gate-to-channel (oxide) capacitance (even for relatively small devices):

$$C_g = C_{ox} \cdot W \cdot L + C_{overlap} \cong C_{ox} \cdot W \cdot L \quad (\text{since } C_{ox} \cdot W \cdot L \gg C_{overlap})$$

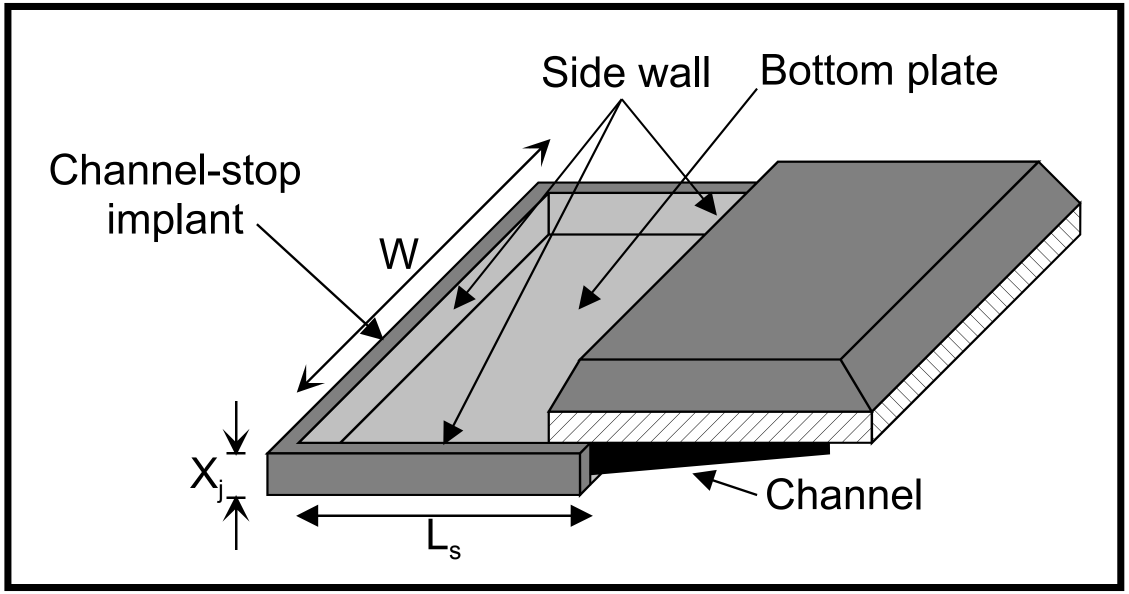
- Source and drain (C_{sb} and C_{db}) capacitances are diffusion capacitances composed of:
  - Bottom-plate capacitance:

$$C_{bottom} = C_j \cdot W \cdot L_s$$

  - Sidewall capacitance:

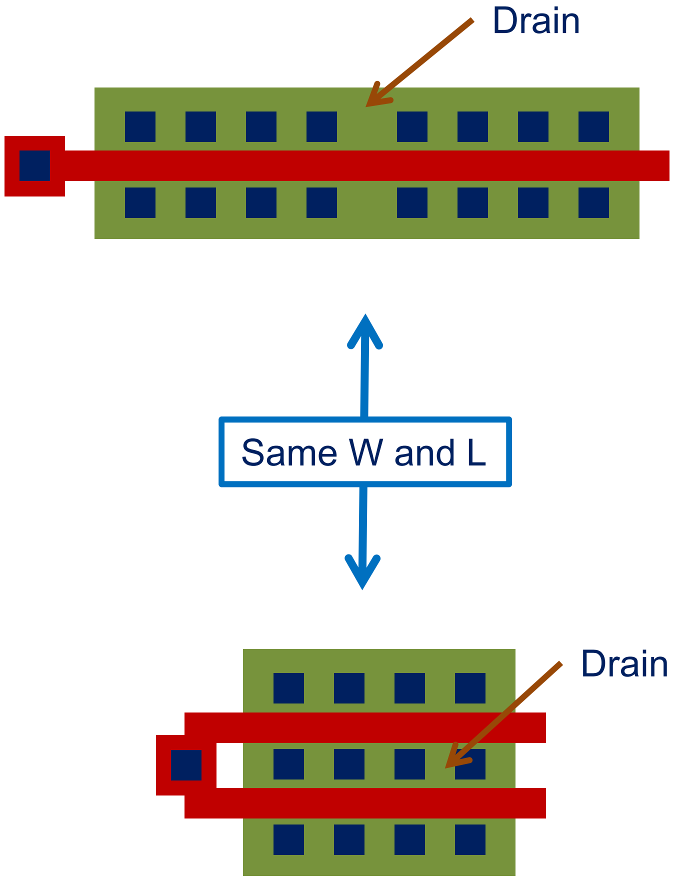
$$C_{sw} = C_{jsw} \cdot (2\,L_s + W)$$

  - In small devices the sidewall capacitance dominates
  - Increasing W decreases the relative importance of the sidewall capacitance
  - Large devices have thus relatively smaller source and drain parasitic capacitances



Paulo.Moreira@cern.ch

Transistor Layout: Multi-finger Devices



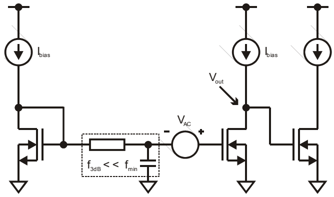
Multi-finger devices have:

- Smaller drain diffusion capacitance:
  - Basically “no” sidewall drain capacitance:
    - Channel on both sides of the drain diffusion
    - A small fraction still remains at the transistor edges:
      - Mainly when an odd number of fingers is used
      - So, no odd number of fingers in the designs!
- Less than half the bottom plate capacitance:
  - Design rules “impose” larger diffusion on the periphery then that in between gates
- Smaller gate resistance:
  - In the example:
    - Two ½ width gates
    - Drive ½ the gate capacitance each
    - Leading to “¼ “ RC delay
  - Gate resistance can be further decreased by contacting on both sides of the gate

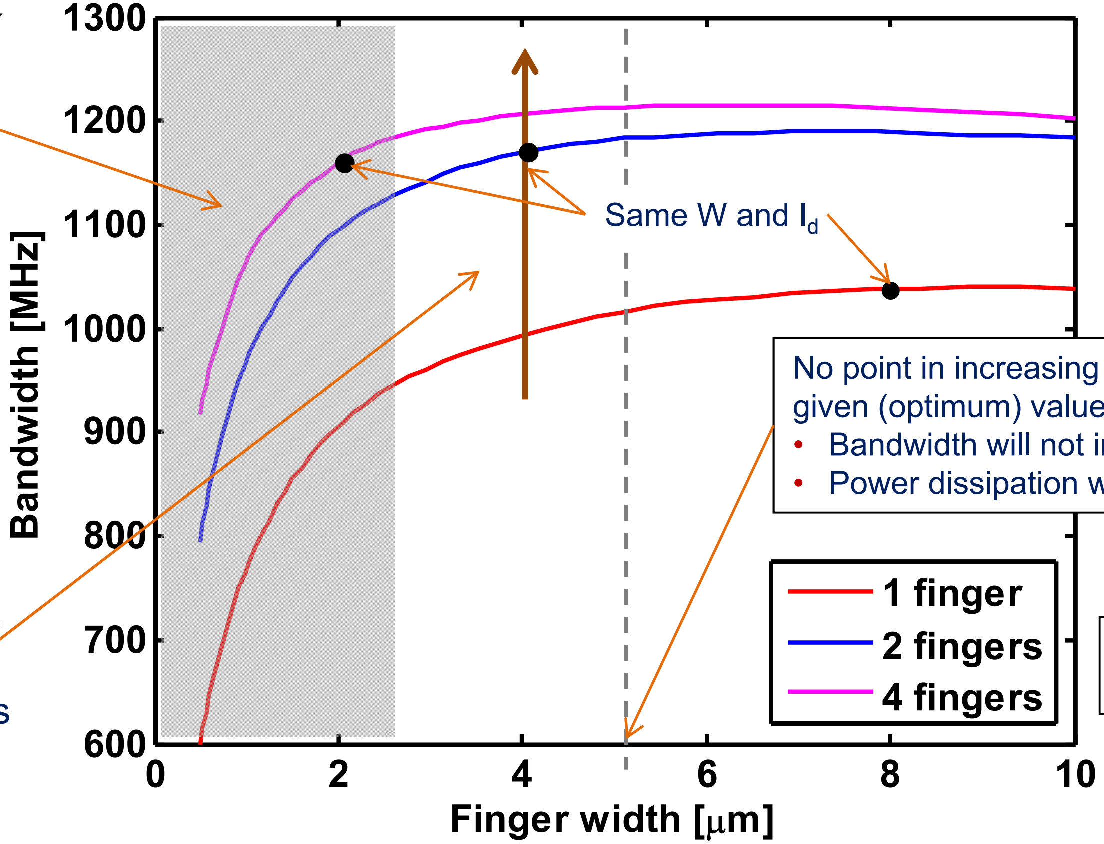
"Available" Bandwidth: Multi-Finger Devices

$J_d = 0.42\ mA/\mu m \rightarrow I_d = NF \times W \times 0.42\ mA/\mu m$

Example:
NF = 4, W = 4 μm
→ $I_{ds} = 6.7\ mA$



NMOS, L = 120 nm, J = 0.42 mA/μm



C_d dominated by C_{SW} in this region

Same W and I_d

No point in increasing W beyond a given (optimum) value:

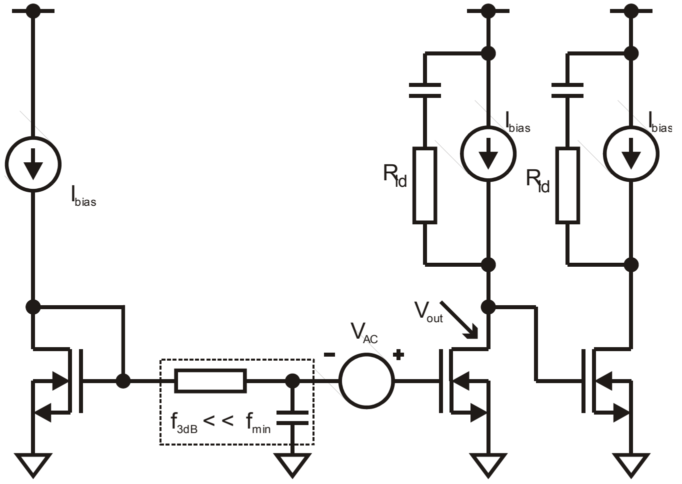
- Bandwidth will not improve but
- Power dissipation will increase!

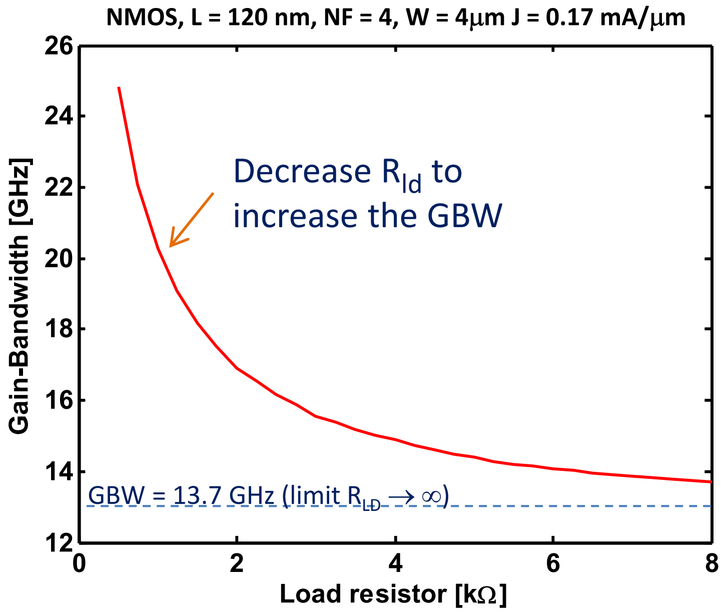
Unless interconnect capacitances dominate

- R_g and C_d decrease
- I_d increases
- C_{SW}/C_{BOT} decreases

- R_g, C_d and I_d increase
- C_{SW}/C_{BOT} decreases

BW Broadening: Resistive Load (1/2)



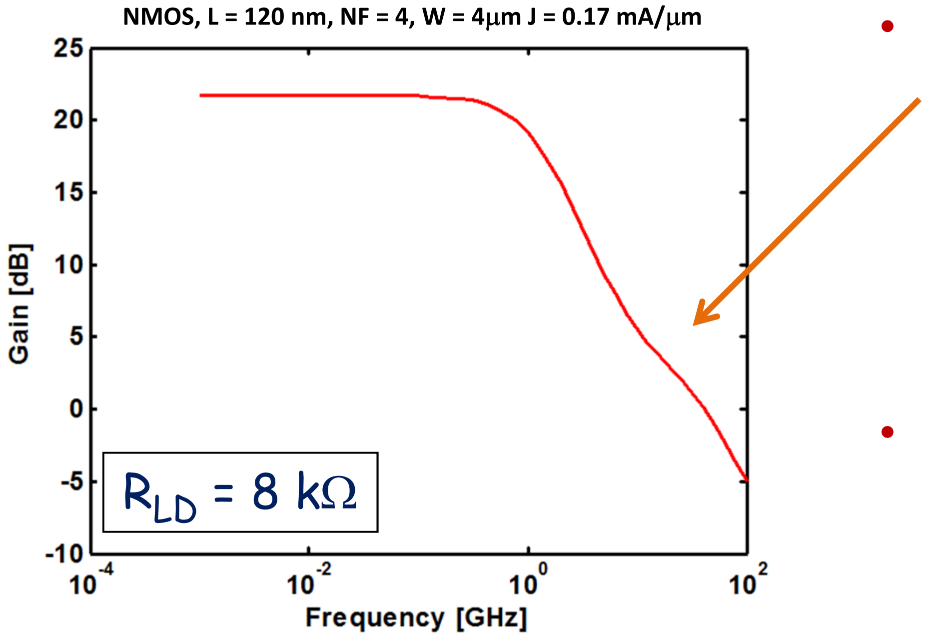


Add a load resistor R_{ld} in parallel with "R_{ds}":

- The output pole will increase:
  - $f_{3dB} = 1/[2\pi (R_{ld} \parallel R_{ds}) C_{gs}]$
- The gain will decrease:
  - $G = g_m (R_{ld} \parallel R_{ds})$
- To a first approximation the gain bandwidth shouldn't change:
  - $GBW = g_m /[2\pi C_{gs}]$
  - (44.8 GHz in the example considered)
- <u>However</u>, the above neglects the Miller capacitance:
  - $GBW = g_m /[2\pi (C_{gs} + C_{Miller})]$
  - GBW = 13.7 GHz
- Since the gain decreases as the load resistance decreases the Miller capacitance (due to C_{gd}) also decreases
- Due to the lower gain, the Miller capacitance is smaller than that of the basic CS stage and the GBW is thus higher than that of the basic CS stage
- In practice the current source (represented in the figure) is avoided altogether and the transistor biased through the load resistor. This avoids the output capacitance C_{db} of an additional current source
  - The parasitic capacitance of the load resistor must be taken into account

Paulo.Moreira@cern.ch

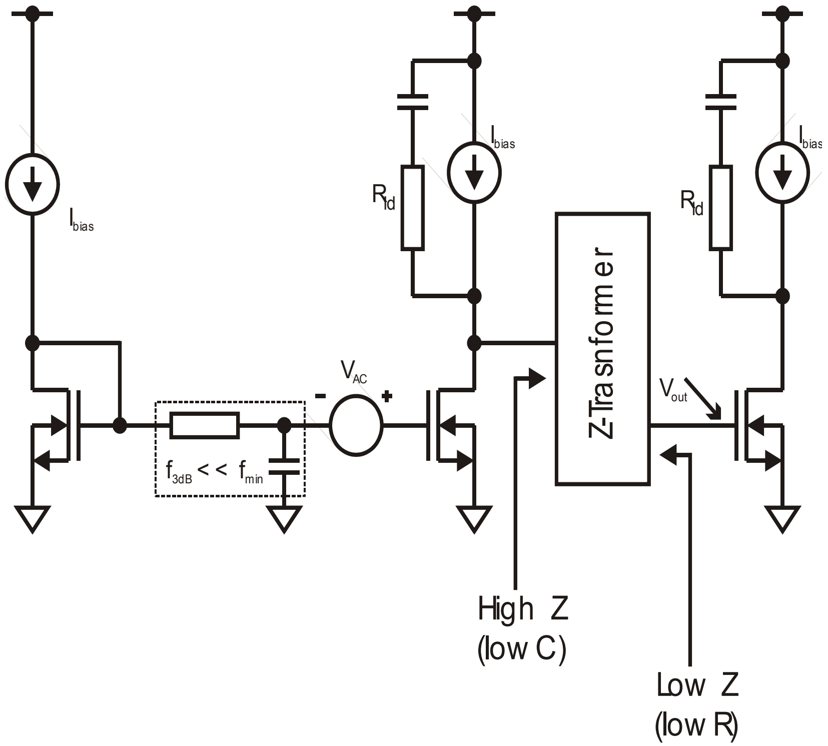
BW Broadening: Resistive Load (2/2)



- The Gain-Bandwidth is:
  - CS: 13.7 GHz
- A zero exists in the transfer function due to the frequency dependence of the Miller capacitance.
  - The next stage transfer function has at minimum a single pole
- Both corner frequencies are dependent on the badly controlled C_{gd} and on the open-loop gain of the next stage.

Paulo.Moreira@cern.ch

BW Broadening: Impedance “transforming”

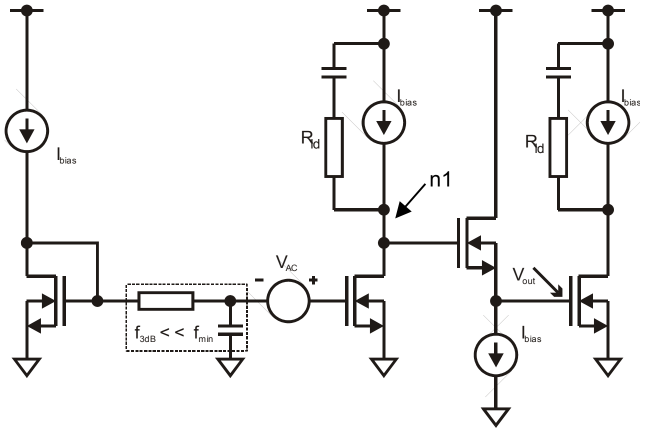


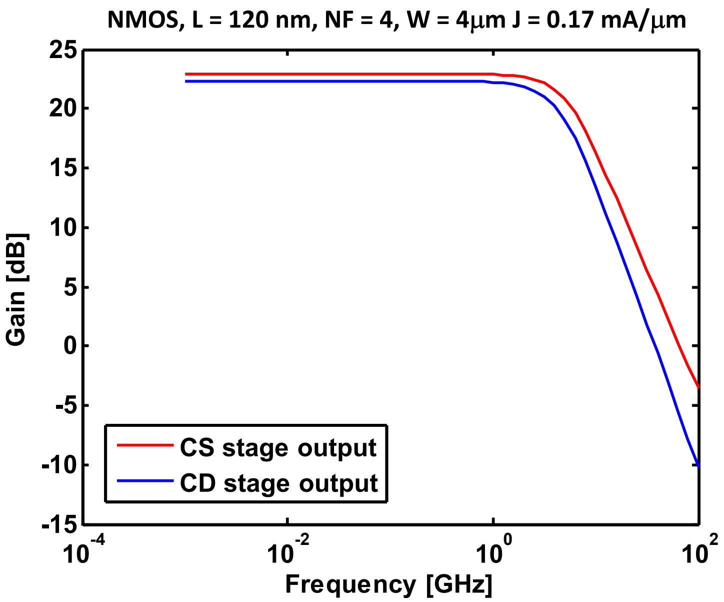
To achieve a high Gain-Bandwidth product:

- Isolate the 1st Gain stage as much as possible from the 2nd:
  - That is, lower the capacitive loading of the 1st stage
- Drive the second stage with as low as possible impedance:
  - That is, move as high as possible the pole introduced by C_{gs} and the Miller capacitance of the 2nd stage
- In other words, transform the impedance that stage 1 represents to stage 2 and vice versa.
- A common-drain (CD) stage can do that job. It has a:
  - High input impedance: $Z_{in} \cong Z[C_{gs}, C_{gd}]$
    - No Miller multiplied capacitance
    - Bootstrapped C_{gs}
  - Low output impedance $Z_{out} \cong 1/g_m$

Paulo.Moreira@cern.ch

BW Broadening: CS-CD Gain “Stage” (1/2)

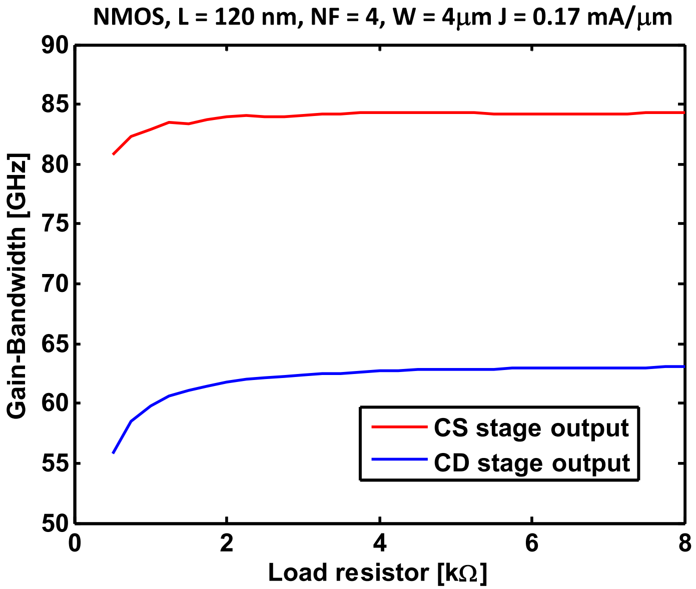


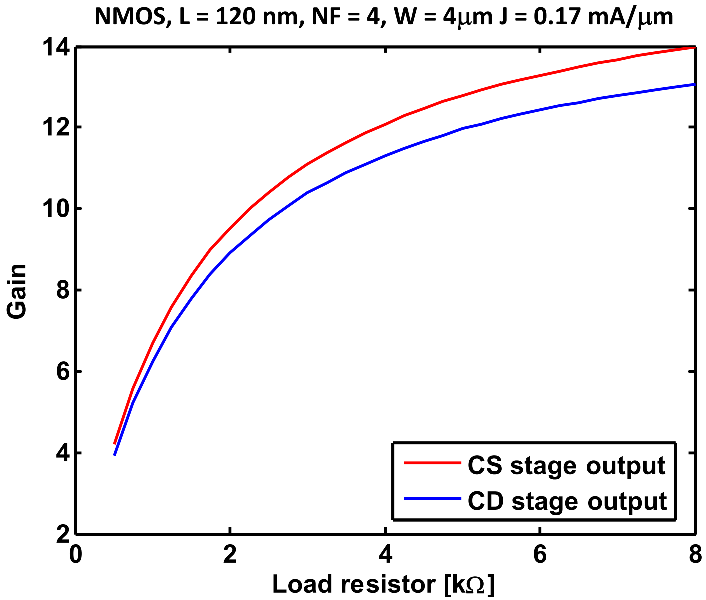


- The Gain-Bandwidth has considerably increased:
  - CS: 13.7 GHz
  - CS-CD: 63.1 GHz
- The frequency response of the CS-CD “stage” is almost “single pole”:
  - The Miller multiplied capacitance C_{gd} is driven by the low impedance of the common drain stage ($1/g_m$)
  - The second stage pole is pushed to a very high frequency
  - The dominant pole is now:
    - $f_{3dB} = 1/[2\pi (R_{ld} \,||\, R_{ds}) C_{n1}]$
    - Where C_{n1} is the “bootstrapped” gate capacitance of the CD stage plus parasitics
- The CD buffer stage has some drawbacks:
  - Additional power consumption
  - Harder to bias for low supply voltages:
    - Additional V_{gs} drop.
    - $V_{dd} \leq 1.5$ V for 0.13 μm CMOS

Paulo.Moreira@cern.ch

BW Broadening: CS-CD Gain “Stage” (2/2)

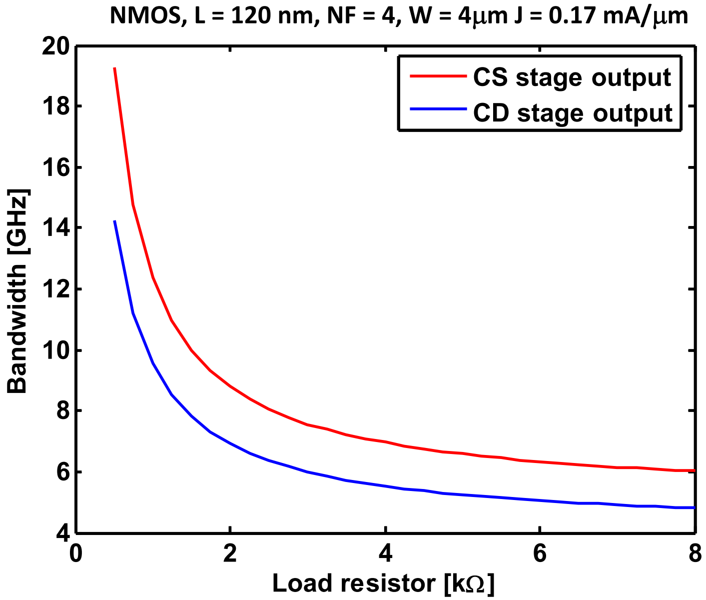




Compared with CS stage, the gain-bandwidth product has “dramatically” increased and is practically independent of the load resistance as it would be the case for a CS stage if the Miller capacitance wouldn’t be present.

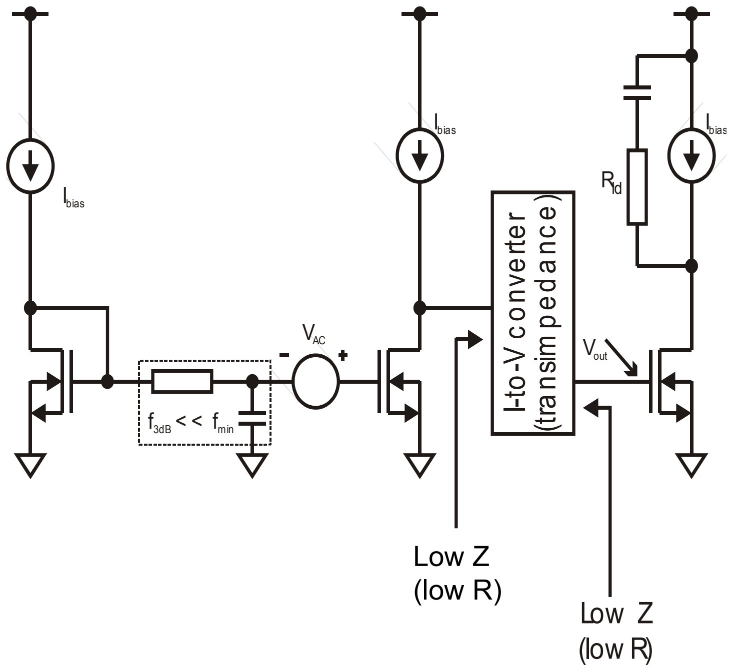
In the simulations:

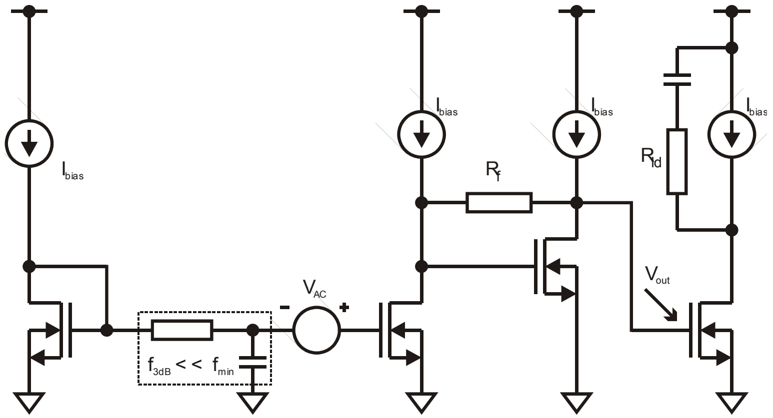
R_{LD} = 1 kΩ:
GBW = 59.8 GHz
G = 6.3
BW = 9.5 GHz



Paulo.Moreira@cern.ch

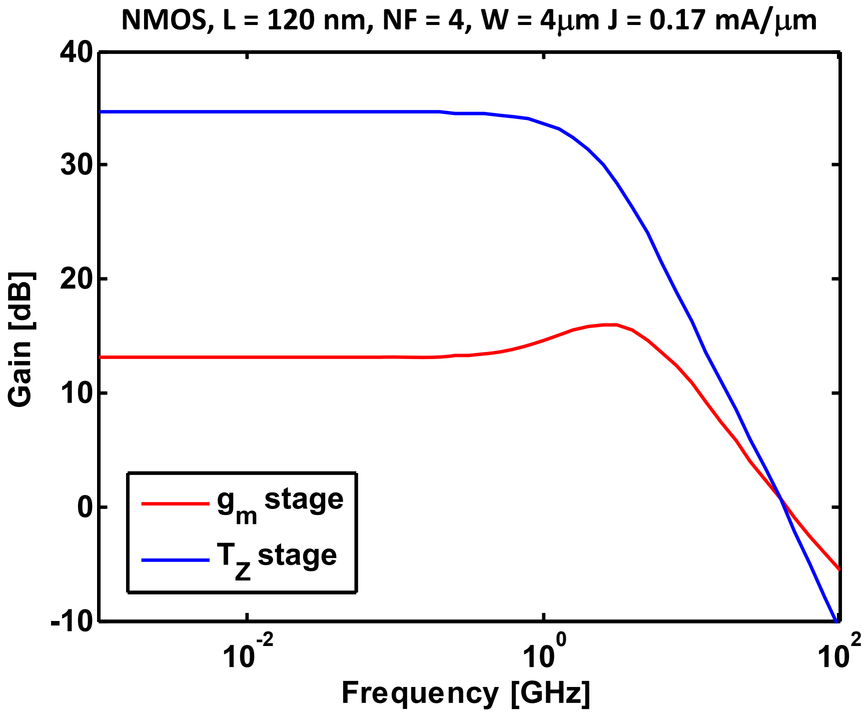
BW Broadening: Cherry & Hooper Gain “Stage” (1/3)





- Cherry and Hooper introduced in 1963 a technique that allows to achieve Gain-Bandwidth products of the order of 90 % of f_T for bipolar transistor circuits. The principle can be applied to MOS circuits
- The basic ideas are:
  - All the nodes in the signal path are low impedance
  - Voltage amplification is achieved by a cascade of two stages: a G_m stage followed by a transimpedance (T_Z) stage
  - For maximum signal transfer the stages must alternate in a succession of high-impedance / low-impedance stages
  - For bipolar transistors due to the relative low impedance of the base, this was implemented as alternating series-feedback / shunt-feedback stages
  - MOS are inherently high input impedance stages so a series feedback stage is strictly speaking not necessary
  - In MOS circuits series-feedback can be used to set accurately the gain. However, this will cost power and will increased noise (reduction of the first stage effective G_m).

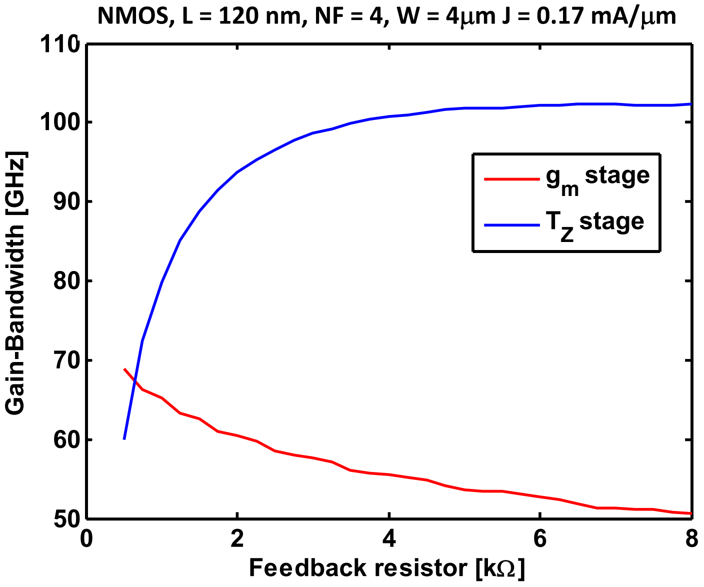
BW Broadening: Cherry & Hooper Gain "Stage" (2/3)

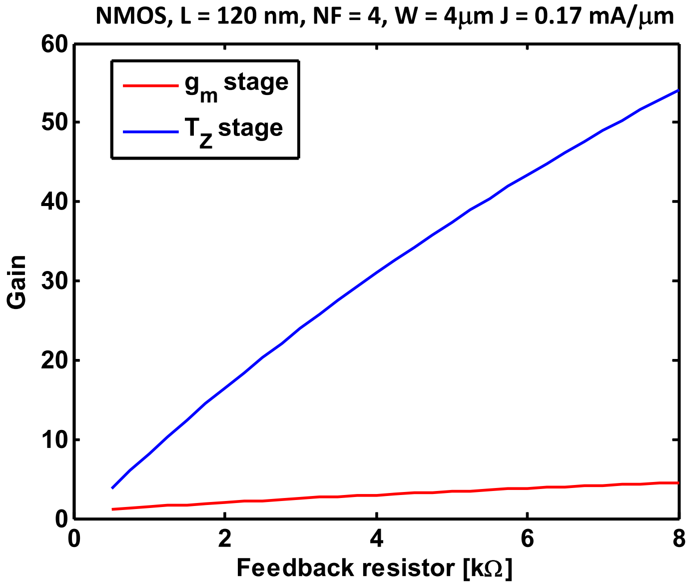


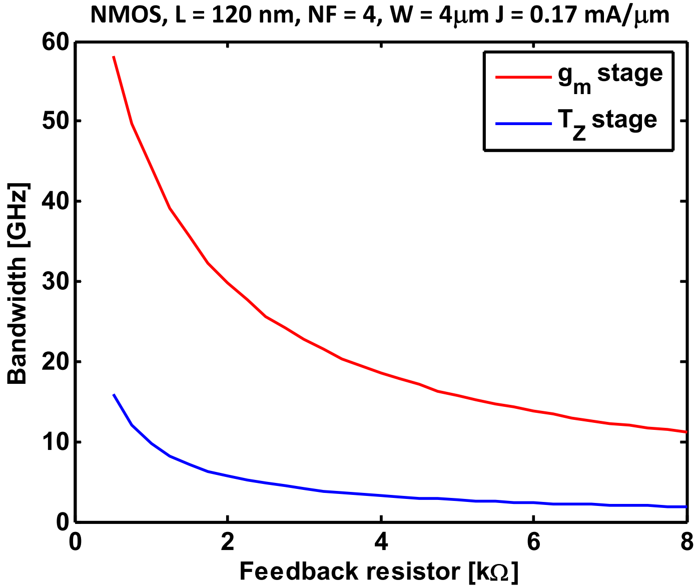
- The Gain-Bandwidth has increased again:
  - CS: 13.7 GHz
  - CS-CD: 63.1 GHz
  - CH: 102.3 GHz
- The frequency response is now dominated by two poles not too distant ($f_1 > f_2/10$)
  - The transfer function displays a second order roll-off
- Like the CS-CD stage the CH stage has some drawbacks:
  - Additional power consumption
  - Not easy to bias either

Paulo.Moreira@cern.ch

BW Broadening: Cherry & Hooper Gain “Stage” (3/3)







R_{LD} = 1 kΩ:

GBW = 79.9 GHz

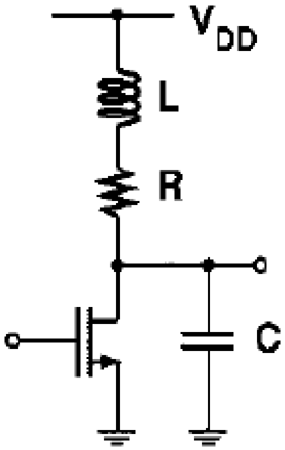
G = 8.2

BW = 9.8 GHz

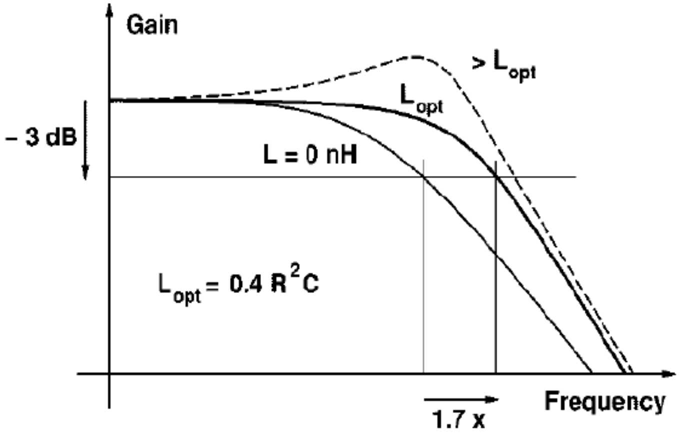
BW Broadening: Inductive Peaking (1/2)

- Add an inductor in series with the load resistor
- The bandwidth can be extended up to 1.85 times
- For optimum group delay response the BW gain is 1.6 times
- The circuit displays a second order transfer function
  - The frequency response is characterized by the ratio **"m"**

$$L = m.R^2.C$$

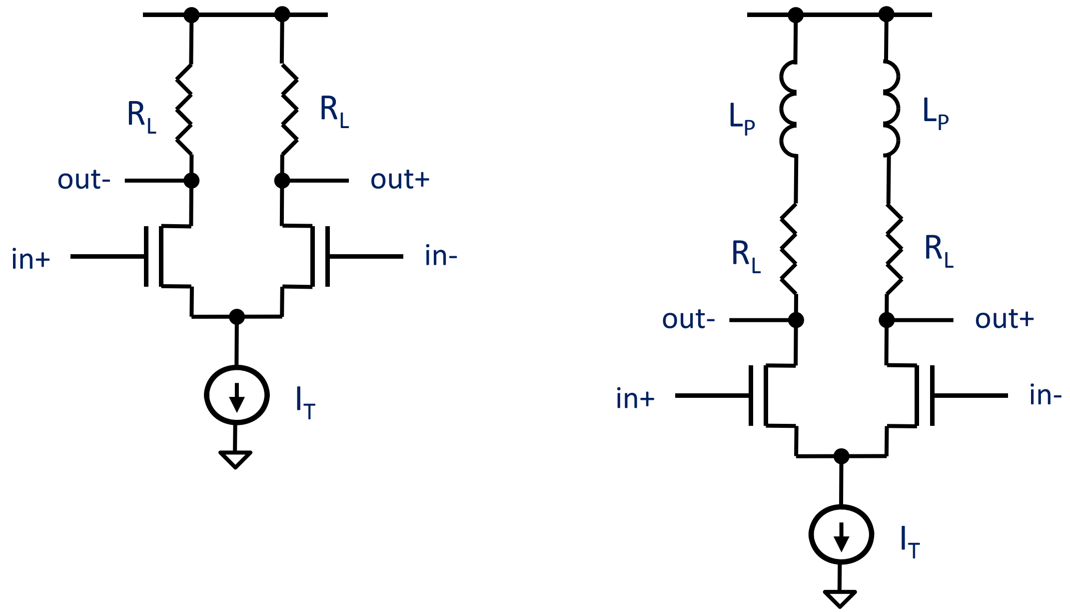


Factor m	Normalized f_{3dB}	Response
0	1.00	No shunt peaking
0.32	1.60	Optimum group delay
0.41	1.72	Maximally flat
0.71	1.85	Maximum bandwidth



Paulo.Moreira@cern.ch

BW Broadening: Inductive Peaking (2/2)

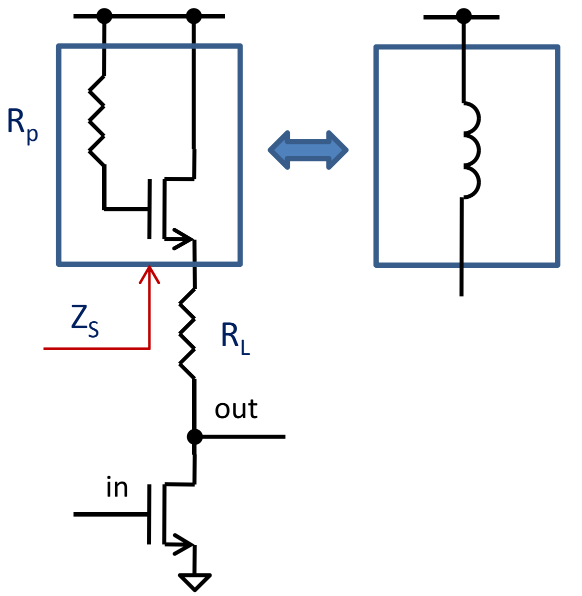


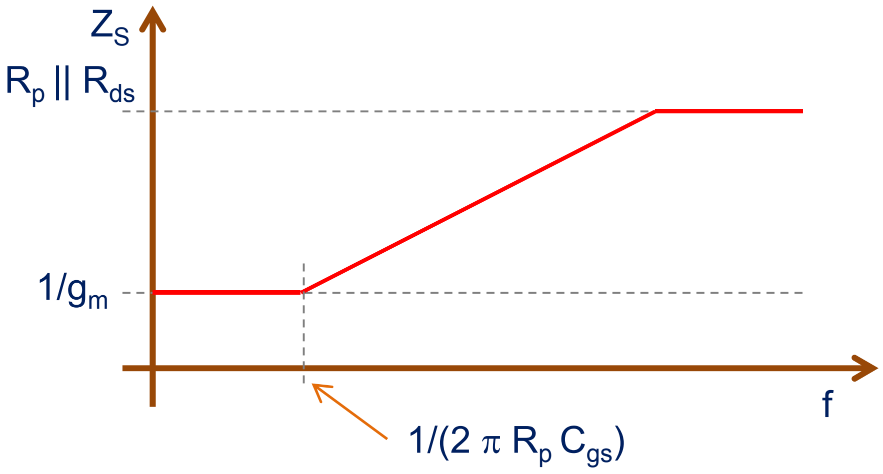
Power optimization:

- For a CML stage the bandwidth can be enhanced by increasing the tail current:
  - By reducing R_L at fixed ΔV
- Or, for by using inductive peaking for a given current and R_L
- It should then be clear that a specified bandwidth can be achieved with lower current if inductive peaking is used.
  - See “active inductive” peaking example later.

BW Broadening: Active “Inductive” Peaking (1/3)

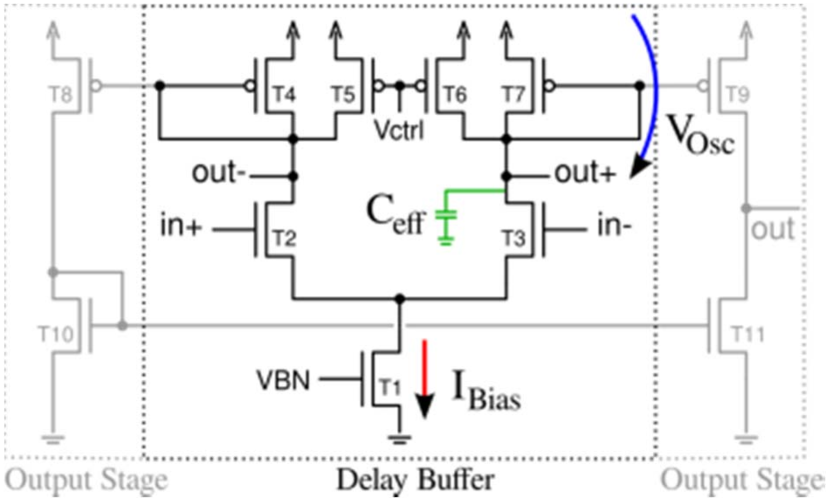
- The main drawback of inductive peaking is the large area required by the peaking inductors
- The impedance “looking” at the source of a Common Gate stage can be “made inductive”
- That is, it has a zero in the transfer function if the gate is “driven” by a resistor (R_p)
- By choosing the value of R_p it is possible to place the zero so that it compensates for the pole in the gain stage
- Bandwidth improvements up to 34% are possible
- It has however some disadvantages:
  - The voltage headroom to process signals is decreased
  - The noise performance is also slightly degraded

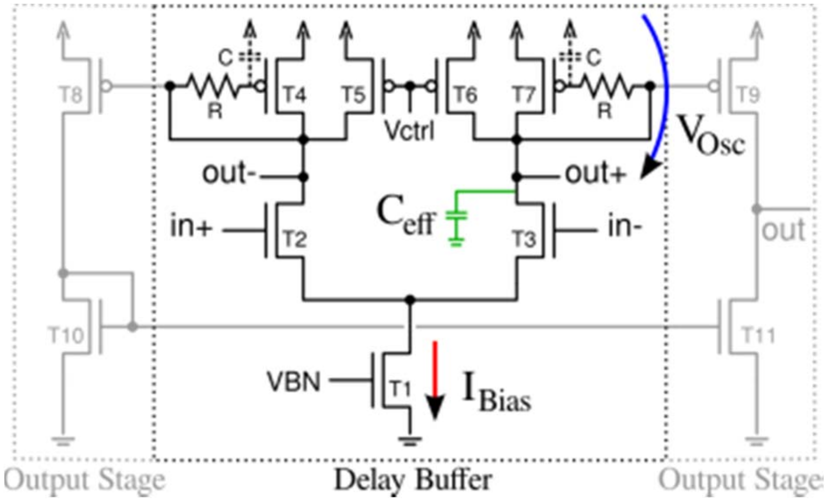


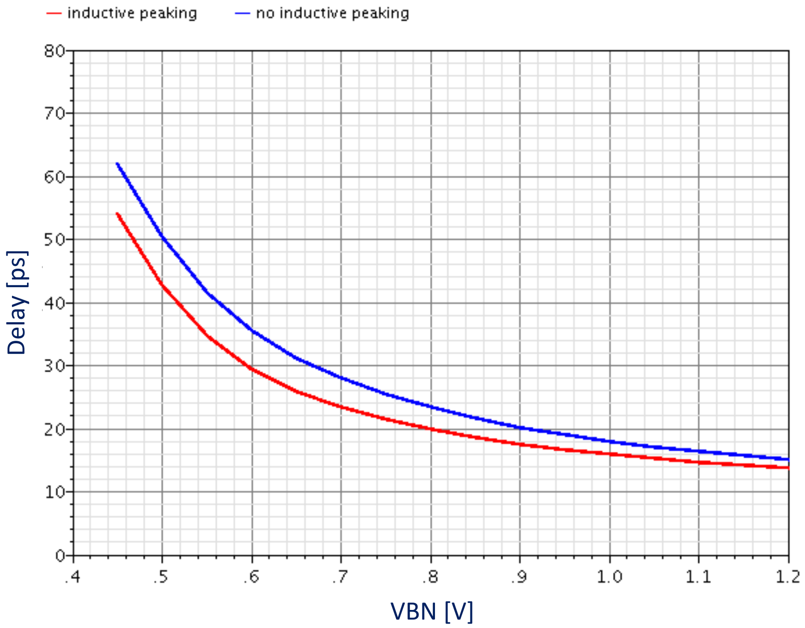


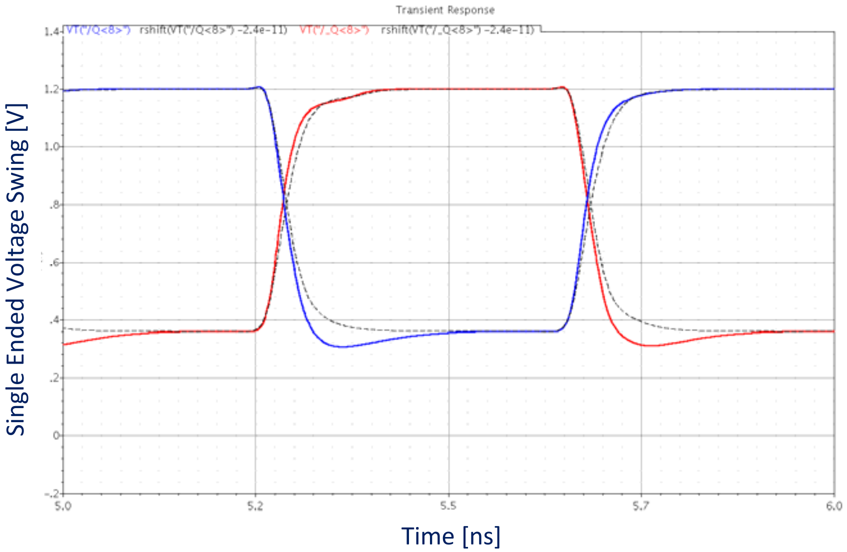
Paulo.Moreira@cern.ch

BW Broadening: Active “Inductive” Peaking (2/3)



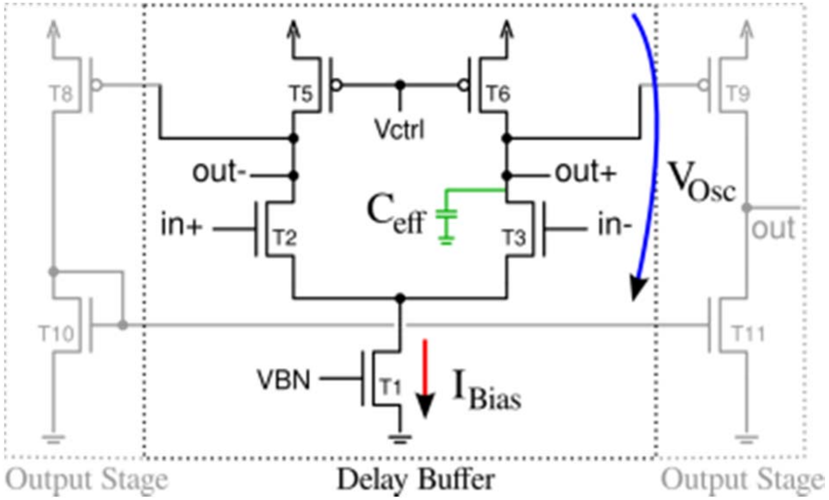


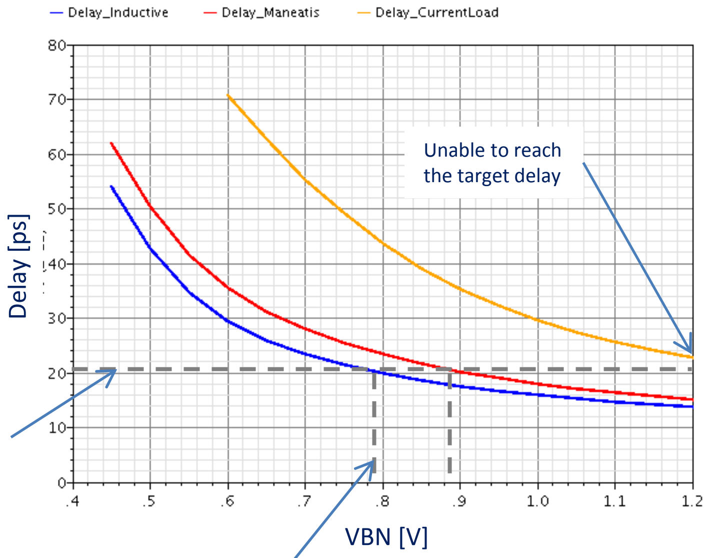




Paulo.Moreira@cern.ch

BW Broadening: Active “Inductive” Peaking (3/3)

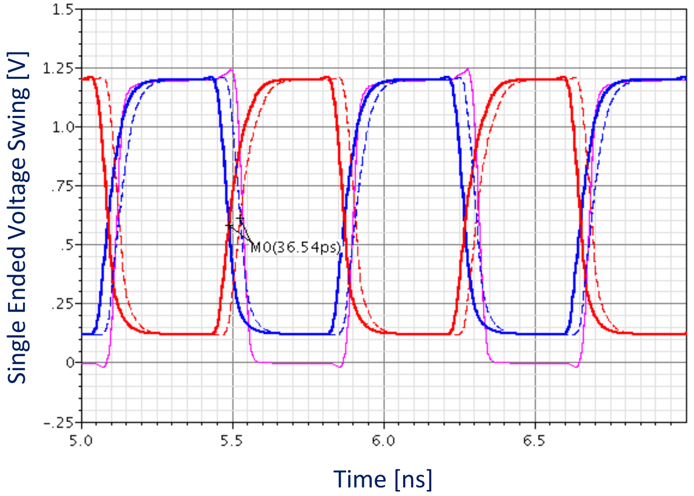




Unable to reach the target delay

Target = 20 ps

For the same delay “inductive” peaking gives a power consumption advantage



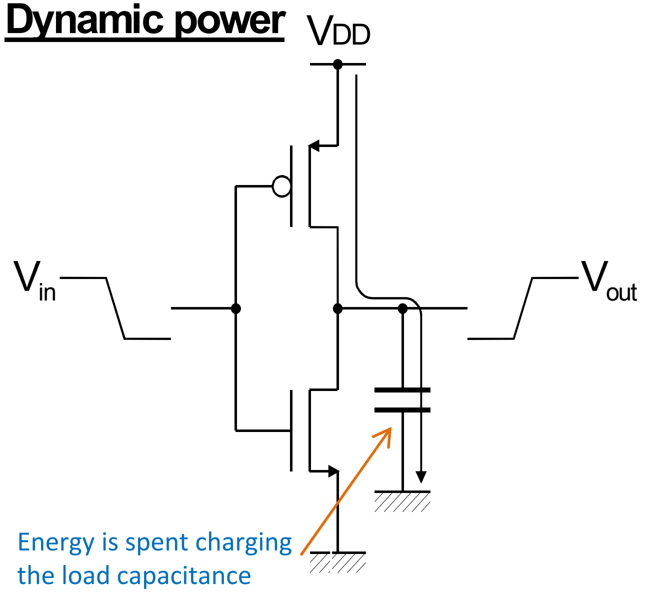
POWER DISSIPATION IN CMOS CIRCUITS

Paulo.Moreira@cern.ch

Power Dissipation and CMOS Circuits

- No matter the complexity of the circuit (logic gate, Flip-Flop, …), a CMOS circuit can be “seen” as:
  - current source and a current sink driving
  - a capacitance (other circuits and interconnects)
- The beauty of CMOS circuits is that:
  - Energy is only dissipated during the “0” → “1” and “1” → “0” transitions.
  - A “standing still circuit” consumes no power!
    - If we ignore leakage currents, of course!

Dynamic power



Energy is spent charging the load capacitance

Transition “0”-to- “1” → Charge is transferred to C: $Q = C \cdot V_{DD}$

Transition “1”-to- “0” → Charge transferred from C to ground

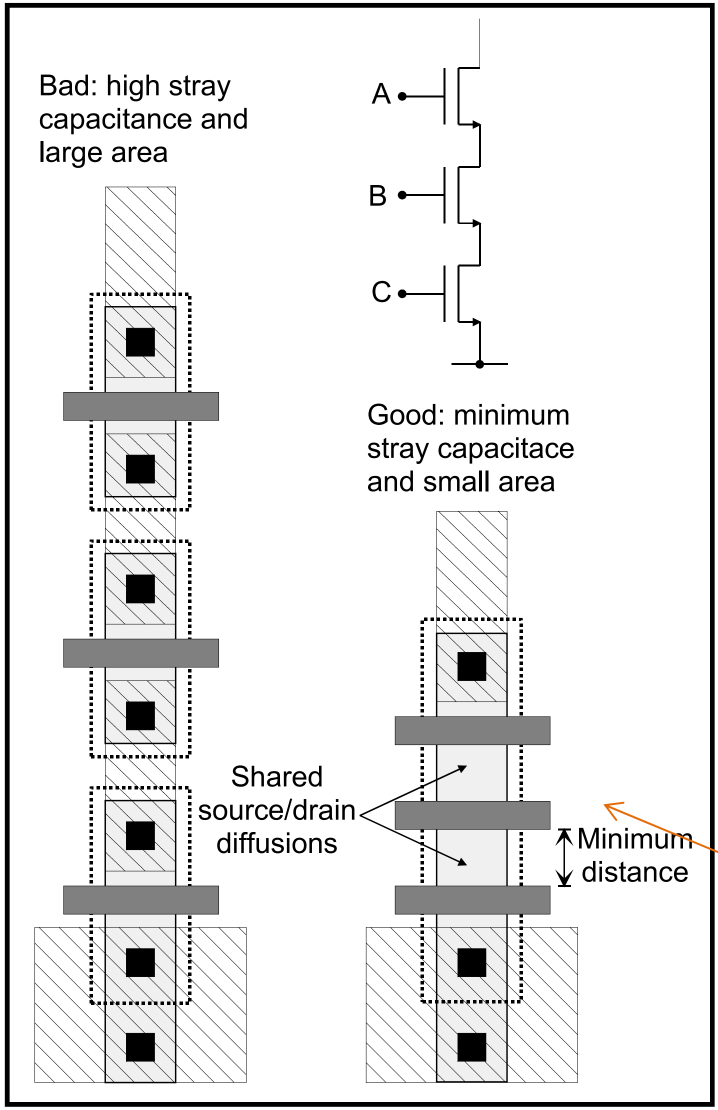
For a “0”-to-“1”-to-“0” cycle the supply current is: $i = f \cdot C \cdot V_{DD}$

The power consumption is thus: $P = i \cdot V_{DD} = f \cdot C \cdot {V_{DD}}^2$

Power dissipation depends on:
- Number of transitions per second (frequency)
- Square of the signal swing (Supply voltage)
- Load capacitance.

Unavoidable:
- High frequency operation is synonymous of
- high power dissipation!

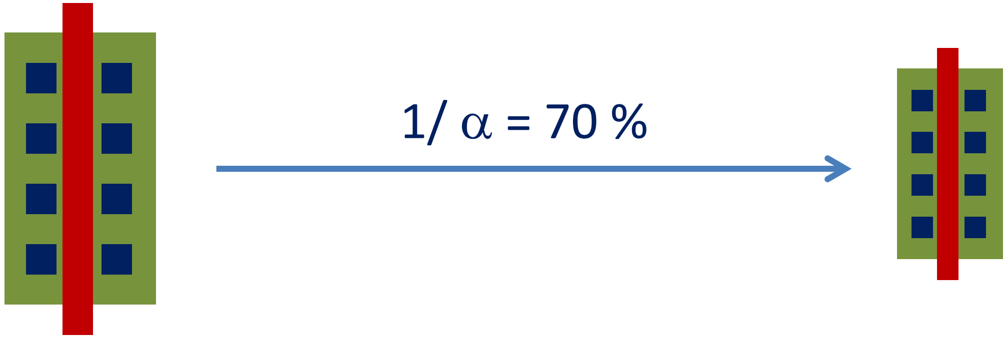
Reducing CMOS Power Dissipation



- High frequency = High power consumption
  - Power consumption is proportional to ***f***
  - Always run the circuits at the lowest useful frequency!
- When a circuit is not "playing a useful role" stop the clock for that circuit or sub-circuit:
  - No transitions "no" power consumption!
- If compatible with the operation speed and noise margins use low supply voltages!
  - Power consumption is proportional to V^2
- Design circuits with "minimum" transistor sizes whenever possible!
  - Power consumption is proportional to ***C***
  - However, make sure the rise/fall times are short to avoid the "short-circuit current"
- Optimize the circuit for low gate count.
- Minimize the interconnect capacitance.
- Optimized the circuit/gate layout to reduce parasitic capacitances.

Can Technology Scaling Help to Keep it Cool?

- Technology scaling has a threefold objective:
  - Increase the transistor density
  - Reduce the gate delay
  - Reduce the power consumption
- Between two technology generations, the objectives are:
  - Doubling of the transistor density
  - Reduction of the gate delay by 30% (43% increase in frequency)
  - Reduction of the power by 50% (at 43% increase in frequency)
- Ideally, CMOS technologies evolution (scaling):
  - All the device dimensions (lateral and vertical) are reduced by $1/\alpha$
  - Concentration densities are increased by α
  - Device voltages reduced by $1/\alpha$
- Typically:
  - $1/\alpha = 0.7$ (30% reduction in the dimensions)
  - $\alpha = 1.41$
- In Practice a lot more than that is going on!



Paulo.Moreira@cern.ch

A “Practical” Case

- Two non-consecutive generations: 130 nm → [90 nm] → 65 nm
  - One generation: $\alpha = 1.43$
  - Two generations: $\beta = \alpha^2 = 2$

$\beta = \alpha \times \alpha$

- The actual scaling between the two technology nodes considered is:
  - $\beta = \beta_L = \beta_W = \alpha^2 = 130\text{ nm} / 65\text{ nm} = 2$
  - $\beta_V = 1.5\text{ V} / 1.2\text{ V} = 1.25$
  - $\beta_{OX} = 3.03\text{ nm} / 2.69\text{ nm} = 1.13$ (Low Power flavour)
  - $\beta_{OX} = 3.03\text{ nm} / 2.00\text{ nm} = 1.52$ (General Purpose flavour)

Deviation from the “ideal” scaling “law”

- Let’s try to understand how much gain in power dissipation can be obtained if a circuit is ported over two generations under the following conditions:
  - Geometries are scaled according to the technology scaling:
    - That is W and L are reduced according to scaling $\alpha = 1.43$ per generation
  - Supply reduced according the practical scaling
  - Frequency:
    - 1st Case: Profit to increase the frequency by a factor of 2
    - 2nd Case: In HEP (LHC) 40 MHz is a kind of “Good given” frequency so let’s keep the operation frequency constant!

Paulo.Moreira@cern.ch

How Much Can We Get From Technology Scaling? (1/2)

CMOS Logic

- Capacitance scaling:
  - $C = \varepsilon_{ox} / t_{ox} \times W \times L$
  - $1 \rightarrow \beta_{ox}/\beta^2 = 0.28$ (LP)
  - $1 \rightarrow \beta_{ox}/\beta^2 = 0.38$ (GP)
- 1^{st} Case: f scales as $1 \rightarrow 2$ (between two nonconsecutive generations, $\beta_f = 0.5$ $(1/\beta_f = 2)$)
  - Ideal
    - Power: $C \times V^2 \times f$
    - One generation scaling: $1 \rightarrow (1/\alpha)^2 = 0.5$
    - Two generations scaling $1 \rightarrow (1/\alpha)^4 = 0.25$
  - Actual two generations
    - $1 \rightarrow \beta_{OX} / (\beta_f \times \beta^2 \times \beta_V^2) = 0.36$ (LP)
    - $1 \rightarrow \beta_{OX} / (\beta_f \times \beta^2 \times \beta_V^2) = 0.49$ (GP)
- 2^{nd} Case: $f \rightarrow 1$ $(\beta_f = 1)$
  - Ideal
    - Power: $C \times V^2 \times 1$
    - One generation scaling: $1 \rightarrow (1/\alpha)^3 = 0.35$
    - Two generations scaling: $1 \rightarrow (1/\alpha)^6 = 0.12$
  - Actual
    - $1 \rightarrow \beta_{OX} /(\beta^2 \times \beta_V^2) = 0.18$ (LP)
    - $1 \rightarrow \beta_{OX} /(\beta^2 \times \beta_V^2) = 0.24$ (GP)

- Interconnect capacitances will prevent such a good results!
- Leakage currents as well (not taken into account here!)

Even though this is a “first order” analysis, it clearly shows that moving up in the technology node gives a clear advantage in terms of power consumption! Within a technology node the choice of “flavour” also makes a difference.

Paulo.Moreira@cern.ch

How Much Can We Get From Technology Scaling? (2/2)

Current Mode Logic

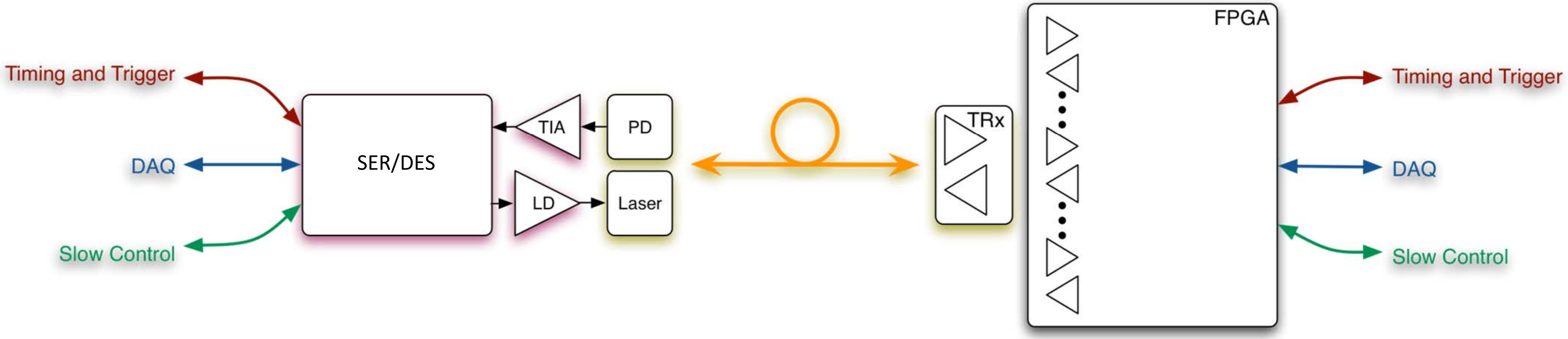
- "Fixed" current buffers (e.g. LVDS, SLVS,…):
  - Current and voltage signal levels stay the same:
    - Specified current and load impedance
  - Only supply voltage scales!
  - Power scaling:
    - Power: $I \times V \rightarrow 1/\beta_V = 0.8$ (LP/GP)
- Fast Current Mode (and high bandwidth "Analogue"):
  - Ideal:
    - Power: $V_{DD} \times W \times J_{opt}$
    - Keeping the current density constant
    - One generation scaling: $1 \rightarrow 1/\alpha^2 = 0.49$
    - Two generations scaling: $1 \rightarrow 1/\alpha^4 = 0.24$
  - Actual:
    - $1 \rightarrow 1/\beta_V \times 1/\beta_W \times 1 = 0.4$ (LP/GP)

Paulo.Moreira@cern.ch

OPTOELECTRONICS
DATA TRANSMISSION LINKS

Paulo.Moreira@cern.ch

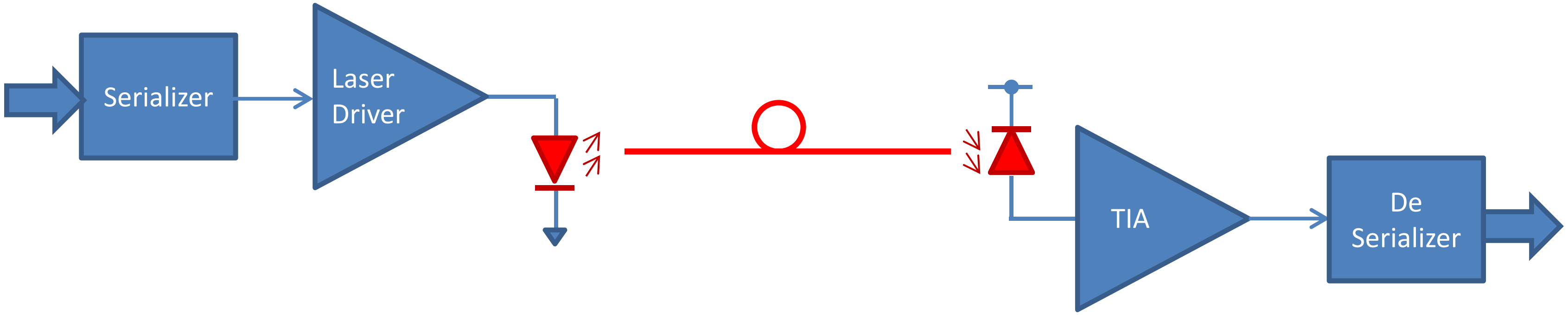
HEP Data Link



On-Detector
Custom Electronics & Packaging
Radiation Hard

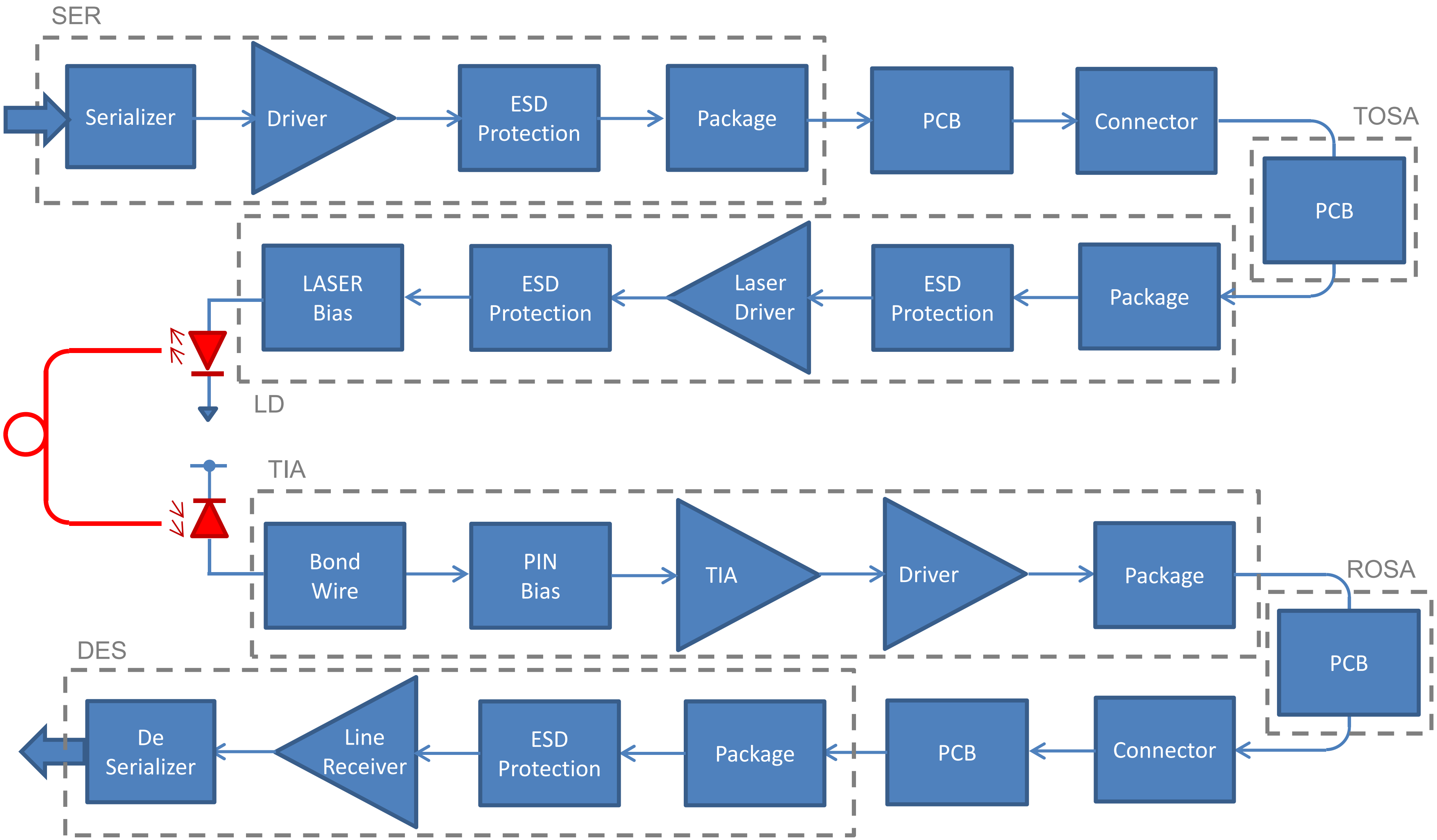
Off-Detector
Commercial Off-The-Shelf (COTS)
Custom Protocol

“Typical” Signal Path for Data Transmission Systems (1/2)



Paulo.Moreira@cern.ch

“Typical” Signal Path for Data Transmission Systems (2/2)



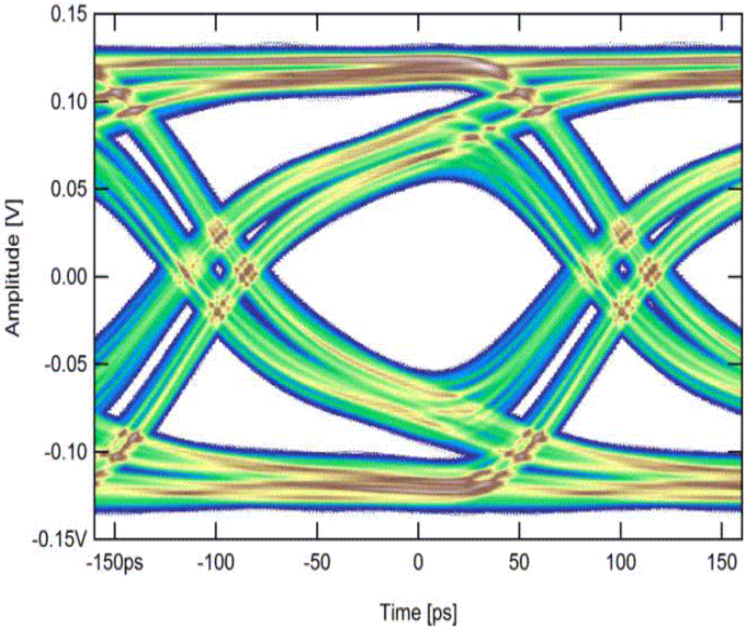
Paulo.Moreira@cern.ch

High Speed Data Link Design

- The design of high speed data links requires careful design and modelling of all the elements in the signal path!
- Don’t overlook “innocent” components:
  - Laser – bias circuit
  - PIN – bias circuit
  - ESD protection
  - Package
  - Connectors
  - PCB
- It is not enough to have high performance optoelectronics and ASICs!
- Equally important are:
  - Package selection / design
  - PCB design
  - Connectors
- In your simulation test benches make sure to use realistic signal sources and loads:
  - At high frequencies there is no such a thing as a ‘zero’ impedance voltage source or ‘infinite’ impedance current source.

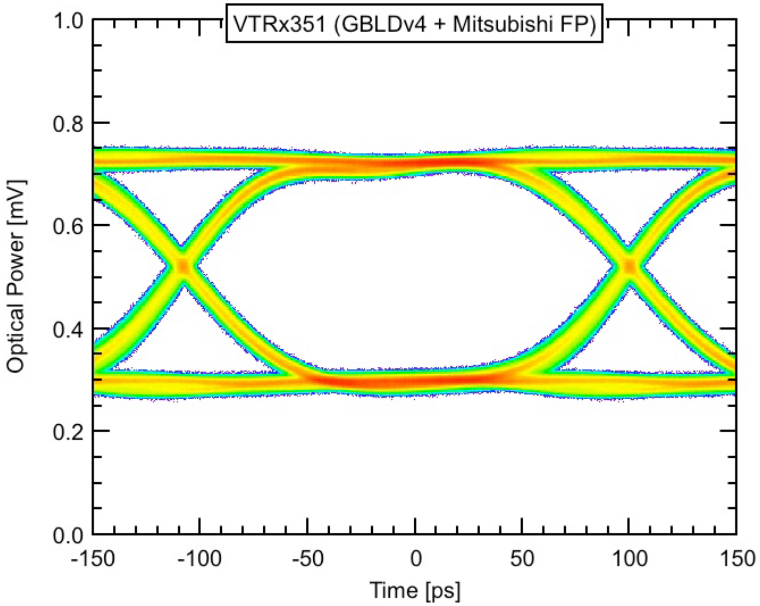
Paulo.Moreira@cern.ch

Your Simulation Test Bench is Important!



- Source impedance inaccurately modelled:
  - ESD impact unrevealed by simulations!
- Jitter independent of the pre-emphasis settings:
  - Bandwidth limitation not in the output node
  - Either in an internal or input node
- Once the modelling problem was detected:
  - Simulations accurately reproduce the response
  - Input identified as the BW limiting node
- "RC" time constant of the source impedance plus ESD protection capacitance dominated the response!
  - An input matching network was clearly needed!

- Once an input matching network was included in the ASIC the circuit revealed excellent performance!



ESD PROTECTION CIRCUITS

Introduction

What is ESD Protection?

- Electrostatic Discharge (ESD) event!
- ESD events occur during:
  - Handling by humans:
    - Human Body Model (HBM)
  - Handling by machines:
    - Machine Model (MM)
    - Charged Device Model (CDM)
- All pins must have ESD protection!
- Including:
  - High speed
  - Sensitive analog
- Failing to protect against ESD will most likely result in poor yield
- ESD protection must be included early in the design phase:
  - Its impact on circuit performance must be considered

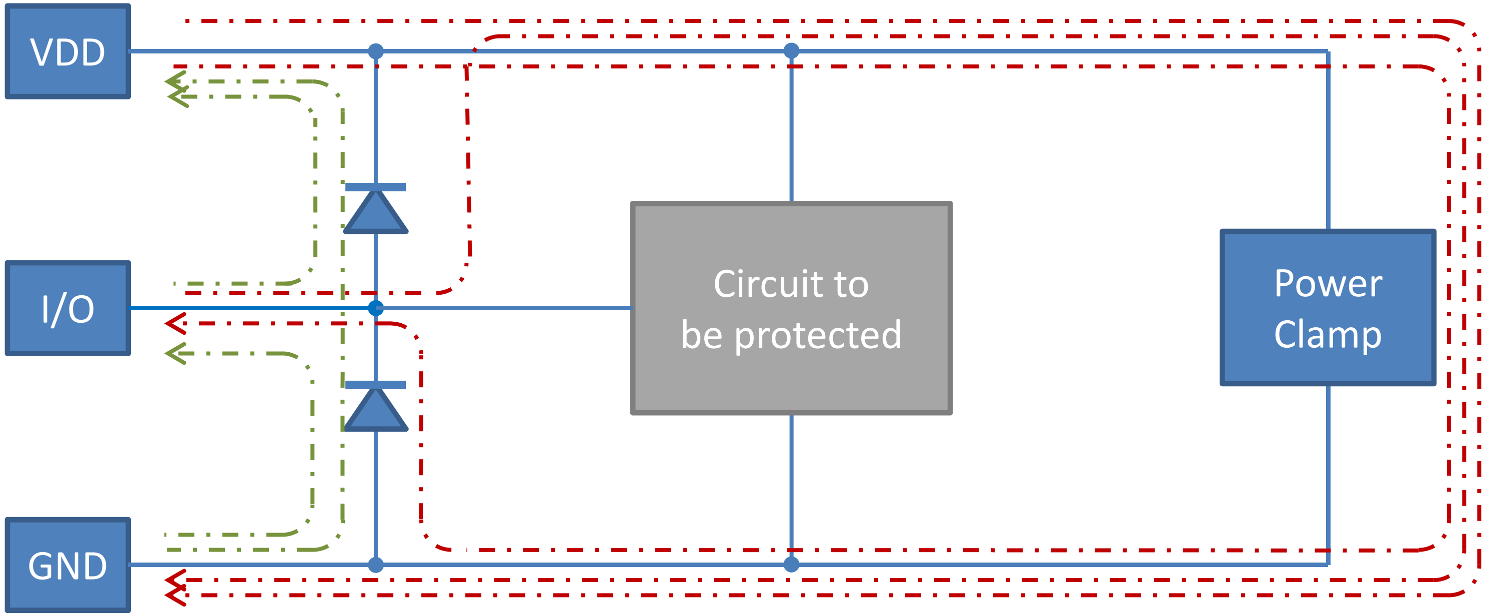
Paulo.Moreira@cern.ch

ESD Models

Characteristics	HBM	MM	CDM
Equivalent circuit	Series 1.5 kΩ + 100 pF	Series 0.5-1.0 μH + 200 pF	Field plate to chip capacitance only
Test voltage	2000 V	200 V	500 V
Discharge Path	Between ANY two pins	Between ANY two pins	One pin only (Discharge Pin)
Simulates	Human discharging through chip	Metal tool discharging through the chip	Charged chip discharging to ground
Discharge Waveform	Exponential Decay Time Constant = 150 ns	11-16 MHz damped Oscillation	~1 GHz damped Oscillation
On-Chip stress characteristics	Lowest Current and Voltage, longest duration I(HBM)=V(HBM)/1.5 kΩ	Intermediate Current, Voltage and duration	Highest Current, Voltage and shortest duration
Failure Mechanisms	Thermal failures: MOSFET snapback, Interconnect fusing, gate oxide damage	Junction damage, Interconnect fusing, gate oxide damage	Gate oxide damage and interconnect damage

Paulo.Moreira@cern.ch

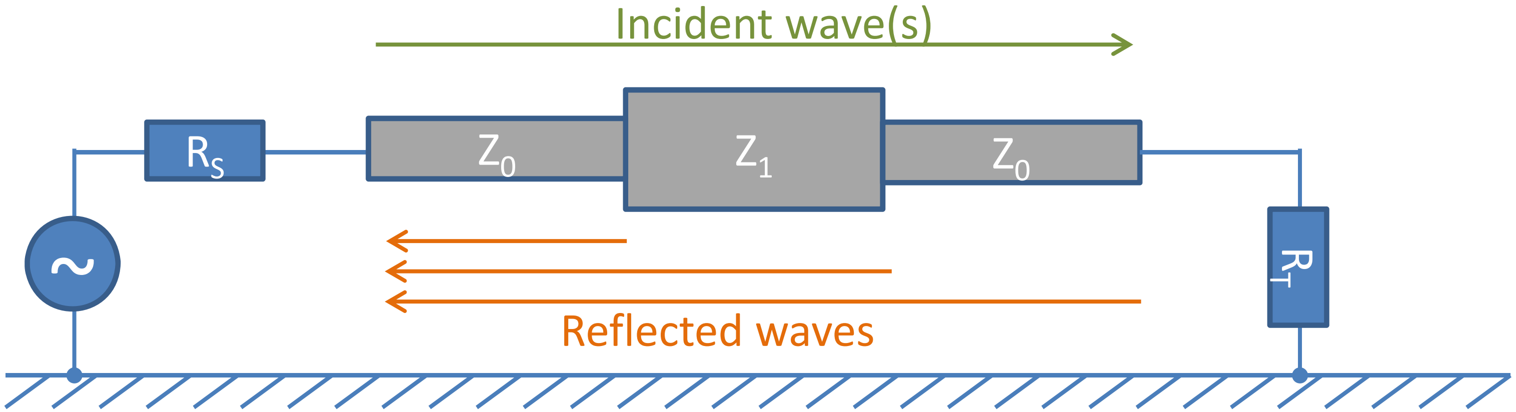
ESD Protection Circuits



- ESD voltages can exceed the breakdown voltage of any structure on the ASIC
- The current will flow through the path(s) of least resistance
- ESD protection provides a preferred discharge current path designed to carry high currents
- The discharge path must carry a high current while developing a low voltage drop or the current will be re-routed through sensitive circuits
- To be effective (low impedance) diodes have to have relatively large areas and thus relatively large capacitances:
  - Example of HBM protection diodes:
    - 2 x 50 µm long, 0.72 µm wide diodes to ground
    - 4 x 50 µm long, 0.72 µm wide diodes to supply
    - Equivalent parasitic capacitance: ~400 fF

Paulo.Moreira@cern.ch

Why Does it Matter?



- For multi Gb/s systems the wave propagation nature of electrical signals needs to be considered!
- Waves travel along transmission lines and are partially reflected at points where the characteristic impedance changes:
  - Transmission lines are intrinsically bidirectional
  - The important quantity is power (and not voltage and current independently)!
- Waves are fully absorbed (no reflection) at the termination if the termination has the same impedance of the line!
- Otherwise, waves are partially reflected:
  - $\Phi = 0^\circ$ if $R_{TERM} > Z_0$
  - $\Phi = 180^\circ$ if $R_{TERM} < Z_0$

Why Does it Matter?

- For purposes of signal propagation, the ESD circuit can be seen as a capacitor shorting the termination impedance
- The transmission line termination is thus frequency dependent and can't match the line impedance at all frequencies
- The incoming wave will be partially reflected over a range of frequencies
- The ratio between the reflected and incident wave amplitudes is the reflection coefficient:

$$\Gamma = \frac{V_r}{V_i} \qquad [1]$$

- The reflection coefficient is related with the impedance of the line and of termination:

$$\Gamma = \frac{Z_L - Z_0}{Z_L + Z_0} \qquad [2]$$

- Or to the return loss (the ratio between the incident and reflected powers)

$$RL(dB) = -20\log|\Gamma| = 10\log\frac{P_i}{P_r} \qquad [3]$$

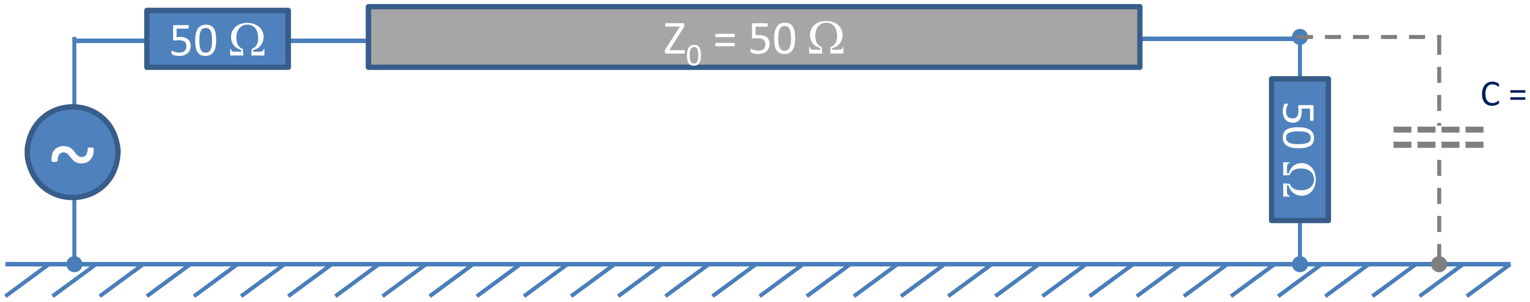
- The higher the return loss the better the load is matched to the source!
  - The name choice was unfortunate!

Paulo.Moreira@cern.ch

Why Does it Matter?

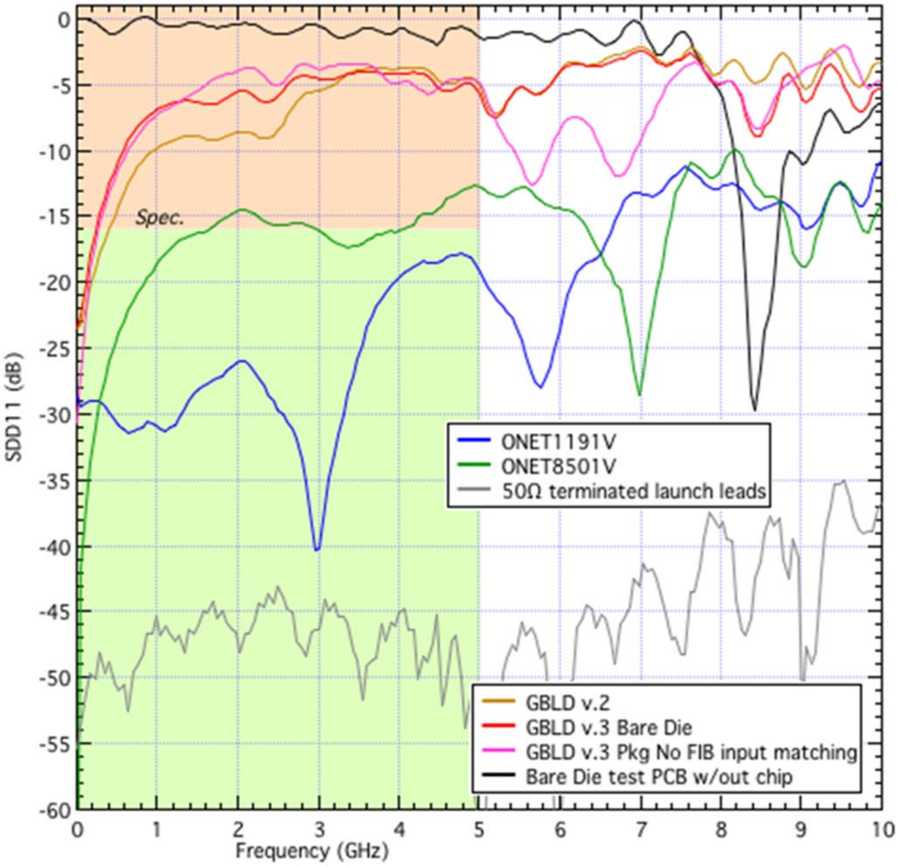
- Example:
  - Signal frequency: 5 GHz
  - 50 Ω transmission line
  - Terminated by a 50 Ω resistor “shorted” by 1 pF parasitic capacitance
  - Capacitor impedance at 5 GHz:
    - 31.8 Ω
  - Termination impedance at 5 GHz (parallel of R and C):
    - 19.4 Ω
  - Reflection coefficient at 5 GHz:
    - \- 0.44
    - The reflected wave has 44% of the amplitude of the incident wave and an 180° phase shift!
  - Return loss at 5 GHz:
    - 7.13 dB
    - 19.3% of the incident power is reflected towards the load.

(we are ignoring here that these are all complex numbers)



Paulo.Moreira@cern.ch

A Real Example



- The on-chip termination resistor is in parallel with:
  - A relatively large capacitance originating from the ESD devices
  - The input stage capacitance
- The ESD capacitance becomes a short at high frequencies
- The return loss does not fulfil the specification (> 16 dB)
- Add a matching network to improve the return loss
- Two methods will be discussed:
  1. Coupled T-Coils
  2. “Inductive compensation”

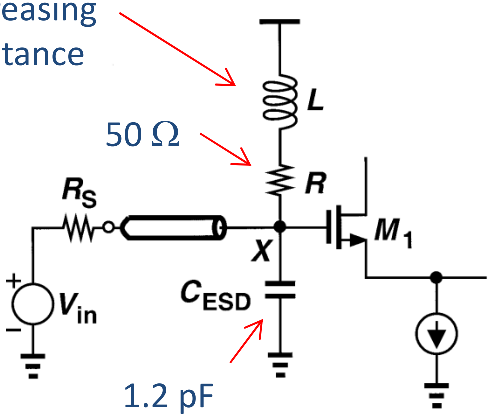
ESD PROTECTION CIRCUITS

Input Matching: Coupled T-Coils

Load C_{ESD} “Compensation”

S. Galal and B. Razavi - JSSC vol. 38, no. 12, p. 2334, Dec. 2003

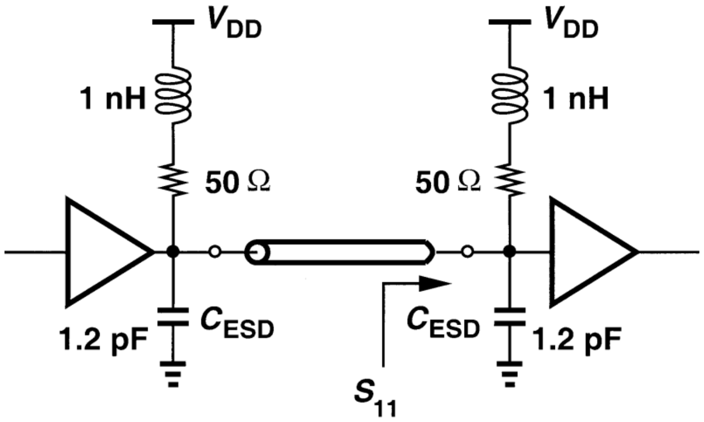
Effectively increases the “load impedance” to try to compensate for the decreasing impedance due to the ESD capacitance (both are electrically in parallel)

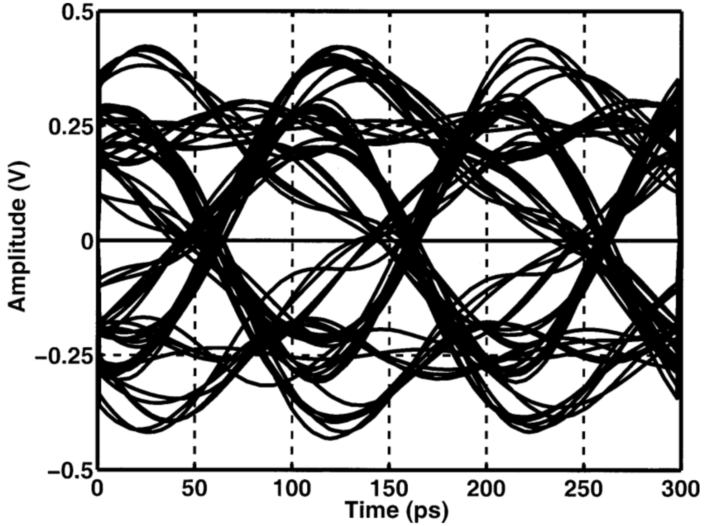


50 Ω

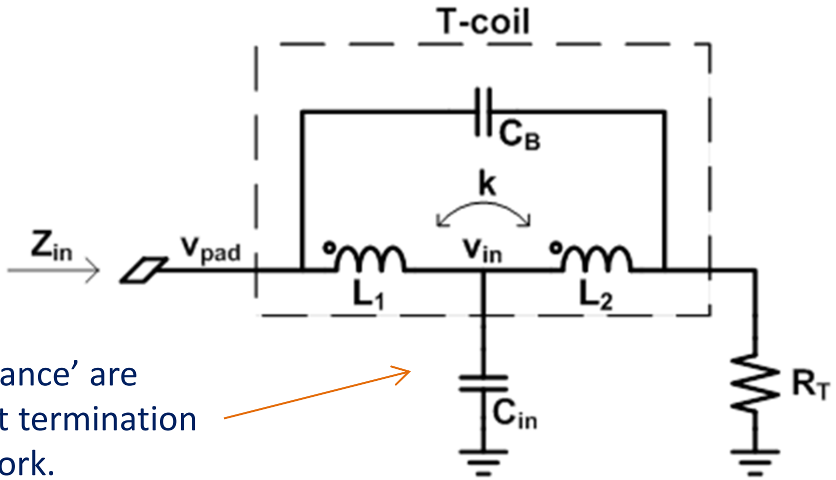
10 Gb/s

1.2 pF





The T-coil Concept

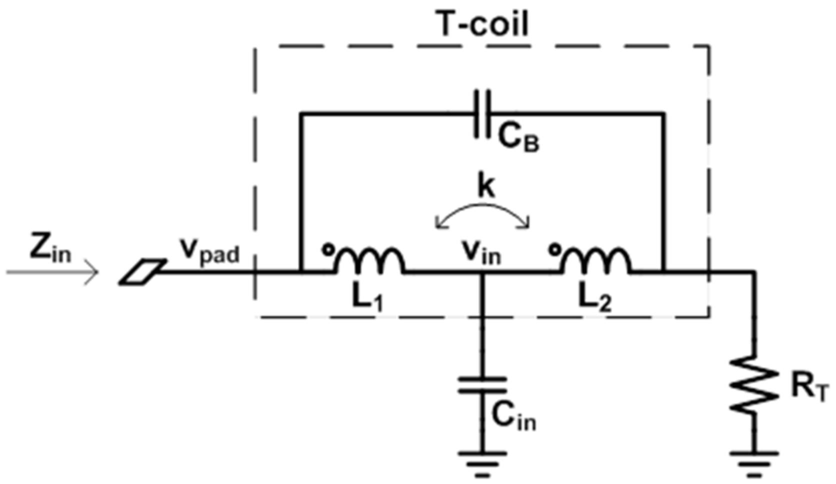


The input and ESD ‘capacitance’ are “separated” from the input termination resistor by the T-Coil network.

- At very low frequencies:
  - $Z_{in} = R_T$ because the inductors act as a short and the capacitors are “open”
- At very high frequencies:
  - $Z_{in} = R_T$ because C_B acts as a short and the inductors are “open”
- $Z_{in} = R_T$ can be guaranteed for all intermediate frequencies by properly choosing:
  - L_1, L_2, k and C_B,.
- In this case, C_{in} (including the ESD capacitance) never influences the overall input impedance so that the return loss ideally is infinite.

Paulo.Moreira@cern.ch

T-Coil Input Impedance (1/2)



Frequency dependent term!

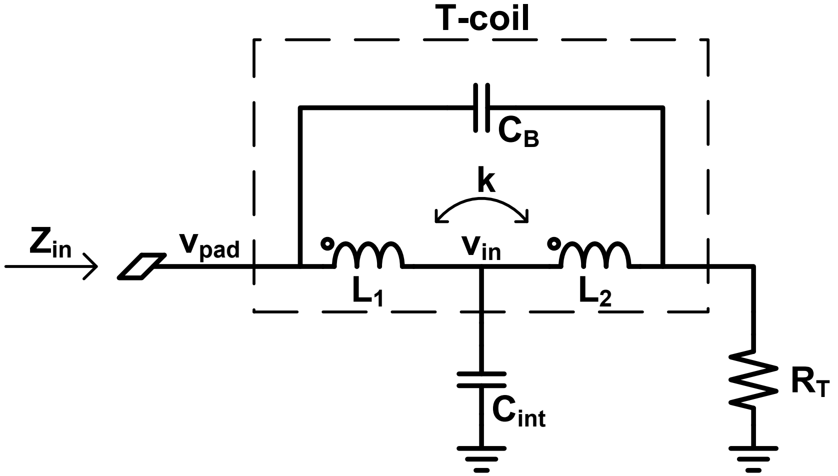
Good matching if = "1"
at all frequencies

Make the "s" coefficients identical
in the numerator and denominator

Termination

$$Z_{in} = R_T \cdot \boxed{\frac{1 + s\cdot\left(\frac{L_1 + L_2 + 2k\sqrt{L_1 L_2}}{R_T}\right) + s^2\cdot\left(C_B\left(L_1 + L_2 + 2k\sqrt{L_1 L_2}\right) + C_{in}L_1\right) + s^3\cdot\frac{C_{in}L_1L_2}{R_T}\left(1-k^2\right) + s^4\cdot C_B C_{in} L_1 L_2\left(1-k^2\right)}{1 + s\cdot R_T C_{in} + s^2\cdot\left(C_B\left(L_1 + L_2 + 2k\sqrt{L_1 L_2}\right) + C_{in}L_2\right) + s^3\cdot R_T C_B C_{in}\left(L_1 + L_2 + 2k\sqrt{L_1 L_2}\right) + s^4\cdot C_B C_{in} L_1 L_2\left(1-k^2\right)}}$$

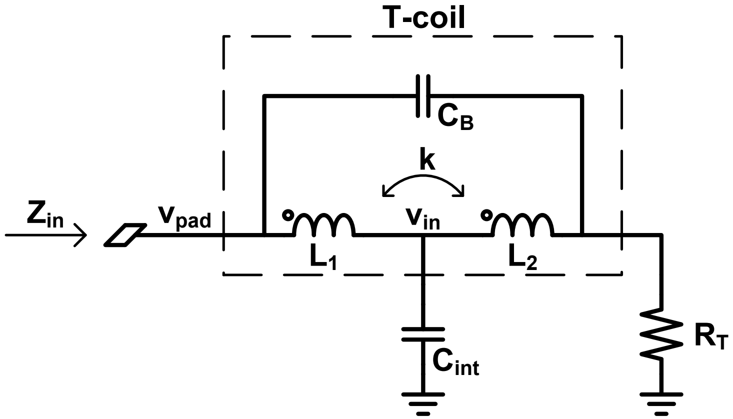
T-Coil Input Impedance (2/2)



$$Z_{in} = R_T \Leftrightarrow \begin{cases} L_1 = L_2 \\ L_1 = \dfrac{R_T C_{in}}{2(1+k)} \\ C_B = \dfrac{C_{in}}{4} \dfrac{1-k^2}{(1+k)^2} \end{cases} \rightarrow \text{choose } L_1 \text{ or } k$$

Paulo.Moreira@cern.ch

T-Coil Transfer Function

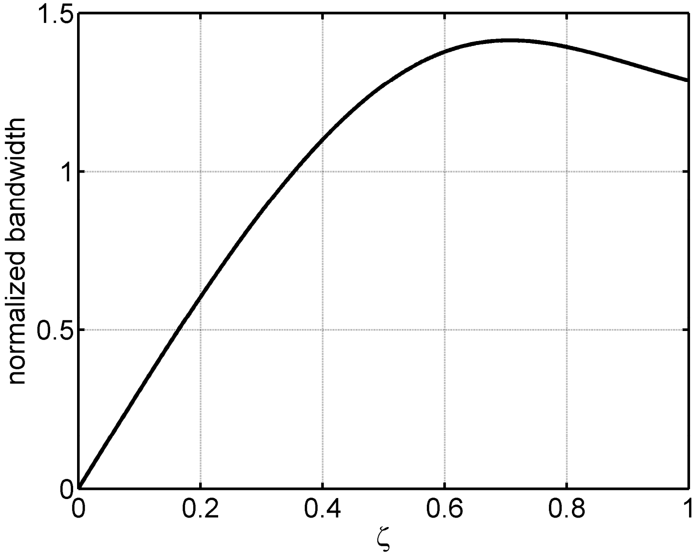


$$\frac{v_{in}}{v_{pad}} = \frac{1}{1 + s\cdot\frac{R_T C_{in}}{2} + s^2\cdot\frac{R_T^2 C_{in}^2}{4}\cdot\frac{1-k}{1+k}}$$

$$= \frac{\omega_n^2}{\omega_n^2 + 2\zeta\omega_n\cdot s + s^2}$$

$$\text{where}\begin{cases}\omega_n = \dfrac{2}{R_T C_{in}}\cdot\sqrt{\dfrac{1+k}{1-k}}\\[2ex] \zeta = \dfrac{1}{2}\cdot\sqrt{\dfrac{1+k}{1-k}}\end{cases}$$

T-coil bandwidth normalized to a simple 50 Ω termination



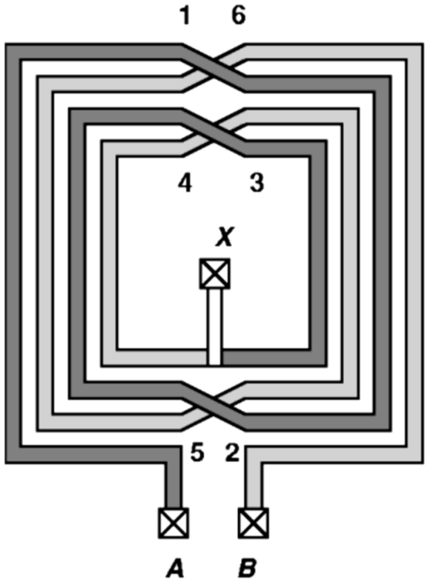
$\zeta = 0.7$ for maximum bandwidth

$\zeta = \frac{\sqrt{3}}{2} \approx 0.866$ for uniform group delay

Paulo.Moreira@cern.ch

T-Coil Implementation

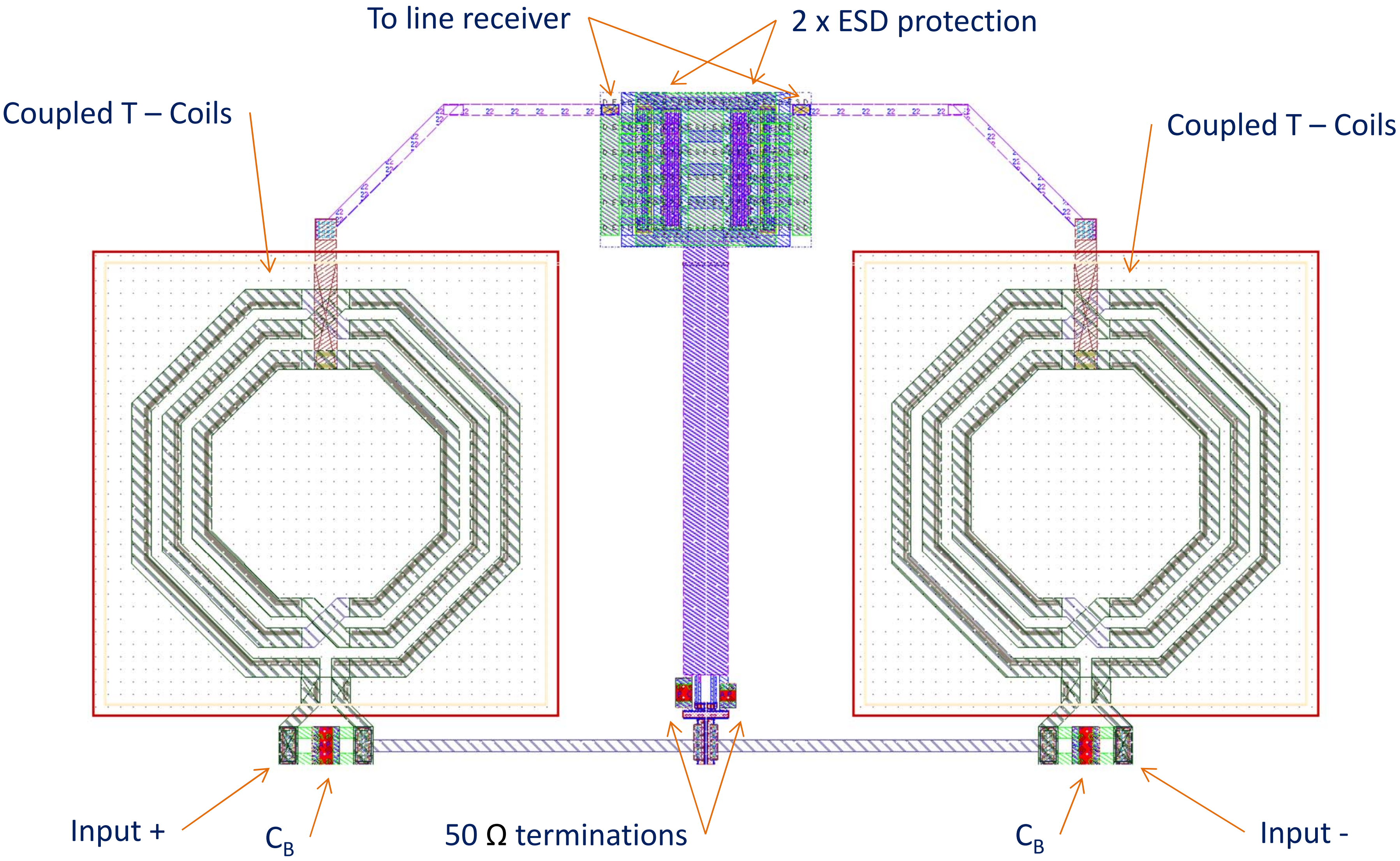
- Coupled inductors are implemented by means of ‘symmetrical inductors’ provided by the design kit
- Coupling factor is adjusted by choosing an appropriate spacing between the turns
- Modeling:
  - Design kit parameterized models
  - LRCK extraction over the inductor region
- Symmetrical inductor:
  - outer diameter: 173 µm
  - coil width: 8.5 µm
  - 3 turns
  - spacing: 5 µm
- Minimal HBM ESD protection is used to keep the input bandwidth high
- Input:
  - Differential 100 Ω terminated
  - Termination implemented by two 50 Ω resistors in series with the mid point to the common-mode voltage



Design parameters	
R_T	50 Ω
C_{IN}	621 fF
L_1	545 pH
L_2	545 pH
k	0.44
C_B	63.2 fF

Paulo.Moreira@cern.ch

Input Network Layout



To line receiver

2 x ESD protection

Coupled T – Coils

Coupled T – Coils

Input +

C_B

50 Ω terminations

C_B

Input -

Paulo.Moreira@cern.ch

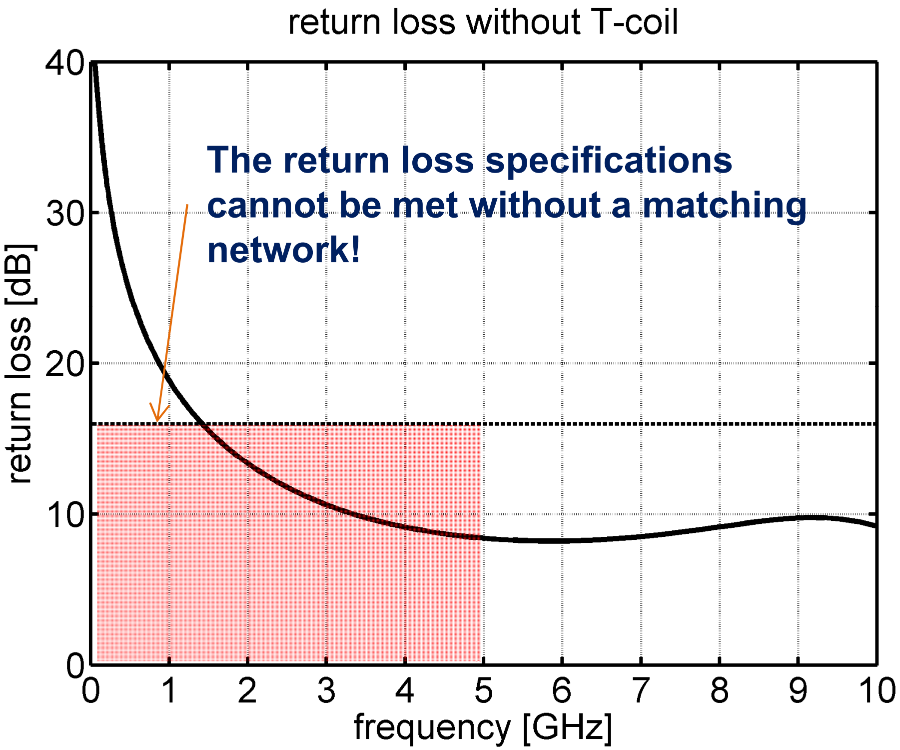
Return Loss Calculation (1/2)

Specification:

- **RL > 16 dB**
- **0 < f < 5 GHz**

RL = 16 dB:

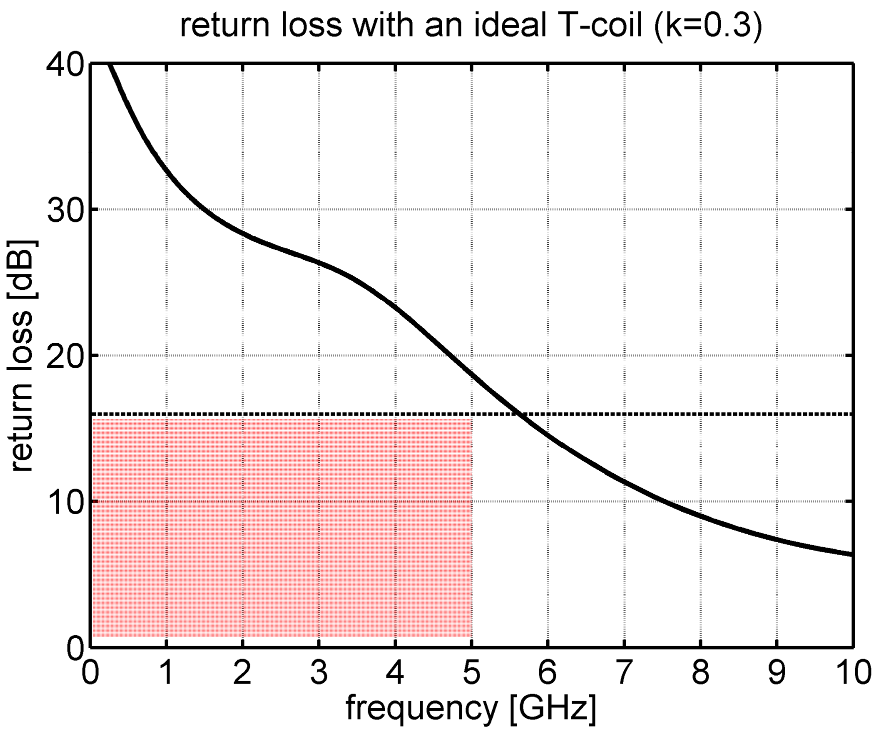
- **2.5% of the power is reflected**
- **Reflected wave has an amplitude which is 15.8% of the incident wave.**

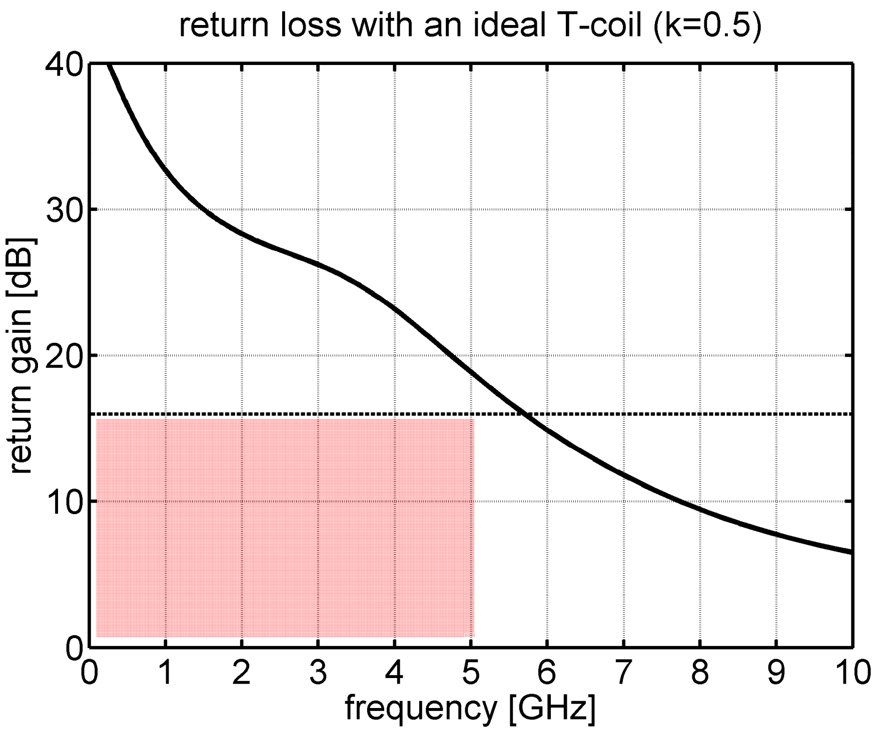


The return loss specifications cannot be met without a matching network!

Paulo.Moreira@cern.ch

Return Loss Calculation (2/2)



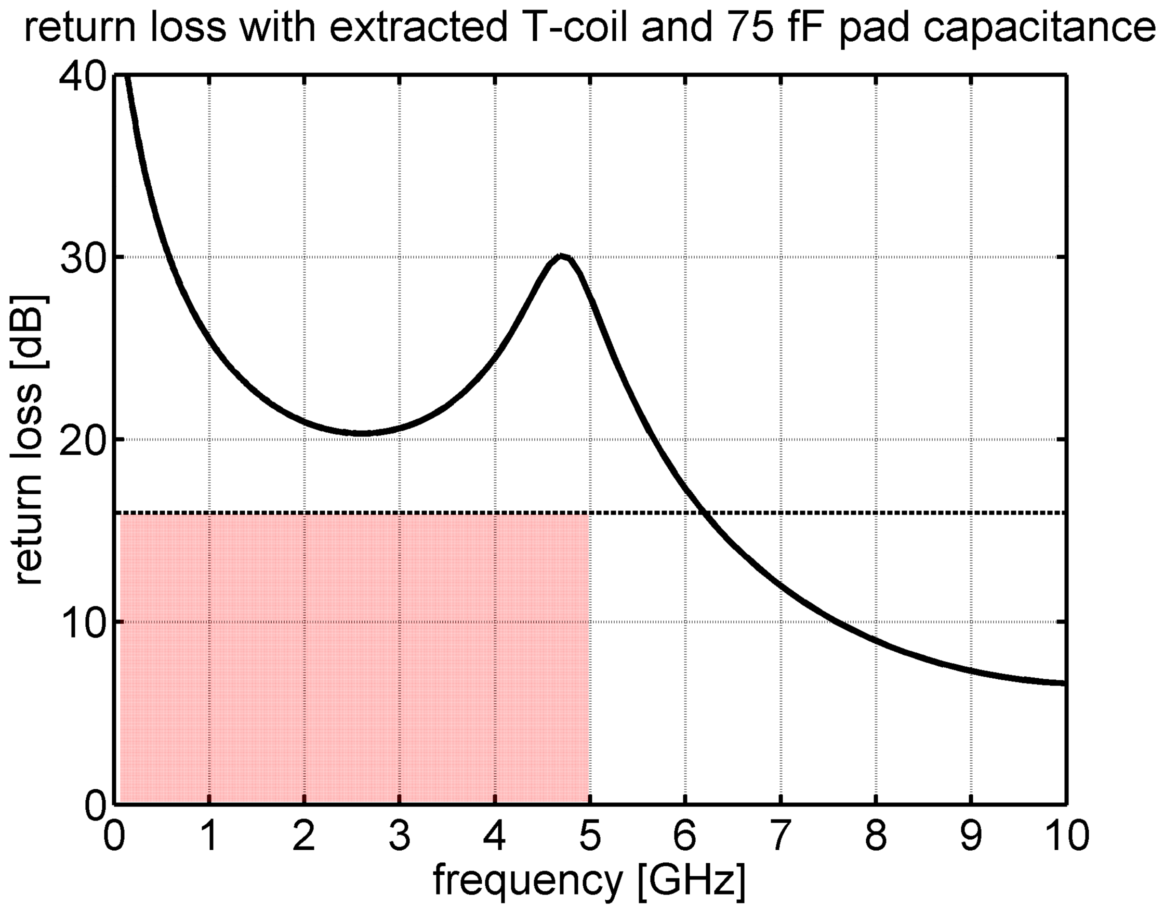


- The return loss specifications can be met with “insertion” of a T-coil.
- The actual value of the coupling factor is not really important as long as the other parameters are changed accordingly.

Paulo.Moreira@cern.ch

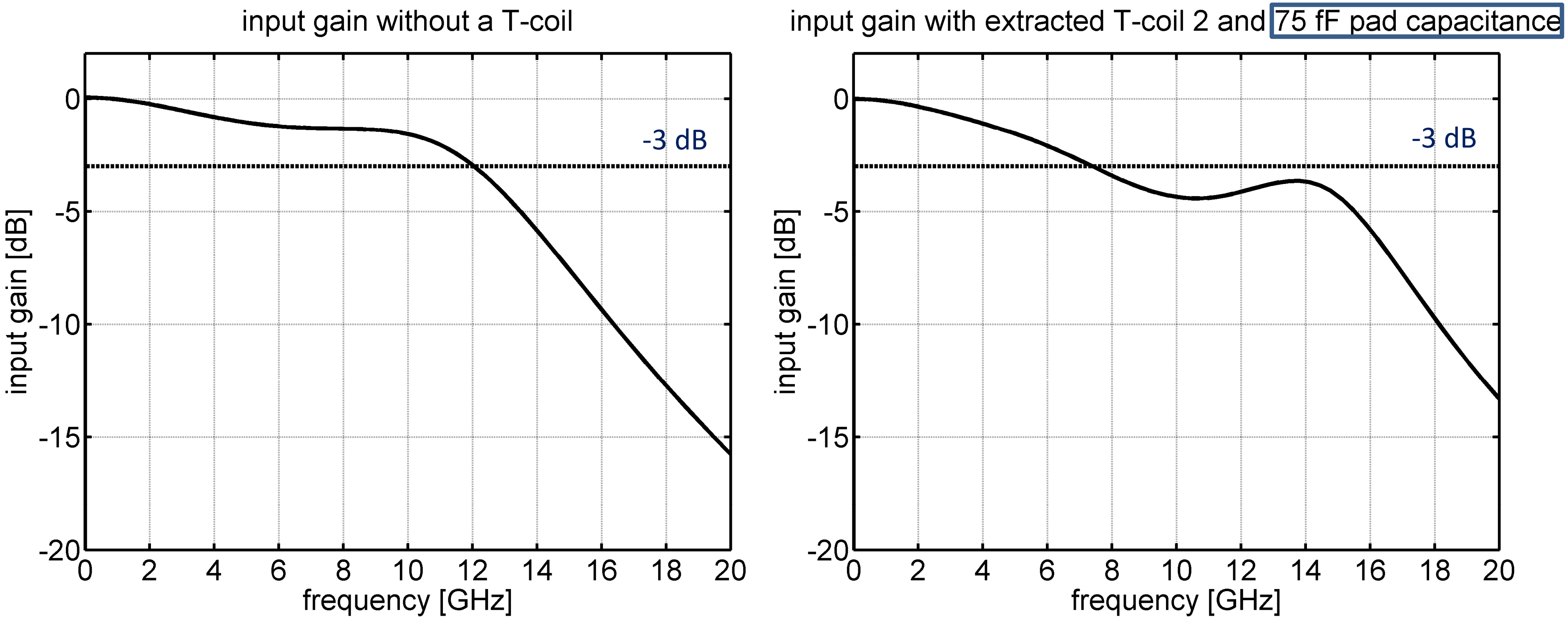
Return Loss Simulation

"Locally" RLCK extraction



Paulo.Moreira@cern.ch

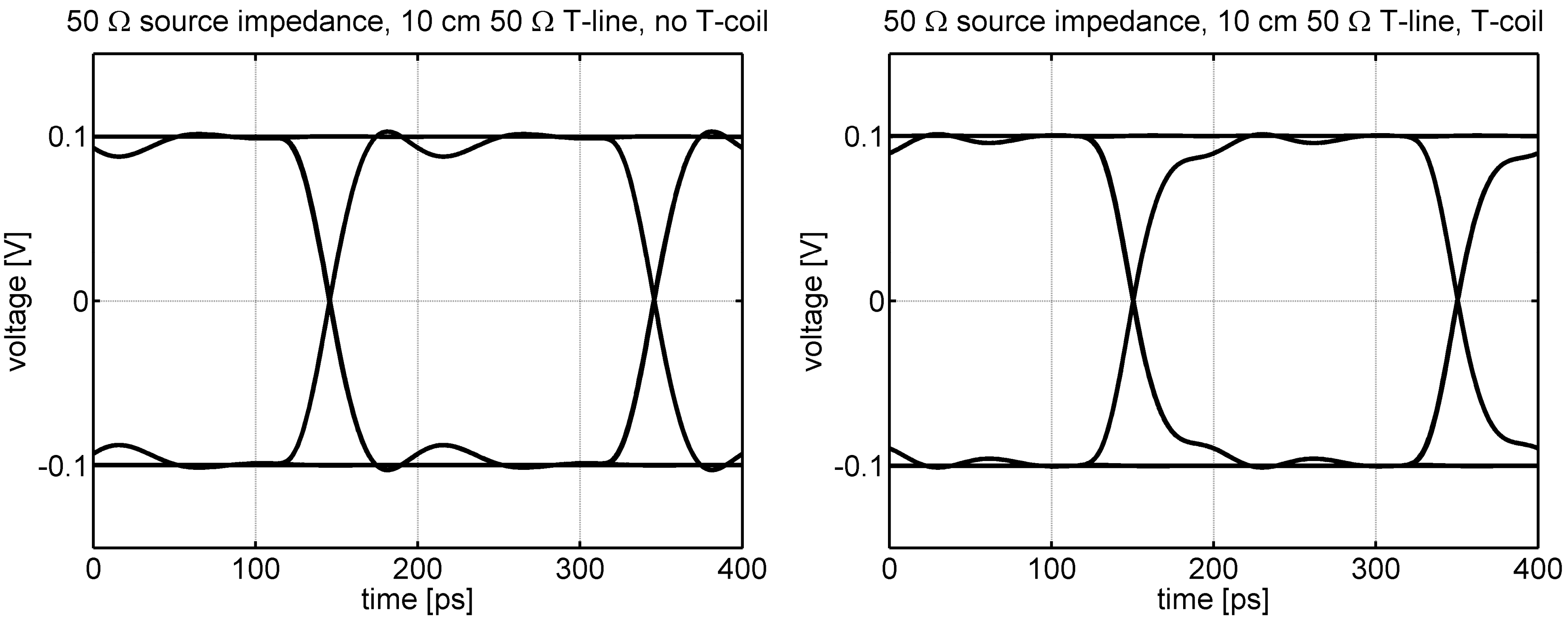
Input Transfer Function Simulation



- The input gain (transfer function from the source to the input gates) is smaller with a T-coil (12 GHz vs. 7 GHz)
- However, for the 5 Gbit/s intended data rate, 7 GHz is more than enough.

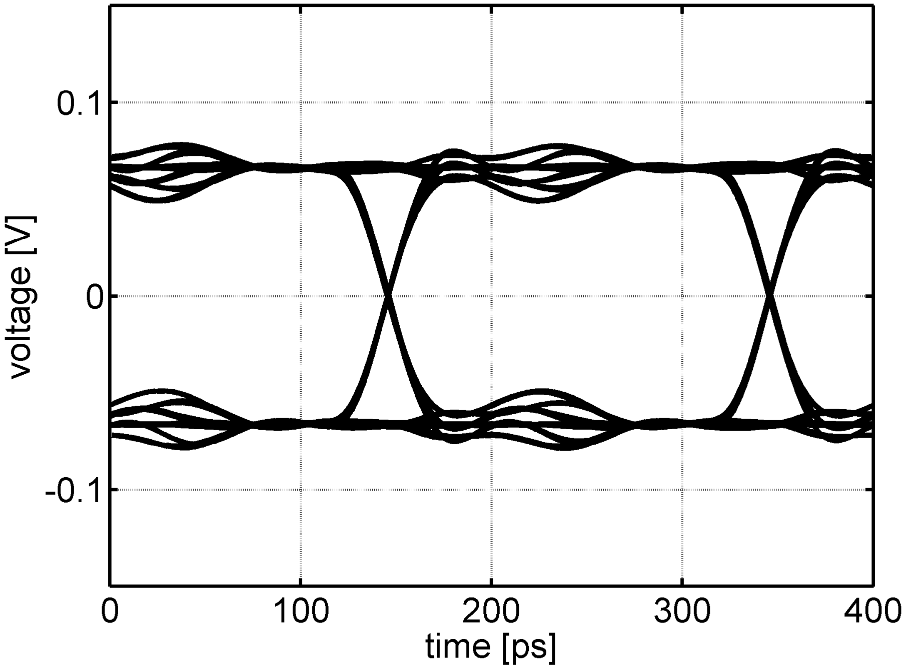
Paulo.Moreira@cern.ch

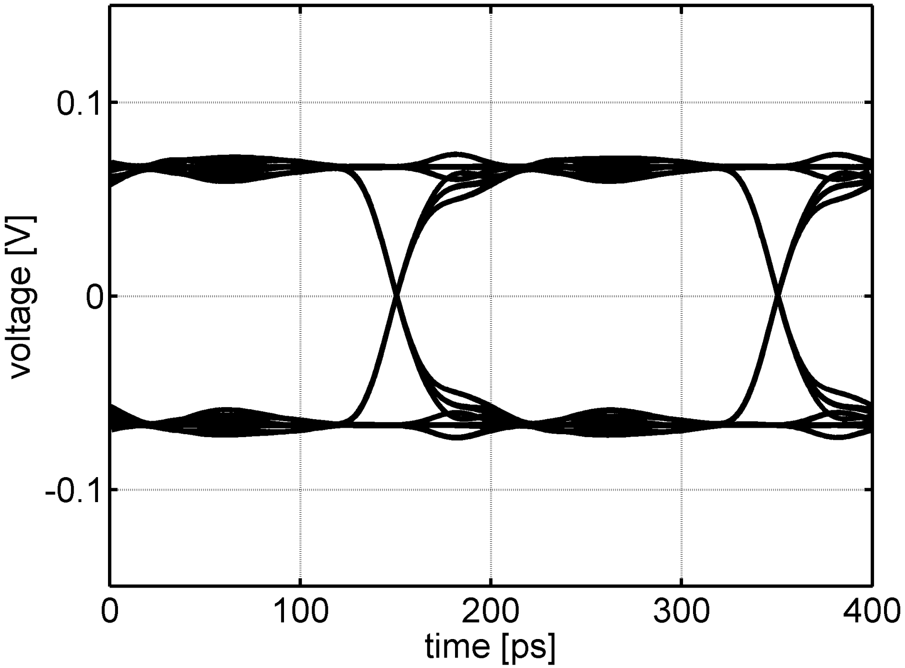
Eye Diagrams with an Ideal Source



Ideal source and ideal T-line → No reflection from the source
→ No advantage in having a T-coil at the input

Eye Diagrams with Non-Ideal Source


100 Ω source impedance, 10 cm 50 Ω T-line, no T-coil


100 Ω source impedance, 10 cm 50 Ω T-line, T-coil

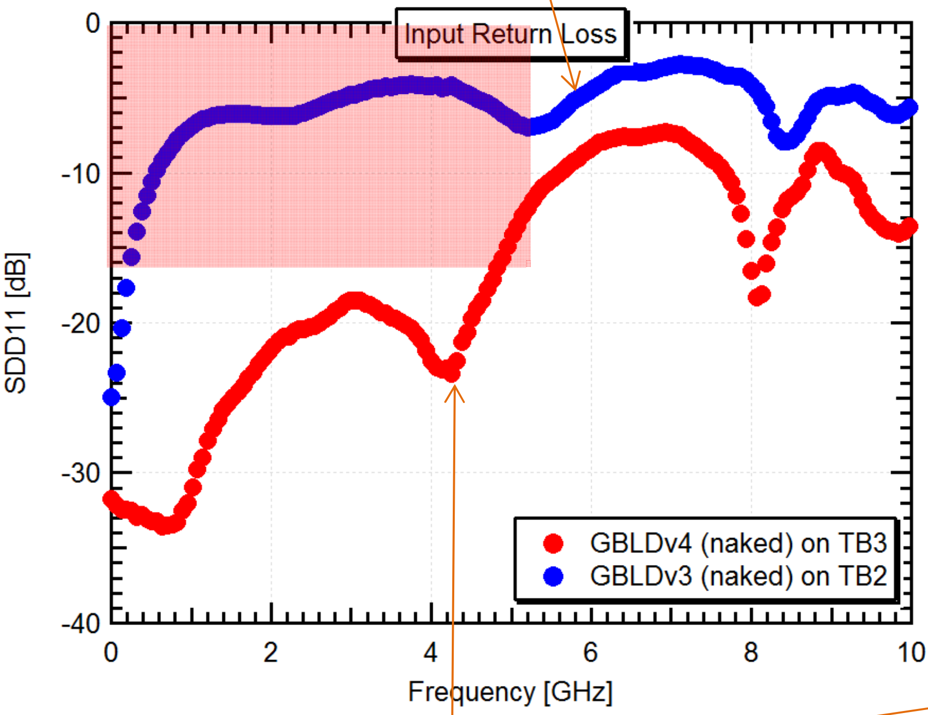
100 Ω source and ideal T-line → significant reflection from the source

→ T-coil reduces the effect of these reflection at the input gates

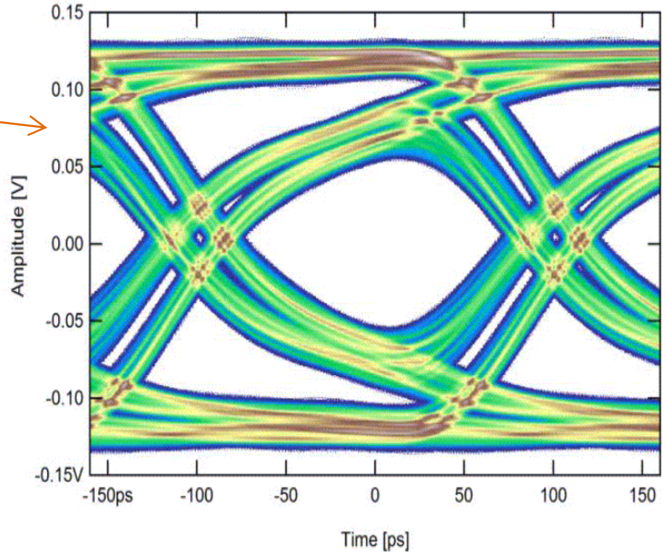
Paulo.Moreira@cern.ch

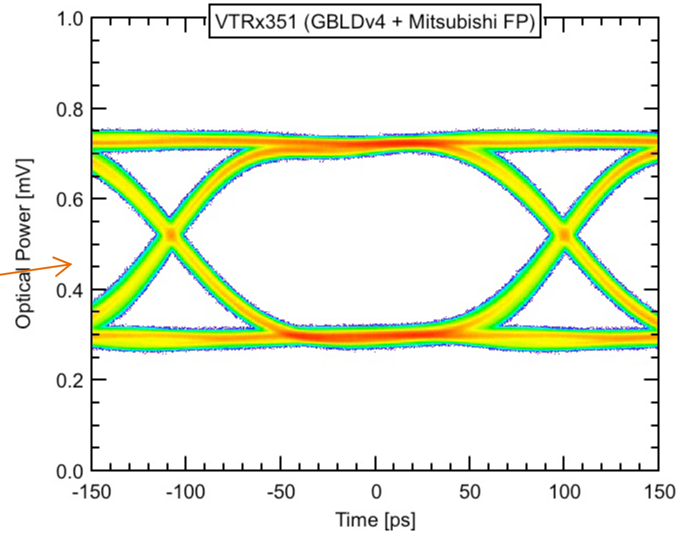
Measurement Results

Without T-Coils



With T-Coils

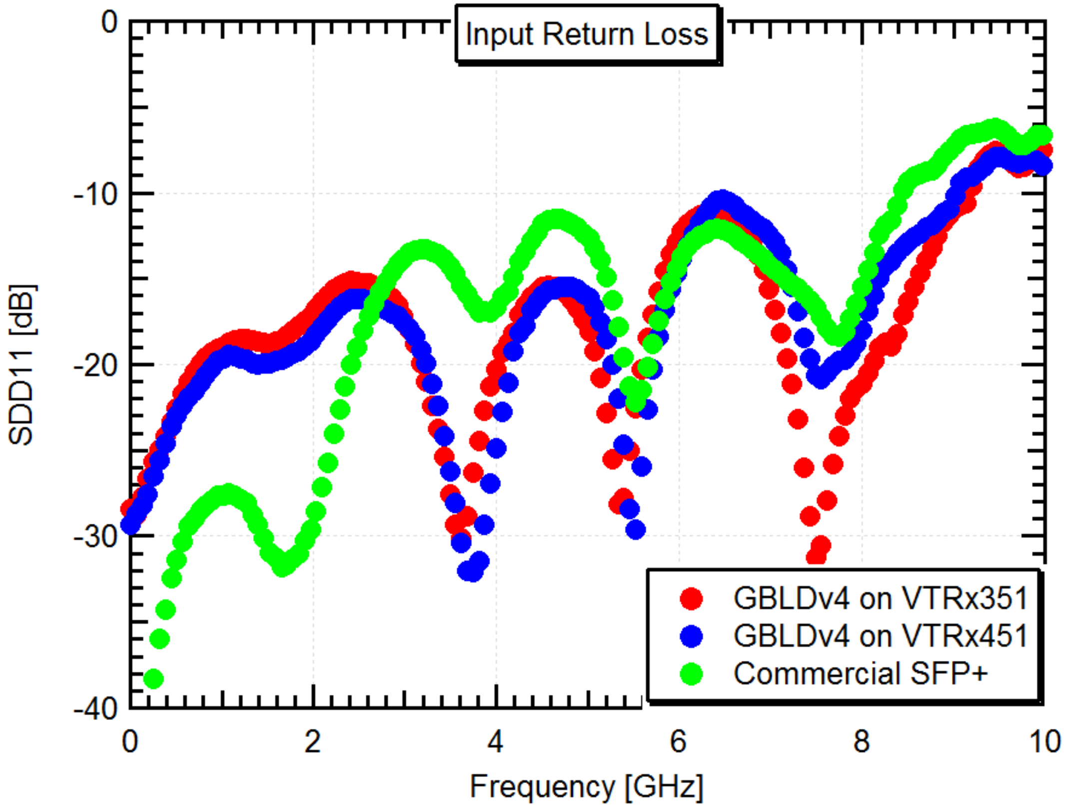




Paulo.Moreira@cern.ch

Measurement Results

Compared with a Commercial Device on the VTRx

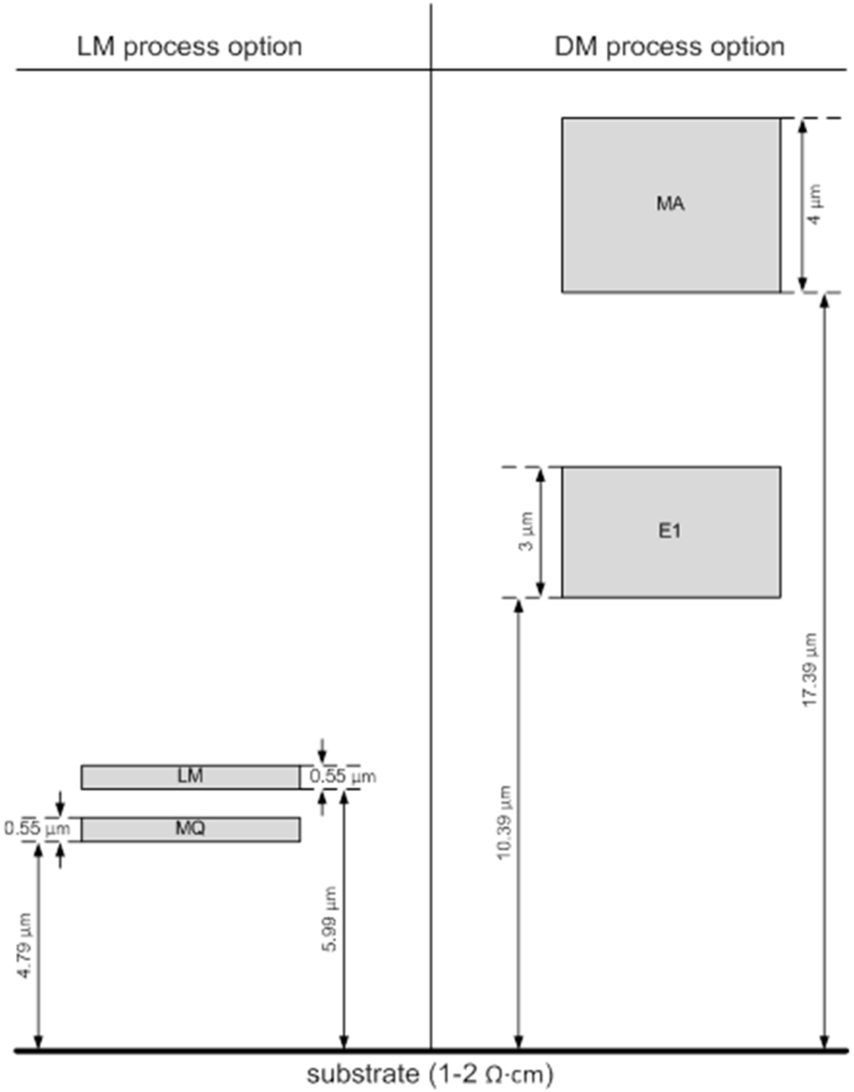


Paulo.Moreira@cern.ch

ESD PROTECTION CIRCUITS

Input Matching: Inductive Compensation

130 nm Process “Flavours”

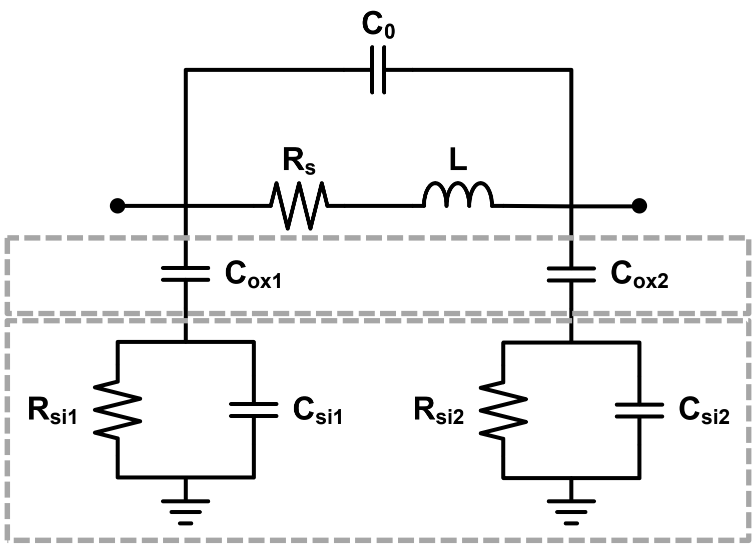


Inductors available in the design kit:

- **DM:**
  - single (MA) and dual layer (E1//MA) inductors, **coupled inductors**
- **LM**:
  - only dual layer (MQ//LM) inductors, **no coupled inductors**
- LM inductors have:
  - higher resistance because of the thinner metal layers
  - Lower Q-factor (not a problem for broadband design)
- Higher oxide capacitance because of the smaller elevation of the traces:
  - Lower self-resonant frequency

LpGBLD implemented in the LM flavour!

Inductor Model



Inductor to substrate →

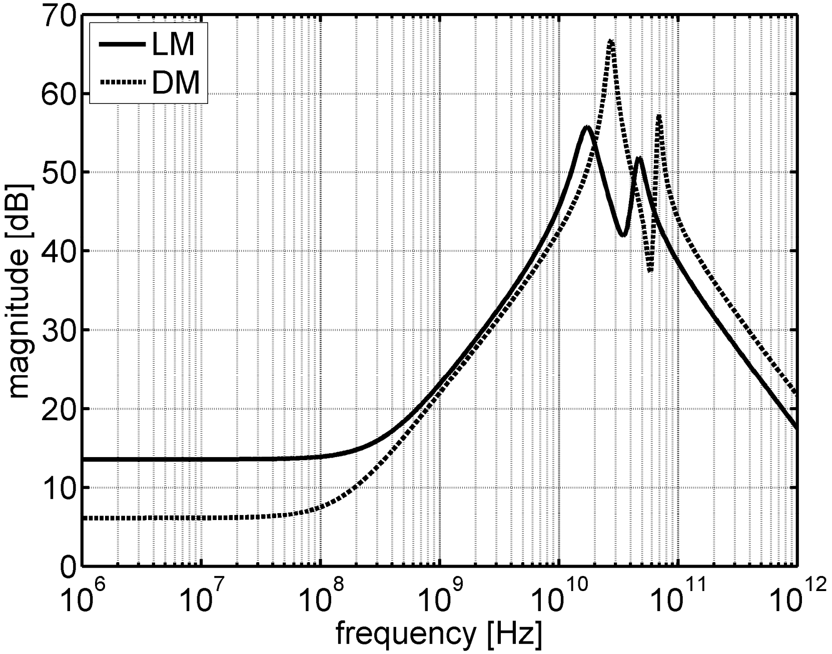
Substrate to substrate contact →

- For the same inductance :
  - C_{ox1} and C_{ox2} are ± 3 times larger in LM than in DM
  - Lower self-resonant frequency
- **BUT:**
  - High-resistivity substrate (± 1.5 $\Omega \cdot$cm)
  - Large R_{si1}/R_{si2} and small C_{si1}/C_{si2}
  - So, self-resonant frequency is not much lower

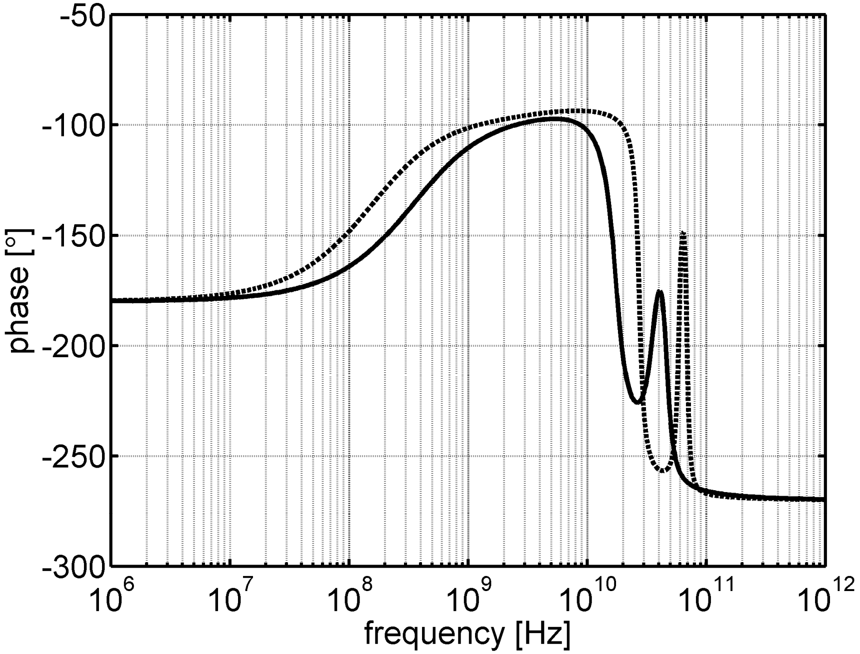
Paulo.Moreira@cern.ch

LM versus DM Inductors

magnitude of the complex impedance



phase of the complex impedance

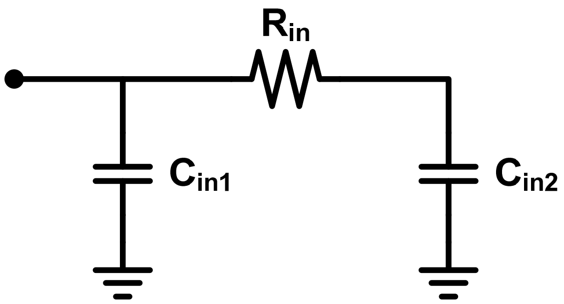


- Higher DC resistance for the LM inductor:
  - Can be mitigated by reducing the load resistance slightly
- Comparable behavior in the range 1 GHz – 8 GHz
  - The region of interest
- Higher self-resonant frequency for DM inductors (above 10 GHz)

Paulo.Moreira@cern.ch

Input Stage Input Impedance Modeling

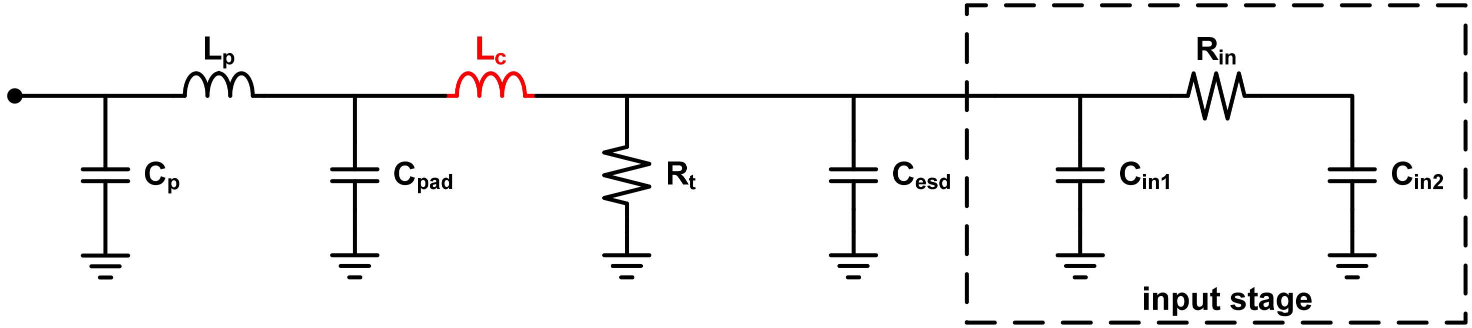
- Because of the inductive loading and the Miller effect, the input impedance of the input stage is not purely capacitive.
- Between DC and 10 GHz, it can be reliably modeled by a C-R-C network:



C_{in1}	35.8 fF
C_{in2}	25.6 fF
R_{in}	1.51 kΩ

Paulo.Moreira@cern.ch

Return Loss Improvement with a Single Inductor

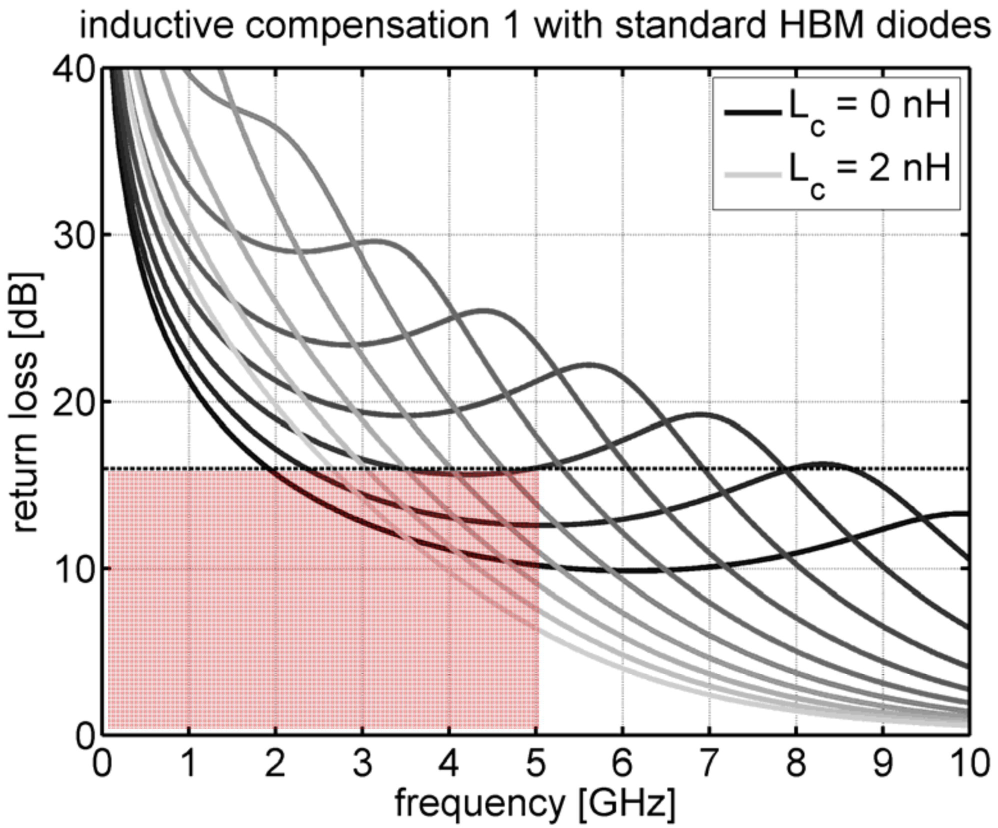


C_p	343 fF	Parasitic capacitance of the package
L_p	750 pH	Bondwire inductance
C_{pad}	75 fF	Capacitance of the bondpad
R_t	50 Ω	On-chip termination resistance
C_{esd}	378 fF	Capacitance of the ESD devices
C_{in1}	35.8 fF	Ad-hoc model for the input impedance of the input stage of the modulator
C_{in2}	25.6 fF	
R_{in}	1.51 kΩ	
L_c	?	Compensation inductor to improve the matching

Paulo.Moreira@cern.ch

Determining the Compensation Inductance

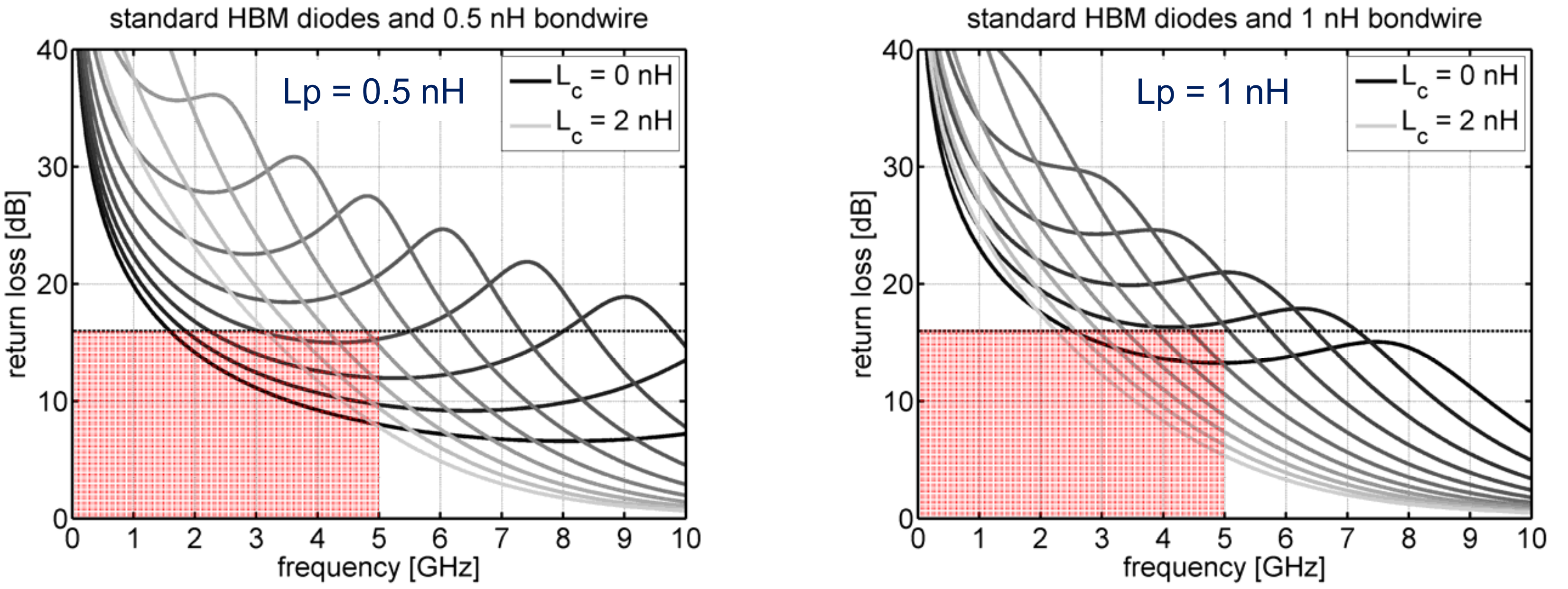
Calculations



- By “tuning” the inductor the return loss can be optimized over the frequency range of interest
- The return loss specification is met with $0.6\ \mathrm{nH} \le L_c \le 1\ \mathrm{nH}$

Paulo.Moreira@cern.ch

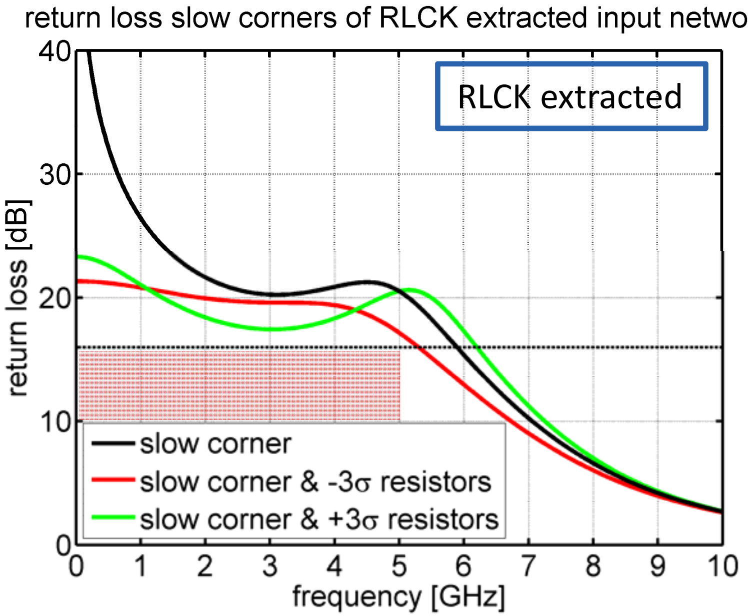
Accounting for the Bondwire Inductance Spread

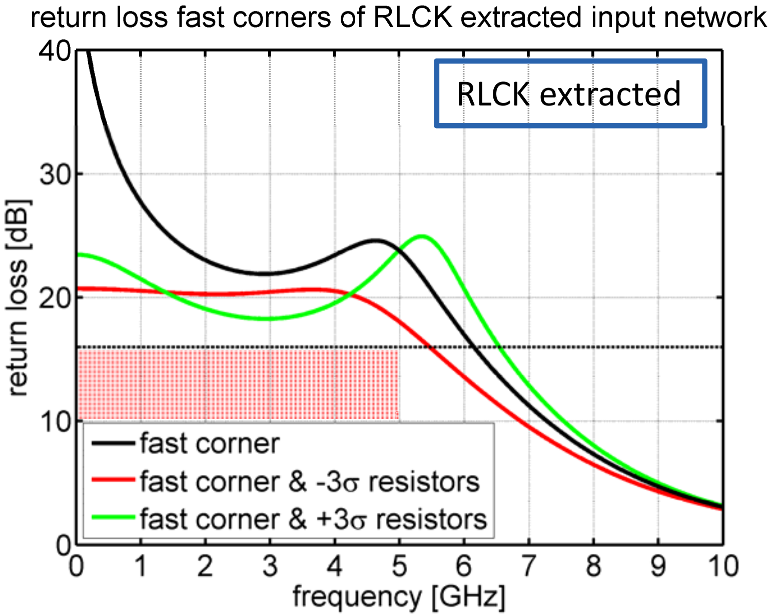


- The bondwire inductance “is also part” of the compensation network
- The value of the bondwire inductance is not well defined!
  - It can be obtained from the package model
  - But a large variation has to be assumed to account for the difference between the package cavity size and the chip size
  - Additionally, variations from the bonding process are expected
  - A range of 0.5 to 1.0 nH was assumed (nominal 0.75 nH)
- A 0.8 nH compensation inductance is OK for ‘all’ bondwire inductances considered
- **But …**

Paulo.Moreira@cern.ch

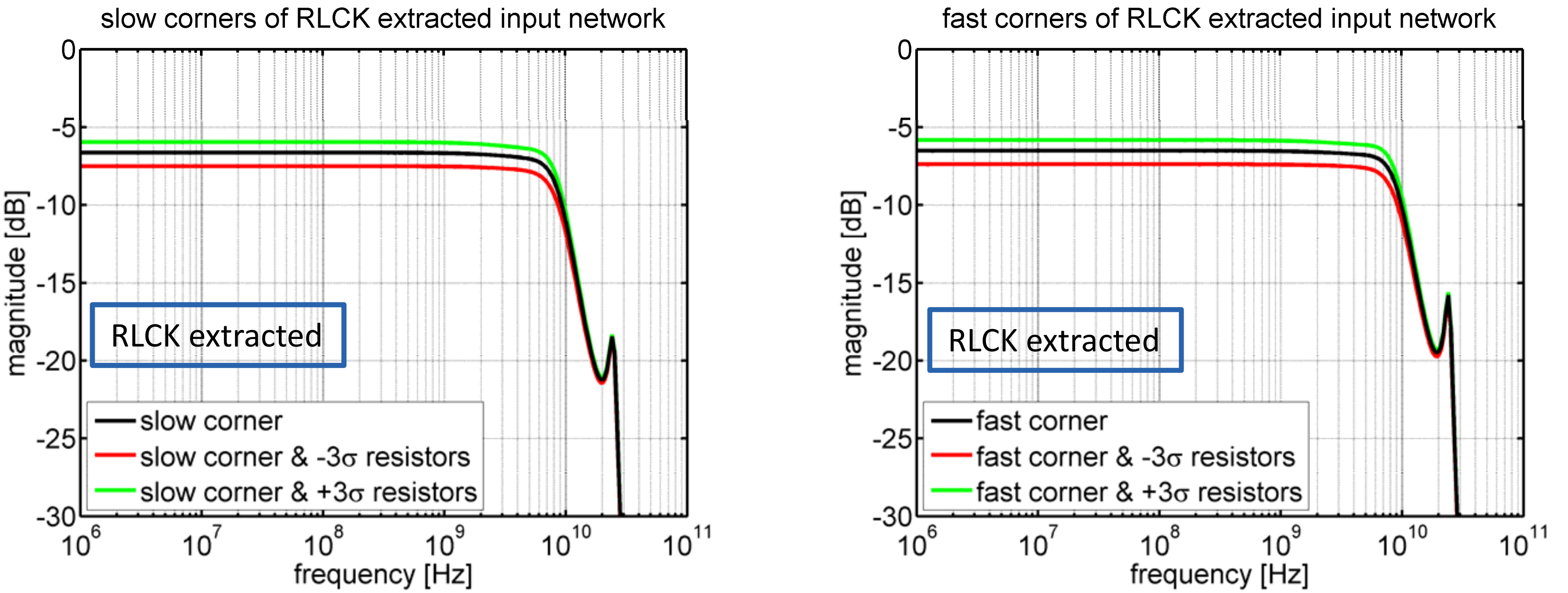
Accounting for the Termination Resistor Spread





- Poly resistors (used for R_T) have a 3σ spread of 15-20%
  - Return loss at DC can get as low as 19 dB
- In **extreme cases**, this resistor variation could push the return loss below 16 dB for frequencies above 5.5 GHz and thus violating the return loss specification
- In order to comply to the return loss specification, the compensation inductance has been decreased slightly to **0.62 nH** so that return loss at lower frequencies is traded with return loss at higher frequencies, giving a flatter overall behaviour.

Input Bandwidth

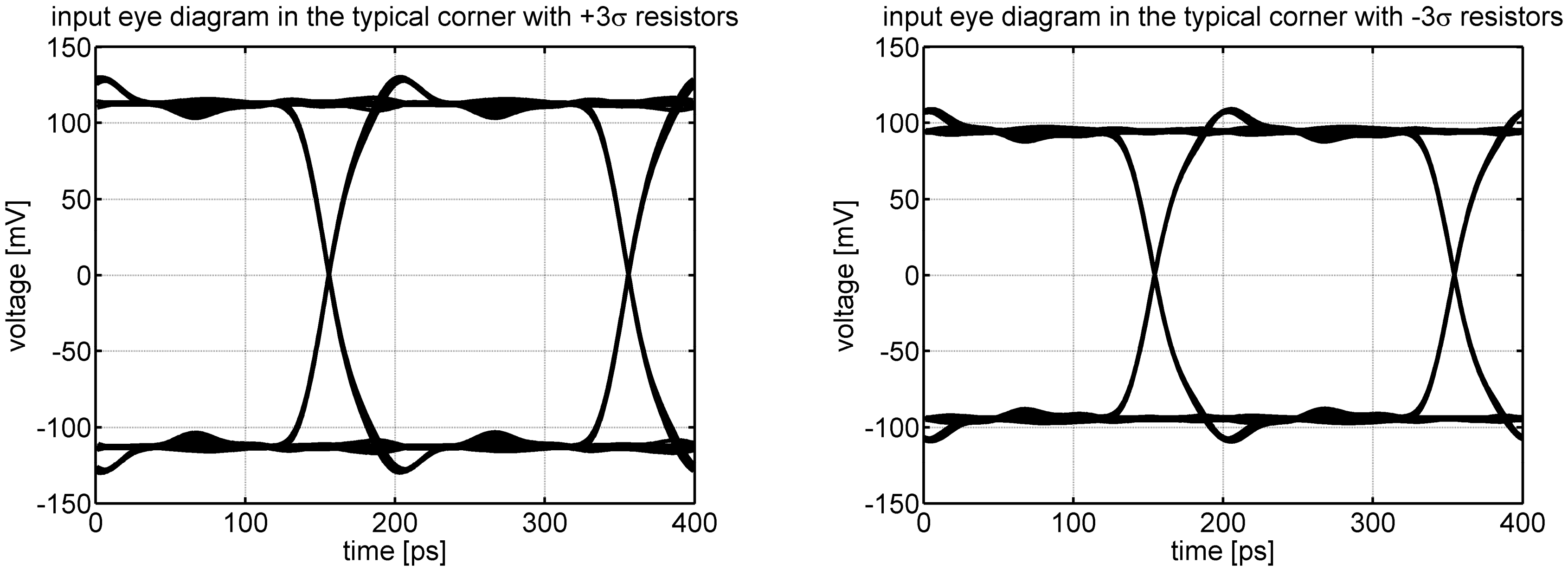


- Input bandwidth always above 9 GHz (7 GHz in GBLD v4)

Paulo.Moreira@cern.ch

Eye Diagrams

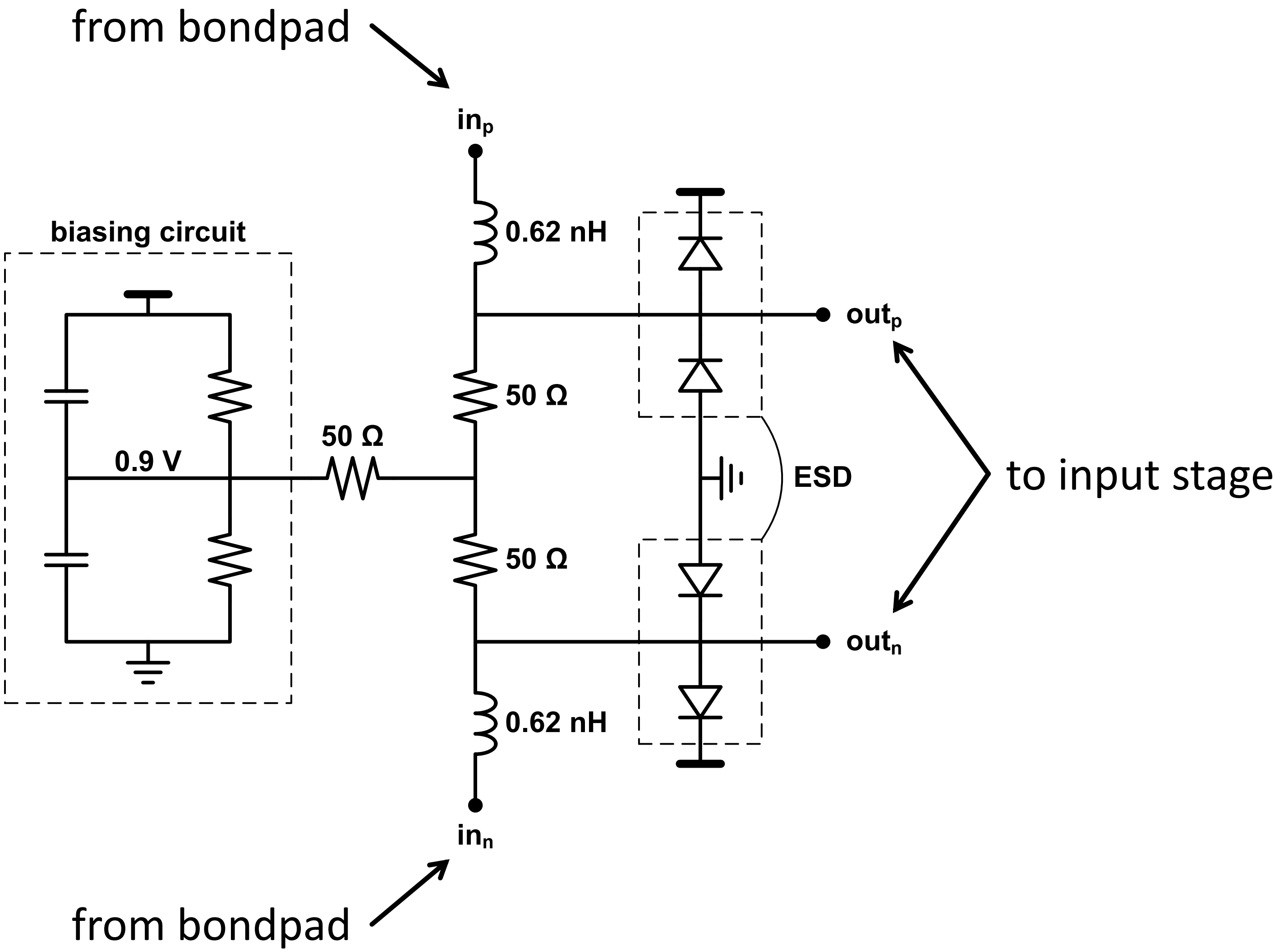
RLCK extracted



- Measured at the input stage inputs
- 50 Ω T-line of 10 cm
- 40 Ω source impedance

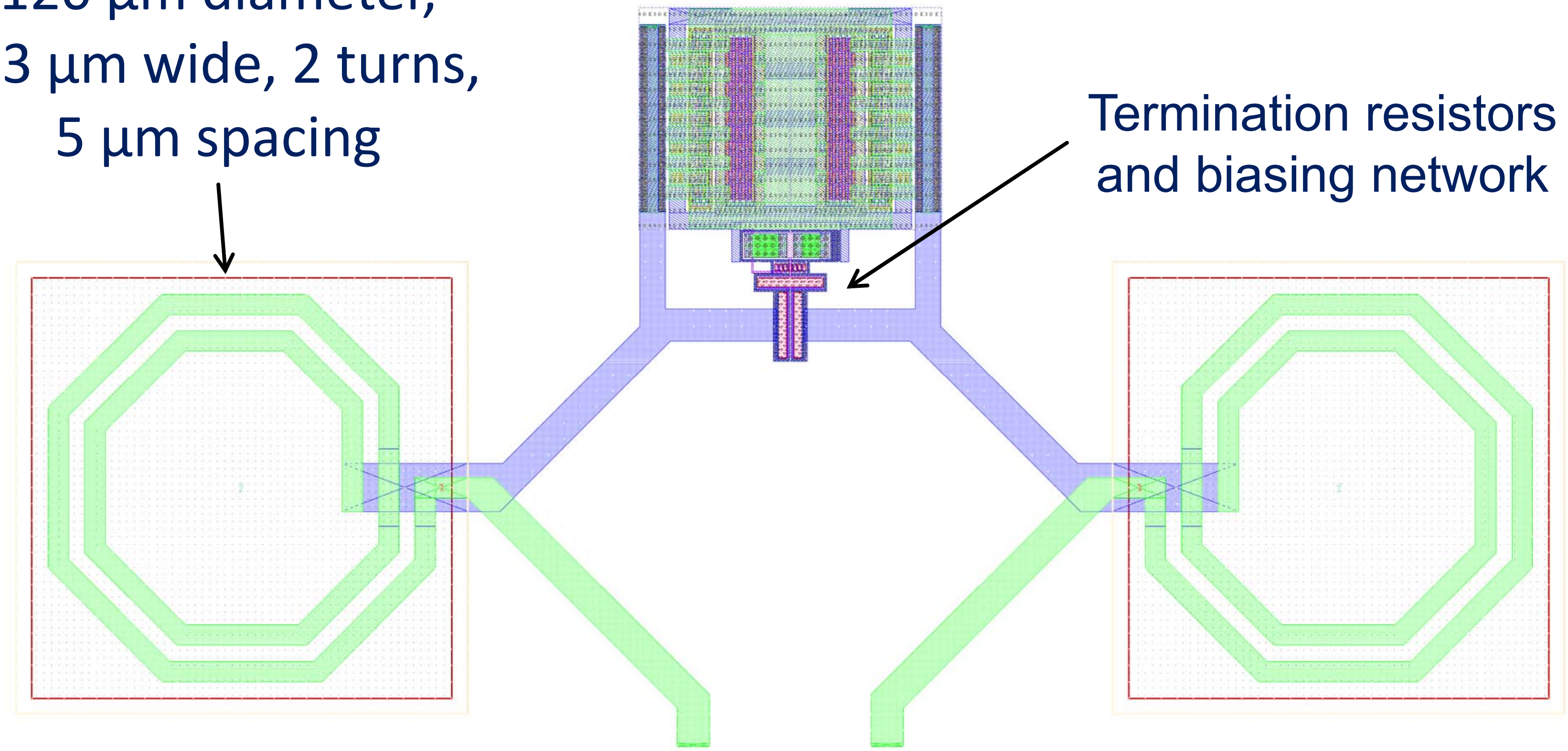
Paulo.Moreira@cern.ch

Input Network Schematic



Paulo.Moreira@cern.ch

Input Network Layout



0.62 nH
120 µm diameter,
6.3 µm wide, 2 turns,
5 µm spacing

Standard HBM diodes
(same as GBLD v4)

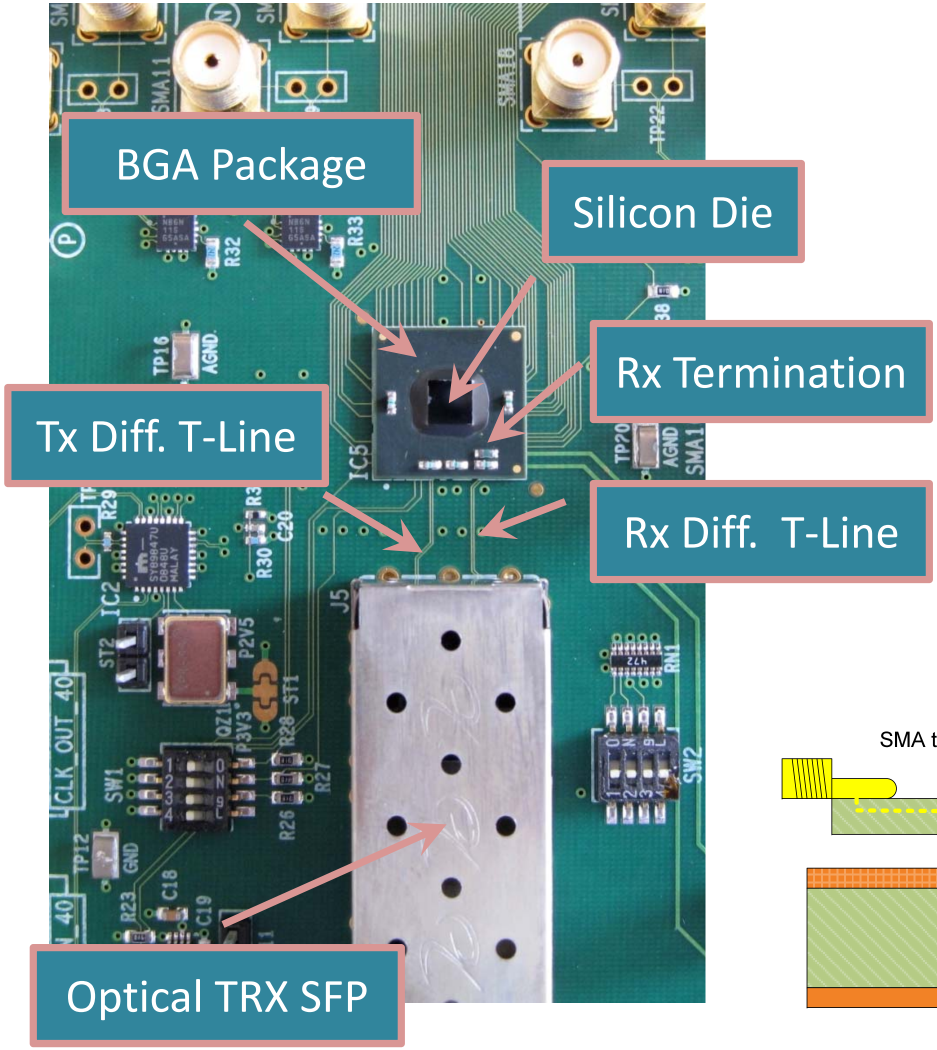
Termination resistors
and biasing network

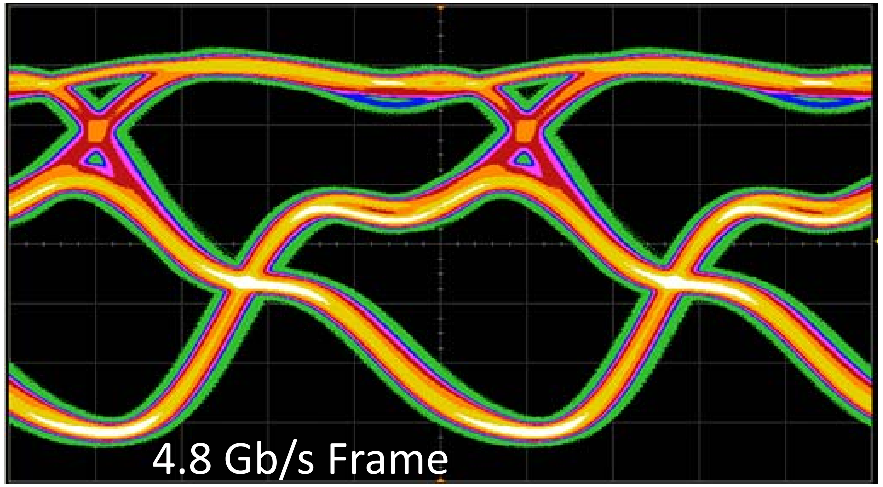
to bondpads

Paulo.Moreira@cern.ch

CONNECTORS
PCB & PACKAGE

Reaching the ASIC



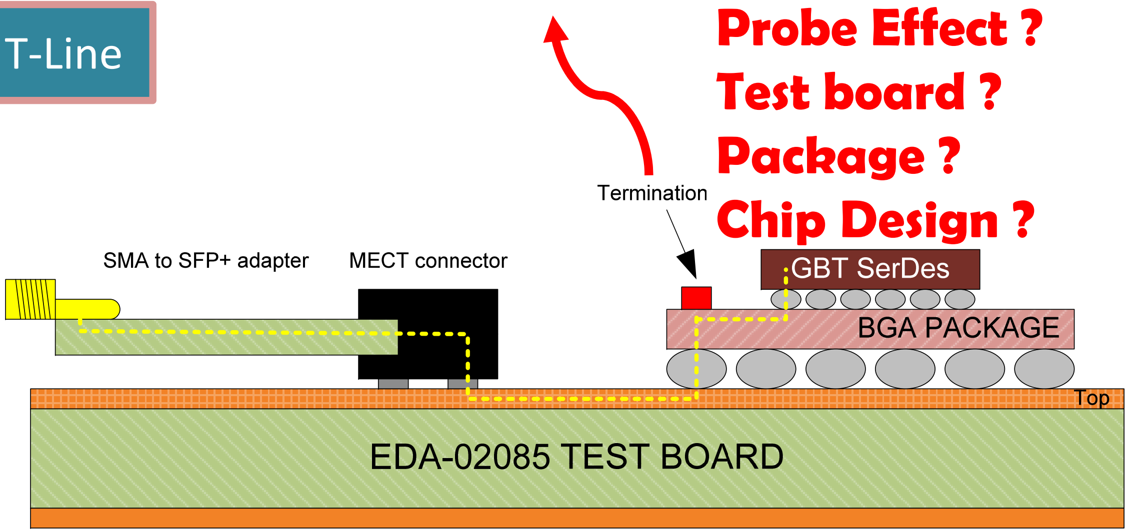


Probe Effect ?

Test board ?

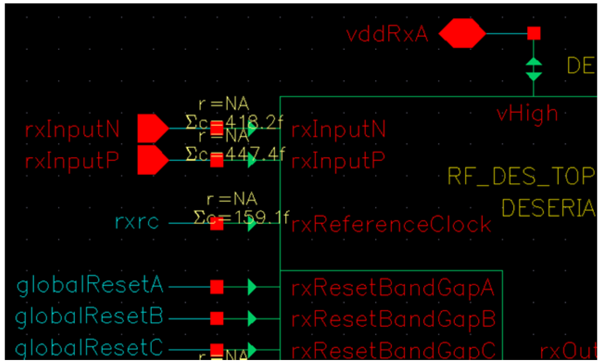
Package ?

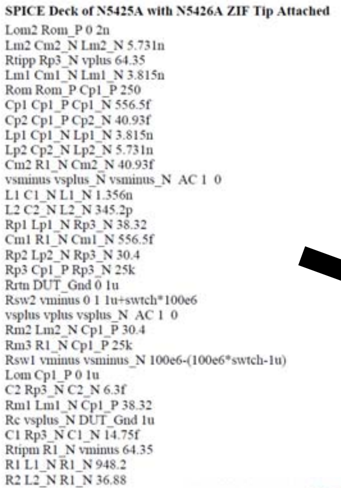
Chip Design ?



Paulo.Moreira@cern.ch

“Build” Models





SPICE Deck of N5425A with N5426A ZIF Tip Attached

```
Lom2 Rom_P 0 2n
Lm2 Cm2_N Lm2_N 5.731n
Rtipp Rp3_N vplus 64.35
Lm1 Cm1_N Lm1_N 3.815n
Rom Rom_P Cp1_P 250
Cp1 Cp1_P Cp1_N 556.5f
Cp2 Cp1_P Cp2_N 40.93f
Lp1 Cp1_N Lp1_N 3.815n
Lp2 Cp2_N Lp2_N 5.731n
Cm2 R1_N Cm2_N 40.93f
vsminus vsplus_N vsminus_N  AC 1  0
L1 C1_N L1_N 1.356n
L2 C2_N L2_N 345.2p
Rp1 Lp1_N Rp3_N 38.32
Cm1 R1_N Cm1_N 556.5f
Rp2 Lp2_N Rp3_N 30.4
Rp3 Cp1_P Rp3_N 25k
Rrtn DUT_Gnd 0 1u
Rsw2 vminus 0 1 1u+swtch*100e6
vsplus vplus vsplus_N  AC 1  0
Rm2 Lm2_N Cp1_P 30.4
Rm3 R1_N Cp1_P 25k
Rsw1 vminus vsminus_N 100e6-(100e6*swtch-1u)
Lom Cp1_P 0 1u
C2 Rp3_N C2_N 6.3f
Rm1 Lm1_N Cp1_P 38.32
Rc vsplus_N DUT_Gnd 1u
C1 Rp3_N C1_N 14.75f
Rtipm R1_N vminus 64.35
R1 L1_N R1_N 948.2
R2 L2_N R1_N 36.88

.AC DEC 200 200k 20G SWEE

.PARAM swtch=1
```

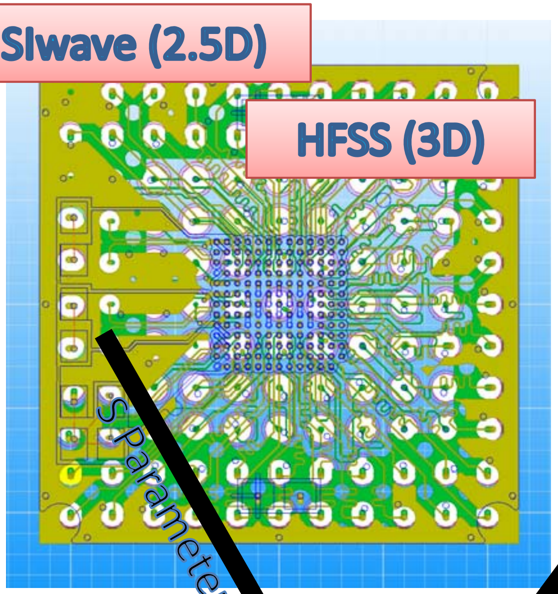
Equivalent circuit

Spice Netlist

HSpice Netlist

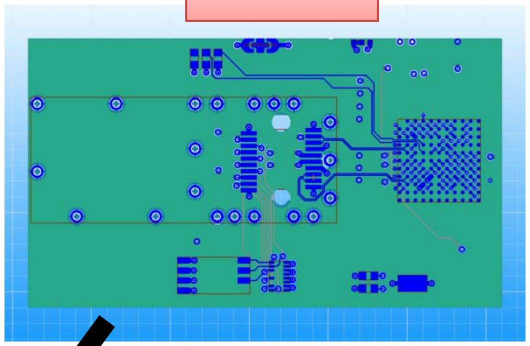
SIwave (2.5D)

HFSS (3D)

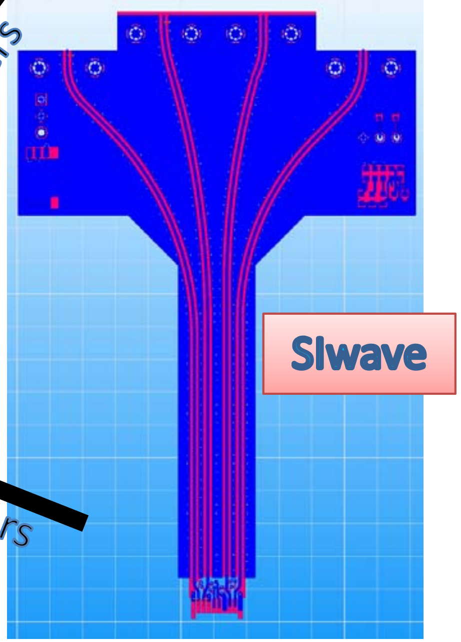


S Parameters

SIwave



S Parameters



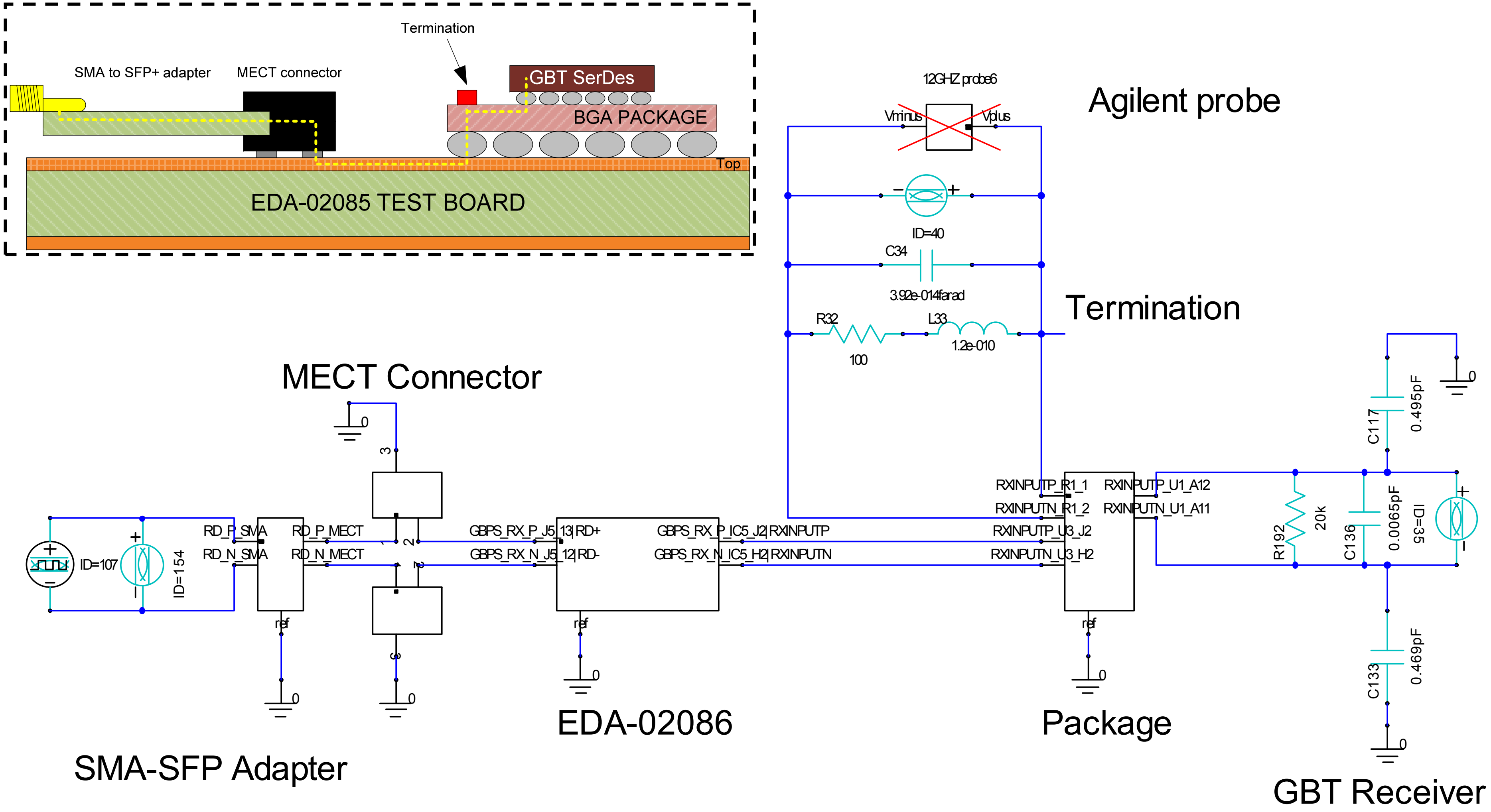
SIwave

S Parameters

Ansoft Designer

Paulo.Moreira@cern.ch

Fully Model the Signal Path



Agilent probe

Termination

MECT Connector

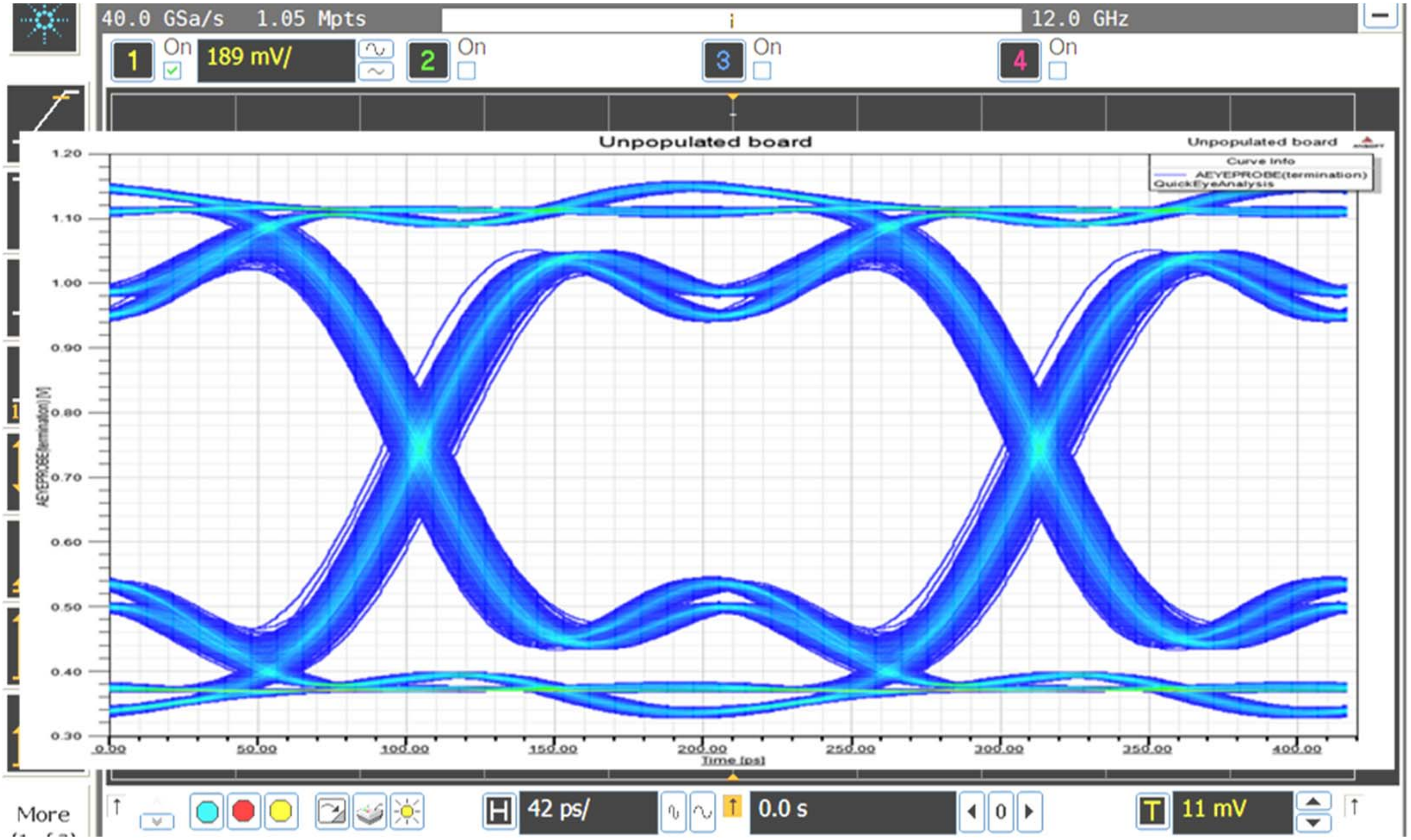
EDA-02086

Package

SMA-SFP Adapter

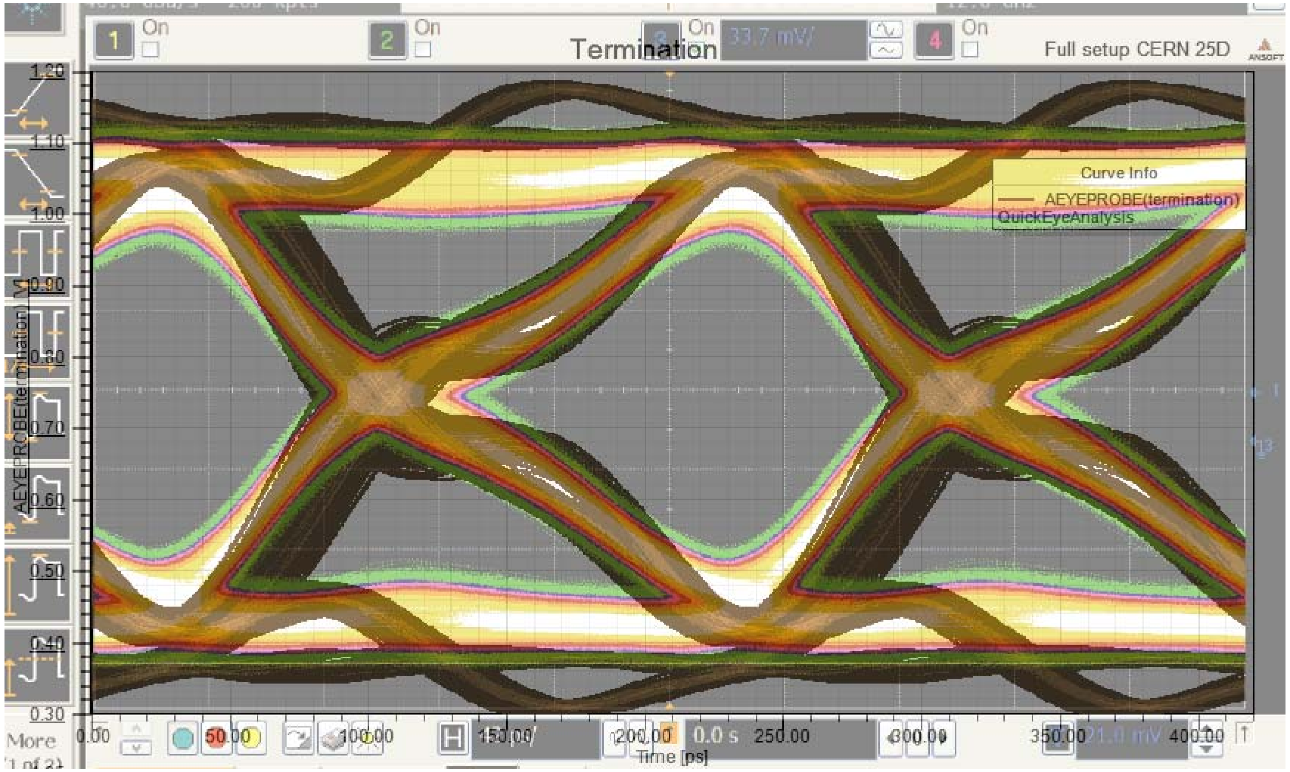
GBT Receiver

Unpopulated Board



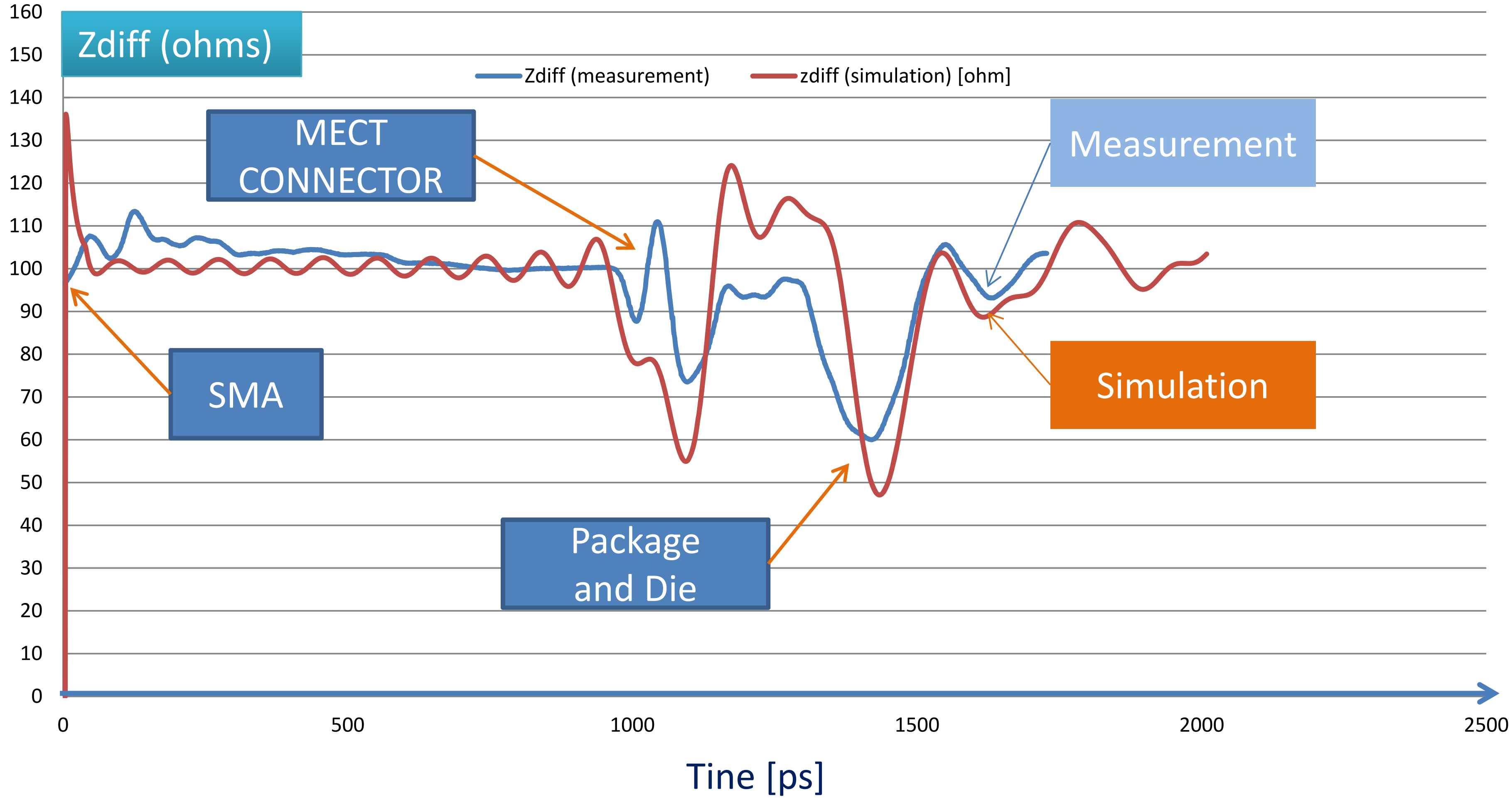
Measurement (yellow) and simulation (blue) with 100 ohms termination at the BGA pads

Fully Populated Board



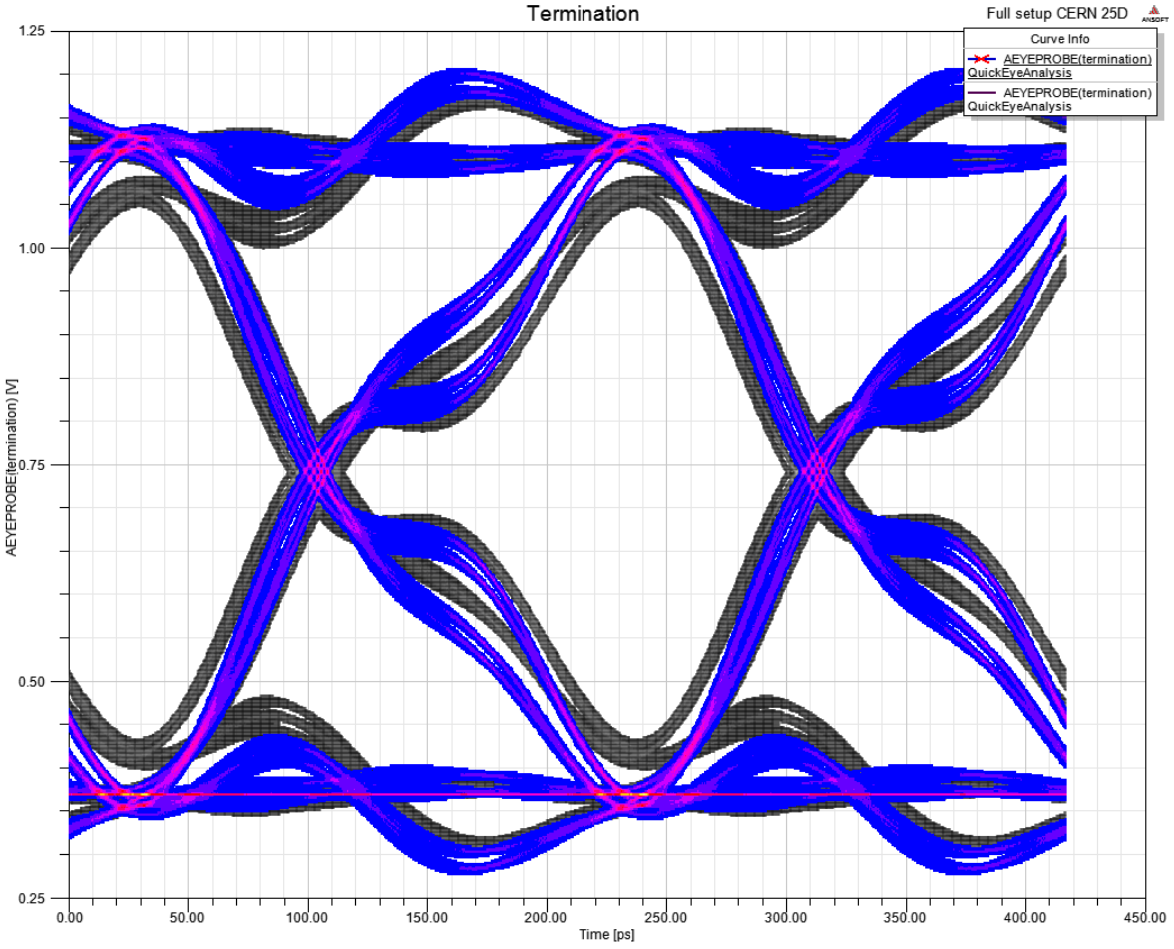
Probe is at the receiver termination , ≈ 3.5 mm from the input buffer.
Measurement (yellow) and Simulation (Brown).

Checking the Models



Paulo.Moreira@cern.ch

Schrodinger In High Frequency Electronics!?



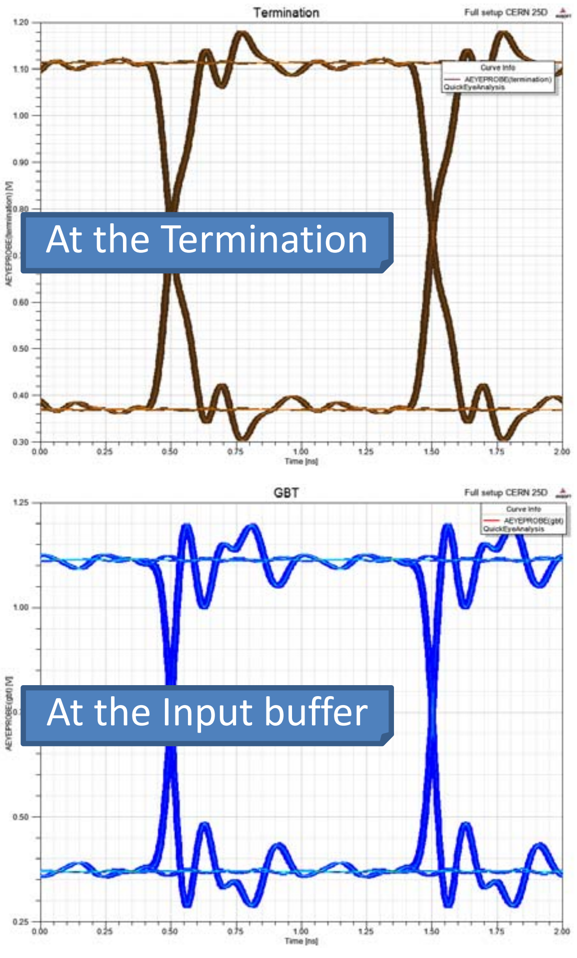
- Can signals be observed without being disturbed?
- Manufacturers provide equivalent electrical modes for their oscilloscope probes
- Simulate to evaluate how much the system is being disturbed
- In our case the loading effect of the probe was small!

With Probe (gray)

No Probe (blue)

“Virtual Probe” - Simulation

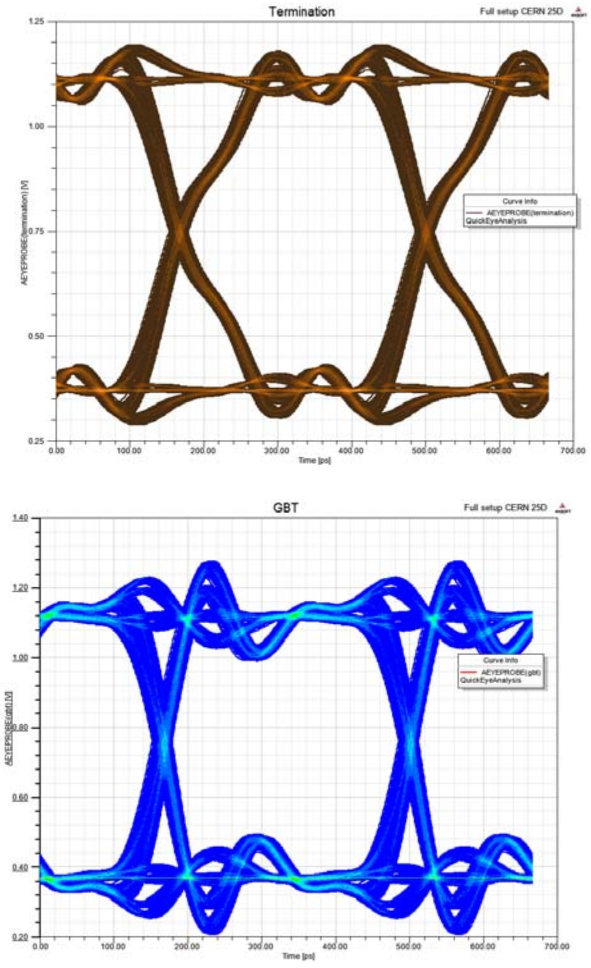
1Gbps



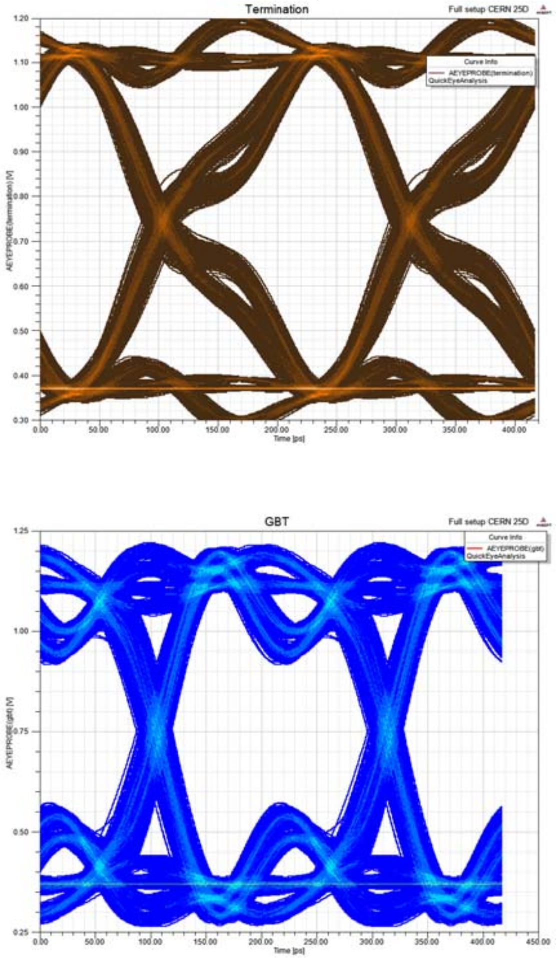
At the Termination

At the Input buffer

3.2Gbps

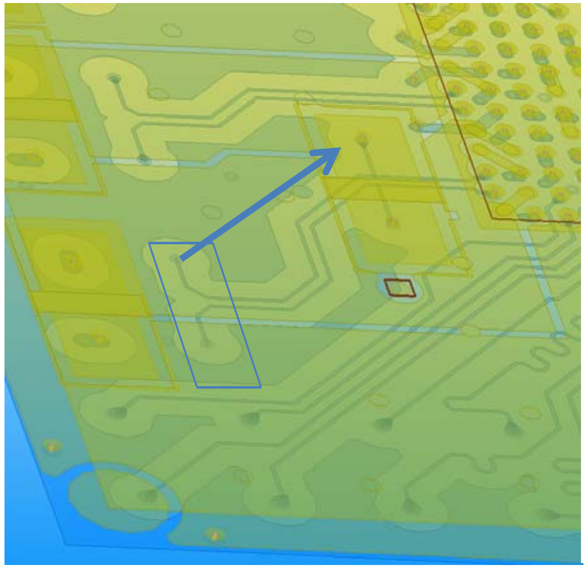


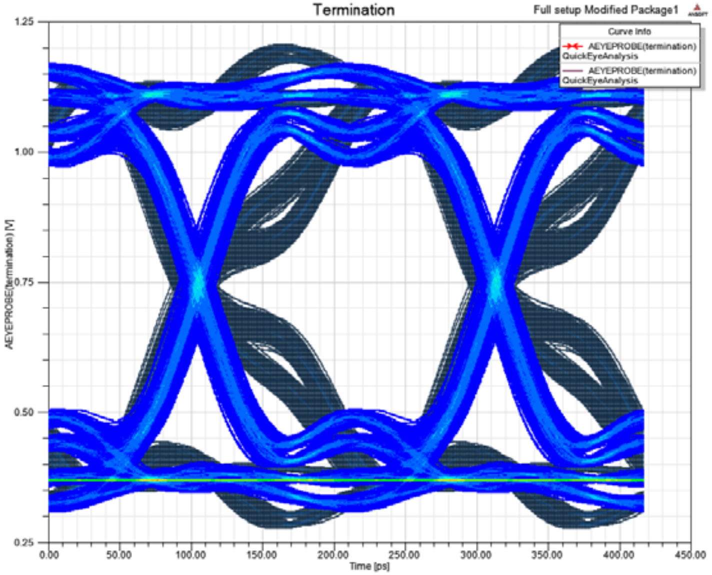
4.8Gbps



Paulo.Moreira@cern.ch

Improving the Package



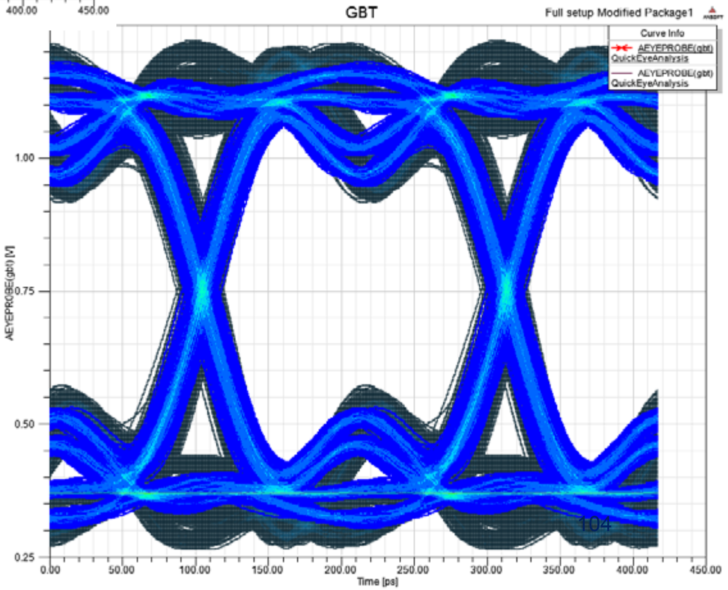


Original design: Dark blue

At the Termination

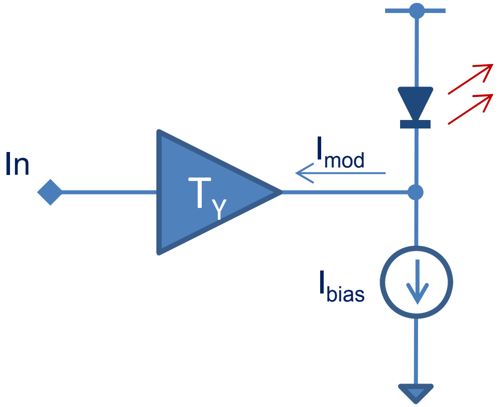
Move the termination close to the input buffer

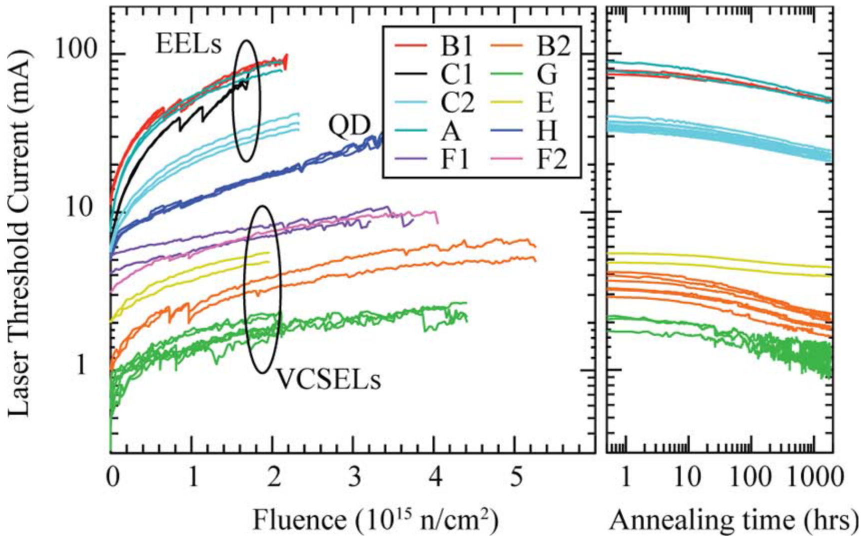
At the Input buffer

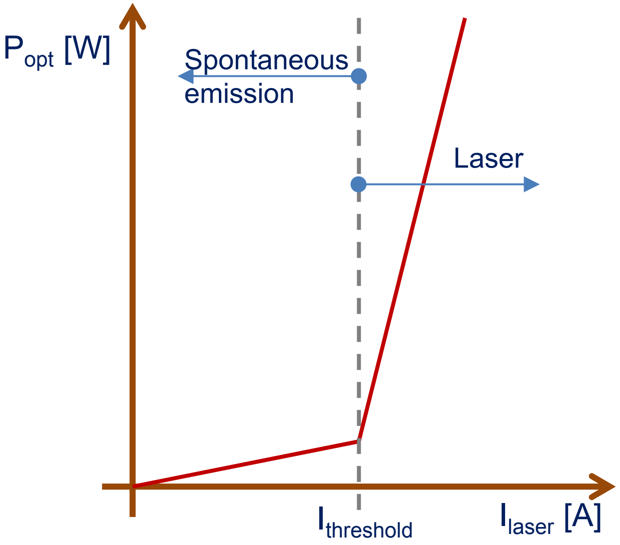


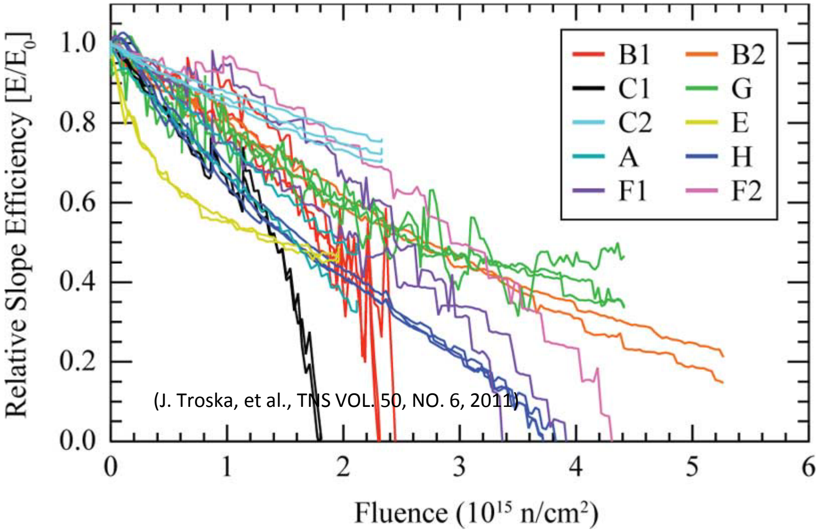
LASER DRIVER DESIGN

LASER-Diodes in HEP









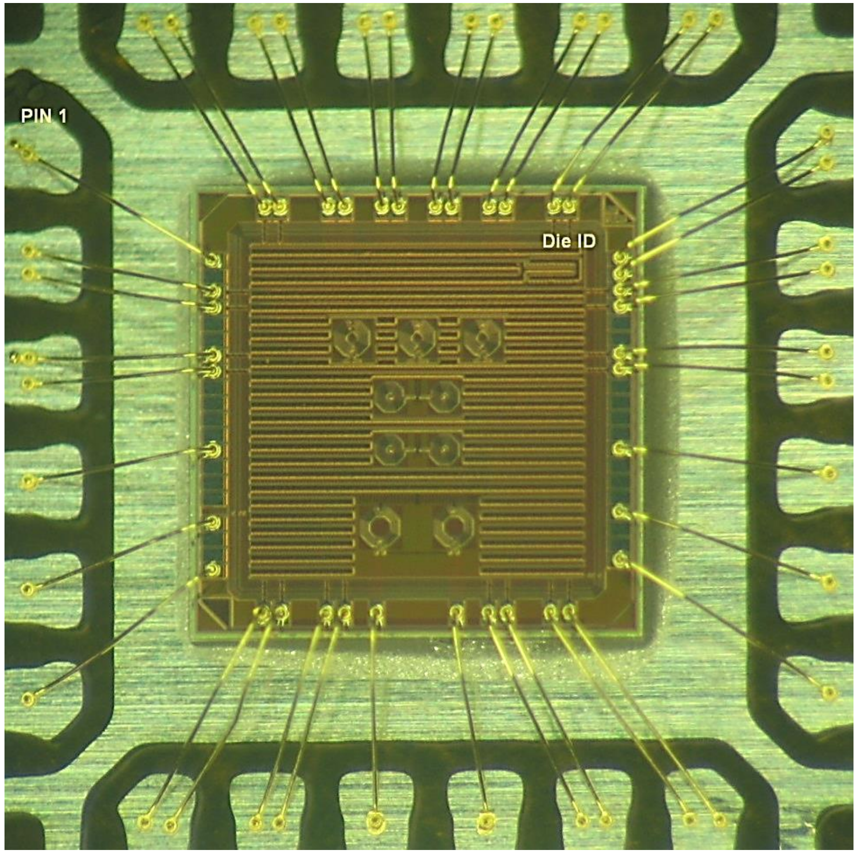
(J. Troska, et al., TNS VOL. 50, NO. 6, 2011)

Paulo.Moreira@cern.ch

GBLD – Giga Bit Laser Driver

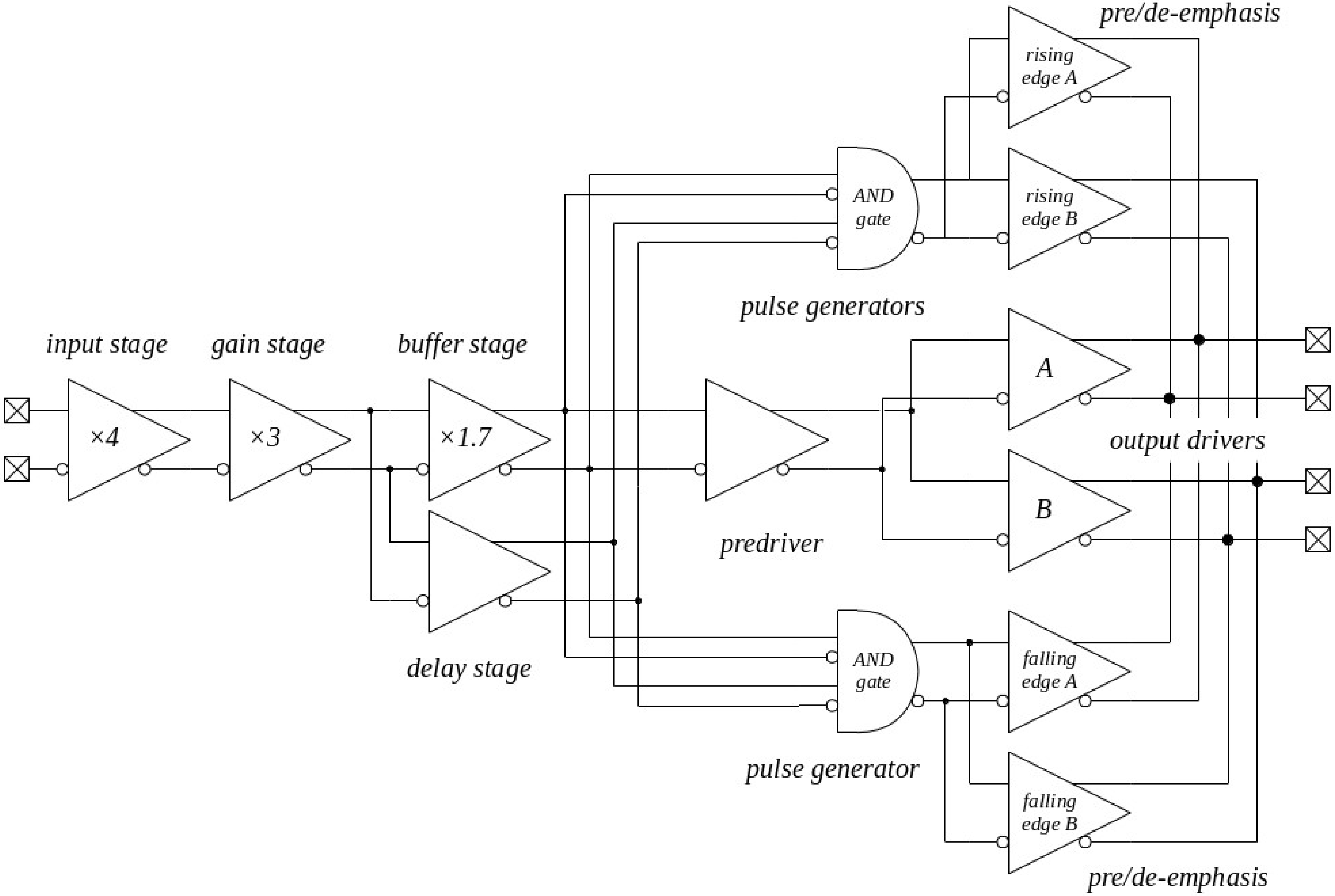
Main specs:

- Bit rate 5 Gb/s (min)
- Modulation:
  - Current sink
  - Single-ended/differential
- Laser modulation current:
  - 2 to 12 mA
- Laser bias:
  - 2 to 43 mA
- “Equalization”
  - Pre-emphasis/de-emphasis
  - Independently programmable for rising/falling edges
- Input return loss
  - > 16 dB for f < 5 GHz
- Supply voltage: 2.5 V
- Die size: 2 mm × 2 mm
- I2C programming interface



Paulo.Moreira@cern.ch

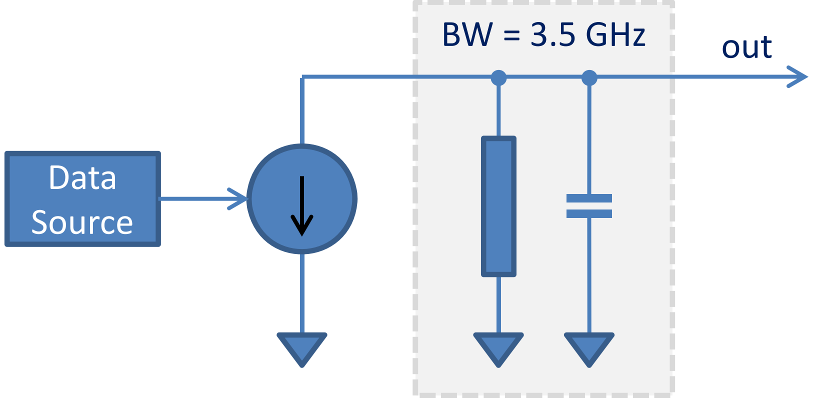
GBLD Architecture

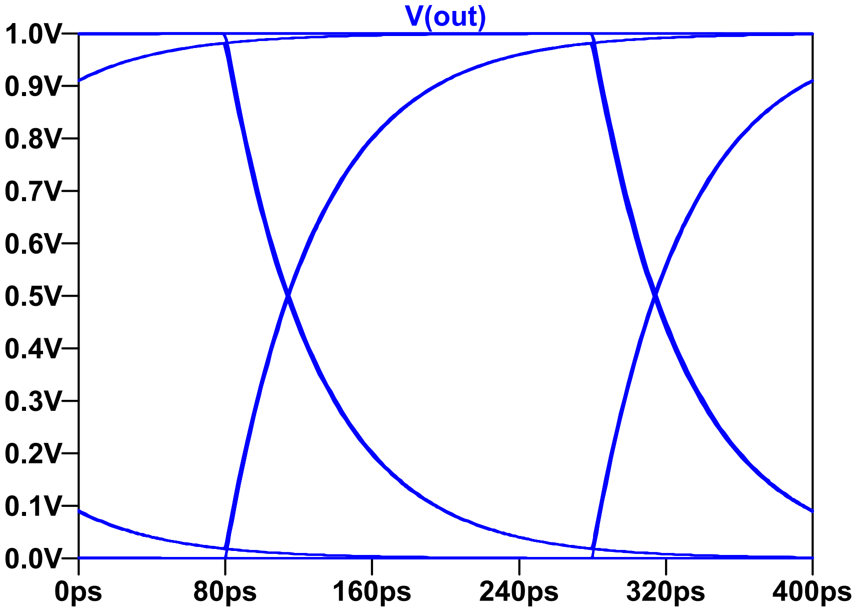


Paulo.Moreira@cern.ch

Pre- De-Emphasis (1/5)

- To transmit NRZ data with little ISI the transmission channel must have a bandwidth of at least (rule of thumb):
  - BW = 0.7 × Bit Rate
- For example, for transmission at 5 Gb/s the bandwidth required is:
  - BW = 0.7 × 5 Gb/s = 3.5 GHz
- For illustration purposes lets suppose that:
  - The laser driver is modelled by an ideal current source
  - And the circuit driven by the laser driver is modelled by a RC network with 3.5 GHz bandwidth:
    - R = 50 Ω
    - C = 0.91 pF
- Simulated eye-diagram:
  - There is very little ISI:
    - The eye is well opened vertically and horizontally
    - The jitter is very low

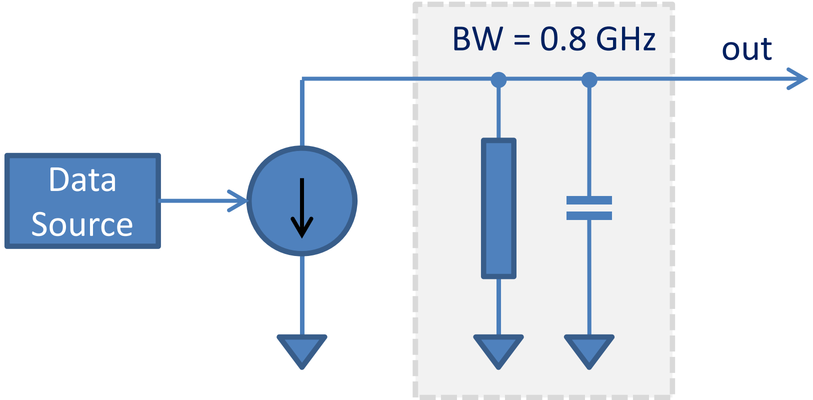


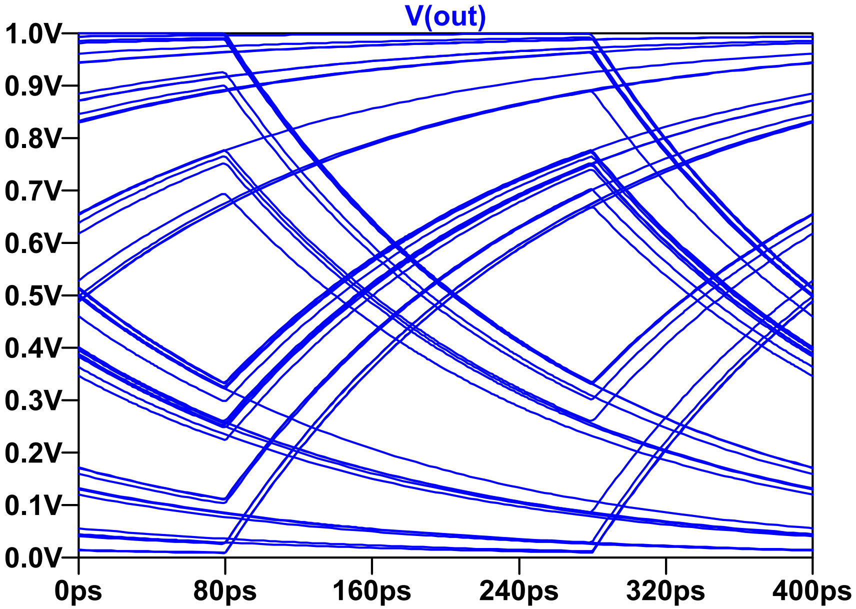


Paulo.Moreira@cern.ch

Pre- De-Emphasis (2/5)

- Lets suppose now that the bandwidth of the circuit being driven is four times lower:
  - BW = 3.5 GHz / 4 = 795 MHz
    - R = 50 Ω
    - C = 3.64 pF
- Simulated eye-diagram
  - There are significant amounts of ISI:
    - A “bit” extends for much longer than a bit period
  - The eye-diagram is almost closed vertically and horizontally
  - Jitter is high
  - The BER would be “prohibitive” for such a system!

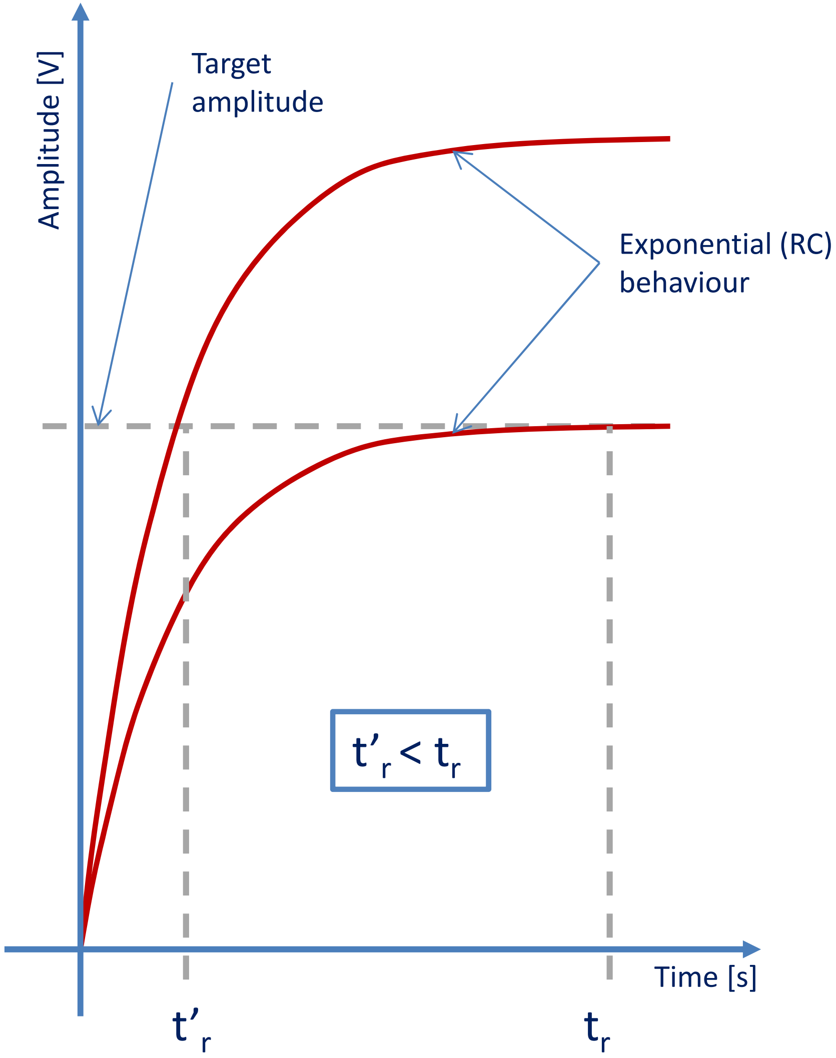




Paulo.Moreira@cern.ch

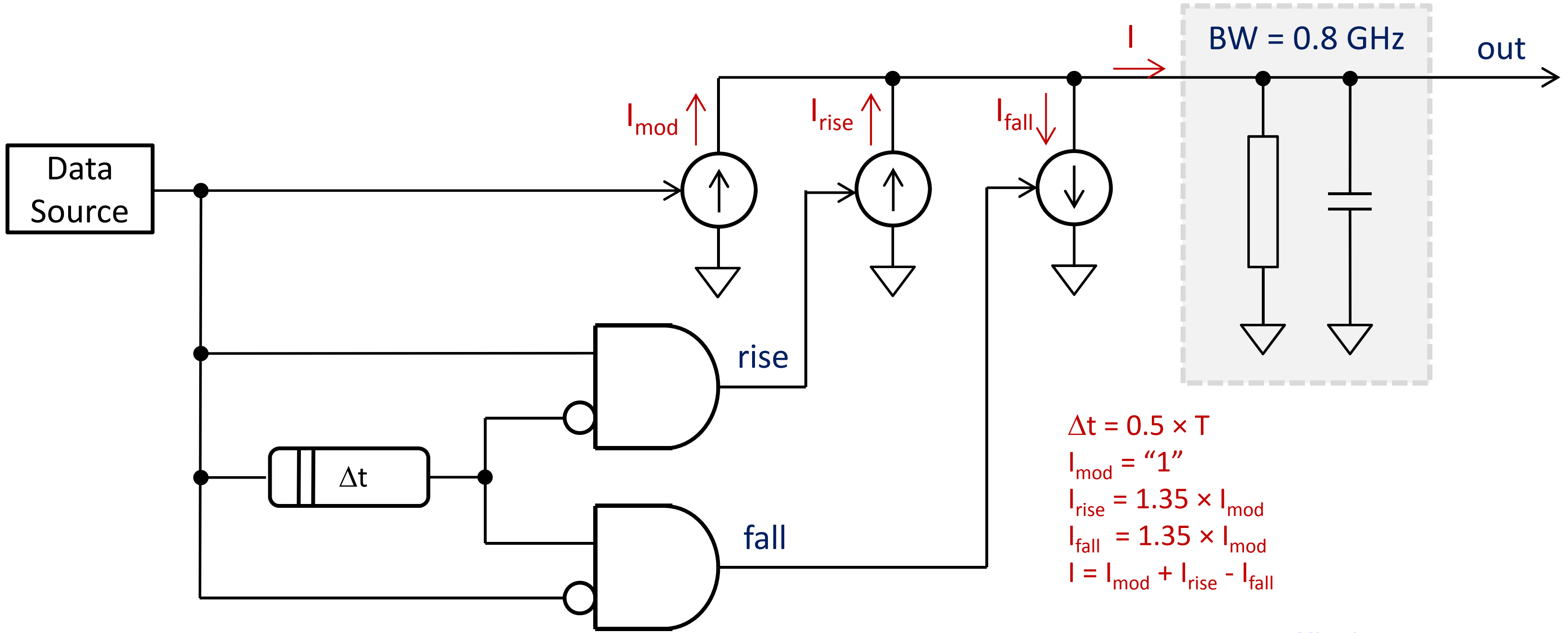
Pre- De-Emphasis (3/5)

- The problem with the low bandwidth is that:
  - For fast successions of “0s”and “1s” the signal has no time to reach the final value before a new “0” or a new “1” is transmitted.
- In an RC type of circuit this can be “easily” overcome:
  - At every “0”-t-“1” or “1”-to-“0” transition drive the circuit to a higher (or lower) voltage than the final one:
    - In our case this is accomplished by using a larger current than the final one
  - Once the voltage reaches the desired amplitude switch to the final current.
- In practice, no level crossing detection is made:
  - Immediately after a transition and for a “short” time a current pulse is added to the “normal” modulation current



Paulo.Moreira@cern.ch

Pre- De-Emphasis (4/5)



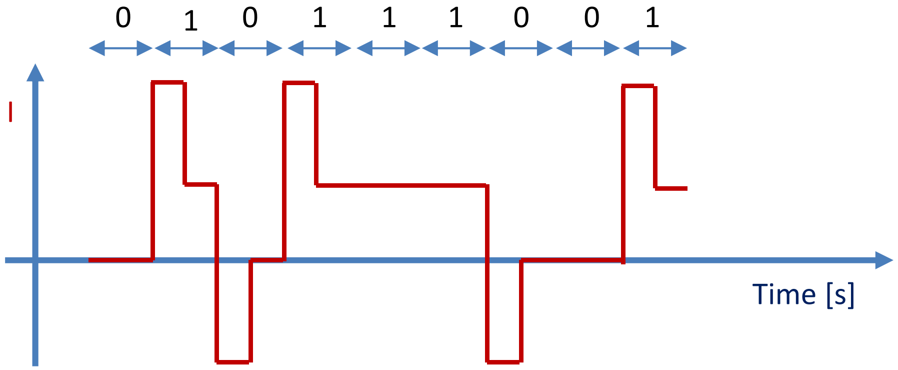
$\Delta t = 0.5 \times T$

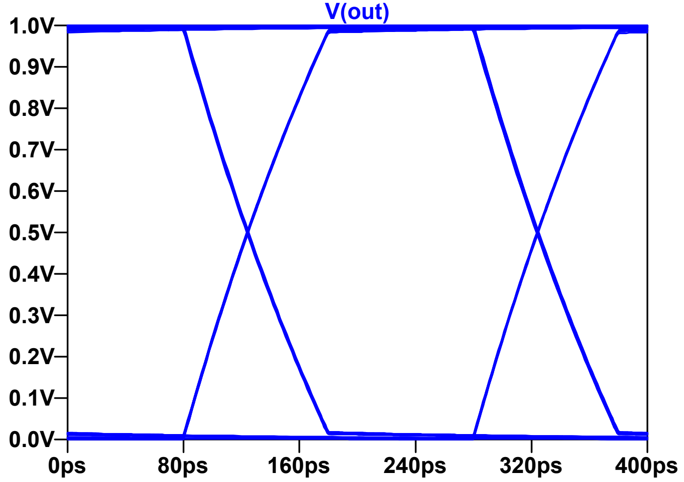
I_{mod} = “1”

$I_{rise} = 1.35 \times I_{mod}$

$I_{fall} = 1.35 \times I_{mod}$

$I = I_{mod} + I_{rise} - I_{fall}$



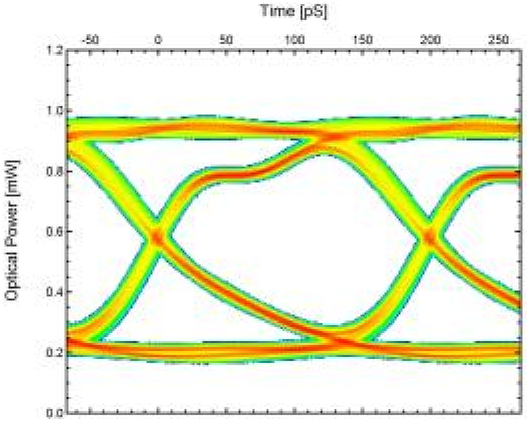


Paulo.Moreira@cern.ch

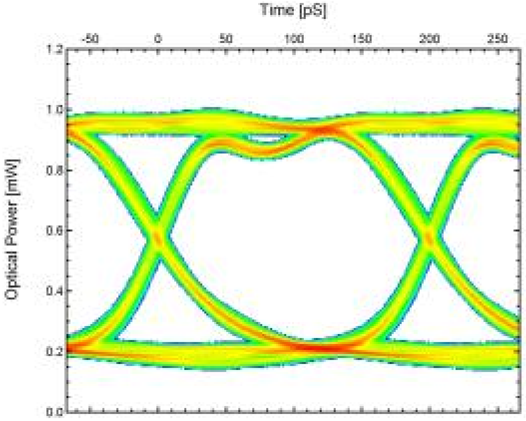
Pre- De-Emphasis (5/5)

All eye-diagrams with:

- Edge Emitting Laser Diode
- I_{bias} = 20 mA
- I_{mod} = 20 mA

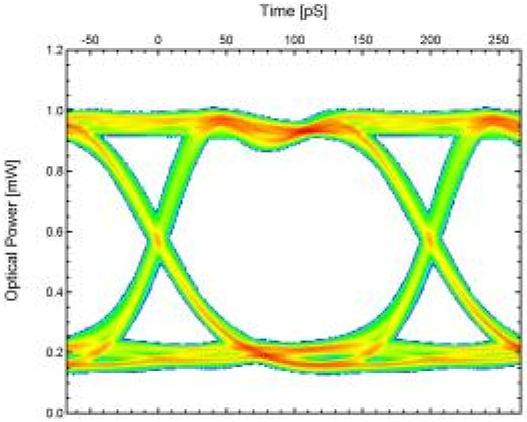


De-Emphasis: $\Delta I_{rise} = \Delta I_{fall} = -10$ mA
→ Rise/fall times increased

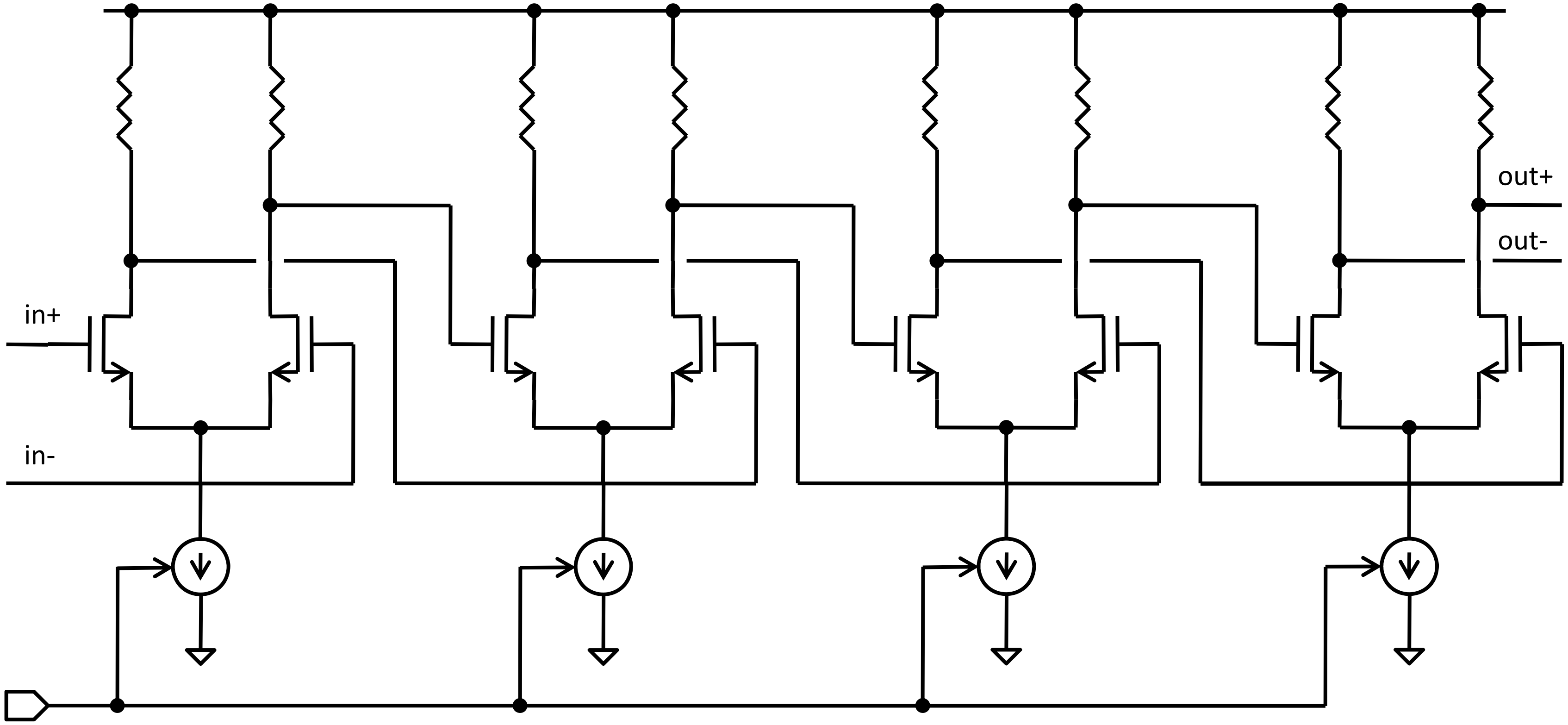


No pre- or de-Emphasis

Pre-Emphasis: $\Delta I_{rise} = \Delta I_{fall} = +13$ mA
→ Rise/fall times decreased

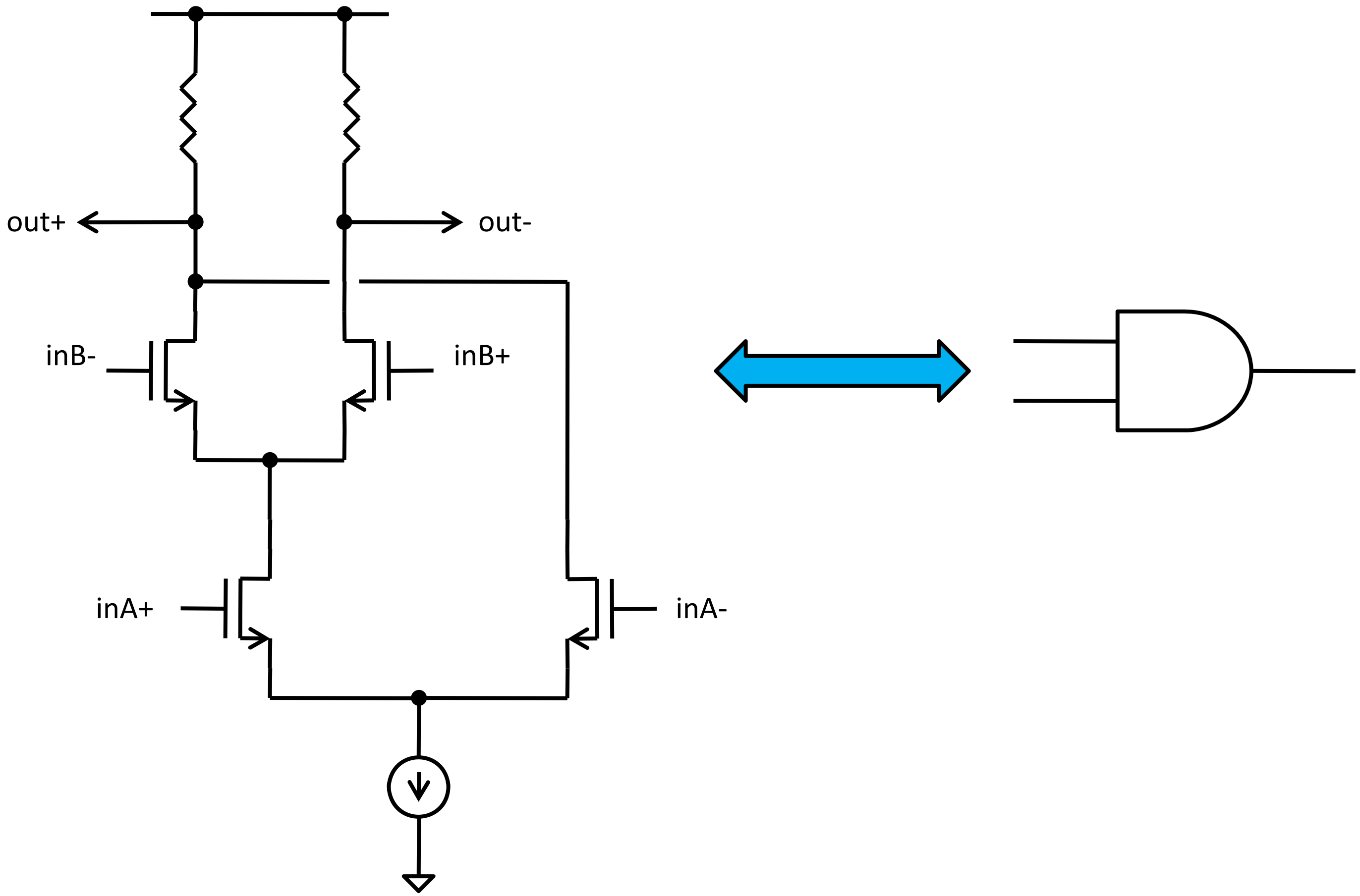


Delay Line



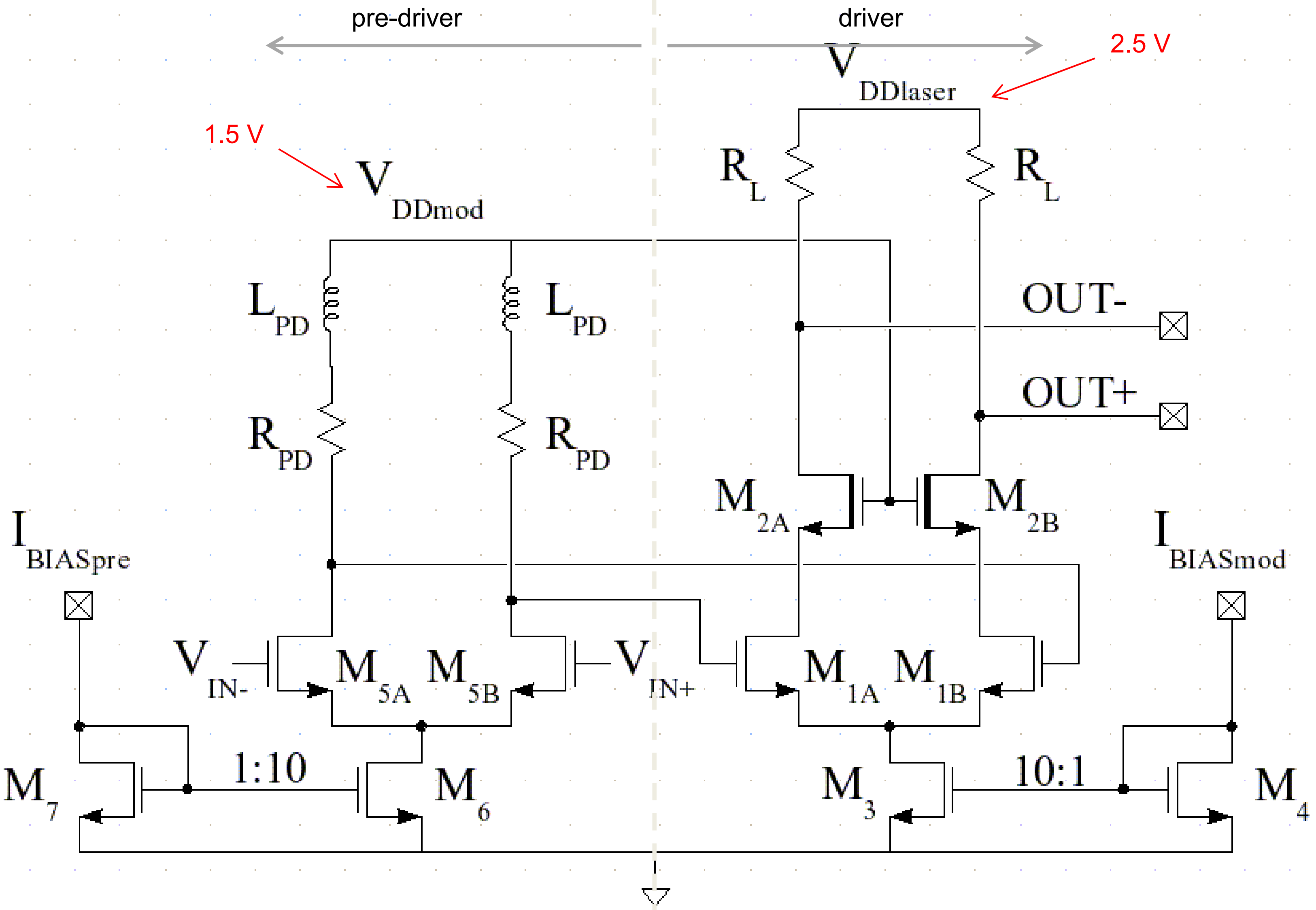
Paulo.Moreira@cern.ch

AND Gate



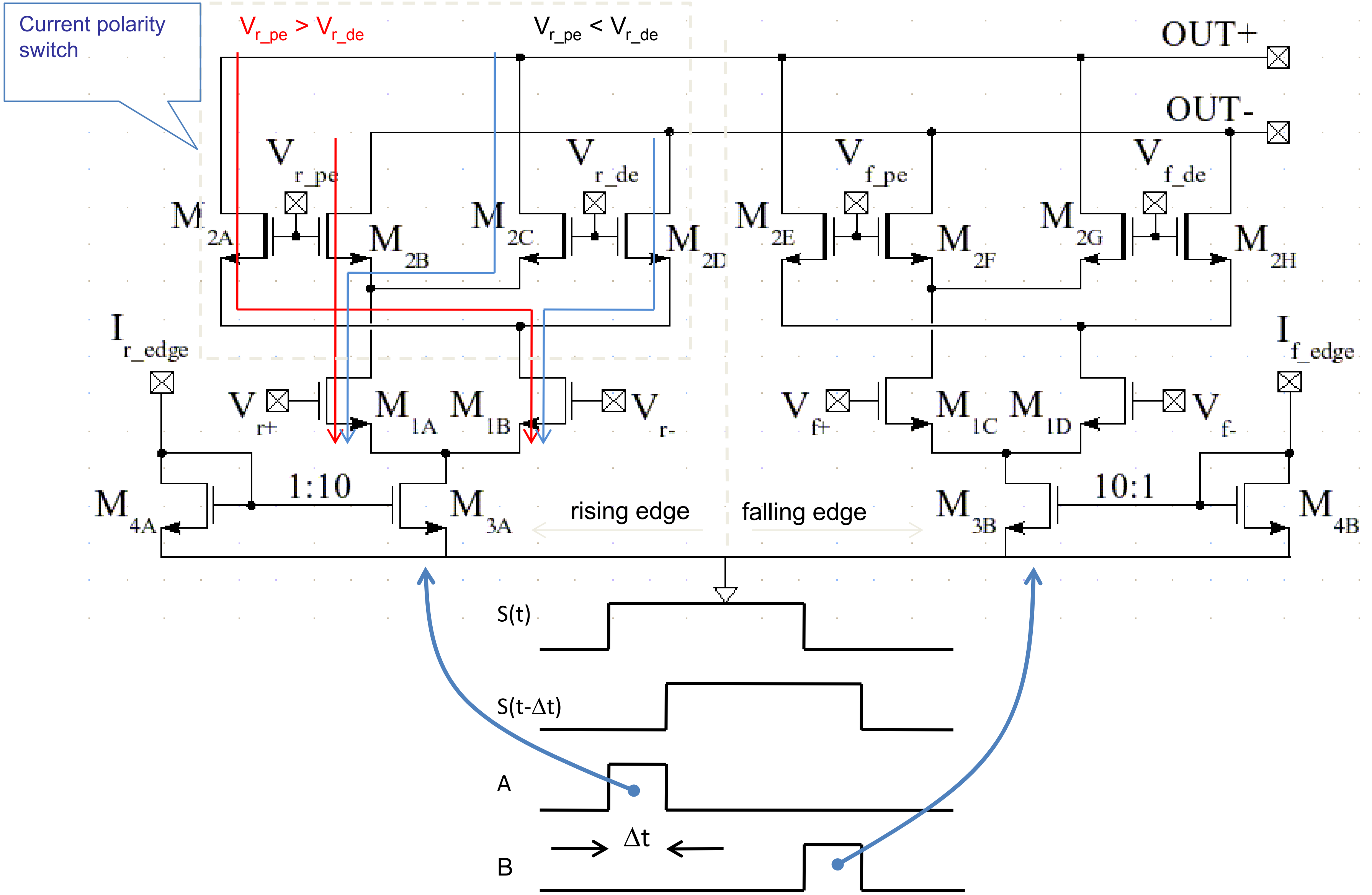
Only if (inA+ > inA-) & (inB+ > inB-) is out+ > out-

Pre-Driver & Driver

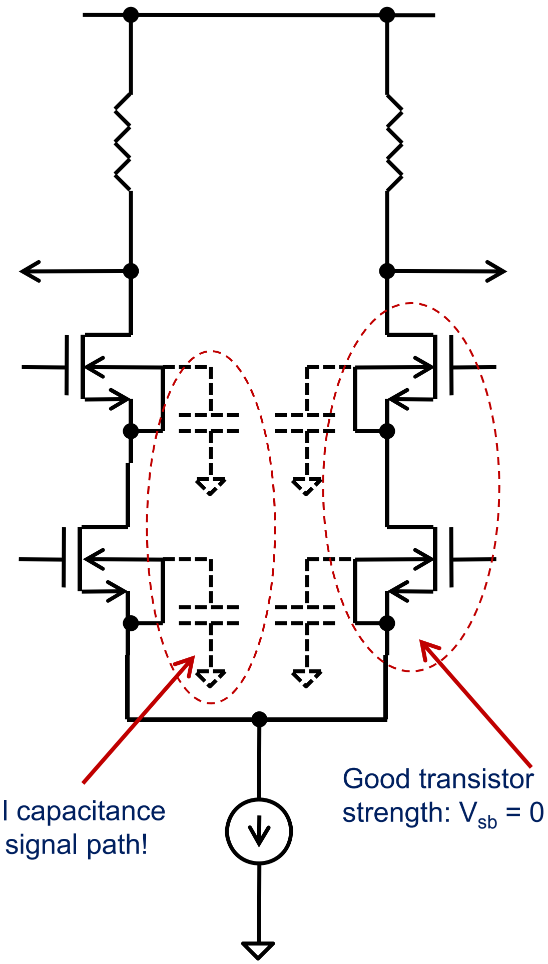


Paulo.Moreira@cern.ch

Emphasis / De-Emphasis Drivers

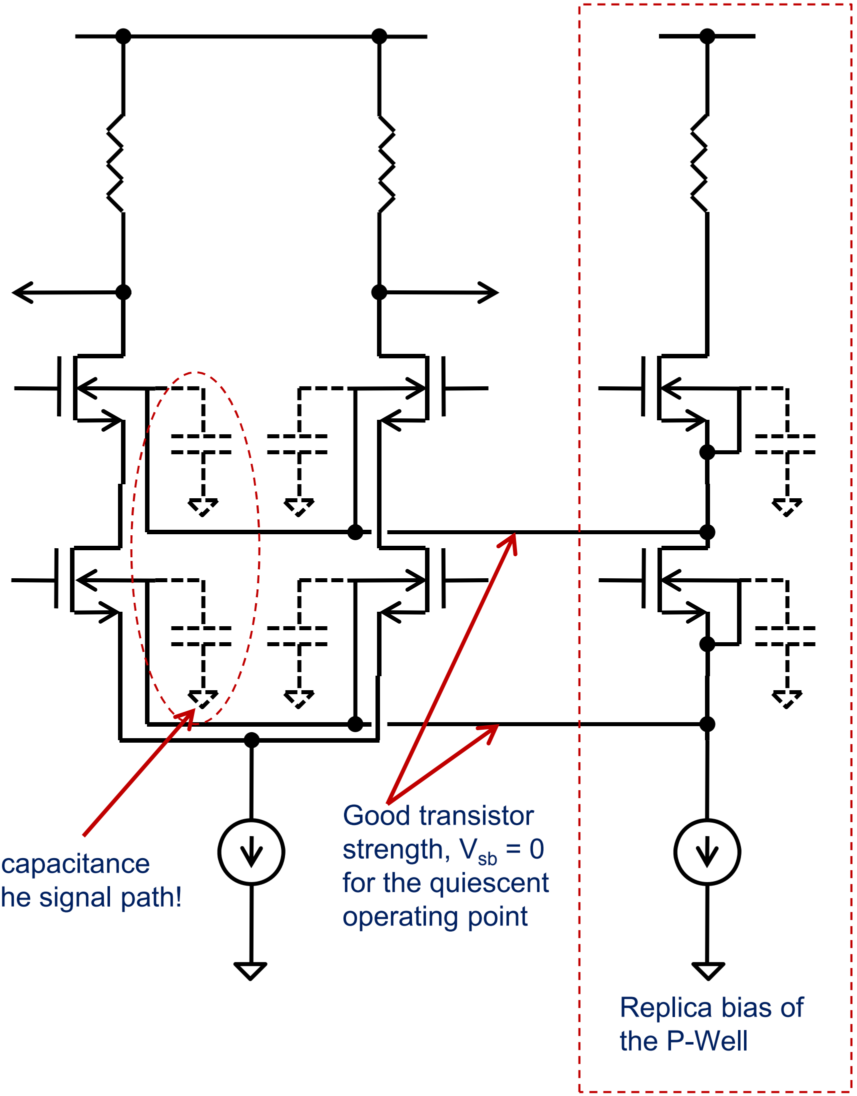


Avoid Parasitic Capacitances in the Signal Path



P-Well capacitance
in the signal path!

Good transistor
strength: $V_{sb} = 0$



P-Well capacitance
out of the signal path!

Good transistor
strength, $V_{sb} = 0$
for the quiescent
operating point

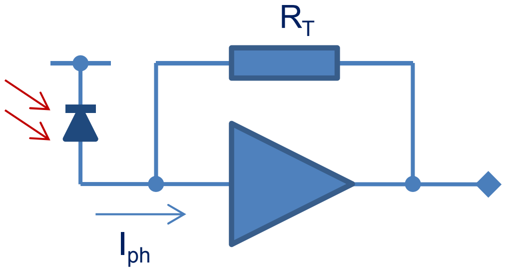
Replica bias of
the P-Well

Paulo.Moreira@cern.ch

PIN – RECEIVER DESIGN

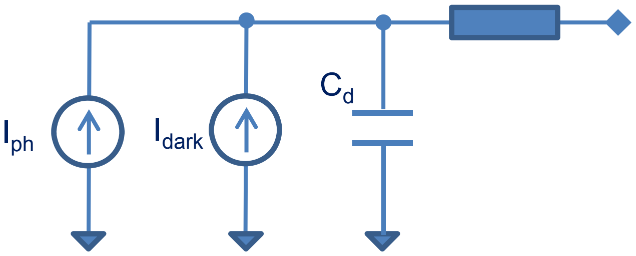
PIN-Diodes In HEP

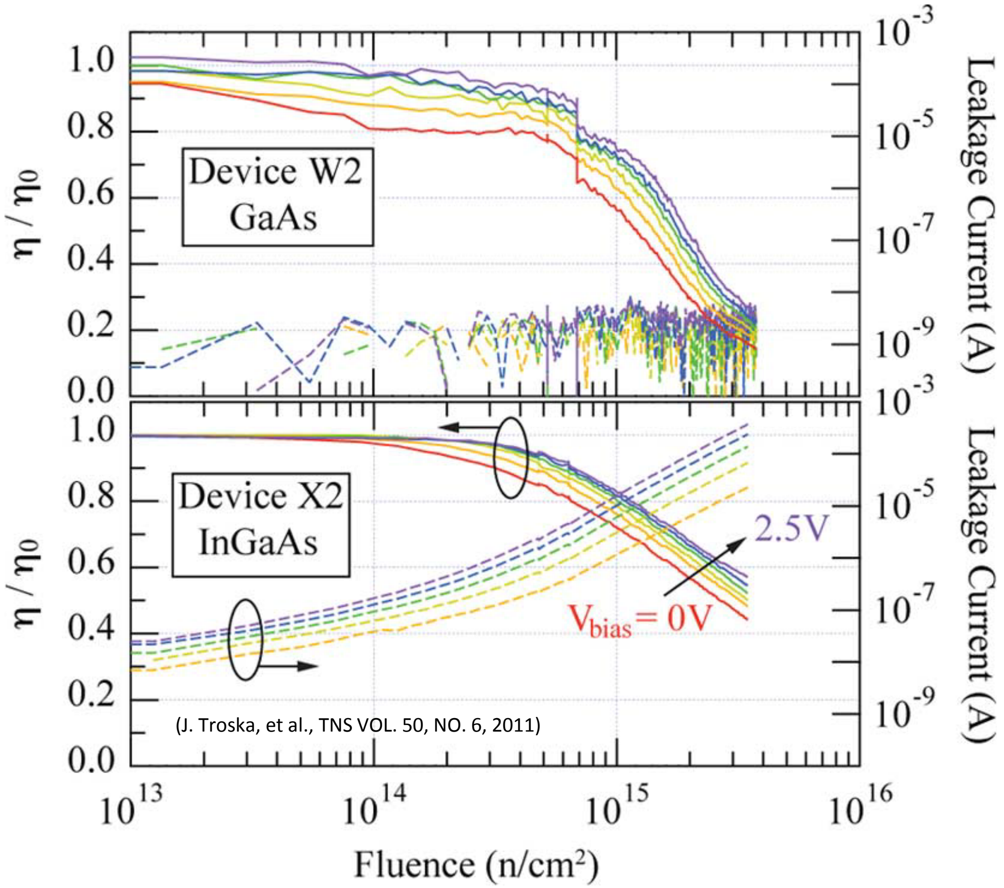
$V_{out} = P_{opt} \times \mathbf{R} \times R_T$



Responsivity [A/W]

$I_{ph} = P_{opt} \times \mathbf{R}$





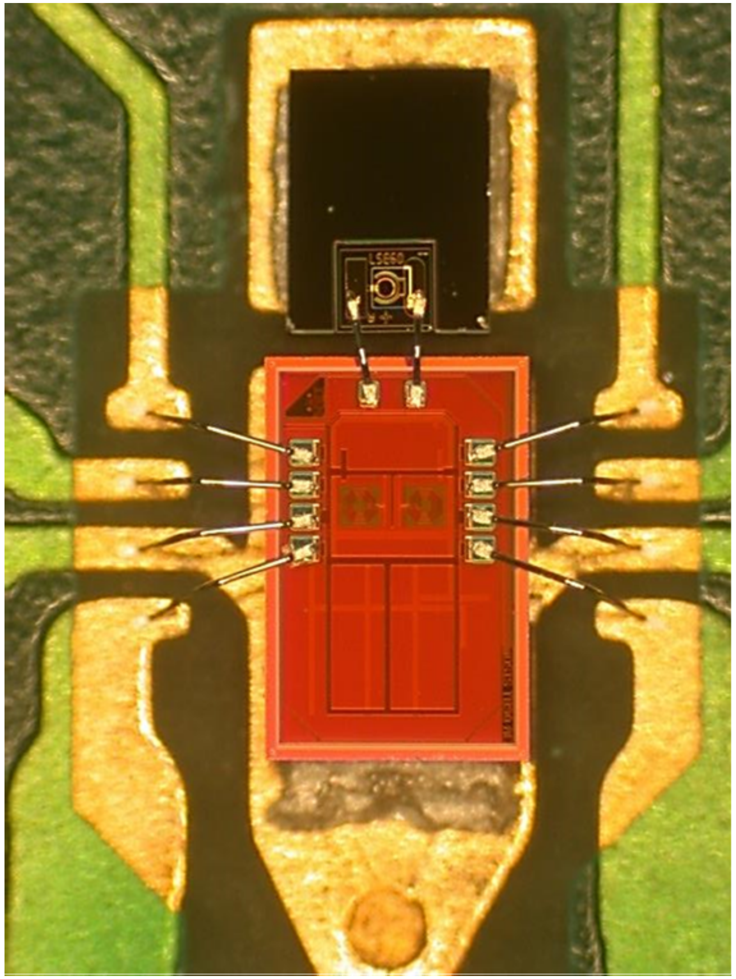
(J. Troska, et al., TNS VOL. 50, NO. 6, 2011)

Paulo.Moreira@cern.ch

GBTIA – Giga Bit Transimpedance Amplifier

Main specs:

- Bit rate:
  - 5 Gb/s (min)
- Sensitivity:
  - 20 µA P-P (10^{-12} BER)
- Total jitter:
  - < 40 ps P-P
- Input overload:
  - 1.6 mA (max)
- Dark current:
  - 0 to 1 mA
- Supply voltage:
  - 2.5 V
- Power consumption:
  - 250 mW
- Die size:
  - 0.75 mm × 1.25 mm



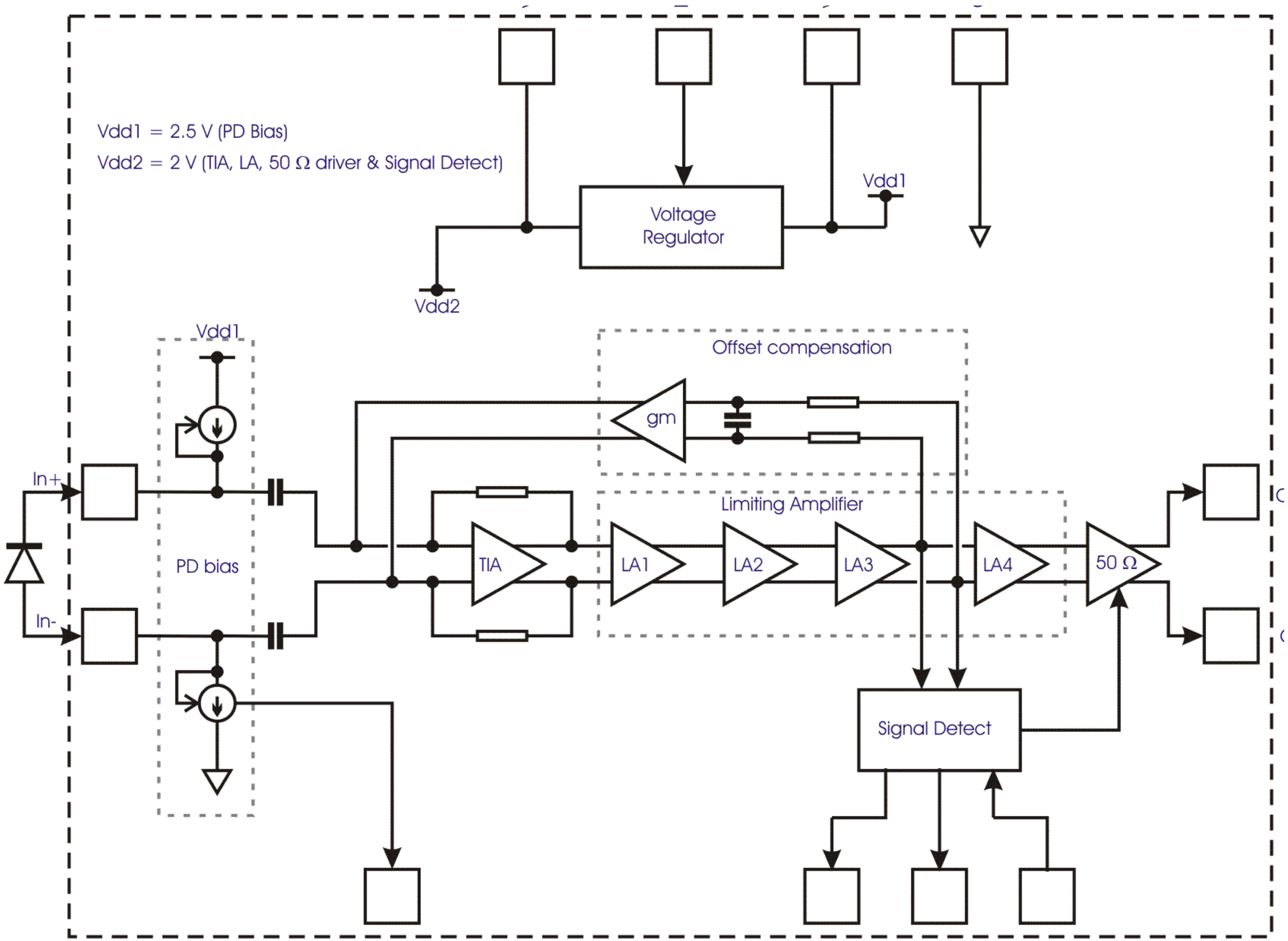
Paulo.Moreira@cern.ch

GBTIA Design Overview

- Integrating the TIA with the LA is potentially risky:
  - Single ASIC
  - Crosstalk noise trough the supply and substrate
- Fully differential architecture:
  - Good tolerance to the supply noise
  - Input noise larger than for single ended
  - The power consumption is higher
- The photodiode is AC coupled to the TIA
  - Dark current deviated from the TIA
  - Low value of the low cut off frequency
  - Parasitics of the integrated coupling capacitances limit the bandwidth
- Transimpedance amplifier (TIA)
  - Wide bandwidth
  - Moderate noise
  - Stability
- Limiting Amplifier and output buffer (LA)
  - High gain
  - Wide bandwidth
  - Offset compensation
- Additional features :
  - Internal voltage regulator (with enable/disable control)
  - Average power indicator

Paulo.Moreira@cern.ch

GBTIA Architecture

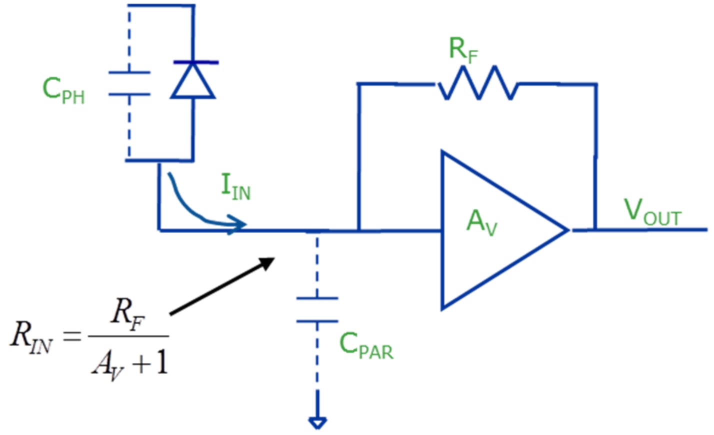


Vdd1 = 2.5 V (PD Bias)
Vdd2 = 2 V (TIA, LA, 50 Ω driver & Signal Detect)

Paulo.Moreira@cern.ch

Transimpedance Amplifier

- Shunt feedback amplifier is widely used for high speed receiver designs
- To increase the bandwidth :
  - Decrease the feedback resistor
  - Increase the amplifier open loop gain
  - Decrease the input node capacitance
- To minimize the thermal noise :
  - Increase the feedback resistor
  - Decrease the input node capacitance
  - Increase the amplifier transconductance



$$R_{IN} = \frac{R_F}{A_V + 1}$$

$$C_T = C_{PH} + C_{PAR} + C_{IN}$$

$$Z_T = \frac{v_{out}}{i_{in}} = \frac{-A_V}{A_V + 1} \frac{R_F}{1 + j\omega C_T \dfrac{R_F}{1 + A_V}}$$

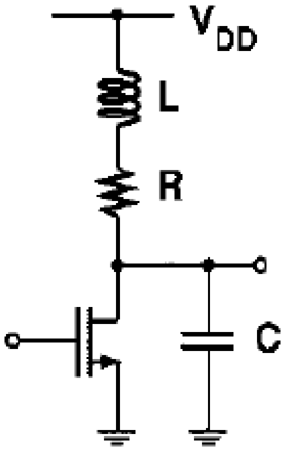
$$BW = \frac{1 + |A_V|}{2\pi . R_F C_T}$$

$$i_{n,in}^{\ 2} = \frac{4kT}{R_F} + \frac{4kT\Gamma}{R_F^{\ 2}} \left(\frac{1}{g_m} + \frac{(2\pi . f . R_F C_T)^2}{g_m} \right)$$

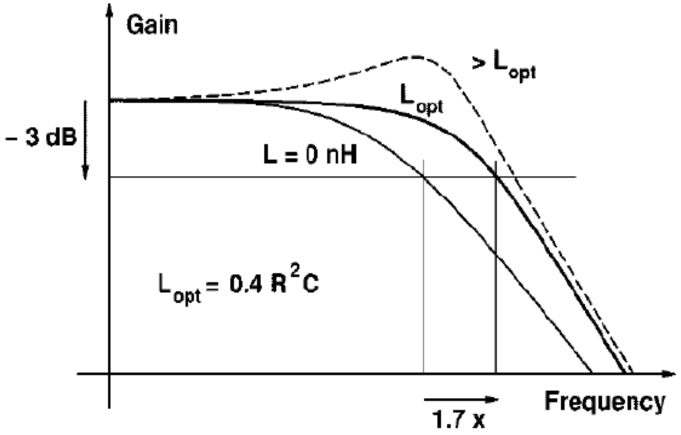
Bandwidth Enhancement by Inductive Peaking

- In order to maintain a low level of noise while keeping a large bandwidth, shunt peaking is used to bust the open loop gain at high frequencies
- Shunt peaking
  - Inductance in series with the load resistance
  - Enhances the bandwidth
  - The frequency response is characterized by the ratio **"m"**

$$L = m.R^2.C$$



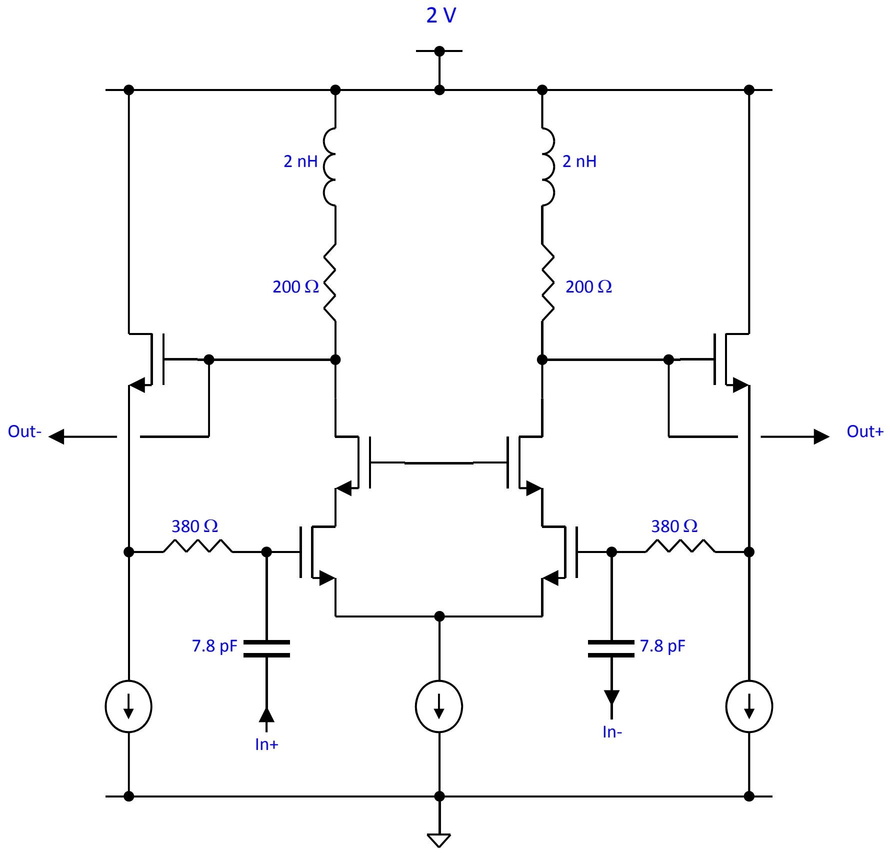
Factor m	Normalized f_{3dB}	Response
0	1.00	No shunt peaking
0.32	1.60	Optimum group delay
0.41	1.72	Maximally flat
0.71	1.85	Maximum bandwidth



Paulo.Moreira@cern.ch

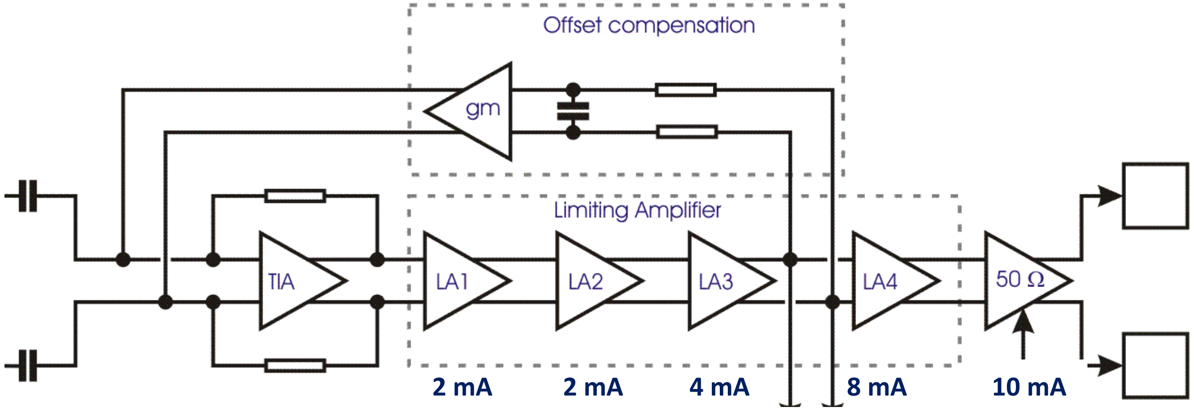
GBTIA Input Stage

- Differential
- Inductive peaking
  - The target bandwidth of 3.5 GHz is achieved for the worst case PVT (simulations including parasitic)
  - High transimpedance gain (R_F=380 Ω)
  - Low level of input referred noise
- Cascode
  - Reduces the Miller effect
- Current density is optimized
  - High current density needed to achieve high cut off frequency for the input transistor
  - Input transistor size optimized for an input capacitance of 700 fF
- 2 V supply required



Paulo.Moreira@cern.ch

Limiting Amplifier Requirements



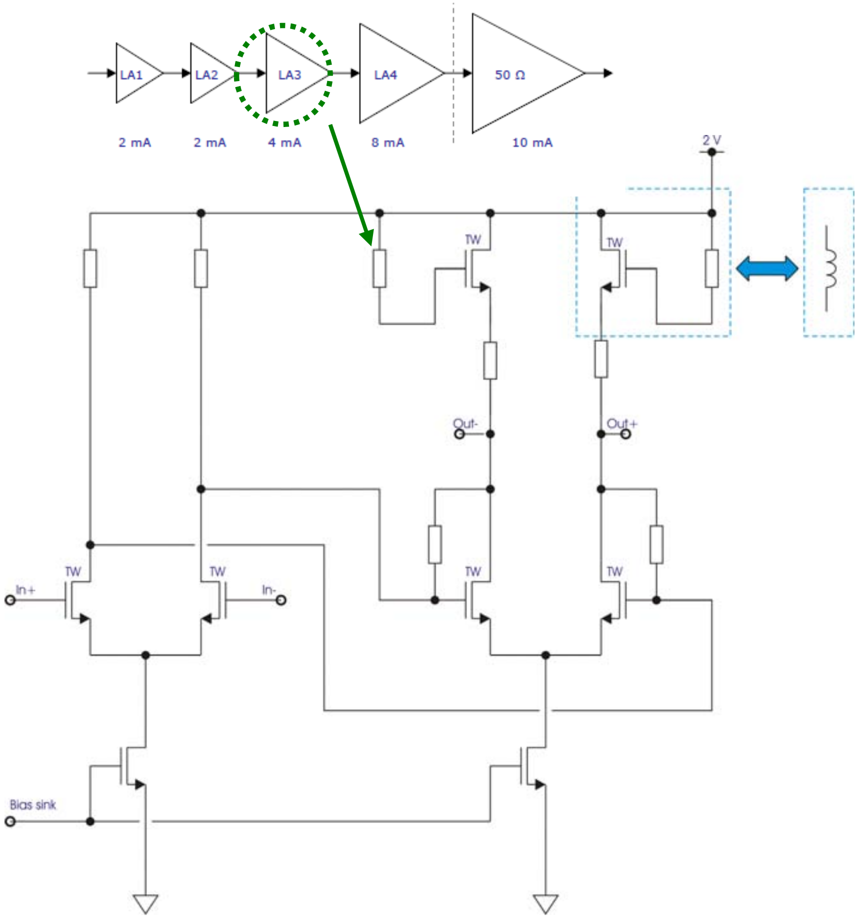
- Considering the sensitivity and the gain of the TIA - preamplifier:
  - In the sensitivity limit the TIA output voltage is 12 mV
  - The next ASIC requires a 400 mV pp input
  - The LA must provide a gain of 40 dB (28 dB in worst-case scenario)
  - The minimum overall bandwidth is 3.5 GHz
  - The noise contribution must be negligible :
    - The input referred noise must be lower than 850 µV RMS (12mV/14) for a BER of 10E-12
- The input capacitance of the LA must be sufficiently low so that it does not reduce the TIA bandwidth
- The number of stages is set to 5 (4 LA + a buffer)
- Offset cancellation is incorporated in LA block to prevent the mismatch in the differential amp from saturating the amplifier and “mask” the small input signals
- In order to maintain a wide bandwidth while delivering a large current to the load, the amplifiers stages in the LA are designed to have increasingly larger size and current:
  - Minimize the load capacitance seen by the previous stage
  - Allow bandwidth extension
- The gain of the first stage (LA1) is higher than the following stages to reduce its noise contribution.

Paulo.Moreira@cern.ch

Limiting Amplifier Stage

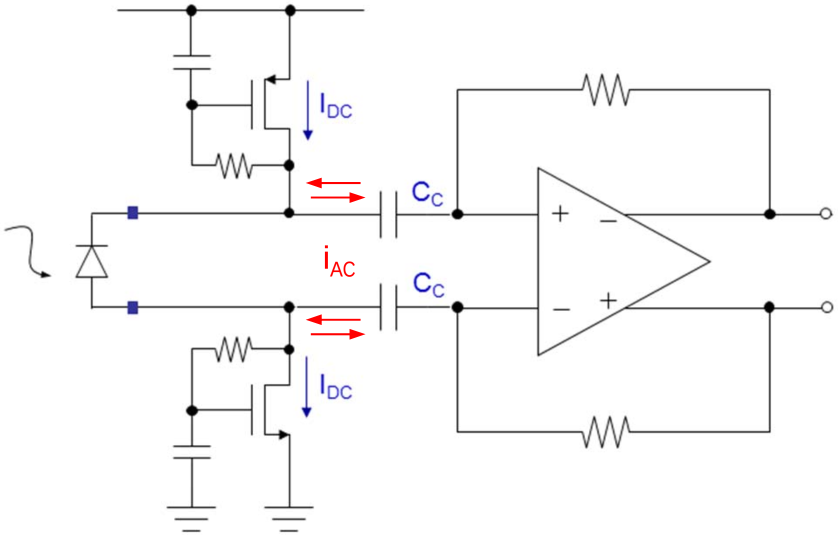
High bandwidth topology for each stage

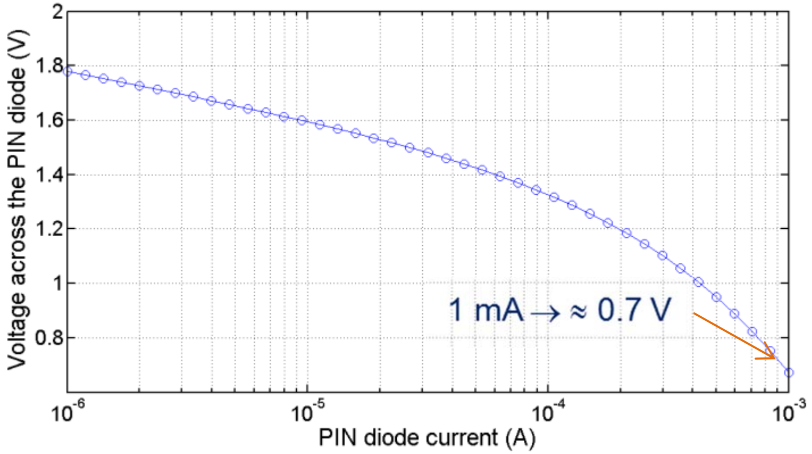
- Cherry and Hooper structure
- A g_m stage followed by shunt-feedback stage
- Second stage uses active “inductors”
- By active inductive peaking, the bandwidth is increased by 34% over a resistive loaded topology.



Pin-Diode Biasing (1/2)

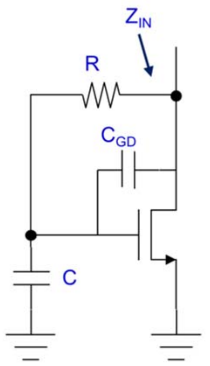
- TID effects result in large DC leakage currents across the PDs:
  - Can be as high as 1 mA!
- A large leakage current causes the PD bias voltage to decrease due to voltage drops in the biasing circuit!
- For high sensitivity and high bandwidth operation it is necessary to keep the reverse biasing voltage across the PD above a minimum:
  - 0.7 V for the devices used
- High receiver sensitivity also requires the impedance of the biasing circuit to be high, so that the signal current flows through the TIA and not the bias circuit
- The last two requirement are incompatible if simple bias resistors are used!
- To overcome the problem two “adaptive” current sources were used to bias the PD
- In the presence of relatively high leakage currents the circuit:
  - Maintains a “high” bias voltage
  - Represents a “high” impedance to the signal

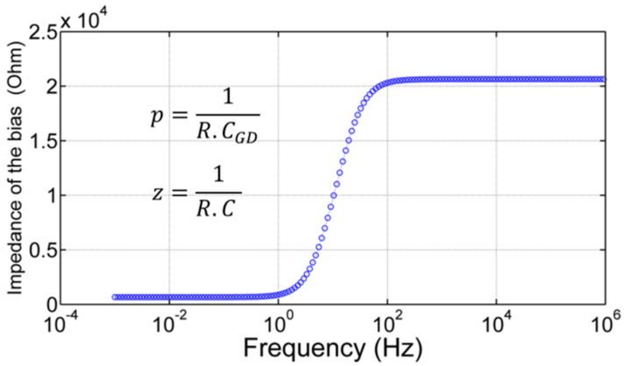


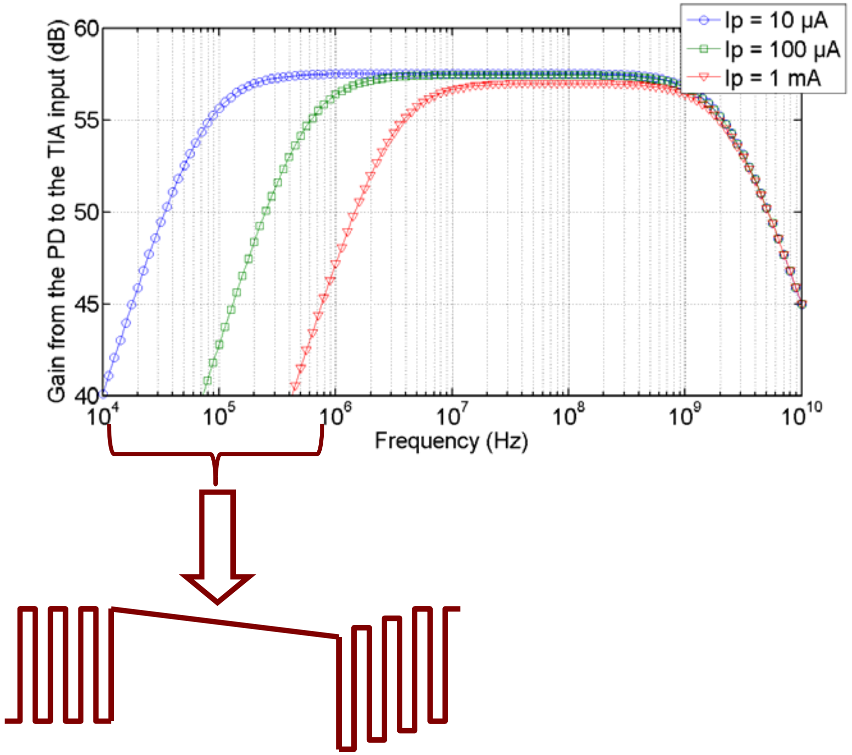


Pin-Diode Biasing (2/2)

- The bias circuit is capable of maintaining a “high” bias voltage:
  - across a large range of leakage currents (up to 1 mA)
- While being high impedance
- And having a low-band cut-off frequency:
  - Below 3 MHz in the worst case.
- A low cut-off frequency is required to avoid DC wander which will close the eye-diagram at the input of the receiver degrading the receiver sensitivity.
- For the codes used (8B/10B and 21-bit linear shift register scrambling) there is no penalty at 5Gb/s:
  - the low cut-off frequency of 3 MHz is a good compromise







GBTIA

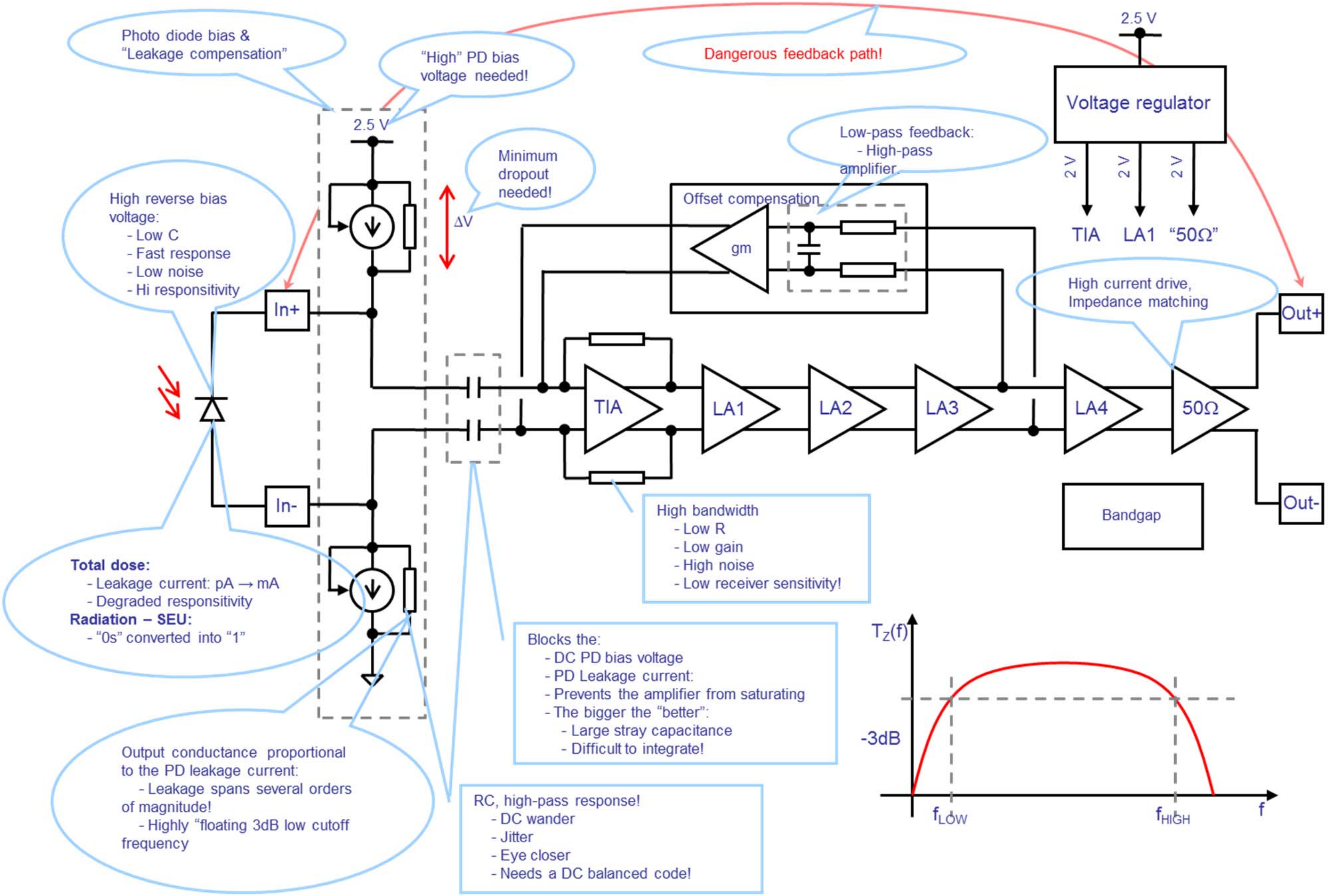


Photo diode bias & "Leakage compensation"

"High" PD bias voltage needed!

Dangerous feedback path!

Minimum dropout needed!

Low-pass feedback:
- High-pass amplifier.

High reverse bias voltage:
- Low C
- Fast response
- Low noise
- Hi responsitivity

High current drive, Impedance matching

High bandwidth
- Low R
- Low gain
- High noise
- Low receiver sensitivity!

Total dose:
- Leakage current: pA → mA
- Degraded responsitivity

Radiation – SEU:
- "0s" converted into "1"

Blocks the:
- DC PD bias voltage
- PD Leakage current:
- Prevents the amplifier from saturating
- The bigger the "better":
  - Large stray capacitance
  - Difficult to integrate!

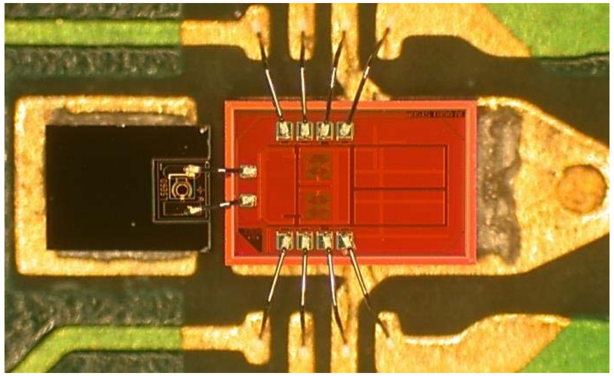
Output conductance proportional to the PD leakage current:
- Leakage spans several orders of magnitude!
- Highly "floating 3dB low cutoff frequency

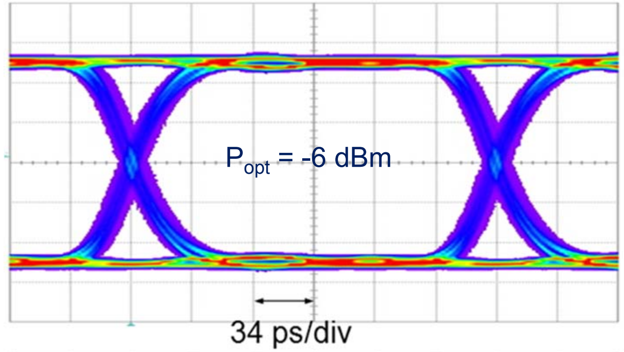
RC, high-pass response!
- DC wander
- Jitter
- Eye closer
- Needs a DC balanced code!

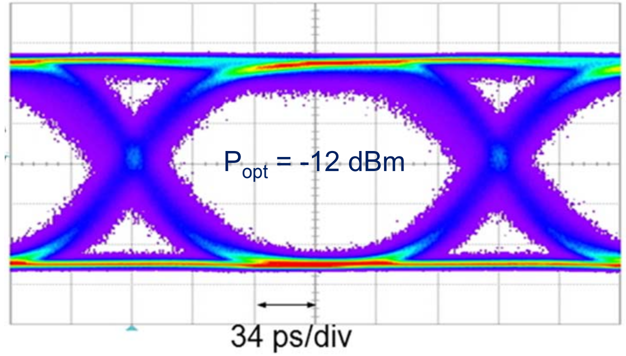
Paulo.Moreira@cern.ch

Measured Eye-Diagram

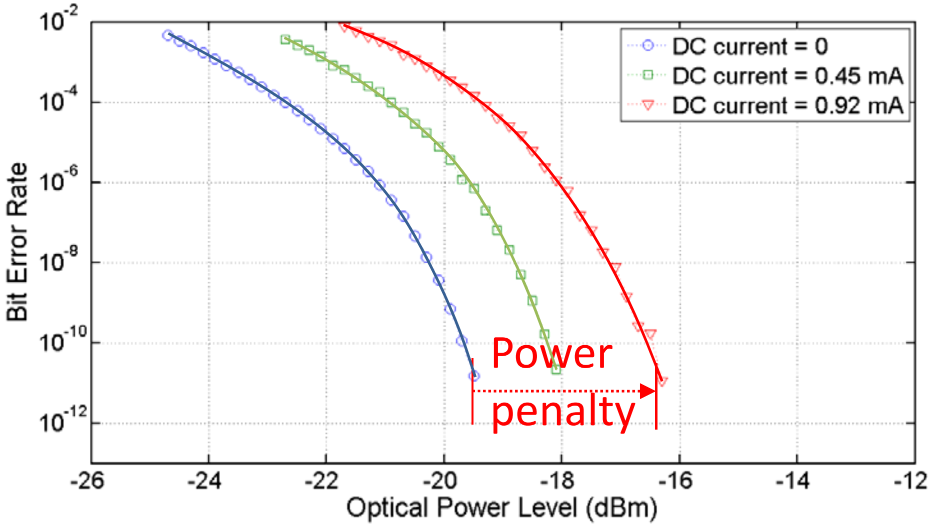
- Photodiode:
  - Responsitivity: 0.9 A/W @ λ = 1310 nm
  - Active area diameter: 60 µm
  - Illuminated on the top
  - Typical equivalent capacitance of 240 fF.
- The chip is wire-bonded to the photodiode:
  - To minimize the wire bond inductance and the input parasitic capacitance, the connection between the TIA and the photodiode is made very short and does not exceed 200 µm
- The eye diagram was measured at 5 Gb/s using a PRBS sequence of length 2^7-1
- For a –6 dBm input, the rise time is 30 ps and the total jitter is below 0.15 UI (30 ps) for a Bit Error Rate (BER) of 10^{-12}
- For –18 dBm input, the jitter is less than 0.55 UI (110 ps) and the rise time is 60 ps
- The output amplitude is virtually independent of the input power:
  - 800 mV differential





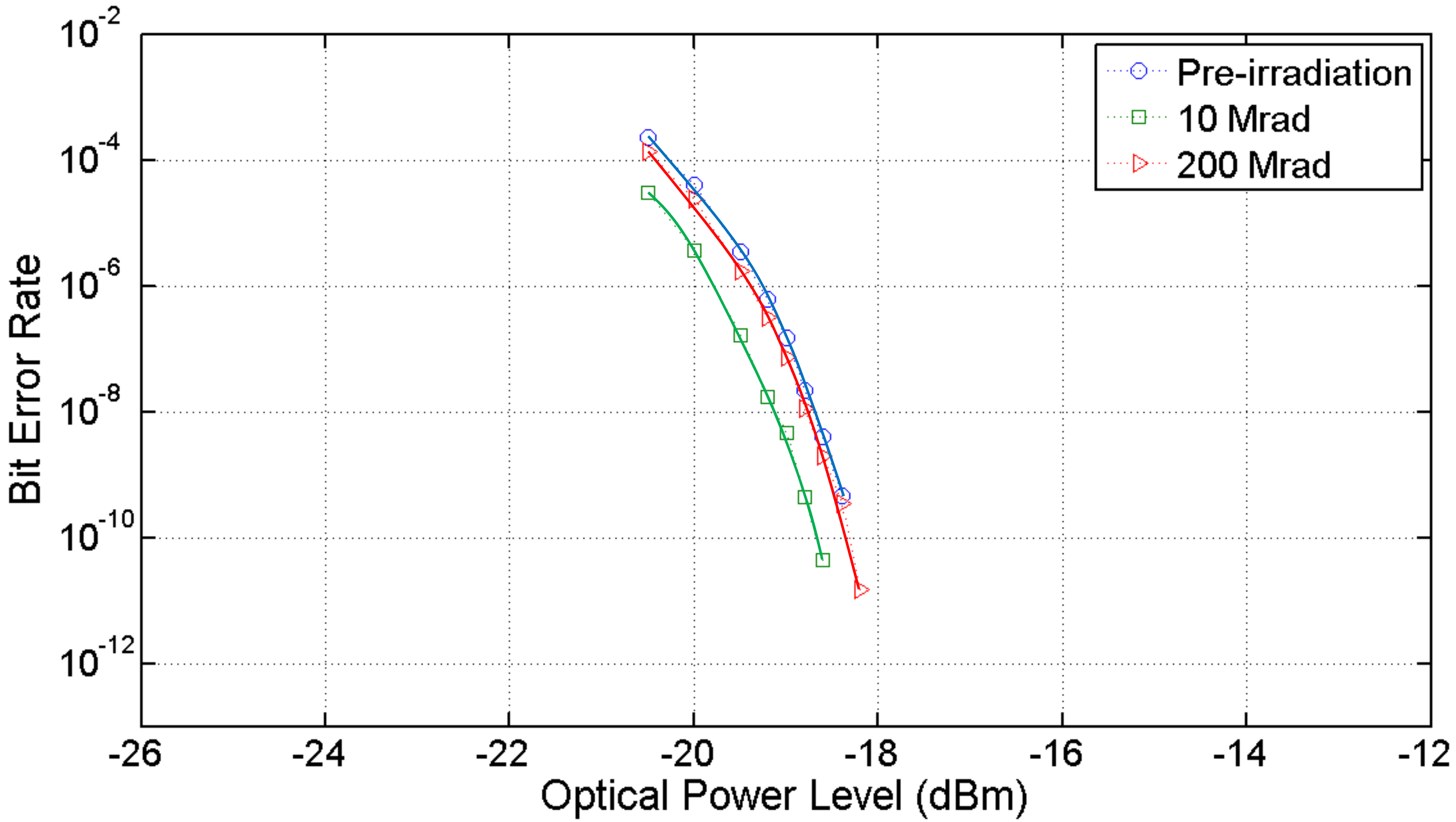


Influence of the optical DC level on the BER



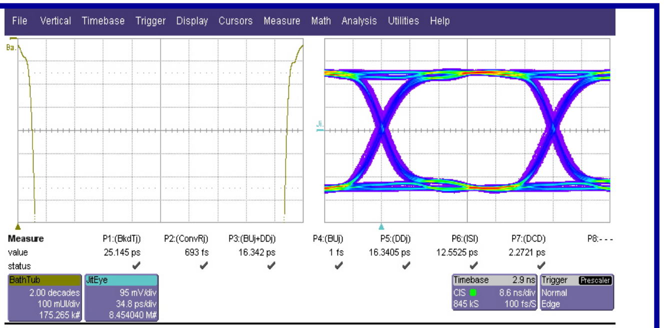
- The PD dark current can increase to 1 mA after 200 Mrad TID irradiation.
- To measure the influence the dark current, the PD was additional illuminated by a “DC” laser source.
- The integrated biasing circuit ensures a sufficient voltage across the PD.
- No noticeable degradation of the BER resulting from a higher low cut-off frequency was observed.
- However, the sensitivity degrades due to shot noise generated by the PD DC current.
  - The power penalty introduced by the shot noise of the DC level is ≈4 dB as expected

BER versus Total Dose



- Only the optical receiver chip (without the PD) was irradiated
- The PD was replaced by a passive network with input capacitance 500 fF.
- The chip was irradiated to a dose of up to 200 Mrad
- Only a marginal variation of the BER is observed

Eye-Diagram versus Total Dose



Measure	P1:(BkdTj)	P2:(ConvRj)	P3:(BUj+DDj)	P4:(BUj)	P5:(DDj)	P6:(ISI)	P7:(DCD)	P8:- - -
value	25.145 ps	693 fs	16.342 ps	1 fs	16.3405 ps	12.5525 ps	2.2721 ps	
status	✓	✓	✓	✓	✓	✓	✓	

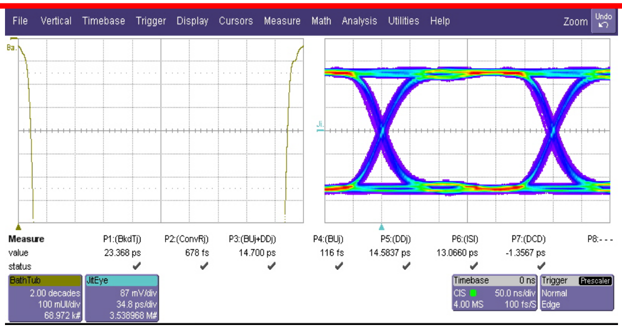
BathTub
2.00 decades
100 mUI/div
175.265 k#

JitEye
95 mV/div
34.8 ps/div
8.454040 M#

Timebase 2.9 ns
CIS 8.6 ns/div
845 kS 100 fs/S

Trigger Prescaler
Normal
Edge

Pre-rad eye diagram (input=500 mV)



Measure	P1:(BkdTj)	P2:(ConvRj)	P3:(BUj+DDj)	P4:(BUj)	P5:(DDj)	P6:(ISI)	P7:(DCD)	P8:- - -
value	23.368 ps	678 fs	14.700 ps	116 fs	14.5837 ps	13.0660 ps	-1.3567 ps	
status	✓	✓	✓	✓	✓	✓	✓	

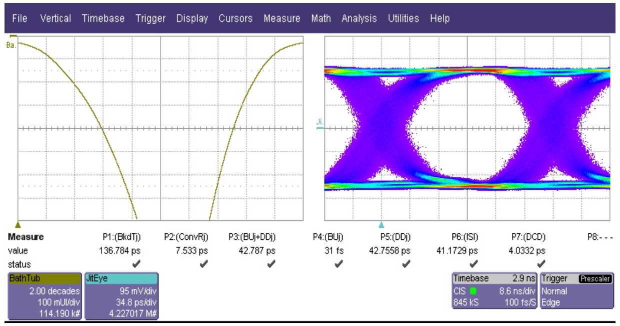
BathTub
2.00 decades
100 mUI/div
68.972 k#

JitEye
87 mV/div
34.8 ps/div
3.538968 M#

Timebase 0 ns
CIS 50.0 ns/div
4.00 MS 100 fs/S

Trigger Prescaler
Normal
Edge

200 Mrad eye diagram (input=500 mV)



Measure	P1:(BkdTj)	P2:(ConvRj)	P3:(BUj+DDj)	P4:(BUj)	P5:(DDj)	P6:(ISI)	P7:(DCD)	P8:- - -
value	136.784 ps	7.533 ps	42.787 ps	31 fs	42.7558 ps	41.1729 ps	4.0332 ps	
status	✓	✓	✓	✓	✓	✓	✓	

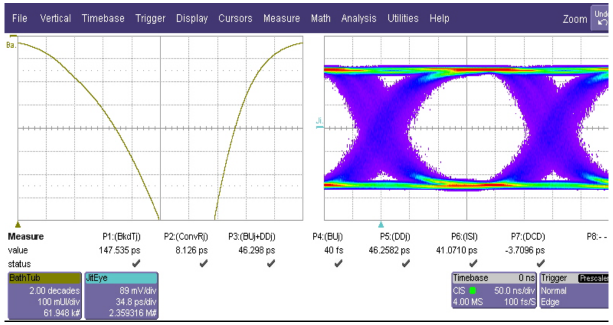
BathTub
2.00 decades
100 mUI/div
114.190 k#

JitEye
95 mV/div
34.8 ps/div
4.227017 M#

Timebase 2.9 ns
CIS 8.6 ns/div
845 kS 100 fs/S

Trigger Prescaler
Normal
Edge

Pre-rad eye diagram (input=50 mV)



Measure	P1:(BkdTj)	P2:(ConvRj)	P3:(BUj+DDj)	P4:(BUj)	P5:(DDj)	P6:(ISI)	P7:(DCD)	P8:- -
value	147.535 ps	8.126 ps	46.298 ps	40 fs	46.2582 ps	41.0710 ps	-3.7096 ps	
status	✓	✓	✓	✓	✓	✓	✓	

BathTub
2.00 decades
100 mUI/div
61.948 k#

JitEye
89 mV/div
34.8 ps/div
2.359316 M#

Timebase 0 ns
CIS 50.0 ns/div
4.00 MS 100 fs/S

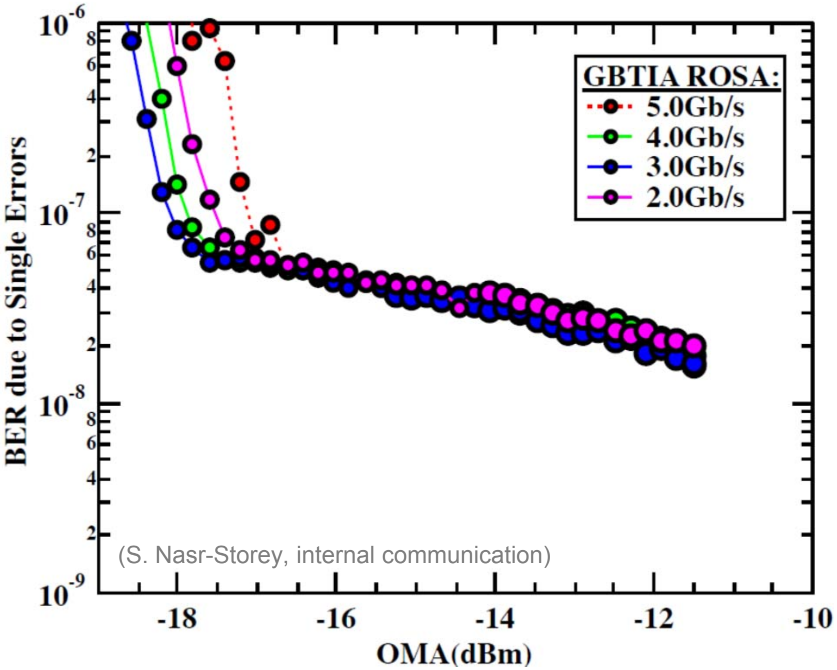
Trigger Prescaler
Normal
Edge

200 Mrad eye diagram (input=50 mV)

Paulo.Moreira@cern.ch

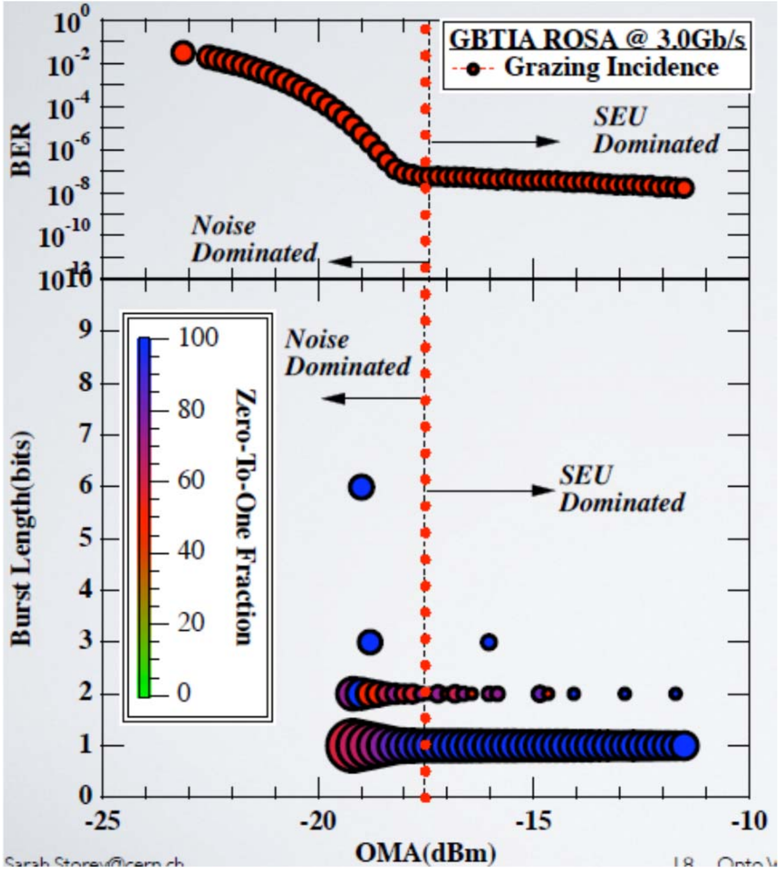
GBTIA: SEUs

- The Bit Error Rate, was measured to be independent of the bit rate
- Test flux:
  - 2 × 10^8 p/cm²/s
- Error rates < 10^{-9} can't be achieved even at high input optical power
- Forward error correction is thus mandatory



(S. Nasr-Storey, internal communication)

- SEU induced error "bursts" in the TIA
  - Biggest majority single or double bit errors
  - Longest observer 5 bits (SEU dominated region)



Paulo.Moreira@cern.ch

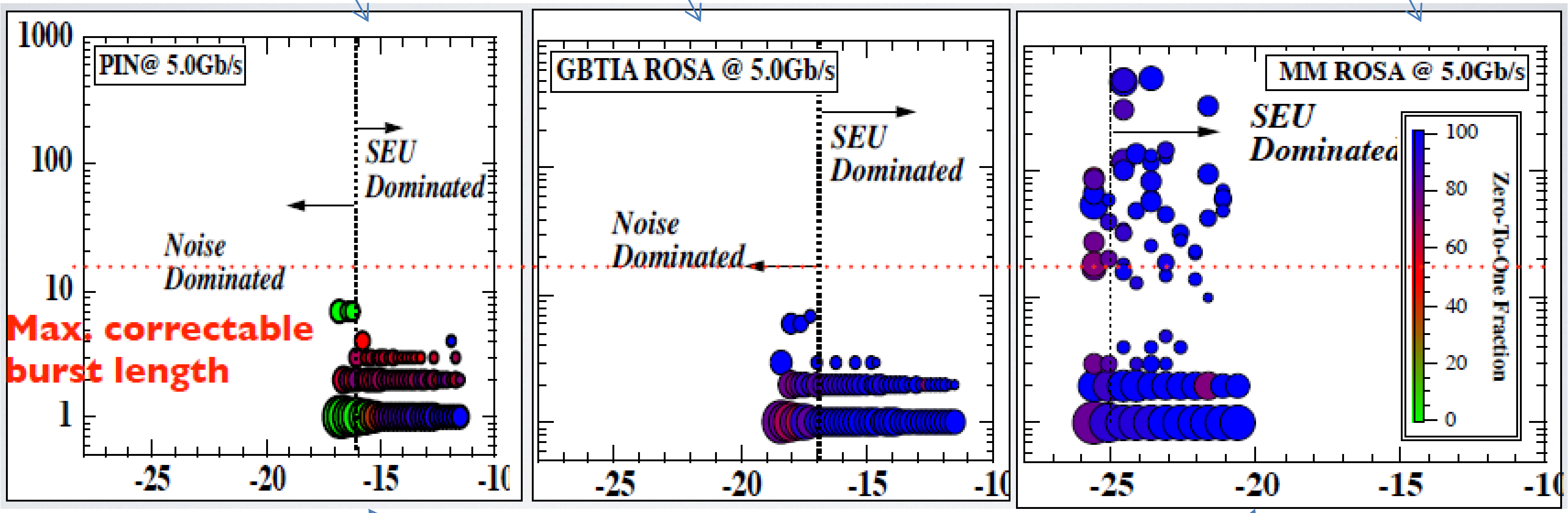
GBTIA vs Commercial PIN-Receiver (1/2)

Same PIN type in the tree cases

Pin with shielded TIA

GBTIA

Commercial PIN-Receiver



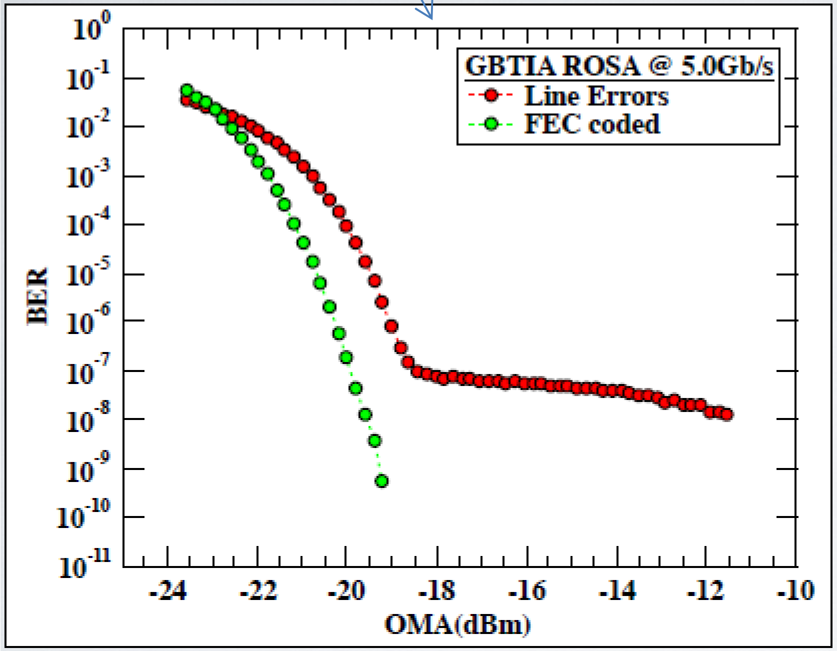
(S. Nasr-Storey, internal communication)

- Shielded PIN receiver and GBTIA display a very similar behaviour!
- Bursts 4 bits long in the SEU dominated region!

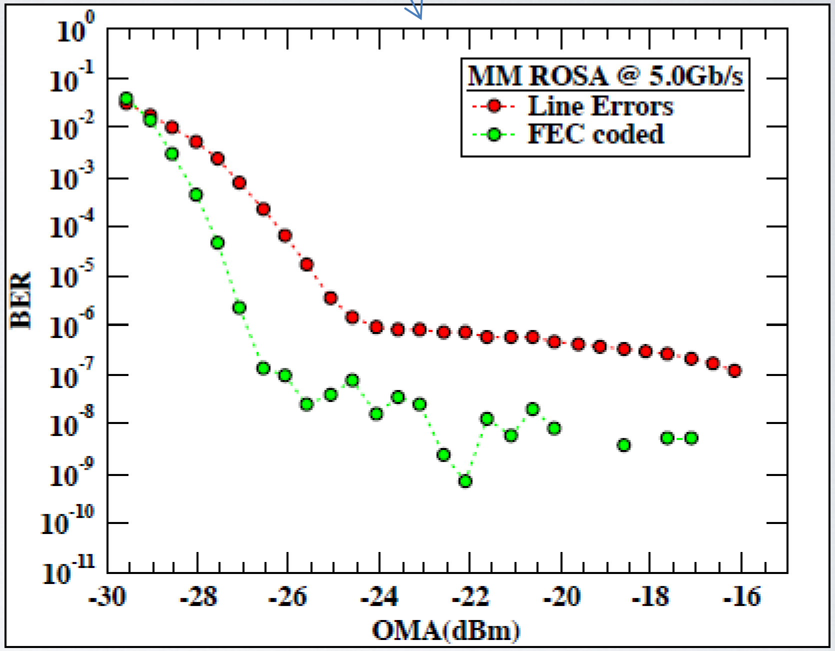
- For the commercial PIN-Receiver bursts longer than 100-bits long were measured!

Paulo.Moreira@cern.ch

GBTIA vs Commercial PIN-Receiver (2/2)



GBTIA



Commercial PIN-Receiver

- All errors corrected in the GBTIA-ROSA
- Error bursts always < 16 bits long

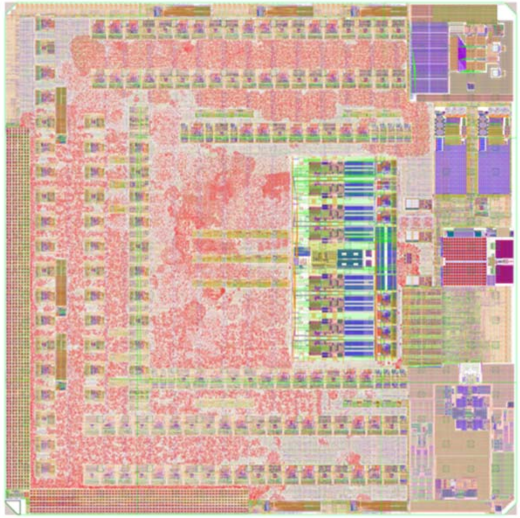
- FEC code “ineffective” for the commercial ROSA
- A good fraction of error bursts > 16 bits long

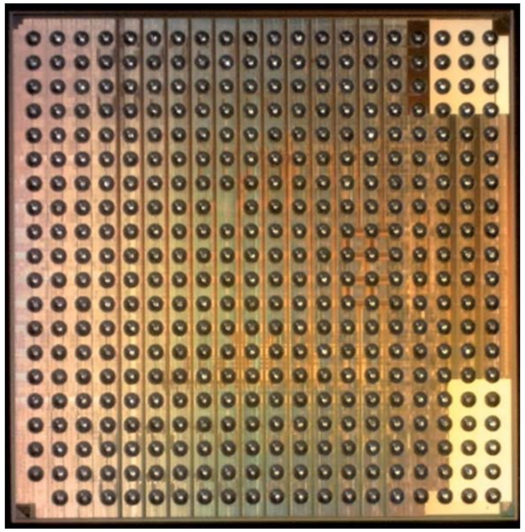
Paulo.Moreira@cern.ch

ASSEMBLING ASICS

The GBTX in Numbers

- ½ million gates
- Approximately:
  - 300 8-bit programmable registers (all TMR)
  - 300 8-bit e-Fuse memory
- Clock tree (chip wide):
  - 9 clock trees (all TMR)
  - Frequencies: 40/80/160/320 MHz
- 7 PLLs:
  - RX: CDR PLL + Reference PLL (2.4 GHz)
  - Serializer PLL (4.8 GHz)
  - Phase-Shifter PLL (1.28 GHz)
  - xPLL (VCXO based PLL, 80 MHz)
  - (2x) ePLL (320 MHz)
- 17 master DLLs:
  - 8 for phase alignment of the e-links
  - 8 for clock de-skewing
- 40 replica delay lines:
  - For phase alignment of the e-links
- 7 power domains:
  - Serializer (1.5V)
  - DESerializer (1.5V)
  - Clock Manager (1.5V)
  - Phase shifter (1.5V)
  - Core digital (1.5V)
  - I/O (1.5V)
  - Fuses (3.3V)



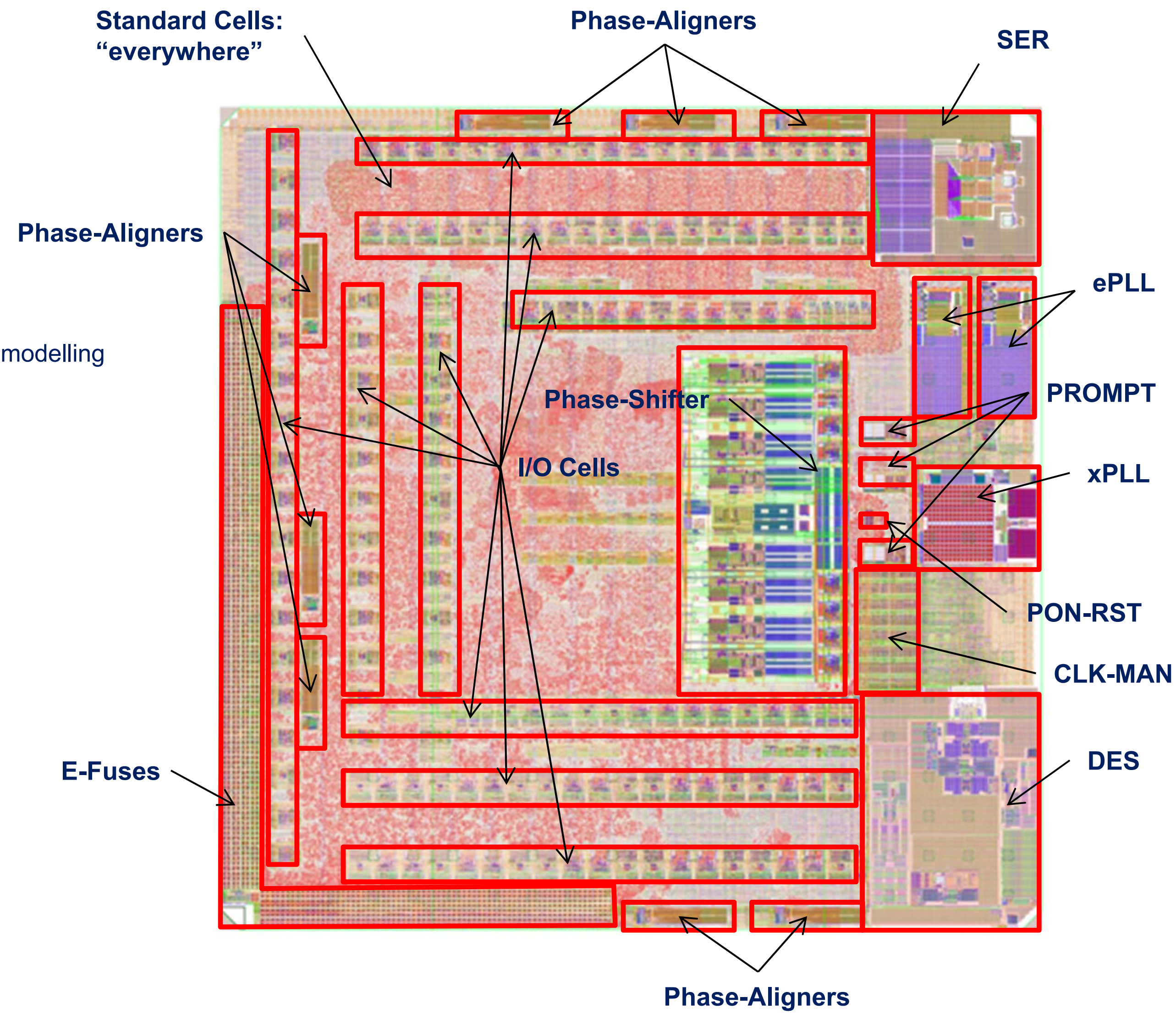


Paulo.Moreira@cern.ch

GBTX – Floor Plan and Modelling

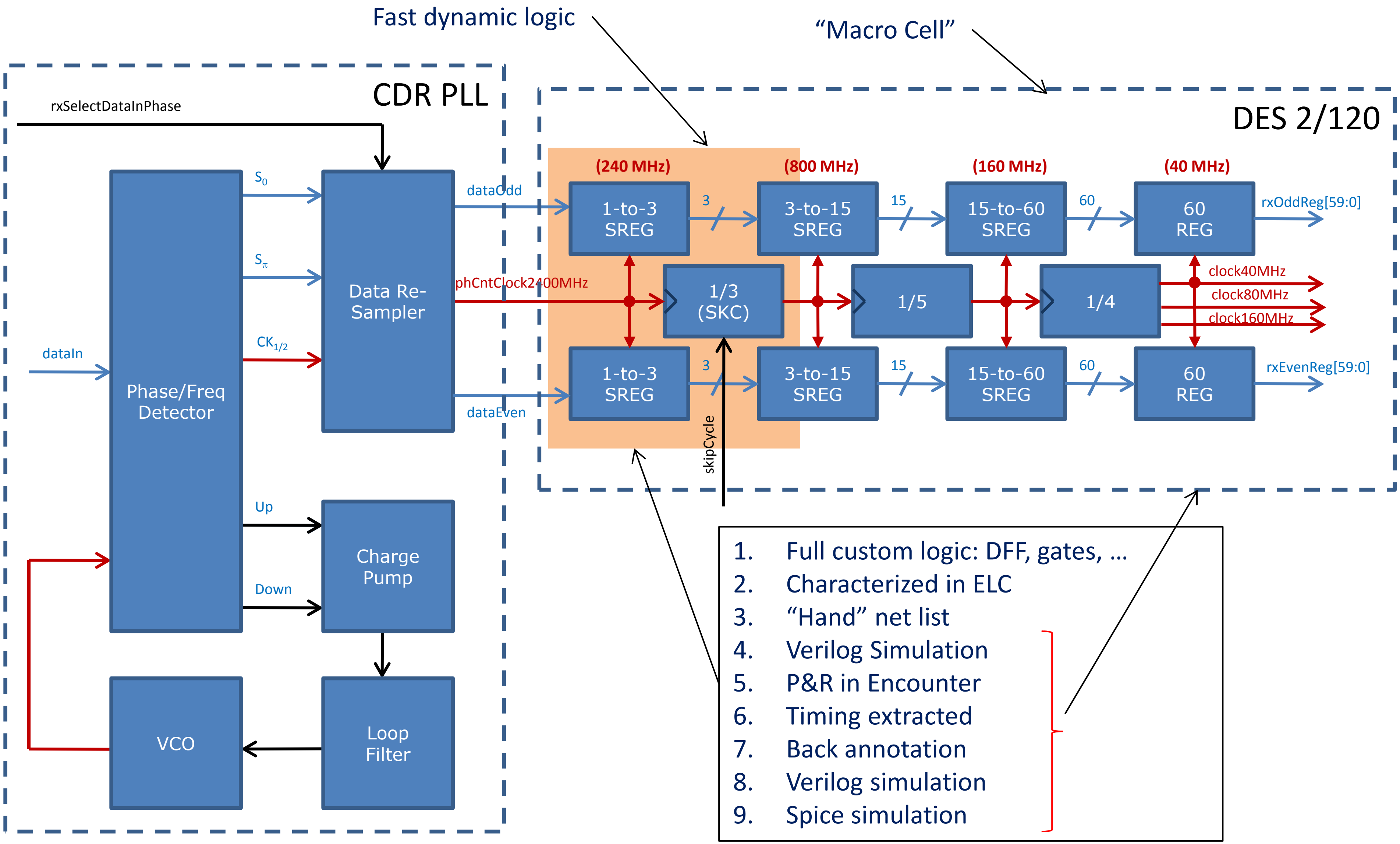
10 “Macro-cells”:

- Serializer:
  - Full custom
  - “Analogue-Verilog" modelling
- Clock Manager:
  - Standard cells
  - Verilog modelling
- DESerializer
  - Full custom
  - + “custom” digital
  - + standard cells
  - “Analogue-Verilog" + Verilog modelling
- XPLL
  - Full custom
  - “Analogue-Verilog" modelling
- EPLL (x2)
  - Full custom
  - “Analogue-Verilog" modelling
- Phase-Shifter
  - Full custom
- Phase Aligners (x8)
  - Full custom
  - “Analogue-Verilog" modelling
- Prompt (x3)
  - Full custom
  - Verilog modelling
- Power on reset
  - Full custom
  - Verilog modelling
- e-Fuses (x300)
  - Foundry IP
  - Verilog modelling



Paulo.Moreira@cern.ch

DESerializer



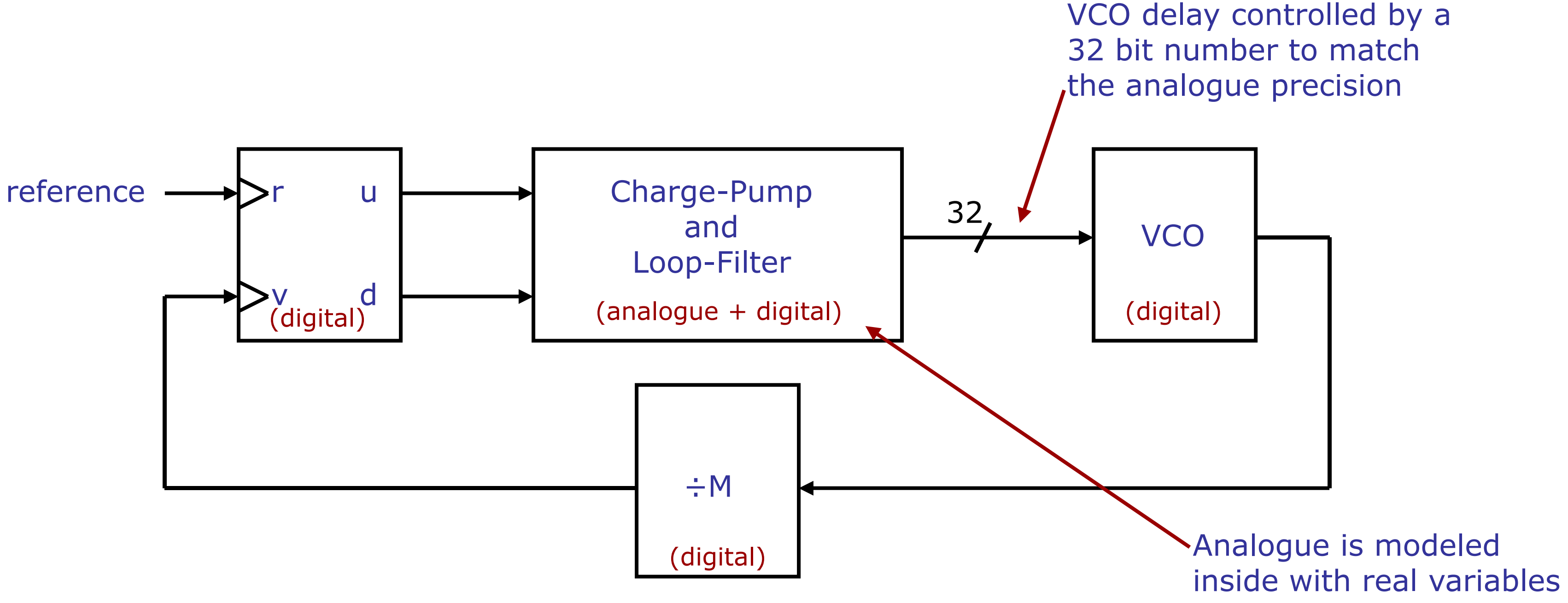
1. Full custom logic: DFF, gates, …
2. Characterized in ELC
3. "Hand" net list
4. Verilog Simulation
5. P&R in Encounter
6. Timing extracted
7. Back annotation
8. Verilog simulation
9. Spice simulation

Paulo.Moreira@cern.ch

“ANALOGUE-VERILOG”

PLL Modeling with (Plain) Verilog

Verilog → Hardware description language for digital circuits



Algorithm:

- Slice the time in very thin intervals (much smaller than T_{VCO})
- Make the time advance in these time increments
- At every time increment do:
    - Update the phase detector outputs
    - Calculate the new filter voltage according to the phase detector state
    - Update the VCO frequency as function of the filter voltage

Phase Detector and VCO

Phase Frequency Detector

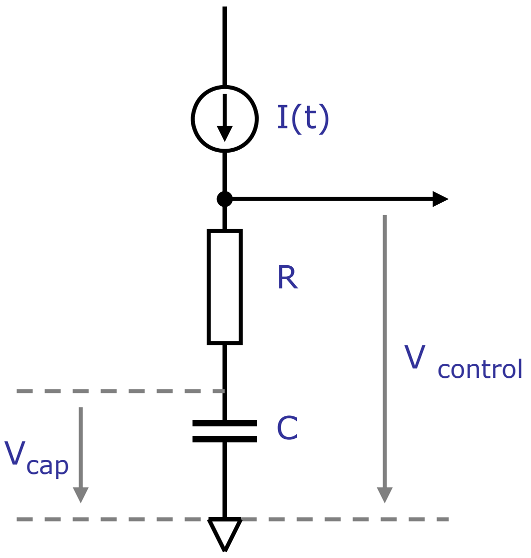
```
`timescale 1 fs / 1 fs
module ThreeStatePD (down, up, r, v);
output
               down,          // Early signal
               up;            // Late signal
input
               r,             // Reference input
               v;             // VCO input
wire           r, v, reset;
reg            up, down;
initial
               begin
               up = 0;
               down = 0;
               end
always @ (posedge r)
               up <= #1 1'b1;
always @ (posedge v)
               down <= #1 1'b1;
always @ (posedge reset)
               begin
               up   <= #1 1'b0;
               down <= #1 1'b0;
               end
assign #1
               reset = up & down;
endmodule
```

VCO

```
`timescale 1 fs / 1 fs
module VCO( delay_control,  vco_output);
input[31:0]    delay_control;
output vco_output;
reg reset;

initial begin
      reset=1;
      #100000;
      reset=0;
      end

delay_element10
   delay1( .in(~vco_output & ~reset), .delay_control(delay_control),
            .out(tap1)),
   delay2( .in(tap1), .delay_control(delay_control), .out(tap2)),
   delay3( .in(tap2), .delay_control(delay_control), .out(tap3)),
   delay4( .in(tap3), .delay_control(delay_control), .out(vco_output));
endmodule

// ---------------------------------------------------------------------- //
`timescale 1 fs / 1 fs
module delay_element10( in, delay_control, out);
input          in;
input[31:0]    delay_control;
output         out;
reg            out;
initial
   out = 1'b0;

always @(in)
               begin
               out <= #( delay_control/8 ) in; // delay =
                          delay_control/(2*number_of_delay_cells);
               end
endmodule
```

Paulo.Moreira@cern.ch

Loop Filter Simulation: R-C filter



- Simulation step: $t_n = t_{n-1} + \Delta t$
- The current I(t)
  - It is "imposed" on the circuit (it is the independent variable)
  - It is controlled by the phase-detector output
- For accuracy the time advances in small increments Δt:
  - Capacitor voltages change very little during a simulation step
  - The time integral of a function in the interval t_{n-1} to t_n can be approximated by:

$$\int_{t_{n-1}}^{t_n} f(t)dt = f(t_{n-1}) \cdot \Delta t \quad [1]$$

- For the simple R-C filter:

$$V_{cap}(t_n) = V_{cap}(t_{n-1}) + \frac{1}{C}\int_{t_{n-1}}^{t_n} I(t)\,dt \quad [2]$$

$$V_{control}(t_n) = R \cdot I(t_n) + V_{cap}(t_n) \quad [3]$$

- Simulation equations:

$$\boxed{V_{cap}(t_n) \cong V_{cap}(t_{n-1}) + \frac{1}{C} \cdot I(t_{n-1}) \cdot \Delta t} \quad [4]$$

$$\boxed{V_{control}(t_n) = R \cdot I(t_n) + V_{cap}(t_n)} \quad [5]$$

Notice that the time increments don't need to be constant (they have to be small). In that case, replace in the equation Δt by: $\Delta t_n = t_n - t_{n-1}$

$$\frac{T_{VCO}}{100} \leq \Delta t \leq \frac{T_{VCO}}{10}$$

Typically

Paulo.Moreira@cern.ch

Charge Pump & Loop Filter

Charge-Pump (includes the loop filter)

```
`timescale 1 fs / 1 fs
module ChargePump(up, down, delay_control);
.
.
.
`define DeltaVProportional (`Icp * `Rfilt)/1.0E-3      // PLL proportional term (in mV)
`define DeltaVIntegral (`Tref * `Icp / `Cfilt)/1.0E-3  // PLL integral term (in mV)
.
.              Variables used in ‘analogue’ computations declared as real
.
real           dv_capacitor,             // Differential capacitor voltage (in mV)
               control_voltage,          // Integral plus proportional control voltage (in mV)
               frequency,                // VCO frequency (in GHz)
               period,                   // VCO period (in ns)
               integral_term,            // loop control integral term (in mV)
               direct_term;              // loop control proportional term (in mV)
initial
               begin
.
.
.
               integral_term = `DeltaVIntegral/`IntegrationPoints;
               direct_term = `DeltaVProportional;
               integral_evaluation_time = 25000000/`IntegrationPoints;
               end
.
.
.
```

Paulo.Moreira@cern.ch

Charge Pump

Main loop, runs at regular time intervals

```
always
          begin
          if ( ~reset )
                    begin
                    // The integral and proportional terms are expressed as voltages
                    if (up)                       // Up refers to frequency not voltage
                              begin
                              if (`InitialFilterVoltage + dv_capacitor > `MinimumFilterVoltage)
                                        dv_capacitor = dv_capacitor - integral_term;
                              end
                    if (down)     // Up refers to frequency not voltage
                              begin
                              if (`InitialFilterVoltage + dv_capacitor < `MaximumFilterVoltage)
                                        dv_capacitor = dv_capacitor + integral_term;
                              end
                    end
          if ((~up & ~down) | (up & down))
                    control_voltage = `InitialFilterVoltage + dv_capacitor;
          else if (up)
                    control_voltage = `InitialFilterVoltage + (dv_capacitor - direct_term);
          else if (down)
                    control_voltage = `InitialFilterVoltage + (dv_capacitor + direct_term);
          frequency = `C0 + control_voltage*(`C1 + control_voltage*(`C2 +control_voltage*`C3));
          period = 1.0/frequency;
          DelayControl = period*1e6; // first convert the real into an integer
          #(integral_evaluation_time);
end
assign
          /* The variable "delay_control" represents the ring oscillator period in fs */
          #1 delay_control = DelayControl + vco_phase_noise;
endmodule
```

Calculate capacitor voltage increment

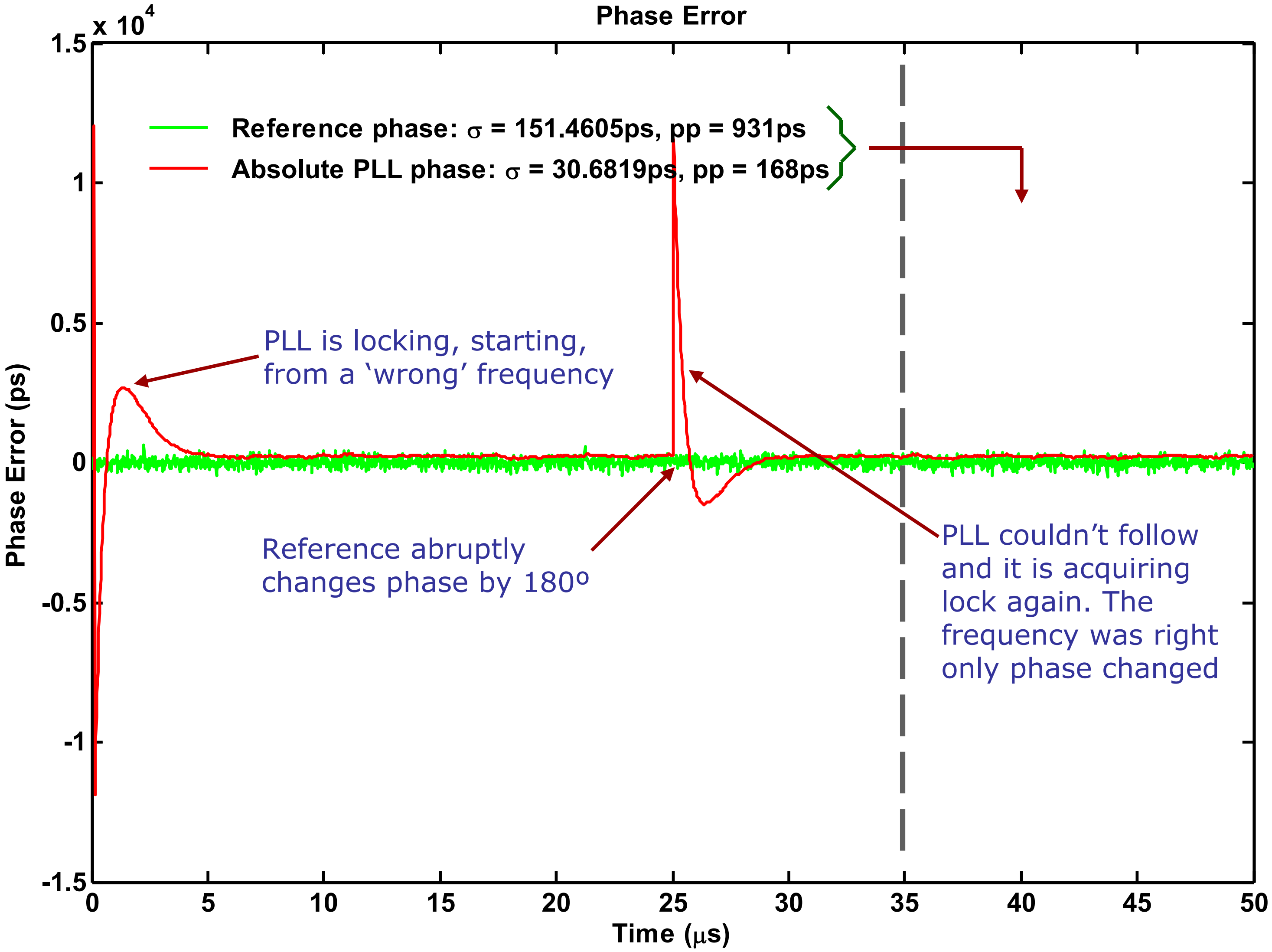
Update the capacitor voltage

Time step control

Update the VCO period

Add VCO phase noise

Simulation Example



PLL is locking, starting, from a 'wrong' frequency

Reference abruptly changes phase by 180º

PLL couldn't follow and it is acquiring lock again. The frequency was right only phase changed

Paulo.Moreira@cern.ch

HANDLING SEUs

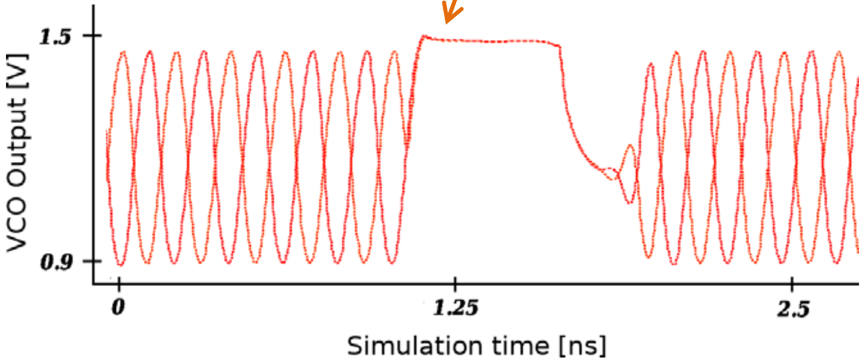
SEUs on Analogue

- In principle, Functional Interrupts (if they are self recovery) in analogue circuits have little consequence overall for a frontend ASIC or a data acquisition system:
  - What do they mean, for example, for an analogue signal representing the charge deposition of a particle?
  - What are the consequences for the overall system of an occasional wrong reading?
- However some SEUs on critical Analogue circuits might be important:
  - For example if a (global) bias circuit is disturbed and returning to the quiescent point takes a significant amount of time this might result on a functional interrupt that lasts long enough to affect the overall system performance!
- There is no universal answer:
  - 1st Make an estimate of how much charge will be typically deposited by radiation
  - 2nd Try to evaluate (realistically) what are the consequences for the full system (not just the circuit in question)
  - 3rd If it is important to make the circuit SEU robust:
    - Choose an appropriate solution
    - Evaluate the consequences for the system and the circuit
- Two generic approaches are often used in critical analogue circuits:
  - For slow circuits like e.g. bias circuits, "artificially" increase the capacitance of the sensitive nodes:
    - This will reduce the voltage disturbance due to the injected charge
    - Recovery might be slow so it important that the voltage disturbance is small
  - For fast circuits ("artificial") increase the bias currents (an thus the transistor sizes) so that the injected current represents a small fraction of the charge already "stored" in the circuit:
    - This results in additional power consumption

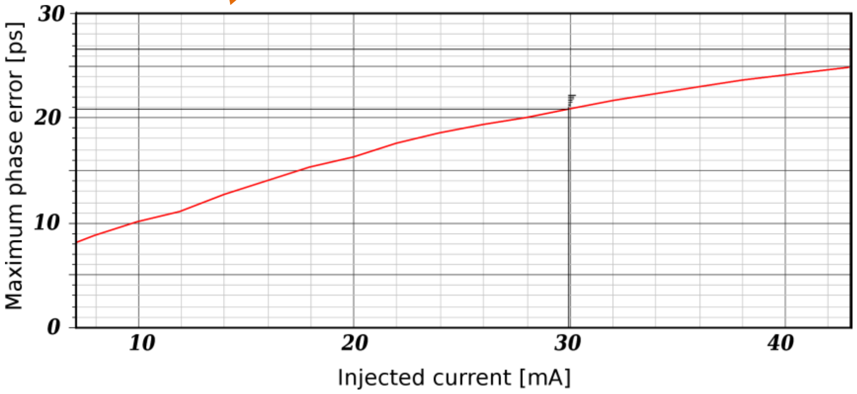
Paulo.Moreira@cern.ch

Example: SEUs in VCOs

- In data transmission systems, SEUs in VCOs are critical:
  - After recovery synchronization is required at the transmitter and/or receiver easily leading to system dead times of the order of the milliseconds!
- Depending on the injected charge, an SEU can:
  - Induce a phase jump
  - Stop the VCO oscillation for a few cycles!
- Example:
  - Technology: 130 nm CMOS
  - Oscillation frequency: 5 GHz
  - Differential delay-cell ring oscillator
  - SEUs modelled by current pulses injecting 300 fC in 10 ps
- Design criteria:
  - The VCO tail current is set so that an SEU causes a phase jump smaller than 10% of the bit period
  - For the PLL this “resembles” like a bit of excess jitter and can be easily handled
- Penalty:
  - High power consumption



Injected charge temporarily blocked the VCO oscillations



The magnitude of the phase jump is a function of the injected charge

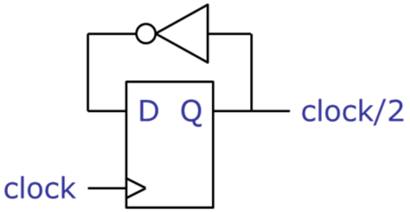
SEUs in Digital Circuits

- Over the years an extensive set of techniques have been developed to protect against SEUs in digital circuits that offer different levels of robustness:
  - One-hot encoded state machine:
    - Monitored by a “watchdog” circuit
  - Dice-cell flip-flops:
    - Depending on the protection level required, it might also need to be monitored
  - Hamming encoding
  - Triple Modular redundancy:
    - Triple voted registers
    - Triple voted registers with triplicated combinatorial logic
    - Triple voted registers with triplicated combinatorial logic and triplicated clock tree
  - …
  - Plus what your imagination can came up with…
- Depending how critical SEU induced errors are systems might require only protection of the state machines and/or the data path
- As for the analogue circuits not all digital circuits need the same level of protection:
  - Evaluate which circuits and to which level they need to be protected
  - In digital, SEU robustness is paid in terms of silicon area, power consumption, simulation and testing complexity!

Paulo.Moreira@cern.ch

Fast TMR Flip-Flops

- Clock dividers in high data rate SERDES circuits typically run at the highest possible speeds in a given technology:
  - They are based on fast flip-flops
- Dynamic flip-flops are intrinsically faster than static flip-flops enabling very high data rates
- However, they are even more sensitive to SEUs than static flip-flops
- It is them mandatory to use TMR techniques to protect the dynamic flip-flops from being upset:
  - E.g. a wrong count sequence can unlock a PLL resulting in a long dead time
- Triple voting adds additional delay in a circuit reducing the operation speed
- However, in dynamic FFs it is possible to vote an internal node:
  - This results in a much smaller speed penalty than if the voting would take place before or at the input stage of the flip-flop

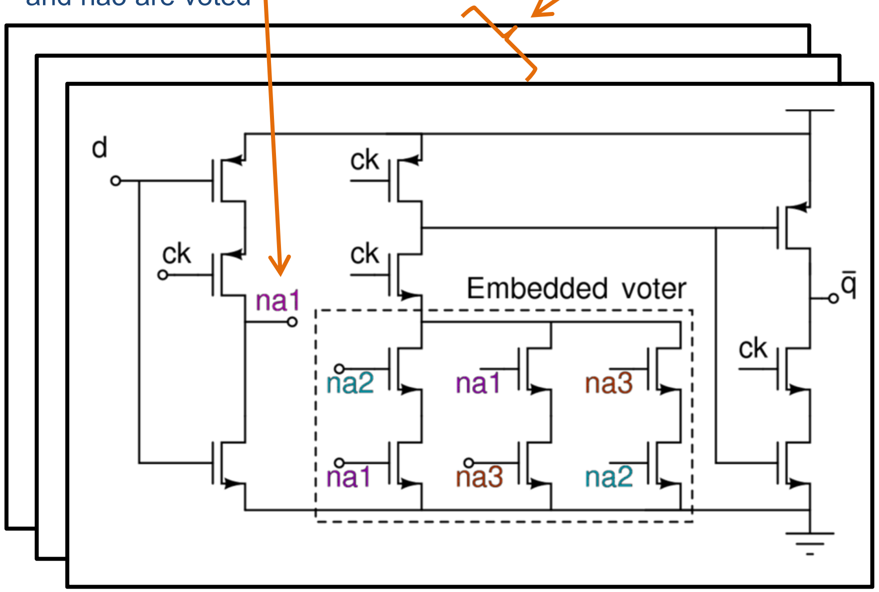


Type	Unprotected	Single voter TMR	Tripple voter TMR	Proposed
Static	1.0	0.9	0.8	-
Dynamic	1.9	1.4	1.1	1.6

Only a small speed penalty, 1.6 x faster than a static FF

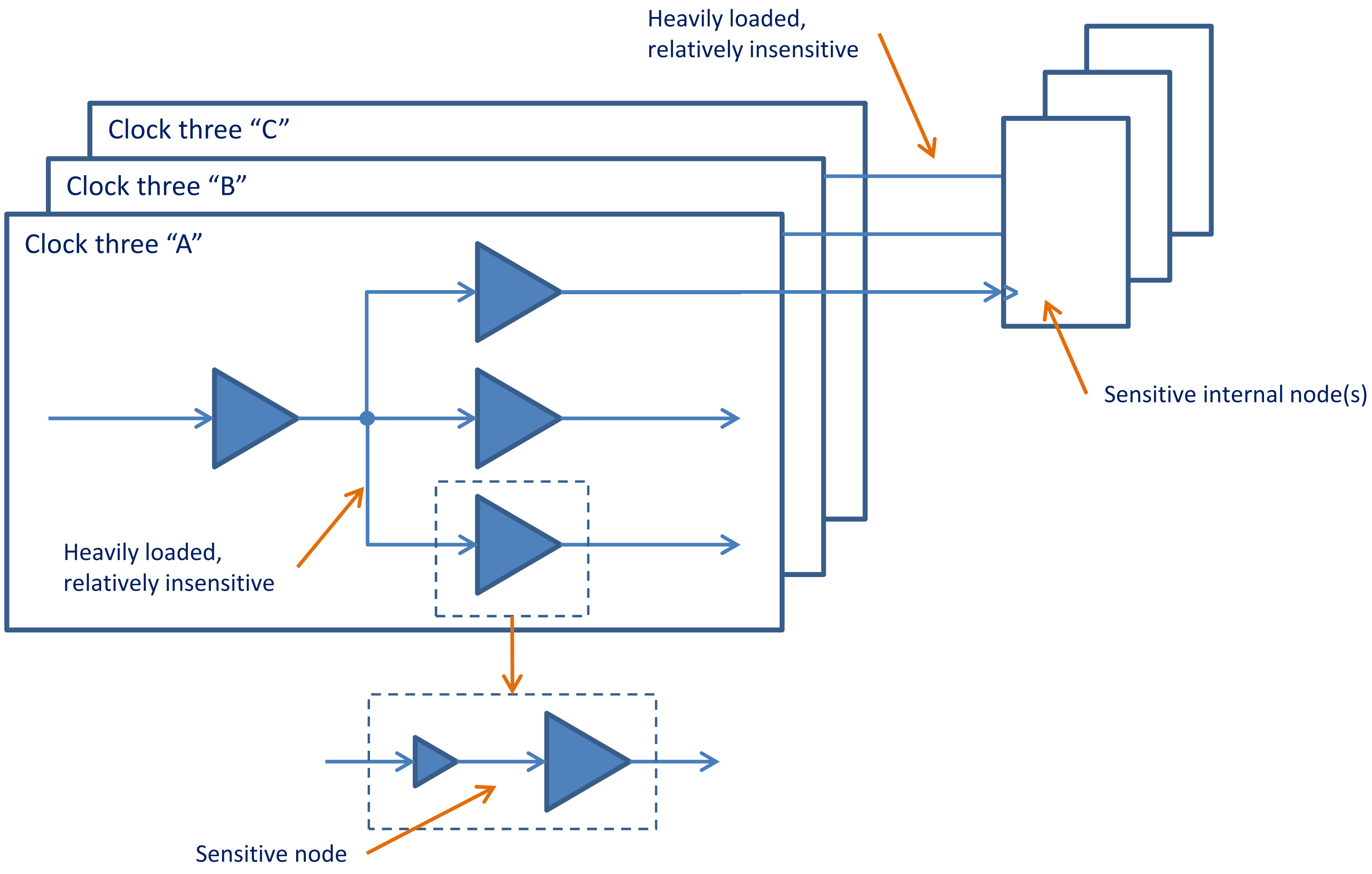
Internal nodes na1, na2 and na3 are voted

TMR FF = 3 x FF "Slice"



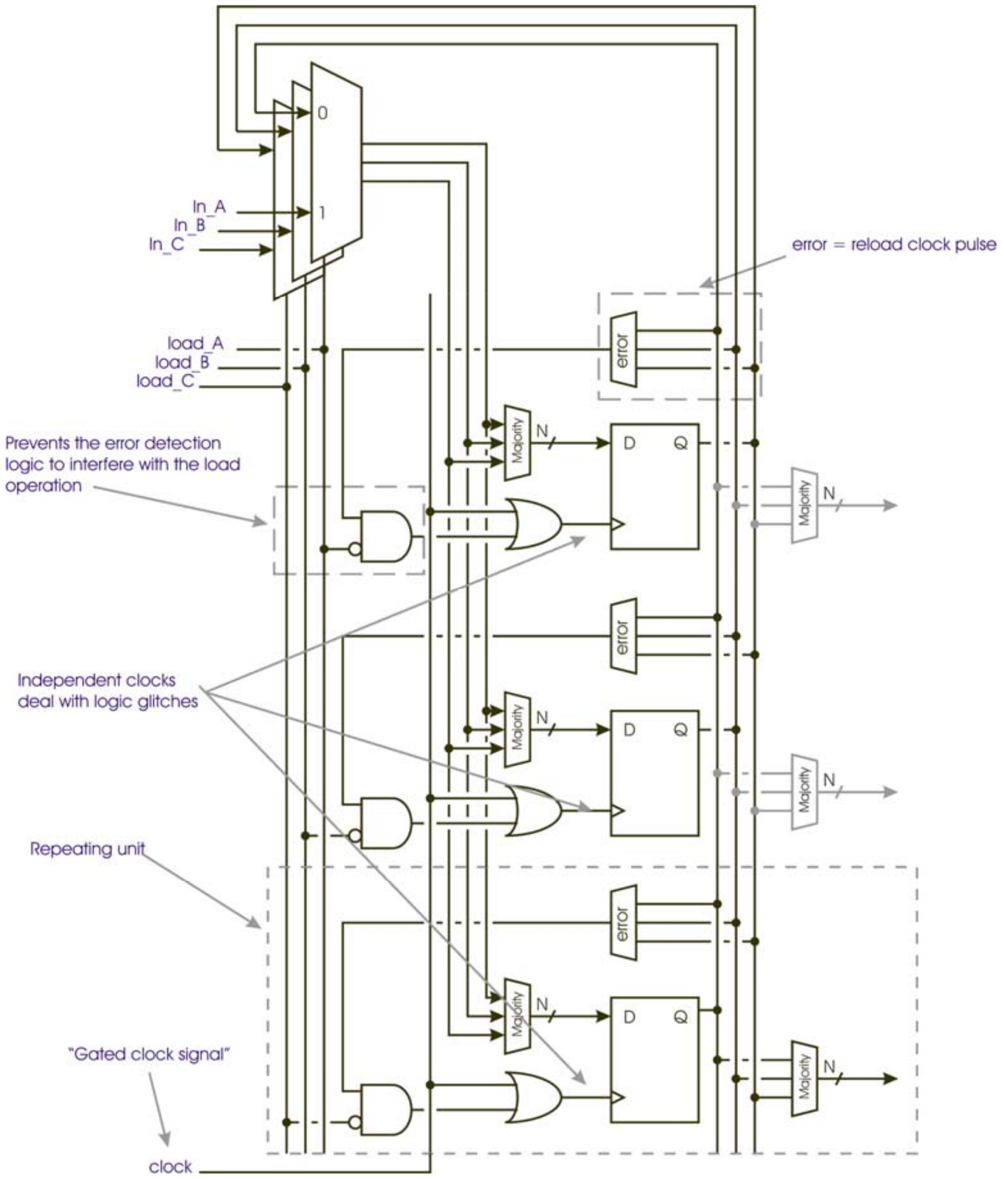
Paulo.Moreira@cern.ch

SEUs: Don’t Forget the Clock Tree



Paulo.Moreira@cern.ch

SEU Recovery in Clock Gated Registers



error = reload clock pulse

Prevents the error detection logic to interfere with the load operation

Independent clocks deal with logic glitches

Repeating unit

"Gated clock signal"