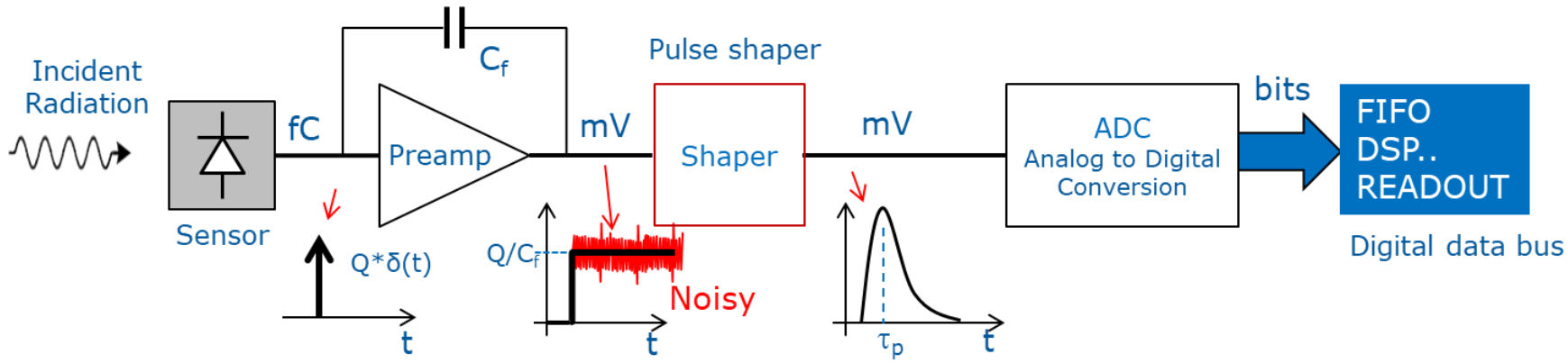
Novel strategies for high-granularity and radiation hardness LGAD sensors and front-end electronics

Michele CASELLE
(KIT)
michele.caselle@kit.edu

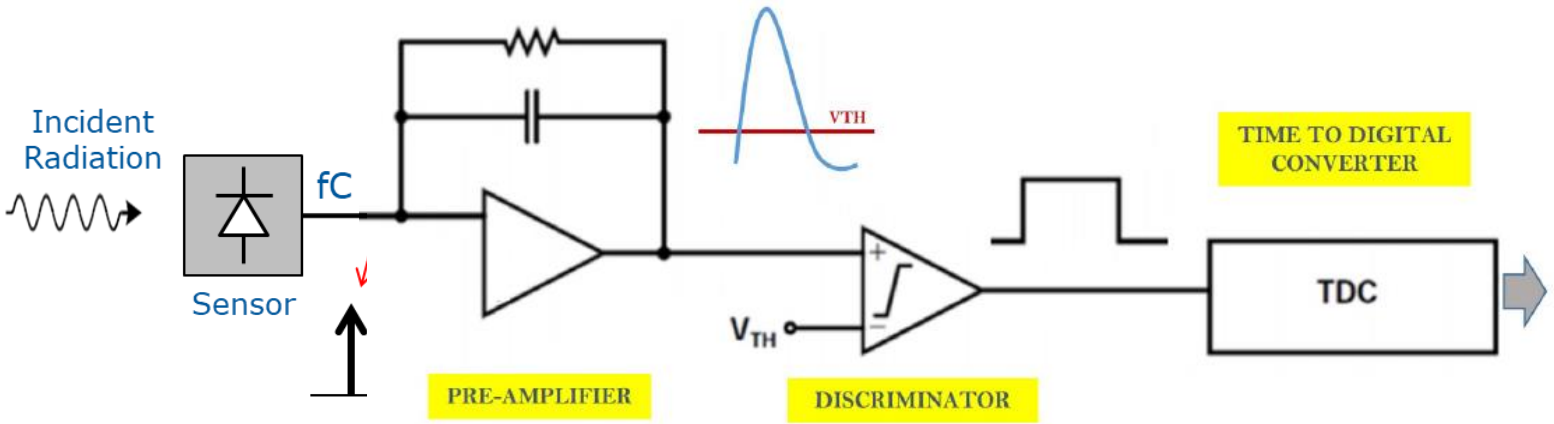
Readout chains

What makes the different between normal sensor and fast sensor?

- Typical front-end readout chain for “slow” silicon sensor in High Energy Physics (HEP)



- Front-end readout chain for “Fast” silicon detector in High Energy Physics (HEP)

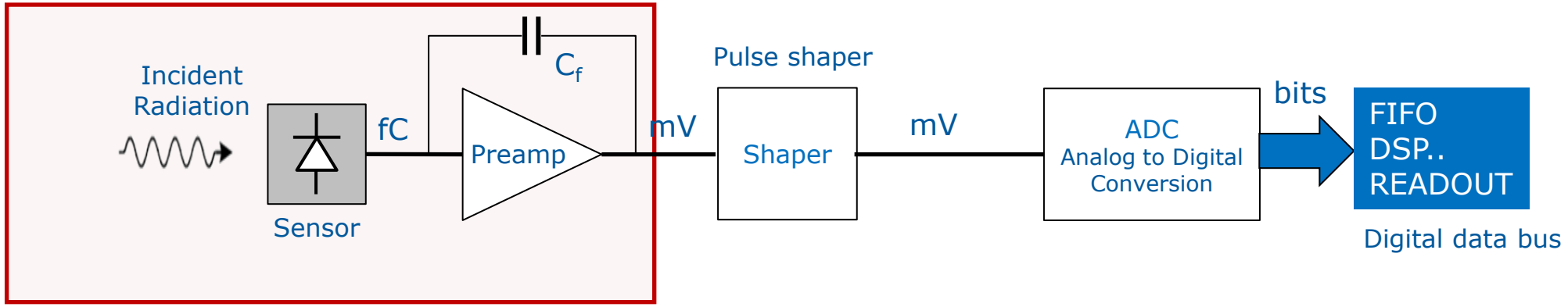


- How can we enhance the time resolution?
- What are the key parameters influencing it?

Readout chains

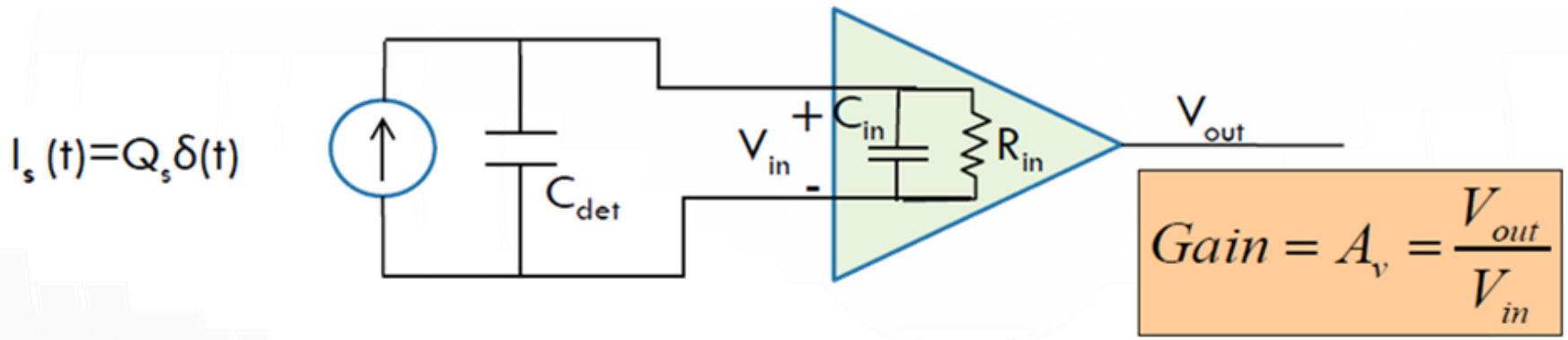
What makes the different between normal sensor and fast sensor?

- Typical front-end readout chain for “slow” silicon sensor in High Energy Physics (HEP)



Voltage with Capacitive Sources

What makes efficient the signal collection from silicon sensor?



- R_{in} is usually very large (because physically connected to the gate of the input transistor)
- I_s is integrated across the total capacitance $\mathbf{C_T}$

$C_T = C_{det} + C_{in}$ ⇒

$$V_{in} = \frac{1}{C_T}\int I_S(t)dt = \frac{Q_S}{C_{det} + Cin}$$

$$V_{OUT} = -A\,Vin = A\frac{QS}{C_{det} + Cin}$$

V_{out} is proportional to Qs

BUT … it also depends on C_{det} !!!

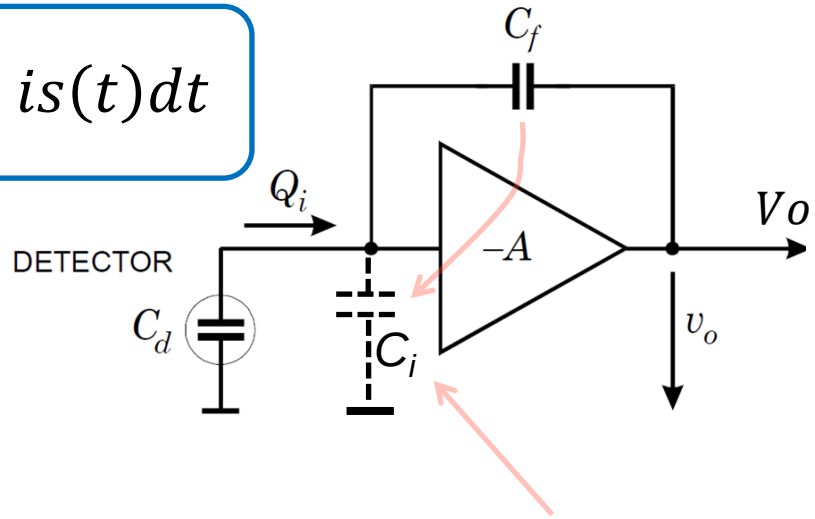
- **This is not desirable in the systems where C_{det} can vary different strip length/width, bias voltage, etc…**

Active integrator based on Charge Sensitive

Most used amplification stage for many sensors

- Inverting voltage amplifier: $v_o = -A\, vi$
- Input impedance: ∞

$$Qi = \int is(t)dt$$

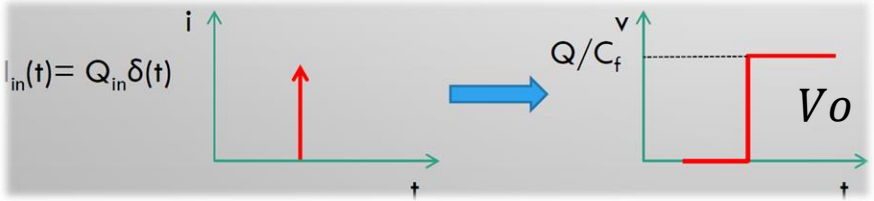


- What is the effective input capacitance seen from the detector
- By Miller theorem, effective input capacitance $C_i = C_f (1+A)$

- Gain of the system:

$$A_Q = \frac{V_o}{Q_i} = \frac{A\, vin}{C_i\, vin} = \frac{A}{C_f\,(1+A)} \cong \frac{1}{C_f} \; if \; A \gg 1$$

- **Set by a well-controlled quantity, the feedback capacitance**



- CSA is a pure integrator
- **Dirac pulse → step function**

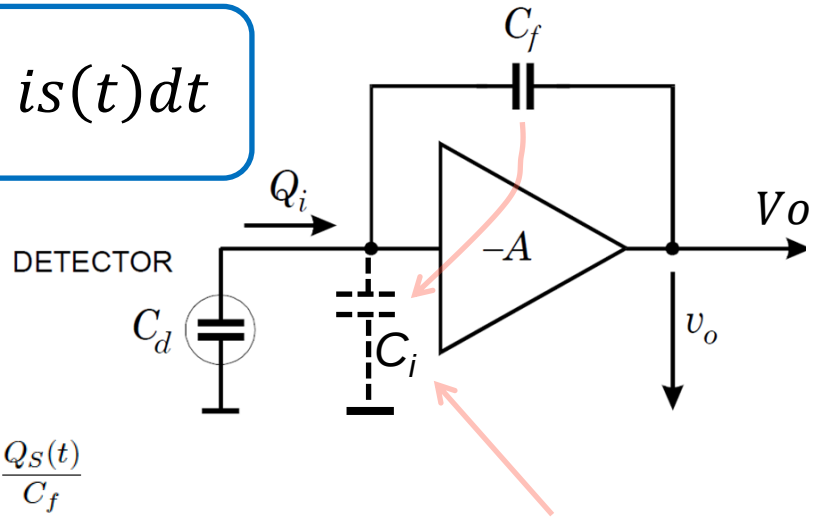
$$\boldsymbol{V_o = -\frac{1}{C_f}\int i_s\,(t)dt = -\frac{Q_i(t)}{C_f}}$$

Active integrator based on Charge Sensitive

Most used amplification stage for many sensors

- Inverting voltage amplifier: $v_o = -A\, vi$
- Input impedance: ∞

$$Qi = \int is(t)dt$$



- What is the effective input capacitance seen from the detector
- By Miller theorem, effective input capacitance $C_i = C_f (1+A)$

- Gain of the system:

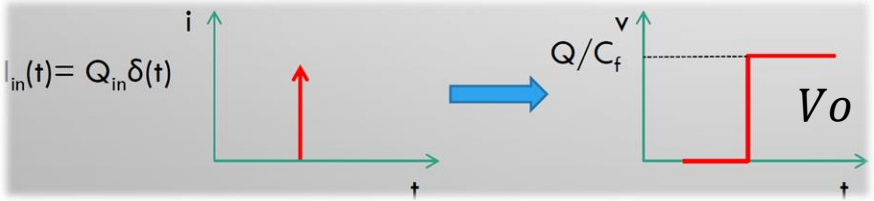
$$v_{out}(t) = -a_0\, v_{in}(t) = -\frac{1}{C_f}\int_0^t i_S\, dt' = -\frac{Q_S(t)}{C_f}$$

$$A_Q = \frac{V_o}{Q_i} = \frac{A\, vi}{C_i\, vi} = \frac{A}{C_f\,(1+A)} \cong \frac{1}{C_f} \; if \; A \gg 1$$

Dynamic input capacitance

BUT … not all the charge goes in the amplifier: a small fraction Q_{det} remains on C_{det} !!!

- **Set by a well-controlled quantity, the feedback capacitance**

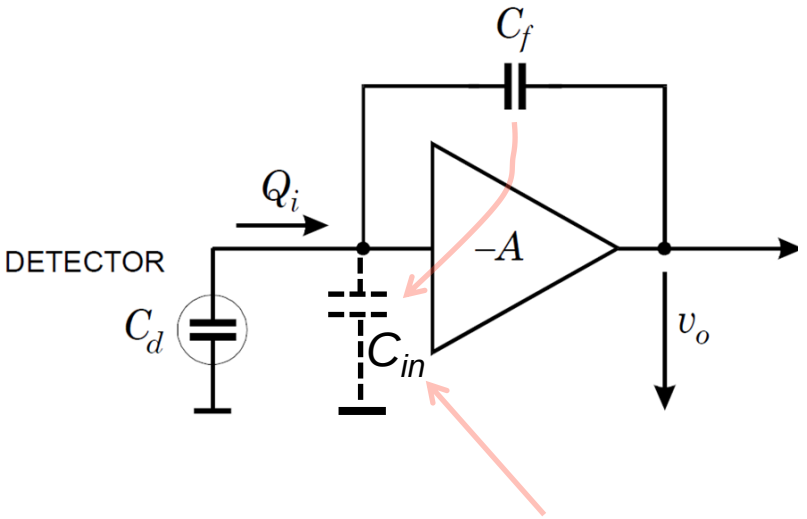


- CSA is a pure integrator
- **Dirac pulse → step function**

$$\boldsymbol{V_o = -\frac{1}{C_f}\int i_s\,(t)dt = -\frac{Q_i(t)}{C_f}}$$

Realistic Charge-Sensitive Preamplifiers

Amplifiers may be too slow to follow the instantaneous detector pulse



- Charge Q_s is divided across two capacitors C_{det} and C_{in}
- What is the fraction of the charge that contributed to the Vo and what is the fraction that will not contribute because remain on the C_{det} ?

$$\frac{Q_{in}}{Q_S} = \frac{Q_{in}}{Q_{in} + Qde_t} = \frac{1}{1 + \frac{Q_{det}}{Q_{in}}} = \frac{1}{1 + \frac{C_{det}}{C_{in}}}$$

- What is the value of Cin ?
- $C_i = C_f\,(1+A)$, where A is the intrinsic voltage gain of the amplifier

$$\frac{Q_{in}}{Q_S} = \frac{Q_{in}}{Q_{in} + Qde_t} = \frac{1}{1 + \frac{Q_{det}}{Q_{in}}} = \frac{\mathbf{1}}{\mathbf{1} + \frac{\boldsymbol{C_{det}}}{\boldsymbol{C_f}\,(\mathbf{1} + \boldsymbol{A})}}$$

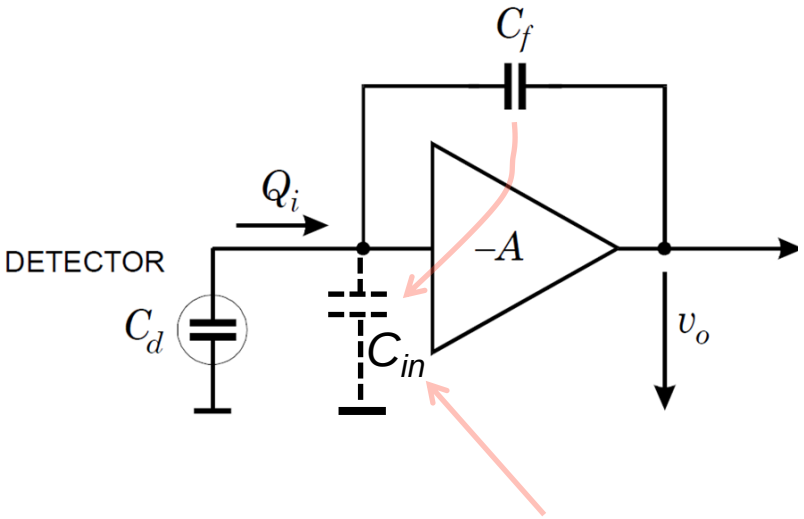
- **Voltage gain "A" play an important role in the change collection & efficiency**
- Example: C_{det}= 160 fF, A=10^2 C_f=16 pF, the fraction of the charge collected is: Q_{in}/Q_s= 0.9 → **10 % lost**
- C_{det}= 160 fF, A=10^3 C_f=16 pF, the fraction of the charge collected is: Q_{in}/Q_s= 0.99 → **1 % lost**

Realistic Charge-Sensitive Preamplifiers

Amplifiers may be too slow to follow the instantaneous detector pulse

- Charge Q_s is divided across two capacitors C_{det} and C_{in}
- What is the fraction of the charge that contributed to the Vo and what is the fraction that will not contribute because remain on the C_{det} ?

$$\frac{Q_{in}}{Q_S} = \frac{Q_{in}}{Q_{in} + Qde_t} = \frac{1}{1 + \frac{Q_{det}}{Q_{in}}} = \frac{1}{1 + \frac{C_{det}}{C_{in}}}$$



- What is the value of Cin ?
- $C_i = C_f\,(1+A)$, where A is the intrinsic voltage gain of the amplifier

$$\frac{Q_{in}}{Q_S} = \frac{Q_{in}}{Q_{in} + Qde_t} = \frac{1}{1 + \frac{Q_{det}}{Q_{in}}} = \frac{\mathbf{1}}{\mathbf{1} + \frac{\boldsymbol{C_{det}}}{\boldsymbol{C_f}\,(\mathbf{1} + \boldsymbol{A})}}$$

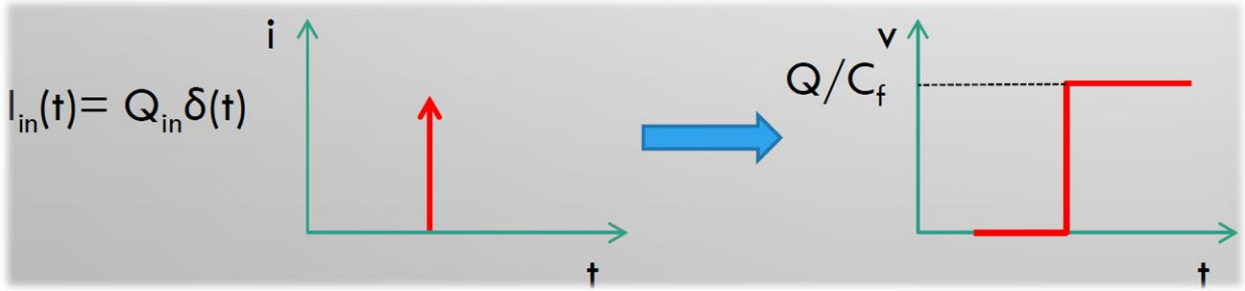
BUT … we are considering an ideal amplifier with infinite bandwidth (infinite speed) !!!

- **Voltage gain "A" play an important role in the change collection & efficiency**
- Example: C_{det}= 160 fF, A=10^2 C_f=16 pF, the fraction of the charge collected is: Q_{in}/Q_s= 0.9 → **10 % lost**
- C_{det}= 160 fF, A=10^3 C_f=16 pF, the fraction of the charge collected is: Q_{in}/Q_s= 0.99 → **1 % lost**

Realistic Charge-Sensitive Preamplifiers

Amplifiers may be too slow to follow the instantaneous detector pulse

- A pulse Dirac of input current ➔ generates a voltage step at the output of the CSA (if ∞ bandwidth (BW)

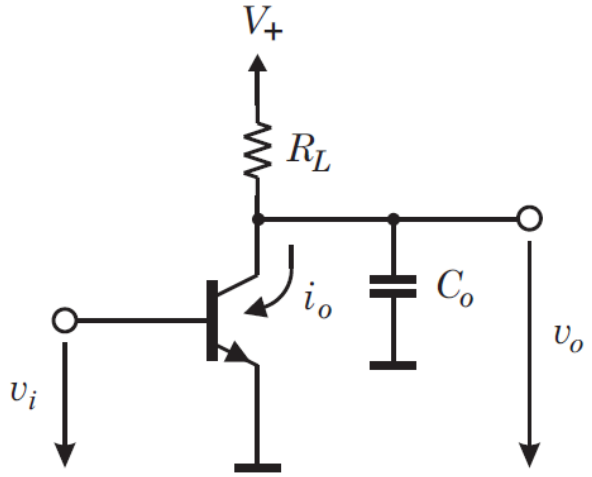


$$V_o = -Qin \int i_s(t)dt = -\frac{Q_{in}}{C_f}$$

- In reality, the volage amplifier is dominated by its **dominant pole** (time constant) due to the internal capacitances in the amplifier must first charge up

Realistic Charge-Sensitive Preamplifiers

Amplifiers may be too slow to follow the instantaneous detector pulse



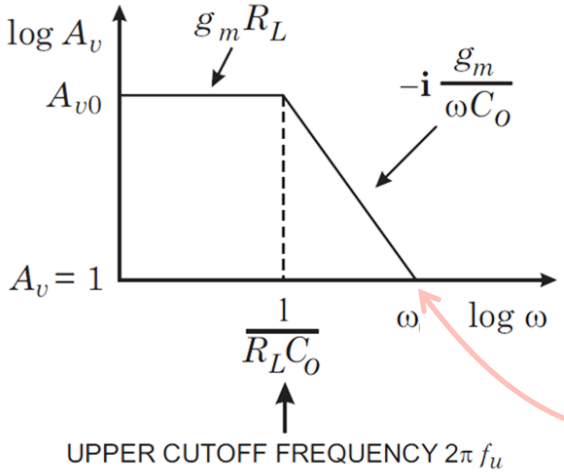
$$\frac{1}{Z_L} = \frac{1}{R_L} + \mathbf{i}\omega C_o$$

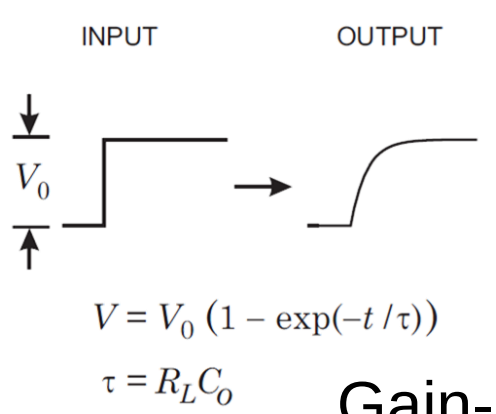
Because R_L is in parallel to C_o

$$A_v = \frac{dv_o}{dv_i} = \frac{di_o}{dv_i} \cdot Z_L \equiv g_m Z_L$$

Transconductance = $\frac{di_o}{dv_i} = g_m$

$$A(\omega) = \frac{A_0}{1 + i\frac{\omega}{\omega_0}} = \frac{g_m R_L}{1 + i\frac{\omega}{\omega_o}}$$ where $\omega_o = 1/\tau = 1/R_L C_o$





$A(\omega)$

low−frequency $= g_m R_L$

high−frequency $A(\omega) = -iA_0\frac{\omega_0}{\omega}$

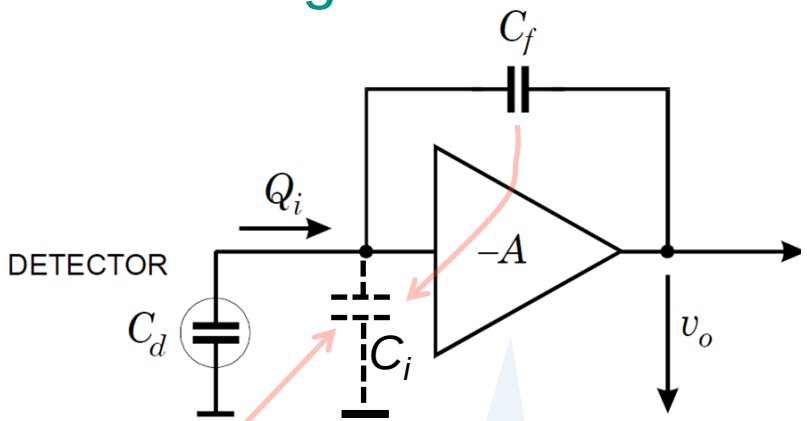
power

Gain-bandwidth product $\boldsymbol{GBW} = \boldsymbol{g_m R_L \frac{1}{R_L C_o}} = \boldsymbol{\frac{g_m}{C_0}}$

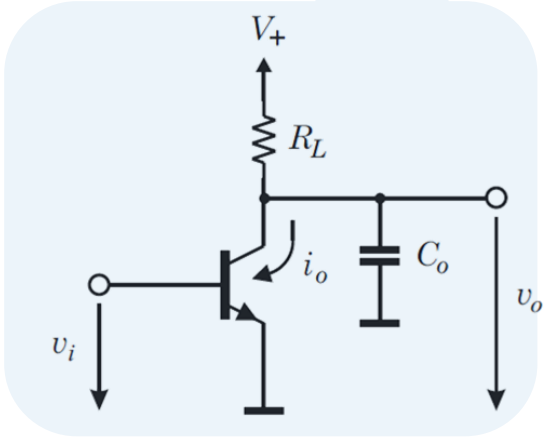
Detector capacitance

Realistic Charge-Sensitive Preamplifiers

For timing



Dynamic input capacitance



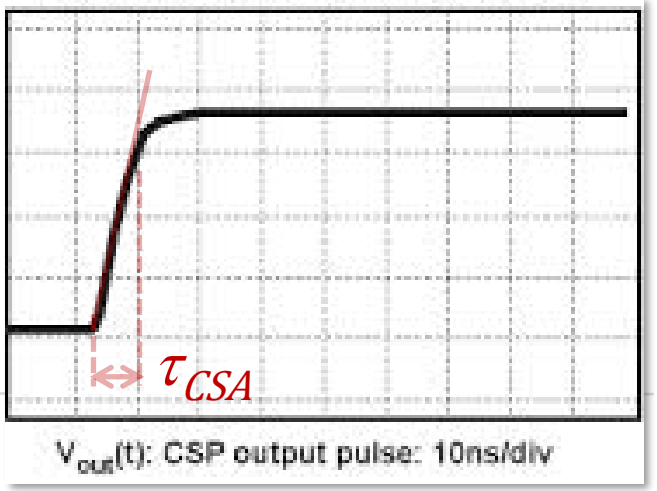
- Effective input capacitance $C_i = C_f (1+A(\omega))$
- Input impedance $Z_{in} = \frac{1}{i\omega Ci}$ $\omega \gg \omega_0$
- High frequency $A(\omega) = -iA_0 \frac{\omega_0}{\omega}$
- Input impedance $Z_{in} = \frac{1}{i\omega Ci} = \frac{1}{i\omega Cf(-iAo\frac{\omega_0}{\omega})} = \frac{1}{C_f A_o \omega_0} = \frac{C_o}{C_f g_m}$

Depending on the size W/L of the input device

The input impedance of the CSA behaves as a resistor

Timing $\tau_{CSA} = C_D Z_{in} = \frac{C_D}{C_f Ao \omega_0} = \frac{C_D}{C_f} \frac{C_o}{g_m}$

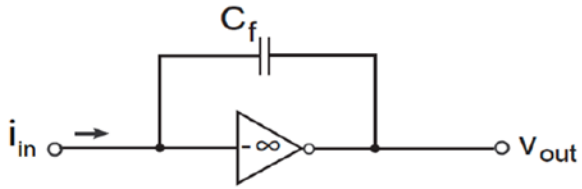
Depends on the **C_D** and **BW** and **g_m** (power)



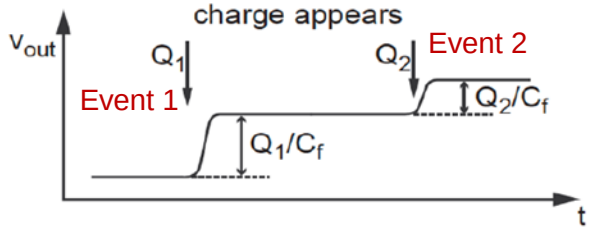
Realistic Charge-Sensitive Preamplifiers

Discharging the C_f and reset

ideal situation w/o reset

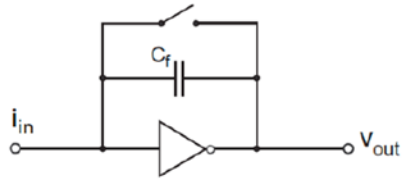


(a) Circuit diagram.

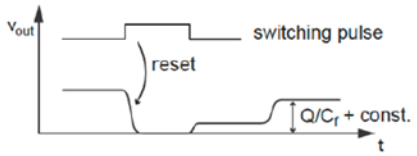


(b) Output voltage.

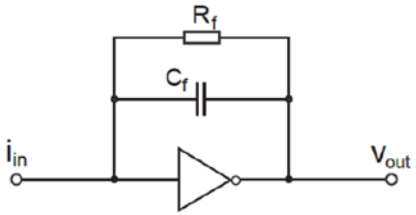
discharging options



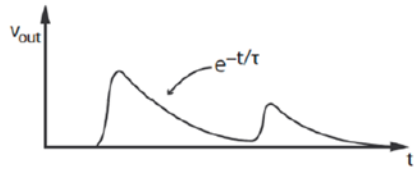
(a) Switch reset: circuit.



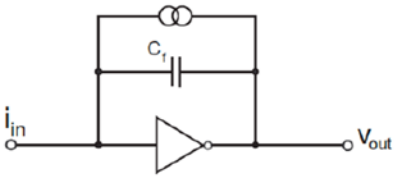
(b) Switch reset: output voltage.

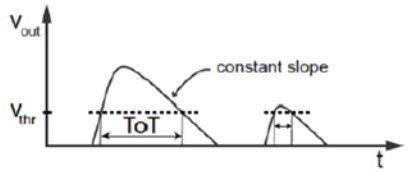


(c) Reset via resistor: circuit.



(d) Reset via resistor: output voltage.





=> “pile-up” => saturation

- Reset circuit is necessary to discharge the C_f

“reset” switch

- Mainly employed for photon sciences, i.e Kalypso detector system

resistor R_f

- Very simple option, mainly employed for fast detector, i.e. timing detectors

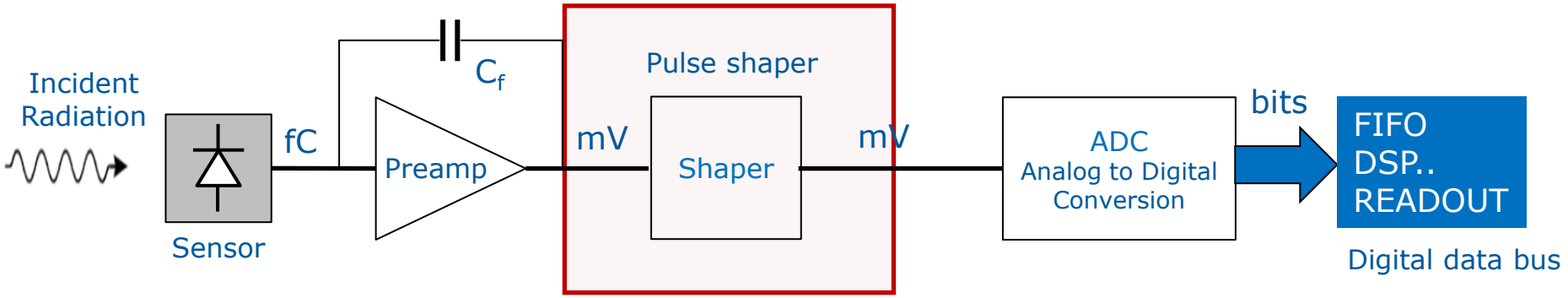
a current source

- Very accurate option, many HEP front-end, i.e. ATLAS, CMS, etc.

Readout chains

What makes the different between normal sensor and fast sensor?

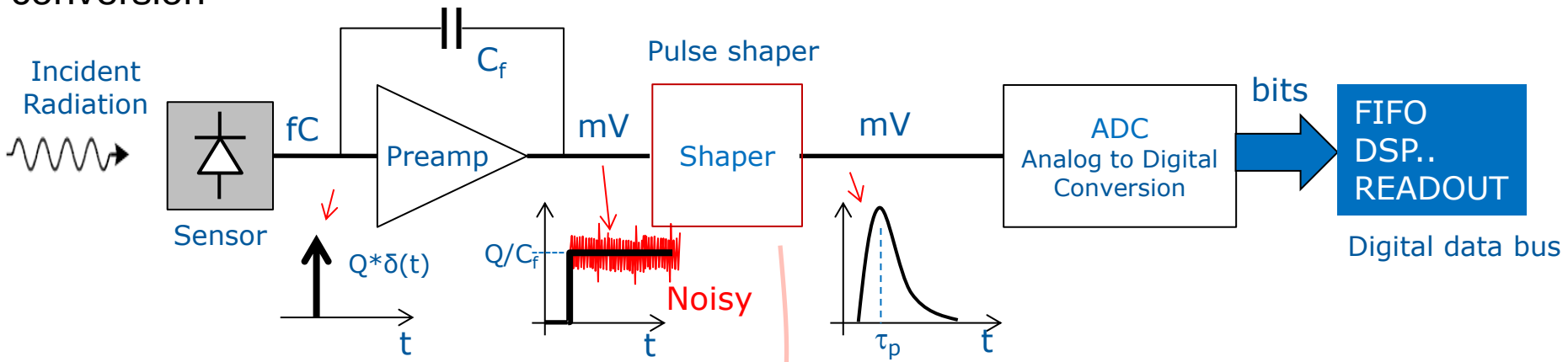
- Typical front-end readout chain for “slow” silicon sensor in High Energy Physics (HEP)



Pulse shaping

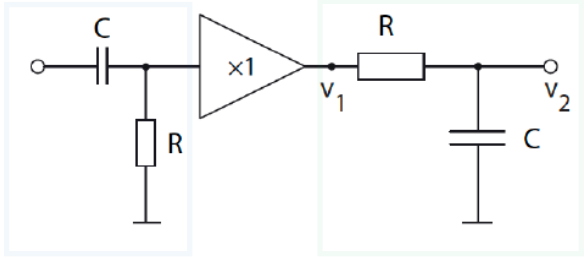
Semi-Gaussian filter based on CR-(RC)^M filter (shaper)

- To reduce the noise generated from the sensor (short noise) and the CSA (thermal noise) before the ADC conversion



High-pass filter

$$H_1(s) = \frac{sRC}{1+sRC}$$



Low-pass filter

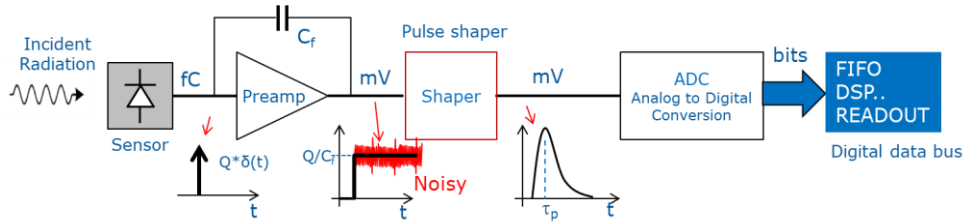
$$H_2(s) = \frac{1}{1+sRC}$$

- What are the main characteristics of the shaper?
- How does the peaking time factor in?

Pulse shaping

Semi-Gaussian filter based on CR-(RC)^M filter (shaper)

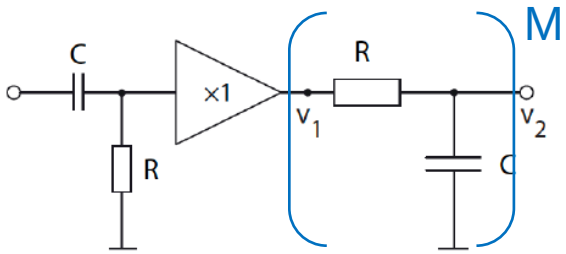
- To reduce the noise generated from the sensor (short noise) and the CSA (thermal noise) before the ADC conversion



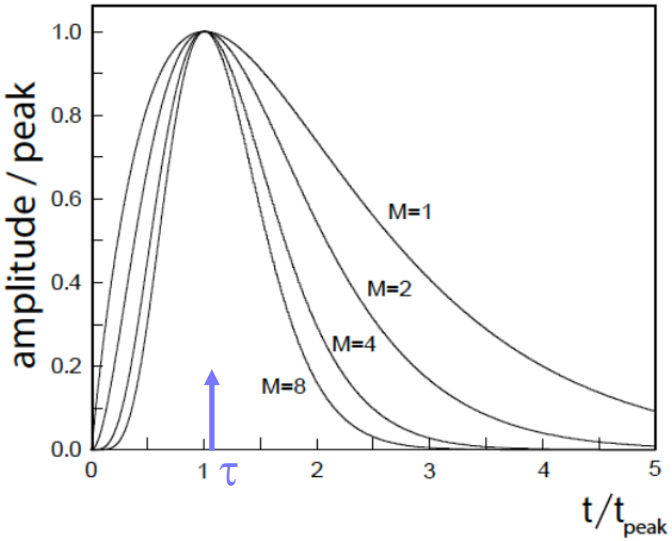
In the time domain $V_{sh}(t) = A\frac{t}{\tau}e^{-\frac{t}{\tau}}$

- A = pulse amplitude = $\frac{Q_i(t)}{C_f}$

- The peaking time is always at t=τ independently on the pulse amplitude (energy released into sensor)
- Where τ = RC



- This is called CR-(RC)^M
- $V_{sh}(t) = A\frac{1}{M!}(\frac{t}{\tau})^M e^{-\frac{t}{\tau}}$

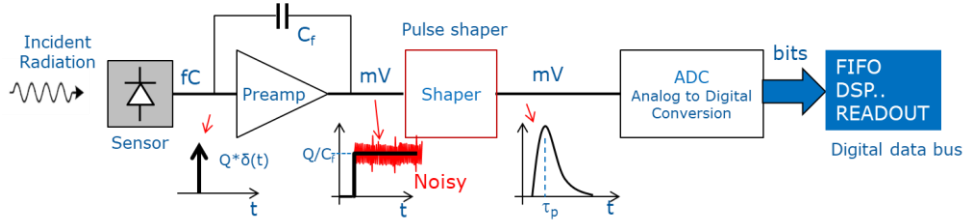


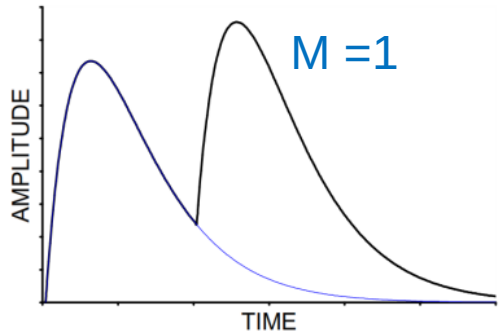
- To improve the pile-up of the signal (see next slide)
- To increase the bandwidth → less noise reduction

Pulse shaping vs pileup

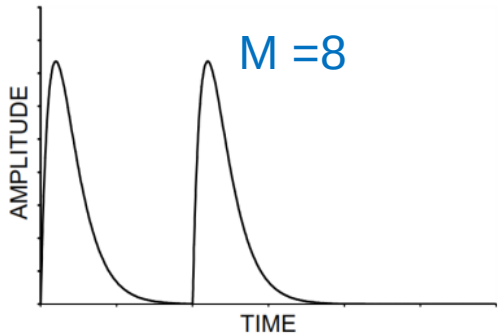
Semi-Gaussian filter based on CR-$(RC)^M$ filter (shaper)

- To reduce the noise generated from the sensor (short noise) and the CSA (thermal noise) before the ADC conversion

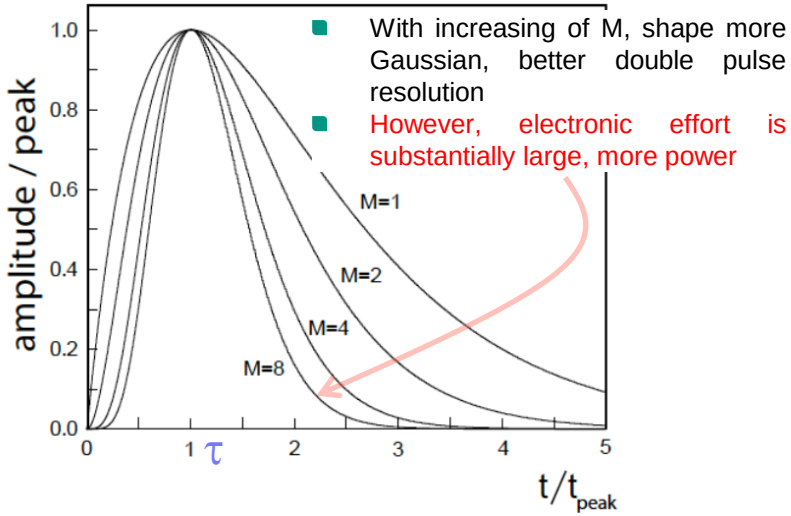




Pile-up



No pile-up



- With increasing of M, shape more Gaussian, better double pulse resolution
- However, electronic effort is substantially large, more power

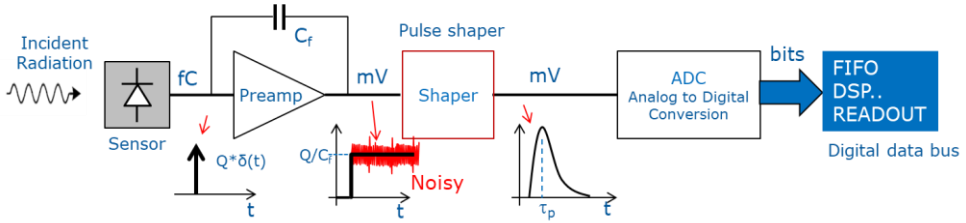
- Improve S/N ratio → restrict bandwidth → increase pulse width
- But increase the pile-up

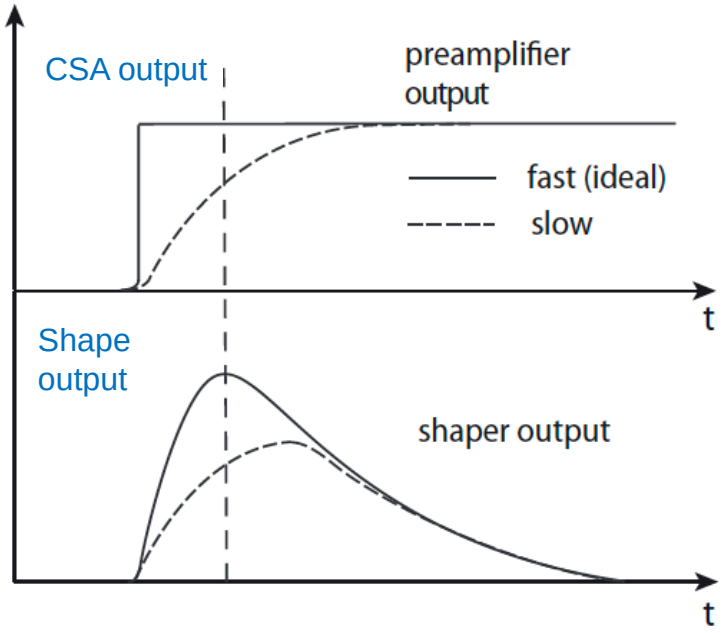
- Reduce the pile-up → decrease pulse width
- Reduce S/N ratio, more compex, more power

- Two conflicting objectives arise: the optimal filter depends on the specific physics application and the pixel or channel geometry of the sensor

Some more things …

“Ballistic deficit” or “shaping loss”





- The term "ballistic deficit" refers to a phenomenon in particle detectors, particularly in systems using charge-sensitive amplifiers and shaping circuits, where the **signal's amplitude is reduced due to incomplete charge collection within the shaping time**. This effect can significantly impact the accuracy of energy measurements in detectors

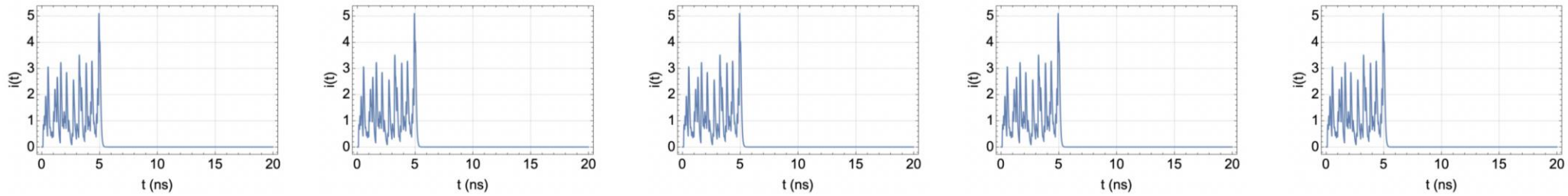
- If signal step at the output of CSA takes much longer than the shaper τ_{CR}
- caused e.g.
  - by a large charge collection time
  - by a large input capacitance
  - by an intrinsically slow preamplifier
- The V_{out} (shaper) is trimmed by the slow rise of the preamplifier output pulse

- **Mitigation strategies, include:** Optimizing Shaping Time, designing detectors with uniform and fast charge collection properties, such as using high-field regions to speed up carrier drift, understanding the detector's response through simulations and measurements to anticipate and compensate for ballistic deficit effects

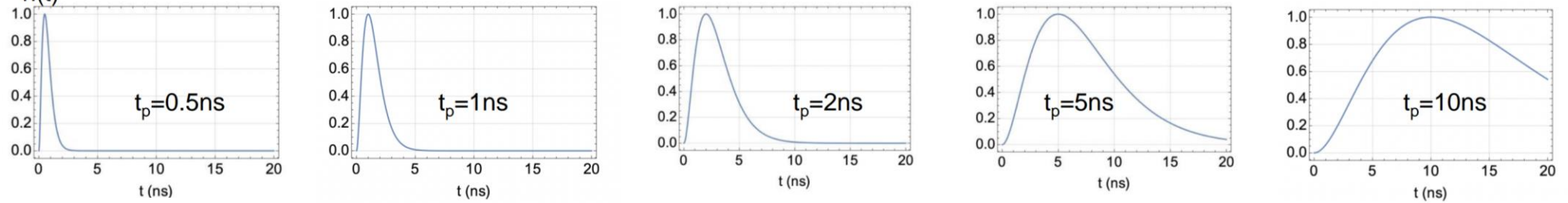
What about the peaking time?

Fast shaper → ballistic deficit

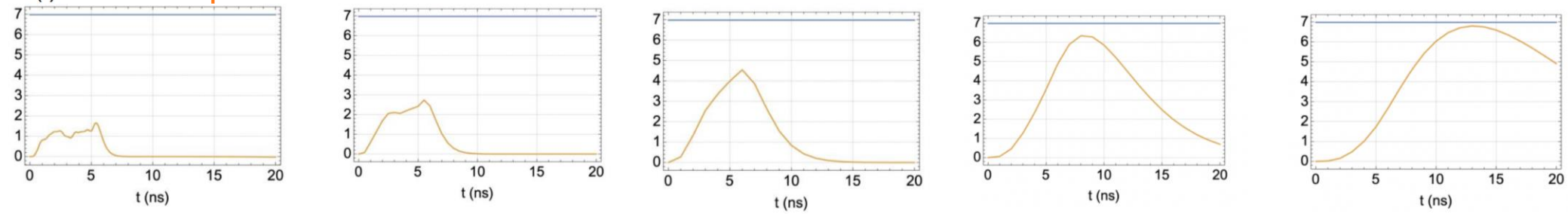
f(t) Current from sensor



h(t) Shaper delta response



v(t) Vout shaper

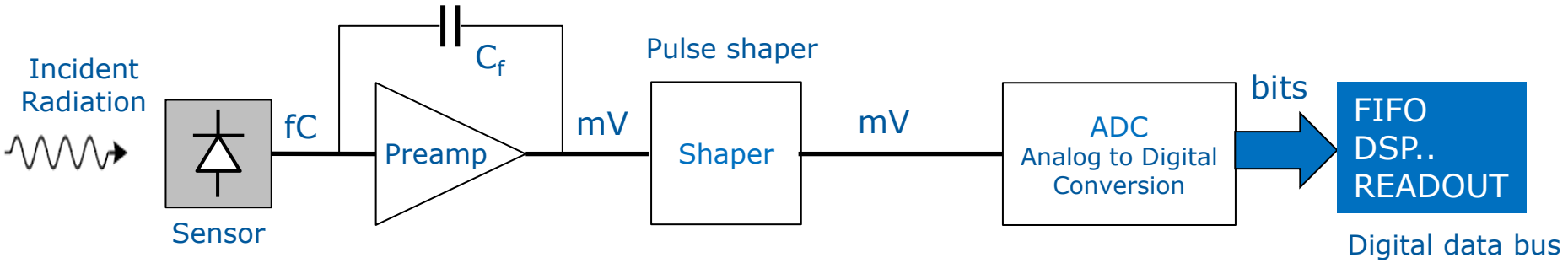


Source: Werner Riegler, CERN

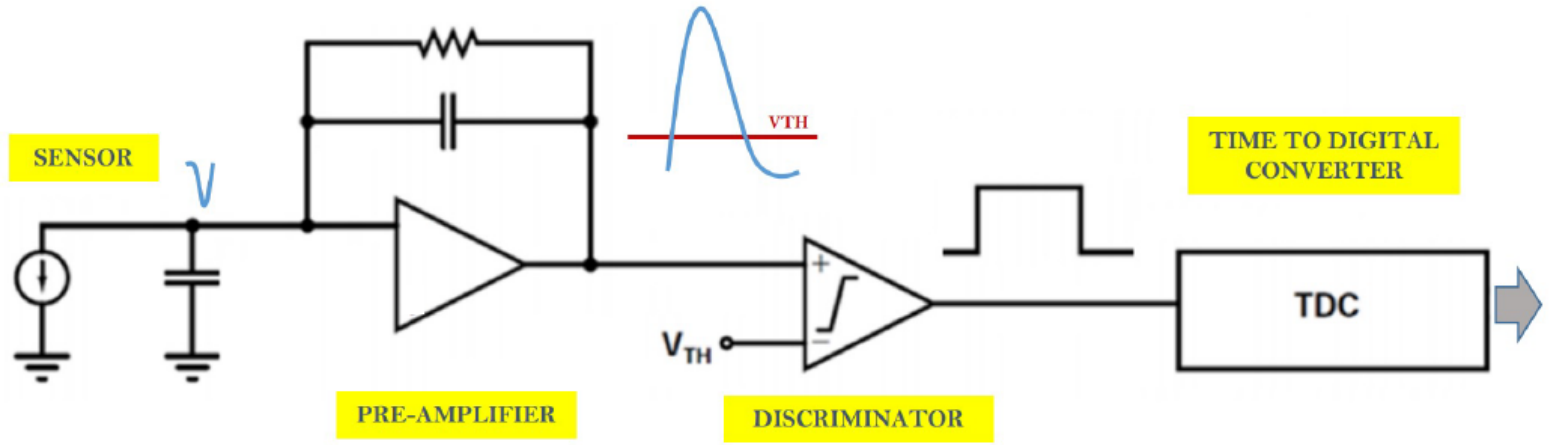
Readout chains

What makes the different between normal sensor and fast sensor?

- Typical front-end readout chain for “slow” silicon sensor in High Energy Physics (HEP)



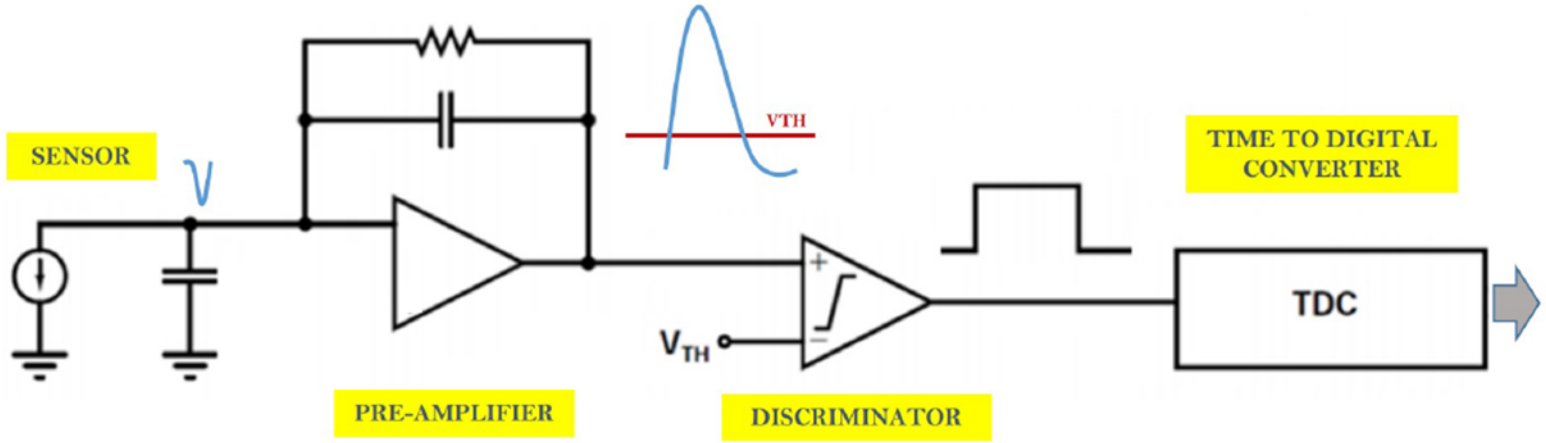
- Front-end readout chain for “Fast” silicon detector in High Energy Physics (HEP)



- How can we improve the time resolution?
- What are the key parameters ?

Readout chains of fast silicon sensors

What makes the different between normal sensor and fast sensor?

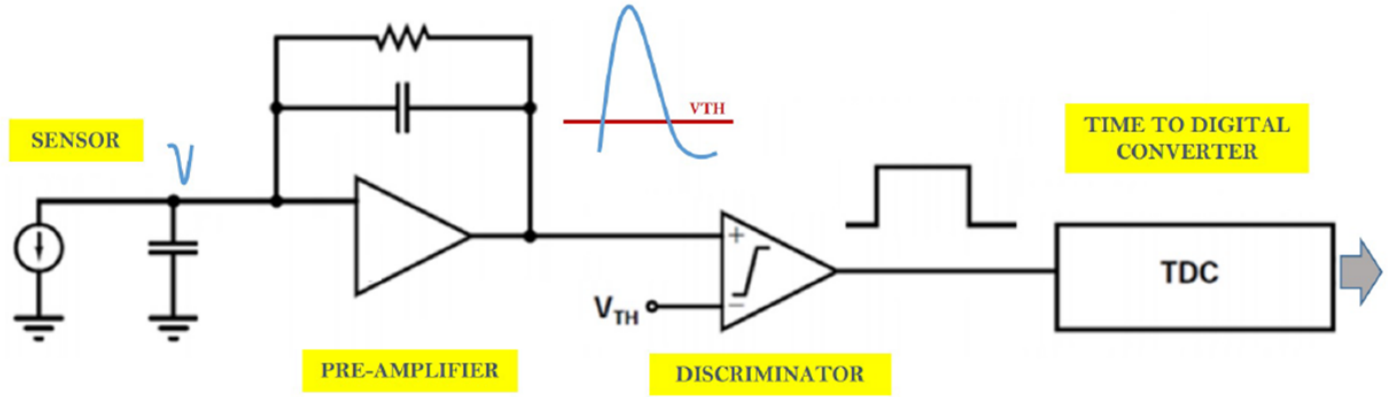


- Readout chain consists of:
  - Preamplifier (first stage, the noise strongly depend on this stage)
  - Fast comparator
  - High resolution TDC

Time resolution in timing detectors

The timing capabilities depends on the signal at the output of the pre-amplifier and by the TDC binning

- Timing measurements is limited by:
  - Jitter → timing resolution
  - Time walk → timing accuracy



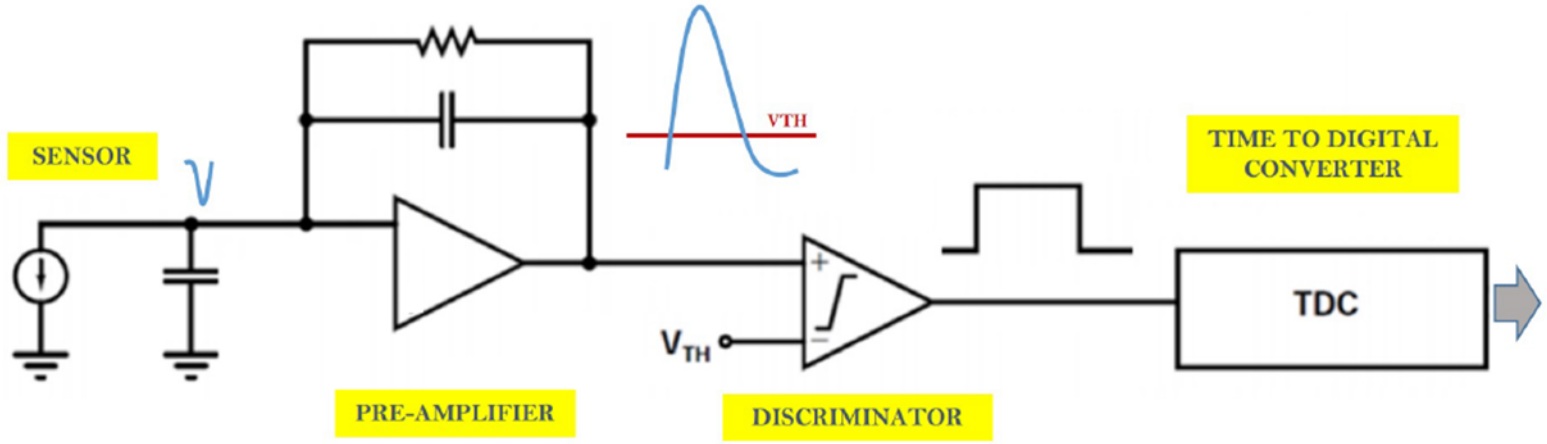
- $\sigma_t^2 = \left(\frac{Noise}{dV/dt}\right)^2 + \sigma_{timewalk}^2 + \left(\frac{LSB}{\sqrt{12}}\right)^2 + \sigma_{sensor}^2 + \sigma_{Landau}^2$

Front-end | TDC | Sensor

Total noise contribution

Time resolution in timing detectors

The timing capabilities depends on the signal at the output of the pre-amplifier and by the TDC binning



- $\sigma_t^2 = (\frac{Noise}{dV/dt})^2 + \sigma_{timewalk}^2 + (\frac{LSB}{\sqrt{12}})^2 + \sigma_{sensor}^2 + \sigma_{Landau}^2$

Front-end | TDC | Sensor

Total noise contribution

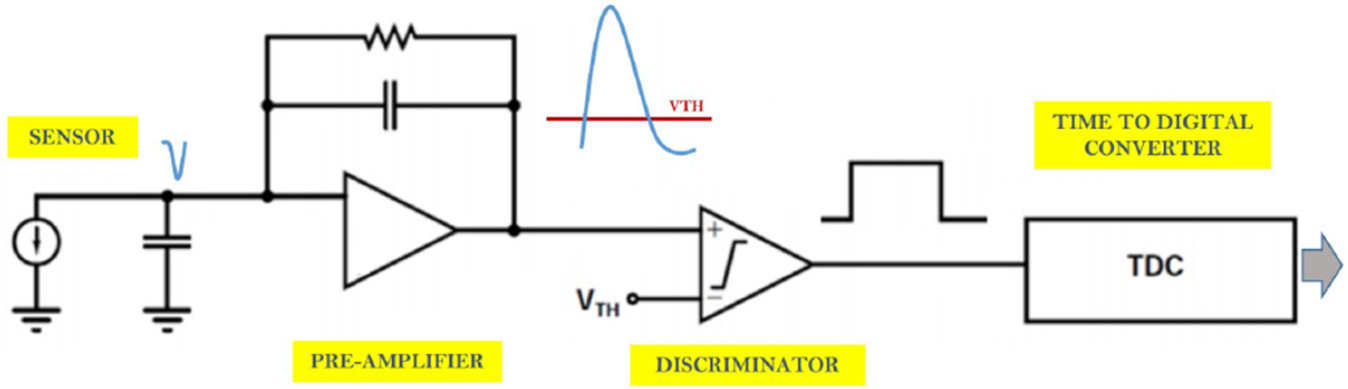
Yesterday lecture

- σ_{Landau}^2 → non-uniform energy deposition
  - Thinner sensor operating E_{sat} condition, reduce this noise contribution
- σ_{sensor}^2 → Weighting field and SNR
  - LGAD or 3D technologies, etc.

Time resolution in timing detectors

The timing capabilities depends on the signal at the output of the pre-amplifier and by the TDC binning

- Timing measurements is limited by:
  - Jitter → timing resolution
  - Time walk → timing accuracy



- $\sigma_t^2 = (\frac{Noise}{dV/dt})^2 + \sigma_{timewalk}^2 + (\frac{LSB}{\sqrt{12}})^2 + \sigma_{sensor}^2 + \sigma_{Landau}^2$

Front-end | TDC | Sensor

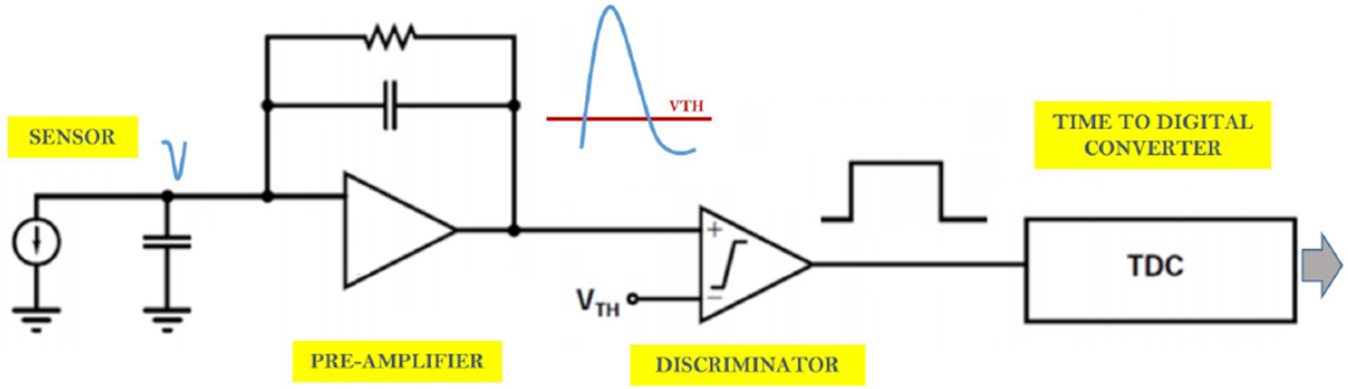
Total noise contribution

- **Noise**: There are two primary sources of noise—sensor noise (shot noise) and thermal noise from the electronics.
- **dV/dt**: This is influenced by several factors, including the sensor characteristics, signal speed, the open-loop gain of the charge-sensitive amplifier (CSA), bandwidth, and other parameters.
- **Time walk**: This effect is tied to the subsequent stage, which is based on the discriminator.

Time resolution in timing detectors

The timing capabilities depends from the signal at the output of the pre-amplifier and by the TDC binning

- Timing measurements is limited by:
  - Jitter → timing resolution
  - Time walk → timing accuracy



- $\sigma_t^2 = \left(\frac{Noise}{dV/dt}\right)^2 + \sigma_{timewalk}^2 + \left(\frac{LSB}{\sqrt{12}}\right)^2 + \sigma_{sensor}^2 + \sigma_{Landau}^2$

Front-end | TDC | Sensor

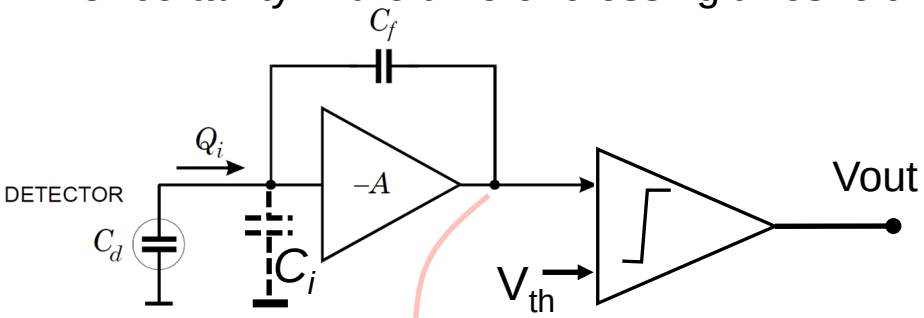
Total noise contribution

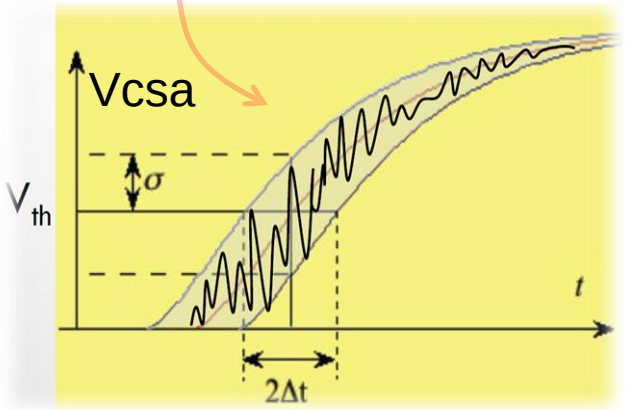
- **Noise**: There are two primary sources of noise—sensor noise (shot noise) and thermal noise from the electronics.
- **dV/dt**: This is influenced by several factors, including the sensor characteristics, signal speed, the open-loop gain of the charge-sensitive amplifier (CSA), bandwidth, and other parameters.
- **Time walk**: This effect is tied to the subsequent stage, which is based on the discriminator.

The jitter

Noise has an impact in the time measurements

- Uncertainty in the time of crossing threshold → *jitter*





- $\Delta t = \frac{\sigma_{noise}}{dV/dt}$

Sensor + electronics noise

Slope of the signal (**Timing τ_{CSA}**)

$$\textbf{Timing } \tau_{CSA} = \frac{C_D}{C_f A o \omega_0} = \frac{C_D}{C_f} \frac{C_o}{g_m}$$

Power

Gain

Bandwidth

- How to decrease jitter?

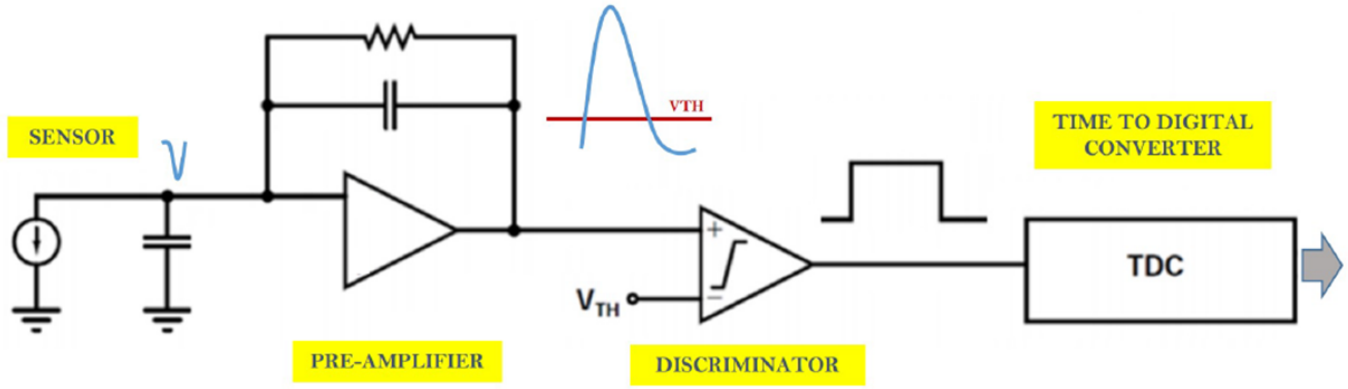
  Conflicting conditions:
  - Decrease σ_{noise} → decrease bandwidth
  - Increase slope → increase bandwidth
- Increasing the A_0 the open loop gain → more power, more complex

- LGADs (Low Gain Avalanche Detectors) offer a significant combined with very short charge collection times and relatively low noise levels. As a result, LGADs are a key sensor technology for addressing the conflicting requirements often encountered with traditional sensors, providing an optimal balance between speed, sensitivity, and noise performance

Time resolution in timing detectors

The timing capabilities depends from the signal at the output of the pre-amplifier and by the TDC binning

- Timing measurements is limited by:
  - Jitter → timing resolution
  - Time walk → timing accuracy



- $\sigma_t^2 = (\frac{Noise}{dV/dt})^2 + \sigma_{timewalk}^2 + (\frac{LSB}{\sqrt{12}})^2 + \sigma_{sensor}^2 + \sigma_{Landau}^2$

Front-end | TDC | Sensor

Total noise contribution

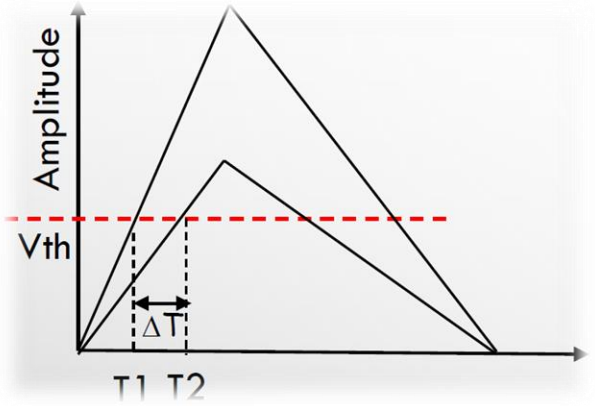
- **Noise**: There are two primary sources of noise—sensor noise (shot noise) and thermal noise from the electronics.
- **dV/dt**: This is influenced by several factors, including the sensor characteristics, signal speed, the open-loop gain of the charge-sensitive amplifier (CSA), bandwidth, and other parameters.
- **Time walk**: This effect is tied to the subsequent stage, which is based on the discriminator.

Time measurements

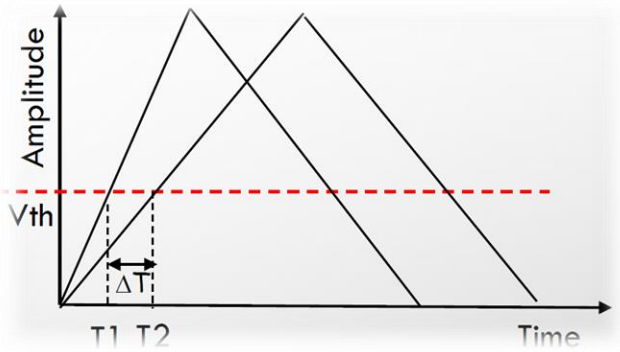
Time walk from standard leading edge discriminator

- The simplest scheme is based on: Leading edge or Threshold discriminator (comparator), when the signal crosses a threshold, the output goes from “low” to “high” level

- In the leading edge discriminators, two pulses with identical shape and time of occurrence, but different amplitude cross the same threshold in different times: **amplitude time walk**



- Even if the input amplitude is constant, time walk can still occur if the shape (rise time) of the pulse changes (for example, for changes in the charge collection time): **rise time walk**

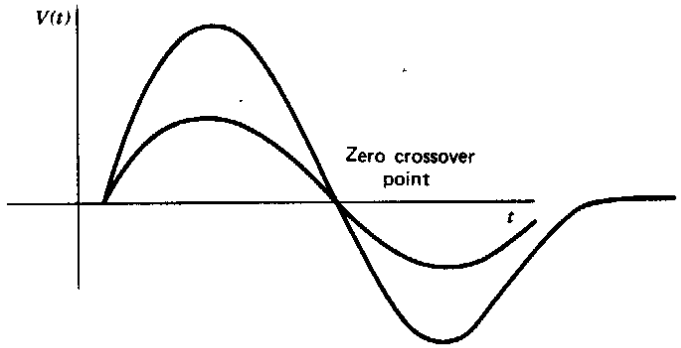


- The sensitivity of leading edge discriminator to time walk is minimized by setting the threshold *as low as possible* but it must be compatible with noise level
- Time walk correction:
  - By time-over-threshold ToT: measure the pulse amplitude and apply correction to timing
  - Hardware: employ the crossover timing, Constant Fraction timing

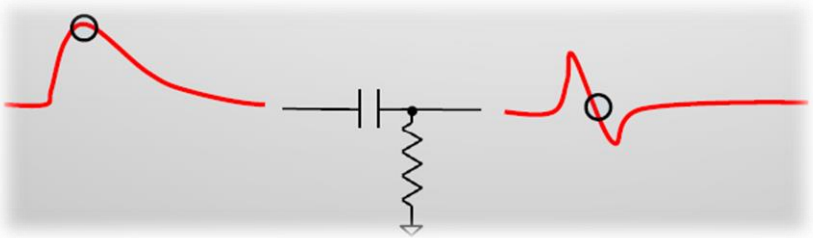
Time walk correction

Crossover timing and Constant Fraction timing

- The crossover timing can greatly reduce the magnitude of the amplitude time walk
- *Hypothesis*: the output of the shaper is a bipolar pulse, and the time of zero-crossing is independent of the pulse amplitude



- If the output of shaper is unipolar, but the peaking time is constant, adding a differentiator (C-R network) we get a bipolar pulse crossing the zero in correspondence of the signal peak



As a derivation function

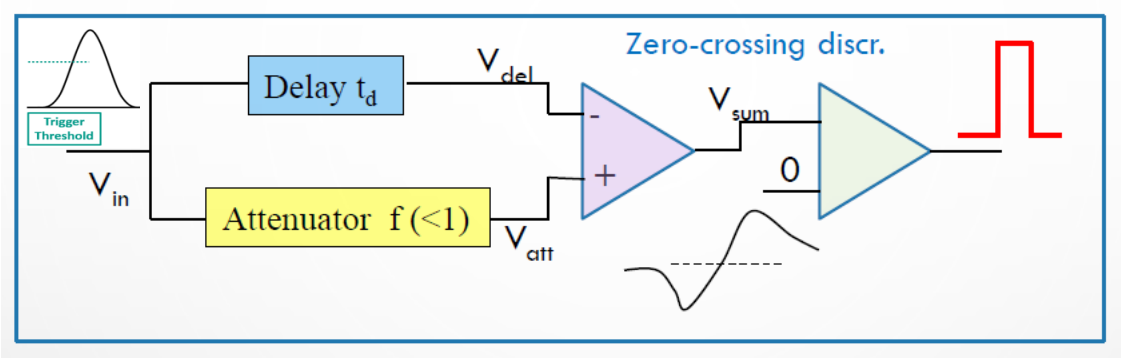
- This method reduce amplitude time walk, but *usually jitter is larger* than leading edge triggering

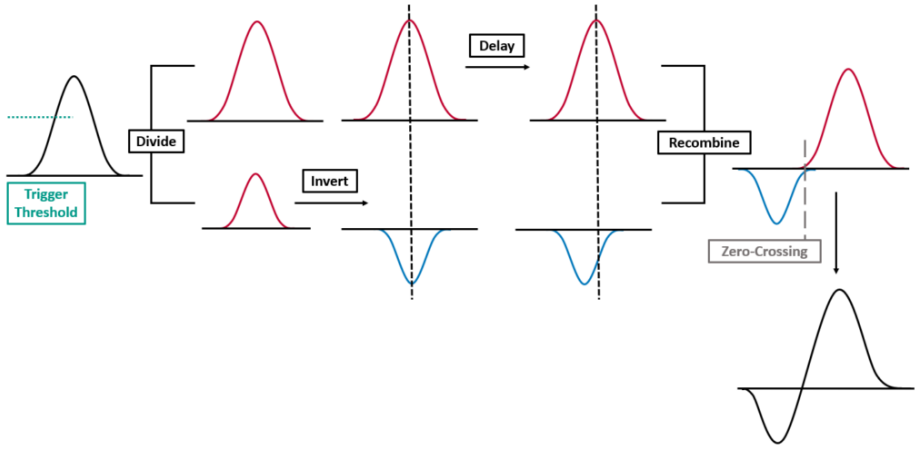
Improve the timing accuracy

reduce the timing resolution

Time walk correction

Constant Fraction timing





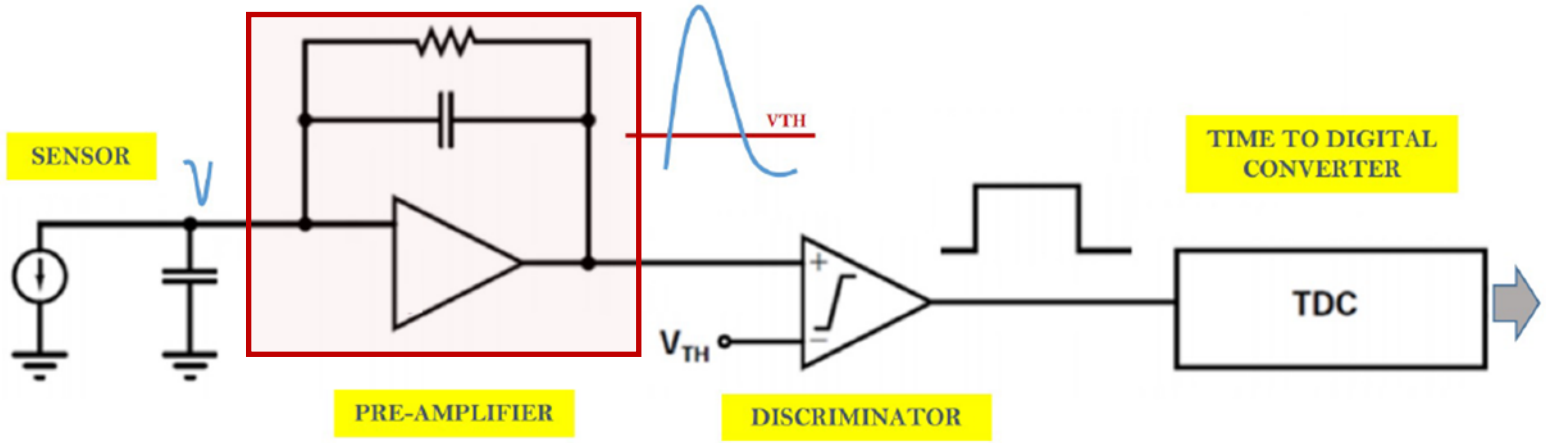
- Summing:
  - inverted and delayed signal, with $t_d > t_{rise}$
  - attenuated signal
- It can be demonstrated that the Constant Fraction Discriminator (CFD) time of resulting bipolar signal is independent of pulse amplitude for all pulses with constant shape
- Resulting jitter for optimal parameters (td, f) is lower than cross-over discriminator technique

hugh the timing accuracy

high the timing resolution

Readout chains of fast silicon sensors

What makes the different between normal sensor and fast sensor?



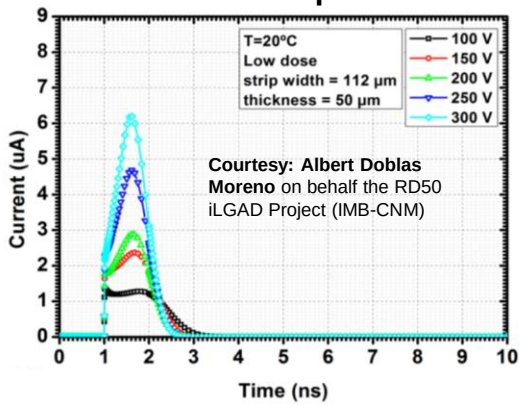
- Readout chain consists of:
  - Preamplifier (first stage, the noise strongly depend from this stage)
  - Fast comparator
  - High resolution TDC

What is the best preamplifier devices

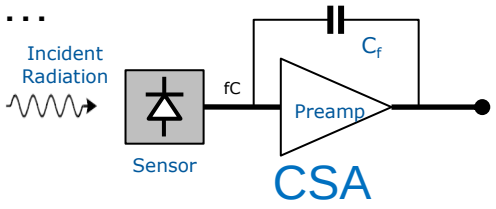
Charge Sensitive *vs* Transimpedance Amplifier

- For planar silicon sensors, the charge collected consists of a limited number of carriers over a period longer than 10 ns, resulting in a very low associated current. Consequently, a Charge-Sensitive Amplifier (CSA) is essential for signal processing.
- In contrast, LGAD sensors generate a significantly larger charge signal over a much shorter time (~1 ns) compared to the >10 ns typical of standard pixel or monolithic sensors. As a result, the associated signal is a fast current pulse

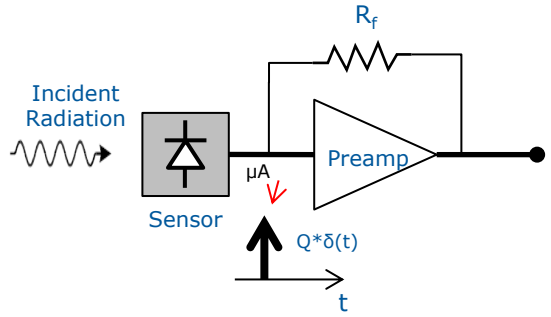
TCAD Simulation 50 µm thick LGAD



- From …



- To …



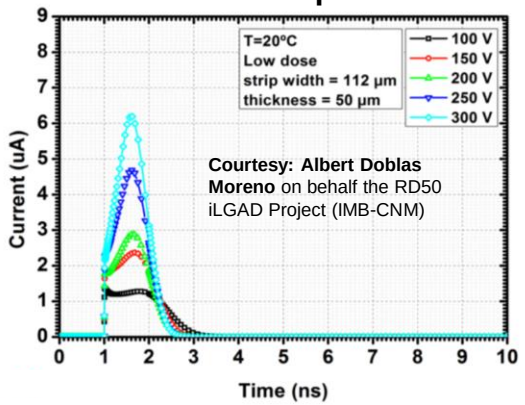
Transimpedance

What is the best preamplifier devices

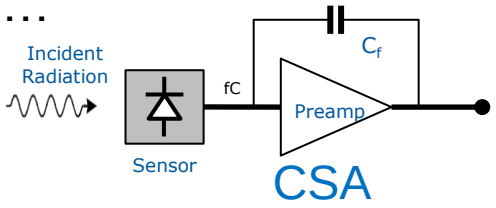
Charge Sensitive *vs* Transimpedance Amplifier

- For planar silicon sensors, the charge collected consists of a limited number of carriers over a period longer than 10 ns, resulting in a very low associated current. Consequently, a Charge-Sensitive Amplifier (CSA) is essential for signal processing.
- In contrast, LGAD sensors generate a significantly larger charge signal over a much shorter time (~1 ns) compared to the >10 ns typical of standard pixel or monolithic sensors. As a result, the associated signal is a fast current pulse

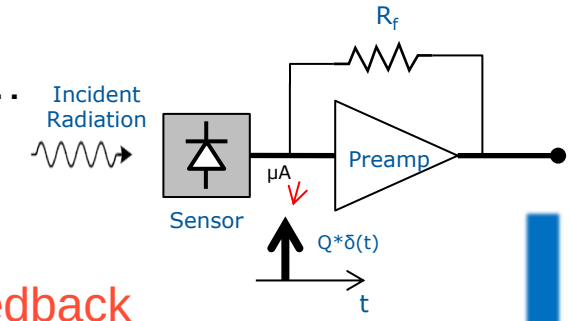
TCAD Simulation 50 µm thick LGAD



- From …



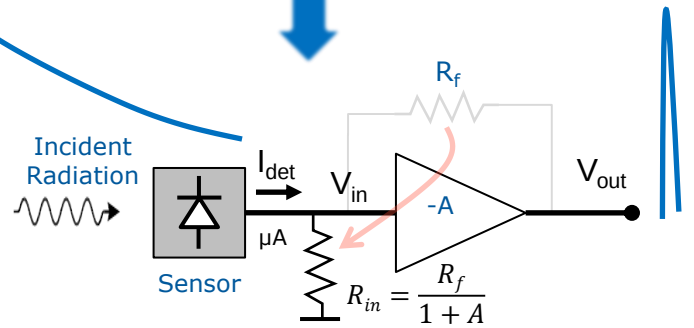
- To …



Feedback

- $V_{out} = -A\,Vin = -A\,Idet * Rin = -A\,Idet\frac{R_f}{1+A} \cong Ide_t\,Rf$

By Miller theorem

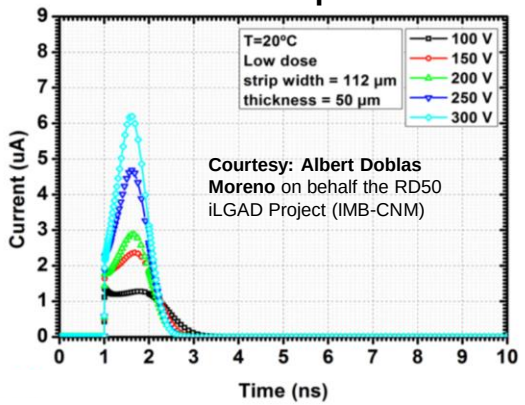


What is the best preamplifier devices

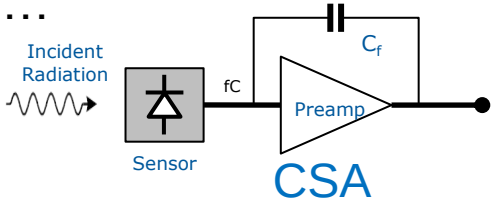
Charge Sensitive *vs* Transimpedance Amplifier

- For planar silicon sensors, the charge collected consists of a limited number of carriers over a period longer than 10 ns, resulting in a very low associated current. Consequently, a Charge-Sensitive Amplifier (CSA) is essential for signal processing.
- In contrast, LGAD sensors generate a significantly larger charge signal over a much shorter time (~1 ns) compared to the >10 ns typical of standard pixel or monolithic sensors. As a result, the associated signal is a fast current pulse

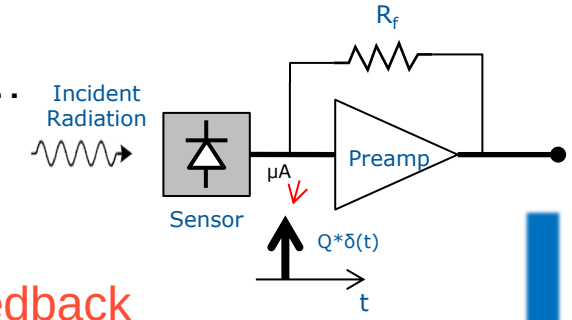
TCAD Simulation 50 µm thick LGAD



- From …



- To …

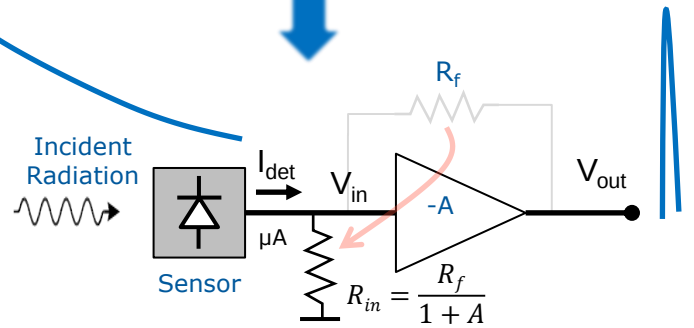


Transimpedance

Feedback

By Miller theorem

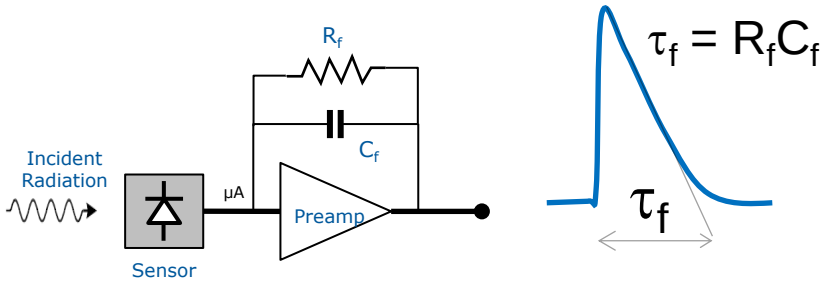
- $V_{out} = -A\,Vin = -A\,Idet * Rin = -A\,Idet\frac{R_f}{1+A} \cong Ide_t\,Rf$



- The *gain of a* **transimpedance amplifier is determined directly by the feedback resistor** R_f. For example if the gain is = 80 dB → R_f = 10 kΩ resulting in an output voltage V_{out}=60 mV, requiring additional power
- The **input impedance** $R_{in} = R_f/(1+A)$, for an open-loop gain A = 40 dB (A=1000), R_{in} ~ 10Ω
- The **time constant**= $\tau = C_{det} * R_{in}$ = is therefore very small, leading to extremely fast pulse response times with a steeper slope.

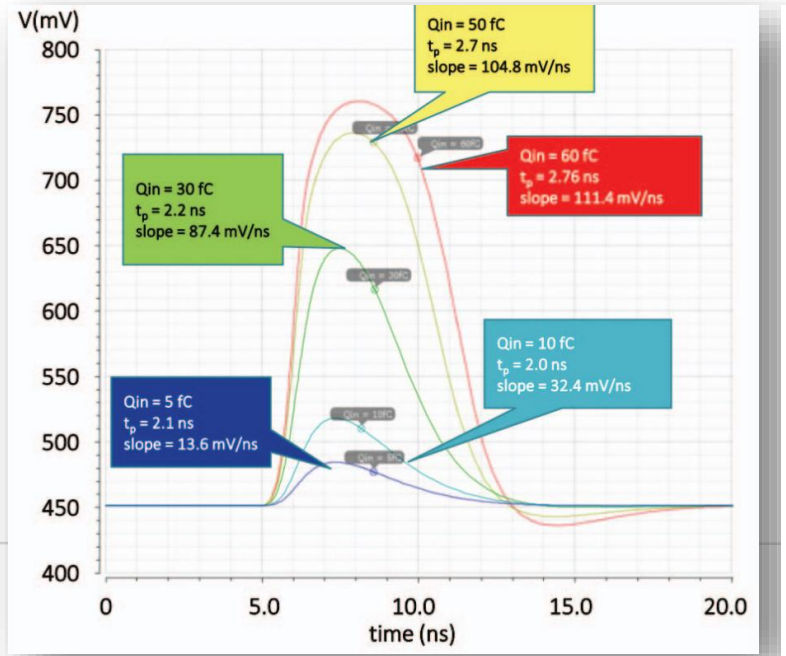
What is the best preamplifier devices

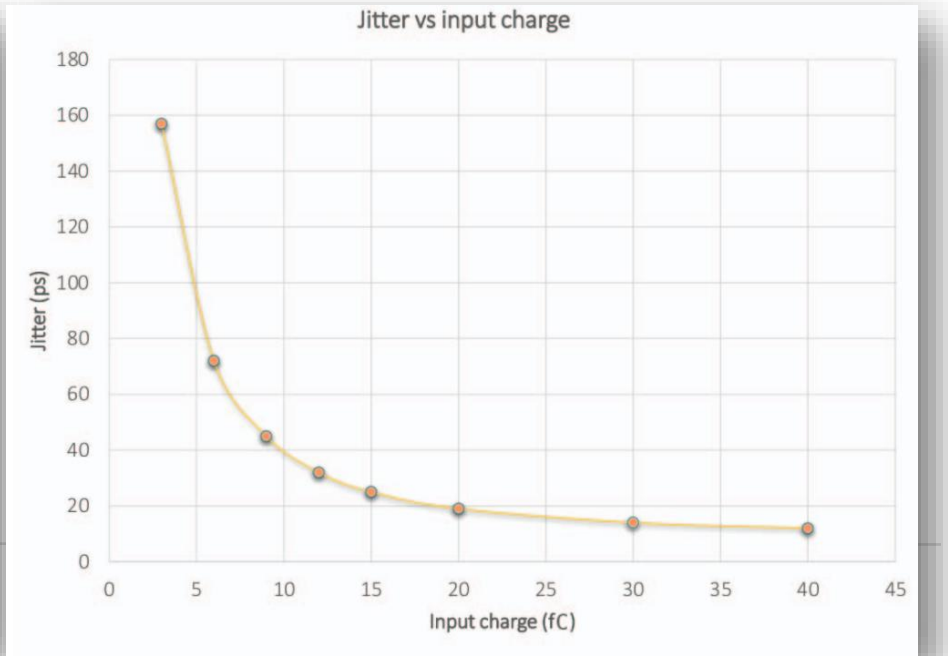
Why not both together ?

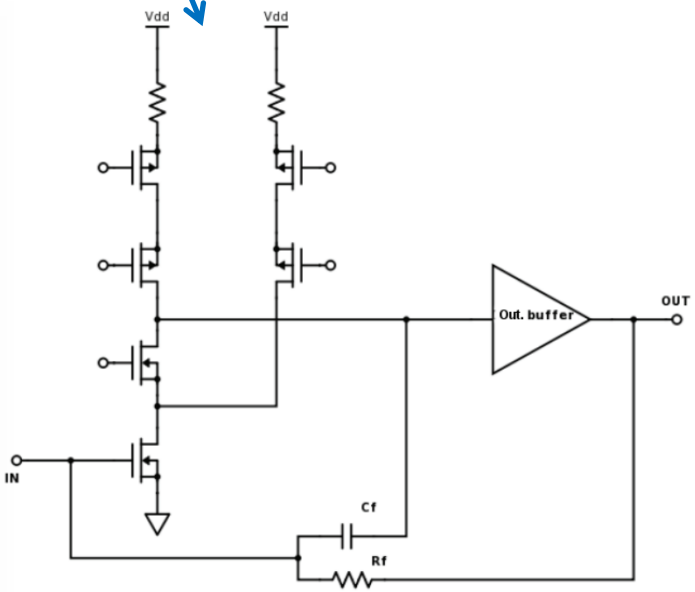


- The gain of the circuit is determined by the feedback capacitance C_f, similar to a Charge-Sensitive Amplifier (CSA). Specifically, $V_{out} = Q_{in}/C_f$, where C_f is typically in the femtofarad range.
- The feedback resistor R_f is used to shape the pulse width and is usually in the range of a few kΩ
- Enable a simple feedback mechanism for charge resetting, and the main amplifier design remains relatively uncomplicated

- TOFFEE: a Fully Custom Amplifier-comparator Chip for Silicon Detectors with Internal Gain https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8069854

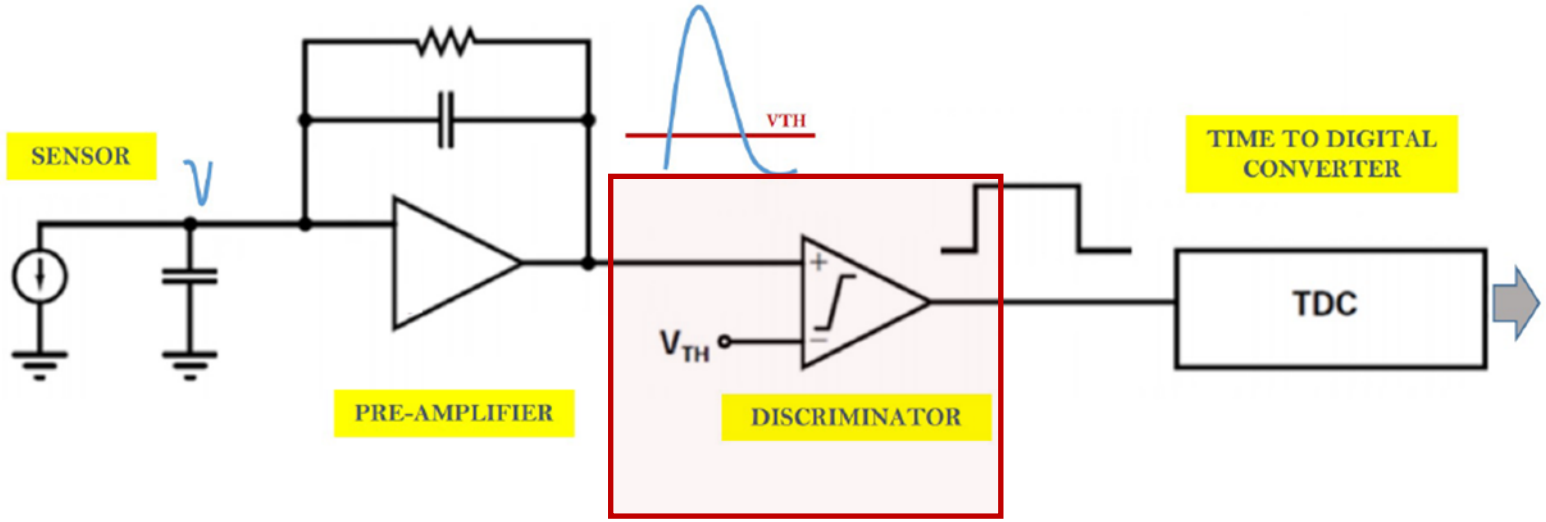






Readout chains of fast silicon sensors

What makes the different between normal sensor and fast sensor?



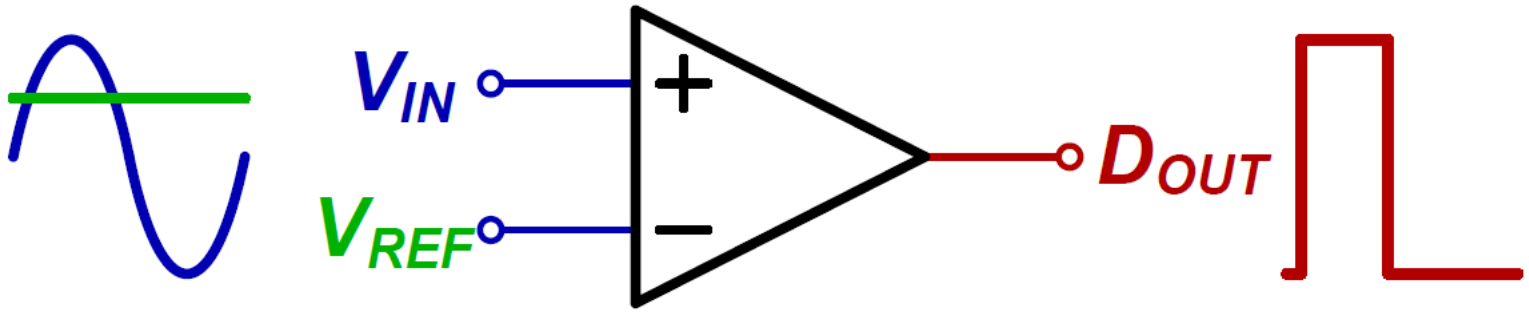
- Readout chain consists of:
  - Preamplifier (first stage, the noise strongly depend from this stage)
  - Fast comparator
  - High resolution TDC

Comparator architecture overview

Applications of comparators

- **Basic building block of an A/D converter**

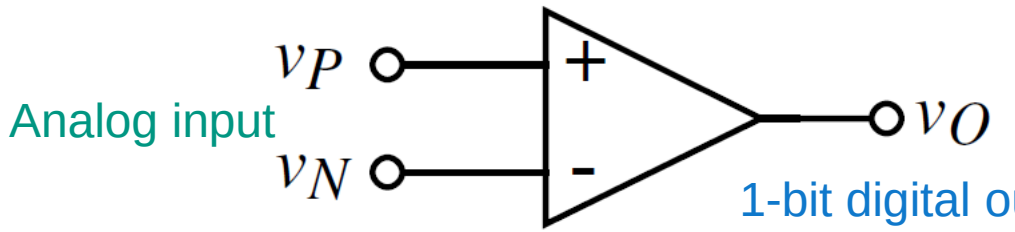
- Output amplifies difference between V_{IN} & V_{REF} with a large gain, output is ‘digital’ at either the positive or negative supply rail



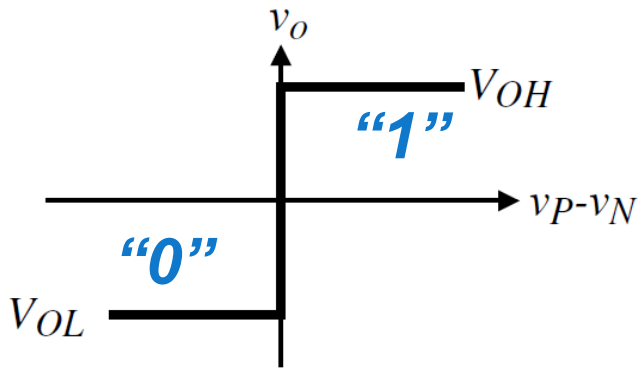
Introduction to the comparator

What is a Comparator?

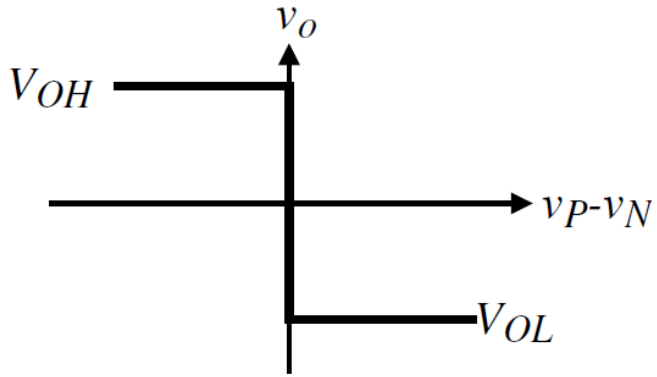
- The comparator is essentially a 1-bit analog-digital converter
  - Input is analog
  - Output is digital
- The comparator output is binary with the two-level outputs:
  - V_{OH} = the high output of the comparator
  - V_{OL} = the low level output of the comparator
- Circuit symbol for a comparator



Voltage transfer function



Noninverting comparator

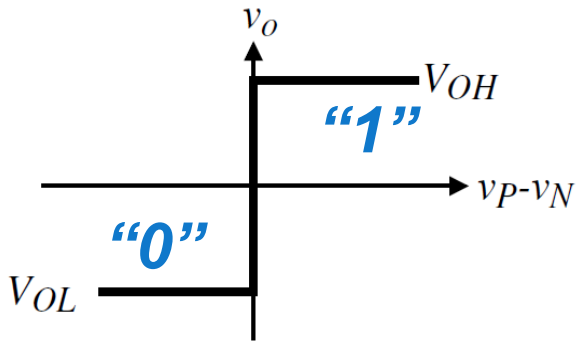


Inverting comparator

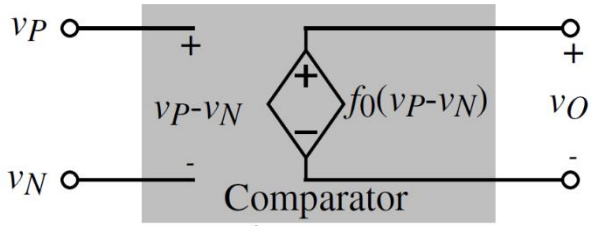
Static characteristics

"*Ideal*" and "*real*" transfer curves

- **Ideal comparator** the output changes states for an input ΔV that approaches zero, which implies a gain of infinity



Model of Noninverting comparator

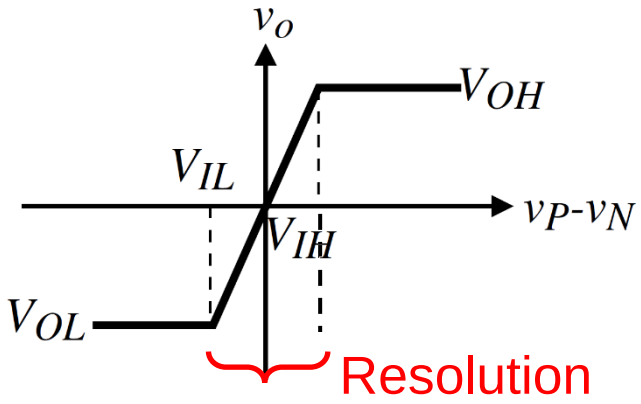


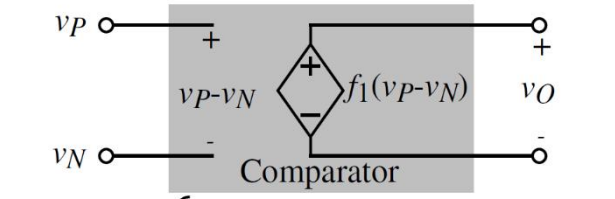
$$f_0(v_P\text{-}v_N) = \begin{cases} V_{OH} \text{ for } (v_P\text{-}v_N) > 0 \\ V_{OL} \text{ for } (v_P\text{-}v_N) < 0 \end{cases}$$

$$\text{Gain} = A_v = \lim_{\Delta V \to 0} \frac{V_{OH}\text{-}V_{OL}}{\Delta V}$$

where ΔV is the input voltage

- First **nonideal** effect of comparator is a **limited** gain Av





$$f_1(v_P\text{-}v_N) = \begin{cases} V_{OH} \text{ for } (v_P\text{-}v_N) > V_{IH} \\ A_v(v_P\text{-}v_N) \text{ for } V_{IL} < (v_P\text{-}v_N) < V_{IH} \\ V_{OL} \text{ for } (v_P\text{-}v_N) < V_{IL} \end{cases}$$

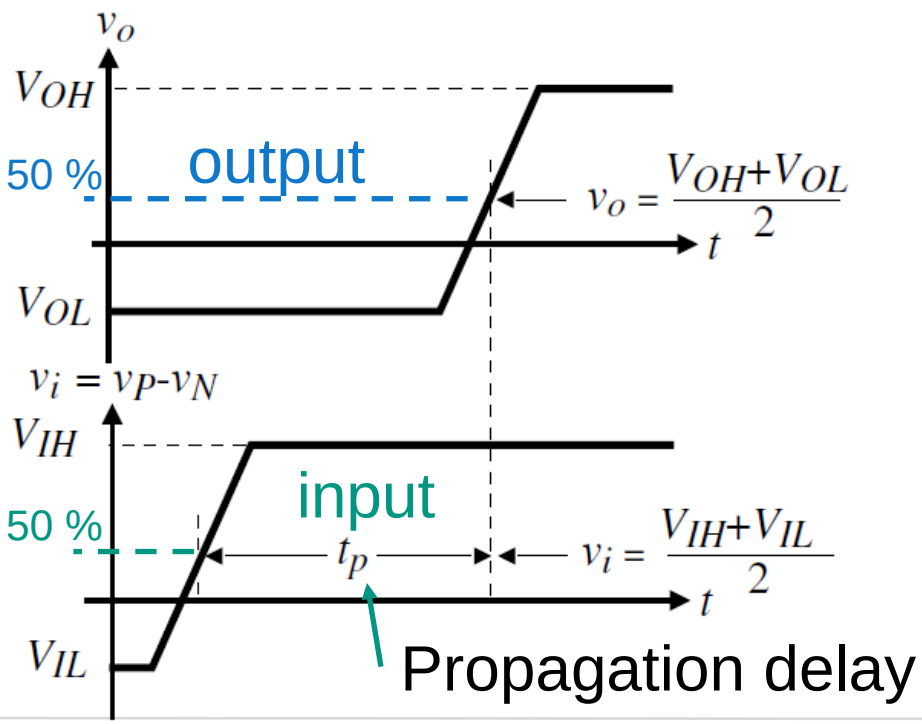
Voltage gain is: $A_v = \dfrac{V_{OH} - V_{OL}}{V_{IH} - V_{IL}}$

- ***DEF***: The minimum amount on input voltage necessary to make the output swing between two binary states is defined: **resolution of the comparator**

Dynamic characteristics

Propagation delay time

- The delay between the input analog signal and digital output response of the comparator is defined **propagation delay time**
- The propagation delay is very important because limits the conversion speed rate of an A/D converters

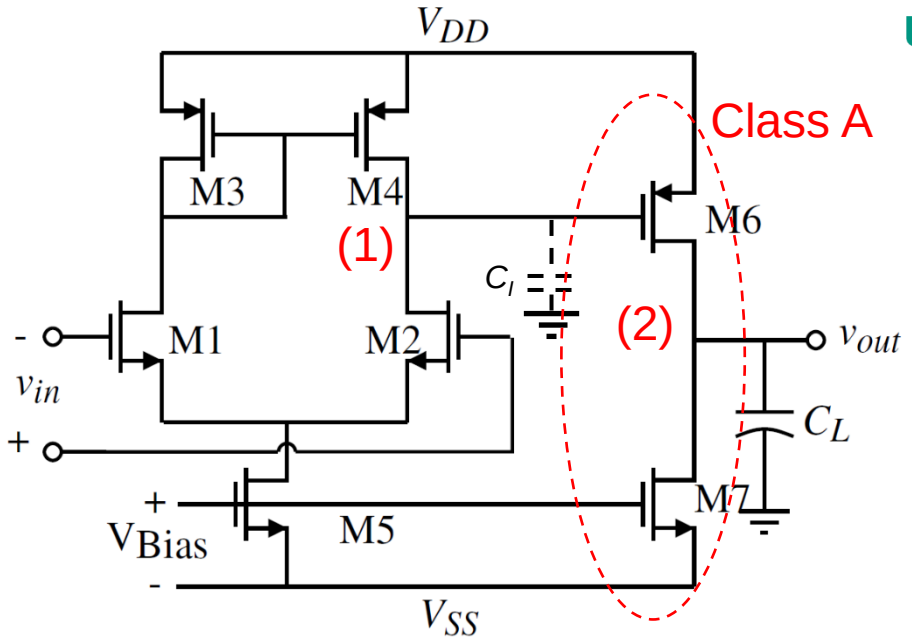


- The propagation delay generally varies as function of the amplitude of the input
- Large input → smaller delay time
- Propagation delay **is large** when the comparator operates in **small-signal mode** and **smaller** when operates in **slew-rate mode**
- The comparator operates in **slewing** or **slew-rate** when further increase in the input will not long effect the delay propagation delay

Two-stage, open-loop comparators

Propagation delay time

- The comparator requires: differential input a high-gain to be able to achieve the desired resolution, the two-stage op amp makes an excellent implementation of the comparator



- The comparator generally works in open-loop mode, therefore it is not necessary any compensation capacitance → it is **preferable not compensate so that it has a large bandwidth possible**
- The small-signal voltage gain: $A_v(0) = \underbrace{\left(\frac{g_{m1}}{g_{ds2}+g_{ds4}}\right)}_{(1)}\underbrace{\left(\frac{g_{m6}}{g_{ds6}+g_{ds7}}\right)}_{(2)}$
- The poles are:

$$p_1 = \frac{-(g_{ds2}+g_{ds4})}{C_I} \quad (1)$$

Output: $$p_2 = \frac{-(g_{ds6}+g_{ds7})}{C_{II}} \quad (2)$$ $C_{II} = C_L + C_{GD6} + C_{GD7}$

- The frequency response is: $A_v(s) = \frac{A_v(0)}{\left(\frac{s}{p_1}+1\right)\left(\frac{s}{p_2}+1\right)}$

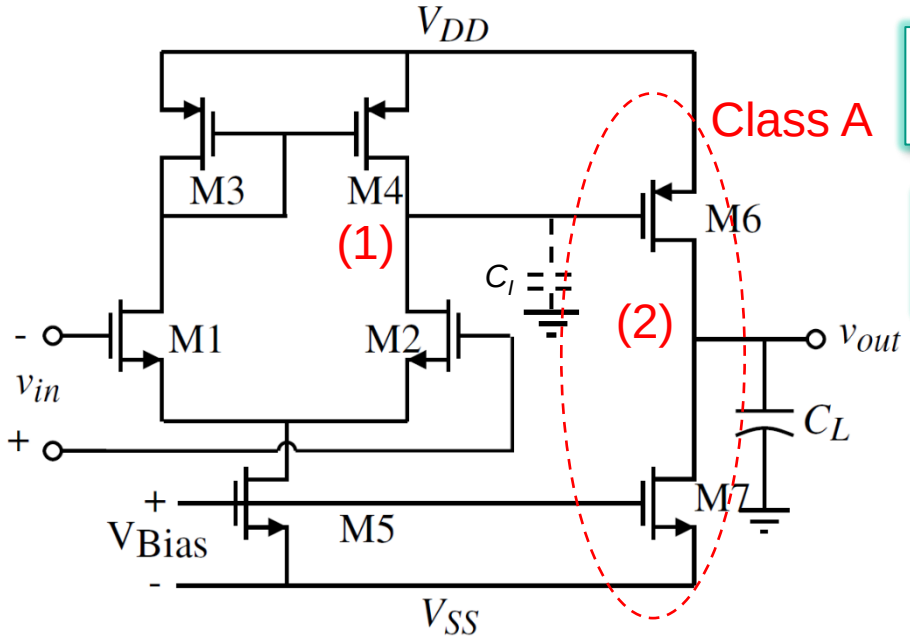
$$v_{out}(t) = A_v(0)V_{in}\left[1 + \frac{p_2 e^{-tp_1}}{p_1 - p_2} - \frac{p_1 e^{-tp_2}}{p_1 - p_2}\right]$$

The time response at step input for small-signal (slew-rate not occurred)

Two-stage, open-loop comparators

Propagation delay time

- Assuming that the slew-rate does not occur

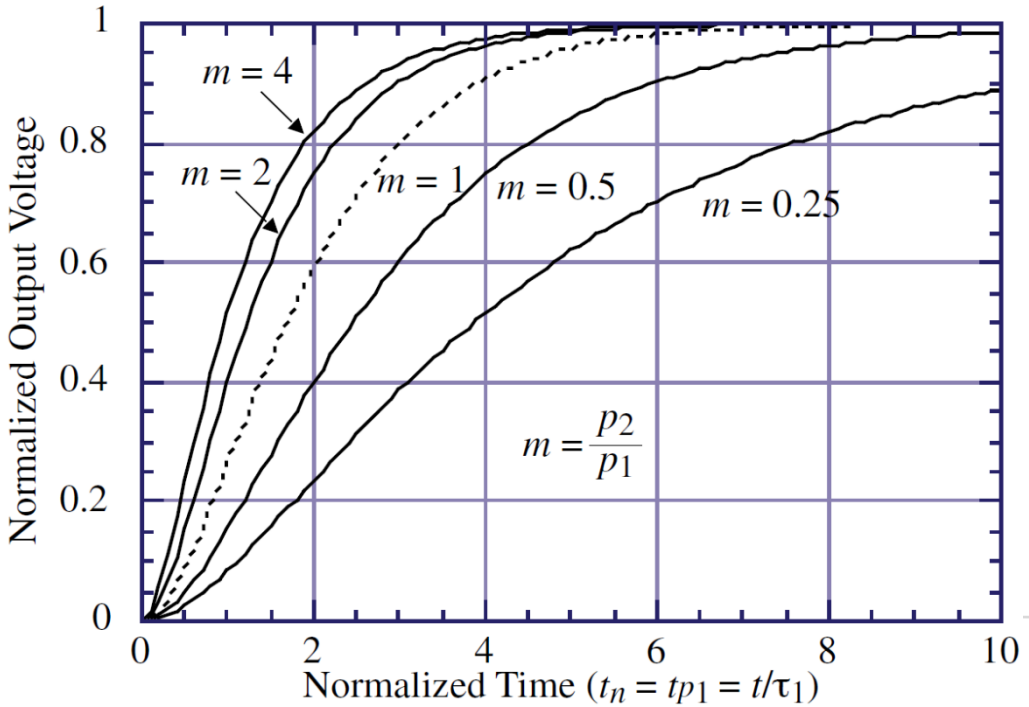


$$v_{out}(t) = A_v(0)V_{in}\left[1 + \frac{p_2 e^{-tp_1}}{p_1 - p_2} - \frac{p_1 e^{-tp_2}}{p_1 - p_2}\right]$$

$$v_{out}'(t_n) = \frac{v_{out}(t)}{A_v(0)V_{in}} = 1 - \frac{m}{m-1}e^{-t_n} + \frac{1}{m-1}e^{-mt_n}$$

- We define m and t_n

$$m = \frac{p2}{p1} \neq 1 \quad \text{and} \quad t_n = tp1 = \frac{t}{\tau_1}$$



- The propagation delay time reduce m > 1 ($p_2 >> p_1$)
- In case of slew-rate mode, the propagation delay are:

$SR^- = I_7/C_{II}$ Negative slew-rate

$SR^+ = (I_6 - I_7)/C_{II}$ Positive slew-rate

Similar to a Class A

Dynamic characteristics

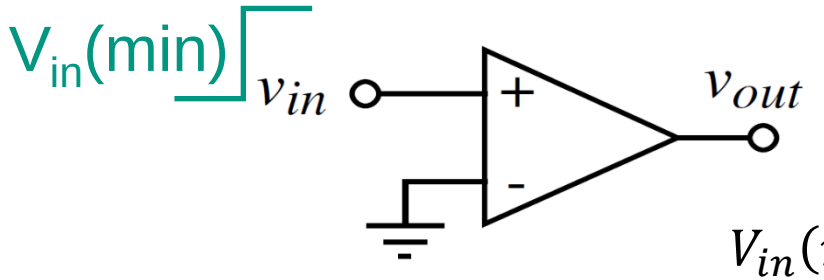
Single pole response

- The small-signal dynamics are characterized by the frequency response of comparator that are similar to the frequency response of a single-pole amplifier

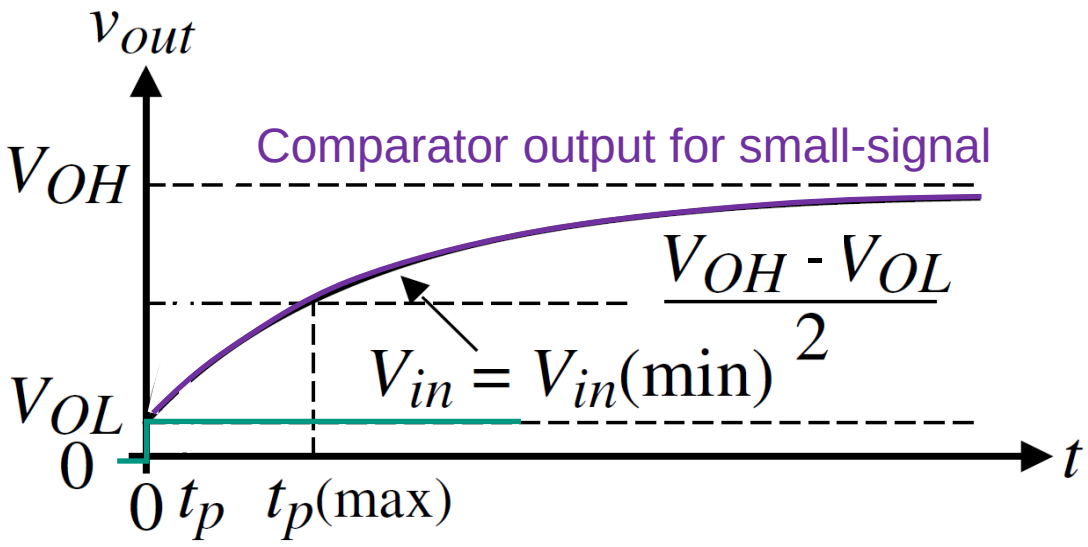
$$A_v(s) = \frac{A_v(0)}{\frac{s}{\omega_c} + 1} = \frac{A_v(0)}{s\tau_c + 1}$$

- $A_v(0)$ is the dc gain
- $\omega_c = 1/\tau_c$ is the -3dB frequency of the single (dominant) pole

- Step response: $v_O(t) = A_v(0)\,[1 - e^{-t/\tau_c}]V_{in}$ (1)



$$V_{in}(min) = \frac{V_{OH} - V_{OL}}{A_v(0)}$$



- The worst propagation delay t_p(max) → when the V_{in} is close to the resolution of the comparator

Comparators

Comparators are divided in three classes

High resolution + High speed = High speed and High resolution

High voltage gain amplification

- Av = very high

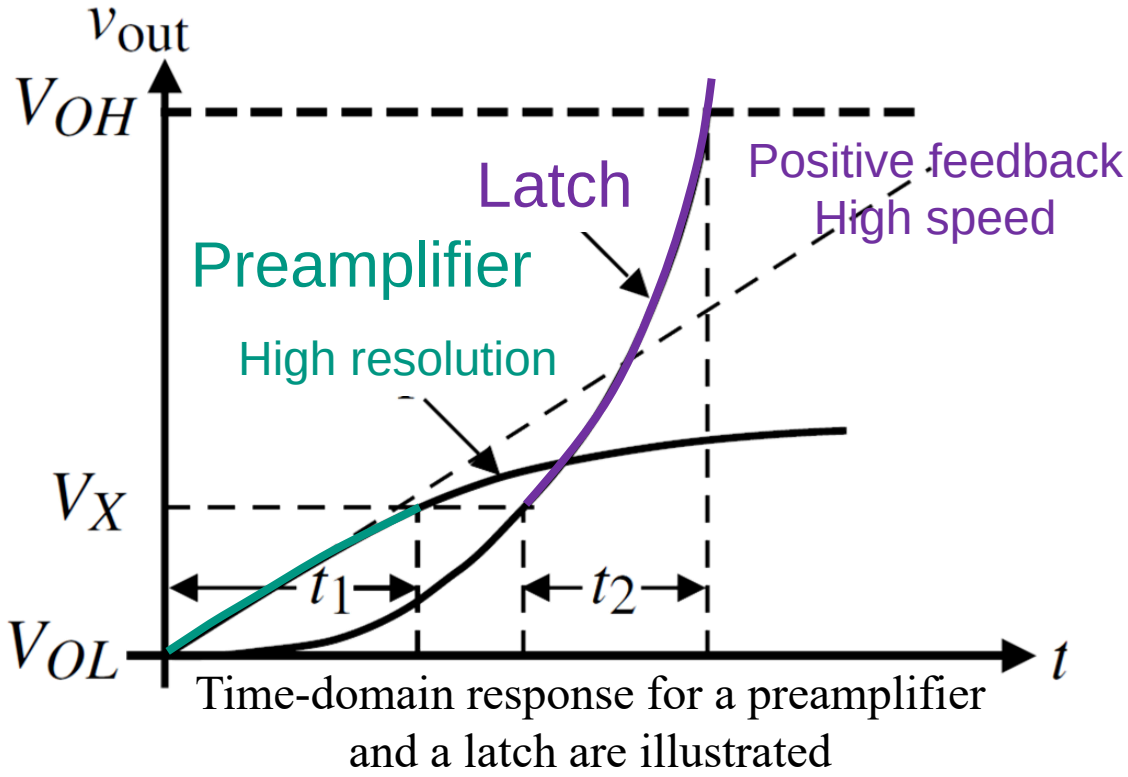
Very low propagation delay

- Bandwidth = very high
- Circuit based on positive feedback

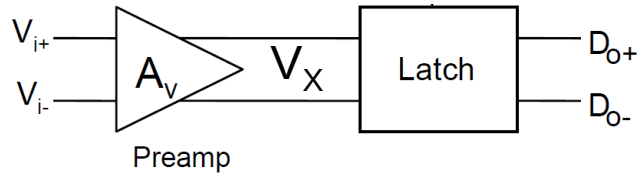
High-speed Comparators

Very fast comparator and very high resolution

- A high-speed comparator should have a propagation delay time as small as possible
- By separating the comparator into a number of cascaded stages with a gain of A_0 and a single pole at $1/\tau$ (very high-frequency)
- The basic principle behind the high-speed comparator is to use a **preamplifier** to build up the input change to a sufficiently large value (V_X) and then apply it to the **latch**. This combines the best aspects of circuits with a negative exponential response (the preamplifier) with circuits with a positive exponential response (the latch). ***Total response time is $t_1 + t_2$***
- Only preamplifier: the transition from V_{OL} to V_{OH} longer than t1 + t2
- Only latch: would require longer time if the input is small



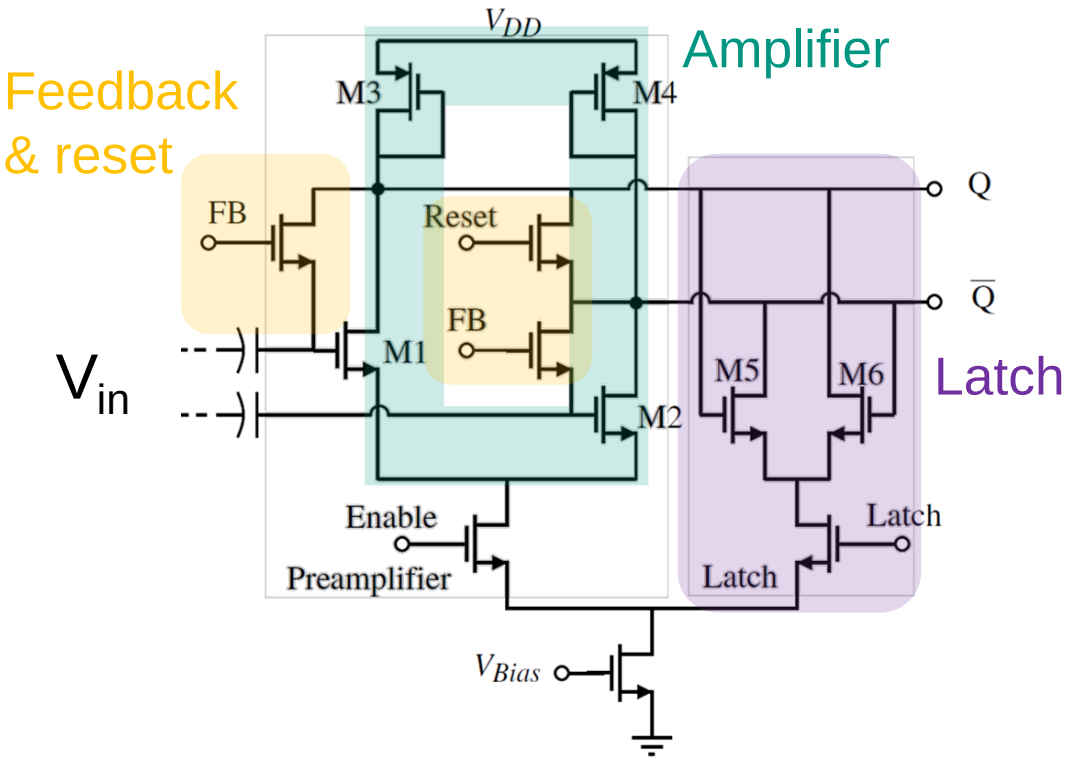
Time-domain response for a preamplifier and a latch are illustrated



High-speed Comparators

High-speed comparators, circuit implementation

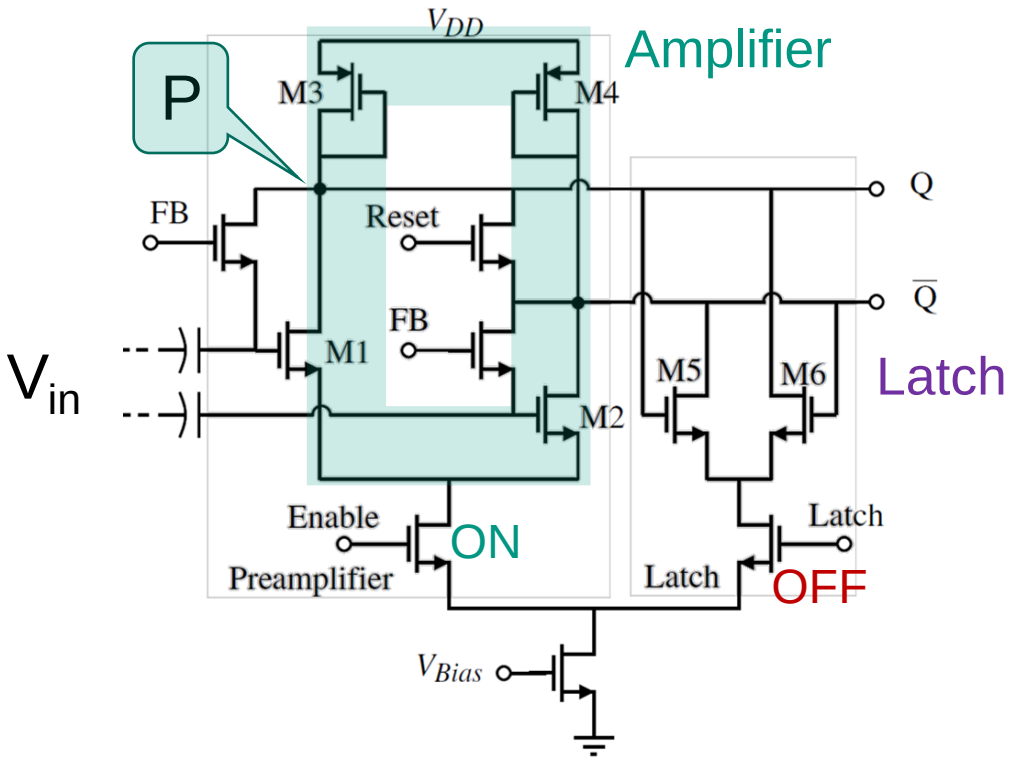
- The low-gain preamplifiers must compromise between a high bandwidth and sufficient gain



High-speed Comparators

High-speed comparators, circuit implementation

- The low-gain preamplifiers must compromise between a high bandwidth and sufficient gain

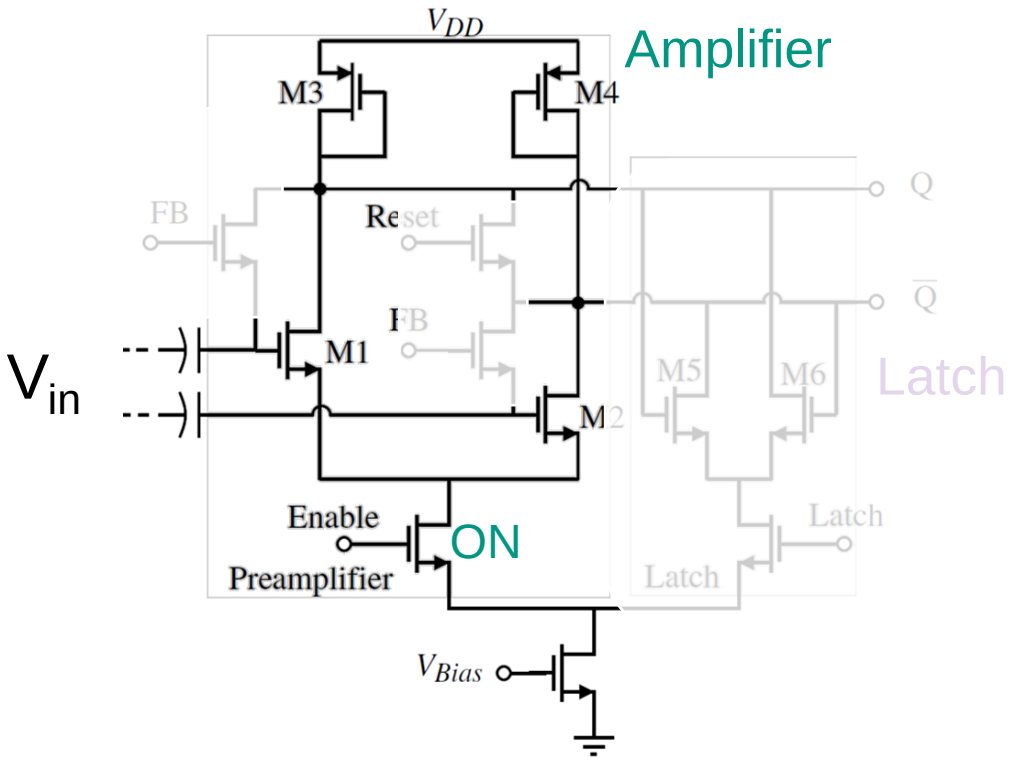


- Gain: $A_v = - g_{m1}/g_{m3} = - g_{m2}/g_{m4} = - \sqrt{\frac{K_N'(W_1/L_1)}{K_p'(W_3/L_3)}}$
- Dominant pole: $p_{dominat} = - g_{m3}/C = - g_{m4}/C$, where C is the total capacitance at the node P
- *Advantage*: high-frequency
- Working principle …

High-speed Comparators

High-speed comparators, circuit implementation

- The low-gain preamplifiers must compromise between a high bandwidth and sufficient gain

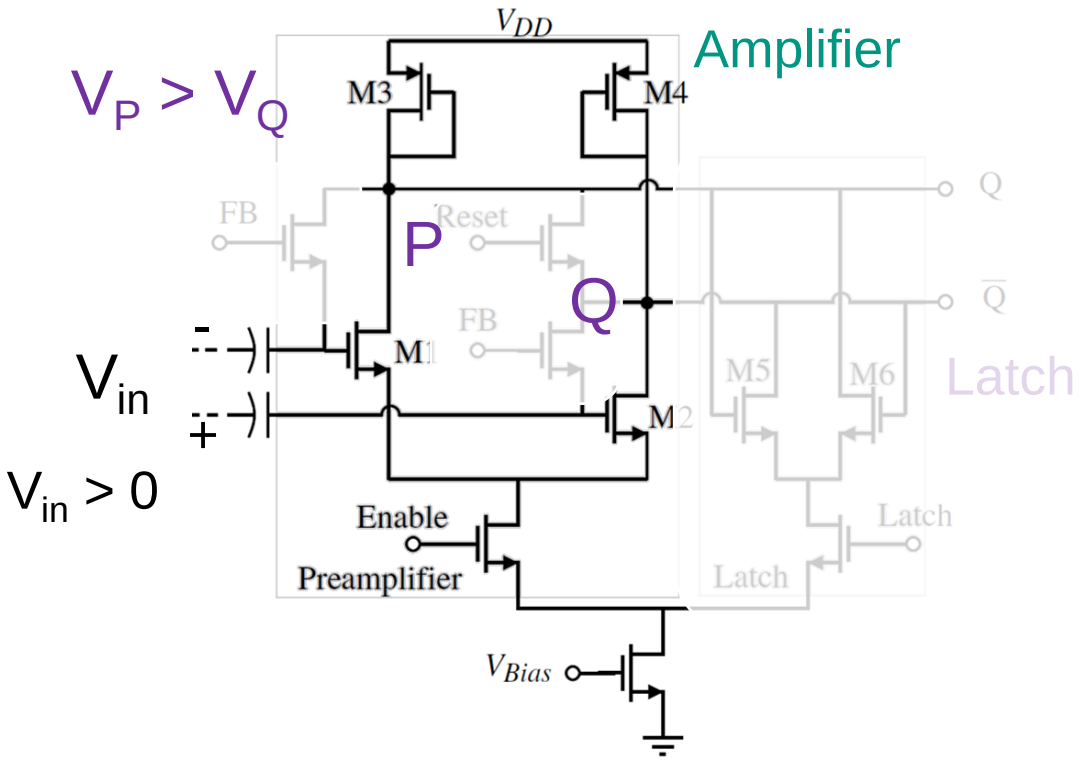


- **Step 1**: Enable Preamplifier = 1 ➔ differential stage biased by V_{bias} MOS
- **Step 2**: Reset mode (FB = 1 & Reset = 1) ➔ Q = $\bar{Q}$, note that M1 and M2 are also connected in a diode mode by the FB and Reset MOS

High-speed Comparators

High-speed comparators, circuit implementation

- The low-gain preamplifiers must compromise between a high bandwidth and sufficient gain

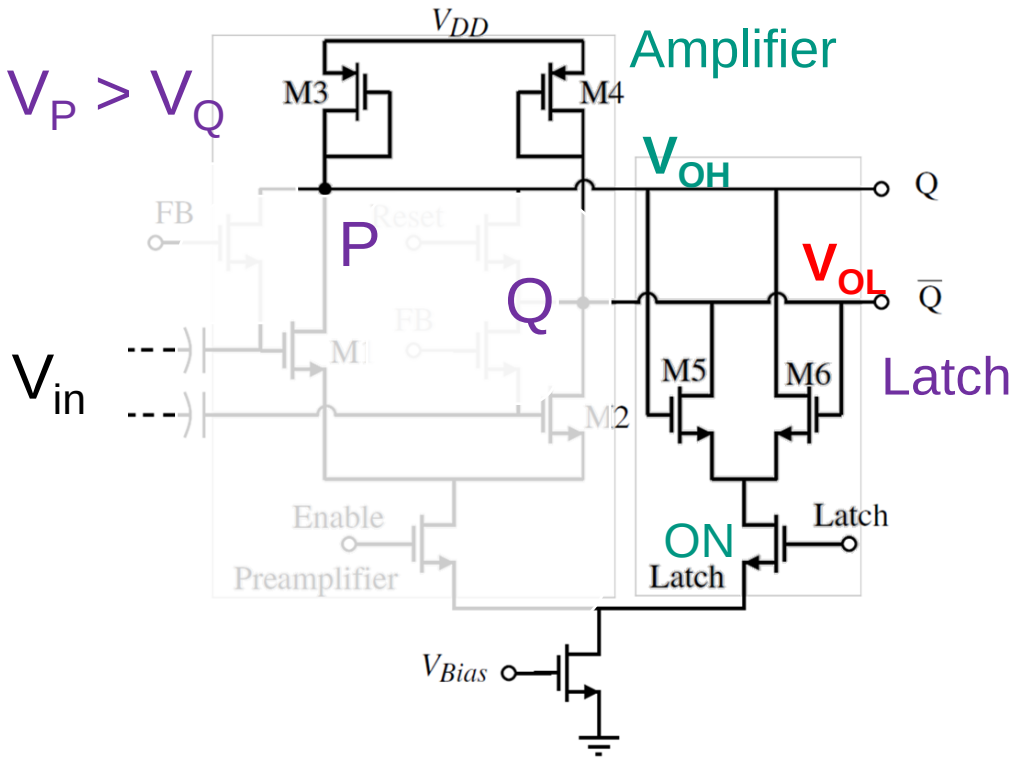


- **Step 1**: Enable Preamplifier = 1 ➔ differential stage biased by V_{bias} MOS
- **Step 2**: Reset mode (FB = 1 & Reset = 1) ➔ Q = $\bar{Q}$, note that M1 and M2 are also connected in a diode mode by the FB and Reset MOS
- **Step 3**: Preamp mode (FB = 0 & Reset = 0) $V_{out\_preamp} = - A_V V_{in}$

High-speed Comparators

High-speed comparators, circuit implementation

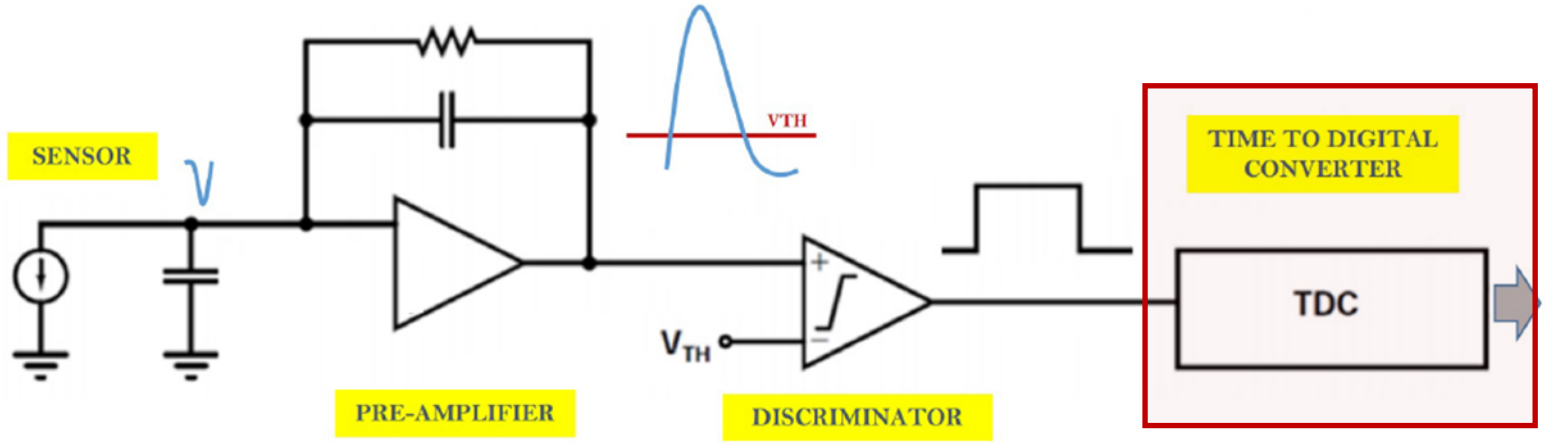
- The low-gain preamplifiers must compromise between a high bandwidth and sufficient gain



- **Step 1**: Enable Preamplifier = 1 ➔ differential stage biased by V_{bias} MOS
- **Step 2**: Reset mode (FB = 1 & Reset = 1) ➔ Q = $\bar{Q}$, note that M1 and M2 are also connected in a diode mode by the FB and Reset MOS
- **Step 3**: Preamp mode (FB = 0 & Reset = 0) $V_{out\_preamp} = A_V V_{in}$
- **Step 4**: latch mode = 1 and Enable Preamplifier = 0, the positive feedback of the latch will set the final voltage levels V_{OH} and V_{OL}

Time resolution in timing detectors

The timing capabilities depends from the signal at the output of the pre-amplifier and by the TDC binning



- $\sigma_t^2 = (\frac{Noise}{dV/dt})^2 + \sigma_{timewalk}^2 + (\frac{LSB}{\sqrt{12}})^2 + \sigma_{sensor}^2 + \sigma_{Landau}^2$

Front-end | TDC | Sensor

Total noise contribution

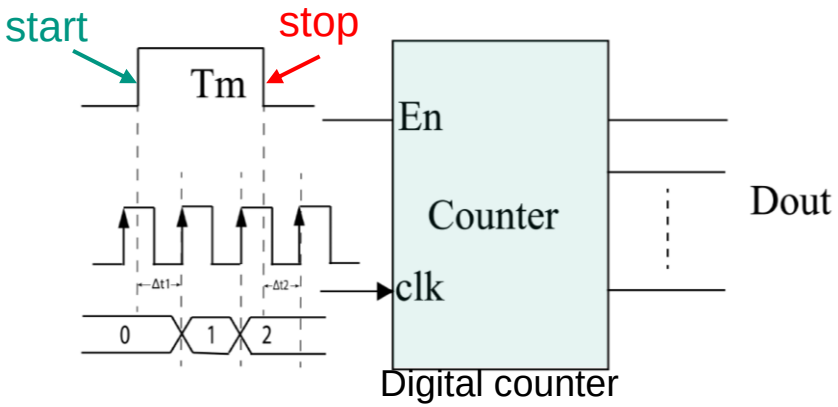
TDC: Its contribution is negligible within the time resolution for LSB < 30 ps

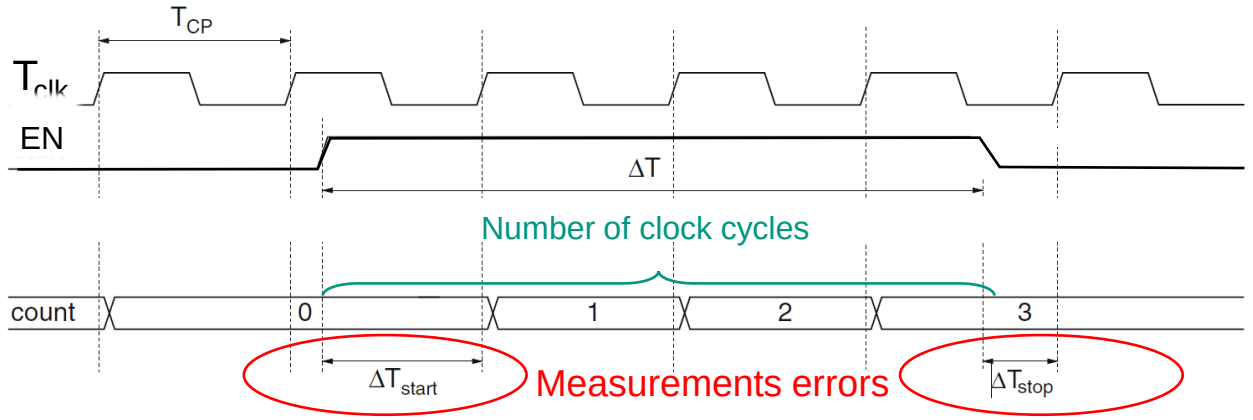
It contributes with few ps of time jitter (uncertainty)

Principle of COUNTER based TDC

The “first” TDC architecture

- Counter is the simplest and most reliable way of time measurement. The conceptual diagram of counter as TDC is as depicted in Figure, ***the resolution depends on the frequency*** of the clock input. The *advantages* of counter remains in its simplicity and reliability





- The simplest technique **to quantize a time interval is to count the cycles of a reference clock**. The measurement interval defined by the **start and stop signal is completely asynchronous to the reference clock signal** → this causes a measurement error ΔT_{start} at the beginning and ΔT_{stop} at the end of the time interval

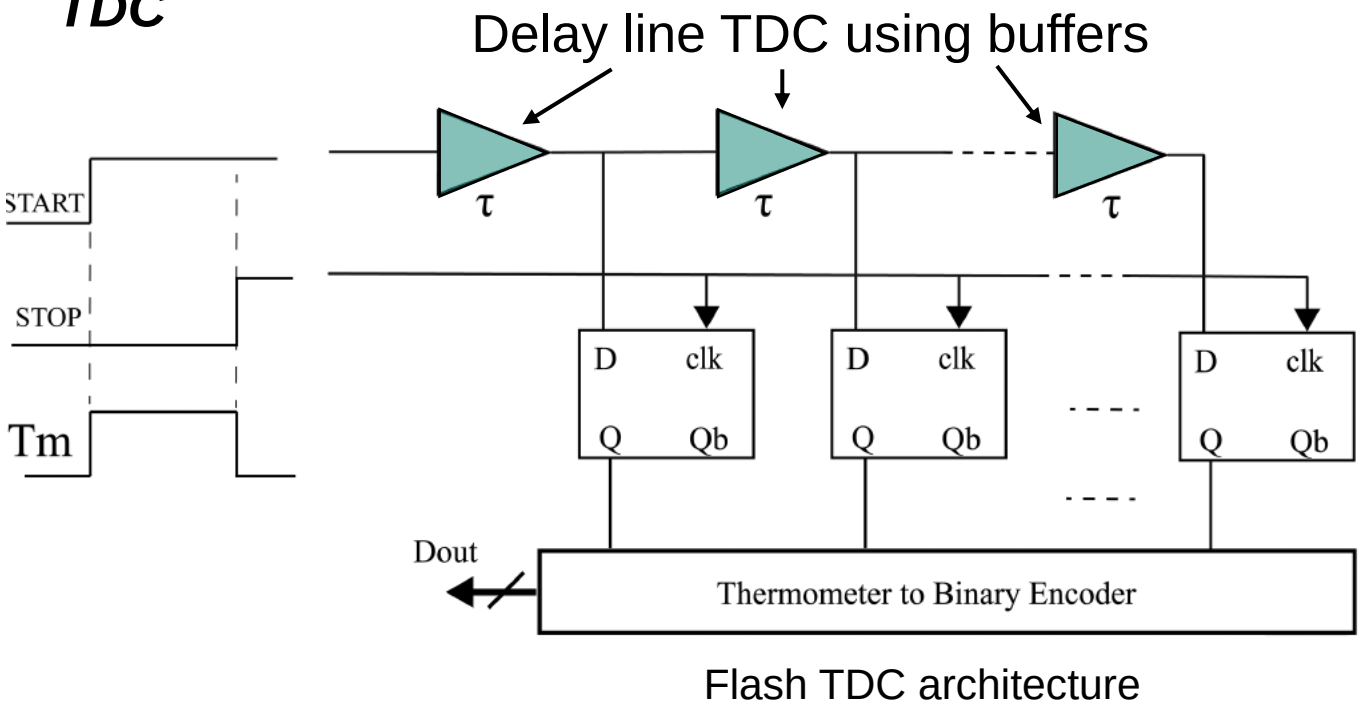
- $\Delta T = N \cdot T_{clk} + (T_{clk} - \Delta T_{stop}) - (T_{clk} - \Delta T_{start})$

Quantization error $\varepsilon_T = \Delta T_{start} - \Delta T_{stop} \in [-T_{clk}; T_{clk}]$

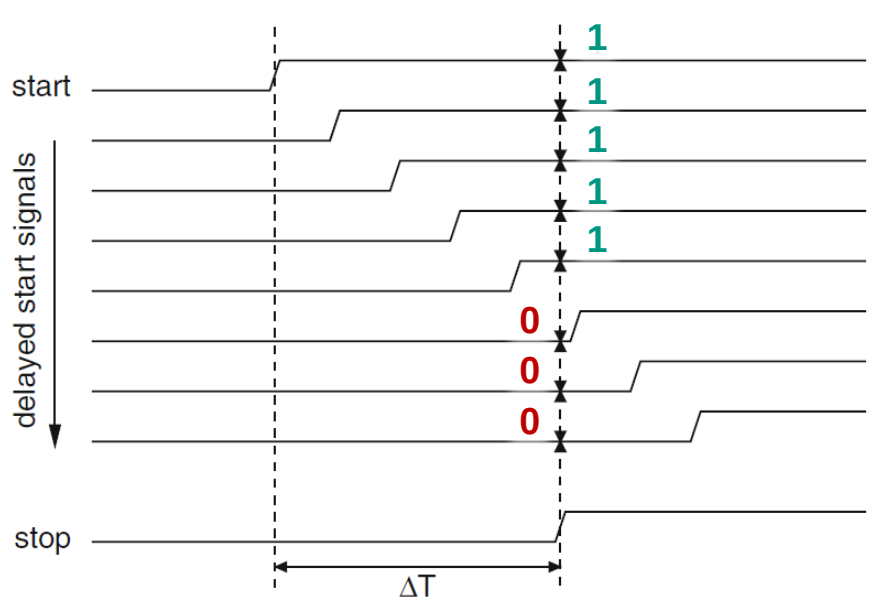
Delay line TDC – Second Generation

FLASH - TDC architecture

- Delay line TDC is also called as ***flash TDC***, ***classical delay line TDC***, or ***tapped delay line TDC***



Flash TDC architecture

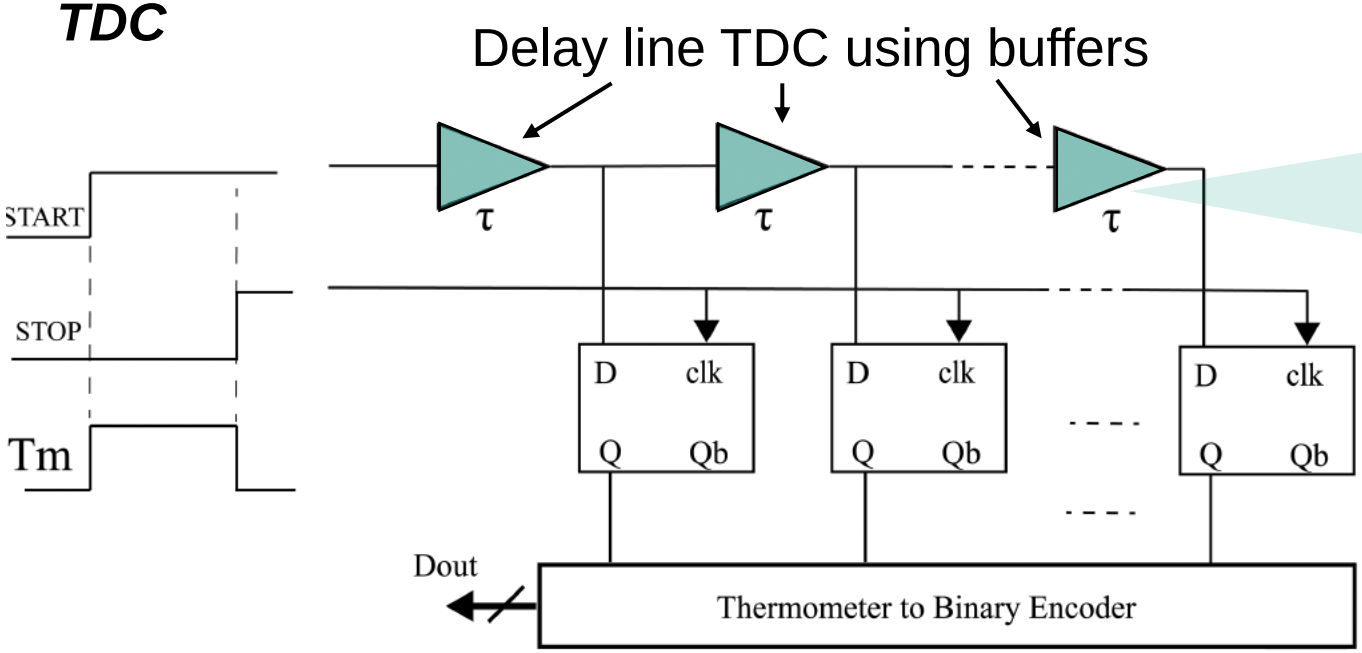


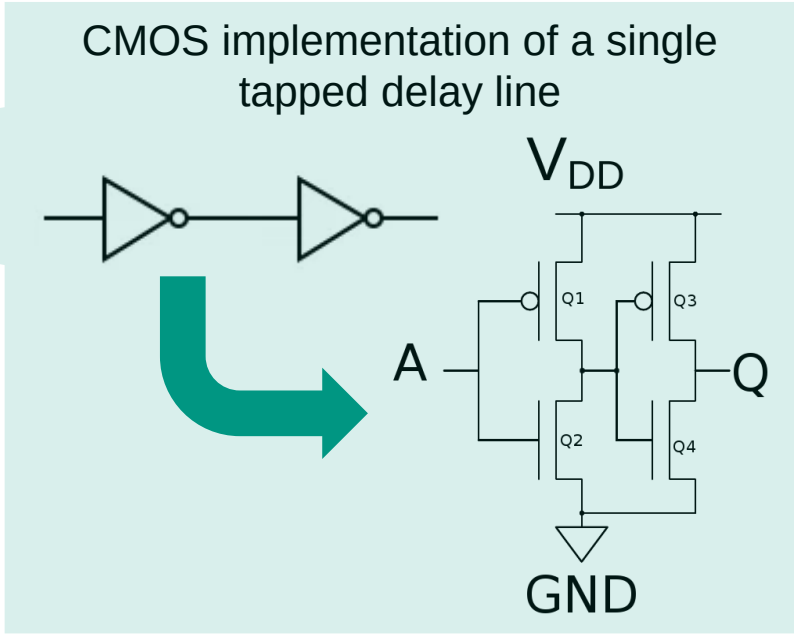
- **Working principle**: delayed versions of the start signal are sampled on the rising edge of the stop signal, resulting in a thermometer code

Delay line TDC – Second Generation

FLASH - TDC architecture

- Delay line TDC is also called as ***flash TDC***, ***classical delay line TDC***, or ***tapped delay line TDC***





- In both CMOS and FPGA technologies the resolution of the TDC is equal to the delay (τ) of the buffer
- The INL and DNL depends on the uniformity of the delays (τ)

Delay line TDC – Second Generation

FLASH - TDC architecture

Flash TDC (Mixed-signal)

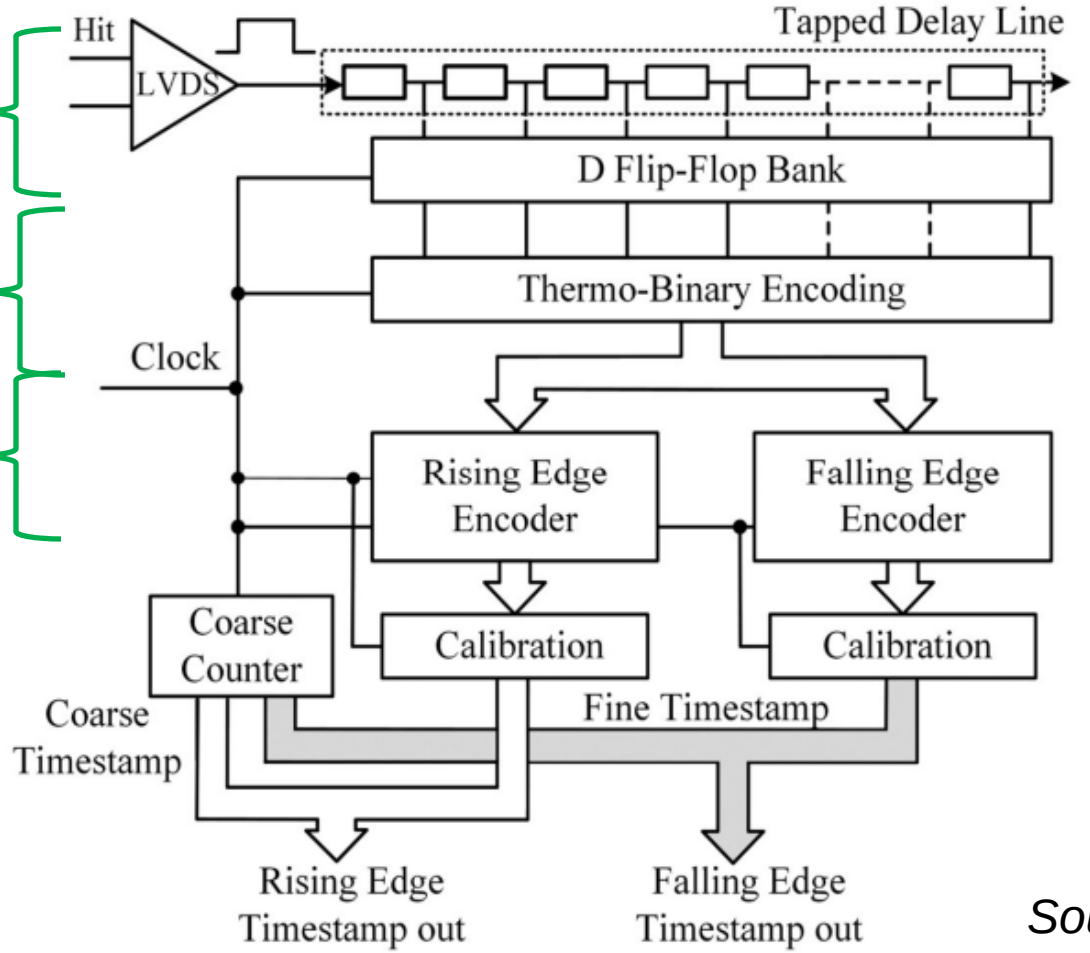
Digital encoding

Rising and Falling edge

+

Errors correction

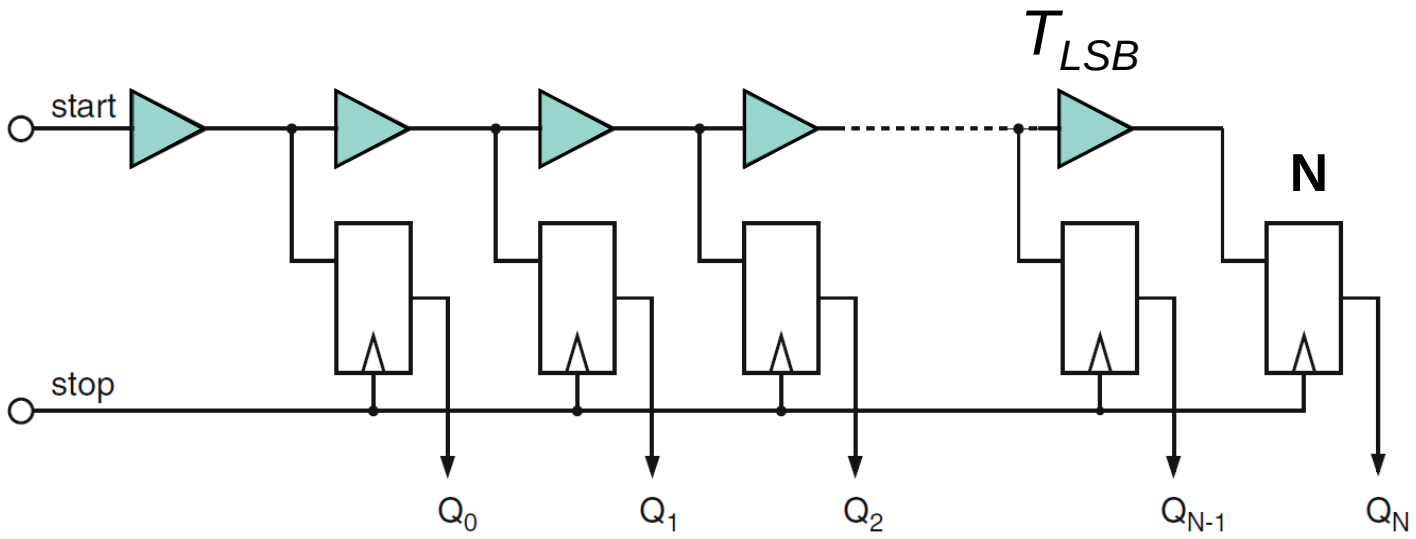
- Coarse timing (by counter)
- Fine timing (by tapped-lines)



Final digital code

Source: Xinchi Xu

TDC resolution and quantization



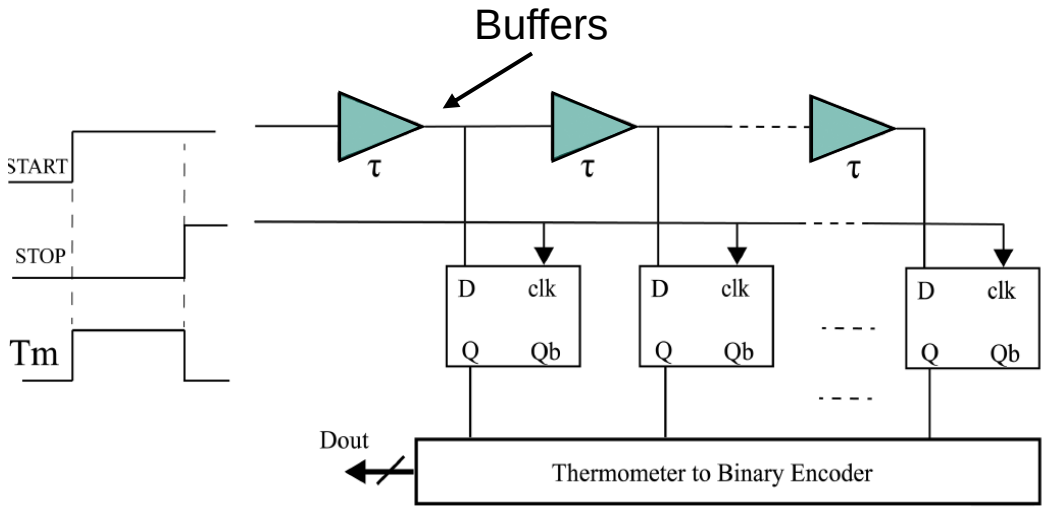
- The number N of all sampling elements with a HIGH output is related to the measurement interval ΔT according to

$$N = \left\lfloor \frac{\Delta T}{T_{LSB}} \right\rfloor$$

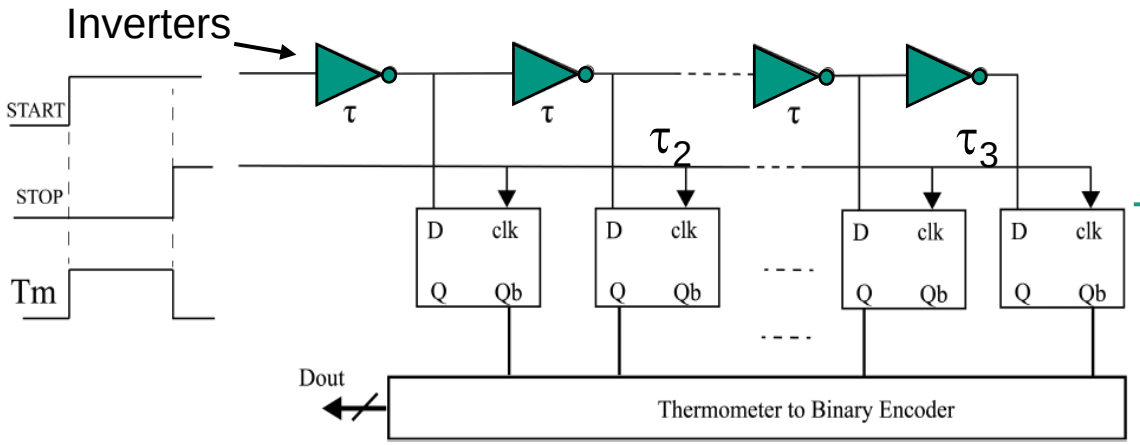
- where T_{LSB} is the delay of a single delay element in the delay line
- The time interval ΔT can be calculated from the number of HIGH outputs by **$\Delta T = N\, T_{LSB} + \varepsilon$**
- where ***ε describes the quantization error*** that arises as a delay element has been either passed by the start signal yet or not

Inverter Based Time-to-Digital Converter

Inverter vs buffer



Buffer based TDC architecture



Delay based TDC architecture

- Replace buffer (two cascaded inverters) with inverter → doubles the resolution but reduces the dynamic range by half
- The thermometer code at the outputs of the sampling elements becomes a pseudo thermometer code with alternating ones and zeros:

111111111111111 ⌋⌈ 000000000000000	buffer TDC
010101010101010 ⌋⌈ 010101010101010	inverter TDC

- The length of the measurement interval is indicated not by a HIGH-LOW transition but by a phase change of the alternation HIGH-LOW sequence

Moreover

Time-to-Digital Converters – Third Generation

TDC with Sub-Gate delay Resolution

- The resolution of time-to-digital converters **is limited by technology to one inverter delay**. Higher resolution can be achieved only by a faster technology.
- **What are the circuit techniques that allow to reach higher resolution TDC?** Thereby, do not depend on a single absolute gate delay
- The ratio between the **technology resolution** and the **actual resolution** is the so called **sub-gate delay interpolation** factor IF:

$$IF := \frac{T_{tech}}{T_{LSB}}$$

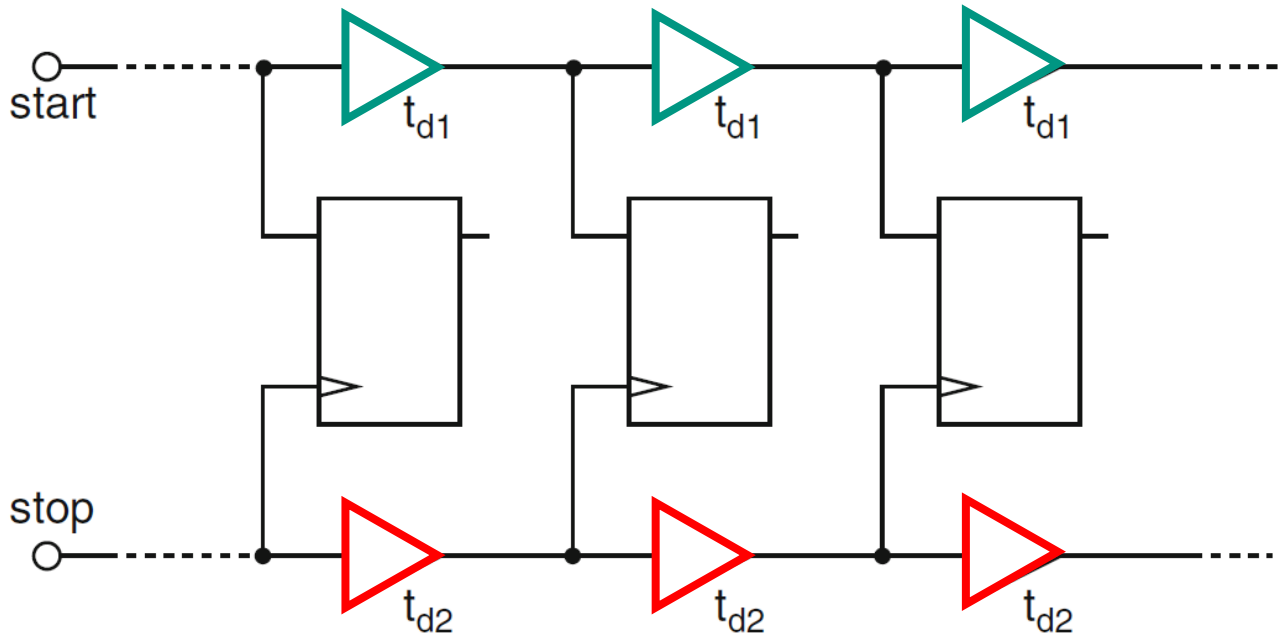
- TDCs with sub-gate delay resolution use sophisticated circuit techniques (resolution enhancement techniques) to circumvent the limitation of the technology delay

One method is Parallel Scaled Delay Elements

Vernier TDC

Advanced Sub-Gate delay Resolution

- A delay-line based TDC that is capable of measuring time intervals with a sub-gate delay resolution is the Vernier TDC

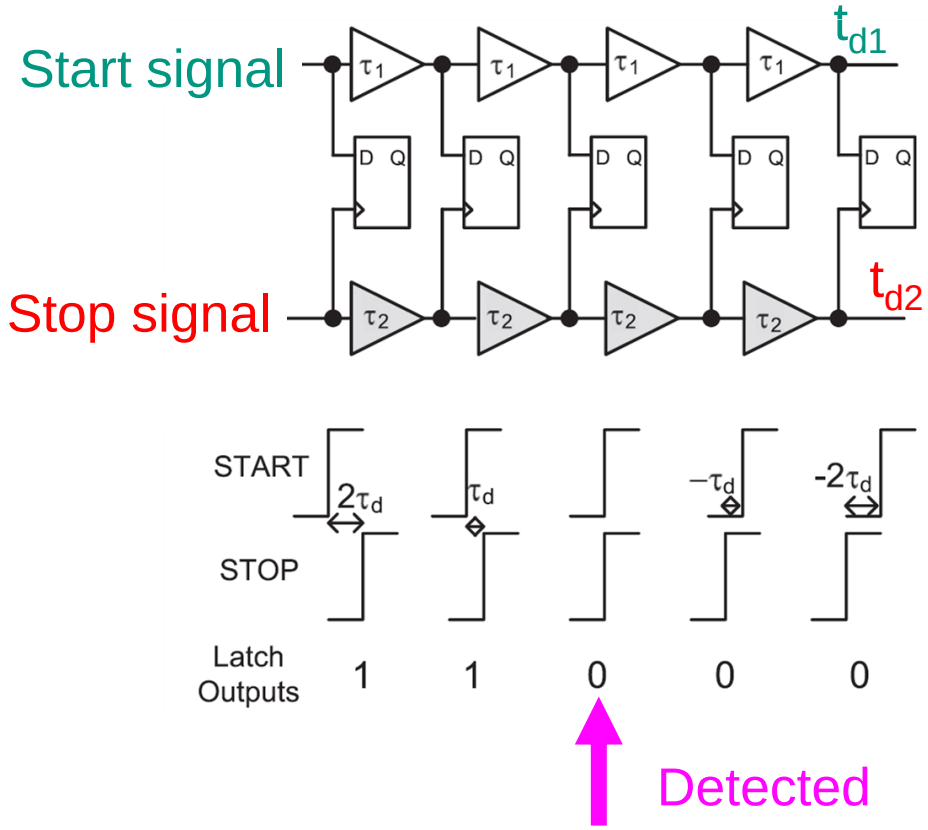


- First delay-line
- $t_{d1} > t_{d2}$
- Second delay-line

- Cut-out of a Vernier TDC is based on a first delay-line for the start signal and a second one for the stop signal. The arrival time at two corresponding nodes is assessed by early–late detectors such as flip-flops

Vernier TDC – working principle

Advanced Sub-Gate delay Resolution



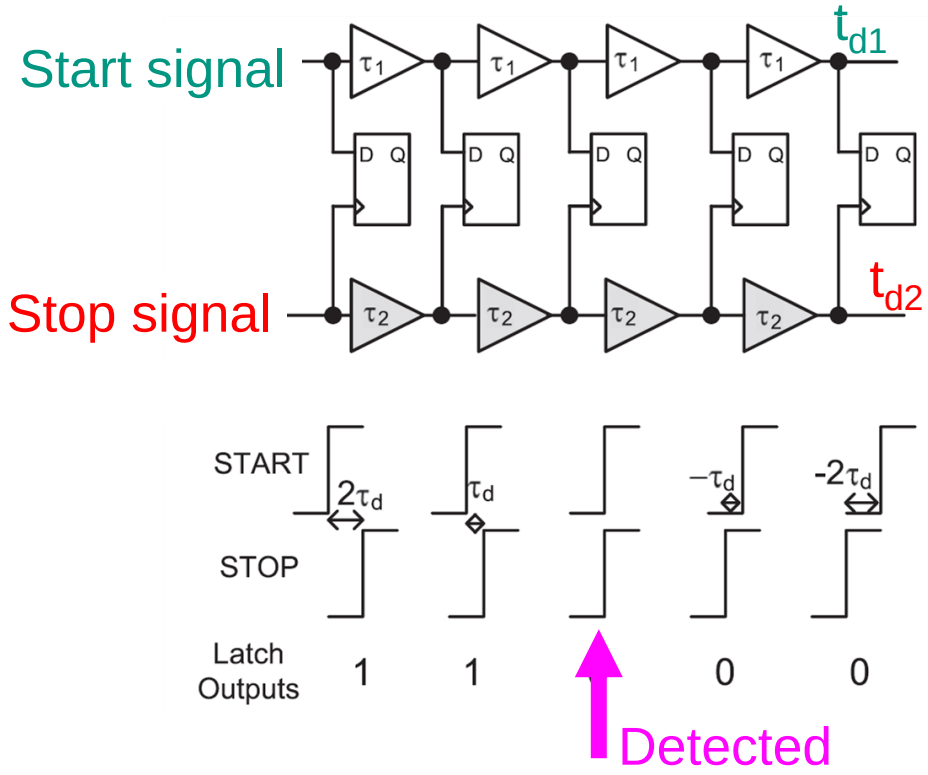
- Thermometer code: 11 0000…

- Step 1: start signal propagates along the first delay-line.
- Step 2: stop signal occurs later, but the delays seen by this signal are smaller
- Step 3: The stop signal chases the start signal
- Step 4: The point (stage) where both signals are in phase is detected by early late detectors (ELDs), usually implemented as flip-flops

- The number of cells necessary for **STOP** signal to surpass the START signal represents the result of TDC conversion

Vernier TDC – working principle

Advanced Sub-Gate delay Resolution



- The granularity in the time domain, i.e. the TDC resolution, is the delay difference $\mathbf{T_{LSB} = t_{d1} - t_{d2}}$ between the elements in the delay-lines, **which can be made arbitrarily small**, therefore **the resolution does not depend on a gate delay**, but on the difference of two gate delays. Consequently, the Vernier TDC provides a circuit technique to overcome the resolution limitations given by a certain technology

- The price for this increased resolution is: the area, the power consumption and the latency. Because the skew between the signals in the first and the second delay line is reduced by T_{LSB} in each stage, a maximum number of N stages required to measure a maximum time interval T_{max} is

$$N = \frac{T_{max}}{T_{LSB}} = \frac{T_{max}}{t_{d1} - t_{d2}}$$

$$A_{core}^{Vernier} = \frac{T_{max}}{T_{LSB}} \left(4A^{inv} + A^{FF}\right)$$

A^{inv} = area of one inverter and
A^{FF} = area of one flip-flop

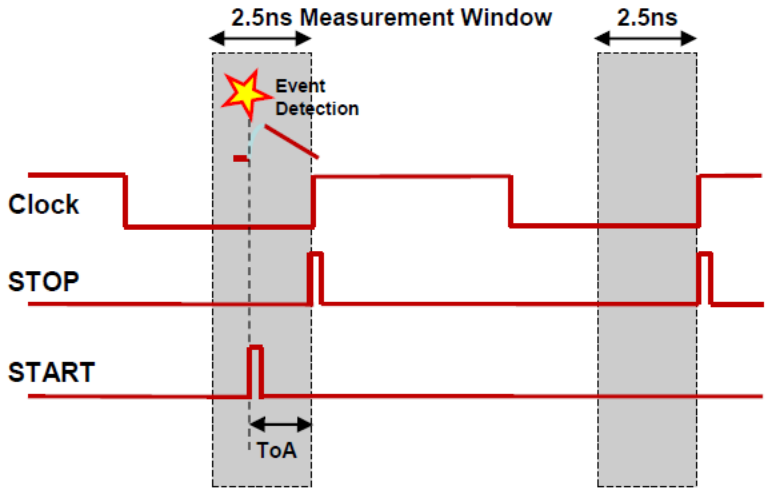
TOA TDC Architecture (Simplified): Vernier Delay Line

TDC Power consumption 0,4 mA *1.2V = 0,5 mW @ 10% occupancy

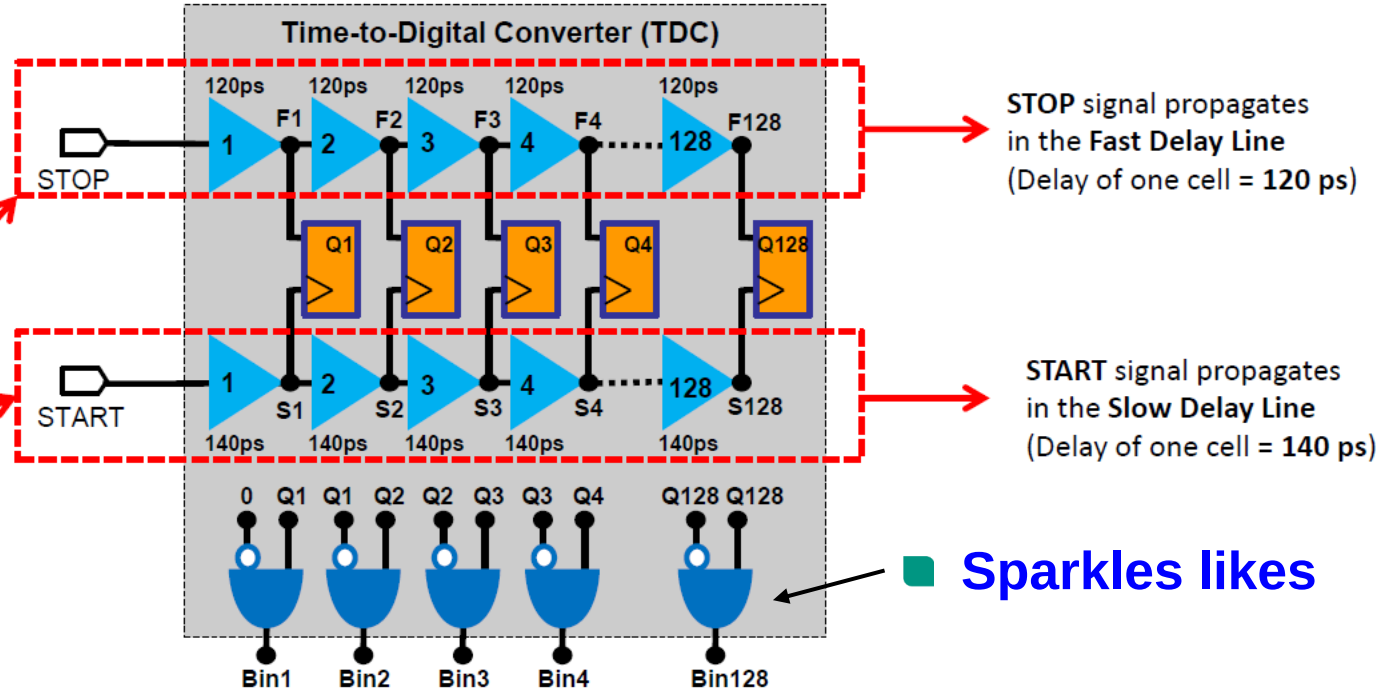
TOA TDC

- **Resolution: 20 ps**
- **Range: 2.5 ns**
- **7 bits**

@ Bojan Markovic, SLAC

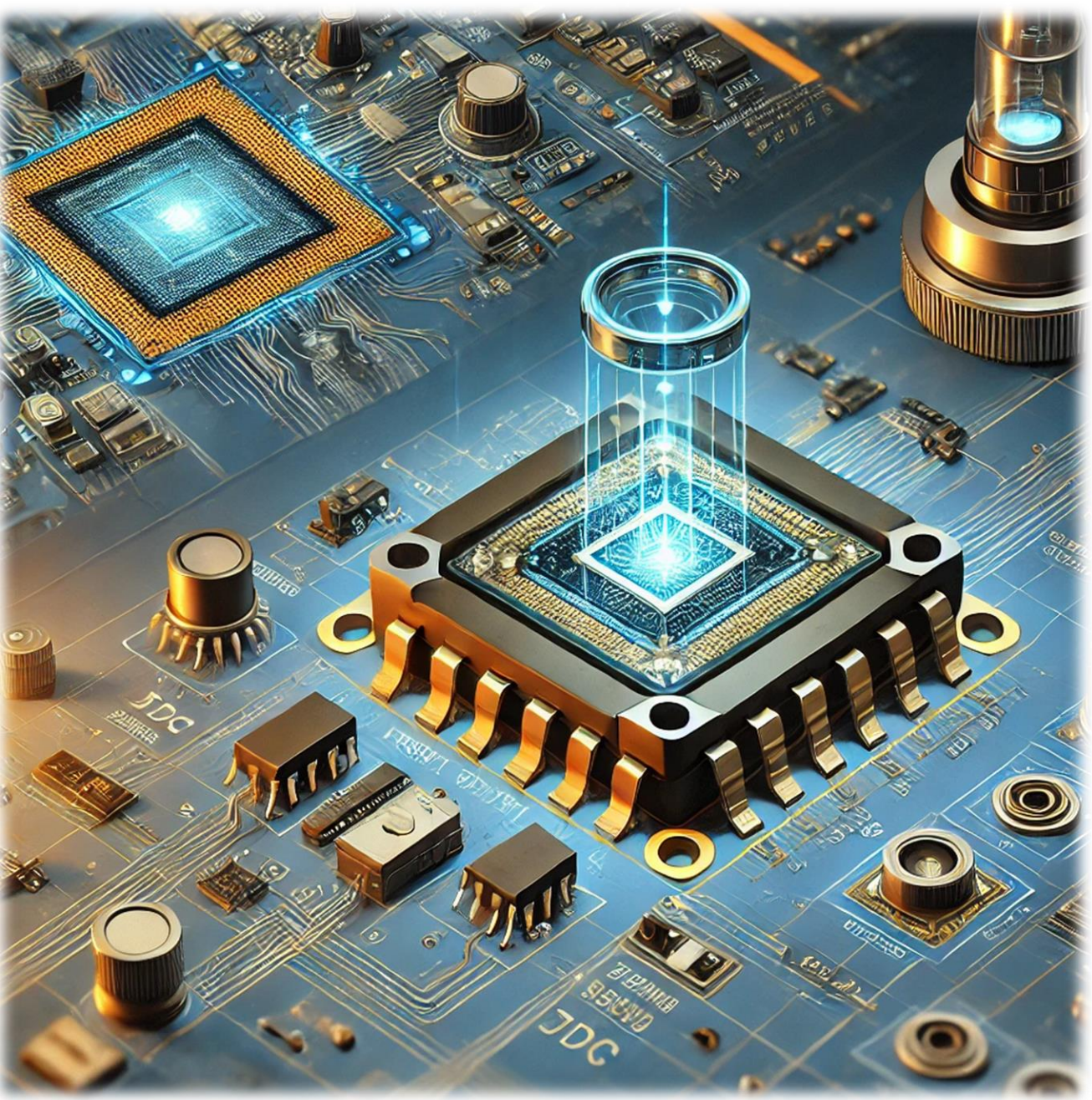


Simplified Block Diagram:



Differential shunt capacitor voltage-controlled delay cells

- The **START** pulse comes first and initializes the TDC operation.
- The **STOP** pulse follows the **START** with a delay that represents the time interval to be digitalized.
- At each tap of the Delay Line the **STOP** signal catches up to the **START** signal by the deference of the propagation delays of cells in Slow and Fast branches of the delay line: i.e. **140ps – 120ps = 20ps** that represents the **LSB** of time measurement.
- The number of cells necessary for **STOP** signal to surpass the **START** signal represents the result of TDC conversion
- Cycling configuration used in order to reduce the total number of Delay Cells.
- TDC range is equal to **128*20ps = 2.56ns**



Karlsruhe Institute of Technology

Thank you so much for your kind attention!

michele.caselle@kit.edu