Sensors, Signals and Noise

COURSE OUTLINE

- Introduction
- Signals and Noise
- Filtering
- Sensors: PD6 – Single-Photon Avalanche Diodes

Single-Photon Counting and Timing with Avalanche Diodes

- Sensitivity limits of APDs in linear amplifying mode
- Limits to Single-Photon Counting with APDs in linear amplifying mode
- Geiger-mode operation of avalanche diodes above the Breakdown Voltage
- Single-Photon Avalanche Diodes SPADs
- Active Quenching Circuit AQC
- SPAD arrays and Silicon PhotoMultipliers (SiPM)
- Integrated systems for photon counting and timing

APDs for Single-Photon Counting (SPC)?

APDs can detect smaller optical pulses than PIN diodes, thanks to the internal gain M.

However, the improvement of sensitivity is much lower than that brought by PMTs with respect to vacuum tube PDs. The reason is that in comparison to PMTs the APD gain M has

1. much lower mean value $\overline{M}$
2. much stronger statistical fluctuations, with relative variance that increases with $\overline{M}$

The **QUESTION** arises:

can we employ linear amplifying APDs instead of PMTs in single photon counting and timing techniques?

And the **ANSWER** is: NO!

More precisely, almost **NO** for silicon APDs and absolutely **NO** for APDs in other materials. In fact, we will now verify that only some special Si-APDs achieve single photon detection, although with marginal performance (detection efficiency lower than APD in analog detection; etc.), and other APD devices are out of the question.

APDs for SPC ?

- The APD output pulses due to a single primary carrier (single-photon pulses) are observed and processed accompanied by the noise of electronic circuitry, arising in the preamplifier and processed by the following circuits.
- A pulse comparator is employed to discriminate SP pulses from noise; pulses higher than the comparator threshold are accepted, lower pulses are discarded.
- The parameters of the set-up (rms noise; pulse amplitude; threshold level) should be ajusted to provide:

    1. Efficient **rejection of noise**, i.e. low probability of false detections due to the noise
    2. Efficient **detection of photon pulses**, i.e. high probability of detecting the SP pulses, which have variable amplitude with ample statistical fluctuations

Noise Rejection in Photon Counting

- With noise amplitude having gaussian distribution (most frequent case) with variance σ_n (rms value), the **noise rejection threshold level must be at least** $\boldsymbol{N_{nr} \geq 2,5\ \sigma_n}$, in order to keep below <1% the probability of false detection
- We have seen that by employing an **optimum filter** for measuring the amplitude of detector pulses we get rms noise (in number of electrons)

$$\sigma_n = \frac{\sqrt{2C_L\sqrt{S_v}\sqrt{S_i}}}{e}$$

e =electron charge and typically:
C_L ≈ 0,1 to 2pF load capacitance;
$\sqrt{S_v}$ ≈ 2 to 5nV $Hz^{-1/2}$ series noise;
$\sqrt{S_i}$ ≈0,01 to 0,1 pA $Hz^{-1/2}$ parallel noise

With high quality APD and preamp we get typically $\sigma_n \approx$ 40 to 120 electrons.
The noise rejection threshold required then is

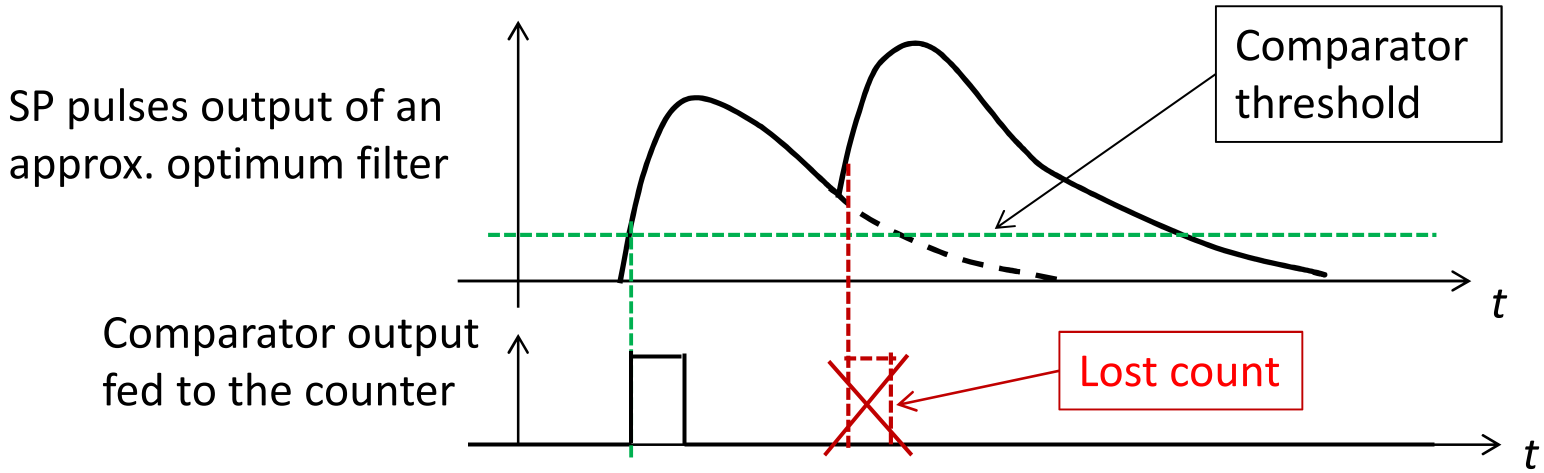
$N_{nr} \geq 2,5\ \sigma_n \approx$ 100 to 300 electrons.

Furthermore, M just higher than N_{nr} is not sufficient for having SP pulses higher than the threshold: we will see that **M much higher than N_{nr} is necessary**.

- We know that the optimum filter (and of course also an approximate optimum) is a low-pass filter and the output pulse has a width (i.e. a reciprocal-bandwidth) of some noise corner time constant T_{nc} . Since in our case T_{nc} ranges from 10ns to a few 100ns, the output pulses are fairly long and this brings drawbacks.

Count losses in Photon Counting

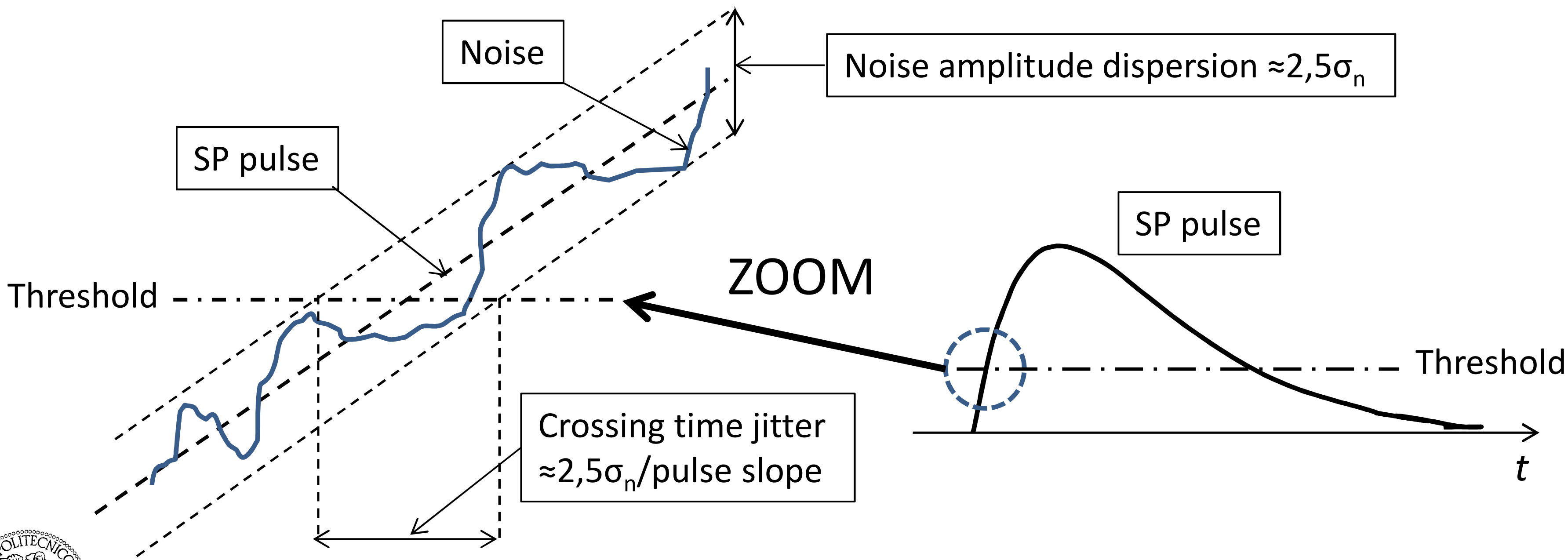
- In photon counting **the finite width of the SP pulse causes count losses**. When the time interval between two photons is shorter than the output pulse width, pulse pile-up occurs (i.e. the two pulses overlap), the comparator is triggered only once and one count is recorded instead of two



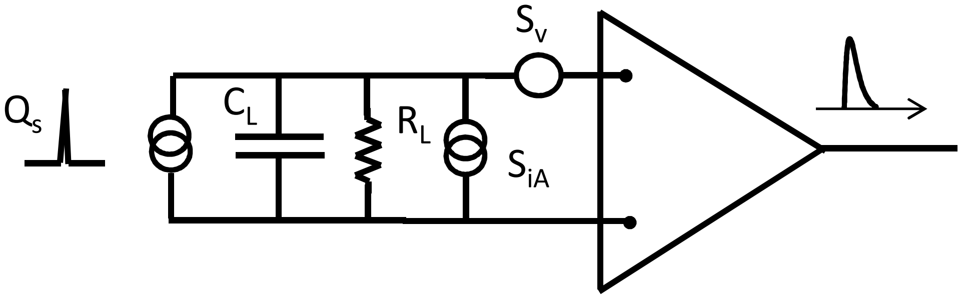
- Photons occur randomly in time, hence the probability of pulse pile-up increases when the pulse width is increased.
- In conclusion, the percentage of lost counts increases as the pulse-width is increased. The width of the SP pulses should be minimized, in order to achieve efficient photon-counting with minimal percentage of lost counts.

Time-jitter in Photon Timing

- In photon timing, the arrival time of the pulse is marked by the crossing time of the threshold of a suitable circuit by the SP pulse.
- The noise causes **time jitter** (statistical dispersion) of the threshold crossing time
- A quantitative analysis is not reported here, but it is evident that the time jitter is proportional to the noise and **inversely proportional to the pulse rise slope**.
- A fairly **long** $\boldsymbol{T_{nc}}$ implies reduced pulse bandwidth and reduced slope of the pulse rise, hence **wide time jitter**.



Photon Counting with wide-band electronics



$C_L \approx 2pF$; $R_L \approx 1k\Omega$

Amplifier band limited by a pole with $T_A=R_LC_L$

$\sqrt{S_{iR}} = \sqrt{4kT/R_L} \approx 4\ pA\ Hz^{-1/2}$ *(white)*

$\sqrt{S_{iA}} \approx 0,1\ pA\ Hz^{-1/2}$ *(white)*

$\sqrt{S_v} \approx 2nV\ Hz^{-1/2}$ *(white)*

- For reducing count-losses and time jitter, we must process the APD pulses with filter bandwidth wider than the optimum filter. However, this implies higher noise, hence higher threshold level and higher gain required to the APD.
- Let's consider a typical wide-band amplifier configuration, with dominant noise due to a low-resistance load $R_L \approx 1k\Omega$. We have a low-pass filtering with two poles (load circuit and amplifier) with equal time constant $T_A=T_L=R_LC_L$.
- With δ-like SP detector pulse of charge Q_s , the SP output pulse is

$$v_S = \frac{Q_s}{C_L}\frac{t}{T_L}e^{-\frac{t}{T_L}} \qquad \text{with maximum} \qquad V_S = \frac{Q_s}{C_L}\frac{1}{e}$$

Photon Counting with wide-band electronics

The output noise is
$$\overline{v_n^2} \approx S_{iR} R_L^2 \frac{1}{8T_L} = 4kTR_L \frac{1}{8T_L}$$

and the S/N ratio is
$$\frac{S}{N} = \frac{V_s}{\sqrt{\overline{v_n^2}}} = \frac{Q_s}{C_L}\frac{1}{e}\frac{\sqrt{8T_L}}{R_L\sqrt{S_{iR}}} = \frac{Q_s}{C_L}\frac{1}{e}\frac{\sqrt{8T_L}}{\sqrt{4kTR_L}}$$

The rms noise referred to the detector output is in terms of charge is

$$\sqrt{\overline{q_n^2}} = \frac{e}{\sqrt{8}}\sqrt{S_{iR}}\sqrt{T_L} \approx 1,7 \cdot 10^{-16}\ C$$

and in electron number

$$\sigma_n = \frac{\sqrt{\overline{q_n^2}}}{q_{el}} \approx 1055\ electrons$$

With this wide-band electronics, the necessary **noise-rejection threshold level thus is**

$N_{nr} \approx 2,5\ \sigma_n \approx 2600$ electrons.

Furthermore, M just higher than N_{nr} is not sufficient for having SP pulses higher than the threshold: we will see that **M much higher than N_{nr} is necessary**.

Efficiency in the detection of SP pulses

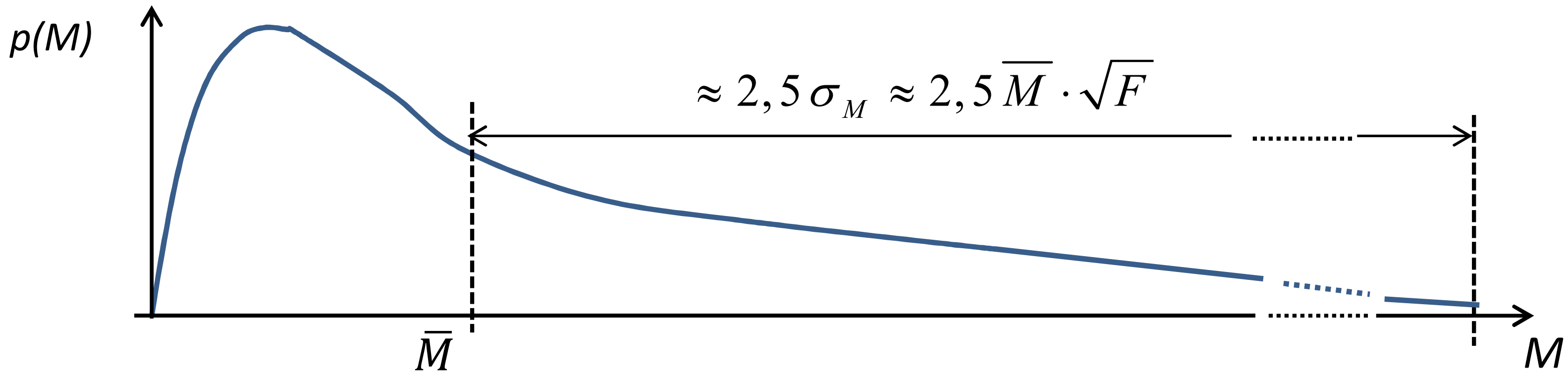
- If the APD gain M were constant for all SP pulses, it would be sufficient to have M just higher than the noise rejection threshold level N_{nr} , but this is not the case.
- The gain M has strong statistical fluctuations, hence a high excess noise factor F>>1, which is directly related to the relative variance of M

$$F = 1 + v_M^2 = 1 + \sigma_M^2 \Big/ \left(\overline{M}\right)^2$$

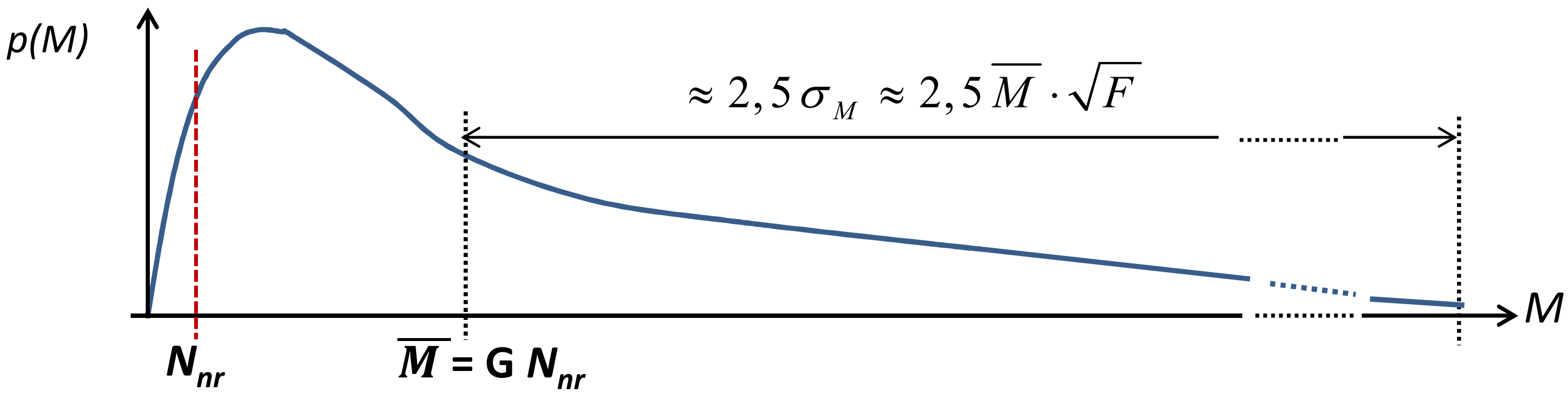
- The statistical M distribution thus has variance σ_M remarkably greater than the mean value $\overline{M}$

$$\sigma_M = \overline{M}\sqrt{F-1} \approx \overline{M} \cdot \sqrt{F}$$

- This implies that M has a strongly asymmetrical statistical distribution, with most of its area below the mean value $\overline{M}$ and a long "tail" above it



Efficiency in the detection of SP pulses

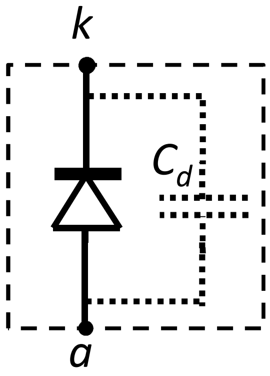


- Therefore, with a mean gain $\overline{M}$ just above the noise rejection threshold a major percentage of the SP pulses is rejected. This downgrades the photon detection efficiency, i.e. the basic performance of the detector.
- In order to limit the reduction of detection efficiency due to the threshold, the mean gain $\overline{M}$ should be higher than the noise rejection threshold N_{nr} by a factor G>>1
- In the most favorable case (special Si-APD with optimum filtering), the value of $\overline{\boldsymbol{M}}$ necessary for attaining the noise rejection threshold N_{nr} is near to the maximum available APD gain, but there is still some margin. In other cases (regular Si-APDs with wideband electronics) there is no margin at all.
- **CONCLUSION:** photon counting with linear amplifying APDs is possible only with special Si-APDs and with photon detection efficiency strongly reduced with respect to that obtained with the same APDs by measuring the analog current signal.

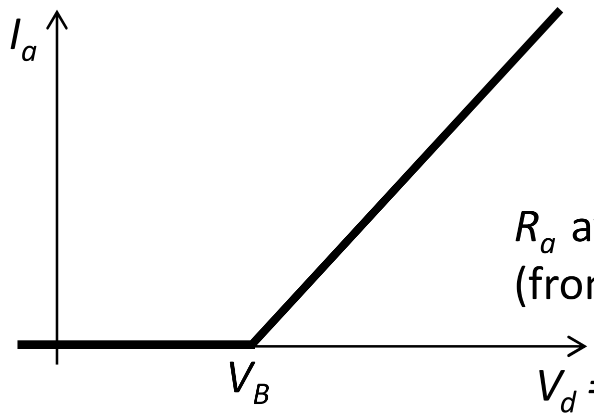
Avalanche diodes above V_B

- We have seen that the positive feedback inherent in the avalanche multiplication of carriers causes strong limitations to the internal gain of APDs in linear operation mode, thus ruling out the possibility of employing them instead of PMTs in single photon counting and timing.
- However, the positive feedback makes possible a radically different operation mode of some avalanche diodes, which working in this mode at voltage **above** the Breakdown Voltage V_B, turn out to be valid single-photon detectors.

Avalanche Diode



Reverse bias I-V characteristics

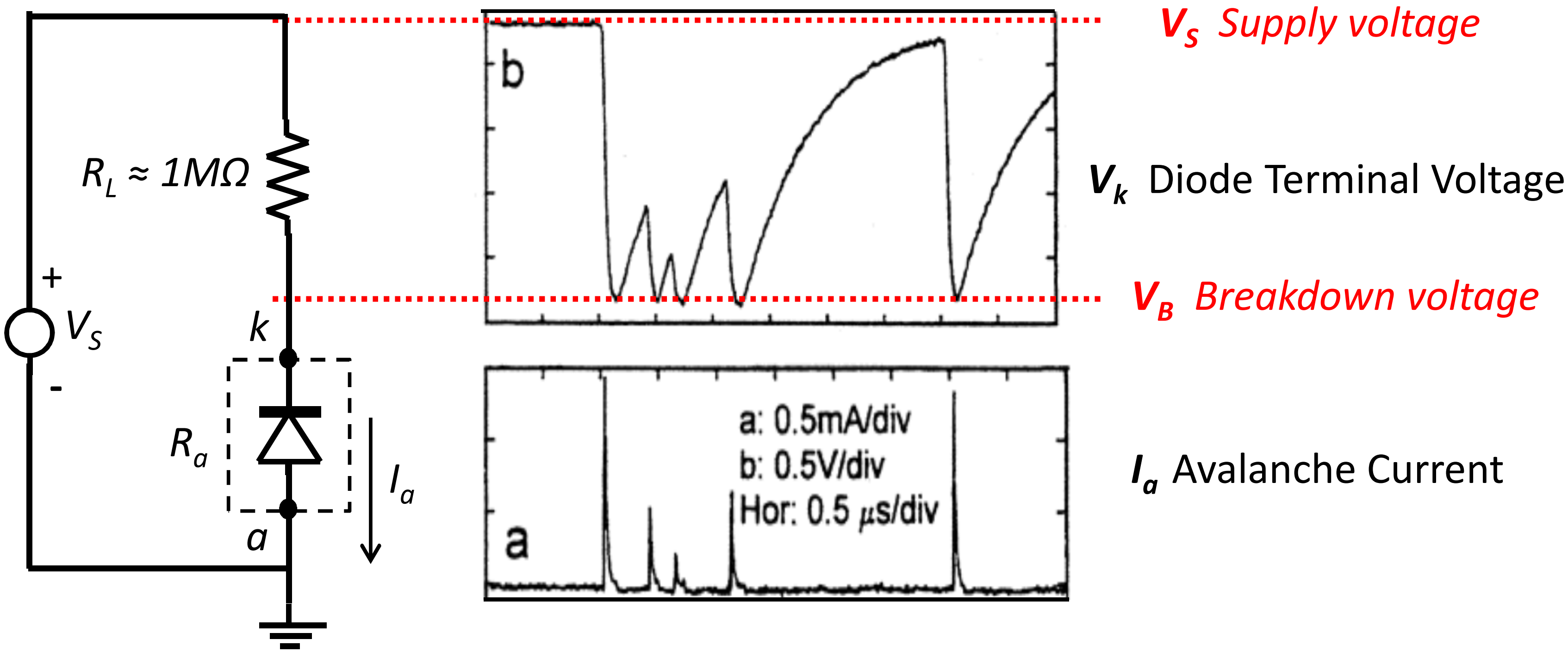


$$R_a = \frac{dV}{dI_a}$$

R_a avalanche diode resistance (from ≈100 Ω to some kΩ)

V_B Breakdown Voltage

Diode biased at $V_s > V_B$ with high load R_L

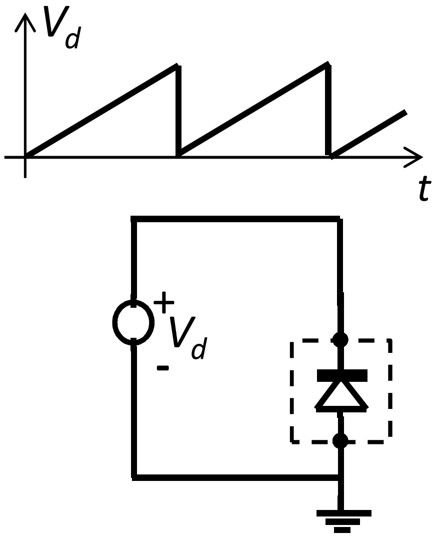


In tests of avalanche diodes the power dissipation can be limited by a **high** load R_L , which limits the current to $I_a \approx (V_S - V_B)/R_L$. **Some diode samples**, however, instead of this steady avalanche current show high-amplitude random pulses:

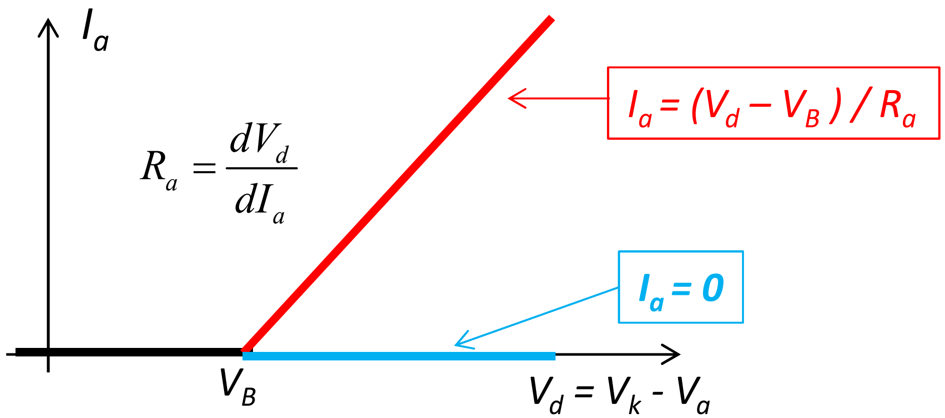
- Fast falls of $\boldsymbol{V_k}$ down to $\boldsymbol{V_B}$, followed by slow exponential recovery towards $\boldsymbol{V_S}$
- Fast current pulses with peak proportional to the amplitude of the voltage fall

With illuminated junction, the repetition rate of pulses increases with the light intensity

I-V characteristics above V_B



Reverse bias I-V characteristics
traced with repetitive **FAST VOLTAGE SCANNING**



The I-V characteristics is currently acquired with a «curve tracer» that applies to the device a repetitive fast voltage scan (scan time typically 10ms). For a diode with the pulsed behavior described, a **bistable** behavior is observed above breakdown $V_d > V_B$:

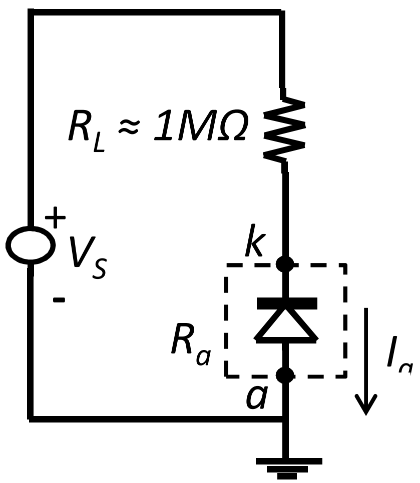
a) in some scans a self-sustaining full avalanche current flows: $I_a = (V_d - V_B)/R_a$

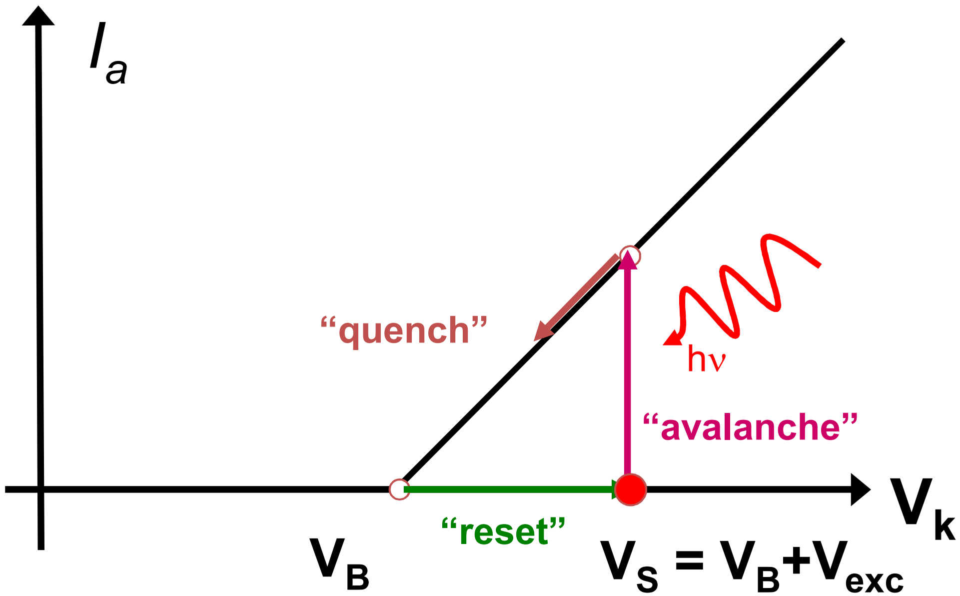
b) in other scans the current is nil : $\boldsymbol{I_a = 0}$

We know that at $V_d > V_B$ a self-sustaining avalanche can be **started even by a single free carrier** entering in the high field region: the **I-V branch with I_a=0 above V_B** thus shows that in some scans this does NOT occur.

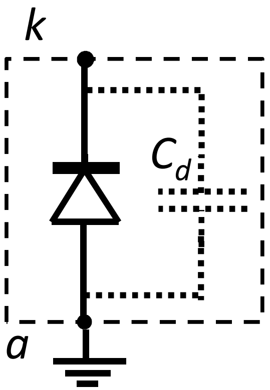
Geiger mode operation

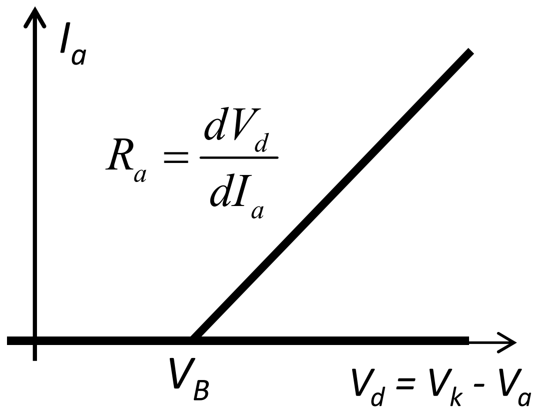
- Bias voltage $\mathbf{V_S}$ ***ABOVE*** breakdown $\mathbf{V_B}$ (with excess bias $\mathbf{V_{exc}}$):
  no current flows in quiescent state
- Single photon switches on avalanche → macroscopic current flows
- It's a triggered-mode avalanche: detector with "***BISTABLE inside***"
- Avalanche quenched by pulling down diode voltage $V_d \approx V_B$
- diode voltage V_d then reset to $\mathbf{V_S}$

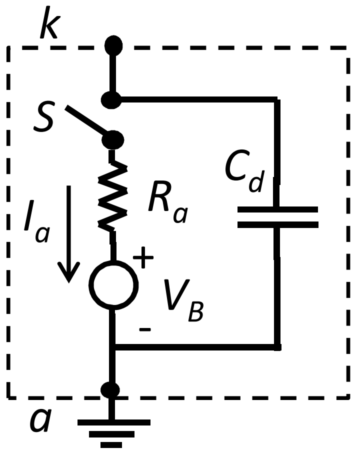




Equivalent Circuit of Diode above Breakdown





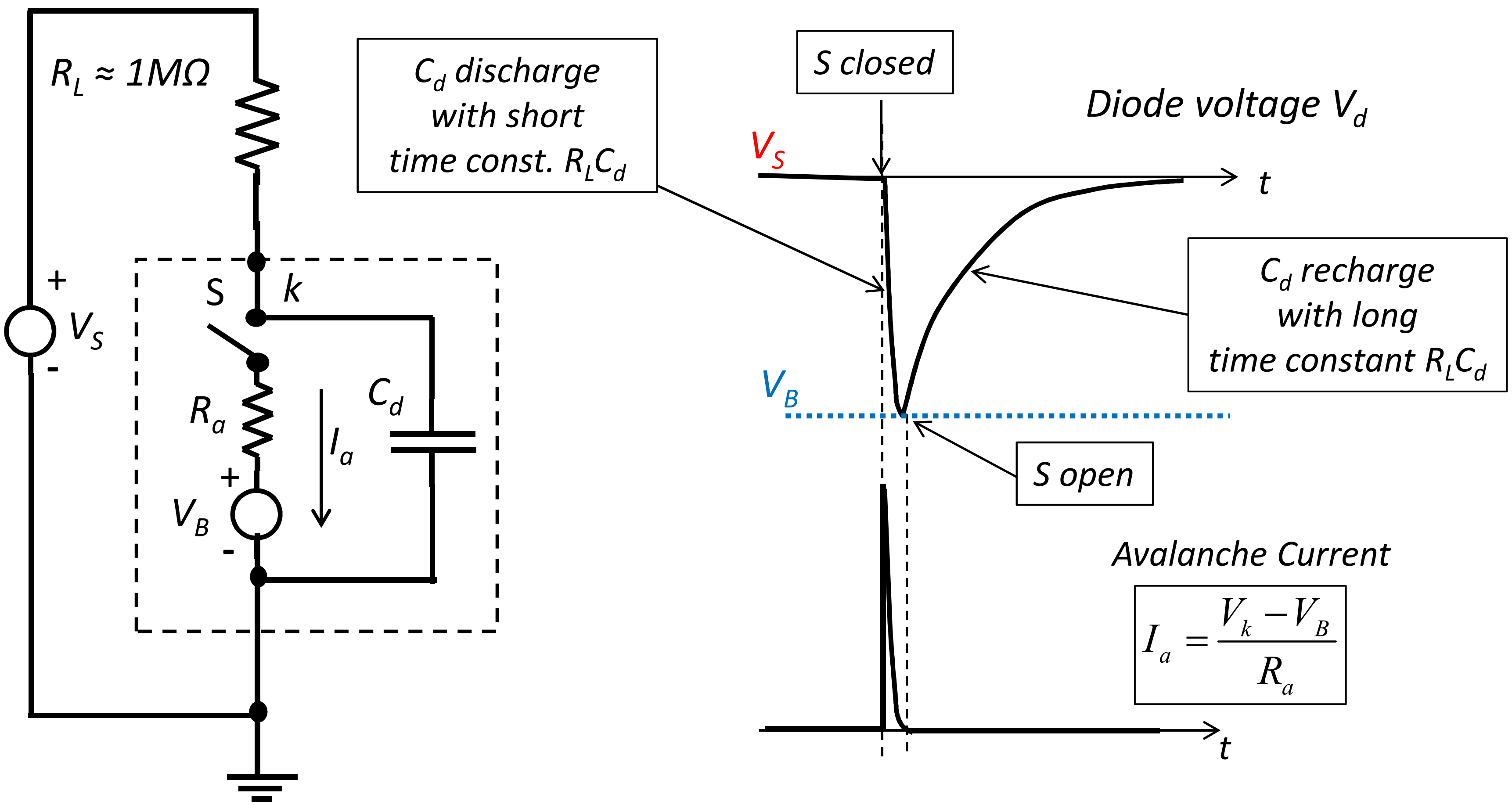


Equivalent Circuit

The equivalent circuit of the diode provides a quantitative understanding of the diode operation and confirms that the pulses observed correspond to single carriers generated in the device, spontaneously or by the absorption of single photons

- at $V_d > V_B$ the switch S can be closed or open; when it is closed, the avalanche current flows. At $V_d \leq V_B$ it is always open.
- **Closing the switch** is the equivalent of **triggering the avalanche** in the diode. Therefore, S is closed when a carrier injected or generated in the high field region succeeds in triggering the avalanche
- S then is open when the avalanche current is quenched (i.e. terminated) by the decrease of the diode voltage down to $\boldsymbol{V_d \approx V_B}$

Passive Quenching Circuit

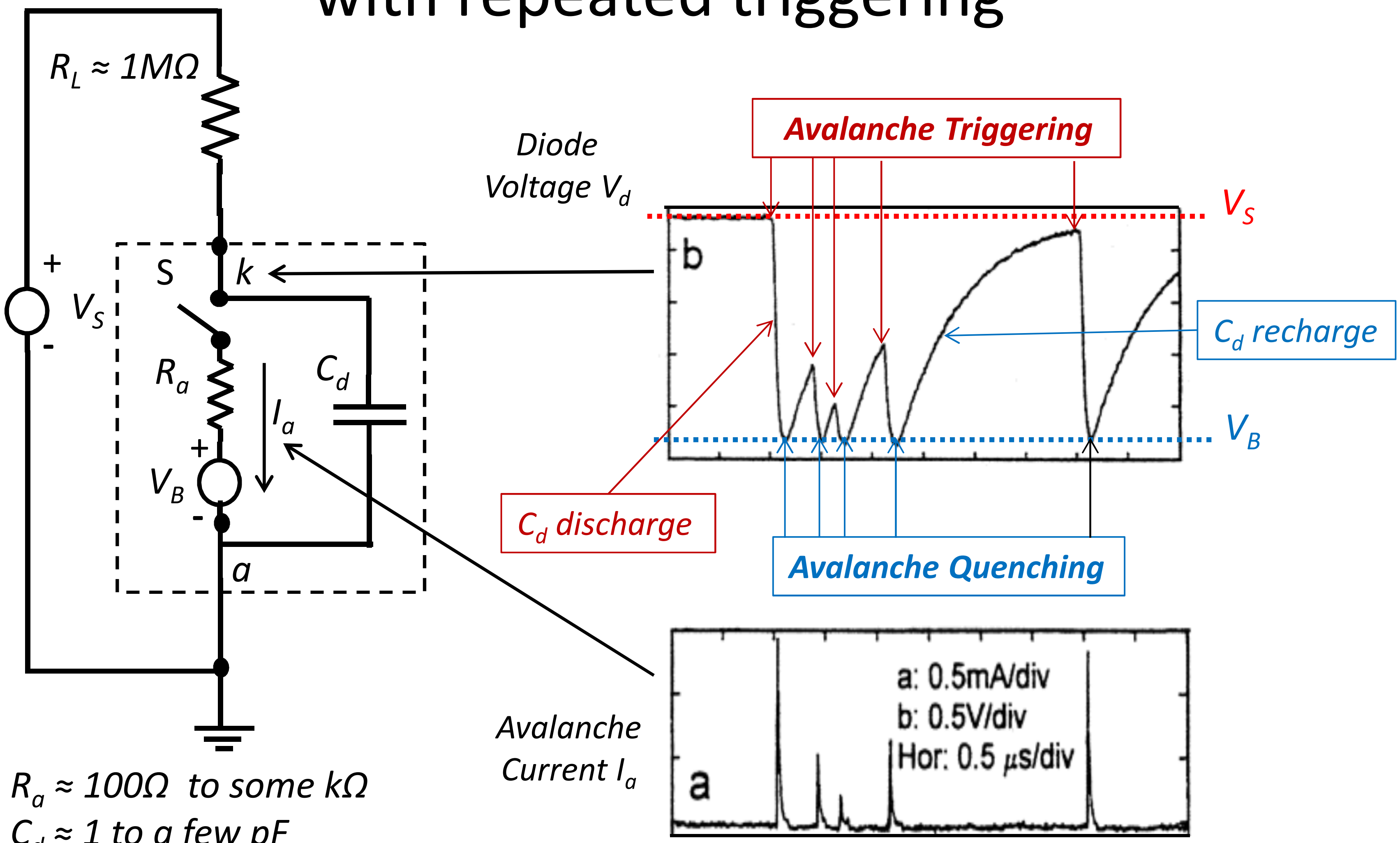


$$I_a = \frac{V_k - V_B}{R_a}$$

$R_a \approx 100\Omega$ to some $k\Omega$
$C_d \approx 1$ to a few pF
$T_a = R_aC_d \approx 100$ps to few ns
$T_L = R_LC_d \approx 1$ to some μs

When the diode voltage goes down to V_B the avalanche is no more self-sustaing. The avalanche is thus quenched by the action of R_L and the circuit is called **Passive Quenching Circuit (PQC)**

Passive Quenching Circuit with repeated triggering



$R_a \approx 100\Omega$ to some $k\Omega$

$C_d \approx 1$ to a few pF

$R_aC_d \approx 100ps$ to few ns

$R_LC_d \approx 1$ to some µs

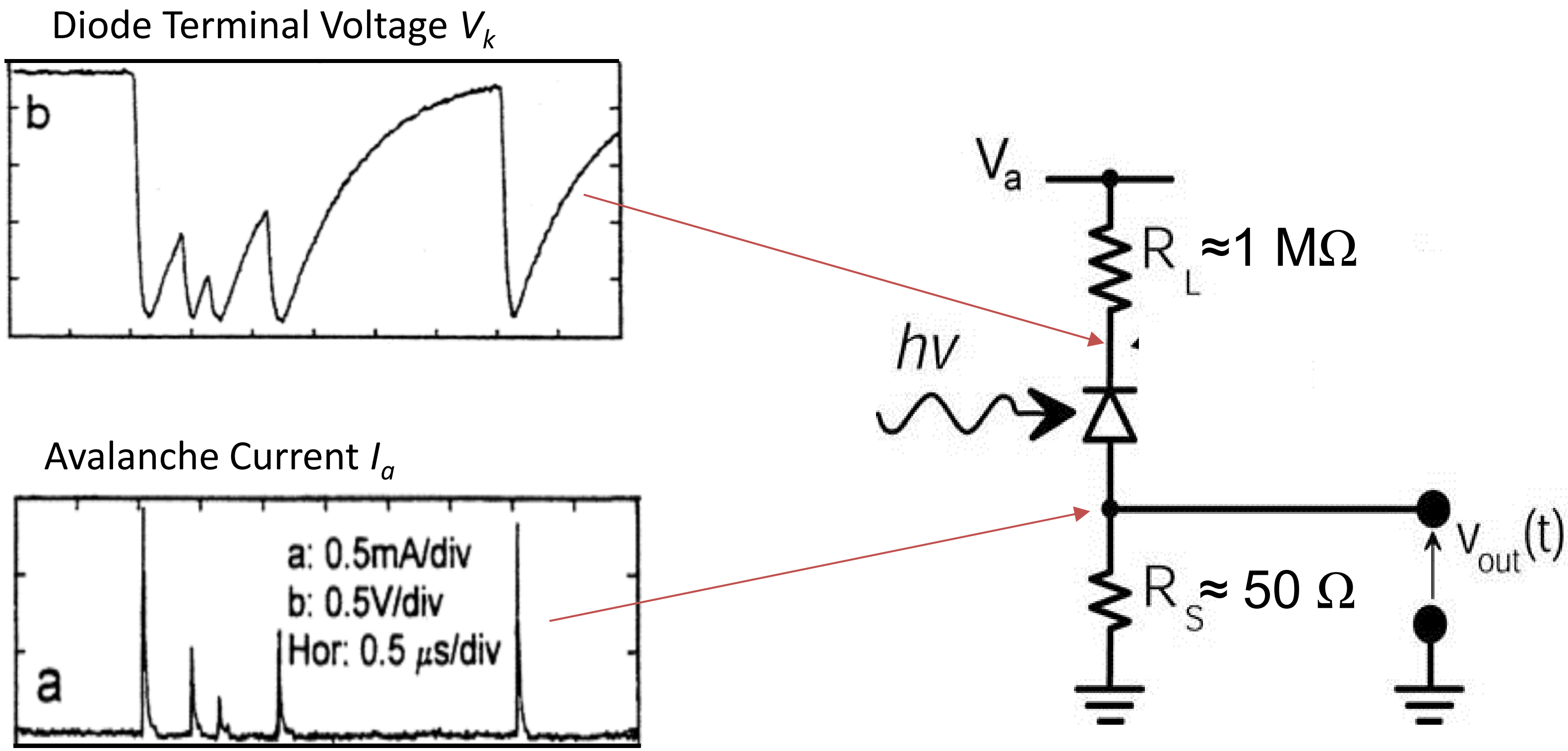
Operation with Passive Quenching

- In order to be able to operate in Geiger mode above the breakdown voltage, a diode should have uniform properties over the sensitive area: in particular, it must be free from defects causing local field concentration and lower breakdown voltage (the so-called microplasmas, due to metal precipitates, higher dopant concentration, etc.)
- Such avalanche diodes, operating above the breakdown voltage in Geiger mode, generate macroscopic pulses of diode voltage and current in response to single photons. They are therefore called **Single-Photon Avalanche Diodes SPAD**.
- Pulses are produced in SPADs also by the spontaneous thermal generation of single carriers in the diode junction and constitute a **dark count rate (DCR)** similar to that observed in PMTs. **Low DCR is a basic requirement** for an avalanche diode to be employed as SPAD.
- Various parameters characterizing the **detector performance strongly depend on the diode voltage**: probability of avalanche triggering, hence the photon detection efficiency; amplitude of the avalanche current pulse; delay and time-jitter of the electrical pulse with respect to the true arrival time of the photon; etc.
- In a passive-quenching circuit, after each quenching the diode voltage slowly recovers from the breakdown voltage V_B to the supply level V_S.

Operation with Passive Quenching

- In photon counting with an avalanche diode in a PQC, count losses are caused by the gradual recovery of the detection efficiency from nil to the correct level after each quenching.
- A correction equation for such losses is not known: it is a case very different from random pulse counting with a constant known deadtime after each event, where the count losses can be accurately corrected by a well known statistical equation
- In photon timing with an avalanche diode in PQC, for photons arriving during a voltage recovery the arrival time measured on the electrical output pulse suffers increased delay and time-jitter with respect to the operation at the correct diode voltage. This effect progressively degrades the time resolution as the pulse counting rate is increased
- In conclusion, the application to photon counting and timing of avalanche diodes in Geiger mode with a PQC has very limited interest. It is restricted to favorable cases with very small probability of occurrence of an event during recovery transients, which can last several microseconds.
- In other words, with avalanche diodes in PQC photon counting and/or photon timing is possible in practice only in simple lucky cases with very low total counting rate; that is, **cases with low dark-count rate, low count-rate of background photons and low count-rate of the signal photons**

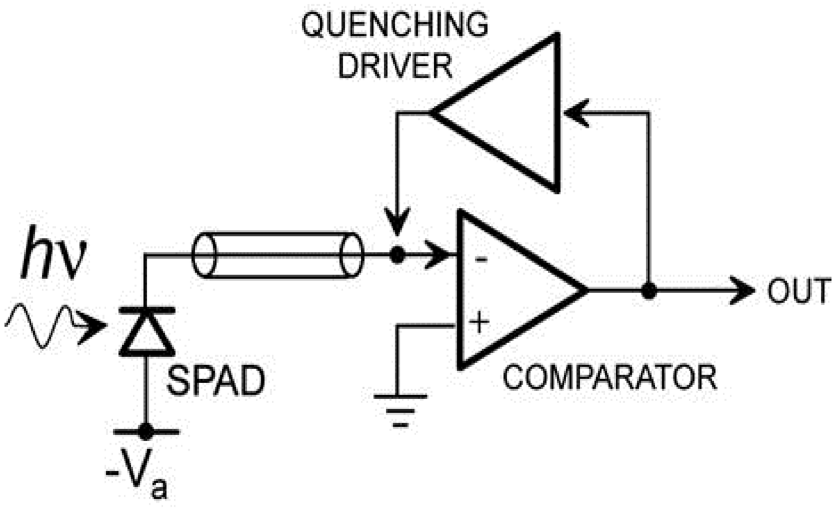
Passive quenching is simple...

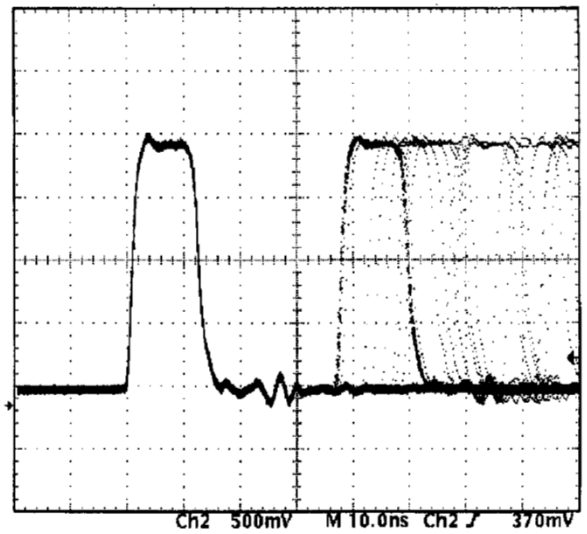


… but suffers from

- long, not well defined deadtime
- low max counting rate < 100kc/s
- photon timing spread
- et al

Principle of Active Quenching Circuits (AQC)





Output Pulses

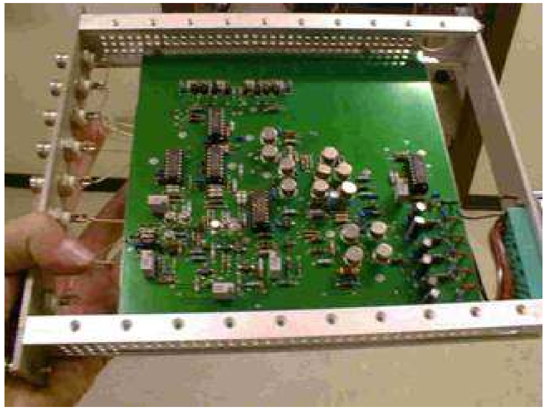
by providing

- short, well-defined deadtime
- high counting rate > 1 Mc/s
- good photon timing
- standard output

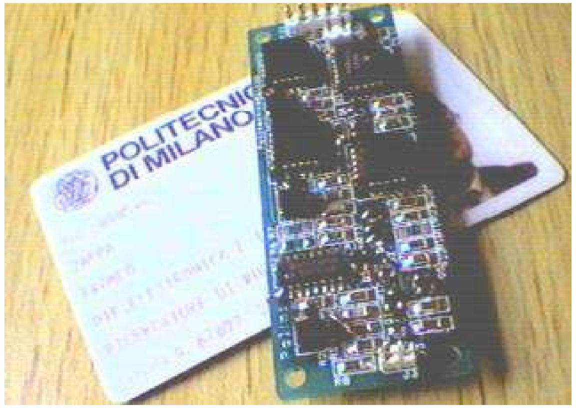
opened the way to SPAD applications

Active Quenching Circuit Evolution

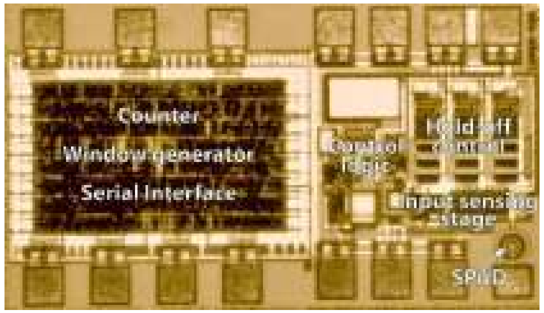
Earlier AQC modules in the 80's



Compact AQC modules in the 90's

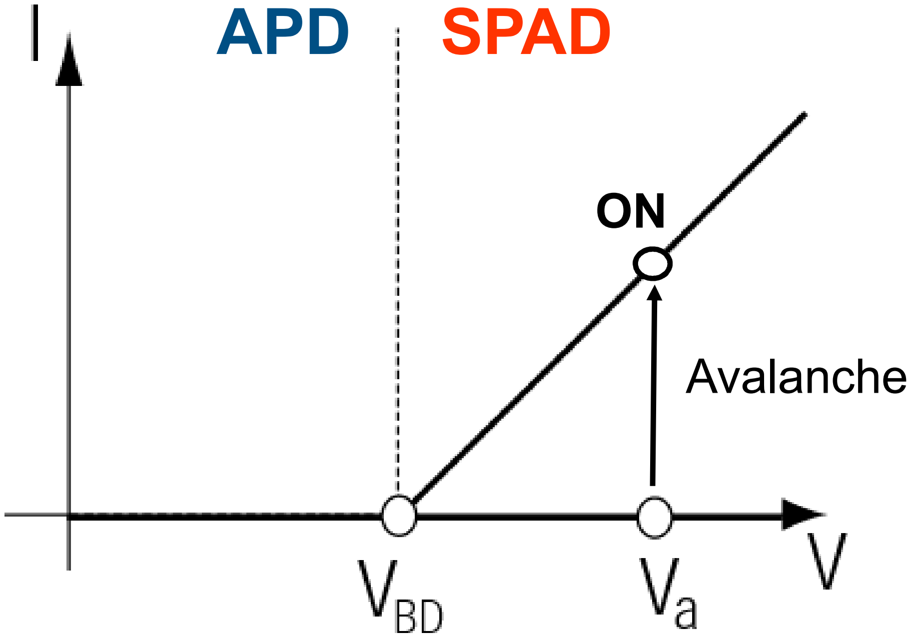


Integrated AQCs in early 2000's



Today: Monolithic chips for Single Photon Counting and Timing

SPADs are different from APDs



Avalanche **P**hoto**D**iode	**S**ingle-**P**hoton **A**valanche **D**iode
• Bias: slightly ***BELOW*** breakdown	• Bias: well ***ABOVE*** breakdown
• Linear-mode: it's an ***AMPLIFIER***	• Geiger-mode: it's a ***BISTABLE !!***
• ***Analogue*** output	• ***Digital*** output
• Gain: ***limited*** < 1000	• Gain: ***meaningless !!***

Why Single Photon Counting

- Direct digital detection
- Overcomes the limit of analog photodetectors, i.e. the circuit noise
- Noise only from the statistics of dark-counts and photons
- Measurement of light intensity with ultra-high sensitivity

and with precise **photon-timing**

Time-Correlated Single Photon Counting (TCSPC)

→ measurement of ultrafast waveforms with ultra-high sensitivity

Single Photon Detectors

Semiconductor SPADs vs. PMTs - Photomultiplier Tubes

- **microelectronic advantages:**

  miniaturized, low voltage, etc.

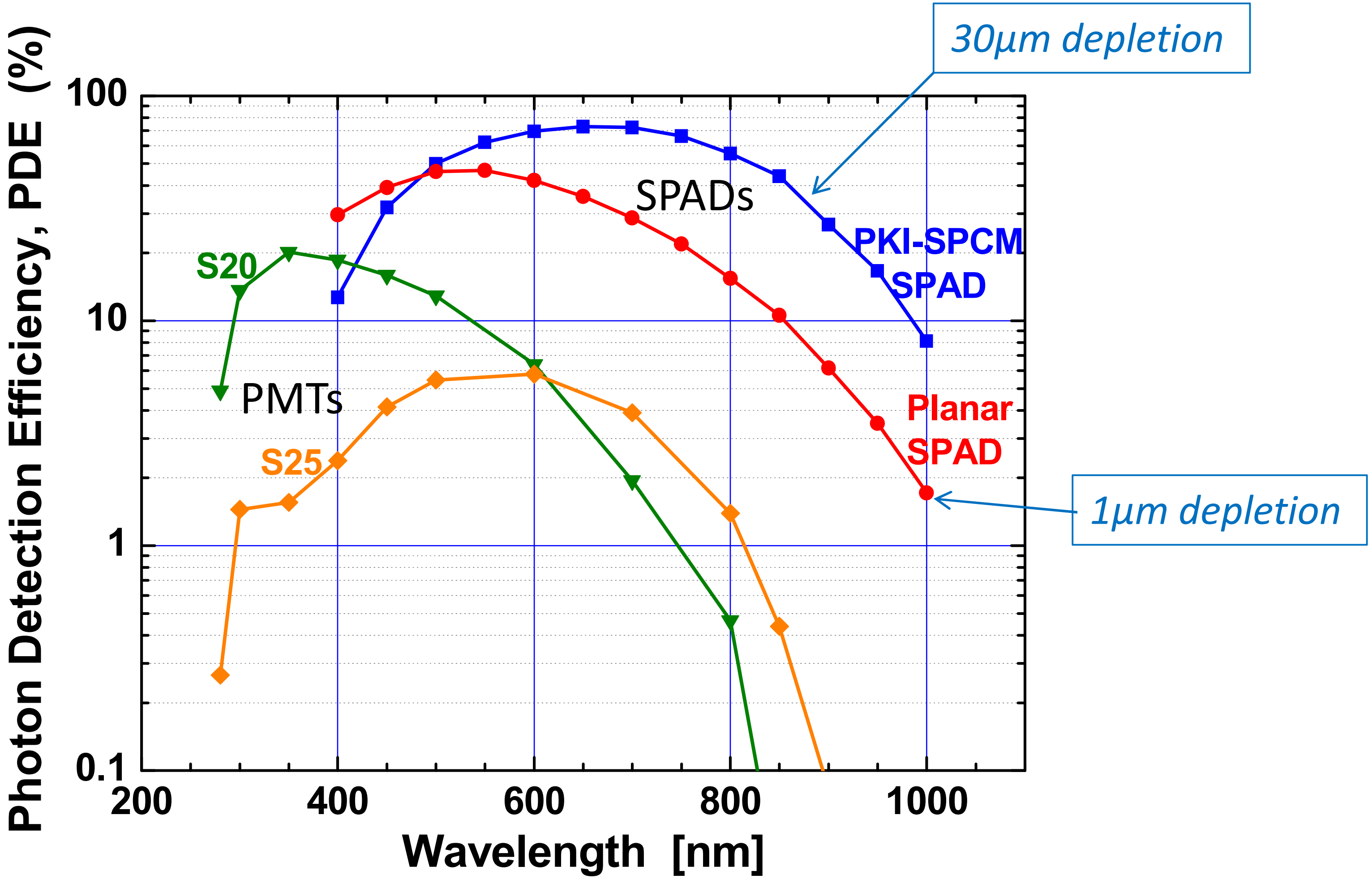
- **improved performance:**

  higher Photon Detection Efficiency

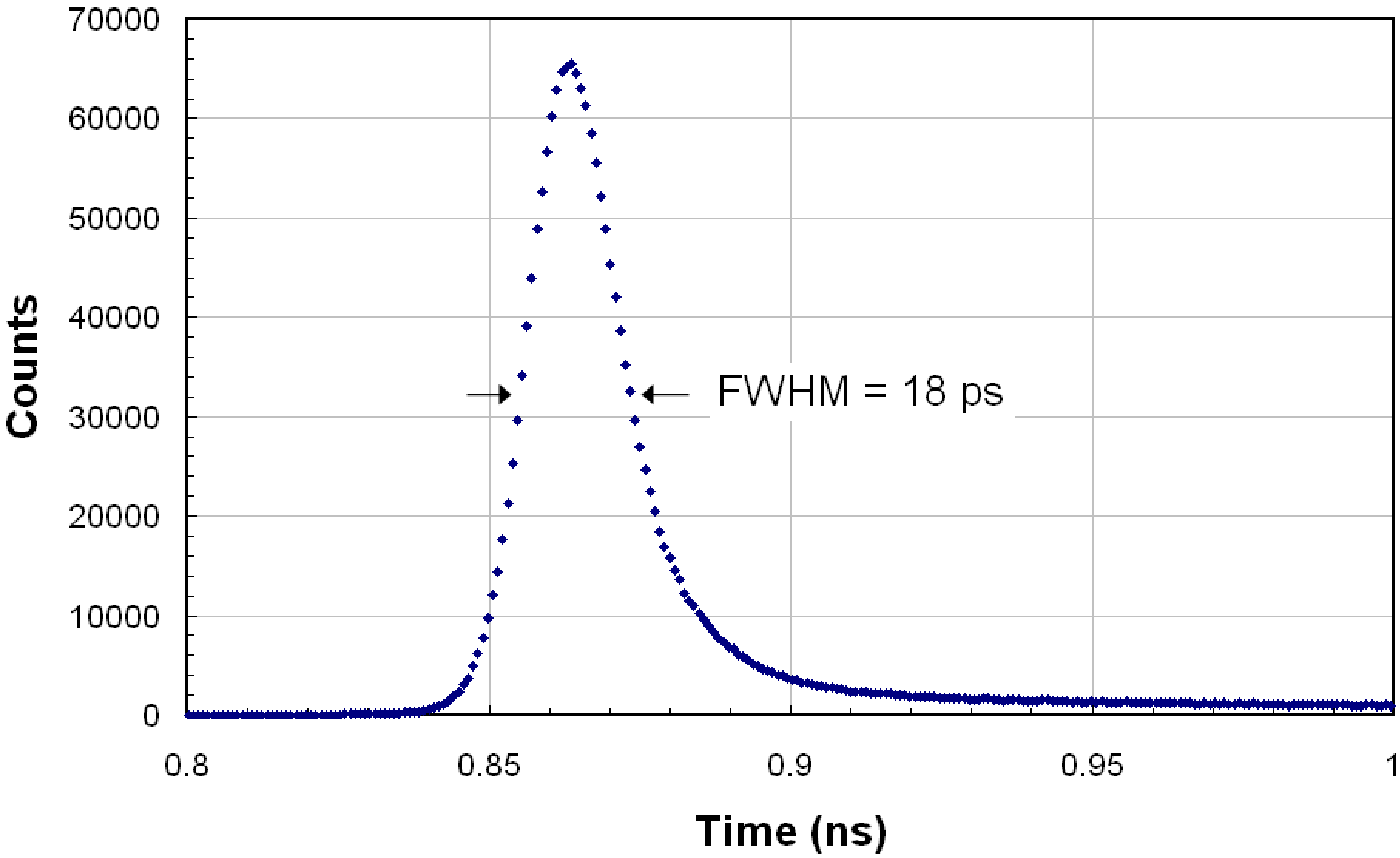
  better photon timing

  comparable or lower noise (dark counting rate)

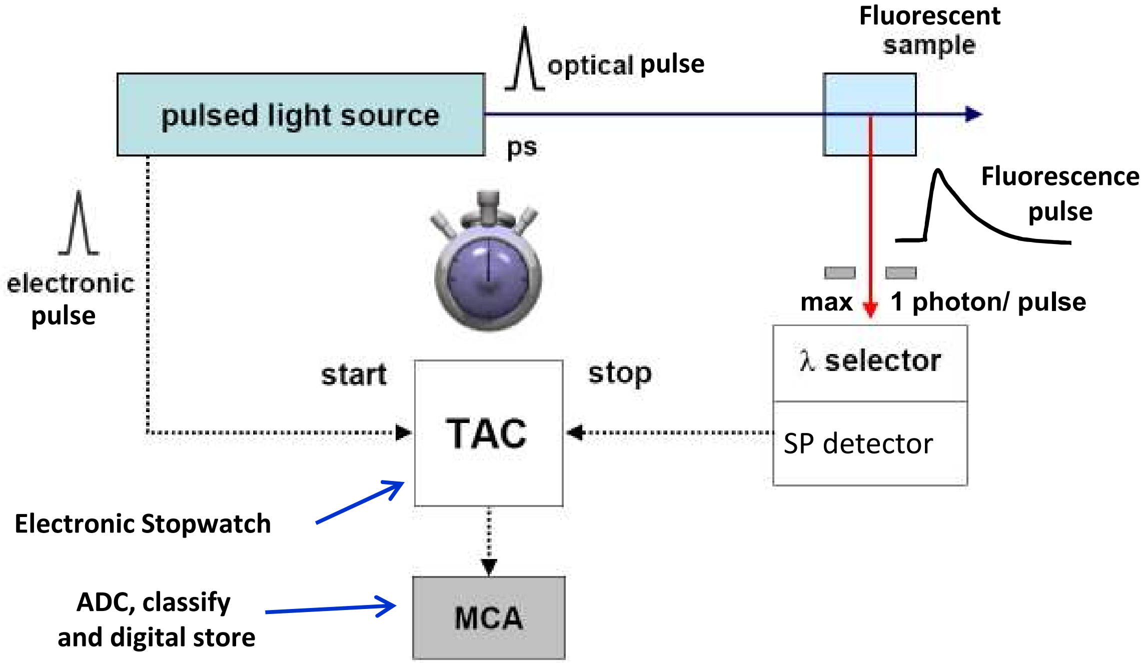
Silicon SPADs vs PMTs: Photon Detection Efficiency

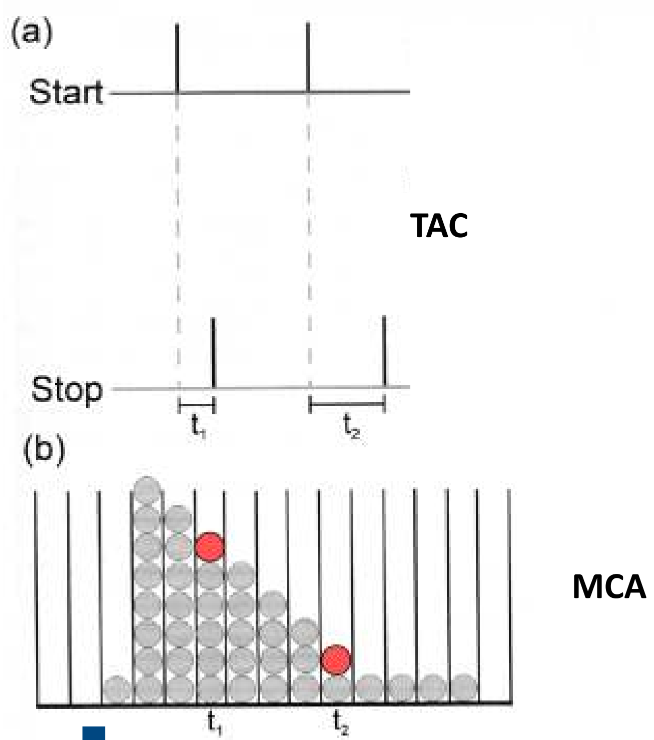


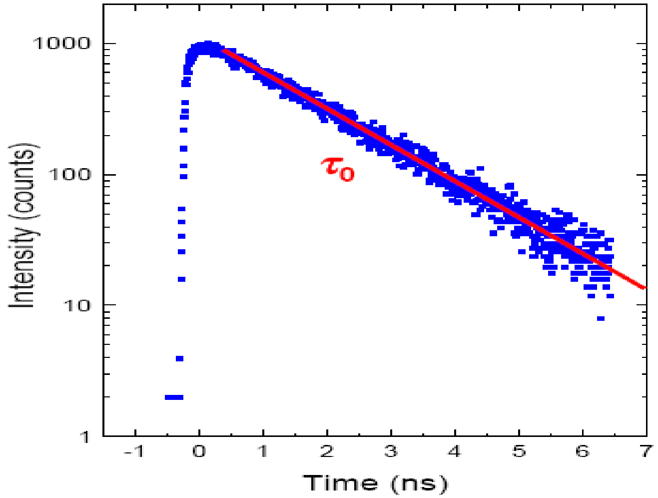
Timing Jitter of Fast Planar SPAD



Time Correlated Single Photon Counting (TCSPC)





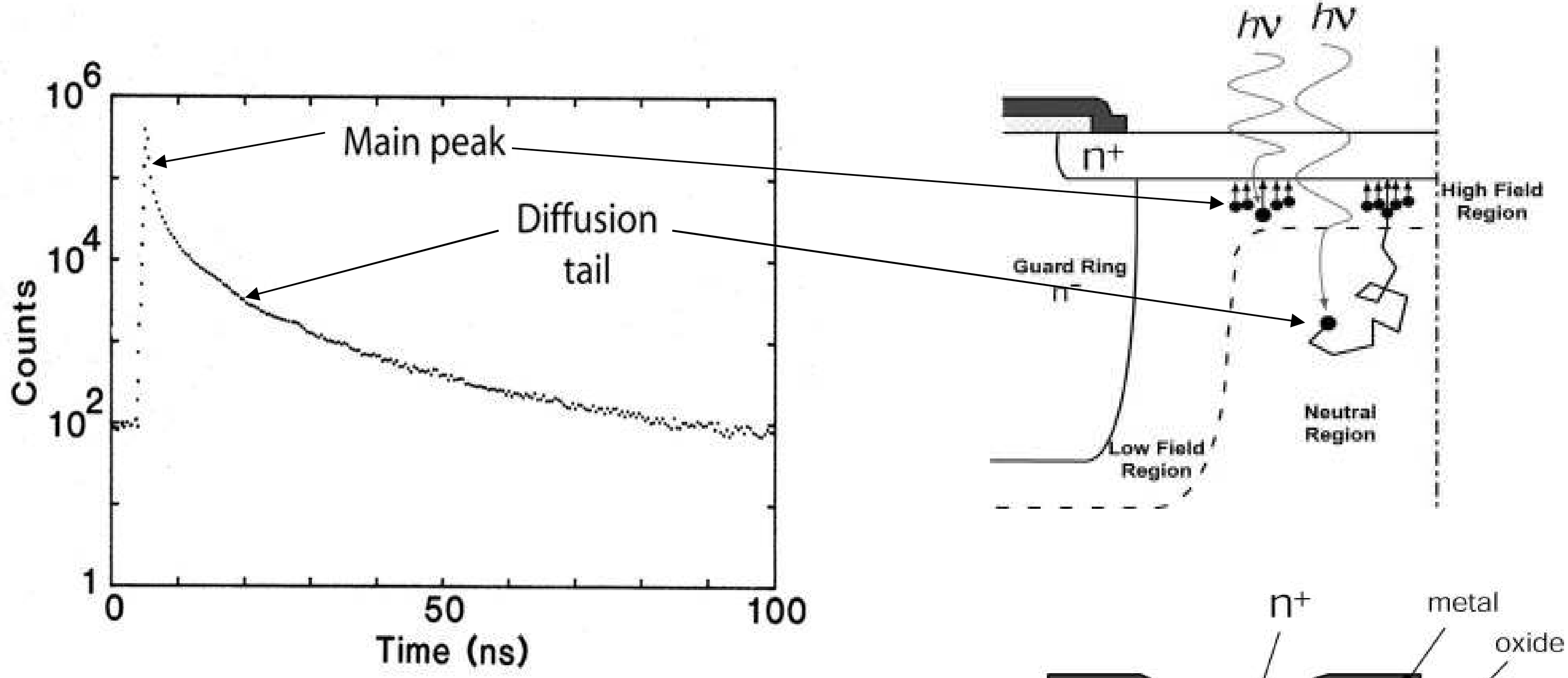


Hystogram of many trials

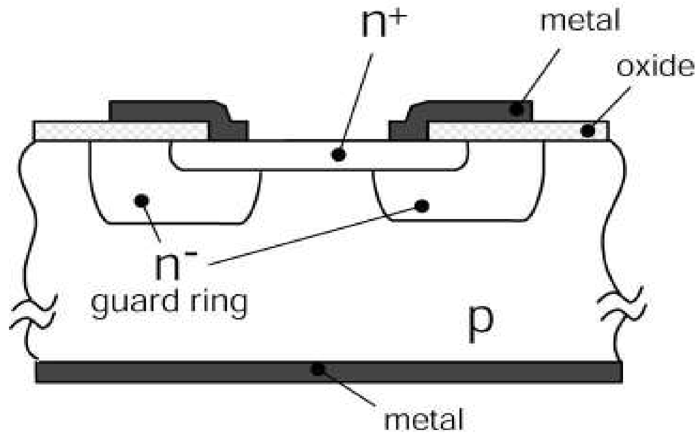
≡

fluorescence decay curve

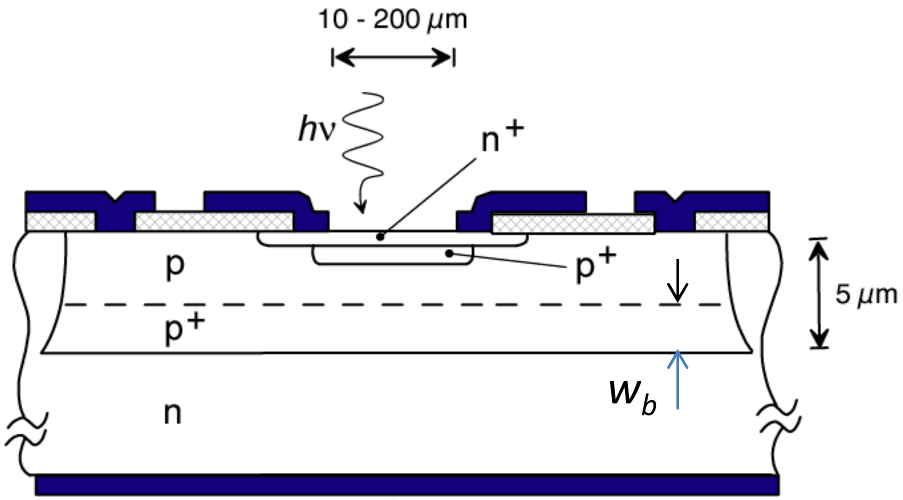
Prototype SPAD structure: diffusion tail

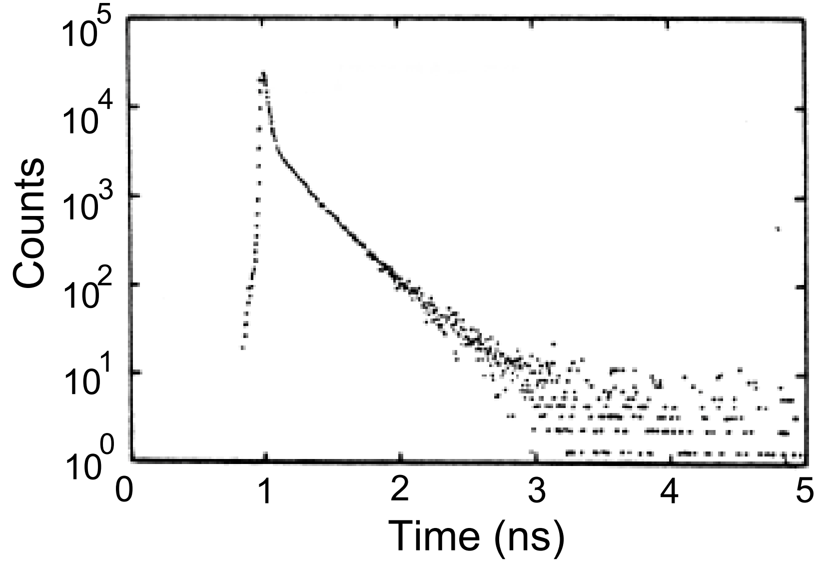


Prototype planar SPAD structure
with deep diffused guard ring
on bulk p-substrate (no epitaxy)



p-p$^+$-n Double-Epitaxial SPAD structure



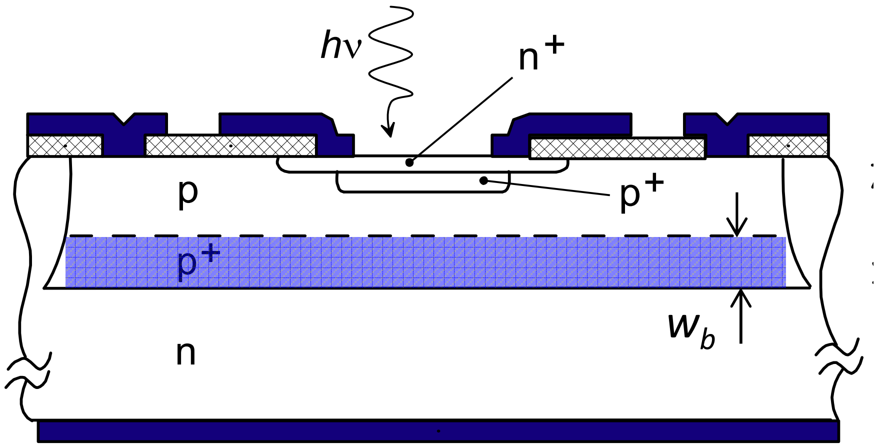


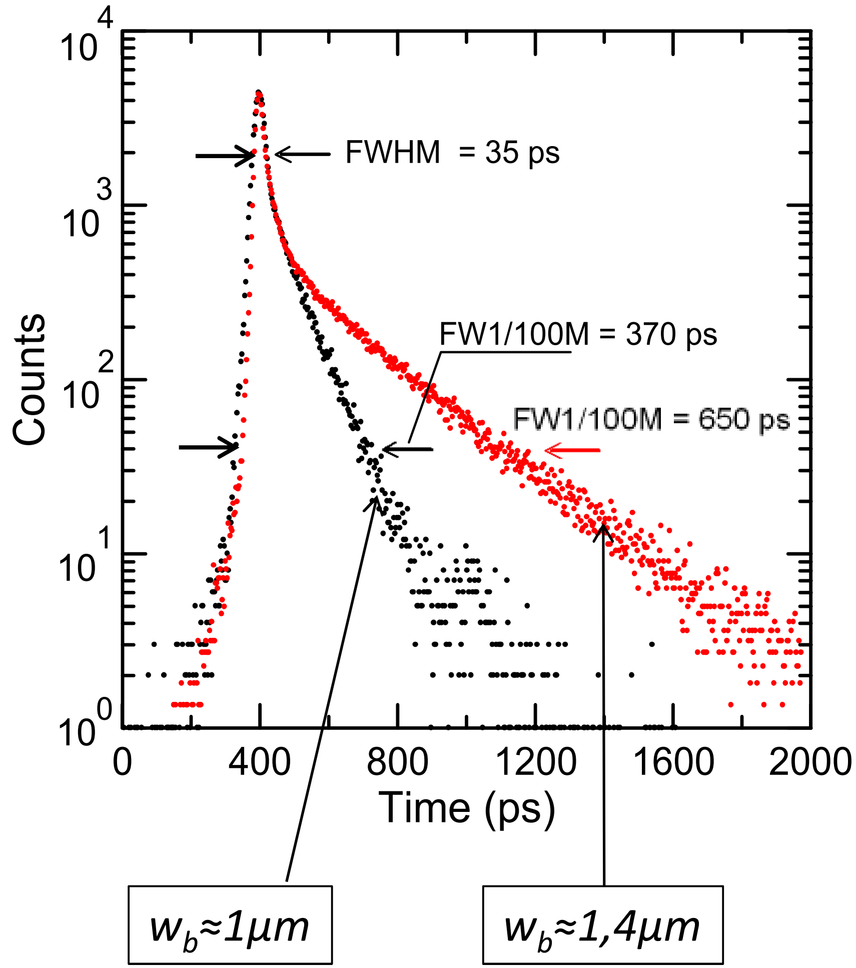
- Short diffusion tail with clean exponential shape
- Active area defined by p+ implantation
- No guard-ring (uniform QE)
- Adjustable V_{BD} and E-field
- Isolated diode structure **SUITABLE** for integration in monolithic systems (array detectors etc.)

w_b neutral p-layer thickness
τ diffusion tail lifetime

$$\tau = \frac{w_b^2}{\pi^2 D_n}$$

Custom SPAD technology

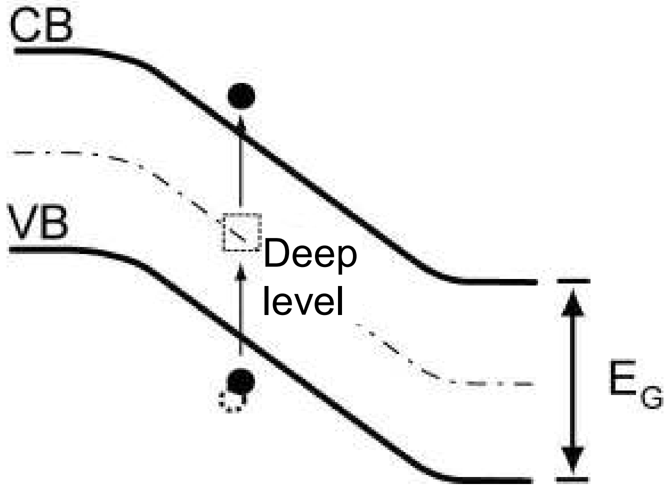




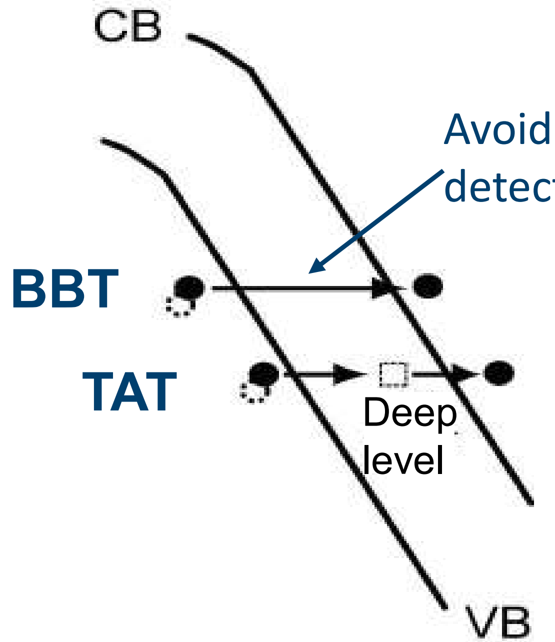
Bottom epi-layer thickess w_b can be adjusted for achieving shorter diffusion tail

Dark Count Rate

Thermal generation via deep levels
(@ low field $F < 10^5$ V/cm)



Field-enhanced generation



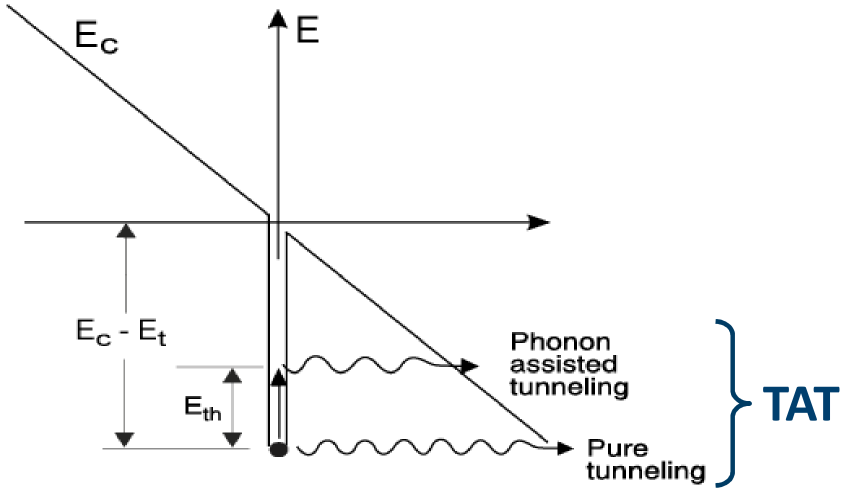
Avoided by suitable detector design !

Thermal generation and tunneling of carriers in the depletion region

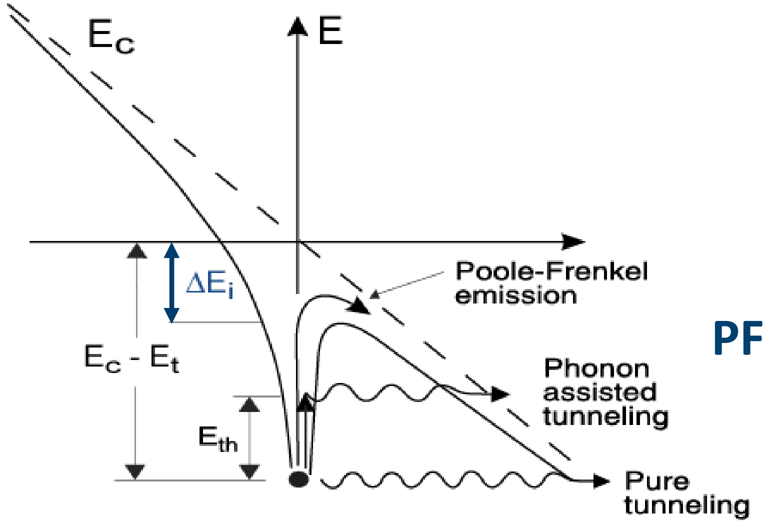
- Deep levels (traps) are mainly due to transition metal impurities
- Fe, Cu, Ti or Ni are usually found in silicon in concentrations of **~10^{11} - 10^{12}cm^{-3}** (unintentional contaminants)

Field-enhanced generation

Dirac well



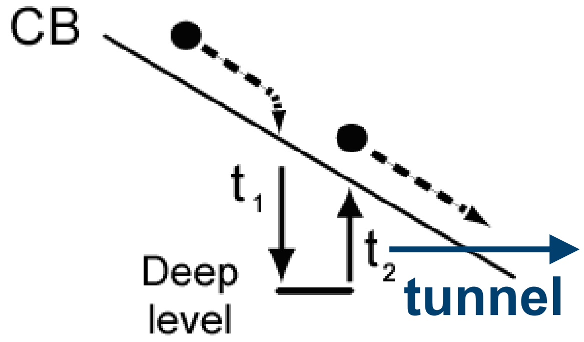
Coulomb well



- Phonon-assisted tunneling
  - barrier width decreased

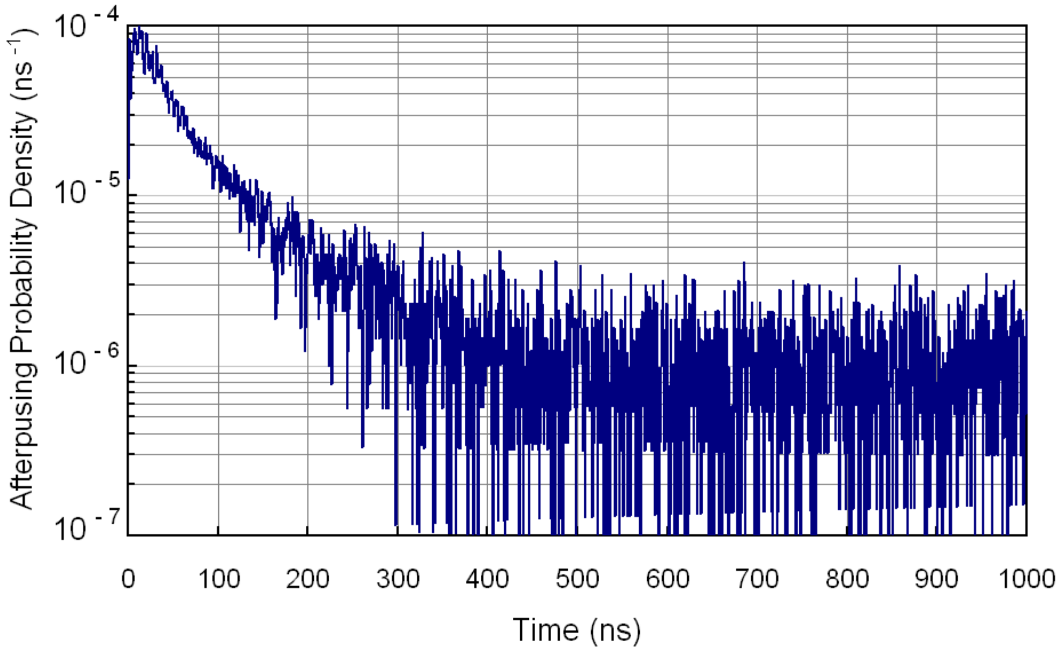
- Poole-frenkel effect
  - barrier height lowered

Afterpulsing



Afterpulsing Effect

- Carriers trapped during avalanche
- Carriers released later re-trigger the avalanche



Characterization of afterpulsing

- Time Correlated Carrier Counting (TCCC) method
- Afterpulsing negligible after 1 μs
- **Total afterpulsing probability:**

  < 1% @ room temperature

Challenges in SPAD development

Microelectronic Technology

- **Strict control** of transition metal contamination
  - ultra-clean fabrication process (defect concentration **< 10^9 cm^{-3}**)
  - suitable gettering processes **compatible** with device structure

Device design

- **Electric field engineering**

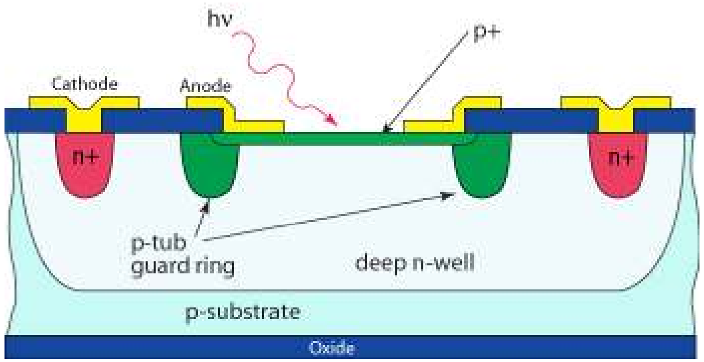
  avoids BB tunneling and reduces field-enhanced generation, with impact on:
  - → dark count rate
  - → dark count decrease with temperature
  - → photon detection efficiency
  - → photon timing jitter

Front-end electronics

- **Low-level sensing of the avalanche current →** avoids or reduces trade-off between timing jitter and active area diameter
- **Application-specific** electronics

SPADs in Standard CMOS technology

High-quality SPADs can now be produced
with industrial High-Voltage CMOS technologies.

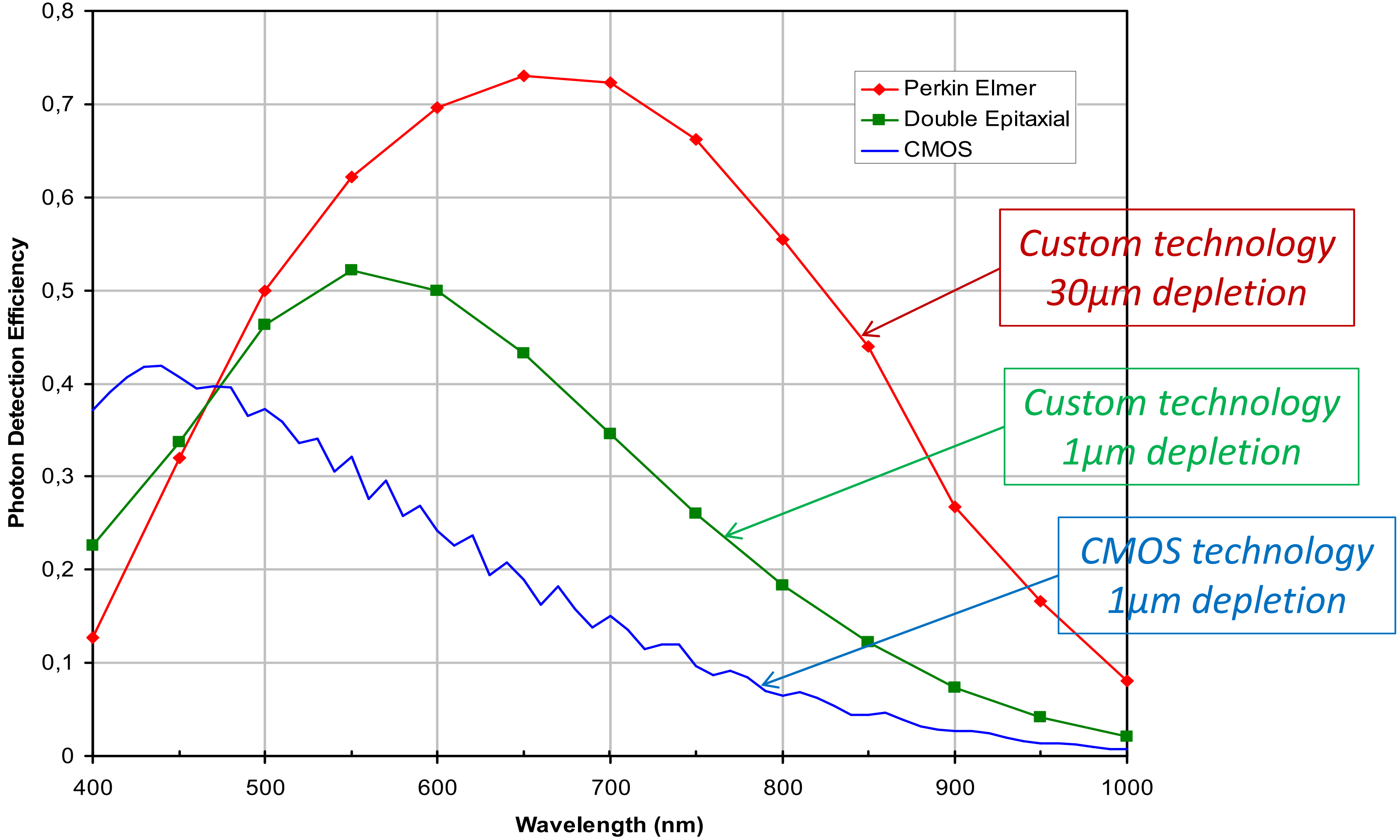


Some limitations have to be faced

- p^+n junction → hole-initiated avalanche → lower PDE
- Guard ring necessary
- no flexibility, device designers cannot modify the process
- the evolution of the technology is driven by circuit requirements, not by detectors!

but it is possible to integrate SPADs with circuits
and develop monolithic integrated systems

PDE Photon Detection Efficiency



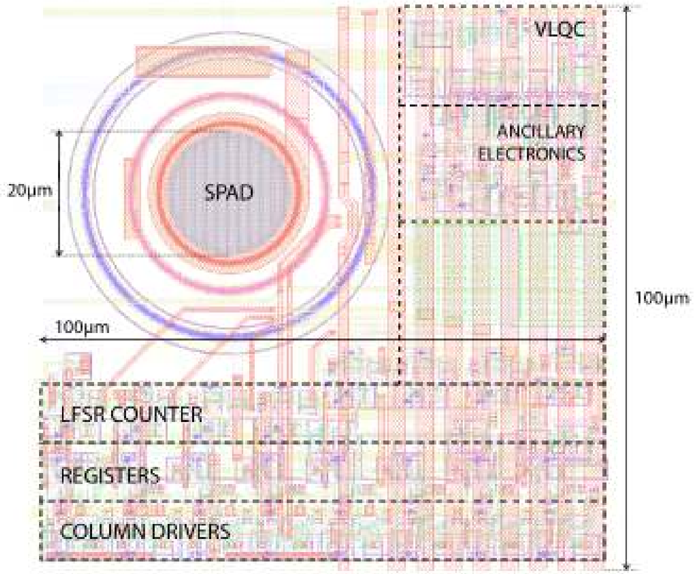
SPAD arrays

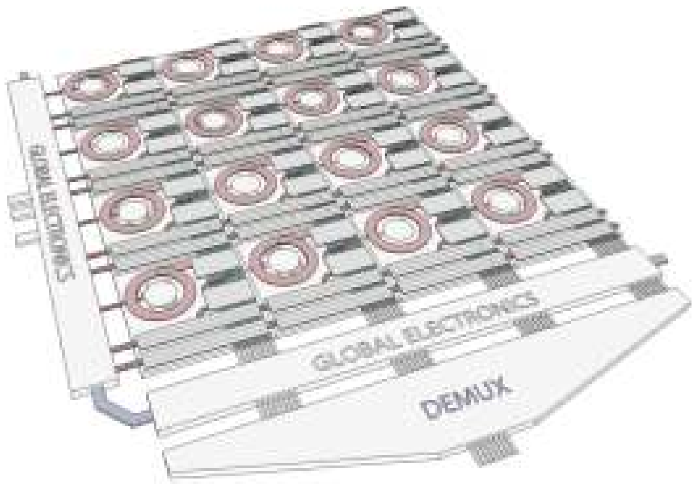
Two approaches in detector technology

- **Dense** arrays → **standard CMOS** technology
  - **small** pixel diameter (< 50µm, higher dark count rate density)
  - **large** number of pixels (>100 pixel)
  - **smart** pixels (in-pixel electronic circuitry)

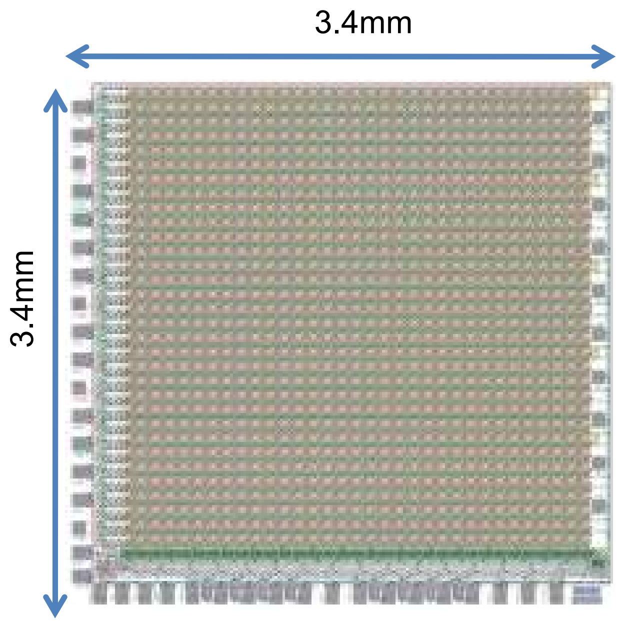
- **High-Quality-pixel** arrays → **Custom** technology
  - **wide** pixel diameter (> 100µm)
  - **low or moderate** number of pixels (< 100 pixel)
  - limitations due to off-chip electronics

SPAD Arrays in HV-CMOS technology

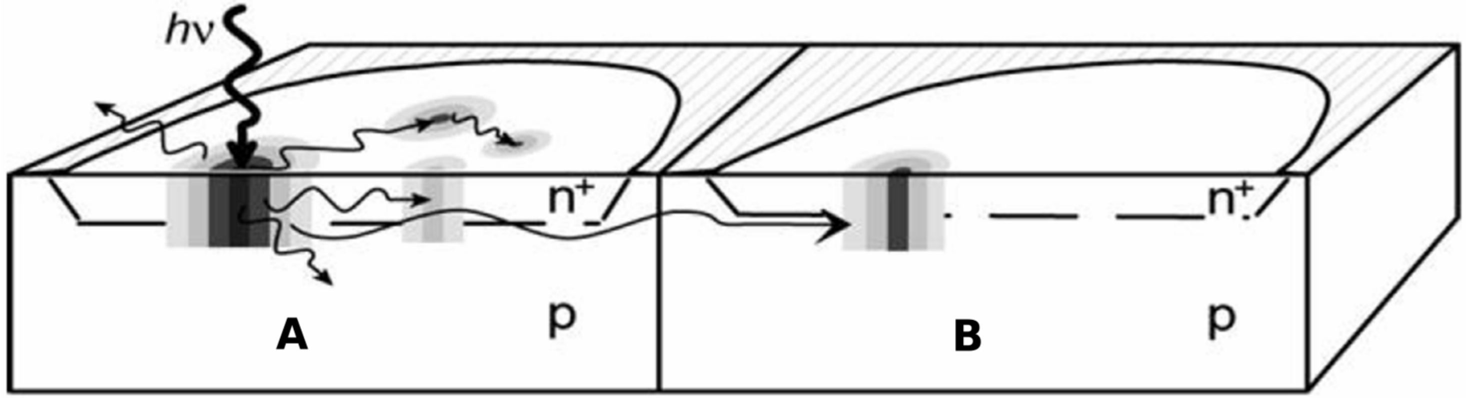




- Smart-pixel
  - ✓SPAD + AQC + counting electronics + register
- Fully parallel operation
  - ✓1024 pixel Single-Photon Imager
    - High frame rate single photon imaging
  - ✓can also act as a “Single pixel” large area detector
    - Low dead time, high count rate and photon number resolution
- Up to 100kframe/s for a 32x32 array
- No dead time between frames

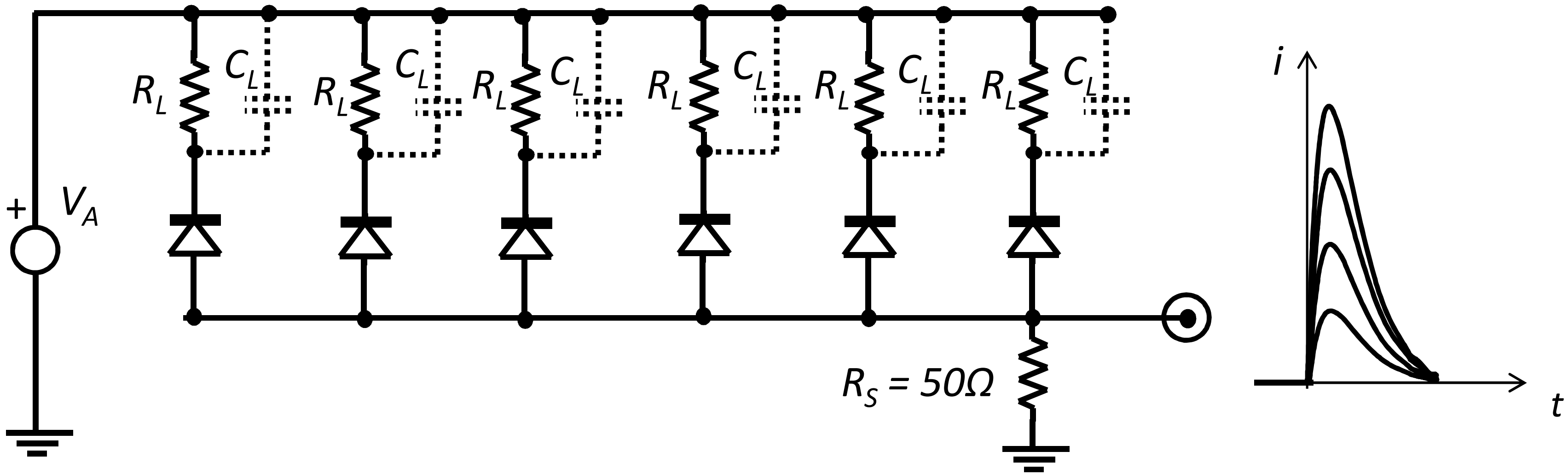


Optical Crosstalk in Arrays



- An impinging **photon triggers a primary avalanche** in a pixel (A)
- **Secondary photons** are emitted by the hot electrons of the avalanche current
- These photons propagate through the bulk silicon and can **trigger a secondary avalanche** in another pixel (B)
- The filling factor (Active area/ total area) is limited for limiting the crosstalk effect

Silicon PhotoMultipliers (SiPM)



This detector is a SPAD array where

- each pixel has an individual integrated quenching resistance $R_L \approx 100k\Omega$.
- each pixel has a very small individual load capacitance $C_L \approx 100\ fF$
- All pixels have a common ground terminal, connected to a low resistance external load, typically $R_S = 50\Omega$. The pixel currents all flow in this terminal, they are added

The detector pixels are thus

a) individually triggered by incident photons,

b) individually quenched by the discharge of the pixel capacitance

c) individually reset by the recharge of C_L with short time constant $R_L C_L \approx 10ns$

Silicon PhotoMultipliers (SiPM)

- The signal charge at the common output is proportional to the number of incident photons (at least as long as the light intensity on the detector is low enough to have negligible probability of more than one photon arriving on a pixel at the same time)
- Each pixel is a digital SPAD detector, but the pixel ensemble provides an analog information about the number of incident photons. The operation is indeed fairly similar to that of PMTs with microchannel plate multiplier. The detector was indeed conceived and is currently denoted as «Silicon PhotoMultiplier» SiPM.

With respect to PMTs, SiPMs offer various advantages

a) The typical properties of microelectronic devices (miniaturization; low voltage and low power; ruggedness; etc.)

b) remarkably higher detection efficiency, particularly in the red spectral range

c) operation insensitive to magnetic fields, which are detrimental for PMTs

However, SiPMs have also drawbacks with respect to PMTs

1. active area not as wide as PMTs
2. lower filling factor, with corresponding reduction of the photon detection efficiency
3. Fairly high dark current, that is, much higher dark current density over the active area

SPAD for the Near InfraRed (NIR)

Silicon absorbs up to λ = ***1.1μm***

⬇

Smaller bandgap required for working at longer λ

⬇

Mandatory:

- ***Deep cooling (< 220 K)***
  for limiting thermal carrier generation

  &

- ***Limitation to electric field***
  for avoiding tunnel-assisted generation

SPAD for the Near InfraRed (NIR)

•$In_{0.53}Ga_{0.47}As$ works up to $\lambda \sim 1.7\mu m$ because $E_g \sim 0.75$ eV

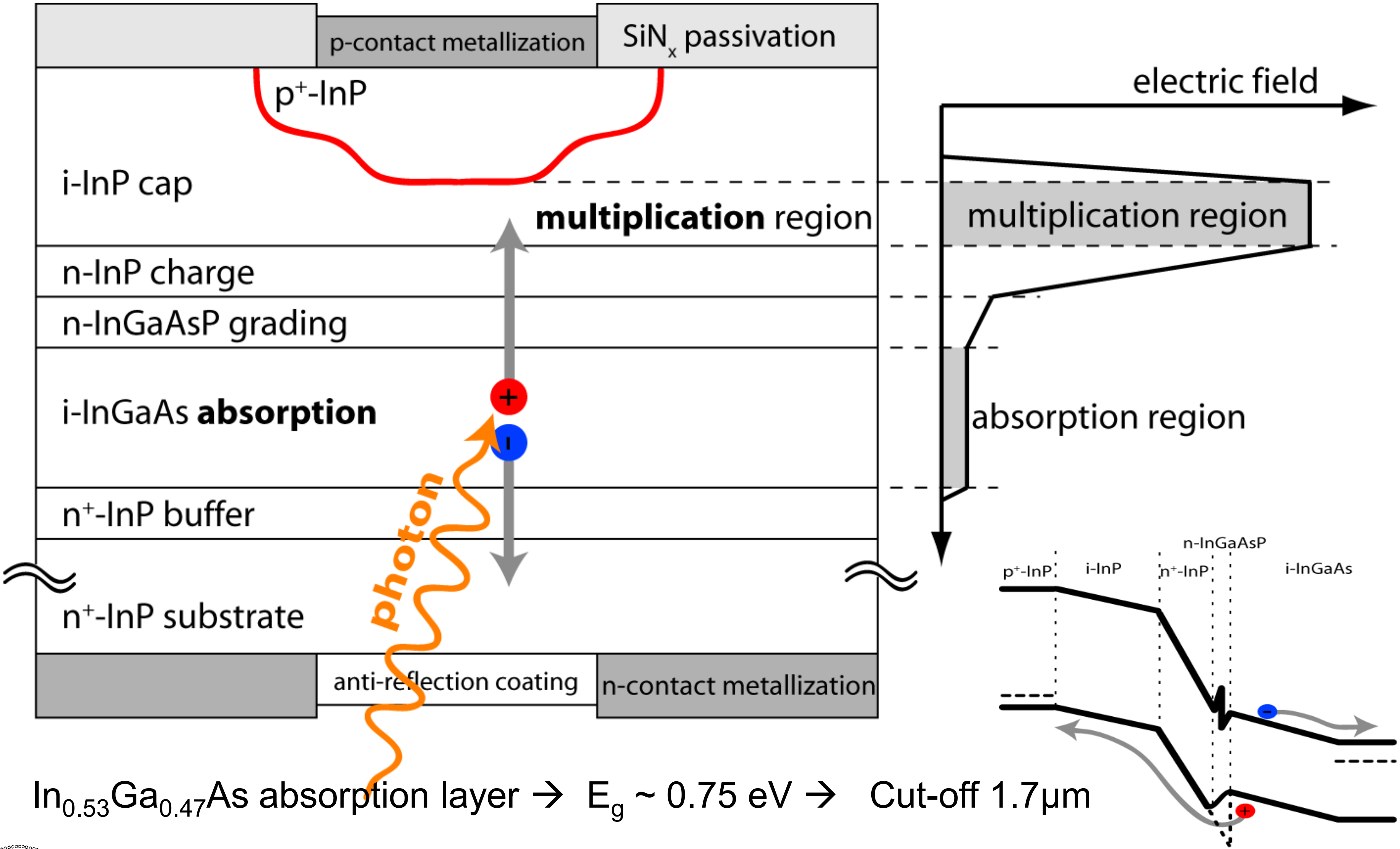
but

- it must be cooled
- it is unsuitable for avalanche

⬇

- *Separate Absorption and Multiplication* (SAM)
  - heterostructure device

Photon absorption and carrier collection



$In_{0.53}Ga_{0.47}As$ absorption layer → E_g ~ 0.75 eV → Cut-off 1.7μm

Photon Detection Efficiency of NIR detectors