Modern Particle Physics Detectors I

Theory
Applications
Practice

Lesson 4: Silicon Photo-Multipliers
Si Technology

Alessandro Bravar
http://dpnc.unige.ch/PhD

14 December 2017

What is a Si-PM?

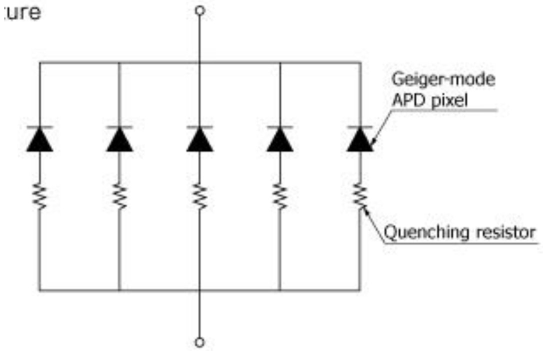
The Si-PM is a type of photon-counting device using multiple APD (avalanche photodiode) pixels operating in Geiger mode.

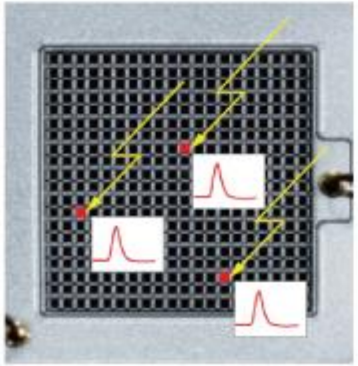
The Si-PM features a high multiplication ratio (gain), high photon detection efficiency, fast response, excellent time resolution, and wide spectral response range.

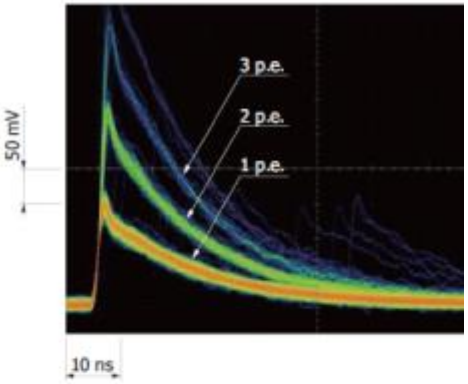
It is immune to magnetic fields, highly resistant to mechanical shocks, and will not suffer from "burn-in" by incident light saturation.

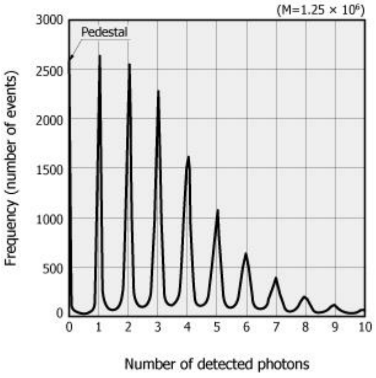
It has a wide range of applications and fields including medical diagnosis, academic research, and measurements.

It is used for low light detection, single photon to a few 1k photons possible.



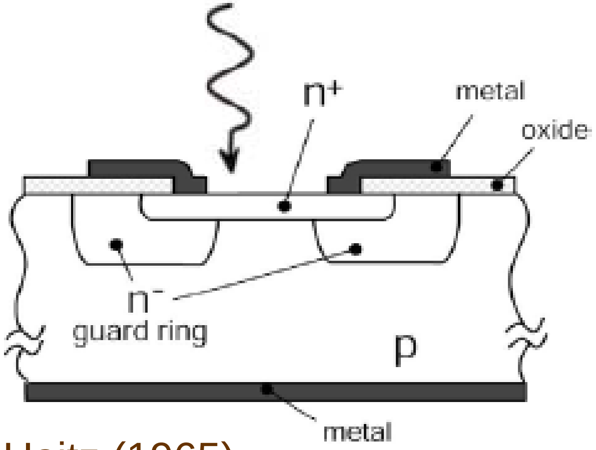




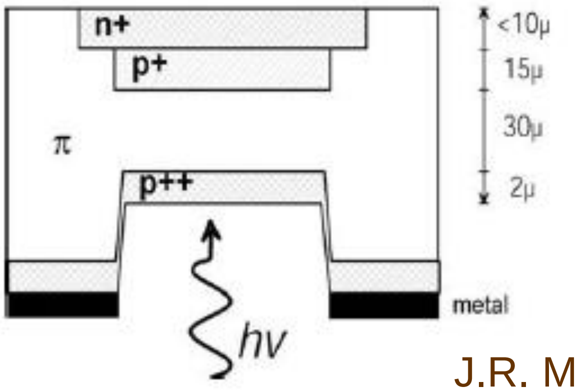


Geiger Mode Avalanche Photodiode (G-APD)

First single photon detectors operated in Geiger-mode

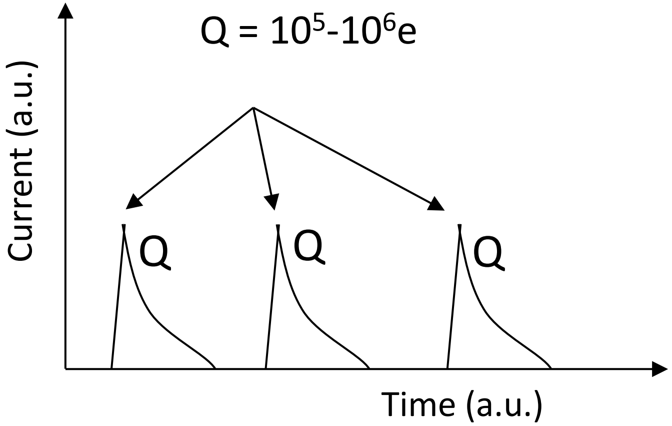


R.H. Haitz (1965)



J.R. McIntire (1966)

Standard output signal
$Q = 10^5$-10^6e



Binary device

if one or more simultaneous photons fire the GM-APD, the output is anytime a standard signal

$Q \sim C(V_{bias} - V_{BD}) \quad \rightarrow \quad gain = C(V_{bias} - V_{BD})/q$

The GM-APD does not give information on the light intensity

Single GM-APD to SiPM

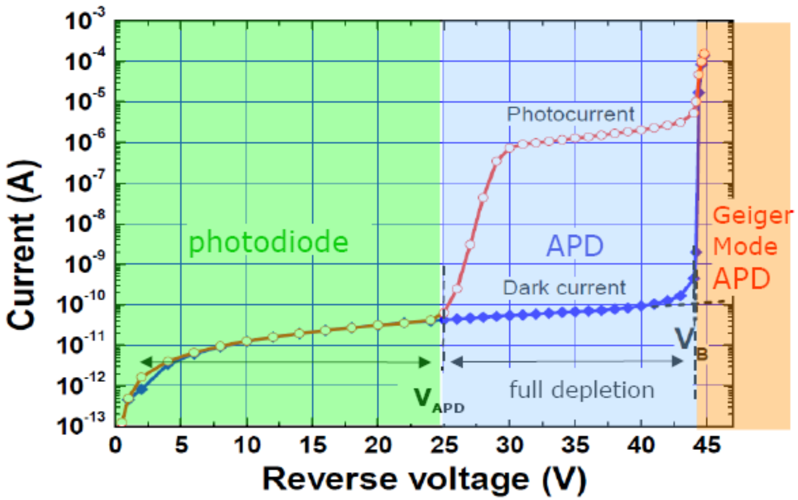
With increasing electrical field the signal gives no information about the light intensity. A matrix structure made from GM-APDs was first proposed by Golovin and Sadygov (1989)

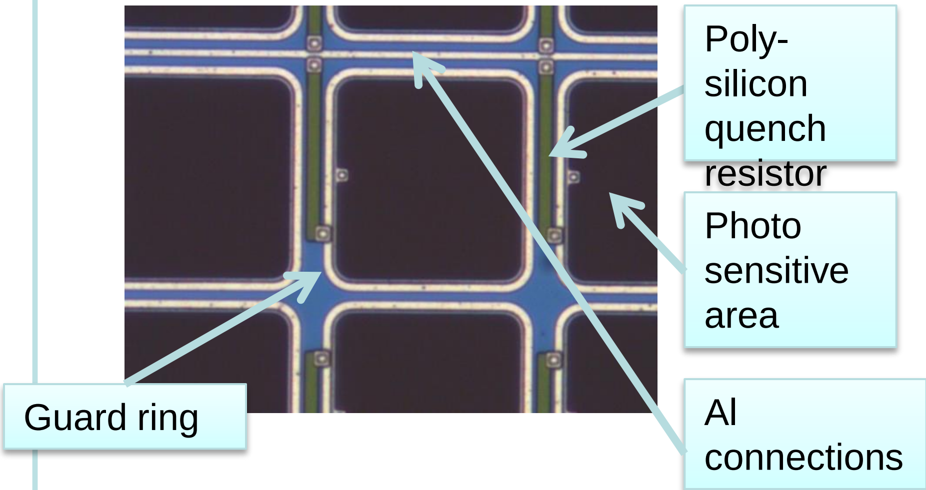
The SiPM is an array of GM-APD micro-cells called a pixel, with a typically cell size of 50x50μm^2
(from 5x5 μm^2 to 100x100 μm^2).

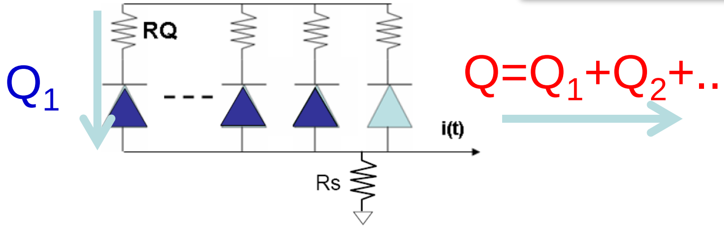
The signal from pixels is connected in parallel with resistors (100kΩ - MΩ) used for decoupling and quenching of the avalanche.

The signal from each pixel is identical, the total signal is proportional to the total number of fired pixels.

A single or multiple photons on a pixel creates the identical signal.







Geiger Mode Avalanche Photodiode (G-APD)

Use APDs operating in limited Geiger mode

Advantage: single photon counting
very high intrinsic gain (~10^6)

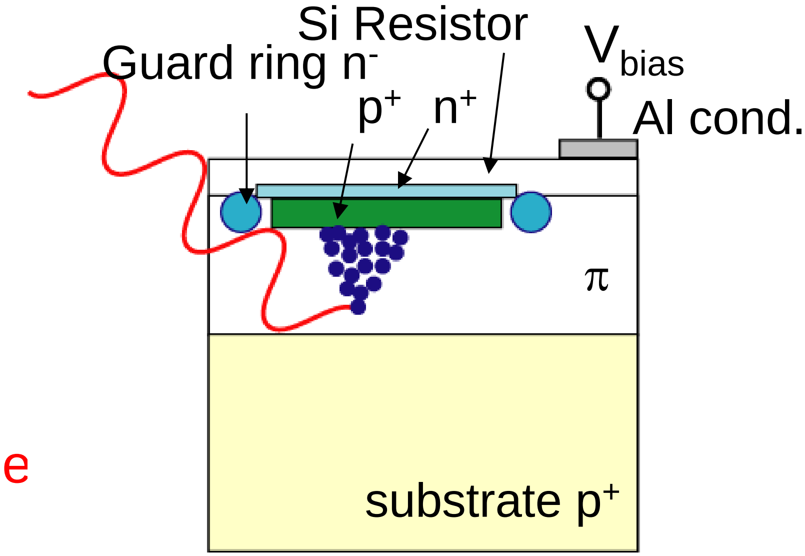
disadvantage: no dynamic range at all, however …

dark counts limits the area to < 200 μm Ø

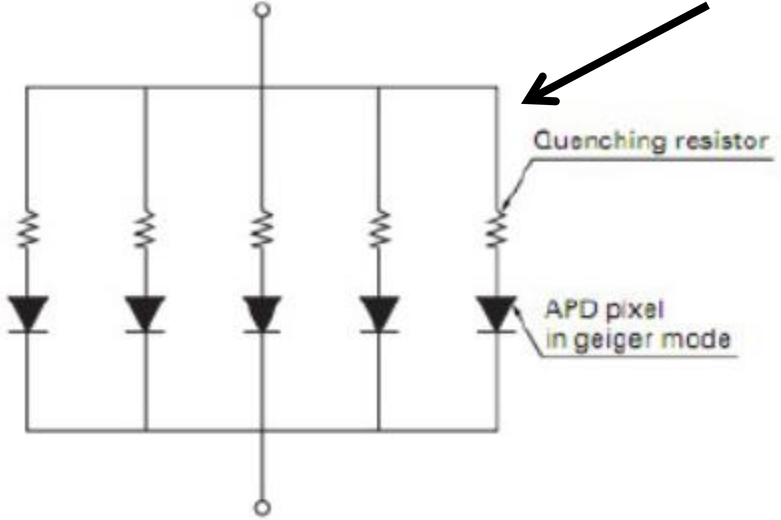
combine many small APD pixels onto the same substrate with a common anode

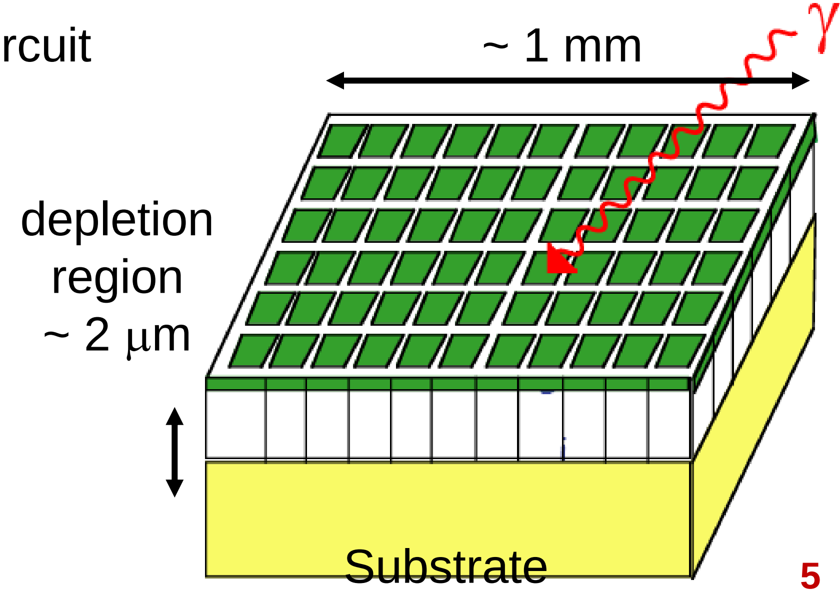
→ gain dynamic range

→ single photon resolution



equivalent circuit

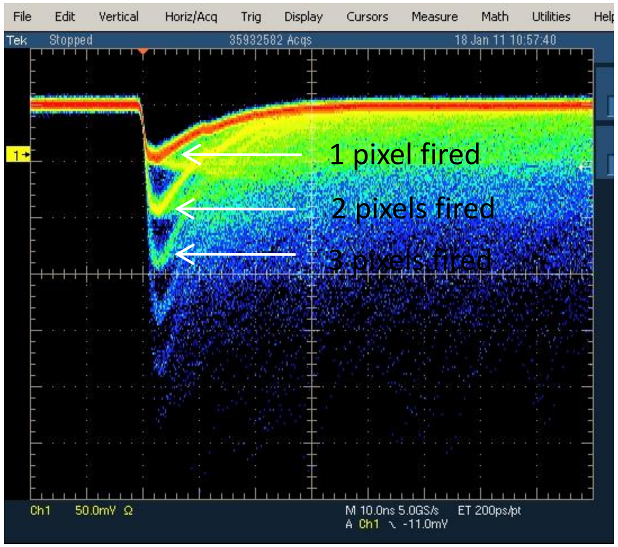


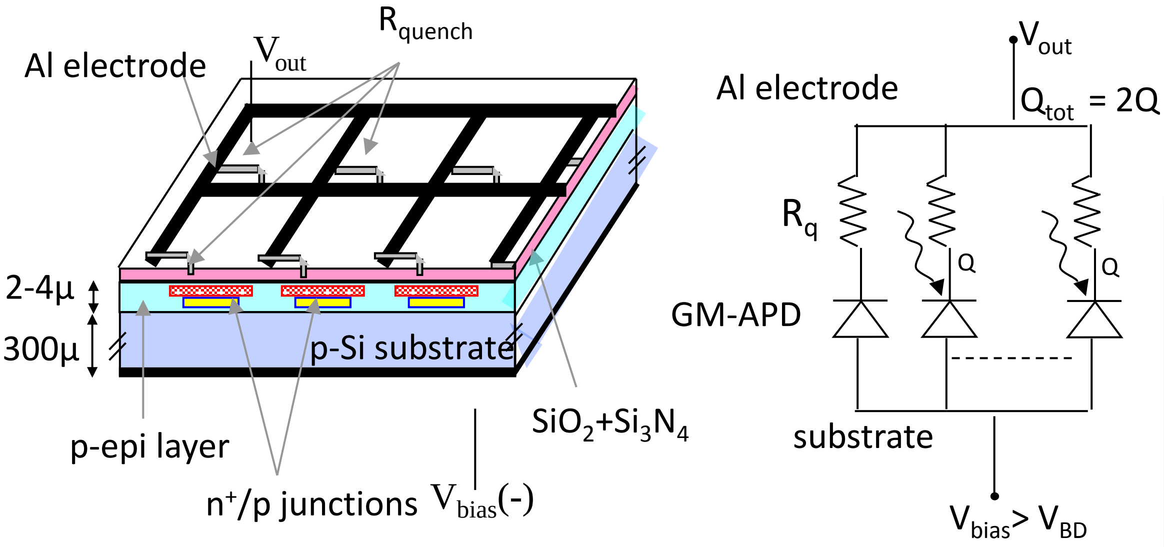


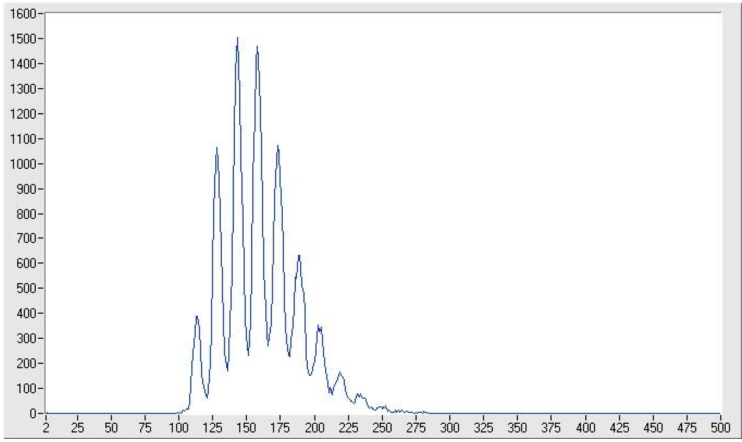
Silicon Photo-Multiplier (Si-PM)

matrix of n pixels connected in parallel
(e.g. few 100 - 1000 / mm^2)
on a common Si substrate

each pixels = GM-APD in series with R_{quench}







key developers :
V. Golovin, Z. Sadygov

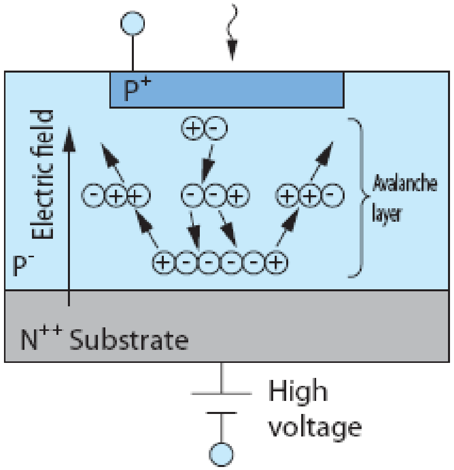
Quasi-analog device

output signal proportional to the number of “excited cells”, $Q \sim \Sigma Q_i$
$= \# \gamma$ if $\#\gamma < \#$ pixels!

MPPC – GM-APDs

Multi-pixel Avalanche PhotoDiodes operated in Geiger mode

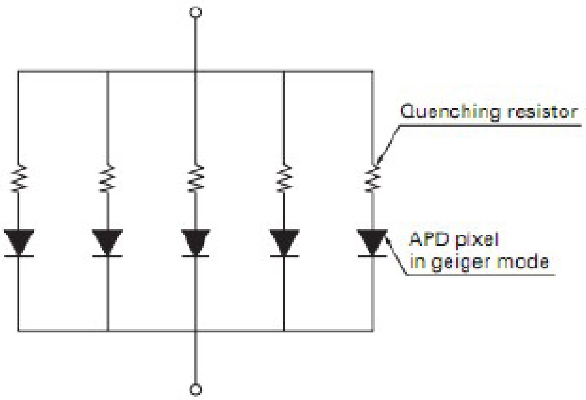
(Single pixel G-APDs developed long time ago (~ 1963), however not able to operate in multi-photon mode, sensitive area limited by dark current, etc.)



γ absorption / ionization

avalanche multiplication

Equivalent circuit



Geiger avalanched quenched by individual pixel resistors

high gain ~O(PMT) and high efficiency ~O(1-2 x PMT), 100 – 20,000 pixels / mm^2

each pixel works as a binary device, for low N_γ the device behaves as an analog detector

The GAPD produces a standard signal when any of the cells goes to breakdown.

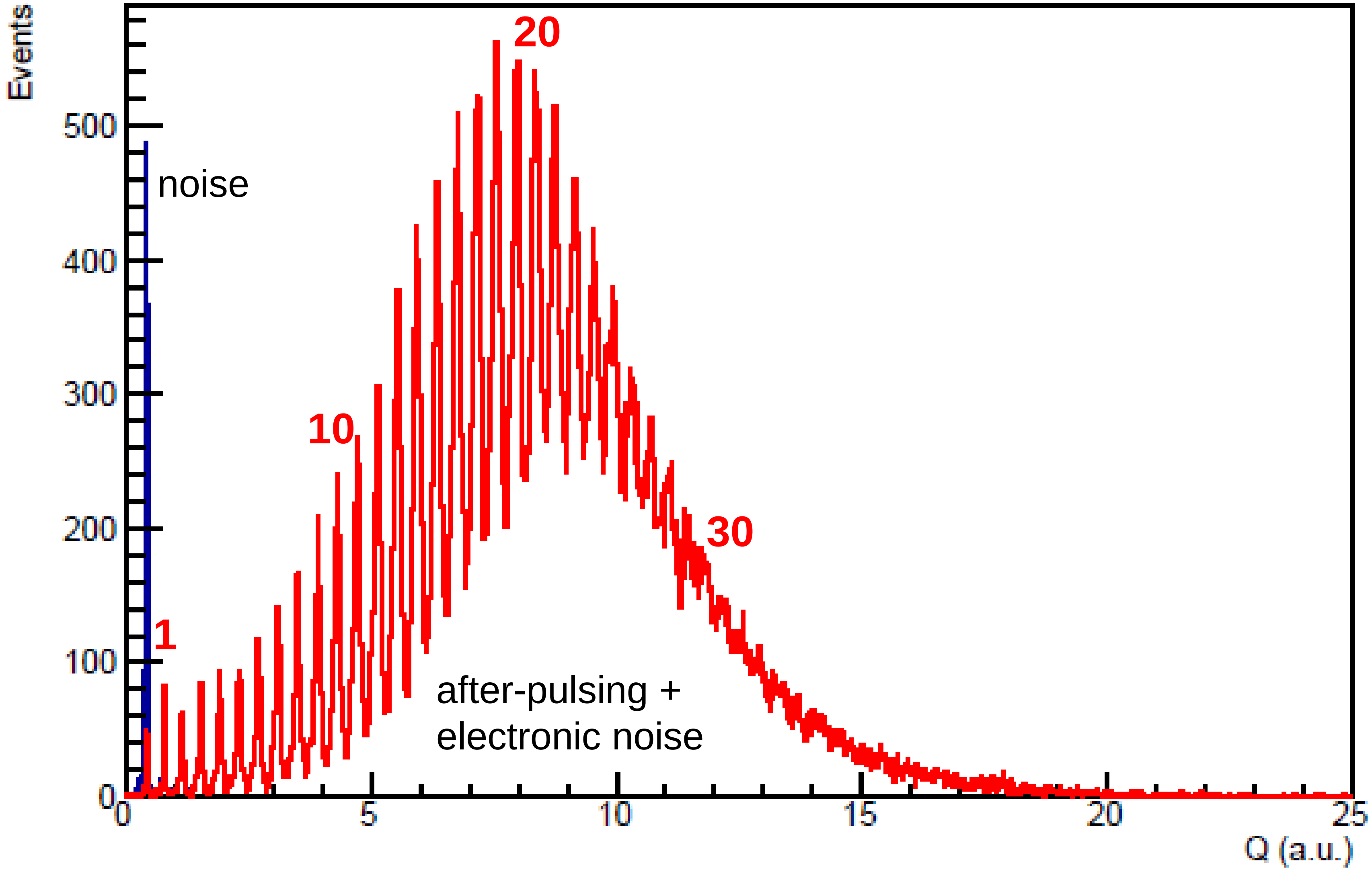
The amplitude A_i is proportional to the capacitance of the cell times the overvoltage

$$A_i \sim C \cdot (V - V_b) \quad \text{->} \quad G = C \cdot (V - V_b)/q$$

When many cells fire at the same time the output is the sum of the standard pulses

$$A = \Sigma A_i$$

Single Photon Counting



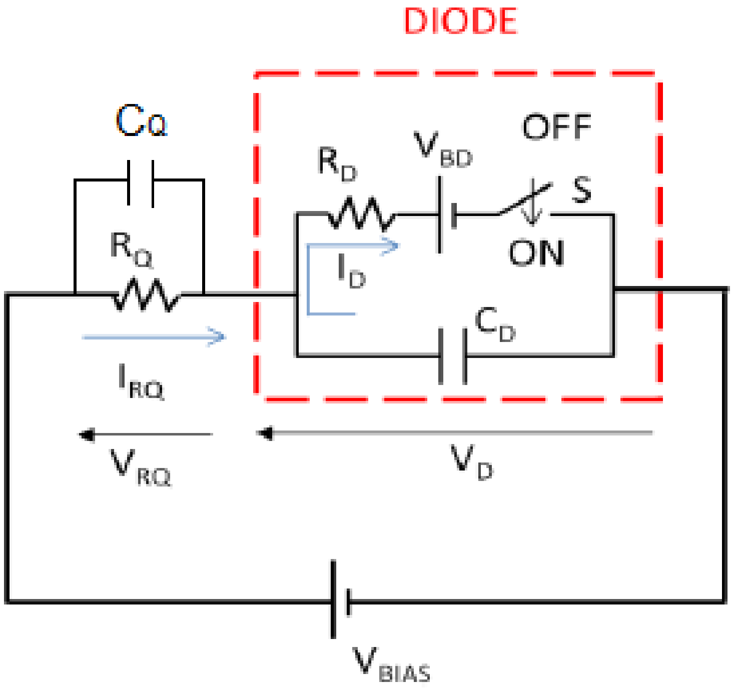
Linearity: The distance between the peaks is constant !

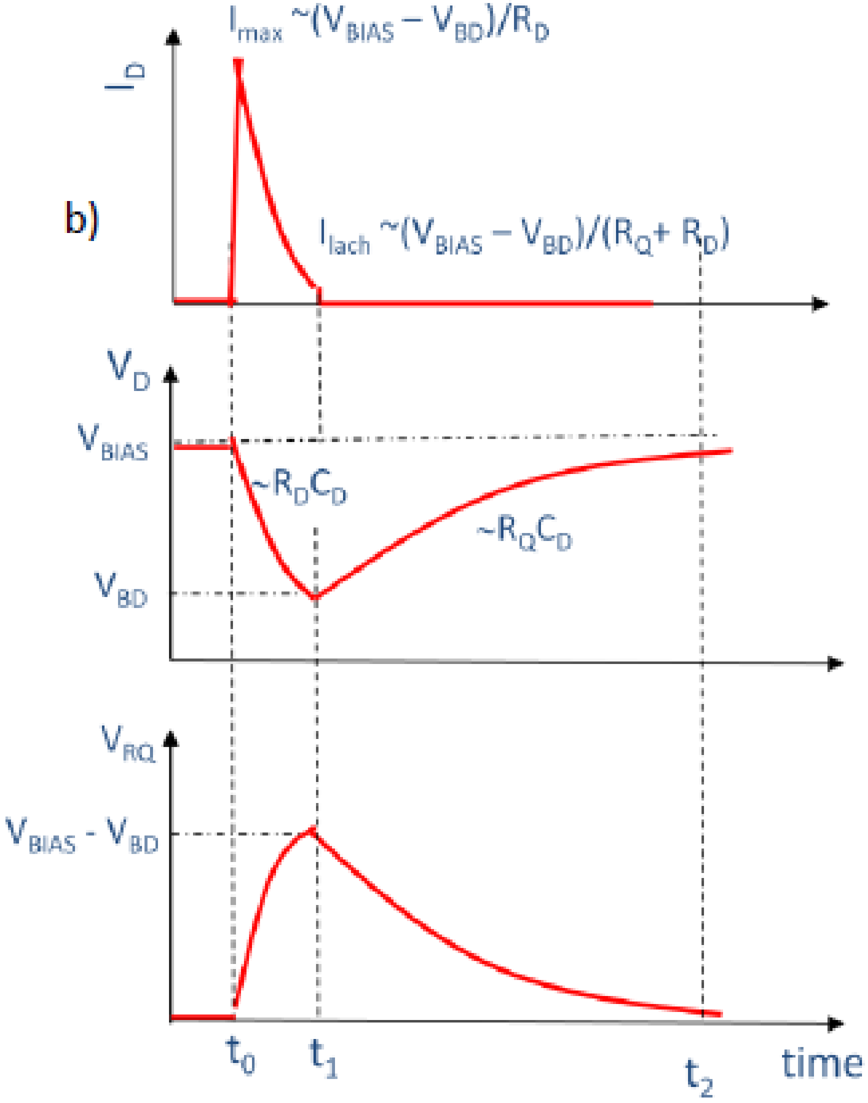
Si-PM Equivalent Circuit

Single cell

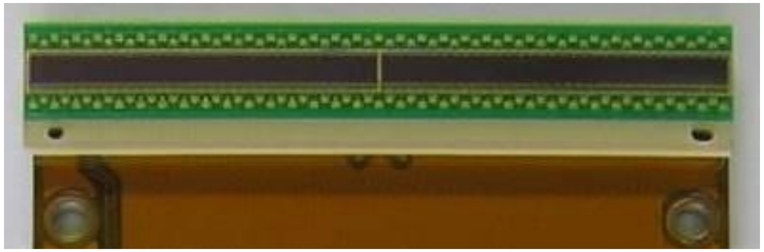
The time characteristics of the signal are determined by

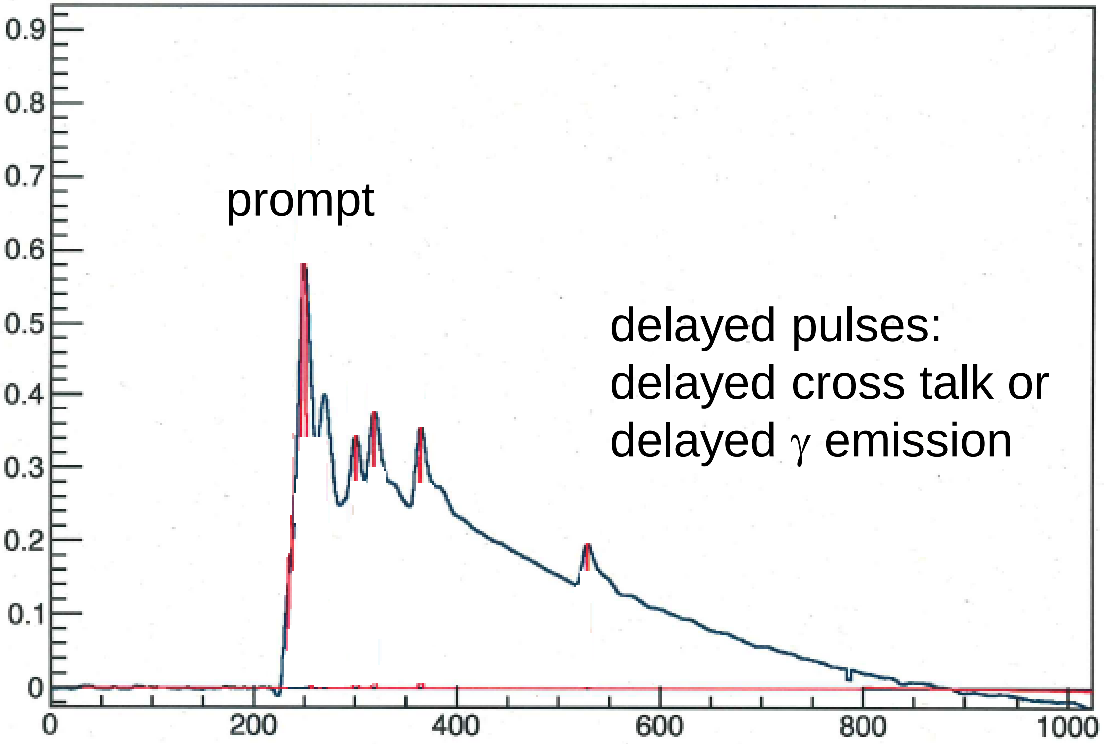
- detector capacitance C_D
- quench resistance R_Q
- detector resistance R_D
- quench capacitance C_Q





Pulse Shapes

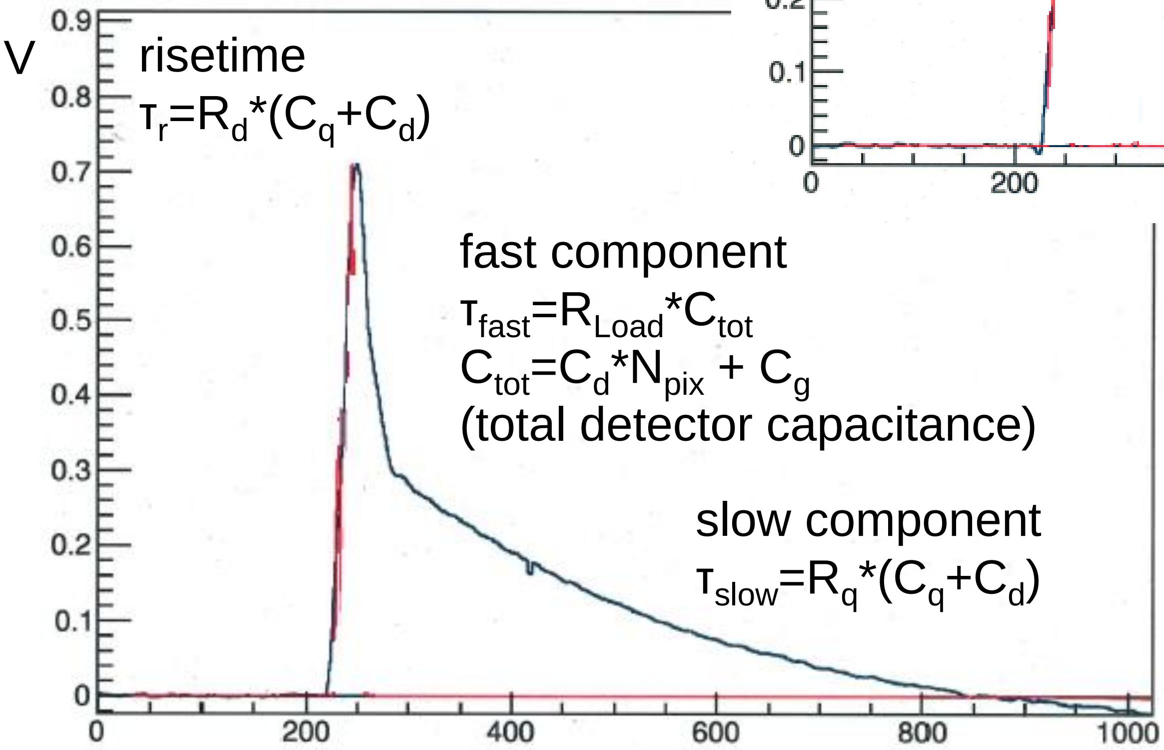




prompt

delayed pulses:
delayed cross talk or
delayed γ emission

200 ns



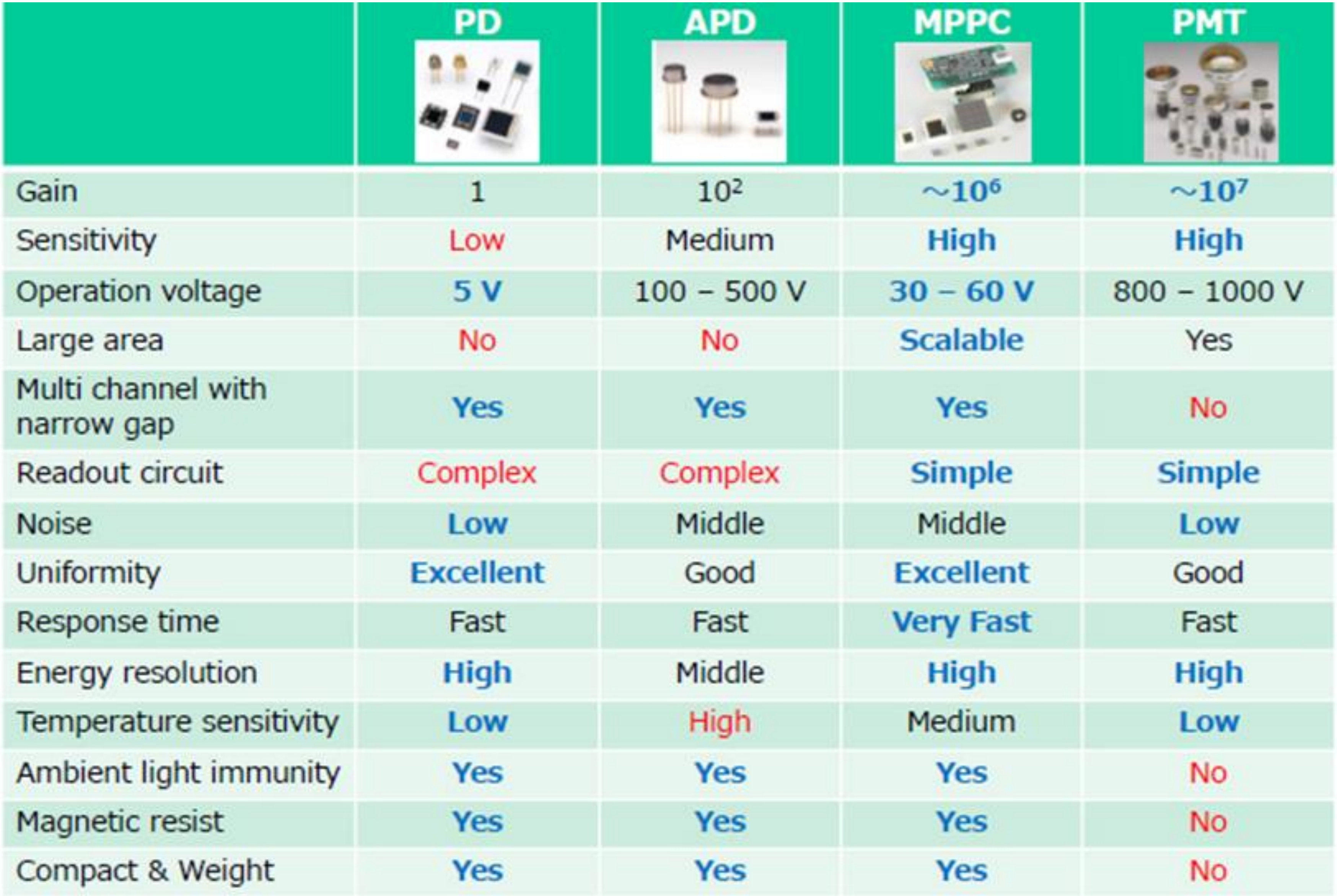
risetime
$\tau_r = R_d * (C_q + C_d)$

fast component
$\tau_{fast} = R_{Load} * C_{tot}$
$C_{tot} = C_d * N_{pix} + C_g$
(total detector capacitance)

slow component
$\tau_{slow} = R_q * (C_q + C_d)$

200 ns

Si-PM Comparison



	PD	APD	MPPC	PMT
Gain	1	10^2	$\sim 10^6$	$\sim 10^7$
Sensitivity	Low	Medium	High	High
Operation voltage	5 V	100 – 500 V	30 – 60 V	800 – 1000 V
Large area	No	No	Scalable	Yes
Multi channel with narrow gap	Yes	Yes	Yes	No
Readout circuit	Complex	Complex	Simple	Simple
Noise	Low	Middle	Middle	Low
Uniformity	Excellent	Good	Excellent	Good
Response time	Fast	Fast	Very Fast	Fast
Energy resolution	High	Middle	High	High
Temperature sensitivity	Low	High	Medium	Low
Ambient light immunity	Yes	Yes	Yes	No
Magnetic resist	Yes	Yes	Yes	No
Compact & Weight	Yes	Yes	Yes	No

PMT vs Si-PM

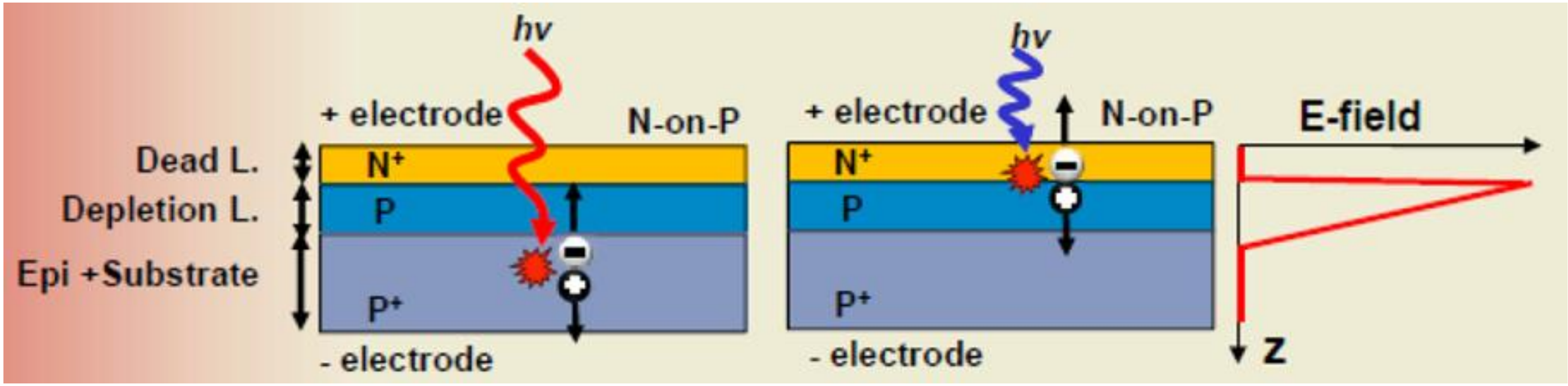
	PMT	MPPC
Gain	10^4~10^7	10^4~10^7
Photon Detection Eff.	0.1 ~ 0.2	0.2~ 0.5
Response	Fast	Fast
Photon counting	Yes	Great
Bias voltage	~ 1000 V	~ 20 - 90 V
Size	Small - Big	Compact
B field	Sensitive	Insensitive
Cost	Expensive	Not expensive (area!)
Dynamic range	Good	Determined by # of pixels
Long-term Stability	Good	Good
Robustness	Decent	Good
Radiation hardness	Good	Acceptable
Noise	Quiet	Noisy (order of 10 kHz)

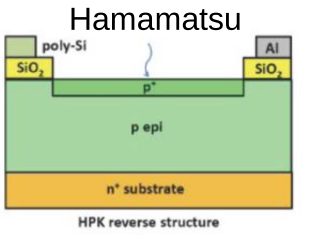
n-on-p Devices vs *p-on-n* Devices

longer wavelength → deeper penetration depth

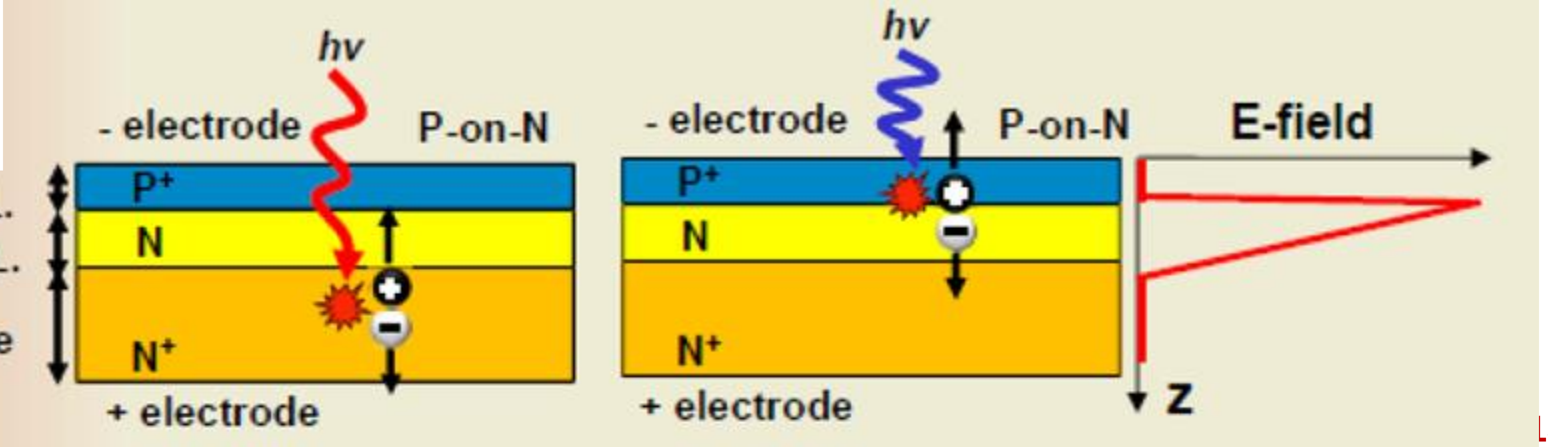
electrons: higher mobility → higher ionization probability

n-on-p higher QE at longer wavelengths





p-on-n higher QE at shorter wavelengths

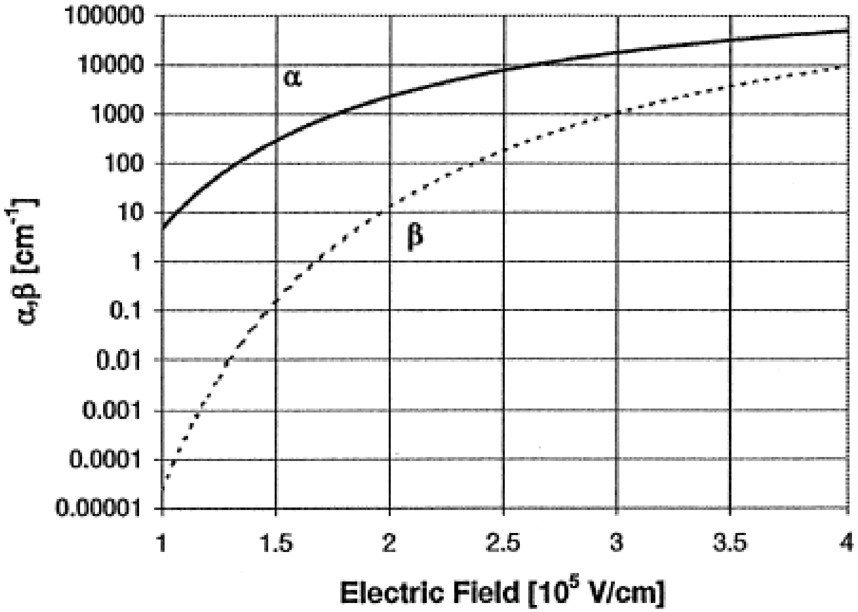


Avalanche in a p-n Junction APD

Applying a high electric field on a p-n junction may cause an avalanche multiplication of electrons and holes created by absorbed light.
(K.G. McKay, K. J. McAffe"Electron multiplication in silicon and germanium", Phys.Rev. 91 (1953))

Ionization coefficients for electrons and holes in Si (at room temperature)

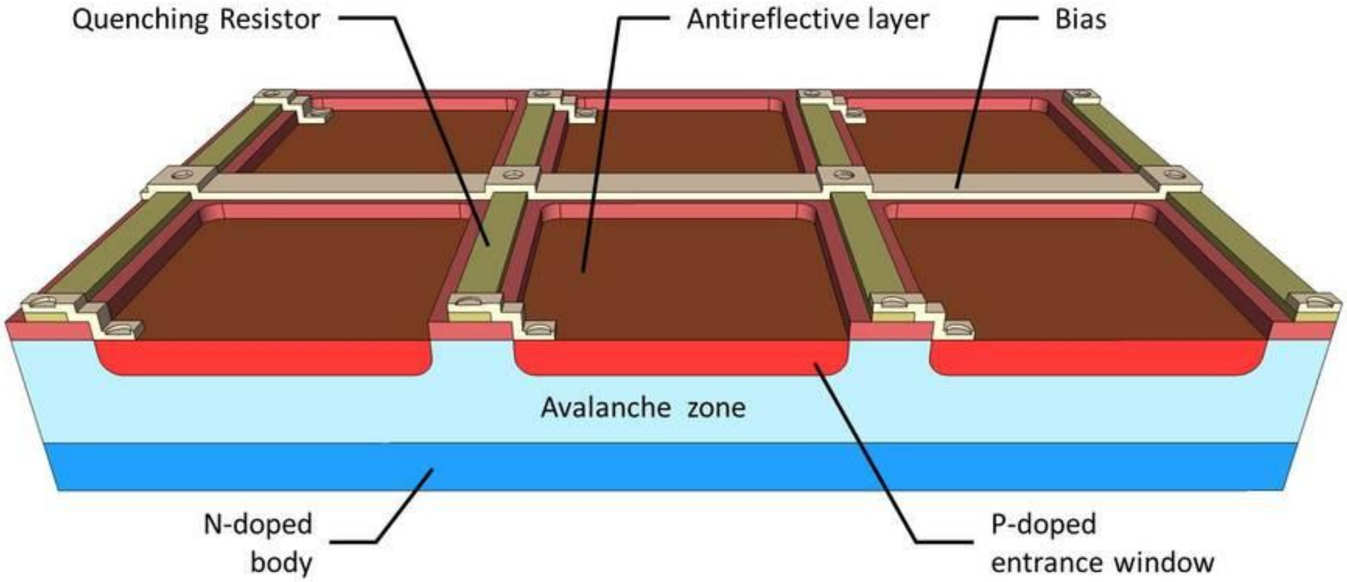
Strong dependence of the electric field



Avalanche multiplication is a stochastic process, it creates multiplication noise
ENF for APD: $F=k*M+(1-k)(2-1/M)$
$k=\beta/\alpha$ (k-factor)
M-average multiplication coefficient
β-ionization coefficient of holes
α-ionization coefficient of electrons
(R.J. McIntyre, IEEE Tr. ED-13 (1972) 164)

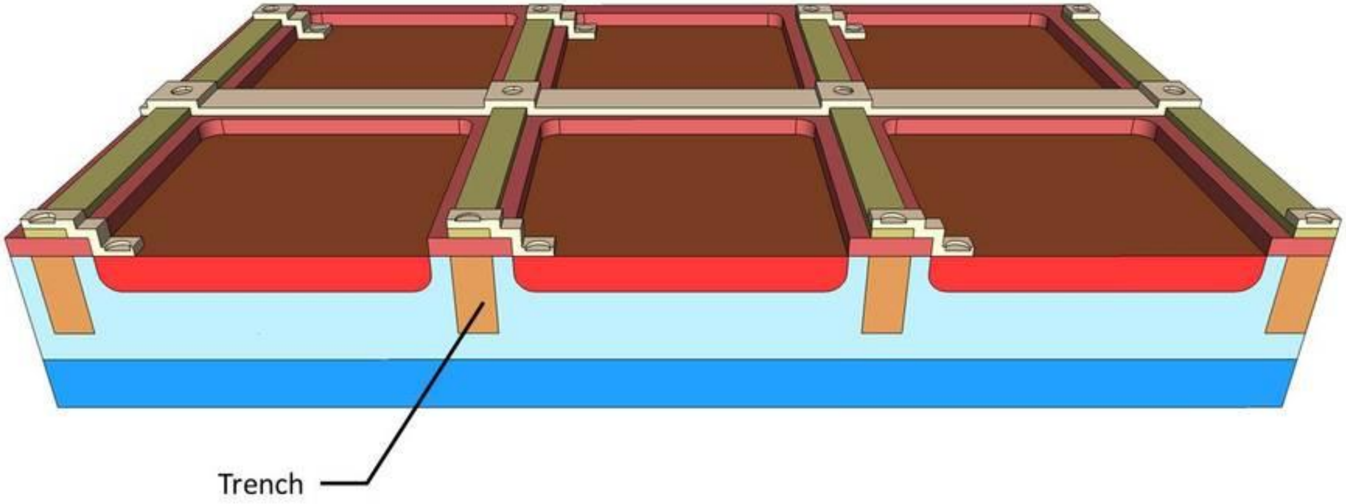
S. Vinogradov "Analytical models of probability distribution and excess noise factor of Solid State Photomultiplier signals with crosstalk"

Basic Si-PM Structure (KETEK): p-on-n



p-on-n

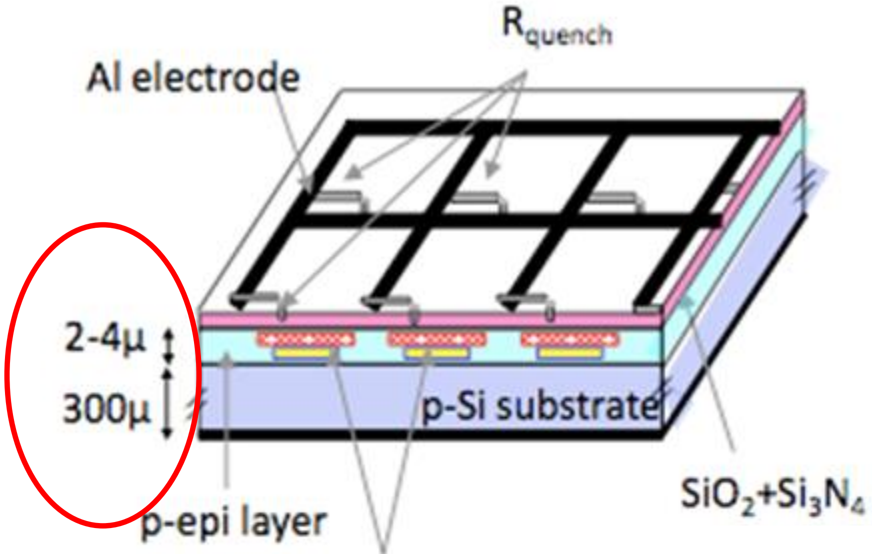
standard design

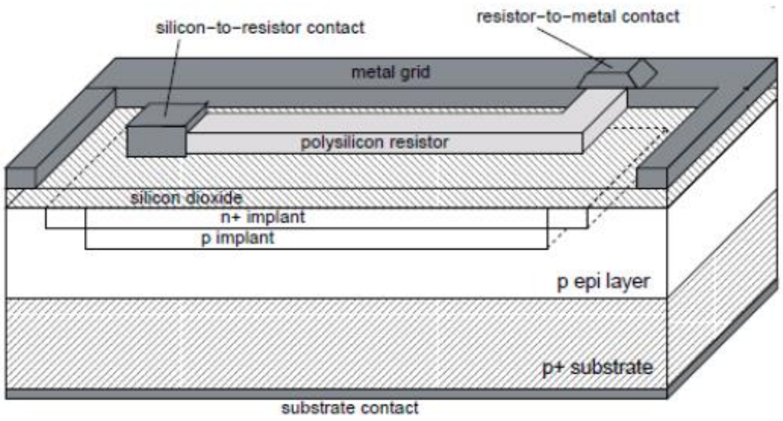


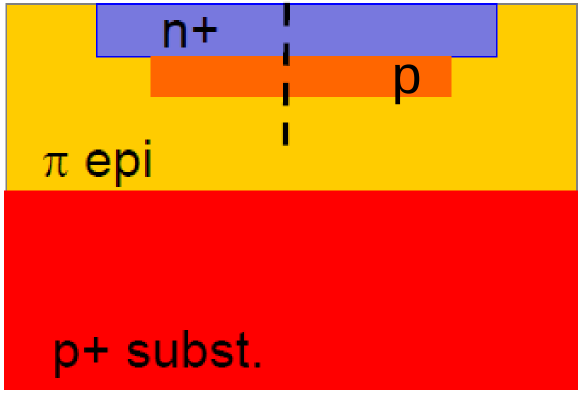
p-on-n

design with trenches
(to reduce x-talk)

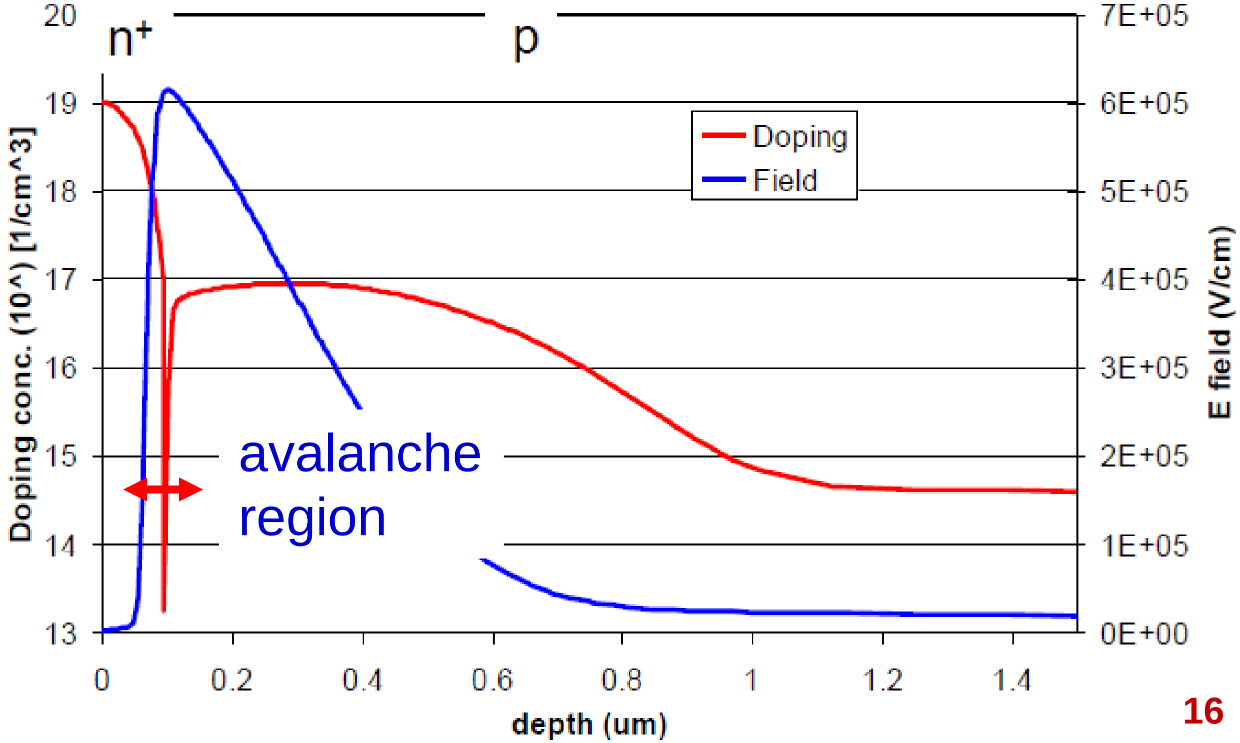
Basic Si-PM Structure (FBK): n-on-p







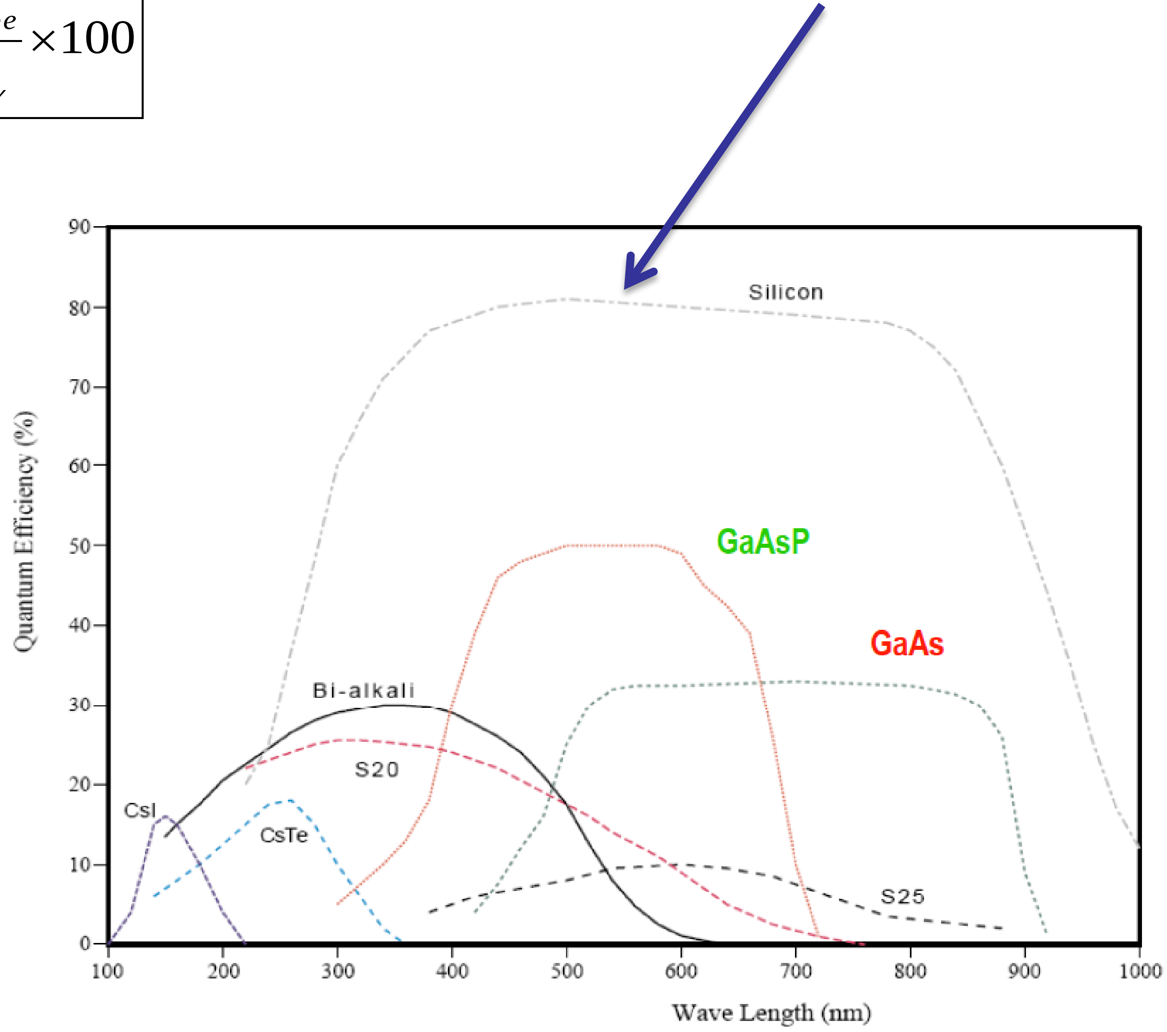
electric field “inside” the Si-PM



Quantum Efficiency

Silicon is at least two times better than photo cathodes, but this is not yet the PDE !

$$\boxed{QE(\%) = \frac{N_{pe}}{N_{\gamma}} \times 100}$$



Quantum Efficiency

Two factors influence QE:

1) Transmittance of the entrance window

ARC (anti reflective coating),

dielectric on silicon surface (SiO_2)

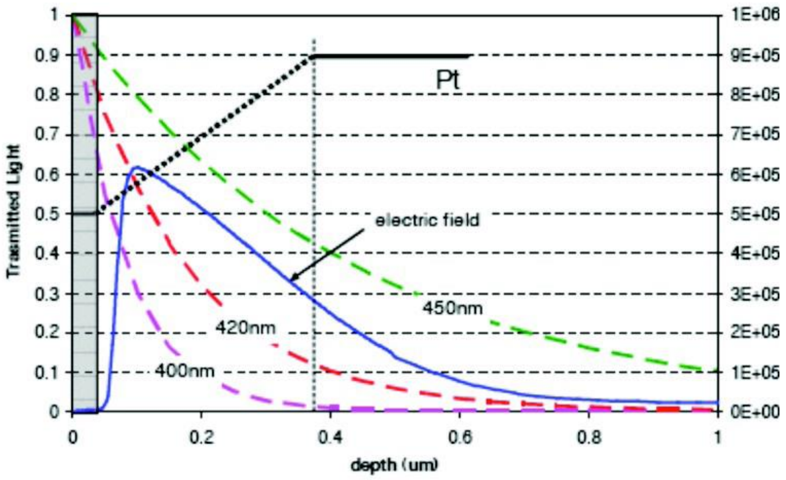
2) Probability of a photon to generate a e-h pair in the active layer

QE optimization:

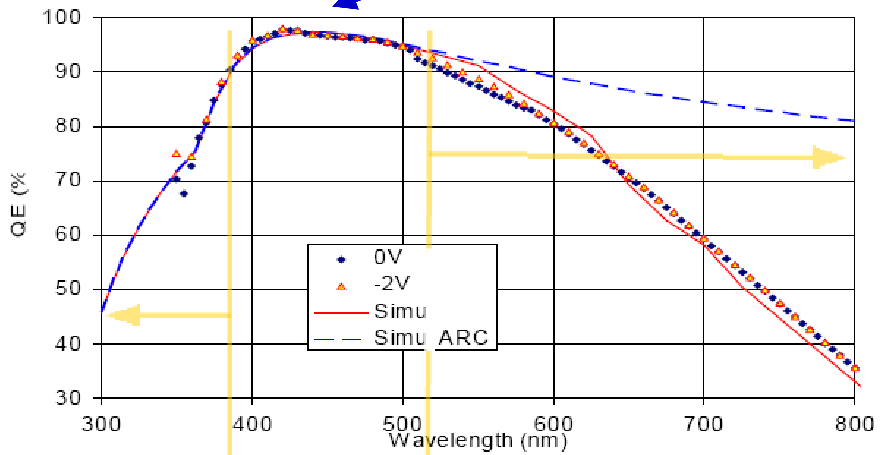
ARC

Shallow junction for short wavelength

Thick epi layers for long wavelength



Is almost 100% !



Reduced by
ARC Transmittance

Reduced by the
small π layer thickness

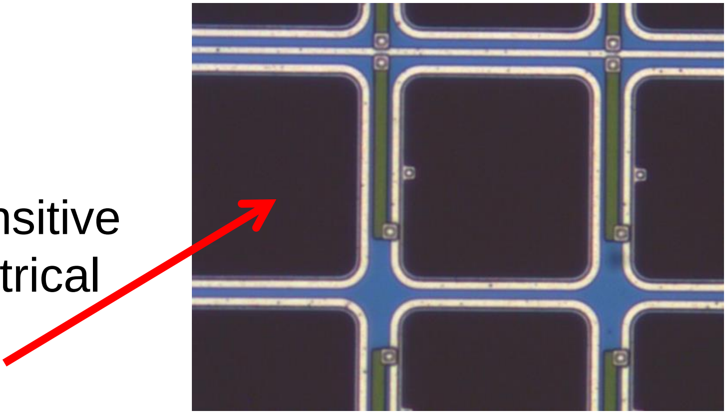
Photon Detection Efficiency (PDE)

Definitions:

1. The radiant sensitivity (S) [A/W]: is the ratio between the output current from a PD and the input radiant power at a given wavelength. S is related to quantum efficiency QE by:

$$\boxed{QE(\%) \approx 0.124 \times \frac{S(A/W)}{\lambda(nm)}}$$

2. The Fill Factor (GFF) is the ratio between the sensitive surface and the detector surface also called geometrical efficiency $\varepsilon_{\mathrm{goem}}$.



3. Collection Efficiency or avalanche trigger probability $\varepsilon_{\mathrm{AT}}$ is the probability to transfer the primary PE or e/h to the amplification stage.

4. Photon Detection Efficiency (PDE) is the probability that a single photon trigger a detectable output (this is the overall quantity)

$$\boxed{\mathrm{PDE} = \mathrm{QE} \times \varepsilon_{\mathrm{goem}} \times \varepsilon_{\mathrm{AT}}}$$

Electrical Models for a GM-APD

Passive quenching studied in detail in the '60 to model micro-plasma instabilities McIntrye JAP 32 (1961), Haitz JAP 35 (1964)

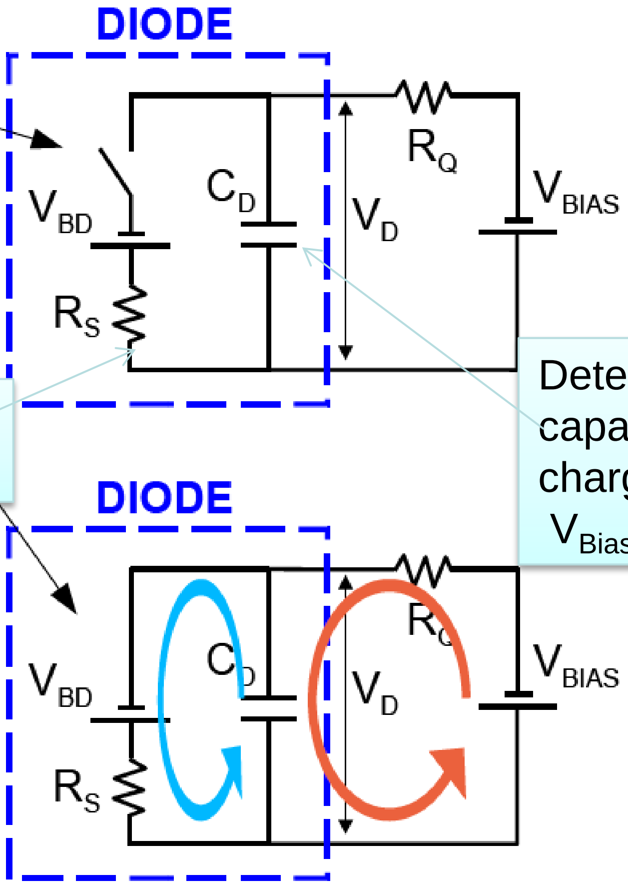
The Geiger-Mode APD can be modeled with an electrical circuit and two probabilities:

- Switch OFF = micro-plasma non-conducting
- Switch ON = micro-plasma conducting

- C_d diode capacitance (some 10fF)
- R_s series resistance (~1kΩ)
- R_q quenching resistance (> 300kΩ)
- $V_{bd} < V_{bias}$ (few % relative)

- P_{01} turn-ON
  Probability that a carrier traversing the high field region trigger an avalanche

- P_{10} turn-OFF
  Probability that number of carriers in the high field region fluctuates to 0



Series resistance during the avalanche

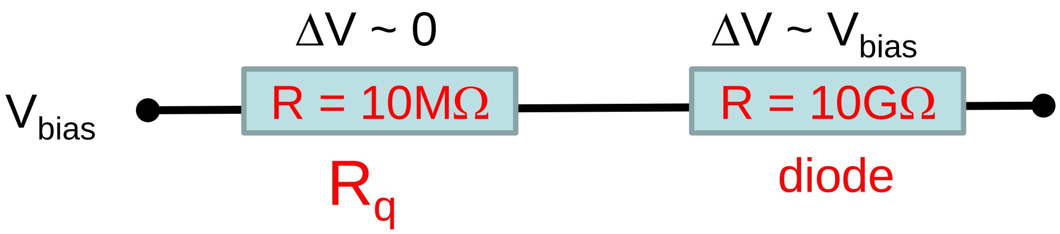
Detector capacitor charged to $V_{Bias}-V_{BD} = V_D$

Internal/external currents

Quenching

How to stop the avalanche?

Switch of the bias!

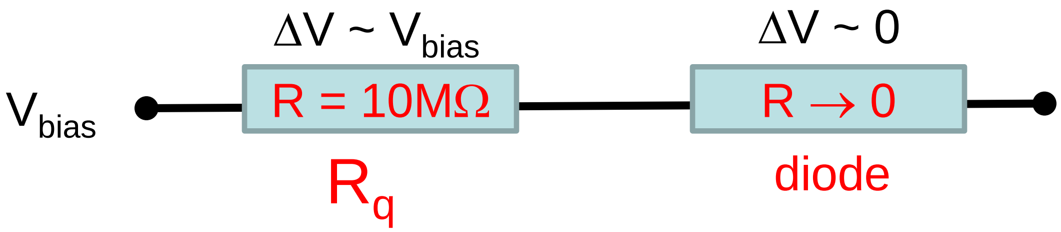


$\Delta V \sim 0$ $\Delta V \sim V_{bias}$

V_{bias} R = 10MΩ R = 10GΩ

R_q diode

during the avalanche, the diode becomes conductive: R → 0



$\Delta V \sim V_{bias}$ $\Delta V \sim 0$

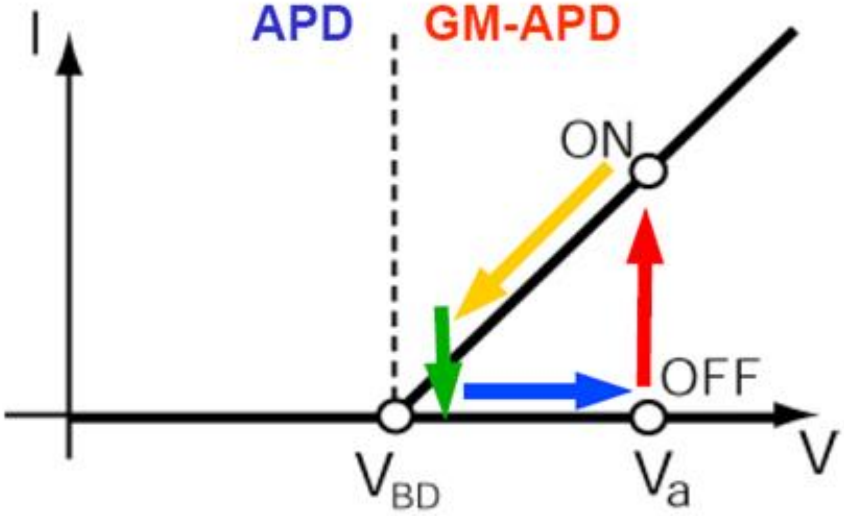
V_{bias} R = 10MΩ $R \rightarrow 0$

R_q diode

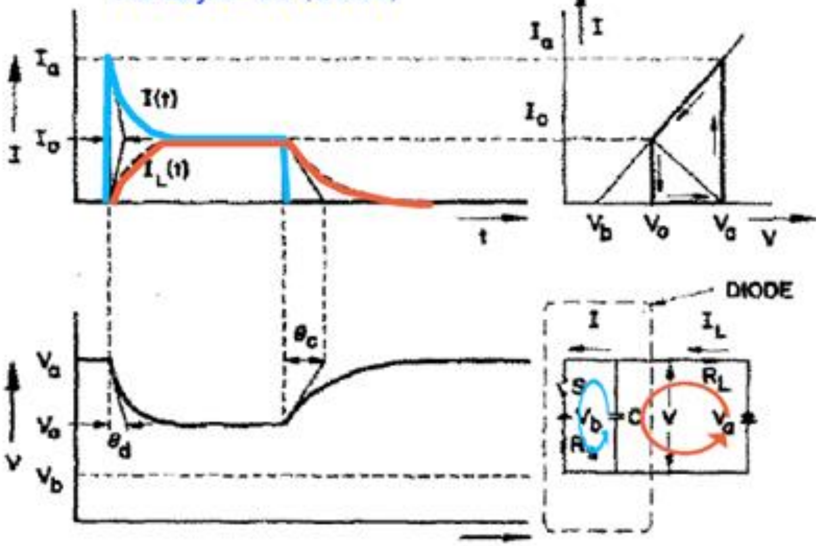
and the (bias) voltage drops to zero across the quench resistor R_q,
so that there is no bias on the diode.

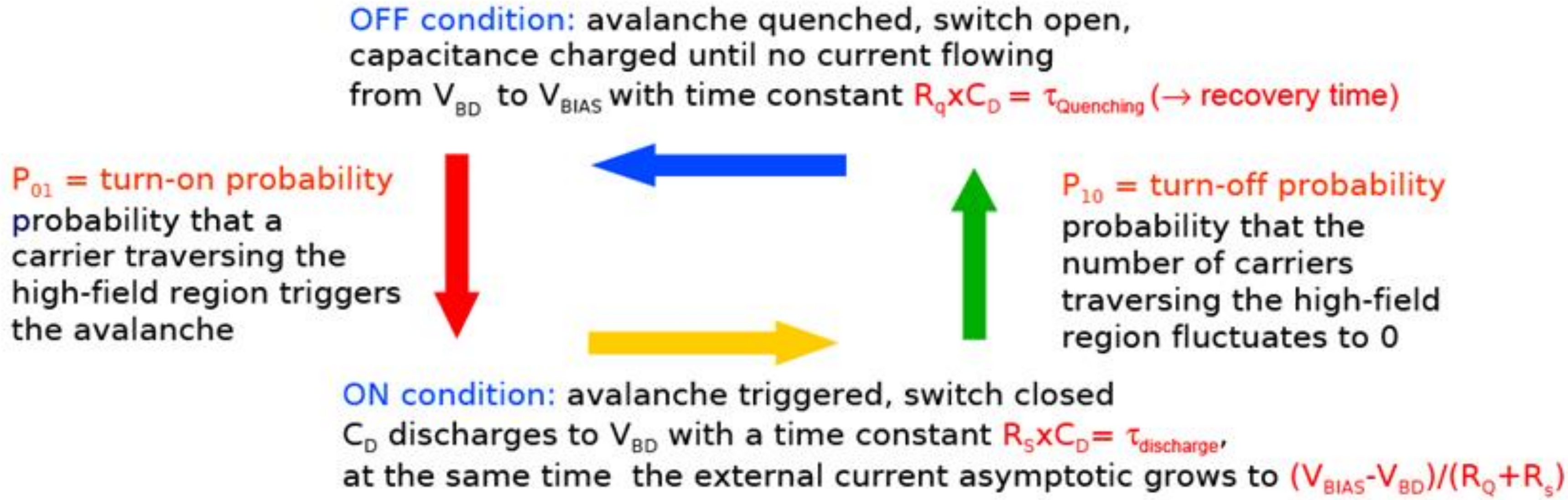
The avalanche stops,
and the device starts recovering and getting ready for the next “event”.

Operation principle of a GM-APD



Haitz JAP 35 (1964)





OFF condition: avalanche quenched, switch open, capacitance charged until no current flowing from V_{BD} to V_{BIAS} with time constant $R_q x C_D = \tau_{Quenching}$ (→ recovery time)

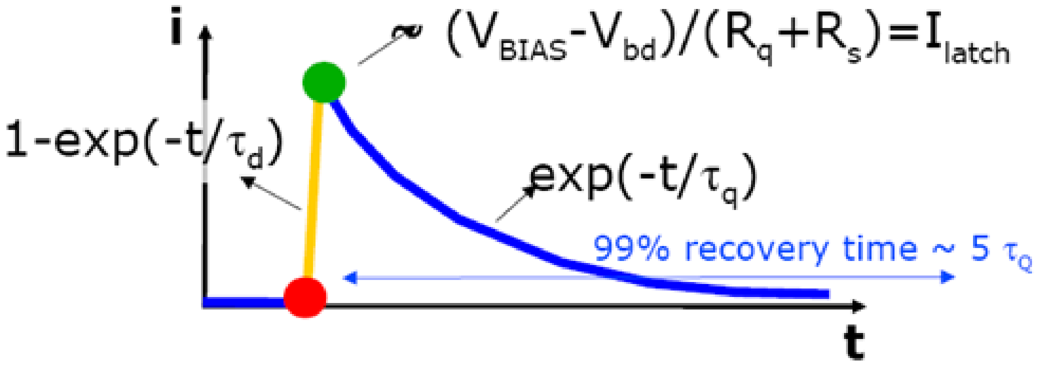
P_{01} = turn-on probability
probability that a carrier traversing the high-field region triggers the avalanche

P_{10} = turn-off probability
probability that the number of carriers traversing the high-field region fluctuates to 0

ON condition: avalanche triggered, switch closed
C_D discharges to V_{BD} with a time constant $R_S x C_D = \tau_{discharge}$,
at the same time the external current asymptotic grows to $(V_{BIAS}-V_{BD})/(R_Q+R_s)$

Passive Quenching

If R_Q is high enough the internal current is so low that statistical fluctuations may quench the avalanche



Haitz JAP 35 (1964)

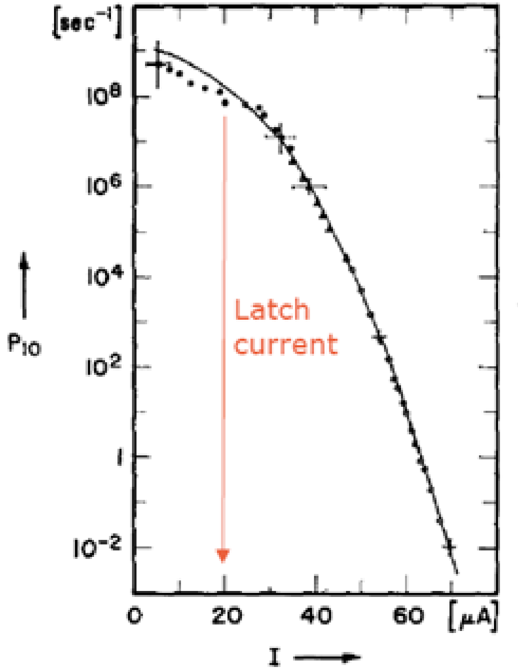


FIG. 2. Turnoff probability per second as function of pulse current.

The leading edge of the signal is much faster than trailing edge:

1. $\tau_d = R_s C_d << R_q C_d = \tau_q$
2. turn-off mean time is very short (if R_q is sufficiently high, $I_{latch} \sim 20\mu A$)

The charge collected per event is the area under the exponential which is determined by circuital elements and bias.

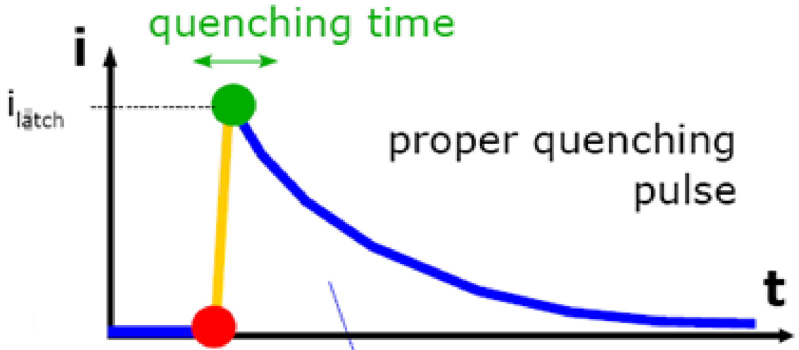
➡ It is possible to define a GAIN (discharge of a capacitor)

$$G = \frac{I_{max}\cdot\tau_q}{q_e} = \frac{(V_{bias}-V_{bd})\cdot\tau_q}{(R_q+R_s)\cdot q_e} = \frac{(V_{bias}-V_{bd})\cdot C_d}{q_e}$$

➡ Gain fluctuations in GM-APD are smaller than in APD essentially because electrons and holes give the same signal

Limits of Passive Quenching

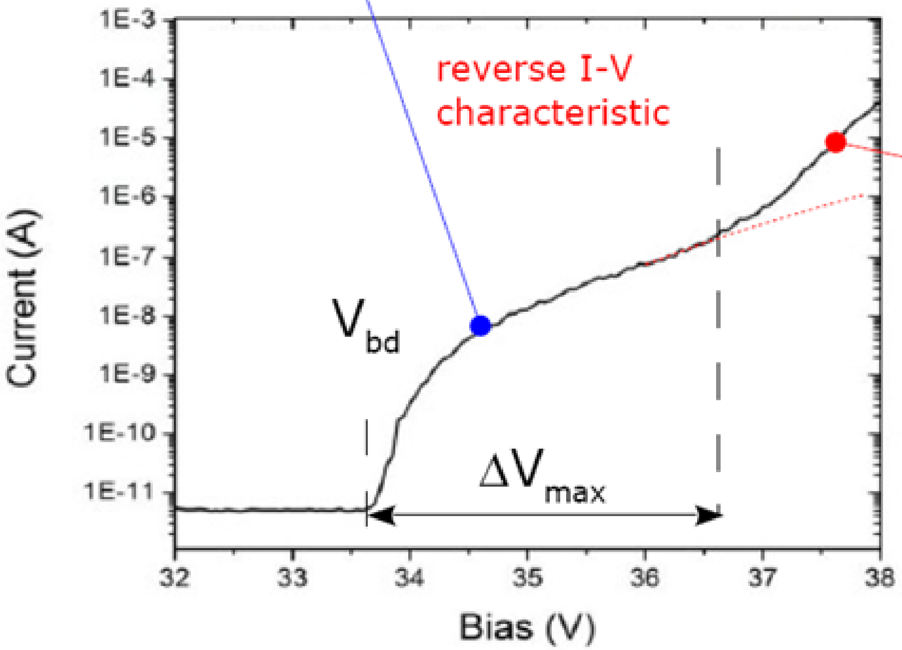
Proper value of quenching resistance Rq is crucial to let the internal current decrease to a level such that statistical fluctuations may quench the avalanche → sub-ns quenching time → crucial to have well defined gain

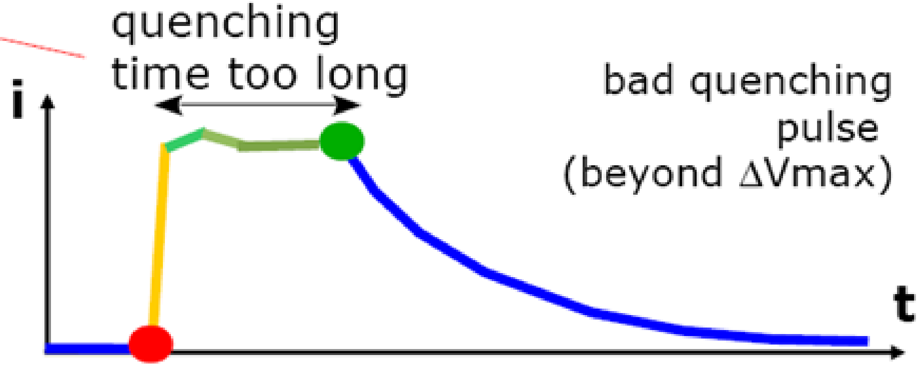


Given Rq the proper quenching regime is for ΔV in the range:

$$0 < \Delta V < R_q I_{latch}$$

where as a rule of thumb
$I_{latch} \sim 20\mu A \rightarrow \Delta V_{max} \sim$ a few Volts (typically)





Operation limit for ΔV

Operative ΔV limited by:
1) $I_{latch} \sim 20\mu A \rightarrow \Delta V < I_{latch} R_q$ (non-quenching regime)
2) Dark Count Rate (DCR) acceptable level ← PDE vs ΔV ← E field shape
3) V_{bd}^{edge} edge breakdown (usually some 10V above V_{bd})

A practical method for estimating the operative range (limited by effetc 1) is to measure the ratio R_I of the measured dark current I_D to the dark current I'_D calculated from the measured dark rate and pixel count spectra:

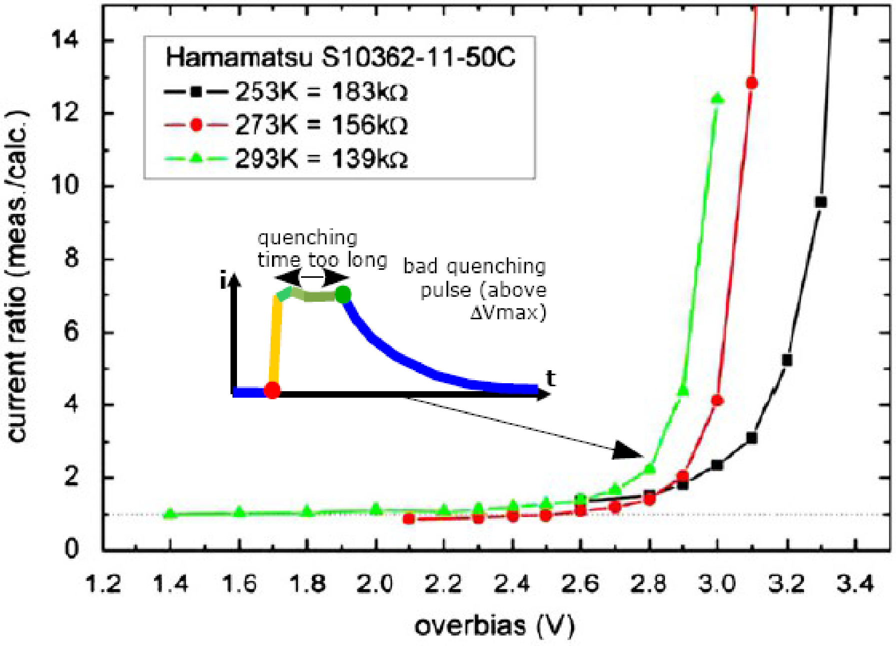
after Jendrysik et al NIM A 2011
doi:10.1016/j.nima.2011.10.007

$$R_I = \frac{I_D}{I'_D = DCR \cdot \bar{N} \cdot G \cdot q_e}$$

where $\bar{N}$ is the average N of fired cells

Non-quenching regime for values of ΔV when R_I deviates significantly from 1

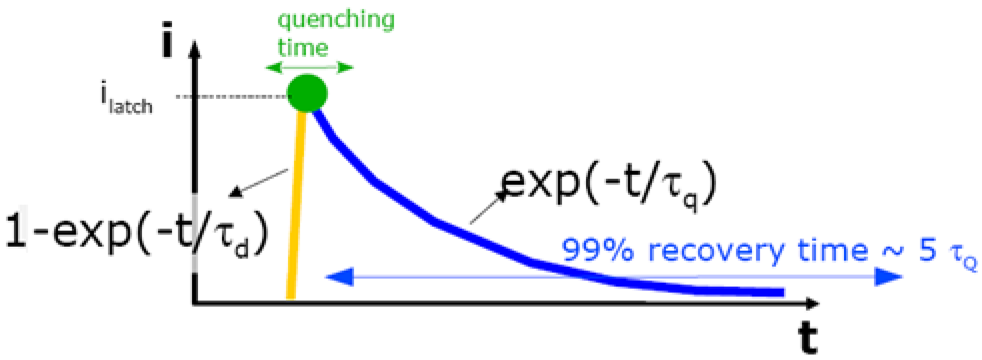
Jendrysik et al suggest
R_I=2 as reasonable threshold



Single GM-APD Model

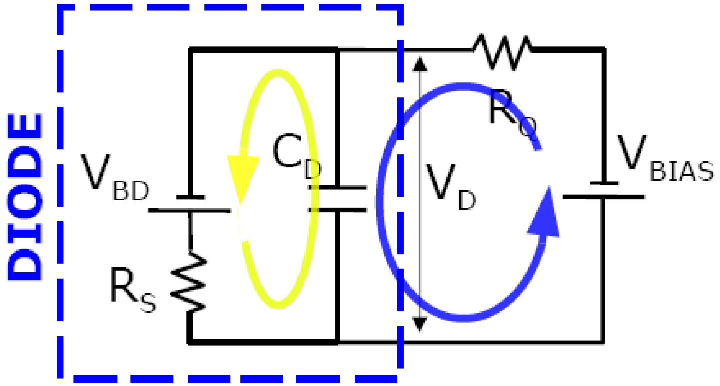
Fast discharge of the junction capacitor C_D via R_S given by the avalanche plasma serial resistance:

$\tau_d=C_D*R_S$, $C_D\approx0.1$pF, $R_S\approx1$kΩ, $\tau_d\approx100$ps



Slow recharge via R_q gives fall time of $\tau_q=C_D*R_q$, $R_q\approx150$-1000kΩ, $\tau_d\approx15$-100ns

currents **internal** / **external**



Recovery time is temperature dependent (R_q)

Rise time less temperature dependent (R_S)

Gain=$C_D*\Delta V$ (T independent)

Over-voltage $\Delta V=(V_{bias}-V_{BD})$

Si-PM Detailed Model

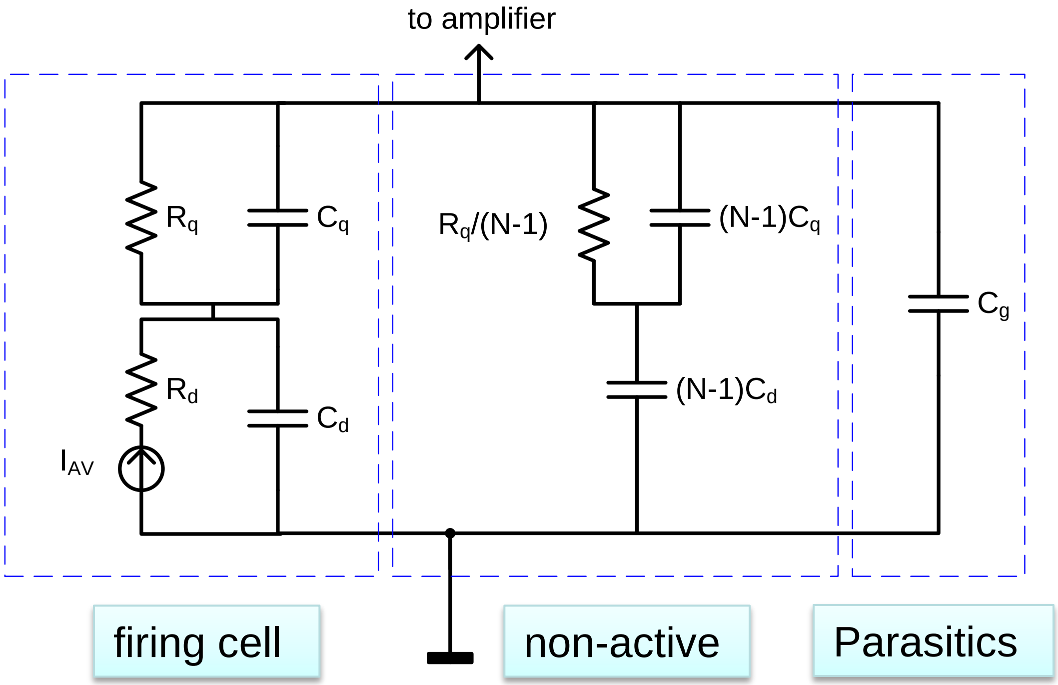
To understand the pulse shape of the SiPM, a high frequency equivalent circuit is used.

In a high frequency (AC) model the bias power supply is replaced as a short cut (bias voltage blocking capacitor (100nF to few μF) are short cuts for high frequency.

The resistor R_q has a parasitic capacitor C_q in parallel.

The non-active (not fired cells) (N-1) of the SiPM are connected.

The connectivity between pixels is modeled as a capacitor C_g which can be as important as the sum of C_d for small pixel devices.



The current due to the avalanche of a firing cell is very fast (100ps).
The parasitic capacitance Cq is supplying current into the non-firing cell.

The charge from the non-firing cells will generate a current flowing into the amplifier and decay with a time constant.

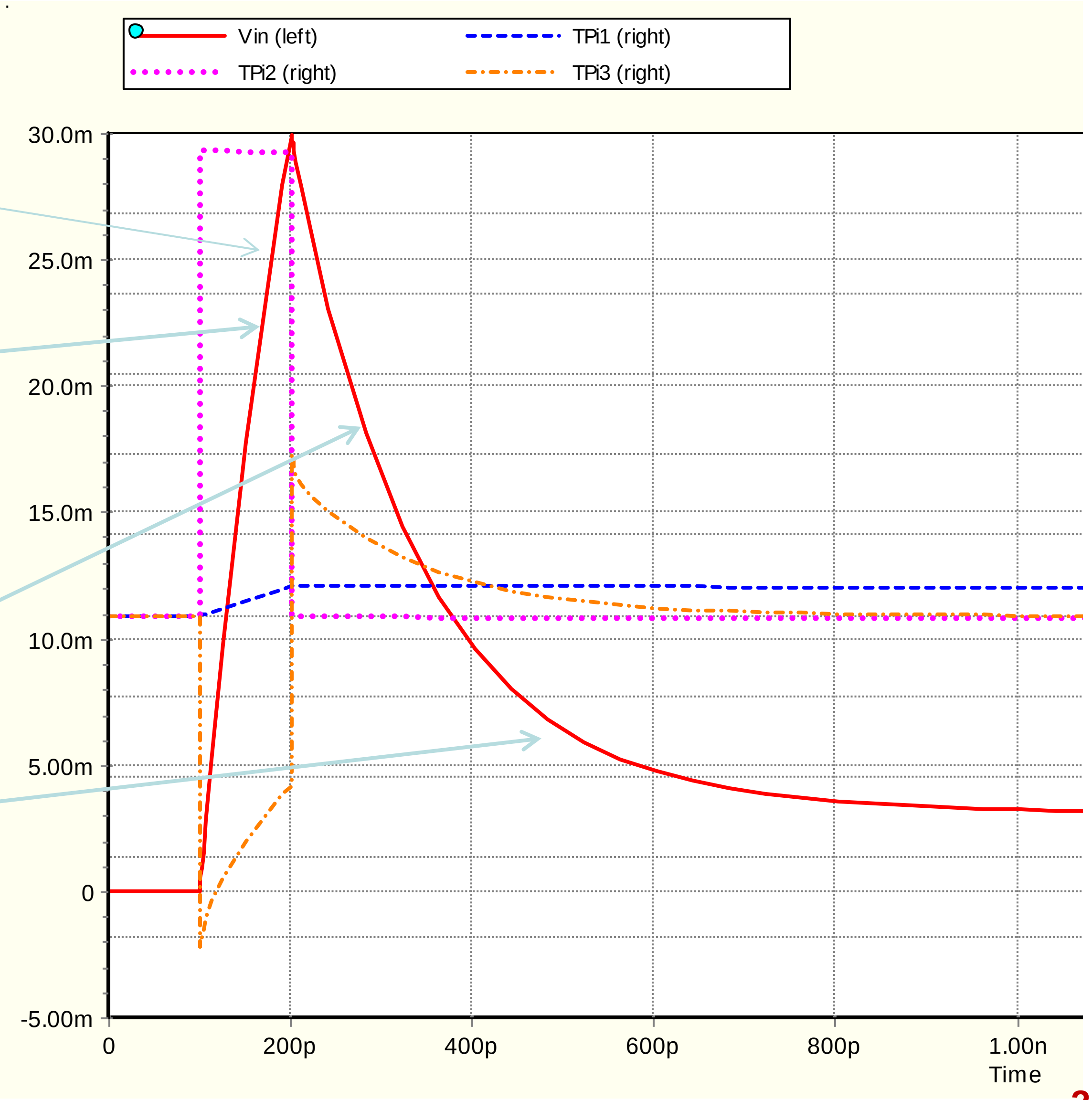
Si-PM Simulation Model

Output voltage on the SiPM

Very fast rise-time given with a time constant $\tau_{d(rise)}=R_d * (C_q+C_d)$

Fast decay time
$C_{tot}=C_d*N_{pix} + C_g$
(total detector capacitance)
$\tau_{fast(fall)}=R_{Load}*C_{tot}$

Slow decay time
$\tau_{slow(fall)}=R_q*(C_q+C_d)$



Si-PM Simulation Model

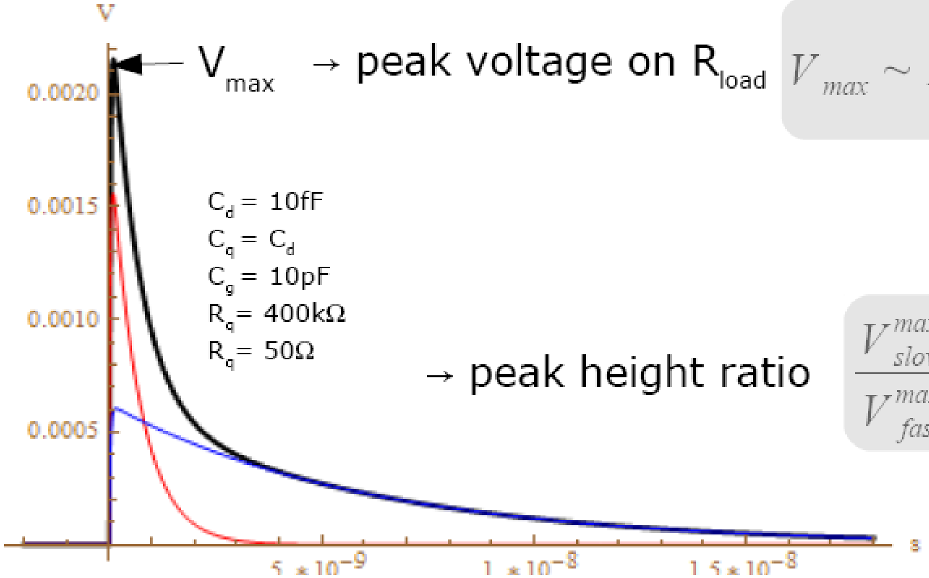
$$V(t) \simeq \frac{Q}{C_q + C_d}\left(\frac{C_q}{C_{tot}}\, e^{\frac{-t}{\tau_{fast}}} + \frac{R_{load}}{R_q}\frac{C_d}{C_q + C_d}\, e^{\frac{-t}{\tau_{slow}}}\right) = \frac{Q\,R_{load}}{C_q + C_d}\left(\frac{C_q}{\tau_{fast}}\, e^{\frac{-t}{\tau_{fast}}} + \frac{C_d}{\tau_{slow}}\, e^{\frac{-t}{\tau_{slow}}}\right)$$

Note: valid for
low impedance load
$R_{load} << Rq$

- $\tau_{fast} = R_{load}\ C_{tot}$
- $\tau_{slow} = R_q\,(C_q + C_d)$

→ gain $G = \int dt \frac{V(t)}{q_e R_{load}} = Q/q_e = \boxed{\frac{\Delta V\,(C_d + C_q)}{q_e}}$ independent of R_q

→ charge ratio $\boxed{\frac{Q_{slow}}{Q_{fast}} \sim \frac{C_d}{C_q}}$



V_{max} → peak voltage on R_{load} $\boxed{V_{max} \sim R_{load}\left(\frac{Q_{fast}}{\tau_{fast}} + \frac{Q_{slow}}{\tau_{slow}}\right)}$ dependent on R_q (increasing with $1/R_q$)

→ peak height ratio $\boxed{\frac{V^{max}_{slow}}{V^{max}_{fast}} \sim \frac{C_d\, C_{tot}\, R_{load}}{C_q^2\, R_q}}$ increasing with C_d and $1/R_q$

“Corsi” Parameters

model	size (mm^2)	# pixels	pixel size (μm^2)	C_d (fF)	C_q (fF)	C_g (pF)	C_t (pF)	R_q (kΩ)
S13360-1325	1.3 × 1.3	2668	25 × 25	18	4.9	10	60	784
S13360-1350	1.3 × 1.3	667	50 × 50	87	8.9	4	60	292
S13360-1375	1.3 × 1.3	285	75 × 75	179	42	0.1	60	393
S13360-3050	3.0 × 3.0	3600	50 × 50	85	16.8	19	320	301
S13360-6050	6.0 × 6.0	14400	50 × 50	83	10	137	1280	485
S12571-10	1.0 × 1.0	10000	10 × 10				35	
S12571-15	1.0 × 1.0	4489	15 × 15	7	3	2	35	1193

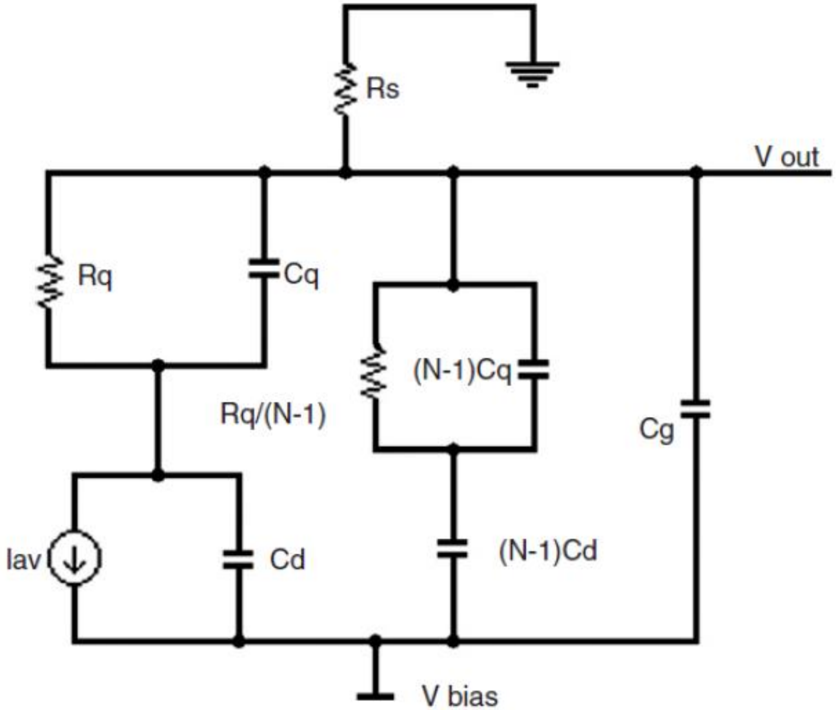
C_d – diode (pixel) capacitance

C_q – “quenching resistor” capacitance

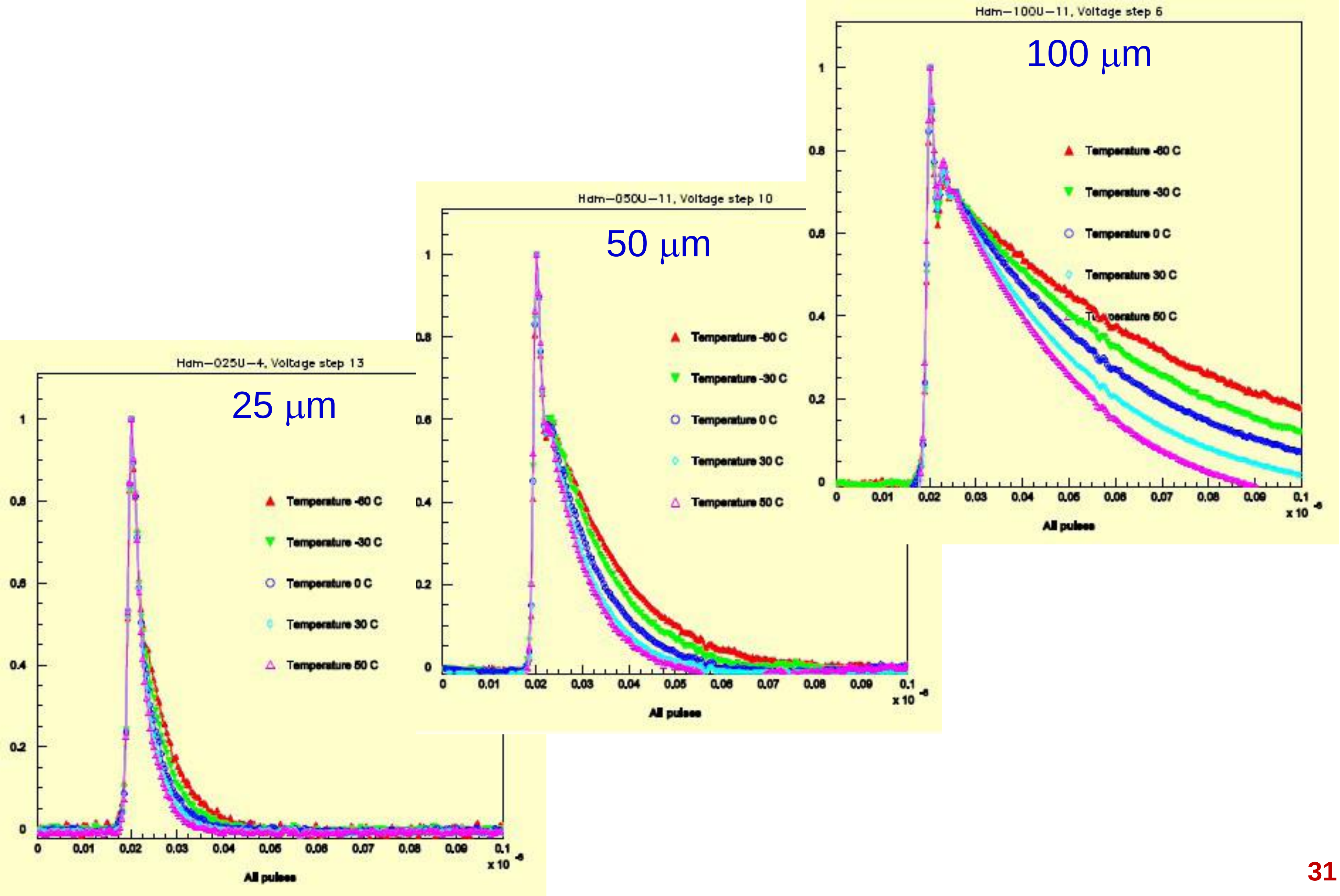
C_g – parasitic capacitance

C_t – terminal capacitance

R_q – quenching resistor



Pulse shape for different pixel sizes



Pulse Shape Analysis

Recording the dark noise pulses with an oscilloscope and determine with statistical analysis

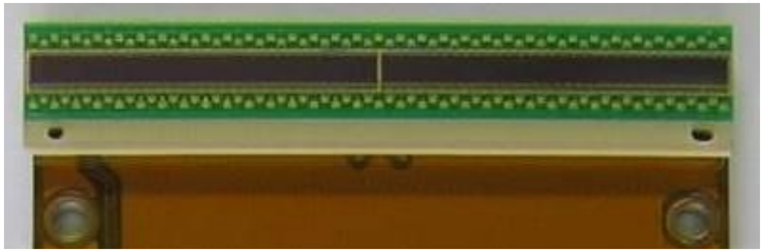
- Break down voltage V_{BD}
- Correlated noise probabilities: $p_{x\text{-talk}}$, p_{AP}, $p_{Dx\text{-talk}}$
- Time constants: τ_{slow}, τ_{rec}

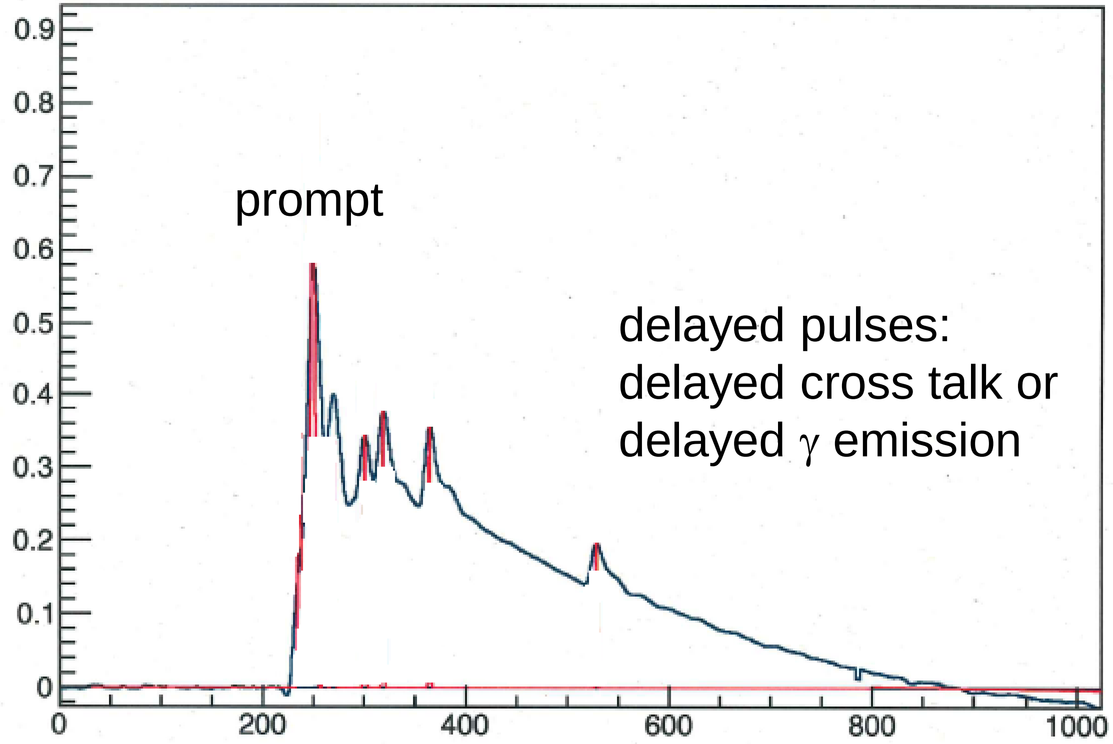
Keep T=const, scan ΔV

Difficulties:

- All parameters are T and ΔV depended
- Signal quality has to be high, use fast amplifier and avoid noise or long connections

Pulse Shapes

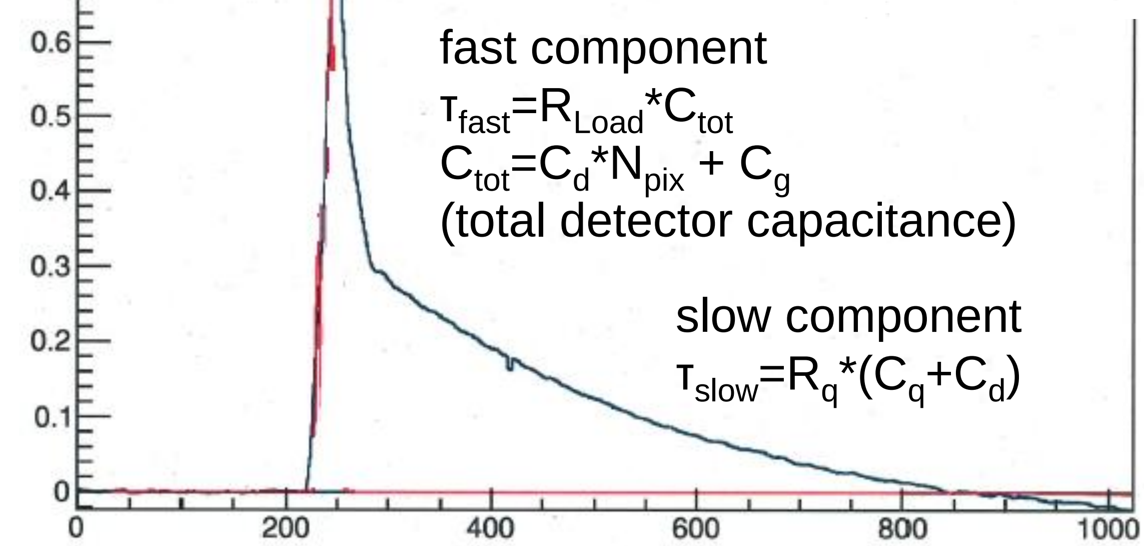




prompt

delayed pulses:
delayed cross talk or
delayed γ emission

200 ns



risetime
$\tau_r = R_d * (C_q + C_d)$

fast component
$\tau_{fast} = R_{Load} * C_{tot}$
$C_{tot} = C_d * N_{pix} + C_g$
(total detector capacitance)

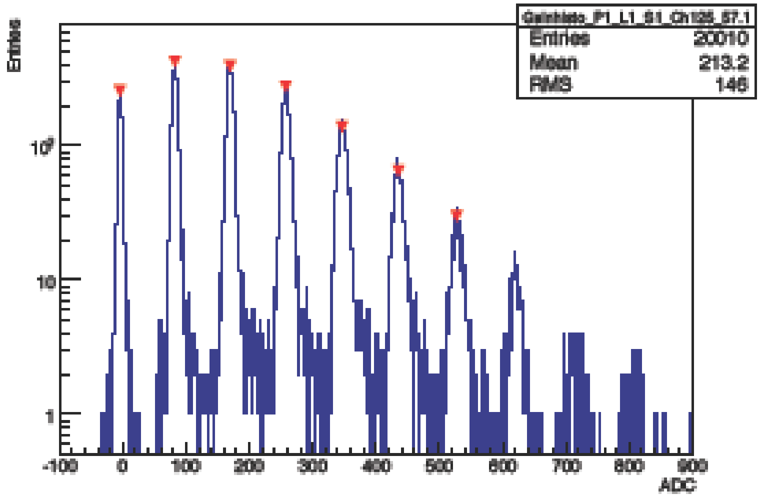
slow component
$\tau_{slow} = R_q * (C_q + C_d)$

200 ns

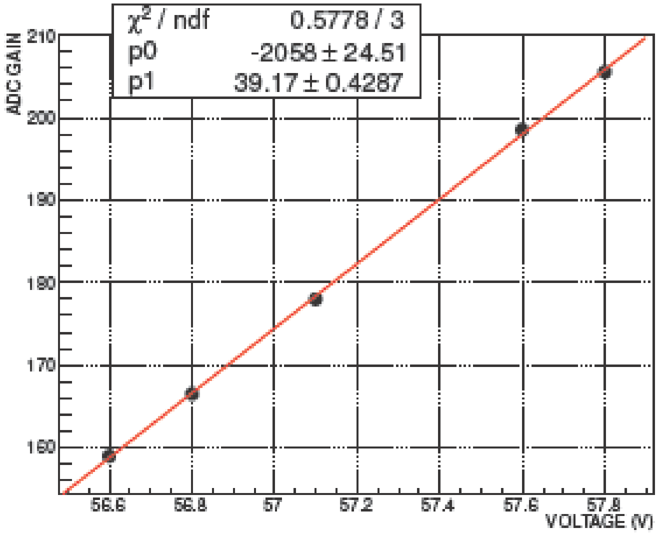
V_{BD} Measurement (1)

The most precise V_{BD} measurement can be obtained by recording a histogram of the photon peak charge integration at different V_{bias}. Calculation of the V_{BD} by a linear fit and extrapolation to zero gain.

signals measured with an ADC



Fit of gain measured at different V_{bias}

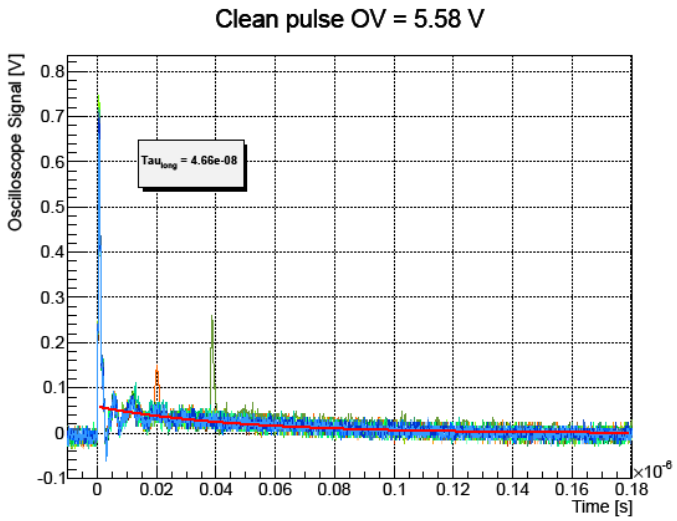


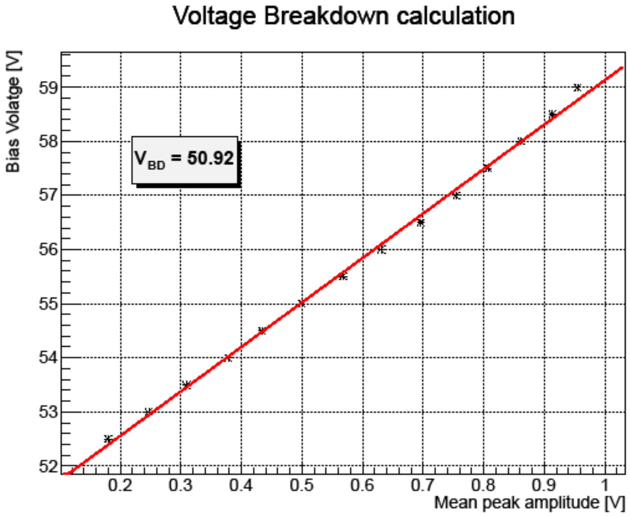
V_{BD} Measurement (2)

Use pulse analysis from oscilloscope recording. Record the dark pulses at different V_{bias}. Measure the peak amplitude and calculate the mean, reject all pulses with double (x-talk) amplitude. Fit the amplitude and calculate the V_{BD} at zero amplitude. This calculation is not as precise as 1 but it allows to calculate ΔV at the same time as the correlated noise probabilities at any temperature.

Extract the mean peak amplitude at different V_{bias}.

Calculate V_{BD} from the linear fit.





V_{BD} Measurement (3)

Applying a reverse bias voltage allows to determine the breakdown voltage (V_{BD}). A spread due to manufacturing difference is expected for V_{BD}.

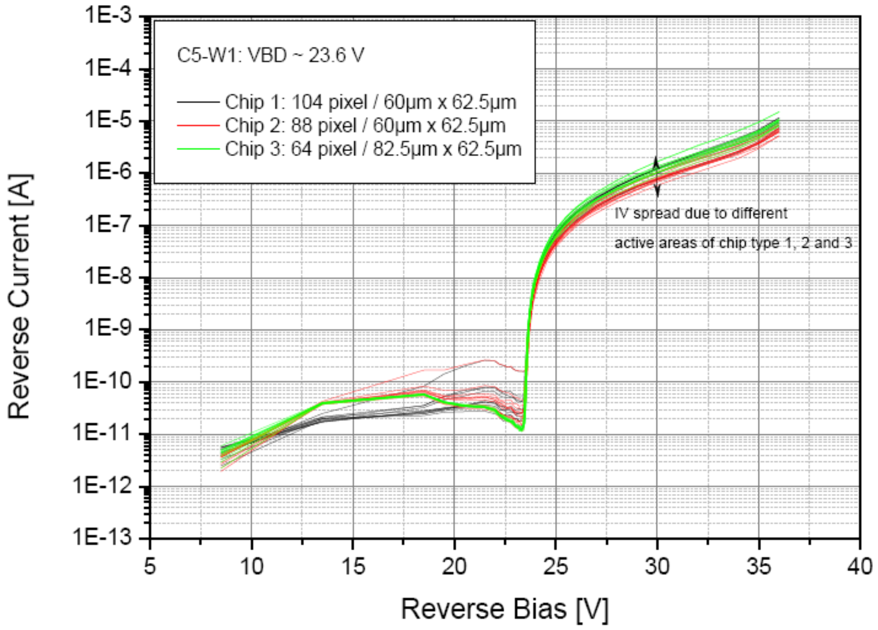
V_{BD} is the voltage where amplification starts and can be extracted from the I-V curve. Already by "eye" one expects V_{BD} to be where the steep slope sets in.

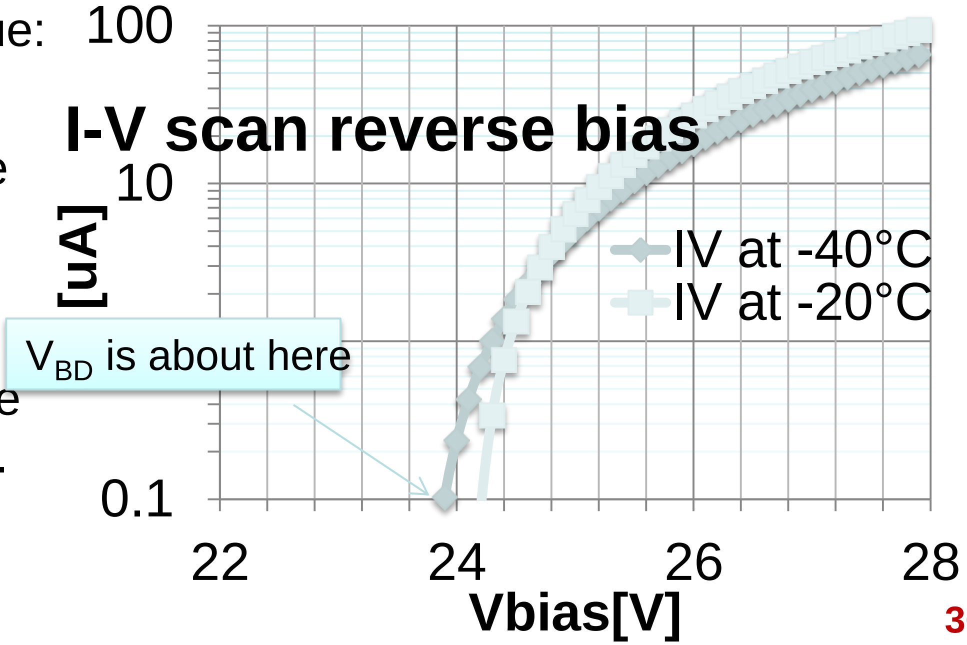
Calculation can be done based on the assumption that an exponential law for the current and the over-voltage is true:

$$I = \alpha (V_{bias} - V_{BD})^n$$

Differentiation allows to define V_{BD} where the calculated quantity is zero. The quantity follows a linear law and can easily be fitted. The increase of this quantity below V_{BD} is due to a leakage current not produced by amplification.

$$\left[\frac{d(\log I)}{dV}\right]^{-1} \propto (V - V_{BD})$$

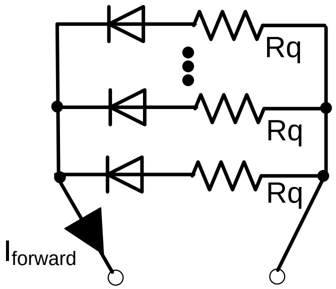




V_{BD} is about here

I-V Characteristics: Forward Biasing

Applying a voltage in the forward direction shows the diode forward characteristics overlapped with the serial quench resistor. All pixels are connected in parallel and therefore the resistor characteristics is the parallel resistor of all quench resistors.



The resistor R_q is temperature dependent. With lower temperature the resistance increases and therefore change of recovery time, after-pulse and pulse shape are the consequences.

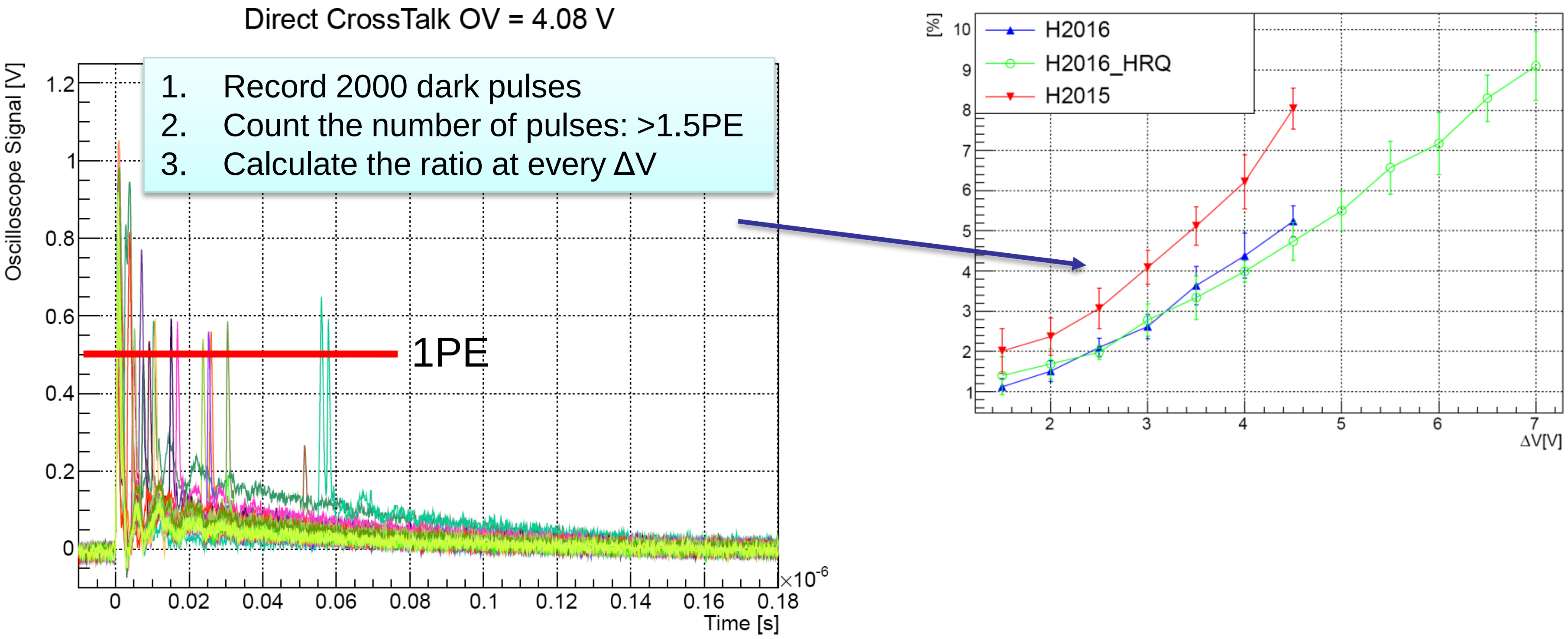
Ca

The quench resister can be calculated from the slope of the curve.
$(R_q/N_{pix})=\Delta U/\Delta I = 2.0V/1.4\ mA = 1.43k\Omega$
➔ $R_q = 137k\Omega$ with $N_{pix} = 96$

Correlated noise probabilities: $p_{x\text{-talk}}$

X-talk events (pulses) can be selected using a time window and a threshold on the amplitude.

Usually the x-talk threshold was set to 1.5 PE. For fast detectors, a small time difference between the original dark pulse in one pixel and the x-talk induced pulse from a neighbouring pixel can be delayed by a few 100ps with the result that the amplitude of the pulse never reaches 1.5 PE.



1. Record 2000 dark pulses
2. Count the number of pulses: >1.5PE
3. Calculate the ratio at every ΔV

Dark Noise Rate (DCR)

If a thermal fluctuation creates an e-h pair in the depletion region, a charge carrier can enter the avalanche region and trigger an avalanche that is indistinguishable from one resulting from photon absorption.

A less likely but still possible scenario is when a pair is thermally produced in the avalanche region itself. In either case, the resulting pulse is referred to as dark noise. The occurrence of dark noise is random in time and uncorrelated to the pulses resulting from photon absorption.

The rate of dark noise depend on:

- Over-voltage
- Temperature
- Pixel size
- Detector area

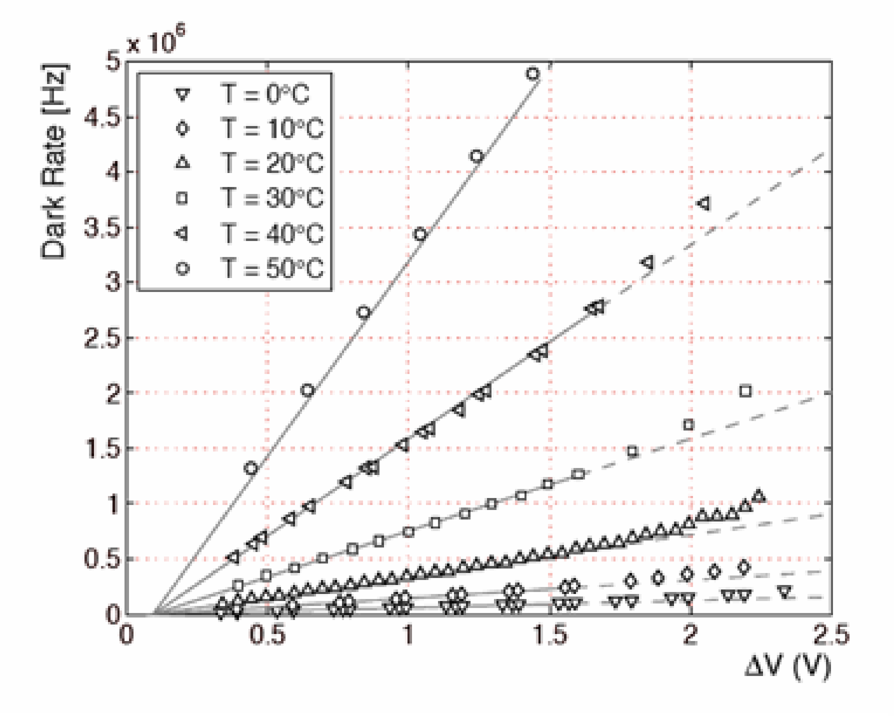
DCR is the number of pulses higher than 0.5 PE per second. It is a frequency.

To compare DCR for different devices the following bench mark point was established in the community:
Threshold: 0.5PE
Temperature: 25°C
Area: 1mm^2

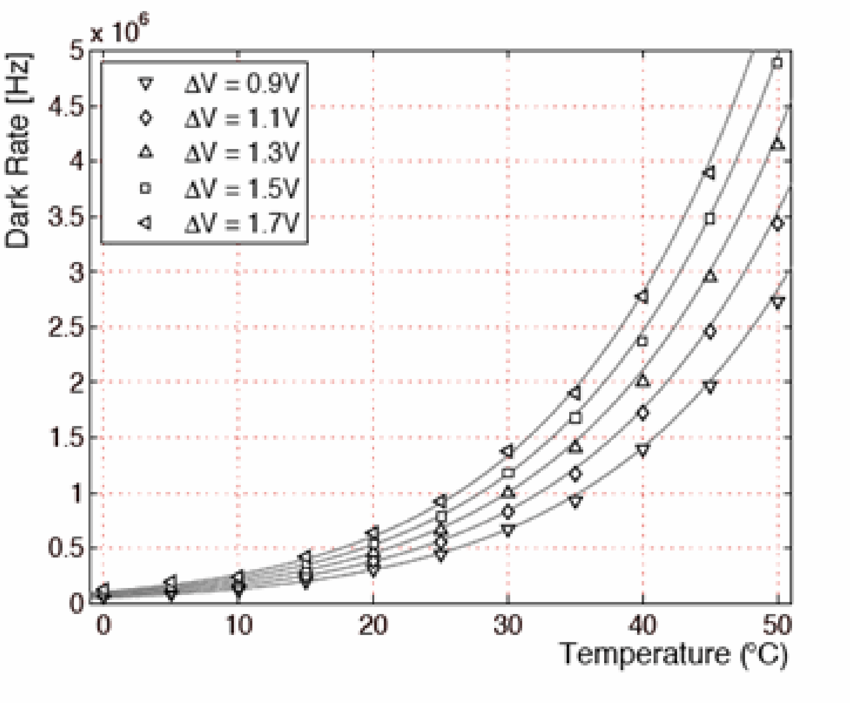
DCR (ΔV, T)

A strong over-voltage and temperature dependence for the DCR is observed.

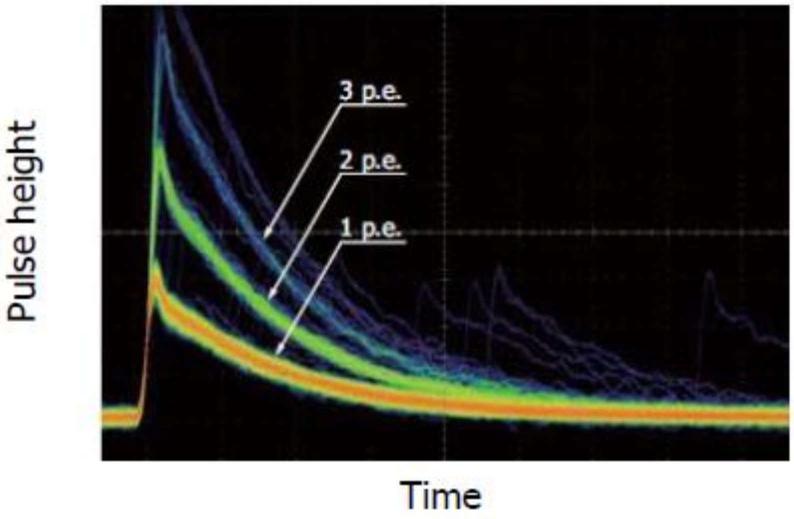
Linear increase with ΔV in limitted range

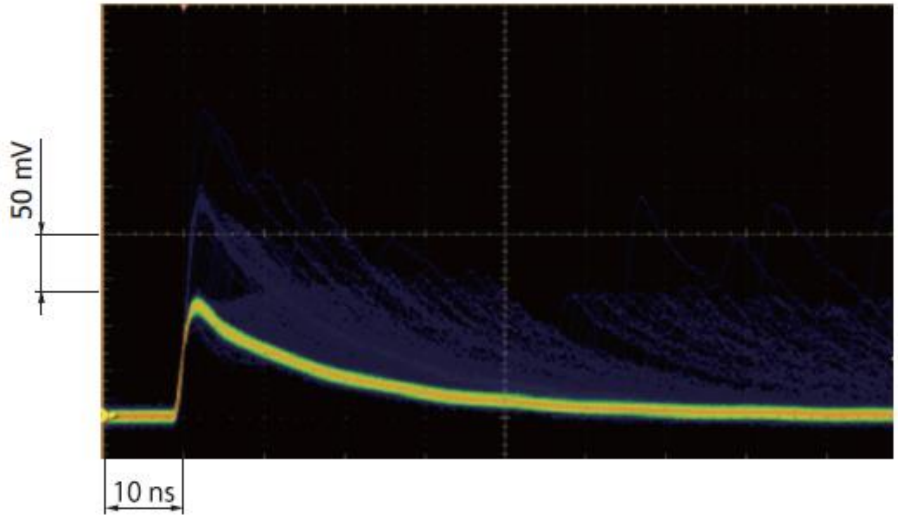


Exponential increase with T



Pixel to Pixel Cross-talk





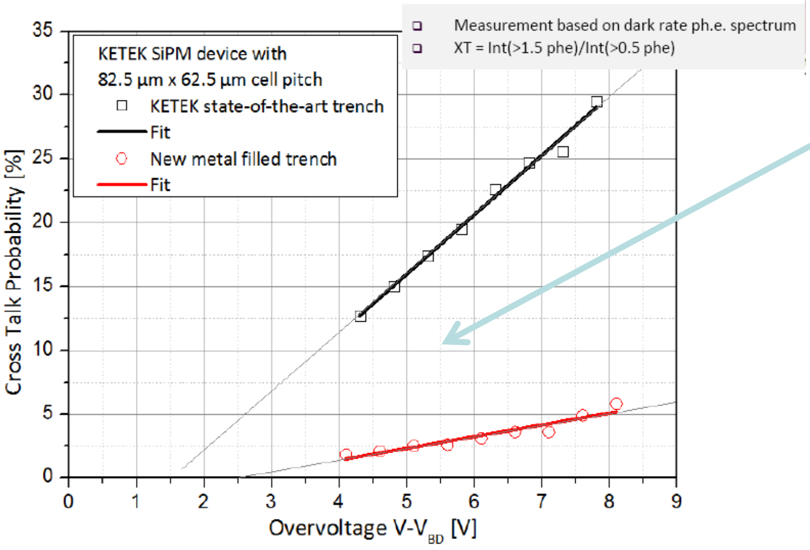
Si-PM with large cross-talk

old generation
series 12571-2

Si-PM with low cross talk

new generation
series 13360

Pixel to Pixel Cross-talk

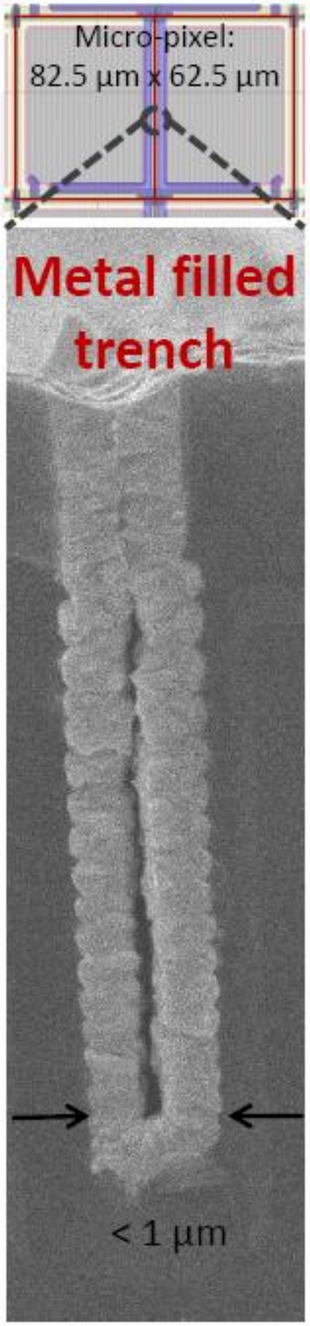


- Measurement based on dark rate ph.e. spectrum
- XT = Int(>1.5 phe)/Int(>0.5 phe)

x-talk is given as a probability in percent. It is proportional to the gain. Therefore:

$$P_{x-talk} \propto \Delta V$$

$$P_{x-talk} \propto Gain \propto A_{pix}$$

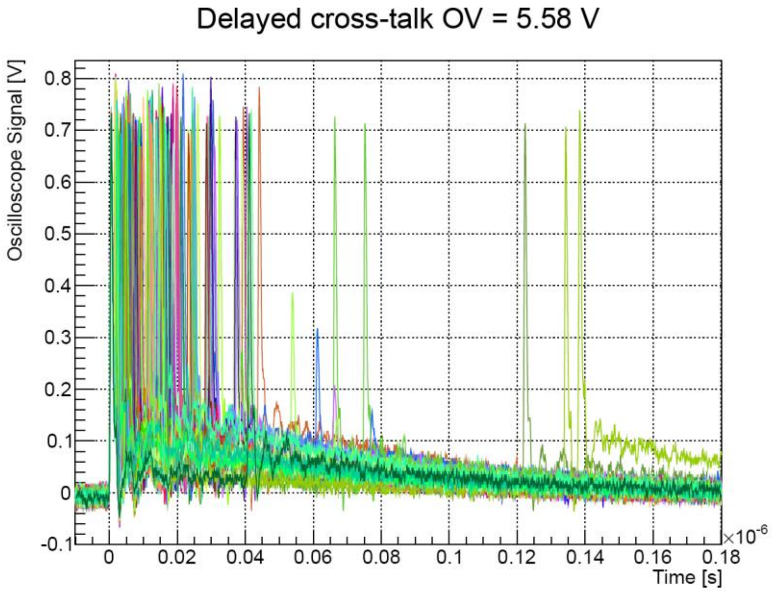


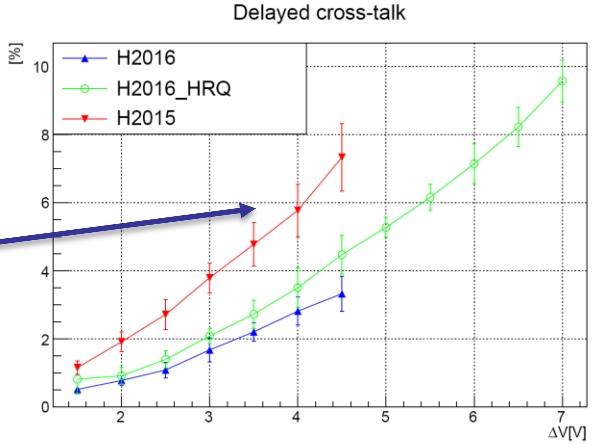
- X-talk is produced by photons crated during the avalanche in one pixel or electrons migrating through the bulk.
- Photons traversing the boundary of a pixel can produce almost instantaneously a secondary avalanche in a neighboring pixel.
- Electrons migration produces delayed x-talk because they drift.
- X-talk is correlated noise! DCR and photon induced avalanches generate both x-talk.
- X-talk is indistinguishable from signal or DCR!
- To reduce the optical x-talk, the active area of each pixel has to be isolated such that the photons produced in the avalanche region are contained within a pixel.

Delayed x-talk: $p_{Dx\text{-}talk}$

Delayed x-talk occurs when free charge carriers produced in the avalanche in one pixel drift via the bulk into a neighbouring cell. This phenomenon is not seen for all detector manufactures. It has been observed in Hamamatsu devices. The pulses can be selected using the a threshold of 0.85 PE.

1. Record 2000 dark pulses
2. Count the number of pulses: >0.85PE
3. Calculate the ratio at every ΔV



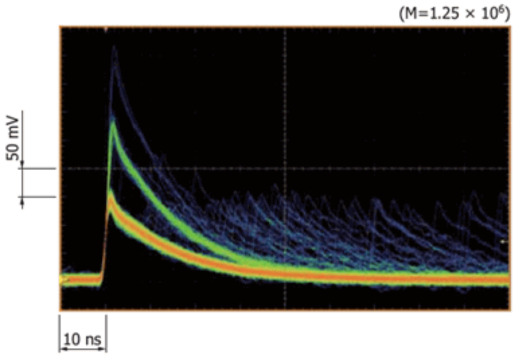


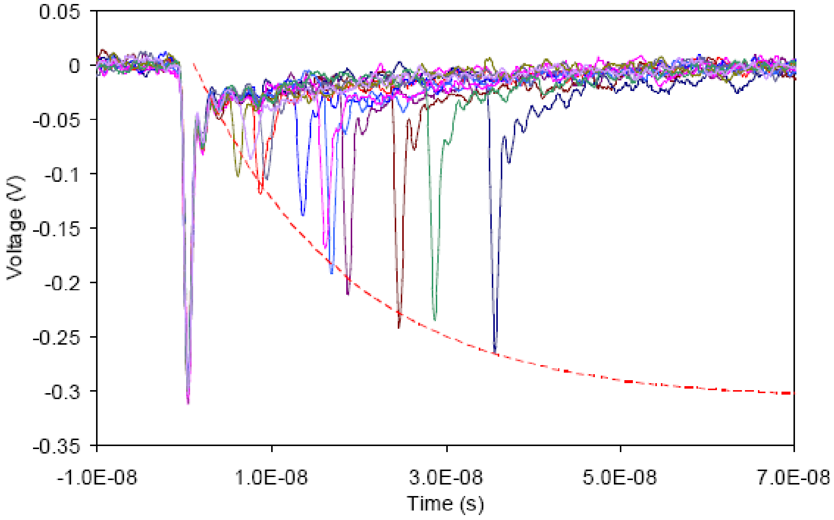
Note: Delayed x-talk and random dark pulses can not be separated.
With the DCR the expected number of random dark pulses can be calculated.
Example: f_{DCR}=100kHz, t_{window}=180ns, N_{pulses}=1000, N_{dark}=18

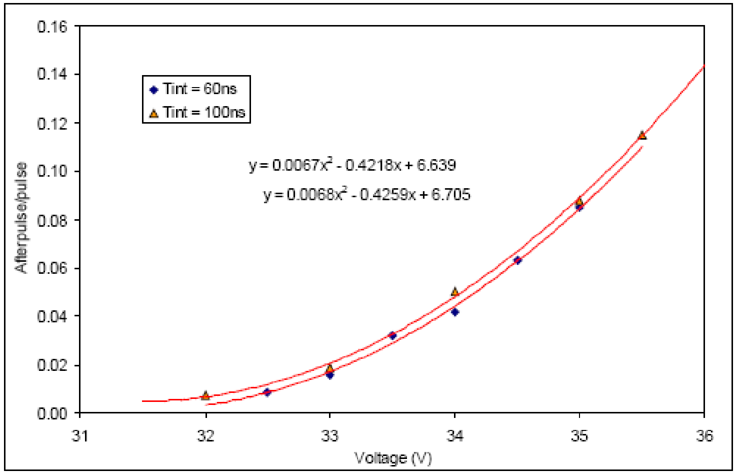
After-pulsing

After-pulsing occurs when a secondary avalanche forms in a pixel that is recovering from a discharge.

The secondary avalanche is due to the release of trapped charge — electrons — at some time after the primary avalanche.







Events with after-pulse measured on a single micropixel.

After-pulse probability increases with the bias

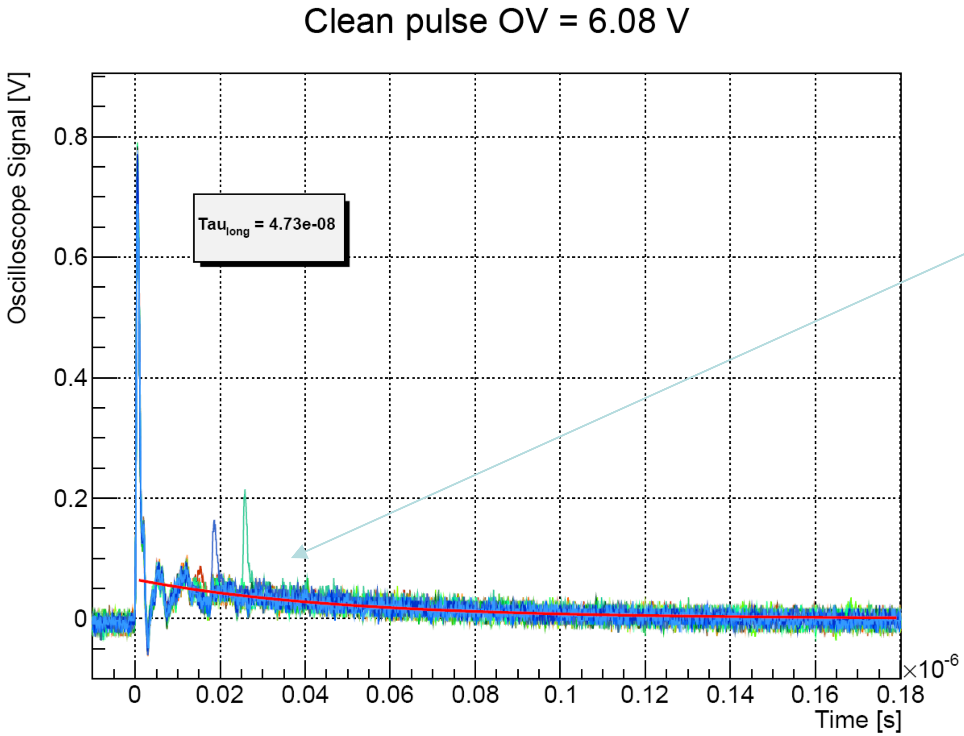
Recovery time

The amplitude of the after-pulse increases with the delay and reaches the full amplitude after a characteristic recovery time

τ_{slow}

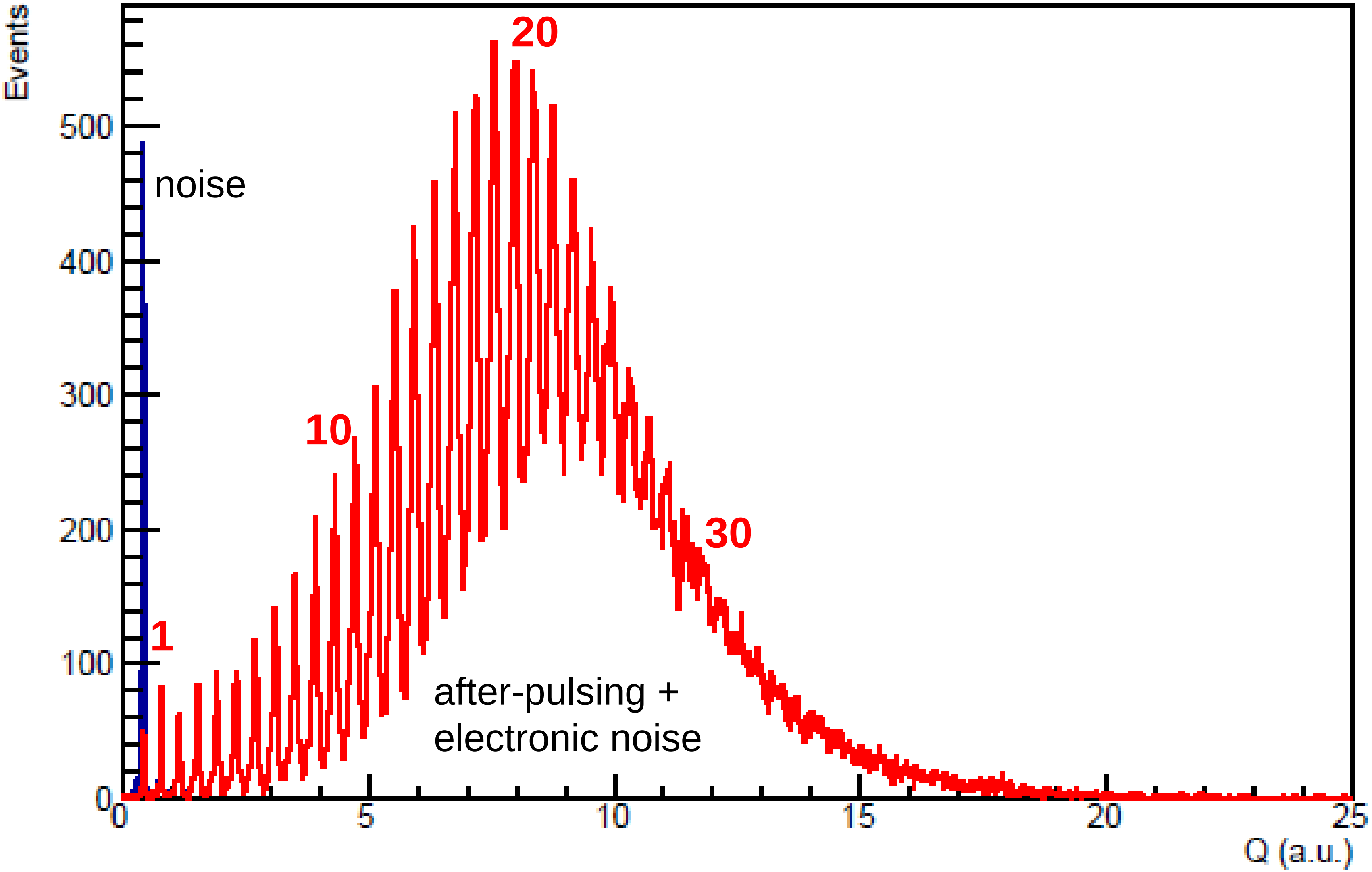
Clean pulses to extract τ_{slow}

Reject all pulses which deviates from single dark pulses to all for a fit of the slow time constant.
The time constant changes with T , small variations with ΔV can be observed.



$$\tau_{slow} = R_q(C_q + C_d)$$

Single Photon Counting



Linearity: The distance between the peaks is constant !

Limited Linear Range

The SiPM arrays have a limited number of pixels – for high photon flux, several photons hit one pixel at the same time causing saturation or non-linear response. To calculate the number of incident photons from the number of hit pixels the following calculation can be done.

- $N_{\text{fired pix}}$: Mean number of fired pixels
- N_{photon} : Number of incident photons
- N_{tot} : Total number of pixels on the detector
- dN_{photon} : increase of incident photons
- $dN_{\text{fired pix}}$: increase of fired pixels

$$dN_{fired\ pix} = dN_{photon} \cdot PDE \cdot \left(1 - \frac{N_{fired\ pix}}{N_{tot}}\right) \Rightarrow \int \frac{dN_{fired\ pix}}{1 - \frac{N_{fired\ pix}}{N_{tot}}} = \int PDE \cdot dN_{photon}$$

$$\boxed{N_{fired\ pix} = N_{tot}\left(1 - e^{\frac{-N_{photon} \cdot PDE}{N_{tot}}}\right)}$$

Best working point is when $N_{\text{fired pix}} \ll N_{\text{tot}}$

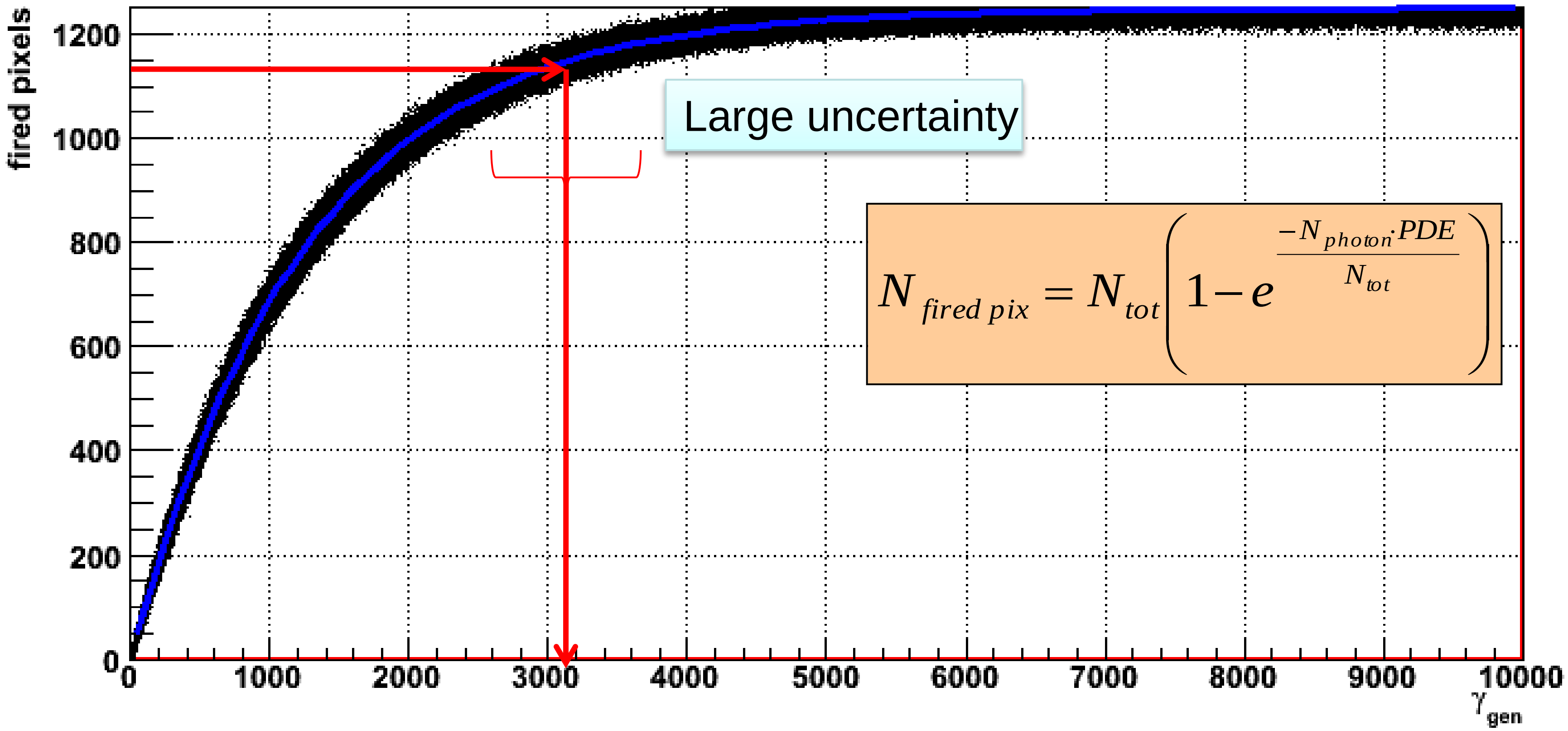
Limited Linear Range

Response simulated for SiPM with N_{tot}=1280 (Monte-Carlo)

A liner response ensures small uncertainty on the measured quantity!

$$for: \frac{N_{fired\ pix}}{N_{tot}} = 0.1 \rightarrow \frac{N_{photon} \cdot PDE}{N_{tot}} = 0.105$$

$$for: \frac{N_{fired\ pix}}{N_{tot}} = 0.5 \rightarrow \frac{N_{photon} \cdot PDE}{N_{tot}} = 0.693$$



Large uncertainty

$$N_{fired\ pix} = N_{tot}\left(1 - e^{\frac{-N_{photon} \cdot PDE}{N_{tot}}}\right)$$

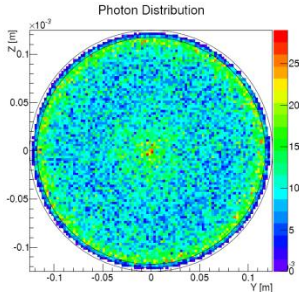
Saturation

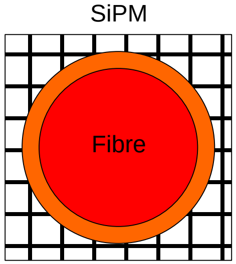
In applications with fibers N_{tot} used is not well known. Only the illuminated pixels have to be counted!

Exit angle of fiber light can be used to spread light over a large detector surface (1mm fibre and 3x3mm^2 square detector, distance between fibre and detector surface e.g. 2mm).

Use small pixel size to reach high dynamic range but small pixel size and high PDE is difficult to achieve.

If the pixel recovery time is of the same order as scintillation light emission time constant, pixels can partially or completely recover. Each pixel can fire several times for one light pulse detectio (CMS HCAL uses this). Short recovery time and small pixels can be achieved ($\tau_q = C_D * R_q$).





Partial used because of non-homogenous light emission at the fibre edge N_{tot}=??? !

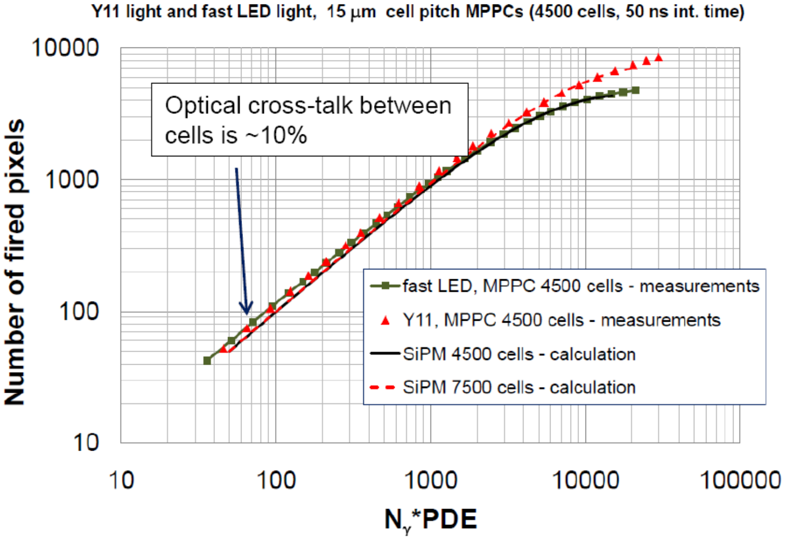


Photo Detection Efficiency

Photo-Detection Efficiency (P.D.E.) :

combined probability to produce a photoelectron and to detect it

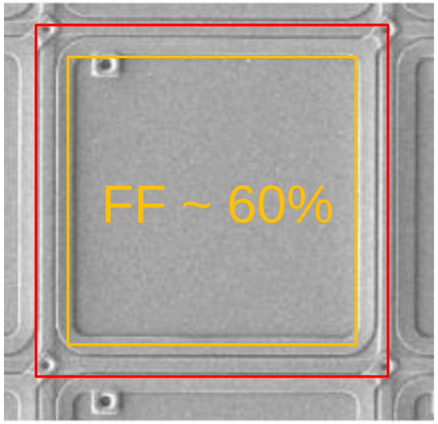
$$PDE(\lambda, \Delta V, T, \ldots) = \varepsilon_{geom} \cdot QE(\lambda) \cdot P_{trig}(\lambda, \Delta V)$$

$DV = V_{bias} - V_{BD}$

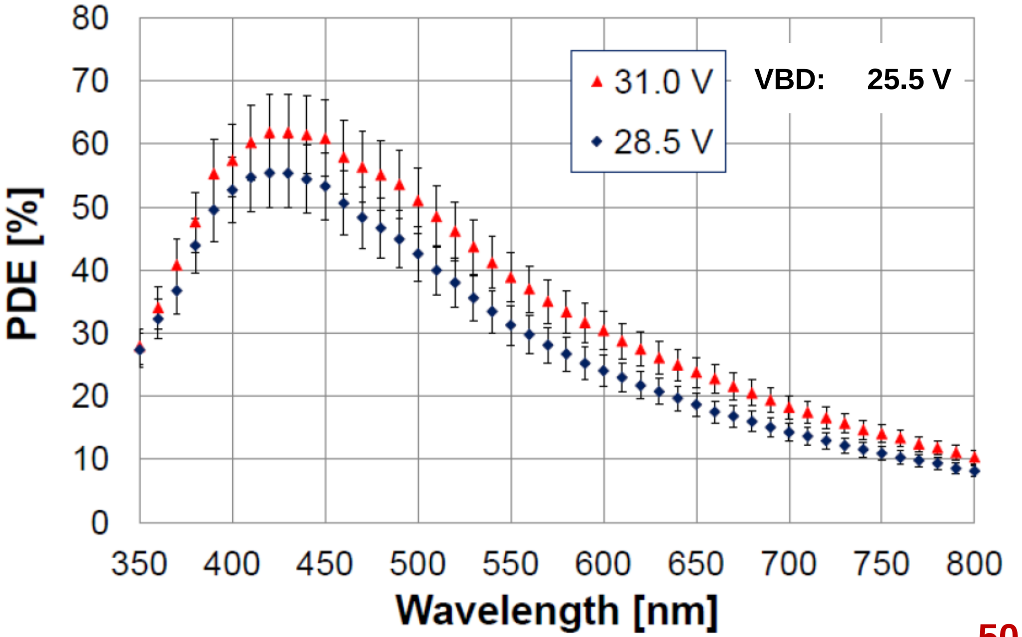
ε_{geom}: fill factor = sensitive area / total area

P_{trig}: avalanche triggering probability, higher the ΔV, higher the P_{trig} (> 95%)

QE: quantum efficiency

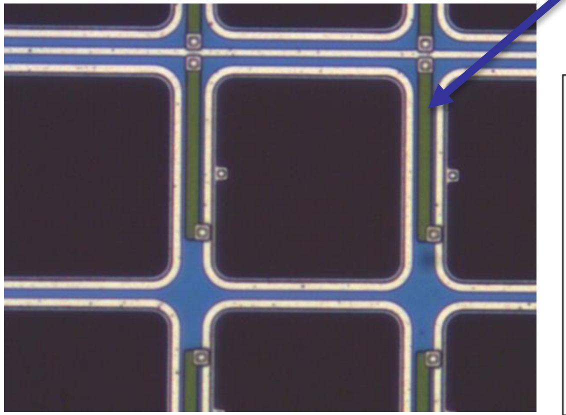


50 μm cell

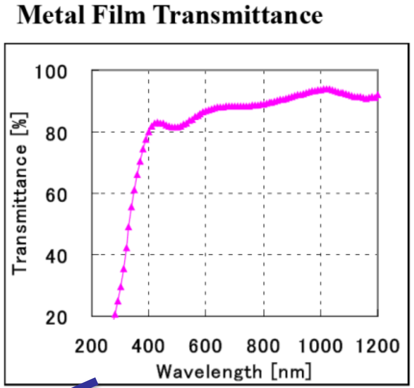


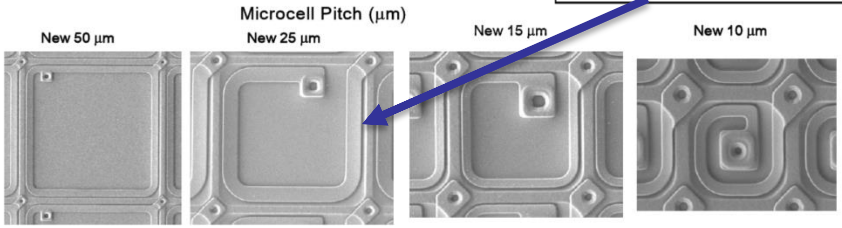
Maximize PDE (ε_{geom})

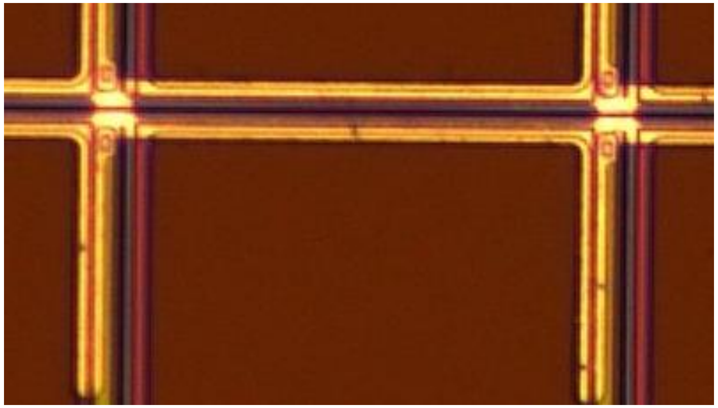
- Metal lines, signal connections
  - Minimize the number of pixels!
- Poly-Silicon quench resistors can be replaced by transparent thin metal film
- Use optimized trench structures using very high aspect ratio trenches



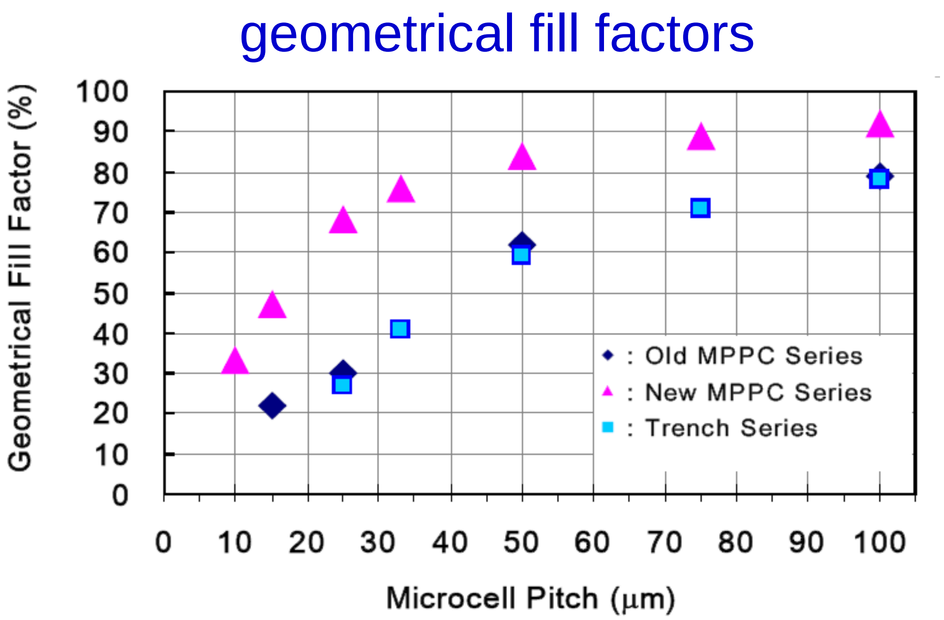
Polysilicon R_q

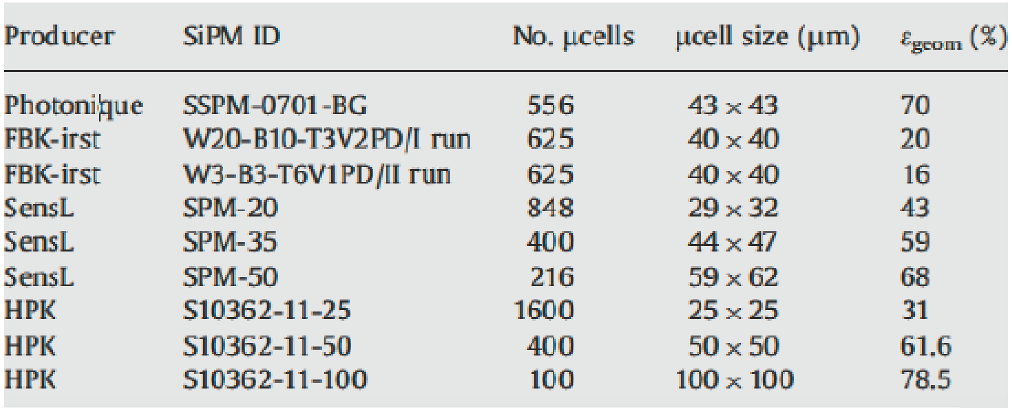






Maximize PDE (ε_{geom})

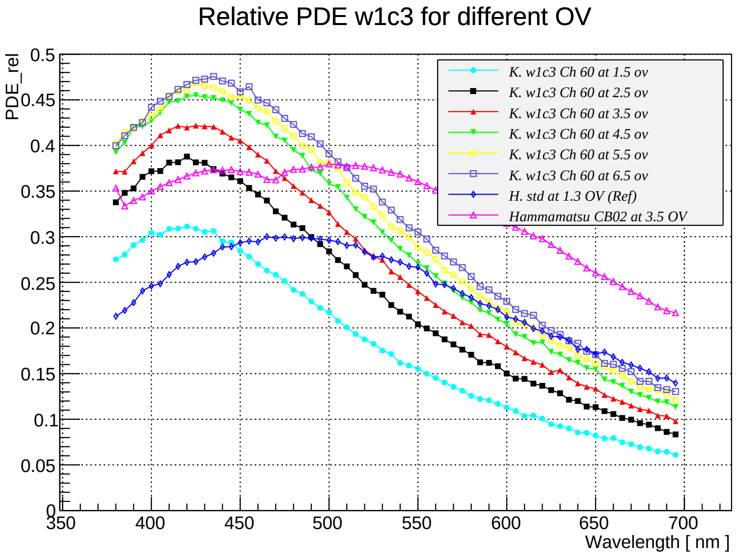




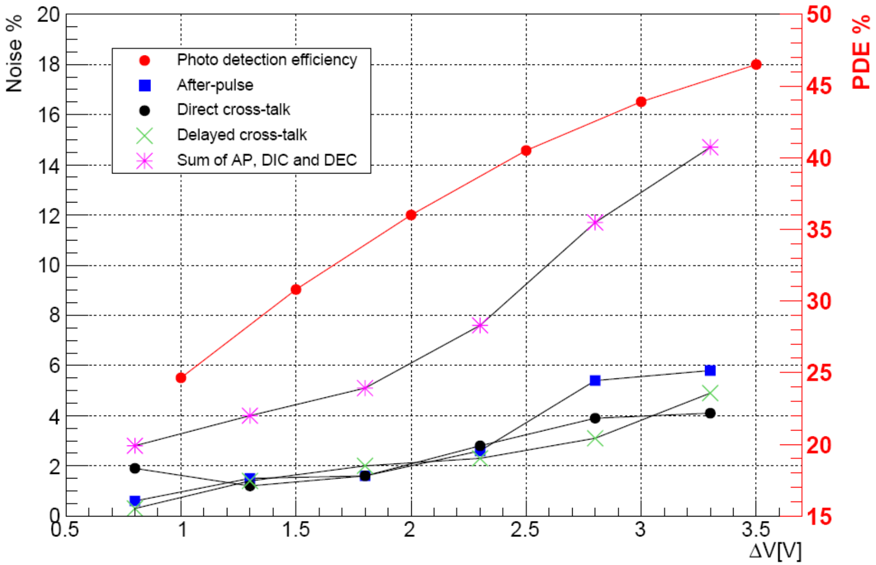
Producer	SiPM ID	No. μcells	μcell size (μm)	ε_{geom} (%)
Photonique	SSPM-0701-BG	556	43 × 43	70
FBK-irst	W20-B10-T3V2PD/I run	625	40 × 40	20
FBK-irst	W3-B3-T6V1PD/II run	625	40 × 40	16
SensL	SPM-20	848	29 × 32	43
SensL	SPM-35	400	44 × 47	59
SensL	SPM-50	216	59 × 62	68
HPK	S10362-11-25	1600	25 × 25	31
HPK	S10362-11-50	400	50 × 50	61.6
HPK	S10362-11-100	100	100 × 100	78.5

HPK	H2016_HRQ	Trench Thin Film	LCT05 104	60x60μm	72%
HPK	H2014	No trench Polysilicon	96	60x60μm	62%

Maximize PDE (ε_{AT})



Optimize thickness of photon absorption region (additional epi layer for KETEK) shows slightly green shift



Increase ΔV , trenches allow higher ΔV without excessive correlated noise

Si-PM Characteristics

Advantages

☺ high gain (10^5 - 10^6)
☺ work with low voltage (< 100V)
☺ low power consumption (< 50μW / mm^2)
☺ fast (timing resolution ~ 100 ps RMS for single photons)
☺ insensitive to magnetic field (tested up to 10 T)
☺ high photon detection efficiency (30 - 50% blue-green)
☺ excess noise factor close to 1
☺ compact and rugged
☺ tolerate accidental illumination
☺ cheap: produced n standard CMOS process

Possible drawbacks

☹ high dark count rate (DCR) at room temperature
  10 kHz – 100 kHz / mm^2
  thermal carriers, cross-talk, after-pulses
☹ temperature dependence (but relatively small)
  V_{BD}, G, R_q, DCR
☹ nonlinear response against input light (saturation)

How Do We Make All This ?

Ultraclean silicon processing to avoid any contaminations causing an increase of the dark count rate:

1. p^+ or n^+ Si-wafer substrate (~300 μm)
2. grow a p or n ultraclean epitaxial layer (few μm)
3. film formation - oxidation (SiO_2 layer)
4. implant the p – n^+ (or n – p^+) junctions (diode)
5. test the wafer (I-V)
6. metalization (add Al electrodes), add quenching resistors (poly-silicon), etc.
7. add the antireflective coating (Si_3N_4)
8. wafer dicing (cut out the single sensors)
9. bonding and packaging

Si Technology

Semiconductor industry has developed rapidly (13% per year over the last 20 years) and now represents a 3×10^{11} \$ market, which turns out to be ~ 10% of the world GDP (Gross Domestic Product).

Particle physics cannot influence this development for silicon detector fabrication, can - on the contrary - adapt mainstream semiconductor technology (which are developed, tuned, checked using huge investments) to the research needs.

There is one main difference → the use of high resistivity silicon wafers
(to allow for large depletion depths)

The rest of the production process consists of lithographic steps, etching, doping profiles, implantation and diffusion which are similar (and use similar tools) to the semiconductor industry but have been tuned for our needs over decades of development

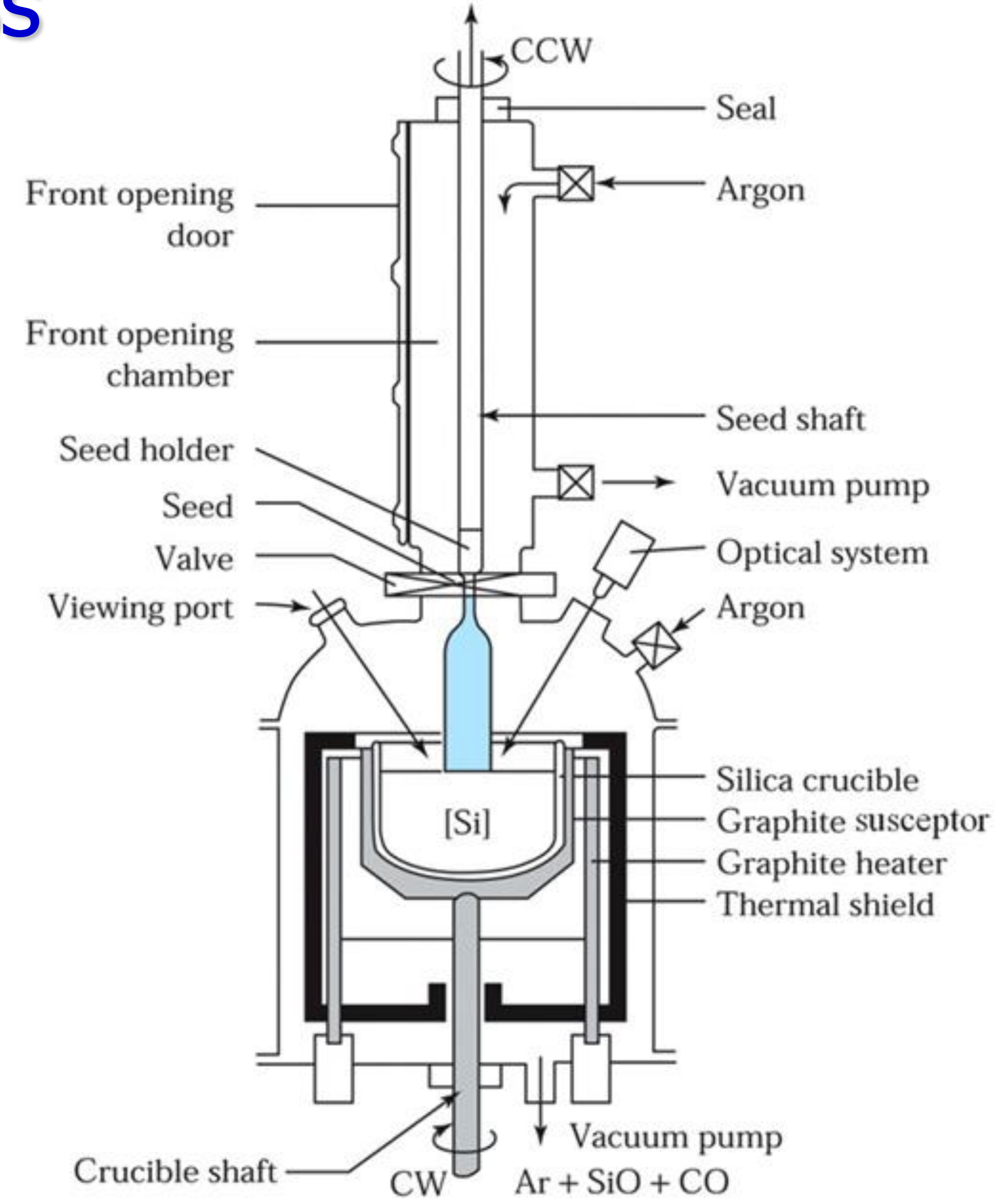
Another significant difference is the production volume, the detector surface required by particle or nuclear physics experiments is O[100 m^2/yr and quite fluctuating], while semiconductor industry is many orders of magnitude larger (~10^6 m^2 for IC and ~10^8 m^2 for solar cells → cost/cm^2 and wafer size)

Czochralski Process

growing of Si-monocrystals
from molten silicon

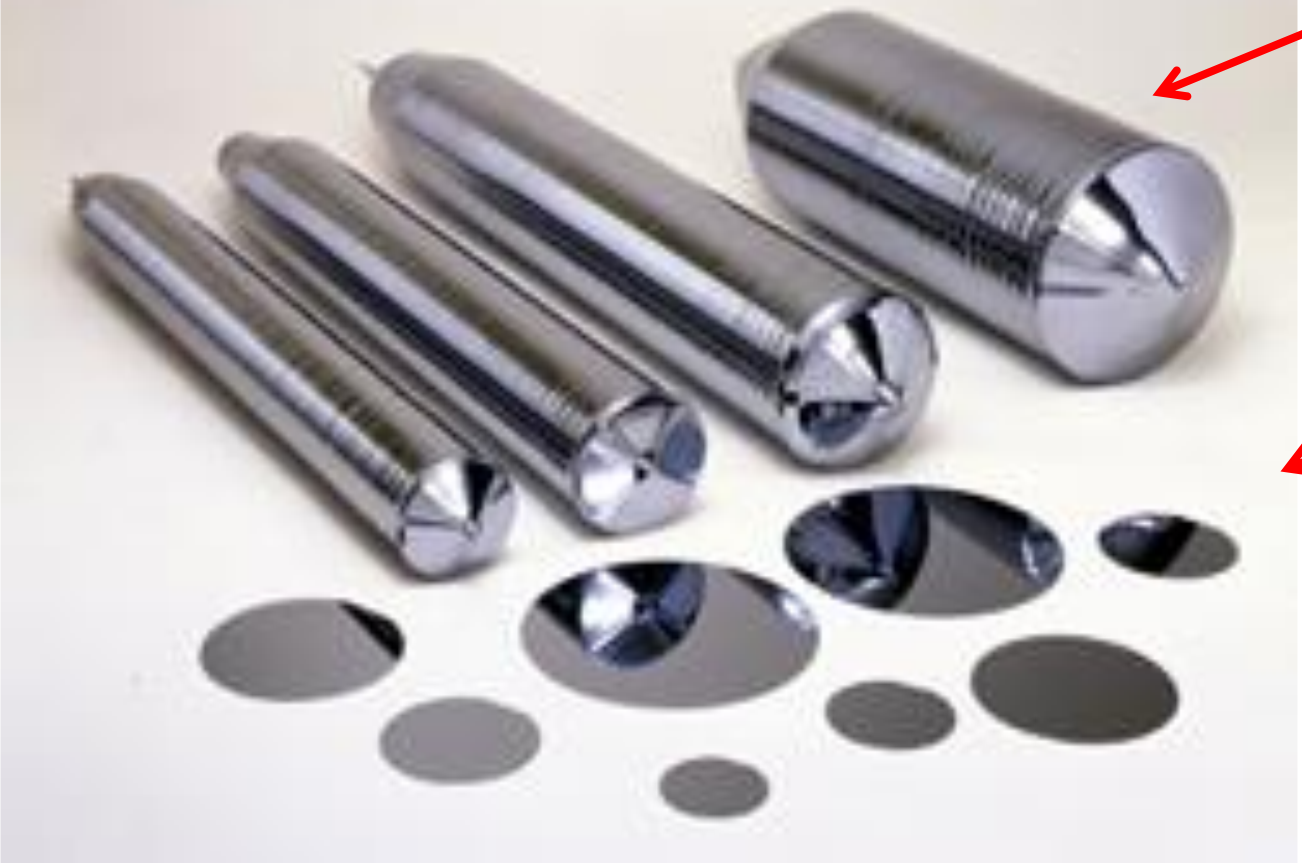
orientation determined
by seed crystal

doping applied directly
for a p-type silicon at 10^{16} B / cm^3
add 8.85 mg of B to 100 Kg of Si



Silicon Wafers

poly-silicon
(“raw” material)



Si rods

up to 300 mm diameter
and 200 kg weight

Si wafers

cut out the wafers
with a diamond saw
and polish

few 100 μm thickness

Float-Zone Process

For the production of highly-pure silicon (very high resistivity)

orientation determined by seed crystal

Melt very pure sand (SiO_2) together with coke (~1800°C)

$$SiO_2 + 2C \rightarrow Si + 2CO$$

Grind the “metallurgical grade silicon” (98% Si)
and expose it to hydrochloric gas

$$Si + 3HCl\,(gas) \rightarrow SiHCl_3 + H_2$$

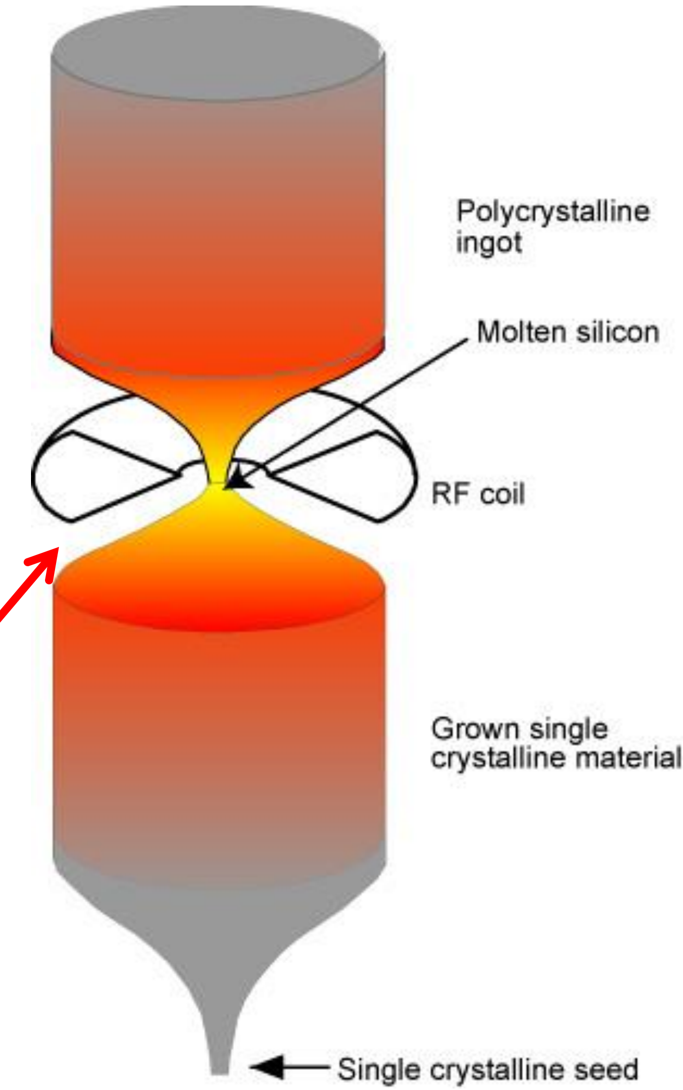
trichlorsilane boils at 31.7°C and can thus be distilled and purified

deposit silicon in a Chemical Vapour Deposition process

$$SiHCl_3 + H_2 \rightarrow Si + 3HCl$$

cast silicon into a poly-crystalline silicon rod

Using a single Si crystal seed,
melt the vertically oriented rod onto the seed
using RF power and “pull” the mono-crystalline ingot

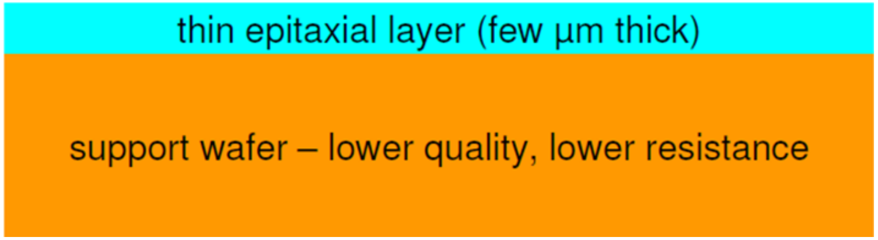


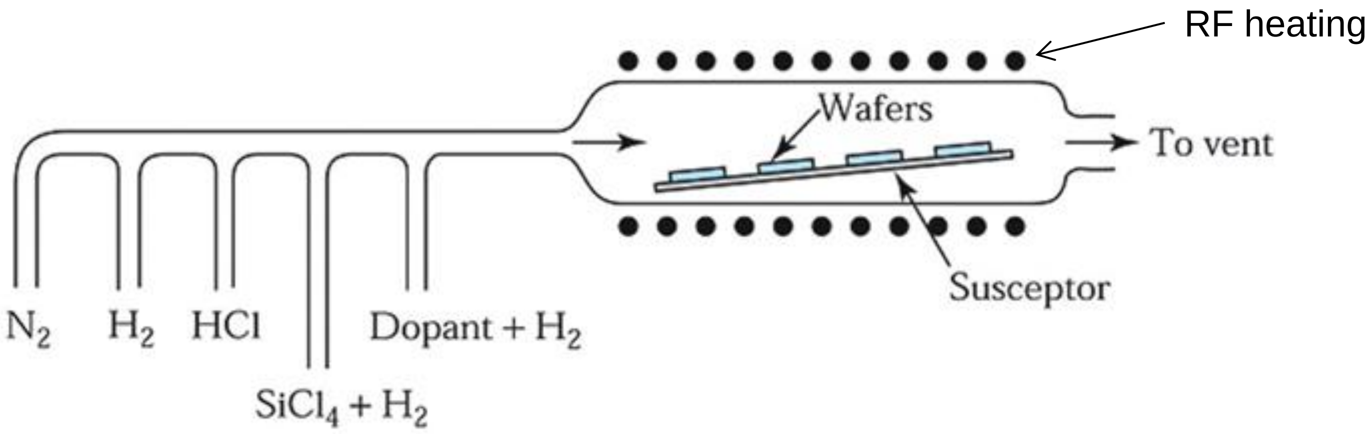
Epitaxial Growth

Growth of single-crystal silicon layers on a single-crystal silicon substrate: precipitation of atomic silicon layers from gaseous phase at high temperatures

$$SiCl_4\ \text{(gas)} + 2H_2\ \text{(gas)} \rightarrow Si\ \text{(solid)} + 4HCl\ \text{(gas)}$$

to obtain an ultra-pure silicon layer on lower quality silicon substrate.
The epitaxial layer assumes crystal structure of the substrate.
Dopants can be added to the gaseous phase.

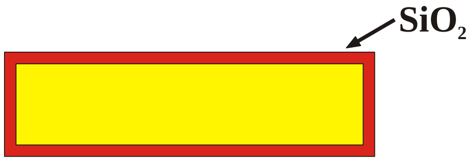




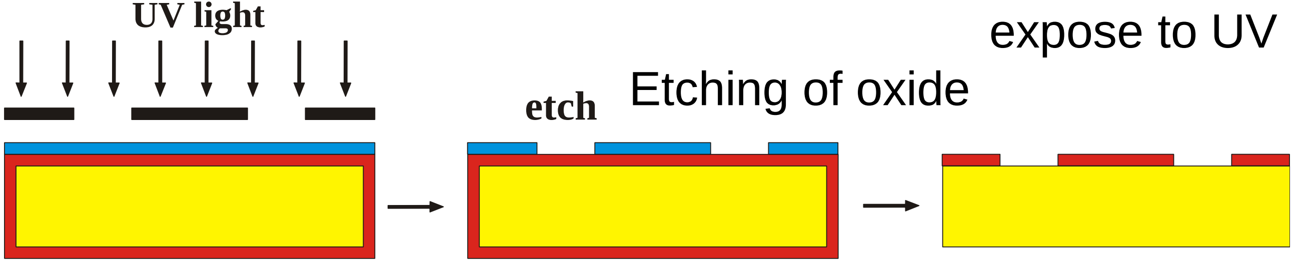
Silicon Sensor Production

Step by step production sequence (schematic)

Polished n-type silicon wafer (typical $\rho \sim$ 1-10 KΩcm)

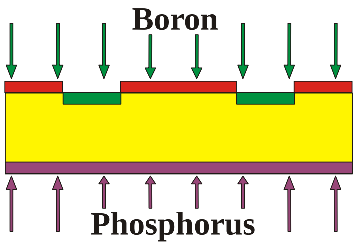


Thermal oxidation (800-1200°C)



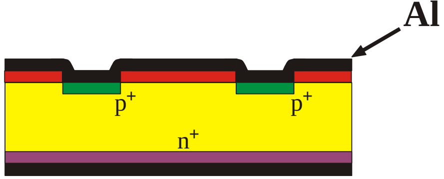
Photolithograpy (coat with photo resist; align mask,
expose to UV light, develop photoresist)

Etching of oxide

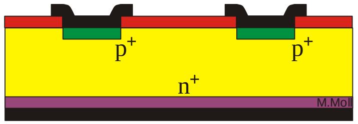


Doping with boron and phosphor by implantation (or by diffusion)

Annealing to cure radiation damage and activate dopants
- p^+ n junction on front side
- n n^+ ohmic contact on back side



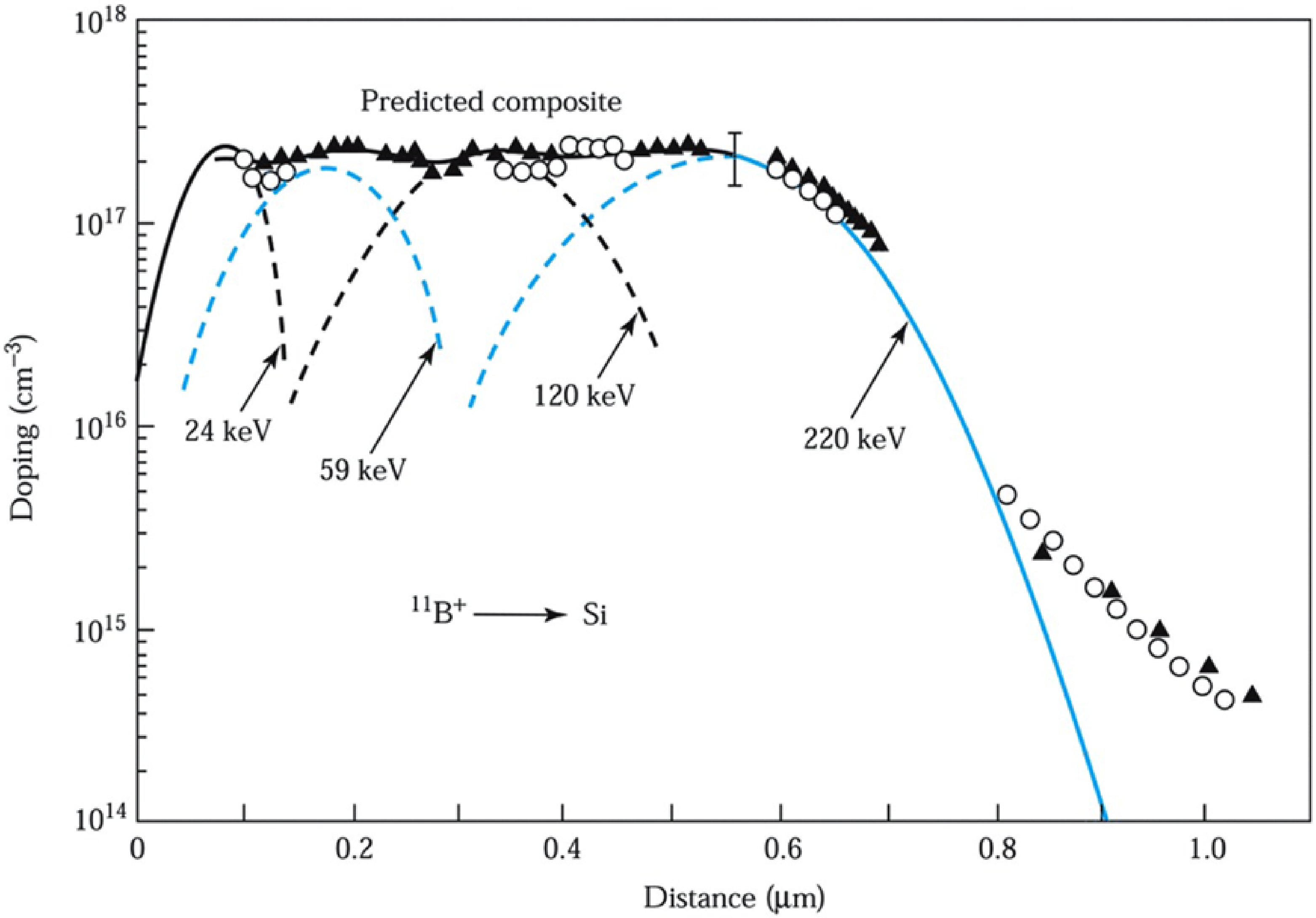
Metalization, i.e. aluminize surface e.g. by evaporation



Pattern metal for diode contacts

Doping Implantation

range of implanted ions (doping profile) as a function of the boron energy (p-type)



In Some More Detail

Thermal oxydation: the first step in ~any wafer processing → it protects Si with a thin (0.15 to 2 μm) layer of SiO_2. This is an insulator grown on the wafer surface storing the wafer in an oxygen atmosphere (900° < T < 1200° C)

Layer deposition: if other insulators (e.g. Si_3N_4) or polysilicon (e.g. to create resistive chains) are needed. They can be deposited through CVD (chemical vapour deposition), i.e. wafer in appropriate hot gases.

Photolitographic steps: this is the key step to give the shape to the electrodes (and the diodes) → photoresist is spun onto the wafer surface (~ 1 μm thickness) → baked at 100° C → apply a pattern mask (Cr on glass) in contact (including reference marks) → UV exposure transfer the pattern to the photoresist → develop and wash-out the exposed part → the wafer is now selectively protected by the photoresist

Etching: is used to copy the structure into the underlying layers (i.e. etching SiO_2). It can be done with HF (hydrofluoridic acid) or a plasma. It can be dry or wet, the last one less expensive is used for detectors (it underetches, but structures are coarse)

Doping: here we come to the hart of the fabrication process

Diffusion: this is one way of doping. The wafer (with the SiO_2 pattern) is exposed to a high T gas (800 < T < 1200 C) which enters the pure silicon. The gas (e.g. PH_3, B_2H_6 or AsH_3) brings the dopant to the Si with a doping profile having the max on surface. Dopant also diffuse laterally in Si (0.8 depth)

Implantation: this is mostly used, costs more but is more precise and more easily controllable. Ions of doping atoms are accelerated and shot to the Si wafer (but stopped by SiO_2) and also by photoresist as this is room-T operation and photoresist can stay). The implantation dose can be well controlled (reproducibility). The ion penetration (therefore doping profile) can also be tuned precisely.

Every implantation is followed by a thermal treatment as the doping atoms are on regular places in the lattice and thermal vibrations should “shake-them-in”. Moreover this thermal cycle also cures some of the local defects (like cluster damage) which has been created by the stopping ions. This “annealing” cycle also help diffusion of dopants and can be used to drive them into shallow junctions.

Metallization: the last step. Used to provide a low-resistivity connection between parts in the same detector or to form bond pads (for the connection with the outside). Most used element is Al as it adheres well to the SiO_2 and has low R. Al-layer thickness ~ 1 μm (evaporated or sputtered).

Last step: wafer dicing (to single out sensors from wafers).

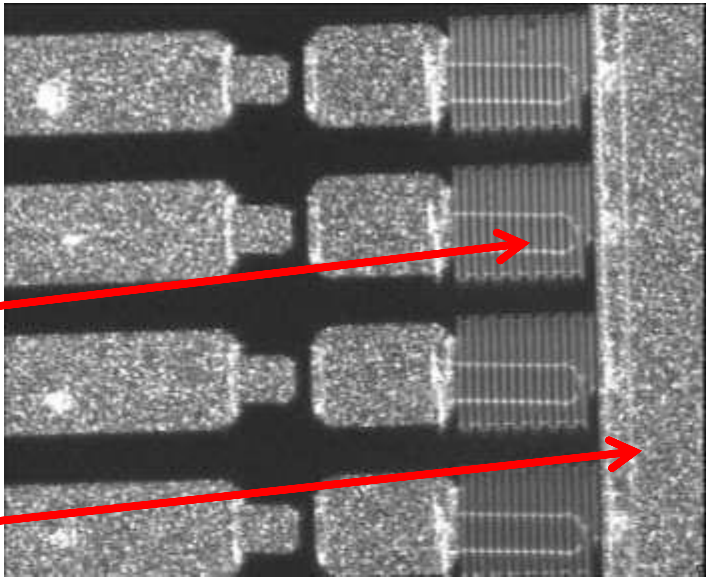
Poly-Silicon

In single crystal silicon the crystal lattice of the entire sample is continuous and unbroken
(no grain boundaries)

Poly-crystalline silicon (poly-silicon) is a material consisting of multiple small silicon crystals
(can be recognized by a visible grain, a “metal flake effect) → may tune R

Deposited by (e.g.) plasma-enhanced chemical vapor deposition (PECVD) of amorphous silicon at ~ 300° C

Sheet resistance of up to $R_s \approx 250\ k\Omega/\mu$.
Up to $R \approx 20\ M\Omega$ is achieved
(→ winding poly structures are deposited)



Floating p+ are grounded (referenced) through deposition of poly-crystalline silicon between p+ implants and a common bias line.