The SiPM Physics and Technology - a Review -

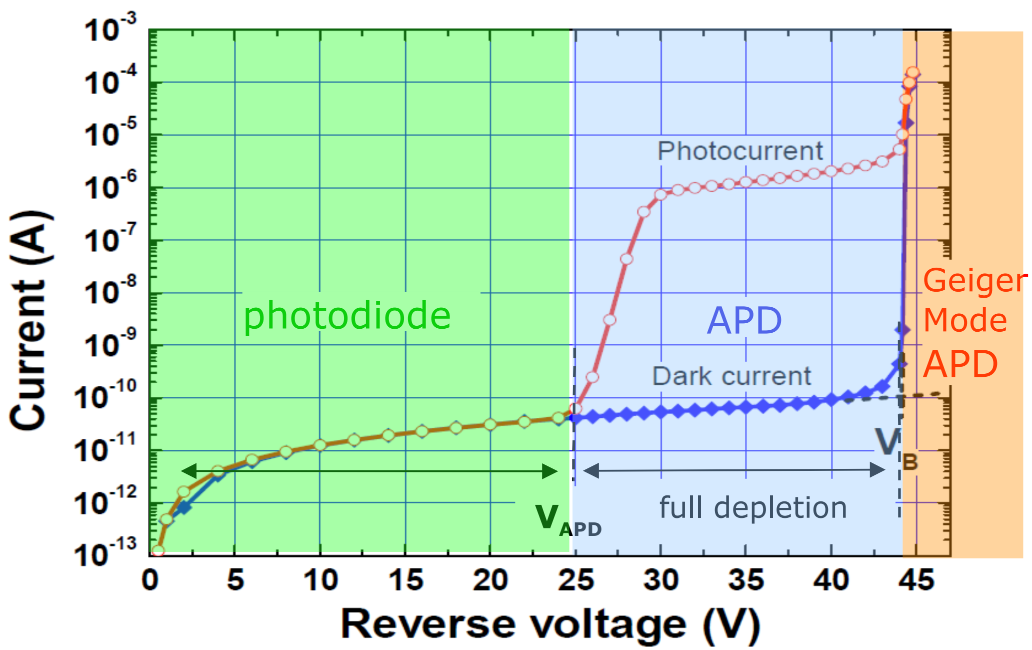
G.Collazuol
Department of Physics and Astronomy, University of Padova and INFN

Overview

- Introduction
  - Key physics and technology features
    - I-V characteristics
      - Device response
        - Noises
          - Photo-detection efficiency
            - Timing properties
              - Summary and Future

The silicon PM: array of GM-APD

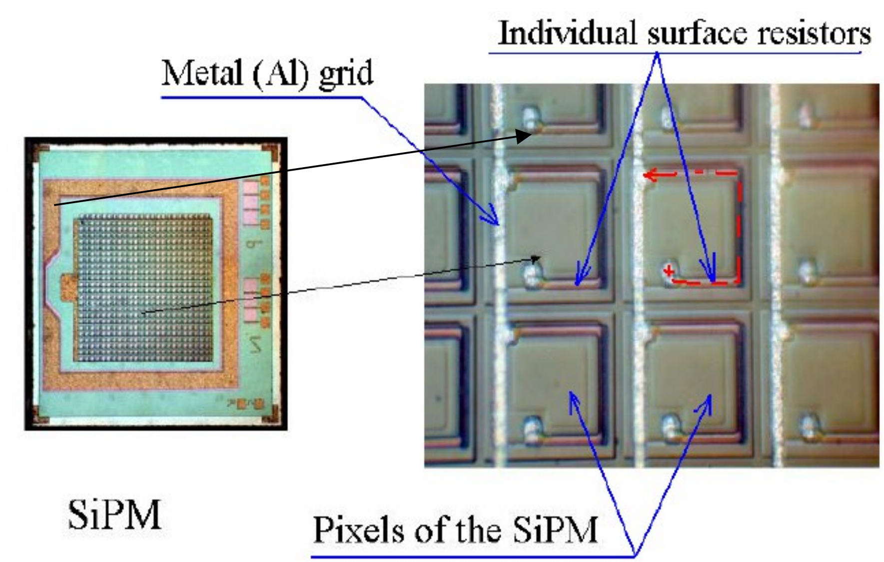
Single GM-APD gives no information on light intensity → MATRIX structure
first proposed in the late '80-ies by Golovin and Sadygov



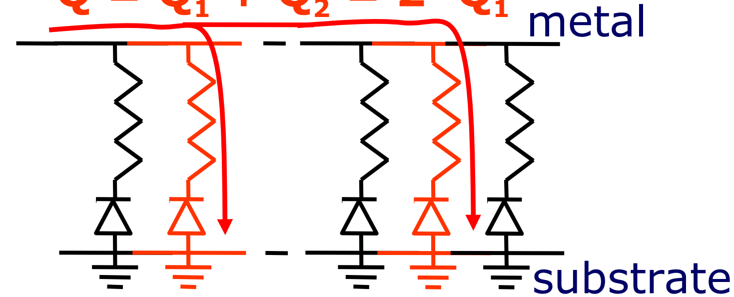
A SiPM is segmented in tiny GM-APD cells and connected in parallel trough a decoupling resistor, which is also used for quenching avalanches in the cells

Each element is independent and gives the same signal when fired by a photon

In principle output charge is proportional to the number of of incident photons



$Q = Q_1 + Q_2 = 2*Q_1$



Σ digital signals → analog signal !!!

A bit of history

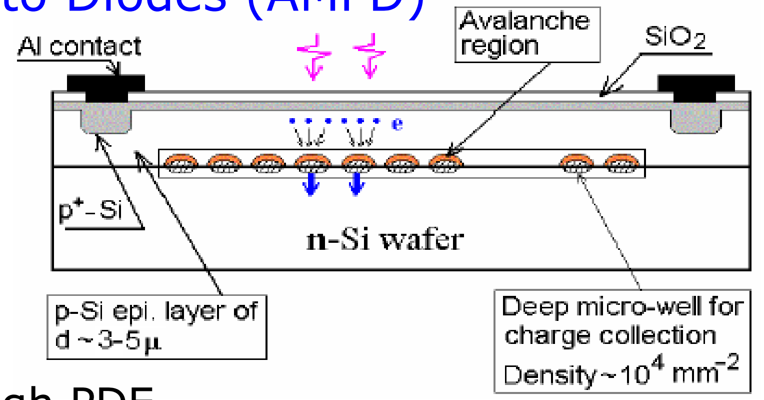
Pioneering work since late 80-ies at Russian institutes

Investigations of various multi-layer silicon structures with local micro-plasma suppression effect to develop low-cost GM-APD arrays

Early devices ageing quickly, unstable, noisy

Sadygov – JINR/Micron (Dubna)

Avalanche Micro-channel/pixel Photo Diodes (AMPD)

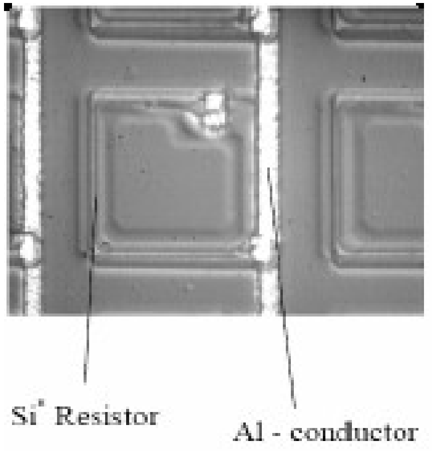


- high PDE
- very high density of micro-cells

eg Sadygov, NIMA 567 (2006)

Dolgoshein - MePhi/Pulsar (Moscow)

Poly-silicon resistor

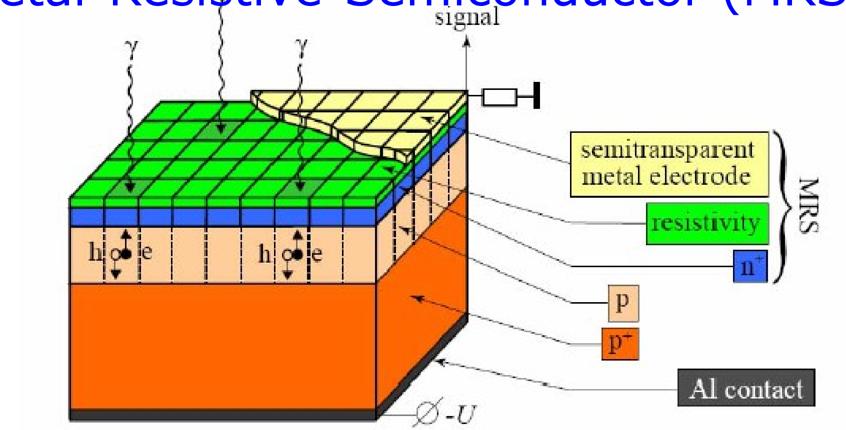


- Low fill-factor
- Simple fabrication technology

e.g., Dolgoshein, NIMA 563 (2006)

Golovin - Obninsk/CPTA (Moscow)

Metal-Resistive-Semiconductor (MRS)



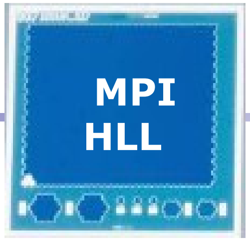
- High fill factor
- Good pixel to pixel uniformity

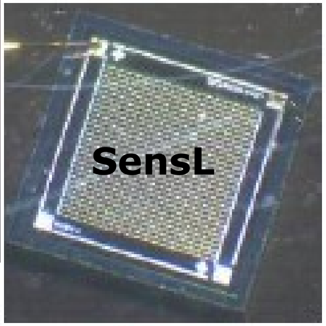
e.g., Golovin NIMA 539 (2005)

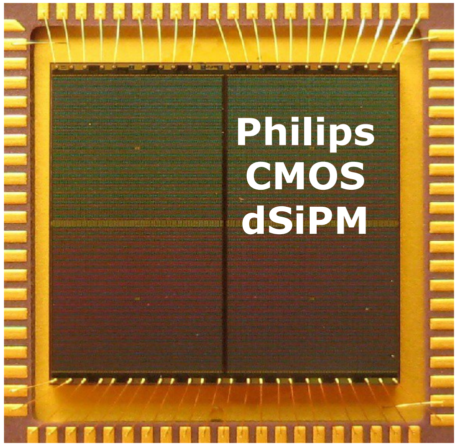
Today

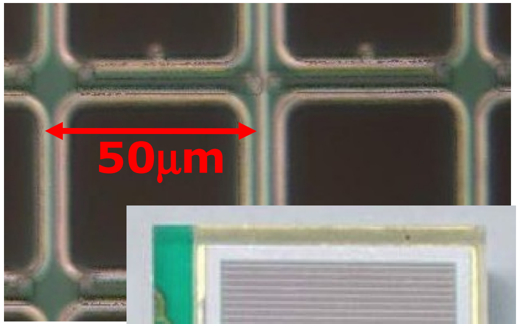
Many institutes/companies are involved in SiPM development/production:

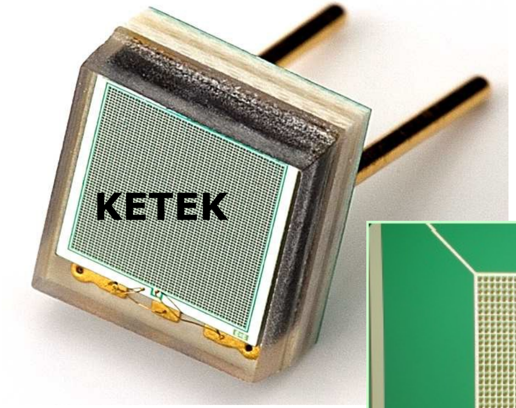
- **CPTA**, Moscow, Russia
- **MePhi/Pulsar** Enterprise, Moscow, Russia
- **Zecotek,** Vancouver, Canada
- **Hamamatsu HPK**, Hamamatsu, Japan
- **FBK-AdvanSiD**, Trento, Italy
- **ST Microelectronics**, Catania, Italy
- **Amplification Technologies** Orlando, USA
- **SensL**, Cork, Ireland
- **MPI-HLL**, Munich, Germany
- **RMD**, Boston, USA
- **Philips**, Aachen, Germany
- **Excelitas** tech. (formerly Perkin-Elmer)
- **KETEK**, Munich, Germany
- **National Nano Fab Center**, Korea
- **Novel Device Laboratory** (**NDL**), Bejing, China
- **E2V**
- **CSEM**

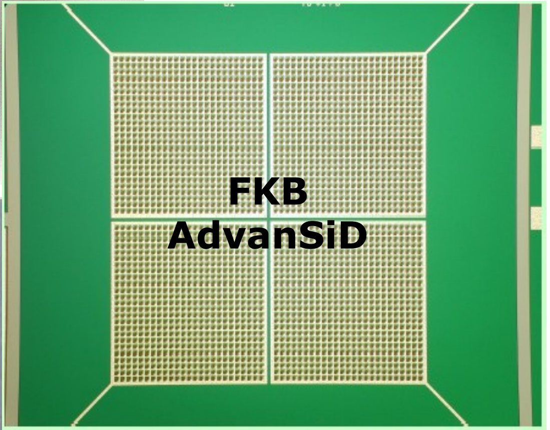


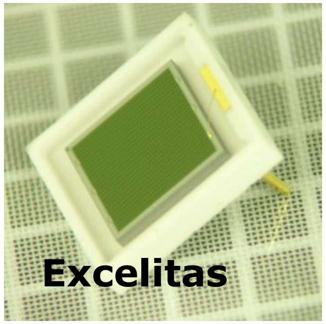




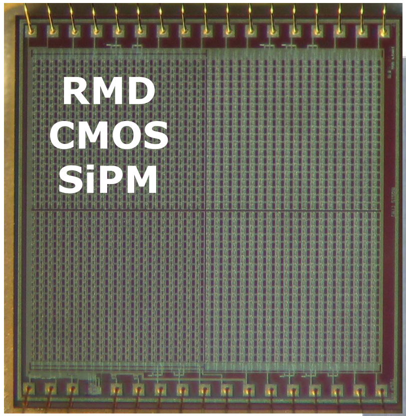


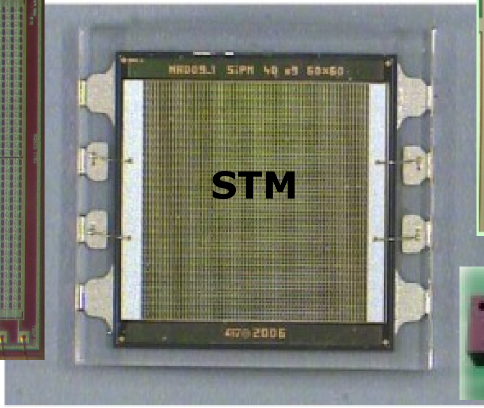


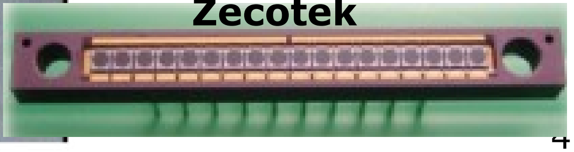




Amplification Technologies (DAPD)







G.Collazuol - PhotoDet 2012

Physics & Technology
Key features

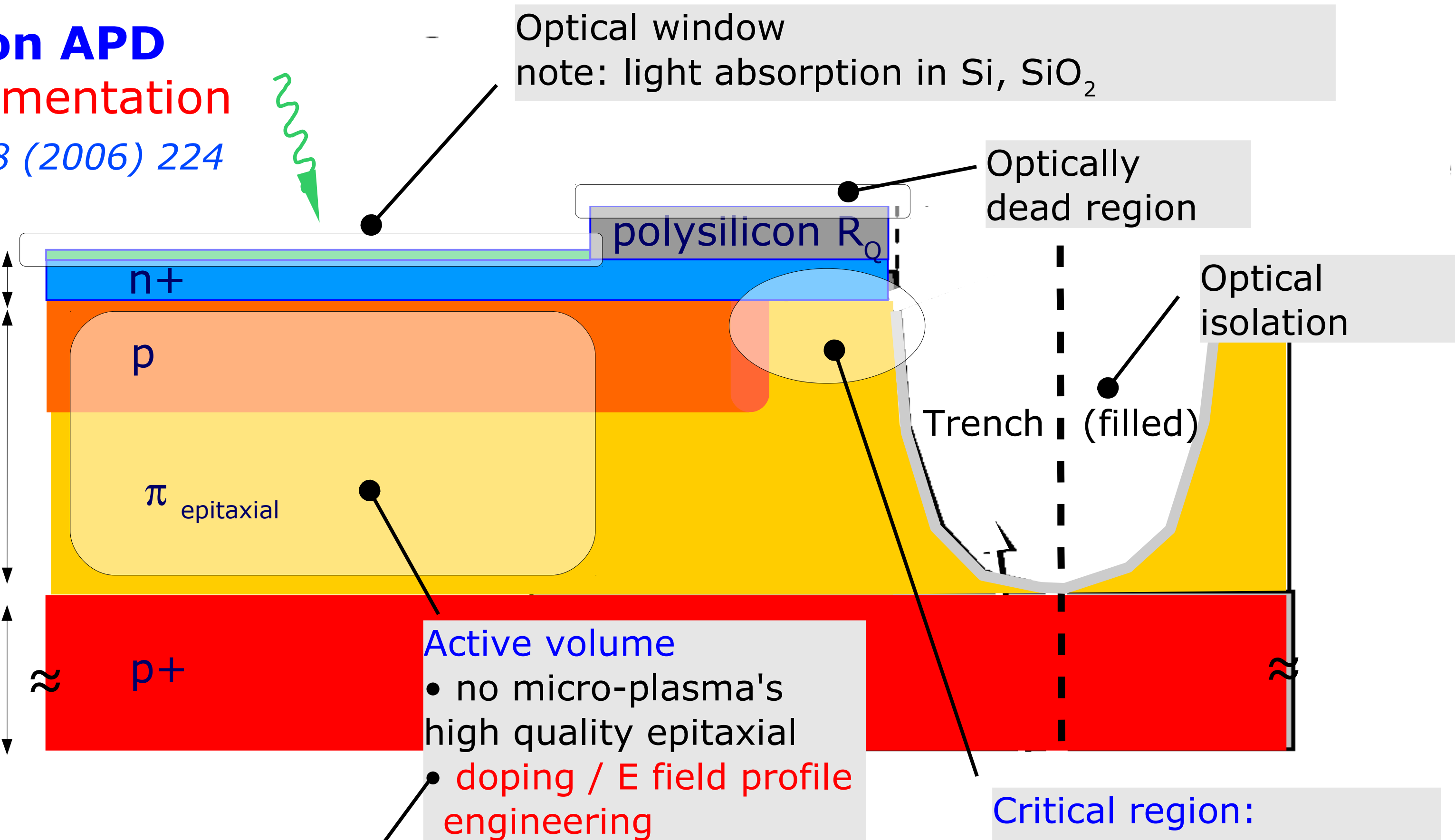
- Closeup of a cell – Custom vs CMOS
- Guard Ring and Optical isolation
- Operation principles of GM-APD and quenching modes

Close up of a cell – custom process

Shallow-Junction APD

Example of implementation

C.Piemonte NIM A 568 (2006) 224



- n^+ on p abrupt junction structure
- Anti-reflective coating (ARC)
- Very thin (100nm) n^+ layer: "low" doping
  → minimize Auger and SHR recombination
- Thin high-field region: "high" doping p layer
  → limited by tunneling breakdown
  → fixes V_{BD} junction well below V_{BD} at edge
- R_Q by poly-silicon
- Trenches for optical insulation (cross-talk)
- Fill factor: 20% - 80%

Optimization for blue light (420nm)

CMOS vs Custom processes

"Standard" CMOS processes

- shallow implant depths
- high doping concentrations
- shallow trench isolation (STI)
- deep well implants (flash extension)
- no extra gettering and high T annealing
- non optimized optical stacks
- design rule restrictions

high E field (low V_{bd})

tunneling
lattice stress
(defects/traps)

high DCR

limited PDE
(often p-on-n)

limited timing performances
(long diffusion tails)

Recent progresses in CMOS APDs due to:

1) **high voltage (flash) extension** often available in **standard** processes
   - deep wells (needed for the high voltages used in flash memories)

2) **Additional processes** (custom) available:
   - buried implants
   - deep trench isolation
   - optical stack optimization

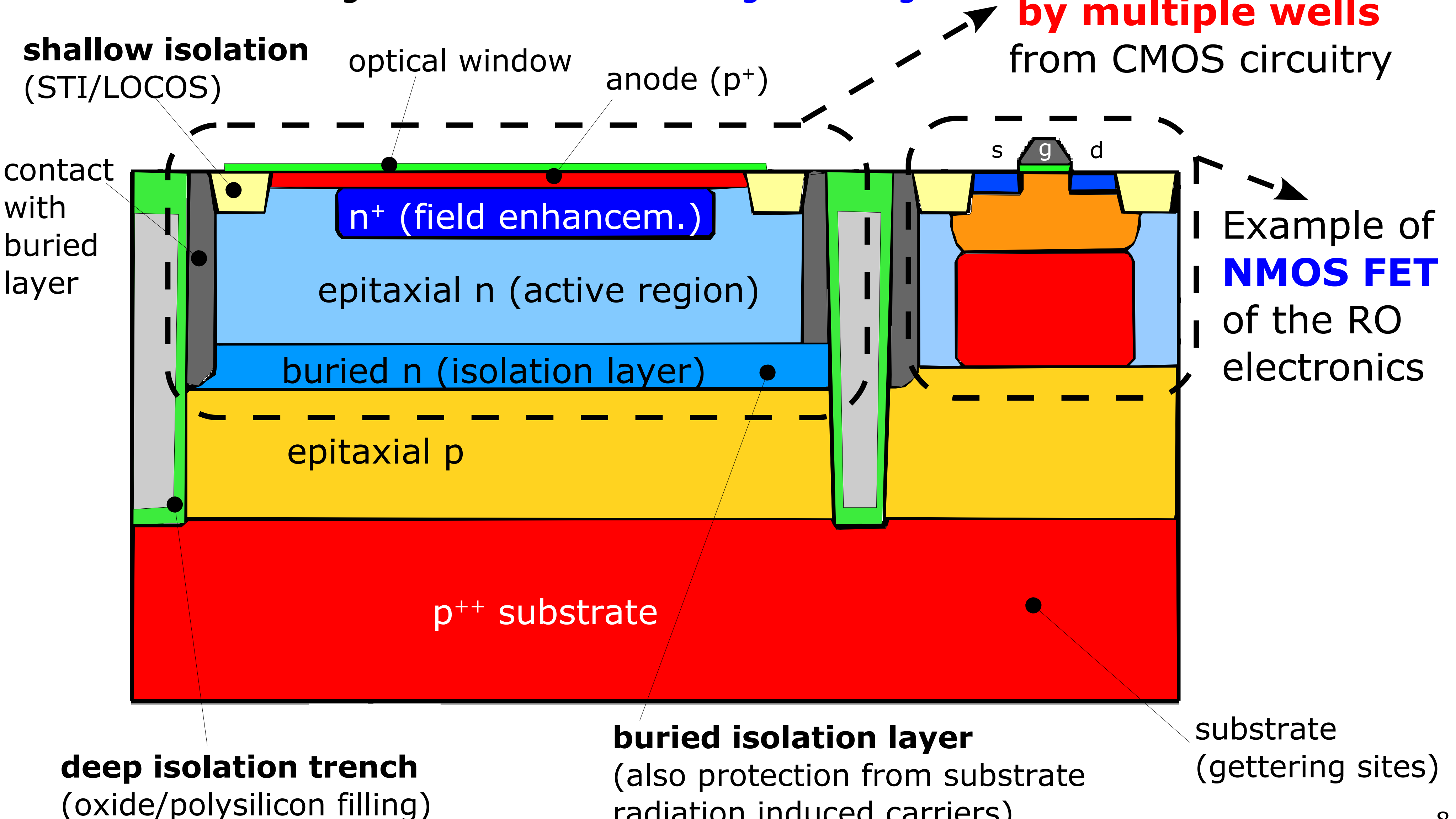
Key elements for CMOS SiPMs
- **APD cell isolation** from CMOS circuitry
- **guard ring** (again)

Close up of a CMOS cell

APD integration into CMOS

Example of implementation *T.Frach in US patent 2010/0127314*

Note • extended CMOS processes exploited
• careful design of cell isolation and guard ring

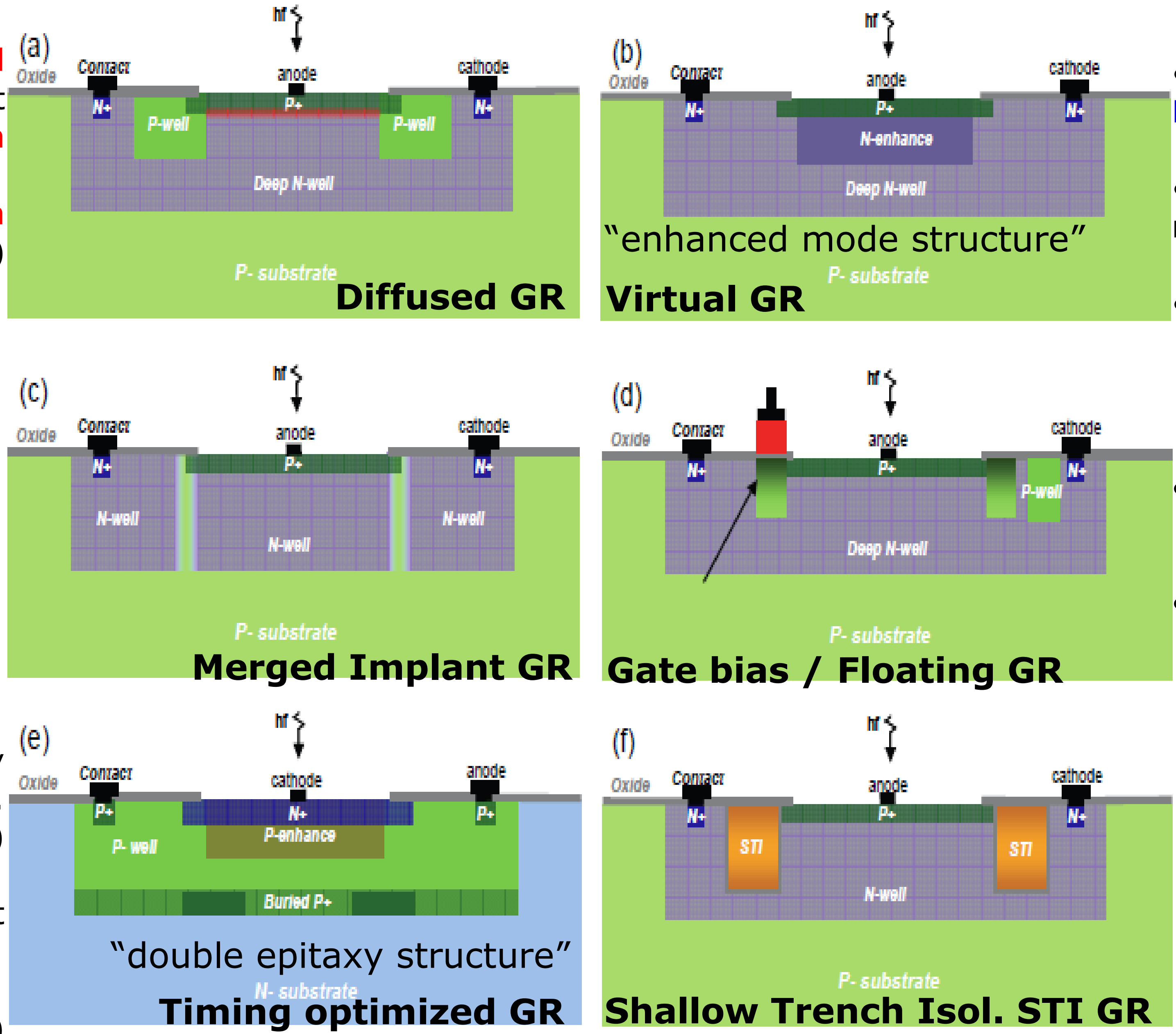


The Guard Ring structure

- high E field structure, not uniform
- neutral region (timing tails)
- limited fill factor

- alternative to Diffused GR
- difficult to implement

- developed by S.Cova and coll. (fully custom)
- state of the art SPAD timing and PDE (red enhanced)



- well tuned high E field structure
- no additional neutral regions
- fill factor less limited

- less commonly exploited
- careful modeling required

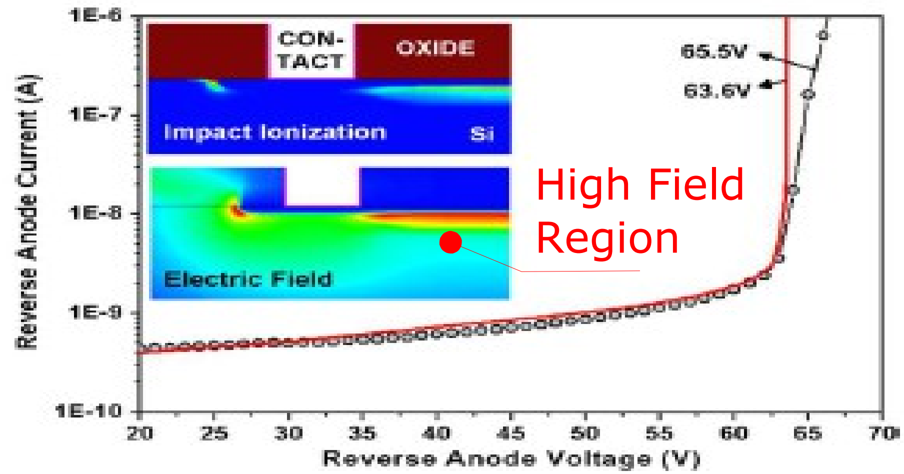
- physically blocks and confines the high E field in active region
- might cause high DCR due to
  - tunneling
  - etching induced defects/traps

from “Avalanches in Photodiodes” G.F.Dalla Betta Ed., InTech Pub. (2011)

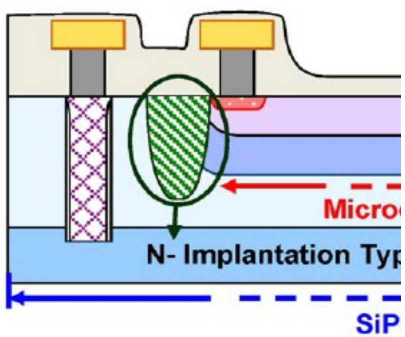
Guard Ring structures in SiPM

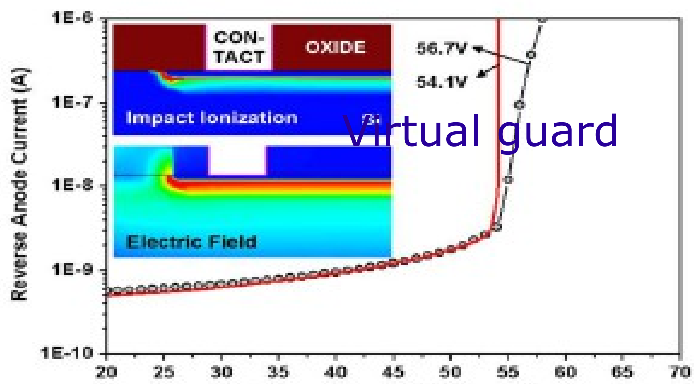
Sul et al, IEEE EDL 31 2010 "G.R. Structures for SiPM"

Virtual guard ring most often used

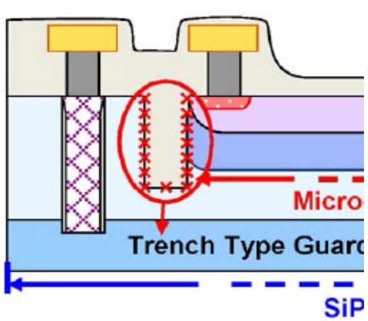


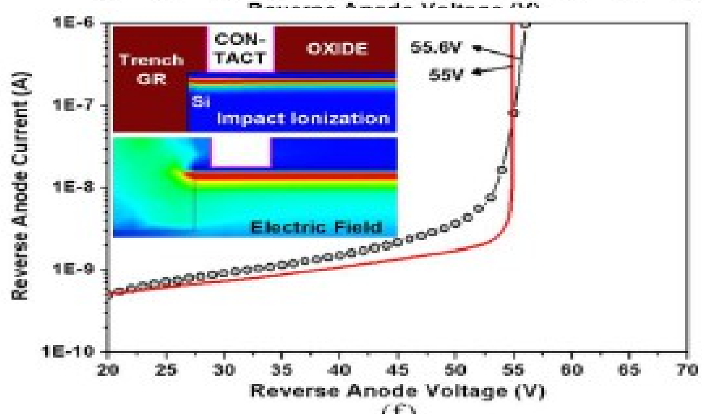
Implant / Gate bias



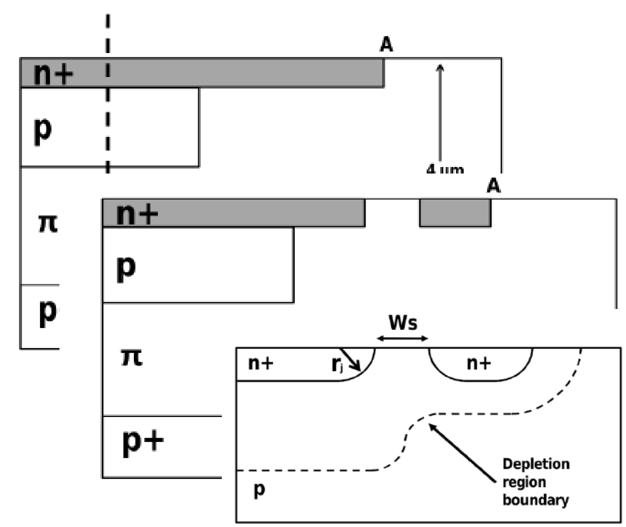


Trench type





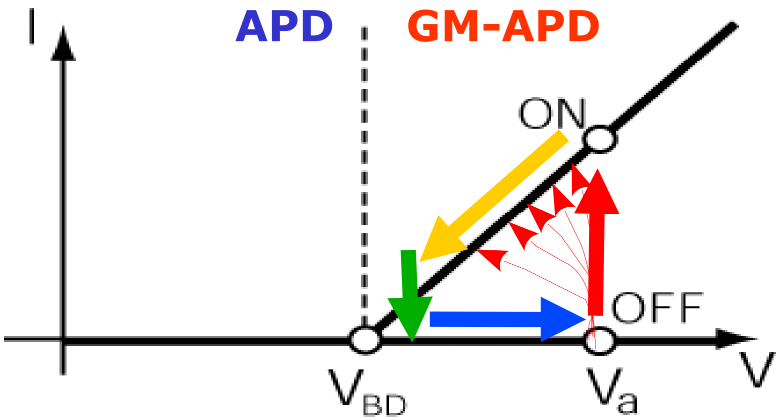
Maresca et al. Proc. of SPIE Vol. 8072 "Floating field ring ... to enhance fill factor of SiPM"



Operation principle of a GM-APD

Avalanche processes in semiconductors are studied in detail since the '60 for modeling micro-plasma instabilities

McIntyre JAP 32 (1961), Haitz JAP 35 (1964) and Ruegg IEEE TED 14 (1967)



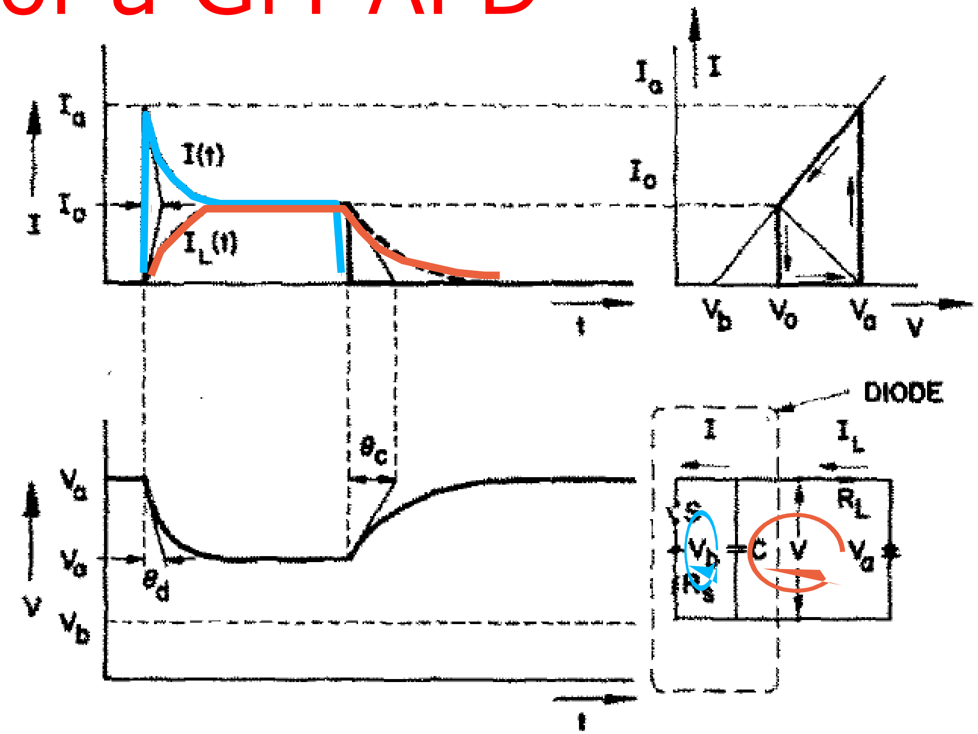
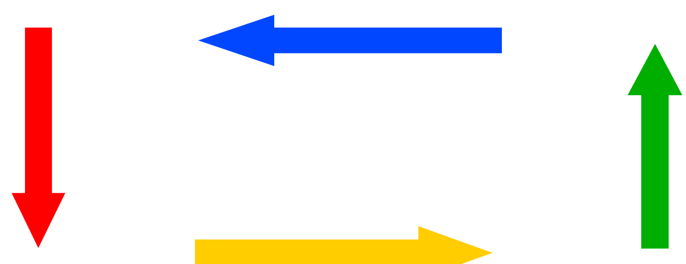


FIG. 3. Shape of current pulse for $\theta_d \ll \tau_1(I_0)$.

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

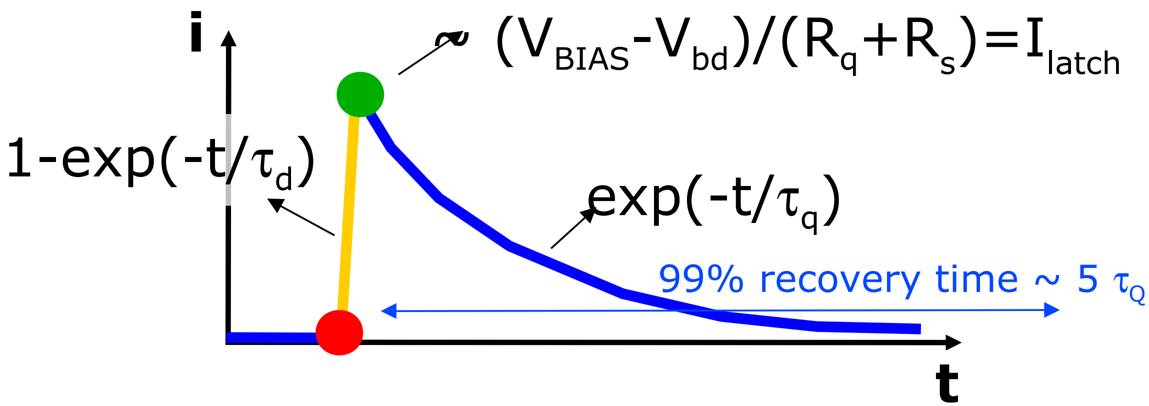


P_{10} = turn-off probability
probability that the number of carriers traversing the high-field region fluctuates to 0

ON condition: avalanche triggered, switch closed
C_d discharges to V_{bd} with a time constant $R_d x C_d = \tau_{discharge}$,
at the same time the external current asymptotic grows to $(V_{bias}-V_{bd})/(R_q+R_d)$

Passive Quenching

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



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

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

Haitz JAP 35 (1964)

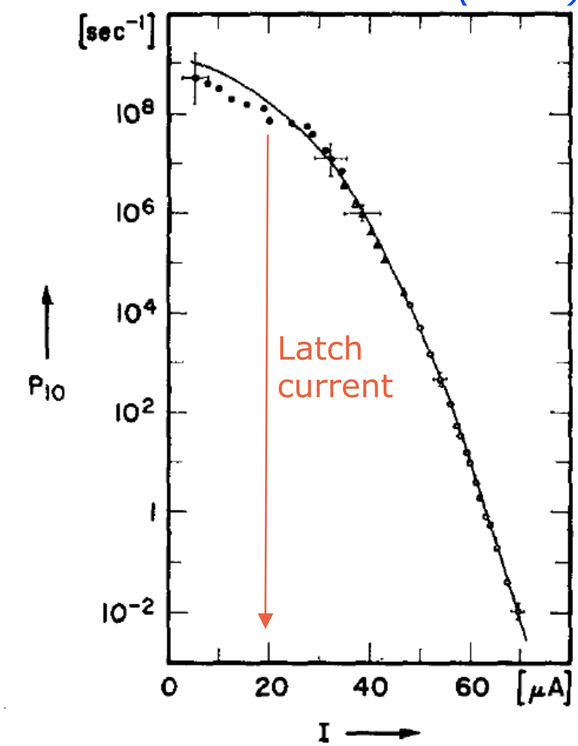


FIG. 2. Turnoff probability per second as function of pulse current.

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

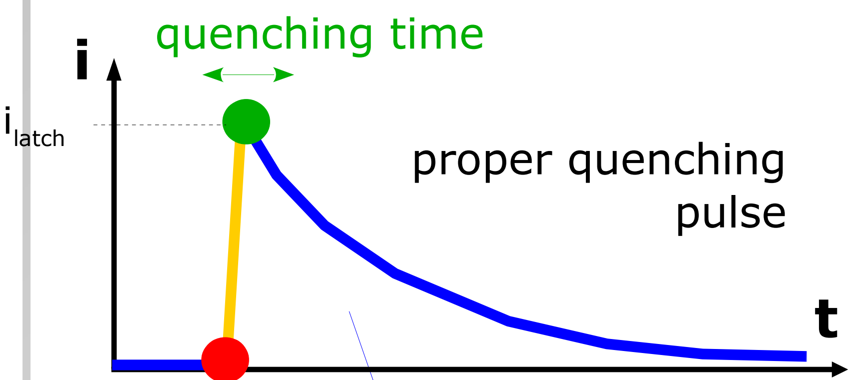
It is possible to define a GAIN (discharge of a capacitor)

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

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

Passive Quenching Regime

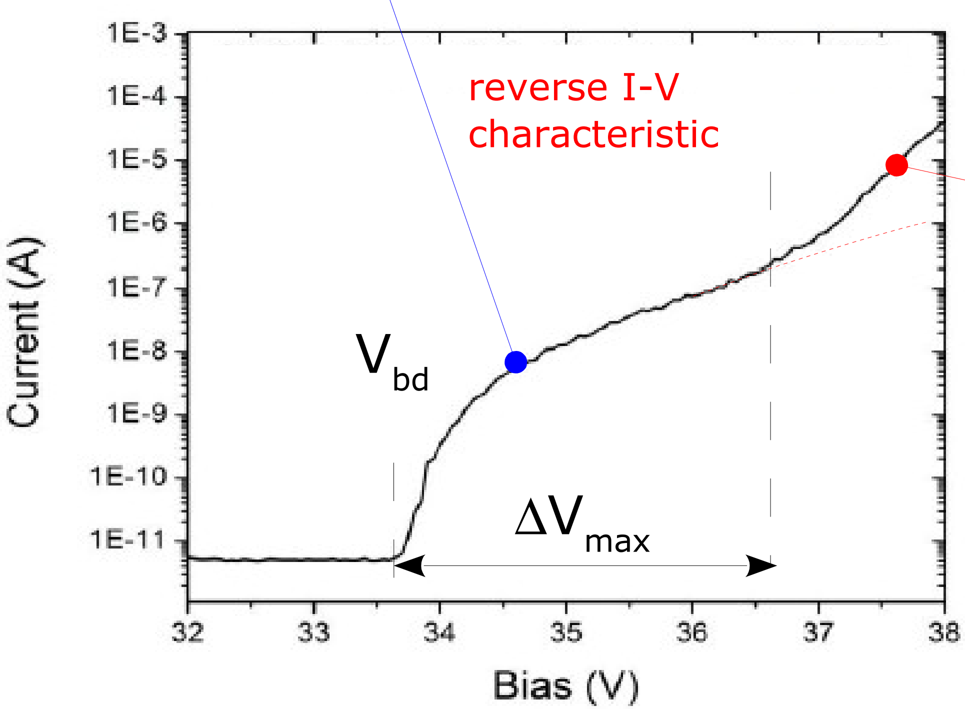
Proper value of quenching resistance Rq is crucial to let the internal current decrease to a level such that statistical fluctuations may quench the avalanche → sub-ns quenching time → crucial to have well defined gain

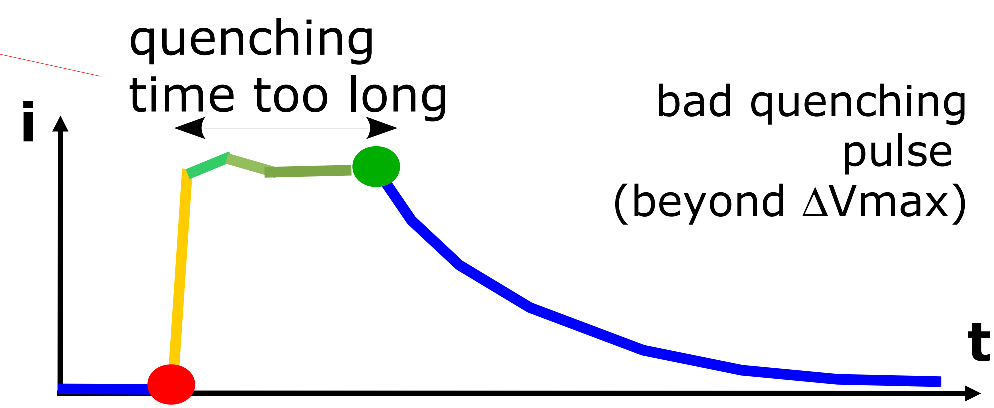


Given Rq the proper quenching regime is for ΔV in the range:

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

where as a rule of thumb
$I_{latch} \sim 20\mu A$ → ΔV_{max} ~ a few Volts (typically)





Operative ΔV Range – I_{dark}/DCR

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:

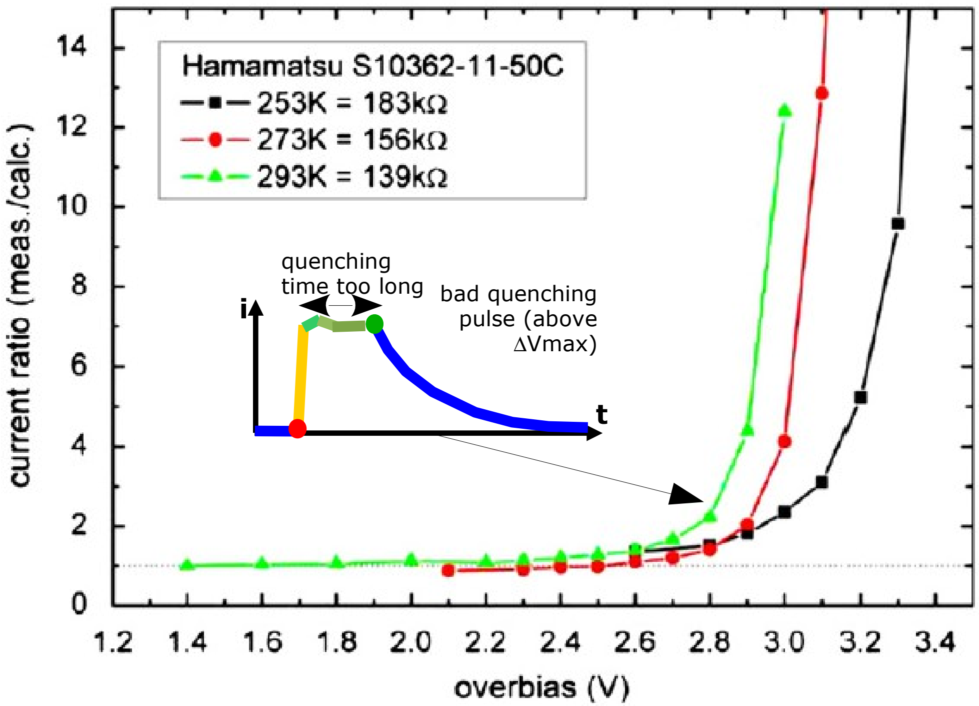
after Jendrysik et al NIM A 2011
doi:10.1016/j.nima.2011.10.007

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

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

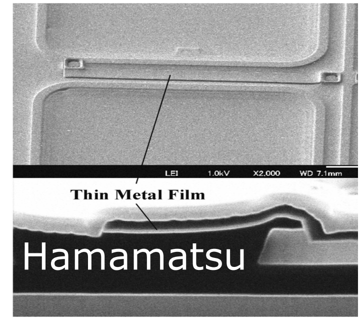


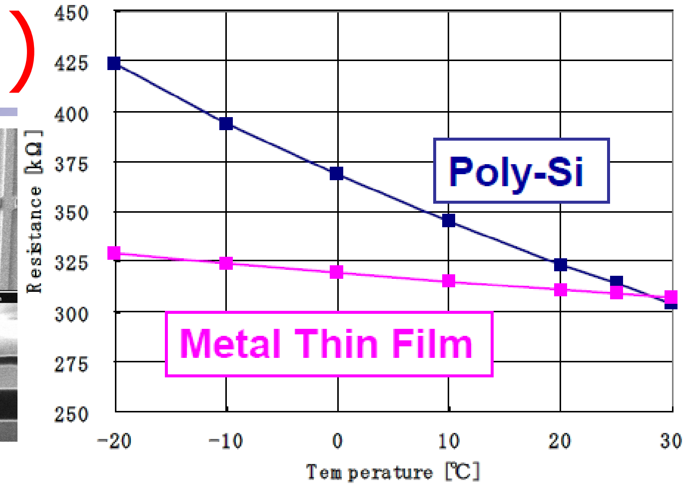
G.Collazuol - PhotoDet 2012

Passive Quenching (Resistive)

1) common solution: poly-silicon

2) alternative: metal thin film
→ higher fill factor
→ milder T dependence





Nagano IEEE NSS-MIC 2011

3) alternative principle: bulk integrate resistor
→ flat optical window → simpler ARC
→ fully active entrance window
→ high fill factor (constraints only from guard ring and X-talk)
→ diffusion barrier against minorities
→ less X-talk
→ positive T coeff. (R~ $T^{+2.4}$)
→ production process simplified → cost

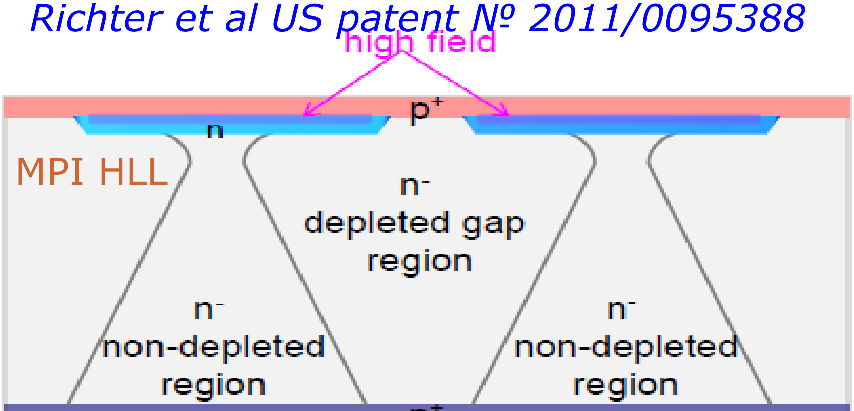
pro

contra

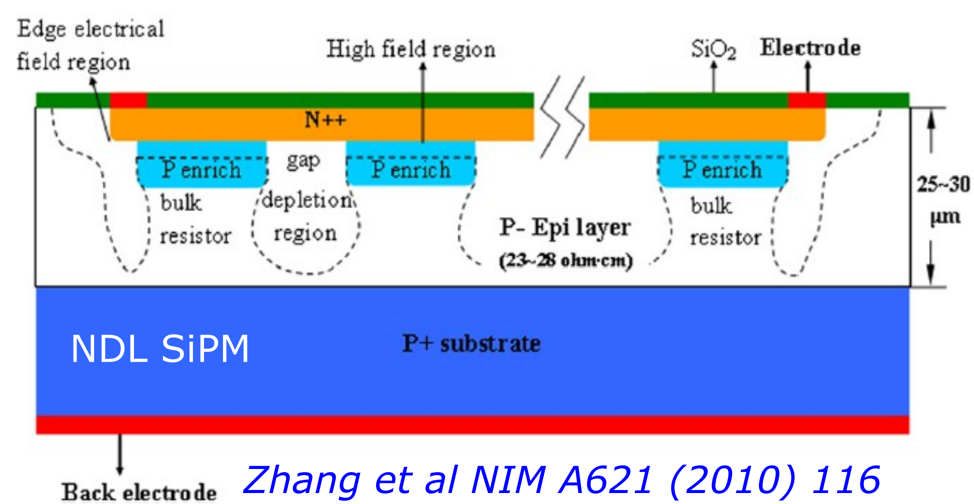
← Rq matching constraints cells' pitch/wafer thickness

← vertical R is JFET
→ non-linear I-V
→ long recovery

Ninkovic et al NIM A610 (2009) 142
and NIM A628 (2011) 407
Richter et al US patent № 2011/0095388



principle proved

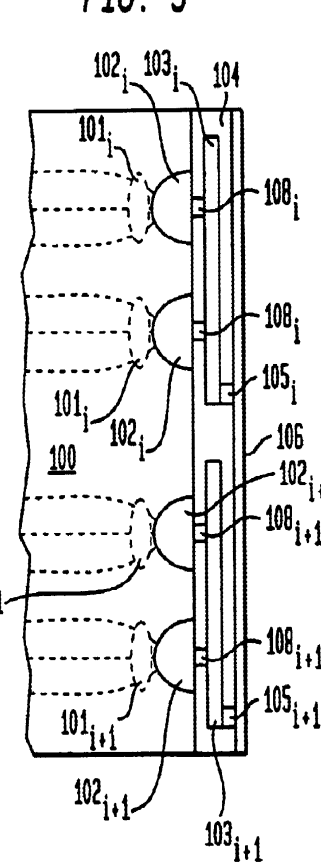


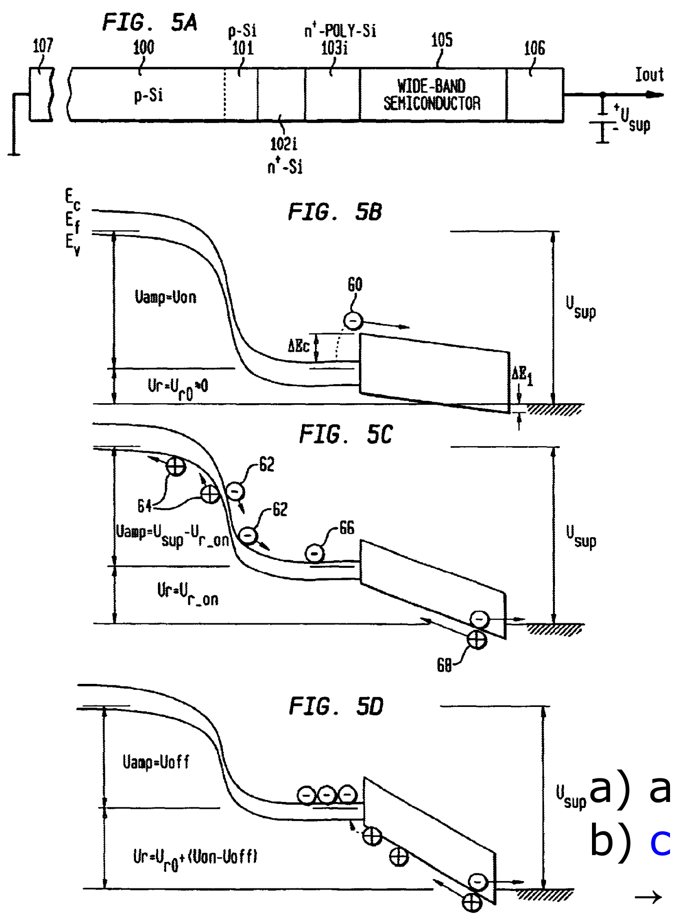
Zhang et al NIM A621 (2010) 116

Passive Quenching (Capacitive)

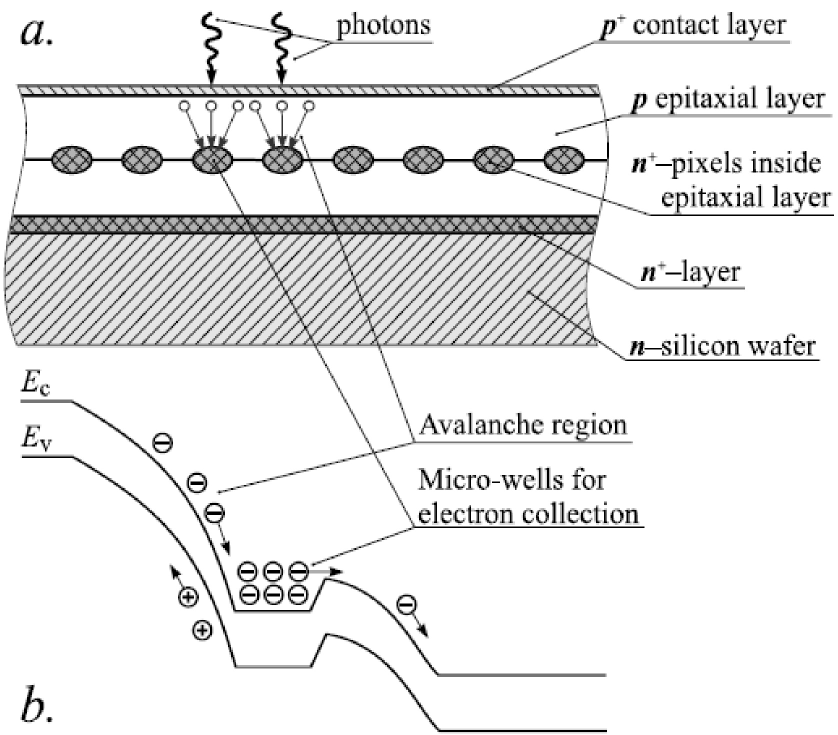
Quenching feedback due to charge accumulated by means by semiconductor barriers

AmplificationTechnologies
Shushakov et al US Patents
№ 2004/6885827 and № 2011/7899339





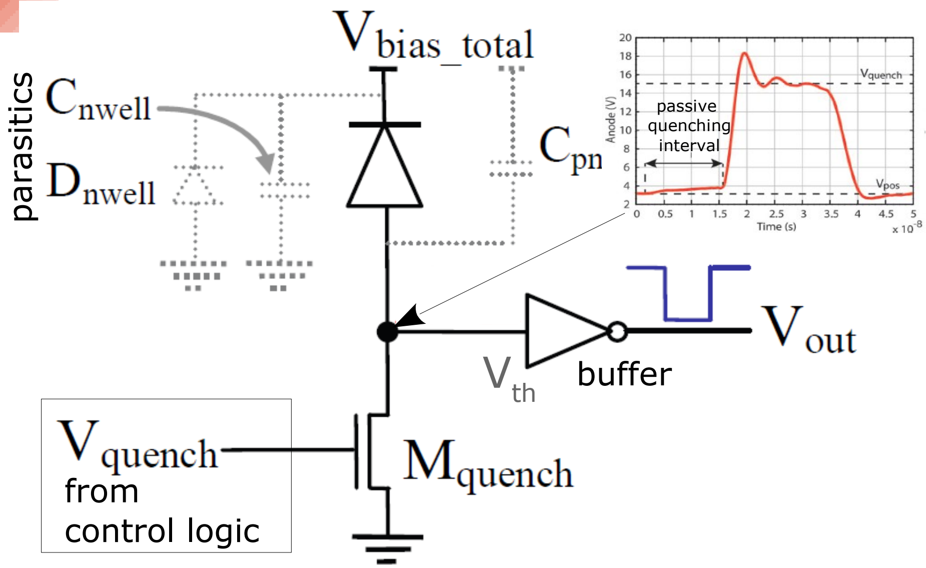
Zecotek
Sadygov et al arXiv 1001.3050
Sadygov RU Patents № 1996/2102820 and № 2006/2316848



a) avalanche at internal high field regions
b) charges accumulated in isolated potential wells
→ E field reduced (locally) → avalanche quenched
→ Fast signal induced (capacitive) outside
c) potential wells discharge slowly by tunneling
(discharge must be delayed for good quenching)
→ high E field recovered

Note: induced signal is fast (ns) but recovery quite slow (ms) (non exponential)

Active Quenching



Basic circuit elements:

1) quench circuit to detect and stop the avalanche and restore bias conditions

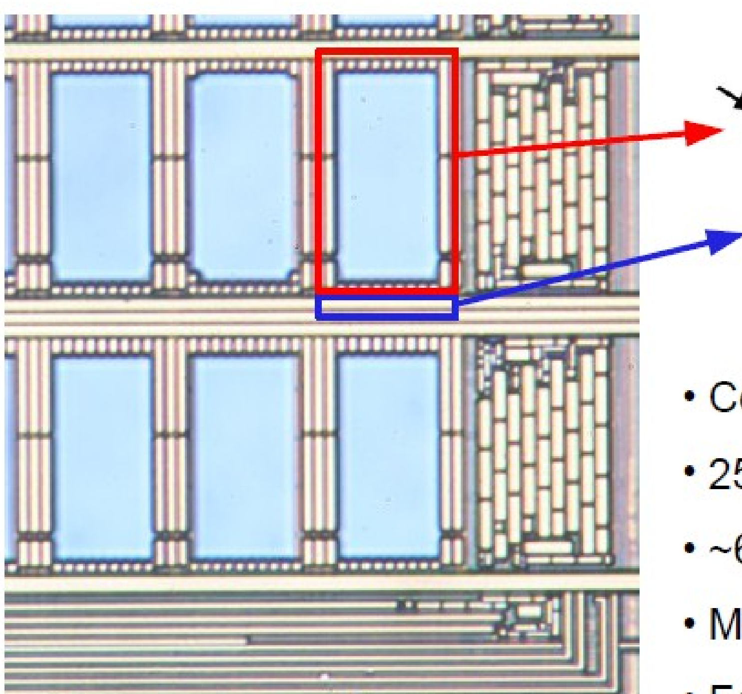
2) buffer (low capacitive load) for isolating the APD from the external electronics capacitance

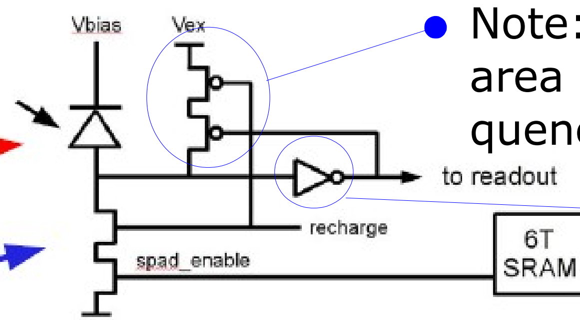
Configuration with anode to ground potential is best: only C_{det} is involved → minimum RC load

→ minimum quenching dead-time

→ minimum charge flow in APD (less after-pulses)

(in addition n-well regions (cathode) can be shared among many cells)





- Note: use of PMOS to minimize the area wrt NMOS for the same target quenching resistance
- buffer → simple inverter as input signal is already digital

dSiPM cell electronics

- Cell electronics area: 120µm²
- 25 transistors including 6T SRAM
- ~6% of total cell area
- Modified 0.18µm 5M CMOS
- Foundry: NXP Nijmegen

T.Frach at LIGHT 2011

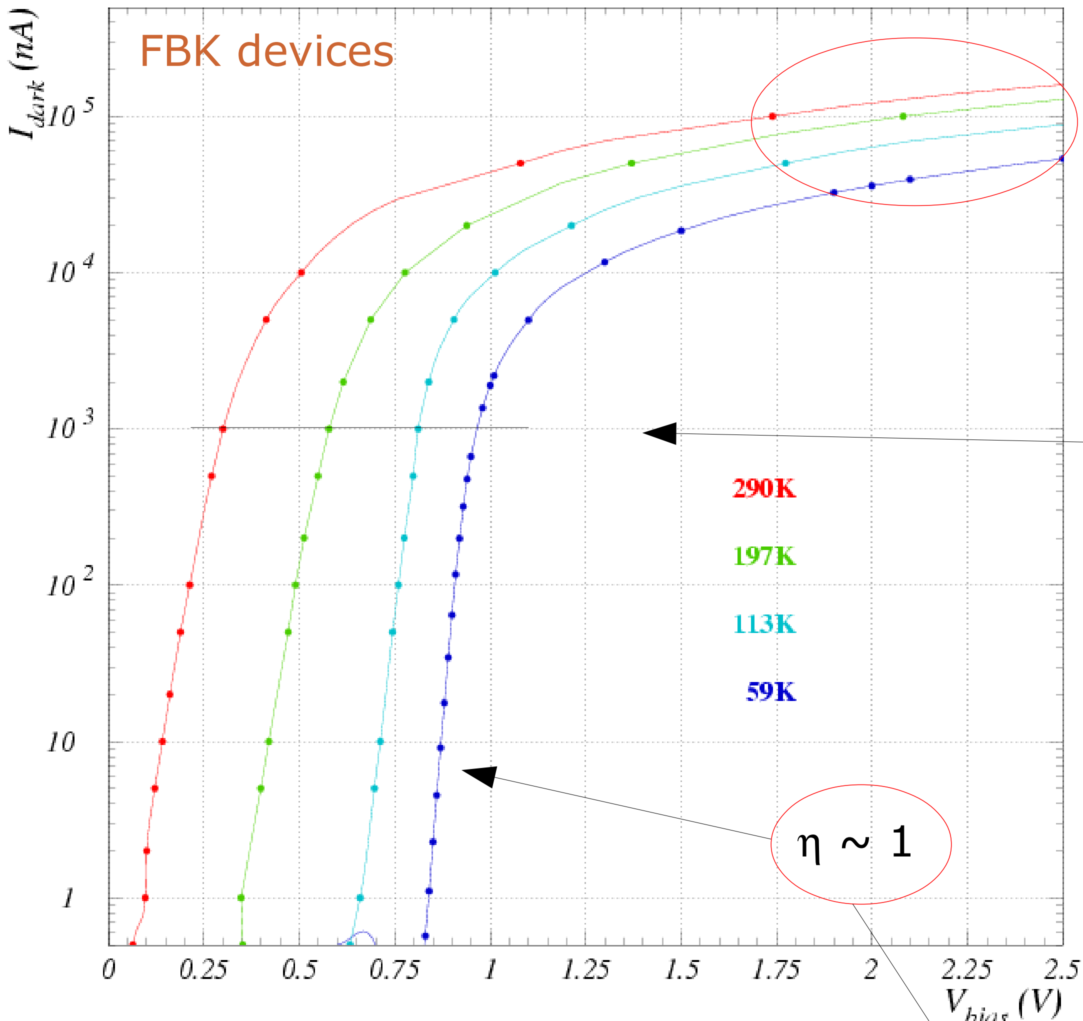
- Cell area ~ 30x50µm²
- Fill Factor ~ 50%

I-V characteristics

- Information from Forward current →
  - Rq
  - junction Temperature
  - ...

- Information from Reverse current →
  - breakdown V_{bd}
  - T coefficient
  - ...

I-V characterization: forward bias



3 **Ohmic** behavior at high current

Linear fit → $R_{series} \sim R_q / N_{cells}$

2 **Voltage drop** (V_d) decreases linearly with T decreasing (e.g. at 1μA)

1 **Forward current**

$$I_{forward} \sim C(\eta) A(T)[\exp(\frac{q V_d}{\eta k T}) - 1]$$

Shockley et al. Proc. IRE 45 (1957)

η ideality factor

Diffusion current dominating: η → 1

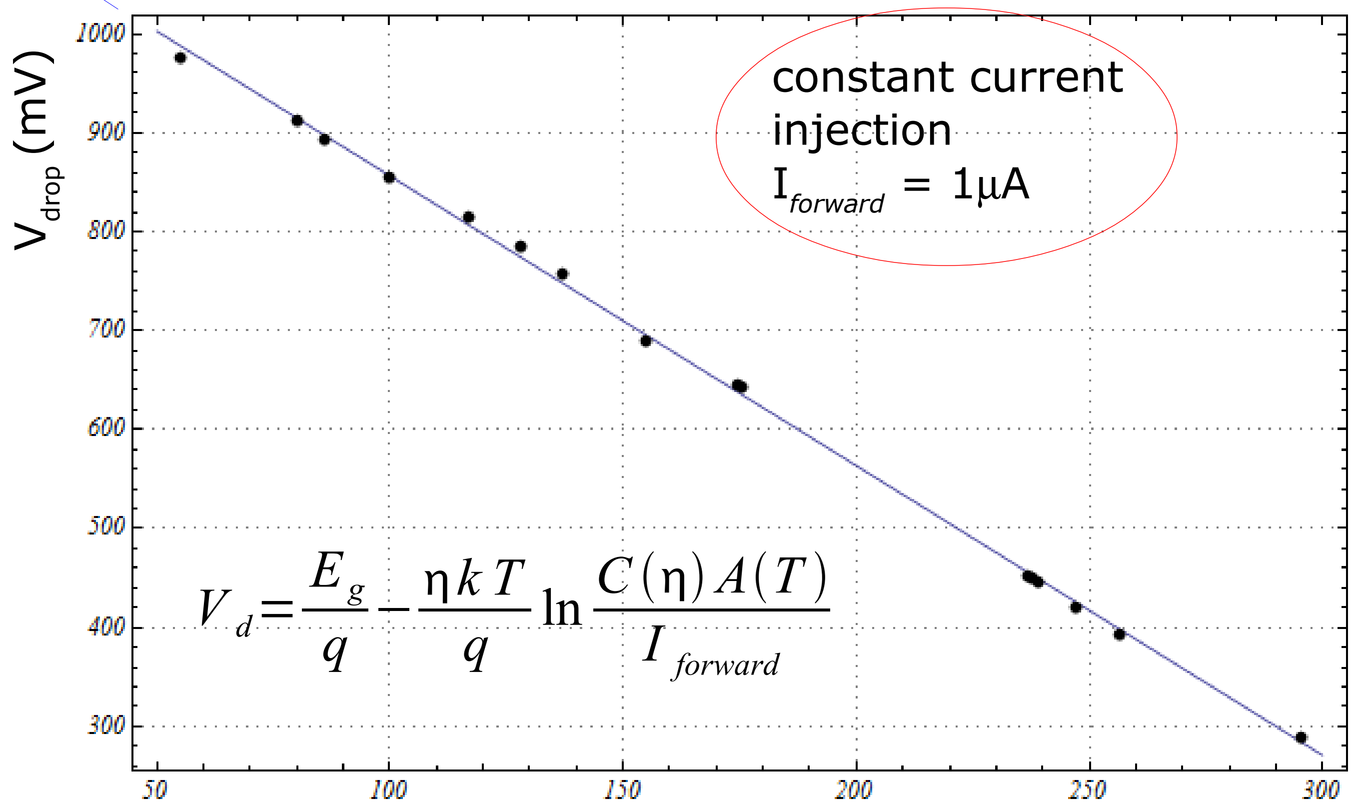
Recombination current dominating: η → 2

Forward I-V → Junction Temperature probe

Voltage drop at fixed forward current → precise **measurement of junction T**…

for T→0 ideally V_d→ Eg
(freeze-out effects apart)

… otherwise not trivially measured !



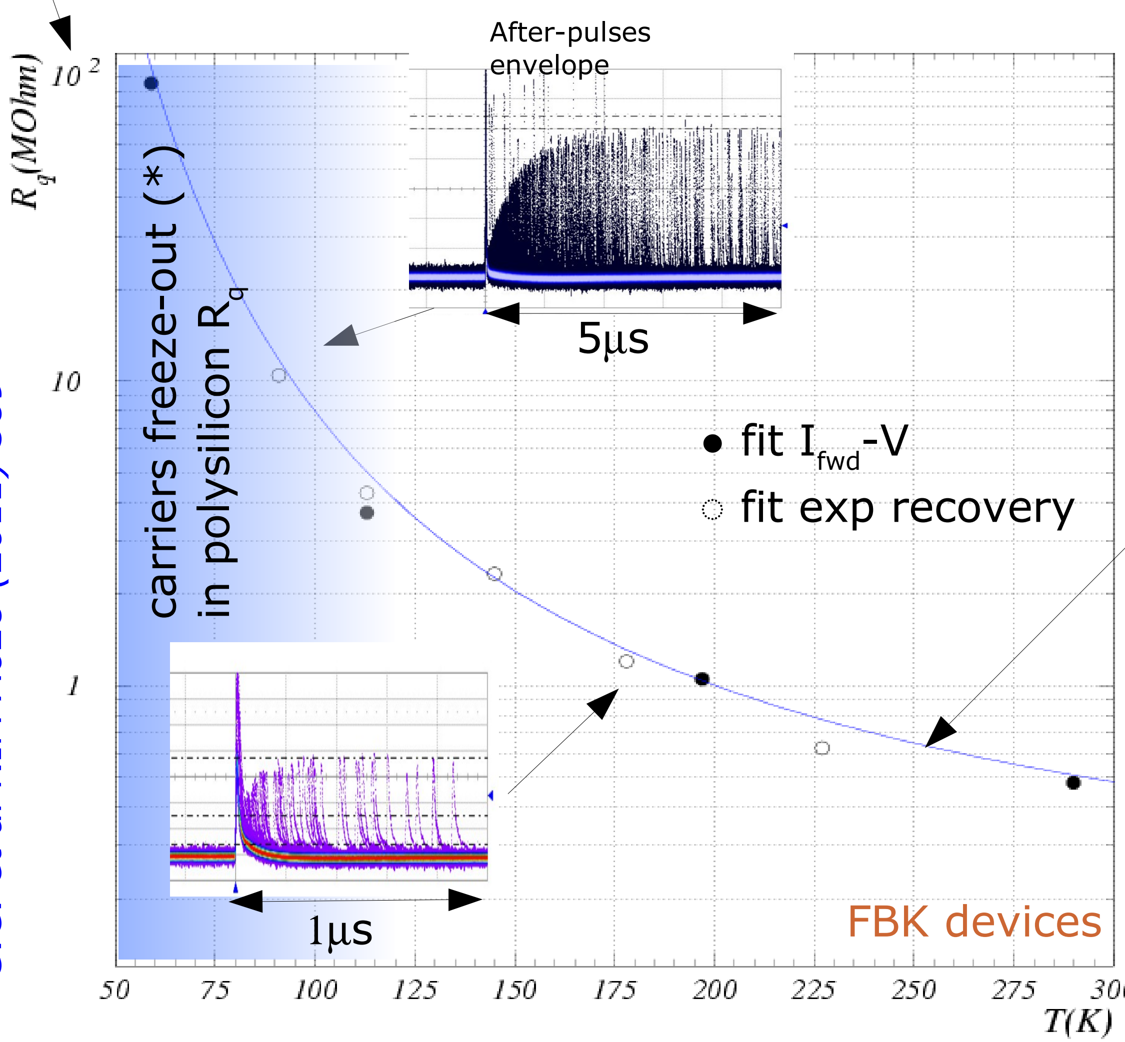
- (almost) linear dependence with slope $dV_{drop}/dT|_{1\mu A}$ ~ -3mV/K
(we don't see freeze-out effects down to 50K)

- direct and precise **calibration/probe** of junction(s) Temperature

Forward I-V → Series Resistance (vs T)

Two ways for measuring series resistance (R_s)

1) Fit at high V of forward characteristic
2) Exponential recovery time (afterpulses envelope)



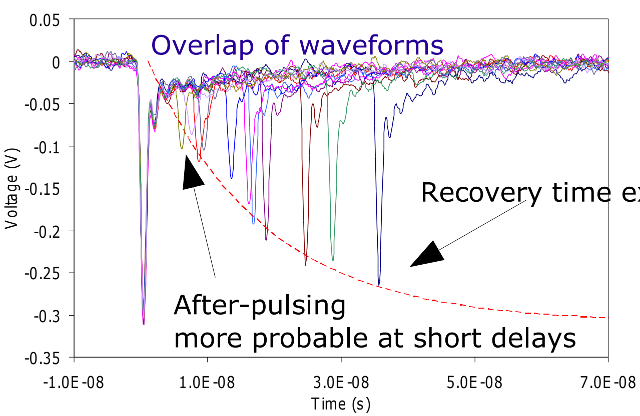
G.C. et al NIM A628 (2011) 389

Measurements (1) and (2) consistent → **dominant effect from quenching resistor R_q**
(→ series R bulk gives smaller contribution)

Empirical fit:

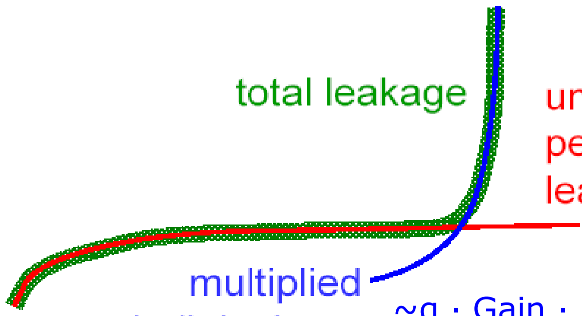
$$R_q(T) \sim 0.13\,(1+300/T\,e^{300/T})\,M\,\Omega$$

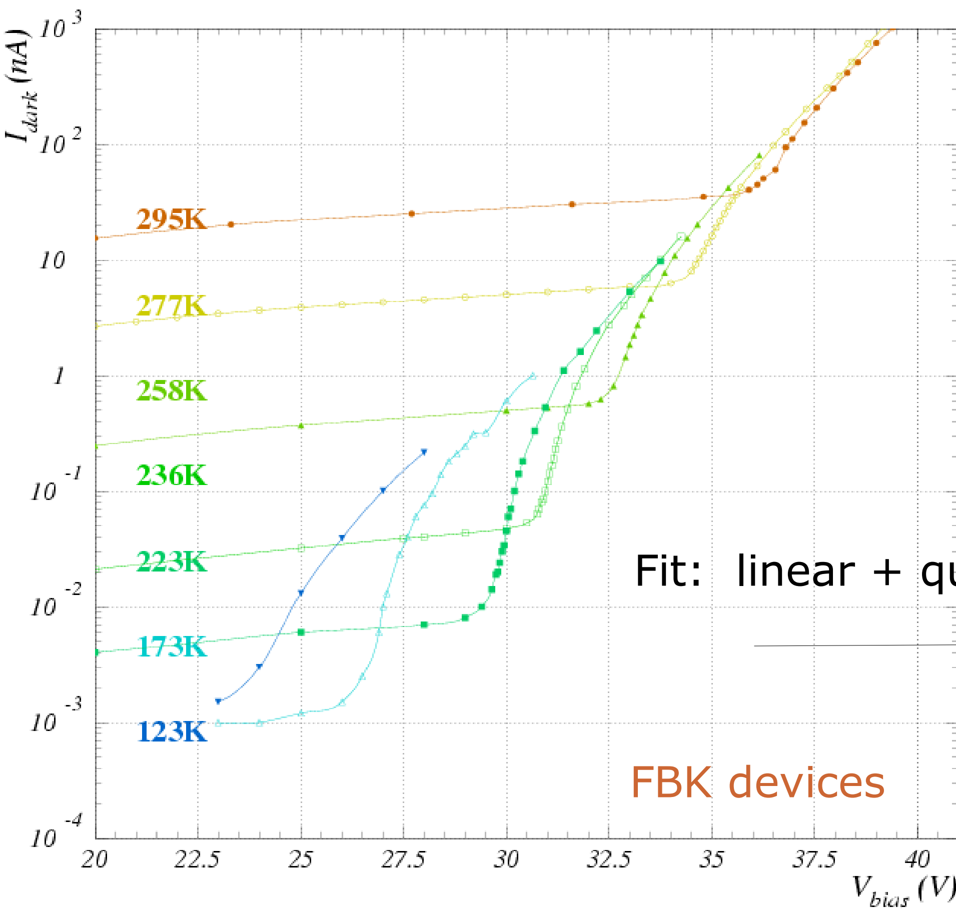
Afterpulses envelope



Note: SiPM for low T applications must have appropriate quenching R (not quenching at room T !)

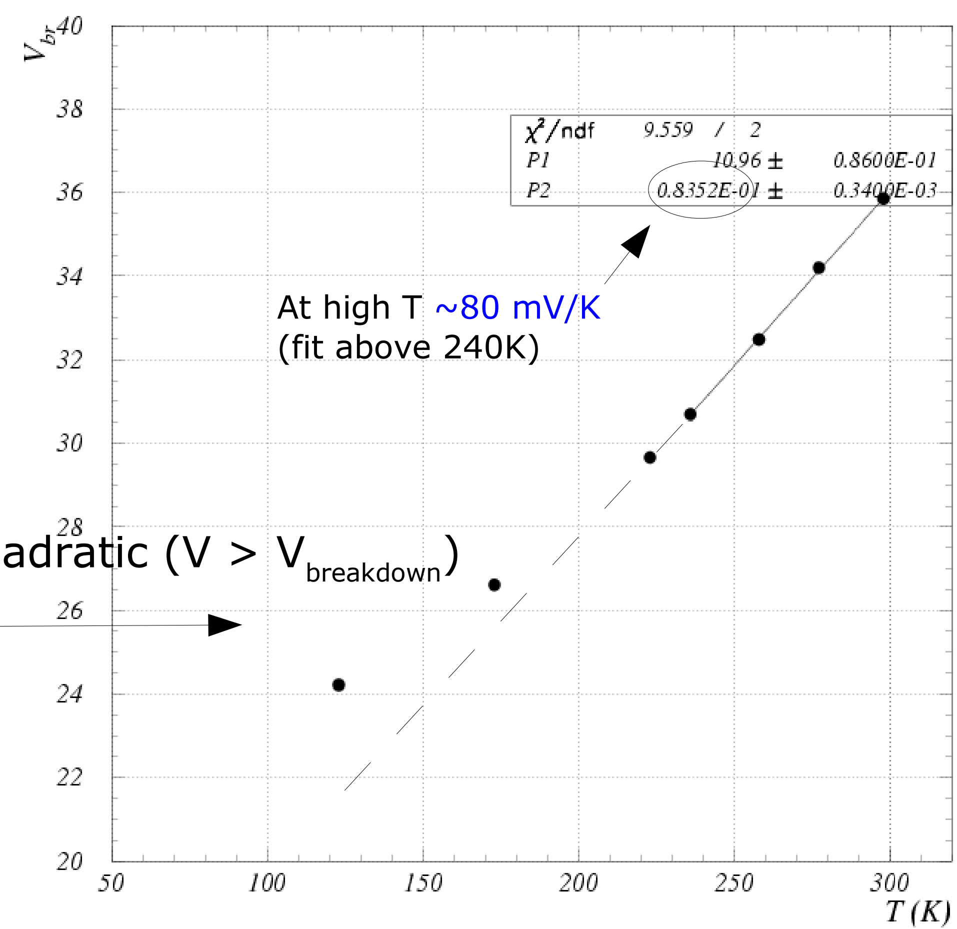
Reverse I-V





Fit: linear + quadratic (V > $V_{breakdown}$)

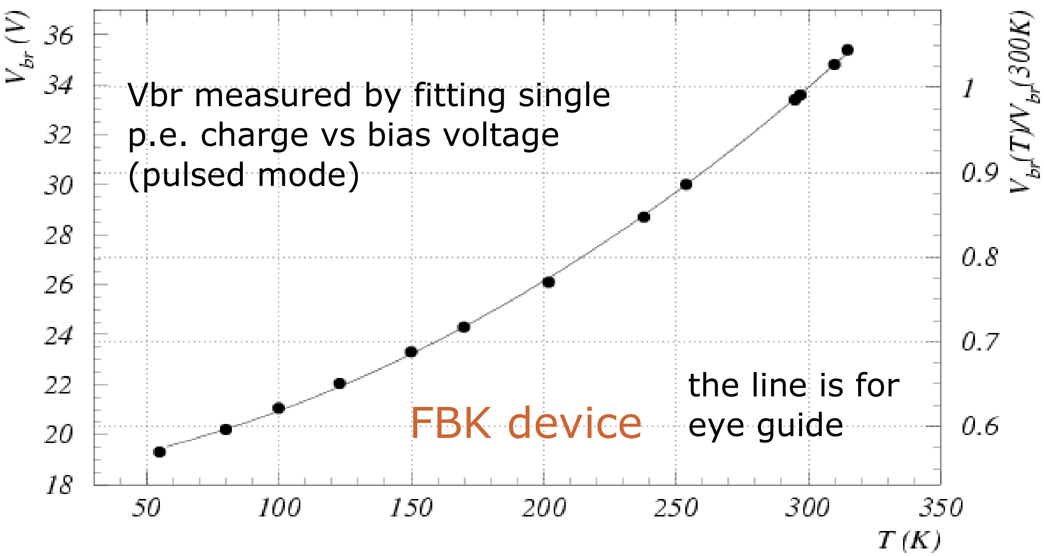
V_{bd} dependence on T



Breakdown voltage decreases at low T due to larger carriers mobility
→ larger ionization rate (electric E field fixed)

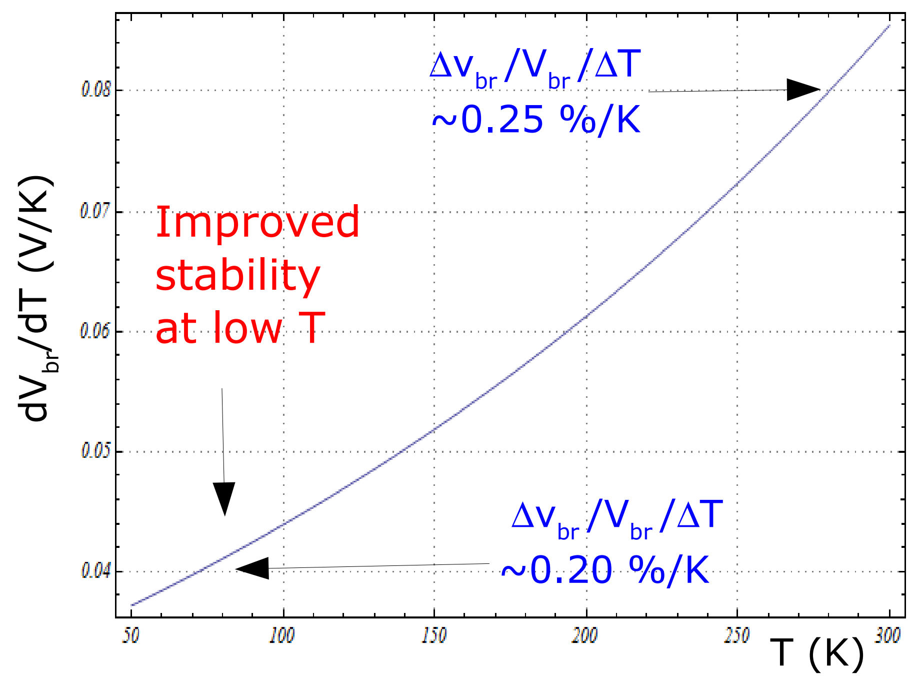
V_{bd} vs T → T coefficient (ΔV stability)

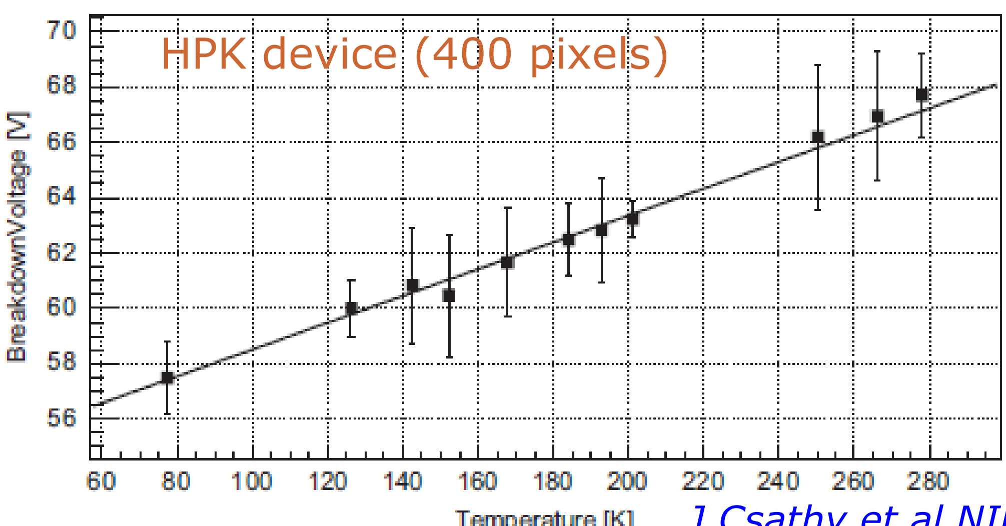
Breakdown Voltage



G.C. et al NIM A628 (2011) 389

Temperature coefficient

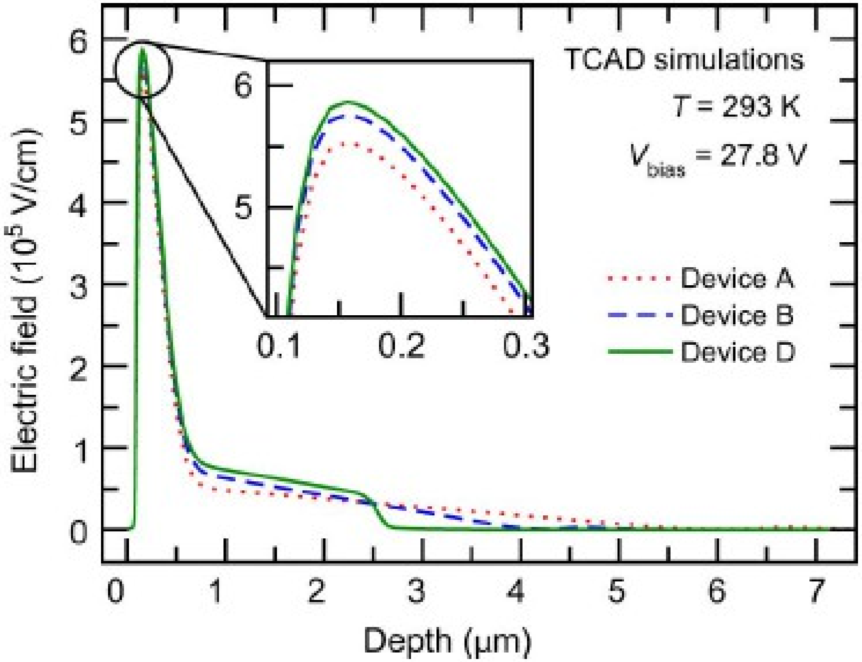




J.Csathy et al NIM A 654 (2011) 225

Fig. 6. Breakdown voltage as a function of temperature of the MPPC with 400 pixels.

Depletion layer → V_{bd} dependence on T



Serra et. al. (FBK) IEEE TNS 58 (2011) 1233
"Experimental and TCAD Study of Breakdown Voltage Temperature Behavior in n+/p SiPMs"

Note: precise agreement simulation/data is not trivial at all. Definition of ionization coefficients is device dependent...

Narrow depletion layer (high background doping(*) or thin epitaxial layer)
→ minimize V_{bd} dependence on T
→ gain stability $\frac{\delta V_{bd}/V_{bd}}{\delta T} = \frac{\delta G/G}{\delta T}$

(*) resulting in epitaxial layer not fully depleted at V_{bd}

Trade off:
→ PDE (thickness)
→ minimum gain (capacity) against after-pulses and cross-talk

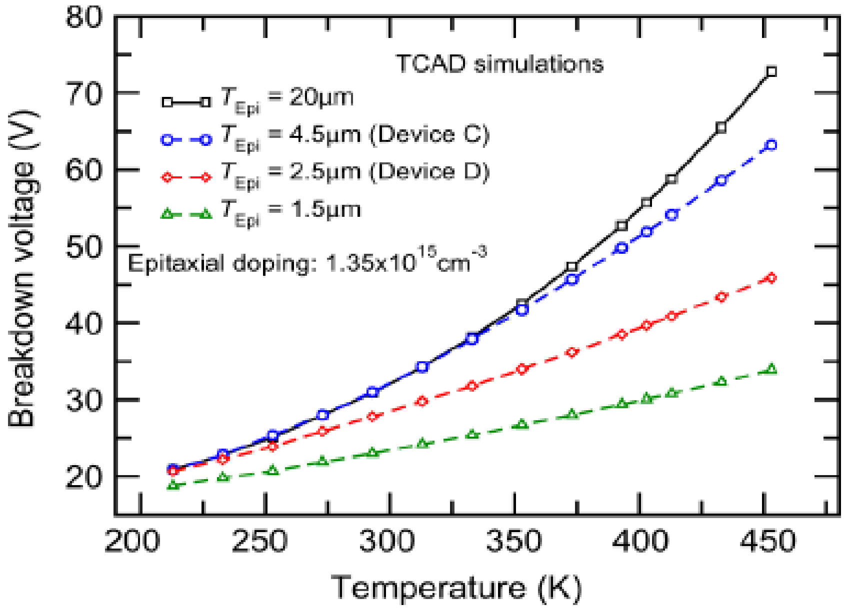


Fig. 9. TCAD simulated $V_{\rm BD}$ in the GM-APDs of this work (see Table I) in an extended temperature range. Two additional epitaxial layer thickness are considered (20 μm, 1.5 μm) to emphasize the impact of the depletion layer width on the $V_{\rm BD}$ vs. temperature characteristic.

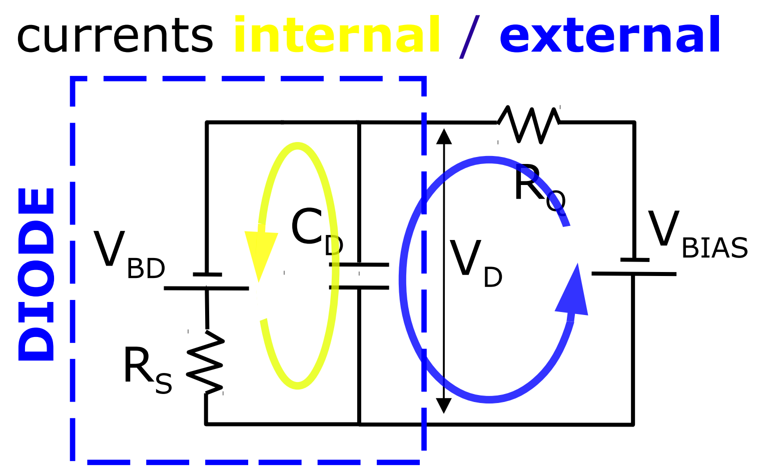
Pulse shape, Gain and Response

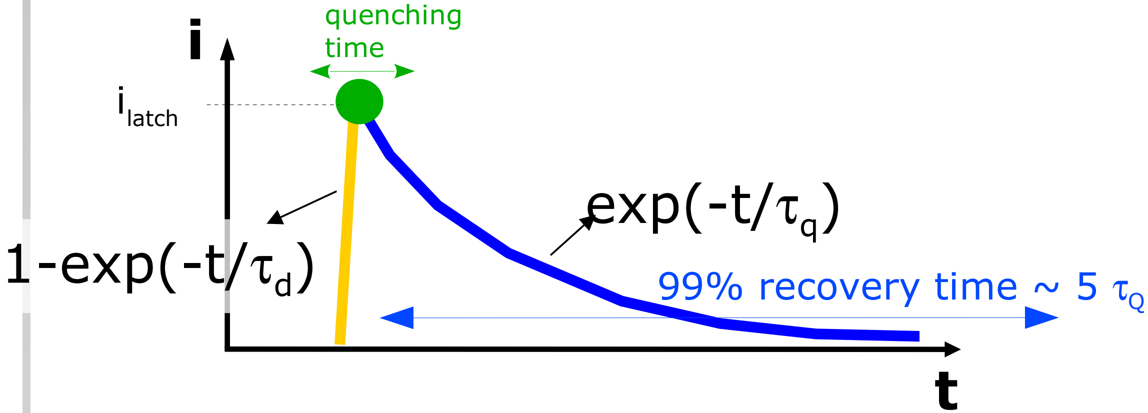
- Detailed electrical model
- Pulse shape
- Gain and Gain fluctuation
- Response non-linearity

(mostly for passive mode)

Basic electrical model

Fast Capacitor (cell) discharge and slow recharge (roughly speaking)





Rise time — Fall time (recovery)

$\tau_d = R_d C_d \ll \tau_q = R_q C_d$

Recovery time: T dependence due to R_q; C_d is independent of T

Rise time: T dependent (to lesser extent) due to R_d

Gain~C ΔV → independent of T

at fixed Over-Voltage ($\Delta V = V_{bias} - V_{bd}$)

SiPM equivalent circuit (detailed model)

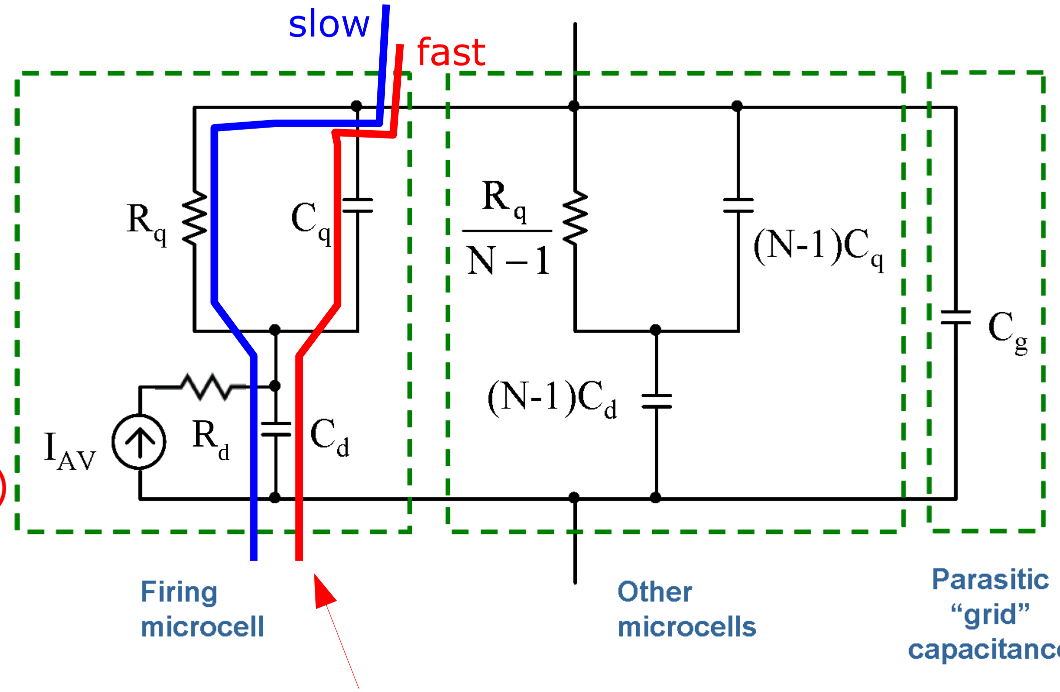
Single cell model → $(R_d||C_d)+(R_q||C_q)$

SiPM + load → $(||Z_{cell})||C_{grid} + Z_{load}$

Signal = slow pulse ($\tau_{d\ (rise)}, \tau_{slow\ (fall)}$) +
+ fast pulse ($\tau_{d\ (rise)}, \tau_{fast\ (fall)}$)

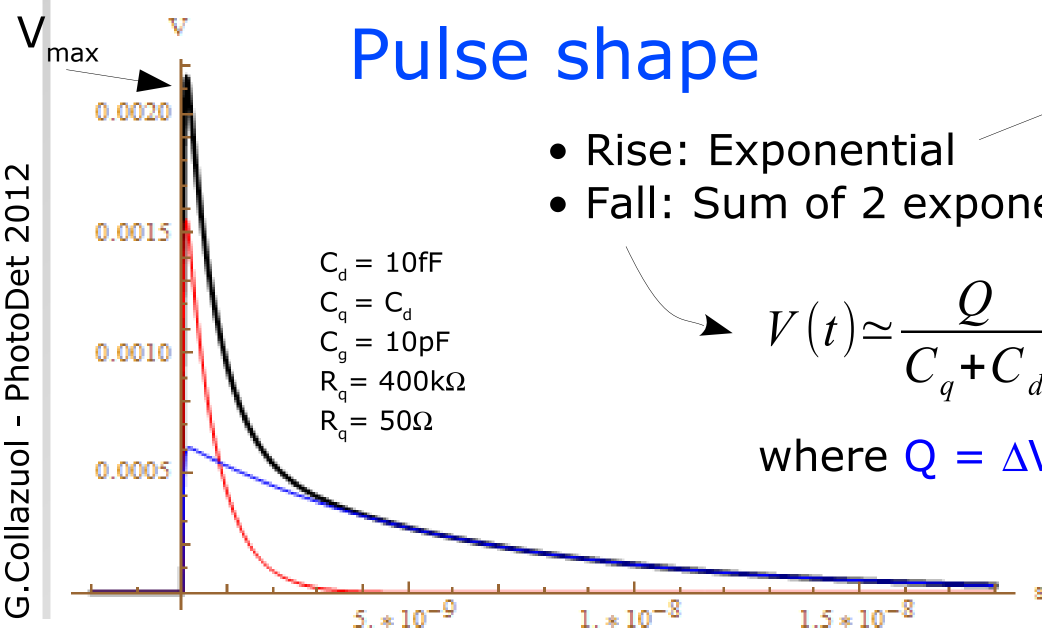
- $\tau_{d\ (rise)} \sim R_d(C_q+C_d)$
- $\tau_{fast\ (fall)} = R_{load}\ C_{tot}$ (fast; parasitic spike)
- $\tau_{slow\ (fall)} = R_q\ (C_q+C_d)$ (slow; cell recovery)

F.Corsi, et al. NIM A572 (2007) 416
S.Seifert et al. IEEE TNS 56 (2009) 3726



Cq → fast current supply path in the beginning of avalanche

Pulse shape



- Rise: Exponential → Sp.Charge Rd x Cd,q filtered by parasitic inductance, stray C, ... (Low Pass)
- Fall: Sum of 2 exponentials

$$V(t) \simeq \frac{Q}{C_q+C_d}\left(\frac{C_q}{C_{tot}} e^{\frac{-t}{\tau_{FAST}}} + \frac{R_{load}}{R_q}\frac{C_d}{C_q+C_d} e^{\frac{-t}{\tau_{SLOW}}}\right) \quad \text{for } R_{load} << Rq$$

where $Q = \Delta V\ (C_q+C_d)$ is the total charge released by the cell

→ 'prompt' charge on C_{tot} is $Q_{fast} = Q\ C_q/(C_q+C_d)$

Pulse shape

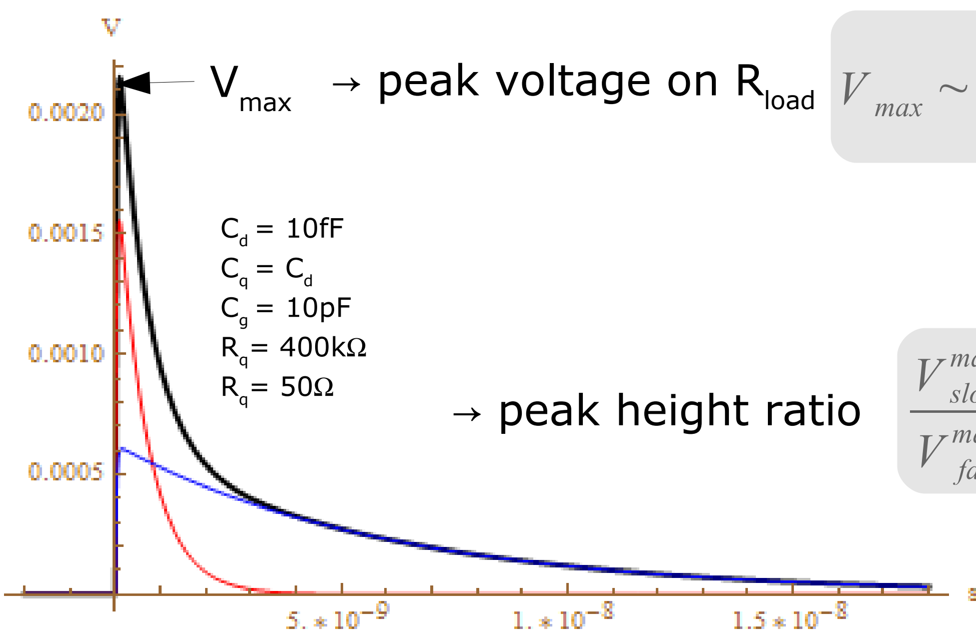
$$V(t) \simeq \frac{Q}{C_q+C_d}\left(\frac{C_q}{C_{tot}}\, e^{\frac{-t}{\tau_{fast}}} + \frac{R_{load}}{R_q}\frac{C_d}{C_q+C_d}\, e^{\frac{-t}{\tau_{slow}}}\right) = \frac{Q\,R_{load}}{C_q+C_d}\left(\frac{C_q}{\tau_{fast}}\, e^{\frac{-t}{\tau_{fast}}} + \frac{C_d}{\tau_{slow}}\, e^{\frac{-t}{\tau_{slow}}}\right)$$

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

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

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

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



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

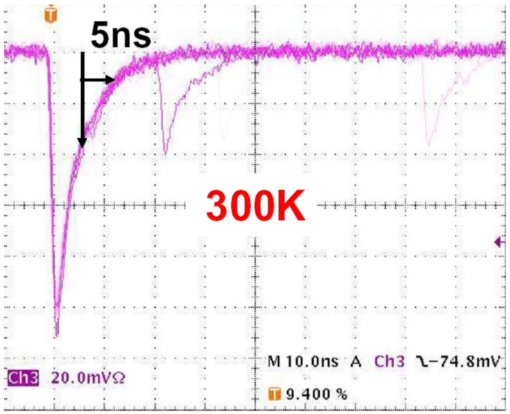
→ peak height ratio $\quad \frac{V^{max}_{slow}}{V^{max}_{fast}} \sim \frac{C_d\, C_{tot}\, R_{load}}{C_q^2\, R_q}$ increasing with C_d and $1/R_q$

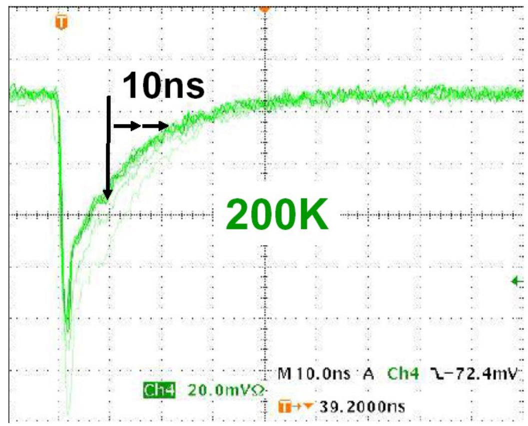
Pulse shape: dependence on Temperature

The two current components behave differently with Temperature

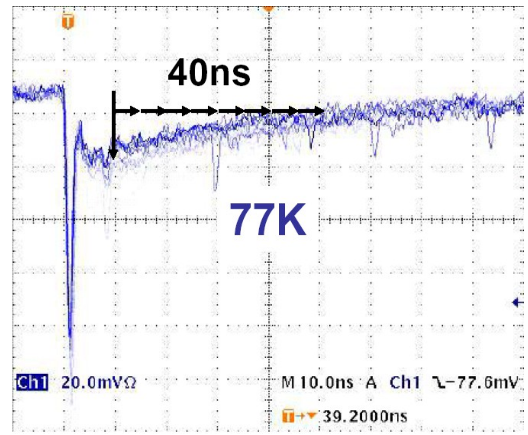
→ fast component is independent of T because C_{tot} couples to external R_{load}

→ slow component is dependent on T because $C_{d,q}$ couple to $R_q(T)$

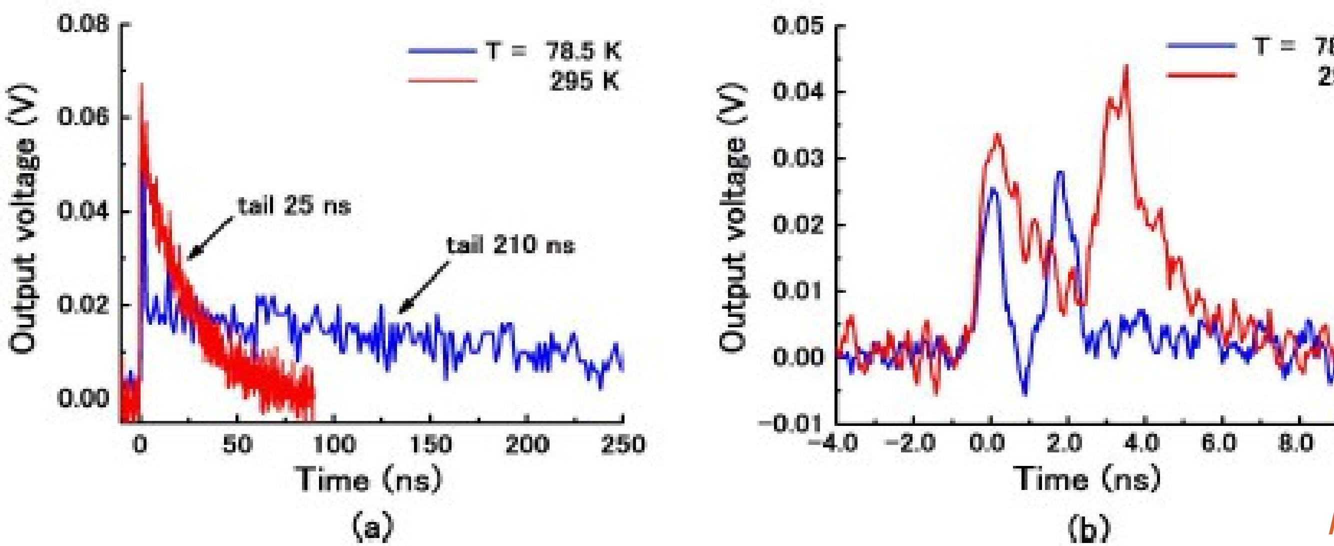




H.Otono, et al. PD07



HPK MPPC



high pass filter / shaping
→ recover fast signals

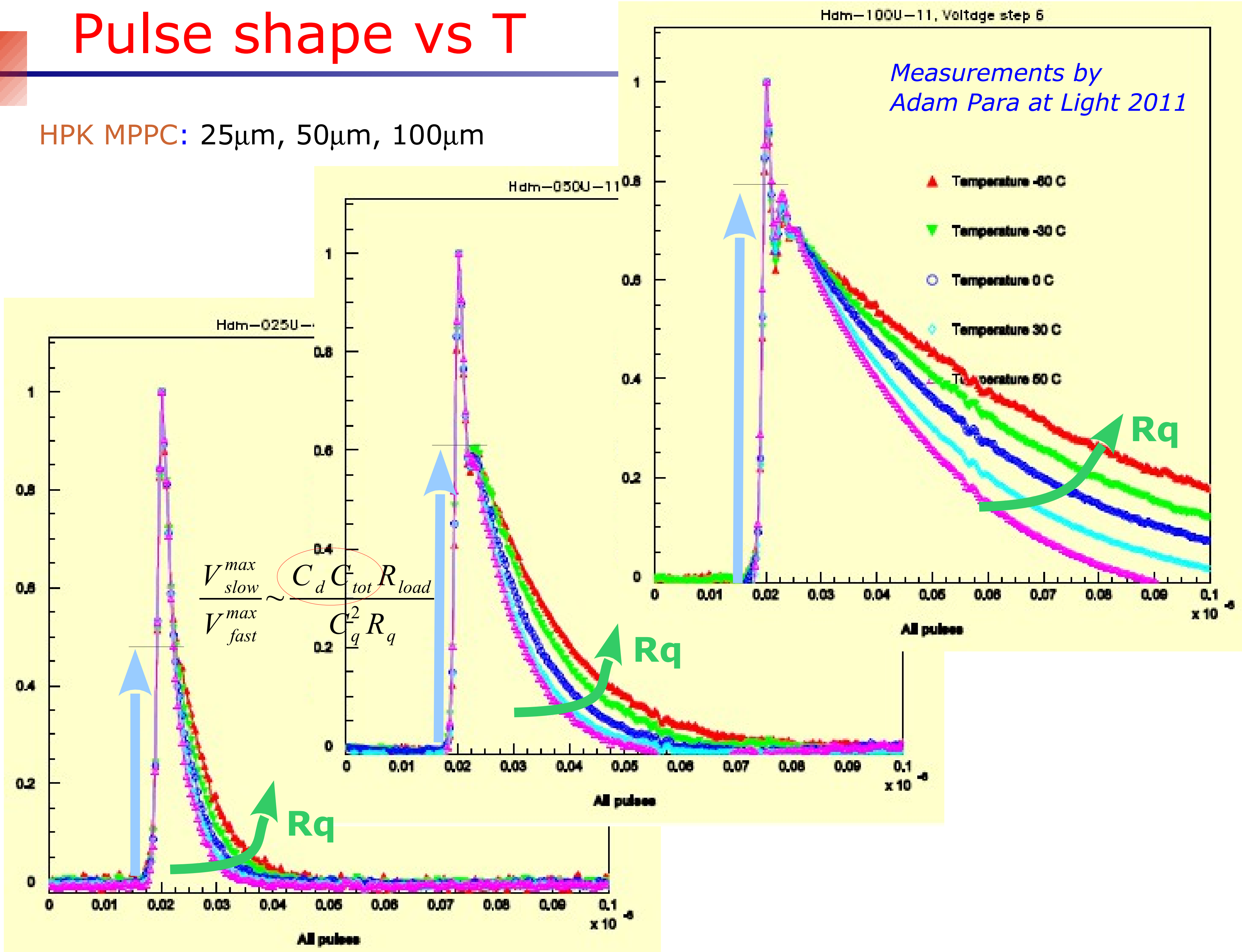
HPK MPPC

Fig. 2. (a) Output signals from the MPPC when no high-pass filter is used, and (b) output signals from the high-pass filter when two pulses were generated successively.

Akiba et al Optics Express 17 (2009) 16885

Pulse shape vs T

HPK MPPC: 25μm, 50μm, 100μm



$$\frac{V_{slow}^{max}}{V_{fast}^{max}} \sim \frac{C_d C_{tot} R_{load}}{C_q^2 R_q}$$

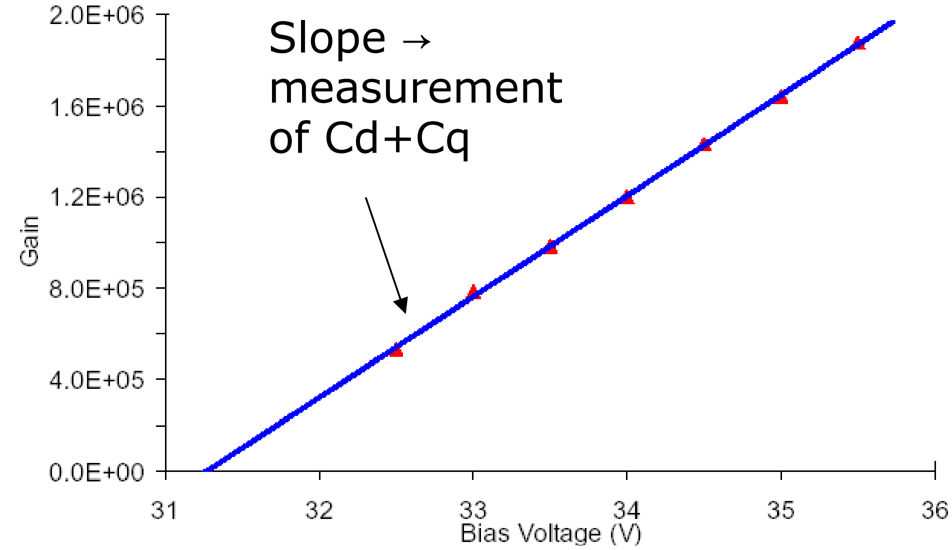
Gain and its Fluctuations

$$G = \Delta V \left(C_q + C_d\right) / q_e$$

→ Gain is linear if ΔV in quenching regime

but

there are many sources for non-linearity of response (non proportionality)

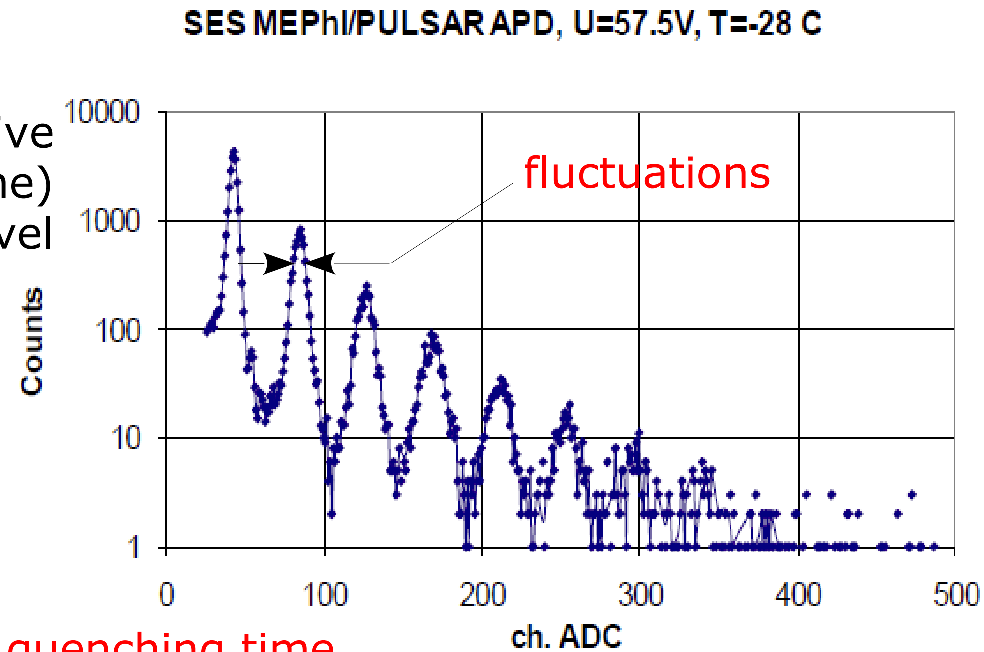


SiPM gain fluctuations (intrinsic) differ in nature compared to APD where the statistical process of internal amplification shows a characteristic fluctuations

$$\frac{\delta G}{G} = \frac{\delta V_{bd}}{V_{bd}} \oplus \frac{\delta C_{dq}}{C_{dq}}$$

cell to cell uniformity (active area and volume) control at % level

- doping densities (Poisson): $\delta V_{bd} \geq 0.3V$ *Shockley, Sol. State Ele. 2 (1961) 35*
- doping, epitaxial, oxide (processing): $\delta V_{bd} \sim O(0.1V)$



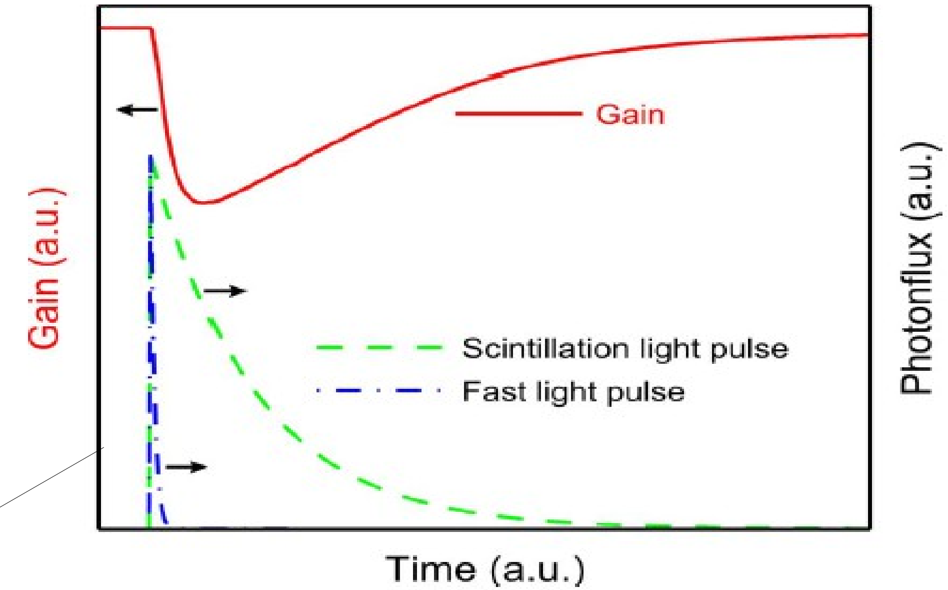
In addition δG might be due to fluctuations in quenching time

… and of course after-pulses contribute too (not intrinsic → might be corrected)

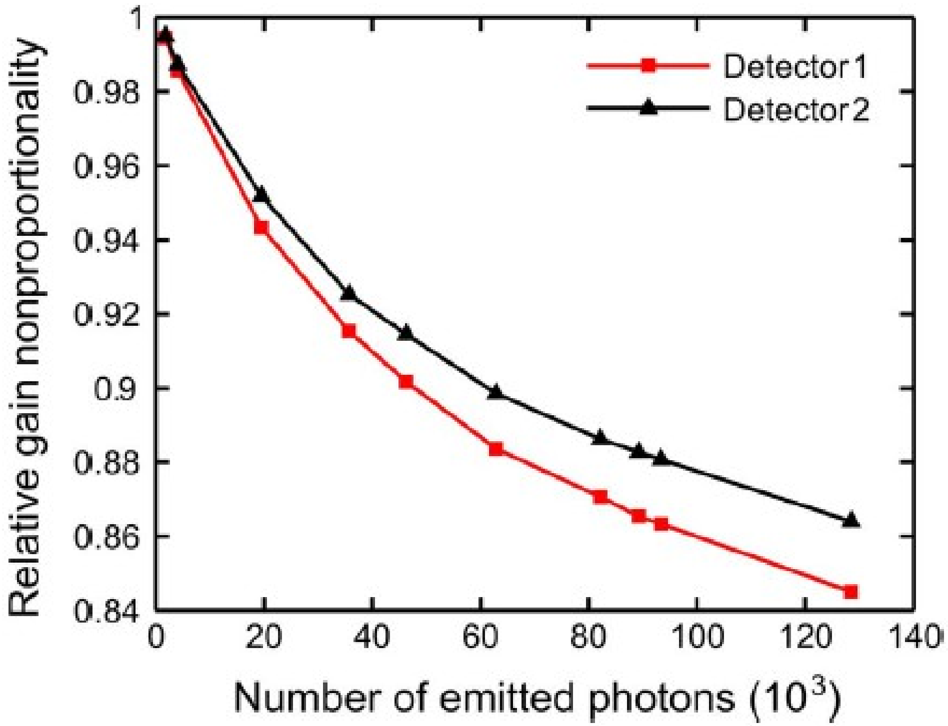
Response Non-Linearity

Non-proportionality of charge output w.r.t. number of photons (i.e. response) at level of several % might show up even in quenching regime (negligible quenching time), depending on ΔV and on the intensity and duration of the light pulse.

Main sources are:
- finite number of pixels
- finite recovery time w.r.t. pulse duration
- after-pulses, cross-talk
- drop of ΔV during the light pulse due to relevant signal current on (large) series resistances (eg ballast)



T.van Dam IEEE TNS 57 (2010) 2254
Detailed model to estimate non-lin. corrections



Finite number of cells is main contribution in case number of photons ~ O(number of cells) (dynamic range not adequate to application)

→ saturation $n_{fired} = n_{all}\left(1 - e^{-\frac{n_{phot.}PDE}{n_{all}}}\right)$

→ loss of energy resolution

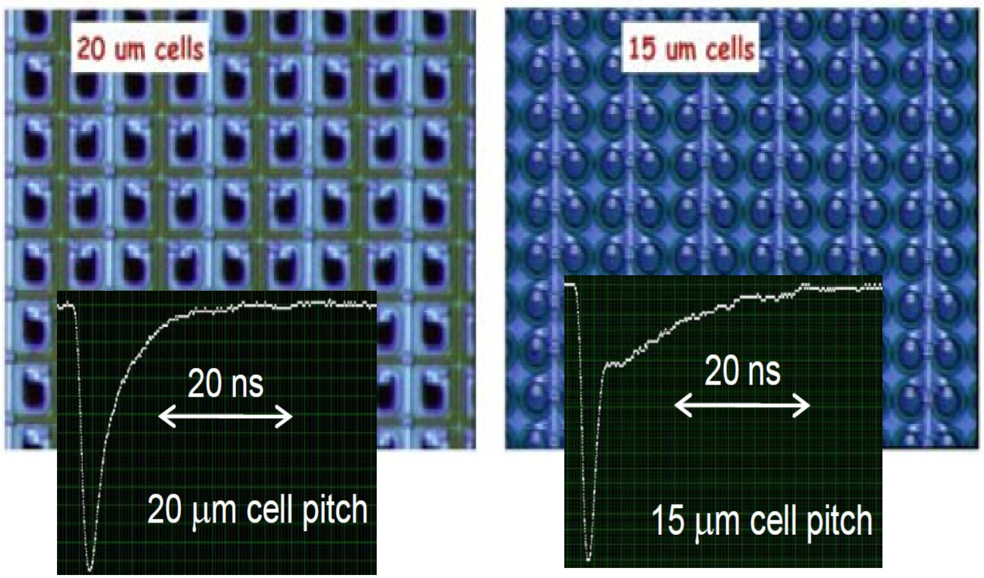
see Stoykov et al JINST 2 P06500 and Vinogradov et al IEEE NSS 2009 N28-3

New high dynamic range SiPMs

Different types available or in preparation:

- tiny cells
  → HPK, FBK, NDL, MPI-LL
- micro cells
  → Zecotek, AmplititationTech

Latest MPPC tiny cell by Hamamatsu

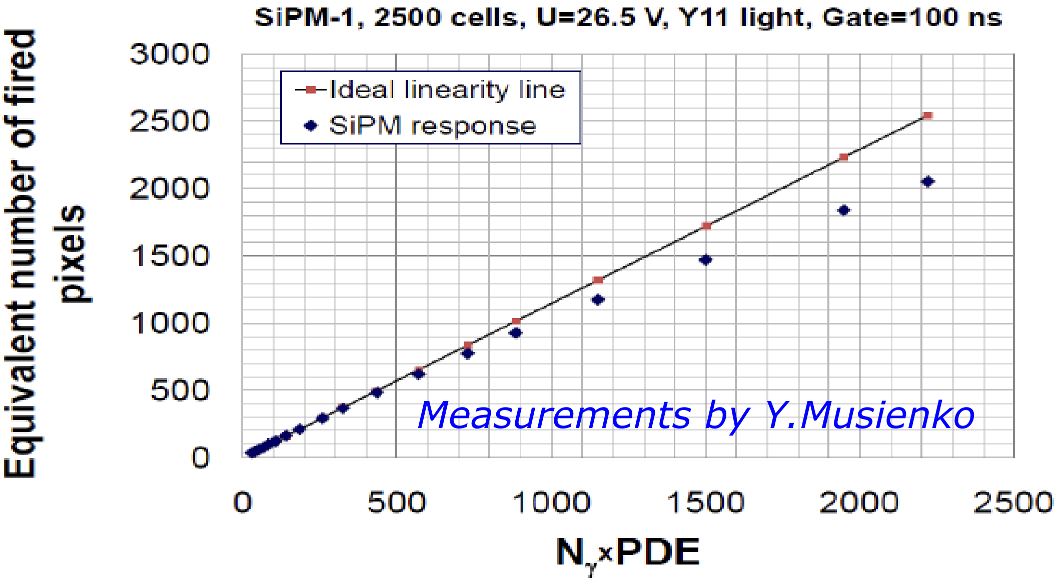


SiPMs NDL (Bejiing)

Zhang et al NIM A621 (2010) 116
Han at NDIP 2011

- type: n-on-p, Bulk Rq
- high cell density (10000/mm²)
- fast recovery (5ns)
- low gain

→ dynamic range (high cell density, fast recovery)
→ radiation hardness (low gain)



Measurements by Y.Musienko

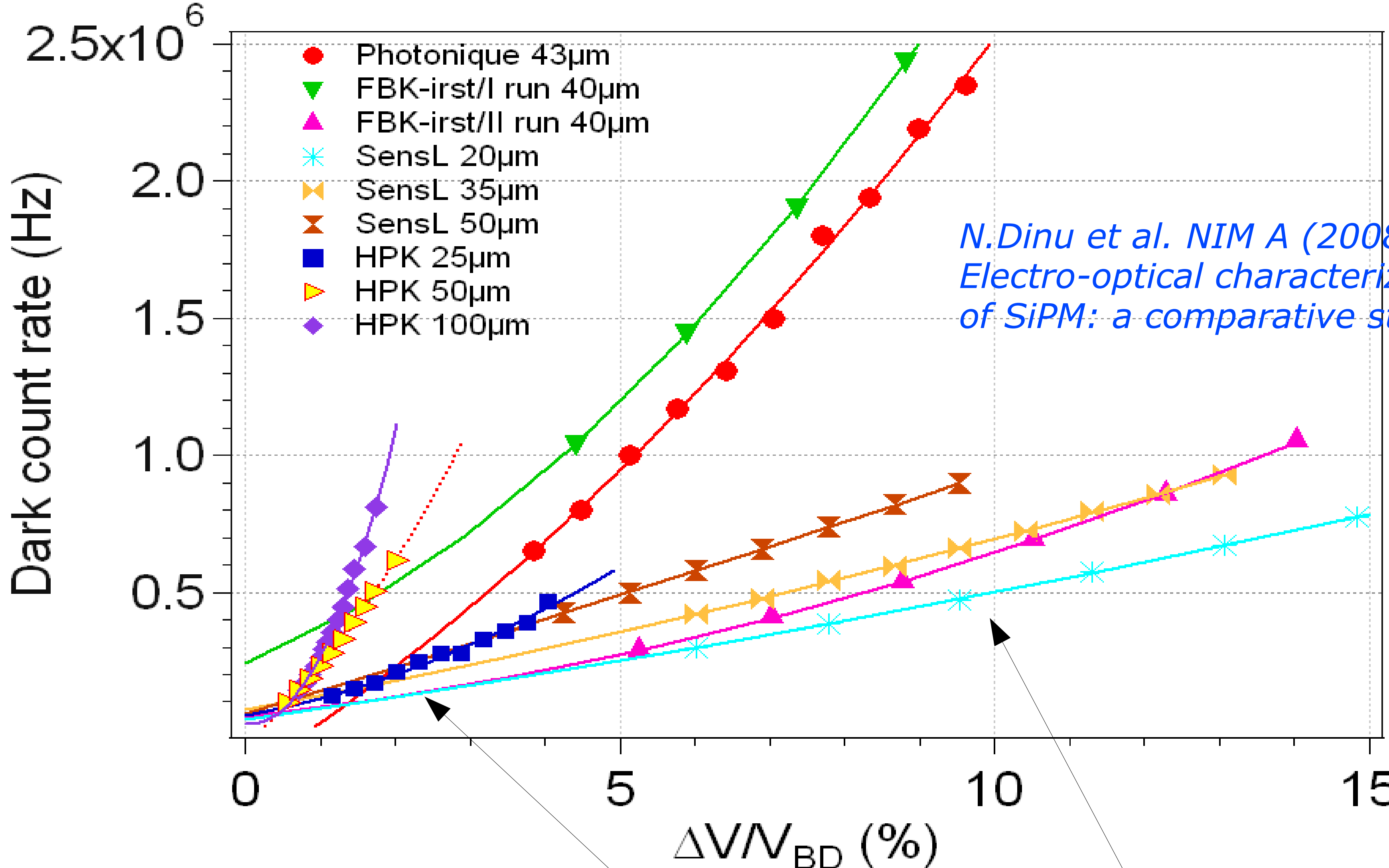
Noise sources:

Dark counts — pulses triggered by non-photo-generated carriers (thermal / tunneling generation in the bulk or in the surface depleted region around the junction)

After-pulsing — carriers can be trapped during an avalanche and then released triggering another avalanche

Cross-Talk "optical" — photo-generation during the avalanche discharge. Some of the photons can be absorbed in the adjacent cell possibly triggering new discharges

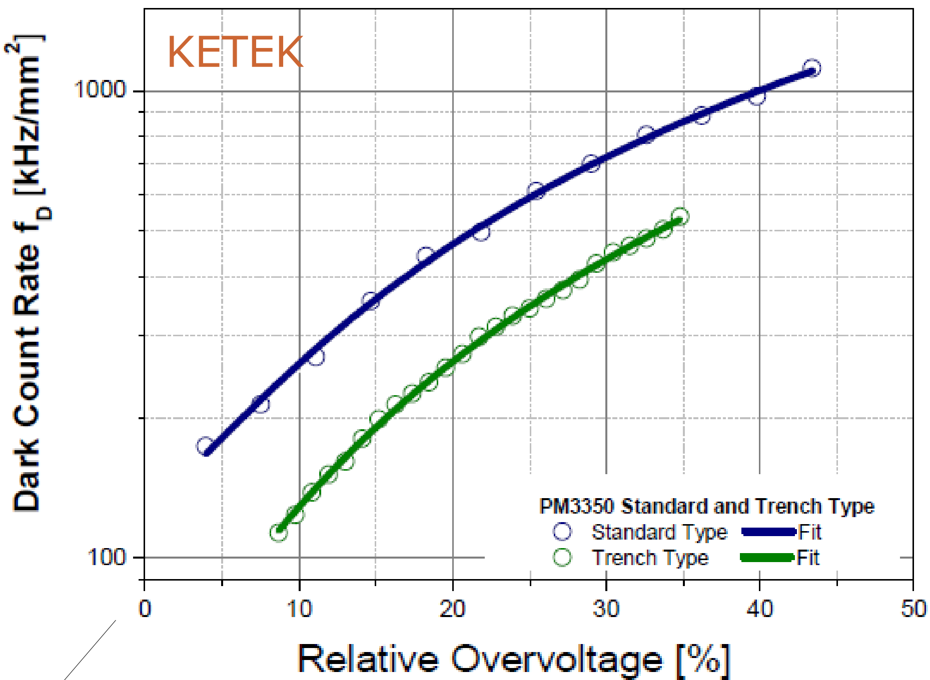
Dark Count Rate



- DCR → linear dependence due to $P_{01} \propto \Delta V$ (→ same as PDE vs ΔV)
  → non-linear at high ΔV due to cross-talk and after-pulsing → $\propto \Delta V^2$
- DCR scales with active surface (not with volume: high field region dominating)

Dark Count Rate

KETEK PM 3350 (p^+-on-n, shallow junction)
3x3mm² active area pixel size 50x50 μm²



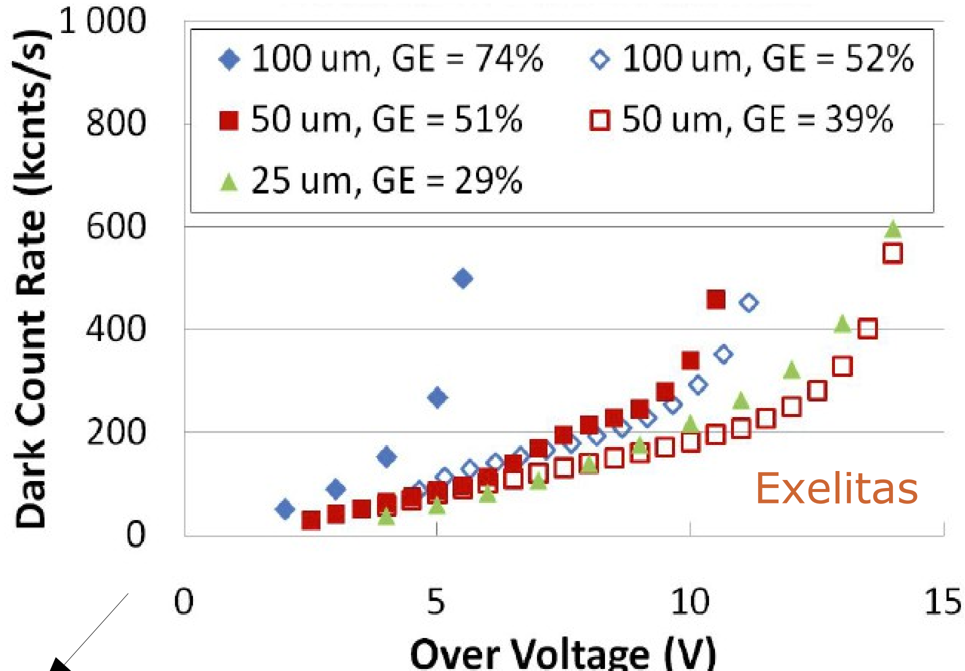
V_{bd} ~ 25V

F.Wiest – AIDA 2012 at DESY

Critical issues:

- quality of epitaxial layer
- gettering techniques
- Efield engineering (low T)

Exelitas 1st generation SiPM 2011
(p^+-on-n) 1x1mm²



V_{bd} ~ 140V

P.Berard – NDIP 2011

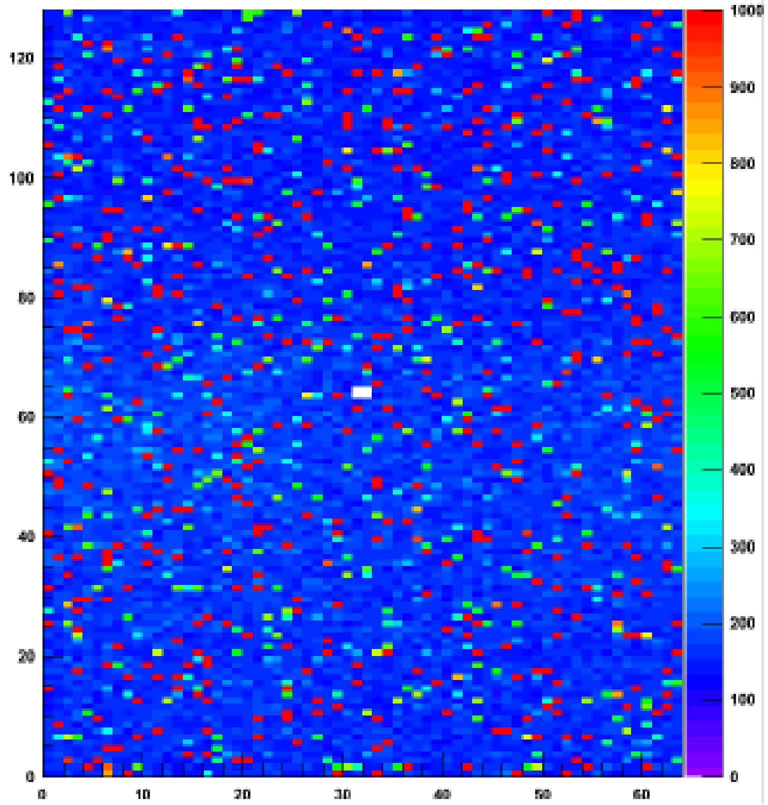
Latest Hamamatsu devices reached ~80kHz/mm²

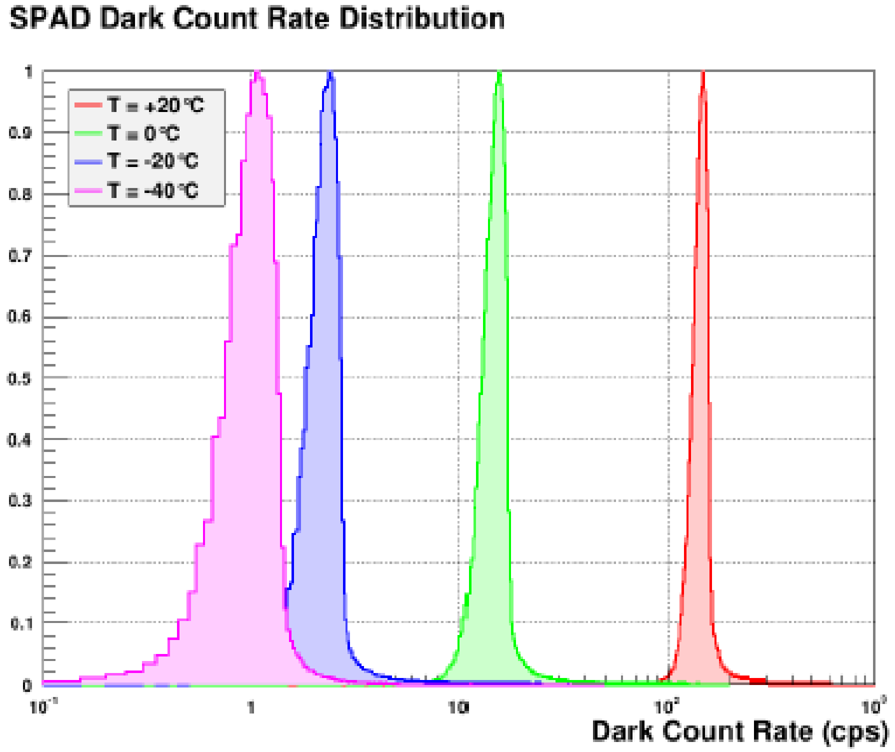
HPK claiming for additional improvements coming (*HPK at LIGHT 2011*)

Dark Count Rate

dSiPM

Control over individual SPADs enables detailed device characterization





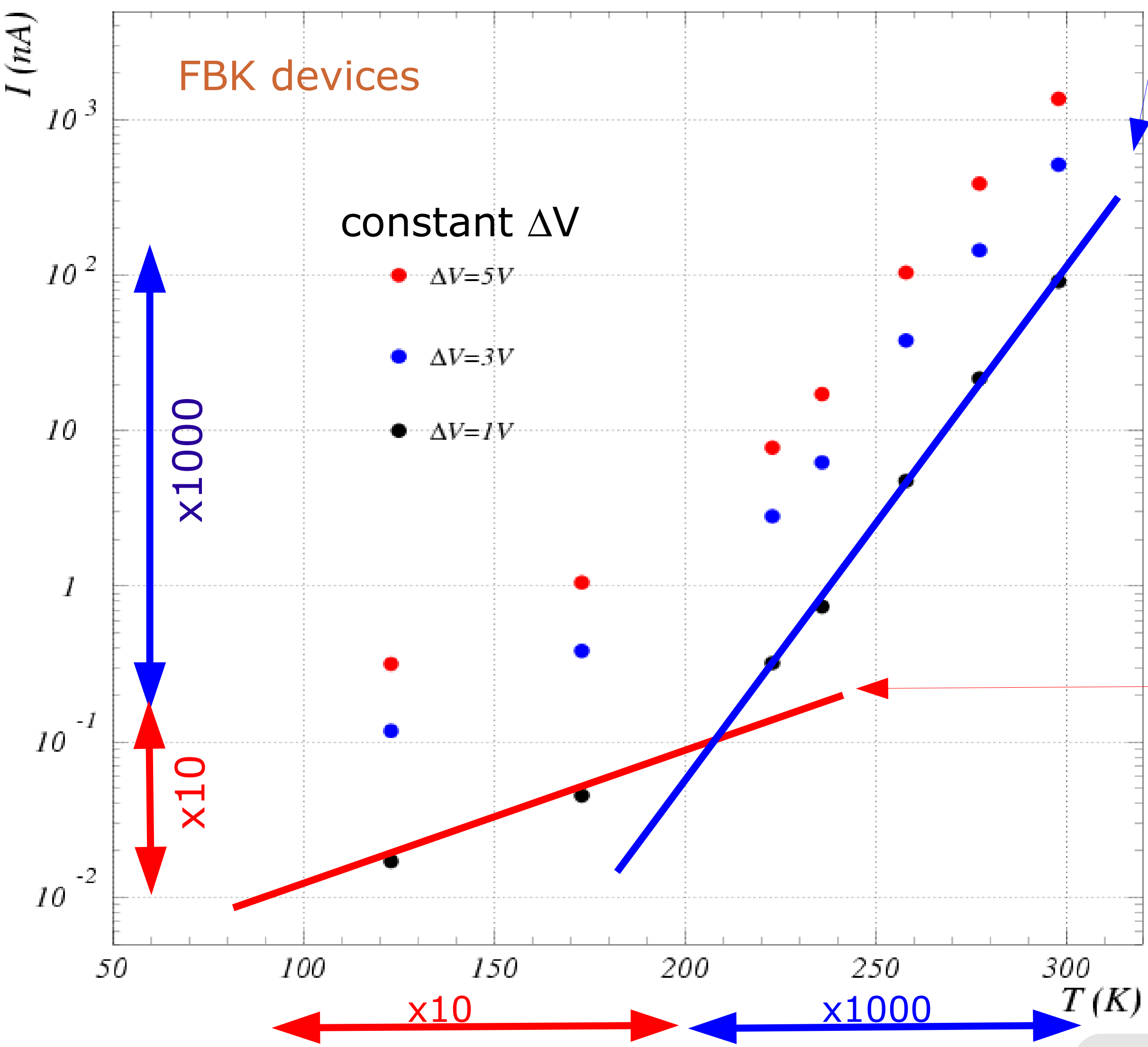
- Over 90% good diodes (dark count rate close to average)
- Typical dark count rate at 20°C and 3.3V excess voltage: ~150cps / diode
- Low dark counts (~1-2cps) per diode at -40°C

T.Frach at NDIP 2011

Dark current vs T sources of DCR

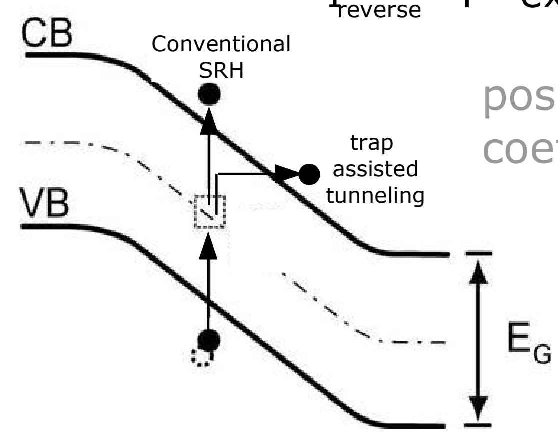
contribution to DCR from diffusion of minority carriers negligible below 350K

Noise mainly comes from the high E Field region (no whole depletion region)



1) Generation/Recombination SRH noise (enhanced by trap assisted tunneling)

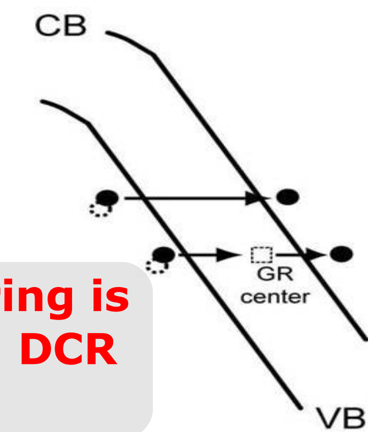
$$I_{reverse} \sim T^{1.5} \exp \frac{-E_{act}}{K_B T}$$



positive T coefficient

2) Band-to-band Tunneling noise (strong dependence on the Electric field profile)

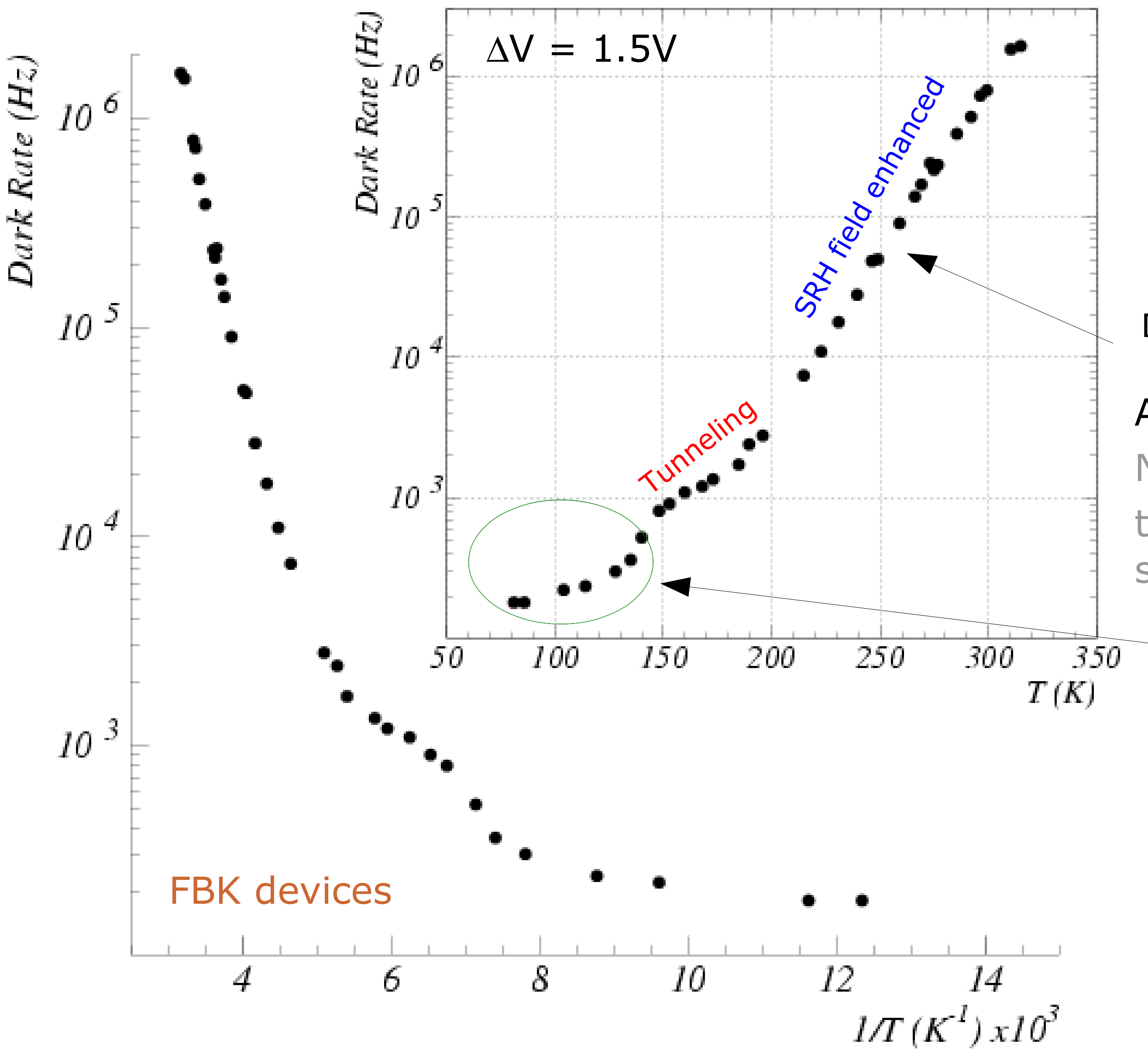
negative T coefficient



Tunneling noise dominating for T<200K (FBK devices have E field quite peaked)

Efield engineering is crucial for min. DCR (esp. at low T)

Dark Count Rate vs T (constant ΔV)



Measurement of **counting rate of ≥1p.e.** at fixed ΔV=1.5V (→ constant gain)

$$DCR \sim T^{1.5} \exp\frac{-E_{act}}{2K_B T}$$

Activation energy E_{act}~0.72eV

Note: E_{act} should be ~ E_g but tunneling makes effective gap smaller

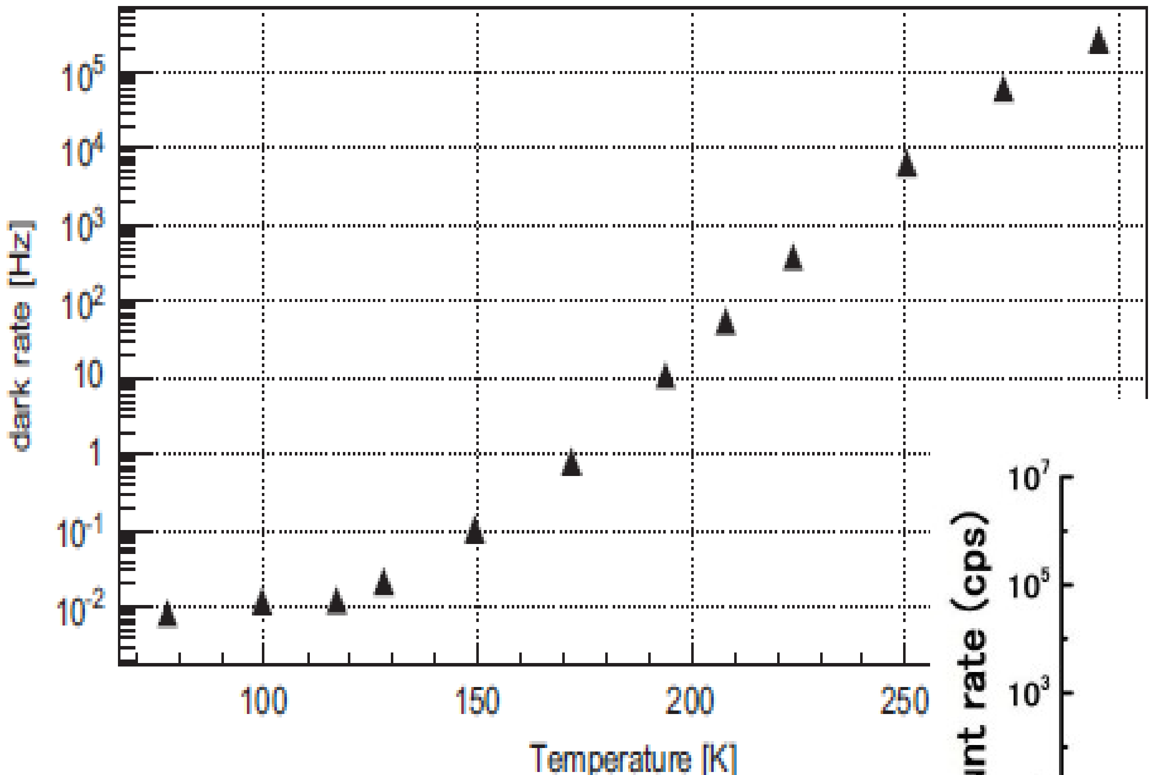
Additional structure carriers freeze-out (?)

(carrier collection losses at very low T due to ionized impurities acting as shallow traps → drop in PDE)

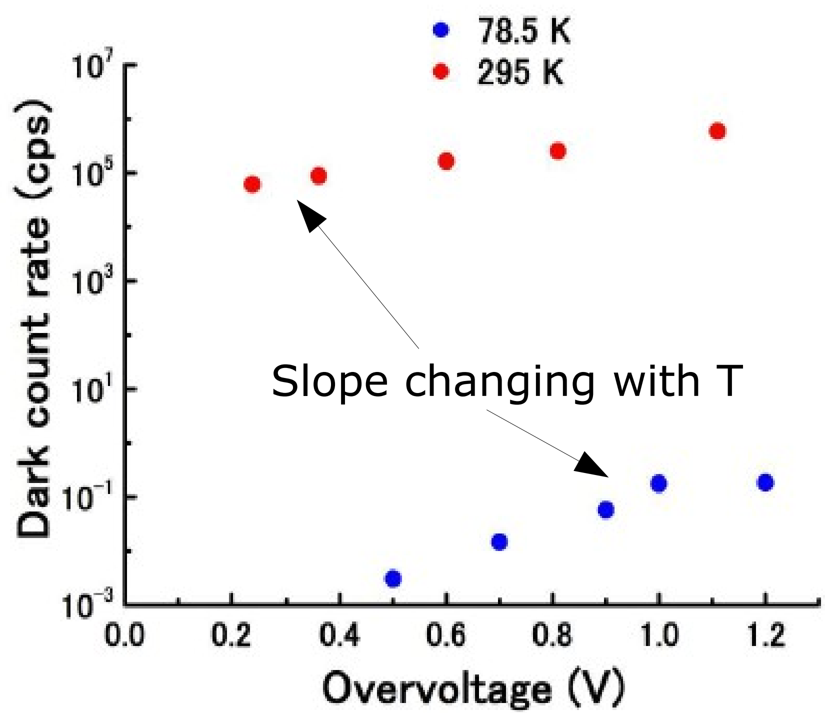
G.C. et al NIM A628 (2011) 389

Dark Count Rate vs T

Hamamatsu
(100μm pixels)

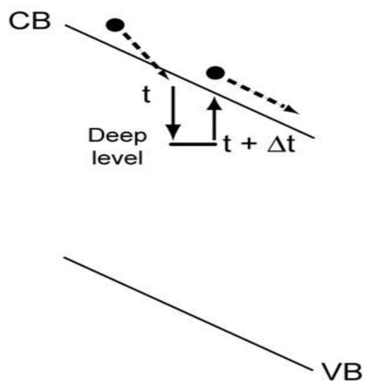


J.Csathy et al NIM A 654 (2011) 225



Comprehensive MPPC characterization at low T → *Akiba et al Optics Express 17 (2009) 16885*

After-Pulsing Carrier trapping and delayed release



$$P_{afterpulsing}(t) = P_c \cdot \frac{\exp(-t/\tau)}{\tau} \cdot P_{01} \quad \propto \Delta V^2$$

~Few % level at 300K

avalanche triggering probability
$\propto \Delta V(t)$

τ : trap lifetime
depends on trap level position

P_c : trap capture probability
$\propto$ carrier flux (current) during avalanche $\propto \Delta V$
$\propto$ N traps

quadratic dependence on ΔV

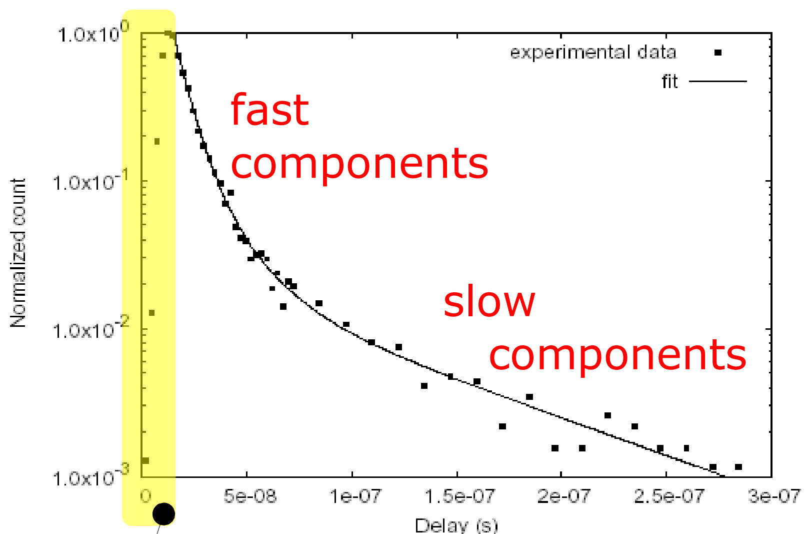
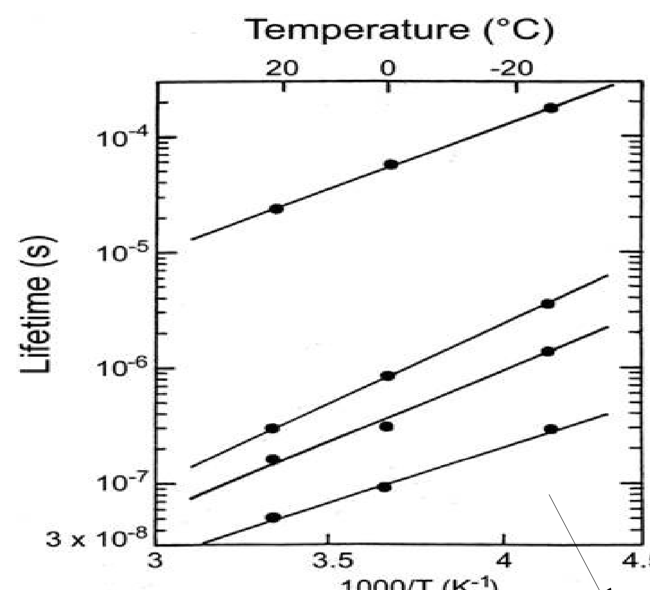


Fig. 10. Spectrum of the delay time from the primary pulse to the after-pulse.



S.Cova, A.Lacaita, G.Ripamonti, IEEE EDL (1991)

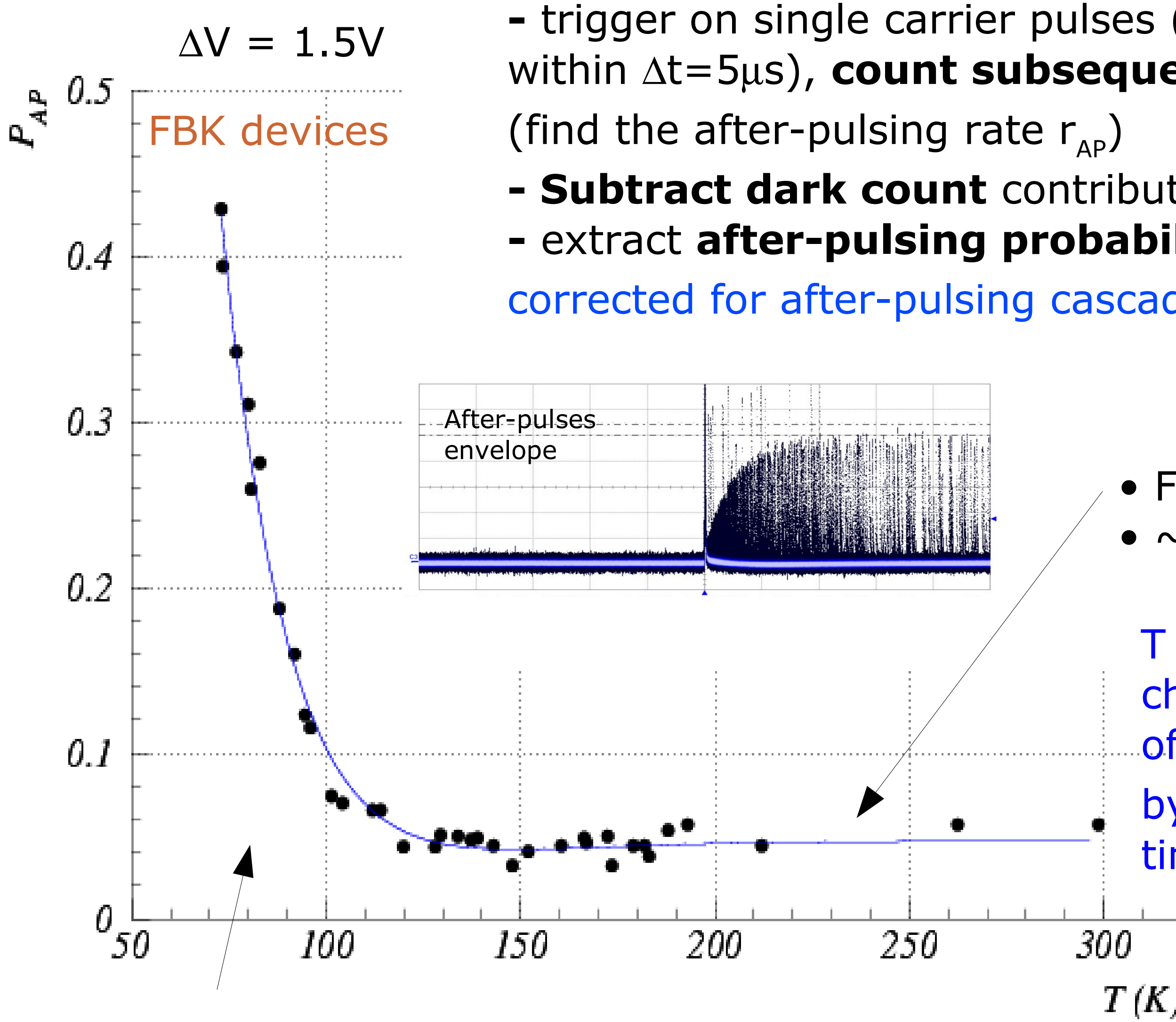
not trivial dependence on T

Only partially sensitive to after-pulsing during recovery
ie recovery hides After-pulses (does not cancel them)

After-Pulses vs T (constant ΔV)

Measurement by waveform analysis:
- trigger on single carrier pulses (with no preceding pulses within Δt=5μs), **count subsequent pulses** within Δt=5μs (find the after-pulsing rate r_{AP})
- **Subtract dark count** contribution
- extract **after-pulsing probability P_{AP}** corrected for after-pulsing cascade

$$P_{AP} = \frac{r_{AP}}{1 + r_{AP}}$$



- Few % at room T
- ~constant down to ~120K

T decreasing: increase of characteristic time constants of traps (τ_{traps}) compensated by increasing cell recovery time (R_q)

- several % below 100K

T<100K: additional trapping centers activated possibly (?) related to onset of carriers freeze-out

→ Analysis of life-time evolution vs T of the various traps (at least 3 types at T_{room})

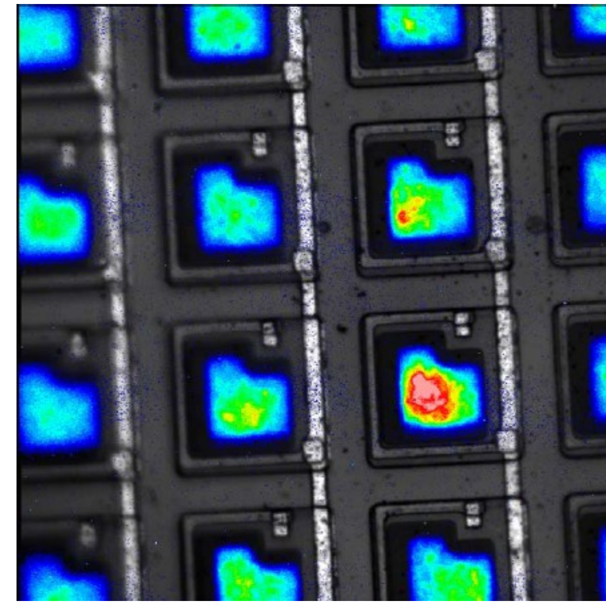
G.C. et al NIM A628 (2011) 389

G.Collazuol - PhotoDet 2012

Optical cross-talk Avalanche luminescence (NIR)

Carriers' luminescence (spontaneous direct relaxation in the conduction band) during the avalanche: probability 3.10^{-5} per carrier to emit photons with E> 1.14 eV

A.Lacaita et al. IEEE TED (1993)

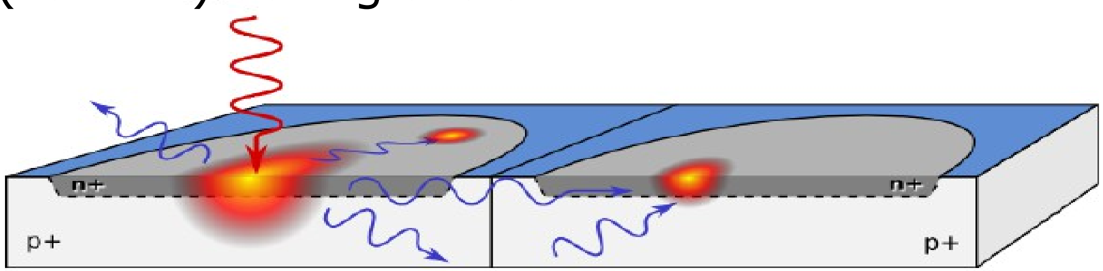


N.Otte, SNIC 2006

Photons can induce avalanches in neighboring cells. Depends on distance between high-field regions

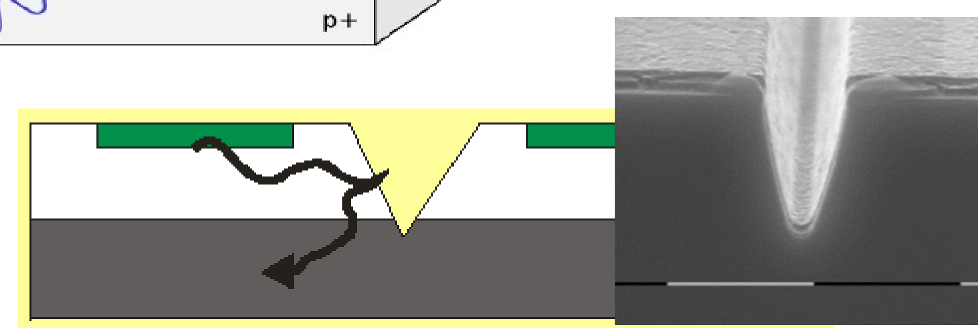
ΔV^2 dependence on over-voltage:

- carrier flux (current) during avalanche $\propto \Delta V$
- gain $\propto \Delta V$



Counteract:

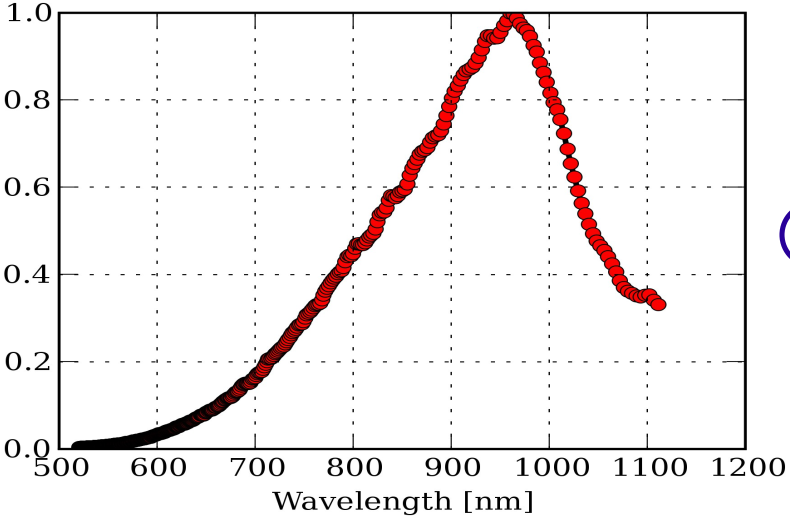
- optical isolation between cells by trenches filled with opaque material
- low over-voltage operation helps



It can be reduced to a level below % in a wide ΔV range

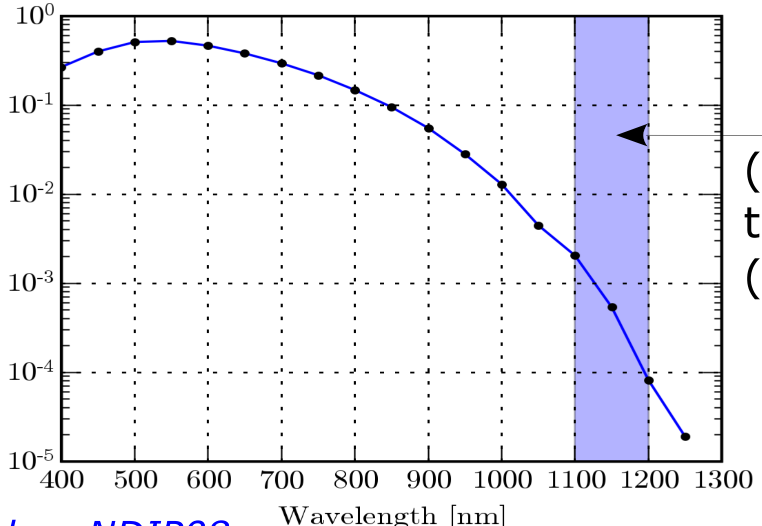
Optical cross-talk:reflections from the bottom

Measured Emission spectrum



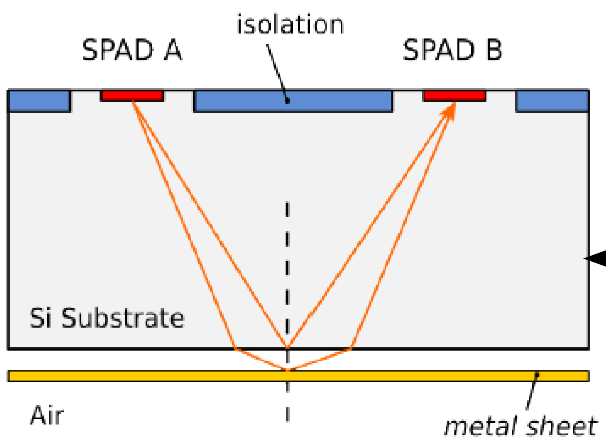
⊗

PDE



(1) Cross-talk due to narrow λ range (<100nm)

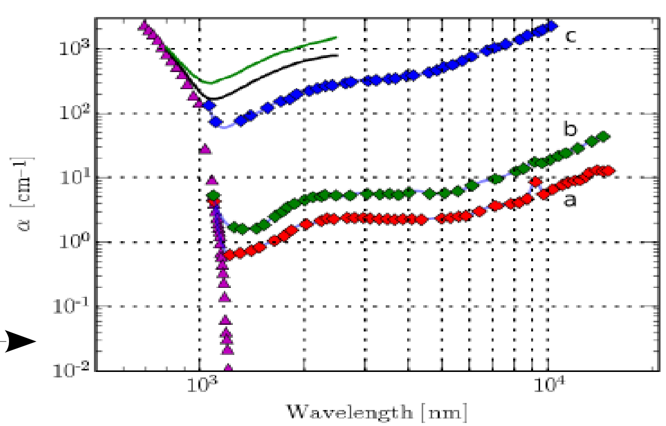
A.Ingargiola – NDIP08
Rech et al Proc. of SPIE Vol. 6771 677111-1



(2) Main component due to total reflection internal from the bottom (substrate)

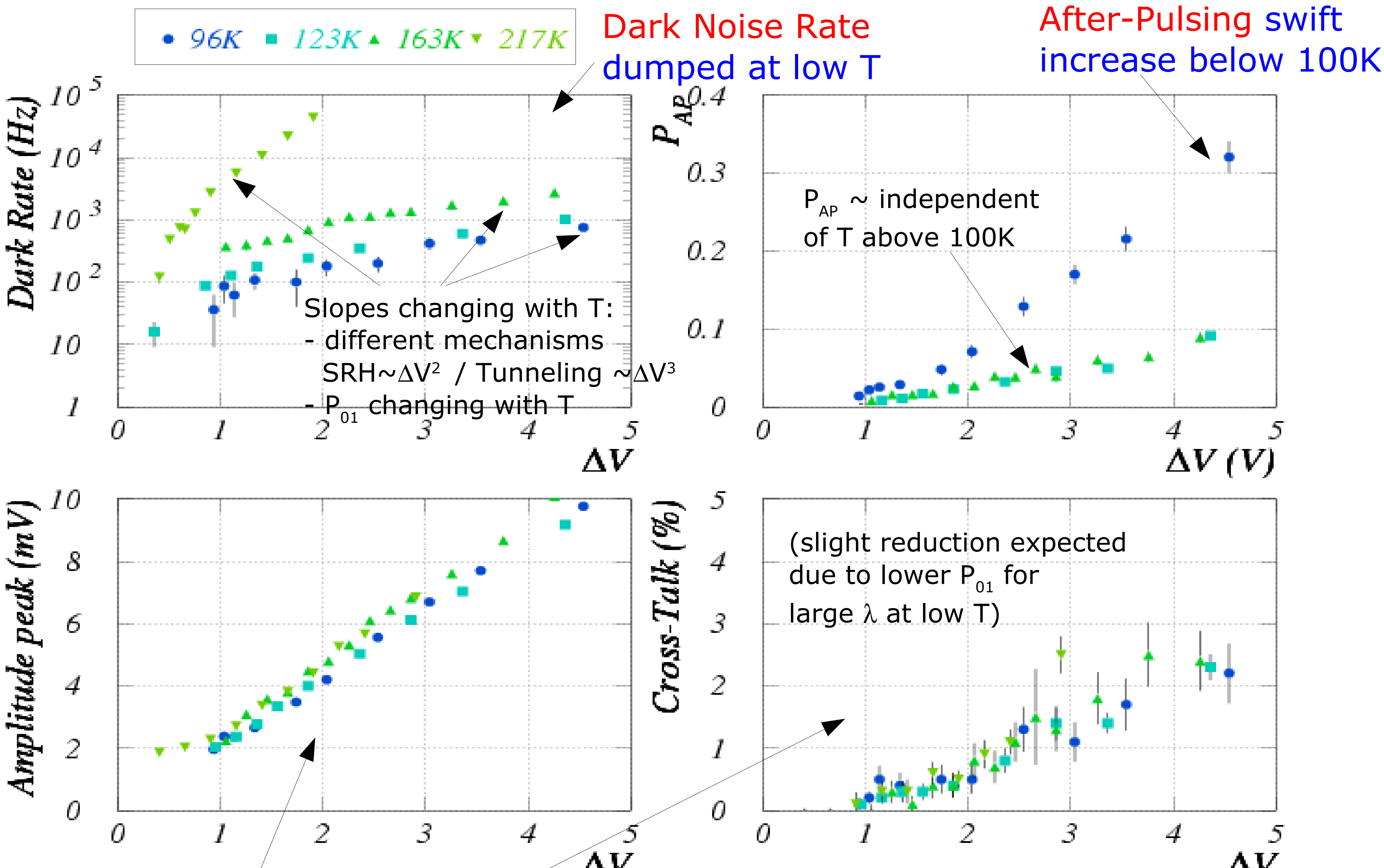
(3) Isolation implants are sufficient to stop direct component

Silicon absorption coefficients:



→ Crosstalk can't be eliminated simply by means of trenches
→ Main contribution to crosstalk comes from bottom reflections (using trenches)

DCR, AP, Gain, X-talk vs ΔV (various T)



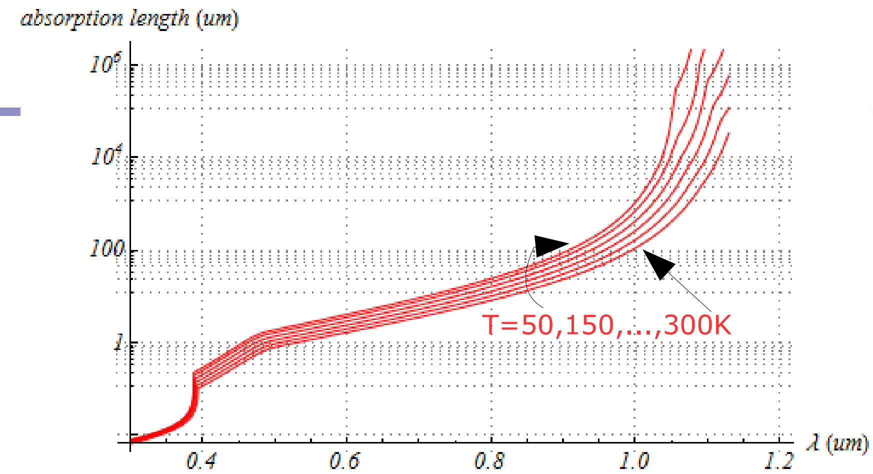
Gain and Cross-Talk are independent of T

FBK devices

G.C. et al NIM A628 (2011) 389

Photo-Detection Efficiency (PDE)

PDE = QE · P_{01} · FF



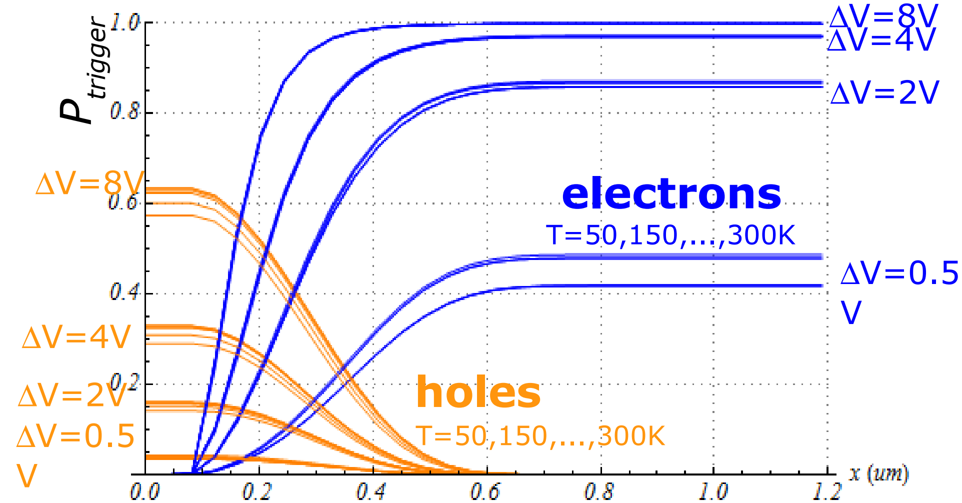
Rajkanan et al, Solid State Ele 22 (1979) 793

QE: carrier Photo-generation

probability for a photon to generate a carrier that reaches the high field region

→λ and T dependent

→ ΔV independent if full depletion at V_{bd}

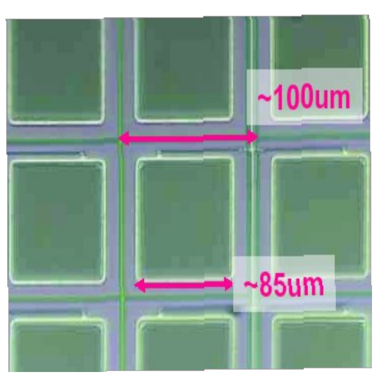


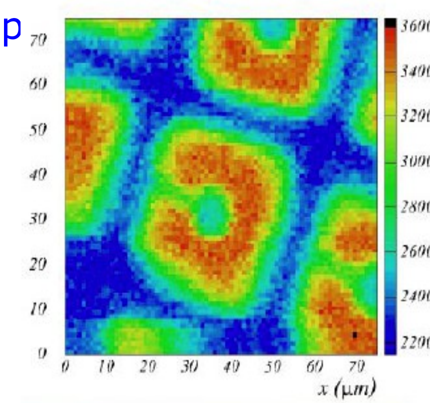
P_{01} : avalanche triggering probability

probability for a carrier traversing the high-field to generate the avalanche

→λ, T and ΔV dependent

Hamamatsu MPPCs close up



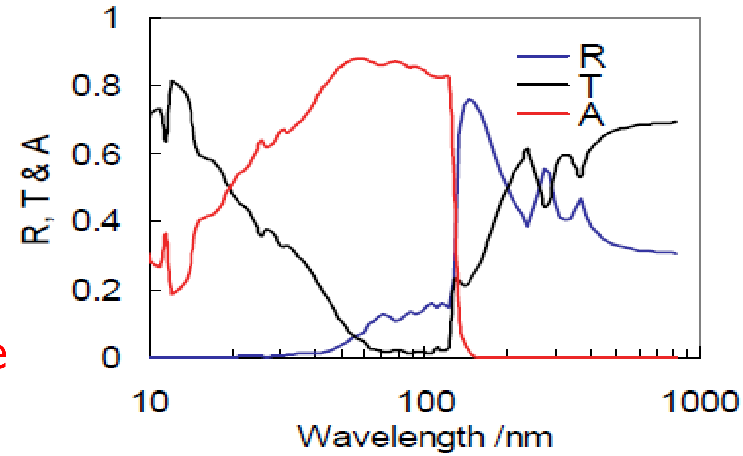


FF: geometrical Fill Factor

fraction of dead area due to structures between the cells, eg. guard rings, trenches

→ mild ΔV dependence (cell edges)

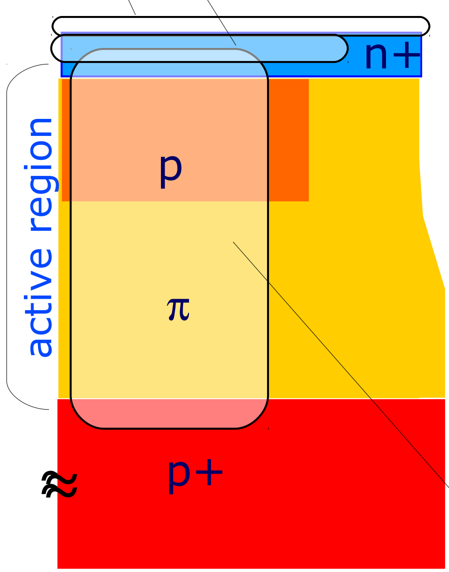
QE

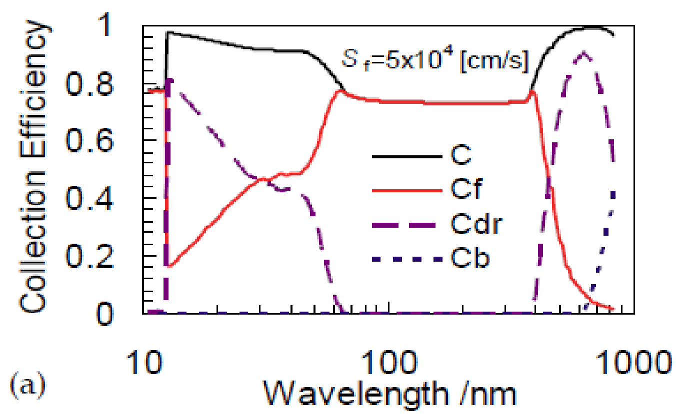


calculation for 30nm SiO_2 on Si layer

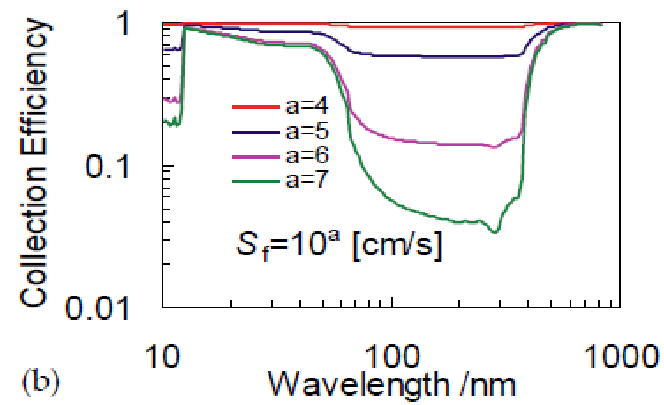
- optical T,A,(R) of the entrance window (dielectric on top of silicon surface)
  → angular and polarization dependence

- carrier recombination loss: collection efficiency front, depl. region, back





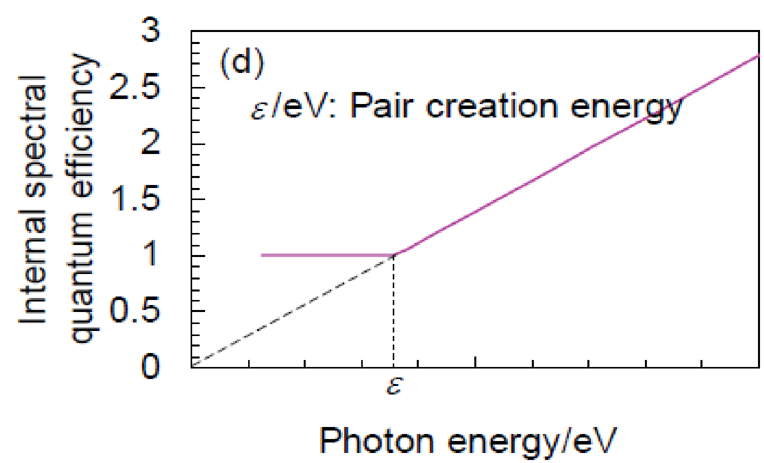
(a)



(b)

→ front region critical for 60nm < λ < 400nm
→ C eff. depends on surface recombination velocity S_f
→ freeze-out at low T

- internal quantum efficiency: probability to photo-generate an e-h pair ~ photon E (above threshold)



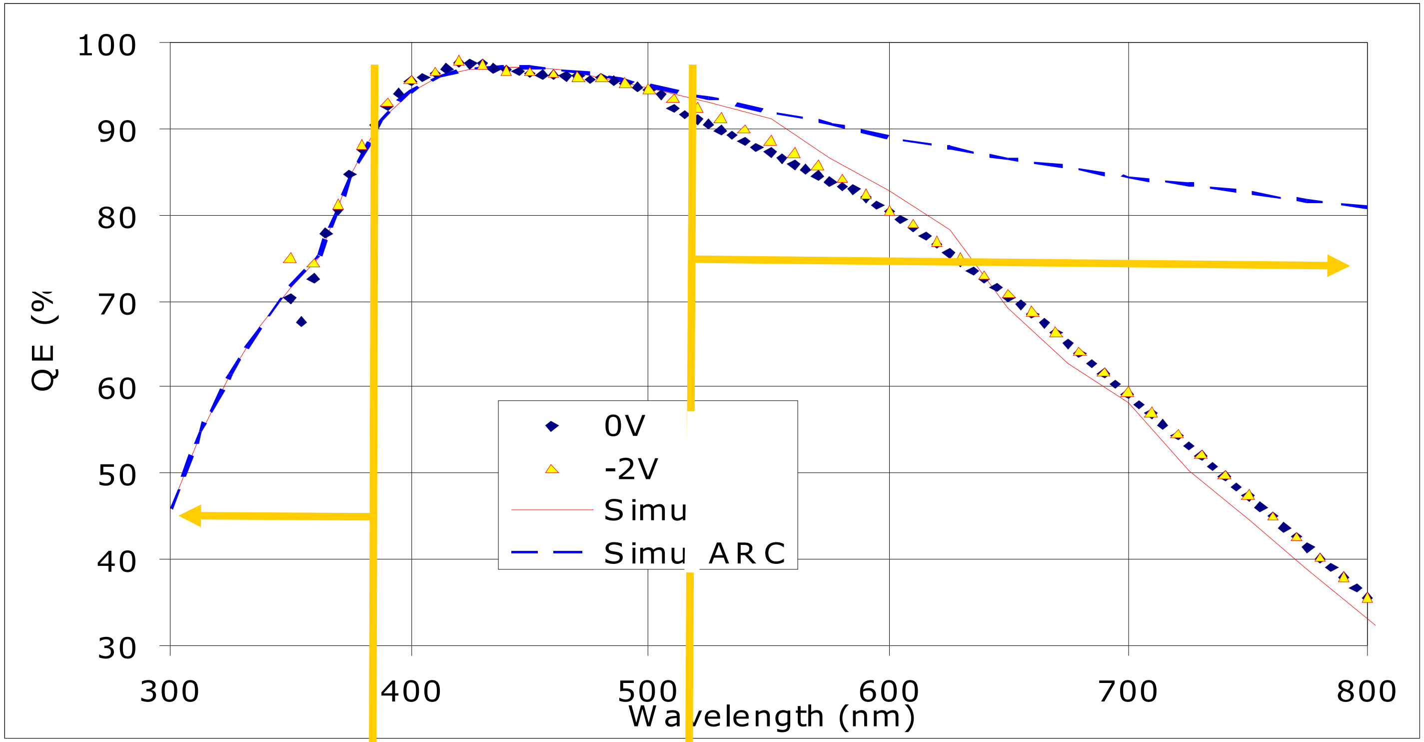
eg of QE optimization (blue)

- Anti-reflective coating (ARC)
- Shallow junctions for short λ
- Thick epi layers for long λ

QE single cell

FBK single cell

photo-voltaic regime (V_{bias}~ 0 V)

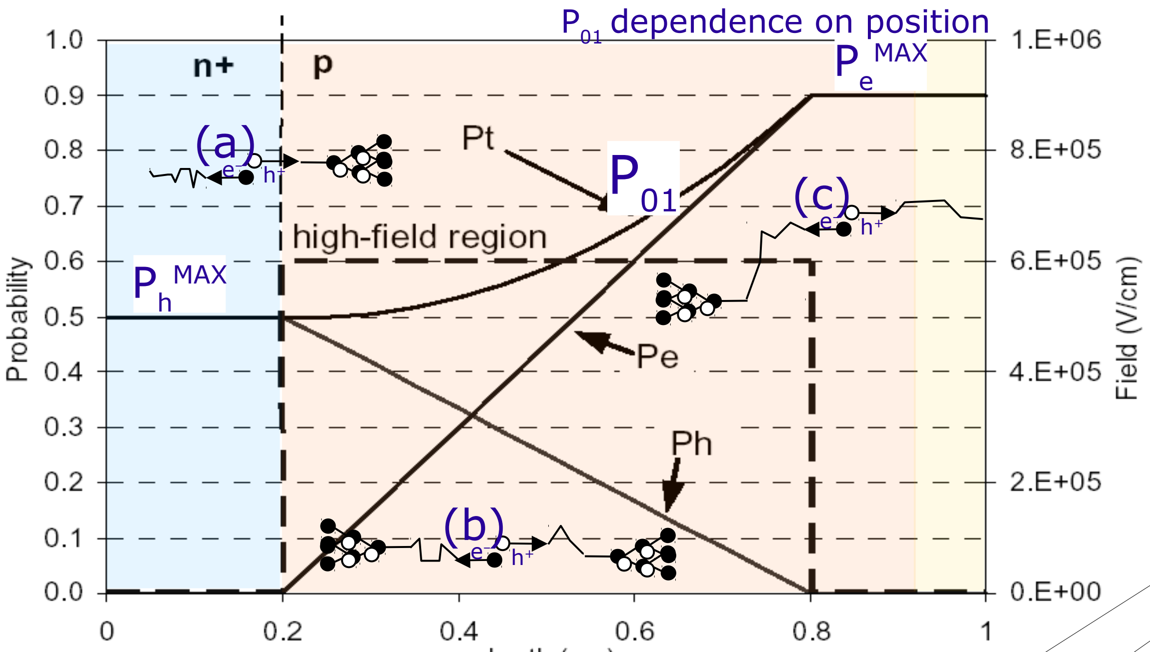


limited by
ARC Transmittance
&
Superficial
Recombination

limited by the
small π layer thickness

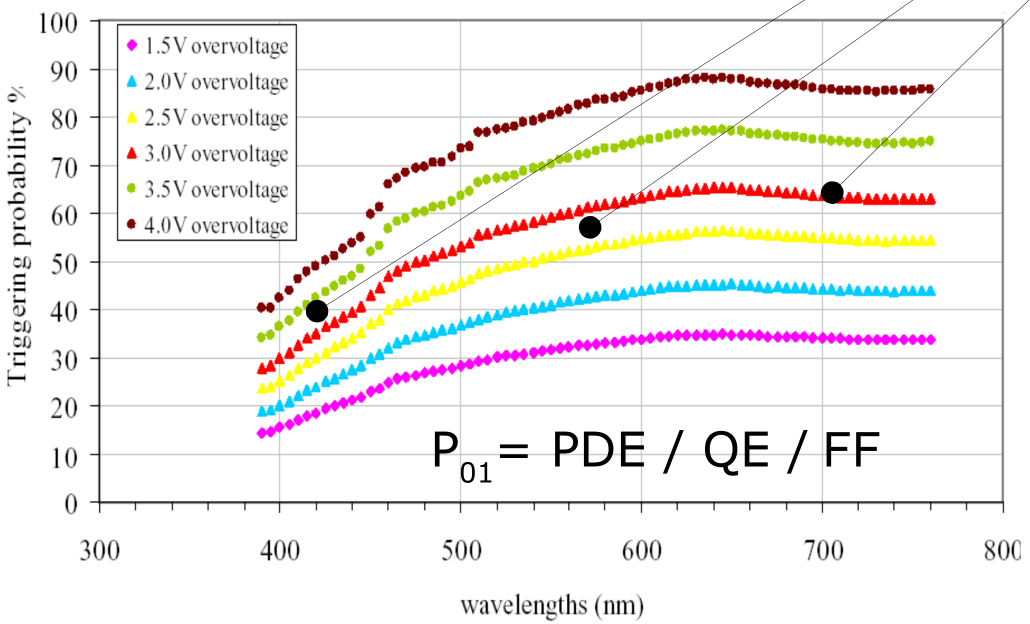
Most critical issue for **Deep UV SiPM**
note: reduced superficial recombination
in n-on-p wrt p-on-n

Avalanche trigger probability (P_{01})

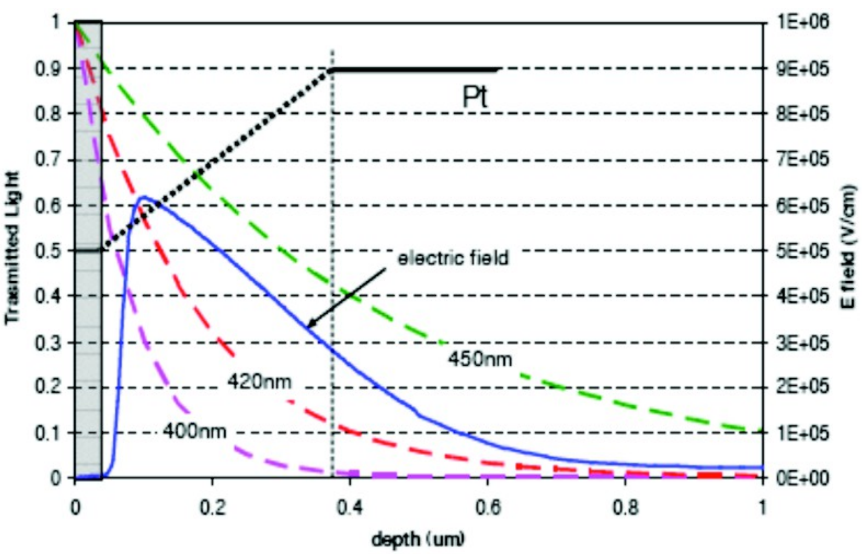


Probability calculations after *W.Oldham et al. IEEE TED (1972)*

Example with constant high-field:
(a) only holes trigger the avalanche
(b) both electrons and holes trigger
(c) only electrons trigger



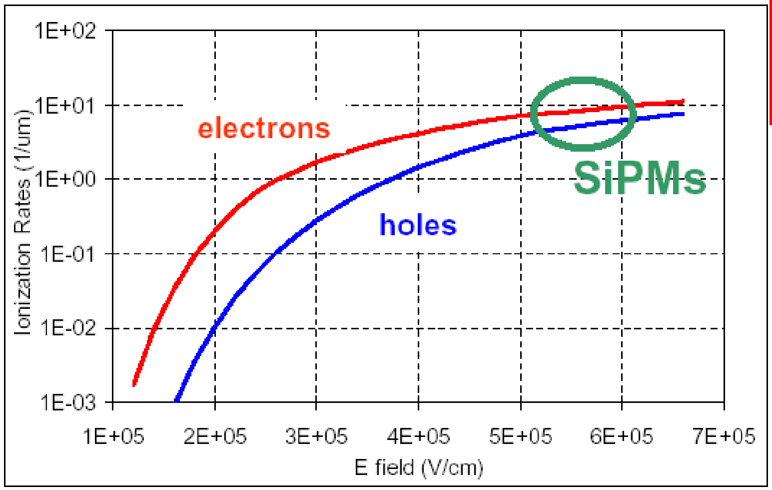
$$P_{01} = PDE / QE / FF$$

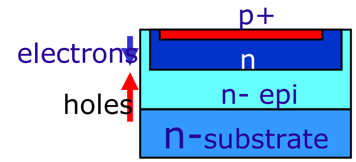


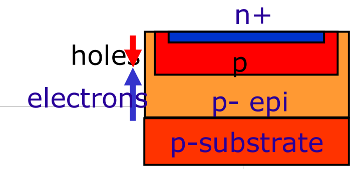
PDE vs ΔV

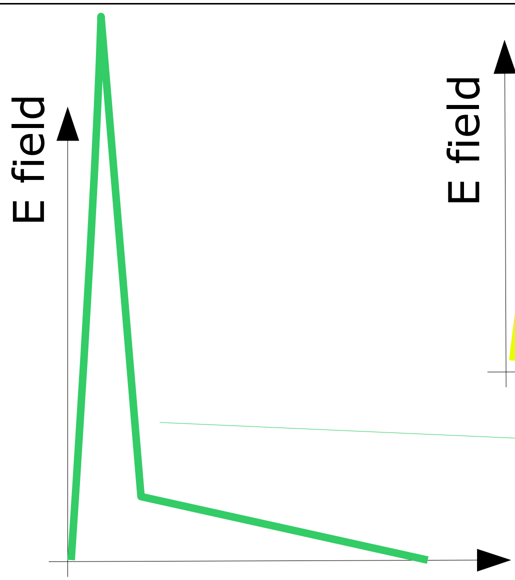
Ionization rate in Silicon → P_{01} optimization (n-on-p)

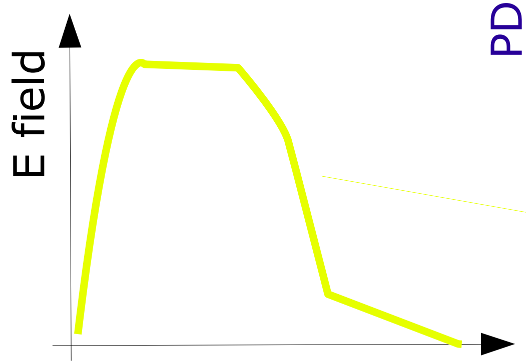
- high over-voltage
- photo-generation in the p-side of the junction

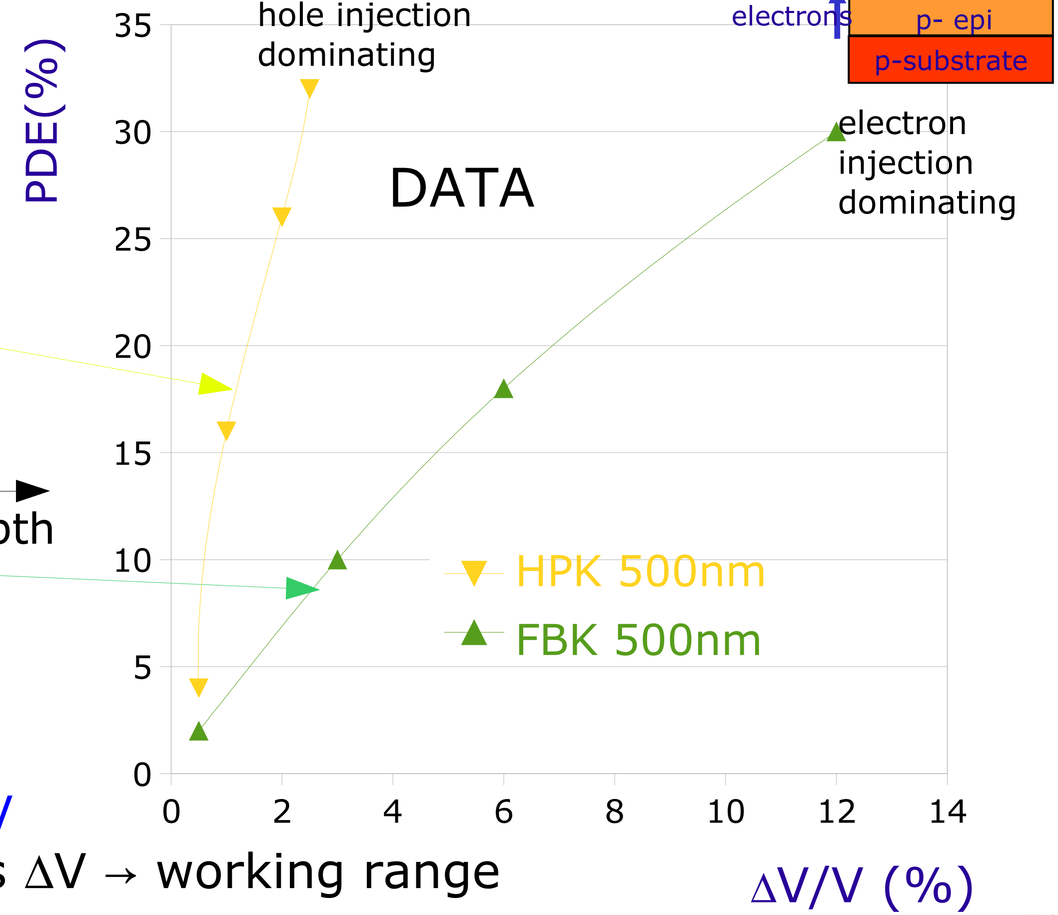












E field profile → the slope of PDE vs ΔV

note: P_{01} fixes also the slope of DCR vs ΔV → working range

G.Collazuol - PhotoDet 2012

PDE VS λ (shape)

N.Dinu et al. NIM A (2008)

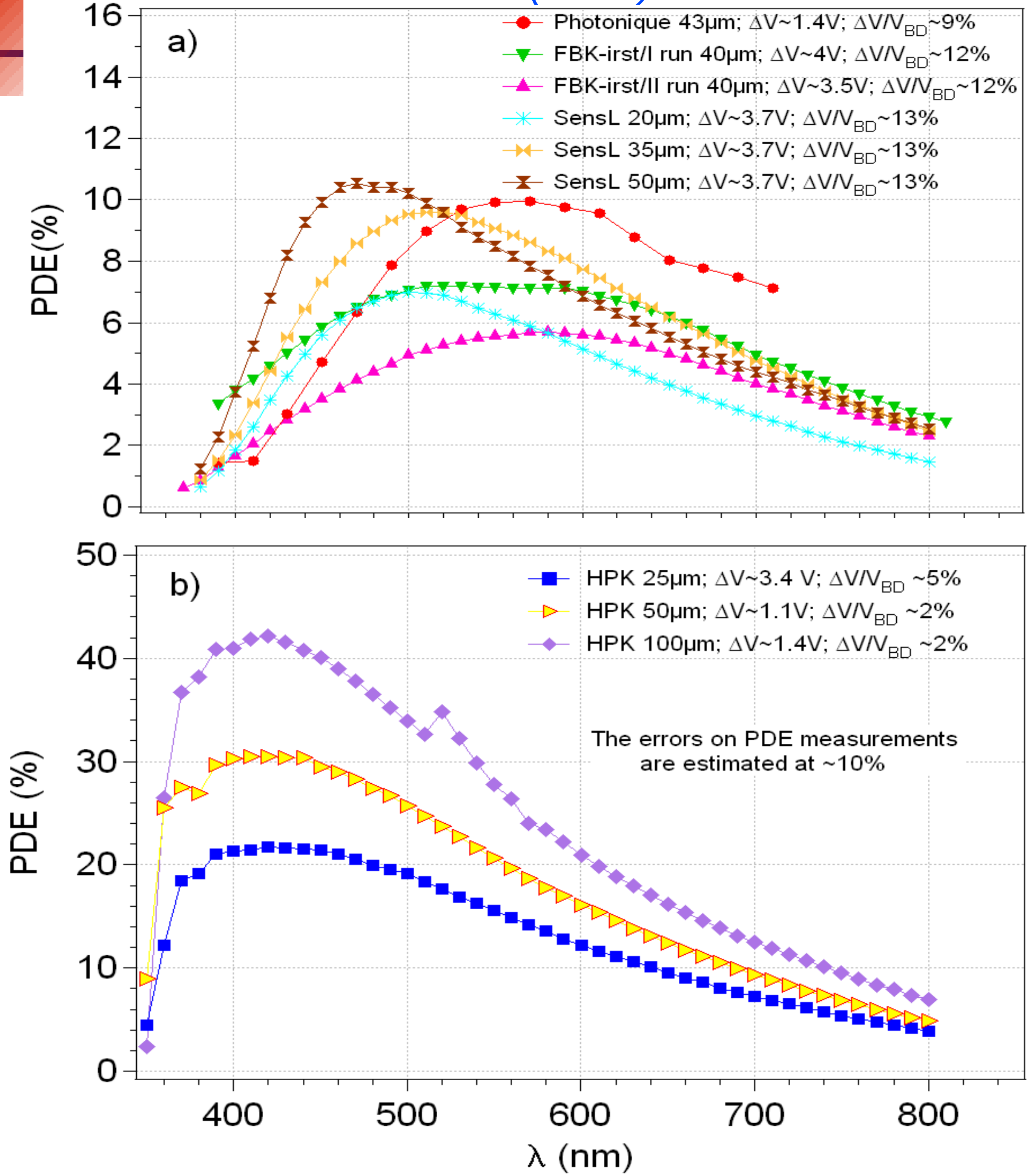
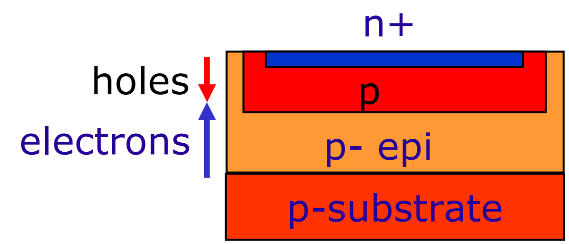
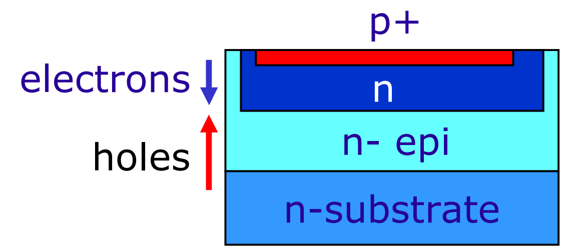


Fig. 5a) The PDE vs. λ of the Photonique, FBK-irst and SensL devices and b) HPK

n-on-p structures



p-on-n structure

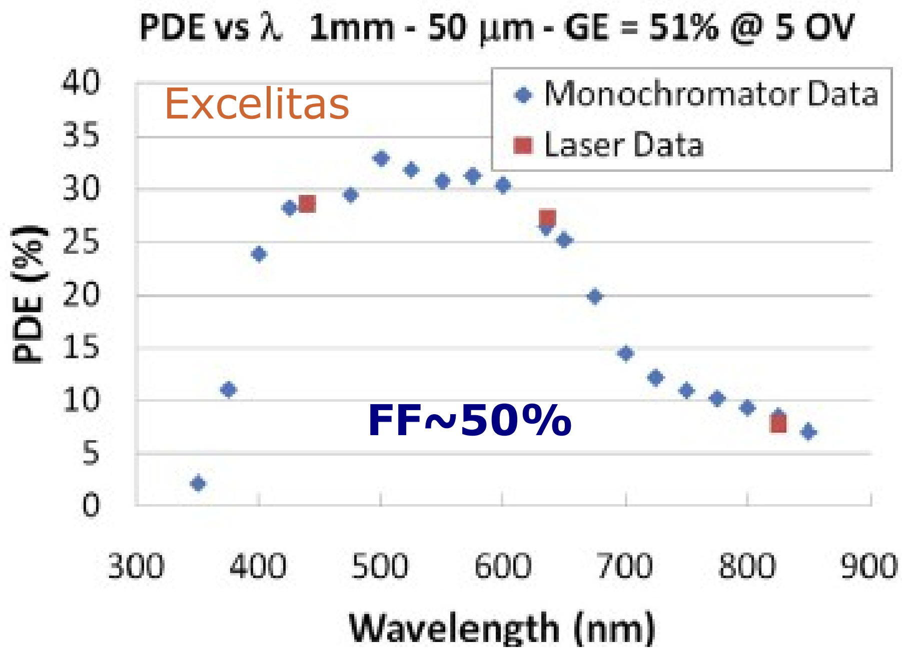


Note: geometrical fill factor included

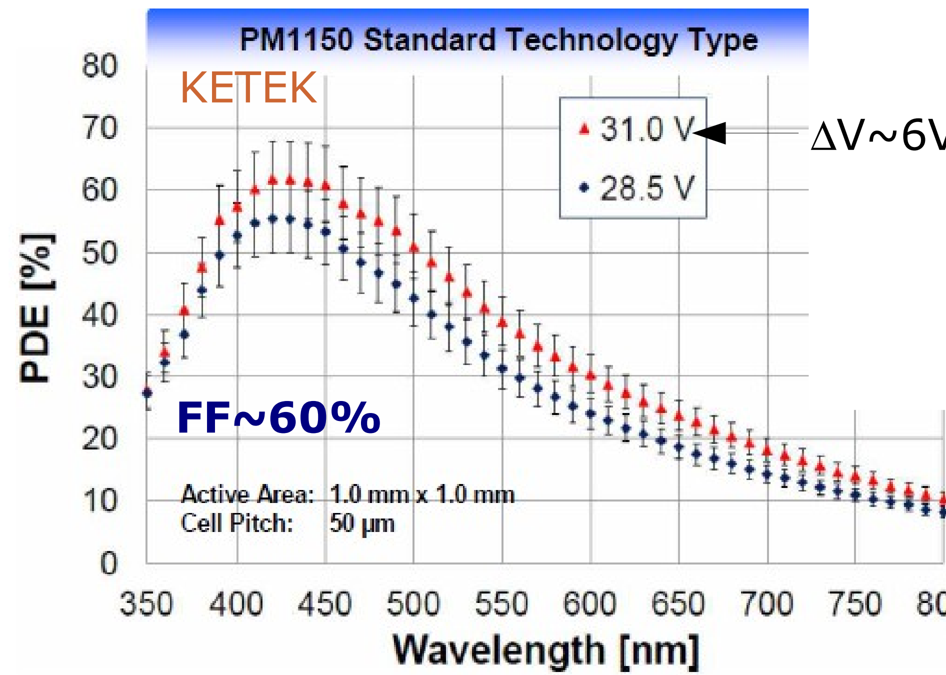
Improving PDE

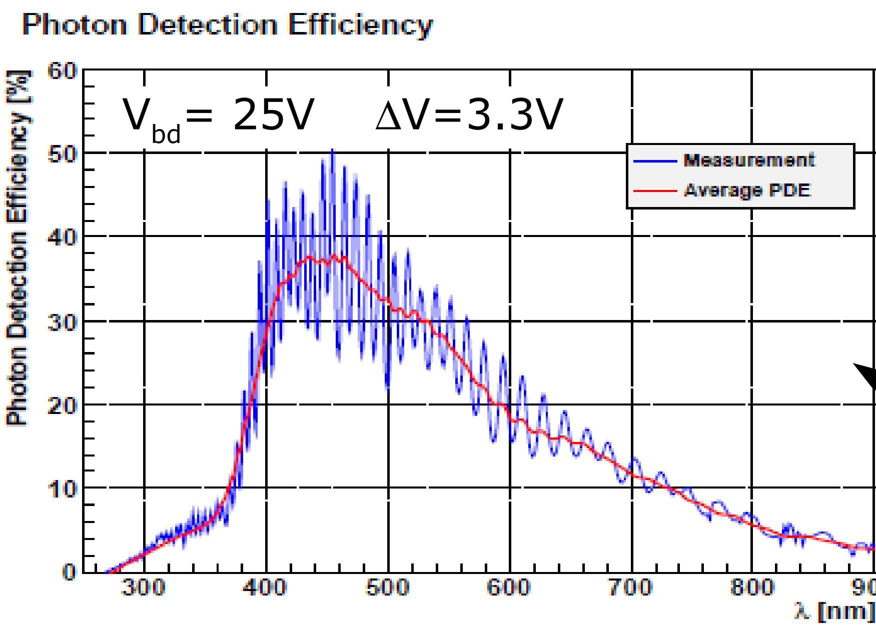
→ PDE peak constantly improving for many devices
→ every manufacturer shape PDE for matching target applications
→ UV SiPM eg from MePhi/Excelitas (see *E.Popova at NDIP 2011*)
→ DUV SiPMs in development too

Barlow – LIGHT 2011



F.Wiest – AIDA 2012 at DESY





T.Frach 2012 JINST 7 C01112

dSiPM (latest sensor 2011)
→ up to now no optical stack optimization
→ no anti-reflecting coating
→ potential improvement up to 60% peak PDE (*Y.Haemish at AIDA 2012*)

PDE vs T (ΔV constant)

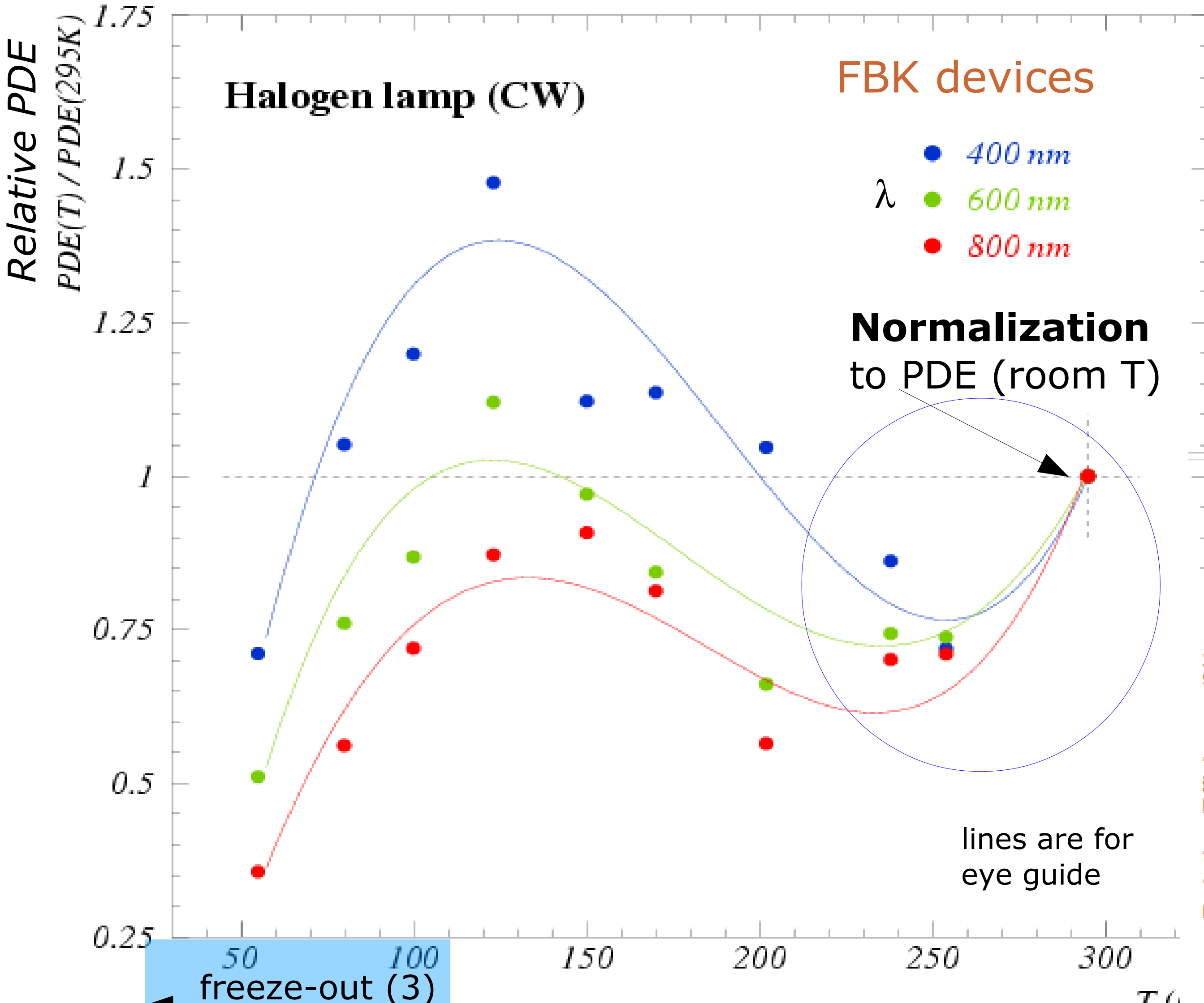
G.C. et al NIM A628 (2011) 389

When T decreases:

1) silicon E_{gap} increasing
→ larger attenuation length
→ lower QE (for larger λ)

2) mobility increasing
→ larger impact ionization
→ larger trigg. avalanche P_{01}

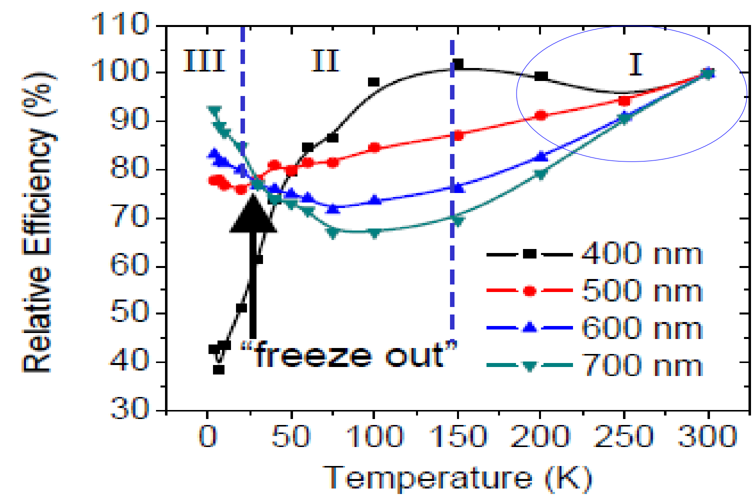
3) carriers freeze-out
onset below 120K
→ loss of carriers



freeze-out (3)

??? interplay between (1) and (2): modulation
… drop in 250<T<300 not well understood
(common feature with APDs')

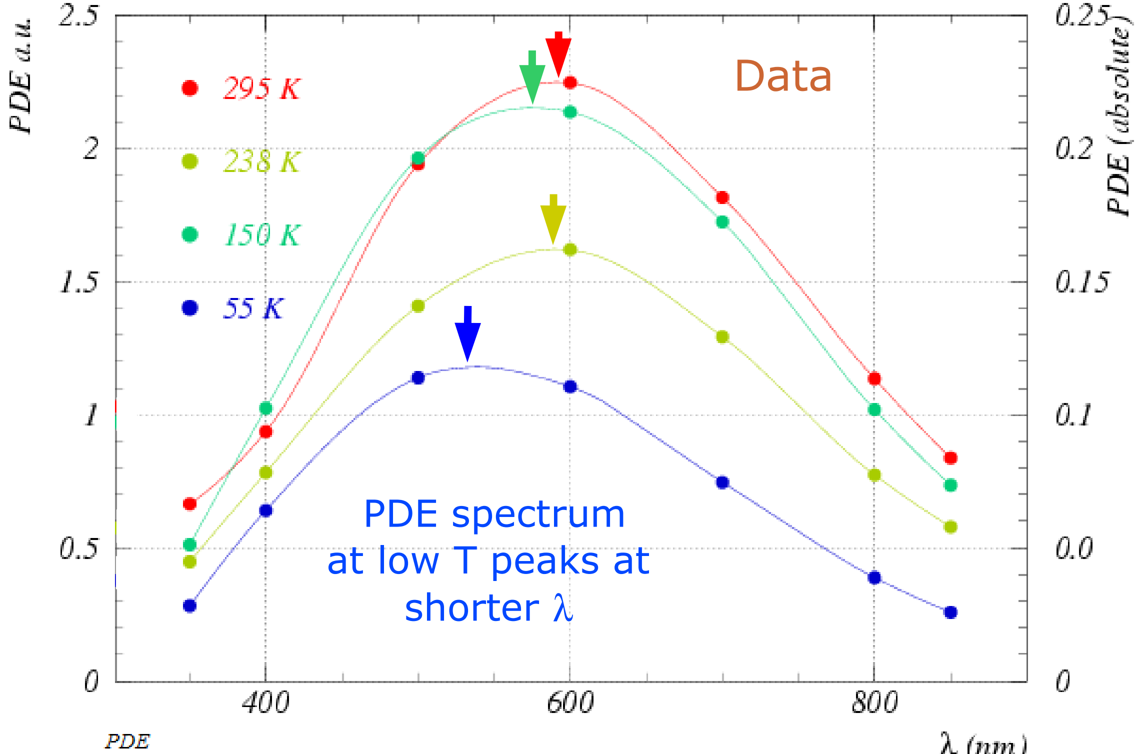
RMD APD at 400nm < λ < 700nm
Johnson et al, IEEE NSS 2009

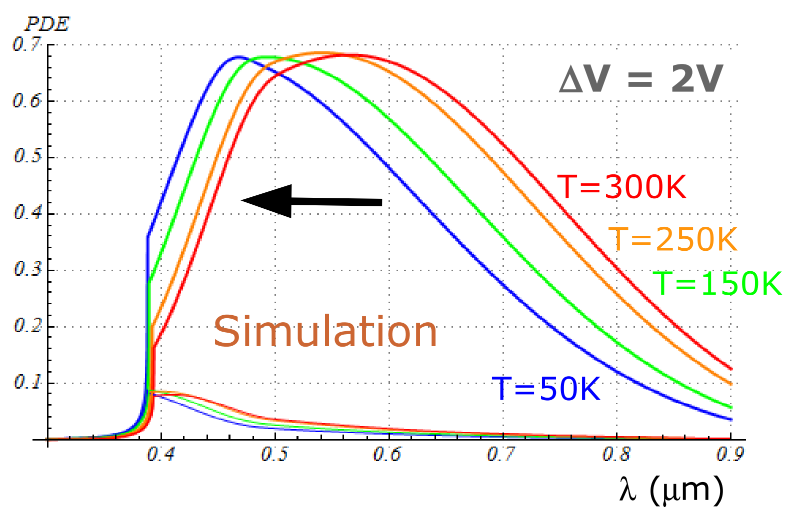


Additional effects in APD
(depletion region depends on T, ...)

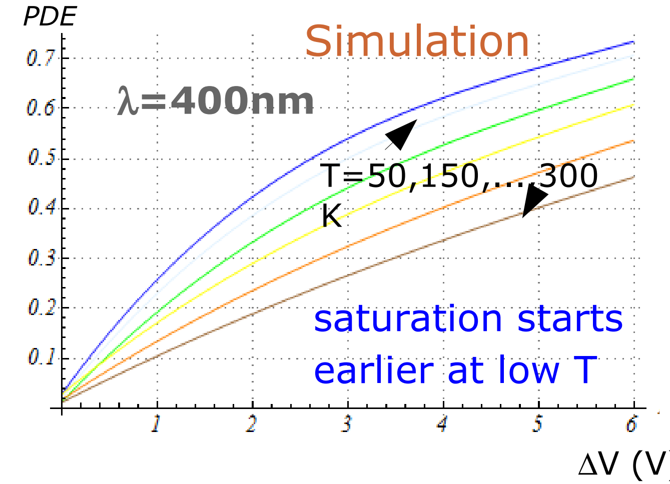
PDE dependences, changing with T

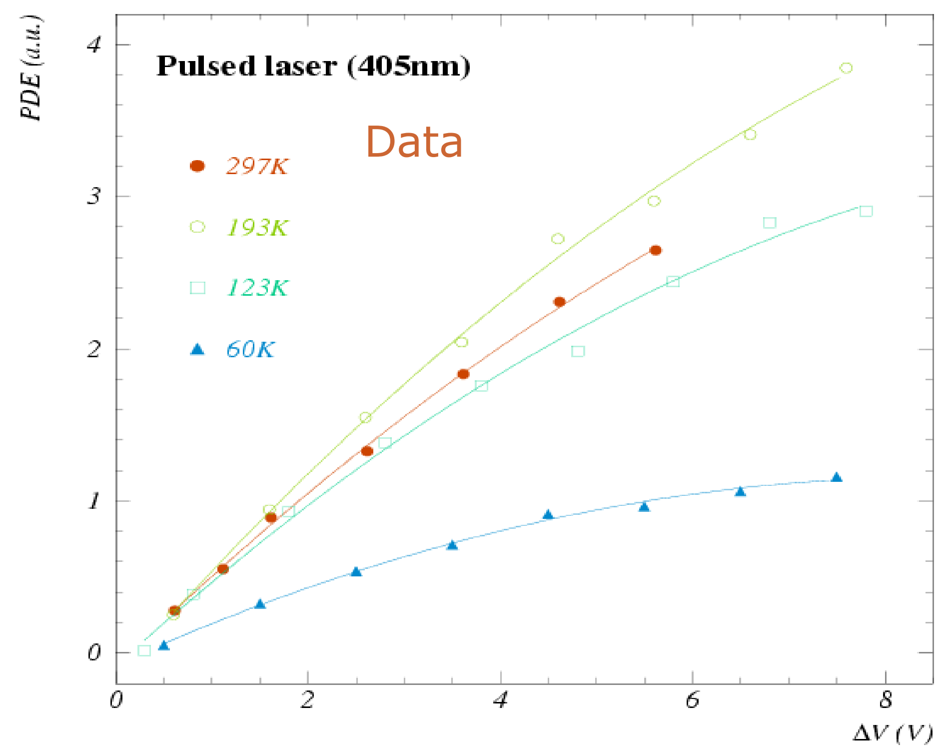
PDE vs λ (ΔV constant)





PDE ΔV vs (λ constant)





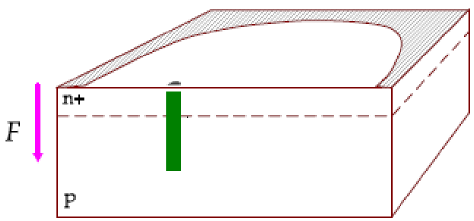
Timing

- SiPM are intrinsically very fast
  → jitter (gaussian) below 100ps, depending on ΔV

but also → non-gaussian tails up to O(ns), depending on wavelength

- Timing measurement:
  → use of fast signal shape component
  → use waveform, better than CFD (much than ToT)

GM-APD avalanche development



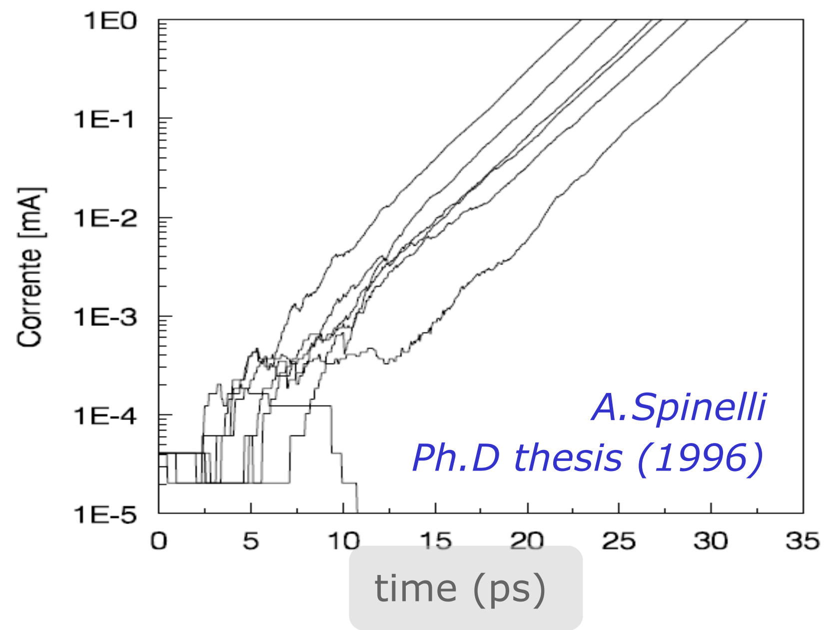
Longitudinal multiplication

- Duration ~ few **ps**
- Internal current up to ~ few **μA**

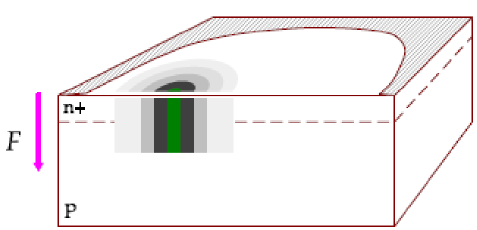
(1) Avalanche "seed": free-carrier concentration rises exponentially by **"longitudinal" multiplication**

(1') Electric field locally lowered (by **space charge R effect**) towards breakdown level

Multiplication is self-sustaining
Avalanche current steady until new multiplication triggered in near regions



x100



Transverse multiplication

- Duration ~ few **100ps**
- Internal current up to ~ several **10μA**

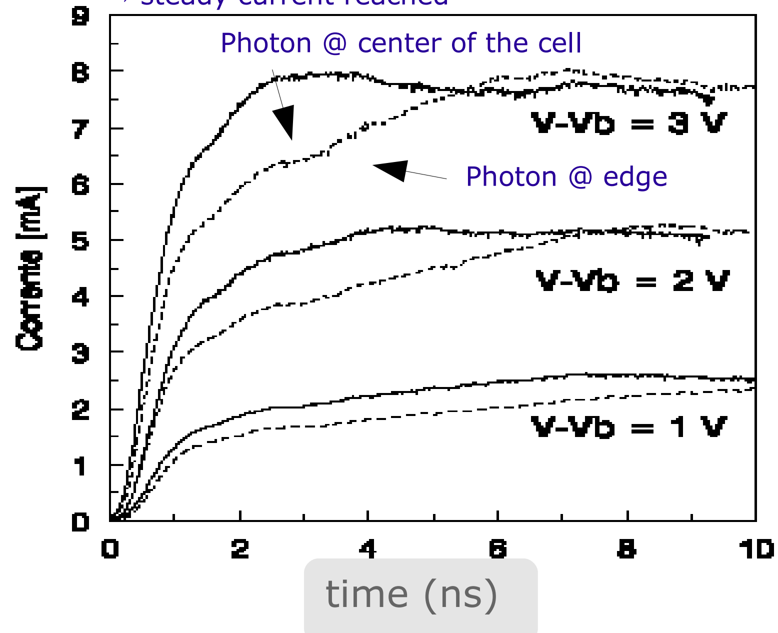
(2) Avalanche spreads "transversally" across the junction

(**diffusion speed ~up to 50μm/ns** enhanced by multiplication)

(2') Passive quenching mechanism effective after transverse **avalanche size ~10μm**

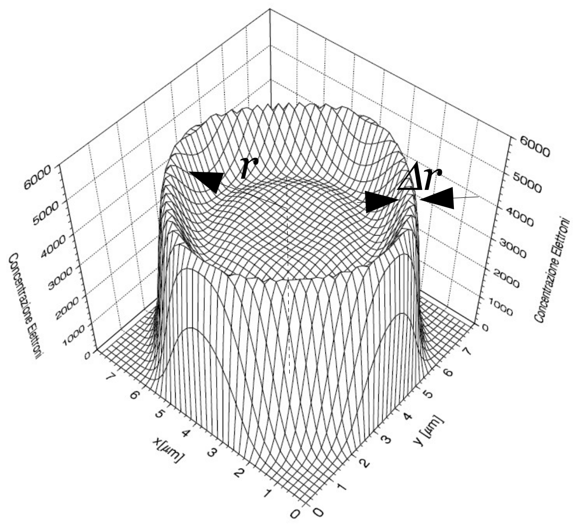
(if no quench, avalanche spreads over the whole active depletion volume → avalanche current reaches a final saturation steady state value)

Simulation w/o quenching:
→ steady current reached



GM-APD avalanche transverse propagation

Avalanche transverse propagation by a kind of shock wave: the wavefront carries a high density of carriers and high E field gradients (inside: carriers' density lower and E field decreasing toward breakdown level)



$$\frac{dS}{dt}=\frac{d}{dt}2\pi r(t)\Delta r=2\pi v_{diff}\Delta r=4\pi\Delta r\sqrt{\frac{D}{\tau}}$$

Rate of current production: $$\frac{dI}{dt}=\frac{dI}{dS}\frac{dS}{dt}\sim\frac{\sqrt{D}}{R_{sp}\sqrt{\tau}}$$

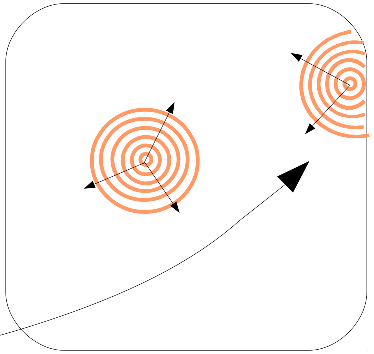
$$\frac{dI}{dS}=J=\frac{V_{bias}}{R_{sp}(S)}$$

S = surface of wavefront (ring of area $2\pi r\Delta r$)
$R_{sp}(S)$ = space charge resistance $\sim w^2/2\varepsilon v \sim O(50\ k\Omega\ \mu m^2)$
$v_{diff} \sim$ O(some 10 μm/ns)
D = transverse diffusion coefficient ~ O(μm^2/ns)
τ = longitudinal (exponential) buildup time ~ O(few ps)

Internal current rising front:
the faster it grows, the lower the jitter
dI/dt → understand/engineer timing features of SiPM cells

$$\tau\sim\frac{1}{1-(E_{max}/E_{breakdown})^n}$$

→ timing resolution improves at high V_{bias}
→ E field profile affects τ and R_{sp} (wider E field profile → smaller R) (should be engineered when aiming at ultra-fast timing)
→ T dependence of timing through τ and D
→ slower growth at GAPD cell edges → higher jitter at edges
reduced length of the propagation front



SiPM cell

GM-APD timing jitter: fast and slow components

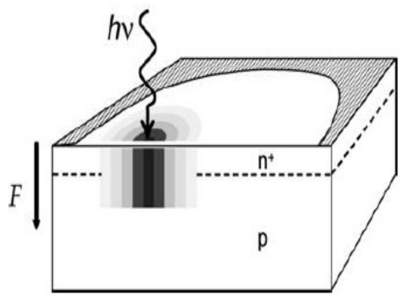
1) Fast component: gaussian with time scale O(100ps)

Statistical fluctuations in the avalanche:

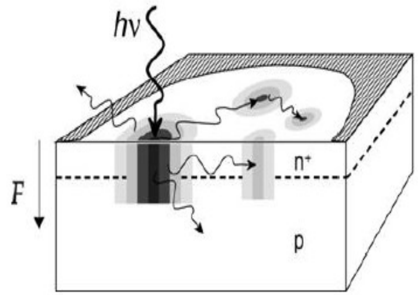
- **Longitudinal** build-up (minor contribution)
- **Transversal** propagation (main contribution):

  - via multiplication assisted diffusion (dominating in few μm thin devices)
    A.Lacaita et al. APL and El.Lett. 1990

  - via photon assisted propagation (dominating in thick devices – O(100μm))
    PP.Webb, R.J. McIntyre RCA Eng. 1982
    A.Lacaita et al. APL 1992



Multiplication assisted diffusion



Photon assisted propagation

Fluctuations due to
a) impact ionization statistics

b) variance of longitudinal position of photo-generation: finite drift time even at saturated velocity
note: saturated ve ~ 3 vh
(n-on-p are faster in general)

→ Jitter at minimum → **O(10ps)**
(very low threshold → not easy)

Fluctuations due to
c) variance of the transverse diffusion speed v_{diff}

d) variance of transverse position of photo-generation: slope of current rising front depends on transverse position

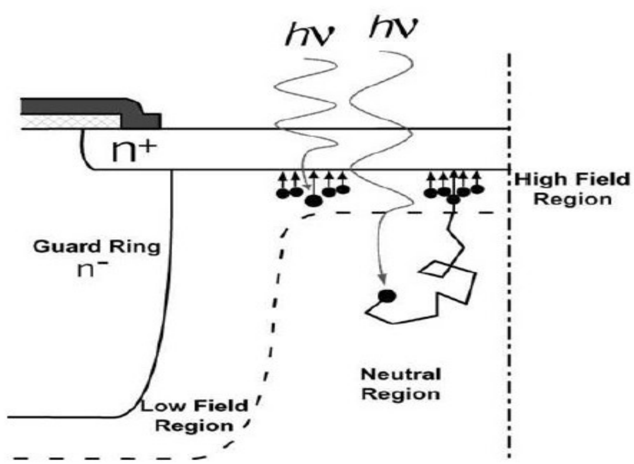
→ Jitter → **O(100ps)**
(usually threshold set high)

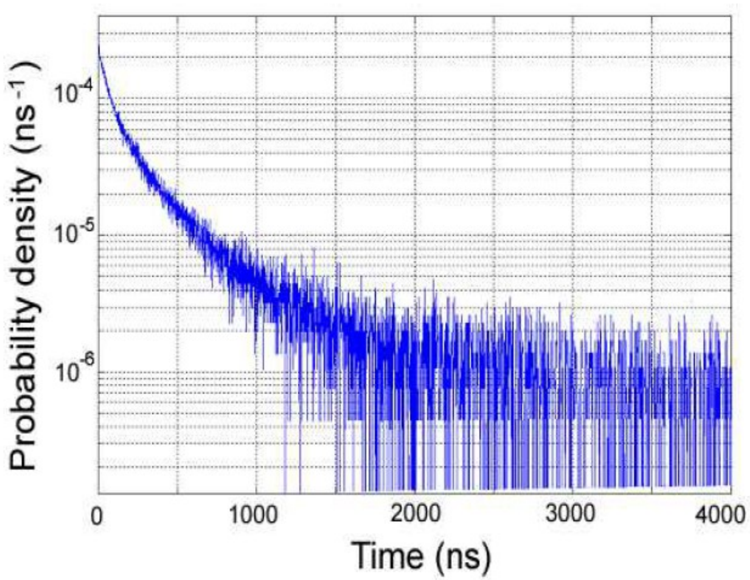
GM-APD timing jitter: fast and slow components

2) Slow component: non-gaussian tails with time scale O(ns)

Carriers photo-generated in the neutral regions above/beneath the junction and reaching the electric field region by diffusion

G.Ripamonti, S.Cova Sol.State Electronics (1985)





S.Cova et al. NIST Workshop on SPD (2003)

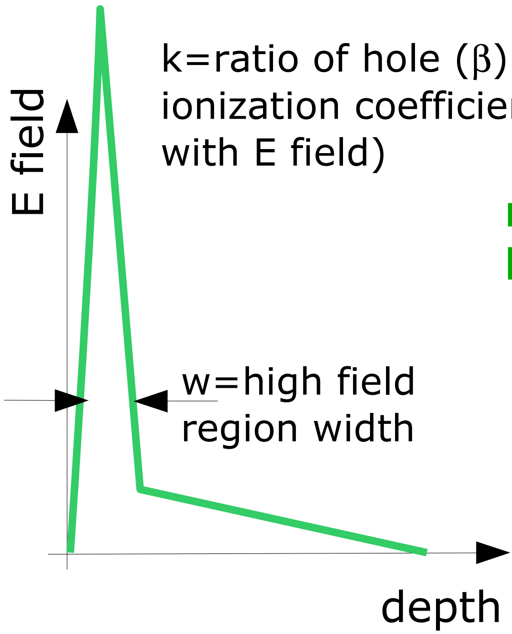
tail lifetime: $\tau \sim L^2 / \pi^2 D$ ~ up to some ns
L = effective neutral layer thickness
D = diffusion coefficient

→ **Neutral regions** underneath the junction : timing tails for long wavelengths
→ **Neutral regions** in APD entrance: timing tails for short wavelengths

PDE vs timing optimization

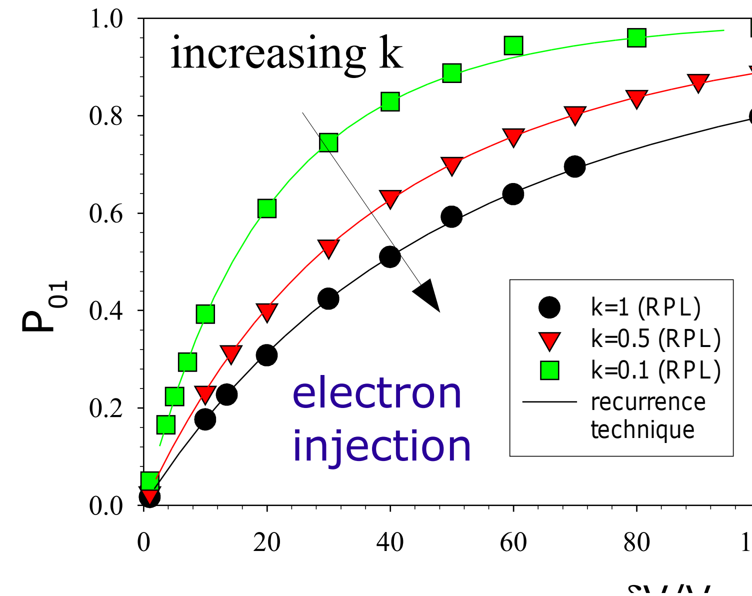
C.H.Tan et al IEEE J.Quantum Electronics 13 (4) (2007) 906

k=ratio of hole (β) to electron (α) ionization coefficient (increasing with E field)

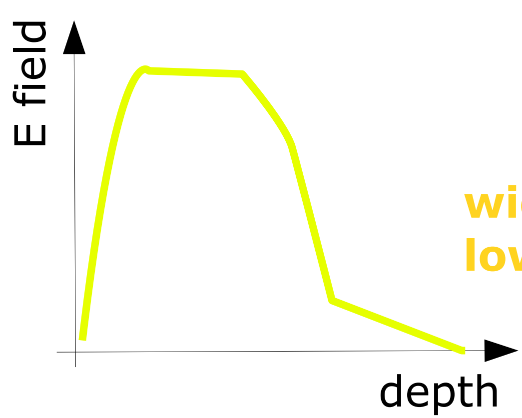


narrow avalanche region, high E:
- **small w**
- **high k = b/a**

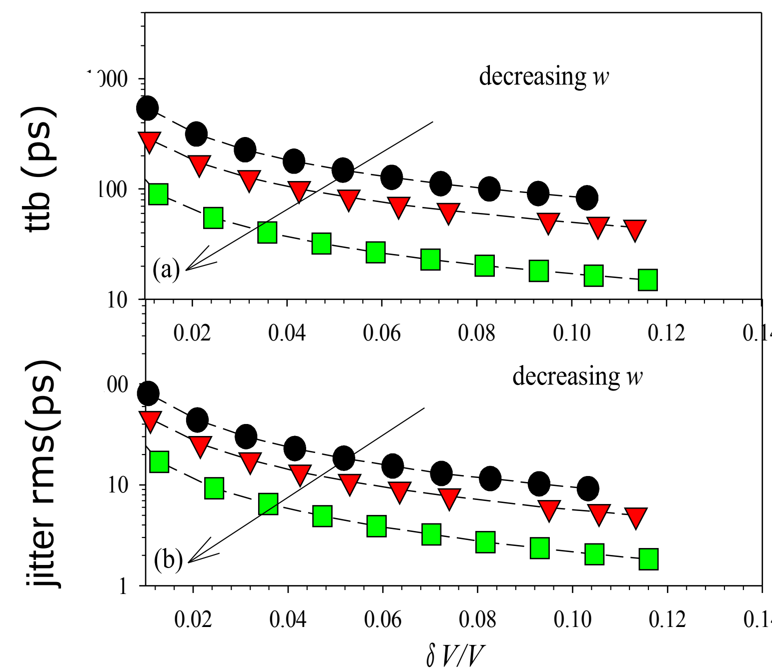
better for **TIMING**



better for **PDE**



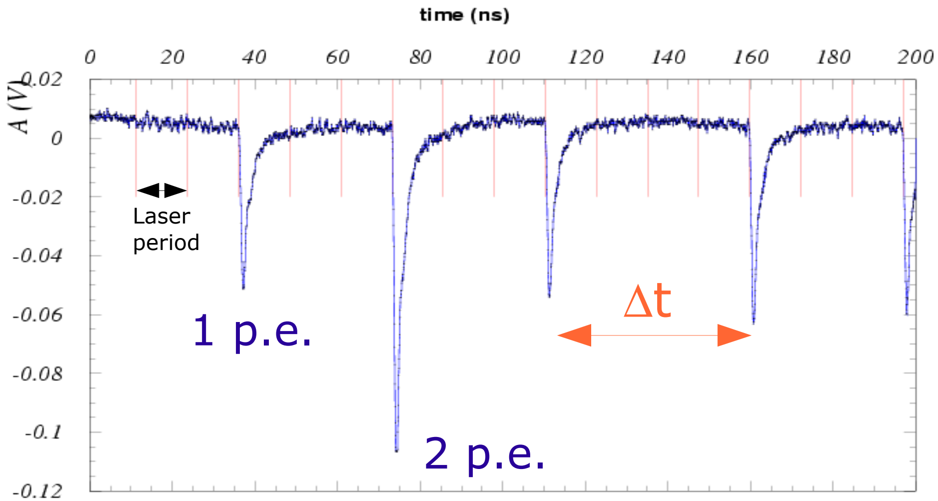
wide avalanche region, low E:
- **wide w**
- **small k = b/a**



Plots are courtesy of C.H.Tan

Waveform analysis: optimum timing filter

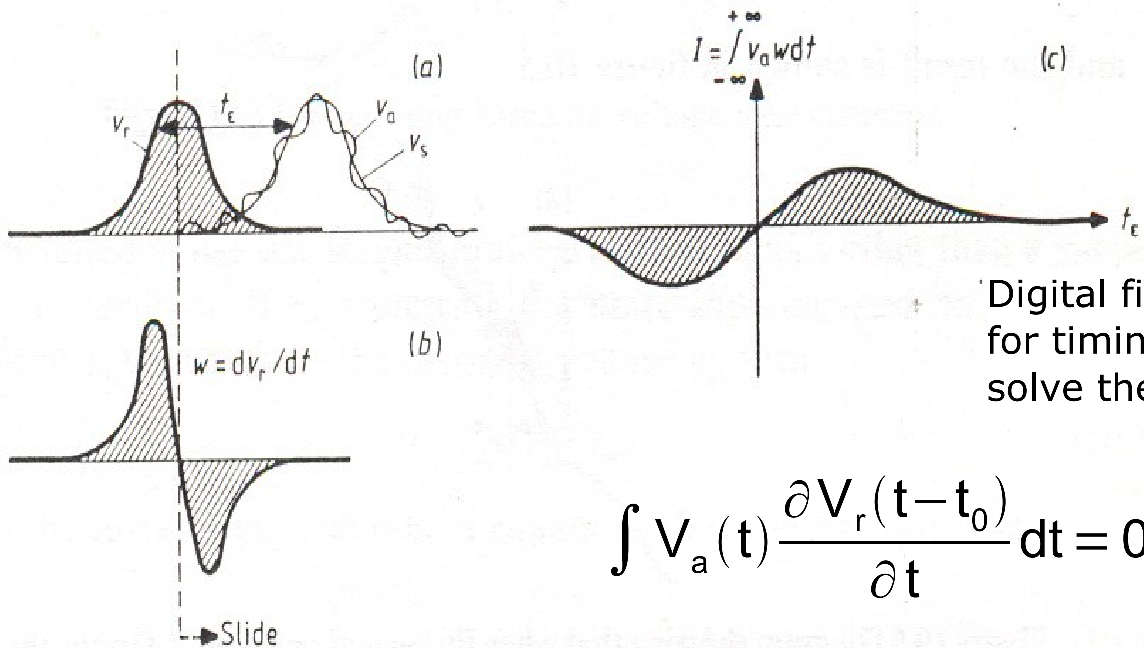
Example of intrinsic SPTR measurement from Δt of consecutive pulses by laser shots



Different algorithms to reconstruct the time of the pulses:

- ✗ parabolic fit to find the peak maximum
- ✗ CFD (digital)
- ✗ average of time samples weighted by the waveform derivative
- ✔ digital filter: weighting by the derivative of a reference signal
  → optimum against (white) noise (if signal shape fixed)

G.C. et al NIMA 581 (2007) 461



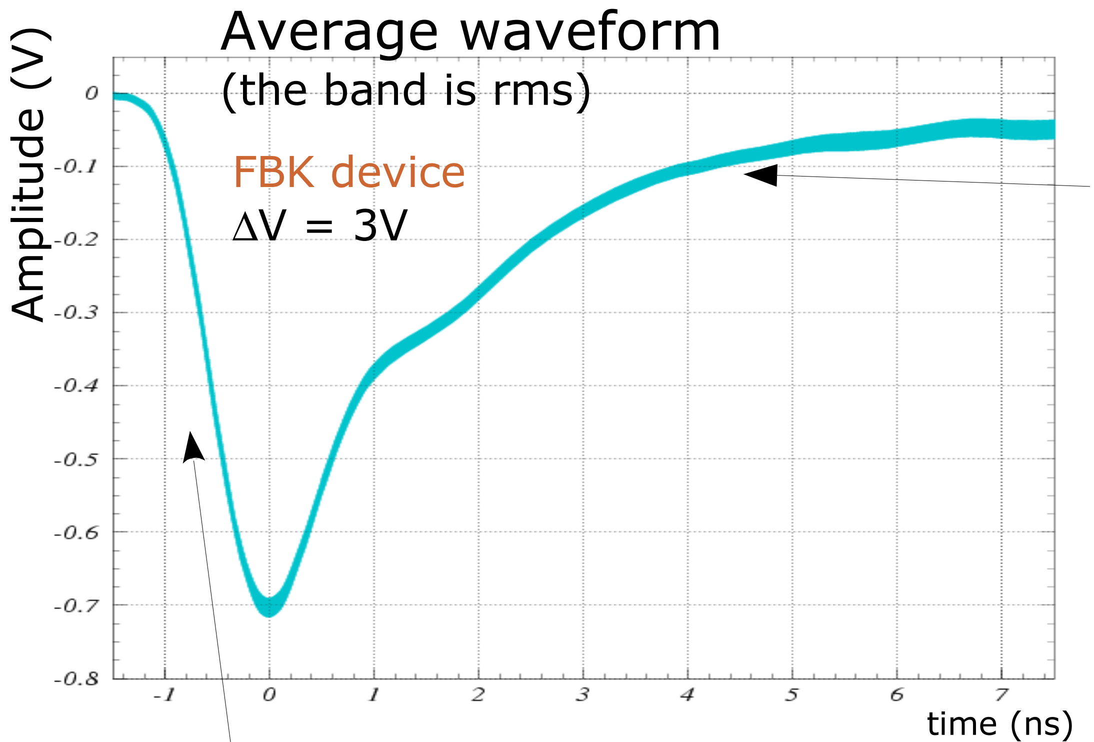
Digital filter to minimize N/S for timing measurements: solve the following equation on t_0 :

$$\int V_a(t)\frac{\partial V_r(t-t_0)}{\partial t}dt = 0$$

Va = measured signal (includes noise)
Vr = reference signal
t_0 = reference time

see e.g. Wilmshurst "Signal recovery from noise in electronic instrumentation"

Waveform (single p.e.)



Falling signal shape fluctuates considerably (due eg to after-pulses) → signal tail is non useful for timing, if not detrimental

note: using Time-over-Threshold method for slew correction might lead to worse resolution

Reminder:

$$\frac{dI}{dt} \sim \frac{\sqrt{D}}{R_{sp}\sqrt{\tau}}$$

$$\tau \sim \frac{1}{1-(E_{max}/E_{breakdown})^n}$$

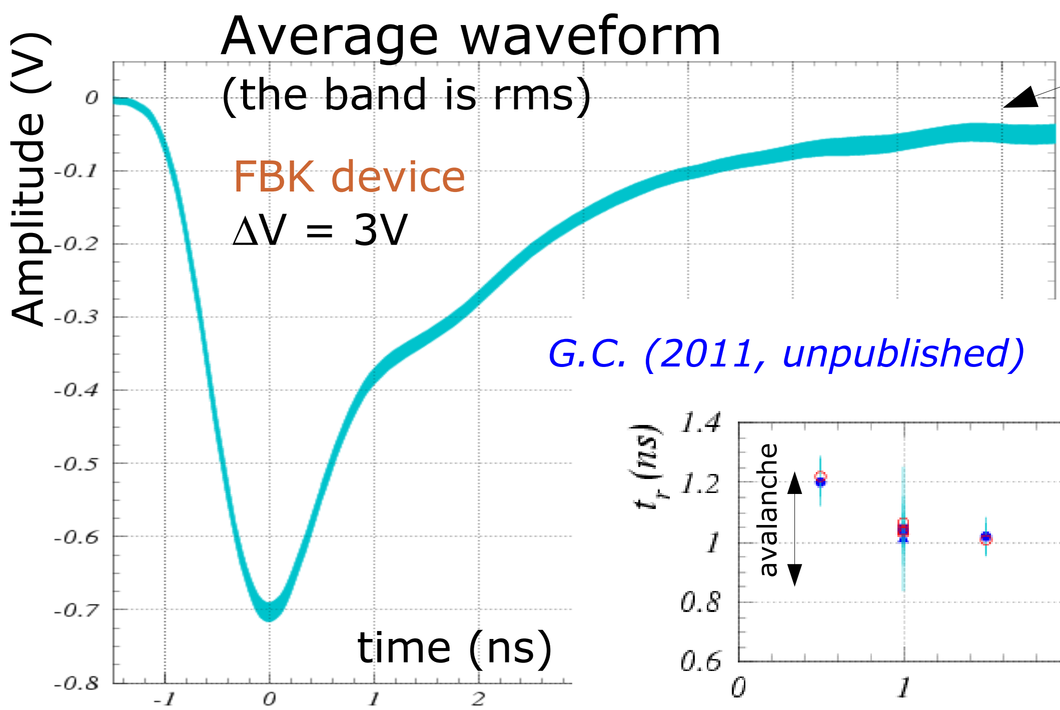
Rise-time depends on **ΔV**, **T** and **impact position** ie **signal shape is not constant**, then:

1) CFD method only partially effective in canceling time walk effects

2) any digital timing filter should account for shape variations (ΔV, T)

For comparison about waveform method and various digital algorithms
see Ronzhin et al NIM A 668 (2012) 94

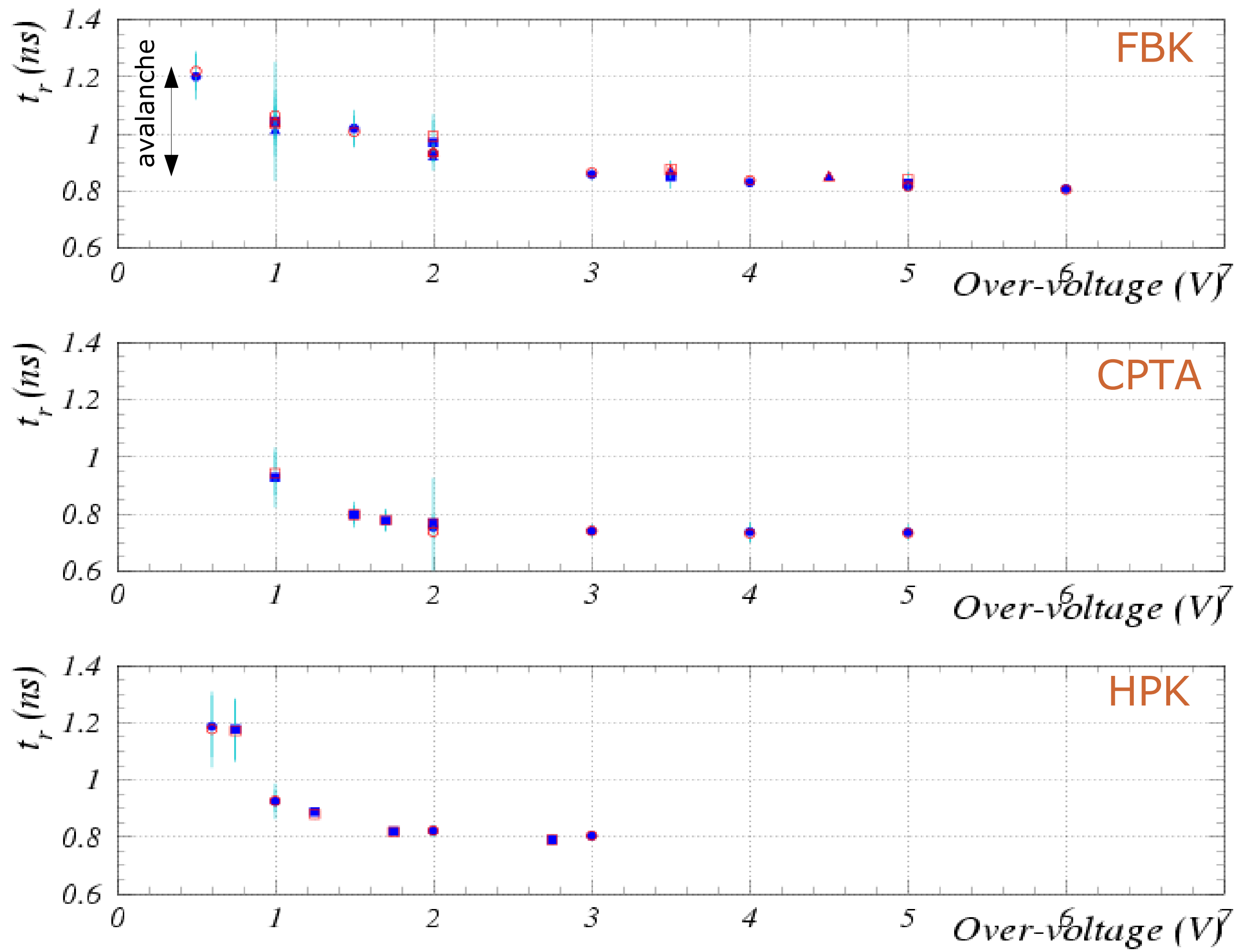
Waveform analysis: 1 p.e. reference signal



Additional contribution to rms (after-pulses)

Rise time (10%-90%)
(dominated by electronics contribution)

G.C. (2011, unpublished)



Reminder:

$$\frac{dI}{dt} \sim \frac{\sqrt{D}}{R_{sp}\sqrt{\tau}}$$

$$\tau \sim \frac{1}{1-(E_{max}/E_{breakdown})^n}$$

For comparison about rise-time of HKP devices see P.Avella et al doi:10.1016/j.nima.2011.11.049

Single Photon Time Resolution = gaussian + tails

Time resolution of SiPM is not just a gaussian, but gaussian + tails
(in particular at long wavelengths)

G.C. et al NIMA 581 (2007) 461

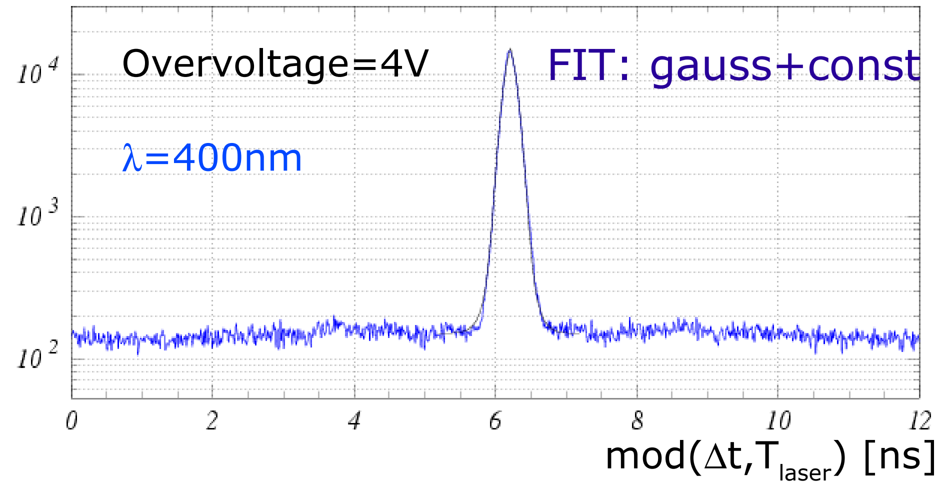
Gaussian + Tails (long λ)

rms ~ 50-100 ps

~ exp (-t / O(ns))
contrib. several %
for long wavelengths

Data at **λ=400nm**

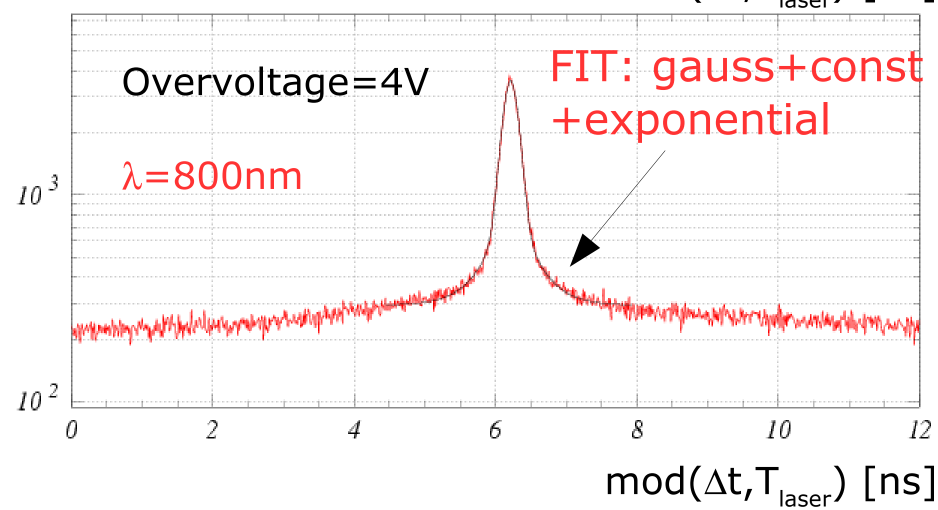
A simple **gaussian component** fits fairly



Data at **λ=800nm**

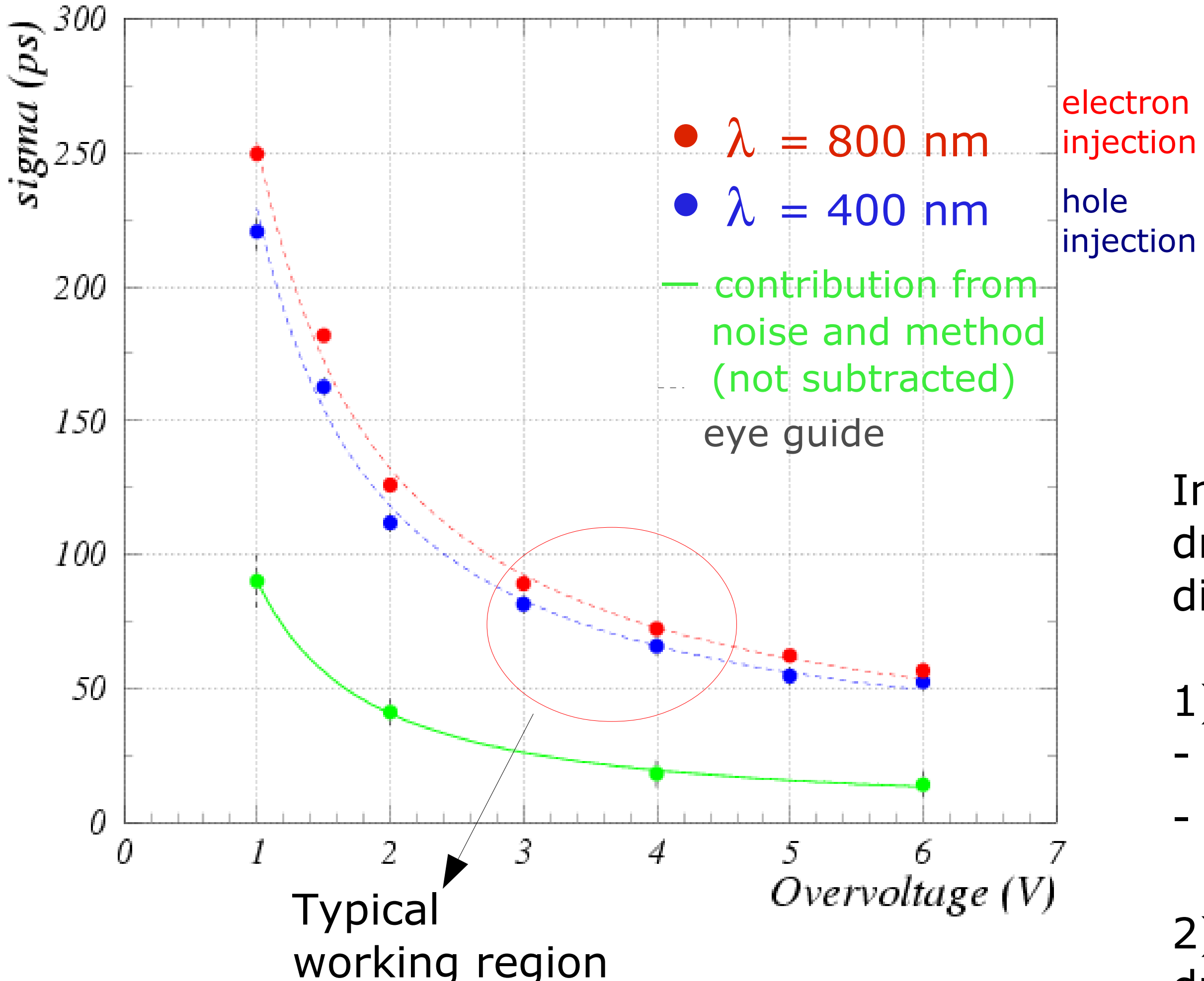
fit gives reasonable χ^2 in case of an **additional exponential term** exp(-|Δt|/τ) summed with a weight

- **τ ~ 0.2÷0.8ns** (depending on device) in rough agreement with diffusion tail lifetime: $\tau \sim L^2 / \pi^2 D$ wher L is the diffusion length
- Weight of the **exp. tail ~ 10%÷30%** (depending on device)

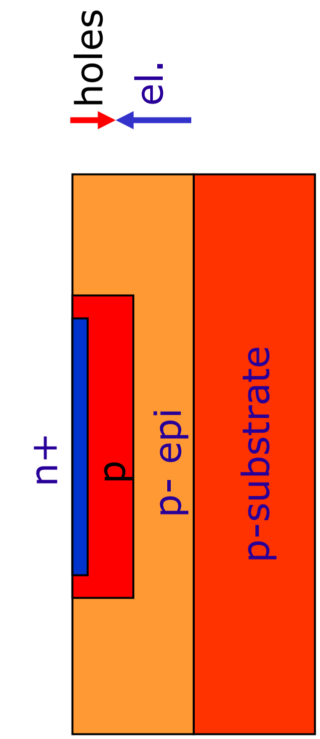


Distributions of the difference in time between successive peaks

SPTR: FBK devices – shallow junction



G.C. et al NIMA 581 (2007) 461



In general due to drift, resolution differences

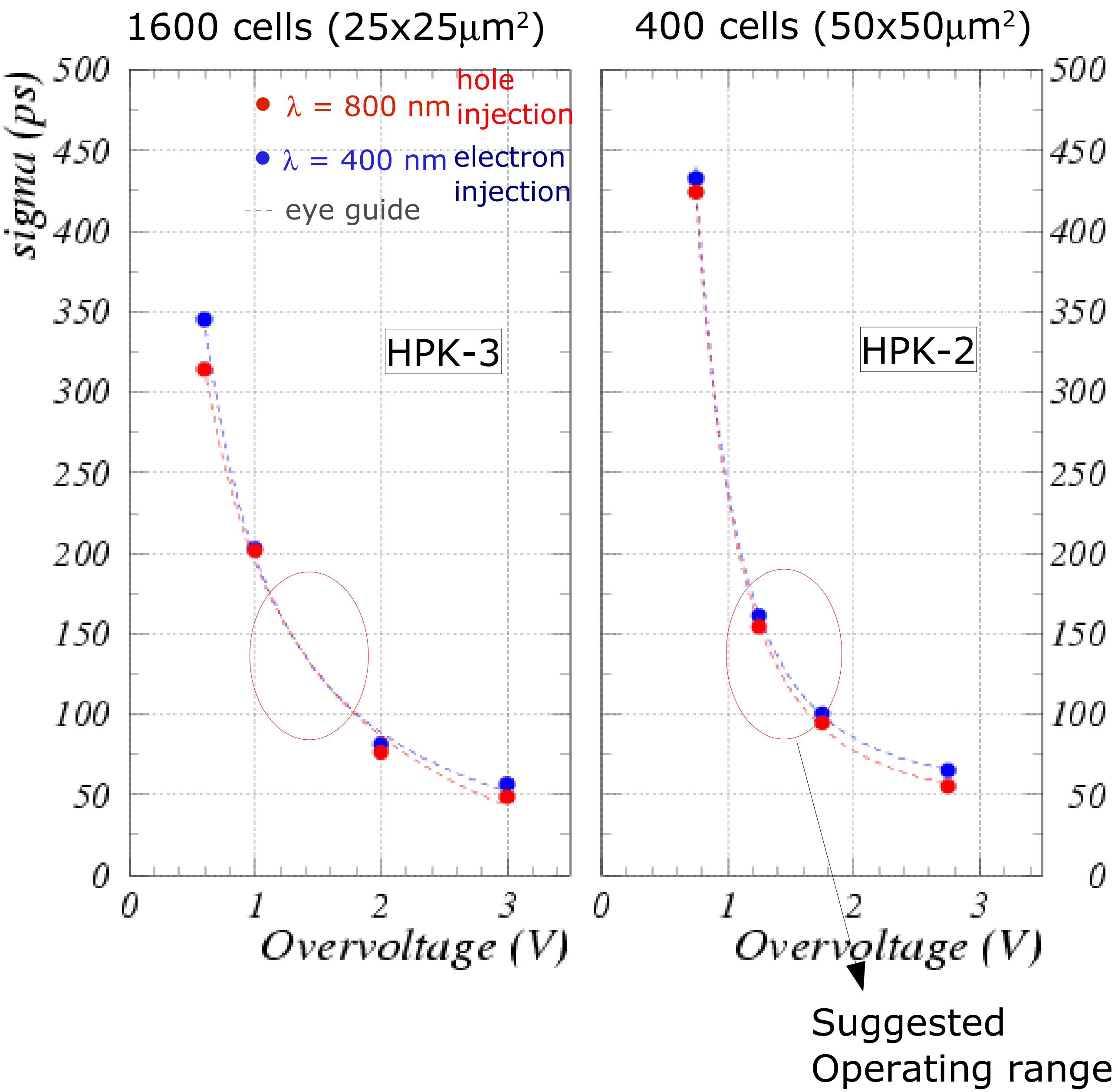
1) high field junction position
- shallow junction: $\sigma_t^{red} > \sigma_t^{blue}$
- buried junction: $\sigma_t^{red} < \sigma_t^{blue}$

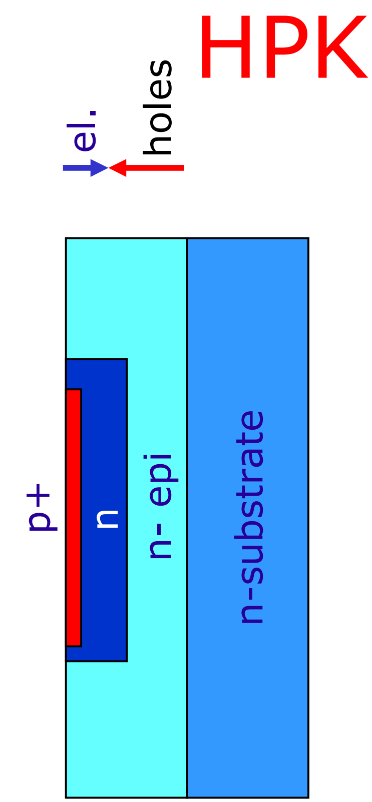
2) n^+-on-p smaller jitter than p^+-on-n due to electrons drifting faster in depletion region (but λ dependence)

3) above differences more relevant in thick devices than thin

NOTE: good timing performances kept up to 10MHz/mm^2 photon rates

SPTR: Hamamatsu

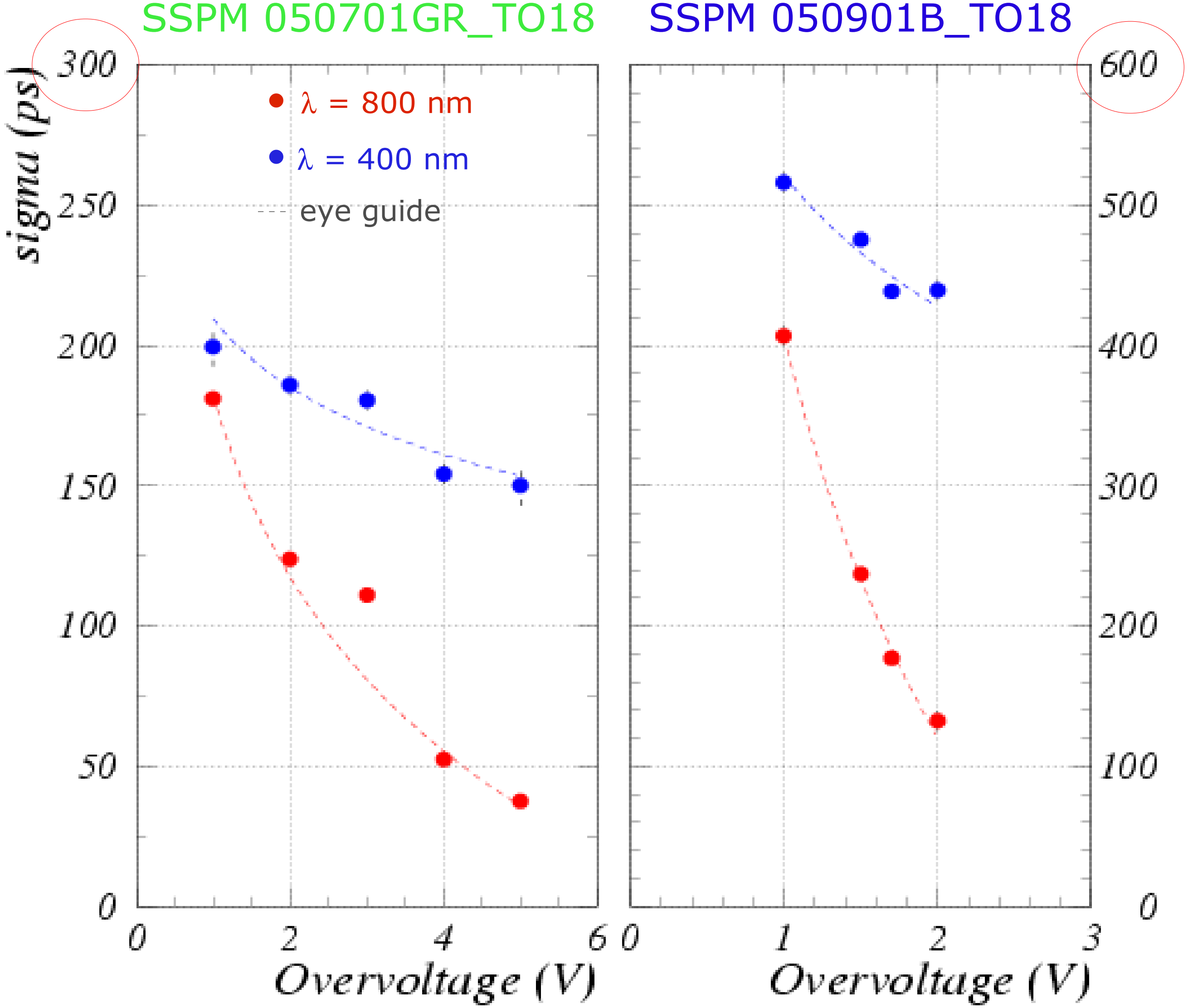




SPTR: CPTA/Photonique – thick structures

a) Green-Red sensitive
SSPM 050701GR_TO18

b) Blue sensitive
SSPM 050901B_TO18

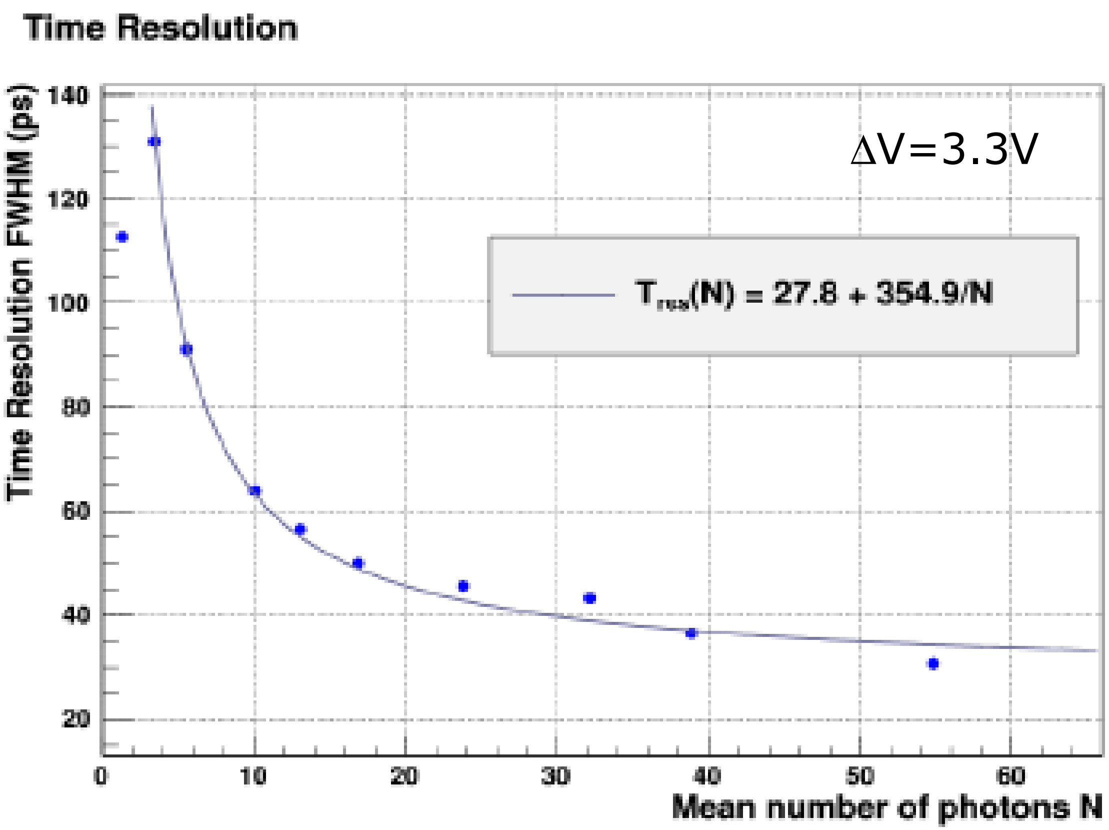


- thick structures
- deep junctions

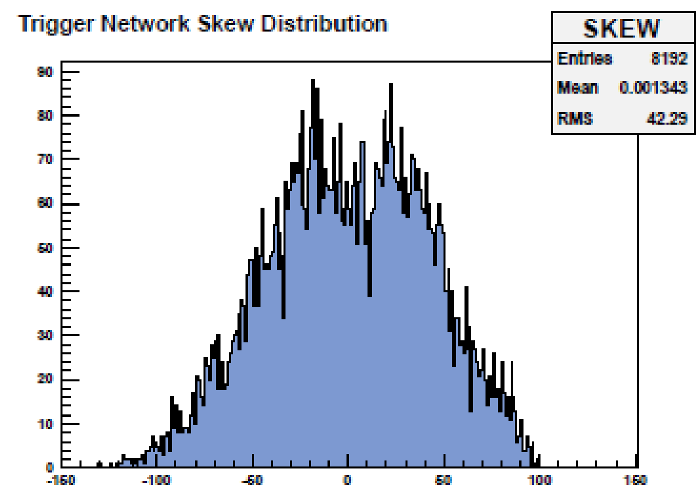
a) n^+-on-p
→ electrons drift

b) p^+-on-n
→ holes drift ($v_e/3$)

dSiPM timing resolution



T.Frach at LIGHT 2011

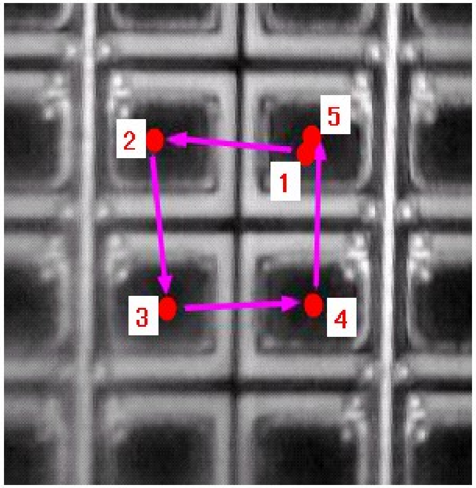


- Sensor triggered by attenuated laser pulses at first photon level
- Laser pulse width: 36ps FWHM, λ = 410nm
- Contribution to time resolution (FWHM):

  SPAD: 54ps, trigger network: 110ps, TDC: 20ps

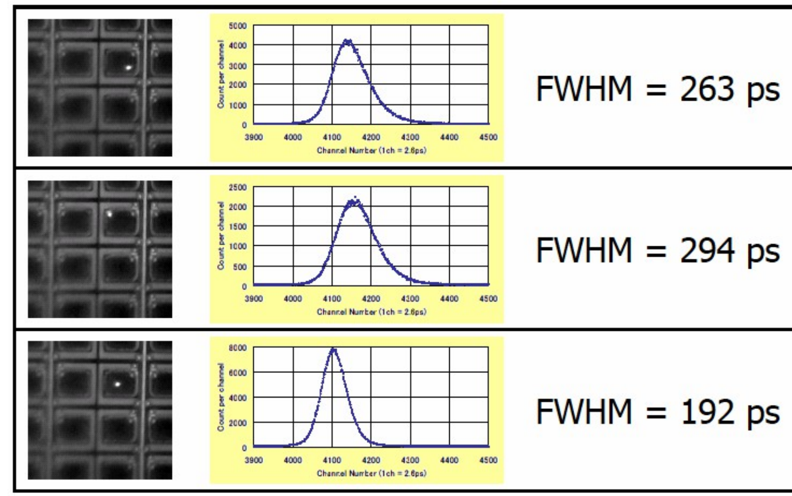
- Trigger network skew currently limits the timing resolution

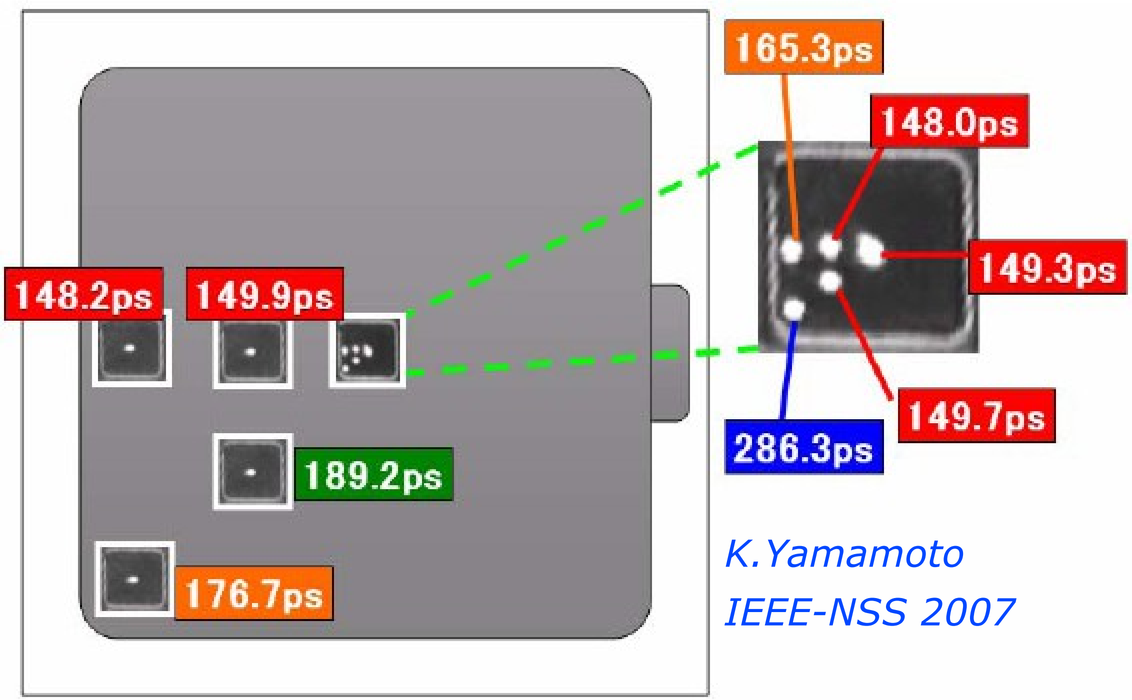
SPTR: position dependence → cell size



	FWHM (ps)	FWTM (ps)
1	199	393
2	197	389
3	209	409
4	201	393
5	195	383

K.Yamamoto PD07





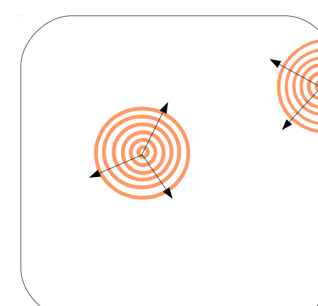
K.Yamamoto
IEEE-NSS 2007

Larger jitter if photo-conversion at the border of the cell

Due to:
1) slower avalanche front propagation

2) lower E field at edges

→ cfr PDE vs position



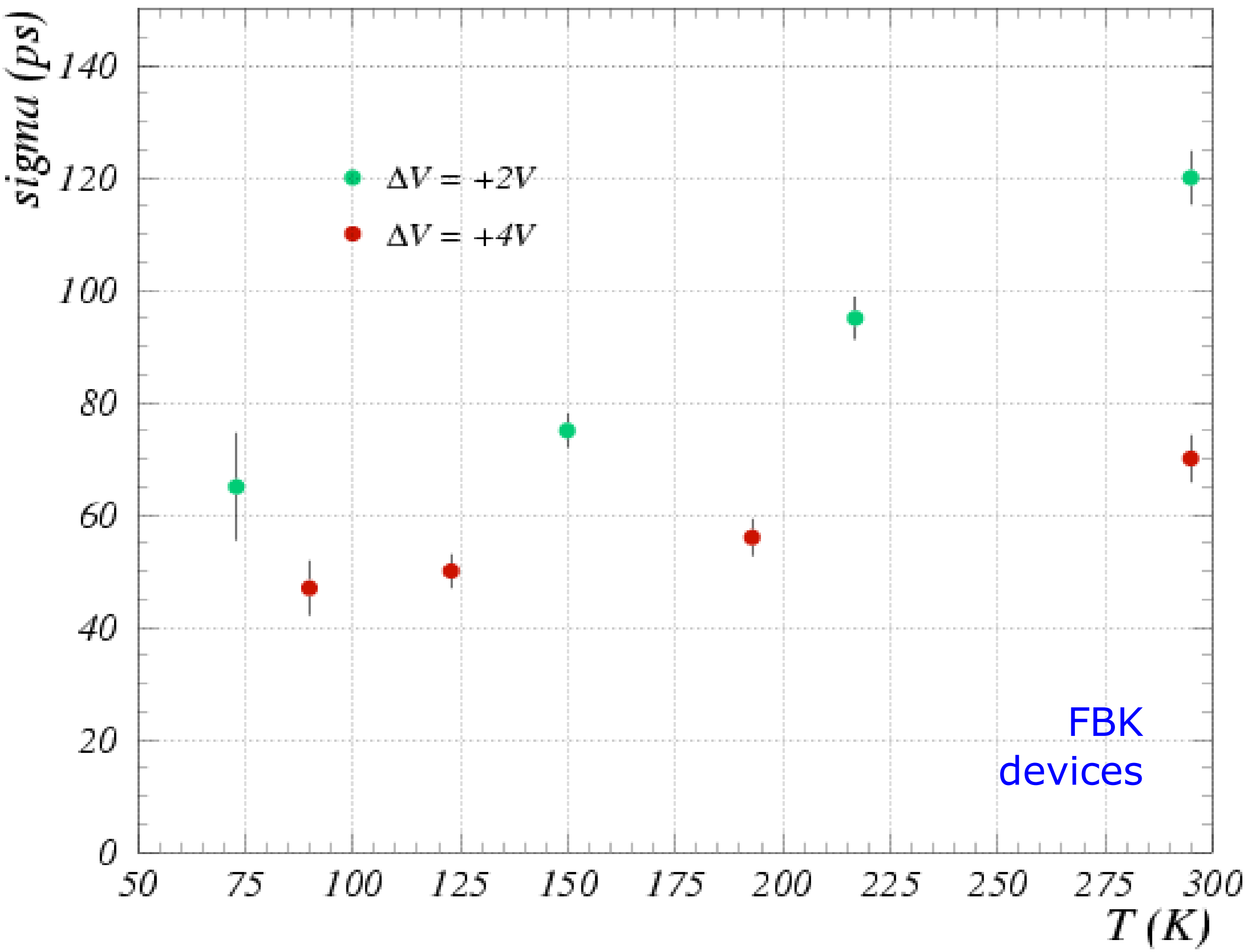
Data include the system jitter (common offset, not subtracted)

SPTR: timing at low T

Timing: improves at low T

Lower jitter at low T due to **higher mobility**: → a) avalanche process is faster; b) reduced fluctuations

(Over-voltage fixed)

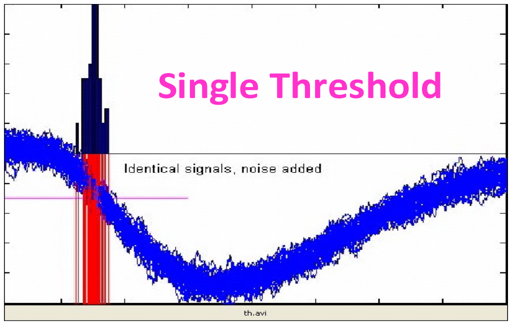


G.C. (2011, unpublished)

Note:

$$\frac{dI}{dt} \sim \frac{\sqrt{D}}{R_{sp}\sqrt{\tau}}$$

Optimizing signal shape for timing



Timing by (single) threshold:

→ time spread proportional to 1/rise-time and noise

$$\sigma_{time} = \frac{\sigma_{amplitude}}{\frac{df(t)}{dt}}$$

Timing with optimum filtering:

→ best resolution with f'(t) weighting function

$$\sigma^2_{time} = \frac{\sigma^2_{amplitude}}{\int dt \left[\frac{df(t)}{dt}\right]^2}$$

Pulse sampling and Waveform analysis:

Sample, digitize, fit the (known) waveform
→ get time and amplitude

$$\sigma^2_{time} = \frac{\sigma^2_{amplitude}}{N_{samples} \int dt \left[\frac{df(t)}{dt}\right]^2}$$

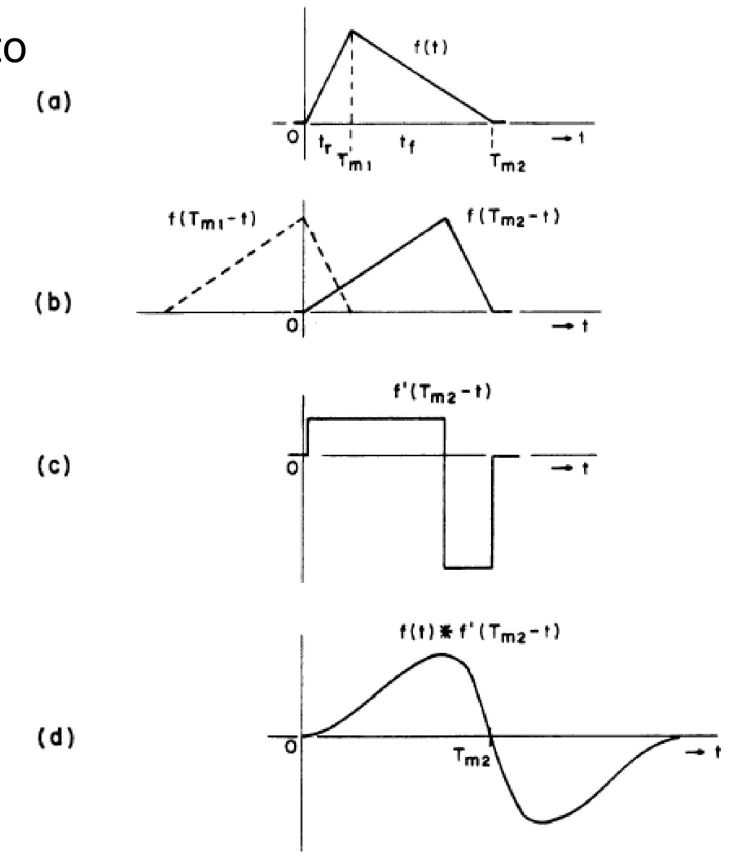


Fig. 7. Optimum filter for timing in presence of white noise (method of derivation).
(a) signal waveform
(b) optimum filter for amplitude measurements.
(c) optimum filter for timing - derivative of (b).
(d) output waveform.

V.Radeka IEEE TNS 21 (1974)...

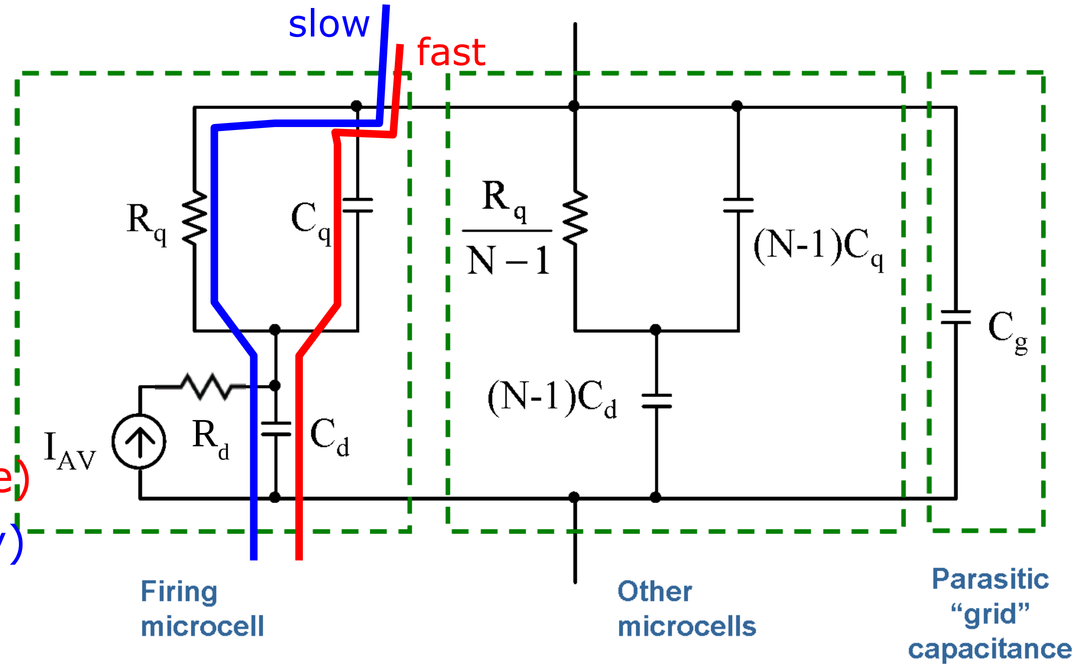
Optimizing signal shape for timing

Single cell model → $(R_d||C_d)+(R_q||C_q)$

SiPM + load → $(||Z_{cell})||C_{grid} + Z_{load}$

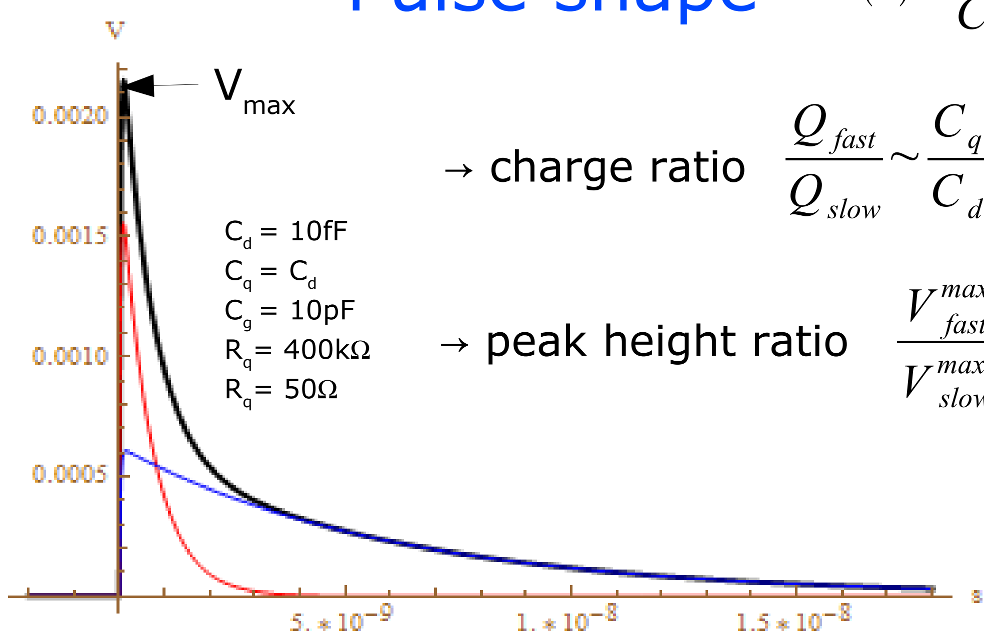
Signal = slow pulse ($\tau_{d\ (rise)}$, $\tau_{q\text{-slow}\ (fall)}$) +
+ fast pulse ($\tau_{d\ (rise)}$, $\tau_{q\text{-fast}\ (fall)}$)

- $\tau_{d\ (rise)} \sim R_d(C_q+C_d)$
- $\tau_{q\text{-fast}\ (fall)} = R_{load}\, C_{tot}$ (fast; parasitic spike)
- $\tau_{q\text{-slow}\ (fall)} = R_q\,(C_q+C_d)$ (slow; cell recovery)



Pulse shape

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



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

→ peak height ratio $\frac{V^{max}_{fast}}{V^{max}_{slow}} \sim \frac{C_q^2 R_q}{C_d C_{tot} R_{load}}$

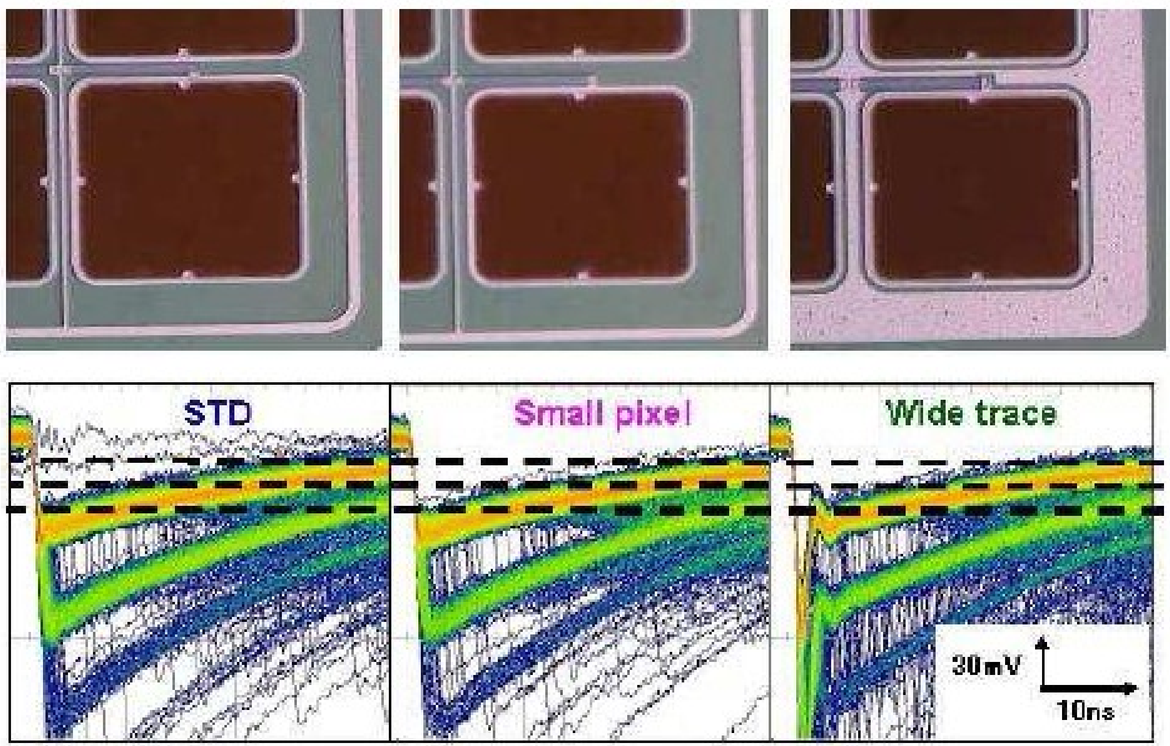
increasing with Rq and $1/R_{load}$ (and Cq of course)

Increasing C_q/C_d or/and R_q/R_{load}
→ spike enhancement
→ better timing

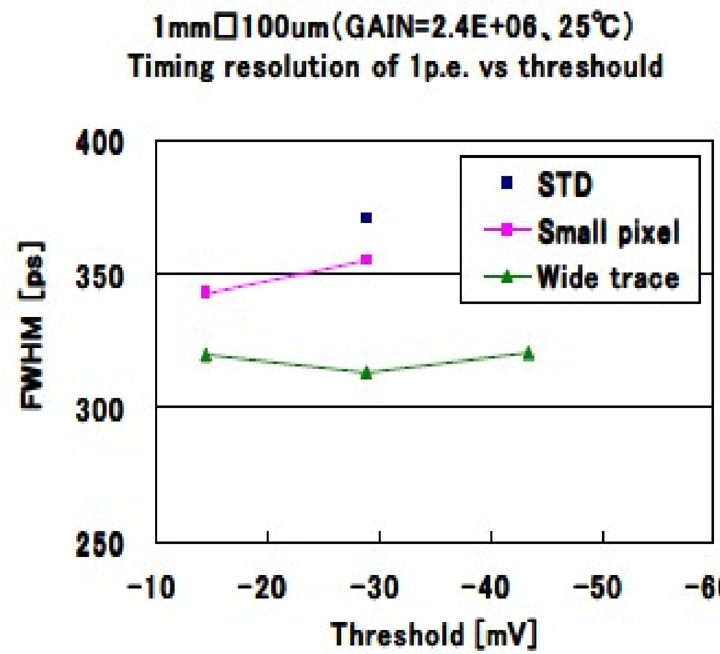
Optimizing signal shape for timing (SPTR)

→ peak height ratio $\frac{V_{fast}^{max}}{V_{slow}^{max}} \sim \frac{C_q^2 R_q}{C_d C_{tot} R_{load}}$

Enhancing C_q does improve timing performances



Yamamura et.al. at PD09



Analogous method for timing optimization proposed in C.Lee et al NIM A 650 (2010) 125 "Effect on MIM structured parallel quenching capacitor of SiPMs"

Note:
The steep falling front of the fast peak could be exploited too for optimum timing

$$\sigma_{time}^2 = \frac{\sigma_{amplitude}^2}{N_{samples} \int dt [f'(t)]^2}$$

Summary

Significant development of SiPMs over the last few years and **new players**

- **Operative ΔV over-bias range**:
  from 2V (eg HPK) to 10V (eg FKB) depending on E field profile and Rq

- **T coefficient**: low, below 0.3%/°C for many devices
  → might be lower, but tradeoff against PDE and noise

- **Pulse Shape** and **Gain**: tuned for matching application requirements (tradeoffs)
  → photon counting and timing vs energy measurement (signal spike, E_{field} profile)

- **Dynamic range:** Large, up a few x10000 pixels (eg NDL, Zecotek)
  → improved **radiation hardness** (not covered in this review) is relevant bonus
  → trade-off with Fill Factor

- **PDE**: up to 60% for blue-green light (eg. KETEK)
  → easily tuned to match applications (but only in visible optical range)

- **DCR** at T room can be < 100kHz/mm^2 (eg. Hamamatsu)

- **Cross-Talk**: can be as low as 1% in operative range (eg. FBK, MePhi/Pulsar)

- **After-Pulsing:** still at some % level for many devices
  → exploiting higher Rq "just" to hide A-P is not a good practice...
  → Digital SiPM is prone too, though less affected (active quenching)

- **Timing**: intrinsically fast, SPTR < 50ps in operative range
  → but mind the diffusion non-gaussian tails in temporal response (long λ)

- **Calibration:** precise, thanks to existing detailed operative models

Still missing and Future threads

- Avalanche **detailed physical models are still missing**. In particular for
  → **ultra-fast timing** applications there is room for device improvement
  → techniques for reducing **long timing tails** might be exploited

- Physical models might be of help also in further reducing **DCR** and **A-P**
  → eg: E field engineering for reducing tunneling

- **PDE**: expected soon are
  → improvements the **UV, VUV, EUV region**
  → devices with **through vias** → coupling with scintillators, fast imaging !

- GM-APD arrays for **NIR, IR sensitivity**: **different semiconductors**
  → InGaAs GM-APD arrays from AmplificationTechnologies do exist but... small area, noise and cost (!)

- **DCR**:
  → expected in 2012 a factor **x3 improvement** → larger area devices will follow
  → in the mean time devices tuned for working at **cryogenic T** easy to devise

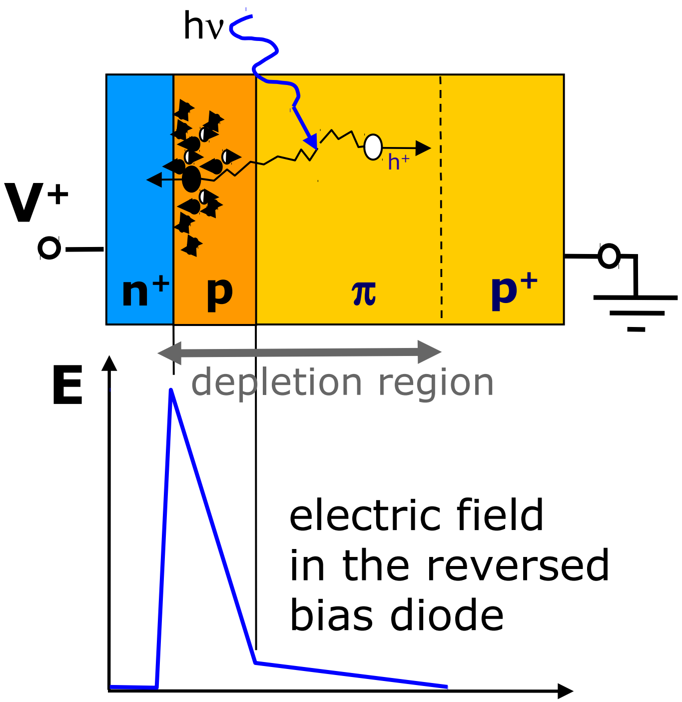
- **Low T**: SiPM perform ~ideally in the range 100K < T < 200K
  → Rq should be tuned shorter recovery (ad hoc devices)
  → lower gain (small cells) might be desired to mitigate after-pulses

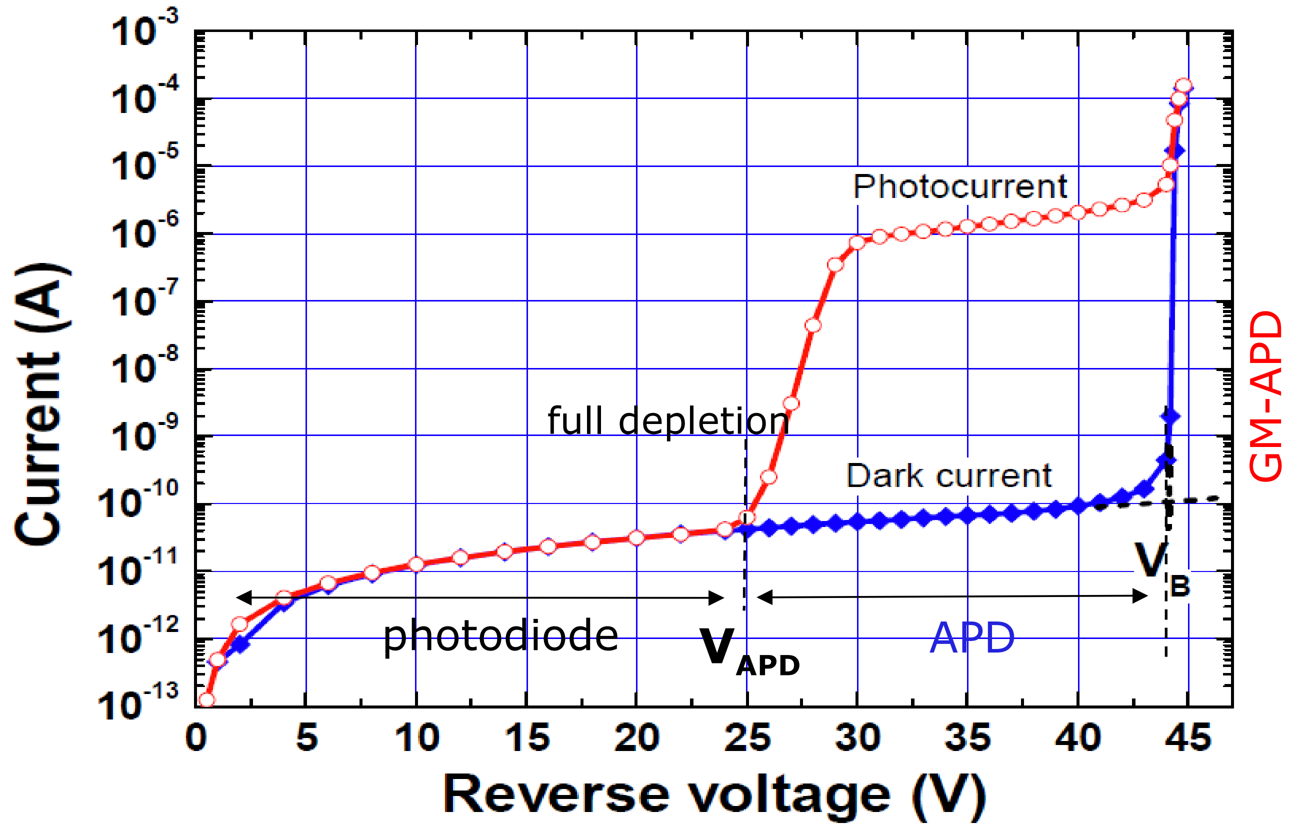
Thanks for your attention

Additional material

The building block of a SiPM: GM-APD

Reverse biased junction





APD: Linear-Proportional Mode

- Bias BELOW V_{BD} ($V_{APD} < V < V_{BD}$)
- It's an AMPLIFIER
- Multiplication: in practice limited to 10^4 by fluctuations
- No single photo-electron resolution

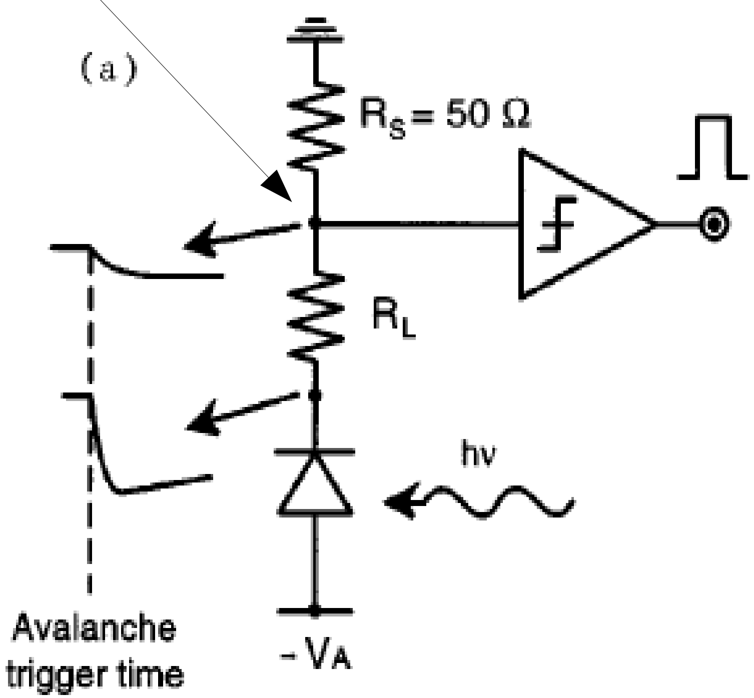
…except at low T with slow electronics,
Dorokhov et.al. J.Mod.Opt. 51 (2004)

GM-APD: Geiger Mode

- Bias ABOVE V_{BD} ($V-V_{BD}$ ~a few volts)
- It's a TRIGGER (BINARY) device
- Multiplication: ∞… in practice limited by macroscopic parameters (R,C)
- Limited by dark count rate
- Single photo-electron resolution
- Need Reset (Feedback - Quenching)

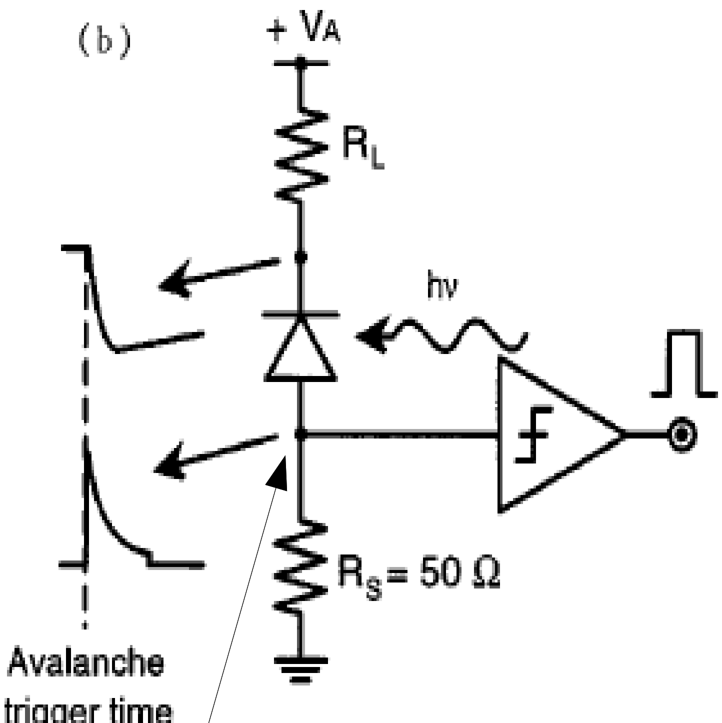
Readout Mode

High Z node



Voltage Mode

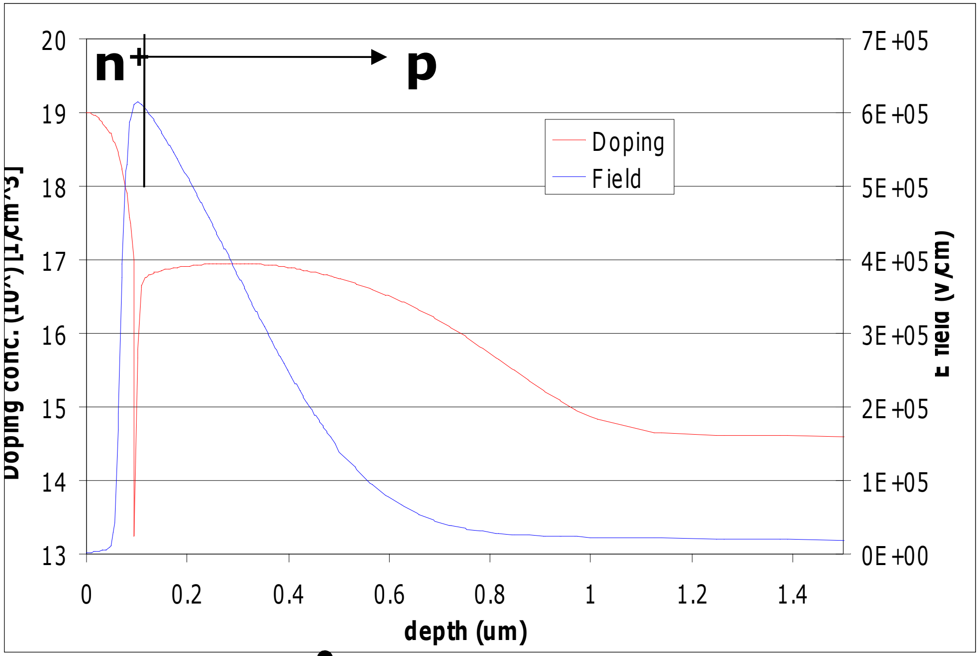
Current Mode



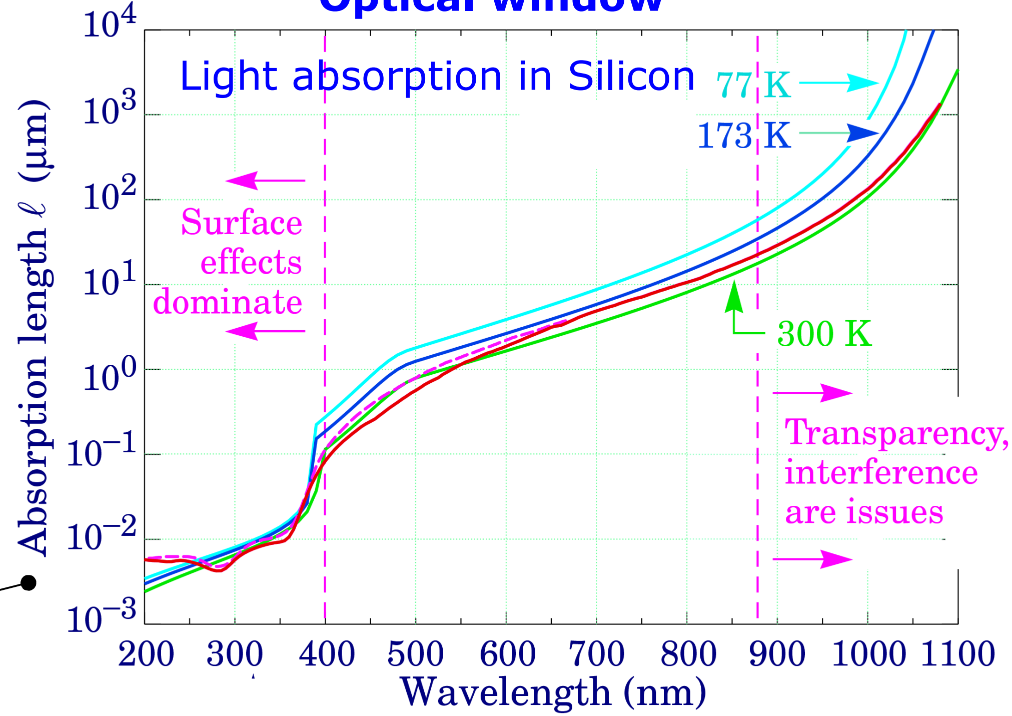
Low Z node

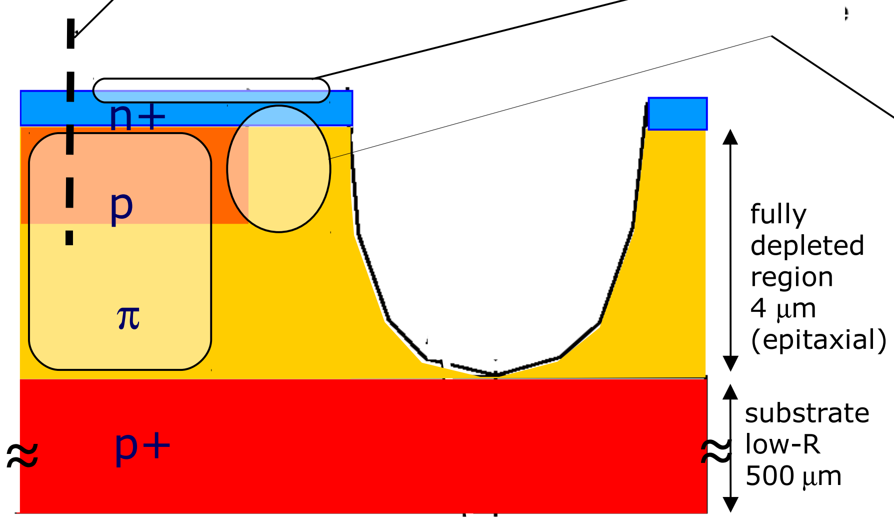
Key elements in SiPM cell

Doping and Field profiles



Optical window

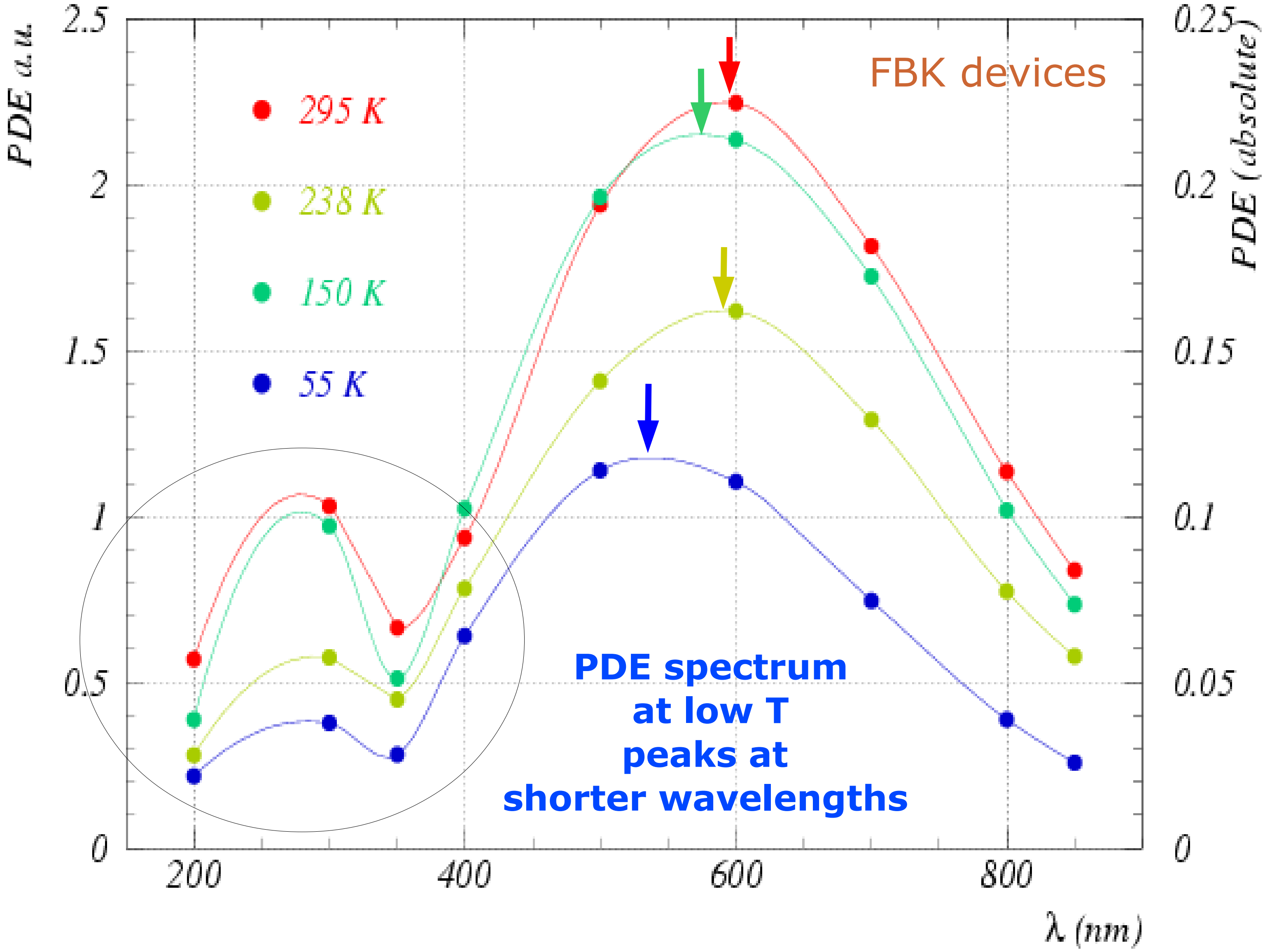




Guard Ring:

→ for avoiding early edge breakdown
→ for isolating cells
→ for tuning E field shape
→ has important impact on fill factor (more than Rq and metal grid)

PDE vs λ (ΔV fixed, various T)



G.C. et al NIM A628 (2011) 389

RPL model: fast simulation

"Statistics of Avalanche Current Buildup Time in Single-Photon Avalanche diodes"
C.H.Tan, J.S.Ng, G.J.Rees, J.P.R.David (Sheffield U.)
IEEE J.Quantum Electronics 13 (4) (2007) 906

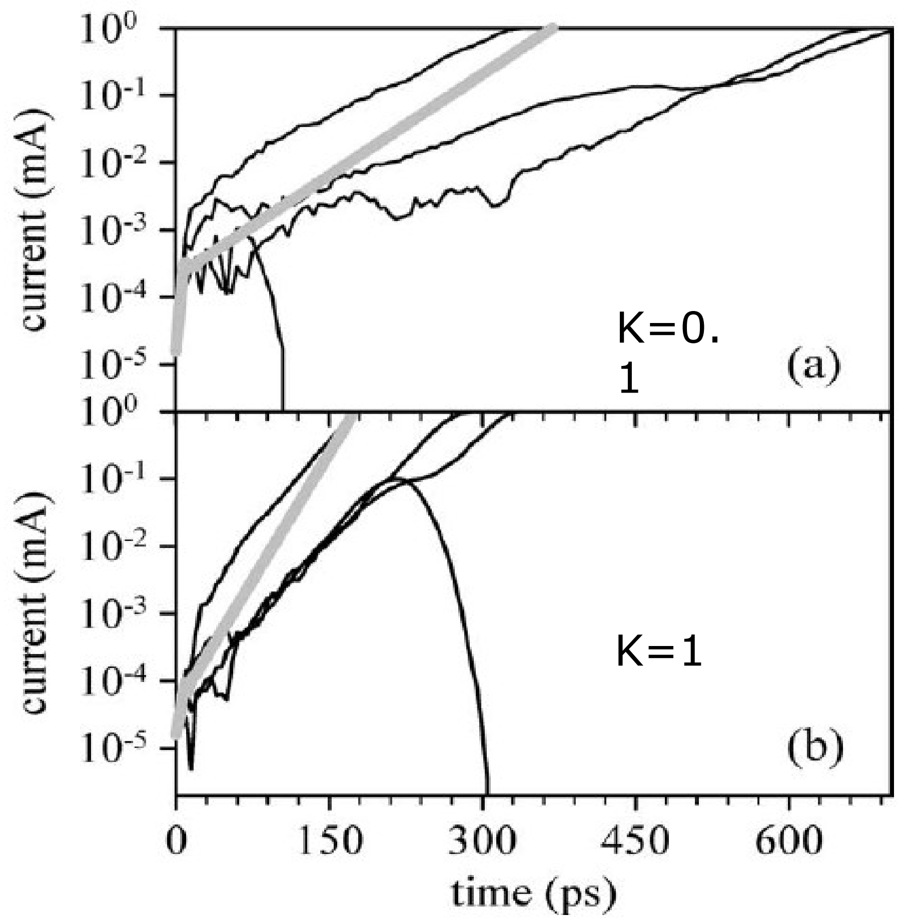
Numerical model (MC): Random distribution of impact ionization Path Length (RPL)

Analysis of breakdown probability, breakdown time and timing jitter as functions of avalanche region width (w), ionization coefficient ratio ($k=\beta_{hole}/\alpha_{electron}$) and dead space parameter (d) (uniform E field, constant carrier velocity)

1) increasing k:
- improves timing performances
- but breakdown probability P_{br} increases slowly with overvoltage

1a) hole injection results in better timing than electron injection (in Si devices)

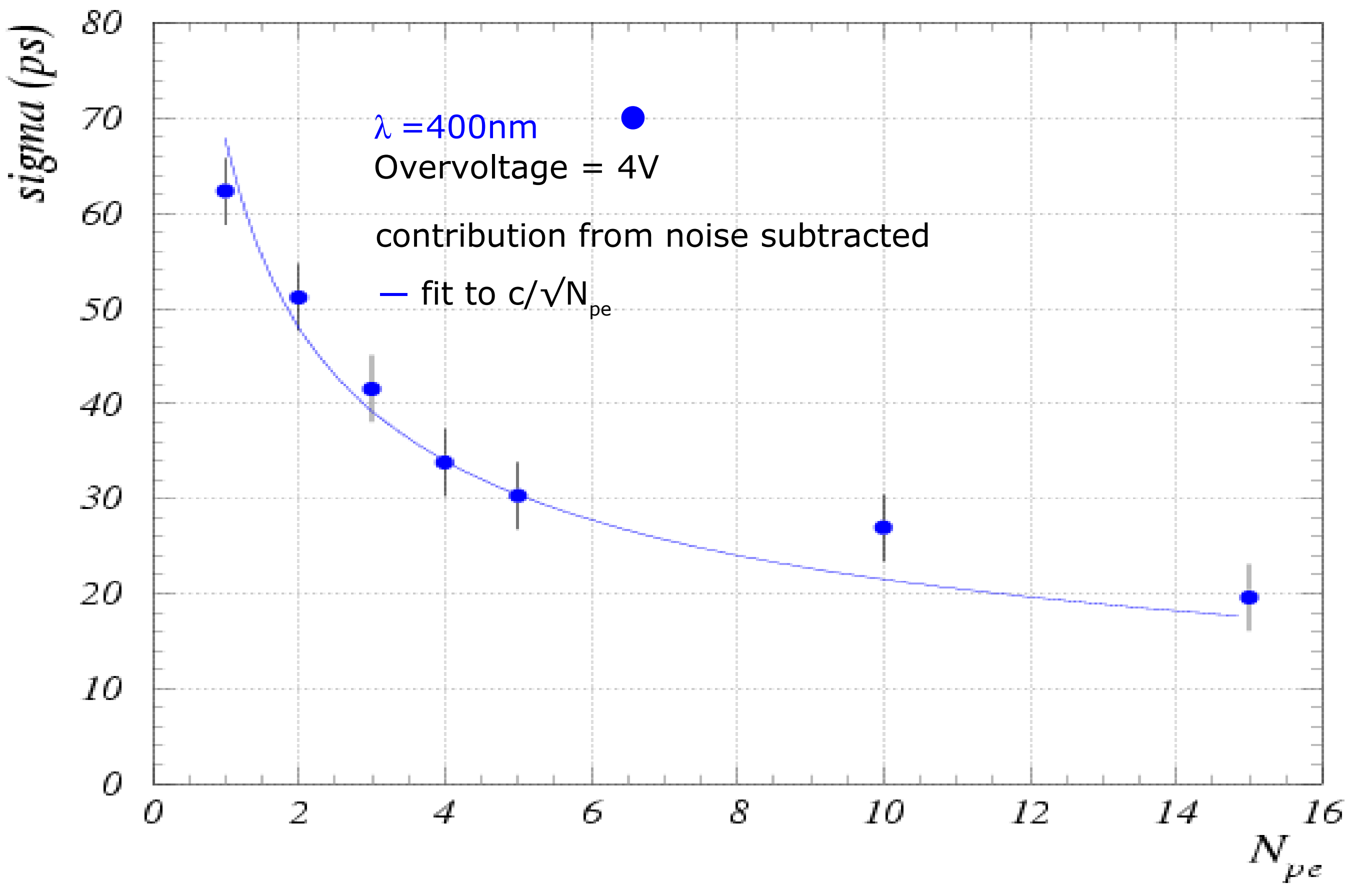
2) dead space effects worsen timing performances (the more at small k) Important for devices with small w



Many photons (simultaneous)

Dependence of SiPM timing on the number of simultaneous photons

Poisson statistics: $\sigma_t \propto 1/\sqrt{N_{pe}}$



N of simultaneous photo-electrons

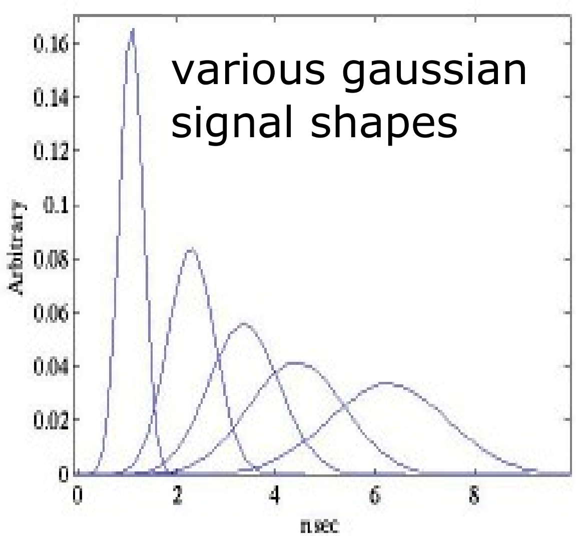
G.Collazuol - PhotoDet 2012

Signal shape for timing - many photons

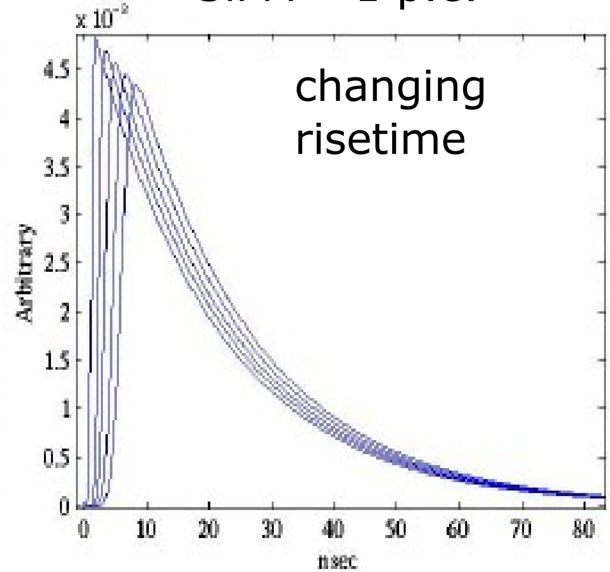
Single p.e. signal slow falling-time component $\tau_{fall} = R_q\ (C_d + C_d)$

strongly affects multi-photon signal risetime

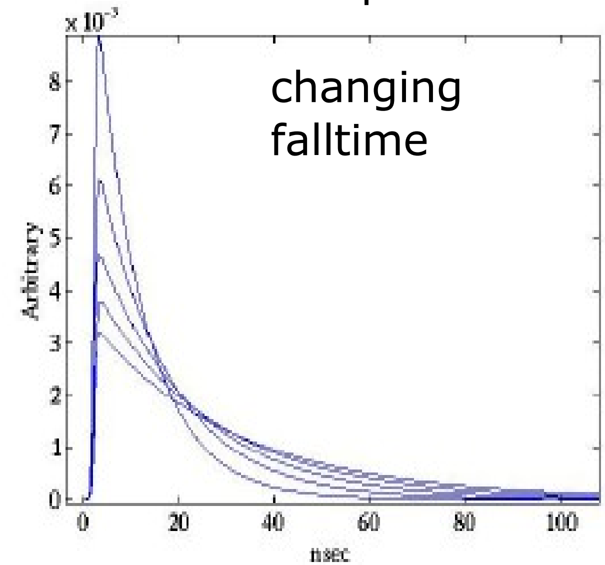
PMT – 1 p.e.



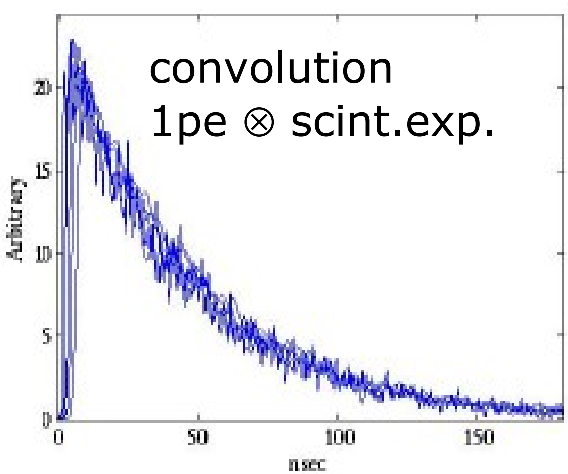
SiPM – 1 p.e.



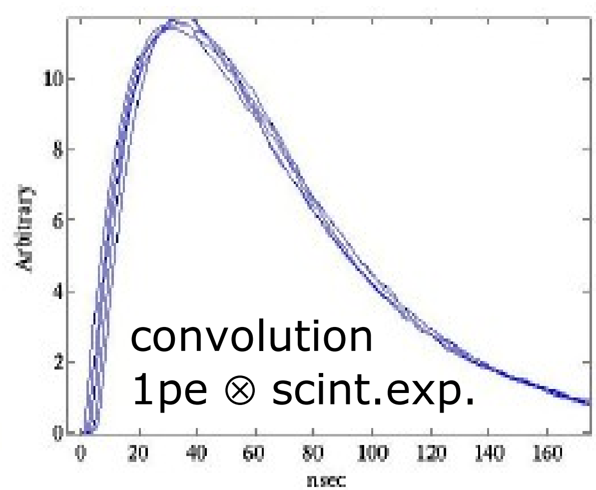
SiPM – 1 p.e.



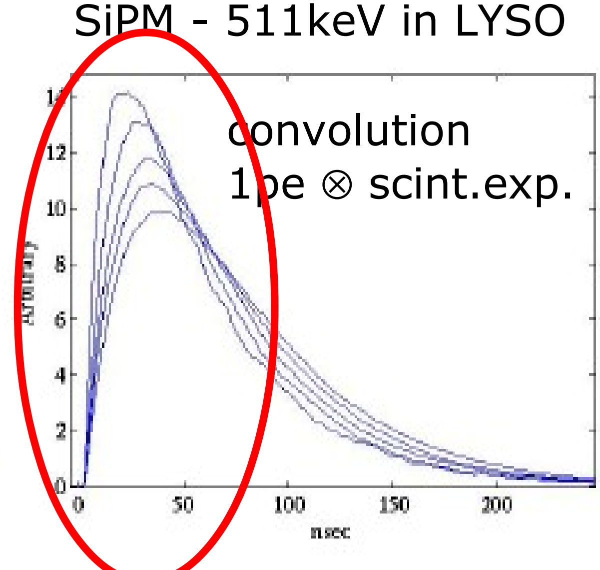
PMT - 511keV in LYSO



SiPM - 511keV in LYSO



SiPM - 511keV in LYSO



C.L.Kim Procs of Sci. 2009 010 (PD09)

Optimizing shape for timing - many photons

→ peak height ratio $\frac{V_{fast}^{max}}{V_{slow}^{max}} \sim \frac{C_q^2 R_q}{C_d C_{tot} R_{load}}$

Enhancing C_q and R_q does improve timing performances

FBK devices type:

- Active area: 4x4mm^2;
- Cell size: 67x67µm^2;
- Fill factor: 60%;
- C_Q+C_D: about 180fF;
- R_Q: 1.1M;
- Dark noise rate:
  ~100MHz at DV> 4V

C.Piemonte et al IEEE TNS (2011)

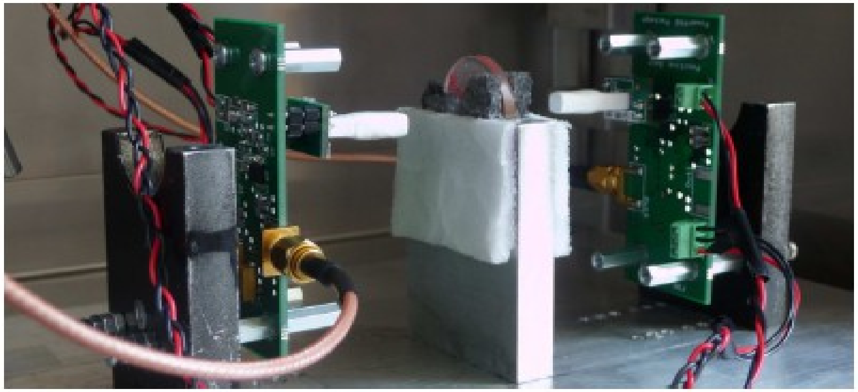
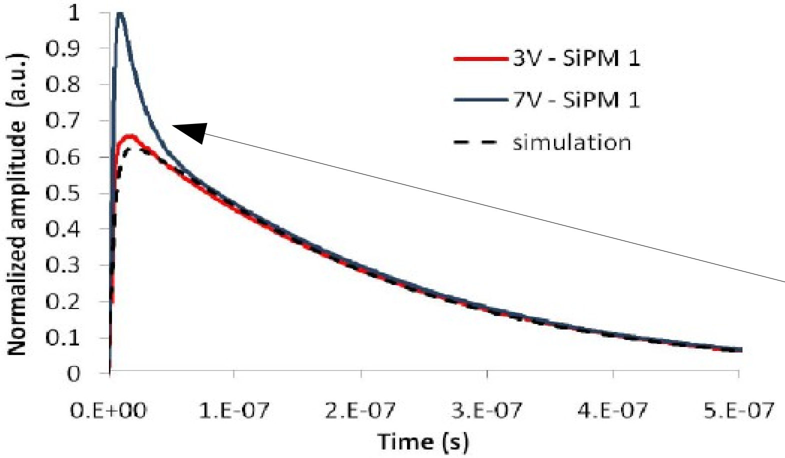


Fig. 2. Test set-up consists of two similar gamma ray detectors (LYSO crystal + SiPM) in coincidence. A ^{22}Na source (disc in the middle) was used to generate two opposite 511keV photons in coincidence.



- Signal rise-time < 5ns
- CRT ~320ps (*) FWHM triggering at 5% height

Both are much better than for different structures with high C_{tot} and/or lower Cq, Rq (risetime up to several x 10ns, CRT > 400ps)

??? peak shape is not scaling with ΔV (non linearity in the F.Corsi etal electrical model) Can be corrected → energy resol. ~11%

(*) ~40% from light propagation in crystals

Radiation damage

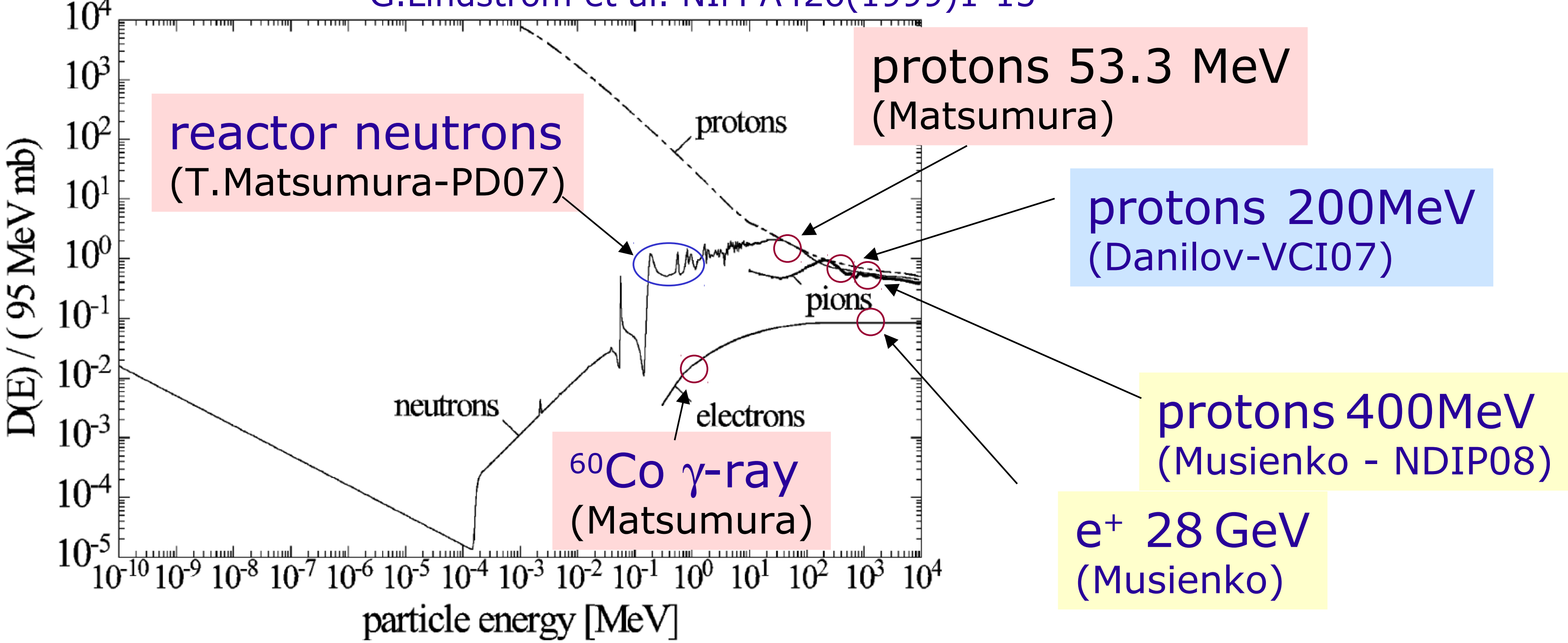
Note:
→ small cells smaller charge flow (small gain, high dynamic range)
→ small epi-layed width

Radiation damage: two types

- Bulk damage due to Non Ionizing Energy Loss (NIEL) ← neutrons, protons
- Surface damage due to Ionizing Energy Loss (IEL) ← γ rays
  (accumulation of charge in the oxide (SiO2) and the Si/SiO2 interface)

Assumption: damage scales linearly with the amount of Non Ionizing Energy Loss (NIEL hypothesis)

G.Lindstrom et al. NIM A426(1999)1-15



Examples of radiation tolerances for HEP and space physics

ATLAS inner detector ... 3×10^{14} hadrons/cm^2/10 year

~ 10^4 hadrons/mm^2/s

General satellites ... ~ 10 Gy/year

Expectations:

protons / γ-ray ~ 100

protons / neutrons ~ 2~10

Radiation damage: effects on SiPM

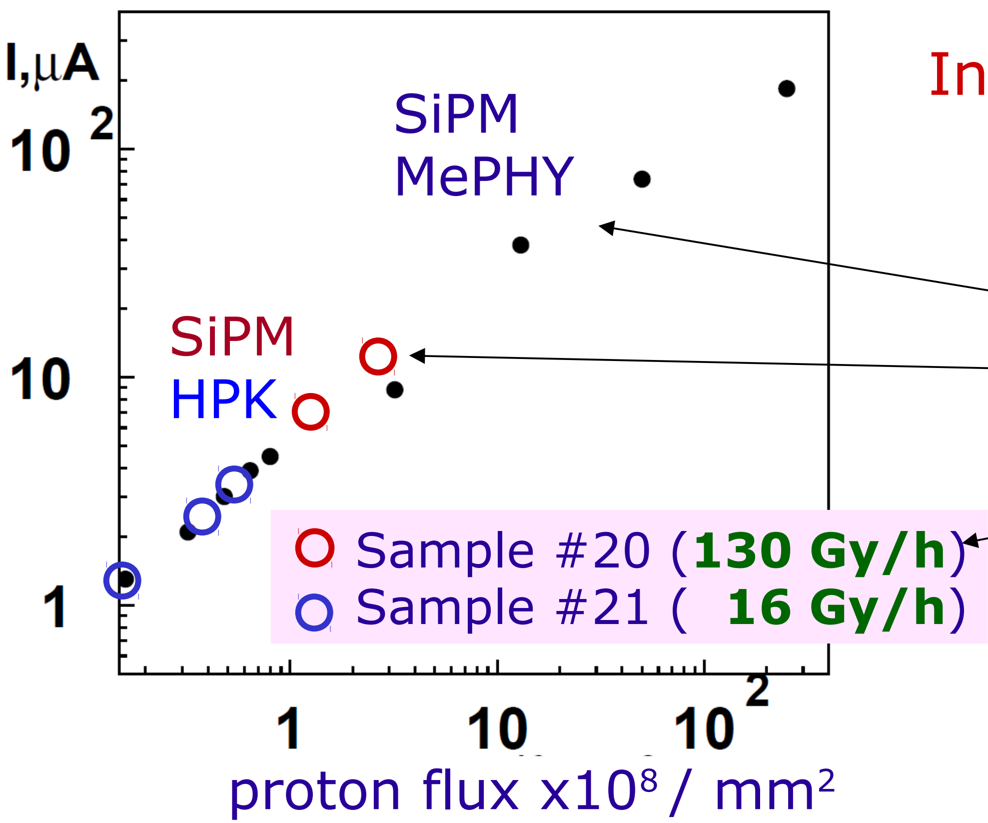
1) Increase of dark count rate due to introduction of generation centers

Increase (ΔR_{DC}) of the dark rate:
$\Delta R_{DC} \sim P_{01}\ \alpha\ \Phi_{eq}\ Vol_{eff} / q_e$
where $\alpha \sim 3 \times 10^{-17}$ A/cm is a typical value of the radiation damage parameter for low E hadrons and $Vol_{eff} \sim Area_{SiPM} \times \varepsilon_{geom} \times W_{epi}$

NOTE:
The effect is the same as in normal junctions:
- independent of the substrate type
- dependent on particle type and energy (NIEL)
- proportional to fluence

2) Increase of after-pulse rate due to introduction of trapping centers

→ loss of single cell resolution → no photon counting capability



Indications from measurements:

1) no dependence on the device
similar effects found for SiPM from MePHY (Danilov) and HPK (Matsumura) (normaliz. to active volume)

2) no dependence on dose-rate
HPK (Matsumura)

3) n similar damage than p
4) p $x10^1$-10^2 more damage than γ

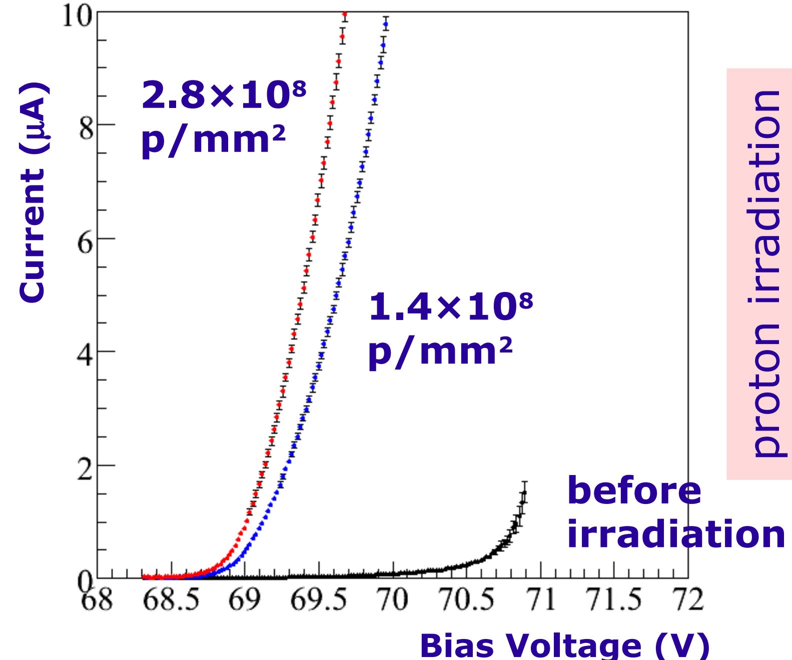
G.Collazuol - PhotoDet 2012

Damage comparison

2.3×10^5 p/mm^2/s (130 Gy/h)

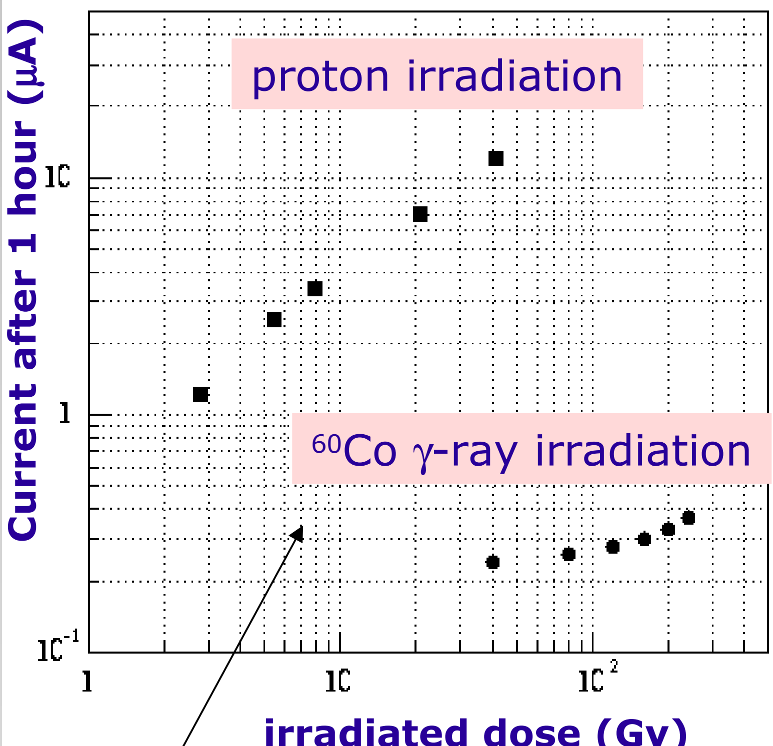
I_{leak} @ (V_{op}, 1.4×10^8 p/mm^2) = 6.7 μA

Damage effect …
almost the same for protons and neutrons



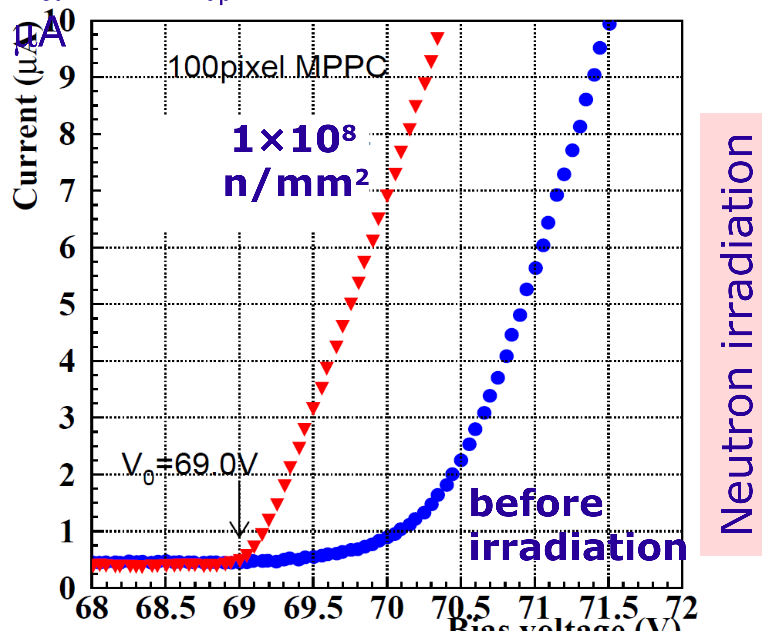
proton irradiation

HPK devices
T.Matsumura – PD07



4.2×10^5 n/mm^2/s

I_{leak} @ (V_{op}, 1.0×10^8 n/mm^2) = 8.5 μA



Neutron irradiation

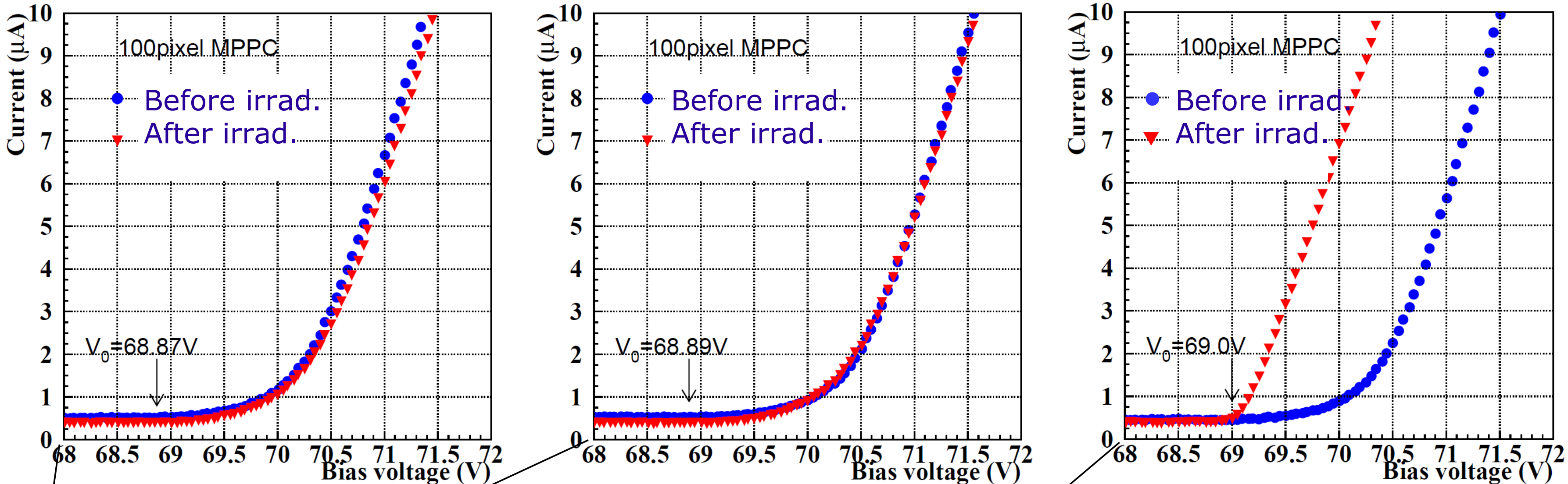
T.Matsumura – PD07

Damage effect …
1~2 orders larger with protons than γ-ray irradiation

G.Collazuol - PhotoDet 2012

Radiation damage: neutrons (0.1 -1 MeV)

8.3×10^4 n/mm^2 | 3.3×10^5 n/mm^2 | 1.0×10^8 n/mm^2

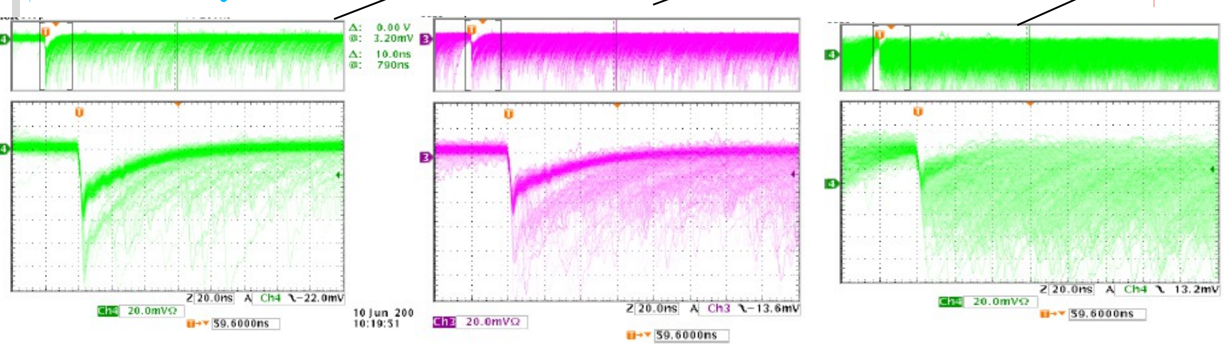


T.Matsumura – PD07

10^5 n/mm^2 | 10^6 n/mm^2 | 10^7 n/mm^2 | 10^8 n/mm^2 | 10^9 n/mm^2 | 10^{10} n/mm^2

n dose

No significant change

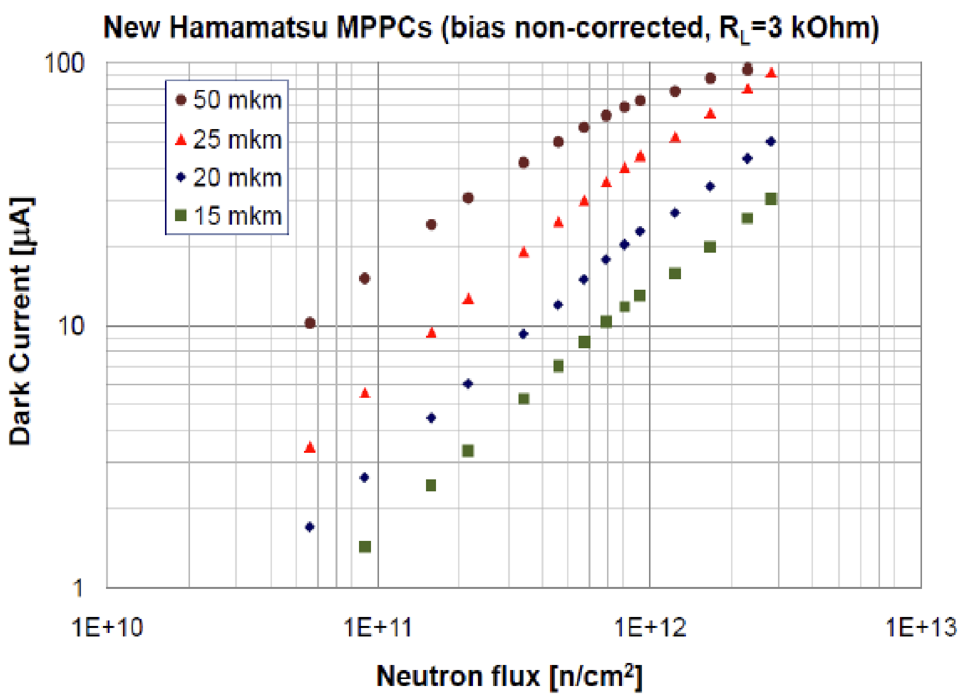


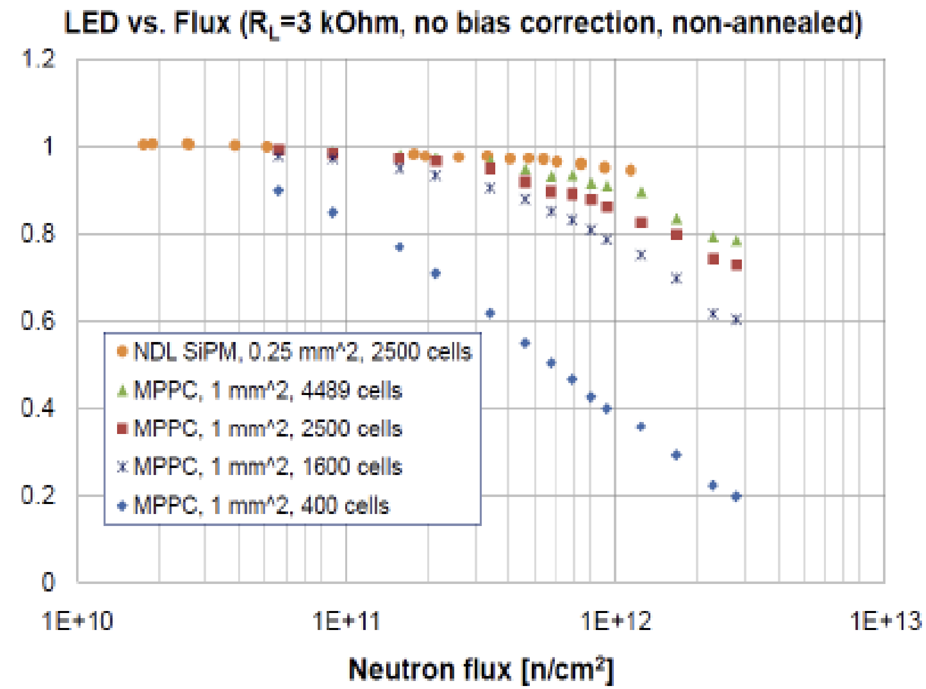
I-V drastically change.
Signal pulse is still there,
but continuous pulse height.
(No photon-counting capability)

No signal

Nakamura at NDIP08

Radiation damage: neutrons 1 MeV E_{eq}





- No change of V_{bd} (within 50mV accuracy)
- No change of R_q (within 5% accuracy)
- I_{dark} and DCR significantly increase

SiPMs with high cell density and fast recovery time can operate up to $3*10^{12}$ n/cm^2 ($\delta G < 25\%$)

Y.Musienko at SiPM workshop CERN 2011