APDs and Geiger-mode APDs for Cherenkov detectors

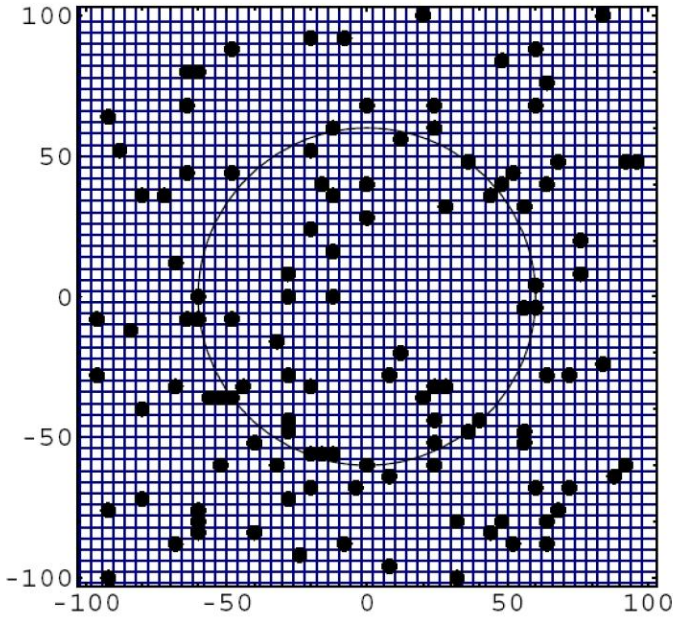
Dieter Renker

Paul Scherrer Institute, Villigen, Switzerland

What is needed?

Single photon detection with:

- High detection efficiency
- Low noise
- Excellent timing
- High rate capability
- Operation in high magnetic field



P. Križan, Light07

Outline

Avalanche photodiodes (APDs)

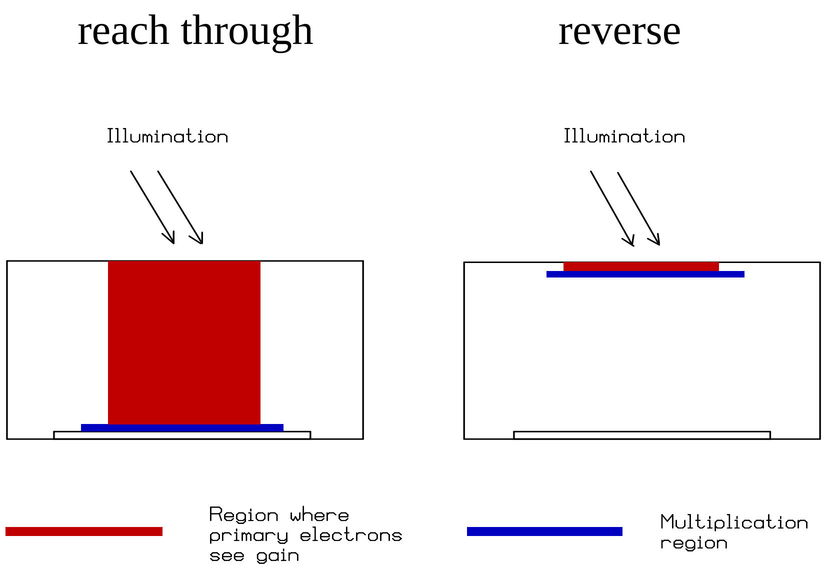
- Limited sensitivity to single photons

Hybrids with APDs

Geiger-mode avalanche photodiodes (G-APDs)

- High Gain and saturation
- Photo detection efficiency PDE
- Optical crosstalk
- Dark count rate
- Afterpulsing probability
- Recovery time
- Temperature dependence
- Fast response
- Radiation hardness
- Conclusion

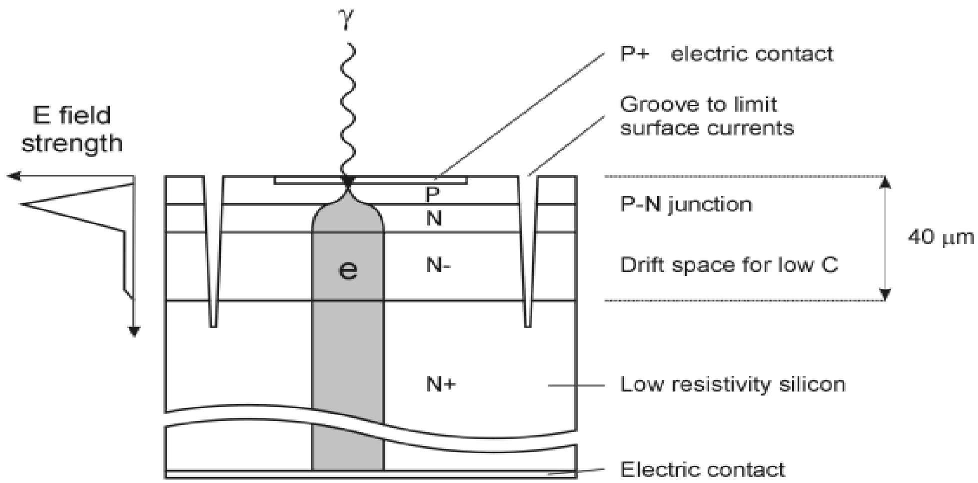
Types of APDs: “reach through” or “reverse”

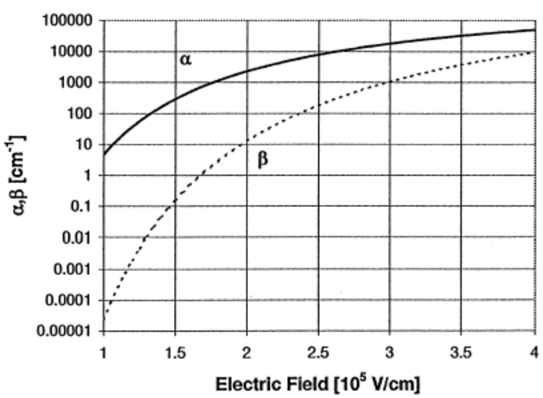


The APD’s with “reverse” structure have improved characteristics:

- high speed
- good visible and blue response
- reduced temperature dependence
- small excess noise factor
- small nuclear counter effect
- better radiation resistance

Basic APD Structure (CMS version)





Ionization coefficients α for electrons and β for holes

Photons create electron-hole pairs in the thin p-layer on top of the device and the electrons induce avalanche amplification at the p-n junction.
Holes contribute little because of their much smaller ionization coefficient.

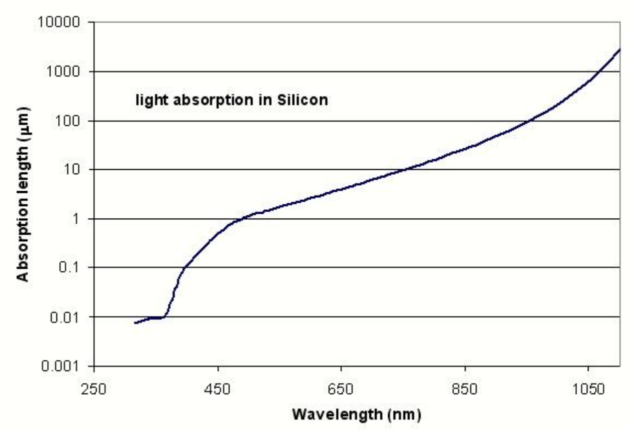
Electrons created by ionising particles traversing the bulk are not amplified.

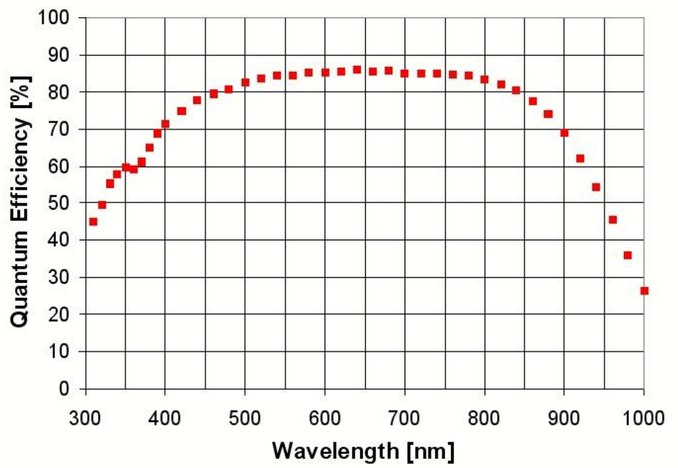
$\rightarrow d_{eff} \sim 6\ \mu m$

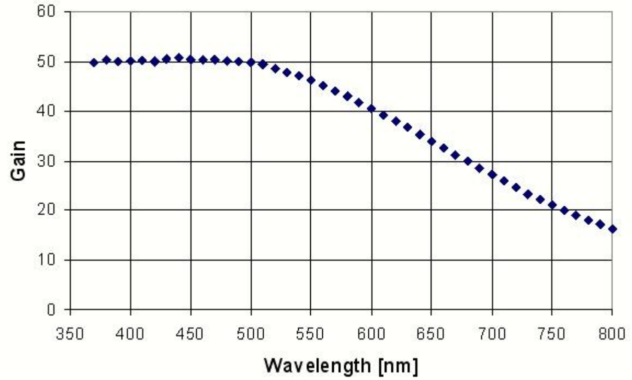
50 times smaller NCE than in a PIN diode.

Photo detection efficiency (PDE)

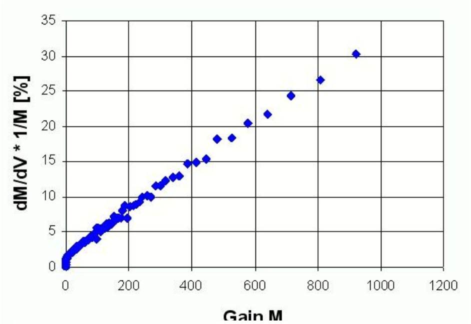
For an APD the PDE is equal to the Quantum efficiency (QE) – no correction for collection efficiency etc.

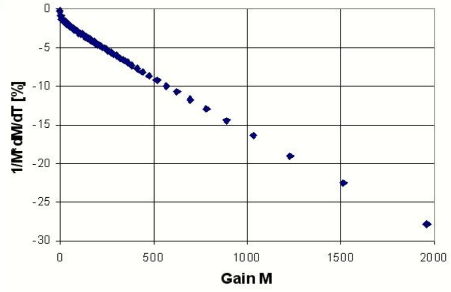


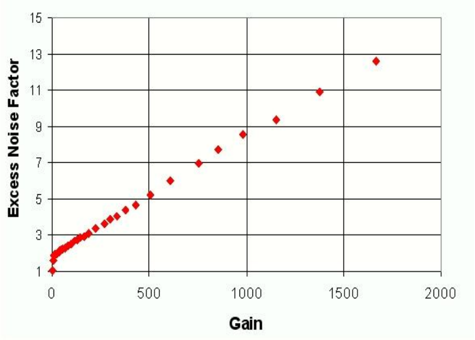




The gain of an APD is limited







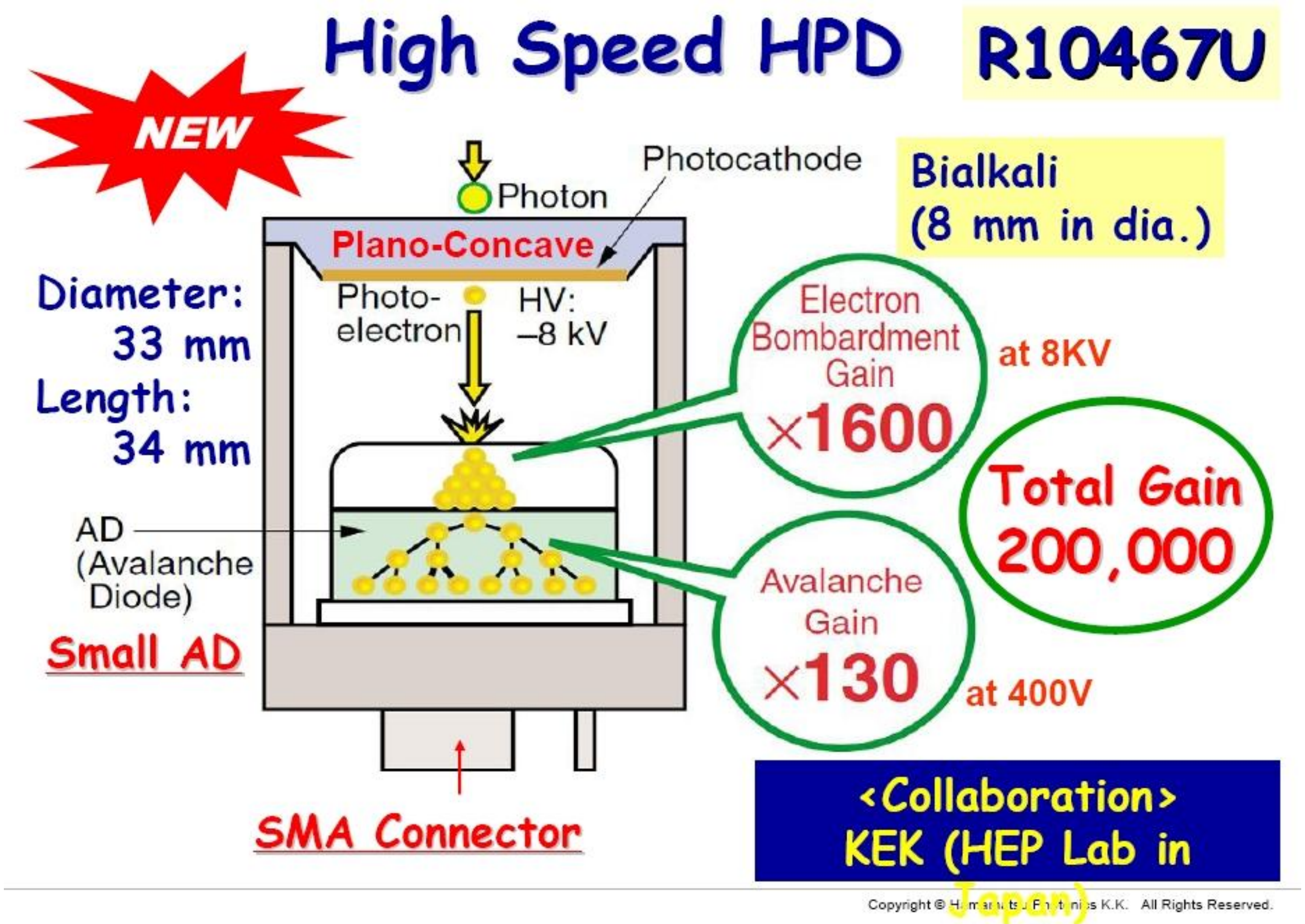
High gain (several 100) is only possible when the environment is very stable. The excess noise factor becomes large.

Even at operation with high gain only light flashes with 20 or more photons can be separated from the noise.

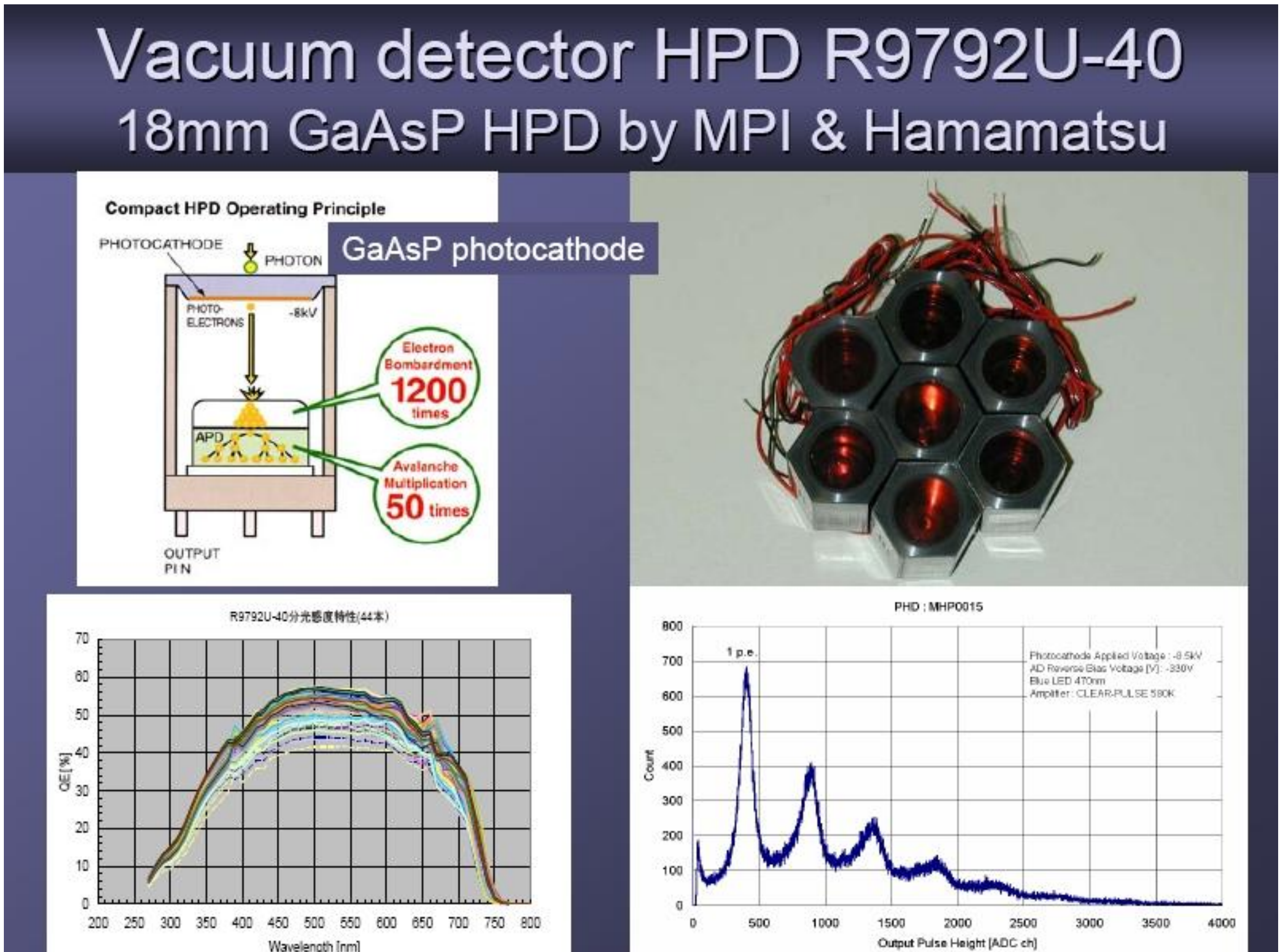
$$ENC^2 \approx 2q \cdot \left(\frac{I_{ds}}{M^2} + I_{db} \cdot F \right) \cdot \tau + 4kTR_s \cdot \frac{C^2}{M^2} \cdot \frac{1}{\tau}$$

Only operation at cryogenic temperatures allows single photon detection.

APDs are very useful in hybrid PMTs



HPDs with high QE are available

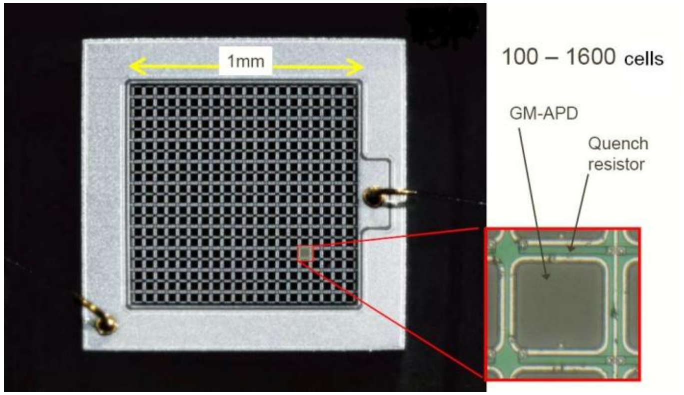


Geiger-mode APDs: Principle of operation

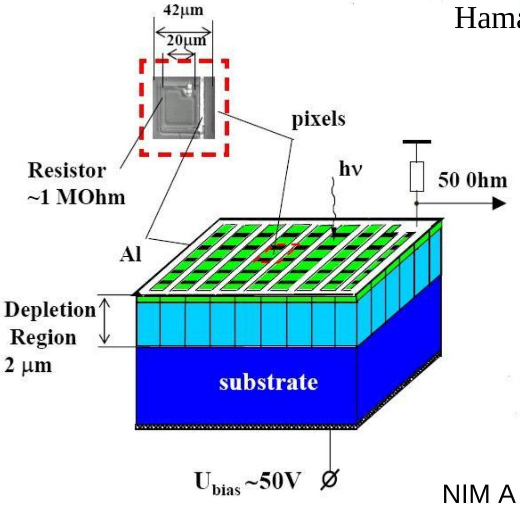
In principle an APD could be operated in a supercritical state above breakdown and could detect single photons but it would not stay long in this state and the recovery time would be much longer than the time between consecutive generation of free carriers (dark counts).

Way out:
Subdivide the APD into many cells and connect them all in parallel via an individual limiting (quenching) resistor.

Key personalities in this development are V. Golovin and Z. Sadygov.



Hamamatsu MPPC



NIM A 504 (2003) 48

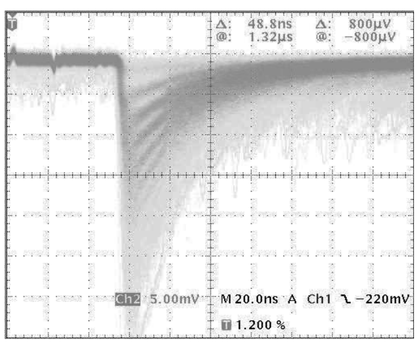
High gain

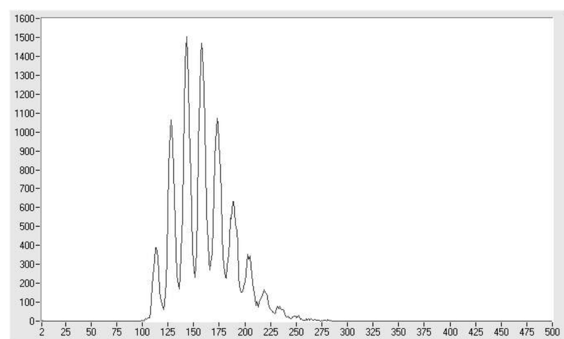
G-APDs produce a standard signal when any of the cells goes to breakdown. The amplitude A_i is proportional to the capacitance of the cell times the overvoltage.

$A_i \sim C \cdot (V - V_b)$ $(V - V_b)$ we call "overvoltage"

V is the operating bias voltage and V_b is the breakdown voltage.
When many cells are fired at the same time, the output is the sum of the standard pulses.

$A = \sum A_i$ sometimes non-linear because of crosstalk





Oscilloscope picture of the signal from a G-APD (Hamamatsu 1-53-1A-1) recorded without amplifier (a) and the corresponding pulse height spectrum (b).

High gain

G-APDs behave like PMTs and some people call them Silicon Photomultiplier, SiPM.

The gain is in the range of 10^5 to 10^7. Single photons produce a signal of several millivolts on a 50 Ohm load. No or at most a simple amplifier is needed.

Pickup noise is no more a concern (no shielding).

There is no nuclear counter effect – even a heavily ionizing particle produces a signal which is not bigger than that of a photon.

Since there are no avalanche fluctuations (as we have in APDs) the excess noise factor is small, could eventually be one.

Grooms theorem (the resolution of an assembly of a scintillator and a semiconductor photodetector is independent of the area of the detector) is no more valid.

Is the gain too high?

Assume a G-APDs with an area of 3x3 mm^2 and a cell size of 100x100 μm^2 (Hamamatsu S10362-33-100) in a high rate environment with a flux of $5x10^6$ /s·mm^2 .

The gain is $2x10^6$	
Then the current will be	16.0 μA
Add the dark count rate of 10 MHz	3.2 μA
Add the afterpulses of ~10%	1.9 μA
Add the crosstalk of ~20%	4.2 μA
Σ	25.3 μA
→ ~ 2 mW	

This is enough for a self-heating of the G-APDs which are very sensitive to temperature changes.

There is no good way to reduce the gain for a given device as we will see later.

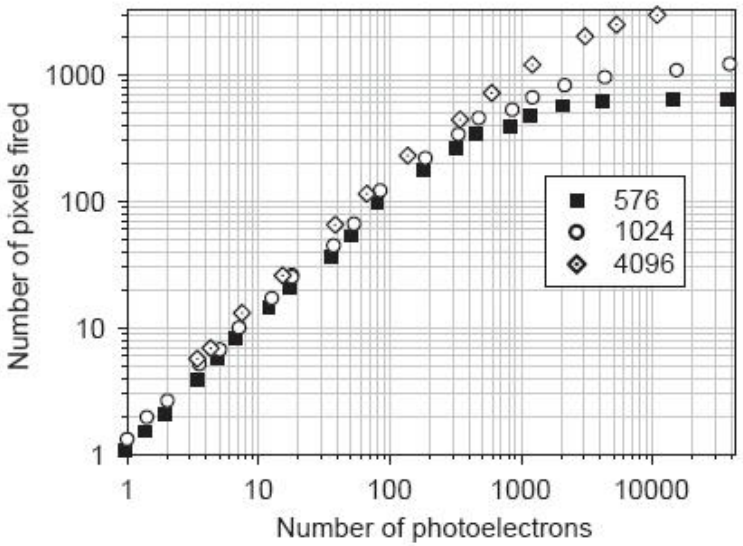
Saturation

The output signal is proportional to the number of fired cells as long as the number of photons in a pulse (N_{photon}) times the photo detection efficiency PDE is significantly smaller than the number of cells N_{total}.

$$A \approx N_{firedcells} = N_{total} \cdot (1 - e^{-\frac{N_{photon} \cdot PDE}{N_{total}}})$$

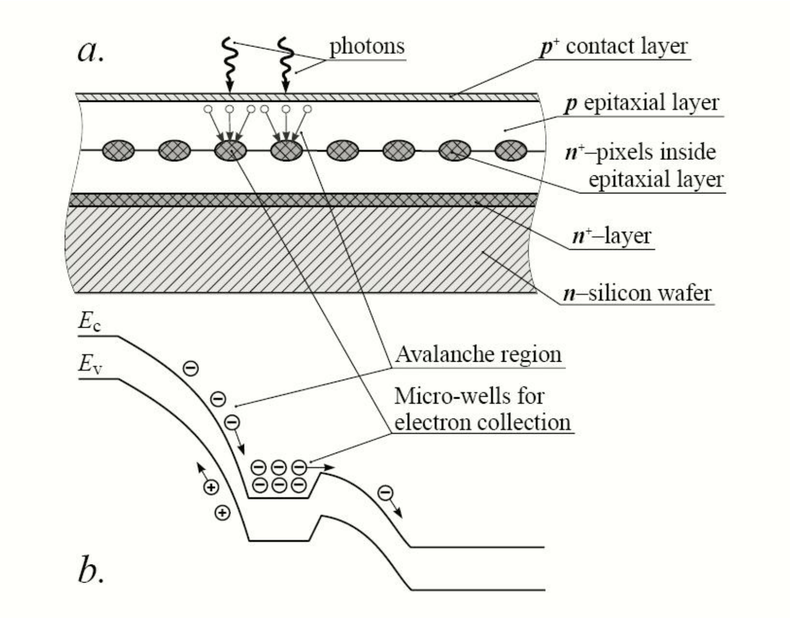
2 or more photons in 1 cell look exactly like 1 single photon.

When 50% of the cells fire the deviation from linearity is 20%.



NIMA 540 (2005) 368

Devices with almost no saturation

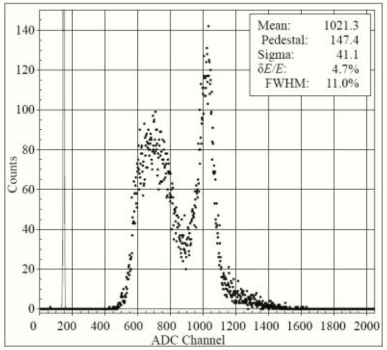


Taken from Z. Sadygov's contribution to NDIP08

Zecotek produces 3x3 mm² G-APD's based on ideas of Z. Sadygov (JINR, Dubna).

The devices have a very high density of cells: 15,000 and 40,000 cells/mm².

The gain of 10^4 to 10^5 is relative low.



^{137}Cs spectrum recorded with a MAPD-3A and LFS crystal. Both produced by Zecotek.

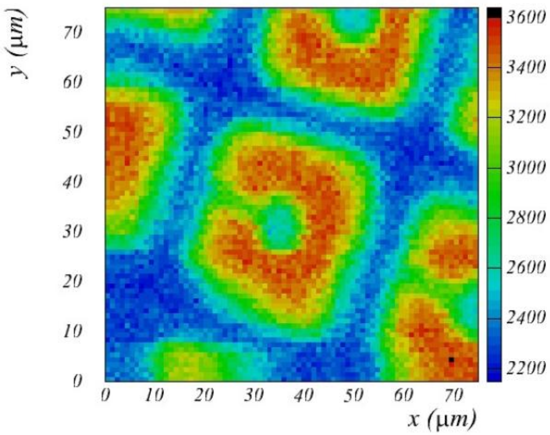
Photon Detection Efficiency (PDE)

The photon detection efficiency (PDE) is the product of quantum efficiency of the active area (QE), a geometric factor (ε, ratio of sensitiv to total area) and the probability that an incoming photon triggers a breakdown ($P_{trigger}$)

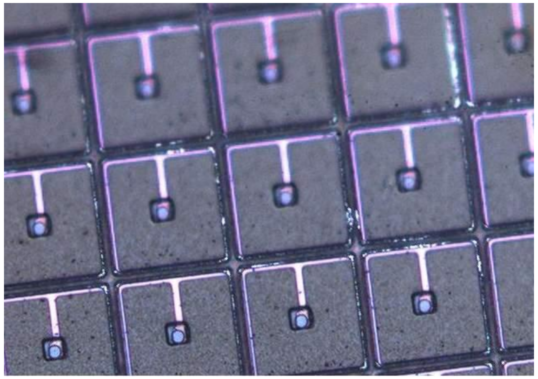
$$PDE = QE \cdot \varepsilon \cdot P_{trigger}$$

The QE is maximal 80 to 90% depending on the wavelength.

ε, the geometric factor has been optimized by most producers. There is little room left for further improvements.



P. Križan, Light07



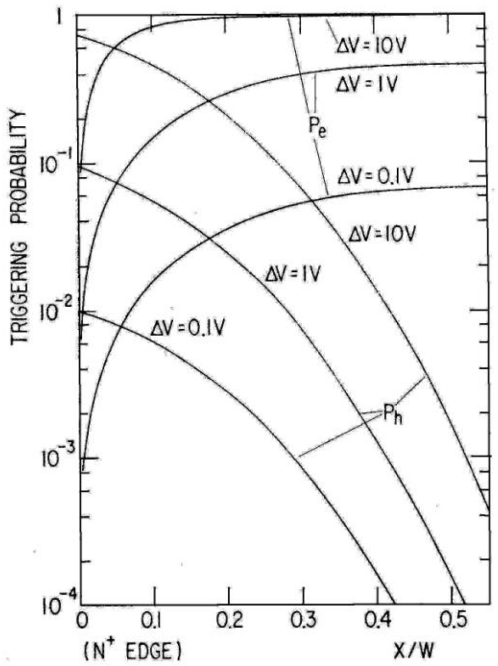
SSPM_0606BG4MM from Photonique/CPTA

PDE

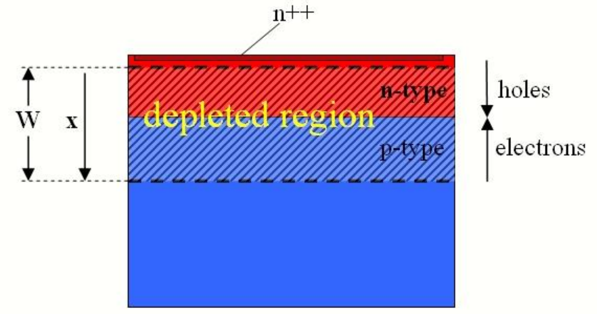
The triggering probability depends on the position where the primary electron-hole pair is generated and it depends on the overvoltage.

Electrons have in silicon a better chance to trigger a breakdown than holes. Therefore a conversion in the p+ layer has the highest probability to start a breakdown.

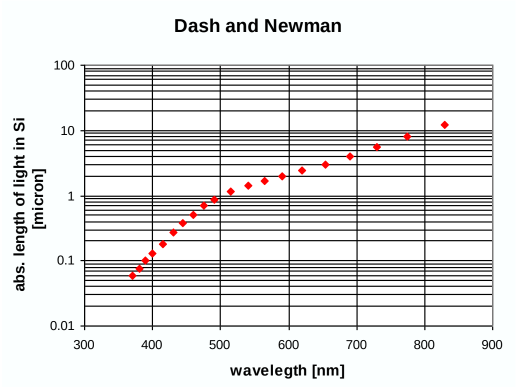
Operation at high overvoltage is favored.



W.G. Oldham et al., IEEE TED 19, No 9 (1972)



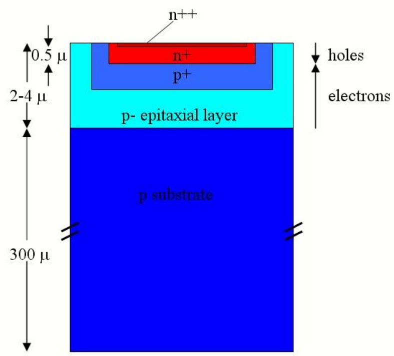
PDE

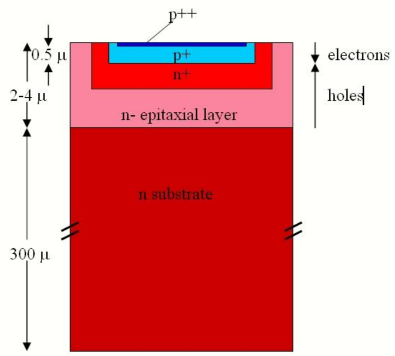


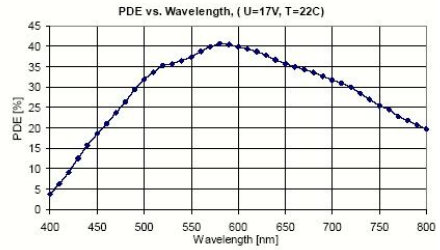
Photons with short wavelengths will be absorbed in the very first layer of Si and create there an electron-hole pair.

In a structure with a n-type substrate (right) the electrons drift towards the high field of the p-n junction and trigger with high probability a breakdown. A G-APD made on a n-type substrate will be preferential sensitive for blue light.

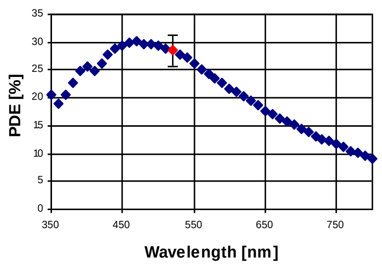
A G-APD made on a p-type substrate (left) needs long wavelengths for the creation of electrons in the p-layer behind the junction and will have the peak sensitivity in the green/red.







Photonique/CPTA
(SSPM_0710G9MM)

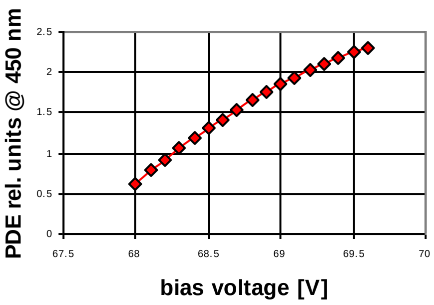


Hamamatsu
(PSI-33-050C)

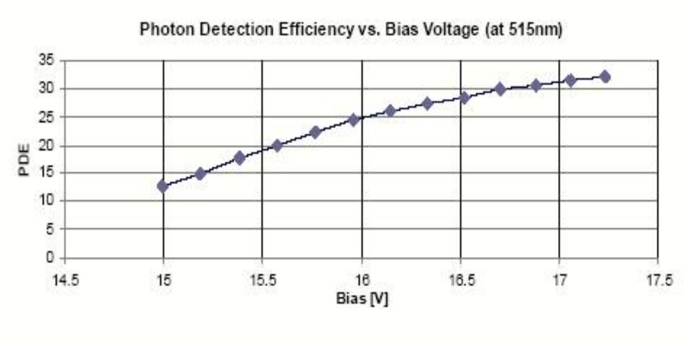
PDE

Current devices don’t reach a triggering probability of 1

Two examples:



Hamamatsu S10362-33-050
(blue sensitive)

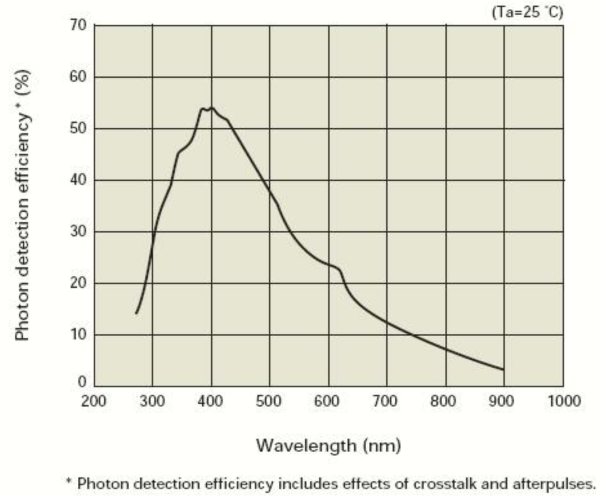


Photonique/CPTA SSPM_0710G9MM
(green sensitive)

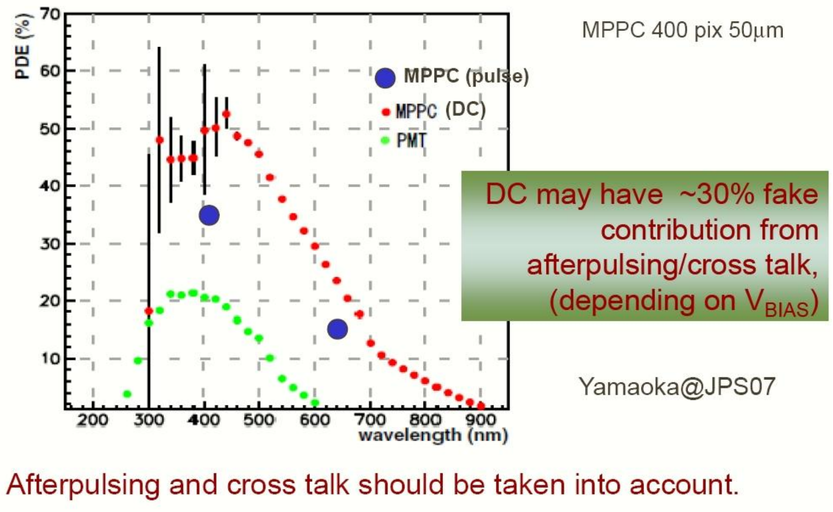
The highest possible PDE is wanted → operation at high overvoltage
A given G-APD has to work at almost the highest possible gain.

PDE

Don't take data sheets too serious



Hamamatsu data sheet for S10362-33-050



Optical crosstalk

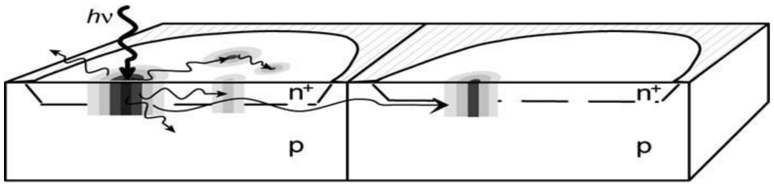
Hot-Carrier Luminescence:

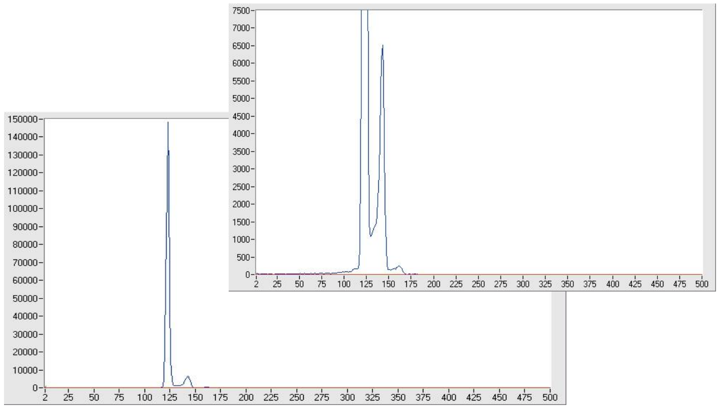
10^5 carriers in an avalanche breakdown emit in average 3 photons with an energy higher than 1.14 eV. (*A. Lacaita et al, IEEE TED (1993)*)

It turns out that only photons with wavelengths between 850 and 1100 nm contribute because shorter wavelengths will be absorbed in the original cell and longer are not absorbed at all.

When these photons travel to a neighboring cell they can trigger a breakdown there.

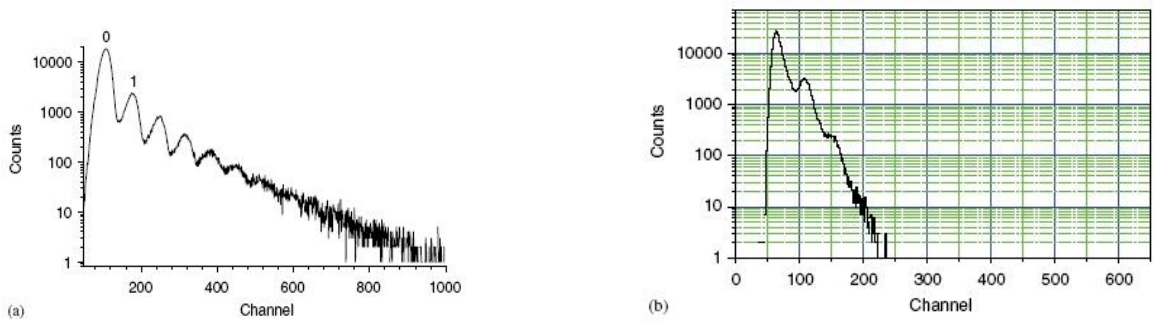
Optical crosstalk acts like avalanche fluctuations in a normal APD. It is a stochastic process. We get the excess noise factor back (F = 1 + crosstalk probability)





Hamamatsu 1-53-1A-1, cell size 70 x 70 μm

Optical Crosstalk



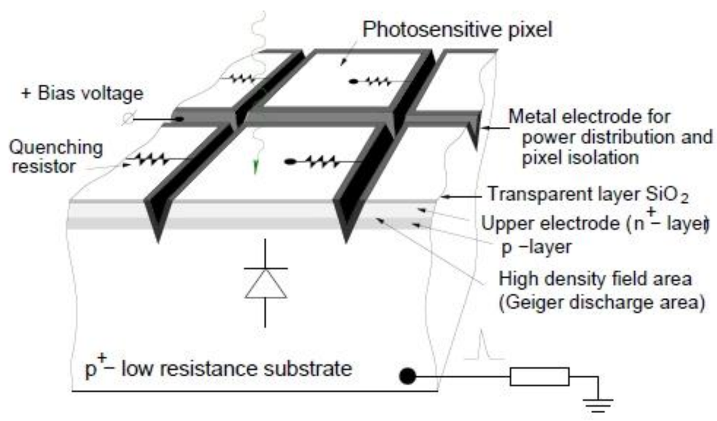
NIMA 567 (2006) 78

Optical crosstalk for a 1x1 mm^2 G-APDs produced by MEPHI/Pulsar, measured as the dark count pulse height distribution: no suppression (a); with suppression of the optical crosstalk (b) by grooves.

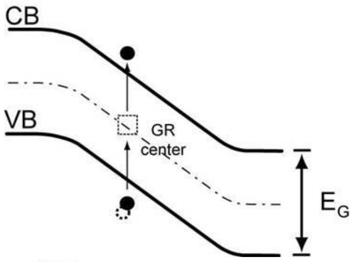
There is a concern that trenches need space and reduce the geometric factor ε.

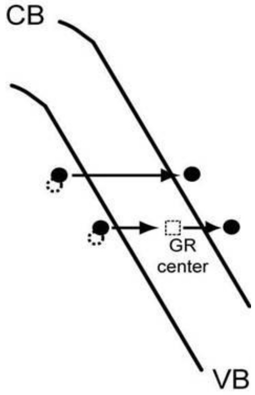
An smart way to deal successfully with this problem was found by Photonique/CPTA. They fill the trenches with Al which acts now in two ways: connect the cells to the bias voltage and isolate them optically.

Another solution is to reduce the gain



Dark count rate

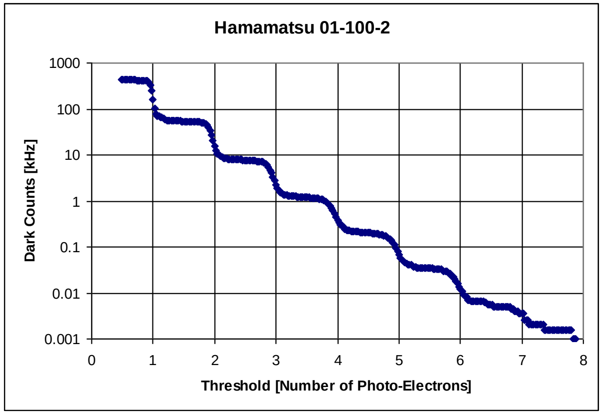




A breakdown can be triggered by an incoming photon or by any generation of free carriers. The latter produces dark counts with a rate of 100 kHz to several MHz per mm2 at 25°C when the threshold is set to half of the one photon amplitude.

Thermally generated free carriers can be reduced by cooling (factor 2 reduction of the dark counts every 8°C) and by a smaller electric field (lower gain).

Field-assisted generation (tunneling) is a relative small effect. It can only be reduced by a smaller electric field (lower gain).



The dark count rate falls rapidly with increasing threshold with steps that depend on the crosstalk probability (~12% for the G-APD shown)

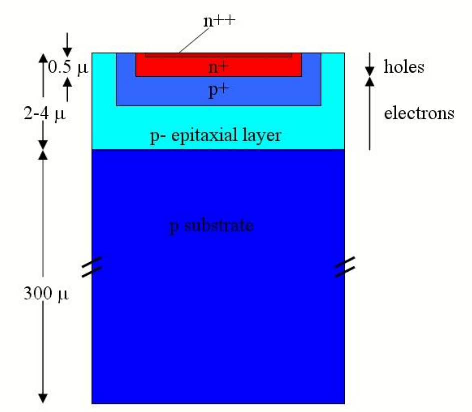
Dark count rate

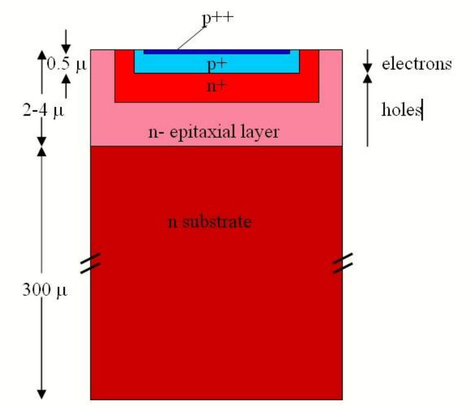
In first order the thermal generation of carriers is proportional to the depleted volume which is for every cell the area times the thickness of all the layers on top of the low ohmic substrate. In the p-type layers the electrons and in the n-type layers the holes drift towards the high field region of the junction. The electrons will trigger there a breakdown with higher probability than the holes.

In a G-APD with a n-type substrate the volume of the p-layer is much smaller than in the G-APD on a p-type substrate.

Devices with a n-type substrate should show smaller dark count rates.

Hamamatsu and Zecotek produce G-APDs on n-type substrates and their devices have significant smaller dark count rates than devices from the other producers.



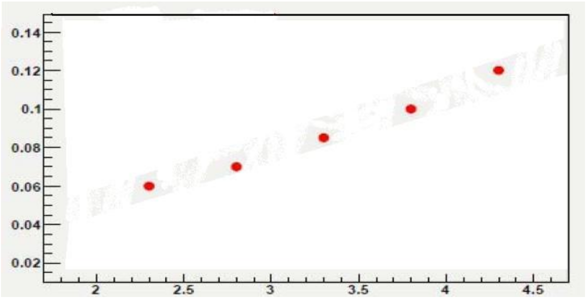


Afterpulses

In the silicon volume, where a breakdown happens, a plasma with high temperatures (few thousand degree C) is formed and deep lying traps in the silicon are filled. Carrier trapping and delayed release causes afterpulses during a period of several 100 nanoseconds after a breakdown.

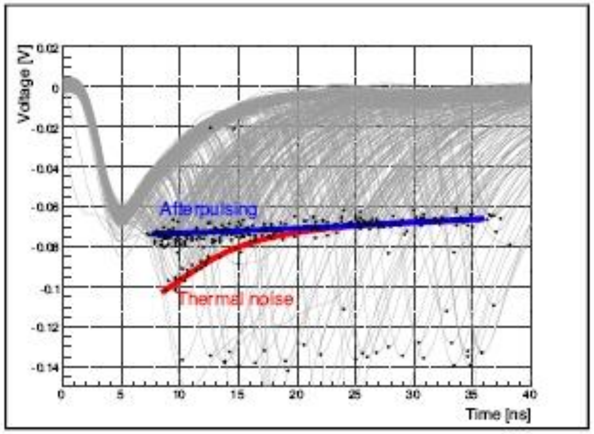
The probability for afterpuses increases with higher overvoltage (higher gain).

Afterpulse probability [%]

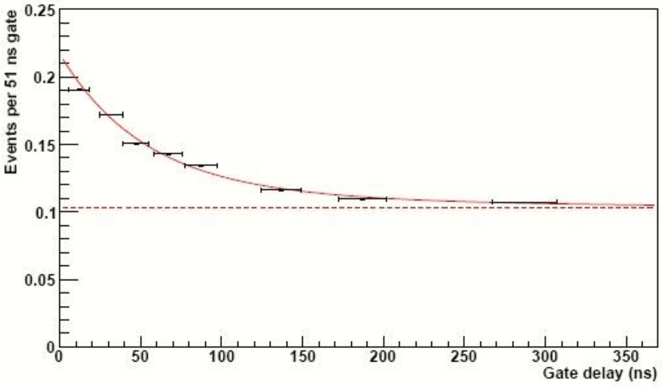


Overvoltage [V]

H. Otono, PD07



H. Oide, PoS (PD07) 008



Hamamatsu S10362-33-050C: afterpulses in a delayed gate. The total probability is ~20%.

The dashed line indicates normal dark counts.

Recovery time

The recovery time is related to the $R_Q C_C$ time constant (quenching resistor and cell capacitance).

The signal decay should have the same time constant.

Type Hamamatsu 1x1 mm²	1600 cells	400 cells	100 cells
Overvoltage [V]	3.3	2.7	0.87
Recovery time [ns]	~ 4	~ 9	~ 33
Pulse decay time [ns]	~ 5	~ 9	~ 35

H. Oide, PoS (PD07) 008

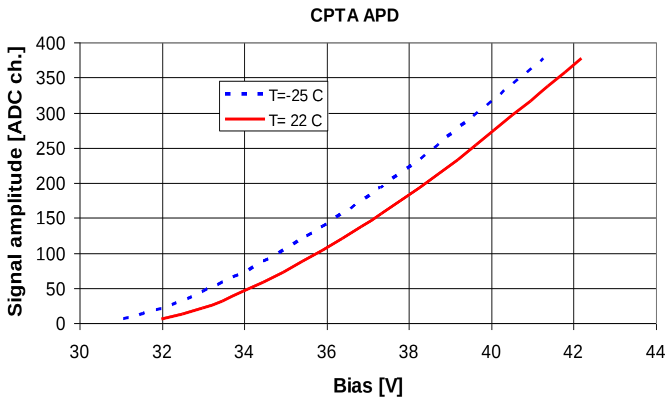
Afterpulses can prolong the recovery time because the recharging starts anew.

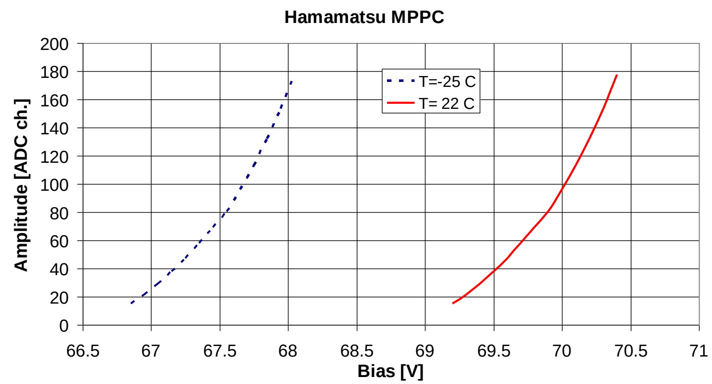
Polysilicon resistors are used up to now which change their value with the temperature. Therefore there is a strong dependence of the recovery time on the temperature.

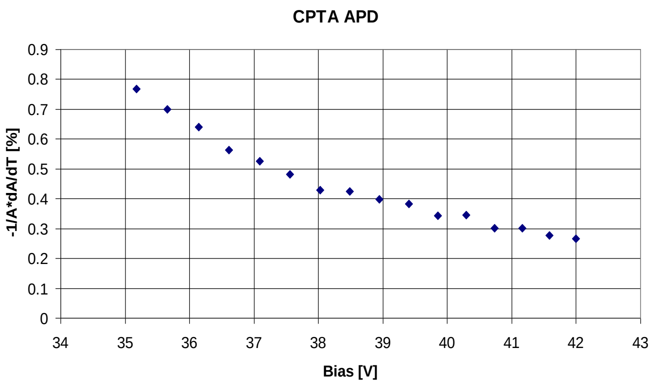
Go to a metal alloy with high resistivity like FeCr. Will be done by Hamamatsu.

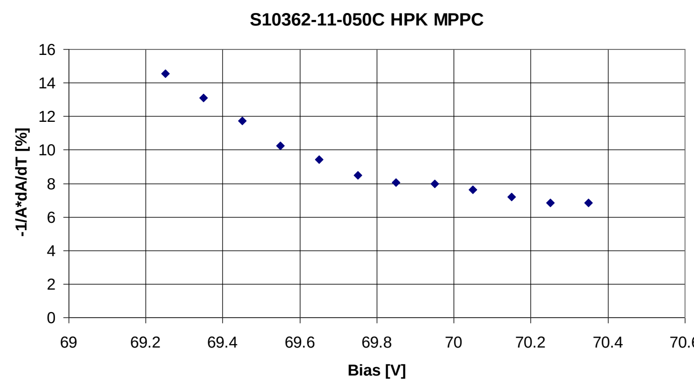
Temperature dependence

$k_T = dA/dT * 1/A$ [%/°C]









Y. Musienko, PD07

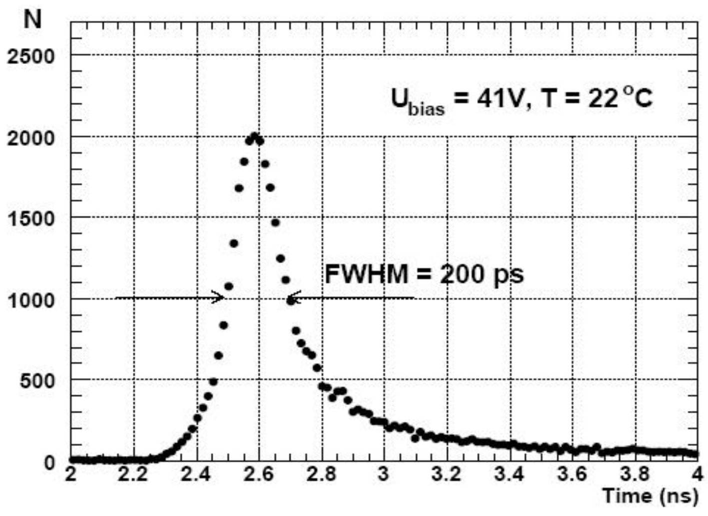
Timing

The active layers of silicon are very thin (few μm), the avalanche breakdown process is fast and the signal amplitude is big. We can therefore expect very good timing properties even for single photons.

Fluctuations in the avalanche are mainly due to a lateral spreading by diffusion and by the photons emitted in the avalanche.

A. Lacaita et al., Apl. Phys. Letters 62 (1992) A. Lacaita et al., Apl. Phys. Letters 57 (1990)

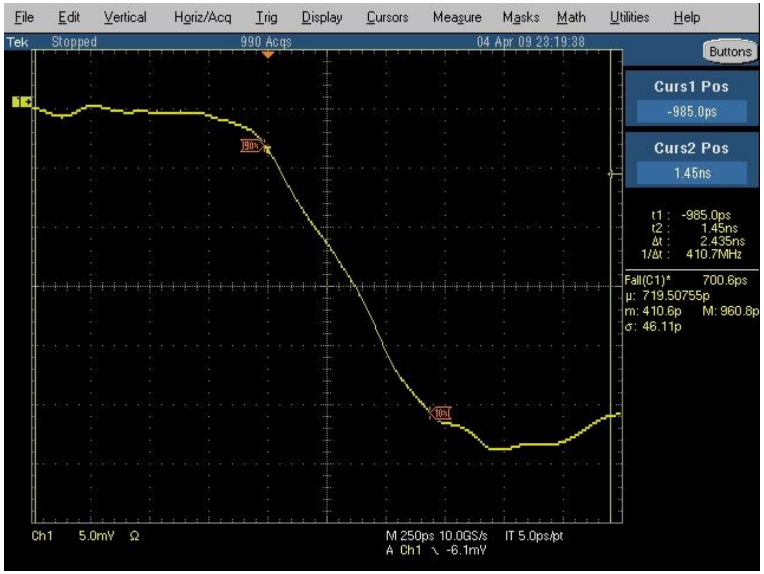
High overvoltage (high gain) improves the time resolution.

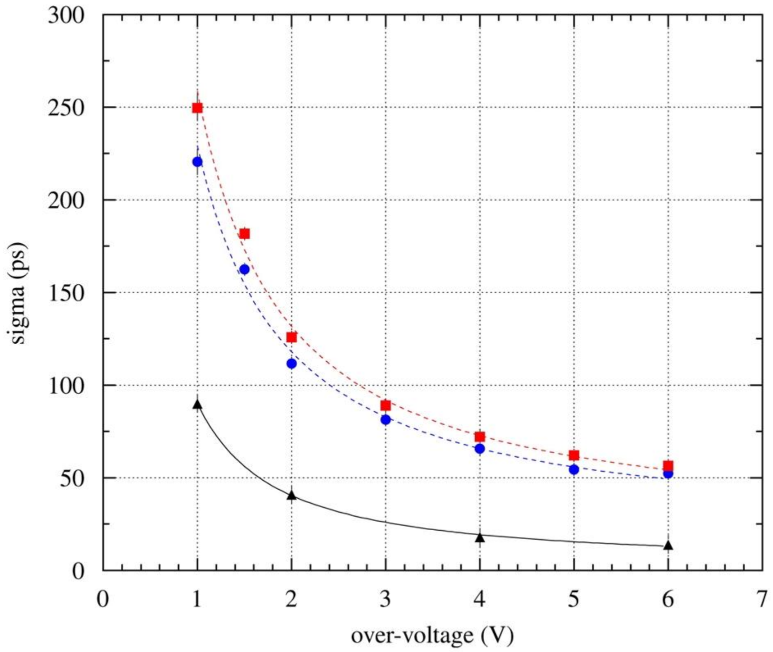


Contribution from the laser is 37 ps FWHM
G-APD from CPTA/Photonique

physics/0606037

Timing





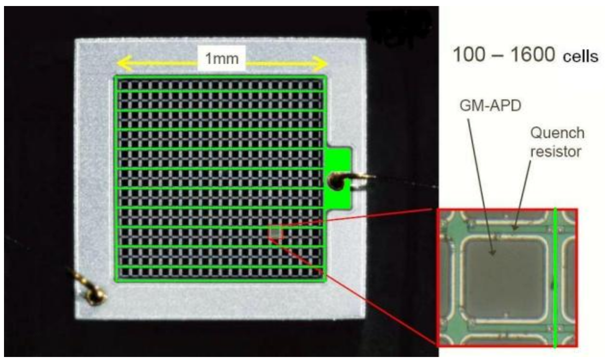
Zecotek 3x3 mm^2 MAPD response to a 35 psec FWHM laser pulse (λ=635 nm)

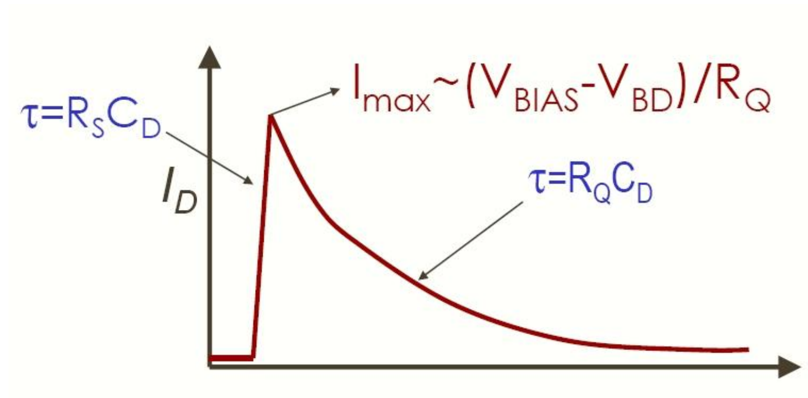
NIMA 581 (2007) 461

~700 psec rise time was measured (limited by circuitry)

Y. Mousienko (INR/Boston N.E.U.)

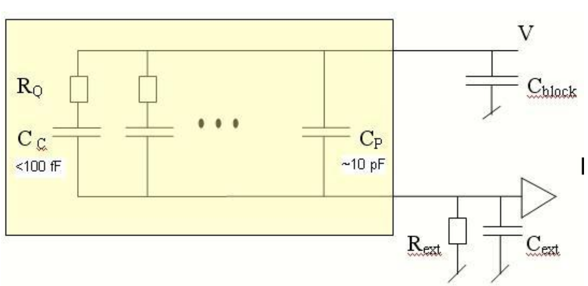
Pulse shape

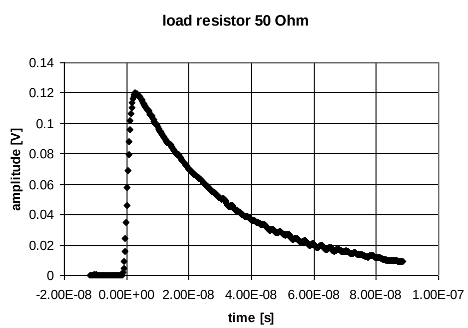


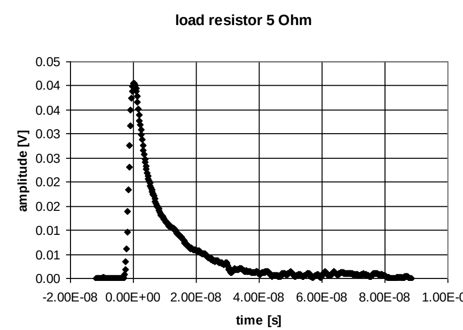


The usual interpretation of the pulse shape

Simple equivalent circuit:



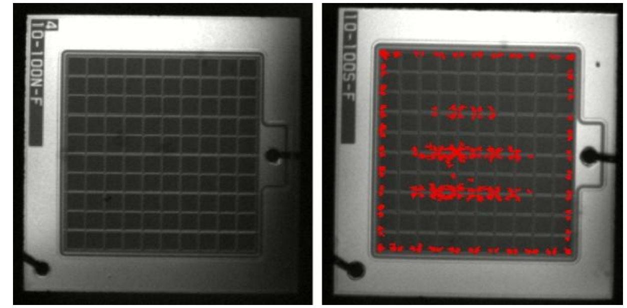




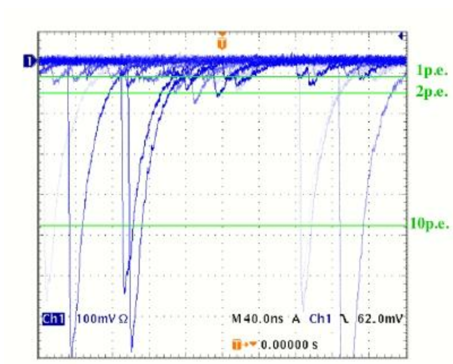
Radiation hardness

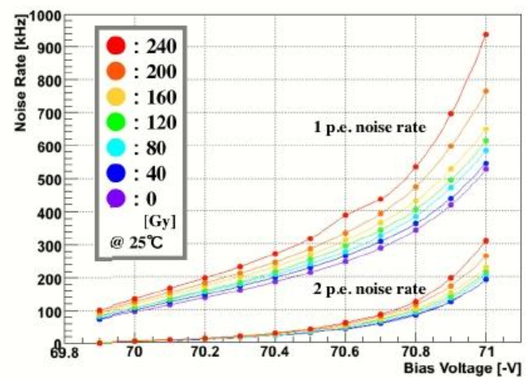
γ-rays have little effect on the bulk of silicon but they can break the SiO_2 bonds. Charges are accumulated at the SiO_2-Si interface and change the field distribution which can generate currents on the surface.

Very large dark count pulses have been observed after irradiation with 240 Gy from a ^{60}Co source.



Infrared light emitted before and after irradiation seen with an infrared sensitive microscope.





All figures from Matsubara, Pos (PD07) 032

Radiation hardness

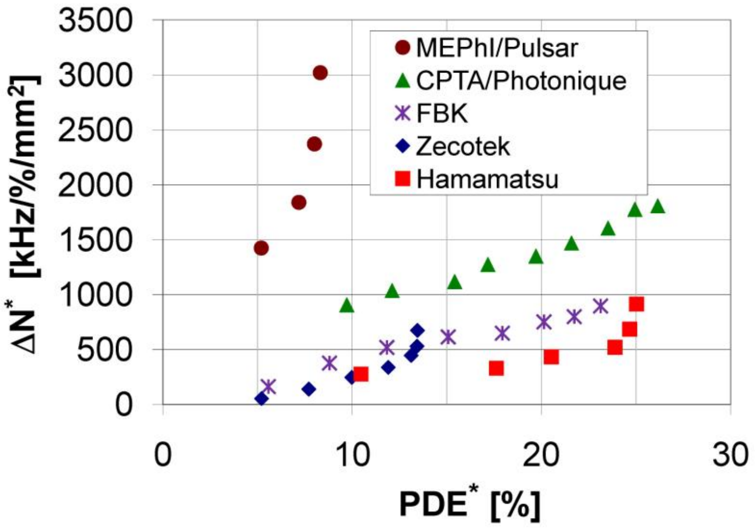
Massive particles (electrons, protons, neutrons etc.) can displace Si atoms from their lattice position and cause defects which act as generation/recombination centers for free carriers.

The defects are created in the whole volume. Therefore the argument used in the discussion of the dark count rate holds as well:

G-APDs made on a n-type substrate will show the smallest increase of the dark counts.

No change, which is not compatible with the measurement errors, has been reported for the PDE, the breakdown voltage, the gain, the crosstalk probability the recovery time constant and the value of the quenching resistors.

The rate of afterpulses is increased – more traps



Increase of the dark count rate after irradiation with 10^{10} protons/cm^2 (82 MeV) as function of PDE* which is the measured PDE corrected for the crosstalk.

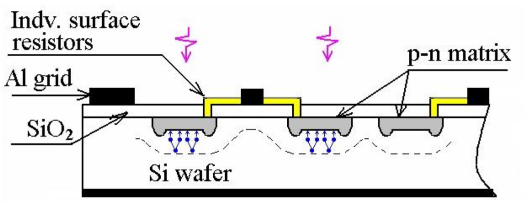
The fluence is equivalent to a fluence of $2\cdot10^{10}$ neutrons/cm^2 (NIEL factor difference ~2)

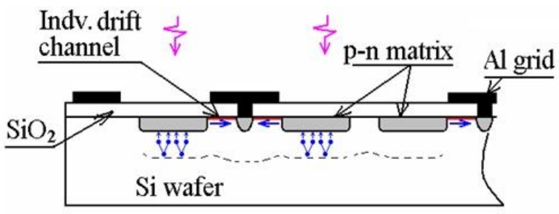
More Properties

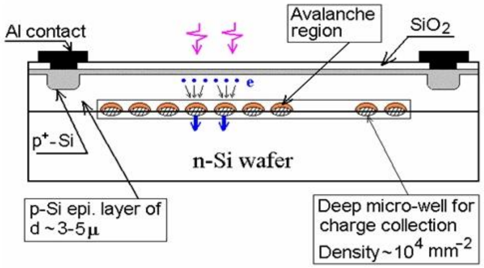
There are more features which are not mentioned yet:

- G-APDs work at low bias voltage (~50 V),
- have low power consumption (< 50 μW/mm^2),
- are insensitive to magnetic fields (tested up to 7 T),
- are compact, rugged and show no aging,
- tolerate accidental illumination,
- could be cheap because they are produced in a standard MOS process

Types of G-APDs







- CPTA/Photonique (Moscow/Geneva)
- MEPhI/Pulsar (Moscow, Russia)
- Amplification Technologies (Orlando, USA)
- Hamamatsu Photonics (Hamamatsu, Japan)
- SensL (Cork, Ireland)
- RMD (Boston, USA)
- MPI Semiconductor Lab. (Munich, Germany)
- FBK-irst (Trento, Italy)
- STMicroelectronics (Catania, Italy)
- ……

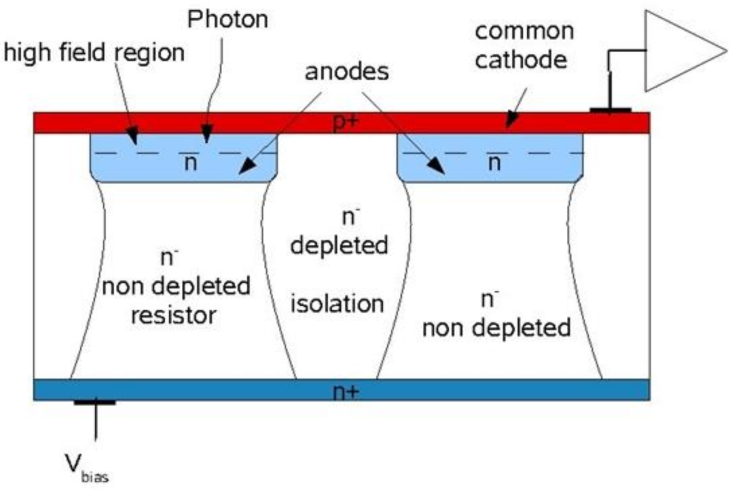
- Z. Sadygov (JINR, Dubna, Russia)

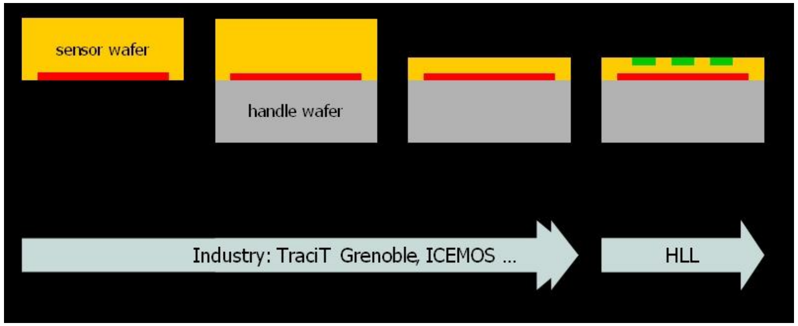
- Zecotek (Singapore)

Every producer uses its own name for this type of device: MRS APD, MAPD, SiPM, SSPM, SPM, DAPD, PPD,

SiMPI: G-APD with bulk integrated quench resistors from MPI-HLL

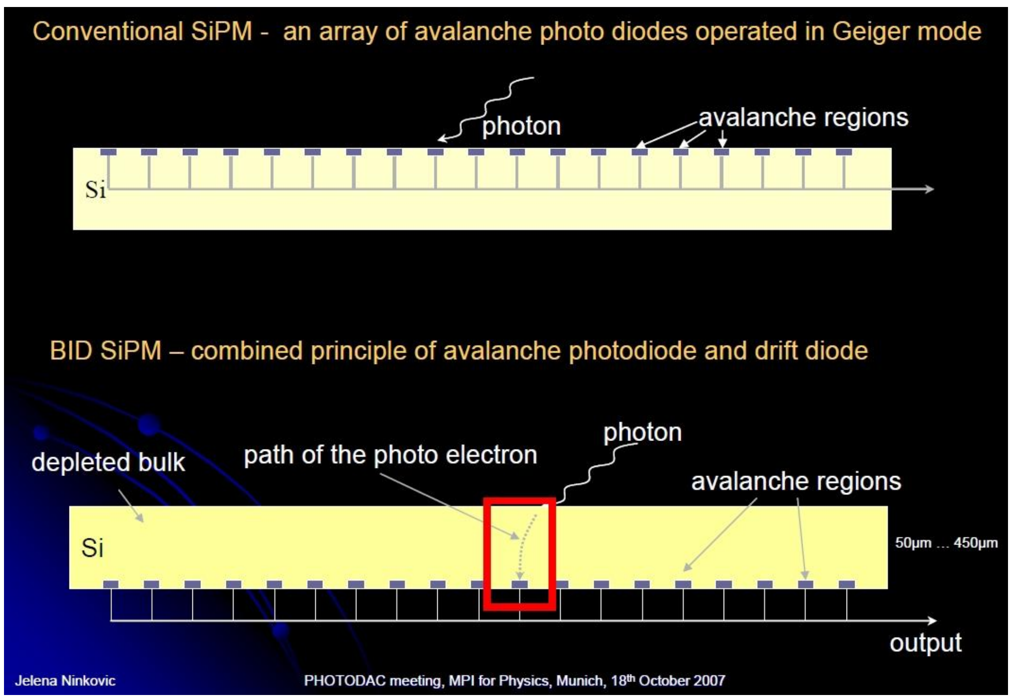
- geometric factor ~75 % - no resistor on the surface → PDE ~ 60 %
- free entrance window for light, no metal necessary within the array
- allows engineering of the entrance window
- improved radiation hardness – no lateral high field regions on the surface

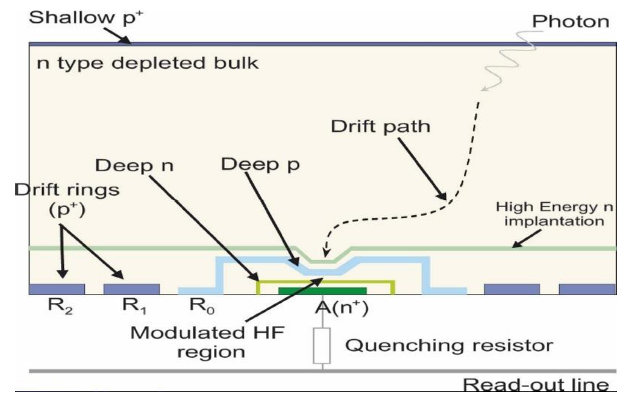




BID-SiPM: Back illuminated G-APD from MPI-HLL

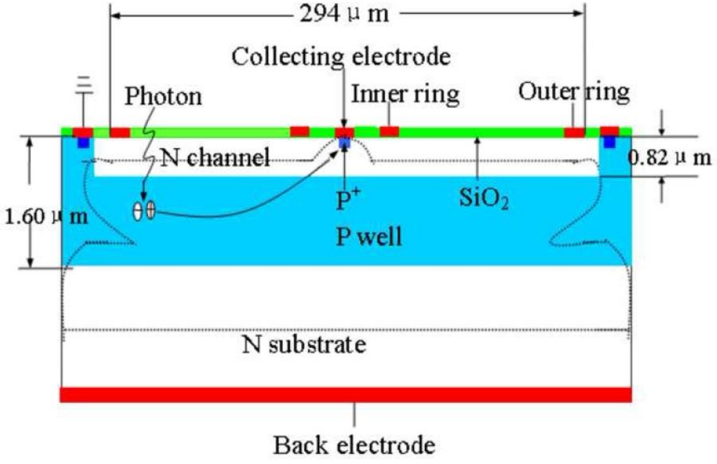
- combination of drift diode and G-APD – geometric factor 100% → QE > 80%
- direct coupling of entrance window to scintillator (no wire bonds)
- bump bonding to a readout chip possible – signal detection for each cell

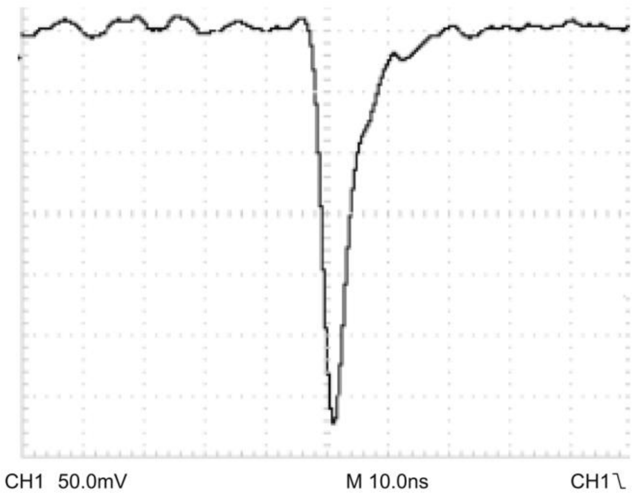




Front illuminated G-APD from Beijing

- combination of drift diode and G-APD
- small capacitance of the amplification structure
- simple CMOS process





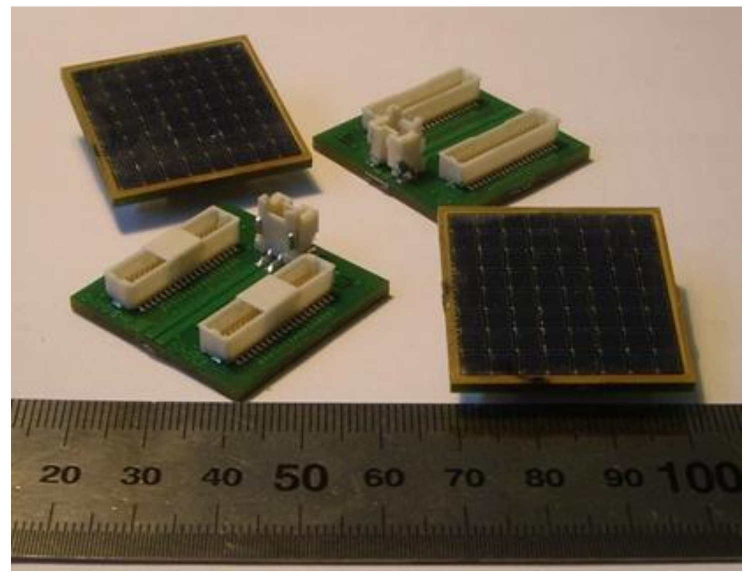
G. Wu et al., Demonstration of an avalanche drift detector with front illumination, NIMA

Price

For the time being the price is prohibitive – some 50 € for a 3x3 mm^2 device.

This is a factor of 30 higher than the same sensitive area of a PMT and a factor of 5 higher than a normal, linear APD (which is more difficult to produce)

8 inch wafer produced by Zecotek



8x8 array of 3 mm^2 G-APDs from Zecotek

Where do we stand?

Significant progress in development of G-APDs over the last 2-3 years:

- High PDE of 30-40% for blue-green light (CPTA/Photonique, Hamamatsu, Zecotek)
- Reduction of dark count at room temperature to 300 kHz/mm^2 (Hamamastu, Zecotek)
- Low cross-talk <1-3% (CPTA/Photonique, STMicroelectronics)
- Low temperature coefficient of ~0.3%/C (CPTA/Photonique)
- Fast timing ~50 ps (RMS) for single photons (all)
- Large dynamic range with 15 000 – 40 000 pixels/mm^2 (Zecotek)
- Large area of 3x3 mm^2 (CPTA/Photonique, Hamamatsu, FBK, SensL, Zecotek, STMicroelectronics…)

Future

The development of G-APDs is ongoing. What can we expect in a couple of years?

- PDE >50-60% for 350-650 nm light
- dark count rate <100 kHz/mm^2 at room temperature
- optical crosstalk <1%
- active area >100mm^2
- high DUV light sensitivity (PDE(128 nm ~20-40%)
- G-APD arrays:6x6, 8x8 ...
- very fast CCDs operated in Geiger mode□
- radiation hard G-APDs -up to $10^{14} \div 10^{15}$ n/cm^2 (new materials: diamond?, GaAs?, SiC?, GaN? …)
- production cost $<1\$/\text{mm}^2$

Conclusion

When single photons need to be detected with very short resolving times, a domain where up to now only photomultiplier tubes could be used, the Geiger-mode APD's promise to perform very well. Their high PDE makes them in some applications even superior to PMT's. In addition they could allow the construction of cost effective detectors because they potentially are cheap due to the standard CMOS production process and because they need only simple electronic circuits, no shielding, little space and have low power consumption.

Geiger-mode APDs are grown out of their infancy and found applications in particle physics, astrophysics, medical imaging and other fields. There is still room for improvements and furthermore the devices can be tailored to user-specific requirements.

Production on a p- or n-doped substrate

G-APDs are produced in the following way: First a thin layer (2 to 4 µm) is grown epitaxial on a wafer (typically 300 µm thick) with low resistivity which is needed for the handling but has no other function. The first step towards a p-n junction is the diffusion of dopants with high concentration into the epitaxial layer. So far all layers, the wafer (substrate), the epitaxial layer and the heavily doped volume created by diffusion, are of the same type of silicon, n-type or alternatively p-type. The junction is then formed some 0.5 µm below the surface by a shallow diffusion (or ion implantation) with the opposite dopant. In a final step an extremely thin but heavily doped layer is created on the surface which distributes the potential uniformly over the whole area of the diode. Further production details like the passivation of the surface with SiO_2, the creation of polysilicon quenching resistors and the aluminization are omitted here.