Silicon Photomultiplier – Concepts, Characteristics, Prospects

Christian Jendrysik*
Max Planck Institute for Physics
E-mail: jendrysik@hll.mpg.de

Silicon Photomultipliers are cheap and efficient photon detectors with the capability of single photon counting. Therefore, they become an attractive alternative for the widely used vacuum photomultiplier tubes. Over the last few years, many different approaches were presented and the technological development led to significant improvements in the device characteristics. In this paper the basic concepts of the detectors are briefly presented, followed by a review of the characteristics, advantages and drawbacks as well as future prospects of those promising devices.

22nd International Workshop on Vertex Detectors
September 15th-20th, 2013
Lake Starnberg, Germany

*Speaker.
1. Introduction

In many applications conventional photomultiplier tubes are the most common low light level detectors up to now. Within the last few years Silicon Photomultipliers (SiPMs) attracted a lot of attention. SiPMs are arrays of small independent avalanche photodiodes operated a few volts above their breakdown voltage \( V_{bd} \) (Geiger-mode, G-APD), which is typically in the order of 20 V to 100 V. If a photon is absorbed in the sensitive volume, an avalanche breakdown can occur (Geiger discharge) resulting in a high intrinsic amplification of the initial single electron-hole pair. In addition, the Geiger-mode operation of a SiPM provides fast timing properties in the sub-ns range. Moreover, those detectors are robust and compact devices, capable of single photon detection and insensitive to magnetic fields.

In conventional SiPMs the quenching of the avalanche breakdown, which is required in order to be sensitive to subsequent photons, is realized by a resistor \( R_Q \) connected in series to each cell (passive quenching). An avalanche breakdown results in a voltage drop at the resistor which reduces the applied voltage at the diode to \( V_{bd} \). If the remaining current through the diode is sufficiently low, there is a certain probability that no new charge carriers are generated by impact ionization and the avalanche is quenched [1, 2]. Afterwards, the cell needs a certain time to recover to the externally applied bias voltage \( V_{bias} \) (recovery time). The cells of the array are all connected in parallel. This results in a signal which is the sum of all simultaneously fired cells.

In addition to this conventional concept, various SiPM approaches have been developed in order to improve the device characteristics (e.g. MAPDs from Zecotek [3] with micro-well structures or SiPMs with bulk-integrated quench resistors [4, 5], which will be discussed below). The properties and prospects of SiPMs, with a focus on non-CMOS based approaches, will be briefly reviewed in this paper.

2. SiPM Characteristics

2.1 Gain and Dynamic Range

The gain \( G \) of a SiPM is defined as the number of charges generated in a Geiger discharge and is proportional to the cell capacitance times the applied overbias voltage \( V_{ob} \):

\[
G = C_{cell} \cdot V_{ob} = (C_D + C_Q) \cdot (V_{bias} - V_{bd}),
\]

where \( C_D \) is the capacitance of the G-APD and \( C_Q \) is the coupling capacitance in parallel to the quench resistor. Eq. 2.1 already shows that the gain is linearly increasing with the applied overbias voltage. Depending on the cell size values of \( G \) can be in the order of \( 10^5 \) – \( 10^7 \). Thus, a signal from a single photon absorption produces a signal of several millivolt on a 50 \( \Omega \) load resistor which simplifies the requirements on the readout electronics. The fluctuations of the Geiger discharge are small which improves the excess noise factor in comparison to avalanche photodiodes in proportional mode (\( V_{bias} \) slightly below breakdown) [6].

Due to the parallel readout of all cells, and the fact that the signal for one or more photons impinging on one cell is the same (quasi digital output), the signal of the SiPM is proportional to the number of incident photons \( N_{ph} \) as long as the number per pulse is significantly smaller than...
the total number of cells $N_{total}$ (see Fig. 1)

$$N_{fired} = N_{total} \left[ 1 - \exp \left( -\frac{N_{ph}}{N_{total}} \right) \right]. \tag{2.2}$$

This behavior is due to an increasing probability for two or more photons hitting the same cell of the array with increasing photon numbers.

In order to increase the dynamic range of a SiPM, a non-conventional SiPM design is provided by MAPDs from Zecotek [3]. A double n-p-n-p junction with micro-well structures is located at a depth of 2–3 µm below the surface, with the multiplication regions just in front of these wells. This technique allows a quenching of the avalanche without an additional quench resistor by specially designed potential barriers. In this way a pixel number of $1.35 \times 10^5$ on an area up to $3 \times 3$ mm$^2$ is possible, which significantly improves the SiPM dynamic range.

### 2.2 Dark Counts and Afterpulsing

Thermal generation of electron-hole pairs in the depletion region can also trigger the Geiger discharge of an avalanche diode – known as dark counts. The resulting SiPM signal cannot be distinguished from a photon induced pulse. The dark count rate (DCR) is a function of the temperature and approximately decreases by a factor of two for a temperature decrease of 7 degrees. It also depends on the overbias voltage, since the probability of triggering a breakdown increases with $V_{ob}$ [8]. In addition, the DCR is strongly influenced by the quality of the wafer material. Improvements of this property require a minimization of the defect state density in the avalanche region, which is a technological challenge regarding the optimization of wafer production, implantation and annealing processes.
During an avalanche breakdown, charge carriers also can be trapped by defect states within the depletion region. The trapped charge is released after a characteristic time and can cause a delayed second avalanche breakdown event faking a photon signal, and thus degrading the resolution of the SiPM. This effect is called afterpulsing. The contribution of afterpulsing depends on the average trapping time $\tau_i$ (a fast and a slow contribution is reported in Ref. [9]) and the recovery time of the SiPM cell. Increasing the overbias voltage results in a non-linear increase of the afterpulsing probability due to the voltage dependence of the gain and the trigger probability.

In its newest series of MPPCs (S12571) Hamamatsu claims a reduced afterpulsing probability (few percent) compared to previously marketed products and dark rates in the order of 100 kHz/mm$^2$ [10]. The same orders of magnitude in the DCR are achieved by e.g. Excelitas [11], and KETEK [12].

2.3 Optical Cross Talk

In an avalanche breakdown, in average one photon with an energy of $E \geq 1.12$ eV (band gap energy of silicon) is generated per $3 \cdot 10^5$ charge carriers crossing the diode junction [13]. Due to the high multiplication of SiPMs in the order of $10^6$ approximately 30 photons are emitted per Geiger discharge, which can propagate in the silicon and initiate an additional breakdown event in neighboring cells (optical cross talk, OCT). Such an event cannot be distinguished from a second, simultaneously incident photon.

Since the gain, and therefore the number of charge carriers, as well as the probability to trigger an avalanche are functions of the overbias voltage [8], the optical cross talk also shows a dependency on $V_{ob}$. This is illustrated in Fig. 2. The voltage dependence of the OCT is one limiting factor of the maximum applicable overbias, since the dynamic range of the device is decreased by

![Figure 2](image-url): Optical cross talk of SiPM with bulk-integrated quench resistor (pitch 130 µm) as a function of the applied overbias voltage. An increase in the gap results in a lower fill factor and gain and thus a reduced OCT. Image from [14].
the increased occupancy of cells due to cross talk. Furthermore, fluctuations in the number of the cross talk events degrades the resolution of the SiPM.

Without OCT suppression, the probability of cross talk is usually in the range of 20% to 40% at the recommended operation voltages. For a suppression of OCT, optical trenches between adjacent cells filled with opaque material can be implemented [15]. The suppression is even more improved if an additional buried pn-junction is added in the bulk material. In this case, minority charge carriers from the bulk, generated by cross talk photons, are prevented from diffusing into the avalanche region (slow OCT contribution). Devices with optical trenches showed improvements in the OCT probability of one order of magnitude (e.g. Ref. [12]). SiPMs with bulk-integrated quench resistors benefit from an inherent diffusion barrier for minorities from the bulk (high field implant) [4]. It suppresses the slow contribution of OCT and makes a second pn-junction obsolete. The remaining cross talk can be attributed to reflections in the device. Their avoidance requires further improvements of technology and design.

2.4 Recovery Time and Timing Properties

During an avalanche breakdown the bias voltage at the diode drops to the breakdown voltage. Once the avalanche is quenched, the cell recovers to \( V_{\text{bias}} \) with a characteristic time constant \( \tau_{\text{rec}} \). For passively quenched SiPMs, the recovery time is in first approximation defined by the product of the quench resistance \( R_Q \) and the diode capacitance \( C_D \), since the equivalent circuit of the SiPM during recovery is dominated by this RC element [16]. Values of \( R_Q \) are in the order of a few hundreds of k\( \Omega \) and the diode capacitances scale with the pixel size in the order of a few hundred fF. As a result, \( \tau_{\text{rec}} \) can be in the range of a few 10 ns to \( \mu s \). In order to minimize the temperature dependence of the resistance, Hamamatsu started to implement high resistive metal alloys instead of polysilicon resistors [17].

Due to the avalanche breakdown, which is a fast process with a large signal amplitude, the SiPM offers good timing properties even at the single photon level. With a rise time in the ns and sub-ns regime for conventional devices, time resolutions below 200 ps FWHM were reported (see Fig. 3) for single photon detection [18, 19, 20]. With respect to the dependence on the number of photoelectrons \( N_p,e \), a \( 1/\sqrt{N_p,e} \) behavior of the time resolution was observed [18].

An interesting approach for improvements of SiPM timing properties was developed by SensL [21]. For achieving fast output pulses with a steep rise time, they introduced an additional derivatively coupled electrode to each cell (fast terminal readout). With their current B Series Sensors, they claim rise times around 300 ps and pulse widths below one ns. This offers improved multi photon resolution and coincidence resolving times [22].

2.5 Photon Detection Efficiency

For a silicon photomultiplier the photon detection efficiency (PDE) is a function of the wavelength and the overbias voltage and is determined by:

\[
PDE(\lambda, V_{\text{ob}}) = QE(\lambda) \cdot FF 
\cdot \varepsilon_G(V_{\text{ob}}, \lambda),
\]

with the internal quantum efficiency \( QE \), the geometrical fill factor \( FF \) and the trigger or Geiger efficiency \( \varepsilon_G \). The internal quantum efficiency is proportional to the absorption coefficient and
Figure 3: Single photon time resolution of the three different MPPCs as a function of the overbias voltage. The 100 \( \mu \text{m} \) measurement is limited because of the rapid increase in dark counts at higher voltages. All SiPM signals were readout with the discriminator amplifier NINO. A resolution of \( \sigma \approx 80 \) ps is obtained. Image from [18].

therefore dependent on the wavelength of the incident photon. Especially for detection of UV and blue light with very short penetration depths, the fabrication of shallow top layers is a technological challenge in order to avoid electron-hole recombination in the non-depleted region close to the surface. In addition, the surface reflection of light due to the differences in the refraction indices is a further limiting factor. By engineering entrance windows with optimized antireflective coatings, the reflection losses can be reduced below 10%.

The fill factor is the ratio of the photon-sensitive area of the SiPM and the whole area of the array. Usually, the dead area of the device is needed for the integration of the quenching mechanism and contacts as well as the suppression of optical cross talk. Therefore, the \( FF \) value generally depends inversely on the cell size of the array. Furthermore, inhomogeneous doping processes can affect the effective geometrical fill factor of the SiPM, as shown in Ref. [23]. Latest values for fill factors are in the range of 70–80% for a cell size of around 100 \( \mu \text{m} \) \([10, 12]\).

The trigger efficiency defines the probability, that the generated electron-hole pair triggers an avalanche breakdown while drifting through the high field region of the detector. This parameter is a function of the electric field strength distribution in the depletion region and thus of the applied overbias voltage and the exact doping profile \([8]\). Furthermore, \( \varepsilon_G \) depends on the position of the photon absorption and therefore on the wavelength as well as on the type of the charge carrier that triggers the avalanche. In silicon, electrons have a higher trigger probability compared to holes due to their larger ionization coefficients, what makes them the preferred charge carriers to initiate an avalanche breakdown. This fact determines the basic structure of the sensor. Devices with a PDE peak wavelength in the blue and sensitivity in the UV regime are built with a p-doped implant on a n-type substrate. In this case, the photons are most probably absorbed close to the surface.
Silicon Photomultiplier – Concepts, Characteristics, Prospects

Christian Jendrysik

Figure 4: Photon detection efficiency for a SiPM from KETEK with a pitch size of 50 µm. The peak sensitivity is around 400 nm and reaches values of larger than 50%. Image adapted from [24].

(absolute depth 100–500 nm) and the electrons drift through the whole high field region of the device. For the detection of longer wavelengths, the structure is inverted, so that again the covered distance by electrons in the high field region is maximized.

Currently the best PDE of commercially available devices is obtained by KETEK (pitch: 50 µm) with values above 50% for a wavelength around 400 nm [24] (cf. Fig. 4). Hamamatsu’s MPPCs and SiPMs from Excelitas are in the range of 30–40% [11, 10]. The achievement of high fill factors, which is crucial for a high PDE, and a simplification of the production process are the main goals of the development of SiPMs with bulk-integrated quench resistors [4, 5]. In this approach, a high-ohmic silicon bulk material is utilized for the realization of the required quench resistance. An additional advantage of this concept is the free entrance window for light which allows an easy connection to scintillator crystals or bump-bonding techniques. However, since the resistor cross-section is related to the cell size, this approach requires thin wafer material adjusted to the cell size, in order to realize the resistor matching. As a result of the bulk integration of $R_Q$, a transistor-like behavior with a non-linear IV-curve is obtained which results in a longer recovery time compared to conventional series resistors. The SiMPi prototypes achieved PDE values in the range of 35% with a pitch size of 100 µm [25]. However, those devices have no optimized entrance windows yet and were limited in the maximum feasible overbias voltage by the high dark count rate and the quenching condition for the avalanche.

3. Active Quenching

An alternative method to stop the avalanche breakdown is to use active quenching, where after a breakdown event the bias voltage of the diode is temporarily reduced below $V_{bd}$. Compared to passive quenching, faster recovery times can be achieved but, in turn, this method requires dedicated electronics for each individual cell. One approach to realize this is the integration of CMOS based circuits as in the case of the digital SiPM [26]. In this design, the cell electronics are
integrated at the cell level which in addition allows the digitization of the signals of individual cells. Furthermore, cells with a high dark count rate can be switched off, which drastically reduces the dark count rate of the device. With active quenching, the cells are kept inactive for a certain amount of time which reduces contributions from temporarily trapped charge carries (afterpulsing). In addition, an outstanding timing performance was reported (50 ps FWHM). More details on CMOS based Geiger-mode avalanche photodiode arrays can be found in Ref. [27].

4. Future Prospects

More recently, arrays of G-APDs become an attractive alternative for particle tracking. The generation of a few tens of electron-hole pairs by a minimum ionizing particle passing the depletion region allows to reduce the overbias voltage, since the trigger probability is increased in comparison to a single photon absorption event. As a result, the dark count rate of the device as well as the contribution from OCT and afterpulsing is significantly reduced. Advantages of G-APD based tracking is the high detection efficiency of minimum ionizing particles and the fast signal generation of the devices (sub-ns) due to the Geiger discharge, which is required for future linear collider machines [28]. In order to achieve a sufficient spatial resolution, a cell size in the range of 25 \( \mu \text{m} \) is required and each pixel needs an individual readout. An approach for a CMOS based tracking with G-APDs is presented in these proceedings [27]. Also the semiconductor laboratory of the Max-Planck-Society is working on a SiPM with single-cell readout for particle tracking, which is based on a bump-bonding technique [14].

5. Summary and Conclusion

Silicon Photomultipliers are very promising candidates to replace photomultiplier tubes in many fields of photon detection (e.g. medical imaging, astroparticle and particle physics, biophotonics) and their application is constantly growing. They have the potential of detecting wavelengths from UV to infrared depending on the chosen material. The fast technological development, mainly driven by medical applications, already improved many properties of this rather novel detector like optical cross talk, dark count rate and afterpulsing. Also the photon detection efficiency values already reach or even outperform conventional photomultiplier tubes.

In addition to conventional devices with polysilicon or metal alloy resistors for passive quenching, different approaches have been developed in order to improve device characteristics like dynamic range, and fill factors. With respect to applications in particle detection (vertex detector) a SiPM with single cell readout and active quenching is becoming increasingly popular.

Although there has been a fast progress in this new photodetector concept there is still room for further improvements in key parameters like, dark count rate, optical cross talk, PDE, and radiation hardness of SiPMs.

References

Silicon Photomultiplier – Concepts, Characteristics, Prospects

Christian Jendrysik


