Silicon Photomultiplier: Technology Improvement and Performance

Roberto Pagano, Sebania Libertino, Domenico Corso, and Salvatore Lombardo
Istituto per la Microelettronica e Microsistemi
CNR
VIII Strada Z.I. 5, 95121, Catania, Italy
roberto.pagano@imm.cnr.it,
sebania.libertino@imm.cnr.it,
domenico.corso@imm.cnr.it,
salvatore.lombardo@imm.cnr.it

Giuseppina Valvo, Delfo Sanfilippo, Giovanni. Condorelli, Massimo Mazzillo, Angelo Piana, Beatrice Carbone, and Giorgio Fallica
Sensor Design Group
STMicroelectronics
Stradale Primosole, 50 95121 Catania, Italy
giusy.valvo@st.com, delfo.sanfilippo@st.com,
giovanni.condorelli@st.com, massimo.mazzillo@st.com,
angelo.piana@st.com, beatrice.carbone@st.com,
giorgio.fallica@st.com

Abstract— **Our main results on the study of both single pixels and Silicon Photomultiplier arrays produced by STMicroelectronics in Catania are reviewed. Our data, coupled to an extensive simulation study, show that the single pixel technology is close to its ultimate physical limit. The distribution of dark current in large arrays follows a Poissonian law. Cross talk effects are strongly reduced by the presence of optical trenches surrounding each pixel of the array. Finally, we demonstrate that these devices can also be used as single photon counters also without a complex amplification stage.**

Keywords – Silicon Photomultiplier; dark count; trenches.

I. INTRODUCTION

The ability to detect single photons represents the ultimate goal in optical detection. To achieve such sensitivity a number of technologies have been developed and refined to suit particular applications. These include: Photomultiplier Tubes (PMTs), Microchannel Plate Photomultiplier Tubes (MCPMT), Hybrid Photon Detector (HPD), p-i-n photodiodes, linear and Geiger mode Avalanche Photo Diodes (APDs), etc. [1], [2], [3], [4], [5]. The need for ever more sensitive, compact, rugged, and inexpensive optical sensors in the visible region of the spectrum continues today, and it is particularly acute in the fields of the biological sciences, medicine, astronomy, and high energy physics. Applications such as fluorescence and luminescence photometry, absorption spectroscopy, scintillation readout, light detection and ranging, and quantum cryptography require extremely sensitive optical sensors often in adverse environments, such as high magnetic fields, and where space is limited.

In many of these applications, the PMT has become since mid-1930's the detector of choice almost without a convincing alternative. However, PMT presents some disadvantages: it is fragile, it requires high operating voltage (higher than 100 V), and it can not operate without a shielding protection in magnetic environment. Since its inception in the 1980's, the so-called Silicon Photomultiplier (SiPM) has begun to rival the PMT in many of its parameters such as gain, photon detection efficiency, and timing [6], [7], [8], [9], [10], [11]. The SiPM has all the additional benefits of silicon technology such as compactness, reliability, ruggedness, high volume of production, and long term stability. Although all the previous motivations would be sufficient to explore this alternative to PMT, it is the low production cost of silicon technology that attracts the most and has led to the efforts that finally has enabled the realization of this photodetector.

The SiPM major drawback is the relatively large dark current [11], due to the combination of a diffusion current produced at the quasi-neutral regions at the boundaries of the device active region and of generation of carriers due to point defects and/or metallic impurities in the active area depletion layer emitting carriers through the Schockley-Hall-Read (SHR) mechanisms, eventually boosted by the Poole-Frenkel effect [12].

In this paper, after a detailed discussion on the principle of operation of SiPMs presented in Section II, two different pixel design technologies of SiPM developed by STMicroelectronics are discussed. They consist in a n^+ on p silicon structure. The device active part is the same in both pixels. They differ in the doping of the epitaxial layer and in the starting substrate (n-type Si for the first device and p-type Si for the second device). The differences between the two technologies lead to significant differences in the dark count rate (*DC*) measured at temperatures higher than 10°C. In Section III the two device structures, the experimental setup, and the simulation environment used are discussed. The current–voltage characteristics in forward and reverse bias of the two pixels for temperatures ranging from -25°C to 65°C are presented and discussed comparing the measured data with electrical simulations in Section IV. Moreover, electrical and optical performance of SiPM devices suitable for large scale fabrication in a VLSI production line are reviewed in Section V. Finally, the conclusions are outlined in Section VII.

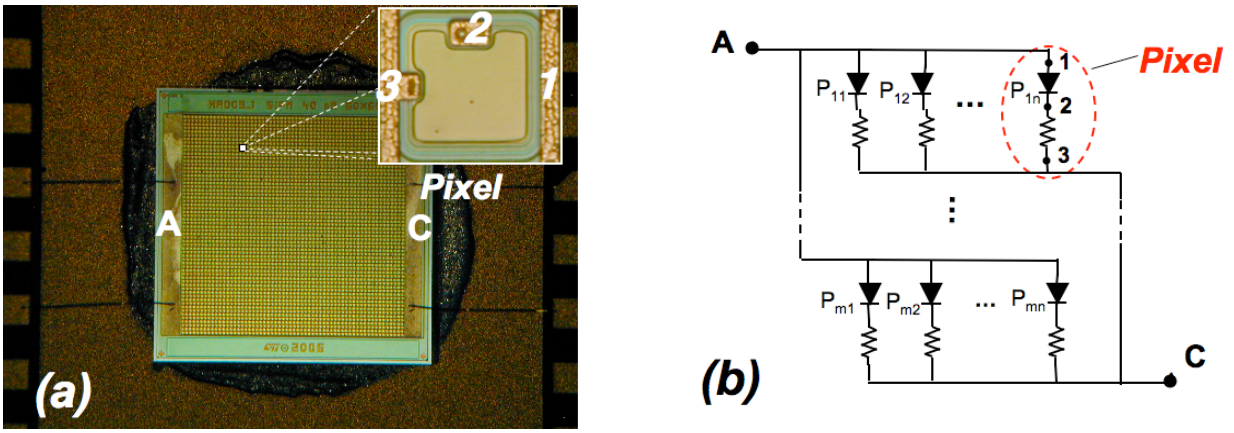

Figure 1. (a) Microphotograph of a 64×64 pixels SiPM with 6.5 mm^2 active area produced by STMicroelectronics.The pixel, shown in the inset, has an active area of 40 µm × 40 µm. (b) Schematic circuit diagram of a SiPM with n×m pixels. The pixel, enclosed by the dashed line, is composed of an SPAD and a quenching resistor. Note that node 1, 2 and 3 are the same in Fig. (a) and (b).

II. PRINCIPLE OF OPERATION

The principle of operation of SiPMs is inspired to the demand of information on the exact determination of the arrival time and density of a very low photon flux. Due to the quantum nature of light, a low photon flux is composed by few photons distributed in time and space. A dense array of space distributed micro-devices (the pixels), individually able to detect the arrival time of a single photon can, in principle, resolve the time and the space distribution of the impinging photons. This is the basic SiPM operation principle. In a SiPM the pixels are electrically connected in parallel forming a matrix of n× m adjacent sensors (see Fig. 1*a*). Each pixel, known in literature as Single Photon Avalanche Diode (SPAD) or Geiger Mode Avalanche Photodiode (GM-APD) [13], [14], consists of a p-n junction suitably doped in order to have avalanche breakdown in a well defined active area with an integrated quenching resistance in series, as shown in the schematic picture of Fig. 1*b*. The active area is formed by creating an enriched well zone, generally doped by ion implantation followed by thermal processing for dopant activation and defect annealing. This dopant local enrichment generates regions where the vertical junction electric field is higher, and these become the pixel active areas for photon detection [15]. The p-n junction devices are operated in Geiger mode, that is, they are biased above the junction breakdown voltage (*BV*). Above *BV* the device can stay in a quiescent state for a relative long time, up to ms, depending on the technology quality process (low defect density) and the operating condition (temperature and voltage) [16]. Then, when the device is in quiescent state, the active volume (active area times the depleted region) is characterized by an electric field well above the breakdown field. However, the p-n junction does not go into avalanche breakdown. In such condition, the absorption of a single photon in the active volume will trigger, through the generation of an electron-hole pair, the onset of the junction avalanche with a probability depending on the operating voltage [17]. A macroscopic current pulse flows through the junction resulting in a strong amplification of the single photon arrival. The amplification value, usually indicated as Gain (*G*), is above 10^6 electrons per pulse. This large gain is the cause of the strong SPAD sensitivity.

The avalanche process is a self-sustaining process and to quench it the integration of a resistor in each pixel is required. In our device the resistor is connected in series to the cathode of the p-n junction (see Fig. 1). The quenching mechanism introduced by the resistor acts as follow: the avalanche following the photon absorption causes a rapid increase of the current flowing through the p-n junction as well as through the external resistor. The voltage drop across the series resistor decreases the voltage applied to the p-n junction below the *BV*, forcing the avalanche quenching and the consequent extinction of the current flux. Once the avalanche is quenched, a recharge time is required to restore the pixel to the original condition of electric field above *BV*, making the pixel ready to the detection of a new photon [18]. Therefore, the detection of a photon by a single pixel results in a current pulse, which can then be easily measured by an external circuit. However, a single pixel works as a digital photon sensor, that is, it can not detect multiple photon arrival. This task is accomplished by the full array. In fact, the current detected by the overall SiPM is simply the sum of the currents produced by the various pixels. Hence, this device compared to the original design of the SPAD has the advantage of having a response, which is in a relatively large dynamic range, proportional to the flux of photons impinging on the detector at the same time [6], [7], [8], [9], [10], [11]. The SiPM Gain is about 10^6, similar to the one of single pixel.

The capability to measure low photon fluxes is limited by the device *DC*. In fact, the single pixel can have a breakdown event even if it does not detect a photon because of thermal generation of electron-hole pairs within the depletion region assisted by defects (SHR mechanism) and / or of minority carrier diffusion from the depletion region boundaries. Such events result in current spikes having the same features of the "real" counts due to photon arrival. This determines a lower limit to the photon count rate. A high performance SiPM

must have a very low *DC* rate. Typical value is in the order of 1MHz/mm^2 at room temperature.

Another limiting factor to the device operation is the cross-talk effect. The cross-talk is a noise contribution common to all pixelated devices. A pulse current produced by a pixel, due to a photon detection event or to a primary dark noise event, can induce one or more adjacent pixels to experience the avalanche breakdown. Then, the corresponding output pulse current of the SiPM has an amplitude peak proportional to the sum of the pixels involved in the single photo-detection and in the correlated cross-talk phenomena. This noise contribution is detrimental for all the applications where a single photon resolution is required.

The cross-talk noise has two different physical origins: optical and electrical. The optical cross-talk is due to the photons generation by radiative emission from the hot carriers produced during an avalanche discharge. In an avalanche multiplication process, on average 3 photons, with energy higher than the silicon band gap (1.12eV), are emitted every 10^5 carriers [19]. These emitted photons can travel to a neighboring pixel and trigger a breakdown there.

The electrical cross-talk can occur when carriers, generated during the avalanche breakdown in a pixel, can travel along the epitaxial layer, common to all pixels, and reach the neighboring pixels triggering there a new avalanche breakdown [20]. Some strategies have been studied to reduce the cross-talk between neighboring pixels. The first is to increment the distance between adjacent pixels. This approach has a detrimental effect on the geometrical fill factor of the SiPM. The second strategy consists in fabricating grooves, filled with optical absorbing material, all around each pixel. These grooves, commonly named trenches, prevent from optical and electrical coupling between pixels. The reduction of the geometrical fill factor with such design is mild while the effect on the cross-talk noise is considerable. The devices studied in this paper are fabricated using the second approach and the beneficial effects provided by the trench presence are discussed in Section V-A.

An important feature of SiPM, as one would expect from a semiconductor device, it is the long term stability of its parameters (*BV*, *G*, *DC,* etc.). In many applications, in fact, the variation of such parameters in the course of time may be a problematic issue. As reported by other authors SiPMs have no aging and show stable parameters even if exposed to elevated temperature for long time [10], [21], [22].

III. EXPERIMENTAL

Electrical characterization was performed at wafer level using a Cascade Microtech Probe Station *11000*. The samples were cooled using a Temptronic *TPO 3200A* ThermoChuck that provides a stable temperature, within 0.1°, between -60°C and 200°C. Current *vs.* voltage measurements (I-V) were acquired using an *HP4156B* precision semiconductor parameter analyzer with an integration time of 1s. The *DC* and the gain were measured using a Tektronix *DPO 7104* Digital Oscilloscope (OSC) with 1 GHz bandwidth and 20 Gsa/s measuring the voltage drop through a 50 Ω resistor connected between the cathode of the pixel and the ground. The I-V characteristics have been measured on more than 30 devices, for both single pixels and SiPM arrays of both technologies.

Optical measurements were carried out using a *Cube* laser diode with a wavelength of 659 nm by Coherent working from continuous wave to 6 ns pulses. Modulation was achieved using an external trigger (Agilent *81110A* 165/330 MHz). The device was biased and the signal acquired using the source-meter and oscilloscope already mentioned.

A. Device structures

In this work, two different technologies are compared. They differ for few, but important, characteristics. The full device fabrication details can be found in [23]. In this paper, we want to focus our attention on similarities and

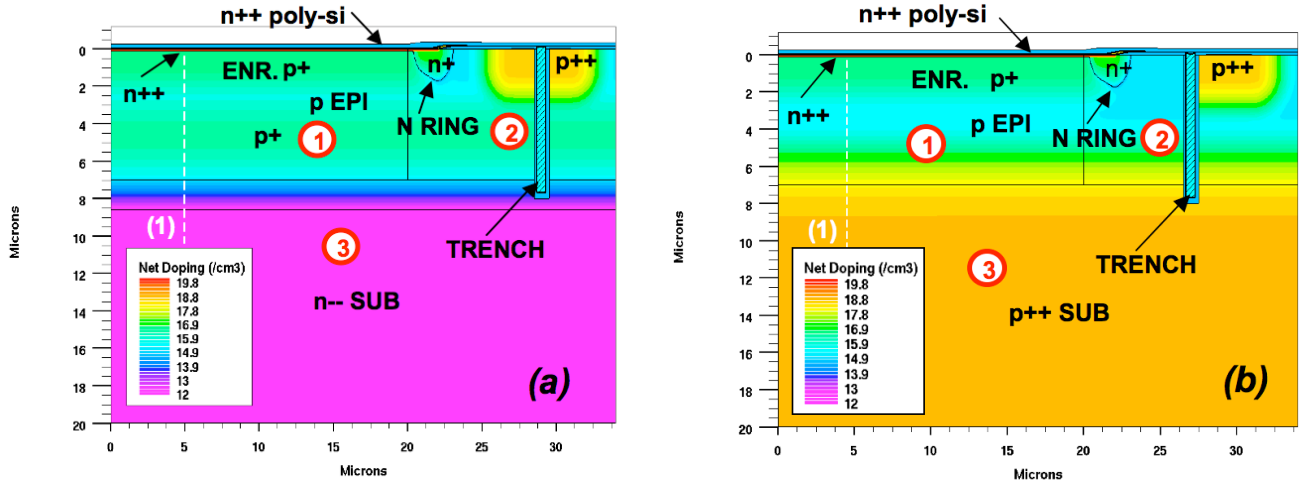

Figure 2. Schematic cross-section of a SiPM pixel. (a) Device 1: double epitaxial layer, n-substrate, trenches crossing sinker diffusion (b) Device 2: single epitaxial layer, p-substrate, isolated sinker diffusion.

differences between the two technologies. They have the same device active part and guard ring, fabricated as discussed in [23], the same *BV* (-28.2±0.3V at 25°C) and the same active area, of 40×40 µm^2.

Fig. 2 shows a half cross section of the two studied technologies, Fig. 2*a* and 2*b* for device 1 and 2, respectively. The main differences between the two technologies are the doping of the epitaxial layer and the starting substrate. In the first technology, device 1, a double epitaxial layer, first p^+ layer, then followed by a p-type Si, is grown on a low doped n-type (100) oriented Si substrate. In the second technology, device 2, only a single p-type epitaxial Si layer is grown on a highly doped p^+-type (100) Si substrate. In both cases, deep optical trenches are fabricated for the optical and electrical isolation between the pixels [24]. In device 2 the optical trench is closer to the active region than in device 1 (see Fig. 2). As a result, the thermal diffusion of the p^{++} implanted sinker needed for the anode contact, is shielded by the trenches.

In both devices, the same anti-reflecting coating, poly-silicon quenching resistance and metal contact are integrated as discussed in [23]. These elements are not included in the cross sections of Fig 2.

Fig. 2 also shows three regions enumerated as 1, 2 and 3: region 1 is the central epitaxial layer below the active area of the pixel (0 µm < x < 20 µm and 0 µm < y < 7 µm for both devices); region 2 is the border epitaxial region of the pixel (x > 20 µm and 0 µm < y < 7 µm) and region 3 is the substrate (y > 7 µm). They have been identified for simulation purposes. It allows to define trap energy and lifetime separately for each region, to better simulate the real structure.

Fig. 3 shows the comparison between the data (symbols) and the simulated results (lines) of the final net doping along the cut line 1 (dashed with lines in Fig. 2) for device 1 (blue solid line) and device 2 (red dashed line). The experimental doping profile was obtained by the spreading resistance measurements for both device 1 (blue squares) and device 2 (red circles). The simulated profiles follow quite well the experimental data. It is important to stress that the profiles of the two devices overlaps in the full active region. The main differences are in the region below 2 µm. Part of it is highlighted in Fig. 3 (EPI).

The structural differences described so far are at the base of the electrical behavior shown in the following sections.

B. Simulation parameters

Electrical simulations were obtained using a 2-D drift-diffusion solver developed by Silvaco Co. LTD [25].

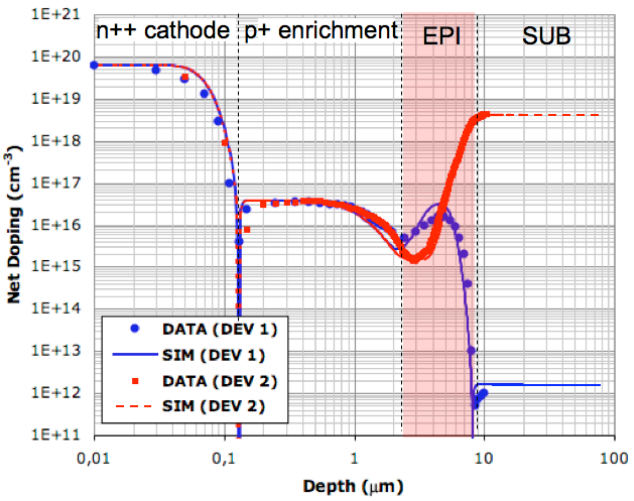

Figure 3. Doping profile of device 1 (blue) and device 2 (red) experimentally measured (symbols) and simulated (lines).

The adopted model is the drift-diffusion approximation including standard SHR generation/recombination, Auger recombination, band gap narrowing, Coulomb scattering, and SHR surface recombination. The parameters of the TCAD simulations are those connected to the minority carrier lifetime in the three device regions above described. The SHR generation (*G*) / recombination (*R*) adopted model consists of the following equations:

$$G - R = \frac{pn - n_{ie}^2}{\tau_p \left[n + n_{ie} \exp\left(\frac{E_T}{kT}\right)\right] + \tau_n \left[p + n_{ie} \exp\left(-\frac{E_T}{kT}\right)\right]} \quad (1)$$

where:

$$\tau_n = \frac{\tau_{n0}}{1 + N_D / N_{\mathrm{Re}f}}, \tau_p = \frac{\tau_{p0}}{1 + N_D / N_{\mathrm{Re}f}} \quad (2)$$

where p, n, and n_{ie} are the hole, electron, and intrinsic carrier concentration, E_T and kT are the trap and the thermal energy, N_D is the local dopant concentration, and N_{Ref}=5×10^{6}cm^{-3}. We have assumed that τ_{n0}=τ_{p0}=τ_0. In the model we assume three different values for τ_0 and E_T in the three above defined device regions, i.e., τ_{01}, τ_{02}, τ_{03}, E_{T1}, E_{T2}, and E_{T3} resumed in Table I. In the same table are also resumed the experimental activation energies discussed in Section IV-B.

TABLE I. SIMULATION PARAMETERS AND EXPERIMENTAL ACTIVATION ENERGIES

Device	Simulation							Experimental	
	τ_{01} (s)	τ_{02} (s)	τ_{03} (s)	E_{T1} (eV)*	E_{T2} (eV)*	E_{T3} (eV)*	SR (cm/s)	E_{a1} (eV)	E_{a2} (eV)
1	10^{-3}	10^{-5}	10^{-5}	0.06	0.2	0	100	0.57	1.18
2	10^{-3}	10^{-5}	10^{-3}	0.06	0.2	0	150	0.59	1.12

* Energies values are with respect to the midgap of the Si energies bandgap.

IV. SINGLE PIXEL TECHNOLOGY

In this section, the experimental data of the two single pixel technologies described previously are discussed and compared with the electrical simulation for both forward and reverse bias.

A. Forward current

In this paragraph, the pixel forward regime will be discussed. The study of the pixel behavior in forward, even if this is not the regime for photon-detection, is functional to understand the causes leading the differences in the *DC* of the two devices. Moreover, the simulations presented in this paper are the results of the best fit obtained from forward and reverse currents at different temperatures and for different geometries as discussed in the following.

The forward current for both devices has a dominant component at the perimeter of the active area. This effect has been observed in pixels of both types having different active areas (A_{ACT}) and dead areas (A_{DEAD}). The dead area of a pixel is the area surrounding the active region as shown in the inset of Fig. 4. In the same figure, the projection of the measured I-V in the ideal diode regime to the y axis (V=0) (symbol), i.e, the pre-factor I_0 of the Schockley diode equation [26], is reported for pixels with three different A_{ACT} and four different A_{DEAD} compared with simulations (dashed line). The data clearly show that I_0 is almost independent from A_{ACT} and has strong dependence on A_{DEAD}. This geometrical information has been taken into account in the electrical simulation defining the physical parameters discussed in Section III-B. Moreover, a surface recombination model [27], with velocity S_R at the boundary between silicon and oxide, is included in region 2. The final parameters, almost the same for both devices, are summarized in Table I.

Such high difference between the carrier lifetime (electron and hole) in region 1 and in region 2 produces a preferential current path at the perimeter of the p-n junction, as suggested by the experimental data. This effect is clearly visible in Fig. 5 that shows the 2D distribution of the total current density (J_{tot}) at 25°C and for a forward bias of 0.3V

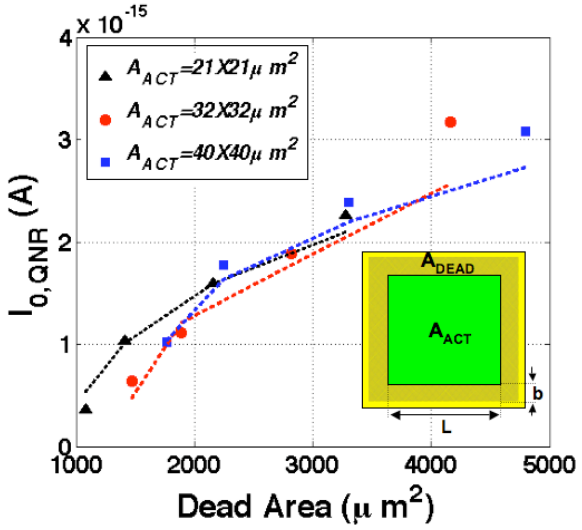

Figure 4. Measured (symbols) and simulated (dotted line) I_0 as function of the active area and the dead area for device 1 @ 25°C. The inset is a plane view of a pixel showing the active area and the dead area.

in both device 1 and 2 (Fig 5 *a* and *b*, respectively). The dashed circle in Fig. 5 highlights the interested region.

Fig. 6 shows the measured I-V (symbols) of device 1 (Fig. 6*a*) and device 2 (Fig. 6*b*) at three different temperatures: -25°C (circles), 25°C (triangles) and 65°C (squares). The measured data are compared with the simulated I-V (dashed line). Two regimes can be clearly observed: the ideal diode following the Schockley law at low voltages (linear region) and the resistive regime due to the integrated quenching resistor R_Q at higher voltages. Actually, the current of device 1 at high voltages deviates from the simulated current (in the range 0.4V - 0.5V depending on temperature). This is due to a parasitic Schotky diode at the anode contact that has been removed in device 2. In the simulation this effect has not been considered. The effect of the R_Q was simulated including an ideal resistor at the cathode contact equal as the measured value in both devices (220 kΩ at 25°C).

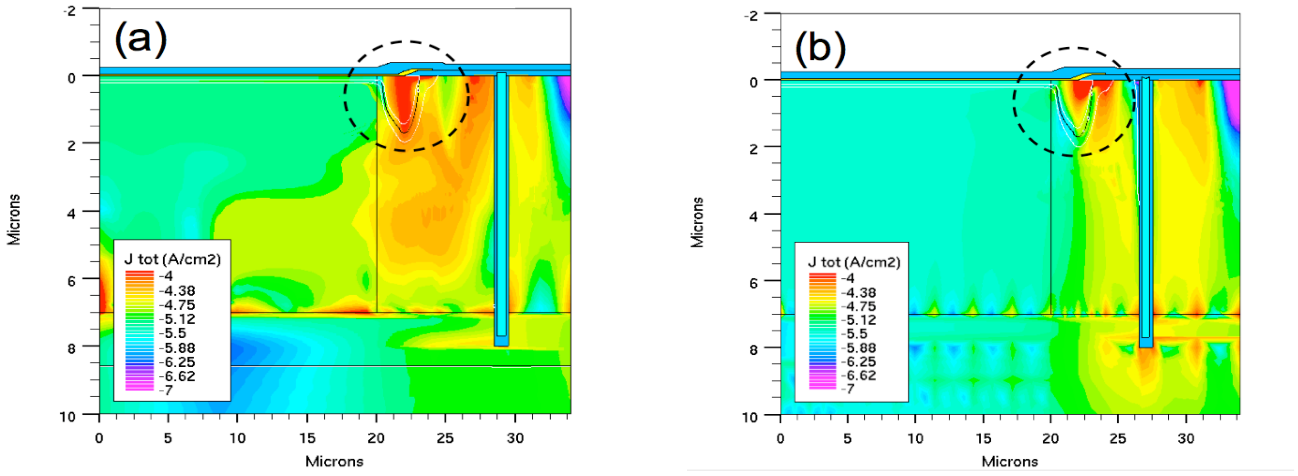

Figure 5. Distribution of the total current density (Jtot) at 25°C and at V=0.3V of (a) device 1 and (b) device 2.

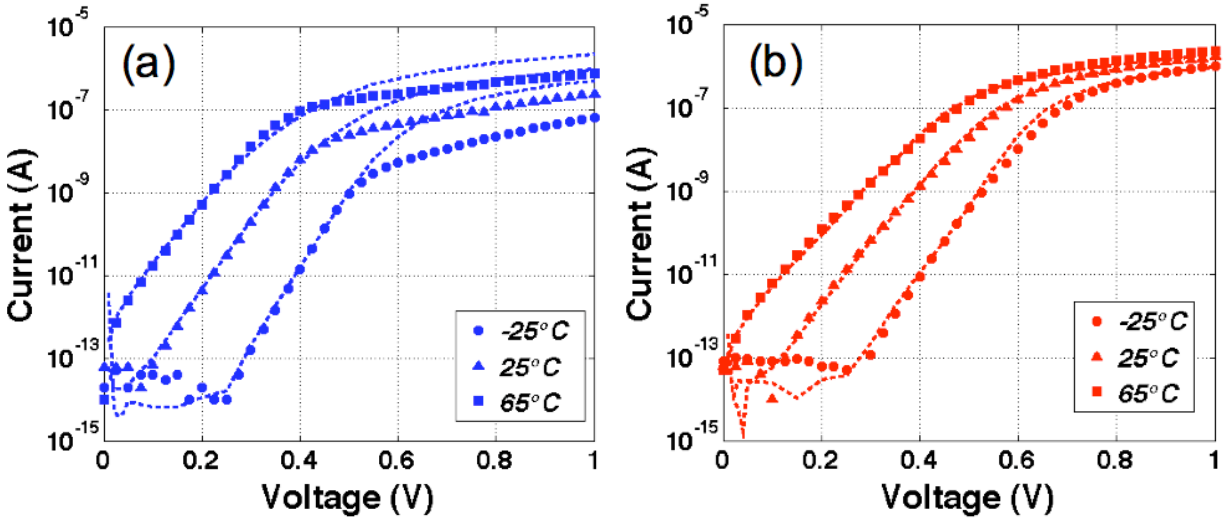

Figure 6. Measured (symbols) and simulated (dotted line) IV in forward polarization at three temperatures of (a) Device 1 and (b) Device 2.

The simulation shows a very good agreement with the experimental data for both devices and deviates only in device 1 as just discussed. The simulations have been carried out using the parameters summarized in Table I.

B. Reverse current and Dark Count

Fig. 7 shows the dark currents as a function of voltage at three different temperatures, -25°C (circles), 25°C (triangles) and 65°C (squares) for a single pixel belonging to device 1 technology (blue symbols) compared to the dark current of a pixel with the structure of device 2 (red symbols). The *BV* of the two pixels is the same, -28.2 V at 25°C, with a temperature coefficient of -29mV/°C.

The leakage currents, i.e., the currents at voltage below the *BV*, are nearly the same for the two kind of devices in the full range of temperature, ~10pA at 25°C. However, the currents at voltage above *BV* increase with a different rate with respect to the temperature. At -25°C the dark currents (circles) are roughly the same while, by increasing the temperature, they show remarkable differences. At 25°C and voltages of -32V (+ 3.8V over-voltage, *OV*) the dark current in the device 1 is one order of magnitude higher than that of the device 2. At 65°C the difference increases approaching two orders of magnitude. When the pixel works as photon detector it is biased above breakdown and the dark currents define the lower limit to the photon rate detectable. The understanding of the physical origin of these currents is an important achievement to improve the device technology. Although the measurements show steady-state I-V curves, the time resolved analysis of the current at the oscilloscope, at a fixed bias above *BV*, reveals that the time averaged current of Fig. 7 is a random sequence of current spikes.

Fig. 8 shows a trace of a single pixel dark current at *OV*=+3.8V, at 25°C in a time window of 1ms acquired with the OSC. Five current spikes with ~90 µA amplitude randomly distributed in time are clearly visible. The

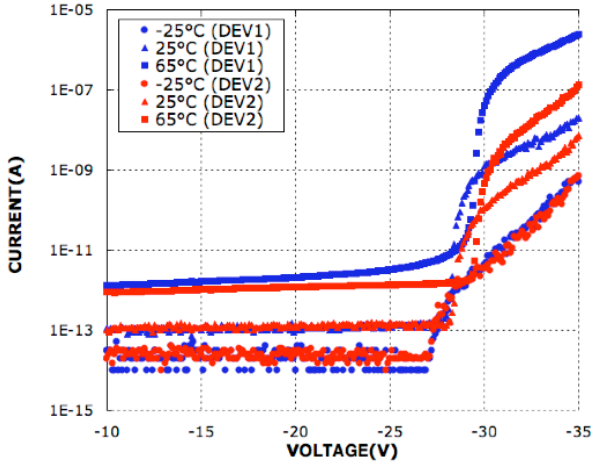

Figure 7. I-V in dark and in reverse bias at -25°C (circle), 25°C (triangle) and 65°C (square) of device 1 (blue) and of device 2 (red).

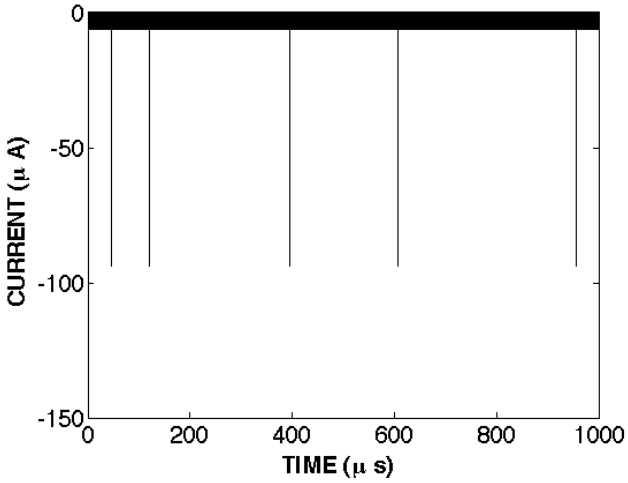

Figure 8. Dark Current v.s time at 25°C and at OV=+3.8V.

frequency of these spikes is the *DC* of the pixel. These dark counts are attributed to generation inside the depleted region of the junction and / or diffusion from quasi neutral boundaries of a single free carrier which triggers the avalanche. The integrated current signal of a dark count in a short time window, typically 50-100 ns, divided by the electron charge q, is usually referred as the gain of the pixel (G). It was demonstrated in a previous work [28] that the steady-state dark current at any temperature and voltage condition of Fig. 7 is the product of q, G and DC, in symbols:

$$I(V,T) = q \times G(V,T) \times DC(V,T) \quad (3)$$

It is clear that the difference between the dark currents of the two devices at 25°C and 65°C (Fig 7) is necessarily due to a difference or in the G or in the DC.

Fig. 9 shows the measured G of device 1 (blue symbol) and of device 2 (red symbol) at voltages higher than the BV and at three temperatures: -25°C (circles), 25°C (triangles) and 65°C (squares). G was measured as described in [28] integrating the mean dark pulse. As the data show, the gain is nearly the same for both devices at all the investigated temperatures. This is expected because $G = 1/q \times C \times OV$, C is the junction capacitance, and the values of C and OV are the same for both devices.

The DC of the two devices is shown in Fig. 10 for the same voltage and temperature ranges investigated for the G (Fig 9). It was measured counting the dark pulses in a time window of 1s. Blue symbols are used for the DC of device 1 and red symbols for the DC of device 2. At -25°C (circles) both devices have the same DC (~10 hz at V=-31V). At 25°C (triangles) the DC of device 1 is ~ 10 times the DC of device 2 and at 65°C the difference becomes ~ 2 order of magnitudes, roughly the same difference observed in the dark current of Fig. 8.

To better understand the DC behavior with respect to the temperature, the DC of both devices has been measured at fixed OV in a temperature range of 100°C, from -25°C to 85°C.

The Arrhenius plot of the DC at a fixed OV=+3V, i.e. the Napieran logarithm of the DC vs. $1/kT$, where k is the Boltzmann constant and T the temperature in Kelvin, is shown in Fig. 11. The experimental data for device 1 (blue symbols) are compared with those of device 2 (red symbols). Lines are the simulation results, as discussed in the following. The slope of the ln(DC) as a function of $1/kT$ provides the DC activation energy E_a [29]. Two different slopes are recognized from the plot: for temperature lower than ~ 10°C for device 1 and ~ 40°C for device 2 the activation energy is $E_{a1} \sim E_G/2$, where E_G is the silicon forbidden energy bandgap (1.12eV). At higher temperature (>10°C for device 1 and >40°C for device 2) the slope of the Arrhenius plot provides an activation energy $E_{a2} \sim E_G$. The experimental values of E_{a1} and E_{a2} are summarized in Table I. Similar values are found regardless of OV. E_{a1} value indicates that at low temperature the DC of both devices is due to SHR generation-recombination defects located inside the depleted region of the p-n junction. The physical explanation of $E_{a2}=E_G$ is that the diffusion of minority carriers from the boundary of the depleted region is the prevalent effect causing the DC at high temperature. Now it is clear that the larger reduction of the DC of device 2 with respect to the DC of device 1 at temperatures higher than 10°C (Fig. 10) is due to a reduction of the diffusion current in device 2. However, it is still unclear why it happens. At a first glance, one may expect that a reduction of the diffusion current at the perimeter of the active area could cause a reduction in the DC, as already observed in the forward regime [1]. This hypothesis could be suggested also by the different device architecture at the borders, region 2 in Fig. 2, for the two devices. Device 1 shows a large p-type dopant

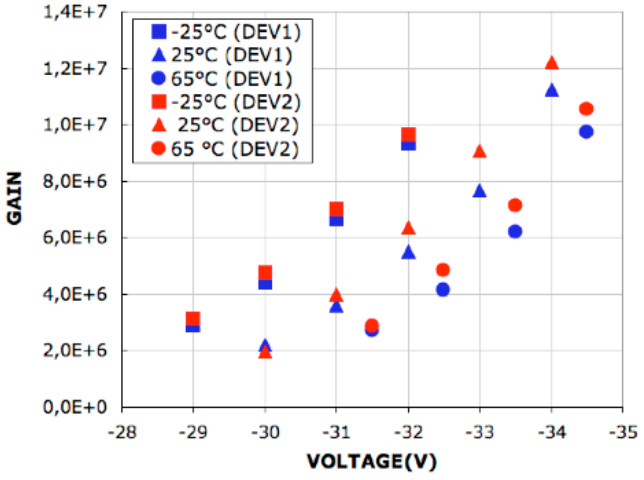

Figure 9. Gain v.s. voltage at -25°C (circle), 25°C (triangle) and 65°C (square) of device 1 (blue symbols) and of device 2 (red symbols).

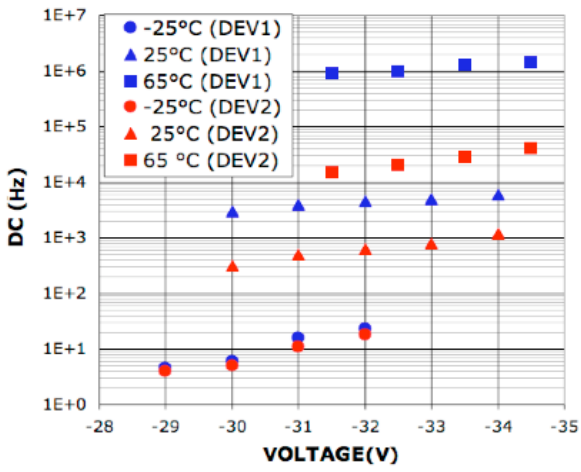

Figure 10. Dark count rate v.s. voltage at -25°C (circle), 25°C (triangle) and 65°C (square) of device 1 (blue) and of device 2 (red).

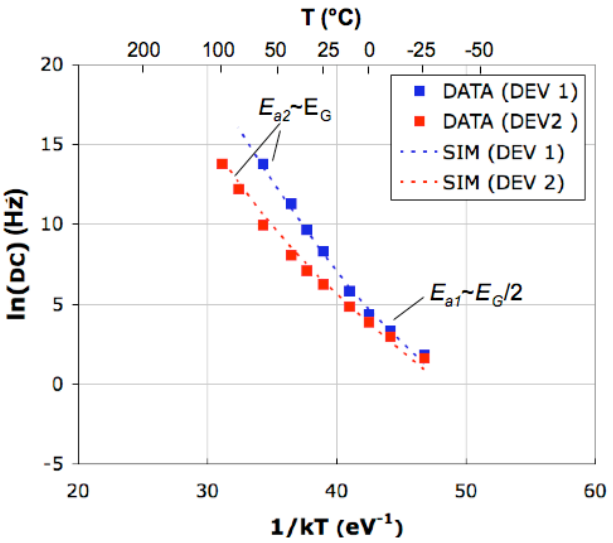

Figure 11. Comparison of the Arrhenius plot of DC measured (symbol) and simulated (dotted lines) at constant OV=+3 V (10%) for device 1 (blue) and device 2 (red).

concentration (2×10^{18}/cm^3), while device 2 has the epitaxial Si concentration value (~1×10^{15}/cm^3). Since the Auger effect [30] is a relevant recombination mechanism at high dopant concentration, a different effective lifetime in the periphery of the two devices could explain the lower diffusion current. A more careful inspection of the results, supported by our electrical simulations, allowed us to obtain a different conclusion. The electrical behavior of both devices was simulated at different temperatures and reverse bias conditions varying the devices physical parameters in the three defined region (see Fig. 2). τ_0 was varied in the range 10^{-7} – 10^{-3} s and E_T in the range 0 – 0.3 eV from to $E_G/2$.

The simulation shown in Fig. 12 refers to 65 °C since at this temperature only the diffusion regime is present. It shows the current density distribution of device 1 at OV=+3V. Even if the 2D drift-diffusion simulator can not reproduce the exact value of J_{TOT} above BV, i.e., it can not simulate the device operation in the Geiger Mode, it gives important information. At voltages above the BV, the current flows preferentially at the center of the device. This behavior was found for all the explored parameter set and for both devices, since it is due only to the electrical field distribution.

Fig. 13 shows the 2D simulation of the electrical field of device 1 at 65°C and +3V of OV (same conditions of Fig. 12). The electrical field at the lateral border is well below the junction breakdown value. It is negligible with respect to the maximum value in the active region, in which, the field is above the value needed for avalanche breakdown. Even if the diffusion current in forward bias or in reverse bias at voltages below BV has it maximum value at the border of the device junction, above BV the probability to trigger an avalanche in this region is too low. The results above described allowed us to conclude that the large reduction of the diffusion current between the two devices, is due to differences in the region 1 physical characteristics. The simulations do not allow us to obtain information on the dark current value above breakdown, but can discriminate between the different components of the leakage current below breakdown. It should be reminded that experimentally, the leakage current below breakdown is due to three components: minority carrier diffusion and SHR generation in region 1, and perimeter current [31]. Since the first two components contribute also to the dark current, the main difference between the two reverse bias regimes (below and above breakdown) is due to the presence of perimeter current below breakdown. In fact, a carrier coming from the perimeter has a probability close to zero to trigger an avalanche [32]. Moreover, the sum of the first two current components is well below the experimentally measured value, demonstrating that the leakage current below breakdown is entirely dominated by the perimeter component. Similar considerations and results hold for device 2.

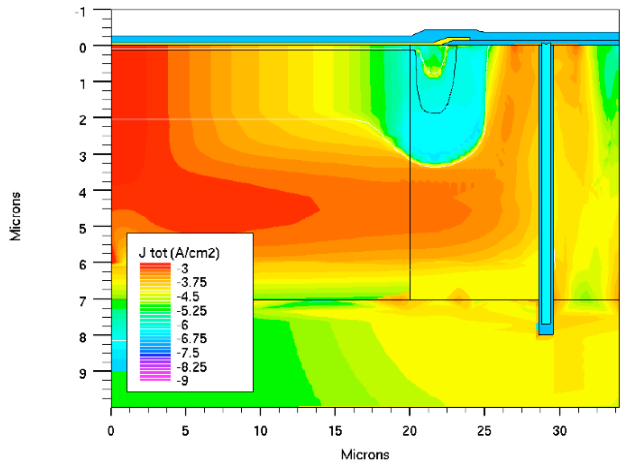

Figure 12. Total current density distribution (Jtot) at 65°C and at OV=+3V of device 1.

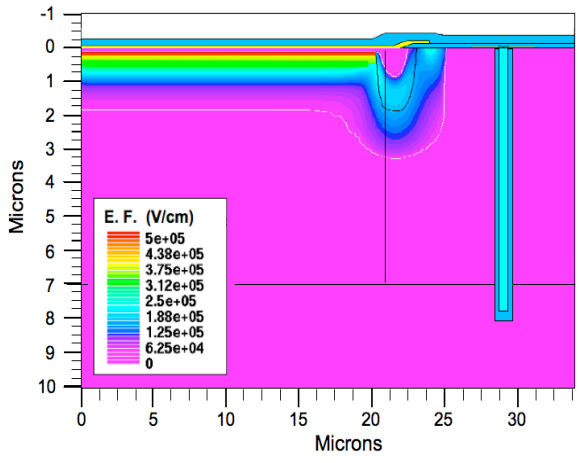

Figure 13. Electrical field distribution at 65°C and at OV=+3V of device 1.

Fig. 14 shows the simulated total current density at 65°C and at -20V of device 1.

The simulated *DC* of Fig. 11 was then obtained considering the J_{TOT} value taken in the center of the depletion layer in region 1 for V=-20V, as shown by the cut line 1 in Fig. 14. Moreover, a uniform triggering probability, P_t, of 0.35 at *OV*=+3V was assumed, as calculated in [33] for a similar device. In symbols:

$$DC=A_{ACT}\times P_t\times J_{TOT}/q. \quad (4)$$

The best fit parameters obtained predict the same τ_0 and E_T for both devices. In region 1, τ_0=1ms and the SHR trap energy is E_T=60 meV below midgap, i.e., 0.54 eV, while in region 2 τ_0=10 µs and E_T=200 meV.

The lower simulated J_{TOT} in the active area of device 1 with respect to that of device 2 (not shown) is due to a large difference in the diffusion component of the *DC* for the two devices at temperatures higher than 10°C. The simulation results are shown in Fig. 11 with the dashed line for device 1 and solid line for device 2.

The differences in the diffusion current passing trough the p-n junction between the two devices in region 1 at 65° must be due to different contribution of the diffusion current components. J_{TOT} is the sum of the diffusion electron current (J_{e-}) and the diffusion hole current (J_{h+}) The first is due to minority electron carriers diffusing from the p bulk to the n^{++} cathode, the latter is due to the diffusion of minority hole carriers from the cathode to the p bulk. The J_{TOT} (squares), J_{e-} (continuous line) and J_{h+} (dashed line) of device 1 (blue) and of device 2 (red) along the cut line 1 of Fig. 14 in the first 2 µm of depth at 65°C and for a voltage polarization of -20V are summarized in Fig. 15. In device 1, J_{TOT}~J_{e-} while in device 2 J_{TOT}~J_{h+} as shown in Fig. 15.

The reduction of the J_{e-} current in device 2 is the cause of the strong reduction of the diffusion current. In fact, J_{h+} is the same in both devices being only due to the doping profile of the n^{++} cathode. As discussed in [34], the J_{e-} decrease in device 2 is mainly due to the doping profile in the epitaxial region, shadowed area in Fig. 3. In fact, minority carriers in this layer (electrons) have a different gradient in their concentration in the substrate direction in the two devices. The gradient is higher in device 2 with respect to device 1, leading to a diffusion of electrons toward the substrate higher than in device 2. As a result, the net diffusion current of electron toward the cathode is reduces below the hole diffusion current flowing in the opposite direction.

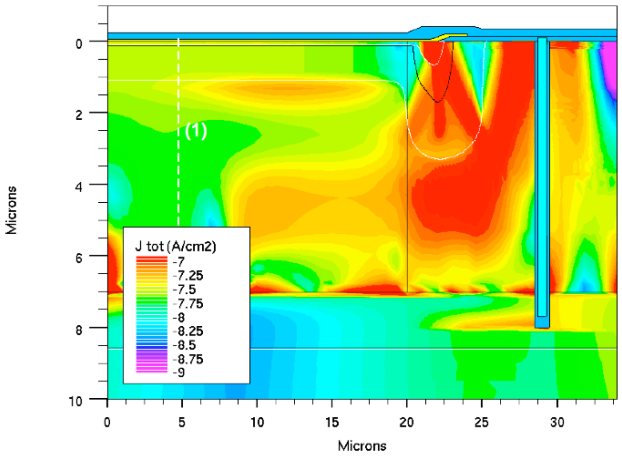

Figure 14. Total current density distribution (Jtot) at 65°C and at V=-20V of device 1. The current inside the depletied region along the cut line 1 is the value cosidered for the DC simulation .

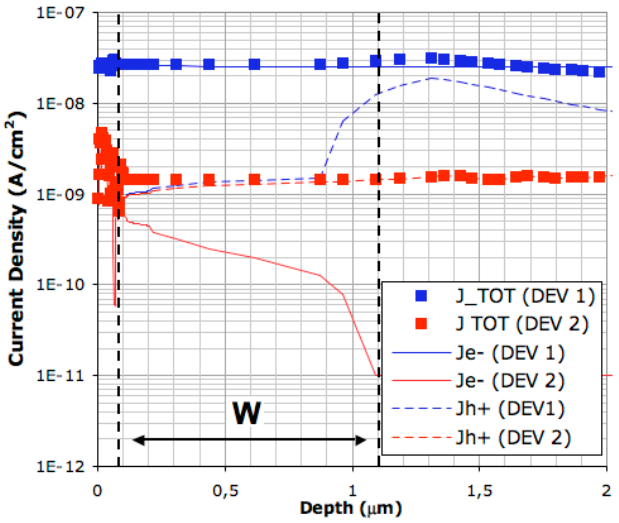

Figure 15. JTOT (squares), Je- (continuous line) and Jh+ (dashed line) at 65°C and at -20V along the cut line 1 of fig. 14 of device 1 (blue) and device 2 (red). W is the length of the depleted region of the p-n junction.

Finally, Fig. 16 shows the comparison of the experimental *DC* (symbols) and the *DC* simulated without SHR generation and recombination (dotted line) at *OV*=+3 V with respect to 1/*kT* for device 2. The comparison shows that at room temperature the experimental *DC* is close to its minimum physical level due to *DC* diffusion component. It is to note that the diffusion process of minority carriers is an intrinsic property of p-n junctions and cannot be avoided. In device 2, diffusion current has been reduced to its minimum value, being dominated by the cathode design. In order to achieve a further reduction of the hole diffusion, a different architecture must be designed. An improvement of the *DC* at room temperature can be reached reducing the defect concentration in the depleted region. We estimated the presence of about 1.6±1.3 defects/cm^{-3} in both devices hence a further reduction is a difficult goal to achieve.

V. SiPM PROPERTIES

In this section, electrical and optical performance of SiPM full array are presented and discussed.

A. Dark count and Cross Talk

The *DC* in a SiPM is conventionally defined as the frequency of dark pulses exciding half of the amplitude of the signal produced by one photo-electron (p.e.) [35].

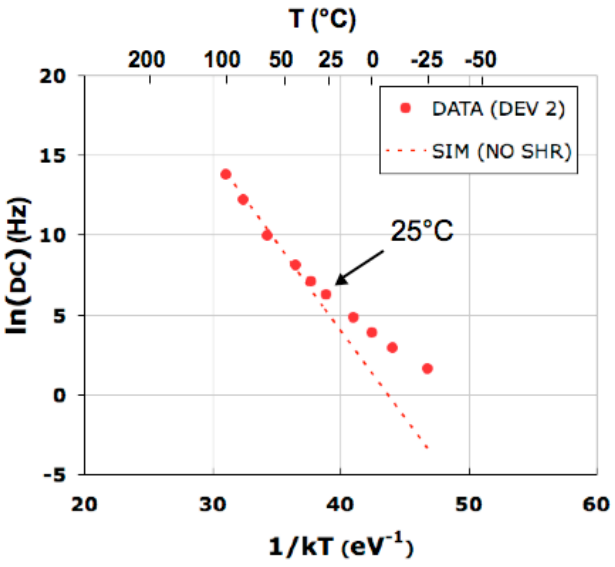

Figure 16. Comparison of the Arrhenius plot of DC measured (symbols) and simulated (dotted lines) without SHR G-R at constant OV=+3 V of device 2.

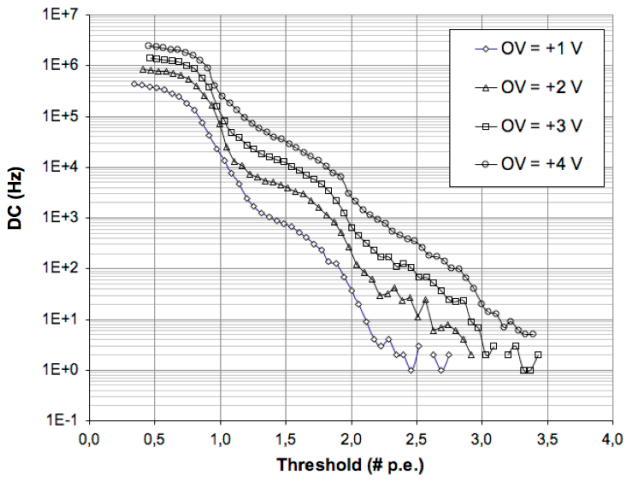

Figure 17. Measured DC at different OV of a 20×20 pixels SiPM as a function of the photoelectron signal amplitude threshold.

Fig. 17 shows the DC measured at room temperature for a 20×20 pixels SiPM (the pixel A_{ACT} is 40×40 µm^2) at different *OV*s, from +1V (diamonds) to +4V (circles) as a function of the normalized photo-electron threshold. The DC at 0.5 p.e. threshold level varies from 400 kHz to 3 MHz in the measured *OV* range. This value is roughly the *DC* rate due to only one pixel in breakdown, being the contribution of two or more pixel in breakdown at the same time to the *DC* rate at least one order of magnitude lower. In fact, at 1.5 p.e. threshold level (two pixels in breakdown simultaneously) the *DC* decreases of about three order of magnitudes at the lowest *OV* (~1 kHz at OV=+1V). The decrease is even more pronounced at 2.5 p.e (three pixels in breakdown simultaneously), being ~2 Hz at +1 *OV*.

The strong reduction in the *DC* value from 0.5 p.e. to 1.5 p.e. is due to different factors. First, the probability of simultaneous avalanche in two different pixels is lower than the probability of a single count. Moreover, the second pixel avalanche may be related to the first pixel one. In fact, there is a finite Cross Talk Probability (*CP*) for each device, strongly related to the array layout. The *CP* can be roughly quantified as:

$$CP = \frac{DC_{1.5}}{DC_{0.5}} \quad (5)$$

where $DC_{0.5}$ and $DC_{1.5}$ are the dark count rate at 0.5 and at 1.5 of the photoelectron signal level threshold. It should be stressed that using this approximation two pixels going in breakdown simultaneously are considered correlated.

The effect of the trench presence in the SiPM array is clearly visible by the inspection of Fig. 18 *a*. In fact, in figure the dark counts of two SiPM both having 20×20 pixels biased to the same *OV* (+2V) are compared. The red triangles are the data obtained from a SiPM with trenches, while the blue squares are from a SiPM without trenches around the pixels. Despite of the fact that the two devices have the same *DC* for single pixel breakdown, they strongly differ in the *DC* for two pixels in avalanche at the same time, being *CP* 0.7% for the device with trenches and 7% for the array without trenches. The trenches presence reduces the two pixels *DC* of one order of magnitude.

The *CP* probability was measured as a function of the *OV* for the two devices described before and the results are summarized in Fig. 18 *b*. The difference is one order of magnitude in the full range of operation. It could be inferred that the *CP* for the array with trenches we measured is actually the probability that two uncorrelated single events occur at the same time. All the devices discussed from now on are arrays with trenches.

B. Dark current in large device

The data so far shown in Section IV refers only to the best single pixels investigated. Since SiPM are an array of pixels, their performances are not exactly the best pixel performances multiplied by the number of pixels in the array. The dark currents can be worsened by the presence of randomly distributed defects that cause a distribution of performances around a mean value. The relationship among the dark currents in single pixels and in complete SiPM arrays can be summarized by the data (points) and simulations (lines) compared in Fig. 19 and already reported in [36]. We modeled the dark current of a single pixel as:

$$I_D = q\left(\frac{N_{Def}}{\tau} + \frac{A_{Pixel}}{\tau_i}\right)G \quad (6)$$

where q is the elementary charge, N_{Def} is the number of carrier generating defects per pixel in the active volume, τ is the average time for carrier generation event by one defect,

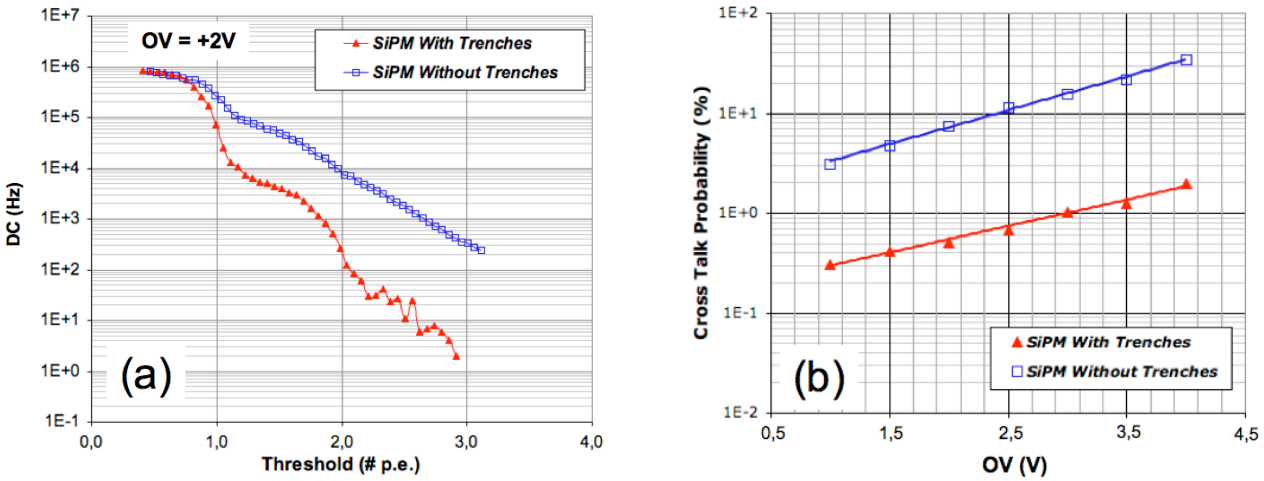

Figure 18. (a) Comparison of the dark count rate as a function of the photoelectron signal amplitude treshold and (b) of the cross talk probability vs. overvoltage of a 20×20 pixels SiPM with trenches (red closed symbols) and a 20×20 pixels SiPM without trenches (blue open symbols).

A_{Pixel} is the single pixel active area, τ_i is the average time per unit area for the intrinsic carrier generation due to diffusion from the quasi neutral regions to the active volume, and G is the gain. The I_D of the overall SiPM devices is simply the sum of the currents of single pixels as above modeled, with no contribution of extrinsic defects providing high leakage paths. In particular, Fig. 19 shows frequency histograms comparing the dark currents measured at room temperature of single pixels and SiPM arrays for a total of 952 devices at OV of 2, 3, and 4 V. The SiPM device contains 4096 pixels, so the respective currents of SiPM to single pixel should stay in ratio of about 4,000, as actually found.

To model I_D in the present devices, we should observe that the term N_{Def}/τ dominates at room temperature [36], hence the I_D statistics should essentially coincide with the N_{Def} statistics. The N_{Def} statistics is a Poisson statistics, hence the probability dP to have a DC between I_D and I_D+dI_D is:

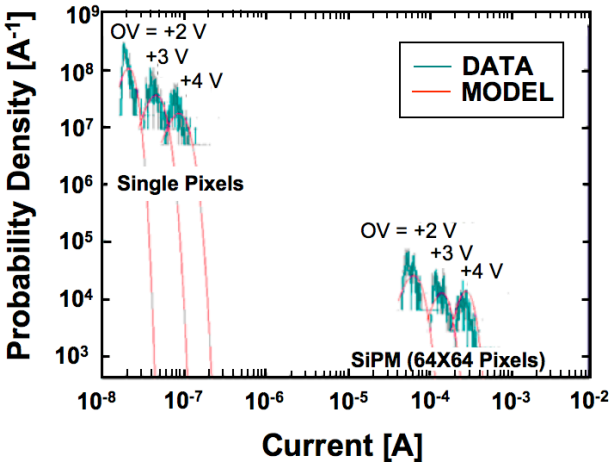

Figure 19. Probability density as a function of the output current at OV of 2V, 3V and 4V, for both single pixels and arrays (having 4096 pixels). The solid red lines are the model results.

$$\frac{dP}{dI_D} = N \exp\left[\frac{m_{I_D}}{\sigma^2_{I_D}}\left(I_D \log(m_{I_D}/I_D) + (I_D - m_{I_D})\right)\right] \quad (7)$$

where N is a normalization constant, m^2_{ID} is the statistical average of the dark current and σ^2_{ID} is the variance. In the case of the SiPM arrays the same expression holds. Fig. 19 reports also the model curves, which show a good match with the experimental data. The model predicts that the combination of statistical parameters σ^2_{ID}/m^2_{ID} should be equal to $q/\tau G$ or 4096× $q/\tau G$, for the single pixel and the SiPM array, respectively.

C. SiPM operation under illumination

Once defined the array performances in dark, measurements under a low illumination were carried out. The device was biased at +2V OV and the pulsed laser (6ns pulses) light was defocused in order to reduce the photon density. The device response is summarized in Fig. 20. In particular, Fig. 20*a* shows the persistence signal acquired on the OSC. The signal is due to the current spikes provided by one to 6 pixels (clearly identified) fired at the same time. More pixels have been fired by with a lower probability during the acquisition time (about 15 min). The signal width is limited by the oscilloscope resolution. This measurement can be made quantitative as shown in Fig. 20 *b*, where counts vs. the signal integrated charge (in 20 ns time range) is reported. The simultaneous firing of up to 8 pixels has been detected. The Gaussian distribution of each peak in Fig. 20 *b* is a clear evidence of the good pixel uniformity in terms of performances. Moreover, the distance between the peaks provide the information on the array gain [37]. It has been estimated as 10^6 at +2V OV. It should be stressed that the light signal, down to one photon count, has been acquired using only a digital oscilloscope, hence this device can be used without an external amplification circuit.

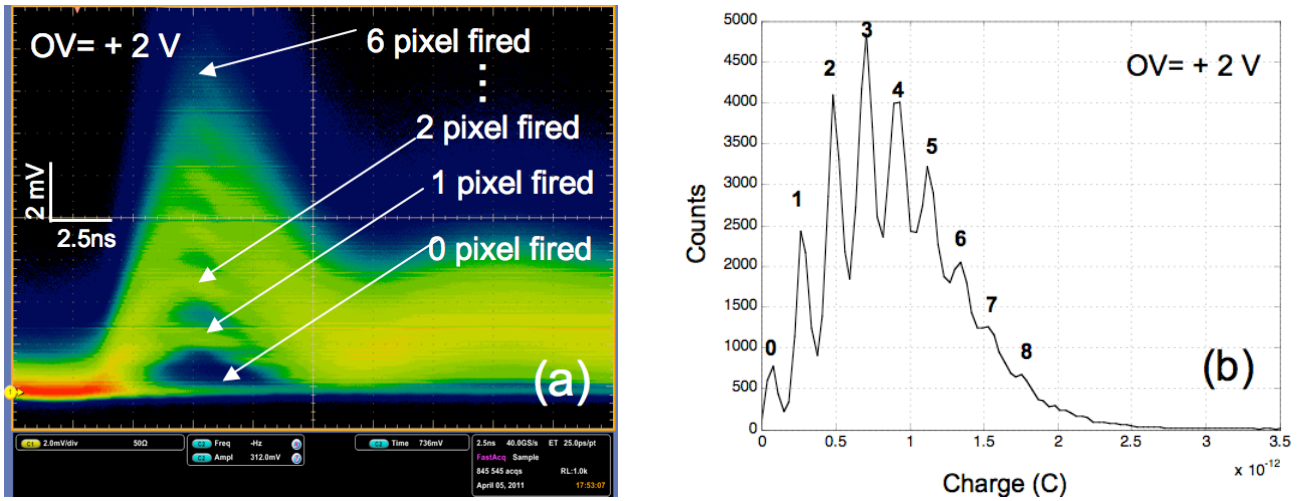

Figure 20. (a) Image of a persistence on a digital OSC and (b) charge distribution of the pulses a 20×20 pixels SiPM for a low intensity nano second laser light at OV=+2V.

VI. CONCLUSIONS

We reviewed our main results on the study of both single pixels and SiPM arrays produced by STMicroelectronics in Catania. Our data coupled to an extensive simulation study, show that the single pixel technology is close to its ultimate physical limit. The *DC* is dominated by diffusion of minority carriers from the cathode for temperatures down to 40°C. At lower temperatures, SHR generation is the main *DC* source. The only improvement in the single pixel technology could be provided by a further reduction in the defect concentration that, up to now, has been estimated to be ~ 1.6±1.3 defects/cm^{-3} in the best device. Obviously, the single pixel performances have a spread, due to the very low defect concentration needed to obtain the "best" device.

Not all the pixels forming the SiPM array are the "best" pixel, their dark current is distributed around a mean value. We found that it follows a Poissonian distribution perfectly mirroring the defect random distribution on the wafer. Hence, SiPM arrays exhibit performances worsened by the presence of defects placed according to a Poissonian distribution.

The presence of optical trenches surrounding each pixel strongly improves the SiPM performances, reducing the cross talk probability of one order of magnitude with respect to arrays without trenches.

The *DC* of SiPM arrays having the latest pixel design technology described at room temperature is in the order of 1 MHz/mm^2 at OV=+3V (~10%), close to that reported by other scientists. The *CP*, thanks to the fabrication of the trenches, is lower than 2% till OV=+4V (~15%), the lowest value, to our knowledge, reported in literature.

Finally, we have shown that these devices can be used as single photon counters also without a complex amplification stage.

ACKNOWLEDGMENT

This work has been partially funded by STMicroelectronics and by the national project MIUR-PON "Hyppocrates – Sviluppo di Micro e Nano-Tecnologie e Sistemi Avanzati per la Salute dell'uomo" (PON02 00355).

REFERENCES

[1] R. Pagano et al., "Improvement of the diffusive component of dark current in SiPM pixels", The Third International Conference on Sensor Device Technologies and Applications, SENSORDEVICES 2012, ISBN: 978-1-61208-208-0.

[2] M. D. Eisaman, J. Fan, A. Migdall, and S. V. Polyakov, "Invited Review Article: Single-photon sources and detectors", Rev. Sci. Instrum., vol. 82, no. 071101, 2011, pp. 1-25.

[3] D. Renker and E. Lorenz, "Advances in solid state state photon detectors", Journ. Instrum., vol. 4, no. P04004, 2009, pp. 1–56.

[4] T. Iijima, "Status and perspectives of vacuum-based photon detectors", Nucl. Instrum. Methods Phys. Res. A, vol. 639, no. 1, 2011, pp. 137–143.

[5] J. C. Campbell et al., "Recent Advances in Avalanche Photodiodes", IEEE Journ. Selec. Topic Quant. Electr., vol. 10, no. 4, 2004, pp. 777-788.

[6] G. Bondarenko, B. Dolgoshein, V. Golovin, Ilyin, R. Klanner, and E. Popova, "Limited Geiger-mode silicon photodiode with very high gain", Nucl. Phys. B Proc. Suppl., vol. 61 B, 1998, pp. 347-352.

[7] P. Buzhan et al., "Silicon photomultiplier and its possible applications," Nucl. Instrum. Meth. Phys. Res. A, vol. 504, no. 1–3, 2003, pp. 48–52.

[8] N. Otte et al., "The Potential of SiPM as Phototn Detector in Astroparticle Physics Experiments like MAGIC and EUSO", Nucl. Phys. B Proc. Suppl., vol. 150, 2006, pp. 144-149.

[9] V. D. Kovaltchouk, G.J. Lolos, Z. Papandreou, and K. Wolbaum, "Comparison of a silicon photomultiplier to a traditional vacuum photomultiplier" Nucl. Instrum. Meth. Phys. Res. A, vol. 538, no. 1–3, 2005, pp. 408–415.

[10] B. Dolgoshein, et al. "Status report on silicon photomultiplier development and its applications", Nucl. Instrum. Meth. Phys. Res. A, vol. 563, 2006, pp. 368-376.

[11] D. Renker, "Geiger-mode avalanche photodiodes, history, properties and problems", Nucl. Instrum. Meth. Phys. Res. A, vol. 567, no.1, 2006, pp. 48-56.

[12] S. Cova, A. Lacaita, and G. Ripamonti, "Trapping Phenomena in Avalanche Photodiodes on Nanosecond Scale" IEEE Electr. Dev. Lett., vol. 12, 1991, pp. 685-687.

[13] M. Ghioni, A. Gulinatti, I. Rech, F. Zappa, and S. Cova, "Progress in Silicon Single-Photon Avalanche Diodes", IEEE Jour. Sel. Top. Quant. Electr.,vol. 13, no. 4, 2007, pp. 852-862.

[14] B. F. Aull et al., "Geiger-Mode Avalanche Photodiode for Three Dimensional Imaging", Lincoln Lab. Jour., vol. 12, no. 2, 2002, pp. 335-350.

[15] E. Sciacca et al., "Silicon Planar Technology for Single-Photon Optical Detectors", IEEE Trans. Electr. Dev., vol. 50, 2003, pp. 918-925.

[16] F. Zappa, S. Tisa, A. Tosi, and S. Cova, "Principles and features of single-photon avalanche diode arrays" Sensors and Actuators A, vol. 140, 2007, pp. 103-112.

[17] W. G. Oldham, R. R. Samuelson, and P. Antognetti, "Triggering Phenomena in Avalanche Diode", ", IEEE Trans. Electr. Dev., vol. 19, no. 9, 1972, pp. 1056-1060.

[18] S. Cova, M. Ghioni, A. Lacaita, C. Samori, and F. Zappa, "Avalanche phototdiodes and quenching circuits for single-photon detection", Appl. Opt., vol. 35, no. 12, 1996, pp. 1956-1976.

[19] A. Lacaita, F. Zappa, S. Bigliardi, and M. Manfredi, "On the bremsstrahlung origin of hot-carrier-induced photons in silicon devices", IEEE Trans. Electr. Dev., vol. 40, no. 3, 1993, pp. 577-582.

[20] J. Briare and K. S. Krisch., "Principles of Substrate Crosstalk generation in CMOS circuits", IEEE Trans. Comp. Aided Des. Integr. Circ. And Sys., vol. 19, no. 6, 2000, pp. 645-653.

[21] O. Mineev et al.," Scintillator counters with multi-pixel avalanche photodiode readout for the ND280 detector of the T2K experiment", Nucl. Instrum. Meth. Phys. Res. A, vol. 577, no. 3, 2007, pp. 540-551.

[22] M. Danilov," Novel photo-detectors and photo-detector systems", Nucl. Instrum. Meth. Phys. Res. A, vol. 604, no. 1-2, 2009, pp. 183-189.

[23] M. Mazzillo et al., "Silicon Photomultiplier Technology st STMicroelectronics", IEEE Trans. Nucl. Sci., vol. 56, 2009, pp. 2434-2442.

[24] E. Sciacca et al., "Arrays of Geiger Mode Avalanche Photodiodes", IEEE Phot. Tech. Lett., vol. 18, no. 15, 2006, pp. 1633-1635.

[25] [Online 13.06.2013]. Available at http://www.silvaco.com.

[26] W. Shockley, "The Theory of p-n Junction in Semiconductors and p-n Junction Transistors", Bell Sys. Tech. Jour., vol. 48, 1949, pp. 435-489.

[27] W. N. Grant, "Electron and Hole Ionization Rates in Epitaxial Silicon at High Electric Fields", Solid-State Electr., vol. 16, 1973, pp. 1189-1203.

[28] R. Pagano et al., "Dark Current in Silicon Photomultiplier Pixels: Data and Model", IEEE Trans. Electr. Dev., vol. 59, no. 9, 2012, pp. 2410-2416.

[29] W. J. Kindt and H. W. van Zeijl, "Modelling and Fabrication of Geiger mode Avalanche Photodiodes", IEEE Trans. Nucl. Sci., vol. 45, no. 3, 1998, pp. 715-719.

[30] L. Passari and E. Susi, "Recombination mechanisms and doping density in silicon", J. Appl. Phys., vol. 54, no. 7, 1983, pp 3935-3937.

[31] A. Poyai, E. Simeon, C. Claeys, and A. Czerwinski, "Silicon substrate effects on the current-voltage characteristics of advanced p-n junction", Mat. Sci. Eng. B, vol. 73, no. 1-3, 2000, pp. 191-196.

[32] J. C. Jackson, P. K. Hurley, B. Lane, and A. Mathewson, "Comparing leakage currents and dark counts rate in Geiger-mode avalanche photodiodes", App. Phys. Lett., vol. 80, no. 22, 2002, pp. 4100-4102.

[33] M. Mazzillo et al., "Quantum Detection Efficiency In Geiger Mode Avalanche Photodiode," IEEE Trans. Nucl. Sci., vol. 55, no. 6, 2008, pp. 3620–3625.

[34] R. Pagano et al., "Silicon photomultiplier device architecture with dark current improved to the ultimate physical limit", App. Phys. Lett., vol. 102, no. 183502, 2013, pp. 1-4.

[35] P. Finocchiaro et al., "Characterization of a novel 100-channel silicon photomultiplier – part I: noise", IEEE Trans. Electr. Dev., vol. 55, no. 10, 2008, pp. 2757-2764.

[36] R. Pagano et al., "Silicon Photomultipliers: Dark Current and its Statistical Spread", Sensors & Transducers Journal, vol. 14, no. 1, 2012, pp. 151-159.

[37] P. Finocchiaro et al., "Characterization of a novel 100-channel silicon photomultiplier – part II: Charge and Time", IEEE Trans. Electr. Dev., vol. 55, no. 10, 2008, pp. 2765-2773.