Dark Count Rate Modeling in Single-Photon Avalanche Diodes

Aymeric Panglosse, Philippe Martin-Gonthier, Olivier Marcelot, Cédric Virmontois, Olivier Saint-Pé, Pierre Magnan

Dark Count Rate Modeling in Single-Photon Avalanche Diodes

Aymeric Panglosse, *Student Member, IEEE*, Philippe Martin-Gonthier, *Member, IEEE*, Olivier Marcelot, *Member, IEEE*, Cédric Virmontois, *Senior Member, IEEE*, Olivier Saint-Pé, *Member, IEEE*, and Pierre Magnan, *Member, IEEE*

***Abstract*— In this paper, we present a model to simulate accurately the Dark Count Rate (DCR) for Single-Photon Avalanche Diodes (SPAD) in Complementary Metal-Oxide Semiconductor (CMOS) technology. The model development has been driven by the necessity to comply with the specifications of SPAD used for future space LIDAR applications. To evaluate the DCR, the model is based on a combination of measurements to acquire data related to trap population, Technology Computer-Aided Design (TCAD) simulations and a Matlab routine.**

***Index Terms*— Single photon avalanche diode (SPAD), complementary metal-oxide semiconductor (CMOS), modelling and simulations, dark count rate (DCR), technology CAD (TCAD), Matlab.**

I. Introduction

IN THE field of space science, spaceborne LIDAR (Light Detection And Ranging) are active optical instruments embedded on-board satellites. These instruments are used for a multitude of applications and offer excellent measurement accuracy and altitude resolution. Numerous LIDAR types are employed to study atmosphere chemical content, to compute the wind velocity or the distance between the satellite and the scene. To probe a scene, a laser emits an optical pulse, and the reflected and back-scattered light is analysed. LIDAR require photo-detectors with low photon flux detection capability, good time resolution and a small dead-time, defined as the minimum time delay between two detection events. Moreover, space applications add some constraints, for instance, the power is limited on board the satellites or probing with high accuracy is needed, at significant distance from the emitter. In contrast to other photo-detectors, SPAD fulfill these requirements. Their high sensitivity allows very low flux detection down to photon counting, reducing the readout noise of the sensor and allowing lower illumination and power consumption. Furthermore, SPAD offer low dead-time, the best ones performing under ten nanoseconds [1]. High apertures of optical systems require minimal optical spots around $150 - 200$ µm, consequently the integration into the space instrument requires large SPAD.

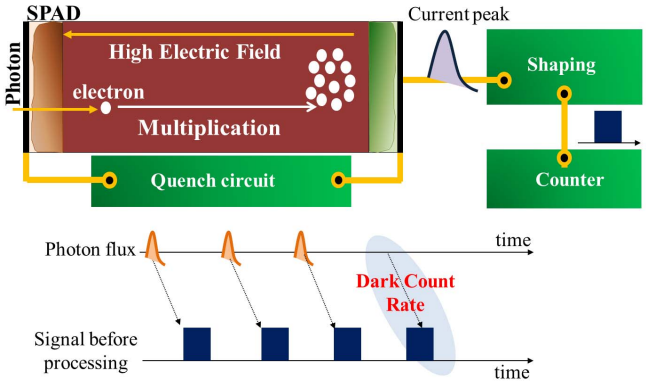

Fig. 1. Simplified SPAD schematic with its electronic and DCR illustration. The holes are not included in this schematic and were omitted for clarity.

SPAD is biased beyond its breakdown voltage in the so-called Geiger-mode. A high electric field generated by the bias voltage, accelerates the photo-carriers which are then multiplied by impact ionization. This multiplication leads to an electronic avalanche and the photodiode breakdown, characterized by a rapid current increase. A high impedance path is required to quench the SPAD and to avoid the photodiode destruction. The current peak is then shaped and processed. The pulse processing depends on the targeted application, for example, in the case of photon counting, each pulse increments a counter. Our study focuses on one of the SPAD performance parameter which is used to evaluate the SPAD quality: the Dark Count Rate (DCR). In the dark, charges can be generated by thermal or tunneling effects within, or around, the high field region leading to an unwanted avalanche and a false count called dark count. False counts can also be generated by the afterpulsing phenomenon: charges captured by deep level traps in the junction depletion layer are released within the high field region with a fluctuating delay. This phenomenon is characterized by its probability of occurrence and characteristic time distribution. Figure 1 shows a simplified SPAD schematic with its electronic and DCR phenomenon illustration.

SPAD fabrication in CMOS process enables a high integration density with the possibility to combine the detector

This work was supported by CNES and Airbus Defence and Space. This article was recommended by Associate Editor B. Gosselin. *(Corresponding author: Aymeric Panglosse.)*

Aymeric Panglosse is with the Department of Electronic Optic and Signal, ISAE-Supaero, University of Toulouse, 31055 Toulouse, France (e-mail: aymeric.panglosse@isae-supaero.fr).

Philippe Martin-Gonthier, Olivier Marcelot, and Pierre Magnan are with ISAE-Supaero, University of Toulouse, 31055 Toulouse, France.

Cédric Virmontois is with the French National Space Agency, CNES, 31400 Toulouse, France.

Olivier Saint-Pé is with Airbus Defence and Space, 31400 Toulouse, France.

Digital Object Identifier 10.1109/TCSI.2020.2971108

and the driving electronics on the same chip [2]. The CMOS technology is mature and many technology nodes have been integrated with success, from 0.8 µm to an advanced 40 nm technology [3] and 28 nm using fully depleted silicon on insulator (FDSOI) CMOS technology [4]. Analysis and simulation are fundamental steps in CMOS process work-flow, to study and predict performance parameters as design, fabrication and characterization are time consuming and costly. The SPAD performance parameters are not directly simulated by commercial TCAD simulation tools. Nonetheless, TCAD tools simulate accurately electrical and physical parameters as the electric field or the breakdown voltage and its outputs can be used to compute the performance parameters. In the literature, some models aim to evaluate performance parameters but are based on theoretical calculations to evaluate electrical parameters and/or do not completely consider the CMOS process influence as in [5], or as in [6] where band-to-band tunneling is not modeled. Some others are based on TCAD, using its outputs and analytical or numerical model to obtain the parameters. The process is entirely generated and simulated, based on literature data and the simulation tool is calibrated by adjusting implantation conditions with literature breakdown voltages [7]. This method, although efficient, raises questions on the accuracy of process quality simulation and trap information acquisition. The main difficulty to model accurately the DCR and the afterpulsing probability is to obtain data about the process quality and physical parameters related to the targeted process as the trap energies or the band-to-band tunneling parameters. These data can only be obtain by measurements. The innovative aspect of this work consists in developing a model using TCAD simulations and measurements performed on structures of the targeted CMOS process which bring information related to trap population. This approach allows to get as close as possible to the reality of the process and to improve prediction. Furthermore, the model should be able to predict the DCR of large SPAD used in space LIDAR. In the following parts, we will give a detailed description of the model in section II, in section III we describe the set of SPAD used for the measurements and model validation, and finally, section IV presents simulation and measurement results with their interpretations.

II. Dark Count Rate Model Description

The SPAD DCR modelling necessitates accurate electric field evaluation, which requires the effective doping profiles of the targeted SPAD process. In this work, a commercial 130 nm technology is used. An overview of the model is illustrated in figure 2 and the SPAD layout will be fully explained in part III. By combining experimental doping profiles measured on dedicated structures and the SPAD architecture (layout), the TCAD software rebuilds the doping profile in two dimensions. Then the tool Sentaurus Device provides the electrical parameters. Among them, the electric field profile and the depletion region width are needed for the DCR calculation. Further details about simulation conditions will be given in parts III and IV. Subsequently, these parameters are injected in a Matlab routine with the addition of measurements

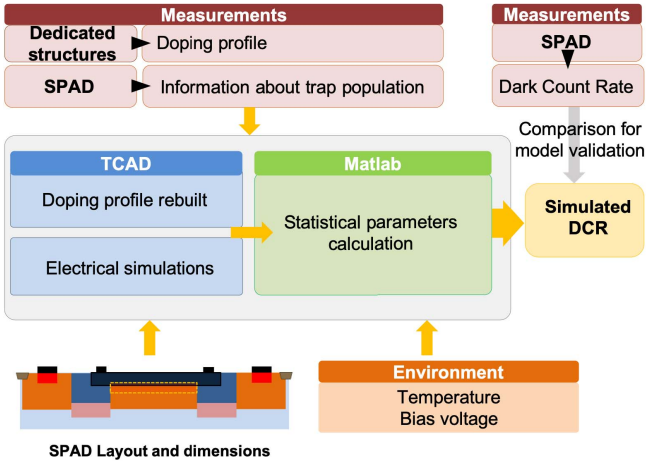

Fig. 2. Overview of the modelling and simulation method to evaluate SPAD DCR.

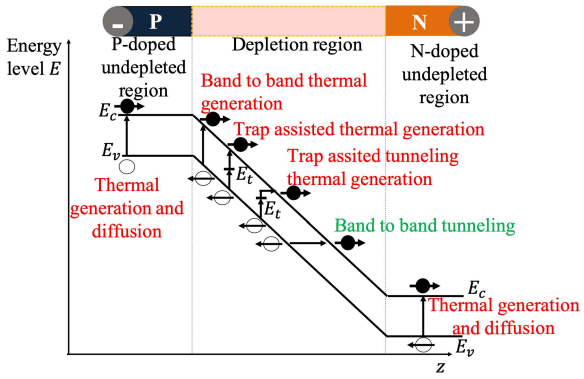

Fig. 3. DCR contributions illustration for a reverse biased p-n junction. Setting aside the afterpulsing contribution, DCR is composed of thermal and band-to-band tunneling contributions.

on the targeted CMOS process, SPAD size and information related to the environment as temperature or bias voltage. With analytical and numerical methods, the Matlab routine evaluates key statistical parameters: ionization coefficients, generation rates or avalanche triggering probabilities. Afterward, these parameters are used to calculate the DCR for any temperature, SPAD dimension and architecture.

Setting aside the afterpulsing contribution, the DCR consists of two contributions illustrated in figure 3:

- a thermal contribution, which relates to charges thermally released. These charges could be generated directly within the high field depletion region or within the undepleted region and diffuse to the depletion region to initiate an electronic avalanche. In the case of our SPAD, we observed that the thermal generation within the undepleted regions has a minor influence on the DCR. Nevertheless, it might not be the case of all SPAD. Within the depletion region, band-to-band thermal generation is unlikely due to the bandgap energy of 1.2 eV. The main thermal contribution comes from thermal generation assisted by trap and thermal generation with trap assisted tunneling. These contributions can be modeled by the Shockley-Read-Hall (SRH) with trap-assisted-tunneling (TAT) generation rate. The thermal contribution has a

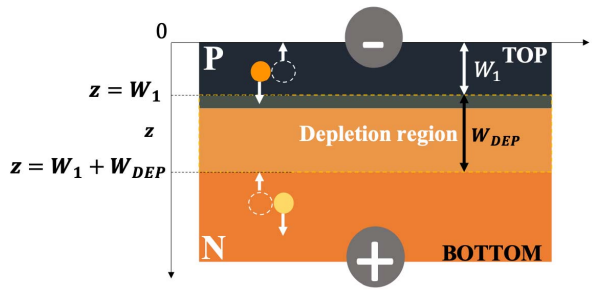

Fig. 4. Cross section for DCR modelling. TOP region corresponds to the undepleted region on top of the depletion region, and BOTTOM region stands for the undepleted region below the depletion region.

strong dependence with lattice defect within the diode and decrease greatly with temperature;

- a band-to-band tunneling contribution. Due to the high electric field of the depletion region, charges are transferred from the valence band to the conduction band by tunneling effect. This contribution does not vary to any great extent with temperature. At low temperature, it is the main DCR contribution, as the thermal contribution is negligible. This contribution can be modeled by the band-to-band-tunneling (BTBT) generation rate.

Dark count rate is expressed in count per second (cps). In our model, DCR will be evaluated for each region of the cross-section illustrated in figure 4. TOP region corresponds to the undepleted region on top of the depletion region, and BOTTOM region stands for the undepleted region below the depletion region. The DCR of the top, depletion and bottom region are respectively named DCR_{Top}, DCR_{Dep}, DCR_{Bot}:

$$DCR = DCR_{Top} + DCR_{Dep} + DCR_{Bot}. \tag{1}$$

A. Avalanche Triggering Probability Modelling

Avalanche triggering probability is required to model the DCR. Our TCAD tool does not implement this parameter calculation, so it will be implemented externally using a Matlab routine. The expression of this probability depends on the model used to describe the impact ionization effect. In our case, we use the McIntyre local model [8]. This model assumes that the ionization rates α_n and α_p, which are defined as the mean number of charges generated by distance unit travelled by an electron or a hole, depend only on the local electric field. The McIntyre local model use has been motivated by the good agreement between the simulated and the measured breakdown voltage of our SPAD presented in part III. α_n and α_p are computed using the electric field profile provided by TCAD and the Van Overstraeten and De Man coefficients [9].

The probability P_p that either an electron or a hole triggers an avalanche is given by equation 2 [10]:

$$P_p(z) = P_{pe}(z) + P_{ph}(z) - P_{pe}(z)\,P_{ph}(z), \tag{2}$$

with $P_{pe}(z)$ and $P_{ph}(z)$ the probability to trigger an avalanche, for respectively an electron and a hole, starting from the position z within the depletion region. These probabilities can be found by solving the following system, with the initial conditions $P_{pe}(W_1 + W_{Dep}) = 0$ and $P_{ph}(W_1) = 0$:

$$\frac{dP_{pe}}{dz} = (1 - P_{pe})\,\alpha_n\left(P_{pe} + P_{ph} - P_{pe}P_{ph}\right), \tag{3}$$

$$\frac{dP_{ph}}{dz} = -(1 - P_{ph})\,\alpha_p\left(P_{pe} + P_{ph} - P_{pe}P_{ph}\right). \tag{4}$$

This system of equations needs to be solved numerically. In our case, the Matlab routine is used.

B. Dark Count Rate Modelling in the Undepleted Top Region

A charge generated within the undepleted top region must diffuse to the upper limit of the depletion region without any recombination to trigger an avalanche. This avalanche will be triggered with the probability $P_e = P_{pe}[W_1]$. The DCR, expressed per unit area, is:

$$DCR_{Top} = P_e \times \frac{J_n[W_1]}{q}, \tag{5}$$

with $J_n[W_1]$, the electron current density at the upper limit of the depletion region and q the elementary charge. The modelling of the current density is introduced below.

C. Dark Count Rate Modelling in the Undepleted Bottom Region

Symmetrically, the DCR, per unit area, within the undepleted bottom region is given by:

$$DCR_{Bot} = P_h \times \frac{J_p\left[W_1 + W_{Dep}\right]}{q}, \tag{6}$$

with $J_p\left[W_1 + W_{Dep}\right]$ and $P_h = P_{ph}\left[W_1 + W_{Dep}\right]$, respectively the hole current density and the probability for a hole to trigger an avalanche at the lower limit of the depletion region.

Current density J_n and J_p are used in equations 5 and 6 to evaluate dark count from the undepleted regions at the upper and lower limits of the depletion region. These densities are driven by the continuity equation:

$$\begin{cases} \nabla.\overrightarrow{J_n} = qG_{net} + q\dfrac{\partial n}{\partial t} \\ -\nabla.\overrightarrow{J_p} = qG_{net} + q\dfrac{\partial p}{\partial t}, \end{cases} \tag{7}$$

with n and p, respectively the electron and hole density. G_{net} is the electron-hole generation rate:

$$G_{net} = G_{SRH,TAT} + G_{BTBT} + G_{ionization}. \tag{8}$$

With $G_{SRH,TAT}$ and G_{BTBT} respectively the Shockley-Read-Hall (SRH) with trap-assisted-tunneling (TAT) and the band-to-band-tunneling (BTBT) generation rates. These rates modelling are given in part II-E and II-F. $G_{ionization}$ is the generation rate due to ionization effect:

$$G_{Ionisation} = \alpha_n n v_n + \alpha_p p v_p, \tag{9}$$

with v_n and v_p respectively the electron and hole recombination velocity. In silicon $v_n = 2.573 \times 10^6\ \mathrm{cm \cdot s^{-1}}$ and $v_p = 1.93 \times 10^6\ \mathrm{cm \cdot s^{-1}}$.

Inside the undepleted regions, as there is negligible electric field, the main contribution to the current density is thermal, leading to weak current. The DCR of these regions depends

on the volume of said region and SPAD with very narrow undepleted regions have less contribution, to the extent that, some modelling method does not take it into account [2], [7]. Other models as Kindt [11] incorporate this contribution to its model based on theoretical structures leading to simplified equation, which can be solved analytically. In our case, the continuity equation has been solved using the TCAD tool fed by doping profile measurements to be as close as possible to the reality of the process and improve prediction.

Accurate current density modelling is also important as we use it to evaluate the breakdown voltage with the TCAD tool. This evaluation allows to gauge if the chosen ionization coefficients and the model is accurate enough to describe the avalanche phenomenon inside the device.

D. Dark Count Rate Modelling Within the Depletion Region

A charge generated by thermal or tunneling effect directly within the depletion region at position z triggers an avalanche and a dark count with probability $P_p(z)$.Then, the DCR is:

$$DCR_{Dep} = DCR_{SRH,TAT} + DCR_{BTBT}, \tag{10}$$

with:

$$DCR_{SRH,TAT} = \int_{W_1}^{W_1+W_{Dep}} P_p\left(z\right) G_{SRH,TAT}(z)dz, \tag{11}$$

and:

$$DCR_{BTBT} = \int_{W_1}^{W_1+W_{Dep}} P_p\left(z\right) G_{BTBT}(z)dz. \tag{12}$$

E. Shockley-Read-Hall With Trap-Assisted-Tunneling Generation Rate Modelling

Shockley-Read-Hall with trap-assisted-tunneling generation rate is given by:

$$G_{SRH,TAT} = \frac{np - n_{i,eff}^2}{\tau_p\left(n + n_1\right) + \tau_n\left(p + p1\right)}, \tag{13}$$

with:

$$n_1 = n_{i,eff} \exp\left[\frac{E_{trap}}{kT}\right], \tag{14}$$

and:

$$p_1 = n_{i,eff} \exp\left[\frac{-E_{trap}}{kT}\right]. \tag{15}$$

$E_{trap} = E_t - E_i$, with E_t the energy level of the dominant generation center and E_i the intrinsic Fermi level. Then the energy E_{trap} is not a function of z [6], and will be used as a fitting parameter as explained in section II-G. In our case, we estimate only one level of trap energy, because of the good correlation discussed in section IV, however some devices could show several distinct trap energies. τ_n and τ_p are minority carriers lifetimes and the modelling of these parameters will be given in the next paragraph. k is the Boltzmann constant and T, the temperature.

$n_{i,eff}$ is the effective intrinsic density:

$$n_{i,eff} = n_i \exp\left(\frac{E_{bgn}}{2kT}\right), \tag{16}$$

TABLE I

PARAMETERS USED IN THE DOPING LEVEL AND TEMPERATURE DEPENDENCE FOR MINORITY CARRIERS LIFETIME MODELLING [20]

Parameter	Electrons	Hole	Unit
τ_{min}	0	0	sec
τ_{max}	1×10^{-5}	3×10^{-6}	sec
N_{ref}	1×10^{16}	3×10^{16}	cm^{-3}
γ	1	1	1
T_α	-1.5	-1.5	1

with E_{bgn} the bandgap narrowing energy. The Bennett-Wilson model will be used for this parameter [12]. n_i is the intrinsic density and depends on the temperature:

$$n_i\left(T\right) = \sqrt{N_c\left(T\right) N_V\left(T\right)} \exp\left(\frac{-E_g\left(T\right)}{2k_B T}\right), \tag{17}$$

with $E_g\left(T\right)$ the bandgap energy as a function of the temperature given in [13], [14].

Minority carriers lifetime is given by:

$$\tau_{n,p} = \tau_{dop} \frac{f\left(T\right)}{1 + \Gamma_{TAT}\left(F\right)}. \tag{18}$$

$\tau_{n,p}$ has a doping level dependence via τ_{Dop}, a temperature dependence via $f(T)$ and an electric field dependence via the electric field enhancement parameter $\Gamma_{TAT}\left(F\right)$ used for the high electric field F, which models the trap assisted tunneling phenomenon with the Hurkx model [15].

Doping level dependence is modelled by the Scharfetter relation. This relation is based on experimental data and the theoretical conclusion that trap density of fundamental acceptor-type defects, such as divacancy or complex vacancy obtained with ionic implantation, is strongly correlated to the doping density [16]–[18]:

$$\tau_{dop}\left(N_{A,0} + N_{D,0}\right) = \tau_{min} + \frac{\tau_{max} - \tau_{min}}{1 + \left(\frac{N_{A,0}+N_{D,0}}{N_{ref}}\right)^{\gamma}}. \tag{19}$$

$N_{A,0}$ and $N_{D,0}$ are chemically active acceptor concentrations computed using the TCAD. The other parameters are given in table I.

Temperature dependence of minority carriers lifetime is given by the Schenk relation [19]. Similarly to the Scharfetter relation, this relation arises from a model for which the parameters have been fitted based on experimental data:

$$f\left(T\right) = \left(\frac{T}{T_0}\right)^{T_\alpha}, \tag{20}$$

with $T_0 = 300$ K and the parameter T_α is given in table I.

F. Band-to-Band Tunneling Generation Rate Modelling

The generation rate for the band-to-band tunneling phenomena is given by the Hurkx tunneling model [15]:

$$G_{BTBT} = B|F|^{\frac{5}{2}} \exp\left(\frac{-F_{0T}}{F}\right), \tag{21}$$

with:

$$F_{0T} = F_0 \left(\frac{E_g(T)}{E_{g300}}\right)^{\frac{3}{2}}. \tag{22}$$

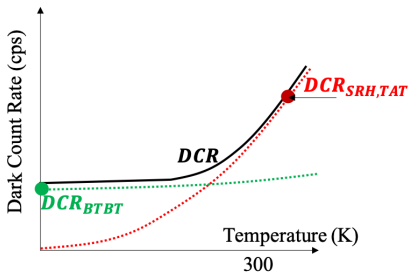

Fig. 5. Theoretical evolution of thermal and tunneling contributions to the DCR as a function of temperature.

$B = 4 \times 10^{14}$ cm$^{-0.5}$V$^{-2.5}$ s^{-1} is a constant based on the carriers' effective mass and electron-phonon interaction. $E_g(T)$ is the bandgap energy at the temperature T and E_{g300} the bandgap energy at 300K. F_0 is a fitting parameter.

G. Thermal and Tunneling Contribution Separation

Figure 5 shows the theoretical evolution of thermal and tunneling contributions to the DCR as a function of the temperature.

As mentioned hereinbefore, thermal contributions decrease greatly with temperature drop. Hence, at low temperature, the band-to-band tunneling contribution will be dominant and a change in slope is observed.

By taking a sufficiently low temperature point, the DCR will only be dominated by the tunneling contribution ($DCR_{SRH,TAT} \ll DCR_{BTBT}$). Consequently, performing a DCR measurement at low temperature and using F_0 as a fitting parameter gives the tunnelling contribution DCR_{BTBT}. At first glance, it is difficult to choose a temperature point, nevertheless, most of the literature studies on this issue, with various CMOS processes, show that a temperature of 30 K below ambient temperature is sufficient to reach the regime dominated by tunneling [7], [21], [22]. In our case, we performed measurements down to 150 K in order to work in the tunneling mode.

Once the band-to-band tunneling contribution is obtained, the thermal contribution can be obtained at higher temperature, in our case at 325K, where DCR is dominated by thermal contribution. The energy E_t of the dominant generation center is used as a fitting parameter to compute $DCR_{SRH,TAT}$.

H. Model Implementation

Figure 6 shows the DCR modelling implementation.

Doping profile measurements performed on dedicated structures and the SPAD layout were injected in the TCAD tool which rebuilt the doping profile in two dimensions and ran the electrical simulation. The TCAD tool evaluates current density, charge concentrations and lifetime, electric field profile and depletion region width. These data flow into the Matlab routine which calculates statistical parameters such as the ionization coefficients and the avalanche triggering probabilities by solving the system of equations 3 and 4 numerically. Finally, the generation rate G_{BTBT} and $G_{SRH,TAT}$ are computed, followed by both DCR contributions DCR_{BTBT} and $DCR_{SRH,TAT}$, using the two DCR measurements at low and ambient temperature to fit F_0 and E_t. External contributions such as temperature and bias voltage

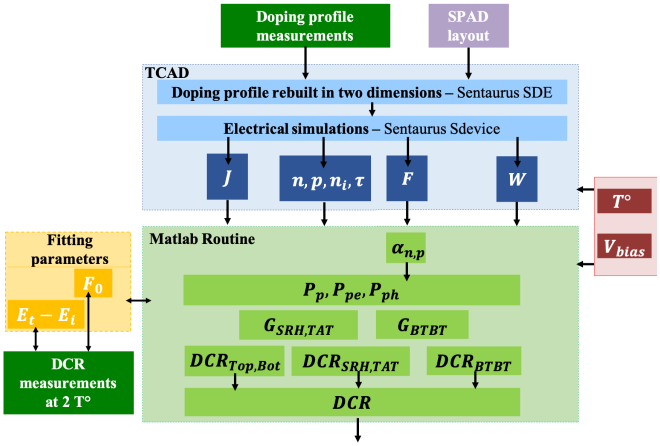

Fig. 6. DCR modelling based on TCAD, Matlab routine, measurements and external data.

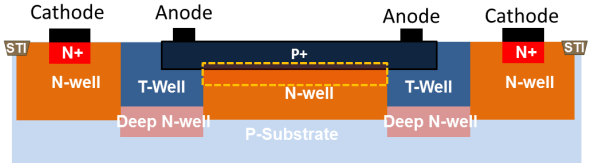

Fig. 7. Cross section of the SPAD under investigation.

are also used to feed the TCAD tool and Matlab routine. DCR is given per unit area; to obtain the SPAD DCR for a given area S, the DCR per unit area has to be multiplied by area S. Thereby, the DCR could be computed for any SPAD area, including large devices used in space LIDAR applications.

$$DCR_{Total} = S \times DCR. \tag{23}$$

III. SPAD Description and Model Adaptation

The set of SPAD used to validate the DCR model was fabricated using a mixed signal 130 nm CMOS process. A cross section is presented in figure 7 and shows a typical SPAD architecture for deep sub-micron CMOS process. The avalanche multiplication region is formed by the P+/N-Well junction. To avoid Premature Edge Breakdown (PEB), a P-guard-ring made with T-Well is used on top of the Deep N-Well (DNW) implant. The SPAD is circular to avoid high electric field concentration in the angles and curvatures [23] and thus another source of premature breakdown. Moreover the circular shape allows to perform TCAD simulations of a 3D device by making use of the rotational symmetry. The Deep N-Well allows contact between the two external N-well implants and the central N-well region. Also, the Deep N-Well and N-Well implants isolate the SPAD from charges that could be generated by the surrounding electronics.

The SPAD design and fabrication was done without the precise knowledge of the doping profile. The post-fabrication TCAD simulation fed by doping profile measurements shows an undesired electric field profile due to a high doping level of the T-Well and N-Well implants. Our SPAD were integrated onto a multi-project wafer (MPW), which imposed some restrictions on the implants placement. In particular, the T-Well implants must be generated on top of the Deep N-Well and are more doped than we expected. Consequently, a large

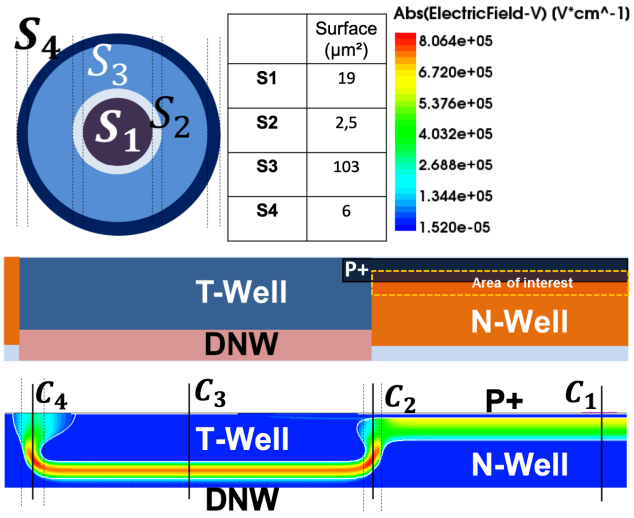

Fig. 8. Electric field profile of the SPAD investigated with the four zones considered to model the device. The table include the areas of each zones.

PEB is observed within the guard ring, causing a high DCR, as presented hereafter. This, has to be taken into account in the measurements analysis.

The SPAD is split into four zones with four surfaces presented in figure 8. The high electric field within the guard ring tends to indicate that the rings enter breakdown early, these observations will be discussed further in section IV. Consequently, it is highly probable that the region of interest for our SPAD does not enter breakdown and is rather at the limit of a linear avalanche diode profile. Consequently, it is important to note that the model will be validated against dark counts coming from the guard ring, for the most part, and not the area of interest of our SPAD.

A cross-section is made in each zone, showed in figure 8, in which the DCR per unit area is evaluated using equation 1. The result is then extended to the whole surface by multiplying this data by the surface of the considered zone. The total DCR is the sum of the rates computed for each area and is given in count per second (cps).

A passive quench circuit composed of an external resistance was used to perform DCR measurements. Connections to this resistance add parasitic capacitances leading to high recharge times, typically more than 1 µs. With this high time constant, experimental data [2], [21], [24] show that the measured DCR can be considered free from afterpulsing contributions, since the trap lifetime is generally much smaller than 1 µs.

This dead time τ_m also hides some events related to thermal and tunneling effects. The relationship between the measured DCR DCR_{meas} and the corrected DCR DCR_{cor}, including thermal and tunnelling contributions hidden by the dead time is [21]:

$$DCR_{cor} = DCR_{meas} + DCR_{meas}^2 \times \tau_m. \quad (24)$$

IV. Simulation and Measurement Results and Interpretations

DCR measurements are done on two SPAD with the same design coming from two different dies, and the mean DCR at 150 K and 325 K are selected to evaluate the fitting

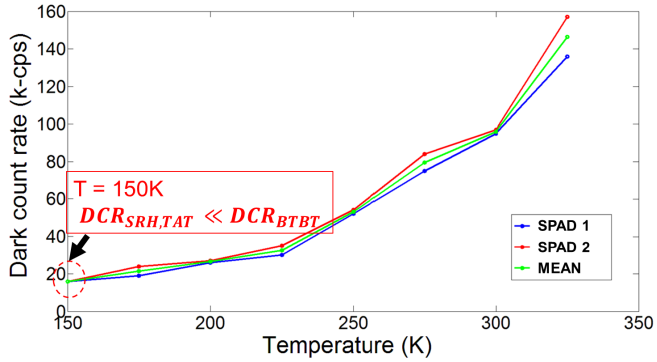

Fig. 9. Measured DCR as a function of temperature on two identical SPAD of two different chips. The mean value at 150 K is used to compute the fitting parameter for the BTBT modelling.

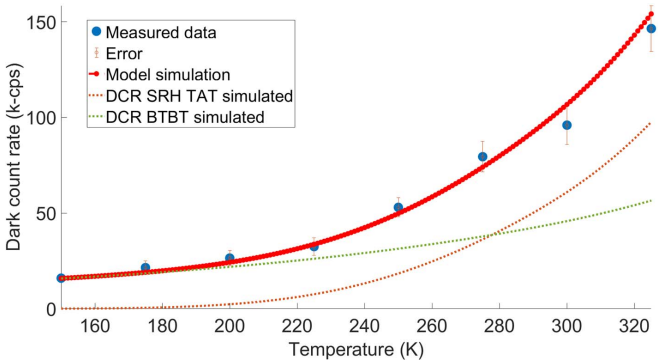

Fig. 10. Measured and simulated DCR as a function of temperature. The simulated DCR shows good agreement with the measured values. The error bar for measured data is displayed in thin dashed line.

parameters, as shown in figure 9. The values obtained are $E_{trap} = 0.19$ eV and $F_0 = 2.2 \times 10^7$ V·cm^{-1}.

These values are credible since, after identification using [25], E_{trap} value corresponds to the neutral divacancy defect family, which is commonly observed after ionic implantation steps. It corresponds as well, to a lesser extent, to the vacancy + phosphorus defect family. This is also consistent, as the main noise source comes from the guard ring and a large part of its depletion region is located within the Deep N-Well implant doped with phosphorus, in the same way as the N-Well implant. The value of F_0 is in good agreement with the range observed in literature where $1.9 \times 10^7 < F_0 < 3 \times 10^7$ V·cm^{-1} [26].

The measured breakdown voltage is 9.4 V and the measurements were done with an excess voltage above breakdown of 300 mV, which represents approximately 3 % of the breakdown voltage value. This ratio was used for all the following simulations. Above this value, it is difficult to measure DCR around ambient temperature, the high count rates and dead time causing the SPAD count rate to saturate.

Figure 10 depicts the measured and simulated DCR as a function of temperature, with each separate contribution. Contributions from top and bottom regions are not displayed since they are negligible, both in the guard ring region and in the area of interest, according to our model prediction. Figure 11 displays the results in logarithmic scale, so it is more convenient to observe the separation between thermal and tunneling contribution at low temperature.

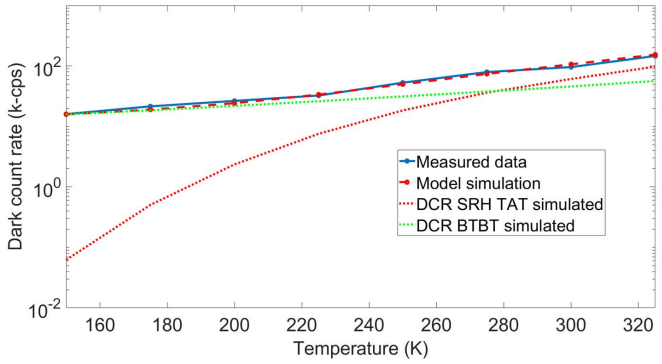

Fig. 11. Simulated DCR as a function of temperature in logarithmic scale. The separation between the two contributions is clearly demonstrated at low temperature. The error bar for measured data is not displayed for clarity.

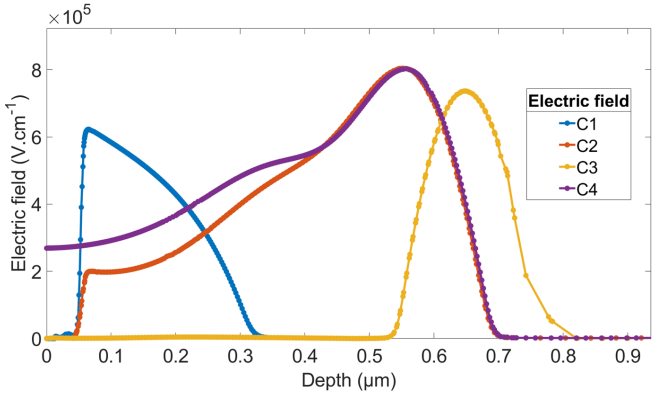

Fig. 12. Electric field profile for each zone. The high DCR is mainly due to high electric field leading to premature edge breakdown within the guard ring.

The DCR goes from 15 kcps at 150 K to 250 kcps at 325 K. The Shockley-Read-Hall with trap-assisted-tunneling and band-to-band tunneling contributions are also plotted and as expected, at low temperature (where $DCR_{BTBT} \simeq 15$ kcps and $DCR_{SRH,TAT} \simeq 0.1$ kcps), the DCR is dominated by band to band tunneling. At higher temperature, (where 325 K, $DCR_{SRH,TAT} \simeq 90$ kcps and $DCR_{BTBT} \simeq 40$ kcps), the DCR is dominated by thermal contribution. Figures 10 and 11 shows very good prediction from the model since the measurements and simulation results are very close, indeed the simulated DCR shows good agreement with the measurements with a mean error less than 10 %.

For each DCR measurement at one specific temperature, we performed multiple acquisitions with an oscilloscope to compute the mean count. Then, the measurement errors, displayed on figure 10, represent the standard deviation obtained. Therefore, our error bar only takes into account the the statistical variations in the count measurements.

Compared to other SPAD DCR in literature, in the same technology node [27], the DCR is very high and it is due to premature edge breakdown, as mentioned previously.

Figure 12 shows the electric field profile for each zone. As anticipated, the guard ring zones presents the highest electric fields, in particular the corners (C2 and C4, figure 8).

Figure 13 and 14 illustrate respectively the Shockley-Read-Hall with trap assisted tunneling $G_{SRH,TAT}$ and band-to-band tunneling G_{BTBT} generation rates. In both generation rates,

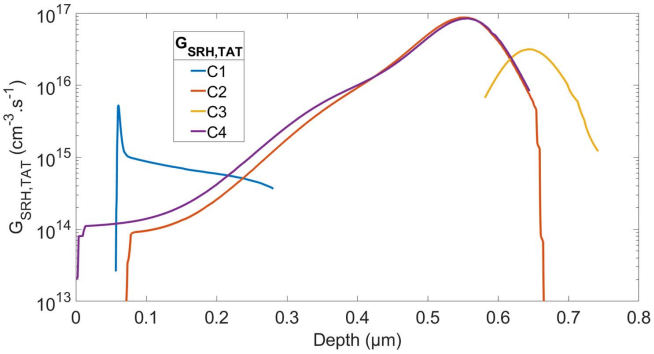

Fig. 13. Shockley-Read-Hall generation rate for each zone, at $T = 300$ K. The trap assisted tunneling effect leads to high generation rate inside the high electric field zones.

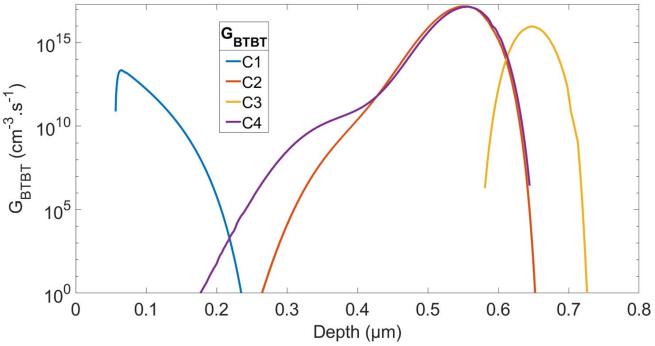

Fig. 14. Band-to-band tunneling generation rate for each zones, at $T = 300$ K. High tunneling rates are observed inside the guard ring.

we observed that tunneling effects have a major influence since the highest values are found in the highest electric field regions. Concerning $G_{SRH,TAT}$, the high electric field leads to high trap assisted tunneling within the corners regions 2 and 4, and similarly, charges generation by band-to-band tunneling will reach its maximum value in these regions.

This high electric field has an impact on ionization coefficient as they rely directly on its value with an exponential dependency. Consequently, it will affect the avalanche triggering probabilities. Figures 15 and 16 show the avalanche triggering probability for a electron-hole pair P_p, an electron P_{pe}, or a hole P_{ph}, starting at position z within the depletion region, for each zones. For an electron, this probability reaches its maximum value at the upper limit of the depletion region, as an electron generated at this position has the full depletion region width to diffuse, and create an avalanche. Most of ionization will occur in this side of the depletion region, especially since the electric field will be at its maximum value around this upper limit. The deeper the electron is generated in the depletion region, its probability to cause an avalanche decreases since both the electric field and the distance to travel to the lower bound is reduced. Near the lower boundary, carriers will drift with a low probability to ionize. The same reasoning applies for holes. P_{pe} is higher than P_{ph} since electrons ionization is higher than holes in silicon. As we can notice, avalanche triggering probabilities are much higher for zones C2 and C4 than region C1. With a weak excess bias voltage of 3 % of the breakdown voltage, avalanche triggering probabilities reach almost 100 % for

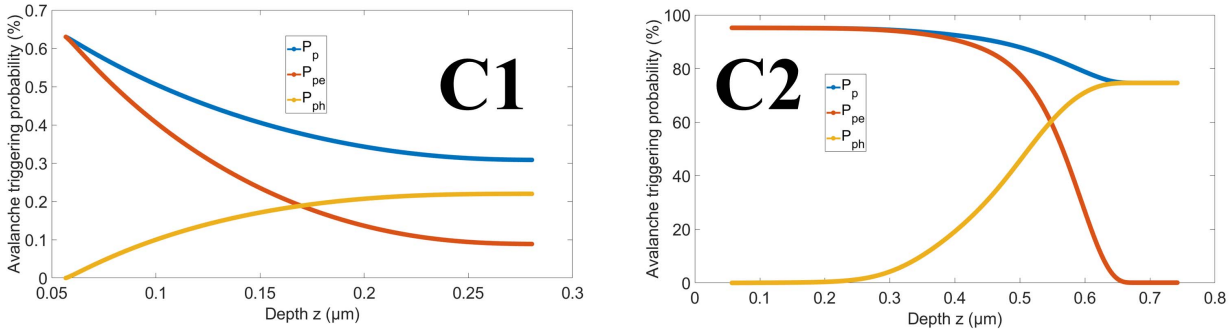

Fig. 15. Avalanche triggering probability for a electron-hole pair P_p, an electron P_{pe} or a hole P_{ph} starting at position z within the depletion region, for cross-section C1 and C2. Van Overstraeten and De Man ionization coefficients were used [9], at T = 300 K.

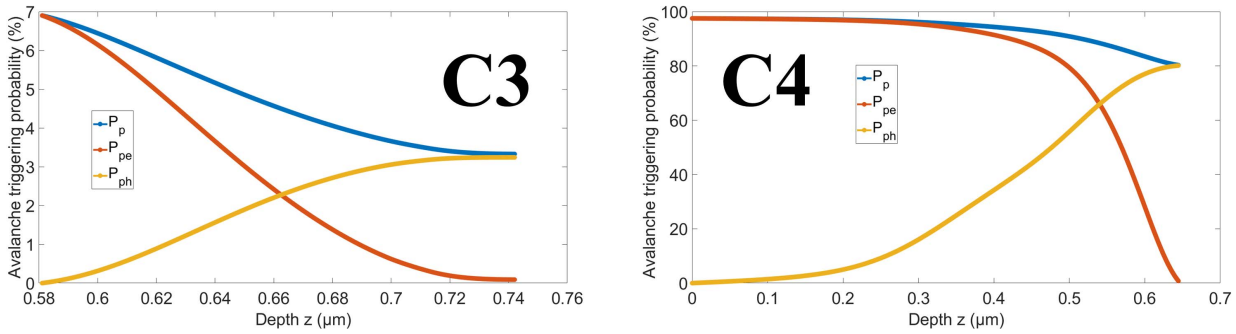

Fig. 16. Avalanche triggering probability for a electron-hole pair P_p, an electron P_{pe} or a hole P_{ph} starting at position z within the depletion region, for cross-section C3 and C4. Van Overstraeten and De Man ionization coefficients were used [9], at T = 300 K.

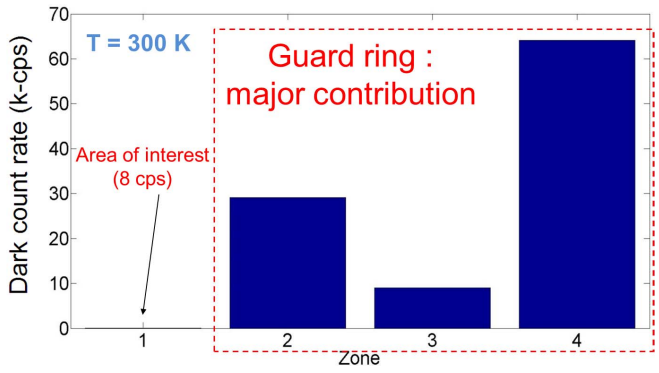

Fig. 17. DCR contribution for each zones.

regions C2 and C4, whereas it is less than 1 % within the area of interest and less than 10 % within the guard ring. These results clearly show that most ionization will happen in the corners zones C2 and C4, and, to a lesser extent, within the guard ring, zone C3. The avalanche triggering probability is very low within the area of interest.

According to equations 11 and 12, the DCR is a combination of the avalanche triggering probabilities and generation rates. As seen previously, these contributions are dominant in the high electric field region, hence we expect to encounter the highest DCR values within these high field regions. This is observed in figure 17 showing the simulated DCR for each zone. As expected, the guard ring, especially its corners, presents the highest DCR values. As discussed previously in section III, the rings seem to enter Geiger mode before the area of interest of the SPAD, which is more likely at the limit of the linear mode. This observation is supported by figures 15 and 16, showing an avalanche triggering probability less than 1 % within the area of interest and almost 100 % for the guard ring, and by figure 17, pointing out the weak DCR contribution of 8 cps, for the SPAD interest region. Therefore, the model is validated for the dark count generated predominantly by the guard rings.

V. Conclusion

In this paper, we developed a model to predict the dark count rate in CMOS SPAD, over a large range of temperature. The main strength of the method is that the model uses only two fitting parameters, obtained by two measurements at low and ambient temperature. The model can be adjusted to any SPAD dimension, in particular for large devices required for spaceborne LIDAR. Comparison between simulated and measured data shows very good results with high correlation between simulated and measured data.

The model is validated with a SPAD in which the dark count is generated, for the most part, within the guard rings and not the area of interest.

The premature edge breakdown within the rings implies high DCR. PEB is present as the chip was fabricated prior to the doping profile measurements. Besides, the external quenching used implies significant recharge times which conceal the afterpulsing phenomenon.

To go further and to improve our SPAD, our perspectives are to design and fabricate new test chips. The set of SPAD of these circuits will be designed using doping profile measurements with a particular attention to the electric field profile, to avoid PEB and improve DCR. Furthermore, the SPAD, the quenching and the shaping electronic will be co-integrated to reduce parasitic capacitances and the recharge time.

We use the local model to evaluate ionization coefficients and avalanche triggering probabilities, this model could show its limits for the most advanced technology nodes and to deal with every CMOS processes, a version implementing the non-local model is also considered. Other SPAD performance parameters modelling, as photon detection probability, afterpulsing probability and timing jitter are also under consideration.

ACKNOWLEDGMENT

The authors would like to thank V. Lalucaa and S. Rolando for the chip design and valuable discussions and advise.

REFERENCES

[1] A. Eisele *et al.*, "185 MHz count rate 139 dB dynamic range single-photon avalanche diode with active quenching circuit in 130 nm CMOS technology," in *Proc. Int. Image Sensor Workshop*, Jun. 2011, pp. 278–280.

[2] A. Rochas, "Single photon avalanche diodes in CMOS technology," Ph.D. dissertation, Microelectron. Microsyst. Inst., École Polytechnique Fédérale de Lausanne, Lausanne, Lausanne, Switzerland, 2003.

[3] S. Pellegrini *et al.*, "Industrialised SPAD in 40 nm technology," in *IEDM Tech. Dig.*, Dec. 2017, pp. 1–16.

[4] T. C. de Albuquerque *et al.*, "Integration of SPAD in 28nm FDSOI CMOS technology," in *Proc. 48th Eur. Solid-State Device Res. Conf. (ESSDERC)*, Sep. 2018, pp. 82–85.

[5] Y. Kang, H. X. Lu, Y.-H. Lo, D. S. Bethune, and W. P. Risk, "Dark count probability and quantum efficiency of avalanche photodiodes for single-photon detection," *Appl. Phys. Lett.*, vol. 83, no. 14, pp. 2955–2957, Oct. 2003.

[6] W. Kindt and H. Van Zeijl, "Modelling and fabrication of Geiger mode avalanche photodiodes," *IEEE Trans. Nucl. Sci.*, vol. 45, no. 3, pp. 715–719, Jun. 1998.

[7] Y. Xu, P. Xiang, X. Xie, and Y. Huang, "A new modeling and simulation method for important statistical performance prediction of single photon avalanche diode detectors," *Semicond. Sci. Technol.*, vol. 31, no. 6, Jun. 2016, Art. no. 065024.

[8] R. J. McIntyre, "Multiplication noise in uniform avalanche diodes," *IEEE Trans. Electron Devices*, to be published.

[9] R. Van Overstraeten and H. De Man, "Measurement of the ionization rates in diffused silicon PN junctions," *Solid-State Electron.*, vol. 13, no. 5, pp. 583–608, 1970.

[10] W. Oldham, R. Samuelson, and P. Antognetti, "Triggering phenomena in avalanche diodes," *IEEE Trans. Electron Devices*, vol. 19, no. 9, pp. 1056–1060, Sep. 1972.

[11] W. Kindt, "Geiger-mode avalanche photodiode arrays," Ph.D. dissertation, Electron. Instrum. Lab., TU Delft, Delft, The Netherlands 1999.

[12] H. S. Bennett and C. L. Wilson, "Statistical comparisons of data on band-gap narrowing in heavily doped silicon: Electrical and optical measurements," *J. Appl. Phys.*, vol. 55, no. 10, pp. 3582–3587, May 1984.

[13] Y. Varshni, "Temperature dependence of the energy gap in semiconductors," *Physica*, vol. 34, no. 1, pp. 149–154, Jan. 1967.

[14] W. Bludau, A. Onton, and W. Heinke, "Temperature dependence of the band gap of silicon," *J. Appl. Phys.*, vol. 45, no. 4, pp. 1846–1848, Apr. 1974.

[15] G. Hurkx, D. Klaassen, and M. Knuvers, "A new recombination model for device simulation including tunneling," *IEEE Trans. Electron Devices*, vol. 39, no. 2, pp. 331–338, Feb. 1992.

[16] J. G. Fossum, "Computer-aided numerical analysis of silicon solar cells," *Solid-State Electron.*, vol. 19, no. 4, pp. 269–277, Apr. 1976.

[17] J. G. Fossum and D. S. Lee, "A physical model for the dependence of carrier lifetime on doping density in non-degenerate silicon," *Solid-State Electron.*, vol. 25, no. 8, pp. 741–747, 1982.

[18] J. G. Fossum, R. P. Mertens, D. S. Lee, and J. F. Nijs, "Carrier recombination and lifetime in highly doped silicon," *Solid-State Electron.*, vol. 26, no. 6, pp. 569–576, 1983.

[19] A. Schenk, "A model for the field and temperature dependence of Shockley-Read-Hall lifetimes in silicon," *Solid-State Electron.*, vol. 35, no. 11, pp. 1585–1596, Nov. 1992.

[20] D. Roulston, N. Arora, and S. Chamberlain, "Modeling and measurement of minority-carrier lifetime versus doping in diffused layers of n^+-p silicon diodes," *IEEE Trans. Electron Devices*, vol. 29, no. 2, pp. 284–291, Feb. 1982.

[21] B.-L. Bérubé, "Conception de matrices de diodes à avalanche à photon unique sur circuits intégrés CMOS 3D," Ph.D. dissertation, Dept. Elect. Comput. Sci., Univ. de Sherbrooke, Sherbrooke, QC, Canada 2014.

[22] L. Pancheri and D. Stoppa, "Low-noise single photon avalanche diodes in 0.15 μm CMOS technology," in *Proc. Eur. Solid-State Device Res. Conf. (ESSDERC)*, Sep. 2011, pp. 179–182.

[23] S. Sze, *Physics of Semiconductor Devices*, 2nd ed. Hoboken, NJ, USA: Wiley, 1981.

[24] M. W. Fishburn, "Fundamentals of CMOS single-photon avalanche diodes," Ph.D. dissertation, Electron. Instrum. Lab., Delft Univ. Technol., Amsterdam, The Netherlands, 2012.

[25] M. Lanoo and J. Bourgoin, *Point Defects in Semiconductors I, Theoretical Aspects* (Solid-State Sciences), vol. 22. Berlin, Germany: Springer, 1981, p. 2.

[26] G. Hurkx, "On the modelling of tunnelling currents in reverse-biased pn junctions," *Solid-State Electron.*, vol. 32, no. 8, pp. 665–668, Aug. 1989.

[27] C. Niclass, M. Gersbach, R. Henderson, L. Grant, and E. Charbon, "A single photon avalanche diode implemented in 130-nm CMOS technology," *IEEE J. Sel. Topics Quantum Electron.*, vol. 13, no. 4, pp. 863–869, Jul. 2007.

Aymeric Panglosse (Student Member, IEEE) received the M.S degree in electrical engineering from ESIEE Paris, France, in 2014, and the master degree in electronics and telecommunication for aerospace applications from ISAE-SUPAERO, Toulouse, France, in 2015. He is currently pursuing the Ph.D. degree with CNES, the French National Space Agency, Airbus Defence and Space, and ISAE-SUPAERO, Toulouse. His doctoral research focuses on performance parameters modeling of single-photon avalanche diodes for space LIDAR applications.

Philippe Martin-Gonthier (Member, IEEE) received the M.S. degree in electrical engineering from ENSERB, Bordeaux, France, in 1998, the Diplôme d'Etudes Approfondies degree in integrated circuit design from ENSAE, Toulouse, France, in 2001, and the Ph.D. degree in electrical engineering from the Université de Toulouse, Toulouse, in 2010.

Since 1998, he has been a Scientist with the Image Sensor Research Team, ISAE-SUPAERO, Université de Toulouse. His doctoral research focused on noise in submicrometer CMOS sensors and design techniques. His research interests include solid-state image sensors, integrated circuit design, noise, and particularly random telegraph signal in pixel arrays.

Olivier Marcelot (Member, IEEE) received the Ph.D. degree from Paul Sabatier University, Toulouse, France, in 2007. In 2008, he was with Synopsys, Zurich, Switzerland, as a Development Engineer. In 2009, he joined Espros AG, Sargans, Switzerland, as a Device Engineer. Since 2011, he has been with the ISAE-SUPAERO, Toulouse, where he is currently a Research Scientist of physic of photodetector.

Cédric Virmontois (Senior Member, IEEE) received the engineering degree in physics from INSA, Toulouse, France, in 2008, and the Ph.D. degree in microelectronic from ISAE-SUPAERO, Toulouse, in 2012. He is currently a Detection Chain Architect with the Science Payload and Imaging Division, CNES, the French National Space Agency, Toulouse. His work involves the development of imagers for future space missions.

Olivier Saint-Pé (Member, IEEE) is currently an Airbus Defence and Space expert in photon detectors. Since 1986, he has been involved in numerous projects requiring high-performances detectors/focal planes. He has authored or coauthored over 50 scientific/technical articles. Since 2003, he has been the Chair of the Focal Plane Technologies Session with SPIE Remote Sensing Symposium. He is a member of the CNES Technical Competence Center and the ESA Photonics Working Group.

Pierre Magnan (Member, IEEE) received the degree in electrical engineering from the University of Paris, France, in 1980. After being a Research Scientist involved in CMOS design until 1993, he moved to image sensors research at ISAE-SUPAERO, Toulouse, France, where he is currently a Full Professor and the Head of the Image Sensor Research Group.