Modeling Negative Feedback in Single Photon Avalanche Diodes

Majeed M. Hayat*[a], David A. Ramirez[a], Graham J. Rees[c], Mark A. Itzler[b]
[a]Center for High Technology Materials and ECE Dept., University of New Mexico, 1313 Goddard St., SE, Albuquerque, NM, USA 87106;
[b]Princeton Lightwave, Inc., 2555 Route 130 South Cranbury, NJ, USA 08512;
[c]Electronic & Electrical Engineering, University of Sheffield, Sir Frederick Mappin Building, Mappin Street, Sheffield S1 3JD, United Kingdom

ABSTRACT

Recently, considerable attention has been placed upon exploiting the negative-feedback effect in accelerating the quenching time of the avalanche current in passively quenched single-photon avalanche-diode (SPAD) circuits. Reducing the quenching time results in a reduction in the total charge generated in the SPAD, thereby reducing the number of trapped carries; this, in turn, can lead to improved after-pulsing characteristics. A passively quenched SPAD circuit consists of a DC source connected to the SPAD, to provide the reverse bias, and a series load resistor. Upon a photon-generated electron-hole pair triggering an avalanche breakdown, current through the diode and the load resistor rises quickly reaching a steady state value, after which it can collapse (quench) at a stochastic time. In this paper we review recent analytical and Monte-Carlo based models for the quenching time. In addition, results on the statistics of the quenching time and the avalanche pulse duration of SPADs with arbitrary time-variant field across the multiplication region are presented. The calculations of the statistics of the avalanche pulse duration use the dead-space multiplication theory (DSMT) to determine the probability of the avalanche pulse to quench by time t after the instant s at which the electron-hole pair that triggers the avalanche was created. In the analytical and Monte-Carlo based models for the quenching time, the dynamic negative feedback, which is due to the dynamic voltage drop across the load resistor, is taken into account. In addition, in the Monte-Carlo simulations the stochastic nature of the avalanche current is also considered.

Keywords: avalanche photodiode, passive quenching, photon counting, single-photon detection, quenching time, impact ionization, avalanche breakdown, negative feedback

1. INTRODUCTION

In recent years there has been an increased interest in high-sensitivity sensing of photonics signals in applications for which the signals are extremely weak. Such applications span a wide range of the electromagnetic wavelength spectrum from ultraviolet (10 - 400 nm) to the long-wave infrared (8-12 μm). Applications include satellite laser ranging, gated 2D and 3D imaging, deep-space laser communication, time-resolved photon counting, quantum key distribution, quantum imaging, photon-counting 3D integral imaging, and quantum cryptography. Under these photon-starved conditions, conventional photodetectors that linearly transform an optical intensity to an electrical current cannot be employed as the photon rates are very low and the resulting photocurrent is totally overwhelmed by Johnson noise arising from resistive elements in the detector as well as the pre-amplifiers and other electronic components used by the detection system. A common approach for detecting sparsely arriving photons is based on employing a non-linear detection scheme, whereby the absorption of a photon results in a huge, saturated current that can be easily detected by electronic circuitry without ambiguity. One of the devices that are commonly used in such non-linear and saturated-mode detection is the avalanche photodiode (APD) [1], also referred to as a single-photon avalanche diode (SPAD).

Advanced Photon Counting Techniques IV, edited by Mark A. Itzler, Joe C. Campbell,
Proc. of SPIE Vol. 7681, 76810W · © 2010 SPIE · CCC code: 0277-786X/10/$18
doi: 10.1117/12.851914

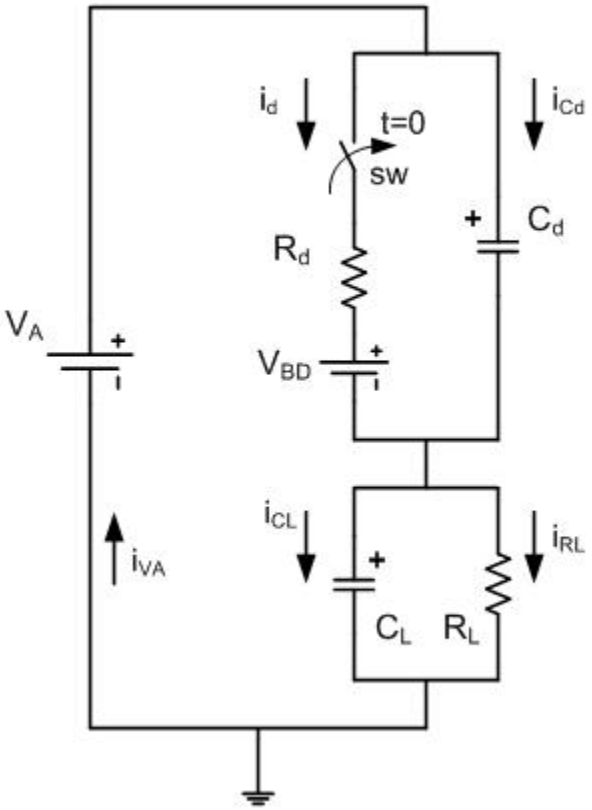

Figure 1. Model for a passive quenching SPAD circuit. I_d represents the self-sustaining current through the multiplication region of the APD; I_N is additive Johnson noise, combining the effects of all resistive elements in the circuit; R_d is the equivalent dynamic resistance of the APD; C_d is the APD's junction capacitance; R_L is the load resistor; and C_L is the parasitic capacitance.

Essentially, an APD is a reversed biased *pn* or *pin* device that operates by converting each electron-hole pair, resulting from the absorption of a photon, to a large number of electron-hole pairs via a cascade of impact ionizations. In its simplest form, the external circuit of an APD consists of a DC source connected to the SPAD (to provide the reverse bias) and a series load resistor.

The type of detector technology used for single-photon detection is determined by the range of wavelengths of the specific application. While silicon SPADs have already shown very good performance in various applications in the 400–900 nm range their performance is degraded drastically when they are operated beyond 1 μm. For applications in the wavelength range 1.06–1.65 μm, devices with a narrower bandgap than silicon, mainly III-V compounds, are utilized. These include InGaAs/InP separate absorption-multiplication SPADs, which are similar to telecommunication APDs. Avalanche photodiodes for mid-infrared (MIR, 3–25 μm range) detection are at their infancy. One approach developed by Krishna *et al*. [2] is based upon incorporating avalanche gain in a quantum dots-in-a-well (DWELL) detector the device through the monolithic inclusion of APD in the DWELL device. Such device is called the quantum dot avalanche photodiode (QDAP) [3]. In the QDAP, an intersubband quantum dot (QD) detector is coupled with an APD through a tunnel barrier. The tunnel barrier reduces the dark current while the avalanche section supplies the photocurrent with internal gain. In this three-terminal device, the applied bias of the QD-detector and the APD section of the QDAP are controlled separately. This feature permits the control of the responsivity and dark current of the QD detector independently of the operating avalanche gain. When operated in Geiger mode the QDAP has the potential for use as a single-photon detector.

In a linear-mode operation of an APD, for which the device is biased below a certain threshold known as the breakdown voltage, the cascade of impact ionizations resulting from each parent carrier pair ends in a finite, stochastic time called the avalanche buildup time. The total number of offspring carriers generated via impact ionizations during the avalanche buildup time constitutes the multiplication factor (gain) by which the photocurrent is amplified (compared to a detector with no gain, such as a pin photodiode).

When the applied bias is above the breakdown voltage, a mode termed Geiger operation, the number of impact ionizations increase in time yielding, in principle, an infinite multiplication factor. However, in an actual device, as shown in Fig. 1, the avalanche current saturates to a level (tens or hundreds of micro Amps) governed by the power supply and the resistive elements [4]. This current is referred to as the self-sustaining or persistent current. Depending upon the value of the applied bias, this persistent current may terminate at a stochastic time, after which the diode behaves like an open circuit. After a recovery period, the voltage across the APD reaches once again the value of the voltage supply and the APD is ready for another avalanche trigger and series of impact ionizations. This type of Geiger-mode operation is referred to as the passive-quenching mode, as the persistent current is allowed to terminate spontaneously. In other configurations, the self-sustaining avalanche current is brought to quenching by the active reduction of the bias across the diode via some external circuit [4]. This latter configuration is referred to as active quenching [4]. The time measured from the onset of the self-sustaining avalanche current to its termination is referred to as the SPAD's quenching time, which is stochastic. Having a good model for the quenching time is important in determining the time resolving capacity and duty cycle of SPAD circuits, as well as in controlling the total amount of charge that flows through the diode as this quantity affects an important performance limitation known as after-pulsing.

In this paper we review recent analytical and Monte-Carlo based models for the quenching time. In the description of these models we will draw freely from our recent work reported in [5]. In addition, results on the statistics of the quenching time and the avalanche pulse duration of SPADs with arbitrary time-variant field across the multiplication region are presented. The calculations of the statistics of the avalanche pulse duration use the dead-space multiplication theory (DSMT) to determine the probability of the avalanche pulse to quench by time t after the instant s at which the electron-hole pair that triggers the avalanche was born. In particular, the model determines the quenching probability as a function of time for an arbitrary time-varying electric field profile across the multiplication region. As such, the calculation of the statistics of the duration of the avalanche pulse under time-variant fields constitutes a very valuable tool that can help the design of biasing strategies for SPADs used in application of signal communications. In the analytical model for the quenching time, the dynamic negative feedback, which is due to the dynamic voltage drop across the load resistor, is taken into account, albeit without considering the stochastic fluctuations in the avalanche pulse. In the second model, Monte-Carlo simulation is used to generate impact ionizations with the inclusion of the effects of negative feedback. The latter model is based on simulating the impact ionizations inside the multiplication region according to a dynamic bias voltage that is a function of the avalanche current it induces. In particular, it uses the time evolution of the bias across the diode to set the coefficients for impact ionization. As such, this latter model includes both the negative feedback and the stochastic nature of the avalanche current.

2. BACKGROUND

Upon triggering an avalanche breakdown, current through the diode and the load resistor rises quickly and a voltage is consequently built up across the load resistor. This reduces the bias across the SPAD, thereby weakening the impact ionization process. Consequently, a drop in the current flow occurs in the SPAD leading to a reduction of the voltage across the load resistor. Next, after some RC-limited delay, the voltage is restored across the SPAD and the current increases once again. This cycle may repeat itself giving rise to a fluctuating avalanche current. To reiterate, the fluctuation in the current is as a result of two factors: the avalanche process being stochastic and more importantly, the quasi-periodic fluctuations in the bias of the SPAD resulting from the feedback received from the voltage of the load resistor. The latter effect, referred to as feedback, plays a key role in forcing the self-sustained current around a fixed pre-determined value (excess bias divided by the load resistor) without allowing it to grow indefinitely. The avalanche current eventually collapses via such a statistical, quasi-periodic fluctuation, thereby quenching the avalanche current until another trigger occurs.

Prior efforts in modeling SPAD circuits have been limited to either a deterministic or probabilistic analysis that captures the stochastic nature of impact ionization. Both approaches ignore the effects of feedback on impact ionization.

According to the deterministic model [4] and upon triggering avalanche breakdown, current through the diode and the load resistor can rise quickly until the excess bias is built up across the resistor, leaving the SPAD biased precisely at breakdown. In contrast to the deterministic model, the probabilistic analysis [6] provides a method to calculate the statistics of the quenching time under the same assumptions on the bias-voltage after triggering an avalanche.

3. MODELS FOR PASSIVE QUENCHING

In Fig. 2 below, we show an equivalent circuit for the SPAD passive-quenched circuit of Fig. 1.

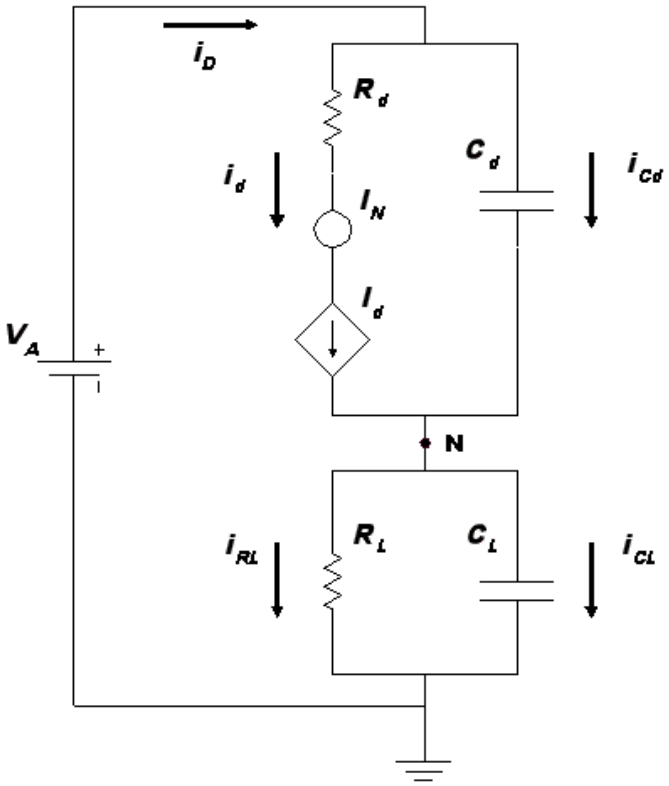

Figure 2. Equivalent circuit of a SPAD circuit. I_d represents the self-sustaining current through the multiplication region of the APD. Note that that the total current produced by the diode is I_D.

3.1 Non-stochastic closed-loop model

This model captures the feedback from the load resistor while ignoring the stochastic element of the avalanche persistent current. Once an avalanche is triggered (while the diode is biased above breakdown), the average avalanche current grows exponentially according to the theory of average impulse response of APDs above breakdown [6]. This growth tends to discharge the capacitor, and therefore, reduce the junction voltage (voltage across C_d), which, in turn, causes the avalanche current to drop. The reduction in the avalanche current continues and the junction bias drops below the breakdown voltage. After this point of time, the DC source begins to recharge the capacitor with the appropriate time constant, causing the avalanche current to increase once again, and so on. The repetition of this process can yield an oscillatory behavior, where the current inside the diode oscillates about the steady state value. Moreover, the field inside the multiplication region of the APD oscillates above and below the breakdown threshold.

The type of behavior described above can actually be seen in an analytical model similar to that reported in [4] if we replace the short connection of the diode, as done traditionally [4], by a voltage-dependent current source. The latter will allow the current to grow exponentially (in time) whenever the junction bias is above breakdown; the growth rate is

proportional to the excess bias beyond the breakdown voltage. On the other hand, the current is allowed to decay exponentially fast when the bias is below breakdown at a rate proportional to the excess bias.

An analytical model capturing all of the above mentioned factors has been developed. Figure 3 shows the calculated voltage across the load resistor according to this non-stochastic, closed-loop model. The oscillations are indeed centered about the steady state current through the diode.

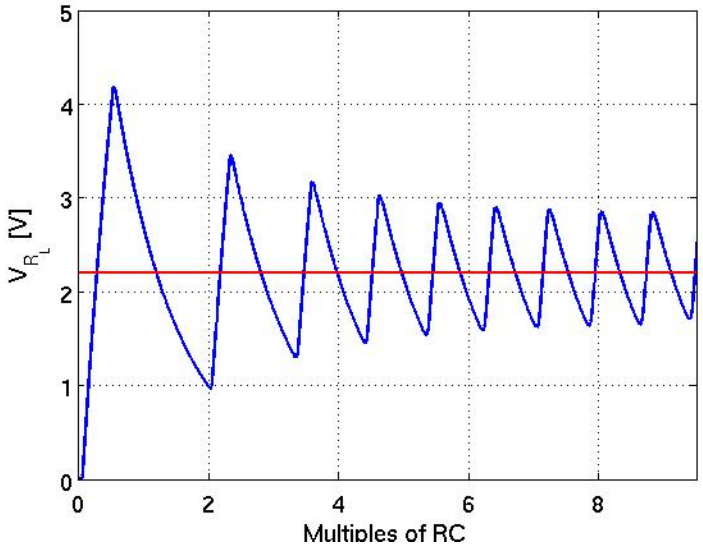

Figure 3. Calculated voltage across the load resistor as a function of time (in terms of multiples of the overall RC time constant). Here, $R \sim R_L$ and $C \sim C_L$. The red straight line is the steady state voltage, resulting simply from voltage division between R_L and R_d.

We note that the oscillatory behavior is governed by two factors: (1) how quickly the junction capacitor can be discharged, which, in turn, depends on the RC time constant and on the growth rate of the avalanche pulse; and (2) how fast the change in the junction voltage can alter the avalanche current in the diode. The latter effect, characterized by a delay factor, δ, is akin to a delay brought about by an inductor, which induces oscillations.

The oscillatory behavior described above has indeed been observed experimentally by Princeton Lightwave Inc. A representative measurement is shown in Fig. 4.

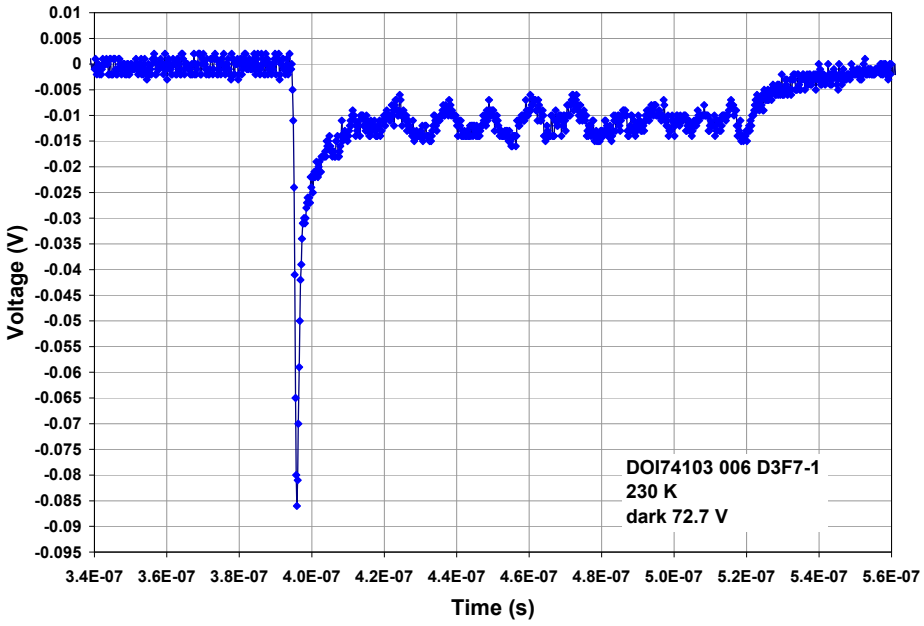

Figure 4. Measured AC-coupled voltage across the load resistor as a function of time. The negative peak marks the trigger of an avalanche. The eventual rise marks the quenching instant and the start of the re-charge of the junction capacitor. The mid portion shows the oscillation due to the feedback.

3.2 Stochastic closed-loop model

To capture the feedback from the load resistor as well as stochastic element of the avalanche persistent current, we will resort to a Monte-Carlo model. An analytical model would be intractable due to the closed-loop nature of the problem. In this model, the carrier dynamics and multiplication are simulated instantaneously via a custom-made real-time Monte-Carlo simulator. In particular, the ionization coefficients of the APD are made dependent on the field across the APD's capacitor, C_d. As the carriers multiply stochastically, the stochastic current supplied by the diode is calculated using Ramo's theorem according to the carries inside the multiplication region. The current evolution in the entire RC-circuit, including the APD's avalanche current is simulated over sufficiently small time increments (much smaller to the carriers' transit times in the multiplication region). We also allow for a delay representing the non-instantaneous response of the field as the voltage across the diode is changed due to feedback.

In Figure 5 below we show two realizations of the calculated voltage across the load resistor according to this stochastic, closed-loop model. The oscillations are indeed centered about the steady state current through the diode. Figure 6 shows the voltage across the capacitor C_d as a function of the transit time of the electron and hole carries in an InP multiplication region of 1600 nm. It can be seen that the voltage across C_d oscillates about the breakdown voltage shown in the red curve. The current through the diode as function of the transit time is shown in fig. 7 for two realizations. This figure also shows an amplified view of three specific times in the simulation. More specifically, fig. (7.b) shows the stochastic quenching of one avalanche pulse (red curve) and the fluctuation of another pulse that does not quench (blue curve). In fig. (7.c) the pulse that did not quench last cycle is able avoid quenching one more time. Finally, fig. (7.d) shows the stochastic quenching of the second pulse.

The Monte-Carlo simulation study has led to some preliminary conclusions: (1) when persistence occurs, quenching always occurs when the field across the multiplication region is in the high cycle (above breakdown); (2) when the delay factor described earlier is reduced below a transit time, then either quenching occurs right after the first cycle or persistence continues indefinitely. In the latter case, the feedback corrects instantaneously leading to quenching.

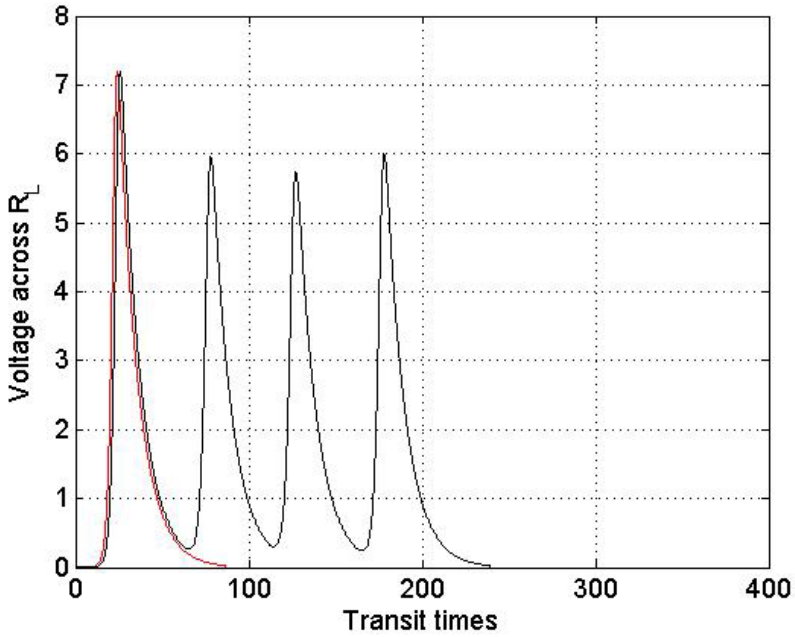

Figure 5. Two Monte-Carlo simulations of the voltage across the loads resistor as a function of the transit time of the electron and hole carriers in the multiplication region. In this example, there are approximately 11 transit times in each RC time constant. For the red curve, quenching occurs after the initial cycle while for the black curve quenching occurs after the fourth cycle.

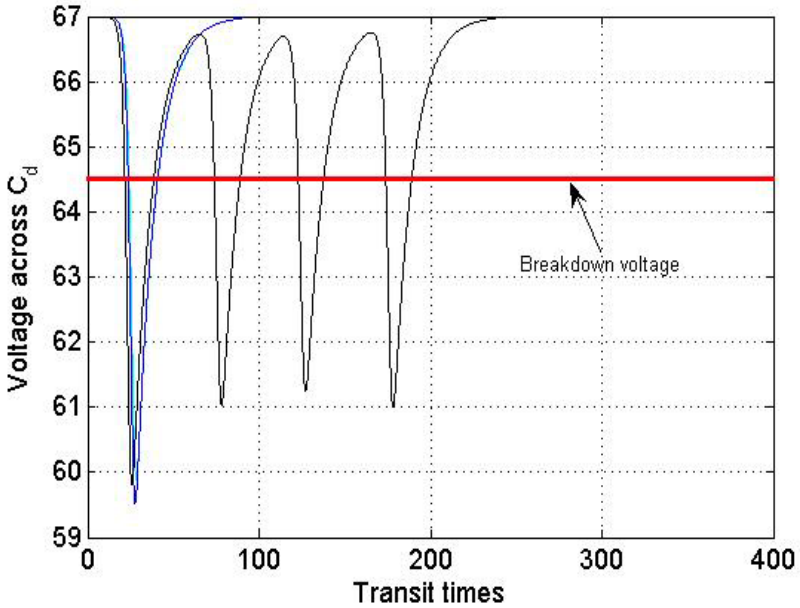

Figure 6. Monte-Carlo simulations of the voltage across the capacitor C_d as a function of the transit time of the electron and hole carriers in an InP multiplication region of 1600 nm. The breakdown voltage (red curve) is also shown. Note that in both cases quenching occurs when the field is in the high cycle (above breakdown).

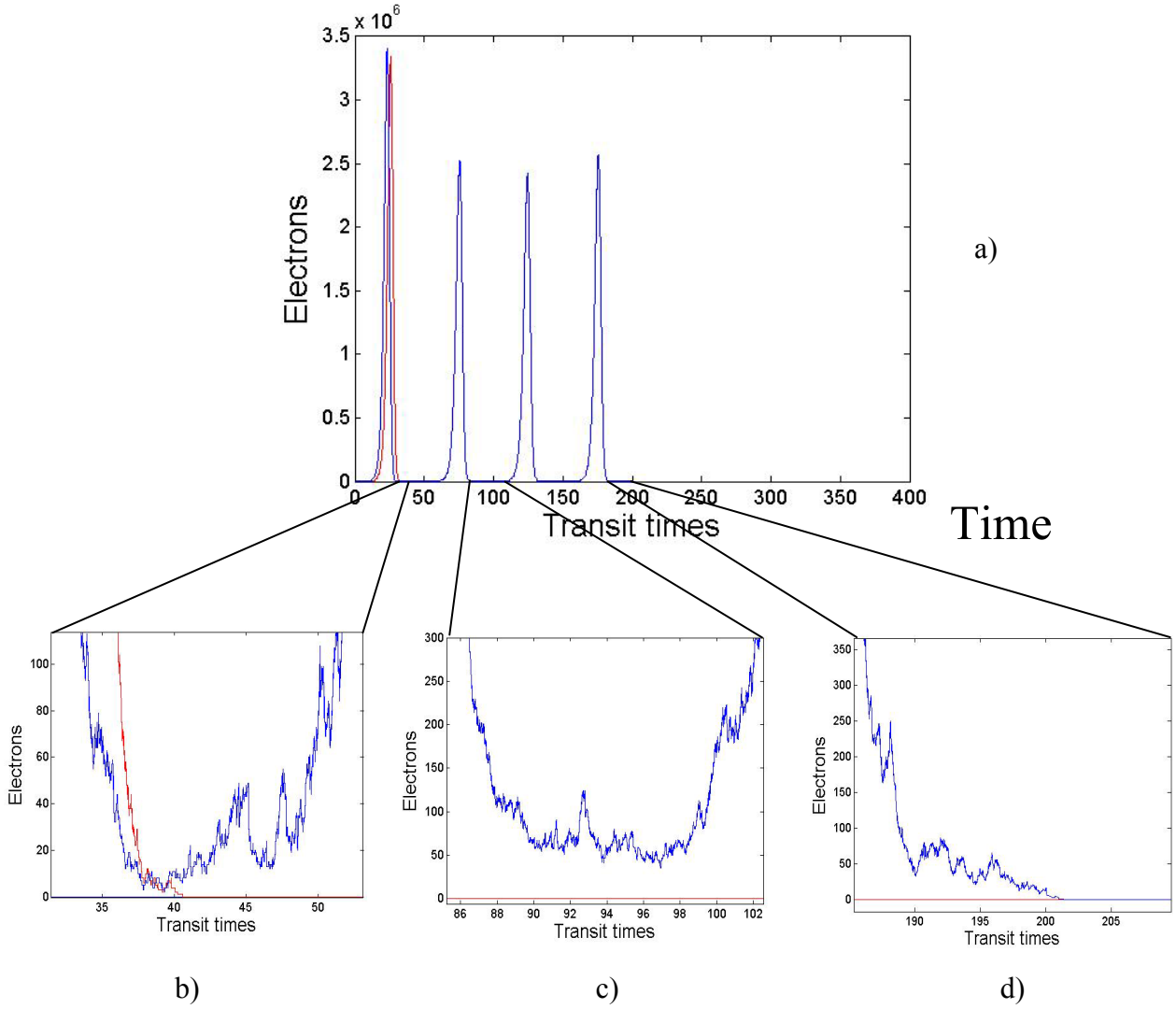

Figure 7. Two Monte-Carlo simulations of the current through the diode as a function of the transit time of the electron and hole carries in the multiplication region. The three bottom curves show an amplified view of three different times: b) shows the stochastic quenching of one avalanche pulse (red curve) and the fluctuation of another pulse that does not quench (blue curve), in (c) the pulse that did not quench last cycle is able avoid quenching one more time, and d) shows the stochastic quenching of the second pulse.

4. MODELING THE STATISTICS OF QUENCHING TIME UNDER DYNAMIC PERSISTENT CURRENT

The objective is to determine the statistics of the quench time, T_Q, which is defined as the time from avalanche trigger to the self termination of the self-persisting current. For simplicity of the analysis, we will adopt the approximation that the self persisting current in the diode, $i(t)$, is oscillatory from its onset, i.e. we will neglect the transient segment and assume it follows the periodic pattern.

We start by observing that the voltage across the diode, $v_{cd}(t)$ shown in fig. 6, oscillates about the breakdown voltage. At the same time, the current inside the diode, $i(t)$ shown in fig. 7, also varies periodically about the steady-state value I_{ss}

$=\Delta V/(R_L+R_d)$. Let I^+ and I^- be the average current over the period when $v_{cd}(t)$ is above the breakdown voltage (positive period) and below the breakdown voltage (negative period), respectively. Similarly, let V^+ and V^- be the average voltages $v_{cd}(t)$ over the positive and negative periods, respectively. Further let P^+ be the probability all I^+/q carrier pairs present in the positive period vanish within the positive period and P^- the probability all I^-/q carriers present in the negative period vanish within the negative period. Finally, let τ^+ and τ^- denote the durations of the positive and negative periods, respectively. Note that $\tau^+ + \tau^- = \tau$, is the period of the oscillation.

With the above quantities available from the model-based calculations, we can calculate the probability $P(t)$ that quenching occurs before time t as follows. Note that quenching does not occur if and only if no quenching occurs in any of the positive and negative periods. As there are $t/(2\tau)$ positive and $t/(2\tau)$ negative periods, respectively, we can write

$$P(t) = 1 - \left(1 - P^+\right)^{t/2\tau} \left(1 - P^-\right)^{t/2\tau} \tag{1}$$

If we take t to be a multiple of the period of oscillation τ, i.e., $t = n\tau$, then we have

$$P(n\tau) = 1 - \left(1 - P^+\right)^{n/2} \left(1 - P^-\right)^{n/2}. \tag{2}$$

The probability that quenching occurs precisely at the nth period is ($n = 0$ corresponds to first period)

$$P_n = \left[1 - \left(1 - P^+\right)\left(1 - P^-\right)\right]\left[\left(1 - P^+\right)\left(1 - P^-\right)\right]^{n/2}. \tag{3}$$

However, results from Monte-Carlo simulations show that quenching occurs only in the positive cycle, which means that $P^-=0$. Then, by letting $P= P^+$ we can recast (3) as

$$P_n = P\left(1 - P\right)^{n/2} \tag{4}$$

And, the mean and variance of the random period, N_Q, of the quenching time are respectively

$$\begin{aligned} &< N_Q >= \frac{\left(1-P\right)}{P} \\ &\mathrm{var}(N_Q) = \frac{\left(1-P\right)}{P^2}, \end{aligned} \tag{5}$$

In addition, the mean and variance of the time to the actual quenching time are respectively

$$\begin{aligned} &< T_Q >= \tau \frac{\left(1-P\right)}{P} \\ &\mathrm{var}(T_Q) = \tau^2 \frac{\left(1-P\right)}{P^2} \end{aligned} \tag{6}$$

5. STATISTICS OF AVALANCHE PULSE DURATION UNDER TIME-VARYING FILED

In order to apply the DSMT to calculate the avalanche-pulse duration for the carriers generated in the SPAD, we use (i) the nonlocalized ionization coefficients, also called enabled ionization coefficients (the ionization coefficient assumed once the carrier travels the dead-space distance), and the threshold energies for each material, and (ii) the electric-field profile through the device.

In the calculations it is assumed that the voltage across the multiplication region has a duration of t=10 transit times, after which the voltage is zero, as shown in fig. 8. It is also assumed that the electron that triggers the avalanche is injected at the start of the multiplication region (x=0). First, we examine the constant-field case. Figure 9 shows the calculated probability of the avalanche pulse to quench by time t for an electron born at time s under a constant field

across the multiplication region. Two cases were simulated: (i) when the electron that triggers the avalanche is created at time s=0 (red curve), and (ii) when the electron that triggers the avalanche is created at time s~45 % of the total duration of the applied voltage (blue curve). It is apparent that under a constant field the probability of pulse termination is not affected by the time at which the avalanche is triggered. On the other hand, when we examine the time-varying-field case, shown in fig. 10, we obtain more informative results. Figure 11 shows the calculated probability of the avalanche pulse to quench by time t for an electron born at time s under a time-varying field. The same two cases were simulated: (i) when the electron that triggers the avalanche is created at time s=0 (red curve), and (ii) when the electron that triggers the avalanche is created at time s~45 % of the total duration of the applied voltage (blue curve). It can be clearly seen that since the voltage across the device is time-dependent (fig. 10) the probability of pulse termination will depend on the time at which the electron was created. As expected, fig. 11 shows that an avalanche pulse created at times where the electric field is above breakdown (red curve) has a smaller probability of termination compared to an avalanche created at times where the field is smaller (blue curve).

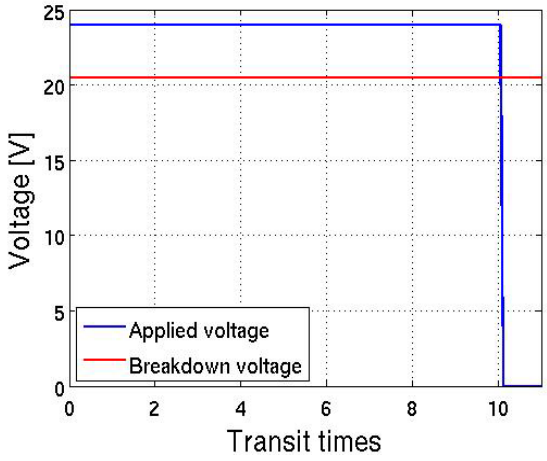

Fig. 8: Applied voltage as a function of time (blue curve) and breakdown voltage (red curve) of a SPAD with an InP multiplication region of 400 nm.

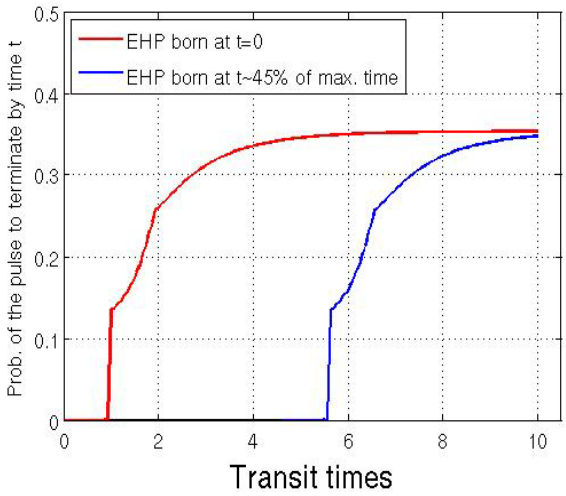

Fig. 9: Probability of pulse termination as a function of time for: (a) an electron created at time s=0 (red curve), and (b) an electron created at time s~45 % of the maximum time (blue curve).

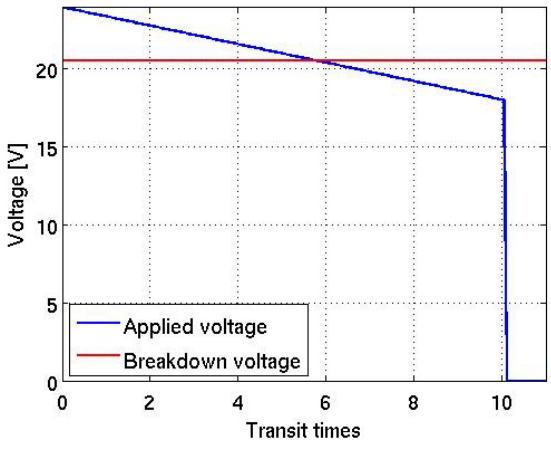

Fig. 10: Applied voltage as a function of time (blue curve) and breakdown voltage (red curve) of a SPAD with an InP multiplication region of 400 nm.

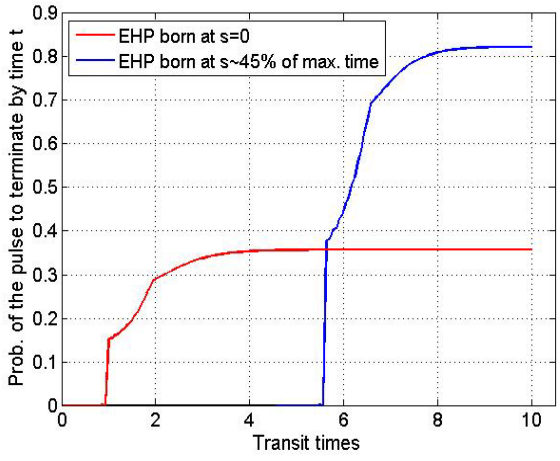

Fig. 11: Probability of pulse termination as a function of time for: (a) an electron created at time s=0 (red curve), and (b) an electron created at time s~45 % of the maximum time (blue curve).

6. CONCLUSIONS

We review analytical and Monte-Carlo based models for the quenching time. Preliminary results suggest that feedback must be incorporated in the analysis. Our results also suggest that there is an oscillatory behavior as a result of the feedback, an attribute that has been observed experimentally.

It is expected that by combining the Monte-Carlo based model with the calculations of the time-variant-field quenching probability will help us design biasing strategies for SPADs for applications in signal communications.

This work was supported by the National Science Foundation under award ECS-0601645 and in part by NASA under grant NNG06LA04C.

REFERENCES

[1] Saleh, B. E. A. and Teich, M. C., [Fundamentals of Photonics, Second Edition], New York: Wiley (2007).
[2] Krishna, S., Kwon, O., and Hayat, M. M., "Theoretical investigation of quantum-dot avalanche photodiodes for mid-infrared applications," IEEE J. Quantum Electronics, 32, 1468 – 1473 (2005).
[3] Ramirez, D. A., Shao, J., Hayat, M. M., and Krishna, S., "Demonstration of the quantum dot avalanche photodiode (QDAP)," Proc. SPIE 7320 (2009).
[4] Cova, S., Ghioni, M., Lacaita, A., Samori, C.,, and Zappa, F., "Avalanche photodiodes and quenching circuits for single-photon detection," Appl. Opt., 35, 1956–1976 (1996).
[5] Hayat, M. M., Itzler, M. A., Ramirez, D. A., and Rees, G. J., "Model for Passive Quenching of SPADs," SPIE Photonics West 2010, San Francisco, California, Jan. 22-27 (2010).
[6] Hayat, M. M., Rees, G. J., Ramirez, D. A., and Itzler, M. A., "Statistics of self-quenching time in single photon avalanche diodes," IEEE-LEOS 2008 Annual Meeting: Photodetectors and Imaging, Long Beach, CA, Nov. 9-13 (2008).