Model for Passive Quenching of SPADs

Majeed M. Hayat*[a], Mark A. Itzler[b], David A. Ramirez[a], Graham J. Rees[c]

[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

Infrared single-photon avalanche photodiodes (SPADs) are used in a number of sensing applications such as satellite laser ranging, deep-space laser communication, time-resolved photon counting, quantum key distribution and quantum cryptography. 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 three recent analytical and Monte-Carlo based models for the quenching time. In the first model, the applied bias after the trigger of an avalanche is assumed to be constant at the breakdown bias while the avalanche current is allowed to be stochastic. In the second model, 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 third 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 indices. 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.

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).
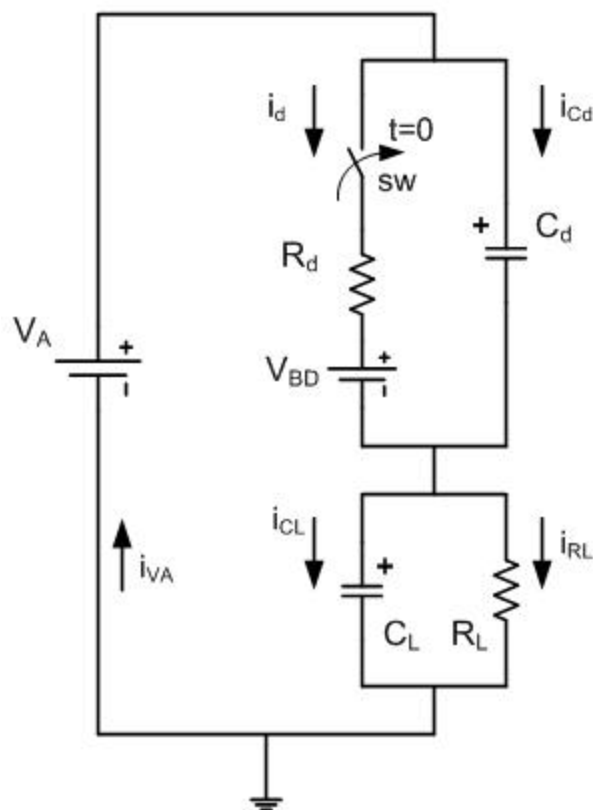
Figure 1. Model for a passive quenching SPAD circuit. I_d represents the self-sustaining current through the multiplication region of the APD; I_N is additive Johnson noise, combining the effects of all resistive elements in the circuit; R_d is the equivalent dynamic resistance of the APD; C_d is the 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, the 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 three recent analytical and Monte-Carlo based models for the quenching time. The first analytical model is that reported in [5]. In this model the bias after the trigger of an avalanche is assumed to be constant at precisely the breakdown bias, which is consistent with [4], and the avalanche current is allowed to be stochastic according to [5,6]. In the second analytical model, the negative feedback is taken into account, albeit no stochastic fluctuations are taken into account. In the third model, Monte-Carlo simulation is used to generate impact ionizations with the inclusion of the effects of negative feedback. This model is based on simulating the impact ionizations inside the multiplication region according to a dynamic bias voltage, for which no analytical treatment is available. In particular, it uses the time evolution of the bias across the diode to control the impact ionization coefficients in the Monte-Carlo phase. As such, this model includes both the negative feedback and the stochastic nature of the avalanche current.

2. DYNAMICS OF PASSIVE QUENCHING

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 [5] 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.
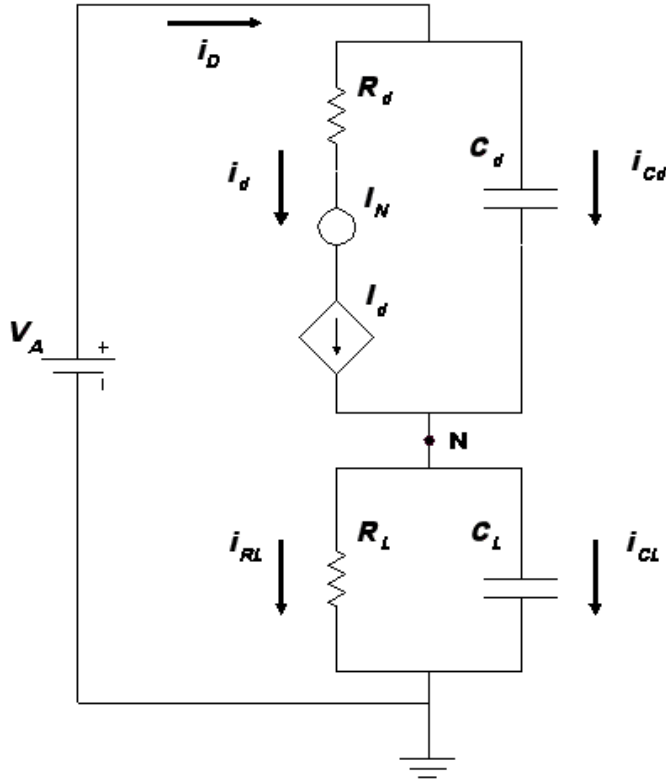

Figure 2. Equivalent circuit of a SPAD circuit. I_d represents the self-sustaining current through the multiplication region of the APD. Note that that the total current produced by the diode is I_D.

3.1 Model I: Stochastic open-loop model

Here we describe a model that assumes that after the trigger of the avalanche, the field remains at the breakdown-field level until the persistent current is quenched due to stochastic fluctuations that are inherent in the impact ionization process. This model is consistent with [4], where it is assumed that upon an initial trigger of an avalanche, the diode acts as a closed switch and the average steady state current flowing through it is governed by the voltage supply and the resistive elements. Here we will review the approach for calculating the statistics of the duration of the stochastic persistent avalanche current, I_d, while drawing freely from our recent work reported in [5]. We begin by noting that the persistent current is generated by electrons and holes distributed throughout the multiplication region; each carrier generates its own individual avalanche pulse. The totality of these contributions gives rise to the stochastic persistent current I_d.

For the avalanche current to self-quench, each of these individual avalanche pulses must terminate independently. The probability that this happens before time t elapses is given by $F_I(t) = \prod_i F_{e,h}(z_i, t)$, where the product is over all electrons and holes, situated at z_i in the multiplication region, and $F_{e,h}(z_i, t)$ is the probability that an electron (hole) injected at z_i will give rise to an avalanche pulse which terminates before time t has elapsed. It has been argued theoretically in [5] that when the SPAD is biased precisely at breakdown and prior to the avalanche current's

collapsing, $F_{e,h}(z_i,t) \approx 1 - f_{e,h}(z_i)/t$. It is worth noting that this asymptotic behavior is different from those corresponding to below or above breakdown, for which the asymptotic behavior is exponential [6]. By utilizing the appropriate spatial distribution of electrons and holes at breakdown (which we have derived but have not included here), it has been shown that the probability distribution function $F_I(t)$ is given by [5]:

$$F_I(t) \approx \exp\left(-\frac{T}{t}\right), \tag{1}$$

where $T = \dfrac{IC\tau_0^2 J}{q}$, $J = 2/\ln(k)\left(2/\ln(k) + \dfrac{1+k}{1-k}\right)$, q is the electronic charge, C is a dimensionless constant, τ_0 is the average electron and hole transit time across the multiplication region, $k = \beta/\alpha$ is the hole-to-electron ionization-coefficient ratio, and I is the steady state current flowing through the diode, given by $V_a/(R_d+R_L)$. We observe that this probability distribution function has an infinite mean; however, its median is $t_{1/2} = T/ln(2)$. Interestingly, the median of the quenching time grows linearly with the steady state current. Figure 3 shows the dependence of $t_{1/2}$ on the ionization coefficient ration, suggesting that SPADs with k values close to unity have a better quench-time performance than those with larger or smaller k values.
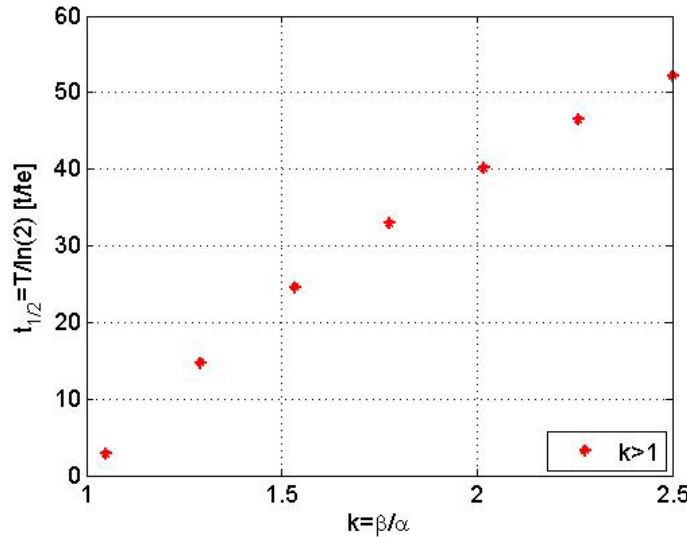

Figure 3. Calculated median of the quenching time as a function of the ionization-coefficient ratio k [5].

One concern about this model is that it suggests that the quenching time has memory. Namely, the form of the distribution function $F_I(t)$ suggests that conditionally on the quenching time being greater than t, the probability that quenching occurs diminishes as a function of s following t, which is unrealistic. However, this shortcoming of the model is a direct result of our open-loop assumption; the model neglects the presence of feedback from the voltage drop on the load resistor, which tends to alter the voltage across the APD. In such an open-loop scenario, a long quenching time implies that the avalanche current is large; thus, there is less likelihood for quenching to occur at longer times. As such, the memoryless property observed under the open-loop assumption is expected albeit not realistic since in actuality we have a closed-loop scenario. (Note that, on average, the avalanche current is still a constant under the open-loop assumption.)

3.2 Model-II: 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 4 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.
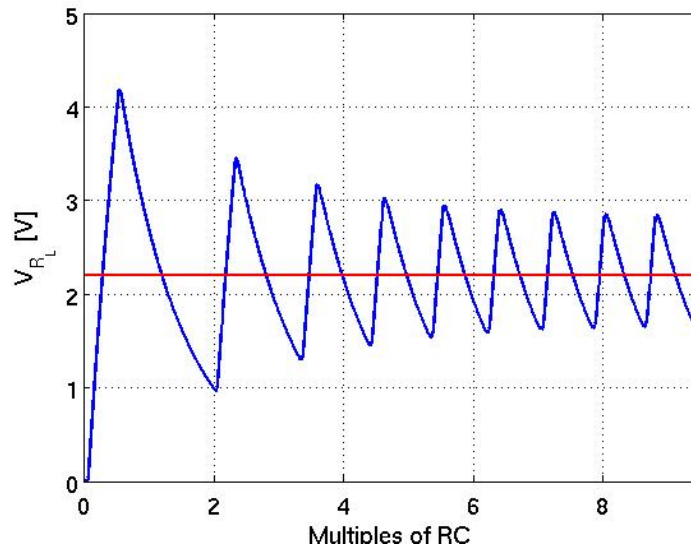

Figure 4. Calculated voltage across the load resistor as a function of time (in terms of multiples of the overall RC time constant). Here, $R \sim R_L$ and $C \sim C_L$. The red straight line is the steady state voltage, resulting simply from voltage division between R_L and R_d.

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. With such periodicity in the electric field inside the multiplication region, we would expect the quenching to occur in cycles for which the field is below breakdown, which, in turn, implies that quenching time is memoryless. This attribute, which is consistent with intuition and experiment, overcomes the shortcoming of Model-I.

The oscillatory behavior described above has indeed been observed experimentally by Princeton Lightwave Inc. A representative measurement is shown in Fig. 5.
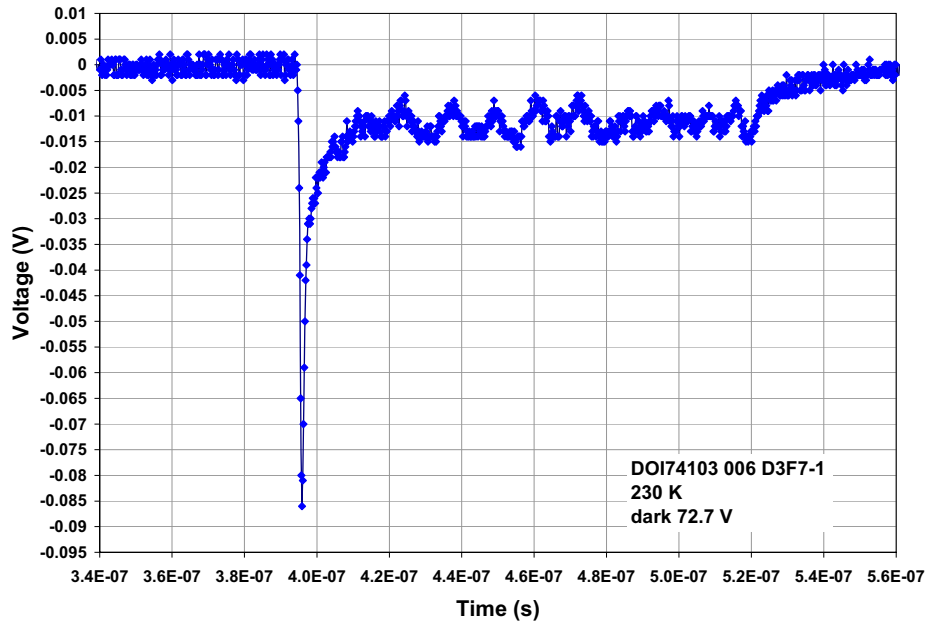
Figure 5. Measured AC-coupled voltage across the load resistor as a function of time. The negative peak marks the trigger of an avalanche. The eventual rise marks the quenching instant and the start of the re-charge of the junction capacitor. The mid portion shows the oscillation due to the feedback.

3.3 Model-III: 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. I 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 6 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.

The Monte-Carlo simulation study has led to some preliminary conclusions: (1) when persistence occurs, quenching always occurs when the field is in the low cycle; (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.
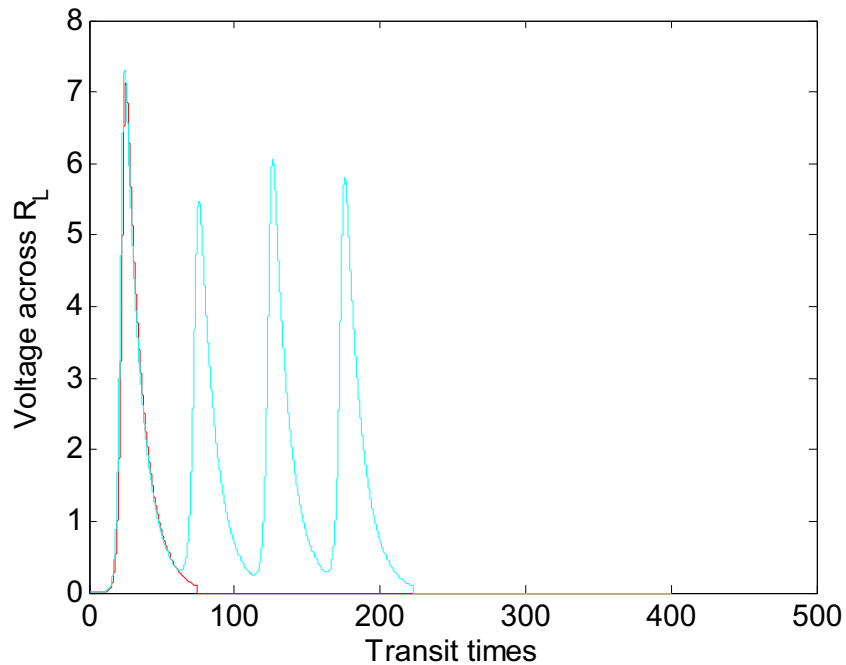
Figure 6. Two Monte-Carlo simulations of the voltage across the loads resistor as a function of the transit time of the electron and hole carriers in the multiplication region. In this example, there are approximately 11 transit times in each RC time constant. For the red curve, quenching occurs after the initial cycle while for the blue curve quenching occurs after the fourth cycle. Note that in both case quenching occurs when the field is in the low cycle.

4. 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.

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., 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).
[6] Groves, C., Tan, C. H., David, J. P. R., Rees, G. J., and Hayat, M. M., "Exponential time response in analogue and Geiger mode avalanche photodiodes," IEEE Trans. Electron Devices, 52, 1527-1534 (2005).