Multiplication theory for dynamically biased avalanche photodiodes: new limits for gain bandwidth product

Majeed M. Hayat* **and David A. Ramirez**

Center for High Technology Materials and Electrical and Computer Engineering Department, University of New Mexico, Albuquerque, New Mexico 87106, USA

hayat@ece.unm.edu

Abstract: Novel theory is developed for the avalanche multiplication process in avalanche photodiodes (APDs) under time-varying reverse-biasing conditions. Integral equations are derived characterizing the statistics of the multiplication factor and the impulse-response function of APDs, as well as their breakdown probability, all under the assumption that the electric field driving the avalanche process is time varying and spatially nonuniform. Numerical calculations generated by the model predict that by using a bit-synchronous sinusoidal biasing scheme to operate the APD in an optical receiver, the pulse-integrated gain-bandwidth product can be improved by a factor of 5 compared to the same APD operating under the conventional static biasing. The bit-synchronized periodic modulation of the electric field in the multiplication region serves to (1) produce large avalanche multiplication factors with suppressed avalanche durations for photons arriving in the early phase of each optical pulse; and (2) generate low avalanche gains and very short avalanche durations for photons arriving in the latter part of each optical pulse. These two factors can work together to reduce intersymbol interference in optical receivers without sacrificing sensitivity.

OCIS codes (040.0040) Detectors; (040.1345) Avalanche photodiodes (APDs); (040.5160) Photodetectors; (060.4510) Optical communications; (060.2330) Fiber optics communications.

References and links

1. B. E. A. Saleh and M. C. Teich, *Fundamentals of Photonics* (Wiley-Interscience, 2007).
2. P. Bhattacharya, *Semiconductor Optoelectronic Devices* (Prentice Hall, 1996).
3. R. G. Smith and S. D. Personick, "Receiver Design for Optical Fiber Communication Systems," *Semiconductor Devices for Optical Communication*, H. Kressel ed. (Springer-Verlag, 1980).
4. B. L. Kasper and J. C. Campbell, "Multigigabit-per-second avalanche photodiode lightwave receivers," J. Lightwave Technol. **5**(10), 1351–1364 (1987).
5. T. Nakata, I. Watanabe, K. Makita, and T. Torikai, "InAlAs avalanche photodiodes with very thin multiplication layer of 0.1 μm for high-speed and low-voltage-operation optical receiver," Electron. Lett. **36**(21), 1807–1808 (2000).
6. P. Sun, M. M. Hayat, B. E. A. Saleh, and M. C. Teich, "Statistical correlation of gain and buildup time in APDs and its effects on receiver performance," J. Lightwave Technol. **24**(2), 755–768 (2006).
7. D. S. G. Ong, J. S. Ng, M. M. Hayat, P. Sun, and J. P. R. David, "Optimization of InP APDs for high-speed lightwave systems," J. Lightwave Technol. **27**(15), 3294–3302 (2009).
8. K. Makita, T. Nakata, K. Shiba, and T. Takeuchi, "40Gbps waveguide photodiodes," NEC J. Adv. Tech. **2**, 234–240 (2005).
9. Y. M. Kang, H. D. Liu, M. Morse, M. J. Paniccia, M. Zadka, S. Litski, G. Sarid, A. Pauchard, Y. H. Kuo, H. W. Chen, W. S. Zaoui, J. E. Bowers, A. Beling, D. C. McIntosh, X. G. Zheng, and J. C. Campbell, "Monolithic germanium/silicon avalanche photodiodes with 340 GHz gain-bandwidth product," Nat. Photonics **3**(1), 59–63 (2009).
10. W. S. Zaoui, H.-W. Chen, J. E. Bowers, Y. Kang, M. Morse, M. J. Paniccia, A. Pauchard, and J. C. Campbell, "Frequency response and bandwidth enhancement in Ge/Si avalanche photodiodes with over 840 GHz gain-bandwidth-product," Opt. Express **17**(15), 12641–12649 (2009).

#160303 - $15.00 USD Received 20 Dec 2011; revised 2 Mar 2012; accepted 9 Mar 2012; published 22 Mar 2012

11. Y. Kang, Z. Huang, Y. Saado, J. Campbell, A. Pauchard, J. Bowers, and M. J. Paniccia, "High performance Ge/Si avalanche photodiodes development in Intel," Opt. Fiber Comm. Conf. & Expo. (OFC/NFOEC), 1–3 (2011).
12. J. C. Campbell, S. Demiguel, F. Ma, A. Beck, X. Guo, S. Wang, X. Zheng, X. Li, J. D. Beck, M. A. Kinch, A. Huntington, L. A. Coldren, J. Decobert, and N. Tscherptner, "Recent advances in avalanche photodiodes," IEEE J. Sel. Top. Quantum Electron. **10**(4), 777–787 (2004).
13. D. C. Herbert and E. T. R. Chidley, "Very low noise avalanche detection," IEEE Trans. Electron. Dev. **48**(7), 1475–1477 (2001).
14. R. J. McIntyre, "Multiplication noise in uniform avalanche diodes," IEEE Trans. Electron. Dev. **13**(1), 164–168 (1966).
15. M. M. Hayat, B. E. A. Saleh, and M. C. Teich, "Effect of dead space on gain and noise of double-carrier-multiplication avalanche photodiodes," IEEE Trans. Electron. Dev. **39**(3), 546–552 (1992).
16. M. M. Hayat and B. E. A. Saleh, "Statistical properties of the impulse response function of double carrier multiplication avalanche photodiodes including the effect of dead space," J. Lightwave Technol. **10**(10), 1415–1425 (1992).
17. G. Agrawal, *Fiber-Optic Communication Systems* (Wiley, 2002).
18. M. M. Hayat, O.-H. Kwon, Y. Pan, P. Sotirelis, J. C. Campbell, B. E. A. Saleh, and M. C. Teich, "Gain-bandwidth characteristics of thin avalanche photodiodes," IEEE Trans. Electron. Dev. **49**(5), 770–781 (2002).
19. N. Yasuoka, H. Kuwatsuka, M. Makiuchi, T. Uchida, and A. Yasaki, "Large multiplication-bandwidth products in APDs with a thin InP multiplication layer," Proc. IEEE Laser & Electro Opt. Soc. Ann. Meeting LEOS' 2003, 999–1000 (2003).
20. N. Namekata, S. Adachi, and S. Inoue, "1.5 GHz single-photon detection at telecommunication wavelengths using sinusoidally gated InGaAs/InP avalanche photodiode," Opt. Express **17**(8), 6275–6282 (2009).
21. J. Zhang, P. Eraerds, N. Walenta, C. Barreiro, R. Thew, and H. Zbinden, "2.23 GHz gating InGaAs/InP single-photon avalanche diode for quantum key distribution," Proc. SPIE **7681**, 76810Z1–76810Z8 (2010).
22. M. A. Saleh, M. M. Hayat, B. E. A. Saleh, and M. C. Teich, "Dead-space-based theory correctly predicts excess noise factor for thin GaAs and AlGaAs avalanche photodiodes," IEEE Trans. Electron. Dev. **47**(3), 625–633 (2000).
23. J. S. Ng, C. H. Tan, J. P. David, G. Hill, and G. J. Rees, "Field dependence of impact ionization coefficients in $In_{0.53}Ga_{0.47}As$," IEEE Trans. Electron. Dev. **50**(4), 901–905 (2003).
24. F. Osaka, T. Mikawa, and T. Kaneda, "Electron and hole ionization coefficients in (100) oriented $Ga_{0.33}In_{0.67}As_{0.70}P_{0.30}$," Appl. Phys. Lett. **45**(3), 292–293 (1984).
25. M. M. Hayat, W. L. Sargeant, and B. E. A. Saleh, "Effect of dead space on gain and noise in Si and GaAs avalanche photodiodes," IEEE J. Quantum Electron. **28**(5), 1360–1365 (1992).
26. M. M. Hayat, O.-H. Kwon, S. Wang, J. C. Campbell, B. E. A. Saleh, and M. C. Teich, "Boundary effects on multiplication noise in thin heterostructure avalanche photodiodes," IEEE Trans. Electron. Dev. **49**(12), 2114–2123 (2002).
27. C. H. Tan, P. J. Hambleton, J. P. R. David, R. C. Tozer, and G. J. Rees, "Calculation of APD impulse response using a space- and time-dependent ionization probability distribution function," J. Lightwave Technol. **21**(1), 155–159 (2003).
28. L. J. J. Tan, J. S. Ng, C. H. Tan, and J. P. R. David, "Avalanche noise characteristics in submicron InP diodes," IEEE J. Quantum Electron. **44**(4), 378–382 (2008).

1. Introduction

Avalanche photodiodes (APDs) are the photodetectors of choice in direct-detection high-speed lightwave communication systems. An APD can provide high optoelectronic gains through a cascade of impact ionizations, a feature that is not present in the simpler *pin* photodetector [1, 2]. The gain offered by the APD improves the receiver sensitivity as it amplifies the photocurrent, thereby reducing the relative effect of Johnson noise in the preamplifier stage of an optical receiver and improving the receiver's sensitivity [3, 4]. In addition to their high sensitivity, APD-based receivers are highly cost effective compared to receivers that employ a combination of optical pre-amplification and a *pin* photodetector [5]. However, the APD's avalanche duration, the time it takes for the chain of all impact ionizations to complete each time a photon-generated carrier triggers and avalanche, can limit the APD's gain-bandwidth product and lead to intersymbol interference (ISI) in optical receivers [6]. Since both the gain and the avalanche duration are stochastic, statistical fluctuations of these quantities must also be factored in when assessing the avalanche-gain versus ISI-penalty tradeoff [6, 7].

While separate absorption and multiplication (SAM) InP-InGaAs APDs have been successfully deployed in 10 Gbps lightwave systems, their utility at higher bit rates has been limited due to their long avalanche durations. Meanwhile, numerous efforts have been reported in developing APDs for 40 Gbps systems. For example, Makita *et al*. [8]

#160303 - $15.00 USD Received 20 Dec 2011; revised 2 Mar 2012; accepted 9 Mar 2012; published 22 Mar 2012

demonstrated a sensitivity of −19dBm at 40 Gbps with a bit-error rate of 10^{-10}, providing approximately a 9dBm improvement over conventional *pin* diode. This was achieved by including a trans-impedance amplifier with tunable response to boost the gain-bandwidth-product (GBP) from 140 to 270 GHz and operating the APDs with avalanche gain values of 3 to 10. Notably, the biggest breakthrough came as Si/Ge APDs were demonstrated in 2009. Kang *et al*. [9] reported a GBP of 340GHz, followed by Zaoui *et al*. [10] demonstrating a Si-Ge APD with a GBP of 840 GHz operating at 1.31 μm. The high GBP associated with the Si multiplication layer is attributed to the favorable ionization properties of Si (the electron-to-hole ionization coefficient ratio, *k*, is much less than unity in Si). However, these devices have lower normal-incidence sensitivity compared with InGaAs-InP APDs of the same absorption-layer thickness [11] due to a combination of high dark currents, resulting from the low bandgap and high level of intrinsic carrier concentration of Ge, and reduced quantum efficiency of Ge at longer telecom wavelengths due to its low absorption coefficient at 1.55 μm ($3.5\text{x}10^2$ cm^{-1} for Ge compared to $6.1\text{x}10^3$ cm^{-1} for InGaAs). To compensate for the reduced photoresponsivity of Ge at longer telecom wavelengths, waveguide structures have been proposed and unity-gain bandwidths of 23 GHz and 29.5 GHz have been achieved for evanescent-coupled and butt-coupled Si/Ge APDs, respectively [11]. While Si/Ge APDs show promise due to their compatibility with SOI CMOS and the almost transit-time-limited avalanche duration associated with Si (since hole ionization is minimal), and despite all numerous other efforts in the past two decades or so that targeted new APD materials and structures [12], to date there are no commercial APDs for 1.55 μm 40-Gbps communication.

In this paper we present a novel theory to rigorously analyze the APD's performance under dynamic biasing and project substantial improvements in the GBP offered by dynamic biasing compared to static biasing. While sinusoidal biasing has been reported by Herbert and Chidley [13] as a way to reduce excess noise in APDs, to the best of our knowledge the concept of dynamic biasing has not been explored for linear-mode operation as a way to improve bandwidth. We will show that sinusoidal biasing of the APD (where the bias is synchronous with the optical pulse stream and has a properly selected DC level, peak-to-peak value and phase) can offer up to five-fold increase the effective gain-bandwidth product. The sinusoidal waveform is selected here for its ease of practical implementation (not considered in this paper) as well as its minimal distortion at high speeds. In making such prediction, we present the first theory for the avalanche multiplication process under dynamic electric fields, which is a major expansion beyond the theories for impact ionization under the assumption of spatially non-uniform yet static electric fields [14, 15]. Specifically, in this paper we extend the recurrence theory for the statistics of the stochastic gain [15] and the impulse-response function [16] of APDs to the general setting where the electric field is dynamic and arbitrary. The dynamic impact ionization theory is equally applicable to the Geiger-mode operation of APDs. To this end, we also derive recursive equations that characterize the breakdown probability under dynamic biasing.

The remainder of this paper is organized as follows. In Section 2 we give a qualitative description of the potential benefits of dynamic biasing in improving the gain-bandwidth product. In Section 3 we extend the key concepts related to impact ionization to a dynamic-electric-field setting. The avalanche multiplication theory under dynamic electric fields is presented in Section 4 while additional generalizations are presented in the Appendix. Numerical results are presented in Section 5, and the conclusions are drawn in Section 6.

2. Potential for breaking the traditional gain-bandwidth limits through dynamic biasing

In the traditional linear-mode setting, an APD is operated under a static (fixed) reverse bias, which is typically selected to yield the optimal multiplication factor that maximizes the receiver sensitivity [17] as it balances speed, avalanche amplification (and hence Johnson-noise suppression), and excess noise resulting from the impact ionization process being stochastic. For values of the mean gain that are of interest to optical communication (typically >10), the bandwidth of an APD is limited by the avalanche duration – the time needed for all the impact ionizations to terminate. It is well known that there is a tradeoff between the

#160303 - $15.00 USD Received 20 Dec 2011; revised 2 Mar 2012; accepted 9 Mar 2012; published 22 Mar 2012

APD's avalanche-duration-limited bandwidth and its avalanche multiplication factor [18]. For example, for InGaAs/InP APDs the GBP has been limited to around 170 GHz [19]. However, if bit-synchronous, periodic dynamic biasing is considered in lieu of the traditional static bias, as presented in this paper, then the impact ionization process may be modulated to achieve a substantial improvement in the GBP and therefore reduced ISI.

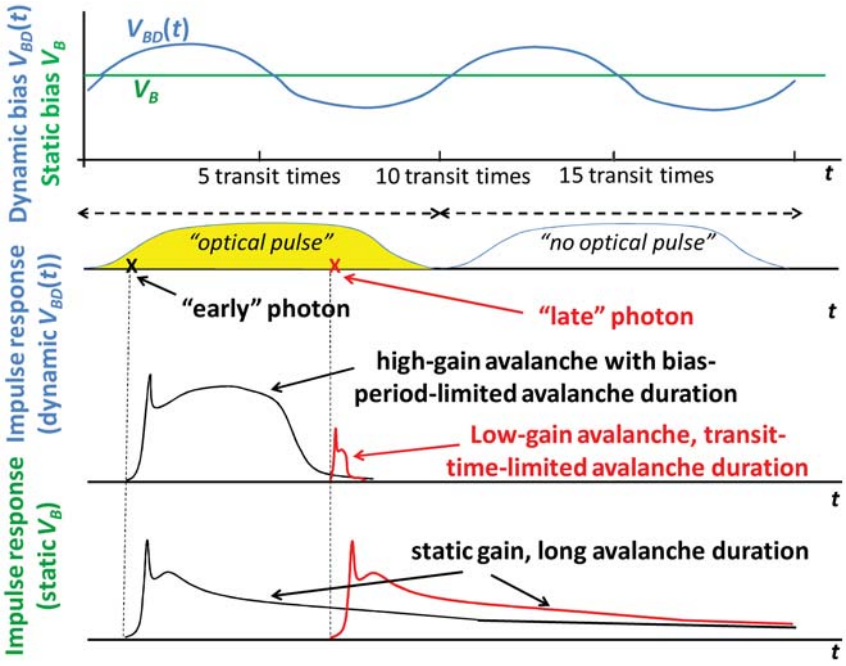

Fig. 1. Schematic of the proposed dynamic biasing approach (blue curve) repeated periodically over optical-pulse intervals. The green straight line represents the traditional constant bias. The periodic change in the reverse bias from the first to the second half of the optical-pulse period causes (1) photons that arrive early in the pulse window to trigger high gains but pulse-limited durations, and (2) late photons to trigger avalanches with low gains and almost transit-time limited avalanche durations. Such modulation of impact ionization results in a much higher average GBP compared to the conventional biasing scheme.

Before presenting the predictions based on mathematical analysis, we will motivate the consideration of a dynamic-biasing scheme by projecting its properties in a qualitative manner. Consider Fig. 1, which depicts a sinusoidal biasing scheme (blue curve) that is frequency-matched and synchronized with the optical bit stream. The photocurrent generated by each optical pulse under such dynamic biasing has the following two properties. First, photons that arrive *early* in the optical pulse experience a period of high electric field in the multiplication region of the APD, where they can generate a strong avalanche current in the early phase of the optical-pulse interval, as depicted schematically by the black curve corresponding to the dynamic-bias response in Fig. 1. Next, as a low electric field period follows the high-field phase within the same optical pulse, carriers in the multiplication region undergo a much weakened impact ionization process, which leads to the termination of the avalanche current with a high probability. Thus, high-gain avalanche pulses are triggered by the "early" arriving photons in each optical pulse, albeit with a much reduced avalanche duration due to the weakening of the impact ionization in the second phase of the bias period. Second, photons that arrive *late* in the optical pulse period are still detected as the APD remains reverse biased throughout the bias period. However, the resulting avalanche gain is very low and the avalanche durations they trigger are very short, as depicted schematically by the red curve correspondig to the dynamic bias response in Fig. 1. The avalanche pulses in the traditional constant-biasing scheme are also shown schematically in Fig. 1 for comparison.

#160303 - $15.00 USD Received 20 Dec 2011; revised 2 Mar 2012; accepted 9 Mar 2012; published 22 Mar 2012

We would like to mention that while a sinusoidal-gating approach has been proposed for Geiger-mode APDs in the context of gated photon counting [20, 21] its rationale is different from the linear-mode dynamic biasing presented here. The purpose of sinusoidal-gating Geiger-mode operation is to force quenching of the avalanche pulse after each detection-gate (high cycle of the sinusoidal bias) and therefore minimize the total number of multiplications, which, in turn, would reduce afterpulsing. However, the rationale is totally different from the linear-mode dynamic biasing approach proposed here. Specifically, photon counting with sinusoidal-gating is a binary detection problem: the APD is responsive to only one photon per gate. In contrast, in the proposed linear-mode dynamic biasing approach each and every photon in the optical pulse that is absorbed by the photodetector contributes to the analog photocurrent.

3. Impact ionization under dynamic electric fields

Suppose that a time-varying bias, $V_{BD}(t)$, $t \geq 0$, is applied to an APD. Consider a charge-depleted multiplication region of the APD extending from $x = 0$ to $x = w$, as shown in Fig. 2, with the convention that the electric field is pointing in the negative x direction. Let $\mathrm{E}(x,t)$ denote the dynamic electric field in the multiplication region at position x and at time t. If the field is spatially uniform, then $\mathrm{E}(x,t) \equiv \mathrm{E}(t) = V_{BD}(t)/w$. Suppose that a parent hole (electron) is created at an arbitrary location x in the multiplication region of the APD, and assume that the field is sufficiently high so as conduction-band electrons and valence-band holes travel at their material-specific saturation velocities, v_e and v_h, respectively. As the hole travels the multiplication region, it can impact ionize at a stochastic location, say ξ, and at time $\tau = (\xi - x)/v_h$. The ability of a carrier to impact ionize is probabilistic and depends upon the time- and position-dependent electric field as well as the carrier's history of prior impact ionizations (detailed description deferred to Subsection 3.1). Upon impact ionization, the parent hole is replaced with two valence-band offspring holes and a conduction-band offspring electron. Each offspring carrier proceeds, as its parent, to further impact ionize, and so on. As this process continues, it may or may not terminate, depending on the field, device and material properties. The stochastic *dynamic multiplication factor* is the total number of electron-hole pairs generated as a result of a parent carrier in the presence of the dynamic field; it can be either finite or infinite. On the other hand, the *dynamic avalanche duration* is the time measured from the creation of the parent carrier to the time when all carriers have exited the multiplication region. In Subsections 4.1 and 4.2 we will derive equations that enable us to calculate the statistics of the mean multiplication factor and the mean impulse response function (integral equations for the excess-noise factor and the breakdown probability under the dynamic fields are presented in the Appendix). However, before doing so we will need to extend the notions of the ionization coefficient, dead space, and the probability density function of the carrier's free path (prior to ionization) to a dynamic-electric-field setting.

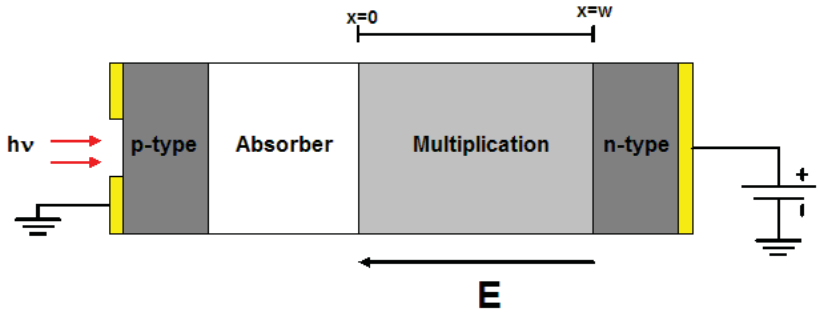

Fig. 2. Schematic of a separate-absorption-multiplication (SAM) APD.

We will assume throughout this paper that the change in the electric field is slow with respect to time for the carrier populations to reach equilibrium with the electric field for each

#160303 - $15.00 USD Received 20 Dec 2011; revised 2 Mar 2012; accepted 9 Mar 2012; published 22 Mar 2012

carrier so that the ionization coefficient of a carrier depends adiabatically on the instantaneous electric field E(x,t) at the point where and when it impact ionizes. We define $\alpha(x,t)$ and $\beta(x,t)$ as the electron and hole *time-varying* non-localized ionization coefficients associated with carriers at location x in the multiplication region and at time t. These are the ionization coefficients for those carriers that have already traveled the dead space, which is the minimum distance a carrier must travel before it acquires sufficient energy to effect an impact ionization. Following the model for non-localized ionization coefficients under a static electric field [22] and by replacing the static field with its dynamic counterpart, the dynamic ionization coefficients are given by

$$\alpha(x,t) = A_e \exp\left(-[E_{c,e} / \mathrm{E}(x;t)]^{m_e}\right), \tag{1a}$$

$$\beta(x,t) = A_h \exp\left(-[E_{c,h} / \mathrm{E}(x;t)]^{m_h}\right), \tag{1b}$$

where the material-specific constants A, E_c, and m, are known for various III-V materials [22–24].

The probability density function of the carrier's free path in a dynamic-field setting depends upon the non-local time-varying ionization coefficients as well as the carrier's history. In particular, a newly born carrier cannot impact ionize before travelling a dead space. Consider an electron and hole created at location x and of age s relative to the launch instant of the dynamic electric field (at $t = 0$), and let X_e and X_h be their stochastic free-path distances to their *first* impact ionization. As it turns out, the *age* of a carrier will play a key role in the formulation of the theory for avalanche multiplication under dynamic fields, as described in Section 4. Under a dynamic electric field, the probability density function of the location, ξ, of the first ionization by a parent carrier of age s and at location x is zero before the dead space is travelled. After the dead space, however, it is exponential with a nonuniform rate that not only depends upon the path of the carrier from its birth location x to ξ but also on the history of the time- and space-varying electric field from the birth instant of the carrier to the instant of its first impact ionization. Along these lines, we can extend the shifted exponential model for the carrier's free path [25] and obtain the *age-dependent* probability density function of X_e and X_h as $h_e(\xi;x,s)$ and $h_h(\xi;x,s)$, respectively,

$$h_e(\xi;x,s) = \begin{cases} \alpha(\xi, s+(\xi-x)/v_e)\exp\left\{-\int_{x+d_e(x,s)}^{\xi} \alpha(\sigma, s+(\sigma-x)/v_e)d\sigma\right\}, & \xi \geq x + d_e(x,s) \\ 0, & \text{otherwise} \end{cases} \tag{2a}$$

and

$$h_h(\xi;x,s) = \begin{cases} \beta(\xi, s+(x-\xi)/v_h)\exp\left\{-\int_{\xi}^{x-d_h(x,s)} \beta(\sigma, s+(x-\sigma)/v_h)d\sigma\right\}, & \xi \leq x - d_h(x,s) \\ 0, & \text{otherwise} \end{cases}, \tag{2b}$$

where $d_e(x,s)$ and $d_h(x,s)$ represent the *age-dependent* dead spaces for an electron and hole, respectively, that were created at position x and of age s. Here, $h_e(\xi;x,s)\Delta\xi$ approximates the probability that an electron born at location x and with age s (relative to the launch instant of the electric field) impact ionizes for the first time anywhere in the location ξ to $\xi + \Delta\xi$. The electron's age-dependent dead space is computed by equating the ionization threshold energy to the energy gained from the ballistic transport of the carrier in the time- and space-varying electric field. Thus, the age-dependent dead space is the minimum d value that satisfies the equation

$$q\int_{x}^{x+d} \mathrm{E}(y, s+(y-x)/v_e)dy = \mathcal{E}_{th,e}(x+d), \tag{3a}$$

where for any $0 \leq x \leq w$, $\mathcal{E}_{th,e}(x)$ is the ionization threshold energy for electrons for the material at position x in the multiplication region. Similarly, the hole's age-dependent dead space is computed as the minimum d value that satisfies the equation

#160303 - $15.00 USD Received 20 Dec 2011; revised 2 Mar 2012; accepted 9 Mar 2012; published 22 Mar 2012

$$q\int_{x-d}^{x} \mathrm{E}(y, s+(x-y)/v_h)dy = \mathcal{E}_{th,h}(x-d), \tag{3b}$$

where $\mathcal{E}_{th,h}(x)$ is the ionization threshold energy for holes for the material at position x in the multiplication region. (Note that Eqs. (3a) and (3b) above respectively collapse to Eqs. (4) and (5) in [26] when the field is static.) Note that given the knowledge of the material composition in the multiplication region and the electric field profile and its evolution in time, the age-dependent dead space can be calculated for all x and s.

4. Avalanche multiplication theory under dynamic electric fields

The recurrence theory for the avalanche multiplication, including the statistics of the gain and impulse-response function, under non-uniform, static electric fields was originally formulated by Hayat *et al*. in [15, 16] and later extended to accommodate stochastic carrier velocity by Tan *et al*. [27]. Here we introduce a generalization of the recurrence theory under dynamic electric fields. As we alluded to before, the age of the parent carrier that triggers the avalanche process is key in modeling the avalanche multiplication process when the field is allowed to be time varying because carriers with different ages will experience different dynamical electric field ahead of them as they travel the multiplication region. Specifically, if we assume a causal and spatially non-uniform electric-field, $\mathrm{E}(x,t)$, launched at time $t = 0$, then an electron born at location x with age 0 will experience this time-varying field in its lifetime. In contrast, if an electron is born at location x with age s (relative to the launch time of the field at $t = 0$) then it will experience in its lifetime a *clipped* version of the dynamic field starting at $t = s$, namely $\mathrm{E}(x,t)u(t-s)$, where $u(.)$ is the unit-step function. To take the carrier's "age" into account while modeling the avalanche multiplication process, we must formulate a model in which the ionization probability is parameterized by the age of the carrier at the time when it triggers the avalanche process. We term such formulation an *age-dependent* analysis. In what follows (including the Appendix), we will derive sets of age-dependent recurrence equations that enable us to calculate the mean gain, the excess noise factor, the probability mass function of the gain, the mean of the impulse-response function, as well as the breakdown probability, all under a dynamic electric field.

4.1. Mean gain

We define $Z(x,s)$ ($Y(x,s)$) as the totality of all electrons and holes, including the parent carrier, initiated by an electron (hole), injected at location x with age s. Note that by convention $Z(w,s) = Y(0,s) = 1$ since an electron (hole) placed at the right (left) edge of the multiplication region will exit the multiplication region without ionizing. Now consider a parent electron-hole pair at location x and of age s. The age-dependent stochastic multiplication factor, $M(x,s)$, defined as the total number of electron-hole pairs generated as a result of an electron-hole pair whose initial location in the multiplication region is x and whose ages are s, is simply

$$M(x,s) = 0.5\left[Z(x,s) + Y(x,s)\right]. \tag{4}$$

Note that in the special case of a SAM InGaAs-InP APD, for which holes are injected at the edge of the multiplication region at $x = w$ and with age $s \geq 0$ (recall that the age is always measured with respect to the launch instant of the dynamic electric field at time $t = 0$), then the stochastic age-dependent gain is given by

$$G(s) = M(w,s) = 0.5\left[1 + Y(w,s)\right]. \tag{5}$$

We now derive equations that allow us calculate the statistics of the quantities $Z(x,s)$ and $Y(x,s)$. Once we find the first and second moments of Z and Y we can relate them to the age-dependent mean gain via Eqs. (4) and (5), and to the age-dependent excess noise factor as shown in the Appendix. Suppose that the first ionization for a parent electron of age s positioned at location x occurs at location $X_e = \xi$, where $x \leq \xi \leq w$. Note that the instant of this

#160303 - $15.00 USD Received 20 Dec 2011; revised 2 Mar 2012; accepted 9 Mar 2012; published 22 Mar 2012

ionization (measured from the launch instant of the dynamic electric field at time $t = 0$) is $s + \tau$, where $\tau = (\xi - x)/v_e$ is the transit time of this electron in the multiplication region. Note that the two offspring electrons at ξ, who are born with common age $s + \tau$, will independently generate $Z_1(\xi, s + \tau)$ and $Z_2(\xi, s + \tau)$ carriers, respectively. On the other hand, the offspring hole, whose age at birth is also $s + \tau$, will generate $Y(\xi, s + \tau)$ carriers independently of $Z_1(\xi, s + \tau)$ and $Z_2(\xi, s + \tau)$. Thus, conditional on the event that the first impact ionization for the parent electron occurs at location ξ, the sum $Z_1(\xi, s + \tau) + Z_2(\xi, s + \tau) + Y(\xi, s + \tau)$ will simply amount to $Z(x,s)$. Since we can always express the mean of $Z(x,s)$ as the expectation of the conditional mean (given that the first ionization of the parent electron occurs at location X_e), we can write $\mathrm{E}[Z(x,s)]$ as an iterated expectation $\mathrm{E}[\mathrm{E}[Z(x,s) \mid X_e]]$ by conditioning first on the location of the first ionization. Hence,

$$\begin{aligned} \mathrm{E}\left[Z(x,s)\right] &= \mathrm{E}[\mathrm{E}\left[Z(x,s) \mid X_e\right]] \\ &= \mathrm{E}\left[Z_1\left(X_e, s + (X_e - x)/v_e\right) + Z_2\left(X_e, s + (X_e - x)/v_e\right) + Y\left(X_e, s + (X_e - x)/v_e\right)\right], \end{aligned} \tag{6}$$

where the symbol "E" denotes expectation. Now if we define the notation for the mean of the quantities as $z(x,s) = \mathrm{E}[Z(x,s)]$ and $y(x,s) = \mathrm{E}[Y(x,s)]$, then the expression on the right hand side of Eq. (6) can be cast as

$$z(x,s) = \mathrm{E}\left[2z\left(X_e, s + (X_e - x)/v_e\right) + y\left(X_e, s + (X_e - x)/v_e\right)\right]. \tag{7}$$

By using the pdf in (2a) to carry out the averaging over X_e in (7) while taking into account the scenario when the parent electron may not impact ionize at all, we obtain the integral equation

$$\begin{aligned} &z(x,s) = \int_w^\infty h_e(\xi; x, s)\,d\xi + \int_x^w [2z(\xi, s + (\xi - x)/v_e) + y(\xi, s + (\xi - x)/v_e)]\,h_e(\xi; x, s)\,d\xi \\ &0 \le x \le w, \quad s \ge 0. \end{aligned} \tag{8a}$$

The first term on the right side of Eq. (8a) captures the case when the parent electron exits the multiplication region without impact ionizing.

We can repeat the above argument while starting from a parent hole at ξ and of age s instead of a parent electron. In this case, we realize that the stochastic location of the first ionization of the parent hole, X_h, can be in the interval $[0, x]$ (instead of $[x, w]$ as in the case of the parent electron). The resulting integral equation for $y(x,s)$ is

$$\begin{aligned} &y(x,s) = \int_x^\infty h_h(\xi; x, s)\,d\xi + \int_0^x [2y(\xi, s + (x - \xi)/v_h) + z(\xi, s + (x - \xi)/v_h)]\,h_h(\xi; x, s)\,d\xi \\ &0 \le x \le w, \quad s \ge 0. \end{aligned} \tag{8b}$$

The pair of linear coupled integral equation in (8) can be solved numerically, e.g., by the method of iterations.

The age-dependent mean multiplication factor, $m(x,s) = \mathrm{E}[M(x,s)]$, can be calculated using the identity $m(x,s) = 0.5[z(x,s) + y(x,s)]$, and the age-dependent mean gain, $g_a(s) = \mathrm{E}[G(s)]$, is simply $m(w,s)$ in the case of a hole-injection APD. As a special case, the usual mean gain, g, of an APD under a static bias is simply $g = g_a(0)$.

It is also possible to derive integral equations characterizing the excess noise factor, the probability mass function of the multiplication factor, as well as the probability of breakdown under dynamic biasing. Since these quantities are not used in gain-bandwidth product calculations and the results presented in Section 5, they are included for completeness in the Appendix.

4.2. Mean impulse response function

We begin by defining $I_e(t,x,s)$, the age-dependent *stochastic* impulse-response function at time t initiated by an *electron* injected at location x and with age s. Similarly, $I_h(t, x, s)$ is the stochastic age-dependent impulse-response function at time t, initiated by a *hole* injected at location x with age s. Mathematically, if we define $i_e(t,x,s)$ and $i_h(t,x,s)$ as the mean quantities of $I_e(t,x,s)$ and $I_h(t,x,s)$, respectively, then we can write that *conditional* mean of $I_e(t,x,s)$ given that the first ionization of the parent electron occurs at location X_e as

#160303 - $15.00 USD Received 20 Dec 2011; revised 2 Mar 2012; accepted 9 Mar 2012; published 22 Mar 2012

$$\mathrm{E}\left[I_e\left(t,x,s\right) \mid X_e \right] = 2i_e(t, X_e, s+\left(X_e - x\right)/v_e) + i_h(t, X_e, s+\left(X_e - x\right)/v_e), \quad (9a)$$

where $x \le X_e \le w$. On the other hand, when no ionization occurs (when $X_e > w$), then

$$\mathrm{E}[I_e\left(t,x,s\right) \mid X_e > w] = \left(qv_e / w\right)\left\{u\left(t\right) - u\left(t - \left(w-x\right)/v_e\right)\right\}, \quad (9b)$$

which is simply a rectangular pulse of duration equal to the transit time of the parent electron at x as it drifts across the remainder of the multiplication region. When we average Eq. (9) over all possible values of X_e, we obtain the following integral equation

$$\begin{aligned} i_e(t,x,s) &= (qv_e / w)\left[u(t) - u\left(t-(w-x)/v_e\right)\right]\int_w^\infty h_e(\xi;x,s)d\xi \\ &+\int_x^{\min(w,x+v_e t)}\left[2i_e\left(t-(\xi-x)/v_e,\xi,s+(\xi-x)/v_e\right)+i_h\left(t-(\xi-x)/v_e,\xi,s+(\xi-x)/v_e\right)\right]h_e(\xi;x,s)d\xi \end{aligned} \quad (10a)$$

A similar integral equation can be obtained for $i_h(t,x,s)$:

$$\begin{aligned} i_h(t,x,s) &= (qv_h / w)\left[u(t) - u(t - x/v_h)\right]\int_x^\infty h_h(\xi;x,s)d\xi \\ &+\int_{\max(0,x-v_h t)}^{x}\left[2i_h\left(t-(x-\xi)/v_h,\xi,s+(x-\xi)/v_h\right)+i_e\left(t-(x-\xi)/v_h,\xi,s+(x-\xi)/v_h\right)\right]h_h(\xi;x,s)d\xi \end{aligned} \quad (10b)$$

The two coupled integral equations can be solved numerically using a simple iterative approach. The mean impulse-response function, $i(t,s)$, (in the case of a hole injection to the multiplication region at $x = 0$) for a photon absorbed at time s is then obtained using $i(t,s) = i_h(t,w,s)$.

We note that when the field is static, the dependence on the age variable s will absent and all the integral equations in this section collapse to their static-field counterparts. For example, Eqs. (10a) and (10b) collapse respectively to Eqs. (5) and (6) in [27], and Eqs. (8a) and (8b) collapse respectively to (10) and (11) in [15].

4.3. Pulse response, pulse-response bandwidth and pulse-integrated mean gain

We saw earlier that under dynamic biasing, the mean impulse response function, $i(t,s)$, is dependent on the birth time s of the photogenerated parent carrier triggering the avalanche (or equivalently on the arrival time of the absorbed photon if we ignored the time it takes the carrier to be transported from the absorber to the edge of the multiplication region). Now in on-off keying optical communication, photons arrive randomly within each optical pulse according to a temporal distribution that is governed by the pulse's power profile within the bit. Hence, the appropriate quantity to look at when assessing ISI would be the *pulse-response* function, $i_p(t)$, rather than an impulse response. If we assume that the time origin (namely the launch time of the electric field) is the beginning of the bit, we can express $i_p(t)$ as

$$i_p(t) = \int_0^T i(t,s)\, p_{ph}(s)ds, \quad (11)$$

where $p_{ph}(s)$ is the probability density function of arriving photons within an optical pulse of duration T. Note that if the electric field is static, then $i(t,s)$ is simply $i(t\text{-}s)$, and $i_p(t)$ would become simply the convolution of $i(t)$ and $p_{ph}(t)$. An alternative way to view $i_p(t)$ is to regard it as a photon-arrival time averaged impulse response. Since early and late photons have long and short impulse responses, respectively, it would make sense to look at the average of the impulse response functions over all possible photon arrival times within each received optical pulse. In any event, by calculating the 3dB bandwidth of the Fourier transform of $i_p(t)$, we can obtain the pulse-response bandwidth, B_p, which combines the APD's avalanche-duration limited bandwidth with the bandwidth of the optical pulse in each bit of duration T. By defining the *pulse-averaged gain*,

$$\overline{g}_p = T^{-1}\int_0^T g_a(s)ds \quad (12)$$

#160303 - $15.00 USD Received 20 Dec 2011; revised 2 Mar 2012; accepted 9 Mar 2012; published 22 Mar 2012

which is simply the average of the age-dependent mean gain, $g_a(s)$, we can introduce the *pulse-integrated gain-bandwidth product*, GBP_p, as

$$GBP_p = g_p B_p. \tag{13}$$

Note that GBP_p collapses to the usual gain bandwidth product whenever the biasing is static.

We end this section by reiterating that under bit-synchronous periodic biasing, the bit-integrated photocurrent (or charge per bit) is proportional to the product of the pulse-gain factor $\overline{g}_p$ and the average number n of detected photons in the optical pulse in each bit. In other words, under bit-synchronous periodic biasing, the output (charge) of an integrate-and-dump receiver *remains proportional* to the energy in the optical pulse in each bit. Thus, while the dynamically biased APD may not be directly applicable to simple analog detection since the APD's avalanche gain is time variant, it is a perfect fit to digital communications as the integrate-and-dump receiver maintains its linearity.

5. Results

We will limit our results to sinusoidal biasing

$$V_{BD}(t) = B + C\sin(2\pi f_c t + \phi), \tag{14}$$

where f_c is set to be equal to the bit transmission rate, $f_c = R$. The quantities B, C and ϕ are free parameters that can be selected to control the overall multiplication factor and also to maximize the benefit of dynamic biasing in minimizing the tail of the pulse response of the APD. As described earlier, it is implicitly assumed that the bias signal is synchronous with the optical pulse stream. In practice, a clock recovery circuit and a phase lock loop must be employed to maintain synchronization. In addition, for simplicity we shall assume a spatially uniform electronic field, $E(t) = V_{BD}(t)/w$, in the multiplication region of the APD.

5.1. Mean impulse-response function under sinusoidal biasing

In our calculations, we considered a SAM APD, as shown in Fig. 2, with an InP multiplication layer of width w = 200 nm. The ionization parameters for InP were extracted from [28] and they are listed for convenience in Table 1. The electron and hole saturation velocities are assumed as 6.7 x 10^6 cm/s. For reference, we first calculated the mean impulse-response function, triggered by a hole injected at position $x = w$ = 200 nm, under a static bias of V_B = 14.30 V. The 3dB bandwidth was then extracted from the Fourier transform of the impulse response. The calculations of the mean impulse-response function were performed according to the integral Eqs. (6), (10), (11) and (12) in [16]; these equations were solved numerically using a simple linear-iteration method. Figure 3 shows five mean impulse-response functions triggered by holes of different trigger instants (i.e., with different ages). Naturally, aside from the time delays corresponding to the instants at which the holes trigger the avalanche, the impulse-response functions are all identical in shape. We also calculated the mean gain by solving the integral Eqs. (10) and (11) in [15]; these yielded a mean gain of g = 28 for this device under the same constant bias of V_B = 14.30 V. The usual gain-bandwidth product for this device is found to be 238 GHz, which is the same as the pulse-integrated gain-bandwidth product, GBP_p, since the field is static in this case.

Table 1. Ionization Parameters for InP [28]

	A [cm^{-1}]	E_c [V/cm]	m	E_{th} [eV]
Electron	1.41×10^6	1.69×10^6	1.23	2.8
Hole	2.11×10^6	1.77×10^6	1.15	3.0

The calculation of the age-dependent impulse response function in the case of the dynamic bias is shown in Fig. 4 (top) for different values of the age variable s. These curves were obtained by solving (10a) and (10b) numerically using the method of iterations. The dynamic electric field profile is also shown in Fig. 4 (bottom). The transit time is simply v/w = 2.985 ps. Note that the period of dynamic bias is equal to 5.5 transit times. (Here, the width of the

#160303 - $15.00 USD Received 20 Dec 2011; revised 2 Mar 2012; accepted 9 Mar 2012; published 22 Mar 2012

optical pulse is 8.3 ps, as in a 60-Gbps NRZ bit stream.) The sinusoidal-biasing parameters were selected as follows: B = 13 V, C = 6 V, and $\phi = 0$. These parameters were chosen, in part, so that the pulse-integrated gain, $\overline{g}_p$, is approximately 28, making the static and dynamic biasing schemes equivalent in our example from a "total multiplication-gain" perspective. For example, the curve with s = 4 transit times (green curve) corresponds to a mean impulse response due to a parent hole of age 4 transit times after the launch of the bias at time $t = 0$. On the other hand, the curve with $s = 0$ transit times (red curve) corresponds to a parent hole of age 0. It is important to note the change in the shape of the impulse response function as a result of dynamic biasing. Unlike the static-bias case, the impulse response corresponding to $s = 0$ no longer peaks at the parent hole's transit time (i.e., one transit time after the trigger instant) but instead it peaks at a later time in response to the increase in the instantaneous electric field in the multiplication region. Another interesting point is that the weakest impulse response is obtained when the age of the hole is approximately s = 2.6 transit times (magenta curve). This observation suggests that the sinusoidal dynamic bias should be delayed appropriately in order for the last photon in the pulse to have the weakest impulse response. The point to be made here is that the time delay in the dynamic bias must be optimized to produce the best pulse-integrated gain-bandwidth product. We have found empirically that a phase angle of $\phi = \pi/3$ gives good results.

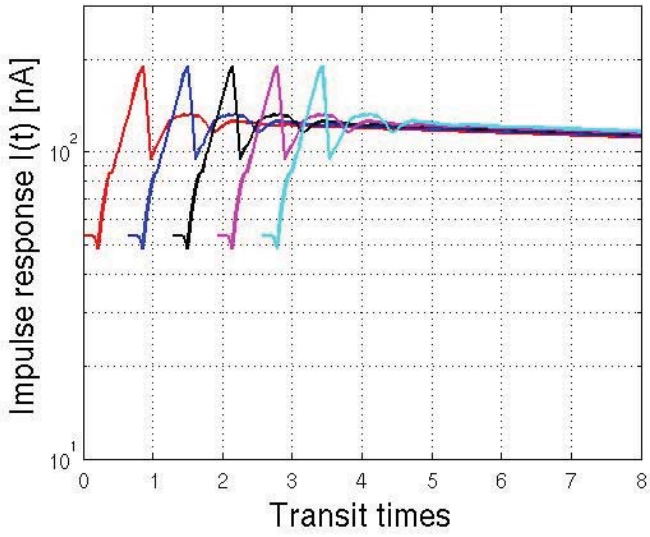

Fig. 3. Mean impulse-response functions triggered by a hole under a constant electric field of V_B = 14.30 V as a function of the integer multiples of the transit time, which is simply v/w = 2.985 ps. Different curves correspond to different ages (in transit times) of the initiating hole: red: age $s = 0$; blue: age s = 0.75; black: age s = 1.25; magenta: age s = 1.9; and cyan: age s = 2.6. The mean gain is calculated as 28 for each case. As expected these curves are simply shifted versions of one another.

Figure 5 shows the calculated age-dependent impulse-response function under sinusoidal biasing using the same dc and ac voltages as those used in Fig. 4 but with the phase angle selected as $\phi = \pi/3$. It is to be noted that due to the combination of the modulating field and the initial time delay in the bias waveform, for small values of the parent-hole's age (or equivalently for early photons) the tail of the impulse response is far shorter than that for the static-bias impulse response. Moreover, the gain associated with small s values is quite high as seen below. This is due to the rise in the field initially, where a high gain is built up, followed by a drop in the field causing the shortening of the impulse response as the probability of the avalanche terminating is high. For example, when $s = 0$ the age-dependent

#160303 - $15.00 USD Received 20 Dec 2011; revised 2 Mar 2012; accepted 9 Mar 2012; published 22 Mar 2012

bandwidth is 62 GHz and the age-dependent mean gain is 82, while in the static-bias case the bandwidth is 8.5 GHz and the gain is 28. Meanwhile, if we look at larger age values (corresponding to photons arriving late in the pulse), we will see that the gain is generally small and so is the bandwidth. For example, at approximately s = 5 the age-dependent bandwidth is 23 GHz while the age-dependent mean gain is 3. This is because carriers have a reduced probability of impact ionizing due to the low field in the second half of the pulse.

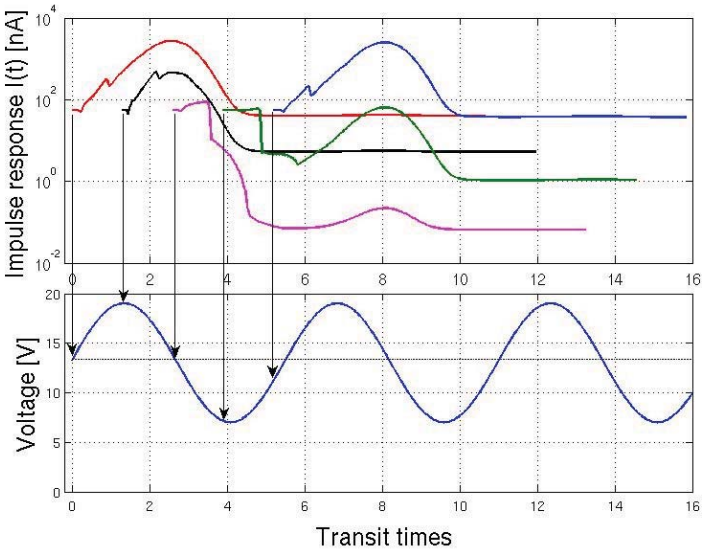

Fig. 4. Calculated age-dependent impulse response function under a sinusoidal dynamic bias. Different curves correspond to different ages (in transit times) of the initiating hole: red: age s = 0; black: age s = 1.4; magenta: age s = 2.6; and green: age s = 4. The dynamic-biasing parameters used are: B = 13 V, C = 6 V and ϕ = 0.

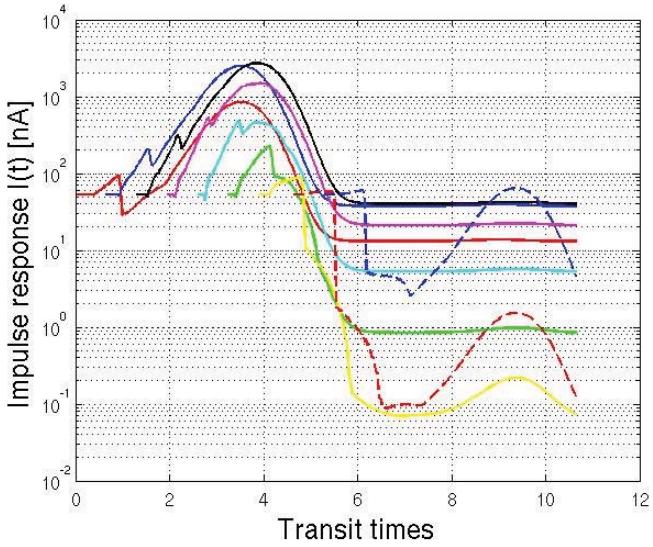

Fig. 5. Calculated age-dependent impulse response function under dynamic biasing using the following values for the dynamic-bias parameters: B = 13 V, C = 6 V and $\phi = \pi/3$.

#160303 - $15.00 USD Received 20 Dec 2011; revised 2 Mar 2012; accepted 9 Mar 2012; published 22 Mar 2012

5.2. Gain-bandwidth-product improvement

To see the net effect of the dynamic biasing scheme used in Fig. 5 on the pulse-integrated gain bandwidth product, we calculated the mean pulse-response function, $i_p(t)$, defined in (11). Figure 6 shows $i_p(t)$ once with the sinusoidal dynamic-field profile used in Fig. 5, and once with the static reverse bias. (For simplicity, in this example $p_{ph}(s)$ in (11) is assumed to be constant at $1/T$, namely, we assume a uniformly distributed random stream of photons). Note that this is simply the average of the age-dependent impulse responses. We observe that in comparison to $i_p(t)$ for the static bias, the tail of $i_p(t)$ in the sinusoidal-bias case is much reduced. The pulse-integrated bandwidth in the sinusoidal-bias case is 43.3 GHz, compared to 8.5 GHz in the static-bias case. The pulse-integrated gain in all three cases is about 28. Hence, the calculations predict an enhancement in the pulse-integrated gain-bandwidth product from 238 GHz in the static-bias case to a pulse gain-bandwidth product of 1169 GHz in the dynamic-bias case. This shows that a sinusoidally biased APD with the bias parameters described earlier can increase the pulse gain-bandwidth product of an APD by a factor of 5 compared to the same APD operated under the conventional static biasing scheme. We expect the results do extend to thicker multiplication regions provided that the ratio of the bit duration to the transit time is kept constant.

While the excess noise factor is not calculated in this paper, Herbert and Chidley [13] have shown via Monte-Carlo simulations that the excess noise factor can be reduced as long as the period of the sinusoidal bias is in excess of 4 times the transit time of each carrier, which would allow the avalanche process to quench. (The work of Herbert and Chidley followed their earlier papers on IMPATTs, which too exhibit a dynamic field in their multiplication region.) In fact, they showed that this reduction is insensitive to the ratio β/α, which is totally different from the behavior in the constant-field scenario. In the future we will further investigate the behavior of the overall excess noise factor analytically by using the recursive method for calculating the age-dependent excess noise factor as described in the Appendix.

#160303 - $15.00 USD Received 20 Dec 2011; revised 2 Mar 2012; accepted 9 Mar 2012; published 22 Mar 2012

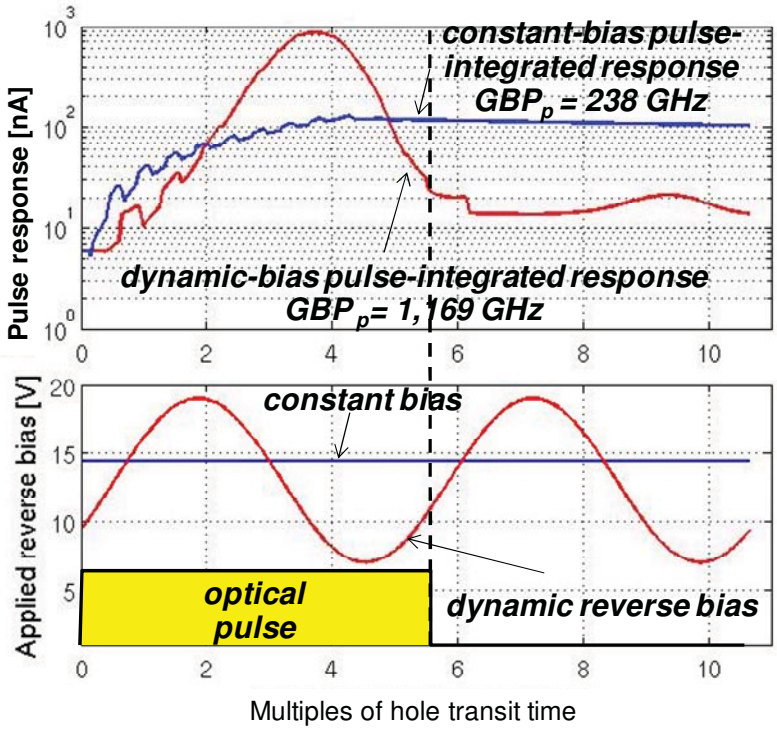

Fig. 6. Calculated pulse response due to a 8.3-ps rectangular optical pulse with the shown sinusoidal-dynamic bias function. The pulse response corresponding to a conventional static bias is also shown for comparison. A five-fold enhancement in the pulse-integrated gain-bandwidth-product is predicted.

6. Conclusions

Although static biasing has been the norm in operating linear-mode APDs in optical receivers for decades, there is much to be gained from modulating the impact ionization process, by means of dynamic biasing, so as to minimize the avalanche duration while maintaining a high muiltiplication factor. In fact, the calculations carried out in this paper have shown that static biasing may have been an overlooked factor in preventing APDs from being used in 40 Gbps lightwave systems. Here, we propose a bit-synchronous sinusoidal biasing scheme that offers two main advantages over static biasing in on-off-keying optical receivers. First, photons arriving early in an optical pulse generate avalanches with very high gains and avalanche durations that almost vanish by the end of the bit interval. Second, photons that arrive in the later part of an optical pulse generate low gains and very short avalanche durations. Consequently, high overall gain is achieved without inducing long avalanche durations and intersymbol interference. To rigorously justify the benefits of dynamic biasing, we have developed the first theory for the avalanche multiplication process under dynamic electric fields. The model allows us to analytically calculate the statistics of the avalanche multiplication factor, the impulse response function and the breakdown probability under arbitrary, dynamic electric fields. Our calculations predict that by using a bit-synchronous sinusoidal biasing scheme the pulse-integrated gain-bandwidth product of an InGaAs-InP SAM APD can be improved by a factor of 5 compared to the same APD operated under the conventional static biasing.

The proposed approach introduces a new paradigm for the operation and design of linear-mode APDs for high-speed lightwave systems and adds a new dimension to the traditional material- and structure-based approaches. Our approach is essentially APD-agnostic and can

#160303 - $15.00 USD Received 20 Dec 2011; revised 2 Mar 2012; accepted 9 Mar 2012; published 22 Mar 2012

be used to improve the gain-bandwidth product of *any* APD that has a poor avalanche-duration performance beyond traditional limits inherited from the conventional static biasing. Another projected benefit of dynamic biasing is that it may allow the relaxation of the often stringent requirements on the minimum width of the multiplication region, as normally done to enhance the APD speed. This, in turn, would lead to a reduction of the electric field in the multiplication region, which reduces tunneling current. Future efforts will focus on implementation of the proposed dynamic scheme as well as the rigorous analysis of the receiver sensitivity under bit-synchronous dynamic biasing.

Collaborative work is currently underway to demonstrate the benefits of dynamic biasing in a high-speed optical receiver. To this end, the dynamic-bias signal is generated from the clock pulse that is obtained by the clock/data recovery circuit on the receiver. Using an adjustable delay line, the generated dynamic bias must be properly aligned with the center of the data eye pattern. Care must be taken in designing the transimpedance amplifier, since the dynamic bias signal will be injected into the amplification path. Elimination of the dynamic bias at the output of the APD can be achieved by a number of methods that are being investigated. Commercially available high voltage high electron mobility transistors (e.g., GaN HEMTs) can be used to produce the necessary voltage swing for the APD bias driver.

Appendix

For completeness, we present for the first time the theory for calculating the excess noise factor, the probability mass function of the gain, and the breakdown probability all under arbitrary, dynamic biasing.

A.1. Age-dependent excess noise factor

Following the notation used in Section 4, we begin by deriving integral equations for the second moments $z_2(x,s) = \mathrm{E}[Z(x,s)^2]$ and $y_2(x,s) = \mathrm{E}[Y(x,s)^2]$. Note that the second moment of the multiplication factor, $m_2(x,s) = \mathrm{E}[M(x,s)^2]$, can be related to $z_2(x,s)$ and $y_2(x,s)$ using the identity $m_2(x,s) = 0.25[z_2(x,s) + y_2(x,s) + 2z(x,s)\ y(x,s)]$, and the age-dependent excess noise factor, $F_a(s) \equiv m_2(w,s) / g_a(s)^2$, is given by

$$F_a(s) = \frac{y_2(w,s) + 2y(w,s) + 1}{\left[y(w,s) + 1\right]^2}. \tag{A1}$$

Next, we proceed to derive integral equations characterizing $z_2(x,s)$ and $y_2(x,s)$. Since we can always express the second moment of $Z(x,s)$ as the expectation of the conditional second moment of $Z(x,s)$ given that the first ionization of the parent electron (triggering $Z(x,s)$) occuring at location X_e, we can write

$$\begin{aligned}
\mathrm{E}\left[Z(x,s)^2\right] &= \mathrm{E}[\mathrm{E}\left[Z(x,s)^2 \mid X_e\right]] \\
&= \mathrm{E}\left[\left\{Z_1\left(X_e, s+(X_e-x)/v_e\right) + Z_2\left(X_e, s+(X_e-x)/v_e\right) + Y\left(X_e, s+(X_e-x)/v_e\right)\right\}^2\right] \\
&= \mathrm{E}\left[Z_1\left(X_e, s+(X_e-x)/v_e\right)^2 + Z_2\left(X_e, s+(X_e-x)/v_e\right)^2 + Y\left(X_e, s+(X_e-x)/v_e\right)^2\right] \\
&\quad +2\mathrm{E}\left[Z_1\left(X_e, s+(X_e-x)/v_e\right) Z_2\left(X_e, s+(X_e-x)/v_e\right) + 2Z_1\left(X_e, s+(X_e-x)/v_e\right) Y\left(X_e, s+(X_e x)/v_e\right)\right] \\
&\quad +2\mathrm{E}\left[Z_2\left(X_e, s+(X_e-x)/v_e\right) Y\left(X_e, s+(X_e-x)/v_e\right)\right],
\end{aligned} \tag{A2}$$

which simplifies to

$$z_2(x,s) = \mathrm{E}\left[2z_2\left(X_e, s+(X_e-x)/v_e\right) + y_2\left(X_e, s+(X_e-x)/v_e\right) + 2z\left(X_e, s+(X_e-x)/v_e\right)^2 + 4z\left(X_e, s+(X_e-x)/v_e\right) y\left(X_e, s+(X_e-x)/v_e\right)\right]. \tag{A3}$$

Upon writing down the averaging over X_e explicitly in terms of the pdf of X_e, we obtain

$$\begin{aligned}
z_2(x,s) &= \int_w^\infty h_e(\xi; x, s)\, d\xi + \int_x^w [2z_2(\xi, s+(\xi-x)/v_e) + y_2(\xi, s+(\xi-x)/v_e) \\
&\quad + 4z(\xi, s+(\xi-x)/v_e)\, y(\xi, s+(\xi-x)/v_e) + 2z(\xi, s+(\xi-x)/v_e)^2]\, h_e(\xi; x, s)\, d\xi, \\
&0 \le x \le w, \quad s \ge 0.
\end{aligned} \tag{A4a}$$

Similarly, we can obtain a recursive equation for $y_2(x,s)$:

#160303 - $15.00 USD Received 20 Dec 2011; revised 2 Mar 2012; accepted 9 Mar 2012; published 22 Mar 2012

$$y_2(x,s) = \int_x^\infty h_h(\xi;x,s)d\xi + \int_0^x [2y_2(\xi, s+(x-\xi)/v_h) + z_2(\xi, s+(x-\xi)/v_h) + 4z(\xi, s+(x-\xi)/v_h)y(\xi, s+(x-\xi)/v_h) + 4y(\xi, s+(x-\xi)/v_h)^2]h_h(\xi;x,s)d\xi, \quad (A4b)$$
$$0 \le x \le w, \quad s \ge 0.$$

The pair of linear coupled integral equation in (A4) can be solved numerically once the quantities $z(x,s)$ and $y(x,s)$ have already been computed by first solving the pair of equations in (8).

A.2. Age-dependent breakdown probability

Here we develop recursive equations that characterize the probability that a parent carrier triggers breakdown, that is, the probability that infinitely many offspring carriers are generated. To this end, and following the general recursive approach [15], we notice that if an electron at position x in the multiplication region and of age s (relative to the launch instant of the dynamic electric field) impact ionizes for the first time at location ξ, then the probability that the parent electron generates a *finite* number of offspring carriers is precisely the product of the probabilities that each of the two offspring electrons and offspring hole created at ξ with age $s + (\xi\text{-}x)/v_e$ generates a *finite* number of offspring carriers. (Implicit in this statement is that each carrier acts independently of the other carries, which is a correct assumption since in this paper we do not include any feedback effect from the created charges on the electric field.) Note that in the special case when the parent electron exits the multiplication region without ionizing, the conditional probability that produces a finite number of carriers is trivially equal to one.

Mathematically, define $P_Z(x,s) = \mathrm{P}\{Z(x,s) < \infty\}$ and $P_Y(x,s) = \mathrm{P}\{Y(x,s) < \infty\}$, and note that $\mathrm{P}\{Z(x,s) < \infty \mid X_e = \xi\} = \mathrm{P}\{Z(\xi, s + (\xi\text{-}x)/v_e)) < \infty\}^2\, \mathrm{P}\{Y(\xi, s + (\xi\text{-}x)/v_e.)) < \infty\}$, while $\mathrm{P}\{Z(x,s) < \infty \mid X_e > w\} = 1$. By averaging the above equation over all possible X_e we obtain

$$P_Z(x,s) = \int_w^\infty h_e(\xi;x,s)d\xi + \int_x^w P_Z\left(\xi, s+(\xi-x)/v_e\right)^2 P_Y\left(\xi, s+(\xi-x)/v_e\right) h_e(\xi;x,s)d\xi \quad (A5a)$$
$$0 \le x \le w, \quad s \ge 0.$$

A similar argument can lead to the equation

$$P_Y(x,s) = \int_x^\infty h_h(\xi;x,s)d\xi + \int_0^x P_Y\left(\xi, s+(x-\xi)/v_h\right)^2 P_Z\left(\xi, s+(x-\xi)/v_h\right) h_h(\xi;x,s)d\xi \quad (A5b)$$
$$0 \le x \le w, \quad s \ge 0.$$

Once $P_Z(x,s)$ and $P_Z(x,s)$ are calculated by solving the nonlinear coupled integral equations in (A5) numerically (using iterations, for example), the breakdown probability is calculated. For example, for a hole-injection APD, the age-dependent breakdown probability for a photon absorbed at time s, is simply

$$P_B(s) = 1 - P_Z(w,s). \quad (A6)$$

A.3. Probability mass function of the age-dependent multiplication factor

Following the same conditioning approach as in A.1, one can derive integral equations for the probability mass function of the processes $Z(x,s)$ and $Y(x,s)$. Here we state the equations without proof. Let $f_Z(x,s,m) = \mathrm{P}\{Z(x,s) = m\}$ and $f_Y(x,s,m) = \mathrm{P}\{Y(x,s) = m\}$, $m = 1, 2, 3, \ldots$. Then, they must satisfy the following coupled integral equations

$$f_Z(x,s,m) = \delta_{m-1}\int_w^\infty h_e(\xi;x,s)d\xi + \int_x^w f_Z(\xi, s+(\xi-x)/v_e, m) * f_Z(\xi, s+(\xi-x)/v_e, m) * f_Y(\xi, s+(\xi-x)/v_e, m) \times h_e(\xi;x,s)d\xi, \quad 0 \le x \le w, \quad s \ge 0, \quad m = 1,2,3,\ldots \quad (A7a)$$

and

#160303 - $15.00 USD Received 20 Dec 2011; revised 2 Mar 2012; accepted 9 Mar 2012; published 22 Mar 2012

$$f_Y(x,s,m) = \delta_{m-1}\int_x^\infty h_h(\xi;x,s)d\xi + \int_0^x f_Y(\xi,s+(x-\xi)/v_h,m) * f_Y(\xi,s+(x-\xi)/v_h,m) * f_Z(\xi,s+(x-\xi)/v_h,m) \times h_h(\xi;x,s)d\xi, \quad 0 \le x \le w, \quad s \ge 0, \quad m = 1,2,3,\ldots \tag{A7b}$$

where δ_k = if k=0 and it is zero otherwise, and "*" denotes discrete convolution with respect to the variable m. An alternative form of (A7) can be obtained by taking the discrete (time) Fourier transform with respect to the discrete variable m, which yields the characteristic function. More precisely, define

$$F_Z(x,s,\omega) = \sum_{m=0}^{\infty} f_Z(x,s,m)e^{j\omega m}, F_Y(x,s,\omega) = \sum_{m=0}^{\infty} f_Y(x,s,m)e^{j\omega m}, \quad -\pi < \omega \le \pi, \text{and obtain}$$

$$F_Z(x,s,\omega) = e^{j\omega}\int_w^\infty h_e(\xi;x,s)d\xi + \int_x^w F_Z\left(\xi,s+(\xi-x)/v_e,\omega\right)^2 F_Y\left(\xi,s+(\xi-x)/v_e,\omega\right) \times h_e(\xi;x,s)d\xi, \quad 0 \le x \le w, \quad s \ge 0, \quad -\pi < \omega \le \pi \tag{A8a}$$

and

$$F_Y(x,s,\omega) = e^{j\omega}\int_x^\infty h_h(\xi;x,s)d\xi + \int_o^x F_Z\left(\xi,s+(x-\xi)/v_h,m\right)^2 F_Z\left(\xi,s+(x-\xi)/v_h,m\right) \times h_h(\xi;x,s)d\xi, \quad 0 \le x \le w, \quad s \ge 0, \quad -\pi < \omega \le \pi \tag{A8b}$$

Once the quantities characteristic functions are solved for numerically, the probability mass functions can be obtain by simply perform a Fourier-transform inversion, which can be implemented approximately by using the inverse discrete Fourier transform.

Acknowledgments

The authors are thankful to Dr. Payman Zarkesh-Ha for many valuable discussions regarding the implementation of the dynamic biasing. This work was supportted in part by the National Science Foundation under the Smart Lighting ERC.

#160303 - $15.00 USD Received 20 Dec 2011; revised 2 Mar 2012; accepted 9 Mar 2012; published 22 Mar 2012