ALLEGRO
microsystems

SENSITIVITY ANALYSIS OF APD PHOTORECEIVERS

By Andrew Huntington, PhD
Allegro MicroSystems

TABLE OF CONTENTS

INTRODUCTION

This paper presents a methodology to determine photoreceiver performance in a time-of-flight system. The Allegro test devices presented as examples in this paper do not reflect a current Allegro product. Rather, the example devices presented in this paper are intended to help system designers who seek to understand the myriad factors that interplay to affect photoreceiver performance.

BACKGROUND

APDs are photodetectors that can be regarded as the semiconductor analog of photomultiplier tubes (PMTs). One important difference is that the photocathode of an APD is not physically separated from the current gain medium, so an APD typically uses primary photocarriers more efficiently than a PMT. For the same reason, the quantum efficiency of an APD does not degrade over the lifetime of the detector. Another difference is that the multiplication process in an APD is normally bi-directional, so it has different characteristics than a PMT in which the gain process is uni-directional.

Linear-mode APDs are used in optical receivers for applications such as optical communications and time-of-flight measurements, which benefit from the internal photocurrent gain, fast response, compact size, durability, and low cost of the APD. The gain of a linear-mode APD improves the signal-to-noise ratio of a photoreceiver by boosting the signal photocurrent relative to circuit noise sources downstream in the signal chain.

InGaAs APD Structure

The structure of an APD and the manufacturing technique used to fabricate an APD depend on the semiconductor alloy employed in the APD. This technical note is primarily concerned with short-wavelength infrared (SWIR)-sensitive APDs with InGaAs absorbers; among other common types of APDs, silicon APDs sensitive to visible light and HgCdTe APDs sensitive in the mid- and long-wavelength infrared (MWIR / LWIR) are structurally dissimilar. Common InGaAs APD configurations include: a mesa-isolated APD with an InAlAs multiplier on the cathode side of the absorber (Figure 1); and a planar APD with an InP multiplier on the anode side of the absorber (Figure 2).
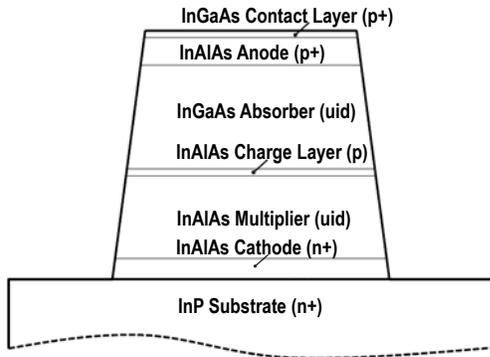
Figure 1: Typical InGaAs/InAlAs mesa APD.
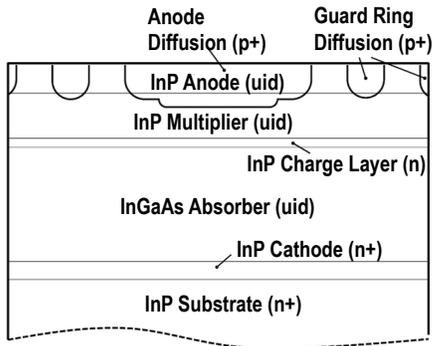
Figure 2: Typical InGaAs/InP planar APD.

Both styles of APD employ the separate absorption, charge, and multiplication (SACM) layer design that divides light absorption and charge carrier multiplication functions into distinct layers separated by a space charge layer that keeps the electric field strength in the absorber much lower than in the multiplier. The purpose of the SACM design is to minimize electric-field-driven tunnel leakage in the comparatively narrow-bandgap InGaAs absorber. Placement of the multiplication layer relative to the absorber is determined by the differing propensity of electrons and holes to impact-

ionize in any given alloy. Electrons drift toward the cathode and holes drift toward the anode, so the multiplier is placed on the side of the absorber toward which the carrier type with the higher ionization rate drifts. The junction of a mesa-isolated APD is formed epitaxially during wafer growth, whereas planar APDs are formed by diffusion of one dopant type into an epitaxially grown wafer containing the other dopant type. Whereas the lateral extent of the junction of a mesa APD is defined by physically etching away the epitaxial material outside its footprint, patterning of the diffusion that forms a planar APD defines its footprint. Planar APDs often use guard-ring diffusions outside the main anode diffusion to reduce the curvature of the depletion region under the perimeter of the device in order to reduce electric field strength there. Similarly, mesa APDs are formed with side-walls that slope gradually outward from the top of the mesa to its base because this geometry avoids localized concentration of the electric field lines at the mesa perimeter.

Avalanche Gain and Gain Distribution

The slope of an APD gain curve as a function of reverse bias limits the gain at which the APD can be used. The slope of the gain curve is a challenge because mean avalanche gain (M) increases asymptotically in the vicinity of the APD breakdown voltage (V_{br}) according to the empirical relation:

Equation 1:

$$M = \left[I - \frac{V}{V_{br}} \right]^{-n},$$

where the parameter n controls how quickly the avalanche gain rises as V approaches its vertical asymptote at V_{br}. Equation 1 demonstrates how an APD characterized by a large n value become impractical to operate at high gain because V/V_{br} cannot be controlled adequately. This relation holds for all APDs in which both carrier types (electrons and holes) can initiate impact ionization. [1]

Avalanche noise imposes a separate limit on the useable gain of an APD. In the limit of high avalanche gain, the sensitivity of a hypothetical photoreceiver that employs an ideal "noiseless" APD is limited by the shot noise on the optical signal itself. However, most APDs generate multiplication noise in excess of the shot noise already present on the optical signal; this excess multiplication noise intensifies with increase in avalanche gain such that, for any given level of downstream amplifier noise, there is a limit to how much avalanche gain is useful. Increasing the avalanche gain beyond the optimal value increases the shot noise faster than the amplified signal photocurrent, which degrades the

---

[1] Photon number histogram for ~100-µJ pulses from an Allegro 1535-nm Er:glass, passively Q-switched, diode-pumped solid state laser as measured with a thermoelectric sensor.

signal-to-noise ratio (SNR).

Excess multiplication noise results from the stochastic nature of the impact ionization process that amplifies the APD primary current. After avalanche multiplication, each primary carrier injected into an APD multiplier may yield a different number of secondary carriers. For most linear-mode APDs, the statistical distribution of n output carriers resulting from an input of a primary carriers is that derived by Robert J. McIntyre [2]:

Equation 2:

$$P_{McIntyre}(n) = \frac{a\,\Gamma\left[\dfrac{n}{1-k}+1\right]}{n\,(n-a)!\times\Gamma\left[\dfrac{n\,k}{1-k}+1+a\right]} \times \left[\frac{1+k(M-1)}{M}\right]^{a+\frac{n\,k}{1-k}} \times \left[\frac{(1-k)(M-1)}{M}\right]^{n-a},$$

where k is the ratio of hole-to-electron impact ionization rates, M is average gain, and Γ is the Euler gamma function.

McIntyre's distribution is far from Gaussian for small inputs (i.e., for a small number of primary photocarriers injected into the multiplier), with a pronounced positive skew (Figure 3).

Comparison of McIntyre and Gaussian Distributions Having the Same Mean and Standard Deviation

- McIntyre for 1 e⁻ input
- Gaussian for 1 e⁻ input
- McIntyre for 10 e⁻ input
- Gaussian for 10 e⁻ input
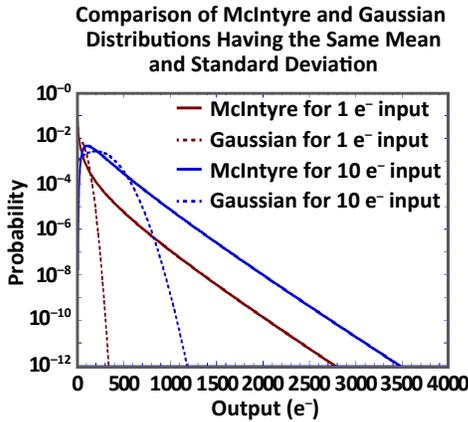
Figure 3: Comparison of McIntyre (solid) to Gaussian (dashed) output distributions from inputs of 1 (red) and 10 (blue) primary electrons, for a k = 0.2; M = 20 APD.

For larger inputs, the McIntyre distribution approximates a Gaussian shape near its mean due to the central limit theorem, and avalanche noise can be quantified for analysis with other common circuit noise sources by computing the variance of the gain.[3] The Burgess variance theorem [4],[5] gives the variance of the multiplied output n, for a primary carriers generated by a Poisson process and injected into a multiplier characterized by a mean gain M and random per-electron gain variable m [6]:

Equation 3:

$$\begin{aligned}\mathrm{var}(n) &= M^2\,\mathrm{var}(a) + \langle a\rangle\,\mathrm{var}(m) \\ &= M^2\,F\langle a\rangle \qquad [(\mathrm{e}^-)^2],\end{aligned}$$

where the excess noise factor F is defined as:

Equation 4:

$$F \equiv \frac{\langle m^2\rangle}{M^2}.$$

The noise factor is described as an "excess" because it is an elementary property of variances that, when a random variable is scaled by a constant factor, its variance is scaled by the square of the constant. Thus, if the gain were a constant $m = M$, rather than a random variable, then var($M \times a$) = M^2var(a) = $M^2\langle a\rangle$, which is smaller than Equation 3 by a factor of F. [7]

For most linear-mode APDs, the excess noise factor has the gain-dependence derived by McIntyre for thick, uniform junctions [8]:

Equation 5:

$$F = M\left[1 - (1-k)\left(\frac{M-1}{M}\right)^2\right].$$

[2] R. J. McIntyre, "The Distribution of Gains in Uniformly Multiplying Avalanche Photodiodes: Theory," IEEE Trans. Electron. Devices, vol. ED-19, no. 6, pp. 703-713, 1972.

[3] The Gaussian approximation does not hold very well far from the mean, so the full McIntyre distribution must be used to realistically model the performance characteristics that are sensitive to the tails of the output distribution, such as false alarm rate.

[4] R. E. Burgess, "Homophase and heterophase fluctuations in semiconductor crystals," Discussions of the Faraday Society, vol. 28, pp. 151-158, 1959.

[5] R. E. Burgess, "Some topics in the fluctuation of photo-processes in solids," J. Phys. Chem. Solids, vol. 22, pp. 371-377, 1961.

[6] M. C. Teich, K. Matsuo, and B. E. A. Saleh, "Excess Noise Factors for Conventional and Superlattice Avalanche Photodiodes and Photomultiplier Tubes," IEEE J. Quantum Electron., vol. QE–22, no. 8, pp. 1184-1193, 1986.

[7] It is important to note that, whereas $n = M \times a$ is the idealized case of constant gain, $n \neq m \times n$. The reason is that m is a per-electron random gain variable which takes on different values for each electron enumerated by a particular value of a. For more details, see the section "Burgess Variance Theorem for Multiplication and Attenuation"."

[8] R. J. McIntyre, "Multiplication noise in uniform avalanche photodiodes," IEEE Trans. Electron. Devices, vol. ED-13, pp. 164–168, 1966.

Equation 3 and Equation 5 were used to calculate the variances of the Gaussian distributions plotted in Figure 3. Note that although the McIntyre and Gaussian distributions have the same mean and standard deviation, they diverge significantly at output levels far from the mean.

In Equation 5, the parameter k is the same ratio of hole-to-electron impact ionization rates appearing in Equation 2. When $k > 0$, k is the slope of the excess noise curve as a function of gain, in the limit of high gain (Figure 4). For single-carrier multiplication, $k = 0$, and $F \rightarrow 2$ in the limit of high gain. Another feature of single-carrier $k = 0$ multiplication is that avalanche breakdown cannot occur. Without participation of one carrier type, all impact ionization chains must eventually self-terminate, because all carriers of the type capable of initiating impact ionization soon exit the multiplying junction. The gain curve of a $k = 0$ APD does not exhibit the vertical asymptote described by Equation 1, enabling stable operation at higher gain than a $k > 0$ APD.

Excess Noise Factor vs. Gain
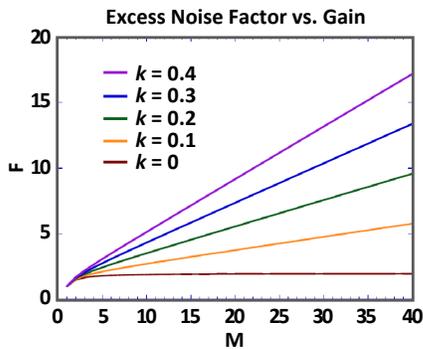
Figure 4: Plot of Equation 5 showing how the excess noise factor of most APDs increases linearly with avalanche gain in the limit of high gain, with a slope equal to k.

McIntyre distributions for APDs operating at the same average gain ($M = 20$) and illuminated by the same signal strength ($a = 10$ primary photoelectrons) but differing in k are plotted in Figure 5.

Comparison of McIntyre Distributions with Different Values of k Having the Same Mean and Primary Input
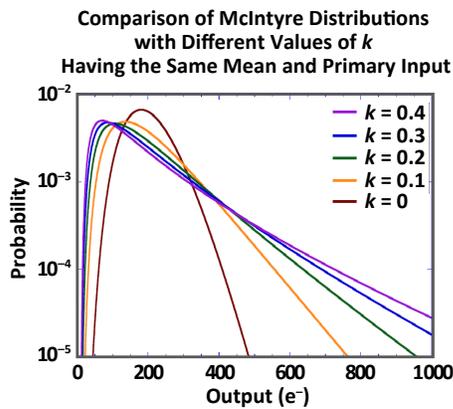
Figure 5: McIntyre distributions for input of 10 primary electrons, corresponding to the excess noise factors at $M = 20$ in Figure 4.

These distributions correspond directly to the excess noise factor values at the $M = 20$ vertical slice through the curves of Figure 4. The full McIntyre distributions illustrate the practical meaning of different values of k and F. For the same input signal strength and the same average gain, an APD with lower k (and F) will:

- Have a higher probability of detecting the signal; and
- Have a lower probability of generating a false alarm.

These statements assume that the APD is employed in a photoreceiver circuit equipped with a binary decision circuit that rejects signals below a certain detection threshold, and that the mean signal photocurrent is larger than the mean dark current. In this common scenario, a single detection threshold is simultaneously in the high-output tail of the dark current distribution and comfortably lower than the bulk of the probability density of the photocurrent distribution, such that the longer tail of the high k distribution increases the probability of false alarm but reduces signal detection probability by decreasing the median output value of the distribution. The detection threshold is employed to reject false alarms arising from circuit noise, and the APD dark current is one component. At the same time, the detection threshold must not be set so high that it also rejects outputs arising from valid photocurrent signals. An output distribution with a higher median for a given input is desirable because the high median will allow the detection threshold to be set higher without sacrifice to signal detection efficiency. On the other hand, a reduced likelihood of very-high-output events will help minimize the false alarm rate arising from "lucky" dark current electrons that happen to individually experience very high avalanche gain. Figure 5 illustrates how the median and high-output tail of the McIntyre distribution vary with k for an input of 10 primary photoelectrons and an average gain of $M = 20$. Figure 5 demonstrates the qualitative behavior of the McIntyre distribution that affects both signal detection and false alarm probability. As noted earlier, practical threshold detection scenarios require that the input level for a photocurrent signal distribution be larger than the input level for a dark current noise distribution, so that the same threshold level is simultaneously in the tail of the dark current distribution and comfortably below the median of the photocurrent distribution. Thus, when thinking about signal photocurrent detection, the medians of the distributions in Figure 5 are most important; and, when thinking about false alarms from dark current, the high-output tails of the distributions are what matter. The median output levels and probabilities of output exceeding a detection threshold of 1000 e⁻ which correspond to the distributions plotted in Figure 5 are tabulated in Table 1.

Table 1: Median Output and Chance of Exceeding 1000 e⁻ Corresponding to Distributions of Figure 5.

k	Median Output (See [1])	Std. Dev.	Chance of Output > 1000 e⁻ (See [2])
0	195 e⁻	88.3 e⁻	5.24×10^{-13}
0.1	179 e⁻	122.6 e⁻	6.61×10^{-5}
0.2	165 e⁻	149.1 e⁻	1.08×10^{-3}
0.3	154 e⁻	171.6 e⁻	3.45×10^{-3}
0.4	145 e⁻	191.5 e⁻	6.55×10^{-3}

[1] Higher is better for signal detection.

[2] Lower is better for avoiding false alarms.

When reviewing Table 1, it is worth noting that the mean output in all cases is 200 e⁻, and a detection threshold of 1000 e⁻ is, in all cases, greater than four standard deviations above the mean. If the output distributions were Gaussian with the same mean and variances as the actual McIntyre distributions, the chance of an output event exceeding 1000 e⁻ would be orders of magnitude lower. This is why the Gaussian approximation is not very good for calculating quantities that are sensitive to the tail of the output distribution, such as false-alarm rate. [9]

Exceptions to Standard APD Noise Theory

Equation 2 and Equation 5 were derived under the assumption that carriers are always "active"—meaning that carriers are always and everywhere capable of impact ionization. In reality, conservation of energy requires that carriers accumulate kinetic energy in excess of a threshold before they become active: the minimum displacement of a carrier within an applied electric field required to accumulate the impact ionization threshold energy is called its **dead space**. In thick, uniform APD junctions, the carrier dead space is negligible relative to the path length of the carrier through the gain medium, so Equation 5 holds very well. However, important exceptions to the excess noise factor formula of Equation 5 include: APDs in which the carrier dead space is a significant portion of the width of the multiplying junction [10],[11],[12],[13]; those in which a change in alloy composition modulates the impact ionization threshold energy and rate across the multiplying junction [14],[15],[16],[17],[18],[19]; and those made from semiconductor alloys with band structures that combine the traits of single-carrier-dominated multiplication (k ~ 0) with an abrupt carrier dead space (i.e., one in which the probability of impact ionization becomes very high immediately after traversing the dead space), resulting in correlation between successive impact

[9] This topic is discussed further in later sections. Technically, photoreceiver performance depends on the convolution of the APD output distribution with a Gaussian distribution representing amplifier noise. However, the general conclusions about how k and F relate to signal detection and false-alarm performance still hold in a more rigorous analysis.

[10] K. A. Anselm, P. Yuan, C. Hu, C. Lenox, H. Nie, G. Kinsey, J. C. Campbell, and B. G. Streetman, "Characteristics of GaAs and AlGaAs homojunction avalanche photodiodes with thin multiplication regions," Appl. Phys. Lett., vol. 71, no. 26, pp. 3883-3885, 1997.

[11] C. Lenox, P. Yuan, H. Nie, O. Bakelov, C. Hansing, J. C. Campbell, A. L. Holmes, Jr., and B. G. Streetman, "Thin multiplication region InAlAs homojunction avalanche photodiodes," Appl. Phys. Lett., vol. 73, no. 6, pp. 783-784, 1998.

[12] 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. Devices, vol. 47, no. 3, pp. 625-633, 2000.

[13] C. H. Tan, J. P. R. David, G. J. Rees, R. C. Tozer, and D. C. Herbert, "Treatment of soft threshold in impact-ionization," J. Appl. Phys., vol. 90, no. 5, pp. 2538-2543 (2001).

[14] S. Wang, R. Sidhu, X. G. Zheng, X. Li, X. Sun, A. L. Holmes, Jr., and J. C. Campbell, "Low-Noise Avalanche Photodiodes With Graded Impact-Ionization-Engineered Multiplication Region," IEEE Photon. Technol. Lett., vol. 13, no. 12, pp. 1346-1348, 2001.

[15] 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: Theory and Experiment," IEEE Trans. Electron. Devices, vol. 49, no. 12, pp. 2114-2123, 2002.

[16] S. Wang, J. B. Hurst, F. Ma, R. Sidhu, X. Sun, X. G. Zheng, A. L. Holmes, Jr., A. Huntington, L. A. Coldren, and J. C. Campbell, "Low-Noise Impact-Ionization-Engineered Avalanche Photodiodes Grown on InP Substrates," IEEE Photon. Technol. Lett., vol. 14, no. 12, pp. 1722-1724, 2002.

[17] S. Wang, F. Ma, X. Li, R. Sidhu, X. Zheng, X. Sun, A. L. Holmes, Jr., and J. C. Campbell, "Ultra-Low Noise Avalanche Photodiodes With a 'Centered-Well' Multiplication Region," IEEE J. Quantum Electron., vol. 39, no. 2, pp. 375-378, 2003.

[18] O.-H. Kwon, M. M. Hayat, S. Wang, J. C. Campbell, A. Holmes, Jr., Y. Pan, B. E. A. Saleh, and M. C. Teich, "Optimal Excess Noise Reduction in Thin Heterojunction Al0.6Ga0.4As-GaAs Avalanche Photodiodes," IEEE J. Quantum Electron., vol. 39, no. 10, pp. 1287-1296, 2003.

[19] C. Groves, J. P. R. David, G. J. Rees, and D. S. Ong, "Modeling of avalanche multiplication and noise in heterojunction avalanche photodiodes," J. Appl. Phys., vol. 95, no. 11, pp. 6245-6251, 2004.

ionization events. [20],[21],[22],[23],[24],[25] In general, the avalanche statistics of these types of APDs must be computed numerically, either through Monte Carlo modeling or application of recursive methods such as the dead space multiplication theory (DSMT). [26] Some APDs, such as those fabricated from HgCdTe alloys with cutoff wavelength in the mid- or long-wavelength infrared (MWIR/LWIR), do not obey McIntyre-like multiplication statistics at all; others, like InGaAs APDs with thin multipliers, generally follow McIntyre statistics but with a value of k that is smaller than the physical ratio of hole-to-electron impact ionization rate coefficients. Notably, Van Vliet derived a generalized analytic expression for F in which the number of possible impact ionizations per transit of the junction is a free parameter; Van Vliet's expression for F reproduces Equation 5 in the limit of an infinite number of possible ionizations per transit and converges to Lukaszek's [27] expression for F when a single ionization per transit is possible. [28]

ANALOG APD PHOTORECEIVERS

A block diagram of a typical analog APD photoreceiver is shown in Figure 6. Depending on the application, analog output from the photoreceiver might be sampled by a fast analog-to-digital converter (ADC) or run into a binary decision circuit such as a threshold comparator. The photoreceiver circuit in the diagram includes features like overload protection on the transimpedance amplifier (TIA) and a DC cancellation circuit to subtract the APD dark current; however, in essence, the photoreceiver is simply an APD and an amplifier.
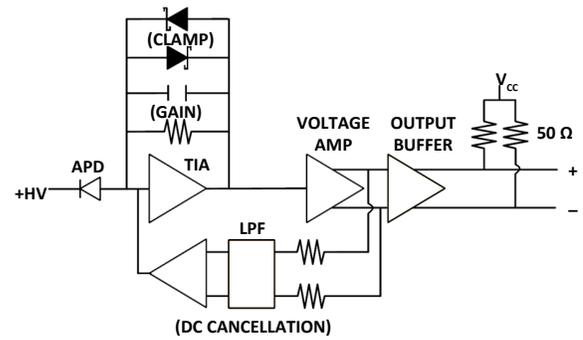
Figure 6: Typical analog APD photoreceiver.

Both capacitive and resistive elements are drawn in the TIA feedback path; however, in practice, TIAs are designed so that one or the other will dominate the amplifier gain. If the feedback is large on bothon the capacitive and resistive elements, the resistive element dominates and the TIA output voltage will be proportional to the instantaneous input current; its gain will be characterized by a transimpedance measured in ohms. The majority of TIAs sold for use with APDs are designed in this way. However, there is also a class of charge amplifier in which both the capacitive and resistive components are small, such that the capacitive element dominates and the TIA output voltage is proportional to the total charge delivered within rolling integration period τ. The capacitive TIA (CTIA) conversion gain is measured in units of reciprocal capacitance, such as V/e⁻. CTIAs for use in pulse-detection systems can be designed to continuously reset themselves by bleeding off the integrated signal charge through a low-pass filter, rendering them sensitive to transients while avoiding the necessity of a hard reset between reception of signal pulses. The time constant of the

[20] C. Vèrié, F. Raymond, J. Besson, and T. Nguyen Duy, "Bandgap spin-orbit splitting resonance effects in Hg1-xCdxTe alloys," J. Cryst. Growth, vol. 59, pp. 342-346, 1982.

[21] B. Orsal, R. Alabedra, M. Valenza, G. Lecoy, J. Meslage, and C. Y. Boisrobert, "Hg0.4Cd0.6Te 1.55-μm Avalanche Photodiode Noise Analysis in the Vicinity of Resonant Impact Ionization Connected with the Spin-Orbit Split-Off Band," IEEE Trans. Electron. Devices, vol. ED-35, pp. 101-107, 1988.

[22] K. A. El-Rub, C. H. Grein, M. E. Flatte, and H. Ehrenreich, "Band structure engineering of superlattice-based short-, mid-, and long-wavelength infrared avalanche photodiodes for improved impact ionization rates," J. Appl. Phys., vol. 92, no. 7, pp. 3771-3777, 2002.

[23] F. Ma, X. Li, J. C. Campbell, J. D. Beck, C.-F. Wan, and M. A. Kinch, "Monte Carlo simulations of Hg0.7Cd0.3Te avalanche photodiodes and resonance phenomenon in the multiplication noise," Appl. Phys. Lett., vol. 83, no. 4, pp. 785-787, 2003.

[24] M. A. Kinch, J. D. Beck, C.-F. Wan, F. Ma, and J. Campbell, "HgCdTe electron avalanche photodiodes," J. Electron. Mater., vol. 33, no. 6, pp. 630-639, 2004.

[25] A. R. J. Marshall, C. H. Tan, M. J. Steer, and J. P. R. David, "Extremely Low Excess Noise in InAs Electron Avalanche Photodiodes," IEEE Photon. Technol. Lett., vol. 21, no. 13, pp. 866-868, 2009.

[26] 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 Devices, vol. 39. pp. 546-552, 1992.

[27] W. Lukaszek, A. van der Ziel, and E. R. Chenette, "Investigation of the transition from tunneling to impact ionization multiplication in silicon p-n junctions," Solid-State Electron., vol. 19, pp. 57-71, 1976.

[28] K. M. Van Vliet, A. Friedmann, and L. M. Rucker, "Theory of Carrier Multiplication and Noise in Avalanche Devices – Part II: Two-Carrier Processes," IEEE Trans. Electron. Devices, vol. ED-20, no. 5, pp. 752-764, 1979.

reset path determines the effective integration period of the CTIA. [29]

Back-of-the-eenvelope receiver sensitivity calculations treat the APD responsivity and the TIA gain as fixed values. However, both are functions of frequency and are subject to saturation. Datasheet values usually correspond to the low-frequency, small-signal limit; in some cases, they are specific to a particular signal pulse shape, taking into account the power spectrum of the pulse. Care is required when attempting to estimate photoreceiver sensitivity to signals with frequency components outside the bandwidth of either the APD or the TIA, or for pulse shapes other than those for which the conversion gain is specified.

Since ideal resistive-feedback TIAs (RTIAs) respond to instantaneous current, but ideal CTIAs respond to integrated charge, APD photoreceivers assembled from either merit separate discussion. In the following sections, both RTIA-centric and CTIA-centric figures of photoreceiver merit are discussed. However, it is important to bear in mind that real TIAs have some degree of mixed character.

MEAN (SIGNAL)

Although phase and frequency modulation can be employed to encode information in an optical signal, this technical note presumes the signal data resides in its intensity, as measured by either its mean optical power (RTIA case) or its mean pulse energy (CTIA case).

RTIA Case

The APD of an RTIA-based receiver converts incident optical power in watts to an output photocurrent in amperes, which the RTIA then converts to a potential in volts. The average power conversion factor of the APD is called its spectral responsivity, R:

Equation 6:

$$R = M \, QE \, \frac{\lambda \, [\mu m]}{1.23985} \quad \left[\frac{A}{W} \right],$$

where QE is the quantum efficiency of the APD at a given wavelength (λ). The RTIA transimpedance is usually quoted in ohms.

Signal can be analyzed at any node in the circuit: at the input of the APD in terms of optical power or laser pulse energy, at the output of the APD (input of the TIA) in terms of current or charge, or at the output of the TIA in terms of potential. However, it is most common to perform calculations at the node between APD and TIA and to transform quantities to the other nodes as needed by applying the appropriate responsivity or conversion-gain factors.

The mean DC photocurrent signal from a continuous-wave (CW) optical signal of average power P_{signal} in watts is:

Equation 7:

$$I_{signal} = P_{signal} \, R \quad [A].$$

Some textbooks use the root-mean-square (RMS) optical power of an intensity-modulated signal for P_{signal} when analyzing optical communications applications, in which case I_{signal} represents the RMS photocurrent. [30]

Gain-Bandwidth Effects Limiting Signal Response

Practically speaking, RTIA-based photoreceivers are seldom used to detect CW optical signals. They are more commonly employed to detect the transition of an intensity-modulated optical signal through a given detection threshold, as in a time-of-flight device that times the arrival of a reflected pulse or a telecommunications receiver that discriminates the binary ones and zeroes of an optical bit stream. The response time of an APD photoreceiver is limited by the individual bandpass characteristics of the APD and the TIA, as well as by collective low-pass filtering associated with the capacitance of the detector and the input capacitance and transimpedance of the TIA.

The fundamental frequency response of an APD depends on its junction transit time and the DC gain at which it operates. Current flows continuously at the terminals of the APD from the time a charge carrier is created in its junction until such time as the carrier is swept to either its anode (in the case of holes) or cathode (in the case of electrons). [31],[32] The photocurrent of the APD cannot keep up with optical signal modulations on time scales shorter than its junction transit time because the carrier population that was generated by

[29]In practice, the conversion gain spectrum of a continuously reset CTIA has a complicated shape, and components of different frequency are amplified by different amounts. Likewise, the reset path may have a complicated bandpass. Thus, the conversion gain and effective signal integration period depend on signal pulse shape and cannot be characterized by a single value. However, conceptually it is helpful to envision the continuously reset CTIA as integrating all input current within a rolling sample period.

[30]P. Bhattacharya, Semiconductor Optoelectronic Devices, Second Edition, (Prentice Hall, Upper Saddle River, NJ, 1997), p. 369.

[31] W. Shockley, "Currents to Conductors Induced by a Moving Point Charge," J. Appl. Phys., vol. 9, no. 10, pp. 635-636, 1938.

[32]S. Ramo, "Currents Induced by Electron Motion," Proc. IRE, vol. 27, no. 9, pp. 584-585, 1939.

one optical power level will still be conducting current when the optical signal has changed to a new level. [33]

The avalanche gain process of an APD extends its impulse response beyond its junction transit time by prolonging generation of new carriers. (Refer to the earlier sketch of a typical mesa-style APD in Figure 1.) Primary photocarriers are generated in the InGaAs absorption layer; the photo-holes drift toward the anode and soon leave the junction. However, the photoelectrons drift toward the cathode by way of the InAlAs multiplier. Impact ionization in the multiplier generates electron-hole pairs. The secondary electrons will be swept from the junction at about the same time as the primary electrons that generated them, because both are drifting out of the multiplier and into the cathode. However, the secondary holes must now transit the entire width of the absorber before they can leave the junction at the anode, which extends the hole drift path and, therefore, the junction transit time. Further, except in the case of $k = 0$ APDs, some of the secondary holes that drift toward the anode may impact-ionize before drifting clear of the multiplier, creating tertiary electrons that drift toward the cathode which may themselves impact-ionize before clearing the junction, etc. Because counterpropagating carriers of either type can generate electron-hole pairs in the multiplier, avalanche multiplication is characterized by chains of impact ionization events. Higher avalanche gain corresponds to impact-ionization chains with more links, which take longer to complete. A tradeoff between avalanche gain and speed results. Not only does APD response roll off at high frequency due to finite junction transit time, but the bandwidth of the APD is also lower for higher DC gain because the gain process itself lasts longer. This effect is illustrated in Figure 7, in which the tail of the APD impulse response grows relative to the peak at higher DC gain.
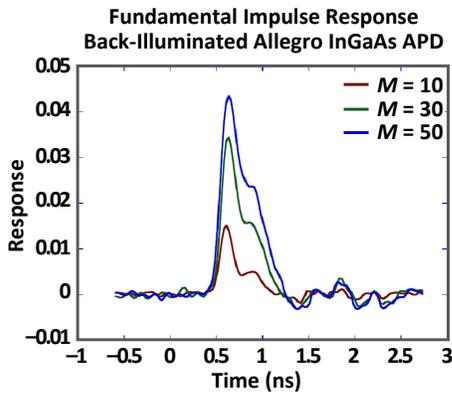
The practical upshot of the tradeoff between APD gain and bandwidth is that the responsivity acting on high-frequency signal components is smaller than the full DC responsivity of an APD calculated in Equation 6. However, fundamental APD response times are generally sub-nanosecond, as in Figure 7, so it is not typical for APD speed to be a factor in applications other than high-bit-rate telecommunications. It is common for the RTIA to limit photoreceiver speed in laser pulse-sensing applications like time-of-flight measurements.

Figure 8 illustrates schematically how laser pulse width would interact with the rise time of an idealized RTIA photoreceiver to reduce its response to fast optical pulses. Assuming laser pulses of equal energy but variable width, the response of the photoreceiver will be stronger to shorter optical pulses as long as the duration that the laser pulse width remains greater than the rise time of the photoreceiver. This is because idealized RTIA-based photoreceivers respond to instantaneous optical power rather than pulse energy, and shorter pulses delivering the same amount of energy have higher peak power. However, if the laser pulse is shorter than the photoreceiver rise time, the receiver response will not reflect the peak power of the optical signal because the driving force is withdrawn before the output has time to slew to a proportional level.
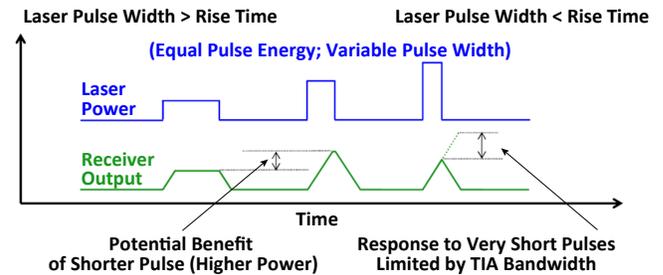
Figure 8: Illustration of how laser pulse width and RTIA photoreceiver rise time affect sensitivity.

Conceptually, a very rough estimate of the decreased response to fast pulses of an RTIA-based photoreceiver can be made by multiplying its responsivity by a correction factor based on the photoreceiver bandwidth (BW) and the pulse width (τ):

Equation 8:

Figure 7: Impulse response of a 75 µm $k = 0.02$ APD operated at three different DC gains.

$$R_{reduced} = R\left[1 - \exp\left(-2\pi\,BW\,\tau\right)\right] \qquad [\text{A/W}].$$

However, in low-duty-cycle applications, real-world RTIA-based photoreceivers often perform much better with short signal pulses than is implied by Equation 8 and its associated

[33]APD rise times are generally faster than fall times, because the APD can respond more-or-less instantaneously to an increase in optical power that adds photocarriers to its junction, but cannot respond to a decrease in optical power until carriers already present in the junction have cleared.

reasoning. In practice, when an RTIA cannot keep up with a fast input current pulse, the charge deposited on its input shifts the potential from virtual ground; current flows in the RTIA feedback resistor until the input potential has been restored to the normal operating point. Because current actually flows in the RTIA feedback resistor until its input has been restored to virtual ground, and not simply for the duration of the photocurrent pulse, the response of an RTIA photoreceiver to short, isolated laser pulses is often much better than implied by Equation 8. This consideration can favor use of lower-noise, low-bandwidth RTIAs in low-duty-cycle applications where absolute sensitivity is the main performance criterion. Low-bandwidth receivers cannot be used in high-duty-cycle applications like optical communications or multi-hit lidar because the slow rise and fall times merge consecutive symbols (pulses). However, in a comparatively low-duty-cycle application like time-of-flight ranging, higher RTIA bandwidth favors improved pulse-timing precision and resolution of pulse returns from objects that are closely spaced in range but is not essential for improving absolute sensitivity to short laser pulses.

A real-world example of a lower-bandwidth TIA delivering superior performance responding to short laser pulses is illustrated in Figure 9, which compares the sensitivity of a 22 MHz RTIA-based APD photoreceiver to a 37 MHz photoreceiver, as a function of laser pulse width. The RTIAs in question are variations of an Allegro time-of-flight application-specific integrated circuit (ASIC), and they differ in 3 dB bandwidth as a result of a difference in transimpedance: the only difference between the two ASICs is that the feedback resistance of the 22 MHz receiver is 1.0 MΩ as compared to 0.5 MΩ for the 37 MHz receiver. At 8 ns, the shortest pulse width tested is faster than the rise time of a 22 MHz amplifier, yet expressed in photons per pulse, [34] absolute receiver sensitivity is superior for shorter pulses versus longer pulses, and for the slower receiver configuration. The two reasons for this are that the amplifier output continues to rise even after the photocurrent pulse from the APD has ended, and because the slower amplifier configuration has both a narrower noise bandwidth and a lower noise spectral density, thanks to higher transimpedance gain.
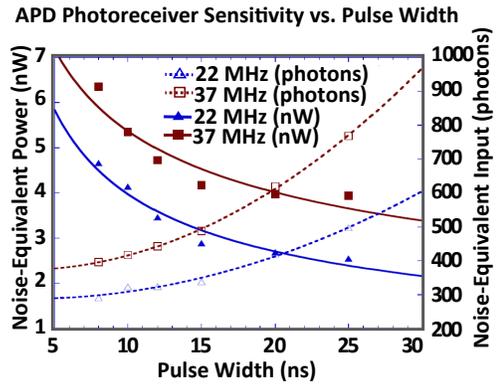
Figure 9: Measurement of how laser pulse width and RTIA photoreceiver bandwidth affect sensitivity (lower is better).

Because CTIA-based photoreceivers respond to integrated charge rather than to instantaneous current, their responsivity does not vary much with laser pulse width. It still takes time for the output of a CTIA photoreceiver to slew to a level that is proportional to the input pulse energy, but the CTIA output must outpace the time constant of the CTIA reset path rather than the laser pulse duration. On the other hand, APD and CTIA bandwidth both factor into the photoreceiver settling time. Although the responsivity of a CTIA-based photoreceiver is largely independent of optical pulse duration, the ability of a CTIA photoreceiver to resolve consecutive pulses that are closely spaced in time depends on high-bandwidth operation (as would an RTIA-based photoreceiver). Further, the finite time constant of the CTIA reset path makes CTIA-based photoreceivers unsuitable for some applications owing to the potential for saturation.

CTIA Case

In a CTIA-based receiver, the APD converts laser pulse energy in joules to an output charge in electrons, which the CTIA then converts to a potential in volts. The average energy-conversion factor of the APD is:

Equation 9:

$$R_{charge} = 5.03411 \times 10^{18} \, M \, QE \, \lambda \, [\mu m] \quad \left[\frac{e^-}{J} \right].$$

The CTIA conversion gain is a reciprocal capacitance usually expressed in V/e⁻.

The mean charge signal from an optical pulse of average energy E_{signal} is:

[34] For an explanation of the sensitivity measures used here, refer to the later sections "Noise-Equivalent Power (*NEP*)" and "Noise-Equivalent Input (NEI)". Expressed in terms of noise-equivalent power, longer signal pulses appear to give better sensitivity. This is not because the laser pulse width is better-matched to the receiver rise times, but is simply because increasing the duration of a laser pulse of a fixed average power increases its pulse energy. From the standpoint of using a fixed laser pulse energy, shorter pulses are superior (and *NEI* is a more relevant measure of sensitivity).

Equation 10:

$$Q_{signal} = E_{signal} \, R_{charge} \qquad [\mathrm{e}^-].$$

VARIANCE (NOISE)

By convention, photoreceiver noise is almost always analyzed at the node between the APD and the TIA. Because this node is at the input of the TIA, it is necessary to refer the TIA output-voltage noise to its input by application of its transimpedance (in the example of the RTIA) or conversion gain (in the example of the CTIA). In general, both the output voltage noise spectral intensity and the TIA gain are functions of frequency, so a rigorous analysis requires numerical methods. However, it is often sufficient for pen-and-paper estimates to approximate an RTIA input-referred noise spectrum as flat across its 3 dB bandwidth; an input noise spectral density in pA/√Hz or an RMS input noise in nA is often specified by the TIA manufacturer. Similarly, the input noise of a continuously reset CTIA is often characterized by a value in RMS electrons. However, scaling a certain number of RMS volts at the CTIA output to a certain number of RMS electrons at the CTIA input requires application of a conversion gain value that is specific to a particular signal pulse shape. The input-referred noise of the CTIA will vary with signal pulse shape even though its output voltage noise does not change.

The fluctuating level at the node between the APD and the TIA can be viewed as a random variable equal to the sum of random variables representing the APD output and the TIA input-referred noise. The APD output is, itself, the sum of random variables for the dark current and photocurrent; and, if there is background illumination, then the photocurrent is further subdivided into signal and background components. Fortunately, none of these random variables are correlated with each other, so the variance of the sum can be calculated as the sum of the variances.

RTIA Case for Conventional InGaAs APDs

In the case of an RTIA-based photoreceiver, the variance of the current at the node between the APD and the TIA is analyzed in terms of spectral intensities. If the input-referred noise of the RTIA is expressed as a spectral density in pA/√Hz, the corresponding spectral intensity (S_{I_TIA}) is the square of the spectral density. Alternatively, if only an RMS input noise current is specified, S_{I_TIA} is found by taking the square root of the ratio of the input noise current over the specified bandwidth. The spectral intensities of different APD noise components are calculated using an

extension of Milatz's theorem outlined by van der Ziel that allows Equation 3 to be recast as a noise spectral intensity theorem [35]:

Equation 11:

$$S_I = 2qM^2 F I_{primary} \qquad \left[\frac{\mathrm{A}^2}{\mathrm{Hz}} \right],$$

where $q = 1.602 \times 10^{-19}$ is the elementary charge in Coulombs and $I_{primary}$ is the primary (i.e., unmultiplied) current in amperes. Technically, Equation 11 only applies in the low-frequency limit, but the typical practice is to consider the APD multiplied shot noise spectrum to be approximately flat across its bandwidth.

The spectral intensity of the current noise at the node between the APD and the TIA is the sum of the individual spectral intensities of the RTIA input noise (S_{I_TIA}), the shot noise on the APD dark current (S_{I_signal}), and the shot noise on the APD photocurrent (S_{I_signal} and $S_{I\_background}$):

Equation 12:

$$S_{I\_total} = S_{I\_TIA} + S_{I\_dark} + S_{I\_background} + S_{I\_signal} \qquad [\mathrm{A}^2/\mathrm{Hz}].$$

For most InGaAs APDs, the majority of the primary dark current is generated in the narrow-bandgap InGaAs absorber, along with the background and signal photocurrent. When that is the case, Equation 11 can be used for the shot-noise spectral intensity, with $I_{primary}$ broken into different current components [36]:

Equation 13:

$$S_{I\_total} = S_{I\_TIA} + 2qMF\left(I_{dark} + I_{background} + I_{signal}\right) \qquad [\mathrm{A}^2/\mathrm{Hz}],$$

where I_{dark} is the dark current in amperes measured across the APD terminals, I_{signal} is the signal photocurrent given by Equation 7, and $I_{background}$ is the background photocurrent, given as:

Equation 14:

$$I_{background} = A \sum_n \left[\Delta\lambda_n \, I_B(\lambda_n) \, R(\lambda_n) \right] \qquad [\mathrm{A}],$$

where A is the area in m² of the receiver optical aperture,

---

[35] Van Der Ziel, A., Noise in Solid State Devices and Circuits (John Wiley & Sons, 1986), pp. 14-18.

[36] Note that Equation 11 is written in terms of an unmultiplied primary current and includes a factor of M^2, whereas multiplied terminal currents appear in Equation 13, so the order of M has been reduced by one.

$\Delta\lambda_n$ is the width in nm of wavelength bin n of a background spectral irradiance data set, $I_B(\lambda)$ is the background spectral irradiance in W m^{-2} nm^{-1} in bin n, and R(λ_n) is the APD spectral responsivity near the center wavelength of bin n given by Equation 6. Note that, in Equation 13, the three separate currents are physically indistinguishable. This means that there is little sensitivity to be gained by minimizing either dark current or background photocurrent once the signal photocurrent dominates both, and that background photocurrent can often be neglected if it is significantly weaker than the APD dark current.

The variance of the current at the node between the APD and the TIA within a bandwidth BW is:

Equation 15:

$$I^2_{noise} = BW\, S_{I\_total} \qquad [\mathrm{A}^2].$$

The standard deviation of the current, I_{noise}, is commonly referred to as the **noise current** of the photoreceiver.

RTIA Case for Multi-Stage Allegro APDs

Equation 13 assumes that all the primary current is generated outside the APD multiplication region, and that the primary photocurrent and primary dark current are subject to the same multiplication process. This is a good assumption for most InGaAs APDs because the InGaAs absorption layer is physically separate from the InP or InAlAs multiplication layer and because dark-current generation tends to be much faster in the narrow-bandgap InGaAs absorber than in the wide-bandgap alloys from which the rest of the APD is fashioned. However, in the special case of a proprietary multi-stage $k = 0.02$ APD by Allegro, noise on dark current must be treated separately from noise on photocurrent.

Internally, the Allegro $k = 0.02$ APD multiplier is divided into seven cascaded multiplying stages and, unlike most InGaAs APDs, the majority of the APD primary dark current is generated inside its multiplier rather than in its InGaAs absorber. This results in the dark current having different gain statistics than the photocurrent because the average avalanche gain experienced by a given current source inside the APD depends on how much of the APD multiplier it traverses. Dark current generated near the side of the multiplier adjacent to the absorber of the diode will experience substantially the same gain as the photocurrent, but dark current generated on the far side of the multiplier will experience very little gain. This can be approximated by summing dark-current noise contributions over the total multiplier.

The first step is to find the net avalanche gain experienced by dark current generated throughout the multiplier. Assuming that both dark current generation and avalanche gain are distributed uniformly across the multiplier, the average gain per stage is:

Equation 16:

$$M_S = \sqrt[stages]{M},$$

where stages is the number of multiplying stages (the example Allegro APD has seven stages), and M is the experimentally accessible avalanche gain measured for photocurrent.

Next, the net avalanche gain experienced by all dark current generated inside the APD multiplier is calculated, assuming that dark current generated in stage i is multiplied in every subsequent stage. The net gain is treated as a uniformly weighted average:

Equation 17:

$$M_{dark} = \frac{1}{stages} \sum_{i=1}^{stages} M_s^{i-1}.$$

Once the net gain experienced by the dark current is known, the primary dark current per stage can be calculated by dividing the multiplied dark current measured at the APD terminals by the net gain and the number of stages [37]:

Equation 18:

$$I_{dp} = \frac{I_{dark}}{stages\, M_{dark}} \qquad [\mathrm{A}],$$

where I_{dark} is the multiplied dark current measured at the APD terminals. Terminal dark-current parameterizations for the example Allegro APDs are given in a later section.

Finally, an expression similar to Equation 11 is summed over all the multiplying stages to find the noise spectral intensity of the APD dark current:

Equation 19:

$$S_{I\_dark} \approx 2q\, I_{dp} \sum_{i=1}^{stages} \left[(M_s^{i-1})^2\, F(M = M_s^{i-1}) \right] \qquad [\mathrm{A}^2/\mathrm{Hz}],$$

[37]Note that a factor of 1/**stages** in Equation 17 cancels with a factor of **stages** in the denominator of Equation 18 so that I$_{dp}$ is equal to the terminal dark current I_{dark} divided by the summation in Equation 17. To illustrate its conceptual origin, the expression is broken into two equations, Equation 17 and Equation 18, in this technical note.

where the notation $F(M = M_s^{i-1})$ means the excess noise factor of Equation 5 calculated with M_s^{i-1} substituted in place of the average avalanche gain measured for the photocurrent. When making calculations for a photoreceiver that uses the example k = 0.02 APD by Allegro, Equation 19 can be substituted into Equation 13 to obtain:

Equation 20:

$$S_{I\_total} \approx S_{I\_TIA} + 2\,q\,I_{dp} \sum_{i=1}^{stages}\left[(M_s^{i-1})^2\ F(M = M_s^{i-1})\right] + 2\,q\,M\,F\left(I_{background} + I_{signal}\right) \qquad [\mathrm{A}^2/\mathrm{Hz}].$$

The series of equations from Equation 16 through Equation 20 were derived to improve correspondence between theory and measurement for photoreceivers assembled from RTIAs and the example Allegro APDs. Table 2 compares noise-equivalent power (NEP) measurements to values calculated using either Equation 13 or Equation 20 for a 200 MHz photoreceiver built from a commercial-off-the-shelf (COTS) RTIA and an example Allegro k = 0.02 APD.

Table 2: NEP Measurements Compared to Models for 200 MHz RTIA, Allegro APD Photoreceiver, Characterized by k = 0.02.

Gain	Measured	Simple Model Equation 13	Distributed Model Equation 20
10	4.4 nW	8.1 nW	6.1 nW
20	3.7 nW	8.0 nW	5.1 nW
30	4.3 nW	8.4 nW	5.0 nW
39	4.5 nW	8.8 nW	5.1 nW
46.5	4.6 nW	9.2 nW	5.1 nW

As can be observed, the measured performance of the photoreceiver is better than predicted by either model, but the distributed dark-current model of Equation 20 is more accurate than the conventional model of Equation 13. Also, the accuracy of Equation 20 improves for M > 30, which is the typical operating point of Allegro APD receivers.

When using an equation within the series from Equation 16 through Equation 20 to model the multi-stage Allegro APD photoreceivers, it should be kept in mind that these equations are strictly valid only for the ideal case of k = 0 in which only electrons can trigger impact ionization. Impact ionization always generates equal numbers of secondary holes and electrons; however, for a k = 0 multiplier, only the secondary electrons cause additional impact ionizations. Because the actual Allegro APD is characterized by $k \approx 0.02$, the model does not treat all of its avalanche physics. The dominance of impact ionization by electrons is implicit in the model because Equation 16, for the gain per stage, treats avalanche as though 100% of the primary carriers multiplied in stage i_0 originate "upstream" ($i < i_0$), implying that they are all electrons. In reality, some secondary holes generated by impact ionization "downstream" in stages $i > i_0$

would also impact-ionize as they pass back through stage i_0. Further, the summation in Equation 19 treats the dark-current shot-noise spectral intensity of a single s-stage multiplier as the sum of the spectral intensities of s different multipliers with stages numbering between 0 and $(s - 1)$. [38] The idea is that primary dark current generated in stage i_0 can avalanche in all the downstream stages $i > i_0$, and i_0 is stepped through all the stages of the multiplier to account for primary dark current generated in each stage. For a given term of the summation, this approach properly treats noise associated with hole feedback involving any of its downstream multiplying stages. However, hole feedback into upstream multiplying stages is not modeled. The exact impact-ionization statistics of a multi-stage k > 0 multiplier have been successfully analyzed using numerical techniques, but the treatment in the equations from Equation 16 through Equation 20 represents a reasonably accurate closed-form approximation that is useful for low k multi-stage APDs.

CTIA Case for Conventional InGaAs APDs

In the case of a CTIA-based photoreceiver, the variance of the electron count at the node between the APD and the TIA is calculated by application of Equation 3 for the variance of the APD multiplied output:

Equation 21:

$$N_Q^2 = N_{CTIA}^2 + N_{dark}^2 + \left(\left\langle a_{signal}\right\rangle + \left\langle a_{background}\right\rangle\right)M^2\,F \qquad [(\mathrm{e}^-)^2],$$

where N_{CTIA} and N_{dark} are respectively the standard deviations of the CTIA input-referred noise and the number of dark current electrons output during the CTIA effective integration period τ; similarly A_{signal} and $A_{background}$ are respectively the number of primary photocurrent electrons generated by signal and background optical power received during τ. F is the excess noise factor calculated from Equation 5.

The input-referred noise of the CTIA, N_{CTIA}, is a characteristic of the CTIA and the laser pulse shape. If N_{CTIA} is not specified by a manufacturer, it can be calculated from a circuit simulation of the CTIA in which the APD capaci-

---

[38] The reason the number of stages ranges between 0 and $(s - 1)$, as opposed to between 1 and s, is that dark current carriers generated in the high-field region of a given multiplier stage do not have sufficient kinetic energy to impact-ionize in that stage; they only become active in the next stage.

tive load on the CTIA input and its mean dark current are modeled, but the shot noise on the APD current is omitted. Alternatively, CTIA conversion gain can be measured using a photoreceiver in which the detector noise contribution is negligible, such as a receiver assembled from a low-leakage p-i-n photodiode. In both cases, the input-referred noise of the CTIA is found by dividing the output voltage noise by the charge-to-voltage conversion gain of the CTIA.

The noise on the multiplied dark current, N_{dark}, depends on the structure of the APD. Most InGaAs APDs generate the majority of their primary dark current in their absorber, alongside the primary photocurrent generated by the optical signal and background. In that case, carriers from primary dark current can be grouped with the primary photo-carriers in Equation 21:

Equation 22:

$$N_Q^2 = N_{CTIA}^2 + \left(\langle a_{dark} \rangle + \langle a_{background} \rangle + \langle a_{signal} \rangle\right)M^2\,F$$
$$= N_{CTIA}^2 + \left[\frac{\tau}{q}\left(I_{dark} + I_{background}\right) + Q_{signal}\right]M\,F \qquad [(e^-)^2],$$

where Q_{signal} is given by Equation 10.

Equation 21 and Equation 22 approximate the shot noise on the signal term as though the signal charge originates from CW illumination rather than a transient laser pulse. The derivation of the excess noise factor from the Burgess variance theorem in Equation 3 assumes that the primary carrier

count results from a Poisson process. This is true of charge integrated over a set time period from steady dark current or from photocurrent from most types of steady background illumination, but laser pulse energy is often not Poisson-distributed from shot to shot. If greater accuracy is desired, the actual distribution of laser shot energy can be empirically measured and used with the full McIntyre distribution of Equation 2. On the other hand, if a noisy optical signal is attenuated by a large factor, a Poisson distribution is recovered. For more details, see "Burgess Variance Theorem for Multiplication and Attenuation".

CTIA Case for Multi-Stage Allegro APDs

The treatment of the dark current shot noise of the example multi-stage Allegro APD in a CTIA-based receiver is closely analogous to that described earlier for the RTIA case. In regard to the the gain and primary dark current per multiplying stage, Equation 16, Equation 17, and Equation 18 apply. An expression similar to Equation 3 is summed over all the multiplying stages to find the variance of the dark current of the example multi-stage Allegro APD:

Equation 23:

$$N_{dark}^2 \approx \frac{\tau}{q}I_{dp}\sum_{i=1}^{stages}\left[(M_s^{i-1})^2\,F(M = M_s^{i-1})\right] \qquad [(e^-)^2].$$

When making calculations for a photoreceiver that uses the $k = 0.02$ model Allegro APD, Equation 23 can be substituted into Equation 22 to obtain Equation 24:

Equation 24:

$$N_Q^2 \approx N_{CTIA}^2 + \frac{I_{dp}}{q}\tau\sum_{i=1}^{stages}\left[(M_s^{i-1})^2\,F(M = M_s^{i-1})\right] + \left[\frac{1}{q}I_{background}\,\tau + Q_{signal}\right]M\,F \quad [(e^-)^2].$$

SENSITIVITY METRICS DERIVED FROM MEAN AND VARIANCE

The sensitivity of an analog APD photoreceiver can be expressed in several forms. These include signal-to-noise ratio (SNR), noise-equivalent power (NEP), and noise-equivalent input (NEI). When the output of an analog APD photoreceiver is run into a decision circuit like a threshold comparator, additional metrics apply, such as optical sensitivity at a given false alarm rate (FAR) or bit error rate (BER). With a decision circuit, analysis can be conducted of the probabilities of true and false positives and negatives, to characterize the probabilities of signal detection (P_D) and

false alarm (P_{FA}), in the form of a parametric plot over detection threshold of P_D versus P_{FA}, called a receiver operating characteristic (ROC), can be prepared.

SNR, NEP, and NEI and are all ways of expressing the standard deviation of a photoreceiver output (the square root of the variance calculated in the preceding sections). SNR compares the mean to the standard deviation, whereas NEP and NEI refer the standard deviation to the APD input.

It is common to calculate FAR, BER, P_D, and P_{FA} based on the mean and standard deviation of the photoreceiver output by assuming it is Gaussian-distributed. However, as

was shown in the introduction (Figure 3), the high-output tail of an APD McIntyre distribution diverges substantially from the Gaussian approximation. When the McIntyre-distributed APD output is convolved with the Gaussian-distributed noise of the TIA, the convolution retains some of the positive skew of the McIntyre distribution. Consequently, when the Gaussian approximation is used, it underestimates FAR, BER, P_D and P_{FA}. For this reason, sensitivity metrics that depend on the tail of the photoreceiver output distribution are discussed in a separate section of this technical note.

Signal-to-Noise Ratio (SNR)

The form of the SNR depends on the node at which it is defined and the convention by which it is defined. Physicists tend to focus on the optical signal measured either in watts of power in the RTIA case or joules of energy (or, equivalently, photon number) in the CTIA case. Electrical engineers are used to dealing with potential signals in circuits measured in volts, such that the power dissipated in an impedance is proportional to the square of the voltage. This can cause confusion because the output voltage of a photoreceiver has a linear relationship to the input optical power or pulse energy, rather than a square relationship. To a physicist thinking about the optical signal, he SNR is the mean output signal voltage divided by its standard deviation; this is because these quantities have a linear relationship to the mean optical power level impinging on the receiver and the equivalent standard deviation found by referring the current and voltage noise sources from the APD and TIA to the receiver input. However, an SNR is sometimes encountered where the SNR is defined as the square of the mean output signal voltage divided by its variance; an SNR defined that way characterizes electrical power dissipation in a load on the photoreceiver output rather than the power of the optical signal itself. Similarly, confusion can arise when measuring power ratios of optical signals in decibels (dB) or optical power in decibels referred to one milliwatt (dBm). Because the power dissipated in an impedance goes as the square of the voltage, electrical engineers are used to applying the conversion [level] dB = 20 log(quantity); however, if the quantity in question is a power and not a voltage, the conversion is [level] dB = 10 log(quantity). In this technical note, the optical-signal-oriented convention is followed, and SNR is defined in terms of the mean and standard deviation rather than their respective squares. For convenience, the SNR expression is evaluated at the node between the APD and the TIA:

In the RTIA case:

Equation 25:

$$SNR = \frac{I_{signal}}{I_{noise}} = \frac{P_{signal}\,R}{\sqrt{BW\,S_{I\_total}}}.$$

For a photoreceiver based on an RTIA and a conventional InGaAs APD, the SNR is:

Equation 26:

$$SNR = \frac{P_{signal}\,R}{\sqrt{BW[S_{I\_TIA} + 2\,q\,M\,F(I_{dark} + I_{background} + P_{signal}\,R)]}}.$$

If a similar RTIA photoreceiver is assembled from the example k = 0.02 multi-stage Allegro APD, the SNR is [39]:

Equation 27:

$$SNR \approx$$

$$\frac{P_{signal}\,R}{\sqrt{BW\left[S_{I\_TIA} + 2\,q\,I_{dp}\sum_{i=1}^{stages}\left[(M_s^{i-1})^2\,F(M=M_s^{i-1})\right] + 2\,q\,M\,F(I_{background} + P_{signal}\,R)\right]}}.$$

The SNR of a CTIA photoreceiver is:

Equation 28:

$$SNR = \frac{Q_{signal}}{N_Q} = \frac{E_{signal}\,R_{charge}}{\sqrt{N_{CTIA}^2 + N_{dark}^2 + (\langle a_{signal}\rangle + \langle a_{background}\rangle)M^2\,F}}.$$

For a photoreceiver based on a CTIA and a conventional InGaAs APD, the SNR is:

Equation 29:

$$SNR = \frac{E_{signal}\,R_{charge}}{\sqrt{N_{CTIA}^2 + \left[\frac{\tau}{q}(I_{dark} + I_{background}) + E_{signal}\,R_{charge}\right]M\,F}}.$$

With the example k = 0.02 multi-stage Allegro APD, the SNR of a CTIA photoreceiver is [39]:

---

[39]For a discussion of the approximations inherent in the denominator of Equation 27 and Equation 30, refer to "RTIA Case for Multi-Stage Allegro APDs".

Equation 30:

$$SNR \approx \frac{E_{signal}\, R_{charge}}{\sqrt{\left[N_{CTIA}^2 + \frac{I_{dp}}{q}\tau \sum_{i=1}^{stages}\left[(M_s^{i-1})^2\, F(M=M_s^{i-1})\right]\right]+\left(\frac{\tau}{q}I_{background}+E_{signal}\,R_{charge}\right)M\,F}}\ .$$

Example *SNR*s calculated using Equation 26 are plotted versus avalanche gain in Figure 10. The hypothetical photo-receiver is assembled from a 200 μm-diameter Allegro APD characterized by $k = 0.2$, and uses either a COTS 622 Mbps TIA or a COTS 2.2125 Gbps TIA; the photoreceiver is band-limited to 200 MHz. Curves are plotted in Figure 10 comparing: *SNR* for optical signal power levels of 10 nW, 100 nW, and 1000 nW at 1550 nm (left); *SNR* for effective ionization-rate ratios of $k = 0$, 0.2 and 0.4 (center); and the *SNR* for receivers assembled from the 622 Mbps TIA or the 2.2125 Gbps TIA (right); the default conditions were $P_{signal} = 100$ nW, k = 0.2, and $S_{I\_TIA} = 4.4 \times 10^{-24}$ A^2/Hz

(the 622 Mbps TIA). Negligible background illumination was assumed. Notice that the optimal gain for maximization of the photoreceiver *SNR* varies for all these situations. Increasing either the optical signal power or the effective ionization rate ratio increases the APD noise contribution, shifting the optimal operating point to lower gain. Increasing the TIA noise contribution shifts the APD optimal operating point to higher gain. Although not shown, a strong background or higher dark current would shift the optimal operating point to lower gain, as would an increase in the APD noise contribution relative to fixed TIA noise.
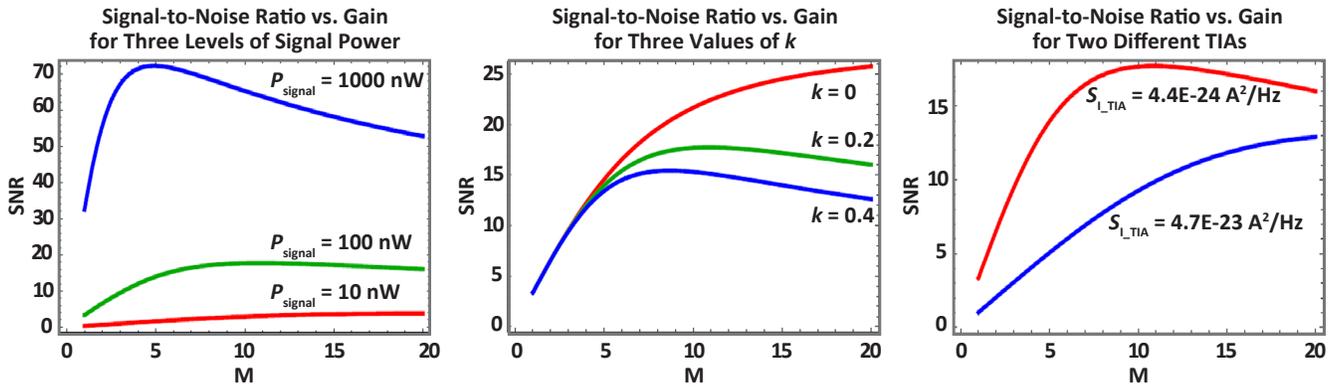
Figure 10: *SNR* vs. M curves calculated for a photoreceiver assembled from a 200 μm k = 0.2 Allegro APD and a COTS TIA, demonstrating how optimal gain depends on signal power (P_{signal}), the APD effective ionization rate ratio (k), and the TIA input-referred noise spectral intensity (S_{I_signal}).

Noise-Equivalent Power (NEP)

NEP can be defined and used with or without consideration of the shot noise on a hypothetical "noise-equivalent" signal. When *NEP* is defined to include the shot noise on a hypothetical noise-equivalent signal, it emphasizes the accuracy with which a photoreceiver can measure analog optical signal power, answering the question, "At what optical signal power will the signal-to-noise ratio of the receiver equal unity?". Because signal shot noise increases with signal strength, *NEP* cannot be used directly to calculate *SNR* at higher signal power. However, *NEP* is useful as a minimum sensitivity benchmark, and the symbol $NEP_{SNR=1}$ is used for this definition.

In contrast, when *NEP* is defined without signal shot noise, it emphasizes the photoreceiver propensity for false alarms

in the absence of a signal, answering the question, "What hypothetical optical signal power would result in an output level that is equal in magnitude to the RMS noise, absent any signal?". Used this way, *NEP* quantifies the photoreceiver noise floor in units that are convenient to compare to the optical signal level characteristic of a given application. For instance, a time-of-flight system can be designed in which a photoreceiver equipped with a threshold comparator times the arrival of laser pulses reflected from a target. The detection threshold must be set high enough that the probability of a false alarm in the absence of a reflected signal, P_{FA}, is negligible. At the same time, computing the pulse detection probability, P_D, requires knowledge of how the reflected signal strength compares to the detection threshold. *NEP* is often used in situations like this to quantify the photoreceiver noise in the absence of a signal because expressing all three quantities—RMS noise level, detection threshold, and mean

signal level—in units of optical power permits easy comparison. Moreover, because false alarms occur in the absence of a signal, it is valid to simply multiply *NEP* by an appropriate factor to set the detection threshold for a desired P_{FA}. [40]

To determine the optical signal power level at which the *SNR* equals unity, equate the numerator to the denominator in Equation 26 or Equation 27, substitute $NEP_{SNR=1}$ for P_{signal}, and solve for $NEP_{SNR=1}$ by using the quadratic formula. In the case of a receiver assembled from a conventional InGaAs APD, corresponding to Equation 26, the $NEP_{SNR=1}$ is:

Equation 31:

$$NEP_{SNR=1} = \frac{q\,M\,F\,BW + \sqrt{(q\,M\,F\,BW)^2 + BW[S_{I\_TIA} + 2\,q\,M\,F\,(I_{dark} + I_{background})]}}{R} \quad [\text{W}].$$

The example $k = 0.02$ multi-stage Allegro APD case corresponding to Equation 27 is [41]:

Equation 32:

$$NEP_{SNR=1} \approx \frac{q\,M\,F\,BW + \sqrt{(q\,M\,F\,BW)^2 + BW\left[S_{I\_TIA} + 2\,q\left(I_{dp}\sum_{i=1}^{stages}[(M_s^{i-1})^2\,F(M=M_s^{i-1})] + M\,F\,I_{background}\right)\right]}}{R} \quad [\text{W}].$$

The form of *NEP* that expresses the photoreceiver noise in the absence of any signal is algebraically simpler, being the standard deviation of the current at the node between the APD and the TIA, referred to the photoreceiver input by application of the APD responsivity:

Equation 33:

$$NEP = \frac{I_{noise}(P_{signal} = 0)}{R} \quad [\text{W}].$$

Referring to Equation 13 and Equation 15 for the variance of the current at the node between the APD and the TIA in a photoreceiver based on a conventional InGaAs APD, the *NEP* without shot noise on the hypothetical noise-equivalent signal is:

Equation 34:

$$NEP = \frac{\sqrt{BW[S_{I\_TIA} + 2\,q\,M\,F\,(I_{dark} + I_{background})]}}{R} \quad [\text{W}].$$

The current variance of a photoreceiver based on a multi-stage $k = 0.02$ Allegro APD is given by Equation 15 and Equation 20; its *NEP* without shot noise on a hypothetical noise-equivalent signal is [41]:

Equation 35:

$$NEP \approx \frac{\sqrt{BW\left[S_{I\_TIA} + 2\,q\,I_{dp}\sum_{i=1}^{stages}[(M_s^{i-1})^2\,F(M=M_s^{i-1})] + 2\,q\,M\,F\,I_{background}\right]}}{R} \quad [\text{W}].$$

---

[40] Challenge remains in the selection of the factor by which the detection threshold is set to exceed the *NEP* because of the divergence of the high-output tail of the McIntyre distribution from the tail of a Gaussian distribution having the same mean and variance. For further discussion, refer to "Avalanche Gain and Gain Distribution" and, particularly, Figure 3. A more accurate treatment of the *FAR* problem is discussed later.

[41] For a discussion of the approximations inherent in the treatment of the $k = 0.02$ APD dark current shot noise appearing in the numerator of Equation 32 and Equation 35), refer to "RTIA Case for Multi-Stage Allegro APDs".

The difference in definition between $NEP_{SNR=1}$ given by Equation 31 and Equation 32, and NEP given by Equation 34 and Equation 35 only becomes relevant when the dark current and TIA noise contribution are both exceptionally small. The two definitions of NEP are differentiated by the quantity $(q\,M\,F\,BW)$ appearing in two places in the numerator of Equation 31 and Equation 32. However, this factor is usually dominated by the terms representing the TIA noise contribution $(BW \times S_{I\_TIA})$, and/or the noise on the dark current, which is $(2\,q\,M\,\bar{F}\,I_{dark}\,BW)$ in Equation 31 or the equivalent summation over multiplier stages in Equation 32. This circumstance may arise in calculations for specialized photon-counting receivers, but that application is more commonly served by CTIA-based photoreceivers, for which an equivalent set of definitions apply to NEI. However, for illustrative purposes, calculations of NEP for a hypothetical photoreceiver in which a 200 µm Allegro APD characterized by k = 0.2 are shown in Figure 11, top, where the APD is operated at −30°C to minimize its dark current, and the noise spectral intensity of the TIA is four orders of magnitude lower than that of the 622 Mbps TIA. In Figure 11, the dashed NEP curves were calculated using Equation 31 for the case that includes shot noise on the noise-equivalent signal, and the solid curves were calculated using Equation 34, which omits signal shot noise. Room-temperature operation of the receiver and normal COTS TIAs are assumed in Figure 11, left and right. The "strong background" mentioned in Figure 11, left, is equivalent to 1 µA of primary photocurrent; a comparison of different APD ionization rate ratios is not shown for the case of zero background; however, in that case, the curves all overlay each other, following the red curve of Figure 11, right, because the TIA noise completely dominates.

Equation 31, Equation 32, Equation 34, and Equation 35 can be converted to spectral densities in W/√Hz by omitting the factor of BW inside the radical.

Cases in which the optimal gain operating point that minimizes NEP is less than the maximum gain are shown in Figure 11, center and left. Similar to the earlier discussion of gain optimization for maximum SNR, the optimal gain is determined by the relative dominance of APD versus TIA noise components. The excess noise factor, F, and the responsivity, R, are both order 1 in M, as per Equation 5 and Equation 6; the terminal dark current, I_{dark}, is at least order 1 (for details, refer to "Parameterization of Terminal Dark Current for Example Allegro APDs"). Consequently, once the APD noise term becomes larger than the TIA noise term, operation at higher avalanche gain will degrade sensitivity because the numerator of the NEP expression increases faster with gain than does the denominator. NEP can be minimized with respect to M to identify an optimal operating point, provided that the application does not depend on the high-output tail of the distribution. However, because the noise distribution of an APD photoreceiver is, compared to the Gaussian distribution with the same mean and variance, skewed with higher probability density at high output, minimum NEP can occur at a gain operating point for which the FAR or BER are not optimal. When an application is sensitive to low-probability false positives, it is best to supplement analysis of NEP with a more rigorous analysis of the actual noise distribution. This is topic is discussed in a later section.

Dashed curves: With noise-equivalent signal shot noise; Solid curves: Without signal shot noise
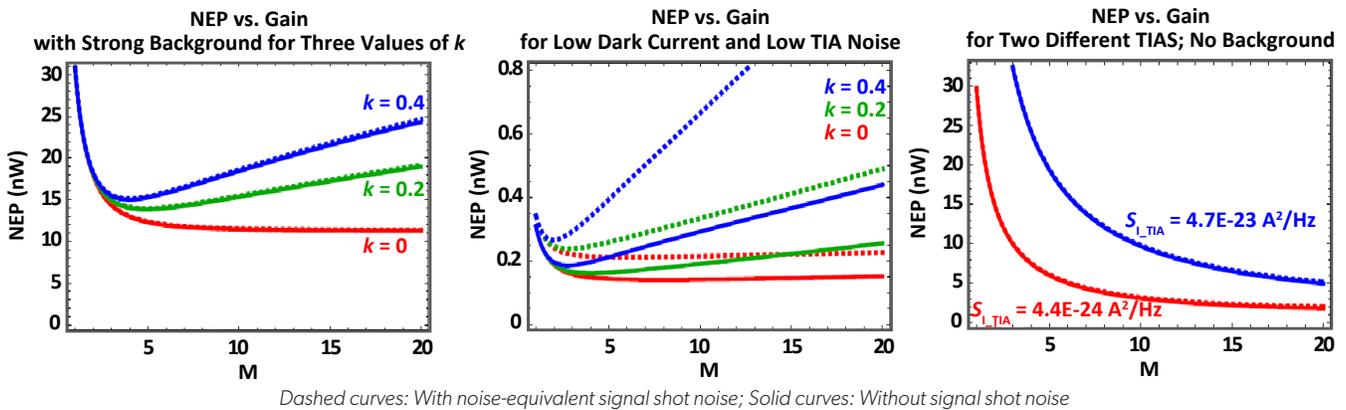
Figure 11: NEP vs. M curves calculated for a photoreceiver assembled from a 200 µm k = 0.2 Allegro APD and a COTS TIA, demonstrating that, except in conditions of exceptionally low dark current and TIA noise, the two alternate definitions of NEP are substantially the same.

Noise-Equivalent Input (NEI)

The acronym *NEI* is used by the imaging community for a different purpose than the meaning here. When discussing passive imagers, *NEI* means noise-equivalent irradiance and is a way of expressing the *NEP* as a spectral irradiance in W m^{-2} nm^{-1}. However, in this technical note, *NEI* is used to represent the signal level in photons that would result in a mean output level of the same magnitude as the RMS noise of a CTIA-based photoreceiver. The noise-equivalent signal is expressed in terms of photons rather than an optical power because the response of a CTIA photoreceiver is proportional to the total number of photons delivered by an optical pulse rather than to its instantaneous optical power during the pulse.

As with *NEP*, there are two alternate definitions of *NEI*. The first definition, $NEI_{SNR=1}$, is the signal level for which the photoreceiver *SNR* is unity. The second definition is the signal level for which the photoreceiver average output will be equal in magnitude to its RMS noise in the absence of an optical signal.

To find $NEI_{SNR=1}$ for a photoreceiver that is assembled from a conventional InGaAs APD and a CTIA, equate the numerator and denominator of Equation 29 and solve for E_{signal}; convert the result to photons by multiplying the energy in joules by $5.034117 \times 10^{18}\,\lambda$ [photons J^{-1} µm^{-1}]:

Equation 36:

$$NEI_{SNR=1} = \frac{MF + \sqrt{(MF)^2 + 4\left(N_{CTIA}^2 + MF\frac{\tau}{q}(I_{dark} + I_{background})\right)}}{2\,M\,QE} \quad \text{[photons]}.$$

In Equation 36, the definition of R_{charge} from Equation 9 was applied to eliminate the wavelength.

The case for a CTIA receiver that uses a $k = 0.02$ Allegro APD is found by solving Equation 30 for E_{signal} with $SNR = 1$ [41]:

Equation 37:

$$NEI_{SNR=1} \approx \frac{MF + \sqrt{(MF)^2 + 4\left[N_{CTIA}^2 + \frac{\tau}{q}\left(I_{dp}\sum_{i=1}^{stages}\left[(M_s^{i-1})^2\,F(M=M_s^{i-1})\right] + I_{background}\,MF\right)\right]}}{2\,M\,QE} \quad \text{[photons]}.$$

Neglecting the shot noise on the hypothetical noise-equivalent signal, the *NEI* of a CTIA-based photoreceiver is:

Equation 38:

$$NEI = 5.03411E18\,\lambda[\mu m]\frac{N_Q(E_{signal}=0)}{R_{charge}} = \frac{N_Q(E_{signal}=0)}{M\,QE}.$$

Using Equation 22 for the charge noise of a CTIA photoreceiver assembled from a conventional InGaAs APD, the *NEI* is:

Equation 39:

$$NEI = \frac{\sqrt{N_{CTIA}^2 + \left[\frac{\tau}{q}(I_{dark} + I_{background})\right]MF}}{M\,QE} \quad \text{[photons]}.$$

Using Equation 24 for the k = 0.02 case, the *NEI* of a CTIA photoreceiver, neglecting signal shot noise, is [41]:

Equation 40:

$$NEI \approx \frac{\sqrt{N_{CTIA}^2 + \frac{I_{dp}}{q}\tau \sum_{i=1}^{stages}\left[(M_s^{i-1})^2 \; F(M = M_s^{i-1})\right] + \left[\frac{1}{q}I_{background} \; \tau + Q_{signal}\right]M \; F}}{M \; QE} \quad [\text{photons}] \; .$$

Unlike *NEP*, the difference in definition between $NEI_{SNR=1}$ and *NEI* is germane in typical applications. Consider the Allegro multi-channel time-of-flight readout integrated circuit (ROIC) characterized by 22 MHz bandwidth, which has an input-referred pixel read noise of about N_{CTIA} = 64 e$^-$ at 235 K. *NEI* plots calculated using Equation 36 and Equation 39 for a 30 μm-diameter conventional InGaAs APD pixel hybridized to one channel of the ROIC are plotted in Figure 12. Negligible background illumination and an effective integration time of τ = 10 ns were assumed. When signal shot noise is omitted (solid curves), the CTIA noise dominates and there is no difference in *NEI* based on the APD impact-ionization coefficient ratio (k). However, the plots of $NEI_{SNR=1}$ are sensitive to k because of the comparatively small noise contributions from APD dark current and CTIA noise. Occasionally a CTIA photoreceiver will be encountered with a sensitivity specification of *NEI* < 1 photon. This is only possible using the form of *NEI* that omits the signal shot noise because $NEI_{SNR=1}$ given by Equation 36 or Equation 37 cannot assume values less than unity, even with a completely noiseless TIA and APD. Thus, a published *NEI* that is less than one does not mean that the CTIA photoreceiver can measure the laser-pulse photon number with sub-photon accuracy; it means that shot noise on the signal charge is guaranteed to be the dominant noise source (and direct calculation of N_Q using the actual value of Q_{signal} would be a better method).
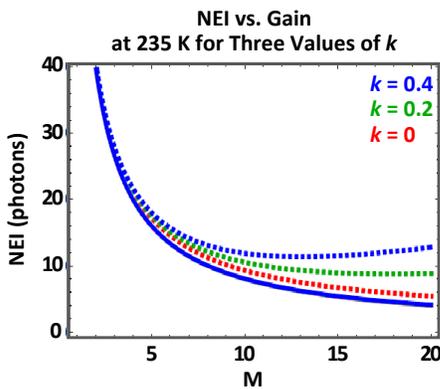
NEI vs. Gain at 235 K for Three Values of k

Dashed curves: with noise-equivalent signal shot noise
Solid curves: without signal shot noise

Figure 12: NEI vs. M curves calculated for a photoreceiver assembled from a 30 μm k = 0.2 Allegro APD pixel and an Allegro multi-channel time-of-flight ROIC characterized by 22 MHz bandwidth, comparing two alternate definitions of NEI for different values of k.

Relationship Between NEP and NEI

Sometimes it is desirable to compare the sensitivity of an RTIA photoreceiver to a CTIA photoreceiver, or to express the sensitivity of one type of receiver in the units characteristic of the other type. The best way to approach the problem is by first calculating the noise-equivalent signal of a given receiver in units that are "native" to the receiver: the optical power for an RTIA photoreceiver or the photon number of an optical pulse for a CTIA photoreceiver. Then, with either *NEP* or *NEI* in hand, specify the photon number that would result in a peak instantaneous optical power equal to the calculated *NEP* or the peak optical power of a pulse having a photon number equal to the calculated *NEI*. Notice that, whereas both *NEP* and *NEI* can themselves be defined without reference to a specific signal pulse shape, the problem of expressing one in terms of the other is fundamentally indeterminate unless the pulse shape is defined. This is because optical pulse energy (photon number) increases linearly with pulse duration at constant optical power.

Consider this example: a 200 MHz RTIA characterized by an input-referred noise spectral intensity of 10^{-24} A^2/Hz and an APD operated at M = 10 with F = 3.5, QE = 80%, 100 nA of dark current, and negligible background illumination. According to Equation 6, Equation 13, and Equation 34, the receiver *NEP* at 1550 nm, without signal shot noise, is approximately 2 nW. Now suppose the signal is a square optical pulse lasting 5 ns. The total energy of a 5 ns pulse of average power equal to the *NEP* is 1×10^{-17} J, or an *NEI* of approximately 78 photons at 1550 nm. However, suppose the point of reference is a 10 ns signal pulse. Because a 5 ns pulse is well within the bandwidth of a 200 MHz receiver (refer to Equation 8), the receiver would be no more responsive to a 2 nW pulse lasting 10 ns, even though twice as many photons would be delivered by such a signal. The *NEI* of the receiver would be worse, without anything about the receiver changing. This is why it is crucial to perform sensitivity comparisons using a specific signal pulse shape that is physically meaningful for a given application.

PHOTORECEIVER OUTPUT DISTRIBUTION

The output of an analog APD photoreceiver is the superposition of the TIA output voltage noise with its voltage response

to the charge or current from the APD. The output of the APD is statistically independent from the TIA noise, so the random variable representing the photoreceiver output is the sum of two independent random variables, and its distribution is the convolution of their individual distributions. As with the earlier treatment of the mean and variance of the photoreceiver output, its distribution is normally analyzed at the node between the APD and the TIA, working in units of electrons. This model presents some difficulties of interpretation because the TIA noise is an analog value characterized by the continuous Gaussian distribution of its output voltage, whereas the APD charge output is quantized and obeys the discrete McIntyre distribution of Equation 2. Further, although the McIntyre distribution applies directly to CTIA-based photoreceivers that sense the total integrated charge delivered by a current pulse, distributions of discrete charge must be related somehow to distributions of instantaneous current in order to analyze RTIA-based photoreceivers.

TIA Input Noise Distribution

In practice, the lack of rigor inherent in using the Gaussian distribution as though it were a discrete distribution is not a serious difficulty. Little accuracy is lost if the random variable n representing the TIA noise is restricted to integer values so that the Gaussian distribution function $P_{TIA}(n)$ can be interpreted as the probability of the TIA noise taking on a value within a band of unit width centered on n. For the purpose of convolving $P_{TIA}(n)$ with the APD output distribution, n represents a quantity of charge in units of electrons:

Equation 41:

$$P_{TIA}(n) = \frac{1}{\sqrt{2\pi\,\mathrm{var}(n)}} \exp\left[-\frac{(n-\bar{n})^2}{2\,\mathrm{var}(n)}\right].$$

In the case of a CTIA, the interpretation of the noise-equivalent input electron count, n, appearing in Equation 41 is straightforward: the output voltage of the CTIA fluctuates with a Gaussian distribution characterized by a particular mean and variance; if those voltages are transformed to the CTIA input by application of its conversion gain, an equivalent number of input electrons results.

The significant challenge is how to relate quantities of electrons to currents, and vice versa, for analysis of RTIA-based photoreceivers. The input-referred current noise of an RTIA expresses its output voltage noise in terms of the magnitude of current from the APD that would result in an output voltage response of the same size. Likewise, an input-referred charge noise for an RTIA must somehow indicate how much mobile charge inside the APD would result in current flow equal in magnitude to the RTIA input-referred current noise.

Therefore, in principle, solving the problem for an APD (how many charge carriers to associate with a particular output current) solves the problem for an RTIA characterized by a particular input-referred noise current.

As will be expanded on below, the product of the TIA input-referred noise current and an effective integration time $t_{transit}$ gives the quantity of charge that, if delivered over the same time span, would produce an APD output current of equal magnitude:

Equation 42:

$$\mathrm{var}(n)_{RTIA} = \frac{t_{transit}^2\, S_{I\_TIA}\, BW}{q^2}.$$

However, it is important to remember that, to the extent a real RTIA has some signal-integrating character, the effective time span that scales between current and charge in Equation 42 is generally longer than the physical junction transit time of the APD.

As a practical matter, it is better to find the RTIA effective input charge noise empirically for the specific signal-pulse shape of interest. Consider, for example, the case of amplitude statistics collected on the output voltage waveform from an analog APD photoreceiver, with the APD operated at unity gain (essentially a p-i-n photodiode). For any APD receiver of practical interest, the TIA noise will dominate the noise contribution from the APD at this operating point, so the standard deviation of the output voltage waveform will be a measure of the TIA output voltage noise. If the unity-gain photoreceiver is then illuminated with optical signal pulses of calibrated energy, chosen to be well above the receiver noise floor, the difference in the mean output-voltage peak height measured between pairs of chosen signal levels divided by the difference in the mean signal charge, found from Equation 9, gives a conversion gain in units of V/e⁻. With the typical caveat that conversion gain depends on input current pulse shape, and the caution that conversion gain is subject to saturation outside the linear dynamic range of the amplifier, the conversion gain arrived at by this measurement can then be used to scale the measured output voltage noise of the RTIA to an equivalent input charge noise.

APD Output Distribution

The McIntyre distribution applies directly to the APD charge output when a transient current pulse completes within the effective integration period (τ) of a CTIA. However, the distribution of the APD instantaneous current output that is relevant to RTIA photoreceivers is harder to calculate accurately.

In principle, the Shockley-Ramo theorem allows calculation of the instantaneous current at the APD terminals from the instantaneous count of electrons and holes within the APD junction, $n_e(t)$ and $n_h(t)$, and their respective saturation velocities, v_{se} and v_{sh} [31],[32],[42]:

Equation 43:

$$i(t) \approx \frac{q}{w} \left[v_{se}\, n_e(t) + v_{sh}\, n_h(t) \right],$$

where w is the junction width. [43] Equation 43 can be recast in terms of junction transit times for electrons and holes, respectively $t_e = w/v_{se}$ and $t_h = w/v_{sh}$:

Equation 44:

$$i(t) \approx q \left[\frac{n_e(t)}{t_e} + \frac{n_h(t)}{t_h} \right].$$

This is the relationship applied in Equation 42 to express the input-referred noise current of an RTIA as a certain number of carriers, although it does not resolve which type of carrier to use for the transit time.

To model the APD output distribution, a further difficulty is that the McIntyre distribution does not give the instantaneous carrier populations of the junction, $n_e(t)$ and $n_h(t)$. It models total output carrier count (n) for a given total input of primary carriers (a), without regard to the time evolution of either population. As discussed earlier in "Gain-Bandwidth Effects Limiting Signal Response", the daughter carriers generated by the impact ionization chain initiated by any given primary carrier are not created simultaneously, and the lifetime of any carrier in the junction depends on its polarity and the location in the junction where it was generated. Moreover, the signal photons that generate primary photocarriers do not arrive at the APD simultaneously. A detailed numerical simulation is required to accurately model APD current statistics. However, some simplifying assumptions that permit a simpler analysis are often applicable.

An APD impulse response function depends on the APD structure and operating point. However, in many cases, the impact-ionization chain triggered by a given primary carrier will complete before the slower of the secondary carriers created by the avalanche process have exited the junction.

For example, the example Allegro InGaAs APDs discussed in this technical note have a thick InGaAs absorber near the anode and a thinner InAlAs multiplier near the cathode (Figure 1). Secondary electrons generated by impact ionization in the multiplier must drift only a short distance before exiting the junction at the cathode, but the secondary holes that are generated along with those electrons must drift all the way back through the thick absorber before exiting the junction at the anode. Hole transport accounts for the majority of the current impulse because of the longer path length traversed by the holes, and the comparatively small difference in saturation drift velocity between holes and electrons. Consequently, for most operating conditions, the impact-ionization process that generates the secondary holes has enough time to complete before the first of the daughter holes have left the junction. The peak of the current impulse therefore tends to correspond to the peak instantaneous hole population in the junction, which also happens to be the total number of secondary holes generated by impact ionization. This relationship does not hold for all APDs in all operating conditions, but it is often the case. When applicable, this argument links the peak of the APD instantaneous current to the McIntyre distribution, justifying the use of Equation 2 to model the distribution of current peak height. It also associates the hole transit time, t_h, with the unspecified transit time $t_{transit}$ appearing in Equation 42 for the input-referred charge variance of an RTIA.

The timing of primary carrier generation and the overlap of current impulses originating from different primary carriers presents a further complication for modeling APD output. If the photons of an optical signal were to arrive in a pulse with a duration somewhat shorter than the junction transit time, the photocurrent pulse height would be relatively well modeled by Equation 2 because all the secondaries generated by all the primaries would be simultaneously present in the junction at some point. However, depending on the APD, the junction transit time is usually sub-nanosecond, whereas most applications involve optical pulses of longer duration. when some, but not all, of the impact-ionization chains overlap in time, determination of the peak height distribution is not an easy problem to address in closed form. To the extent that the peak of an optical pulse is flat and broad compared to the APD junction transit time, some insight can be gained from analysis of steady-state current.

In the case of stable dark current or CW illumination, generation of primary carriers is a Poisson process. Primary-carrier generation is continuous, and the probability that any

[42] 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 Devices, vol. 49, no. 5, pp. 770-781, 2002.

[43] Equation 43 is an approximate statement of the Shockley-Ramo theorem. Technically, the total current is a summation over all the carriers present in the junction of a current contribution from each individual carrier, factoring in its time-dependent velocity. In Equation 43, only the electron and hole populations are presumed to vary in time, and the saturation drift velocities represent averages over time and the population of each carrier type.

given number of primary carriers will be generated within any given time interval depends solely on the duration of that time interval. The probability that an average primary current, $I_{primary}$, will inject a primary carriers into the APD multiplier within the junction transit time $t_{transit}$ is:

Equation 45:

$$P_{Poisson}(a) = \frac{\left(\dfrac{i_{primary}}{q} t_{transit}\right)^a \exp\left(-\dfrac{i_{primary}}{q} t_{transit}\right)}{a!}.$$

Within a time bin of width equal to $t_{transit}$, the output distribution of the APD is the sum over the primary carrier count of McIntyre distributions, weighted by the Poisson distribution of the primary carrier count:

Equation 46:

$$P_{APD}(n) = \sum_a P_{Poisson}(a) \times P_{McIntyre}(n).$$

Equation 46 applies to InGaAs APDs of simple structure in which the majority of the dark current is generated in the same layer as the photocurrent. Equation 45 and Equation 46 also apply to the photocurrent of a multi-stage $k = 0.02$ Allegro APD; however, the dark current distribution requires separate consideration of output from each stage. Equation 16, Equation 17, and Equation 18 can be used to estimate the gain per stage, M_s, and the primary dark current per multiplying stage, I_{dp}, of the $k = 0.02$ APD. Then, for stage i, the distribution of the dark current is approximately equal to Equation 46 where I_{dp} has been substituted for $I_{primary}$ in Equation 45 and $P_{McIntyre}$ is calculated with $M = M_s^{i-1}$. Important comments regarding the accuracy of this approximation are provided in "RTIA Case for Multi-Stage Allegro APDs".

When applying Equation 45 and Equation 46 to steady-state current in a CTIA receiver, such as dark current or background photocurrent, the CTIA effective integration period, τ, is used in place of the junction transit time $t_{transit}$.

Convolution of APD and TIA Distributions

The probability that the APD output and the TIA input-referred noise will sum to a particular quantity of charge, n, is given by the discrete convolution:

Equation 47:

$$P_{RX}(n) = (P_{TIA} \times P_{APD})[n] \equiv \sum_i P_{TIA}(i) P_{APD}(n-i).$$

The discrete random variable n represents the total output of the photoreceiver, referred to the node between APD and TIA.

In the case of the $k = 0.02$ APD, the distributions of the photocurrent and the dark current generated in each multiplier stage are distinct, so the distribution of the photoreceiver output is found by convolving all of the APD-related distributions with the TIA distribution:

Equation 48:

$$P_{RX} = P_{TIA} \times P_{APD\_photo} \times P_{APD\_dark1} \times P_{APD\_dark2} \times \dots$$

Example output distributions calculated using Equation 47 for the hypothetical 200 MHz RTIA photoreceivers assembled from 200 µm-diameter $k = 0.2$ APDs are plotted in Figure 13.
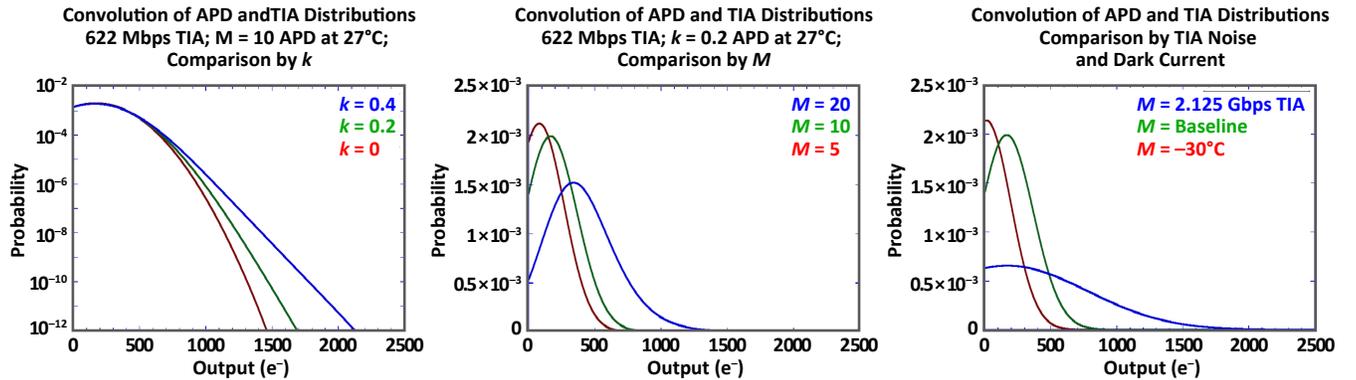
Figure 13: Photoreceiver output distribution functions calculated for a photoreceiver assembled from a 200 µm $k = 0.2$ Allegro APD and a COTS TIA, demonstrating how the shape of the receiver output distribution depends on the APD effective ionization rate ratio (k), the APD mean gain (M), the TIA input-referred noise, and the APD dark current.

The APD steady-state dark current was used to calculate the curves in Figure 13. However, the results would be equivalent for any combination of primary photocurrent and dark current having the same sum. At 27°C, the primary dark current levels were 2.57 nA at $M = 5$, 2.72 nA at $M = 10$, and 3.17 nA at $M = 20$; at -30°C, the primary dark current level was 0.16 nA at $M = 10$. A junction transit time of $t_{transit} = 1$ ns was assumed for the purpose of calculating the TIA noise; this resulted in input-referred charge noise levels of 185 e⁻ for the 622 Mbps TIA and 603 e⁻ for the 2.125 Gbps TIA. Except where specified, the default values used in the calculations were $M = 10$, $k = 0.2$, and $T = 27°C$.

Photoreceiver output distributions (scaled to the node between the APD and the TIA) for effective impact-ionization rate ratios of and $k = 0$, 0.2, and 0.4 are compared in Figure 13, left, which plots the data on semi-logarithmic axes to emphasize that k has a significant impact on the high-output tail of the distribution. However, the use of semi-log axes in Figure 13, left, conceals the other important trend in k: the median of the output distribution shifts to higher output levels as k drops. In most applications, the signal photocurrent is stronger than the dark current, and the photoreceiver detection threshold is set to a value far in the high-output tail of the dark current distribution, but below the median of the signal photocurrent distribution. APD photoreceivers assembled from APDs with low values of k are advantageous because, for a given detection threshold, they have both a lower false alarm rate and a higher signal detection efficiency than those assembled from APDs with higher values of k.

Note that linear axes are used in Figure 13, center and right, which prevents visual comparison to Figure 13, left. However, the green curves in all three panels correspond to the same baseline set of conditions: $M = 10$; $k = 0.2$; $T = 27°C$; and 622 Mbps TIA. The variation of the output distribution with APD gain is shown in Figure 13, center, for $M = 5$, 10 and 20. As would be expected, the median of the output distribution shifts to higher output values as M increases. However, the distribution also broadens and skews to higher output. The baseline case of the 622 Mbps TIA is compared to the noisier 2.125 Gbps TIA model in Figure 13, right, which also compares operation at 27°C to –30°C to demonstrate the influence of varying APD dark current.

In the introductory section "Avalanche Gain and Gain Distribution", the point was made that the Gaussian approximation of the output distribution of an APD is incorrect for predicting behavior in the high-output tail (Figure 5). That point is revisited for the total photoreceiver output in Figure 14, which analyzes the same case as Figure 13, left. The usual approximation of an APD output distribution is a Gaussian distribution with a variance calculated using Equation 3 and Equation 5. Comparisons of output distributions calculated by Equation 47 using proper McIntyre distributions for the APD (solid curves) to Gaussian approximations (dashed curves) are shown in Figure 14. As Figure 14 emphasizes, the divergence from Gaussian behavior is larger for larger values of k, and is mainly significant for false alarm-related calculations that are sensitive to the high-output tail of the distribution.
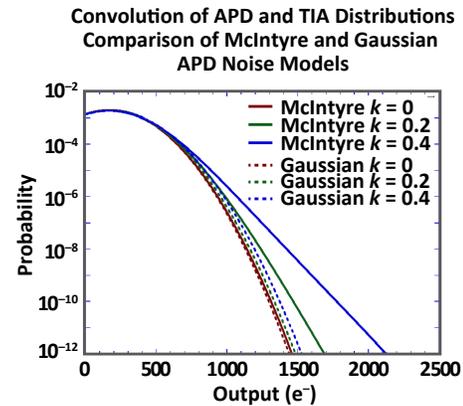
Figure 14: Comparison of photoreceiver output pulse height distributions calculated by convolving TIA noise with either McIntyre (solid) or Gaussian (dashed) APD noise models.

SENSITIVITY METRICS DERIVED FROM OUTPUT DISTRIBUTION

The threshold-detection performance of an APD photoreceiver equipped with a binary decision circuit that registers a detection event if the photoreceiver output signal exceeds a specified detection threshold is illustrated in Figure 15. Output-pulse-height distributions of an analog APD photoreceiver based on the convolution of Equation 47 are plotted for two conditions: with 10 e⁻ of dark current (red); and with 10 e⁻ of dark current plus 50 e⁻ of signal photocurrent (blue). In both cases the APD is characterized by a mean avalanche gain of $M = 10$ and an ionization rate ratio of $k = 0.2$; an input-referred TIA noise of 50 e⁻ is assumed. The dashed black line at an output level of 200 e⁻ represents the detection threshold. In the presence of the optical signal, the shaded area under the blue distribution—its complementary cumulative distribution function (CCDF) at the detection threshold—is equal to the probability of signal detection (P_D): a true positive. Likewise, when the optical signal is not present, the CCDF of the red distribution is equal to the probability of detecting the noise (P_{FA}): a false positive. The areas to the left of the detection threshold are the respective cumulative distribution functions (CDFs), equal to the probabilities of a false negative (for the signal pluse the noise distribution) and of a true negative (for the noise distribution). These distributions are the basis for calculating APD photoreceiver performance metrics, such as optical sensitivity at a given *FAR* or *BER*, and the receiver operating characteristic (ROC).

Threshold Detection of Signal and Noise
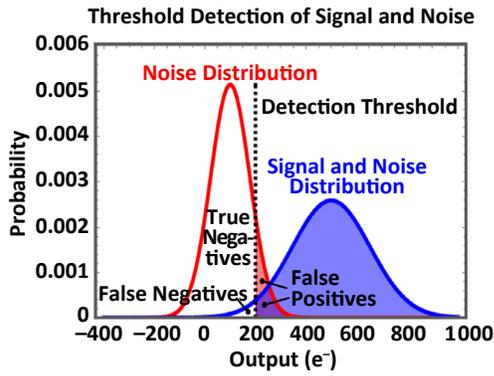
Figure 15: Illustration of the threshold detection of signal and noise.

The CCDFs represented by the shaded areas of Figure 15 can be thought of as the probability per attempt that a fluctuating signal will exceed a given detection threshold; however, the CCDFs do not consider the attempt rate or whether or not the decision circuit is in a state where it can register a detection event. Interpretation of the CCDF in terms of a pulse-detection probability is straightforward when the number of attempts is known and the decision circuit is known to be in a receptive state. For instance, if a single laser pulse is incident on an APD photoreceiver, and the pulse width is shorter than the effective signal integration time of the receiver TIA, then that is one attempt; the CCDF at the detection threshold is equal to P_D provided the comparator is ready at the time the signal pulse arrives. However, in the case of a continuous input, such as dark current or a quasi-CW optical signal that persists longer than the receiver effective integration time, the state of the comparator must be considered. Specifically, decision circuits are commonly built so that they register a detection when the noisy wave-

form rises through the detection threshold, but they do not trigger again if the waveform stays above the threshold for a span of time. Accurate analysis of the *FAR* depends on the probability that the waveform is transitioning through the detection threshold with positive slope, not just the probability that it exceeds the detection threshold, given by the CCDF. [44]

False Alarm Rate (FAR)

The *FAR* resulting from Gaussian-distributed noise was definitively analyzed by Stephen O. Rice in his foundational paper "Mathematical Analysis of Random Noise". [45] Rice analyzed a noisy current waveform defined in terms of uncorrelated random variables for its current (ξ) and the slope of its current (η) at every point in time, t. [46] A false alarm occurs when the current transitions through a threshold value, I_{th}, with a positive slope. Rice showed that the probability of this occurring during the infinitesimal time interval $(t, t + dt)$ is [47]:

Equation 49:

$$PDF_{FA} = dt \int_0^\infty \eta\, p(\xi = I_{th}, \eta; t)\, d\eta \quad [\text{Hz}],$$

where $p(\xi = I_{th}, \eta; t)$ is the joint probability distribution of the current and its slope at time t, assuming the random variable for the current has the value I_{th}. Rice's classic result for *FAR* applies to Gaussian-distributed noise, for which $p(\xi = I_{th}, \eta; t)$ is the bivariate normal distribution. In the case of two uncorrelated random variables, the bivariate normal distribution is the product of two single-variable Gaussian distributions:

---

[44] Technically, accurate calculation of PDE also requires considering whether the comparator is armed and ready to register a detection event. This is a significant issue in photoreceiver systems with a long dead time, such as Geiger-mode APD photoreceivers, or when the detection threshold is set far down in the noise. However, most photoreceivers are operated with the detection threshold set far out in the tail of the noise distribution, in which case the probability that the comparator will be unable to respond to a signal pulse due to an immediately preceding false alarm is negligible.

[45] S. O. Rice, "Mathematical Analysis of Random Noise," Bell System Technical Journal, vol. 23 no. 3 & vol. 24 no. 1, pp. 282-332 & pp. 46-156, 1944 & 1945.

[46] Although Rice explicitly analyzes the case of a noisy current waveform, corresponding to the output of an RTIA-based APD photoreceiver referred to the node between APD and TIA, the general mathematical treatment can be adapted to analyze continuously reset CTIA-based photoreceivers.

[47] As will shortly be made explicit, the normalization of $p(\xi = I_{th}, \eta; t)$ gives a factor of A^{-2} Hz^{-1} when ξ is in units of A and η is in units of A/s; multiplication of $p(\xi = I_{th}, \eta; t)$ by η, followed by integration $d\eta$, results in units of Hz. When PDF_{FA} is integrated over a finite time span to find the probability of a positive-slope threshold crossing during that time span, the factor of seconds resulting from integration dt cancels the factor of Hz in PDF_{FA}, resulting in a unitless probability. Rice wrote in terms of integrating PDF_{FA} over the interval of one second to find the expected number of positive-slope threshold crossings per second, which could then be divided by one second to find the *FAR*. Equivalently, if the *FAR* is understood to be the probability density of false alarms that are uniformly distributed in time—a quantity that can be measured by counting false alarms during a suitable sample period and dividing by that sample period—then PDF_{FA} (without the differential dt) is the *FAR*.

Equation 50:

$$p(\xi, \eta; t)_{Rice} = \frac{1}{2\pi \sqrt{\mathrm{var}(\xi)\,\mathrm{var}(\eta)}} \exp\left[-\frac{1}{2}\left(\frac{[\xi - \bar{\xi}]^2}{\mathrm{var}(\xi)} + \frac{[\eta - \bar{\eta}]^2}{\mathrm{var}(\eta)}\right)\right] \quad [\mathrm{A}^{-2}\,\mathrm{Hz}^{-1}].$$

Noting that the average slope, $\bar{\eta}$, must be zero to ensure that $I(t)$ does not diverge, substitution of Equation 50 in Equation 49 gives:

Equation 51:

$$\begin{aligned}
PDF_{FA\_Rice} &= dt\,\frac{1}{2\pi\sqrt{\mathrm{var}(I)\,\mathrm{var}(\eta)}} \exp\left[-\frac{1}{2}\left(\frac{[I_{th} - \bar{I}]^2}{\mathrm{var}(I)}\right)\right]\int_0^\infty \eta\,\exp\left[-\frac{\eta^2}{2\,\mathrm{var}(\eta)}\right]d\eta \\
&= dt\,\frac{\mathrm{var}(\eta)}{2\pi\sqrt{\mathrm{var}(I)\,\mathrm{var}(\eta)}} \exp\left[-\frac{1}{2}\left(\frac{[I_{th} - \bar{I}]^2}{\mathrm{var}(I)}\right)\right] \\
&= \frac{dt}{2\pi}\sqrt{\frac{\mathrm{var}(\eta)}{\mathrm{var}(I)}}\,\exp\left[-\frac{1}{2}\left(\frac{[I_{th} - \bar{I}]^2}{\mathrm{var}(I)}\right)\right] \quad [\mathrm{Hz}].
\end{aligned}$$

The *FAR* is Equation 51 without the differential dt:

Equation 52:

$$FAR_{Rice} = \frac{1}{2\pi}\sqrt{\frac{\mathrm{var}(\eta)}{\mathrm{var}(I)}}\,\exp\left[-\frac{1}{2}\left(\frac{[I_{th} - \bar{I}]^2}{\mathrm{var}(I)}\right)\right] \quad [\mathrm{Hz}].$$

Rice relates the variance of the current and its slope to its autocorrelation function at zero time lag as:

Equation 53:

$$\mathrm{var}(I) = \psi_0 \equiv \lim_{t\to\infty}\frac{1}{t}\int_0^t I(t)\,I(t+\tau)\,dt\bigg|_{\tau=0} \quad [\mathrm{A}^2]; \text{ and}$$

Equation 54:

$$\mathrm{var}(\eta) = -\psi_0'' \equiv -\frac{\partial^2}{\partial\tau^2}\psi\bigg|_{\tau=0} \quad [\mathrm{A}^2/\mathrm{s}^2].$$

The autocorrelation function is itself related to the spectral intensity, S_I,[48] of the noisy current, by inversion of the Wiener-Khinchin theorem [49],[50],[51]:

---

[48] This is S_I total – the total noise current spectral intensity of the photoreceiver, referred to the node between APD and TIA, previously given by Equation 12. Although S_I cancels out in the *FAR* for Gaussian-distributed noise, it is used for the modified calculation for McIntyre-distributed noise, discussed shortly.

[49] N. Wiener, "Generalized Harmonic Analysis," Acta Mathematica, vol. 55, pp. 117-258, 1930.

[50] A. Khinchin, "Korrelationstheorie der stationären stochastischen Prozesse," Mathematische Annalen, vol. 109 no. 1, pp. 604–615, 1934.

[51] Van Der Ziel, A., Noise in Solid State Devices and Circuits (John Wiley & Sons, 1986), pp. 10-12.

Equation 55:

$$\psi(\tau) = \int_0^\infty S_I(f) \cos(2\pi f \tau) df \quad [A^2].$$

Therefore:

Equation 56:

$$var(I) = \psi_0 = \int_0^\infty S_I(f) df \quad [A^2]; \text{ and}$$

Equation 57:

$$var(\eta) = 4\pi^2 \int_0^\infty f^2 S_I(f) df \quad [A^2/s^2].$$

Substituting Equation 56 and Equation 57 into Equation 52, the *FAR* for Gaussian-distributed noise is:

Equation 58:

$$FAR_{Rice} = \frac{1}{2\pi} \sqrt{\frac{4\pi^2 \int_0^\infty f^2 S_I(f) df}{\int_0^\infty S_I(f) df}} \exp\left[-\frac{1}{2}\left(\frac{[I_{th} - \bar{I}]^2}{var(I)}\right)\right] \quad [Hz].$$

When the noise spectrum is white (constant S_I) over a finite bandwidth BW, S_I cancels out in the radical and Equation 58 becomes:

Equation 59:

$$FAR_{Rice} = \sqrt{\frac{\frac{1}{3}BW^3}{BW}} \exp\left[-\frac{1}{2}\left(\frac{[I_{th} - \bar{I}]^2}{var(I)}\right)\right] = \sqrt{\frac{1}{3}} BW \exp\left[-\frac{\Delta I_{th}^2}{2 I_{noise}^2}\right] \quad [Hz].$$

Equation 59 is the expression for *FAR* found in most references, such as the RCA/Burle Electro-Optics Handbook. [52] In Equation 59, the symbol ΔI_{th} is the excess of the detection threshold above the mean current level, and I_{noise} is the standard deviation of the current, as in Equation 15.

Calculating *FAR* with better accuracy at threshold levels set high in the tail of the output distribution of an APD photoreceiver requires the Gaussian distribution of ξ assumed by Rice to be replaced with the convolution of the APD McIntyre-distributed output with the Gaussian-distributed TIA noise, $P_{RX}(n)$, given by Equation 47. $P_{RX}(n)$ is an electron count distribution (referred to the node between the APD and the TIA). However, it can be used for the current distribution through a change of variable. As previously discussed in "APD Output Distribution", Ramo's theorem says that the terminal current of the APD is a mono-tonic function of the instantaneous carrier population, which can be approximated as equal to n:

Equation 60:

$$\xi = \frac{q}{t_{transit}} n \quad [A].$$

Following the rule for change of variable of a probability density function, the current distribution is:

---

[52] Burle Industries, Inc., TP-135 Electro-Optics Handbook, (Burle Industries, Inc., Tube Products Division, 1974), p. 110.

Equation 61:

$$p(\xi)_{RX} = \left| \frac{d}{d\xi} n(\xi) \right| P_{RX}\left[n(\xi) \right] = \frac{t_{transit}}{q} P_{RX}\left(n = \frac{t_{transit}}{q} \xi \right) \quad [\mathrm{A}^{-1}].$$

The joint probability distribution of the current and its slope, equivalent to Equation 50, is:

Equation 62:

$$p(\xi, \eta; t)_{McIntyre} = \frac{1}{\sqrt{2\pi \operatorname{var}(\eta)}} \frac{t_{transit}}{q} P_{RX}\left(n = \frac{t_{transit}}{q} \xi \right) \exp\left[-\frac{\eta^2}{2 \operatorname{var}(\eta)} \right] \quad [\mathrm{A}^{-2}\,\mathrm{Hz}^{-1}].$$

Substituting the modified joint probability distribution into Equation 49 gives:

Equation 63:

$$PDF_{FA\_McIntyre} = dt \int_0^\infty \eta \, p(\xi = I_{th}, \eta; t) \, d\eta \quad [\mathrm{Hz}]$$

$$PDF_{FA\_McIntyre} = dt \frac{1}{\sqrt{2\pi \operatorname{var}(\eta)}} \frac{t_{transit}}{q} P_{RX}\left(n_{th} = \frac{t_{transit}}{q} I_{th} \right) \int_0^\infty \eta \, \exp\left[-\frac{\eta^2}{2 \operatorname{var}(\eta)} \right] d\eta$$

$$= dt \frac{\operatorname{var}(\eta)}{\sqrt{2\pi \operatorname{var}(\eta)}} \frac{t_{transit}}{q} P_{RX}\left(n_{th} = \frac{t_{transit}}{q} I_{th} \right)$$

$$= \frac{dt}{2\pi} \sqrt{\frac{\operatorname{var}(\eta)}{\operatorname{var}(I)}} \frac{t_{transit}}{q} P_{RX}\left(n_{th} = \frac{t_{transit}}{q} I_{th} \right) \sqrt{2\pi \operatorname{var}(I)} \quad [\mathrm{Hz}].$$

Note that the last line of Equation 63 is multiplied by $1 = \sqrt{2\pi \, var(I)} / \sqrt{2\pi \, var(I)}$ to cast the expression in the same form as Equation 51, whereby the operations of Equation 52 through Equation 58 can be applied to find the *FAR* equivalent to Equation 59:

Equation 64:

$$FAR_{McIntyre} = \sqrt{\frac{2\pi}{3}} \frac{t_{transit}}{q} I_{noise} \, BW \, P_{RX}\left(n_{th} = \frac{t_{transit}}{q} I_{th} \right) \quad [\mathrm{Hz}].$$

The conditions for which the $P_{RX}(n)$ curves of Figure 14 were calculated result in factors in front of P_{RX} of 54.651 GHz for $k = 0$, 55.489 GHz for $k = 0.2$, and 56.315 GHz for $k = 0.4$. The *FAR* calculated using Equation 64 is compared to that calculated using Equation 59 in Figure 16. The more-realistic model reveals that a few standard deviations beyond the mean ($\bar{\eta} = 170.06$; $\sigma_{k=0} = 1188.82$; $\sigma_{k=0.2} = 191.71$; $\sigma_{k=0.4} = 194.57$), *FAR* drops off more slowly with increasing detection threshold—and is much more sensitive to k—than predicted by Rice's model.

False Alarm Rate vs. Detection Threshold
Comparison of FAR Models

Legend:
— Allegro / McIntyre *k* = 0
— Allegro / McIntyre *k* = 0.2
— Allegro / McIntyre *k* = 0.4
···· Rice / Gaussian *k* = 0
···· Rice / Gaussian *k* = 0.2
···· Rice / Gaussian *k* = 0.4
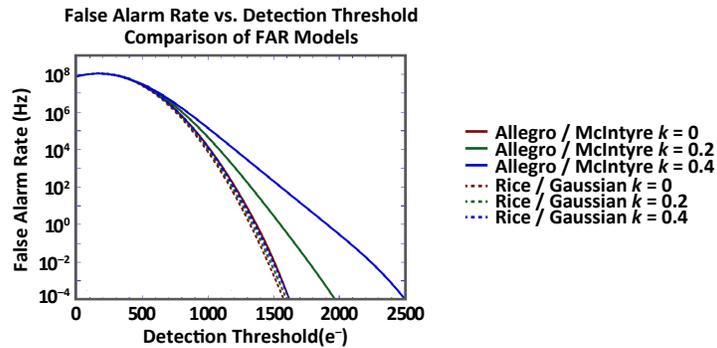
Figure 16: Comparison of FAR calculated by convolving TIA noise with McIntyre APD noise (solid), or using the Rice [45] FAR model (dashed).

Equation 64 can be applied to calculate the *FAR* of either RTIA- or CTIA-based photoreceivers. In the CTIA case, the effective integration period τ is used in place of the junction transit time t_{transit}, in which case the product of noise current and integration time, scaled by the elementary charge, can be recognized as the total charge noise (N_Q) given by Equation 22 or Equation 24, in the absence of an optical signal:

Equation 65:

$$FAR_{McIntyre} = \sqrt{\frac{2\pi}{3}}\, N_Q\, BW\, P_{RX}\left(n_{th} = \frac{t_{transit}}{q}\, I_{th}\right)\Bigg|_{Q_{signal}=0} \quad [\text{Hz}].$$

Bit Error Rate (BER)

The *BER* of a digital optical communications link is defined in terms of overlapping distributions similar to the diagram of Figure 15. [53] In the illustration of binary signal detection and bit errors in Figure 17, the amplitude distribution of the signal level coding a binary "0" is red, and the distribution of the signal level coding a binary "1" is blue.
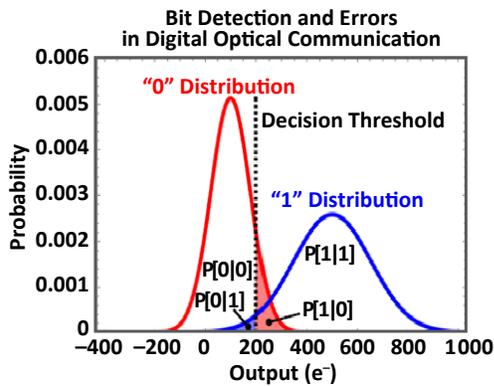
Figure 17: Illustration of binary signal detection and bit errors.

A bit error occurs when a "0" is sent but the receiver registers a "1", or when a "1" is sent but the receiver registers a "0"; the probabilities of these errors are respectively written $P[1|0]$ and $P[0|1]$. $P[1|0]$ is the CCDF of the "0" distribution, whereas $P[0|1]$ is the CDF of the "1" distribution, both evaluated at the decision threshold, n_t:

Equation 66:

$$P[1|0] = 1 - \sum_{n=-\infty}^{n_t} P_{RX}(n_0); \text{ and}$$

Equation 67:

$$P[0|1] = \sum_{n=-\infty}^{n_t} P_{RX}(n_1),$$

where n_0 and n_1 are discrete random variables that represent the effective carrier count at the node between the APD and the TIA, calculated for primary photocurrent levels corresponding to optical-signal level coding for binary "0" or "1" values of the. $P_{RX}(n)$ is the distribution of the photoreceiver output, referred to this node, and is calculated according to Equation 47 for conventionally structured APDs and according to Equation 48 for multi-stage $k = 0.02$ APDs. When making calculations for conventionally structured APDs using Equation 47, the primary current used in Equation 45 and Equation 46 to compute the APD output distribution for convolution with the TIA noise is the sum of the primary dark current and photocurrent; when making calculations for $k = 0.02$ APDs using Equation 48, the primary photocurrent and dark current are treated in separate distributions, which are subsequently convolved as described in "Convolution of APD and TIA Distributions".

The primary photocurrent is found from the optical power incident on the APD using $M = 1$ in Equation 6 and Figure 7. Because optical communication signals are usually generated by modulating a CW laser, the optical-power level coding a binary "0" value, P_0, is generally defined relative to the power level coding a binary "1" value, P_1:

Equation 68:

$$P_0 = P_1\, 10^{-\frac{ER}{10}},$$

where *ER* is the extinction ratio of the modulator in dB (typically 15 dB to 20 dB for Mach-Zehnder interferometer-type lithium niobate electro-optic modulators). In communications applications, optical-signal power is normally specified on a logarithmic scale relative to 1 mW, whereas the equations of this technical note are scaled in standard units (watts). To convert between the two:

Equation 69:

$$P_{watts} = 1\,\text{mW} \times 10^{\frac{P_{dBm}}{10}}.$$

The frequency with which "0" and "1" bits occur within a binary sequence must be known to calculate the *BER* and also the sensitivity at a given *BER*, since this determines the

[53]R. Ramaswami, K. N. Sivarajan and G. H. Sasaki, Optical Networks – A Practical Perspective, 3rd ed. (Elsevier, Burlington, MA, 2010), pp. 264-269.

weighting of both the error rate and the average optical power. If R_1 is the rate of occurrence for transmission of "1" and $(1 - R_1)$ is the rate of occurrence of "0", the BER is:

Equation 70:

$$BER = R_1 P[0|1] + (1 - R_1)P[1|0].$$

It is common to specify the sensitivity of an optical communications receiver in terms of the average optical-signal power required to achieve a benchmark BER (e.g., 10^{-12}) given a benchmark binary sequence (e.g., PRBS23, a pseudorandom $2^{23} - 1$ bit binary sequence). The average power P_{av} is related to P_1, ER, and R_1 by:

Equation 71:

$$P_{av} = R_1 P_1 + (1 - R_1) P_1 \, 10^{-\frac{ER}{10}}.$$

Often, binary sequences for which $R_1 = 0.5$ are used, in which case, if ER is on the high side (e.g., > 15 dB), $P_{av} \approx 0.5\, P_1$.

When the detector is a simple (non-avalanche) photodiode, the output distribution of the photoreceiver is Gaussian, and convenient analytic formulas for the CCDF and CDF apply to $P[1|0]$ and $P[0|1]]$. Assuming $R_1 = 0.5$, the optimal decision threshold is very close to [36]:

Equation 72:

$$n_t\Big|_{optimum} \approx \frac{\langle n_1 \rangle \sqrt{\mathrm{var}(n_0)} + \langle n_0 \rangle \sqrt{\mathrm{var}(n_1)}}{\sqrt{\mathrm{var}(n_0)} + \sqrt{\mathrm{var}(n_1)}}.$$

The mean and standard deviation of n_0 and n_1 appearing in Equation 72 are the photoreceiver signal and noise under the "0" and "1" signal conditions, which can be calculated as described in "Mean (Signal)" and "Variance (Noise)".

In the case of a photoreceiver with Gaussian-distributed output, if the decision threshold is set as in Equation 72, the BER is [36]:

Equation 73:

$$BER = \frac{1}{2}\mathrm{erfc}\left[\frac{\langle n_1 \rangle - \langle n_0 \rangle}{\sqrt{2}\left(\sqrt{\mathrm{var}(n_0)} + \sqrt{\mathrm{var}(n_1)}\right)} \right].$$

If the TIA noise of the receiver dominates the shot noise on the dark current and photocurrent of the detector (including the photocurrent shot noise when receiving a "1"), and if the modulator extinction ratio is large, then the BER can be approximated in terms of the signal-to-noise ratio of the receiver, as defined in Equation 25 [54]:

Equation 74:

$$BER \approx \frac{1}{2}\mathrm{erfc}\left(\frac{SNR}{2\sqrt{2}} \right).$$

Equation 74 is often used for quick back-of-the-envelope estimates because of its simplicity. $BER = 10^{-9}$ corresponds to $SNR \approx 12$; $BER = 10^{-12}$ corresponds to $SNR \approx 14$. Because Equation 74 is predicated on the dominance of $\langle n_1 \rangle$ in Equation 73, the sensitivity at a given BER is found by applying Equation 25 to solve for the optical power, P_1, that results in the specified SNR; Equation 71 is then used to find the corresponding average signal power, which is the sensitivity at that BER.

For several reasons, Equation 74 is not accurate for APD-based photoreceivers. First, as discussed previously in the context of FAR, the distribution of the APD output is not Gaussian, and divergence of the tail of the distribution from the Gaussian approximation several standard deviations away from its mean can significantly impact $P[1|0]$. Also, the skewness of the APD output distribution means that Equation 72 for the optimal decision threshold is less accurate for APD-based photoreceivers than for p-i-n photoreceivers. Second, neglecting the shot noise on the APD photocurrent, using $\sqrt{\mathrm{var}(n_0)} \approx \sqrt{\mathrm{var}(n_1)}$ to simplify the form of Equation 74 results in a bad approximation. In practice, the extra signal shot noise when a "1" is being received affects both the optimal decision threshold and the bit-error probabilities.

A more accurate calculation of BER based on the proper distributions involves directly calculating $P[1|0]$ and $P[0|1]$ according to Equation 66 and Equation 67, using either Equation 47 or Equation 48 for $P_{RX}(n)$, depending on the APD internal structure. For a given optical-power level coding a "1" (P_1) and a given extinction ratio (ER), BER depends on the APD gain operating point (M) and effective ionization rate ratio (k), as well as the threshold of the decision circuit (n_t). To find the BER sensitivity, $P_{RX}(n_0)$ and $P_{RX}(n_1)$ are calculated numerically for a fixed value of P_1, across a range of M values. For each value of M, BER is minimized with respect to n_t. The M value giving the lowest BER is the optimal gain setting for that value of P_1. In this way, a plot of optimal BER versus average optical signal power can be developed by

[54] P. Bhattacharya, Semiconductor Optoelectronic Devices, Second Edition, (Prentice Hall, Upper Saddle River, NJ, 1997), p. 386.

stepping through values of P_1, using Equation 71 to convert P_1 to average power; the average power for which a particular BER is achieved is the receiver sensitivity at that BER.

Generating $P_{RX}(n)$ is computationally intensive, whereas optimizing n_t is comparatively fast, so an effort should be made to economize on the number of M values tested. One efficient approach is to calculate the gain that maximizes the ratio:

Equation 75:

$$C = \frac{I_{signal\,1} - I_{signal\,0}}{\sqrt{I_{noise\,1}^2 + I_{noise\,0}^2}}.$$

When considering the gain that maximizes the ratio in Equation 75, it is important to keep in mind that it is difficult to operate conventionally structured InGaAs APDs above $M = 20$, or $k = 0.02$ APDs above $M = 50$. I_{signal} and I_{noise} are respectively calculated according to Equation 7 and Equation 15, [55] and the ratio in Equation 75 is essentially the SNR, where the "signal" is the difference in optical power between the "1" and "0" levels. Within the Gaussian approximation, maximizing C will nearly minimize BER, so it is a good starting point for numerical optimization. In general, the Gaussian approximation underestimates the high-output tail of the photoreceiver output distribution, so it will tend to underestimate $P[0|1]$. Optimal gain operating points are often lower than found by maximizing Equation 75, whereas optimal decision thresholds are often higher than Equation 72.

Receiver Operating Characteristic (ROC)

The ROC of an APD photoreceiver equipped with a binary decision circuit is a plot of the true-positive rate (TPR) against the false-positive rate (FPR) under a specified signal condition. The TPR and FPR should be defined for maximal relevance to the physical problem being solved. For instance, suppose that a simple time-of-flight system is configured to look for returns from targets within a range of 5 km, from which the maximum round-trip travel time of the laser pulse would be approximately 33.36 µs. It would be interesting to know both the probability that any given target return will be detected, and the probability that a confounding false alarm will occur during the time span within which a target return is expected. Because there ought to be one return per target for every transmitted laser pulse, the TPR is the pulse detection probability, P_D, as calculated from the CCDF of $P_{RX}(n)$ with optical signal present. However, the raw false-alarm

probability, P_{FA}, calculated from the CCDF of $P_{RX}(n)$ in the absence of an optical signal, is not the natural definition of FPR for this scenario. P_{FA} gives the probability of false alarm per attempt, but not per time interval; P_{FA} alone does not indicate how likely a false alarm will occur while the receiver is waiting to detect a signal return. Instead, the natural definition of FPR for this system is the probability of at least one false alarm occurring during the range gate. Because false alarms are uniformly distributed in time, Poisson statistics can be applied to calculate the probability of zero false alarms occurring during the range gate τ, with the expected value of the number of alarms equal to $FAR \times \tau$:

Equation 76:

$$FPR = 1 - \exp(-FAR \times \tau).$$

FAR is calculated as in Equation 64 for RTIA-based photoreceivers or Equation 65 for CTIA receivers.

Both P_D and FAR are functions of the detection threshold, so the ROC is generated as a parametric plot by varying the detection threshold. Example ROCs calculated for a time-of-flight receiver assembled from a 75 µm-diameter Allegro APD characterized by $k = 0.2$ and an Allegro single-channel time-of-flight application-specific integrated circuit (ASIC) characterized by 37 MHz bandwidth is shown in Figure 18, which assumes an average signal return strength of 100 photons and a 5 km range gate. The calculation was made for a 1550 nm laser pulse of Gaussian shape, of 4 ns full width at half maximum (FWHM), for which the effective signal integration time of the ASIC was 8.2 ns and input-referred charge noise was 314 e⁻.
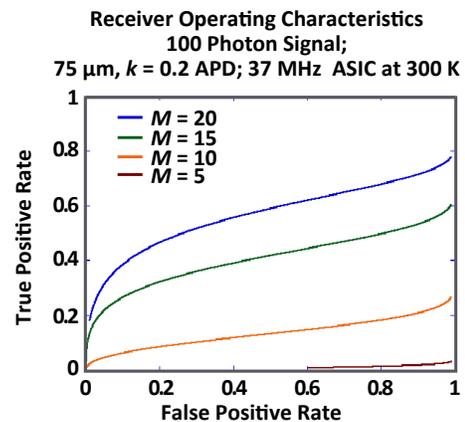
Figure 18: Example ROCs for a time-of-flight receiver with APD operated at four different gains, assuming a 100 photon signal and 5 km range gate.

---

[55]For multi-stage $k = 0.02$ APDs, use Equation 20 for S_{I_total} in Equation 15.

PARAMETERIZATION OF TERMINAL DARK CURRENT FOR EXAMPLE ALLEGRO APDS

This paper presents a methodology to determine photoreceiver performance in a time-of-flight system. The Allegro test devices presented as examples in this paper do not reflect a current Allegro product. Rather, the example devices presented in this paper are intended to help system designers who seek to understand the myriad factors that interplay to affect photoreceiver performance.

Allegro $k = 0.2$ APD

Parameterizations for the Allegro $k = 0.2$ model APDs are shown in Equation 77. These parameterizations are accurate in the range $1 \leq M \leq 20$; however, they diverge from empirical measurements for $M > 20$. It is important to remember that the parameterizations are fit to average device behavior, but dark current varies somewhat from part to part.

Equation 77:

- 75 µm: $I_{dark} = \mathrm{Exp}[0.05 \times (T - 27°C)] \times (-0.080665 + 0.29786M - 0.0076941M^2 + 0.00010214M^3)$ [nA]
- 200 µm: $I_{dark} = \mathrm{Exp}[0.05 \times (T - 27°C)] \times (-0.7902 + 2.7376M - 0.104M^2 + 0.001701M^3)$ [nA]

Allegro $k = 0.02$ APD

The parameterizations for the Allegro $k = 0.02$ model APDs are: shown in Equation 78. These parameterizations are accurate in the range $1 \leq M \leq 50$ but diverge from empirical measurements for $M > 50$. As with the $k = 0.2$ parts, it is important to remember that the parameterizations are fit to average device behavior, but dark current varies from part to part.

Equation 78:

- 75 µm: $I_{dark} = \mathrm{Exp}[0.0234 \times (T - 27°C)] \times (-28.262 + 32.557M - 0.29065M^2 + 0.0036571M^3)$ [nA]
- 200 µm[56]: $I_{dark} = \mathrm{Exp}[0.0234 \times (T - 27°C)] \times (-103.42 + 69.477M + 2.425M^2 - 0.040039M^3)$ [nA]

BURGESS VARIANCE THEOREM FOR MULTIPLICATION AND ATTENUATION

The Burgess variance theorem [4],[5] is applied to introduce the APD excess-noise factor in the introductory discussion in "Avalanche Gain and Gain Distribution", and in connection with attenuation of noisy optical signals toward the end of "CTIA Case for Conventional InGaAs APDs". In this section, derivations of the Burgess variance theorem for these two applications are described, and the theorem is applied to treat attenuation of an optical signal generated by a pulsed laser with large variability of pulse energy.

Derivation

In the case of avalanche gain, a fluctuating output electron count, n, is conceived of as resulting from a fluctuating per-electron discrete gain, m, that is summed over a fluctuating input electron count, a. In the case of attenuation of a noisy optical signal, a fluctuating output photon count, p, is thought of as the result of a fluctuating per-photon binary transmission outcome, t, that is summed over a fluctuating input photon count, b. In the following derivation, the {n, a, m} variable set is explicitly used for avalanche multiplication, remembering that the substitutions {n, a, m} $\rightarrow$ {p, b, t} can be made to analyze the attenuation problem. In general, the same treatment applies to any situation in which the discrete random outcome of a fluctuating number of trials is summed, but different expressions result from the statistics of the different physical processes governing the discrete per-trial outcomes.

If avalanche multiplication was a deterministic process characterized by a constant number of output electrons per input electron, M_{const}, then $n = a \times M_{const}$, and per the basic rule of computing, the variance of the product of a constant with a random variable, $\mathrm{var}(n) = M_{const}^2 \mathrm{var}(a)$. However, when the per-electron discrete gain is, itself, a random variable, the product $a \times m$ does not correspond to n because a single value of m does not multiply every electron of an a-electron input current pulse.

---

[56] The functional form of this fit is different from the other three APDs, with the M^2 coefficient positive and the M^3 coefficient negative; this is not a typo.

Rather, every electron of the fluctuating quantity a is multiplied by a potentially different value of the fluctuating gain m, and var(n) is computed from the statistics of a and $n|a$ (n given a).

Begin with the definition of variance:

Equation 79:

$$\mathrm{var}(n) \equiv \left\langle n^2 \right\rangle - \left\langle n \right\rangle^2 .$$

The task is to calculate $\langle n \rangle$ and $\langle n^2 \rangle$.

Presume there exist discrete distributions for a and for $n|a$. The number of trials (input electron count) might be Poisson-distributed, but it could be anything. The distribution of the output (n) for a given number of trials depends on the physical process. In the case of optical transmission, if exactly b input photons are incident on an attenuator characterized by an average transmission probability T, the transmitted photons will obey a binomial distribution because transmission of each individual photon constitutes a successful Bernoulli trial:

Equation 80:

$$P_{binomial}\left(p|b\right) = \binom{b}{T} T^p \left(1 - T\right)^{b-p} .$$

Likewise, in the case of avalanche multiplication, $n|a$ obeys the McIntyre distribution given in Equation 2.

Assuming the distribution functions $P(a)$ and $P(n|a)$ exist, the expected values can be written symbolically as $\langle n \rangle$ and $\langle n^2 \rangle$ (the mean and mean square). [57] Because $n|a$ is the sum of a random variables, each distributed as m, the expected value $\langle n|a \rangle$ can be rewritten as the expected value of the sum of a random variables, m_i. The linearity of the expectation operator can be applied to write the expectation of the sum as the sum of the individual expectations, as:

Equation 81:

$$\left\langle n \right\rangle = \sum_{a=0}^{\infty}\sum_{n=0}^{\infty} P(a) \cdot P(n|a) \cdot n = \sum_{a=0}^{\infty} P(a) \cdot \left\langle n|a \right\rangle = \sum_{a=0}^{\infty} P(a) \cdot \left\langle \sum_{i=1}^{a} m_i \middle| a \right\rangle = \sum_{a=0}^{\infty} P(a) \cdot \sum_{i=1}^{a} \left\langle m_i \middle| a \right\rangle$$

$$= \sum_{a=0}^{\infty} P(a) \cdot a \left\langle m|a \right\rangle = \sum_{a=0}^{\infty} P(a) \cdot a \cdot M = M \times \left\langle a \right\rangle .$$

In the last line of Equation 81, the mean number of output electrons per input electron, given a input electrons, is written as $\langle m|a \rangle$. The subsequent substitution $M = \langle m|a \rangle$ explicitly assumes that the average gain-per-electron is not a function of the number of input electrons. The equivalent assumption for the case of optical attenuation is that the average per-photon transmission probability $T = \langle t|b \rangle$ is independent of optical signal strength. Therefore, it should be noted that the Burgess variance theorem assumes there is no saturation of the process governing m, which is not always the case for avalanche multiplication or optical absorption.

---

[57]Note that, depending on the details of the specific processes, the limits of the second summation may be physically restricted. For instance, in the specific case of transmission through an attenuator, values of p that are larger than b are not physically possible, so the upper limit of the second summation would be limited to b rather than infinity. On the other hand, in the case of avalanche multiplication, the lower limit could not be smaller than a. However, it is equivalent to regard the contingent probability $P(p|b)$ or $P(n|a)$ to be zero for some values of b or a, and to write the summation from zero to infinity.

The mean square is given by:

Equation 82:

$$\left\langle n^2 \right\rangle = \sum_{a=0}^{\infty} \sum_{n=0}^{\infty} P(a) \cdot P(n|a) \cdot n^2 = \sum_{a=0}^{\infty} P(a) \cdot \left\langle n^2 | a \right\rangle = \sum_{a=0}^{\infty} P(a) \cdot \left[\mathrm{var}(n|a) + \left\langle n | a \right\rangle^2 \right]$$

$$= \sum_{a=0}^{\infty} P(a) \cdot \left[\mathrm{var}\left(\sum_{i=1}^{a} m_i \Big| a \right) + (a \cdot M)^2 \right] = \sum_{a=0}^{\infty} P(a) \cdot \left[\sum_{i=1}^{a} \mathrm{var}(m_i | a) + (a \cdot M)^2 \right]$$

$$= \sum_{a=0}^{\infty} P(a) \cdot \left[a \, \mathrm{var}(m) + a^2 \, M^2 \right] = \left\langle a \right\rangle \mathrm{var}(m) + \left\langle a^2 \right\rangle M^2,$$

where the definition of variance has been used to rewrite $\langle n^2|a \rangle = \mathrm{var}(n|a) + \langle n|a \rangle^2$, and the result, $\langle n|a \rangle = a{\cdot}M$, that was found in Equation 81 has been applied. In the second line of Equation 82, $\mathrm{var}(n|a)$ has been written as the summation of a random variables, each distributed as m, and it has been assumed that each of these random variables is statistically uncorrelated, such that the variance of their sum is equal to the sum of their respective variances. Since the discrete random gain variables are identically distributed, their variances are identical, allowing the collection of terms in the final line of Equation 82. As with the assumption in Equation 81 that the mean per-electron gain $\langle m \rangle$ is independent of the number of input electrons (a), the Burgess variance theorem does not strictly apply to situations in which the gain statistics of different trials are correlated, or in which $\mathrm{var}(m)$ depends on a.

Substitution of Equation 81 and Equation 82 into Equation 79 gives the form of the Burgess variance theorem originally presented in Equation 3:

Equation 83:

$$\mathrm{var}(n) \equiv \left\langle n^2 \right\rangle - \left\langle n \right\rangle^2 = \left\langle a \right\rangle \mathrm{var}(m) + \left\langle a^2 \right\rangle M^2 - M^2 \left\langle a \right\rangle^2 = M^2 \, \mathrm{var}(a) + \left\langle a \right\rangle \mathrm{var}(m).$$

When the Burgess variance theorem is explicitly written for optical transmission, m is generated by a Bernoulli process in which the photon is either transmitted with probability T or blocked with probability $(1 - T)$. Written in the $\{p, b, t\}$ variable set, and using the variance of the Bernoulli distribution for $\mathrm{var}(m) \rightarrow \mathrm{var}(t) = T \times (1 - T)$, the Burgess variance theorem can be written for optical attenuation as:

Equation 84:

$$\mathrm{var}(p) = T^2 \, \mathrm{var}(b) + \left\langle b \right\rangle \mathrm{var}(t) = T^2 \, \mathrm{var}(b) + \left\langle b \right\rangle T (1 - T).$$

Based on the mean and variance of the number of input photons, respectively $\langle b \rangle$ and $\mathrm{var}(b)$, and the average transmission probability of the attenuator, T, Equation 84 can be used to calculate the variance of the transmitted signal.

Application to Attenuation of a Noisy Optical Signal

It is common in the analysis of photoreceiver sensitivity to assume that the amplitude of the optical signal is Poisson-distributed, but this is not always the case. For instance, Figure 19 shows the pulse amplitude histogram measured for a ~100 µJ passively Q-switched Er:glass diode-pumped solid-state (DPSS) laser emitting at 1535 nm. The standard deviation of the laser pulse energy is only approximately 1%, but with an average amplitude of approximately 8×10^{14} photons, 1% of the mean pulse energy is approximately 2×10^{12} photons. If the laser output were Poisson-distributed, its variance would equal its mean; instead, its variance is more than 6×10^9 times the mean. Indeed, the width of a Poisson-distributed signal would not be perceivable on the x-axis of Figure 19.

Gaussian Fit to DPSS Laser Pulse-Amplitude Histogram
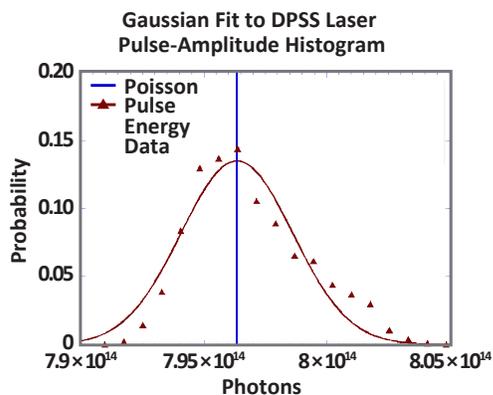
Figure 19: Photon number histogram for ~100 µJ pulses from an Allegro 1535 nm Er:glass, passively Q-switched, diode-pumped solid-state laser as measured with a thermoelectric sensor.

Because the noise analysis presented in this technical note generally proceeds from the assumption that the input opti-cal signal is Poisson-distributed, if this is not the case in an actual laboratory measurement, these equations will under-estimate the total measured noise. When that happens, a photoreceiver's accurate measurement of the noise on an optical signal can be misinterpreted as poor photoreceiver sensitivity, when in fact, a significant component of the total noise does not originate in the photoreceiver at all.

However, despite the large variation in laser pulse energy, a sensitive optical receiver is seldom used to measure a 100 µJ pulse. Typically, optical signals reach photoreceiv-ers after significant attenuation. Various processes like atmospheric scintillation can affect signal amplitude sta-tistics at the receiver, but the Burgess variance theorem of Equation 84 can be used to understand how simple attenuation affects pre-existing noise on an optical signal. Consider the case of the unattenuated 100 µJ laser, for which $\mathrm{var}(b) \approx 5.19 \times 10^{24}$ and $\langle b \rangle \approx 7.96 \times 10^{14}$. Consider two different cases: attenuation to 100 photons (for which $T = 100$ and $\langle b \rangle \approx 1.26 \times 10^{-13}$); and attenuation to 100,000 photons (for which $T = 1 \times 10^5$ and $\langle b \rangle \approx 1.26 \times 10^{-10}$). In the case of attenuation to 100 photons, Equation 84 gives a variance that is essentially indistinguishable from that of a Poisson-distributed signal (variance of about 100 photons equal to the mean). This is because attenuation by a very large factor cuts the magnitude of the noise that is originally on the opti-cal signal until it is negligible, but the attenuation process itself adds its own binomially distributed noise. In the limit of a large number of input photons and a low transmission probability, the binomial distribution converges with the Poisson distribution. On the other hand, attenuation to 100,000 photons results in almost twice the variance of a Poisson-distributed signal. Therefore, when attempting to calculate the sensitivity of a photoreceiver, it is important to know how much noise is originally on the input optical signal, and to consider how much it may be attenuated.

Revision History

Number	Date	Description	Responsibility
–	December 13, 2021	Initial release.	A. Huntington

MCO-0000xxx
P0181