Sensitivity Analysis of APD Photoreceivers
Andrew Huntington, Ph.D.
September 2016

Contents
Introduction .................................................................................................................................................. 2
InGaAs APD Structure ............................................................................................................................... 2
Avalanche Gain and Gain Distribution ...................................................................................................... 3
Exceptions to Standard APD Noise Theory ............................................................................................... 6
Analog APD Photoreceivers .......................................................................................................................... 7
Mean (Signal) ................................................................................................................................................ 7
    RTIA Case ................................................................................................................................................ 8
    Gain-Bandwidth Effects Limiting Signal Response .................................................................................... 8
    CTIA Case ................................................................................................................................................. 11
Variance (Noise) .......................................................................................................................................... 11
    RTIA Case For Conventional InGaAs APDs .............................................................................................. 11
    RTIA Case for Multi-Stage Siletz APDs .................................................................................................... 12
    CTIA Case for Conventional InGaAs APDs ............................................................................................... 14
    CTIA Case for Multi-Stage Siletz APDs ..................................................................................................... 15
Sensitivity Metrics Derived from Mean and Variance ................................................................................ 16
    Signal-to-Noise Ratio (SNR) ..................................................................................................................... 16
    Noise-Equivalent Power (NEP) ................................................................................................................ 18
    Noise-Equivalent Input (NEI) ................................................................................................................... 20
    Relationship Between NEP and NEI ........................................................................................................ 22
Photoreceiver Output Distribution ............................................................................................................. 22
    TIA Input Noise Distribution .................................................................................................................. 23
    APD Output Distribution ........................................................................................................................ 24
    Convolution of APD and TIA Distributions .............................................................................................. 26
Sensitivity Metrics Derived from Output Distribution ................................................................................ 27
    False Alarm Rate (FAR) ............................................................................................................................ 28
    Bit Error Rate (BER) ................................................................................................................................. 31
    Receiver Operating Characteristic (ROC) ................................................................................................. 34
Parameterization of Terminal Dark Current for Voxtel APDs ..................................................................... 35
Burgess Variance Theorem for Multiplication & Attenuation .................................................................... 35
    Derivation ............................................................................................................................................... 36
    Application to Attenuation of a Noisy Optical Signal .............................................................................. 38
References .................................................................................................................................................. 39
**Introduction**

APDs are photodetectors that can be regarded as the semiconductor analog of photomultiplier tubes (PMTs). One important difference is that APDs don’t have a photocathode that is physically separate from their current gain medium, and so they typically use primary photocarriers more efficiently than PMTs. For the same reason, the quantum efficiency of an APD does not degrade over the lifetime of the detector. A second difference is that the multiplication process in an APD is normally bi-directional, so it has different statistics 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 laser range-finding which benefit from the APD’s internal photocurrent gain, fast response, compact size, durability, and low cost. A linear-mode APD’s gain 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 manufacturing techniques used to fabricate APDs differ depending upon the semiconductor alloys used, as does the device structure. This technical note is primarily concerned with short-wavelength infrared (SWIR)-sensitive APDs with InGaAs absorbers; among other common types of APD, silicon APDs sensitive to visible light and HgCdTe APDs sensitive in the mid- and long-wavelength infrared (MWIR / LWIR) are structurally dissimilar. Two common InGaAs APD configurations are sketched in Figure 1 and Figure 2: 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). Both styles of APD employ the separate absorption, charge, and multiplication (SACM) layer design which 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 a mesa APD’s junction 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 sidewalls 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’s gain curve as a function of reverse bias limits the gain at which it can be used. The slope of the gain curve is an issue because mean avalanche gain ($M$) increases asymptotically in the vicinity of the APD’s breakdown voltage ($V_{br}$) according to the empirical relation

$$M = \left(1 - \frac{V}{V_{br}}\right)^{-n},$$  \hspace{1cm} (1)

which holds for all APDs in which both carrier types (electrons and holes) can initiate impact ionization.$^1$ In Eq. (1), the parameter $n$ controls how quickly the avalanche gain rises as $V$ approaches its vertical asymptote at $V_{br}$; stable operation of APDs characterized by large values of $n$ becomes impractical at high gains because $V/V_{br}$ cannot be adequately controlled.

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 employing 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 increasing 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, degrading the signal-to-noise ratio (SNR).

Excess multiplication noise results from the stochastic nature of the impact ionization process that amplifies the APD’s primary current. After avalanche multiplication, each primary carrier injected into an APD’s 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:\cite{2}

$$P_{\text{McIntyre}}(n) = \frac{a \Gamma\left[\frac{n}{1-k} + 1\right]}{n (n-a)! \times \Gamma\left[\frac{n}{1-k} + 1 + a\right]} \times \left[1 + \frac{k(M-1)}{M}\right]^{\frac{n-k}{1-k}} \times \left[\frac{(1-k)(M-1)}{M}\right]^{n-a},$$  \hspace{1cm} (2)

where $k$ is the ratio of hole-to-electron impact ionization rates, $M$ is the average gain, and $\Gamma$ 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). 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. The Burgess variance theorem\textsuperscript{3,4} 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 \( m \):\textsuperscript{5}

\[
\text{var}(n) = M^2 \text{var}(a) + \langle a \rangle \text{var}(m)
\]

\[
= M^2 F \langle a \rangle [\text{e}^{-2}]
\]

where the excess noise factor \( F \) is defined as:

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

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 was a constant \( m=M \) rather than a random variable, \( \text{var}(M \times a) = M^2 \text{var}(a) = M^2 \langle a \rangle \), which is smaller than Eq. (3) by a factor of \( F \).\textsuperscript{†}

For most linear-mode APDs, the excess noise factor has the gain-dependence derived by McIntyre for thick, uniform junctions:\textsuperscript{6}

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

Eq. (3 & 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 Eq. (5), the parameter \( k \) is the same ratio of hole-to-electron impact ionization rates appearing in Eq. (2). When \( k>0 \), it 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 Gaussian approximation doesn’t hold very well far from the mean, so the full McIntyre distribution has to be used to realistically model things like false alarm rate which are sensitive to the tails of the output distribution.

\textsuperscript{†} It is important to note that whereas \( n=M \times a \) in the idealized case of constant gain, \( n \neq m \times a \). 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 \( a \). See the section Burgess Variance Theorem for Multiplication & Attenuation for more details.
the multiplying junction. The gain curve of a $k=0$ APD does not exhibit the vertical asymptote described by Eq. (1), enabling stable operation at higher gain than a $k>0$ APD.

McIntyre distributions for APDs operating at the same average gain ($M=20$) and illuminated by the same signal strength ($\alpha=10$ primary photoelectrons) but differing in $k$ are plotted in Figure 5. 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;
- 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 but comfortably lower than the bulk of the photocurrent distribution’s probability density, such that the longer tail of the high-$k$ distribution increases the probability of false alarm but reduces signal detection probability by decreasing the distribution’s median output value. The detection threshold is employed to reject false alarms arising from circuit noise, of which the APD’s 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 one to set the detection threshold higher without sacrificing 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 remarked above, 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 but 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, but 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 below:

<table>
<thead>
<tr>
<th>$k$</th>
<th>Median Output (Higher is Better for Signal Detection)</th>
<th>Std. Dev.</th>
<th>Chance of Output &gt;1000 e- (Lower is Better for Avoiding False Alarms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>195 e-</td>
<td>88.3 e-</td>
<td>5.24E-13</td>
</tr>
</tbody>
</table>
When reviewing Table 1, it’s 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 4 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 great for calculating quantities like false alarm rate which are sensitive to the tail of the output distribution.*

## EXCEPTIONS TO STANDARD APD NOISE THEORY

Eq. (2 & 5) were derived under the assumption that carriers are always “active” – i.e. 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 a carrier’s path length through the gain medium, so Eq. (5) holds very well. However, important exceptions to the excess noise factor formula of Eq. (5) include APDs in which the carrier dead-space is a significant portion of the width of the multiplying junction, those in which a change in alloy composition modulates the impact ionization threshold energy and rate across the multiplying junction, 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 ionization events. 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). Some APDs, like those fabricated from HgCdTe alloys with cutoff wavelength in the mid- or long-wavelength infrared (MWIR/LWIR) don’t 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 Eq. (5) in the limit of an infinite number of possible ionizations per transit and converges to Lukaszek’s expression for \(F\) when a single ionization per transit is possible.25

---

* This discussion is taken further in later sections. Technically, photoreceiver performance depends upon the convolution of the APD’s 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.

---

<table>
<thead>
<tr>
<th>0.1</th>
<th>179 e-</th>
<th>122.6 e-</th>
<th>6.61E-5</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2</td>
<td>165 e-</td>
<td>149.1 e-</td>
<td>1.08E-3</td>
</tr>
<tr>
<td>0.3</td>
<td>154 e-</td>
<td>171.6 e-</td>
<td>3.45e-3</td>
</tr>
<tr>
<td>0.4</td>
<td>145 e-</td>
<td>191.5 e-</td>
<td>6.55e-3</td>
</tr>
</tbody>
</table>

---

** Figure 6: A typical analog APD photoreceiver.**
Analogue APD Photoreceivers

A block diagram of a typical analogue APD photoreceiver is shown in Figure 6. Depending upon the application, analogue output from the photoreceiver might be sampled by a fast analogue-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 off the APD’s dark current, but in essence, the photoreceiver is just an APD and an amplifier.

Both capacitive (C) and resistive (R) elements are drawn in the TIA’s feedback path, but in practice, TIAs are designed so that one or the other will dominate the amplifier’s gain. If the feedback R and C are both large, the resistive element dominates and the TIA’s output voltage will be proportional to the instantaneous input current; its gain will be characterized by a transimpedance measured in \( \Omega \). 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’s output voltage is proportional to the total charge delivered within a rolling integration period \( \tau \). The capacitive-feedback TIA’s (CTIA’s) 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 reset path is what determines the CTIA’s effective integration period.¹

Back-of-the-envelope receiver sensitivity calculations treat the APD’s responsivity and the TIA’s gain as fixed values, but 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 that pulse’s power spectrum. Care is required when attempting to estimate photoreceiver sensitivity to signals with frequency components outside the bandwidth of either APD or TIA, or for pulse shapes other than that 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 will be 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’s information resides in its intensity, as measured by either its mean optical power (RTIA case) or its mean pulse energy (CTIA case).

¹ In practice, the conversion gain spectrum of a continuously-reset CTIA has a complicated shape, amplifying different frequency components by different amounts. Likewise, the reset path may have a complicated bandpass. Thus, the conversion gain and effective signal integration period depend upon signal pulse shape, and cannot be characterized by a single value each. However, conceptually it is helpful to envision the continuously-reset CTIA as integrating all input current within a rolling sample period.
RTIA CASE

The APD of an RTIA-based receiver converts incident optical power in Watts to an output photocurrent in Amps, which the RTIA then converts to a potential in Volts. The APD’s average power conversion factor is called its spectral responsivity, \( R \):

\[
R = M \frac{\lambda}{1.23985} \frac{[\text{W}]}{[\text{A}]} \frac{[\text{A}]}{[\text{W}]} = R = M \frac{\lambda}{1.23985} \frac{[\text{A}]}{[\text{W}]},
\]

where \( QE \) is the APD’s quantum efficiency at a given wavelength \( (\lambda) \). The RTIA’s transimpedance is usually quoted in \( \Omega \).

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 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 \( P_{\text{signal}} \) in Watts is:

\[
I_{\text{signal}} = \frac{P_{\text{signal}}}{R} \frac{[\text{A}]}{[\text{W}]}.
\]

Some textbooks use the root-mean-square (RMS) optical power of an intensity-modulated signal for \( P_{\text{signal}} \) when analyzing optical communications applications, in which case \( I_{\text{signal}} \) represents the RMS photocurrent.\(^{26}\)

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 laser range-finder that times the arrival of a reflected pulse or a telecommunications receiver discriminating 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 APD and TIA as well as by collective low-pass filtering associated with the detector’s capacitance and the TIA’s input capacitance and transimpedance.

The fundamental frequency response of an APD depends upon its junction transit time and the DC gain at which it operates. Current flows continuously at the APD’s terminals from the time a charge carrier is created in its junction until such time as the carrier is swept to either its anode (for holes) or cathode (for electrons).\(^{27,28}\) The APD’s photocurrent cannot keep up with optical signal modulations on time scales shorter than its junction transit time because the carrier population that was generated by one optical power level will still be conducting current when the optical signal has changed to a new level.\(^{29}\)

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 but the photoelectrons drift toward the cathode by way

\( ^{29} \) 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.
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 which generated them, since 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, extending the junction transit time by extending the hole drift path. 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; the bandwidth of the APD is lower for higher DC gain because the gain process itself lasts longer. This effect can be seen in Figure 7, in which the tail of the APD’s impulse response grows relative to its peak at higher DC gain.

The practical upshot of the tradeoff between APD gain and bandwidth is that the responsivity acting upon high-frequency signal components is smaller than the full DC responsivity of an APD, calculated above in Eq. (6). However, fundamental APD response times are generally sub-nanosecond, as in Figure 7, so the APD’s speed normally isn’t 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 range-finding.

Figure 8 illustrates schematically how laser pulse width would interact with an idealized RTIA photoreceiver’s rise time to reduce its response to fast optical pulses. Assuming laser pulses of equal energy but variable width, the photoreceiver’s response will be stronger to shorter optical pulses as long as the laser pulse width remains greater than the photoreceiver’s rise time. That’s 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’s rise time, the receiver’s 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. Conceptually, a very rough estimate of an RTIA-based photoreceiver’s decreased response to fast pulses can be made by multiplying its responsivity by a correction factor based on the photoreceiver’s bandwidth ($BW$) and the pulse width ($\tau$):

$$R_{\text{reduced}} = R\left[1 - \exp\left(-2\pi BW \tau\right)\right] \quad \text{[A/W]}.$$  (8)

However, in low duty cycle applications, real world RTIA-based photoreceivers often perform much better with short signal pulses than is implied by Eq. (8) and its associated reasoning. In practice, when an RTIA can’t 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’s feedback resistor until the input’s potential has been restored to its normal operating point. Since current actually flows in the RTIA’s 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 Eq. (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 can’t be used in high duty cycle applications like optical communications or multi-hit LADAR because the slow rise and fall times merge consecutive symbols (pulses). However, in a comparatively low duty cycle application like laser range-finding, 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 from the standpoint of 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 10, 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 Voxel’s model VX-809 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, 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’s 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.

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

---

* Refer to the later sections on Noise-Equivalent Power (NEP) and Noise-Equivalent Input (NEI) for an explanation of the sensitivity measures used here. 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 receivers’ 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).
CTIA’s 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 APD’s average energy conversion factor is:

\[
R_{\text{charge}} = 5.03411 \times 10^{18} M QE \lambda \left[ \frac{\text{e}^-}{\text{J}} \right]. \tag{9}
\]

The CTIA’s conversion gain is a reciprocal capacitance usually expressed in V/e-.

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

\[
Q_{\text{signal}} = E_{\text{signal}} R_{\text{charge}} \left[ \text{e}^- \right]. \tag{10}
\]

**Variance (Noise)**

By convention, photoreceiver noise is almost always analyzed at the node between APD and TIA. Since this node is at the input of the TIA, it is necessary to refer the TIA’s output voltage noise to its input by application of its transimpedance (RTIA case) or conversion gain (CTIA case). In general, both the output voltage noise spectral intensity and the TIA’s 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’s input-referred noise spectrum as flat across its 3 dB bandwidth; an input noise spectral density in pA/rt-Hz or an RMS input noise in nA is often specified by a TIA’s manufacturer. Similarly, the input noise of a continuously-reset CTIA is often characterized by a value in RMS electrons. However, it is well to remember that scaling a certain number of RMS Volts at the CTIA’s output to a certain number of RMS electrons at its 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 APD and TIA can be viewed as a random variable equal to the sum of random variables representing the APD’s output and the TIA’s input-referred noise. The APD’s 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 APD and TIA is analyzed in terms of spectral intensities. If the RTIA’s input-referred noise is expressed as a spectral density in pA/rt-Hz, the corresponding spectral intensity \(S_{TIA}\) is just the square of the spectral density. Alternatively, if only an RMS input noise current is specified, \(S_{TIA}\) 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 us to recast Eq. (3) as a noise spectral intensity theorem:

\[
S_I = 2 q M^2 F I_{\text{primary}} \left[ \frac{A^2}{\text{Hz}} \right], \tag{11}
\]
where \( q = 1.602 \times 10^{-19} \text{ C} \) is the elementary charge in Coulombs and \( I_{\text{primary}} \) is the primary (i.e. unmultiplied) current in Amps. Technically, Eq. (11) only applies in the low-frequency limit, but it is usual practice to take the APD’s multiplied shot noise spectrum as being approximately flat across its bandwidth.

The spectral intensity of the current noise at the node between APD and TIA is the sum of the individual spectral intensities of the RTIA’s input noise (\( S_{I_{\text{RTIA}}} \)), the shot noise on the APD’s dark current (\( S_{I_{\text{dark}}} \)), and the shot noise on the APD’s photocurrent (\( S_{I_{\text{signal}}} \) and \( S_{I_{\text{background}}} \)):

\[
S_{I_{\text{total}}} = S_{I_{\text{RTIA}}} + S_{I_{\text{dark}}} + S_{I_{\text{background}}} + S_{I_{\text{signal}}} \quad \text{[A}^2/\text{Hz].} \tag{12}
\]

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, Eq. (11) can be used for the shot noise spectral intensity, with \( I_{\text{primary}} \) broken into different current components:

\[
S_{I_{\text{total}}} = S_{I_{\text{RTIA}}} + 2 q M F(I_{\text{dark}} + I_{\text{background}} + I_{\text{signal}}) \quad \text{[A}^2/\text{Hz],} \tag{13}
\]

where \( I_{\text{dark}} \) is the dark current in Amps measured across the APD’s terminals, \( I_{\text{signal}} \) is the signal photocurrent given by Eq. (7), and the background photocurrent \( I_{\text{background}} \) is:

\[
I_{\text{background}} = A \sum_{n} [\Delta \lambda_n I_B(\lambda_n) R(\lambda_n)] \quad \text{[A].} \tag{14}
\]

In Eq. (14) for the background photocurrent, \( A \) is the area in m\(^2\) of the receiver’s optical aperture, \( \Delta \lambda_n \) is the width in nm of wavelength bin \( n \) of a background spectral irradiance dataset, \( I_B(\lambda) \) is the background spectral irradiance in W m\(^{-2}\) nm\(^{-1}\) in bin \( n \), and \( R(\lambda_n) \) is the APD’s spectral responsivity near the center wavelength of bin \( n \) given by Eq. (6). Note that in Eq. (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’s dark current.

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

\[
I_{\text{noise}}^2 = BW S_{I_{\text{total}}} \quad \text{[A}^2]. \tag{15}
\]

The standard deviation of the current, \( I_{\text{noise}} \), is commonly referred to as the photoreceiver’s “noise current”.

### RTIA Case for Multi-Stage Siletz APDs

Eq. (13) assumes that all the primary current is generated outside the APD’s 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 Vtxtel’s Siletz model APD, noise on dark current must be treated separately from noise on photocurrent.

Internally, the Siletz APD’s multiplier is divided into seven cascaded multiplying stages, and unlike most InGaAs APDs, the majority of the Siletz APD’s 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

---

\* Note that Eq. (11) is written in terms of an unmultiplied primary current and includes a factor of \( M^2 \) whereas multiplied terminal currents appear in Eq. (13), so the order of \( M \) has been reduced by one.
APD depends upon how much of the APD’s multiplier it traverses. Dark current generated near the side of the multiplier adjacent to the diode’s absorber 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:

\[ M_s = \sqrt[\text{stages}]{M} \]  

(16)

where \( \text{stages} \) is the number of multiplying stages (Voxtel’s Siletz APD has 7) 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’s 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:

\[ M_{\text{dark}} = \frac{1}{\text{stages}} \sum_{i=1}^{\text{stages}} M^{i-1}_s. \]  

(17)

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’s terminals by the net gain and the number of stages:

\[ I_{dp} = \frac{I_{\text{dark}}}{\text{stages} M_{\text{dark}}}. \]  

(18)

where \( I_{\text{dark}} \) is the multiplied dark current measured at the APD’s terminals. Terminal dark current parameterizations for Voxtel APDs are given in a later section.

Finally, an expression similar to Eq. (11) is summed over all the multiplying stages to find the noise spectral intensity of the Siletz APD’s dark current:

\[ S_{I_{\text{dark}}} = 2 q I_{dp} \sum_{i=1}^{\text{stages}} \left( M^{i-1}_s \right)^2 F(M = M^{i-1}_s) [\text{A}^2/\text{Hz}], \]  

(19)

where the notation \( F(M = M^{i-1}_s) \) means the excess noise factor of Eq. (5) calculated with \( M^{i-1}_s \) substituted in place of the average avalanche gain measured for the photocurrent. When making calculations for a photoreceiver that uses the Siletz model APD, Eq. (19) can be substituted into Eq. (13) to obtain:

\[ S_{I_{\text{total}}} = S_{I_{\text{TIA}}} + 2 q I_{dp} \sum_{i=1}^{\text{stages}} \left( M^{i-1}_s \right)^2 F(M = M^{i-1}_s) + 2 q M F(I_{\text{background}} + I_{\text{signal}}) [\text{A}^2/\text{Hz}]. \]  

(20)

Eq. (16-20) were derived to improve correspondence between theory and measurement for photoreceivers assembled from RTIAs and Voxtel’s Siletz APDs. Table 2 compares noise-equivalent...
power (NEP) measurements to values calculated using either Eq. (13) or Eq. (20) for a 200-MHz photoreceiver built from a COTS RTIA and a Siletz APD.

### Table 2: NEP measurements compared to models for a 200-MHz RTIA/Siletz APD Photoreceiver.

<table>
<thead>
<tr>
<th>Gain</th>
<th>Measured</th>
<th>Simple Model Eq. (13)</th>
<th>Distributed Model Eq. (20)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>4.4 nW</td>
<td>8.1 nW</td>
<td>6.1 nW</td>
</tr>
<tr>
<td>20</td>
<td>3.7 nW</td>
<td>8.0 nW</td>
<td>5.1 nW</td>
</tr>
<tr>
<td>30</td>
<td>4.3 nW</td>
<td>8.4 nW</td>
<td>5.0 nW</td>
</tr>
<tr>
<td>39</td>
<td>4.5 nW</td>
<td>8.8 nW</td>
<td>5.1 nW</td>
</tr>
<tr>
<td>46.5</td>
<td>4.6 nW</td>
<td>9.2 nW</td>
<td>5.1 nW</td>
</tr>
</tbody>
</table>

As can be seen, the photoreceiver’s measured performance is better than predicted by either model, but the distributed dark current model of Eq. (20) is more accurate than the conventional model of Eq. (13). Also, the accuracy of Eq. (20) improves for \( M > 30 \), which is the typical operating point of Siletz APD receivers.

When using Eq. (16-20) to model Siletz APD photoreceivers it should be kept in mind that these equations are strictly valid only for the ideal \( k=0 \) case in which only electrons can trigger impact ionization. Impact ionization always generates equal numbers of secondary holes and electrons, but for a \( k=0 \) multiplier, only the secondary electrons cause additional impact ionizations. The actual Siletz APD is characterized by \( k=0.02 \), so the model does not treat all of its avalanche physics. The dominance of impact ionization by electrons is implicit in the model because Eq. (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 of the 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 Eq. (19) treats the dark current shot noise spectral intensity of a single \( s \)-stage Siletz multiplier as the sum of the spectral intensities of \( s \) different multipliers with stages numbering between 0 and \( s-1 \). 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 of Eq. (16-20) is 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 APD and TIA is calculated by application of Eq. (3) for the variance of an APD’s multiplied output:

---

* The reason the number of stages ranges between 0 and \( s-1 \) as opposed to 1 and \( s \) is that dark current carriers generated in the high-field region of a given multiplier stage don’t have sufficient kinetic energy to impact-ionize in that stage; they only become active in the next stage.
\[ N_Q^2 = N_{CTIA}^2 + N_{dark}^2 + \left( \langle a_{signal} \rangle + \langle a_{background} \rangle \right) M^2 F \quad \text{[e}^{-2}\text{]}, \]  

(21)

where \( N_{CTIA} \) and \( N_{dark} \) are respectively the standard deviations of the CTIA’s input-referred noise and the number of dark current electrons output during the CTIA’s effective integration period \( \tau \); similarly \( a_{signal} \) and \( a_{background} \) are respectively the number of primary photocurrent electrons generated by signal and background optical power received during \( \tau \). \( F \) is the excess noise factor calculated from Eq. (5).

The input-referred noise of the CTIA, \( N_{CTIA} \), is a characteristic of the CTIA and the laser pulse shape. If \( N_{CTIA} \) isn’t specified by a manufacturer, it can be calculated from a circuit simulation of the CTIA in which the APD’s capacitive load on the CTIA input and its mean dark current are modeled, but the shot noise on the APD’s current is omitted. Alternatively, CTIA conversion gain can be measured using a photoreceiver in which the detector’s 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 CTIA’s charge-to-voltage conversion gain.

The noise on the multiplied dark current, \( N_{dark} \), depends upon 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 photocarriers in Eq. (21):

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

(22)

where \( Q_{signal} \) is given by Eq. (10).

Eq. (21-22) 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 Eq. (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 Eq. (2). On the other hand, if a noisy optical signal is attenuated by a large factor, a Poisson distribution is recovered – see the section Burgess Variance Theorem for Multiplication & Attenuation for more details.

**CTIA CASE FOR MULTI-STAGE SILETZ APDS**

The treatment of the dark current shot noise of a Siletz APD in a CTIA-based receiver is closely analogous to that described earlier for the RTIA case. Eq. (16-18) concerning the gain and primary dark current per multiplying stage apply. An expression similar to Eq. (3) is summed over all the multiplying stages to find the variance of the Siletz APD’s dark current:

\[
N_{dark}^2 = \frac{\tau}{q} I_{dp} \sum_{i=1}^{\text{stage}} \left[ (M^{i+1})^2 F(M = M^{i-1}) \right] \quad \text{[e}^{-2}\text{]},
\]

(23)

When making calculations for a photoreceiver that uses the Siletz model APD, Eq. (23) can be substituted into Eq. (22) to obtain:

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

(24)
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 such as optical sensitivity at a given false alarm rate (FAR) or bit error rate (BER) apply. With a decision circuit, one can also analyze the probabilities of true and false positives and negatives to characterize the probabilities of signal detection (PD) and false alarm (PFA), preparing a parametric plot over detection threshold of PD versus PFA called a receiver operating characteristic (ROC).

SNR, NEP, and NEI are all ways of expressing the standard deviation of a photoreceiver’s 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’s input.

It is common to calculate FAR, BER, PD, and PFA based on the mean and standard deviation of the photoreceiver’s output by assuming it is Gaussian-distributed. However, as was shown in the introduction (Figure 3), the high-output tail of an APD’s McIntyre distribution diverges substantially from its Gaussian approximation. When the McIntyre-distributed APD output is convolved with the Gaussian-distributed noise of the TIA the convolution retains some of the McIntyre distribution’s positive skew. Consequently, when the Gaussian approximation is used, it underestimates FAR, BER, PD and PFA. For this reason, sensitivity metrics which depend upon the tail of the photoreceiver’s output distribution are discussed in a separate section of this technical note.

**SIGNAL-TO-NOISE RATIO (SNR)**

The form of the SNR depends upon at which node it is defined and by which convention. Physicists tend to focus on the optical signal measured either in Watts of power in the RTIA case or Joules of energy\(^*\) 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 a photoreceiver’s output voltage 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, the SNR is the mean output signal voltage divided by its standard deviation because these quantities have a linear relationship to the mean optical power level impinging upon the receiver and the equivalent standard deviation found by referring the current and voltage noise sources from the APD and TIA to the receiver’s input. However, one sometimes encounters an SNR 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’s 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 powers 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, we follow the optical-signal-oriented convention, and define SNR in terms of the mean and standard deviation rather than their respective squares. For convenience, the SNR expression is evaluated at the node between APD and TIA:

In the RTIA case:

\(^*\) Or, equivalently, photon number.
\[ SNR = \frac{I_{signal}}{I_{noise}} = \frac{P_{signal} R}{\sqrt{BW S_{total}}}. \]  

(25)

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

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

(26)

If a similar RTIA photoreceiver is assembled from a Siletz APD, the \( SNR \) is:

\[ SNR \approx \frac{P_{signal} R}{\sqrt{BW \left[ S_{TIA} + 2 q I_{dp} \sum_{i=1}^{\text{stages}} \left(M_{s}^{i-1}\right)^{2} F (M = M_{s}^{i-1}) \right] + 2 q M F (I_{background} + P_{signal} R)}}. \]

(27)

The \( SNR \) of a CTIA photoreceiver is:

\[ SNR = \frac{Q_{signal}}{N_{Q}} = \frac{E_{signal} R_{charge}}{\sqrt{N_{CTIA}^{2} + \left(\langle a_{signal}\rangle + \langle a_{background}\rangle\right) M^{2} F}}. \]

(28)

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

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

(29)

With a Siletz APD, the \( SNR \) of a CTIA photoreceiver is:

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

(30)

Example \( SNRs \) calculated using Eq. (26) are plotted versus avalanche gain in Figure 10. The hypothetical photoreceiver is assembled from a 200-μm-diameter Deschutes APD, and uses either a Maxim MAX3658 or MAX3277 RTIA; the photoreceiver is band-limited to 200 MHz. Curves are plotted comparing \( SNR \) for optical signal power levels of 10, 100, and 1000 nW at 1550 nm (left), comparing \( SNR \) for effective ionization rate ratios of \( k = 0, 0.2 \) and 0.4 (center), and comparing \( SNR \) for receivers assembled from the MAX3277 TIA versus the MAX3658 (right); the default conditions were \( P_{signal} = 100 \) nW, \( k = 0.2 \), and \( S_{I,TIA} = 4.4E-24 \text{ A}^{2}/\text{Hz} \) (the MAX3658 TIA). Negligible background illumination was assumed. Notice that the optimal gain which maximizes the photoreceiver’s \( SNR \) varies for all these situations. Increasing either the optical signal power or the effective ionization rate ratio increases the APD’s noise contribution, shifting the optimal operating point to lower gain. Increasing the TIA’s noise contribution shifts the APD’s optimal operating point to higher gain. Although not shown, a strong

---

* Refer to the section RTIA Case for Multi-Stage Siletz APDs for a discussion of the approximations inherent in the denominator of Eq. (27 & 30).
background or higher dark current would shift the optimal operating point to lower gain, as either would increase the APD’s noise contribution relative to fixed TIA noise.

**NOISE-EQUIVALENT POWER (NEP)**

There are two ways to define and use \( NEP \) – with and 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?” Since signal shot noise increases with signal strength, \( NEP \) cannot be used directly to calculate \( SNR \) at higher signal powers. However, \( NEP \) is useful as a minimum sensitivity benchmark. We will use the symbol \( NEP_{SNR=1} \) for this definition.

In contrast, when \( NEP \) is defined without signal shot noise, it emphasizes the photoreceiver’s 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’s noise floor in units that are convenient to compare to the optical signal level characteristic of a given application. For instance, one may design a laser range-finding system 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, one needs to know how the reflected signal strength compares to the detection threshold in order to compute the pulse detection probability, \( P_D \). \( NEP \) is often used in situations like this to quantify the photoreceiver’s 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, since 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 \( P_{FA} \).

To find the optical signal power level at which \( SNR \) equals unity, equate the numerator to the denominator in Eq. (26 or 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 Eq. (26), the \( NEP_{SNR=1} \) is:

\[
\text{SNR vs. } M \text{ curves calculated for a photoreceiver assembled from a 200-μm Deschutes APD and a COTS TIA, demonstrating how optimal gain depends upon signal power (} P_{signal} \text{), the APD’s effective ionization rate ratio (} k \text{), and the TIA’s input-referred noise spectral intensity (} S_{ITIA} \text{).}
\]
\[ NEP_{\text{SNR=1}} = \frac{q M F BW + \sqrt{(q M F BW)^2 + BW (S_{\text{I\!\!\!TIA}} + 2q M F (I_{\text{dark}} + I_{\text{background}}))}}{R} \]  

The Siletz APD case corresponding to Eq. (27) is:

\[ NEP_{\text{SNR=1}} \approx \frac{q M F BW + \sqrt{(q M F BW)^2 + BW \left(S_{\text{I\!\!\!TIA}} + 2q \sum_{\text{stages}} (M_{i-1})^2 F(M = M_{i-1}) + M F I_{\text{background}}\right)}}{R} \]  

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

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

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

\[ NEP = \sqrt{BW \left(S_{\text{I\!\!\!TIA}} + 2q M F (I_{\text{dark}} + I_{\text{background}})\right)} \]  

The current variance of a photoreceiver based on a multi-stage Siletz APD is given by Eq. (15 & 20); its \( NEP \) without shot noise on a hypothetical noise-equivalent signal is:

\[ NEP \approx \frac{\sqrt{BW \left(S_{\text{I\!\!\!TIA}} + 2q I_{\text{dp}} \sum_{\text{stages}} (M_{i-1})^2 F(M = M_{i-1}) + 2q M F I_{\text{background}}\right)}}{R} \]  

The difference in definition between \( NEP_{\text{SNR=1}} \) given by Eq. (31 & 32) and \( NEP \) given by (Eq. 34 & 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 Eq. (31 and 32), but this factor is usually dominated by the terms representing the

Figure 11: \( NEP \) vs. \( M \) curves calculated for a photoreceiver assembled from a 200-μm Deschutes 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.
TIA’s noise contribution \((BW \times S_{I_{TTA}})\) and/or the noise on the dark current – \(2 q M F I_{\text{dark}} BW\) in Eq. (31) or the equivalent summation over multiplier stages in Eq. (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, the center panel of Figure 11 shows calculations of \(NEP\) made for a hypothetical photoreceiver in which a 200-μm Deschutes 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 Maxim MAX3658. In Figure 11, the dashed \(NEP\) curves were calculated using Eq. (31) for the case that includes shot noise on the noise-equivalent signal, and the solid curves were calculated using Eq. (34), which omits signal shot noise. The left and right panels of Figure 11 both assume room-temperature operation of the receiver and normal COTS TIAs. The “strong background” mentioned in the left panel of Figure 11 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, but in that case the curves all overlay each other, following the red curve of the right-hand panel, since the TIA’s noise completely dominates.

Eq. (31, 32, 34 & 35) can be converted to spectral densities in W/rt-Hz by omitting the factor of \(BW\) inside the radical.

The left and center panels of Figure 11 show cases in which the optimal gain operating point that minimizes \(NEP\) is less than the maximum gain. 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 Eq. (5 & 6); the terminal dark current \(I_{\text{dark}}\) is at least order 1 (refer to the later section Parameterization of the Terminal Dark Current of Voxtel APDs for details). Consequently, once the APD’s noise term becomes larger than the TIA’s 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 upon the high-output tail of the distribution. However, because the noise distribution of an APD photoreceiver is skewed, with higher probability density at high output than the Gaussian distribution with the same mean and variance, 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 done in a later section.

### NOISE-EQUIVALENT INPUT (NEI)

The acronym NEI is used by the imaging community for a different purpose than our meaning here. When discussing passive imagers, NEI means \(noise-equivalent irradiance\) and is just a way of expressing the \(NEP\) as a spectral irradiance in W m\(^{-2}\) nm\(^{-1}\). However, we use the acronym \(NEI\) 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, \(NEI_{\text{SNR}=1}\), is the signal level for which the photoreceiver’s \(SNR\) is unity. The second definition is the signal level for which the photoreceiver’s average output will be equal in magnitude to its RMS noise in the absence of an optical signal.

---

**VoxTELOPTO**

A Division of Voxtel Inc.  15985 NW Schendel Ave. Ste.200  Beaverton, OR 97006  Phone: 971-223-5846  Fax: 503-296-2962  www.voxtel-inc.com
To find $\text{NEI}_{\text{SNR}=1}$ for a photoreceiver that is assembled from a conventional InGaAs APD and a CTIA, equate numerator and denominator of Eq. (29) and solve for $E_{\text{signal}}$; convert the result to photons by multiplying the energy in Joules by 5.034117E18 $\lambda$ [photons J$^{-1}$ μm$^{-1}$]:

$$\text{NEI}_{\text{SNR}=1} = \frac{M F + \sqrt{(M F)^2 + 4 \left( N_{\text{CTIA}}^2 + M F \frac{\tau}{q} (I_{\text{dark}} + I_{\text{background}}) \right)}}{2 M \text{QE}} \quad \text{[photons].}$$  \hspace{1cm} (36)

In Eq. (36), the definition of $R_{\text{charge}}$ from Eq. (9) was applied to eliminate the wavelength.

The case for a CTIA receiver that uses a Siletz APD is found by solving Eq. (30) for $E_{\text{signal}}$ with $\text{SNR}=1$:

$$\text{NEI}_{\text{SNR}=1} = \frac{M F + \sqrt{(M F)^2 + 4 \left( N_{\text{CTIA}}^2 + \frac{\tau}{Q} \sum_{i=1}^{\text{stages}} (M_i)^2 F(M = M_i^{-1}) + I_{\text{background}} \ M F \right)}}{2 M \text{QE}} \quad \text{[photons].}$$  \hspace{1cm} (37)

Neglecting the shot noise on the hypothetical noise-equivalent signal, the $\text{NEI}$ of a CTIA-based photoreceiver is:

$$\text{NEI} = 5.0341118 \lambda \mu m \frac{N_q(E_{\text{signal}} = 0)}{R_{\text{charge}}} = \frac{N_q(E_{\text{signal}} = 0)}{M \text{QE}}. \quad \text{(38)}$$

Using Eq. (22) for the charge noise of CTIA photoreceiver assembled from a conventional InGaAs APD, the $\text{NEI}$ is:

$$\text{NEI} = \sqrt{N_{\text{CTIA}}^2 + \frac{\tau}{q} (I_{\text{dark}} + I_{\text{background}}) \ M F} \quad \text{[photons].} \hspace{1cm} (39)$$

Using Eq. (24) for the Siletz case, the $\text{NEI}$ of a CTIA photoreceiver, neglecting signal shot noise, is:

$$\text{NEI} = \sqrt{N_{\text{CTIA}}^2 + \frac{\tau}{q} \sum_{i=1}^{\text{stages}} (M_i^{-1})^2 F(M = M_i^{-1}) + \frac{1}{Q} I_{\text{background}} \ \tau + Q_{\text{signal}}} \ M F \quad \text{[photons].} \hspace{1cm} (40)$$

Unlike $\text{NEP}$, the difference in definition between $\text{NEI}_{\text{SNR}=1}$ and $\text{NEI}$ is germane in typical applications. Consider Voxtel’s model VX-806 readout integrated circuit (ROIC), which has an input-referred pixel read noise of about $N_{\text{CTIA}}=64$ e- at 235 K. NEI plots calculated using Eq. (36 & 39) for a 30-μm-diameter conventional InGaAs APD pixel hybridized to one channel of the VX-806 ROIC are plotted in Figure 12. Negligible background illumination and an effective integration time of $\tau=10$ ns were assumed. When signal shot noise is omitted (solid curves), the CTIA noise dominates and there’s no difference in $\text{NEI}$ based on the APD’s impact ionization coefficient ratio ($k$). However, the plots of $\text{NEI}_{\text{SNR}=1}$ are sensitive to $k$ because of the comparatively small noise contributions from APD dark current and CTIA noise.

* Refer to the section \textit{RTIA Case for Multi-Stage Siletz APDs} for a discussion of the approximations inherent in the treatment of Siletz APD dark current shot noise appearing in the numerator of Eq. (37 & 40).
Occasionally one will encounter a CTIA photoreceiver sensitivity specification for which \( \text{NEI} < 1 \) photon. This is only possible using the form of \( \text{NEI} \) which omits the signal shot noise, as \( \text{NEI}_{\text{SNR}=1} \) given by Eq. (36 or 37) cannot assume values less than unity, even with a completely noiseless TIA and APD. Thus, a published \( \text{NEI} \) that is less than one does not mean that the CTIA photoreceiver can measure 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 one would be better served by calculating \( N_Q \) directly, using the actual value of \( Q_{\text{signal}} \)).

**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 that receiver’s “native” units: an optical power for an RTIA photoreceiver or the photon number of an optical pulse for a CTIA photoreceiver. Then, with either \( \text{NEP} \) or \( \text{NEI} \) in hand, specify the photon number which would result in a peak instantaneous optical power equal to the calculated \( \text{NEP} \) or the peak optical power of a pulse having a photon number equal to the calculated \( \text{NEI} \). Notice that whereas both \( \text{NEP} \) and \( \text{NEI} \) 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 \( 10^{-24} \text{ A}^2/\text{Hz} \) and an APD operated at \( M=10 \) with \( F=3.5, QE=80\% \) and 100 nA of dark current; background illumination is negligible. According to Eq. (6, 13 & 34), this receiver’s \( \text{NEP} \) at 1550 nm, without signal shot noise, is about 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 \( \text{NEP} \) is \( 1E-17 \) J, or an “\( \text{NEI} \)” of about 78 photons at 1550 nm. But suppose our point of reference was a 10-ns signal pulse, instead. Since a 5-ns pulse is well within the bandwidth of a 200-MHz receiver, 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 receiver’s “\( \text{NEI} \)” 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’s 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’s noise, so the random variable representing the photoreceiver’s 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’s output, its

\[
\text{Figure 12: } \text{NEI} \text{ vs. } M \text{ curves calculated for a photoreceiver assembled from a 30-μm Deschutes APD pixel and a VX-806 ROIC, comparing two alternate definitions of } \text{NEI} \text{ for different values of } k.
\]
distribution is normally analyzed at the node between APD and TIA, working in units of electrons. This model presents some difficulties of interpretation, since the TIA’s noise is an analog value characterized by the continuous Gaussian distribution of its output voltage, whereas the APD’s charge output is quantized and obeys the discrete McIntyre distribution of Eq. (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 have to 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’s 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’s output distribution, \( n \) represents a quantity of charge in units of electrons:

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

(41)

In the case of a CTIA, the interpretation of the noise-equivalent input electron count \( n \) appearing in Eq. (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’s 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’s 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 upon below, the product of the TIA’s input-referred noise current and an effective integration time \( t_{\text{transit}} \) gives the quantity of charge which, if delivered over the same time span, would produce an APD output current of equal magnitude:

\[
\text{var}(n)_{RTIA} = \frac{t_{\text{transit}}^2 S_{TIA} \text{BW}}{q^2}.
\]

(42)

However, one should bear in mind that to the extent a real RTIA has some signal-integrating character, the effective time span which scales between current and charge in Eq. (42) is generally longer than the physical junction transit time of the APD.

As a practical matter, it is better to find the RTIA’s effective input charge noise empirically for the specific signal pulse shape of interest. Suppose one collected amplitude statistics 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’s 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’s output voltage noise. If one then illuminates the unity-gain photoreceiver with optical signal pulses of calibrated energy, chosen to be well above the receiver’s...
noise floor, the difference in mean output voltage peak height measured between pairs of chosen signal levels divided by the difference in mean signal charge, found from Eq. (9), gives a conversion gain in units of V/e-. With the usual caveat that conversion gain depends upon 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’s charge output when a transient current pulse completes within the effective integration period (τ) of a CTIA. However, the distribution of the APD’s instantaneous current output that is relevant to RTIA photoreceivers is harder to calculate accurately.

In principle, the Shockley-Ramo theorem allows one to calculate the instantaneous current at an APD’s terminals from the instantaneous count of electrons and holes within its junction, \( n_e(t) \) and \( n_h(t) \), and their respective saturation velocities, \( v_{se} \) and \( v_{sh} \):

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

where \( w \) is the junction width. Eq. (43) can be recast in terms of junction transit times for electrons and holes, \( t_e = w/v_{se} \) and \( t_h = w/v_{sh} \):

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

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

For modeling the APD’s 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 \( \alpha \), without regard to the time evolution of either population. As discussed earlier in the section on *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 upon its polarity and where in the junction it was generated. Moreover, the signal photons which generate primary photocarriers do not arrive at the APD simultaneously. A detailed numerical simulation is required to accurately model APD current statistics, but there are some simplifying assumptions which often apply that permit a simpler analysis.

An APD’s impulse response function depends upon its structure and operating point, but 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, Voxel’s InGaAs APDs usually 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 which are generated along with those electrons must drift all the way back through the thick absorber before

___

* Eq. (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 Eq. (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.
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 which 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’s instantaneous current to the McIntyre distribution, justifying the use of Eq. (2) to model the distribution of current peak height. It also associates the hole transit time, $t_{\text{transit}}$, with the unspecified transit time $t_{\text{transit}}$ appearing in Eq. (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 arrived in a pulse lasting somewhat less than the junction transit time, the photocurrent pulse height would be relatively well modeled by Eq. (2) because all the secondaries generated by all the primaries would be simultaneously present in the junction at some point. However, depending upon the APD, the junction transit time is usually sub-nanosecond, whereas most applications involve optical pulses of longer duration. Working out the peak height distribution when some, but not all, of the impact ionization chains overlap in time 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’s 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 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 $i_{\text{primary}}$ will inject $a$ primary carriers into the APD’s multiplier within the junction transit time $t_{\text{transit}}$ is:

$$ P_{\text{Poisson}}(a) = \left( \frac{i_{\text{primary}}}{q} t_{\text{transit}} \right)^a \exp \left( - \frac{i_{\text{primary}}}{q} t_{\text{transit}} \right) \frac{1}{a!}. $$ \hspace{1cm} (45)

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

$$ P_{\text{APD}}(n) = \sum_a P_{\text{Poisson}}(a) \times P_{\text{McIntyre}}(n). $$ \hspace{1cm} (46)

Eq. (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. Eq. (45 & 46) also apply to the photocurrent of a multi-stage Siletz APD, but the dark current distribution requires separate consideration of output from each stage. Eq. (16-18) can be used to estimate the gain-per-stage, $M_s$, and the primary dark current per multiplying stage, $I_{dp}$, of a Siletz APD. Then, for stage $i$, the distribution of the dark current is approximately equal to Eq. (46) where $I_{dp}$ has been substituted for $i_{\text{primary}}$ in Eq. (45) and $P_{\text{McIntyre}}$ is calculated with $M = M_s^{-1}$. Make note of the comments regarding the accuracy of this approximation which appear in the section on Variance (Noise) for the RTIA Case for Multi-Stage Siletz APDs.
When applying Eq. (45 & 46) to steady state current in a CTIA receiver, such as dark current or background photocurrent, the CTIA’s effective integration period $\tau$ is used in place of the junction transit time $t_{\text{transit}}$.

**CONVOLUTION OF APD AND TIA DISTRIBUTIONS**

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

$$P_{RX}(n) = (P_{\text{TIA}} * P_{\text{APD}})(n) \equiv \sum_{i} P_{\text{TIA}}(i) P_{\text{APD}}(n-i).$$  \hspace{1cm} (47)

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 Siletz APD, the distributions of the photocurrent and the dark current generated in each multiplier stage are distinct, so the distribution of the photoreceiver’s output is found by convolving all of the APD-related distributions with the TIA’s distribution:

$$P_{RX} = P_{\text{TIA}} * P_{\text{APD, photo}} * P_{\text{APD, dark}_1} * P_{\text{APD, dark}_2} * \ldots$$  \hspace{1cm} (48)

Example output distributions calculated using Eq. (47) for hypothetical 200-MHz RTIA photoreceivers assembled from 200-μm-diameter Deschutes APDs are plotted in Figure 13. The APD’s steady state dark current was used to calculate the curves in Figure 13, but the results would be equivalent for any combination of primary photocurrent and dark current having the same sum. The primary dark current levels at 27°C were 2.57 nA at $M=5$, 2.72 nA at $M=10$, and 3.17 nA at $M=20$; at -30°C and $M=10$ the primary dark current was 0.16 nA. A junction transit time of $t_{\text{transit}}=1$ ns was assumed for the purpose of calculating the TIA’s noise, resulting in input-referred charge noise levels of 185 e- for the MAX3658 and 603 e- for the MAX3277. Except where specified, the default values used in the calculations were $M=10$, $k=0.2$ and $T=27°C$.

The left-hand panel of Figure 13 compares photoreceiver output distributions (scaled to the node between APD and TIA) for effective impact-ionization rate ratios of $k=0$, 0.2, and 0.4 on semi-logarithmic axes to emphasize that $k$ has a big impact on the high-output tail of the distribution. However, the use of semi-log axes in the left panel of Figure 13 conceals the other important trend in $k$, which is that 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’s 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 the other panels of Figure 13 use linear axes, preventing visual comparison to the left-hand panel. However, the green curves in all three panels correspond to the same baseline set of conditions: \( M=10; k=0.2; T=27^\circ\text{C}; \) MAX3658 TIA. The center panel of Figure 13 shows how the output distribution varies with the APD’s gain, for \( M=5, 10 \) and 20. As would be expected, the median of the output distribution shifts to higher output values as \( M \) increases, but the distribution also broadens and skews to higher output. The right-hand panel of Figure 13 compares the baseline case of a MAX3658 TIA to the noisier MAX3277 model, and also compares operation at 27°C to -30°C to demonstrate the influence of varying APD dark current.

In the introductory section on Avalanche Gain and Gain Distribution, the point was made that the Gaussian approximation of an APD’s output distribution gets the behavior in the high-output tail wrong (Figure 5). Figure 14 revisits that point for the total photoreceiver output, analyzing the same cases as the left-hand panel of Figure 13. The usual approximation of an APD’s output distribution is a Gaussian distribution with a variance calculated using Eq. (3 & 5). Figure 14 compares output distributions calculated by Eq. (47) using proper McIntyre distributions for the APD (solid curves) to Gaussian approximations (dashed curves). 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.

**Sensitivity Metrics Derived from Output Distribution**

Figure 15 illustrates threshold detection by an APD photoreceiver equipped with a binary decision circuit that registers a detection event if the photoreceiver’s output signal exceeds a specified detection threshold. Output pulse height distributions of an analog APD photoreceiver based on the convolution of Eq. (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 + 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 false alarm rate (FAR) or bit error rate (BER), and receiver operating characteristic (ROC).

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, but they don’t 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 upon an APD photoreceiver, and the pulse width is shorter than the effective signal integration time of the receiver’s TIA, then that is one attempt; the CCDF at the detection threshold is equal to $P_D$ provided one assumes that the comparator is ready at the time the signal pulse arrives. However, in the case of a continuous input like dark current or a quasi-CW optical signal that persists longer than the receiver’s 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 waveform rises through the detection threshold, but they don’t trigger again if the waveform stays above the threshold for a span of time. Accurate analysis of the FAR depends upon the probability that the waveform is transitioning through the detection threshold with positive slope, not just the probability it exceeds the detection threshold, given by the CCDF.*

**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”. Rice analyzed a noisy current waveform defined in terms of uncorrelated random variables for its current ($\xi$) and the slope of its current ($\eta$) at every point in time, $t$. 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:‡

* Technically, accurate calculation of $PDE$ also requires considering whether the comparator is armed and ready to register a detection event. This is a very 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.

† Although Rice explicitly analyzed 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.

‡ As will shortly be made explicit, the normalization of $p(\xi=I_{th};\eta; t)$ gives a factor of $A^2$ Hz$^{-1}$ when $\xi$ is in units of A and $\eta$ is in units of A/s; multiplication of $p(\xi=I_{th};\eta; t)$ by $\eta$, 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 which 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.
where \( p(\xi = I_\text{th}, \eta; t) \) is the joint probability distribution of the current and its slope at time \( t \), assuming the random variable for the current has the value \( I_\text{th} \). Rice’s classic result for FAR applies to Gaussian-distributed noise, for which \( p(\xi = I_\text{th}, \eta; t) \) is the bivariate normal distribution. In the case of two uncorrelated random variables, the bivariate normal distribution is just the product of two single-variable Gaussian distributions:

$$ p(\xi, \eta; t)_{\text{Rice}} = \frac{1}{2\pi \sqrt{\text{var}(\xi) \text{var}(\eta)}} \exp \left[ -\frac{1}{2} \left( \frac{(\xi - \bar{\xi})^2}{\text{var}(\xi)} + \frac{(\eta - \bar{\eta})^2}{\text{var}(\eta)} \right) \right] \quad [A^2 \text{ Hz}^{-1}] \tag{50} $$

Noting that the average slope \( \bar{\eta} \) has to be zero in order that \( I(t) \) not diverge, substitution of Eq. (50) in Eq. (49) gives:

$$ PDF_{\text{FA, Rice}} = dt \frac{1}{2\pi \sqrt{\text{var}(I) \text{var}(\eta)}} \exp \left[ -\frac{1}{2} \left( \frac{(I_\text{th} - \bar{I})^2}{\text{var}(I)} \right) \right] \int_0^\infty \eta \exp \left[ -\frac{\eta^2}{2 \text{ var}(\eta)} \right] d\eta $$

$$ = \frac{dt}{2\pi} \frac{\text{var}(\eta)}{\text{var}(I)} \exp \left[ -\frac{1}{2} \left( \frac{(I_\text{th} - \bar{I})^2}{\text{var}(I)} \right) \right] [\text{Hz}] \tag{51} $$

The FAR is just Eq. (51) without the differential \( dt \):

$$ \text{FAR}_{\text{Rice}} = \frac{1}{2\pi} \frac{\text{var}(\eta)}{\text{var}(I)} \exp \left[ -\frac{1}{2} \left( \frac{(I_\text{th} - \bar{I})^2}{\text{var}(I)} \right) \right] [\text{Hz}] \tag{52} $$

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

$$ \text{var}(I) = \psi_0 = \lim_{t \to \infty} \frac{1}{t} \int_0^t I(t) I(t + \tau) dt \quad [A^2], \tag{53} $$

and

$$ \text{var}(\eta) = -\psi_0 \equiv -\frac{\partial^2}{\partial \tau^2} \psi \bigg|_{\tau=0} \quad [A^2/s^2]. \tag{54} $$

The autocorrelation function is itself related to the spectral intensity, \( S_\text{t} \), of the noisy current, by inversion of the Wiener-Khintchine theorem:

$$ \psi(\tau) = \int_0^\infty S_\text{t}(f) \cos(2\pi f \tau) df \quad [A^2], \tag{55} $$

so

$$ \text{var}(I) = \psi_0 = \int_0^\infty S_\text{t}(f) df \quad [A^2], \tag{56} $$

and

$$ \text{var}(\eta) = 4\pi^2 \int_0^\infty f^2 S_\text{t}(f) df \quad [A^2/s^2]. \tag{57} $$

Substituting Eq. (56 & 57) into Eq. (52), the FAR for Gaussian-distributed noise is:

\[ * \]

This is \( S_{\text{I total}} \) – the total noise current spectral intensity of the photoreceiver, referred to the node between APD and TIA, previously given by Eq. (12). Although \( S_\text{t} \) cancels out in the FAR for Gaussian-distributed noise, we will shortly make use of it for the modified calculation for McIntyre-distributed noise.
\[
FAR_{Rice} = \frac{1}{2\pi} \sqrt{\frac{4\pi^2}{\int_0^\infty f^2 S_f(f) df}} \exp\left[-\frac{1}{2} \left(\frac{(I_{th} - \bar{I})^2}{\text{var}(I)}\right)\right] [\text{Hz}].
\] (58)

When the noise spectrum is white (constant \(S_f\)) over a finite bandwidth \(BW\), \(S_f\) cancels out in the radical and Eq. (58) becomes:

\[
FAR_{Rice} = \sqrt{\frac{1}{3} \frac{BW^3}{BW}} \exp\left[-\frac{1}{2} \left(\frac{(I_{th} - \bar{I})^2}{\text{var}(I)}\right)\right] = \sqrt{\frac{1}{3} \frac{BW}{BW}} \exp\left[-\frac{\Delta I_{th}^2}{2I_{noise}^2}\right] [\text{Hz}].
\] (59)

Eq. (59) is the expression for \(FAR\) found in most references, such as the RCA/Burle Electro-Optics Handbook. In Eq. (59), the symbol \(\Delta I_{th}\) is the excess of the detection threshold above the mean current level, and \(I_{noise}\) is the standard deviation of the current, as in Eq. (15).

Calculating \(FAR\) with better accuracy at threshold levels set high in the tail of an APD photoreceiver’s output distribution requires using in place of the Gaussian distribution of \(\xi\) assumed by Rice the convolution of the APD’s McIntyre-distributed output with the Gaussian-distributed TIA noise, \(P_{RX}(n)\), given by Eq. (47). \(P_{RX}(n)\) is an electron count distribution (referred to the node between APD and TIA), but it can be used for the current distribution through a change of variable. As previously discussed in the section APD Output Distribution, Ramo’s theorem says that the APD’s terminal current is a monotonic function of the instantaneous carrier population, which we approximate as equal to \(n\):

\[
\xi = \frac{q}{t_{transit}} n [\text{A}].
\] (60)

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

\[
p(\xi)_{RX} = \frac{d}{d\xi} n(\xi) P_{RX} [n(\xi)] = \frac{t_{transit}}{q} P_{RX} \left(n = \frac{t_{transit}}{q} \xi\right) [\text{A}^{-1}].
\] (61)

The joint probability distribution of the current and its slope, equivalent to Eq. (50), is:

\[
p(\xi, \eta; t)_{McIntyre} = \frac{1}{\sqrt{2\pi \text{var}(\eta)}} \frac{t_{transit}}{q} P_{RX} \left(n = \frac{t_{transit}}{q} \xi\right) \exp\left[-\frac{\eta^2}{2\text{var}(\eta)}\right] [\text{A}^{-2} \text{Hz}^{-1}].
\] (62)

Substituting the modified joint probability distribution into Eq. (49) gives:

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

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

\[
= dt \frac{\text{var}(\eta)}{\sqrt{2\pi \text{var}(\eta)}} \frac{t_{transit}}{q} P_{RX} \left(n = \frac{t_{transit}}{q} I_{th}\right)
\]

\[
= dt \frac{\text{var}(\eta)}{2\pi \text{var}(I)} \frac{t_{transit}}{q} P_{RX} \left(n = \frac{t_{transit}}{q} I_{th}\right) \sqrt{2\pi \text{var}(I)} [\text{Hz}].
\] (63)
Note that the last line of Eq. (63) was multiplied by $1 = \sqrt{2\pi \text{var}(I)} / \sqrt{2\pi \text{var}(I)}$ to cast the expression in the same form as Eq. (51), whereby the operations of Eq. (52-58) can be applied to find the FAR equivalent to Eq. (59):

$$\text{FAR}_{\text{McIntyre}} = \sqrt{\frac{2\pi}{3}} \frac{t_{\text{transit}}}{q} I_{\text{noise}} BW P_{\text{RX}} \left( n_{th} = \frac{t_{\text{transit}}}{q} I_{th} \right) \quad \text{[Hz]}. \quad (64)$$

The conditions for which the $P_{\text{RX}}(n)$ curves of Figure 14 were calculated result in factors in front of $P_{\text{RX}}$ 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 Eq. (64) is compared to that calculated using Eq. (59) in Figure 16. The more realistic model reveals that a few standard deviations beyond the mean ($\bar{n} = 170.06; \sigma_{k=0} = 188.82; \sigma_{k=0.2} = 191.71; \sigma_{k=0.4} = 194.57$) FAR drops off more slowly with increasing detection threshold – and is much more sensitive to $k$ – than predicted by Rice’s model.

Eq. (64) can be applied to calculate the FAR of either RTIA- or CTIA-based photoreceivers. In the CTIA case, the effective integration period $\tau$ is used in place of the junction transit time $t_{\text{transit}}$ 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 Eq. (22 or 24), in the absence of an optical signal:

$$\text{FAR}_{\text{McIntyre}} = \sqrt{\frac{2\pi}{3}} N_Q BW P_{\text{RX}} \left( n_{th} = \frac{t_{\text{transit}}}{q} I_{th} \right) \quad \text{[Hz]}. \quad (65)$$

**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. 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. 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 $n_c$:

$$P[1|0] = 1 - \sum_{n_t = -\infty}^{n_t} P_{\text{RX}} (n_t), \quad (66)$$

and
\[ P(0|1) = \sum_{n_0} P_{RX}(n_1), \]  
(67)

where \( n_0 \) and \( n_1 \) are discrete random variables which represent the effective carrier count at the node between APD and TIA, calculated for primary photocurrent levels corresponding to the optical signal levels coding binary “0” or “1” values. \( P_{RX}(n) \) is the distribution of the photoreceiver’s output, referred to this node, and is calculated according to Eq. (47) for conventionally structured APDs and according to Eq. (48) for multi-stage Siletz APDs. When making calculations for conventionally structured APDs using Eq. (67), the primary current used in Eq. (45 & 46) to compute the APD’s output distribution for convolution with the TIA’s noise is the sum of the primary dark current and photocurrent; when making calculations for Siletz APDs using Eq. (48), the primary photocurrent and dark current are treated in separate distributions which are subsequently convolved, as described in the section Convolution of APD and TIA Distributions.

The primary photocurrent is found from the optical power incident on the APD by setting \( M=1 \) in Eq. (6 & 7). Since 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, \( P_1 \):

\[ P_0 = P_1 \times 10^{\frac{ER}{10}}, \]
(68)

where \( ER \) is the extinction ratio of the modulator in dB (typically 15 – 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:

\[ P_{\text{Watts}} = 1 \text{ mW} \times 10^{\frac{P_{\text{dBm}}}{10}}. \]
(69)

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 weighting of both the error rate and 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:

\[ BER = R_1 P[0|1] + (1-R_1) P[1|0]. \]
(70)

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:

\[ P_{av} = R_1 P_1 + (1-R_1) P_1 \times 10^{\frac{ER}{10}}. \]
(71)

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), \( P_{av}=0.5 P_1 \).

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: \(^{36}\)

\[ n_1|_{\text{optimum}} = \frac{\langle n_1 \rangle \sqrt{\text{var}(n_0)} + \langle n_0 \rangle \sqrt{\text{var}(n_1)}}{\sqrt{\text{var}(n_0)} + \sqrt{\text{var}(n_1)}}. \]
(72)
The mean and standard deviation of \( n_0 \) and \( n_1 \) appearing in Eq. (72) are the photoreceiver’s signal and noise under the “0” and “1” signal conditions, which can be calculated as described in the sections on Mean (Signal) and Variance (Noise).

In the case of a photoreceiver with Gaussian-distributed output, if the decision threshold is set as in Eq. (72), the BER is:

\[
BER = \frac{1}{2} \text{erfc} \left( \frac{\langle n_1 \rangle - \langle n_0 \rangle}{\sqrt{2 \left( \sqrt{\text{var}(n_0)} + \sqrt{\text{var}(n_1)} \right)}} \right). \tag{73}
\]

If the receiver’s TIA noise dominates the shot noise on the detector’s dark current and photocurrent (including the photocurrent shot noise when receiving a “1”), and if the modulator’s extinction ratio is large, then the BER can be approximated in terms of the receiver’s signal-to-noise ratio, as defined in Eq. (25):

\[
BER \approx \frac{1}{2} \text{erfc} \left( \frac{\text{SNR}}{2\sqrt{2}} \right). \tag{74}
\]

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

There are several reasons why Eq. (74) is not accurate for APD-based photoreceivers. First, as discussed previously in the context of FAR, the distribution of the APD’s output is not Gaussian, and divergence of the distribution’s tail from the Gaussian approximation several standard deviations away from its mean can significantly impact \( P[1|0] \). Also, the skewness of the APD’s output distribution means that Eq. (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’s photocurrent in order to equate \( \sqrt{\text{var}(n_0)} \approx \sqrt{\text{var}(n_1)} \) to simplify the form of Eq. (74) is 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 Eq. (66 & 67), using either Eq. (47 or 48) for \( P_{RX}(n) \), depending upon the APD’s internal structure. For a given optical power level coding a “1” \( (P_1) \) and a given extinction ratio (ER), BER depends on the APD’s 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 built up by stepping through values of \( P_1 \), using Eq. (71) to convert \( P_1 \) to average power; the average power for which a particular BER is achieved is the receiver’s 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 at which the ratio
is maximized, keeping in mind that it’s difficult to operate conventionally structured InGaAs APDs above \( M=20 \) or Siletz APDs above \( M=50 \). \( I_{\text{signal}} \) and \( I_{\text{noise}} \) are respectively calculated according to Eq. (7) and Eq. (15), and the ratio in Eq. (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’s output distribution, so it will tend to underestimate \( P[1|0] \). Optimal gain operating points are often lower than found by maximizing Eq. (75), whereas optimal decision thresholds are often higher than Eq. (72).

### compares alternate calculations of optimal decision threshold and gain as functions of average signal power for an APD photoreceiver assembled from a 75-μm-diameter Deschutes model APD and the MAX3277 TIA. ### compares the two calculations of BER at optimal threshold and gain as functions of average signal power. Substantial divergence from the Gaussian approximation of Eq. (74) is evident.

#### 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 true and false positive rates should be defined for maximal relevance to the physical problem being solved. For instance, suppose that a simple laser range-finding (LRF) 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. Since there ought to be one return per target for every transmitted laser pulse, the TPR is just the pulse detection probability, \( P_D \), as calculated from the CCDF of \( P_{\text{RX}}(n) \) with optical signal present. However, the raw false alarm probability, \( P_{FA} \), calculated from the CCDF of \( P_{\text{RX}}(n) \) 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 doesn’t 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. Since false alarms are uniformly distributed in time, one applies Poisson statistics to

\[ C = \frac{I_{\text{signal}} \negmedspace 1 - I_{\text{signal}} \negmedspace 0}{\sqrt{I_{\text{noise}} \negmedspace 1^2 + I_{\text{noise}} \negmedspace 0^2}} \]  

(75)

For multi-stage Siletz APDs, make sure to use Eq. (20) for \( S_{\text{total}} \) in Eq. (15).

† As of this draft of the Technical Note, plots for this section have not yet been generated.
calculate the probability of zero false alarms occurring during the range gate \( \tau \), with the expected value of the number of alarms equal to \( FAR \times \tau \):

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

(76)

\( FAR \) is calculated as in Eq. (64) for RTIA-based photoreceivers or Eq. (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. Figure 18 shows example ROCs that were calculated for a LRF receiver assembled from a 75-μm-diameter Deschutes APD and the model VX-809E application-specific integrated circuit (ASIC), assuming 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 VX-809E’s effective signal integration time was 8.2 ns, and its input-referred charge noise was 314 e-.

**Parameterization of Terminal Dark Current for Voxel APDs**

The parameterizations for Deschutes model APDs are accurate in the range 1≤M≤20 but diverge from empirical measurements for M>20. Bear in mind that the parameterizations were fit to average device behavior, but dark current varies somewhat from part to part.

**Deschutes**

75 μm: \[ I_{\text{dark}} = \exp[0.05 \times (T-27^\circ C)] \times (-0.080665 + 0.29786 M - 0.0076941 M^2 + 0.00010214 M^3) \] [nA]

200 μm: \[ I_{\text{dark}} = \exp[0.05 \times (T-27^\circ C)] \times (-0.7902 + 2.7376 M - 0.104 M^2 + 0.001701 M^3) \] [nA]

(77)

The parameterizations for Siletz model APDs are accurate in the range 1≤M≤50 but diverge from empirical measurements for M>50. As with the Deschutes parts, bear in mind that the parameterizations were fit to average device behavior, but dark current varies from part to part.

**Siletz**

75 μm: \[ I_{\text{dark}} = \exp[0.0234 \times (T-27^\circ C)] \times (-28.262 + 32.557 M - 0.29065 M^2 + 0.0036571 M^3) \] [nA]

200 μm: \[ I_{\text{dark}} = \exp[0.0234 \times (T-27^\circ C)] \times (-103.42 + 69.477 M + 2.425 M^2 - 0.040039 M^3) \] [nA]

(78)

**Burgess Variance Theorem for Multiplication & Attenuation**

The Burgess variance theorem\(^3\,^4\) is applied to introduce the APD’s excess noise factor in the sub-section of the *Introduction* titled *Avalanche Gain and Gain Distribution*, and is mentioned in connection to attenuation of noisy optical signals at the end of the sub-section titled *CTIA Case for Conventional*
InGaAs APDs of the section on Variance (Noise). In this section, derivations of the Burgess variance theorem for these two applications is described, and the theorem is applied to treat attenuation of an optical signal generated by a pulsed laser with large pulse energy variability.

**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 resulting from a fluctuating per-photon binary transmission outcome \( t \) that is summed over a fluctuating input photon count \( b \). In the following derivation, we will explicitly use the \{n, a, m\} variable set for avalanche multiplication, remembering that we can make the substitutions \{n, a, m\} \rightarrow \{p, b, t\} 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, \( M_{\text{const}} \), then \( n = a \times M_{\text{const}} \) and by the basic rule for computing the variance of the product of a constant with a random variable, \( \text{var}(n) = M_{\text{const}}^2 \text{var}(a) \). However, when the per-electron discrete gain is itself a random variable, the product \( a \times m \) doesn’t correspond to \( n \) because a single value of \( m \) doesn’t multiply every electron of an \( a \)-electron input current pulse. Rather, every electron of the fluctuating quantity \( a \) is multiplied by a potentially different value of the fluctuating gain \( m \), and \( \text{var}(n) \) is computed from the statistics of \( a \) and \( n|a \) (\( n \) given \( a \)).

Begin with the definition of variance:

\[
\text{var}(n) \equiv \langle n^2 \rangle - \langle n \rangle^2. \tag{79}
\]

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:

\[
P_{\text{binomial}}(p|b) = \binom{b}{T} T^p (1-T)^{b-p}. \tag{80}
\]

Likewise, in the case of avalanche multiplication, \( n|a \) obeys the McIntyre distribution given earlier in Eq. (2).

Assuming the distribution functions \( P(a) \) and \( P(n|a) \) exist, we can symbolically write the expected values \( \langle n \rangle \) and \( \langle n^2 \rangle \) (the mean and mean square).\(^*\) Since \( n|a \) is the sum of \( a \) random variables, each

\(^*\) Note that depending upon 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.
distributed as \( m \), the expected value \( \langle n|a\rangle \) can be rewritten as the expected value of the sum of \( a \) random variables \( m_i \). Applying the linearity of the expectation operator to write the expectation of the sum as the sum of the individual expectations, we get:

\[
\langle n \rangle = \sum_{a=0}^{\infty} \sum_{n=0}^{\infty} P(a) \cdot \langle n|a\rangle \cdot n = \sum_{a=0}^{\infty} P(a) \cdot \sum_{n=0}^{\infty} m_i |a\rangle = \sum_{a=0}^{\infty} P(a) \sum_{i=1}^{a} m_i |a\rangle \\
= \sum_{a=0}^{\infty} P(a) \cdot a \langle m|a\rangle = \sum_{a=0}^{\infty} P(a) \cdot a \cdot M = M \times \langle a \rangle \]  

(81)

In the last line of Eq. (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.

The mean square is given by:

\[
\langle n^2 \rangle = \sum_{a=0}^{\infty} \sum_{n=0}^{\infty} P(a) \cdot \langle n|a\rangle \cdot n^2 = \sum_{a=0}^{\infty} P(a) \cdot \langle n^2|a\rangle = \sum_{a=0}^{\infty} P(a) \cdot \left[ \text{var}(n|a) + \langle n|a\rangle^2 \right] \\
= \sum_{a=0}^{\infty} P(a) \cdot \left[ \text{var} \left( \sum_{i=1}^{a} m_i |a\rangle + (a \cdot M)^2 \right) \right] = \sum_{a=0}^{\infty} P(a) \cdot \left[ \sum_{i=1}^{a} \text{var}(m_i |a\rangle + (a \cdot M)^2 \right] \\
= \sum_{a=0}^{\infty} P(a) \cdot \left[ a \text{var}(m) + a^2 M^2 \right] = \langle a \rangle \text{var}(m) + \langle a^2 \rangle M^2  

(82)

where the definition of variance has been used to rewrite \( \langle n^2|a\rangle = \text{var}(n|a) + \langle n|a\rangle^2 \) and the result \( \langle n|a\rangle = a \cdot M \) that was found in Eq. (81) has been applied. In the second line of Eq. (82), \( \text{var}(n|a) \) 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 Eq. (82). As with the assumption in Eq. (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 \( \text{var}(m) \) depends on \( a \).

Substitution of Eq. (81 & 82) into Eq. (79) gives the form of the Burgess variance theorem originally presented in Eq. (3):

\[
\text{var}(n) \equiv \langle n^2 \rangle - \langle n \rangle^2 = \langle a \rangle \text{var}(m) + \langle a^2 \rangle M^2 - M^2 \langle a \rangle^2 = M^2 \text{var}(a) + \langle a \rangle \text{var}(m) .  

(83)

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 \( \text{var}(m) \rightarrow \text{var}(t)=T\times(1-T) \), the Burgess variance theorem can be written for optical attenuation as:

\[
\text{var}(p) = T^2 \text{var}(b) + \langle b \rangle \text{var}(t) = T^2 \text{var}(b) + \langle b \rangle T (1-T) .  

(84)
Based on the mean and variance of the number of input photons, respectively $\langle b \rangle$ and $\text{var}(b)$, and the average transmission probability of the attenuator, $T$, one can use Eq. (84) 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 about 1%, but with an average amplitude of about 8E14 photons, 1% of the mean pulse energy is around 2E12 photons. If the laser’s output was Poisson-distributed, its variance would equal its mean; instead, its variance is more than 6E9 times the mean. Indeed, the width of a Poisson-distributed signal would not be perceivable on the x-axis of Figure 19.

Since the noise analyses presented in this technical note generally proceed from the assumption that the input optical 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, one seldom uses a sensitive optical receiver to measure a 100-μJ pulse. Typically, optical signals reach photoreceivers after significant attenuation. Various processes like atmospheric scintillation can affect signal amplitude statistics at the receiver, but the Burgess variance theorem of Eq. (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 $\text{var}(b)=5.19E24$ and $\langle b \rangle=7.96E14$. Consider two different attenuations – to 100 photons (for which $T=100/\langle b \rangle=1.26E-13$ and to 100,000 photons (for which $T=1E5/\langle b \rangle=1.26E-10$. In the case of attenuation to 100 photons, Eq. (84) gives a variance that is essentially indistinguishable from that of a Poisson-distributed signal (variance of about 100 photons equal to the mean). The reason is that attenuation by a very large factor cuts the magnitude of the noise that is originally on the optical 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.
References

26. S. Ramo, “Currents Induced by Electron Motion,” *Proc. IRE*, vol. 27, no. 9, pp. 584-585, 1939.