Statistical Correlation of Gain and Buildup Time in APDs and Its Effects on Receiver Performance

Peng Sun, Majeed M. Hayat, Senior Member, IEEE, Fellow, IEEE, Bahaa E. A. Saleh, Fellow, IEEE, Malvin C. Teich, Life Fellow, IEEE

Abstract. We report a novel recurrence theory that enables us to calculate the exact joint probability distribution function (PDF) of the random gain and the random avalanche-buildup time in avalanche photodiodes (APDs) including the effect of dead space. Such calculations reveal a strong statistical correlation between the gain and the buildup time for all widths of the multiplication region. To facilitate the calculation of the photocurrent statistics in the presence of this correlation, the impulse-response function of the APD is approximately modeled by a function of time whose pre-specified shape is appropriately parameterized by two random variables: the gain and the buildup time. The evaluation of the variance of the photocurrent under this model leads to the definition of the shot-noise-equivalent bandwidth of the APD that captures the statistical correlation between the gain and the buildup time. It is shown that the shot-noise-equivalent bandwidth in GaAs APDs is greater, by approximately 30%, than the traditional buildup-time-limited 3-dB bandwidth, which is calculated from the mean of the impulse-response function. A thorough analysis of the performance of APD-based integrate-and-dump digital receivers reveals that the strong correlation between the gain and the buildup time accentuates intersymbol-interference noise and thus adversely affects receiver sensitivity at high transmission rates beyond previously known limits.

Keywords. Avalanche photodiodes, buildup time, bandwidth, gain, correlation, impulse response function, shot-noise, intersymbol-interference, bit-error rate, receiver sensitivity.

This work was supported by the National Science Foundation through awards ECS-0334813 and ECS-0196569.

P. Sun and M. M. Hayat are with the Department of Electrical and Computer Engineering, The University of New Mexico, Albuquerque, NM 87131–0001 USA (E-Mail: {pengsun,hayat}@ece.unm.edu).

B. E. A. Saleh and M. C. Teich are with the Department of Electrical and Computer Engineering, Boston University, Boston, MA 02215–2421 USA (E-Mail: {besaleh,teich}@bu.edu).
1. Introduction

Avalanche photodiodes (APDs) are widely-used photodetectors in many high-speed optical receivers including those deployed in 10-Gbps/channel lightwave systems. The popularity of APDs is due to their ability to provide high internal optoelectronic gains. In high-speed systems, the gain that an APD provides translates into improved receiver sensitivity, as the gain combats the Johnson noise in the preamplifier stage of an optical receiver. This benefit, however, comes at the often tolerable expense of an increase in shot noise by a factor called the APD’s excess noise factor, $F$, which is a measure of the gain uncertainty [1–4]. In addition, the APD’s avalanche buildup time, which is the duration of the APD’s single-photo-excited impulse response function, limits the APD’s bandwidth and causes intersymbol-interference (ISI) in digital communication. The buildup time has heretofore been the factor that limits the use of APDs in 40-Gbps systems to this date.

Advances in APD technology, such as the development of APDs that may feature evanescent light coupling into very thin multiplication regions (where both carrier transit time and avalanche buildup time are significantly reduced without much sacrifice in coupling efficiency) may be a promising path for bringing APDs to the 40-Gbps arena [5–14]. On the other hand, to know the exact device requirements for specified system performance requires (yet to be developed) exact theoretical models for assessing the receiver performance at high speeds, where factors such as ISI and the stochastic nature of the APD’s impulse response function play critical roles. In this paper we exactly compute the joint statistics of the random gain and random buildup time in APDs for the first time and incorporate these statistics in the analysis of APD-based receivers at high transmission speeds. This theory is expected to impact the way we translate system requirements (e.g., those for next-generation lightwave systems) to device requirements. Moreover, this work may also impact the design and performance of emerging strategies (such as equalization and for forward-error correction) used to mitigate ISI and polarization-mode-dispersion in high-speed lightwave systems.

The APD’s impulse-response function is a stochastic process, with a random duration (viz., avalanche buildup time) and a random area representing the multiplication factor, or gain. Moreover, the random gain and random buildup time are statistically correlated. One manifestation of this correlation is the APD’s gain-bandwidth product (GBP), traditionally
calculated as the product between the mean gain and the 3-dB bandwidth (calculated from the Fourier transform of the average impulse response function). However, the GBP conveys the correlation in the gain and the buildup time in a very limited way—being the product of mean quantities. In particular, it does not capture the inherent correlation in the fluctuations in the gain and the buildup time. For example, even if an APD is operated to yield a certain fixed mean gain, specific realizations of the avalanche process that exhibit high gains are accompanied with long buildup times (those realizations result in greater ISI and adversely affect receiver sensitivity), and vice versa. Consequently, at transmission rates near the APD’s bandwidth, where buildup-time uncertainty is expected to affect ISI, it is plausible to suspect that the statistical correlation between the random gain and the random buildup time may play an intricate role in receiver performance.

To understand the complex interplay between the buildup time and the gain, one must appeal to the statistical analysis of the APD’s impulse-response function. For example, to calculate the variance of the photocurrent, knowledge of the second moment of the impulse-response function is required at each time. Further, in order to calculate the variance of the integrated photocurrent, as required when assessing the bit-error rate (BER) of an integrate-and-dump receiver, the autocorrelation function of the impulse response is also necessary [15,16]. Although recurrence relations for these quantities are available, their solutions require intensive computing, especially for the autocorrelation function [16]. Moreover, calculation of higher-order statistics (e.g., the probability distribution of the photocurrent, which is required in the calculation of the probability distribution of the output of an integrate-and-dump-receiver) is practically impossible with existing models.

In this paper, we also provide a novel approximate method for calculating the statistics of the impulse-response function using our exact knowledge of the joint statistics of the random gain and buildup time. The rationale is to first approximate the random impulse-response function by a function of time whose pre-specified shape is parameterized by the random gain and the random buildup time. A simple example of such a function is a rectangular random-duration function whose duration is the buildup time and its area is proportional to the gain (more complex functions may also be considered). The exact calculation of the joint probability distribution function (PDF) of the gain and buildup time is accomplished
by developing a novel recursive theory that generalizes the existing recursive techniques for computing the marginal PDFs of the gain [17,18] and the buildup-time [19]. The computations are carried out for homojunction APDs under the simplifying assumptions of a constant electric field (in the multiplication region) and a constant drift velocity for carriers. We emphasize that the theory developed here incorporates the dead-space effect, which is inherent in the process of cascaded impact ionizations. Dead space is the minimum distance a newly generated carrier must travel before it becomes capable of effecting an impact ionization. It is well known that this effect becomes increasingly more important in thin multiplication layers, where the fast build up time is accompanied by reduced excess noise factor.

We also utilize this avalanche multiplication theory to analyze the receiver performance in an on-off keying optical receiver. We particularly investigate the effect of the statistical correlation in the gain and the buildup time on the receiver performance in high-speed transmission settings.

2. Joint Probability Distribution of Gain and Buildup Time

Consider an APD with a multiplication region extending from \( x = 0 \) to \( x = w \). Assume that an electron is injected into the multiplication region from the left at \( x = 0 \), thereby starting the avalanche process. Upon entering the multiplication region, the injected electron travels to the right a random distance \( X_e,0 \) after which it impact ionizes, resulting in two offspring electrons and a hole. The probability density function (pdf) of the free path \( X_e,0 \) of the injected electron is denoted by \( h_{e,0}(\xi) \). For example, according to the dead-space multiplication theory (DSMT) [20], \( h_{e,0}(\xi) = \alpha(\xi)e^{-\int_{\xi}^{\xi+d_e,0}\alpha(y)dy}u(\xi-d_e,0) \), where, \( \alpha(\xi) \) is the position-dependent nonlocalized ionization coefficient, \( d_e,0 \) is the dead space associated with the injected electron, and \( u(x) \) is the unit step function (i.e., \( u(x) = 1 \) if \( x \geq 0 \) and zero otherwise). After each impact ionization, a newly-created electron (resp., hole) born at position \( x \) travels a free distance \( X_e \) to the right (resp., a distance \( X_h \) to the left) before it impacts ionizes. Let \( h_e(\xi|x) \) and \( h_h(\xi|x) \) denote the pdfs of \( X_e \) and \( X_h \), respectively. [In the DSMT model [20], \( h_e(\xi|x) = \alpha(x + \xi)e^{-\int_{\xi}^{\xi+d_e(x)}\alpha(y)dy}u(\xi - d_e(x)) \) and \( h_h(\xi|x) = \beta(x - \xi)e^{-\int_{\xi}^{\xi+d_h(x)}\alpha(x-y)dy}u(\xi - d_h(x)) \), where \( d_e(x) \) and \( d_h(x) \) are respectively the dead space associated with the electron and holes that are created at \( x \) and \( \beta(\xi) \) is the
position-dependent nonlocalized ionization coefficient for holes.]

For a linear-mode operation of the APD, the cascade of impact ionizations terminate at some finite, random time $T$, called the avalanche buildup time, yielding a net random gain $G$. Note that if we consider the electrical current, or the APD’s impulse-response function due to the injected electron, resulting from the carriers generated by the cascade of impact ionizations, then $T$ is precisely the random duration of the impulse-response function. Our goal is to determine $f_{G,T}(m,t) = P\{G = m, T \leq t\}$, the joint PDF of the random variables $G$ and $T$.

We first define certain intermediate quantities that are essential in the formulation of the recursive (renewal) equations that would ultimately yield $f_{G,T}(m,t)$. For each $x \in [0,w]$, let $Z(x)$ be the total number of carriers (electrons and holes, collectively) generated by a parent electron positioned at $x$ and let $T_e(x)$ be the random time until all these carriers exit the multiplication region. Similarly, we define $Y(x)$ as the total number of carries generated as a result of a parent hole positioned at $x$ and let $T_h(x)$ denote the time until all these carries exit the multiplication region. Note that $G = 0.5(Z(0) + 1)$ and $T_e(0) = T$. Let us now define the joint PDF of $Z(x)$ and $T_e(x)$ by $f_e(m,t;x) = P\{Z(x) = m, T_e(x) \leq t\}$. Similarly define $f_h(m,t;x) = P\{Y(x) = m, T_h(x) \leq t\}$. Note that as a special case, $f_{G,T}(m,t) = f_e(2m - 1,t;0)$. We now invoke a renewal argument that will allow us to recursively characterize $f_e(m,t;x)$ and $f_h(m,t;x)$.

2.1. Recursive equations

In avalanching, once a parent carrier impacts ionizes, the regenerated parent carrier and the offspring carriers independently repeat a similar process as their parent. Suppose that the parent electron at $x$ first impacts the ionization after traveling a distance $X_e = \xi \in [0,w-x]$ from $x$. In such an event, there would be two newly-created carriers (an electron and a hole). Now the key observation is that conditional on the occurrence of this initial ionization, the event that $T_e(x)$ is less than or equal to $t$ and $Z(x) = m$ is precisely the event that the responses due to the newly-created carries all terminate in the remaining time (i.e., $t$ less the electron transport time from $x$ to $x + \xi$, or simply $t - \xi/v_e$) and that the three carriers collectively generate $m$ total offsprings within the remaining time of $t - \xi/v_e$. 

5
Here, $v_e$ is the electron saturation velocity in the multiplication region. By means of this renewal argument and by using standard properties of conditional expectations and sums of independent random variables, we can see that by conditioning on the occurrence of the initial ionization at $\xi$, the conditional PDF $f_{e}(m, t; x | \xi)$ can be written as

$$f_{e}(m, t; x | \xi) = f_{e}(m, t - \xi/v_e; x + \xi) * f_{e}(m, t - \xi/v_e; x + \xi) * f_{h}(m, t - \xi/v_e; x + \xi),$$

(1)

where $*$ denotes discrete convolution (in the variable $m$). Similarly, the conditional PDF $f_{h}(m, t; x | \xi)$ can be written as

$$f_{h}(m, t; x | \xi) = f_{h}(m, t - \xi/v_h; x - \xi) * f_{h}(m, t - \xi/v_h; x - \xi) * f_{e}(m, t - \xi/v_h; x - \xi),$$

(2)

where $v_h$ is the hole saturation velocity in the multiplication layer.

The condition on the location $\xi$ of the first impact ionization can now be removed by averaging over all possible locations $\xi$ in the interval $[0, w - x]$. However, it is possible that the parent electron (or hole) born at $x$ does not impact ionizes at all, corresponding to the event $\{X_e > w - x\}$ (or $\{X_h > x\}$). In particular, if $X_e > w - x$ (or $X_h > x$), then $T_e(x) = (w - x)/v_e$ (or $T_h(x) = x/v_h$) and $Z(x) = 1$ (or $Y(x) = 1$).

With the above considerations and after performing an average over all possible locations of the first ionization, we obtain the following coupled recursive equations which together characterize $f_{e}(m, t; x)$ and $f_{h}(m, t; x)$:

$$f_{e}(m, t; x) = g_{e}(x, t)\delta_{m-1} + \int_{0}^{w-x} \{f_{e}(m, t - \xi/v_e; x + \xi) * f_{e}(m, t - \xi/v_e; x + \xi) * h_{e}(\xi|x)\} d\xi,$$

(3)

$$f_{h}(m, t; x) = g_{h}(x, t)\delta_{m-1} + \int_{0}^{x} \{f_{h}(m, t - \xi/v_h; x - \xi) * f_{h}(m, t - \xi/v_h; x - \xi) * h_{h}(\xi|x)\} d\xi,$$

(4)

where, $\delta(k)$ is equal to 1 if $k = 0$ and zero otherwise,

$$g_{e}(x, t) = (1 - \int_{x}^{w} h_{e}(\xi|x) d\xi) u(t - \frac{w - x}{v_e})$$

(5)

and

$$g_{h}(x, t) = (1 - \int_{0}^{x} h_{h}(\xi|x) d\xi) u(t - \frac{x}{v_h}).$$

(6)
The convolutions in the above equations can be converted into products by taking the $z$-transform ($0 \leq |z| \leq 1$) of $f_e(m, t; x)$ and $f_h(m, t; x)$ with respect to the variable $m$. The transformed quantities, denoted by $F_e(z, t; x)$ and $F_h(z, t; x)$, respectively, will satisfy
\begin{align}
F_e(z, t; x) &= g_e(x, t)z + \int_0^{w-x} F_e^2(z, t - \frac{\xi}{v_e}; x + \xi)F_h(z, t - \frac{\xi}{v_e}; x + \xi)h_e(\xi|x) \, d\xi \tag{7}
\end{align}
and
\begin{align}
F_h(z, t; x) &= g_h(x, t)z + \int_0^{x} F_h^2(z, t - \frac{\xi}{v_h}; x - \xi)F_e(z, t - \frac{\xi}{v_h}; x - \xi)h_h(\xi|x) \, d\xi. \tag{8}
\end{align}

We next introduce a numerical recipe for computing $F_e$, $F_h$, $f_e$ and $f_h$.

2.2. Numerical solutions

Calculation of $F_e$ and $F_h$: For any fixed $z$, we can solve equations (7) and (8) numerically using the simple iterative method described below. We first select a maximum limit $t_{\text{max}}$ for the range of the buildup time to be considered. We then select a mesh size for the time $t$ and the space $x$ allowing the discretization of the function $F_1$ and $F_2$, and hence, converting the integrals into summations. We set the zeroth iteration of the functions $F_e(z, t; x)$ and $F_h(z, t; x)$, denoted by $F_e^{(0)}(z, t; x)$ and $F_h^{(0)}(z, t; x)$, respectively, to be $g_e(x, t)z$ and $g_h(x, t)z$ for $0 \leq x \leq w$ and $0 \leq t \leq t_{\text{max}}$. Next, the first iterates, $F_e^{(1)}(z, t; x)$ and $F_h^{(1)}(z, t; x)$, are computed by substituting the zeroth iterates into the right-hand side of (7) and (8). Subsequent iterates are generated in the same way. This procedure is continued until the iterates converge uniformly in $t$ and $x$. More precisely, we stop the iteration process when the maximum relative change from the $(n - 1)$st to the $n$th iterate in the functions $F_e^{(n)}(z, t; x)$ and $F_h^{(n)}(z, t; x)$ is below a predefined tolerance level. At such point, the procedure for computing $F_e$ and $F_h$ is complete and we set them to the respective values at the last round of iterations.

Calculation of $f_{G,T}$: Let $F_{G,T}(z, t)$ be the $z$-transform of $f_{G,T}(m, t)$ with respect to $m$. Then, since $f_{G,T}(m, t) = f_e(2m - 1, t; 0)$, we obtain
\begin{align}
F_{G,T}(z, t) &= \sqrt{z}F_1(\sqrt{z}, t; 0),
\end{align}
which we need to convert back to the $m$ domain. Since $G$ is an integer-valued random variable, we can apply the efficient technique proposed by Gubner and Hayat [21] to recover the $f_{G,T}(m, t)$ from $F_{G,T}(z, t)$ with $z$ restricted to the unit circle.
Figure 1 shows an example of the joint PDF $f_{G,T}(m,t)$ calculated for a homojunction GaAs APD with a 160-nm multiplication layer exhibiting an average gain of 10.46. The electron and hole saturation velocities are assumed to be $0.67 \times 10^7$ cm/s (from Ref. [22]) and the nonlocalized ionization coefficients for electrons and holes and their respective ionization threshold energies are taken from Ref. [23]. The joint probability density function (pdf), shown in Fig. 2, is obtained by taking the derivative of $f_{G,T}(m,t)$ with respect to $t$. Recall that if we restrict the random gain realization to unity (the cross section corresponding to $m = 1$ in Fig. 1), then the PDF of $T$ becomes a step function with a jump at the electron transit time $w/v_e$. Consequently, the joint pdf of $T$ is expected to exhibit a Dirac impulse at the electron transit time. These two features are observed in Figs. 1 and 2, respectively. Figure 2 also shows that as the realized value $m$ of the random gain increases, the pdf of the $T$ (i.e., cross section of the joint pdf along the t-axis) shifts toward a higher buildup-time mean and a larger spread about the mean. These are clear indications of the statistical correlation between the gain and the buildup time as well as the increase in the buildup-time uncertainty at higher realizations of the gain.

Similarly, the probability mass function of $G$ for various realizations $t$ of the random buildup time $T$ shifts toward a higher mean (accompanied by a larger spread) as $t$ increases. This also confirms the correlation between the gain and the buildup time. Figure 3 (bottom graph) shows the correlation coefficient $\rho$ of $G$ and $T$ as a function of the width of the multiplication region while the mean gain is held fixed at 10. The values of correlation coefficients are approximately 0.9 for the entire range of $w$.

We next use the joint PDF of $G$ and $T$ to calculate the statistical properties of the impulse-response function and investigate the effect of the correlation between $G$ and $T$ on ISI and receiver performance in high-speed digital receivers.

### 3. Parametric Modeling of the Impulse Response Function and the Shot-Noise-Equivalent Bandwidth

In this section, we introduce a novel stochastic model for the impulse-response function that facilitates the calculation of its statistics once the joint PDF of the gain and the buildup time has been determined. The rationale is to approximate the impulse-response function
\( I_p(t) \) by a pre-specified shape function of time that is parameterized by the random gain \( G \) and the random buildup time \( T \). In this fashion, the randomness of the impulse-response function, as a stochastic process, is lumped in the random parameters \( G \) and \( T \). This significantly simplifies the complexity of the impulse-response function while maintaining the key features that govern the excess-noise and speed properties of the APD; namely, randomness in the impulse-response function’s area (representing gain uncertainty) and randomness in its duration (representing bandwidth uncertainty). An example of such a shape function is the rectangular random-duration (RD-rectangular or RD-R) with random height \( qG/T \) and random duration \( T \), where \( q \) is the electronic charge as shown in Fig. 4. Note that the area under this function is \( qG \). Another example is the triangular random-duration (RD-triangular or RD-T) model, shown in the same figure, for which the random duration is again \( T \) but a peak occurs at the electron transit time, \( \tau_e \), across the multiplication region. The peak in the RD-triangular model is consistent with the behavior of the true impulse-response function, as all first-generation electrons continue to multiply until they simultaneously exit the multiplication region at precisely \( \tau_e \) units of time past the time of the primary-electron injection. The height of the peak is \( 2qG/T \) so that the area under the shape function is \( qG \). Other shape functions may also be considered.

To see the validity of the proposed RD-rectangular model and RD-triangular model, we calculated the mean impulse-response function, denoted by \( i_p(t) = E[I_p(t)] \), and compared it to the exact theoretical model for the mean impulse-response function [15]. For example, it can easily shown that for the rectangular model,

\[
E[I_p(t)] = \sum_{m=1}^{\infty} \int_t^{\infty} qm\tau^{-1} \frac{\partial}{\partial \tau} f_{G,T}(m, \tau) d\tau, \quad t \geq 0,
\]

which can be easily computed as the joint distribution \( f_{G,T} \) has already been computed. Figure 5 depicts \( i_p(t) \) obtained using both the rectangular and triangular parametric models. Indeed, the area under \( i_p(t) \) is 10.48q, which is in excellent agreement with the theoretical average gain of 10.46q that is calculated independently using Ref. [23]. The comparison with the theoretical prediction of the APD mean impulse-response function by Hayat and Saleh [15] is also shown on this figure (dashed curve). Note that, due to the frequency-response characteristics of the rectangular and triangular waveforms, the rectangular model slightly
underestimates the bandwidth while the triangular model overestimates it. Nevertheless, the approximation is very good in both cases.

### 3.1. Bandwidth

The calculation of the variance of the photocurrent generated by an APD (i.e., the variance of the filtered shot noise [24]) requires knowledge of the second moment of the APD’s impulse-response function of time $t$. In particular, the variance of the photocurrent is given by [25] $q(\eta P/h\nu) \int_{-\infty}^{\infty} \langle I_p^2(t) \rangle dt$, where $P$ is the optical power, $\eta$ is the APD’s quantum efficiency, $h$ is Plank’s constant, and $\nu$ is the photon’s frequency. However, calculation of the second-order statistics of the impulse-response function are generally computationally intensive, with no known closed-form expressions available [15,26]. As a result, simplifying assumptions are often practiced in the calculation of the photocurrent variance that ignore the randomness in the shape of the impulse-response function. For example, one customary approach is to assume that the impulse-response function takes the simplified form of $I_{p,c}(t) = qGh(t)$ (DDmodel, i.e., the deterministic duration model), where $h(t)$ is a unit-area deterministic function that is proportional to the mean impulse-response function [3]. We emphasize that implicit in such simplifications is the absence of the statistical correlation between the gain and buildup time. For example, the DD-model $I_{p,c}(t) = qGh(t)$ implies that both high-gain and low-gain realizations of $I_{p,c}(t)$ yield the same bandwidth, while in actuality, high-gain realizations are accompanied by a long buildup times, and hence low bandwidth, and vice versa. Nevertheless, the type of simplifications described above admit the customary closed-form expression for the variance of the photocurrent given by $2B_c q\langle G \rangle^2 F \eta P/h\nu$, where $B_c$ is the conventional bandwidth of the APD and $F$ is the APD’s excess noise factor ($F = \langle G^2 \rangle / \langle G \rangle^2$). For example, if the DD-model $I_{p,c}(t) = qGh(t)$ is used, then $B_c = \int_{-\infty}^{\infty} h^2(t) dt / 2(\int_{-\infty}^{\infty} h(t) dt)^2$ [3], which approximates the 3-dB bandwidth of the APD. The point being made here is that $B_c$ is calculated solely from the mean impulse-response function and it does not capture the statistical correlation between the gain and the buildup time.

In contrast, if the exact expression for the photocurrent variance is used [viz., $q(\eta P/h\nu) \times \int_{-\infty}^{\infty} \langle I_p^2(t) \rangle dt$], then we can continue to conveniently express the shot-noise variance in the
customary form of \(2qB_{\text{sneq}}(G)^2F\eta P/h\nu\) provided that we use the correct bandwidth \(B_{\text{sneq}}\), which we term the *shot-noise equivalent bandwidth*. This leads to the definition of \(B_{\text{sneq}}\) as

\[
B_{\text{sneq}} = \frac{\int_{-\infty}^{\infty} \langle I_p^2(t) \rangle \, dt}{2q^2\langle G \rangle^2 F}.
\]

We emphasize that in contrast to \(B_c\), \(B_{\text{sneq}}\) includes the effects of fluctuations in the shape (e.g., gain and duration) of the APD’s random impulse-response function. Moreover, a closed-form estimate of the second moment of the APD’s impulse-response function can be obtained approximately using the parametric stochastic models for the impulse-response function introduced in Section 3. For example, with the RD-rectangular model,

\[
\langle I_p^2(t) \rangle = \sum_{m=1}^{\infty} \int_{t}^{\infty} q^2 m^2 \tau^{-2} \frac{\partial}{\partial \tau} f_{G,T}(m, \tau) \, d\tau, \quad t \geq 0,
\]

and the expression for \(B_{\text{sneq}}\) can be simplified to

\[
B_{\text{sneq}} = \frac{\langle G^2/T \rangle}{2\langle G \rangle^2 F} \equiv \frac{\langle G^2/T \rangle}{2\langle G^2 \rangle},
\]

which can be readily evaluated using our knowledge of the joint distribution \(f_{G,T}\) obtained in Section 2.

The point of the preceding analysis is that to incorporate the effect of statistical correlation between the APD’s gain and buildup time on the signal-to-noise ratio of the photocurrent, we must use \(B_{\text{sneq}}\) in place of \(B_{3\text{dB}}\) in the customary SNR expression \(\eta P/2h\nu FB_{3\text{dB}}\).

We have evaluated the shot-noise-equivalent bandwidth \(B_{\text{sneq}}\) for GaAs APDs, for various widths of the multiplication layer, and compared it with \(B_c\) taken as the 3-dB bandwidth. Figure 3 (upper curve) shows the ratio \(r = B_{\text{sneq}}/B_c\) as a function of the width of the multiplication layer for a fixed mean gain of \(\langle G \rangle = 10\). Notably, the shot-noise-equivalent bandwidth exceeds the 3-dB bandwidth by approximately 30%. Thus, our calculations show that the 3-dB bandwidth, which is customarily used in calculating the shot-noise variance, leads to tangible underestimation of the shot noise variance. This is primarily a result of ignoring the statistical correlation between the buildup time and the gain, or equivalently, it results from ignoring the uncertainty in the shape of the impulse-response function.
4. Receiver Performance Analysis

In this section we investigate the effect of the correlation between the buildup time and the gain on the performance of an APD-based receiver in a direct-detection on-off keying (OOK) optical communication system. Since we are particularly interested in the performance at high transmission speeds, we will assume that the buildup time dominates the RC effects and governs the receiver bandwidth characteristics. We first consider the signal-to-noise ratio (SNR) for the output of an integrate-and-dump receiver in the presence of ISI, Johnson noise, and buildup-time and gain uncertainty of the APD. Subsequently, we will examine the receiver sensitivity at high transmission speeds. Our analysis will ignore optical pulse broadening and dispersion effects.

4.1. Statistics of the photocurrent and the integrate-and-dump output

Consider an arbitrary bit of duration $T_b$ from a stream of optical pulses, and let $I(t)$ denote the photocurrent in such a bit. The photocurrent $I(t)$ contains two components: $I_c(t)$, the component due to photons occurring in the present bit, and $I_{ISI}(t)$, the ISI component, which is due to the photons that had arrived in previous bits. More precisely, if we represent the photon arrival times by the sequence $\{t_i\}$ and assume that the current bit extends from 0 to $T_b$, then for $0 \leq t \leq T_b$, the photocurrent can be written as

$$I(t) = \sum_{t_i < 0} I_{p_i}(t - t_i) + \sum_{0 \leq t_i < t} I_{p_i}(t - t_i)$$

$$\triangleq I_{ISI}(t) + I_C(t), \quad (12)$$

where $I_{p_i}(t - t_i)$ is the impulse-response function induced by a photon absorption at time $t_i$. Note that $I_C(t)$ and $I_{ISI}(t)$ are statistically independent since they correspond to disjoint and independent sets of photon absorptions. Suppose that photos are absorbed at a rate of $\phi(\tau)$ photons per second ($\phi(\tau) = \eta P(\tau)/h\nu$, where $P(\tau)$ is received optical power at time $\tau$). By using standard filtered-shot-noise analysis [24], we can express the mean of $I(t)$ in terms of a convolution as follows:

$$\langle I(t) \rangle = \int_{-\infty}^{0} \langle I_p(t - \tau) \rangle \phi(\tau) \, d\tau + \int_{0}^{t} \langle I_p(t - \tau) \rangle \phi(\tau) \, d\tau$$

$$\triangleq \langle I_{ISI}(t) \rangle + \langle I_C(t) \rangle. \quad (13)$$
To see the maximum ISI effect, we consider the case for which all the bits prior to the current bit are ON, i.e., $\phi(\tau) \equiv \phi_1$, for all $\tau$, where $\phi_1$ is the photon absorption rate in a pulse. It follows that

$$
\langle I_C(t) \rangle = \phi_1 \int_0^t \langle I_p(t - \tau) \rangle d\tau = \phi_1 \int_0^t \langle I_p(\tau) \rangle d\tau \tag{14}
$$

$$
\langle I_{\text{ISI}}(t) \rangle = \phi_1 \int_{-\infty}^0 \langle I_p(t - \tau) \rangle d\tau = \phi_1 \int_t^\infty \langle I_p(\tau) \rangle d\tau. \tag{15}
$$

The photocurrent $I(t)$ is fed into a bit-integrator, which yields the integral of the photocurrent synchronously over each bit of duration $T_b$. The integral is used to detect the information (0 or 1) modulated with the incident optical sequences. We denote the output of the integrate-and-dump unit by $\Gamma = \int_0^{T_b} I(t) dt$, which can be further decomposed (see (12),(13),(14) and (15)) into two parts: $\Gamma = \Gamma_C + \Gamma_{\text{ISI}}$. The first component is the net signal resulting from the photocurrent due to the absorbed photons in the present bit, and the second component corresponds to ISI resulting from the photocurrent due to the absorbed photons in the previous bits. The means of $\Gamma$, $\Gamma_C$ and $\Gamma_{\text{ISI}}$ are respectively

$$
\langle \Gamma_C \rangle = \phi_1 \int_0^{T_b} \int_0^t \langle I_p(t - \tau) \rangle d\tau dt, \tag{16}
$$

$$
\langle \Gamma_{\text{ISI}} \rangle = \phi_1 \int_0^{T_b} \int_{-\infty}^0 \langle I_p(t - \tau) \rangle d\tau dt \tag{17}
$$

and

$$
\langle \Gamma \rangle = \phi_1 \int_0^{T_b} \int_{-\infty}^t \langle I_p(t - \tau) \rangle d\tau dt. \tag{18}
$$

The variance of $\Gamma$ is given by

$$
\sigma_\Gamma^2 = \phi_1 \int_0^{T_b} \int_0^{T_b} \int_{-\infty}^{\mu \land \nu} R_{I_p}(\mu - \tau, \nu - \tau) d\tau d\mu d\nu \tag{19}
$$

where $\mu \land \nu$ is the minimum of $\mu$ and $\nu$, $R_{I_p}(\mu, \nu)$ is the autocorrelation function of the impulse-response $I_p(t)$, which is defined as $R_{I_p}(\mu, \nu) = E[I_p(\mu - \xi)I_p(\nu - \xi)]$. (The derivation of (19) utilizes standard analysis of filtered shot-noise processes [24] and will not be presented here.)

We now define the maximal-ISI SNR for the integrate-and-dump output as $\langle \Gamma_C \rangle^2 / \sigma_\Gamma^2$. In the next subsection, we specialize the SNR to two models for the APD’s impulse-response.
function: RD-rectangular model and DD-exponential model. This will permit us to perform a comparison that shows the effect of the uncertainty in the shape of the impulse-response function on the receiver SNR.

4.2. SNR for a deterministic-duration exponential model

We now use a deterministic-duration exponential (DD-exponential or DD-E) model to approximate the APD’s impulse-response function: \( I_{DD}(t) = gGbe^{-bt}u(t) \) (we ignore the electronic charge). Here, \( b \) is the fixed exponential rate consistent with the 3-dB bandwidth \( B_{3dB} \) of the APD \([15,16]\), namely, \( b = 2\pi B_{3dB} \). Note that in this model the statistical correlation between the buildup time and the gain is ignored. Clearly, the mean of this impulse-response function is \( \langle I_{DD}(t) \rangle = \langle G \rangle be^{-bt}u(t) \), and the autocorrelation function can also be calculated yielding

\[
R_{I_{DD}}(\mu, \nu) = \langle G^2 \rangle b^2 e^{-b(\mu+\nu)}u(\mu)u(\nu).
\]

We now substitute \( \langle I_{DD}(t) \rangle \) and \( R_{I_{DD}}(\mu, \nu) \) into equations (16) and (19), respectively, and obtain

\[
\langle \Gamma_C \rangle = \phi_1 \langle G \rangle [T_b - \frac{1}{b}(1 - e^{-bT_b})]
\]

\[
\sigma_I^2 = \phi_1 \langle G^2 \rangle [T_b - \frac{1}{b}(1 - e^{-bT_b})].
\]

Note that only the marginal statistics for the random variable \( G \) is used in above equations. Finally, we replace \( b \) by \( 2\pi B_{3dB} \) and obtain

\[
\text{SNR}_{DD} = \frac{\phi_1 T_b}{F} \left( \frac{2\pi T_b B_{3dB} - 1 + e^{-2\pi T_b B_{3dB}}}{2\pi T_b B_{3dB}} \right). \tag{20}
\]

Note that the first fraction is the SNR of an instantaneous detectors and the second fraction is a scaling factor accounting for ISI effects.

4.3. SNR for the random-duration rectangular model

As introduced in Section 3, the RD-rectangular parametric model for the random impulse-response function is

\[
I_{RD}(t) = (G/T)\{u(t) - u(t - T)\}.
\]
Since the first moment $\langle I_{RD}(t) \rangle$ and the second moment $\langle I_{RD}^2(t) \rangle$ of impulse-response function (below avalanche breakdown) are known to decay exponentially at the same rate [15], $b_{se}$ is referred to the decay rate (we will see later that $b_{se}$ is related with $B_{sneq}$). we approximate them by

$$\langle I_{RD}(t) \rangle \approx ae^{-b_{se}t} \quad \text{and} \quad \langle I_{RD}^2(t) \rangle \approx ce^{-b_{se}t}. \quad (21)$$

If we substitute these approximations in (16), we obtain

$$\langle \Gamma_C \rangle = \phi \frac{a}{b_{se}} \left( T_b - \frac{1}{b_{se}} \left( 1 - e^{-b_{se}T_b} \right) \right). \quad (22)$$

Next, to calculate the second moment of the photocurrent in (19), we need to first determine the autocorrelation function $R_{I_{RD}}(\mu, \nu) = E[I_{RD}(\mu)I_{RD}(\nu)]$, which can be approximated by (see Appendix A for derivation)

$$R_{I_{RD}}(\mu, \nu) \approx ce^{-b_{se}(\mu + \nu)} \quad (23)$$

If we now substitute the above autocorrelation function into (19), we obtain the expression for the variance of the receiver output

$$\sigma^2_{I_{RD}} \approx \phi_1 \frac{2c}{b_{se}^2} \left[ T_b - \frac{1}{b_{se}} \left( 1 - e^{-b_{se}T_b} \right) \right] \quad (24)$$

and the SNR according to the RD-rectangular model can be cast as

$$\text{SNR}_{RD} = \frac{\langle \Gamma_C \rangle^2}{\sigma^2_{I_{RD}}} = \phi \frac{a^2}{b_{se}^2} \left[ T_b - \frac{1}{b_{se}} \left( 1 - e^{-b_{se}T_b} \right) \right] \cdot \frac{2c}{b_{se}}. \quad (25)$$

Our next step is to explore the physical meaning of the parameters $a, b_{se}$ and $c$ based on our rectangular model. To do so, we note that

$$\int_0^\infty \langle I_{RD}(t) \rangle \ dt \approx \int_0^\infty ae^{-b_{se}t} \ dt = \frac{a}{b_{se}} \quad (26)$$

and

$$\int_0^\infty \langle I_{RD}(t) \rangle \ dt = \int_0^\infty \frac{G}{T} \{u(t) - u(t - T)\} \ dt = \langle G \rangle \quad (27)$$
and use the fact that the left side of the above two equations are equal to conclude that $a/b_{se} \approx \langle G \rangle$. Similarly,

$$
\int_0^\infty \langle I_{RD}^2(t) \rangle \, dt \approx \int_0^\infty ce^{-b_{se}t} \, dt = \frac{c}{b_{se}}
$$

and

$$
\langle \int_0^\infty I_{RD}^2(t) \, dt \rangle = \langle \int_0^\infty \frac{G^2}{T^2} \{u(t) - u(t - T) \, dt\} \rangle = \langle \frac{G^2}{T} \rangle,
$$

which results in $c/b_{se} \approx \langle G^2/T \rangle$.

We now proceed to characterize $b_{se}$. Note that the approximations of $\langle I_{RD} \rangle \approx ae^{-b_{se}t}$ is based on the rectangular parameterized model, which takes into account the correlation of $G$ and $T$. Thus, the effective bandwidth produced by its shape (which is $b_{se}/4$) should be identical to the shot-noise-equivalent bandwidth $B_{sneq}$. We therefore have $b_{se} = 4B_{sneq}$. By substituting the approximation $a/b_{se} \approx \langle G \rangle$ and $b_{se} = 4B_{sneq}$ in (25), we can recast the SNR as

$$
\text{SNR}_{RD} = \frac{\phi_1 \langle G \rangle^2 \left(4T_b B_{sneq} - 1 + e^{-4T_b B_{sneq}}\right)}{\langle \frac{G^2}{T} \rangle^2}.
$$

(30)

Note that the joint statistics of $G$ and $T$ is used in the above equation. Finally, recall that $B_{sneq} = \frac{\langle G^2/T \rangle}{2\langle G \rangle^2}$ (see Section 3.1) and define a correction factor $\kappa = 4B_{sneq}/2\pi B_{3dB}$ to obtain

$$
\text{SNR}_{RD} = \frac{\phi_1 T_b}{F} \left(\frac{2\pi \kappa T_b B_{3dB} - 1 + e^{-2\pi \kappa T_b B_{3dB}}}{2\pi \kappa T_b B_{3dB}}\right).
$$

(31)

In the above formula, first fraction is the SNR for an instantaneous detector for which case the ISI is absent; the second fraction represents the correction factor due to ISI. Note that the statistical correlation between the buildup time and the gain is incorporated through the parameter $B_{sneq}$.

More general expressions for the SNR taking into account the randomness of the sequence of ones and zeros (as in a OOK setting) are derived in Appendix B. Results corresponding to these generalized expressions for the SNR will be in the next subsection.

4.4. Results: ISI and gain–buildup-time correlation effects on the SNR

We now examine the ISI-correction factor of the SNR expression given by (31) (the quantity in parenthesis) and observe that it monotonically varies from 1 to 0 as the product
T_b B_{sneq}, which is a measure of the transmission speed relative to the APD’s buildup time, varies from $\infty$ to 0. Recall that it was shown in Section 3.1 that the ratio $r = B_{sneq}/B_{3dB}$ is approximately 1.3 (for the APD parameters considered), which implies that $\kappa < 1$. Consequently, by comparing the SNR expression given in (20) (which corresponds to the case when the correlation between the gain and the buildup time is ignored) and SNR_{RD}, and by using the monotonicity of the ISI factor we conclude that SNR_{DD} < SNR_{RD}. Namely, the correlation between the gain and buildup time adversely affects the SNR. In what follows we will show that this effect becomes more significant at high transmission speeds.

Figure 6 depicts SNR_{DD} and SNR_{RD} as a function of the transmission rates. The signal-to-noise expressions are given in Appendix B (see (53), (54), (59) and (60)). It is assumed that a GaAs APD is being used with multiplication layer width of $w = 100$nm and an average gain of $\langle G \rangle = 10$. Moreover, the average number of photons absorbed by the APD is assumed to be 1000 per 1 bit, and the Johnson noise parameter is $\sigma_J = 500$ noise-counts per bit. We observe that as the transmission rate increases, the effect of ISI becomes progressively more detrimental to the SNR. It is also shown that the relative separation between SNR_{RD} and SNR increases with the transmission rate.

5. Bit Error Rate and Receiver Sensitivity

In this section we investigate the effect of the statistical correlation between the gain and buildup time on the receiver performance. We use the same APD parameters as those used in the previous section. The APD’s 3-dB bandwidth is found to be $B_{3dB} = 29$ GHz. We first consider the RD-rectangular model. Let $\mu_{0,RD}, \sigma^2_{0,RD}, \mu_{1,RD},$ and $\sigma^2_{1,RD}$ denote the mean and variance corresponding to the cases for which the current bit is 0 and 1, respectively. These parameters are (see Appendix B for derivation):

$$\mu_{0,RD} = \frac{1}{2} \frac{n_0 \langle G \rangle}{\kappa \lambda} (1 - e^{-\kappa \lambda}), \quad (32)$$

$$\sigma^2_{0,RD} = \frac{1}{4} \frac{n_0 \langle G \rangle^2}{\kappa \lambda^2} \left(\frac{1 - e^{-2\kappa \lambda}}{1 - e^{-\kappa \lambda}}\right) + \frac{n_0 \langle G \rangle}{2\kappa \lambda} (1 - e^{-\kappa \lambda} - \kappa \lambda e^{-\kappa \lambda}) + \sigma_J^2, \quad (33)$$

$$\mu_{1,RD} = \mu_{0,RD} + \frac{n_0 \langle G \rangle}{\kappa \lambda} (\kappa \lambda - 1 + e^{-\kappa \lambda}), \quad (34)$$

$$\sigma^2_{1,RD} = \sigma^2_{0,RD} + \frac{n_0 \langle G \rangle^2}{\kappa \lambda} (\kappa \lambda - 2 + 2e^{-\kappa \lambda} + \kappa \lambda e^{-\kappa \lambda}), \quad (35)$$
where \( \lambda = 2\pi T_b B_{3dB} \) and \( \kappa \) is the correction factor defined earlier. Note that \( \mu_{0,\text{RD}} \) and \( \sigma^2_{0,\text{RD}} \) (less \( \sigma_J^2 \)) are entirely due to ISI.

In contrast to the above statistics, the corresponding quantities for the DD exponential model for which the correlation between the gain and the buildup time is ignored can be found to be (we omit the derivation)

\[
\begin{align*}
\mu_{0,\text{DD}} & = \frac{n_0 \langle G \rangle}{2} (1 - e^{-\lambda}), \\
\sigma^2_{0,\text{DD}} & = \frac{n_0^2 \langle G \rangle^2}{4} \frac{(1 - e^{-\lambda})^4}{\lambda^2 (1 - e^{-2\lambda})} + \frac{n_0 \langle G \rangle^2}{\lambda} (1 - e^{-\lambda})^2 + \sigma_J^2, \\
\mu_{1,\text{DD}} & = \mu_{0,\text{DD}} + \frac{n_0 \langle G \rangle}{\lambda} (\lambda - 1 + e^{-\lambda}), \\
\sigma^2_{1,\text{DD}} & = \sigma^2_{0,\text{DD}} + \frac{n_0 \langle G \rangle^2}{2\lambda} (2\lambda - (1 - e^{-\lambda})^2).
\end{align*}
\]

As before, \( \mu_{0,\text{DD}} \) and \( \sigma^2_{0,\text{DD}} \) (less \( \sigma_J^2 \)) are contributions of ISI.

To simplify the analysis, we will approximate the output of the receiver by a Gaussian random variable. The BER is then obtained according to

\[
\text{BER} \approx \frac{1}{2} \left\{ \text{erfc} \left( \frac{\theta - \mu_{0,\text{RD}}}{\sqrt{2}\sigma_{0,\text{RD}}} \right) + \text{erfc} \left( \frac{\mu_{1,\text{RD}} - \theta}{\sqrt{2}\sigma_{1,\text{RD}}} \right) \right\},
\]

where \( \theta \) is the optimal decision threshold. Figure 7 depicts the BER as a function of the transmission rate calculated for both the DD-exponential model and RD-rectangular model. Here, \( n_0 \) is assumed 1000 photons per bit and the Johnson noise is held constant at \( \sigma_J = 500 \) noise-counts per bit. The adverse effect of the correlation between the gain and the buildup time is evident in this plot, as seen by the elevated BER curve in the RD-rectangular model case compared to that obtained for the DD-exponential model. This behavior is also consistent with the SNR plots shown in Fig. 6.

Finally, Fig. 8 shows the receiver sensitivity, defined as the minimum \( n_0 \) needed to achieve a BER of \( 10^{-9} \). As shown in the figure, the entire transmission range can be split into two parts: a Johnson noise limited regime at low transmission speeds and an ISI-limited regime at high transmission speeds. The adverse effect of the correlation between the gain and the buildup time is observed in the ISI-limited regime through the higher sensitivity predicted by the RD-rectangular model compared to that obtained for the DD-exponential model. For instance, we observe that in comparison to the predictions of the RD-rectangular model, the
DD-exponential model results in underestimating the sensitivity by 3.4% at transmission speed of 10 Gbps, by 7.2% at 15 Gbps and by 26% at 20 Gbps. Thus, the effect of neglecting the statistical correlation between build-time and the gain in the receiver performance analysis leads to an underassessment of ISI, which, in turn, leads to an overly optimistic assessment of the receiver performance. Moreover, this trend intensifies progressively as transmission rate increases.

6. Conclusion

The statistical correlation between the random gain and the random avalanche buildup-time in APDs is determined for the first time, and the effect of the correlation on the receiver performance is established. We have shown that as a result of this correlation, the photocurrent SNR is diminished by a factor $r = B_{\text{sneq}}/B_{\text{3dB}}$, where $B_{\text{3dB}}$ is the APD’s 3-dB bandwidth and $B_{\text{sneq}}$ is the shot-noise-equivalent bandwidth, which takes into account the statistical correlation between the gain and the buildup time and defined by $B_{\text{sneq}} = \langle G^2/T \rangle / 2\langle G^2 \rangle$. Thus, $B_{\text{sneq}}$ is precisely the bandwidth that needs to be used in calculating the shot-noise variance when photodetection is achieved by means of an APD. For example, our calculations show that $B_{\text{sneq}}/B_{\text{3dB}}$ is approximately 1.3 for a GaAs APD with a multiplication region in the range 30-250 nm.

Additionally, we have derived compact expressions for the output of an integrate-and-dump receiver in an OOK direct-detection systems that includes the effects of ISI and the statistical correlation between the gain and the buildup time. This expression is similar to its counterpart when the gain-buildup-time correlation is ignored, but it includes a correction factor, $\kappa = 4B_{\text{sneq}}/2\pi B_{\text{3dB}}$, which explicitly quantifies the SNR degradation due to this correlation. Notably, our results predict that this correlation adversely affects ISI noise, and hence receiver sensitivity. Moreover, such adverse effect becomes progressively more significant as the transmission rate increases.

The analysis in this paper is carried out under the simplifying assumptions of a constant electric field in the APD’s multiplication region and a constant drift velocity for carriers. The analysis, however, includes the effects of dead space and utilizes non-localized ionization coefficients. Generalization to non-uniform fields, variable carrier drift speed, and
heterostructure APDs can be carried out in a straightforward fashion using the techniques recently reported in [20,27].
Appendix A: Autocorrelation Function of the Random-Duration Parametric Model

For the RD-rectangular model, the autocorrelation function can be expressed as

$$R_{IRD}(\mu, \nu) = \langle \frac{G}{T} \{u(\mu) - u(\mu - T)\} \frac{G}{T} \{u(\nu) - u(\nu - T)\} \rangle$$

$$= \sum_{m=1}^{\infty} \int_{0}^{\infty} \frac{m^2}{\tau^2} \{u(\mu) - u(\mu - \tau)\}\{u(\nu) - u(\nu - \tau)\} \frac{\partial}{\partial \tau} f_{G,T}(m, \tau) d\tau.$$ 

Note that the integrand is 0 unless $\tau \geq \mu \lor \nu$ (the notation $\mu \lor \nu$ denotes the maximum of $\mu$ and $\nu$). Therefore the above integral can be rewritten compactly as

$$R_{IRD}(\mu, \nu) = \sum_{m=1}^{\infty} \int_{\mu \lor \nu}^{\infty} \frac{m^2}{\tau^2} \frac{\partial}{\partial \tau} f_{G,T}(m, \tau) d\tau.$$ (41)

In particular, the mean of $I_{RD}^2(t)$ is given by

$$\langle I_{RD}^2(t) \rangle = R_{IRD}(t, t).$$ (42)

We now recall the approximation $\langle I_{RD}^2(t) \rangle \approx ce^{-b_{out}t}$ and compare equations (41) and (42), we conclude that $R_{IRD}(\mu, \nu)$ can be approximated by

$$R_{IRD}(\mu, \nu) \approx ce^{-b_{out}(\mu \lor \nu)}.$$ (43)

Appendix B: Statistics of Receiver Output in the Presence of Intersymbol Interference

We assume that zero and one occur with equal probability in the transmitted binary sequence $I_b(n)$, and for convenience, we extend $I_b(n)$ to $I_b(t)$ (where $t$ is a continuous variable) by

$$I_b(t) = I_b(n) \quad \text{if } nT_b \leq t < (n + 1)T_b,$$
and the (absorbed) optical signal at the receiver can be represented by \( \phi(t) = \phi_1 I_b(t) \). Note that the covariance function and the mean of \( I_b(t) \) are

\[
\sigma^2_{I_b}(\xi_1, \xi_2) = \begin{cases} 
\frac{1}{4} & \text{if } nT_b \leq \xi_1, \xi_2 < (n + 1)T_b, \text{ for } n = -1, -2, -3, \ldots \\
0 & \text{otherwise,} \\
\frac{1}{2} & \text{if } t \leq 0 
\end{cases}
\]

\[
\langle I_b(t) \rangle = \begin{cases} 
0 & \text{if } t > 0 \text{ and current bit is 0} \\
1 & \text{if } t > 0 \text{ and current bit is 1.}
\end{cases}
\]

The mean value of the receiver output can be calculated by

\[
\langle \Gamma \rangle = \phi_1 \int_0^{T_b} \int_{-\infty}^t \langle I_p(t - \xi) \rangle \langle I_b(\xi) \rangle d\xi dt.
\]

If the current bit being sent is zero, then \( \langle I_b(\xi) \rangle = 0 \) for \( 0 \leq \xi < T_b \), and

\[
\mu_0 = \frac{\phi_1}{2} \int_0^{T_b} \int_{-\infty}^0 \langle I_p(t - \xi) \rangle d\xi dt. \tag{44}
\]

On the other hand, if the current bit being sent is one, then \( \langle I_b(\xi) \rangle = 1 \) for \( 0 \leq \xi < T_b \), and

\[
\mu_1 = \frac{\phi_1}{2} \int_0^{T_b} \int_{-\infty}^0 \langle I_p(t - \xi) \rangle d\xi dt + \phi_1 \int_0^{T_b} \int_{0}^t \langle I_p(t - \xi) \rangle d\xi dt
\]

\[
= \mu_0 + \phi_1 \int_0^{T_b} \int_0^t \langle I_p(t - \xi) \rangle d\xi dt. \tag{45}
\]

To calculate the variance of \( \Gamma \), we will utilize standard analysis of filtered shot-noise processes [24] while considering the stochastic nature of the transmitted signal \( I_b(t) \). This yields

\[
\langle \Gamma^2 \rangle = \int_0^{T_b} \int_0^{T_b} \int_{-\infty}^{\mu} \int_{-\infty}^{\nu} \langle I_p(\mu - \xi_1) \phi_1(\xi_1) I_p(\nu - \xi_2) \phi_1(\xi_2) \rangle \, d\xi_1 d\xi_2 d\mu d\nu
\]

\[
+ \int_0^{T_b} \int_0^{T_b} \int_{-\infty}^{\mu} \int_{-\infty}^{\nu} \langle I_p(\mu - \xi) I_p(\nu - \xi) \phi_1(\xi) \rangle \, d\xi d\mu d\nu
\]

\[
= \phi_1^2 \int_0^{T_b} \int_0^{T_b} \int_{-\infty}^{\mu} \int_{-\infty}^{\nu} \langle I_p(\mu - \xi_1) \rangle \langle I_p(\nu - \xi_2) \rangle \langle I_b(\xi_1) I_b(\xi_2) \rangle \, d\xi_1 d\xi_2 d\mu d\nu
\]

\[
+ \phi_1 \int_0^{T_b} \int_0^{T_b} \int_{-\infty}^{\mu} \int_{-\infty}^{\nu} \langle I_p(\mu - \xi) \rangle \langle I_p(\nu - \xi) \rangle \langle I_b(\xi) \rangle \, d\xi d\mu d\nu.
\]
Since \( \langle I_b(\xi_1)I_b(\xi_2) \rangle = \langle I_b(\xi_1) \rangle \langle I_b(\xi_2) \rangle + \sigma_{I_b}^2(\xi_1, \xi_2) \), and since the set \( \{ (\xi_1, \xi_2) : \xi_1 \neq \xi_2 \} \) has zero area, we obtain

\[
\sigma_{I}^2 = \frac{\phi_1^2}{4} \int_0^{T_b} \int_0^{T_b} \int_{-\infty}^{\mu} \int_{-\infty}^{\nu} \langle I_p(\mu - \xi_1) \rangle \langle I_p(\nu - \xi_2) \rangle \sigma_{I_b}(\xi_1, \xi_2) \, d\xi_1 d\xi_2 d\mu d\nu \\
+ \phi_1 \int_0^{T_b} \int_0^{T_b} \int_{-\infty}^{\mu} \int_{-\infty}^{\nu} R_{I_p}(\mu - \xi, \nu - \xi) \langle I_b(\xi) \rangle \, d\xi d\mu d\nu.
\]  

(46)

We first consider the case for which the current bit is 0; in this case,

\[
\sigma_0^2 = \frac{\phi_1^2}{4} \int_0^{T_b} \int_0^{T_b} \int_{-\infty}^{\mu} \int_{-\infty}^{\nu} \langle I_p(\mu - \xi_1) \rangle \langle I_p(\nu - \xi_2) \rangle \, d\xi_1 d\xi_2 d\mu d\nu \\
+ \phi_1 \int_0^{T_b} \int_0^{T_b} \int_{-\infty}^{\mu} \int_{-\infty}^{\nu} R_{I_p}(\mu - \xi, \nu - \xi) \, d\xi d\mu d\nu,
\]  

(47)

where \( \mathbb{D} = \{ (\xi_1, \xi_2) : nT_b \leq \xi_1, \xi_2 < (n+1)T_b, n = -1, -2, -3, \ldots \} \).

Alternatively, if the current bit is 1, then \( \sigma_{I_b}^2(\xi_1, \xi_2) = 0 \) for \( 0 \leq \xi_1, \xi_2 < T_b \), and we obtain

\[
\sigma_1^2 = \frac{\phi_1^2}{4} \int_0^{T_b} \int_0^{T_b} \int_{-\infty}^{\mu} \int_{-\infty}^{\nu} \langle I_p(\mu - \xi_1) \rangle \langle I_p(\nu - \xi_2) \rangle \, d\xi_1 d\xi_2 d\mu d\nu \\
+ \phi_1 \int_0^{T_b} \int_0^{T_b} \int_{-\infty}^{\mu} \int_{-\infty}^{\nu} R_{I_p}(\mu - \xi, \nu - \xi) \, d\xi d\mu d\nu \\
+ \phi_1 \int_0^{T_b} \int_0^{T_b} \int_{-\infty}^{\mu} \int_{-\infty}^{\nu} R_{I_p}(\mu - \xi, \nu - \xi) \, d\xi d\mu d\nu \\
= \sigma_0^2 + \phi_1 \int_0^{T_b} \int_0^{T_b} \int_{-\infty}^{\mu} \int_{-\infty}^{\nu} R_{I_p}(\mu - \xi, \nu - \xi) \, d\xi d\mu d\nu.
\]  

(48)

We next specialize these results to the DD-exponential model and RD-rectangular models. For the DD-exponential model, we substitute the impulse-response function \( I_{DD}(t) = qGbe^{-bt} \) and the autocorrelation function \( R_{I_{DD}}(\mu, \nu) = (G^2)b^2e^{-b(\mu+\nu)} \) into (44), (45), (47) and (48), use \( b = 2\pi B_{3dB} \) and define \( \lambda = 2\pi T_b B_{3dB} \) to obtain

\[
\mu_{0,DD} = \frac{1}{2} \frac{n_0(G)}{\lambda} (1 - e^{-\lambda}) \quad (49)
\]

\[
\sigma_{0,DD}^2 = \frac{1}{4} \frac{n_0^2(G)^2}{\lambda^2} (1 - e^{-\lambda})^4 + \frac{n_0(G^2)}{\lambda} (1 - e^{-\lambda})^2 + \sigma_J^2 \quad (50)
\]

\[
\mu_{1,DD} = \mu_{0,DD} + \frac{n_0(G)}{\lambda} (\lambda - 1 + e^{-\lambda}) \quad (51)
\]

\[
\sigma_{1,DD}^2 = \sigma_{0,DD}^2 + \frac{n_0(G^2)}{2\lambda} (2\lambda - (1 - e^{-\lambda})^2), \quad (52)
\]
where, $n_0$ is the average number of absorbed photons per 1 bit ($\phi_1 = n_0/T_b$) and $\sigma_j$ is the Johnson noise parameter. The general expression of SNR and ISI for this RD-exponential model are given by

$$\text{SNR}_{DD} = \frac{(\mu_{1,DD} - \mu_{0,DD})^2}{\sigma^2_{1,DD}}$$  \hspace{1cm} (53)

$$\text{ISI}_{DD} = \sigma^2_{0,DD} - \sigma^2_j.$$  \hspace{1cm} (54)

For the RD-rectangular model, in particular, we substitute the approximations (21), (43), $a/b_{se} \approx \langle G \rangle$ and $c/b_{se} \approx (G^2/T)$ into (44), (45), (47) and (48), use $b_{se} = 4B_{sneq}$, $\kappa = 4B_{sneq}/2\pi B_{3dB}$ and $\lambda = 2\pi T_b B_{3dB}$, and obtain

$$\mu_{0,\text{RD}} = \frac{1}{2} \frac{n_0 \langle G \rangle}{\kappa \lambda}(1 - e^{-\kappa \lambda})$$  \hspace{1cm} (55)

$$\sigma^2_{0,\text{RD}} = \frac{1}{4} \frac{n_0 \langle G \rangle^2}{\kappa \lambda^2} \frac{(1 - e^{-\kappa \lambda})^2}{1 - e^{-4\pi \kappa T_b B_{3dB}}} + \frac{n_0 \langle G^2 \rangle}{2\kappa \lambda} (1 - e^{-\kappa \lambda} - \kappa \lambda e^{-\kappa \lambda}) + \sigma^2_j$$  \hspace{1cm} (56)

$$\mu_{1,\text{RD}} = \mu_{0,\text{RD}} + \frac{n_0 \langle G \rangle}{\kappa \lambda} (\kappa \lambda - 1 + e^{-\kappa \lambda})$$  \hspace{1cm} (57)

$$\sigma^2_{1,\text{RD}} = \sigma^2_{0,\text{RD}} + \frac{n_0 \langle G^2 \rangle}{\kappa \lambda} (\kappa \lambda - 2 + 2e^{-\kappa \lambda} + \kappa \lambda e^{-\kappa \lambda}).$$  \hspace{1cm} (58)

Finally, the general expressions of SNR and ISI for this RD-rectangular model are given by

$$\text{SNR}_{\text{RD}} = \frac{(\mu_{1,\text{RD}} - \mu_{0,\text{RD}})^2}{\sigma^2_{1,\text{RD}}}$$  \hspace{1cm} (59)

$$\text{ISI}_{\text{RD}} = \sigma^2_{0,\text{RD}} - \sigma^2_j.$$  \hspace{1cm} (60)
References


Fig. 1. Joint probability distribution function (PDF) of the random gain $G$ and the random buildup time $T$ for a GaAs APD with a 160-nm multiplication layer. The applied electric field in the multiplication region is taken as $5.47 \times 10^3$ kV/cm, yielding a theoretical average gain of $\langle G \rangle = 10.46$.

Fig. 2. Joint probability density function (pdf) of the random gain $G$ and the random buildup time $T$ for the APD considered in Fig. 1. In order to better show the details of the pdf, large peaks have been truncated.

Fig. 3. Bottom: The correlation coefficient $\rho$ of the random gain $G$ and the random buildup time $T$ for a GaAs APD as a function of the multiplication-layer width $w$. Top: the ratio $r$ of shot-noise-equivalent bandwidth $B_{\text{neq}}$ to the 3-dB bandwidth $B_{\text{3dB}}$. The mean gain is held constant at $\langle G \rangle = 10$.

Fig. 4. Models for impulse-response function: the rectangular random-duration (RD-R) model (solid curve) and the triangular random-duration (RD-T) model (dotted curve). Note that the area under each of the models is $qG$. A deterministic-duration exponential (DD-E) is also shown for which the decay rate is deterministic.

Fig. 5. Mean impulse-response function $i_p(t)$ corresponding to the RD-model with a rectangular shape (solid curve), the RD-model with a triangular shape (dotted curve), and the exact model [15] (dashed curve). Note that the rectangular-shape RD-model slightly underestimates the bandwidth while the triangular-shaped RD-model overestimates it.

Fig. 6. Signal-to-noise ratio of the output of the integrate-and-dump receiver as a function of the digital transmission speed for a GaAs APD with a 100-nm multiplication layer and a theoretical mean gain of 10. The solid curve corresponds to the case where the RD-rectangular model is used for the APD’s impulse-response function, while the the dashed curve corresponds to the DDexponential model. The average number of photons is fixed to 1000 photons per bit and the Johnson noise is selected so that $\sigma_J = 500$ noise-counts per bit.

Fig. 7. Bit-error rate as a function of the transmission speed using the RDrectangular model (solid curve) and the DDexponential model (dashed curve). The mean gain of the APD is assumed as 10 and the average number of photons per bit is fixed to 1000 photons per bit.
The Johnson noise is selected so that $\sigma_J = 500$ noise-counts per bit.

**Fig. 8.** Receiver sensitivity as a function of the transmission speed using the RD rectangular model (solid curve) and the DD exponential model (dashed curve). The parameters of the APD and Johnson noise are the same as Fig. 7.
Figure 1.
Figure 2.
Figure 3.
Figure 4.
Figure 5.
Figure 6.
Figure 7.
Figure 8.