Photon Detection Characteristics and Error Performance of SPAD Array Optical Receivers

Elham Sarbazi, Majid Safari and Harald Haas
Li-Fi Research and Development Centre, Institute for Digital Communications
The University of Edinburgh, Edinburgh, EH9 3JL, UK.
Email: {e.sarbazi, majid.safari, h.haas}@ed.ac.uk

Abstract—In this paper a novel photon counting receiver for optical communication applications is proposed. The proposed receiver is a single photon avalanche diode (SPAD) array which can provide a significantly improved detection sensitivity compared to conventional photodiodes. First, the detection statistics and main characteristics of a single SPAD receiver is presented, and the effects of the SPAD dead time, which is introduced by the quenching process, on the counting probability and effective count rate are studied. The approach is then extended to account for SPAD arrays. Using a Gaussian approximation, the counting distribution of a large size SPAD array is derived and effective count rate of arrays with different sizes is evaluated and compared. It is found that even in SPAD arrays, dead time still has a significant role in the maximum achievable count rate, and the fill factor of the array greatly affects the performance and count rate and has to be carefully dealt with. The impact of SPAD background counts and fill factor on the error performance of an on-off keying (OOK) modulation optical communication system is also investigated. It is shown that the bit error rate (BER) depends critically on background counts and improves with increasing fill factor.

Keywords—Single Photon Avalanche Diode (SPAD), SPAD arrays, photon counting, optical receivers, on-off keying (OOK).

I. INTRODUCTION

Visible light communications (VLC) has recently been an area of interest, and new devices have been proposed as potential transmitters and/or receivers for VLC systems. There has been significant progress towards the realization of optical receivers fully integrated with the standard digital CMOS technology. Recent trends towards integrated CMOS high-speed optical receivers have specially employed avalanche photodiodes (APDs), but the maximum achievable gain of an APD is limited due to low sensitivity and the gain-dependent excess noise. This requires the use of intricate high gain transimpedance amplifiers (TIAs), limiting amplifiers (LA) and adaptive equalizers.

To address these challenges, APDs can be used in the so-called ‘Geiger mode’ as single photon avalanche diodes (SPADs). In Geiger mode, the SPAD is biased beyond its breakdown voltage. As a result, due to the high electric field, the absorption of a single photon will initiate an avalanche of charge carriers which leads to a large internal gain. The extremely high gain allows single photon events to be detected effectively and single-photon detection sensitivity can then be improved. However, after each photon detection event, the SPAD needs to be quenched to recover from the excess charge carriers. This quenching process introduces a finite recovery time, known as the ‘dead time’, during which the device does not respond to another incident photon. Two approaches can be followed to recover the SPAD after a successful detection: passive quenching and active quenching. In general, passively quenched circuits show an extended or paralyzable dead time behavior, whereas active quenching generates a short constant or nonparalyzable dead time. In a SPAD device with nonparalyzable deadtime, any photon arriving during the dead time is neither counted nor has any influence on the dead time duration; while for the paralyzable case, any photon arrival occurring during the dead time is not counted but is assumed to extend the dead time period [1].

Various types of SPADs have been successfully employed in a number of applications, including three-dimensional imaging [2], quantum key distribution [3], and deep space laser communications [4]. The high sensitivity and time resolution of SPADs have recently highlighted the potential of employing SPADs as photon counting receivers for VLC systems [5]–[8]. They can be used with the long term aim of power efficient and highly sensitive receivers and are particularly attractive because they are able to closely approach quantum-limited sensitivity in the detection of weak optical signals in long distance communications, such as in the gas extraction industry, or in downhole monitoring systems [9].

Nevertheless, to the best of authors’ knowledge, there is limited published research on the detection statistics of SPAD receivers in literature. In [10], we presented a thorough characterization and detailed analysis of detection statistics of a single SPAD with nonparalyzable dead time, operating as an optical receiver. In this paper, we extend our previous approach and characterize an array of SPADs for optical communication applications. In particular, analytical modelling and simulation results are provided which predict the performance of a SPAD-based array receiver. Throughout this paper, a SPAD device with a nonparalyzable dead time is considered.

The rest of the paper is organized as follows. The photocount statistics and count rate of a single SPAD and a SPAD array are discussed in Section II, and how the count statistics and effective count rate are affected by SPAD dead time is explained. An approximate mathematical model for the count probability of a SPAD array is developed and Monte Carlo methods are employed to verify the validity of the analytical models. Furthermore, the major constraints which limit the achievable count rate of SPAD receivers are addressed. The mathematical counting distribution of the SPAD array derived in Section II is then used in Section III to predict the error performance of an on-off keying (OOK) modulation optical system. Finally, concluding remarks are given in Section IV.
II. PHOTOCOUNT STATISTICS

A. Single SPAD

1) Probability Mass Function (PMF): In the absence of SPAD dead time, the detection of photon arrival events can be modeled as a Poisson arrival process for which the probability of detecting $k$ photons over a time period of $[0, T_b]$ is given by:

$$p_k = \frac{(\lambda T_b)^k e^{-\lambda T_b}}{k!},$$

where the constant $\lambda$ is the average photon arrival rate, hence, $\lambda T_b$ is the average number of photons arriving at the SPAD during the observation time of $T_b$ seconds. The photocount rate $\lambda$ is related to the power of the optical signal by:

$$\lambda = \frac{n_{QE} P_s}{h \nu},$$

where $n_{QE}$ is the quantum efficiency of SPAD; $P_s$ denotes the power of the incident optical signal; $h$ is the Planck’s constant; and $\nu$ represents the frequency of the optical signal.

When the SPAD dead time is considered, however, the actual count statistics can be very different from the photon arrival statistics and the photon counts are no longer Poisson distributed. Any incident photons which arrive after the initial photon event and before the end of the quenching process, go undetected. In this study, a SPAD detector with constant dead time is considered which cannot record counts for a time interval of fixed duration, $\tau$, immediately following the registration of a count. It is assumed that SPAD is ready to operate at the beginning of the counting interval of $[0, T_b]$. Therefore, the maximum observable count during this period is $k_{\text{max}} = \lfloor \delta + 1 \rfloor$, where $\delta = T_b/\tau$ and $\lfloor x \rfloor$ denotes the largest integer that is smaller than $x$. In [10], the detection statistics of a single SPAD was investigated and it was shown that the probability of $k$ photons being detected during the time interval of $[0, T_b]$ is given by [10]:

$$p_K(k) = \begin{cases} \sum_{i=0}^{k} \psi(i, \lambda_k) - \sum_{i=0}^{k-1} \psi(i, \lambda_{k-1}) & k < k_{\text{max}} \\ 1 - \sum_{i=0}^{k-1} \psi(i, \lambda_{k-1}) & k = k_{\text{max}} \\ 0 & k > k_{\text{max}} \end{cases}$$

where $\lambda_k = \lambda(T_b - k\tau)$, $\lambda_{k-1} = \lambda(T_b - (k-1)\tau)$ and the function $\psi(i, \lambda)$ is defined as:

$$\psi(i, \lambda) = \frac{\lambda^i e^{-\lambda}}{i!}.$$

2) First and second moments: The mean and variance of the photocount distribution in (3) are:

$$\mu_K = k_{\text{max}} \sum_{k=0}^{k_{\text{max}}-1} \sum_{i=0}^{k} \psi(i, \lambda_k),$$

$$\sigma_K^2 = k_{\text{max}} \sum_{k=0}^{k_{\text{max}}-1} \sum_{i=0}^{k} (2k - 2k - 1) \psi(i, \lambda_k)$$

$$- \left( \sum_{k=0}^{k_{\text{max}}-1} \sum_{i=0}^{k} \psi(i, \lambda_k) \right)^2.$$

B. Single SPAD with constant dead time

In the absence of any background noise and afterpulsing, the photocount distribution in (3) may be approximated as

$$p_K(k) \approx \psi(k; \lambda_k),$$

for $k \leq k_{\text{max}}$.

3) Approximation of PMF for large mean counts: In the case where the mean count is large, it can be shown that the count distribution of (3) may be approximated as

$$p_K(k) \approx \psi(k; \lambda_k),$$

for $k \leq k_{\text{max}}$.

4) Effective count rate: In a SPAD-based receiver, background noise and dead time losses limit the minimum and/or maximum achievable count rate. While dead time gives restrictions on the highest measurable count rate, noise is the limitation in the low count rate region. The maximum count rate of commercial SPADs is restricted to a few MHz, due to the slow recharging process, also called ‘quenching process’, after a detection event, and it is also affected by afterpulsing, which is an additional source of counting errors and refers to avalanche events that originate from the emission of carriers that were trapped in the multiplication region during previous avalanche events.

SPADs can be considered as a new generation of Geiger-Muller (GM) detectors which have been widely studied in published research [11], [12]. Provided that all the device characteristics concerning noise and afterpulsing are taken into account, a SPAD with constant dead time can be treated as a nonparalyzable GM counter, in which any photon arriving during the dead time is neither counted nor has any influence on the dead time duration. According to the nonparalyzable dead time count rate model, the relationship between the true counting rate (i.e. photon rate), $\lambda$, and the effective count rate (i.e. observed rate), $\lambda'$, is given by [11]:

$$\lambda' = \frac{\lambda}{1 + \lambda \tau_{\text{nonp}}},$$

where $\tau_{\text{nonp}}$ is the nonparalyzable dead time. Note that assuming a time interval of $T_b$ seconds, the maximum predicted count rate for the nonparalyzable case would be $1/\tau_{\text{nonp}}$, which is termed ‘saturation count rate’, meaning that a SPAD is not able to reach count rates higher than this value.
B. SPAD Array

To increase the capacity of the photon counts, an array of SPADs may be considered which outputs the superposition of the photon counts from the individual SPADs. Other than the dead time of the single SPADs, the Fill Factor (FF) of the SPAD array affects the photocount distribution. FF is defined as the ratio of the SPAD total active area to the total array area and it represents the probability that the incoming photons hit the active area.

Figure 1 illustrates the configuration of a rectangular SPAD array consisting of $R \times C$ single SPADs as the cell element of the array. The FF coefficient of this array is given by:

$$C_{FF} = \frac{lw}{(l+g)(w+g)}.$$  (9)

Array elements are indexed with the subscripts $mn$, where $1 \leq m \leq R$ and $1 \leq n \leq C$, to denote their position within the array.

1) Probability distribution, mean and variance: In the absence of dead time, Poisson counting process will be observed at each element of the array. However, when dead time is present and the effect of FF is considered, the PMF in (3) can be rewritten for the $mn$th element of the array as $p_K(k_{mn})$ with parameters

$$k_{max,mn} = \left\lfloor \frac{T_b}{\tau_{mn}} \right\rfloor + 1,$$

$$\lambda'_{mn} = C_{FF} \lambda_{mn}(T_b - k_{mn}\tau_{mn}),$$

where $\lambda_{mn}$ is the average photon arrival rate at $mn$th SPAD and $\tau_{mn}$ is the dead time of the $mn$th element.

Assuming independent statistics for each SPAD in the array, the joint sample function density of the SPAD array can be described as:

$$P(t(n)) = \prod_{m=1}^{R} \prod_{n=1}^{C} p_K(k_{mn}),$$  (10)

where $n \equiv [k_{11}, k_{12}, ..., k_{R(C-1)}, k_{RC}]$.

Considering independent random variables, $K_{mn}$, as the number of photon counts at $mn$th element of the array, a new random variable can be defined as:

$$X = \sum_{m=1}^{R} \sum_{n=1}^{C} K_{mn}.$$  (11)

Therefore, the probability distribution of $X$ is expressed as:

$$p_X(x) = \sum_{k_{11}} \sum_{k_{12}} ... \sum_{k_{R(C-1)}} Pr(n'),$$  (13)

where $n' \equiv [k_{11}, k_{12}, ..., k_{R(C-1)}, x - \sum_{m=1}^{R} \sum_{n=1}^{C} k_{mn}]$.

It is in general challenging to obtain a closed-form expression for (13), nevertheless, an approximate expression for $p_X(x)$ can be obtained when the number of array elements is large. In that case, according to Central Limit Theorem (CLT), the dead time modified counting distribution of a SPAD array can be approximated by a Gaussian distribution:

$$p_X(x) \sim \mathcal{N}(\mu_X, \sigma_X^2),$$  (14)

where

$$\mu_X = \sum_{m=1}^{R} \sum_{n=1}^{C} \mu_{mn},$$

$$\sigma_X^2 = \sum_{m=1}^{R} \sum_{n=1}^{C} \sigma_{mn}^2.$$  

Here, $\mu_{mn}$ and $\sigma_{mn}^2$ are the mean and variance of the photocount distribution of the $mn$th SPAD in the array.

The exact counting distribution in (13), calculated using numerical methods, and the approximate counting distribution obtained in (14) are plotted in Fig. 2 and compared with the Monte Carlo simulation results. In Fig. 3, (14) is plotted for different values of $\delta = \tau/T_b$. As shown, the Monte Carlo simulation results and the Gaussian approximation are perfectly matched and this confirms the validity of the approximation approach. Also note that as the dead time increases, both the mean and variance of the photon counts decrease and this is in total agreement with the analytical approximations.
model in (8) can be modified for a SPAD array of size \( N \) elements increases. Assuming identical array elements and the saturation level of the count rate.

Furthermore, note that throughout this paper, in all simulation results, practical values for SPAD parameters are assumed which are all adopted from [6] and [7].

2) **Effective count rate:** For a SPAD array, the achievable count rate is expected to be improved as the number of array elements increases. Assuming identical array elements and constant photon arrival rate, the nonparalyzable count rate model in (8) can be modified for a SPAD array of size \( R \times C \) elements as:

\[
    \lambda' = \frac{\lambda RC}{1 + \lambda \tau_{\text{nonp}}}.
\]

The comparison between the Monte Carlo simulations and the above dead time modified count rate model for a SPAD array, is given in Fig. 4 where the observed count rate for arrays of different sizes are compared. According to these curves, the saturation count rates are scaled by the size of array compared to a single SPAD. Also note that the dead time has a significant effect on the maximum achievable count rate and it determines the saturation level of the count rate.

### III. PERFORMANCE EVALUATION

SPADs can be used as photon counting receivers in optical communication systems. In a photon counting receiver the dominant noise source is the background counts which mainly arises from dark counts, afterpulsing and ambient light, and will determine the achievable BER. However, the dead time is another limiting factor for the performance of any SPAD-based receiver. In the following, the effect of background counts and the dead time on the performance of a SPAD-based receiver is investigated and the bit error probability for OOK modulation is provided.

In OOK modulation each bit is transmitted by either pulsing the light source on or off during each bit time interval, say \( T_b \), seconds duration, so that one data bit is sent every \( T_b \) seconds. Hence the system transmits at the bit rate \( B = 1/T_b \).

Assuming \( \lambda_s \) and \( \lambda_b \) as the average photon arrival rates from source and background noise, respectively, \( K_s = \lambda_s T_b \) and \( K_b = \lambda_b T_b \) are the contributions to the average count from the signal and background noise counts per bit interval \( T_b \) for each array element. When a “0” bit is transmitted, the average number of photons impinging on each single SPAD per bit time interval is \( K_s \), and when a “1” bit is transmitted, the average number of received photons per bit time interval is \( K_s + K_b \). Therefore, according to (13), \( p_0(x) \) and \( p_1(x) \), the probability that exactly \( x \) photons are counted by the SPAD array in the counting interval of \( T_b \) seconds, when “0” or “1” are sent, respectively, are given by:

\[
    p_0(x) = p(x; \lambda_s) ,
    \quad p_1(x) = p(x; \lambda_s + \lambda_b) .
\]

In this system, decoding is simply achieved by a threshold decoding process. The number of counted photons is compared with a threshold \( x_T \). A decoding error will occur if \( x < x_T \) when a “1” bit is sent, or if \( x > x_T \), when a “0” bit is sent. Hence the probability of error for equally likely bits is [13]:

\[
    P_e = \frac{1}{2} \Pr \{ x > x_T | 0 \} + \frac{1}{2} \Pr \{ x < x_T | 1 \} .
\]

Considering the count probabilities in (16):

\[
    P_e = \frac{1}{2} \sum_{x=x_T+1}^{\infty} p_0(x) + \frac{1}{2} \sum_{x=0}^{x_T} p_1(x) .
\]

In order to calculate the probability of error in (18), the Gaussian approximation in (14) can be applied to \( p_0(x) \) and \( p_1(x) \) so that \( p_0(x) \sim N(\mu_0, \sigma_0^2) \) and \( p_1(x) \sim N(\mu_1, \sigma_1^2) \). Note that the array size is assumed to be sufficiently large, hence, this approximation is valid. Therefore, \( P_e \) can be approximated as:

\[
    P_e \approx \frac{1}{2} \int_{x_T}^{\infty} p_0(x) \, dx + \frac{1}{2} \int_{0}^{x_T} p_1(x) \, dx
    = \frac{1}{2} Q \left( \frac{x_T - \mu_0}{\sigma_0} \right) + \frac{1}{2} Q \left( \frac{\mu_1 - x_T}{\sigma_1} \right).
\]

where, \( Q(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} \, dt \) is the Q-function. The error probability, \( P_e \), highly depends on \( x_T \) which can be selected to yield the lowest probability of occurring an error. This occurs at the value of \( x_T \) where \( dp_e/dx_T = 0 \). It can be shown that the threshold value \( x_T \) which minimizes \( P_e \) is given by (20) which can be further approximated as:

\[
    x_T = \frac{\mu_1 \sigma_0 + \mu_0 \sigma_1}{\sigma_0 + \sigma_1} .
\]

When this threshold is used, the resulting \( P_e \) in (19) is simplified to:

\[
    P_e \approx Q \left( \frac{\mu_1 - \mu_0}{\sigma_1 + \sigma_0} \right) .
\]

In OOK modulation each bit is transmitted by either pulsing the light source on or off during each bit time interval, say \( T_b \), seconds duration, so that one data bit is sent every \( T_b \) seconds. Hence the system transmits at the bit rate \( B = 1/T_b \).

Assuming \( \lambda_s \) and \( \lambda_b \) as the average photon arrival rates from source and background noise, respectively, \( K_s = \lambda_s T_b \) and \( K_b = \lambda_b T_b \) are the contributions to the average count from the signal and background noise counts per bit interval \( T_b \) for each array element. When a “0” bit is transmitted, the average number of photons impinging on each single SPAD per bit time interval is \( K_s \), and when a “1” bit is transmitted, the average number of received photons per bit time interval is \( K_s + K_b \). Therefore, according to (13), \( p_0(x) \) and \( p_1(x) \), the probability that exactly \( x \) photons are counted by the SPAD array in the counting interval of \( T_b \) seconds, when “0” or “1” are sent, respectively, are given by:

\[
    p_0(x) = p(x; \lambda_s) ,
    \quad p_1(x) = p(x; \lambda_s + \lambda_b) .
\]

In this system, decoding is simply achieved by a threshold decoding process. The number of counted photons is compared with a threshold \( x_T \). A decoding error will occur if \( x < x_T \) when a “1” bit is sent, or if \( x > x_T \), when a “0” bit is sent. Hence the probability of error for equally likely bits is [13]:

\[
    P_e = \frac{1}{2} \Pr \{ x > x_T | 0 \} + \frac{1}{2} \Pr \{ x < x_T | 1 \} .
\]

Considering the count probabilities in (16):

\[
    P_e = \frac{1}{2} \sum_{x=x_T+1}^{\infty} p_0(x) + \frac{1}{2} \sum_{x=0}^{x_T} p_1(x) .
\]

In order to calculate the probability of error in (18), the Gaussian approximation in (14) can be applied to \( p_0(x) \) and \( p_1(x) \) so that \( p_0(x) \sim N(\mu_0, \sigma_0^2) \) and \( p_1(x) \sim N(\mu_1, \sigma_1^2) \). Note that the array size is assumed to be sufficiently large, hence, this approximation is valid. Therefore, \( P_e \) can be approximated as:

\[
    P_e \approx \frac{1}{2} \int_{x_T}^{\infty} p_0(x) \, dx + \frac{1}{2} \int_{0}^{x_T} p_1(x) \, dx
    = \frac{1}{2} Q \left( \frac{x_T - \mu_0}{\sigma_0} \right) + \frac{1}{2} Q \left( \frac{\mu_1 - x_T}{\sigma_1} \right).
\]

where, \( Q(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} \, dt \) is the Q-function. The error probability, \( P_e \), highly depends on \( x_T \) which can be selected to yield the lowest probability of occurring an error. This occurs at the value of \( x_T \) where \( dp_e/dx_T = 0 \). It can be shown that the threshold value \( x_T \) which minimizes \( P_e \) is given by (20) which can be further approximated as:

\[
    x_T = \frac{\mu_1 \sigma_0 + \mu_0 \sigma_1}{\sigma_0 + \sigma_1} .
\]

When this threshold is used, the resulting \( P_e \) in (19) is simplified to:

\[
    P_e \approx Q \left( \frac{\mu_1 - \mu_0}{\sigma_1 + \sigma_0} \right) .
\]
Note that the assumption of a Gaussian distribution, $P_e$ depends only on the difference of the photodetected mean values. Thus any contribution to both means, such as from dark current or background noise, would not effect the values. Thus any contribution to both means, such as from $P$ and $\mu_0$ term, these will however contribute to the variances. Defining the signal-to-noise-ratio (SNR) as:

$$SNR = \frac{(\mu_1 - \mu_0)^2}{(\sigma_1 + \sigma_0)^2}$$  \hfill (23)

$P_e$ can also be written as:

$$P_e = Q\left(\sqrt{SNR}\right).$$  \hfill (24)

The probability of error given in (22) is evaluated and compared with simulation results in Fig. 5, using the threshold obtained in (21). Independent count statistics are assumed for each transmitted bit, and it is assumed that the array elements are identical and $T_0 = 1 \mu s$. In this figure, BER is plotted as a function of $K_s$ for different values of $C^\text{FF}$ and $K_s$. As shown, Monte Carlo simulations and analytical models result in perfectly matching curves. Also note that the threshold depends on the average number of received photons from both the source and background noise, and this highlights a technical challenge with the OOK system, as $\lambda_s$ and $\lambda_d$ must be known exactly to optimally set the threshold.

According to this figure, it can also be concluded that the array FF has an important role in the performance of a SPAD-based array receiver where the increase in the array FF improves the system performance.

IV. CONCLUSION

In this paper, a comprehensive study of SPAD-based optical receivers is conducted. The detection statistics and main characteristics of single SPAD and SPAD array receivers are discussed and it is shown that, the counting distribution of a large size SPAD array can be well approximated by Gaussian distribution. The effects of SPAD dead time and array fill factor on the photon counting process and the maximum achievable count rate is also investigated. In addition, the error performance of an OOK modulation optical system is studied and it is concluded that as the background counts increase, a higher signal power is needed to maintain the system performance.

ACKNOWLEDGMENT

Professor Harald Haas acknowledges support by the UK Engineering and Physical Sciences Research Council (EPSRC) under Grant EP/K008757/1.

REFERENCES