On the Information Transfer Rate of SPAD Arrays

On the Information Transfer Rate of SPAD Arrays

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

***Abstract*—In this paper the information transfer rate of a single-photon avalanche diode (SPAD) array is investigated. The SPAD array is modelled as a discrete-time Gaussian channel with signal-dependent mean and variance. The SPAD dead time is a parameter which affects the extent of this signal dependency. The SPAD array channel capacity and the properties of the capacity-achieving input distributions are studied. Using a numerical algorithm, the capacity and the optimal input distributions subject to peak and average power constraints are obtained for various array sizes, dead times and background count levels.**

***Index Terms*—Single-photon avalanche diode (SPAD), SPAD array, photon counting, dead time, capacity.**

I. Introduction

Single-photon avalanche diodes (SPADs) are extremely sensitive devices capable of detecting very weak light signals. Nowadays, they are drawing particular attention in the field of optical wireless communication (OWC), resulting in wider and deeper studies among the scientific research community. Thermal noise is one of the main sources of noise in any electronic circuit. In most light detectors, thermal noise is considerable and significantly limits the sensitivity of the device. SPADs can overcome this noise thanks to their unique operating mechanism, thereby achieving very high sensitivities. This makes them a promising candidate for OWC applications which deal with very weak light levels or involve data transmissions over long distances.

The physical characteristics of the SPAD photodetectors are well studied in the literature [1]–[3]. There are a number of contributory factors playing a role in the SPAD photon counting functionality, among which *dead time* is the most impactful. Dead time is the inevitable periods of inactivity the SPAD experiences during the photodetection. Depending on the physical structure of the SPAD, this dead time can be constant (as in active quenching (AQ) SPADs) or variable (as in passive quenching (PQ) SPADs), in any case, it degrades the SPAD photodetection performance. The device limitations have been examined through extensive experiments [4]–[7]. However, since SPADs have just recently found applications in the OWC domain, they are not fully characterized from a communication perspective. Many questions regarding their performance for OWC applications are still unanswered. Therefore, it is necessary to deeply understand the characteristics and limitations of these devices from a communication theory point of view and analyse their performance in this context.

In [8] and [10], we presented a thorough characterization of single SPADs, where we derived the exact photocount distribution of both AQ and PQ single SPADs under the limits of a short dead time. We investigated the effect of a long dead time on the photocount statistics of AQ single SPADs in [11]. In [8], [10], [11], the reliability performance of single SPADs for OWC applications was specified by system error probabilities. It was found that the existence of dead time results in a considerable increase in the bit error probability. Another key performance metric is the maximum achievable data rate of the OWC system. In SPAD-based systems, the transmission data rate needs to be chosen in such a way that a reliable photon counting performance is ensured for the SPAD photodetector. In these systems, the dead time limits the achievable data rates. In [12], it was shown that in single SPADs, the dead time does not allow data rates higher than a few Mbits/s, as the relative length of dead time compared to the system time slots is the determining factor, rather than the absolute length of dead time.

The existing experimental data suggests that by using SPAD arrays, the effect of dead time can be mitigated; SPAD arrays exhibit lower effective dead times and can achieve higher data rates with improved photon detection performance [13], [14].

In [9], we characterized an AQ SPAD array for OWC applications, where we derived the photocount distribution of SPAD arrays and evaluated the bit error ratio (BER) of an OWC system under the limits of dead time. In this paper, we model the SPAD array as a communication channel, then we assess the limiting effect of dead time by the *input-output information transfer rate* of the corresponding channel. In this context, the relevance of the *channel capacity* as a performance metric is clear. As long as the information transfer rate through the SPAD array channel is less than its capacity, it is possible to make the error probability arbitrarily small with proper modulation and coding schemes. In this paper, we also obtain the channel capacity and the capacity-achieving input distributions for AQ SPAD arrays for various array sizes, dead time values and background count levels.

The remainder of this paper is organized as follows. In Section II, the photon counting statistics of AQ single SPADs and SPAD arrays are briefly discussed and the dead time-modified photocount distribution of SPAD arrays is provided. In Section III, a signal-dependent Gaussian channel model is proposed for the SPAD array. In Section IV, the properties of the SPAD array channel capacity are studied. The capacity-achieving distributions are obtained and the effect of various parameters are discussed. Finally, we provide some conclusions in Section V.

II. Photocount Statistics

A. Single SPAD

The photon counting process of a SPAD photodetector is adversely impacted by its dead time. After each photon detection, the SPAD becomes unresponsive for the duration of its dead time and can not detect any other photons. This leads to some *counting losses*. In the photon counting context, the *photocount* refers to the number of successfully detected photons in the so-called *counting interval*. The number of counting losses is random and depends on the *relative* dead time, i.e., the dead time to the counting interval ratio.

Assume that the photons arrive on the surface of the SPAD photodetector according to a Poisson arrival process and the SPAD is free at the beginning of the counting interval. Consider a counting interval of length T and a constant average photon rate λ (in photons/s). Let the random variable K denote the photocounts in the counting interval and k be an arbitrary realization of K. For an ideal single SPAD without dead time, the photocounts follow a Poisson distribution [15]:

$$p_0(k) = \frac{(\lambda T)^k e^{-\lambda T}}{k!}, \tag{1}$$

The mean and variance of the photocounts are given by [15]:

$$\mu_0 = \sigma_0^2 = \lambda T. \tag{2}$$

In the presence of dead time, the photocounts are no longer Poisson distributed. With a dead time of length τ, the maximum number of photocounts is $k_{\max} = \lfloor T/\tau \rfloor + 1$, where $\lfloor x \rfloor$ denotes the largest integer that is smaller than x. The probability mass function (PMF) of the photocounts in this case is given by [8], [10]:

$$p_K(k) = \begin{cases} \sum_{i=0}^{k} \psi(i,\lambda_k) - \sum_{i=0}^{k-1} \psi(i,\lambda_{k-1}), & k < k_{\max} \\ 1 - \sum_{i=0}^{k-1} \psi(i,\lambda_{k-1}), & k = k_{\max} \\ 0, & k > k_{\max} \end{cases} \tag{3}$$

where the function $\psi(i,u)$ is defined as $\psi(i,u) = u^i e^{-u}/i!$ and $\lambda_k = \lambda(T - k\tau)$. The mean and variance of this PMF are expressed as [8], [10]:

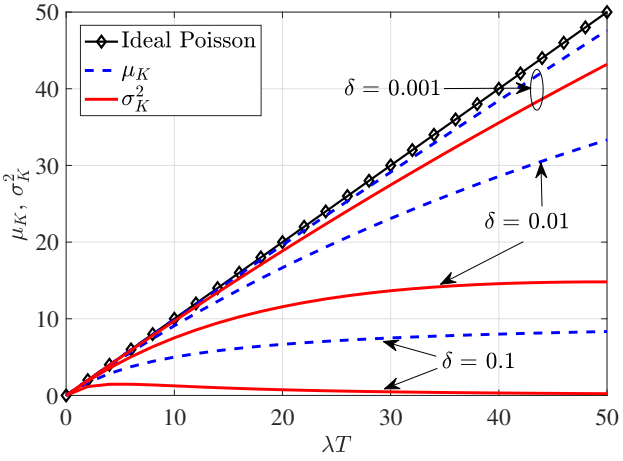
$$\mu_K = k_{\max} - \sum_{k=0}^{k_{\max}-1} \sum_{i=0}^{k} \psi(i,\lambda_k), \tag{4a}$$

$$\sigma_K^2 = \sum_{k=0}^{k_{\max}-1} \sum_{i=0}^{k} (2k_{\max} - 2k - 1)\psi(i,\lambda_k) - \left(\sum_{k=0}^{k_{\max}-1} \sum_{i=0}^{k} \psi(i,\lambda_k) \right)^2. \tag{4b}$$

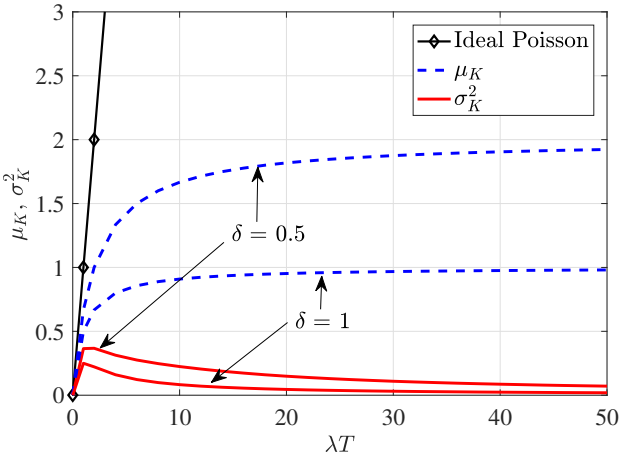
For $\tau \ll T$, the expressions in (4) are approximated as [10]:

$$\mu_K \approx \frac{\lambda T}{1+\lambda\tau}, \tag{5a}$$

$$\sigma_K^2 \approx \frac{\lambda T}{(1+\lambda\tau)^3}. \tag{5b}$$



(a) $\delta = 0.001, 0.01, 0.1$



(b) $\delta = 0.5, 1$

Fig. 1. Comparison of mean and variance of a single SPAD with ideal Poisson distribution, for $T = 1\mu$s and various δ values.

Fig. 1 presents μ_K and σ_K^2 as functions of λT with $T = 1$ μs. Let the dead time ratio be defined as $\delta = \tau/T$. In Fig. 1(a), μ_K and σ_K^2 are plotted for $\delta = 0.001, 0.01, 0.1$, and are compared to an ideal single-photon detector with μ_0 and σ_0^2 defined in (2). In Fig. 1(b), $\delta = 0.5, 1$ are considered. As shown, for $\delta = 0.1, 0.5$ and 1, μ_K and σ_K^2 are significantly lower than those of an ideal detector with Poisson distribution. As λT increases, μ_K tends to a constant value ($= 1/\tau$) and σ_K^2 approaches zero. For $\delta = 0.001$, the gap between the mean and variance is small, such that μ_K and σ_K^2 almost grow with the same slope as μ_0 and σ_0^2. According to Fig. 1, the mean and variance of the photocounts strongly depend on the dead time ratio δ.

B. SPAD Array

The output of a SPAD array is the summation of all photocounts of the individual SPADs over the same counting interval. The fill factor (FF) of the SPAD array also affects

the photocount distribution[1].

Consider an array of N_{array} SPAD elements, and assume independent counting statistics for the array elements (due to negligible crosstalk[2]). Denote by K_i the photocount at the ith SPAD. Thus, the array total photocount is expressed as [9]:

$$Y = \sum_{i=1}^{N_{\text{array}}} K_i. \tag{6}$$

In sufficiently large arrays, according to the central limit theorem (CLT), the photocount distribution of the SPAD array can be approximated by a Gaussian distribution, that is:

$$p_Y(y) \sim \mathcal{N}(\mu_Y, \sigma_Y^2), \tag{7}$$

with,

$$\mu_Y = \sum_{i=1}^{N_{\text{array}}} \mu_{K_i}, \quad \sigma_Y^2 = \sum_{i=1}^{N_{\text{array}}} \sigma_{K_i}^2. \tag{8}$$

Here, μ_{K_i} and $\sigma_{K_i}^2$ are the mean and variance of the photocount distribution of the ith SPAD in the array, respectively.

In the SPAD array, the incident light intensity is divided between the array elements. It is then very unlikely that all the SPADs become inactive at the same time. Therefore, the array can detect a larger portion of the incident light intensity. In addition, the dead time losses decrease as the size of the array increases. In fact, with an array of N_{array} SPADs, the effective dead time reduces to τ/N_{array} [9]. This means that SPAD arrays are more robust to dead time and can tolerate longer dead times, maintaining the required photon counting performance. However, the dead time still limits the maximum achievable count rate and determines the saturation level or peak value of the count rate curve, but the effect is shifted to higher photon rates [9].

III. SPAD Array Communication Channel Model

Consider the SPAD array as a memoryless communication channel with a discrete-time signalling scheme: the channel input is the intensity of the optical signal which can vary between discrete time slots of length T while remaining constant within each time interval. The channel output is the number of detected photons in each time interval, corrupted by background counts of a constant intensity λ_{b} photons/s. Thus, $K_{\text{b}} = \lambda_{\text{b}} T$ is the average background counts per counting interval. Here, we refer to all the noisy counts (arising from dark counts, afterpulsing, and ambient light) as the background counts. Let the random variables X and Y denote the channel input and output, respectively. Further, x and y are arbitrary realizations of X and Y, respectively. Therefore, the channel model for the SPAD array can be expressed as:

$$p_{Y|X}(y|x) = \frac{1}{\sqrt{2\pi\sigma_{Y|X}^2(x)}} \exp \frac{-(y-\mu_{Y|X}(x))^2}{2\sigma_{Y|X}^2(x)} \tag{9}$$

where $x \in \mathbb{R}^+$ and $y \in \mathbb{Z}^+$. Also:

$$\mu_{Y|X}(x) = \sum_{i=1}^{N_{\text{array}}} \mu_{K_i}\left(\frac{x+\lambda_{\text{b}}}{N_{\text{array}}}\right), \tag{10a}$$

$$\sigma_{Y|X}^2(x) = \sum_{i=1}^{N_{\text{array}}} \sigma_{K_i}^2\left(\frac{x+\lambda_{\text{b}}}{N_{\text{array}}}\right). \tag{10b}$$

Note that according (4a) and (4b), the mathematical expressions of μ_{K_i} and $\sigma_{K_i}^2$, and hence $\mu_{Y|X}$ and $\sigma_{Y|X}^2$ are functions of x, i.e., λ. The mean and variance of the SPAD array photon counts are signal-dependent due to the dead time. Therefore, the dead time is the parameter to determine the degree of signal dependency.

IV. Capacity Analysis

In this section, the effect of dead time on the capacity of the SPAD array channel is investigated. Due to practical considerations and device limitations, such as the saturation of SPADs at high intensities [8], [9], the input signal is subject to peak and average power constraints. Furthermore, since X is the light intensity, the constraints are directly imposed on X. In addition, X should be nonnegative. Thus,

$$\begin{aligned} 0 \le X \le \mathcal{A}, \\ \mathbb{E}[X] \le \mathcal{E}, \end{aligned} \tag{11}$$

where $\mathcal{A}$ and $\mathcal{E}$ are the peak and average power, respectively. Without loss of generality, we assume that $0 \le \mathcal{E} \le \mathcal{A}$ and $\mathcal{A}$ is finite.

For the Gaussian channel given in (9), $\mu_{Y|X}$ and $\sigma_{Y|X}^2$ are signal-dependent, unlike the classical Gaussian channels [16]. Such a class of Gaussian channels are termed as conditionally Gaussian (CG) channels [17], [18]. Although the properties of such channels have been studied in the literature, their capacity is not yet known [17], [18]. Nevertheless, it is well known that subject to peak and average power constraints, the channel capacity is achievable and the capacity-achieving distribution is unique and discrete with a finite number of mass points for finite $\mathcal{A}$ and $\mathcal{E}$ [17]. In what follows, some of the findings in the aforementioned reference articles are adopted to study the capacity of SPAD arrays.

Assume an input distribution defined over constellation $\psi_x = \{a_1, a_2, \ldots, a_l\}$, with probability distribution $\psi_p = \{p_1, p_2, \ldots, p_l\}$, where $l = |\psi_x|$ and $0 \le a_1 < a_2 < \cdots < a_l \le \mathcal{A}$. Denote by P_X the corresponding cumulative distribution function (CDF), that is:

$$dP_X = p_1\delta(x-a_1) + p_2\delta(x-a_2) + \cdots + p_l\delta(x-a_l), \tag{12}$$

[1]FF is defined as the ratio of the SPAD array active area to the total array area and is denoted by C_{FF}. Without loss of generality, we assume $C_{\text{FF}} = 1$, i.e. the entire surface of the SPAD is sensitive and therefore, the probability that the incident photons hit the active area is equal to one.

[2]Crosstalk is a phenomenon that takes place in SPAD arrays, when the avalanche in one SPAD triggers an undesired secondary avalanche in a neighbouring SPAD.

where $\delta(\cdot)$ is the Dirac delta function. Also, let $\mathcal{P}_X$ be the set of all input distributions satisfying the constraints defined in (11):

$$\mathcal{P}_X \triangleq \left\{ P_X : \int_0^{\mathcal{A}} dP_X = 1, \mathbb{E}[X] \leq \mathcal{E} \right\}. \tag{13}$$

Let $p_{Y|X}(y|x)$ be the conditional probability of Y given X. For each P_X, denote the corresponding distribution of Y by $p_Y(y; P_X)$, the marginal entropy of Y by $H(Y; P_X)$, the conditional entropy of Y given X by $H(Y|X; P_X)$, and the mutual information between Y and X by $I(P_X)$ [19]:

$$\begin{aligned} p_Y(y; P_X) &= \int_x p_{Y|X}(y|x) dP_X \\ H(Y; P_X) &= -\sum_y p_Y(y; P_X) \log_2 p_Y(y; P_X) \\ H(Y|X; P_X) &= \frac{1}{2} \int_x \log_2 2\pi e \sigma^2_{Y|X}(x)\ dP_X \\ I(P_X) &= \int_x \left[\sum_y p_{Y|X}(y|x) \log_2 \frac{p_{Y|X}(y|x)}{p_Y(y; P_X)} \right] dP_X \end{aligned} \tag{14}$$

And the channel capacity is [19]:

$$C = \max_{P_X \in \mathcal{P}_X} I(P_X)\ . \tag{15}$$

Let the capacity-achieving values of ψ_x, ψ_p, and P_X subject to the constraints $\mathcal{A}$ and $\mathcal{E}$, be denoted by $\psi_x^*(\mathcal{A}, \mathcal{E})$, $\psi_p^*(\mathcal{A}, \mathcal{E})$, and $P_X^*(\mathcal{A}, \mathcal{E})$, respectively. In the following, some of the main properties of the capacity-achieving distribution are summarized [17]:

1) *Existence and uniqueness*: There exists a unique probability measure P_X^* satisfying the bounded-input and average power constraints which maximizes $I(P_X)$.

2) *Necessity and sufficiency*: P_X^* is capacity-achieving if and only if there exists $\epsilon \geq 0$ such that for all $x \in [0, \mathcal{A}]$:

$$Q(x; P_X^*) - I(P_X^*) - \frac{1}{2} \log_2 2\pi e \sigma^2_{Y|X}(x) - \epsilon(x - \mathcal{E}) \leq 0, \tag{16}$$

where,

$$Q(x; P_X) = -\sum_y p_{Y|X}(y|x) \log_2 p_Y(y; P_X). \tag{17}$$

3) *Discreteness*: The capacity-achieving distribution P_X^*, is discrete and consists of a finite set of mass points.

4) *Mass point at zero*: The capacity-achieving distribution always contains a mass point located at zero. That is, $0 \in \psi_x^*(\mathcal{A}, \mathcal{E})$. Therefore,

$$\epsilon = \frac{1}{\mathcal{E}} \left[I(P_X^*) - Q(0; P_X^*) + \frac{1}{2} \log_2 [2\pi e \sigma^2_{Y|X}(x)] \right]. \tag{18}$$

5) *Mass point at peak power*: The capacity-achieving distribution contains a mass point located at $\mathcal{A}$, i.e., $\mathcal{A} \in \psi_x^*(\mathcal{A}, \mathcal{E})$.

Although the above properties of the capacity-achieving distributions for the CG channels are known, closed-form analytical expressions are unknown in general. Therefore, we apply numerical methods in order to compute the capacity and capacity-achieving distributions for the SPAD array channel.

Algorithm 1 Search algorithm for finding the capacity-achieving input distribution.

Input: $\mathcal{A}$, $\mathcal{E}$
Output: C, P_X^*

procedure CAPACITY($\mathcal{A}, \mathcal{E}$)
  $l \leftarrow 2$
  Solve (15) such that $|\psi_x| = l$. ▷ See [16].
  Determine $\epsilon^{(l)}$ according to (18).
  if $\epsilon^{(l)} < 0$ **then**
    $l \leftarrow l + 1$
    go to 3
  end if
  if (16) holds for all $x \in [0, \mathcal{A}]$ **then**
    return C and P_X^*
  else
    $l \leftarrow l + 1$
    go to 3
  end if
end procedure

Using an approach similar to that of [17], the optimal input distribution and the capacity of the SPAD array channel can be obtained via the search algorithm presented in Algorithm 1.

In Algorithm 1, the inputs are $\mathcal{A}$ and $\mathcal{E}$. The algorithm initializes with a binary distribution ($l = 2$). In each iteration, first the optimal P_X which maximizes $I(P_X)$ is obtained using the method presented in [16]. Since a mass point at $x = 0$ always exits, $\epsilon^{(l)}$ is determined using (18). Failure of the necessary condition $\epsilon^{(l)} > 0$ indicates that this P_X is not optimal and the current number of mass points, l, is not sufficient. Thus, the number of mass points should be increased by one, and the distribution function P_X which maximizes the information rate (subject to constraints) should be determined again. If $\epsilon^{(l)} > 0$, then the necessary and sufficient condition in (16) is tested. If it is satisfied, then P_X is the capacity-achieving probability distribution. Otherwise, l is increased by one and the procedure is repeated.

A. Numerical Results

Here, some numerical results on the capacity of the SPAD array channel and the optimal distributions are provided.

Fig. 2 illustrates the SPAD array channel capacity as a function of dead time ratio δ for $\mathcal{A} = 50$, $\mathcal{E} = 20$, $K_\mathrm{b} = 5$, and various array sizes. According to this figure, larger arrays exhibit higher capacities for longer dead times, and are more robust to dead time losses. As can be seen, if $N_\mathrm{array} = 1024$, for the considered values of the parameters $\mathcal{A}$, $\mathcal{E}$ and K_b, the effect of dead time is effectively cancelled. This is because the array does not experience saturation in these operating conditions. If the intensity of the optical signal and/or the background noise level is high such that the array gets saturated (e.g., $N_\mathrm{array} = 16$), longer dead times lead to lower capacity values. Followed by Fig. 2, in Figs. 3(a) and 3(b) the effect of dead time on the optimal mass points (ψ_x^*) and the corresponding probability distributions (ψ_p^*) is presented for one small array ($N_\mathrm{array} = 16$) and one large array ($N_\mathrm{array} = 1024$). In these figures, the parameters $\mathcal{A}$, $\mathcal{E}$,

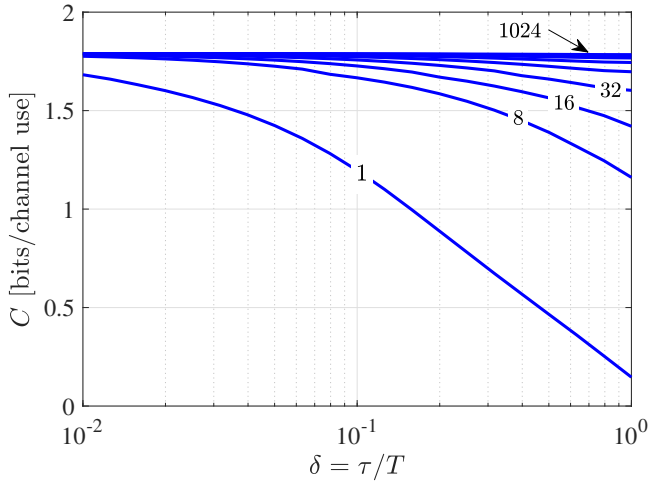
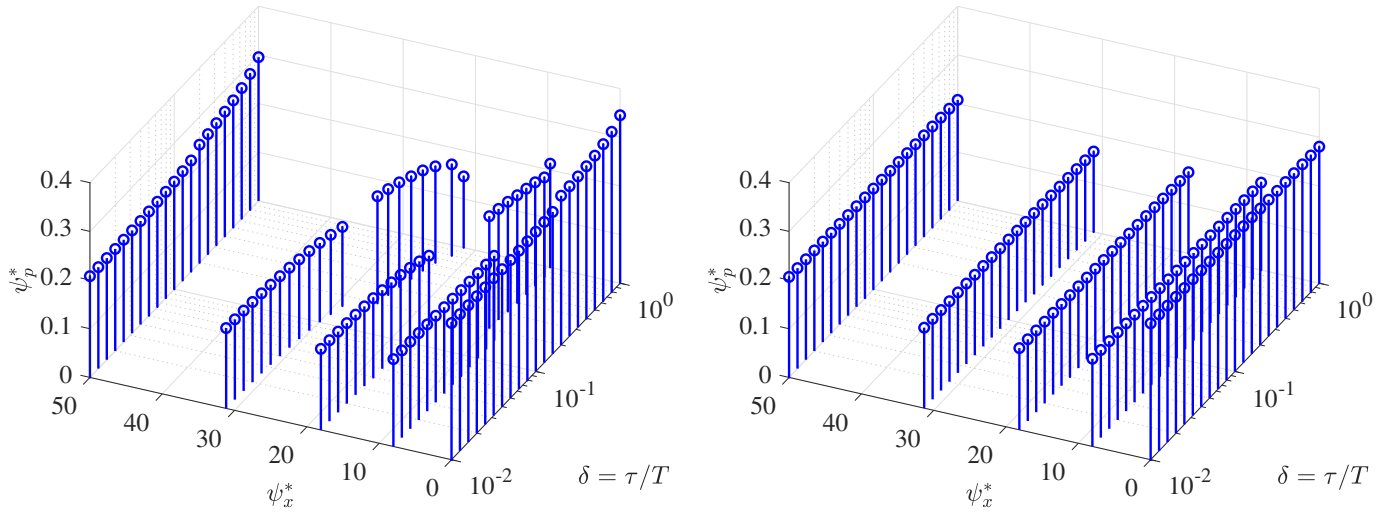


Fig. 2. SPAD array capacity as a function of δ for several array sizes: $\mathcal{A} = 50$, $\mathcal{E} = 20$, and $K_{\rm b} = 5$.



(a) $N_{\rm array} = 16$ (b) $N_{\rm array} = 1024$

Fig. 3. SPAD array capacity-achieving distributions: $\mathcal{A} = 50$, $\mathcal{E} = 20$ and $K_{\rm b} = 5$.

$K_{\rm b}$ are assumed to remain fixed as follows: $\mathcal{A} = 50$, $\mathcal{E} = 20$, $K_{\rm b} = 5$. It is observed that for $N_{\rm array} = 16$, as δ grows, the number of mass points decreases. However, in the case of $N_{\rm array} = 1024$, the capacity-achieving measure (ψ_x^*, ψ_p^*) does not depend on δ. It contains 5 mass points in total, including one at $x = 0$ and one at $x = \mathcal{A}$, for $0.01 \leq \delta \leq 1$.

Fig. 4 plots the channel capacity of an array of 64 SPADs versus $\mathcal{A}$, with the average power constraint $\mathcal{E} = 20$, $K_{\rm b} = 5$, 10 and $\delta = 0.1$. In this figure $|\psi_x^*|$ denotes the total number of the mass points in the capacity-achieving distribution. As $\mathcal{A}$ increases, more mass points are required to achieve the capacity. It should be noted that in the cases where $\mathcal{A} < \mathcal{E}$, the optimization is only subject to the peak power constraint, since the average power constraint is always met.

Fig. 5 illustrates the channel capacity of an array with $N_{\rm array} = 1024$ and $\delta = 1$ versus $\mathcal{E}$, with the peak power constraint $\mathcal{A} = 50$ and $K_{\rm b} = 5$, 10. If $K_{\rm b} = 10$, and $\mathcal{E} < 20$, the optimal input distribution has 5 mass points and as $\mathcal{E}$ increases, $|\psi_x^*|$ decreases to 4. However, for $\mathcal{E} > 20$, the channel capacity remains constant, and the average power constraint becomes ineffective. In this regime, the capacity is solely limited by the peak power constraint. In addition, note that in classical CG channels with the additive Gaussian noise, whose power is a linear function of the channel input (e.g. the CG channel studied in [17]), the larger $\mathcal{E}$ is, the more mass points are required. However, in the case of SPAD arrays the trend is reversed; this is because the variance of the signal-dependent noise decreases for larger values of $\mathcal{E}$.

V. Conclusion

In this study, the information transfer rates of SPAD array photodetectors were studied for OWC applications. A discrete-time Gaussian communication channel model with input-dependent mean and variance was proposed for the SPAD arrays. A numerical algorithm was used to obtain the SPAD array channel capacity and the optimum input distribution subject to average and peak power constraints. It was found that the channel capacity is achievable and the capacity-achieving distribution is discrete with a finite number of mass

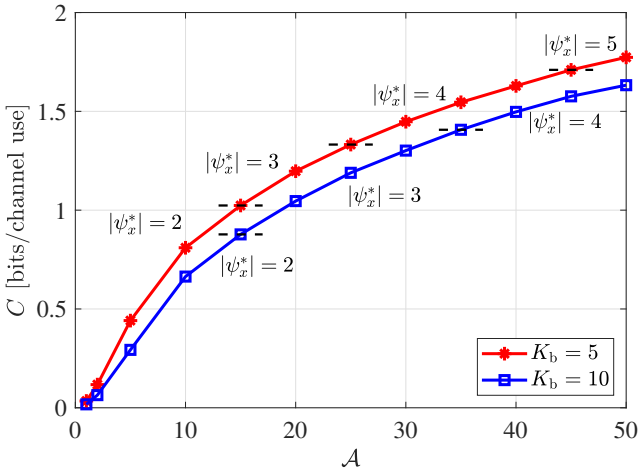


Fig. 4. SPAD array capacity as a function of $\mathcal{A}$: $N_{\text{array}} = 64$, $\delta = 0.1$ and $\mathcal{E} = 20$.

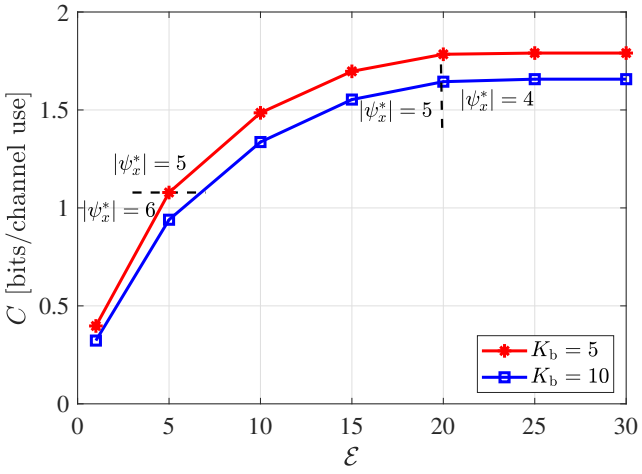


Fig. 5. SPAD array capacity as a function of $\mathcal{E}$: $N_{\text{array}} = 1024$, $\delta = 1$ and $\mathcal{A} = 50$.

points. The results show that, compared with single element SPADs, SPAD arrays can tolerate longer dead times offering higher capacities, and thereby supporting higher data rates for the OWC applications. Depending on the optical signal strength and background noise levels, the effect of dead time may be mitigated or fully cancelled. However, the inherent limitation of dead time remains.

The data rates can be improved to some extent with the help of advanced modulation techniques; but to achieve data rates in the range of Gbits/s, SPAD devices with much shorter dead times are required. SPADs are still a relatively immature technology, thus, in this current stage, they are suitable for scenarios in which the high data rate is not a mandatory requirement. With further technological advances, they may become capable of targeting higher data rates.

The results presented in this work can be used as a benchmark for evaluating the efficiency of practical SPAD-based optical systems. This is particularly required to design efficient modulation schemes, and optimizing the device structure and operating conditions to maximize the achievable data rates.

ACKNOWLEDGEMENT

Professor Haas gratefully acknowledges the support of this research by the Engineering and Physical Sciences Research Council (EPSRC) under an Established Career Fellowship grant, EP/R007101/1. He also acknowledges the financial support of his research by the Wolfson Foundation and the Royal Society. Dr. Safari acknowledges the EPSRC grant ARROW, EP/R023123/1.

REFERENCES

[1] A. Spinelli and A. Lacaita, "Physics and Numerical Simulation of Single Photon Avalanche Diodes," *IEEE Trans. Electron Devices*, vol. 44, no. 11, pp. 1931–1943, Nov. 1997.

[2] R. H. Hadfield, "Single-Photon Detectors for Optical Quantum Information Applications," *Nature Photonics*, vol. 3, no. 12, pp. 696–705, Nov. 2009.

[3] D. Bronzi, F. Villa, S. Tisa, A. Tosi, and F. Zappa, "SPAD Figures of Merit for Photon-Counting, Photon-Timing, and Imaging Applications: A Review," *IEEE Sensors J.*, vol. 16, no. 1, pp. 3–12, Jan. 2016.

[4] H. Dautet, P. Deschamps, B. Dion, A. D. MacGregor, D. MacSween, R. J. McIntyre, C. Trottier, and P. P. Webb, "Photon Counting Techniques with Silicon Avalanche Photodiodes," *Applied Optics*, vol. 32, no. 21, pp. 3894–3900, Jul. 1993.

[5] R. G. Brown, K. D. Ridley, and J. G. Rarity, "Characterization of Silicon Avalanche Photodiodes for Photon Correlation Measurements. 1: Passive Quenching," *Applied Optics*, vol. 25, no. 22, pp. 4122–4126, Nov. 1986.

[6] R. G. Brown, R. Jones, J. G. Rarity, and K. D. Ridley, "Characterization of Silicon Avalanche Photodiodes for Photon Correlation Measurements. 2: Active Quenching," *Applied Optics*, vol. 26, no. 12, pp. 2383–2389, Jun. 1987.

[7] A. Eisele *et al.*, "185 MHz Count Rate, 139 dB Dynamic Range Single-Photon Avalanche Diode with Active Quenching Circuit in 130 nm CMOS Technology," in *Proc. Int. Image Sensor Workshop*, Japan, 2011, pp. 278–281.

[8] E. Sarbazi and H. Haas, "Detection Statistics and Error Performance of SPAD-based Optical Receivers," in *Proc. IEEE 26th Ann. Int. Symp. Personal, Indoor, and Mobile Radio Communications*, Hong Kong, China, Sep. 2015, pp. 830–834.

[9] E. Sarbazi, M. Safari, and H. Haas, "Photon Detection Characteristics and Error Performance of SPAD Array Optical Receivers," in *Proc. IEEE 4th Int. Workshop on Optical Wireless Communications*, Istanbul, Turkey, Sep. 2015, pp. 132–136.

[10] ——, "Statistical Modeling of Single-Photon Avalanche Diode Receivers for Optical Wireless Communications," *IEEE Trans. Commun.*, vol. 66, no. 9, pp. 4043–4058, Apr. 2018.

[11] ——, "The Impact of Long Dead Time on the Photocount Distribution of SPAD Receivers," in *Proc. IEEE Global Communications Conf.*, Abu Dhabi, UAE, Dec. 2018, pp. 1–6.

[12] ——, "On the Information Transfer Rate of SPAD Receivers for Optical Wireless Communications," in *Proc. IEEE Global Communications Conf.*, Washington, DC, USA, Dec. 2016, pp. 1–6.

[13] E. Fisher, I. Underwood, and R. Henderson, "A Reconfigurable Single-Photon-Counting Integrating Receiver for Optical Communications," *IEEE J. Solid-State Circuits*, vol. 48, no. 7, pp. 1638–1650, Jul. 2013.

[14] D. Chitnis and S. Collins, "A SPAD-Based Photon Detecting System for Optical Communications," *IEEE/OSA J. Lightw. Technol.*, vol. 32, no. 10, pp. 2028–2034, May 2014.

[15] R. M. Gagliardi and S. Karp, *Optical Communications*, 2nd ed. New York: Wiley, 1995.

[16] J. G. Smith, "The Information Capacity of Amplitude- and Variance-Constrained Scalar Gaussian Channels," *Inf. Ctrl.*, vol. 18, no. 3, pp. 203–219, 1971.

[17] T. H. Chan, S. Hranilovic, and F. R. Kschischang, "Capacity-Achieving Probability Measure for Conditionally Gaussian Channels with Bounded Inputs," *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 2073–2088, Jun. 2005.

[18] S. M. Moser, "Capacity Results of an Optical Intensity Channel with Input-Dependent Gaussian Noise," *IEEE Trans. Inf. Theory*, vol. 58, no. 1, pp. 207–223, Jan. 2012.

[19] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. John Wiley & Sons, 2012.