A SPAD-Based Photon Detecting System for Optical Communications

Danial Chitnis and Steve Collins

Abstract—A small array of single photon avalanche detectors (SPADs) has been designed and fabricated in a standard 0.18 μm CMOS process to test a new photon detecting system for optical communications. First numerical results are presented which show that using arrays of SPADs reduces the optical power density required at the receiver. Experimental results then show that the new system preserves the photon counting ability of the SPADs. Finally, a simple method is presented which can be used to estimate the size of the array needed to achieve a particular target bit error rate at a specific optical power density. Together these results indicate that by replacing the avalanche photodiode in a receiver with the new system it will be possible to count the received photons.

Index Terms—CMOS, mixed analog digital integrated circuits, optical receivers, poisson channels, single photon avalanche diode.

I. INTRODUCTION

An ideal receiver for plastic optical fibre [1] or visible light communications systems [2] in the edge network is both inexpensive and sensitive. Since the cost of any electronics is reduced by integration the ideal receiver for these systems is an integrated circuit, that is able to detect individual photons. The current state of the art photodetector in sensitive optical receivers are avalanche photodiodes (APDs). Unfortunately, avalanche multiplication creates a gain-dependent excess noise source which limits the maximum APD’s gain to between 10 and 20. However, this additional noise source can be avoided by operating the APD in Geiger-mode as a single photon avalanche detector (SPAD) [3]–[5]. Furthermore, this type of photodetector has been integrated into application specific integrated circuits for a range of applications, including laser ranging [6], three-dimensional imaging [7]–[9], fluorescence imaging [10]–[13] and charged particle detection [14]–[17]. In addition SPADs have been proposed for both intra-chip communications [18], quantum key distribution [19], [20] and deep space laser communications [21]. However, only two groups have previously investigated using SPADs within systems suitable for communicating over the edge network [22], [23]. One of these groups investigated the possibility of using a receiver based upon an array of passively quenched SPADs that form a silicon photomultiplier [22]. Although they predicted that this type of SPAD array would be more sensitive than an APD, to achieve a particular bit error rate (BER) it still required four times more signal power than a photon counting receiver. In the most recent work in this area, the SPAD array was designed to count photons, however, even though the whole receiver has been redesigned it is only compatible with some modulation schemes [23].

In this paper a design for a SPAD-based optical receiver, which is a possible replacement for an APD in an otherwise unchanged receiver, is investigated. In particular, numerical modelling and results from an experimental system are combined to predict the performance of a particular type of SPAD-based receiver. Section II contains some results obtained when SPADs, manufactured on a 0.18 μm CMOS process, are characterised, whilst Section III contains the results of a model that has been used to estimate the BER that might be achieved despite finite SPAD deadtimes and background counts. The SPAD-based receiver that has been manufactured is described at the beginning of Section IV, which also contains results obtained when these receivers were characterised. A model developed to explain some of the results in Section IV is then used in Section V to predict the performance of future systems. Section VI contains concluding remarks.

II. SPAD OPERATION

When an APD is operated in the Geiger-mode the avalanche process is self-sustaining and so any time dependent signal can only be followed if the APD is connected to a load device. If the load device is a resistor the resulting SPAD is said to be passively quenched. Alternatively, if the load device is a transistor the SPAD is actively quenched. In either case when one of these SPADs is awaiting the arrival of a photon, it is said to be in its idle mode, there will be no current flowing through the load device. However, when a photon triggers an avalanche event a large current will flow and the APD bias voltage will decrease until the avalanche process stops. The drop in the voltage across the SPAD is usually converted to a digital pulse using a comparator and as a result SPADs can detect single photons. However, whilst the APD output voltage is below the comparator threshold the comparator output voltage will not change if a second avalanche event occurs. This period is therefore referred to as the deadtime of the SPAD.

SPADs have been fabricated using a UMC 0.18 μm standard CMOS process, available through Europractice, in which the APDs have a photosensitive area with a diameter of 10 μm,

Manuscript received December 12, 2013; revised February 9, 2014 and March 28, 2014; accepted March 28, 2014. Date of publication April 10, 2014; date of current version May 20, 2014. This work was supported by the United Kingdom’s Engineering and Physical Sciences Research Council under Grant EP/G006784/1.

D. Chitnis was with the Department of Engineering Science, University of Oxford, Oxford, OX1 2JD, U.K. He is now with the Department of Medical Physics and Bioengineering, University College London, London, WC1E 6BT, U.K. (e-mail: d.chitnis@ucl.ac.uk).

S. Collins is with the Department of Engineering Science, University of Oxford, Oxford, OX1 2JD, U.K. (e-mail: steve.collins@eng.ox.ac.uk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JLT.2014.2316972

This work is licensed under a Creative Commons Attribution 3.0 License. For more information, see http://creativecommons.org/licenses/by/3.0/
Fig. 1. The theoretical (dashed line) and the experimental (solid lines) relative frequency histograms of the inter pulse time interval statistics for 20 fabricated SPADs. The outer histogram has a logarithmic time axis, and the inner histogram has a linear time axis. Both histograms represent the same set of data. The exponential decay slope in the linear time scale represents the mean time interval, whilst the logarithmic time scale emphasises the shorter time scales of the relative histogram.

within a guard ring with a diameter of 30 μm and a breakdown voltage of approximately 10.4 V [5]. Ideally, avalanche events will only be triggered by a photon. However, even in the dark spontaneous avalanche events occur that create a dark count rate (DCR). Experiments show that the DCR of these SPADs typically varies from 3 kcps at a bias voltage of 11 V to 90 kcps at 12 V. The other possible non-ideal source of avalanche events is after-pulsing. To investigate the probability of after-pulsing the inter-avalanche time interval statistics for the SPADs have been measured by capturing long sequences of pulses generated by individual SPADs using an Agilent MSO6104A oscilloscope which has a 1 ns resolution. Fig. 1 shows the theoretical and experimental histograms of the inter pulse time interval statistics for 20 different SPADs that each have an average count rate of 100 kcps. The fact that the shortest time interval that is observed is 5 ns shows that the deadtime of these actively quenched SPADs can be as small as 5 ns. An important observation from these results is that the match between the experimental and theoretical results means that the probability of after-pulsing is less than 0.7%. This suggests that for a system using On-Off Keying (OOK) after pulsed will only be important if the extinction ratio (ER) of the transmitter is larger than 100.

III. BER MODELLING

SPADs have the potential to function as photon counting receivers. In a well designed photon counting receiver the dominant noise source will be statistical fluctuations in the time between photons and this will determine the achievable BER. However, the performance of any SPAD-based receiver might be limited by the deadtime of the SPADs and/or background counts (arising from a combination of dark counts, afterpulsing, ambient light and/or a transmitter’s finite ER). The effect of the deadtime on the average number of photons per bit time needed to achieve a particular BER for simple OOK modulation therefore has been investigated using numerical simulation. Fig. 2 shows the BER = 10^{-3} contours as a function of the average number of photons detected by an array of SPADs and the ratio between the deadtime of the SPADs in the array and the bit-time of the transmitted data.

![Probability vs Time Interval](image1)

Fig. 2. The BER = 10^{-3} contours as a function of the average number of photons detected by an array of SPADs and the ratio between the deadtime of the SPADs in the array and the bit-time of the transmitted data.

In real systems, the dark counts and after-pulsing in each SPAD, background light and a finite transmitter ER can all contribute to a background count rate that is proportional to the number of SPADs in the receiver. When a 0 is transmitted, the average number of photons detected by each SPAD per bit will be $\lambda_d$, and for an array of $M$ SPADs the average number of photon counts when a 0 is transmitted will be $M \lambda_d$. However, when a 1 is transmitted the average number of photon counts per bit time is $\lambda_{sb} = M \lambda_s + M \lambda_d$. The existence of background counts means that when OOK is being used a threshold photon count is required to distinguish between a 0 and a 1 bit. If the decision threshold is, $n_T$, then for a random bit sequence, $p(T1) = p(T0) = 0.5$, the BER is [24]

$$
\text{BER} = \frac{1}{2} \sum_{k=0}^{n_T} (\lambda_{AS} + M \lambda_b) e^{-(\lambda_{AS} + M \lambda_b)}
+ \frac{1}{2} \sum_{k=n_T}^{\infty} (M \lambda_b)^k e^{-M \lambda_b}
$$

(1)
which suggests that the BER is a function of $M\lambda_b$, $n_T$ and $\lambda_{AS} = M\lambda_s$. However, in a complete receiver it is anticipated that $n_T$ will be selected to ensure that the best possible BER is achieved, and if this is done, the BER for a receiver is a function of $\lambda_{AS}$, $M$ and $\lambda_b$. To study the potential effect of background counts on BER, the decision threshold $n_T$, has been determined for three BERs and different values of both $\lambda_{AS}$ and $\lambda_b$. The resulting BER contours in Fig. 3 show abrupt changes whenever $n_T$ changes. One potentially important observation from these results is that the maximum negligible background count is related to the target BER. In particular, as shown by the dashed line 3, the background count is negligible when $M\lambda_b < 0.1 \times BER$. More importantly, although the number of photons needed to achieve a particular BER increases as the background count increases, the rate of increase is less than linear. This means that although increasing the number of SPADs in an array will increase both the signal and the background counts at the same rate, the achievable BER is expected to improve as the number of SPADs increases. As shown in Fig. 4, this means that increasing the number of SPADs in a receiver will significantly reduce the average number of photons that have to be detected by each SPAD per transmitted bit, $\lambda_s$. This means that despite background counts, increasing the number of SPADs in a receiver will reduce the transmitted optical power needed to achieve a specific BER.

IV. EXPERIMENTAL RESULTS

The results in the previous section show that the impact of both background counts and deadtime can be reduced using an array of SPADs. Arrays of SPADs, in which each SPAD is associated with a digital counter, have previously been proposed for use in an optical communications link. Although this digital readout method exploits the digital output from SPADs, the counting process has to be carefully synchronised to the transmitter clock [23], [25] and a new receiver design is required to recover the clock and data from the transmitted signal. Even then, not all modulation schemes can be supported [23]. To avoid this limitation an alternative approach to creating a SPAD-based receiver has been investigated. In particular, a SPAD-based system has been designed that can function as a replacement for the APDs in existing receivers.

As shown in Fig. 5, each SPAD is actively quenched by an automatic diode reset (ADR) circuit, which includes transistor switches to first hold and then reset the avalanche diode, and two monostables to provide appropriate timings for these switches. An analogue input voltage, $V_{DT}$, to the hold monostable determines the deadtime of the SPAD. To create a system that has an output current the readout method is based upon the current sources used within equally weighted current steering digital-to-analogue converters [26]. In particular, in Fig. 5, these current sources are formed by a pMOS transistor acting as a constant current source connected to a pair of pMOS transistors that steer this current to one of two outputs. The two voltages needed to control this pair of devices are generated by a switch driver circuit that is controlled by the SPAD output. When the SPAD is in its idle mode, most of the current $I_{CS}$ flowing through $M_{CS}$, flows through the negative output. However, when an avalanche event occurs, the switch driver causes more current to flow through $M_{S1}$ than through $M_{S2}$. Once the deadtime has elapsed, the state of the switch driver is changed by the ADR circuit, and the current is steered back to the negative output. Connecting the individual positive and negative outputs creates an output current, $I_{out}$, which is proportional to the number of SPADs in which an avalanche event has occurred within a deadtime. To facilitate experiments the current flowing through each constant current source is controlled by an input voltage, $V_{CS}$,
applied to the gate of these devices. This voltage can be used to ensure that the current generated by a photon is larger than the noise generated in the subsequent stages of the receiver.

To test this circuit idea an integrated circuit has been manufactured that contains an array of SPADs with their associated circuits. In view of the limited area available in the Europractice mini-ASIC used to manufacture the integrated circuit, and the results in Fig. 2, an experimental array of 64 SPADs has been fabricated using a UMC $0.18 \mu m$ standard CMOS process. In order to save area and more importantly increase flexibility, this array has an output current which can be converted to a voltage by an off-chip resistor. Fig. 6 shows a micrograph of the manufactured mini-ASIC and a close up of one SPAD and its associated circuits. In addition to this array, four individual SPADs and associated circuits were included in the mini-ASIC so that these circuits could be characterised. Some results of this characterisation are summarised in Table I.

A parameter that will be critical to the performance of this receiver is the deadtime. The importance of this parameter arises because the instantaneous amplitude of the output signal is proportional to the number currently inactive SPADs. However, if the deadtime is too long compared to the bit time, it will cause inter symbol interference. To investigate the effects of different deadtimes eye diagrams have been captured as the deadtime was varied. For this experiment a 860 nm laser diode with a finite ER has been used as the transmitter [27]. A pseudo-random bit stream was generated in realtime using a linear feedback shift register (LFSR) with a repetition rate of $2^{21} - 1$ [28]. In order to obtain rise and fall times that are relative short compared to the bit time, a bit time of 50 ns was selected for this experiment. To create an output voltage, one of the shared outputs from the current steering circuits output was connected to a 300 $\Omega$ off-chip load resistance. The output current from the current sources was then set to its maximum value, which gives a step change of approximately 30 mV in the output voltage. Although the linearity of the output is affected by setting the maximum current, it is not important in this experiment to have a linear output because the decision threshold for moderate BERs is within the linear region of the resulting output.

To investigate the effect of deadtime the input voltage that controls the SPAD deadtime was varied and an eye diagram for each deadtime was formed by superimposing many traces. Fig. 7 shows three eye diagrams representing the conditions when the deadtime is shorter than the bit time, comparable to the bit time and longer than the bit time. As expected increasing the deadtime

<table>
<thead>
<tr>
<th>Active area diameter</th>
<th>$10 \mu m$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total diameter</td>
<td>$30 \mu m$</td>
</tr>
<tr>
<td>Breakdown voltage ($V_{BD}$)</td>
<td>10.4 V</td>
</tr>
<tr>
<td>Temperature coefficient of $V_{BD}$ (-70 °C to +25 °C)</td>
<td>6.9 mV/°C</td>
</tr>
<tr>
<td>Average DCR when Vex=1.6</td>
<td>90 KHz</td>
</tr>
<tr>
<td>Peak PDP at 490 nm when Vex=1.6</td>
<td>26 %</td>
</tr>
<tr>
<td>PDP at 850nm when Vex=1.6</td>
<td>1.5 %</td>
</tr>
<tr>
<td>Minimum deadtime</td>
<td>5 ns</td>
</tr>
<tr>
<td>Maximum afterpulsing probability</td>
<td>&lt; 1%</td>
</tr>
<tr>
<td>Maximum $I_{CS}$</td>
<td>100 $\mu A$</td>
</tr>
<tr>
<td>Maximum $I_{OUT} (= \sum I_{CS}$ of the array</td>
<td>6.4 mA</td>
</tr>
</tbody>
</table>

Fig. 7. The eye diagram: (a) when the dead time is shorter than the bit time, (b) when dead time is equal to the bit time (c) when dead time is longer than the bit time, (d) The relative frequency histogram for the dark count rate (DCR) at $T_b = 50$ ns (deadtime same as BER experiment in Fig. 9)
from shorter than the bit time, Fig. 7(a), to comparable to the bit time, Fig. 7(b), causes an increase in the mean output current and this creates a clearer eye. However, as shown in Fig. 7(c), if the deadtime is increased further inter symbol interference results in an edge in the eye diagram which closes the significant eye in the top right hand corner of the diagram. The optimum deadtime for this receiver is therefore a deadtime which is approximately equal to the bit time.

With the deadtime fixed another parameter that will influence the best BER is the background count. Fig. 7(d) shows the relative frequency histogram of the output voltage of a SPAD array in the dark when the deadtime is set to the bit time. Although the measured voltages include noise from the measurement system the data clearly shows peaks corresponding to different numbers of SPADs operating in parallel. These results show that in the dark the background count rate is so low that the most common output corresponds to no active SPADs. In fact, the relative frequency of the peaks corresponding to one, two and three current sources operating in parallel can be predicted using Poisson statistics with an average count of \( \lambda_{DCR} = 0.095 \). Under ideal conditions the average signal count to achieve a BER of \( 10^{-3} \) is 6.2. This means that at a bit rate of 20 Mbps the dark count will be less than the background signal from the transmitter unless the transmitter has an ER larger than 63. This suggests that the DCR from these SPADs will be negligible in many systems when compared to the finite ER of the transmitter. More importantly, the fact that Fig. 7(d) shows outputs corresponding to one, two and three current sources operating in parallel demonstrates that this system can count photons.

The eye diagram suggests the operating conditions that will correspond to a good BER. However, a specific experiment is required to determine the best decision threshold and BER, hence a real-time BER testing platform has been constructed, Fig. 8. The results obtained using this equipment are shown in Fig. 9, which also shows the relationship between the BER and \( \lambda_{AS} \) for an ideal photon counting receiver with a total background count of 0.2 counts per bit. Comparing these two results indicates that, even though the SPAD deadtime is the same as the bit time, when only a few photons per bit are required this relatively small array acts as an ideal photon counting receiver. Not surprisingly, as \( \lambda_{AS} \) becomes a significant fraction of the number of SPADs the BER that is achieved is worse than expected from an ideal receiver. This is because under these conditions the deadtime means that the number of SPADs in the array that are idle at any time is significantly less than the number of SPADs in the array. The effect of this reduction on the effective number of SPADs that are capable of detecting a photon at a particular instant has been estimated using

\[
\lambda_{es} = \lambda_{AS} \times \frac{M}{M - \lambda_{AS}}
\]

where \( \lambda_{es} \) is the new estimate of the average number of photons that can be detected. Fig. 9 suggests that this simple model predicts the measured relationship between BER and \( \lambda_{AS} \) quite accurately. Furthermore, when the model and the data diverge the model overestimates the number of photons required. When the SPAD deadtime and bit time are comparable this simple equation can therefore be used to estimate the minimum performance of an optical receiver formed from an array of SPADs.

V. FUTURE PROSPECTS

Practical constraints, in particular the area available in a Europractice mini-ASIC mean that results have been obtained using only a relatively small number of SPADs. However, the results that have been obtained can be used to estimate the probable performance of future, potentially larger, systems. Fig. 10 shows two estimates of the \( \lambda_{AS} \) required to achieve particular BERs for different sizes of SPAD arrays. One of these estimates assumes that the array is an ideal photon counting receiver, whilst the other is based upon Equation 2. These results again show that although the total background count is proportional to the number of SPADs in an array, the average number of detected photons, \( \lambda_{AS} \), required to achieve a BER increases by a factor less than \( M \). This means that the optical power density needed...
to achieve a particular BER is reduced when larger arrays are used.

The experimental results that have been reported were obtained using an 860 nm transmitter because this light source is faster than those available at shorter wavelengths and this simplified the BER experiment. However, as the peak photon detection probability (PDP) of the avalanche diode occurs in 490 nm, once better transmitters become available it will be sensible to use shorter wavelengths to exploit the higher SPAD PDP at shorter wavelengths. When appropriate this change in wavelength will reduce the optical power density required to achieve a target BER with a particular SPAD array. Alternatively, the SPAD design could be changed to one that is more effective at detecting 860 nm photons [29], [30]. However, one problem that will not be solved by these changes is the fact that the light sensitive area of the SPADs is only a fraction of the total area of the receiver. In the future, the fill-factor, that is the ratio of the photosensitive area to the total area of the receiver, can be improved by more careful layout of the SPADs and their associated circuits. However, even then micro-lenses that focus light onto the active area of each SPAD will be useful [31]. Finally, for systems that employ OOK at bit rates of approximately 1 Gbit/s it may be necessary to add a monostable between the SPAD output and the readout circuits so that current can be steered to the relevant output for periods that are shorter than the deadtime [32]. The resulting photon counting receivers will probably be most attractive in systems that can tolerate a relatively high BER and therefore require fewer detected photons per bit. Fortunately, systems that employ forward error correction can tolerate a BER of $10^{-3}$ [2]. Fig. 3 shows that if the average photon count when a zero bit is transmitted is less than one, then this BER can be achieved if the average number of photons detected when a one is transmitted is 15 or less. The results in Fig. 11 suggest that despite the effects of deadtime ideal photon counting behaviour can be achieved for this BER using arrays of approximately 100 SPADs.

VI. CONCLUSION

Optical communications systems often rely upon receivers that include APDs, however, these photodetectors generate additional noise which limits the gain that can be used. This additional noise can be avoided by increasing the bias voltage applied to the APD so that it is operating in its Geiger-mode as a SPAD. Since SPADs can detect single photons it is anticipated that it will be significantly easier to create inexpensive integrated optical receivers using SPADs than using the APDs.

However, a SPAD has a deadtime during which it is unable to detect another photon. Numerical simulation results have been presented which show that despite this non-ideality, by using an array of SPADs operating in parallel, it should be possible to create a sensitive receiver. A receiver based upon an array of SPADs has been described recently in which the outputs from individual SPADs were added together using digital electronics. However, this meant that a new receiver design was required, and, even with the new design the receiver may not be able to support some modulation schemes [23]. To create a receiver front-end that is compatible with existing receiver designs and modulation schemes the approach that has been investigated is to add the outputs from SPADs using a current steering circuit. Experimental results show that with the SPAD layout that has been employed, the DCR and afterpulsing are negligible when compared to other sources of background counts, particularly the finite extinction ratios of transmitters.

Experimental results have also been presented which show that the optimum deadtime is the bit time. In addition, a simple model has been proposed that can be used to estimate the optical power density needed to achieve a specific BER. This model suggests that in the future it will be possible to design a photon counting receiver that achieves a BER of $10^{-3}$, suitable for use in a system using forward error correction, with an array of approximately 100 SPADs.

Like an APD the SPAD-based circuit that has been designed has an output current. This means that it can replace APDs in existing receivers. However, results have been presented which show that it will be possible to ensure that the current that flows when a single photon is detected will be large enough to dominate the noise floor of the rest of the receiver. In the future the design and the layout of the receiver circuit should

Fig. 10. The BER contours as a function of the average number of photons representing a 1 and the number of SPADs for an ideal photon counting receiver (solid line) and for the model that accounts for deadtime (dashed line).

Fig. 11. Contours for BER = $10^{-3}$ versus optical irradiance and size of the array for 850 nm and 490 nm illumination at 100 Mbps. The solid line shows the photon shot noise limited performance and the dashed line includes an estimate of the effect of having a deadtime equal to the bit time.
be optimised, micro-lenses should be added to concentrate light onto its photosensitive area. The result is expected to be a low-cost, photon counting front end for an otherwise conventional optical receiver that operates close to the quantum limit.

REFERENCES


