Research Article | Vol. 27, No. 13 | 24 Jun 2019 | OPTICS EXPRESS 18201
Optics EXPRESS

Dual anode single-photon avalanche diode for high-speed and low-noise Geiger-mode operation

CHULWOO PARK,[1,2] SEOK-BEOM CHO,[1] CHAN-YONG PARK,[3] SOOHYUN BAEK,[3] AND SANG-KOOK HAN[2,*]

[1]*ID Quantique, 6 Hwangsaeul-ro 258beon-gil, Bundang-gu seongnam-si, Gyeonggi-do, 13595, South Korea*
[2]*Departmnet of Electrical and Electronic Engineering, Yonsei University, 50 Yonsei-ro Seodaemun-gu, Seoul, 03722, South Korea*
[3]*Optical Communication R&D Center, Wooriro Co.,Ltd, 102-22, Pyeongdongsandan 6beon-ro, Gwangsan-gu, Gwangju, 62453, South Korea*
**skhan@yonsei.ac.kr*

Abstract: The after-pulsing effect is a common problem in high-speed and low-noise single-photon detection based on single-photon avalanche diodes (SPADs). This article presents a dual anode InGaAs/InP SPAD (DA-SPAD) with two separate anode output ports that can be utilized for discriminating relatively weak avalanche signals, providing a simple and robust configuration of the SPAD-based single-photon detection system. Weak avalanche signals with amplitudes below the amplitude of the parasitic capacitive response of the SPAD were easily detected by the DA-SPAD and a simple subtraction circuit. The gated Geiger-mode performance of the DA-SPAD was also investigated. At a gating frequency of 1 GHz, the detection efficiency was 20.4% with an after-pulse probability of 3.5% at a temperature of –20 °C.

1. Introduction

Single-photon detection techniques in the near-infrared (NIR) range have been widely studied in various fields, such as biomedical research, optical quantum information, and communication [1–3]. Among the various single-photon detection (SPD) technologies, the single-photon avalanche diode (SPAD) is advantageous for constructing a small and low-cost SPD system. An InGaAs-based SPAD operating in the telecom band is especially promising for fiber-optic quantum communications, optical time domain reflectometry, and eye-safe LiDARs [4,5]. Generally, the SPAD operates in Geiger-mode for SPD, in which the bias voltage applied to the SPAD exceeds the breakdown voltage of the SPAD, and the operating modes are classified into gate-mode and free-running mode depending on the on-off repetition method of the Geiger-mode. However, the InGaAs-based SPD is disadvantaged by a relatively high after-pulse probability.

In the after-pulse phenomenon, some of the avalanche carriers are trapped, triggering subsequent avalanches into Geiger-mode. To avoid the false counts due to after-pulses, a dead-time proportional to the lifetime of the trapped carriers is required, which limits the maximum count rate [6].

To avoid these problems, one must lower the after-pulse probability for low-noise high-speed SPD. As the number of trapped carriers that can induce the after-pulse is proportional to the avalanche gain, a bias with a low avalanche gain is desired. However, lowering the amplitude of the avalanche signal reduces the ability to discriminate an avalanche event from other events. Particularly in gate-mode operation, the minimum avalanche amplitude that can be discriminated by a conventional comparator is determined by the capacitive response

#361968 https://doi.org/10.1364/OE.27.018201
Journal © 2019 Received 11 Mar 2019; revised 20 May 2019; accepted 21 May 2019; published 14 Jun 2019

Optics EXPRESS

signal of the SPAD to the gate signal. Consequently, the after-pulse probability is difficult to reduce.

Various approaches can overcome the problems related to weak avalanche discrimination, the most representative being self-differencing (SD) and sinewave gating (SG) techniques [7,8]. Although these schemes differ in principle, both methods eliminate the capacitive response from the SPAD output, achieving high-speed gate-mode operation above 1 GHz. However, both methods require relatively complicated electronic circuits for removing the capacitive response signal (a time-delay circuit in the SD technique, and bandpass filters in the SG technique). In addition, changing the gate frequency in both methods is relatively cumbersome, meaning that the circuits for removing the capacitive response must be altered when the gate frequency is changed.

This research proposes a dual anode SPAD (DA-SPAD), in which the capacitive response of the SPAD is easily removed by a simple subtraction circuit. The DA-SPAD is composed of one cathode and two anode ports monolithically integrated on the same substrate. One of the two anode ports generates a capacitive response signal only, while the other operates as a SPAD, simultaneously generating an amplified signal and a capacitive response signal. The proposed DA-SPAD is a compact, low-complexity, high-flexibility, cost-effective circuit design for capacitive response compensation. The most important advantage might be the robustness of the designed circuit to changes in gate frequency or gate signal waveform, meaning that no circuit modifications are required. This study demonstrates the capacitive response of the DA-SPAD with a simple subtraction circuit. In 1 GHz gate-mode operation, the proposed design efficiently detects single photons with reasonable low after-pulse probabilities applicable to various fields including quantum key distribution.

2. Configuration of the dual anode SPAD (DA-SPAD)

In gate-mode SPD based on SPADs, the capacitive response limits the operation speed of the SPD and the reduction of the after-pulse probability [6]. In SPAD-based SPD, a photon is usually detected as an avalanche signal with higher amplitude than the capacitive response to the SPAD gate signal. Therefore, the SPAD-specific capacitive response signal generated by the input gate signal limits the minimum threshold for discriminating the avalanche signal. Consequently, avalanche signals with amplitudes below the capacitive response signal cannot be discriminated, which limits the reduction of the after-pulse probability.

To overcome this problem, the Toshiba research group has applied the SD technique, which detects the weak avalanche signals by removing the capacitive response signal using the capacitive response of the neighboring gate signal. SD achieves high-speed operation (2 GHz) [9], and is relatively simple and robust, but cannot easily detect consecutive avalanche events at two gates in the neighborhood. It also requires a time delay corresponding to the gate period. The circuit that generates this time delay must be relatively complex to support the variable gate frequency.

By removing the capacitive response signal, the dual anode SPAD (DA-SPAD) detects weak avalanche signals even with a simple circuit configuration. Figure 1 shows the conceptual design and electronic component symbols of the DA-SPAD chip fabricated by a local company (Wooriro; www.wooriro.com) for this research.

Optics EXPRESS

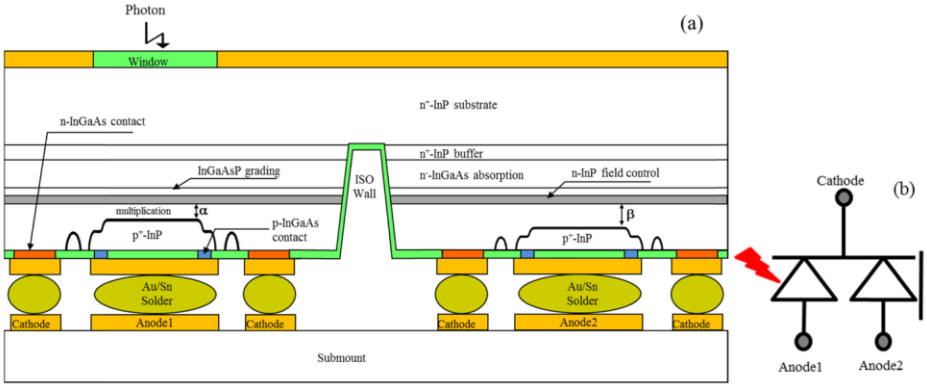

Fig. 1. Conceptual design of the dual anode SPAD: (a) structure and (b) electronic component symbols.

As shown in Fig. 1, the DA-SPAD chip contains two adjacent diodes separated by an isolation wall. The main diode (denoted as Anode 1 in Fig. 1(b)) receives the optical signal and generates the avalanche signal. The auxiliary diode (Anode 2 in Fig. 1(b)) outputs the same capacitive response signal as the main diode but generates no avalanche signal. For this purpose, the designed breakdown voltage of the auxiliary diode is approximately 4 V higher than that of the main diode. In the fabrication process, this is achieved by adjusting the thicknesses of the multiplication layers of the main and auxiliary diodes (denoted α and β, respectively). Difference in the multiplication layer thicknesses applied for the breakdown voltage difference of 4 V can result in difference in capacitances of main and auxiliary diodes, and the capacitance measurement results showed the capacitance difference of about 0.01 pF between the two anodes. To compensate the capacitance difference due to the different thicknesses of the multiplication layers, the diodes were fabricated so that the active area diameter of the auxiliary diode was 1 μm larger than the active area diameter of the main diode. After the fabrication, dozens of samples were randomly selected and their capacitance difference between main and auxiliary diodes were measured, and the measurement results showed that the differences in capacitance values between the two diodes were less than 10 fF or so. Additionally, the auxiliary diode is blinded by a metal mask, while a transparent window is kept over the main diode. The process parameters of the other layers (except the multiplication layer) are identical in the main and auxiliary diodes. The fabricated DA-SPAD chip is assembled on the submount by the flip-chip bonding technique. This chip-on-submount is packaged on a TO46 CAN and subsequently packaged into a coaxial module employing a fiber-pigtail. In the submount with a DA-SPAD, the cathode is structurally and electrically shared with both diodes, so both anodes give the same capacitive responses when the gate signal is applied through the common cathode. In this design, the capacitive response signal is easily canceled by subtracting the output of Anode 2 from that of Anode 1. In addition, as the auxiliary diode has a higher breakdown voltage than the main diode, it operates in linear mode with very low avalanche gain, and generates no avalanche signal even when the main diode operates in Geiger-mode. Therefore, the DA-SPAD-based configuration with the simple subtraction circuit easily discriminates the weak avalanche signals, regardless of the frequency and type of the gate signal. It should be noted that if higher excessive voltage than 4 V, which is the breakdown voltage difference between main and auxiliary diodes in this study, is applied to the main diode, the auxiliary diode also can operate in Geiger-mode. Even in this case, the noise count does not occur due to the avalanche generated in the auxiliary diode. This is because, when avalanche occurs only in the auxiliary diode, a negative pulse is output, in principle, from the subtraction circuit and is not discriminated by the comparator. However, there is a very low probability that avalanche events occur in the main and auxiliary diodes simultaneously, and thus it is impossible to discriminate the

Optics EXPRESS

avalanche of the main diode. To avoid this, the breakdown voltage difference between the two diodes must be designed considering the excess voltage that will be applied to the main diode.

It should be noted that two independent SPAD chips will achieve a similar effect as the proposed DA-SPAD. For example, balanced photodetection techniques have reduced the background noise, enabling the detection of small avalanche signals and reducing the after-pulse probability [10–12]. However, for the two independent SPAD chips, the packaging process of the SPAD module becomes more complicated and the two chips must exhibit almost identical capacitive response characteristics. On the contrary, the electrical and electronic characteristics of the diodes in the proposed DA-SPAD negligibly depend on the position of the fabrication in wafer, because the diodes are directly adjacent to each other. In addition, the configuration of the SPD system (including the packaging of the SPAD module and removal of the capacitive response signal) is greatly simplified in the DA-SPAD.

3. Experimental setup

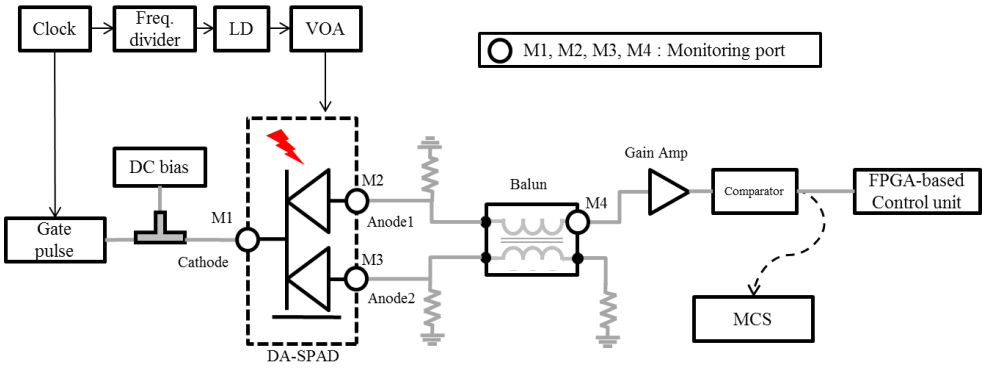

Fig. 2. Experimental setup for testing the cancellation of the capacitive response signal and the Geiger-mode operation of the DA-SPAD-based device. LD: laser diode, ATT: optical attenuator, MCS: multi-scaler, VOA: variable optical attenuator.

The capacitive response of a SPAD operated in gated Geiger-mode was observed to be easily removed by the DA-SPAD, enabling 1 GHz high-speed SPD. Figure 2 schematizes the experimental setup of this study, where “Anode1” and “Anode2” denote the anode ports of the main and auxiliary diodes, respectively. As shown in Fig. 2, the bias voltage for gated Geiger-mode operation was applied to the cathode port of the DA-SPAD by combining the DC voltage and gate signal through the bias Tee. The capacitive response signal of Anode1 was removed from the two anode outputs of the DA-SPAD by a conventional transmission line transformer-type balun. In the present setup, the balun behaved like a differential amplifier with a gain close to 1. One of the two balun outputs was terminated with a load resistor, and several resistors were used for other terminations and impedance matching. As mentioned in the previous section, the auxiliary diode of the DA-SPAD was designed to have the same capacitive response as the main diode, but a higher breakdown voltage than the main diode. Therefore, the auxiliary diode operates in linear mode even when the main diode operates in Geiger-mode. Ideally, the balun outputs a zero-level DC signal when no avalanche occurs in the main diode, and an electrical pulse corresponding to the avalanche signal when an avalanche occurs in the main diode. The open circles in Fig. 2 indicate the monitoring ports for observing the gate input (M1), anode outputs (M2, M3) and balun output (M4).

Optical pulses, required for the performance test of Geiger-mode operation in the DA-SPAD-based device, were generated by a laser diode (LD) and injected into the SPAD. Before injection, the mean photon number per LD pulse was adjusted by a variable optical attenuator. In the Geiger-mode test, a frequency divider is used to set the laser pulse and gate

Optics EXPRESS

frequencies differently for analyzing the after-pulse probability. More information on the Geiger-mode test is detailed in the next section.

4. Experimental results

We first compared the capacitance response signals from the two anode ports of the fabricated DA-SPAD. To effectively remove the capacitive response, the diodes should output almost identical capacitive responses for an arbitrarily shaped gate signal input. In general, when the gate signal is a square wave, the shorter the rising and falling times, the more difficult is the impedance matching of the input and output stages of the SPAD, and the greater the dependence of the diode on the capacitance response characteristics. Here we compared the outputs of Anode1 and Anode2 for two differently shaped square wave gate signals and observed the various frequency components of the capacitive responses of the main and auxiliary DA-SPAD diodes.

Both types of gate signals had the same gating frequency (125 MHz), amplitude (5 V_{pp}) and time duration (2 ns FWHM), but different rising and falling times (~800 ps or 300 ps from 20% to 80% peak level). During these measurements, we monitored the M1, M2 and M3 nodes in Fig. 2 at the room temperature of our laboratory (i.e., the SPAD temperature was not controlled). Panels (a) and (b) of Fig. 3 show the diode output responses to the input square gate signals with rising and falling times of 800 and 300 ps, respectively.

As shown in Fig. 3, the capacitive responses of the two anode ports were negligibly different in both cases of gate input. As expected, the gate signal with the short rising time elicited more higher-frequency components in the response signal than the gate signal with the longer rising time. It should be noted that in the DA-SPAD, similarity of the two anode outputs is more important than the shape of the response. The slight difference between the two anode outputs in response to the shorter rising time (see Fig. 3(b)) is more likely attributable to impedance mismatch of the RF cables and readout circuits used in the measurement than to mismatch of the capacitance responses of the two diodes.

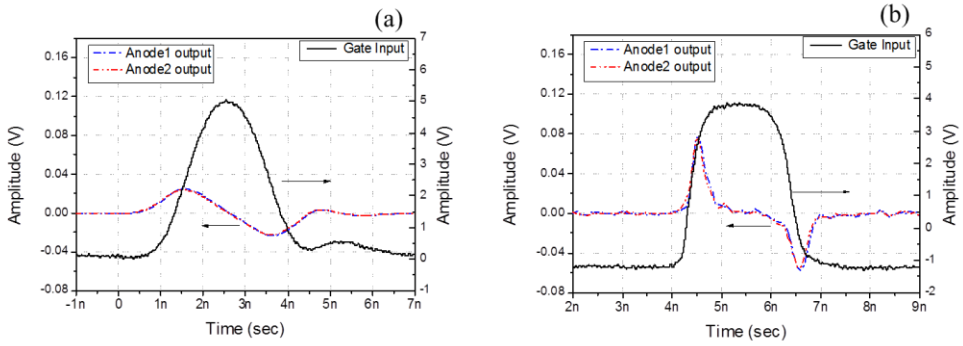

Fig. 3. Capacitance responses of the DA-SPAD to input gate pulses with rise and fall time of (a) 800 ps and (b) 300 ps.

For the two gate signals, we then compared the balun output signals in the cases with and without removal of the capacitive response signal. The signals with the capacitive response removed or not removed were monitored by turning the connection line between Anode2 and the balun input port on and off, respectively. In these measurements, the gate signals and SPAD temperature condition were unchanged from those of the capacitive response comparison. The avalanche signals were observed with the SPAD operated in Geiger-mode. The DC bias voltage was increased, and the output port of the balun was accumulated over several seconds on the oscilloscope screen. Figure 4 shows the measured results. In panels (a) and (c), the captured waveforms were accumulated without removing the capacitive response, and in panels (b) and (d), the capacitive response signals were removed. As no optical pulse

Optics EXPRESS

was input to the SPAD during these measurements, the avalanche signals in Fig. 4 are dark-count events. As shown in the figure, the weak avalanche signals with amplitudes below the amplitude of the capacitive response were easily discriminated by using the DA-SPAD and the simple subtraction circuit.

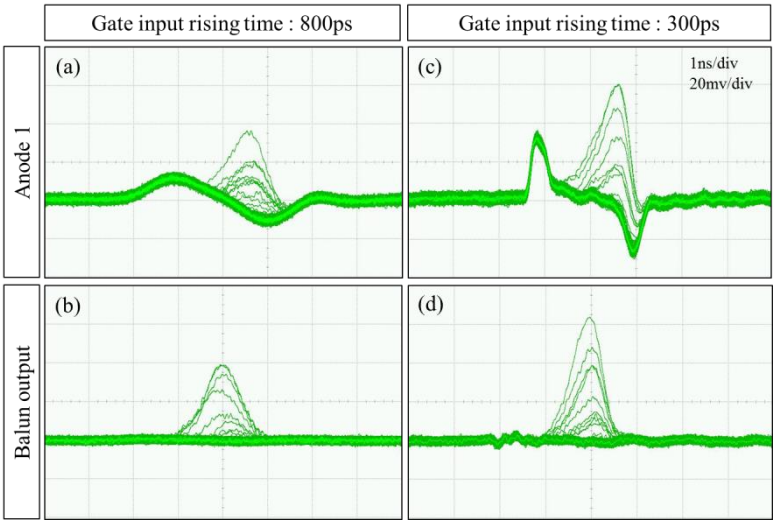

Fig. 4. Avalanche signals with ((b) and (d)) and without ((a) and (c)) cancellation of the capacitive response signals.

Finally, the SPD performance of the fabricated DA-SPAD was observed in Geiger-mode tests. The gating frequency (f_g) was 1 GHz with a gate width of 400 ps (FWHM) and an amplitude of 8 V_{pp}. An LD centered at 1550 nm was synchronously triggered to the gate pulse at a rate (f_l) of 3.125 MHz. The average photon number (μ) was set to 0.1 photons per laser pulse through an optical variable attenuator. The SPAD was maintained at −20 °C by a temperature controller based on a thermo-electric cooler. From the output port of the balun, the avalanche events were discriminated using a conventional gain amplifier and a comparator. The avalanche detection events were counted and analyzed using additional field-programmable gate array (FPGA)-based hardware. The dark-count probability per gate (P_{dark}) was calculated as $P_{dark} = R_{dark} / f_g$, where R_{dark} is the dark-count rate measured without photon injection into the SPAD. The efficiency (η) and after-pulse probability (P_{after}) were respectively calculated as [13,14]

$$\eta = (1/\mu) \times \ln\left[\left(1 - R_{dark}/f_g\right)/\left(1 - R_{\mathrm{det}}^{p}/f_l\right)\right] \quad (1)$$

$$P_{after} = \left[R_{\mathrm{det}} - R_{\mathrm{det}}^{p} - \left(1 - f_l/f_g\right)R_{dark}\right]/R_{\mathrm{det}}^{p} \quad (2)$$

Here, R_{det} is the rate of the detection events at all gates, and R_{det}^{p} is the rate of the detection events occurring at the gates receiving optical pulses only. Both rates were analyzed and monitored by our own FPGA-based SPD control unit. The SPD parameters mentioned above were observed while varying the applied DC bias. The observation results are summarized in Table 1 and plotted in Fig. 5.

When the detection efficiency was below 20%, the after-pulse probability was 3.5% or less, suitable for various SPD applications such as quantum cryptography systems. The count-off time in these measurements was 160 ns, meaning that once an avalanche was triggered, any avalanches during the following 160 ns were ignored [13]. The long count-off time

Optics EXPRESS

relative to the inverse of the fast gate frequency (1 GHz) was attributed to the characteristics of the FPGA-based SPD board used in this study. The FPGA board acquires the avalanche detection events by a "detection and reset" mechanism. After an arbitrary avalanche event, this mechanism ignores the avalanche events until the digital reset signal is generated by the FPGA chip. Therefore, to evaluate the after-pulse probability in a condition of higher counting rate, it is necessary to analyze the after-pulse probability for shorter count-off time than the time applied in this measurement.

Table 1. Performance measures of the DA-SPAD

DC bias (V)	η(%)	P_{dark} (/gate)	P_{after} (%)
52.20	10.19	5.70×10^{-6}	0.69
52.70	15.21	1.07×10^{-5}	1.04
53.27	20.49	2.17×10^{-5}	3.48
53.55	23.04	3.05×10^{-5}	9.31
53.91	25.28	6.72×10^{-5}	21.62

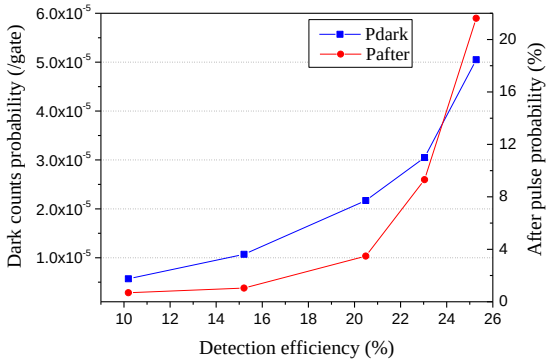

Fig. 5. Dark count and after-pulse probabilities vs detection efficiency of the DA-SPAD.

To analyze the after-pulse probability within a range of count-off times, we measured the after-pulse distribution in the time domain using a multi-scaler (MCS6) manufactured by Fast ComTec. The pulse signals triggering the optical pulse were duplicated and injected into the start port of the multi-scaler, and the comparator output was connected to the stop port of the multi-scaler. In this measurement configuration, the avalanche events discriminated by the comparator could be acquired in the time domain based on the start input. The timing resolution of the multi-scaler was 100 ps, and the histograms were created using the 319-ns acquisition time binned into 1-ns window, starting from the time at which the optical pulse was injected into the gate. The number of sweeps for histograms was 31.25 million times and every avalanche event acquired through the multi-scaler was recorded and accumulated in the time domain. Figure 6(a) shows part of the obtained after-pulse probability distribution in the time domain. Data in Fig. 6(a) indicate the afterpulse probabilities of each gate opened after the optical pulse injection, which were calculated by dividing the number of avalanche events occurred in each gate by the number of avalanche events occurred in the gate where the optical pulse injected. The average number of clicks due to dark count, which was obtained by similar measurement with block of optical pulse injection, was subtracted from the numbers of avalanche events before the division. The after-pulse probability was relatively high up to 5 ns after the optical pulse injection, and exponentially decayed over the measurement period. Figure 6(b) shows the notable part of the estimated after-pulse probabilities as functions of count-off time for different bias voltages. For the Fig. 6(b), all

Optics EXPRESS

the afterpulse probabilities for the given the efficiency, obtained as Fig. 6(a), were added excluding the afterpulse probabilities at the gate within the count-off time region. At the efficiency of 20.49%, the after-pulse probabilities with short count-off times (below 20 ns) were estimated to be much higher than those obtained with 160 ns count-off time. As shown in the figure, the after-pulse probabilities were higher than 6% even for the shorter count-off time than 20 ns. However, when the efficiency was below 10%, the after-pulse probabilities can be reduced to below 2% even with 2 ns deadtime (ignoring one gate after each avalanche event), implying the maximum count rate can be increased up to 500 MHz with lower after-pulse probability than 2%.

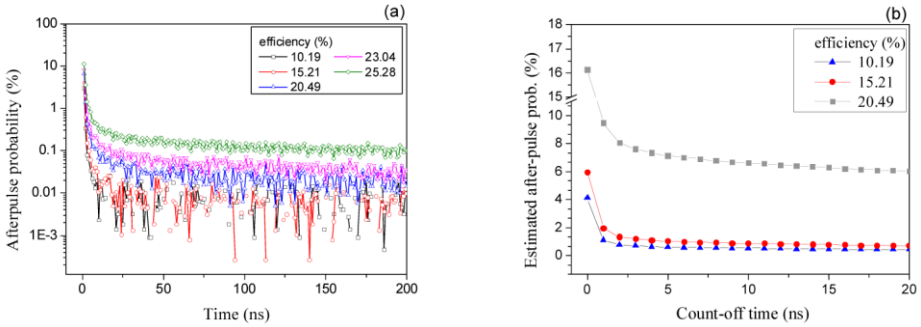

Fig. 6. Time domain analysis of the after-pulse distribution. (a) Measured after-pulse probability distribution. (b) Estimated after-pulse probability versus count-off time at different the efficiencies.

5. Conclusion

This research analyzed a dual anode SPAD (DA-SPAD) with simplified discrimination of weak avalanches. The DA-SPAD easily removed the SPAD capacitive response signal to the gate input. In a 1 GHz gated Geiger-mode test, the proposed DA-SPAD operated in Geiger-mode at high speed and high efficiency, with reasonable low after-pulse probabilities. This result confirms the weak avalanche detectability of the proposed device. The DA-SPAD removes the capacitive response irrespective of the frequency and waveform of the gate signal, enabling dramatic simplification of SPAD-based SPD systems for future biomedical and communication applications.

Funding

ICT R&D Program of MSIP/IITP (1711073835, Reliable crypto-system standards and core technology development for secure quantum key distribution network).

References

1. D. Palubiak, M. M. El-Desouki, O. Marinov, M. J. Deen, and Q. Fang, "High-speed, single-photon avalanche-photodiode imager for biomedical applications," IEEE Sens. J. **11**(10), 2401–2412 (2011).
2. G. Boso, D. Ke, B. Korzh, J. Bouilloux, N. Lange, and H. Zbinden, "Time-resolved singlet-oxygen luminescence detection with an efficient and practical semiconductor single-photon detector," Biomed. Opt. Express **7**(1), 211–224 (2016).
3. A. Migdall and J. Dowling, "Introduction to journal of modern optics special issue on single-photon: detectors, applications, and measurement methods," J. Mod. Opt. **51**(9–10), 1265–1266 (2004).
4. A. L. Lacaita, P. A. Francese, S. D. Cova, and G. Riparmonti, "Single-photon optical-time-domain reflectometer at 1.3 Mum with 5-cm resolution and high sensitivity," Opt. Lett. **18**(13), 1110–1112 (1993).
5. C. Yu, M. Shangguan, H. Xia, J. Zhang, X. Dou, and J. W. Pan, "Fully integrated free-running InGaAs/InP single-photon detector for accurate lidar applications," Opt. Express **25**(13), 14611–14620 (2017).
6. M. Liu, C. Hu, J. C. Campbell, Z. Pan, and M. M. Tashima, "Reduce after-pulsing of single-photon avalanche diodes using passive quenching with active reset," IEEE J. Quantum Electron. **44**(5), 430–434 (2008).
7. Z. L. Yuan, B. E. Kardynal, A. W. Sharpe, and A. J. Shields, "High speed single photon detection in the near infrared," Appl. Phys. Lett. **91**(4), 041114 (2007).

Optics EXPRESS

8. N. Namekata, S. Sasamori, and S. Inoue, "800 MHz Single-photon detection at 1550-nm using an InGaAs/InP avalanche photodiode operated with a sine wave gating," Opt. Express **14**(21), 10043–10049 (2006).
9. Z. L. Yuan, A. W. Sharpe, J. F. Dynes, A. R. Dixon, and A. J. Shields, "Multi-gigahertz operation of photon counting InGaAs avalanche photodiodes," Appl. Phys. Lett. **96**(7), 071101 (2010).
10. A. Tomita and K. Nakamura, "Balanced, gated-mode photon detector for quantum-bit discrimination at 1550 nm," Opt. Lett. **27**(20), 1827–1829 (2002).
11. Z. Lu, W. Sun, Q. Zhou, J. Campbell, X. Jiang, and M. A. Itzler, "Improved sinusoidal gating with balanced InGaAs/InP Single Photon Avalanche Diodes," Opt. Express **21**(14), 16716–16721 (2013).
12. C. Scarcella, G. Boso, A. Ruggeri, and A. Tosi, "InGaAs/InP Single-Photon Detector Gated at 1.3 GHz With 1.5% Afterpulsing," IEEE JST QE **21**(3), 17–22 (2015).
13. J. Zhang, R. Thew, C. Barreiro, and H. Zbinden, "Practical fast gate rate InGaAs/InP single-photon avalanche photodiodes," Appl. Phys. Lett. **95**(9), 091103 (2009).
14. S. B. Cho and S. K. Kang, "Weak avalanche discrimination for gated-mode single-photon avalanche photodiodes," Opt. Express **19**(19), 18510–18515 (2011).