JOURNAL OF MODERN OPTICS
https://doi.org/10.1080/09500340.2020.1817591

Taylor & Francis
Taylor & Francis Group

Ultra-low dead time free-running InGaAsP single-photon detector with active quenching

Junliang Liu[a,b], Yining Xu[b], Yongfu Li[b,c], Yi Gu[d], Zhaojun Liu[a,b] and Xian Zhao[b,c]

[a]School of Information Science and Engineering, Shandong University, Qingdao, People's Republic of China; [b]Key Laboratory of Education Ministry for Laser and Infrared System Integration Technology, Shandong University, Qingdao, People's Republic of China; [c]Center for Optics Research and Engineering, Shandong University, Qingdao, People's Republic of China; [d]Key Laboratory of Infrared Detection and Imaging Technology, Shanghai Institute of Technical Physics, Chinese Academy of Sciences, Shanghai, People's Republic of China

ABSTRACT
High afterpulse probability and the consequent long dead time of free-running InGaAs(P) single-photon detectors have long been the limiting factors to their practical applications. Here we present a free-running InGaAsP single-photon detector for 1.06 μm wavelength with ultra-low dead time and afterpulse probability. With the optimized active-quenching circuit and packaging, the avalanche pulse discrimination level is down to 2.4 mV, and the full-width at half-maximum of the avalanche pulse is as short as 500 ps, which greatly lessens afterpulsing effects. As a result, the dead time of the detector is as low as 35 ns, comparable to free-running silicon detectors, with a low afterpulse probability of 11.6% at 10% detection efficiency, 242 K. The proposed detector provides a high-performance, small-sized, low-power approach to single-photon detection for practical applications such as light detection and ranging and free-space quantum key distribution systems.

ARTICLE HISTORY
Received 28 June 2020
Accepted 24 August 2020

KEYWORDS
Single-photon detectors; single-photon avalanche diodes; avalanche photodiodes; afterpulsing; quenching circuits

1. Introduction

Lower dead time and higher count rate have been the vigorous pursuit for single-photon detectors (SPDs) as they are widely used in applications such as quantum key distribution (QKD) [1] and light detection and ranging (LiDAR) [2] systems. In systems with working distance as long as hundreds kilometres, 1.06 μ m laser is preferred for its higher pulse energy. For the detection of this wavelength, traditional silicon-based SPDs and photo-multiplier tubes are not preferred for its low detection efficiency of about 2%. Recently, superconducting SPDs have demonstrated an excellent overall performance for such applications [3], and yet the requirement of cryostats limits its practical use in mobile and size-limited environments. InGaAsP/InP single-photon avalanche diodes (SPADs) have been favoured for use with 1.06 μ m lasers for its compact size and relatively high detection efficiency of over 10%. However, severe afterpulsing effects in these devices result in a long hold-off (dead) time required, and hence has limited its use in free-running mode and the use with high-repetition-frequency lasers.

Many techniques and devices have been developed to lessen afterpulsing effects in InP-based SPADs. Gated passive quenching circuits have demonstrated excellent afterpulsing performance [4], with a dead time of typically several or tens of nanoseconds. However, they require the incident photon to arrive within the short gate-on time (typically less than 1 ns); otherwise it leads to a low detection efficiency of only about 6% at 1.06 μ m for free-running applications [5]. Small-sized and easy-to-use InGaAs/InP negative-feedback avalanche diode (NFAD) [6] was proposed, followed by a similar device with shorter recovery time [7], where the SPAD is passively but quickly quenched. NFADs have demonstrated an afterpulse probability of 13.2% at 0.6 µs [8], or above 50% at 100 ns [9]. Recently, the SPAD based on InGaAs-GaAs nanowires demonstrates a concept of limiting the avalanche within each nanowire, opening a new gate to high-performance InGaAs SPADs [10]. Besides, Ge-on-Si SPADs work in NIR wavelengths with silicon multiplication layer to achieve considerably reduced afterpulsing effects compared with InGaAs SPADs, but they have a higher dark count rate, and still require cryostats [11].

Nevertheless, conventional InGaAs(P)/InP SPADs have not been exploited to their extreme. A shorter quenching delay can effectively lower the number of avalanche carrier by interrupting the development of

CONTACT Yongfu Li ✉ yfli@sdu.edu.cn 🖃 Key Laboratory of Education Ministry for Laser and Infrared System Integration Technology, Shandong University, 72 Binhai Road, Qingdao 266237, People's Republic of China; Center for Optics Research and Engineering, Shandong University, 72 Binhai Road, Qingdao 266237, People's Republic of China

avalanche earlier, resulting in a lower afterpulse probability [12]. Starting from this concept, we have demonstrated free-running SPDs equipped with a fast active-quenching circuit [13] and a faster one [14]. The latter has demonstrated an afterpulsing performance of 4% at 10% detection efficiency, 1 µs hold-off time, or 20% at 100 ns hold-off time, superior to NFADs.

In this paper, we present a free-running InGaAsP SPD with superior performance and smaller size, striking both ultra-low hold-off time and low afterpulse probability. The improvements attributes to an even faster active-quenching circuit, and the integrated packing of the SPAD and some other key components. Since our previous work has demonstrated an afterpulsing performance better than other InP-based SPDs, we believe that the SPD we present here sets a new milestone.

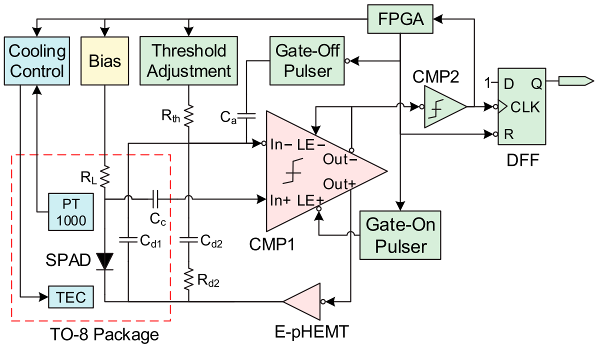

Figure 1. Simplified schematic of the active-quenching circuit. LE: latch enable. CMP: comparator.

2. Detector design

InGaAsP/InP SPAD (PGA-284) from Princeton Lightwave with 80 µm effective diameter is used in the SPD, and its function is under the control of the specially designed active-quencing circuit shown in Figure 1. The SPAD is biased by an programmable negative voltage source through a current-limiting resistor R_L (10 kΩ). The voltage at the anode of the SPAD is compared with an adjustable threshold by a SiGe high-speed comparator (CMP1, HMC674 from Analog Devices) to detect avalanche. Once an avalanche is detected, a low-to-high transition is triggered at the positive output of CMP1, and it is inversely amplified by a GaAs enhancement mode pseudomorphic high-electron-mobility transistor (E-pHEMT, ATF-531P8 from Avago Technologies) to quench the SPAD. After quenching, the negative output of CMP1 goes high and latches CMP1 to its current output state, which holds the SPAD off. Since the comparator itself acts as the hold-off circuit, there are only two active components, CMP1 and the E-pHEMT, in the feedback loop, which minimizes the quenching delay. Meanwhile the signal at the negative output of CMP1 is translated by another fast comparator (CMP2, ADCMP572 from Analog Devices), and is sent to a field-programmable gate array (FPGA) for an adjustable hold-off timing of 35 ns to 300 µs. When the hold-off time elapses, the FPGA triggers the Reset Pulser to generate a short pulse to disable the latch function of CMP1, and the SPAD is reset.

The basic structure and function of the circuit described above is similar to our previous work [14], but we have made several significant optimizations and extensions based on it. First of all, the comparator is upgraded to one that incorporates reduced swing positive emitter-coupled logic (RSPECL) output, with an ultra-low propagational delay of 80 ps. The RSPECL output stage has a low output impedance of only a few ohms rather than 50 Ω. Since the E-pHEMT has a large input capacitance, the low output impedance effectively shortens the quenching delay. Secondly, the 50 Ω bias resistor of the E-pHEMT is distantly connected to the SPAD through an impedance-matched cable, and its output polarity is set to enable the SPAD when no current flows through the bias resistor, which significantly decreases the power consumption and heat dissipation. Thirdly, a gate-off pulser is added to the inverting input of CMP1 to trigger a quenching process when required, and a D-type flip-flop (DFF) register is inserted into the output path to block the triggered false output signal by forcing it to output '0' before gate-off process. The added circuitry enables true gated operation of the SPD without affecting the quenching performance.

Most importantly, we have improved the dummy-capacitor method which prevents false triggering during the reset process previously used by Tosi et al. [15]. Originally, the method requires a dummy capacitor C_{d1} of the same capacitance (0.4 pF for PGA-284) as the SPAD, and it should be placed between the output of the quenching transistor and the inverting input of CMP1 for our case. We add an RC network of C_{d2} (0.7 pF) and R_{d2} (24 kΩ) in parallel with the dummy capacitor. As shown in Figure 2, when resetting the SPAD, adding this RC network slightly increases the amplitude and relaxing time of the signal of the dummy path at the inverting input of CMP1. The improvement allows more tolerance to the noise residue of the reset signal due to the mismatch between the SPAD and the dummy path, at the cost of a slight increase of the minimum hold-off time. Moreover, we create a common electromagnetic environment for the SPAD and the dummy capacitor by putting them close to each other within a hermetically sealed TO-8 package. With all the efforts, the threshold of CMP1 for

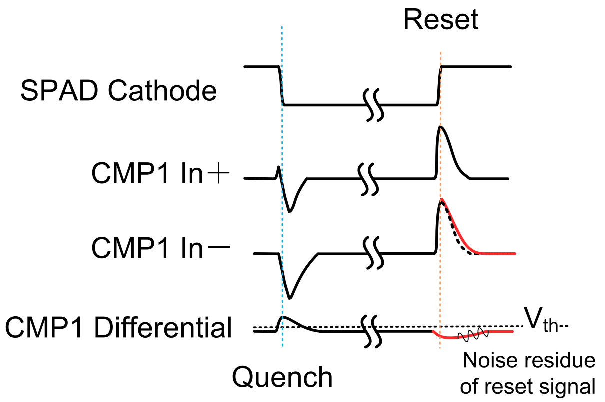

Figure 2. Waveforms at critical nodes of the circuits. CMP1 is the comparator for avalanche current pick-up. For inverting input of CMP1 (In−), the solid line at the time of reset shows the result of the improved circuit, while the dashed line is for the original design.

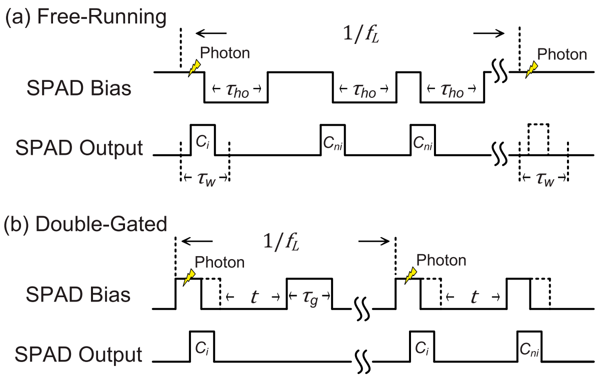

Figure 3. Timing diagram for afterpulsing measurement of (a) free-running and (b) gated operation of the detector. The dashed line in (b) shows the bias of SPAD under the situation when no avalanche occurred within the gate, and the solid line demonstrates that the bias is actively quenched after detection.

avalanche discrimination is down to 2.4 mV. This value is reaching the physical limit of 2 mV, because the hysteresis of the comparator is typically ±1 mV. A 2-stage Peltier thermo-electric cooler (TEC) and a PT1000 resistance temperature detector (RTD) are also installed in the TO-8 package, making the SPD with its heat dissipation structure as compact as fist-sized.

3. Experimental setup

The performance of the proposed SPD is mainly characterized by dark count rate (DCR) and afterpulse probability at different detection efficiencies (PDE) and hold-off times. A time-correlated single-photon counting (TCSPC) system is built for the performance assessment. The photon source used in the experiments is a 1.06 µm pulsed laser diode (PDL 800-B from PicoQuant) with a repetition rate of 39 kHz, and is attenuated to 0.1 photon per pulse on average. The photon source is focused on the active area of the detector through a microscope objective. A dual-channel time-to-digital converter with 36 ps resolution and 3.75 ns dead time, and a gated counter are embedded in the FPGA of the SPD, ready for low-dead-time TCSPC data acquisition.

The total afterpulse probability (P_{tap}) and the DCR (C_d) is measured under free-running mode of the SPD. Taking the hold-off time into consideration, C_d is calculated by

$$C_d = \frac{C_{\mathrm{dr}}}{1 - \tau_{\mathrm{ho}} C_{\mathrm{dr}}}, \tag{1}$$

where τ_{ho} is the hold-off time, and C_{dr} is the raw count rate without illumination.

P_{tap} is evaluated by the increase of dark counts after the detection of a photon, compared with previously acquired number of dark counts. As shown in Figure 3(a), we define C_i as the count rate within the window of illumination in the histogram collected by the TCSPC, τ_w as the width of the window, C_{ni} as the count rate outside the window. Instead of using the straightforward $(C_{ni} - C_{\mathrm{dr}})/C_i$ to calculate P_{tap}, considering the fact that afterpulses may be masked by the dead time induced by any count event, we have modified the equation as follows to avoid an underestimation of the afterpulse probability,

$$P_{\mathrm{tap}} = \frac{C_{ni} - C_d\left[1 - \tau_{\mathrm{ho}}\left(C_{ni} + C_i\right) - \tau_w\left(f_L - C_i\right)\right]}{C_i - \tau_w f_L C_d}, \tag{2}$$

where f_L is the repetition rate of the pulsed laser.

Since the detector also supports gated operation, the time distribution of the afterpulse probability can be assessed by using double-gated method [16], as shown in Figure 3(b). There are two gates that enable the SPAD within a laser period: the first gate is the gate synchronized to the incident photon, and the second gate is opened after a certain time t for the measurement of afterpulses induced by the avalanche generated in the first gate. The width of all gates is set to $\tau_g = 5$ ns. Note that the actual gate width can be shorter than the set value because of the active quenching after the detection within any gate. The afterpulse per unit time can be calculated using

$$P_{\mathrm{ap}}(t) = \frac{1}{\tau_g} \times \frac{C_{ni}(t) - \tau_g f_L C_d}{C_i - \tau_g f_L C_d}, \tag{3}$$

where C_i is the count rate within the illuminated gate, and $C_{ni}(t)$ is the count rate within the non-illuminated gate.

Besides, in order to characterize the avalanche of the SPAD and the active-quenching process of the circuit without any probe insertion, electroluminescence method [14] is adopted. The method requires a TCSPC to record the time difference between the detection output of the SPD under test and the output of another SPD detecting the electroluminescent photons of the SPAD in the SPD under test, and the recorded histogram represents the avalanche waveform. Our previously built InGaAsP SPD and a high-resolution low-jitter time-to-digital converter (quTAG from quTools) act as the TCSPC for the measurement. The electroluminescent photons generated by dark-count avalanche pulses is measured in order to simplify the setup.

4. Results and discussion

The PDE of the detector is up to 35% at an excess bias of 5.6 V, and the DCR is approximately 10 kc/s at 35% PDE. The SPAD device has a quantum efficiency of 75%, and hence the maximum PDE of the detector could possibly be higher than 40% [*17*] if a higher maximum excess bias were achieved. This requires a higher breakdown voltage of the transistor for quenching signal amplification, which is an critical component of the quenching circuit.

However, compound-semiconductor-based, high-speed and easy-to-drive transistors usually have a low maximum operating voltage, e.g. 7 V for the GaAs E-pHEMT used in this work. As for other available transistors, silicon-based ones are not as fast as compound ones; GaN field-effect transistors (FETs) have a much higher breakdown voltage and are fast enough, but commercially available fast GaN FETs are all for medium or high power radio-frequency (RF) applications, with relatively larger input capacitance and lower trans-conductance gain. Using these alternatives may result in an increase of quenching delay, which has a negative impact on afterpulse suppression. Thus, a low-voltage high-speed transistor such as GaAs E-pHEMT serves better as a low-delay quenching signal amplifier, at the cost of lower maximum available PDE.

Besides, a PDE of above 25% requires the hold-off time to be set to 5 µs at minimum to achieve an acceptable afterpulse probability. Since the SPD is designed to run at a low hold-off time starting from 35 ns, only performance specifications with PDE below 25% are demonstrated next.

The DCR of the detector as a function of hold-off time were measured at 10%, 15%, 20%, 25%, respectively. As shown in Figure 4, at 10% PDE, the DCR rises only 200 cps from 800 cps with the decrease of hold-off time from 20 µs to 35 ns, owing to a good suppression of afterpulsing effects. DCR soars when PDE is above 20% and hold-off time is lower than 500 ns because of the aggravation of afterpulsing, especially higher-order afterpulsing effects.

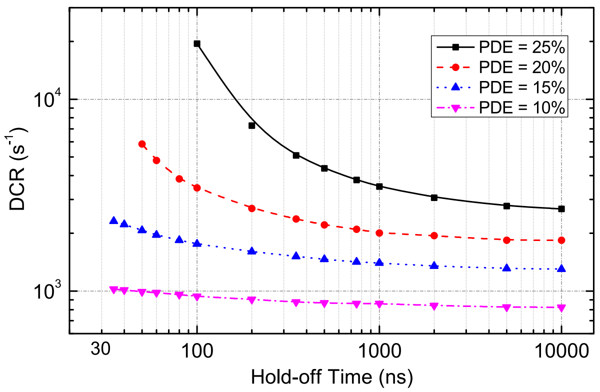

Figure 4. Dark count rate as a function of hold-off time at different photon detection efficiencies, 242 K.

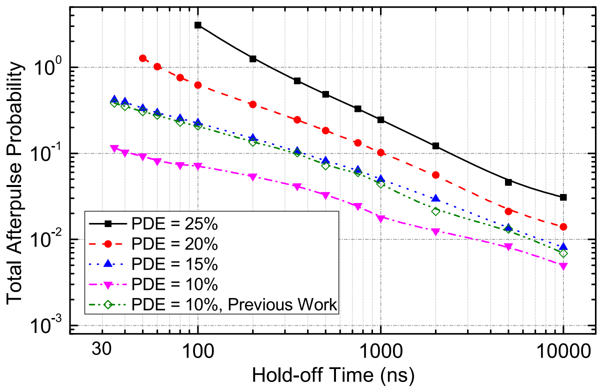

Figure 5. Total afterpulse probability (P_{tap}) as a function of hold-off time at different photon detection efficiencies, 242 K.

In order to further obtain an operable region of the parameters for the detector in free-running mode, we examined the total afterpulse probabilities (P_{tap}) as a function of hold-off time, as shown in Figure 5. We also measured the P_{tap} of our previous work [*14*] at the same temperature and 10% PDE as a comparison, and the result was quite close to the result at 15% PDE of our new design. The total afterpulse probability remains below 30% for our new SPD when PDE is no more than 15% at a hold-off time of 60 ns. A low afterpulse probability of 11.6% was achieved at a hold-off time of 35 ns and 10% PDE. Note that 35 ns is approximately the dead time of commercial silicon SPADs and Sanzaro's work [7], and is lower than that of NFADs which is approximately 100 ns [6]. However, at PDE higher than 20%, hold-off time of 200 ns or below is still not practicable, as P_{tap} could be higher than 100%, bringing more than one afterpulse due to higher-order afterpulsing effects. The optimal PDE is 10% for the full hold-off time range, and is up to 25% when hold-off time is above 2 µs. In general, a PDE of no more than 15% is favourable in low-dead-time free-running applications considering both DCR

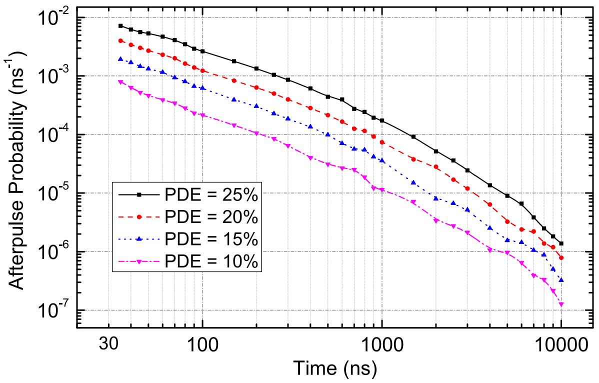

Figure 6. Afterpulse probability per ns as a function of time after previous avalanche at different photon detection efficiencies, 242 K.

and P_{tap}, and parameter optimization depends on specific applications.

Since the new circuitry has enabled gated operation, one is likely to be interested in the afterpulse probability density. The afterpulse probability per nanosecond as a function of time after the previous detection is shown in Figure 6. The afterpulse probability per ns stays below 10^{-2} for the PDE of 25%, and even 10^{-3} for the PDE of 10% within the full operating region. With such low afterpulse probability and flexible gating scheme, the minimum hold-off time of 35 ns is fully functional.

The significant improvement of afterpulsing performance is undoubtedly attributed to the faster discrimination and quenching of the avalanche, which can be proved by the avalanche waveforms of the SPAD shown in Figure 7. The waveform from this work with discrimination levels of 2.4 and 5.8 mV are drawn in solid line and dashed line, respectively, and that of our previous work at its minimum discrimination level of 5.8 mV is in dotted line. The waveforms are aligned to the time of discrimination, defined as time zero. The full-width at half-maximum of the avalanche pulse of this work is as short as 500 ps, and the total duration of the avalanche pulse is approximately 1.2 ns, at its minimum discrimination level of 2.4 mV. Unlike NFAD and similar self-quenching devices, the quenching delay (excluding the time need for detection) of the proposed detector is almost independent to the excess bias, and there is no 'persistent current flow' following the avalanche pulse [6,7], making it ideal for low-dead-time operation.

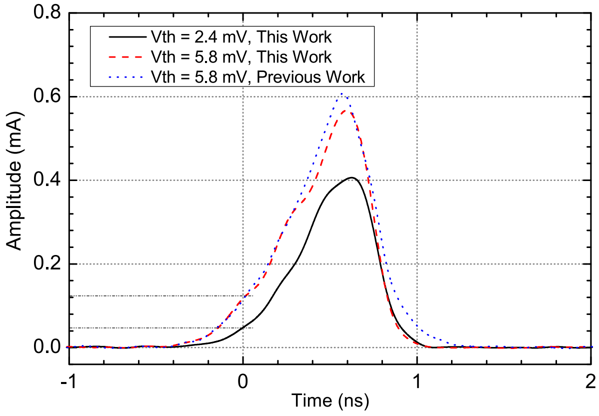

Figure 7. Avalanche waveforms of this work at 2.4 mV and 5.8 mV discrimination levels and our previous work, all at 20% PDE, 291 K.

Compared with our previous work, as can be seen from the waveforms in Figure 7, the benefit of lower discrimination level is apparent: the earlier the avalanche is detected, the earlier it is quenched. Besides, the waveform of our previous work is slightly wider than that of the new one with the same discrimination level, which indicates a faster quenching of the new circuit. This is the merit of having a lower output impedance of the output driver of CMP1.

The faster quenching effectively limits the number of avalanche charge flow, to which afterpulse probability density is proportional [16]. The number of charge is readily obtained by integrating the waveform under each condition. The area under the dashed line is 85.4% of the area under the dotted line, contributed by faster quenching speed. The area under the solid line is 82.3% of the dashed line, contributed by lower discrimination level. In total, the number of avalanche carriers per avalanche of this work is 70.3% of our previous work. Though it appears to be a small reduction in the value, it actually contributes to a large decrease of total afterpulse probability at low hold-off time, e.g. a 70% reduction at 35 ns hold-off time, 10% PDE. This is because higher-order afterpulses are easily triggered by previous afterpulses as the hold-off time decreases, and results in a significantly larger afterpulse probability [18].

In addition, a faster quenching and the integrated cooling also contributes smaller thermal saturation effect, which may degrade the maximum count rate and detection efficiency of the detector [19]. As the minimum dead time is as low as 35 ns, the maximum count rate can be up to 28 Mcps. Nevertheless, the peak current and avalanche duration are as low as approximately 0.4 mA and 1 ns, respectively, thanks to the fast active quenching and the small effective junction capacitance of the SPAD device of 0.4 pF. Thus, the maximum average current and the total power dissipation of the device are estimated to be only 5.6 μ A and 0.47 mW, respectively. The closed-loop cooling is capable of handling the heat dissipation of the SPAD of 0.47 mW and keep a stable working temperature of the SPAD, since the PT1000 RTD is located adjacent to the SPAD device on the same metallized surface of the TEC. Thus, the performance of the detector is less likely to be affected by the thermal saturation effect.

5. Conclusion

In conclusion, we have designed a free-running InGaAsP SPD for 1.06 μ m wavelength with a typical performance of a low afterpulse probability of 11.6% and a low DCR of 1 kcps at 10% PDE and 35 ns hold-off time. Results show that both ultra-low afterpulse probability and hold-off time are achieved owing to the ultra-low discrimination threshold and quenching delay, allowing lower hold-off time. The proposed gating-compatible free-running InGaAsP SPD provides a high-performance, small-sized, low-power approach to 1.06 μ m single-photon detection for applications such as free-space quantum key distribution and wide-range LiDAR, where true free-running SPDs are needed. It is worth mentioning that the circuitry and method can be applied to InGaAs SPADs for 1.31 and 1.55 μ m applications as well, and even better afterpulse performance could be expected since they usually have a much smaller active area.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work is supported by Opening Foundation of the State Key Laboratory of Functional Materials for Informatics, Shanghai Institute of Microsystem and Information Technology, CAS, and in part by National Defense Science and Technology Innovation Special Zone Foundation of China.

References

[1] Yuan ZL, Dixon AR, Dynes JF, et al. Gigahertz quantum key distribution with InGaAs avalanche photodiodes. Appl Phys Lett. 2008;92(20):201104.

[2] Kuzmenko K, Vines P, Halimi A, et al. 3D LIDAR imaging using Ge-on-Si single-photon avalanche diode detectors. Opt Express. 2020 Jan;28(2):1330–1344. Available from: http://www.opticsexpress.org/abstract.cfm?URI=oe-28-2-1330

[3] Marsili F, Verma V, Stern J, et al. Detecting single infrared photons with 93% system efficiency. Nat Photon. 2012;7(3):210–214.

[4] He DY, Wang S, Chen W, et al. Sine-wave gating ingaas/inp single photon detector with ultralow afterpulse. Appl Phys Lett. 2017;110(11):111104.

[5] Scarcella C, Boso G, Ruggeri A, et al. InGaAs/InP single-photon detector gated at 1.3 GHz with 1.5% afterpulsing. IEEE J Sel Topics Quant Electron. 2015 May;21(3):17–22.

[6] Itzler MA, Jiang X, Nyman B, et al. InP based negative feedback avalanche diodes. Proc SPIE. 2009;7222:72221K.

[7] Sanzaro M, Calandri N, Ruggeri A, et al. InGaAs/InP SPAD with monolithically integrated zinc-diffused resistor. IEEE J Quant Electron. 2016;52(7):4500207.

[8] Yu C, Shangguan M, Xia H, et al. Fully integrated free-running InGaAs/InP single-photon detector for accurate lidar applications. Opt Express. 2017 Jun;25(13):14611–14620. Available from: http://www.opticsexpress.org/abstract.cfm?URI=oe-25-13-14611

[9] Itzler MA, Jiang X, Entwistle M, et al. Advances in InGaAsP-based avalanche diode single photon detectors. J Mod Opt. 2011;58:174–200.

[10] Farrell AC, Meng X, Ren D, et al. InGaAs/GaAs nanowire avalanche photodiodes toward single-photon detection in free-running mode. Nano Lett. 2019;19(1):582–590.

[11] Vines P, Kuzmenko K, Kirdoda J, et al. High performance planar germanium-on-silicon single-photon avalanche diode detectors. Nat Commun. 2019;10:1086.

[12] Wayne MA, Restelli A, Bienfang JC, et al. Afterpulse reduction through prompt quenching in silicon reach-through single-photon avalanche diodes. J Lightwave Technol. 2014;32(21):4097–4103.

[13] Liu J, Li Y, Ding L, et al. Fast active-quenching circuit for free-running InGaAs(P)/InP single-photon avalanche diodes. IEEE J Quantum Electron. 2016;52(10):4000306.

[14] Liu J, Zhang T, Li Y, et al. Design and characterization of free-running InGaAsP single-photon detector with active-quenching technique. J Appl Phys. 2017;122(1):013104.

[15] Tosi A, Frera AD, Shehata AB, et al. Fully programmable single-photon detection module for InGaAs/InP single-photon avalanche diodes with clean and sub-nanosecond gating transitions. Rev Sci Instrum. 2012;83(1):013104.

[16] Jensen KE, Hopman PI, Duerr EK, et al. Afterpulsing in geiger-mode avalanche photodiodes for 1.06 μm wavelength. Appl Phys Lett. 2006;88(13):133503.

[17] Comandar LC, Fröhlich B, Dynes JF, et al. Gigahertz-gated InGaAs/InP single-photon detector with detection efficiency exceeding 55% at 1550 nm. J Appl Phys. 2015;117(8):083109.

[18] Wang FX, Chen W, Li YP, et al. Non-Markovian property of afterpulsing effect in single-photon avalanche detector. J Lightwave Technol. 2016;34(15):3610–3615.

[19] Kreger ST, Brown TG. Solid state photon counters for scanned image acquisition: thermal and electronic saturation effects. Proc SPIE. 1997;2984:178–189.