Avalanche discrimination and high-speed counting in periodically gated single-photon avalanche diodes

Alessandro Restelli*[a], Joshua C. Bienfang[a]
[a]Joint Quantum Institute, National Institute of Standards and Technology and University of Maryland

ABSTRACT

We discuss avalanche discrimination in a periodically-gated InGaAs/InP single-photon avalanche diode. We investigate the interrelation between the minimum detectable avalanche charge and the detection efficiency, and we show that the technical solutions we implement can improve performance. Gating the detector at 1.25 GHz, single-photon count rates above 250×10^6 s^{-1} can be obtained while maintaining low afterpulse probability with detection efficiencies larger than 0.10.

Keywords: Single photon detection, self-differencing, afterpulsing, avalanche photodiodes.

1. INTRODUCTION

In low-probability and post-selective experiments such as high-order entanglement and multi-photon coincidence detection, it is necessary to maintain stable experimental conditions for several hours, or even days, to collect a statistically relevant amount of data. In such experiments, every incremental increase in the generation and detection rate can significantly benefit the overall system, and in some cases make an otherwise impractical experiment feasible. These requirements are becoming more and more important as complex quantum communication protocols [1] are proposed and tested. Solutions to increase the maximum single-photon count rate are being pursued for a wide variety of single-photon detectors.

Single-photon avalanche diodes (SPAD) are commonly used in quantum-optics experiments [2]. InGaAs/InP SPADs offer single-photon detection efficiency in the spectral bandwidth between 0.9 μm and 1.7 μm, a region that is rich with possibilities for quantum optics experiments. Traditional limitations in maximum count rate due to afterpulsing [3] in InGaAs/InP detectors have been recently overcome by electronic systems that significantly reduce the avalanche current [4–9], allowing gating the detector at gigahertz rates. In this regime of operation the avalanche current is significantly lower than the displacement current associated with the gate electrical pulse. Discrimination techniques are therefore required to isolate the avalanche current from the gate, and can be facilitated by using a periodic electrical gate. For electrical gates with sinusoidal waveform it is possible to remove the gate signal with a notch filter [4]. For gates with a more spectrally-complex shape a technique called self-differencing was recently introduced and is based on subtracting the signal from a replica delayed by exactly one gate period [5].

The goal of our research is to design a high-speed detection system for high-throughput quantum-optics experiments, and we chose the self-differencing technique to obtain high-speed performance. In section 2 we give a description of the electronic boards that we designed for gating and sensing the SPAD, and we focus on the critical components that allow improved avalanche-discrimination sensitivity. Section 3 discusses the linearity, detection efficiency, and afterpulse probability of the system at high count rates.

2. HARDWARE DESIGN

2.1 System design

The system described in figure 1 is designed for compactness, re-configurability, and flexibility.

*Alessandro.restelli@nist.gov; phone 1 301 975 5686

Advanced Photon Counting Techniques VI, edited by Mark A. Itzler, Joe C. Campbell, Proc. of SPIE
Vol. 8375, 83750Z · © 2012 SPIE · CCC code: 0277-786X/12/$18 · doi: 10.1117/12.919803

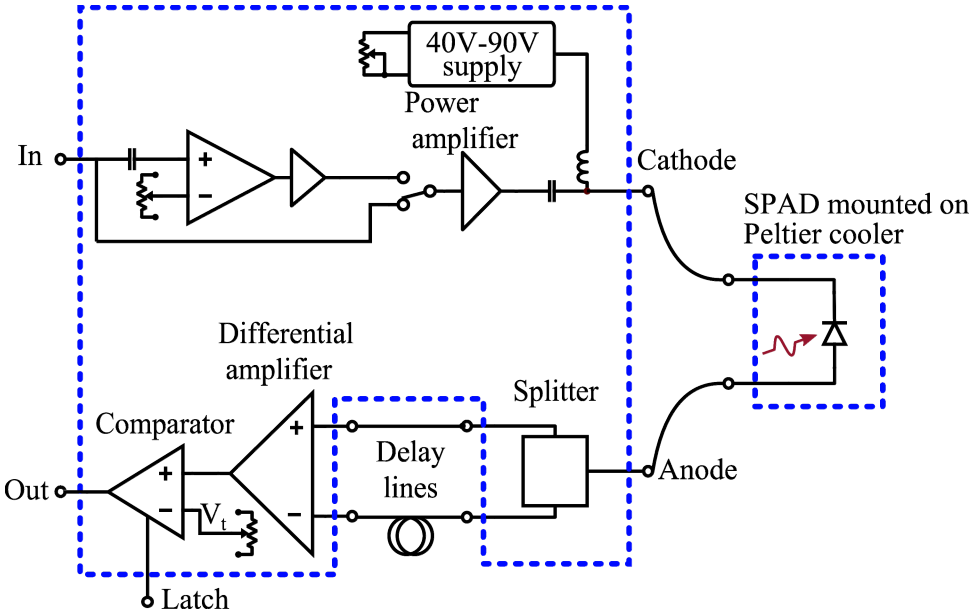


Figure 1. Block schematic of the SPAD gating system. The dashed lines delimit printed circuit boards. The main board contains two identical copies of all the electronics in order to control a second SPAD, here only one instance of the circuitry is shown. Also "Latch" and "Out" are differential lines, but for simplicity they are shown as single-ended here.

A printed circuit board (PCB) is populated with electronics used to form the periodic gate, sense the avalanche current, and discriminate the avalanche signal using the self-differencing technique. Only the delay lines and SPAD are outside the PCB. We chose to construct the delay lines with rigid coaxial cables, which are easily interchangeable with SMA connectors, in order to reconfigure the gate period as desired. The SPAD is connected to the main board through coaxial cables so that it can be conveniently placed inside a dry box and cooled with a triple-stage Peltier cooler down to -30 ˚C.

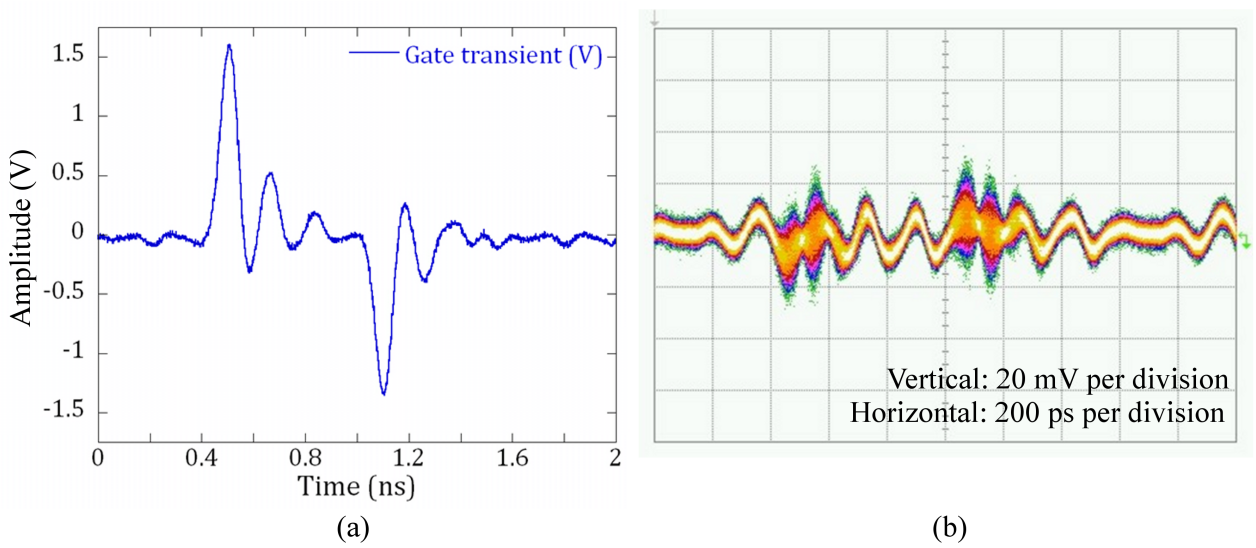
Following figure 1 from the top-left, the board requires an external clock to operate, entering the "in" terminal. To achieve a high degree of sensitivity, the clock signal must be carefully matched to the delay difference between the delay lines. Stability better than 10^{-5} and low jitter are also important for adequate avalanche discrimination. A 10 GHz power amplifier raises the input amplitude from 1 V to up to 12 V peak to peak to bias the SPAD at convenient levels of overvoltage and undervoltage. The power amplifier is a monolithic microwave integrated circuit (MMIC) with good linearity and preserves the shape of the input signal. Thus, when the clock source is sinusoidal, the system applies sinusoidal gates to the SPAD. Alternatively, we can switch in a high-speed comparator followed by 10 GHz-bandwidth preamplifier to form a square gate signal from a sinusoidal input. In this case, changing the comparator threshold allows us to select the duty cycle of the square waveform. A bias tee combines the gate waveform with a DC bias voltage (also generated on the board) that can be set from 40 V to 90 V. In the design of the coaxial connection between the SPAD and the board, significant care was given to minimize the series inductance and to keep the impedance close to 50 Ω all the way to the trimmed terminals of the SPAD.

From the SPAD another coaxial cable brings the anode signal back to the main board where it enters a resistive power splitter. As mentioned before, two external rigid coaxial cables are used to set a delay of one gate period between the two outputs of the power splitter. More details about the delay cables are discussed in the next session. To subtract the two signals we use a differential amplifier. The differential amplifier has a 10 GHz bandwidth, common mode rejection ratio better than 20 dB, and gain greater than 25 dB. The differential inputs are internally terminated to 50 Ω and isolated from each other. This guarantees that signals are not recirculated back into the delay lines toward the power splitter. The amplifier has a differential output, but in the figure 1 only the non-inverting line is shown, connected to a fast comparator for avalanche discrimination. The non-inverting output, not shown in figure 1, is used as signal monitor. The comparator has differential outputs and a latch input that can be conveniently used to enable the comparator only in specific intervals, enabling photon counting conditional on an external signal.

2.2 Minimizing the detectable avalanche current.

The circuit described in section 2.1 operates over a 10 GHz bandwidth and allows flexibility in the choice of gate shape, period, and duration. However, the wide bandwidth requires components with flat amplitude and phase transmission over the entire frequency range, and is more susceptible to phase noise in the clock signal. This concept is clearly illustrated in figure 2. Figure 2(a) shows the signal at the SPAD anode before cancellation, over a 50 Ω load, when a 9 V square gate with 600 ps temporal width is applied to the cathode. During these tests the SPAD is kept at room temperature and biased 15 V below breakdown so that the waveforms shown do not contain avalanches, allowing the exclusive characterization of the cancellation performance. The gate period is 1.6 ns. The delay lines are both constructed with 3.58 mm-diameter rigid coaxial cable.

Figure 2(b) is a screenshot from a sampling oscilloscope connected to the monitor output of the differential amplifier and shows the cancellation performance. The vertical scale in figure 2(b), and in figure 3, are rescaled to compensate for the overall signal gain from the SPAD anode to the output of the difference amplifier. The reported voltages are therefore referred to a 50 Ω load directly connected to the anode. As shown in figure 2(b), the signal after cancellation has two qualitative components that limit the overall cancellation. One is a periodic baseline that we ascribe to differences in phase and amplitude between the two signals from the delay lines. The second component is characterized by voltage fluctuations that occur in locations that correspond with the rising and falling edges of the gate signal. We ascribe this latter component to the effects of phase noise. The steep rising and falling gate edges have a slope of 40 V/ns, so jitter of only 0.8 ps from one period to the next can produce the fluctuation in amplitude of 32 mV peak to peak observed in figure 2(b). One solution to mitigate this latter effect is to reduce the signal bandwidth. To this end we added a 9.5 m-long spool of 2.49 mm diameter flexible coaxial cable between the SPAD anode and the resistive splitter. The cable introduces 8.7 dB attenuation at 10 GHz, 5.4 dB at 5 GHz and only 3.8 dB at 1 GHz. The residual signal after adding the cable is shown in figure 3(a), where the voltage fluctuations, as well as the baseline, are observed to be significantly reduced.



(a) (b)

Figure 2. (a) Signal measured at the SPAD anode when a 600 ps, 9 V peak-to-peak square gate is applied to the cathode. (b) Result of transient cancellation operated with the circuit described in figure 1. The residual signal, captured with a sampling oscilloscope has deterministic components, that we ascribe to the difference in dispersion introduced by the two transmission lines, as well as strong random fluctuations that we ascribe to temporal jitter in the gating signal.

The cancellation shown in figure 3(a) has still an evident baseline periodic component due to the difference in attenuation and total dispersion between the two rigid delay cables. The imbalance between the two cables cannot be addressed by altering the power ratio of the splitter because the attenuation of the cables is frequency dependent. We therefore investigate using rigid cables with different external diameters, 3.58 mm and 2.18 mm, for the long (L) and short (S) delays. Each 50 Ω cable is characterized (by the manufacturer) by an attenuation per unit length, $\alpha(f)$, at a given frequency f. For a given difference in physical length δL, and therefor propagation delay, we use the following

expression for the length of the short delay line that results in a more accurate match of the frequency-dependent losses over our bandwidth of interest:

$$L = \frac{\delta L}{\frac{\alpha_S(f)}{\alpha_L(f)} - 1} \quad . \tag{1}$$

A δL = 34 cm corresponds to a relative delay of 1.6 ns. Using the nominal information provided by the manufacturer to match losses at f = 2.0 GHz we obtain L = 43 cm. Figure 4 shows difference in attenuation for this pair of delay cables, as well as the case of two cables with the same 3.58 mm diameter. As can be seen in figure 4, with our approach to match the losses between different types of cable the difference in attenuation remains below 0.02 dB over a 3 GHz range, while using the same diameter cable results in an attenuation mismatch of more than 0.2 dB at 3 GHz. Figure 3(b) shows the benefit of this approach to the self-differencing setup: the residual peak to peak amplitude is reduced to 9 mV and dominated (once again) by jitter.

To further reduce the minimum detectable avalanche signal it is also possible to change the waveform of the electrical gate to limit high frequency components. The simplest case is to use a sinusoidal gate. Figure 5 shows a comparison between sinusoidal gating and square-wave gating with a 600 ps gate. The overvoltage applied to the SPAD in the two configurations is plotted in figure 5(a). For each waveform the peak-to-peak gate amplitude and the bias voltage are set to drive the SPAD with the same peak overvoltage, with the sinusoidal waveform above breakdown for 600 ps, as with the square gate. In both cases the SPAD is held at a temperature of -20 ˚C. In both measurements the gain of the differential amplifier is 26 dB, and the comparator threshold, V_t in figure 1, for avalanche discrimination is set to the lowest value possible without triggering spurious counts. The difference in performance between the two configurations results in a minimum threshold of only 10 mV for the sinusoidal gate and a threshold of 30 mV for the square gate.

The detection efficiency η is measured using the calibrated pulsed-laser source described in [10] that has a wavelength of 1310 nm and pulse duration less than 60 ps. The detector is expected to have maximum detection efficiency for photons arriving within the applied electrical gate. To find the peak value of detection efficiency the laser pulse is shifted in time with respect the SPAD gate and the observed detection efficiency is plotted in figure 5(b). The efficiency curves for sinusoidal and square gates differ significantly. Due to the lower discrimination threshold, the sinusoidally-gated system is sensitive to photons for a longer amount of time, twice that of the square gate (the FWHM values are 240 ps for the sinusoidal gate and 120 ps for the square gate). The increased sensitivity is also reflected in the per-gate dark-count probability, which is 2×10^{-5} for the sinusoidal gate configuration and 1×10^{-5} for the square gate.

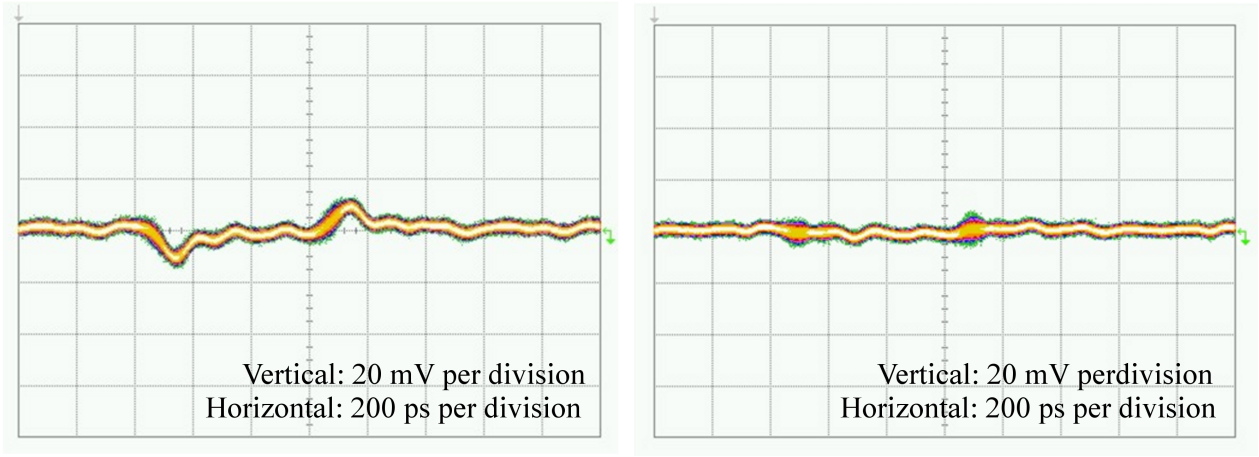


Figure 3. Cancellation signal when the SPAD anode is connected to the main board by a 9.5 m spool of coaxial cable. In (a) the delay lines have the same diameter, while in (b) diameters and lengths of delay lines are chosen to match the attenuation.

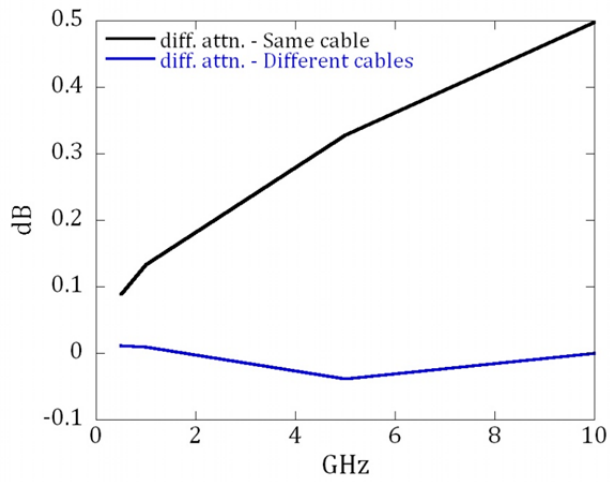


Figure 4. Difference in attenuation as a function of frequency between the two rigid coaxial cables used to introduce a relative delay of 1.6 ns. When the same cable is used, differences in attenuation are significantly larger than when using two different types of cable.

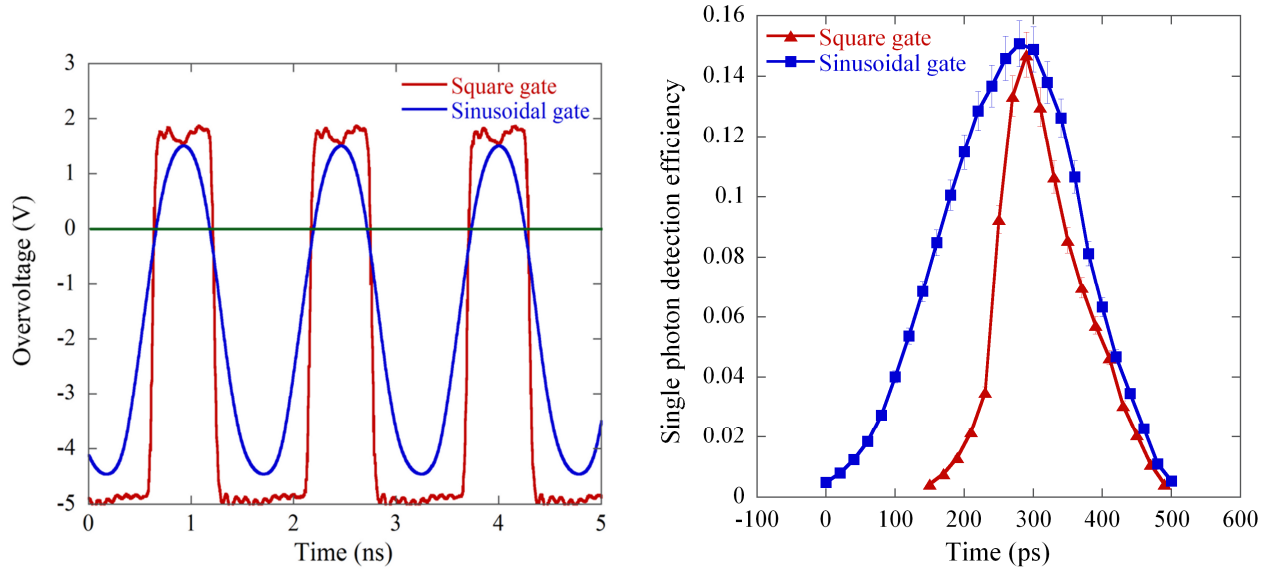


Figure 5. Difference in performance between square and sinusoidal gate. Figure 5(b) shows the detection efficiency as a function of an arbitrary delay between the gate pulse and optical pulse sent to the SPAD.

3. PERFORMANCE AT HIGH COUNT RATES

3.1 Detection efficiency and timing resolution

To characterize the system for high count rates, we use a gate period of 794 ps (1.26 GHz). To achieve the lowest threshold we run the system with a sinusoidal gate. The SPAD is maintained at a temperature of -20 ºC. The gate amplitude (measured on a 50 Ω load) is set to $V_g = 8$ V peak-to-peak and the bias voltage is adjusted for a maximum detection efficiency greater than $\eta = 0.11$. The dark count probability per gate is 1.5×10^{-5}. The detection

efficiency as a function of the delay between attenuated laser pulse and electrical gate, as measured in section 2.2 is shown in figure 6.

To check the fidelity and reliability for count rates exceeding 100 MHz a calibrated laser pulse is sent to the SPAD every fourth gate (at a frequency of 315 MHz). As the mean photon number μ is varied the expected count rate in the illuminated gate is given by:

$$C_{expected}(\mu) = f \cdot \left(1 - e^{-\mu \cdot \eta}\right), \tag{2}$$

where we assume Poisson statistics. To show that even at even at high count rates dark counts and afterpulsing account for a negligible fraction of the count-rate (less than 1.4 %), we monitor the count rate, C_B, observed in the gate that occurs directly before the gate illuminated by the laser. These data are also shown in figure 7.

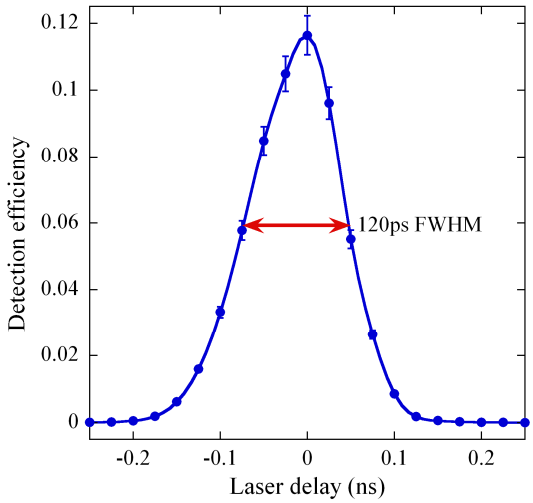


Figure 6. Detection efficiency as a function of arbitrary delay between the gate pulse and optical pulse sent to the SPAD when the SPAD is gated sinusoidally at 1.26 GHz.

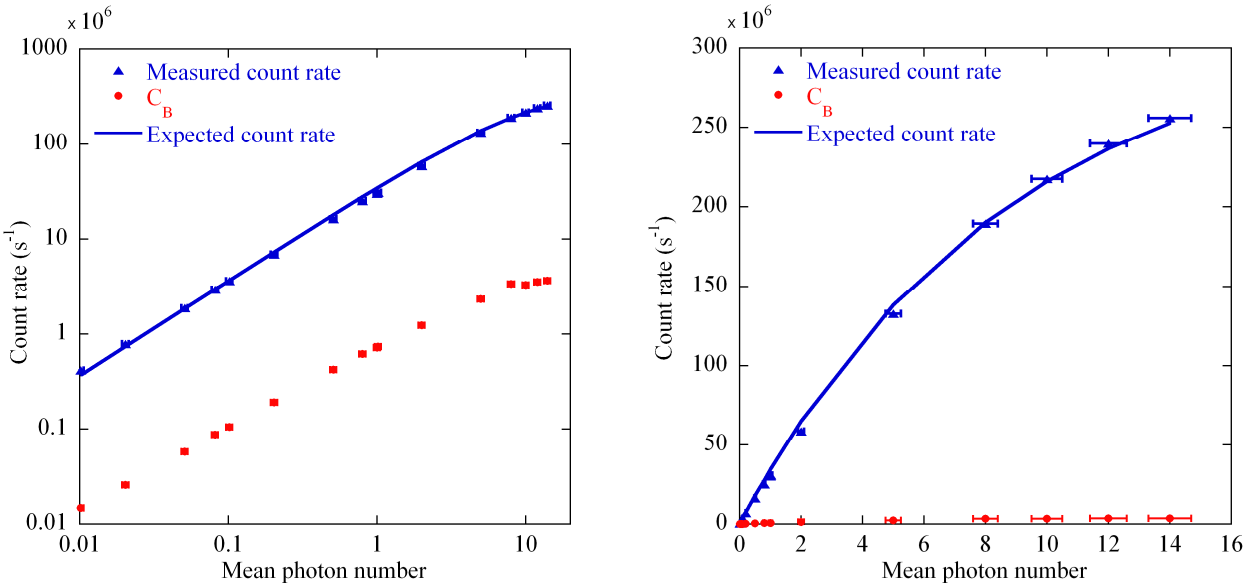


Figure 7. The two plots show in logarithmic (left) and linear (right) scales the measured and expected count rate as a function of mean photon number (blue) when the detector is illuminated in every fourth gate, and the count rate, C_B observed in the non-illuminated gate that precedes each illuminated gate.

3.2 Afterpulse probability

To monitor the excess count rate over more non-illuminated gates, and to study the benefits of applying hold-off times, as suggested in [8], we also illuminated the SPAD once every eighth gate. A convenient way to characterize the afterpulse probability is to plot count rate in the seven non-illuminated gates versus count rate in illuminated gates, as shown in figure 8. To simulate various hold-off times, we ignore counts in the first gate following the illuminated gate (i.e. simulating an 800 ps hold off), the first two gates following the illuminated gate (simulating a 1.6 ns hold off), and the first three gates following the illuminated gate (simulating 2.4 ns hold off). These data are also shown in figure 8, normalized to a single gate. To obtain the experimental data in figure 8, the mean photon number μ is varied from 0.1 to 40, where the count rate in the illuminated gate saturates at 157×10^6 s^{-1}. As the count rate in the illuminated approaches saturation, the count rate in the non-illuminated gates steeply increases. Below saturation there is a linear dependence between the count rates, and this slope gives an estimate of the afterpulse probability per gate. Figure 8(b) shows an expanded portion of figure 8(a), and the estimated afterpulse probability for each hold-off time. The introduction of a hold-off of only 2.4 ns reduces the afterpulse probability from 0.014 to 0.0084 (a 40% reduction). We note that this type of hold-off does not physically turn off the device, so this measurement slightly over-estimates the actual afterpulse probability, but nonetheless gives a clear picture of the benefits of applying even a brief hold off, either in hardware or software.

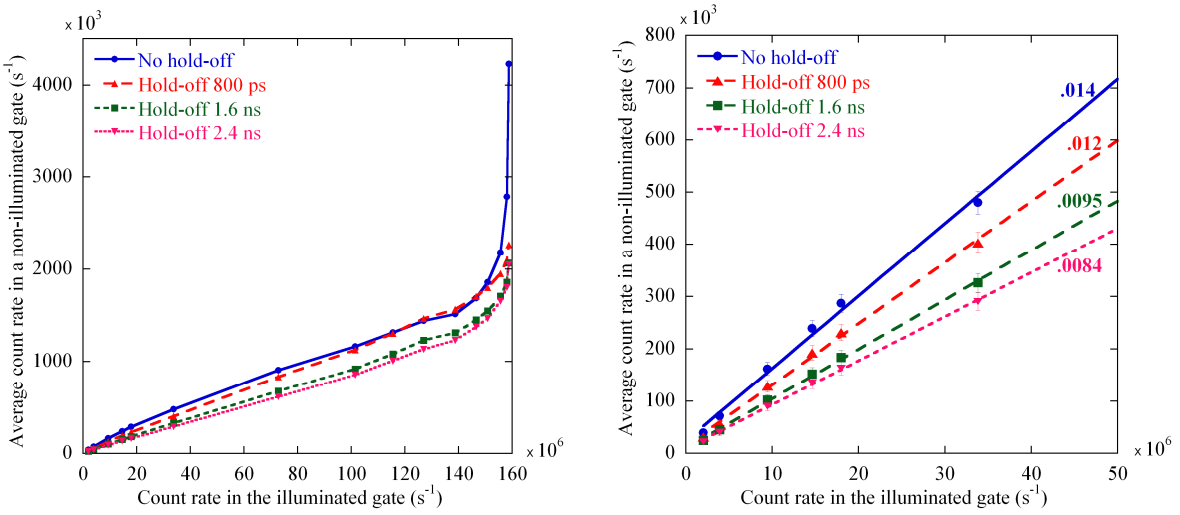


Figure 8. The average count rate observed in a non-illuminated gate versus the count-rate in the illuminated gate. Figure 8(a) shows measurements over the entire range of count rates while figure 8(b) shows a subset of the horizontal scale in (a). Measurements at different hold-off times are reported. The slope gives an estimate of the afterpulse probability per gate.

4. SUMMARY

We discuss a system for sub-nanosecond gating of InGaAs/InP SPADs. The wide bandwidth of our implementation requires careful design of microwave circuits. We illustrated how the design of attenuation-matched delay lines benefits the rejection of the gate signal. We also tested the system by applying gates of different shapes, square and sinusoidal, and we demonstrated the role of high frequency components in limiting the minimum threshold for avalanche-current detection that limits the detection efficiency in short gates. We tested the performance at high count rates, maintaining good linearity up to 250×10^6 s^{-1} with low contributions from afterpulsing.

This work was supported in part by the DARPA InPho program.

REFERENCES

[1] Ali-Khan I., Broadbent C. J., and Howell J. C., "Large-Alphabet Quantum Key Distribution Using Energy-Time Entangled Bipartite States," Phys. Rev. Lett. 98, 060503 (2007).

[2] Hadfield R. H., "Single-photon detectors for optical quantum information applications," Nat Photon 3, 696-705 (2009).

[3] Ben-Michael R., Itzler M. Nyman A., B., and Entwistle M., "Afterpulsing in InGaAs/InP single photon avalanche photodetectors," Digest of the LEOS Summer Topical Meetings 2006, pp. 15-16 (2006).

[4] Namekata N., Sasamori S., and Inoue S., "800 MHz single-photon detection at 1550-nm using an InGaAs/InP avalanche photodiode operated with a sine wave gating," Opt. Express 14, 10043 (2006).

[5] Yuan Z. L., Kardynal B. E., Sharpe A. W., and Shields A. J., "High speed single photon detection in the near infrared," Appl. Phys. Lett. 91, 041114 (2007).

[6] Namekata N., Adachi S., and Inoue S., "Ultra-Low-Noise Sinusoidally Gated Avalanche Photodiode for High-Speed Single-Photon Detection at Telecommunication Wavelengths," IEEE Photon. Technol. Lett. 22, 529-531 (2010).

[7] Yuan Z. L., Sharpe A. W., Dynes J. F., Dixon A. R., and Shields A. J., "Multi-gigahertz operation of photon counting InGaAs avalanche photodiodes," Appl. Phys. Lett. 96, 071101 (2010).

[8] Zhang J., Eraerds P., Walenta N., Barreiro C., Thew R., and Zbinden H., "2.23 GHz gating InGaAs/InP single-photon avalanche diode for quantum key distribution," Proc. SPIE 7681, 76810Z (2010).

[9] Jian Y., Wu E., Wu G., and Zeng H., "Optically Self-Balanced InGaAs/InP Avalanche Photodiode for Infrared Single-Photon Detection," IEEE Photon. Technol. Lett. 22, 173-175 (2010).

[10] Restelli A., Bienfang J. C., and Migdall A. L., "A technique to measure afterpulse probabilities in InGaAs SPADs at nanosecond time scales with sub-picoCoulomb avalanche charge," Proc. SPIE 8033, 80330I (2011).