# Advanced Devices and Circuits for Near Infrared Single Photon Detection

---

**Dissertation**

Presented to

the faculty of the School of Engineering and Applied Science

University of Virginia

---

In Partial Fulfillment

of the requirements for the Degree

**Doctor of Philosophy**

in Electrical Engineering

by

**Zhiwen Lu, B. S., M. E.**

May 2013

APPROVAL SHEET

The dissertation is submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy in Electrical Engineering

_____
Zhiwen Lu

This dissertation has been read and approved by the examining Committee:

_____
Dr. Joe C. Campbell (Advisor)

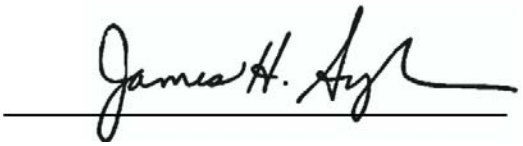
_____
Dr. Archie L. Holmes Jr. (Committee Chair)

_____
Dr. Nicolas Barker

_____
Dr. Stephen G. Wilson

_____
Dr. Xudong Jiang

Accepted for the School of Engineering and Applied Science:



Dr. James H. Aylor, Dean
School of Engineering and Applied Science

(May 2013)

# Abstract

Single photon counting in near infrared (NIR) provides the sensitivity of detecting extremely weak optical signals and enables a wide spectrum of applications. These applications fall into two categories, free space imaging and fiber optical communication. Light detection and ranging and quantum key distribution are the primary representatives of the two types of applications respectively. Especially quantum key distribution has been the major driven force for research in high data rate communications. Single photon avalanche diodes (SPADs) present high detection efficiency in NIR, compact size, low power consumption and low cost compared with their alternative candidates, namely photon multiplier tubes (PMT) and superconducting single photon detectors. Therefore, SPADs, especially InGaAs/InP SPADs become the most practical solution to the single photon detection in NIR.

Afterpulsing effect is the major limiting factor for high data rate single photon counting applications. It arises when the avalanche events happen so frequent that the trapped carriers cause severe accumulated excess dark counts. Recent research has focused on emerging gating techniques that provide narrow biasing pulses for SPADs, such as self-differencing and sine-wave gating. Narrow pulses are beneficial in the fact that the total charge flow during the avalanche is reduced in this manner. However it encounters the challenge to detect small avalanche pulses from various background noises. Differential mode signaling then becomes an effective solution for detecting weak signals. For instance, balanced photodetectors based on differential signaling can improve

system sensitivity by 15-20 dB compared with single detector. In my research study, I have incorporated this idea into both sine-wave gating and pulsed gating schemes and have achieved record result compared with previous results in our group.

In the first part of my research work on single photon counting, I established a sine-wave gating system that has agile laser repetition rate. Characterization based on this system is presented with comparison with previous results. I further developed this system into balanced detection mode, where the gating frequency is also agile thanks to the common mode cancellation mechanism. In the end, I also realized the balanced detection in pulsed gating and have achieved the smallest charge flow with the narrowest available pulse (1.4ns). From the road map of the performance improvement and the reduction in charge flow, one can conclude that reducing charge flow is directly effective in suppressing afterpulsing, especially when the data rate reaches 1 MHz and beyond. To be more specific, the data rate has been increased by 200 times without degradation in performance with 200 times reduction in the total avalanche charge flow.

# Acknowledgements

My deepest gratitude goes to my research advisor, Dr. Joe C. Campbell for his invaluable guidance and support through my graduate study. Without him, I could never have achieved any part of the PhD research program and dissertation. His high dedication and enthusiasm to fundamental and practical research has had and will continue to have great impact on my professional career. His technical and editorial advice has been precious gift to me throughout the entire PhD study. I am very thankful to him for his generously sharing experience and insights in both academic and personal life. I also learned a lot from his group and greatly appreciate the opportunity at his group and the University of Virginia.

I am very grateful to Dr. Mark A. Itzler and Dr. Xudong Jiang from Princeton Lightwave Inc. for their insightful discussion and friendly support on single photon counting avalanche diodes. I would like to thank Dr. Archie Holmes for serving as the chair of my dissertation committee and Dr. Nicolas Barker and Dr. Stephen G. Wilson for being my proposal and dissertation defense committee. They all provide helpful advice during my proposal and dissertation presentations. I also appreciate Dr. John P. R. David and Dr. C.H. Tan at University of Sheffield (UK) for their generous guidance during the initiation of projects and friendly encouragement.

I would like to thank my senior students Dr. Mingguo Liu and Dr. Chong Hu for their excellent work on single photon counting, which serves as a strong starting point for me. Dr. Hu also patiently familiarized me with the standard procedure of single photon

counting and has contributed a lot into our single photon counting set up. I greatly appreciate Dr. Andreas Beling for his helpful discussion and guidance on my initial thoughts of proposal. I want to thank Dr. Xiaoguang Zheng for his supervision and comments on experiments. In addition, Dr. Qiugui Zhou has been a great fellow researcher for me; discussion with him was always constructive and helpful. I deeply appreciate his friendship. I am also very grateful to the group members at UVA, it is them together who have made this research group productive and competitive. My great thanks go to Wenlu Sun, Kejia Li, Lijun Li, Yaojia Chen, Yang Fu, Allen Cross and Angad Sachdeva. Furthermore, I also want to thank the previous group members at UVA, Drs. Han-Din Liu, Huapu Pan, Dion McIntosh, and Zhi Li. I have enjoyed the friendship, collaboration and discussion with them during the past years.

I would like to thank Brenda E. Crider, Yadira Weaver, Dan R. Fetko, Eileen F. Thacker and Gloria R. Walker for their administrative help. Brenda has also brought many cheerful moments and warmness to our group. In addition, I would like to thank Joe Beatrice, Harry H. Wade, Alex M. Lobo, David Durocher and Gary Li for all the technical support.

I also greatly appreciate the friendship I had during the years at UVA. Many happy moments belonged to our friendship: Fang Wang, Kathryn Owen, Deborah Clowney, Darien Wei, Weijie Wang, Viola Gongyu Zhang, Shi Liu, Matthew Baysinger, Lichao Zhang, Manli Ding and many others. Moreover, Matt has kindly proofread all my dissertation chapters and provided many good editorial suggestions.

Last but not the least, I could not thank enough my family. My father has been a wise coach to my technical study; my mom has been the endless encouragement and positive

energy source to me when I was desperate during research study. My sister, Jinwen Lu, also acted as my strong support and guidance in life. In addition, as the youngest child among all my cousins, I great appreciate my senior cousins for their warm encouragement and guidance. I could hardly imagine any achievement without my family.

# Table of Contents

# List of Figures

# List of Tables

# Chapter 1 Introduction

Single photon detection is the technique that provides the ultimate sensitivity for detecting extremely weak optical signals in a broad-spectrum regime. It is widely employed in many applications (Section 1.1). [2, 3] Section 1.2 introduces the candidates for single photon detection: photomultiplier tubes (PMT), single photon avalanche diodes (SPADs), superconducting single photon detectors (SSPDs), quantum-dot field-effect transistor-based detectors (QDOFETs), superconducting tunnel junction (STJ) based detectors, and superconducting transition edge sensors (STES). Section 1.3, the thesis organization, describes the outline of this dissertation.

## 1.1 Applications

Single photon detectors are widely used in infrared detection modules in order to achieve single-photon level detection [2, 3]. These applications at near infrared (NIR) wavelength include quantum key distribution [4, 5], semiconductor circuit diagnosis [6], eye-safe laser ranging (LIDAR) [7], and biological imaging [8]. The ideal single photon detector that could satisfy all the applications would detect every incident photon (100% detection efficiency), produce no false counts (zero dark count rate), work in the full optical spectral range of interest, and have photon energy limited timing resolution (zero timing jitter and zero dead time), and infinite dynamic range. Fortunately, some levels of imperfection are adequate for these applications. In fact, different applications have quite different requirements on single photon detectors. Achieving the specific requirements

that are essential for different applications has been the focus of numerous research efforts.

One of the major driving forces of single photon detectors in the NIR has been quantum information. Figure 1.1 summarizes academic paper publications during the past several decades.

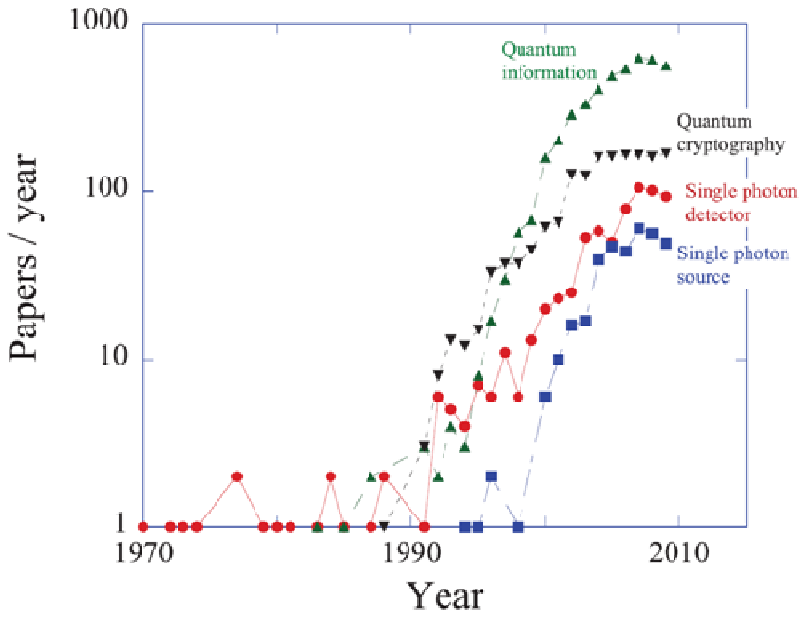



Figure 1.1: Papers published each year. ISI web of knowledge search terms are shown. [3]

Quantum information includes quantum computing and quantum communication. Quantum communication, based on quantum key distribution (QKD), has already been deployed worldwide. QKD achieves secure communication by generating and sharing a secret key with privacy guaranteed by fundamental quantum mechanics [4, 9-11]. In practice the key is transmitted at the same rate as the data transmission rate. Therefore in order to realize high-speed connections, QKD has to operate at a high clock rate. The key information is typically carried on single photons that are generated and transmitted

between two parties via optical fibers. The signal is eventually detected by the single photon detectors at the NIR wavelengths. The system performance strongly relies on the photodetector and other key optical links in the system. In addition, for long distance transmission, the low noise characteristic of the photodetector is also crucial. Currently InGaAsP SPADs are used in QKD systems, however, their performance is not optimized and they are currently the bottleneck for the development of high bit rate quantum communication. This is due primarily to the dead time (recovery time) of the single photon receiver. Dead time is the time that the receiver needs to reset itself to detect subsequent photons following detection of an initial photon. Previously that has been limited to the kHz range due to high afterpulse rates. Afterpulsing is one type of dark count that arises from the frequent onset of avalanche events. Chapter 2 will discuss this concept in detail. Owing to the afterpulsing effect in the InGaAsP material system, the hold off time can be as long as several microseconds, which determines the dead time of a receiver. On the other hand, the gating window (ON time) needs to be as short as possible to avoid excess charge flow during the avalanche. Various quenching and gating techniques have been developed to increase the clock rate and data rate into the GHz range. Chapter 5 will introduce these techniques.

LIght Detection And Ranging (LIDAR), especially eye-safe laser ranging, is another application for single photon detection. LIDAR is a technique that captures 3D images of objects by measuring the time-of-flight of laser pulses. It is realized using SPAD arrays with highly specialized readout integrated circuits (ROICs). The spatial resolution depends on the time resolution (jitter) of the system and the image quality depends on the signal-to-noise (SNR) ratio. The SNR relies on the repetition rate of the

laser pulses and on the dark count rate of the detector. Therefore, low jitter, low noise, and high count rate single photon receivers are desired. However, different from QKD applications, the gating window (ON time) has to be long enough to capture the photons that travel the distance range spanned by the system.

Semiconductor circuit diagnosis detects light emission from hot carriers in the transistors of VLSI circuits and collects signals from such circuits. It also requires detectors with low jitter, low noise, high count rate, high SNR, and even longer gating windows. These characteristics can usually be realized with cryogenic cooling.

Though different applications have different requirements for single photon detectors, some of the common characteristics are low noise, high count rate, low jitter, and low power consumption. The next section will introduce different single photon detector candidates and define their characteristics.

## 1.2 Single Photon Detectors

### 1.2.1 Photomultiplier tubes (PMT)

The PMT is based on the photoelectric effect discovered by Hertz in 1887. It is also the first device that demonstrated single photon detection. It has been commercially available since 1960. Figure 1.2 illustrates the working mechanism of a typical PMT.

In the PMT shown in Figure 1.2, an incident photon hits the photocathode and releases an electron. This electron is then accelerated under the high electric field toward the first dynode and knocks out more electrons. The process repeats itself with more subsequent dynodes till the signal gets amplified to a macroscopic level that is detectable by conventional electronics. The conditions for this process are vacuum environment,

high voltage operation, and designated material for photocathode and dynodes. The efficiency of PMTs is typically in the range 10 % - 40 %, limited by the first and subsequent electron emission rates. Therefore, specially designed photocathode and dynode shapes and material is important, especially for some specific working wavelengths such as NIR. In addition to low efficiencies, PMTs are also limited by their lifetime, reliability and scalability due to the vacuum and high voltage requirements.

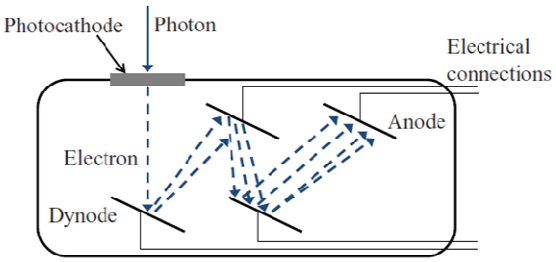



Figure 1.2: A photomultiplier tube, the first detector able to sense a single optical photon, is shown schematically with a transmissive photocathode and just 3 dynodes. The photocathode may be designed to have the photoelectrons emitted from its front or back surface and typically 10 dynodes are used [3].

### 1.2.2 Single Photon Avalanche Diodes (SPADs)

SPADs utilize impact ionization as the amplification mechanism for detecting weak signals. They are usually operated in Geiger mode (GmAPD), which will be discussed in detail in Chapter 3. In Geiger mode, the diode is biased close to or beyond breakdown so that the electric field inside the diode is strong enough to trigger serial impact ionization events. Once the impact ionization events become self-sustaining, sufficient minority carriers are collected by the electrodes to generate macroscopic current signals for detection. The efficiency of some SPADs can be as high as 85%, however these SPADs exhibit higher timing jitter and dark count rates than the best

PMTs. SPAD performance is also limited by an affect called afterpulsing, which is most severe at high detection rates. As a result, afterpulsing greatly limits the repetition rate of a SPAD as will be discussed in Section 2.2.2. A related detector, the intensified photodiode is a hybrid detector that combines a photocathode and an impact multiplication stage for amplification. The combination can overcome some wavelengths that are difficult to achieve high quantum efficiency with PMTs for single photon counting.

### 1.2.3 Superconducting Single Photon Detectors (SSPDs)

Superconducting SPDs are based on narrow superconducting wires biased slightly below the critical current density level. When an incident photon is absorbed it breaks a Cooper pair resulting in a hot electron. This hot electron causes the small region around it

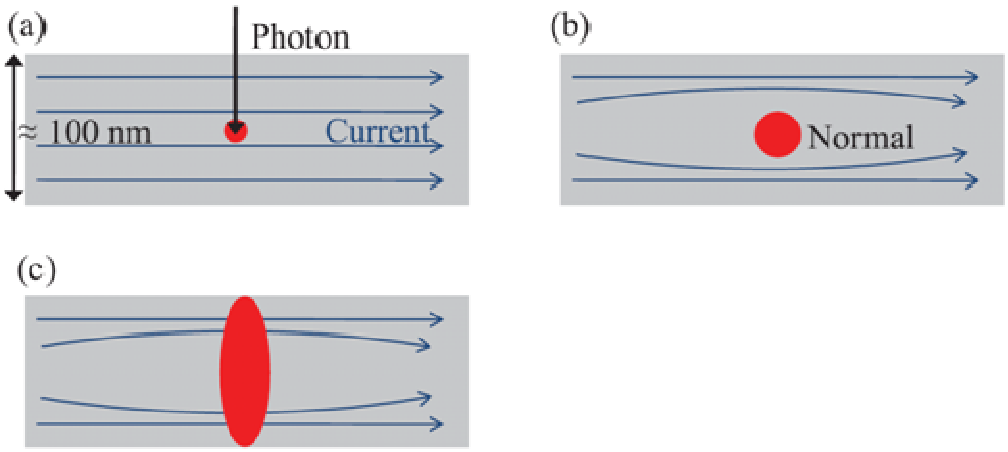



Figure 1.3: A section of a superconducting nanowire single photon detector is shown with a bias current just below the critical current density that would drive the wire normal. (a) An incoming photon creates a small normal region within the nanowire. (b) The superconducting current is expelled from the normal region, increasing the current density in the adjacent areas of the nanowire. (c) That increase in current density is

to revert to a normal resistance state, which repels local current flow. As a result the current density around this hot spot increases and a complete strip of resistive region is formed across the nanowire. This resistive strip produces a voltage spike in the external circuit indicating absorption of a single photon. (Figure 1.3)

Superconducting SPDs exhibit acceptable detect efficiencies and a broad photo-sensitive spectrum. More importantly, they do not suffer from afterpulsing. However, they may enter a stage where the current flow constantly stays above the critical point due to self-heating, and active resetting and lowering the current flow is necessary in this case. Moreover, operation of SPDs requires cryogenic cooling to as low as a few Kelvin, which largely restricts their deployment in practical systems.

### 1.2.4 Superconducting Tunnel Junction (STJ) detector

STJ-based detectors are one the first superconducting detectors that have photon number resolution (PNR) capability. The superconducting tunnel junction structure consists of three layers, two superconducting layers separated by a very thin insulator layer (~1 nm). The absorbed photon generates a large number (~ 1000) of broken Cooper pairs in the first superconducting layer. Since the separating insulating layer is very thin, a small cross bias can create a tunneling current between the two superconducting layers. This current flow is also proportional to the number of photon generated Cooper pairs (quasiparticles). The detector is operated at significantly lower temperature than the superconducting critical temperature; therefore there are fewer other quasiparticles other than photon generated quasiparticles. Thus the single photon detection is accurate with very low noise (<0.1% of photon counts), which is primarily electronic noise.

### 1.2.5 Superconducting Transition Edge Sensor (STES)

STES single photon detectors operate as a bolometer: they detect the temperature change due to the absorption of small amounts of electromagnetic radiation – single photons. The extreme temperature sensitivity is achieved by utilizing material with very small heat capacity, e.g., a thin superconducting layer. The superconducting layer is maintained in the transition temperature region where a small temperature change will trigger a transition between resistive and superconducting states. This increased resistance due to the rising temperature yields current reduction as a negative feedback on the thermal heating of the device and, thus, maintains the device in the transition temperature region. The highest detection efficiency among superconducting detectors is achieved by STES, 95% at 1556 nm. The disadvantages of STES include slow response (~ 100ns), low maximum count rate ~ 100 kHz, and the requirement for much lower operating temperature than the SSPD (~ 100 mK).

### 1.2.6 Quantum-Dot Field-Effect Transistor-Based Detector

Quantum-Dot Field-Effect Transistor-based detectors (QDFETs) are also referred to as Quantum-Dot Optically Gated Field-Effect Transistor-based detectors (QDOGFETs). They use quantum dots as optically addressable floating gates in a GaAs/AlGaAs doped field-effect transistor [12]. Figure 1.4 displays the device diagram. When a photon is absorbed in the AlGaAs active region, an electron-hole pair is produced and separated by the internal electric field. The electrons travel toward the 2-dimensional electron gas (2DEG) and the holes travel toward the quantum dots (QDs) where they get trapped. The accumulated holes in the QDs screen the electric field inside

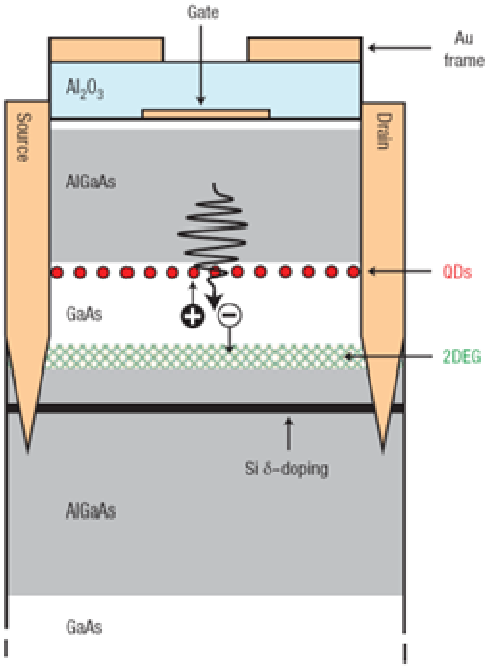



Figure 1.4: Diagram of a QDOGFET. Semiconductor layers (from bottom to top): GaAs substrate and 200 nm GaAs buffer layer, 2.5 mm $Al_{0.2}Ga_{0.8}As$, Si d-doping (~1 × $10^{12}cm^{-2}$), 70 nm $Al_{0.2}Ga_{0.8}As$, 100 nm GaAs, InGaAs QDs (400–500 $\mu m^{-2}$), 200 nm $Al_{0.2}Ga_{0.8}As$,10 nm n-doped (~6 × $10^{17}cm^{-3}$) GaAs cap. The device was fabricated by etching a channel mesa between the Ni/Au/Ge source and drain ohmic contacts and by depositing a semitransparent Pt Schottky barrier gate midchannel. The gated portion of the channel mesa defines the active area (2.0 µm × 2.4 µm). Photo-absorption is limited to the interior of the active area by an opaque Au mask with a 0.7 µm × 0.7 µm transmission window. A ~100-nm-thick transparent layer of $Al_2O_3$ separates the Au frame from the rest of the structure. Layers are not drawn to scale.[12]

the gate region and cause the gate bias to increase in the form of a channel current flow increase in the 2DEG. The channel current flow is proportional to the number of trapped photon-generated holes in the QDs. Thus the magnitude in the increment of channel current indicates the incident photon number. Although the individual current flow change by different numbers of incident photons is not well resolved. In other words, the photon number is only statistically related to the current indicator. The relation between the photon number and the current flow can be extracted by averaging over many laser pulses [12, 13].

## 1.3 Thesis Organization

The motivation for the work reported in this dissertation was to improve the performance of InGaAs/InP single photon avalanche diodes, such as reducing afterpulsing, increasing the data rate, lowering the dark count rate, and increasing photon detection efficiency. Improvements have been achieved on two levels, the device level and the electronics level. The device characteristics are the primary factor with regard to system performance. For instance, afterpulsing is caused by defects in the multiplication layer, which is usually InP. Reducing the trap density in the multiplication layer can significantly suppress the afterpulsing effect. The goal for circuit changes is to reduce the charge flow during avalanche events. This has proven effective for suppressing afterpulsing. This dissertation will focus on improving the single photon detector electronics.

Chapter 2 introduces the figures of merit for characterizing a SPAD; it also briefly describes the experiment set up for the characterization. The quenching circuit is a key part of the counting electronics. The conventional quenching circuits and operation mode will be introduced in Chapter 3. Chapter 4 presents experimental findings on a novel material system for single photon counting – Ge on Si SPADs; this is the first report of single photon counting for a Ge on Si SPAD. In order to upgrade the counting electronics, novel techniques are desired for both biasing and gating the diode. Chapter 5 discusses the emerging techniques in this area. Sine-wave gating is among the emerging gating techniques and will be discussed in Chapter 6 with experimental data and numerical analysis. In this dissertation, I will also report common mode cancellation techniques applied in single photon counting. Chapter 7 discusses the technique utilized in gated

mode. The subsequent chapter updates the result for sine wave gating. The final chapter, Chapter 9, discusses future work and summarizes the dissertation.

# Chapter 2 Characterization

In this chapter, I will introduce the basic experimental setup we used for characterizing single photon avalanche diodes (SPADs) and define the figures of merit for SPADs. The entire experimental setup consists of three components: a mechanical system, an optical system, and an electrical system. The mechanical system is comprised primarily of the cryogenic probe station that provides a low temperature environment for measurements. The optical system includes laser diodes and their drivers, monitoring apparatus (optical oscilloscope), optical fibers, and optical and electrical connections. The electrical system primarily deals with the biasing of the photodiode and, more importantly, detecting the photon count signals with either an oscilloscope or counters (a photon counter and a multichannel analyzer in this paper). In Section 2.1, these subsystems will be presented in detail. The following Section 2.2 will discuss the definitions of the figures of merit used for characterizing SPADs and the methods for calculating them.

## 2.1 Experimental setup

### 2.1.1 Mechanical system

The mechanical system is comprised of two parts, the cryogenic probe station and the vacuum system. The cryogenic probe station is model TTP4 from Lakeshore Inc. It has four arms and up to four probes can be installed in the four arms (Figure 2.1). It also

has a cryogenic Dewar and transfer line, which transports liquid cryogenic fluid (liquid nitrogen in this case) into the bottom of the chamber to cool the stage inside the chamber.

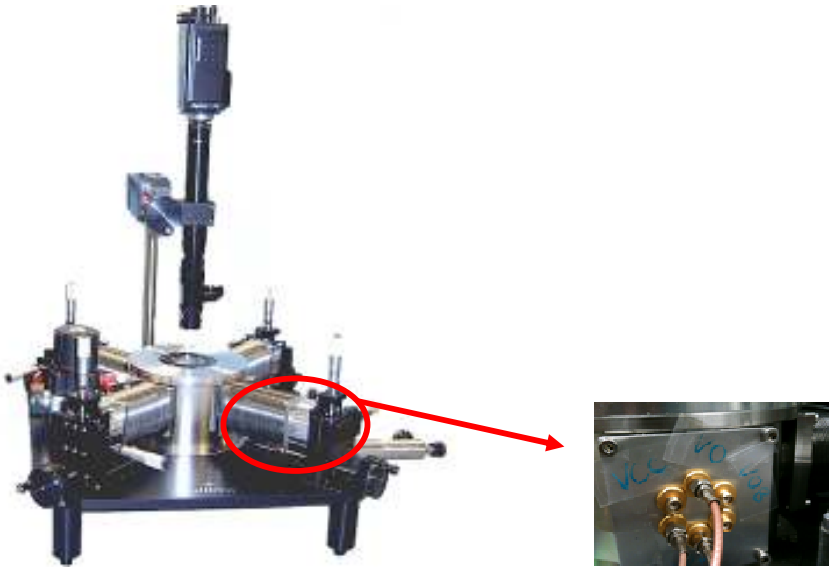


Figure 2.1: Four-arm cryogenic probe station with monitoring camera on the top from Lakeshore Inc. Mechanical pump and cryogenic system are not shown.

In order to characterize opto-electric properties, one of the four arms contains an optical fiber instead of an electric cable/probe. Two of the other arms have point probes that connect to the photodiodes; the fourth arm has been removed and replaced with a metal panel (inset of Figure 2.1) with SMA adapters. This panel is used when more than two electrical connections are needed; six connectors have been installed and they are connected to cables/adapters inside the chamber.

The whole chamber system is sealed and can be pumped to pressure as low as $5x10^{-7}$ torr with a mechanical pump and a turbomolecular pump. Underneath the sealed chamber is another chamber where liquid nitrogen can be introduced to lower the temperature of the probe stage. Two temperature sensors have been installed in the

system; one, sensor A, is inside the liquid nitrogen chamber, below the probe stage and the other, sensor B, is attached to one of the probe arms within the main chamber. The temperature control meter reads both measurements from sensor A and B, and adjusts the internal heater according to the relation between the reading from sensor A and the set point to reach equilibrium between cooling and heating. The reading of sensor B is usually within 10 K higher than that of sensor A when a good equilibrium status is reached. Since the performance of avalanche photodiodes is highly temperature sensitive, especially when the operation is near or beyond breakdown, a well-stabilized temperature environment is very important. After sensors A and B reach their equilibrium stage, it is necessary to wait for a while until the device inside the chamber equilibrates to the same temperature as sensor A. This is even more critical for circuit board mounted devices than for on-wafer devices, whose substrate sits on top of the probe stage. One way to check the stability of device temperature is to repeat the current-voltage measurement several times, ensuring that all the curves show breakdown voltage at the same point. These curves will be used to determine the DC bias point for most of the gated mode operations.

One more note about the low temperature station is the cryogenic efficiency maintenance. Many experiments involve low temperature measurements; some require temperature as low as 77 K, the lowest temperature for a system that uses liquid nitrogen. The vacuum status of the transfer line affects the cooling speed of temperature lowering and the cryogenic consumption. The transfer line is pre-pumped for several days before it is delivered. Later maintenance requires at least 12 hours re-pump if it is leak tight or

being re-evacuated. This type of maintenance is carried out at least once every three months.

### 2.1.2 Optical system

In order to characterize a single photon detector, one needs a single photon source. Figure 2.2 illustrates the optical system that generates single photons at an intensity level of average 0.1 photon/pulse. The two lasers, designated continuous (CW) laser and pulsed laser in Fig. 2.2, are driven by DC and AC biases, respectively. They are coupled into a 2x2 optical switch controlled by a 5 V power supply. By switching on the power supply, the input 1 is switched from output 1 to output 2 while input 2 is switched from output 2 to output 1. When the power supply is turned off, the outputs return to their initial states. The CW laser is used to calibrate the loss from the power meter to the fiber collimator that illuminates the photodiode, since it usually carries a significant amount of light that is much stronger than single photon level and therefore, the signals used to do calibration is much larger than the background noise (which is not the case with pulsed laser source). After the fiber collimator is aligned with the photodiode, output 1 can be switched to the pulsed laser, which functions as the single photon source. In order to obtain single photons, output 1 is split into two portions by a ratio of 9:1. The 90% signal is measured with a power meter and the 10% is transmitted through several attenuators before coupling into the collimator. The spare branch, output 2, is utilized as an optical monitor for both lasers. The optical pulse width can be measured when the laser pulse is displayed on an oscilloscope.

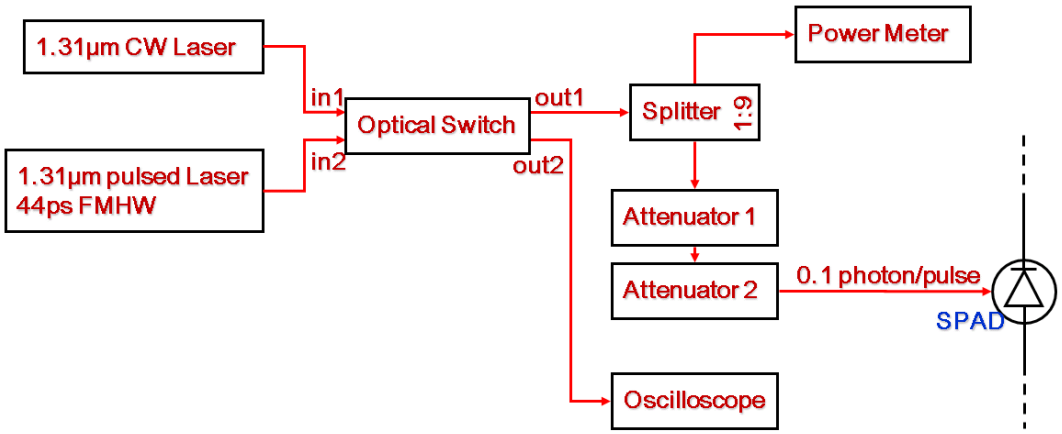



Figure 2.2: Optical setup and paths to achieve photon intensity at single photon level.

### 2.1.3 Electrical system

The electrical system is the most important and complicated part of a single photon counting system, not only because it determines how the device is powered and quenched, but it also differentiates the avalanche pulse signals from the background noise, which can be very close to each other. The biasing and quenching circuit is a complex system itself, which I will discuss in the next chapter. In this section, I will discuss signal processing, the function of the single photon counter.

Two photon counters have been used in this work, a PicoHarp 300 and a Stanford Research 400 (SR 400). Both are used for synchronized photon counting events, in which case, a trigger signal is sent to the counter each time a count is registered.

The PicoHarp 300 is also referred to as a multichannel analyzer. The version we have has two channels, one is the default synchronization channel, and the other is the

counting channel. The approach used to record a single photon event is the start-stop method as shown in Figure 2.3. Each time the counter receives a synch signal it refreshes its internal clock and waits for a potential count. When the count is registered, the clock stops and records that time interval according to the trigger time. This makes one set of data for one count. These data will be processed later with Matlab, yielding the real time tags for each counting event.

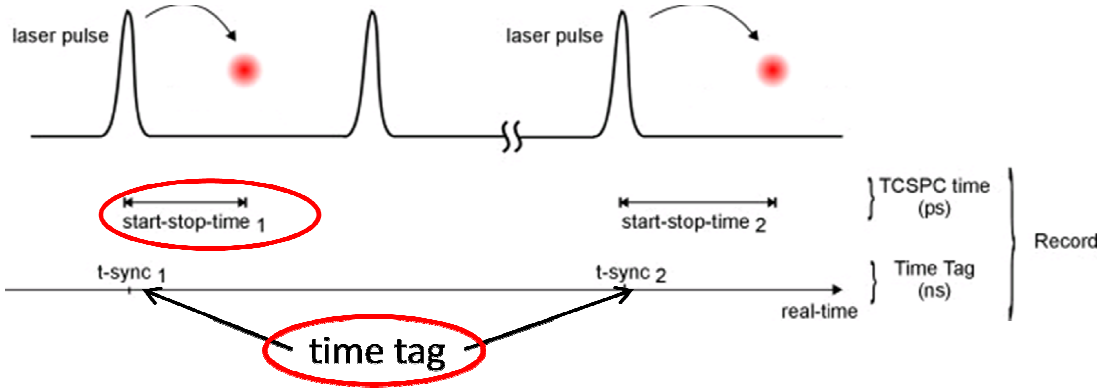



Figure 2.3: Time correlated single photon counting operation of the PicoHarp 300. Red dots are the signals that are recorded by the counting; t-sync is the time tag used for calculating the absolute time of the event later.

Figure 2.3 illustrates that this counter is capable of recording real time events in an efficient way due to the fact that part of the signal processing is done off-line. Combining this feature and the precise built-in electric clock, the PicoHarp 300 has a timing resolution of 4 ps. The histogram mode is very useful for measuring the timing resolution of the photodiode. It is noted that the timing resolution is different from the bandwidth/speed of this instrument. The PicoHarp 300 has a maximum synch rate of 80 MHz. Furthermore its time to digital converter, the time required to reset for the next count, is ~ 80ns,. Therefore, in this work, the highest gating speed is 80 MHz and a saturation effect occurs when the count rate reaches ~ $1x10^5$. Figure 2.4 shows the proof

of the dead time by counting a series of random events; the minimum time interval between two adjacent counts is approximately 80 ns.

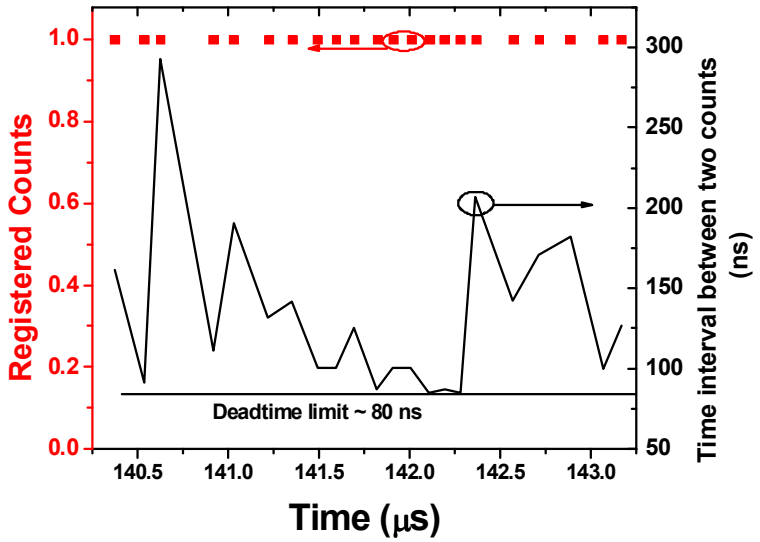



Figure 2.4: Experimental proof of dead time of photon counter PicoHarp 300

Compared to the PicoHarp 300, the SR400 has more flexible counting modes. It has a continuous (CW) and a gated mode. In continuous mode, it counts for each trigger signal for a specified number of triggers. Usually the number of triggers is set the same as the gating repetition rate; therefore it takes one second to finish one period of counting. The gated mode is more precise in determining the time that the count occurs, as both gate width and gate delay can be set independently. As shown in Figure 2.5 (a) and (b), the counting gate width can be set between 5 ns to 10 μs, the smallest gate width is consistent with the 200 MHz bandwidth of the counter. Similar to the PicoHarp, its maximum synchronization rate - 1 MHz is also smaller than its resolution corresponding bandwidth.

In most cases where the gating frequency is the same as the laser repetition rate, the CW mode is equivalent to the gated mode operation. While in sinusoidal gating the laser repetition rate is a submultiple of the gating rate. The gated mode yields two different types of counts: those including a laser pulse are photon count gates, and those without laser pulses are dark count gates. I will discuss this point more in Chapter 6. We note that both counters yield similar device performance as shown in Fig. 2.6.

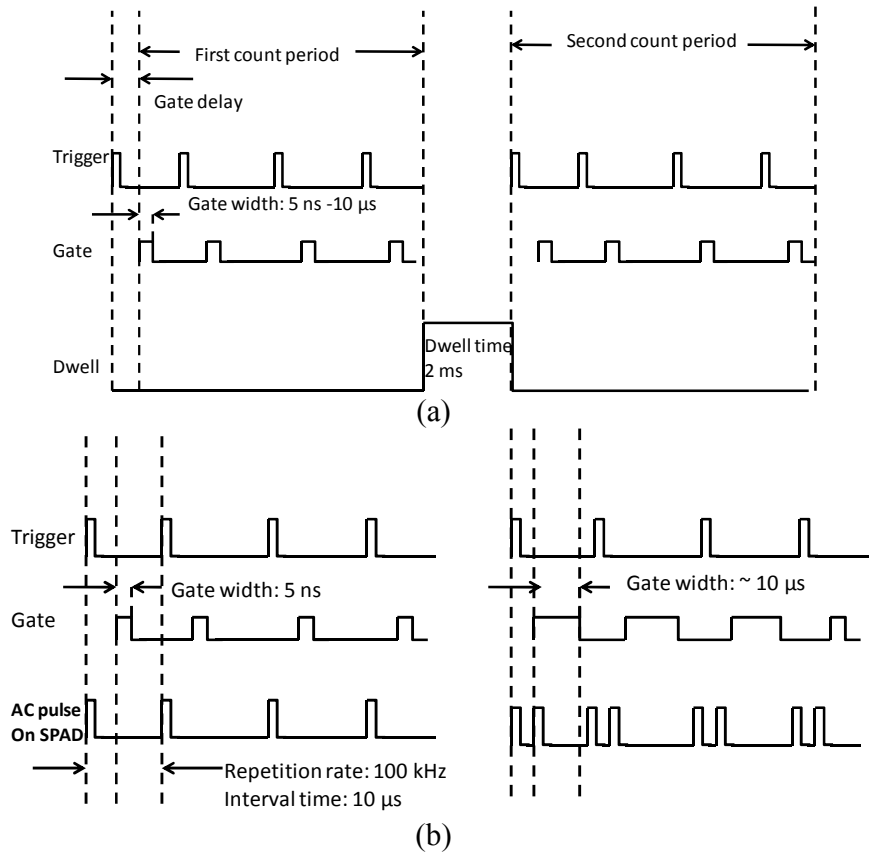



Figure 2.5: Timing diagrams for (a) external trigger counting mode for the SR 400 (b) Single (left) and double (right) pulse measurement scenario.

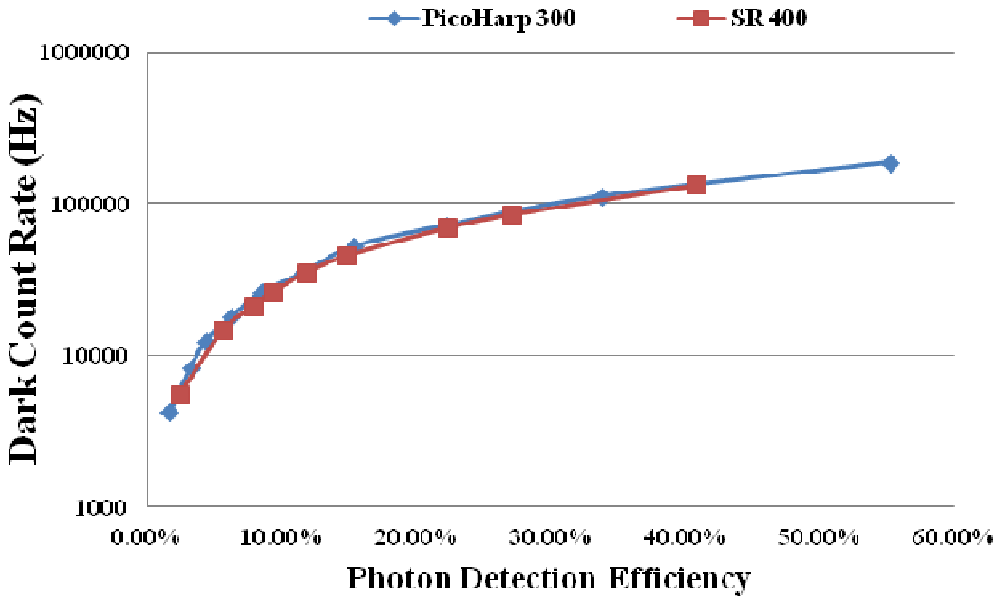



Figure 2.6: Counting results from two different counters under the same conditions.

## 2.2 Figures of merit

### 2.2.1 Dark count rate and Photon detection efficiency

Dark count rate (DCR) refers to counts that are generated by dark carriers instead of photon-generated carriers. The population of dark carriers is largely related to bulk dark current at the desired bias, which is close to or above breakdown. DCR can be reduced by decreasing the dark current, which can be accomplished by lowering the temperature with a thermal electric cooling (TEC) or a low temperature station.

The dark current ($I_d$) consists of unmultiplied dark current ($I_{um}$) and primary multiplied dark current ($I_m$). The total dark current is given by the expression

$$I_d = I_{um} + M \times I_m \quad . \tag{2.1}$$

The unmultiplied dark current is usually surface leakage that does not experience the high electric field in the bulk region. Surface leakage does not experience impact

ionization and, therefore, does not contribute to dark counts. The primary multiplied dark current on the other hand is closely related to DCR. Improved wafer crystalline quality plays an important role in suppressing the primary multiplied dark current since it originates primarily through defects and impurities in the bulk region. A good SPAD exhibits smaller multiplied dark current at unity gain than the unmultiplied dark current as shown in Figure 2.7. A linear fit between the dark current and the multiplication is shown in Figure 2.7, where the slope represents the multiplied dark current component and the intersect value is the unmultiplied dark current component.

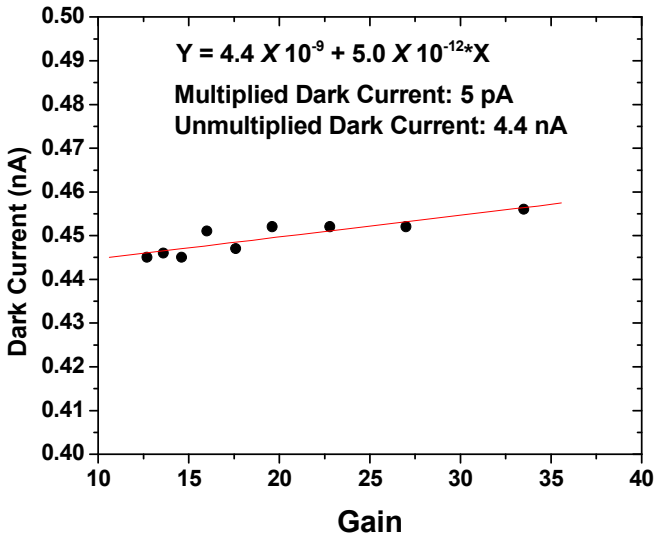



Figure 2.7: Linear fit of dark current vs. gain shows much lower multiplied dark current than unmultiplied dark current. The device is 25 µm InGaAs/InP APD from Princeton Lightwave.

Photon detection efficiency (PDE), which is the probability that an incident photon is registered as a photon count. PDE can be written as

$$PDE = \eta \times P_{av} \tag{2.2}$$

where $\eta$ is the external quantum efficiency and $P_{av}$ is the probability that an electron-hole pair can initiate a chain of impact ionizations that leads to avalanche breakdown. Equations (2.1) and (2.2) were used to calculate the dark count rate (DCR) and photon detection efficiency (PDE). $P_d$ is dark count probability, $P_t$ is total count probability and n is number of photons per pulse (0.1). The effective pulse width ($\tau_e$) was determined by measuring the full width at half maximum of the temporal distribution of the photon counts; it can be determined using the following expression

$$DCR \times \tau_e = -\ln(1 - P_d) \qquad (2.3)$$

$$\text{PDE} = \frac{1}{n}\ln(\frac{1-P_d}{1-P_t}) \qquad (2.4)$$

To measure PDE and DCR, I first measure the counts with and without light illumination. Then I calculate the rates with following equations

$$P_t = 1-(1-P_d)e^{-PDE\bullet\bar{n}} \qquad (2.5)$$

$$P_d = 1-e^{-DCR\bullet\tau_p} \qquad (2.6)$$

where it is assumed that both incident photons and dark carriers obey a Poisson distribution. With measured PDE and $\eta$ one can obtain the avalanche probability, $P_{av}$.

### 2.2.2 Afterpulsing

As stated above, the afterpulse probability is the probability that an emitted carrier that was trapped by a defect or an impurity during a previous avalanche event triggers a dark count. Afterpulsing becomes more severe when the frequency of avalanche events increases. Two time constants are important with respect to afterpulsing. One is the trapped carrier lifetime ($\tau$) and the other is the time interval between two adjacent

avalanche events, i.e., the hold-off time. When the hold-off time is close to or shorter than the trapped carrier lifetime, DCR increases. Therefore a measure of DCR as a function of frequency provides a good indication of the severity of afterpulsing. It should be noted, however, that different techniques are used to measure afterpulsing for different gating schemes. For example, the double pulse measurement [14, 15] has been widely accepted for gated quenching. Recently a generalized double-pulse measurement has been introduced [16, 17] for narrow-pulse-width gated quenching [18, 19]. In this modified double-pulse measurement, the incident laser pulse is temporally aligned with pulses in which photons are incident. These "lit" pulses are interleaved by one or several (up to 128) "dark" gates. The dark count rate is measured in the absence of incident light, while the PDE and afterpulsing are measured with coincidence "lit" gates and interleaved "dark" gates, respectively. By changing the gating rate, shown in Figure 2.8, "dark" gates with different temporal separation can be achieved.

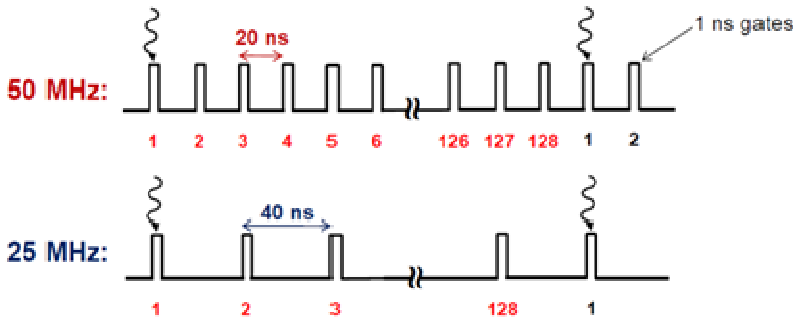



Figure 2.8: Generalized double pulse scheme with 50 MHz and 20 MHz gating rate.

### 2.2.3 Jitter

Jitter, also known as timing resolution, is the uncertainty in time when an avalanche event occurs with respect to the fixed arrival time of incident photon. This randomness also imposes a limitation on the operation speed. Compared with the hold-off time for suppressing afterpulsing (~100ns), the typical jitter for InP/InGaAs SPADs (50 - 100 ps) is negligible in most applications. However, the recent development in fast gating, such as self-differencing and sinusoidal gating, where the gating width for the device is sub-nanosecond, jitter on the level of hundred of picoseconds could be an obstacle for further development [20].

### 2.2.4 Sensitivity and noise equivalent power

Noise equivalent power (NEP) is an important parameter for linear-mode counting. NEP is defined as the signal that gives unity signal-to-noise (SNR) ratio with one hertz bandwidth. NEP can be calculated with PDE and DCR from equation (3) in Ref.[21], and repeated here as equation 2.7. Theoretically, NEP of $1.4 \bullet 10^{-17}$ W/√Hz is necessary for a PDE of 40% and DCR of 1 kHz at a wavelength of 1.5 μm.

$$NEP = \frac{h\nu}{PDE}\sqrt{2 \bullet DCR} \tag{2.7}$$

In Chapter 3, where linear mode operation is discussed, a similar parameter - NEPh will be used. NEPh is derived from NEP by transforming the power into the number of photons at the wavelength of interest. While NEP is a traditional concept for electronic systems, NEPh is more straightforward when dealing with photon number resolution applications.

## 2.3 Conclusion

In summary, I have introduced the experimental systems for characterizing single photon avalanche diodes and the figures of merit for evaluating the performance of the device. More details regarding the electronics for single photon counting will be presented in Chapter 3.

# Chapter 3 Operation Modes and Quenching Circuits

In order to achieve high sensitivity for single photon detection, single photon detectors need to have either very high internal gain, such as single photon avalanche diodes (SPADs) and photomultiplier tubes (PMT); or small thermal capacity/electrical resistivity such as superconducting single photon detectors. Avalanche photodiodes, based on the InGaAsP material system are a mature technology for detecting signals with wavelength from 0.9 -1.6 μm (near infrared). When applied to detect single photons, APDs are typically operated in Geiger mode (Section 3.1) and are referred to as Geiger mode APDs (GmAPDs). In Geiger mode, SPADs exhibit theoretically infinite gain limited by series resistance. Resetting the SPAD after it has been triggered to its "on" state by a photon is the function of the quenching circuit, which will be discussed in Section 3.2. Recent developments focused on reducing afterpulsing in InGaAs/InP SPADs have raised the possibility of linear-mode operation; Section 3.3 will explain the motivation for linear-mode operation and present preliminary simulations.

## 3.1 Geiger Mode Operation

A SPAD operated in Geiger mode functions as a photo-activated switch. However, the term Geiger mode APD is due to the similarity with a Geiger-Muller counter for radiation detection. The following will describe the process of generating a macroscopic signal from the incidence of single photons. When a photon is absorbed by a reverse-biased semiconductor photodiode, the photon generated electron-hole pair (EHP) can diffuse to

the high electric field region. If the electric field is high enough to initiate impact ionization, an EHP pair will be created. These secondary carriers can, in turn, impact ionize creating more EHPs. At breakdown this process become self-sustaining causing the APD to switch to a high current "on" state. When biased beyond the breakdown voltage, the photo current exhibits a saturated value that is determined by the external circuit and the bias conditions. This current level is usually at least several hundred μA, which results in an impulse with an amplitude of 10-100mV on a 50 ohms load. The voltage pulse is then amplified with a low-noise and wide-band amplifier before the signal is input to a photon counter.

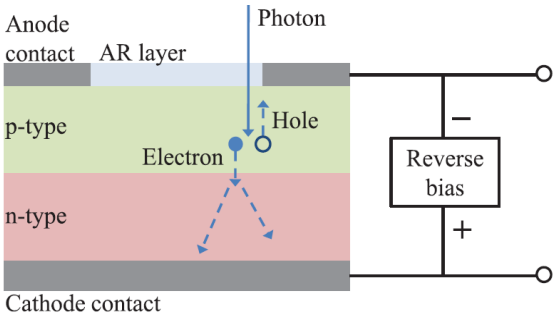



Figure 3.1: A layer-structure of a single-photon avalanche diode and its bias circuit. The reverse voltage is applied to accelerate the electrons toward the multiplication region.

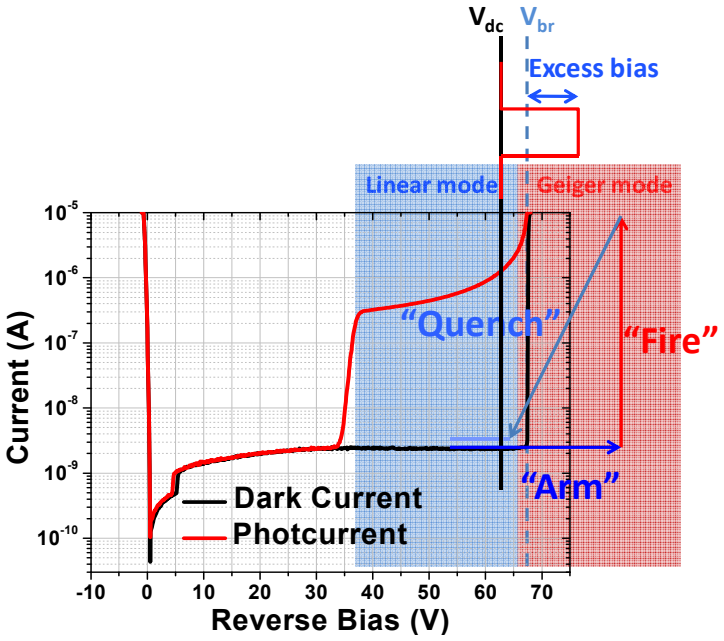



Figure 3.2: Typical current-voltage curve of a SPAD and illustration for linear and Geiger mode operations

The single photon counting process in Geiger mode consists of three steps, arm, fire and quench, as shown in Figure 3.2. In the "arm" stage, the device is biased above the breakdown voltage. Figure 3.2 also illustrates the biasing method for gated-mode operation. In gated mode, the DC bias is below the breakdown voltage and the AC bias increases the bias above breakdown periodically for a short time. In the armed state, the APD is in a state similar to unstable equilibrium. A single photon can initiate the self-sustaining avalanche event that switches the APD to the high-current "on" state. This is referred to as the "fire" step in Fig. 3.2. This triggers the counter to register a photon counting event. The readout circuit also provides a feedback signal that reduces the voltage and quenches the avalanche event. After the device is quenched, it also takes

some time for the device to recharge itself and recover the bias potential on its cathode, at which point it is rearmed to detect again.

It is clear that the detection process depends on the device characteristics and on the readout and quenching circuits. A major limitation for high-speed applications has been the afterpulsing effect (introduced in Chapter 2). There are two approaches to suppress afterpulsing. At the device level decreasing the trap density in the InP material system can reduce afterpulsing. The research in this direction has not been very successful in part because there are no obvious avenues to further improve the material quality of InP. The other method is to reduce the avalanche charge flow during the avalanche event. The latter has been achieved by modifying the quenching and readout circuits. We will see in the following sections how afterpulsing is closely tied to the quenching circuits.

## 3.2 Quenching Circuits

### 3.2.1 Passive Quenching

Passive quenching uses a large resistive load in series with the photodiode as shown in Figure 3.3. During the avalanche, a large voltage drop develops across the resistor $R_L$, which reduces the bias on the photodiode and turns the avalanche event off. Prior to quenching the avalanche current produces an impulse signal on the 50 ohm load for detection. This configuration only requires a single DC bias and is easy to implement. Recently, passive quenching has been realized monolithically by integrating a quenching resistor on the chip with the photodiode; in this manner negative feedback signal can be

provided by the diode itself, further reducing potential stray capacitance in the conventional module level integration.[2, 22]

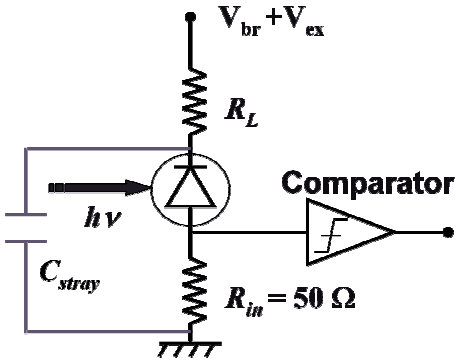



Figure 3.3: Passive quenching circuit configurations

The advantage of passive quenching is simplicity and fast quenching response. However, the recharge can be slow if the resistor $R_L$ is large [23]. The reverse DC bias for the diode is the breakdown voltage plus the excess bias. When a photon is detected, the bias on the SPAD will decrease by $I_{av} \times R_L$, where $I_{av}$ is the avalanche current. After the avalanche has quenched, the circuit will begin to recharge the diode capacitance ($C_d$) and stray capacitance ($C_{stray}$) through the load resistor, $R_L$. The time constant for the recharging is

$$R_L \times (C_d + C_{stray}) \quad . \tag{3.1}$$

$R_L$ is usually several hundred kilo-ohms in order to effectively quench the avalanche event. With a typical value of 1pF for the total capacitance ($C_d+C_{stray}$), the recharge time can be as long as several hundred nanoseconds to a few microseconds. On the other hand, the quenching time can be estimated as the total avalanche charge, $Q_{av}$ divided by the avalanche current $I_{av}$

$$Q_{av}/I_{av} = V_{ex} \times (C_d+C_{stray})/(V_{ex}/R_s) = R_s \times (C_d+C_{stray}) \tag{3.2}$$

where $R_s$ is typically a few hundred ohms. Therefore the quenching time is much shorter than the recharging time.

### 3.2.2 Active Quenching

Active quenching incorporates a feedback loop that controls the quenching and recovery of the diode with high precision timing. The basic idea is to sense the rise of the avalanche pulse and feedback to the diode with a controlled bias-voltage source within nanoseconds. Active quenching has been developed extensively for silicon SPADs using mature complementary metal-oxide-semiconductor (CMOS) integrated circuit (IC) technologies [24, 25]. As a result of monolithic integration and minimized stray capacitance, the hold-off time is adjustable at sub-microsecond level and the quenching time can be as short as 5 ns [26]. Recently, a compact single photon counting module has been developed for InGaAs/InP SPADs using active quenching mode [1]. Sub-nanosecond quenching has also been achieved with a fully programmable photon counting module [1, 27]. A drawback with this module, however, is that the biasing gate is long (~ 10ns) and it has an enforced hold-off time of at least 50 ns. Therefore the maximum count rate is 20 MHz.

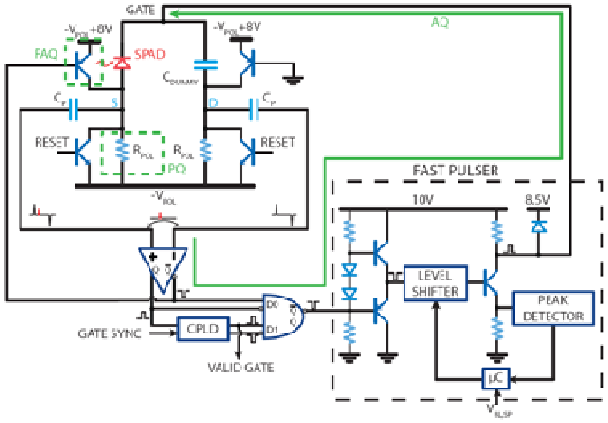



Figure 3.4: Front end electronics and fast pulser. Note the presence of passive quenching (PQ), active quenching (AQ) and fast active quenching (FAQ). At the meantime, differential mode detection is also adopted to minimize afterpulsing.[1, 2]

In order to utilize the fast quenching feature of passive quenching and overcome the slow recharge issue, Dr. Liu in our group developed a circuit that is referred to as "passive quenching with active reset" (PQAR), which combines fast passive quenching with fast active reset to achieve higher speeds and better performance [23]. The reduced average avalanche charge flow was estimated to be approximately 10 pC for each avalanche event [23]. The PQAR was extended to gated mode (gated-PQAR) by Dr. Hu, and has achieved detection of smaller avalanche pulses and higher operation speeds. The total charge flow per avalanche event was reduced by 10 times to approximately 1 pC [15, 28]

### 3.2.3 Gated Quenching

Gated quenching is most appropriate for synchronized applications. For gated quenching the APD is DC biased below breakdown. A periodic AC bias increases the

total bias beyond the breakdown voltage. In this manner, the detection window is well defined and can be synchronized with incident photons. Dark

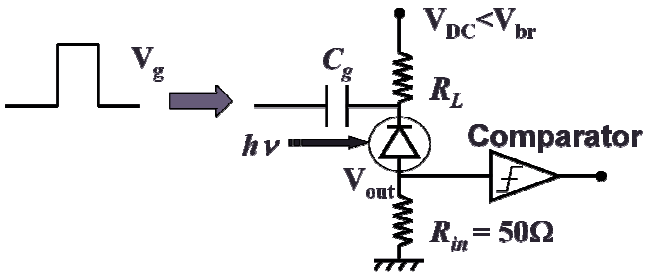



Figure 3.5: Gated quenching circuit configuration

counts and afterpulses are largely suppressed due to the fact that the device is only turned on during the detection windows. One major issue with gated quenching is the capacitive response of the diode. The transient responses occur at voltage spikes at the leading and trailing edges of the biasing gates (Figure 3.6). These spurious spikes can be larger than the avalanche pulse. In order to differentiate the avalanche pulse, excess bias has to be high enough to amplify the avalanche pulse to a significant and detectable magnitude. Figure 3.5 shows a typical avalanche pulse output with gated mode quenching. The peak of the transient at the leading edge of the gate is approximately 100 mV and the avalanche pulse has amplitude of 350 mV. The total avalanche charge during this avalanche event can be calculated as

$$Q = I_{av} \times T_{av} = 350\text{mV}/50\ \Omega \times 1/2 \times 5\ \text{ns} = 17.5\ \text{pC}.$$

Compared with the PQAR and gated-PQAR we mentioned before, gated quenching has larger average charge flow.

We will keep track of avalanche charge throughout this dissertation as an indicator of the improvement of the read out electronics and a key factor for suppression

of afterpulsing. Gated quenching is still used in current quantum key distribution protocols, however with more complicated variations. Chapter 5 will introduce those emerging techniques and their benefits toward reducing avalanche charge and afterpulsing.

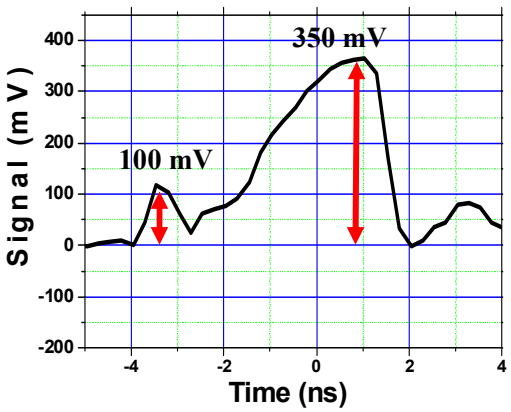



Figure 3.6: Typical output signal from gated quenching with avalanche pulse and transient responses.

## 3.3 Linear Mode

Given the limitations imposed by afterpulsing with Geiger mode InGaAs/InP SPADs, linear mode operation has been studied as an alternative way to realize single photon counting. In linear mode, the output photocurrent of the receiver is proportional to the input optical intensity. This has two advantages: this enables photon number resolution through the amplitude of the output signals and since the APD is always biased below breakdown afterpulsing is eliminated. However these advantages come at cost. A linear mode single photon counter has to cope with various noise sources due to lack of sufficient avalanche gain. The noise sources include the dark current and the excess noise from the APD as well as the noise from the following amplifiers. These noises contribute

to the total background noise from which the avalanche signal must be identified. Thus the noise level is critical consideration for linear mode detection.

The candidates for linear mode counting are APDs with very low excess noise, such as InAlAs APDs [29], InAs APDs [30, 31], HgCdTe APDs [32, 33], superconducting photo detectors, and photomultiplier tubes. These with the exception of InAlAs APDs these photodetectors usually have close to zero excess noise factors and therefore are very good candidates for high gain applications such as photon counting. The avalanche signal and the noise are both amplified through a secondary amplifier, which is usually a transimpedance amplifier (TIA). The receiver, which includes the photodetector and the TIA, are cryogenically cooled to a degree where the total noise is sufficiently small to permit single photon detection.

The necessary total gain from the APD and TIA combination is at least $10^5 – 10^6$. The APD gain is typically several hundred while the gain required from TIA is usually several thousand. They both depend on the conditions under which they are operated. Figure 3.7 shows the simulation of noise equivalent photons (NEPh) as a function gain for APD materials, with effective k in the range 0 to 0.1. The simulation has included contributions from excess noise, dark current, and TIA noise. The assumptions made here are that the output pulse has amplitude of 20 mV, pulse width 1 ns, total charge flow is 0.04 pC, the charge flow in the APD is 0.02 fC, the gain of APD is 125, and the gain of the TIA is 2000. Therefore the total gain for this linear receiver is $2.5 \times 10^5$. This total avalanche charge flow (0.04 pC) is still smaller than that achieved by most of current Geiger-mode gating schemes (~ 0.1 pC) [1, 25, 34]. As we can see in Figure 3.7, the smaller the value of k, the higher the gain at which the APD can be operated. It is well

documented that smaller k is beneficial for APDs operate at high gain. The curve with higher k value increases rapidly because the amplified excess noise overwhelms the avalanche signal. Thus a primary goal for linear mode counting is to design and discover APDs with low excess noise.

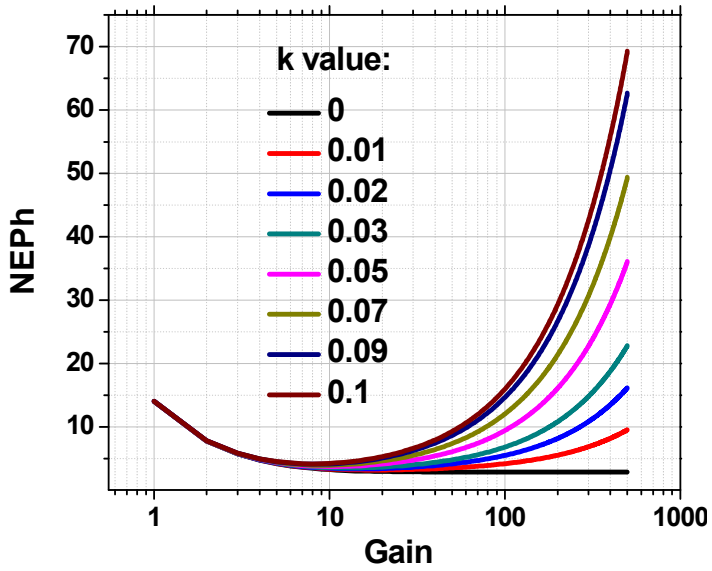



Figure 3.7: The effect of k value on noise equivalent photons for different gains.

The effect of APD dark current is also significant but in a different pattern. Figure 3.8 shows NEPh versus dark current at gain of 50 with different k values. The curve with higher k increases faster with dark current than the curve with smaller k. Dark current higher than 100 pA has a more significant impact of the rise of NEPh.

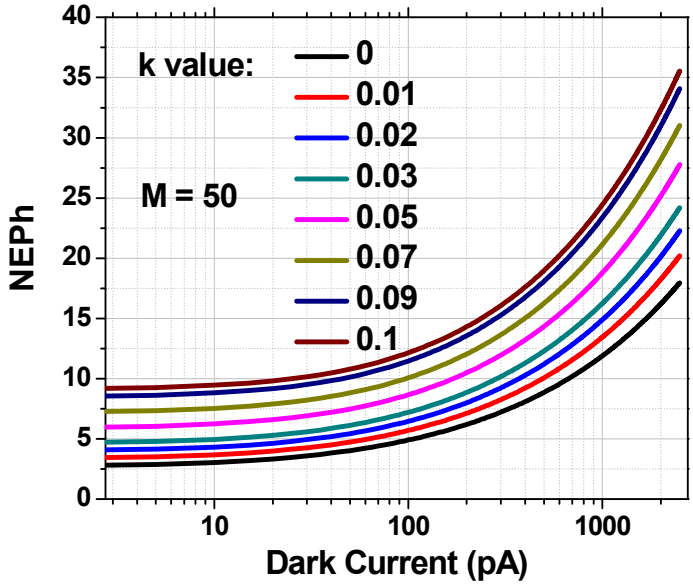



Figure 3.8: Dark current dependency of NEPh for different materials.

## 3.4 Conclusion

In summary, this chapter has introduced the basic idea for single photon counting with SPADs. The implementation of SPADs includes two operation modes, Geiger mode and linear mode.

Geiger mode operation is well developed mode and there are three conventional quenching circuits associated with this mode: passive quenching, active quenching and gated quenching. Passive quenching is the simplest and also provides very fast quenching; its primary disadvantage is slow recharging, which renders it unsuitable for high speed applications. Active quenching is the most complicated, but it does provide precisely controlled detection hold-off windows, however it is more costly due to the complexity of the feedback loop. An additional disadvantage is that it has an enforced hold-off time, which is usually too long to permit high-speed operation. Gated quenching

is widely used to characterize SPADs due to its well-defined parameters and uncomplicated implementation.

The ultimate performance of a single photon-counting receiver depends on the device characteristics and on the quenching and readout circuits. We have seen the impact of quenching circuits in this chapter. In Chapter 5, more sophisticated circuits will be presented. One of the primary motivations for the development of these quenching circuits is to reduce afterpulsing in order to enable high-speed operation.

Linear mode operation is an alternative way to realize single photon counting. It operates below breakdown voltage. However, amplifying the single photon signals without introducing noise is the primary challenge for linear mode. Currently the solution is to combine the gain from an APD and a transimpedance amplifier to achieve single photon sensitivity. Both the APD and the TIA are required to exhibit sufficiently low noise under the single photon counting operation conditions. Cryogenic cooling is necessary for most occasions. Theoretical and experimental research on linear mode is still ongoing. The focus areas include developing low excess noise APDs, designing low noise TIAs, and developing low noise packaging.

# Chapter 4 Geiger-Mode Operation of Ge on Si SPADs

This chapter presents the Geiger mode performance of a Ge on Si separate-absorption-charge-multiplication (SACM) SPAD fabricated by Intel. The operating wavelength is 1.31 μm. Si SPADs have exhibited high detection efficiency and low dark count probability in single photon counting [35, 36], however, the band gap of Si restricts operation to wavelengths <1 μm. One approach to extend the operating wavelength of Si-based SPADs is to utilize an SACM APD in which the multiplication region is Si with adjacent InGaAs [37] or Ge absorption regions [38]. The use of Si for the multiplication region is advantageous owing to its low excess noise factor and favorable avalanche breakdown probability. The benefits of Ge as the absorber include its compatibility with CMOS process technology and its long-wavelength cutoff ($\lambda \sim 1.55$ μm). An advantage of Si-based SPADs is that they can be integrated with CMOS circuits [35, 39]. This facilitates performance improvements in applications such as photon timing and monolithic focal plane array imaging. The challenge presented by Ge on Si for single photon detection is the relatively high dark current, which contributes to high DCR. Ge APDs tend to show high dark current due to the narrow band gap of Ge. More significant for Ge on Si, however, is the 4% lattice mismatch between Ge and Si. Recently, improved epitaxial growth techniques have enabled significant progress in the performance of Ge on Si APDs [38, 40, 41]. In this chapter, single photon detection is reported for Ge on Si SPADs.

The Ge on Si SPADs that were used in this study are normal-incidence structures with 30-µm-diameter active regions. Epitaxial layers of Ge were grown on (100) Si substrates by chemical vapor deposition (CVD) [38]. The wafer has a 1 µm-thick unintentionally doped (uid) Ge absorption layer and a 0.5 µm-thick Si multiplication layer (Figure 4.1). The multiplication layer is also uid to achieve a relatively constant high electric field. And the charge layer (0.1 µm) is p-doped with a concentration of $1.52\times10^{-17}$ $cm^{-3}$. Figure 4.1 also shows the complete circuit for the gated quenching operation.

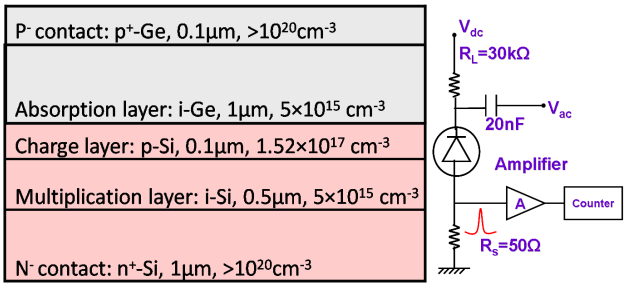



Figure 4.1: Schematic cross section of SACM Ge on Si APD and circuit for gated mode detection (right).

## 4.1 Dark current

For single photon counting, low dark current is very important in order to achieve low dark count rate and high detection efficiency. Several sources could be the origin of the dark current: Generation Recombination (G-R) current, band-to-band tunneling current, trap-assisted band-to-band tunneling, and diffusion current. Diffusion current is generated from the undepleted regions of the device; it is proportional to $n_i^2$, the intrinsic carrier density of the material:

$$I_{\text{diff}} \propto n_i^2 \propto T^3 e^{-\frac{E_g}{kT}} \qquad (4.1)$$

where $E_g$ is the bandgap of the material. Therefore, if diffusion current dominates, the activation energy extracted from a study of the temperature dependence of the dark current will be equal to the bandgap energy. While the G-R current originates from the depleted region of the device, it is proportional to $n_i$:

$$I_{\text{GR}} \propto n_i \propto T^{3/2} e^{-\frac{E_g}{2kT}} \quad . \qquad (4.2)$$

The activation energy extracted from equation 4.2 will be $E_g/2$, i.e., half the bandgap of the material. While the activation energy of tunneling current (band-to-band and trap-assisted) can be calculated from the temperature dependant bandgap and breakdown voltage, it is usually smaller than half of the bandgap [42]. It follows that the measured

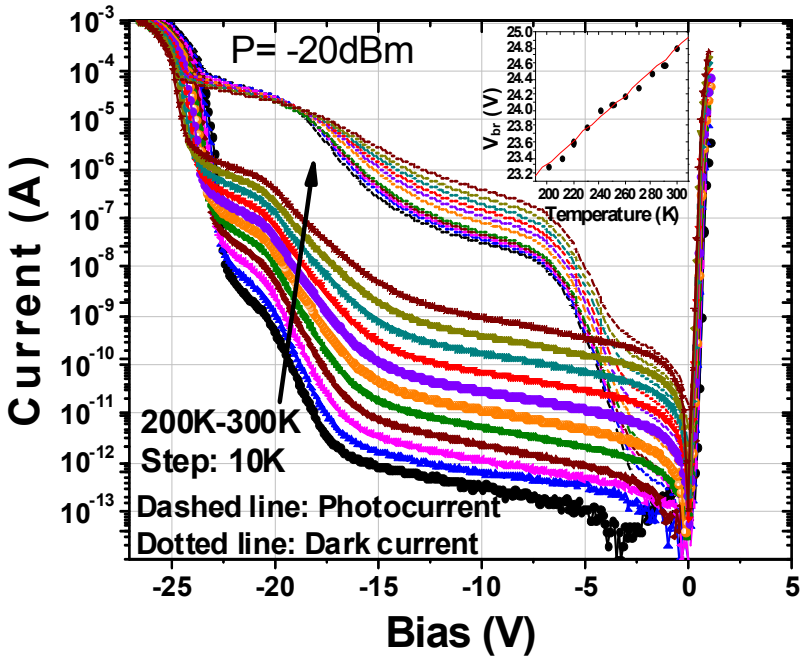



Figure 4.2: Dark current and photo current at different temperatures. Device diameter is 30 μm, photocurrent and dark current at the same temperature are in the same color. The incident light power was -20dBm.

activation energy will provide insight into the dominant contributor to the dark current.

Figure 4.2 shows the photo and dark current of the device under test in the temperature range of 200 K to 300 K. It appears that there are three distinct regions to the dark current curves. Below 15V, the current-voltage slope is smaller than that at higher voltage and the separation between the curves is larger. Reduced temperature dependence is observed at higher bias, a signature of the tunneling component of the dark current. The inserted figure shows the variation of the breakdown voltage with temperature. We somewhat arbitrarily define the breakdown voltage as the bias point where the dark current reaches 100 μA. The breakdown voltage increases with temperature; this is due to the temperature dependent photon scattering effect. The dark current at 200 K before

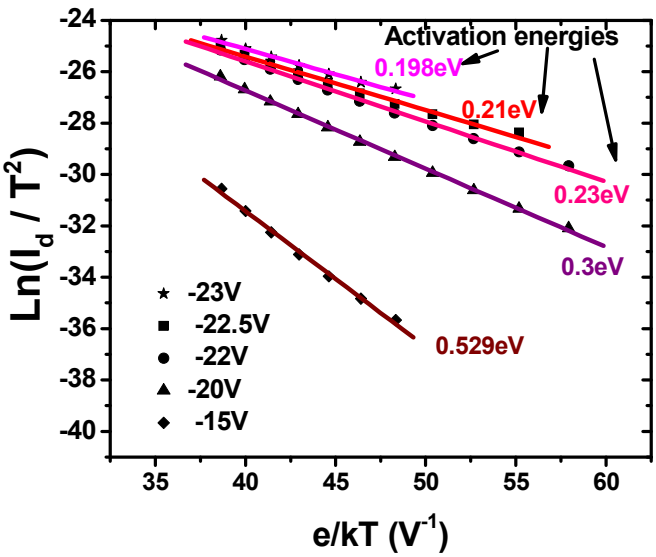



Figure 4.3: Logarithm of $I_d/T^2$ versus e/kT at different reverse bias. The absolute value of the slope of each linear fit equals the activation energy for the corresponding bias.

breakdown is approximately 10 nA, which is two orders of magnitude higher than that of InGaAs/InP SPADs.

The extracted activation energy from the curves in Figure 4.2 is plotted versus bias voltage in Figure 4.3. The activation energy decreases with increasing bias, from

0.53 eV to 0.2 eV. This is consistent with the conclusion that the dark current near breakdown is primarily tunneling and not generation-recombination [38]. The bandgap of Si is 1.12 eV and Ge is 0.66 eV. The total activation energy decreases from half the bandgap of Si to less than half the bandgap of Ge with increasing bias. This indicates that G-R dominates below 20 V; the origin of the G-R current is the Si layer at low biases and the Ge layer at higher biases [43, 44].

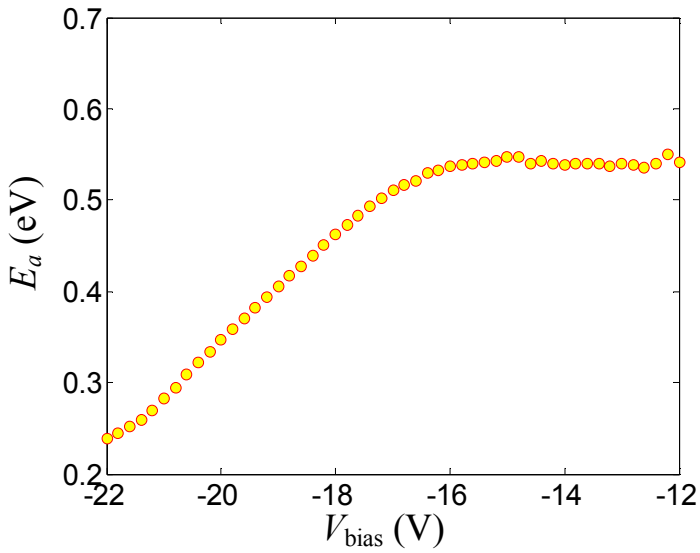



Figure 4.4: Activation energy extracted from temperature dependent dark current versus bias voltage.

Figure 4.4 plots the activation energy from -12 V to -22 V. The activation energy drops from 0.52 eV at low bias to 0.25 eV at 22V.

The change in G-R current is due to variation in the electric field change as the depletion moves from Si into the Ge absorber. In devices with SACM structure, the absorber layer is usually a material with a narrower bandgap than the multiplication layer.

In order to suppress the dark current, the designed electric field in absorber is smaller than that in multiplication layer. This electric field profile is accomplished with a charge layer. When the bias increases, the charge layer between the absorber and multiplication layer will gradually deplete. When the charge layer is totally depleted, the electric field in absorber starts to increase. Punch-through can be defined as the voltage at which the edge of the depletion reaches the absorber. Above the punchthrough voltage the electric field in absorber increases with the increasing bias. The punch-through point can be extracted from capacitance-voltage (CV) measurements (Figure 4.5). There is a drop in CV curve shown in Figure 4.5, at the bias point of 21 V – 22 V. This is the punchthrough point [38]. Comparing with the data in Figure 4.4, it can be seen that the punch-through point is consistent with the point where the activation energy drops.

Figures 4.6 and 4.7 show the photocurrent and gain versus voltage, respectively,

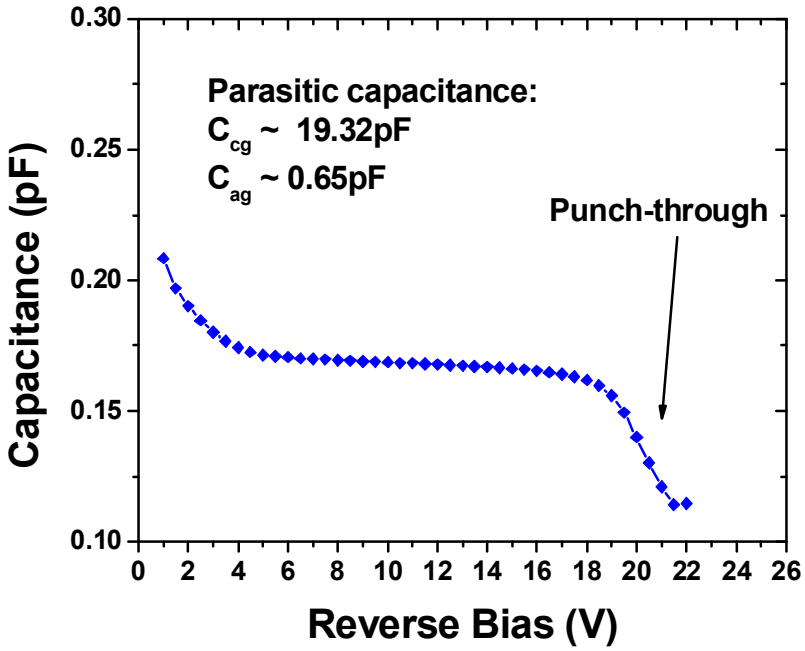



Figure 4.5: CV curve and parasitic capacitance estimation.

for temperature in the range 10 – 60 °C. The photocurrent peak in Fig. 4.6 is due to the space charge effect in the depletion region. From Figure 4.7, it can be seen that the maximum gain occurs at the same voltage, $V_0$, as the photocurrent peak in Figure 4.6. For a given bias voltage $|V| < |V_0|$, the gain increases as the temperature decreases. This is due to the fact that the ionization rate increases as the temperature decreases. The gain increases by approximately 1.3 for a temperature increase of $\Delta T = 10°C$. At higher bias voltages, the gain is more sensitive to the temperature change [43].

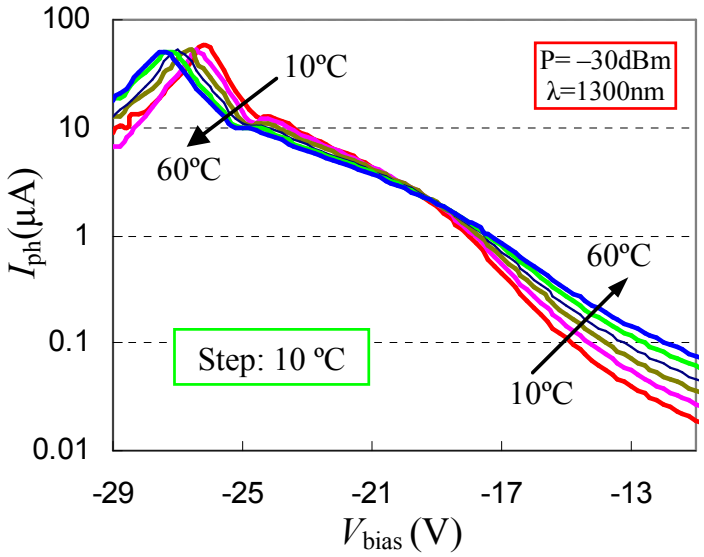



Figure 4.6: The photocurrent $I_{ph}$ versus voltage at different temperatures.

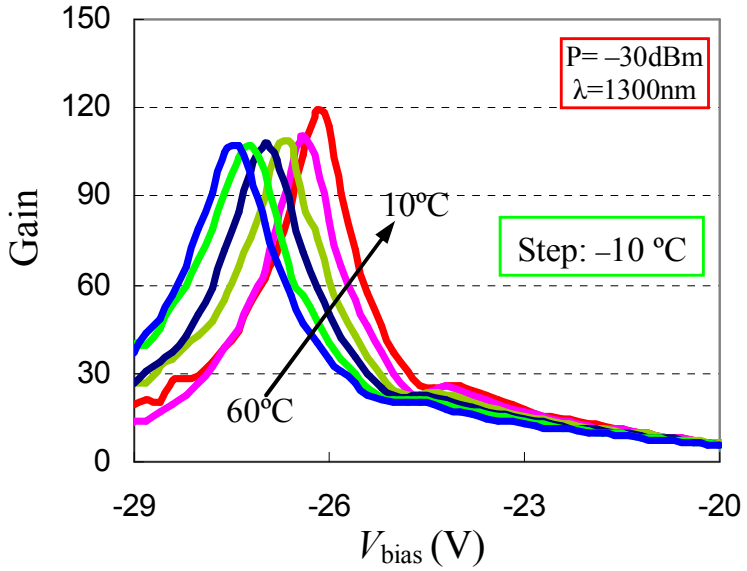



Figure 4.7: Gain versus voltage for temperature in the range 10°C to 60°C.

The dark current is significantly higher than that of commercial InGaAs/InP SPADs. In order to reduce the dark current in Ge on Si APDs it would be beneficial to have high quality Ge growth with fewer defects. An alternative approach would be to design devices with punch-through voltage that is closer to breakdown voltage, in which case, the electric field in the absorber would be lower. The disadvantage of this approach is that the responsivity of an APD with the SACM structure strongly depends on the relation between the punch-through voltage and the breakdown voltage. If the breakdown voltage is lower than the punch-through voltage, the device will breakdown before the electric field in the absorber high enough to aid carrier injection into the multiplication region. In this case, the responsivity would be adversely affected.

## 4.2 Dark count rate and photon detection efficiency

To achieve a lower DCR and acceptable PDE it is necessary to reduce the temperature. For this work, 200K was selected because at lower temperature, the breakdown voltage approaches the punch-through voltage. When the breakdown voltage is lower than the punch-through voltage, the responsivity decreases because fewer carriers surmount the heterojunction barriers, which degrades the PDE.

Another reason for using 200K is that the DCR saturates at lower temperature. Figure 4.8 shows the DCR versus excess bias at 180K and 200K. The dark count rate ceases to decrease for temperature < 200K, this can be explained as follows: At low temperature the thermal generation rate, which decreases with temperature, becomes less than the tunneling rate, a dark current mechanism with weak temperature dependence. This is consistent with the current-voltage measurements that indicate tunneling is the dominant source of dark current at high bias. It is well known that the DCR is strongly

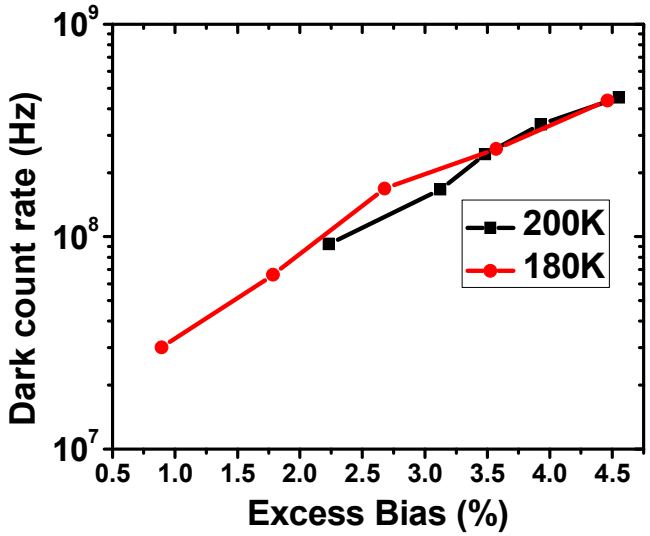



Figure 4.8: Dark count rate versus excess bias at 180 and 200K. Dark count rate is calculated using equation (4.3).

correlated to the dark current. If we compare the dark current of the Ge on Si APDs to that of InGaAs/InP SPADs, [45] we see consistency between the relative dark current and the DCRs of the two types of SPADs.

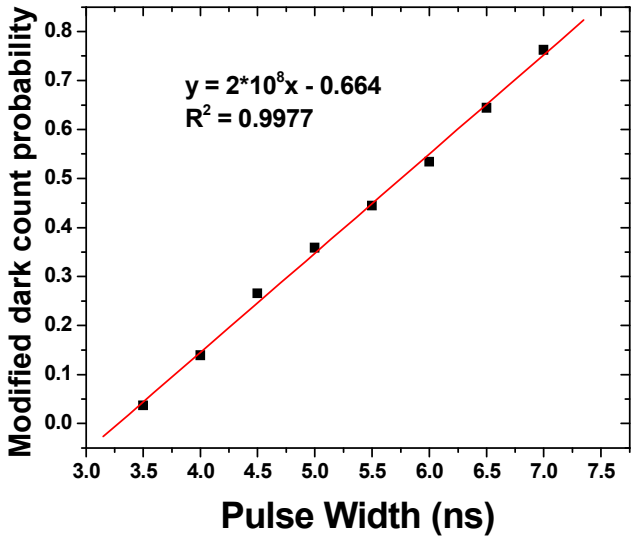



Figure 4.9: Fitting of modified dark count probability of different pulse width for dark count rate.

The optical source for these measurements was a pulsed laser diode that produced 50 ps optical pulses at a wavelength of 1.31 μm. The pulse repetition rate was 100 kHz. The optical pulses were synchronized with the gate pulses by adjusting the delay of the laser, which was attenuated to 1 photon per pulse. The actual number of photon-generated carriers and the number of dark carriers in each pulse obey a Poisson distribution. The dark count probability, $P_d$, is given by the expression stated in Chpater 2 and repeated here as equation (4.3):

$$P_d = 1 - \exp(-DCR \times \tau_e) \qquad (4.3)$$

where $\tau_e$ is the effective pulse width and DCR is the dark count rate. In this work, the difference ($t_d \sim$ 3ns) between the applied pulse width and the effective pulse width is

significant compared with the applied pulse width. Therefore the minimum applied pulse width is constrained by $t_d$. The relatively large value of $t_d$ is due to the large stray capacitance ~ 19pF, compared to that for chip-to-chip bonding (~2 pF). In addition, the avalanche build-up time of the device, the rise and fall times of the applied pulse, and system delay in the circuit can also cause increase of $t_d$. From the measured dark count probability for a given applied pulse width, $\tau_p$, the following expression can be employed to determine the dark count rate and the effective pulse width $\tau_e = \tau_p - t_d$ ,

$$DCR \times \tau_e = -\ln(1 - P_d) \quad . \qquad (4.4)$$

The fitting curve is shown in Figure 4.9. The photon detection efficiency (PDE) can be expressed as

$$\mathrm{PDE} = \frac{1}{n} \ln\left(\frac{1 - P_d}{1 - P_t}\right) \quad . \qquad (4.5)$$

where n is the average number of photons per pulse and $P_t$ is the total count probability.

The PDE can also be defined as the product of the external quantum efficiency (QE) and the breakdown probability, assuming efficient carrier transport from the narrow band gap absorber to the multiplication layer. The responsivity of the Ge on Si APDs is 5.6 A/W at the punch-through voltage ~ 22 V. In this case, the punch-through voltage is close to breakdown and the electric field at punch-through is large enough to produce gain in the multiplication layer. Thus, an estimation of quantum efficiency necessitates decoupling from the avalanche gain. We estimated the quantum efficiency as 50% based on the measured responsivity of 0.55 A/W of p-i-n photodiodes fabricated on the same wafer as the APDs [38]. The p-i-n structure has the same absorbing layer thickness as the SACM structure, which implies the two structures have the same quantum efficiency.

However the SACM structure includes a charge layer, which the p-i-n structure does not. The responsivity and QE can be improved by using a thicker Ge layer [38]. Figure 4.10 shows the DCR versus PDE at 200K. The highest PDE measured was 14%. We anticipate that future implementations with lower dark current and higher responsivity can provide

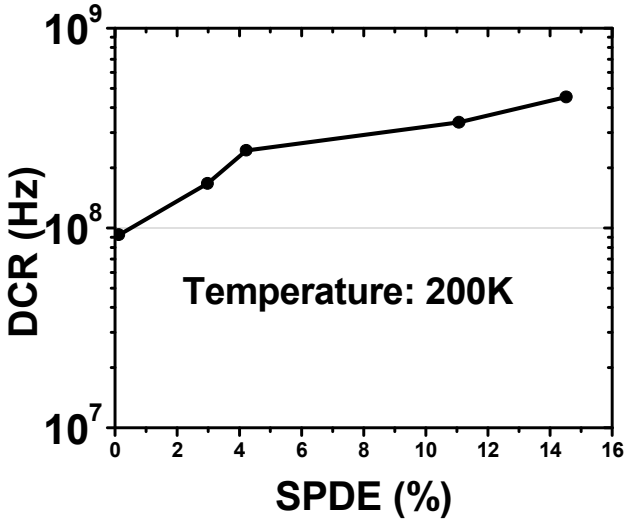



Figure 4.10: Dark count rate versus SPDE at 200K. The excess biases range from 0.5 V to 1.02V.

better single photon detection performance.

Figure 4.11 plots the NEP calculated from DCR and PDE with equation (2.7). The NEP value is approximately two orders larger than that of an InGaAs/InP SPADs. This is again consistent with the two orders higher dark current near breakdown.

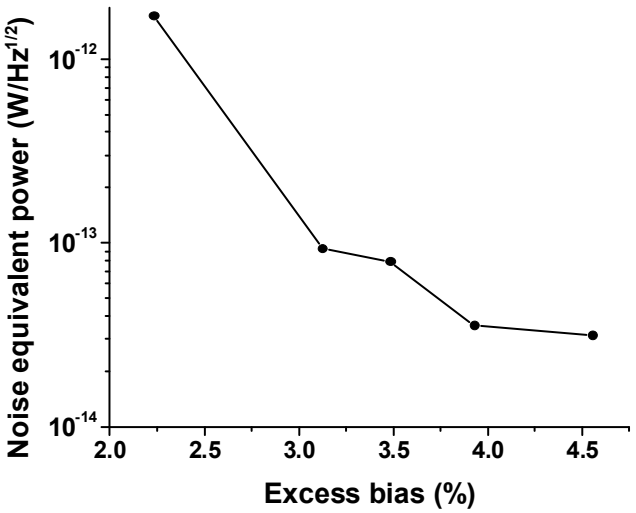



Figure 4.11: Calculated NEP from the measured DCR and PDE.

## 4.3 Afterpulsing and jitter

As stated in Chapter 3, a signature of afterpulsing is an increase in the dark counts when the time between the adjacent pulses, referred to as hold-off time, is less than the emission lifetime of the trapped carriers. When a SAPD is operated under certain excess bias and temperature, the trend of dark count rate versus operation frequency reflects the effect of after pulsing. Figure 4.12 shows the dark count rate versus frequency from 1 kHz to 1 MHz. We note a slight decrease in the dark count rate at a higher frequency. There are two reasons that the dark count rate decreases with increasing operation frequency. One is the RC effect. During an avalanche event, the avalanche current discharges, $C_g$, the capacitor in the bias tee used in gated quenching shown in Figure 4.1. During the hold off time, $C_g$ is recharged by the current through the load resistor $R_L$. The time constant ($R_LC_g$) is large compare to the hold off time (~ 10μs) at high frequencies (~100 kHz). As a result the 20 nF capacitor ($C_g$) is not fully recharged during the hold off

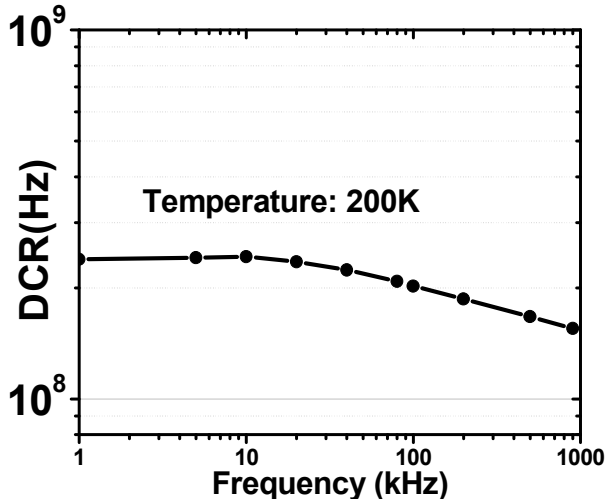



Figure 4.12: Dark count rate versus frequency with at 200K with 1V excess bias.

time; therefore the actual DC bias on the diode is lower than the set value. The other effect is that with a larger duty cycle, the average DC bias level (the reference DC component) is higher than that at lower frequencies. This causes the excess bias to be slightly lower at high frequency than at low frequency and results in a drop in dark count rate versus operation frequency.

The fact that the dark count rate does not increase up to 1 MHz is an indication that after pulsing is not a significant contributor to dark count rate in this frequency range. The after pulsing shown in Ge on Si SPADs is less significant than that in InP based SPADs operating in the same gating mode [8]. Less afterpulsing in Si-based SPADs is primarily due to the high quality of Si as the multiplication layer. Si-based SPADs have the potential to greatly suppress afterpulsing. In addition, the shorter the applied pulse width, the less serious after pulsing. This is due to the fact that there is less charge flow integrated over time during the avalanche for shorter applied pulse widths. At high frequency, the large number of avalanche events caused by dark counts results in a slight lowering of the excess bias, which causes the dark count rate to drop slightly.

Jitter, which is also referred to as timing resolution, is defined as the uncertainty in the time that an avalanche event occurs with respect to the fixed arrival time of incident photons. For this measurement, the pulse width was 6 ns. The temporal

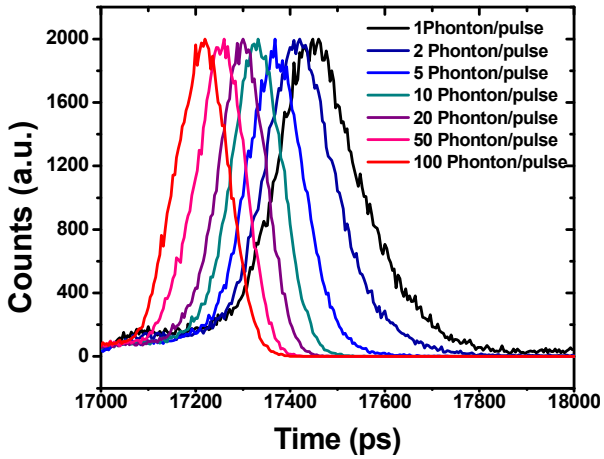



Figure 4.13: Jitter with different number of incident photons, shown in the histograms of the temporal response for a range of incident photon intensities. The absolute time that an avalanche event is detected shifts to the left when the number of incident photon increases.

distribution of avalanche events was measured using a multichannel analyzer with a resolution of 4 ps. The full-width at half-maximum (FWHM) of the laser pulse was 50 ps. The jitter was observed to decrease with increasing excess bias [8]. Figure 4.13 shows the histograms of the temporal response for a range of incident photon intensities. The FWHM of the histograms (jitter) at 1V excess bias with 1 photon per pulse is 195ps. It decreases to 117ps with 20 photons per pulse.

## 4.4 Conclusion

In conclusion, Geiger-mode operation of a 30 μm-diameter Ge on Si SPADs has been characterized using gated-mode quenching. The dark count rate (DCR) at 200K was ~ $10^8$ Hz, with single photon detection efficiency of 14%. After-pulsing was

characterized by dark count rate versus frequency; in the frequency range from 1 kHz to 1MHz, it is not a significant performance factor for these APDs. The jitter at 1 V excess bias was 195 ps with 1 photon per pulse at 200K.

# Chapter 5 Emerging Technology in Counting Electronics

Much of the research on InGaAs/InP SPADs in the past decade has focused on the counting electronics. The demonstrated capability of these emerging systems to suppress afterpulsing has stimulated work on InGaAs/InP SPADs for high-speed applications. The improved detection schemes include self-differencing, sine-wave gating, matched delay line, dummy path, and balanced detection. This chapter will describe these approaches and present state-of-the-art performance for each.

## 5.1 Self-Differencing

Self-differencing was first shown to be effective in suppressing afterpulsing at high gating frequencies in 2007 [46]. The first demonstration of self-differencing for single photon counting with InGaAs/InP SPADs was reported in the same year [47]. Subsequently the gating frequency has been increased to 2 GHz [20]; ultra short dead time was also reported with this technique [48]. The record low avalanche charge flow was 0.035 pC reported in Ref. [20].

The first version of self-differencing circuit is shown in Figure 5.1, where (a)-(d) and (f) show the input and output signals during the signal processing. The basic idea is split the output from the APD. The signal in one arm is inverted and delayed by one gate period. The two signals are then recombined. This effectively reduces the capacitive

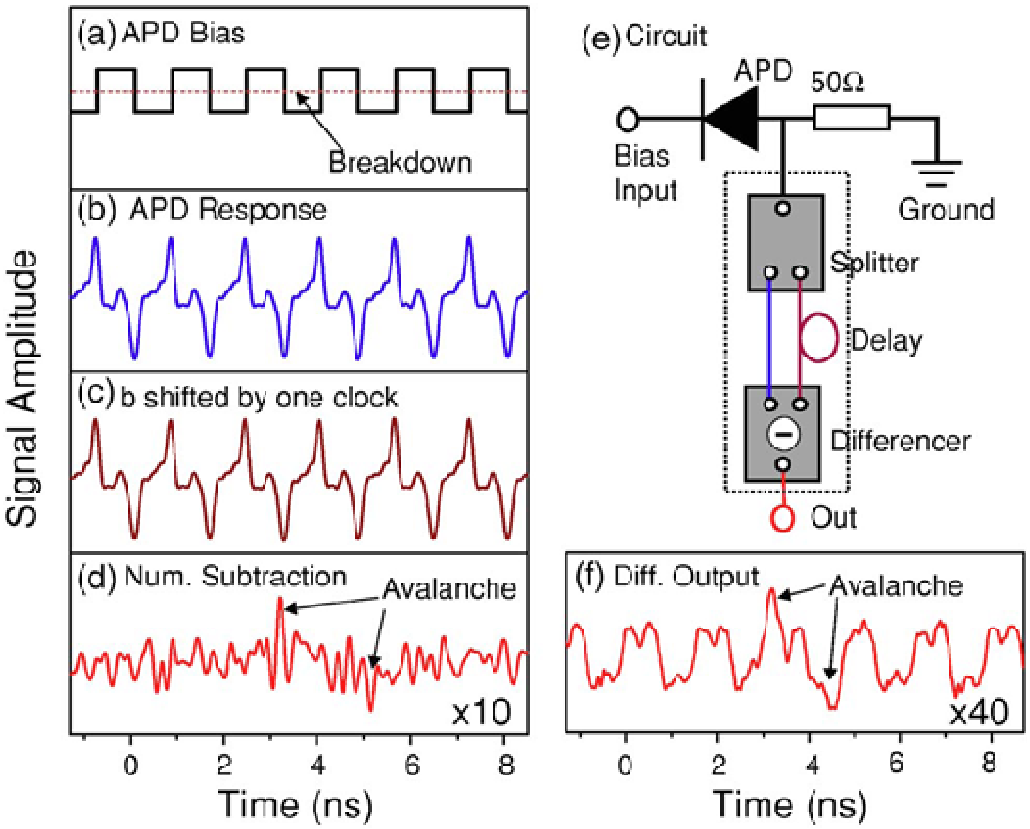



Figure 5.1: (a) Series of biasing square wave gates (solid line) applied to an APD. The dashed line indicates the APD breakdown voltage. (b) An APD response to the square wave gates. Note that no avalanche is visible. (c) Same response as (b), but shifted by a clock period. (d) Numerical subtraction (b) – (c) leaving the avalanche signal visible. Vertical scale in (d) is scaled up by a factor of 10 as compared to (b) and (c) for clarity. (e) An electrical circuit to realize the self differencing. (f) Output of the self-differencer. Vertical scale here is scaled up by a factor of 40. [47]

transient response of the SPAD by 21 dB, which permits the detection of weak avalanche pulses. Figure 5.1 (e) shows the splitter, delay line, and comparison circuit.

The subtraction or cancellation can be done either numerically or experimentally. Figure 5.1 (d) and (f) show the numerical and experimental outputs, respectively. In high-speed applications, numerical signal process is impractical. Therefore differential detection, as illustrated in Fig. 5.1 (e) is preferred. In Figure 5.1 (d) and (f), there are two

avalanche pulses, one positive and one negative. They represents the same original avalanche pulse, the negative pulse originates from the inverted and delayed signal. In the first reported self-differencing circuit, continuous frequency tuning was not possible due to the lack of adjustability in the branches of cancelling signals in the circuit [47]. In Ref. [20], where the best results were reported, the authors achieved better cancellation by adding a potentiometer to one arm of the split signal and in the delay line a line stretcher was added with a delay range of 40 ps, which enabled frequency tuning from 0.987 to 1.033 GHz. The added tunability is very important in order to achieve low background noise. Figure 5.2 (b) shows that the noise background for all three gating frequencies is much smaller compared with that in the first version. (Figure 5.1) This makes the detection of even smaller avalanche pulses possible, which is beneficial for reducing afterpulsing. The best result reported for self-differencing is 23.5% PDE with afterpulse probability of 4.84% and dark count probability of $1.32 \times 10^{-5}$ at a gating frequency of 2 GHz. We note that the laser repetition rate used to characterize the system was 1/64 of the gating rate, i.e., 31MHz.

## 5.2 Sine-wave gating

Sine-wave gating is another gating scheme for GHz-range gating frequency. Compared with conventional gated mode, sine-wave gating replaces the square-wave

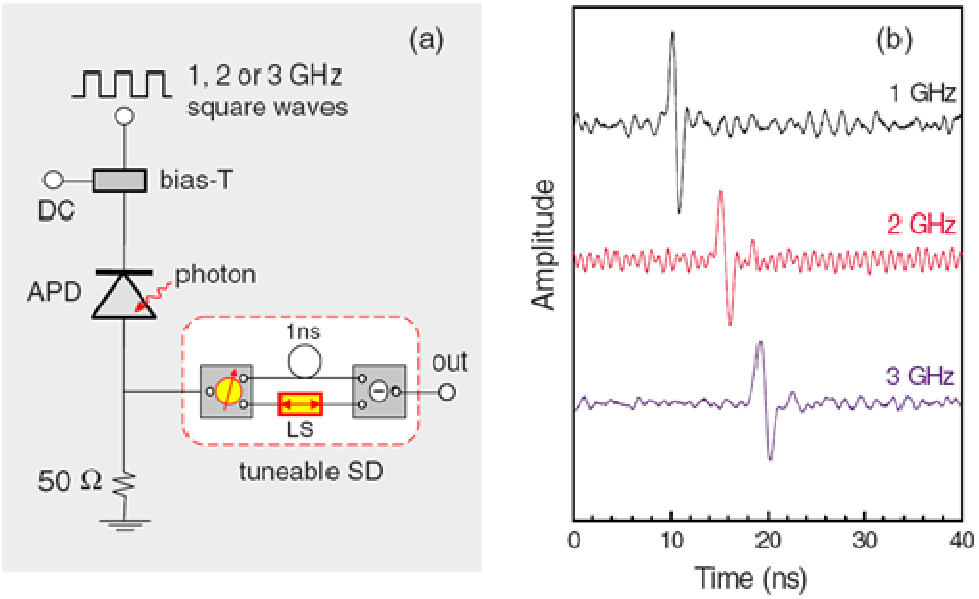



Figure 5.2: (Color online) (a) Schematic for a tunable SD circuit with 1-ns delay; LS: line-stretcher. (b) Avalanche waveforms for dark events recorded by an oscilloscope after the 1-ns SD circuits under different gating frequencies. The waveforms are plotted in the same scale but shifted vertically for clarity. [20]

gates with sine-wave gates. The advantage of sine-wave gating is that the capacitive response of the SPAD is also a sine-wave (cosine) signal with the same frequency as that of the gating signal. Therefore one or several narrow band notch filters can remove it. Similar to square wave gating, by filtering the capacitive response extremely small avalanche pulses are accessible. A typical experimental set up for sine-wave gating is shown in Figure 5.3. [49] The signal generator (SG) produces a gating signal at the designed frequency and the synchronization signal. The gating signal is amplified (HP-AMP) and passed through a band pass filter (BPF) before it is coupled through a

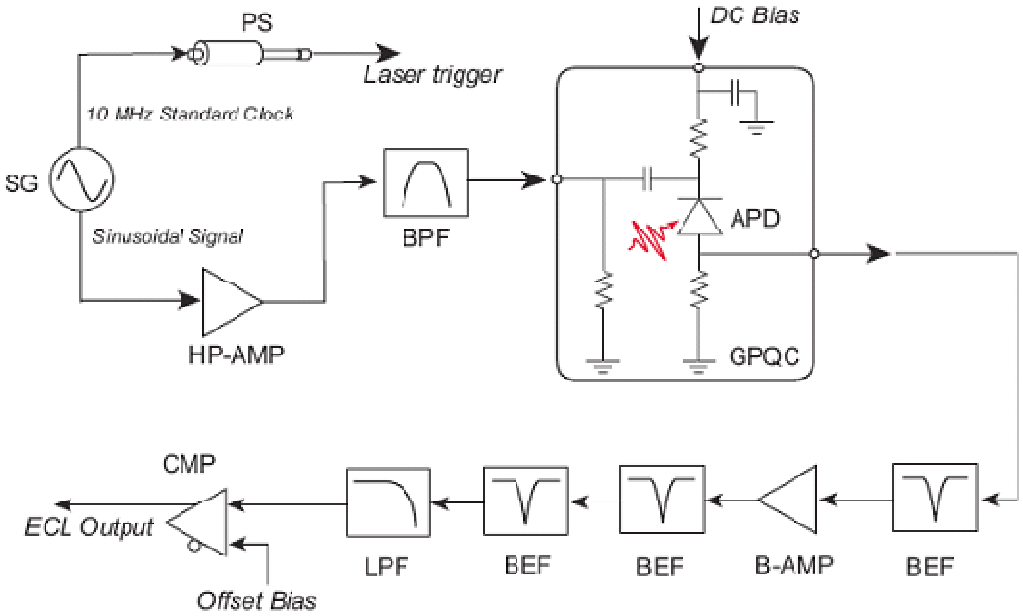
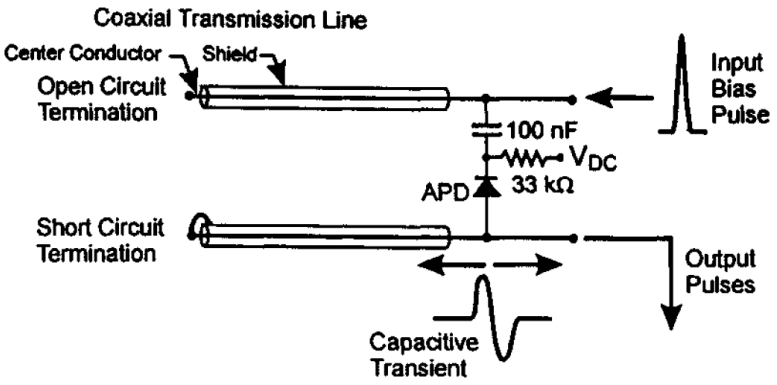



Figure 5.3: Diagram of the single-photon detection circuit. SG: signal generator, HP-AMP: 1 W class high-power amplifier, BPF: bandpass-filter, GPQC: Gated passive quenching circuit, BEF: Band-elimination filter, B-AMP: 3 GHz broadband inverting amplifier, LPF: low-pass-filter, CMP: comparator (discriminator). [49]

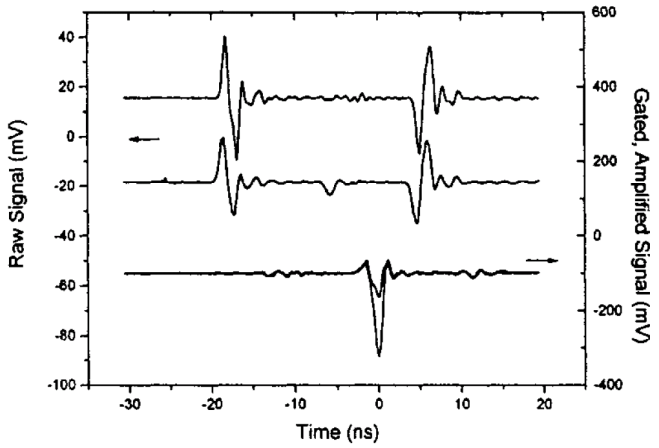
capacitor into the SPAD. The synchronization signal serves as input to the laser diode driver. In this manner, the gating signal and the laser pulses can be synchronized.

To remove the capacitive response in the output signal, several band elimination filters (BEFs) are needed. In Ref. 20 the total rejection at the center frequency was 100 dB and the background noise was ~ 20 mV. Like the self-differencing technique, sine-wave gating has been successful in detecting extremely small avalanche pulses. However, one difference is worth noticing: the frequency of the sine-wave gating determines the sine-wave gate width. As a result, in order to obtain narrow pulses and suppress charge flow, the frequency of sine-wave gating is usually required to be in the GHz range. After the technique was demonstrated at 800 MHz gating rate [50], all subsequent work has been done with GHz gating rates [5, 49, 51, 52]. In order to achieve better capacitive

response cancellation, a combination of self-differencing and sine-wave gating has also been reported with higher laser repetition rate (~77MHz) [53]. The highest laser repetition rate to date (315MHz) used both sine-wave gating and self-differencing [54], The dark count probability was $1.5 \times 10^{-5}$ with PDE of 11% at -20 °C. One key factor to achieve low noise background is to have high filter rejection ratio. The higher the filter rejection, the smaller the noise floor can be, and the smaller the detectable avalanche pulses. In Ref. [54], careful adjustment of the cancelling cables and housing electronics yielded 20 dB filter rejection ratio.




(a)




(b)

Figure 5.4: Single photon detection. (a) Circuit diagram. (b) Oscillograms showing the cancellation of transients (upper trace), the photon signal against a flat baseline (middle trace), and the photon signal and residual transient after amplification and electronic gating (lower traces). [19]

## 5.3 Matched delay line

The matched delay line technique was developed by IBM [19] and developed into a commercial product by Princeton Lightwave Inc. [18]. The highest reported laser repetition rate for this technique is 50 MHz.

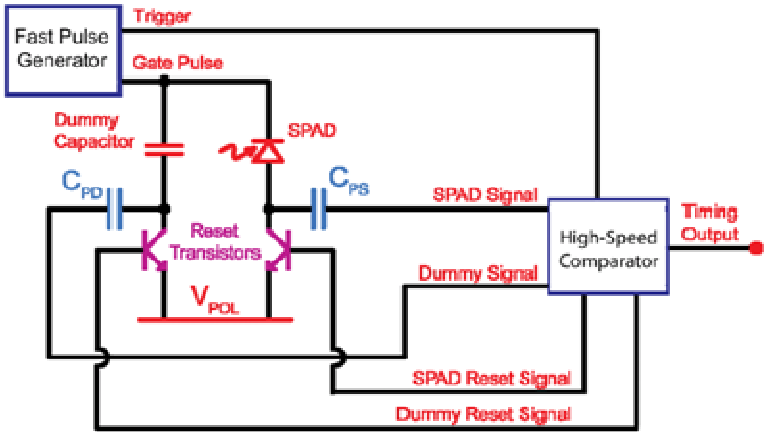
Figure 5.4 (a) shows the circuit diagram of the apparatus in Ref. [19]. The transient cancellation is illustrated in Fig. 5.4(b) [19]. In Figure 5.4 (a), both the anode and cathode of the APD are connected to SMA tee connectors. The cathode side, where the input bias pulse is incident, has an open circuit termination at the end of a 122 cm-long cable. On the anode side, the impulse response is split into two parts: one part travels along another 122 cm-long cable. At the end it is reflected back to the anode with inverted polarity. The signals from the two transmission lines have the same delay but are of opposite polarity. Consequently they cancel out the common mode signal, the capacitive responses of APD, as illustrated in the lower trace of Figure 5.4 (b). The photon induced pulse from the APD (center of the middle trace), on the other hand, is not cancelled. The residual noise background is very small (~ 10 mV). Therefore this circuit is capable of detecting avalanche signal as small as 20 mV. This type of cancellation is also very robust: it is independent on the pulse height, width, and shape. It even works for lossy transmission lines as long as the losses in the two lines are equal and well matched in phase.

From the matching process of this technique, one can see some similarity with self-differencing. Both techniques combine signals with opposite polarity in order to cancel the common mode capacitive response. This idea is quite effective in gated mode operation of SPADs.

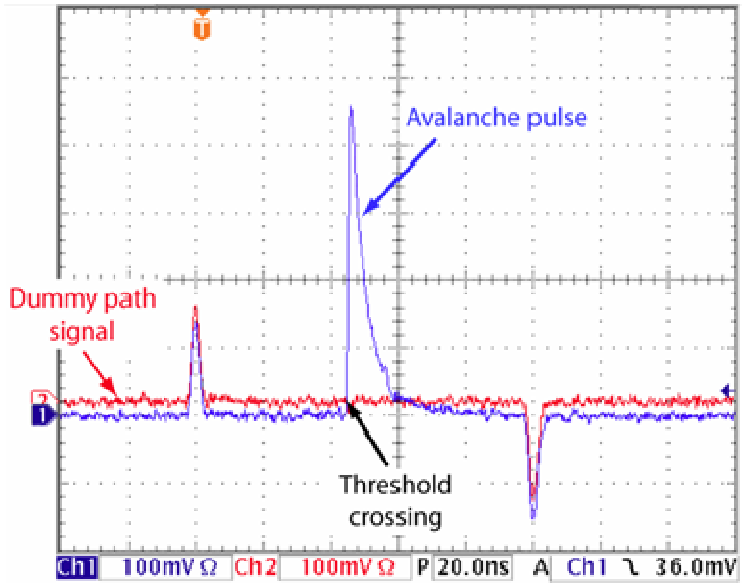
## 5.4 Dummy path

Dummy path was first discussed in S. Cova's review paper published in 1996 [26]. His group successfully implemented this technique with Si and III-V SPADs [55, 56].

Very similar to the idea in matched delay lines, the dummy path is used to generate a spurious spike that is identical in amplitude to the capacitive response of SPADs. The spurious spike is generated by a dummy capacitor that is biased to have the same charge as the SPAD. A circuit diagram is shown in Figure 5.5 (a). The dummy signal and SPAD signal inputs to a high-speed comparator. The comparator also produces a reset signal for both the SPAD and the dummy capacitor. This technique has been integrated into active quenching circuits and is referred to as differential mode sensing [1].




(a)




(b)

Figure 5.5: (a) Fast-gating electronics: the pulse generator provides gate pulses to enable the detector, the differential pick-up network allows precise avalanche detection, and the high-speed comparator drives the reset transistors to rearm the SPAD after each ignition[55]. (b) The signals at the input of the timing electronics.

## 5.5 Balanced detection

Akihisa Tomita and Kazuo Nakamura proposed a balanced configuration for cancelling the transient spikes in the SPAD signals [57]. They utilized the balanced output from two APDs for qubit discrimination. The schematic in Figure 5.6 shows the configuration of the two APDs and the post signal processing circuit. In the balanced configuration two APDs share the same bias and their outputs are identical except only one of the APDs exhibits an avalanche signal. A 180°C hybrid junction combines the two signals and inverts one of the signals to the opposite polarity. The common mode transient spikes therefore are cancelled. The discriminator compares the avalanche

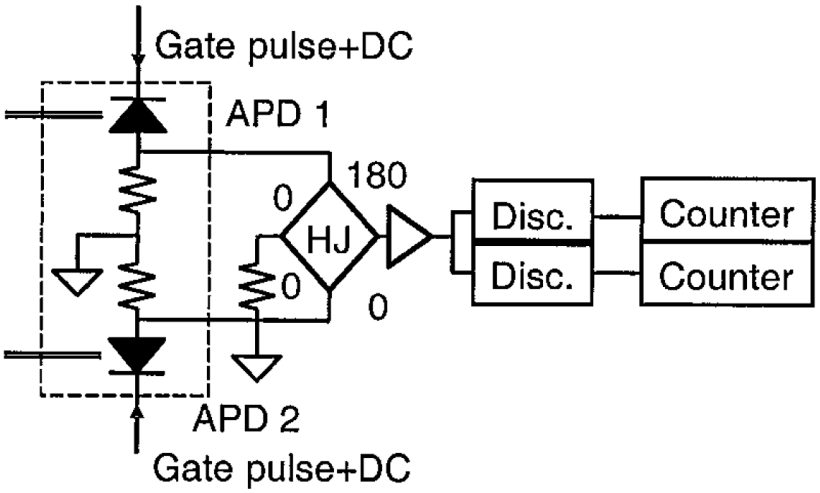



Figure 5.6: Schematic of the photon detector, HJ: hybrid junction, Disc.'s: discriminators. [57]

signals with a preset threshold and sends out a register signal to the counter.

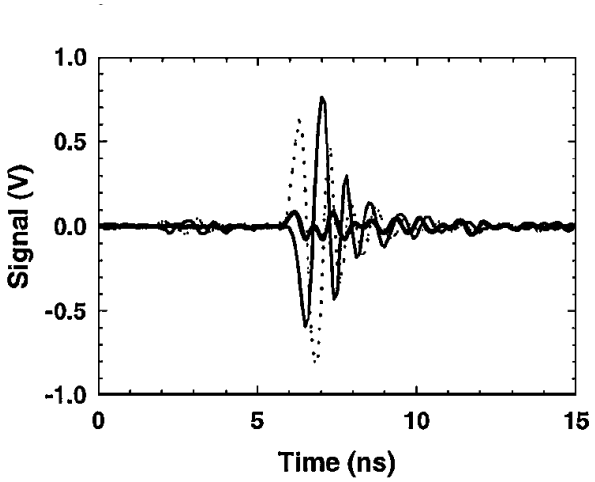



Figure 5.7: Cancellation of the transient spike: thin solid curve, APD 1, dots, APD 2, thick solid curve, differential output of the APD 1 and 2. [57]

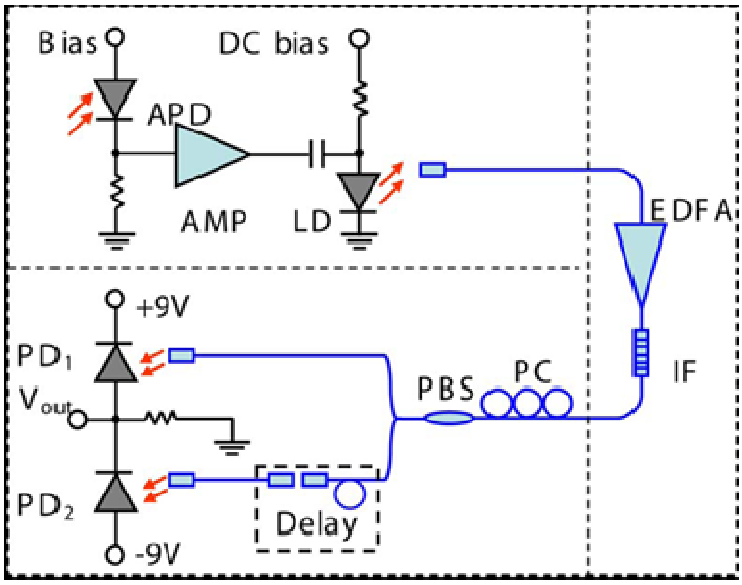



Figure 5.8: Schematic setup of the optical self-balancing single-photon detector. AMP: RF amplifier (10); LD: distributed-feedback laser diode at 1550 nm; IF: inline fiber filter at 1550 nm with pass bandwidth of 3 nm; PC: fiber polarization controller; PBS: fiber polarization beam splitter; and PD1 and 2: pin photodiodes. [58]

Figure 5.7 shows the cancellation result of the two APDs as reported in Ref. [57]. The residual noise signal had amplitude of 100 mV. Compared with the 600 mV transient spikes, the detectable avalanche signals were much smaller (~ 200 mV) due to the cancellation effect.

Another approach is to generate the balanced output optically; Jian *et al*. has demonstrated this idea experimentally in 2010 [58]. In their work, they used the APD signal to drive a laser diode and split the optical signal into two parts for comparison. Compared with electronic balancing circuit, the optical method can be very stable and precisely controlled with various optical components (Figure 5.8) and provide immunity to electromagnetic field of the ambient circuits. A 31 dB common mode rejection ratio has been achieved with this method. Figure 5.9 shows the oscilloscope traces of various signals. Figure 5.9 (e) represents the avalanche signal and the noise background. The detectable avalanche was as small as 20 mV since the noise background was smaller than 10 mV.

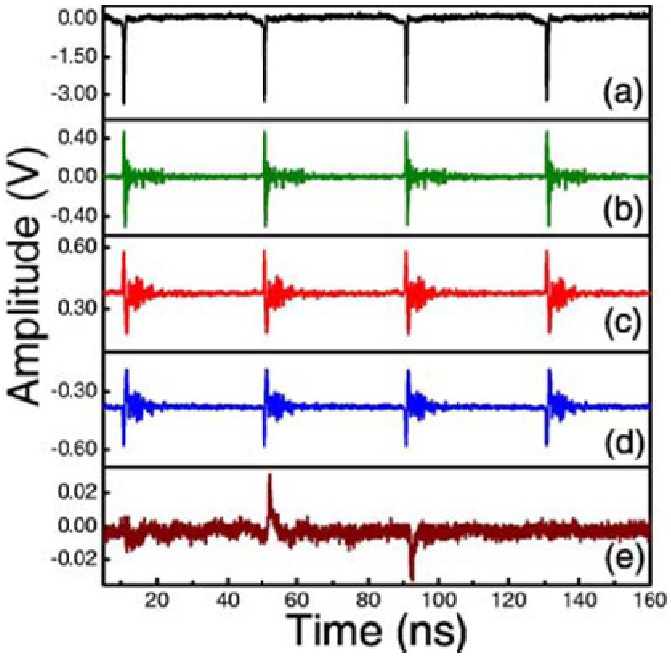



Figure 5.9: (a) Oscilloscope traces of the gating pulse, (b) APD response, (c) PD 1 response, (d) PD 2 response, and (d) the self balanced avalanche signal. The average photon number was 1 photon/pulse from the attenuated laser. [58]

## 5.6 Conclusion

In this chapter, I introduced gating schemes that have recently been developed for high-speed applications. These schemes have successfully addressed the speed limitation due to the afterpulsing effect by reducing the total avalanche charge flow and the background noise in SPAD signals. Though they have different circuit configurations, the fundamentals are similar.

# Chapter 6 Sinusoidal Gating Results and Findings

In this chapter, I report a sinusoidal gating technique that has achieved PDE of 55% with DCR of 15.5 kHz at 240 K. Section 6.1 describes the experimental set up. Sections 6.2 and 6.3 report the figures for merit from the experiment results. Sections 6.4 and 6.5 analyze result and discuss the factors that impact afterpulsing.

## 6.1 Experimental Evaluation

Sinusoidal gating is similar to pulsed gating in that the SPAD is periodically biased above breakdown. Its success in achieving high data transmission rates, with low DCR and afterpulsing is attributable to two factors. First, a high frequency sinusoidal signal generates short sub-nanosecond gates that are beneficial for reducing afterpulsing. Secondly, high signal-to-noise detection is facilitated by the fact that the SPAD output consists of two components, a sinusoidal response at the bias frequency, which can be effectively eliminated with one or more narrow band filters, and a short temporal pulse with a broad frequency spectrum that is relatively unperturbed by the filter. A disadvantage, however, is that the time interval between two adjacent peaks in the bias voltage is also sub-nanosecond for gating frequency in the GHz range. This can have a detrimental effect on afterpulsing. For InP-based SPADs, one of the dominate traps has a lifetime ~ 100 ns [17, 45], which, for GHz range frequencies is much longer than the time interval between two peaks in the sine wave bias. In this work, the afterpulse probability

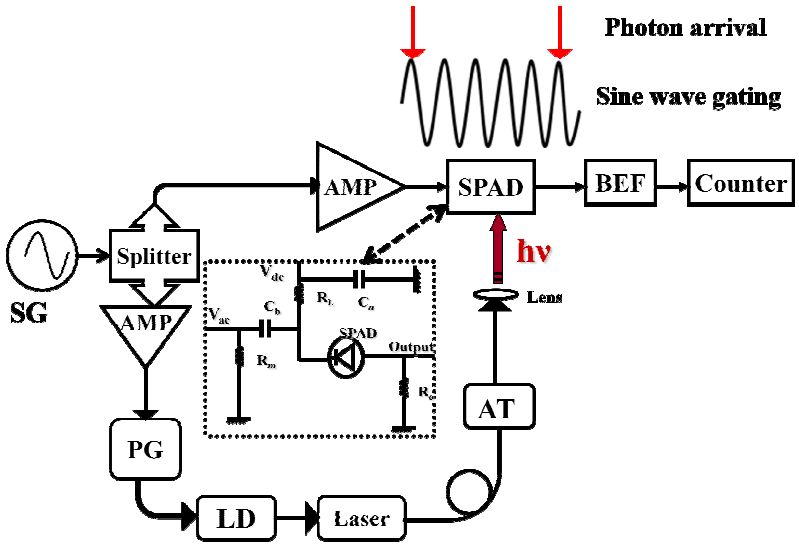


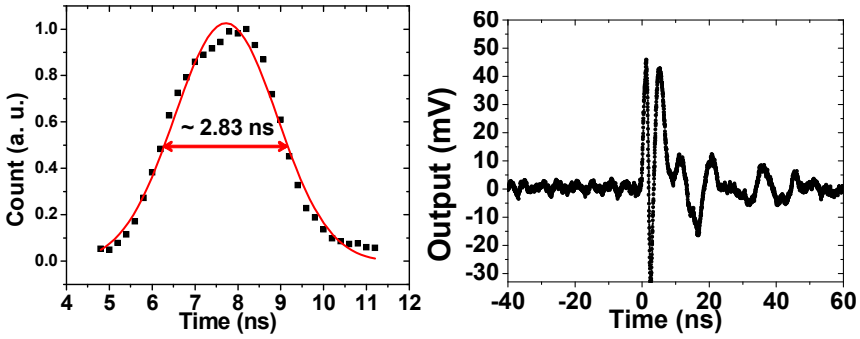

Figure 6.1: Experimental set up for sinusoidal gating. SG: sinusoidal wave generator, AMP: high power/speed amplifier, PG: pulse generator, LD: laser driver BEF: band elimination filter and AT: optical attenuator.

is studied at a lower bias repetition rate of 80 MHz for different laser repetition rates and excess bias levels.

A block diagram of the measurement set up as illustrated in Figure 6.1. The inset (dotted line) shows the circuit board that was used to mount the SPAD and passive components. The master sinusoidal signal was split equally into two parts. One component was amplified to 2.8 V and used as ac bias for SPAD. The other ac signal was divided by a pulse/pattern generator and used to trigger the laser driver. The excess bias was tuned by changing the dc bias on the SPAD.

Figure 6.2 (a) is the temporal distribution of the normalized counts, from which the effective pulse width of the 80 MHz gating pulse can be extracted as the full width of half maximum. Figure 6.2 (b) shows a typical output oscilloscope trace. The amplitude of one avalanche output pulse is approximately one fifth that of an avalanche pulse obtained in gated mode operation. The equivalent charge flow is 0.8 pC, which is much smaller than that in gated mode. The excess bias was selected so that PDE is between 10% - 20%.

The laser repetition rate determines the data transmission rate or the sampling frequency for imaging applications. To realize a 10 MHz laser repetition rate, any gating frequency that can be synchronized with the laser pulses is technically feasible. A primary difference is the effective pulse width, which affects the PDE and afterpulse rate.




(a) (b)

Figure 6.2: (a) effective pulse width of 80 MHz gating frequency; (b) oscilloscope output signal.

Higher gating frequency results in narrower pulse width, which makes it more difficult to synchronize the arrival of the photons with the maximum excess bias point. Although narrower pulse widths also reduce the time that the SPAD is biased above the breakdown voltage and, therefore, decrease the total charge flow, which leads to lower afterpulsing rates. To reduce total charge flow during an avalanche event has been the most successful approach to reduce afterpulsing, which may account for GHz sinusoidal gating of

recently reported QKD systems [42, 51]. However, the impact of the effective pulse width (Figure 6.2 (b)) together with the excess bias amplitude has not been reported. Here I compare the afterpulse probability in sinusoidal gating with that obtained using the PQAR approach, another method to reduce the total avalanche charge. Another interesting phenomenon is that among all the published results on sinusoidal gating, the laser repetition rate is always a fixed value, typically 10 MHz [49-51, 53]. In this work, I used a different method to down convert the frequency of the sinusoidal bias signal to a range of different laser repetition frequencies. The pulse generator divides the trigger signal by an arbitrary integer between 1 and 1000. The pulsed laser was then attenuated to single photon level with a pulse width of 60 ps. The temporal position of the laser pulse was aligned with the sinusoidal peaks by tuning the pulse generator delay. Two notch filters, one at 80 MHz, the fundamental gate frequency, and the other at 160 MHz before connection to an oscilloscope or a multichannel analyzer filtered the output signal. Through changing the laser repetition rate, the PDE can be maximized and a high PDE of 55% was achieved.

Various methods have been utilized to study the afterpulse probability [49-51, 53]. Double pulse measurements have been demonstrated to be an effective and accurate method to measure the afterpulse probability [15, 45]. This method applies two adjacent bias pulses whose time interval can be tuned continuously. The photon is temporally aligned within the first pulse and 100% avalanche is achieved by increasing the intensity of the incident light. The delay of the second pulse is defined as the hold-off time. It is used to measure the dark counts triggered by released dark carriers that were trapped during previous avalanche events, primarily during the previous adjacent avalanche event.

However, it is not straightforward to adapt this technique to sinusoidal gating. For sinusoidal gating, the gating frequency is fixed due to the narrow stop-band notch filter. A different approach was adopted here. To measure the excess dark avalanche events triggered by the released trapped carriers, I set the counter to gated mode, where the counter only counts for a certain period of time, for example, 50ns. In this case, a second temporal alignment is required to obtain the normal count rate, which is to align the counting window together with the incident photon and the appropriate sinusoidal gates. Then the counting window is delayed in time, which aligns the avalanche events that are triggered by trapped carriers instead of incident photons. Thus I obtain a ratio between the afterpulsing avalanche events and the photon triggered avalanche events, which is defined as the afterpulse probability. Note that this refers to the count rate during the delayed counting window after the dark count rate has been subtracted, which is obtained with the same timing condition without light illumination. The result will be reported in Section 6.2.

## 6.2 Dark Count Rate and Photon Detection Efficiency

Different photon flux (photons per pulse) at different laser repetition rates have been studied with the set up in Figure 6.1. Figure 6.3 shows that the count rate is approximately linearly proportional to the photon number from 0.01 to ~ 10 photons per pulse. For repetition rate of 1.6 MHz, the detection range spans 0.002 to 10 photons per pulse, which represents a dynamic range of 37 dB. The count rate is also linear with respect to laser repetition rate. In the following the photon flux is 0.1 photons per pulse if not explicitly stated otherwise.

Two types of SPADs were studied in this work, one is a packaged 40 μm-

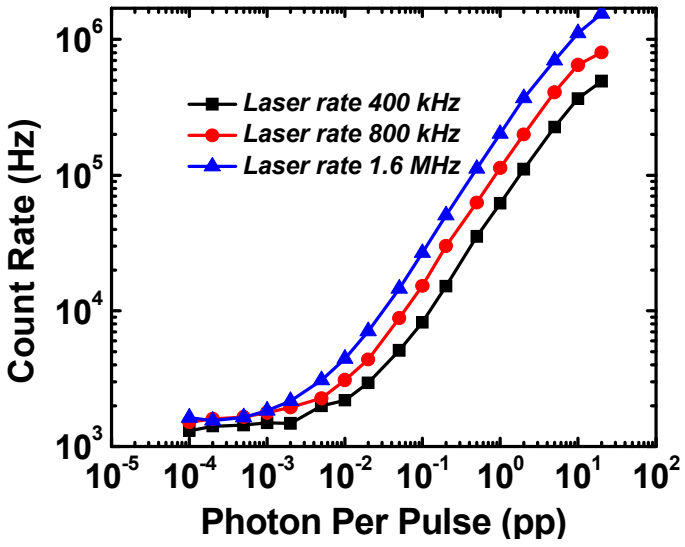



Figure 6.3: Count Rate versus Photon influx at different laser repetition rate.

diameter 2.5 Gb/s telecommunications APD from JDS Uniphase (JDSU) and the other is a 25 μm-diameter APD chip from Princeton Lightwave (PLI). The dark count rate versus

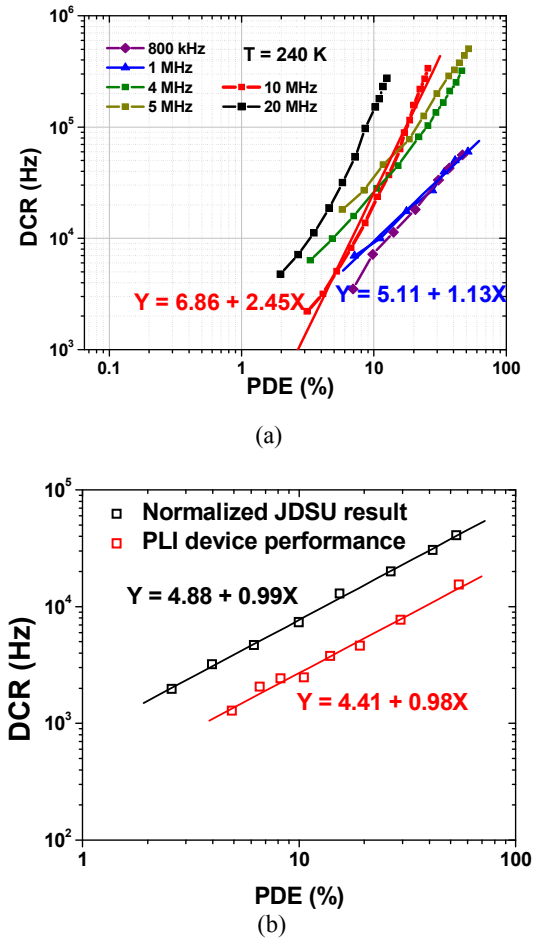
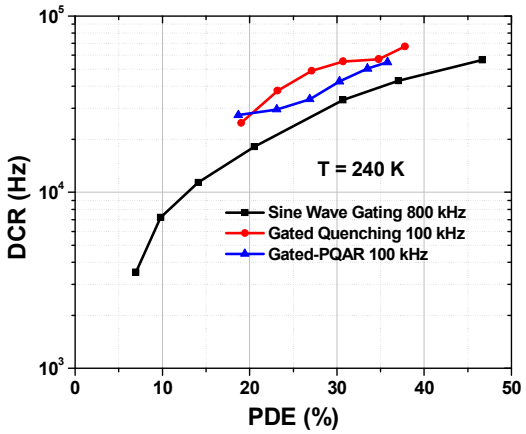


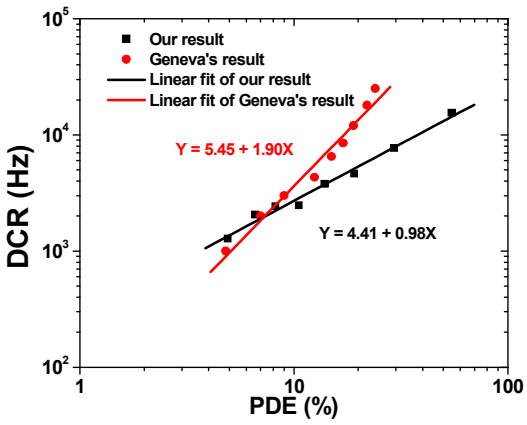

Figure 6.4: (a) DCR versus PDE of JDSU and PLI SPADs at 240 K and 400 kHz laser repetition rate. DCR of the JDSU SPAD was normalized to the same area as the PLI SPAD. (b) DCR versus PDE for JDSU SPAD at different laser repetition frequencies.

photon detection efficiency at 240 K for both devices is plotted in Figure 6.4(a). To facilitate comparisons, the JDSU data was normalized to 25-µm diameter, assuming that the dark count rate scales with area. Figure 6.4 (b) illustrates the effect of changing the laser repetition rate. The photon detection efficiency and dark count rate were determined using the equations 2.3 and 2.4 in Chapter 2. The slope of DCR versus PDE (log-log scale) increases as the laser repetition rate increases up to 10 MHz. It follows that for currently available SPADS higher detection efficiencies are achieved at lower data

transmission rates. Possible explanations are (1) higher laser repetition rate results in higher afterpulse probability and/or (2) incomplete recharge after frequent avalanche




(a)




(b)

Figure 6.5: (a) Comparison of sine wave gating with the PQAR technique [15, 23]. (b) Comparison of sine wave gating in Ref. [53] with present work.

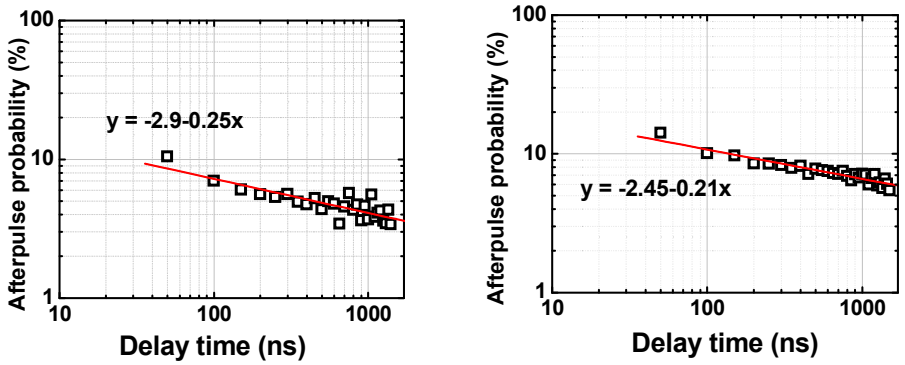
events.

Figure 6.5 compares the sinusoidal gating technique reported here to other receiver configurations. Figure 6.5(a) shows the JDSU SPAD operated in gated quenching, gated PQAR, and sinusoidal gating mode. The sinusoidal gating mode yields lower DCR at the same PDE with higher laser repetition rate [23]. Figure 6.5(b)

compares the PLI results with the sine wave gating receiver in Ref. [53]. In Ref. [53], the authors have reported a parameter called dark count rate per nanosecond, which is obtained by dividing the dark count rate by the gating frequency and the effective pulse width. Assuming the Poisson distribution can be neglected, this yields the dark count rate per nanosecond. I have adjusted the dark count rate per ns to DCR for this comparison. One may note that the gating frequency and laser repetition rate in [53] are 921 MHz and 77 MHz, respectively, whereas in the present work the gating frequency and laser repetition rate are 80 MHz and 400 kHz, respectively. The difference in gating frequency and, therefore, the difference in effective pulse width impact the avalanche charge flow, which affects the afterpulse probability.

## 6.3 Afterpulse Probability

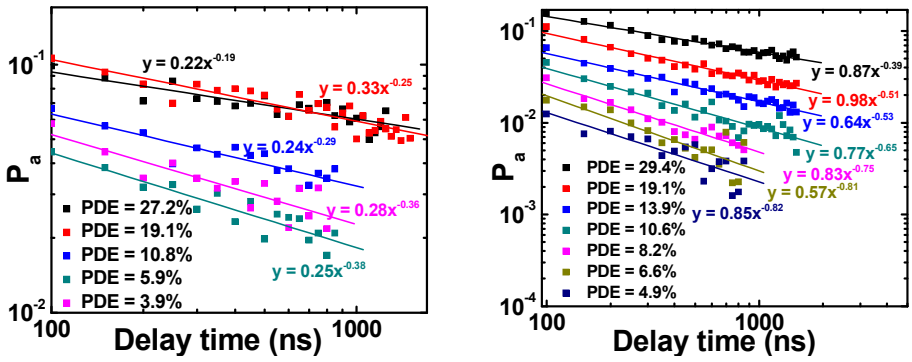
Different methods have been proposed to evaluate afterpulse probability for sinusoidal gating [49, 50, 53]. Essentially they fall into two categories, autocorrelation and the coincidence window method with the latter being used more often. In this work I have used the coincidence window technique. The counter (Stanford Research 400) was set to gated mode, where the main counter is only activated by a trigger signal for a short period of time, e.g. 50 ns, and then kept off until the next trigger. In this mode, a count rate similar to a time histogram can be obtained. The counting process was performed with and without laser illumination, obtaining counts $C_i$ and $C_d$, respectively. For all the windows that are not coincident with the photon arrival, $C_i$-$C_d$ accounts for excess dark counts that are caused by the previous photon induced avalanche events, i.e., afterpulses.




(a) (b)

Figure 6.6: Afterpulse probability for PDE of (a) 10% and (b) 40%.

The afterpulse probability for each window except the one with photon incidence was then obtained by dividing $C_i$-$C_d$ by the number of photon counts obtained within the coincident window. The afterpulse probability, $P_a$, versus delay time, T, is shown in Figure 6.6(a) and (b) for PDE of 10% and 4%, respectively. The data in Figure 6.6 can be




(a) (b)

Figure 6.7: Fits for the detrapping parameter α for (a) JDSU and (b) PLI SPAD.

fit with a power law of the form $P_a \propto T^{-\alpha}$[18]. The index α indicates the severity of afterpulsing. Larger values of α, correspond to faster detrapping and lower afterpulsing for the same delay time. The plots in Figure 6.7(a) and (b) shows power law fits for the JDSU and PLI SPADs, respectively. It can be seen that the PLI device yields larger α for

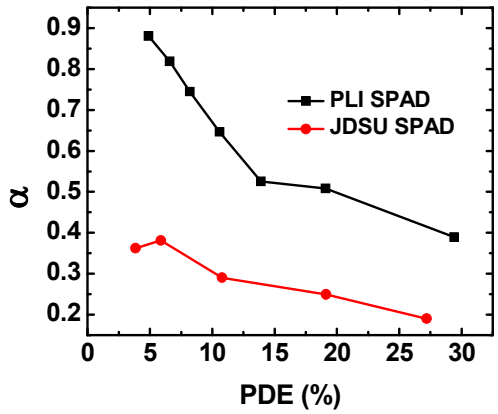



Figure 6.8: Comparison of detrapping parameter α with for the JDSU and PLI SPADs.

the same PDE compared with the JDSU device. The detrapping parameter increases with PDE for the same delay time as illustrated in Figure 6.8.

## 6.4 Analysis of PDE and Afterpulse Probability

The highest PDE of 55% was achieved with laser repetition rate of 400 kHz. A remaining challenge for sinusoidal gating is to achieve high PDE at high transmission

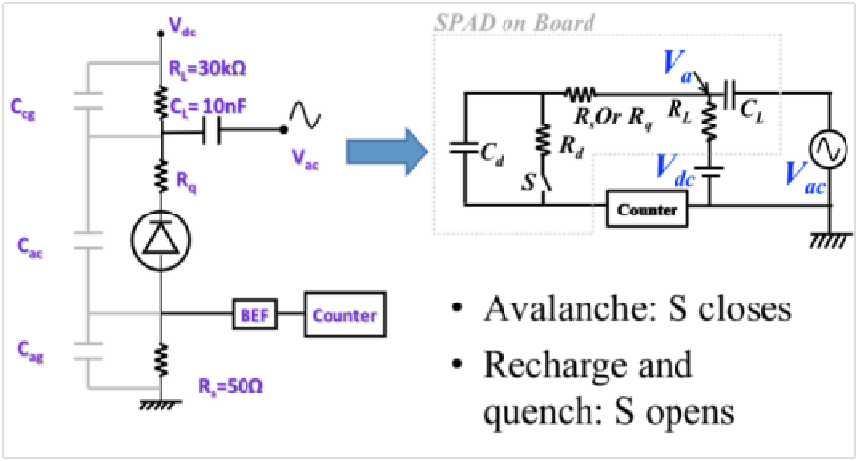



Figure 6.9: Equivalent circuit for SPAD operated in sinusoidal gating mode.

rates. I present an equivalent circuit to study this issue [59]. The model for the SPAD includes one resistor and one capacitor in parallel with a series resistor. The switch S controls the state of the SPAD. When the SPAD is armed for detection, S is open. It closes during an avalanche state and is open again for recharging.

The receiver circuit and equivalent circuit are shown in Figure 6.9. In the initial state, the SPAD is armed; the bias is the sum of the DC bias and the sinusoidal voltage swing. During each period of the sinusoidal wave, the leakage current across the load resistor $R_L$ will slowly discharge the node $V_a$. As long as $R_LC_L$ is much larger than the period of the sine wave, the potential at $V_a$ will not drop significantly so that the excess bias on the SPAD is maintained. When an avalanche event occurs, S closes. The avalanche current, $I_1$ discharges the diode capacitor and $I_2$ discharges the load capacitor until the avalanche is quenched and S opens. In the recharge state, two currents through two time constants, $R_sC_d$ and $R_LC_L$, recharge the potential $V_a$. This also restores charges on the load capacitor and the diode capacitor. The phenomenon I noticed is that when the total count rate reaches the level of $10^5$, the PDE drops significantly. The counting rate corresponds to a separation time interval between two adjacent avalanche events of 10 μs, while the RC time constant to recharge $C_L$ is 300 μs. Once the charge on the load capacitor is constantly below the designed charge level, the potential of $V_a$ drops accordingly, which results in the decrease in SPAD excess bias and lower PDE. It follows that high detection efficiency is difficult to achieve for high laser repetition rate. In addition the excess bias required to achieve the same PDE with high gating frequency is usually higher than that in the present work [49, 50, 53, 60]. The reason for a high excess bias is that with high gating frequencies, the effective pulse width is so narrow that the

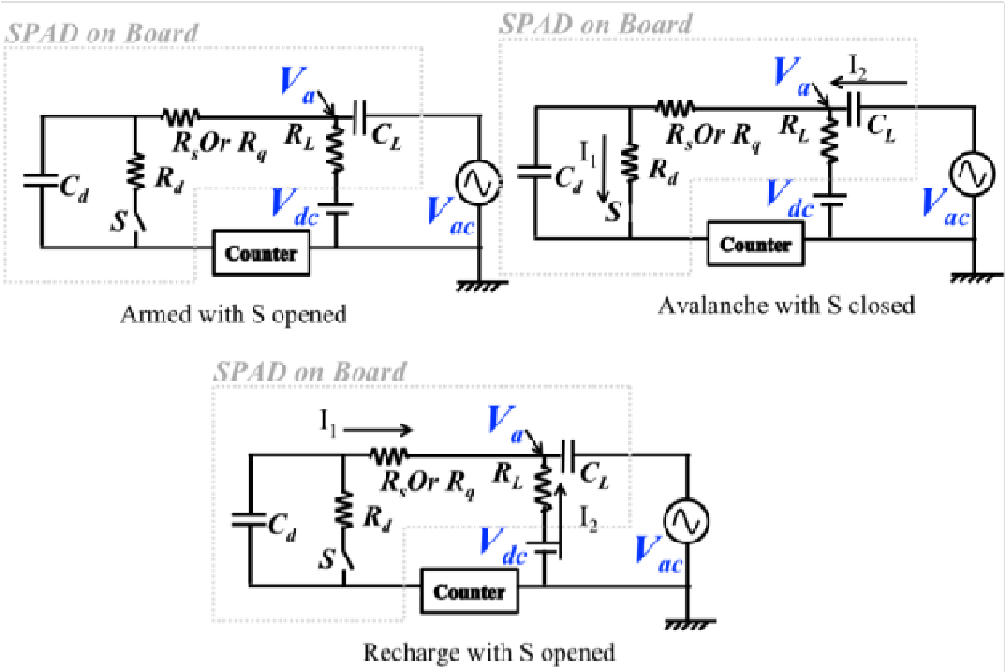



Figure 6.10: Three states illustrated with equivalent circuit.

large excess bias is required to build up an avalanche event quickly. High afterpulsing also adversely affects PDE. With higher afterpulsing rate, the total count rate includes more false positives contributed by afterpulsing. The resulting increase in DCR limits the PDE. Figure 6.11 shows the afterpulse probability versus the hold off time for gated quenching, gated PQAR, and sine wave gating. With similar PDE, the sine wave gated SPAD yields a factor of three lower afterpulse probability. This may be due to reduced charge flow for sine wave gating, which is linked to the integrated applied excess bias. Comparing gated PQAR and sine wave gating, the temporally integrated voltage drops from 24 ns•V to roughly 9.6 ns•V, giving a reduction of roughly 3.

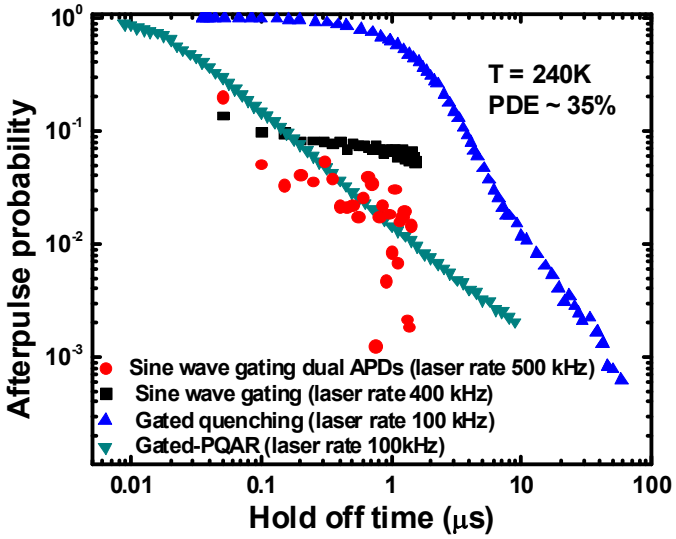



Figure 6.11: Afterpulse probability with different gating scheme and laser repetition rate and similar PDE.

## 6.5 Impacting Factors for Afterpulsing in Sine Wave Gating

It has been reported that the afterpulsing decreases faster than linearly with decreasing applied pulse width [18] where we define the applied pulse width as the time period when the diode is biased above breakdown voltage. An advantage of sinusoidal gating at high frequencies is that it is straightforward to achieve very narrow excess bias pulses, which restricts charge flow and reduces afterpulsing. On the other hand, afterpulsing decreases with lower excess biases. Consequently, a useful parameter is the integrated excess bias over the applied pulse width, which will be designated as $\Delta$. For gated quenching, the applied pulse width is a constant that does not depend on excess bias [18], which is not the case for sinusoidal gating, as illustrated in Figure 6.12. For sine wave gating, the excess bias can be increased by increasing the DC bias with fixed AC voltage swing or increasing the AC voltage swing for fixed DC bias. Figure 6.12

shows the measured afterpulse probability and the applied pulse width versus various peak excess biases.

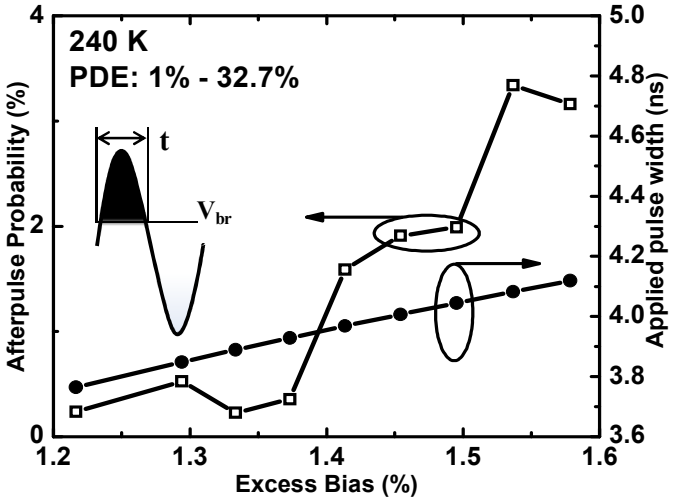



Figure 6.12: Afterpulse probabilities at 1 μs hold off time versus peak excess bias, right axis shows the applied pulse width.

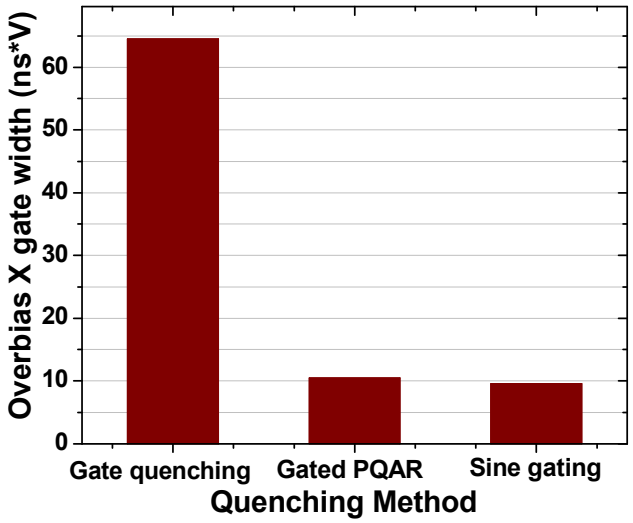



Figure 6.13: Comparison of different gating schemes in respect with the integrated gate width with over bias for data in Figure 6.11.

Figure 6.13 compares computed values of Δ for the different gating schemes in Figure 6.11. For the four gating schemes, Δ (y axis) for gated quenching is the highest, which explains the highest afterpulse probability in Figure 6.11. The comparable value of Δ for the three other gating methods is consistent with the similar afterpulse probabilities shown by the three overlapping trends in Figure 6.11. This indicates that Δ is closely linked to the afterpulsing probability.

Δ is also an indicator for photon detection efficiency as shown in Figure 6.14. Both photon detection efficiency and afterpulse probability increase with Δ; the increase becomes more abrupt as Δ exceeds 9 $ns \bullet V$. Unfortunately it is not trivial to suppress afterpulsing and simultaneously maintain high PDE by adjusting Δ. For sine wave gating Figure 6.12 shows that the applied pulse width increases with increasing peak excess bias. The higher the sine wave gating rate the narrower are the applied excess bias pulses,

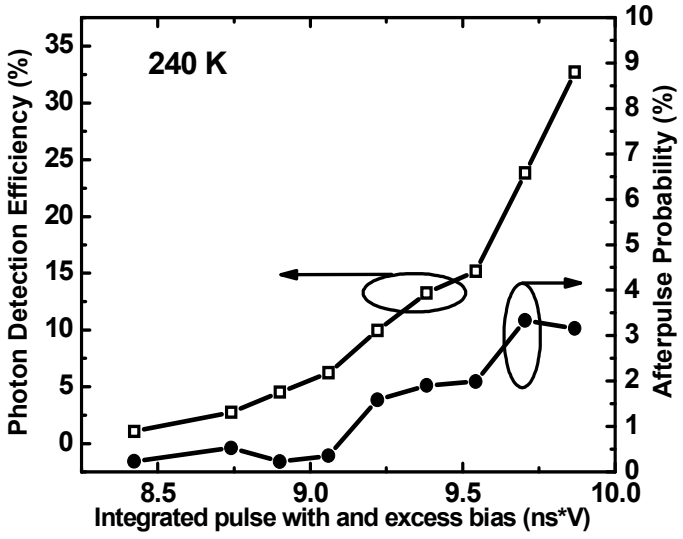



Figure 6.14: Indicator for afterpulse probability – Integrated pulse width with excess bias. Afterpulse probability is the value at hold off time of 1 μs.

which would lead to reduced afterpulsing. However, this is somewhat over simplified because higher gating rates also mean there are more excess bias gates within a given time period. The total time that the SPAD is biased above breakdown must be considered. To elucidate this frequency dependence, Δ has been calculated with the total excess duration in a certain time period (the shaded area shown in Figure 6.15) for different gating frequencies using the following relations

$$\Delta_1 = \int_{\frac{T}{2\pi}\arcsin\alpha}^{\frac{T}{2\pi}(\pi-\arcsin\alpha)} A\sin(\omega t)dt = \frac{2A}{\omega}\sqrt{1-\alpha^2} \tag{6.1}$$

$$\Delta_2 = Area \times f = \frac{A}{\pi}\sqrt{1-\alpha^2} \tag{6.2}$$

Equation (6.1) yields the integrated area for one period ($\Delta_1$) and equation (6.2) is the total value in one second ($\Delta_2$). Equation (6.2) shows that the value of $\Delta_2$ is actually independent of the gating rate. In other words, for the same peak excess bias, the total time that the device is biased above breakdown is independent of the gating frequency. If

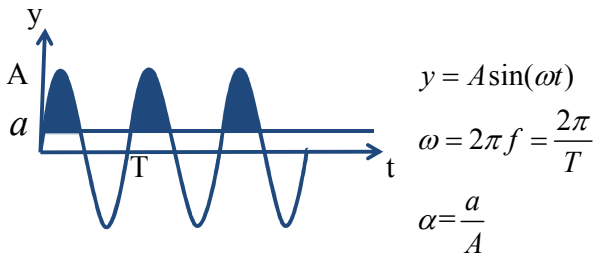



Figure 6.15: Sketch and equations illustrating calculation for over all time when the device is biased beyond breakdown. T: period of sinusoidal wave, A: peak value of the sine wave, a: breakdown voltage, ω: sine wave frequency.

we only consider this averaging effect, lower gating frequency yields higher photon detection efficiency and comparable afterpulse probability. Equation (6.2) also shows that the primary factor that affects the value of Δ is the difference between peak bias and breakdown voltage, which is the excess bias.

Recent re-interpretation of afterpulse behavior has provided insight on the detrapping process in InGaAs/InP SPADs, where the detrapping speed can be studied by fitting the afterpulse probability curve versus hold off time with a power law $P_a \propto T^{-\alpha}$; $P_a$ is the afterpulse probability and $T$ is the hold off time [2, 18, 61]. The larger α, the faster detrapping occurs. The reported α for conventional gated-mode operation with various gate widths is ~ 1 [2, 18], while α obtained from sinusoidal gating for different devices and PDE are all < 1 [61]. The indicated difference in detrapping speed due to different gating schemes is consistent with the data in Figure 6.11.

## 6.6 Conclusion

In this chapter, InGaAs/InP SPADs were operated with 80 MHz sinusoidal gating at wavelength of 1310 nm. Photon detection efficiency of 55% has been achieved with dark count rate of 15.5 kHz at 240 K. The afterpulse probability is also compared with other gating schemes. A factor of three lower afterpulse probabilities has been achieved relative to gated PQAR at a repetition rate of 100 kHz with the same PDE. The lower afterpulsing is explained in terms of the integrated excess bias over time. It is found that the afterpulse probability, Pa, is related to the hold off time through a power law, $P_a \propto T^{-\alpha}$. The parameter, α, is a measure of the detrapping time.

# Chapter 7 High Rate Gating with Common-Mode Cancellation

In this chapter, a novel transient cancellation technique is presented. A laser repetition rate as high as 20 MHz, which is limited by the speed of the laser system, has been achieved. This transient cancellation utilizes a balanced detector configuration and its common mode cancellation mechanism. A balanced detector pair is usually used to detect small signal differences between the two channels. The two photodiodes are both reverse biased. They compete to affect the voltage at the common node ($V_{out}$ in Figure 7.1) by producing a small difference in the current flow. The configuration of balanced detectors enables the cancellation of common mode signals between the two detectors. Compared with a conventional single detector, it has been shown that balanced detectors can improve system sensitivity by 15-20 dB [62]. For the dynamic excess biasing used for single photon detection, opposite AC voltage swings are used for the two diodes. We have utilized this detector configuration to detect the small avalanche current generated by single photons in one diode of the balanced SPAD pair. This approach is similar to differential detection [63] except that it does not require additional components in the output circuitry [57].

In Section 7.1, the principle of common mode cancellation will be introduced. Section 7.2 reports dark count probability and photon detection efficiency. Afterpulsing probability is presented in Section 7.3 followed by the conclusion in Section 7.4.

## 7.1 Principle of Common Mode Cancellation and Experimental Set Up

For gated-mode quenching it is important to distinguish the avalanche pulse from the transient spikes. As discussed in Section 5.3, one technique that has been used very successfully is to use two matched delay lines to cancel the transient spikes [19]. This

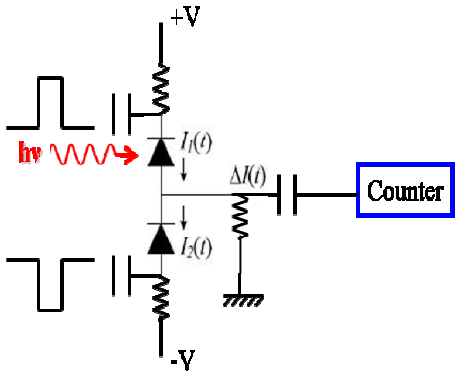



Figure 7.1: Circuit layout and bias configuration for balanced single photon receiver

approach is appropriate for short bias pulses that yield reduced afterpulsing. Another successful technique is self differencing (Section 5.1), which has data transmission up to 315 MHz [64]. This type of common-mode cancellation can also be achieved with balanced detectors [65]. With balanced detection the transient pulses are common to the diode pair and can be suppressed. Balanced receivers[66, 67] can be implemented in modules or at the chip level, which presents the potential for device-level integration.

The gated-mode balanced receiver, shown in Figure 7.1, is the hybrid circuit that I developed for this work. Both diodes are biased with excess bias pulses superimposed on the DC voltages. The amplitude, pulse width, and time delay of the AC voltage pulses are adjusted to achieve the best noise cancellation, which yields minimum dark counts and maximum photon counts. The incident light wavelength was 1310 nm and, if not

explicitly stated otherwise, the optical signal was attenuated to 0.1 photons per pulse. The diode pair was mounted on a copper sub-mount, which has a high thermal conductivity. Measurements were carried out at 240K and 280K. The counter was a PicoHarp 300 multi-channel analyzer that has a maximum synchronization rate of 80 MHz.

The oscilloscope traces in Figure 7.2 illustrate transient spike cancellation with pulse width of 4 ns. Figure 7.2(a) shows the pulse responses of the individual SPADs in the dark. Figure 7.2(b) is the output in the absence of incident photons with both SPADs pulse biased; the residual signal amplitude is < 5 mV. Figure 7.2(c) is the avalanche signal with amplitude of 9 mV. The counter threshold was set to 5 mV in this case. With 1.4 ns pulse width, the calculated charge flow is approximately 0.09 pC, the lowest value achieved in this dissertation.

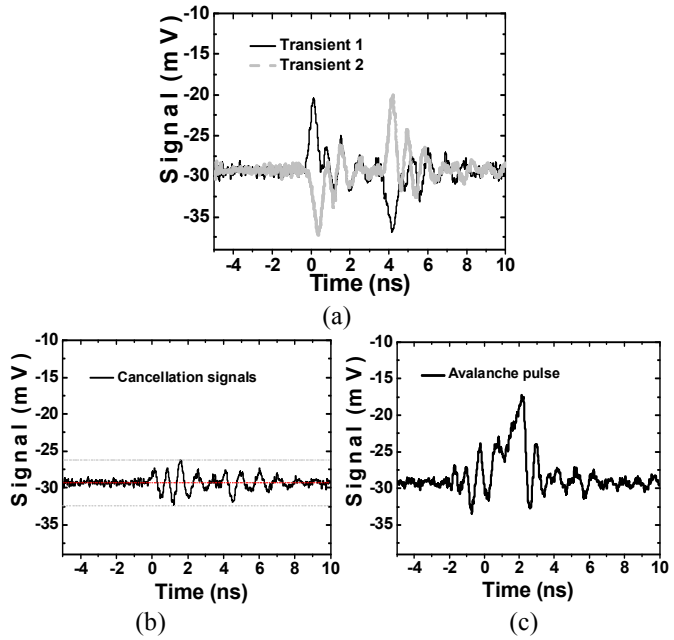



Figure 7.2: Transients (a), cancellation (b) and avalanche signals (c) on Oscilloscope

## 7.2 Detection Efficiency and Dark Count Probability

Balanced detection enables operation with short bias pulses; the shortest pulse available for this work was 1.4ns. Figure 7.3 shows dark count probability (DCP) versus photon detection efficiency (PDE) at 240 K and 280 K for different laser repetition rates. The laser repetition rate is the same as the gating rate; therefore the clock rate equals the maximum count rate.

The pulse width for all 240 K and 280 K measurements was 1.4 ns except the 1 MHz result at 240 K was achieved with a 2.5 ns pulse. At 240 K and 20 MHz repetition rate, the dark count probability is $1.9 \times 10^{-5}$ and detection efficiency is 13%. At 280 K and 20 MHz laser repetition rate, the dark count probability is $1.8 \times 10^{-4}$ with a detection efficiency of 12%. For the same gate width there is no significant increase in DCP for laser rates in the range 1 MHz to 20 MHz for both temperatures. This indicates significant suppression of afterpulsing due to the small avalanche pulse detection. On the other hand, sinusoidal gating with balanced detection exhibited an order of magnitude

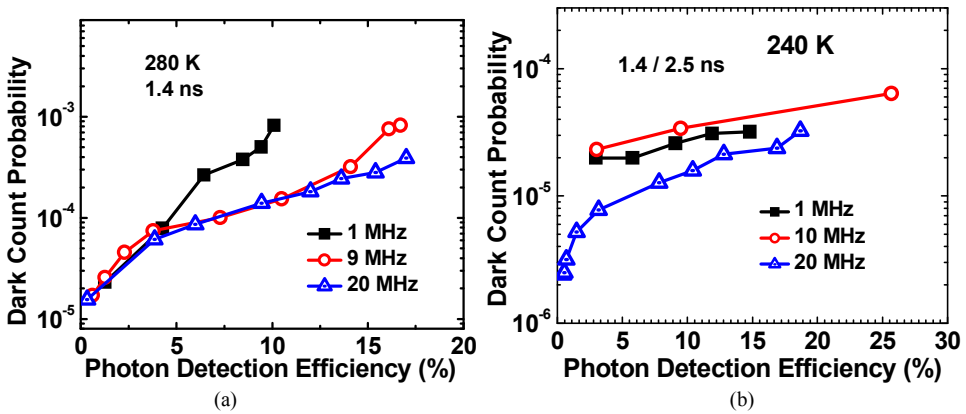
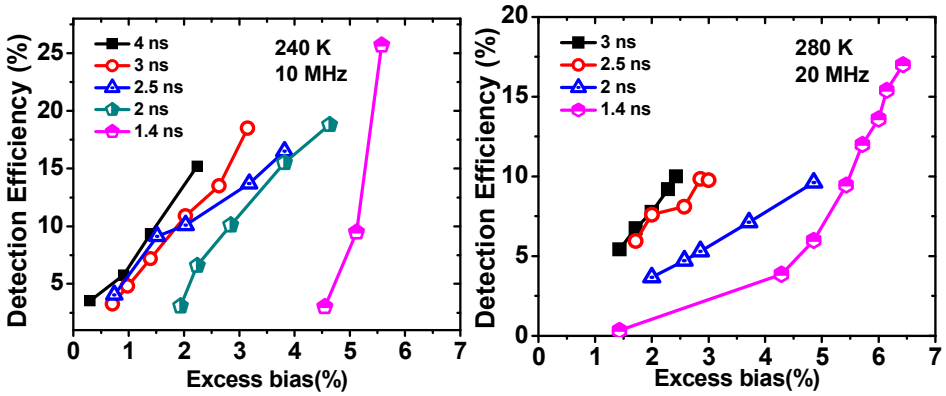



Figure 7.3: Dark count probabilities and photon detection efficiency at 280 K and 240 K with various laser repetition rates

increase in afterpulsing as the repetition rate increased from 1 MHz to 20 MHz [66].

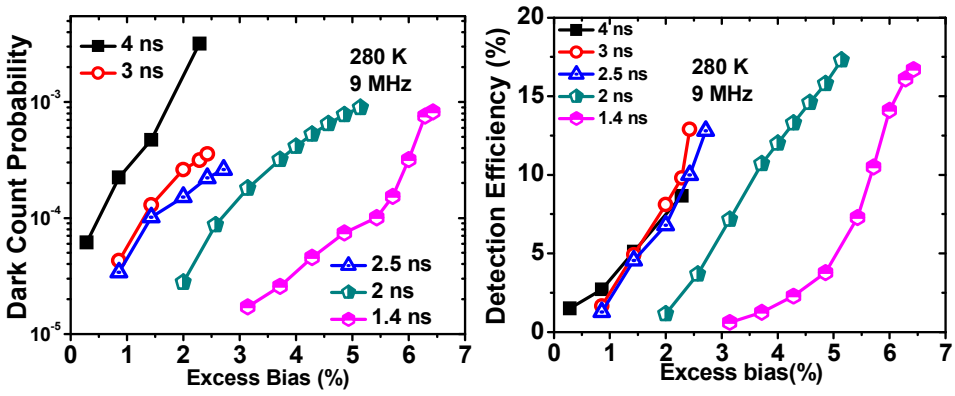
Figure 7.4 provides plots of detection efficiency versus excess bias for different applied pulse widths. For 240K and 280K the best DCP versus PDE results were achieved with the narrowest pulse width, 1.4 ns. The 1.4 ns pulse exhibits the most abrupt rise with increasing excess bias.

Both the applied pulse width and excess bias amplitude have a significant impact on DCP and PDE. For a given excess bias (%), wider pulse width yields a higher dark count probability (Figure 7.5(a)). This is partly due to the longer time period for dark count accumulation. The difference between the applied pulse width and effective pulse width is also important for detection efficiency. Owing to the avalanche build up time and the response time of the electronic circuit there is a delay between the arrival of the photon and the peak of the avalanche pulse[45, 61]. If the applied pulse width is comparable to or smaller than the delay time, the detection efficiency is adversely affected, as shown in the case with 2 and 1.4 ns in Figure 7.5 (b). From these results we estimate that for excess bias of 1-2 % the delay time is approximately 2 ns at 280 K. The data for 20 MHz laser repetition rate (Figure 7.4(b)) also indicates a delay time of 2 ns. Since the avalanche buildup time decreases with excess bias, this also explains the rapid rise of detection efficiency with excess bias for the 1.4 ns pulse. On the other hand, when the applied pulse width is long enough for the avalanche to build up, the same excess bias yields the same detection efficiency for different pulse widths (2.5 - 4 ns in Figure 7.5 (b)).




(a) (b)

Figure 7.4: Detection efficiencies at 240 K and 280 K with 10 and 20 MHz laser repetition rate.




(a) (b)

Figure 7.5: Dark count probability and detection efficiency versus excess biases at 280 K and 9 MHz

## 7.3 Afterpulsing Characterization

Afterpulse probability was characterized with the double pulse method [14]. For this measurement the incident photons during the first pulse trigger an avalanche event. Deep-level traps in the multiplication layer capture some of the carriers. A second pulse measures counts in excess of the dark counts, i.e. afterpulses, which result from the emission of trapped carriers. The temporal separation between the two pulses is the “hold

off time", which is varied during the measurement. The afterpulse probability is the ratio of the counts in the second pulse after subtracting the dark counts that were measured in the absence of light to the photon counts in the first pulse. In this manner the afterpulse probability can be quantified as a function of hold off time.

With balanced detection the modification made for the double pulse measurement is to utilize normal output and complimentary output pulses for both the first and second pulses. Since the detection of the induced afterpulses also requires common mode cancellation the second pulse is also carefully synchronized for the two diodes. Figure 7.6 shows the histogram counts from both pulses at 240 K and 1MHz laser rate with 10% PDE. The counts in the first pulse and the second pulse have pulse separations from 10 ns to 40 ns.

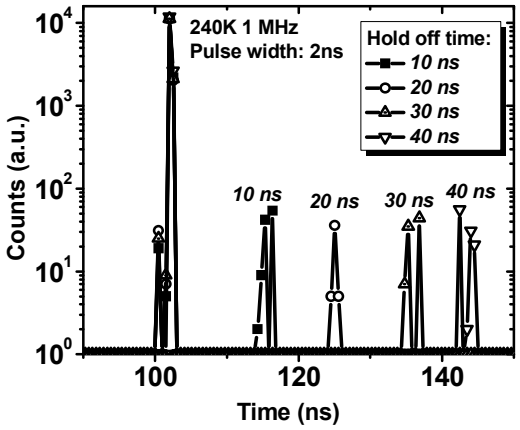



Figure 7.6: Temporal histogram of the photon count rate for double pulse mode at 240 K and 1 MHz laser rate.

The afterpulse probability for different bias pulse widths is plotted in Fig. 7.7. For the 2 ns pulse width, the afterpulse probability is approximately 0.3% with a 20 ns hold off time. This indicates that the afterpulse probability could be < 0.5% for laser repetition rate of 50 MHz. The afterpulse probability versus hold off time does not fit a power law as reported in Ref. [2, 15, 61]. The reason is due to the dark count cancellation effect[67]. When dark counts occur simultaneously in the two SPADs they can cancel one another yielding a reduced dark count rate. While this complicates the determination of the afterpulse probability on hold off time, the net result is a decrease in afterpulse counts.

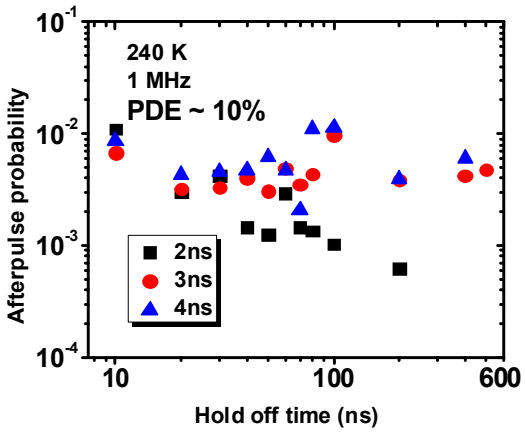



Figure 7.7: Afterpulse probabilities versus delay time at 240 K and 1 MHz laser

## 7.4 Conclusion

In this chapter we demonstrate balanced detection with InGaAs/InP avalanche photodiodes in gated mode for single photon counting at 1310 nm. The advantage of balanced detection is suppression of transient voltage spikes at the leading and trailing edges of the bias pulses, which facilitates detection of short, low-amplitude signal pulses. With a dead time of 50 ns and laser repetition rate of 20 MHz the dark count probability is $1.9 \times 10^{-5}$ and detection efficiency is 13% at 240 K. At 280 K and 20 MHz laser

repetition rate, the dark count probability is $1.8 \times 10^{-4}$ with a detection efficiency of 12%. The afterpulse probability at 240 K and 1 MHz laser repetition rate has been characterized with different hold off times and applied pulse widths. The afterpulse probability for 20 ns hold off time and 10% PDE is 0.3% at 240 K.

# Chapter 8 Balanced Sinusoidal Gating with Phase Shifter

The operation and advantages of sine wave gating have been discussed in Chapter 6. While sinusoidal gating has achieved high PDE and low DCR, one of its limiting factors for some applications is that the operation frequency is fixed owing to the narrow band filters that are required to remove the gating signal from the SPAD output. In this chapter the balanced detector configuration introduced in the last chapter will be adopted with sine wave gating. This approach eliminates the RF filters permitting data transmission at any sub harmonic of the gate frequency. In addition, signal to noise is further improved with added adjustability provided by the phase shifter. With better signal to noise, even smaller avalanche pulses can be detected permitting both total charge flow and dark count to be reduced. In this chapter, Section 8.1 describes the experimental set up and characterization methods. Section 8.2 presents the measured PDE, DCR, and afterpulsing and compares them with conventional the sine wave gating results in Chapter 6. Section 8.3 introduces the improved sine wave system with added phase shifter. Section 8.4 summarizes all the results from sine wave gating with common mode cancellation.

## 8.1 Experimental Set Up

A modified conventional sinusoidal counting system described in [61] has been adopted for this work. Avalanche events were registered with a PicoHarp 300 multi-

channel analyzer, which has a maximum synchronization rate of 80 MHz. Thus all the experiments were performed at or below this gating rate. The counter has an 80 ns dead time, which is the time that the counter requires to reset its electronics after registering a count. To operate the sinusoidal system for balanced APDs, the setup in [61] was modified to provide two opposite phase sinusoidal input signals (Figure 8.1).

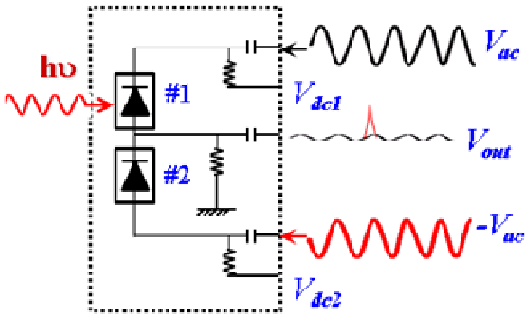



Figure 8.1: Balanced receiver layout.

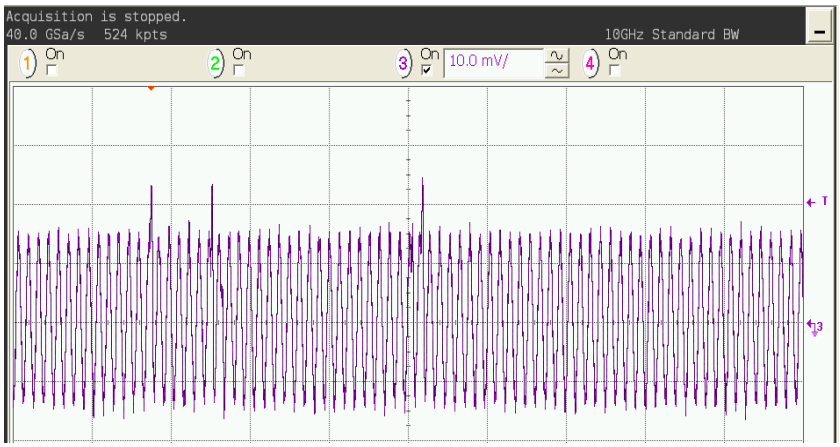



Figure 8.2: Avalanche traces on top of sine wave signals, the residual signal is ~ 15 mV.

Figure 8.2 shows an oscilloscope trace of three avalanche events and the residual background noise. The noise floor is approximately 15 mV, which means the minimum

detectable avalanche signal has to be larger than that. Therefore the total charge flow with this magnitude of avalanche pulse (25 mV) is ~0.5 pC, which is very close with that of the conventional sine wave gating (~0.8 pC).

The measurements and calculations are similar to those described in Chapter 6, except for the balanced configuration I used a new method to characterize afterpulsing. The gating rate was 80 MHz and the laser repetition rate was in the range 400 kHz to 10 MHz (dividing the driving sine wave signal by 200 to 8). The wavelength of incident light was 1310 nm and the light intensity was 0.1 photons per pulse. Figure 8.3 shows the counts histogram from which the photon counts, dark counts and afterpulses, can be extracted. After setting an appropriate threshold and aligning the laser pulse with a peak of the sine wave gate on the illuminated SPAD, a counts histogram was collected using the multi channel analyzer. The threshold was adjusted to be close to but above the noise floor. Figure 4 shows the counts in 14 ns windows; the small peaks are the dark counts, while the largest peak at the 130 ns point in the figure shows the photon counts. Since the repetition rate of laser pulses was 400 kHz, the photon count peaks were separated by 2.5 μs, which explains why only one photon window is shown. In [5, 49, 68], it has been reported that the histogram can be used to calculate afterpulse probability using the relation

$$P_a = \frac{(C_{NI} - C_D)R}{(C_I - C_{NI})} \qquad (8.1)$$

where $C_I$ and $C_{NI}$ represent the average counts per gate for the illuminated and non-illuminated gates, respectively (photon count gates in Figure 8.3). $C_D$ is the dark count per gate with no light incidence for all gates (dark count gates in Figure 8.3). R is the ratio between the gating rate and the laser rate; in the case of Fig. 8.3 the value is 200.

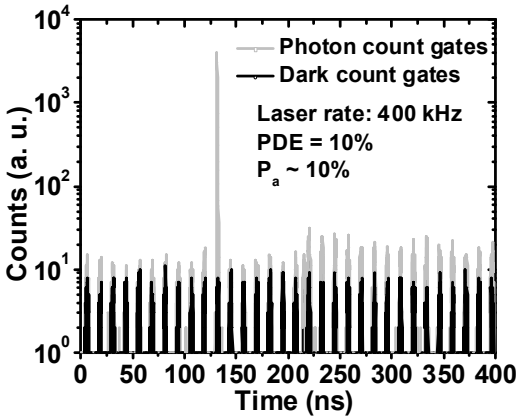



Figure 8.3: Counts histogram collected with integration mode of the multi channel analyzer, the highest peak stands for the gate coinciding with the laser pulse; other small peaks are the dark counts from each gate.

The excess bias was set to achieve 10% PDE; in this case, the photon count rate is $4\times10^3$ $s^{-1}$. The estimated total afterpulse probability is approximately 10%. Note that the dark counts increase due to afterpulsing in the channels after the 210 ns point, which is 80 ns after the photon window at 130 ns. This is due to the 80 ns dead time of the counter.

The effective pulse width, $\tau_e$, was determined by tuning the delay time of the laser pulse and measuring the full width at half maximum of the temporal distribution of the

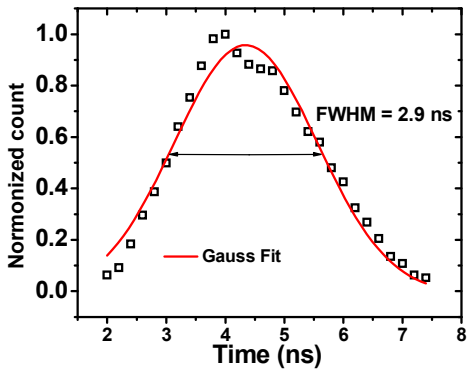



Figure 8.4: Effective pulse width of 70 MHz gating rate with laser pulse width of 40 ps for balanced diodes at 200 K.

photon counts, which is plotted in Fig. 8.4. Assuming a Gaussian distribution the full width at half maximum (FWHM) was 2.9 ns.

## 8.2 Preliminary Result on Balanced Receiver

Figure 8.5(a) shows the DCR versus PDE for a single SPAD and the balanced SPAD pair for different gating frequencies. The balanced receiver exhibits the same DCR and PDE characteristics as the single SPAD receiver, as discussed in Chapter 6. The DCR versus PDE slope increases with the laser repetition rate. Possible explanations are (1) higher laser repetition rate results in higher afterpulse probability and/or (2) incomplete recharge after frequent avalanche events [61]. It was found that for laser repetition rate ≥10 MHz, the PDE is restricted to < 30% due to the fast rising dark count probability [5, 49-51, 53]. The best result with balanced SPADs is 43% PDE and 58 kHz DCR with 1 MHz laser repetition rate. At photon detection efficiency of ~ 10%, the DCR is 9.6 kHz and the equivalent dark count probability is $2.8 \times 10^{-5}$. In Fig. 8.5(b) we also compare the balanced SPADs with gated quenching and gated-PQAR. At 1MHz repetition rate, the DCR of the dual SPADs is slightly lower than that of gated quenching at 100 kHz with the same PDE.

Afterpulsing for the dual SPADs was also characterized using the coincidence window method [51, 53]. For these measurements a Stanford Research SR400 counter was used in gated mode. The temporal position of the counter gate was scanned over several periods of the sine wave gates. The counter gate width was set to be 50 ns. This provided a temporal distribution of counts similar to Figure 8.2 except with more precise

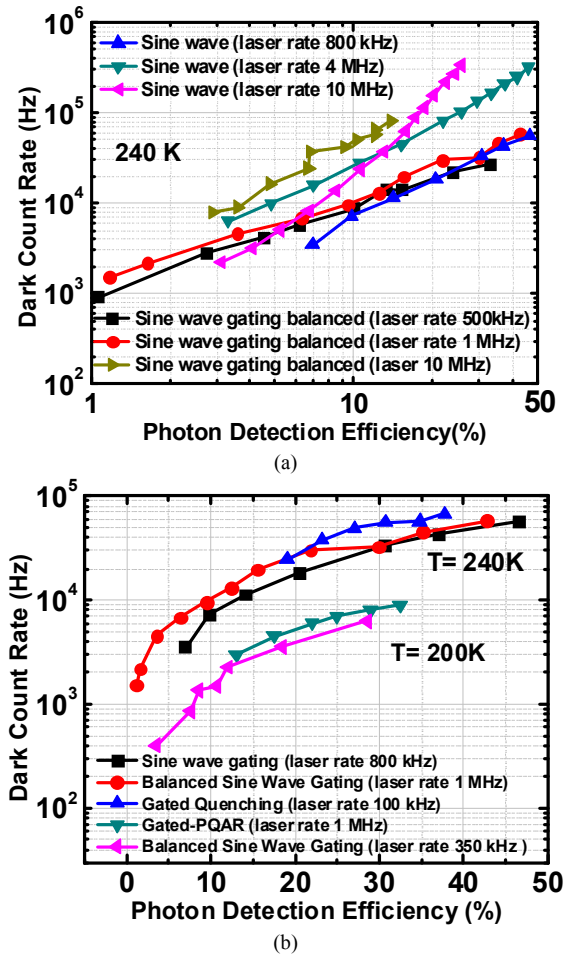



Figure 8.5: (a) DCR Vs PDE results from single and dual SPADs with different laser repetition rate; (b) DCR and PDE comparison with other gating schemes.

count rate information. The afterpulse probability was calculated from the ratio of the afterpulse counts to the photon counts. A relevant phenomenon is the dark count cancellation effect. Figure 8.6 shows that the afterpulse probability of dual SPADs is smaller than that of single SPADs. Also the dual SPAD curve is not as flat as that for sine wave gating of a single SPAD. We attribute this to the cancelling effect between the two SPADs. This is similar to the transient cancellation shown in Figure 8.2. If the two SPADs register dark counts at the same time, the resulting avalanche pulses will cancel,

which results in a decreased number of dark counts. This is more significant for high dark count rates, especially when afterpulsing is high. Therefore the afterpulse probabilities within 1 μs hold off time exhibit a random pattern as a result of the cancellation effect. The afterpulse probability of the dual SPADs for 1 μs hold off time is only 1% with PDE of 35%; owing to the scatter in the data, the afterpulse probabilities at 0.95, 1 and 1.05 μs hold off times were averaged. Overall, the afterpulse probability for sine wave gating of the dual SPADs at 500 kHz is comparable to that for gated PQAR at 100 kHz.

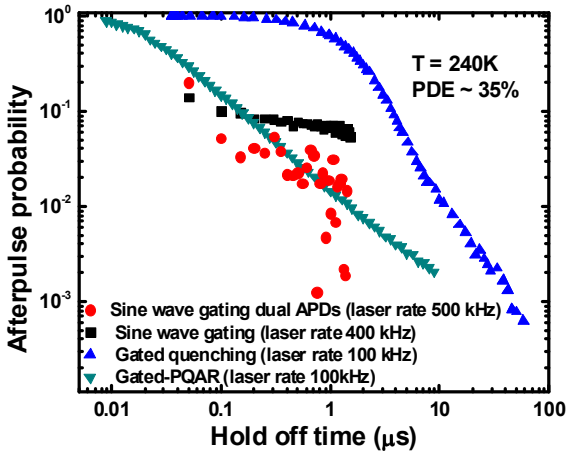



Figure 8.6: Afterpulse probabilities of single and dual SPADs with different gating scheme and laser repetition rate and similar PDE.

## 8.3 Improved Balanced Receiver

As stated several times in previous chapters, low background noise is beneficial for detecting small avalanche pulses that are associated with small charge flow and reduced afterpulsing [20]. For the balanced receive discussed above [66, 67], a significant component of the residual background noise was caused by imperfect phase matching of the two sine-wave signals. Better phase matching has been realized by adding a phase shifter (RF-Lambda RVPT0003MAC) to one branch of the sine-wave signal. The

adjustability of the signal phase allows greater suppression of the capacitive response; the background noise has been significantly reduced. In Figure 8.7, two graphs show the residual background noise from balanced pulsed-mode [34] and sine-wave gating (with an avalanche pulse) using the phase shifter. Both of the curves in Figure 8.7 demonstrate amplitude of noise levels less than 3 mV; five times smaller than sine-wave gating without the phase shifter. The residual background noise is well below the avalanche signal in the right graph for sine-wave gating.

These results were achieved using an 80 MHz gating frequency and no band-stop filters were necessary for balanced detection. The photon counts were recorded as the delay time of the laser pulse was varied relative to the gating signal of the SPADs. The temporal distribution of the photon counts is plotted in Figure 8.8. Assuming a Gaussian distribution, the full width at half maximum (FWHM) was 2.7 ns.

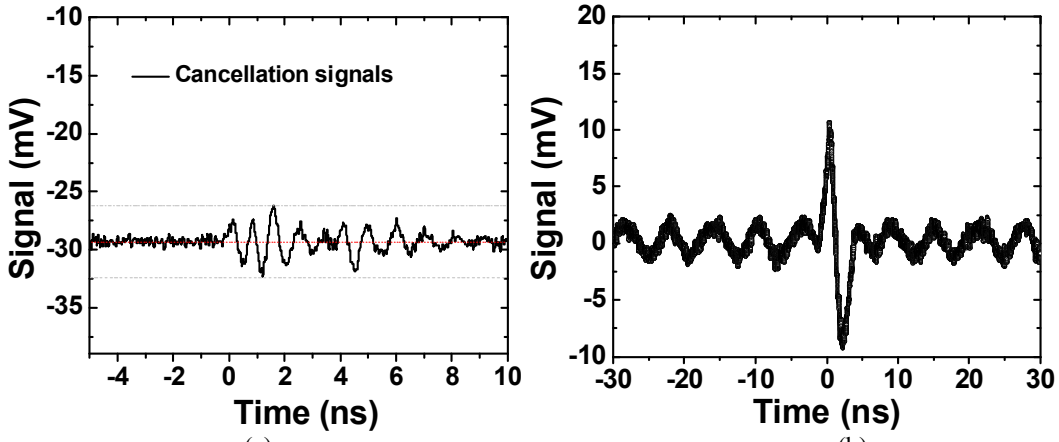



Figure 8.7: Cancellation effect for pulsed gating (a) and sine-wave gating (b); (b) also shows an avalanche pulse at gating frequency of 80 MHz. The signals were captured with oscilloscope without amplification.

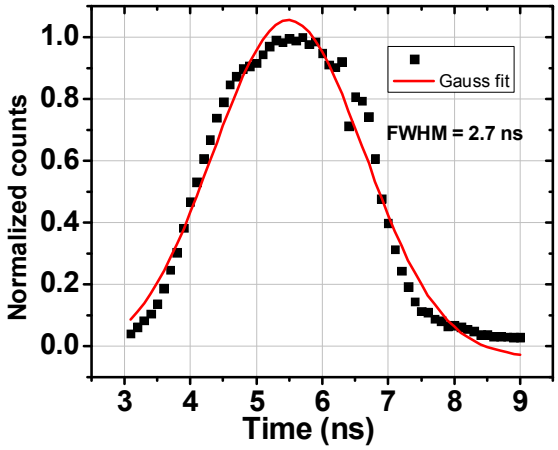



Figure 8.8: Effective pulse width of 80 MHz gating rate with laser pulse width of 40 ps for balanced diodes at 240 K.

Similar to the self-differencing technique, phase matching is crucial to improve the performance of the counting system [20, 69]. The first version of the sine-wave balanced receiver exhibited high residual background noise (~ 15 mV). As a result, acceptable detection efficiencies could only be achieved with avalanche pulses greater than 20 mV. By incorporating phase matching into one of the bias arms, the detectable avalanche pulse level has been reduced to less than 10 mV (Figure 8.7). The lower noise floor enables a lower threshold level, which is beneficial for reducing both afterpulsing and timing jitter.[1]

For a laser repletion rate of 20 MHz at 240 K, the dark count rate for photon detection efficiency of 10% is 8.9 kHz. (Figure 8.9) Figure 8.9 also compares the dark count rate versus photon detection efficiency (PDE) at 1 MHz and 20 MHz for balanced detectors with and without the phase shifter. For a laser repetition rate of 20 MHz, the dark count rate is significantly lower than that for the circuit with the phase shifter. While at 1MHz, no significant reduction in dark count probability (DCP) indicates less severe afterpulsing

effect, i.e., a hold–off time of 1 μs is sufficient for complete emission of the trapped carriers that cause afterpulsing.

Figure 8.10 compares the DCP with PDE at 240K for both pulsed gating and balanced sine-wave gating at 20 MHz laser repetition rate. The phase shifter reduces DCP at 10% PDE by approximately one order of magnitude compared with the original balanced receiver. The pulsed gating result with 2.5 ns pulse width (PW) overlaps with sine-wave gating with phase shifter. This indicates that the extra biasing gates in sine-wave gating are not the primary reason for the degraded performance at high laser repetition rate.[66] On the other hand, pulse width and avalanche charge flow are crucial in terms of reducing dark counts. A quantitative study of the total charge flow shows that there is 0.09 pC charge flow during an avalanche pulse with the 80 MHz sine-wave gating frequency. This is very close with the avalanche charge flow achieved with self-differencing [20]. Figure 8.10 shows that the best performance is achieved with the 1.4 ns pulse width pulsed gating. The reason is simply due to the faster quenching and smaller avalanche pulses associated with the narrower pulses. This is consistent with the excellent performance achieved with GHz sine-wave gating.[5, 70]

In Chapter 6, the sine wave gated single receiver shows that the total time that the device is biased above breakdown is independent of the gating frequency, indicating that gating frequency determined pulse width does not affect total time the device is biased above breakdown. Here we show that even the total time is the same, smaller effective pulse width still plays a very important role in affecting DCR (Figure 8.10). Therefore, higher gating rate in sine wave gating does help reduce DCR in terms of narrower effective pulse width, which is directly related with the total charge flow.

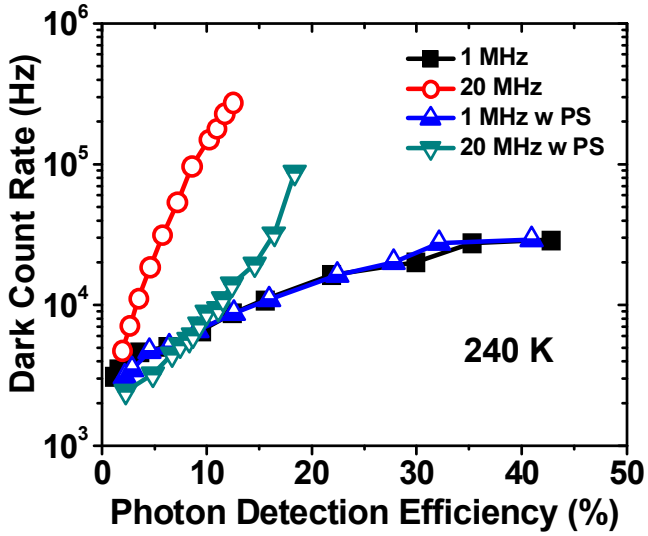



Figure 8.9: Comparison of balanced sine-wave gating results with and without phase shifter at 240K.

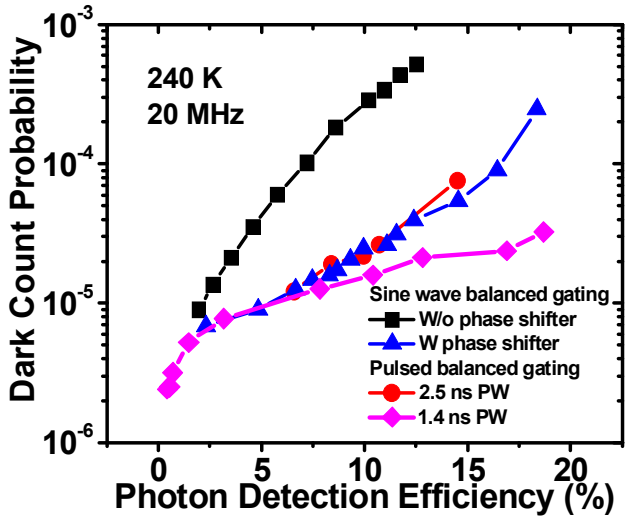



Figure 8.10: Photon counting result at 20 MHz counting rate from pulsed gating and sine-wave gating, both gating schemes are realized with balanced detection.

## 8.4 Conclusion

In this chapter, an InGaAs/InP SPAD pair was operated with 70 MHz sinusoidal gating at wavelength of 1310 nm. At a laser repetition rate of 1 MHz, photon detection efficiency of 43% has been achieved with dark count rate of 58 kHz at 240 K. At photon detection efficiency of ~ 10%, DCR is 9.6 kHz; the equivalent dark count probability is $2.8 \times 10^{-5}$. The afterpulsing probability is compared with gated PQAR and conventional sinusoidal gating. At higher repetition rate and the same PDE, the afterpulse probability has been reduced compared with a sinusoidal gated single SPAD. The lower afterpulsing is explained in terms of the integrated excess bias over time, $\Delta$, and some cancellation of dark counts in the SPAD pair. Parameters that affect afterpulsing in sinusoidal gating have also been discussed. The parameter $\Delta$ provides insight into the nature of sinusoidal gating and its effects on the performance of SPADs. I then demonstrated balanced sine-wave gating single photon counting with a phase shifter included in the AC bias are of one of the SPADs. This yielded improved results. The noise background was 5 times smaller and the dark count probability ($2.5\times10^{-5}$ at PDE of 10%) was reduced by an order of magnitude compared with the non-tunable balanced sine-wave gating result, both at 20 MHz laser repetition rate. The enhanced result also provides direction for further improvement in high-speed single photon counting operations.

# Chapter 9 Conclusions and Future Work

## 9.1 Conclusions

I have established two experimental set ups in pulsed gating and sine wave gating with balanced receiver for high rate single photon counting, enabling data transmission rate up to 20 MHz. Afterpulsing and dark count rate are significantly reduced due to the suppressed charge flow enabled by the small pulse detection. The smallest charge flow (0.09 pC) during avalanche events has been reduced by 200 times compared with the conventional gated mode. Table 9.1 compares all the gating schemes in terms of total charge flow.

| Quenching Method | Charge (pC) |
|---|---|
| **Gated Mode** | **18** |
| **PQAR** | **10** |
| **Gated-PQAR** | **1** |
| **Conventional Sine-Wave Gating** | **0.8** |
| **Balanced Sine-Wave Gating** | **0.5** |
| **Improved Balanced Sine-Wave Gating** | **0.16** |
| **Balanced Pulse Gating** | **0.09** |

Table 9.1: Comparison of total charge flow with different gating schemes.

In reviewing the efforts to reduce the charge flow, the most effective approach is to reduce the pulse width. Sine-wave gating is very effective in generating narrow pulses.

In fact, the smallest reported charge flow was achieved by 2 GHz sine-wave gating with a total charge flow of only 0.035 pC, which is even smaller than can be achieved with linear-mode counting (0.04 pC) [20]. One of the challenges with GHz range sine-wave gating is the relatively small detection efficiency. With sub-nanosecond gating width, timing jitter (~ 400ps for 4 GHz gating rate) also becomes a major issue. The performance tradeoffs associated with narrow gate widths, namely afterpulse rate, jitter, and detection efficiency, determine the optimum gating conditions.

Geiger-mode operation of Ge on Si SPADs is reported for the first time. The preliminary results show that the most critical issue is the relatively high dark current. The DCR and NEP of these devices are two orders of magnitude than those for InGaAs/InP SPADs. Although the performance is not as good as III-V SPADs, Ge on Si SPAD has the advantage of integration with CMOS technology thanks [71]. In addition, the afterpulsing issue in Ge on Si SPADs is less severe than for III-V SPADs.

The limiting factor for Geiger mode operation of III-V SPADs, especially at high data rates, is afterpulsing. Emerging technologies including sine-wave gating and self-differencing have improved the counting electronics so that the total charge flow is reduced, which in turn, has reduced afterpulsing. I have established an experimental system for sine-wave gating with the laser repetition rate variable from kHz to MHz. To further reduce the charge flow, a balanced configuration was introduced for sine-wave gating. This permits the gating frequency to be changed, unlike the conventional sine wave implementation which is restricted by the narrow band notch filter. Further improvement of the balanced sine-wave gated systems includes adding phase tuning in one arm of the AC bias circuitry. This has resulted in reduced residual noise floor. In

addition to sine-wave gating, common mode cancellation using a balanced receiver was also demonstrated for pulsed gating. This has yielded the best results in terms of total charge flow.

I have also developed a new measurement method to characterize afterpulsing in sine-wave gating. Since double pulse measurement is not practical for sine-wave gating, a gated window mode of the counter has been utilized to measure the afterpulsing probability at different hold-off times. Analysis and comparison of afterpulsing under different gating modes were carried out. Suppression of afterpulsing was achieved with the newly implemented counting systems. The reduced total charge flow is consistent with the suppressed afterpulsing at high repetition rates.

## 9.2 Future work

### 9.2.1 Linear Mode Counting

Based on the simulation results in Chapter 3, I have done some preliminary work on linear mode counting and obtained encouraging results. Though it is better to use a device with less abrupt breakdown curve, meaning lower ratio of the ionization coefficients and smaller excess noise, the devices I used for this initial work were InP-based SPADs. Figure 9.1 shows the current-voltage (I-V) and gain curves of the device before and after packaging with a transimpdance amplifier (TIA). The I-V curve after bonding exhibits higher leakage current; however, this part of current will not contribute to dark counts, since it is DC current.

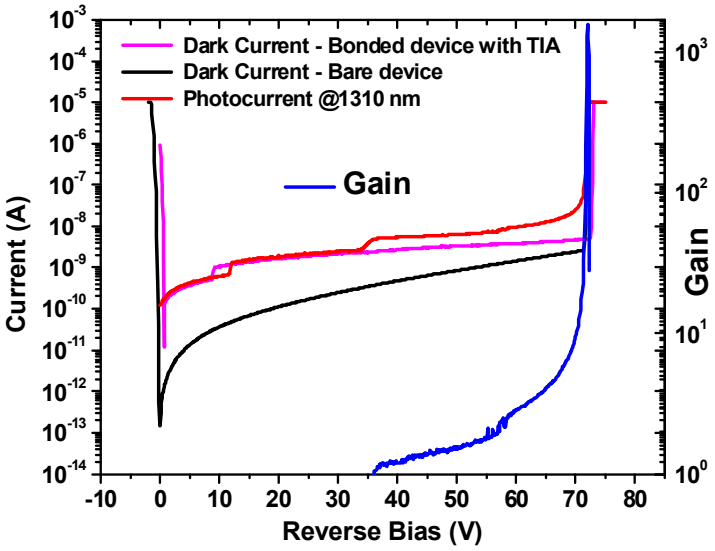



Figure 9.1: Current voltage and gain curve before and after packaging at room temperature.

The region of operation corresponds to gain in the range 100 and 1000 and bias voltages of 72 V and 73 V, respectively. Figure 9.2 shows the impulse response of an avalanche event that has sufficient extinction ratio for detection.

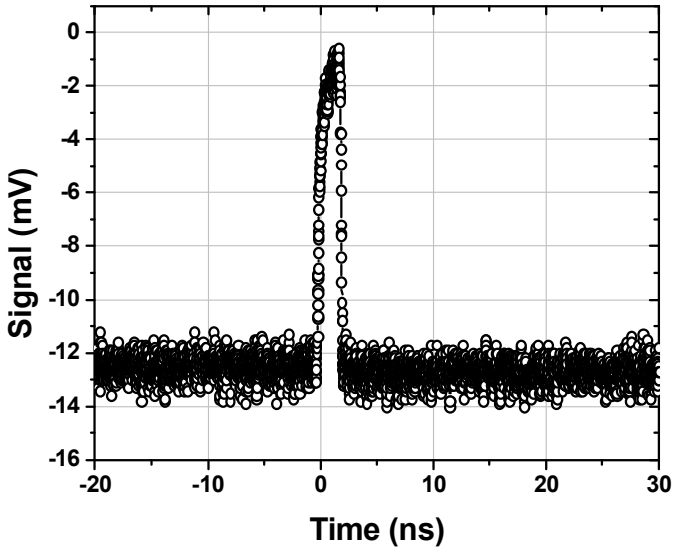



Figure 9.2: Impulse response extracted from output signal on oscilloscope. Reverse bias is 72.6 V.

This demonstrates linear mode counting. However, to improve the low-level detection characteristics, more effort is required. For instance, the multiplication region

determines the noise features of an APD in linear mode. Generally speaking, InAlAs exhibits lower excess noise than InP. Other materials such as HgCdTe and InAs are also candidates for low noise APDs. In addition, the design of the TIA is critical in terms of the overall system noise level. Both SPADs and TIA require further investigation in order to deliver a usable linear counting system.

### 9.2.2 Monolithic Integration of Dual SPADs

For hybrid integration of the balanced SPAD receiver the issue of residual noise signal caused by mismatched APD characteristics is a concern. The AC and DC response of the two devices can be different due to the fact that the two devices are fabricated and packaged separately. The parasitic capacitance and dark/photo currents affect the AC and DC responses, respectively. The ultimate solution to this issue is to integrate the dual SPADs monolithically. (Figure 9.3)

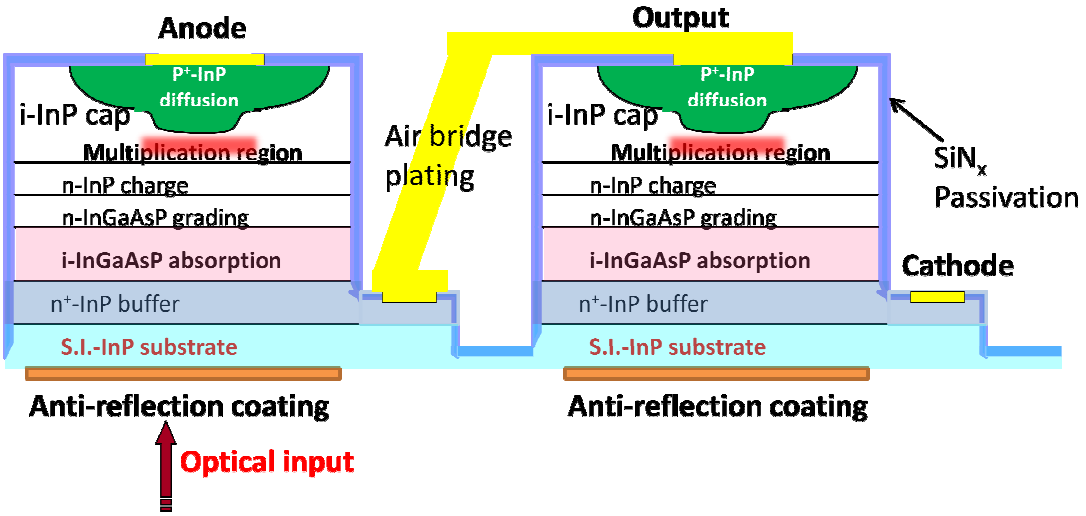



Figure 9.3: Device schematic for the monolithically integrated dual SPADs.

The challenge for monolithic integration is the design and fabrication of the dual SPADs. First of all, the SPADs that we have used to date were fabricated from epitaxial

layers grown on doped substrate. However, for monolithic integration, semi-insulating substrates are required. Secondly, the fabrication of mesa structured SPADs shown in Figure 9.3 is quite complicated, which could result in higher dark current. One alternative solution is to use planar structure devices.

If the on-wafer integrated balanced SPAD receiver becomes standard product, the device chip can be packaged with counting electronics, two bias tees, and a ground resistor. This type of receiver is expected to achieve lower noise and higher speed than the hybrid circuit.

### 9.2.3 SPADs + FET Topology

Another integration idea is based on the PQAR circuit. The PQAR circuit uses two FETs as the sensing and quenching elements. Integration of the two FETs with on the APD chip would reduce parasitic capacitance and improve performance.

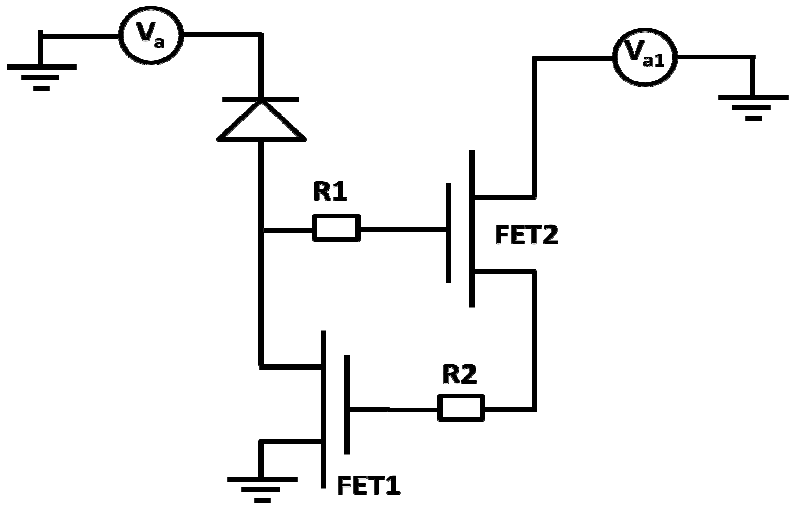



Figure 9.4: Circuit layout for the hybrid integrated SPAD and FETs.

The challenge of realizing the circuit shown in Figure 9.4 is the design of the two FETs and incorporating the resistors. In order to achieve the right timing and quenching

effect, the two FETs and resistors need to have accurate timing at the nanosecond level. There also is risk of device failure due to the failed incomplete quenching. Consequently, this approach requires careful design and implementation to realize the goal.

# Appendix A: Publications

[1] Z. Lu, W. Sun, J. Campbell, X. Jiang, and M. A. Itzler, "Pulsed Gating with Balanced InGaAs/InP Single Photon Avalanche Diodes," Submitted, 2012.

[2] Z. Lu, W. Sun, J. Campbell, X. Jiang, and M. A. Itzler, "Common-mode Cancellation in Sinusoidal Gating with Balanced InGaAs/InP Single Photon Avalanche Diodes," Quantum Electronics, IEEE Journal of, vol. 48, pp. 1505-1510, 2012.

[3] Z. Lu, Y. Kang, C. Hu, Q. Zhou, H.-D. Liu, and J. C. Campbell, "Geiger-Mode Operation of Ge-on-Si Avalanche Photodiodes," Quantum Electronics, IEEE Journal of, vol. 47, pp. 731-735, 2011.

[4] Z. Lu, X. Zheng, W. Sun, J. Campbell, X. Jiang, and M. A. Itzler, "InGaAs/InP Single Photon Avalanche Diodes," 221st ECS Meeting, vol. MA2012-01, p. 808, February 15, 2012 2012.

[5] Z. Lu, W. Sun, X. Zheng, J. Campbell, X. Jiang, and M. A. Itzler, "Balanced InGaAs/InP avalanche photodiodes for single photon detection," Proc. SPIE, pp. 84601H-84601H, 2012.

[6] Z. Lu, W. Sun, J. Campbell, X. Jiang, and M. A. Itzler, "Single photon detection with sine gated dual InGaAs/InP avalanche diodes," in Photonics Conference (IPC), 2012 IEEE, 2012, pp. 98-99.

[7] Z. Lu, X. Zheng, W. Sun, J. C. Campbell, X. Jiang, and M. A. Itzler, "Characterization of Sinusoidal Gating of InGaAs/InP Single Photon Avalanche Diodes," Proc. of SPIE, vol. 8375 - 37, 26 April 2012.

[8] Z. Lu, W. Sun, C. Hu, A. Holmes, J. C. Campbell, Y. Kang, and H.-D. Liu, "Ge on Si and InP/InGaAs single photon avalanche diodes," Proc. of SPIE, pp. 81551E-81551E, 2011.

[9] D. Dai, J. E. Bowers, Z. Lu, J. C. Campbell, and Y. Kang, "Temperature dependence of Ge/Si avalanche photodiodes," in Device Research Conference (DRC), 2010, 2010, pp. 231-232.

[10] W. Sun, Z. Lu, X. Zheng, J. C. Campbell, S. J. Maddox, H. P. Nair, and S. R. Bank, "High-Gain InAs Avalanche Photodiodes," Quantum Electronics, IEEE Journal of, vol. 49, pp. 154-161, 2013.

[11] W. Sun, Z. Lu, X. Zheng, J. C. Campbell, S. J. Maddox, H. P. Nair, and S. R. Bank, "Charge-compensated high gain InAs avalanche photodiodes," in Photonics Conference (IPC), 2012 IEEE, 2012, pp. 169-170.

[12] W. Sun, X. Zheng, Z. Lu, and J. C. Campbell, "Monte Carlo Simulation of InAlAs/InAlGaAs Tandem Avalanche Photodiodes," Quantum Electronics, IEEE Journal of, vol. 48, pp. 528-532, 2012.

[13] Q. Zhou, D. C. McIntosh, Z. Lu, J. C. Campbell, A. V. Sampath, H. Shen, and M. Wraback, "GaN/SiC avalanche photodiodes," Applied Physics Letters, vol. 99, pp. 131110-3, 2011.

[14] W. Sun, X. Zheng, Z. Lu, and J. C. Campbell, "Monte Carlo Simulation of AlxGa(1-x)As (x>0.6) Avalanche Photodiodes," Quantum Electronics, IEEE Journal of, vol. 47, pp. 1531-1536, 2011.

[15] W. Sun, X. Zheng, Z. Lu, B. Chen, A. L. Holmes, and J. C. Campbell, "Numerical simulation of InAlAs/InAlGaAs tandem avalanche photodiodes," in Photonics Conference (PHO), 2011 IEEE, 2011, pp. 280-281.

[16] H.-D. Liu, H. Pan, C. Hu, D. McIntosh, Z. Lu, J. Campbell, Y. Kang, and M. Morse, "Avalanche photodiode punch-through gain determination through excess noise analysis," Journal of Applied Physics, vol. 106, pp. 064507-4, 2009.

[17] X. Bai, P. Yuan, P. McDonald, J. Boisvert, J. Chang, R. Sudharsanan, M. Krainak, G. Yang, X. Sun, W. Lu, Z. Lu, Q. Zhou, W. Sun, and J. Campbell, "Development of low excess noise SWIR APDs," pp. 83532H-83532H, 2012.

# Bibliography

[1] A. Tosi, A. Della Frera, A. B. Shehata, and C. Scarcella, "Fully programmable single-photon detection module for InGaAs/InP single-photon avalanche diodes with clean and sub-nanosecond gating transitions," *Review of Scientific Instruments,* vol. 83, pp. 013104-8, 2012.

[2] M. A. Itzler, X. Jiang, M. Entwistle, K. Slomkowski, A. Tosi, F. Acerbi, F. Zappa, and S. Cova, "Advances in InGaAsP-based avalanche diode single photon detectors," *Journal of Modern Optics,* vol. 58, pp. 174-200, 2011.

[3] M. D. Eisaman, J. Fan, A. Migdall, and S. V. Polyakov, "Invited Review Article: Single-photon sources and detectors," *Review of Scientific Instruments,* vol. 82, pp. 071101-25, 2011.

[4] A. R. Dixon, Z. L. Yuan, J. F. Dynes, A. W. Sharpe, and A. J. Shields, "Continuous operation of high bit rate quantum key distribution," *Appl. Phys. Lett.,* vol. 96, pp. 161102-3, 2010.

[5] Y. Nambu, S. Takahashi, K. Yoshino, A. Tanaka, M. Fujiwara, M. Sasaki, A. Tajima, S. Yorozu, and A. Tomita, "Efficient and low-noise single-photon avalanche photodiode for 1.244-GHz clocked quantum key distribution," *Opt. Express,* vol. 19, pp. 20531-20541, 2011.

[6] F. Stellari, A. Tosi, F. Zappa, and S. Cova, "CMOS circuit testing via time-resolved luminescence measurements and simulations," *Instrumentation and Measurement, IEEE Transactions on,* vol. 53, pp. 163-169, 2004.

[7] B. F. Aull, A. H. Loomis, D. J. Young, A. Stern, B. J. Felton, P. J. Daniels, D. J. Landers, L. Retherford, D. D. Rathman, R. M. Heinrichs, R. M. Marino, D. G. Fouche, M. A. Albota, R. E. Hatch, G. S. Rowe, D. G. Kocher, J. G. Mooney, M. E. O'Brien, B. E. Player, B. C. Willard, Z.-L. Liau, and J. J. Zayhowski, "Three-dimensional imaging with arrays of Geiger-mode avalanche photodiodes," *Proc. SPIE,* vol. 5353, pp. 105-116, 2004.

[8] W. Becker, A. Bergmann, G. L. Biscotti, and A. Rueck, "Advanced time-correlated single photon counting techniques for spectroscopy and imaging in biomedical systems," *Proc. SPIE* vol. 5340, pp. 104-112, 2004.

[9] A. K. Ekert, "Quantum cryptography based on Bell's theorem," *Physical Review Letters,* vol. 67, pp. 661-663, 1991.

[10] F. B. C. H. Bennett, G. Brassard, L. Salvail, and J. Smolin, "Experimental Quantum Cryptography," *J. Cryptol.,* vol. 5, pp. 3-28, 1992.

[11] C. H. Bennett and G. Brassard, "Quantum cryptography: Public key distribution and coin tossing," *Theoretical Computer Science,* 1984.

[12] E. J. Gansen, M. A. Rowe, M. B. Greene, D. Rosenberg, T. E. Harvey, M. Y. Su, R. H. Hadfield, S. W. Nam, and R. P. Mirin, "Photon-number-discriminating detection using a quantum-dot, optically gated, field-effect transistor," *Nature Photonics,* vol. 1, pp. 585 - 588, 2007.

[13] B. E. Kardynał, Z. L. Yuan, and A. J. Shields, "An avalanche‐photodiode-based photon-number-resolving detector," *Nature Photonics,* vol. 2, pp. 425-428, 2008.

[14] S. Cova, A. Lacaita, and G. Ripamonti, "Trapping phenomena in avalanche photodiodes on nanosecond scale," *Electron Device Letters, IEEE,* vol. 12, pp. 685-687, 1991.

[15] C. Hu, X. Zheng, J. C. Campbell, B. M. Onat, X. Jiang, and M. A. Itzler, "Characterization of an InGaAs/InP-based single-photon avalanche diode with gated-passive quenching with active reset circuit," *Journal of Modern Optics,* vol. 58, pp. 201-209, 2011/02/10.

[16] M. A. Itzler, X. Jiang, M. Entwistle, K. Slomkowski, A. Tosi, F. Acerbi, F. Zappa, and S. Cova, "Advances in InGaAsP-based avalanche diode single photon detectors," 2011.

[17] X. Jiang, M. A. Itzler, R. Ben-Michael, and K. Slomkowski, "InGaAsP–InP Avalanche Photodiodes for Single Photon Detection," *IEEE Journal of selected topics in quantum electronics,* vol. 13, pp. 895-905, 2007.

[18] M. A. Itzler, M. Entwistle, and X. Jiang, "High-rate Photon Counting with Geiger-mode APDs," *IEEE Photonics Annual meeting,* vol. S1, 11 Oct. 2011.

[19] D. S. Bethune and W. P. Risk, "An autocompensating fiber-optic quantum cryptography system based on polarization splitting of light," *Quantum Electronics, IEEE Journal of,* vol. 36, pp. 340-347, 2000.

[20] 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.,* vol. 96, p. 071101, 2010.

[21] A. Lacaita, F. Zappa, S. Cova, and P. Lovati, "Single-photon detection beyond 1 μm: performance of commercially available InGaAs/lnP detectors," *Appl. Opt.,* vol. 35, pp. 2986-2996, 1996.

[22] X. Jiang, M. A. Itzler, B. Nyman, and K. Slomkowski, "Negative feedback avalanche diodes for near-infrared single-photon detection," *Proc. of SPIE* vol. 7320, pp. 732011-732011-10, 2009.

[23] M. Liu, C. Hu, J. C. Campbell, P. Zhong, and M. M. Tashima, "Reduce Afterpulsing of Single Photon Avalanche Diodes Using Passive Quenching With Active Reset," *Quantum Electronics, IEEE Journal of,* vol. 44, pp. 430-434, 2008.

[24] F. G. Zappa, M.; Cova, S.; Samori, C.; Giudice, A.C., "An Integrated Active-Quenching Circuit for Single-Phonton Avalanche Diodes," *Instrumentation and Measurement, IEEE Transactions on* vol. 49, 2000.

[25] F. Zappa, A. Lotito, A. C. Giudice, S. Cova, and M. Ghioni, "Monolithic active-quenching and active-reset circuit for single-photon avalanche detectors," *Solid-State Circuits, IEEE Journal of,* vol. 38, pp. 1298-1301, 2003.

[26] S. Cova, M. Ghioni, A. Lacaita, C. Samori, and F. Zappa, "Avalanche photodiodes and quenching circuits for single-photon detection," *Applied Optics,* vol. 35, pp. 1956-1976, 1996.

[27] A. Tosi, A. Bahgat Shehata, A. Della Frera, A. Dalla Mora, S. Tisa, F. Acerbi, and F. Zappa, "Compact detection module based on InGaAs/InP SPADs for near-infrared single-photon counting up to 1.7 μm," pp. 79452L-79452L, 2011.

[28] "Chong Hu_Dissertation."

[29] K. V. W. R. Clark, W. D. Waters, "InAlAs-InGaAs based avalanche photodiodes for next generation eye-safe optical receivers," *Proc. of SPIE,* vol. 6796, 2007.

[30] A. R. J. Marshall, J. P. R. David, and T. Chee Hing, "Impact Ionization in InAs Electron Avalanche Photodiodes," *Electron Devices, IEEE Transactions on,* vol. 57, pp. 2631-2638, 2010.

[31] W. Sun, Z. Lu, X. Zheng, J. C. Campbell, S. J. Maddox, H. P. Nair, and S. R. Bank, "Charge-compensated high gain InAs avalanche photodiodes," in *Photonics Conference (IPC), 2012 IEEE*, 2012, pp. 169-170.

[32] J. V. a. J. McPhate and L. D. a. M. Stapelbroek, "Mid-IR photon counting array using HgCdTe APDs and the Medipix2 ROIC," *Proc. of SPIE,* vol. 6660, 2007.

[33] W. M. Steven Bailey, Jinxue Wang, and Michael Jack, "Advances in HgCdTe APDs and LADAR Receivers," *Proc. of SPIE,* vol. 7660, 2010.

[34] Z. Lu, W. Sun, J. Campbell, X. Jiang, and M. A. Itzler, "Pulsed Gating with Balanced InGaAs/InP Single Photon Avalanche Diodes," *Submitted,* 2012.

[35] F. Zappa, S. Tisa, A. Gulinatti, A. Gallivanoni, and S. Cova, "Monolithic CMOS Detector Module for Photon Counting and Picosecond Timing," *Proceedings of the 34th European Solid-State Device Research Conference,* pp. 341-4, 2004.

[36] I. Rech, A. Gulinatti, F. Zappa, M. Ghioni, and S. Cova, "High-performance silicon single-photon avalanche diode array," pp. 73200H-73200H, 2009.

[37] A. R. Hawkins, W. Wu, P. Abraham, K. Streubel, and J. E. Bowers, "High gain-bandwidth-product silicon heterointerface photodetector," *Applied Physics Letters,* vol. 70, pp. 303-305, 1997.

[38] Y. Kang, H.-D. Liu, M. Morse, M. J. Paniccia, M. Zadka, S. Litski, G. Sarid, A. Pauchard, Y.-H. Kuo, H.-W. Chen, W. S. Zaoui, J. E. Bowers, A. Beling, D. C. McIntosh, X. Zheng, and J. C. Campbell, "Monolithic germanium/silicon avalanche photodiodes with 340 GHz gain–bandwidth product," *Nature Photonics,* vol. 3, pp. 59-63, 2008.

[39] I. Rech, C. Cammi, A. Gulinatti, M. Ghioni, and S. Cova, "Monolithic front-end system for photon timing applications," *LEOS 2009 -22nd Annuall Meeting of the IEEE Lasers and Electro-Optics Society,* pp. 299-300, 2009.

[40] S. Assefa, F. Xia, and Y. A. Vlasov, "Reinventing germanium avalanche photodetector for nanophotonic on-chip optical interconnects," *Nature,* vol. 464, pp. 80-84, 2010.

[41] X. Chunlai, X. Haiyun, C. Buwen, B. Anqi, H. Weixuan, Y. Yude, and W. Qiming, "Si/Ge separated absorption charge multiplication avalanche photodetector with low dark current," in *Group IV Photonics, 2009. GFP '09. 6th IEEE International Conference on*, 2009, pp. 178-180.

[42] M. Liu, "Dissertation_Mingguo Liu," 2008.

[43] D. Dai, J. E. Bowers, Z. Lu, J. C. Campbell, and Y. Kang, "Temperature dependence of Ge/Si avalanche photodiodes," in *Device Research Conference (DRC), 2010*, 2010, pp. 231-232.

[44] Z. Lu, Y. Kang, C. Hu, Q. Zhou, H.-D. Liu, and J. C. Campbell, "Geiger-Mode Operation of Ge-on-Si Avalanche Photodiodes," *Quantum Electronics, IEEE Journal of,* vol. 47, pp. 731-735, 2011.

[45] M. Liu, C. Hu, X. Bai, X. Guo, J. C. Campbell, Z. Pan, and M. M. Tashima, "High-Performance InGaAs/InP Single-Photon Avalanche Photodiode," *Selected Topics in Quantum Electronics, IEEE Journal of,* vol. 13, pp. 887-894, 2007.

[46] R. J. Young, S. J. Dewhurst, R. M. Stevenson, A. J. Shields, P. Atkinson, K. Cooper, and D. A. Ritchie, "Controlling the polarization correlation of photon pairs from a charge-tunable quantum dot," *Applied Physics Letters,* vol. 91, pp. 011114-3, 2007.

[47] 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.,* vol. 91, p. 041114, 2007.

[48] A. R. Dixon, J. F. Dynes, Z. L. Yuan, A. W. Sharpe, A. J. Bennett, and A. J. Shields, "Ultrashort dead time of photon-counting InGaAs avalanche photodiodes," *Appl. Phys. Lett.,* vol. 94, p. 231113, 2009.

[49] N. Namekata, S. Adachi, and S. Inoue, "1.5 GHz single-photon detection at telecommunication wavelengths using sinusoidally gated InGaAsInP avalanche photodiode," *OPTICS EXPRESS,* vol. 17, pp. 6275-6282, 2009.

[50] N. Namekata, S. Sasamori, and S. Inoue, "800 MHz single-photon detection at 1550 -nm using an InGaAsInP avalanche photodiode operated with a sine wave gating," *OPTICS EXPRESS,* vol. 14, pp. 10043-10049, 2006.

[51] J. Zhang, P. Eraerds, N. Walenta, C. Barreiro, R. Thewa, and H. Zbinden, "2.23GHz gating InGaAs/InP single-photon avalanche diode for quantum key distribution," *Proc. of SPIE,* vol. 7681, 28 April 2010.

[52] N. Namekata, S. Adachi, and S. Inoue, "Ultra-Low-Noise Sinusoidally Gated Avalanche Photodiode for High-Speed Single-Photon Detection at

Telecommunication Wavelengths," *IEEE Photonics technology letters,* vol. 22, 2010.

[53] J. Zhang, R. Thew, C. Barreiro, and H. Zbinden, "Practical fast gate rate InGaAs/InP single-photon avalanche photodiodes," *Appl. Phys. Lett.,* vol. 95, p. 091103, 2009.

[54] A. Restelli and J. C. Bienfang, "Avalanche discrimination and high-speed counting in periodically gated single-photon avalanche diodes," in *SPIE*, Baltimore, Maryland, USA, 2012, pp. 83750Z-8.

[55] A. Tosi, A. Dalla Mora, A. Della Frera, A. Bahgat Shehata, A. Pifferi, and D. Contini, "Fast-gated SPAD for ultra-wide dynamic range optical investigations," in *IEEE Photonics Society, 2010 23rd Annual Meeting of the*, 2010, pp. 185-186.

[56] F. Zappa, A. Tosi, and S. Cova, "InGaAs SPAD and electronics for low time jitter and low noise," pp. 65830E-65830E, 2007.

[57] A. Tomita and K. Nakamura, "Balanced, gated-mode photon detector for quantum-bit discrimination at 1550 nm," *Opt. Lett.,* vol. 27, pp. 1827-1829, 2002.

[58] J. Yi, E. Wu, W. Guang, and Z. Heping, "Optically Self-Balanced InGaAs/InP Avalanche Photodiode for Infrared Single-Photon Detection," *Photonics Technology Letters, IEEE,* vol. 22, pp. 173-175, 2010.

[59] M. A. Itzler, X. Jiang, B. M. Onat, and K. Slomkowski, "Progress in self-quenching InP-based single photon detectors," *Proc. of SPIE* vol. 7608, pp. 760829-760829-13, 2010.

[60] N. Namekata, S. Adachi, and S. Inoue, "High-Speed Single-Photon Detection Using 2-GHz Sinusoidally Gated InGaAsInP Avalanche Photodiode," *QuantumCom* pp. 34-38, 2010.

[61] Z. Lu, X. Zheng, W. Sun, J. C. Campbell, X. Jiang, and M. A. Itzler, "Characterization of Sinusoidal Gating of InGaAs/InP Single Photon Avalanche Diodes," *Proc. of SPIE,* vol. 8375 - 37, 26 April 2012.

[62] G. P. Agrawal, "Fiber-Optic Communications Systems," *John Wiley & Sons, Inc.,* vol. 3rd ed. Chp. 10, 2002.

[63] http://en.wikipedia.org/wiki/Differential_signaling.

[64] A. Restelli and J. C. Bienfang, "Avalanche discrimination and high-speed counting in periodically gated single-photon avalanche diodes," Baltimore, Maryland, USA, 2012, pp. 83750Z-8.

[65] Q.-L. Wu, Y. Liu, Z.-F. Han, W. Chen, Y.-M. Dai, G. Wei, and G.-C. Guo, "Gated-mode integrated single photon detector for telecom wavelengths," in *SPIE*, Boston, MA, USA, 2007, pp. 67711A-7.

[66] Z. Lu, W. Sun, J. Campbell, X. Jiang, and M. A. Itzler, "Common-mode Cancellation in Sinusoidal Gating with Balanced InGaAs/InP Single Photon Avalanche Diodes," *Quantum Electronics, IEEE Journal of,* vol. 48, pp. 1505-1510, 2012.

[67] Z. Lu, W. Sun, X. Zheng, J. Campbell, X. Jiang, and M. A. Itzler, "Balanced InGaAs/InP avalanche photodiodes for single photon detection," *Proc. SPIE,* pp. 84601H-84601H, 2012.

[68] Z. L. Yuan, B. E. Kardynal, A. W. Sharpe, and A. J. Shields, "High speed single photon detection in the near infrared," *Applied Physics Letters,* vol. 91, p. 041114, 2007.

[69] Z. L. Yuan, A. R. Dixon, J. F. Dynes, A. W. Sharpe, and A. J. Shields, "Gigahertz quantum key distribution with InGaAs avalanche photodiodes," *Appl. Phys. Lett.,* vol. 92, pp. 201104-3, 2008.

[70] N. Namekata, S. Mori, and S. Inoue, "Quantum key distribution over an installed multimode optical fiber local area network," *Opt. Express,* vol. 13, pp. 9961-9969, 2005.

[71] D. Liang and J. E. Bowers, "Integrated Optoelectronic Devices on Silicon," *MRS Online Proceedings Library,* vol. 1396, pp. null-null, 2012.

# Vita

Zhiwen Lu (卢致文) was born in Xiaogan, Hubei Province, China on Oct. 28, 1985, the daughter of Shengchu Lu （卢胜初）and Shouqing Wang (王首庆). She graduated from Xiaogan High School in Hubei province and entered University of Science and Technology of China in Heifei, Anhui, China. After receiving her Bachelor of Science degree in Physics from University of Science and Technology of China in 2007, she continued her study as a graduate student in the Department of Physics, University of Virginia. In 2008, she joined Dr. Joe C. Campbell's group and transferred to the department of Electrical and Computer Engineering, University of Virginia, where she received his Master degree in Engineering in 2009.

Address: 501 Seymour Rd, Apt. 1

Charlottesville, VA 22903

This dissertation was typed by the author.