A Capacitive Quenching Circuit for Use with Single Photon Avalanche Detectors

Graham Goldberg

A thesis submitted to the University of Sheffield in fulfilment of the requirements for the degree of Master of Philosophy

November 2013

Department of Electronic and Electrical Engineering

Sir Frederick Mappin Building

Mappin Street

Sheffield

S1 3JD

Contents

Acknowledgements

This thesis is dedicated to the memory of Hannah Monchar and Norma Sneeden, without whom this would not have been possible.

I would first like to extend my deepest thanks to Jo Shien Ng and Chee Hing Tan for giving me the opportunity to pursue a Masters degree in Electronic Engineering. The guidance, support and supervision you both provided is very much appreciated. I would like to further thank Jo Shien for the patience and understanding you showed by allowing me to complete a six month Internship in Berlin during my studies.

I am very grateful for the contributions made by my colleagues, you were always happy to take the time to have a discussion with me about my work, offer your insight and advice on the approach I was taking, even if it was not so relatable to your own research. Particular thanks for this go to Richard Taylor, James Green, Matthew Hobbs, Rajiv Gomes, Jacques Durand and Simon Dimler. I would also like to thank David Williams, Pin Jern Ker, Jennifer Ong, Syahrin Idris and Daniel Ong for the assistance you provided.

To my family, especially my parents, brother Peter and grandma Zena, I would like to thank you for always being there for me, raising me and generating the interest I have in further learning. Your unwavering belief that I would achieve this qualification helped to keep me motivated on the days when I was not so sure.

"Happiness lies in the joy of achievement and thrill of creative effort." - Franklin D Roosevelt

Abstract

In this thesis a Capacitive Quenching Circuit (CQC) is modelled using Linear Technology's Simulation Program with Integrated Circuit Emphasis (LT SPICE). Key components are optimised and the entire circuit is modelled. Once satisfied with the simulation, Printed Circuit Board (PCB) design is considered. Computer programs Intelligent Schematic Input System (ISIS) and Advanced Routing and Editing Software (ARES) are used to produce a schematic and a virtual PCB which was submitted for fabrication. The circuit is then constructed and the results are compared to modelled and published data.

Chapter One - Introduction

1.1 Background

Single Photon Avalanche Detectors (SPADs) are a subset of Avalanche Photodiodes (APDs). APDs were researched and developed to replace Photo Multiplier Tubes (PMTs) in numerous applications, most notably optical communications, due to the advantages offered by the solid state components such as smaller dimensions, lower operating voltages, greater robustness and immunity to magnetic fields [01]. APDs amplify photocurrent by a significant factor, allowing them to be used as very sensitive detectors.

SPADs are operated in so-called 'Geiger Mode' with the reverse bias exceeding their avalanche breakdown voltages. When in this mode, absorption of a single photon by a SPAD can result in a large detectable avalanche current pulse that increases rapidly with time [02]. This is possible because one of the two photo-generated carriers, electron or hole, may initiate an avalanche breakdown through successive impact ionisation events. Although APDs also utilise successive impact ionisation events to provide avalanche gains, they operate in linear mode which limits the maximum gain. A large reverse bias is essential to produce a high electric field within the SPAD to facilitate the impact ionisation. Once a photon is detected the resulting avalanche is self-sustaining, meaning that until the voltage is reduced to a level below breakdown, current will continue to flow [02][03].

Disruption of the avalanche process is termed quenching, and involves external components or circuitry [02]. After the detection of a photon the current becomes very large and additional photons will contribute little to the total current, hence it will be difficult to distinguish a one photon event from a multiple photon event. In order to detect subsequent photons the addition of a quenching circuit to the SPAD is required. Also, in the event of a large current, the diode is likely to be damaged by excess heat [03]. Since the SPAD must be operated with a quenching circuit, the design of the quenching circuitry is often integral to

the overall device performance, and often compromises are made depending on the targeted applications.

Some limitations of the current quenching circuits include an upper limit on the operation frequency, formation of high and low pass filters and poor synchronisation with the device. These limit the achievable data rate, distort the over-bias pulses and cause quenching to not take place as soon as an avalanche current begins to build up, respectively. This is mainly caused by active components like fast transistors and comparators that require external power to operate, in order to attempt to detect the avalanche and respond accordingly [04]. A Capacitive Quenching Circuit (CQC) was designed by Dimler *et al.*, to overcome these problems and is discussed further in 1.5.9 [05]. The CQC uses a capacitance to limit the total charge flow during an avalanche event. This work replicated the CQC and demonstrated that it was a viable alternative to the other circuits researched.

1.2 Applications of SPADs

The applications for single photon detectors include high-energy physics experiments [06], laser-based distance measurement [07], testing Very Large Scale Integration (VLSI) circuits [08], photon correlation spectroscopy [09], fluorescent and luminescent decays [10][11], quantum key distribution [12], and optical fibre characterisation [13]. Each of these applications are described briefly below.

Laser-based distance measurement uses a pulsed diode laser in conjunction with photon detectors to calculate the distance between an object and the measurement system. This method was developed for use when some distances could not be measured by conventional means. Short, picoseconds duration laser light pulses strike a target with some light reflecting and scattering back towards the measurement system. The returning photons are detected and their time of arrival is compared to that of photons returning from a reference point whose distance is known. The delay between arrivals of each set of photons is used to accurately calculate the distance of the target [07]. This technique is commonly used for satellite ranging, with distances as far as 30,000 km being measured

using lasers with 532 nm wavelength. Research is being conducted to move the wavelength of operation from the current 350-1064 nm to ≈ 1500 nm, which is safer for human eyes [14].

Complementary Metal Oxide Semiconductor (CMOS) is the technology most widely used on VLSI circuits. When device size reduces and the overall number of devices per circuit increases, traditional methods of testing at individual transistor levels become ineffective. As electrons travel through the Metal Oxide Semiconductor Field Effect Transistor (MOSFET) channel, they experience high electric fields, generating either electron-hole pairs by impact ionisation or near-infrared photons by recombination [15]. The emitted photons can be detected by a suitable SPAD. As the electrons only experience this high electric field whilst the transistor is switching, detection of the luminescence confirms the device is operating correctly. As the test is non-invasive, dynamic and static operation can be examined without increasing current or voltage levels [08].

Data security, which is vital in many aspects of our societies, necessitates encryption of sensitive information to prevent it from easily being read by a third party if intercepted. The encryption key needs to be supplied to the chosen user(s), usually over an unsecured channel, which can, however be intercepted. Quantum Key Distribution (QKD) coupled with Quantum Cryptography ensures security of the key. The key needs to have the same length as the message, be randomly generated, and only used once. To generate a key, light is transmitted along an optical fibre from point A, termed 'Alice', to point B, termed 'Bob', with both points using a polarising filter. Alice positions her filter randomly in one of two states on each clock cycle, for example 45° and 90°. This encodes the data with one state 'high' and the other 'low'. At the other end of the fibre, Bob uses a filter in conjunction with a SPAD to measure the direction of polarisation and receive the data from Alice. As Alice does not disclose the direction she chose, Bob must also select which way to position his filter for each cycle.

On average, there is a 25% probability that Bob will obtain a so-called 'unambiguous' result, which is where the received bit is the same as the transmitted bit. This is because of the 50% likelihood that he detects a photon, coupled with the 50% likelihood that he will choose

the correct orientation of the polariser. If a photon is not detected then an 'ambiguous' result is produced as the variable responsible is unknown. When the transmission has ended Alice and Bob compare the clock cycles during which photons were detected. This provides a subset of information, which is about a quarter of the length of the original broadcast, and is the key. This method is known as the B92 protocol and is quite vulnerable to an intercept-resend attack by a third party, termed 'Eve'. A more secure technique is the BB84 protocol, which uses an additional pair of polarising filters incompatible with the prior two, to produce four possible polarisation states so that Eve cannot guarantee the polarity of any photon she receives. Both protocols utilise fibre optics to transmit the key and operate at wavelengths of between 850 nm to 1550 nm depending on the distance between Alice and Bob [12].

Optical-Time-Domain Reflectometry (OTDR) is a technique used to determine the characteristics of optical fibres. A 'Y-coupler' is connected to one end of a fibre, with a laser attached to one arm and a SPAD to the other. The laser provides pulses of light that travel down the fibre, before a portion of the light is reflected back and some becomes incident on the SPAD. Time delay between emission of light by the laser and detection of light by the SPAD is measured, along with the optical power of the returning light in order to quantify the attenuation experienced [13].

1.3 Optical Communications

Optical fibres have been developed since 1920, with the first generation of optical fibres used for telecommunications having core diameters of either 50 or 62.5 microns and operated at 850 nm [16]. Figure 1.1 shows the three 'windows' within which almost all of the optical communication industry operates. These are chosen due to their low level of attenuation and dispersion at specific wavelengths.

The first band is at 800-900 nm, and was brought into use due to its relatively low losses of ≈ 2 dB/km. For this band, standard Silicon photodiodes and APDs are used in the receiver modules. This band is still used over shorter distances, such as in Local Area Networks

(LANs) [17], due to their low cost. It was soon discovered that the peak that develops after the first window decreases just after 1000 nm to offer a lower attenuation of 0.5 dB at 1300 nm. This second window was favoured by service providers for national systems as fewer repeaters were required. However, photodiodes made with materials other than Silicon had to be developed as Silicon does not respond to 1300 nm light [16].

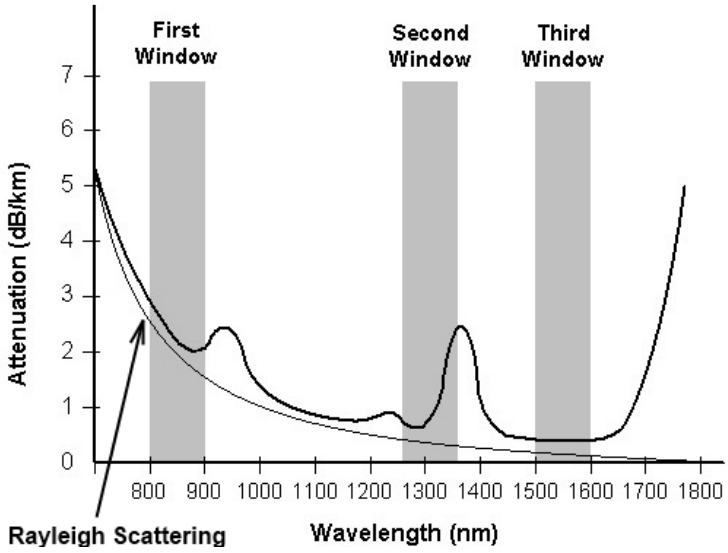


Figure 1.1 - Attenuation In Optical Fibres For Varying Wavelengths

Towards the latter part of the second window, the attenuation rapidly increases due to absorption by hydroxyl ions from water contamination during manufacturing, within the optical fibre material. After ≈ 1400 nm this absorption is no longer significant, resulting in a third window at 1500-1600 nm with a loss as low as 0.2 dB/km [17][18]. This wavelength has the added advantage that if the fibres are doped with Erbium, they naturally act like optical amplifiers. This is particularly attractive to communication providers serving a dense population, as this can help to restore the attenuated signal as it branches out across the network [16][18].

These three windows determine the main wavelengths of operations for optical communications. The second and third windows in turn drive the development of long wavelength photodiodes, APDs and SPADs. These detectors utilise different materials, in various combinations and with increasingly complex structures and external circuitry, as higher performance and efficiency are pursued [19].

1.4 Photo-detectors

Photo-detectors can be categorised into two groups depending on how they interact with light, external photoemission and internal photoemission. External photoemission occurs when photons interact with either a metal or semiconductor in a vacuum, causing the release of electrons into the space. PMTs utilise this phenomena. In internal photoemission, photons excite electrons from the valence band to the conduction band in the material, resulting in a photocurrent. This type of photoemission can be sub-categorised into photovoltaic and photoconductive, with solar cells and photodiodes being examples respectively [20].

1.4.1 Photo Multiplier Tubes

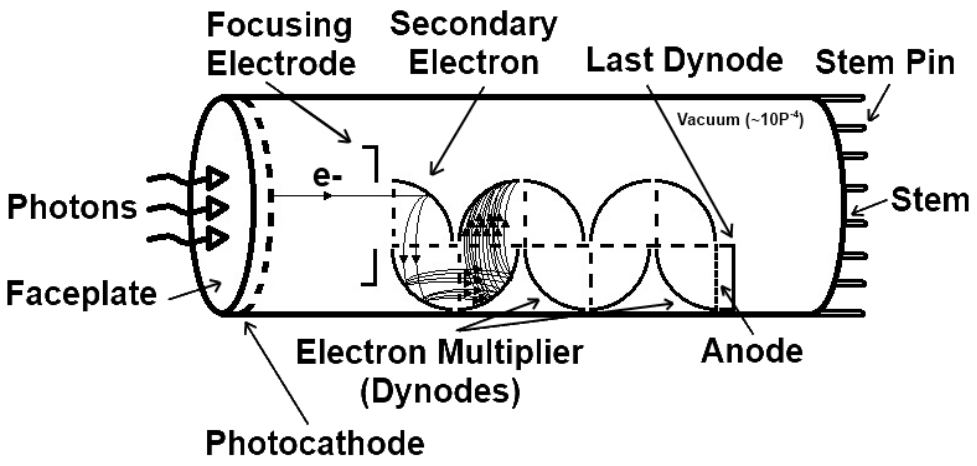


Figure 1.2 - Head-On Photo Multiplier Tube

In a PMT, shown schematically in Figure 1.2, light enters through a window at either the end or side of the detector. This light then becomes incident on a photocathode, which releases excited electrons into the vacuum chamber. These electrons are then channelled by the focusing electrodes into the electron multiplier, which has a series of dynodes, each with negative bias, in order to multiply the number of electrons. The negative bias for a given dynode is a function of its distance from the photocathode, with the furthest having the smallest negative bias. As the electron travels through the multiplier, it is accelerated away from the photocathode, colliding with the less negative dynode. This energy causes new

electrons to be emitted from the first dynode and accelerated towards the sequential one, eventually producing a large number of electrons to be collected by the anode. The position of the first dynode has to be carefully calculated in order to minimise the chance of missed collisions with emitted electrons from the photocathode.

With up to 19 dynodes, typical gains offered by PMTs are 10^8, or 160 dB. The secondary emission process in the electron multiplier does not add noise, with the exception of shot noise, resulting in a high signal to noise ratio [21][22]. In general, PMTs can be designed to detect photons from visible to near ultraviolet wavelengths, with performance mainly affected by the materials used for the photocathode and the window. There are eight main compositions used to make the photocathode, usually compounds of alkali metals designed to have a very low work function, so that even photons with low energies can interact with the material to produce an electron. Typical photocathode response curves are shown in Figure 1.3. Bi-Alkali is usually a combination of Sb-Rb-Cs or Sb-K-Cs, with Multi-Alkali having the addition of Na [22].

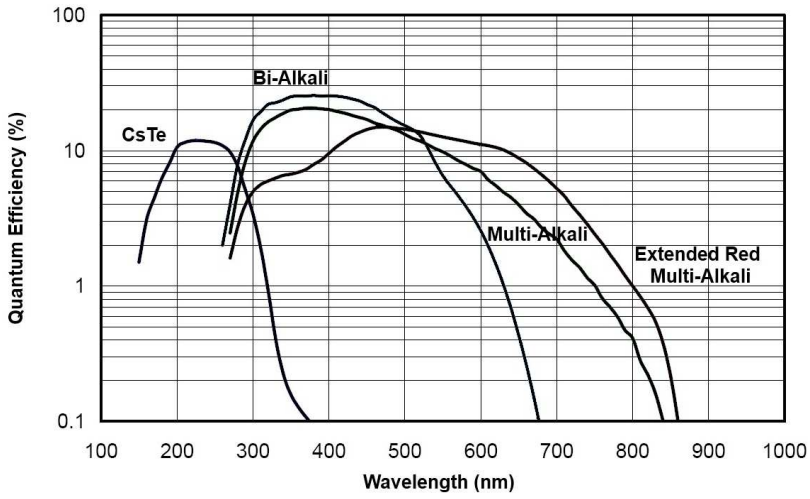


Figure 1.3 - Response Curves of Different Photocathodes at Varying Wavelengths

A Bi-Alkali photocathode can achieve a peak quantum efficiency of 25% at 400 nm, but is insensitive to wavelengths longer than 680 nm. Multi-Alkali offers a larger range than Bi-Alkali and can detect up to 850 nm, near infra-red, at the cost of a lower peak quantum efficiency of 20% at 400 nm. Special processing can allow detection of wavelengths up to 930 nm, however, quantum efficiency is reduced drastically above 700 nm [21][22].

Quantum efficiency is the probability that an incoming photon will generate an electron-hole pair within a photo-sensitive device. It can be expressed as $\eta = \frac{I_p \times P_{opt}}{h\nu \times q}$, where I_p is the incident optical power, and P_{opt} is the optical power [23][24].

In spite of their excellent signal to noise ratio, there are significant disadvantages associated with PMTs that make them unsuitable for some applications. Gradual deterioration of the dynodes means electrons can suffer from 'drift', which relates to a short-term variation in their trajectory through the electron multiplier, or 'life', which is a long-term variation. This can affect the secondary electron emission and cause fluctuations in the gain [22]. PMTs are also subject to problems caused by hysteresis following the application of an input. This is due to the deviation of electrons within the multiplier and can result in charge building on the ceramic attachments for the dynodes. In some cases this can last for several seconds which is not ideal for high speed applications [25].

The limited range of wavelengths that PMTs are able to detect is another significant disadvantage, since optical fibre communications require operations in the windows of 800-900 nm, 1300 nm and 1500-1600 nm [16][17]. Figure 1.3 shows that the photocathode materials currently available have a limit of 850 nm or 930 nm, depending on whether standard or special processing is used. This makes them unsuitable in all except the first window, where they exhibit a quantum efficiency of 0.1-0.3%, which is very poor [22].

1.4.2 Hybrid Photo-detectors (HPDs)

One way to circumvent the shortcomings of the PMT is to replace the electron multiplier with an Avalanche Photodiode, to be discussed further in section 1.4.3, which also provides multiplication of the electrons. Such PMTs are termed HPDs. As there are no dynodes to experience a gradual loss of performance, and the electrons are accelerated only from the photocathode to the photodiode, the 'drift' and 'life' exhibited by standard PMTs are greatly reduced. Thus the HPD only displays minor fluctuations in gain during the operation lifetime, provided it is not subjected to magnetic fields or vibrations. Spatial uniformity is now no longer dependent on the position of the photocathode relative to the dynodes and the

procedure the emitted electrons undergo is simpler, providing a more uniform output and reduced hysteresis.

As the range of photocathode materials is the same for HPDs as conventional PMTs, there is no improvement on the operational wavelengths, making both devices unsuitable for communications. In addition, there are still issues caused by the large sizes and high operating voltages, which proves problematic in areas with poor access and where supplying power is difficult, as with trans-Atlantic cables. For these reasons replacements of PMTs are sought, ideally with a high signal to noise ratio at the 1310 nm and 1550 nm wavelength windows [26][27].

1.4.3 Avalanche Photodiodes

The Avalanche Photodiode is a photoconductive device that relies on internal photoemission. APDs are used in a variety of fields, with optical communications currently being the most noteworthy. The small size of the device in comparison to the PMT is favourable when designing receivers at the end of optical fibres, as it allows many devices to occupy a small area. To describe its operation, a reverse-biased p-i-n diode can be considered, such as the one shown in Figure 1.4. This forms a basic APD with the i-region being the multiplication region. When a reverse bias applied across it, the free electrons in the i-region drift towards the n^+ side of the device, and the holes drift toward the p^+ cladding. As they travel with the field in opposite directions, they gain energy from the electric field and their kinetic energies increase [28].

If the electric field is sufficiently high, some of the carriers will acquire sufficient energies to initiate impact ionisation. In an electron-initiated impact ionisation event, the initiating electron gives up some of its energy to promote a new electron from the valence band to the conduction band. As this leaves behind a hole in the valence band, there are three free final carriers, two electrons and one hole. Successive impact ionisation events can result in a measurable multiplication of the initial number of carriers. Increasing the reverse bias further will cause the multiplication factor to become a very large value and eventually

result in avalanche breakdown. The voltage at which avalanche breakdown occurs is termed the breakdown voltage, V_B. In impact ionisation, a minimum energy E_g must be attained prior to impact ionisation occurring, due to energy losses through scattering, as well as the surpassing of 'dead time' and 'dead space', which is the minimum time and distance that the carriers must adhere to [29].

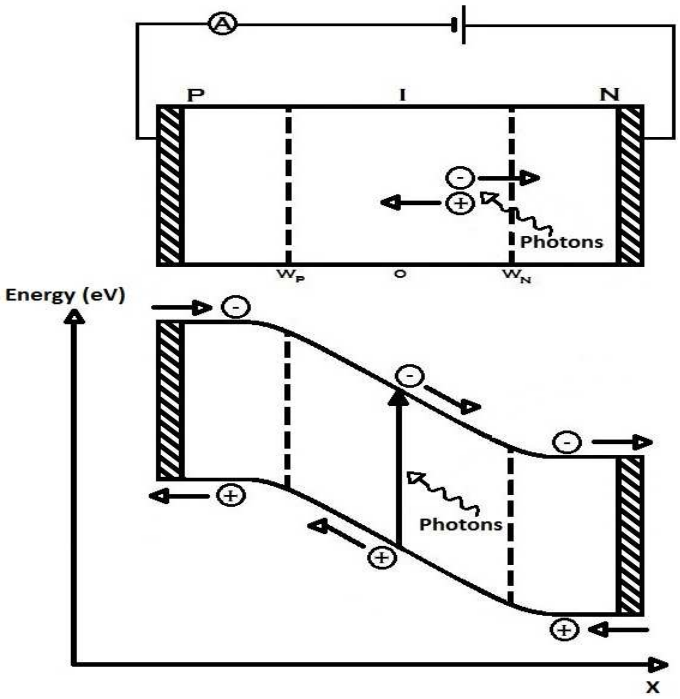


Figure 1.4 – Structure and Band Diagram of an Avalanche Photo-Diode

Commercial APDs have more complex structures than the p-i-n diode shown in Figure 1.4, often taking the form of a Separate Absorption and Multiplication (SAM) APD, in order to provide greater sensitivity. The SAM APD was modified from the original InGaAs APDs, as the unmodified versions suffered from high dark currents due to tunnelling [27][30]. A basic diagram of the regions that make up the SAM APD and typical electric field profile during operation are shown in Figure 1.5. Using material with a narrow bandgap for the absorption region allows long wavelengths to be detected which is an important quality for optical communications, as discussed in 1.3 A low field in the absorption region means that impact ionisation is unlikely to occur, minimising dark counts. Increasing the thickness of this region improves absorption. A high field in the multiplication region with sufficient magnitude for impact ionisation to occur will give rise to a high gain from successive impact ionisation

events. By separating the absorption and multiplication regions in this manner, tunnelling is reduced as the width of the barrier through which the carriers must tunnel is greatly increased [27][29].

By applying a voltage greater than V_B to an APD, the APD is operated in Geiger mode, which differs from the linear mode in which APDs operate. In Geiger mode it is possible for a single photon to generate an avalanche breakdown by the impact ionisation mechanism. Such an APD is therefore also termed a Single Photon Avalanche Detector (SPAD). It is mainly used in very weak signal applications discussed in section 1.2. As with standard APDs, SPADs have an internal gain due to this avalanche that will amplify the detected signal [31]. By operating the device in this manner, the internal gain can be increased in comparison to the p-i-n diodes and APDs, which have gains of almost zero and up to several hundred respectively [27]. SPADs have been demonstrated with gains of up to $1x10^6$, which produces a few milliamps of current as short as a few nanoseconds after photon detection. These characteristics, in conjunction with their small size and relatively low required voltage supply, allow for these devices to be used in place of PMTs [23][32].

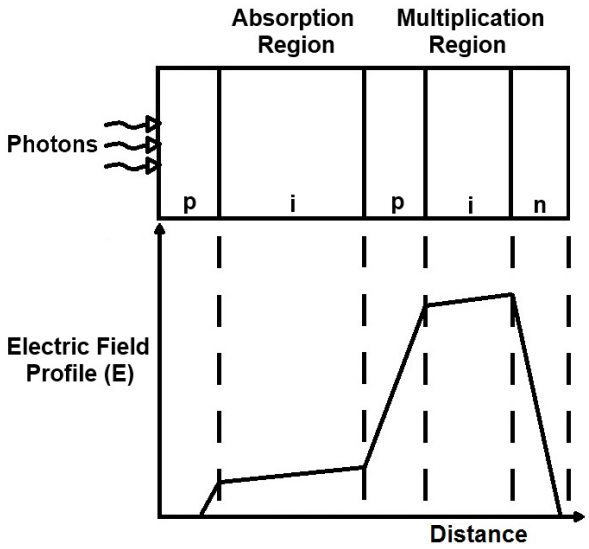


Figure 1.5 - Separate Absorption and Multiplication Avalanche Photo-Diode

Silicon tends to produce better performing devices in Geiger mode than compounds such as Indium Phosphate (InP) or Indium Gallium Arsenide (InGaAs), because of the higher quality crystal and greater maturity of the research. This has made Silicon SPADs the detectors of

choice in wavelengths to which the material is sensitive (600-1000 nm), and there have been reported quantum efficiencies up to 70% at ≈ 700 nm [31][33]. For longer wavelengths such as near infrared, Silicon is ineffective at detecting photons due to the band-gap of the material. Thus near infrared SPADs made with InP, InGaAs and Indium Aluminium Arsenide (InAlAs) are commonly used because they offer a higher quantum efficiency, 20-30% at 1550 nm [35][36], than PMTs [34]. Although this is still much lower than that of Silicon devices at wavelengths shorter than 800 nm, research is keenly being conducted to improve this to meet the need of the 1550 nm window for optical communications, as discussed in section 1.3 [37].

In ultraviolet (UV) and shorter wavelength applications, PMTs have historically been used almost exclusively as there was no alternative detector available. Without the inclusion of highly expensive filters to prevent visible light from becoming incident on the photocathode, PMTs demonstrated poor sensitivity and low quantum efficiency [38]. Silicon Carbide (4H-SiC) has been investigated as a possible replacement as the material demonstrates the ability to detect UV light at power levels in the femto Watt range and offers a low dark current [39]. However, as these devices also require filters to operate within the UV scale [40]. Even with quantum efficiencies over 40% at 280 nm [41][38] they are not an ideal solution. Gallium Nitride (GaN) and Aluminium Gallium Nitride (AlGaN) have similar band gaps to SiC and APDs have been produced that operate in both Linear and Geiger Mode [42]. A common term with nitride detectors is ‘solar-blind’, which means that their absorption terminates at around 280 nm, before entering the visible-light portion of the spectrum and removes the need to use filters in addition to the devices. These materials, GaN and AlGaN, were expected to have a low dark current due to the direct band-gap of 3.4 - 6.2 eV, but devices were reported to have unacceptably high dark currents, possibly due to wafer defects that formed during growth [43].

One of the disadvantages of using a SPAD is that a breakdown event can be caused by a 'dark carrier', which is not generated by absorption of a photon but through thermal or tunnelling processes. Such breakdown events are described as ‘dark counts or events’, to distinguish them from the 'photon counts or events'. The Dark Count Rate (DCR) can be decreased by lowering the device temperature and/or reducing the active area, which both

reduce the dark current [44]. Reducing the electric field will also reduce the dark current and hence DCR, but this will be at the expense of detection efficiency [45].

After-pulsing is another issue for these devices. Carriers of a previous breakdown event can be trapped in the depletion region of the SPAD and then spontaneously released, initiating another avalanche event, even when no photon has arrived. This phenomenon is more common within compound semiconductor devices, though it is not exhibited there exclusively, and limits the operation frequency of near infrared SPADs to a few mega-Hertz [31][46]. The probability of after-pulsing can be reduced by increasing the SPAD temperature, which unfortunately increases the number of dark carriers and hence the DCR. Another method is to allow the trapped carriers in the SPAD to be released before applying the next over-bias pulse to it. This is called 'hold-off' time or 'dead' time, as it is unable to detect a photon in this period. This is achieved through the use of external circuitry that keeps the voltage across the device below that of breakdown so that an avalanche cannot be generated. However, the maximum counting rate remains limited [33][45]. Another alternative is to limit the number of charge carriers flowing during a breakdown event, which will ensure fewer carriers being trapped in the SPAD. This is always implemented, although to different extents, because a SPAD is nearly always operated with a quenching circuit which limits the avalanche current in a breakdown event. The various methods of quenching are discussed in section 1.5 [47].

Research has also been conducted into the statistical analysis of SPAD behaviour. Verilog and other computer simulation software were utilised to produce accurate behavioural models for SPADs during breakdown [48][49]. This work looked at the probability that an avalanche would be caused by a dark carrier, how many carriers would be produced, the likelihood that the carriers may be become trapped and then if they would be responsible for an after-pulse. The effect that temperature has on these interactions was another area examined by the research group [50].

1.5 Quenching Circuits

As explained in sections 1.1 and 1.4.3, quenching circuits are essential for the practical use of SPADs. They are designed to stop the rapidly increasing avalanche current, and this is often achieved by decreasing the over-bias. The different types of quenching circuits include passive, active, mixed passive and active, as well as gated modes, all of which are described in the following sections. The ideal circuit would lower the voltage applied to the device as soon as a photon has been detected, then wait for the avalanche current to diminish to insignificant level before restoring the over-bias to the device. The avalanche current is quenched as soon as it begins to flow, so the number of carriers travelling through the SPAD is low, which in turn limits after-pulsing. This would lead to a short hold-off time, maximising the SPAD's maximum operation frequency. In the case of gated mode operation, the circuit would also not distort the applied over-bias pulses. Each of the circuits reviewed below will have their 'Free-Running Mode' operation examined. This is where the device has the excess-bias applied constantly unless it has detected either a photon or a phonon. This is not always practical, especially for SPADs with a high DCR, as it would be difficult to distinguish between the photon counts from the numerous dark counts.

1.5.1 Passive Quenching

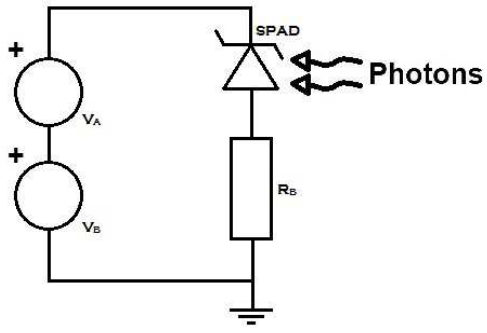


Figure 1.6 - Schematic of a Passive Quenching Circuit

The most basic of the quenching circuits is the Passive Quenching circuit (PQC), depicted in Figure 1.6. This consists of a large 'ballast' resistor, R_B, of 1 kΩ to 100 kΩ typically, placed in series with the SPAD. The entire applied bias is dropped across the SPAD when no avalanche

current is flowing. When an avalanche occurs, significant current will flow through the SPAD and R_B, producing a significant voltage drop across R_B. This decreases the voltage across the SPAD, restricting the current from growing to damaging proportions. The SPAD presents a junction capacitance, C_J, and when connected to the circuit a parasitic or stray capacitance, C_P, that forms between the SPAD and ground. During breakdown, the avalanche current discharges C_P and C_J, forming an RC circuit. Hence, the voltage across the device will exponentially decrease with time with a time constant of $\tau = (C_P + C_J) \times \frac{R_D R_B}{R_D + R_B}$, where R_D is the series resistance of the device, leaving a residual current through the diode. The final value of this current can be approximated using $I_f \approx \frac{V_A}{R_B}$, where I_f is the final value of the current and V_A is the voltage applied above breakdown. Its value is generally < 100 µA because a high I_f will lead to a self-sustaining avalanche, causing the device to be unable to detect further photons.

A large value of R_B will ensure a small I_f. As the current decreases, the number of carriers passing through the SPAD will also decrease. This increases the probability that after some amount of time there will be no carriers remaining in the junction to impact ionise, allowing the SPAD to 'self-quench'. However, a large R_B leads to a longer recharge time constant, or time required for the bias across the SPAD to reach $V_B + V_A$ again, an undesirable quality as it increases the SPAD's dead time [03]. On the other hand, if R_B is too small, the bias across the SPAD will recover so quickly that there will not have been sufficient time to release the trapped carriers, leading to significant after-pulsing. It is important to note that passive quenching does not reduce the bias below V_B, unlike other forms of quenching circuits to be discussed later.

The main factors that influence total charge flow during a breakdown event include the magnitude of V_A, quenching time and the parasitic capacitance. Since V_A can be altered easily, it is not discussed here. The total quenching time can be given as the time taken for the circuit to quench the current and restore the over-bias from the initiation of a breakdown. To reduce quenching time when using a passive quenching circuit, R_B can be increased to decrease I_f so that the quenching threshold, I_S, is reached sooner. The resistor can also be integrated with the device to minimise parasitic capacitance, between the

device and resistor, to a few pF [04], which will in turn reduce the total charge flow. Even with these methods to improve circuit performance, with time constants ≈ 200 ns and total reset times ≈ 1 μs, the reset time is too long for most applications. As a result, use of PQCs is restricted for applications where there is limited space for the quenching circuit [51][52][53]. In pursuing a better alternative to PQCs, research has been conducted into active quenching and other methods of resetting the over-bias voltage by using additional, non-passive components, such as operational amplifiers and transistors [03].

1.5.2 Resetting the Over-Bias Voltage

In the previous section on passive quenching, it was explained that the choice of R_B is a trade-off between quick restoration of the voltage across the SPAD to $V_B + V_A$ after an avalanche event and efficient quenching of the avalanche current. To de-couple this trade-off, a Passive Quenching and Active Reset (PQAR) circuit that incorporates a transistor in parallel with R_B was demonstrated [54]. In a PQAR circuit, after the avalanche current has been quenched, the bias across the device will gradually increase towards $V_B + V_A$, as in a standard PQC. Following the 'hold-off time' the transistor is activated for a few nanoseconds to rapidly restore the operating voltage, allowing this to be completed at a faster rate than in standard PQCs. By using this method it is possible to take advantage of a larger value resistor which will quench the device sooner, without the repercussion of a slow reset. The hold-off time is the amount of time certain quenching circuits 'pause', in order to allow carriers trapped within the diode to be released. This follows the quenching of the device before the over-bias is set across it again, and is often included in the quoted dead time.

1.5.3 Active Quenching

Active Quenching Circuits (AQC), such as the one depicted in Figure 1.7, operate by detecting the change of voltage across the SPAD via a 'sensing resistor', R_S, during an avalanche event. When this happens, the supply voltage to the device is quickly changed from $V_B + V_A$ to below V_B, thus rapidly quenching the avalanche current. After the hold-off time, the bias across the device is increased back to $V_B + V_A$, ready to detect another

incoming photon via the method discussed above [55]. Compared to PQCs, AQCs offer shorter dead-times and a faster voltage reset above V_B as the procedure does not have a recharge time constant. As the hold-off time is directly linked to after-pulsing, there is a compromise in its duration such that the mean release time of the traps is suitably exceeded. If this is not the case, then the probability of after-pulsing greatly increases. This does mean, however, that the dead-time of the device can be well-defined [56].

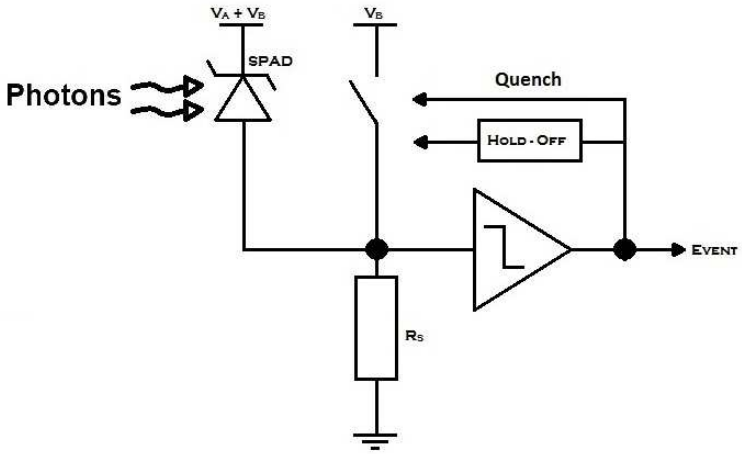


Figure 1.7 - Schematic of an Active Quenching Circuit

The total quenching time of an AQC is not necessarily less than that of a PQC, due to the AQC first having to detect the avalanche event and then react. The total quenching time is affected by the proximity of the detector to the AQC, the peak value of the avalanche current and the sensitivity of the AQC. If too much time elapses from the breakdown event to the response and action of the AQC, then the total charge in the avalanche pulse, Q, will be high [04] as $Q = I_0 \times t \equiv \frac{V_A}{R_B} \times t$, where I_0 is the peak current and t is the total quenching time. Total charge flow can thus be greater in AQCs than in PQCs. To reduce the quenching time, research has been done into Integrated Active Quenching Circuits (IAQCs), where shorter electrical connections between the SPAD and circuit components ensured earlier detection of avalanche events and faster quenching [47][57].

Hold-off time for each SPAD is usually adjusted using information from measurements taken using either Time Correlated Carrier Counting (TCCC) or a Multi Channel Analyser (MCA). TCCC uses trap emission statistics to determine the initial minimum hold-off time, and then

requires the user to manually lengthen this value by repeatedly running the circuit, until after-pulsing reaches a minimum or negligible level. The MCA uses an Analogue to Digital Converter (ADC) in conjunction with autocorrelation to produce a histogram of the SPAD breakdown event. Autocorrelation is best described as the comparison of one signal at two different points in time, and is shown mathematically, $G_{XX}(\tau) = E[X(t_1)X(t_2)] = E[X(t_1)X(t_1 + \tau)]$, *where* $\tau = t_2 - t_1$. By using this technique in combination with the multiple channels of the MCA, it is possible to show the correlation of discrete pulses for a given hold-off time. As with TCCC, the user must then vary the hold-off time so as to best reduce the after-pulsing, which is displayed with greater frequency in earlier channels. Both of these methods are very time consuming and make circuits quite inflexible once they have been calibrated for use with a particular device, unless the process described above is repeated each time the detector is replaced [56].

To overcome this issue, an Auto-tuning Module capable of allowing the user to select the optimal hold-off time for a SPAD in less than twenty seconds was reported [58]. For each run, the device is activated and quenched repeatedly using an AQC, with the outputs of the circuit fed into a Field Programmable Gate Array (FPGA) as well as an MCA to record the histogram. The FPGA increases the hold-off time from the previous value after each 'run'. This continues until after-pulsing is no longer detected during the hold-off time, to yield the optimal hold-off time.

1.5.4 Mixed Active-Passive Quenching

Mixed Active-Passive Quenching Circuits (MAPQCs) combine passive and active quenching to overcome some of the limitations that the circuits have individually. A MAPQC tends to be an AQC with the large resistor in series with the SPAD, as shown in Figure 1.8. If an avalanche current is building up, there will be a voltage drop across R_B, which restricts the growth of avalanche current, achieving 'quasi-quenching'. Tens of nanoseconds later, the comparator will sense that an avalanche event has taken place, so will lower the voltage applied to the SPAD to below V_B. The remainder of the process then takes place as in an AQC, with the diode restored to its quiescent state after a suitable hold-off time by the

'active loop' [47]. By including the quenching resistor in the AQC, the total charge in each avalanche pulse is greatly reduced since the avalanche current is restricted sooner than it would have been in standard AQCs [59].

One disadvantage of AQCs that MAPQCs do not rectify is the reset of the over-bias, which still takes place by the Active Reset mechanism to avoid the slow operation seen in PQCs. However, this leaves the device unprotected if another photon is absorbed during resetting as the circuit is unable to react to a second event until resetting has been completed. Integration of the circuit with a SPAD has the same advantages seen in PQCs and AQCs, such as the reduction in parasitic capacitance due to shorter connections between diode and circuitry which helps to shorten the reset time. With shorter reset time, total charge from events triggered during recharge is reduced, as is the probability that events will occur during this time. However, these benefits come at the expense of occupying an area as large as a few square millimetres, which makes the design unsuitable for use with arrays [04].

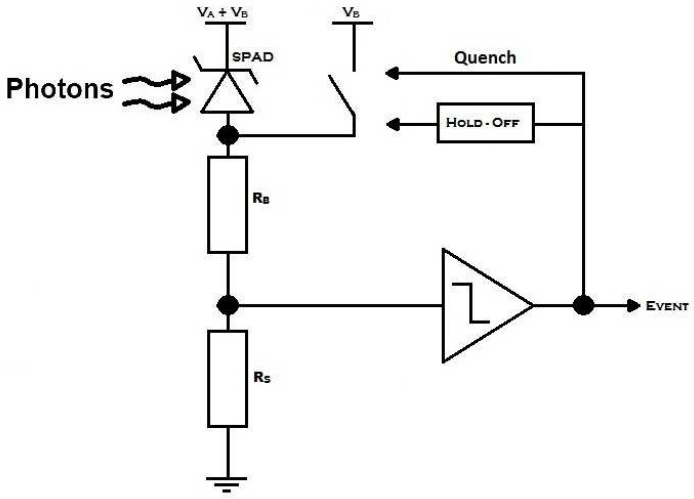


Figure 1.8 – Schematic of a Mixed Active-Passive Quenching Circuit

Thus far, all of the AQCs discussed use active components to restore the over-bias across the device, but this is not the only method. Some circuits use active components for quenching and passive components for reset only. It is important to point out that the resistor, R_B, is intended to be used for restoring the voltage across the SPAD and not for quenching, otherwise many of the disadvantages of PQCs would become present in this design. For this reason, the value chosen for R_B can be smaller to reduce the time taken for

recharging the SPAD. Although not as rapid as active reset, with times quoted from 10-20 ns compared to a few nanoseconds, the passive part removes the overshoot and ringing present in active recharge [60]. Another approach has been examined using an active driver to raise the voltage close to, but not reaching, the desired level before deactivating; allowing the passive reset to take over and restore the remaining bias much more smoothly [03].

1.5.5 Gated Mode

Gated Mode (GM) is often the favoured method of operation for SPADs that have high DCRs and/or severe after-pulsing problems, due to poorer crystal quality and/or narrower band gaps, such as InGaAs/InP SPADs. Operating the SPAD in gated mode is the most effective way to reduce these undesirable effects, especially with very short over-bias pulses. This is because when an avalanche event happens, the SPAD is only allowed to conduct for a very short period of time, as the end of the over-bias pulse will quench the avalanche current, limiting the charge generated by the avalanche. The over-bias pulses are synchronised to the expected arrivals of photons so that the device is more likely to break down due to photon detection than it is to dark carriers [61].

In GM the SPAD is reverse-biased with a DC voltage smaller than V_B, which is superimposed with AC over-bias pulses whose widths range from less than 100 ps to tens of nanoseconds. The AC over-bias pulse height should be sufficient to exceed V_B. If very narrow over-bias pulses are used, then a quenching circuit is not needed for the SPAD [62]. In fact, if a passive quenching circuit is used in conjunction with GM and very narrow pulses, the resistor, R_B, could distort the over-bias pulses to the extent that avalanche breakdown becomes unlikely. This is because DC coupling would form a low pass filter with C_P and R_B, significantly slowing down the pulse edges. AC coupling however, would form a high pass filter, meaning that the pulse would reach a maximum value and begin to reduce in magnitude [04].

High-speed gating also produces current spikes that flow through the SPAD on the leading and trailing edge of the gates, due to charging and discharging of the junction capacitance of

the SPAD. These capacitive transient pulses are large enough to prevent detection of the avalanche pulses. Unfortunately, increasing the over-bias pulse height will increase both the height of the capacitive transients and the avalanche pulses. A better way to deal with the capacitive transients is to attempt to minimise them using cancellation techniques [37], such as coaxial cable reflections and capacitive balancing.

1.5.6 Coaxial Cable Reflections

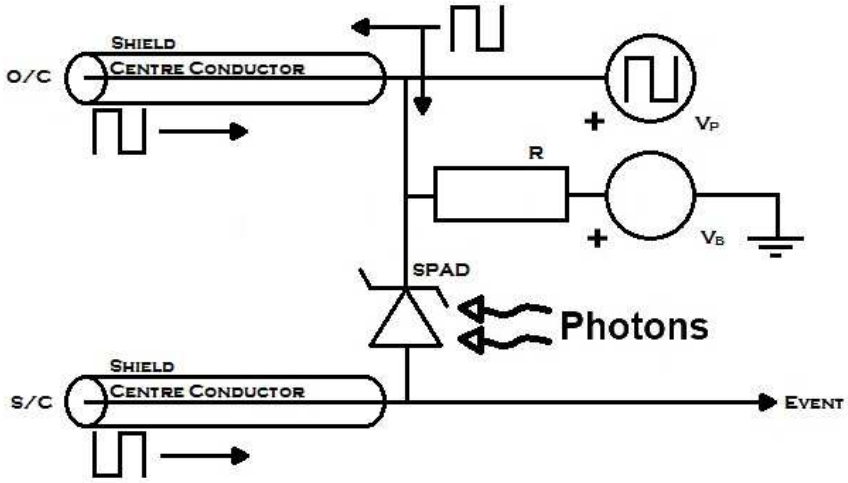


Figure 1.9 – Schematic of a Coaxial Cable Reflection

Capacitive transients can be minimised by using the coaxial cable reflection cancellation technique depicted in Figure 1.9, where the SPAD has an open-circuit coaxial cable connected to its Cathode and a short-circuit coaxial cable to its Anode. When the over-bias pulses of pulse height V_P are applied to the SPAD, the capacitive transient due to the leading edge of the over-bias pulse travels both through the device and along the open-circuit cable. The signal that travels into the device then travels both along the short-circuit cable and to the output of the circuit. The pulses travelling along the coaxial cables are reflected once they reach their relative terminations, with the short-circuit termination inverting the reflection and the open-circuit termination producing a non-inverted reflection. The time taken for the signals to travel back along the coaxial cables is carefully matched with the falling edge of the pulse, such that the second pulse combines with the signal from the open-circuit coaxial cable, passes through the SPAD and is then cancelled by the inverted signal from the short-circuit coaxial cable [63].

This method leaves one spike caused by the capacitance of the device which, providing the magnitude of it is known, can show the detection of a photon even if the interaction occurs as the pulse passes through the device. However, the coaxial cable length must to be carefully matched to the pulse width, and varying the width or duty cycle of the over-bias pulses will require a change in the coaxial cable length.

1.5.7 Capacitive Balancing

Another way of removing capacitive transients from the SPAD signal is to use a capacitive balancing scheme. This involves using two 'identical' SPADs so the capacitance will be almost the same value in both components. Although it is unlikely to obtain two identical SPADs because SPADs are extremely sensitive to minute variations in thickness and doping levels, those from the same batch are likely to give the electrically similar characteristics [64]. Over-bias pulses, superimposed onto a DC bias, are applied to both SPADs, with the outputs of each connected to a 180° hybrid ring coupler.

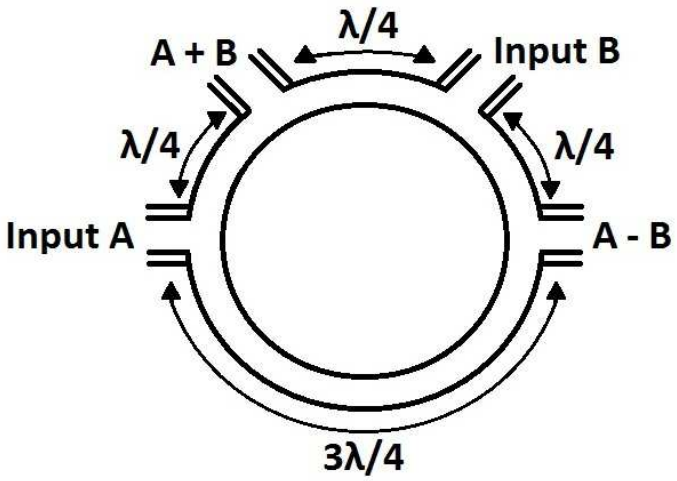


Figure 1.10 - Schematic of a 180° Hybrid Ring Coupler

Shown in Figure 1.10 this is a four port network, ringed in shape, which has all four ports on one side each separated by *λ/4*, with the remaining port-less side *3λ/4* in length. Depending on how connections are made, takes one input and produces from it two outputs, either both in phase or 180° out of phase with one another; or takes two inputs and provides from them two outputs, with one output the sum of the inputs, and the other output the

difference between the inputs [65]. The circuit is connected so that the two SPADs provide the inputs to the hybrid coupler, which has its output ports connected to two discriminators. The transients produced by the over-bias pulses are common to both inputs and they are removed before they reach the discriminators. This means the avalanche pulse will be easier to detect amongst the noise, without the unwanted spikes, so it is possible to reduce the threshold voltage in the discriminators with the application of the hybrid coupler [64].

This method effectively cancels the capacitive transients from the SPAD signal, with superior performance compared to the coaxial cable reflection cancellation technique. However, the hybrid coupler is also inflexible to changes in pulse width, as the circumference of the ring is chosen based on the wavelength of the signal travelling through it. As before, if the pulse-width or duty-cycle is changed then a hybrid coupler of different dimensions would be required. A more flexible approach is to use an Instrumentation Amplifier, also known as a Differential Amplifier, which behaves in a similar manner to the hybrid coupler, but does not require modifying when the gating duration is changed.

1.5.8 Gated Mode and Quenching Circuits

As the duration of each over-bias pulse is increased, it provides a longer period for the SPAD to detect a photon and it becomes necessary to quickly quench the avalanche current before the end of the over-bias pulse [66]. A Gated Active Quenching Circuit (GAQC) can be used as it behaves similarly to the AQC for free-running mode. A Gated Passive Quenching Circuit (GPQC) is an alternative method, but if gating is applied to the SPAD by superposition on the DC bias, the low pass filter formed between the quenching resistor and the SPAD's capacitance, C_P and C_J, causes significant distortion of both rising and falling edges of the gate. This reduces the speed that the over-bias pulses reach their final values by as much as microseconds, and still affects quenching and reset times, as seen in section 1.5.1 [67].

Applying the over-bias pulses separately with a ‘bias-tee’ through a capacitor, C_g, forms a high pass filter with the rest of the circuit, so once the gate reaches its maximum value it

decays until the end of the pulse. This imposes a limit on the maximum duration of the pulse and removes the 'flat top' from the wave. The additional capacitor also causes capacitive division of the applied gate voltage as it is connected in parallel to C_P and C_J. This is particularly noticeable if C_g is roughly equal to the device capacitance as, $V_o = V_i \times \frac{C_g}{(C_g+C_p+C_J)}$. To overcome this, a C_g much larger than $C_P + C_J$ is desirable, but it is worth noting that the avalanche current is only quenched before the gate ends by recharging C_g, so a value should be chosen with care [03].

For the rectangular pulses used in the GM examples so far, capacitive transients in the SPAD signal are unavoidable. Although some methods of removing them were discussed above, they still greatly increase the total charge for each avalanche event. As, $I = C\frac{dV}{dt}$, the larger the over-bias pulse and the faster it is applied, the larger this capacitive transient is. To avoid capacitive transients completely, Sinusoidal Gating, which does not involve a large change in voltage over a very small duration, has been reported. [68] This was shown to reduce the charge per avalanche event from ≈10^7 to ≈10^5, reducing the probability of after-pulsing and allowing the repetition rate of the gates to be increased [69].

1.5.9 Capacitive Quenching

Another way to overcome the capacitive response of rectangular pulses is to use a Capacitive Quenching Circuit (CQC) [05]. This design uses a capacitor to quench the avalanche current instead of the resistor seen in standard GPQCs. By doing this, the total charge for each interaction is fixed to the value of the capacitor, C_B, offering a lower total charge flow than previous circuits and a reduced after-pulsing probability. The lack of a large resistance in series with C_P removes the filtering effects shown in this section and 1.5.1, allowing undistorted over-bias pulses to be applied to the SPAD. Another advantage is that the circuit only utilises passive components, meaning there is no concern about a delay in sensing and reacting to an avalanche event as seen in GAQCs.

1.6 Aims, Objectives and Thesis Overview

The aim of this Thesis is to demonstrate that the CQC is a viable alternative to the other quenching circuits currently used. Chapter 1 is the introduction to the work, comprising of an extensive literature review. Section 1.4 categorised and discussed the different types of single photon detection in detail, highlighting their strengths and weaknesses. From this, an informed choice was made on the area that warranted further investigation, which was the SPAD. As these devices are commonly operated in conjunction with a quenching circuit, section 1.5 compared and contrasted various designs to show which had the stronger performance. The remainder of this Thesis will be concerned with the design, implementation and testing of the CQC, due to the advantages it offers over the alternative circuits.

Chapter 2 is the modelling section, where the CQC was simulated using Linear Technology's Simulation Program with Integrated Circuit Emphasis (LT SPICE), a piece of computer software designed to model the behaviour of electronic components. This allowed the circuit design to be separated into three main sections, with each examined and tested within the program to confirm correct operation. When this was completed the entire design was simulated, with key components' values being altered to ensure the most suitable choice had been made.

Chapter 3 examines the steps that were taken to design, fabricate and test a Printed Circuit Board (PCB) once satisfied with the simulation results. With high speed operation, the layout and overall size of the PCB is hugely important. A pair of computer programs named Intelligent Schematic Input System (ISIS) and Advanced Routing and Editing Software (ARES) were used to create the circuit schematic and then place the components onto a virtual PCB, which was optimised before being submitted for fabrication. Chapter 3 also describes the tests conducted on the circuit and compares the data with the simulation results from Chapter 2 and the data reported in [05].

Chapter 4 draws conclusions on the work carried out and offers recommendations for further work on this topic.

Chapter Two - Modelling

This chapter will examine the three key sections of the CQC, describing first the predicted operation based on theory, before discussing the results from modelling them, and the entire circuit, in LT SPICE. The theoretical and modelled values will be used in Chapter 3 of this Thesis to compare with the behaviour of the physical circuit.

2.1 Details of the Circuit

Figure 2.1 is a circuit diagram of the CQC. This is operated in GM, a type of operation discussed at length in sections 1.5.1, 1.5.5 and 1.5.8. A negative bias, V_{DC}, is applied to the Device Under Test (DUT), at a magnitude lower than V_B. This is applied across both R_1 and D_1, which limit the current in the case that V_{DC} is too large or of incorrect polarity, respectively. The output of the CQC was observed exclusively from the point labelled "Out" until the addition of the Active Probe in section 2.3.

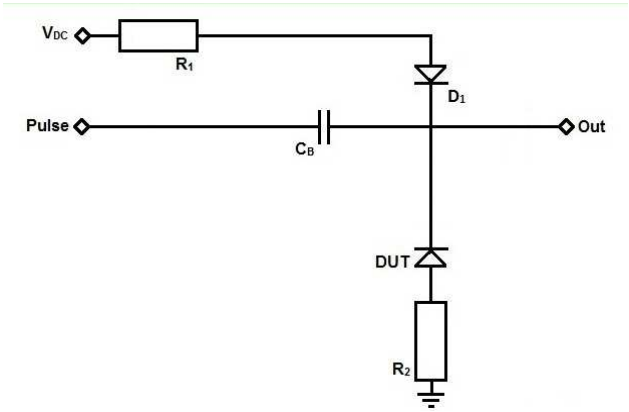


Figure 2.1 – Schematic of the Capacitive Quenching Circuit

Illustrated in Figures 2.2(a) and (b) are sketches of the voltage-time waveforms that depict the expected output of the CQC without and with an avalanche event, respectively. On initial application of the negative bias, without the presence of a pulse, C_B charges and the voltage across it becomes equal to V_{DC}, as seen in Region 1. It should be noted that this only

occurs when the circuit is first connected to the DC voltage, and does not take place on each duty cycle. In Region 2, a pulse is applied through C_B across the device to some level V_A above V_B, with both of these stages common to Figures 2.2(a) and (b). Should no avalanche event take place, the voltage level will remain at V_A+V_B across the DUT until the end of the pulse, when it will return to V_{DC}, as depicted in Regions 3 and 4 of Figure 2.2(a).

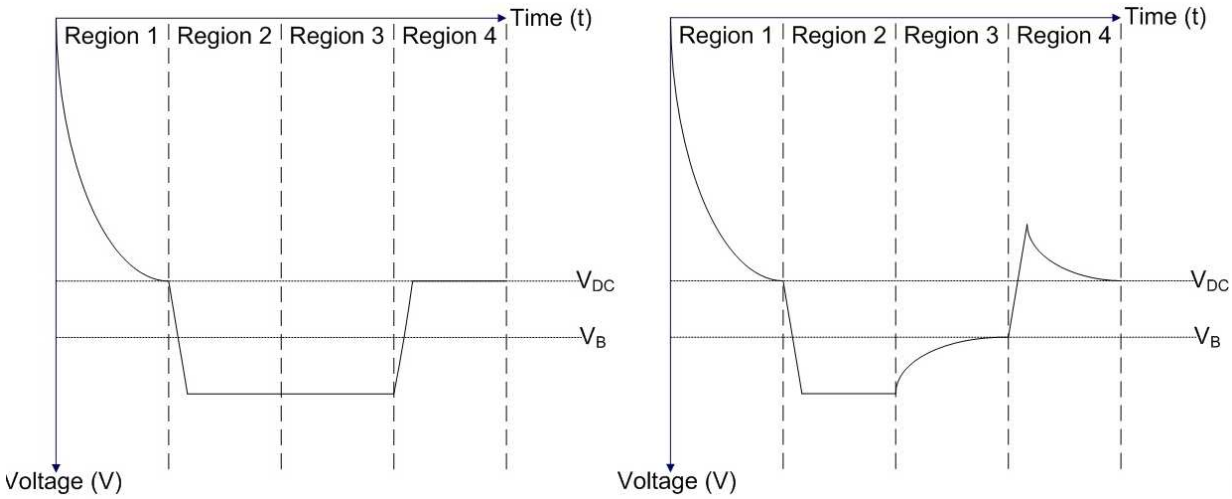


Figure 2.2 - A Sketch of the Output Voltage-Time Waveform of the CQC (a) Not Experiencing an Avalanche Event and (b) Experiencing an Avalanche Event

If an avalanche event occurs while the pulse is applied, as in Region 3 of Figure 2.2(b), there is an exponential decay across the DUT as C_B, C_P and C_J discharge and current flows through the DUT into C_B. The time taken for the voltage drop to reach V_B, or the quenching time, is mainly dependent on these capacitances. It is worth noting that the delay between the application of the pulse and the avalanche event shown in Region 3 of Figure 2.2(b) was included for explanatory purposes. In reality avalanche breakdown is a spontaneous event that may occur at any time during the application of the pulse. At the end of the pulse the voltage settles at V_{DC}, and the capacitance discharged in quenching the avalanche event is restored by the DC bias. This is shown in Region 4 of Figure 2.2(b), where the recharging time constant is $\tau \approx R_1(C_B + C_P + C_J)$.

To aid in taking measurements from the DUT additional components are often used in conjunction with it and the CQC. Figure 2.3 depicts one possible configuration which is

composed of three main component parts - the CQC, the Active Probe and the Differential Amplifier. The modelling process was simplified by first examining only the CQC, adding the Active Probe and Differential Amplifier later in sections 2.3 and 2.4 respectively. By taking this approach, correct operation and the effect of varying component values could be investigated without the outcome being masked or altered by supplementary circuitry.

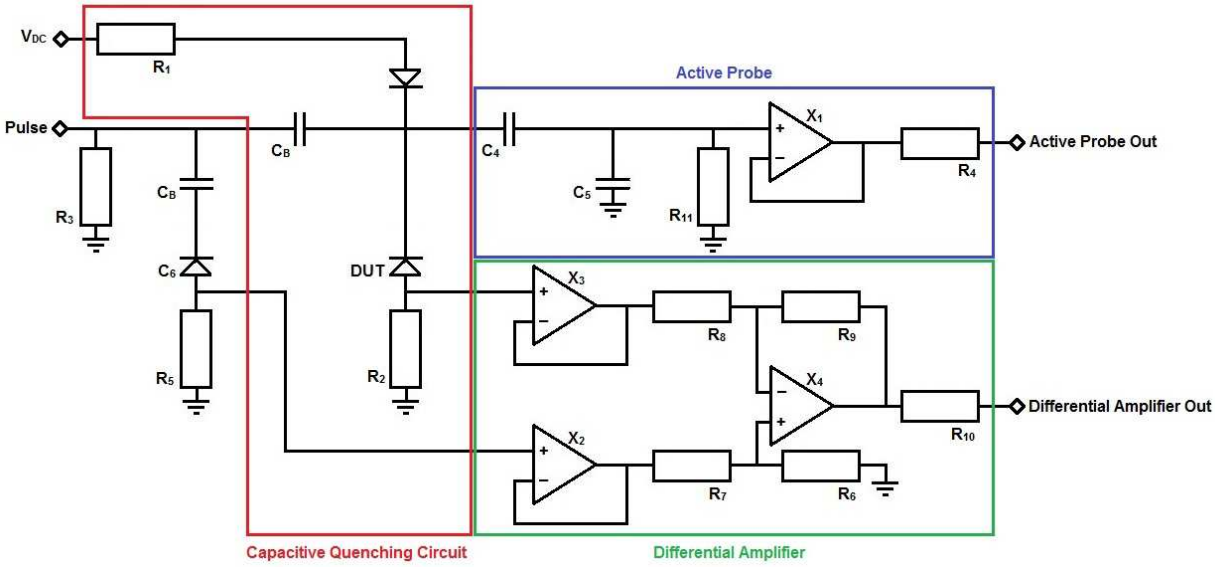


Figure 2.3 - Measurement Circuit Configuration Schematic

2.2 Modelling the CQC

To confirm the predicted operation, the circuit was modelled using LT SPICE, a circuit simulation software. The software did not have a component model for a SPAD, so a Zener diode [70] was used as the DUT instead. As pulses were applied across the Zener diode, the output characteristic exhibited similarities to the waveforms that are seen during the charge and discharge of a capacitor. This suggested that there was a large capacitance present in the circuit that was misshaping the pulses. A brief investigation revealed that the source of this capacitance was the Zener diode as the its default value of C_J was 150 pF, when the C_J of a typical SPAD [71] was only 2 pF. It was possible to open the model file [72] and manually change this value to 2 pF, thereby solving the problem and producing the desired shape of the pulses. Figure 2.4(a) shows the voltage output across the CQC and modified DUT, from 0 V to -8 V, when -8 V pulses were applied. -8 V was chosen for the magnitude of the pulses as

a large number of values could have been selected and -8 V allowed direct comparison with *Dimler et al.* R_1 and R_2 were taken to be 1 kΩ and 50 Ω respectively, with C_B being 27 pF. This demonstrates a good qualitative fit between modelled data in Figure 2.4(a) and expected data in Figure 2.2(a). The reduction in absolute pulse height is due to stray capacitance in the circuit. As LT SPICE is not equipped with a default SPAD model and additional modelling methods to represent an avalanche event were found to be too complex and time consuming, a Zener diode was used in place of a SPAD.

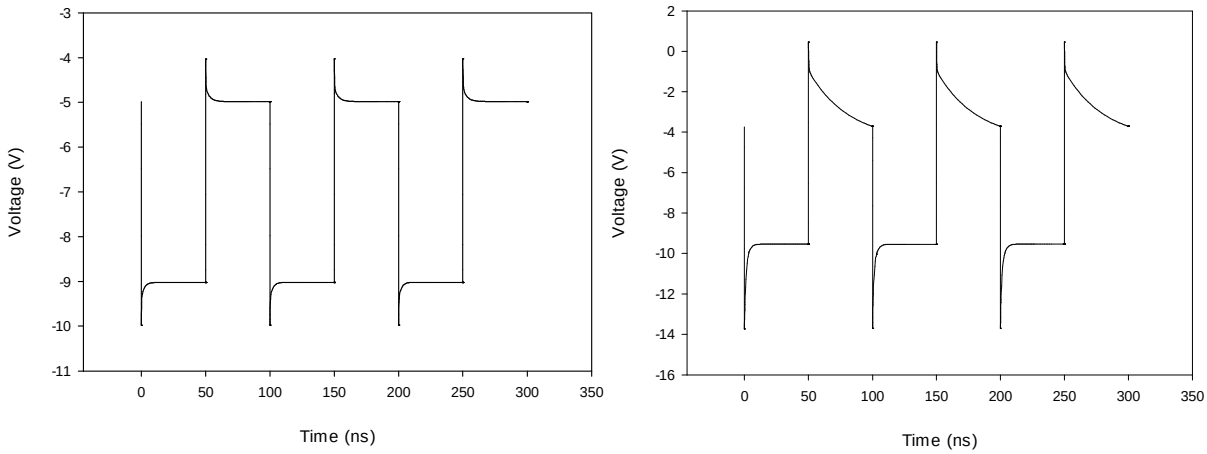


Figure 2.4 - LT SPICE Output Waveform of the CQC (a) When -8 V Pulses Were Applied Without a DC Bias and (b) When -10 V Pulses Were Applied In Addition to a -5 V DC Bias

Figure 2.4(b) shows the effect on the output voltage when -10 V over-bias pulses were applied in addition to a -5 V DC bias across the CQC and DUT over a range of 0 V to -14 V. The Zener diode has V_B = -10 V, when the applied voltage exceeds this threshold value the DUT experiences Zener breakdown [73][74]. This drives the DUT into conduction and the voltage falls to V_B. -10 V pulses with a -5 V bias were selected to ensure that the Zener diode exceeded its breakdown voltage, although other variations of values could have been used. Circuit behaviour from the avalanche event to recovery can be seen in Figure 2.2(b), the expected voltage-time waveform during an avalanche event. In conclusion, Figure 2.4(b) demonstrates the effect that DUT conduction has on the circuit and demonstrates that it is able to react accordingly during an avalanche event.

Following this, the optimal value for C_B was investigated. If too small a value is chosen, the avalanche event will be quenched before the avalanche current becomes large enough to be easily detected. Although this allows the achievement of a short response time and a high operation frequency, the avalanche pulse can be hidden amongst transient voltage spikes generated by the application of over-bias pulses. These are mentioned in 1.4.6 and 1.4.9 and will be discussed further in section 2.4, and can both mask avalanche events and cause false counts to take place. On the other hand, if the chosen value for the capacitor is too large, the response time and operation frequency will be reduced and the avalanche current possibly allowed to grow too large before quenching, damaging or destroying the device.

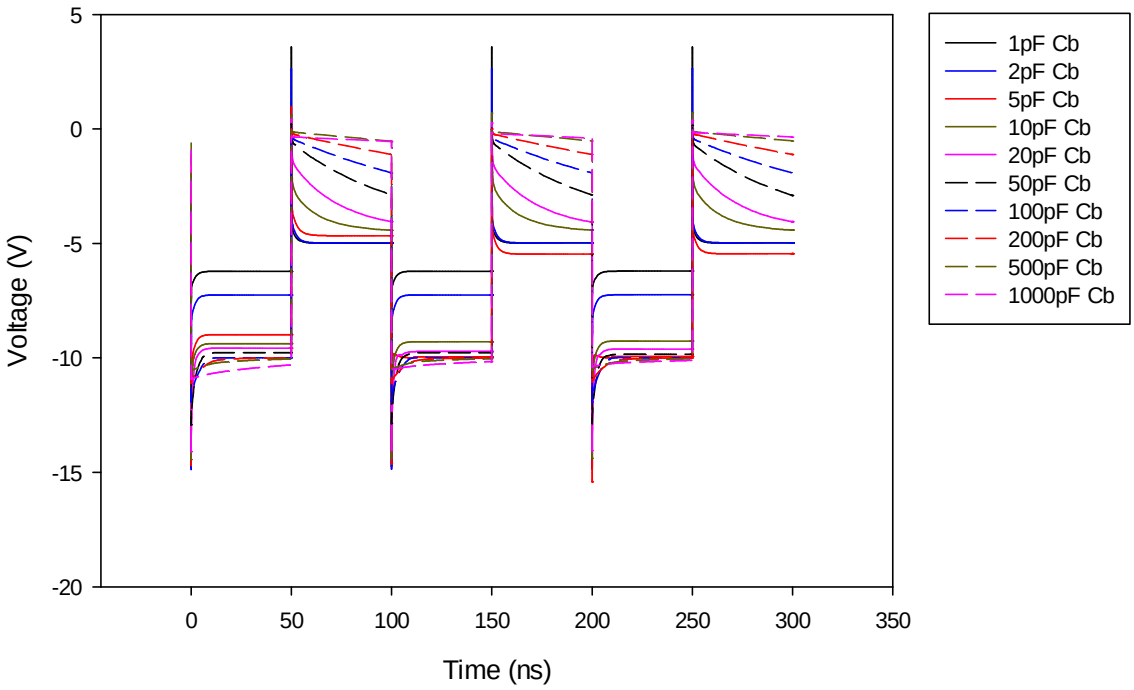
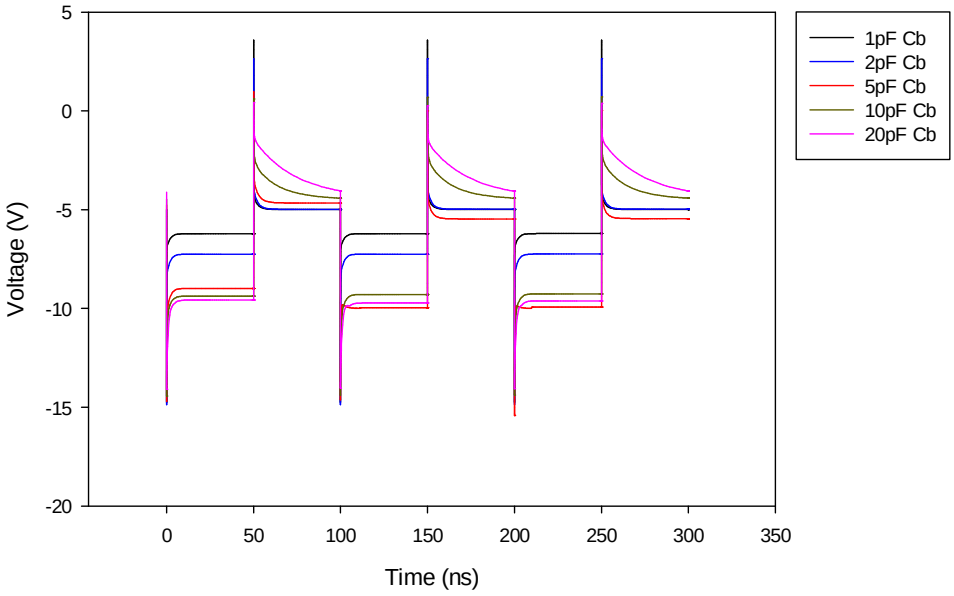


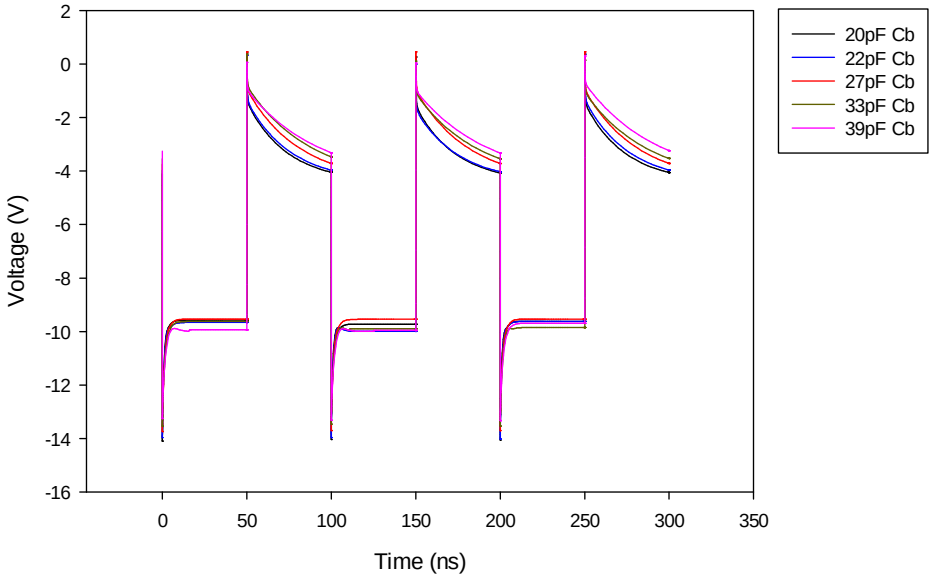
Figure 2.5 - LT SPICE Output Waveforms of the CQC When Ranging C_B From 1 pF to 1000 pF With Over-Bias Pulses of -10 V Applied In Addition to a -5 V DC Bias

To gain an overview of the effect that altering the magnitude of C_B had on the circuit output its value was increased from 1 pF to 1000 pF logarithmically, as shown in Figure 2.5, with over-bias pulses of -10 V applied across the CQC and DUT in addition to a -5 V DC bias. These magnitudes were chosen in preference to the -8 V pulses used earlier as they accentuated the effect that C_B had on the output. The values of R_1, R_2 and other components remained unchanged unless stated otherwise. As predicted, as the value of C_B was increased, response to a break down event was reduced and recharge time was increased, becoming particularly

noticeable above 50 pF. When the value of C_B was set below 20 pF, in addition to the quenching time being too short, the recharge of the circuit may take place too quickly, increasing the probability of after-pulsing, discussed in 1.3.6. Figure 2.6(a) shows the values of C_B only between the values of 1 pF and 20 pF for added clarity.



(a)



(b)

Figure 2.6 - LT SPICE Output Waveforms of the CQC (a) When Ranging C_B From 1 pF to 20 pF and (b) When Ranging C_B From 20 pF to 39 pF All With Over-Bias Pulses of -10 V Applied In Addition to a -5 V DC Bias

Using a selection of standard capacitor values [75], the magnitude of C_B was increased from 20 pF to 39 pF, as depicted in Figure 2.6(b). As the output of the CQC remained fairly consistent with each magnitude any of these values of C_B would be suitable. 27 pF was selected due to its availability.

2.3 Modelling the Active Probe

Once the CQC was behaving as expected and the optimum value for C_B had been chosen, the Active Probe was added to the model. This part of the circuit is included for monitoring the over-bias pulses within an external piece of equipment, such as an oscilloscope, without additionally loading the CQC. It is composed of an Operational Amplifier (OPA2356) [76], X_1, a resistor, R_{11}, and two capacitors, C_4 and C_5. The Op-Amp chosen was a high speed device, able to tolerate the fast pulses applied across the DUT. C_4 and C_5 were used to form a capacitive potential divider which acted to ensure that the magnitude of the pulses did not exceed the maximum input value for the Op-Amp. As $V_O = V_I \times \frac{C_4}{C_4 + C_5}$, values were initially chosen of 1 pF and 47 pF respectively to give a ratio ≈ 1:50, affording some protection to the Op-Amp should larger than expected pulses mistakenly be applied. The voltage source to X_1 is grounded through R_{11}, which combined with C_4 and C_5 forms a high-pass filter. A value of 100MΩ was chosen for R_{11} such that the cut-off was ≈80 kHz, preventing the transmission of fluctuations to the DC supply.

It is worth noting that the capacitive divider equation only holds true when the Active Probe is considered independently. Once used in conjunction with the CQC, C_B is then in series with C_4 which causes the equation to become $V_O = V_I \times \frac{\left(\frac{C_B \times C_4}{C_B + C_4}\right)}{\left(\frac{C_B \times C_4}{C_B + C_4}\right) + C5}$. For the chosen value of C_4, 1 pF, the effect this has is limited as $\frac{(27 \times 1)}{(27+1)}$ is ≈ 1. By completing the calculation the multiplication factor is found to be $\frac{1}{49.74}$, giving an expected peak output value of -0.1608 V when the applied over-bias pulses are -8 V. The Multiplication Factor is the resultant fraction by which the input voltage, V_I, is multiplied to produce the seen output voltage, V_O.

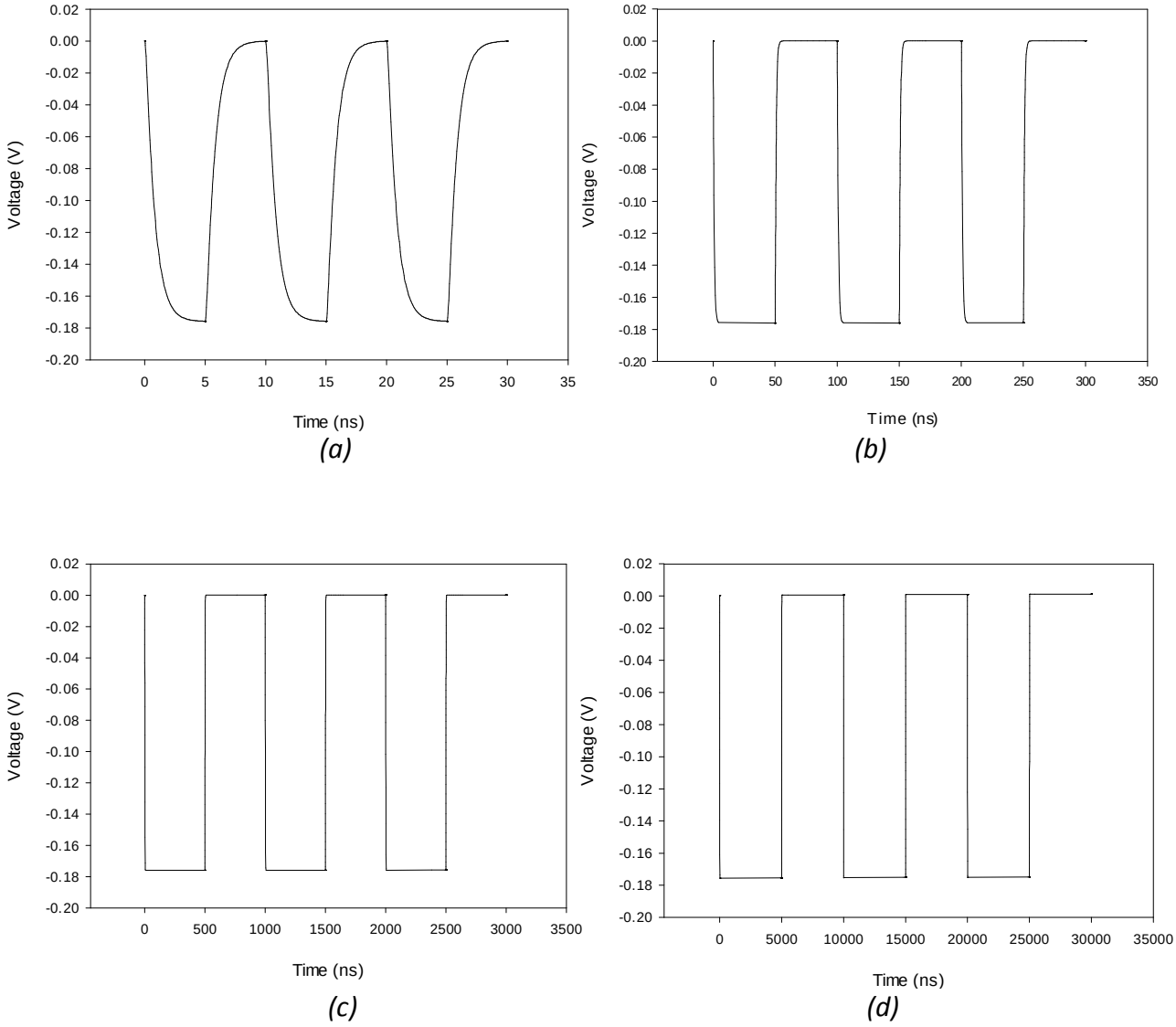


Figure 2.7 - LT SPICE Output Waveforms of the Active Probe During the Application of -8 V Pulses with (a) 5 ns, (b) 50 ns, (c) 500 ns and (d) 5000 ns Pulse Width

-8 V pulses with widths of 5 ns, 50 ns, 500 ns and 5000 ns were applied across the CQC so that the effect of the varying duration could be observed as well as the overall behaviour of the Active Probe, as displayed in Figures 2.7 (a), (b), (c) and (d). These measurements were taken from the point "Active Probe Out", shown in Figure 2.3, until section 2.4 when the Differential Amplifier was examined. The pulse height measured on the output had reduced to ≈ -0.16 V, showing the capacitive division was behaving as expected. The maximum value of these pulses was -0.1546 V in the raw data, giving a multiplication factor of $\frac{1}{51.75}$. This is a comparable value to the expected multiplication factor calculated earlier and confirmed

that the correct operation was being observed. Figures 2.7 (b), (c) and (d) display square pulses with flat tops, the desired output, whereas Figure 2.7 (a) shows a less ideal square wave, with the capacitances in the circuit rounding the leading and trailing edges of the pulses. This is most likely a result of the fast, short pulses applied across the circuit, demonstrating that the maximum operating frequency is being approached. Both of the expected and modelled values will be compared with values taken from the real Active Probe in sections 3.2 and 3.4.

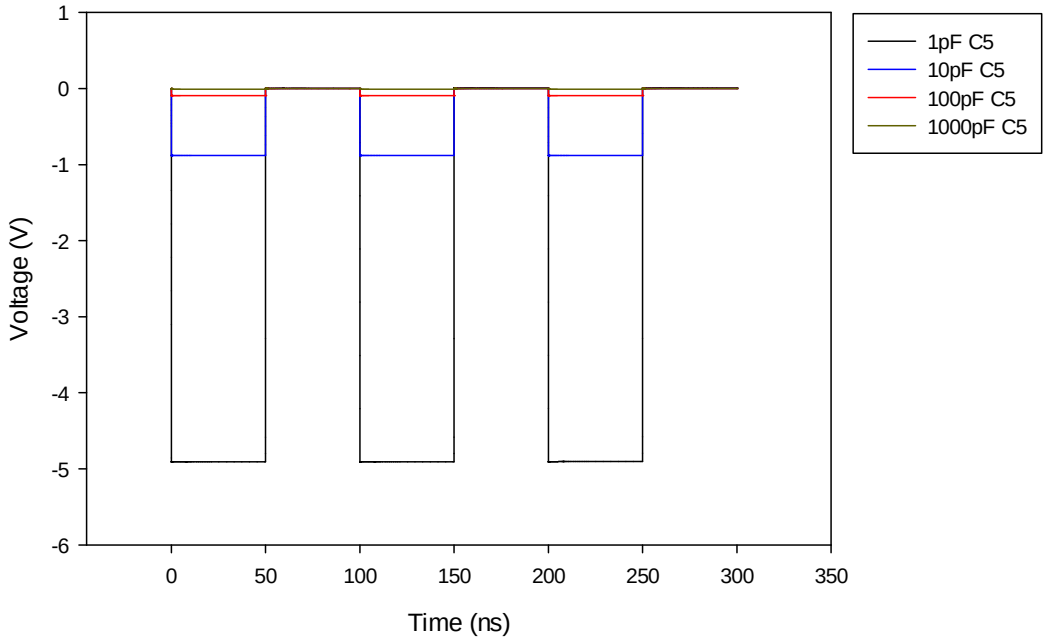


Figure 2.8 - The LT SPICE Active Probe Output for C_5 = 1, 10, 100 and 1000 pF During Applied Voltage Pulses of -10 V

C4	***C5***	***Expected Output (V)***	***Multiplication Factor***
1pF	1pF	-4.91	1/2.04
1pF	10pF	-0.88	1/11.37
1pF	100pF	-0.10	1/104.70
1pF	1000pF	-0.01	1/1038.04

Table 2.1 - Expected Output Voltages (V_O) And Associated Multiplication Factors for C_5 = 1, 10, 100 and 1000 pF During Applied Voltage Pulses of -10 V

To view the effect different magnitudes of C_4 and C_5 imposed on the Active Probe Output, their values were stepped through a range of capacitances, 1, 10, 100 and 1000 pF whilst pulses of -10 V were applied. Figure 2.8 shows the LT SPICE outputs when C_5 was increased and C_4 fixed at 1 pF. By increasing C_5 the denominator becomes larger, dominating the resultant of the fraction, and the output voltage, V_O, decreases proportionately. Table 2.1 displays the values of V_O and the associated multiplication factors that were calculated using the above equation to verify the LT SPICE output and to compare with previous values for the Active Probe.

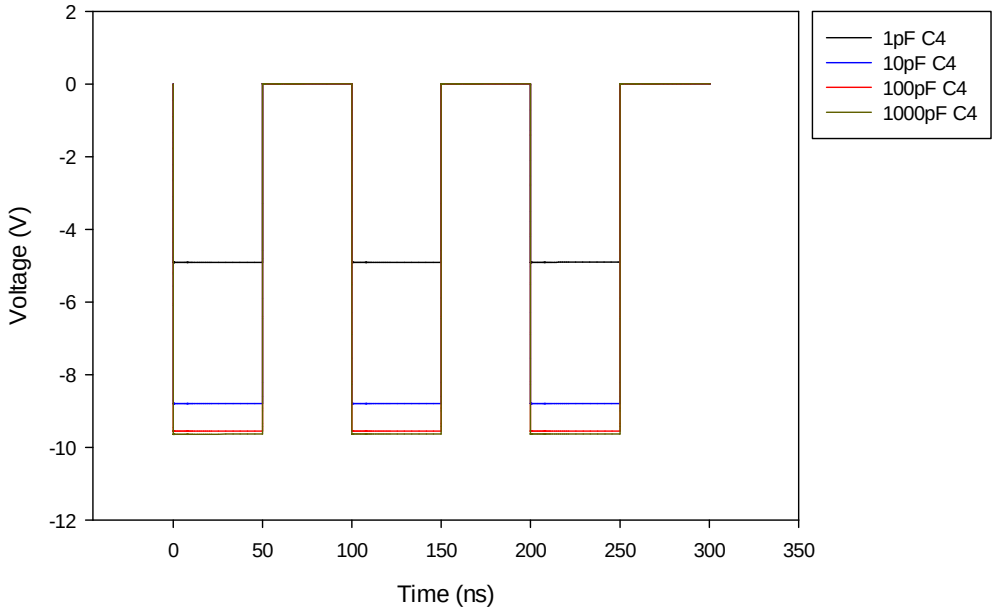


Figure 2.9 - The LT SPICE Active Probe Output for C_4 = 1, 10, 100 and 1000 pF During Applied Voltage Pulses of -10 V

C4	***C5***	***Expected Output (V)***	***Multiplication Factor***
1pF	1pF	-4.91	1/2.04
10pF	1pF	-8.79	1/1.14
100pF	1pF	-9.55	1/1.05
1000pF	1pF	-9.63	1/1.04

Table 2.2 - Expected Output Voltages (V_O) And Associated Multiplication Factors for C_4 = 1, 10, 100 and 1000 pF During Applied Voltage Pulses of -10 V

C_4 then had its value incremented through the same magnitudes as C_5 had previously, this time with C_5 kept at a constant value of 1 pF. After the initial step, C_4 became much larger than C_5 meaning that the numerator and denominator became a closer and closer approximation to 1, with C_5 having minimal effect on either the value of the denominator or V_O. The LT SPICE output waveforms are shown in Figure 2.9 with the calculated values for both V_O and the multiplication factors shown in Table 2.2.

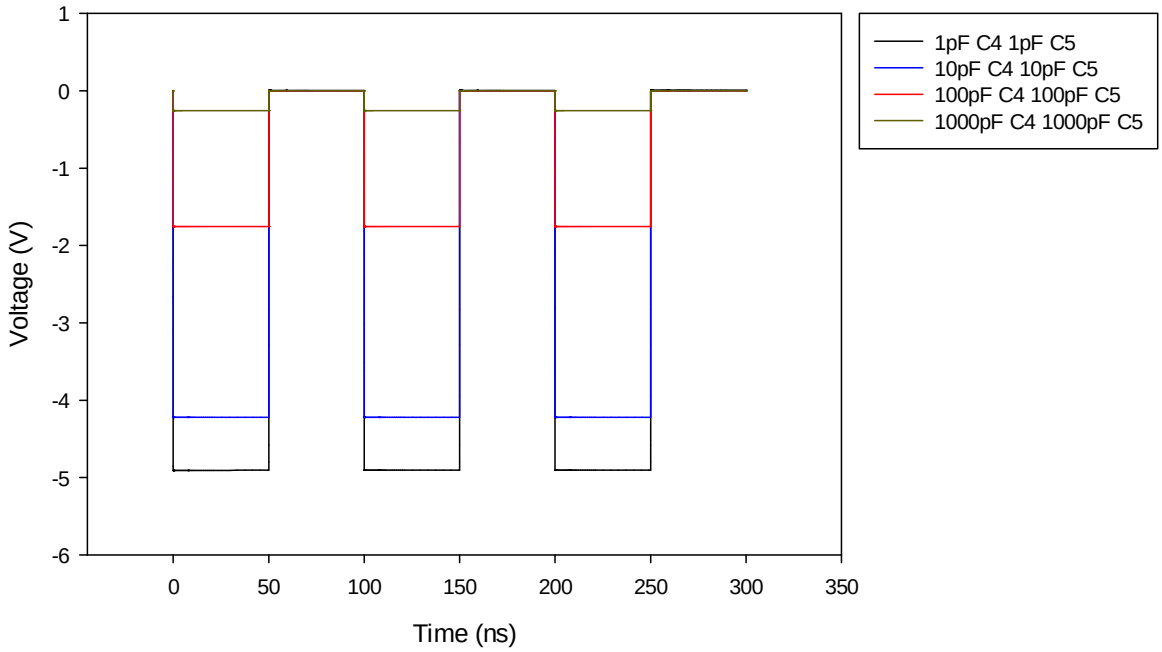


Figure 2.10 - The LT SPICE Active Probe Output for C_4 and C_5 = 1, 10, 100 and 1000 pF During Applied Voltage Pulses of -10 V

C4	***C5***	***Expected Output (V)***	***Multiplication Factor***
1pF	1pF	-4.91	1/2.04
10pF	10pF	-4.22	1/2.37
100pF	100pF	-1.75	1/5.70
1000pF	1000pF	-0.26	1/39.04

Table 2.3 - Expected Output Voltages (V_O) And Associated Multiplication Factors for C_4 and C_5 = 1, 10, 100 and 1000 pF During Applied Voltage Pulses of -10 V

Finally, C_4 and C_5 were increased simultaneously, which resulted in the denominator being increased faster than the numerator, as C_5 was incremented at the same rate that C_4 was. This produced a smaller V_O each time the magnitudes of C_4 and C_5 were incremented but with a gradual decrease, unlike that seen in Figure 2.8(a). Figure 2.10 and Table 2.3 show the LT SPICE outputs and calculated values respectively. Having viewed the effect that different magnitudes had on the output of the Active Probe, it was decided that the initial values for C_4 and C_5, 1 pF and 47 pF respectively, would be used in the circuit once it was constructed. This was due in part to an attempt to keep the capacitance of the circuit at a minimum, and also due to the maximum input of the Op-Amp chosen being -2.5 V, so a ratio of 1:50 would allow a maximum voltage of -125 V to be applied to the SPAD.

2.4 Modelling the Differential Amplifier

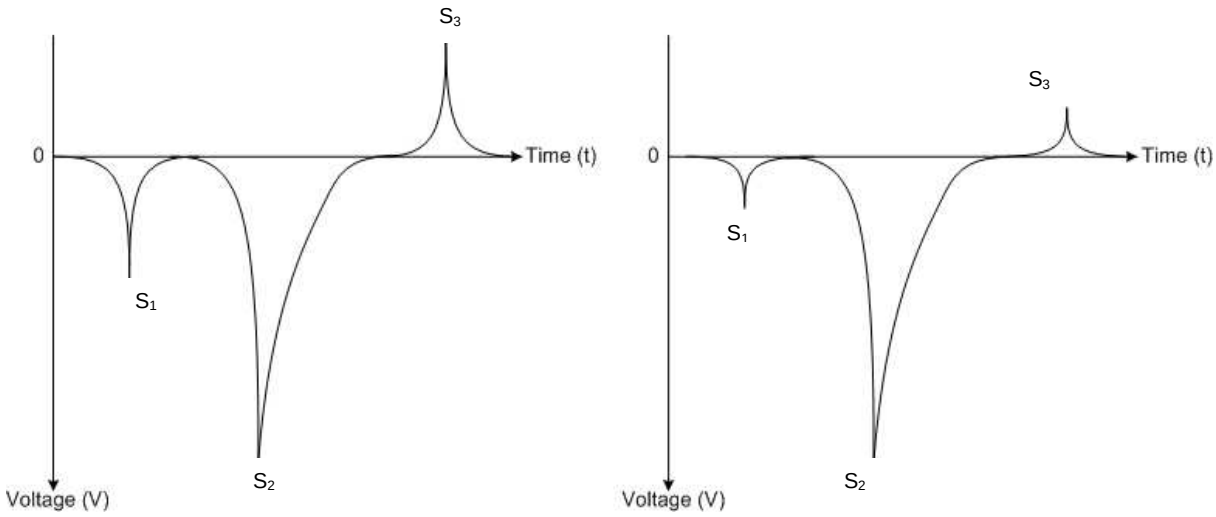


Figure 2.11 - A Sketch of the Outputs of the CQC with Transient Voltage Spikes, S_1 and S_3, and Avalanche Pulse, S_2, (a) Without and (b) With the Differential Amplifier

Section 1.4.6 describes that operating in GM often produces transient voltage spikes as the applied pulses charge and discharge the internal capacitance of the APD. These spikes cause false counts within the discriminator by being mistaken for avalanche pulses and can also mask avalanche events [61]. The Instrumentation Amplifier, also known as the Differential

Amplifier, is introduced in section 1.4.8 and was the chosen solution to be used with the CQC.

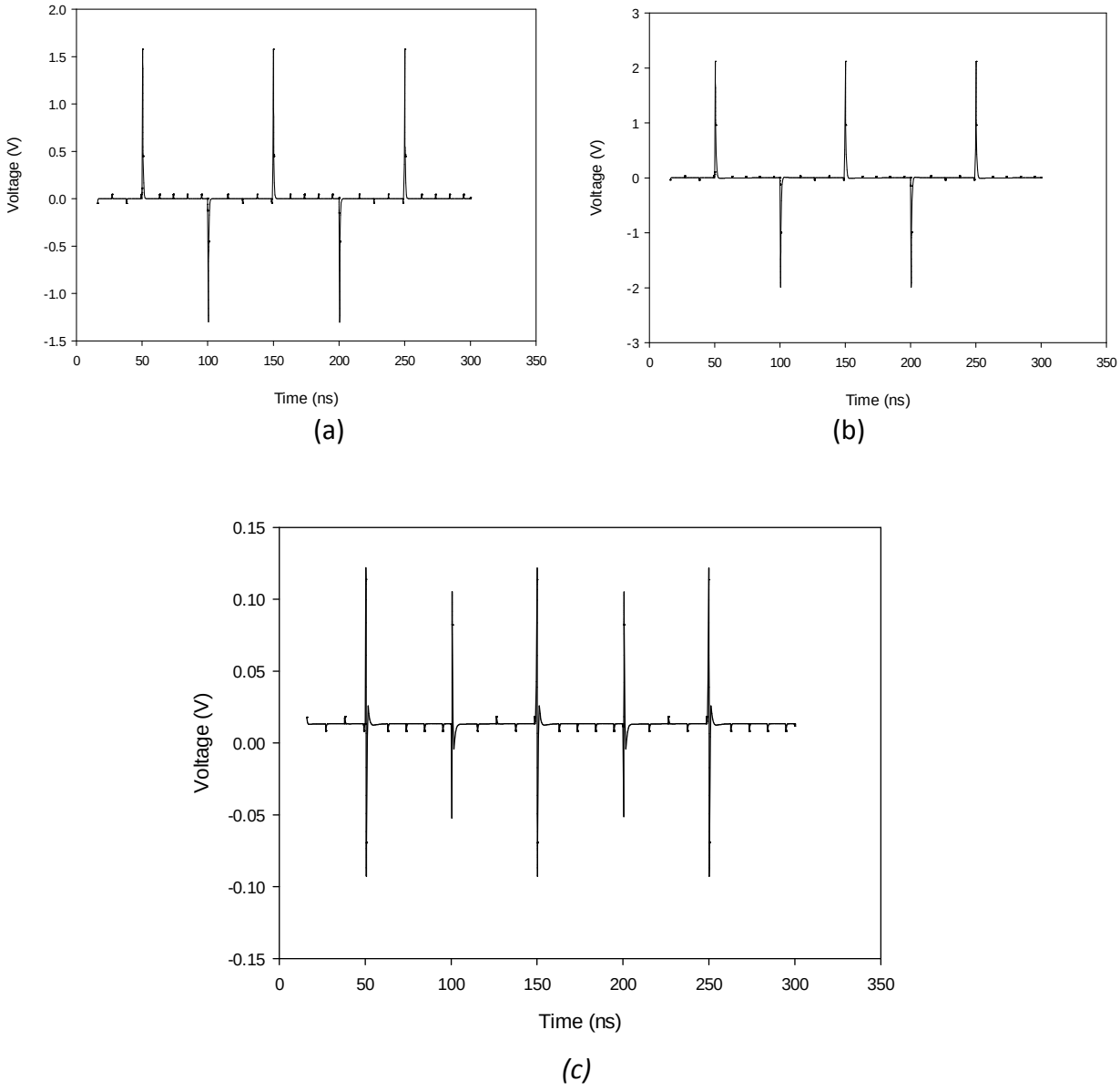


Figure 2.12 - LT SPICE Waveforms Displaying the Transients Present on (a) the DUT, (b) C_6 and (c) the Output of the Differential Amplifier With -8 V Pulses Applied

Shown schematically in Figure 2.3, the Differential Amplifier is made from three Op-Amps (LMH6609) [77], X_2, X_3 and X_4, and a network of resistors, R_6-R_{10}, which compare the DUT output with that from C_6, a capacitive element chosen to mimic the behaviour of this device [37]. C_6 can be either a second 'identical' APD, to keep the electrical characteristics of the two devices as similar as possible [64], or a trimming capacitor whose capacitance can be

adjusted to match that of the DUT. In the case where another APD is used, the breakdown of C_6 is prevented by allowing only the over-bias pulses to be applied across it, meaning the potential difference never becomes high enough for an avalanche event to occur.

When pulses are applied to the circuit in this configuration, transient voltage spikes are produced in both the DUT and C_6, which are connected to the non-inverting inputs of X_3 and X_2 respectively. These Op-Amps are present to buffer the signals before they reach X_4, which subtracts one signal from the other. As the avalanche pulse, S_2, will only be present on the output from the DUT, the effect of this circuit is to reduce or remove the voltage spikes, S_1 and S_3, from the output as shown in Figures 2.11(a) and (b) whilst keeping S_2 its original magnitude.

Figures 2.12(a), (b) and (c) show modelled results of the Differential Amplifier in LT SPICE during the application of -8 V pulses across the CQC. Figures 2.12(a) and (b) show the voltage across the DUT and C_6, with spikes ≈ 1.5 and 2.0 V, respectively. Once passed through the Differential Amplifier the magnitudes reduced to ≈ 100 mV, as shown in Figure 2.12(c), demonstrating correct transient cancellation by the Differential Amplifier.

Chapter Three - Construction of the CQC

This chapter will discuss the steps taken to design the PCB layout and construct the circuit used in the laboratory, before presenting test results obtained from it. They will then be compared to the modelling results from Chapter 2 and the data in [05]..

3.1 Schematic Capture and Board Layout

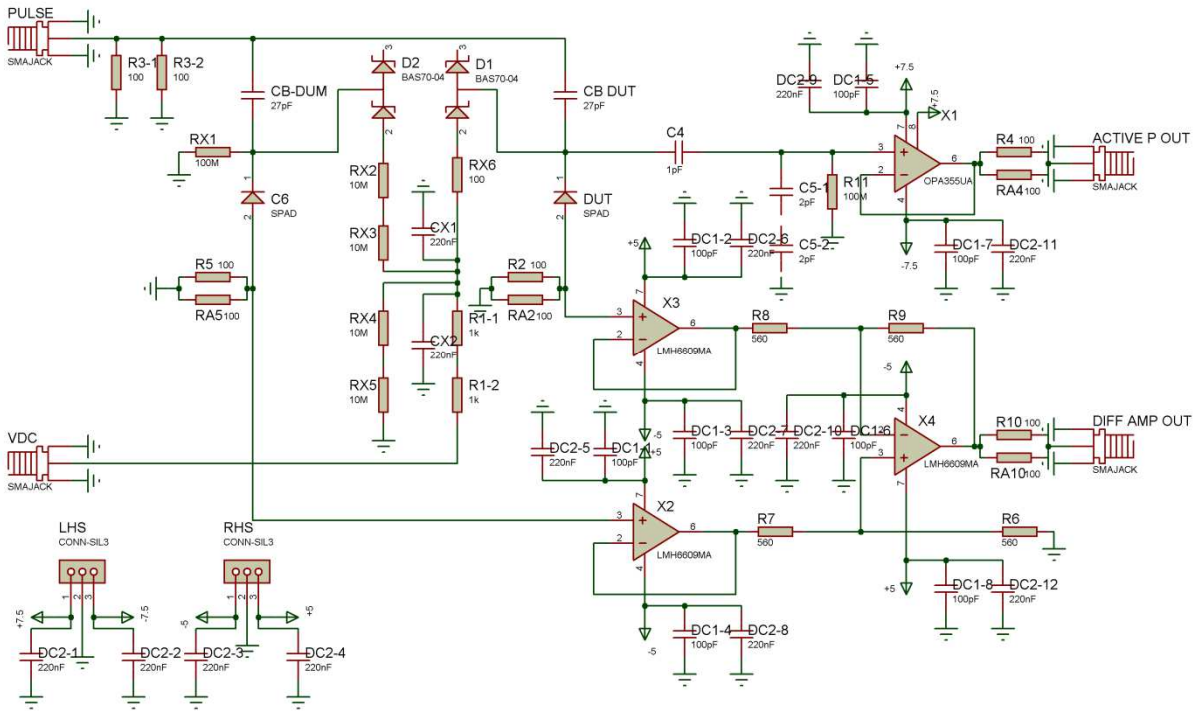
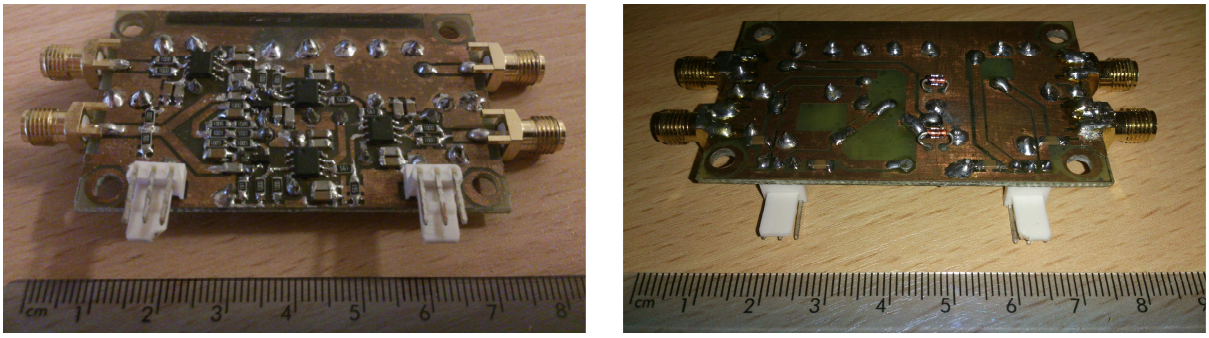


Figure 3.1 - ISIS Circuit Schematic

The PCB was designed using Proteus, a piece of computer software comprising of two core programs, ISIS and ARES, which are able to communicate with one another and streamline the process. Firstly, the schematic was input into ISIS, which told Proteus how the various components were connected together, as depicted in the schematic in Figure 3.1. Although the facility to perform interactive simulations and tests was available in addition to the schematic capture; as LT SPICE had already provided detailed information on the circuit's

behaviour and ISIS would not provide anything extra, it was not considered advantageous to repeat the circuit modelling with ISIS.

In Dimler *et al.*, [05], the theory behind the CQC is introduced and discussed, however the circuit layout is not examined in detail. The PCB produced for this work was designed and optimised independently. Green *et al.* [78] disclose a number of high frequency techniques that help to improve the operating performance of high frequency circuits, such as the use of multiple capacitors of varying magnitudes to de-couple nodes to ground [78]. For this reason there are additional components shown in Figure 3.1 that are not shown in Figure 2.3. These techniques were applied further when the layout was designed in ARES, with Figures 3.2(a) and (b) depicting the PCB that was fabricated and constructed.



(a) *(b)*

Figure 3.2 - Photographs of the Produced PCB Depicting the (a) Top and (b) Bottom View

Additionally, three design features unique to this circuit were implemented. The first of the unique considerations laid out the track for the over-bias pulses so that the fork was electrically symmetrical as shown in Figure 3.2(a). This meant that the DUT and C_6 would receive as close to identical inputs as possible, without a delay to either device. Next, the op-amps, X_2 and X_3, were placed as close as possible to the DUT and C_6 so that the signal generated in the DUT would travel the least distance before being input into the amplifier. Finally, the Differential Amplifier was laid out as symmetrically as possible to keep the inputs to X_4 in phase, maximising the cancellation effect. Silicon diodes (1N148) [79] were chosen for use as the DUT and C_6 throughout this chapter in place of APDs or SPADs. This allowed circuit operation to be examined as well as a comparison with previously modelled data and work by Dimler *et al.* [05] without the possibility of avalanche events.

3.2 Active Probe Calibration and Testing

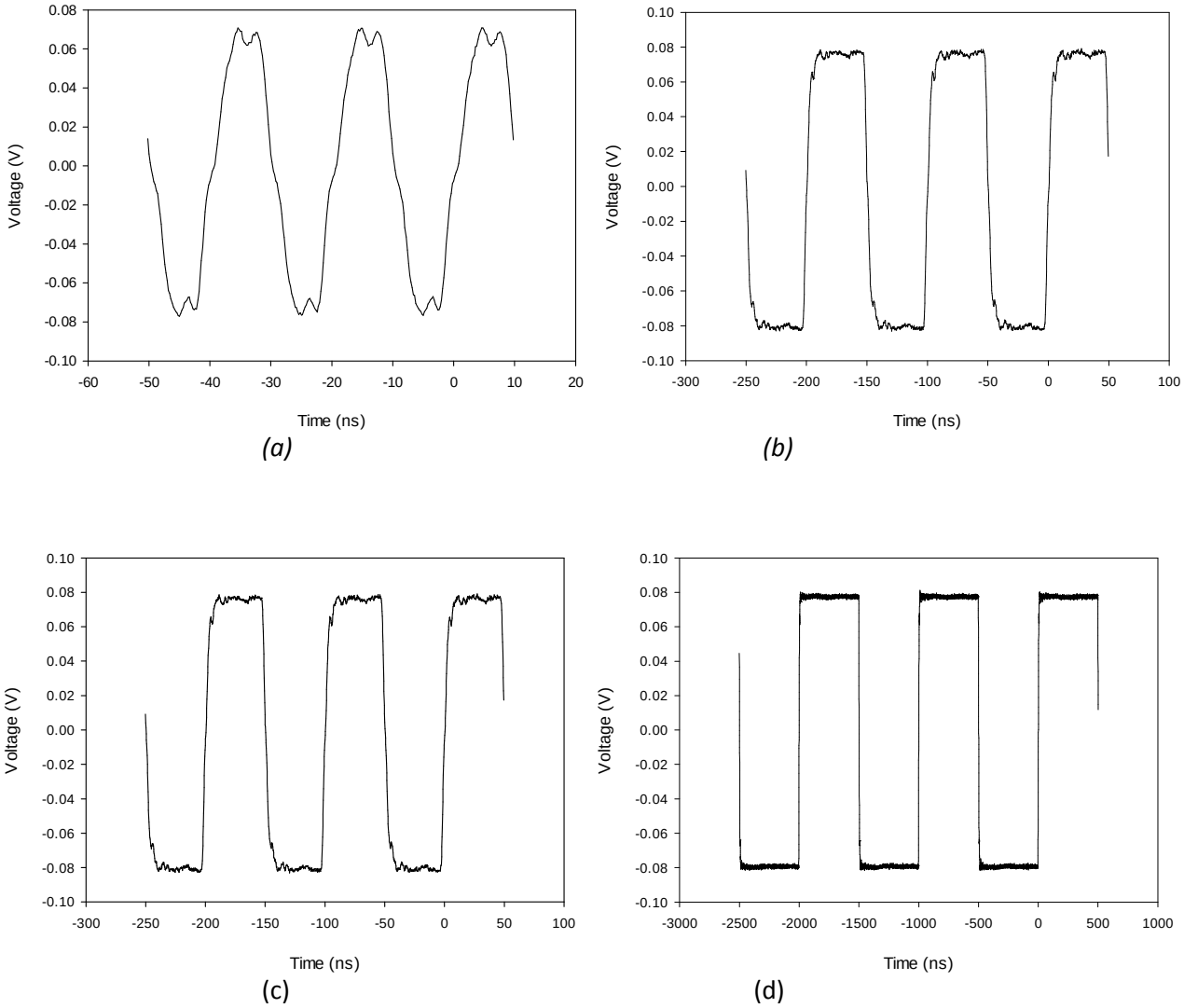
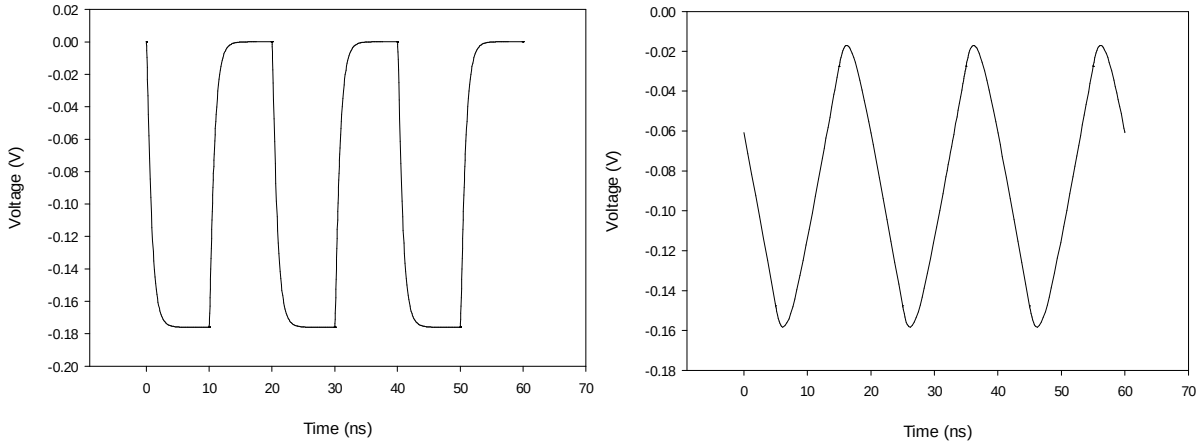


Figure 3.3 - Measured Output of the Active Probe During the Application of -8 V Pulses with (a) 10 ns, (b) 50 ns, (c) 500 ns and (d) 5000 ns Pulse Width

Figures 3.3(a), (b), (c) and (d) show the "Active P Out" output for -8 V applied at the point "Pulse" shown in Figure 3.1 with pulse widths of 10 ns, 50 ns, 500 ns and 5000 ns, respectively, giving output waveforms ≈ 160 mV. Figure 3.1 will be used when referring to points on the circuit until the end of the document, unless stated otherwise. This confirms effective capacitance division within the circuit. Section 1.4.10 states that one of the advantages of the CQC is that it does not distort the shape of the over-bias pulses as they are applied across it as in PQCs and AQCs. The waveforms in 3.3 (b), (c) and (d) show that

holds true and is consistent with the expected outputs shown in section 2.3, substantiating that the Active Probe can be used to monitor the over-bias pulses applied across the device. These results are also consistent with Dimler *et al.* [05] which showed over-bias pulses of -8 V over-bias pulses with widths of 50 ns, 500 ns and 5000 ns.



(a) *(b)*

Figure 3.4 -LT SPICE Output Waveform of the Active Probe During the Application of -8 V Pulses with 10 ns Pulse Width and (a) 2 ps Rise Time (b) 5 ns Rise Time

Figure 2.7(a) shows the LT SPICE output with 5 ns pulses applied. Limitations of the signal generator meant the shortest pulses that could be produced were 10 ns, so modelling was repeated for 10 ns pulse widths, results are shown in Figure 3.4(a). Discrepancies between 3.3(a) and 3.4(a) can be explained by the difference between the rise/fall time of the applied pulses. Modelling had been conducted with a 2 ps rise/fall time, whereas the shortest signal generator rise/fall time was 5 ns. For pulses equal to 10 ns, the rise/fall time becomes equal to the pulse width, yet for longer pulses the rise/fall time becomes increasingly insignificant. This causes the triangular shape that can be seen in 3.3(a). By setting the rise and fall times in LT SPICE to 5 ns each, it is possible to demonstrate this in the modelling work, shown in 3.4(b).

To determine the Active Probe multiplication factor, pulses were applied to the point "Pulse" and observed at "Active P Out". The magnitude of the applied pulses was increased from -0.5 to -10.0 V. Figure 3.5 shows the magnitude of the applied input pulses and the magnitude of the observed output pulses. The gradient gives the multiplication factor of

$\frac{1}{54.4}$, which is within 10% of the theoretical value and within 5% of the value modelled in section 2.3. The difference between the expected values and the experimental value is most likely a result of stray capacitance present in the circuit.

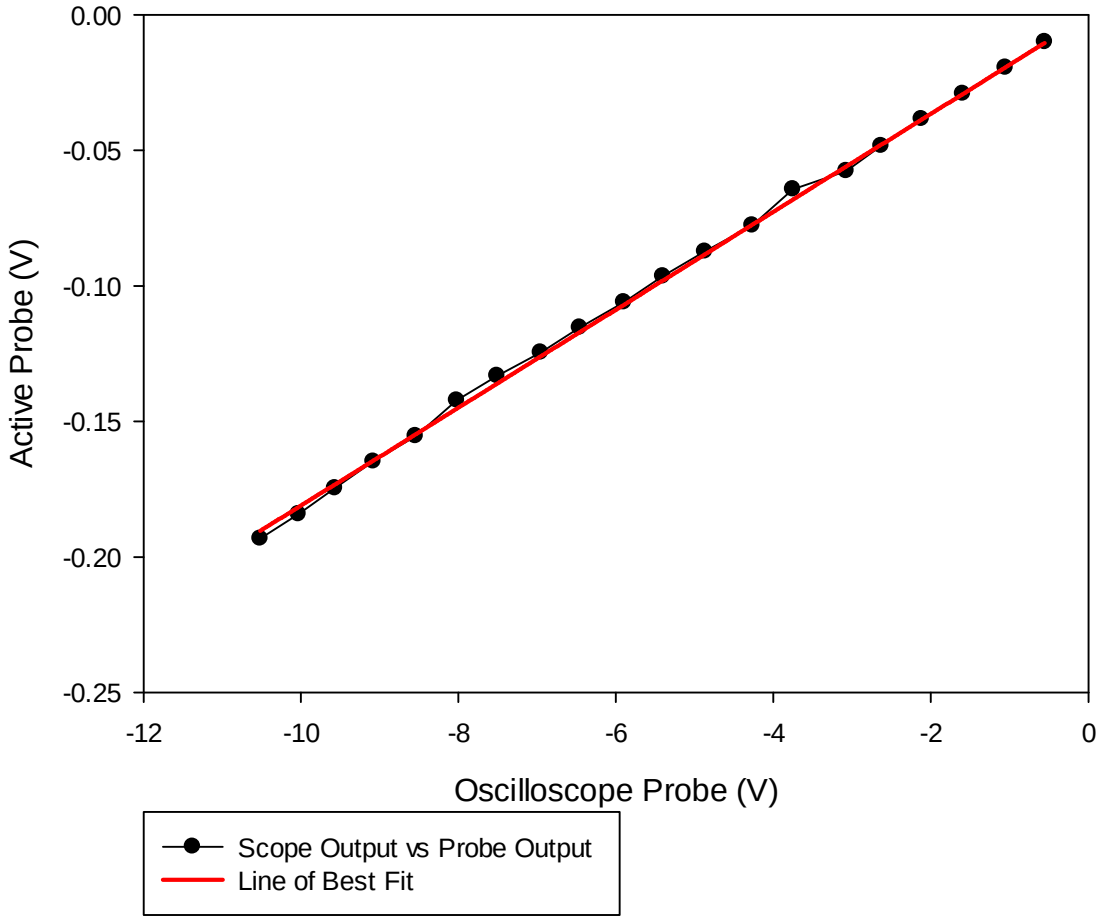


Figure 3.5 - Measured Output of the Active Probe Against Input Pulses From -0.5 to -10.0 V

3.3 Transient Cancellation

Figures 3.6(a) and (b) show the output of the Differential Amplifier observed at the point "Diff Amp Out" without and with the transient cancellation circuitry, respectively. For Figure 3.6(a), the input signal to X_2 was connected to the ground terminal of the voltage source, removing one of the inputs to X_4 and disconnecting the transient cancellation portion of the circuit. The transients shown appear at the leading and trailing edges of the over-bias pulse, returning to pre-transient values during the flat portion of the pulse and at its termination.

For Figure 3.6(a) the voltage spikes are ≈ 50 mV and the value of ground is 0 mV. Similar transients can be seen in Dimler *et al.*, [05] where they claim they were able to reduce the transients present on their output from 150.4 mV to <10 mV.

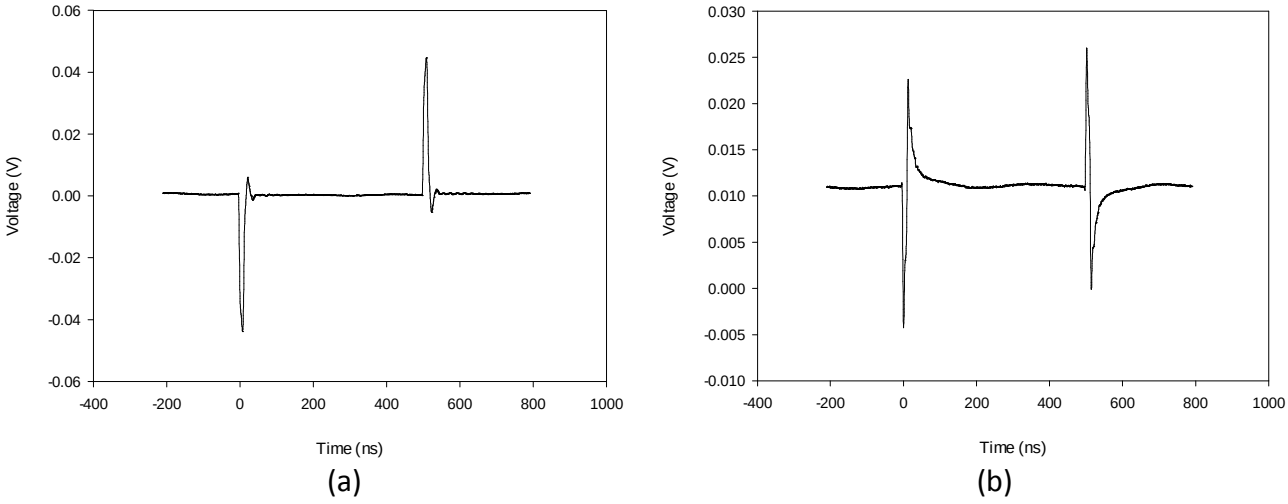


Figure 3.6 - Measured Output of the Differential Amplifier (a) Without and (b) With the Transient Cancellation Circuitry

Figure 3.6(b) displays the output when the connection to the ground terminal of the voltage source was removed from X_2, reconnecting the transient cancellation circuitry and reducing the voltage spikes from ≈ 50 mV to ≈ 15 mV. This is consistent with the modelled data displayed and discussed in section 2.4. Reducing the magnitude of the transients results in a lower probability that the discriminator will generate a false count, allowing the thresholds of the discriminator to be set lower, hence reducing the magnitude of the minimum detectable avalanche event and total charge flow. The difference in the value of ground, which is 0 mV in Figure 3.6(a) and ≈ 12 mV in Figure 3.6(b), is due to the removal of the additional ground connection to the circuit board. When the PCB was designed, it was considered that two connections from the ground plane to the ground terminal on a voltage source would be sufficient, but as can be seen a third connection may be required to prevent this deviation from 0 mV. However, as the ≈ 12 mV translation is present across the entire waveform and does not affect the absolute magnitude of the voltage spikes, its impact on the circuit is negligible.

3.4 Review of Data

In section 2.3, the Active Probe portion of the circuit was modelled in LT SPICE, allowing the behaviour to be examined, optimised and compared to the work by Dimler *et al.*, [05] prior to constructing a PCB. -8 V pulses with widths of 5, 50, 500 and 5000 ns were applied across the DUT and the output waveforms shown graphically in Figures 2.7(a), (b), (c) and (d). The pulses with widths of 50, 500 and 5000 ns gave an output of undistorted square waves, consistent with [05]. The 5 ns pulses had no basis for comparison, yet the output waveforms displayed rounding on the leading and trailing edges of the wave due to the capacitive response of the circuit, demonstrating that the maximum operating frequency of the circuit was being approached. The result of this effect would be that for a portion of the gate the device would not be biased at the desired level, causing the drawbacks discussed in 1.5.8.

Section 3.2 repeated the 50, 500 and 5000 ns measurements taken in section 2.3 with the Active Probe on the physical PCB. These are displayed in Figures 3.3(b), (c) and (d) and are a close fit to the modelled results and previous work [05], although some noise is present on the top and bottom of the pulse. The lack of 5 ns measurements was due to the inability of the signal generator to provide shorter pulses than 10 ns, so modelling and experimental work was carried out for pulses with widths of 10 ns with the outcomes shown in Figures 3.4(a) and 3.3(a), respectively. Figure 3.3(a) does not resemble any of the modelling waveforms and appears almost triangular in shape. Investigation determined this was because the minimum rise/fall time of the signal generator was equal to 5 ns, or the duration of 10 ns pulses. By increasing the rise/fall time in LT SPICE to match that of the signal generator a comparable output was modelled as Figure 3.4(b) depicts, confirming the behaviour of the physical PCB.

The findings in sections 2.3 and 3.2 demonstrated the Active Probe was a suitable choice for monitoring the pulses applied across the DUT. Although undesirable effects were displayed at pulses of widths <50 mV, this was not a shortcoming of the Active Probe circuitry but a result of the CQC nearing its maximum operation frequency, as the shape of the pulses were

affected by capacitive response and the rise/fall time of the signal generator. The pulses of widths >50 mV demonstrated a good qualitative fit to data shown in Dimler *et al.*, [05].

Sections 2.3 and 3.2 also examined the capacitive division portion of the Active Probe, with hand calculations and extensive modelling conducted in section 2.3 before calibration took place in section 3.2. The desired capacitive division was ≈ 1:50, with hand calculations giving a multiplication factor of $\frac{1}{49.74}$ and modelling data providing one of $\frac{1}{51.75}$. The difference between these values is due to LT SPICE taking into account the capacitance on the input of the Op-Amp and other such details considered negligible for the purpose of hand calculations. Section 3.2 investigated the capacitive division further, stepping the pulse height from -0.5 to -10.0 V and observing the output of the Active Probe, with the results shown in Figure 3.5. Calculating the gradient produced a multiplication factor of $\frac{1}{54.40}$, with the discrepancy between this and the modelling value a result of stray capacitance.

LT SPICE was used to model the Differential Amplifier portion of the circuit as described in section 2.4. Figures 2.12(a) and (b) display the voltage spikes across the DUT and C_6 with magnitudes of ≈ 1.5 and 2.0 V, respectively. Although these values seemed large, particularly in comparison to the ≈ 150 mV transients in previous work [05], they were reduced by the Differential Amplifier to ≈ 100 mV as can be seen in Figure 2.12(c), showing expected behaviour. Section 3.3 displays the output from the constructed PCB, where the transients were reduced from ≈ 50 mV to ≈ 15 mV in Figures 3.6(a) and (b), respectively. Dimler *et al.*, [05] displayed transients ≈ 150 mV prior to cancellation and claim these were reduced to <10 mV following cancellation. Differences between modelled and experimental values in sections 2.4 and 3.3 could be a result of the high speed techniques adopted from Green *et al.*, [78] when designing the PCB layout and the improved symmetry of the Differential Amplifier on the PCB. Figures 2.12(c) and 3.6(b) display a shift from 0 for the steady state level as ground is not equal to 0 in this case.

Chapter Four - Conclusions and Further Work

4.1 Conclusions

Through a comprehensive literature review of photon detection methods and associated quenching circuits, capacitive quenching circuits were identified as an area that warranted further investigation. Computer modelling has been carried out to confirm the behaviour dictated by theory and optimise the value of key components. After this process, a physical circuit has been constructed which allowed practical experimentation to be conducted.

Research had shown that passive quenching circuits are plagued by the low-pass filter formed by the RC combination of the ballast resistor and the SPAD as it deforms the over-bias pulses. The removal of the resistor in favour of a capacitor solved this problem, allowing the square waves applied by the signal generator to be propagated to the DUT cleanly, without compromising the quenching ability of the circuit. These pulses were monitored by the Active Probe, added to prevent additionally loading the CQC and to export the waveforms to an external piece of equipment. The Active Probe and its performance are examined in sections 2.3 and 3.2.

The Differential Amplifier has been included in addition to the CQC to reduce the transient voltage spikes produced during the application and discontinuation of the over-bias pulses. Sections 2.4 and 3.3 confirm that the Differential Amplifier demonstrated a good ability to reduce their presence on the output of the CQC. In the absence of the Differential Amplifier, the voltage transients were ≈ 50 mV, falling to ≈ 15 mV with its inclusion, as shown in Chapter 3. The lower magnitude of these spikes allows for the discriminator to amend its threshold voltage and thereby reduce the magnitude of the minimum detectable avalanche event, without the risk of increasing the number of false counts.

In conclusion, the work in this thesis has shown that the CQC offers a viable alternative to the other forms of quenching circuits currently available.

4.2 Recommendations for Further Work

Should research into capacitive quenching circuits be continued from the point at which this thesis concludes, this section provides recommendations for further work that could be done.

The circuit described and tested in Chapter 3 utilised silicon diodes (1N148) [79] for the Device Under Test and the capacitor C_6, to allow circuit operation to be compared with the modelled data without any avalanche events to contend with. APD and SPAD devices could be fabricated and characterised, eventually replacing the silicon diodes for dark count measurements. The work conducted by Dimler *et al.* [05] used Perkin Elmer C30902S SPADs [71], these SPADs could be used following tests with the aforementioned fabricated devices as a means of comparison.

Temperature variation was not focused on in this thesis as Dimler *et al.*, [05] obtained their measurements at 300 K. However, as described in section 1.4.3, lower temperatures have a positive impact on the Dark Count Rate [44][45], suggesting that experimentation in this area might be of interest. This could be attempted using a Peltier or Thermoelectric Cooler [80][81][82], a vacuum Dewar flask [83][84] or a cryostat [85][86][87], depending on the desired temperature range. As described by Liu *et al.*, [86] it is possible to insert the SPAD and associated quenching circuit directly into the cryostat, providing that the internal dimensions are sufficient to accommodate this, and the components are not cooled beyond their minimum operating temperatures. As the SPAD used by Dimler *et al.* [05][71] and the Op-Amps used for the Active Probe [76] and Differential Amplifier [77] have minimum operating temperatures of 233 K, a Peltier Cooler would be recommended as minimal modifications to the existing design would be required in order to test to this temperature.

Finally, the circuit could be used to attempt to detect single photons.

References

[01] G E Stillman and C M Wolfe, “Physical Properties of Semiconductors”, Prentice Hall, 1989, pp291-355.

[02] M A Itzler, X Jiang, R Ben-Michael, B Nyman and K Slomkowski, “Geiger-Mode APD Single Photon Detectors”, Optical Fiber Communication/National Fiber Optic Engineers Conference, pp1-3, May 2008.

[03] S Cova, M Ghioni, A Lacaita, C Samori and F Zappa, “Avalanche Photodiodes and Quenching Circuits for Single Photon Detection”, Journal of Applied Optics, **35**, Number 12, pp1956-1976, April 1996.

[04] A Gallivanoni, I Rech, M Ghioni, "Progress In Quenching Circuits for Single Photon Avalanche Diodes", IEEE Transactions on Nuclear Science, **57**, Number 6, pp3815-3823, December 2010.

[05] S J Dimler, J S Ng, R C Tozer, G J Rees and J P R David, “Capacitive Quenching Measurement Circuit for Geiger-Mode Avalanche Photodiodes”, Journal of Selected Topics in Quantum Electronics, pp919-925, July 2007.

[06] S Vasile, P Gothoskar, D Sdrulla and R Farrell, “Photon Detection With High Gain Avalanche Photodiode Arrays”, IEEE Nuclear Science Symposium, **1**, pp339-342, November 1997.

[07] S Pellegrini, G S Buller, J M Smith, A M Wallace and S Cova, “Laser-based Distance Measurement Using Picosecond Resolution Time-Correlated Single-Photon Counting”, IOP Journal of Measurement Science and Technology, **11**, Number 6, pp712-716, March 2000.

[08] F Stellari, A Tosi, F Zappa and S Cova, “CMOS Testing via Time-Resolved Luminescence Measurements and Simulations”, IEEE Transactions on Instrumentation and Measurement, **53**, Number 1, pp163-169, February 2004.

[09] M Ghioni, S Cova, I Rech and F Zappa, “Monolithic Dual-Detector for Photon-Correlation Spectroscopy With Wide Dynamic Range and 70-ps Resolution”, IEEE Journal of Quantum Electronics, **37**, Number 12, pp1588-1593, December 2001.

[10] S Cova, A Longoni, A Andreoni and R Cubeddu, “A Semiconductor Detector for Measuring Ultraweak Fluorescence Decays with 70-ps FWHM Resolution”, IEEE Journal of Quantum Electronics, **19**, Number 4, pp630-634, April 1983.

[11] T A Louis, G Ripamonti and A Lacaita, “Photoluminescence Lifetime Microscope Spectrometer Based on Time Correlated Single Photon Counting With An Avalanche Diode Detector”, AIP Review of Scientific Instruments, **61**, Number 1, pp11-22, August 1989.

[12] K J Gordon, V Fernandez, P D Townsend and G S Buller, “A Short Wavelength GigaHertz Clocked Fiber-Optic Quantum Key Distribution System”, IEEE Journal of Quantum Electronics, **40**, Number 7, pp900-908, July 2004.

[13] A L Lacaita, P A Francese and S D Cova, “Single-Photon Optical-Time-Domain Reflectometer at 1.3µm with 5-cm Resolution and High Sensitivity”, Optics Letters, **18**, Number 13, pp1110-1112, July 1993.

[14] H Kunimori, B Greene, K Hamal and I Prochazka, “Centimetre Precision Eye-Safe Satellite Laser Ranging Using a Raman-Shifted Nd:YAG Laser and Germanium Photon Counter”, IOP Journal of Optics A: Pure and Applied Optics, **2**, Number 1, pp1-4, August 1999.

[15] F Stellari, F Zappa, S Cova and L Vendrame, “Tools for Non-Invasive Optical Characterization of CMOS Circuits”, IEDM Technical Digest, International Electron Devices Meeting, **1**, 1, pp487-490, December 1999.

[16] J Hecht, "A Short History of Fiber Optics" Reproduction of "Fiber Optics Technician's Handbook" by J Hayes, Delmar Publishers, Albany, New York. Also excerpted from "City of Light: The Story of Fiber Optics", Oxford University Press, New York, 1999.

[17] J Crisp, "Introduction to Fiber Optics", 2nd Edition, Biddles Ltd, Chapter 3, pp19-20, 2001.

[18] M Azadeh, "Fiber Optics Engineering", Springer, Chapter 5, pp126-155, 2009.

[19] J C Campbell, "Recent Advances in Telecommunications Avalanche Photodiodes", Journal of Lightwave Technology, **25**, Number 1, pp109-121, Jan 2007.

[20] Hamamatsu Photonics K.K. Editorial Committee, "Photomultiplier Tubes, Basics and Applications" Third Edition, (Editorial Chief Toshikazu Hakamata), Introduction, August 2007.

[21] Hamamatsu Photonics K.K. Editorial Committee "Photomultiplier Tubes, Construction and Operating Characteristics, Connections to External Circuits" Catalogue (Editorial Chief Toshikazu Hakamata), December 2003 Revised Edition.

[22] Hamamatsu Photonics K.K. Editorial Committee, "Photomultiplier Tubes, Basics and Applications" Third Edition, (Editorial Chief Toshikazu Hakamata), Chapter 2, Basic Principles of Photomultiplier Tubes, August 2007.

[23] S. M. Sze, "Semi-Conductor Devices, Physics and Technology" 2nd Edition, John Wiley & Sons, Inc, Chapter 9, pp290-991, 1981.

[24] Definition of Quantum Efficiency in The Photonics Dictionary, http://www.photonics.com/Directory/Dictionary/Definition.aspx?type=2&DictionaryID=6488

[25] Hamamatsu Photonics K.K. Editorial Committee, "Photomultiplier Tubes, Basics and Applications" Third Edition, (Editorial Chief Toshikazu Hakamata), Chapter 11, Hybrid Photo-Detector (HPD), August 2007.

[26] G Barbarino, R de Asmundis, G de Rosa, G Fiorillo and S Russo, "A New Design for a High Gain Vacuum Photomultiplier: The Silicon PMT Used as Amplification Stage", Science Direct, Nuclear Physics B, **197**, pp52-56, 2009.

[27] N Dinu, R Battison, M Boscardin, G Collazuol, F Corsi, GF Dalla Betta, A del Guerra, G Llosa, M Ionica, G Levi, S Marcatili, C Marzocca, C Piemonte, G Pignatel, A Pozza, L Quadrani, C Sbarra and N Zorzi, "Development of the First Prototypes of Silicon Photo Multiplier (SiPM) at ITC-irst", Science Direct, Nuclear Instruments and Methods in Physics Research A, **572**, pp422-426, 2007.

[28] X Zhang, Junli Wan, C Yan, F Dong, "The Development and Application of Single Photon Detectors", Proceedings of SPIE, **7055**, pp70550V-1-70550V-8, 2008.

[29] G J Rees and J P R David, “Nonlocal Impact Ionization and Avalanche Multiplication”, IOP Journal of Physics D: Applied Physics, **43**, Number 24, pp1-17, June 2010.

[30] M. Azadeh, "Fiber Optics Engineering", Springer, Chapter 6, pp156-184, 2009.

[31] A Pauchard, G Ribordy, A Roches, L Widmer, R Thew and W Becker, "Quantum Detectors: Single Photon Detectors Take Quantum Leaps", Laser Focus World, **42**, Number 5, pp79-64, 2006.

[32] E Sciacca, A C Giudice, D Sanfilippo, F Zappa, S Lombardo, R Consentino, C Di Franco, M Ghioni, G Fallica, G Bonanno, S Cova and E Rimini, "Silicon Planar Technology for

Single Photon Optical Detectors", IEEE Transactions On Electron Devices, **50**, Number 4, pp918-925, April 2003.

[33] Y S Kim, V Makarov, Y C Jeong and Y H Kim, "Silicon Single Photon Detector With 5Hz Dark Counts", Conference on Lasers and Electro-Optics (CLEO) and Conference on Quantum Electronics and Laser Science (QELS), Article Number 5224894, 2009.

[34] S Pellegrini, R E Warburton, L J J Tan, J S Ng, A B Krysa, K Groom, J P R David, S Cova, M J Robertson and G S Buller, "Design and Performance of an InGaAs-InP Single Photon Avalanche Diode Detector", IEEE Journal of Quantum Electronics, **42**, Number 4, April 2006.

[35] A Karlsson, M Bourennane, G Ribordy, H Zbinden, J Brendel, J Rarity and P Tapster, “A Single Photon Counter for Long Haul Telecom”, IEEE Circuits and Devices Magazine, **15**, Number 6, pp34-40, November 1999.

[36] 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, **13**, Number 4, pp895-905, July 2007.

[37] G Wu, C Zhou, X Chen and H Zeng, “High Performance of Gated-Mode Single Photon Detector at 1.55μm”, Optics Communications, **265**, Number 1, pp126-131, September 2006.

[38] X Bai, H D Liu, D C McIntosh and J C Campbell, “High-Detectivity and High-Single-Photon-Detection-Efficiency 4H-SiC Avalanche Photodiodes”, Journal of Quantum Electronics, **45**, Number 3, March 2009.

[39] X Bai, D McIntosh, H Liu and J C Campbell, “Ultraviolet Single Photon Detection With Geiger-Mode 4H-SiC Avalanche Photodiodes”, Photonics Technology Letters, **19**, Number 22, pp1822-1824, November 2007.

[40] E Cicek, Z Vashaei, C Bayram, R McClintock, M Razeghi and M P Ulmer, “Comparison of Ultraviolet APDs Grown on Free-Standing GaN and Sapphire Substrates”, SPIE Proceedings, **7780**, pp77801P-1-77801P-11, August 2010.

[41] A Vert, S Soloviev, A Bolotnikov and P Sandvik, “Silicon Carbide Photomultipliers and Avalanche Photodiode Arrays for Ultraviolet and Solar-blind Light Detection”, Sensors Conference, pp1893-1896, October 2009.

[42] S Verghese, K A McIntosh, R J Molnar, L J Mahoney, R L Aggarwal, M W Geis, K M Molvar, E K Duerr, and I Melngailis, “GaN Avalanche Photodiodes Operating in Linear-

Gain Mode and Geiger-Mode", Transactions on Electron Devices, **48**, Number 3, pp502-511, August 2002.

[43] M Razeghi, "Short-Wavelength Solar-Blind Detectors – Status, Prospects, and Markets", Proceedings of the IEEE, **90**, Number 6, pp1006-1014, June 2002.

[44] E A Dauler, P I Hopman, K A McIntosh, J P Donnelly and E K Duerr, "Scaling of Dark Count Rate with Active Area in 1.06µm Photon-Counting InGaAsP/InP Avalanche Photodiodes", Applied Physics Letters, **89**, Number 11, pp111102-1-111102-3, September 2006.

[45] M A Itzler, X Jiang, R Ben-Michael, B Nyman, K Slomkowski, "Geiger Mode APD Single Photon Detectors", Conference on Optical Fiber Communication (OFC) and National Fiber Optic Engineer Conference (NFOEC), Article Number 4528353, 2008.

[46] J C Campbell, "Advances In Single Photon Detectors", Conference Proceedings - International Conference on Indium Phosphide and Related Materials, Article Number 5012461, pp 140-145, 2009.

[47] F Zappa, A Lotito, A C Giudice, S Cova and M Ghioni, "Monolithic Active-Quenching and Active-Reset Circuit for Single Photon Avalanche Detectors", Journal of Solid-State Circuits, **38**, Number 7, pp1298-1301, July 2003.

[48] I Rech, A Ingargiola, R Spinelli, I Labanca, S Marangoni, M Ghioni and S Cova, "A New Approach to Optical Crosstalk Modeling in Single-Photon Avalanche Diodes", Photonics Technology Letters, **20**, Number 5, pp330-332, March 2008.

[49] A D Mora, A Tosi, S Tisa and F Zappa, "Single-Photon Avalanche Diode Model for Circuit Simulations", Photonics Technology Letters, **19**, Number 23, pp1922-1924, December 2007.

[50] G Giustolisi, R Mita and G Palumbo, "Verilog – A Modeling of SPAD Statistical Phenomena", IEEE International Symposium on Circuits and Systems, pp773-776, May 2011.

[51] J A Richardson, L A Grant and R K Henderson, "Low Dark Count Single-Photon Avalanche Diode Structure Compatible With Standard Nanometer Scale CMOS Technology", IEEE Photonics Technology Letters, **21**, Number 14, pp1020-1022, July 2009.

[52] M A Marwick and A G Andreou, "Single Photon Avalanche Photodetector With Integrated Quenching Fabricated In TSMC 0.18µm 1.8V CMOS Process", Electronics Letters, **44**, Issue 10, pp643-644, May 2008.

[53] M A Marwick and A G Andreou, "Fabrication and Testing of Single Photon Avalanche Detectors in the TSMC 0.18µm CMOS Technology", Information Sciences and Systems - Conference Proceedings, pp741-744, March 2007.

[54] M Liu, C Hu, J C Campbell, Z Pan and M M Tashima, “Reduce Afterpulsing of Single Photon Avalanche Diodes Using Passive Quenching With Active Reset”, Journal of Quantum Electronics, **44**, Number 5, pp430-434, May 2008.

[55] R Mita and G Palumbo, “High-Speed and Compact Quenching Circuit for Single-Photon Avalanche Diodes”, Transactions on Instrumentation and Measurement, **57**, Number 3, pp543-547, March 2008.

[56] D Cronin and A P Morrison, “Intelligent System for Optimal Hold-Off Time Selection in an Active Quench and Reset IC”, Journal of Selected Topics in Quantum Electronics, **13**, Number 4, pp911-918, July-August 2007.

[57] F Zappa, M Ghioni, S Cova, C Samori and A C Giudice, “An Integrated Active-Quenching Circuit for Single Photon Avalanche Diodes”, Transactions on Instrumentation and Measurement, **49**, Number 6, pp1167-1175, December 2000.

[58] D Cronin, L D’Andrea and A P Morrison, “Hold-Off Time Autotuning for a Geiger-Mode Avalanche Photodiode”, Lasers and Electro-Optics Society (LEOS), pp957-958, October 2005.

[59] M Ghioni, S Cova, F Zappa and C Samori, “Compact Active Quenching for Fast Photon Counting With Avalanche Photodiodes”, Review of Scientific Instruments, **67**, Number 10, pp3440-3448, October 1996.

[60] N S Nightingale, “A New Silicon Avalanche Photodiode Photon Counting Detector Module For Astronomy”, Experimental Astronomy, **1**, Number 6, pp407-422, 1991.

[61] Z Wei, P Zhou, J Wang, C Liao, J Guo, R Liang and S Liu, “An Integral Gated Mode Single Photon Detector At Telecom Wavelengths”, Journal of Physics D: Applied Physics, **40**, pp6922-6925, 2007.

[62] C Hu, X Zheng, J C Campbell, B M Onat, X Jiang and M A Itzler, “High-Performance InGaAs/InP-Based Single Photon Avalanche Diode With Reduced Afterpulsing”, SPIE Proceedings, **7681**, pp76810S-1-76810S-7, April 2010.

[63] D S Bethune and W P Risk, “An Autocompensating Fiber-Optic Quantum Cryptography System Based on Polarization Splitting of Light”, Journal of Quantum Electronics, **36**, Number 3, pp340-347, 2000.

[64] A Tomita and K Nakamura, “Balanced, Gated-Mode Photon Detector for Quantum-Bit Discrimination at 1550nm”, Optics Letters, **27**, Number 20, pp1827-1829, 2002.

[65] D M Pozar, “Microwave Engineering”, 3rd Edition, John Wiley and Sons, Chapter 7, pp352-353, 2005.

[66] F Zappa, A Lacaita and S Cova, "Nanosecond Single-Photon Timing With InGaAs/InP Photodiodes", Optics Letters, **19**, Issue 11, pp846-848, January 1994.

[67] A D Mora, A Tosi, F Zappa, S Cova, D Contini, A Pifferi, L Spinelli, A Torricelli and R Cubeddu, “Fast-Gated Single-Photon Avalanche Diode for Wide Dynamic Range Near Infrared Spectroscopy”, **16**, Number 4, pp1023-1030, July 2009.

[68] S B Cho and S K Kang, “Weak Avalanche Discrimination for Gated-Mode Single-Photon Avalanche Photodiodes”, Optics Express, **19**, Number 19, pp18510-18515, 2011.

[69] N Namekata, S Sasamori and S Inoue, “800MHz Single-Photon Detection at 1550nm Using an InGaAs/InP Avalanche Photodiode Operated With A Sine Wave Gating”, Optics Express, **14**, Number 21, pp10043-10049, October 2006.

[70] Leshan Radio Company Ltd, "Zener Voltage Regulator Diodes", BZX84C2V4LT1 series datasheet, December 2004.

[71] Perkin Elmer Optoelectronics, "Silicon Avalanche Photodiodes", C30902 series datasheet, 2008.

[72] Giuseppe Massobrio and Paolo Antognetti, "Semiconductor Device Modeling with SPICE", 2nd Edition, R R Donnelley and Sons, 1993.

[73] K G McKay, "Avalanche Breakdown in Silicon", Physical Review, **94**, Number 4, pp877-884, May 1954.

[74] R B Fair and H W Wivell, "Zener and Avalanche Breakdown in As-Implanted Low-Voltage Si n-p Junctions", IEEE Transactions on Electron Devices, **23**, Issue 5, pp512-518, May 1976.

[75] J David Irwin and R Mark Nelms, "Basic Engineering Circuit Analysis", 10th Edition, John Wiley and Sons, 2005, pp247.

[76] Burr-Brown Products from Texas Instruments, "200MHz, CMOS Operational Amplifier", OPA 356 and OPA2356 series datasheet, 2001.

[77] National Semiconductor, "900MHz Voltage Feedback Op Amp", LMH6609 series datasheet, 2010.

[78] J E Green, R C Tozer and J P R David, "Stability in Small Signal Common Base Amplifiers", IEEE Transactions on Circuits and Systems, **99**, pp1-10, June 2012.

[79] NXP Semiconductors, "High-Speed Diodes", 1N148 and 1N4448 series datasheet, 2004.

[80] G K White, "Experimental Techniques in Low-Temperature Physics", 3rd Edition, Oxford Science Publications, 1979, pp161**.**

[81] D Stucki, G Ribordy, A Stefanov and H Zbinden, "Photon Counting for Quantum Key Distribution with Peltier Cooled InGaAs/InP APD's", Journal of Modern Optics, **48**, Issue 13, pp1967-1981, June 2001.

[82] L Xu, E Wu, X Gu, Y Jian, G Wu and H Zeng, "High-speed InGaAs/InP-based Single-photon Detector with High Efficiency", Applied Physics Letters, **94**, pp161106, April 2009.

[83] G K White, "Experimental Techniques in Low-Temperature Physics", 3rd Edition, Oxford Science Publications, 1979, pp34**.**

[84] R B Gomes, "Compound III-V Semiconductor Avalanche Photodiodes for X-ray Spectroscopy", PhD Thesis, University of Sheffield, pp135-137, June 2012.

[85] G K White, "Experimental Techniques in Low-Temperature Physics", 3rd Edition, Oxford Science Publications, 1979, pp171**.**

[86] 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", IEEE Journal of Selected Topics in Quantum Electronics, **13**, Number 4, pp887-894, August 2007.

[87] S Takeuchi, J Kim, Y Yamamoto and H H Hogue, "Development of a High-Quantum-Efficiency Single-Photon Counting System", Applied Physics Letters, **74**, Number 8, pp1063-1065, February 1999.