Title
Development of shallow trench isolation bounded single- photon avalanche detectors for acousto-optic signal enhancement and frequency up-conversion

Author
Hsu, Mark J.

Publication Date
2010

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Development of Shallow Trench Isolation Bounded Single-Photon Avalanche Detectors for Acousto-optic Signal Enhancement and Frequency Up-conversion

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in

Electrical Engineering (Photonics)

by

Mark J. Hsu

Committee in charge:

Professor Sadik C. Esener, Chair
Professor David Hall
Professor Yu-Hwa Lo
Professor Robert Mattrey
Professor Deli Wang

2010

The dissertation of Mark J. Hsu is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California, San Diego

2010

DEDICATION

To my family.

TABLE OF CONTENTS

Contents

LIST OF TABLES

LIST OF FIGURES

LIST OF EQUATIONS

ACKNOWLEDGEMENTS

Throughout the course of my Ph.D., I have had the great privilege of working with many brilliant minds and people who have had a profound impact on me as a researcher and a person. First and foremost, I would like to give great thanks to my adviser, Dr. Sadik Esener, who guided and mentored me through my Ph.D. degree, molding me into the researcher I am today. He taught me how to think creatively and to always keep the end goal in mind. He allowed me to explore different research areas to find my true interests, from optics and optoelectronics, to plasmonics and fluorescence, to finally single photon detectors and ultrasound contrast agents, and has always put me in the best possible position to succeed. Dr. Esener mentorship has been instrumental in reshaping the way I think, to have a vision and a clear direction, so that I'm able to truly understand the purpose of my work. It is through Dr. Esener's training that I have the tools necessary to succeed in future endeavors, both personally and professionally.

I would also like to thank Dr. Yu-Hwa Lo for his guidance with my single photon detector work. Without his technical expertise, the frequency up-conversion project presented in this dissertation would not have come to fruition. His knowledge and practical attitude towards engineering kept me grounded and always looking forward to new and exciting results.

My work with ultrasound and ultrasound contrast agents would not have been possible without the guidance of Dr. Robert Mattrey and Dr. David Hall. Dr. Mattrey

provided incredible insight to the needs and challenges in the clinic, helping me create new ideas and new directions for next generation in vivo imaging techniques. With his expertise in areas originally unfamiliar to me, he showed me the ropes, with an incredibly eventful day in the basement of UCSD Moore's Cancer Center getting microbubbles to oscillate. It is through much of his creative thinking that device concepts were conceived.

Some of my most productive work was working with Dr. Hall who has been instrumental in my Ph.D. degree. Under his guidance, I have become adept in formulating theories, particularly with in vivo optical imaging. Through his out of the box thinking, he has shown me how to systematically prove or disprove hypothesis and theories. Further, Dr. Hall has taught me how to be an effective researcher, to leave no stone unturned, and that scientific curiosity can lead to very powerful concepts. I truly appreciate Dr. Hall's time and effort spent mentoring me and helping me understand the true essence of scientific research.

I would like to thank my whole dissertation committee, Dr. Esener, Dr. Lo, Dr. Mattrey, Dr. Hall, and Dr. Deli Wang for taking the time out of their busy schedules to offer me great advice and assistance throughout my Ph.D. years.

I owe special thanks to Dr. Hod Finkelstein, who was essential in developing my interests in electrical engineering applied towards biology. I am thankful to Dr. Finkelstein for introducing to me the power of single photon detectors and allowing me to collaborate with him on a complete design cycle, from device concept, to modeling and simulation, to tape-out, to characterization, and finally to application. I

am forever indebted to Dr. Finkelstein for guiding me to become an effective and efficient researcher. He has advised me on many theoretical and experimental concepts and methodologies and has taught me to persevere through even the most difficult obstacles. It is through collaboration with Dr. Finkelstein that many of my papers were published. Dr. Finkelstein has continued to give me much professional advice, and I consider him a great friend.

I would also like to thank some of my colleagues in the Esener research group. I would like to thank Dr. Matthias Gross and Dr. Sanja Zlatanovic for their patience and guidance, particularly in my early years, with putting the right tools together in my optical characterization setups. Through their assistance, I was able to design highly efficient setups to get the best possible results. Further, I would like to thank Dr. Mohammad Eghtedari for his collaboration on the acousto-optic project. Working in the basement of Moore's Cancer Center and EBU-I, through many days of frustration, Dr. Eghtedari has demonstrated what determination really means. He has helped me in molding the acousto-optic project into what it is today, so that we did not continue to lose valuable time on a dead-end. I will always reflect fondly on my time working with Dr. Eghtedari. And I'll always be impressed by Dr. Eghtedari's self-taught electronics expertise.

I would also like to thank Michael Benchimol for his assistance on my work with frequency up-conversion and also fluorescence microbubbles. Through months of trial and error, Mike was able to develop dye-loaded microbubbles for fluorescence modulation. I wish him the best as he continues this work. I would also like to thank

Carolyn Schutt for helping with imaging of the fluorescent microbubbles. I would like to thank Dr. Stuart Ibsen for his technical expertise, in particular with ultrasound. Further, I would like to thank Dr. Andrew Goodwin for providing microbubble samples and also for spending much time with me discussing new ideas, development of theories, and helping me with my presentations. I would also like to thank Sifang You and James Cheng of the Lo group for collaboration on the frequency up-conversion project and design of the IR-SPAD. I would also like to thank Priscilla Haase and Ronnie Pait for their administrative support and for helping me get all the equipment and materials on time for my experiments.

Very importantly, I would like to thank my family. My parents have been always supportive of my dream of pursuing my Ph.D. and have continued to support me through all my ups and downs. They have been a constant in my life and have always stood by me no matter what the circumstance. I would like to thank my older brother Richard for his sound advice and support, and my little brother Michael for always keeping things fun and interesting. And I would like to thank Karen, for always keeping me grounded, for always keeping me joyful, and for always keeping optimistic and hopeful in all things.

Chapter 3, in part, is a reprint of the material as it appears in M. J. Hsu, H. Finkelstein, and S. Esener, "A CMOS STI-bound Single Photon Avalanche Photodiode with 27ps Timing Resolution and a Reduced Diffusion Tail," IEEE Electronic Device Letters, vol. 30, no. 6, 2009. The dissertation author was the primary investigator and author of this paper.

Chapter 4, in part, is a reprint of the material as it appears in: M. J. Hsu, D. Hall, S. Farshchi-Heydari, S. Esener, R. Mattrey, "Novel Acousto-optic Signal Enhancement using Microbubbles," World Molecular Imaging Congress (New Optical Instrumentation), Nice, France 2008. The dissertation author was the primary investigator and author of this proceeding.

Chapter 5, in part, is currently being prepared for submission for publication of the material as it appears in: M. J. Hsu, M. Benchimol, C. Schutt, D. J. Hall, R. F. Mattrey, S. Esener, "Fluorescence Modulation of DiI-loaded Microbubbles," in preparation for the Journal of Biomedical Optics; M. J. Hsu, M. Eghtedari, A. P. Goodwin, D. J. Hall, R. F. Mattrey, S. Esener, "Characterization of Individual Ultrasound Microbubble Dynamics with Novel Light Scattering System," in preparation for the Journal of the Acoustical Society of America. The dissertation author was the primary investigator and author of this material.

VITA AND PUBLICATIONS

1981	Born, Boston, MA.
2003	B. Sc. in Engineering, Electrical Engineering, Princeton University, Princeton, New Jersey.
2007	M. Sc., Electrical Engineering (Photonics), University of California, La Jolla, CA.
2010	Ph.D., Electrical Engineering (Photonics), University of California, La Jolla, CA.

- M. J. Hsu, M. Benchimol, C. Schutt, D. J. Hall, R. F. Mattrey, S. Esener, "Fluorescence Modulation of DiI-loaded Microbubbles," in preparation for the Journal of Biomedical Optics.
- M. J. Hsu, M. Eghtedari, A. P. Goodwin, D. J. Hall, R. F. Mattrey, S. Esener, "Characterization of Individual Ultrasound Microbubble Dynamics with Novel Light Scattering System," in preparation for the Journal of the Acoustical Society of America.
- M. J. Hsu, H. Finkelstein, and S. Esener, "A CMOS STI-bound Single Photon Avalanche Photodiode with 27ps Timing Resolution and a Reduced Diffusion Tail," IEEE Electronic Device Letters, vol. 30, no. 6, 2009.
- M. J. Hsu, D. Hall, S. Farshchi-Heydari, S. Esener, R. Mattrey, "Novel Acousto-optic Signal Enhancement using Microbubbles," World Molecular Imaging Congress (New Optical Instrumentation), Nice, France 2008.

- A. P. Goodwin, M. J. Hsu, M. Eghtedari, Y. Shi, J. N. Cha, R. F. Mattrey, and S. Esener, "Stimulus-Responsive Microbubbles for Site-Specific Ultrasound Contrast Enhancement," Gordon Research Conference; Chemistry of Supramolecules and Assemblies, Waterville, ME, 2009.

- D. J. Hall, M. J. Hsu, S. Esener, R. F. Mattrey, "Detection of Ultrasound-modulated Photons and Enhancement with Ultrasound Microbubbles," Photons Plus Ultrasound: Imaging and Sensing 2009, SPIE Vol. 7177, San Jose, CA 2009.

- H. Finkelstein, M. J. Hsu and S. Esener, "Dual-junction single-photon avalanche diode," Electronics Letters, vol. 43, no. 22, 2007.

- H. Finkelstein, M. J. Hsu, S. Zlatanovic, S. Esener, "Performance trade-offs of ultrafast single-photon avalanche photodiodes," Review of Scientific Instruments, vol. 78, no. 10, 2007.

- H. Finkelstein, M. J. Hsu, S. Esener, "STI-bounded single-photon avalanche diode in a deep-submicron CMOS technology," IEEE Electron Device Letters, vol. 27, no. 11, 2006.

- H. Finkelstein, M. J. Hsu, S. Esener, "An ultra-fast Geiger-mode single photon avalanche diode in 0.18 μm CMOS technology," Advanced Photon Counting Techniques, SPIE Vol. 6372, Boston, MA, 2006.

- H. Finkelstein, M. J. Hsu, S. Esener, “A compact single-photon avalanche diode in a deep-submicron CMOS technology,” International Conference of Solid-State Devices and Materials, Yokohama, Japan, 2006.
- Dogariu, M. J. Hsu, L.J. Wang, “Reducing Far-Field Diffraction by Structured Apertures,” Optics Communications, vol. 220, Issues 4-6, 2003.

ABSTRACT OF THE DISSERTATION

Development of Shallow Trench Isolation Bounded Single-Photon Avalanche Detectors for Acousto-optic Signal Enhancement and Frequency Up-conversion

by

Mark J. Hsu

Doctor of Philosophy in Electrical Engineering
(Photonics)

University of California, San Diego, 2010

Professor Sadik C. Esener, Chair

This dissertation describes the development of the first CMOS single photon avalanche diode (SPAD) fabricated in a deep-submicron commercial CMOS process. Single photon detection is often necessary for high-sensitivity, high dynamic range time-resolved optical measurements in diverse applications in medicine, biology, military, and optical communication links. Single photon avalanche diode (SPAD) detectors have become the device of choice and have made strong gains in recent years. They are unique in that they provide digital information of the arrival of an

individual photon impinging on the detector, thus being a very powerful tool when time-of-arrival information and timing resolution are crucial. Timing precision of the detector will improve contrast in fluorescence lifetime imaging and resolution in laser ranging applications.

Traditionally, single photon detectors have been fabricated using custom processes because of the conditions under which the devices operate under. Because of the need to sustain high currents and high electric-fields, special custom fabrication processes have been developed. These fabrication processes have great benefits such as low-noise, high detection efficiencies, low jitter, and tailored spectral responses towards longer wavelengths. However, these fabrication techniques are often undesirable due to increased capacitance from off-chip quenching, recharging, and processing circuitry, resulting in longer detector dead times and slower sampling rates. Furthermore, large-scale production and scalability to arrays is impractical. There has been a trend towards using Complementary Metal-Oxide-Semiconductor (CMOS) technology for constructing SPAD pixels and arrays with integrated circuitry to overcome these limitations. Though commercial CMOS technologies are by nature, generic, and are not designed for SPAD devices, they can still offer considerable advantages in certain areas where custom processes lack.

This dissertation describes the modeling, simulation, and full characterization of a CMOS STI-bounded Single Photon Avalanche Diode (SPAD). State-of-the-art sampling rates, dead-time, and jitter performance are characterized, and the device is compared to traditional diffused-guard ring structures for solid-state SPADs. Further, novel applications of the CMOS SPAD for acousto-optic signal enhancement and frequency up-conversion of 1550nm are described. An optical scatter system for acoustic characterization of ultrasound responsive microbubbles and particles is designed and developed. Further, a novel method of fluorescence modulation using dye-loaded microbubbles is demonstrated.

1. INTRODUCTION

1.1. Principles of CMOS Single-Photon Detection

Single photon detection is often necessary for high-sensitivity, high dynamic range time-resolved optical measurements in diverse applications in medicine [1-8], biology [9-13], military [14-17], and optical communication links [18-25]. Single photon avalanche diode (SPAD) detectors have become the device of choice and have made significant progress in recent years. They are unique in that they provide digital information of the arrival of an individual photon impinging on the detector, thus being a very powerful tool when time-of-arrival information and timing resolution are crucial.

Traditionally, single photon detectors have been fabricated using custom processes because of the need to sustain high currents and high electric-fields. These custom fabricated devices have great benefits with low-noise, high detection efficiencies, low jitter, and tailored spectral responses towards longer wavelengths. However, custom fabrication techniques are often undesirable due to increased capacitance from off-chip quenching, recharging, and processing circuitry, resulting in longer detector dead times and slower sampling rates. Furthermore, large-scale production and scalability to arrays is impractical. There has been a trend towards using generic Complementary Metal-Oxide-Semiconductor (CMOS) technology for constructing SPAD pixels and arrays with integrated circuitry to overcome these limitations. Though commercial CMOS technologies are by nature, generic, and are not designed for SPAD devices, they can still offer considerable advantages in certain areas where custom processes lack.

CMOS has become a standard technology for digital logic circuits such as microprocessors, microcontrollers, static RAM, as well as analog circuits such as image sensors, data converters, and transceivers for wireless communications. In the image sensor market, CCDs have long been superior in image quality with low noise characteristics, good sensitivity, high dynamic range, and high uniformity. However, manufacturing costs remain comparatively high. In the past two decades, CMOS has become the technology of choice for inexpensive low-power image sensors with fast readout rates over CCDs. The emergence of CMOS transistor logic has allowed CMOS image sensors to compensate for inherent physical limitations through integrated circuitry for image acquisition and on-chip processing. With development of newer architectures, CMOS sensors have gained in performance and reduced the gap in performance with CCDs.

In this dissertation, I will further add to this trend towards CMOS technologies for single photon avalanche diode (SPAD) detectors with improved performance and demonstrate new emerging applications where CMOS SPADs are most effective, particularly with respect to timing resolution and fast sampling rates. Through a complete design and fabrication cycle—device simulation, circuit simulations, layout, and device characterization—I will show the power of CMOS SPADs with integrated circuitry to overcome inherent physical limitations and improve on various aspects over custom fabricated and previous CMOS SPAD implementations.

In this chapter, I will give an overview of various SPAD applications and the current state-of-the-art in single photon detection. Many applications for single photon detection are based on time-correlated single photon counting (TCSPC) where a probability distribution of single photon detection is generated after an excitation pulse. Through time-correlation, a photon count histogram can be constructed and utilized in a number of powerful applications. With a solid understanding of current challenges in single photon detection and implementations, we will see where improvements can be made and how to address them with novel designs and device structures.

1.2. Applications of Single-Photon Detectors

1.2.1. Molecular Imaging and Single-Molecule Spectroscopy

Molecular imaging and spectroscopy have gained increased relevance in recent years, particularly in studies of biomolecular interactions and dynamics in live cells and tissues. Through the development of confocal microscopes, two-photon microscopes, and fluorescent probes, specific protein interactions and cell functions can be studied. Furthermore, development of such proteins has increased molecular probe applications in monitoring gene expression, detection of protein-protein interactions, membrane potential, and single-molecule detection [26][27]. Green fluorescent protein (GFP), in particular, has had a strong impact on imaging methods in biophysics, neuroscience, and developmental and cell biology [28][29][30].

One of the key components in laser-scanning confocal and two-photon microscope systems is the optical detector [31]. Typically, these systems use laser excitation and wavelength filter-sets for specific fluorophore excitation-emission spectra. Furthermore, the excitation intensity is usually several orders of magnitude greater than the fluorescence emission. Hence, number of photons emitted per fluorophore molecule is low, requiring highly sensitive, low noise optical detectors to obtain good image resolution and contrast, with fast image acquisition speeds [32].

Traditionally, laser-scanning confocal and two-photon microscopes use photomultiplier tubes (PMTs) as optical detectors whose current output is integrated over time and converted to a voltage which corresponds to the number of photons collected on that pixel. When photon flux levels are medium to high, integration methods are acceptable due to sufficient signal-to-noise ratios (SNR) and signal-to-background ratios (SBR).

Detection of single-molecule dynamics has proven to be difficult because only one or a few fluorescent molecules are observed at a given time, leading to low-photon flux rates, short acquisition times, and ultimately poor SNR and poor SBR. Furthermore, photobleaching becomes an issue when averaging over long acquisition times with detectors that exhibit long dead-times. Fast photon-counting detectors are necessary for such applications to overcome poor SNR and to shorten acquisition times.

1.2.2. Time-correlated Single Photon Counting

Time-correlated single photon counting (TCSPC) is a powerful tool for waveform reconstruction when signal levels are low such that single or few photons impinge on the optical detector during each measurement cycle. Using a periodic light source, waveforms are reconstructed by measuring the time-of-arrival of individual photons (Figure 1.1) over many sample periods. This method is particularly powerful in applications with low light levels and high repetition rate signals.

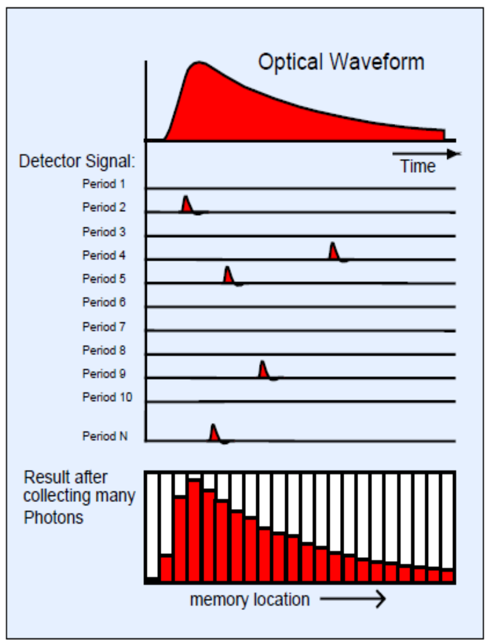


Figure 1.1 Time-Correlated Single Photon Counting [33]

The detector records the time-of-arrival of randomly distributed pulses due to the detection of individual photons. When a photon is detected during the measurement cycle, the detection event is recorded into a time-of-arrival histogram. After many measurement cycles, the target waveform is produced. The important parameters for

TCSPC measurements are sensitivity, timing resolution, accuracy, and sampling speed.

Sensitivity is defined as the light intensity needed for a signal equal to the detector dark count:

$$S = \frac{\sqrt{R_d \cdot \frac{N}{T}}}{Q}$$

Equation 1.1 Sensitivity for TCSPC

where R_d is the dark count rate, N is the number of time channels, Q is the quantum efficiency of the detector, and T is the total integration time. Sensitivity is mainly limited by the dark count rate of the detector.

Timing resolution is of particular importance since single photon counting is now not dependent on the width of the detector impulse response, but on the detection of the leading edge of the pulse. Better timing resolution results in more accurate waveform reconstruction.

Accuracy is essentially the signal-to-noise ratio of the measurement. TCSPC is powerful in that many measurement cycles can be collected till sufficient SNR is achieved. For N number of photons, SNR=$\sqrt{N}$.

Recording speed is limited by the pulse repetition rate and also the detection sampling rate. This is one of the bottlenecks of TCSPC measurements. Faster pulsed laser sources, faster pulse counting cards, and faster detector sampling rates will improve slow acquisition times.

1.2.3. Fluorescence Lifetime Imaging Microscopy (FLIM)

Fluorescent molecules not only have characteristic excitation/emission spectra, but also characteristic lifetime decays. Fluorescence lifetime imaging microscopy (FLIM) is an effective method that produces high contrast images by mapping fluorescence lifetime decays of fluorescent molecules [9-12]. The fluorescent lifetime of the particle is used instead of intensity, so images aren't intensity-dependent, which can be a problem with tissue absorption. A fluorophore is excited and subsequently drops to its ground state with a probability determined by its decay characteristics through radiative and non-radiative processes. Through the spontaneous emission of photons, the fluorophore will emit a photon at a characteristic time after it has been excited with an exponential distribution:

$$F(t) = F_0 e^{-t/\tau}$$

Equation 1.2

and t is time, τ is fluorescence lifetime, and F_0 is initial fluorescence intensity at t=0. Figure 1.2 compares a steady state intensity image and a fluorescence lifetime image demonstrating the increased contrast of looking at lifetime instead of intensity alone.

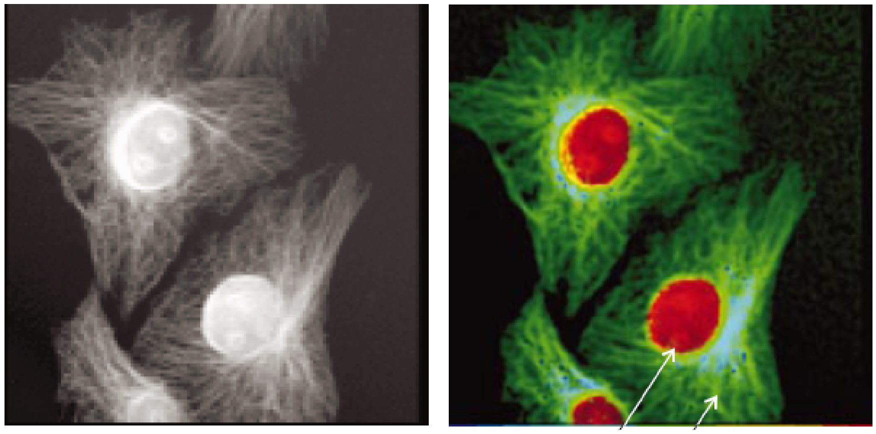


Figure 1.2 Steady State Intensity Image (left) and Fluorescence Lifetime Image (right) [34]

Furthermore, fluorophore lifetime can be affected by local properties, such as pH, ion concentrations, or oxygen saturation by fluorescence quenching. Hence, combining intensity and lifetime information is a powerful tool to differentiate various fluorescent markers or the same fluorophore in different local environments.

To produce high quality images in FLIM, optical detectors with excellent timing resolution are necessary, particularly with sub 100ps fluorescence lifetimes. With environment-sensitive dyes, better detector timing resolution will result in better image contrast with improved lifetime decay measurements. Furthermore, high detection efficiencies are needed to detect few numbers of photons. To limit photobleaching of the fluorophore, fast acquisition times are needed to minimize exposure to the excitation laser light.

1.2.4. Time-Domain Optical Imaging

Time-domain diffuse optical tomography utilizes scattering of light to gain insight of the tissue through which the photons have passed, in particular, solid tumors in the breast or brain [1-5]. It has benefits over other mainstream medical imaging techniques such as MRI, X-ray/CT, and PET since it is non-ionizing, non-invasive, portable, and comparatively low-cost.

Time-domain optical imaging typically uses near infrared (NIR) light because of its deeper penetration depth in tissue [6]. Time-domain optical imaging systems traditionally have a pulse excitation source and detector, either in transmission or reflection geometry. As the laser pulse propagates through the tissue, the photons undergo multiple scattering events, spreading into ballistic photons which travel in an unscattered straight path, snake photons which travel in a zigzag pattern along the straight path, and diffuse photons which undergo multiple random scatter events. Light propagation is governed by the diffusion equation [7]:

$$\frac{1}{c}\frac{\partial\phi(r,t)}{\partial t} - D\nabla^2\phi(r,t) + \mu_a\phi(r,t) = S(r,t)$$

Equation 1.3 Diffusion Equation for Light Propagation

with

$$D = \frac{1}{3\mu'_s}$$

Equation 1.4 Diffusion Coefficient

where $\phi(r,t)$ is the photon fluence rate, D is the diffusion coefficient, μ'_s is the reduced scattering coefficient, μ_a is the linear absorption coefficient, v is the speed of light in the medium, and $S(r,t)$ is the source term.

Light propagation depends on both the tissue absorption and early-arriving photons ballistic photons which arrive earliest because they travel along the straightest path through the tissue, providing the highest spatial resolution in low scattering media. Single photon detectors are useful in time-domain optical systems because of their sensitivity to a low number of photons. With increased detection efficiency of early-arriving photons, more precise optical properties can be obtained, resulting in more accurate forward and inverse propagation models and better image quality. The challenges for single photon detectors in time-domain optical imaging are detection efficiency, timing resolution, and sensitivity at longer wavelengths.

1.2.5. PET/CT

Positron Emission Tomography (PET) and X-ray Computed Tomography (CT) have merged functional and anatomical information together to localize functional information in tissue. PET produces a three-dimensional image by detecting gamma rays emitted indirectly by positron-emitting radiotracer. Image reconstruction is often aided by an X-ray CT scan done on the same machine [35-39].

A radioactive isotope is typically injected in the bloodstream and allowed to circulate. After allowing the isotope to circulate for some time, the patient is placed in the scanner. Through positron emission decay, the radioisotope emits a positron. After traveling a few mm, the positron collides with an electron, which results in annihilation of the two to produce a pair of gamma photons which travel in exact opposite directions. These gamma photons impinge on a scintillating crystal such as LSO or LYSO and create a burst of photons shifted towards UV-visible wavelengths. Optical detectors situated opposite each other on a detector ring look for coincidence photon detections or temporal pairs (Figure 1.3). If two detection events occur on opposite detectors in a given time window (on the order of a few nanoseconds), then the event is recorded. Through these coincidence detection events, image reconstruction is carried out.

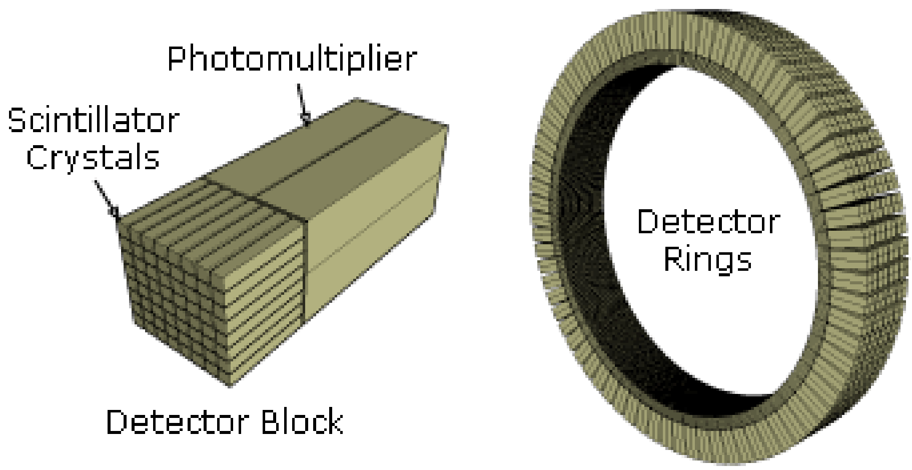


Figure 1.3 Scintillator Detector Block for PET/CT Imaging [40]

PMTs have generally been used, but recently Si Photomultipliers (SiPMs) and avalanche photodiodes (APDs) have been implemented [8]. For good PET/CT images, we must know precisely the location of the gamma interaction point in the scintillator crystal. Reducing the uncertainty at the point of interaction will increase image resolution. It is essential for the single photon detectors to have low noise, high sensitivity, good timing resolution, and good energy resolution to produce high quality PET/CT images. Increased timing resolutions result in reduction of the position uncertainty of the annihilation event.

1.2.6. Classical Optical Communication Links

Single-photon detectors have been used in classical optical links [18-21] (Figure 1.4), particularly with pulse position modulation (PPM) communication links [41-43]. PPM uses optical pulses of uniform height and width and modulates the pulse sequence in time. *M* message bits are transmitted using single-pulses for each bit in one of 2^M possible time-bins. This has advantages over pulse amplitude modulation (PAM) and pulse duration modulation (PDM) because it only requires the detection of the photon event, not the amplitude or duration, which enables better noise immunity. The information that is useful with PPM is the photon's time-of-arrival. Average system power for PPM is much lower than PDM, but at the expense of bandwidth.

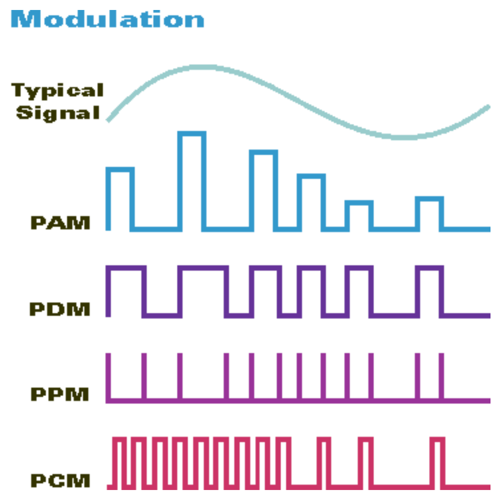


Figure 1.4 Optical modulation techniques [44]

PPM is also used extensively in deep-space communications for both uplink and downlink transmissions [45],[46] because high peak-powers can be used in short duration pulses, providing good SNR over background radiation and thermal noise. Timing-resolution or jitter of the optical detector is one of the main limiting factors in transmission bit-rates. Full-width-at-half-maximum (FWHM) is generally used as the metric for measuring timing-resolution. However, because of single photon counting detector pixel structures and architectures, a so-called diffusion tail, described in Section 2.5.6, increases the timing uncertainty of the photons arrival time, particularly at the full-width-at-hundredth-maximum (FW(1/100)M) and the full-width-at-thousandth-maximum (FW(1/1000)M). The reduction of this long tail serves to provide potential benefits in increasing bit-rates and reduction of bit-errors via inter-symbol interference.

1.2.7. Quantum Key Distribution (QKD)

Quantum Key distribution (QKD) was developed for secure communication links using quantum mechanics. With a proper "key," two parties can communicate securely using a shared random bit string to code and decode bit strings. The security in QKD systems relies on fundamentals of quantum mechanics which exposes the presence of an eavesdropper interfering with the communication link. More specifically, the measuring of a quantum system disturbs it, resulting in increased bit-error-rates. Significant bit-errors present in the system indicate the presence of a third-party eavesdropper and the communication link is subsequently closed or moved.

There have been several protocols developed for QKD systems, notably BB84 [47] which originally utilized photon polarization states. In typical quantum cryptography systems, Alice and Bob are denoted as the sender and receiver, respectively, and are transmitting to each other through a quantum communication channel, usually through optical fiber or free-space. Quantum communication involves the transmission of quantum states (qbits) usually in the form of single photons with polarization [22],[48],[49],[23].

There has been movement towards increasing key generation rates and distribution over longer distances. Bit-errors caused by detector dark counts are one of the main limiting factors in key-distribution distances. Two effective approaches to increasing distribution distances and reducing bit-error-rates are to reduce detector dark count rates and increasing system timing resolution (jitter). Reducing detector dark counts reduces the number of falsely recorded bits, resulting in increased bit-error-rates. Increasing timing resolution (reducing timing jitter) allows for increased bit transmission rates and faster key generation.

1.2.8. Laser Detection and Ranging (LIDAR)

Laser detection and ranging (LIDAR) is often associated with military applications, though the concept is used in archaeology, geology, astronomy, biology, imaging, and 3D mapping [14-17]. LIDAR measures scattered light to find information of distance targets, usually through active illumination with laser pulses. It is also referred to as laser radar, since laser light is used instead of radio waves as in conventional radar.

LIDAR was developed to overcome limitations of conventional radar, particularly for non-metallic objects which don't readily reflect microwave/radio frequencies and can be considered "invisible" to radar. Using active illumination with pulsed laser sources, light is scattered off the target and detected with an optical detector. This scattered light is termed "backscatter," and based on Rayleigh, Mie, and Raman scattering.

The ability to focus and collimate laser beams provides excellent spatial resolution while the ability to pulse a laser for short time-durations provides excellent timing resolution. Backscattered light is then collected and focused onto a detector pixel or array, usually through a scanning system, and time-of-flight information is recorded. Since few of the backscattered photons can be expected to reflect back to the detector, particularly at long distances of tens of kilometers, or into deep space, high detection efficiencies are necessary for optimal LIDAR system design. Furthermore, reducing timing-uncertainty of the photons' arrival time allows for increased distance precision. Moreover, tracking moving targets requires fast acquisition and short data processing times.

1.3. State-of-the-art Single Photon and Low-light Level Detectors

1.3.1. Photomultiplier Tubes and Micro-Channel Plates

Photomultiplier Tubes (PMTs) are vacuum tubes which utilize the photoelectric effect and internal gain mechanisms to produce a current in response to low light levels [50]. Incident photons on the photocathode produce electrons. These electrons are accelerated and focused on the first dynode and multiplied via secondary electron emission. These secondary electrons are accelerated towards the second dynode and subsequently multiplied. After multiple stage dynode amplification, the electrons are collected by the final anode and the current is read out.

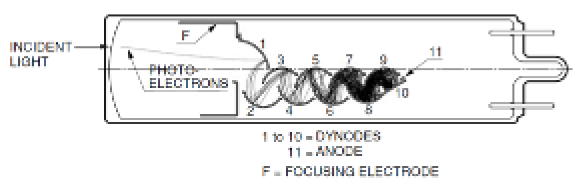


Figure 1.5 Construction of a photomultiplier tube [50]

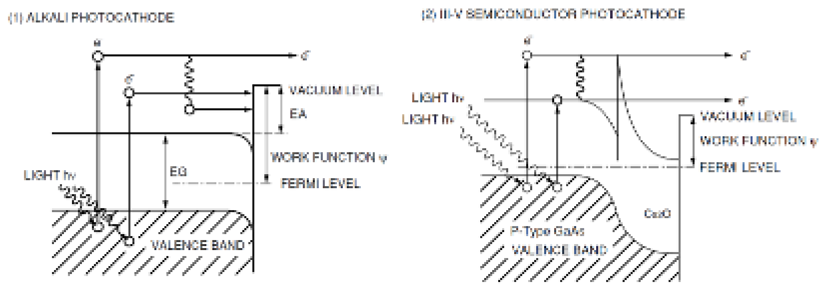


Figure 1.6 Band diagrams for alkali (1) and III-V semiconductor (2) photocathodes

The photocathode is a semiconductor, typically alkali or a III-V compound semiconductor, and can be characterized with band models (Figure 1.6). The quantum efficiency of the PMT is determined by the efficiency of converting photons into detectable electrons and is described by the probability process:

$$\eta(\upsilon) = (1-R)\frac{P\upsilon}{k}\cdot\left(\frac{1}{1+1/kL}\right)\cdot Ps$$

Equation 1.5 Quantum efficiency

where *R* is the reflection coefficient, *k* is the full absorption coefficient of photons, *P* is the probability that light absorption may excite electrons to a level greater than the vacuum level, *L* is the mean escape path length, *Ps* is the probability that electrons reaching the photocathode surface may be released into the vacuum, and *v* is the frequency of light. Spectral response of PMT materials is typically most suited for UV-visible with a steep drop-off in the NIR/IR wavelengths.

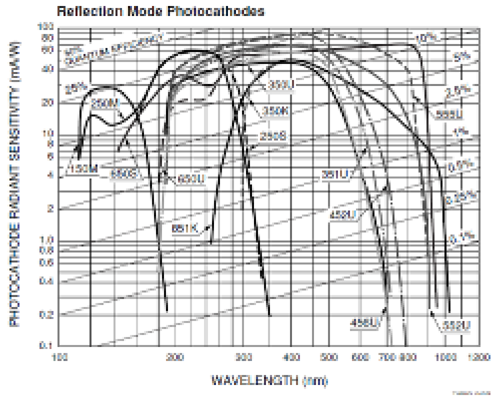


Figure 1.7 Typical spectral response characteristics of reflection mode photocathodes [50]

Dynode design has a strong impact on the timing spread of the electrons incident on the final anode and becomes the limiting factor in the PMT timing resolution. As soon as the large number of electrons hits the anode, the charge must be quickly removed to prevent surface-charge effects.

More recently, microchannel plates (MCPs) have been used to improve PMT performance by replacing conventional discrete dynode stages, offering greater timing resolution down to picoseconds as well as photon counting modes. A MCP is made up of a 2D array of glass capillary tubes bundled to form a thin disk (Figure 1.8) [50]. They are advantageous over discrete dynode stages in that they exhibit more compact sizes, faster time response, two-dimensional detection with high spatial resolution, can operate in high magnetic fields, are sensitive to charged particles, UV, x-rays, gamma rays, and neutrons. Furthermore, they have low power consumption.

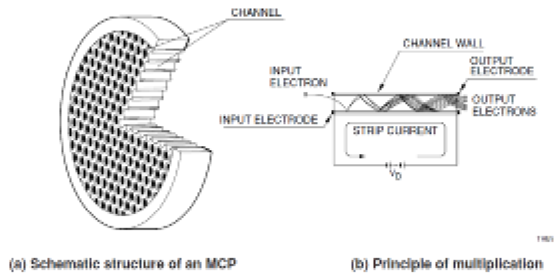


Figure 1.8 Schematic structure of MCP (a) and its method of multiplication (b) [50]

For medium-high light levels, the PMT integrates the current for adequate SNR. At low photon flux levels, integration of current is unsuitable due to poor noise performance with multiplication noise, thermal noise, and afterpulsing. In this regime, different detection methods must be used, such as single photon counting, to overcome poor SNR. Cooled photon counting PMTs have been developed to reduce thermally generated dark current, though their long dead-times have prevented them from being used more widely in microscope systems. Nonetheless, PMTs remain the technology of choice in many imaging systems because of their maturity and availability. However, because they are fragile, are not scalable to large arrays, and have poor spectral response at longer wavelengths, other technologies may be more suitable for certain applications.

1.3.2. Charged-Coupled Devices

Charged-Coupled Devices (CCDs) are silicon-based devices that convert light energy into electronic charge. These photogenerated electrons are accumulated in a potential well and subsequently transferred across the chip through registers and output to an amplifier. The transfer gate applies a voltage to block the stored charges from overflowing, and when the integration time is complete, the transfer gate unblocks the potential wells, allowing the stored charged be read-out.

CCDs suffer from different noise mechanisms, namely shot noise, dark current noise, photoelectron noise, and read-out noise. Noise sources such as shot noise and read-out noise can be reduced by increasing integration times. Dark current can be reduced by either using higher quality silicon with fewer defects or by cooling. Highly sensitive CCDs are generally cooled to reduce the effects of dark current.

CCD have historically provided better image quality over CMOS sensors, but through recent advancements in CMOS sensor technologies, the gap in image quality has shrunk considerably. Other variations of the CCD for low light level detection include the electron-multiplying CCD (EMCCD), which utilizes a gain register in between the shift register and the output amplifier. In each gain register, amplification is achieved through impact ionization, similar to an avalanche photodiode. Because of this, read-out noise becomes negligible, but it introduces amplification noise which is by nature a statistical process.

Similarly, the intensified CCD (ICCD) has been developed for better sensitivity in low light situations by implementing an image intensifier (such as an MCP) on the CCD through a phosphor screen and fiber-optic bundle [51]. ICCDs offer advantages in improved sensitivity, improved SNR, and the ability for time-gated operation. However they suffer from drawbacks such as MCP noise at high gains, reduced MCP sensitivity at NIR/IR wavelengths, variable MCP sensitivity at short time-windows, low spatial resolution, and are fragile and bulky.

1.3.3. Silicon Photomultipliers (SiPM)

Silicon photomultipliers (SiPM) have been developed as an alternative to PMTs and are silicon devices that typically consist of 100-1000 parallel avalanche photodiodes (APDs) on a single chip (Figure 1.9) [52-56]. An avalanche in any one of the diodes will result in a current pulse at the output. If multiple avalanche events occur simultaneously, the output will be an integration of these pulses, corresponding to the number of incident photons on the detector.

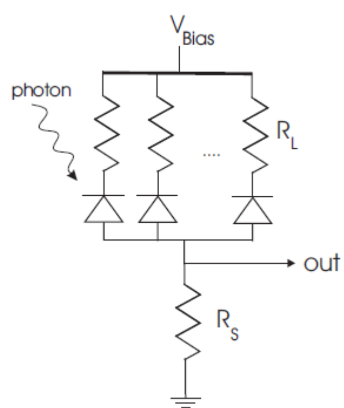


Figure 1.9 Schematic of SiPM [53]

SiPM have shown great promise for nuclear medicine applications, in particular with PET/CT systems. Because of their high gain, low noise levels, insensitivity to magnetic fields, and large areas, SiPM are rapidly becoming the detector of choice in PET systems. Furthermore, because of their fast timing properties

(<100ps jitter), they become particularly suitable when time-of-flight information is needed. However, fast signal read-out remains an issue due to the lack of circuitry integrated on-chip, requiring high performance ASICs and electronics [53].

1.3.4. Superconducting Nanowire detectors

NbN based superconducting nanowires have been developed as single photon detectors [57-59] and have shown promise in delivering sub 30ps timing resolutions with greater than 5% detection efficiencies at 1550nm. Each detector consists of a nanowire, cryogenically cooled below the critical temperature to 2-4K to reach superconductivity. When a photon is absorbed in the nanowire, a local hot spot is created where superconductivity is disrupted. The hot spot creates a local increase in resistivity resulting in a measureable voltage drop. This technique provides extremely good timing resolution, of less than 30ps, comparable to the best detection technologies developed.

Resetting of nanowire detectors requires cooling of the hot spot back to superconductivity, which is by nature a slow process, which results in slow detector recharge times. Furthermore, the only know technique for patterning these nanowire structures is electron beam lithography, which is a serial technique, preventing scalability towards large scale integration at the present time.

1.4. Single-Photon Avalanche Diodes

Avalanche diodes are highly sensitive semiconductor detectors that use an internal gain mechanism to convert photons into an electrical signal. This is done through high-electric fields that initiate impact ionizations and create a so-called avalanche of charge that flows through the diode. The physical processes for photoabsorption and charge multiplication will be described in Chapter 2. I will describe state-of-the-art avalanche photodiodes (APDs) and Geiger-mode single photon avalanche diodes (GM-SPADs) in the next subsections.

1.4.1. Avalanche Photodiodes (APDs)

Avalanche Photodiodes (APDs) are semiconductor-based photodetectors that utilize an internal gain mechanism through applying a reverse bias voltage. Because of this internal gain mechanism, APDs are useful in detecting low light levels and used in a variety of applications. APD structures are typically p-n junctions or p-i-n junctions (Figure 1.10) [60].

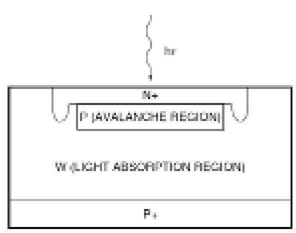


Figure 1.10 Cross-section of an APD [60]

When a photon enters the photodiode, an electron-hole pair is generated if the energy of the photon is greater than the bandgap of the material. When the electron-hole pair is generated in the depletion region of the p-n junction, they are subsequently separated from each other by the electric field generated by the reverse bias voltage, with electrons drifting towards the n+ side and holes drifting towards the p+ side. Because of the high electric field, the electrons and holes carriers are accelerated and can collide with the crystal lattice with enough energy to create a secondary electron-hole pair. This phenomenon is termed impact ionization. These secondary electron-hole pairs can generate subsequent electron-hole pairs in a so-called avalanche process. Impact ionization begins at electric field strengths of $>10^5$ V/cm [60].

APDs suffer from various noise sources such as shot noise, multiplication noise (excess noise factor), photoelectron noise, and dark current. The physics of each of these noise sources will be described in more detail in Chapter 2. Silicon APDs are sensitive from ~450nm to 1100nm, with peak quantum efficiency in the 600-800nm range and low multiplication noise. Germanium (Ge) APDs are sensitive up to 1.7um, but suffer from high multiplication noise. InGaAs APDs are sensitive up to 1.6um and have less multiplication noise than Ge APDs. Furthermore, InGaAs APDs have high absorption at telecom wavelengths at 1550nm. Unfortunately, APDs do not have sufficient gains for single photon detection. Further, they exhibit poor timing resolutions because of their avalanche multiplication mechanism.

1.4.2. Geiger-mode Single Photon Avalanche Diodes (GM-SPAD)

Similar to an APD, a Geiger-mode single photon avalanche diode (GM-SPADs) utilizes impact ionization as its internal gain mechanism. Much of the pioneering work with the avalanche phenomenon in p-n junctions came from Shockley laboratories in the 1960's [61],[62]. P-n junctions were designed and fabricated to have uniform behavior over a small area and reverse biased above the breakdown voltage. The physical phenomenon of avalanche breakdown was studied in great depth by Haitz et al [62-64]. Furthermore, McIntyre and Webb observed generation of single-photon induced pulses when biased above breakdown [65].

GM-SPADs operate at a reverse bias above the breakdown voltage, whereas APDs operate below the breakdown voltage. When biased above the breakdown voltage, by definition, electron-hole pairs are generated on average faster than they can be extracted. Hence, when a photon impinges on the photodiode and is absorbed in the depletion region, an electron-hole pair will be generated, and a self-sustaining runaway avalanche process will be occur. Immediate quenching of the runaway charge is necessary to prevent damage to the photodiode and this forms a short current pulse indicating the detection of a single photon. This avalanche pulse acts as a binary signal, whether or not a photon was detected, unlike the analog signal obtained from electron multiplication in APDs. Quenching can be achieved through a simple passive quenching resistor or through more complex active quenching circuits. Important SPAD figures-of-merit include detection probability, spectral response, dark count rate, timing precision (jitter), sampling speed, active area and fill factor. Sampling speed and timing jitter are of particular importance with time-of-flight applications with which SPADs excel. These will be described in more detail in Chapter 2. In the following subsections, I will describe previous implementations of solid-state single photon detectors as well as current state-of-the-art.

1.4.2.1. SPAD devices fabricated in custom processes

Solid-state SPAD devices were originally developed in custom fabrication processes and have shown good performance particularly with detection efficiencies, low jitter, and tailored spectral responses. There have been two major implementations with such custom fabricated devices: reach-through bulk devices (Figure 1.11) [66-68] and surface devices [61],[69],[70]. Early SPAD implementations were designed by Haitz et al. and McIntyre and Webb. McIntyre and Webb developed a SPAD in a dedicated reach-through process technology with ultralow doped p-type silicon substrates (Figure 1.11). Their devices had thick depletion regions of 20-100µm and high breakdown voltages of 100-500V, resulting in good detection efficiencies, but poor timing jitter. Improvements to this first reach-through SPAD was implemented in Perkin-Elmer's (formerly RCA ElectroOptics) Slik™ device [66]. This device was also manufactured in a dedicated process with large active areas of 200µm and high detection efficiencies and low dark count rates of 150 counts per second (cps). However, they require high operating voltages of 300-400V and due to their large diameter, they suffer from poor timing resolutions of >300ps FWHM.

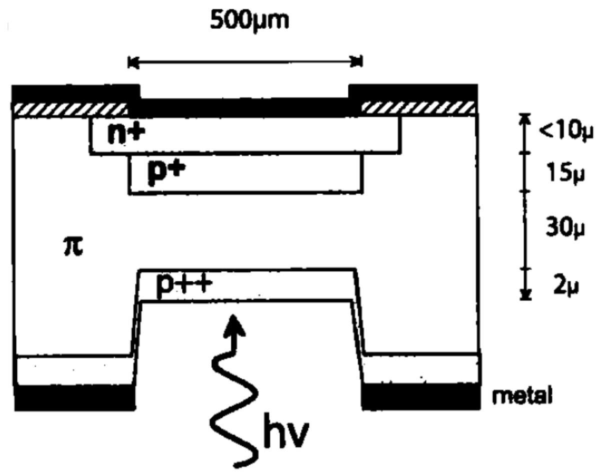


Figure 1.11 Cross-section of reach-through SPAD developed by McIntyre et al. [65],[71]

Further, Haitz et al. developed devices using a planar technology with a diffused guard ring structure (Figure 1.12) to prevent premature edge breakdown. The diffused guard ring is a lighter doped region where the electric field at the curved regions is spread out over a large area, thereby eliminating premature edge breakdown. They utilize a thin junction depletion layer and with a low breakdown voltage of 15-50V. The diffused guard ring structure has become the guard ring structure of choice in current SPAD architectures in commercial and non-commercial process technologies.

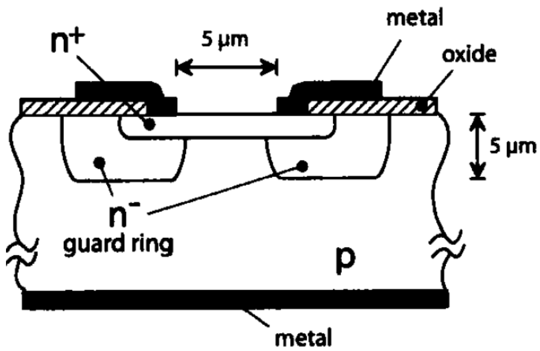


Figure 1.12 Cross-section structure of planar SPAD developed by Haitz [61],[62]

Cova et al. developed a planar SPAD in a double-epitaxy process to overcome drawbacks of the diffused guard ring structure (Figure 1.13) [69]. The epi-substrate p-n junction limits the neutral region from where electron-hole pairs are collected, thereby reducing the associated diffusion tail in timing resolution. These devices had 1µm thick depletion regions and exhibited good detection efficiencies of 40% at 500nm with a breakdown voltage of 20V. Further, they exhibited low dark count rates at room temperature of a few kcps or lower with little afterpulsing.

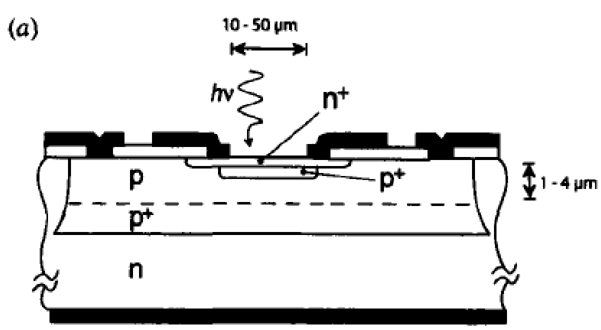


Figure 1.13 Cross-section of Cova SPAD designed with double-epitaxy [69]

Surface SPAD implementations have a few drawbacks, namely reduced detection efficiencies due to thinner depletion regions and spectral responses shifted towards shorter wavelengths due to shallow junction depths. However, they succeed in reduced jitter, reduction in dead-times due to smaller junction capacitances, reduction in breakdown voltages, and the potential for large-scale integration with on-chip processing circuitry. As noted earlier, Haitz's planar diffused guard ring structure has become the basis for many SPAD implementations in standard CMOS technologies today.

1.4.3. High Voltage Commercial CMOS

CMOS technologies are not specifically designed for operation above breakdown voltages, so implementation of SPADs in CMOS is challenging. Further, implementing SPADs in commercial CMOS technologies requires adherence to fixed process design rules. In 2002, Rochas demonstrated implementation of a SPAD with a similar structure to Haitz's, utilizing a diffused guard-ring structure in an Austria Microsystems 0.8μm high-voltage CMOS (HV-CMOS) technology (Figure 1.14) [72],[73].

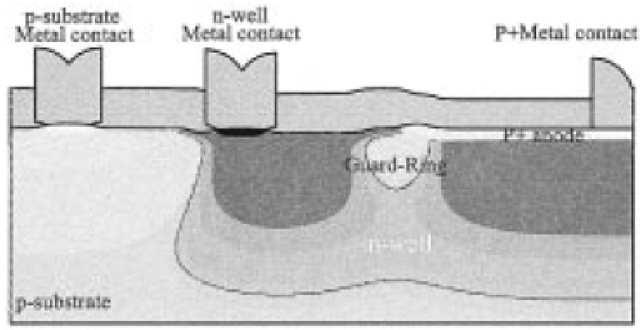


Figure 1.14 Doping profile of Rochas design implemented in HV-CMOS process [73]

Doping profiles were inverted compared to Haitz's design to be compatible with the p-substrate used in CMOS processes. A shallow p+ source/drain implant and n-well form the diode terminals. To make contacts to the n-well terminal, a triple-ring structure, consisting of a deep n-well, n-well, and n+ contact implants were used. The diffused guard-ring was implemented using a p-well structure.

The 0.8µm CMOS Si-gate technology offers two metal layers, two poly layers, and requires 2.5V and 5.5V VDD. Because of the diffused guard-ring, higher reverse bias voltages are possible without inducing diode damage, allowing for operation above breakdown at 25V. Because of the shallow junction of 0.3µm, peak detection efficiency was 30% at 450nm. The Rochas design exhibited low dark count rates of 100cps for a 7µm diameter active-area, dead-times of 30-50ns, and a 60ps FWHM jitter, but with a long 2ns FW(1/100)M diffusion tail.

Following Rochas's first CMOS implementation, several groups have implemented SPAD designs in other HV-CMOS processes. Charbon et al. have developed SPADs in 0.35µm HV-CMOS processes for applications in biometrics and biological applications [74-76]. A number of other groups have designed CMOS SPADs with diffused guard-ring structures with integrated circuitry targeting specific applications, also in high-voltage processes [77],[78].

1.4.4. Low voltage Commercial CMOS

Recently, there has been movement towards implementation in low voltage CMOS technologies with smaller feature sizes of 0.18µm and smaller, pioneered at UCSD by Hod Finkelstein and me [79] that utilizes a novel Shallow Trench Isolation (STI) guard ring structure. This design was implemented to address drawbacks of the triple-well diffused guard ring structure, namely:

1) A diffused guard ring suffers from poor fill factors, because much of the pixel is dedicated towards the p-well and n-well guard-rings, while the active area is only at the p+ to n-well interface. Low fill factors from the pixel alone result in poor sensitivity of the photodetector when quantum efficiency and fill factor are taken into account together.
2) The diffused guard ring is also very wide and still exhibits an electric field, albeit weaker than the high-field region. Nonetheless, electron-hole pairs are capable of drifting towards the high-field region and initiate impact ionization, resulting in greater timing spread. This is typically characterized by a long diffusion tail in the jitter distribution, and becomes detrimental in applications such as fluorescence lifetime imaging and pulse position modulation links.

Chapters 3 will describe the design, simulation, and characterization of the UCSD designed SPAD in a standard low-voltage 0.18µm CMOS process, the first of its kind, and how its novel STI guard ring structure addresses the above challenges of the triple-well diffused guard ring structure for SPADs.

Following the design of our prototype SPAD test chip, various CMOS SPADs and SPAD arrays have been fabricated in other low voltage 0.18µm CMOS processes [80-82]. Further, recently implementations in 130nm processes have been realized [83],[84], as well as in 130nm CMOS image sensor technologies [85], which utilizes a an STI-guard ring structure with passivation for improved performance characteristics (Figure 1.15, Figure 1.16). Further, Richardson et al. have devised a method that uses both STI and a diffusion ring with promising performance [80] (Figure 1.17).

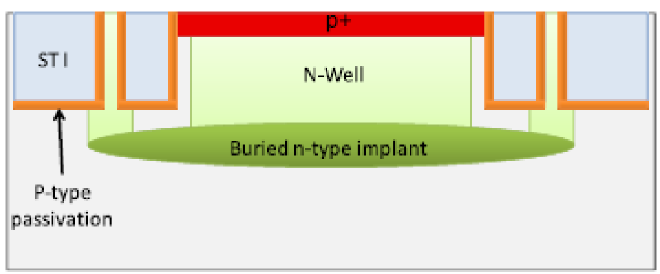


Figure 1.15 Cross-section of STI-bound SPAD implemented in 130nm CMOS process with p-type passivation [85]

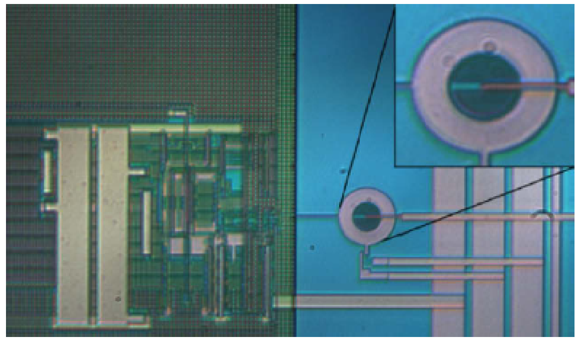


Figure 1.16 Photomicrograph of STI-bound SPAD implemented in 130nm process [85]

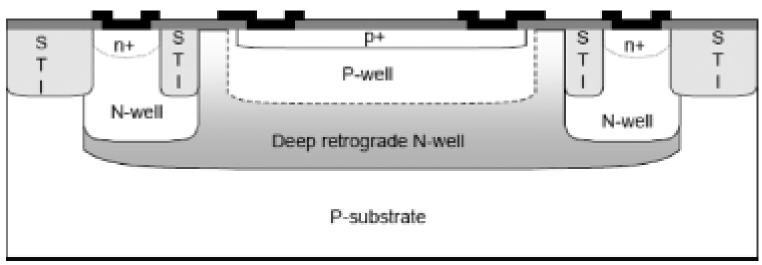


Figure 1.17 SPAD Cross-section implemented with STI- with diffusion ring by Richardson et al. [80]

1.5. Aims and Challenges

Single photon detectors have been developed for various applications in biology, medicine, communications, and distance ranging, and are continuing to improve in performance and cost. The choice in the type of detector is crucial for optimal performance for each application. CMOS SPAD detectors stand to improve in areas where traditional single photon detectors, such as PMTs, lack, namely robustness, size, scalability, power requirements, cost savings, timing response, room temperature operation, and low noise. Moreover, detector performance and its surrounding integrated circuitry can be tailored for optimized performance in various applications. Figures of merit include detection efficiency, spectral response, dark count rate, timing resolution, dynamic range, robustness, and power dissipation.

Custom dedicated SPAD fabrication processes have produced SPADs with good performance, but they have shortcomings in scalability, large junction capacitances, afterpulsing, in part due to their inability to integrate on-chip processing circuitry.

This dissertation develops and characterizes a novel SPAD device, with an emphasis on noise characteristics, spectral response, and timing jitter. The following areas will be explored to address the design and characterization challenges:

1) Electric-field simulations to determine electric-field distribution over planar and curved junctions. Knowing the electric-field distribution over the whole device will help with proper device design for planar breakdown.
2) Circuit simulations for optimal fast read-out circuitry are required for proper electrical read-out of ultrafast photo-induced pulses and compatibility with pulse-counting acquisition boards.
3) Characterization setups for detection efficiency, spectral response, and timing jitter. It is of the utmost importance to know the exact number of photons incident on the photodiode to make accurate calculations of the above metrics. Furthermore, it is of great importance to have the correct data acquisition cards and oscilloscopes with enough precision and bandwidth to measure the high-speed nature of the SPAD test devices.

Furthermore, after design and characterization, various SPAD applications will be explored for frequency up-conversion, acousto-optic imaging enhancement, optical characterization of ultrasound contrast agents, and fluorescence modulation using dye-loaded microbubbles.

1.6. Dissertation Outline

The outline of this dissertation is as follows. Following the overview of single photon detection and the state-of-the-art in low light level optical detection, we outline the background physics of solid state single photon avalanche diodes in Chapter 2. Device operation and various figures of merit will be described. Chapter 3 will describe advances of a UCSD designed CMOS Shallow Trench Isolation (STI) SPAD from device concept, electric-field simulations, circuit simulations, and a complete device characterization. Also in chapter 3, we will describe proof-of-concept of a frequency up-conversion hybrid device for detection of 1550nm light. In Chapter 4, we describe work done in the area of acousto-optic signal detection with various optical detectors, including the CMOS SPAD, and signal enhancement with ultrasound microbubbles. In Chapter 5, the design of a light scattering system is described and utilized to characterize the acoustic response of various ultrasound sensitive microbubbles and particles and detection of fluorescence modulation with dye-loaded microbubbles. We conclude this dissertation with a summary and outlook for future research directions for the work described here.

2. PHYSICS AND OPERATION OF SINGLE PHOTON AVALANCHE DIODES

2.1. Introduction

In this chapter, we will describe the physics and operating principles of solid-state single photon avalanche diodes. We will understand avalanche breakdown mechanisms in semiconductor p-n junctions and examine how curvatures in the junction cause premature avalanche breakdown. Further, we will explain how guard rings are used to prevent premature breakdown. Following, we will examine avalanche quenching and junction recharge methods. Lastly, we will identify the key figures of merit for SPAD performance. Knowledge of the physics of SPAD operation will allow a solid understanding of the STI-bound CMOS SPAD that we have developed and characterized.

2.2. Basic Properties of the p-n Junction

2.2.1. P-n Junctions at Thermal Equilibrium (zero applied voltage bias)

The p-n junction is the basis for semiconductor SPAD devices. It is formed by joining a p-type and n-type semiconductor in contact with one another forming a junction at their interface. Often, p-n junctions are formed on a single substrate by doping via diffusion of dopants, ion implantation, or grown with epitaxy. When a p-n junction is formed, a space-charge region is formed at the junction interface. This is attributed to electrons from the n-region diffusing towards the p-region leaving behind positively charged donors in the n-region. Likewise, holes in the p-region diffuse towards the n-region leaving behind negatively charged acceptors (Figure 2.1). These positive and negatively charged regions form a space-charge region or depletion layer (Figure 2.2) [86].

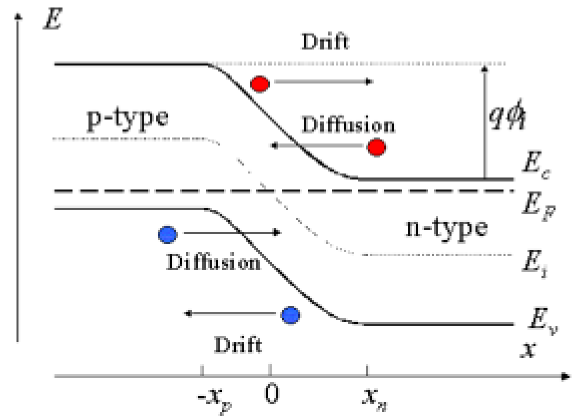


Figure 2.1 Energy band diagram of p-n junction at thermal equilibrium [87]

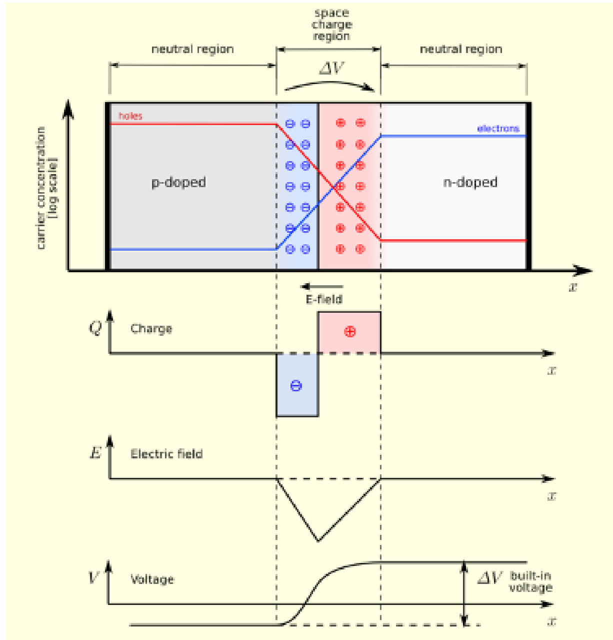


Figure 2.2 A p-n junction at thermal equilibrium with zero bias voltage. Plots for carrier concentration, charge distribution, electric field, and voltage [86].

Derivations for space-charge densities, electric field distribution, and built-in voltage are not in the scope of this dissertation and can be found in Sze [88].

2.2.2. Reverse-biased p-n Junction

Under reverse bias, the p-type region is attached to the negative terminal, while the n-type region is attached to the positive terminal. Because of this bias configuration, holes in the p-region are pulled away from the junction while electrons in the n-region are likewise pulled away from the junction. This results in a widening of the depletion region. An increased potential barrier proportional to the difference between the Fermi energy in the n-type and p-type quasi-neutral regions is created resulting in high resistance to charge flow thus minimizing the current that can flow across the junction. This is illustrated in the energy band diagram (Figure 2.3).

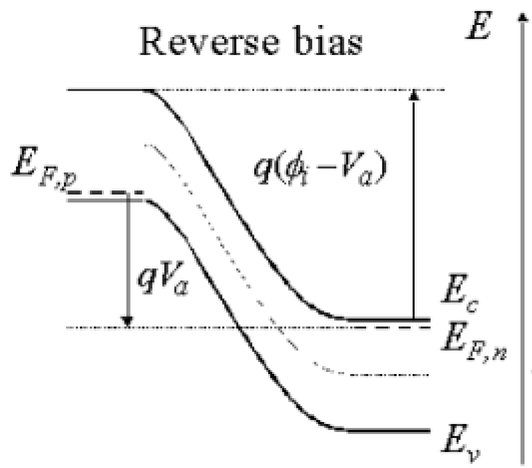


Figure 2.3 Energy band diagram for a reverse-biased p-n junction [87]

The strength of the electric-field in the depletion region increases with reverse bias voltage. When the electric field is increased beyond a critical level, the p-n junction will undergo breakdown and current will flow. Breakdown can be due to a number of processes, namely thermal instabilities, Zener breakdown (tunneling) or avalanche breakdown, which will be described in the next subsection.

2.3. Breakdown Mechanisms in p-n Junctions

The strength of the electric field in the depletion region increases with increased applied reverse bias voltage. There is a critical point where the electric-field strength is sufficiently strong to induce breakdown in the p-n junction and current begins to flow. There are a few physical mechanisms attributed to junction breakdown, namely thermal instabilities, Zener effect (tunneling breakdown), and avalanche breakdown. These breakdown processes are non-destructive and reversible as long as the current flow does not reach levels where thermal damage can occur. It is in this regime in which solid-state SPADs operate and require quenching to stop the runaway current flow, usually with an in-series current limiting resistor. Quenching mechanisms will be described in Section 2.5.2**.**

2.3.1. Thermal Instability

At high reverse-bias voltages, high current flow causes temperature increases in the p-n junction. This temperature increase increases the reverse current that flows through the junction [88]. Because of the heat dissipation, a negative differential resistance is observed. Further, the current varies with $T^{3+\gamma/2} \exp^{-\frac{E_g}{kT}}$. This effect is more prominent with semiconductors with relatively small bandgaps, such as Ge.

2.3.2. Zener Breakdown (Tunneling)

Quantum mechanical tunneling of carriers through the bandgap is the dominant breakdown mechanism for highly doped p-n junctions. An electron with kinetic energy less than the barrier potential energy can tunnel from one potential well to another (Figure 2.4).

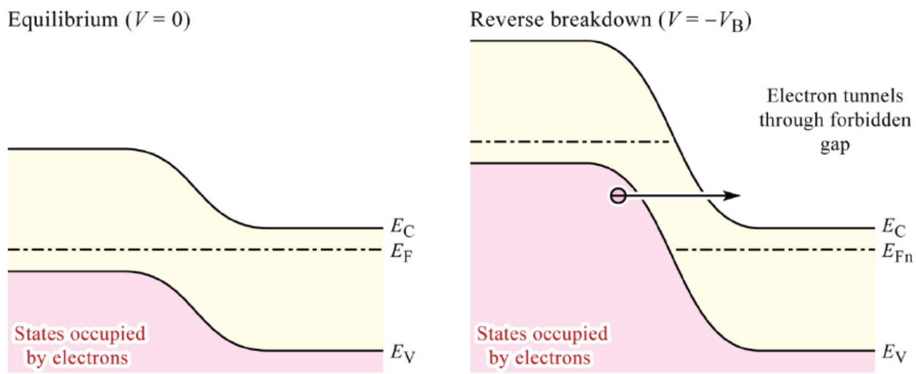


Figure 2.4 Junction breakdown through electron tunneling [89]

This is based on Heisenberg's uncertainty principle:

$$\Delta E \Delta t \leq h$$

Equation 2.1 Heisenberg's Uncertainty Principle

If the product of the energy barrier height and the time required for tunneling is smaller than Planck's constant *h*, then tunneling will occur. This will occur when the barrier height ΔE is low or if the tunnel barrier thickness is thin. A large number of filled states and a large number of empty states separated by a potential barrier are necessary for tunneling breakdown. The depletion region width is related to the doping profile by:

$$W_D \propto \left(\frac{1}{N_A} + \frac{1}{N_D} \right)^{1/2}$$

Equation 2.2 Depletion region width

It follows that heavy doping makes a narrow depletion width. Lightly doped p-n junctions also exhibit tunneling, albeit at very high reverse-bias voltages. Tunneling current density is given by Equation 2.3 with a tunneling probability equal to the exponential term:

$$J_t = \frac{\sqrt{2m^*}q^3EV}{4\pi^2\hbar^2E_g^{1/2}}\exp\left(-\frac{4\sqrt{2m^*}E_g^{3/2}}{3qE\hbar}\right)$$

Equation 2.3 Tunneling current density [88]

2.3.3. Avalanche Breakdown (Impact Ionization)

Avalanche breakdown or impact ionization is the operating principle for avalanche photodiodes and Geiger-mode single photon avalanche diodes, and has been studied extensively [61],[64],[90-94]. With sufficiently high electric-fields, electron and hole carriers are accelerated to a speed that upon collision with the crystal lattice, an electron is kicked out of its orbit thereby ionizing the atom and creating a secondary electron-hole pair. These secondary electron-hole pairs will create further electron-hole pairs by impact ionization leading to carrier multiplication and high current flow through the junction (Figure 2.5). When electron-hole pairs are continually generated in a self-sustaining manner, avalanche breakdown occurs.

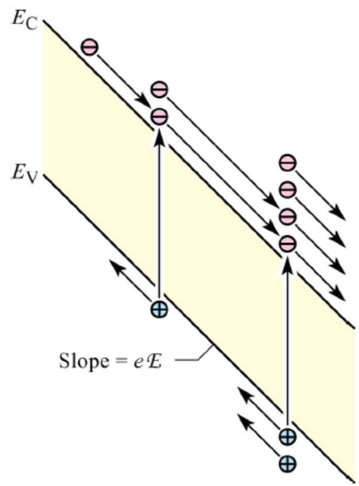


Figure 2.5 Band diagram of impact ionization in a reverse biased p-n junction [89]

The electron-hole pair generation rate during impact ionization is given by:

$$G = \alpha_n n \upsilon_n + \alpha_p p \upsilon_p$$

Equation 2.4 Electron-hole pair generation rate [88]

where α_n is the electron ionization rate (the number of electron-hole pairs generated by an electron per unit distance) and α_p is the hole ionization rate ((the number of electron-hole pairs generated by a hole per unit distance). Both α_n and α_p are strongly dependent on electric field and can be expressed as:

$$\alpha(E) = \frac{qE}{E_I} \exp\left\{- E_I / \left[E(1 + e / E_p) + E_{kT}\right]\right\}$$

Equation 2.5 Ionization Coefficient [88]

where E_I: is the high field, effective ionization threshold energy

E_{kT}: thermal scattering threshold fields

E_p: optical-phonon scattering threshold fields

E_I: ionization scattering threshold fields

Knowing the ionization rates of electrons and holes is necessary to choose the optimal semiconductor material. For silicon, electrons are much more likely to initiate impact ionization than holes which leads to low multiplication noise for APDs. Photodiodes tend towards materials with either strong ionization rates of electrons or holes, but not both. Figure 2.6 illustrates two hypothetical materials where only electrons can initiate impact ionizations (top) electrons and holes initiate impact ionization (bottom) [68]. Both diagrams have the same average gain. However, in the top graph, with many electrons in the high-field region, many impact ionizations occur in parallel, making up for any electrons that don’t have sufficient energy to generate a secondary electron-hole pair. In the bottom diagram, where both electrons and holes are equally effective generating secondary electron-hole pairs, the ionizations will tend to grow in series because of the equal probability of an electron or hole generating a secondary pair. Though the overall gain is the same, this takes a much longer time, which leads to statistical variations in the gain resulting in greater multiplication noise. Hence it is necessary to choose proper materials for optimal gain-bandwidth characteristics.

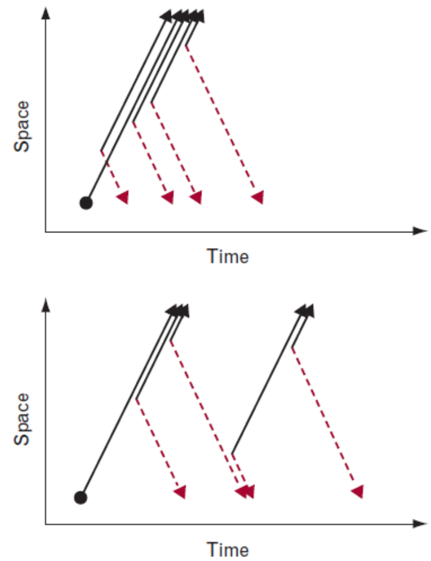


Figure 2.6 Hypothetical impact ionization diagram. Top, only electrons initiate impact ionization. Bottom, both electrons and holes initiate impact ionization equally [68].

Analytical expressions for avalanche multiplication and junction breakdown for an abrupt p-n junction are derived by Sze [88]. The junction breakdown voltage can be obtained numerically and can be represented with the following expression with known boundary conditions for an abrupt one-sided junction:

$$V_B(breakdown\_voltage) = \frac{E_m W}{2} = \frac{\varepsilon_s E_m^2}{2q}(N_B)^{-1}$$

Equation 2.6 Breakdown voltage for abrupt one-sided junction

An approximate expression for planar junction breakdown for semiconductors with bandgap E_g (eV) at room temperature and N (cm^{-3}) background doping is given by [88]:

$$V_B \cong 60\left(E_g/1.1\right)^{3/2}\left(N_B/10^{16}\right)^{-3/4}$$

Equation 2.7 Approximation for breakdown voltage for an abrupt planar junction

2.3.4. Avalanche Breakdown in Non-planar Junctions

Cylindrical or spherical junctions exhibit higher electric field densities and higher current densities, leading to lower breakdown voltages at these junctions compared to planar junctions. The voltage potential and electric field at non-planar p-n junction scan be calculated from Poisson's equation, following Sze [88]:

$$\frac{1}{r^n}\frac{d}{dr}\left[r^n E(r)\right] = \frac{\rho(r)}{\varepsilon_s}$$

Equation 2.8

where *n*=1 for a cylindrical junction and *n*=2 for a spherical junction. *E(r)* is given by:

$$E(r) = \frac{1}{\varepsilon_s r^n}\int_{r_j}^{r} r^n \rho(r)dr + \frac{A}{r^n}$$

Equation 2.9

where r_j is the radius of curvature of the junction and the constant *A* must be adjusted to satisfy the breakdown condition. The breakdown voltage for a cylindrical junction is:

$$V_{CY} = \left[\frac{1}{2}\left(\eta^2 + 2\eta^{6/7}\right)\ln\left(1 + 2\eta^{-8/7}\right) - 2\eta^{6/7}\right]V_B$$

Equation 2.10 Breakdown voltage for a cylindrical junction

while the breakdown voltage for a spherical junction is:

$$V_{SP} = \left[\eta^2 + 2.14\eta^{6/7} - \left(\eta^3 + 3\eta^{13/7}\right)^{2/3}\right]V_B$$

Equation 2.11 Breakdown voltage for a spherical junction

where V_B is the breakdown voltage for a planar junction with the same doping concentrations and $\eta \equiv \frac{r_j}{W_m}$.

2.3.5. Avalanche Breakdown Temperature Dependence

Thus far, we have discussed avalanche breakdown at room temperature. Temperature dependence of avalanche breakdown voltage can be described using the optical-phonon mean free path [88]:

$$\lambda = \lambda_0 \tanh\left(\frac{E_p}{2kT}\right)$$

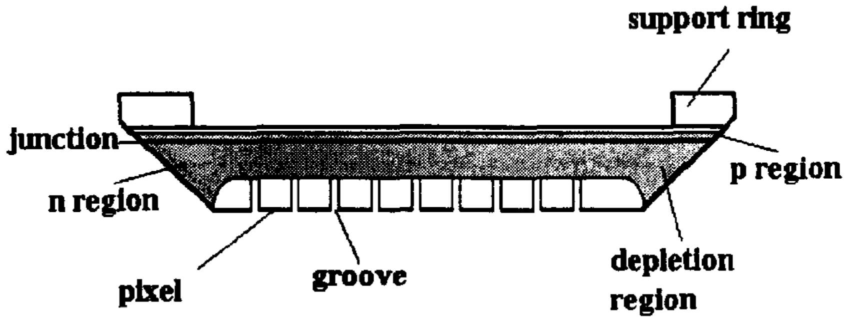
Equation 2.12 Optical-phonon mean free path

An increase in temperature results in hot carriers losing some of their energy to optical-phonons as they traverse the high electric-field region of the depletion region. The optical-phonon mean free path, λ, decreases with increasing temperature leading to hot-carriers losing more energy over a shorter path length given the same field-strength. It follows that carriers must overcome a greater voltage potential to initiate impact ionization with increased temperature.

2.4. Solid-State Guard Rings for High-Field Devices

Solid-state guard rings are used for a number of reasons in high-field devices such as device isolation and the elimination of premature junction breakdown in curved regions. More specifically, guard-rings are implemented in avalanche photodiodes operation since these devices need to sustain high instantaneous current spikes. Device isolation becomes of particular importance when designing arrays of pixels where electrical crosstalk is an issue. Further, one of the biggest challenges in high-field devices is fabricating a planar junction. Achieving a planar p-n junction with high-field devices is critical for uniform breakdown characteristics over the detection area. Planar junctions have been achieved with a number of different guard-ring structures in both dedicated custom processes, as well as, in commercial CMOS processes.

Custom fabrication has the advantage of being able to tailor their process flow to achieve highly planar junctions. Planar junctions have been achieved a number of ways such as physically cleaving the device [95],[96] to create a beveled-edge leaving just the p-n junction remaining (Figure 1.2). This method, while effective, is not practical for mass production.



2.7 Beveled planar p-n junction [95]

Similarly, mesa etching physically planarizes the junction by etching vertically and isolating neighboring pixels by filling the etched holes with a dielectric and has been demonstrated with arrays, albeit with poor fill-factors [97]. Fabrication methods such as the double-epitaxy and a straddled junction have been demonstrated to create planar junctions by extending the shallow implant edges past the deep implant thereby creating a planar junction at the shallow-deep implant interface [69],[98].

The most common structure for guard-rings in SPADs is the diffused-ring structure. A low-doped ring is implanted at the edges of the high-doped implants, overlapping with the curved regions where one would normally get premature junction breakdown due to high current densities at these regions. By implanting a low-doped ring, the high electric-field is spread over a larger area, reducing the current densities at the regions of highest curvature. Because of the lightly doped profile at the edges, breakdown voltages here are much higher here compared to at the highly-doped p+ to n-well junction planar interface. This type of processing is compatible with standard CMOS technologies and SPAD device implementations have been demonstrated [80-83],[99-101].

As mentioned previously, this dissertation involves the design, characterization, and development of a novel Shallow Trench Isolation (STI)- guard-ring structure which will be described in detail in Chapter 3. Following our work, STI-guard ring structures with passivation have been implemented [85].

2.5. Geiger-Mode Single Photon Avalanche Diodes

2.5.1. Principle of Operation

Geiger-mode Single Photon Avalanche Diodes (GM-SPADs) are p-n junctions biased above the breakdown voltage. The term “Geiger-mode” is derived from Geiger counters which measure ionizing radiation, such as nuclear radiation. Instead of detecting ionization radiation, Geiger-mode SPADs detect individual photons.

When a p-n junction is biased above the breakdown voltage, by definition, electron-hole pairs are generated on average faster than they can be extracted. Hence, an electron-hole pair, photo- or thermally generated, can trigger a strong avalanche current pulse through impact ionizations because of the extremely high electric-fields. This avalanche is self-sustaining and therefore needs to be quenched in order to prevent irreversible damage to the diode. A current limiting resistor or circuit is used to stop the avalanche and reduce the junction bias to below the breakdown voltage. Because of the high instantaneous current densities, a current spike is output corresponding to a photon detection event, producing a binary signal, whether or not a photon was detected. After the junction is biased below breakdown, the junction recharges back to above the breakdown voltage to reset and wait for the next detection event.

Table 2.1 Comparison of APD and SPAD characteristics

Figure of Merit	APD	SPAD	Implications
Bias	Below breakdown voltage	Above breakdown voltage	Need for quenching circuitry in SPADs
Gain	Linear	"Infinite"	Multiplication noise in APD but not in SPAD
Output	Proportional to # of impinging photons	Binary signal corresponding to detection event, regardless of # of impinging photons	# of photons resolvable with APDs as long as above the noise floor of the detector.
Avalanche Buildup	Slow	Fast	Timing precision better with SPAD
Noise	Different noise mechanisms which will be illustrated in Section 2.5.5		

2.5.2. Avalanche Quenching and SPAD Recharge Methods

When a photon is incident on a SPAD and is absorbed and creates an electron-hole pair in the high electric-field of the depletion region, an avalanche occurs through self-sustaining impact ionizations resulting in a runaway current. To prevent irreversible thermal damage to the diode, an avalanche quenching mechanism must be used which temporarily drop the diode's voltage bias below breakdown levels. After quenching, the SPAD must be reset back above breakdown through a recharging mechanism. In this section, we will illustrate various quenching and recharge methods.

2.5.2.1. Passive Quenching and Recharge

A simple way to quench the avalanche in SPADs and reduce the voltage bias of the device below breakdown is to use a high-load series resistor (Figure 2.8). A model based on [102],[103] is used for equivalent circuit models for SPADs (Figure 2.9). The SPAD is reverse-biased with an excess voltage above breakdown with $V_{ex} = V_d - V_b$, where V_d and V_b are the diode bias voltage and breakdown voltage respectively, through a high-load resistor R_L with a small resistor R_S (usually 50Ohms). The diode junction has a capacitance C_d with parasitic capacitance C_p in parallel. Avalanche triggering corresponds to closing of the switch, or discharging of the junction capacitor.

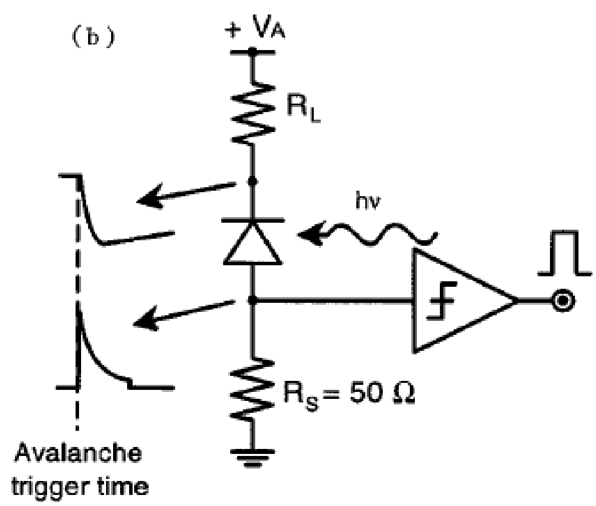


Figure 2.8 Schematic Diagram of a SPAD with passive quenching circuit [103]

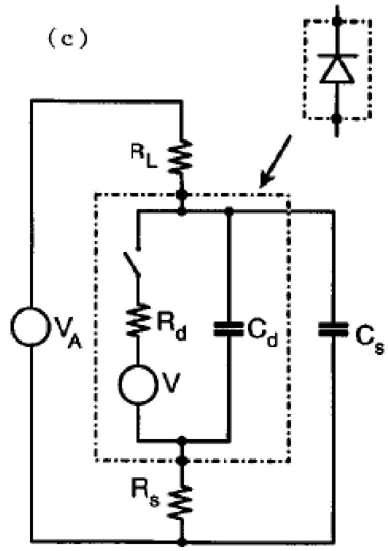


Figure 2.9 Equivalent circuit model for passively quenched SPAD [103]

The current flow is given by:

$$I_d(t) = \frac{V_d(t) - V_B}{R_d} = \frac{V_{ex}(t)}{R_d}$$

Equation 2.13

An avalanche current spike induces a voltage drop across the load resistor and decreases the voltage towards the breakdown voltage, resulting in a subsequent decrease in the avalanche current. With the reduction in the junction current and reduction in the overbias voltage, the probability of impact ionization in the junction effectively goes to zero resulting in quenching of the avalanche.

It follows that the diode voltage and current flow drop exponentially towards:

$$V_f = V_B + R_d I_f$$

$$I_f = \frac{V_A - V_B}{R_d + R_L} \cong \frac{V_E}{R_L}$$

Equation 2.14 Asymptotic Steady-state values for diode quenched voltage and current

The characteristic quenching time constant τ_q is set by the total parallel capacitance and resistance:

$$\tau_q = (C_d + C_p)\frac{R_d R_L}{R_d + R_L} \cong (C_d + C_p)R_d$$

Equation 2.15 Quenching Time Constant

The voltage necessary for quenching can be approximated by the voltage necessary to bring the diode to the breakdown voltage:

$$V_q = V_b + I_q R_d$$

Equation 2.16

Following the avalanche quenching, when there is no current flow in the junction, the diode voltage bias is brought back up to above breakdown and the junction is recharged. Further, the total charge per avalanche pulse is important to know when considering trapping effects on afterpulsing. Total charge per avalanche pulse is given by:

$$Q_{pc} = (V_A - V_q)(C_d + C_s) \cong V_{ex}(C_d + C_s) \cong I_f \tau_r$$

Equation 2.17

where τ_r is the characteristic junction *RC* recharge time constant:

$$\tau_r = R_L\left(C_d + C_s\right)$$

Equation 2.18

The recharge time constant has implications with diode reset speeds which puts a limit on the sampling speed of the device. If a photon arrives before the SPAD has been given sufficient time to recharge, the SPAD will be prematurely discharged, resulting in variable detection efficiencies and varying pulse heights (which effects pulse counting edge thresholds and in effect timing precision). Junction capacitance is fixed as are the resistor values determined by the avalanche current flow. To reduce the *RC* recharge time constant, one can reduce parasitic capacitance and thermal effects of external quenching resistors by using integrated quenching resistors [104].

2.5.2.2. Active Quenching and Recharge

Active quenching and recharge circuitry has been developed to improve on slow response times from passive quenching and recharge. Several methods have been explored where the circuit senses the avalanche pulse and immediately drops to voltage bias below breakdown with a controlled voltage source and then resets the device by bringing the voltage back above breakdown [77],[105-108]. Active quenching is useful for SPADs with slow response times, particularly those with large junction capacitances. Active recharge has been used to speed up the device reset time by using separate quenching and recharge resistors. The active quench circuitry senses the onset of an avalanche pulse and quenches it. Following this, the quenching resistor is disconnected and a smaller resistor is connected to recharge the junction, thereby reducing the characteristic recharge time determined by Equation 2.18. Using active quenching and recharge circuitry, it is possible to tune the delay between quenching and recharge to optimize noise performance, particularly in the form of releasing of charge traps to reduce afterpulsing effects [109].

2.5.2.3. Negative-feedback quenching

There has been development of a self-quenching method where a resistive layer is utilized to form charge accumulation at the resistive layer to silicon interface [110],[111]. This is a nice method where external circuitry is unneeded and may have implications for improving SPAD arrays, particularly in non-CMOS processes where circuit integration is difficult.

2.5.3. SPAD Figures of Merit

We have described several applications for single photon detection in Chapter 1. It is critical to recognize the challenges and needs for each of these applications when designing SPADs and to tailor various specifications for the best possible performance. However, it is difficult to optimize every parameter simultaneously, necessitating trade-offs in performance. In this section, I will describe various figures of merit for SPAD performance and how they relate to various SPAD applications

2.5.4. Single Photon Detection Probability and Spectral Response

To register a single photon detection event, several conditions must be met. First, a photon must impinge on the active area of the detector. Next, the photon must be absorbed to create an electron-hole pair. Following this, the photogenerated carriers must initiate impact ionization resulting in a current pulse. In the following subsections, I will outline factors that influence each of these conditions.

2.5.4.1. Transmittance from air to silicon

In standard CMOS technology process flows, p-n junction diodes in silicon are buried deep under a layer of silicon dioxide film. This non-absorbing film can act as a reflector and has a wavelength dependent transmission curve depending on its thickness. Further, multiple reflections can be present in the oxide layer creating a cavity for constructive and destructive interference. Anti-reflection coatings are designed to maximize transmittance of light to the junction and can be tuned to specific wavelength ranges. The transmittance for air-to-silicon through an oxide layer is given by [112]:

$$T = \frac{4n_{Si}}{\left[(n_{Si}+1)\cos\delta - \frac{k_{Si}}{n_{Ox}}\sin\delta\right]^2 + \left[\frac{(n_{Si}+n_{Ox})}{n_{Ox}}\sin\delta - k_{Si}\cos\delta\right]^2}$$

Equation 2.19 Transmittance for air-to-silicon through silicon dioxide layer

where n_{Ox}, n_{Si}, and k_{Si} are the real and imaginary parts of the silicon index of refraction (Figure 1.2) and δ is the phase change on the traversal through the oxide layer of thickness *d* given by:

$$\delta = \frac{2\pi}{\lambda} n_{Ox} d$$

Equation 2.20 Phase change from silicon dioxide layer traversal

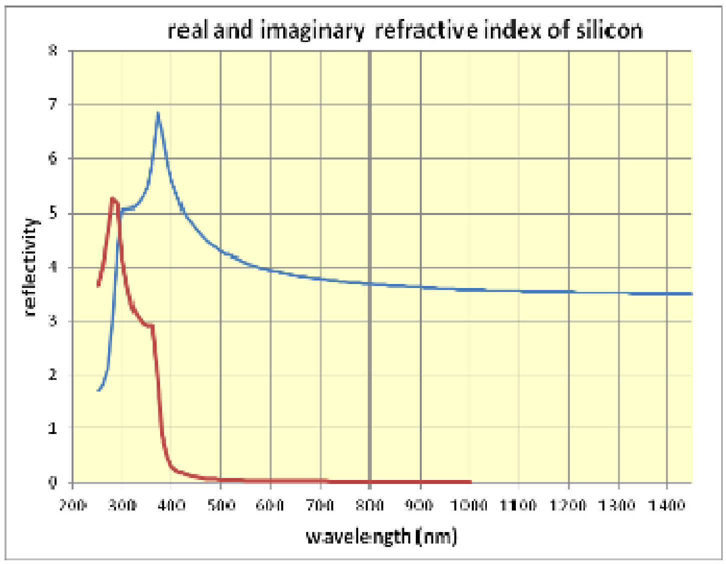


Figure 2.10 Real and imaginary refractive index of silicon [113]

2.5.4.2. Absorption in silicon

After the incident photons penetrate the silicon dioxide layer, they will reach the silicon to be either absorbed or cross the material altogether. The fundamental condition for photon absorption is for the photon to have enough energy to excite an electron to surpass the bandgap energy of the material. Since silicon is an indirect band-gap semiconductor material, transition from the valence band to the conduction band requires phonon interaction for conservation of momentum (Figure 2.11).

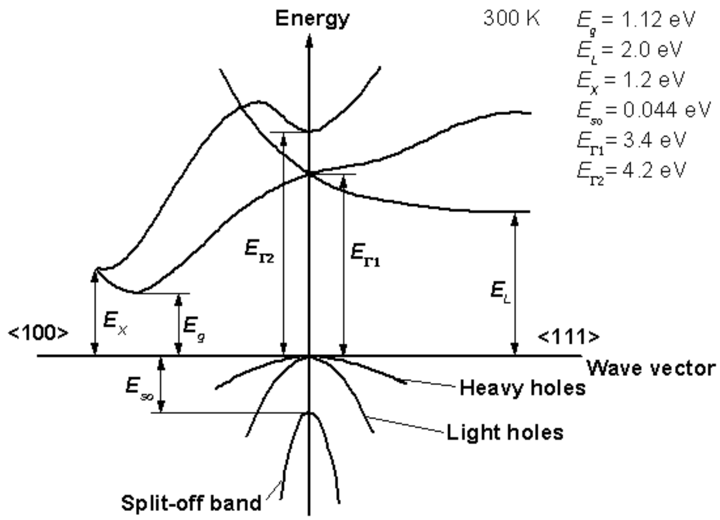


Figure 2.11 Energy Band Diagram for Silicon [114]

Silicon has a bandgap energy of 1.12eV resulting in a cutoff wavelength of approximately 1100nm.

$$\lambda_c = \frac{1.24}{E_g} \cong 1.1\mu m$$

Figure 2.12 Cutoff wavelength for Silicon

Because of the need for suitable phonon interaction, the absorption coefficient increases from the NIR to visible-UV wavelengths (Figure 2.13). It follows that the absorption depth is wavelength dependent, shown in Figure 2.14, where the absorption depth is defined by where the light intensity has fallen to *1/e* (36%) of its original value.

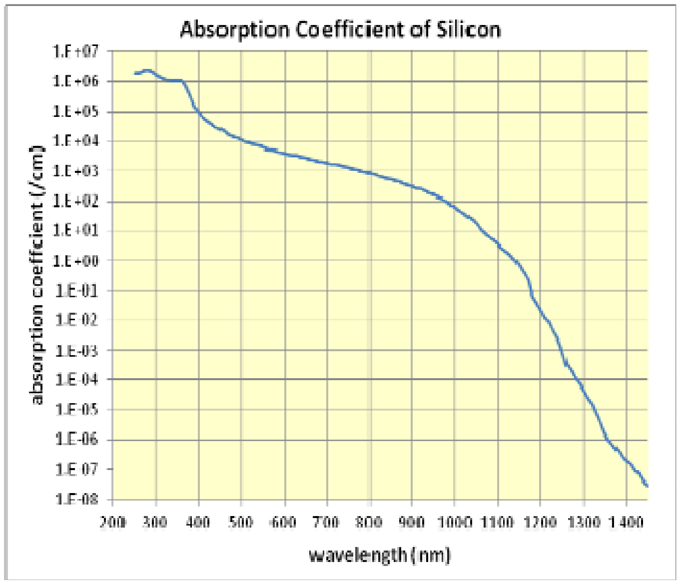


Figure 2.13 Absorption Coefficient of Silicon [113]

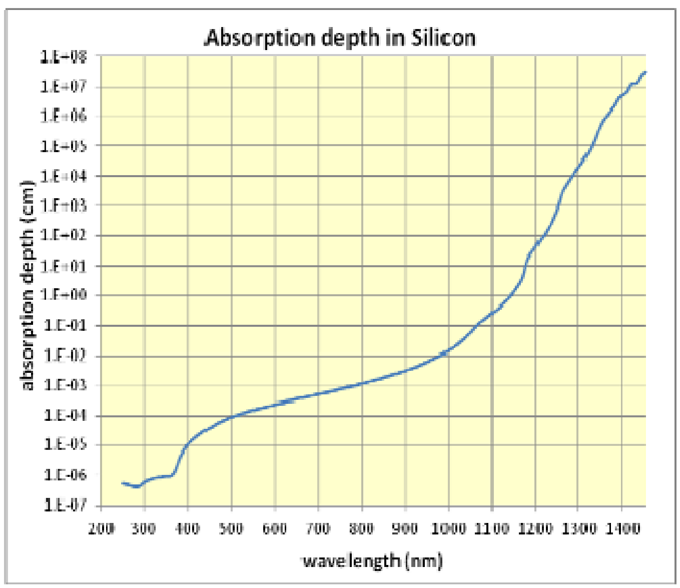


Figure 2.14 Absorption depth in Silicon [113]

Further, the probability of photon absorption at depth z and wavelength λ follows a cumulative distribution function given by:

$$P(z) = 1 - e^{-a(\lambda)z}$$

Equation 2.21 Probability for photon absorption of wavelength λ and depth z

2.5.4.3. Surface recombination

When an electron-hole pair is generated, there is a chance that they will recombine with each other (or another electron and hole) preventing them from initiating impact ionization. This can be a direct effect where the electron relaxes from the conduction band to an empty hole in the valence band, or it can be through an indirect transition using traps or defect centers in the crystal lattice.

Surface recombination has implications with SPAD devices, particularly those with thin depletion regions. High energy photons will be absorbed at shallower regions in the silicon where there is no high electric-field and will likely recombine before the carriers are able to diffuse to the high-field region to initiate impact ionization. This reduces the avalanche probability for shorter wavelengths, particularly those in the UV range as those carriers will need to rely on diffusion to generate an avalanche pulse. The probability of a carrier diffusing into the depletion region is given by:

$$P_{e_{z\to z_0}}(z) = \frac{\sinh\left(\frac{z}{L_e}\right)L_e s + \cosh\left(\frac{z}{L_e}\right)D_e}{\sinh\left(\frac{z_0}{L_e}\right)L_e s + \cosh\left(\frac{z_0}{L_e}\right)D_e} \cong \frac{z}{z_0}$$

Equation 2.22

where L_e is the electron diffusion length, s is the surface recombination velocity, and D_e is the diffusion constant of the electrons. Equation 2.22 can be approximated to $\frac{z}{z_0}$ when surface recombination and diffusion length are larger than the junction depth at z_0.

2.5.4.4. Avalanche Breakdown probability

Avalanche breakdown probability is dependent on the electric-field strength (or excess voltage). As electric-field strength increases, the probability of avalanche breakdown approaches 100%. Further, higher doped junctions will achieve maximum avalanche breakdown probability at lower excess voltages. Low doped p-n junctions have high breakdown voltages of tens of volts [66], compared to just a few volts for highly doped junctions [115-117]. However, avalanche breakdown probability saturates even though the depletion region continues to expand at high excess bias voltages.

2.5.4.5. Single Photon Detection Probability

The calculation of the single photon detection probability of a junction combines the aforementioned factors of transmittance, wavelength, depth, and is expressed as:

$$DP(\lambda) = T(\lambda)\left[P_e(z_0)\int_0^{z_0} \left(a(\lambda)\exp(-a(\lambda)z)\right)dz + \int_{z_0}^{z_w} \left(a(\lambda)\exp(-a(\lambda)z)P_p(z)\right)dz \right]$$

Equation 2.23 Single Photon Detection Probability

where we combine the transmittance, photon absorption and diffusion to the depletion region from outside the high-field region at the p+ anode, and photon absorption at the depletion region.

2.5.5. Noise

SPAD noise comes in the form of false counts, where non-photo-generated electron-hole pairs can initiate impact ionization. Free carriers can be generated thermally, through band-to-band tunneling, and through afterpulsing effects, and produce false counts in the photodiode. Dark count rates in SPADs vary with overbias voltage due to increased avalanche probability and also high-field effects. Further, the different noise mechanisms have different temperature dependencies. In the following subsections, I will describe the various noise sources in SPADs and their implications on design and performance.

2.5.5.1. Band-to-band Thermal Generation

Dark current from thermally generated electron-hole pairs is present in photodetectors of all sorts. In semiconductors, thermal generation of carriers is attributed to band-to-band transitions, via either a direct band-to-band transition or via trap-assistance by generation-recombination centers described by Shockley-Read-Hall (SRH) theory [118]. A direct transition is unlikely, particularly in indirect bandgap materials such as silicon with a bandgap of 1.12eV, even at high temperatures. Trap-assisted generation-recombination is a much more likely phenomenon. SRH theory uses four basic processes to illustrate trap-assisted thermal generation/recombination (Figure 2.15):

(a) electron capture

(b) electron emission

(c) hole capture

(d) hole emission

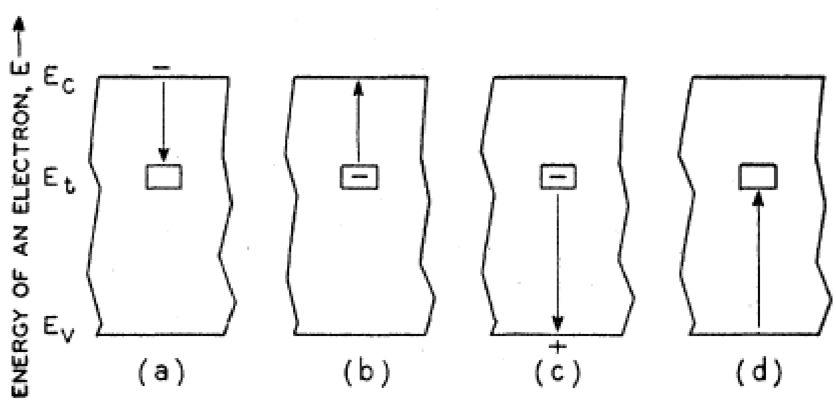


Figure 2.15 Trap-assisted recombination [118]

As with dark current in regular photodiodes, thermal generation of carriers increases with temperature.

2.5.5.2. Field-assisted Tunneling

At very high electric fields, direct band-to-band transitions may occur without trap assistance. The Poole-Frenkel effect enhances transitions from the valence band to the conduction band in high electric fields. As an indirect bandgap semiconductor, tunneling in silicon diodes requires phonon interactions for conservation of momentum. Tunneling has a weak dependence on temperature, putting a lower limit of how much cooling the detector will benefit dark count rates in SPADs.

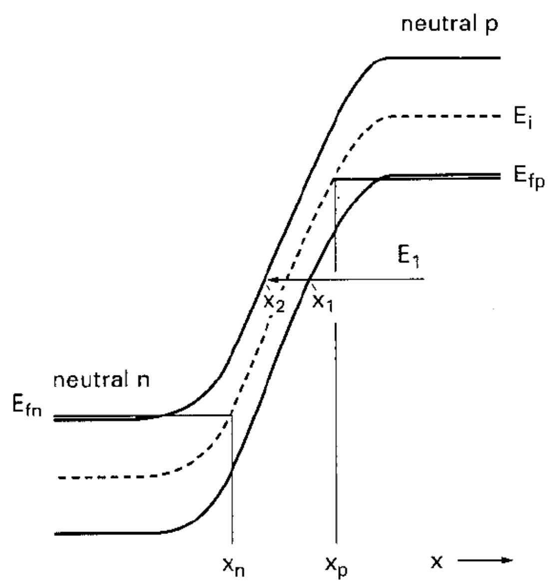


Figure 2.16 Energy band diagram of reverse-biased p-n junction with band-to-band tunneling transition [119]

Hurkx et al. describe a model for band-to-band tunneling and derive an expression for tunneling generation rate as [119]:

$$R_{Tunn}(z) = -B|F|^{5/2} D\left(F, E, E_{fn}, E_{fp}\right)\exp\left(-F_0/|F|\right)$$

Equation 2.24 Tunneling generation rate

where F is the electric field, F_0 is the electric field at room temperature (1.9E17 V/cm), B is a temperature independent constant 4E14 $cm^{-0.5}V^{-2.5}s^{-1}$. The function $D(F, E, E_{fn}, E_{fp})$ accounts for relative position of the Fermi levels E_{fn}, E_{fp} in the neutral regions on the n and p side, respectively.

2.5.5.3. Afterpulsing

During each avalanche pulse, millions of electron-hole pairs traverse the junction. Carriers may be trapped by deep levels in the depletion region at trapping centers in the forbidden band and be released at a characteristic time after the initial avalanche pulse (Figure 2.17). The release of these trapped carriers may initiate a secondary avalanche pulse correlated to the first one, and is termed an afterpulse [120],[121].

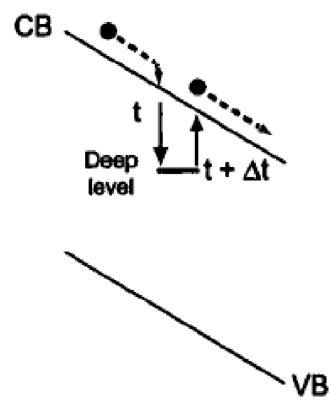


Figure 2.17 Trap release from forbidden band [121]

Afterpulsing is a statistical phenomenon with a characteristic emission probability density function Figure 2.18 and a temperature-dependent trap lifetime Figure 2.19. Trap lifetime is modeled using the Arrhenius equation [122]:

$$\tau_d = \frac{1}{\sigma v N} e^{E_a / kT}$$

Equation 2.25 Arrhenius equation for Trap Lifetime

where T is the temperature, σ is the trap cross-section, $v=v(T)$ is the average thermal velocity of carriers to be trapped, and $N=N(T)$ is the effective density of states of the relevant band. It is assumed that $v \propto T^{1/2}$ and $N \propto T^{3/2}$, making $\sigma v N \propto T^2$. Figure 2.19 shows that a temperature reduction of tens of Kelvin increases the lifetime by an order of magnitude for a custom silicon photodetector [121].

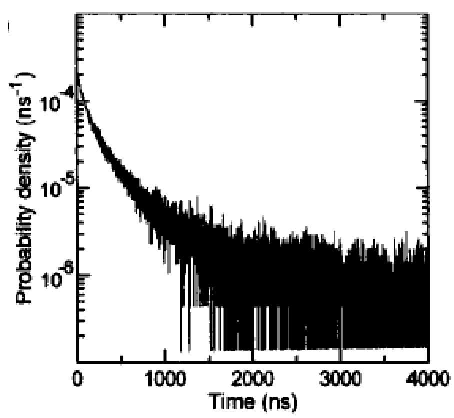


Figure 2.18 Probability density function of trap release [121]

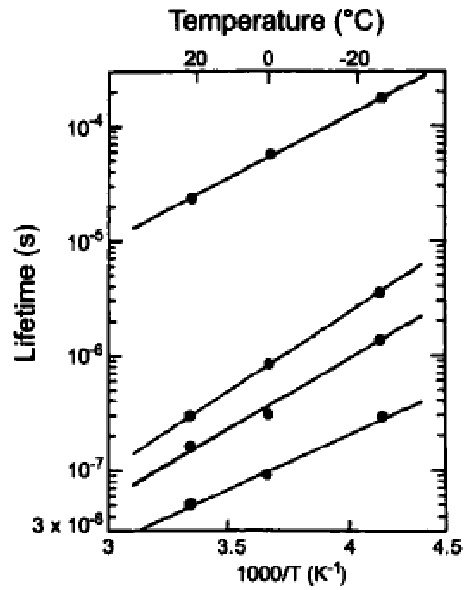


Figure 2.19 Arrhenius plot of temperature dependence on trap lifetime [121]

Since trap release has a characteristic decay and is not dependent on whether an avalanche pulse has occurred or not, the probability of trap release at time *t* is the sum of all previous trapped charges released at a probability following their characteristic trap release time:

$$P(t)\Delta t = B\Delta t + \sum_i A_i \Delta t \exp(-t/\tau_i)$$

Equation 2.26 Carrier emission probability [120]

where Δt is the time width of each interval in the histogram, *B* is the background rate due to thermally generated carriers, and A_i is the initial number of charges trapped after avalanche pulse *i*. This probability distribution function allow the optimal hold-off time for trap release to be determined.

In addition, afterpulsing can act as a positive feedback loop producing self-propagating avalanche pulses and a stark increase in the dark count rate, unless avalanche probability is small. Trapping centers are unlikely to be saturated, thus the number of charges trapped is linearly proportional to the amount of charge that flows through the junction. Therefore, afterpulsing can be reduced by reducing junction capacitance or be allowing traps to be released before resetting the detector above breakdown. Active quenching and recharge circuitry described in Section 2.5.2.2 has been developed to have a variable hold-off time for junction reset to allow the majority of trapped carriers to be released. Table 2.2 summarizes the dependencies of each type of SPAD noise:

Table 2.2 Noise dependencies in SPAD

	Shockley-Read-Hall (SRH)	**Tunneling**	**Afterpulsing**
Correlation	Uncorrelated	Uncorrelated	Correlated
Distribution	Poisson	Poisson	Exponential
Voltage dependence	Strong dependence	Very strong dependence	Strong dependence
Temperature dependence	Strong dependence	Weak dependence	Inverse dependence

2.5.6. Timing Precision (Jitter)

SPADs detect the arrival of photons and knowing their precise arrival time is important for many applications. Timing precision is affected by the device structure as well as sense amplifier design. CMOS SPADs traditionally use a diffused guard-ring structure to prevent edge breakdown by extending the region across which the electric field develops [123]. This guard-ring structure leads to increased jitter due to the lateral drift and diffusion of electron-hole pairs created by photon absorption in the low-field and neutral regions (Figure 2.20) [121]. This results in a main photon arrival time distribution peak where photons are absorbed in the high-field region and immediately initiate an avalanche pulse and a diffusion tail where photons absorbed in the low-field and neutral regions diffuse to the high-field region to initiate an avalanche pulse (Figure 2.21).

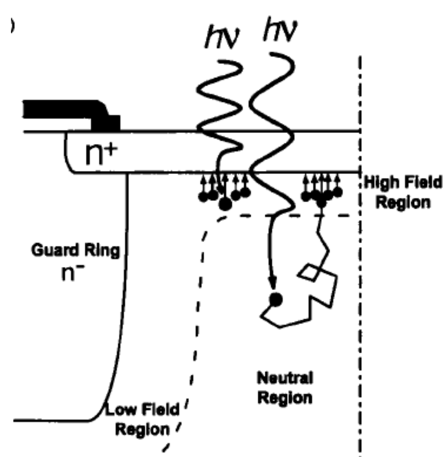


Figure 2.20 Photon absorption in different regions of SPAD [121]

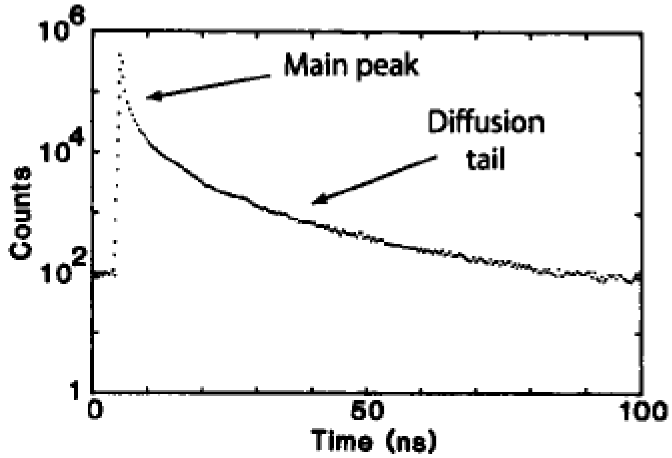


Figure 2.21 Photon time-of-arrival distribution [121]

The resulting diffusion tail limits SPAD timing performance, e.g., in QKD systems [124] and in classical pulse position modulation optical links [20], where diffusion tails limit bit-error-rates and bit-rates, respectively, and in high resolution FLIM [125], where better timing resolution translates into better image contrast. Therefore, it is necessary to optimize SPAD design not only for a minimal Full-Width-at-Half-Maximum (FWHM), but also for a minimal Full-Width-at-Hundredth-Maximum (FW(1/100)M). A jitter-optimized CMOS SPAD has recently been reported with a FWHM jitter of 36 ps using a high-voltage (HV) CMOS process albeit with a 1.1ns FW(1/100)M [77].

Several methods can be used to reduce a SPADs timing jitter namely, localizing the SPAD high-field region to smaller areas, reducing/eliminating the low-field region for carrier diffusion, optimize avalanche sensing circuitry to better detect the onset of the avalanche pulse, and utilize high precision integrated time-to-digital converters.

2.5.7. Dynamic Range and Dead Time

Dynamic range and Noise Equivalent Power (NEP) in SPADs is dark count rate limited as is defined by the incident power needed to have a SNR of 1:

$$NEP = \frac{P_\lambda}{DP}\sqrt{\frac{DCR}{\Delta t}}$$

Equation 2.27 Noise Equivalent Power

where P_λ is the optical power, *DP* is the detection probability, *DCR* is the dark count rate, and Δt is the integration time. Dark count rate is assumed to follow Poisson statistics.

Dead-time is the time during which the SPAD resets. Before the voltage is brought back above breakdown, the SPAD does not have an adequate electric field to initiate impact ionization, thereby making it “blind” during the dead-time. This puts an upper limit on count rates and the maximum photon flux the detector can handle. Dead-times also have implications on noise characteristics mentioned in Section 2.5.5. In particular, longer dead-times are used to allow trapped charges to be released to reduce afterpulsing effects on the dark count rate.

When incident photon rates increase, photons may arrive during the detector dead-time and must be taken into account. Further, due to junction recharge times, non-uniform detection probabilities become an issue when a photon reaches the junction before it has fully recharged.

3. STI-BOUND CMOS SPAD DEVELOPMENT – CONCEPT, MODELING, AND CHARACTERIZATION

3.1. Concept

Various SPAD designs have been described in Chapter 2. We have developed a new type of SPAD in a deep submicron CMOS process which eliminates the timing jitter diffusion tail with a SiO_2 Shallow Trench Isolation (STI) guard-ring [126],[127]. This isolation trench is constructed early in the fabrication process and is traditionally used to prevent punch-through and latch-up in CMOS circuits [128]. The edges of the drain implant are confined by the oxide trench preventing any lateral diffusion and subsequent formation of curved edges, thus eliminating the need for a diffused guard-ring structure. A cross-section schematic of the device is shown in Figure 3.1 [126]. Two generations of test chips were fabricated in an IBM 0.18μm CMOS technology through the MOSIS service. No design rules were violated during the layout. Further, design of various pixel sizes was possible, providing the power of CMOS design for single-pixel applications where large area detectors are desirable for easy alignment, or small area pixels for better spatial resolution and formation of scalable arrays.

With smaller feature sizes (and reduced junction capacitances) and integrated quenching and recharge, we have demonstrated state-of-the-art performance with the STI-bound SPAD in parameters such as short dead-times (faster recharge and quenching), fill-factors [129],[126],[127] and timing precision [130]. Further, integrated active quench circuitry for dark count suppression [131] and demonstration of a dual-junction SPAD where the deep n-well to p-substrate is used as a second junction for spectral response shifted towards longer wavelengths [132]. In this chapter, I will describe the development of this SPAD through device simulation, electrical and optical characterization, and application development, with a focus on timing precision.

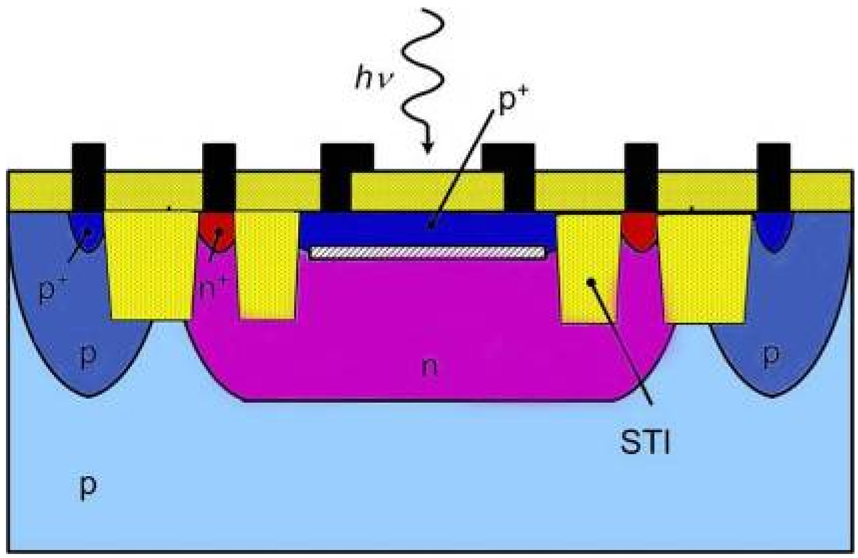


Figure 3.1 Cross-section of STI-Bound SPAD

3.2. Device Modeling

3.2.1. Physical Modeling

Electric field simulations were done with the ISE-TCAD software to confirm electric field distributions and planar breakdown at the p-n junction. Due to unavailability of exact CMOS process parameters, certain assumptions and approximations were utilized in simulation. A comparison of the electric-field of an STI-bound SPAD and a SPAD with a diffused guard-ring are shown:

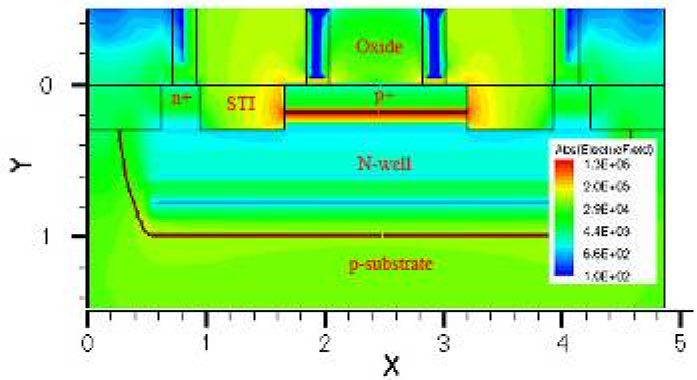


Figure 3.2 Electric field distribution – STI-bound SPAD

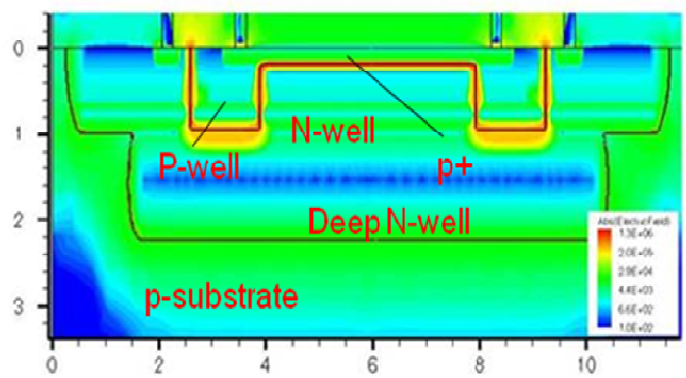


Figure 3.3 Electric field distribution – SPAD with diffused guard-ring

The STI-bound device model was formed with a shallow p+ implant surrounded by an n-well, with STI truncating the curved regions. Further, shallow n+ implants were formed as contact pads to the n-well. The diffused guard-ring SPAD was formed with a p+ implant also surrounded by an n-well. However, p-well implants are used for the guard-ring and a deep n-well is formed for electrical connection to the n-well. Because the STI guard-ring truncates the curved regions of the p+ implant, there is little possibility for premature breakdown at the curved regions. The diffused guard ring SPAD works by using low-doped p-well implants to distribute the high-field region over a larger area, thereby preventing premature breakdown of the junction at the curved regions. Because of the large area of the low doped p-well implants, fill-factors remain limited. Further, there remains a low-field region in the p-well implants where carriers generated in that region can diffuse towards the high-field region and initiate impact ionization. This has implications in timing jitter as outlined in Section 2.5.6. Since the STI guard-ring eliminates the diffused guard ring and hence the low-field region, the diffusion of carriers in the low-field region is eliminated.

3.2.2. Electrical Modeling

Electrical modeling of the SPAD was based on methods developed in [102] where the SPAD acts as a switch to charge and discharge a capacitor, with its characteristic *RC*-time constant. Novel electrical modeling was out of the scope of this dissertation. Two electrical output buffers were designed—a source-follower and an inverter chain.

The source-follower, also known as common drain, operates as a voltage buffer. Figure 3.4 illustrates a source-follower schematic where the output load is driven at the source allowing the source potential to follow the gate voltage with near unity gain. This provides an analog output of the avalanche charge flow through the junction, giving a clear view of the avalanche buildup, quenching, and recharge mechanisms.

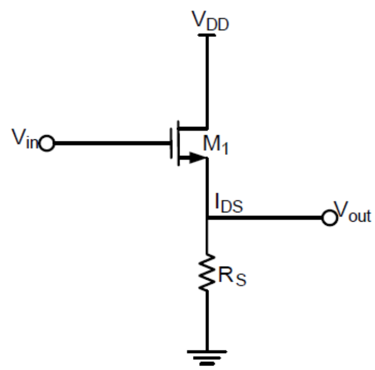


Figure 3.4 Source-follower schematic

An inverter chain output buffer was designed to generate uniform output pulses for better compatibility with pulse counting electronics. It is important to design inverters with high slew rates such that the leading edge timing precision is maintained. The importance of these output buffers becomes more apparent when analyzing SPAD jitter characteristics in Section 3.3.5.

3.3. Device Characterization

3.3.1. Introduction

Accurate characterization of single photon detectors is challenging. Not only do you need to know the precise number of photons impinging on the detector active area, it is essential to have fast electronics to be able to count pulses and high-bandwidth oscilloscopes to extract precise timing information. Two generation of test chips were fabricated. The first generation test chip was designed to prove junction planarity in an IBM 0.18μm RFCMOS process with various test structures, packaged in a 44-pin plastic lead chip carrier (PLCC) package.

The second generation test chip was designed with new test structures for measurement of different SPAD figures of merit and packaged in a 121-pin plastic grid array (PGA) package. Later the second generation test chip was repackaged in a high speed quad flat no-lead (QFN) glass top package for high speed jitter measurements. The second generation test chip had several improvements over the first, namely the removal of a metal auto-fill pattern on top of the diode which is used to reduce mechanical stress across the wafer which previously blocked the active area from clear light exposure. Further, a cobalt silicide layer was removed, which is shown to have strong absorbance at longer wavelengths. In this section, I will describe the characterization setup design and figures of merit outlined in Section 2.5.3 for our SPAD test structures, with an emphasis on characterization of the second generation test chip and comparison of a diffused guard-ring bound SPAD compared to an STI-bound SPAD. All printed circuit boards (PCBs) were self designed and manufactured though the *ExpressPCB* service.

A digital Agilent E3631A DC power supply was used to bias the SPAD devices. Initially, pulse counting was done with a Tektronix 3032B sampling oscilloscope controlled with Labview. Further, a pulsed excitation source, a Spectra-Physics DUO-210 <30Hz repetition rate pulsed dye laser, was used initially for proof-of-concept demonstration, despite the slow repetition rates.

Because of the large number of pulses necessary for meaningful counting statistics and slow data acquisition speeds of the Labview-Oscilloscope interface, a dedicated pulse counting/multiscalar, Becker-Hickl MSA-1000, was obtained for the second generation test chip development. The MSA-1000 uses 128-bit memory for a dead-time free accumulation of subsequent sweeps with 1ns time bins (Figure 3.5) [133]. The MSA-1000 counts all pulses whose amplitude is greater than a selectable discriminator threshold and stores them in subsequent memory locations, with a maximum recording time of 131μs. In addition, a pulsed diode laser (BHL-600) was used with a 20MHz repetition rate at 650nm with <10ps pulse-to-pulse jitter.

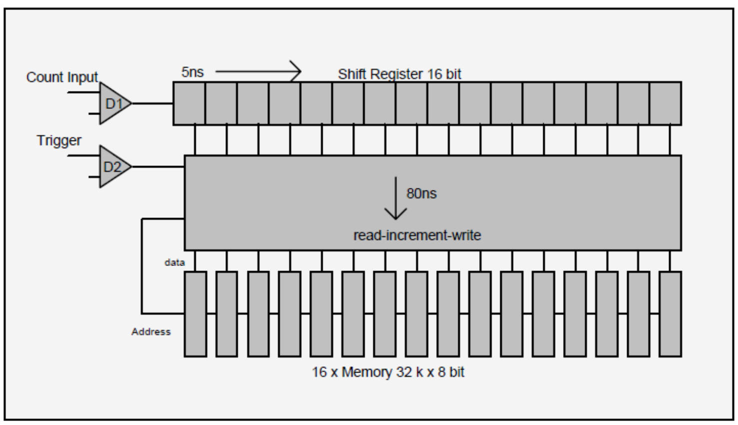


Figure 3.5 Block Diagram of MSA module [133]

3.3.2. Temperature dependence on breakdown voltage

Breakdown voltage temperature dependence was characterized with a TE-cooler coupled to the SPAD package with thermal paste (Artic Silver) with a thermistor coupled to the package surface with thermal epoxy to monitor the temperature in real-time. An ILX Lightwave LDT-5910 temperature controller was used to control the TE-cooler. An integrated on-chip resistor with a known temperature coefficient was simultaneously monitored as a validation for the chip temperature. The SPAD voltage bias was swept with an Agilent E3631A power supply and output pulses were monitored on a TDS3032B oscilloscope. Vdd was fixed at 3.3V and the p+ voltage was swept to find at what bias avalanche pulses begin to appear at each temperature point (Figure 3.6). We can see the breakdown voltage follows temperature trends outlined in Section 2.3.5, where breakdown voltage increases with increased temperature due to the reduction of the optical phonon mean free path. Temperature measurements were only able to be carried out to 277K, due to non-vacuum experimental conditions. A 10mV/K increase in breakdown voltage is measured.

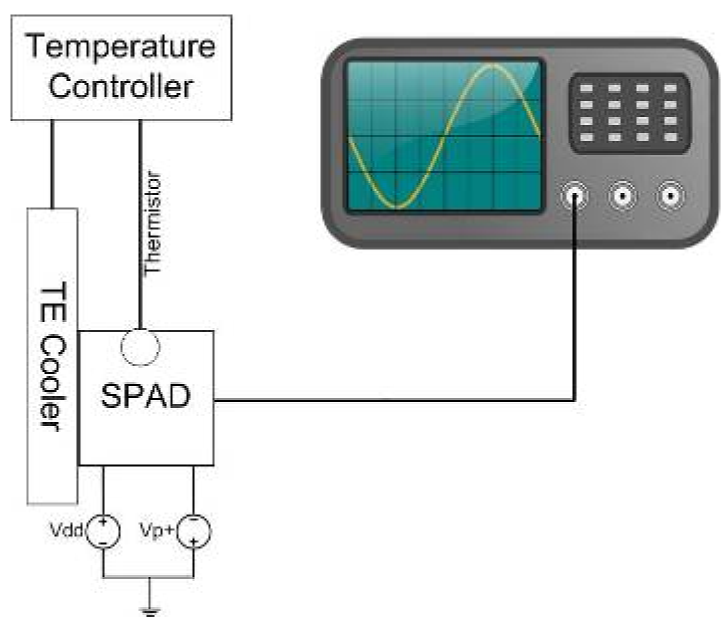


Figure 3.6 Schematic of SPAD Temperature Measurement

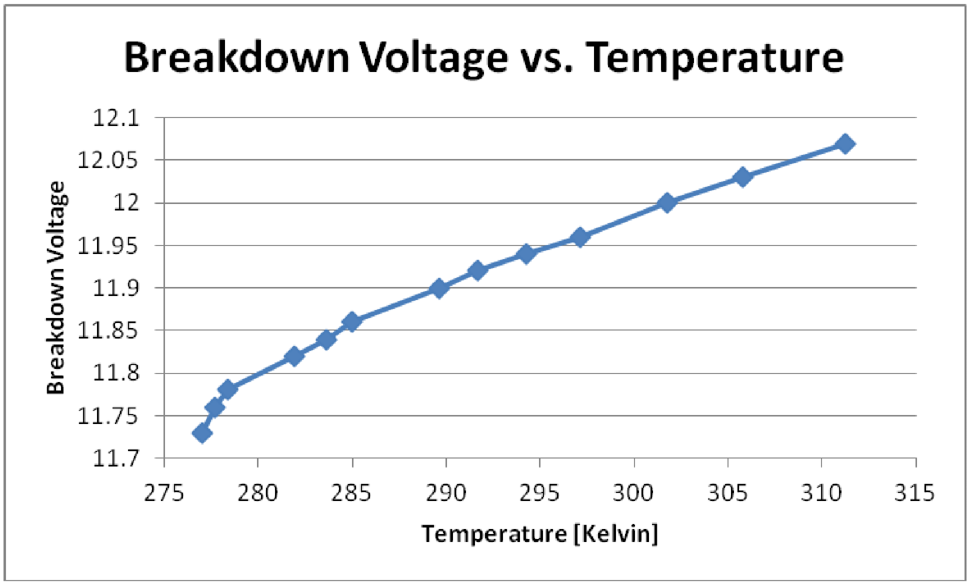


Figure 3.7 Breakdown Voltage vs. Temperature

3.3.3. Noise and Dark Count Rate

Noise characteristic measurements of the SPAD were done in a dark environment where the test chip was shielded from light by a metal cover. Pulse counting was done with the MSA-1000 on STI-bound SPAD and diffused guard-ring SPAD, both with inverter output buffers. Since dark counts are not correlated with light pulses, the MSA was operated in a non-triggered mode with 0ns holdoff time (Figure 3.8).

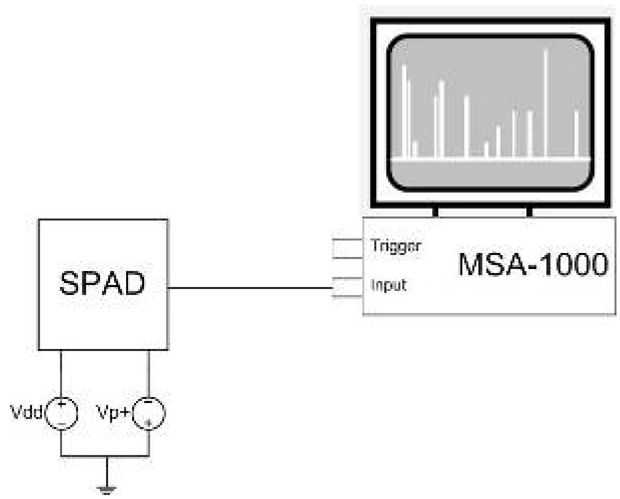


Figure 3.8 Dark Count Rate Measurement Setup

Dark count rate is measured by counts per second (cps) vs. voltage bias and calculated using:

$$DCR = \frac{(Average\#Counts)}{(BinSize)(Total\#Samples)}$$

Equation 3.1 Dark Count Rate

where the bin size was 1ns and total # samples was 10^7. Average # of counts was taken as the average over the full sample window of 131k time bins. Figure 3.9 shows the dark count rate as a function of junction bias for both an STI- and diffused guard-ring SPAD with identical inverter outputs, with a significant difference in behavior of the two devices. In particular, there is a large difference in operating voltage where the diffused SPAD begins registering dark counts at a voltage 0.4V lower than the STI-SPAD. The expected breakdown voltage should be the same due to identical doping profiles at the p+-nwell junction. It is not entirely clear the source of this difference, but it is indicative of premature breakdown in the diffused junction. It is possible that localized premature breakdown is occurring at the contacts at the n-well and p+ implants. This is supported by the initial high slope of its DCR at low device bias.

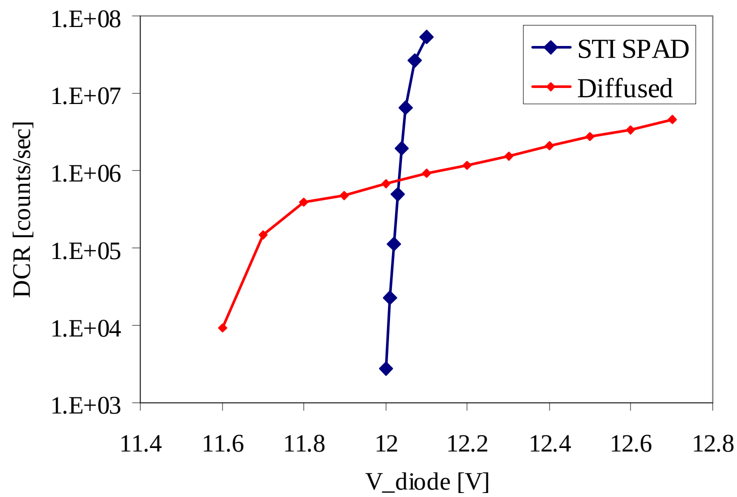


Figure 3.9 Dark Count Rate vs. diode bias for STI SPAD and Diffused SPAD

Further, the slopes of the dark count rates are very different between the two devices, where the STI-bound SPAD increases dramatically over a small voltage range. This stark increase in the dark count rate, puts an upper limit on operating voltage, where dark counts saturate the detector preventing it from fully recharging, thereby reducing its effective detection efficiency at high voltage biases. The gradual slope of the diffused SPAD can be attributed to a possible lower junction field at equivalent voltage biases compared to the SPAD due to a small electric potential that develops across the long n-well to deep n-well to n-well path. Thus, a slightly lower junction field strength perhaps reduces the number of impact ionizations from dark carriers.

Temperature characterization of the dark count rate provides insight on the dominant noise mechanism is the SPAD, namely Shockley-Read-Hall, tunneling, or afterpulsing summarized in Table 2.2. Figure 3.10 shows dark count rate vs. overbias of the STI-bound SPAD at various temperatures. Temperatures were measured using an on-chip polysilicon resistor with a known temperature coefficient. Heat was applied using a heat gun while constantly monitoring the polysilicon resistor for stability throughout the measurements.

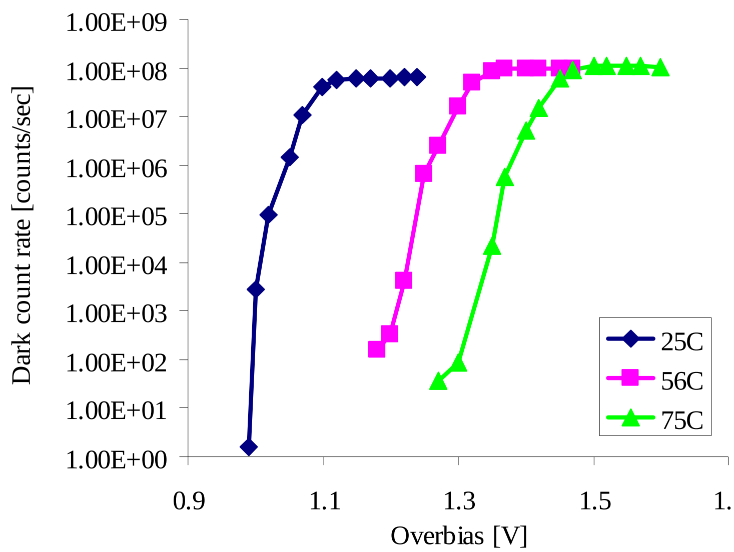


Figure 3.10 Dark Count Rate vs. Overbias of STI-bound SPAD at varying temperatures

A TE-cooled measurement of the dark count rate was also carried out in the same setup as outlined when measuring breakdown voltage vs. temperature. The SPAD is set at a fixed bias 1% above its breakdown voltage. We can see a reduction in dark count rate of 6x when cooled to 277K compared to room temperature (Figure 3.11). It seems that Figure 3.10 indicates the dominant noise is afterpulsing, due to the release of greater numbers of trapped charges during the dead time. However, Figure 3.11 indicates that thermal noise is the dominant source, particularly below room temperature.

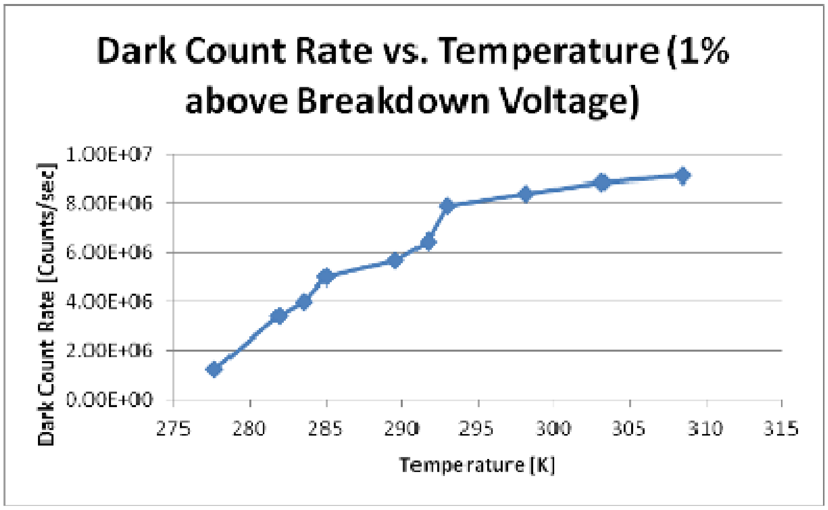


Figure 3.11 Dark Count Rate vs. Temperature at a fixed 1% above breakdown

Autocorrelation of avalanche pulses can be used to demonstrate the presence of afterpulsing. Autocorrelation was done in two ways—first with Labview data acquisition of pulse arrival times processed with a Matlab script (Figure 3.12) and second using an avalanche pulse to trigger the MSA pulse counter and recording the subsequent pulses (Figure 3.13). Autocorrelation of time-of-arrivals processing in Matlab used time-series to calculate correlation of all subsequent pulses. In both cases, it is clear that an autocorrelation peak is present at about the 20ns time bin, indicating a dead-time of >20ns is necessary for adequate trap release for suppression of afterpulsing effects at room temperature.

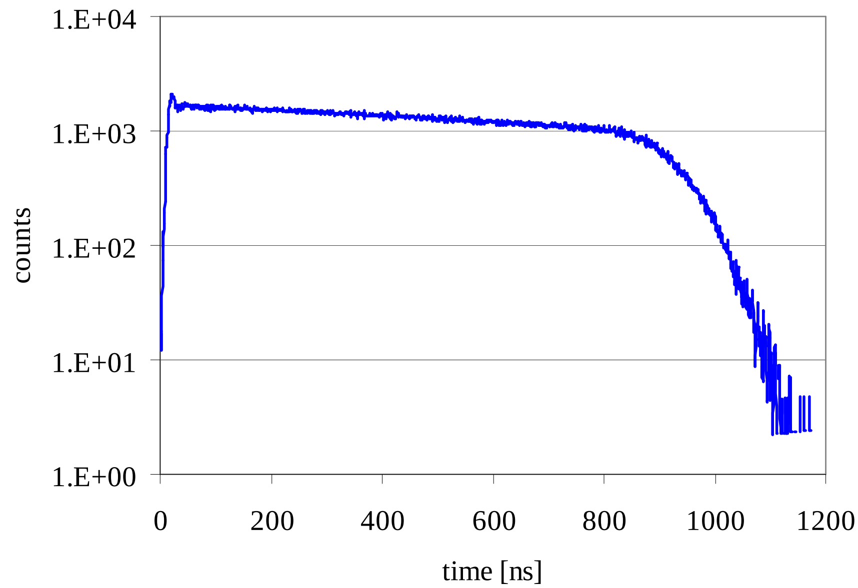


Figure 3.12 Autocorrelation of SPAD using Labview and Matlab

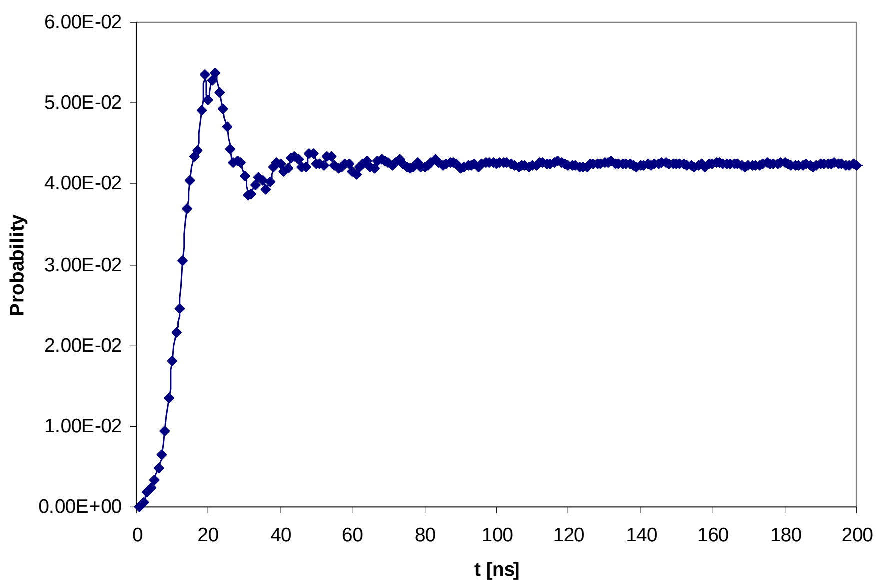


Figure 3.13 Time-of-arrival histogram of SPAD triggering MSA-1000 pulse counter with dark pulse

3.3.4. Detection Efficiency and Spectral Response

Detection efficiency and spectral response characterization of SPADs is challenging because it is necessary to know the exact photon flux incident on the detector active area. Further, to do single photon detection efficiency measurements, on average one photon should be incident on the photodetector at a time. We use a 650nm pulsed diode laser (BHL-600) as our light source and connect its sync output to the trigger signal of the MSA pulse counter (Figure 3.14). Because the beam profile of the diode laser output is elliptical, a 25μm pinhole spatial filter is used to transform the beam into a uniform circular Gaussian profile. A subsequent lens collimates the beam and passes it through a wavelength calibrated beam splitter. In one arm of the beam splitter, an optical power meter constantly monitors the reflected power such that the transmitted power is always known. A calibrated neutral density (ND) filter is placed in line for beam attenuation. A focusing microscope objective is used to focus the beam to a spot on the SPAD test chip. The SPAD PCB is mounted on a 3D micrometer stage which is then mounted on a 2D *Physique Instruments M227.25* piezoelectric stage. The piezoelectric stage is used to move the SPAD precisely into the laser focus. We expect the highest photoresponse when the active area is placed at the focus of the laser (i.e. highest photon flux per area). On the other arm of the beamsplitter, a CCD camera is mounted to monitor the laser beam location for rough alignment to the test structure of interest.

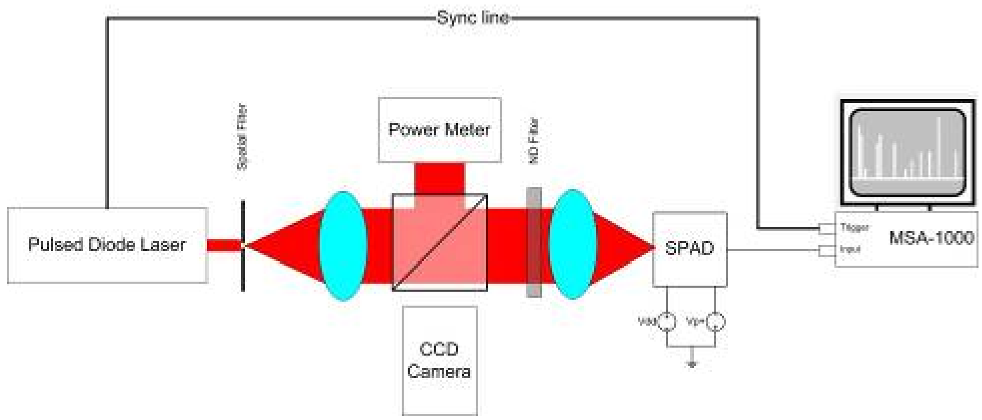


Figure 3.14 Schematic for pulsed excitation detection efficiency measurements

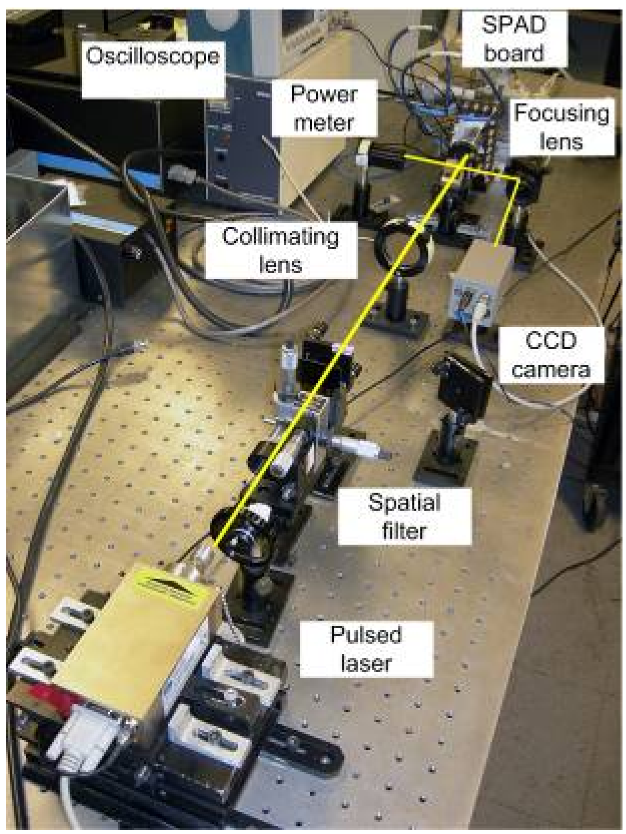


Figure 3.15 Photo of detection efficiency setup

To know the photon flux incident on the SPAD active area, one can overfill the SPAD with a known laser profile and known power distribution where the detector is placed. Alternatively, one can under-fill the SPAD such that all the pulsed photons are incident on the active area. For measurement of the detection efficiency with the pulsed excitation source, we under-fill the SPAD so that all pulsed photons are incident on the active area of the detector of 7μm. Spectral response measurements were done by over-filling the SPAD active area due to the difficulty in focusing the coherent light source to μm sized beams.

We first find the pixel visually with the CCD and roughly align it using the 3D micrometer stage while watching the SPAD output on a Tektronix TDS3032 oscilloscope. Once the beam is centered on the pixel, we move the output cable to the MSA1000 pulse counter. We then control the piezoelectric stage to move the pixel forward and backward to find the position with the highest photoresponse. We move the pixel in the horizontal direction with the piezoelectric stage and record the number of photocounts. If the laser spot under-fills the active area, we should expect a planar region with respect to photoresponse. If it is over-filling the SPAD, we should expect to see a convolution of the Gaussian beam profiler with a square area (step function) (Figure 3.16). The planar response across the junction indicates uniform breakdown characteristics in the planar region of the active area.

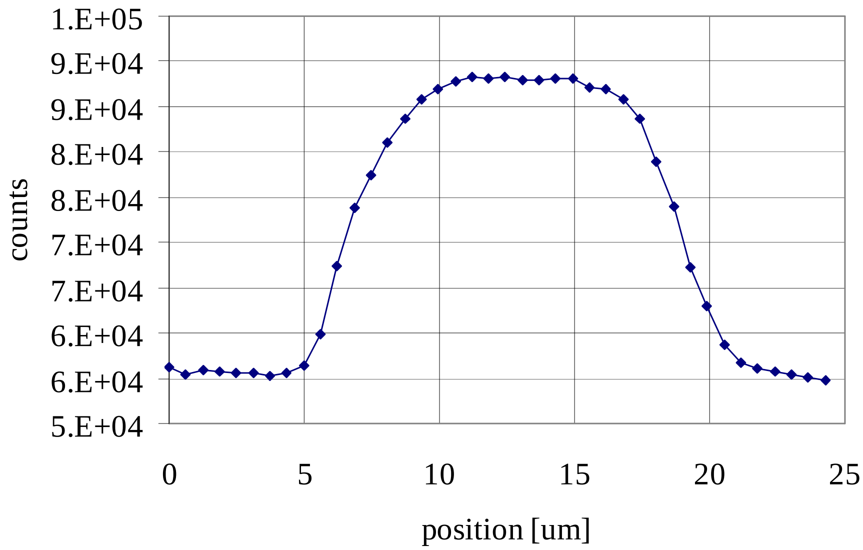


Figure 3.16 Position sweep of pixel with focused laser spot

We then move the pixel back to the center position of the beam for the optimal point of planar photoresponse and attenuate the laser power such that, on average, one photon is incident on the active area per pulse. Using time correlated single photon counting, a photon count histogram is obtained and photon detection statistics can be extracted. The laser is pulsed at 20MHz, resulting in photons impinging on the detector every 50ns. Non-correlated avalanche pulses register as a flat DC background. Signal-to-noise ratio (SNR) is proportional to the square root of the number of recorded photons ($N^{1/2}$), thus by simply acquiring more samples, increased SNR is obtained. This method becomes particularly important in applications where photon flux is low. This measurement setup was used to determine the single photon detection efficiency of 650nm for all structures on the SPAD test chip.

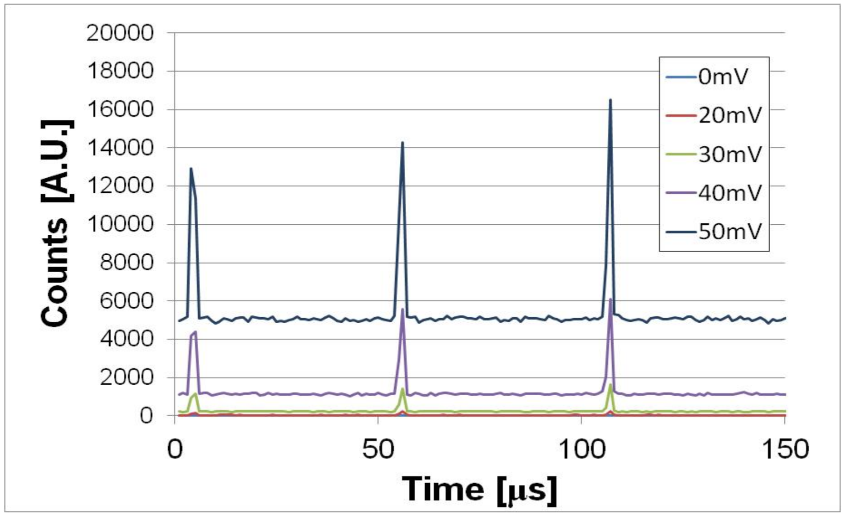


Figure 3.17 Time-Correlated Single Photon Counting Histogram

While detection efficiency measurements at a single wavelength is suitable using a pulsed excitation source, determining the spectral response is not as trivial, due to the need to sweep the wavelength of the excitation light. We use a continuous wave (CW) mercury arc lamp light source with broadband emission with calibration peaks. The lamp is output fed through a Digikrom 240 Monochromator with a 600 grooves/mm diffraction grating tunable from 350nm to 700nm. Because slits are used at the input and output port of the monochromator, a cylindrical lens is used to shape the beam more uniformly and then passed through an adjustable iris to create a circular beam profile. A beamsplitter is placed in series with a microscope objective to focus the light. The beamsplitter's spectral response is calibrated by measuring the split ratio with two calibrated power meters at each arm, with one after the microscope objective. The SPAD PCB then replaces that power meter so that incident photon flux can be monitored with the other power meter at all times. Photon flux is kept constant for all wavelengths by adjusting the monochromator slit size and is attenuated such that the average inter-arrival times of photons are more than twice the dead-time of the device and 10 times slower than the dark count rate.

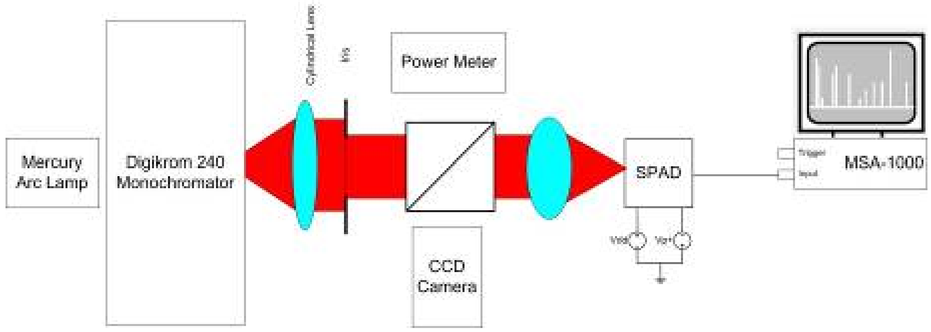


Figure 3.18 Spectral Response Setup Schematic

The relative spectral response curve is obtained with this method. The absolute spectral response curve is obtained by normalizing the relative detection efficiencies by the single photon detection efficiency obtained using the pulsed diode laser source at 650nm. The spectral response of an STI-bound SPAD compared to a diffused guard-ring SPAD is shown in Figure 3.19.

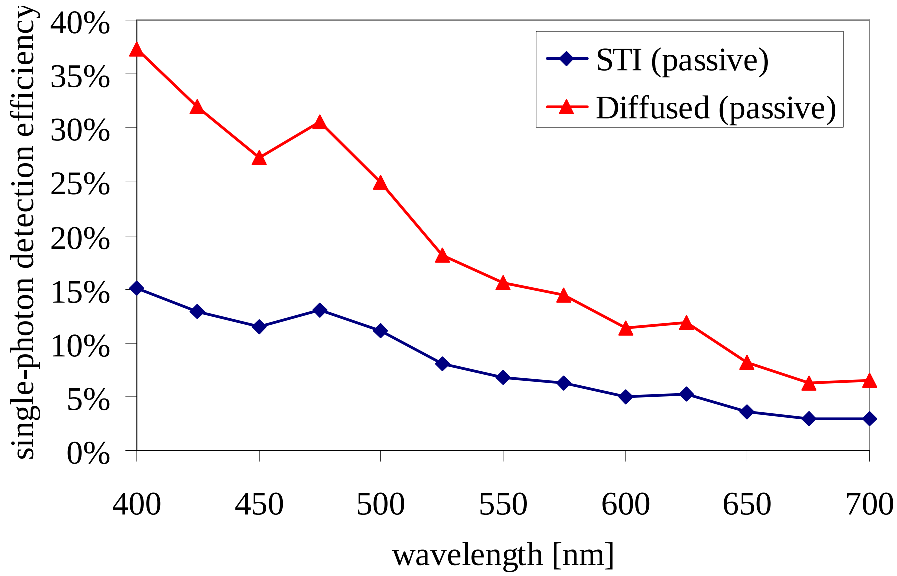


Figure 3.19 Spectral Response of STI-bound SPAD compared to diffused guard-ring SPAD

It is worth noting the higher detection efficiency of the diffused guard-ring SPAD. This is attributed to the noise characteristics of the STI-bound SPAD, where because of the high dark count rate of the device, it seldom has a chance to fully recharge, thereby reducing its effective single photon detection efficiency. Because the diffused SPAD is able to fully recharge, its sensitivity is greatly increased.

3.3.5. Timing Precision (Jitter)

A test chip was fabricated in an IBM 0.18-μm CMOS technology through the MOSIS service, packaged in a high speed Quad Flat No-Lead (QFN) glass-top package with ultra-short wiring to minimize capacitance and inductance. Two different output buffers were used—a source-follower and an inverter chain. The source-follower buffer uses a simple circuit which outputs a voltage proportional to the current flow, thus allowing a direct observation of the avalanche pulse. The inverter chain output buffer generates uniform-amplitude pulses for improved compatibility with pulse counting electronics.

We characterized devices with a source-follower output buffer to examine the temporal spread of an STI-bound device. We used an Agilent 86100A 20GHz sampling oscilloscope in conjunction with a Becker-Hickl BHL-600 pulsed laser at 650nm with <10 ps jitter. The oscilloscope was triggered using the laser's electrical sync signal and the SPAD output was fed into the electrical input. The oscilloscope recorded a time-of-arrival histogram of leading edges at a set voltage threshold.

STI-bound SPADs of 2µm x 2µm and 14µm x 14µm active areas are characterized to determine the mechanism with the greatest impact on the device jitter, namely longitudinal diffusion or lateral drift. Both devices were biased 7% beyond their breakdown voltage of 11V. The instrument response functions (IRFs) for both devices are shown in Figure 3.20. The jitter of the STI guard-ring SPADs is shown in

Table 3.1 and compared to previously reported SPAD devices.

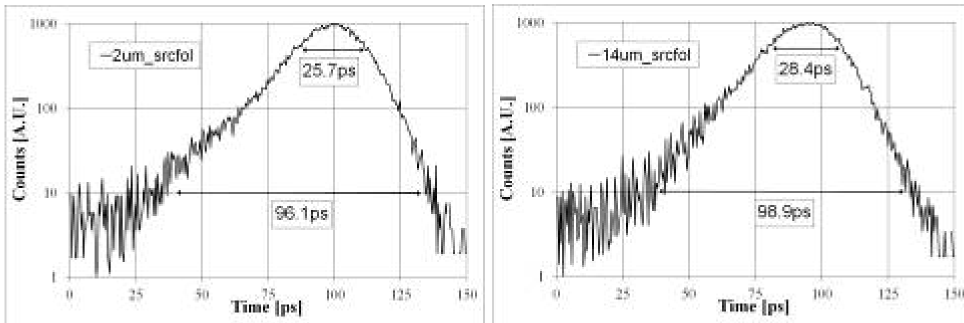


Figure 3.20 IRF of STI-bound SPADs of 2µm x 2µm (a) and 14µm x 14µm (b) active areas with source-follower output buffer biased 7% above its breakdown voltage

The nearly identical pulse arrival time histograms of the 2µm and 14µm devices demonstrate that lateral avalanche spreading, which was reported to be the dominant jitter source in larger devices [134],[135], is not the dominant jitter mechanism in these devices. The symmetric time-of-arrival histogram is typical of avalanches seeded in the high-field region; the lack of a long diffusion tail indicates that lateral drift and longitudinal diffusion are virtually eliminated.

The jitter performance of the STI-bound SPAD demonstrates a significant reduction of the FW(1/100)M compared to previously published SPAD device performance. The measured FWHM for the STI-bound SPAD is comparable to the best reported results of both custom fabricated and CMOS SPADs. The FW(1/100)M of the STI-bound SPAD is 3x better than that of the best reported CMOS SPAD devices and comparable to that of custom SPADs.

Table 3.1 Comparison of IRFs of STI-bound CMOS SPAD and prior SPAD art

Device	Process	Guard-ring	FWHM	FW(1/100)M
2 μm STI-	0.18μm CMOS	STI	26.7 ps	96.1 ps
14 μm STI-	0.18μm CMOS	STI	27.4 ps	98.9 ps
Cova [136]	Custom Double Epitaxial	Double Epitaxy	28 ps	Not shown
Spinelli [137]	Custom Silicon	Diffused	35 ps	950 ps
Tisa [77]	0.8μm HV CMOS	Diffused	36 ps	1.1 ns
Tisa [138]	0.35μm HV CMOS	Diffused	39 ps	617 ps
Niclass [139]	0.35μm HV CMOS	Diffused	80 ps	290 ps
Niclass [84]	130nm CMOS	Diffused	144 ps	(>4 ns)
Gersbach [85]	130nm CMOS Image Sensor	STI w/ passivation	125ps	1ns

In order to further demonstrate the benefit of the STI guard-ring in terms of timing resolution and diffusion tail, we compared a STI-bound SPAD with a traditional diffused guard-ring SPAD fabricated on the same die. Both test devices had 7μm x 7μm active areas and identical sampling circuitry, and were characterized with a setup similar to the one described earlier, with the sampling oscilloscope replaced with a Picoquant Picoharp300 TCSPC pulse counter. The laser's electrical trigger served as the "START" input to the PicoHarp300 and the SPAD output was routed to the "STOP" channel input of the counter. The devices were also biased at 7% beyond their breakdown voltage.

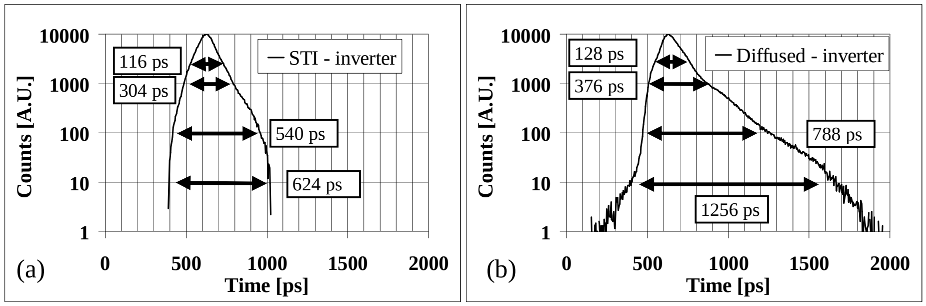


Figure 3.21 IRF of the STI-bound SPAD (a) and the diffused guard-ring SPAD (b) with inverter output buffers biased at 7% above their breakdown voltage.

The STI-bound SPAD IRF (Figure 3.21) exhibits a factor of two improvement in the FW(1/1000)M, 1256ps to 624ps, comparable with results from cooled superconducting SPADs [140]. The longer IRFs shown in Figure 3.21 compared to those in Figure 3.20 with source-follower output buffers are due to a suboptimal sensing threshold. Ideally, the sensing threshold should be positioned low at the onset of the avalanche pulse [137]. The source-follower output stage follows the diode's n-well terminal voltage, making it possible to search for the optimal sensing threshold. The inverter stage has a fixed threshold defined by the relative sizes of its transistors. Consequently, its sensing threshold could not be tuned in the current test chip and the respective IRF was longer.

In addition to the low jitter of the STI-bound SPAD device, it also exhibits shorter dead-times and higher fill factor than diffused-ring SPADs [131]. This comes at the expense of a relatively high dark count rates (in the 10^4-10^6 counts per second) and afterpulsing, possibly due to interface states at the SiO_2-silicon boundary. Time gating and active recharge, which are commonly used with SPADs can reduce the effects of this noise to acceptable levels.

3.4. Single Photon Frequency Up-conversion via Hot-carrier Electroluminescence

3.4.1. Introduction

Single photon detection of infrared (IR) light is of particular interest in optical communication links, in vivo optical imaging, eye-safe LIDAR, optical time-domain reflectography, and in semiconductor failure analysis. IR single photon detectors should have high sensitivity, short dead-times for fast bit-rates in the tens to hundreds of MHz, low noise, and high timing precision. Other factors such as active area, scalability to arrays, pixel pitch, and cost are important. In addition, low power consumption and operation at room temperature are desirable.

IR sensitive detectors have been implemented in several ways. As described in Section 1.3.4, superconducting nanowires used as microcalorimeters in response to single photons have been developed and have demonstrated good quantum efficiencies, excellent timing precision, and low dark count rates [57-59],[141-145]. However, they require special fabrication processes and cryogenic cooling to <4K making them impractical for large scale integration. Further large arrays have not been demonstrated yet.

Solid-state IR sensitive SPADs require low bandgap absorbing materials and have been fabricated with planar geometries using III-IV semiconductor materials with separate absorption and multiplication regions. Photons are absorbed in a thick lightly doped InGaAs layer and then swept away by an electric field toward a high electric-field InP multiplication region for impact ionization. Similar to silicon SPAD described earlier, IR SPADs need to quench the avalanche current and is typically done with off-chip external active or passive quenching circuit, usually in silicon. The off-chip quenching and recharge circuitry is the main source of capacitance in the device.

Recently, there has been development of a new type of quenching scheme which uses a resistive layer deposited on top of the p-n junction [146] eliminating the need for external quenching and has been implemented at UCSD by Zhao et al. with InGaAs/InAlAs materials [147]. This device uses a transient carrier buffer (TCB) to form an energy barrier as a negative feedback mechanism where charges accumulate at the resistive layer and create a charge screening effect which quenches the avalanche current. The carriers slowly escape the barrier which leads to device recovery. Integration of a self-quenching resistive layer eliminates the need for external circuitry and reduces overall junction capacitance resulting in faster recharge times and reduced afterpulsing effects, a common problem with III-V detectors. Further, power dissipation is reduced.

There have been trends to up-convert IR photons to visible wavelengths for detection with silicon photodetectors. Using non-linear sum frequency generation, wavelength conversion can be achieved with periodically poled lithium niobate (PPLN) waveguides coupled to silicon SPADs with good jitter performance of 40ps FWHM [148]. Peak up-conversion efficiencies of 5-7% after losses have been demonstrated for single detectors, though with the use of a 980nm diode laser with 300mW power to pump the non-linear crystal.

Recently, a novel read-out scheme has been introduced by Finkelstein et al [149] where electrical interconnection between the IR SPAD and CMOS SPAD is not required resulting in low power consumption and the prospect of scalability to arrays. Instead, optical coupling is achieved by exploiting a normally adverse byproduct of recombination of carriers and photoemission. Figure 3.22 shows the device concept where an IR SPAD detects IR photons and re-emits secondary high energy photons at visible wavelengths which are subsequently detected with a CMOS SPAD.

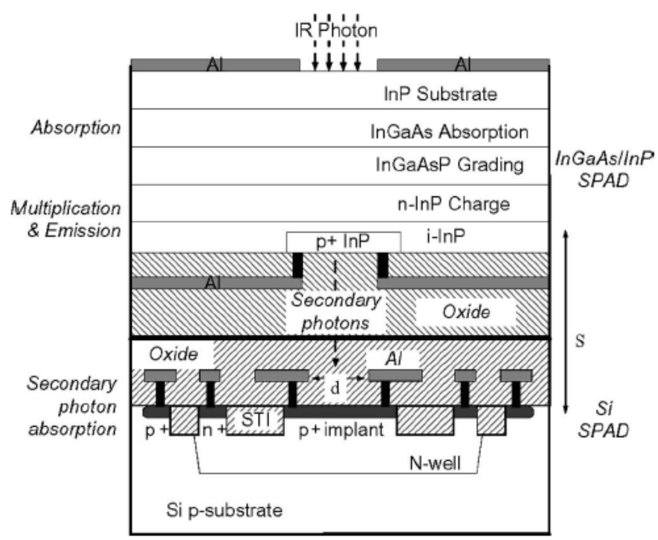


Figure 3.22 Up-conversion schematic [149]

This effect is termed hot-carrier luminescence and is the source of optical crosstalk between adjacent pixels in photodiode arrays. During each avalanche pulse, millions of carriers traverse the junction. Statistically, a portion of these hot carriers will recombine and emit a photon. For direct bandgap materials such as InP and GaAs, the photon emission rate is given by:

$$R_d(\eta\omega) \propto \eta\omega\left(\eta\omega - E_{gd}\right)^{1/2} f(E)\left[1 - f\left(E - \eta\omega\right)\right]$$

Equation 3.2 Photon emission rate for direct bandgap materials

where $\eta\omega$ is the emitted photon's energy, E_{gd} is the direct bandgap, E is the electron energy above the conduction band, and $f(E)$ and $\left[1 - f\left(E - \eta\omega\right)\right]$ are the hot electron and hole distributions, respectively, which are temperature dependent [150]. Because the hot carriers have excess energy above the bandgap, the photon energy is up-converted to a higher energy photon. Total up-conversion efficiency is dependent on a number of factors such as single IR photon detection efficiency of the IR SPAD, photon conversion efficiency, wavelength shift of the up-converted photons, detection efficiency and spectral response of CMOS SPAD. Further, the up-converted photons must be optically coupled with the CMOS SPAD active area despite many of them being self-absorbed in the upper layers of the IR SPAD.

This scheme stands to simplify manufacturing issues for hybrid devices by eliminating the need for wires and reducing parasitic capacitances. Further it reduces avalanche charge, resulting in lower power consumption and decreased afterpulsing. One of the challenges of hybrid devices is determining an efficient optical coupling scheme as well as optimal operating conditions. In the following sections we will describe free-space SPAD-to-SPAD optical coupling techniques. First we will demonstrate free space up-conversion using silicon SPADs, which was done while a 30μm III-V IR SPAD was under development. Second, we will demonstrate proof of concept of an IR-SPAD to CMOS SPAD uplink at room temperature. Third, we will show design of an uplink system where an IR-SPAD can be cooled for optimal dark current characteristics while coupling that light to a CMOS SPAD running at room temperature.

3.4.2. Emission Spectrum and Yield Measurements

The expected light emission yield from the silicon SPAD is expected to be lower due it being an indirect bandgap material, requiring both energy and momentum conservation through phonons. Emission spectrum and yield measurements are necessary to calculate the system efficiency. Electroluminescence yield is the number of emitted photons per hot carrier. To determine the absolute electroluminescence yield and emission spectrum, a *Princeton Instruments* Intensified cooled CCD is used with a monochromator grating. Spectral response calibration of the cooled CCD camera was done using a broadband halogen lamp and also a mercury arc lamp with known spectral peaks.

The emission spectrum of an avalanching Si SPAD is measured and compared with published absolute emission spectrum of silicon [151] (Figure 3.24). Following, the absolute collection efficiency of the CCD is determined.

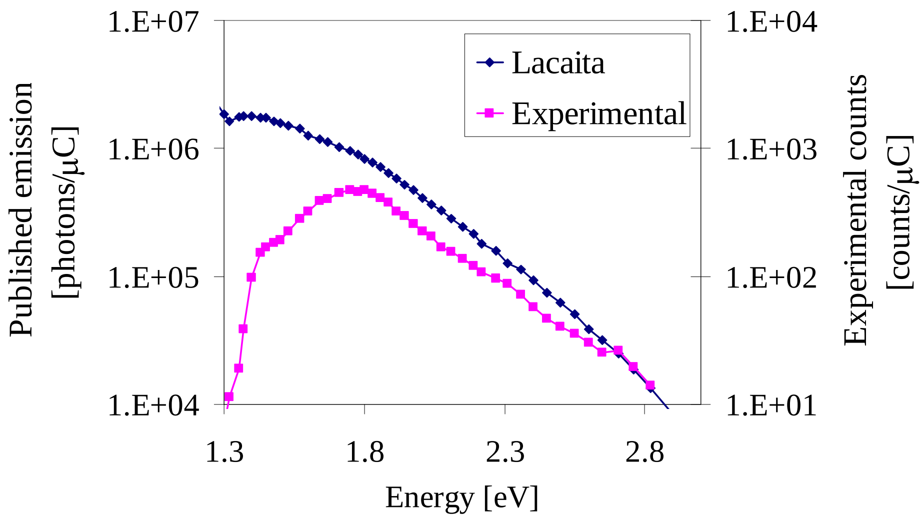


Figure 3.23 Emission Energy Spectrum of Silicon SPAD compared to published results

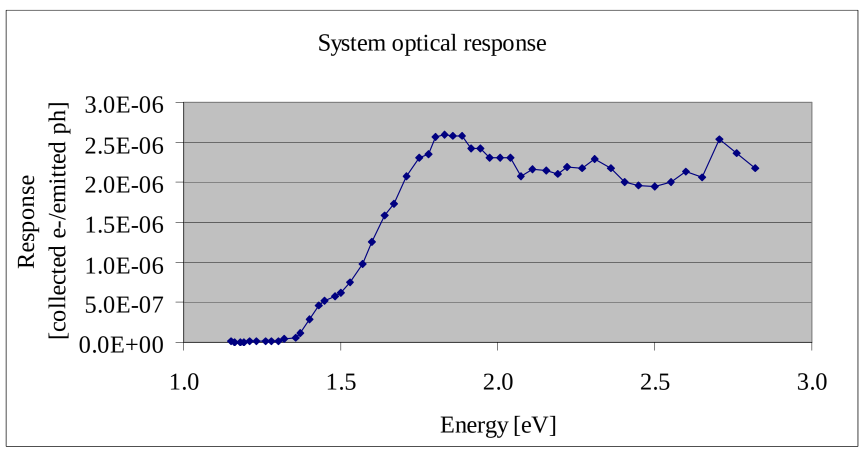


Figure 3.24 Absolute System Collection Efficiency

By normalizing the collected photoelectrons by the number of emission photons, the absolute system collection efficiency is obtained (Figure 3.24) showing strong components at 1.7eV to 2.8eV.

With a calibrated electroluminescence measurement system, the electroluminescence from an InGaAs/InAlAs SPAD device provided by Zhao et al. (Figure 3.25) is measured. An intensity image of the 30μm SPAD electroluminescence is shown in Figure 3.27. The uniform intensity distribution indicates uniform breakdown at the planar regions of the device.

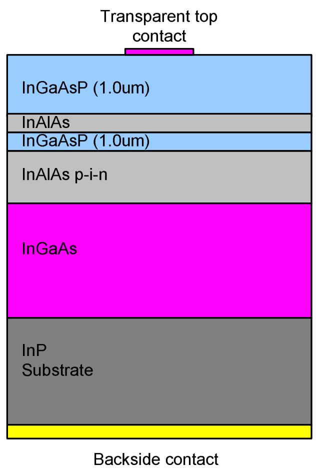


Figure 3.25 InGaAs/InAlAs SPAD device cross-section

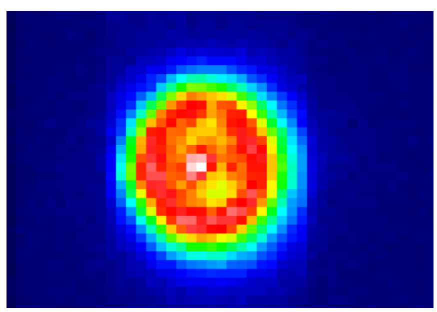


Figure 3.26 Electroluminescence Intensity of InGaAs/InAlAs SPAD

The electroluminescence spectrum of the IR sensitive SPAD was measured with a spectral peak at 1.24eV, which is lower than predicted. This was most likely due to photon self absorption at the top layers of the device. However, an electroluminescent yield of 3.3×10^{-5} photons/carrier at the surface was measured, indicating capability for high-efficiency up-conversion.

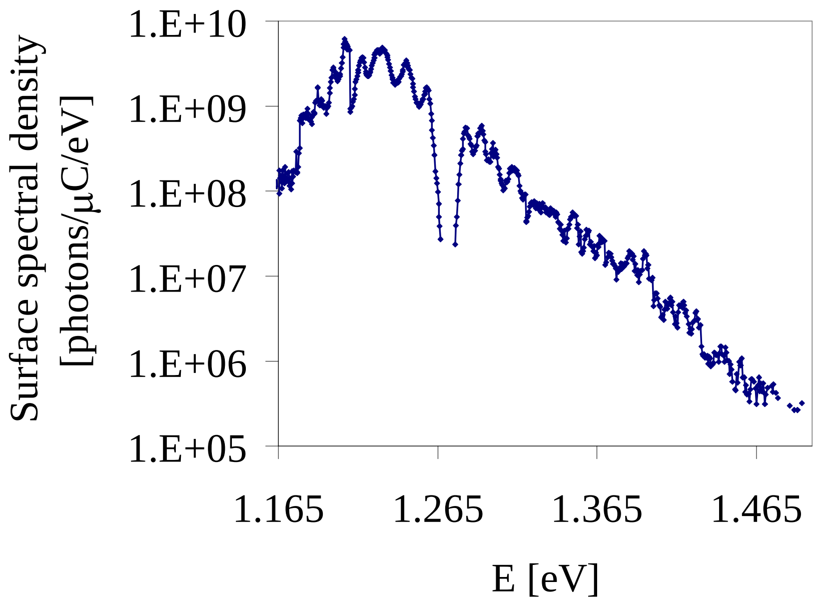


Figure 3.27 Electroluminescence Surface Spectral Density of InGaAs/InAlAs SPAD

3.4.3. Free Space Up-conversion Link: Silicon-to-Silicon

Initial proof-of-concept was demonstrated with a free space optical up-conversion link, with two 7µm CMOS SPADs used—one as a photon source and the other as a detector. In this setup, diffused junction SPADs were used because of their lower dark count characteristics and higher detection efficiencies compared to the STI-bound SPAD. As described earlier, recombination in the high electric-field region can result in emission of photons, termed hot-carrier electroluminescence (Figure 3.28). The challenge is to efficiently couple light emission from one SPAD to the other.

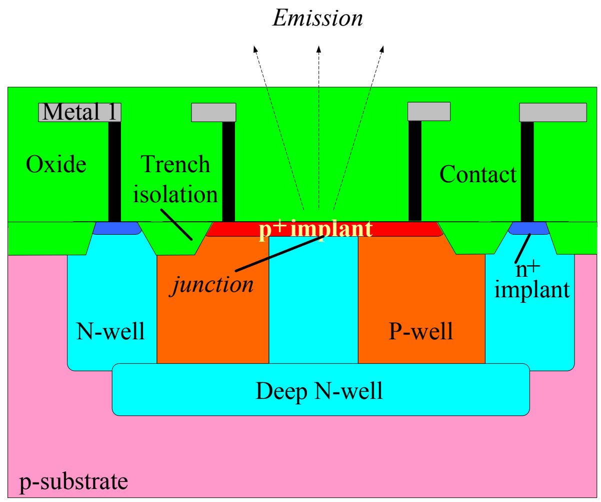


Figure 3.28 Hot-carrier electroluminescence from Si SPAD

A free space two lens optical system is developed for precise alignment of the 7μm pixels. Alignment of what is essentially a point source to a point detector is challenging. A method of referencing is employed using a sensitive integrating Si CCD Beamprofiler (Coherent LaserCam) where the CMOS SPAD receiver is mounted on the beam profiler and onto a 3D micrometer stage which is in turn mounted on a 2D motorized piezoelectric stage. The displacement between the mounted CMOS SPAD pixel and an individual pixel on the beamprofiler is determined using a fixed laser source and moving the detectors a known distance with the motorized translation stage.

Following, the source SPAD is biased several volts above breakdown such that there is sufficient photoemission for detection with the beamprofiler. The beamprofiler and SPAD source pixel are both moved into the foci of the lens system to optimize the light collection and focusing efficiency. The beamprofiler-SPAD mount is moved to the known beamprofiler pixel-to-SPAD pixel displacement with the motorized stage and micrometer stage. Because of errors in the displacement, optimal pixel alignment is not trivial. Hence, even after moving the SPAD into the supposed focus of the lens system, slight displacements of microns are present which result in minimal to zero optical coupling. Because of the low total photoemission, a DC measurement of total photon counts is unsuitable because of high dark count rates of the detector. Hence a time-correlated method must be used to improve SNR and to move the receive pixel to search for the source emission.

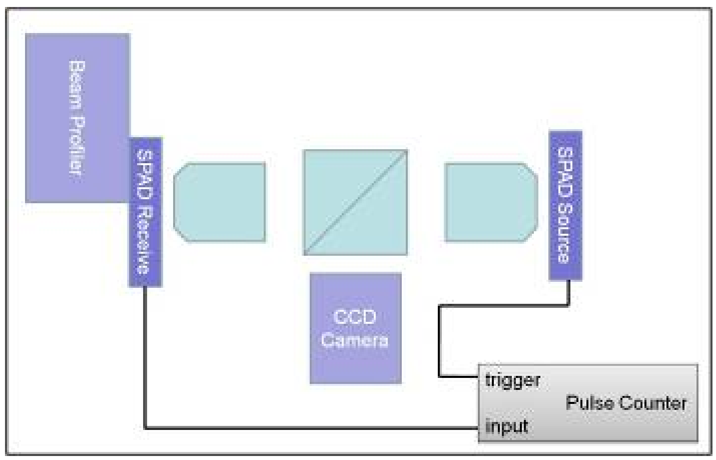


Figure 3.29 SPAD-to-SPAD optical link schematic

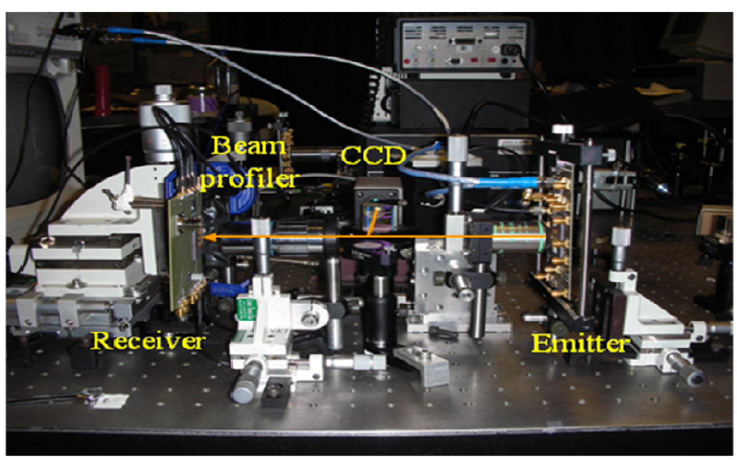


Figure 3.30 Photograph of optical up-conversion link

As proof-of-concept of an optical link using hot carrier electroluminescence, one can collect photons from an avalanche pulse, regardless of whether it is photon induced or not. In the silicon SPAD-to-SPAD uplink, we use a dark pulse from the source SPAD to trigger the pulse counter and then look for correlated avalanches in the receiver SPAD. Because of the high dark count rate of the SPAD devices, long time-correlated acquisition times are necessary to pull the signal out of the noise. Figure 3.31 shows correlated avalanche pulses from the receiver SPAD with a strong afterpulse peak following the primary pulse. The background pulses are subtracted since they are uncorrelated to the triggering avalanche pulse. To demonstrate that we are not observing electrical crosstalk through the electronics or power supplies, a screen was placed in the optical link to show that it was indeed an optical signal being detected. An internal efficiency of $1x10^{-4}$ detected events per source SPAD avalanche was measured.

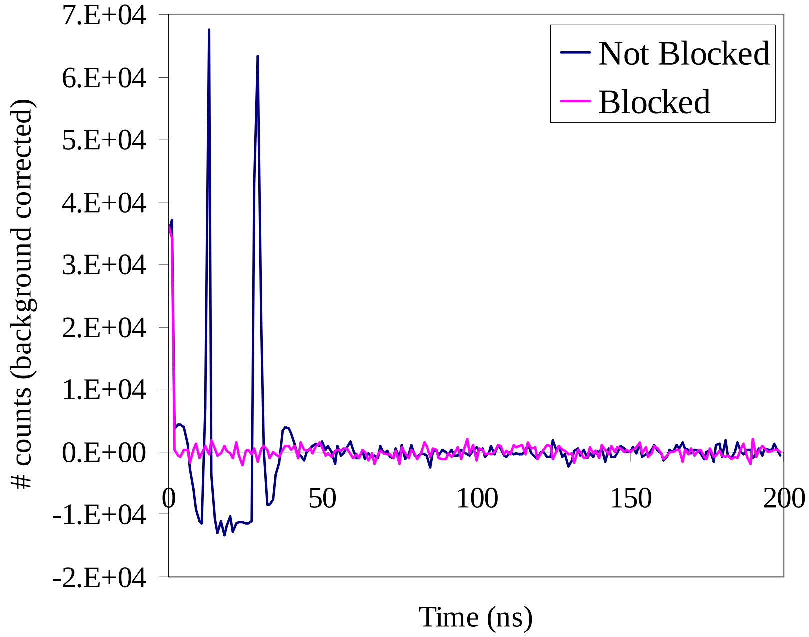


Figure 3.31 Time correlated counts from Si-to-Si uplink

3.4.4. Free Space Up-conversion Link: InGaAs/InAlAs-to-Silicon

Following the demonstration of a free space Si-to-Si up-conversion link, a III-V SPAD to Si SPAD uplink is demonstrated. The experimental setup is similar, except an active 1550nm laser source (IdQuantique id300) is used to illuminate the IR SPAD (Figure 3.32). A pattern generator (Agilent 81110A) is used to drive the laser source at 250MHz with 4ns pulse widths. The pulse generator is then synced with the pulse counting board and records correlated pulses at the CMOS SPAD output.

A 30μm InGaAs/InAlAs with a resistive TCB quenching layer fabricated by Zhao et al. was used in this demonstration. Similar to the Si-to-Si uplink alignment,

alignment of these two SPAD devices proved to be challenging. A similar referencing method was used in conjunction with time-correlation measurements and a 250MHz bit-rate was demonstrated (Figure 3.33). The uncorrelated avalanche pulses were subtracted for a clearer view of the optical signal.

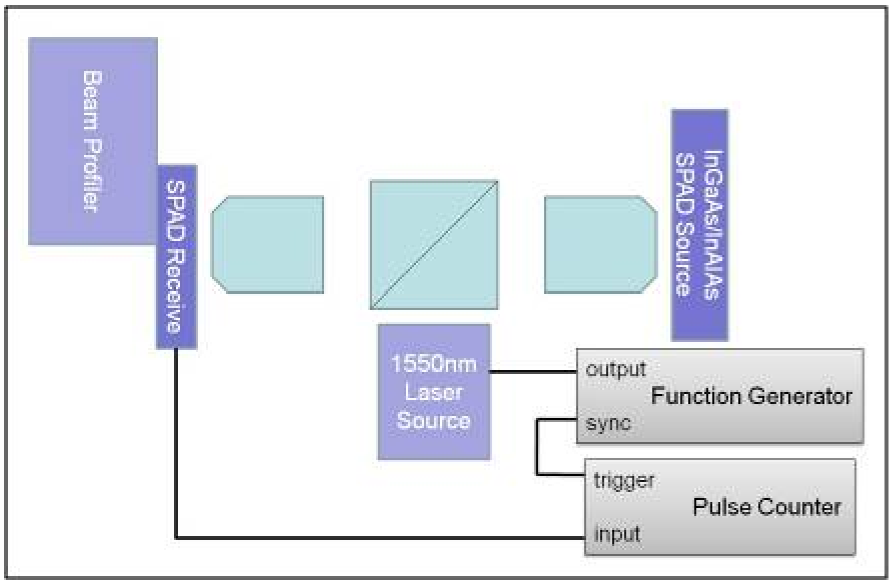


Figure 3.32 IR SPAD to Si SPAD up-conversion optical link

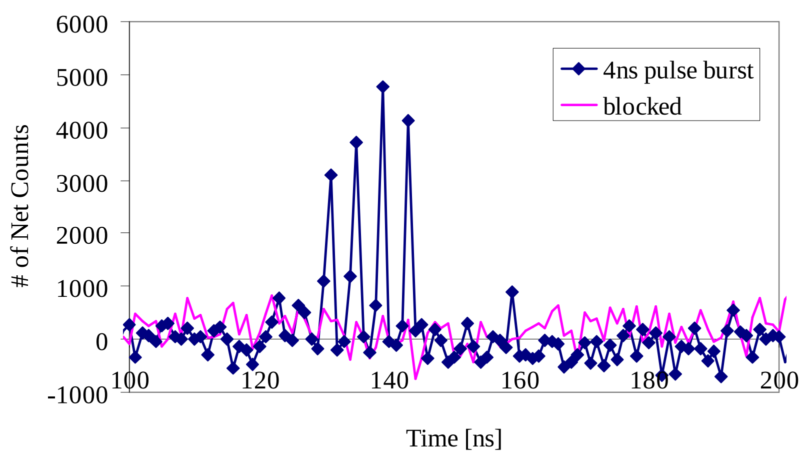


Figure 3.33 250MHz bit rate up-conversion link

Similarly, a screen was place in between the optical link to demonstrate that the signal was optical and not an artifact from electrical crosstalk. This was the fastest up-conversion rate possible with our current hardware and is comparable to superconducting SPADs. Further, room temperature operation results in high dark currents for the IR SPAD, preventing the SPAD from fully recharging resulting in poor detection efficiencies. In addition, emission efficiency is reduced due to self-absorption at the upper layers, leading to the need for an improved design.

3.4.5. Cooled Up-conversion Optical Link Development

Following Zhao et al., a second generation TCB device has been developed by You et al. [152]. III-IV SPADs suffer from high dark currents and need to be cooled. It has been determined that an operating temperature of 240K is suitable for best performance as shown in Figure 3.34. Because of strong afterpulsing effects with the CMOS SPAD described in Section 3.3.3, a hybrid cooling scheme is desirable, where the IR SPAD is cooled and the CMOS SPAD is operated at room temperature.

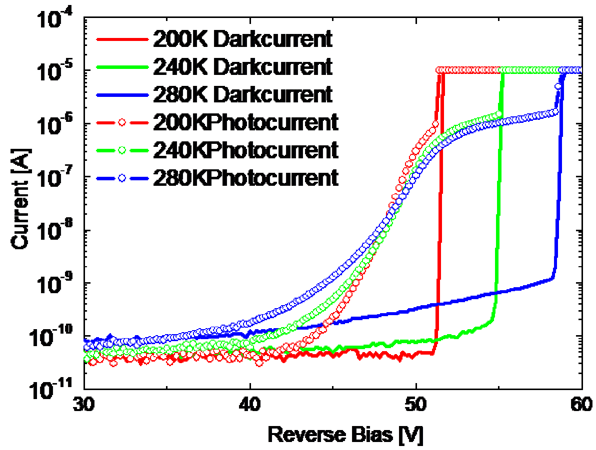


Figure 3.34 Temperature dependent I-V characteristics of TCB device [152]

A MMR Technologies Low Temperature Microprobe (LTMP) system with a clear optical window is used. The LTMP chamber houses a micro miniature refrigerator and two micromanipulators for probing mounted devices on the refrigerator. The LTMP is used in conjunction with a programmable temperature controller, a dry nitrogen source, gas filter, and high vacuum pump. The MMR refrigerators operate using the Joule-Thomson effect where a gas such as nitrogen is allowed to expand through a porous plug or fine capillary tube at high pressure resulting in gas cooling. This effect is relatively small at 0.1K/atm at ambient temperatures. However, this effect can be magnified by allowing the expanded cooled gas to pass through a countercurrent heat exchanger, thus precooling the incoming high pressure gas. This feedback cooling process continues until the gas is liquefied or the temperature drop is limited by the heat load of the device mounted on the refrigerator stage [153].

The vacuum pump is first attached to the vacuum port of the LTMP chamber and ensured to be able to pump the chamber down to <5mTorr. High pressure nitrogen gas is used for cold finger operation. The outlet of the gas is attached to the inlet of the gas filter/dryer unit for purification and moisture removal. Using the programmable temperature controller, temperatures of 80K are capable of being realized.

Similar methods of free space optical coupling are used, with a slightly different lens configuration due to the vertical orientation of the optical window (Figure 3.35 and Figure 3.36). Unfortunately, at the time of writing of this dissertation, a suitable IR-SPAD with high detection efficiencies and satisfactory electroluminescence was not yet realized.

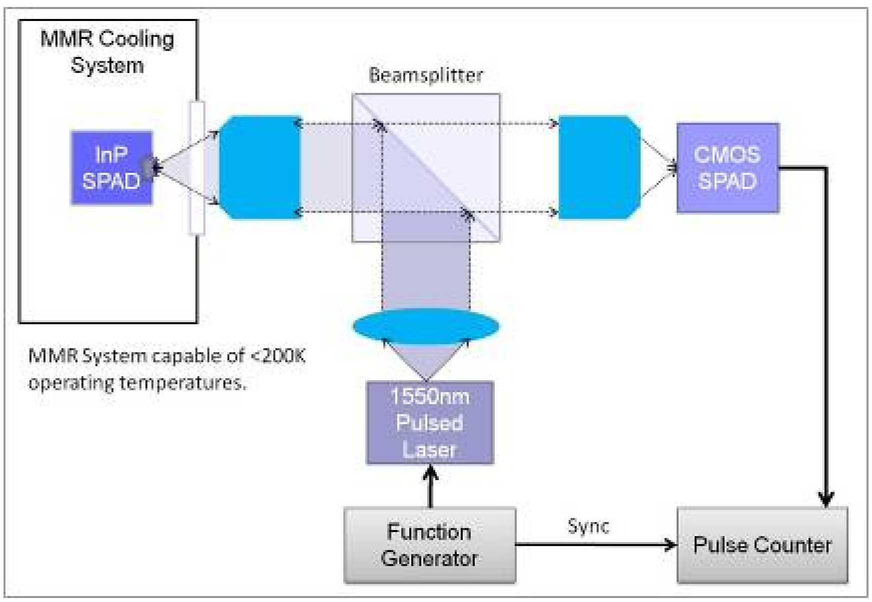


Figure 3.35 Hybrid Free Space Up-conversion Schematic

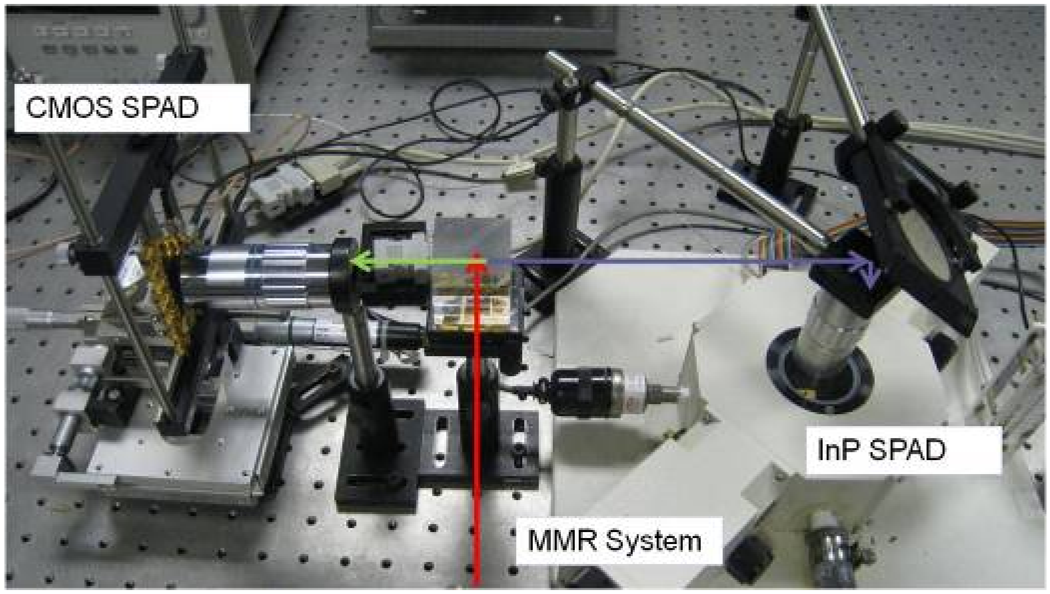


Figure 3.36 Hybrid Free Space Up-conversion Photograph

3.5. Conclusions

In this chapter we have described the design, modeling, and characterization of a CMOS SPAD fabricated in a 0.18μm deep-submicron technology with a novel planarizing guard ring structure. Device modeling was done using ISE-TCAD simulation software for electric field distributions at device breakdown. An STI-bound SPAD is compared to a diffused guard-ring SPAD. We demonstrate state-of-the-art performance in sampling rates and timing precision. A peak spectral sensitivity at 400nm is measured, due to a shallow and thin depletion region. In addition, noise characteristics are measured to be unexpectedly high and are attributed to afterpulsing from charges trapped in the oxide guard ring.

Further, in this chapter, we have demonstrated proof-of-concept of free space up-conversion of 1550nm using a novel optical readout technique. A previously adverse effect is turned into an optical coupling scheme. Using a III-IV IR sensitive SPAD which integrates quenching via a resistive layer on top of the p-n junction eliminating the need for external quenching circuitry, junction capacitances are significantly reduced, resulting in shorter dead-times and reduced afterpulsing effects. Recombination of hot electrons results in photon emission at energies greater than the bandgap energy, resulting in a wavelength shift towards visible wavelengths. These up-converted photons are then collected and focused onto a silicon CMOS SPAD with integrated processing circuitry. 250MHz bit-rates are demonstrated in a free-space optical link at room temperature. Potential for simplified manufacturing flow of hybrid devices and arrays is considered for future development. Further, a hybrid optical setup is developed for next generation device characterization.

Acknowledgement:

This chapter, in part, is a reprint of the material as it appears in the following publications:

- M. J. Hsu, H. Finkelstein, and S. Esener, "A CMOS STI-bound Single Photon Avalanche Photodiode with 27ps Timing Resolution and a Reduced Diffusion Tail," IEEE Electronic Device Letters, vol. 30, no. 6, 2009.

- Up-conversion Imager Project Status Update from September 29, 2007 (combined effort from Esener and Lo research groups)
- UCSD CMOS Compatible IR-SPAD Project Update from February 1, 2010 (combined effort from Esener and Lo research groups)

4. ACOUSTO-OPTIC SIGNAL ENHANCEMENT

4.1. Introduction

In the past two decades researchers have used ultrasound to modulate photons in an effort to improve spatial resolution of optical imaging, termed acousto-optic imaging. When photons pass through a volume of insonated tissue, they are modulated (tagged) at the ultrasound frequency due to changes in the local refractive index, optical absorption, and scattering properties. After the photons are collected with an optical detector, modulated photons can subsequently be separated from non-modulated photons through electronic filtering, optical interferometers, lock-in amplifiers, or time-averaging. Since ultrasound can be focused to mm-sized voxels, high spatial resolution images can be reconstructed using tagged photons. The detection of modulated photons in acousto-optic imaging, particularly in a focused ultrasound beam where the voxel size is small, is challenging because few tagged photons are collected compared to untagged photons, resulting in low signal-to-background ratios (SBR). We explore various methods of improving acousto-optic SNR through the use of the CMOS SPAD and through the novel use of ultrasound microbubbles.

4.2. Acousto-optic Imaging in Turbid Media

Optical imaging is a powerful tool that can provide sensitive contrast of tissue properties such as cancer-characteristics like angiogenesis, oxy- deoxy- hemoglobin concentrations, hyper-metabolism, and pleomorphism. More specifically, this information can be extracted by measuring the local optical scattering and attenuation properties of the tissue at the cellular and sub-cellular level. However, because of the highly scattering nature of tissue, high resolution optical imaging is limited to depths of 1 mm [154]. Techniques such as diffuse-optical tomography (DOT) [155-157] and optical coherence tomography (OCT) [158-162] have been developed to improve in vivo optical imaging depths and resolutions, respectively. Ultrasound imaging provides good spatial resolution, but poor contrast because of its reliance on the mechanical properties of the tissue and making early stage tumor detection difficult. There has been movement to improve on the limited depth and resolution of optical imaging by combining optics and ultrasound to achieve optical contrast at ultrasound spatial resolutions, namely acousto-optic tomography and photo-acoustic tomography.

Table 4.1 Summary of various in-vivo optical imaging techniques

	OCT	**DOT**	**US**	**AOT/PAT**
Contrast	+	++	-	++
Resolution	10µm	5mm	150µm	150µm
Imaging Depth	1mm	5cm	3cm	3cm
Speckle Artifacts	Strong	None	Strong	None
Scattering Coefficient	100 cm^{-1}	100 cm^{-1}	0.3 cm^{-1}	0.3 cm^{-1}

Acousto-optic (AO) imaging is based on tagging photons in the ultrasound focal zone by changing the local optical properties of the tissue with pressure waves. Several groups have led efforts in AO imaging from concept development and imaging of tissue phantoms [163-166], typically with the use of long coherence light sources. Various other methods have been used for AO imaging such as a lock-in detectors [167], parallel detection [168], and frequency sweeping [169]. Each of these techniques relies on coherent light.

Three mechanisms are attributed to acousto-optic modulation of light in turbid media and are outlined in Figure 4.1.

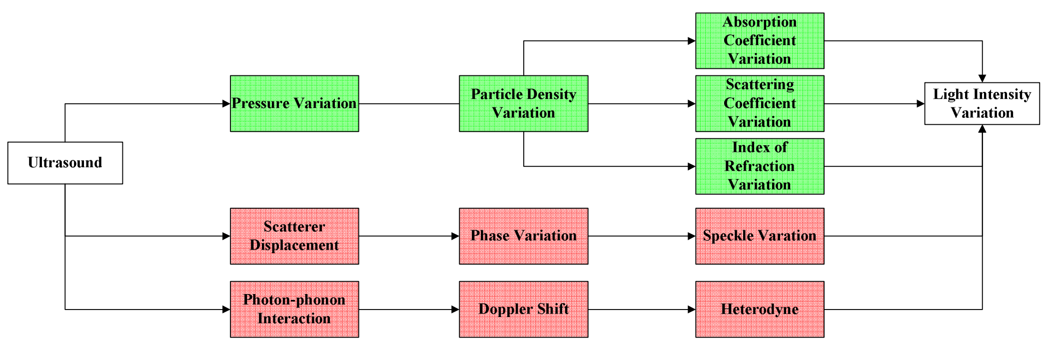


Figure 4.1 Mechanisms for Acousto-Optic Modulation in Turbid Media

The first mechanism is modulation of local optical properties of the media via pressure waves at the ultrasound focus and has been modeled by Mahan et al. [170]. As the ultrasound propagates through the turbid media, positive and negative pressure waves compress and expand the scattering media, respectively, leading to local density variations. These local density variations induce changes in the local optical attenuation, scattering, and index of refraction resulting in an intensity modulation of light.

The second mechanism is based on ultrasound induced displacement of optical scatterers. The displacement of optical scatterers modulates the optical path lengths of the photons that traverse the region, resulting in an intensity differential. Leutz et al. model this mechanism [163], though only valid when the scattering mean path is much greater than the acoustic wavelength. The third mechanism is based on modulation of the index of refraction resulting in an optical phase change between consecutive scattering events. In both the second and third mechanism, multiply scattered light accumulates along the light propagation path and results in modulation of the speckle intensity. Further, these latter two mechanisms require the use of coherent light and need complex detection schemes such as Fabry-Perot interferometers or lock-in amplifiers to detect the modulation at the ultrasound driving frequency. In the next subsection, an analytical model based on Wang [171] will be presented for ultrasound modulation of coherent light, namely the second and third mechanisms described above.

4.2.1. Ultrasound Modulation of Coherent Light

An acoustic plane wave is assumed to be incident on a uniformly scattering medium. A derivation of this analytical solution can be found in [171]. The autocorrelation function of the electric field can be expressed as:

$$G_1(\tau) = \int_0^\infty p(s)\left\langle E_s(t)E_s^*(t+\tau)\right\rangle ds$$

Equation 4.1 Autocorrelation function of electric field

where E_s is the electric field of the scattered light along path length s and *p(s)* is the probability density function (pdf) of s. The autocorrelation function only takes into account contributions of the acoustic wave; Brownian motion is considered negligible. The scattering medium is approximated to be made up of non-interacting spherical Rayleigh scatterers. The autocorrelation becomes the following after encountering s/l scattering events along path s.

$$\left\langle E_s(t)E_s^*(t+\tau)\right\rangle = \left\langle \exp\left\{-i\left[\sum_{j=1}^{s/l+1}\Delta\phi_{nj}(t,\tau)+\sum_{j=1}^{s/l}\Delta\phi_{dj}(t,\tau)\right]\right\}\right\rangle$$

Equation 4.2

where $\Delta\phi_{nj}(t,\tau)=\phi_{nj}(t+\tau)-\phi_{nj}(t)$ and ϕ_{nj} is the phase variation induced by the modulation of the index of refraction along the *jth* free path and $\Delta\phi_{dj}(t,\tau)=\phi_{dj}(t+\tau)-\phi_{dj}(t)$ and ϕ_{dj} is the phase variation induced by the path length displacement of the *jth* scatterer following the *jth* free path. The phase variation from the *jth* free path is given by:

$$\phi_{nj}(t)=\int_0^{l_j} k_0\Delta n(r_{j-1},s_j,\theta_j,t)ds_j$$

Equation 4.3 Phase variation

where l_j is the length of the *jth* free path, k_0 is the optical wave vector, r_j is the location of the *jth* scatterer, s_j is the distance along the *jth* free path, and θ_j is the angle between the optical wave vector of the *jth* free path and acoustic wave vector k_a. Δn is the modulated index of refraction:

$$\Delta n(r_{j-1}, s_j, \theta_j, t) = n_0 \eta k_a A \sin(k_a \cdot r_{j-1} + k_a s_j \times \cos\theta_j - \omega_a t)$$

Equation 4.4 Modulated index of refraction

Integrating Equation 4.3 yields:

$$\phi_{nj}(t) = 2n_0 \eta k_a A \sin(k_a \cdot r_{j-1} + k_a l_j \cos\theta_j / 2 - \omega_a t) \times \sin(k_a l_j \cos\theta_j / 2) / \cos\theta_j$$

Equation 4.5

The phase variation from the *jth* scattering event is given by:

$$\phi_{dj}(t) = -n_0 k_0 \left(\hat{k}_{j+1} - \hat{k}_j\right) \cdot A \sin(k_a \cdot r_j - \omega_a t)$$

Equation 4.6

where $\hat{k}_j$ is the unit optical wave vector of the *jth* free path and ***A*** is the acoustic amplitude vector. Equation 4.1 becomes the following in the diffusion limit:

$$G_1(\tau) = \int_0^\infty p(s) \exp\left\{ -\frac{2s}{l} \left(\delta_n + \delta_d\right) (n_0 k_0 A)^2 [1 - \cos(\omega_a \tau)] \right\} ds$$

Equation 4.7

A coherent optical beam is assumed to be incident on a slab thickness of *L* with light transmission detected at a point. The function *p(s)* is given by diffusion theory with a zero boundary condition. After integration, Equation 4.7 becomes:

$$G_1(\tau) = \frac{\frac{L}{l} \sinh\left[\sqrt{\varepsilon[1 - \cos(\omega_a \tau)]}\right]}{\sinh\left[\frac{L}{l}\sqrt{\varepsilon[1 - \cos(\omega_a \tau)]}\right]}$$

Equation 4.8

with $\varepsilon = 6(\delta_n + \delta_d)(n_0 k_0 A)^2$. Finally, the speckle modulation intensity at frequency $n\omega_a$ is calculated based on the Wiener-Khinchin theorem:

$$I_n = \int_0^{T_A} \cos(n\omega_a \tau) G_1(\tau) d\tau / T_a$$

Equation 4.9

with acoustic period T_a. Modulation depth is defined as the ratio between the intensity of the fundamental frequency compared to the non-modulated intensity:

$$M = \frac{I_1}{I_0}$$

Equation 4.10

4.3. Detection of Ultrasound-modulated photons

Acousto-optic imaging suffers from poor SNR because of the small number of ultrasound tagged photons compared to non-modulated background photons. In this section, we compare the use of various traditional optical detectors (APD and PMT) and a new type of SPAD detector using a TCSPC detection method and ultrasound microbubbles to enhance the local acousto-optic effect.

4.3.1. Detection of Ultrasound-modulated Light with an Avalanche Photodiode (APD)

A 65 mm wide tank, filled with 0.65% intralipid was illuminated using a laser diode (5 mW at 650 nm, coherence length ~5m). The intralipid provided a reduced optical scatter coefficient of 0.2 mm^{-1} which is 20% of an in vivo reduced scatter coefficient of 1 mm^{-1}. Light was detected in transmission mode with an amplified silicon photodetector (PDA36A, ThorLabs), amplified and filtered (Stanford Research Systems, SR560) to remove DC and low frequency noise (<10 KHz) and recorded on a digital scope (Tektronix, TDS 2012B). A function generator (Stanford Research Systems, DS345) and RF Amplifier (ENI, 240L) delivered to a single-element US transducer (Panametrics, V303, 3cm focal length) a 1 MHz sine wave (20 cycles at 50Vpp) every 50μs (2% duty cycle). The transducer was submerged in water and the light was passed through the US focal zone (Figure 4.2a). The local US effect on water was sufficient to modulate the optical properties that were detected in the optical signal and displayed on the scope (Figure 4.2b). A Fourier transform of this signal (Figure 4.2c) shows a peak at 1 MHz (the US frequency) where the amplitude is proportional to the magnitude of the AO effect. When, the measurement was repeated without US there was no modulation detected (not shown). Note that transmission through 65-mm of media that scatters light at the same order as biological tissues is similar to imaging backscattered photon from ~3cm depth as would be expected when imaging the breast.

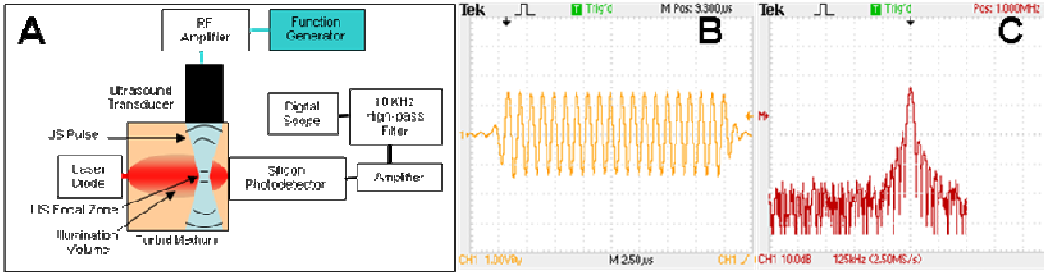


Figure 4.2 We used the experimental setup for AO measurements of US-modulated coherent light (a) to acquire the modulated optical signal detected and recorded on the digital scope (b). When the signal was Fourier transformed (c) peak frequency was the same as the US transmit frequency demonstrating US modulation of the optical signal.

4.3.2. Detection of Ultrasound-modulated Light with a Photomultiplier Tube (PMT)

Detection of AO modulated photons has traditionally been done with photomultiplier tubes due to their good quantum efficiencies and high gains. We use a similar setup shown in Figure 4.1a and replace the silicon photodetector with a Hamamatsu HC-125 Analog PMT with 8MHz bandwidth (Figure 4.3a). We use 532nm light and place an aperture in front of the PMT to reduce the total photon flux to prevent detector saturation. A clear ultrasound-modulated light signal is shown in the time-domain (Figure 4.3b) and the frequency domain (Figure 4.3c).

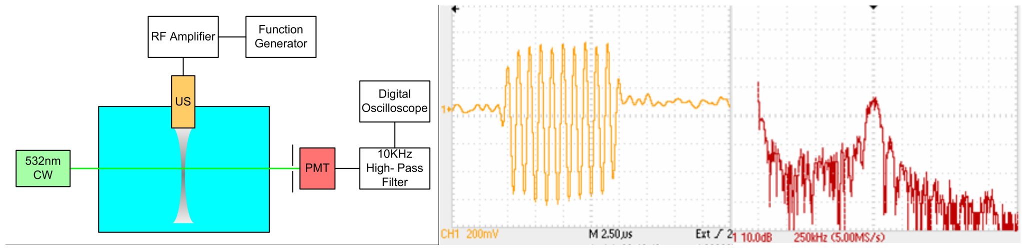


Figure 4.3 Setup for detection of ultrasound-modulated light with PMT (a), time-domain AO signal (b), and Fourier transform of signal (c)

Real-time AO imaging will inevitably require photodetector arrays. Large arrays of PMTs, however, are not practical due their inability to be cost-effectively scaled as well as their bulky and fragile nature. In the next section we present preliminary results demonstrating feasibility of AO imaging with a CMOS SPAD capable of being scaled towards megapixel arrays for a fraction of the cost of a single PMT detector.

4.3.3. Detection of Ultrasound-modulated Light with a CMOS SPAD

The number of modulated photons are few compared to the number of unmodulated photons. Time-correlated single photon counting (TCSPC) is a powerful tool to improve poor SNR. It operates on the principle of detecting few photons from a periodic signal over multiple sample periods to reconstruct the waveform from the individual time measurements. Utilization of a CMOS is desirable, particularly when multiple source-detector pairs are used when scalability is needed for real-time AO imaging with diagnostic ultrasound systems.

A proof-of-concept setup schematic is shown in Figure 4.4 and photo in Figure 4.5. A 532nm cw laser diode is used with <5mW laser power. A Panametrics 1MHz unfocused V303 single element ultrasound transducer is used. A water bath is used with scatterer solution (Intralipid) titrated in with a known optical reduced scattering coefficient. A reduced scattering coefficient of 1 mm^{-1} is targeted to simulate biological tissue. A transmission geometry is used with a collection lens and microscope objective to focus the light onto a 7μm STI-bound CMOS SPAD operated at 3% above its breakdown voltage. For low scatter coefficients of <0.15 mm^{-1}, an OD1 filter was used to attenuate the incident light so not to saturate the SPAD detector. The US transducer is pulsed at a 10kHz repetition rate and synced with the pulse counting electronics. The SPAD output is then input to the pulse counter board.

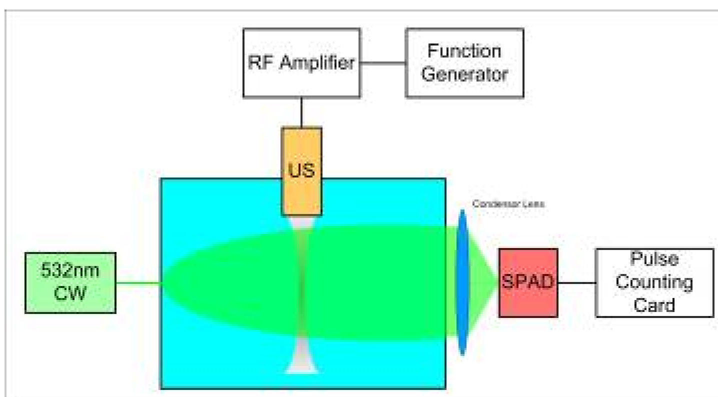


Figure 4.4 Schematic of AO-SPAD setup

Figure 4.5 Photo of AO-SPAD setup

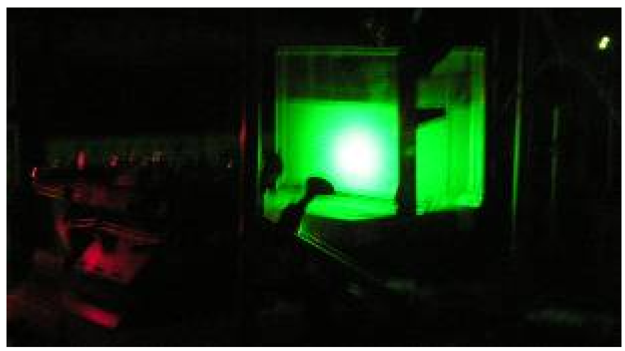


Figure 4.6 Illuminated Intralipid Bath

A modulated signal is detected with modulation depth defined as the signal amplitude divided by the background DC level (Figure 4.7). The background is an accumulation of non-modulated photons and the background dark count rate of the SPAD detector.

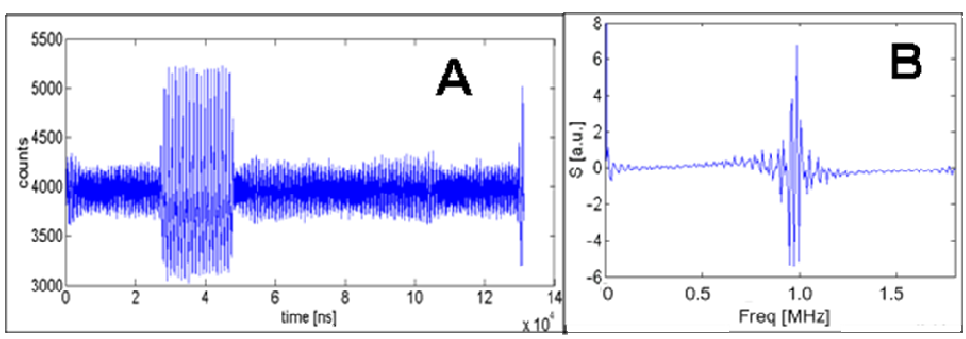


Figure 4.7 (a) photon counting histogram collected from SPAD mirrors the US modulation frequency; (b) FFT of histogram showing peak at 1MHz

The modulation depth is plotted vs. % of the target μ_s' of $1mm^{-1}$ in Figure 4.8. The upward slope of the blue curve at low concentrations indicates that focusing of the non-scattered laser light was not optimal as increasing the optical scattering most likely resulted in a greater number of photons impinging on the detector active area. As expected, at increased scatter coefficients, the modulation depth decreases as a result of the modulated photons being overwhelmed by the background unmodulated photons. As a proof-of-concept, we have shown the ability for a SPAD to be used as the optical detector in AO modulation. Through the use of TCSPC, an ultrasound modulated signal was recovered through a μ_s' of up to 0.2 mm^{-1}.

The 7-µm SPAD tuned towards ultraviolet was used for this experiment. Not only was the laser light at 532nm, but also the 5mW power was 10 times less than had been used in previous AO implementations. Despite these issues and the fact that the active detection area is only 7µm, we were able to observe the modulated optical signal. The larger surface area of a 2D array should capture many more photons increasing SNR. The AO effect can be increased by using higher ultrasound pressures, higher intensity laser excitation, better light collection, and averaging over longer periods.

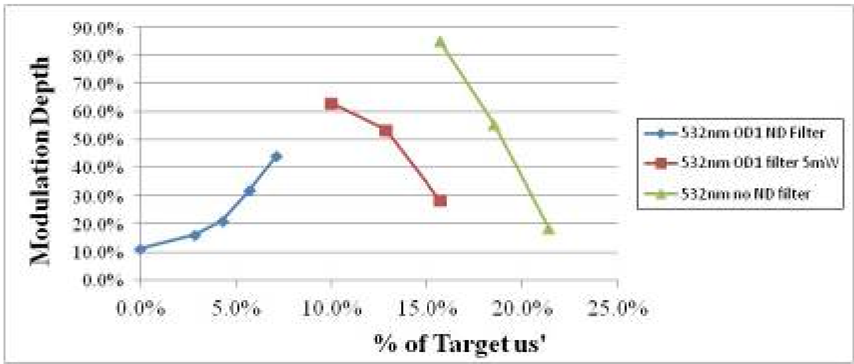


Figure 4.8 Modulation Depth vs. % Target μ_s'

4.4. Acousto-Optic Signal Enhancement with Ultrasound Contrast Microbubbles

One of the novel ideas proposed here, is the use of microbubbles to increase photon modulation to improve AO SNR. Microbubbles are common, FDA-approved diagnostic ultrasound contrast agents (UCAs), and have also been explored as externally-activated gene and drug delivery agents [172-174]. They are usually made up of a perfluorocarbon gas core with a lipid or albumin shell. Because the AO effect is in part dependent on local refractive index changes and in part dependent on the degree of oscillation of optical scatterers, we aimed to test the hypothesis that microbubbles will provide greater modulation and higher SNR. We hypothesize that the effect of optical attenuation that is dramatically increased in the presence of microbubbles.

We shine light (650 nm) through the ultrasound focal zone of an in vitro phantom and collect modulated photons with an avalanche photodetector (APD). We observe a significant improvement in SNR (>15 dB) with the addition of microbubbles in the focal zone by performing a Fast Fourier Transform of the APD signal in real-time. Not only do we demonstrate a SNR enhancement due to microbubble oscillation at the ultrasound fundamental frequency, but non-linear oscillations of the microbubbles modulate the photons at higher order harmonic frequencies too. The specific microbubble harmonic signature would allow greater differentiation of microbubbles from tissue and dramatically increase sensitivity, as is the case with harmonic ultrasound imaging. Furthermore, the use of FDA-approved microbubbles facilitates direct translation to the clinic where a significant increase in SNR would greatly enhance in vivo acousto-optic imaging.

We submerged a 2.25MHz US transducer (Panametrics, V306) in a water tank oriented perpendicular to the wall of a square 12.5x12.5x68mm quartz cuvette such that the US focal zone was within the cuvette (Figure 4.9). It was driven by a continuous sine wave at 1.5 MHz using a signal generator and power amplifier. We passed a 5mW 650nm CW laser light through the US focal zone perpendicular to the US wave that was collected by an avalanche photodetector (APD) with 10dB built-in gain. The optical signal was sent through a high-pass filter (>10KHz), amplified, converted into frequency domain, and displayed synchronized to an external trigger from the function generator.

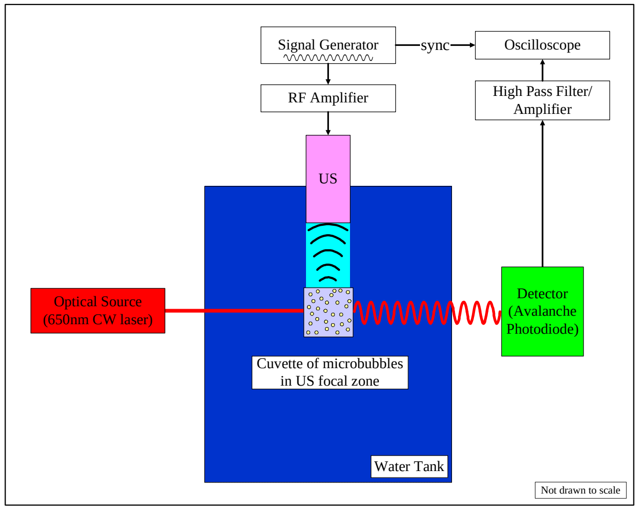


Figure 4.9 Schematic of AO enhancement with microbubbles experiment

We prepared Definity (FDA approved US contrast agent, Bristol-Myers Squibb Medical Imaging) per manufacturer's instruction. We filled a 1-mL syringe and placed it vertically with the opening pointing down for one hour to allow the microbubbles to settle. We then pushed the most dependent 0.1ml into 10ml of saline, discarded 0.4ml, pushed the next 0.1ml into another 10ml of saline, discarded 0.35ml and resuspended the remaining 0.05ml with 0.95 ml of saline, mixed the suspension and pushed 0.1ml into a 3rd 10ml of saline. This resulted in three suspensions of small (least buoyant), medium, and large (most buoyant) microbubble suspensions. The optical signal was recorded when the cuvette was filled with saline and then again with each of the 3 suspensions following thorough rinsing between exchanges.

Figure 4.10a shows AO modulation at the fundamental frequency in all suspensions; however, note that (1) there is an increase in the SNR at the fundamental frequency for all three microbubble suspensions; and (2) the appearance of ultraharmonics at the 3 and 4.5MHz for the medium and large microbubbles. Figure 4.11 shows the dB gain of the modulated photons shown in Figure 4.10 when microbubbles were added to saline. Note that the SNR increased by 16 and 19dB at the fundamental frequency, by ~13 dB at 3MHz, and again by 16 and 19dB at 4.5MHz for the medium and large microbubbles, respectively.

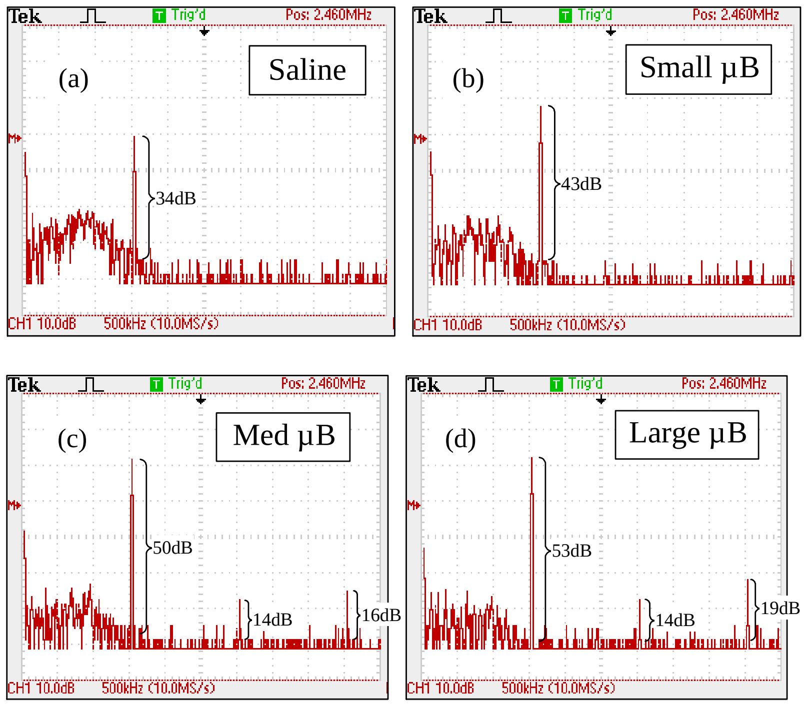


Figure 4.10 Fourier transform of APD output from (a) saline, (b) small, (c) medium, and (d) large microbubble suspensions

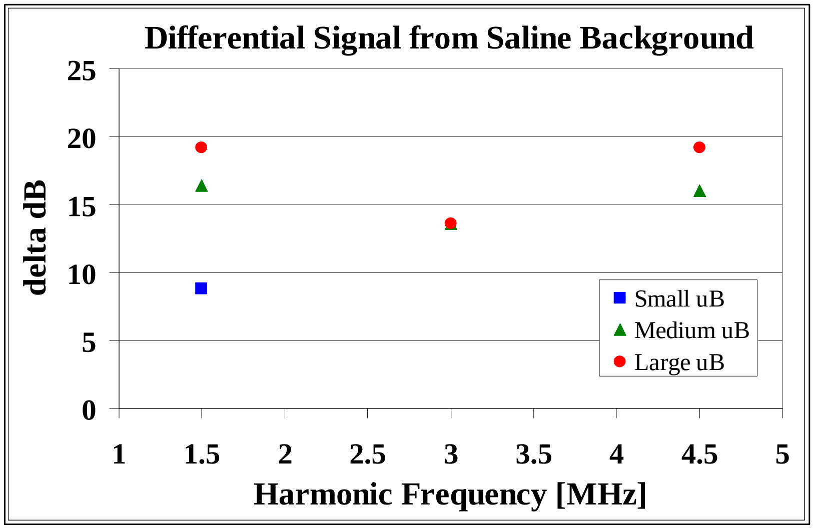


Figure 4.11 dB gain of modulated photon signal at the fundamental and harmonic frequencies of the microbubble suspensions relative to saline

In an effort to further understand the linear and non-linear behavior of ultrasound microbubbles, fundamental studies of microbubbles were done and are described in the next chapter using a light scattering tool with pulsed ultrasound.

4.5. Conclusion

In this chapter, we have compared the use of various optical detectors for detection of acousto-optic modulated light. Acousto-optic imaging suffers from poor signal-to-background ratios, resulting in the need for complex detection schemes utilizing lock-in amplifiers or optical interferometers. Detection in scattering media is difficult. We demonstrate the use of time-correlated single photon counting using a CMOS STI-bound SPAD for acousto-optic signal detection in a scattering media of 0.2 mm^{-1}, 20% of the target reduced scattering coefficient. Improved collection efficiencies, larger optical detector areas, higher laser powers, and higher ultrasound pressure intensities will aid the reconstruction of the optical waveform. Further, we demonstrated the use of FDA-approved ultrasound microbubbles as a method to increase the acousto-optic effect in the region of interest for enhancement of the fundamental driving and harmonic frequencies.

Acknowledgement:

This chapter, in part, is a reprint of the material as it appears in the following publications:

- M. J. Hsu, D. Hall, S. Farshchi-Heydari, S. Esener, R. Mattrey, "Novel Acousto-optic Signal Enhancement using Microbubbles," World Molecular Imaging Congress (New Optical Instrumentation), Nice, France 2008.

- D. J. Hall, M. J. Hsu, S. Esener, R. F. Mattrey, "Detection of Ultrasound-modulated Photons and Enhancement with Ultrasound Microbubbles," Photons Plus Ultrasound: Imaging and Sensing 2009, SPIE Vol. 7177, San Jose, CA 2009.

5. CHARACTERIZATION OF ULTRASOUND CONTRAST AGENTS WITH LIGHT SCATTERING

5.1. Introduction

It is necessary to know the optimal acoustic parameters of ultrasound contrast agents (UCAs) to observe specific contrast agents while reducing background noise. For therapeutic applications, knowing the specific acoustic properties of the UCA of interest will enable tuning the acoustic system for more precise and efficient drug and gene delivery while minimizing adverse effects to surrounding tissues. During insonation, microbubbles oscillate, emitting linear and non-linear backscatter at low to medium acoustic pressures, and cavitate at high pressures [175]. Non-linear echoes may be detected with high SNRs [176], providing excellent contrast for ultrasound imaging. Important parameters include acoustic scatter, stability, resonance frequency, and cavitation thresholds.

Common methods for UCA characterization include measurement of acoustic scatter and attenuation [177-180]. However, acoustic resolution limits the observation of individual bubble oscillations and bandwidth-limited acoustic receivers limit harmonic detection capabilities for free microbubbles, though acoustic properties of individually bound microbubbles have recently been explored [181]. Optical imaging techniques, such as high-speed cameras and streak cameras [182-186], have been developed to obtain visual information of oscillating microbubbles and produce high quality images for a direct observation of an individual microbubble’s response to ultrasound. However, because of the large amount of data processed per video, only a few cycles of ultrasound may be recorded at most. Furthermore, aforementioned high-speed cameras are expensive, which prohibits their widespread use for high throughput microbubble characterization.

A low-cost alternative to the aforementioned characterization techniques is the use of light scattering, which can provide an indirect view of the interaction of ultrasound with an individual free microbubble [187],[188]. Furthermore, due to the small amount of data per oscillation cycle, high throughput characterization of microbubbles is possible. The light scattering intensity of an individual microbubble is dependent on factors such as particle size, optical index of refraction of the core and shell, and shell thickness [189]. Mie scattering of coated spheres and optimal scattering angles for air-filled bubbles is described by Marston [190].

Marston et al. recommend an observation near the critical angle of 83^{o} for an air bubble in water due to a nearly R^2 dependence on the scattered light intensity as a function of radius. The scattering angle is also affected by the microbubble shell which induces a shift in the critical angle by approximately:

$$\Delta = \frac{2h}{a}\left[\left(n_w^2 - 1\right)^{-1/2} - \left(n_s^2 - 1\right)^{-1/2}\right] \text{ (radians)}$$

Equation 5.1 Shift in critical angle by microbubble shell

where the coating thickness-to-radius ratio is *h/a* and n_s and n_w are the respective optical indices of refraction for the shell and water. Ultrasound microbubbles typically have either a lipid or albumin shell with indices of refraction of 1.48 and 1.50, respectively. It has been shown that the shell thickness for a 1μm microbubble has a small but negligible effect on the scattering angle and is shown as the dotted curve in Figure 5.2. A microbubble's size changes with ultrasound pressure results in a traversal up and down the relative scatter intensity curve.

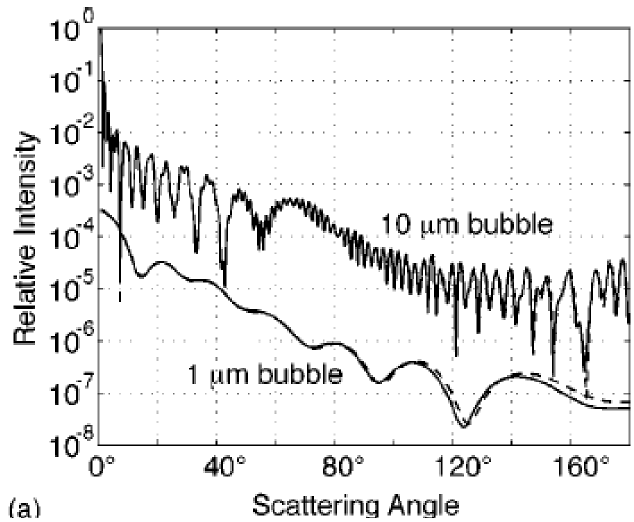


Figure 5.1 Relative light scatter intensity as a function of scattering angle [187]

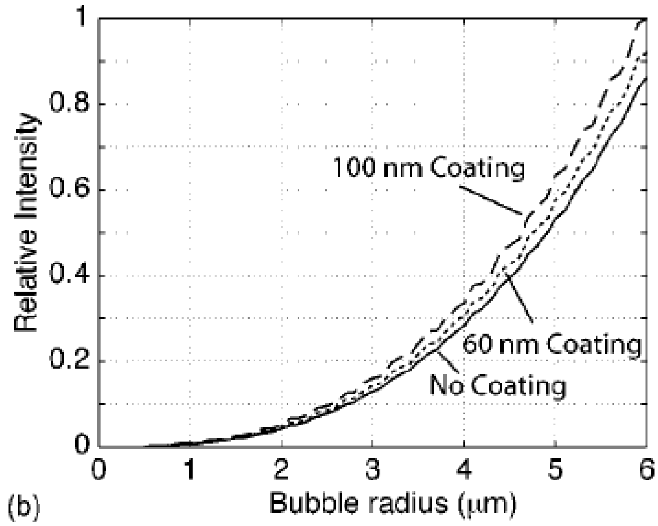


Figure 5.2 Relative light scatter intensity as a function of bubble radius at 80^o [187]

Microbubble dynamics models of coated microbubbles based on the Rayleigh-Plesset equations have been described [191-193]. Each model has its own merits. A model presented by Morgan et al. [193] uses the following to describe microbubble dynamics:

$$\rho R\ddot{R} + \frac{3}{2}\rho\dot{R}^2 = \left(P_0 + \frac{2\sigma}{R_0} + \frac{2\chi}{R_0}\right)\left(\frac{R_0}{R}\right)^{3\gamma}\left(1 - 3\frac{\dot{R}}{c}\right) - \frac{4\mu\dot{R}}{R} - \frac{2\sigma}{R}\left(1 - \frac{\dot{R}}{c}\right)$$

$$-\frac{2\chi}{R}\left(\frac{R_0}{R}\right)^2\left(1 - 3\frac{\dot{R}}{c}\right) - 12\varepsilon\mu_{sh}\frac{\dot{R}}{R(R-\varepsilon)} - \left(P_0 + P_{drive}(t)\right)$$

Equation 5.2 Morgan model for microbubble dynamics

where R is the radius of the microbubble, R_0 is the initial radius of the microbubble, P_0 is the ambient pressure, $P_{drive}(t)$ is the acoustic driving pressure, ρ is the liquid density, γ is the ratio of specific heats, c is the speed of sound, σ is the surface tension coefficient, χ is the shell elasticity, μ is the fluid shear viscosity, μ_{sh} is the microbubble shell shear viscosity, and ε is the shell thickness. Further, each microbubble has a characteristic resonance frequency determined by:

$$f_r = \frac{1}{2\pi}\sqrt{\frac{3\gamma}{\rho R_0^2}\left(P_0 + \frac{2\sigma}{R_0} + \frac{2\chi}{R_0}\right) - \frac{2\sigma + 6\chi}{\rho R_0^3} - \frac{\left(4\mu + 12\varepsilon\mu_{sh} / R_0\right)^2}{\rho^2 R_0^4}}$$

Equation 5.3 Microbubble Resonance Frequency

A pressure induced size change of the UCA will induce a change in the scattered light intensity which combines Mie theory with microbubble models [187]. Guan *et al*. have demonstrated the use of optical scattering to observe microbubble size changes in response to pulsed ultrasound [187] and fits them to models while Tu *et al*. compare various bubble dynamic models to estimate shell parameters [188]. Herein, we present an improvement on the optical scatter concept presented by Guan *et al*. to investigate individual microbubble dynamics in response to pulsed ultrasound in real-time with Fourier analysis of raw experimental data.

Using a stable laser source with increased intensity, a dark optical environment, and low-noise optical detector and electronics, we increase the signal-to-noise ratio (SNR) of the system by a factor of 10 at comparable ultrasound intensities from previously published work with SNR defined as the ratio of signal amplitude to the variance of the noise. With improved SNR, individual microbubble dynamics in response to a single ultrasound pulse train can be acquired with ease without the need of averaging, permitting an indirect observation of the individual microbubble's ability to expand and contract with consecutive ultrasound cycles without any latency between ultrasound pulses. In addition, using optical detector with 10MHz bandwidth, we are able to detect multiple higher order harmonic signatures simultaneously and carry out Fourier analysis of raw experimental data. We demonstrate the power of this tool by comparing the linear and non-linear behavior of different clinically relevant microbubble formulations at varying ultrasound intensities.

5.1.1. Materials and Method

5.1.1.1. Experimental Setup

We have improved on a light scattering technique to characterize dynamics of various perfluorocarbon gas-filled microbubbles. Commercial Definity (Lantheus Medical Imaging, N. Billerica, MA, USA) microbubbles with a phospholipid shell and perfluoropropane gas core, have been characterized using both acoustic and high-speed video techniques [194]. In-house prepared albumin-shelled microbubbles, also with a perfluorocarbon gas core, were formulated as comparison. Albumin-shelled microbubbles are stiffer [183] than phospholipid-shelled microbubbles, making them good models for comparison.

The experimental setup is shown in Figure 5.3. We used a two liter black water bath for a clean sound field with transparent optical windows for light delivery and collection. A Verdi-V5 continuous-wave 532nm laser source (Coherent Inc., Santa Clara, CA, USA) was focused with a 20x objective with a power density greater than 5μW/cm^2 at the laser focus. We used a 2.25MHz focused single-element Panametrics

V306 ultrasound transducer (Olympus NDT Inc., Waltham, MA, USA), with its focus confocal with the laser focus. The ultrasound was pulsed at a 1Hz repetition rate, with 15 cycles per pulse at 2.25MHz, which was within the resonance frequency range of Definity microbubbles in the 1.1-3.3μm size range [195],[196]. Transducer positioning and ultrasound pressure were calibrated with a HNC-0200 needle hydrophone (Onda Corp., Sunnyvale, CA, USA).

Scattered light was collected at 75^o to 95^o, which includes the critical angle of 83^o for air bubbles in water [197], focused and collimated onto a low dark current R6095 photomultiplier tube (PMT) optical detector (Hamamatsu Corp., Bridgewater, NJ, USA) with a 10MHz bandwidth. The PMT signal was then passed through a low noise M7279 amplifier (Hamamatsu Corp., Bridgewater, NJ, USA), held at constant gain, to a high-speed digital TDS2020 sampling oscilloscope (Tektronix Inc., Beaverton, OR, USA) which was controlled using Labview software on a personal computer. Ultrasound contrast agents were delivered using a PHD 3000 syringe pump (Harvard Apparatus, Holliston, MA, USA) and a capillary tube placed in-line with the laser optical axis and outside the ultrasound interaction region. The ultrasound contrast agents were diluted such that at most one microbubble is present in the laser focus at a time, which was confirmed using a standard video-rate 1322 CCD camera (Cohu Inc., San Diego, CA, USA).

The ultrasound signal was synchronized with the oscilloscope, which continuously monitored both the DC scatter intensity and AC modulation. For a modulated light scattering event to occur, a microbubble must flow to the laser focus to create the scattering event, interact with the pulsed ultrasound, and change its size to produce a change in the light scattering intensity. When these conditions were met, the waveform was saved and converted from PMT voltage to relative intensity and then to size.

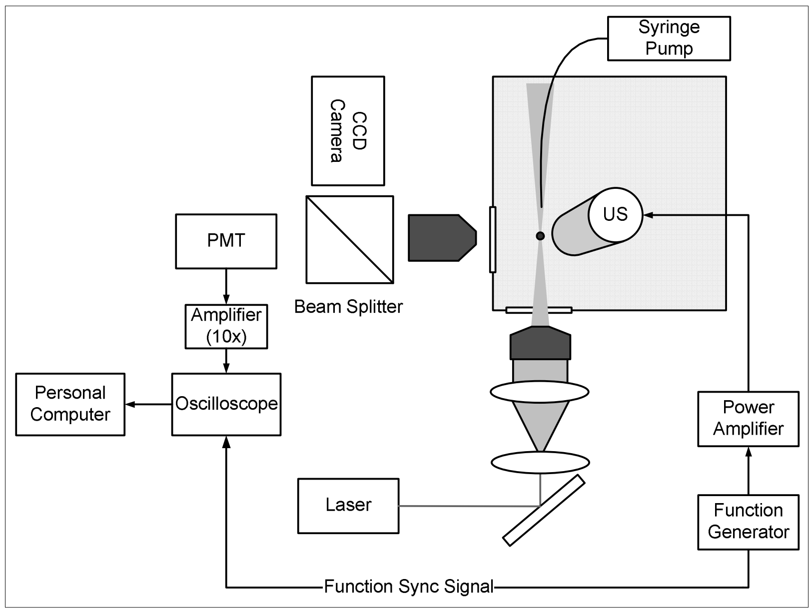


Figure 5.3 Schematic of experimental setup.

5.1.1.2. Microbubble Preparation

Commercial Definity microbubbles were diluted to a concentration on the order of 10^6 microbubbles/mL with a mean diameter of 1.1-3.3 microns. To produce a more monodispersed sample, the microbubbles were fractionated via floatation to separate the stock solution into different size distributions [198].

As a model for rigid shelled bubbles, bovine serum albumin (BSA) microbubbles were prepared. Formulation was based on published techniques using probe sonication [199],[200]. BSA is a 67 kDa protein containing 35 thiol groups, 17 of which are paired as disulfide bonds. Probe sonication causes the disulfide bonds to split apart and reform with other BSA molecules. This process produces microbubbles stabilized by a layer of denatured protein that can be as thick as 200 nm. Briefly, a 2.5 wt% solution of BSA in phosphate buffered saline was probe sonicated continuously for 2 min in a centrifuge tube under a headspace of perfluorobutane. The bubbles were then centrifuged at 400 *g* for 4 min to concentrate the bubbles at the top of the tube and the liquid was removed via pipette. Fresh PBS was added and the mixture was agitated to resuspend the bubbles. The bubbles were washed five times in this manner to ensure a clean suspension. The separated microbubble population was determined via optical microscopy to have a similar size distribution to that of the Definity sample.

Solid 1.8 μm monodispersed silica microspheres (Bangs Laboratories Inc., Fishers, IN, USA) were used as controls because of their inability to change size at diagnostic ultrasound pressures.

5.1.2. Results

With a 10 fold increase in SNR of the detection system, we used light scattering to compare microbubble dynamics of commercial Definity phospholipid microbubbles to in-house prepared albumin-shelled microbubbles. When there was no particle present, as expected, there is no optical signal leading to a quiet background.

At 100kPa, both the Definity and albumin microbubbles exhibited a linear response. Figure 5.4 (a) shows the microbubble size change of an individual Definity microbubble. An initial radius of 2μm is used based on average size. The approximately uniform modulation amplitude indicates that the phospholipid-shelled microbubbles are able to sustain linear oscillations over these 15 cycles without bubble destruction or significant gas leakage. Figure 5.4 (b) shows the root-mean-squared (rms) Fourier transform of the relative intensity signal from the solid silica microspheres, albumin-shelled, and Definity microbubbles and shows the presence of the fundamental driving frequency. Due to the greater shell stiffness of the albumin microbubbles, the microbubble oscillation is dampened and results in a smaller relative intensity. As expected, silica microspheres do not change size and exhibit no light scattering oscillation.

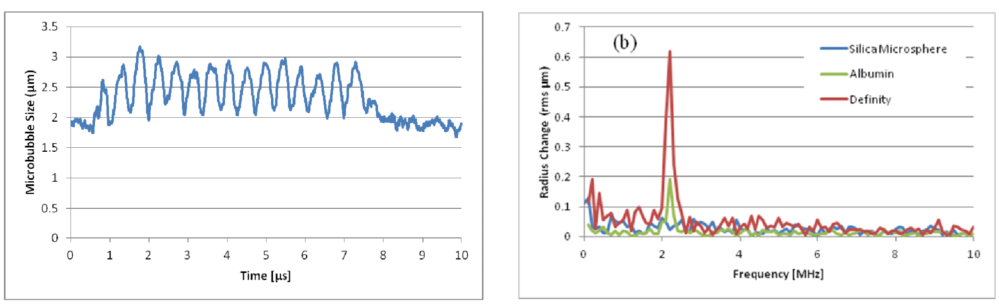


Figure 5.4 Time domain waveform of symmetric size change of Definity microbubble at 100kPa (a). Corresponding frequency response at 100kPa of silica microsphere, albumin microbubble, and Definity microbubbles (b).

At an increased pressure of 200kPa, a non-linear scatter intensity change was observed, which can be attributed to the asymmetric expansion and collapse of the microbubble gas core [175]. Figure 5.5 shows the asymmetry in the time-domain waveform (a) and the second, third, and fourth harmonics in the Fourier domain (b). Due to the greater shell stiffness of the albumin bubbles, the gas core was not compressed and expanded enough to have a strong asymmetric response. As expected, solid silica microspheres again exhibited zero oscillation at this pressure.

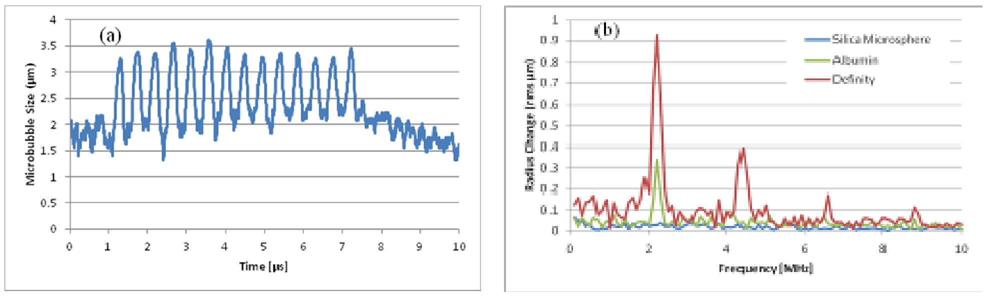


Figure 5.5. Time domain waveform of asymmetric size change of Definity microbubble at 200kPa (a). Corresponding frequency response at 200kPa of silica microsphere, albumin microbubble, and Definity microbubbles with the presence of higher harmonics (b).

Microbubble dynamics were subsequently explored. Small amounts of perfluorocarbon gas escape via diffusion during each shell expansion and contraction due to the stretching and breaking of lipid shell [201]. Under increased non-destructive pressures, the shell will be compromised and gas is released at a faster rate. The loss of gas resulted in an increasingly non-acoustically responsive microbubble, reducing its ability to expand and contract. Figure 5.6 shows the size change of a microbubble undergoing increased gas release at 260kPa.

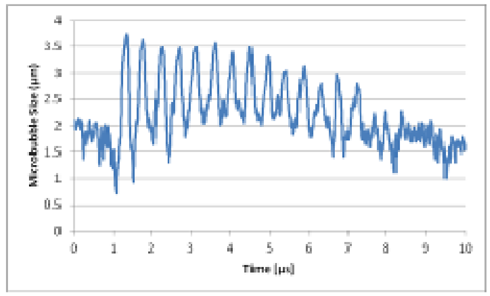
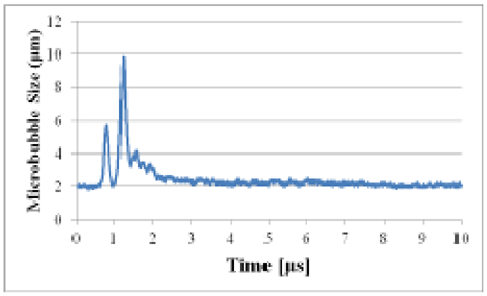


Figure 5.6. Time-domain waveform of size change showing dynamics of Definity microbubble with sucessive ultrasound cycles at 260kPa.

At high pressures, one can demonstrate instantaneous microbubble destruction or cavitation. Characterization of microbubble destruction is of particular interest to determine cavitation thresholds. Microbubble cavitation is marked by one or two large oscillations followed by immediate microbubble destruction or collapse. The disappearance of the optical signal indicates destruction in the phospholipid shell and consequent release of the encapsulated perfluorocarbon gas due to the loss of the gas-water index mismatch. With light scattering, we get an instantaneous, albeit indirect, view of the microbubble destruction. Figure 5.7 demonstrates the instantaneous destruction of the Definity microbubble at 1.3MPa in the time domain (a) and its corresponding frequency response (b). The initial microbubble oscillations and ensuing implosion are captured in a single signal sweep. The broadened frequency dsitribution is due to the few number of oscillation cycles and broadband frequency spread associated with microbubble cavitation.



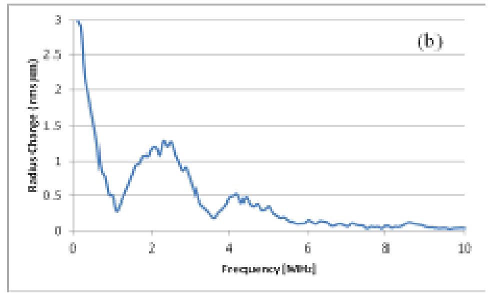
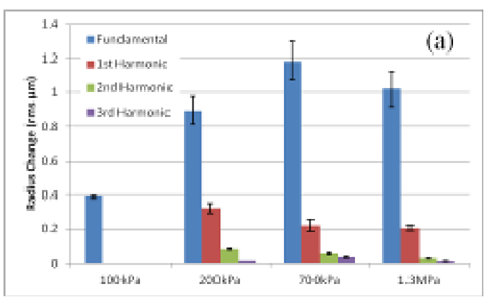


Figure 5.7 Time-domain (a) and corresponding Fourier response (b) waveform of size change of Definity microbubble cavitation at 1.3MPa.

Figure 5.8 summarizes the average acoustic response and peak harmonic signatures of Definity (a) and albumin-shelled (b) microbubbles at 2.25MHz with 15 cycles per pulse for multiple microbubbles. At 1.3MPa, instant microbubble cavitation was observed resulting in spreading of the frequency spectrum and a reduction in the peak amplitude at the fundamental and harmonic frequencies. Stiffer albumin-shelled microbubbles were not observed to have cavitation under these conditions, with their energy distribution remaining at the fundamental and harmonic frequencies.



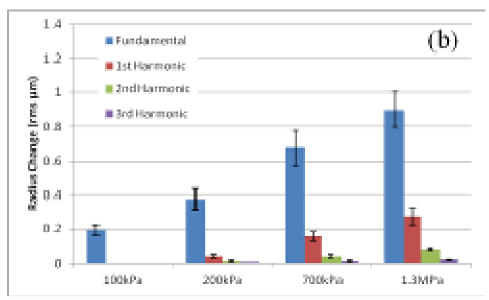


Figure 5.8 Harmonic frequency peak intensity vs acoustic pressure for Definity (a) and albumin-shelled (b) microbubbles.

We have demonstrated the strength of the system with increased signal-to-noise ratios by characterizing microbubble formulations at varying ultrasound intensities. Due to the low background noise of the system, high SNR is attained allowing for single sweep acquisition of microbubble dynamics in real-time. All results were shown to be repeatable over thousands of microbubbles and ultrasound cycles within the expected microbubble size spread. Due to the sensitivity and accuracy of this light scattering tool for ultrasound contrast agent characterization, it will be used for further characterization of custom-formulated microbubbles and ultrasound sensitive particles.

5.1.3. Liposome with Encapsulated Microbubble Characterization

Characterization of microbubbles encapsulated in liposomes (“mothership”) has become of particular interest in our group for novel ultrasound activated drug delivery platforms. Knowledge of their specific acoustic properties is needed to understand the optimal parameters for payload delivery. Below is a time-domain image of a mothership when insonated with 15 pulses of 2.25 MHz (Figure 5.9). The time-domain waveforms show that the microbubbles can sustain oscillations through the 15 cycles of ultrasound. The outer liposome shell acts as a thick shell around the microbubble, dampening the oscillations, resulting in the ability to sustain multiple cycles of ultrasound without cavitation. Controls for free microbubbles and free liposomes were carried out demonstrating distinctly different acoustic responses with cavitation at high pressures for microbubbles akin to observations shown in section 5.1.2, and no response with liposomes due to the absence of an air-to-water interface. Further studies are needed to understand the complete dynamics of mothership behavior in response to different ultrasound frequencies and pressures. This light scattering system should provide the means to do further studies.

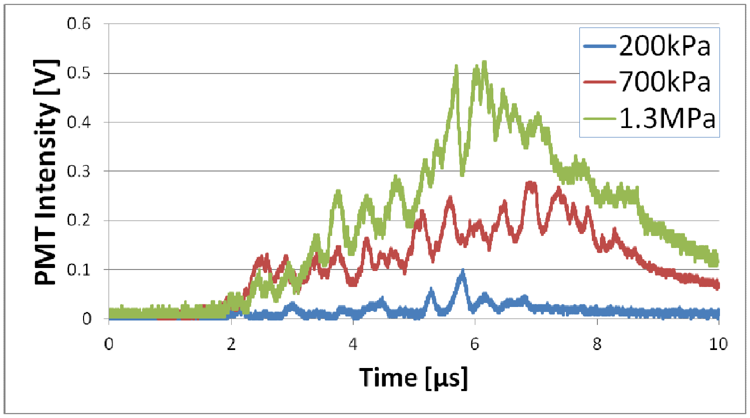


Figure 5.9 Light scatter Characterization of Microbubble encapsulated in Liposome

5.2. Fluorescence Modulation with Dye-loaded Microbubbles

Acousto-optic imaging has gained interest in recent years as a method to increase spatial resolutions of optical imaging. Two out of the three modulation mechanisms outlined in section 4.2 require the use of coherent light. Further, complex detection schemes are needed because of poor signal-to-background ratios. Recently, there has been movement towards modulating fluorescence. Acousto-optic modulation of fluorescent light is of particular interest to further localize the optical signal at the region of interest. By doing this, the overwhelming background of the emission light can be filtered out optically, leaving just the fluorophore emission light for increased signal-to-background ratios. In this section, mechanisms of fluorescence modulation will be described. A novel use of dye-loaded microbubbles is demonstrated as a method to achieve modulation of fluorescence light. The use of a dye-loaded microbubble stands to improve acousto-optic imaging techniques. First, the microbubble itself can be imaged with pulsed ultrasound to know its exact location. Second, local optical properties around the microbubble can be better extracted with better knowledge of the types of tissue the photons must scatter through to reach the optical detector. Further, modulation of fluorescence with dye loaded microbubbles provides insight in lipid behavior during expansion and contraction cycles.

5.2.1. Theory of Fluorescence Modulation in Solution

Fluorescence acousto-optic imaging was first demonstrated by Kobayashi et al. where they present detection of fluorescent microspheres embedded in scattering tissue phantoms [202]. The physical mechanism of fluorescence modulation is not clear, particularly when it is claimed that it is due to the modulation of refractive index, which would indicate a coherent effect, while fluorescence emission is by nature incoherent. Hall et al. also demonstrated detection of fluorescence modulation using a quadrature method [203].

The physical mechanisms of ultrasound modulation of fluorescence in turbid media is described by Yuan et al. [204] and are attributed to a modulation of the local fluorophore concentration when concentrations are low. At high concentrations, a modulation of a self-quenching effect is considered which leads to a nonlinear intensity modulation. Further, Yuan et al. consider fluorescence resonance energy transfer as another possible mechanism for fluorescence modulation. Interestingly, they attribute the effect seen by Kobayashi et al. as a simple wavelength change, where the source photons are modulated before the fluorescent target and then emitted at the fluorescence wavelength. Figure 5.10 shows a schematic of this concept where S and D are the optical source and detector, respectively. The solid curves show the optical path of the excitation light that pass through the ultrasound focal zone while the dotted curves indicate the optical path of the fluorescent light. This type of modulation is akin to a local speckle modulation at the fluorescent target region.

The modulation of fluorescence can be attributed to the following mechanisms and analytical expressions are derived based on Yuan [204]:

1) Modulation of the local fluorophore concentration in a low concentration and a high concentration (self-quenching) regime.
2) Modulation of excitation light before the fluorophore.
3) Modulation of fluorescence emission.

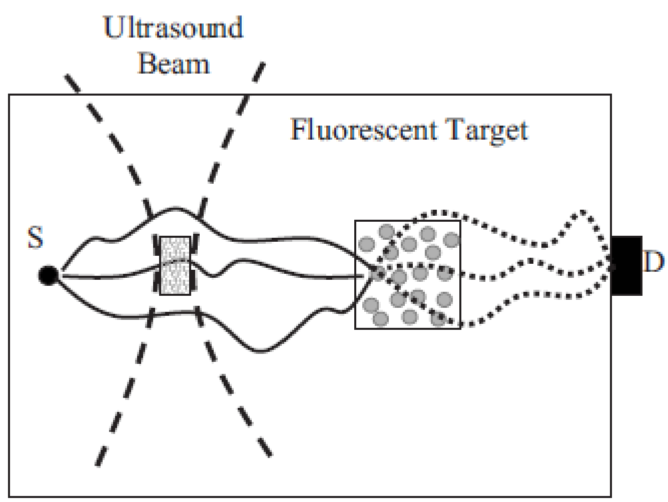


Figure 5.10 Schematic of experimental setup similar to Kobayashi [204]

Local changes in fluorophore concentration at the ultrasound focal zone can be expressed as:

$$N(r,t) = N_0 + N_1(r)\exp(-i\omega_s t) + N_1^*(r)\exp(i\omega_s t)$$

Equation 5.4

where *N(r,t)* is the fluorophore concentration at position *r* and time *t*, N_0 is the average fluorophore concentration in the ultrasound focal zone. $N_1(r)$ is the complex amplitude of the fluorophore concentration at position *r*:

$$N_1(r) = K' N_0 P_1 \exp(ikr)$$

Equation 5.5

where $K' = \dfrac{k}{\rho_0 c(k\upsilon_0 - \omega_s)} \cdot \dfrac{1+\tilde{\varepsilon} - i\left[\tilde{\varepsilon} + \frac{2}{3}\tilde{\varepsilon}^2\right]}{1+\tilde{\varepsilon} - i\left[\tilde{\varepsilon} + \frac{4}{9}\left(\rho + \frac{1}{2}\right)\tilde{\varepsilon}^2\right]} + \beta$, $\tilde{\varepsilon} = \sqrt{\dfrac{\omega_s a^2}{2\gamma}}$, *a* is the radius of the microsphere, γ is the fluid kinematic viscosity, and $\rho = \rho_P / \rho_0$ is the ratio between the microsphere and the average density of the surrounding fluid.

Modulation of excitation light is considered another possible mechanism for fluorescence modulation. An excitation fluence rate can be expressed as:

$$U^{ex}(r_s,r,t) = U_0^{ex}(r_s,r) + U_1^{ex}(r_s,r)\exp(-i\omega_l t) + U_1^{ex*}(r_s,r)\exp(-i\omega_l t)$$

Equation 5.6 Excitation fluence rate

where ω_l is the modulation angular frequency of the excitation light. $U_1^{ex}(r_s,r,t)$ is the excitation photon fluence rate at position *r* at time *t* with a point source at position

r_s. $U_0^{ex}(r_s, r)$ is the average fluence rate of the excitation light at position r, which is considered to be the DC component of the light. $U_1^{ex}(r_s, r)$ is the AC component of the excitation light with a complex conjugate $U_1^{ex*}(r_s, r)$. According to diffusion theory [205],

$$U_0^{ex}(r_s, r) = \frac{S_{DC}}{D_{ex}} G_0^{ex}(r_s, r)$$

$$U_1^{ex}(r_s, r) = \frac{S_{AC}}{D_{ex}} G_1^{ex}(r_s, r)$$

Equation 5.7

where $D_{ex} = \frac{1}{3\mu'_{s\_ex}}$ is the diffusion coefficient at the excitation wavelength and μ'_{s_ex} is the reduced scattering coefficient. S_{DC} and S_{AC} are the respective dc and ac components of the excitation source. G_0^{ex} and G_1^{ex} are the Green's functions for the dc and ac components, respectively.

The emission strength by fluorophores at position r and time t can be expressed by [206]

$$S^{fl}(r_s, r, t) = \phi \Gamma N_e(r, t)$$

Equation 5.8

where □ is the fluorophore quantum efficiency, Γ is the total decay rate from the excited state to the ground state through radiative and non-radiative decay. $N_e(r,t)$ is the total number of fluorophores at position r and time t:

$$\frac{\partial N_e(r,t)}{\partial t} = -\Gamma N_e(r,t) + \varepsilon U^{ex}(r_s, r, t)\left[N(r,t) - N_e(r,t)\right]$$

Equation 5.9

The propagation of the emission light from position *r* to the detector results in a photon fluence rate of:

$$U^{fl}(r_s, r_d, \omega_i, t) = \int_{\Omega} \frac{S^{fl}(r_s, r_d, \omega_i, t)}{D_{fl}} G^{fl}(r, r_d, \omega_i) dr^3$$

Equation 5.10

where $D_{fl} = \dfrac{1}{3\mu'_{s\_fl}}$ and G^{fl} is the Green's function describing the propagation of fluorescence light through the media.

5.2.2. Fluorescence Modulation with Dye-loaded Microbubbles

Modulation of fluorescence has been demonstrated by Kobayashi et al. and has been followed by Yuan et al., albeit with the use of a lock-in amplifier [207]. Further, the use of microbubbles to enhance fluorescence modulation has been developed experimentally by Yuan [208], albeit with questions regarding the physical phenomena that they have observed. In addition, dye-loaded microbubbles have been analyzed theoretically by Yuan [209] with FRET pairs.

Fluorescence concentration quenching is a well known effect [210-212] and has been used in a number of applications including studies of protease and nuclease activities [213], membrane fusion [214], protein-dimer formation [215], and single molecule protein-folding [216]. Lipophilic fluorescent dyes can be attached to the shell of phospholipid microbubbles. As described in Section 5.1, microbubbles expand and contract in response to insonation. Under insonation, dye-loaded microbubbles' surface concentration can change by an order of magnitude or greater, resulting in a dramatic change in the effective surface dye concentration. With the right surface concentrations, microbubble expansions and contractions can generate periods of reduced and increased quenching efficiency, respectively, resulting in an intensity modulation of the fluorescence intensity (Figure 5.11). Yuan postulates that a 100% modulation efficiency may be achieved [209]. It is important to have a very homogeneous loading of the dye on the microbubble surface so that a maximum concentration differential occurs with each size oscillation cycle. The depth of AO imaging with microbubbles will be limited to the quenching efficiency, intensity of the fluorescent signal as well as the light collection and sensitivity of the optical detector. Time-correlated single photon counting will be needed to extract the low signal levels, particularly after tissue scattering of the optical signal has taken place. In the next section, we will describe our breakthroughs in fabrication of fluorescently-loaded microbubbles and detection of a modulated fluorescent signal.

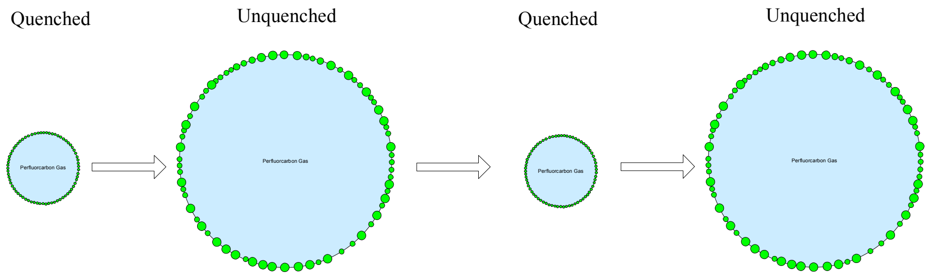


Figure 5.11 Quenched to unquenched cycle resulting from microbubble expansion

5.2.2.1. Fluorescent Microbubble Preparation

Microbubble preparation is outlined in this section. Stock solutions of the lipids and the dye were prepared by dissolving the dry powder in chloroform. 1 mg of DSPC and 1 mg of mPEG-DSPE were combined in a 1.5 mL microcentrifuge tube. An appropriate amount of DiI was added to the same tube. The solution was vortexed at low speed while applying an argon stream to evaporate all of the chloroform and create a lipid film.

500uL of DPBS was added to the lipid film, and the tube was vortexed at high speed for 1 minute. The tube was immediately placed into a heating block which was at 70°C. After 1 hour, the tube was removed and placed directly on ice for 5 minutes. A 3 mL syringe was filled with 100 µL perfluorohexane. A 22 gauge needle was attached and bent at a 60° angle to allow the syringe to be pumped without ejecting liquid. By pumping the syringe, the headspace of the tube was filled with perfluorohexane gas.

The XL2000 probe sonicator (QSonica LLC., Newtown, CT) was operated at the liquid/gas interface for 3 seconds at level 20 to produce microbubbles. The bubbles were then put back on ice. To purify the bubbles and wash away the excess lipid and dye, the bubbles were centrifuged at 1000 rpm for 3 minutes. The subnatant was removed with a syringe (approximately 450 μL), and the bubbles were resuspended with 500 μL DPBS. This was repeated once more.

5.2.2.2. Fluorescence Detection Setup

An optical scattering system similar to the one described in Section 5.1.1.1 is used and shown in Figure 5.12. There are several differences in the setups, namely the use of two filters (a 546nm long pass and 570nm band pass) for a high rejection ratio of the excitation light. Because the filters are interference filters, normal incidence of photons is necessary for high rejection ratios. Therefore, two long black optical barrels are used to ensure only straight photons can pass through the filters and reach the optical detector. Again, the data is recorded with Labview through a Tektronix TDS3032 sampling oscilloscope.

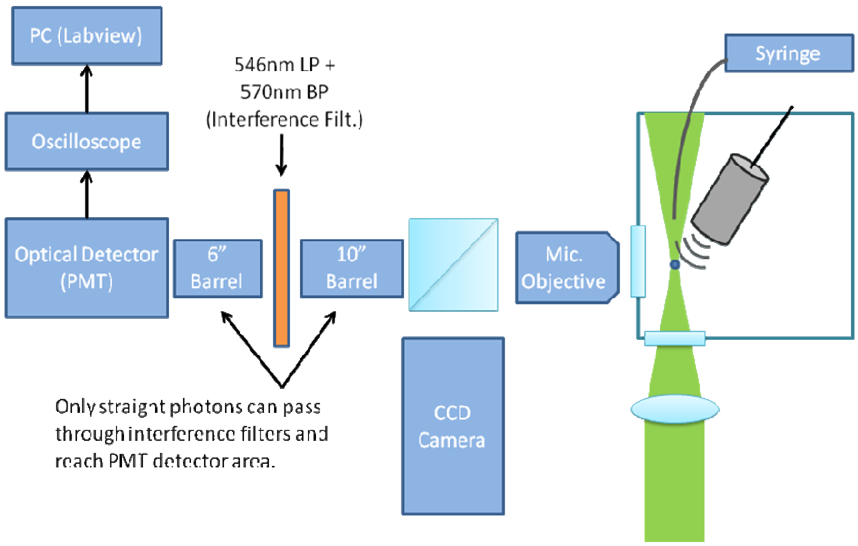


Figure 5.12 Schematic for Fluorescence Modulation Detection

Initial preliminary studies were done with a single PMT detector, with and without the fluorescent filters. Representative waveforms are shown with their corresponding Fourier transforms. Figure 5.13 shows time domain studies and corresponding Fourier transforms of the microbubble sample without fluorescent filters demonstrating the ability for the microbubble sample to oscillate easily, with the presence of a weak harmonic in some cases. Figure 5.14 shows time domain studies and Fourier transforms of the same microbubble sample with the fluorescent filters placed in the optical beam path with an increased PMT gain with clear oscillations are present, albeit much weaker. The variance in time-domain waveform intensities is attributed to a polydispersed microbubble sample and perhaps uneven loading of the fluorescent dye on the lipid shell. To test whether we were detecting leakage light, a control sample of non-fluorescently loaded microbubbles was run through the system demonstrating microbubble oscillations without the wavelength filter and no detected oscillations with the wavelength filter placed in line.

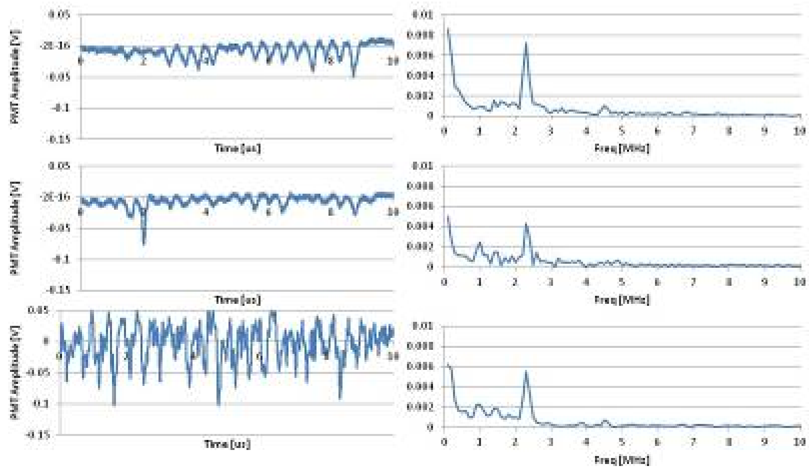


Figure 5.13 0.48 mol % DiI without wavelength filters

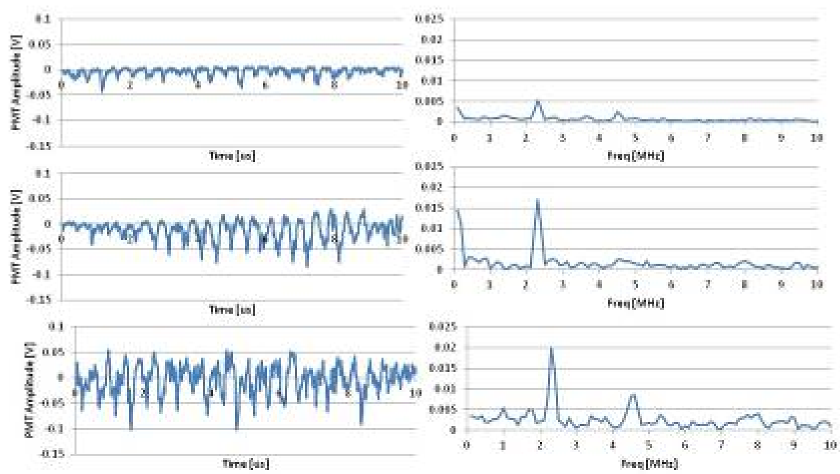


Figure 5.14 0.48 mol % DiI with wavelength filters

After demonstrating that fluorescence modulation is possible with dye-loaded microbubbles, further studies were carried out with an improved setup (Figure 5.15). Two PMTs are used to simultaneous monitor both the fluorescence signal and optical scatter of excitation light to correlate microbubble size changes to fluorescence modulation intensities. Both optical signals are recorded simultaneously with Labview.

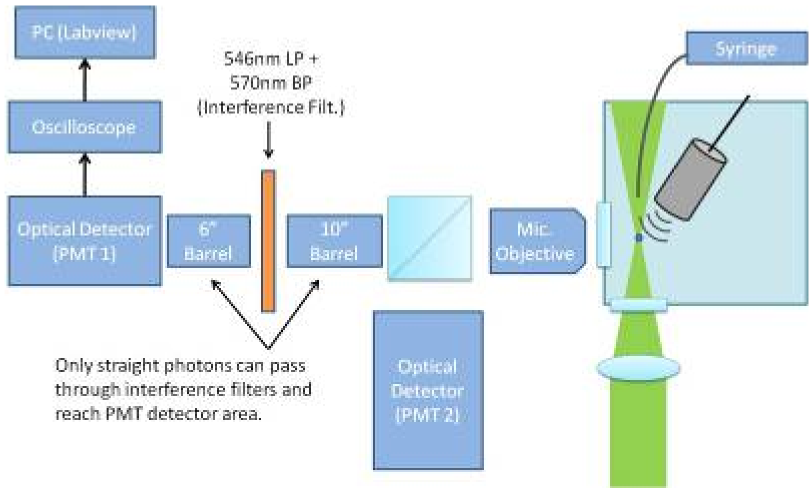


Figure 5.15 Dual Detector Fluorescence Modulation Detection Setup

To test the concentration quenching effect of dye-loaded microbubbles, DiI was titrated to create microbubble samples with six different surface concentrations ranging from 0.25 mol% to 8 mol%. Static fluorescent microscope images are shown in Figure 5.16 under identical imaging conditions (same acquisition times and gain settings). It is noted that the images shown here are intensity images and are not normalized with the total amount of surface-loaded fluorophore. However, the 8 mol% image is clearly dimmer than the 4mol%. The normalized intensity curve shown in Figure 5.17 shows efficient quenching at high concentrations indicating a concentration quenching mechanism on the microbubble surface, with slight variation at low concentrations, perhaps due to errors in dye uptake on the microbubble shell. The amount of fluorescence modulation will depend on how much the quenching efficiency can be modulated which will be a function of initial concentration and microbubble size change. A microbubble size change of 3x results in a surface concentration change of almost an order of magnitude. It is hypothesized that the greatest contrast in quenching efficiency will be with microbubbles with surface concentrations around 1mol% to 4mol%, where the slope of the concentration quenching curve is the steepest.

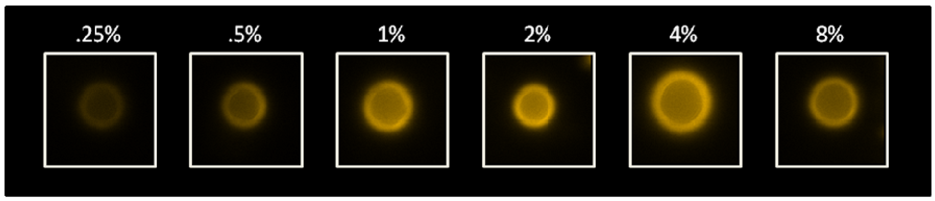


Figure 5.16 Fluorescent microscope images of DiI-loaded Microbubbles

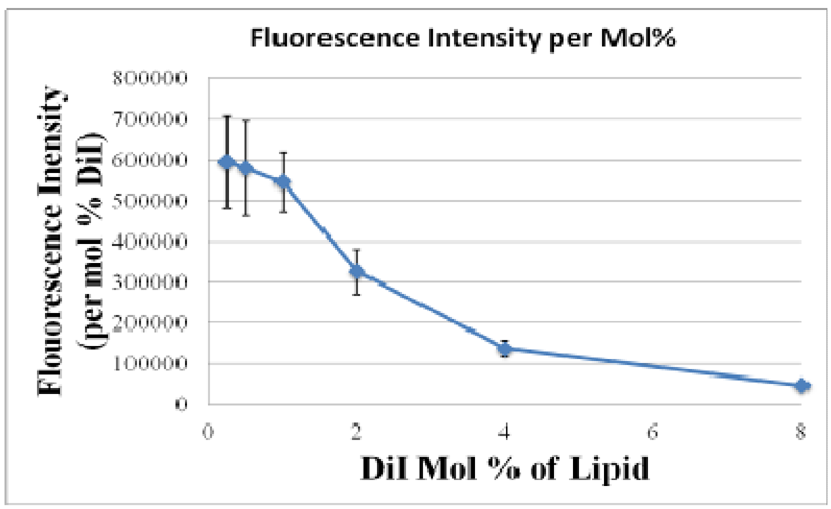


Figure 5.17 Self-quenching curve of DiI-loaded microbubbles

Each of the six DiI-loaded microbubble samples were flowed through the optical scatter system with identical measurement conditions (fixed PMT gain and laser power) and measured at ultrasound pressure intensities of 260kPa, 700kPa, and

1.3MPa at 2.25MHz with 25 cycles. Figure 5.18 is a representative plot using the dual detector scheme with the fluorescent modulation in blue and the scattered excitation light intensity modulation in blue. Fourier transforms of the many fluorescent modulation signals are acquired, averaged, and plotted as a function of concentration and ultrasound pressure in Figure 5.19. For all three ultrasound pressures, the greatest fluorescence intensity modulation occurred with microbubbles with a surface concentration of 2 mol%, which is in agreement with our measured concentration quenching curve (Figure 5.17). Further, it is interesting to note the large difference from 2 mol% to 4 mol%, indicating high sensitivity to loaded dye concentrations. Controls of non-DiI-loaded microbubbles were flowed through the system with no fluorescent response. Further, non-DiI-loaded microbubbles in a fluorescent solution with concentrations comparable to loaded-microbubbles were flowed through the system, again showing no fluorescence modulation signal. High ultrasound intensities are used in this experiment without resulting in microbubble cavitation. This is attributed to the polydispersed nature of the microbubble sample, and most likely the insonation frequency is off the resonance frequency of the microbubbles. Further, a high concentration of microbubble was used, perhaps resulting in multiple microbubbles being delivered into the ultrasound-light interaction area at a time.

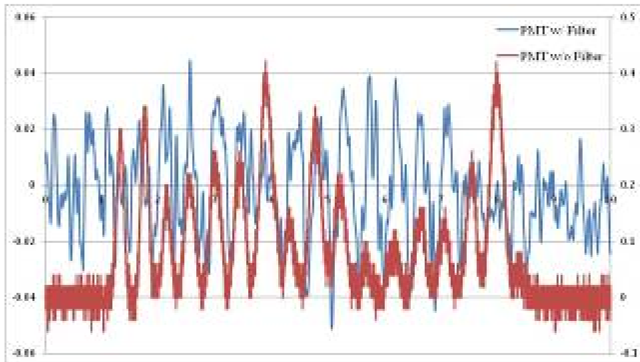


Figure 5.18 Dual detector plot

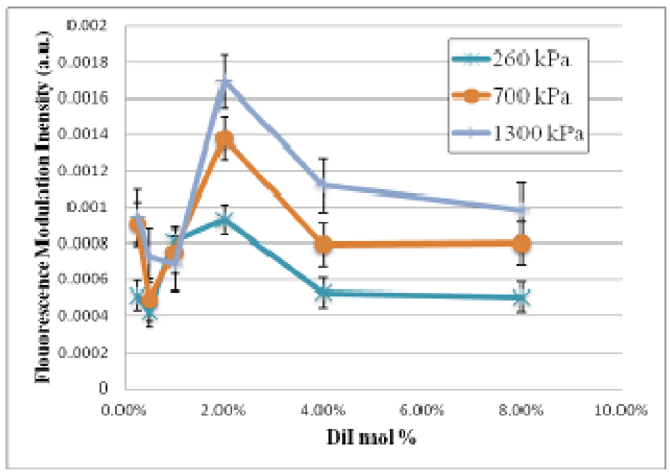


Figure 5.19 Fluorescence Modulation Intensity vs. DiI mol%

We have demonstrated the use of dye-loaded microbubbles as a method for modulating incoherent fluorescence light. Dye-loaded microbubbles can provide a local optical beacon when used in conjunction with pulsed ultrasound, with the potential for use in a number of in-vivo applications. The presence of harmonic signatures in the fluorescence intensity modulation can also help enhance weak acousto-optic signals. Preliminary studies indicate the effect of surface concentration quenching of DiI on phospholipid microbubbles. Enhanced fluorescence modulation should be possible with more efficient self-quenching dyes or perhaps a FRET pair to utilize a $\frac{1}{r^6}$ relation. Further studies are required to fully understand the physics of concentration quenching on the microbubble surface. In addition, the use of brighter near-IR emitting fluorescent dyes stands to benefit in vivo applications. Light collection in turbid media of the weak fluorescence signal will be challenging. However, since the majority of the emitted photons from the microbubble surface are being modulated with a dark optical background, signal-to-background ratios are increased compared to previous AO methods. Further studies are necessary, particularly in scattering media.

5.3. Conclusion

We report an optical tool for characterization of ultrasound contrast microbubble dynamics in real-time. We acquired relative light scatter intensity oscillations and converted them to microbubble size changes in response to pulsed ultrasound insonation. Due to high signal-to-noise ratios, with at least a 10 fold increase over previously published work, individual microbubble dynamics were acquired at various ultrasound intensities corresponding to linear and non-linear oscillation and microbubble cavitation regimes. In addition, higher-order harmonic signatures are acquired simultaneously using a single optical detector. Individual commercial Definity microbubbles were compared to in-house formulated albumin-stabilized microbubbles to demonstrate the system’s sensitivity and accuracy over thousands of microbubbles and ultrasound cycles. Fundamental frequency and harmonic signatures were observed from Fourier analysis of raw non-averaged experimental data. Phospholipid Definity microbubbles were characterized showing linear oscillations at 100kPa, non-linear oscillations at 200kPa and greater, and microbubble cavitation at 1.3MPa.

Damped oscillations were observed with albumin-shelled microbubbles and as expected, no oscillations were observed with solid silica microspheres. Further, sustained acoustic response from microbubbles encapsulated in liposomes was observed. With high signal-to-noise ratios, light scattering combined with pulsed ultrasound allows for eased characterization and design of custom formulated microbubbles and ultrasound contrast agents in the future.

Further, fluorescence modulation by ultrasound is demonstrated using dye-loaded microbubbles for the first time as a potential method to improve signal-to-background ratios for acousto-optic imaging. DiI-loaded microbubbles are fabricated with various surface concentrations and measured with fluorescence microscopy. The aforementioned light scattering system is modified for fluorescence detection. Fluorescent microbubbles are flowed through the scattering system with acoustic insonation. A modulated fluorescence signal is detected for each dye surface concentration and plotted as at different ultrasound intensities.

Acknowledgement:

This chapter, in part, is a reprint of the material as it appears in the following publications:

- M. J. Hsu, M. Benchimol, C. Schutt, D. J. Hall, R. F. Mattrey, S. Esener, "Fluorescence Modulation of DiI-loaded Microbubbles," in preparation for the Journal of Biomedical Optics.

- M. J. Hsu, M. Eghtedari, A. P. Goodwin, D. J. Hall, R. F. Mattrey, S. Esener, "Characterization of Individual Ultrasound Microbubble Dynamics with Novel Light Scattering System," in preparation for the Journal of the Acoustical Society of America.

6. CONCLUSION

6.1. Dissertation Summary and Original Contributions

This dissertation presented the development of a single photon avalanche diode (SPAD) fabricated in a standard 0.18μm deep-submicron CMOS process. A new efficient guard ring structure is modeled, simulated, characterized and is compared to a diffused-guard ring structure to demonstrate a new efficient junction planarizing method. We demonstrated many benefits of this new structure, from fast sampling speeds, low jitter, and high fill factors, which stand to benefit applications where fast and precise timing is important as outlined in Chapter 1 such as fluorescence lifetime imaging and time-correlated single photon counting.

After a complete device characterization and comparison to prior art, we utilize the SPAD in various applications. We demonstrated proof-of-concept of a free-space single photon frequency up-conversion scheme of 1550nm using a hybrid device where secondary hot-carrier electroluminescence, typically an adverse byproduct of avalanche diodes, is exploited. Further, acousto-optic imaging is explored with various traditional low-light photodetectors, including the CMOS SPAD with time-correlated single photon counting. Acousto-optic signal enhancement is demonstrated with the use of ultrasound contrast microbubbles with enhancement at both fundamental and harmonic frequencies. Furthermore, a light scattering system is developed for characterization of microbubbles and various other acoustically responsive particles. Finally, we introduced dye-loaded microbubbles and demonstrated a method for modulation of fluorescence emission by ultrasound.

The original contributions of this work are summarized below:

- Development of electrostatic and electrical simulations of a SPAD device in a deep-submicron technology.
- Development of a complete characterization environment of a single-photon avalanche diode for detection efficiency, spectral response, dead-time, cross-talk, and jitter.
- Comparison of the STI guard ring to the traditional diffused guard-ring structure.
- Proof-of-concept of a free-space single photon frequency up-conversion scheme based on hot-carrier electroluminescence.
- Acousto-optic signal detection with various optical detectors.
- Acousto-optic signal enhancement with ultrasound microbubbles with higher-order harmonics.
- Design and development of an optical scatter system to characterize acoustic responsive particles—commercial FDA-approved microbubbles, in-house fabricated microbubbles, and microbubble encapsulated liposomes.
- Experimental measurement of fluorescence modulation with dye-loaded microbubbles.

6.2. Outlook and Future Research

The outlook for single photon detectors improvements is application driven. While it is difficult to optimize all performance characteristics, certain key metrics of interest can be optimized in a single device. SPADs can be designed to be optimal for certain target applications. Reducing noise characteristics is very desirable in most applications, but can be made up for using a time-correlated single photon counting technique. Large active areas are desirable for systems where optical alignment is difficult. Low jitter is important in applications where precise timing resolution is necessary, for example in pulse-position modulation optical links, time-of-flight applications, three-dimensional imaging in LIDAR, biometrics, and fluorescence lifetime imaging (FLIM). Scaling to arrays is desirable for fast imaging applications which could revolutionize FLIM imaging that is currently limited by slow acquisition rates.

There has already been movement based on our work with the STI-bound SPAD for new device architectures using a STI- guard ring with diffused passivation [85]. New system architectures for fast pulse processing, particularly with large arrays can be improved for faster sampling rates and shortened image acquisition times. Further, new processing circuitry can be developed for optimization of dark count rates and reduction of jitter in the sensing amplifiers.

Hybrid architectures for frequency up-conversion of 1550nm light have many areas for improvement. Faster, more efficient IR-sensitive SPADs can continue to be developed. Detection of IR/NIR light is needed for classical optical links as well as in vivo optical imaging. Further, new device coupling schemes can be designed for optimization of up-conversion efficiency.

Acousto-optic imaging suffers from poor signal-to-noise ratios (SNR). Further developments in improving SNR can be realized using various methods outlined in Chapter 4. Signal localization with fluorescence microbubbles is promising for increasing signal-to-background ratios. Despite the few number of collected photons, further improvements can be made with the photon collection and also with the fluorescence quenching efficiency with improved design of the dye-loading, or perhaps going with a FRET pair.

SPAD development, hybrid devices for frequency up-conversion, and acousto-optic signal enhancement are all promising areas with each of their own technological challenges. With the groundwork set, further research in these areas can be continued to realize their potential.

REFERENCES

[1] V. Ntziachristos, A.G. Yodh, M. Schnall, and B. Chance, “Concurrent MRI and diffuse optical tomography of breast after indocyanine green enhancement,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 97, 2000, pp. 2767-2772.

[2] R. Choe, A. Corlu, K. Lee, T. Durduran, S.D. Konecky, M. Grosicka-Koptyra, S.R. Arridge, B.J. Czerniecki, D.L. Fraker, A. DeMichele, B. Chance, M.A. Rosen, and A.G. Yodh, “Diffuse optical tomography of breast cancer during neoadjuvant chemotherapy: A case study with comparison to MRI,” *Medical Physics*, vol. 32, 2005, pp. 1128-1139.

[3] W. Mo and N. Chen, “Fast time-domain diffuse optical tomography using pseudorandom bit sequences,” *Optics Express*, vol. 16, 2008, pp. 13643-13650.

[4] B. Pogue, M. Testorf, T. McBride, U. Osterberg, and K. Paulsen, “Instrumentation and design of a frequency-domain diffuse optical tomography imager for breast cancer detection,” *Optics Express*, vol. 1, 1997, pp. 391-403.

[5] A.B. Konovalov, V.V. Vlasov, A.G. Kalintsev, O.V. Kravtsenyuk, and V.V. Lyubimov, “Time-domain diffuse optical tomography using analytic statistical characteristics of photon trajectories,” *Quantum Electronics*, vol. 36, 2006, pp. 1048-1055.

[6] A.H. Hielscher, A.Y. Bluestone, G.S. Abdoulaev, A.D. Klose, J. Lasker, M. Stewart, U. Netz, and J. Beuthan, “Near-infrared diffuse optical tomography,” *Disease Markers*, vol. 18, 2002, pp. 313-337.

[7] S. Lam, F. Lesage, and X. Intes, “Time Domain Fluorescent Diffuse Optical Tomography: analytical expressions,” *Optics Express*, vol. 13, 2005, pp. 2263-2275.

[8] V.C. Spanoudaki, A.B. Mann, A.N. Otte, I. Konorov, I. Torres-Espallardo, S. Paul, and S.I. Ziegler, “Use of single photon counting detector arrays in combined PET/MR: Characterization of LYSO-SiPM detector modules and comparison with a LSO-APD detector,” *Journal of Instrumentation*, vol. 2, 2007, pp. P12002-P12002.

[9] T. Oida, Y. Sako, and A. Kusumi, “Fluorescence lifetime imaging microscopy (flimscopy). Methodology development and application to studies of endosome fusion in single cells,” *Biophysical Journal*, vol. 64, 1993, pp. 676-685.

[10] P.I. Bastiaens and A. Squire, “Fluorescence lifetime imaging microscopy: spatial resolution of biochemical processes in the cell,” *Trends in Cell Biology*, vol. 9, 1999, pp. 48-52.

[11] R. Yasuda, “Imaging spatiotemporal dynamics of neuronal signaling using fluorescence resonance energy transfer and fluorescence lifetime imaging microscopy,” *Current Opinion in Neurobiology*, vol. 16, 2006, pp. 551-561.

[12] P.J. Verveer, F.S. Wouters, A.R. Reynolds, and P.I. Bastiaens, “Quantitative imaging of lateral ErbB1 receptor signal propagation in the plasma membrane,” *Science (New York, N.Y.)*, vol. 290, 2000, pp. 1567-1570.

[13] F.S. Wouters and P.I. Bastiaens, “Fluorescence lifetime imaging of receptor tyrosine kinase activity in cells,” *Current Biology: CB*, vol. 9, 1999, pp. 1127-1130.

[14] C. Niclass and E. Charbon, “A single photon detector array with 64x64 resolution and millimetric depth accuracy for 3D imaging,” 2005.

[15] A. Yu, “A spaceborne lidar for high-resolution topographic mapping of the Earth's surface,” *SPIE Newsroom*, 2010.

[16] M. Pfennigbauer and A. Ullrich, “Applicability of single photon detection for laser radar,” *e & i Elektrotechnik und Informationstechnik*, vol. 124, 2007, pp. 180-185.

[17] I. Prochazka, K. Hamal, and B. Sopko, “Recent achievements in single photon detectors and their applications,” *Journal of Modern Optics*, vol. 51, 2004, p. 1289.

[18] M.A. Albota and E. Dauler, “Single photon detection of degenerate photon pairs at 1.55 μm from a periodically poled lithium niobate parametric downconverter,” *Journal of Modern Optics*, vol. 51, 2004, p. 1417.

[19] M.A. Albota and F. Wong, “Efficient single-photon counting at 1.55 mm by means of frequency upconversion,” *Optics Letters*, vol. 29, 2004, pp. 1449-1451.

[20] M.A. Albota, B. Robinson, D. Caplan, S. Hamilton, and D. Boroson, “Photon-counting 1.55 μm optical communications with pulse-position modulation and a multimode upconversion single-photon receiver,” 2008.

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

[22] A. Beveratos, R. Brouri, T. Gacoin, A. Villing, J. Poizat, and P. Grangier, “Single Photon Quantum Cryptography,” *Physical Review Letters*, vol. 89, 2002, p. 187901.

[23] V. Kurochkin, A. Zverev, Y. Kurochkin, I. Ryabtsev, and I. Neizvestny, “Using single-photon detectors for quantum key distribution in an experimental fiber-optic communication system,” *Optoelectronics, Instrumentation and Data Processing*, vol. 45, 2009, pp. 374-381.

[24] E. Diamanti, H. Takesue, T. Honjo, K. Inoue, and Y. Yamamoto, “Performance of various quantum-key-distribution systems using 1.55-mum up-conversion single-photon detectors,” *Physical Review A*, vol. 72, 2005, p. 052311.

[25] S. Guha, J.L. Habif, and M. Takeoka, “PPM demodulation: On approaching fundamental limits of optical communications,” *1001.2447*, 2010.

[26] D.M. Chudakov, S. Lukyanov, and K.A. Lukyanov, “Fluorescent proteins as a toolkit for in vivo imaging,” *Trends in Biotechnology*, vol. 23, 2005, pp. 605-613.

[27] B.J. Baker, H. Mutoh, D. Dimitrov, W. Akemann, A. Perron, Y. Iwamoto, L. Jin, L.B. Cohen, E.Y. Isacoff, V.A. Pieribone, T. Hughes, and T. Knöpfel, “Genetically encoded fluorescent sensors of membrane potential,” *Brain cell biology*, vol. 36, 2008, pp. 53-67.

[28] R. Yuste, “Fluorescence microscopy today,” *Nat Meth*, vol. 2, 2005, pp. 902-904.

[29] J. Livet, T.A. Weissman, H. Kang, R.W. Draft, J. Lu, R.A. Bennis, J.R. Sanes, and J.W. Lichtman, “Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system,” *Nature*, vol. 450, Nov. 2007, pp. 56-62.

[30] M. Gunkel, F. Erdel, K. Rippe, P. Lemmer, R. Kaufmann, C. Hörmann, R. Amberger, and C. Cremer, “Dual color localization microscopy of cellular nanostructures,” *Biotechnology Journal*, vol. 4, 2009, pp. 927-938.

[31] L. Li and L.M. Davis, “Single photon avalanche diode for single molecule detection,” *Review of Scientific Instruments*, vol. 64, 1993, p. 1524.

[32] R.K.P. Benninger, W.J. Ashby, E.A. Ring, and D.W. Piston, “Single-photon-counting detector for increased sensitivity in two-photon laser scanning microscopy,” *Optics Letters*, vol. 33, 2008, pp. 2895-2897.

[33] Becker and Hickl, “What is Time Correlated Single Photon Counting,” 2002.

[34] W. Becker, “Lifetime Imaging Techniques for Optical Microscopy,” 2003.

[35] Arion F. Chatziioannou, “Molecular imaging of small animals with dedicated PET tomographs,” *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 29, 2002, pp. 98-114.

[36] T. Beyer, D.W. Townsend, T. Brun, P.E. Kinahan, M. Charron, R. Roddy, J. Jerin, J. Young, L. Byars, and R. Nutt, “A Combined PET/CT Scanner for Clinical Oncology,” *J Nucl Med*, vol. 41, 2000, pp. 1369-1379.

[37] P.E. Kinahan, D.W. Townsend, T. Beyer, and D. Sashin, “Attenuation correction for a combined 3D PET/CT scanner,” *Medical Physics*, vol. 25, 1998, pp. 2046-2053.

[38] J.S. Newman, I.R. Francis, M.S. Kaminski, and R.L. Wahl, “Imaging of lymphoma with PET with 2-[F-18]-fluoro-2-deoxy-D-glucose: correlation with CT.,” *Radiology*, vol. 190, 1994, pp. 111-116.

[39] D.W. Townsend, J.P. Carney, J.T. Yap, and N.C. Hall, “PET/CT Today and Tomorrow,” *J Nucl Med*, vol. 45, 2004, pp. 4S-14.

[40] “PET-detectorsystem 2.png,” *http://en.wikipedia.org/wiki/File:PET-detectorsystem_2.png*.

[41] B.S. Robinson, A.J. Kerman, E.A. Dauler, R.J. Barron, D.O. Caplan, M.L. Stevens, J.J. Carney, S.A. Hamilton, J.K. Yang, and K.K. Berggren, “781 Mbit/s photon-counting optical communications using a superconducting nanowire detector,” *Optics Letters*, vol. 31, 2006, pp. 444-446.

[42] K. Quirk and J. Gin, “Optical PPM Combining Loss for Photon Counting Receivers,” *MILCOM*, Los Alamitos, CA, USA: IEEE Computer Society, 2006, pp. 1-4.

[43] R.G. Lipes, “Pulse-Position-Modulation Coding as Near-Optimum Utilization of Photon Counting Channel with Bandwidth and Power Constraints,” *Deep Space Network Progress Report*, vol. 56, 1980, pp. 108-113.

[44] "Optical Modulation Techniques," *http://www.l-3com.com/tw/tutorial/modulation.html*.

[45] T. Yan and C. Chen, "Design and development of a baseline deep-space optical PPM transceiver," *Free-Space Laser Communication Technologies XI*, G.S. Mecherle, Ed., San Jose, CA, USA: SPIE, 1999, pp. 154-169.

[46] R.A. Cryan and R. Unwin, "Optical Space Communications Employing Pulse Position Modulation," *Advanced Modulation and Coding Techniques for Satellite Communications, IEEE Colloquium on*, 1992.

[47] C. Bennet and G. Brassard, "Quantum Cryptography: Public Key Distribution and Coin Tossing," *International Conference on Computers, Systems and Signal Processing*, Bangalore, India: 1984.

[48] H. Takesue, S.W. Nam, Q. Zhang, R.H. Hadfield, T. Honjo, K. Tamaki, and Y. Yamamoto, "Quantum key distribution over a 40-dB channel loss using superconducting single-photon detectors," *Nat Photon*, vol. 1, 2007, pp. 343-348.

[49] C. Gobby, Z.L. Yuan, and A.J. Shields, "Quantum key distribution over 122 km of standard telecom fiber," *Applied Physics Letters*, vol. 84, 2004, p. 3762.

[50] Hamamtsu, "Photomultiplier Tubes: Basics and Applications."

[51] Andor, "ICCD Detectors," 2003.

[52] P. Buzhan, A. Ilyin, V. Kantserov, V. Kaplin, A. Karakash, A. Pleshko, E. Popova, S. Smirnov, and Y. Volkov, "AN ADVANCED STUDY OF SILICON PHOTOMULTIPLIER," *ICFA Instrumentation Bulletin*.

[53] P. Eraerds, M. Legre, A. Rochas, H. Zbinden, and N. Gisin, "SiPM for fast Photon-Counting and Multiphoton Detection," *Optics Express*, vol. 15, 2007, pp. 14539-14549.

[54] S. España, L. Fraile, J. Herraiz, J. Udías, M. Desco, and J. Vaquero, "Performance evaluation of SiPM photodetectors for PET imaging in the presence of magnetic fields," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 613, 2010, pp. 308-316.

[55] D.R. Schaart, H.T. van Dam, S. Seifert, R. Vinke, P. Dendooven, H. Löhner, and F.J. Beekman, "A novel, SiPM-array-based, monolithic scintillator detector for PET," *Physics in Medicine and Biology*, vol. 54, 2009, pp. 3501-3512.

[56] A. Del Guerra, N. Belcari, M. Giuseppina Bisogni, G. LLosa, S. Marcatili, G. Ambrosi, F. Corsi, C. Marzocca, G. Dalla Betta, and C. Piemonte, “Advantages and pitfalls of the silicon photomultiplier (SiPM) as photodetector for the next generation of PET scanners,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 617, 2010, pp. 223-226.

[57] S. Miki, M. Fujiwara, M. Sasaki, B. Baek, A.J. Miller, R.H. Hadfield, S.W. Nam, and Z. Wang, “Large sensitive-area NbN nanowire superconducting single-photon detectors fabricated on single-crystal MgO substrates,” *Applied Physics Letters*, vol. 92, 2008, p. 061116.

[58] V. Anant, A.J. Kerman, E.A. Dauler, J.K.W. Yang, K.M. Rosfjord, and K.K. Berggren, “Optical properties of superconducting nanowire single-photon detectors,” *Optics Express*, vol. 16, 2008, pp. 10750-10761.

[59] S. Miki, M. Takeda, M. Fujiwara, M. Sasaki, and Z. Wang, “Compactly packaged superconducting nanowire single-photon detector with an optical cavity for multichannel system,” *Optics Express*, vol. 17, 2009, pp. 23557-23564.

[60] Hamamtsu, “Chacteristics and use of Si APD (Avalanche Photodiode).”

[61] A. Goetzberger, B. McDonald, R.H. Haitz, and R.M. Scarlett, “Avalanche Effects in Silicon p-n Junctions. II. Structurally Perfect Junctions,” *Journal of Applied Physics*, vol. 34, 1963, p. 1591.

[62] R.H. Haitz, “Mechanisms Contributing to the Noise Pulse Rate of Avalanche Diodes,” *Journal of Applied Physics*, vol. 36, 1965, p. 3123.

[63] R.H. Haitz, “Model for the Electrical Behavior of a Microplasma,” *Journal of Applied Physics*, vol. 35, 1964, p. 1370.

[64] R. Haitz and A. Goetzberger, “Avalanche noise study in microplasmas and uniform junctions,” *Solid-State Electronics*, vol. 6, 1963, pp. 678-680.

[65] P. Webb and R. McIntyre, “Single photon detection with avalanche photodiodes,” *Amer. Phys. Soc. Bull*, 1970, p. 813.

[66] H. Dautet, P. Deschamps, B. Dion, A.D. MacGregor, D. MacSween, R.J. McIntyre, C. Trottier, and P.P. Webb, “Photon counting techniques with silicon avalanche photodiodes,” *Applied Optics*, vol. 32, 1993, pp. 3894-3900.

[67] F. Zappa, S. Tisa, S. Cova, P. Maccagnani, D.B. Calia, G. Bonanno, M. Belluso, R. Saletti, and R. Roncella, "Pushing technologies: single-photon avalanche diode arrays," *Advancements in Adaptive Optics*, D.B. Calia, B.L. Ellerbroek, and R. Ragazzoni, Eds., Glasgow, Scotland, United Kingdom: SPIE, 2004, pp. 1200-1210.

[68] B.F. Aull, A.H. Loomis, D.J. Young, R.M. Heinrichs, B.J. Felton, P.J. Daniels, and D.J. L, "Geiger-mode avalanche photodiodes for three-dimensional imaging," *LINCOLN LABORATORY JOURNAL*, 2002, pp. 335--350.

[69] A. Lacaita, M. Ghioni, and S. Cova, "Double epitaxy improves single-photon avalanche diode performance," *Electronics Letters*, vol. 25, 1989, pp. 841-843.

[70] D.M. Taylor, J.C. Jackson, A.P. Morrison, A. Mathewson, and J.G. Rarity, "Characterization of novel active area silicon avalanche photodiodes operating in the geiger mode," *Journal of Modern Optics*, vol. 51, 2004, p. 1323.

[71] R. McIntyre, "The distribution of gains in uniformly multiplying avalanche photodiodes: Theory," *IEEE Transactions on Electron Devices*, vol. 19, 1972, pp. 703-713.

[72] A. Rochas, "Single Photon Avalanche Diodes in CMOS Technology," EPFL, 2003.

[73] A. Rochas, A. Pauchard, P. Besse, D. Pantic, Z. Prijic, and R. Popovic, "Low-noise silicon avalanche photodiodes fabricated in conventional CMOS technologies," *IEEE Transactions on Electron Devices*, vol. 49, 2002, pp. 387-394.

[74] C. Niclass, M. Sergio, and E. Charbon, "A single photon avalanche diode array fabricated in deep-submicron CMOS technology," *Proceedings of the conference on Design, automation and test in Europe: Proceedings*, Munich, Germany: European Design and Automation Association, 2006, pp. 81-86.

[75] C. Niclass, A. Rochas, P. Besse, R. Popovic, and E. Charbon, "A 4 [mu]s integration time imager based on CMOS single photon avalanche diode technology," *Sensors and Actuators A: Physical*, vol. 130-131, 2006, pp. 273-281.

[76] C. Niclass, A. Rochas, P. Besse, and E. Charbon, "Toward a 3-D camera based on single photon avalanche diodes," *Selected Topics in Quantum Electronics, IEEE Journal of*, vol. 10, 2004, pp. 796 - 802.

[77] S. Tisa, A. Tosi, and F. Zappa, "Fully-integrated CMOS single photon counter," *Optics Express*, vol. 15, 2007, pp. 2873-2887.

[78] C. Stapels, W. Lawrence, F. Augustine, and J. Chistian, "Characterization of a CMOS Geiger photodiode pixel," *Electron Devices, IEEE Transactions on*, vol. 53, 2006, pp. 631 - 635.

[79] H. Finkelstein, M. Hsu, and S. Esener, "STI-Bounded Single-Photon Avalanche Diode in a Deep-Submicrometer CMOS Technology," *Electron Device Letters, IEEE*, vol. 27, 2006, pp. 887-889.

[80] 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*, vol. 21, 2009, pp. 1020-1022.

[81] T. Baechler, S. Neukom, C. Lotto, and N. Blanc, "Single-Photon Resolution CMOS Integrating Image Sensors," *Procedia Chemistry*, vol. 1, 2009, pp. 1355-1358.

[82] M. Marwick and A. Andreou, "Single photon avalanche photodetector with integrated quenching fabricated in TSMC 0.18um 1.8V CMOS process," *Electronics Letters*, vol. 44, 2008, pp. 643-644.

[83] C. Niclass, M. Gersbach, R. Henderson, L. Grant, and E. Charbon, "A 130-nm CMOS single-photon avalanche diode," *Proceedings of SPIE*, Boston, MA, USA: 2007, pp. 676606-676606-9.

[84] C. Niclass, M. Gersbach, R.K. Henderson, L. Grant, and E. Charbon, "A Single Photon Avalanche Diode Implemented in 130nm CMOS Technology," 2007.

[85] M. Gersbach, J. Richardson, E. Mazaleyrat, S. Hardillier, C. Niclass, R. Henderson, L. Grant, and E. Charbon, "A low-noise single-photon detector implemented in a 130 nm CMOS imaging process," *Solid-State Electronics*, vol. 53, 2009, pp. 803-808.

[86] "p-n junction," *http://en.wikipedia.org/wiki/P-n_junction*.

[87] "p-n junctions," *http://ecee.colorado.edu/~bart/book/book/chapter4/ch4_2.htm*.

[88] S. Sze, *Physics of Semiconductor Devices*, Wiley & Sons, Inc., 1981.

[89] E. Schubert, "Reverse-Biased PN Junctions," 2010.

[90] A.G. Chynoweth, "Ionization Rates for Electrons and Holes in Silicon," *Physical Review*, vol. 109, 1958, p. 1537.

[91] A.G. Chynoweth and K.G. McKay, "Photon Emission from Avalanche Breakdown in Silicon," *Physical Review*, vol. 102, 1956, p. 369.

[92] K.G. McKay and K.B. McAfee, "Electron Multiplication in Silicon and Germanium," *Physical Review*, vol. 91, 1953, p. 1079.

[93] S.L. Miller, "Avalanche Breakdown in Germanium," *Physical Review*, vol. 99, 1955, p. 1234.

[94] J.L. Moll and R. Van Overstraeten, "Charge multiplication in silicon p-n junctions," *Solid-State Electronics*, vol. 6, 1963, pp. 147-157.

[95] E. Gramsch, M. Szawlowski, S. Zhang, and M. Madden, "Fast, high density avalanche photodiode array," *IEEE Transactions on Nuclear Science*, vol. 41, 1994, pp. 762-766.

[96] R.J. Locker, "A new ionizing radiation detection concept which employs semiconductor avalanche amplification and tunnel diode element," *Applied Physics Letters*, vol. 9, 1966, p. 227.

[97] J.P. Donnelly, K.A. McIntosh, D.C. Oakley, A. Napoleone, S.H. Groves, S. Vernon, L.J. Mahoney, K. Molvar, J. Mahan, J.C. Aversa, E.K. Duerr, Z.L. Liau, B.F. Aull, and D.C. Shaver, "1-um Geiger-mode detector development," *Laser Radar Technology and Applications X*, G.W. Kamerman, Ed., Orlando, FL, USA: SPIE, 2005, pp. 281-287.

[98] D.M. Taylor, J.C. Jackson, A.P. Morrison, A. Mathewson, and J.G. Rarity, "Characterization of novel active area silicon avalanche photodiodes operating in the geiger mode," *Journal of Modern Optics*, vol. 51, 2004, p. 1323.

[99] J. Baudot, W. Dulinski, M. Winter, R. Barbier, E. Chabanat, P. Depasse, and N. Estre, "Photon detection with CMOS sensors for fast imaging," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 604, 2009, pp. 111-114.

[100] E. Charbon, "Towards large scale CMOS single-photon detector arrays for lab-on-chip applications," *Journal of Physics D: Applied Physics*, vol. 41, 2008, p. 094010.

[101] F. Zappa, A. Tosi, A.D. Mora, F. Guerrieri, and S. Tisa, “Single-photon avalanche diode arrays and CMOS microelectronics for counting, timing, and imaging quantum events,” *Quantum Sensing and Nanophotonic Devices VII*, M. Razeghi, R. Sudharsanan, and G.J. Brown, Eds., San Francisco, California, USA: SPIE, 2010, pp. 76082C-15.

[102] R.H. Haitz, “Model for the Electrical Behavior of a Microplasma,” *Journal of Applied Physics*, vol. 35, 1964, p. 1370.

[103] S. Cova, M. Ghioni, A. Lacaita, C. Samori, and F. Zappa, “Avalanche photodiodes and quenching circuits for single-photon detection,” *Applied Optics*, vol. 35, 1996, p. 1956.

[104] A. Rochas, “First passively-quenched single photon counting avalanche photodiode element integrated in a conventional CMOS process with 32ns dead time,” *Proceedings of SPIE*, Quebec City, Canada: 2003, pp. 107-115.

[105] S. Cova, A. Longoni, and A. Andreoni, “Towards picosecond resolution with single-photon avalanche diodes,” *Review of Scientific Instruments*, vol. 52, 1981, pp. 408-412.

[106] S. Cova, A. Longoni, and G. Ripamonti, “Active-Quenching and Gating Circuits for Single-Photon Avalanche Diodes (SPADs),” *IEEE Transactions on Nuclear Science*, vol. 29, 1982, pp. 599-601.

[107] A. Gallivanoni, I. Rech, D. Resnati, M. Ghioni, and S. Cova, “Monolithic active quenching and picosecond timing circuit suitable for large-area single-photon avalanche diodes,” *Optics Express*, vol. 14, 2006, pp. 5021-5030.

[108] M. Stipcevic, “Active quenching circuit for single-photon detection with Geiger mode avalanche photodiodes,” *Applied Optics*, vol. 48, 2009, pp. 1705-1714.

[109] H. Finkelstein, M.J. Hsu, S. Zlatanovic, and S. Esener, “Performance trade-offs in single-photon avalanche diode miniaturization,” *Review of Scientific Instruments*, vol. 78, 2007, p. 103103.

[110] D. Bisello, “Silicon avalanche detectors with negative feedback as detectors for high energy physics,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 367, 1995, pp. 212-214.

[111] V. Golovin and V. Saveliev, “Novel type of avalanche photodetector with Geiger mode operation,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 518, 2004, pp. 560-564.

[112] R. Wolffenbuttel, “Integrated Silicon Colour Sensors,” Delft University, 1998.

[113] M.A. Green and M.J. Keevers, “Optical properties of intrinsic silicon at 300 K,” *Progress in Photovoltaics: Research and Applications*, vol. 3, 1995, pp. 189-192.

[114] HTE Labs, “HTE Labs - Si-Silicon, physical constants at 300K, silicon basic parameters, silicon properties.”

[115] W.J. Kindt, “Geiger mode avalanche photodiode arrays: For spatially resolved single photon counting,” 1999.

[116] S. Vasile, P. Gothoskar, R. Farrell, and D. Sdrulla, “Photon detection with high gain avalanche photodiode arrays,” *IEEE Transactions on Nuclear Science*, vol. 45, 1998, pp. 720-723.

[117] F. Zappa, M. Ghioni, S. Cova, L. Varisco, B. Sinnis, A. Morrison, and A. Mathewson, “Integrate Array of Avalanche Photodiodes for Single-Photon Counting,” *Proceedings of the 27th European Solid-State Device Research Conference ESSDERC '97*, 1997, pp. 600-603.

[118] W. Shockley and W.T. Read, “Statistics of the Recombinations of Holes and Electrons,” *Physical Review*, vol. 87, 1952, p. 835.

[119] G.A.M. Hurkx, D. Klaasen, and M. Knuvers, “A new recombination model for device simulation including tunneling,” *IEEE Transactions on Electron Devices*, vol. 39, 1992, pp. 331-338.

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

[121] S. Cova, M. Ghioni, A. Lotito, I. Rech, and F. Zappa, “Evolution and prospects for single-photon avalanche diodes and quenching circuits,” *Journal of Modern Optics*, vol. 51, 2004, pp. 1267-1288.

[122] K.E. Jensen, P.I. Hopman, E.K. Duerr, E.A. Dauler, J.P. Donnelly, S.H. Groves, L.J. Mahoney, K.A. McIntosh, K.M. Molvar, A. Napoleone, D.C. Oakley, S. Verghese, C.J. Vineis, and R.D. Younger, "Afterpulsing in Geiger-mode avalanche photodiodes for 1.06 μm wavelength," *Applied Physics Letters*, vol. 88, 2006, p. 133503.

[123] A. Rochas, M. Gani, B. Furrer, P.A. Besse, R.S. Popovic, G. Ribordy, and N. Gisin, "Single photon detector fabricated in a complementary metal–oxide–semiconductor high-voltage technology," *Review of Scientific Instruments*, vol. 74, 2003, p. 3263.

[124] H. Takesue, E. Diamanti, C. Langrock, M.M. Fejer, and Y. Yamamoto, "10-GHz clock differential phase shift quantum key distribution experiment," *Optics Express*, vol. 14, 2006, pp. 9522-9530.

[125] A. Spinelli, M. Ghioni, S. Cova, and L. Davis, "Avalanche detector with ultraclean response for time-resolved photon counting," *IEEE Journal of Quantum Electronics*, vol. 34, 1998, pp. 817-821.

[126] H. Finkelstein, M. Hsu, and S. Esener, "An ultra-fast Geiger-mode single photon avalanche diode in 0.18 μm CMOS technology," Boston, MA, USA: 2006.

[127] H. Finkelstein, M. Hsu, and S. Esener, "STI-bounded single-photon avalanche diode in a deep-submicron CMOS technology," *IEEE Electron Device Letters*, vol. 27, 2006.

[128] C.Y. Chang and S.M. Sze, *ULSI Technology*, Mcgraw-Hill College, 1996.

[129] H. Finkelstein, M. Hsu, and S. Esener, "A compact single-photon avalanche diode in a deep-submicron CMOS technology," Yokohama, Japan: 2006.

[130] M.J. Hsu, H. Finkelstein, and S.C. Esener, "A CMOS STI-Bound Single-Photon Avalanche Diode With 27-ps Timing Resolution and a Reduced Diffusion Tail," *IEEE Electron Device Letters*, vol. 30, 2009, pp. 641-643.

[131] H. Finkelstein, M.J. Hsu, S. Zlatanovic, and S. Esener, "Performance trade-offs in single-photon avalanche diode miniaturization," *Review of Scientific Instruments*, vol. 78, 2007, p. 103103.

[132] H. Finkelstein, M. Hsu, and S. Esener, "Dual-junction single-photon avalanche diode," *Electronics Letters*, vol. 43, 2007, pp. 1228-1229.

[133] Becker and Hickl, "MSA-1000."

[134] A. Lacaita, S. Cova, M. Ghioni, and F. Zappa, "Single-photon avalanche diode with ultrafast pulse response free from slow tails," *IEEE Electron Device Letters*, vol. 14, 1993, pp. 360-362.

[135] A. Lacaita, S. Cova, A. Spinelli, and F. Zappa, "Photon-assisted avalanche spreading in reach-through photodiodes," *Applied Physics Letters*, vol. 62, 1993, p. 606.

[136] S. Cova, A. Lacaita, M. Ghioni, G. Ripamonti, and T.A. Louis, "20-ps timing resolution with single-photon avalanche diodes," *Review of Scientific Instruments*, vol. 60, 1989, p. 1104.

[137] A. Spinelli and A.L. Lacaita, "Physics and numerical simulation of single photon avalanche diodes," *IEEE Transactions on Electron Devices*, vol. 44, 1997, pp. 1931-1943.

[138] S. Tisa, F. Guerrieri, and F. Zappa, "Variable-load quenching circuit for single-photon avalanche diodes," *Optics Express*, vol. 16, 2008, pp. 2232-2244.

[139] C. Niclass, M. Sergio, and E. Charbon, "A single photon avalanche diode array fabricated in 0.35-μm CMOS and based on an event-driven readout for TCSPC experiments," 2006, p. 23.

[140] G.N. Gol'tsman, A. Korneev, I. Rubtsova, I. Milostnaya, G. Chulkova, O. Minaeva, K. Smirnov, B. Voronov, W. Słysz, A. Pearlman, A. Verevkin, and R. Sobolewski, "Ultrafast superconducting single-photon detectors for near-infrared-wavelength quantum communications," *physica status solidi (c)*, vol. 2, 2005, pp. 1480-1488.

[141] E.A. Dauler, A.J. Kerman, B.S. Robinson, J.K.W. Yang, B. Voronov, G. Goltsman, S.A. Hamilton, and K.K. Berggren, "Photon-number-resolution with sub-30-ps timing using multi-element superconducting nanowire single photon detectors," *Journal of Modern Optics*, vol. 56, 2009, p. 364.

[142] M. Tarkhov, J. Claudon, J.P. Poizat, A. Korneev, A. Divochiy, O. Minaeva, V. Seleznev, N. Kaurova, B. Voronov, A.V. Semenov, and G. Gol'tsman, "Ultrafast reset time of superconducting single photon detectors," *Applied Physics Letters*, vol. 92, 2008, p. 241112.

[143] A. Kerman, E. Dauler, B. Robinson, R. Barron, D. Caplan, and M. Stevens, "Superconducting nanowire photon-counting detectors for optical communications."

[144] P. Feautrier, E.L. Coarer, R.E.D. Lamaestre, P. Cavalier, L. Maingault, J. Villégier, L. Frey, J. Claudon, N. Bergeard, M. Tarkhov, and J. Poizat, "High-speed superconducting single photon detectors for innovative astronomical applications," *Journal of Physics: Conference Series*, vol. 97, 2008, p. 012087.

[145] D. Rosenberg, S.W. Nam, P.A. Hiskett, C.G. Peterson, R.J. Hughes, J.E. Nordholt, A.E. Lita, and A.J. Miller, "Quantum key distribution at telecom wavelengths with noise-free detectors," *Applied Physics Letters*, vol. 88, 2006, p. 021108.

[146] V. Golovin and V. Saveliev, "Novel type of avalanche photodetector with Geiger mode operation," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 518, 2004, pp. 560-564.

[147] K. Zhao, S. You, J. Cheng, and Y. Lo, "Self-quenching and self-recovering InGaAs⁄InAlAs single photon avalanche detector," *Applied Physics Letters*, vol. 93, 2008, p. 153504.

[148] R.T. Thew, S. Tanzilli, L. Krainer, S.C. Zeller, A. Rochas, I. Rech, S. Cova, H. Zbinden, and N. Gisin, "Low jitter up-conversion detectors for telecom wavelength GHz QKD," *New Journal of Physics*, vol. 8, 2006, pp. 32-32.

[149] H. Finkelstein, M. Gross, Y. Lo, and S. Esener, "Analysis of hot-carrier luminescence for infrared single-photon upconversion and readout," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 13, 2007, pp. 959-966.

[150] M. Lahbabi, "Analysis of electroluminescence spectra of silicon and gallium arsenide p–n junctions in avalanche breakdown," *Journal of Applied Physics*, vol. 95, 2004, p. 1822.

[151] A.L. Lacaita, F. Zappa, S. Bigliardi, and M. Manfredi, "On the bremsstrahlung origin of hot-carrier-induced photons in silicon devices," *IEEE Transactions on Electron Devices*, vol. 40, 1993, pp. 577-582.

[152] S. Esener and Y. Lo, "UCSD CMOS Compatible _IR-SPAD Project Update," Feb. 2010.

[153] MMR Technologies, "Low Temperature Microprobe Operating Manual."

[154] W. Cheong, S. Prahl, and A. Welch, "A review of the optical properties of biological tissues," *IEEE Journal of Quantum Electronics*, vol. 26, 1990, pp. 2166-2185.

[155] F. Gao, H. Zhao, and Y. Yamada, "Improvement of Image Quality in Diffuse Optical Tomography by use of Full Time-Resolved Data," *Applied Optics*, vol. 41, 2002, pp. 778-791.

[156] A.P. Gibson, J.C. Hebden, and S.R. Arridge, "Recent advances in diffuse optical imaging," *Physics in Medicine and Biology*, vol. 50, 2005, pp. R1-R43.

[157] Y. Xu, N. Iftimia, H. Jiang, L.L. Key, and M.B. Bolster, "Three-dimensional diffuse optical tomography of bones and joints," *Journal of Biomedical Optics*, vol. 7, 2002, pp. 88-92.

[158] B.E. Bouma and G.J. Tearney, *Handbook of optical coherence tomography*, Informa Health Care, 2002.

[159] W. Drexler, Y. Chen, A. Aguirre, B. Povaz□ay, A. Unterhuber, and J.G. Fujimoto, "Ultrahigh Resolution Optical Coherence Tomography," *Optical Coherence Tomography*, 2008, pp. 239-279.

[160] J.G. Fujimoto, "Optical coherence tomography," *Comptes Rendus de l'Académie des Sciences - Series IV - Physics*, vol. 2, 2001, pp. 1099-1111.

[161] M.R. Hee, J.A. Izatt, E.A. Swanson, D. Huang, J.S. Schuman, C.P. Lin, C.A. Puliafito, and J.G. Fujimoto, "Optical Coherence Tomography of the Human Retina," *Arch Ophthalmol*, vol. 113, 1995, pp. 325-332.

[162] C.A. Puliafito, M.R. Hee, C.P. Lin, E. Reichel, J.S. Schuman, J.S. Duker, J.A. Izatt, E.A. Swanson, and J.G. Fujimoto, "Imaging of macular diseases with optical coherence tomography," *Ophthalmology*, vol. 102, 1995, pp. 217-229.

[163] W. Leutz and G. Maret, "Ultrasonic modulation of multiply scattered light," *Physica B: Condensed Matter*, vol. 204, 1995, pp. 14-19.

[164] F.A. Marks, H.W. Tomlinson, and G.W. Brooksby, "Comprehensive approach to breast cancer detection using light: photon localization by ultrasound modulation and tissue characterization by spectral discrimination," 1993.

[165] L. Wang, S.L. Jacques, and X. Zhao, "Continuous-wave ultrasonic modulation of scattered laser light to image objects in turbid media," *Optics Letters*, vol. 20, 1995, pp. 629-631.

[166] M. Kempe, M. Larionov, D. Zaslavsky, and A.Z. Genack, "Acousto-optic tomography with multiply scattered light," *Journal of the Optical Society of America A*, vol. 14, 1997, pp. 1151-1158.

[167] A. Lev, Z. Kotler, and B.G. Sfez, "Ultrasound tagged light imaging in turbid media in a reflectance geometry," *Optics Letters*, vol. 25, 2000, pp. 378-380.

[168] J. Li, G. Ku, and L.V. Wang, "Ultrasound-Modulated Optical Tomography of Biological Tissue by Use of Contrast of Laser Speckles," *Applied Optics*, vol. 41, 2002, pp. 6030-6035.

[169] G. Yao, S. Jiao, and L.V. Wang, "Frequency-swept ultrasound-modulated optical tomography in biological tissue by use of parallel detection," *Optics Letters*, vol. 25, 2000, pp. 734-736.

[170] G.D. Mahan, W.E. Engler, J.J. Tiemann, and E. Uzgiris, "Ultrasonic tagging of light: Theory," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 95, 1998, pp. 14015-14019.

[171] L.V. Wang, "Mechanisms of Ultrasonic Modulation of Multiply Scattered Coherent Light: An Analytic Model," *Physical Review Letters*, vol. 87, 2001, p. 043903.

[172] E.C. Unger, T.O. Matsunaga, T. McCreery, P. Schumann, R. Sweitzer, and R. Quigley, "Therapeutic applications of microbubbles," *European Journal of Radiology*, vol. 42, May. 2002, pp. 160-168.

[173] P.A. Frenkel, S. Chen, T. Thai, R.V. Shohet, and P.A. Grayburn, "Dna-loaded albumin microbubbles enhance ultrasound-mediated transfection in vitro," *Ultrasound in Medicine & Biology*, vol. 28, Jun. 2002, pp. 817-822.

[174] R. Bekeredjian, P. Grayburn, and R. Shohet, "Use of ultrasound contrast agents for gene or drug delivery in cardiovascular medicine," *J Am Coll Cardiol*, vol. 45, Feb. 2005, pp. 329-335.

[175] M. Postema and G. Schmitz, "Ultrasonic bubbles in medicine: Influence of the shell," *Ultrasonics Sonochemistry*, vol. 14, Apr. 2007, pp. 438-444.

[176] P. Jiang, Z. Peng, and J.C. Lazenby, "A new tissue harmonic imaging scheme with better fundamental frequency cancellation and higher signal-to-noise ratio," *Ultrasonics Symposium, 1998. Proceedings., 1998 IEEE*, 1998, pp. 1589-1594.

[177] J.N. Marsh, M.S. Hughes, G.H. Brandenburger, and J.G. Miller, "Broadband measurement of the scattering-to-attenuation ratio for Albunex® at 37°C," *Ultrasound in Medicine & Biology*, vol. 25, Oct. 1999, pp. 1321-1324.

[178] N. de Jong, L. Hoff, T. Skotland, and N. Bom, “Absorption and scatter of encapsulated gas filled microspheres: theoretical considerations and some measurements,” *Ultrasonics*, vol. 30, Mar. 1992, pp. 95-103.

[179] D. Zhang, X.F. Gong, J.H. Liu, L.Z. Shao, X.R. Li, and Q.L. Zhang, “The experimental investigation of ultrasonic properties for a sonicated contrast agent and its application in biomedicine,” *Ultrasound in Medicine & Biology*, vol. 26, Feb. 2000, pp. 347-351.

[180] J.M. Correas, E. Kurtisovski, S.L. Bridal, A.L. Amararene, O. Hélénon, and G. Berger, “Optimizing an ultrasound contrast agent's stability using in vitro attenuation measurements,” *Investigative Radiology*, vol. 37, Dec. 2002, pp. 672-679.

[181] M.R. Sprague, E. Chérin, D.E. Goertz, and F.S. Foster, “Nonlinear Emission from Individual Bound Microbubbles at High Frequencies,” *Ultrasound in Medicine & Biology*, vol. 36, Feb. 2010, pp. 313-324.

[182] A.L. Klibanov, K.W. Ferrara, M.S. Hughes, J.H. Wible, J.K. Wojdyla, P.A. Dayton, K.E. Morgan, and G.H. Brandenburger, “Direct video-microscopic observation of the dynamic effects of medical ultrasound on ultrasound contrast microspheres,” *Investigative Radiology*, vol. 33, Dec. 1998, pp. 863-870.

[183] P.A. Frenkel, K.E. Morgan, A.L. Klibanov, G.H. Brandenburger, and K.W. Ferrara, “Optical and acoustical observations of the effects of ultrasound on contrast agents,” *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on*, vol. 46, 1999, pp. 220-232.

[184] J.E. Chomas, P.A. Dayton, D. May, J. Allen, A. Klibanov, and K. Ferrara, “Optical observation of contrast agent destruction,” *Applied Physics Letters*, vol. 77, 2000, pp. 1056-1058.

[185] V. Garbin, D. Cojoc, E. Ferrari, E. Di Fabrizio, M.L.J. Overvelde, S.M. van der Meer, N. de Jong, D. Lohse, and M. Versluis, “Changes in microbubble dynamics near a boundary revealed by combined optical micromanipulation and high-speed imaging,” *Applied Physics Letters*, vol. 90, Mar. 2007, pp. 114103-3.

[186] M. Postema, A. van Wamel, F.J. ten Cate, and N. de Jong, “High-speed photography during ultrasound illustrates potential therapeutic applications of microbubbles,” *Medical Physics*, vol. 32, Dec. 2005, pp. 3707-3711.

[187] J. Guan and T.J. Matula, “Using light scattering to measure the response of individual ultrasound contrast microbubbles subjected to pulsed ultrasound in vitro,” *The Journal of the Acoustical Society of America*, vol. 116, Nov. 2004, pp. 2832-2842.

[188] J. Tu, J. Guan, Y. Qiu, and T.J. Matula, “Estimating the shell parameters of SonoVue microbubbles using light scattering,” *The Journal of the Acoustical Society of America*, vol. 126, Dec. 2009, pp. 2954-2962.

[189] D.L. Kingsbury and P.L. Marston, “Mie scattering near the critical angle of bubbles in water,” *Journal of the Optical Society of America*, vol. 71, Mar. 1981, pp. 358-361.

[190] P.L. Marston and S.C. Billette, “Scattering of light by a coated bubble in water near the critical and Brewster scattering angles,” *The Journal of the Acoustical Society of America*, vol. 80, Dec. 1986, p. S59.

[191] M.S. Plesset and A. Prosperetti, “Bubble Dynamics and Cavitation,” Nov. 2003.

[192] C.C. Church, “The effects of an elastic solid surface layer on the radial pulsations of gas bubbles,” *The Journal of the Acoustical Society of America*, vol. 97, Mar. 1995, pp. 1510-1521.

[193] K.E. Morgan, J.S. Allen, P.A. Dayton, J.E. Chomas, A.L. Klibaov, and K.W. Ferrara, “Experimental and theoretical evaluation of microbubble behavior: effect of transmitted phase and bubble size,” *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on*, vol. 47, 2000, pp. 1494-1509.

[194] S. Yang, D.E. Kruse, P.A. Dayton, and K.W. Ferrara, “High-frequency dynamics of ultrasound contrast agents,” *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on*, vol. 52, 2005, pp. 1981-1991.

[195] Definity Prescribing Information, “Definity Prescribing Information 2008,” 2008.

[196] T.G. Leighton, *The Acoustic Bubble*, Academic Press, 1997.

[197] P.L. Marston and D.L. Kingsbury, “Scattering by a bubble in water near the critical angle: interference effects,” *Journal of the Optical Society of America*, vol. 71, Feb. 1981, pp. 192-196.

[198] S. Kvåle, H.A. Jakobsen, O.A. Asbjørnsen, and T. Omtveit, “Size fractionation of gas-filled microspheres by flotation,” *Separations Technology*, vol. 6, Oct. 1996, pp. 219-226.

[199] M.W. Grinstaff and K.S. Suslick, “Air-filled proteinaceous microbubbles: synthesis of an echo-contrast agent,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 88, 1991, pp. 7708-7710.

[200] L. Hoff, P.C. Sontum, and J.M. Hovem, “Oscillations of polymeric microbubbles: effect of the encapsulating shell,” *The Journal of the Acoustical Society of America*, vol. 107, Apr. 2000, pp. 2272-2280.

[201] J.E. Chomas, P. Dayton, J. Allen, K. Morgan, and K.W. Ferrara, “Mechanisms of contrast agent destruction,” *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on*, vol. 48, 2001, pp. 232-248.

[202] M. Kobayashi, T. Mizumoto, Y. Shibuya, M. Enomoto, and M. Takeda, “Fluorescence tomography in turbid media based on acousto-optic modulation imaging,” *Applied Physics Letters*, vol. 89, 2006, p. 181102.

[203] D.J. Hall, U. Sunar, and S. Farshchi-Heydari, “Quadrature Detection of Ultrasound-Modulated Photons with a Gain-Modulated, Image-Intensified, CCD Camera,” *The Open Optics Journal*, vol. 2, 2008, pp. 75-78.

[204] B. Yuan, J. Gamelin, and Q. Zhu, “Mechanisms of the ultrasonic modulation of fluorescence in turbid media,” *Journal of Applied Physics*, vol. 104, 2008, p. 103102.

[205] X.D. Li, M.A. O'Leary, D.A. Boas, B. Chance, and A.G. Yodh, “Fluorescent diffuse photon density waves in homogeneous and heterogeneous turbid media: analytic solutions and applications,” *Applied Optics*, vol. 35, 1996, pp. 3746-3758.

[206] J.R. Lakowicz, *Principles of fluorescence spectroscopy*, Springer, 2006.

[207] B. Yuan and Y. Liu, “Ultrasound-modulated fluorescence from rhodamine B aqueous solution,” *Journal of Biomedical Optics*, vol. 15, 2010, pp. 021321-6.

[208] B. Yuan, Y. Liu, P.M. Mehl, and J. Vignola, “Microbubble-enhanced ultrasound-modulated fluorescence in a turbid medium,” *Applied Physics Letters*, vol. 95, 2009, p. 181113.

[209] B. Yuan, “Ultrasound-modulated fluorescence based on a fluorophore-quencher-labeled microbubble system,” *Journal of Biomedical Optics*, vol. 14, 2009, pp. 024043-11.

[210] K.D. Chaudhuri, "Concentration quenching of fluorescence in solutions," *Zeitschrift für Physik A Hadrons and Nuclei*, vol. 154, 1959, pp. 34-42.

[211] R.F. Chen and J.R. Knutson, "Mechanism of fluorescence concentration quenching of carboxyfluorescein in liposomes: energy transfer to nonfluorescent dimers," *Analytical Biochemistry*, vol. 172, 1988, pp. 61-77.

[212] S.I. Druzhinin, A.B. Demyashkevich, and B.M. Uzhinov, "Concentration dependence of the quenching of the fluorescence of acridine cation in aqueous solutions," *Journal of Applied Spectroscopy*, vol. 30, 1979, pp. 238-239.

[213] L.J. Jones, R.H. Upson, R.P. Haugland, N. Panchuk-Voloshina, M. Zhou, and R.P. Haugland, "Quenched BODIPY Dye-Labeled Casein Substrates for the Assay of Protease Activity by Direct Fluorescence Measurement," *Analytical Biochemistry*, vol. 251, 1997, pp. 144-152.

[214] D. Hoekstra, T. de Boer, K. Klappe, and J. Wilschut, "Fluorescence method for measuring the kinetics of fusion between biological membranes," *Biochemistry*, vol. 23, 1984, pp. 5675-5681.

[215] H. Wendt, C. Berger, A. Baici, R.M. Thomas, and H.R. Bosshard, "Kinetics of folding of leucine zipper domains," *Biochemistry*, vol. 34, 1995, pp. 4097-4107.

[216] X. Zhuang, T. Ha, H.D. Kim, T. Centner, S. Labeit, and S. Chu, "Fluorescence quenching: A tool for single-molecule protein-folding study," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 97, 2000, pp. 14241-14244.