To the Graduate Council:

I am submitting herewith a dissertation written by Michael Joseph Paulus entitled "Investigation of a New Positron Emission Tomography Detector Using Bismuth Germanate, Avalanche Photodiodes and Low Noise, Wide Band CMOS Preamplifiers." I have examined the final copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Electrical Engineering.

James M. Rochelle

James M. Rochelle, Major Professor

We have read this dissertation
and recommend its acceptance

T.V. Blalock

E.A. Kennedy

Donald W. Bouldin

Accepted for the Council:

C.W. Minkel

Associate Vice Chancellor and
Dean of The Graduate School

Investigation of a New Positron Emission Tomography Detector Using Bismuth Germanate, Avalanche Photodiodes and Low Noise, Wide Band CMOS Preamplifiers

A Dissertation
Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Michael J. Paulus
May, 1995

DEDICATION

This dissertation is dedicated, with love, to my wife

Margaret M. Paulus

whose love and support carried me through this project.

ACKNOWLEDGMENTS

This research and manuscript preparation could not have been completed without the support and encouragement of a number of people. I am very fortunate to be employed at CTI, an exciting research oriented company committed to making Positron Emission Tomography a widely available medical imaging modality. It is a privilege to participate in the development a tool whose sole purpose is to save human life. I am especially grateful to my management, Ron Nutt, Wilfried Loeffler, Mark Andreaco and Clif Moyers for supporting and encouraging this research. I am also indebted to my technical mentors, Dave Binkley and Mike Casey, with whom I had many helpful discussions, and to Mark Long, Brian Swann, Steve Hudson and Kevin Behel for their assistance in the collection of the data presented here.

I am grateful to the National Institute of Mental Health for funding this research with Phase I and Phase II SBIR support (grants# 1R43MH50, and 2R43MH50).

It has been a wonderful opportunity to study in the Electrical Engineering Department of the University of Tennessee. I am grateful to my dissertation committee, Dr. J. Rochelle, Dr. T. V. Blalock, Dr. D. Bouldin, Dr. W. Bugg and Dr. E. J. Kennedy for their guidance and patience as this manuscript came together in the waning days of the semester.

A great many people offered much needed encouragement during the course of my studies. I am grateful to my parents, T. J. and Sue Paulus, for their love and encouragement as I sought to follow their example of building a loving and nurturing family while simultaneously pursuing a demanding career and seeking a graduate degree at night. I am grateful to my brother, Rick, for his frequent notes of encouragement sent via e-mail from the frozen Northwest, and to my parents-in-law, Ed and Peg Moy, for their thoughtful cards and notes of support. Thanks are also due to my neighbor and friend, Nance Ericson, who kept my lawn from becoming an object of neighborhood contempt during the final months of this project.

And finally, I am most grateful to my wife, Margaret, and my daughters, Emily and Paige. My family is my greatest source of joy. Without their love, support and patience I could not have reached this day.

ABSTRACT

A new Positron Emission Tomography (PET) detector is investigated which replaces conventional photomultiplier tubes (PMTs) with avalanche photodiodes (APDs). As a part of the investigation a low-noise, wide-band CMOS preamplifier array is designed and tested with APD based detectors.

The fundamental theory of APD operation is first reviewed and the governing equations are derived. The frequency dependent behavior of the APD is calculated and a new treatment of the frequency response of the undepleted p-region based on the classic model for graded base transistors is presented.

A modeling methodology is developed to simulate APD performance based on the device fabrication procedure. Using this methodology, the performance characteristics of the beveled-edge and reach-through APD structures are compared to determine which device is most appropriate for use in a PET detector. The model results indicate the beveled-edge APD has superior characteristics for PET due to its lower excess-noise factor and lower dark current. Commercially available beveled-edge and reach-through APDs are experimentally evaluated to confirm the model predictions.

A new wide-band, low-noise CMOS preamplifier topology is developed for use with APD based PET detectors. A 6-element prototype preamplifier array is designed, fabricated and evaluated in a commercial 2 μm CMOS fabrication line. The preamplifiers in the array have rise-times of 7 ns and an equivalent input noise voltages of 1.1 nV/rt-Hz with an input load capacitance of 6 pF due to the package. When coupled to a 5 mm diameter beveled-edge APD and 3 x 3 x 30 mm^3 BGO crystal, the amplifier and detector have a measured pulse height resolution of 14% (511 keV) and a measured timing resolution of 9.1 ns FWHM.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PART 1

INTRODUCTION

Background

The goal of this study is to investigate the potential for replacing the photomultiplier tubes currently used in Positron Emission Tomography (PET) detectors with silicon avalanche photodiode arrays. PET is a medical imaging technique that detects coincident 511 keV γ-rays in order to map the position and density of a positron emitting species in the body. This section reviews the process by which the 511 keV γ-rays are emitted, provides a brief overview of PET, describes the existing photomultiplier tube based detector technology and provides an introduction to avalanche photodiode (APD) technology.

Positron Emission Tomography

Positron Emission

The γ-rays detected by a PET scanner are generated by β^+ radioactive decay processes occurring within the patient. This section presents an overview of the β^+ decay process following the discussions presented in [1], and [2]. The following definitions and symbols are used:

Atomic Number (Z):	The number of protons (or electrons) in an atom.
Mass Number (N):	The total number of protons and neutrons in an atom.
Nucleon:	A subatomic particle which resides in the atomic nucleus.
Proton (P):	A nucleon with a mass of 1.673×10^{-27} kg and a charge of $+1.602 \times 10^{-19}$ C.
Neutron (N):	An uncharged nucleon with a mass of 1.675×10^{-27} kg.
Electron (e^- or β^-):	A subatomic particle which orbits the nucleus with a mass of 9.11×10^{-31} kg and a charge of -1.602×10^{-19}C. β^- denotes an electron which is emitted during a radioactive decay process.
Positron (e^+ or β^+):	A subatomic particle with a mass of 9.11×10^{-31} kg and a charge of $+1.602 \times 10^{-19}$C. β^+ denotes a positron which is emitted during a radioactive decay process.
Neutrino (ν or $\overline{\nu}$):	A subatomic particle emitted during β^- or β^+ decays. The

	neutrino has no charge and little or no rest mass. The neutrino (ν) is emitted during a β^+ decay while the antineutrino ($\overline{\nu}$) is emitted during a β^- decay.
Parent nucleus:	A radioactive nucleus which has not yet undergone radioactive decay.
Daughter nucleus:	The new nucleus which results from a radioactive decay.

Within an atomic nucleus, nucleons exert attractive forces on their nearby neighbors. These *hadronic* forces are stronger than the electric forces of repulsion between protons and are responsible for binding nuclei together. For every stable nucleus of atomic number Z there exists a minimum number of neutrons required to compensate for the repulsive forces of the protons; for lighter nuclei there is typically one neutron for every proton while heavier nuclei typically have more neutrons than protons. If a nucleus has fewer than its optimal number of neutrons, it will relax to a more stable state by converting one or more of its excess protons to neutrons. This is accomplished via a β^+ decay, in which a proton is converted to a neutron as described by the expression

$$P \rightarrow N + \beta^{+} + \nu + energy. \tag{1-1}$$

The resulting daughter nucleus will have the same mass number as the parent but its atomic number will be reduced by one. The energy term in Equation 1-1 is the kinetic energy imparted to the neutrino and the positron plus the typically small recoil energy of daughter nucleus. The ratio of the positron energy to the neutrino energy is a random variable described by the Fermi probability distribution function where the mean positron energy is approximately half the total energy released in the decay.

After emission the positron loses its kinetic energy via coulombic interactions with nearby electrons. Most of the radioactive species used in PET have maximum kinetic energies on the order of a few MeV, limiting their range of travel to a few millimeters. When the positron comes to rest, it quickly combines with a nearby electron and annihilates, converting the masses of the positron and electron to electromagnetic energy in the form of two 511 keV γ-rays emitted 180° to one another. (The angle of emission will deviate slightly from 180° if the positron annihilates before coming completely to rest.) β^+ decays are therefore readily observed experimentally by detecting the presence of the two simultaneously emitted (coincident) 511 keV γ-rays.

Description of the Positron Emission Tomography (PET) System

Positron Emission Tomography is a medical imaging modality based upon the detection of coincident 511 keV γ-rays emitted by compounds labeled with positron emitting isotopes. Commonly used tracer nuclei include ^{11}C, ^{13}N, and ^{15}O, which may be substituted for their stable isotopes, and ^{18}F which is frequently used as a hydrogen substitute. Many physiologically important compounds such as drugs, glucose or oxygen are readily labeled using these isotopes. The PET scanner measures the tracer compounds' locations and concentrations, enabling the physician to monitor their activity *in vivo*. Thus, unlike other medical imaging modalities which map the structure of the body, PET maps metabolic processes occurring in the body, enabling the physician to detect physiological abnormalities before they induce structural variations detectable with other medical imaging techniques.

Conceptually, the PET scanner is simply a ring of γ-ray detectors surrounding the patient, as shown in Figure 1-1*. When two detectors in the ring simultaneously (within a few nanoseconds) observe 511 keV γ-rays, a coincident event is said to have occurred and the associated positron annihilation is known to have occurred along the axis between the two detectors. After many thousands of coincident events have been observed and recorded, a map of the radioisotope concentration can be constructed. A

* All figures may be found at the end of each part.

typical PET scan of a healthy human brain obtained using 2-[^{18}F]fluoro-2-deoxy-glucose (FDG) is shown in Figure 1-2. FDG, a glucose compound which has been tagged with ^{18}F, is used to map the rate at which glucose is absorbed by the body. The dark areas of the brain scan represent regions of rapid glucose absorption, indicating a high metabolic rate in those regions. The lighter areas indicate regions of slower glucose absorption.

The first laboratory PET scanner was developed by Phelps and Hoffman in 1975 [3] and the first commercial scanner was manufactured by EG&G ORTEC in 1978 [4]. The current PET market is dominated by two major manufacturers [5]with several other smaller corporations also offering scanners. Approximately 150 PET systems are in operation world-wide today.

Current Detector Technology

PET scanners detect γ-rays in a two step scintillation process in which the γ-ray energy is first converted to visible light in a scintillator, typically Bismuth Germanate ($Bi_4Ge_3O_{12}$ or BGO) and then the visible light is detected using a photomultiplier tube or avalanche photodiode. The following review of scintillation detector technology is based on discussions presented in Ref. [1]. BGO is an organic scintillator with a band structure somewhat similar to a semiconductor. The atomic shells of the crystal components are separated by the crystal formation so that a band of unfilled upper allowed electron states, S_1, is separated from the filled lower allowed electron states, S_0, by a forbidden gap. Unlike semiconductors, however, the S_0 and S_1 states do not form continuous valence and conduction bands in which carriers can freely travel but instead form an array of closely spaced discrete energy levels. When an energetic charged particle passes through the scintillator it excites S_0 electrons into S_1 via coulombic interactions. The excited electrons relax back to the S_1 states through photon emissions where the photon emission rate is a decaying exponential function with a 300 ns time constant.

Because the S_0 and S_1 levels are not true energy bands, the emitted light is not monochromic; instead a roughly gaussian shaped spectrum of light is emitted with a maximum at 480 nm and a range of approximately 350 nm to 700 nm. The γ-rays of interest in PET are uncharged particles, however, and cannot directly excite electrons into the BGO S_1 states. Instead, the γ-rays transfer their kinetic energy to electrons through *photoelectric absorption* or *Compton scattering* and the resultant energetic electrons excite the S_0 electrons into the S_1 states. Photoelectric absorption is the preferred γ-ray interaction in which the γ-ray's energy is fully absorbed by an atom within the scintillator. The atom sheds the energy absorbed from the γ-ray by ejecting an energetic electron with kinetic energy equal to the energy of the γ-ray minus the electron binding energy. The vacancy left by the electron is quickly filled by a nearby free electron, and a characteristic x-ray is emitted. The x-ray normally undergoes a second photoelectric absorption through which a second photoelectron is emitted with kinetic energy approximately to the binding energy of the first photoelectron. Thus, a photoelectric absorption typically results in the emission of two photoelectrons: the primary photoelectron carries the majority of the energy of the γ-ray and the secondary photoelectron carries the remainder. The kinetic energy of the two photoelectrons is converted to visible light in the scintillator as described above, and the detector output due to photoelectric absorption appears as a gaussian-shaped "photopeak" as shown in Figure 1-3.

The Compton scattering process occurs between the incident γ-ray and an electron in the scintillator. The γ-ray is deflected by the electron and transfers some of its kinetic energy to it, where the amount of energy remaining with the scattered γ-ray, $h\nu'$, is related to the angle of deflection by the expression

$$h\nu' = \frac{h\nu}{1 + \frac{h\nu}{m_o c^2}(1 - \cos\theta)}, \tag{1-2}$$

where h is Plank's constant, ν is the γ-ray frequency, m_o is the rest mass of an electron, c is the speed of light in a vacuum and θ is the γ-ray angle of deflection. For 511 keV γ-rays, Equation 1-2 reduces to

$$h\nu' = \frac{h\nu}{2 - \cos\theta}, \tag{1-3}$$

indicating a 511 keV γ-ray undergoing a Compton scattering event will impart no more than half of its energy to the recoil electron. As shown in Figure 1-3, γ-rays which undergo single Compton scattering interactions are detected as low energy events with a continuum of possible energies ranging from 0 to 255 keV. In detector geometries typically used for PET, many γ-rays undergo multiple Compton interactions causing a broadening of the Compton continuum.

In most PET applications, γ-rays which are detected through Compton scattering interactions in the scintillator are rejected because they are indistinguishable from γ-rays which scatter before reaching the detector. Consequently, scintillators which absorb γ-rays via photoelectric absorption more readily than through Compton scattering are preferred. High Z scintillators have this property. The probability per scintillator atom that a γ-ray will undergo photoelectric absorption is approximately proportional to $Z^{4.5}$ while the probability per scintillator atom that a γ-ray will undergo a Compton scattering event increases linearly with Z. Therefore, a high Z scintillator will collect more total γ-rays than a lower Z scintillator, and the ratio of photoelectric absorptions to Compton scattering events will be greater in a high Z scintillator. It is this consideration which has made BGO the most widely used scintillator for PET; the bismuth component has an atomic number of 83, giving BGO the largest probability per atom of any commonly available scintillator that an incident γ-ray will undergo photoelectric absorption.

The major properties of BGO as reported in [1] are compared with those of Thallium Activated Sodium Iodide (NaI(Tl)) in Table 1-1. NaI(Tl) is the most commonly used scintillator for general γ-ray spectroscopy and is the standard against which other scintillators are compared. Note that although BGO has a significantly higher specific gravity than NaI(Tl), the light delivered from BGO to a typical PMT is only 13% of that delivered by NaI(Tl).

Table 1-1.

Comparison of NaI(Tl) and BGO Properties [1].

Property	**NaI(Tl)**	**BGO**
Specific Gravity	3.67	7.13
Wavelength of Maximum Emission λmax (nm)	415	505
Index of Refraction at λmax.	1.85	2.15
Principal Decay Constant (ns)	230	300
Total Light Yield in Photons/MeV	38000	8200
Absolute Scintillation Efficiency for Fast Electrons	11.3%	2.1%
Relative γ-ray Pulse Height with Bialkali PMT	1.00	.13

The light detection device coupled to the BGO must convert an extremely small (typically on the order of 1000 photons) light output into a useful electrical signal. PMTs have traditionally been selected for this application because they detect the scintillator light efficiently and they have tremendous internal gain with very little associated noise. The goal of this study is to assess the potential for replacing PMTs with APDs; a brief overview of PMT characteristics is presented here in order to establish a bench-mark against which the APD may be compared.

A schematic diagram of a head-on type PMT similar to that typically used for PET is shown in Figure 1-4. The PMT consists of a photocathode, focusing electrodes, a dynode string and an anode; physically it is approximately four inches long and has a surface area of about one square inch. The photocathode has a band structure much like a semiconductor such that when an incident photon is absorbed an electron is excited into the conduction band. The electron diffuses across the photocathode and, if it has sufficient energy, it escapes the photocathode into the vacuum region below. The quantum efficiency, η, is the average number of electrons emitted by the photocathode per photon. PMTs used for PET typically have quantum efficiencies on the order of 20% for the wavelength of light emitted by BGO. Thus, if a scintillator coupled to a PMT emits N photons upon absorbing an incident nuclear particle, the photocathode will on the average emit ηN photoelectrons.

The emitted electrons (frequently called photo-electrons) are accelerated into the first dynode, D1, by the focusing electrodes. When the energetic electrons are absorbed by the first dynode, the imparted energy results in the re-emission of more than one electron from the same surface. This process, called secondary-emission multiplication, is the gain mechanism for PMTs. The secondary emission ratio,

$$\delta = \frac{number\ of\ electrons\ emitted}{number\ of\ electrons\ absorbed}, \tag{1-4}$$

is roughly proportional to field between the dynodes and is typically on the order of 4 to 5 for PMTs used in PET. The electrons emitted by the first dynode are absorbed and reemitted by the second dynode and so on so that current gains in excess of 10^6 are readily attained with 10 dynode PMTs.

The two primary sources of noise in a PMT are dark-current noise and statistical noise associated with the multiplication process. In the absence of an input light flux, there exists a finite probability that electrons will be spontaneously emitted from the photocathode or any one of the dynodes. Each spontaneously emitted electron will be swept down the chain of dynodes to the anode, inducing secondary emissions in each dynode it encounters. Thus, spontaneous emission results in a standing dark current, I_d, which is always present. The dark current typically has a magnitude on the order of 10^{-9} A and contributes full shot noise to the signal.

For low light output scintillators such as BGO, the dark-current noise is dwarfed by the statistical noise of the multiplication process. The secondary emission process is well described by the Poisson density function where the mean number of secondary electrons is proportional to the number of incident electrons, N_i, and the standard deviation of the number of secondary electrons is proportional to $\sqrt{N_i}$. The signal to noise ration of a signal leaving a dynode is therefore

$$SNR = \frac{N_i}{\sqrt{N_i}} = \sqrt{N_i}. \tag{1-5}$$

The first dynode sees the photoelectrons prior to any multiplication, so it contributes the most statistical noise to the signal. It is common to assume the statistical noise is dominated by the first dynode, giving rise to the following approximate expression for the signal to noise ratio of a PMT coupled to a scintillator:

$$SNR \approx \frac{\eta N}{\sqrt{\eta N}} = \sqrt{\eta N}, \tag{1-6}$$

where N is the number of photons emitted by the scintillator and η is the PMT quantum efficiency. Note that the signal-to-noise ratio is directly proportional to $\sqrt{\eta}$.

The first PET scanners used simple detectors consisting of discrete BGO crystals coupled to individual PMTs (Figure 1-5a). The primary advantage of this detector is its simplicity: Each crystal has its own PMT and electronics, so no decoding is required to determine which detector channel absorbs the γ-ray. The key disadvantages of this detector are its cost and size. The high cost of PMTs and their electronics make the detector impractical for commercial systems. Also, the smallest available PMTs have diameters of about 10 mm, setting the detector resolution well above the theoretical PET resolution limit of ~2 mm.

The block detector was introduced by CTI in 1986 [6] and is currently used in some form on every commercial PET scanner built today (Figure 1-5b). The CTI/Siemens block detector consists of four PMTs coupled to a single block of BGO which has been partially sawed into 64 pixels. When the block absorbs a γ-ray, the 4 PMT signals are decoded to determine which of the 64 pixels was illuminated. The block concept has the advantages of using far fewer PMTs than the earlier detectors and of having BGO pixels which are much smaller than the minimum available PMT area.

Despite these advantages, block detectors have drawbacks. While less expensive than the small PMT detector, block detectors account for up to 30% of the overall cost of a high resolution PET system. Furthermore, current calibration technology limits commercial 4 PMT block detectors to no more than 64 pixels, although recent research indicates blocks with as many as 256 pixels may one day be feasible[7]. Because the number of pixels per block is limited, any reduction in the pixel size increases the total number of PMTs in the scanner. Such a reduction in pixel size does not significantly increase the BGO cost since the volume of BGO used does not change, but it does increase the PMT costs, because the cost per PMT is largely independent of size. Therefore, the primary source of the increased cost for high resolution (small pixel) PET systems is the cost of the additional PMTs.

Block detectors have performance limitations as well. Information from all four photomultiplier tubes is used in the block decoding scheme, so each time one of the crystals detects an event, all 64 channels are disabled while the event is processed. The system throughput is therefore lower than for individually coupled crystals. (Stated another way, a block based system has a higher dead time than a system with individually coupled crystals). Increased dead time translates directly to longer scan times and reduced statistical quality in the images. A second disadvantage is due to imperfections in the decoding scheme. A finite probability exists that detected events will be assigned to the wrong pixel, causing a reduction in the spatial resolution of the PET image. The probability for decoding errors increases as the number of detectors per block is increased. These disadvantages of the presently used block detector scheme provide strong motivation for research into alternative detector schemes.

The Avalanche Photodiode

Basic Avalanche Photodiode Structure

In its simplest form, an APD is a reverse-biased p-n junction with a sufficiently large field in the space-charge region to cause impact ionization (Figure 1-6). APDs detect photons in a four step process: 1) The photon is absorbed by the diode, resulting in the generation of a hole-electron pair. The depth at which the photon is absorbed is a strong function of the photon wavelength; short wavelength light (ultra-violet) is absorbed very near the surface of the diode while longer wavelength light is absorbed deeper in the diode. For this discussion it is assumed the photons are absorbed in the undepleted p-region. 2) The majority of the photon-generated electrons (photoelectron) diffuse across the undepleted p-region into the charge region. The photoelectrons which recombine with holes before reaching the space-charged region do not contribute to the APD current, and the photo-generated holes are prevented from entering the space charge region by the polarity of the electric field. 3) If the electric field in the space charge region is large enough, the electron will collide with atoms in the silicon lattice with sufficient force to produce additional hole-electron pairs. This processes, called impact-ionization, is the gain mechanism for the APD. Depending on the magnitude of the field, the photo-electron and its "children" may undergo many ionizing collisions before all of the carriers are swept from the space charge region, giving rise to current gains of more than 1000 in currently available APDs. 4) Finally, the photo-electron and the additional holes and electrons generated via impact-ionization are collected in the undepleted n and p regions at the edges of the space charge region.

Quantum Efficiency

Like the PMT, the multiplication process in an APD is a Poisson process. If the internal noise in the APD is negligible, the signal to noise ratio of the APD output can be no better than

$$SNR = \frac{\eta N}{\sqrt{\eta N}} = \sqrt{\eta N}, \tag{1-7}$$

where N is the number of photons emitted by the scintillator and η is the APD quantum efficiency. As with the PMT, the signal-to-noise ratio is directly proportional to $\sqrt{\eta}$.

Several factors influence η, including the efficacy of the anti-reflective coating on the diode surface, the photon energy and the photon depth-of-interaction. For PET applications the first two factors are not limiting: Most modern APDs have very high quality anti-reflective coatings, and the scintillators commonly used in PET emit ultra-violet light with hν~ 2.5 eV, which is more than twice the minimum energy required to generate an electron-hole pair.

The photon depth-of-interaction does, however, place limitations on the APD quantum efficiency. Assume the APD is illuminated by a monochromatic light source with $h\nu > E_g$ and a photon flux of Φ_0 (photons/cm^{-2}). Upon entering the APD, the photon flux is attenuated by absorption following the expression [8]

$$\Phi(x) = \Phi_o \exp(-\alpha_s x), \tag{1-8}$$

where x (cm) is the position in the APD as shown in Figure 1-6 and α_S (cm^{-1}) is the silicon absorption coefficient which is a strongly decreasing function of the light wavelength. BGO emits light at approximately 480 nm, where α_S(480 nm) = $1.4\text{x}10^4$ cm^{-1}. Thus, from Equation 1-8, 87% of the photons emitted by BGO are absorbed in the first 1.4μm of the APD. This shallow depth-of-interaction for photons emitted by BGO limits the acceptable thickness of the p+ contact layer at the top of the diode because virtually all carriers generated in the p+ region recombine before reaching the undepleted p-region. APDs which have been optimized for ultra-violet have p+ contacts which are only a few tenths of a μm thick, allowing for quantum efficiencies of up to 65% for 480 nm light. While 65% quantum efficiency is much less than the near 100% quantum obtained with APDs for 1μm light, it is none-the-less a significant improvement over the 22% quantum efficiency typically obtained with photomultiplier tubes.

APD Gain

APD gain is generated via impact ionization of both electrons and holes. The parameters α and β are the electron and hole ionization coefficients, respectively, and are defined as the number of ionizations per cm in a semiconductor at a given field and temperature. The following empirical expressions for α and β at room temperature have been developed [9]:

$$\alpha(\varepsilon) = 2300 \cdot \exp\{-6.18[(\tfrac{2\cdot 10^5 V/cm}{\varepsilon}) - 1]\} \quad cm^{-1}, \tag{1-9}$$

$$\beta(\varepsilon) = 13 \cdot \exp\{-13.2[(\tfrac{2\cdot 10^5 V/cm}{\varepsilon}) - 1]\} \quad cm^{-1}. \tag{1-10}$$

From these expressions it is clear the ionization coefficients are strongly dependent on the electric field. The electric field in an APD varies as a function of position, so it is convenient to express the ionization coefficients as position dependent terms $\alpha(x)$ and $\beta(x)$ for a given device at a given bias voltage. The gain of an APD may be calculated by integrating the position dependent ionization coefficients over the trajectories of all carriers generated in the APD through photo-excitation and ionizing collisions [10]:

$$M(x) = 1 + \int_x^L \alpha M(x')dx' + \int_0^x \beta(x')dx \tag{1-11}$$

The 1 in the right-hand side of Equation 1-11 considers the initial hole-electron pair generated at x, the first integral considers all of the "children" pairs generated by the electron and the second integral considers all of the "children" pairs generated by the hole. Equation 1-11 is a general expression which is valid for photons generated anywhere in the APD structure of Figure 1-6. The expression will be developed further in Chapter 2 and will be solved numerically using models for several different APD structures.

Commercially available APDs typically have gains ranging between 10 and 1000, much lower than the gains of 10^6 which are routinely obtained with PMTs. The APD gain deficit must be made up in the signal processing electronics, placing stringent noise and gain requirements on the front-end circuitry.

APD Noise

McIntyre[10] has shown the white noise current power spectral density, $i_n^2/\Delta f$, of an avalanche photodiode is given to very close approximation by the expression

$$\frac{i_n^2}{\Delta f} = 2qI_{inb}M^2(0)\left[k_{eff}M(0) + 2(1-k_{eff}) - \frac{1-k_{eff}}{M(0)}\right] + 2qI_{ds} \tag{1-12}$$

where $M(0)$ is the gain seen by an electron-hole pair generated at x=0, k_{eff} is an effective weighted average of the ratio of the hole and electron ionization rates, q is the fundamental electron charge, I_{inb} is the sum of the signal current and the bulk dark current (i.e. the dark current subject to internal multiplication), and I_{ds} is the surface dark current (i.e. the dark current which is not subject to internal multiplication). Equation 1-12 is frequently rewritten as

$$\frac{\overline{i_n^2}}{\Delta f} = 2q\left(I_{inb}M(0)^2 F + I_{ds}\right), \tag{1-13}$$

where

$$F = k_{eff}M(0) + 2(1-k_{eff}) - \frac{1-k_{eff}}{M(0)} \tag{1-14}$$

is called the *excess noise factor* because it describes the amount by which the mean squared current noise is increased above normal shot noise. The term $2qI_{inb}F$ in Equation 1-13 is the input equivalent noise current of the APD and $M^2(0)$ is the gain-squared transfer function. The surface dark current is a parasitic current flowing outside the APD.

The excess noise in an APD is associated with the statistical nature of the multiplication process. The dependence on the ratio of hole and electron ionization rates may be intuitively understood by considering a packet of carriers traversing the APD structure of Figure 1-6. If a photon is absorbed near the top of the APD, the photogenerated electron will be swept across the high field region near x_j where it will undergo avalanche multiplication, and the hole will be swept back to the APD contact without undergoing multiplication. Each time the electron undergoes an ionizing collision, the newly generated electrons will be swept toward the bottom contact and the holes will be swept toward the top contact. If $k_{eff}=0$, only electrons will be multiplied and the carrier packet will only pass through the APD structure once. In this case, $F=2$, indicating only one new Poisson process has been imposed on the signal. If, however, $k_{eff} > 0$, holes will also undergo multiplication. In this case, some of the holes generated in the avalanche processes will undergo additional avalanche processes as they are swept toward the top contact. The electrons created in the hole avalanche processes will then be swept back toward the bottom contact where they will again be multiplied, creating new hole-electron pairs. Thus, a positive feedback loop is created which increases the number of avalanche processes occurring as a signal propagates through the APD, thus increasing both the gain and the excess noise factor of the APD. As $k_{eff}\rightarrow 1$, the multiplication process becomes unstable and avalanche breakdown occurs. The expression for the noise spectral density of an arbitrary APD structure will be fully developed in Chapter 2.

The Avalanche Photodiode Array

Avalanche Photodiode Arrays have been fabricated from the basic APD structure by segmenting the back side contacts and undepleted n-regions (Figure 1-7). This segmentation has been accomplished by etching or sawing grooves in the back side of an APD [11,12,13] or by diffusing a p-type grid into the back side of an APD [14,15]. In both cases, no topside processing of the APD is required since conduction in the depletion region is virtually all perpendicular to the plane of the APD due to the very high fields in the depletion region. Conduction in the drift region may have a lateral diffusion component, but it is not sufficient to provide significant cross-talk between the devices. Recent measurements on an APD array have shown performance similar to that of discrete APDs with cross-talk between adjacent devices on the order of 1% [12,13].

Review of the Literature

Avalanche multiplication is a well understood phenomenon which is observed to some extent in virtually all p-n junctions operating at high reverse-bias voltages [8]. It is most frequently observed as a breakdown phenomenon, such as in "Zener" diodes operating above about 7 volts. (Although these reference diodes are generally called Zener diodes, the break down mechanism is predominately avalanche multiplication rather than Zener tunneling for diodes designed to break down above about 7 volts.[8]) Avalanche multiplication is also exploited in Impact Ionization Avalanche Transit Time (IMPATT) diodes to develop large gains at microwave frequencies[8]. and is observed as a parasitic substrate current in short-channel MOSFETs [8].

The earliest observation of controlled avalanche gain was reported by McKay and McAfee in 1953[16]. These researches observed internal current gain in p-n junctions biased at large reverse voltages and illuminated with white light or α-particles. In 1964 Huth expanded upon the work of McKay and McAfee to develop the first beveled-edge APD for nuclear particle detection [17]. The beveled edge APD, shown schematically in Figure 1-8a, is one of the two major APD structures in use today.

In 1965 Johnson independently reported signal amplification in PIN photodetectors biased in the avalanche region [18,19]; Johnson was apparently unaware of the work of McKay and McAfee, claiming his measurements as the first of their kind. Following Johnson's work, Ruegg invented the other APD structure in wide use today, the reach-through APD [20] (Figure 1-8b) and developed an expression for calculating the gain of an APD. In 1967 McIntyre further developed the theory for calculating APD gain and developed an expression for calculating the internal noise of an APD [10]. The beveled-edge APD community, which initially focused on nuclear particle detection, and the reach-through APD community, which initially focused on optical communication, worked independently and did not cite the work of one another until the mid 1970s. Interestingly, the beveled-edge and reach-through structures have been modified very little since their introduction.

The Beveled-Edge APD

The beveled edge diode is the simpler of the two structures [9,11-13,17,21-39]. It is fabricated by first performing a deep Gallium (p-type) diffusion into a lightly doped n-type silicon wafer (typically 50 Ω-cm). After the diffusion, the top-side of the wafer is lapped until the metallurgical junction is approximately 20-30μm from the wafer surface. Thin p+ and n+ layers are added to the top and bottom of the structure to provide for electrical contacts. The edges of the diode are beveled to angles of 6-10° relative to the surface plane to reduce the electric field along the sides of the structure, and metallic contacts are deposited on the edge of the top surface and across the entire bottom surface. The earliest diodes had surface areas on the order of 20 mm^2, and with current processing techniques beveled-edge APDs with surface areas in excess of 1500 mm^2 have been fabricated [28].

The quality of beveled edge APDs was improved by the introduction of neutron transmutation doped (NTD) wafers [26]. NTD wafers are fabricated by irradiating slightly p-type silicon with high energy neutrons. When a silicon nucleus absorbs a neutron, the excess neutron decays to a proton following reaction:

$$^{30}Si\,(n,\gamma)^{31}Si \underset{2.6h}{\overset{\beta^-}{\Rightarrow}} {}^{31}P\,. \tag{1-15}$$

The daughter species is an n-type donor, and because the probability of neutron absorption is very uniform across the silicon, the resultant n-type doping profile is far more uniform than could be obtained with traditional doping techniques. The availability of highly uniform n-type starting material made possible the development of beveled edge APDs with surface areas similar to those of PMTs.

Initially the beveled edge diode was considered as a direct detector for alpha particles, beta particles, low-energy protons and low-energy x-rays as well as for .88 and 1.06 μm laser pulses. Beveled-edge APDs were found to offer improved signal-to-noise ratios over unity gain semiconductor detectors for low energy particles, but did not achieve wide-spread use as direct particle detectors because the internal gain was strongly dependent on the particle depth of interaction. The dependence of the gain on the particle depth of interaction may be understood by considering the electric field profile of a beveled edge APD shown in Figure 1-9. (Figure 1-9 also shows the electric field profile of a reach-through APD, discussed below.) The field is highest at the metallurgical junction, and, due to the exponential dependence of avalanche probabilities to the electric field, virtually all avalanche multiplication occurs within a few μm of the junction. Both electrons and holes are susceptible to impact ionization, but the probability of a hole undergoing impact ionization is much less than that of an electron under normal bias conditions. When an incident particle deposits the majority of its energy in the space charge region above the metallurgical junction, the excited electrons are swept across the metallurgical junction and the excited holes are swept back to the p contact. This is the optimal condition for signal amplification because the electrons have the maximal opportunity for avalanche multiplication. Conversely, if the majority of an incident particle's energy is absorbed below the metallurgical junction, the holes are swept across the metallurgical junction and the electrons will be carried forward to the n contact. Because holes have a much lower probability for avalanche multiplication, the signal receives much less amplification. If the particle deposits its energy both above and below the metallurgical junction, the hole-electron pairs created above the junction are preferentially amplified over those created below the junction. Because the depth of interaction is related to the particle energy, the reach-through APD gain is biased toward low-energy particles.

Beveled-edge APDs were not initially considered as light detectors for scintillators because most scintillators emit light in the near-UV range (300-500 nm). Short wavelength light is absorbed very near the silicon surface, and the early p+ contacts were rather thick. Consequently, most of the scintillator light was collected in the p+ contact where the photo-excited electron-hole pairs quickly recombined before they could diffuse to the avalanche region. In 1983, however, Huth collaborated with a group from Radiation Monitoring Devices, Inc. (RMD) to fabricate new beveled edge diodes with very thin (~0.1μm) p+ contacts and anti-reflective coatings optimized for the near-UV light emitted by scintillators[27,28].

RMD was the first corporation to develop a commercial line of beveled-edge APDs [11,30-36]. The basic APD structure remained the same as that developed by Huth, but the anti-reflective coatings and p+ layer thickness were optimized for several different applications including scintillation detectors [30,33,35], laser pulse detectors [5,10] and low energy x-ray detectors [11,32,34]. The RMD detector line includes discrete silicon APDs ranging in size from 4 mm^2 to 90 mm^2. As low energy x-ray detectors the RMD devices have pulse height resolutions of 11.5% full width at half maximum (FWHM) for 5.9 keV x-rays emitted by ^{55}Fe [34]. RMD has also coupled its

APDs to scintillators and measured the pulse height resolutions of those detectors with 662 keV γ-rays emitted from a ^{137}Cs source. They reported room temperature pulse height resolutions of 6.9%[35], 8.5%[30] and 40%[30] FWHM for APDs coupled to CsI(Tl), NaI(Tl) and BGO, respectively. RMD also reported the first beveled edge APD array which was fabricated by etching grooves in the back side of a large area APD to obtain six 0.5 mm x 0.5 mm pixels [11].

Huth also collaborated with a second corporation, Advanced Photonix, Inc. to develop a commercial line of APDs. Unlike RMD, which has a small manufacturing capability and whose market is primarily in the research community, Advanced Photonix was formed to mass produce APDs for commercial applications. Advanced Photonix developed a new manufacturing technique to improve the yield of the mechanical grinding process required to form the beveled edge [37]. Using this new process, discrete APDs with areas up to 200 mm^2 are manufactured in volume and packaged for commercial sale [39]. Advanced Photonix reports a room temperature pulse height resolution of 10.8% FWHM for 5.9 keV x-rays emitted by ^{55}Fe [39]. They also report pulse height resolutions of 6.05% and 7.7% for APDs coupled to CsI(Tl) and NaI(Tl), respectively, where the detectors were illuminated with 662 keV γ-rays emitted from a ^{137}Cs source [39]. Additionally, Advanced Photonix has published detailed application notes [9,38] describing theoretical analyses of the noise and time of response characteristics of their APDs.

Advanced Photonix has also pursued the commercial development of APD arrays, reporting on the characteristics of prototype 9 and 64-element arrays [12,13]. In both arrays, the inter-pixel cross-talk is approximately 1% and the pulse height resolutions are similar to those of their discrete APDs.

The Reach-Through APD

Ruegg [20] proposed the reach-through APD [10,14,15,40-61] as an optimized structure for sensitivity to a wide range of light wavelengths and gain stability. Shown in Figure 1-8b, the reach-through APD is fabricated in a lightly doped p-type (π) silicon wafer. The p-n junction across which the avalanche gain is developed is formed on the back side of the APD by first performing a deep p-type diffusion, followed by a shallow n+ diffusion. A p+ contact layer is then diffused into the APD top side and metal contacts are formed on the top and bottom of the structure. Because the metallurgical junction is near the bottom of the structure, the electric field profile of a reach-through APD is much different from that of a beveled edge APD (Figure 1-9). The lightly doped p-region above the junction is fully depleted under normal operating conditions, so a constant electric field (~20,000 V/cm) appears throughout most of the structure. The field peaks at the junction near the bottom of the structure where virtually all of the avalanche multiplication occurs. Ruegg cited the following advantages to the structure:

1. Light absorbed anywhere above the p-n junction sees the same signal amplification. (Since the p-n junction is near the bottom of the structure, the reach-through APD is less sensitive to wavelength, or particle energy, than the beveled-edge structure which has its p-n junction near the top of the structure.)
2. Because the lightly doped p-region above the p-n junction is fully depleted when the APD is biased in its normal operating region, carriers generated in this region travel toward the p-n junction with limiting velocities, thus assuring a fast transient response.
3. Due to the low doping levels and high purity of the material used, recombination in the depletion region may be assumed to be negligible. The collection efficiency for

light with a penetration depth shorter than the thickness of the structure can, therefore, be expected to be close to unity.

4. Because most of the bias voltage falls across the depletion region, the electric field at the p-n junction is relatively insensitive to small changes in the bias voltage.

Because of these advantages, the reach-through structure became the preferred topology for optical communication researchers. A group with RCA Research Laboratories adopted this structure and used it to develop the currently used theory of operation for APDs. (The "McIntyre Theory" will be developed in Chapter 2). In 1970 McIntyre showed experimentally that reach-through APDs could be used in place of PMTs in laser signal detection applications [43] and in 1972 McIntyre [40] and Conradi [41] presented a pair of papers which theoretically and experimentally established the theory for gain distribution in APDs. In 1974 Conradi[42] developed the theory for temperature effects in reach-through APDs and Webb and Jones [44] showed that reach-through APDs could be used as detectors for low energy x-rays and γ-rays. That same year Webb, McIntyre and Conradi[45] published a comprehensive review of the theory of operation for APDs. The publication of this review marked the end of the early development of APD technology and the beginning of RCA's commercialization of the device. RCA, which later became EG&G Optoelectronics Division, developed an extensive line of APD detectors [63] ranging in size from 0.05 mm^2 to 100 mm^2 with structures optimized for x-ray spectroscopy [46], scintillation detection [47,48,49,50] and fiber optic applications [63]. As low energy x-ray detectors the RCA devices have reported pulse height resolutions as low as 10.8% FWHM for 5.9 keV x-rays emitted by ^{55}Fe [46] and as scintillation detectors the RCA devices have reported pulse height resolutions as low as 10.4% FWHM with NaI(Tl) scintillators [47] and 9.5% with small

BGO scintillators [51]. Additionally, RCA developed a reach-through APD array technology in which the individual pixels were isolated by diffusing a grid of p-type isolation regions into the back side of a standard reach-through structure [14,15]. Unlike Advanced Photonix, RCA has not developed APD arrays with pixel sizes greater than 0.3 mm on a side.

In the mid 1980s, a group led by Lecomte from Université de Sherbrooke in Québec, Canada, began collaborating with RCA to develop a BGO-APD detector for use in Positron Emission Tomography [50-59]. The resulting detector module consisted of two discrete 4 mm x 4 mm reach-through APDs coupled to 3 mm x 5 mm x 25 mm BGO crystals. This group has also developed hybrid and printed circuit board-based electronics to process the detector signals in a small PET scanner. Their detectors have typical room temperature pulse height resolutions of 34% FWHM and typical single detector timing resolutions of 10.5 ns [53]. Lecomte and co-workers have reported on a small bore research tomograph built using these detectors. Although the Sherbrookè group's work is very significant in that it marks the first time APDs have been used in PET detectors, their detectors do not compare favorably with currently used PET block detectors. The Sherbrookè detector has poorer pulse height resolution than the CTI/Siemens block detector (typically <25% FWHM for 511 keV γ-rays)[64] and has markedly poorer single detector timing resolution than the CTI/Siemens block detector (typically 3.2 ns FWHM) [65]. Additionally, because the Sherbrookè PET detector is based on discrete APDs and because their electronics are based on hybrid and printed circuit board technologies, their detector and electronics are significantly more expensive to manufacture than a PMT based block detector and electronics. For these reasons, additional research and design work beyond that reported by the Sherbrookè group is required before APD based detectors will be viable alternatives to traditional block detectors.

Problem

Recent work in APD based PET detectors has shown that APDs can be used as light detection devices in PET detectors [50-59]. However, the only APD based detectors which have been used in PET to date are based on discrete APDs and hybrid and printed circuit board based electronics. These APD based detectors do not yet have adequate price/performance characteristics to challenge the block detector in commercial PET systems. One potential means for improving the performance and cost-effectiveness of APD based PET detectors is to use monolithic arrays of APDs instead of the currently used discrete devices and to develop custom integrated signal processing circuitry for the detector arrays. Monolithic arrays of APDs have not previously been used as PET detectors, however, and it will be necessary to theoretically and experimentally evaluate these devices in order to determine whether they can have performance characteristics similar to those of the currently used PMT based block detectors.

The key detector performance parameters in PET are spatial resolution (the accuracy with which the position of a concentration of positron emitting isotopes can be located), efficiency (the percentage of available coincident events which are successfully processed by the scanner), and stability (the ability of the scanner to maintain performance over time and variations in ambient conditions). For good spatial resolution, the detector and its electronics must accurately determine which BGO pixel absorbs the majority of an incident γ-ray's energy and must be able to distinguish between γ-rays which reach the detector with full energy and those which lose a portion of their energy due to scatter in the body. Additionally, the detector must be able to precisely determine the time of arrival of a γ-ray in order for the scanner to accurately identify which γ-rays were emitted in coincidence. For high efficiency, the detector and electronics must be able to process events in the shortest possible time in order minimize the detector "dead time". In order to assess the acceptability of APD based detectors against these scanner requirements, the following issues must be addressed:

1. Which of the currently available APD array structures (beveled-edge or reach-through) is most suitable for use in PET? In particular, which structure has the lowest noise, highest gain and best time of response?
2. Is the signal-to-noise ratio in an APD coupled to BGO and its electronics sufficient to permit accurate discrimination between γ-rays reaching the detector with their full energy and those which lose a portion of their energy through scatter before reaching the detector? In other words, is the energy resolution of an APD based detector comparable to that of a PMT based block detector?
3. Do currently available APDs have sufficiently high bandwidth and sufficiently low noise to provide timing resolution comparable to that of a PMT based block detector when used with BGO?
4. Is it possible to reliably produce monolithic APD arrays with the required performance characteristics and at an acceptable cost to justify using APD arrays in place of photomultiplier tubes in PET detectors.

Scope and Organization of the Dissertation

This dissertation project is part of a broader multi-year detector development program sponsored by the National Institute of Health aimed at designing, fabricating and evaluating an APD-based detector array consisting of 64 BGO pixels, a 64 element monolithic APD array and a 64 channel CMOS preamplifier array. This dissertation project embodies the first phase of the detector development project in which issues 1, 2 and 3 listed above are addressed. The task of addressing issue 4 has been subcontracted to Advanced Photonix, Inc., a prominent avalanche photodiode manufacturer. The tasks completed during this study are summarized below in the order in which the material is presented in the dissertation.

Chapter 2 reviews the APD theory of operation and derives the fundamental expressions for APD gain and noise. A methodology for calculating the gain and noise of an arbitrary APD structure is presented. Expressions describing the APD frequency response are developed including a new formalism for calculating the frequency response of the undepleted region beneath the top APD contact.

Chapter 3 investigates the performance of beveled edge and reach through APDs. A computer modeling methodology based upon the theory developed in Chapter 2 is developed which calculates the expected performance characteristics of the two APD structures from their basic fabrication processes. The model predicts a beveled-edge APD will have superior noise characteristics to a reach-through APD. This finding is verified by testing commercially available beveled-edge and reach-through APDs.

From the results of this study, Advanced Photonix was selected to manufacture a beveled-edge APD array for use in the new PET detector. Chapter 4 presents the array design in light of the constraints imposed by the PET application and develops an equivalent circuit model for calculating the parasitic effects of fabricating multiple APD elements on the same silicon substrate. The expected timing and pulse height resolution of elements in the array are calculated. Advanced Photonix is currently refining their fabrication process to meet the design requirements.

In Chapter 5, the design of a 6 channel prototype CMOS preamplifier array suitable for use with APDs is described and the fabricated circuits are evaluated. The preamplifier topology is believed to be new; it improves the traditional cascode topology by introducing a local feedback loop around the cascode device to significantly improve its input and output impedance. The preamplifier has a measured 10-90% rise-time of 7 ns and a high corner (-3dB) frequency of 40 MHz with an input capacitance of approximately 6 pF due to the package. The equivalent input noise voltage of the preamplifier is calculated to be approximately 1 nV/$\sqrt{\text{Hz}}$ and the gain bandwidth product is calculated to be approximately 600 MHz.

In Chapter 7 the performance of the preamplifier array coupled to discrete BGO/APD detectors is presented. The detector/amplifier system has a pulse height resolution of 12% FWHM for 511 keV gamma rays. This is believed to be the best pulse height resolution reported to date for APDs coupled to BGO crystals with appropriate dimensions for use with PET. This pulse height resolution is superior to the resolution of commercially available block detectors (typically ~20%) and is adequate for the PET application. The measured timing resolution (9 ns FWHM) is poorer than the ~5 ns FWHM resolution predicted by the model and is only marginally adequate for use in a PET scanner. The poorer-than-expected timing resolution is attributed to flaws in the APD manufacturing process. Recent improvements in the fabrication process are expected to yield improved timing resolution [66]. Chapter 6 also presents data obtained with a prototype Advanced Photonix APD array and briefly investigates the potential for using a newly developed scintillator in place of BGO in future APD based detectors.

References

1. G. F. Knoll, *Radiation Detection and Measurement,* 2nd Edition, Wiley, New York, 1989.

2. P. A. Tipler, *Physics,* 2nd Edition, Worth, New York, 1982.

3. M. E. Phelps, E. J. Hoffman, S. C. Huang, and M. M. Ter-Pogossian, "Applications of Annihilation Coincidence Detection to Transaxial Reconstruction Tomography," *J. Nucl. Med.*, **16**, pp. 210-224, 1975.

4. C. W. Williams, M. C. Crabtree, and S. G. Burgiss, "Design and Performance Characteristics of a Positron Emission Axial Tomograph - ECAT II," *IEEE Trans. Nucl. Sci.,* **NS-26**, pp. 619-627, 1979.

5. Siemens Medical Systems and GE Medical Systems together supply more than 90% of the total PET market.

6. M. E. Casey and R. Nutt, "A Multicrystal Two Dimensional BGO Detector System for Positron Emission Tomography," *IEEE Trans. Nucl. Sci.,* **NS-33**, pp. 460-463, 1986.

7. Joel G. Rogers, Ronald Nutt, Mark Andreaco, and C.W. Williams, "Testing 144- and 256-crystal BGO Block *Detectors," Conference Record of the 1993 IEEE Symposium on Nuclear Science*, San Francisco, October, 1993.

8. S. M. Sze, *Physics of Semiconductor Devices,* 2nd Edition, Wiley, New York, 1981.

9. "Noise Characteristics of Advanced Photonix Avalanche Photodiodes," Advanced Photonix Application Note API/NOIS/1291/B, 1991.

10. R.J. McIntyre, "Multiplication Noise in Uniform Avalanche Diodes," *IEEE Trans. Elec. Dev.*, **ED-13**, pp. 164-168, 1966.

11. M.R. Squillante, R. Farrell, J.C. Lund, F. Sinclair, G. Entine and K.R. Keller, "Avalanche Diode Low Energy X-ray and Nuclear Particle Detector," *IEEE Trans. Nucl. Sci.*, **NS-33**, pp. 336-339, 1986.

12. E. Gramsch, S. Zhang, M. Madden, M. Lindberg and M. Szawlowski, "Performance of a 9-Element Avalanche Photodiode Array," *Conference Record of the 1992 IEEE Symposium on Nuclear Science, Orlando*, October, 1992.

13. Ernesto Gramsch, S. Zhang, M. Madden, M. Lindberg and M. Szawlowski, "High Density Avalanche Photodiode Array," *Conference Record of the 1992 IEEE Symposium on Nuclear Science, Orlando*, October, 1992.

14. Robert J. McIntyre and Paul P. Webb, "Avalanche Photodiode Array," U.S. Patent 4,458,260, July 3, 1984.

15. Murray Trakalo, Paul P. Webb, Pierre Poirier, and Robert J. McIntyre, "Avalanche Photodiode Thirty-two-element Linear Array with Minimal Dead Space," *Applied Optics*, **26**, pp. 3694-3599, 1987.

16. K. G. McKay and K. B. McAfee, "Electron Multiplication in Silicon and Germanium," *Physical Review,* **91**, pp. 1079-1084, 1953.

17. G.C. Huth, J.B. Trice and R.A. McKinney, *Rev. Sci. Inst.* **35**, p1220, 1964.

18. K. M. Johnson, "Photodiode Signal Enhancement Effect at Avalanche Breakdown Voltage," *International Solid-State Circuits Conference, Digest of Technical Papers,* p. 64, 1964.

19. K.M. Johnson, "High-Speed Photodiode Signal Enhancement at Avalanche Breakdown Voltage," *IEEE Trans. Elec. Dev.*, **ED-12**, 1965.

20. H.W. Ruegg, "An Optimized Avalanche Photodiode," *IEEE Trans. Elec. Dev.*, **ED-14**, p. 239, 1967.

21. G.C. Huth, J.B. Trice and J.A. Shannon, *IEEE Trans. Nucl. Sci.,* **NS-12**, pp. 275-280, 1965.

22. G.C. Huth, "Recent Results Obtained with High Field Internally Amplifying Semiconductor Radiation Detectors", *IEEE Trans. Nucl. Sci.,* **NS-13**, pp. 36-42, 1966.

23. Robert J. Locker and Gerald C. Huth, "A New Ionizing Radiation Detection Concept Which Employs Semiconductor Avalanche Amplification and the Tunnel Diode Element", *Applied Physics Letters,* **9**, pp. 227-230, 1966.

24. P.A. Johnson, G.C. Huth, and R.J. Locker, "A Solid State X-ray Fluorescence analyzer for use in Hostile Environments", *IEEE Trans. Nucl. Sci.*, **NS-16**, p. 432, 1969.

25. G.C. Huth and Russell A. McKinney, "Radiation Detector," U.S. Patent 3,449,177, June 10, 1969.

26. V.L. Gelezunas, W. Siebt and G. Huth, "Uniform Large-Area High-Gain Silicon Avalanche Radiation Detectors from Transmutation Doped Silicon," *Applied Physics Letters,* **30**, pp. 118-120, 1977.

27. G. Reiff, M.R. Squillante, H.B. Serreze, G. Entine and G.C. Huth, "Large Area Silicon Avalanche Photodiodes: Photomultiplier Tube Alternate," *Material Research Society Proceedings, Nucl. Rad. Det. Symp.,* **16**, pp. 161-170, 1983.

28. G. Entine, G. Reiff, M.R. Squillante, H.B. Serreze, S. Lis, and G. Huth, "Scintillation Detectors Using Large Area Silicon Avalanche Photodiodes," *IEEE Trans. Nucl. Sci.*, **NS-30**, pp. 431-435, 1983.

29. Gerald C. Huth, "Silicon Avalanche Photodiode Array," U.S. Patent 5,021,854, June 4, 1991.

30. M.R. Squillante, G. Reiff, and G. Entine, "Recent Advances In Large Area Avalanche Photodiodes," *IEEE Trans. Nucl. Sci.*, **NS-32**, pp. 563-566, 1985.

31. R. Burnham and D. Scaff, "Detection of Deep-red Low-level Light Pulses," *Applied Optics* **25**, pp. 1514-1518, 1986.

32. K. Keller, R. Farrell, and K. Epps, "The Large Area Avalanche Photodiode as an X-ray Detector," *Proceedings of the Sensors Expo*, pp. 483-487, 1987.

33. R. Farrell, F. Olschner, E. Frederick, L. McConchie, K. Vanderpuye, M.R. Squillante, and G. Entine, "Large Area Silicon Avalanche Photodiodes for Scintillation Detectors," *Nuclear Instruments and Methods in Physics Research*, **A288**, pp. 137-139, 1990.

34. R. Farrell, K. Vanderpuye, G. Entine and M.R. Squillante, "High Resolution, Low Energy Avalanche Photodiode X-ray Detectors," *IEEE Trans. Nucl. Sci.*, **NS-38**, pp. 144-147, 1991.

35. K.M. James, M.J. Masterson, and R. Farrell, "Performance Evaluation of New Large-area Avalanche Photodiodes for Scintillation Spectroscopy," *Nuclear Instruments and Methods in Physics Research*, **A313**, pp. 196-202, 1992.

36. F. Martin, R. Farrell, and G. Entine, "Measurements of the Operating Characteristics of a Large-area Avalanche Photodiode," *Optical Engineering*, **31**, pp. 48-52, 1992.

37. Jan S. Iwanczyk, "Light Response Avalanche Diode," U.S Patent 5,057,892, October 15, 1991.

38. "Response Time Characteristics of Advanced Photonix Avalanche Photodiodes," Advanced Photonix Application Note API/TIME/0792/A, 1992.

39. M.J. Szawlowski, S. Zhang, A. DeCecco, M. Madden, M. Lindberg and E. Gramsch, "Performance of Large Area Avalanche Photodiodes," *Conference Record of the 1992 IEEE Symposium on Nuclear Science, Orlando*, October, 1992.

40. Robert J. McIntyre, "The Distribution of Gains in Uniformly Multiplying Avalanche Photodiodes: Theory," *IEEE Trans. Elec. Dev.*, **ED-19**, pp. 703-713, 1972.

41. Jan Conradi, "The Distribution of Gains in Uniformly Multiplying Avalanche Photodiodes: Experimental," *IEEE Trans. Elec. Dev.*, **ED-19**, pp. 713-717, 1972.

42. J. Conradi, "Temperature Effects in Silicon Avalanche Diodes," *Solid State Electronics*, **17**, pp. 99-106, 1974.

43. R. J. McIntyre, "Comparison of Photomultipliers and Avalanche Photodiodes for Laser Applications," *IEEE Trans. Elec. Dev.*, **ED-17**, 1970.

44. P. Webb and A.R. Jones "Large Area Reach-Through Avalanche Diodes for Radiation Monitoring," *IEEE Trans. Nucl. Sci.*, **NS-21**, pp. 151-157, 1974.

45. P.P. Webb, R.J. McIntyre, and J. Conradi, "Properties of Avalanche Photodiodes," *RCA Review*, **35**, 1974.

46. P.P. Webb and R.J. McIntyre, "Large Area Reach-Through Avalanche Diodes for X-Ray Spectroscopy, *IEEE Trans. Nucl. Sci.*, **NS-23**, pp. 138-144, 1976.

47. G. A. Petrillo, R.J. McIntyre, R. Lecomte, G. Lamoureux, and D. Schmitt, "Scintillation Detection with Large-Area Reach-Through Avalanche Photodiodes," *IEEE Trans. Nucl. Sci.*, **NS-31**, pp. 417-423, 1984.

48. Robert J. McIntyre and Paul P. Webb, "Silicon Avalanche Photodiode with Low K_{eff}," U.S. Patent 4,463,368, July 31, 1984.

49. Alexander W. Lightstone, Paul P. Webb and Robert J. McIntyre, "Avalanche Photodiode," U.S Patent 4,654,678, March 31, 1987.

50. A.W. Lightstone , R.J. McIntyre, R. Lecomte, and D. Schmitt, "A Bismuth Germanate-Avalanche Photodiode Module Designed for use in High Resolution Positron Emission Tomography," *IEEE Trans. Nucl. Sci.*, **NS-33**, pp. 456-459, 1986.

51. R. Lecomte, D. Schmitt, A.W. Lightstone and R.J. McIntyre, "Performance Characteristics of BGO-Silicon Avalanche Photodiode Detectors for PET," *IEEE Trans. Nucl. Sci.*, **NS-32**, pp. 482-486, 1985.

52. C. Carrier, C. Martel, D. Schmitt and R. Lecomte, "Design of a High Resolution Positron Emission Tomograph Using Solid State Scintillation Detectors," *IEEE Trans. Nucl. Sci.*, **NS-35**, pp. 685-690, 1988.

53. Roger Lecomte, Charles Martel and Christian Carrier, "Status of BGO-Avalanche Photodiode Detectors for Spectroscopic and Timing Measurements," *Nucl Instr. and Meth.*, **A278**, pp. 585-597, 1989.

54. C. Carrier and R. Lecomte, "Recent Results in Scintillation Detection with Silicon Avalanche Photodiodes," *IEEE Trans. Nucl. Sci.*, **NS-37**, pp. 209-214, 1990.

55. R. Lecomte, J. Cadorette, A. Jouan, M. Heon, D. Rouleau and G. Gauthier, "High Resolution Positron Emission Tomography with a Prototype Camera Based on Solid State Scintillation Detectors," *IEEE Trans. Nucl. Sci.*, **NS-37**, pp. 805-811, 1990.

56. M. Heon, C. Carrier, J. Cadorette, P. Richard, D. Rouleau, S. Rodrigue and R. Lecomte, "A New Stationary Sampling Scheme for Multi-Layer Positron Tomographs," *IEEE Trans. Med. Imag.*, **11**, pp. 1628-1634, 1992.

57. J. Cadorette, S. Rodrigue and R. Lecomte, "Tuning of Avalanche Photodiode PET Camera", *Conference Record of the 1992 IEEE Symposium on Nuclear Science, Orlando*, October, 1992.

58. R. Lecomte, M. Bentourkia, P.Msaki, J. Cadorette and M. Heon, "Potentials of Multispectral Acquisition in Positron Emission Tomography," *Conference Record of the 1992 IEEE Symposium on Nuclear Science, Orlando*, October, 1992.

59. R. Lecomte, J. Cadorette, S. Rodrigue, M. Heon, D. Rouleau, P. Richard, M. Bentourkia and P. Msaki, "A PET Camera Simulator with Multispectral Acquisition Capabilities," *Conference Record of the 1992 IEEE Symposium on Nuclear Science, Orlando*, October, 1992.

60. A.R. Jones, "A Portable Gamma-Ray Survey Meter Using Silicon Amplifying Radiation Detectors," *IEEE Trans. Nucl. Sci.*, **NS-26**, p. 425, 1969.

61. A.R. Jones, "Comparison of Two Types of Amplifying Radiation Detector for use in Portable Gamma Radiation Monitors," *IEEE Trans. Nucl. Sci.*, **NS-20**, pp. 528-533, 1973.

62. Raymond R. Raylman and Gary D. Hutchins, "Use of Plastic Scintillator with Large Area Avalanche Photodiodes," *Conference Record of the 1990 IEEE Symposium on Nuclear Science*, Arlington, October, 1990.

63. EG&G Optoelectronics Catalog, EG&G Optoelectronics Canada, 23832 Rockfield Blvd., Suite 235, Lake Forest, CA 92630-2896.

64. J. G. Rogers, A. J. Taylor, M. F. Rahimi, R. Nutt, M. Andreaco, C. W. Williams, "An Improved Multicrystal 2-D BGO Detector for PET," *Conference Record of the 1991 IEEE Symposium on Nuclear Science*, Santa Fe, October, 1991.

65. D.M. Binkley, Michael J. Paulus, Michael E. Casey, Ronald Nutt, Wilfried Loeffler, J. Clif Moyers and James M. Rochelle, "A Custom CMOS Integrated Circuit for PET Tomograph Front-End Applications," *Conference Record of the 1993 IEEE Nuclear Science Symposium and Medical Imaging Conference.*

66. Joe Boitsvert, Advanced Photonixs Inc, private communication.

Figures

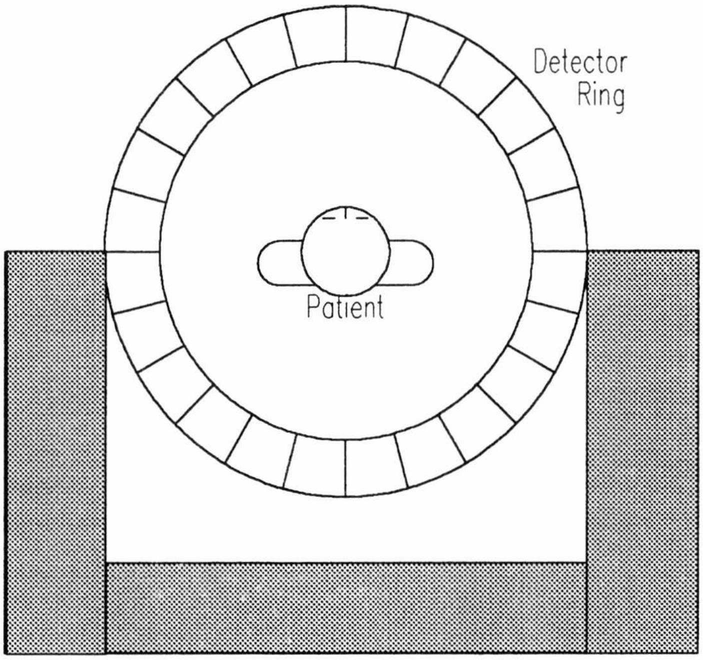

Figure 1-1. Schematic diagram of a PET scanner.

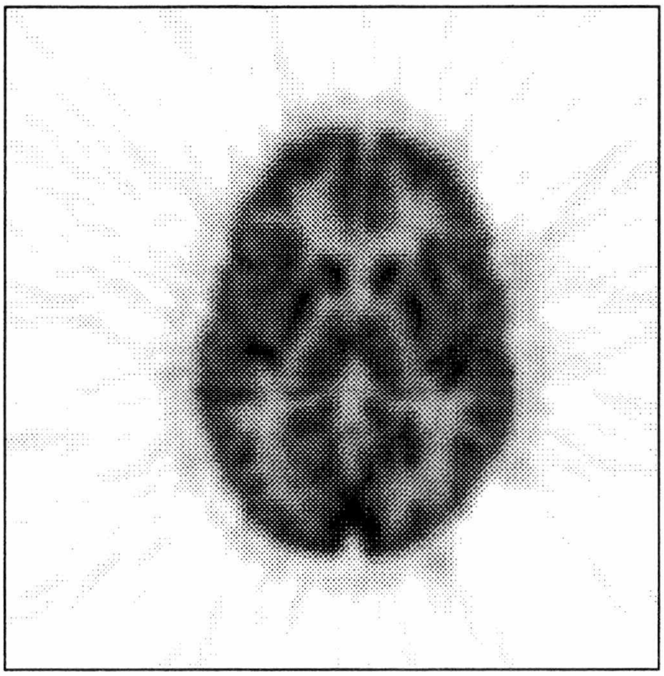

Figure 1-2. Typical PET scan of the human brain using FDG as a tracer compound.

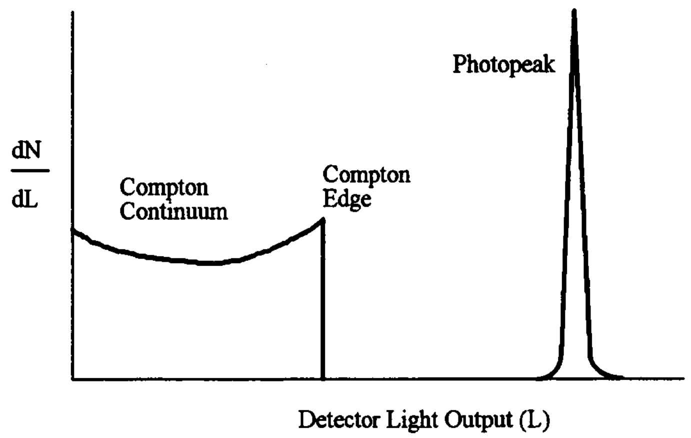

Figure 1-3. Ideal scintillation detector response to a monoenergetic beam of γ-rays.

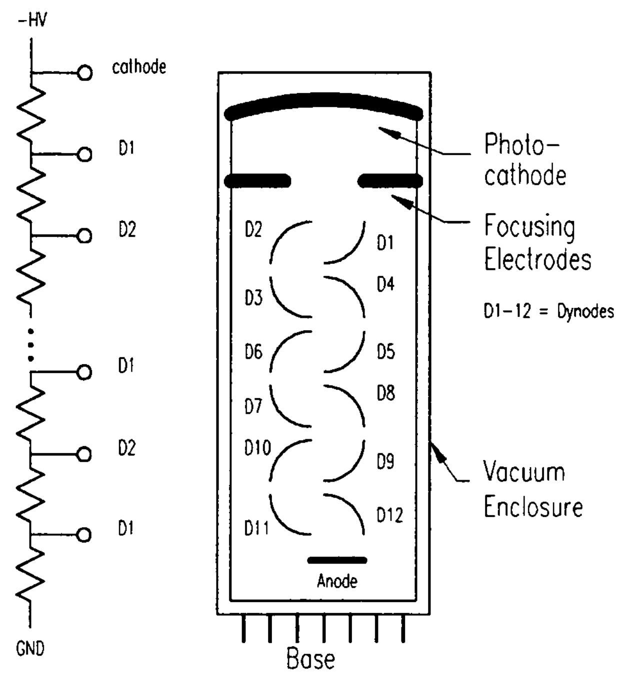

Figure 1-4. Schematic diagram of a photomultiplier tube and its accompanying bleeder network.

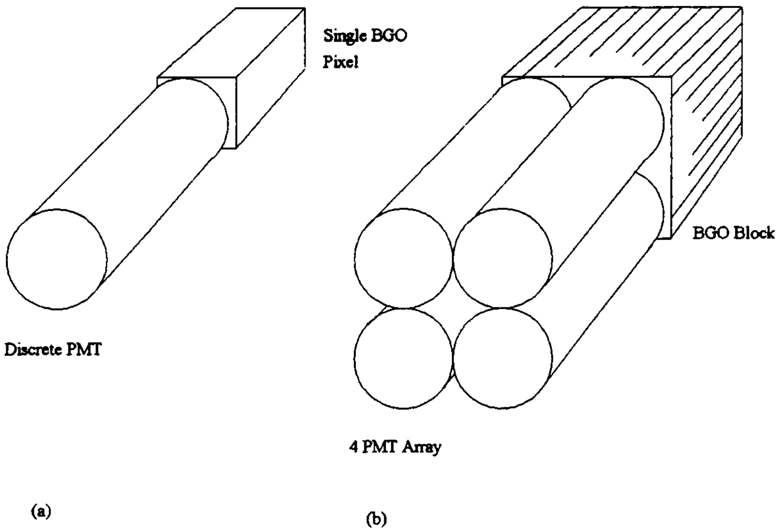

Figure 1-5. Schematic drawing of a) the single PMT-single BGO crystal detector and b) the CTI/Siemens block detector.

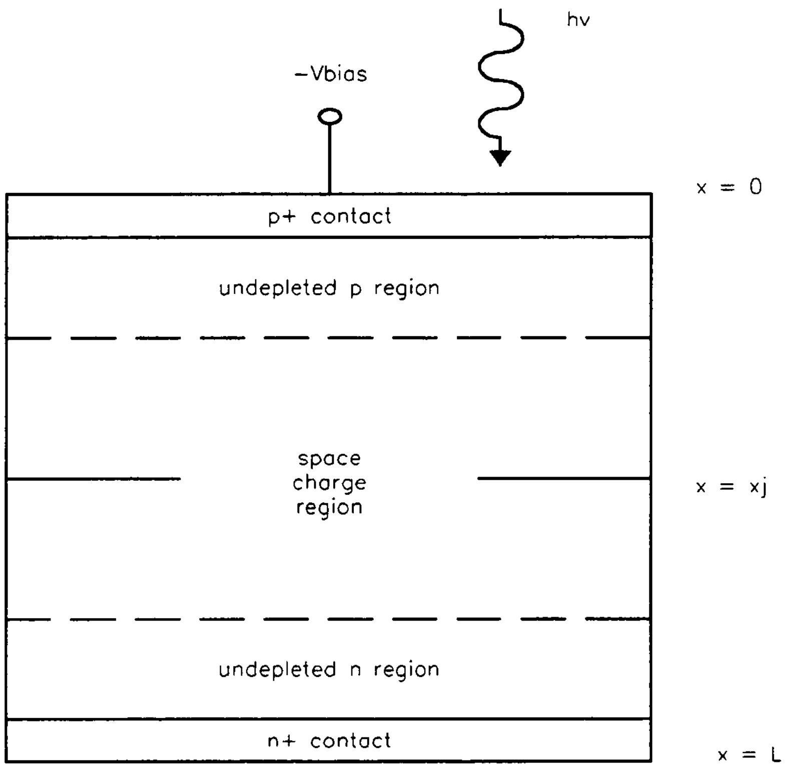

Figure 1-6. The basic avalanche photodiode structure.

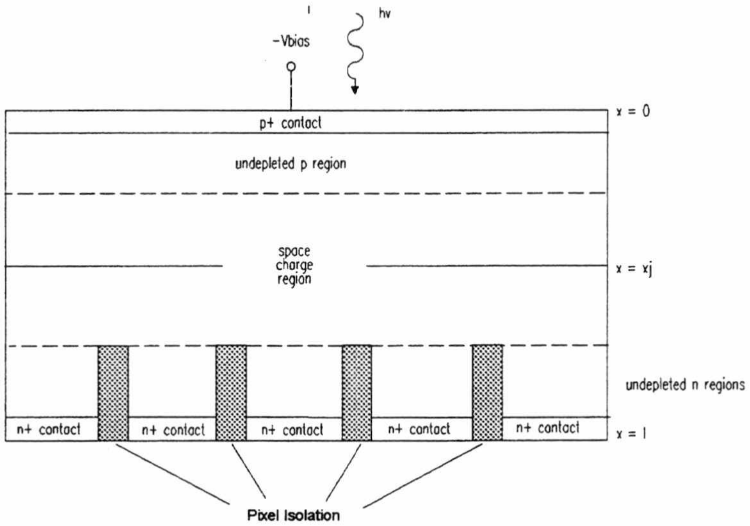

Figure 1-7. The basic avalanche photodiode array structure.

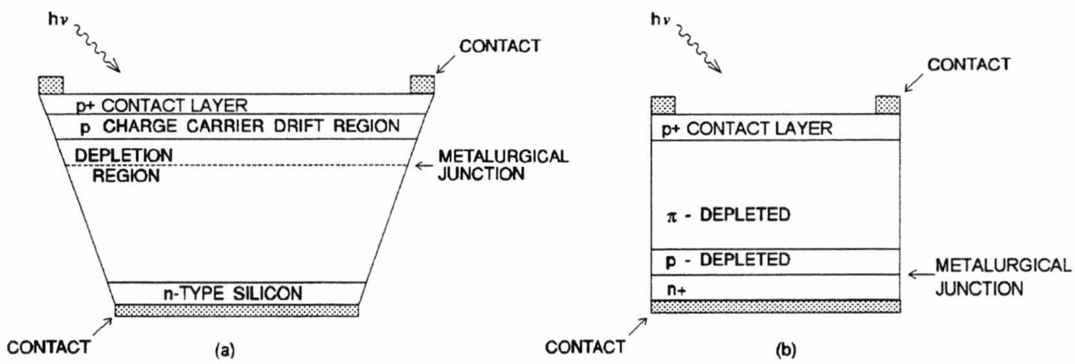

Figure 1-8. Comparison of beveled-edge and reach-through APD structures.

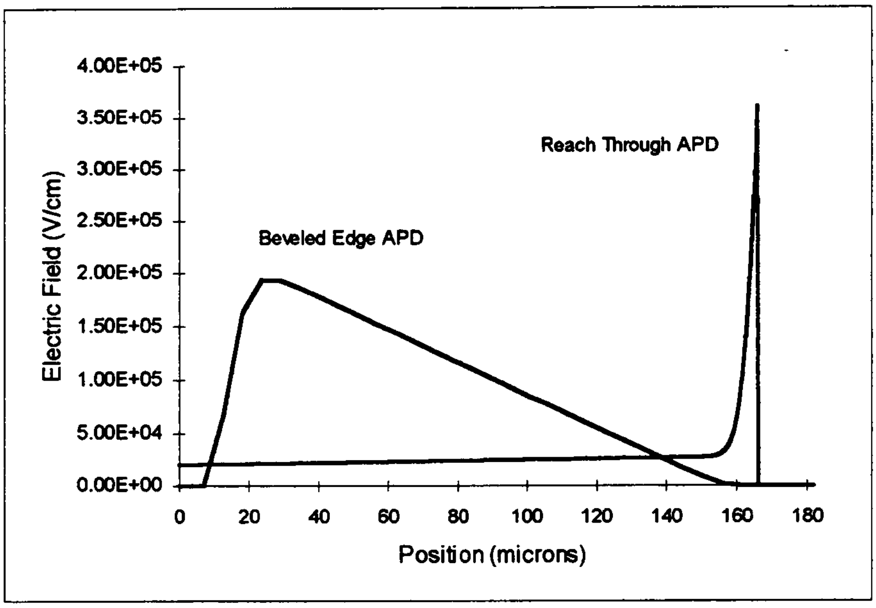

Figure 1-9. Comparison of the electric field profiles of beveled-edge and reach-through APDs.

PART 2

DEVELOPMENT OF AVALANCHE PHOTODIODE THEORY

Overview

This chapter develops the basic theory of operation for Avalanche Photodiodes. The general expressions for APD gain and noise are derived and solved numerically for an ideal APD structure. Frequency dependent transfer functions are developed for the reach-through and beveled edge APD. The material developed in this chapter serves as the foundation for the theoretical and experimental evaluation of the two major APD processing technologies presented in Chapter 3.

Current Flow in a Uniform Electric Field

Before considering the current flow in an APD, the simpler case of current flow in a semiconductor biased with a uniform electric field is examined. Consider a metal-semiconductor-metal capacitor of width d as shown in Figure 2-1, where a constant electric field, E, exists in the semiconductor between the plates. An electron injected into the semiconductor from the negative plate will drift toward the positive plate with a constant velocity, $v=\mu E$, where μ is the electron mobility in the semiconductor. (The diffusion current component is assumed to be negligibly small). The drifting electron causes a current, $I(t)$, to flow in the capacitor bias circuitry (Figure 2-2), where the current pulse width is equal to the drift time,

$$T = \frac{d}{v}, \tag{2-1}$$

and the current pulse amplitude is

$$I(t) = \frac{q}{T}, \quad t < T. \tag{2-2}$$

The total charge transferred between the plates is

$$Q_T = \int_0^T I(t)dt = q. \tag{2-3}$$

Now consider the case where an electron-hole pair is generated somewhere between the plates as shown in Figure 2-3. This is analogous to a reverse biased PIN photodetector in which an incident photon generates a carrier pair in the depletion region. The electron drifts to the positive plate while the hole drifts to the negative plate. Assuming the electron and hole have equal and constant drift velocities, the current pulse has a duration

$$T' = \frac{d'}{v}, \tag{2-4}$$

where d' is the distance traveled by the carrier farthest from its destination plate. As shown in Figure 2-4, the current pulse height is

$$I = \frac{q}{T'}, \quad t < T, \tag{2-5}$$

and the charge transferred is

$$Q_T = \int_0^{T'} I(t)dt = q. \tag{2-6}$$

Thus, for each electron-hole pair generated, one fundamental electron charge is transferred between the plates.

Consider next the case where an electron-hole pair is generated at *t=0* somewhere between the plates, and the electron subsequently undergoes avalanche multiplication at some later time *t=t'* (Figure 2-5). This is analogous to an APD illuminated by a single photon. Prior to the multiplication *(t<t')* the current pulse amplitude follows Equation 2-5. Assuming the avalanche multiplication occurs instantaneously and all carriers drift at the same velocity, the current pulse amplitude after multiplication is given by

$$I(t) = \frac{q}{T'} + \frac{Nq}{T''}, \quad t' < t < T, \tag{2-7}$$

where T' is the time at which the last of the initial carriers reaches a plate, N is the number of additional electron-hole pairs generated by avalanche multiplication, and

$$T'' = \frac{d''}{v}, \tag{2-8}$$

where d" is the distance traveled by the multiplied carriers farthest from their destination plate. Assuming the initial carriers reach the plates before the last multiplied carriers, the current after the initial carriers have been collected is given by

$$I(t) = \frac{Nq}{T''}, \quad T' < t < t' + T''. \tag{2-9}$$

The current pulse is shown schematically in Figure 2-6. The total charge transferred between the plates is

$$Q_T = \int_0^{T'} I(t)dt = \int_0^{t'} \frac{q}{T'}dt + \int_{t'}^{T'} \left(\frac{q}{T'} + \frac{qN}{T''}\right)dt + \int_{T'}^{t'+T''} \frac{qN}{T''} = q(N+1). \tag{2-10}$$

The gain of the structure, M, is defined as

$$M = \frac{\#\ initial\ carriers}{\#\ collected\ carriers} = N + 1. \tag{2-11}$$

Several conclusions about current flow in an APD may be drawn from these examples. First, when a charge impulse is injected into a depletion region, a current pulse will flow through the device whose width and amplitude are determined by the drift time through the depletion region. Second, in most APD structures a delay exists between the time charge is injected into the structure and multiplication occurs. This delay is an important parameter in the APD frequency response. Finally, the gain of a multiplying structure is defined as the number of carriers collected at the terminals divided by the initial number of carriers.

Gain

Consider next the carrier transport in the arbitrary one dimensional APD structure shown in Figure 2-7. The p-n junction is reversed biased so that an electric field, $E(x) \geq 0$, exists between the APD terminals such that a free electron will drift toward x=0 while a free hole will drift toward $x=w$. A free carrier drifting across the

APD has a finite probability of colliding with an atomic nucleus in the crystal structure with sufficient energy to create a new electron-hole pair. Such a collision is called *impact ionization*, and the number of electron-hole pairs generated by a free carrier per unit distance is called the *ionization coefficient.* Field dependent electron and hole ionization coefficients, $\alpha(E)$ and $\beta(E)$, are shown in Figure 2-8 [1], and empirical expressions are given in Equations 2-12 and 2-13 [2].

$$(E) = 2300 \cdot \exp\left\{-6.18 \cdot \left[\left(\frac{2 \cdot 10^5 V / cm}{E}\right) - 1\right]\right\} \quad cm^{-1}. \tag{2-12}$$

$$\beta(E) = 13 \cdot \exp\left\{-613.2 \cdot \left[\left(\frac{2 \cdot 10^5 V / cm}{E}\right) - 1\right]\right\} \quad cm^{-1}. \tag{2-13}$$

The first step in determining the gain of the arbitrary APD in Figure 2-7 is to calculate the electric field in the structure by solving the one dimensional Poisson equation,

$$\frac{dE(x)}{dx} = \frac{\rho(x)}{\varepsilon(x)}, \tag{2-14}$$

where $\rho(x)$ is the net charge present at x and $\varepsilon(x)$ is the position dependent dielectric constant. Position dependent ionization coefficients, $\alpha(x)$ and $\beta(x)$ are then obtained by substituting $E(x)$ into Equations 2-12 and 2-13. The position dependent ionization coefficients implicitly assume that any electron (hole) at x undergoes $\alpha(x)$ ($\beta(x)$) ionizing collisions per unit distance regardless of the path it takes before reaching x. This implies any carrier at x has reached its maximum drift velocity. This assumption is not exactly valid for an electron-hole pair created at x because any newly generated carrier must travel some distance before reaching maximum velocity. However, the error is small for most practical APD structures because the time required for a newly generated carrier to reach maximum drift velocity is small relative to the mean time between ionizing collisions [3].

When a photon is absorbed at x in the general APD (Figure 2-7), the generated electron and hole drift toward $x=0$ and $x=w$, respectively where they undergo $\alpha(x)$ and $\beta(x)$ ionizing collisions per unit distance, respectively. Any new electron-hole pairs generated at x' will also undergo ionizing collisions, so the gain is the sum of the initial carriers plus the carriers generated by the initial electron-hole pair and the carriers generated by all subsequent ionizing collisions. This is described by the transcendental gain equation [4,5]

$$M(x) = 1 + \int_0^x \alpha(x')M(x')dx' + \int_x^w M(x')\beta(x')dx'. \tag{2-15}$$

The 1 in the right hand side of Equation 2-15 describes the initial electron-hole pair generated by the photon, the first integral describes the electron-hole pairs generated by electron impact ionizations and the second integral describes the electron-hole pairs generated by hole impact ionizations. Because $M(x)$ is a function of the gains of all electron-hole pairs generated during the avalanche process, Equation 2-15 cannot be solved directly. However, it can be manipulated to obtain a closed form expression as follows. First take the derivative of both sides of Equation 2-15 to obtain

$$\frac{dM(x)}{dx} = \alpha(x)M(x) - \beta(x)M(x), \tag{2-16}$$

or [4]

$$\frac{dM(x)}{dx} = [\alpha(x) - \beta(x)] \cdot M(x). \tag{2-17}$$

If $M(x)$ is assumed to be of the form [4]

$$M(x) = A \cdot \exp(u(x)), \tag{2-18}$$

then Equation 2-17 may be rewritten as

$$A \cdot \exp(u(x)) \cdot \frac{d(u(x))}{dx} = [\alpha(x) - \beta(x)] \cdot A \cdot \exp(u(x)), \tag{2-19}$$

or

$$\frac{d(u(x))}{dx} = [\alpha(x) - \beta(x)]. \tag{2-20}$$

The following expressions satisfy Equation 2-20:

$$u(x) = \int_0^x [\alpha(x') - \beta(x')]dx, \tag{2-21}$$

and

$$u(x) = -\int_x^w [\alpha(x') - \beta(x')]dx. \tag{2-22}$$

Substituting Equation 2-21 into Equation 2-18,

$$M(x) = A \cdot \exp(\int_0^x [\alpha(x') - \beta(x')]dx). \tag{2-23}$$

Evaluating Equation 2-23 at $x=0$ yields

$$M(0) = A, \tag{2-24}$$

and substituting Equation 2-24 back into Equation 2-23 yields

$$M(x) = M(0) \cdot \exp(\int_0^x [\alpha(x') - \beta(x')]dx). \tag{2-25}$$

Similarly, substituting Equation 2-22 into Equation 2-17 and evaluating at $x=w$ yields

$$M(x) = M(w) \cdot \exp(-\int_x^w [\alpha(x') - \beta(x')]dx). \tag{2-26}$$

Equations 2-25 and 2-26 are equally valid expressions for the gain of an arbitrary APD structure. Equation 2-26 is selected in order to obtain the expressions developed in [4]. Substituting Equation 2-26 into Equation 2-15 yields

$$M(w) \cdot \exp(-\int_x^w [\alpha(x') - \beta(x')]dx) = 1 + \int_0^x \alpha(x')M(x')dx' + \int_x^w M(x')\beta(x')dx. \tag{2-27}$$

Evaluating Equation 2-27 at $x=w$ yields

$$M(w) = 1 + \int_0^w \alpha(x) M(x) dx, \tag{2-28}$$

or

$$M(w) = 1 + M(w) \cdot \int_0^w \alpha(x) \cdot \exp(-\int_{x'}^w [\alpha(x) - \beta(x)] dx') dx, \tag{2-29}$$

which reduces to

$$\frac{1}{M(w)} = 1 - \int_0^w \alpha(x) \cdot \exp(-\int_{x'}^w [\alpha(x) - \beta(x)] dx') dx. \tag{2-30}$$

Recalling from Equation 2-26

$$M(w) = \frac{M(x)}{\exp(-\int_x^w [\alpha(x') - \beta(x')] dx)}, \tag{2-31}$$

and substituting into Equation 2-30 yields

$$\frac{\exp(-\int_x^w [\alpha(x') - \beta(x')] dx)}{M(x)} = 1 - \int_0^w \alpha(x) \cdot \exp(-\int_{x'}^w [\alpha(x) - \beta(x)] dx') dx \tag{2-32}$$

or

$$M(x) = \frac{\exp(-\int_x^w [\alpha(x') - \beta(x')] dx)}{1 - \int_0^w \alpha(x) \cdot \exp(-\int_{x'}^w [\alpha(x) - \beta(x)] dx') dx}. \tag{2-33}$$

Because the numerator of Equation 2-33 appears frequently in APD gain and noise calculations, it is useful to define it as a separate variable,

$$G(x) = \exp(-\int_x^w [\alpha(x') - \beta(x')] dx), \tag{2-34}$$

or

$$G(x) = \frac{\exp \int_x^w \beta(x')dx}{\exp \int_x^w \alpha(x')dx}. \tag{2-35}$$

As shown below, $G(x)$ may be considered an approximate expression for the APD gain normalized to $G_{max}=1$. Equation 2-33 is now written as

$$M(x) = \frac{G(x)}{1 - \int_0^w \alpha(x) \cdot G(x)dx}. \tag{2-36}$$

The next step in the derivation is to simplify the denominator of Equation 2-36. Let k_1 be defined as a weighted ratio of average ionization coefficients given by

$$k_1 = \frac{\int_0^w \beta(x)G(x)dx}{\int_0^w \alpha(x)G(x)dx}. \tag{2-37}$$

Multiplying the numerator and denominator of Equation 2-36 by $1-k_1$ yields

$$M(x) = \frac{G(x)}{1 - \int_0^w \alpha(x')G(x')dx'} \cdot \frac{1-k_1}{1-k_1} = \frac{G(x)\cdot(1-k_1)}{D}. \tag{2-38}$$

Now consider the denominator of Equation 2-38, D, separately.

$$D = 1 - \int_0^w \alpha(x')G(x')dx' - k_1 + k_1 \int_0^w \alpha(x')G(x')dx. \tag{2-39}$$

$$D = 1 + \int_0^w \beta(x')G(x')dx' - \int_0^w \alpha(x')G(x')dx' - k_1. \tag{2-40}$$

Combining terms and substituting Equation 2-34 for $G(x')$ yields

$$D = 1 + \int_0^w [\beta(x') - \alpha(x')]dx' \cdot \exp\left\{ -\int_{x'}^w [\alpha(x'') - \beta(x'')]dx'' \right\} dx' - k_1, \tag{2-41}$$

which is evaluated to obtain

$$D = \exp - \int_0^w [\alpha(x') - \beta(x')] dx' - k_1, \tag{2-42}$$

or

$$D = G(0) - k_1. \tag{2-43}$$

Substituting Equation 2-43 into the denominator of Equation 2-38 yields the frequency-independent, or low-frequency expression for the gain of an APD.

$$M(x) = \frac{G(x)(1-k_1)}{G(0)-k_1}. \tag{2-44}$$

This is the McIntyre expression for APD gain. For practical APD structures $k_1 \ll 1$, so a reasonable approximation is

$$M(x) \approx \frac{G(x)}{G(0)-k_1}, \tag{2-45}$$

and

$$M_{\max} \approx \frac{1}{G(0)-k_1}. \tag{2-46}$$

Example APD Gain Calculation

As an example, consider the case of an abrupt p-n junction with equal doping levels in the p and n regions where the doping levels and bias voltage are chosen such that the peak electric field = 2×10^5 when the device is fully depleted. Such an ideal electric field profile is shown in Figure 2-9. Equations 2-12 and 2-13 are solved for the electric field profile to obtain the ionization coefficient profiles shown in Figure 2-10. The ionization coefficients are exponentially dependent upon the electric field and have maximum values at the metallurgical junction (x = 0.01 cm). Note the electron ionization coefficient, $\alpha(x)$, is more than two orders of magnitude greater than the hole ionization coefficient, $\beta(x)$, at the junction. However, $\beta(x)$ is a more rapidly increasing function of E so the ratio β/α increases with increasing E. This ratio is important in determining the APD noise in the next section.

The ionization coefficient profile shown in Figure 2-10 is used to calculate $G(x)$ using Equation 2-35 (Figure 2-11). $G(x)$ has a maximum value of 1 at the light entrance (p-contact) side of the device and has a minimum value of 0.0039 at the bottom of the device ($x = 0$). G(x), $\alpha(x)$ and $\beta(x)$ are used to calculate k_l in Equation 2-37; $k_l = 0.0019$. From Equation 2-46, $M_{max} = 1/(0.0039\text{-}0.0019) = 500$. This is verified by solving Equation 2-44 to obtain the gain profile shown in Figure 2-12.

Figure 2-12 provides insight into the expected sensitivity and dark current of the APD. The gain is highest for electron-hole pairs generated on the light-entrance side of the device above the junction where the electron drifts down through the high field region at the junction and the hole drifts up to the top contact. (Recall from Figure 2-8 electrons have much higher ionization coefficients than holes). When an electron-hole pair is generated below the junction, the hole drifts through the high-field region and the gain is much lower. For maximal sensitivity, the junction must be positioned well below the mean absorption depth for incident photons. Thus, if the incident light is red (long wavelength and deeper absorption) the junction must be deep in the structure. If the incident light is blue (short wavelength and shallower absorption) the junction can be closer to the top of the structure.

Of course, along with the photo-generated electron-hole pairs of interest, the APD has unwanted thermally generated electron-hole pairs. These thermally generated carriers are the source of the APD bulk dark current, I_{db}. (The other dark current component, I_{ds}, is due to conduction along the surface of the APD). Although pairs are thermally generated throughout the depletion region, only those pairs generated in the p-region above the junction cause electrons to drift through the high-field (multiplication) region. Thus, to minimize the bulk dark current, the junction must be located as close to the top of the APD as possible. The optimal APD design therefore places the junction deep enough in the structure to achieve the required sensitivity to photo-generated electron-hole pairs but close enough to the top to minimize the dark current due to thermally-generated electron-hole pairs.

Noise Spectral Density

The noise theory summarized here was first developed by McIntyre in 1966 [4]. McIntyre's derivation is first reviewed in order to understand the simplifying assumptions embodied in the noise expressions which are widely used today. The first step in the derivation is recognition that all current flow in a reverse biased APD is due to the creation of electron-hole pairs through optical generation, thermal generation or impact ionization. Each of these pair-producing processes generates shot noise.

Consider again the arbitrary APD structure shown in Figure 2-7. The first step in the noise calculation is to develop an expression for the shot noise generated at an arbitrary point x in the structure. Since the generation processes create both electrons and holes, the generation rate is determined by counting either holes or electrons. In this derivation the holes are counted. The number of pairs generated in the differential slice beginning at x and ending at $x+dx$ is determined from the increase in hole current at $x+dx$ over the hole current at x. This differential increase in hole current, $dI_p(x)$, is expressed as

$$dI_p(x) = \left[\alpha(x)I_n(x) + \beta(x)I_p(x) + q \cdot g(x)\right]dx, \tag{2-47}$$

where $\alpha(x)I_n(x)$ is the charge generation rate per unit length due to impact ionizations caused by electrons, $\beta(x)I_p(x)$ is the charge generation rate per unit length due to impact ionizations caused by holes, q is the fundamental electron charge and $g(x)$ is the sum of the thermal and optical charge generation rates. Because thermal generation, optical generation and impact ionization are each independent random processes which inject discrete charge carriers into a depletion region, the current generated by these processes must support full shot noise, given by

$$\frac{di_n^{\,2}(x)}{\Delta f} = 2q \cdot dI_p(x), \tag{2-48}$$

where $di_n^{\,2}$ is the differential noise current due to the electron-hole generation processes occurring at x. This differential element of noise current experiences gain as described by Equation 2-44. Thus, the noise current appearing at the APD terminals due to the generation processes at x is obtained by multiplying $di_n^{\,2}(x)/\Delta f$ by $M^2(x)$. The total noise spectral density is obtained by integrating over all x to obtain

$$\frac{i_n^{\,2}}{\Delta f} = \int_0^w \frac{di_n^{\,2}(x)}{\Delta f} \cdot M^2(x)dx, \tag{2-49}$$

or

$$\frac{i_n^{\,2}}{\Delta f} = 2q\int_0^w M^2(x)dI_p(x)dx, \tag{2-50}$$

where $i_n^{\,2}(x)/\Delta f$ is the total APD noise spectral density. Equation 2-50 is the basic expression for APD noise spectral density. The remainder of this discussion is devoted to rewriting the expression in a more useful form.

Equation 2-50 is evaluated using the technique of partial integration to obtain

$$\frac{i_n^{\,2}}{\Delta f} = 2q\left[M^2(x)\cdot I_p(x)\right]_0^w - 2q\int_0^w I_p(x)\cdot\frac{d}{dx}\left(M^2(x)\right)dx, \tag{2-51}$$

or

$$\frac{i_n^{\,2}}{\Delta f} = 2qM^2(w)I_p(w) - 2qM^2(0)I_p(0) - 2q\int_0^w I_p(x)\cdot 2M(x)\frac{d}{dx}(M(x))dx. \tag{2-52}$$

Recall the expression for $M(x)$, Equation 2-33, which is given again below:

$$M(x)=\frac{\exp(-\int_x^w[\alpha(x')-\beta(x')]dx)}{1-\int_0^w\alpha(x)\cdot\exp(-\int_{x'}^w[\alpha(x)-\beta(x)]dx')dx}. \tag{2-33}$$

The derivative of *M(x)* is given by

$$\frac{dM(x)}{dx}=\frac{\exp(-\int_x^w[\alpha(x')-\beta(x')]dx)\cdot[\alpha(x)-\beta(x)]}{1-\int_0^w\alpha(x)\cdot\exp(-\int_{x'}^w[\alpha(x)-\beta(x)]dx')dx}=M(x)\cdot[\alpha(x)-\beta(x)]. \tag{2-53}$$

Substituting into Equation 2-52 yields

$$\frac{i_n^2}{\Delta f}=2qM^2(w)I_p(w)-2qM^2(0)I_p(0)-2q\int_0^w I_p(x)\cdot 2M^2(x)\cdot[\alpha(x)-\beta(x)]dx. \tag{2-54}$$

Thus far only the hole current, I_p,, has been considered. The total current in the APD, I, is now defined as the sum of the hole and electron current, I_n. The electron current is then expressed as

$$I_n=I-I_p. \tag{2-55}$$

Substituting Equation 2-55 into Equation 2-36 yields

$$(\alpha(x)-\beta(x))I_p=\alpha(x)I+qg(x)-\frac{dI_p}{dx}. \tag{2-56}$$

Substituting Equation 2-56 into Equation 2-54 yields

$$\frac{i_n^2}{\Delta f}=2q\left[M^2(w)I_p(w)-M^2(0)I_p(0)-2\int_0^w M^2(x)\cdot\left[\alpha(x)I+qg(x)-\frac{dI_p}{dx}\right]dx\right], \tag{2-57}$$

or,

$$\frac{i_n^2}{\Delta f}=2q\left[M^2(w)I_p(w)-M^2(0)I_p(0)-2\int_0^w M^2(x)\cdot[\alpha(x)I+qg(x)]dx+2\int_0^w M^2(x)\cdot\frac{dI_p}{dx}dx\right] \tag{2-58}$$

Substituting Equation 2-50 into Equation 2-58 yields

$$\frac{i_n^2}{\Delta f} = 2q\left[M^2(w)I_p(w) - M^2(0)I_p(0) - 2\int_0^w M^2(x)\cdot\left[\alpha(x)I + qg(x)\right]dx\right] + 2\frac{i_n^2(x)}{\Delta f}, \tag{2-59}$$

or

$$\frac{i_n^2}{\Delta f} = 2q\left[M^2(0)I_p(0) - M^2(w)I_p(w) + 2\int_0^w M^2(x)\cdot\left[\alpha(x)I + qg(x)\right]dx\right]. \tag{2-60}$$

Substituting $I_p(w) = I(w) - I_n(w)$ yields

$$\frac{i_n^2}{\Delta f} = 2q\left[M^2(0)I_p(0) + M^2(w)I_n(w) + 2\int_0^w M^2(x)qg(x)dx + 2I\int_0^w M^2(x)\alpha(x)dx - I(w)M^2(w)\right]. \tag{2-61}$$

The model assumes all current flow is due to optical, thermal or impact-ionization generation process with no current injected at the contacts, or $I_p(0) = I_n(w) = 0$. Following this argument, Equation 2-61 is reduced to

$$\frac{i_n^2}{\Delta f} = 2q\left[2\int_0^w M^2(x)qg(x)dx + 2I\int_0^w M^2(x)\alpha(x)dx - IM^2(w)\right], \tag{2-62}$$

which is the general expression for APD noise current.

Several simplifications are made to obtain the more familiar closed-form expression for APD noise. First assume all incident light is absorbed near $x=w$. Under this condition, only electrons pass through the high field region and undergo avalanche multiplication. This is a reasonable approximation for APDs coupled to blue-light emitting scintillators such as BGO. Second, assume the thermally generated carriers which form the dark current are all generated at $x=0$. With these approximations, the generation current, $qg(x)$, may be treated as an input current injected at $x=w$,

$$\int_0^w qg(x)M^2(x)dx = I_{in}M^2(w), \tag{2-63}$$

Additionally, the current *I(w)* is treated as the input current multiplied by the gain *M(w)*, or

$$I(w) = I_{in}M(w). \tag{2-64}$$

Note that I_{in} is the composite input current to the device representing the sum of the unmultiplied photo-generated current and the unmultiplied dark current. Substituting Equations 2-63 and 2-64 into Equation 2-62 yields

$$\frac{i_n^{\,2}}{\Delta f} = 2q\left[2I_{in}M^2(w) - I_{in}M^3(w) + 2M(w)I_{in}\int_0^w M^2(x)\alpha(x)dx\right], \tag{2-65}$$

or

$$\frac{i_n^{\,2}}{\Delta f} = 2qI_{in}M^2(w)\left[2 - M(w) + \frac{2}{M(w)}\int_0^w M^2(x)\alpha(x)dx\right]. \tag{2-66}$$

Equation 2-66 has the form of normal shot noise multiplied by the quantity in parentheses. The quantity in parentheses represents the factor by which the APD noise exceeds shot noise and is termed the *excess noise factor, F*. The next step is to express this excess noise factor in terms of experimentally measurable parameters.

The Excess Noise Factor

Equation 2-28 is rewritten as

$$M(w) = 1 + \int_0^w [\alpha(x) - \beta(x)]dx + \int_0^w \beta(x)M(x)dx. \tag{2-67}$$

Substituting Equations 2-17 and 2-37 into Equation 2-67 yields

$$M(w) = 1 + \int_0^w \frac{dM(x)}{dx}dx + k_1\int_0^w \alpha(x)M(x)dx. \tag{2-68}$$

Substituting Equation 2-28 into Equation 2-68 yields

$$M(w) = 1 + \int_0^w \frac{dM(x)}{dx} dx + k_1[M(w) - 1], \tag{2-69}$$

and evaluating the integral yields

$$M(w) = 1 + M(w) - M(0) + k_1[M(w) - 1], \tag{2-70}$$

or

$$M(0) = k_1[M(w) - 1] + 1. \tag{2-71}$$

Equation 2-71 is used later in the derivation and also provides a useful second definition of k_1.

Consider now the integral in the general expression for APD noise spectral density,

$$\int_0^w M^2 \alpha(x) dx = \int_0^w [\alpha(x) - \beta(x)] M^2(x) dx + \int_0^w \beta(x) M^2(x) dx. \tag{2-72}$$

By defining a second weighted ratio of average ionization coefficients as

$$k_2 = \frac{\int_0^w \beta(x) M^2(x) dx}{\int_0^w \alpha(x) M^2(x) dx}, \tag{2-73}$$

Equation 2-72 is rewritten as

$$\int_0^w M^2 \alpha(x) dx = \int_0^w [\alpha(x) - \beta(x)] M^2(x) dx + k_2 \int_0^w \alpha(x) M^2(x) dx, \tag{2-74}$$

or

$$\int_0^w M^2 \alpha(x) dx = \frac{1}{(1 - k_2)} \int_0^w [\alpha(x) - \beta(x)] M^2(x) dx. \tag{2-75}$$

Recalling Equations 2-35 and 2-36, Equation 2-75 is written as

$$\int_0^w M^2(x)\alpha(x)dx = \frac{1}{(1-k_2)}\int_0^w [\alpha(x)-\beta(x)] \frac{\exp\left[-2\int_x^w [\alpha(x')-\beta(x')]dx'\right]}{\left[1-\int_0^w [\alpha(x')G(x')dx']\right]^2} dx, \tag{2-76}$$

or

$$\int_0^w M^2(x)\alpha(x)dx = \frac{1}{(1-k_2)\left[1-\int_0^w [\alpha(x')G(x')dx']\right]^2}\int_0^w [\alpha(x)-\beta(x)]\exp\left[-2\int_x^w [\alpha(x')-\beta(x')]dx'\right]dx. \tag{2-77}$$

Integrating the right hand side of Equation 2-77 yields

$$\int_0^w M^2(x)\alpha(x)dx = \frac{\exp\left[-2\int_w^w [\alpha(x')-\beta(x')dx']\right]-\exp\left[-2\int_0^w [\alpha(x')-\beta(x')dx']\right]}{2(1-k_2)\left[1-\int_0^w [\alpha(x')G(x')dx']\right]^2}. \tag{2-78}$$

Recalling the expression for M(x), Equation 2-36, Equation 2-78 is simplified to

$$\int_0^w M^2(x)\alpha(x)dx = \frac{M^2(w)-M^2(0)}{2(1-k_2)}. \tag{2-79}$$

Equation 2-71 is now substituted into Equation 2-79 to obtain

$$\int_0^w M^2(x)\alpha(x)dx = \frac{M^2(w)-\{k_1[M(w)-1]+1\}^2}{2(1-k_2)}. \tag{2-80}$$

After some algebraic manipulation, Equation 2-80 becomes

$$\int_0^w M^2(x)\alpha(x)dx = \frac{(1-k_1^2)}{2(1-k_2)}M^2(w)+\frac{(k_1^2-k_1)M(w)}{1-k_2}+\frac{k_1(2-k_1)-1}{2(1-k_2)}. \tag{2-81}$$

Substituting into Equation 2-66 yields the new expression for APD noise spectral density,

$$\frac{i_n^2}{\Delta f} = 2qI_{in}M^2(w)\left\{2 - M(w) + \frac{2}{M(w)}\left[\frac{(1-k_1^2)}{2(1-k_2)}M^2(w) + \frac{(k_1^2-k_1)M(w)}{1-k_2} + \frac{k_1(2-k_1)-1}{2(1-k_2)}\right]\right\}. \quad (2\text{-}82)$$

After some algebraic manipulation, Equation 2-82 simplifies to

$$\frac{i_n^2}{\Delta f} = 2qI_{in}M^2(w)\left\{\left(\frac{k_2-k_1^2}{1-k_2}\right)M(w) + 2\left[1-\left(\frac{k_1-k_1^2}{1-k_2}\right)\right] - \frac{1}{M(w)}\left[\frac{1-k_1}{1-k_2} - \left(\frac{k_1-k_1^2}{1-k_2}\right)\right]\right\}. \quad (2\text{-}83)$$

This is the noise spectral density presented by McIntyre [5,6] for electron current injected at $x=w$. McIntyre notes that for most practical applications the following simplifications are possible. Let the effective weighted ratio of electron and hole ionization rates be defined as

$$k_{eff} = \frac{k_2 - k_1^2}{1-k_2}. \quad (2\text{-}84)$$

Now, assume $k_1 \approx k_2$, so that

$$k_{eff} = \frac{k_1 - k_1^2}{1-k_2}, \quad (2\text{-}85)$$

and

$$\frac{1-k_1}{1-k_2} \approx 1. \quad (2\text{-}86)$$

With these approximations, Equation 2-83 is presented in its most commonly expressed form:

$$\frac{i_n^2}{\Delta f} = 2qI_{in}M^2(w)\left\{k_{eff}M(w) + 2\left[1-k_{eff}\right] - \frac{1-k_{eff}}{M(w)}\right\}, \quad (2\text{-}87)$$

or

$$\frac{i_n^2(x)}{\Delta f} = 2qI_{in}M^2(w)F, \quad (2\text{-}88)$$

where the excess noise factor is

$$F = k_{eff}M(w) + 2\left[1-k_{eff}\right] - \frac{1-k_{eff}}{M(w)}. \quad (2\text{-}89)$$

At unity gain, $F = 1$, at high gains with non negligible values of k_{eff} $F=k_{eff}M(w)$, and at high gains with negligible values of k_{eff}, F=2. Because noise spectral density and gain may be experimentally determined, values for F and k_{eff} are frequently reported by APD manufacturers.

The excess noise factor is intuitively understood as follows[6]. If the APD gain mechanism is a deterministic (noiseless) process, the APD noise factor simply follows the standard shot noise expression and the excess noise factor is unity. However, the gain, M, is actually a random process with a wide probability distribution. Consequently, the mean square gain, $<M^2>$ is greater than the square of the mean gain, $<M>^2$. (Note that the $M^2(\omega)$ term in Equation 2-88 is actually $<M>^2$). The excess noise factor is therefore a correction factor for the difference between $<M^2>$ and $<M>^2$, or $F = <M^2>/<M>^2$.

Frequency Response

Thus far only the frequency independent or low frequency behavior of the APD has been considered. In this section a model is developed to describe its frequency dependent behavior. Figure 2-13 shows the three regions which contribute to the APD frequency response. Region I, the undepleted p-region, exists only in the beveled-edge structure. Region II, the high field region, is common to both the reach-through and beveled-edge APD, and Region III is the external circuitry. Transfer functions are developed for each of these regions and the over-all frequency response is obtained by convolving the transfer functions.

The Undepleted P-Region (Beveled Edge APD)

Because the beveled-edge APD junction is formed by a deep donor diffusion from the top surface of the silicon wafer, the structure has a graded impurity profile much like that of the base in a double-diffused PNP transistor. When the beveled-edge APD is properly biased, approximately 10 μm of p-region at the top surface remains undepleted and carrier transport across this region is due to both drift and diffusion. The derivation of the transfer function for this region presented here relies heavily on the treatment of the frequency dependent base transport factor presented in [7].

Electric Field Profile in the Undepleted P-region

The beveled edge APD junction is formed by performing a closed ampoule, long gallium diffusion into high resistivity n-type silicon. In this type of diffusion the surface concentration is held constant at the solid solubility limit of the doping species, C_s, and the rest of the doping profile, $N_a(x)$ is determined by the time, t, and temperature, T, of the diffusion process and the diffusion coefficient, D, of the doping species.

$$N_a(x) = C_s \, erfc\left(\frac{x}{2\sqrt{Dt}}\right). \tag{2-90}$$

Over a narrow region, such as the undepleted portion of the beveled-edge APD p-diffusion, Equation 2-90 is well approximated by

$$N_a(x) \approx p_p(x) = A\exp(-Bx), \tag{2-91}$$

where A and B are constants which can be determined using a numerical curve fitting algorithm. For the doping levels typically used in beveled-edge APDs, the free hole concentration in the undepleted p-region, $p_p(x)$ is approximately equal to the doping concentration.

At equilibrium, the current density in the undepleted p-region is given by

$$J = q\mu_n n_p(x)E + qD_n \frac{dn_p(x)}{dx} + q\mu_p p_p(x)E - qD_p \frac{dp_p(x)}{dx} = 0. \tag{2-92}$$

Because $p_p(x) >> n_p(x)$ and $-qD_p[dp_p(x)/dx] >> -qD_n[dn_p(x)/dx]$, Equation 2-92 reduces to

$$J \approx q\mu_p p_p(x)E - qD_p \frac{dp_p(x)}{dx} = 0. \tag{2-93}$$

Solving for the electric field yields

$$E = \frac{D_p}{\mu_p} \cdot \frac{1}{p_p(x)} \frac{dp_p(x)}{dx} = \frac{kT}{q} \cdot \frac{1}{p_p(x)} \frac{dp_p(x)}{dx}. \tag{2-94}$$

Substituting Equation 2-91 and its derivative into Equation 2-94 yields

$$E = -B\left(\frac{kT}{q}\right) = constant. \tag{2-95}$$

Thus, for the calculation of the transfer function in the undepleted p-region the electric field is treated as a constant.

Frequency Dependent Transfer Function of the Undepleted P-region

Consider the schematic diagram of the undepleted p-region shown in Figure 2-14. For simplicity the very thin p^+ contact is omitted. A sinusoidal photon flux, $\phi(t) = \phi\exp(j\omega t)$, is incident on the structure and all of the electron-hole pairs are assumed to be generated at $x=0$. (This is a reasonable approximation for the blue light emitted by BGO in which 87% of the photons are absorbed in the first 1.4 μm.). With a sinusoidal forcing function, the minority carrier distribution, $n_p(x)$ has a quiescent component and a frequency dependent component[7]:

$$n_p(x) = n_{pdc}(x) + n_{pac}\exp(j\omega t) \tag{2-96}$$

The continuity equation for the undepleted p-region is [7]

$$\frac{dn_p(x)}{dt} = j\omega n_{pac} e^{j\omega t} = D_n\left(\frac{d^2 n_{pdc}(x)}{dx^2} + e^{j\omega t}\frac{d^2 n_{pac}(x)}{dx^2} + \frac{\eta}{W}\frac{dn_{pdc}(x)}{dx} + \frac{\eta}{W}e^{j\omega t}\frac{dn_{pac}(x)}{dx}\right), \tag{2-97}$$

where $\eta = qEW/kT$ is the built in electric field parameter and W is the width of the undepleted p-region. At equilibrium the dc component of the continuity equation is

$$\frac{dn_{pdc}(x)}{dt} = 0 = D_n\left(\frac{d^2 n_{pdc}(x)}{dx^2} + \frac{\eta}{W}\frac{dn_{pdc}(x)}{dx}\right), \tag{2-98}$$

so Equation 2-97 reduces to

$$\frac{dn_p(x)}{dt} = j\omega n_{pac} = D_n\left(\frac{d^2 n_{pac}(x)}{dx^2} + \frac{\eta}{W}\frac{dn_{pac}(x)}{dx}\right). \tag{2-99}$$

The boundary conditions for the continuity equation are $n_{pac}(0)$ = *constant* and $n_{pac}(W) = 0$. Solving Equation 2-99 for these boundary conditions yields

$$n_{pac}(x) = \phi\left(\frac{\exp(\lambda_2 W + \lambda_1 x) - \exp(\lambda_1 W + \lambda_2 x)}{\exp(\lambda_2 W - \exp(\lambda_1 W)}\right), \tag{2-100}$$

where

$$\lambda_1 = -\frac{\eta}{2} + \sqrt{\left(\frac{\eta}{2W}\right)^2 + \frac{j\omega}{D_n}}, \tag{2-101}$$

and

$$\lambda_2 = -\frac{\eta}{2} - \sqrt{\left(\frac{\eta}{2W}\right)^2 + \frac{j\omega}{D_n}}. \tag{2-102}$$

Equation 2-13 is used to obtain the current densities at $x=0$ and $x=1$:

$$J_n(0) = qD_n\left(\left.\frac{dn_{pac}(x)}{dx}\right|_{x=0} + \frac{\eta}{W}n_{pac}(0)\right), \tag{2-103}$$

$$J_n(W) = qD_n\left(\left.\frac{dn_{pac}(x)}{dx}\right|_{x=W}\right). \tag{2-104}$$

The ratio of currents at $x=W$ and $x=0$ is the transfer function of the undepleted p-region:

$$H(\omega) = \frac{J_n(W)}{J_n(0)} = \frac{\exp\left(-\frac{\eta}{2}\right)}{\cosh\left(\sqrt{\left(\frac{\eta}{2}\right)^2 + \frac{j\omega W^2}{D_n}}\right) - \frac{\frac{\eta}{2}}{\sqrt{\left(\frac{\eta}{2}\right)^2 + \frac{j\omega W^2}{D_n}}}\sinh\left(\sqrt{\left(\frac{\eta}{2}\right)^2 + \frac{j\omega W^2}{D_n}}\right)}. \tag{2-105}$$

A single pole approximation for Equation 2-105 is

$$H(\omega) = \frac{1}{1 + \frac{j\omega}{\omega_0}}, \tag{2-106}$$

where [7]

$${}_0 = \frac{2.43 D_n}{W^2}\left[1 + \left(\frac{\eta}{2}\right)^{\frac{4}{3}}\right]. \tag{2-107}$$

The general transfer function and single pole approximation are compared in Figures 2-15 and 2-16 where the frequency scales in these figures are normalized to $(W^2/D_n) = 1$. For frequencies below ω_0, the single approximation is reasonably accurate. At higher frequencies, however, the single pole approximation is rather poor. For commercially available beveled edge APDs, the undepleted p-region typically has a width of $W \approx 10$ µm, $E \approx 50$ V/cm, $\mu_n \approx 1000$ cm^2/V-s, $D_n \approx 25$ cm^2/s and $\eta \approx 2$. Thus, a typical -3dB frequency for a commercially available beveled-edge APD given by Equation 2-107 is approximately 120 x 10^6 rad/sec or 20 MHz. A 10 µm undepleted p-region will therefore limit the APD rise-time to approximately 18 ns. Similarly, a 5 µm undepleted p-region will limit the APD rise-time to approximately 5 ns. This analysis clearly shows that the undepleted p-region thickness must be minimized to achieve the short rise-times required for high resolution timing performance.

The High-Field Region

Both the beveled-edge and reach-through APD structures have high-field regions in which the APD gain is generated (Figure 2-17). The frequency dependent transfer function developed here is based on the parallel plate capacitor analogy presented at the beginning of this chapter. The following assumptions are made.

1. Assume electrons are injected into the high-field region at x=0. For the beveled-edge structure this implies electrons are injected from the un-depleted p-region. For the reach-through structure this implies the photo-generated electron-hole pairs are all generated at x=0 (the same assumption was made in developing the transfer function for the beveled-edge undepleted p-region).
2. Assume all free carriers travel at a field independent saturation velocity. For electrons, $v_{s,e} \approx 10 \times 10^6$ cm/sec. For holes, $v_{s,h} \approx 6 \times 10^6$ cm/sec.

3. Assume all multiplication occurs instantaneously at the metallurgical junction *($x = Xj$).*

With these assumptions the high field region is well described by the parallel plate capacitor model and the output current due to a unit charge impulse injected at $x=0$ is shown in Figure 2-6 and described by

$$\frac{i_{out}(t)}{\delta i(t)} = \frac{1}{T'}u(t) + N\frac{1}{T''}u(t-t') - \frac{1}{T'}u(t-T') - N\frac{1}{T''}u(t-t'-T''), \qquad (2\text{-}108)$$

where $t' = X_j / v_{s,e}$ is time required for the injected electron to reach the metallurgical junction, $T' = W/v_{s,e}$ is the time required for the injected electron to traverse the entire high field region, and $T'' = X_j/v_{s,h}$ or $(W-X_j)/v_{s,e}$ (whichever is greater) is the time required for all of the multiplied carriers to reach the edges of the depletion region. Because Equation 2-108 is the output current due to a current impulse, it represents a time-domain impulse response, $h(t)$, for the high-field region.

For the APDs under consideration here, the gain ($M=N+1$) is much greater than unity. Consequently $M \approx N$ and the unmultiplied current may be neglected, yielding the simpler expression

$$h(t) \approx \frac{M}{T''}\left(u(t-t') - u(t-t'-T'')\right). \qquad (2\text{-}109)$$

It is shown by convolving Equation 2-109 with $u(t)$ that the high-field region has a step input rise time equal to T'' and a delay equal to $t' + T''$. The frequency dependent transfer function is obtained by taking the Laplace transform of Equation 2-109 to obtain

$$H(\omega) = \frac{M}{sT''}\left[\exp(-s(t'+T'')) - \exp(-st')\right], \qquad (2\text{-}110)$$

or

$$H(\omega) = \frac{M}{T''\omega}\left\{\left[\sin(\omega(t'+T'')) - \sin(\omega t')\right] + j\left[\cos(\omega(t'+T'') - \cos(\omega t')\right]\right\}. \qquad (2\text{-}111)$$

For commercially available beveled-edge APDs, the junction is approximately 25 μm below the undepleted p-region and 150 μm above the undepleted n-region. The delays in Equations 2-109 through 2-111 are then $t' \approx (25 \times 10^{-4}$ cm)/(10×10^{6} cm/sec) = 250 ps and $T'' \approx (150 \times 10^{-4}$ cm)/(10×10^{6} cm/sec) = 1.5 ns. Thus, the high field region limits the typical beveled-edge APD rise-time to approximately 1.5 ns and introduces approximately 1.75 ns of delay. Note that the rise-time limitation imposed by the high-field region is less than half the 5 ns rise-time limitation imposed by a hypothetical 5 μm wide undepleted p-region.

For a typical reach-through APD the junction is directly above the bottom contact, approximately 150 μm below the top contact. For this structure typical delays are $t' \approx (150 \times 10^{-4}$ cm)/(10×10^{6} cm/sec) = 1.5 ns and $T'' \approx (150 \times 10^{-4}$ cm)/(6×10^{6} cm/sec) = 2.5 ns. Thus, the high field region limits the typical reach-through APD rise time to approximately 2.5 ns and introduces approximately 4 ns of delay. Unlike the beveled-edge structure, the reach-through APD frequency response is largely determined by the width of the high field region.

A single pole approximation for Equation 2-111 which has been used elsewhere to model the transit time across the base-collector depletion region of microwave bipolar transistors [1] is

$$H(\omega) = M\left(\frac{1}{1+j\omega/\omega_1}\right), \qquad (2\text{-}112)$$

where the -3db frequency is $\omega_1=2/T''$. The general expression and single pole approximation are compared in Figures 2-18 and 2-19 where the data is normalized to $M=1$, $t'=1$ and $T''=1$. These plots indicate Equation 2-112 provides a reasonably accurate model of the transfer function magnitude, but significantly underestimates the phase shift in the high field region.

External Circuitry

The output of an APD is a current which is injected into a read-out preamplifier. One simple readout circuit might be a series load resistor, R_L coupled to a low capacitance oscilloscope probe as shown in Figure 2-13. This circuit has a single pole response with a time constant

$$\tau = R_L\left(C_{APD} + C_{PROBE} + C_{STRAY}\right). \tag{2-113}$$

A more practical readout circuit would be an ideal charge sensitive preamplifier with a single dominant pole at $\omega=0$. For the remainder of this discussion the frequency response of the APD is assumed to dominate.

Composite Transfer Function

Assuming the read-out electronics do not limit the APD frequency response, the transfer function of the reach-through APD is determined solely by the high-field region and given by Equations 2-109 and 2-111. The beveled-edge APD transfer function is determined by both the undepleted p-region and the high field region and is given by

$$H(\omega) = \frac{M}{T''\omega} \cdot \frac{\exp\left(-\frac{\eta}{2}\right) \cdot \left\{\left[\sin(\omega(t'+T'')) - \sin(\omega t')\right] + j\left[\cos(\omega(t'+T'') - \cos(\omega t')\right]\right\}}{\cosh\left(\sqrt{\left(\frac{\eta}{2}\right)^2 + \frac{j\omega W^2}{D_n}}\right) - \frac{\frac{\eta}{2}}{\sqrt{\left(\frac{\eta}{2}\right)^2 + \frac{j\omega W^2}{D_n}}} \sinh\left(\sqrt{\left(\frac{\eta}{2}\right)^2 + \frac{j\omega W^2}{D_n}}\right)}. \tag{2-114}$$

For first order approximations the single pole models are invoked to obtain

$$H(\omega) = M \frac{1}{\left(1 + {}^{j\omega}\!/_{\omega_0}\right)\left(1 + {}^{j\omega}\!/_{\omega_1}\right)}. \tag{2-115}$$

where ω_0 is given in Equation 2-107 and $\omega_1 = 2/T''$.

Conclusions

The avalanche photodiode theory of operation is reviewed and the fundamental expressions for gain and noise are derived. Sample gain and noise calculations are presented for an ideal APD structure. A three part model for the APD frequency response is described and frequency dependent transfer functions are developed for the reach-through and beveled-edge structures. The frequency response is found to be dominated by the carrier transit time across the high-field depletion region, the RC time constant associated with the diode junction capacitance and the external circuitry and, in the case of the beveled-edge APD, the transit time across the undepleted p-region.

In Chapter 3 the equations developed in this chapter are incorporated into a simulation model which predicts the performance characteristics of APDs based on their fabrication process.

References

1. S. M. Sze, *Physics of Semiconductor Devices,* 2nd Edition, New York: John Wiley and Sons, 1981.

2. "Noise Characteristics of Advanced Photonix Avalanche Photodiodes," Advanced Photonix Application Note API/NOIS/1291/B, 1991.

3. J. McIntyre, private communication.

4. J. McIntyre, "Multiplication Noise in Uniform Avalanche Diodes," *IEEE Trans. Elec. Dev.*, **ED-13**, pp. 164-168, 1966.

5. Robert J. McIntyre, "The Distribution of Gains in Uniformly Multiplying Avalanche Diodes: Theory," *IEEE Trans. Elec. Dev.*,**ED-19**, pp. 703-713, 1972.

6. P. Webb, R. J. McIntyre, and J. Conradi, "Properties of Avalanche Photodiodes," *RCA Review,* **35**, pp. 234-278, 1974.

7. Lindmayer and C. Y. Wrigley, *Fundamentals of Semiconductor Devices,* New York: D. Van Nostrand Company, Inc., pp. 152-171, 1965.

Figures

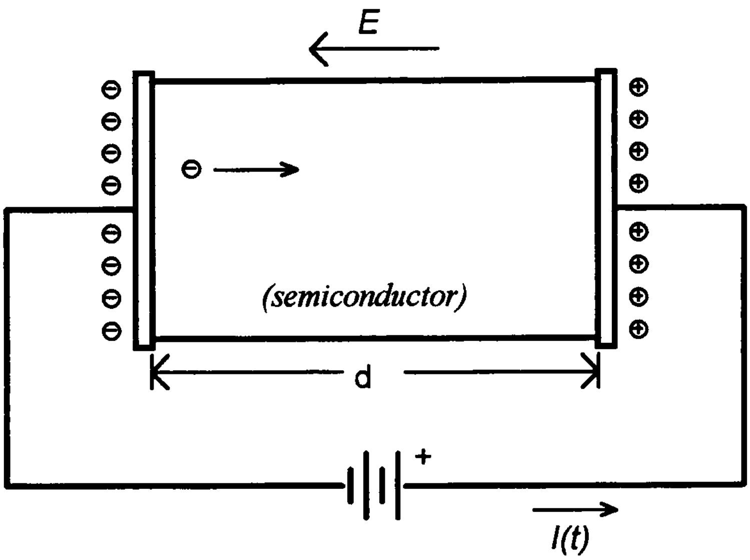

Figure 2-1. Example 1: schematic diagram of a metal-semiconductor-metal capacitor. An electron is injected from the negative plate at time $t = 0$ and traverses the space between the plates with a constant velocity.

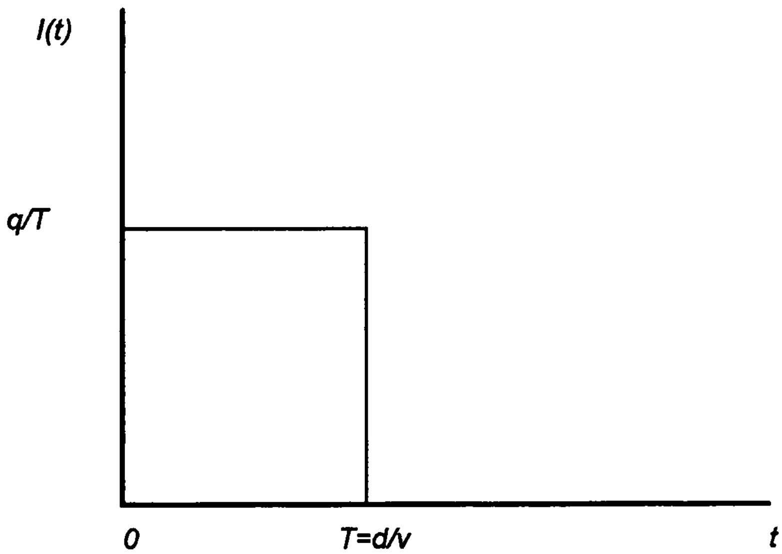

Figure 2-2. Current flow due to the injection of an electron from the negative plate of the metal-semiconductor-metal capacitor. The electron traverses the distance between the plates, d, with a constant velocity, v.

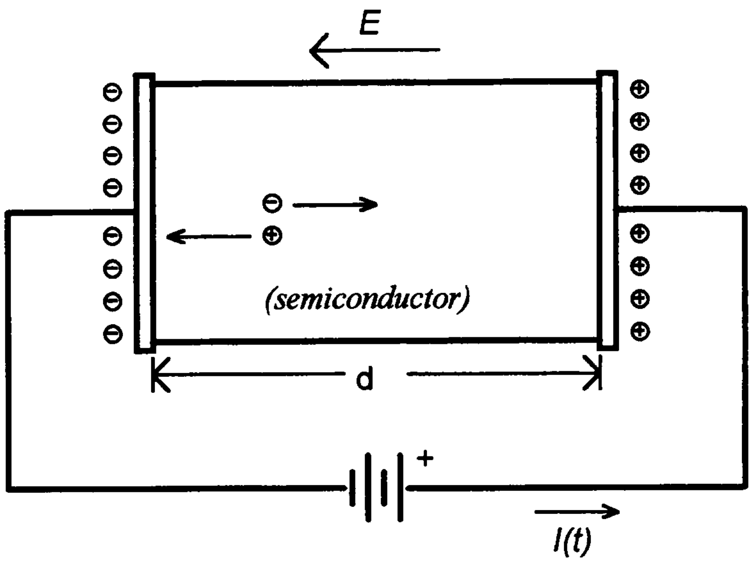

Figure 2-3. Example 2: schematic diagram of a metal-semiconductor-metal capacitor. An electron-hole pair is generated between the plates at time $t = 0$, and the carriers traverse the space between the plates with a constant velocity.

Figure 2-4. Current flow due to the generation of an electron-hole pair between the capacitor plates. The carriers traverses the distance between the plates, d, with a velocity, v. The pulse width and height are set by the time required for the last carrier to reach a plate, T'.

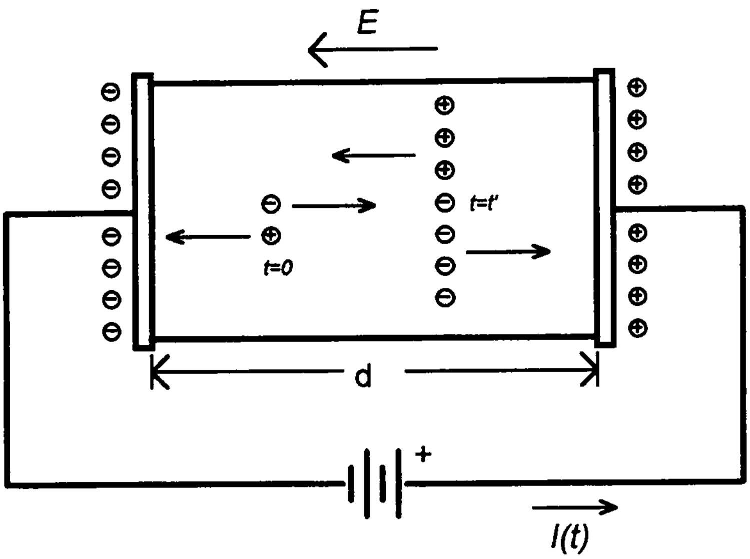

Figure 2-5. Example 3: schematic diagram of a metal-semiconductor-metal capacitor. An electron-hole pair is generated between the plates at time $t = 0$ and subsequently undergoes impact ionization at time $t = t'$.

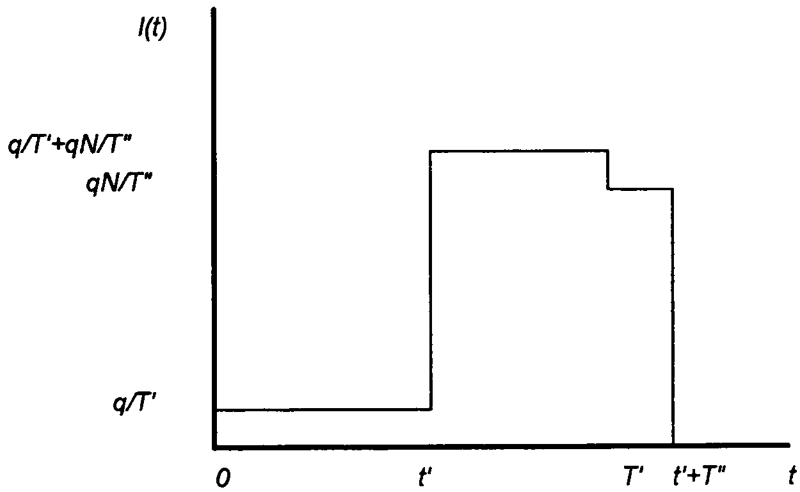

Figure 2-6. Current flow due to the generation of an electron-hole pair between the capacitor plates at $t=0$ followed by avalanche multiplication at $t=t'$. The last initial carrier reaches a plate at time $T'=d'/v$, where d' is the distance traveled by the last carrier. The last multiplied carrier reaches a plate at time $t'+T''$, where $T''=d''/v$, and d'' is the distance traveled by the last multiplied carrier.

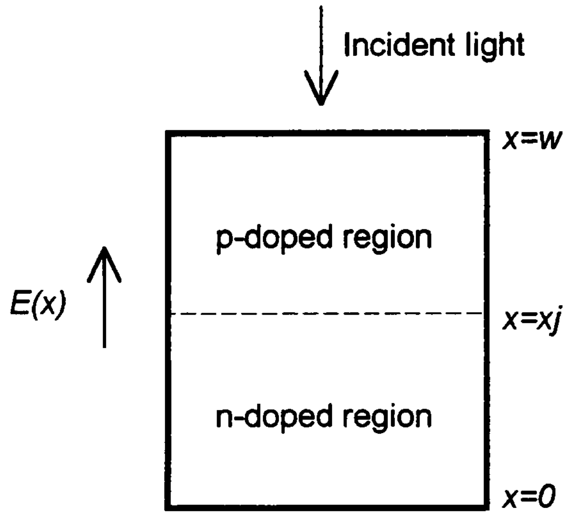

Figure 2-7. Arbitrary APD structure used to develop APD gain and noise model.

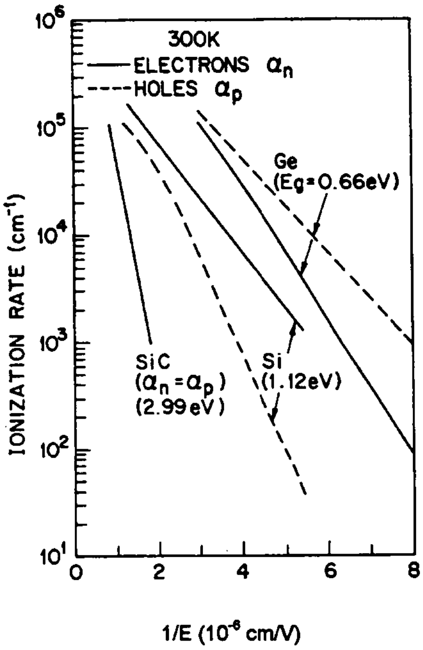

Figure 2-8. Measured electron and hole ionization coefficient values. Source: S.M. Sze, *Physics of Semiconductor Devices*, 2nd Edition, New York: John Wiley and Sons, 1981, p 47.

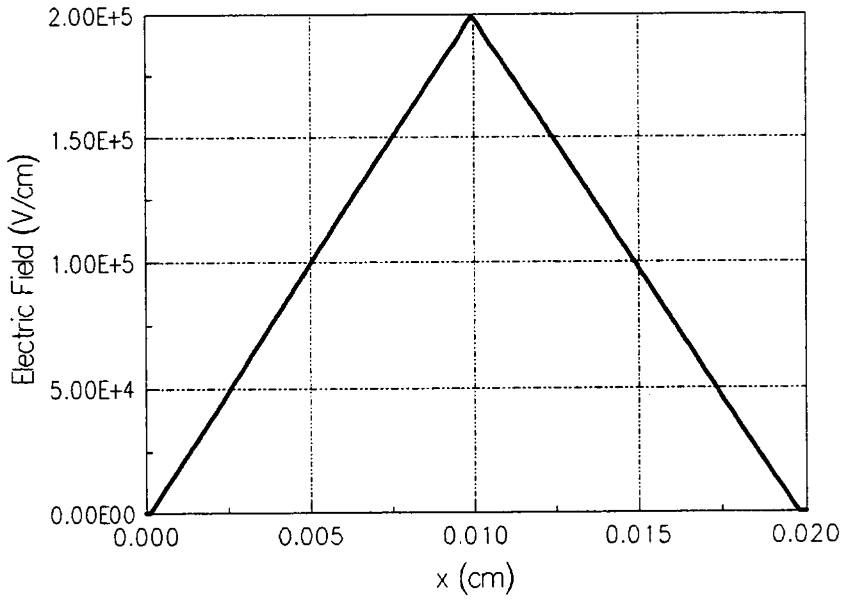

Figure 2-9. Example electric field profile. The metallurgical junction occurs at x=0.01 cm; light enters the structure from the right.

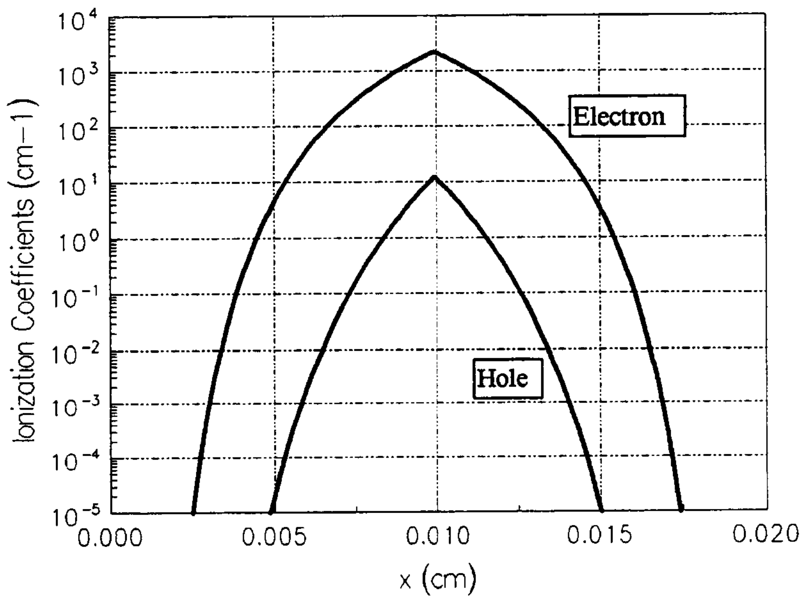

Figure 2-10. Calculated ionization coefficients for example electric field profile.

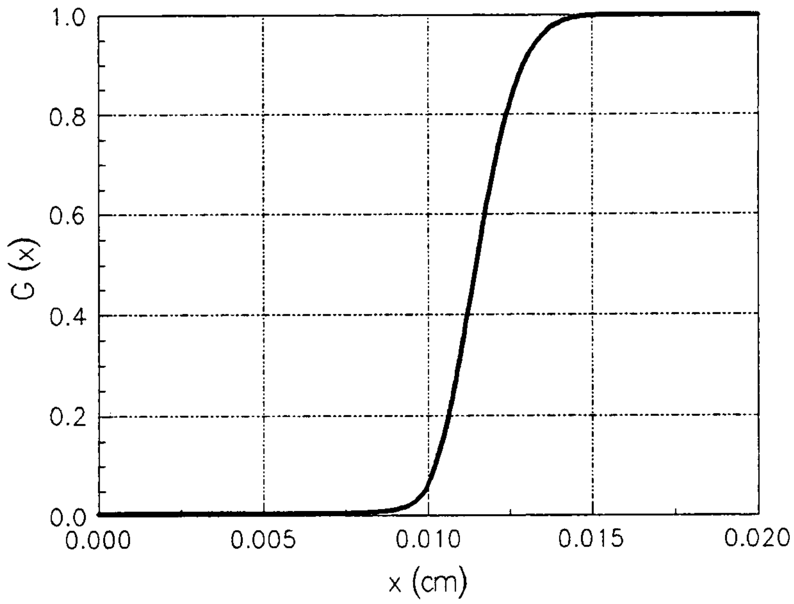

Figure 2-11. Calculated G(x) for example electric field profile. Light enters structure from the right.

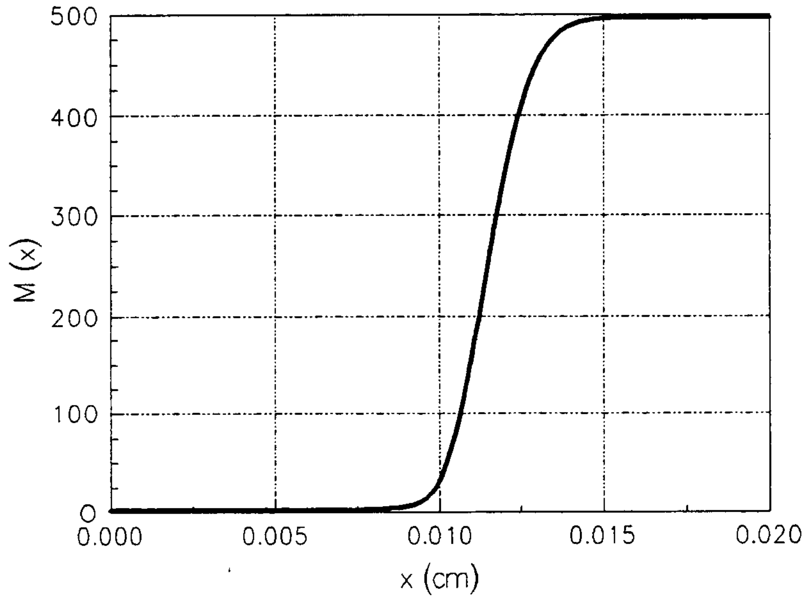

Figure 2-12. Calculated Position dependent gain (M(x)) for example electric field profile. Light enters structure from the right.

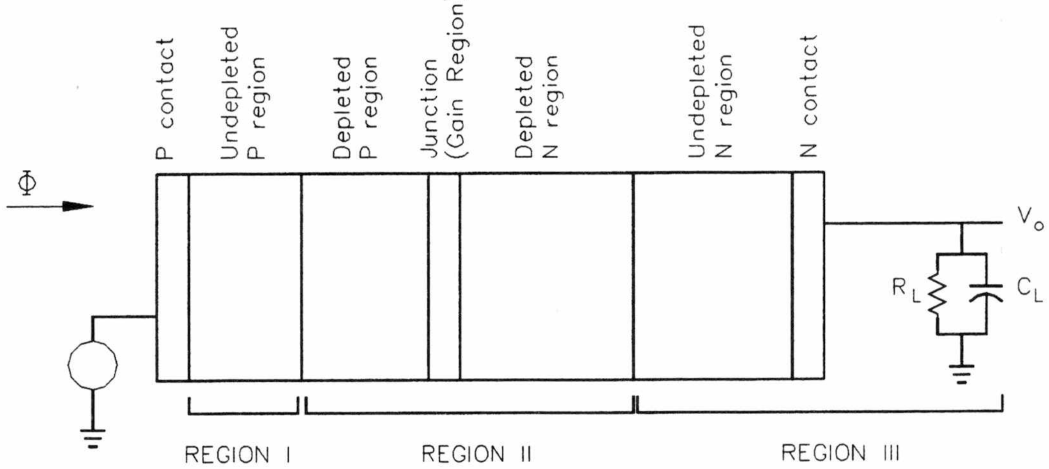

Figure 2-13. Schematic diagram of APD showing the three regions which dominate the frequency response.

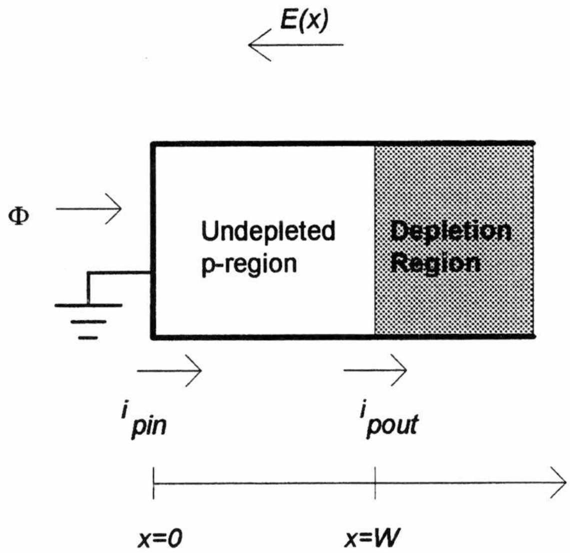

Figure 2-14. Schematic representation of undepleted p-region and edge of depletion region.

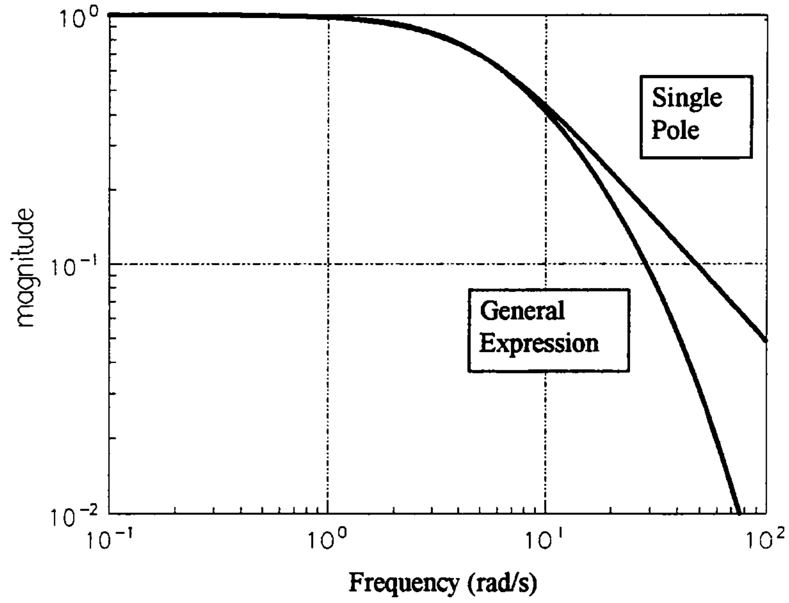

Figure 2-15. Comparison of general expression for p-region frequency dependent signal magnitude with single pole approximation.

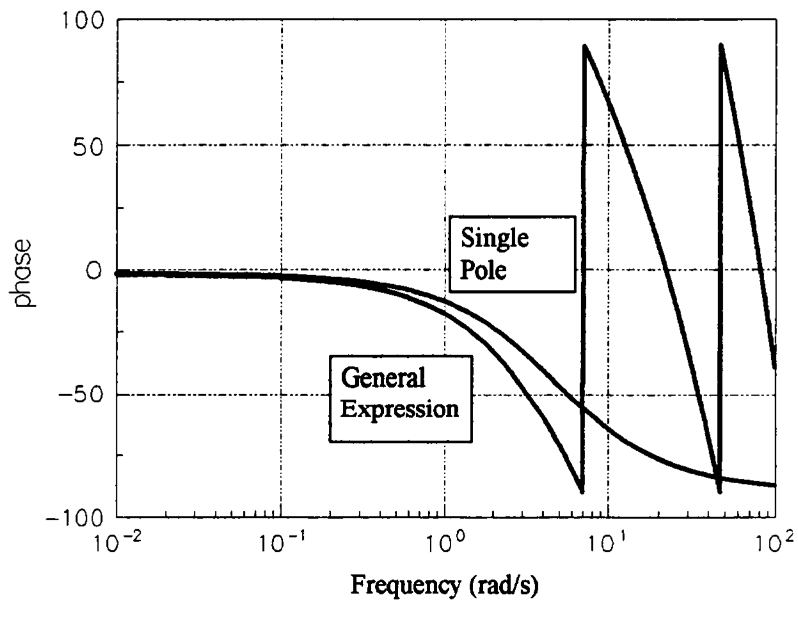

Figure 2-16. Comparison of general expression for p-region frequency dependent phase shift with single pole approximation.

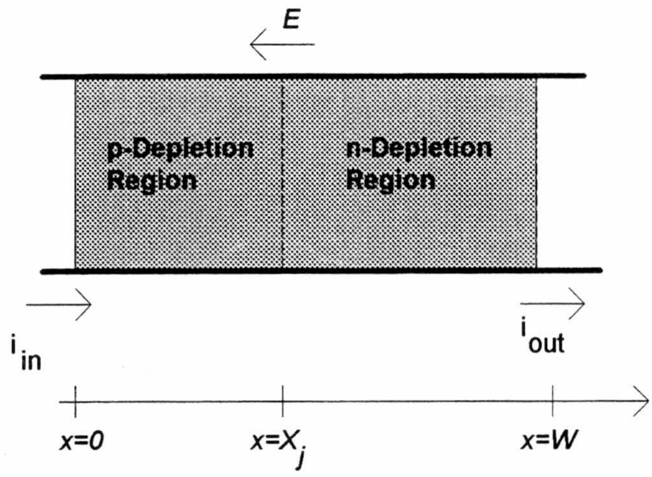

Figure 2-17. Schematic representation of the high field region.

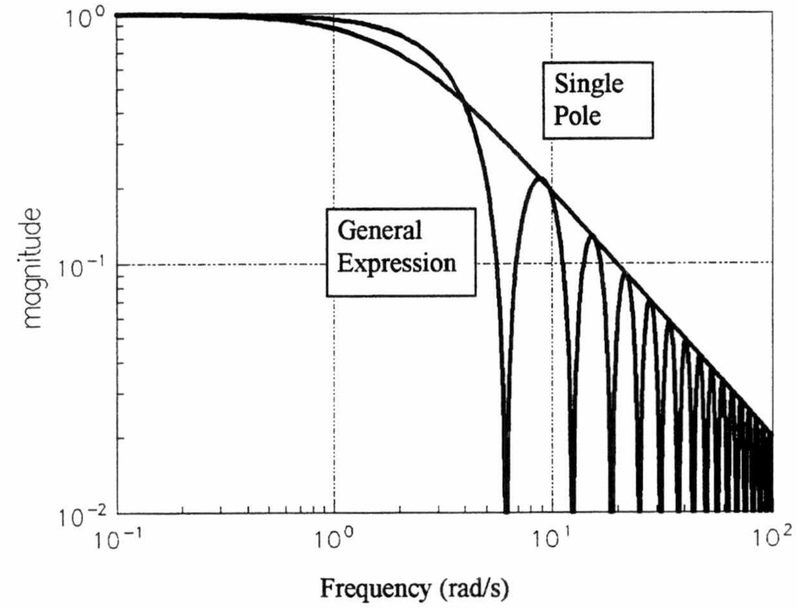

Figure 2-18. Comparison of general expression for drift region frequency dependent signal magnitude with single pole approximation.

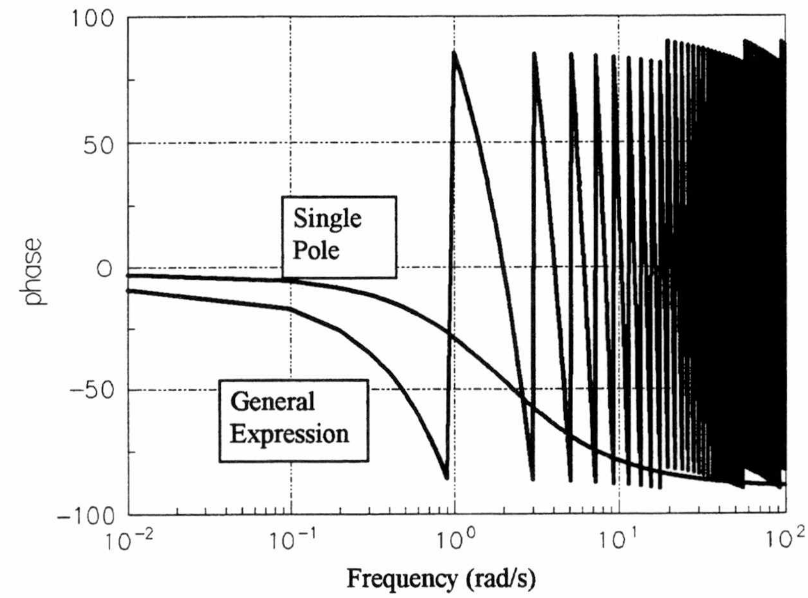

Figure 2-19. Comparison of general expression for drift region frequency dependent phase shift with single pole approximation.

PART 3

THEORETICAL AND EXPERIMENTAL COMPARISON OF DISCRETE BEVELED-EDGE AND REACH-THROUGH APDS

Overview

In Chapter 2 the expressions for APD gain and noise are derived and the frequency dependent transfer function is calculated. In this chapter the expressions developed in Chapter 2 are incorporated into a simulation model which calculates the expected performance of an APD structure based on its fabrication process.

The beveled-edge and reach through structures are the two most widely used APD profiles. This chapter compares the structures and relates their structural differences to their performance characteristics. Typical doping profiles for the two structures are calculated using the semiconductor fabrication simulator SUPREM [1] and fabrication data found in the literature. From these doping profiles, the bias dependent electric field profiles are calculated using the Poisson equation solver FISH1D [2]. The field profiles are in turn used to calculate the ionization coefficients, gain profiles and noise characteristics of the structures using the theory developed in Chapter 2. The calculated data is compared with experimental energy and timing resolution data obtained using commercially available beveled-edge and reach-through discrete APDs coupled to BGO scintillators. The experimental data obtained with the discrete APDs are used to select the manufacturer of a custom APD array for use in a prototype PET detector.

Theoretical Investigation

Beveled Edge APD

The beveled edge APD is the simpler of the two major APD structures. Its gain producing junction is formed by a single deep diffusion of gallium into a lightly doped n-type substrate, and the majority of the fabrication challenges are in the mechanical formation and passivation of the "beveled" edges and in the deep lapping processes required to minimize the undepleted p-region thickness.

Beveled-Edge APD Fabrication Procedure

Although the details of current beveled-edge APD fabrication techniques are trade secrets, much can be learned about the general fabrication steps from the early literature and from the manufacturers' patents [3,4,5]. The key process steps are shown schematically in Figure 3-1.

The starting material is typically a neutron transmutation doped (NTD) silicon wafer, which is obtained by irradiating high resistivity silicon with thermal neutrons. N-type doping is achieved through fractional transmutation of silicon into phosphorous as follows [6,7]:

$$Si_{15}^{30} + neutron \rightarrow Si_{14}^{31} + \gamma\text{-}ray \rightarrow P_{15}^{31} + \beta\text{-}ray. \tag{3-1}$$

Since the penetration range of neutrons in silicon is very long (~1 m) the doping is very uniform. For example, resistivity variations are typically ±15% for conventionally doped silicon and approximately ±0.5% for NTD silicon [6,7]. The NTD wafers used for beveled-edge APDs typically have ~40 Ω-cm resistivities, or ~1 x 10^{14} cm^{-1} doping concentrations.

The first process step is the deep gallium diffusion in which a bare NTD wafer is sealed in a quartz ampoule along with a solid gallium source. The ampoule is then loaded into a furnace where the gallium diffuses into the wafer on all sides as shown in Figure 3-1(1) [8]. The next step is the removal of the gallium doped silicon from all surfaces except the top, and the removal of approximately 50 μm of doped silicon from the top surface (Figure 3-1(2)). The amount of silicon removed from the top surface determines the thickness of the undepleted p-region. Shallow p+ and n+ diffusions are then performed on the top and bottom surfaces, respectively, to create the ohmic contacts (Figure 3-1(3)). Metal (typically gold or a gold alloy) is then deposited on the silicon surface to form metal contacts as shown in Figure 3-1(4). At this point the individual APD elements are cut from the wafer and the edges of each element are mechanically beveled to create the characteristic "beveled-edge" profile. As shown in Figure 3-1(5), the beveling process increases the distance between the top and bottom contact along the surface of the APD, thus reducing the surface electric field in the device relative to the bulk electric field. The angle of the bevel is very important: the angle must be shallow enough to prevent large leakage currents or breakdown along the APD surface, yet, it must be steep enough to prevent the APD from being overly fragile at the edges.

The final step is passivation of the beveled edge. The purpose of the passivation step is threefold. First, the passivation layer should provide a diffusion barrier to prevent impurities on the surface of the APD from diffusing into the silicon. Second, the passivation layer should prevent charged ions from collecting directly on the silicon surface. Charged ions can induce conductive channels on the silicon surface, forming leakage current paths. Finally, the passivation layer should neutralize the dangling bonds at the surface of the silicon. Polyimide is frequently used because it has the advantages of being easily applied and fairly thick (> 50 μm), thus meeting requirements 1 and 2 above. However, since polyimide is a deposited coating rather than a grown insulator such as silicon dioxide, it does not effectively neutralize the dangling bonds. A better passivation which is under consideration for future APD designs [8] combines a thin thermally grown oxide layer designed to passivate the silicon surface states with a thicker polyimide coating designed to provide mechanical protection and to keep impurities from reaching the silicon surface.

SUPREM Simulation of the Beveled-Edge Fabrication Process

From a one-dimensional modeling perspective, the process steps which determine the intrinsic characteristics of the APD are the gallium diffusion, the surface lapping step and the p+ contact diffusion. These process steps were modeled with SUPREM using the estimated process parameters given in Table 3-1.

The SUPREM-calculated doping profile for the model beveled-edge APD is shown in Figure 3-2. The background phosphorus doping concentration ($N_D = 1 \times 10^{14}$ cm^{-1}) corresponds to ~40 Ω-cm n-type silicon. After the lapping step, the gallium concentration at the surface is approximately 2×10^{17} cm^{-1} and the junction is approximately 30 μm below the silicon surface. The p+ (boron) contact layer is approximately 0.5 μm thick.

Calculation of the Beveled-Edge APD Performance Parameters

Once the one-dimensional doping profile is known, the bias-dependent electric field may be calculated by solving the one-dimensional Poisson equation,

$$\frac{dE(x)}{dx} = \frac{\rho(x)}{\varepsilon}, \qquad (3\text{-}2)$$

where $E(x)$ is the position dependent electric field, $\rho(x)$ is the net charge at x and ε is the dielectric constant of silicon. This equation was solved using FISH1D, a one-dimensional Poisson Equation Solver developed at Purdue University [2]. The FISH1D-calculated electric field profile for a calculated gain of 210 (V = 1475V) is shown in Figure 3-3. The peak electric field is approximately -2×10^5 V/cm and occurs at the metallurgical junction. An expanded view of the electric field profile is shown in Figure 3-4; the undepleted region is approximately 6 μm wide and has a relatively constant built in electric field of ~-60 V/cm.

Table 3-1.
Estimated Fabrication Parameters for Beveled-Edge APD SUPREM Simulations.

Step	**Dopant**	**Time**	**Temperature**	**Depth**
Deep p-diffusion	Gallium	12.5 hours	1350 C	$x_j \geq 75\mu m$
Lap top surface				remove ~50 μm
Shallow p^+-diffusion	Boron	20 minutes	1100 C	$x_j < 0.5\mu m$

The position dependent ionization coefficients are calculated from the electric field profile using Equations 2-12 and 2-13 (Figure 3-5). The electron ionization coefficient ($\alpha(x)$) has a maximum value of 3000 cm^{-1} and the hole ionization coefficient ($\beta(x)$) has a maximum value of 100 cm^{-1}. Using these calculated profiles $\alpha(x)$ and $\beta(x)$, the APD performance parameters introduced in Chapter 2, are then obtained (Table 3-2):

Table 3-2.
Calculated Performance Parameters of a Beveled-Edge APD Biased at 1475V.

Parameter	Value	Equation
k_1	.0035	2-37
k_2	.0030	2-73
k_{eff}	.0030	2-84
M_{max}	210	2-44
F	2.63	2-89

The position dependent ionization coefficients are used to obtain the calculated position dependent gain profile shown in Figure 3-6. The most important feature of the gain profile is the narrow region of high gain sensitivity above the junction. Photons absorbed in the ~25 μm above the junction see the full gain of the device whereas photons absorbed below the junction see only a fraction of the total gain. This is because a photon absorbed above the junction will send an electron through the peak electric field while a photon absorbed below the junction will send a hole below through the peak electric field. As shown in Figure 3-5, the electron ionization coefficient is more than an order of magnitude greater than the hole ionization coefficient. For long wavelength light, the placement of the peak electric field in the beveled-edge structure is a decided disadvantage since photons will be deposited on both sides of the junction. For the short wavelength light emitted by BGO and other scintillators of interest for PET, however, virtually all of the photons are absorbed in the first few microns of the structure where they see the full gain of the device.

For minimization of dark current, the near-surface location of the junction in the beveled-edge device offers a clear advantage. The dark current is given by

$$I_d = I_s + M_p I_{gp} + M_n I_{gn}, \tag{3-3}$$

where I_s is the surface leakage current, M_p and M_n are the gains seen by electron-hole pairs generated in the p and n depletion regions and I_{gp} and I_{gn} are the thermal generation currents in the p and n depletion regions. Because $M_p >> M_n$, $M_p I_{gp}$ dominates the dark current expression at high gains. In order to minimize the bulk dark current at high gains, it is necessary to minimize the depleted p-region width. The placement of the beveled edge junction near the top surface of the diode should therefore lead to lower bulk dark current. It should be noted, however, that the surface component of the total APD dark current, I_s, is not affected by the location of the metallurgical junction and can dominate the overall dark current in a poorly fabricated device, particularly at low gains [9].

Figure 3-7 shows the calculated gain vs. bias voltage curve for the beveled-edge APD. (Note that the voltage scale has a minimum value of 1000V.) As should be expected by the exponential dependence of the ionization coefficients on the bias voltage (Equation 2-12), the gain-voltage curve has roughly exponential shape. For the PET application, the APD would typically be biased for a gain greater than 200, well above the knee of the gain curve. Here the device has a strong sensitivity to the bias voltage (~2.5%/V at M=200) which must be considered when the APD bias circuitry is designed.

The excess noise factor was calculated as a function of the APD gain using Equation 2-89 and the previously calculated $k_{eff}=0.0030$ (Figure 3-8). For gains above ~10, $F \approx 2 + k_{eff} \cdot M$, indicating the importance of minimizing k_{eff} in the APD design and of choosing the minimum acceptable gain when operating the APD. The procedure for choosing the APD gain for optimal signal to noise ratio is presented in the *Experimental Procedure* portion of this chapter.

Reach-Through APD

The reach-through APD concept is shown schematically in Figure 1-8(b). The gain producing junction is formed by two diffusion steps into the back side of the wafer, yielding a device with higher sensitivity, lower operating voltage and a somewhat higher excess noise factor than typical beveled-edge devices. A typical fabrication procedure is described below [10].

Reach-through APD Fabrication Procedure

Unlike the beveled-edge APD manufacturers, the principal developer of the reach-through APD, EG&G Optoelectronics, presents detailed descriptions of their fabrication processes in their patent disclosures. The process presented here follows EG&G Patent 4,463,368 which was issued in 1984. EG&G has subsequently enhanced their structure by adding guard ring structures to reduce the surface leakage current [11]. The 1984 structure is chosen for this discussion, however, because it is simpler and because the added guard ring structures do not affect intrinsic performance parameters. The key process steps are shown schematically in Figure 3-9.

The starting material for the reach-through APD is very high resistivity (>5000 Ω-cm) p-type silicon. The p-side of the APD junction is formed by a boron implantation into the back side of the wafer (Figure 3-9(1)), followed by a long drive-in diffusion (Figure 3-9(2)). The n-side of the junction (Figure 3-9(3))is then formed by a short pre-diffusion of phosphorus into the back side of the wafer followed by a short drive-in diffusion. The boron top contact (Figure 3-9(4)) is typically formed by a simultaneous deposition and drive-in of boron from a boron nitride wafer at a temperature of about

1000 C for approximately 30 minutes. The top side metal contacts are typically formed by sequential deposition of Cr and Au layers while the bottom metal contacts are typically formed by sequential deposition of Al, Cr and Au layers (Figure 3-9(5)). Also shown in Figure 3-9(5) are p+ channel stop regions which are either implanted or diffused several micrometers into the substrate. The channel stops are actually formed at the beginning of the fabrication procedure and are designed to prevent the depletion region formed at the edges of the n+ region from extending to the sidewalls of the APD.

The edges of the reach-through APD are also passivated, but the demands on this step are less stringent than those imposed on the beveled-edge passivation step because the junction does not extend to the sides of the device. Consequently, the sidewalls do not support the full electric field of the device and surface leakage currents are more easily minimized.

SUPREM Simulation of the Reach-Through Fabrication Process

For the purposes of developing a one-dimensional SUPREM model, the key process steps are the boron implant and drive-in and the phosphorus diffusion. The process parameters used in the SUPREM model are summarized in Table 3-3.

The SUPREM-calculated one-dimensional doping profile for the reach-through APD is shown in Figure 3-10. The modeled background p-doping concentration is 10^{12} cm^{-3}, the back-side boron diffusion extends ~40 μm into the substrate and the metallurgical junction appears ~4 μm above the bottom surface. Note that the doping concentrations at the junction (~$2x10^{15}$ cm^{-3}) are an order of magnitude greater than the concentrations at the beveled-edge junction while the background doping concentration in the substrate is two orders of magnitude lower than that of the beveled-edge APD. These differences cause the electric field profile of the reach-through APD to be markedly different from that of the beveled-edge structure.

Table 3-3.
Estimated Fabrication Parameters for Reach-Through APD SUPREM Simulations.

Step	Dopant	Time/Dose	Temperature/Energy	Depth
P Implant	boron	dose=$6x10^{12}$ cm^{-2}	energy=100 keV	
P drive-in	boron	time=42 hr	temp = 1250C	~40 μm
N predeposition	phosphorus	time=10 min	temp=1135C	
N drive-in	phosphorus	time=3.3 hr	temp=1135C	~4 μm

Calculation of the Reach-Through APD Performance Parameters

The FISH1D calculated electric field profile for the beveled-edge APD biased at a gain of 210 is shown in Figure 3-11. The electric field peak appears near the metallurgical junction, approximately 4 μm above the bottom of the diode. Because the doping concentrations near the junction are much greater than those of the beveled-edge structure, the maximum field (~$2.75x10^5$ V/cm) is greater than that of the beveled edge device and the peak is considerably narrower. Because the native substrate has a very low doping concentration, the depletion region extends, or "reaches through" the entire structure, preventing the presence of rise-time limiting undepleted regions.

The position dependent ionization coefficients are calculated from the electric field profile using Equations 2-12 and 2-18 (Figure 3-12). (Note that the x-axis of Figure 3-12 is expanded to show only the lower half of the device). The peak ionization coefficients (α_{max}≈20,000 cm^{-1}, β_{max}≈1,000 cm^{-1}) are considerably higher than those of the beveled-edge APD, and, more importantly, the ratio of the ionization coefficients is an order of magnitude lower than that of the beveled-edge APD. The excess noise factor is strongly dependent on the ratio of the ionization coefficients, so the reach-through APD should have a significantly higher excess noise factor than the beveled edge APD.

The position dependent ionization coefficients are used to calculate the remainder of the APD performance parameters summarized in Table 3-4 for a reach-through APD biased at 550V. The equations used to obtain the parameters are also summarized in Table 3-4. The calculated position dependent gain profile for V=550V is shown in Figure 3-13. Unlike the beveled-edge APD, the high gain region of the reach-through APD extends across virtually all of the device. This configuration does offer improved sensitivity to longer wavelength light, but it is of no significant advantage for the PET application. Because most of the depletion region does lie in the high-gain region, however, the reach-through APD should have a greater bulk dark current than the beveled edge device.

Table 3-4.
Calculated Performance Parameters of a Reach-Through APD Biased at 550V.

Parameter	**Value**	**Equation**
k_1	.0207	2-37
k_2	.0084	2-73
k_{eff}	.0081	2-84
M_{max}	210	2-44
F	3.64	2-89

The electric field profile was calculated using FISH1D for a number of bias voltages to obtain the calculated reach-through APD gain vs. bias voltage plot shown in Figure 3-14. The bias voltages are significantly lower than those for the beveled-edge structure because the higher doping concentrations at the junction provide a greater built-in electric field. The shape of the gain curve near the expected operating point for PET detectors (M>200) and the sensitivity of the gain to small changes in bias voltage are similar to those of the beveled-edge device. The calculated excess noise factor vs. gain curve (Equation 2-89) is shown in Figure 3-15. As expected, the calculated excess noise factor is considerably higher than that of the beveled-edge APD due to the lower ratio of electron and hole ionization coefficients.

Summary of Simulation Data

The key characteristics of the beveled edge and reach through APD structures are summarized below:

1. The beveled-edge junction is formed by a single deep diffusion into the top surface of the device, placing the junction approximately 25 μm below the diode top contact. The reach-through junction is formed by two diffusions into the back surface of the device, placing the junction approximately 4 μm above the bottom contact.
2. The impurity concentrations at the junction are an order of magnitude lower in the beveled-edge APD than in the reach-through APD. The beveled-edge device therefore has a much lower built-in field at the junction and requires a larger bias voltage to achieve a given gain. For a gain of ~210 the calculated bias voltages are 1475V and 550V for the beveled-edge and reach-through structures, respectively.
3. The beveled-edge structure has a lower excess noise factor than the reach-through structure due to the lower electric field in the multiplication region.
4. The beveled-edge structure should have a lower bulk dark current than the reach-through structure due to its narrower depletion region above the metallurgical junction.
5. The reach-through structure does not have a rise-time limiting undepleted p-region; its depletion region extends through virtually the entire structure.
6. Both structures have wide depletion regions (>100 μm), yielding capacitances less than 100 pF/cm^2.

Based upon this study, the beveled-edge APD is predicted to be more suitable than the reach-through structure for the PET detector application. The key advantages of the beveled-edge structure for PET are lower excess noise factor and lower bulk dark current. For the beveled-edge structure to be acceptable, however, the undepleted p-region thickness must be minimized. Calculations presented in Chapter 2 indicate the rise-time can be less than 5 ns if the undepleted region is made narrower than 5 μm.

Experimental Investigation

To verify the theoretical analysis, beveled-edge APDs from Advanced Photonix (API) and Radiation Monitoring Devices, Inc. (RMD) and a modular reach-through APD/BGO detector were experimentally evaluated. This section presents a background discussion on pulse-height resolution, experimental pulse-height resolution data, a background discussion on timing resolution and experimental timing data. The experimental results are then compared with the theoretical predictions presented above.

Pulse Height Resolution: Background and Example Calculation

A typical readout circuit for measuring APD pulse height spectra including the dominant noise sources is shown in Figure 3-16. The circuit consists of the APD coupled to a scintillator, a charge sensitive preamplifier and a CR-RC shaping amplifier. When the APD detects a light pulse in the scintillator, the output is a voltage pulse with amplitude proportional to the number of photons in the scintillator light pulse. The accuracy with which the system measures the total light output is defined as the system pulse height resolution (PHR) or energy resolution,

$$PHR = \frac{\sigma_{vo}}{\langle v_{o,peak} \rangle}, \qquad (3\text{-}4)$$

where σ_v is the standard deviation or rms noise of the output peak and $\langle v_{o,peak} \rangle$is the average output pulse height. Equation 3-4 expresses the PHR as a normalized standard deviation; a more common presentation of the parameter is as a full width at half maximum (FWHM) value,

$$FWHM_{vpeak} = 2.35 \cdot \frac{\sigma_{vo}}{\langle v_{o,peak} \rangle}. \qquad (3\text{-}5)$$

Expressions for $\langle v_{o,peak} \rangle$ and $\sigma_{v\,o}$ are developed below and expected values for the FWHM are calculated as functions of the APD gain and noise, the preamplifier noise, and the shaping amplifier time constant.

Assume the light pulse emitted by the scintillator in Figure 3-16 is an impulse, $\langle N \rangle \delta(t)$. If the bandwidth of the APD is much greater than that of the shaping amplifier, the APD output current, $\langle i(t) \rangle$, is

$$\langle i(t) \rangle = \eta q \langle N \rangle \langle M \rangle \delta(t), \tag{3-6}$$

where η is the APD quantum efficiency, q is the fundamental electron charge, $\langle N \rangle$ is the average number of photons per light pulse and $\langle M \rangle$ is the average gain per pulse of the APD. In frequency space the APD current is

$$\langle I(s) \rangle = \eta q \langle N \rangle \langle M \rangle, \tag{3-7}$$

where $s = j\omega + \sigma$ is the Laplace operator.

The transfer function of the preamplifier and shaping amplifier for an input current is

$$H_i(s) = \frac{A}{C_F \tau_{sh}} \left(\frac{1}{s + \frac{1}{\tau_{sh}}} \right)^2, \tag{3-8}$$

where A is the gain of the shaping amplifier, C_F is the feedback capacitor on the charge sensitive preamplifier and τ_{sh} is the shaping amplifier time constant. The output voltage due to a light pulse is $\langle V_o(s) \rangle = I(s)H(s)$, or

$$\langle V_o(s) \rangle = \frac{\eta q \langle N \rangle \langle M \rangle A}{C_F \tau_{sh}} \left(\frac{1}{s + \frac{1}{\tau_{sh}}} \right)^2. \tag{3-9}$$

Taking the inverse Laplace transform of Equation 3-9 yields

$$\langle v_o(t) \rangle = \frac{\eta q \langle N \rangle \langle M \rangle A}{C_F \tau_{sh}} \cdot t \exp\left(\frac{-t}{\tau_{sh}} \right), \tag{3-10}$$

which may be evaluated to obtain the average output pulse height,

$$\left\langle v_{o,peak} \right\rangle = \frac{\eta q \langle N \rangle \langle M \rangle A}{C_F} \cdot \exp(-1). \tag{3-11}$$

Note that the initial assumption that the scintillator output is a delta function yields an expression for $\langle v_{o,peak} \rangle$ which is independent of τ_{sh}.

The standard deviation of the output noise voltage is determined by the APD dark current, the preamplifier equivalent input noise current and noise voltage and the variance of the product NM. Let us first consider the contribution to the mean-squared output noise voltage due to the parallel noise currents[12]. The transfer function seen by the parallel noise current sources is given in Equation 3-8. Making the substitution $s=j\omega$, the power transfer function may be shown to be

$$\left| H_i(\omega) \right|^2 = \left(\frac{A}{C_F} \right)^2 \left(\frac{\tau_{sh}^2}{1 + 2\omega^2 \tau_{sh}^2 + \omega^4 \tau_{sh}^4} \right). \tag{3-12}$$

For white noise sources it is convenient to use the single sided gain-squared bandwidth,

$$G^2 BW_i = \frac{1}{2\pi} \int_0^{\infty} \left| H(\omega) \right|^2 d\omega = \frac{\tau_{sh} A^2}{8 C_f^2}, \tag{3-13}$$

so that the mean squared output voltage due to a white noise current with single sided power spectral density $i_n^2/\Delta f$ is given by

$$\sigma_{vi}^2 = \frac{i_n^2}{\Delta f} \cdot G^2 BW = \frac{i_n^2}{\Delta f} \left(\frac{\tau_{sh} A^2}{8 C_f^2} \right). \tag{3-14}$$

In Chapter 2 the APD dark current noise spectral density was shown to be

$$\frac{i_{nd}^2}{\Delta f} = 2qI_{db} \langle M \rangle^2 F + 2qI_{ds}, \tag{3-15}$$

where I_{db} is the bulk dark current subject to internal multiplication and I_{ds} is the surface dark current not subject to internal multiplication. From Equation 3-14, the variance of the output voltage due to the APD dark current and the preamplifier input noise current, $i_{na}^2/\Delta f$, is

$$\sigma_{vi}^2 = \left(2qI_{db}\langle M\rangle^2 F + 2qI_{ds} + \frac{i_{na}^2}{\Delta f}\right)\cdot\left(\frac{\tau_{sh}A^2}{8C_F^2}\right). \qquad (3\text{-}16)$$

The second major noise source is the preamplifier equivalent series noise voltage. The transfer function of the input series noise source to the output is

$$H_v(s) = \frac{As}{\tau_{sh}}\cdot\left(\frac{C_F + C_D}{C_F}\right)\cdot\left(\frac{1}{s + \frac{1}{\tau_{sh}}}\right)^2, \qquad (3\text{-}17)$$

and the power transfer function for the series noise voltage is

$$\left|H_v(\omega)\right|^2 = A^2\left(\frac{C_F + C_D}{C_F}\right)^2\left(\frac{\tau_{sh}^2}{1 + 2\omega^2\tau_{sh}^2 + \omega^4\tau_{sh}^4}\right). \qquad (3\text{-}18)$$

Integrating over all positive frequency yields the gain-squared bandwidth for the series input noise voltage,

$$G^2 BW_v = \frac{1}{2\pi}\int_0^\infty \left|H_v(\omega)\right|^2 d\omega = \frac{A^2}{8\tau_{sh}}\cdot\left(\frac{C_F + C_D}{C_F}\right)^2. \qquad (3\text{-}19)$$

The variance of the output voltage due to the preamplifier input noise voltage, $e_{na}^2/\Delta f$, is then

$$\sigma_{vv}^2 = \frac{e_{na}^2}{\Delta f}\left[\frac{A^2(C_D + C_F)^2}{8C_F^2\tau_{sh}}\right]. \qquad (3\text{-}20)$$

The final major contribution to the output noise voltage is the statistical variation in the number of photons emitted by the scintillator and the APD gain seen by each photon. McIntyre has shown that product ηNM is a gaussian random variable having a probability density function of [13]

$$f(\eta NM) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(\frac{-\left(\eta NM - \langle\eta NM\rangle\right)^2}{2\sigma^2}\right), \qquad (3\text{-}21)$$

where

$$\sigma^2 = \langle\eta N\rangle\langle M\rangle^2 F \qquad (3\text{-}22)$$

represents the statistical variance in the number of electrons in the input signal. The mean-squared noise charge in the signal is obtained by multiplying Equation 3-16 by q^2.

The charge gain of the system is given by

$$H_Q = \frac{\left\langle v_{o,peak} \right\rangle}{q\langle \eta N \rangle \langle M \rangle} = \frac{A}{C_F} \exp(-1), \tag{3-23}$$

and the charge power gain is

$$H_Q{}^2 = \left(\frac{A}{C_F} \right)^2 \exp(-2). \tag{3-24}$$

The variance in the peak output voltage due to the variance in the quantity ηNM is then given by

$$\sigma_{\eta NM}{}^2 = \left(\frac{qA}{C_F} \right)^2 \left(\langle \eta N \rangle \langle M \rangle^2 F \right) \exp(-2). \tag{3-25}$$

The total output noise voltage is obtained by taking the square root of the sum of Equations 3-16, 3-20 and 3-25 to obtain

$$\sigma_{vo} = \sqrt{\sigma_{vi}{}^2 + \sigma_{vv}{}^2 + \sigma_{\eta NM}{}^2}. \tag{3-26}$$

The pulse height resolution of the system in Figure 3-16 may be calculated by substituting Equations 3-26 and 3-11 into Equation 3-5.

To illustrate the theoretical effects of each of the noise sources, the PHR for an Advanced Photonix 17 mm diameter APD was calculated as a function of the APD gain and the shaping time constant. The APD parameters used in the calculations were taken from [14] and the preamplifier parameters are typical of commercially available charge sensitive preamplifiers. These parameters are summarized in Table 3-5.

Table 3-5.
Summary of Parameters used in PHR Calculations.

Parameter	Value
Scintillator photon yield	N=1500
APD Quantum Efficiency	η=0.85
APD Capacitance	C_D=150 pF
APD k_{eff}	k_{eff}=0.0017
Un-multiplied APD bulk dark current	I_{db}=0.5 nA
APD surface dark current	I_{ds}=200 nA
Shaping Amplifier Gain	A=100
Preamplifier feedback capacitance	C_F=1 pF
Preamplifier input noise current	$i_{na}/\sqrt{\Delta f}$=50fA/√Hz
Preamplifier input noise voltage	$e_{na}/\sqrt{\Delta f}$=2nV/√Hz

Figure 3-17 shows the calculated pulse height resolution of the 17 mm diameter APD and readout circuit as a function of gain due to (A) all of the sources of statistical variation, (B) the parallel noise current only, (C) the series noise voltage only and (D) the statistical variation of NM. The APD noise contribution to curve C of Figure 3-17 was calculated as a function of gain using the k_{eff} value given in Table 3-5 and Equation 2-89. An 0.5 μs shaping time constant was selected for this curve because commercially available shaping amplifiers such as the ORTEC 570 typically have minimum shaping time constants of 0.5 μs. The composite PHR has a minimum value of 17.7% at M=130. At low gains the PHR is dominated by the amplifier noise and the APD dark current as these noise components are independent of the APD gain. At high gains the PHR is dominated by the APD bulk dark current noise which increases approximately as $M^{3/2}$ while the signal amplitude increases only as M. The dominant noise source over most of the gain range is the parallel noise current, principally the APD bulk dark current. Any step which reduces the APD bulk dark current, such as reducing the APD area, reducing the p-depletion region thickness above the junction or lowering the device operating temperature, will improve the system PHR.

Figure 3-18 shows the calculated pulse height resolution as a function of the CR-RC shaping time constant with the APD biased for a gain of 100. Again, the PHR depicted by (A) is due to all of the sources of statistical variation, (B) is due to the parallel noise current only, (C) is due to the series noise voltage only and (D) is due to the statistical variation of *NM*. At short shaping time constants the PHR is dominated by the series noise voltage of the preamplifier while at long time constants the PHR is dominated by the parallel noise current of the preamplifier and APD. The optimal calculated PHR (14.8%) for an impulse signal occurs at the time constant for which the series noise voltage and parallel noise current are equal (τ_{sh}=170ns.) In practice the finite decay time constant of the scintillator (~300 ns for BGO) makes the optimal shaping constant somewhat longer than 170 ns.

By way of comparison, the pulse height resolution which would be expected for a PMT with 20% quantum efficiency coupled to a scintillator yielding 1500 photons would be 13.5%, where the resolution is dominated by the BGO statistics.

Pulse Height Resolution: Experimental Data

The manufacturer's specifications for the three APDs evaluated are summarized in Table 3-6. The beveled edge devices have lower total dark currents than the reach-through device and the values reported by API and EG&G indicate the beveled-edge device also has a lower k_{eff}. The k_{eff} values reported by the manufacturers are lower than those predicted by the simulations, probably due to fabrication enhancements not considered in the calculations.

The test system used to measure the APD PHR is shown in Figure 3-19. The APDs were soldered directly to the input FET of the ORTEC 142A preamplifier to

Table 3-6.
Summary of Manufacturers' Data for APDs Under Study.

Parameter	API	RMD	EG&G
Part Number	VG-17-10	SH2S	C30994E
Structure	beveled-edge	beveled-edge	reach-through
Area	20 mm^2	4 mm^2	9 mm^2
Capacitance	16 pF	5 pF	12 pF
keff	0.0015	not reported	0.006
Dark current	75 nA	60 nA	200 nA

minimize parasitic capacitance. The shaping amplifier is an ORTEC 572 (CR-RC^4 semi-gaussian) modified to include a 250 ns time constant and the multi-channel analyzer is a Nucleus PCA-II.

Direct comparison of the three devices was complicated by their different surface areas and the geometry's of the BGO crystals available for the experiment. The 5 mm diameter opening in the Advanced Photonix packaged permitted good coupling between the APD and the 3 x 3 mm^2 BGO crystals available for this study. A 3 x 3 x 25 mm^3 Teflon wrapped crystal was used for the measurements. The RMD device area is less than half the area of the available crystals, so a shorter 3 x 3 x 10 mm^3 Teflon wrapped crystal was used to compensate for the light lost due to the area mismatch. (The 10 mm crystal emits approximately 75% more light than the 25 mm crystal). The EG&G device came pre-packaged with a 3 x 3 x 20 mm^3 aluminum foil wrapped BGO crystal. Despite these differences, the experimental measurements provide insight into the relative qualities of the different discrete APDs.

For all three of the devices studied, the best pulse height resolution was obtained with a 250 ns shaping time constant, the minimum shaping time constant available. A plot of measured PHR vs. bias voltage for the API beveled-edge APD using the 250 ns shaping time constant is shown in Figure 3-20. The minimum measured PHR is 14.2%

obtained at a bias of 2560V, where M $\approx$ 100 (Figure 3-21). The optimal operating voltage of the API device is higher than the bias voltage calculated for the "model" beveled-edge APD because API currently uses a higher resistivity starting material and a deeper gallium diffusion than the early devices described in the literature [8]. The lack of smoothness in the measured data in Figure 3-20 is attributed to the non-gaussian shape of the energy peak; the MCA calculates the FWHM via a gaussian fit of the peak. Several interesting features of the spectrum in Figure 3-21 should be noted. The "bump" in left hand side of the energy peak is the x-ray escape peak located 76 keV below the 511 keV photopeak. The noise floor is approximately 40 keV and the peak observed at 170 keV is the back-scatter peak. The 14.2% measured PHR and low noise floor indicate the API beveled-edge APD has both a low k_{eff} and a low bulk dark current. The measured PHR also agrees well with the theoretical PHR obtained in the sample calculation (Figure 3-18).

The measured PHR vs. bias voltage for the RMD beveled-edge APD using a 250 ns shaping time constant is shown in Figure 3-22. The best PHR (17.2%) is measured at 1440V. The close agreement between the RMD operating voltage and that predicted for the modeled beveled-edge APD suggests the model process parameters closely match the RMD process. The noise floor of the RMD device is approximately 100 keV, and the x-ray escape peak and back-scatter peak are much less clearly resolved in the PHR spectrum (Figure 3-23). The poorer resolution of the RMD device suggest the structure has a higher k_{eff} and bulk dark current than the API device.

The measured PHR vs. bias voltage for the EG&G reach-through APD/BGO module using a 250 ns shaping time constant is shown in Figure 3-24. The best PHR (19.1%) is observed at 520V. Again, the close agreement between the EG&G beveled-edge operating point and that of the model suggests close agreement between the model and EG&G's process. The noise floor for the reach-through device is approximately 180 keV, masking the back-scatter peak, and the x-ray escape peak is masked by the noise in the energy peak. As predicted by the model, the EG&G device has the poorest PHR, suggesting a higher k_{eff} and bulk dark current than the API and RMD devices.

General agreement between the predictions of the model and the experimental data is observed. In particular, the API beveled-edge device shows significantly better pulse height resolution and lower noise floor than the other devices. Care must be taken in reaching general conclusion from this data, however, because both the device geometry's and the crystal geometry's are different for each of the APDs studied.

Timing Resolution: Background and Example Calculations

If the output of the signal processing circuitry shown in Figure 3-16 is connected to an ideal voltage discriminator with threshold voltage v_{th}, the resolution with which the discriminator can determine the time that $v_o(t)$ crosses v_{th} is given by

$$\sigma_t = \frac{\sigma_{vo}}{\left. {dv}/{dt} \right|_{v_o = v_{th}}}, \tag{3-27}$$

where σ_{vo} is the standard deviation of the timing measurement distribution, σ_v is the rms noise voltage at the discriminator input and dv/dt is the mean signal slope at the discriminator input when the signal crosses v_{th}. As with the pulse height resolution, the timing resolution is typically specified as a FWHM value,

$$t_{FWHM} = 2.35\sigma_t. \tag{3-28}$$

The first step in obtaining an expression for the timing resolution is to calculate the signal slope at the timing point ($v_o = v_{th}$). Unlike the pulse height resolution measurement which samples virtually all of the scintillator output pulse, the timing measurement typically is concluded in the first few nanoseconds of the signal.

Consequently it is convenient to model the output of the APD/BGO detector as a step function [15],

$$i(t) = q\langle r\rangle\langle M\rangle u(t) \tag{3-29}$$

where

$$\langle r\rangle = \frac{\eta\langle N\rangle}{\tau_{BGO}} \tag{3-30}$$

is the initial photon rate (attenuated by the quantum efficiency) entering the APD and τ_{BGO} (~300 ns) is the decay time constant of the light emitted by BGO.

The time domain transfer function of the charge sensitive preamplifier and CR-RC shaping amplifier is

$$h_i(t) = \frac{A}{C_F\tau_{sh}}\cdot t\exp\left(\frac{-t}{\tau_{sh}}\right). \tag{3-31}$$

The frequency domain output of the preamplifier and shaping amplifier due to the assumed step function output of the detector is

$$V_o(s) = \frac{q\langle r\rangle\langle M\rangle}{s}\cdot\frac{A}{C_F\tau_{sh}}\left(\frac{1}{s+\frac{1}{\tau_{sh}}}\right). \tag{3-32}$$

It is convenient to calculate *dv/dt* in the frequency domain by recalling the Laplace operation

$$\mathbf{L}\left(\frac{dv}{dt}\right) \Leftrightarrow sV(s) \tag{3-33}$$

where v(t) and V(s) are a Laplace transform pair. Thus, the Laplace transform of the derivative of the output signal is

$$\mathbf{L}\frac{dv}{dt} = s\left[\frac{q\langle r\rangle\langle M\rangle}{s}\cdot\frac{A}{C_F\tau_{sh}}\cdot\left(\frac{1}{s+\frac{1}{sh}}\right)^2\right] = \frac{q\langle r\rangle\langle M\rangle A}{C_F\tau_{sh}}\cdot\left(\frac{1}{s+\frac{1}{sh}}\right)^2. \tag{3-34}$$

Note that Equation 3-34 has the same form as the impulse response of the preamplifier and shaping amplifier. Performing the inverse Laplace transform yields

$$\frac{dv}{dt} = \frac{q\langle r\rangle\langle M\rangle A}{C_F} \cdot \frac{t}{\tau_{sh}} \exp\left(\frac{-t}{\tau_{sh}}\right), \tag{3-35}$$

which may be evaluated to obtain the maximum slope of the output signal,

$$\left.\frac{dv}{dt}\right|_{\max} = \frac{q\langle r\rangle\langle M\rangle A}{C_F} \cdot \exp(-1). \tag{3-36}$$

The maximum slope occurs when $t = \tau_{sh}$, and interestingly, the maximum slope is independent of the shaping time constant. Thus, increasing the system bandwidth by reducing τ_{sh} has no effect on the maximum signal slope so long as $\tau_{sh} << \tau_{BGO}$ so that the initial assumption of a step signal input valid.

The next step is to develop an expression for the mean-squared noise at the output of the shaping amplifier. The contributions to the output noise voltage due to the parallel noise currents at the preamplifier input, σ_{vi}, and due to the series equivalent input noise voltage of the preamplifier, σ_{vv}, were calculated in the discussion of pulse height resolution and are repeated below:

$$\sigma_{vi}^2 = \left(2qI_{db}\langle M\rangle^2 F + 2qI_{ds} + \frac{i_{na}^2}{\Delta f}\right) \cdot \left(\frac{\tau_{sh} A^2}{8 C_F^2}\right). \tag{3-16}$$

$$\sigma_{vv}^2 = \frac{e_{nv}^2}{\Delta f}\left[\frac{A^2 (C_D + C_F)^2}{8 C_F^2 \tau_{sh}}\right]. \tag{3-20}$$

The expression previously derived for the output noise contribution due to the detector statistics is not valid for timing, however, because only a fraction of the total signal has been processed when the timing measurement is made. Instead, the *time variant* noise of the detector must be calculated following the procedure outlined by Blalock and Nowlin [16].

The "statistical noise" due to the random nature of the scintillator photon rate, r, and the APD gain, M, may be treated as a white noise superimposed on the mean output signal of the APD/BGO detector. Because this noise source only exists in the presence of the signal, it is useful to model the detector noise as a zero-mean time-invariant white noise source, σ_{irM}, which is modulated by the function *m(t)* and shaped by the system transfer function, *h(t)* as shown in Figure 3-26. The detector signal is being treated as a step function, so $m(t) = u(t)$. The output noise of the system due to the statistics of the photon emission rate and APD gain is now a time variant noise voltage, $\sigma_{vrM}(t)$ given by [16]

$$\sigma_{vrM}{}^2(t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(u)h(v)m(t-u)m(t-v)R_{irM}(u-v)dudv, \tag{3-37}$$

where $R_{irM}(\tau)$ is the autocorrelation of the input noise current. The limits of integration in Equation 3-37 may be modified by assuming *h(t)* is causal (*ie* $h(t) = 0$ for $t < 0$) to obtain

$$\sigma_{vrM}{}^2(t) = \int_0^{\infty}\int_0^{\infty} h(u)h(v)m(t-u)m(t-v)R_{irM}(u-v)dudv. \tag{3-38}$$

The integration limits may be further modified by noting $m(t) = u(t)$ so that $m(t-u) = 0$ for $u > t$ and $m(t-v) = 0$ for $v > t$ to obtain

$$\sigma_{vrM}{}^2(t) = \int_0^{t}\int_0^{t} h(u)h(v)R_{irM}(u-v)dudv. \tag{3-39}$$

The input noise current has been defined as a white noise source, so the autocorrelation function is known to have the form

$$R_{irM}(\tau) = \sigma_{irM}{}^2\delta(t). \tag{3-40}$$

Substituting Equation 3-40 into Equation 3-39 yields

$$\sigma_{vrM}{}^2(t) = \sigma_{irM}{}^2\int_0^{t}\int_0^{t} h(u)h(v)\delta(u-v)dudv. \tag{3-41}$$

Let us now consider the inner integral of Equation 3-41. By the sifting property of the delta function, the inner integral has a solution equal to the integrand evaluated at $u=v$, or

$$\int_0^t h(u)h(v)\delta(u-v)du = h^2(v). \tag{3-42}$$

Substituting Equation 3-42 back into Equation 3-41 yields the expression for the time variant mean squared output voltage due to the statistics of the scintillator and APD gain,

$$\sigma_{vrm}{}^2(t) = \sigma_{irM}{}^2 \int_0^t h^2(v)dv. \tag{3-43}$$

Note that Equation 3-43 is identical to the well known Campbell's theorem for shot noise.

An expression for σ_i^2 may be obtained by rewriting McIntyre's density function for the output of an APD/scintillator detector (Equation 3-21) in terms of the scintillator photon rate, r,

$$f(rM) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(rM - \langle rM\rangle)^2}{\sigma^2}\right] \tag{3-44}$$

where

$$\sigma^2 = \langle r\rangle\langle M\rangle^2 F. \tag{3-45}$$

As before, σ represents a variance in the photon emission rate. The variance in Equation 3-45 may be converted into a mean-squared current by multiplying the quantity by q^2 to obtain

$$\sigma_{irM}{}^2 = q^2\langle r\rangle\langle M\rangle^2 F. \tag{3-46}$$

Substituting Equations 3-46 and 3-31 into Equation 3-43 yields

$$\sigma_{vrM}^2(t) = q^2 \langle r \rangle \langle M \rangle^2 F \cdot \left(\frac{A}{C_F \tau_{sh}} \right)^2 \int_0^t \left[t \exp\left(\frac{-t}{\tau_{sh}} \right) \right]^2 dv, \tag{3-47}$$

and evaluating the integral yields the expression for the time varying mean-squared output noise voltage due to the detector statistics,

$$\sigma_{vrM}^2(t) = \frac{q^2 \langle r \rangle \langle M \rangle^2 F A^2}{C_F^2 \tau_{sh}^2} \left[\frac{\tau_{sh}^3}{4} - \frac{\left(2t^2 \tau_{sh} + 2t\tau_{sh}^2 + \tau_{sh}^3\right)}{4} \exp\left(\frac{-2t}{\tau_{sh}} \right) \right]. \tag{3-48}$$

For timing at the maximum signal slope, let $t = \tau_{sh}$ to obtain

$$\sigma_{vrM}^2(t = \tau_{sh}) = \frac{q^2 \langle r \rangle \langle M \rangle^2 F A^2}{C_F^2 \tau_{sh}^2} \left[\frac{\tau_{sh}^3}{4} - \frac{5\tau_{sh}^3}{4} \exp(-2) \right], \tag{3-49}$$

or

$$\sigma_{vrM}^2(t = \tau_{sh}) \approx \frac{q^2 \langle r \rangle \langle M \rangle^2 F A^2}{C_F^2 \tau_{sh}^2} \left(0.081 \tau_{sh} \right). \tag{3-50}$$

The total output noise voltage is obtained by taking the square root of the sum of Equations 3-16, 3-20 and 3-49 to obtain

$$\sigma_{vo} = \sqrt{\sigma_{vi}^2 + \sigma_{vv}^2 + \sigma_{vrM}^2}. \tag{3-51}$$

The timing resolution of the system in Figure 3-16 may be calculated by substituting Equations 3-36 and 3-51 into Equation 3-28.

The effects of each of the noise sources is illustrated by calculating the timing resolution for the 17 mm diameter Advanced Photonix APD described in the pulse height resolution section of this chapter. As before, the timing resolution is calculated as a function of the APD gain and the shaping time constant. The parameters used in the calculations are summarized in Table 3-5.

Figure 3-27 shows the calculated timing resolution for the 17 mm diameter APD and readout circuit as a function of gain due to (A) all of the noise sources, (B) the parallel noise current only, (C) the series noise voltage only and (D) the statistical variation of $r{\cdot}M$ (Equation 3-47). The calculated timing resolution is strongly dependent on the APD gain, with the best resolution of 14 ns FWHM occurring at the maximum APD gain of 1000. The timing resolution is dominated by the equivalent noise voltage of the preamplifier due to the large capacitance (150 pF) of the APD.

Figure 3-28 shows the calculated timing resolution as a function of the CR-RC shaping time constant for an APD gain of 500. Again the plot shows the calculated resolution due to (A) all of the noise sources, (B) the parallel noise current only, (C) the series noise voltage only and (D) the statistical variation of $r{\cdot}M$. The optimal timing resolution (23 ns FWHM) occurs at a shaping time constant of 18.5 ns. This shaping constant represents the optimum compromise between the noise contribution of all three noise sources.

Modern PET scanners typically use detectors with timing resolutions on the order of 3 ns FWHM. From the data in Figures 3-27 and 3-28 it is clear that no combination of bias voltage and shaping time constant exists for which the 17 mm diameter APD timing resolution approaches 3 ns. The area of the APD used in this model is much larger than would be used in a PET detector since the scintillators used in a PET detector typically have areas on the order of 10 mm^2. To evaluate the effects of reducing the APD area, the expected timing resolution of a 5 mm diameter APD was calculated by scaling the data in Table 3-5. Assuming the detector capacitance and bulk dark current are proportional to the device area and the surface current is proportional to the device perimeter, a 5 mm diameter APD should have a capacitance of C_D=13 pF, a bulk dark current of I_{db}=43 pA and a surface dark current of I_{ds}=60 nA. The new detector capacitance is on the order of the input capacitance of most low noise preamplifiers (~10 pF) so an effective detector capacitance of C_D=23 pF is used.

Figure 3-29 shows the calculated timing resolution for the 5 mm diameter APD and readout circuit as a function of gain where again the timing resolution due to (A) all of the noise sources, (B) the parallel noise current only, (C) the series noise voltage only and (D) the statistical variation of $r{\cdot}M$ are shown As before, the calculated timing resolution is strongly dependent on the APD gain, with the best resolution (5.9 ns FWHM) occurring at an APD gain of 700. Once again the timing resolution is dominated by the equivalent noise voltage of the preamplifier, but the effect is greatly diminished due to the reduction in the detector capacitance. From this data it is clear that the APD must be operated near its maximum gain and the APD capacitance and preamplifier input capacitance must be minimized to achieve optimal timing resolution.

Figure 3-30 shows the calculated timing resolution as a function of the CR-RC shaping time constant for an APD gain of 500. A much shorter shaping time constant is required for optimal timing with the 5 mm device, however, due to the reduction in diode capacitance. The optimal timing resolution (5.7 ns FWHM) occurs at a shaping time constant of 6 ns. This data shows that timing resolutions on the order of 5 ns may be achieved with commercially available APDs and suggests that with further reductions in the APD capacitance and in the preamplifier input noise voltage timing resolutions similar to those achieved with PMT based detectors may be attained.

Comparison of the PHR and timing resolution calculations shows that the optimal APD bias voltages and shaping amplifier time constants are quite different for the two applications. Of course, the conflict in optimal shaping constant is easily resolved by placing two shaping amplifiers in parallel at the output of the charge sensitive preamplifier, but the APD bias voltage cannot be simultaneously optimal for both PHR and timing. Because timing resolution is the more critical parameter for PET, it is expected that the APD will be biased for optimal timing at the expense of the pulse height resolution.

Timing Resolution: Experimental Data

To verify the calculated data, timing histograms were experimentally measured for the three test APDs (Table 3-6) in coincidence with a plastic scintillator coupled to a photomultiplier tube. The test system used to measure the timing resolution is shown in Figure 3-31. The system is a standard coincidence measurement system [17] consisting of a fast plastic channel, a BGO channel which includes the device under test and an energy channel. The fast plastic detector consists of a plastic scintillator coupled to an RCA 8575 PMT and biased with an ORTEC 265 tube base. The PMT output is fed into the 50 Ω input of a Tennelec 455 constant fraction discriminator (CFD), and the output of CFD triggers the START input of the ORTEC 457 time to amplitude converter (TAC). Plastic scintillators typically have timing resolutions less than 250 ps for 511 keV gamma rays [18] and are frequently used in test systems with BGO because their contribution to the system timing resolution is negligibly small, making the measured timing resolution that of the BGO channel alone.

The BGO/APD channel has the APD soldered directly to the input FET of an ORTEC 142A charge sensitive preamplifier. The preamplifier output is fed into a CR-RC shaping amplifier with variable time constant. The shaping amplifier is fed into a Tennelec 455 CFD whose output passes through a variable delay line into the STOP channel of the TAC. The TAC output feeds a Nucleus PCA-II multi channel analyzer.

As expected, the best timing resolution is observed when the APD is biased for maximal gain, near its break-down voltage, and when a short shaping time constant ($\leq$ 20 ns) is used. Under these conditions, the APD dark current pulses have amplitudes comparable to the signal pulses. An energy channel is therefore added to the BGO/APD side of the system to discriminate between the noise pulses and the signal. The energy channel consists of a modified ORTEC 572 shaping amplifier with 250 ns shaping time constant whose output drives an ORTEC 553 single channel analyzer (SCA). The SCA is configured to provide an output pulse when the energy pulse height is 511 keV ±50 keV. The SCA output drives an ORTEC 416A gate and delay which in turn gates the MCA. Thus the MCA only generates an output when the BGO/APD channel sees a 511 keV photoelectric event. The energy channel also prevents the system from responding to the 1.274 MeV gamma rays emitted by the ^{22}Na source.

Figure 3-32 is a plot of the measured timing resolution vs. bias voltage for the standard 5 mm diameter API APD coupled to a 3 x 3 x 25 mm^3 Teflon wrapped BGO crystal. The optimal shaping time constant for the 5 mm diameter API APD is found to be 2.2 ns, the shortest time constant available for the CR-RC amplifier. The best measured timing resolution is 9.0 ns FWHM obtained when the APD is biased at 2700V, 15 V below the measured breakdown voltage. The measured timing resolution of the 5 mm diameter APD device is poorer than the 5.7 ns FWHM resolution predicted by the simulations (Figure 3-30). Discussions with Advanced Photonix [19] revealed that their standard antireflective (AR) coating process introduces shallow traps at the silicon/AR coating interface. The traps collect between 10 and 50 % of the photogenerated carriers and release them with a 100-500 ns time constant. Advanced Photonix delivered a second prototype APD fabricated using a different AR coating process. Preliminary measurements showed the device to a PHR of 15% FWHM at a bias voltage of 2340 V and timing resolution of 7.6 ns FWHM at a bias voltage of 2400 V (Figure 3-33). The improved timing resolution is closer to the value predicted by the model and is believed to be the best timing resolution reported to date for APDs coupled to BGO crystals suitable for PET (length > 20 mm). (Carrier and Lecomte have reported a timing resolution of 6.5 ns FWHM using a 4 mm long BGO crystal [20].) The device failed during the early stages of the test, however, due to a permanent breakdown on the surface of the beveled edge. This device was not discussed in the PHR section of this report because it failed before a full set of PHR data could be obtained.

The measured timing resolution vs. bias voltage for the RMD device coupled to a 3 x 3 x 10 mm^3 Teflon wrapped BGO crystal is shown in Figure 3-34. The optimal shaping time constant for the RMD device is 20 ns, suggesting this device has a somewhat longer rise-time than the API device, perhaps due to a wider undepleted p-region. The best measured timing resolution (14.8 ns FWHM) is obtained at 1475V, 3 V below the measured breakdown voltage.

Like the API devices, the EG&G BGO/APD module exhibits the best timing resolution at a 2.2 ns shaping time constant. The measured timing resolution vs. bias voltage for the EG&G module is shown in Figure 3-35 where the best value (10.5 ns FWHM) is obtained at a bias of 550V, 10V below breakdown.

The experimental pulse height and timing resolution data is summarized in Table 3-7. The best resolutions were obtained with the API beveled-edge devices. This finding is consistent with the data from the modeled structures which indicated the beveled-edge structure has inherently superior noise characteristics. The measured pulse height resolution for the RMD beveled-edge device is poorer than that of the API device but superior to that of the EG&G device, suggesting the RMD device has a k_{eff} between those of the API beveled-edge devices and the RMD reach-through device. The timing resolution of the RMD device is poorer than that of the other two devices. This suggests the undepleted p-region in the RMD structure is wider than that of the API device, limiting the bandwidth of the device. Interestingly, although the pulse-height resolution of the EG&G device is the poorest of the structures studied and its timing resolution is somewhat poorer than that of the API device, it is the only one of the devices to have been used in PET. Lecomte *et al* have built a prototype PET scanner using the EG&G APD/BGO module studied here and have reported promising initial results [21-24].

Reliability

For APDs to be viable PET detectors, they must be able to operate continuously for many years without significant performance degradation. The first API and RMD devices studied failed during the initial weeks of the test and were replaced by the manufacturers. During the course of this investigation, the replacement API and RMD devices also suffered performance degradation due to increased surface leakage current. This degradation could be partially reversed with a 50C over-night bake. Both RMD and API believe the degradation is largely due to limitations in their packaging technology and expect future devices to have improved reliability. In contrast, the EG&G detector has been operated intermittently over the last 3 years with no performance degradation.

Table 3-7.
Summary of Measured Data for Discrete APDs.

Parameter	**API 5 mm dia.**	**API 5x5 mm^2**	**RMD**	**EG&G**
BGO size	3x3x25 mm^3	3x3x25 mm^3	3x3x10 mm^3	3x5x20 mm^3
$V_{Breakdown}$	2715 V	2420 V	1478 V	560 V
PHR V_{bias}	2560 V	2340 V	1440 V	520 V
Best PHR	14.2% FWHM	15% FWHM	17.2% FWHM	19.1 %
Timing V_{bias}	2700 V	2400 V	1475 V	550 V
Timing τ_{sh}	2.2 ns	2.2 ns	20 ns	2.2 ns
Best Timing Resolution	9.0 ns FWHM	7.6 ns FWHM	14.8 ns FWHM	10.5 ns

Vendor Selection

The goal of the experimental comparison of the three devices was to make an unbiased comparison of the technologies of the three major APD vendors. As noted above, however, each device was studied with a different scintillator and, as a result, some of the differences in the measured data must be due to differences in the scintillator and in the coupling between the scintillator and the APDs. Despite this uncertainty in the measurements, the data provides clear evidence that the API devices have the best available performance characteristics.

On the other hand, the API devices suffered from significant reliability problems. As described above, one of the two devices studied failed in the first two weeks of operation and the second suffered some performance degradation over the two-month investigation period.

In selecting the array vendor, the clear performance advantages of the API devices were weighed against the observed reliability problems. Because this study is intended to be a proof-of-principle investigation rather than a design-for-production, and because API expressed confidence that the reliability problems will be eliminated with improved packaging, Advanced Photonix was selected to develop a 64-element custom APD array for use in the prototype PET detector.

Conclusions

The two most commonly used APD structures, beveled-edge and reach-through, have been theoretically modeled and experimentally evaluated for use with BGO. Both the theoretical and experimental data indicate the beveled-edge APD has inherently lower noise characteristics than the reach-through structure due to its lower effective ionization coefficient ratio, k_{eff}, and its lower bulk dark current. APD pulse height resolution is dominated by the APD noise, and the superior noise characteristics of the reach-through APD are clearly seen in pulse height resolution data taken with commercially available APDs of both profiles.

APD timing resolution is determined by the APD and preamplifier noise, the detector statistics, and rise-time characteristics. If the undepleted p-region of the beveled-edge APD is not minimized (as is suspected to be the case with the RMD device), the rise-time will be increased, leading to a degradation in timing resolution. Experimentally, the timing resolution of the API beveled-edge APDs is found to be superior to that of the EG&G reach-through APD, while the timing resolution of the RMD beveled-edge APD is found to be poorer than that of the reach-through APD. The disparity between the performance of the two beveled-edge APD timing performance is attributed to differences in the undepleted p-region thickness.

Despite their performance advantages, both of the commercially available beveled-edge APDs exhibited reliability problems. These problems must be resolved before beveled-edge APDs may be incorporated into commercial PET detectors. Based upon the theoretical and experimental data, the manufacturer of the best beveled-edge APD, Advanced Photonix, Inc., was selected to fabricate a custom 64-element APD array for use in a prototype PET detector. The details of the custom APD array design are presented in Chapter 4.

References

1. S. E. Hansen and M. D. Deal, Ed., *Suprem-IV.GS Two Dimensional Process Simulation for Silicon and Gallium Arsenide*, Stanford University, 1993.
2. *FISH1D 2.0 User's Manual*, Purdue Research Foundation, 1988.
3. G. C. Huth, J. B. Trice and R. A. McKinney, *Rev. Sci. Inst.*, **35**, p. 1220, 1964.
4. G. C. Huth, J. B. Trice and J. A. Shannon, *IEEE Trans. Nucl. Sci.*, **NS-12,** pp. 275-280, 1965.
5. Gerald C. Huth and Russell A. McKinney, "Radiation Detector", U.S. Patent 3,449,177, June 10, 1969.
6. Gerald C. Huth, "Silicon Avalanche Photodiode Array," U.S. Patent 5,021,854, June 4, 1991.
7. S. M. Sze, *Physics of Semiconductor Devices*, 2nd Edition, New York: John Wiley and Sons, 1981.
8. Shawn Fagen, Advanced Photonix APD Array Product Manager, private communication.
9. Robert J. McIntyre, private communication.
10. Robert J. McIntyre and Paul P. Webb, "Avalanche Photodiode Array," U.S. Patent 4,458,260, July 3, 1984.
11. Alexander W. Lightstone, Paul P. Webb and Robert J. McIntyre, "Avalanche Photodiode," U.S. Patent 4,654,678, March 31, 1987.
12. W. Marshall Leach, Jr., "Fundamentals of Low-Noise Analog Circuit Design," *Proc. IEEE*, **82**, pp. 1515-1538, 1994.
13. P. P. Webb, R. J. McIntyre, and J. Conradi, "Properties of Avalanche Photodiodes," *RCA Review*, **35**, pp. 235-278, 1974.
14. "Noise Characteristics of Advanced Photonix Avalanche Photodiodes," Advanced Photonix Application Note API/NOIS/1291/B, 1991.
15. David M. Binkley, *Development and Analysis of Non-Delay-Line Constant-Fraction Discriminator Timing Circuits, Including A Fully Monolithic CMOS Implementation*, Ph. D. Dissertation, The University of Tennessee-Knoxville, 1992.
16. T. V. Blalock and C. H. Nowlin, "Time Variant Filters for Generating Bipolar Pulses with Signal-to-Noise Ratios of Unipolar Pulses in Nuclear Pulse Amplifiers," *The Review of Scientific Instruments*, **40**, pp. 1093-1100, 1969.
17. "Principles and Applications of Timing Spectroscopy," EG&G ORTEC Application Note AN-42.
18. T. J. Paulus, EG&G ORTEC, private communication.
19. Joe Boisvert, Advanced Photonix Member of Technical Staff, private communication.
20. C. Carrier and R. Lecomte, "Timing Performance of Scintillators Read Out by Silicon Avalanche Photodiodes," *Nucl. Instr. and Meth.*, A299, pp. 115-118, 1990.
21. R. Lecomte, D. Schmitt, A.W. Lightstone, and R. J. McIntyre, "Performance Characteristics of BGO-Silicon Avalanche Photodiode Detectors for PET," *IEEE Trans. Nucl. Sci.*, **NS-32**, pp. 482-486, 1985.

22. C. Carrier, C. Marel, D. Schmitt and R. Lecomte, "Design of a High Resolution Positron Emission Tomograph Using Solid State Scintillation Detectors," *IEEE Trans. Nucl. Sci*, **NS-35**, pp. 685-690, 1988.

23. R. Lecomte, J. Cadorette, A. Jouan, M. Heon, D. Rouleau, and G. Gauthier, "High Resolution Positron Emission Tomography with a Prototype Camera Based on Solid State Scintillation Detectors," *IEEE Trans. Nucl. Sci.*, **NS-37**, pp. 805-811, 1990.

24. R. Lecomte, J. Cadorette, S. Rodrigue, M. Heon, D. Rouleau, P. Richard, M. Bentourkia, and P. Msaki, "A PET Camera Simulator with Multispectral Acquisition Capabilities," *IEEE Trans. Med. Imag.*, **11**, pp. 1628-1634, 1992.

Figures

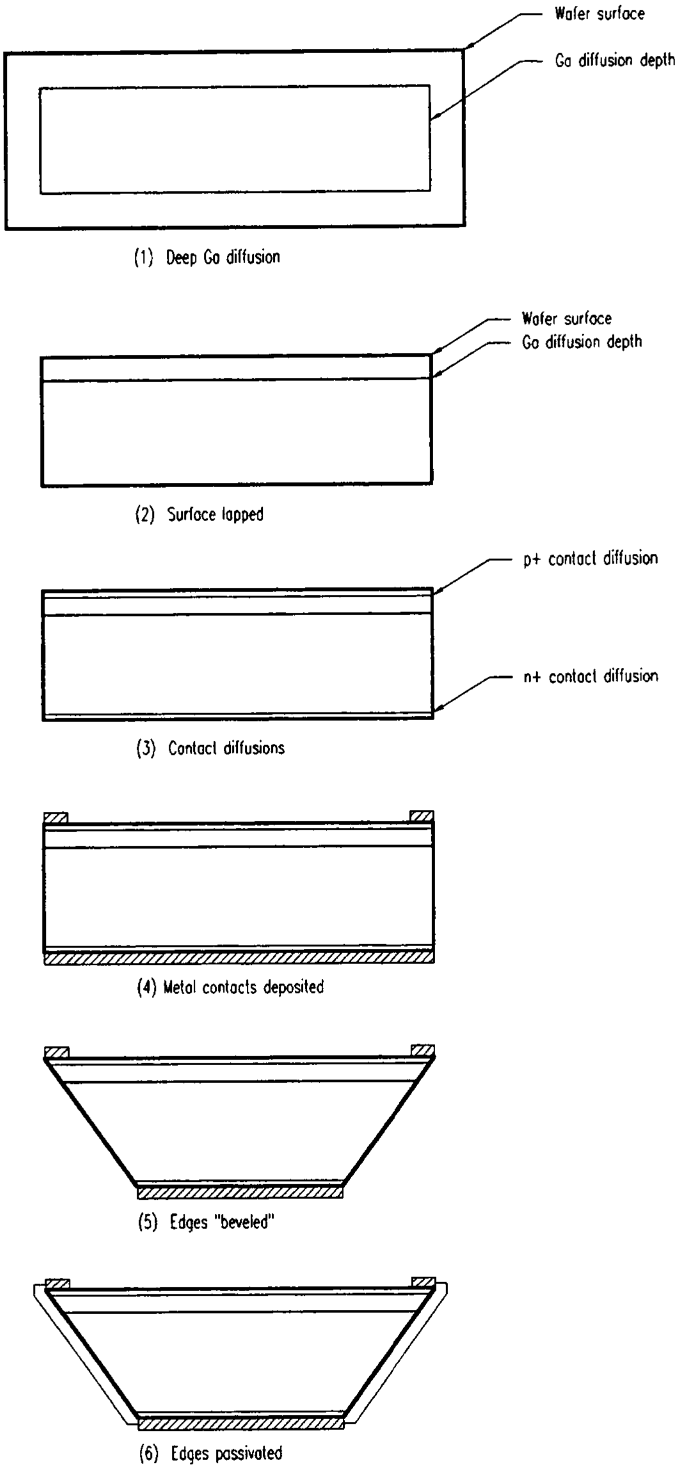

Figure 3.1 Basic fabrication steps for the beveled-edge APD.

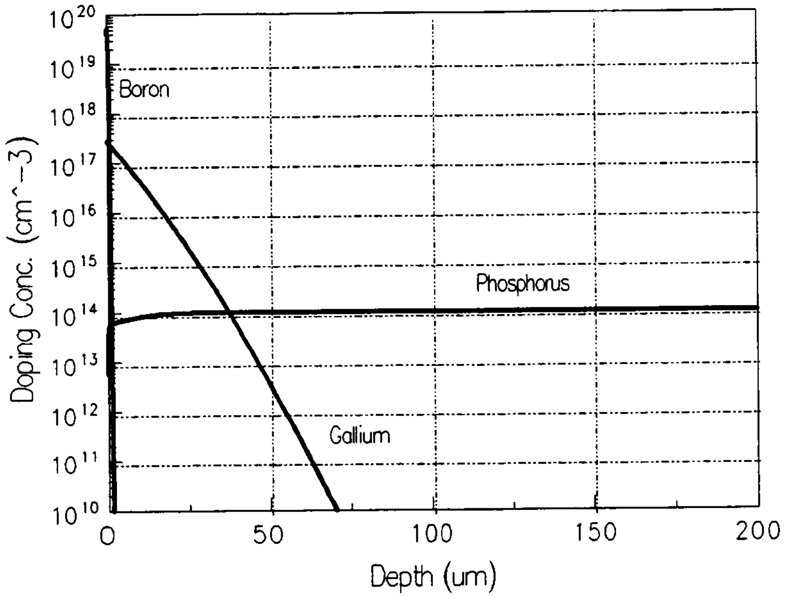

Figure 3-2: SUPREM calculated doping profile for beveled-edge APD.

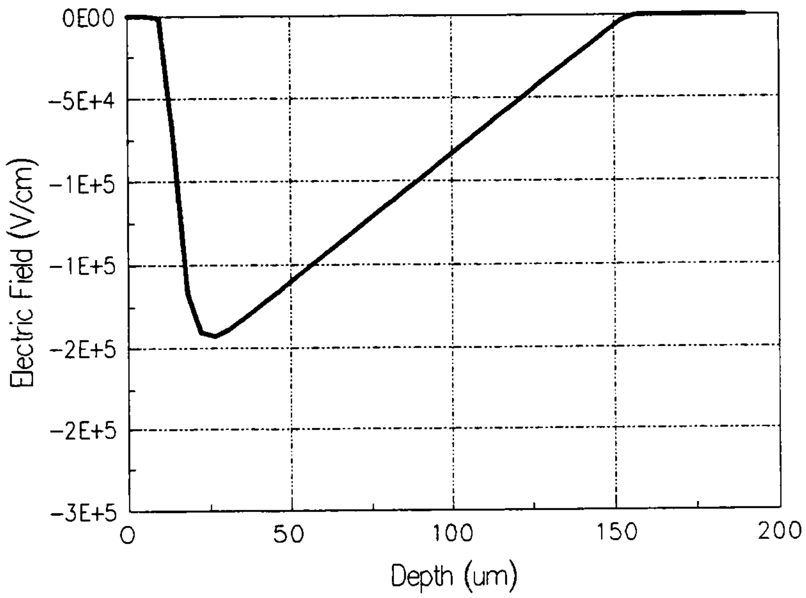

Figure 3-3. Calculated electric field profile for a typical beveled-edge APD biased at 1475V. The calculated gain for this structure is M≈210.

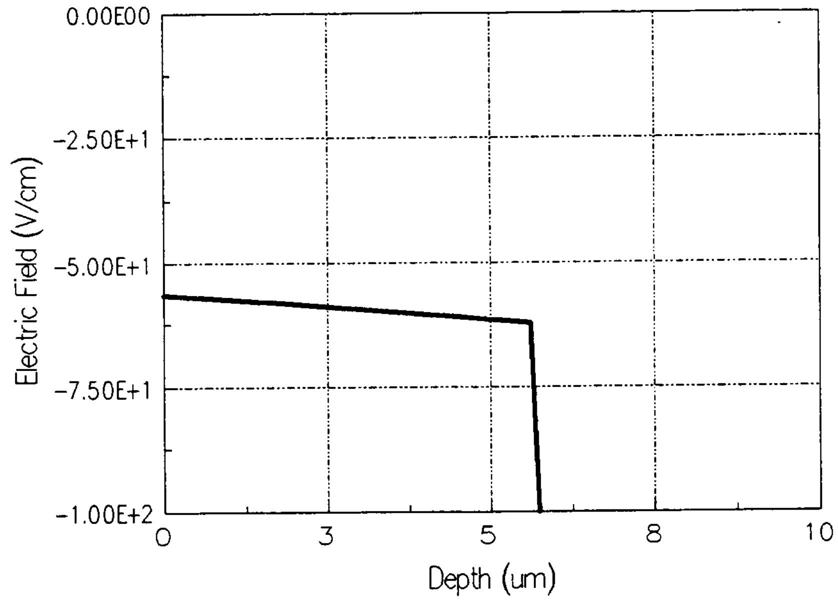

Figure 3-4. Expanded view of calculated electric field profile for beveled-edge APD showing the undepleted p-region. The APD is biased at 1475V for a gain of 210.

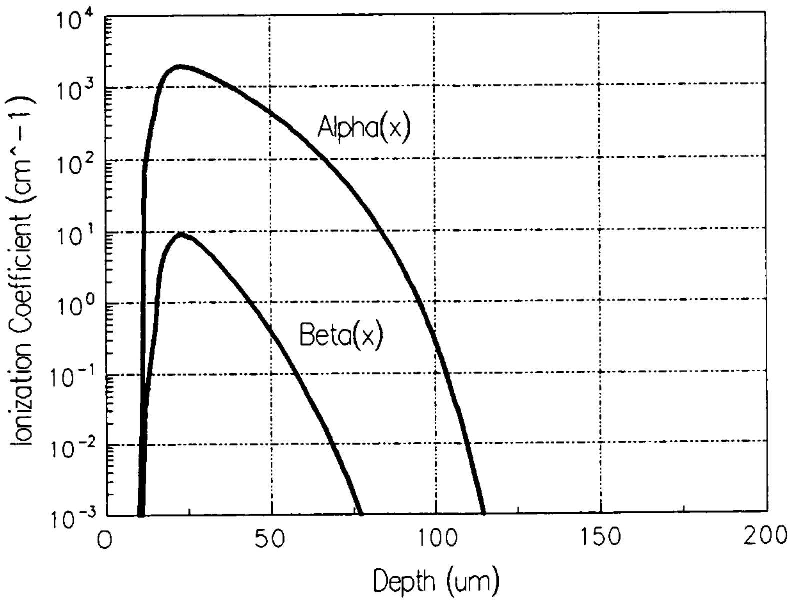

Figure 3.5. Calculated ionization coefficients for beveled-edge APD biased for M≈210 (V=1475V).

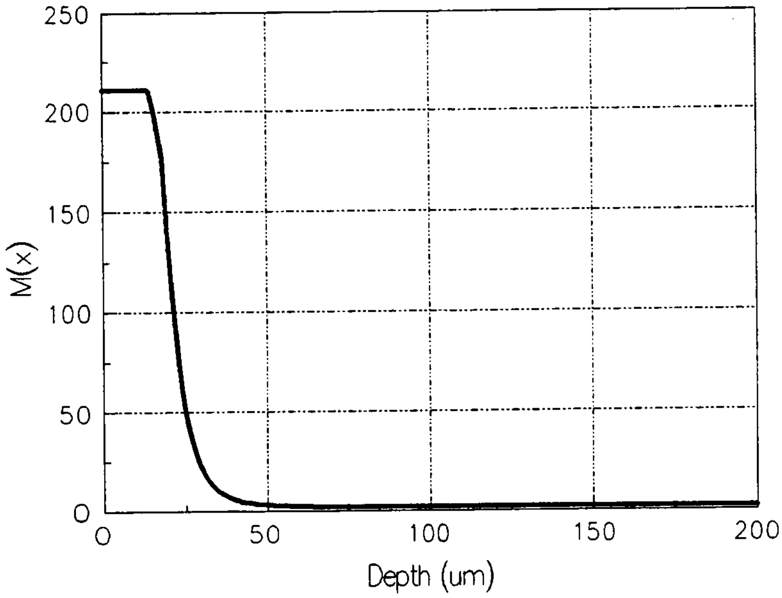

Figure 3-6. Calculated gain profile for beveled-edge APD. V=1475V.

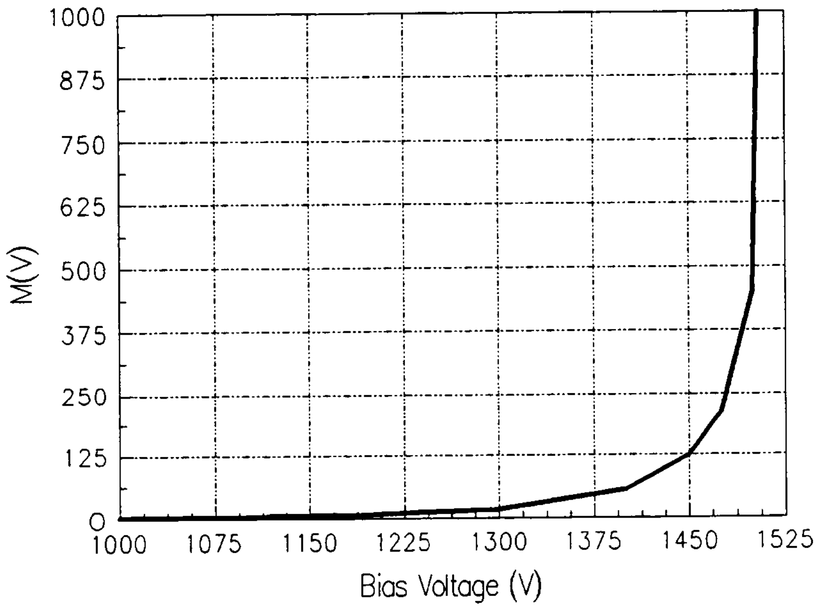

Figure 3-7. Calculated gain vs. bias voltage curve for the beveled-edge APD.

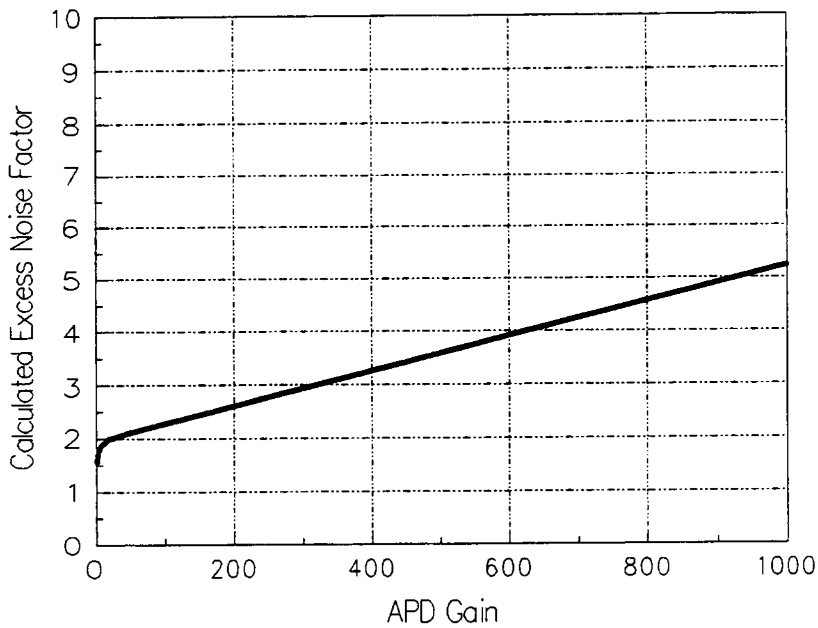

Figure 3-8. Calculated excess noise factor (F) vs. APD gain curve for the beveled-edge APD.

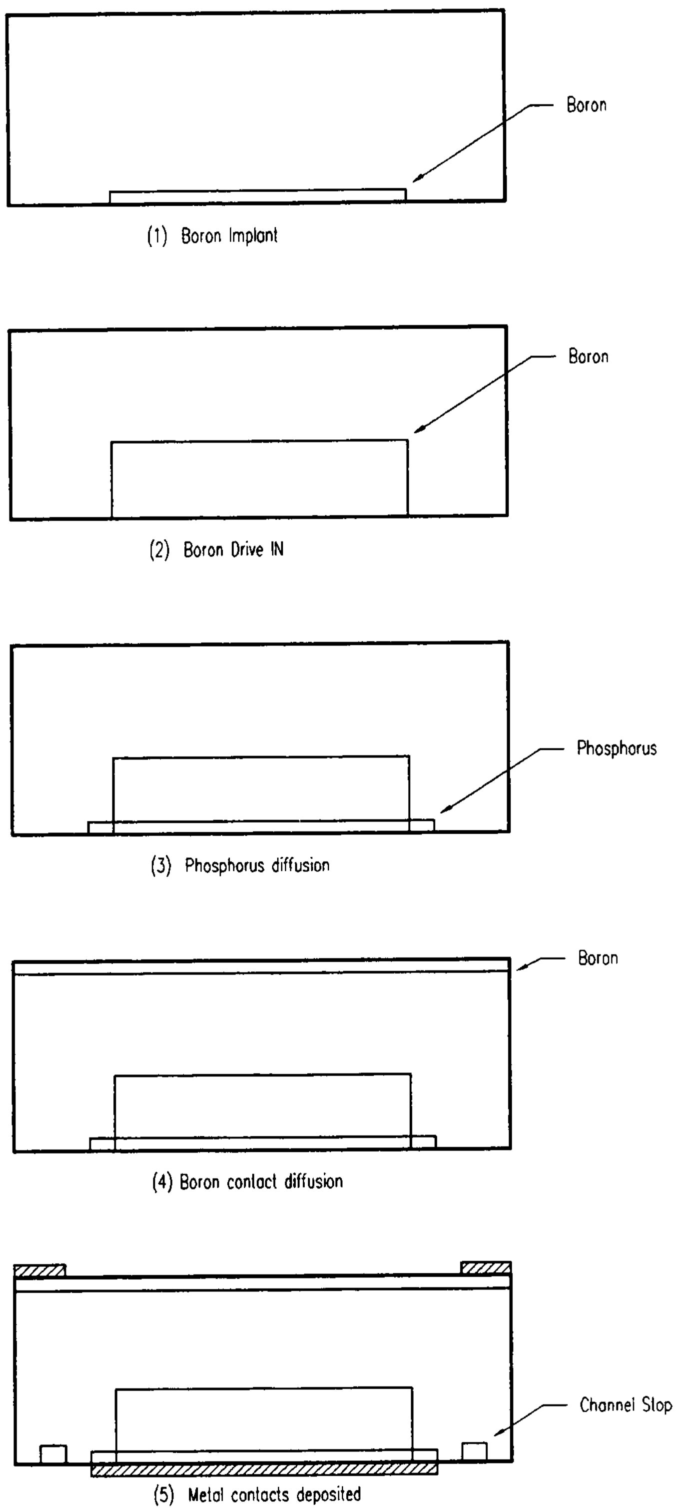

Figure 3-9. Basic fabrication steps for the reach-through APD.

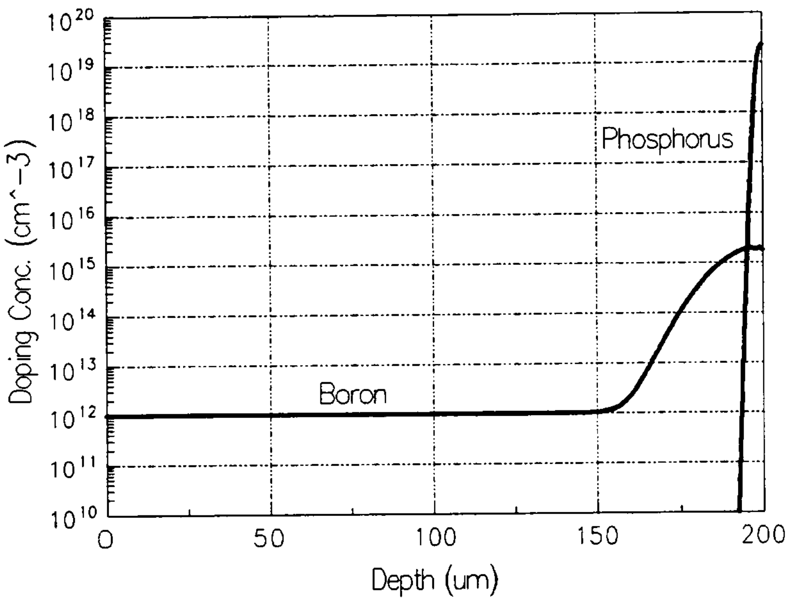

Figure 3-10 SUPREM calculated doping profile for reach-through APD.

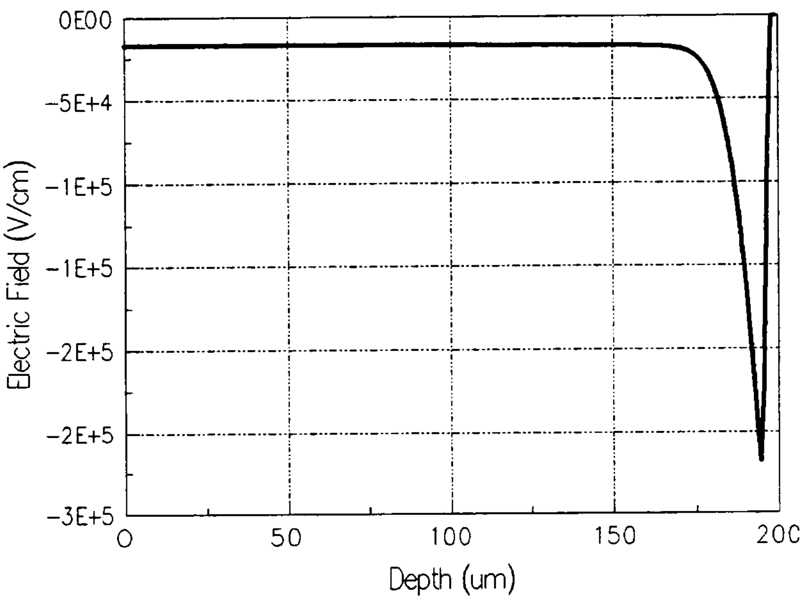

Figure 3-11. Calculated electric field profile for a typical reach-through APD biased at 550V. The calculated gain for this structure is M≈210. V = 550V.

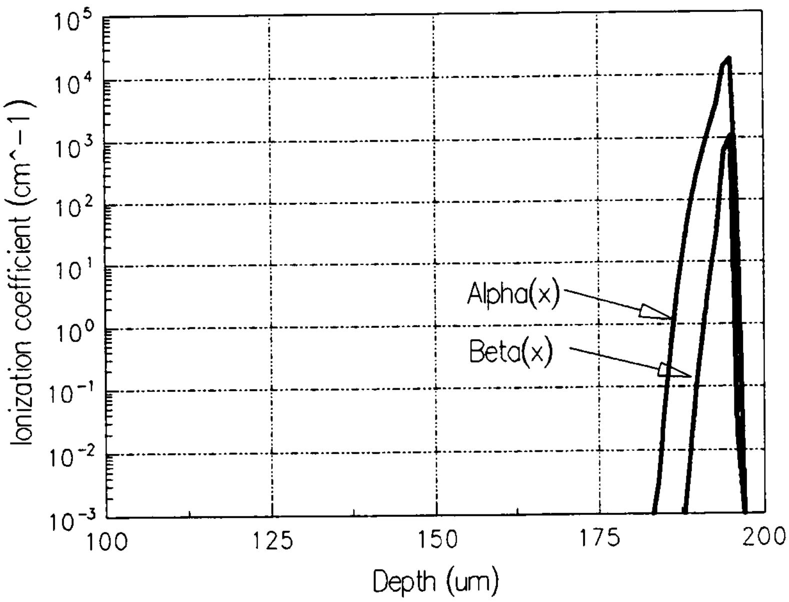

Figure 3-12. Calculated ionization coefficients for reach-through APD biased for M≈210. V = 550V. Note the scale is expanded to show the bottom 100μm of the APD.

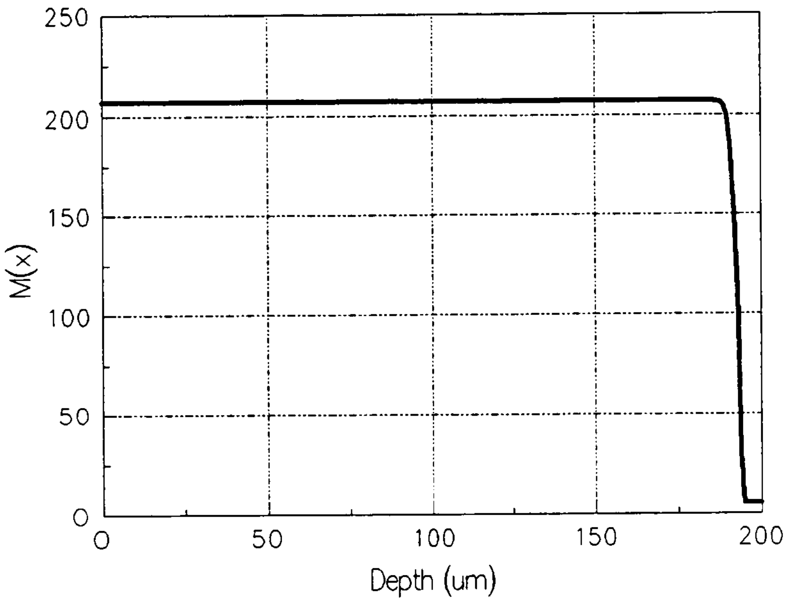

Figure 3-13. Calculated gain profile for reach-through APD. V = 550V.

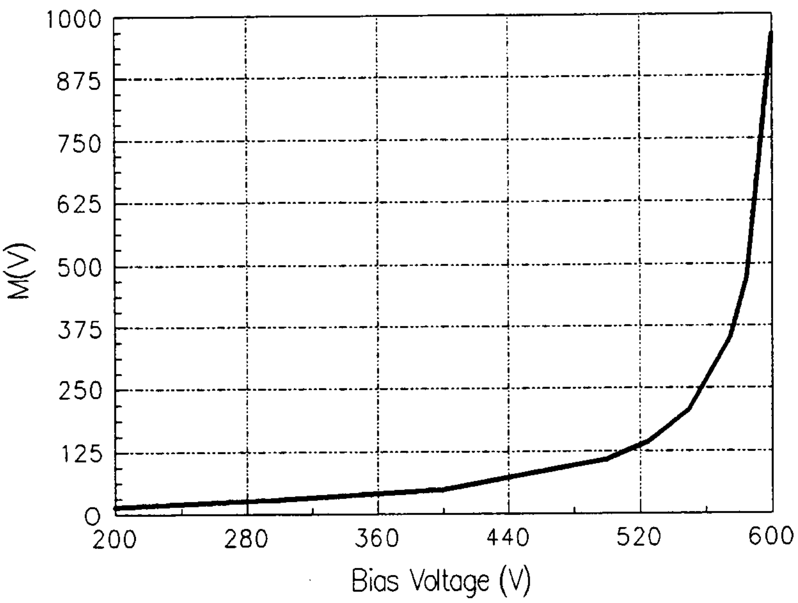

Figure 3-14. Calculated gain vs. bias voltage curve for the reach-through APD.

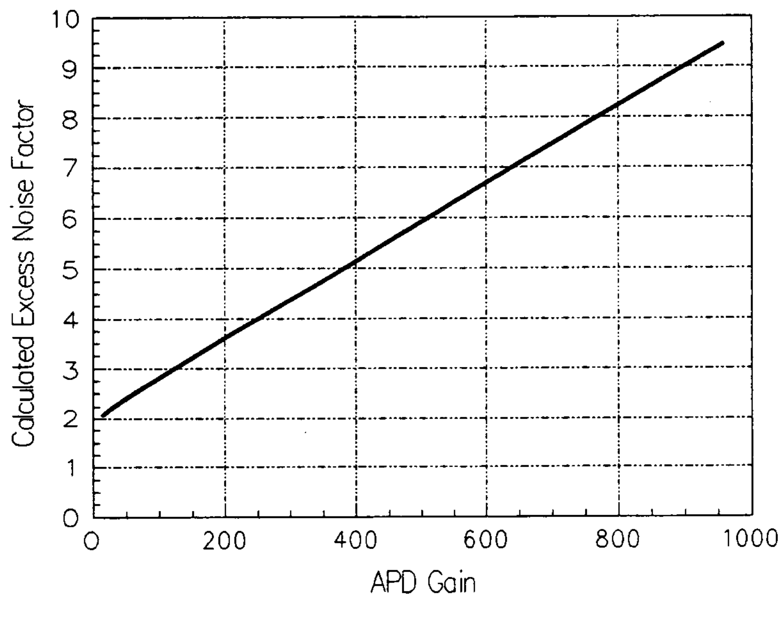

Figure 3-15. Calculated excess noise factor (F) vs. APD gain for the reach-through APD.

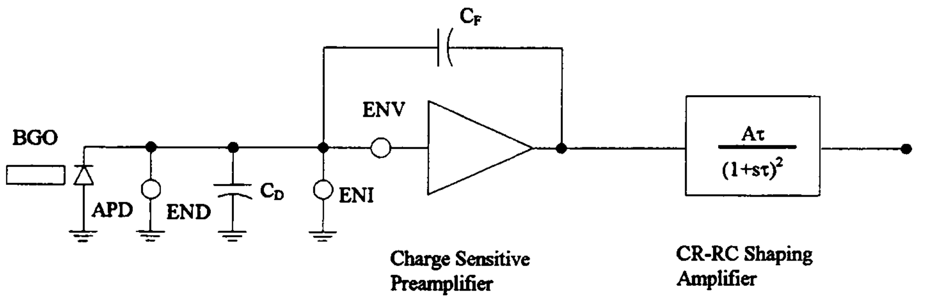

Figure 3-16. Typical preamplifier and shaping amplifier configuration for pulse height resolution and timing resolution measurement.

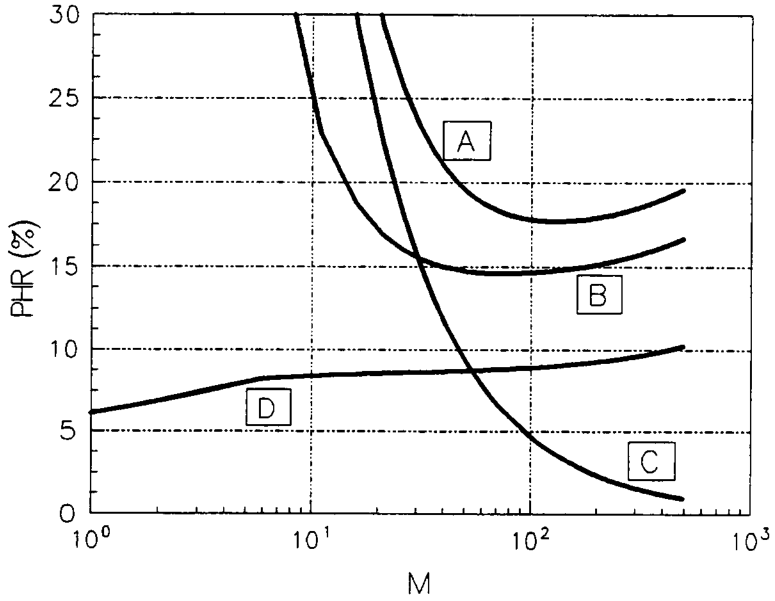

Figure 3-17. Calculated PHR vs. APD gain for a 17 mm diameter Advanced Photonix APD with a CR-RC shaping time constant of 0.5 μs. A) Overall PHR, B) PHR due to parallel noise current, C) PHR due to series noise voltage, D) PHR due to statistical variation in N·M.

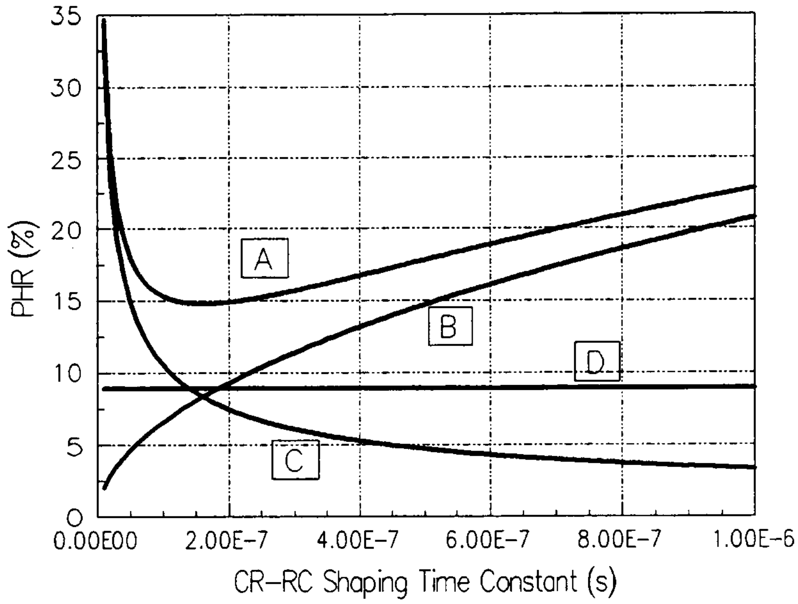

Figure 3-18. Calculated PHR vs. CR-RC shaping time constant for a 17 mm diameter Advanced Photonix APD biased for a gain of 100. A) Overall PHR, B) PHR due to parallel noise current, C) PHR due to series noise voltage, D) PHR due to statistical variation in N·M.

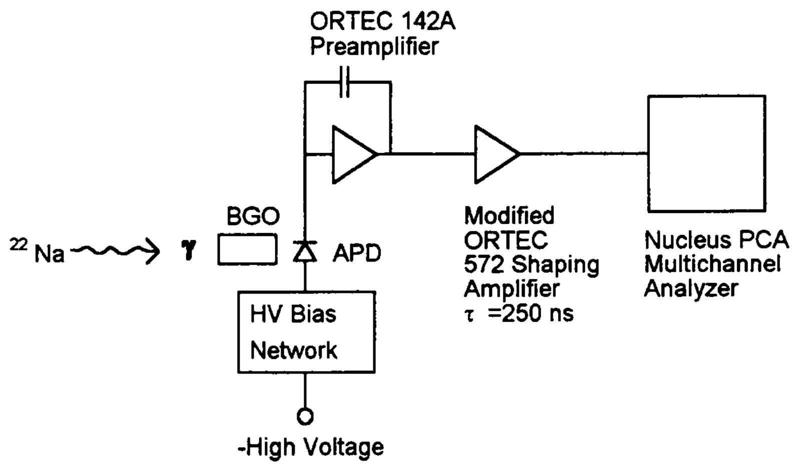

Figure 3-19. Test apparatus used to measure APD pulse height resolution.

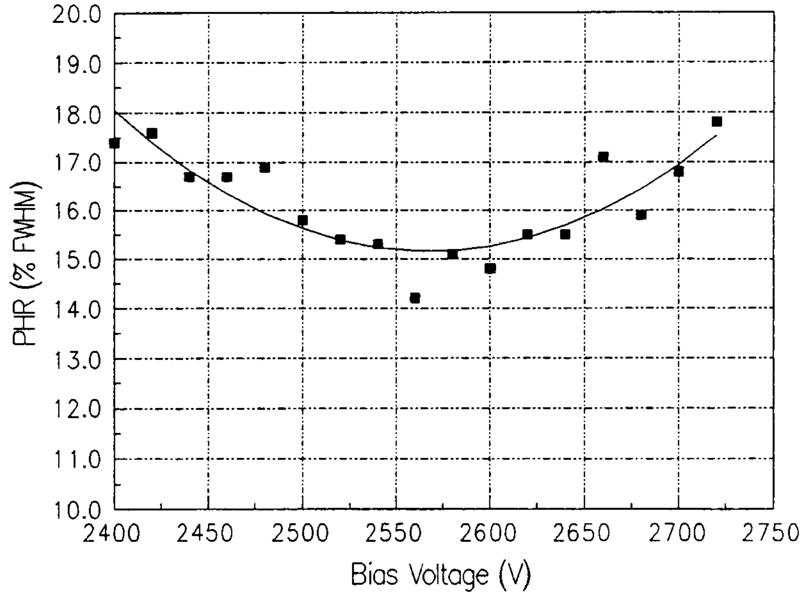

Figure 3-20. Measured PHR vs. bias voltage for API APD. Shaping time constant = 250 ns. Scintillator = 3x3x25 mm^3 BGO.

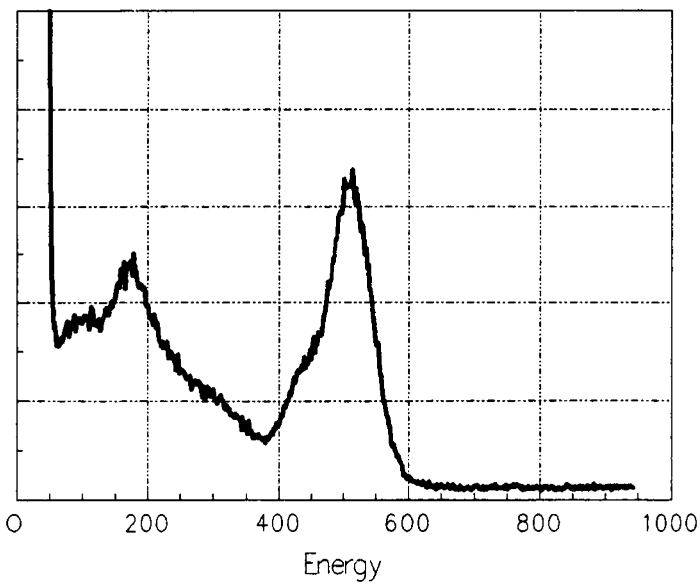

Figure 3-21. Best measured API ^{22}Na (511 keV) PHR. V_{bias}=2560V, PHR = 14.2%, Noise Floor = 40 keV. Scintillator = 3x3x25 mm^3 BGO.

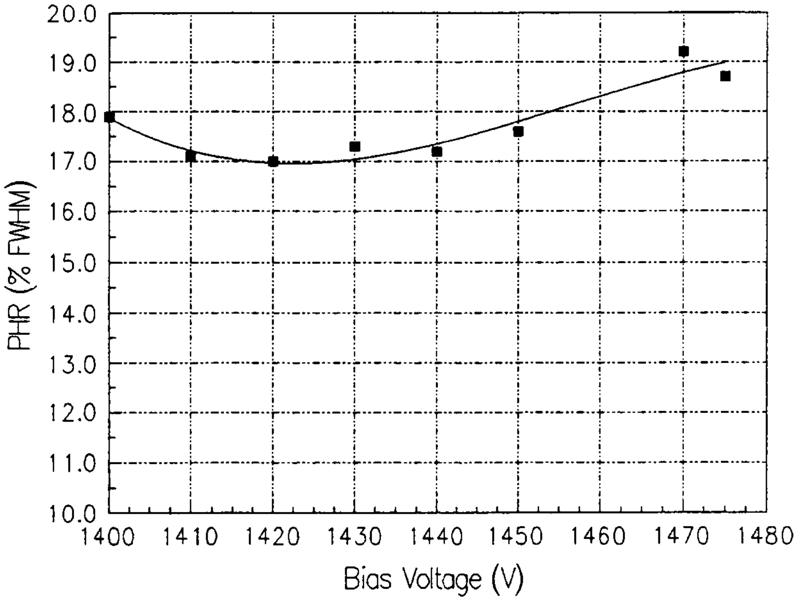

Figure 3-22. Measured PHR vs. bias voltage for RMD APD. Shaping time constant = 250 ns. Scintillator = 3x3x10 mm^3 BGO.

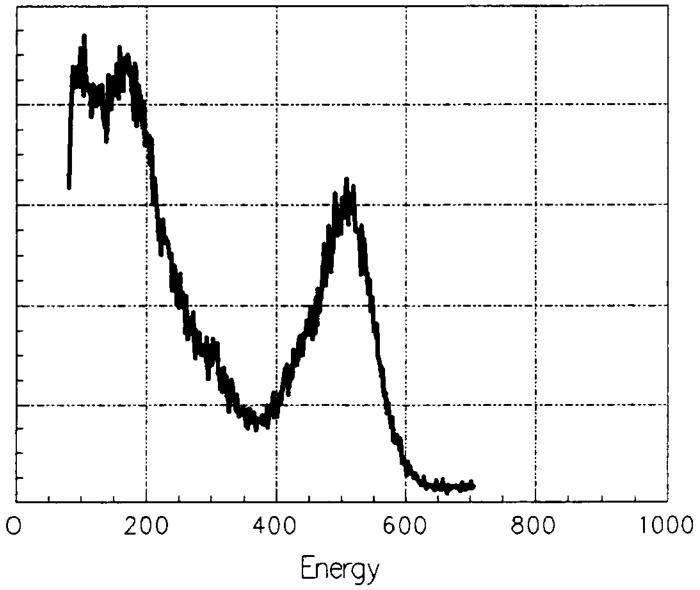

Figure 3-23. Best measured RMD ^{22}Na (511 keV) PHR. V_{bias}=1440V, PHR = 17%, Noise Floor = 100 keV. Scintillator = 3x3x10 mm^3 BGO.

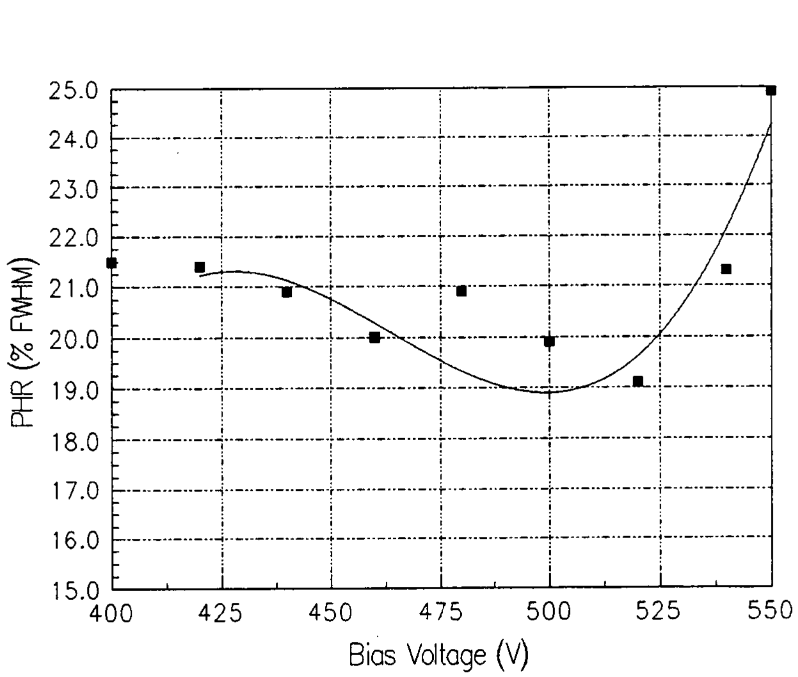

Figure 3-24. Measured PHR vs. bias voltage for EG&G APD/BGO module. Shaping time constant = 250 ns. Scintillator = 3x5x20 mm^3 BGO.

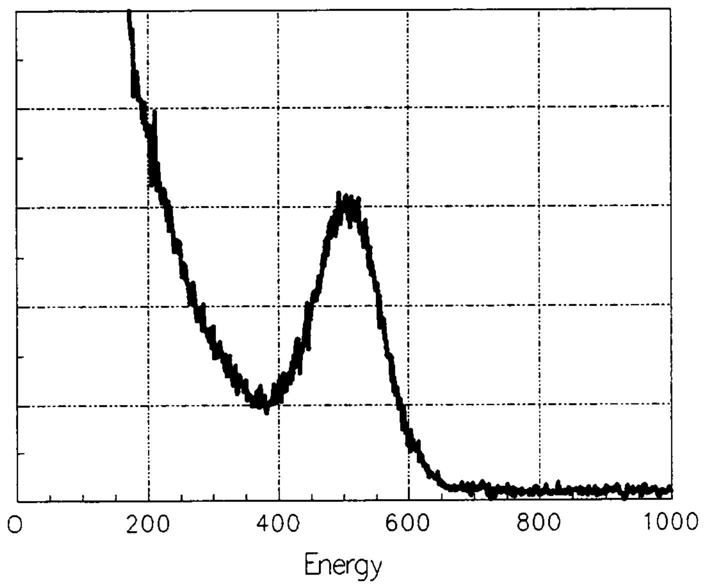

Figure 3-25. Best measured EG&G ^{22}Na (511 keV) PHR. V_{bias}=520V, PHR = 19%, Noise Floor = 180 keV. Scintillator = 3x5x20 mm^3 BGO.

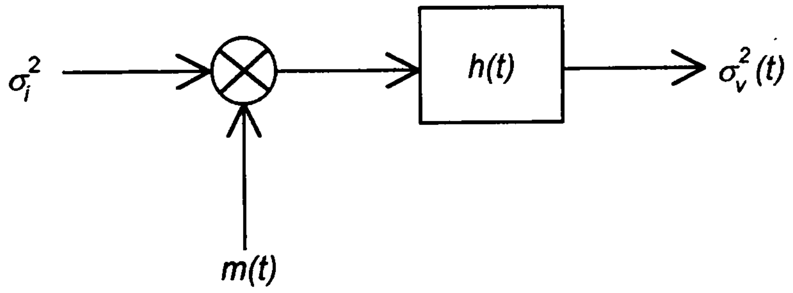

Figure 3-26. Schematic diagram of APD/BGO time variant noise model.

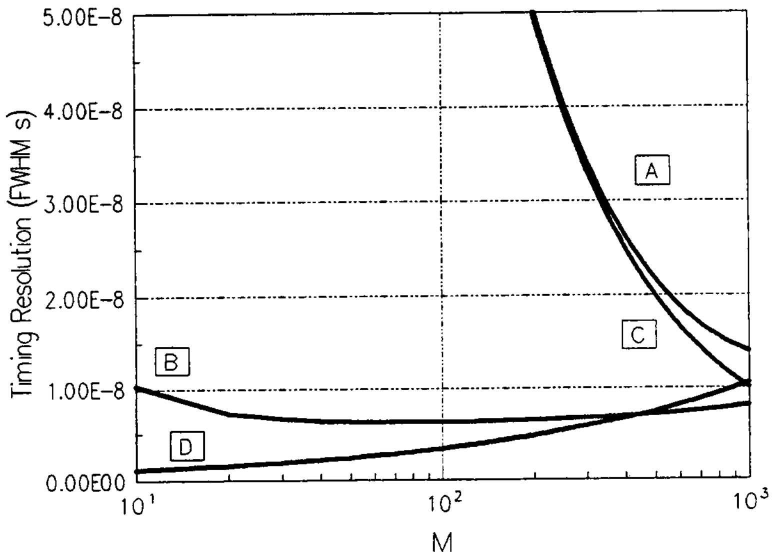

Figure 3-27. Calculated timing resolution vs. APD gain for a 17 mm diameter Advanced Photonix APD with a CR-RC shaping time constant of 10 ns.
A) Overall timing resolution, B) timing resolution due to parallel noise current, C) timing resolution due to series noise voltage, D) timing resolution due to statistical variation in r·M.

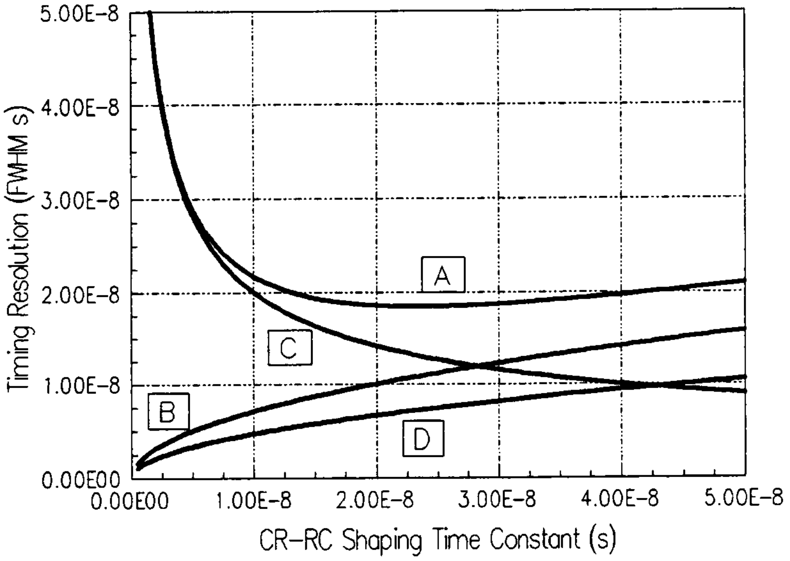

Figure 3-28. Calculated timing resolution vs. shaping time constant for a 17 mm diameter Advanced Photonix APD with an APD gain of 500.
A) Overall timing resolution, B) timing resolution due to parallel noise current, C) timing resolution due to series noise voltage, D) timing resolution due to statistical variation in r·M.

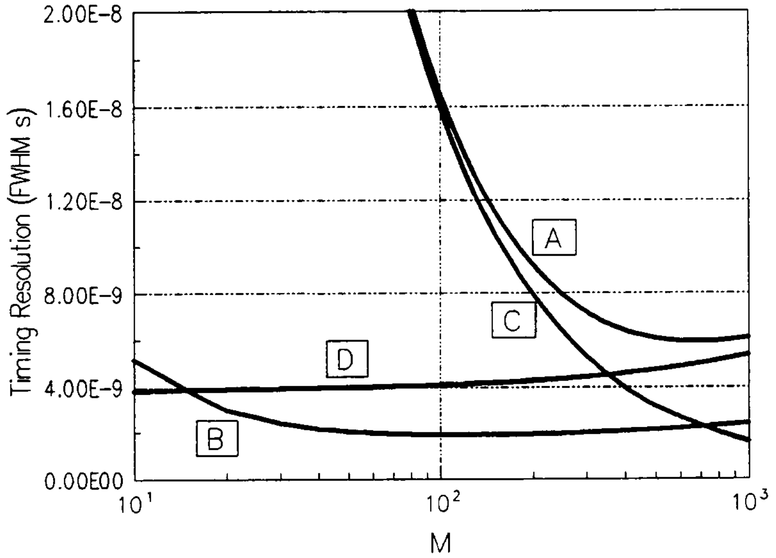

Figure 3-29. Calculated timing resolution vs. APD gain for a 5 mm diameter Advanced Photonix APD with a CR-RC shaping time constant of 10 ns.
A) Overall timing resolution, B) timing resolution due to parallel noise current, C) timing resolution due to series noise voltage, D) timing resolution due to statistical variation in r·M.

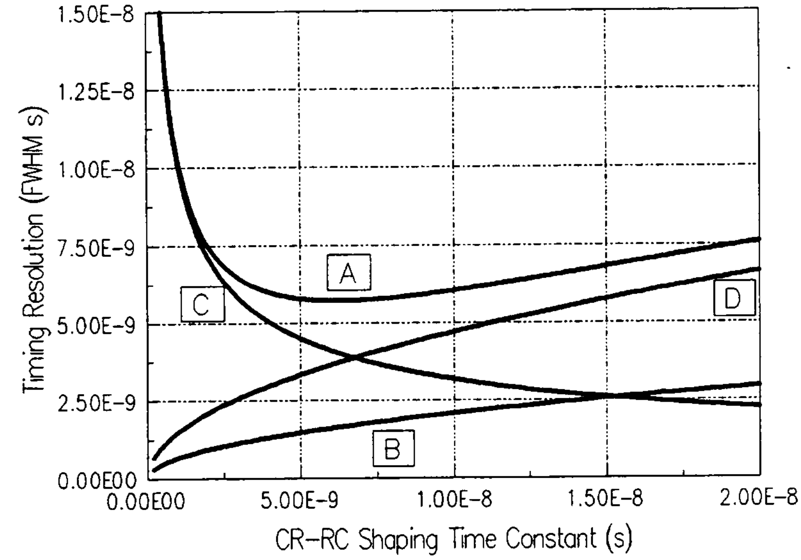

Figure 3-30. Calculated timing resolution vs. shaping time constant for a 5 mm diameter Advanced Photonix APD with an APD gain of 500.
A) Overall timing resolution, B) timing resolution due to parallel noise current, C) timing resolution due to series noise voltage, D) timing resolution due to statistical variation in r·M.

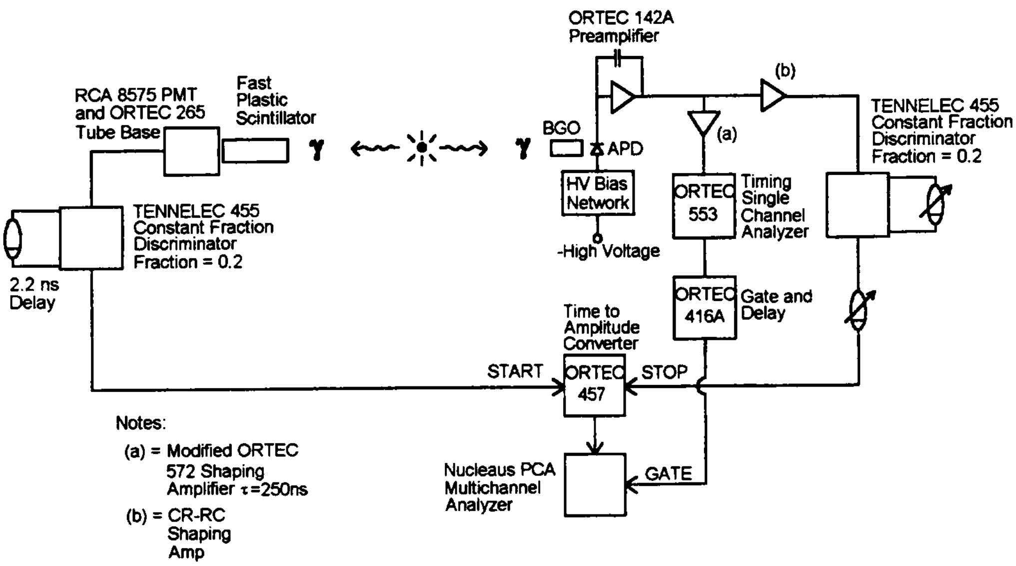

Figure 3-31. Test system used to measure the timing resolution of the APD coupled to BGO in coincidence with a PMT coupled to a plastic scintillator.

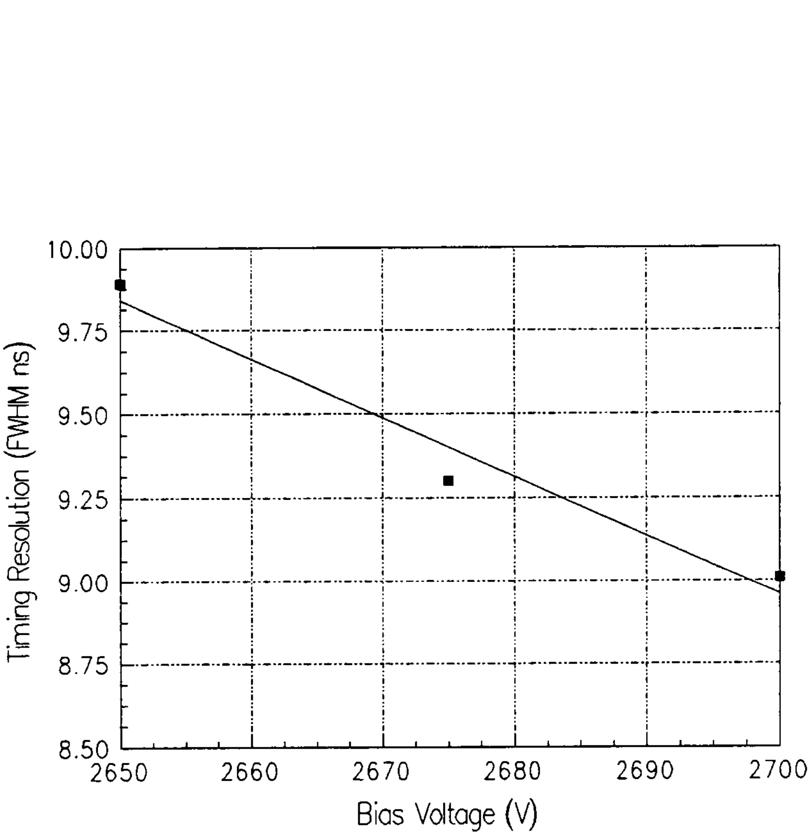

Figure 3-32. Measured timing resolution vs. bias voltage for 5 mm diameter. round API APD (Table 3-6). The shaping time constant is 2.2 ns and the scintillator is 3 x 3 x 25 mm^3 BGO.

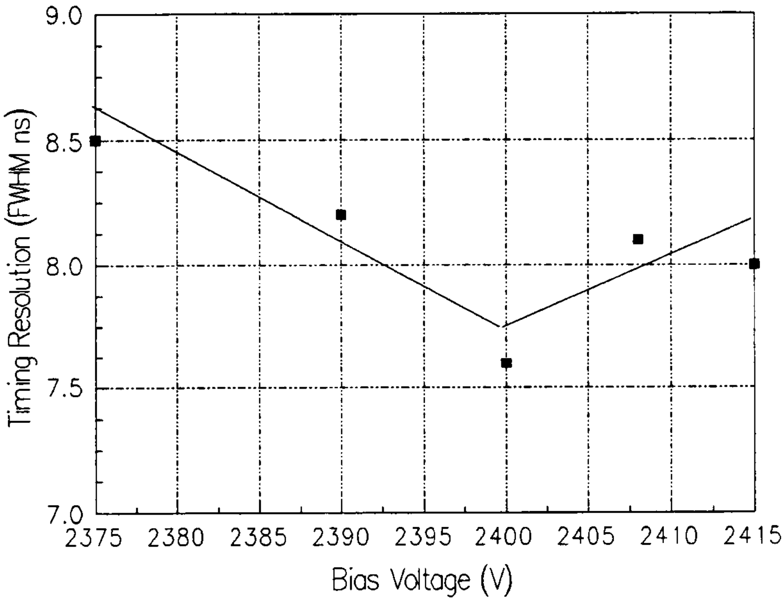

Figure 3-33. Measured timing resolution vs. bias voltage for 5 x 5 mm^2 square API APD. The shaping time constant is 2.2 ns and the scintillator is 3 x 3 x 30 mm^3 BGO.

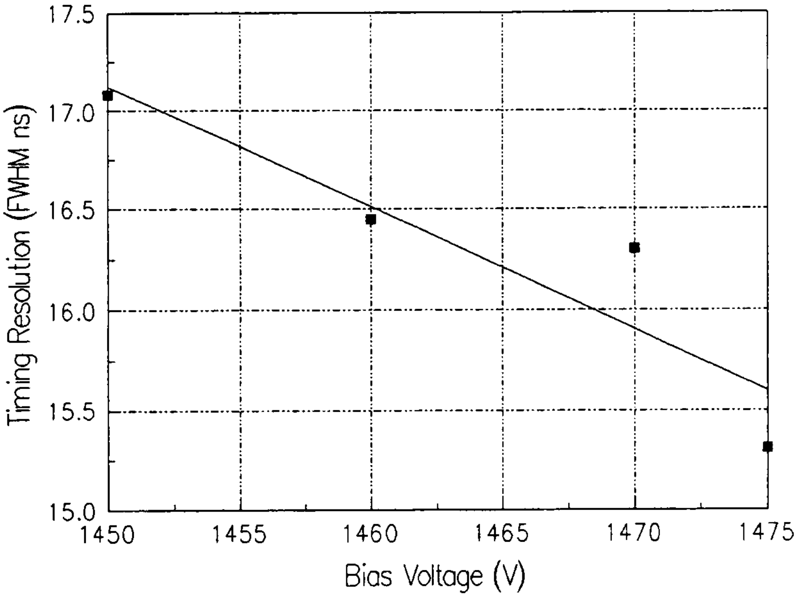

Figure 3-34. Measured timing resolution vs. bias voltage for the RMD APD (Table 3-6). The shaping time constant is 20 ns and the scintillator is 3 x 3 x 10 mm³ BGO.

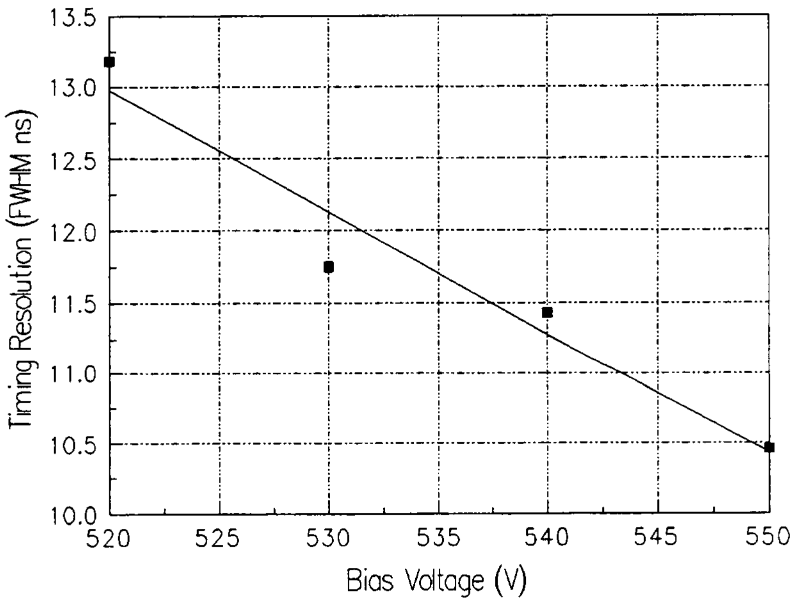

Figure 3-35. Measured timing resolution vs. bias voltage for the EG&G APD/BGO module (Table 3-6). The shaping time constant is 2.2 ns and the scintillator is 3 x 5 x 20 mm³ BGO.

PART 4

DESIGN AND ANALYSIS OF A 64 ELEMENT APD ARRAY

Overview

The ultimate goal of this project is to develop a new PET detector using a monolithic APD array. The first step toward this goal was to review the APD's basic theory of operation and to calculate its frequency dependent behavior (Chapter 2). In Chapter 3, the theory was used to model commercially available discrete APDs to determine which fabrication technology is most appropriate for use in PET. The model predictions were experimentally validated using APDs from three different manufacturers. Based upon this evaluation, Advanced Photonix, Inc. (API) was selected to fabricate a 64-element APD array customized for PET. In this chapter the design criteria for the array are developed, an equivalent circuit for the array is presented and the expected performance of the array is calculated. The design requirements have been communicated to Advanced Photonix which is currently working to fabricate the array.

Custom APD Array Design Criteria

API was selected to fabricate the custom APD array because API uses the beveled-edge technology which simulations show to have inherently superior noise characteristics compared to the reach-through structure and because API's devices have been experimentally proven to have the best available performance for the PET application. Advanced Photonix has patented a technique for segmenting large-area APDs into monolithic arrays of discrete pixels suitable for use in a PET detector [1-4]. Shown in Figure 4-1, the API array is formed from a single large area APD by sawing or etching grooves in the back side (n-side) of the device. The grooves isolate discrete pixels in the undepleted n-region of the device and form a guard-ring around the outer edge of the device. These features are discussed in detail below. An interesting characteristic of all API APDs is the silicon support ring on the top side of the wafer [5].

The support ring is cut from a second heavily doped p-type silicon wafer and then bonded to the base wafer, forming the top side electrical contact and providing mechanical support for the beveled-edge.

Pixel Isolation

The objective of the pixel isolation scheme is to divide a single large area APD into an array of independent APD pixels with a common p^+ top contact. The API groove technology accomplishes this by dividing the undepleted n-region into electrically isolated islands. The groove depths are chosen so that the depletion regions associated with the surface states at bottoms of the grooves merge with the depletion region of the p-n junction, pinching off the channels between the back-side contacts. The API arrays are designed to have inter-pixel resistances greater than 10 MΩ when properly biased [6]. Note that the p-side of the array is not segmented. The API pixel isolation technique relies on the high electric field in the depletion region to minimize carrier transport parallel to the plane of the array so that virtually all of the charge generated by a photon absorbed above a given back-side contact is collected by that contact. The grooves do not introduce dead regions in the array. Photons collected in the region above the grooves deliver their signal to the nearest adjacent contact.

Selection and control of the groove depth is a critical part of the API design and fabrication process. If the grooves are too shallow, resistive paths exist between the backside contacts, introducing inter-pixel crosstalk and thermal noise. If the grooves are too deep, they extend into the p-n junction depletion region, causing premature breakdown of the array [6].

The pixel isolation technique also segments the array capacitance. From basic electrostatic theory, the small signal capacitance is given by

$$C = \frac{dQ}{dV}, \tag{4-1}$$

where Q is the stored charge on the capacitor and V is the potential difference across the device. For a reverse biased diode, Q is proportional to the diode area. When the back side of the APD is divided into electrically isolated pixels, the charge associated with each pixel is isolated from the charge associated with the other pixels. Consequently, the capacitance per pixel is determined by the area of the pixel rather than by the area of the entire array. The per-pixel capacitance of the custom API array is calculated below in the development of the APD array equivalent circuit.

Guard Ring

As described in Chapter 3, the dark current flowing along the surface of the beveled edge, I_{ds}, is an important source of noise in an APD. For the discrete API devices studied in Chapter 3, the surface current is summed with the signal current at the back side contact, degrading the signal-to-noise ratio. In the API APD array design, the electrical contact at the edge of the back side is isolated from the signal current flowing in the pixel contacts. This edge contact, called the guard ring (Figure 4-1), collects the surface current and routes it directly to ground, resulting in an improved signal-to-noise ratio. Although this improvement can be significant, the guard ring does introduce a dead region along the perimeter of the APD array. The size of the guard ring must therefore be minimized in order to obtain the largest possible active region on the APD surface.

Physical Design of the Custom APD Array

Three parameters were considered in developing the physical specifications for the custom APD array: the pixel area, the active area of the APD array and the total area of the APD array. These parameters are discussed below, and the array dimensions are summarized in Table 4-1.

Pixel Size

The highest resolution PET scanner currently manufactured by CTI uses BGO block detectors with approximately 4.5 x 4.5 x 30 mm^3 pixels. Future PET scanners will use scintillator pixels with smaller facial areas to improve the scanner spatial resolution. For this reason, the custom APD array is designed to accommodate an array of 3 x 3 x 30 mm^3 pixels.

Active Area of the APD Array

In order to minimize the number of components in the PET scanner, it is desirable to maximize the number of pixels per APD array. However, in order to achieve acceptable yield from the APD fabrication process, the active area of the APD array must be kept relatively small. Based upon the yield parameters of the API process, the custom array was designed to have sixty-four 3 mm x 3 mm elements, giving the array a total active area of approximately 1 in^2. This is the largest area APD array known to be fabricated to date for any application [6].

Total Area of the APD Array

The software algorithms which generate PET images require the pixel-to-pixel spacing of the detectors to be constant around the detector ring. Consequently, the detectors must be designed so that no gaps appear in the pixel spacing at the interfaces between detector modules. As shown in Figure 4-1, however, the APD array has dead regions around its perimeter due to the presence of the silicon support ring, beveled edge and guard ring contact. The maximum acceptable dead area of the array is determined by the geometry of the PET scanner gantry.

All of CTI's current PET scanners are designed with ~80 cm diameter detector rings at the face of the scintillators. Figure 4-2 shows three APD/BGO detector modules positioned along such an 80 cm ring. If the detector modules are packed so that the BGO pixel-to-pixel spacing at the front surface of the detector modules is constant around the ring, the spacing between the BGO arrays at the rear of the modules will be approximately 1.8 mm. Thus, the APD arrays are permitted to extend approximately 0.9 mm beyond the edge of the BGO arrays.

The requirement that the dead ring around the APD array be less than 0.9 mm wide was found to be incompatible with API's current fabrication technology [6]. Consequently, the specifications were relaxed to permit a dead ring of 2 mm for the prototype APD array. Although the wider dead ring is unacceptable for use in a full size PET scanner, it is acceptable for a smaller (<36 cm diameter) gantry such as might be used in a brain scanner or in a research scanner for use with laboratory animals. Future improvements in the array fabrication technology are expected to reduce the dead ring to less than 0.9 mm.

Physical Specifications

Based upon the considerations described above, the physical dimensions of the custom APD array were specified as shown in Figure 4-3. The array design calls for a 64 pixel device with center-to-center pixel spacing of 3 mm; the mechanical specifications agreed upon by API are summarized in Table 4-1.

Table 4-1.

APD Array Mechanical Specifications.

1. Pixel pads are square with a 3 mm center-to-center pitch.
2. The nominal gap between the pads is 300 μm.
3. The metallic guard ring is less than 1 mm in width.
4. The edge bevel is less than 1 mm in width.
5. The silicon front side support ring is 2 mm wide or less.
6. The backside pixel pads are metalized with Ti-Pt-Au.
7. The front surface is coated with a 55 nm anti-reflection coating.
8. The beveled edges are passivated with polyimide.

Equivalent Circuit

Although the objective of the pixel isolation technique is to create an array of independent APD elements, some parasitic coupling between pixels is inevitable. In order to evaluate the electrical crosstalk between pixels due to parasitic coupling as well as to calculate the expected performance of a pixel in the custom array, an equivalent circuit for a prototype element in the APD array is required.

Consider the schematic diagram of three pixels in the APD array shown in Figure 4-4, where D1 is the signal carrying pixel and D2 and D3 are adjacent pixels which may be effect by electrical cross-talk. For the custom array, the dimensions shown in Figure 4-4 have the values given in Table 4-2 [2,6].

The parameters of interest for the equivalent circuit (Figure 4-5) are the depletion region capacitance, C_{depl}, the per-pixel series resistance due to the undepleted n-region above the pixel contacts, R_s, the inter-pixel capacitance, C_{ip}, and the inter-pixel resistance, R_{ip}. Values for these parameters are calculated below:

Table 4-2.
Various Process Dependent Dimensions for the Custom APD Array.

Parameter	Symbol	Value
Depletion Region Depth	W	0.25 mm
Groove to Depletion Region Spacing	W'	15 Å
Groove Depth	W''	0.3 mm
Pixel Width	P	2.7 mm
Groove Width	G	0.3 mm

Depletion Region Capacitance

API APDs are fabricated using nominally 55 Ω-cm transmutation doped n-type wafers and typically have depletion region widths of approximately 250 μm when biased for gains on the order of 200 [2]. Neglecting edge effects, the depletion region capacitance is given by [7]:

$$C_{depl} = \frac{\varepsilon_{Si}}{W} P^2, \tag{4-2}$$

where $\varepsilon_{Si} \approx 1.05$ pF/cm is the dielectric constant of silicon. Using the values in Table 4-2,

$$C_{depl} \approx 3.1\ pF. \tag{4-3}$$

Series Resistance

The undepleted n-regions above the bottom contacts introduce series resistances in each pixel. The series resistance is given by

$$R_s = \frac{\rho W''}{P^2} \tag{4-4}$$

where $\rho = 55$ Ω-cm is the resistivity of the bulk material. Using the values in Table 4-2, the per-pixel series resistance in the custom APD array is

$$R_s \approx 25\Omega. \tag{4-5}$$

Inter-pixel Capacitance

In the current API technology, the grooves are not filled with any dielectric medium. Neglecting edge effects, the capacitance between any two pixels is determined by the surface area of the contact island sidewalls and the groove widths:

$$C_{ip} = \frac{\varepsilon_0 pW''}{G}, \tag{4-6}$$

where $\varepsilon_0 \approx 88.5$ fF/cm is the dielectric constant for free space. Using the values in Table 4-2,

$$C_{ip} = 24\, fF \tag{4-7}$$

The total parasitic capacitance between a pixel and its four nearest neighbors is

$$C_{ip,tot} = 4 \times C_{ip} \approx 0.1\, pF. \tag{4-8}$$

Inter-pixel Resistance

API designs its arrays such that the impedance between any pixel and its nearest neighbors is greater than 10 MΩ [6]. To understand the physical meaning of this specification, let us assume the channel between pixels is approximately 15 Å. The resistance between a pixel and one of its nearest neighbors is given by

$$R_{ip} = \frac{\rho G}{W' p}. \tag{4-9}$$

Using the values in Table 4-2,

$$R_{ip} = 40\, M\Omega. \tag{4-10}$$

The resistance between a pixel and its four nearest neighbors is

$$R_{ip,tot} = \frac{R_{ip}}{4} = 10\, M\Omega. \tag{4-11}$$

Thus, the requirement that the $R_{ip,tot}$ be greater than 10 MΩ implies the inter-pixel channel is less than 15 Å, or that the channel is essentially pinched off.

Electrical Cross Talk and Additional Noise Sources

The calculated parameters for the equivalent circuit shown in Figure 4-5 are summarized in Table 4-3. If the preamplifiers (Amp1-Amp5 in Figure 4-5) are designed with large open loop gains and current summing feedback nodes at the inputs, the amplifier input impedances will be small relative to R_{ip} and R_s. The percentage of the signal current in D1 collected by the amplifiers associated with D2-D5 is determined by the ratio of the parallel interpixel impedances to the sum of the parallel interpixel impedances and R_s,

$$\%Crosstalk = \frac{\dfrac{R_{ip}}{4\left(1+j\omega R_{ip}C_{ip}\right)}}{\dfrac{R_{ip}}{4\left(1+j\omega R_{ip}C_{ip}\right)}+R_s}\times 100\% \tag{4-12}$$

PSPICE analysis of the equivalent circuit shows the crosstalk between D1 and its 4 nearest neighbors to be less than 0.01% for all shaping time constants of interest.

In addition to being a potential source of electrical crosstalk, R_{ip} is a source of additional thermal noise current. The thermal noise added to the signal by $R_{ip,tot}$ is

$$i_{n,Riptot} = \sqrt{\frac{4kT}{R_{ip,tot}}} = 41\frac{fA}{\sqrt{Hz}}, \tag{4-13}$$

which is negligible relative to the multiplied bulk dark current noise of the APD (typically >1 pA/√Hz).

Table 4-3.

Calculated Equivalent Circuit Parameters for the Custom APD Array.

Parameter	Symbol	Value
Depletion Region Capacitance	C_{depl}	3.1 pF
Series Resistance	R_s	25 Ω
Inter-pixel capacitance	C_{ip}	24 fF
Inter-pixel Resistance	R_{ip}	40 MΩ

R_s also adds an equivalent noise voltage in series with the preamplifier input. For $R_s = 25\Omega$, the added noise voltage is

$$e_{n,Rs} = \sqrt{4kTR_s} = 0.64\frac{nV}{\sqrt{Hz}}. \qquad (4\text{-}14)$$

This added noise voltage is somewhat less than the 1-2 nV/√Hz typically found in low-noise preamplifiers, but it is sufficiently large to degrade the timing performance at short shaping time constants. In comparison, the 5 mm diameter API APD considered in Chapter 3 has an area approximately 2.7 times greater than that of a pixel in the custom array, giving the 5 mm diameter device a series resistance of ~8.4 Ω and an equivalent noise voltage of ~0.37 nV/√Hz.

Expected APD Performance

The expected pulse-height resolution and timing resolution of a pixel in the custom APD array was calculated using the procedure described in Chapter 3, where the APD signal is modeled as a current impulse. The preamplifier and shaping amplifier parameters used in this analysis are the same as those used in Chapter 3, but the APD capacitance was modified based on the calculations presented above. Also, extra noise sources due to R_{ip} and R_s sere added to the model to treat the parasitic effects of adjacent pixels in the array. The parameters used in the analysis are summarized in Table 4-4.

Pulse Height Resolution

Although the data in Chapter 3 shows the optimal shaping time constant for pulse height resolution to be less than 200 ns, the laboratory equipment available to analyze the APD array can only process pulses with peaking times greater than or equal to 500 ns. Thus, the pulse height resolution is calculated as a function of APD gain using the

Table 4-4.

Summary of Parameters Used in PHR Calculations.

Parameter	**Value**
Scintillator photon yield	N=1500
APD Quantum Efficiency	η=0.85
APD Depletion Capacitance per pixel	C_{depl}=3.1 pF
Added Capacitance due to Grooves per pixel	$C_{ip,tot}$= 0.1 pF
APD k_{eff}	k_{eff}=0.0017
Un-multiplied APD bulk dark current per pixel	I_{db}=20 pA
APD surface dark current	I_{ds}=0
Shaping Amplifier Gain	A=100
Preamplifier feedback capacitance	C_F=1 pF
Preamplifier input noise current	$i_{na}/\sqrt{\Delta f}$=50 fA/√Hz
Noise current due to R_{ip}	$i_{n,Riptot}$=41 fA/√Hz
Preamplifier input noise voltage	$e_{na}/\sqrt{\Delta f}$=2 nV/√Hz
Noise voltage due to R_s	$e_{n,Rs}/\sqrt{\Delta f}$=0.65 nV/√Hz

parameters in Table 4-4 and a shaping time constant of 500 ns (Figure 4-6). Figure 3-18 shows that increasing the time constant from 200 ns to 500 ns will increase the PHR (obtained with a current impulse) by approximately 3%.

Comparison of Figure 4-6 with Figure 3-17 (calculated PHR for a 17 mm diameter discrete APD) shows a marked improvement in the calculated PHR due to the reduction in parallel noise current attained with a reduced pixel area and elimination of the surface dark current. The PHR of the larger APD analyzed in Chapter 3 is dominated by the series noise current over most of the APD gain range, while the PHR of the smaller array pixel analyzed here is dominated by the statistics of the BGO scintillation process and multiplication in the APD. This improvement offers two significant advantages. First, the calculated optimal PHR is improved from approximately 17% FWHM with the larger device to approximately 10% with the array pixel. Second, the PHR is much less dependent on the APD gain. As shown in Figure

4-6, the PHR minimum remains relatively constant over a gain range of approximately 10 to 500.

Timing Resolution

Figure 4-7 shows the calculated timing resolution of an APD and the read-out circuitry as a function of APD gain with the CR-RC shaping time constant set to 5 ns. As in Chapter 3, the APD/BGO signal is modeled as a current step function. Comparison of Figure 4-7 with Figure 3-28 (calculated timing resolution for a 5 mm diameter APD) shows that the array timing resolution should also be superior to that of the discrete device. In this case, the timing resolution improvement comes as a result of the reduced capacitance of the pixel. As with the discrete APD, the optimal calculated timing resolution (~4.1 ns FWHM) occurs at high APD gains (>600).

Figure 4-8 shows the calculated timing resolution of a pixel in the array as a function of CR-RC shaping time constant, where the APD is biased for a gain of 500. As with the discrete APDs, the optimal calculated timing resolution (~4.9 ns FWHM) is achieved with a very short shaping time constants (~5 ns).

Conclusions

In this chapter, the details of the custom APD array design are presented and the expected performance of the array is calculated. PSPICE simulations show that the individual elements of the array should behave as independent APDs, with PSPICE simulations showing the expected electrical crosstalk to be less than 0.01% for both pulse-height resolution and timing resolution measurements. The pulse-height resolution of the pixels in the array is expected to be superior to that of the discrete APDs studied in Chapter 3 (~10% FWHM) due to the reduction in active area and to the removal of the surface dark current from the signal path by the guard ring. The timing resolution of the pixels in the array is also expected to be superior to that of the discrete APDs (~4.9 ns FWHM) due to the reduction in the per-pixel capacitance. In Chapter 5, the design and experimental characterization of a CMOS preamplifier array suitable for use with the custom APD array is presented.

References

1. Gerald C. Huth, "Silicon Avalanche Photodiode Array," U. S. Patent 5,021,854, June 4, 1991.
2. Ernesto Gramsch, S. Zhang, M. Madden, M. Lindberg and M. Szawlowsky, "High Density Avalanche Photodiode Array", Advanced Photonix, Inc., internal report.
3. E. Gramsch, S. Zhang, M. Madden, M. Szawlowsky and M. Lindberg, "Performance of a 9-element Avalanche Photodiode Array," *IEEE NS Conference Record,* p 186, October 25-31, 1992, Orlando, Florida.
4. The Avalanche Photodiode Catalog, Advanced Photonix, Inc., Camarillo, California.
5. Jan S. Iwanczyk, "Light Responsive Avalanche Diode," U. S. Patent 5,057,892, October 15, 1991.
6. Michael Madden, Advanced Photonix Vice President for Research and Development, private communication.
7. M. Sze, *Physics of Semiconductor Devices,* 2nd Edition, New York: John Wiley and Sons, 1981.

Figures

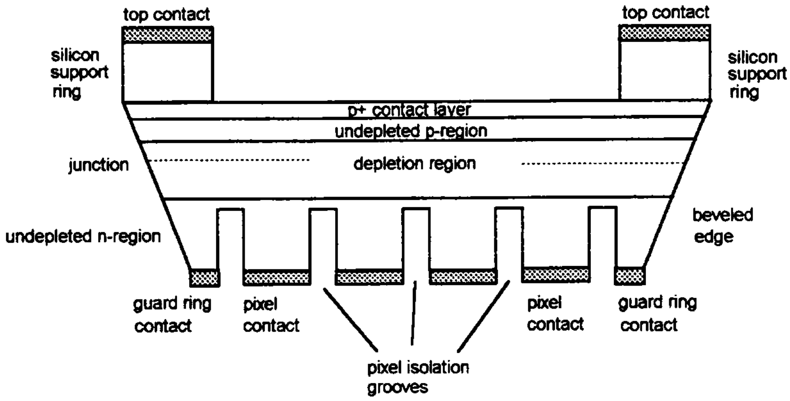

Figure 4-1. Schematic representation of the Advanced Photonix beveled-edge APD array.

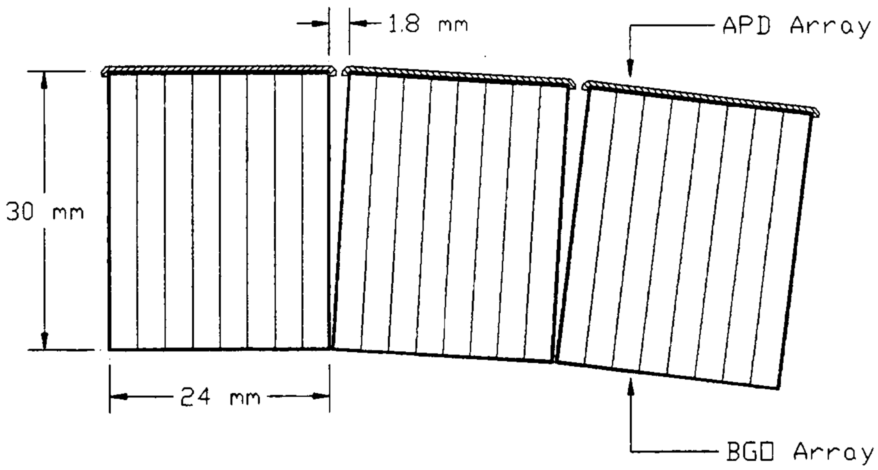

Figure 4-2. Scale drawing of 3 APD/BGO detectors in an 80 cm diameter ring. The diameter of the detector ring determines the maximum distance which the APD array may extend beyond the edge of the BGO array.

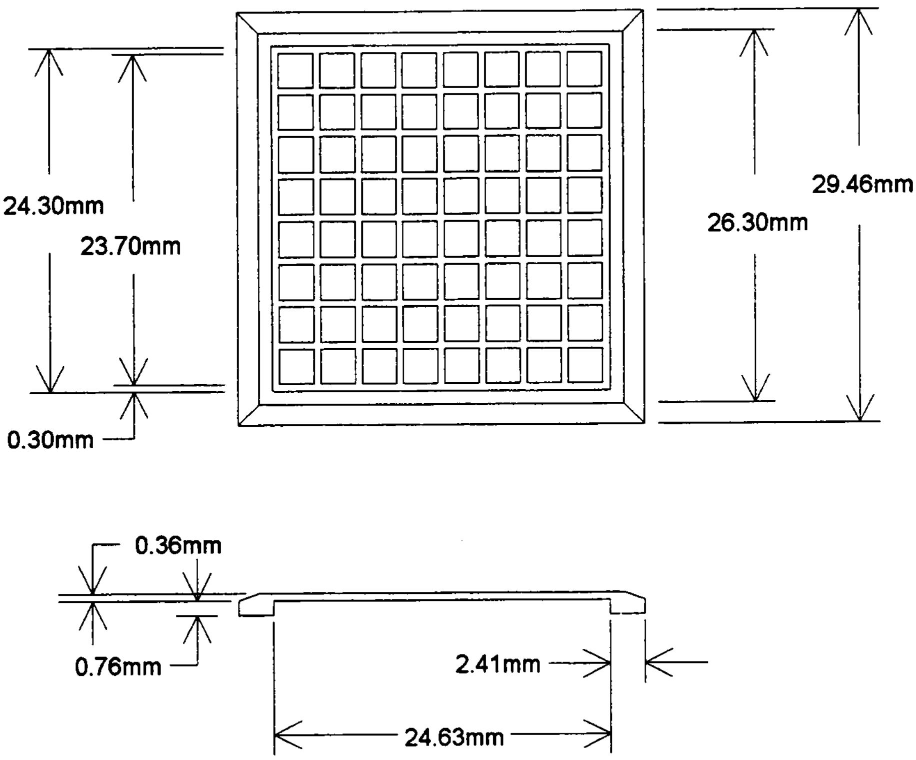

Figure 4-3. Mechanical dimensions of the custom APD array.

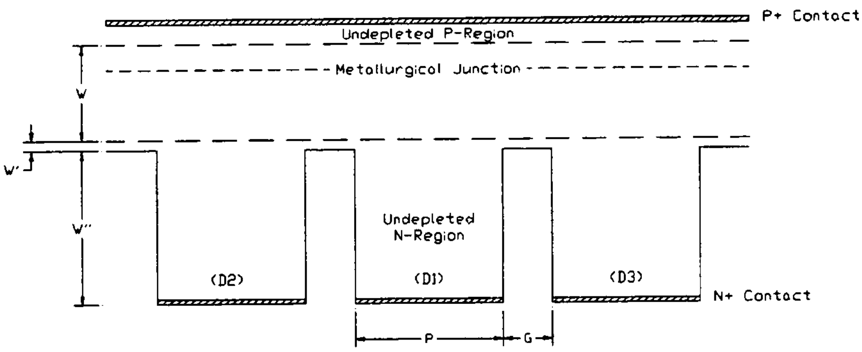

Figure 4-4. Schematic diagram of a pixel in the custom APD array. The dimensions shown are input parameters in the equivalent circuit model.

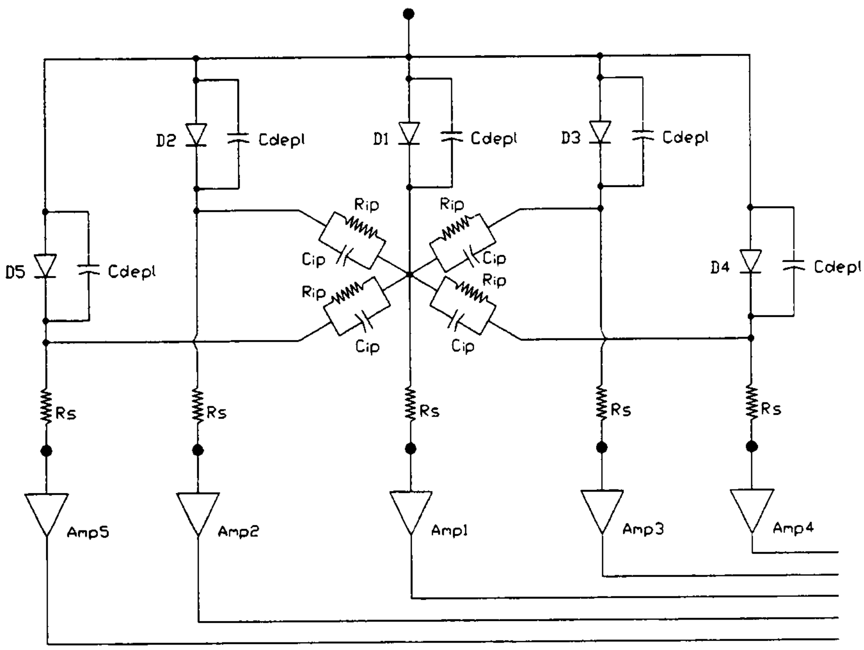

Figure 4-5. Equivalent circuit model of a pixel in the custom APD array.

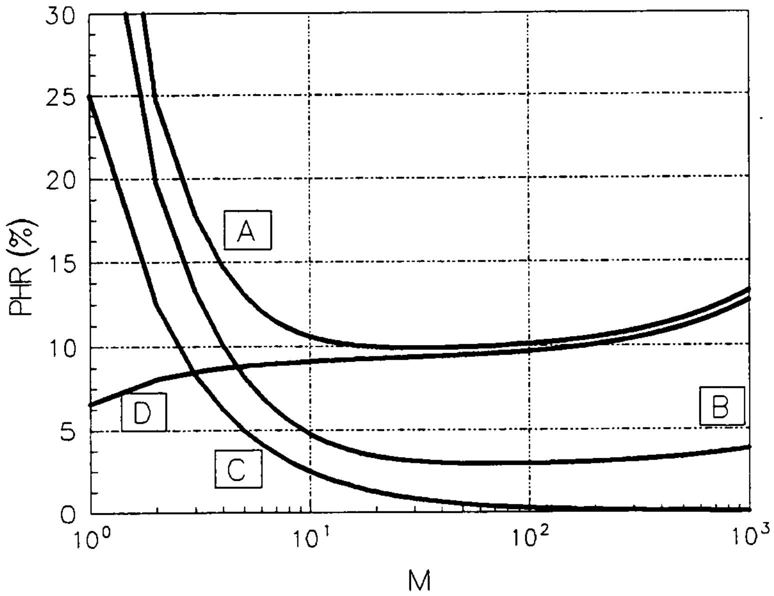

Figure 4-6. Calculated pulse height resolution vs gain for one 3mm x 3mm pixel in the custom APD array. CR-RC shaping time constant = 500 ns. A) Overall PHR, B) PHR due to parallel noise current, C) PHR due to series noise voltage, D) PHR due to statistical variation in N·M.

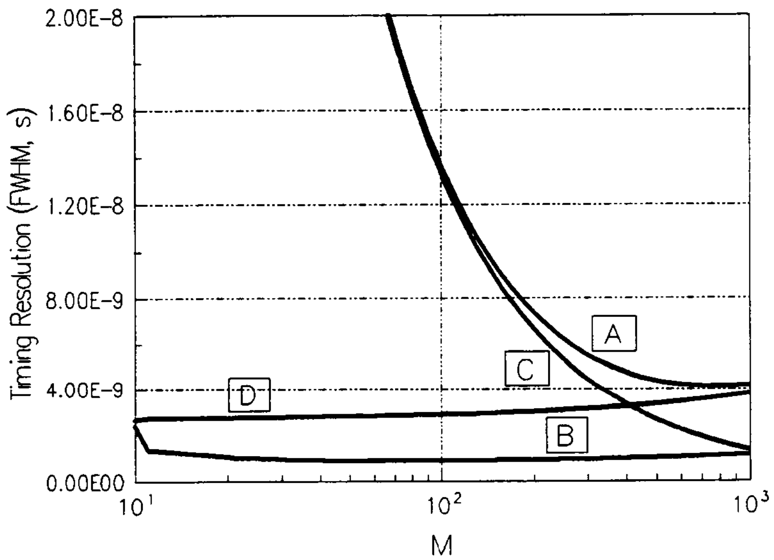

Figure 4-7. Calculated timing resolution vs gain for one 3mm x 3mm pixel in the custom APD array. CR-RC shaping time constant = 5 ns. A) Overall timing resolution, B) timing resolution due to parallel noise current, C) timing resolution due to series noise voltage, D) timing resolution due to statistical variations in r·M.

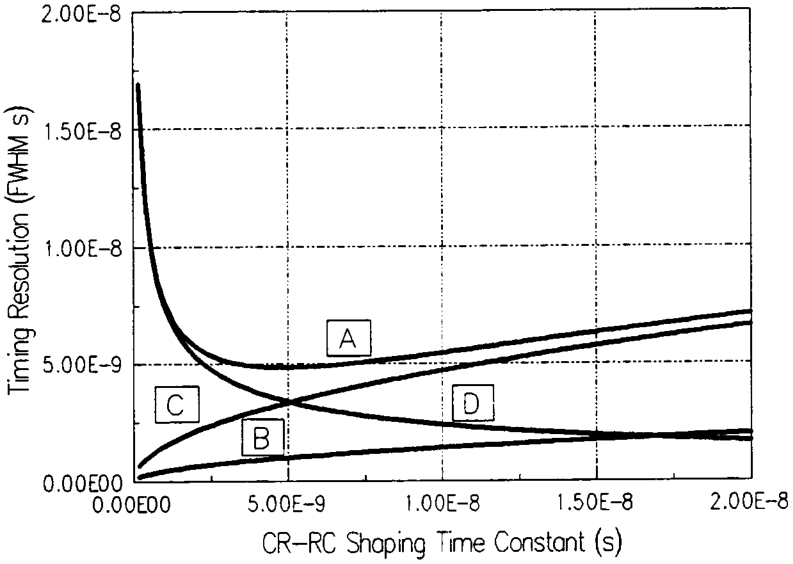

Figure 4-8. Calculated timing resolution vs CR-RC shaping time constant one 3mm x 3mm pixel of the custom APD array. APD gain = 500.

PART 5

DESIGN AND EVALUATION OF A LOW-NOISE, WIDE-BAND CMOS PREAMPLIFIER

Introduction

The previous chapters have been devoted to the analysis, design and evaluation of APD structures. As the data has shown, the performance of the APD/BGO detector is strongly influenced by the characteristics of the readout circuitry which processes the detector signal. The most important component in the readout circuit is the preamplifier, which must simultaneously have adequately low noise characteristics to permit accurate observation of the detector signals and have sufficient bandwidth to permit accurate measurement of the detector signal's time of arrival. This chapter details the design and evaluation of a custom CMOS preamplifier array designed to process the APD signals. The chapter begins with a discussion of basic MOSFET noise characteristics, followed by a discussion of the design procedure used in the selection of the MOS fabrication technology and basic amplifier topology. A detailed design description is then presented followed by a description of the amplifier's measured performance characteristics.

MOS Noise Sources

Although the principle of the surface field-effect transistor was first proposed in the early 1930s [1,2] and the first thermal oxide MOSFETs were fabricated in 1960 [3], the study of MOS noise characteristics is still an area of active research. The development of a universal noise model for MOSFETs has been hampered by the strong dependence of the noise characteristics on the fabrication process. As a consequence, the noise models found in simulation programs are in general less accurate than the large and small signal transfer function models [4, 5, 6]. Additionally, because the noise parameters are difficult to extract, they are less readily available than the transfer function model parameters. In this section the most widely accepted noise models are explored and some of the parameters found in the literature are reviewed.

Channel Thermal Noise

The classic channel thermal noise model presented here was developed for long channel devices operating in strong inversion and saturation. The basic source of noise considered by the model is the random motion of free carriers in the induced channel of a MOSFET which gives rise to thermal (or Johnson) noise analogous to the noise found in any resistive element. This noise is generally described by [7, 8]

$$\frac{i_d^2}{\Delta f} = 4kt\frac{\mu^2 W^2}{L^2 I_{DS}}\int Q_I^2(x)dx, \tag{5-1}$$

where $i_d^2/\Delta f$ is the power spectral density of the thermal noise current flowing in the channel, k is the Boltzman constant, T is the absolute temperature, μ is the channel surface mobility, W is the gate width, L is the gate length, I_{DS} is the dc current flowing in the channel and $Q_I(x)$ is the charge per unit length in the inversion channel. The inversion channel charge is generally written as [9]

$$Q_I = C_{ox}\left(V_{GS} - V_T(x) - V(x)\right), \tag{5-2}$$

where C_{ox} is the gate oxide capacitance, V_{GS} is the gate-source voltage, $V_T(x)$ is the position dependent threshold voltage and $V(x)$ is the position dependent channel potential. Invoking the gradual (long) channel approximation [10], $V(x)$ can be approximated as

$$V(x) = \frac{x}{L}V_{DS}, \tag{5-3}$$

where V_{DS} is the drain source voltage. Making the variable substitution $dx=(L/V_{DS})dV$ and substituting Equations 5-2 and 5-3 into Equation 5-1 yields

$$\frac{i_d^2}{\Delta f} = 4kT\frac{\mu^2 W^2}{L^2 I_{DS}}\int_0^{V_{DS}}\left(C_{ox}\left(V_{GS} - V_T - V\right)\right)^2 dV, \tag{5-4}$$

where the position dependence on V_T has been neglected for simplicity. Substituting the long-channel expression for I_{DS} in the non-saturated region of operation [11] yields,

$$I_{DS} = \frac{\mu C_{ox} W}{L}\left(\left(V_{GS} - V_T\right) - \frac{V_{DS}}{2}\right) \cdot V_{DS}, \tag{5-5}$$

and evaluating the integral yields

$$\frac{i_d^{\,2}}{\Delta f} = 4kT\mu C_{ox}\frac{W}{L}\frac{2}{3}\left(\frac{3\left(V_{GS} - V_T\right)V_{DS} - 3\left(V_{GS} - V_T\right)^2 - V_{DS}{}^2}{2\left(V_{GS} - V_T\right) - V_{DS}}\right). \tag{5-6}$$

At the onset of saturation $V_{DS} = V_{GS}\text{-}V_T$. Making this substitution yields

$$\frac{i_d^{\,2}}{\Delta f} = 4kT\frac{2}{3}\left[\mu C_{ox}\frac{W}{L}\left(V_{GS} - V_T\right)\right]. \tag{5-7}$$

The term in brackets is recognized as the standard expression for the device transconductance, g_m, so that Equation 5-7 reduces to the standard expression

$$\frac{i_d^{\,2}}{\Delta f} = 4kT\frac{2}{3}g_m. \tag{5-8}$$

Equation 5-8 is widely cited as the general expression for thermal channel noise in MOSFETs. Yet, it has been derived for the very specific case where V_T is position independent (lightly doped substrate), $V_{DS}=V_{GS}\text{-}V_T$ (onset of saturation), and I_D follows the simple long-channel expression. Because these conditions are rarely met in modern short-channel CMOS devices, an empirical noise factor, γ, is frequently used in place of the 2/3 term to obtain

$$\frac{i_d^{\,2}}{\Delta f} = 4kT\gamma g_m. \tag{5-9}$$

Reported γ values for saturated PMOS devices operating in strong inversion with substrate doping levels less than 10^{16} cm^{-3} range between 0.7 and 2.1 [5,12-16]. Possible reasons for this range in reported values will be explored below.

Several variations on Equation 5-9 exist in the literature. For example, Allen and Holberg[11] include the bulk-source transconductance term to provide some consideration of the noise due to the substrate resistance,

$$\frac{i_n^2}{\Delta f} = 4kT\frac{2}{3}(gm + gmbs). \tag{5-10}$$

As will be discussed below, however, the noise resistance due to substrate resistance is somewhat more complicated than this expression suggests and is typically treated as a separate noise source.

Van der Ziel [17] writes Equation 5-9 in terms of the channel conductance at $V_{DS} = 0$, g_{do}, to obtain

$$\frac{i_n^2}{\Delta f} = 4kT\gamma g_{do}. \tag{5-11}$$

Equation 5-11 has the advantage of being applicable in both the saturation and triode regions of operation. (Equation 5-9 fails below $V_{DS} = V_{DSAT}$ because $g_m \to 0$ as $V_{ds} \to 0$).

The preceding discussion has been limited to devices operating in strong inversion. In weak inversion, the MOSFET exhibits BJT-like behavior where the drain current is saturation is approximately given by [11],

$$I_D \approx \frac{W}{L} I_{DO} \exp\left(\frac{qV_{GS}}{nkT}\right). \tag{5-12}$$

I_{DO} is a process dependent parameter and n is the subthreshold slope factor which has an ideal value of 1. The subthreshold slope factor is strongly process dependent, however, and can have values in excess of 2 for processes with poor oxide interface characteristics [18]. Taking the derivative of Equation 5-12 with respect to V_{GS} yield the transconductance for a MOSFET in weak inversion,

$$g_m = \frac{qI_D}{nkT}, \tag{5-13}$$

which may be rewritten to obtain an expression for I_D in terms of g_m,

$$I_D \approx \frac{nkTg_m}{q}. \tag{5-14}$$

The channel noise of a MOSFET in weak inversion is frequently modeled as shot noise associated with injection of carriers over the potential barrier between the source and channel to obtain [9]

$$\frac{i_n^2}{\Delta f} = 2qI_D = 2nkTg_m. \tag{5-15}$$

Rewriting Equation 5-15 in the form of Equation 5-8 yields

$$\frac{i_n^2}{\Delta f} = 4kT\left(\frac{n}{2}\right)g_m. \tag{5-16}$$

Thus, for an ideal MOSFET (n=1) operating in weak inversion, $\gamma = 1/2$. As with the case of strong inversion, g_{do} is frequently substituted for g_m in equation 5-16 to obtain an expression which is valid in both the saturation and triode region [12].

The transition between weak and strong inversion is not abrupt. In the intermediate region of operation, moderate inversion, the MOSFET is not well described by either the strong or weak inversion model. Typically semi-empirical models are used to bridge the gap between the two regions [9]. The noise factor for an ideal MOSFET in moderate inversion thus will lie somewhere between 1/2 and 2/3. Equation 5-11 may be used for all regions of MOSFET operation if the appropriate values of γ are chosen [12].

$$\begin{array}{ll} \gamma = 0.5 & \text{weak inversion} \\ 0.5 \le \gamma \le 0.67 & \text{moderate inversion} \\ \gamma = 0.67 & \text{strong inversion} \end{array} \tag{5-17}$$

Experimental values for the noise factor are typically somewhat larger than the ideal values in Equation 5-17[14,15]. Furthermore, calculated data plots presented by Tedja *et al* [15] show that for a given drain current, g_{do} is typically larger than g_m, even at the onset of strong inversion, so the noise factor for Equation 5-9 should be somewhat greater than that for Equation 5-11.

Several factors contribute to higher-than-expected measured values of γ in Equations 5-8 and 5-11. Klassen and Prins [19] have shown that the assumption of a position independent V_T is only valid for the case of very low substrate doping, N_S. Their theoretical and experimental curves show that γ increases from the ideal value of 2/3 at $N_S = 10^{14}$ to approximately 1 at $N_S = 10^{15}$ and approximately 1.5 at $N_S = 10^{16}$. Above $N_S = 10^{16}$, γ increases rapidly, reaching a value of approximately 10 at $N_S = 10^{17}$. Klassen and Prins also show that the assumption of a position independent V_T is invalid for gate oxide thicknesses above 200 nm. Virtually all modern CMOS processes have gate oxides less than 50 nm, however, so this effect may be neglected.

Wang *et al* [5] show that the noise also increases for $V_{DS} > V_{D,SAT}$ in short channel devices due to the reduction in the effective gate length. This effect is treated in Equation 5-8 because g_m also increases as V_{DS} increases, but it is not treated in Equation 5-11.

Jindal [20] has shown that the noise factor increases from approximately 1 at $V_{DS} = V_{D,SAT}$ to greater than 2 for $V_{DS} = V_{D,SAT} + 2.5V$ in 1.25 μm NMOS devices due to hot electron effects. Hot carrier effects are known to occur to a somewhat lesser extent in PMOS devices as well[21].

Two other independent thermal noise sources exist in MOS devices which can also increase the *measured* noise factor. The polysilicon gate typically has a resistivity on the order of 25 Ω/□ [22]. The thermal noise due to the poly-gate modulates V_{GS}, inducing a noise current in the channel. Chang and Sansen have shown the thermal noise current due to the gate resistance to be given by [12]

$$\frac{i_d^2}{\Delta f} = 4kT\frac{R_g}{12n^2}g_m^2 \tag{5-18}$$

when both sides of the gate are shorted tied together and where R_g is the total gate resistance and n is the number of gate fingers. When only one side of the gate fingers are tied together the noise current increases to [12]

$$\frac{i_d^2}{\Delta f} = 4kT\frac{R_g}{3n^2}g_m^2. \tag{5-19}$$

For most multi-finger devices the effective gate resistance is much less than $1/g_m$ so that the noise due to the gate resistance is negligibly small with respect to the channel thermal noise.

A second white noise source exists due to the resistivity of the substrate. Jindal [13] has shown the channel noise current due to the substrate resistance to be given by

$$\frac{i_d^2}{\Delta f} = 4kTR_{sub}g_{mbs}^2, \tag{5-20}$$

where R_{sub} is the effective substrate resistance. R_{sub} may be minimized by careful plugging of the substrate and by operating the MOSFET with a low bulk-source reverse bias. Unfortunately, Jindal [13] notes that g_{mbs} increases as the bulk-source bias is reduced, so that the optimal V_{BS} represents a compromise between minimizing R_{sub} with a small V_{BS} and minimizing g_{mbs} with a large V_{BS}.

In the end, it is clear that the expressions given by Equations 5-8 and 5-11 cannot be blindly applied to all devices. Instead the designer must either measure the thermal noise characteristics of a particular process or choose a conservative value for the noise factor (2, for example) when developing a low-noise circuit.

Flicker Noise

Low frequency MOSFET noise is dominated by flicker or *1/f* noise. The research community has never held any illusion of having a universally applicable theory for flicker noise. Instead, the most commonly used flicker noise expressions are written in terms of process dependent parameters which must be empirically determined.

Two different theories have achieved prominence in describing the *1/f* noise. The Hooge model [23] assumes the *1/f* noise is attributable to fluctuations in the mobilities of free carriers when they collide with the crystal lattice. The current noise spectral density expression for this model is [12]

$$\frac{i_d^2}{\Delta f} = \alpha_l q \mu \frac{(V_{GS} - V_T) I_{DS}}{L^2 f}, \tag{5-21}$$

where α_l is the process dependent Hooge *1/f* noise parameter, q is the fundamental electron charge, and μ is the effective channel mobility. In saturation I_{DS} varies as $(V_{GS}-V_T)^2$, so the mobility fluctuation model predicts an $I^{3/2}$ dependence for the noise current spectral density.

The second widely accepted model is the number fluctuation model first proposed by McWhorter [24]. This model assumes the *1/f* noise is due to trapping and detrapping of carriers in the traps located at the Si-SiO_2 interface. The expression derived for this model is

$$\frac{i_d^2}{\Delta f} = \frac{K_F I_{DS}}{C_{ox} L^2 f}, \tag{5-22}$$

where K_F is a process dependent constant. Note that Equation 5-22 is dependent on I_{DS} rather than $I_{DS}^{3/2}$. Typical values for K_F are on the order of 10^{-32} C^2/cm^2 for PMOS devices and 3-4x$10^{-32}C^2/cm^2$ for NMOS devices [25]. Recent experimental data [26] suggests the flicker noise in PMOS devices is dominated by mobility fluctuations

(Equation 5-21) while the flicker noise in NMOS devices is dominated by carrier density fluctuations (Equation 5-22). Others have suggested that both mechanisms for *1/f* noise are simultaneously present in many devices and therefore use a combination of the two models [27]. The widely used circuit simulators SPICE[28] and PSPICE[29] attempt to reconcile the two models by introducing a variable exponential term, α, in Equation 5-22 to obtain

$$\frac{i_d^2}{\Delta f} = \frac{K_F I_{DS}{}^{\alpha}}{C_{ox} L^2 f}. \tag{5-23}$$

Of course, the appropriate value of K_F varies with the choice of α in Equation 5-23. The circuit simulator HSPICE [30] makes the substitution

$$g_m{}^2 = 2\,\mu C_{ox} \frac{W}{L} I_{DS} \tag{5-24}$$

to obtain

$$\frac{i_d{}^2}{\Delta f} = \frac{K_f g_m{}^2}{C_{ox} W L f^{AF}} \tag{5-25}$$

where $K_f = K_F/2\mu C_{ox}$ is the modified flicker noise parameter. Equation 5-25 has the advantage of explicitly presenting the reciprocal dependence of the flicker noise on the MOSFET area. The term *AF* permits the modeled frequency dependence of the noise to be set empirically.

Although flicker noise parameters have been published for many CMOS processes, the designer is once again forced to either measure the *1/f* noise characteristics of a particular process or use appropriately conservative values for K_F or α_I when developing a circuit.

Complete Noise Expression

Although Equation 5-11 has been shown to be a more accurate predictor of CMOS thermal noise, the term g_{do} is difficult to measure. Consequently, for this discussion the approximation given in Equation 5-9 shall be used to describe MOSFETs operating in saturation. Additionally, the more widely used number fluctuation model (Equation 5-22) shall be used to model the 1/f noise. A complete expression for MOSFET channel noise current is then obtained by summing Equations 5-9, 5-22, 5-19 and 5-20 to obtain

$$\frac{i_d^{\,2}}{\Delta f} \approx 4kT\gamma g_m + \frac{K_F I_{DS}}{C_{ox} L^2 f} + 4kT\frac{R_g}{3n^2} g_m^{\,2} + 4kTR_{sub} g_{mbs}^{\,2}. \tag{5-26}$$

It is often convenient to translate the channel noise current to an equivalent noise voltage at the gate by dividing Equation 5-26 by $g_m^{\,2}$ to obtain

$$\frac{e_n^{\,2}}{\Delta f} \approx \frac{i_d^{\,2}}{\Delta f} \cdot \frac{1}{g_m^{\,2}} = 4kT\frac{\gamma}{g_m} + \frac{K_F I_{DS}}{C_{ox} L^2 g_m^{\,2} f} + 4kT\frac{R_g}{3n^2} + 4kTR_{sub}\frac{g_{mbs}^{\,2}}{g_m^{\,2}}. \tag{5-27}$$

The *1/f* term in Equation 5-27 may be simplified by recalling the square-law (long channel) expression for MOSFET transconductance,

$$g_m = \sqrt{2\mu C_{ox} I_{DS} \frac{W}{L}} \tag{5-28}$$

and substituting into Equation 5-27 to obtain the general expression for the MOSFET equivalent input noise voltage,

$$\frac{e_n^{\,2}}{\Delta f} \approx 4kT\frac{\gamma}{g_m} + \frac{K_F}{2\mu C_{ox}^{\,2} WLf} + 4kT\frac{R_g}{3n^2} + 4kTR_{sub}\frac{g_{mbs}^{\,2}}{g_m^{\,2}}. \tag{5-29}$$

Equation 5-29 is the basic design equation which will be used in the following sections to develop a low noise preamplifier array for use with the APD/BGO detector.

Basic Amplifier Topology

Overview

In the previous section the noise sources associated with CMOS transistors were reviewed and the standard expression for the MOSFET equivalent noise voltage was presented. In this section, two commonly used preamplifier topologies are compared and a basic topology is selected for the preamplifier array design.

The input to the preamplifier is provided by the APD which may be modeled as a high-impedance current source with a load capacitance appearing at the output. High output impedance signal sources require low input impedance preamplifiers to prevent the load capacitance from limiting the signal bandwidth. The two most commonly used preamplifier topologies meeting this requirement are the charge-sensitive preamplifier (Figure 5-1) and the transimpedance preamplifier (Figure 5-2)[12]. The charge sensitive topology is usually preferred for detectors with low signal levels because the feedback capacitor does not introduce noise current at the amplifier input [31]. Binkley *et al* [16] have suggested that the transimpedance amplifier may be used with APDs with very good results, however, because the noise introduced by the feedback resistor can be made small with respect to the excess noise generated by the APD. The purpose of the analysis presented here is to examine this argument and to choose the most appropriate topology for the amplifier array design.

In this section the signal and noise transfer functions for the two preamplifier topologies are presented and the expected pulse height resolution and timing resolution of the circuits using two preamplifier topologies when driven by a typical APD are calculated using the circuit simulator PSPICE [29].

Charge Sensitive Topology

The charge sensitive preamplifier configuration (Figure 5-1) consists of an inverting voltage-sensitive amplifier with a feedback capacitor, C_F. In practice, a reset element such as a large resistor or a switch would be connected in parallel with C_F to close the dc loop of the amplifier. The amplifier noise is modeled by the input equivalent noise voltage source, *ENV*, and the input equivalent noise current source, *ENI*. The APD noise current is modeled by *END* and the APD capacitance is modeled by C_D.

The preamplifier will normally drive a shaping amplifier for both pulse height resolution and timing resolution measurements. A CR-RC shaping amplifier is selected for both measurements in this analysis, and its transfer function is also indicated in Figure 5-1. The various signal and noise transfer functions for the charge sensitive preamplifier and shaping amplifier are listed in Table 5-1, along with the expressions for the transimpedance preamplifier discussed below. The charge sensitive preamplifier transfer function and noise bandwidth expressions in Table 5-1 are written in terms of the preamplifier feedback capacitance (C_F), and the CR-RC shaping amplifier gain (A) and shaping time constant (τ). The expressions for $V_{o,peak}$ and *dV/dt(max)* describe the peak signal height and peak signal slope for a BGO light pulse of N photons coupled to an APD with a quantum efficiency of η, and a gain of M. The expressions for $V_{o,peak}$ and *dV/dt(max)* assume the APD/BGO signal may be modeled as an impulse function for pulse height resolution measurements and as a step function for timing resolution measurements. Although these approximations offer valuable insight into the expected behavior of the detector and preamplifier, they are not strictly valid for the APD/BGO detector due to the 300 ns decay time constant associated with the BGO light output.

Table 5-1.

Summary of Important Equations Describing Signal Processing and Noise Characteristics of Ideal Charge Sensitive and Transimpedance Preamplifiers.

Parameter	**Charge Sensitive Preamplifier**	**Transimpedance Preamplifier**
H(s)	$\frac{-1}{C_F}\left[\frac{A\tau}{(1+s\tau)^2}\right]$	$\frac{-R_F A}{(1+s\tau)^2}$
h(t)	$\frac{-A}{C_F\tau}\cdot t\exp\left(\frac{-t}{\tau}\right)$	$\frac{-R_F A}{\tau^2}\cdot t\exp\left(\frac{-t}{\tau}\right)$
$\lvert H(\omega)\rvert^2$ (noise current)	$\left(\frac{A}{C_F}\right)^2\left(\frac{\tau^2}{1+2\omega^2\tau^2+\omega^4\tau^4}\right)$	$\frac{(R_F A)^2}{1+2\omega^2\tau^2+\omega^4\tau^4}$
Noise Power Bandwidth (noise current)	$\frac{1}{8}\left(\frac{A}{C_F}\right)^2\tau$	$\frac{1}{8}\frac{(R_F A)^2}{\tau}$
$\lvert H(\omega)\rvert^2$ (noise voltage)	$\left(\frac{C_F+C_D}{C_F}\right)^2\left(\frac{A^2\tau^2\omega^2}{1+2\omega^2\tau^2+\omega^4\tau^4}\right)$	$A^2\left(\frac{1+\omega^2 R_F^2 C_D^2}{1+2\omega^2\tau^2+\omega^4\tau^4}\right)$
Noise Power Bandwidth (noise voltage)	$\frac{1}{8}A^2\left(\frac{C_F+C_D}{C_F}\right)^2\frac{1}{\tau}$	$\frac{1}{8}A^2\left(\frac{\tau^2+R^2C_D^2}{\tau^3}\right)$
$V_o$peak (impulse input)	$\frac{q\eta\langle N\rangle\langle M\rangle A}{C_F}\exp(-1)$	$\frac{q\eta\langle N\rangle\langle M\rangle R_F A}{\tau}\exp(-1)$
$\left.\frac{dV}{dt}\right\rvert_{max}$ (step input current of N/τ_{bgo} A)	$\left(\frac{q\eta\langle N\rangle\langle M\rangle A}{\tau_{BGO}C_F}\right)\exp(-1)$	$\left(\frac{q\eta\langle N\rangle\langle M\rangle R_F A}{\tau_{BGO}\tau_{SH}}\right)\exp(-1)$

In order to more accurately calculate the contributions of the amplifier and APD noise to the overall PHR and timing resolution of the charge sensitive preamplifier and shaping amplifier, the circuit shown in Figure 5-1 is modeled using PSPICE. The preamplifier is modeled as an ideal gain element ($A_v > 10^6$, bandwidth $\rightarrow \infty$, $C_{in} = 0$) with an equivalent input noise voltage of 2 nV/√Hz and an equivalent input noise current of 50 fA/√Hz. The BGO signal is modeled as a current pulse with a 1 ns rise time and a decaying exponential tail with a 300 ns time constant. The APD model is based on the published parameters for a 5 mm diameter Advanced Photonix APD biased for a gain of 200 [31,32]. The model parameters are summarized in Table 5-2. Because the analysis is focused on the noise of the preamplifier, the statistical noise of the BGO signal and APD gain process are not considered.

Figure 5-3 shows the PSPICE calculated rms output noise voltage of the charge sensitive preamplifier and CR-RC shaping amplifier due to (A) the noise current sources ENI and END, (B) the noise voltage source ENV, and (C) all three noise sources as a function of shaping time constant, where the shaping time constant range is selected for optimal pulse height resolution (100 ns - 1μs). The parallel noise current sources are dominant over the entire range and increase with increasing shaping time constant. Figure 5-4 shows the PSPICE calculated peak voltage and pulse height resolution of the charge sensitive preamplifier and CR-RC shaping amplifier due to the electronic noise sources as a function of shaping time constant. Unlike the data presented in Chapter 3 for an impulse function input, the peak voltage increases with increasing shaping constant due to the long (300 ns) decay of the BGO signal. The best "electronic noise" PHR (4.6%) is obtained for a shaping time constant of 200 ns. Note that the optimal time constant is strongly influenced by the 300 ns decay time constant of the BGO signal. If the BGO signal was a true impulse function, the optimal time constant would be much shorter (41 ns).

Table 5-2.
Summary of Model Parameters Used in PSPICE Simulations.

Parameter	**Value**
Scintillator photon yield	N=1500
BGO decay time constant	τ_{BGO}=300 ns
APD Quantum Efficiency	η=0.85
APD Capacitance	C_D=23 pF
APD k_{eff}	k_{eff}=0.0017
APD Gain	M = 200
Un-multiplied APD bulk dark current	I_{db}=43 pA
APD surface dark current	I_{ds}=60 nA
APD noise current (M=200)	END=1.16 pA/√Hz
Shaping Amplifier Gain	A=100
Preamplifier input noise current	ENI=50fA/√Hz
Preamplifier input noise voltage	ENV=2nV/√Hz
Charge sensitive preamplifier feedback capacitance	C_F=1 pF
Transimpedance preamplifier feedback resistance	R_F=50 kΩ
Input noise current due to feedback resistance	$ENIR$=576fA/√Hz

Figure 5-5 again shows the rms output noise voltage due to (A) the noise current sources ENI and END, (B) the noise voltage source ENV, and (C) all three noise sources as a function of shaping time constant, where in this case the shaping time constant range is selected for optimal timing resolution (1 ns - 50 ns). At these short shaping time constants the noise is dominated by the series noise voltage, indicating the importance of minimizing the amplifier noise voltage for timing measurements.

Figure 5-6 shows the PSPICE calculated output slope, *dV/dt (max)* and the timing resolution due to the noise sources as a function of the shaping time constant. At these short shaping time constants the step function approximation is reasonably valid. The maximum slope changes by less than 15% over the entire shaping time constant range and best timing resolution (5.99 ns FWHM) is obtained with the shaping time constant which minimizes the total output noise, near 35 ns. In practice the preamplifier will be coupled to an APD with a large noise current. Consequently, the optimal shaping time constant for timing measurements will be less than 35 ns for the detector/preamplifier system.

Transimpedance Topology

The transimpedance preamplifier configuration (Figure 5-2) consists of an inverting voltage-sensitive amplifier with a resistive feedback element, R_F. In addition to the noise sources described for the charge sensitive preamplifier, the transimpedance preamplifier has a noise current source, *ENIR*, due to the thermal noise of the feedback resistor. Because the preamplifier does not integrate the signal, a 2 pole low pass (RC-RC) shaping amplifier is used in place of the CR-RC amplifier. The various signal and noise transfer functions are listed along with those for the charge sensitive preamplifier in Table 5-1 in terms of the preamplifier feedback resistance (R_F), and the RC-RC shaping amplifier gain (A) and shaping time constant (τ).

Unlike the charge-sensitive preamplifier, the peak output voltage of the transimpedance preamplifier is a function of the shaping time constant. This difference occurs because the charge sensitive feedback impedance, $1/sC_F$,decreases with increasing frequency while the transimpedance feedback impedance is frequency independent. For the two topologies to have equivalent output signals with a given shaping time constant, the feedback elements must be selected such that $\tau/C_F=R_F$ or $\tau=C_FR_F$.

As before, the circuit shown in Figure 5-2 is modeled using PSPICE and the model parameters presented in Table 5-2. The feedback resistor value (50 kΩ) is chosen such that its noise current is approximately half that of the APD. Figure 5-7 shows the rms output noise voltage due to (A) the noise current sources *ENI*, and *END*, (B) the noise voltage source *ENV*, (C) the noise current due to the feedback resistor, *ENIR*, and (D) all three noise sources as a function of shaping time constant, where the shaping time constant range is selected for optimal pulse height resolution (100 ns - 1μs). As before, the parallel noise current sources are dominant over the entire range. Figure 5-8 shows the PSPICE calculated peak voltage and pulse height resolution due to the noise sources as a function of shaping time constant. Unlike the charge sensitive preamplifier, both the signal and noise decrease with increasing time constant. As with the charge sensitive preamplifier, the best PHR (5.1%) is obtained with a 200 ns shaping time constant. The "electronic noise" PHR is somewhat poorer in the transimpedance preamplifier due to the additional noise of the feedback resistor.

Figure 5-9 shows the rms output noise voltage due to (A) the noise current sources *ENI* and *END*, (B) the noise voltage source *ENV*, (C) the noise current *ENIR*, and (D) all three noise sources over a shaping time constant range of 1 ns to 50 ns. As with the charge sensitive preamplifier, the noise is dominated by the series noise voltage. Figure 5-10 shows the PSPICE calculated output slope, *dV/dt (max)* and the timing resolution due to the noise sources as a function of the shaping time constant. Again the optimal "electronic noise" timing resolution (6.15 ns FWHM) is obtained with a shaping time constant of 35 ns, and the transimpedance preamplifier timing resolution is slightly poorer than that of the charge sensitive preamplifier.

Other Considerations

Two other considerations influence the choice between the charge sensitive and transimpedance preamplifier topologies. As mentioned above, a reset device, typically a several hundred megohm resistor, must be added to the charge sensitive preamplifier feedback path to close the dc loop. This resistor introduces a low-frequency pole which introduces some undershoot in the impulse response. For accurate pulse-height resolution measurements using a charge sensitive preamplifier, a "pole-zero" compensation network must be added to the shaping amplifier. On the other hand, the transimpedance preamplifier has no need for "pole-zero" compensation, but the feedback resistor introduces a pole in the loop transfer function at the input. For the 50 kΩ resistor and 23 pF detector capacitance considered above, the input pole would have a corner frequency of 5.5 MHz. In order to achieve rise times less than 10 ns, the input pole would have to become the dominant pole in the loop transfer function and the secondary poles would be required to have corner frequencies greater than ~1GHz.

Selected Topology

As reported by Binkley et al [16], the transimpedance preamplifier has only slightly poorer noise characteristics than the charge sensitive when used with APD based detectors. For detectors with low noise currents, such as PIN diodes, however, the transimpedance preamplifier is clearly less desirable due to the added noise current of the feedback resistor. For this design, the charge sensitive topology was selected due to its slightly superior noise performance with APD based detectors and due to its potential for use with other non APD based detectors with lower noise currents.

Design

Circuit Configuration

Traditionally, a cascode topology is used for low-noise, wide-band pulse processing readout circuits[16, 34, 35]; an example of a folded cascode stage based on the circuit of [16] is shown in Figure 5-11. When selecting a cascode topology, the designer typically has the following objectives:

1. Reduction of amplifier input capacitance. The low impedance at the input device drain (*M1* in Figure 5-11) established by the cascode device (*M2* in Figure 5-11) reduces the Miller effect on the *M1* gate-drain capacitance.
2. Development of large loop gain. The high output impedance at the drain of *M2* increases the amplifier open-loop gain for improved amplifier gain stability and linearity. In general, cascode preamplifiers are designed with a single gain node (at the drain of *M2*) in order to achieve near single-pole characteristics.
3. Development of near single-pole frequency response. The low impedance at the drain of *M1* prevents the formation of a second dominant pole, preserving the single-pole behavior of the amplifier.

Unfortunately, these objectives are not always realized with CMOS cascode circuits which use input FETs with large *W/L* ratios. In this section the folded cascode topology of [16] is evaluated against the objectives presented above and an alternate topology which better meets those objectives is proposed. The new topology forms the basis for the CMOS design presented in the next section.

Consider first the *AC* equivalent circuit shown in Figure 5-11. The input PMOS device has a large *W/L* ratio (2400 μm/2 μm) and is biased with a large drain current (5 mA) in order to obtain a large g_m for low noise and high loop gain. The NMOS cascode device has a *W/L* ratio of 200 μm/2 μm and is biased with a 200 μA drain current. In order to achieve a large loop gain, the drain of *M2* must look into a large impedance, so a load impedance, R_{load}, of 1 MΩ is assumed. For this analysis, the body effects on the devices will be neglected and the devices will be assumed to have the small signal parameters listed in Table 5-3, which are based on extracted SPICE model parameters from a typical ORBIT Semiconductor 2 μm NWELL process [36].

Table 5-3.
Small Signal Parameters Used in Cascode Circuit Analysis.

Device	W/L Ratio	g_m	g_{ds}	r_{ds}
M1	2400/2	19 mS	1.5 mS	667 Ω
M2	200/2	1.5 mS	66 μS	15 kΩ

Analysis of the MOSFET small signal equivalent circuit model shows the input impedance (into the source) of *M2* to be given by

$$R_{in,M2} = \frac{r_{ds2} + R_{load}}{g_{m2}\, r_{ds2} + 1}. \tag{5-30}$$

Substituting the values listed in Table 5-3 and letting R_{load} = 1MΩ yields

$$R_{in,M2} = 43k\Omega. \tag{5-31}$$

The input impedance of the common gate transistor is much larger than the drain resistance of *M1*, so instead of having a reduced voltage gain as desired, *M1* develops the maximum achievable low-frequency voltage gain for a common source amplifier, $A_{max} = -g_{m1}r_{ds1} = -12.7$. Consequently the objectives of minimizing the input Miller capacitance and of preventing the occurrence of a significant pole at the drain of *M1* are not achieved.

The output impedance of the cascode amplifier may be shown to be

$$R_{out,M2} = r_{ds1} + r_{ds2} + g_{m2} r_{ds1} r_{ds2}. \tag{5-32}$$

Substituting the data from Table 5-3 yields

$$R_{out,M2} = 31k\Omega. \tag{5-33}$$

Thus, the output impedance of *M2* is only a factor of 2 larger than the intrinsic r_{ds2}, and is actually less than the input impedance.

The voltage gain of the cascode amplifier may be shown to be

$$A_v = \frac{-g_{m1} R_{load} \left(r_{ds1} + g_{m2} r_{ds1} r_{ds2} \right)}{r_{ds1} + r_{ds2} + R_{load} + g_{m2} r_{ds1} r_{ds2}}. \quad (5\text{-}34)$$

Again letting R_{load} = 1MΩ and substituting the values from Table 5-3 yields

$$A_v = -289. \quad (5\text{-}35)$$

Although this gain is sufficiently large to permit the amplifier to function, it is lower than might be desired for a charge-sensitive amplifier coupled to a detector with a large capacitance since the loop gain is reduced by the factor $(C_F + C_d + C_{in})/C_F$, where C_F is the feedback capacitance, C_d is the detector capacitance and C_{in} is the amplifier input capacitance. For example, if the detector has a capacitance of 20 pF, the preamplifier has an input capacitance of 5 pF and the feedback capacitor has a value of 1 pF, the low-frequency loop gain will be reduced to -11.

Based upon this analysis, it is evident that the design objectives of a low input impedance and high output impedance for the cascode device are not fully realized in the simple CMOS circuit. An alternate topology which better meets these objectives is presented below.

Consider now the circuit shown in Figure 5-12. As before, *M1* (*W/L* = 2400μm/2μm) is the input device, *M2* (*W/L* = 200μm/2μm) is the cascode device, and a 1 MΩ load impedance is assumed to appear at the drain of *M2*. Also, as before, *M1* and *M2* are biased with a 5 mA and 200 μA drain currents, respectively. A local feedback loop is formed around *M2* by *M4* and *M5*, and *M3*, *M6* and *M7* serve to bias the circuit. Although this circuit topology is believed to be new, the configuration of *M2-M7* may be recognized as the "super-regulated cascode current mirror" of [37], and the concept of driving a regulated current mirror with a common source amplifier has been presented elsewhere [34].

For the purposes of this analysis, the effects of modulation of the bulk-source potential are neglected and all of the NMOS aspect ratios and drain currents are set to 200μm/2μm and 200 μA, respectively. With these assumptions, the small signal parameters of the transistors for a typical ORBIT NWELL process are summarized in Table 5-4.

Table 5-4.
Small Signal Parameters Used in Improved Cascode Circuit Analysis.

Device	W/L Ratio	g_m	g_{ds}	r_{ds}
M1	2400/2	19 mS	1.5 mS	667 Ω
M2-7	200/2	1.5 mS	66 μS	15 kΩ

The input impedance of the regulated cascode stage (*M2* source) with respect to ground may be calculated using Blackman's theorem [38],

$$R_{in,M2} = R_{in,no\ feedback}\left(\frac{1-T_{sc}}{1-T_{oc}}\right), \tag{5-36}$$

where $R_{in,no\,feedback}$ is the input impedance calculated for the simple cascode circuit, T_{sc} is the loop gain of the regulated current mirror with the source of *M2* shorted to ground and T_{oc} is the loop gain of the regulated current mirror with the source of *M2* open. It is immediately evident that $T_{sc} = 0$. T_{oc} is given by

$$T_{oc} = A_{vM5M4} \cdot A_{vM2}, \tag{5-37}$$

where A_{vM5M4} is the voltage gain of the cascode stage formed by *M5* and *M4* and A_{vM2} is the voltage gain of the source follower stage formed by *M2*. As before, the voltage gain of the *M5-M4* cascode stage is given by

$$A_{vM5M4} = \frac{-g_{m5}R_{ibias}(r_{ds5} + g_{m4}r_{ds4}r_{ds5})}{r_{ds4} + r_{ds5} + R_{ibias} + g_{m4}r_{ds4}r_{ds5}}. \tag{5-38}$$

Letting $R_{ibias} = 1$ MΩ and using the values in Table 5-4 yields

$$A_{vM5M4} = -386. \tag{5-39}$$

The gain of the common source amplifier formed by *M2* is given by

$$A_{vM2} = \frac{g_{m2}}{g_{m2} + g_{ds2} + g_{ds3} + R_{load} g_{ds2} g_{ds3}}. \tag{5-40}$$

Note that g_{ds1} does not appear in Equation 5-40 because this is an *open circuit* loop gain calculation. The effects of *M1* on the loop gain are considered when calculating the circuit gain. Letting R_{load} = 1 MΩ and using the values in Table 5-4 yields

$$A_{vM2} = 0.25. \tag{5-41}$$

Substituting Equations 5-39 and 5-41 into Equation 5-37 yields the open circuit loop gain,

$$T_{oc} = A_{vM5M4} \cdot A_{vM2} = -96.7. \tag{5-42}$$

The input impedance without the effects of the feedback loop is

$$R_{in,no\,feedback} = \frac{r_{ds2} + R_{load}}{g_{m2} r_{ds2} + 1} \Bigg\| r_{ds3}, \tag{5-43}$$

or

$$R_{in,no\,feedback} = 43k\Omega \| 15k\Omega = 11k\Omega. \tag{5-44}$$

Substituting Equations 5-42 and 5-44 into Equation 5-36 yields

$$R_{in,M2} = 11k\Omega \frac{1-0}{1+96.7} = 113\Omega. \tag{5-45}$$

Thus, the local feedback reduces the low-frequency cascode input impedance by a factor of 380, and the input impedance is now much less than r_{ds1}. The voltage gain of *M1* is given by

$$A_{vM1} = -g_{m1}\left(r_{ds1} \| R_{in}\right) = -1.86, \tag{5-46}$$

much less than the gain of -12.7 obtained with the simple cascode topology.

Likewise, the output impedance is obtained using Blackman's theorem,

$$R_{out,M2} = R_{out,no\,feedback}\left(\frac{1-T_{sc}}{1-T_{oc}}\right), \tag{5-47}$$

where, in this case, T_{sc} is the loop gain with the drain of *M2* shorted to ground and T_{oc} is the loop gain with the drain of *M2* open. As before, the loop gain is given by the product $A_{vM5M4} \cdot A_{vM2}$. Opening the drain of *M2* effectively increases R_{load} to infinity, and from Equation 5-40, as $R_{load} \to \infty$, $A_{vM2} \to 0$ and thus $T_{oc} \to 0$. The short circuit loop gain is given by

$$T_{sc} = A_{vM5M4} \cdot A_{vM2} = A_{vM5M4} \cdot \left(\frac{g_{m2}}{g_{m2} + g_{ds1} + g_{ds2} + g_{ds3}} \right). \tag{5-48}$$

The expression for A_{vM2} differs from Equation 5-40 in that g_{ds1} is included since the source of *M2* is no longer open circuited and in that the R_{load} term is neglected since R_{load} = 0 for the short circuit calculation. Using the parameters in Table 5-4 yields

$$T_{sc} = A_{vM5M4} \cdot A_{vM2} = -386 \cdot 0.48 = -185. \tag{5-49}$$

Neglecting the effects of feedback, the output impedance is given by

$$R_{out,no\ feedback} = r_{ds2} + r_{ds3} \| r_{ds1} + g_{m2} r_{ds2} \left(r_{ds3} \| r_{ds1} \right). \tag{5-50}$$

Evaluating Equation 5-50 yields

$$R_{out,no\ feedback} = 30k\Omega, \tag{5-51}$$

which is essentially the output impedance calculated for the simple cascode circuit. Substituting Equations 5-49 and 5-51 into Equation 5-47 yields the output impedance,

$$R_{out,M2} = 30k\Omega(1+185) = 5.6M\Omega, \tag{5-52}$$

which is approximately 180 times greater than the low-frequency output impedance of the simple cascode circuit.

The voltage gain of the amplifier is given by

$$A_v = A_{v,ideal} \left(\frac{-T}{1-T} \right), \tag{5-53}$$

where $A_{v,ideal}$ is the gain of an ideal cascode circuit and T is the loop gain of the local feedback loop. The ideal gain is simply

$$A_{v,ideal} = -g_{m1} R_{load} = -(19mS)(1M\Omega) = -19 \times 10^3 . \qquad (5\text{-}54)$$

The loop gain is given by

$$T = \left(\frac{-g_{m5} R_{ibias} \left(r_{ds5} + g_{m4} r_{ds4} r_{ds5} \right)}{r_{ds4} + r_{ds5} + R_{ibias} + g_{m4} r_{ds4} r_{ds5}} \right) \cdot \left(\frac{g_{m2}}{g_{m2} + g_{ds1} + g_{ds2} + g_{ds3} + R_{load} g_{ds2} \left(g_{ds1} + g_{ds3} \right)} \right), \qquad (5\text{-}55)$$

where the first term is A_{vM5M4} and the second term is A_{vM2}. Note that for the overall loop gain the A_{vM2} term includes the effects of both *M1* and R_{load}. Solving Equation 5-55 yields

$$T = (-386) \cdot (0.014) = -5.4 . \qquad (5\text{-}56)$$

Note the significant reduction in the loop gain caused by the large load resistance and the low output impedance of *M1*. Substituting Equation 5-54 and 5-56 into Equation 5-53 yields

$$A_v = 19 \times 10^6 \frac{-5.4}{1 - 5.4} = -16 \times 10^3 . \qquad (5\text{-}57)$$

Thus, the low frequency voltage gain is increased by a factor of 55 over that of the simple cascode circuit.

Based upon this analysis, the improved cascode circuit was adopted for this design. The circuit design is presented below. As indicated by the simulated circuit performance, the local feedback loop is stable without additional compensation, and its bandwidth is adequately high to prevent the feedback loop from having any noticeable effect on the simulated circuit behavior.

Circuit Design

A six channel preamplifier array was designed for use with avalanche photodiodes. The circuit schematic for one of the preamplifiers is shown in Figure 5-13. Based upon the discussion presented above, the following design choices were made:

1. The charge-sensitive topology was selected over the transimpedance topology for its lower noise and to maintain the option of using the preamplifier with other detectors with lower noise currents than APD based detectors.
2. The improved cascode topology was selected over the standard cascode topology to reduce the input capacitance, increase the low-frequency open loop gain, and to achieve near single-pole performance.

The design was fabricated in the ORBIT Semiconductor 2 μm foundry. The 2 μm process was selected primarily because the MOSIS prototype fabrication service [36] has thoroughly documented the process and has published measured simulation model parameters for a number of recent fabrication runs. Some improvement in bandwidth and noise characteristics may be attained in future design iterations by using ORBIT's newer 1.2 μm process.

Traditionally an NWELL process is selected for analog circuits to optimize the performance characteristics of the NMOS devices. NMOS devices are the preferred gain elements for most analog designs due to the superior surface mobility of electrons. In the NWELL process, the NMOS transistors are fabricated in the low-doped native substrate, giving them the highest possible surface mobility and transconductance and the lowest possible body bias effects and surface area. The NWELL process was selected for the preamplifier array design because the array will ultimately be incorporated into a larger analog system. As a result of this choice, the input PMOS device is located in a well. It is not clear from the literature what effect placing the input device in a well will have on the noise performance. On one hand, evidence exists to suggest the thermal noise performance will be somewhat degraded due to the higher doping in the well [19]. On the other hand the substrate potential has been shown to affect MOSFET noise[13], and this configuration permits the substrate potential to be empirically optimized. For the simulations presented below, the input device well is tied to V_{DD}; in the circuit implementation the input device well contact is brought out to a pin to permit experimental optimization of the well bias voltage.

The input device, *M1*, is a 7500 μm /2 μm PMOS transistor. *M1* was laid out as an array of 30-250 μm/2 μm transistors with significant substrate plugging between the smaller transistors to minimize the noise due to the substrate resistance. An extra PMOS cascode transistor (*M2*) was placed between *M1* and the improved cascode circuit (*M3, M8, M12, M11*) to further reduce the load at the drain of *M1*, increase the loop gain of the local feedback, and to permit independent control of V_{DS1}. The drain current of *M1* and *M2* is set by a simple current mirror (*M18* and *M5*) and can be independently selected to achieve optimal noise performance. For the calculations presented below $Id_{M1} = Id_{M2} = 10$ mA.

The cascode device with local feedback is M3 (400 μm/2 μm NMOS), and the local feedback is formed by *M8*, *M12*, and *M11*. The load at the drain of *M3* (I_{load} in Figure 5-12) is formed by a super-regulated cascode current mirror [37] *using M6, M7, M9 and M10*. The bias current for the improved cascode circuit (I_{bias2} in Figure 5-12) is formed with a simple cascode current mirror (*M13* and *M14*). The dominant node is Node 12, and the poly-poly compensation capacitor, C_{comp} was inserted to set the *total* capacitance at Node 12 to ~6.5 pF. The class A output driver is an NMOS source follower (*M4*) biased with a simple current mirror (*M7*). All of the transistors marked with (*) are common to all six preamplifiers on the chip.

Open Loop Gain

The open loop gain is calculated following the analysis presented above. The PSPICE operating point data listed in Table 5-5 is used for all of the hand calculations. As before, the effects of bulk-source modulation are neglected for the hand analysis.

In order to calculate the loop gain, the output impedances of each of the current mirrors must be calculated. The output impedance of the cascode mirror formed by *M13* and *M14* (R_{bias1}) is

$$R_{bias1} = r_{ds14} + r_{ds13} + g_{m14} r_{ds13} r_{ds14} = 1.10\ M\Omega, \tag{5-58}$$

and the output impedance of the cascode mirror formed by *M15* and *M16* (R_{bias2}) is

$$R_{bias2} = r_{ds15} + r_{ds16} + g_{m15} r_{ds15} r_{ds16} = 467\ k\Omega. \tag{5-59}$$

The impedance looking into the super-regulated cascode current mirror (*M6*, *M7*, *M9* and *M10*) is obtained using Blackman's theorem (Equation 5-36). Neglecting the feedback, the impedance looking into the drain of *M7* is

$$R_{load1,nf} = r_{ds7} + r_{ds6} + g_{m7} r_{ds6} r_{ds7} = 71.6\ k\Omega. \tag{5-60}$$

Table 5-5.
Operating Point Data used for Hand Calculations.

Device	g_m	g_{ds}	r_{ds}	Transistor Type
M1	44 mS	3.68 mS	272 Ω	PMOS
M2	20 mS	2.01 mS	498 Ω	PMOS
M3	4.03 mS	152 μS	6.58 k Ω	NMOS
M4	6.17 mS	171 μS	5.85 k Ω	NMOS
M5	14.8 mS	830 μS	1.20 k Ω	NMOS
M6	1.53 mS	209 μS	4.78 k Ω	PMOS
M7	3.45 mS	262 μS	3.82 k Ω	PMOS
M8	1.48 mS	82 μS	12.2 k Ω	NMOS
M9	1.43 mS	142 μS	7.04 k Ω	PMOS
M10	2.61 mS	257 μS	3.89 k Ω	PMOS
M11	3.25 mS	125 μS	8.00 k Ω	NMOS
M12	1.48 mS	82 μS	12.2 k Ω	NMOS
M13	1.22 mS	45 μS	22.3 k Ω	PMOS
M14	2.36 mS	50 μS	20.1 k Ω	PMOS
M15	3.51 mS	97 μS	10.3 k Ω	NMOS
M16	1.49 mS	81 μS	12.3 k Ω	NMOS
M17	1.60 mS	66 μS	15.1 k Ω	NMOS

The open circuit loop gain, T_{oc}, is 0, the short circuit loop gain is

$$T_{sc} = \frac{-g_{m9}R_{bias2}\left(r_{ds9} + g_{m10}r_{ds9}r_{ds10}\right)}{r_{ds9} + r_{ds10} + R_{bias2} + g_{m10}r_{ds9}r_{ds10}} \cdot \frac{g_{m7}}{g_{m7} + g_{ds7} + g_{ds6}} = -84.6, \tag{5-62}$$

and the closed-loop impedance of the super-regulated cascode current mirror is

$$R_{load1} = R_{load1,nf}\left(1 - T_{sc}\right) = 6.1\,M\Omega. \tag{5-63}$$

The impedance looking into the input cascode circuit (*M1,M2*) is

$$R_{out,M2} = r_{ds1} + r_{ds2} + g_{m2}r_{ds1}r_{ds2} = 3.48\,k\Omega. \tag{5-64}$$

The output impedance of the improved cascode circuit (looking into the drain of *M3*) is also calculated using Blackman's theorem. The open loop impedance is

$$R_{load2,nf} = r_{ds3} + r_{ds8} \| rds5 \| R_{out,m2} + g_{m3}\left(r_{ds3}\right)\left(r_{ds8} \| r_{ds5} \| R_{out,m2}\right) = 29.4\,k\Omega. \tag{5-65}$$

As before, the open circuit loop gain is 0 and the short circuit loop gain is

$$Tsc = \frac{-g_{m12}R_{bias1}\left(r_{ds12} + g_{m11}r_{ds11}r_{ds12}\right)}{r_{ds11} + r_{ds12} + R_{bias1} + g_{m11}r_{ds11}r_{ds12}} \cdot \frac{g_{m3}}{g_{m3} + g_{ds3} + g_{ds5} + g_{ds8} + 1/R_{out,M2}} = -279. \tag{5-66}$$

The output impedance of the improved cascode circuit is then

$$R_{load2} = R_{load2,nf}\left(1 - T_{sc}\right) = 8.23\,M\Omega. \tag{5-67}$$

The input impedance of the improved cascode circuit is also calculated using Blackman's theorem. The open loop impedance looking into the source of *M3* and the drain of *M8* is

$$R_{inM3,nf} = \frac{r_{ds3} + R_{load1}}{g_{m3}r_{ds3} + 1} \Bigg\| r_{ds8} = 11.6\,k\Omega. \tag{5-68}$$

In this case the short circuit loop gain is 0 and the open circuit loop gain is given by

$$T_{oc} = \frac{-g_{m12}R_{bias1}\left(r_{ds12} + g_{m11}r_{ds11}r_{ds12}\right)}{r_{ds11} + r_{ds12} + R_{bias1} + g_{m11}r_{ds11}r_{ds12}} \cdot \frac{g_{m3}}{g_{m3} + g_{ds3} + g_{ds8} + R_{load1}\left(g_{ds3}g_{ds8}\right)} = -18.7. \tag{5-69}$$

The input impedance of the improved cascode circuit is given by

$$R_{inM3} = \frac{R_{inM3}}{1 - T_{oc}} = 590\Omega. \tag{5-70}$$

With knowledge of the current mirror impedances, the gain of each stage in the preamplifier may be calculated. The load impedance at the drain of *M1* is

$$R_{inM2} = \frac{r_{ds2} + R_{inM3} \| r_{ds5}}{g_{m2}r_{ds2} + 1} = 81.2\Omega. \tag{5-71}$$

The voltage gain of *M1* common source stage is

$$A_{vM1} = -g_{m1}\left(r_{ds1} \| R_{inM2}\right) = -2.75. \tag{5-72}$$

Thus, the gate-drain capacitance of *M1* will be Miller-multiplied by a factor of 3.75. The voltage gain of the *M1*, *M2* cascode stage is

$$AvM1M2 = \frac{-g_{m1}\left(r_{ds1} + g_{m2} r_{ds1} r_{ds2}\right)\left(r_{ds5} \| R_{inM3}\right)}{r_{ds1} + r_{ds2} + r_{ds5} \| R_{inM3} + g_{m2} r_{ds1} r_{ds2}} = -13.4. \tag{5-73}$$

This is the second highest gain node in the circuit and is the source of a high-frequency pole in the loop gain.

The overall low frequency loop gain of the local feedback loop is given by

$$T = \frac{-g_{m12} R_{bias1}\left(r_{ds12} + g_{m11} r_{ds11} r_{ds12}\right)}{r_{ds11} + r_{ds12} + R_{bias1} + g_{m11} r_{ds11} r_{ds12}} \cdot \frac{g_{m3}}{g_{m3} + g_{ds3} + g_{ds8} + g_{ds5} + \frac{1}{R_{outM2}} + R_{load1}\left(g_{ds3}\left(g_{ds8} + g_{ds5} + \frac{1}{R_{outM2}}\right)\right)}, \tag{5-74}$$

or

$$T = -1.34. \tag{5-75}$$

Using standard feedback theory, the gain to the dominant node of the preamplifier is

$$A_{vdn} = A_{vdnideal} \frac{-T}{1-T}, \tag{5-76}$$

where

$$A_{vdnideal} = -g_{m1} R_{load1} = -244 \times 10^3. \tag{5-77}$$

Substituting Equations 5-75 and 5-77 into Equation 5-76 yields

$$A_{vdn} = -244 \times 10^3 \frac{1.34}{1+1.34} = -140 \times 10^3. \tag{5-78}$$

The gain of the *M5* source follower (still neglecting the effects of bulk-source modulation) is given by

$$A_{vM4} = \frac{g_{m4}}{g_{m4} + g_{ds4} + g_{ds17}} = 0.96. \tag{5-79}$$

The no-load open loop gain of the preamplifier may now be written as

$$A_{vol} = A_{vdn} A_{vM5} = -134 \times 10^3, \tag{5-80}$$

which is in very close agreement with the PSPICE calculated open loop gain of -139×10^3.

Several observations may be made about the preamplifier topology from this analysis. First, although the local feedback greatly reduces the input impedance of *M3* ($R_{in,M3}$ = 590 Ω), the input impedance of the improved cascode state is still nearly twice the drain-source impedance of *M1* (r_{dsm1} = 272 Ω). For this reason, an extra cascode device, *M2*, is added to the circuit to reduce the load at the drain of *M1* to 81.2 Ω. Second, the local feedback has the desired effect of generating a very high open loop gain (A_{vol} > 100 dB). While this improvement is important for maintaining linearity and gain stability, the frequency response is determined by the gain-bandwidth product of the circuit. The gain-bandwidth product is set by the second dominant pole appearing at the drain of *M2*. Circuit stability is achieved by adding capacitance to the dominant node in order to position the second dominant pole near the 0 dB crossover of the loop transfer function.

Stability

The next step is to analyze the frequency response of the preamplifier. The dominant node capacitance is set to ~6.5 pF to achieve a dominant pole corner frequency of

$$f_{-3db,p1} = \frac{1}{2\pi}\frac{1}{\left(R_{load1} \| R_{load2}\right)C_{dom}} = 7\,kHz\,. \tag{5-81}$$

This hand calculated value is in close agreement with the 6.4 kHz corner frequency shown in the post-layout PSPICE calculated open loop bode plot (Figure 5-14). Assuming the preamplifier is well modeled by a single pole transfer function, the gain-bandwidth of the preamplifier is

$$GBW = A_{vol} f_{-3db,p1} = \left(134\times10^{3}\right)\left(6.4\times10^{3}\right) = 858\,MHz\,. \tag{5-82}$$

The non-dominant pole occurs at the drain of *M2*. The total impedance at this node is

$$R_{p2} = R_{oM2} \| r_{ds5} \| R_{inM3} = 355\Omega, \tag{5-83}$$

and the total capacitance is

$$C_{p2} \approx C_{d2} + C_{d5} + C_{d8} + C_{d3} + C_{gs12} = 2.5pF. \tag{5-84}$$

The corner frequency of the second open loop pole is therefore approximated by

$$f_{-3db,p2} \approx \frac{1}{2\pi}\frac{1}{R_{p1}C_{p1}} = 180MHz. \tag{5-85}$$

The second pole near 200 MHz is clearly evident in the post-layout PSPICE plot of Figure 5-14.

The loop transfer function is given by

$$T(\omega) = A_{vol}(\omega)\beta(\omega), \tag{5-86}$$

where $A_{vol}(\omega)$ is the preamplifier open loop gain with poles at 6.4 kHz and 180 MHz, and $\beta(\omega)$ is the feedback gain given by

$$\beta(\omega) = \frac{1 + j\omega R_F C_F}{1 + j\omega R_F\left(C_F + C_D + C_{in}\right)}. \tag{5-87}$$

PSPICE simulations show the preamplifier input capacitance, C_{in} to be approximately 10 pF, and R_F and C_F are selected to be 500 MΩ and 1 pF, respectively. The smallest anticipated detector and stray capacitance at the input is 5 pF, so in the worst case (for stability) the feedback network has a pole at 23 Hz , a zero at 318 Hz, and a gain of 1/16 above 318 Hz. The loop transmission bode plot is shown in Figure 5-15, where the load capacitance was set to 10 pF. The 0 dB magnitude crossover frequency is 47 MHz, and the phase margin is 47°. This relatively conservative phase margin was selected to insure the fabricated amplifier would be stable over the full range of variations in the ORBIT fabrication process.

Rise Time

The expected closed-loop preamplifier corner frequency is approximately given by

$$f_{-3db,closed\,loop} = GBW \cdot T_{mid} = 858MHz \cdot \frac{1}{16} = 54MHz. \qquad (5\text{-}88)$$

Assuming the preamplifier is well modeled by a single pole transfer function, the expected rise-time is

$$t_{10-90} = 2.2\frac{1}{2\pi\ f_{-3db,closed\,loop}} = 6.5ns. \qquad (5\text{-}89)$$

For the PSPICE simulations, a BGO pulse of 1500 photons, an APD quantum efficiency of 0.85 and an APD gain of 500 were assumed. For this configuration, the input charge is approximately 100 fC, inducing a 100 mV voltage step at the preamplifier output. Figure 5-16 shows the positive and negative going preamplifier responses to a 100 fC impulse; the simulated rise time is approximately 5.8 ns for both polarity signals. The simulated rise-time is somewhat shorter than the value predicted by Equation 5-89 due to peaking in the simulated transient response. The rise-time may be improved by increasing the feedback capacitance to reduce the amplifier phase margin. For example, if the feedback capacitor is set to 2 pF, the simulated rise-time is ~3.5 ns with a 30% overshoot.

Slew Rate

The positive going slew-rate is determined by the drain current of *M6* (1.48 mA) and the dominant node capacitance (6.5 pF), or

$$SR_{pos} = \frac{I_{DM6}}{C_{dom}} = 228\frac{V}{\mu s}. \qquad (5\text{-}90)$$

The negative going slew rate is determined by the drain current of *M17* (1 mA) and the load capacitance (10 pF), or

$$SR_{neg} = \frac{ID_{M17}}{C_{load}} = 91\frac{V}{\mu s}. \tag{5-91}$$

The PSPICE calculated large signal transient responses to a 2 pC input charge are shown in Figure 5-17. The simulated slew-rates agree with the calculations, and these plots demonstrate the preamplifier has a dynamic range of approximately ±2V. The long settling time in the negative going signal is due to clipping in the over-shoot portion of the output signal.

Noise Analysis

A noise analysis of the preamplifier design was performed using BSIM and Level 2 PSPICE models [36]. The BSIM model generally provides a more accurate description of MOSFET circuit performance [10], but the PSPICE implementation of the BSIM flicker noise model has been shown to yield physically unrealistic results using published values of K_F[39]. For this study the BSIM model was used to calculate the output noise voltage due to all noise sources except the MOSFET flicker noise and the Level 2 model was used to model the output noise due to the MOSFET flicker noise.

Figure 5-18 shows the PSPICE calculated output noise voltage spectra for C_d = 5 pF, C_{load} = 10 pF, C_F = 1 pF and R_F = 500 MΩ. The BSIM model output is shown in Curve A and the Level 2 model output is shown in Curve B. The BSIM and Level 2 models were extracted from devices from the same ORBIT by the MOSIS prototyping service [36]. The flicker noise parameters ($K_{F,NMOS} = 3.5x10^{-28}\ C/m^2$, $K_{F,PMOS} = 1.0x10^{-29}\ C/m^2$) were extracted from devices fabricated in the ORBIT 2 μm PWELL process [40]. The appropriate K_F values for the ORBIT NMOS process may be slightly different, but no published values were found in the literature.

The BSIM model shows that 82% of the preamplifier output noise power is due to white thermal channel noise is generated by *M1* and *M5*. Using the PSPICE (BSIM) operating point data, the input white noise voltage due to *M1* is given by

$$env_{M1} = \sqrt{4kT \cdot \frac{2}{3g_{m1}}} = 0.50 \frac{nV}{\sqrt{Hz}}, \tag{5-92}$$

and the input white noise voltage due to *M5* is given by

$$env_{M5} = \sqrt{4kT \cdot \frac{2}{3} \frac{g_{m5}}{g_{m1}^2}} = 0.30 \frac{nV}{\sqrt{Hz}}, \tag{5-93}$$

yielding a total input white noise voltage due to *M1* and *M5* of

$$env_{M1,M5} = \sqrt{env_{M1}^2 + env_{M5}^2} = 0.58 \frac{nV}{\sqrt{Hz}}. \tag{5-94}$$

The BSIM PSPICE simulation calculates a total white noise voltage for the preamplifier of 0.64 nV/√Hz. Below the amplifier cutoff frequency, the white input noise voltage is translated to the output by the transfer function

$$\frac{V_{no}}{\Delta f} \approx ENV \left(\frac{C_F + C_{in} + C_d}{C_F} \right) = (0.64)(16) = 10.2 \frac{nV}{\sqrt{Hz}}. \tag{5-95}$$

The < 11 nV/√Hz output noise voltage floor is clearly seen in the BSIM calculated output noise spectrum (Figure 5-18, Curve A) above 100 kHz.

The primary source of equivalent noise current for the preamplifier is the 500 MΩ feedback resistor which has injects a noise current of

$$ENI = \sqrt{\frac{4kT}{R_F}} = 5.76 \frac{fA}{\sqrt{Hz}} \tag{5-96}$$

into the preamplifier input terminal. The input noise current is translated to the preamplifier output by the transfer function

$$\frac{V_{no}}{\Delta f} = \frac{ENI}{2\pi C_F f} = \frac{917 \times 10^{-6}}{f} \frac{V}{\sqrt{Hz}}. \tag{5-97}$$

The *1/f* (not flicker) component in the output spectrum is also clearly evident in the BSIM calculated output noise spectrum (Figure 5-18, Curve A).

The Level 2 PSPICE simulation output shows the flicker noise at the preamplifier output to be predominately due to the simple NMOS current mirror, *M5*. From Equation 5-22, *M5* has a channel flicker noise current of

$$\frac{{i_n}^2}{\Delta f}=\frac{\left(3.5\times10^{-28}\right)\left(10\times10^{-3}\right)}{\left(8.7\times10^{-4}\right)\left(4\times10^{-12}\right)f}=\frac{1\times10^{-15}}{f}\frac{A^2}{Hz}, \tag{5-98}$$

where C_{ox}=8.7x10^{-4} F/m^2 (t_{ox} = 397Å) and K_F=3.5x10^{-28} C/m^2. The *M5* channel current noise is referred to the gate of *M1* by the relationship

$$ENV^2{}_{M5,flicker}=\frac{{i_n}^2/\Delta f}{\left({g_{m1}}^2\right)}=\frac{519\times10^{-15}}{f}\frac{V^2}{Hz} \tag{5-99}$$

The equivalent input flicker noise voltage due to *M5* is translated to the preamplifier output by Equation 5-97 to obtain

$$\frac{V_{no,M5,flicker}}{\Delta f}=\frac{11.53\times10^{-6}}{\sqrt{f}}\frac{V}{\sqrt{Hz}}. \tag{5-100}$$

Comparison of Equation 5-100 with Figure 5-18, Curve B, shows that *M5* contributes virtually all of the flicker noise present in the preamplifier output noise voltage spectrum. The simulated flicker/white noise corner is approximately 1 MHz. A major focus of future design iterations will be the development of an improved current sink with reduced flicker noise.

Experimental Results

The preamplifier array was fabricated in the ORBIT Semiconductor foundry using ORBIT's 2 μm N-well process. The preamplifier arrays are packaged in 40-pin ceramic DIP packages and evaluated in a hand-wired prototype test circuit (Figure 5-19). The preamplifier bias currents are set by external resistors and the reference voltages V_{M1WELL} and V_{BIAS} are established with external potentiometers. The HFA1100 serves as a low input capacitance (C_{in}≈2 pF), wideband (f_{-3db}≈530 MHz) buffer amplifier with a gain of +2. The 2kΩ bias resistor at the non-inverting input of the HFA1100 is selected as a compromise between maximizing the buffer input impedance and minimizing the output offset voltage due to the input bias currents. The input noise voltage (4 nV/√Hz) of the HFA1100 is small with respect to the output noise voltage of the CMOS preamplifier.

Selection of Bias Currents and Voltages

The optimal settings for I_{BIAS1}, V_{BIAS} and V_{M1WELL} were determined using the test circuit shown in Figure 5-20. The bias settings were chosen to minimize the rms noise measured at the output of the ORTEC 570 where the front panel switch was set for a shaping time constant of 2 μs. (As discussed in Chapter 6, the optimal shaping time constant for the preamplifier with no detector noise is 2 μs.) The characteristics of the ORTEC 570 are discussed in detail later in this chapter. Figure 5-21 shows the measured noise charge as a function of I_{BIAS1} where V_{BIAS} = -3.5V and $V_{M1WELL} = V_{DD}$ = 5V. The optimal value of I_{BIAS1}=2.6 mA corresponds to a drain current of ~26 mA in the input MOSFET (*M1*). For Id_{M1} < 26 mA, $g_{m,M1}$ is reduced, causing an increase in the preamplifier equivalent input noise voltage. For Id_{M1} > 26 mA, the output impedance of the *M1-M2* cascode stage is reduced, degrading the loop gain of the super-regulated current mirror and reducing the overall open loop gain of the preamplifier. The reduction in open loop gain causes the noise contributions of transistors other than *M1* and *M5* to become important.

Figure 5-22 shows the measured output noise charge as a function of V_{BIAS}, where I_{BIAS1} = 2.6mA and $V_{M1WELL} = V_{DD}$ = 5V. The measured output noise charge is minimal for V_{BIAS}=-3.0V. PSPICE analysis of the circuit shows that for V_{BIAS}=-3.0V and I_{BIAS1}=2.6 mA, $V_{DS,M1}$ is ~400 mV greater than $V_{DS,SAT}$. As the magnitude of V_{BIAS} decreases, *M1* goes out of saturation and the output noise charge increases markedly. As the magnitude of V_{BIAS} increases, $V_{DS,M1}$ increases, causing an increase in the noise of *M1* [20].

The measured output noise charge vs V_{M1WELL} is shown in Figure 5-23. The pulse height resolution is found to be only weakly dependent on the well voltage, with a shallow minimum appearing in the curve for V_{M1WELL}=2V. Based upon this data, the amplifier was biased for I_{BIAS1} = 2.6 mA, V_{BIAS} = -3.0V, and V_{M1WELL} = +2V.

Frequency Response

The ac behavior of the preamplifier was evaluated by measuring the transient response of the circuit to a current impulse and by measuring the output signal magnitude as a function of frequency using a spectrum analyzer. The test circuit used for the ac measurements is shown in Figure 5-24. For the transient response measurements, the circuit is driven with a 1 ns voltage step terminated with a 1 pF capacitor in series with the preamplifier input. For a 50 mV input voltage step, 50 fC of charge is injected into the summing node at the preamplifier input. The charge impulse is integrated by the preamplifier, yielding a 50 mV voltage step at the preamplifier output which is measured with an oscilloscope.

For the swept frequency measurements, the preamplifier is driven by the voltage source of an HP3589 spectrum analyzer. Again, the input voltage is terminated with a 1 pF capacitor so that the current injected into the input summing node of the preamplifier is given by

$$\left|I_{in}(f)\right| = V_{source}(f) \cdot 2\pi f C_s, \tag{5-101}$$

where C_s is the series capacitance at the preamplifier input. The transfer function of the preamplifier is given by

$$|H(f)| = \left|\frac{1}{2\pi f C_F} \cdot \frac{-T(f)}{1-T(f)}\right| \tag{5-102}$$

where C_F is the feedback capacitor and $T(f)$ is the frequency dependent loop gain of the preamplifier. The magnitude of the output voltage is given by

$$|V_O(f)| = |I_{in}(f)| \cdot |H(f)| = V_{source}(f)\left|\frac{-T(f)}{1-T(f)}\right|. \tag{5-103}$$

The frequency response of the preamplifier is established by the open-loop corner frequency (f_o) and open-loop gain (A_o) of the preamplifier. Assuming the preamplifier may be modeled as a single pole system, the loop gain of the preamplifier is given by

$$T(f) = \frac{A_O}{1 + j\,{}^{f}\!/_{f_o}} \cdot \left(\frac{C_F}{C_F + C_{in} + C_D}\right) \tag{5-104}$$

where C_{in} is the preamplifier input capacitance and C_D is the external "detector" capacitance. The closed loop gain of the preamplifier when driven through the 1pF series capacitor, C_s is

$$A_{cl} = -\frac{C_F}{C_s}\left(\frac{-T}{1-T}\right) \approx 1 \cdot \frac{1}{1 + j\dfrac{f}{A\left(\dfrac{C_F}{C_F + C_{in} + C_D}\right) f_o}}. \tag{5-105}$$

Thus, the closed loop -3dB frequency is approximately given by

$$f_{-3db,cl} = Af_o\left(\frac{C_F}{C_F + C_{in} + C_d}\right). \tag{5-106}$$

The measured transient and frequency response for the detector for C_D= 0, 5, 10, 15 and 22 pF are shown in Figures 5-25 through 5-34. The best rise-time and bandwidth are obtained for C_D= 0 pF ($t_{10\text{-}90}$ = 7 ns, $f_{-3db,cl}$ = 40 MHz). These values are somewhat lower than the PSPICE predicted values for C_{in} = 10 pF and C_d = 5 pF: $t_{10\text{-}90}$ = 5.8 ns, $f_{-3db,cl}$ = 85 MHz (with +2dB peaking at 50 MHz in the calculated bode plot).

The measured $f_{-3db,cl}$ values are summarized in Figure 5-35. In order to obtain an estimate of the gain-bandwidth product of the preamplifier (GBW=$A_o f_o$), the data is fit to Equation 5-106 where it is assumed

$$C_{in} = C_{preamp} + C_{package} + C_{parasitic} = 10pF + 4pF + 2pF = 16pF\,. \tag{5-107}$$

The value C_{preamp}=10pF is based on the post-layout PSPICE simulations, and the values for $C_{package}$ and $C_{parasitic}$ are estimated. The least-squares fitting algorithm estimates GBW = 605 MHz, which is 30% lower than the value predicted by Equation 5-88 (858 MHz). Because A_o and f_o appear as a product term in Equation 5-106, it is not possible to determine whether the poorer-than-expected GBW is due to a reduction in the open-loop gain or in the dominant pole frequency.

Preamplifier Noise

Measured Output Noise Spectral Density

The output noise voltage of the preamplifier was measured as a function of detector capacitance using the test circuit shown in Figure 5-36. The ORTEC 9305 preamplifier was included in the circuit to keep the output noise voltage above the noise floor of the spectrum analyzer. The measured output noise spectra are shown in Figure 5-37. The narrow peaks appearing in the spectra at 25, 50, 75 and 100 MHz are attributed to pick-up of an external electromagnetic signal. The peaks also appear in the measured spectra of a commercially available charge-sensitive preamplifier (ORTEC 142A). The magnitude of the peaks could be reduced with aggressive shielding of the test circuit but it could not be eliminated from the spectra.

At low frequencies, the noise is dominated by the 500 MΩ feedback resistor, yielding a low-frequency 1/f characteristic in the curves. The expected contribution of the feedback resistor to the output noise spectrum is drawn on the plot. Over most of the frequency range, the output noise is dominated by flicker noise, giving the spectrum a roughly $1/\sqrt{f}$ slope. The magnitude of the flicker noise is consistent with the Level 2 simulations shown in Figure 5-18, indicating the input NMOS current source (*M5*) is the dominant flicker noise source. The extrapolated white/flicker noise corners are approximately between 800 kHz and 1 MHz.

The white-noise plateau due to *ENV* is evident in each of the measured spectra. The lowest measured white output noise voltage (20 nV/√Hz) is obtained with 0 pF added to the input terminal. The measured white output noise voltage values are plotted as a function of C_d in Figure 5-38 and fit to Equation 5-98 to obtain estimates of *ENV* (1.07nV/√Hz) and C_{in} (16.6 pF). The fitted value for C_{in} is consistent with the estimated value given in Equation 5-107, while the fitted value for *ENV* is somewhat larger than the PSPICE calculated value (0.64 nV/√Hz) for the ideal case of $\gamma = 2/3$. An estimate of the effective noise factor for the preamplifier may be obtained by taking the ratio of the PSPICE value and the experimental value to obtain

$$\gamma_{eff} = \frac{1.08}{0.64} \cdot \frac{2}{3} = 1.13\,. \tag{5-108}$$

This estimated noise factor is consistent with the measured noise factors found in the literature [5,12-16].

Conclusions

A six-element wide-band, low-noise CMOS preamplifier array was designed and fabricated for use with APD/BGO PET detectors. As background for the design discussion, a summary of MOS noise characteristics was presented and the noise characteristics of transimpedance and charge sensitive preamplifiers were compared. A new cascode topology was presented which uses local feedback to improve the performance of the cascode element.

A detailed description of the preamplifier design and measured characteristics was presented. The preamplifier has a measured closed loop bandwidth of 40 MHz and a 10-90% rise time of 7 ns. Analysis of experimental data indicates the preamplifier has a gain-bandwidth product of approximately 600 MHz and an equivalent input noise voltage of approximately 1 nV/√Hz.

In Chapter 6 the amplifier performance with an APD detector is characterized and the status of the APD array development as of this writing is discussed.

References

1. J. E. Lilienfeld, United States Patent 1,745,175, 1930.

2. O. Heil, British Patent 439,457, 1935.

3. D. Kahng and M. M. Atalla, "Silicon-Silicon Dioxide Field Induced Surface Devices," IRE Solid State Devices Research Conference, Pittsburgh, Pa., 1960.

4. Robert M. Fox, "Comments on Circuit Models for MOSFET Thermal Noise," *IEEE J. Solid-State Circuits,* **SC-28**, pp. 184-185, 1993.

5. Bing Wang, James R. Hellums, and Charles G. Sodini, "MOSFET Thermal Noise Modeling for Analog Integrated Circuits," *IEEE J. Solid-State Circuits,* **SC-29**, pp. 833-835, 1994.

6. Germano Nicollini, Davide Pancini, and Sergio Pernici, "Simulation-Oriented Noise Model for MOS Devices," *IEEE J. Solid-State Circuits,* **SC-22**, pp. 1209-1212, 1987.

7. A. G. Jordan and N. A. Jordan, "Theory of Noise in Metal Oxide Semiconductor Devices," *IEEE Trans. Electron Devices,* **ED-12**, pp. 148-156, 1965.

8. C. T. Sah, S. Y. Wu and F. H. Hielscher, "The Effects of Fixed Bulk Charge on the Thermal Noise in Metal-Oxide-Semiconductor Transistors," *IEEE Trans. Electron Devices,* **ED-13**, pp. 410-414, 1966.

9. Y. P. Tsividis, *Operation and Modeling of the MOS Transistor.* New York: McGraw-Hill, 1987.

10. John P. Uymura, *Circuit Design for CMOS VLSI.* Kluwer Academic Publishers, 1992.

11. Phillip E. Allen and Douglas R. Holberg, *CMOS Analog Circuit Design.* New York: Hold, Rinehart and Winston, Inc., 1987.

12. Zhong Yuan Chang and Willy M. C. Sansen, *Low-Noise Wide-Band Amplifiers in Bipolar and CMOS Technologies*, Kluwer Academic Publishers, 1991.

13. R. P. Jindal, "Distributed Substrate Resistance Noise in Fine-Line NMOS Field-Effect Transistors," *IEEE Trans. Electron Devices,* **ED-32**, pp. 2450-2453, 1985.

14. S. Tedja, H. H. Williams, J. Van der Spiegel, F. M. Newcomer and R. Van Berg, "Noise Spectral Density Measurements of a Radiation Hardened CMOS Process in the Weak and Moderation Inversion," *IEEE Trans. Nuclear Science,* **NS-39**, pp 804-808, 1992.

15. Suharli Tedja, Jan Van der Spiegel, and Hugh H. Williams, "Analytical and Experimental Studies of Thermal Noise in MOSFETs," *IEEE Trans. Electron Devices,* **ED-41**, pp. 2069-2075, 1994.

16. David M. Binkley, James M. Rochelle, Michael J. Paulus and Michael E. Casey, "A Low-Noise, Wideband, Integrated CMOS Transimpedance Preamplifier for Photodiode Applications," *IEEE Transactions on Nuclear Science,*. **NS-39**, pp. 747-752, 1992.

17. A. van der Ziel, *Noise in Solid State Devices and Circuits.* New York: John Wiley, 1986.

18. Phillip E. Allen, Georgia Institute of Technology EE6311 Course Notes, 1991.

19. F. M. Klassen and J. Prins "Thermal Noise of MOS Transistors," *Philips Res. Repts* **22**, pp 505-514, 1967.

20. R. P. Jindal, "Hot-Electron Effects on Channel Thermal Noise in Fine-Line NMOS Field-Effect Transistors," *IEEE Trans. Electron Devices,* **ED-33**, pp. 1395-1397, 1986.

21. Paul Heremans, Rudi Bellens, Guido Groeseneken and Herman E. Maes, "Consistent Model for the Hot-Carrier Degradation in n-Channel and p-Channel MOSFETs," *IEEE Trans. Electron Devices,* **ED-35**, pp. 2194-2209, 1988.

22. *Foresight User's Manual*, ORBIT Semiconductor, Inc., Sunnyvale, California, 1991.

23. F. N. Hooge, "1/f noise," *Physica* **83B**, pp 14-23, 1976.

24. A. L. McWhorter, *Semiconductor Surface Physics.* Philadelphia, Univ. Pennsylvania Press, 1956.

25. Willy Sansen, "Integrated Low-Noise Amplifiers in CMOS Technology," *Nucl. Inst. and Meth. in Physics Research,* **A253**, pp 427-433, 1987.

26. Jimmin Chang, A. A. Abidi and C. R. Viswanathan, "Flicker Noise in CMOS Transistors from Subthreshold to Strong Inversion at Various Temperatures," *IEEE Trans. Electron Devices,* **ED-41**, pp. 1965-1971, 1994.

27. Hiroaki Mikoshiba "1/f Noise in n-Channel Silicon-Gate MOS Transistors," *IEEE Trans. Electron Devices,* **ED-29**, pp. 965-970, 1982.

28. L. W. Nagel, "SPICE2, A Computer Program to Simulate Semiconductor Circuits," Electron. Res. Lab., Univ. of Calif., Berkeley, Memo ERL-M520, May 1975.

29. *PSPICE User's Guide,* Microsim Corporation, Irvine California, 1989.

30. *HSPICE User's Manual,* Meta-Software, Inc., Campbell, California, 1992.

31. P. W. Nicholson, *Nuclear Electronics.* New York: John Wiley, 1974.

32. "The Avalanche Photodiode Catalog," Advanced Photonix, Inc., Camarillo, California.

33. "Noise Characteristics of Advanced Photonix Avalanche Photodiodes," Advanced Photonix Application Note API/NOIS/1291/B, 1991.

34. W. W. Moses, I. Kipnis, and M. H. Ho, "A 16-Channel Charge Sensitive Amplifier IC For a PIN Photodiode Array Based PET Detector Module," *IEEE Transactions on Nuclear Science,*. **NS-41**, pp. 1469-1472, 1994.

35. J. K. Millard, T. V. Blalock, and N. W. Hill, "A Broadband Current Preamplifier Configuration for Obtaining High Resolution Energy and Time Information from Nuclear Radiation Detectors," *IEEE Transactions on Nuclear Science,*. **NS-19**, pp. 388-395, 1972.

36. MOSIS report on N07S ORBIT fabrication run, September 18, 1990.

37. James M. Rochelle, David M. Binkley, and Michael J. Paulus, "Fully Integrated Current-Mode CMOS Gated Baseline Restorer Circuits," *Conference Record of the 1994 Nuclear Science Symposium Conference Record,* Norfolk, Virginia, November 1994.

38. J. F. Pierce and T. J. Paulus, *Applied Electronics.* New York: Bell and Howell, 1972.

39. E.J. Kennedy, private communication.

40. Christopher M. Stillo, Robert M. Fox, D. Scott Langford and Daniel J. Ferris, "Electrical Characterization of MOSIS-Fabricated Circuit Elements," *Conference Record of the 1991 Custom Integrated Circuits Conference,* San diego, California, May 1991.

Figures

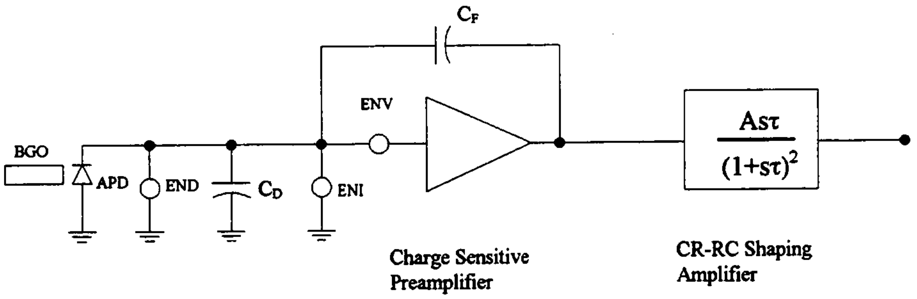

Figure 5-1. Schematic diagram of charge sensitive preamplifier and shaping amplifier.

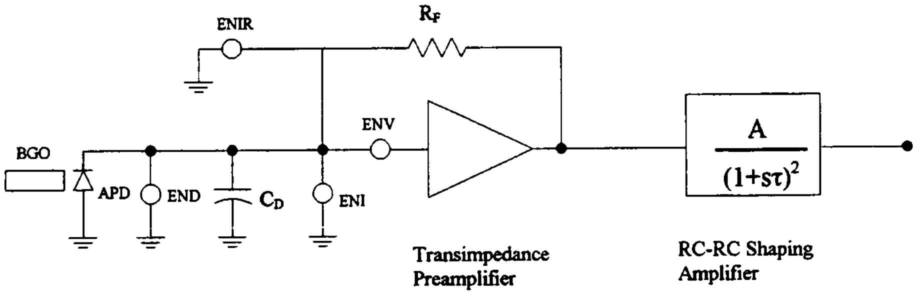

Figure 5-2. Schematic diagram of transimpedance preamplifier and shaping amplifier.

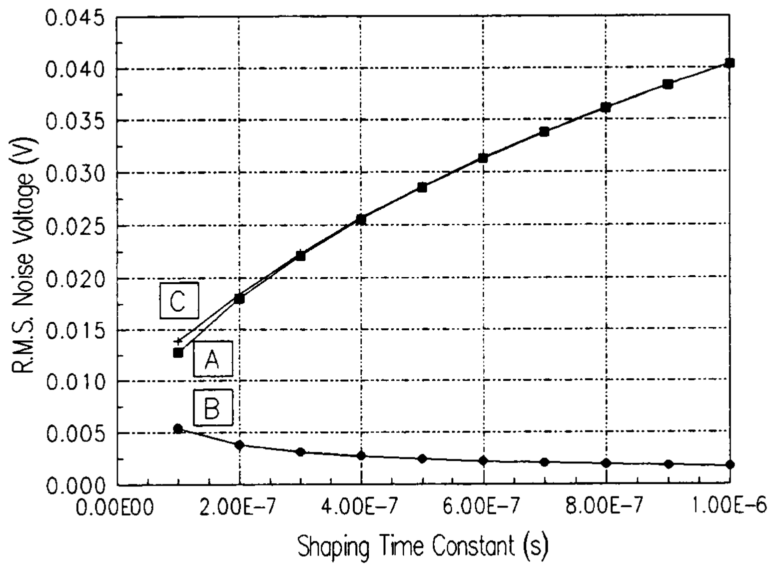

Figure 5-3. Contributions of (A) parallel noise current sources, (B) series noise voltage, and (C) all noise sources to the rms output noise voltage as a function of shaping time constant for the charge sensitive preamplifier in PHR mode.

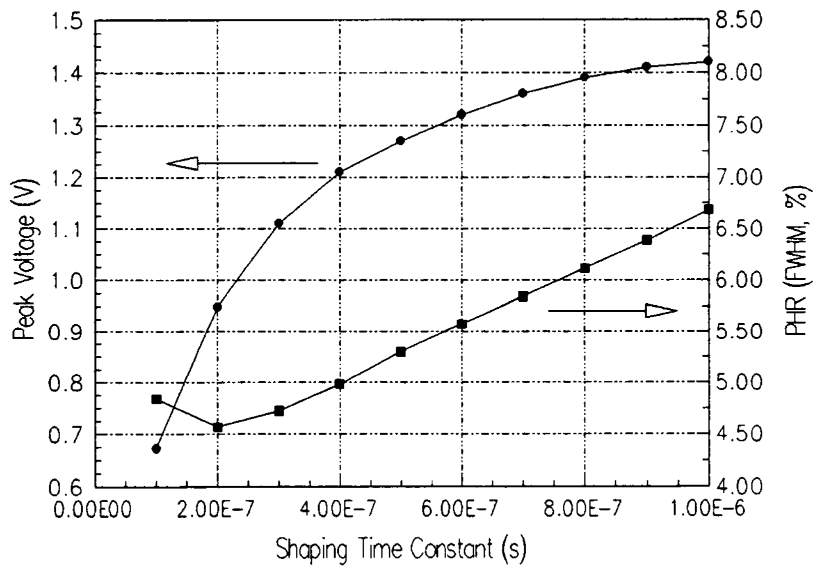

Figure 5-4. Calculated peak voltage and pulse height resolution due to electronic noise for the charge sensitive preamplifier.

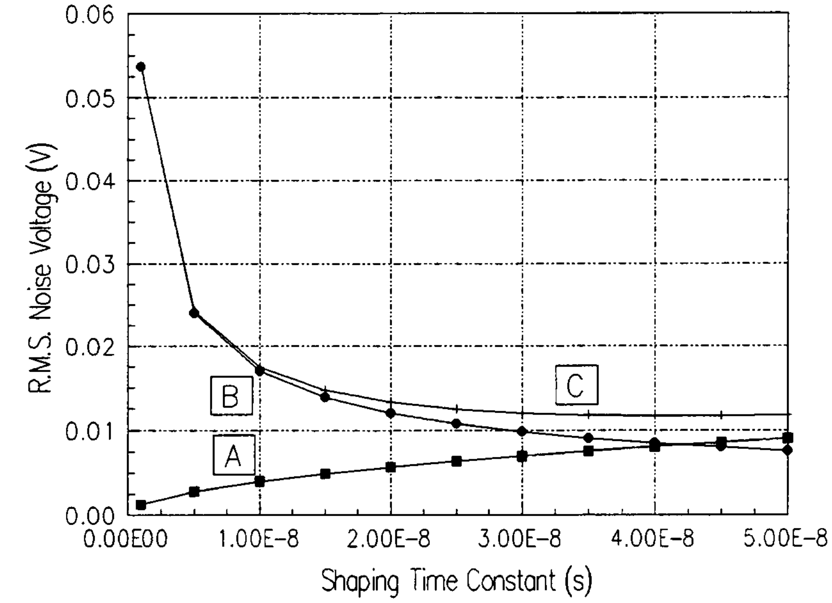

Figure 5-5. Contributions of (A) parallel noise current sources, (B) series noise voltage, and (C) all noise sources to the rms output noise voltage as a function of shaping time constant for the charge sensitive preamplifier in timing mode.

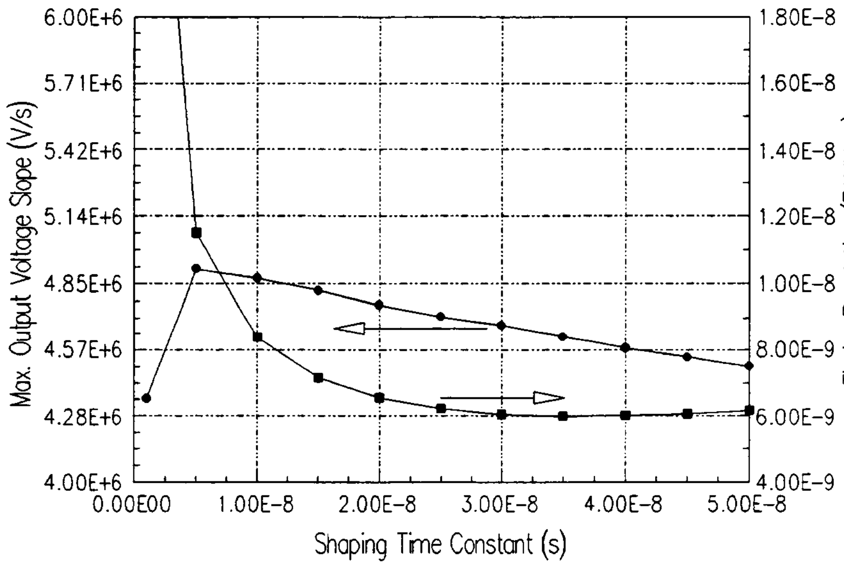

Figure 5-6. Calculated maximum signal slope and timing resolution due to electronic noise for the charge sensitive preamplifier.

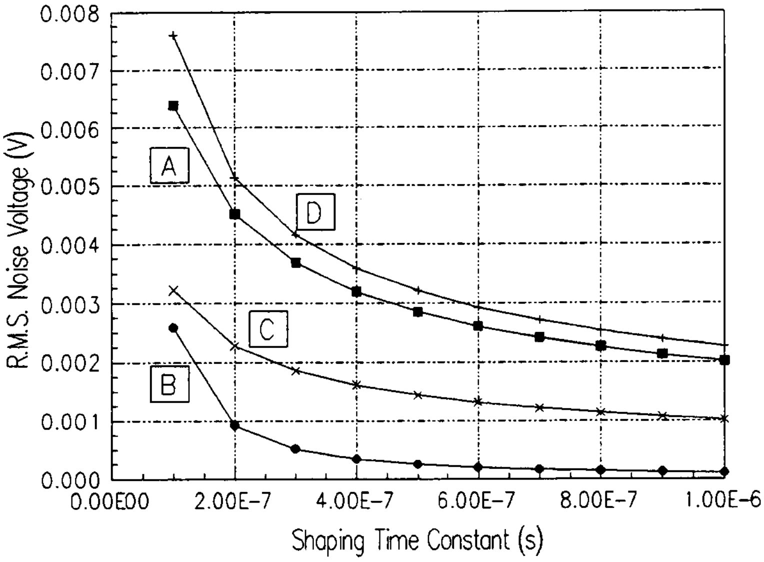

Figure 5-7. Contributions of (A) parallel noise current sources, (B) series noise voltage, (C) the feedback resistor and (D) all noise sources to the rms output noise voltage as a function of shaping time constant for the transimpedance preamplifier in PHR mode.

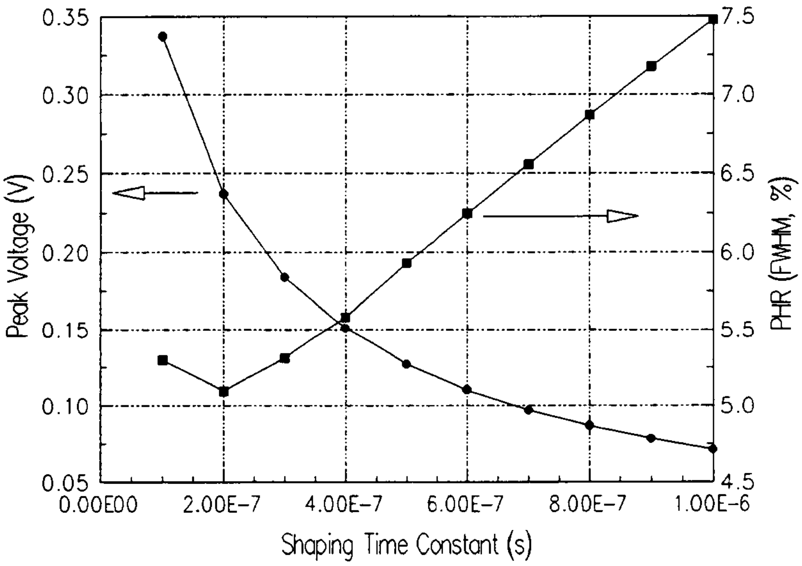

Figure 5-8. Calculated peak voltage and pulse height resolution due to electronic noise for the transimpedance preamplifier.

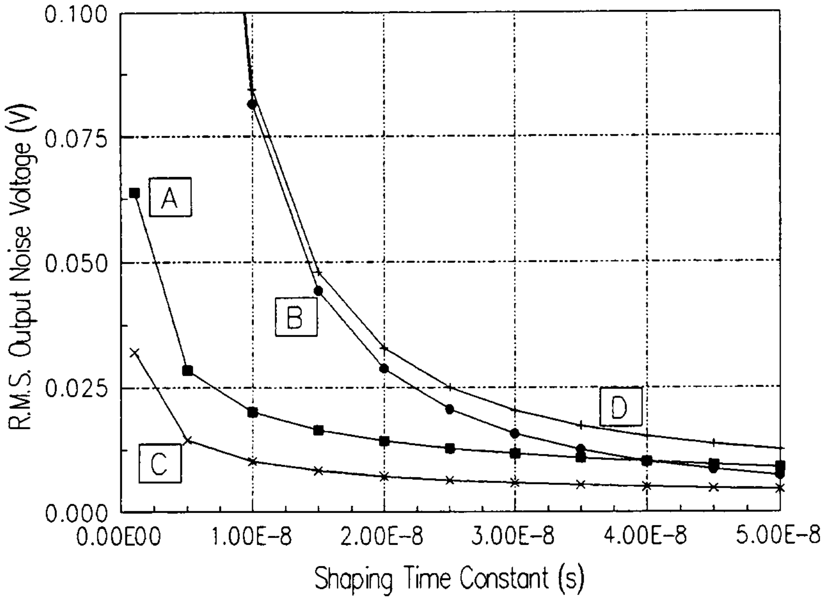

Figure 5-9. Contributions of (A) parallel noise current sources, (B) series noise voltage, (C) the feedback resistor and (D) all noise sources to the rms output noise voltage as a function of shaping time constant for the transimpedance preamplifier in timing mode.

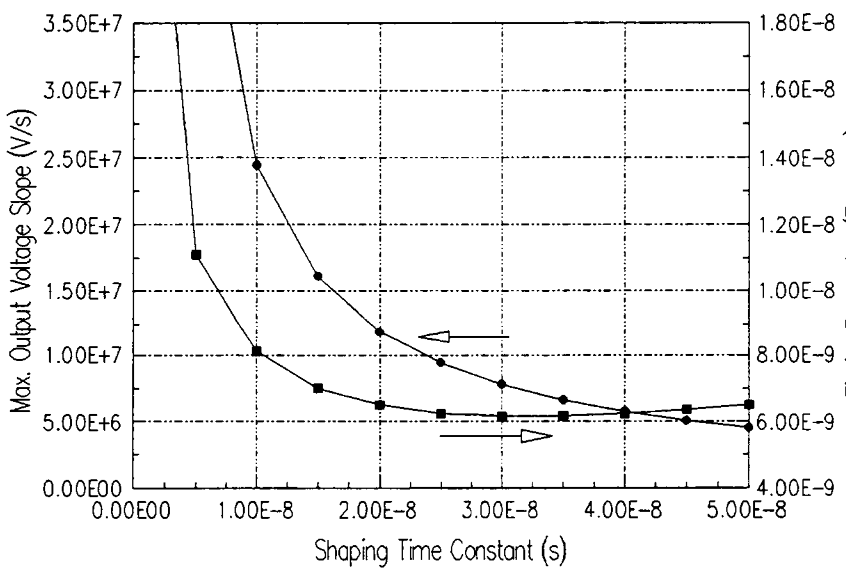

Figure 5-10. Calculated maximum signal slope and timing resolution due to electronic noise for the transimpedance preamplifier.

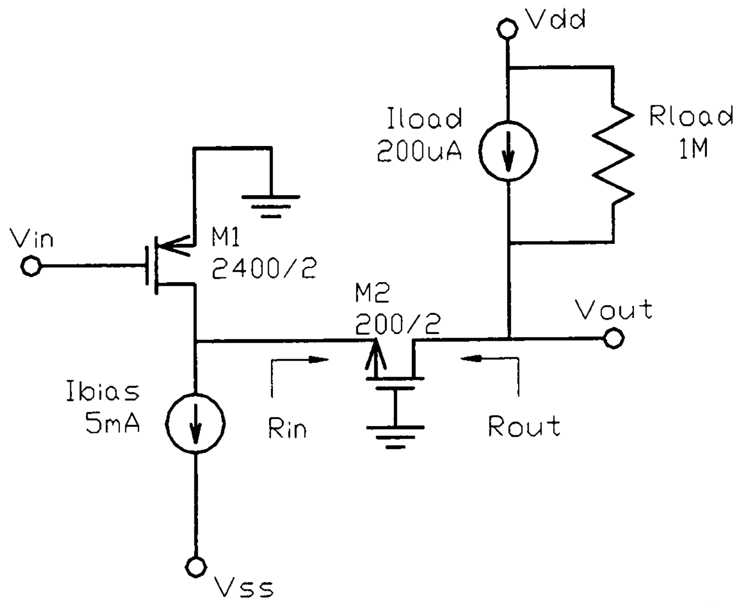

Figure 5-11. AC equivalent circuit of the folded cascode topology in a low-noise CMOS preamplifier presented by Binkley *et al* [16].

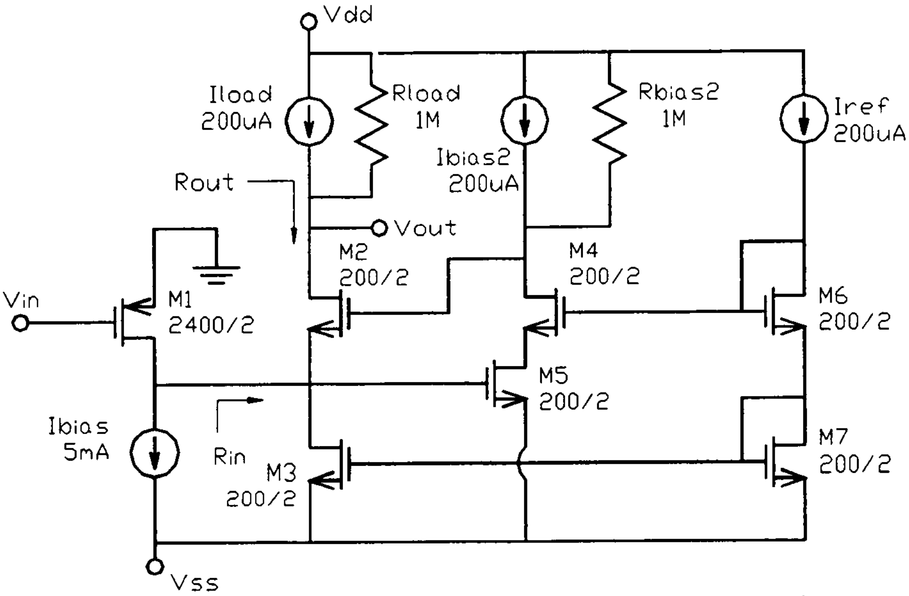

Figure 5-12. Improved cascode topology employing local feedback. Transistor geometries and bias currents match those in [16].

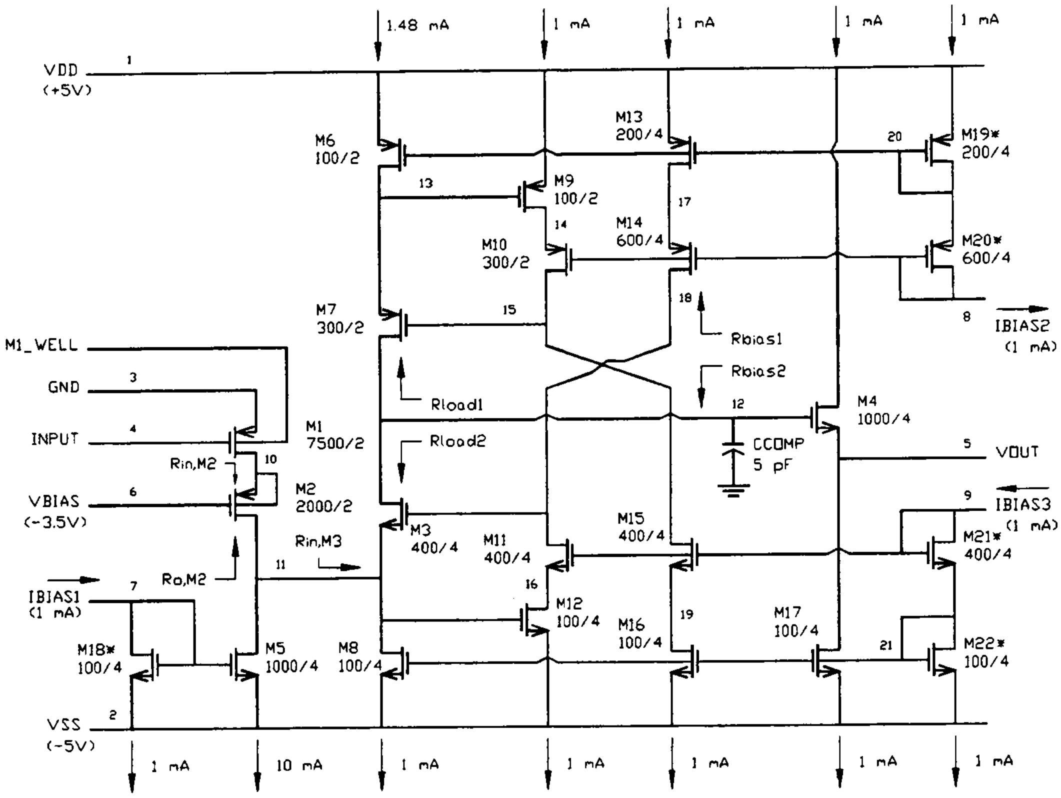

Figure 5-13. Schematic diagram of charge sensitive preamplifier design developed for use with an APD array.

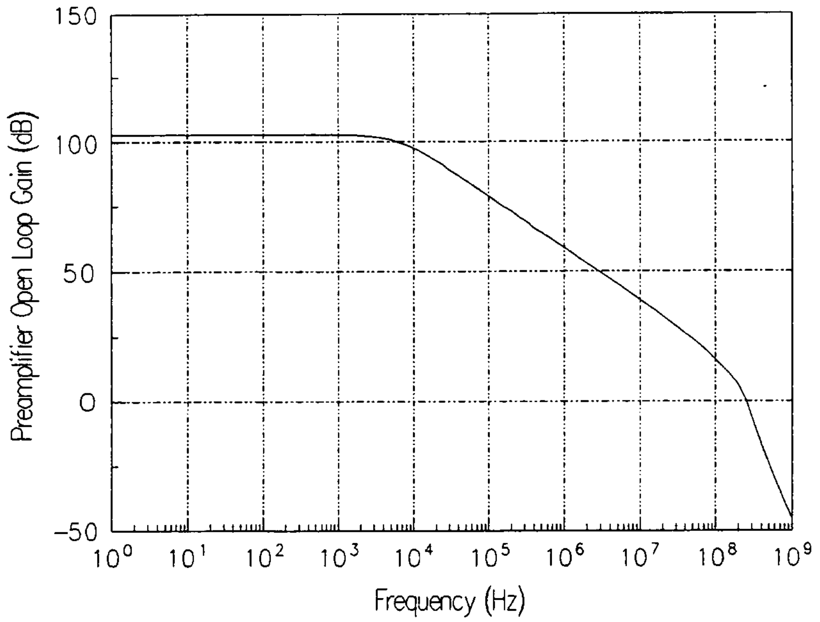

Figure 5-14. Magnitude plot of the PSPICE calculated preamplifier open loop gain.

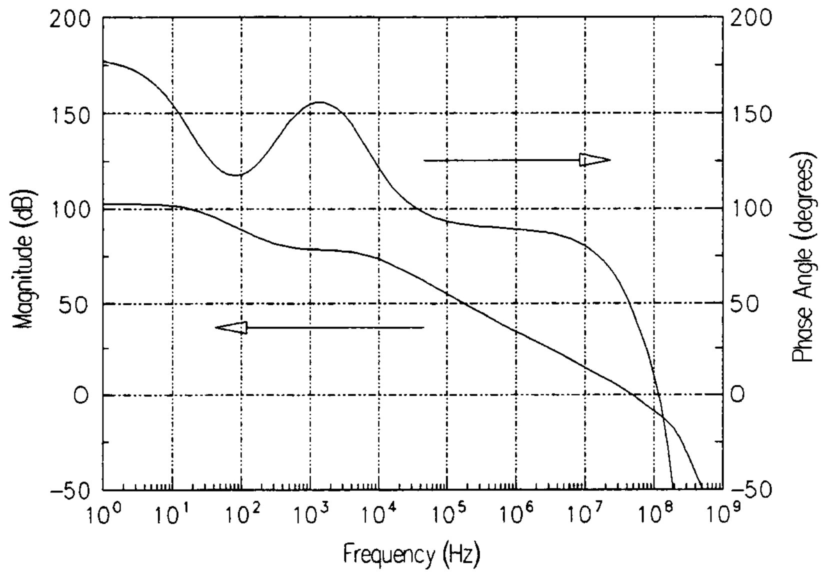

Figure 5-15. Bode plot of preamplifier and feedback network loop gain. The 0 dB crossover occurs at 47 MHz, where the phase margin is 47°.

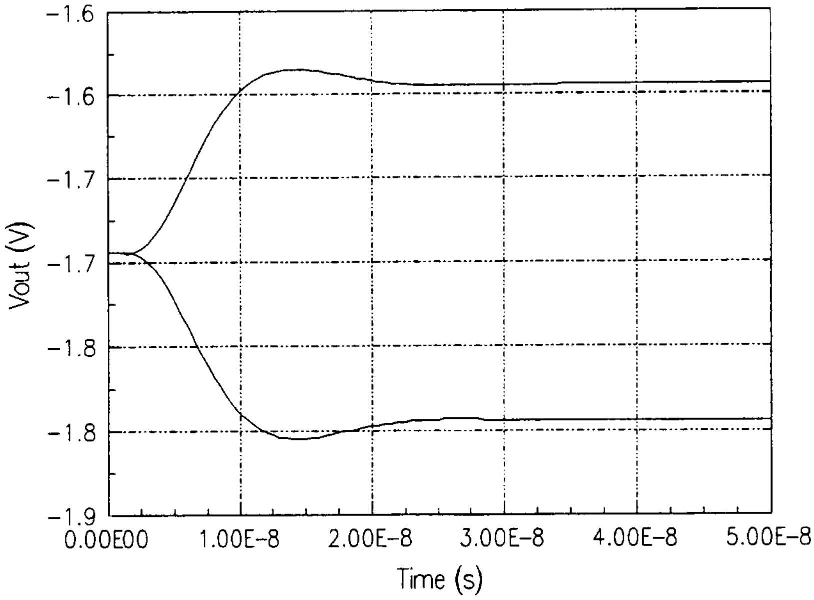

Figure 5-16. PSPICE calculated transient response to 100 fC (small signal) charge impulse.

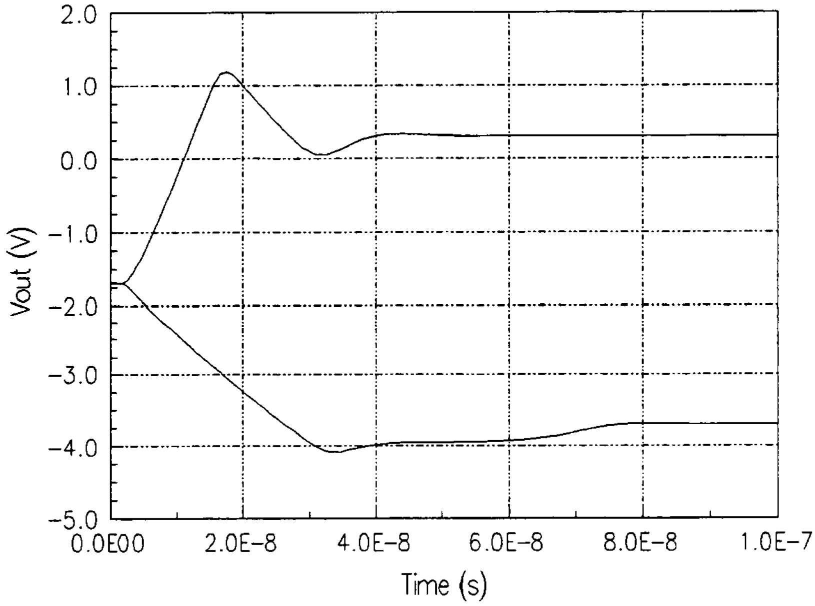

Figure 5-17. PSPICE calculated transient response to 2 pC (large signal) charge impulse. In both curves the rise-time is slew-rate limited.

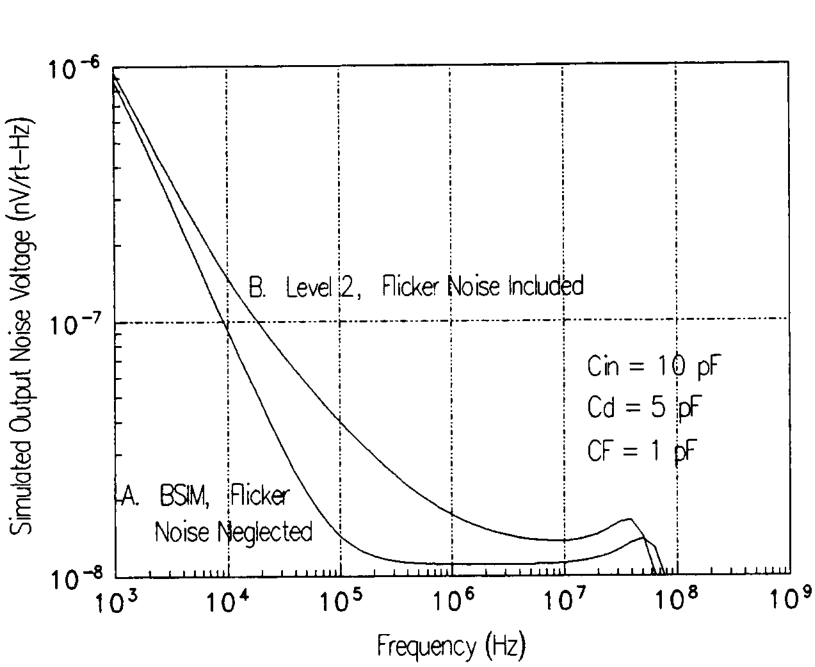

Figure 5-18. PSPICE calculated output noise voltage spectral density plots. Plot A was obtained using BSIM model parameters without flicker noise coefficients. Plot B was obtained using Level 2 model parameters with $K_F(PMOS)=1.0x10^{-29}$ C/m^2 and $K_F(NMOS)=3.5x10^{-28}$ C/m^2.

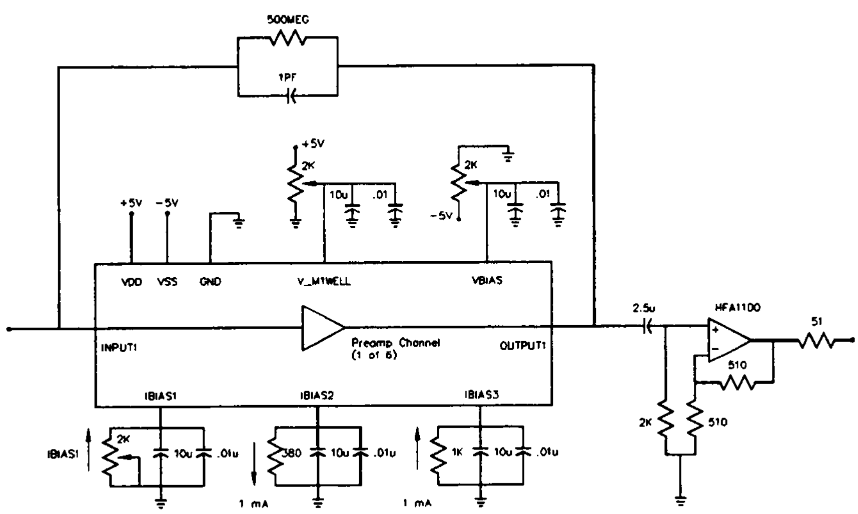

Figure 5-19. Test circuit for evaluation of the CMOS preamplifier array.

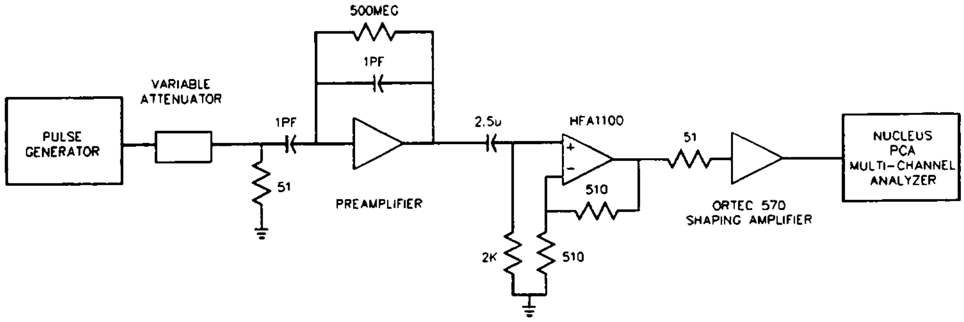

Figure 5-20. Test circuit for measuring the output noise charge of the preamplifier using a 10,000 electron input test pulse.

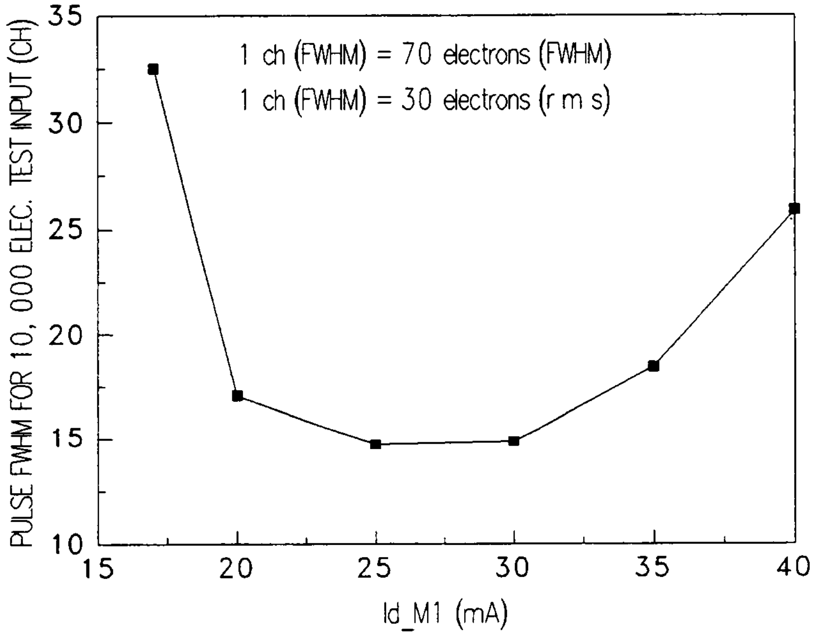

Figure 5-21. Measured output noise charge as a function of Id_{M1} obtained using a 10,000 electron test pulse and an ORTEC 570 shaping amplifier with a 2 μs shaping time constant. V_{BIAS} = -3.5V and V_{M1WELL} = +5V.

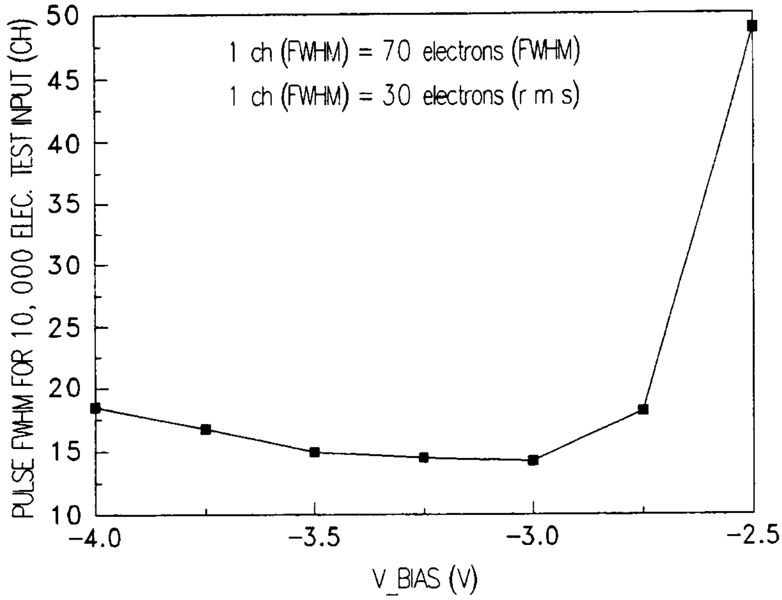

Figure 5-22. Measured output noise charge as a function of V_{BIAS} obtained using a 10,000 electron test pulse and an ORTEC 570 shaping amplifier with a 2 μs shaping time constant. Id_{M1}=25 mA and V_{M1WELL} = +5V.

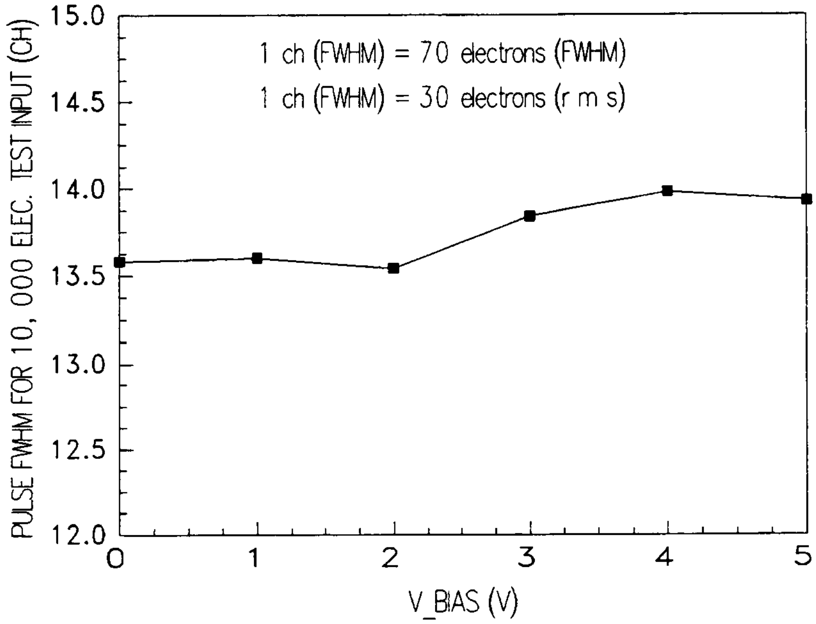

Figure 5-23. Measured output noise charge as a function of V_{M1WELL} obtained using a 10,000 electron test pulse and an ORTEC 570 shaping amplifier with a 2 μs shaping time constant. Id_{M1}=25 mA and V_{BIAS} = -3V.

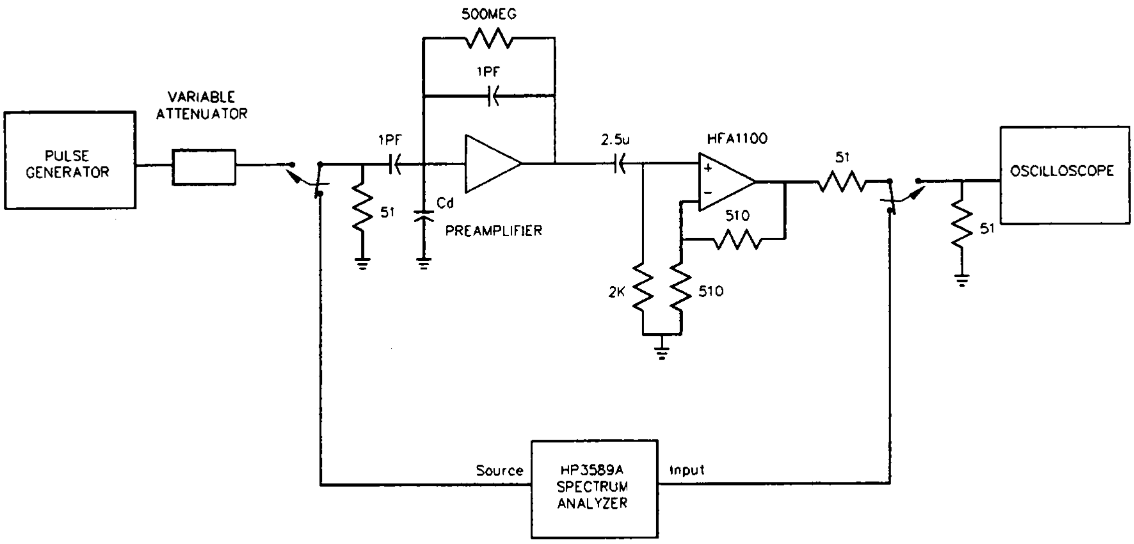

Figure 5-24. Test circuit used for measuring the preamplifier rise-time and closed-loop unity-gain frequency response. For rise-time measurements the circuit is driven with the pulse generator and the output is read by the oscilloscope. For frequency-response measurements the input is driven with the spectrum analyzer source voltage and the output is read by the spectrum analyzer input.

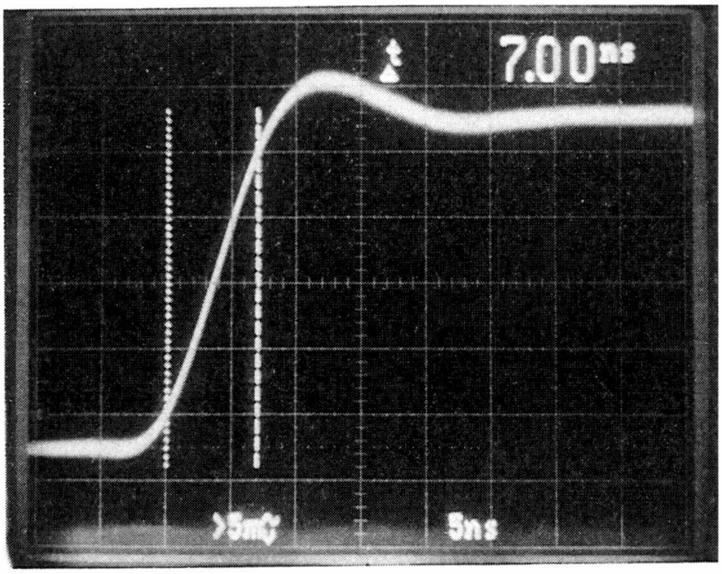

Figure 5-25. Measured transient response of the CMOS preamplifier with no additional input capacitance. $t_{10\text{-}90} = 7$ ns.

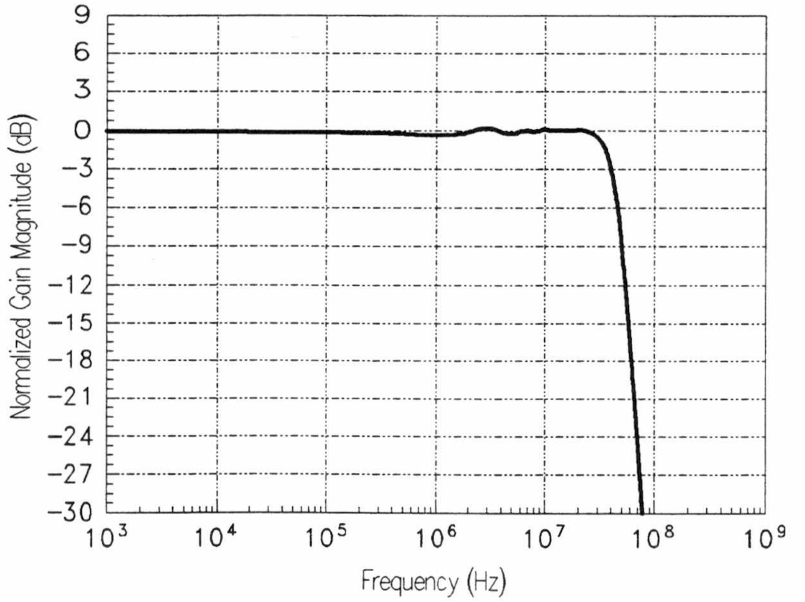

Figure 5-26. Measured frequency response of the preamplifier with no additional input capacitance. f_{3dB} =40 MHz.

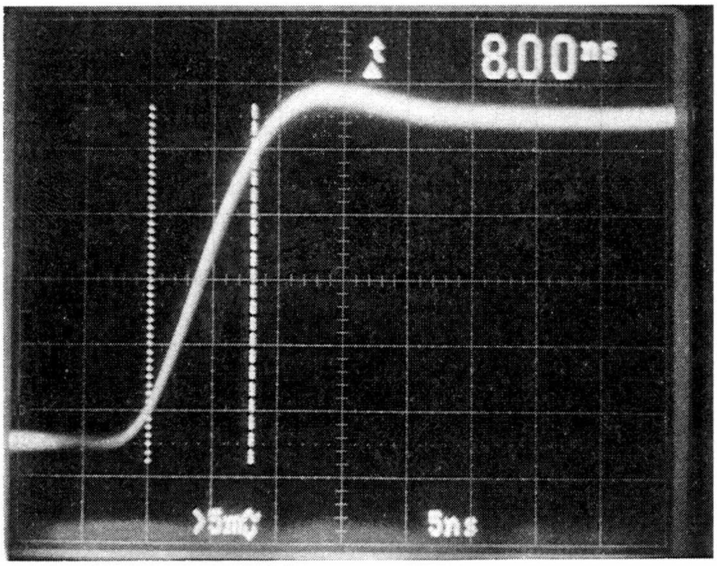

Figure 5-27. Measured transient response of the CMOS preamplifier with 5 pF additional input capacitance. $t_{10\text{-}90} = 8$ ns.

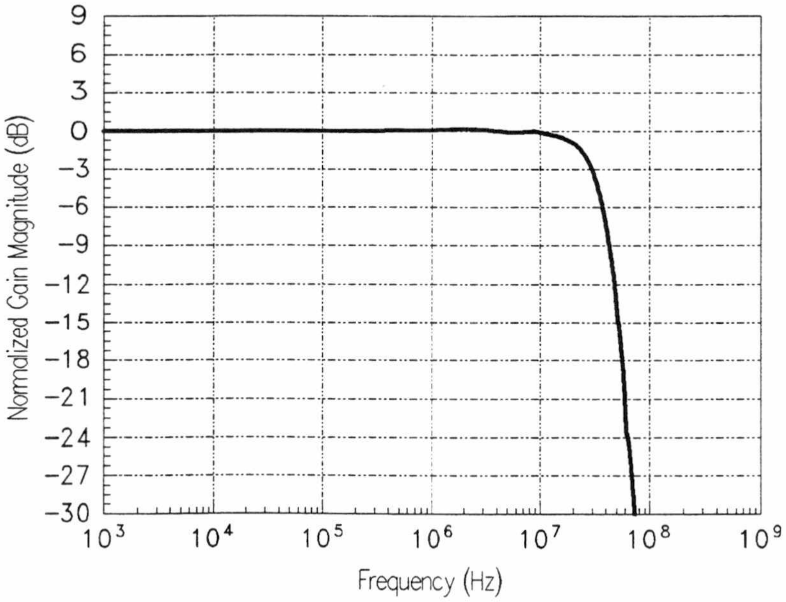

Figure 5-28. Measured frequency response of the preamplifier with 5 pF additional input capacitance. f_{3dB} =30 MHz.

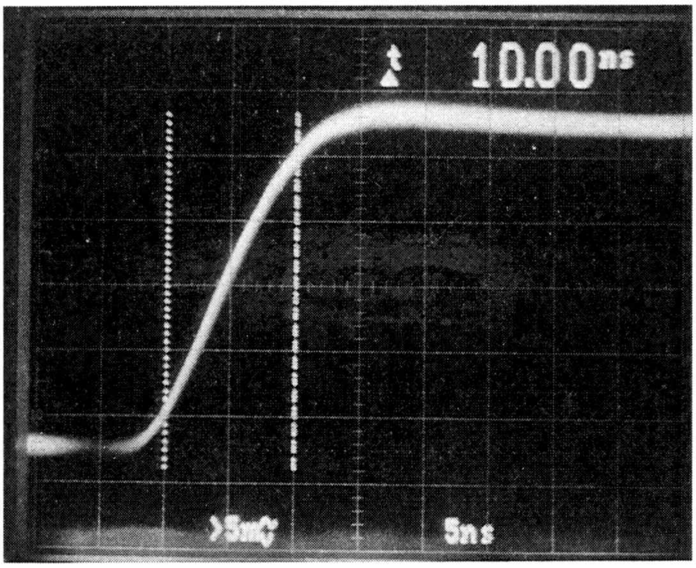

Figure 5-29. Measured transient response of the CMOS preamplifier with 10 pF additional input capacitance. $t_{10\text{-}90} = 10$ ns.

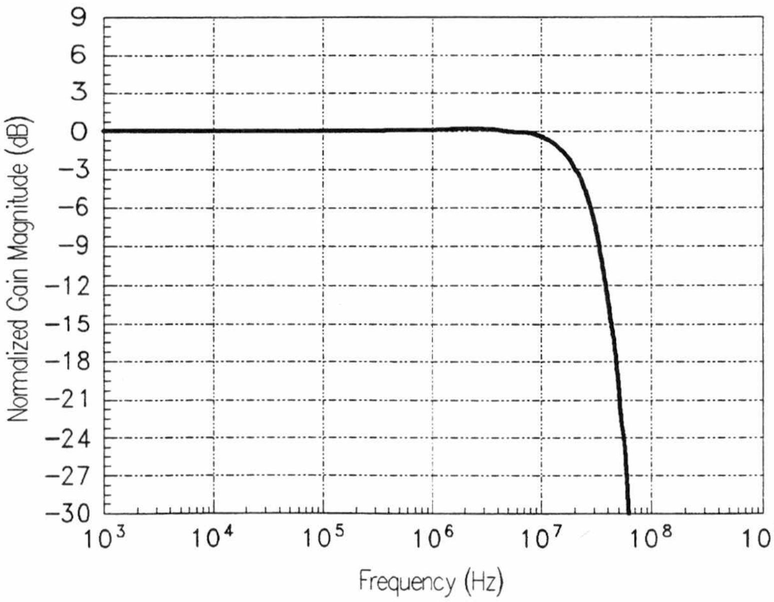

Figure 5-30. Measured frequency response of the preamplifier with 10 pF additional input capacitance. f_{3dB} =21 MHz.

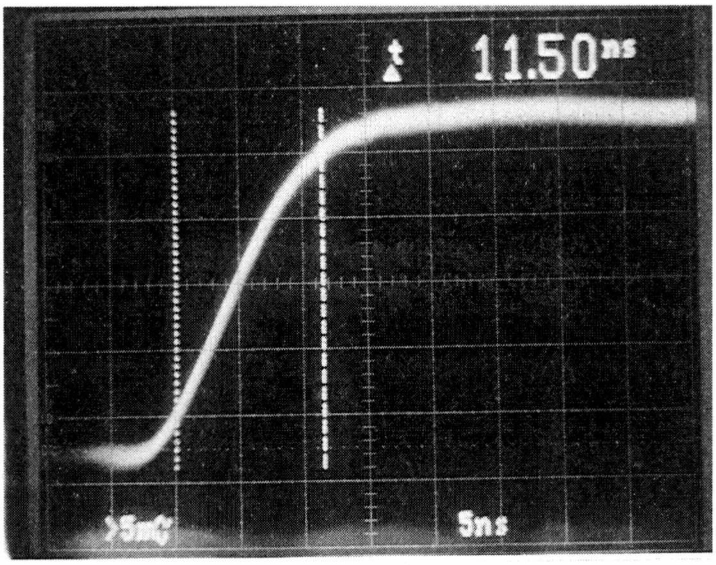

Figure 5-31. Measured transient response of the CMOS preamplifier with 15 pF additional input capacitance. $t_{10\text{-}90} = 11.5$ ns.

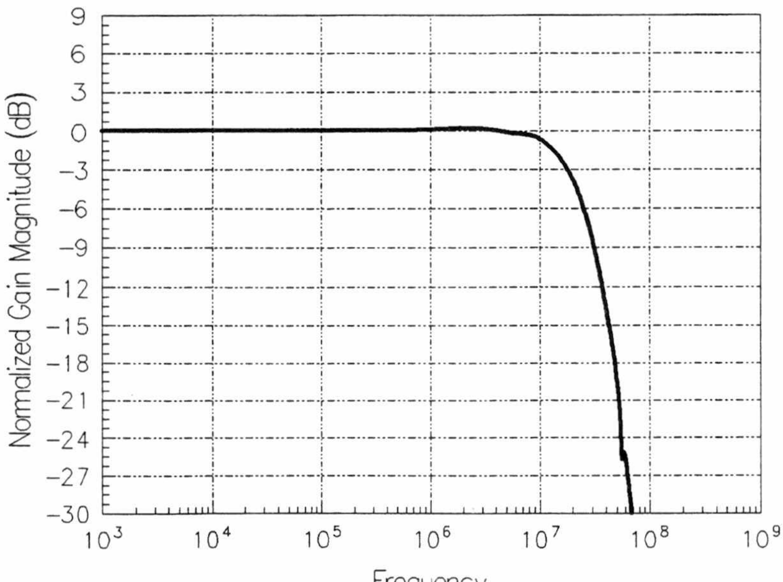

Figure 5-32. Measured frequency response of the preamplifier with 15 pF additional input capacitance. f_{-3dB} =18 MHz.

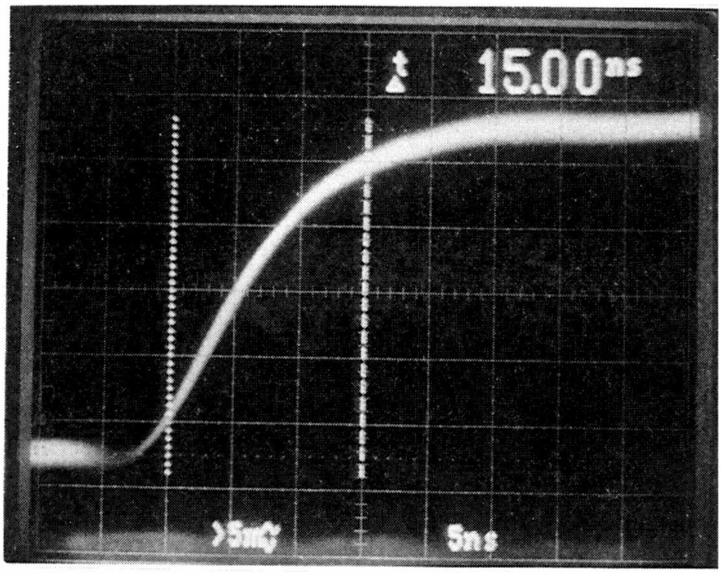

Figure 5-33. Measured transient response of the CMOS preamplifier with 22 pF additional input capacitance. $t_{10\text{-}90} = 15$ ns.

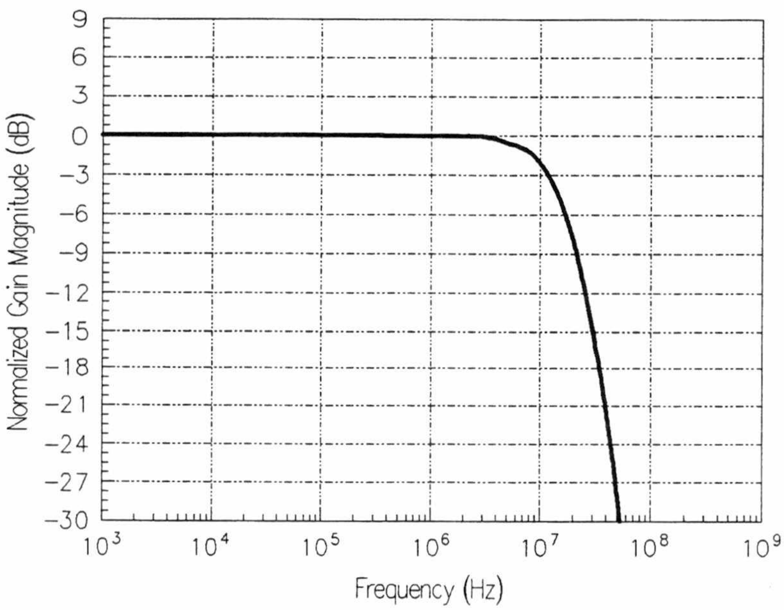

Figure 5-34. Measured frequency response of the preamplifier with 22 pF additional input capacitance. $f_{-3dB} = 14$ MHz.

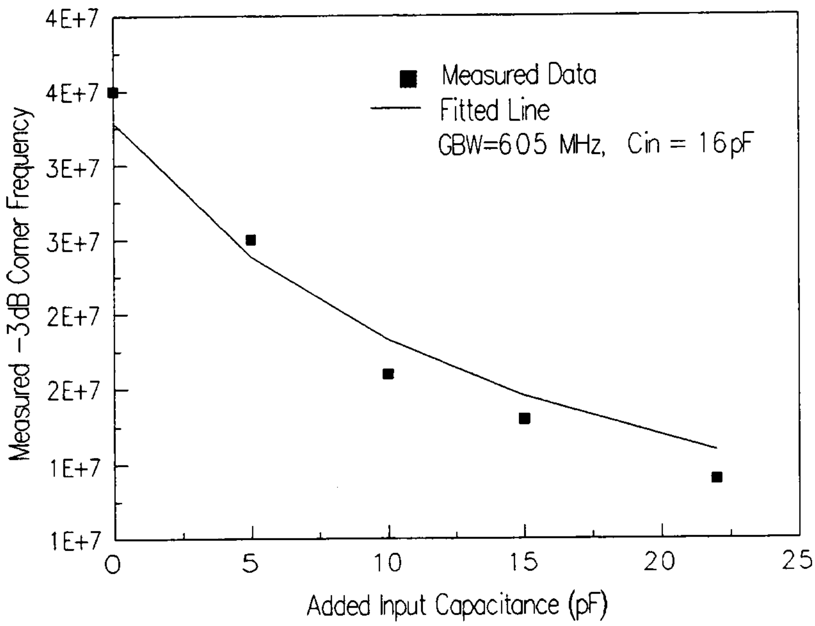

Figure 5-35. Calculation of the preamplifier gain-bandwidth product based on the measured preamplifier corner frequency as a function of input capacitance.

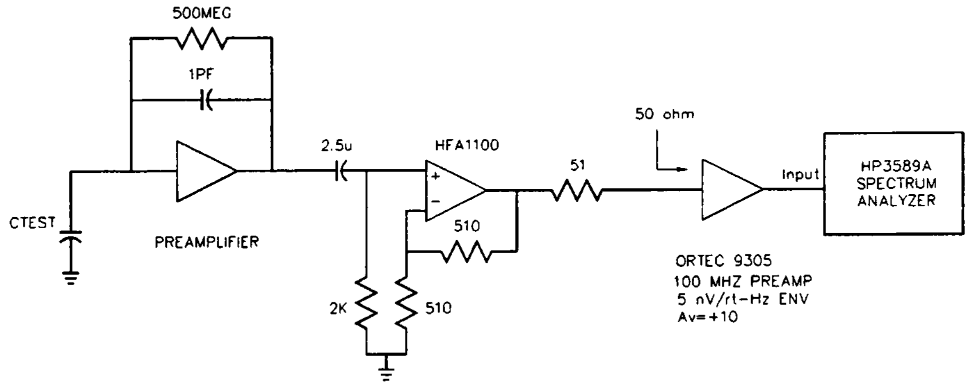

Figure 5-36. Test circuit used to measure the output noise voltage spectrum of the preamplifier.

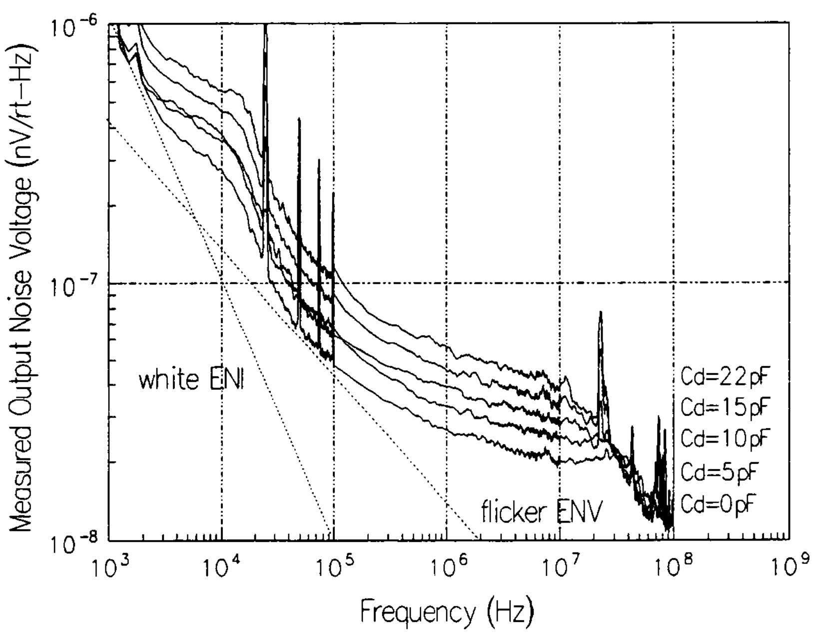

Figure 5-37. Measured output noise spectral density plots for the preamplifier with various test input capacitances.

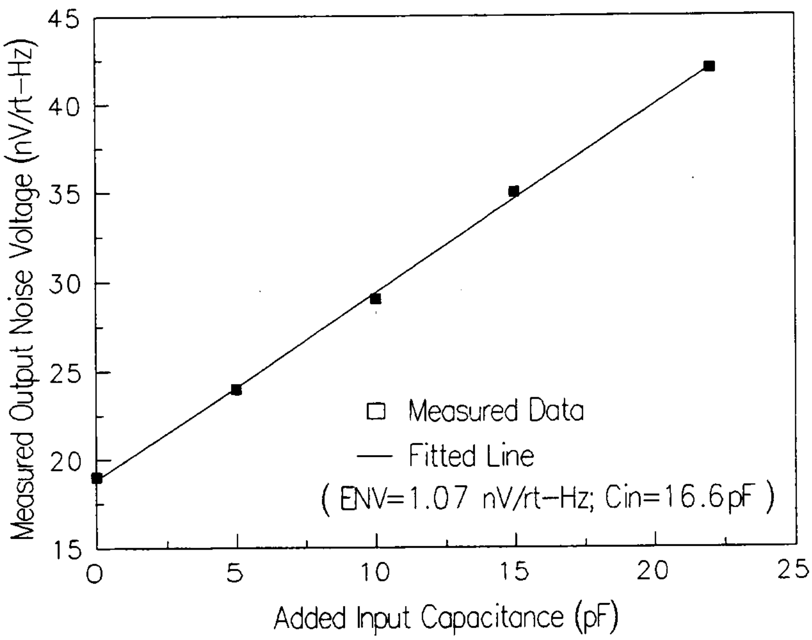

Figure 5-38. Calculation of the preamplifier equivalent input noise voltage (ENV) and input capacitance based on the measured preamplifier high frequency output noise voltage.

PART 6

EVALUATION OF THE CMOS PREAMPLIFIER WITH APD/BGO DETECTORS AND EVALUATION OF A PROTOTYPE APD ARRAY

Overview

In the previous chapters, the historical evolution of APD based detectors has been reviewed, the theory of operation for APDs has been developed, the performance of discrete BGO/APD detectors has been theoretically and experimentally evaluated and the design and evaluation of a low-noise CMOS preamplifier array has been presented. In this chapter the performance characteristics of the CMOS preamplifier coupled to an APD/BGO detector are presented and a prototype APD array is evaluated.

Unfortunately, after a two year development effort, Advanced Photonix was unable to fabricate the 64 element APD array described in Chapter 4. API was able to deliver a smaller array, however, which had been developed for a different application. Although the prototype array performance is inadequate for use in a PET system, evaluation of the array provides insight into the expected performance of an optimized array. The poor performance of the array also highlights the technical challenges associated with developing a reliable large area APD array.

The first section of this chapter is devoted to measuring the contributions of the CMOS preamplifier noise to the pulse height resolution and timing resolution of the integrated module. As a part of this investigation the differences between the noise characteristics of CR-RC and $CR\text{-}RC^4$ shaping amplifiers are reviewed. The second section presents experimental pulse height resolution and timing resolution data obtained with a discrete 5 mm diameter API APD and the CMOS preamplifier. The third section presents experimental data obtained with a prototype four element APD array and the fourth section briefly discusses a new development in scintillator technology which may significantly improve the potential performance of APD based PET detectors.

CMOS Preamplifier Equivalent Input Noise Charge and Timing Jitter

In this section the measured preamplifier equivalent input noise charge and timing jitter are presented. For background, the first sub-section compares the characteristics of the ORTEC 570 CR-RC[4] shaping amplifier used in the PHR measurements described below with the characteristics of the CR-RC shaping amplifier used to obtain the theoretical PHR data presented in Chapter 3. The equivalent input noise charge measurements are presented in the second sub-section and the timing jitter measurements are presented in the third sub-section.

Comparison of CR-RC and CR-RC⁴ Shaping Amplifiers

The theoretical PHR data presented in Chapter 3 is developed using a simple CR-RC shaping amplifier model. The measured PHR data presented below is obtained using an ORTEC 570 CR-RC[4] shaping amplifier, however, so in order to evaluate the measured data in terms of the previously derived theoretical data, it is necessary to review the differences between the ORTEC 570 and an ideal CR-RC amplifier.

Consider first the case of a charge-sensitive preamplifier coupled to a simple CR-RC shaping amplifier. The preamplifier output noise power spectral density due to a white equivalent input noise current (ENI) is given by

$$S_I(\omega) = ENI^2 \cdot \frac{1}{\omega^2 C_F{}^2} \quad \frac{A^2}{Hz}. \tag{6-1}$$

The mean-squared noise voltage at the output of a unity gain CR-RC amplifier due to ENI is

$$\left\langle V_{no,I}{}^2 \right\rangle = \frac{ENI^2}{C_F{}^2} \cdot \frac{\tau}{8} \quad V^2, \tag{6-2}$$

indicating, the output mean-squared noise due to ENI increases as τ increases. Similarly, the output noise power spectral density due to the equivalent input noise voltage is

$$S_V(\omega) = \left(ENV_{white}{}^2 + ENV_{flicker}(\omega)^2 \right) \cdot \left(\frac{C_F + C_d}{C_F} \right)^2 \quad \frac{V^2}{Hz}, \tag{6-3}$$

where ENV_{white} is the white component of the equivalent input noise voltage and $ENV_{flicker}(\omega)$ is the 1/f component of the equivalent input noise voltage. For simplicity, input capacitance of the preamplifier is lumped with the detector capacitor, C_d. The mean-squared noise voltage at the output of a unity gain CR-RC amplifier due to ENV_{white} is

$$\left\langle V_{no,VW}{}^{2}\right\rangle = ENV_{white}{}^{2}\cdot\left(\frac{C_F+C_d}{C_F}\right)^{2}\cdot\left(\frac{1}{8\tau}\right)\ V^{2}, \tag{6-4}$$

indicating the mean squared output noise voltage due to ENV_{white} decreases as τ increases. The flicker input equivalent noise voltage ($ENV_{flicker}$)is described by

$$ENV_{flicker}(\omega)^{2} = \frac{K}{\omega}\ \frac{V^{2}}{Hz}, \tag{6-5}$$

and the mean-squared output noise voltage due to $ENV_{flicker}$ is

$$V_{no,VF}{}^{2} = \frac{K}{4\pi}\left(\frac{C_F+C_d}{C_F}\right)^{2}\ V^{2}, \tag{6-6}$$

indicating the mean squared output noise voltage due to the preamplifier flicker noise does not change as a function of the shaping time constant.

The total mean squared output noise voltage is obtained by summing Equations 6-2, 6-4 and 6-6,

$$\left\langle V_{no}{}^{2}\right\rangle = \frac{ENI^{2}}{C_F{}^{2}}\frac{\tau}{8}+\frac{K}{4\pi}\left(\frac{C_F+C_d}{C_F}\right)^{2}+\frac{ENV_{white}{}^{2}}{8\tau}\left(\frac{C_F+C_d}{C_F}\right)^{2}\ V^{2}, \tag{6-7}$$

or

$$\left\langle V_{no}{}^{2}\right\rangle = A\tau+B+C\tau^{-1}. \tag{6-8}$$

The optimal shaping time constant for minimum output noise is obtained by differentiating Equation 6-8,

$$\frac{\partial V_{no}{}^{2}}{\partial\tau} = A-\frac{C}{\tau^{2}} = 0, \tag{6-9}$$

or

$$\tau_{opt}{}^{2} = \frac{A}{C} = \frac{ENV_{white}{}^{2}\left(C_F+C_d\right)^{2}}{ENI^{2}}. \tag{6-10}$$

For $\tau = \tau_{opt}$, the CR-RC shaping amplifier transfer function has a maximum gain at

$$\omega_{o,opt}^2 = \frac{1}{\tau_{opt}} = \frac{ENI^2}{ENV_{white}^2 (C_F + C_d)^2}. \qquad (6\text{-}11)$$

Substituting Equation 6-11 into Equation 6-3 yields the noise power spectral density at the preamplifier output due to ENI at $\omega = \omega_{o,opt}$,

$$S_I(\omega_{o,opt}) = ENV_{white}^2 \cdot \left(\frac{C_F + C_d}{C_F} \right)^2. \qquad (6\text{-}12)$$

Comparison of Equation 6-12 with Equation 6-3 shows that when $\omega = \omega_{o,opt} = 1/\tau_{opt}$, the noise power spectral density due to ENI and ENV_{white} are equal. Thus, the optimal (pulse height resolution) shaping time constant for a charge sensitive preamplifier coupled to a CR-RC shaping amplifier corresponds to the ENI-ENV_{white} preamplifier output noise corner frequency. $ENV_{flicker}$ increases the mean-squared output noise voltage but does not change the optimum shaping time constant.

The ORTEC 570 is a CR-RC^4 shaping amplifier where the CR^4 stage is implemented with two Sallen-Key filters. For simplicity, this analysis assumes all poles and zeroes in the ORTEC 570 transfer function are purely real so that

$$H(s) = \frac{sA\tau}{(1 + s\tau)^5}, \qquad (6\text{-}13)$$

where A is the shaping amplifier gain and τ is the shaping time constant. The mean squared output noise voltage of the CR-RC^4 shaping amplifier (A=1) due to ENI at the charge sensitive preamplifier input is

$$\left\langle V_{no,I}^2 \right\rangle = \frac{ENI^2}{C_F^2} \cdot \left(\frac{35\tau}{512} \right) \ V^2. \qquad (6\text{-}14)$$

As with the CR-RC shaping amplifier, the mean squared output noise voltage due to ENI is directly proportional to τ. Similarly, the mean squared noise voltage at the CR-RC^4 amplifier output due to ENV_{white} is given by

$$\left\langle V_{no,VW}^2 \right\rangle = ENV_{white}^2 \cdot \left(\frac{C_F + C_d}{C_d} \right)^2 \cdot \left(\frac{5}{512\tau} \right) \ V^2. \qquad (6\text{-}15)$$

Once again, the mean squared output noise voltage due to ENV_{white} is inversely proportional to τ. The mean squared output noise voltage at the CR-RC^4 amplifier output due to the equivalent input flicker noise voltage is

$$\left\langle V_{no,1/F}{}^2 \right\rangle = \frac{K}{16\pi}\left(\frac{C_F + C_d}{C_F}\right)^2, \tag{6-16}$$

which again is independent of the shaping time constant. As with the CR-RC shaping amplifier, the three components of output noise voltage have the form of Equation 6-9 and the optimal shaping time constant is the value for which the mean square output noise voltage due to ENI and ENV_{white} are equal. Substituting Equations 6-15 and 6-16 into 6-9 and evaluating yields

$$\tau_{opt,cr-rc^4} = \frac{1}{\sqrt{7}\cdot 2\pi f_c}, \tag{6-17}$$

where f_c is the ENI-ENV_{white} preamplifier output noise spectral density corner frequency. Thus, the optimal CR-RC^4 shaping time constant is a factor of $1/\sqrt{7}$ smaller than the optimal CR-RC shaping time constant.

By comparing Equations 6-4, 6-6, 6-14 and 6-15 it is evident that the total mean-squared noise of the preamplifier and CR-RC^4 shaping amplifier due to ENI and ENV_{white} is a factor of 0.206 lower than the mean-squared output noise of the preamplifier and CR-RC shaping amplifier. However, the peak pulse amplitude of the preamplifier and CR-RC^4 shaping amplifier is a factor of 0.53 lower than that of the preamplifier and CR-RC shaping amplifier so the ratio of the pulse height resolutions of the two shaping amplifiers for a given ENV and ENI is

$$\frac{PHR_{cr-rc^4}}{PHR_{cr-rc}} = \frac{\sqrt{0.206}}{0.53} = 0.85. \tag{6-18}$$

Comparison of Equations 6-6 and 6-16 indicates the mean squared output noise of the preamplifier and CR-RC^4 shaping amplifier is a factor of 4 lower than that of the preamplifier and CR-RC shaping amplifier, so the ratio of the PHR of the two shaping amplifiers for a given input flicker noise voltage is

$$\frac{PHR_{cr-rc^4}}{PHR_{cr-rc}} = \frac{\sqrt{0.25}}{0.53} = 0.94. \tag{6-19}$$

Thus, the PHR measured with the ORTEC 570 will be lower than (superior to) the PHR which would be measured with a CR-RC shaping amplifier by a factor between 0.85 and 0.94.

The swept frequency response of the ORTEC 570 is shown in Figure 6-1 for each of the front panel shaping time constant settings, and the measured frequencies of peak output signal magnitude are summarized in Table 6-1. For a CR-RC^4 shaping amplifier with all real poles, the frequency of peak output signal magnitude occurs at $f_p = 1/(4\pi\tau)$. The calculated time constants associated with the measured ORTEC 570 peak frequencies are listed in the third column of Table 6-1. The noise corner associated with ENI and ENV_{white} may be experimentally estimated by determining which time constant provides the best signal-to-noise ratio. The fourth column of Table 6-1 shows the approximate noise corner frequency which would exist for each time constant if that time constant were found to be optimal.

Table 6-1.
Summary of ORTEC 570 Filter Amplifier Characteristics.

Front Panel Switch Setting	**Peak Frequency**	**Approximate CR-RC^4 Shaping Time Constant**	**Related ENI-ENV_{white} Noise Corner**
0.5 μs	300 kHz	250 ns	238 kHz
1.0 μs	160 kHz	475 ns	127 kHz
2.0 μs	85 kHz	891 ns	68 kHz
3.0 μs	60 kHz	1.3 μs	48 kHz
6.0 μs	35 kHz	2.2 μs	28 kHz
10.0 μs	23 kHz	3.3 μs	18 kHz

Pulse Height Resolution

As discussed in Chapter 3, preamplifier noise can be an important source of pulse height resolution loss in an APD based detector. The input equivalent noise charge of the CMOS preamplifier and ORTEC 570 shaping amplifier was measured using the test circuit shown in Figure 6-2. The pulse generator and attenuator settings are selected to generate a 1.6 mV step voltage at the attenuator output so that a charge of

$$Q_{in} = \frac{1.6\,mV \cdot 1\,pF}{q} = 10{,}000\,electrons \tag{6-20}$$

is injected into the preamplifier input. The input equivalent noise charge of the preamplifier and shaping amplifier was measured for a 10,000 electron test pulse at each shaping time constant setting with added input capacitances of 0 pF, 5 pF, 10 pF and 22 pF. The data are shown in Figure 6-3. For C_{in} = 0 and 5 pF, the optimal shaping time constant setting is 2 μs, which corresponds to an ENI-ENV$_{white}$ noise corner of approximately 64 kHz (Table 6-1). Extrapolation of the white high frequency component of the measured output noise spectra in Figure 5-38 shows that the ENI-ENV white noise corner does occur near 60 kHz for C_{in} = 0 pF and 5 pF. For C_{in} = 10 pF and 22 pF the optimal shaping time constant setting is 3 μs, which corresponds to an ENI-ENV$_{white}$ noise corner of approximately 43 kHz (Table 6-1). Examination of Figure 5-38 shows that, because the high frequency noise is higher, the ENI-ENV$_{white}$ crossover frequency is reduced to a value near 40 kHz.

For the APD application, the total noise current at the preamplifier input is much greater than that of the preamplifier alone due to the multiplication noise of the APD, causing the ENI-ENV$_{white}$ noise corner to occur at higher frequencies. The theoretical data presented in Chapter 3 shows the optimal CR-RC shaping time constant for an APD detector to be on the order of a few hundred nanoseconds.

Based on the measured data shown in Figure 6-3, the preamplifier and shaping amplifier equivalent input noise charge is always less than 2800 electrons FWHM. A typical BGO crystal emits approximately 1500 electrons per 511 keV pulse. If the APD has an 85% quantum efficiency and a gain of 200, the total input signal is approximately 255,000 electrons. Thus, the PHR associated with the preamplifier alone is less than 2.6%. Measured and experimental data presented in Chapter 3 indicate the expected pulse height resolution of an APD/BGO detector is typically greater than 10%, so the noise of the preamplifier should have only a small effect on the detector pulse height resolution.

Timing Resolution

The calculated data in Chapter 3 shows that the input noise voltage (ENV) of the preamplifier strongly influences the timing resolution achievable with an APD/BGO detector. For example, Figure 3-30 shows that if a 5 mm diameter Advanced Photonix APD (M=500, C_d=16 pF) is coupled to a preamplifier with a 10 pF input capacitance and a purely white ENV of 2 nV/$\sqrt{\text{Hz}}$, the optimal timing resolution is achieved with a shaping time constant of 5 ns, where the output noise due to the preamplifier contributes approximately 3.75 ns FWHM to the timing spectrum width and the noise of the APD contributes approximately 4 ns FWHM to the timing spectrum width. Adding these components in quadrature yields a net timing resolution of 5.5 ns FWHM. In this section the contribution of the preamplifier noise to the system timing spectrum is measured and compared with the data of Chapter 3.

In Chapter 3 it was shown that the APD should be biased for a gain of 500 or greater or optimal timing resolution. If the APD has a quantum efficiency of 85% and the BGO emits 1500 photons per 511 keV event, the APD/BGO detector (M=500) will inject approximately 640,000 electrons into the preamplifier input. The input charge pulse will have a 300 ns decay constant associated with the light emission characteristics of BGO.

Figure 6-4 shows the test circuit used to evaluate the preamplifier timing resolution when excited with a 640,000 electron pulse with a 350 ns decay time constant. (A 350 ns decay constant was selected to make use of an available capacitor.) A -10V step voltage with a ~1 ns rise time drives a -40 dB attenuator which has a 50Ω output impedance. The 0.007 μF capacitor at the attenuator 50Ω output forms a 455 kHz pole so that the -100 mV voltage step has a 350 ns time constant. The -100 mV voltage step drives a 1 pF capacitor in series with the preamplifier so that a 640,000 electron pulse with 350 ns decay constant is injected into the preamplifier input. The preamplifier output drives a CR-RC shaping amplifier with variable time constant which in turn drives a Tennelec 455 CFD. The CFD delay is chosen to be 4 ns longer than the CR-RC shaping time constant to insure true CFD timing [1]. The CFD output drives the STOP input of an ORTEC 457 time to amplitude converter.

The pulse generator simultaneously drives an RC network which differentiates the pulse to present a -10V, 3 ns wide pulse to the ORTEC 457 START input. The variable delay in the START signal path is used to calibrate the MCA. A Nucleus PCA2 multi channel analyzer is used to histogram the TAC output signal.

The measured timing resolution is shown in Figure 6-5 as a function of the CR-RC shaping time constant with no additional capacitance inserted at the preamplifier input. The measured timing jitter associated with the preamplifier noise is less than 3.5 ns FWHM for most time constants. Unlike the calculated data of Chapter 3, however, the timing resolution due to the preamplifier noise does not improve at longer time constants. The theoretical data presented in Chapter 3 is based on the assumption that the timing resolution is determined by the maximum signal slope rather than by the signal rise-time. The slope of a CR-RC shaping amplifier driven by a step voltage is relatively independent of the shaping time constant while the rise-time is strongly dependent on the shaping time constant. The measured data suggests the CFD used here triggers more accurately on the smaller amplitude, shorter rise-time signals generated by the CR-RC amplifier when short time constants are employed.

Figure 6-6 shows the measured timing resolution obtained as a function of added input capacitance for a shaping time constant of 10 ns. As expected, the timing resolution degrades as the input capacitance is increased due to the increase in the output noise of the preamplifier. The calculated timing resolution curve shown in Figure 6-6 is obtained using the measured and calculated noise parameters presented in Chapter 6 (ENV_{white} = 1.07 nV/√Hz, ENI = 5.76 fA/√Hz, and $K_{F,M5} = 3.5 \times 10^{-28}$ C/m^2) and PSPICE simulated output signal slopes. The calculations show that 97% of the mean squared noise voltage at the shaping amplifier (τ_{sh} = 10 ns) is due to the white component of the preamplifier equivalent input noise voltage. Despite having a noise corner approaching 1 MHz, the flicker component of the equivalent noise voltage contributes only 3 % of the mean squared noise voltage. The calculated and measured curves show reasonable agreement, although the measured data has a weaker dependence on C_{in} than the calculated data. The difference in the curve slopes may be due in part to the fact that the calculations did not consider the variations in the preamplifier noise bandwidth as a function of input capacitance.

Evaluation of the CMOS Preamplifier Coupled to an APD/BGO Detector

In the previous section the contributions of the preamplifier noise to the system pulse height resolution and timing resolution were assessed. In this section the performance characteristics of the preamplifier and a discrete APD/BGO detector are presented. The measured data is compared with the data of Chapter 3 obtained with a similar APD/BGO detector and an ORTEC 142A charge sensitive preamplifier.

Pulse Height Resolution Measurements

Pulse height resolution data was obtained using the circuit shown in Figure 6-7. The scintillator is a Teflon wrapped 3 x 3 x 25 mm^3 BGO crystal and the APD is a 5 mm diameter Advanced Photonix (API) device (C_d = 16 pF). As discussed in Chapter 5, the preamplifier is biased with I_{BIAS} = 2.6 mA, V_{BIAS} = -3V and V_{M1WELL} = +2V. The ORTEC 570 shaping amplifier is configured for a shaping time constant of 0.5 μs, the shortest available shaping time constant. The measured PHR is shown in Figure 6-8 as a function of the APD bias voltage. The optimal PHR (12.0% FWHM) is obtained at a bias voltage of 2550V, which corresponds to an APD gain of approximately 100. The optimal PHR spectrum is shown in Figure 6-9. Both the escape peak and the back-scatter peak are resolved in the spectrum and the noise floor is approximately 70 keV. This is believed to be the best pulse height resolution reported to date for an APD coupled to a BGO crystal with a geometry suitable for a PET detector (length > 20 mm). The PHR measurement obtained with the CMOS preamplifier compares favorably with the measurements presented in Chapter 3 for a similar detector where an optimum PHR of 14.2% FWHM was obtained at a bias voltage of 2560V using an ORTEC 142A preamplifier.

Timing Resolution Measurements

Timing resolution measurements were made using the circuit shown in Figure 6-10. The test circuit is a standard coincidence timing measurement system [1] and the APD, scintillator and preamplifier configuration is identical to that used in the PHR measurements. The system START signal is generated with a fast plastic scintillator coupled to an RCA 8575 PMT in an ORTEC 265 tube base. The PMT output drives a TENNELEC 455 CFD which in turn drives the START input of an ORTEC 457 time to amplitude converter.

The APD/BGO detector and preamplifier generate the system STOP signal. The CMOS preamplifier drives an ORTEC 9305 preamplifier which serves as a wide-band line driver for both the timing channel and energy qualification channel. The APD/BGO timing channel is formed by an ORTEC 579 CR-RC filter amplifier which drives a TENNELEC 455 CFD. The CFD output passes through a variable delay line used for system calibration into the STOP input of the time to amplitude converter. The time to amplitude converter output is histogrammed by a Nucleus PCA2 multi channel analyzer.

An energy qualification channel is included in the circuit to ensure all recorded events are associated with 511 keV γ-rays. This channel consists of an ORTEC 570 shaping amplifier, an ORTEC 553 single channel analyzer and an ORTEC 416A gate and delay unit. The energy qualification window is 511 ±50 keV.

The measured timing resolution is shown in Figure 6-11 as a function of APD bias voltage. The optimal CR-RC integration time constant is 10 ns, the shortest available time constant on the ORTEC 579. The optimal CR-RC differentiation time constant is 50 ns. The longer differentiation time constant offers the advantage of increasing the signal amplitude and the disadvantage of increasing the shaping amplifier noise bandwidth. The 50 ns differentiation time constant provides the most timing resolution improvement at low APD bias voltages where the signal amplitude is smallest. At higher bias voltages (>2600 V) the difference between the timing resolution obtained with τ_{diff} = 50 ns and τ_{diff} = 10 ns is negligible.

The best timing resolution (9.08 ns FWHM) is obtained at the highest APD bias voltage of 2700 V, which corresponds to an APD gain of between 300 and 400. Advanced Photonix has recently determined that biasing their devices for gains above 200 ($V_{bias} \approx$ 2600V) can lead to degraded performance and ultimately failure of the device. For this reason the APD was biased above 2600V for only a short period of time and was not biased above 2700V at all. The best timing spectrum obtained at a bias voltage of 2700V is shown in Figure 6-12.

The measured timing resolution again compares favorably with the data presented in Chapter 3, where a timing resolution of 9 ns FWHM was obtained using a 5 mm API APD biased at 2700 V, a 3 x 3 x 25 mm^3 BGO crystal and an ORTEC 142A preamplifier. However, as discussed in Chapter 3, the measured timing resolution of 9 ns is significantly poorer than the ~5 ns FWHM predicted by the model. The discrepancy between the measured and modeled data is attributed to traps in the APD anti-reflective coating/silicon interface which degrade the APD transient response.

Evaluation of a Four Element Prototype APD Array

Although Advanced Photonix has been unable to deliver the 64 element array designed for the first APD/BGO PET detector, they were able to deliver a smaller prototype array for evaluation. The prototype array performance is inadequate for use with PET, but it does provide insight into the way in which the API array technology provides isolation between pixels.

The prototype array is designed for a non-PET application calling for 64 pixels with an inter pixel spacing of 1.18 mm. For this study the array is reconfigured for use with 3x3 mm^3 BGO crystals as shown in Figure 6-13, where the outer pixels are tied together to form a guard ring and the inner pixels are connected to form four 3.52 x 3.52 mm^2 "super-pixels". As discussed in Chapter 4, the pixels are isolated with etched backside grooves which pinch off the channels between pixels when the APD is biased. Figure 6-14 shows the measured inter-pixel resistance as a function of bias voltage. At low bias voltages the inter pixel resistances are less than 1 kΩ, and the array functions electrically as a single large-area APD. The pixels are isolated for bias voltages above 2200 V, with inter-pixel resistances approaching 10 MΩ.

Figure 6-15 shows the measured pulse height resolution of each of the four "super-pixels" measured using the CMOS preamplifier array and an ORTEC 570 shaping amplifier. The optimal pulse height resolution is obtained with a bias voltage of 2400 V, near the break-down voltage of the array. The pulse height spectra are dominated by the dark current noise of the APD array which is attributed to charge injection from the isolation grooves [2]. The noise in pixel 4 completely obscures the BGO photopeak. Efforts to obtain timing spectra with the array were unsuccessful. The poor performance of this prototype array is indication that significant technical challenges still must be conquered before a reliable APD array suitable for use with PET is commercially available.

An Alternative Scintillator for Use with APD Based PET Detectors

The data presented in this Chapter and in Chapter 3 shows that discrete APD/BGO detectors have superior pulse height resolution compared to conventional PMT/BGO block detectors (block detector PHRs are typically ~ 20% FWHM) but have poorer timing resolution (block detector timing resolutions are typically < 3.5 ns FWHM). Recently a new scintillator, Lutetium Oxyorthosilicate (LSO), has become available which greatly improves the timing performance of APD based detectors. LSO emits approximately four times the light of BGO and has a decay time constant of approximately 40 ns (vs 300 ns for BGO)[3]. If LSO becomes widely available, it may be a viable alternative to BGO in APD based PET detectors.

Figure 6-16 shows a pulse-height resolution spectrum obtained with the 5 mm diameter Advanced Photonix APD (V = 2400V) coupled to a 3x3x22 mm^3 LSO crystal using the test circuit shown in Figure 6-7. Despite the higher light output of the LSO crystal, the 14.3% pulse height resolution is slightly poorer than that obtained with BGO. The broader than expected pulse height resolution is attributed to imperfections in the crystal structure [4]. Figure 6-17 shows a timing spectrum obtained with the 3x3x22 mm^3 LSO crystal and 5 mm diameter APD (V = 2700 V) obtained using the test circuit of Figure 6-10 (t_{int}=10 ns, t_{diff}=50 ns). The measured timing resolution (1.57 ns FWHM) is vastly superior to that obtained with BGO (9.08 ns) due to the increased light output and shorter decay time constant of LSO. This encouraging early data suggests the future of APD based PET detectors may lie with the use of LSO as the scintillation material of choice.

Conclusions

In this chapter the system level performance of the CMOS preamplifier coupled to an APD/BGO detector are presented, a prototype APD array is evaluated and a new scintillator which may offer improved performance in future APD based detectors is introduced. The first section presents the measured equivalent input noise charge and timing jitter of CMOS preamplifier. The equivalent input noise charge is less than 2800 electrons FWHM for all shaping time constants and input capacitances considered and less than 1400 electrons FWHM for a shaping time constant of 2 μs and an input capacitance of 10 pF. The measured preamplifier timing jitter is less than 4 ns FWHM for shaping time constants less than 80 ns and detector capacitances less than 22 pF and is approximately 3 ns FWHM for a shaping time constant of 10 ns and a detector capacitance of 10 pF. The second section presents measured pulse height resolution and timing resolution data obtained using the CMOS preamplifier and a discrete APD/BGO detector. The best measured pulse height resolution (12% FWHM) and timing resolution (9.08% FWHM) compare favorably with the data presented in Chapter 3 for a similar detector and a commercially available charge sensitive preamplifier. The third section presents measured data obtained with a prototype APD array. The measurements indicate the array pixels are fully isolated for bias voltages greater than 2100V. The dark current noise of the APD significantly degrades the measured PHR obtained with the array. The final section briefly introduces a new scintillator, LSO, which significantly improves the timing resolution of the APD based detector. Measured pulse height resolution (14.2%

FWHM) and timing resolution (1.57 ns FWHM) data obtained using a 5 mm diameter Advanced Photonix APD and the CMOS preamplifier offer encouragement that LSO may be the optimal scintillator for future APD based detectors.

References

1. "Principles and Applications of Timing Spectroscopy," EG&G ORTEC Application Note AN-42.

2. Joe Boisvert, Advanced Photonix Member of Technical Staff, private communication.

3. C. L. Melcher and J. S. Schweitzer, "Cerium Doped Lutetium Oxyorthosilicate: a Fast Efficient New Scintillator," *IEEE Trans. Nucl. Sci.*, **NS-39**, pp. 228-232, 1992.

4. Mark S. Andreaco, private communication.

Figures

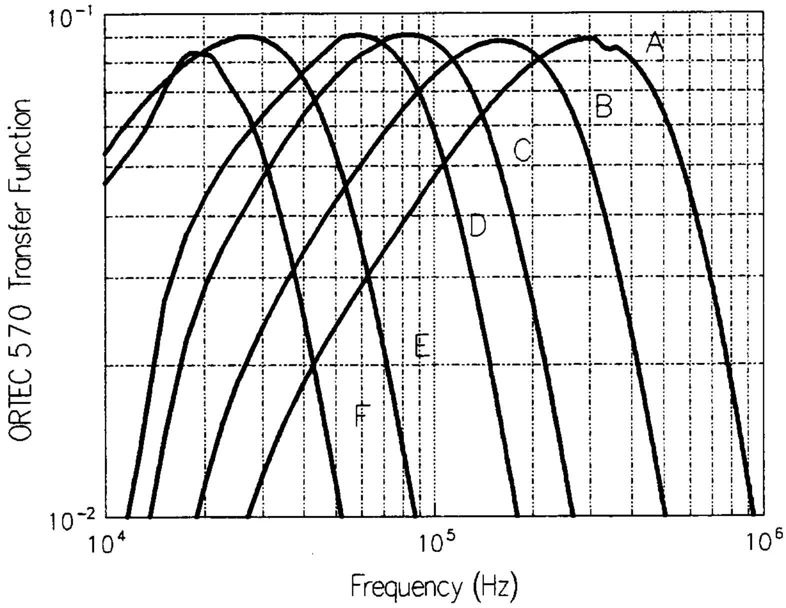

Figure 6-1. Measured frequency response of the ORTEC 570 shaping amplifier. Shaping time constant: (A) 0.5 μs, (B) 1.0 μs, (C) 2.0 μs, (D) 3.0 μs, (E) 6.0 μs, (F) 10.0 μs.

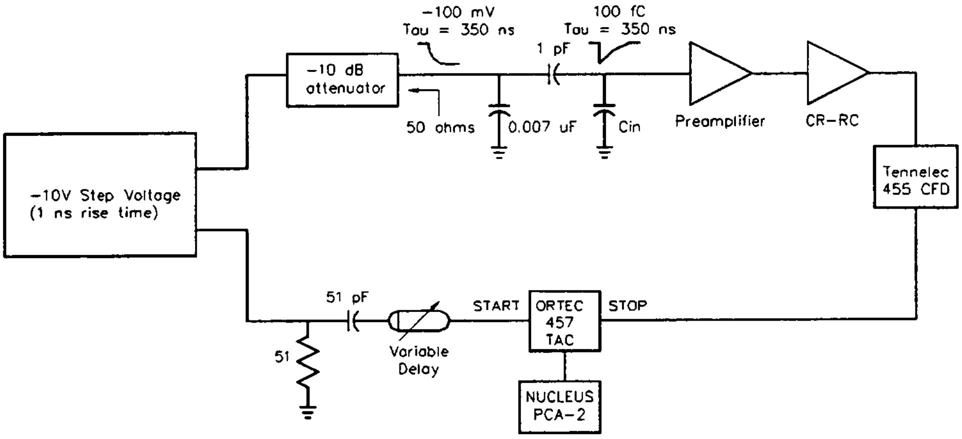

Figure 6-2. Test circuit for measuring the pulse height resolution of the preamplifier using a 10,000 electron input test pulse.

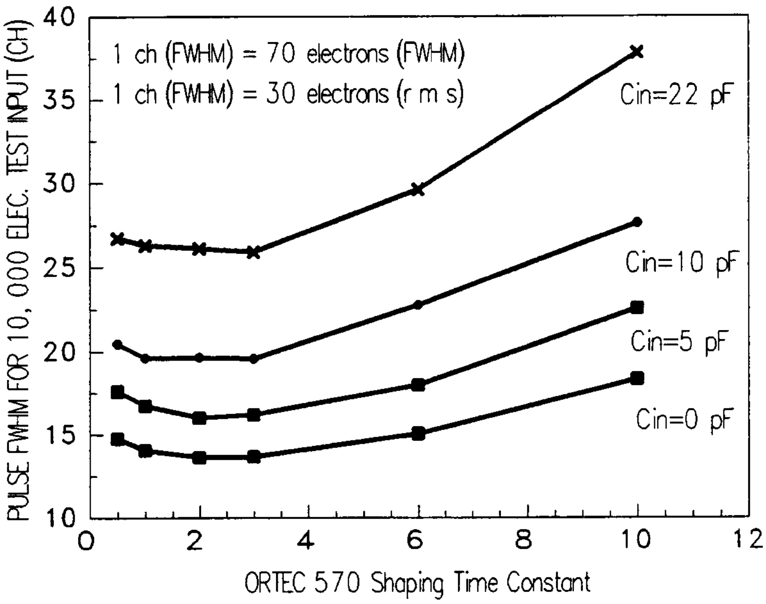

Figure 6-3. Measured preamplifier output noise charge as a function of shaping time constant obtained using a 10,000 electron test pulse and an ORTEC 570 shaping amplifier.

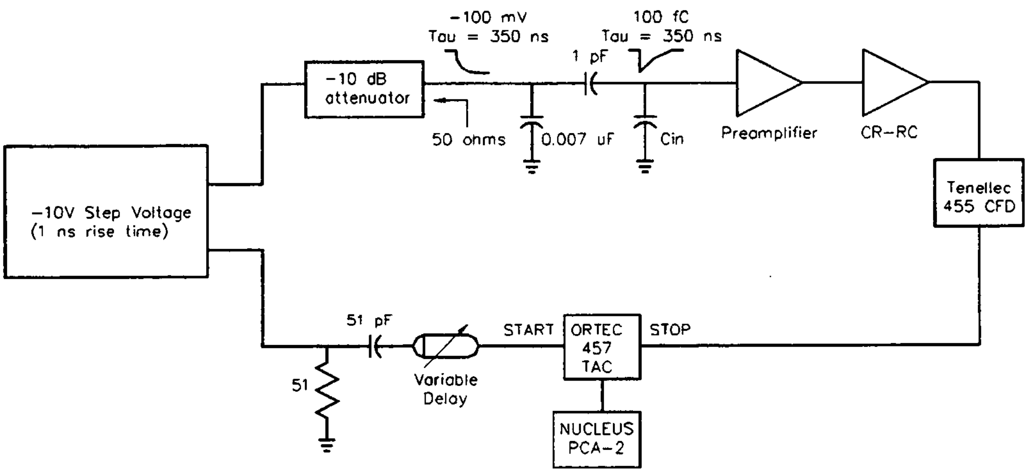

Figure 6-4. Test circuit used to measure the intrinsic timing resolution of the CMOS preamplifier.

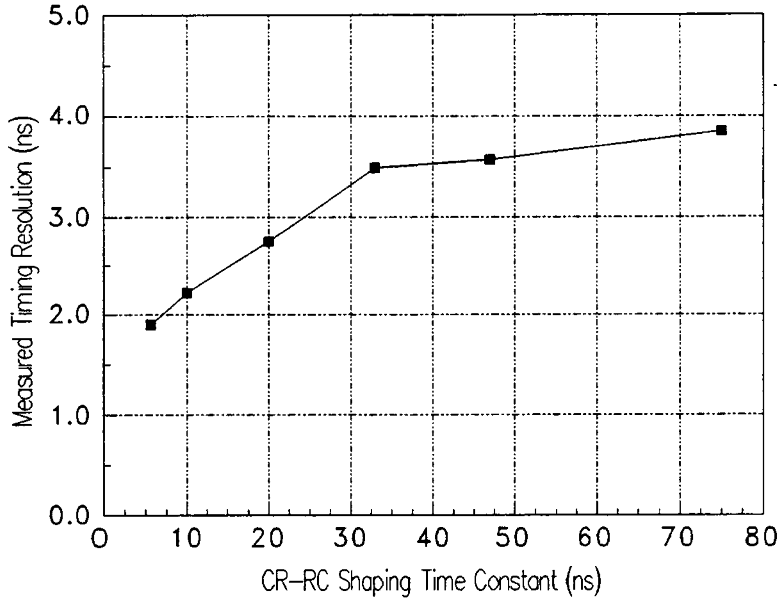

Figure 6-5. Measured timing resolution of the CMOS preamplifier array as a function of the CR-RC shaping amplifier time constant obtained using the test fixture of Figure 6-4. No additional capacitance was added to the preamplifier input for these measurements.

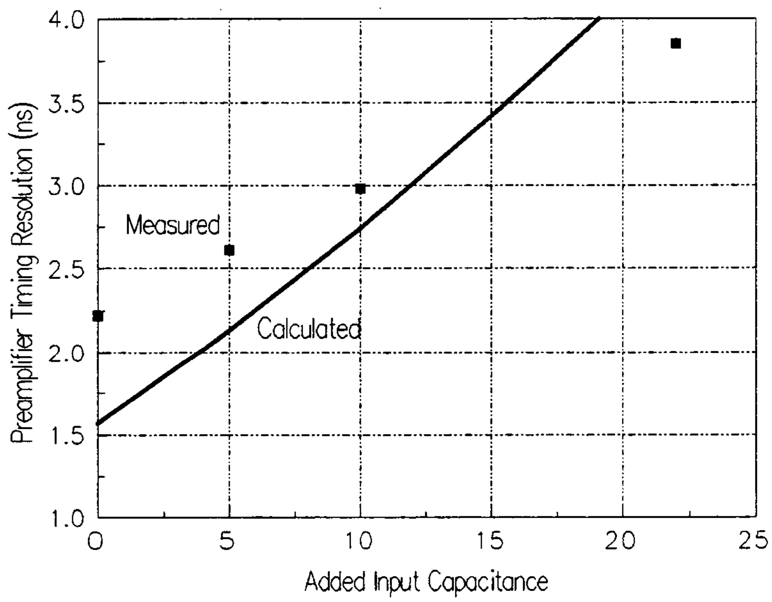

Figure 6-6. Measured and calculated timing resolution of the CMOS preamplifier array as a function of the added input capacitance obtained using the test fixture of Figure 6-4. CR-RC shaping time constant = 10 ns.

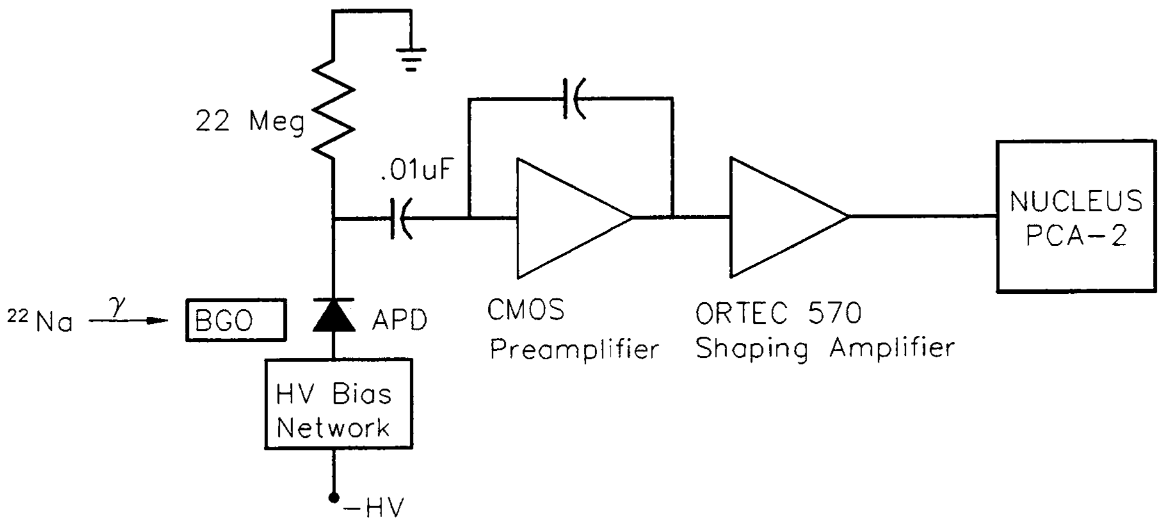

Figure 6-7. Pulse height resolution test circuit incorporating the CMOS preamplifier array.

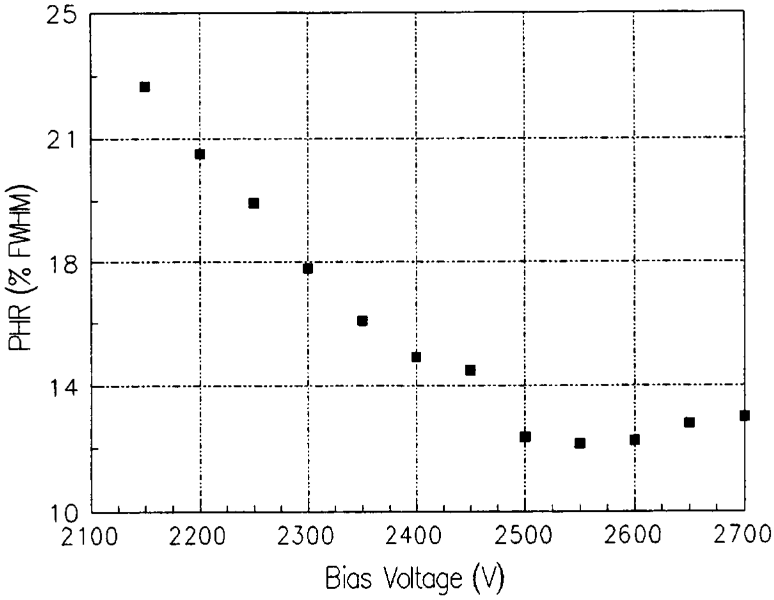

Figure 6-8. Measured PHR vs bias voltage using the CMOS preamplifier and a 5 mm diameter Advanced Photonix APD. Shaping time constant = 0.5 us, scintillator = 3 x 3 x 25 mm^3 BGO.

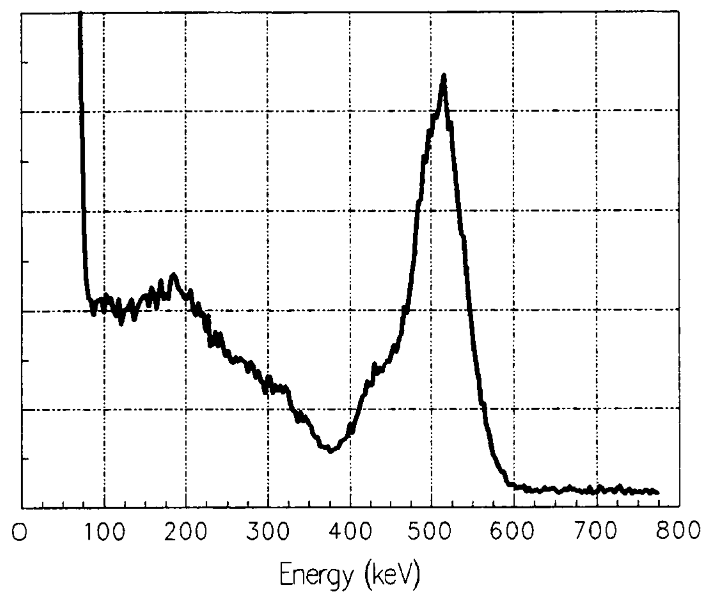

Figure 6-9. Best measured PHR (12%) obtained with a 5 mm diameter API APD and the CMOS preamplifier. Vbias = 2550 V, shaping time constant = 0.5 μs.

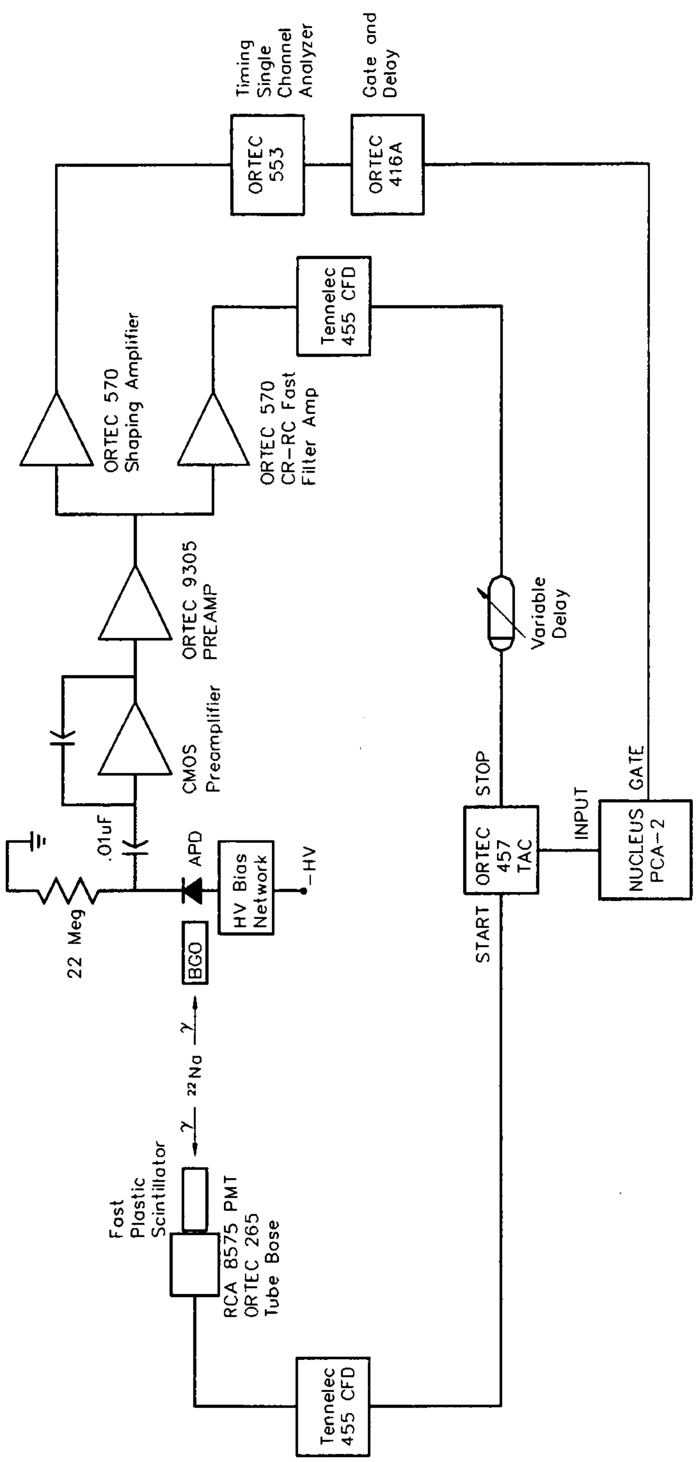

Figure 6-10. Timing resolution test circuit incorporating the CMOS preamplifier array.

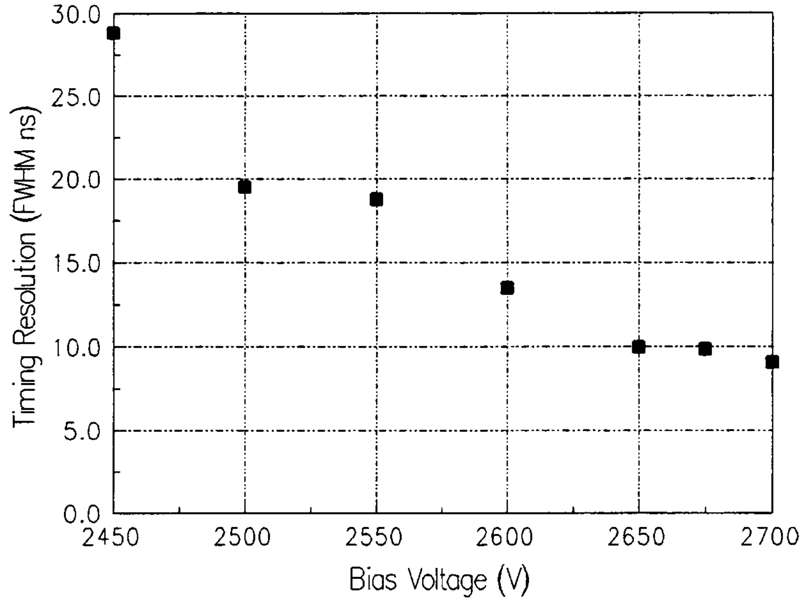

Figure 6-11. Measured timing resolution vs bias voltage using the CMOS preamplifier and a 5 mm diameter Advanced Photonix APD. Integration shaping time constant = 10 ns, differentiation shaping time constant = 50 ns.

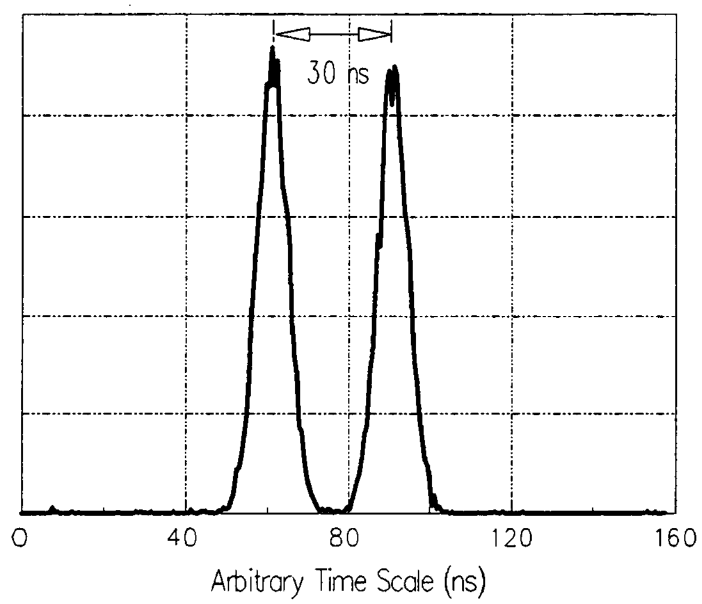

Figure 6-12. Best measured timing resolution (9.08 ns) obtained with a 5 mm diameter API APD and the CMOS preamplifier. Vbias = 2700 V, τ_{int} = 10 ns, τ_{diff} = 50 ns.

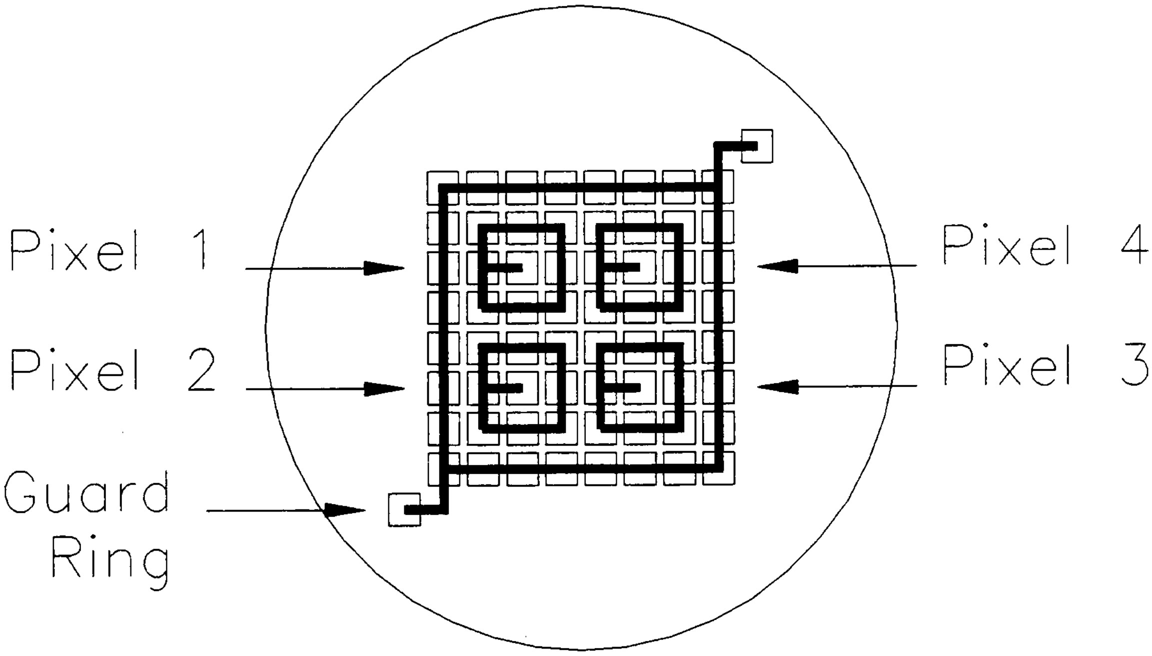

Figure 6-13. 64 element APD array with 1.3 mm x 1.3 mm elements configured as a 4 element array with 3.8 mm x 3.8 mm elements.

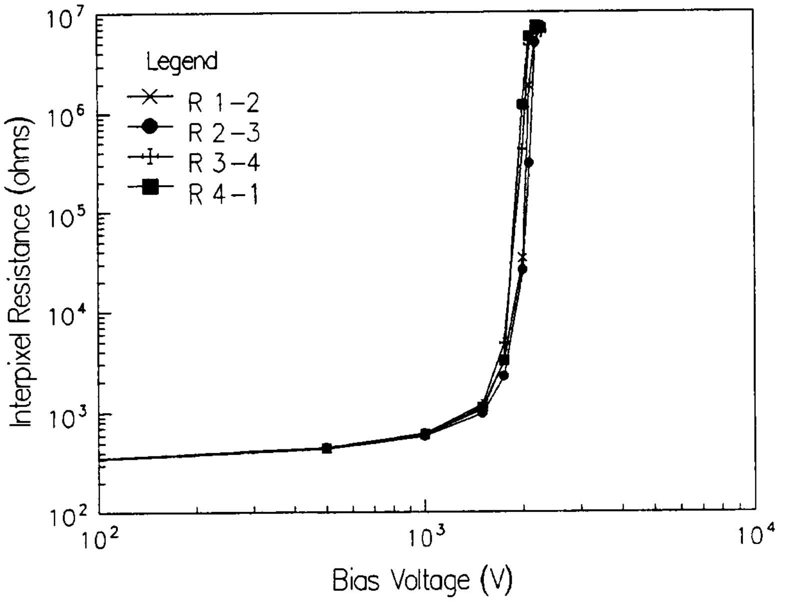

Figure 6-14. Measured inter-pixel resistance as a function of bias voltage. The pixels are isolated for V>2100V.

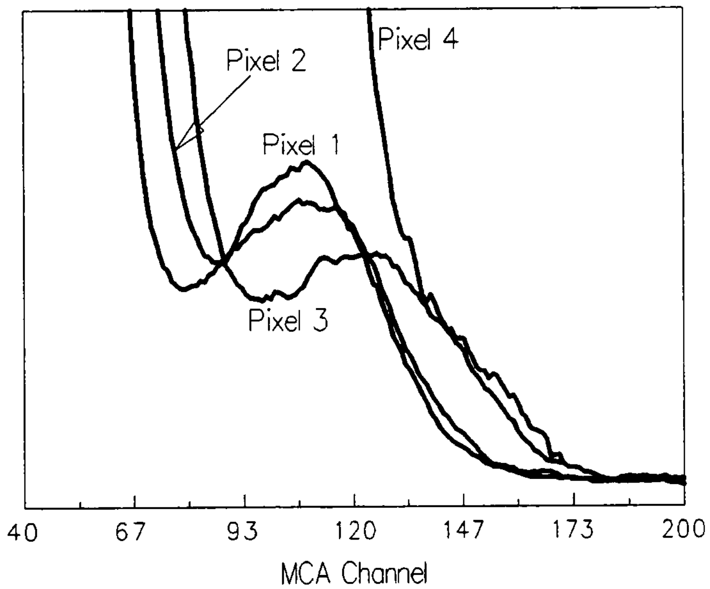

Figure 6-15. Measured PHR of pixels 1-4 of the 4 element API APD array. Vbias = 2400V, shaping time constant = 0.5 μs. Pixel 4 photopeak is obscured by the noise.

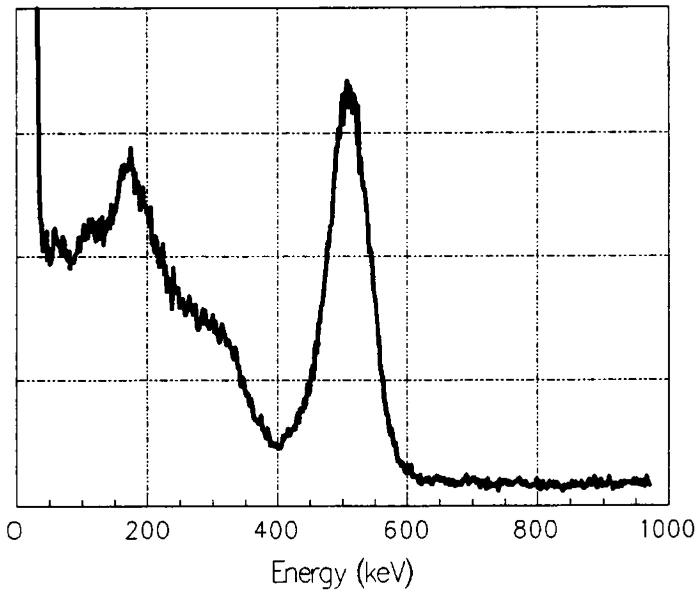

Figure 6-16. Measured PHR (14.3%) obtained with a 5 mm diameter API APD, the CMOS preamplifier and a 3x3x22 mm^3 LSO crystal. Vbias =2400V and τ_{sh} = 0.5 μs.

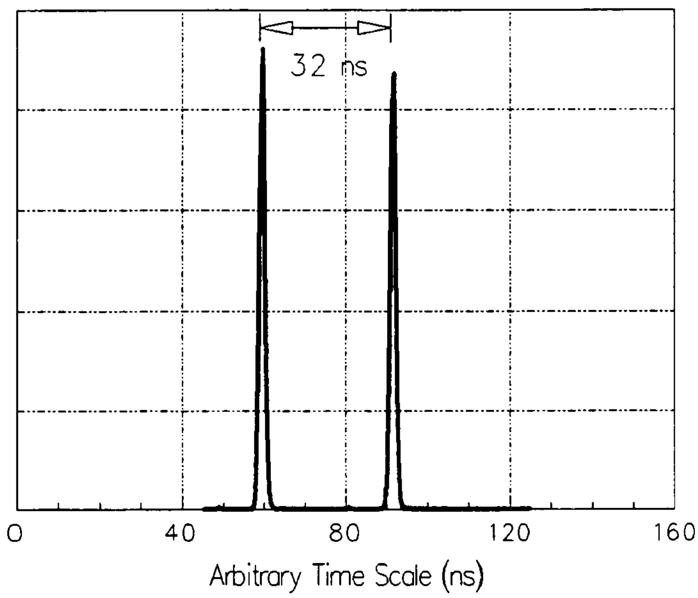

Figure 6-17. Measured timing resolution (1.57 ns FWHM) obtained with a 5 mm diameter API APD, the CMOS preamplifier and a 3x3x22 mm^3 LSO crystal. Vbias = 2700 V, τ_{int}=10 ns and τ_{diff} = 50 ns.

PART 7

CONCLUSIONS

Summarizing Discussion

The objective of this study was to investigate the potential for replacing the photomultiplier tubes currently used in Positron Emission Tomography detectors with silicon avalanche photodiodes. The work described in this report is part of an ongoing CTI solid-state detector project sponsored by the National Institute of Mental Health.

In Chapter 1, the positron emission tomography system was described and the photomultiplier tube based block detector technology currently used in virtually all commercial PET systems was reviewed. The block detector design was shown to have performance advantages normally associated with PMT based detectors (high gain and low noise) as well as the cost and packing fraction advantages of using only four photomultiplier tubes to read the signals from up to 64 BGO pixels. The principal disadvantages of the block detector were shown to be the high cost of photomultiplier tubes (even when only 4 PMTs are used per 64 detector pixels), system dead time associated with the position decoding process and loss of position resolution due to errors in the decoding scheme. A new avalanche photodiode based detector was proposed to address these disadvantages. The basic APD and APD array structures were introduced and the standard expressions for APD gain and noise were presented. A review of the literature was presented and the two most commonly used APD structures, the beveled-edge and reach-through structures, were identified. Finally, the purpose and scope of the project were defined.

In Chapter 2, the fundamental avalanche photodiode theory of operation was reviewed. The classic APD gain and noise expressions first developed by McIntyre were derived and the assumptions used to obtain the commonly used closed form gain and noise expressions were identified. A three step model for the APD frequency response was developed. The APD frequency response was found to be dominated by the carrier transit time across the high-field depletion region, the RC time constant associated with the diode junction capacitance and the external circuitry and, in the case of the beveled-edge APD, the transit time across the undepleted p-region.

In Chapter 3, the beveled-edge and reach-though APD structures were theoretically and experimentally evaluated. Typical fabrication procedures for the two APD structures found in the literature were simulated using the semiconductor fabrication simulator SUPREM. The SUPREM calculated doping profiles were used as inputs for the Poisson equation solver, FISH1D, which calculated the bias dependent electric field of the two APD structures. The field profiles were used to calculate the APD gain and excess noise factor following the theory developed in Chapter 2. The simulations showed the beveled-edge APD to have lower noise than the reach-through APD and thus superior pulse-height and timing resolution. The simulated data was verified experimentally using commercially available reach-through and beveled-edge APDs. The best measured timing resolution (7.6 ns FWHM) and pulse height resolution (14.2% FWHM) were obtained using two different Advanced Photonix beveled-edge devices.

The design requirements for a 64 element APD array suitable for use in a PET detector were presented in Chapter 4. Advanced Photonix was selected to fabricate the array based up on the performance of their discrete devices described in Chapter 3. The Advanced Photonix APD array fabrication procedure was described and an equivalent circuit model for the array was presented. The expected performance characteristics of the array were calculated using the theory developed in Chapter 2. Unfortunately, as of this writing Advanced Photonix has been unable to deliver the APD Array.

In Chapter 5, the design of a 6 channel CMOS preamplifier array suitable for use with APDs was presented and the fabricated circuits were evaluated. The preamplifier topology was believed to be new; it improved the traditional cascode topology by introducing a local feedback loop around the cascode device to significantly improve its input and output impedance. The preamplifier had a measured 7 ns 10-90% rise-time and 40 MHz -3dB high corner frequency with an input capacitance of approximately 6 pf due to the package. Extrapolation of the bandwidth vs. input capacitance indicated the preamplifier had a gain bandwidth product of approximately 600 MHz. Extrapolation of measured output noise spectral density data indicated the preamplifier had an equivalent input noise voltage of approximately 1 nV/√Hz and an input capacitance (including the ceramic DIP package) of 16 pF.

In Chapter 6 the performance characteristics of the CMOS preamplifier and APD were evaluated and the performance of a prototype APD array was discussed. The equivalent input noise charge of the CMOS preamplifier and a CR-RC ^{4} shaping amplifier was found to be approximately 1400 electrons FWHM for a shaping time constant of 2 μs and a detector capacitance of 10 pF. The timing jitter due to the preamplifier noise was approximately 3 ns FWHM for an input test signal modeling to the output of an APD/BGO PET detector biased for an APD gain of 500, a CR-RC shaping time constant of 10 ns and a detector capacitance of 10 pF.

The CMOS preamplifier was incorporated into pulse height and timing resolution test systems using an Advanced Photonix 5 mm diameter APD (C_d = 16 pF) and a 3 x 3 x 3 mm^3 BGO crystal. The measured pulse height resolution (12% FWHM) was believed to be the best PHR obtained to date using a BGO crystal with a suitable geometry for PET (length > 20 mm). The measured timing resolution (9.08 ns FWHM) was virtually identical to the timing resolution obtained with a similar detector coupled to an ORTEC 142A preamplifier. Discussions with the Advanced Photonix indicated the timing performance was limited by carrier trapping at the interface between the silicon substrate and the anti-reflective coating. Recent improvements in the anti-reflective coating process should reduce the timing resolution to approximately 5 ns FWHM. Pulse height and timing resolution data obtained with a recently developed scintillator, cerium doped lutetium oxyorthosilicate (LSO) were also presented and briefly discussed. The measured pulse height resolution (14.2% FWHM) and timing resolution (1.57 ns FWHM) suggested LSO may be a viable replacement for BGO in future APD based PET detectors.

Suggestions for Further Study

The presentation of this dissertation marks the end of the first phase of a solid state detector development effort at CTI. As a result of this study, the theory of operation of APD based detectors has been summarized and simulation models for APD/BGO detectors have been developed. The basic feasibility of using APDs in PET detectors has been demonstrated although the timing performance has been poorer than expected due to imperfections in the APD fabrication technology. The first component of a CMOS readout system, the charge sensitive preamplifier, has been designed and evaluated.

The next phase of the development effort will focus on building a detector module suitable for use in a prototype PET scanner. The principal milestone in this phase will be the development of a reliable APD array, a goal which has proven to be elusive to date. Additionally, the remainder of the readout system including a pulse height (energy) channel and a timing channel must be integrated with the preamplifier array into a single ASIC. This system may be designed in a 1.2 μm process to achieve the required packing density. As a part of this ASIC design effort, improved current mirror topologies will be explored to reduce the output flicker noise of the preamplifier.

An exciting parallel path of study will focus on the use of LSO in future detectors. The higher light output and shorter decay constant of LSO relative to BGO offers the potential for designing APD based detector modules with superior pulse height resolution and substantially superior timing resolution relative to currently used PMT/BGO detectors.

VITA

Michael Joseph Paulus was born in Knoxville, Tennessee on April 12, 1963. He earned the Eagle Scout rank in 1979 and graduated from Bearden High School (Knoxville) in 1981. He attended the University of Tennessee in Knoxville where he was a member of the Sigma Phi Epsilon Social Fraternity and an active duty member of The United States Air Force under the College Senior Engineering Program. He worked as a summer intern with Technology for Energy Corporation and North American Philips Corporation. Upon graduation with a Bachelor of Science Degree in Electrical Engineering in 1985 he attended Officer Training School at Lackland Air Force Base where he was commissioned as a Second Lieutenant. He served for four years in the Wright Aeronautical Laboratories, Wright-Patterson Air Force Base, where he performed basic research on novel gallium arsenide devices. While in the Air Force he earned a Masters of Science Degree in Electrical Engineering from the University of Dayton; his master's thesis is entitled "Investigation of Short Period AlAs/GaAs Superlattices as Conduction Band Barriers."

After completing his Air Force tour in 1989 he joined the research staff of the Georgia Tech Microelectronics Research Institute where he and his co-workers brought on line a new Class 10/Class 100 clean room facility. While a member of the Georgia Tech staff he began his doctoral studies.

In 1991 he joined the research and development staff of CTI, Inc., where he is presently a senior development engineer. At CTI he has designed high-speed analog and digital CMOS integrated circuits and performed basic research on new detector technologies. He is the currently the principal investigator for a Phase 2 Small Business Innovative Research program sponsored by the National Institute of Mental Health. He completed his doctoral studies at the University of Tennessee while at employed by CTI.

In 1989 he married the former Margaret Anne Moy of Ithaca, New York. The Pauluses have two daughters, Emily and Paige.