DELAY-LINE 3D POSITION SENSITIVE RADIATION DETECTION

by

Manhee Jeong

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Nuclear Engineering and Radiological Science)
in the University of Michigan
2012

Doctoral Committee:

Assistant Research Scientist Mark D. Hammig, Co-Chair
Professor David K. Wehe, Co-Chair
Professor Emeritus Ronald F. Fleming
Associate Professor Jamie Dean Phillips

To my family.

ACKNOWLEDGEMENTS

I would first like to thank Dr. Mark D. Hammig for guiding my education over the past five years. I have appreciated his willingness to help when I needed it most, allow me to try new ideas when I had them, and help me grow as an engineer. I expect that I will still be learning from him years from now. Professors David Wehe and Zhong He of the University of Michigan have taught me in courses and provided much technical insight and career guidance over the past few years. Furthermore, I would like to thank the committee members: Professor Ronald Fleming and Professor Jamie Phillips for the careful reviews of my dissertation and insightful suggestions.

Along this line, Peggy Gramer and Caroline Joaquin, also of the University of Michigan, have provided countless hours of assistance with financial issues and navigating Rackham's sea of requirements. The Korea Science and Engineering Foundation Grant (No. 2005-215-D00399), the Domestic Nuclear Detection Office (DNDO) of the U.S. Dept. of Homeland Security and the Defense Threat Reduction Agency (DTRA) of the U.S. Dept of Defense have provided me great flexibility in pursuing my education via a Graduate Research Fellowship and Scholarship. I am appreciative of their support and their willingness to continue to fund me.

The members of my research group have always been very supportive and helpful: Paul Barton, Geehyun Kim, Taehoon Kang and Inyong Kwon. Thanks for years of help on the Shockley-Ramo Theorem, Geant4, LabVIEW, TCAD, Sentaurus, and just being good friends. Jooseuk Kim, Gaemin Lee and Sungjin Hwang have provided countless instances of help with topics such as math and computer programming, in addition to great conversations and good friendships. My experience at Michigan would not have been half as enjoyable without them.

Finally, I could not have completed this dissertation and the years of education leading up

to its culmination without the unconditional support of my family: my parents and my sister. Above all I would like to thank Jeongmin and my children, Ethan and Ian, who have been by my side every step of the way for more than a decade showing her love and support.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

LIST OF APPENDICES

Appendix

CHAPTER I

INTRODUCTION

Over the past several decades, semiconductor radiation detectors have been studied due to their fast response and their crystalline properties, possessing a band structure that is neither totally conducting nor insulating. Silicon (Si) detectors can be operated at room temperature, but their low atomic number and high leakage currents have limited their application to, principally, low-energy x-ray and charged-particle detection [1]. However, high-resistivity Si in large volumes and with good charge carrier transport properties has been produced and achieved success as a radiation detector material over the past few years due to its relatively low cost as well as the availability of well-established processing technologies. One application of that technology is in the fabrication of various position-sensing topologies from which the incident radiation's direction can be determined. The resulting image can be enhanced if the information content of the interaction is maximized. In particular, if one can determine the direction of the recoil electron or ion, then the event reconstruction can be well-informed. The determination of that direction depends on an accurate measurement of the *initial* positional distribution of the electron-hole (e-h) products of the charge passage. Although extrinsic effects such as electronic noise can impact the measurement, that determination is ultimately limited by the stochastic and nonlinear effects that muddle the electron-hole transport and consequently, the induced electrode signals [2]. The goal of the research is to image the transit of individual electrons and holes, the realization of which will provide benefits in two ways. First, if we can sense, with high precision, the location at which the carrier is collected, then we can study the non-uniformity of the junction formation as a function of position. Typically, some gross parameter is used to characterize the junctions formed on a detector's surface, but detailed studies have shown that the junction formation is highly non-uniform, a fact which complicates the fabrication of robust contacts [3]. More

topically, if one can image individual electron-hole pairs that are created via the passage of radiation through solids, then one can greatly enhance the directional resolution of semiconductor detectors. From either the dynamics of the induced charge creation, or from the charge collection following drift, one can deploy techniques in which the charge cloud formation indicates the direction of the incident radiation. For instance, if the charge cloud can be used to indicate the direction of the Compton electron, then the back-projected cone can be reduced to an arc-segment, enhancing image resolution. The resolution loss that accompanies reduced count rates in coded-aperture imaging can also be avoided if the individual photoelectrons can, in effect, be imaged. In order to optimize the design of the detector capable of charge-cloud mapping, we: (a) numerically modeled the formation of the charge cloud, (b) placed that charge distribution in the electric field environment of the detector and derived the induced current signals on the bounding electrodes by tracking the diffusion and drift of the carriers through the detector bulk, (c) fabricated meander line detectors which have low leakage current and high break down voltage, and (d) acquired energy spectrum and position information with the amplitude variation and time difference methods

1. 1 History of the Room-temperature Semiconductor Detectors

Since the advent of applying planar technology to semiconductor radiation detector manufacturing [4,5] numerous detector designs have been invented and developed for a wide range of applications both in particle and photon detection. The possibility to process micro-structured silicon diode detectors was first utilized in vertex detection in high energy physics experiments [6,7]. Vertex detectors consisting of single sided and later double sided silicon strip detectors arranged in a cylindrical shape around the beam collision point provided position information of the passing ionizing particles with a resolution better than 10 mm [8,9]. The success of silicon detectors in position sensitive particle detection in accelerator experiments soon lead to their application in other fields of research, such as medicine and industry. Strip detectors as such were applied to autoradiography [10,11] and tested for x-ray imaging [12,13]. Silicon drift chambers and depleted MOSFET detectors were developed for energy sensitive detection in x-ray spectroscopy [14,15], and a fully depleted CCD x-ray imaging detector has been applied in astrophysics [16]. Pixel detectors, which have the advantage over strip detectors of

unambiguous position information and higher signal to noise ratios, were proposed and have been developed for both particle tracking and x-ray imaging [17,18]. Since digital x-ray imaging offers several advantages over conventional analog film-based imaging, an immense effort has been invested in research and development of digital x-ray imagers both in the fields of medicine and industry. Commercial state of the art large area systems generally rely on flat panel technology and apply either indirect conversion of x-ray energy with a scintillator [19,20] or direct conversion with amorphous selenium [21]. Small field detectors which may be combined for scanning configurations are mainly based on optically coupled CCDs [22], but directly-converting semiconductor systems, CMOS and bump bonding technology have also been developed [23,24]. Directly converting systems yield higher spatial resolution compared to indirect systems [25]. Crystalline semiconductors combined with bump bonding and CMOS electronics allow smaller pixel sizes than flat panel technology but large continuous image areas are impossible with current CMOS technology. The efficient detection of higher energy x-rays and gamma-rays benefits from a high atomic number (Z) conversion material. Studies on GaAs [26] and HgI_2 [27] have been presented but the most promising high Z materials to date are CdTe and CdZnTe which have been proposed to be used in medical gamma imaging instead of scintillators in order to improve both the spatial and energy resolution of gamma cameras [28]. High absorption efficiency offers advantages also in real time and scanning imaging or generally in any application where the image acquisition speed or radiation dose is critical.

1.2 Semiconductor Physics for Radiation Detection

1.2.1 Junctions in semiconductor: PN diode

Semiconductor radiation detectors are based on the favorable properties that are created near p-n junctions. In Fig. 1.1a and 1.1b the p- and n- type semiconductors are shown forming a junction between them and the positions of the Fermi level, the conduction and valance bands, and vacuum energy level. The electron affinity $e\chi$, defined as the energy difference between the conduction band and vacuum level, is also shown along with the work function ($e\phi_{sp}$ or $e\phi_{sn}$). The work function represents the energy required to remove an electron from the semiconductor to the free vacuum level and is the difference between the vacuum level and the Fermi level. In the absence of any

applied bias, there is no current in the system and in the absence of any current flow, the Fermi level is uniform throughout the structure. In the depletion region, which extends a distance W_p in the p-region and a distance W_n in the n-region, an electric field exists. Any electrons or holes in the depletion region are swept away by this field. Thus a drift current exists that counterbalances the diffusion current, which arises because of the difference in electron and hole densities across the junction.

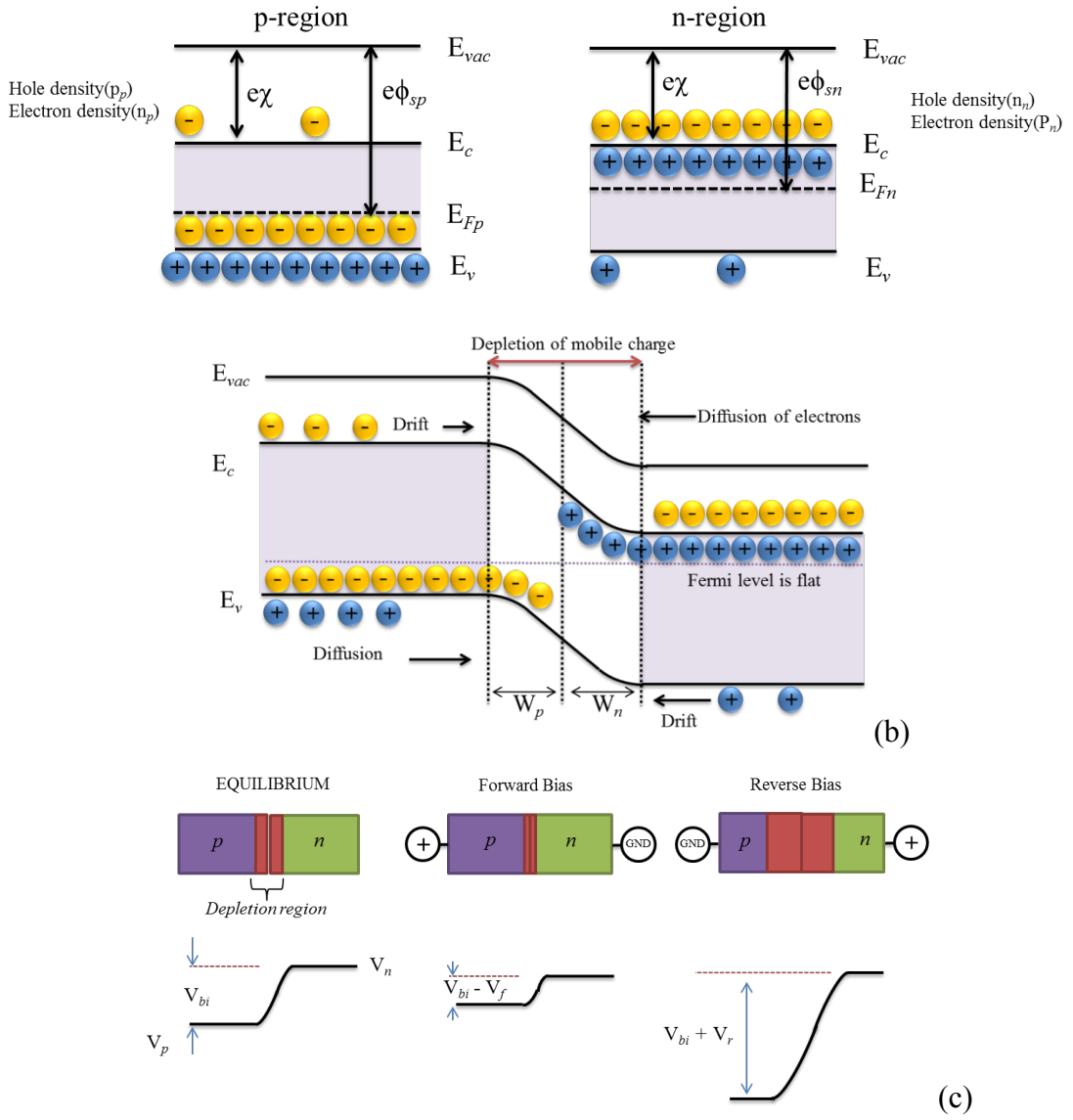


Fig. 1.1: (a) The p- and n type regions before junction formation. (b) A schematic of the junction and the band profile showing the vacuum level and the semiconductor bands. The Fermi level is flat in the absence of current flow. (c) The biasing of the p-n diode in equilibrium, forward and reverse bias cases and their voltage profiles.

All the current components flowing in the detector are composed of electron drift and diffusion current as well as the hole drift and diffusion currents. When there is no applied bias, these current cancel each other individually. The hole current density is

$$J_p(x) = e\left[\mu_p p(x)E(x) - D_p\frac{dp(x)}{dx}\right] \tag{1.1}$$

Note that the mobility μ_p is actually field-dependent and the drift current density in the depletion region is really $ep(x)v_s(x)$ and independent of E (electric field) at high fields, where p(x) is hole distribution, v_s is the saturated velocity. This is because the field in the depletion region is very large even under equilibrium. Similarly, the diffusion coefficient is not constant but is field-dependent. The ratio of μ_p and D_p is given by the Einstein relation $\mu=(e/kT)D$ where k is the Boltzmann constant, T is absolute temperature and e is electric charge (1.6×10^{-19}C). A built-in voltage, V_{bi}, is produced between the n and p side of the structure. The built-in voltage is given by

$$eV_{bi} = E_g - (E_c - E_F)_n - (E_F - E_V)_p \tag{1.2}$$

where the subscript n and p refer to the n-side and p-side of the detector. The built-in voltage depends logarithmically on doping level,

$$V_{bi} = \frac{kT}{e}ln\left(\frac{N_aN_d}{n_i^2}\right), \tag{1.3}$$

where N_a and N_d are the acceptor and donor concentrations and n_i is the intrinsic carrier concentration. Fig. 1.1c shows the biasing of a p-n diode in equilibrium, forward and reverse bias cases and their voltage profiles. When an external potential is applied, thermal equilibrium no longer holds. With positive bias applied to the p-region and the negative bias to the n-region, the potential barrier is reduced and the flow of electron and holes across the junction increased. When the opposite polarity is applied, reverse biased, the potential barrier is increased and the width of the depletion grows. The pn-junction is asymmetric with the respect to current flow. The dependence of detector current versus voltage is given by the "Shockley equation"

$$I = I_0(e^{\frac{eV}{kT}} - 1) \quad (1.4)$$

where $I_0 = eA\left(\frac{D_p p_n}{L_p} + \frac{D_n n_p}{L_n}\right)$, A is the detector area, D_i is the diffusion coefficient and L_i is the diffusion length. Under forward bias the current increases exponentially with the applied bias. Under reverse bias, the current simply tends towards the value $-I_0$. Since the depletion region is a volume devoid of mobile carriers, it forms a capacitor, in which the undepleted p- and n-regions are the electrodes and the depletion region is the dielectric. The electric field in the depletion region will sweep mobile carriers to the electrodes, so the detector forms an ionization chamber. The depletion region formed by thermal diffusion is rather thin and the width of the depletion region can be increased by applying a reverse bias voltage.

1.2.2 Junctions in semiconductor: Schottky barrier diode

Metals play three important roles in electronic circuits (1) as interconnects, they provide the pathway to pass electrical signals to and from a device, (2) as Schottky barriers, they provide junctions that can provide rectification, possessing a built-in electric field, and (3) as ohmic contacts, they allow electrons or holes to enter and leave the semiconductors with little resistivity. The metal-semiconductor junction can be an active device capable of a strongly nonlinear response. The Schottky barrier diode has characteristics that are essentially similar to those of the p-n diode except that for many applications it has a much faster response.

In order to clarify the properties of the Schottky barrier junction, a metal semiconductor structure is shown in Fig. 1.2, which shows the band profiles of a metal and semiconductor. Fig. 1.2a shows the band profile and Fermi level positions when the metal and semiconductor are separated. When brought into contact, the energy-level diagram changes to that shown in Fig. 1.2b. Due to the built-in potential at the junction, a depletion region of width W is created.

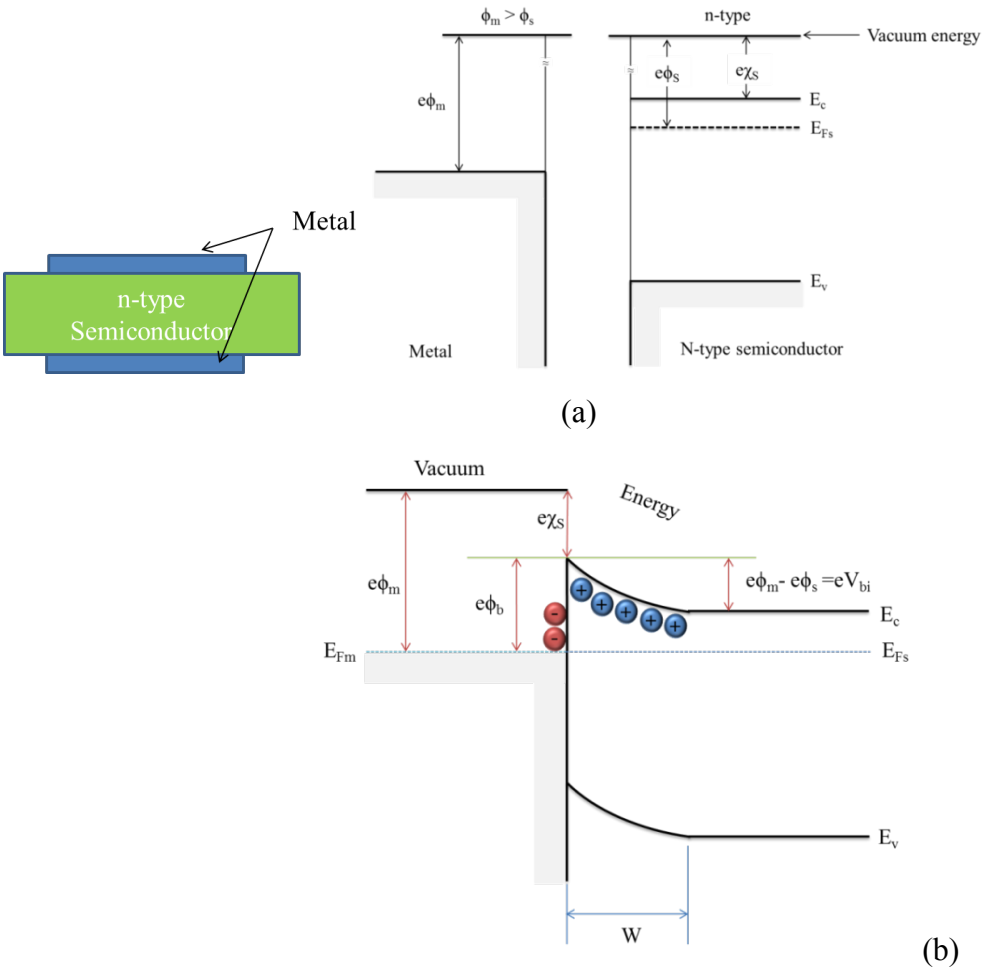


Fig. 1.2: (a) A schematic of a Schottky barrier junction and the various energy levels in the metal and the semiconductor with respect to the vacuum level. (b) The junction potential produced when the metal and semiconductor are brought together.

1.2.3 Signal charge generation

The charge pairs are formed in the sensitive detector volume and the average signal charge is expressed as $Qs = (E/E_i)\cdot e$, where E is the absorbed energy, E_i the energy required to form a charge pair, and e the electronic charge. In the detector the absorbed energy must exceed the bandgap to form mobile charge carriers. Specifically, the gap energy of Si is 1.12eV, so particles with greater energy can be detected. The ionization energy E_i is proportional to the bandgap, so higher bandgap materials yield less signal charge. In semiconductor detectors the electrical signal is formed directly by ionization. Incident quanta impart sufficient energy to individual atomic electrons to form electron-

hole (e-h) pairs. For a charged particle track traversing the detector, the energy loss E (or the integrated signal charge Q_s) will increase with detector thickness. Minimum ionizing particles average about 80 e-h pairs per μm path length in Si. The detector should be sufficiently thick to provide good detection efficiency for the Compton events which can have multiple sequential interactions. The signal is formed when the liberated charge carrier move, which changes the induced charge on the detector electrodes.

1.2.4 Charge collection

Carrier transport can proceed through diffusion or drift. Thermal energy causes carriers to move in a random direction, but in the presence of a concentration gradient they collide more frequently in the direction of higher concentration, so the net motion is in the opposite direction. The concentration profile spreads out with time forming a Gaussian distribution with the variance $\sigma = \sqrt{Dt}$, where D is a material dependent diffusion constant. In the presence of an electric field, the carriers move parallel to the electric field, which is called drift. A potential is applied between the electrodes to accelerate the charge carriers. After a short equilibration time carrier transport becomes non-ballistic and the velocity does not depend on the duration of acceleration, but only on the magnitude of the local electric field. Thus, the velocity of the carriers at position x depends only on the local electric field E(x) and the carrier velocity become $\vec{v}(x) = \mu\vec{E}(x)$, where μ is the mobility. The mobility is related to the diffusion constant through the Einstein relation $\mu = (e/kT)D$ where k is the Boltzmann constant, and T is the absolute temperature. At 293 K, the numerical value of kT/e is 0.0253V. In silicon, electrons and holes have different motilities, 1350 and 480cm^2/Vs at 300K, respectively [1]. For comparison, the thermal velocity of an electron in Si at room-temperature is about 10^7cm/s, so the carrier motion is the superposition of a substantial random thermal motion and the drift due to the electric field.

1.2.5 Charge carriers migration in electric field vs. position resolution

Many semiconductor detectors are operated with electric field values sufficiently high to result in saturated drift velocity for the charge carriers. If the electric field is high enough to reach the saturation velocity, the time required to collect the carriers over

typical wafer thickness of 0.5 mm will be under 5 ns; thus, semiconductor detectors can be fast-responding devices. Also, the position resolution of the detector is determined by the electrode geometry. Detectors for tracking applications have strip or pixel type electrodes. Despite the gaps between the electrodes, the detectors can still have 100% efficiency. The electrical segmentation is determined by the electrode pitch, rather than the width. Because the response function of a detector is essentially box-like, the position resolution of single detector is equal to the strip pitch, p; however, the differences between the measured and true positions have a Gaussian distribution with the standard deviation $\sigma^2 = \int_{-p/2}^{p/2} \frac{x^2}{p} dx = \frac{p^2}{12}$ [29], so the root mean square (RMS) resolution is the strip pitch divided by $\sqrt{12}$. In addition to the carriers' drift, they will undergo the influence of diffusion. Without diffusion, all charge carriers will drift to the collecting electrodes following the electric field lines. However, in reality, the charge carriers are subject to thermal diffusion, which spreads the charge cloud transversely as the charges drift. Thus, the RMS width can be expressed as: $\sigma = \sqrt{2Dt} \approx \sqrt{\frac{2kT}{e}\frac{d^2}{V_b}}$, where d is the detector thickness and V_b is the bias voltage. Typically, for the 500μm thick detector with 100V bias voltage at 293K, the traverse diffusion will be $\sigma \approx 11.24$μm [30].

1.3 Principle of delay-line Position Sensitive Radiation Detectors

1.3.1 Lateral Position Sensing based principally on Amplitude Variations

As will be explained in more detail below, Fig. 1.3 shows a schematic of the current signal derived from a semiconductor detector upon the creation of a *single* radiation-induced charge carrier. The initial abrupt rise is due to the finite time-of-creation and subsequent acceleration of the carrier, the extent of which is generally neglected during radiation measurements- and for good reason as that time is typically less than a picosecond compared with typical timing resolutions on the order of a nanosecond or longer. Furthermore, most readout systems are designed to survey the charge dynamics via the *integrated* current signal- that is, via a voltage waveform- for which the sharp rise is less apparent. Nevertheless, fast timing can enable fine position sensing both laterally and in depth, and we are therefore seeking to optimize the time

resolutions of our detectors and amplifying components. It is therefore useful to consider those temporal variations that are normally neglected.

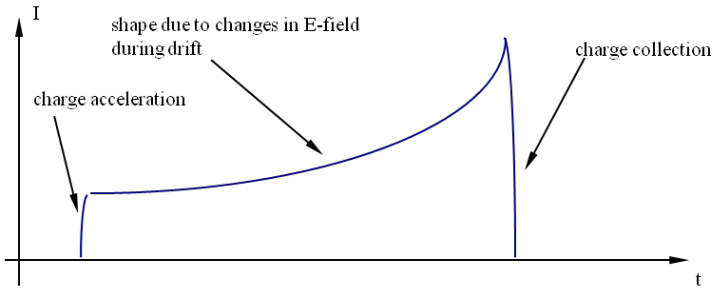


Fig. 1.3: Schematic of current pulse formed on the electrode of a semiconductor detector by a charge carrier.

Specifically, one can employ Newtonian mechanics and electrodynamics to gauge the charge's acceleration time after creation in the electric field (E) inside a semiconductor detector. Equation 1.5 shows time, t_{sat}, required for the carrier of effective mass m^* and charge q to reach the drift velocity in silicon (v_{drift}).

$$t_{sat} = \frac{m^* v_{drift}}{qE} \tag{1.5}$$

For a saturated field strength of 10^7 V/m [29] (at which the charges reach the saturated drift velocity of ~10^5 m/s), t_{sat} is equal to 0.067 ps for electrons and 0.046 ps for holes, miniscule values that justify the exclusion of charge acceleration in the consideration of the charge drift. However, if smaller field strengths are applied, then longer acceleration times can result. For instance, in order to safely avoid electrical discharge phenomena in the air, we typically apply 500 volts or less to our silicon detectors of thickness 550 μm, resulting in a electric field of ~10^6 V/m; that is, about an order of magnitude less than the saturated drift field. At that voltage, t_{sat} for the electrons increases to 0.92 ps and that for the holes increases to 0.21 ps. Thus, if one employs a detector with timing resolution on the order of a picosecond or less, then the effect of charge acceleration may play a role in the position determination; however, we expect that such fast timing is not required to achieve the goal of recoil-electron tracking.

Similarly, the "charge collection time" at the end of the pulse in Fig. 1.3 is also typically ignored. That period corresponds to the time during which the electric field changes abruptly from its drift value in the semiconductor to its (nominal) negation in the electrode conductor, a time that depends on the material details of the interface. For typical drift speeds and skin depths, that time is also less than a picosecond. The position-information content of the current signal is therefore typically found in both the shape and temporal extent of the waveform during drift. Recall that our goal is to examine the signal and determine the distribution of the radiation-induced charge immediately following the interaction of the quantum. For instance, in a planar or strip detector, the duration of the Fig. 1.3 current pulse can be used to gauge the depth at which the charge originated, and the shape of the signal can identify its lateral position relative to a strip.

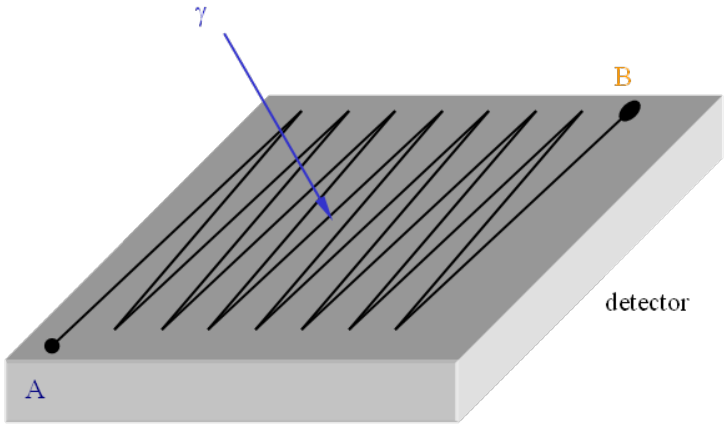
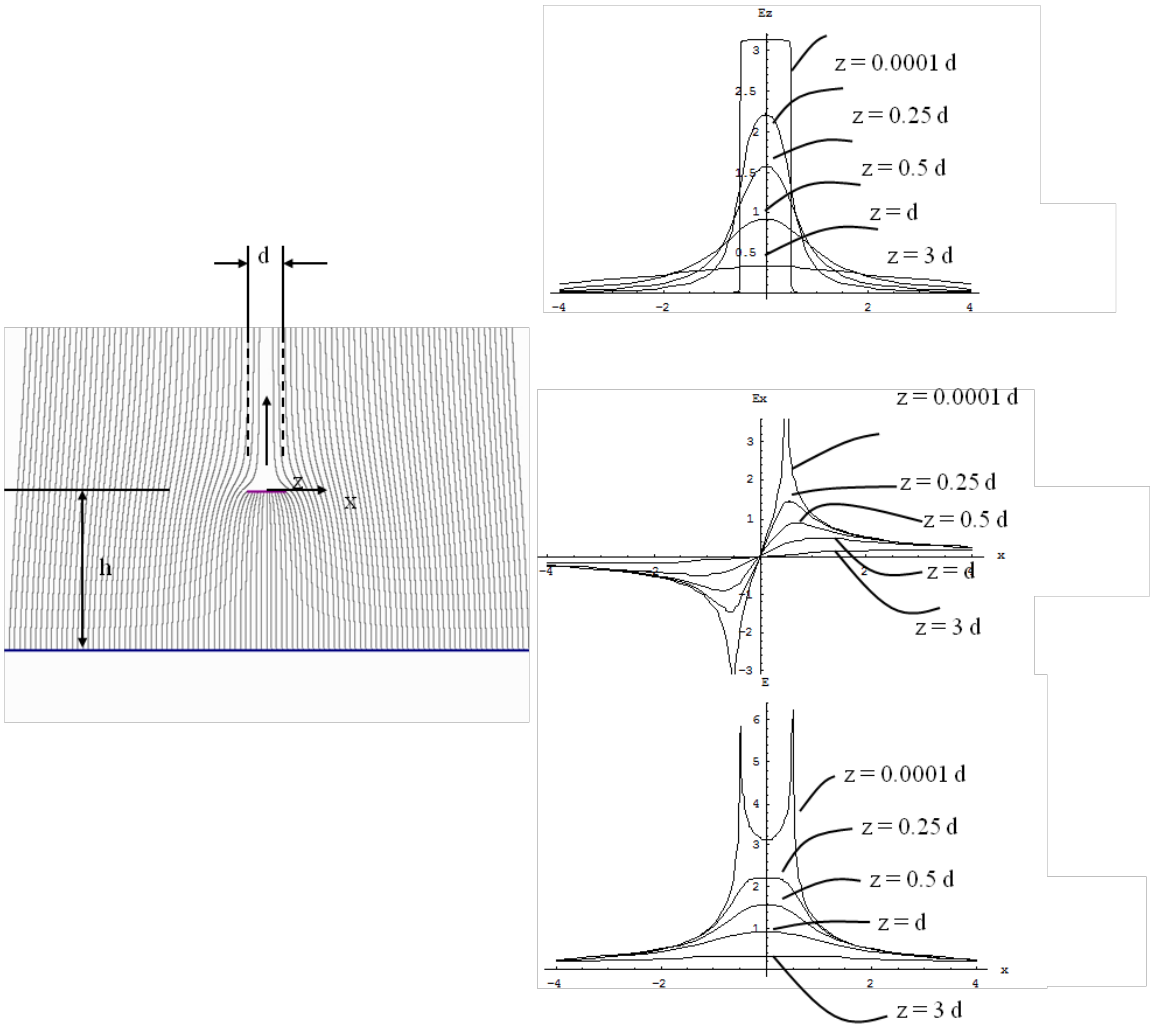


Fig. 1.4: Schematic of semiconductor detector on which a non-uniform electrode is patterned on, at least, one of its surfaces.

Specifically, consider the geometry in Fig. 1.4 and recall that we are initially considering strip geometries because we intend to assess the degree to which one can readout the radiation's position (and energy) information with a minimum number of electrical readouts, one means of achieving such is to vary the electrical field across the detector so that each position provides a unique waveform signature. In this section, we will discuss the means by which we are assessing the relationship between the electrical signal and the charge's initial position.



(a) (b)

Fig. 1.5: (a) Electric field lines for a single strip (purple) above an infinite plane (blue), and (b) the lateral electric field (E_x), the vertical electric field (E_z) and the electric field magnitude (E) as a function of position both laterally and vertically.

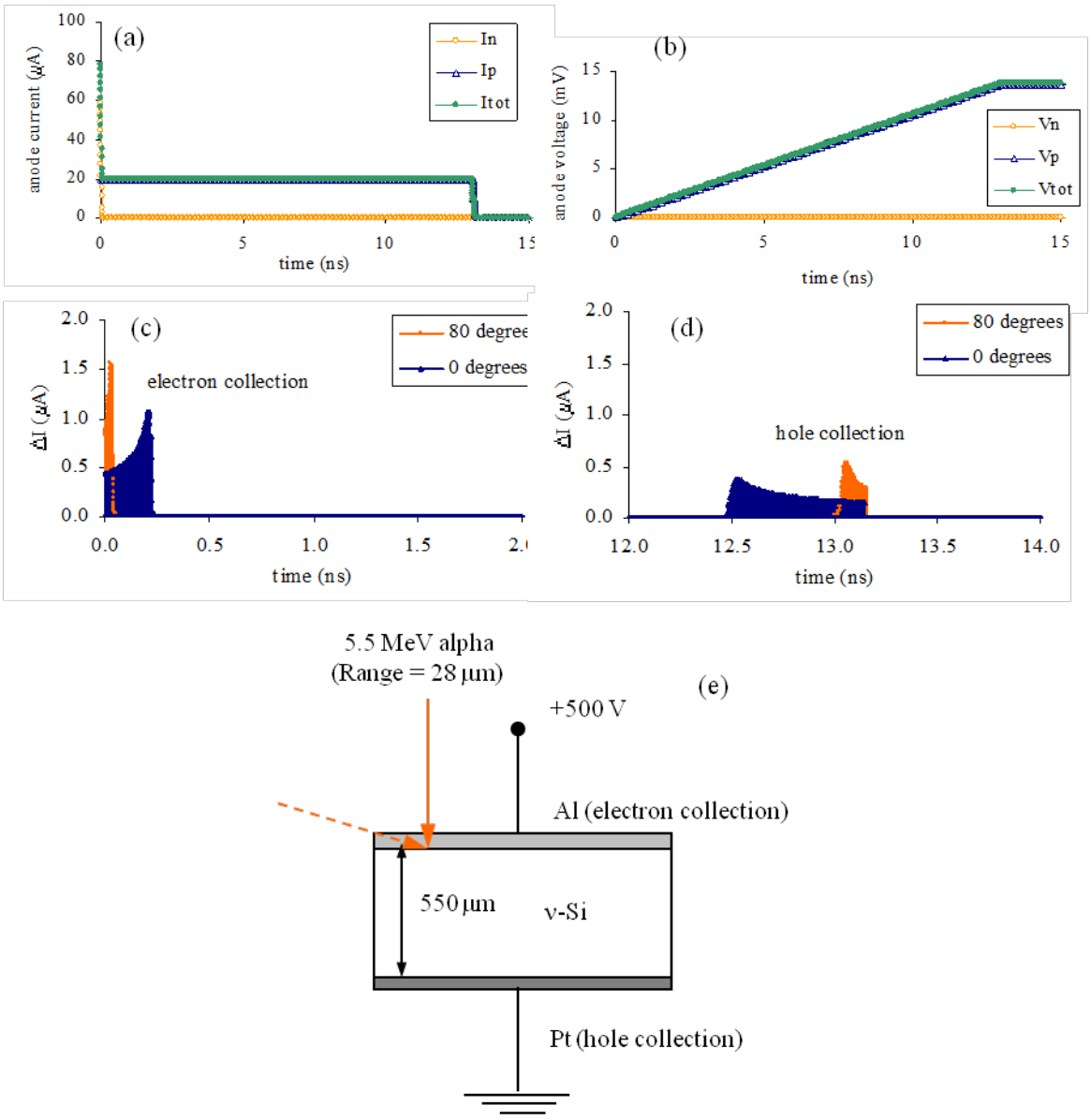


Fig. 1.6: (a) Current induced on the anode following the alpha interaction event shown in (e). (b) The resulting voltage from integral of charge on the capacitance of the detector. (c, d) The charge distributions in depth as measured during electron collection (c) and hole collection (d).

By employing the equivalent of Coulomb's law for continuous charge distributions [31], we can analytically calculate the electric field within a region when bounded by electrical contacts of various geometric designs. For instance, the result for a single strip is shown in Eq. 1.6 and plotted in Fig. 1.5, using Mathematica and a custom C code.

$$\vec{E}(x,z) = \frac{V}{\left[\pi h + 2\left\{h\tan^{-1}\frac{d}{2h} + \frac{d}{4}\ln\left(1+\left(\frac{2h}{d}\right)^2\right)\right\}\right]}\left[\begin{array}{l}\hat{i}\ln\left[\frac{(x+d/2)^2+z^2}{(x-d/2)^2+z^2}\right] \\ +\hat{j}\left\{\tan^{-1}\left(\frac{x+d/2}{z}\right)-\tan^{-1}\left(\frac{x+d/2}{z}\right)-\pi\right\}\end{array}\right] \tag{1.6}$$

Once the electric field is known, then one can calculate the current or charge induced on the conductors in the environment using the Shockley-Ramo Theorem [4]. That is, the instantaneous current induced on a given electrode is given by,

$$i = q\vec{v} \bullet \vec{E}_0 \tag{1.7}$$

where q is the carrier's charge, v is it's instantaneous velocity, and E_o is the weighting field. The current signal can therefore be used to extract the energy and positional information.

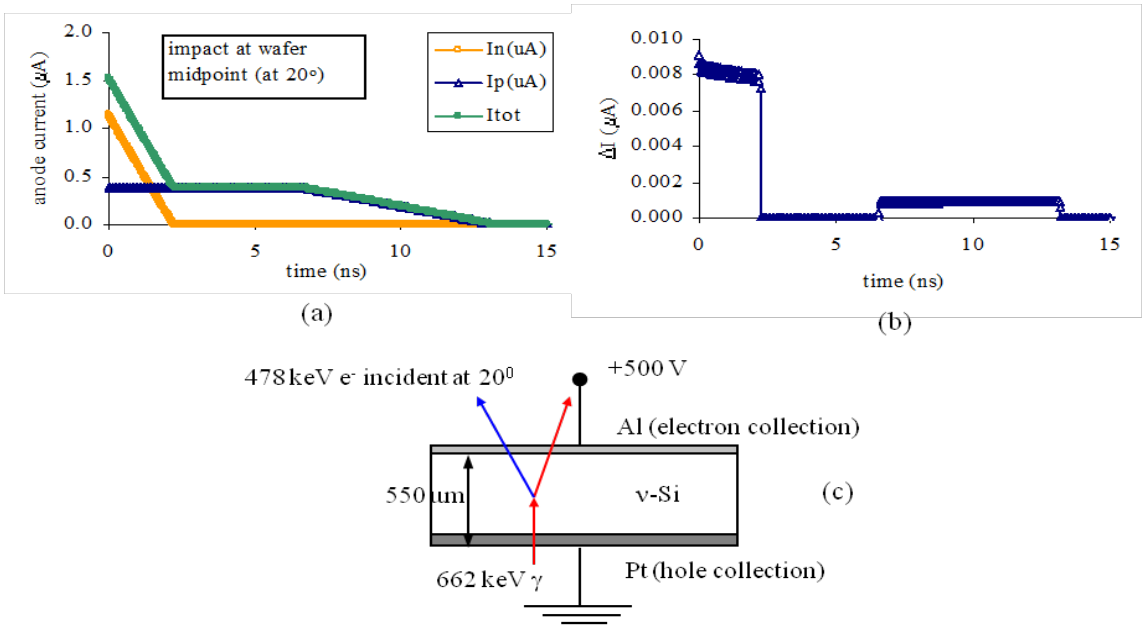


Fig. 1.7: (a) Current induced on the anode following the alpha interaction event shown in (c). (b) The charge distributions in depth (via the time).

Specifically, using the planar detector in Fig. 1.6e as an example, the anode current and voltage are shown in Figs. 1.6a and 1.6b for an 5.485 MeV alpha particle (from ^{241}Am) incident on the anode at two different angles of incidence. Because the electric field is uniform, the current signal is characterized by a constant value during the charge-cloud drift and punctuated by a rapid variation during charge collection at either bounding electrode. As shown in Figs. 1.6c and 1.6d, the change in the current signal provides a direct measure of the initial charge distribution as well as its position (in depth). The current signal can also be used to extract the position information for interactions; specifically, those resulting from gamma-rays and neutrons. An example is shown in Fig. 1.7, which shows the current and charge-distribution signals for a 662 keV gamma-ray scattering at the midpoint of the silicon detector.

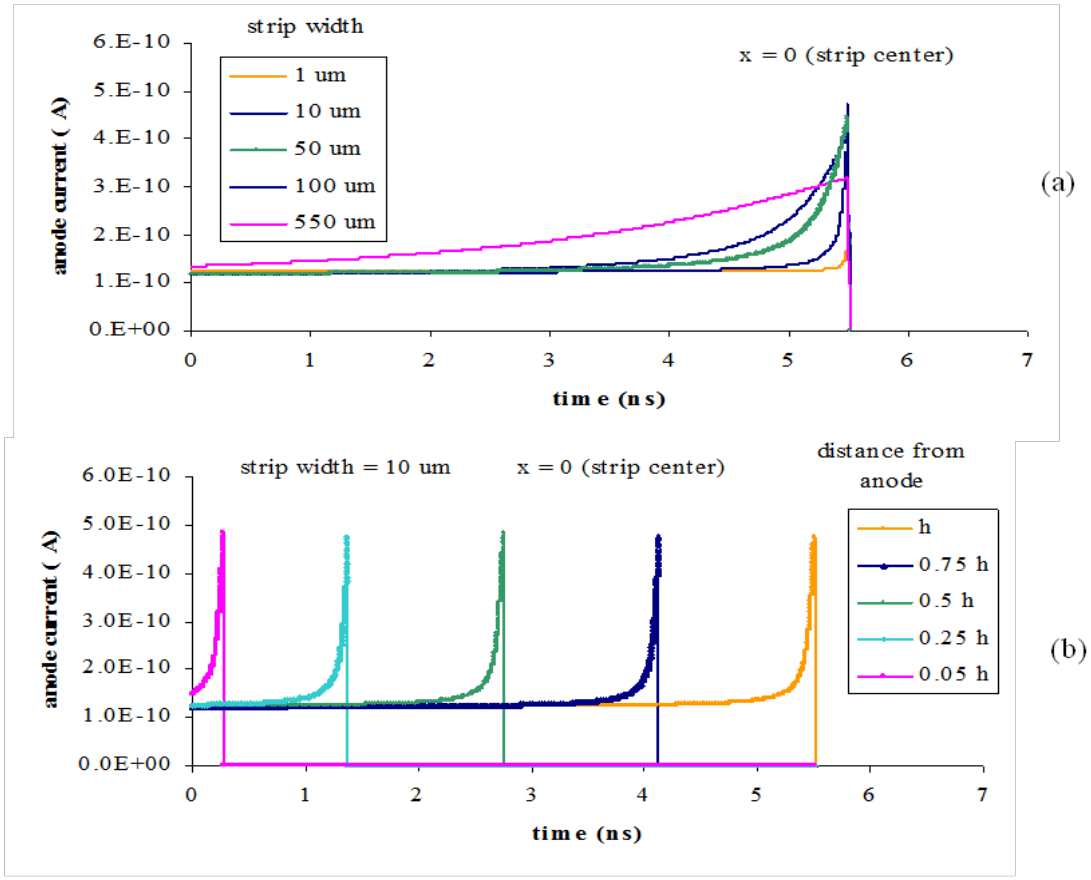


Fig. 1.8. (a) Variation in the single-carrier response as the width of the strip is varied. Note that the initial position of the carrier is assumed to be at the cathode and at the strip center. (b) Variation in the response for a single carrier starting directly under the strip as its initial depth is varied.

We examined comparable variations for a partitioned electrode with the goal of finding the electrode pattern that optimizes the position determination. In order to understand the gross variation in the signal as the pattern is varied, consider the single strip in Fig. 1.8 (and Fig. 1.5a) and the resulting current characteristics in Figs. 1.8a and 1.8b for a single charge carrier. As the strip width is decreased (Fig. 1.8a), the field becomes increasingly uniform in the detector bulk- as reflected in the flat response during drift- and near the surface the sharpness of the response becomes more pronounced due to the enhanced concentration in the electric field, a small-pixel effect which is used to good effect in single-carrier devices (such as CZT). As shown in Fig. 1.8c, the duration of the signal (along with its shape, in general) reflects the depth at which the interaction occurs as in the planar case; however, there are slight variations in the temporal width of the signal as the initial charge location is varied laterally.

For instance, Fig. 1.9 shows the manner in which the electron's current characteristic varies as its initial position is varied laterally. If a single electron is injected near the cathode (via a laser pulse for instance), then the shape and duration of the current-pulse induced during drift will vary, although the variation in the pulse-duration is less pronounced than that revealed in the pulse-shape, as shown in Fig. 1.9(c). Note that the two pulses corresponding to the charge starting at $x = 1$ d and $x = 1.25$ d reflect that the charge skims by the electrode without being collected. In a real detector, these pulses will typically terminate when the charge encounters traps at the surface, but they still reveal the initial position of the charge carrier. In fact, the current characteristic provides a direct measure of the trajectory of the charge-carrier during its motion if the electric field is known. Since that trajectory can be used to determine the initial position of the carrier, the initial charge distribution can thus be mapped, as desired.

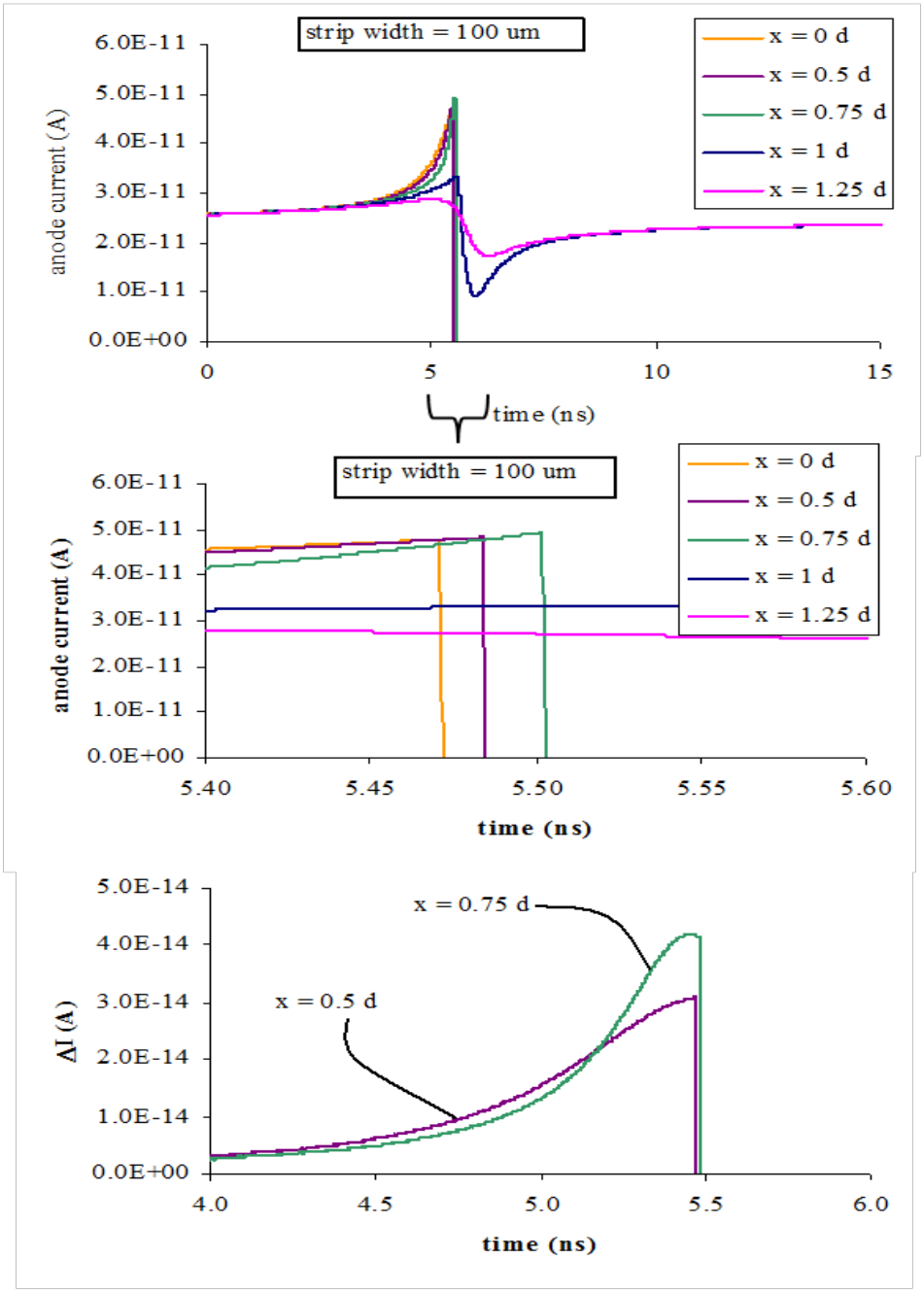


Fig. 1.9: (a, b) Variation in the single-carrier response as the charge's initial lateral position is varied. Note that the initial depth of the carrier is assumed to be at the cathode. (c) ΔI characteristic functions for two lateral positions for a strip width of 10 μm (note different than that assumed in a and b).

This procedure is complicated, of course, when multiple charge carriers all contribute to the signal. Nevertheless, one can extract the charge distribution from the current signal by

again using the change in the current, although for non-planar geometries, the charge is not characterized by a delta-function in the current curve but instead by a characteristic function depending on its trajectory, two examples of which are shown in Fig. 1.9c. Thus, during the charge-cloud reconstruction, one must deconvolve the individual charges (or batches of charges) using characteristic functions like those shown in Fig. 1.9c. Using these general principles, we therefore began the examination of various electrode partitioning geometries and charge states using both analytical and numerical techniques.

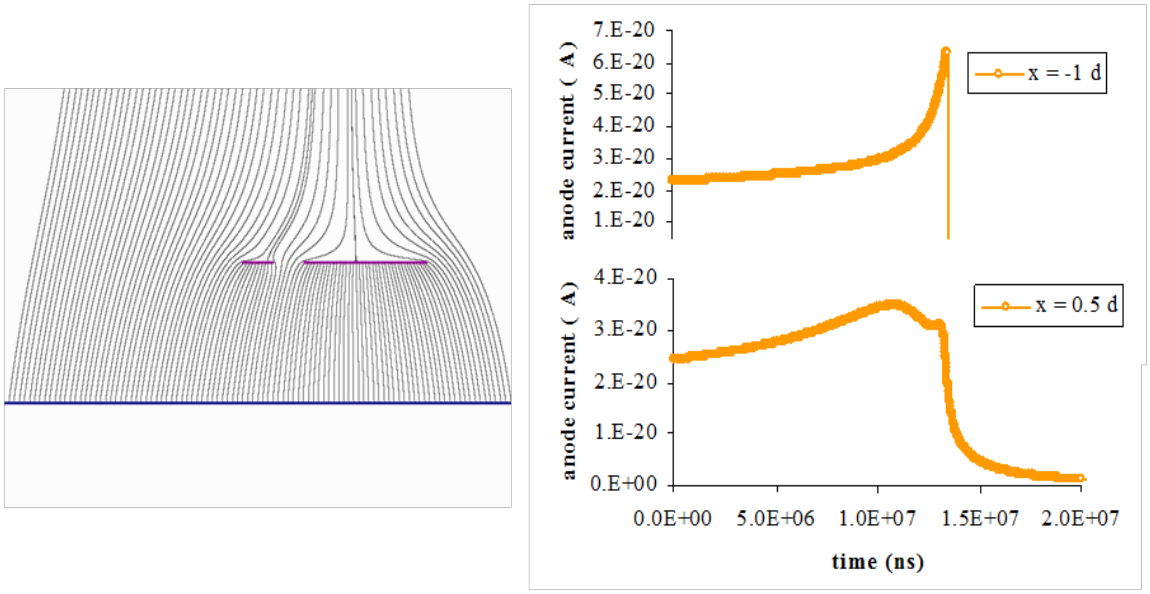


Fig. 1.10: E-field pattern for two strip pattern- noting that the width of the second strip is 4 times the width of the first strip, and the gap is equal to the width of the first strip- as well as some representative single-carrier pulse shapes. Note that the origin for x is at the midpoint of the smaller strip with width d.

For instance, Fig. 1.10 shows a two electrode partition and some representative current pulses for the case in which the strips have the same low charge density. That is, note that the current pulses in Fig. 1.10 correspond to only 1 V applied at the anode plane, the result of which is that the drift velocity is small, the effect of which is to significantly reduce the pulse magnitude and increase its duration. In fact, from a detector design standpoint, it is important to rapidly drift the carriers to the contact not only for

timing optimization but also to raise the pulse amplitude and the S/N ratio. Roughly speaking, only about 70 V is required to fully deplete our 0.5 mm thick silicon detectors, but to saturate the drift velocities- and in particular, the hole velocity- another order of magnitude of voltage is required, an extra velocity that can be important when one is striving to maximize the carrier sensitivity.

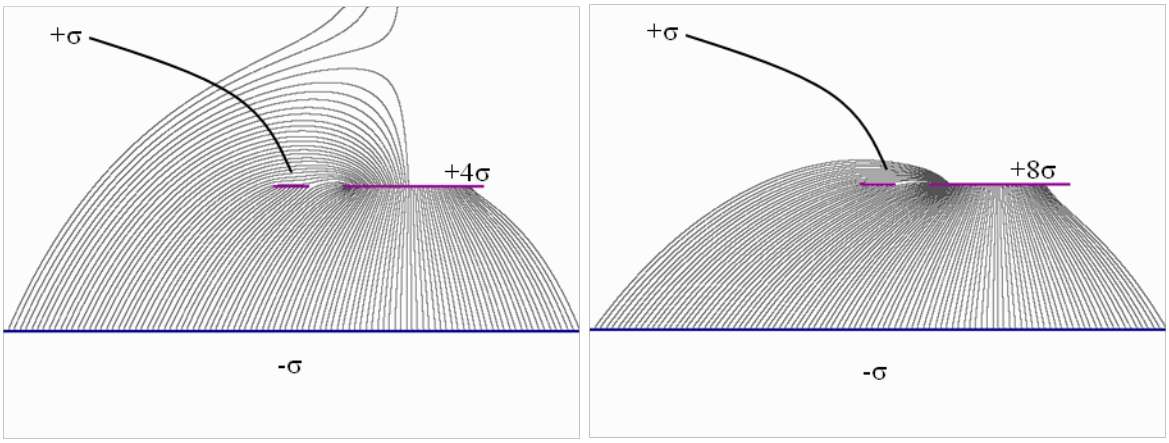


Fig. 1.11: Example of E-field patterns for two-strip electrodes.

Fig. 1.11 provides other examples of different geometries and charge states that were examined. The preliminary conclusion from our analytical efforts is that one has limited leeway in significantly increasing the detector's position resolution by altering the geometry of the bounding electrodes alone. The underlying physical equation- particularly Poisson's equation- averages the scalar potential and smoothes the electric field. Thus, if a carrier is relatively distant from the regions of charge concentration, the electric-field variation is weak from point-to-point. One would prefer to have a precipitous variation in the current characteristic as the position is varied, but this can only be achieved if charge-concentrations are distributed in depth as well as laterally. One can certainly envision fabrication techniques to realize such a design, but we focused on alternative methods that, if successful, will allow us to achieve micron-scale position resolution in a fast simple detector.

Before discussing these timing-based methods, first note that amplitude-based techniques can alone be used to determine the direction of the recoil electron, particularly in the case of gamma-ray interactions, for which the range of the resulting product-electron can be 100's of microns. For neutron detectors which convert the incident neutron into ions however, one would prefer finer resolution. Our numerical examinations, discussed in Chapter II, are being developed in order to provide accurate predictions of the pulses which come from ions, gamma-rays, and various neutron interactions.

1.3.2 Lateral Position Sensing based principally on Timing

The most precise assessment of the nature of the incident quantum, in terms of its energy, position-of-interaction, and incident direction, will use all of the information content in the signal; that is, it will exploit both the timing and amplitude information. Nevertheless, most methods lean on either one or the other and then use the complementary measure to refine the measurement. We plan on taking this approach and in this section; we will discuss a method of achieving fine lateral position information based principally on timing. In Section 1.3.1, we discussed fast timing sequences- the acceleration-time of the carrier, for instance- that are typically ignored because of the system's limited timing resolution. Another such effect is the transit time of the electrical pulse information along the contacts, which is generally small because the signals transit at some significant fraction of the speed of light. Nevertheless, one can use microelectronic fabrication techniques to construct geometries for which this transit time is non-negligible and further, provides a measure of the position-of-interaction.

Consider Fig. 1.12, in which our zig-zag pattern of Fig. 1.4 is replaced by a rectilinear electrode that snakes back-and-forth across the face of the detector. As with a conventional surface-barrier detector, the detector is depleted via the bounding electrodes so that if ionizing radiation interacts in the detection volume, then the resulting charge carriers will drift toward the contacts under the influence of the electric field. The resulting induced current pulse will transit toward the readout contacts A and B, resulting in pulses such as those shown in Fig. 1.12b (assuming proper termination and transmission-line design). The time difference between the pulses can then be used to measure the position at which the charge was collected. If the timing uncertainty is

significantly less than the transit time along a single strip, then the position can be isolated to one strip of the pattern or better. If the strips are closely separated, then the lateral resolution can be made quite small. In particular, using conventional photolithographic techniques at the University of Michigan, we can fabricate strips and gaps that are a fraction of a micron, resulting in sub-micron position resolution. If an identical pattern is deposited on the bottom of the wafer in an orthogonal configuration, then one can determine the position in both lateral directions. Hence, the pulse duration (of either signal) provides a measure of the depth of interaction, and the pulse time-difference provides a measure of the lateral position.

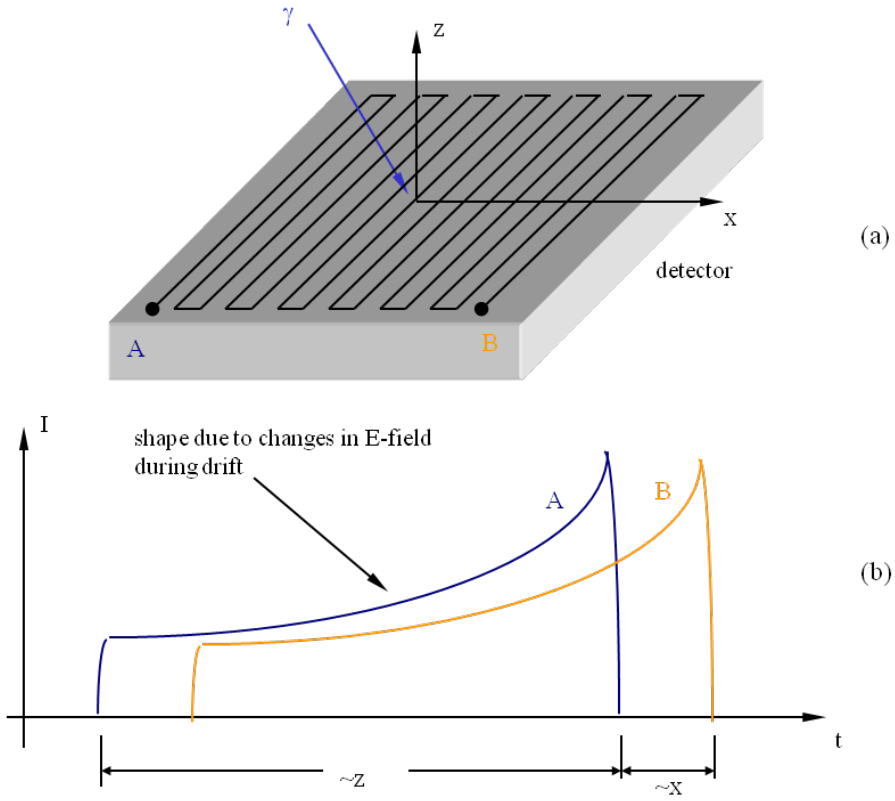


Fig. 1.12: Schematic of minimal-readout time-based position-sensing method.

The implementation of such a scheme depends on the propagation velocity of the pulse, the length of strips, and the timing resolution of the readout system. The pulse propagation velocity depends on the materials used in the transmission line and its surroundings as well as its geometric layout. For instance, in order to slow down the

pulse, a microstrip line can be surrounded by a material with high relative permittivity, or the strip can be made with a multitude of kinks and curves in order to retard the propagation. Let us consider the simple system shown in Fig. 1.12a, in which a straight metallic microstrip is deposited directly on our silicon substrate. In that case, the signal propagation velocity for a microstrip with a width in the 1-10 μm range is 0.4 c, where c is the speed of light in vacuum [32].

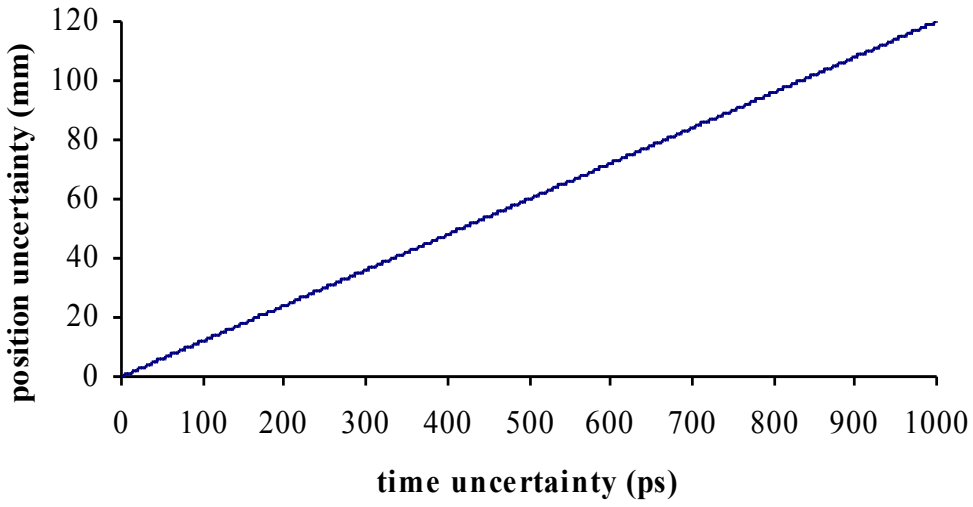


Fig. 1.13: For a 1-10 μm microstrip deposited on silicon, the relationship between the position uncertainty and the time uncertainty.

We can therefore relate the position uncertainty (σ_x) to the timing uncertainty (σ_t), as shown in Fig. 1.13. For the sake of clarity, a 1 ns time resolution- which is readily achievable- results in a position uncertainty of around 10 cm (or 4"). If the system has 100 ps time resolution, which again is achievable with standard components, then the charge collection point can be isolated to within 1 cm. In state-of-the-art systems, 10 ps or even 1 ps (or better) timing resolution is achieved, resulting in σ_x–values of approximately 1 mm and 100 μm, respectively. Thus, one does not need an inordinately long strip-length in order to isolate the position of charge-induction to a single strip. Finally, if one bounds the microstrip with neighboring electrodes, then the group velocity

decreases by orders of magnitude due to inductive effects we confirmed experimentally and will be discussed later.

1.4 Microstrip Design for Silicon Detectors

The design of microstrip lines on a detector substrate differs from typical design configurations because of charge creation within the semiconductor and their eventual distribution to other lines on the detector surface. Once the charges are created due to the gamma ray interaction within the Si substrate, the charge cloud drifts towards the electrodes when the bias voltage is maintained considerably high and the electron drift time is also small.

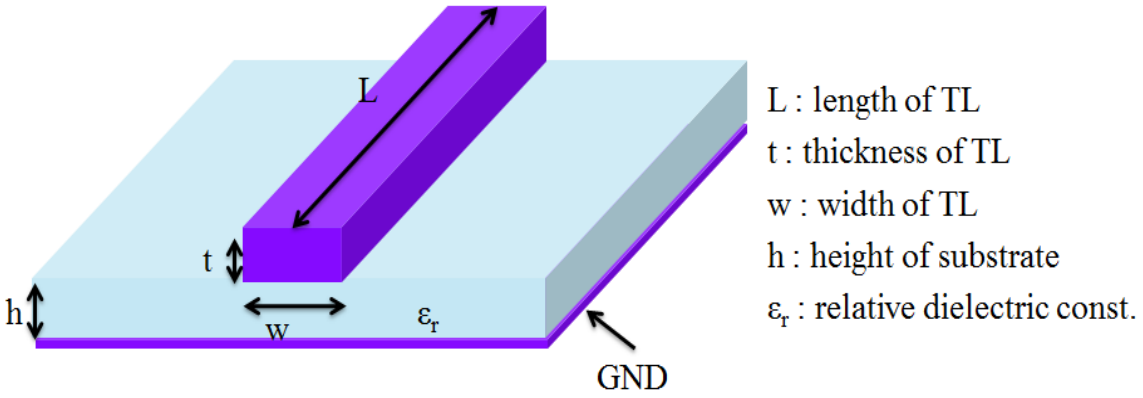


Fig. 1.14: Single Microstrip line.

The microstrip line consists of a single conductor trace on one side of a dielectric substrate and a single ground plane on its opposite side. The microstrip structure is open at the top. This open configuration makes the microstrip very convenient for its use in microwave hybrid integration (MIC). A microstrip is capable of propagating DC or AC at frequencies up to some tens of gigahertz. All microstrip lines are dispersive if the frequency of a signal exciting a microstrip line is doubled and the phase constant or wave number β ($= 2\pi / \lambda_g$) is not exactly doubled, which is termed dispersion. Methods of microstrip signal analysis may be classified into three groups; Quasi-static methods, dispersion methods and full wave analysis [33,34]. In the quasi-static methods, the nature of the mode of propagation is considered to be a pure transverse electromagnetic (TEM),

and microstrip responses are calculated from the capacitance and inductance of the structure. In the second group, called dispersion models, the effective dielectric constant and characteristic impedance are analyzed in frequency. In the third group, called full-wave analysis models, the scattering parameters are analyzed in frequency [32-34]. The general geometry of a microstrip is shown in Fig. 1.14. The most important dimensional parameters are the microstrip width w, height h, and the relative permittivity of the substrate ε_r. The characteristic impedance of a microstrip line is principally a function of its width and the thickness of the substrate. The physical length depends upon the wavelength, which is a function of the width, substrate permittivity and signal frequency.

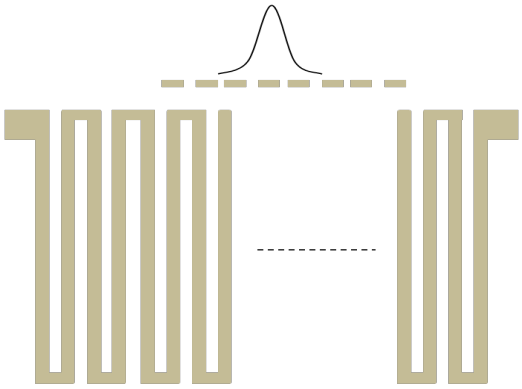


Fig. 1.15: Meander Lines. Charge cloud gets collected by one of these lines and some fraction gets collected on the neighboring lines. But because of the tight spacing cross talk or dispersion between the lines interfere with the charge pulses on the neighboring lines.

The structure for the present analysis is a meandering microstrip line whose line length is large (around 1 cm) but w/h (width to height ratio) and s/h (spacing to height ratio) is small. The resistivity of the silicon in use is 10 kΩ-cm which is quite low because 500 μm thick Si will only present a maximum resistance of 500 Ω for 1 cm^2 of area [35]. A finite value of conductance (instead of an ideal value of zero) the microstrip line analysis becomes more complex with impedance now having reactive part instead of just the resistive part.

Complications are avoided by reverse biasing the p-n structure of Si, so that the electrons essentially drift toward the electrode and the lines now can be treated as a

transmission line structure without having to deal with large values of conductance. The position resolution of a detector is determined by the electrode geometry. The field lines remain parallel in the detector and bend along the surface to end on the electrodes as in Fig. 1.15. Hence it is the pitch that determines the electrical segmentation rather than the width of the electrodes. Thus, for a single strip, the position resolution is same as the pitch. But when one has multiple strips, the difference between measured and true positions follow a Gaussian distribution which is directly related to the pitch again. Thermal diffusion is another factor due to which the charge cloud spreads transversely which is independent of the mobility of electrons and holes.

Any transverse spread is turned to our advantage as one can find the charge distribution over the central strip and its neighbors to improve the position resolution beyond what can be obtained in a single strip case which is limited by the strip geometry; namely, the pitch. When the charges drift towards neighboring strips and again drift from there to read outs on both ends, there is the simultaneous movement of pulses on the strip lines and hence cross talk analysis becomes more complex. The proposed size of the wafer for the present research is around 3" by 3". So when the charges drift along they essentially go through a delay line and arrive at the ends with a constant delay in an ideal case. But at higher frequencies when the line length becomes comparable to the wavelength, coupling between the lines (which is a function of frequency) now also depends on the line length. This means that some frequencies which exhibit more coupling will interfere with charge pulses on the neighboring strips induced by diffusion. During the course of this analysis we'll be evaluating each of these factors along with the effects of corner discontinuities for an optimal design of a microstrip structure with read outs at both ends.

1.4.1. Single Microstrip line analysis

For simplicity and to determine the line dimensions, let us consider a single microstrip. The most general way a transmission line can be represented is through lumped equivalents which are distributed all along the length of the line.

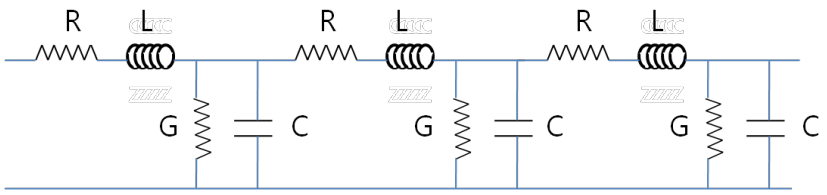


Fig. 1.16: Equivalent Circuit of a Transmission line.

Referring to Fig. 1.16, at lower frequencies the inductive effects (= Lω) of the line diminish; hence, the line behaves like an RC line. But at higher frequencies the inductive effects of the line begin to dominate to such an extent that under normal values of line resistances, the value of R can hence be neglected. The line at high frequencies essentially behaves like an LC line. We have all along assumed that the conductance of the substrate is negligible which in our case of detectors is not true unless the doped Si is reverse biased and the electrons only drift across it. For a lossless transmission line, both the resistance of the signal line and the line to ground become negligible and behaves as though it has a distributed impedance given by the value $\sqrt{L/C}$ and phase coefficient $\beta = \omega\sqrt{L/C}$. This is derived from two important equations of a transmission line [34]:

$\gamma = \sqrt{(R + j\omega L)(G + j\omega)}$, where γ the propagation constant = α+jβ where α is attenuation constant and β is phase coefficient. So when R and G are negligible the line offers zero attenuation of the signal and the only change in the signal will be its phase.
The characteristic impedance can be expressed as:

$$Z_0 = \sqrt{\frac{L}{C}}\left\{1 - i\left[\frac{R}{2\omega L} - \frac{G}{2\omega C}\right]\right\} \tag{1.8}$$

Hence for negligible values of R and G the distributed impedance of the line becomes resistive (with resistance $Z_o = (L/C)^{1/2}$) and independent of frequency. To achieve the above criteria, we have to make $R \leq \omega L/100$ and $G \leq \omega C/100$. The propagation time of electromagnetic waves on a microstrip line with dielectric constant ε is given by $T_p = 1.017(0.475 \cdot \varepsilon + 0.67)^{0.5}$ [32]. So for Si we get a propagation time of 8.45 ns/m. We can neglect conductance assuming a reverse bias condition. But for aluminum electrodes the

low resistivity, $2.8\text{x}10^{-8}$ Ohms–m, does not help in the resistance of the line when its dimensions are small, such as in our case have a line width of 10 microns. This sets a minimum limit on the thickness of the Al line. Also it means that the thickness of the Al electrode has to be huge in order to avoid the resistive effects of the line. A balance can be obtained by increasing the line width slightly so that the thickness can be reduced. As stated earlier the position resolution expected is less than the pitch because of interpolated charges on the neighboring strips.

Consider a conventional microstrip line whose physical structure is shown in Fig.1.14. If the losses and metal thickness are neglected, the characteristic impedance Z_o of this microstrip line can be estimated using the following equations [36]:

$$Z_0 = \begin{cases} \frac{60}{\sqrt{\varepsilon_{eff}}} \ln\left(\frac{8\text{h}}{\text{w}} + \frac{\text{w}}{4\text{h}}\right) \Omega & \text{if } \frac{w}{L} \leq 1 \\ \frac{120\pi}{\sqrt{\varepsilon_{eff}}} \left[\frac{1}{\left(\frac{w}{h}\right) + 1.393 + 0.667 \ln\left(\frac{w}{h} + 1.444\right)} \right] \Omega & \text{if } \frac{w}{L} \geq 1 \end{cases} \tag{1.9}$$

where ε_{eff} is the effective dielectric constant given by

$$\varepsilon_{eff} = \frac{\varepsilon_r + 1}{2} + \frac{\varepsilon_r - 1}{2} \left(\sqrt{1 + \left(\frac{12h}{w}\right)}^{-1} + 0.04\left(1 - \frac{w}{h}\right)^2 \right) \tag{1.10}$$

The propagation velocity v_p or wave speed of a lossless microstrip line is calculated using

$$v_p = \frac{c}{\sqrt{\varepsilon_{eff}}} \tag{1.11}$$

where c is the velocity of light ($c \approx 3.0 \times 10^8$ m/s).

In order to calculate the inductance per length, in a lossless transmission line, we have

$$\beta = \omega\sqrt{LC} \text{ and } Z_0 = \sqrt{\frac{L}{C}} \text{ since } v_p = \frac{\omega}{\beta}. \tag{1.12}$$

Then by substituting

$$v_p = \frac{1}{\sqrt{LC}} \tag{1.13}$$

Therefore to define L and C in terms of v_p and Z_o, and substituting onto v_p we find two important equations.

$$\text{L} = {Z_0}^2 C \qquad \text{and} \qquad Z_0 C = \frac{1}{v_p} = \beta \tag{1.14}$$

Using these equations, we calculated characteristic impedance (Z_0), inductance per unit length (L) and capacitance (C) per unit length (meter) of the 10, 25, 50 and 100 µm width of microstrip as shown in Table 1.1.

Table 1.1: Calculated characteristic impedance (Z_0), inductance (L), capacitance (C) and propagation time (T_p) per unit length (meter) of the 10, 25, 50 and 100 µm width of microstrip.

	10µm	25µm	50µm	100µm
Z_0 (Ω)	135.7443	114.9127	98.6976	82.3501
L (µH/m)	1.1793	1.0065	0.8719	0.7358
C (pF/m)	64.0011	76.2188	89.5103	108.5010
T_p(ns/m)	1.2806	1.2703	1.2594	1.2452

The speed of light slows down in any other medium by the square root of the relative dielectric coefficient of the medium. The signal propagation time (T_p) is the inverse of this figure. The propagation speed (v_p) for a microstrip trace poses the problem that the trace is in a mixed environment. The medium underneath the trace is the silicon dielectric. The medium above the trace is air. So the electromagnetic (EM) wave travels through this mixed medium at a speed somewhere between that of the speed of light and the propagation speed in microstrip line. Two things that contribute to an increase in the concentration of field lines underneath the trace are bringing the trace closer to the plane and increasing the trace width. Note that increasing the trace thickness has a minor effect on propagation speed, but the effect is much smaller than with the other variables. Each of these will cause the propagation speed to slow down.

Note that, in the limit, the propagation speed for a microstrip trace is the same as for a strip-line trace. The limit is reached with an infinitely wide trace or a zero separation between trace and plane. Under any other conditions, the propagation speed increases. Therefore, we should think of the microstrip propagation speed as some factor of the propagation speed for the same trace in a strip-line environment surrounded by a material with the same dielectric coefficient.

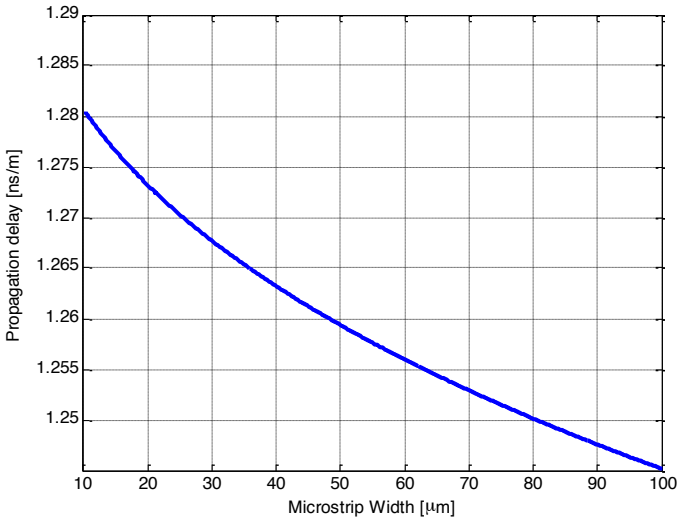


Fig. 1.17: Propagation time as a function of frequency for the 50 μm width microstrip line.

On the other hand, the group velocity is the velocity of a modulated waveform's envelope and describes how fast information propagates. It is the velocity at which the energy or information in the waveform moves. The changes in modulation on a signal as it travels down a microstrip are related to this value v_g. Thus group velocity can't be more than the speed of light in a vacuum, c. For a lossless, dispersionless line, the group velocity and phase velocity are the same as below [34];

$$v_g v_p = c^2 \tag{1.15}$$

It is to borne in mind that the phase velocity is the velocity at which a field component, such as the electric field, travels down the guide, while the group velocity is the velocity at which energy travels down the guide. We derive an expression that is a measure of the dispersion a modulated signal would undergo as it travels down a waveguide of length L. Toward this end, we consider the group delay, t_g, which is the time that it takes for a change in signal to propagate. The group delay is simply length divided by speed, or

$$t_g = \frac{L}{v_g} \tag{1.16}$$

Using Eq. 1.15 and 1.16, we calculated propagation time of the 10, 25, 50 and 100 μm width of microstrip as shown in Table 1.1.

1.4.2 Dispersion Analysis of a Microstrip line

When a microstrip has a very simple geometric structure, the electromagnetic fields involved are complex. The dielectric–air interface modifies the mode of propagation in a microstrip. The mode of propagation on a microstrip is almost transverse electric and magnetic (TEM) or quasi-TEM. The microstrip is a type of microwave circuit interconnection. This "waveguide" will support the "quasi-TEM" mode, which like regular TEM modes, has no non-zero cutoff frequency. TEM is when the electromagnetic waves propagate with the same phase velocity, because the wave is in one medium. A microstrip on the other hand propagates within two mediums: the substrate and air. As we know the wave will propagate with different speed in the two mediums. Therefore it is called quasi-TEM. This is why a dielectric constant for a microstrip is between 1(air) and that of the material. Fig. 1.18 shows the electric and magnetic field on the microstrip in TEM mode and quasi-TEM mode.

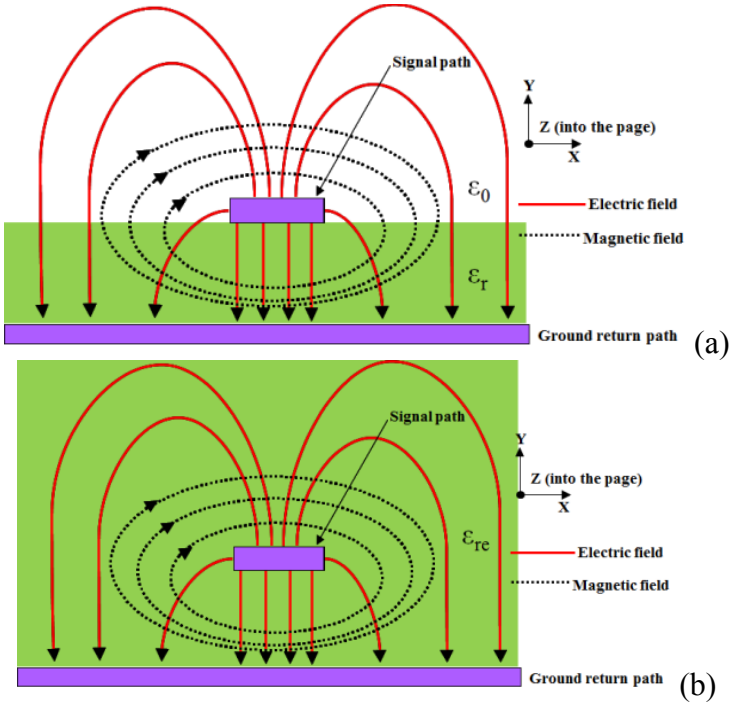


Fig. 1.18: Electric field and magnetic field in TEM mode (a) and quasi-TEM mode (b).

The effective microstrip permittivity ε_{eff} (f) increases with frequency and the wave is progressively slowed down [38]. The relationship between frequency dependent effective microstrip permittivity and phase velocity is given by:

$$\varepsilon_{eff}(f) = \left\{\frac{c}{v_p(f)}\right\}^2 \tag{1.17}$$

The frequency-dependent effective microstrip permittivity may also be written as a function involving ε_{eff} and Z_0 as follows: $\varepsilon_{eff}\left(f, h, \varepsilon_r \varepsilon_{eff} Z_0\right)$. The basic features regarding the microstrip dispersion are as follows: (a) ε_{eff} (f) always increases with frequency, (b) ε_{eff} (f=0) = ε_{eff} , and (c) ε_{eff} (f→ ∞) = ε_r.

An improved design formula has been developed by Kirschning and Jansen [39]. Their approach begins with a function that bears a close resemblance to Getsinger's, but they have a new and reportedly more accurate frequency-dependent denominator term. Also, their formula covers a much wider range of permittivities, aspect ratios and frequencies than those considered by Getsinger [40] or Edwards and Owens [41]. Kirschning and Jansen's basic expression is (with frequency f in GHz and thickness h in cm):

$$\varepsilon_{eff}(f) = \varepsilon_r - \left(\frac{\varepsilon_r - \varepsilon_{eff}}{1 + P(f)}\right) \tag{1.18}$$

and the form of the denominator frequency function is

$$\mathrm{P(f)} = P_1 P_2 \{(0.1844 + P_3 P_4) 10 fh\}^{1.5763} \tag{1.19}$$

where

$$P_1 = 0.27488 + \left[0.6315 + \frac{0.525}{(1 + 0.157fh)^{20}}\right]\left(\frac{w}{h}\right) - 0.065683 \exp\left(-\frac{8.7513w}{h}\right)$$

$$P_2 = 0.33622[1 - \exp(-0.03442\varepsilon_r)]$$

$$P_3 = 0.3363\exp\left(-\frac{4.6w}{h}\right)[1 - \exp(-fh/3.87^{4.97})]$$

$$P_4 = 1 + 2.751[1 - \exp\left(-(\frac{\varepsilon_r}{15.916})^8\right)]$$

These researchers used computer-based matching in conjunction with the available hybrid mode results (derived from full electromagnetic numerical computations) to force the correct asymptotic behavior of the function P(f). An accuracy of better than 0.6% is claimed for all frequencies up to 60 GHz (although a full check only appears to have been conducted up to 30 GHz). The validity ranges are:

$1 \leq \varepsilon_r \leq 20$, $0.01 \leq w/h \leq 100$,

and $0 \leq \frac{h}{\lambda_0} \leq 0.13$ where λ_0 is the free space wavelength at operation frequency f.

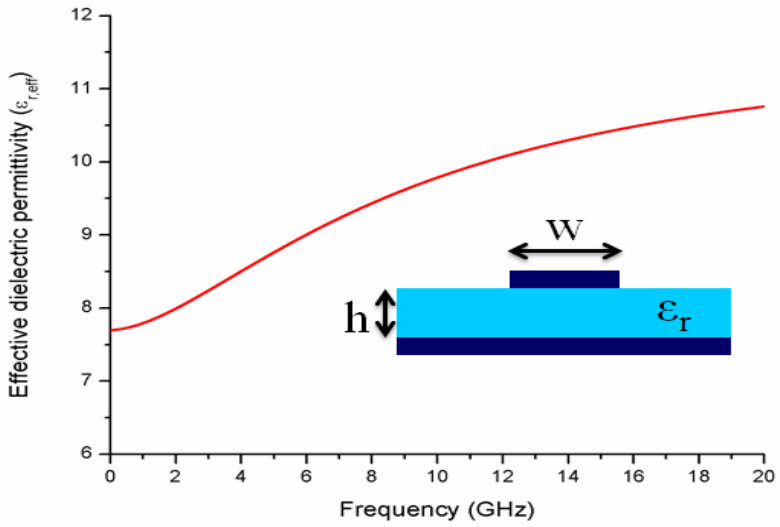


Fig. 1.19: Effective dielectric permittivity as a function of frequency.

When the frequency increases, the quasi-TEM analysis of the microstrip will no longer remain valid and a full-wave analysis becomes necessary. At higher frequencies the quasi-TEM mode still exists but the longitudinal field components E_z and H_z become important, i.e. the mode loses its quasi-TEM character. At the same time this mode becomes dispersive: the signal speed changes and depends on frequency. This can be expressed by also making the effective dielectric permittivity frequency dependent. Fig. 1.19 shows this effective dielectric permittivity for the microstrip depicted in the same figure. The metal strip is infinitely thin, w = 0.1mm, h = 0.5mm and ε_r = 11.7 for the

silicon. The value of $\varepsilon_{r,eff}$ at $\omega = 0$ is 7.7. When the frequency increases the effective dielectric permittivity also increases. This is a consequence of the fact that at higher frequencies field lines tend to concentrate in the dielectric. For very high frequencies no field lines are found in the air. Hence $\varepsilon_{r,eff}$ must asymptotically tend towards $\varepsilon_{r,eff} =$ 11.7.

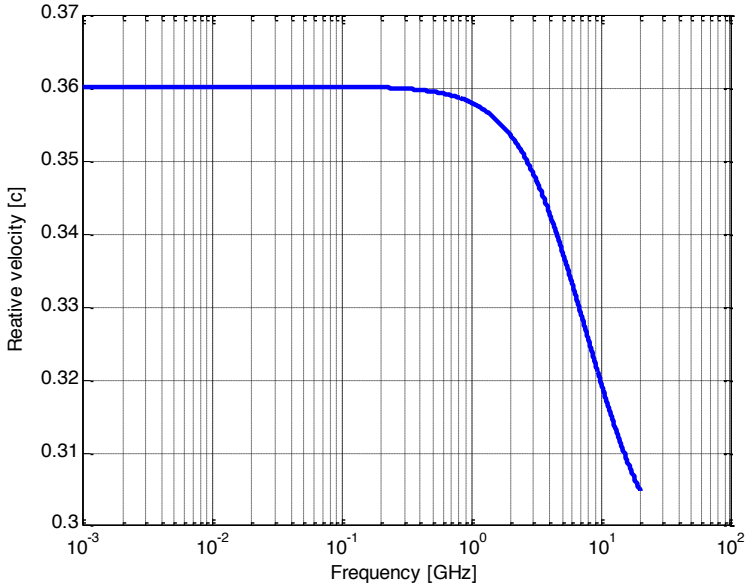


Fig. 1.20: Relative phase velocity as a function of frequency for the 50 μm width microstrip line.

Once the effective dielectric constant of a microstrip is determined, the guided wavelength of the quasi-TEM mode of microstrip is given by

$$\lambda_g = \frac{\lambda_0}{\sqrt{\varepsilon_{r,eff}}} \tag{1.20}$$

where λ_0 is the free space wavelength at operation frequency f. The associated propagation constant β and phase velocity v_p can be determined by Eqs. 1.11-1.13.

With these equations, the phase velocity as a function of frequency and microstrip's line width are shown in Fig. 1.20 and 1.21.

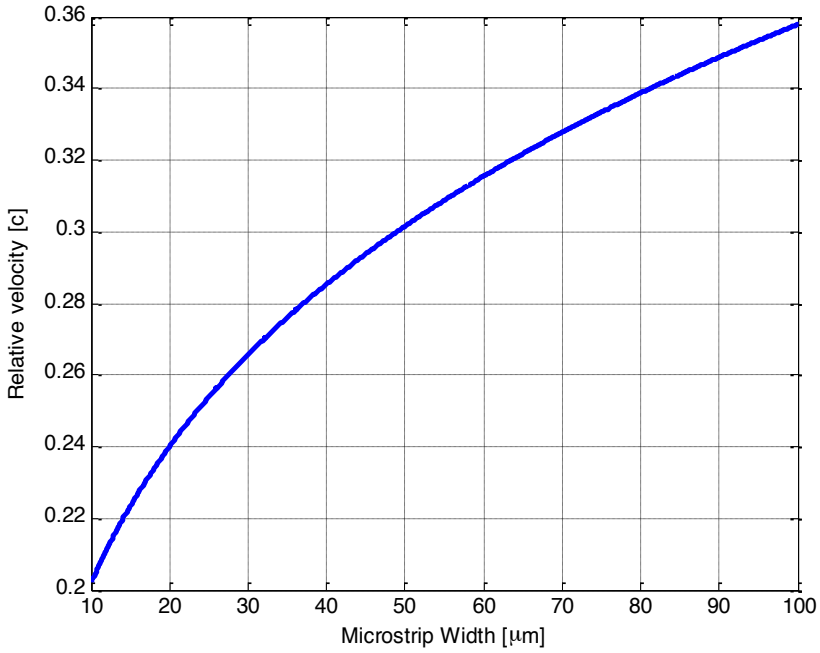


Fig. 1.21: Relative group velocity as a function of microstrip line width.

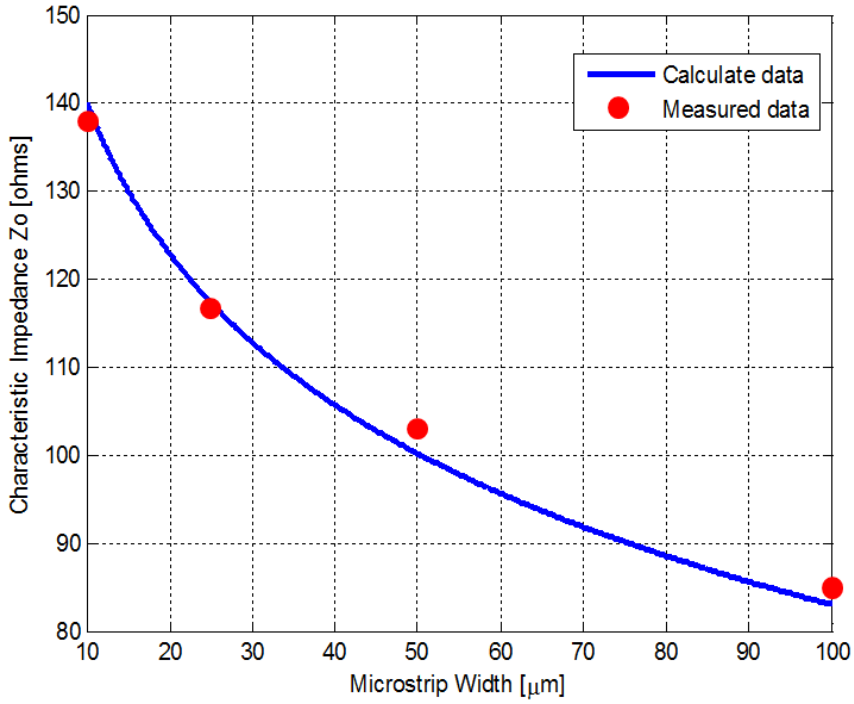


Fig. 1.22: Calculated (blue line) and measured (red circle) characteristic impedance of microstrip line as a function microstrip line width.

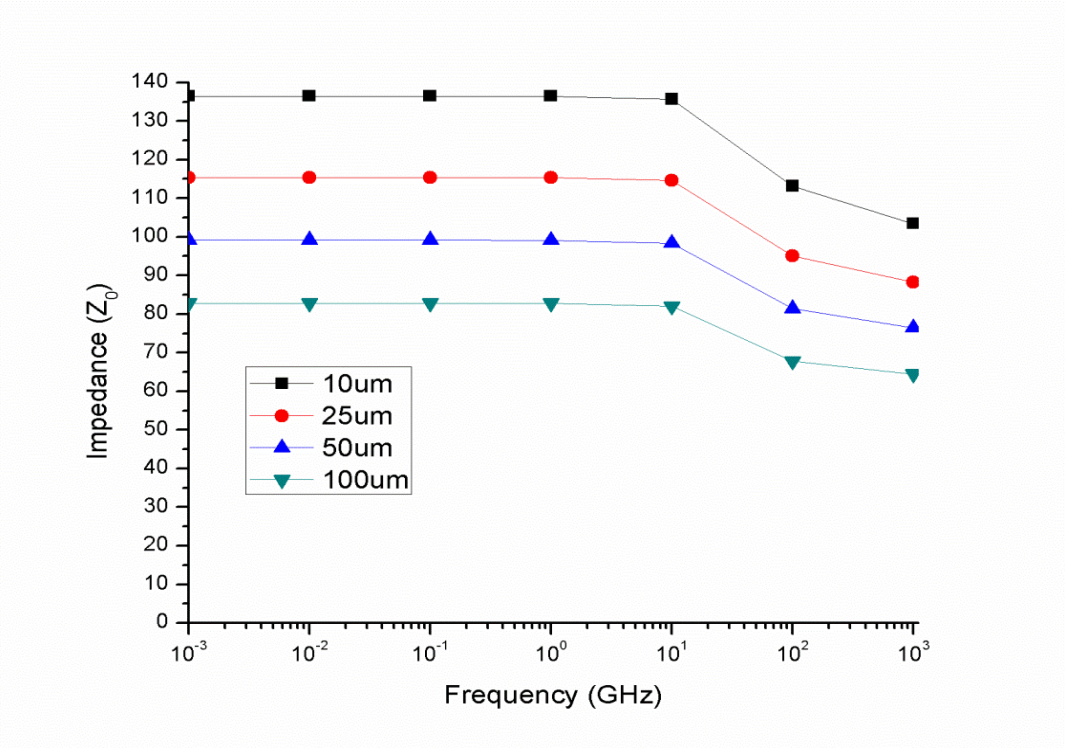


Fig. 1.23: Characteristic impedance for the different microstrip line widths as a function of frequency.

In addition, the characteristic impedances of a microstrip line as a function of microstrip line width and frequency are shown in Fig. 1.22 and Fig. 1.23. From the fabricated meander line detectors impedances are measured and compared with calculated data in Fig. 1.22. Generally speaking, there is dispersion in microstrips; namely, the phase velocity and characteristic impedance are not a constant but depend on frequency.

1.4.3 Coupled Microstrip lines and effects

Figure 1.24a illustrates the cross section of a pair of coupled microstrip lines under consideration in this section, where two microstrip lines of width (w) are in the parallel- or edge-coupled configuration with a separation *s*. This coupled line structure supports quasi-TEM modes, i.e., the even- and odd-mode, as shown in Fig. 1.24b and 1.24c. For an even-mode excitation, both microstrip lines have the same voltage potentials or carry the same sign charges, say the positive ones, resulting in a magnetic wall at the symmetry plane, as Fig. 1.24b shows. They propagate with different phase velocities because they

are not pure TEM modes, which mean that they experience different permittivities. Therefore, the coupled microstrip lines are characterized by the characteristic impedances as well as the effective dielectric constants.

In a static approach similar to the single microstrip, the even- and odd-mode characteristic impedances and effective dielectric constants of the coupled microstrip lines are obtained in terms of the even-mode capacitances, denoted by C_e. As shown in Fig. 1.24b and c, the even-mode capacitances C_e can be expressed as [42]

$$C_e = C_p + C_f + C'_f \quad \text{and} \quad C_o = C_p + C_f + C_{gd} + C_{ga} \tag{1.21}$$

In these expressions, C_p denotes the parallel plate capacitance between the strip and the ground plane, and hence is simply given by

$$C_p = \varepsilon_0 \varepsilon_r w/h \tag{1.22}$$

C_f is the fringe capacitance as if for an uncoupled single microstrip line, and is evaluated by

$$2C_f = \sqrt{\varepsilon_{eff}}/(cZ_0) - C_p \tag{1.23}$$

where Z_0 is the characteristic impedance of a microstrip line with the same ratio of w/h and can be calculated by Eq. 1.9.

The term C_f' accounts for the modification of fringe capacitance C_f of a single line due the presence of another line. An empirical expression for C_f' is given below

$$C'_f = \frac{C_f \sqrt[4]{\varepsilon_0/\varepsilon_{eff}}}{1+A\left(\frac{h}{s}\right)\tanh\left(\frac{10s}{h}\right)} \tag{1.24}$$

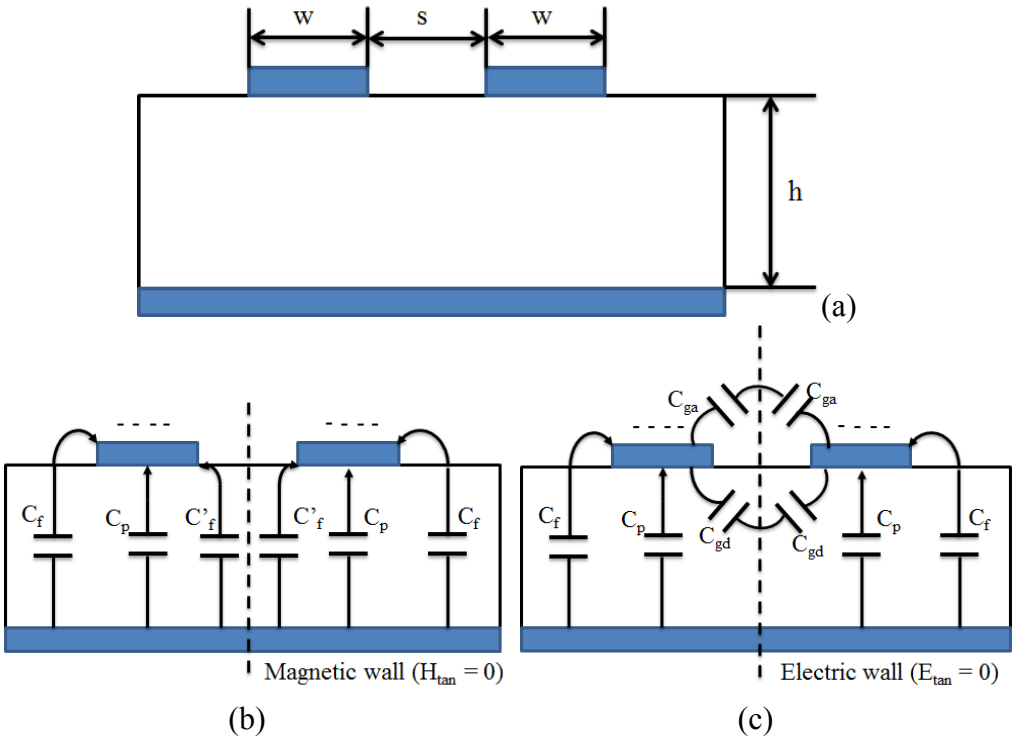


Fig. 1.24: Configuration of coupled microstrip line (a) and Quasi-TEM modes of a pair of coupled microstrip lines for the even mode (b) and odd mode (c).

If the coupled segments are in a strip line configuration, the following set of equations are used

$$C_e(\varepsilon_r) = \frac{\sqrt{\varepsilon_r}}{cZ_e} \tag{1.25}$$

where, c is the velocity of light and Z_e is the characteristic impedance of even-mode and is

$$Z_e = \frac{30\pi}{\sqrt{\varepsilon_r}} \frac{K(k'_e)}{K(k_e)} \tag{1.26}$$

where the function K(k) and K(k') are the complete elliptic function and its complement and their ratio is given by

$$\frac{K(k_e{}')}{K(k_e)} = \begin{cases} \frac{1}{\pi} ln\left(2\frac{1+\sqrt{k_e'}}{1-\sqrt{k_e'}}\right) & if\ 0 \le k_e < \frac{1}{\sqrt{2}} \\ \frac{\pi}{ln\left(2\frac{1+\sqrt{k_e}}{1-\sqrt{k_e}}\right)} & if\ \frac{1}{\sqrt{2}} \le k_e < 1 \end{cases} \tag{1.27}$$

$$k_e = tanh\left(\frac{\pi w}{4h}\right)\tanh\left[\frac{\pi}{4}\left(\frac{w+s}{h}\right)\right], \qquad k'_e = \sqrt{1-{k_e}^2} \tag{1.28}$$

For the odd-mode, C_{ga} and C_{gd} represent, respectively, the fringe capacitances for the air and dielectric regions across the coupling gap. The capacitance C_{gd} can be found from the corresponding coupled microstrip line geometry, with the spacing between the ground planes given by h. A closed-form expression for Cgd is

$$C_{gd} = \frac{\varepsilon_o \varepsilon_r}{\pi} ln\left[coth\left(\frac{\pi s}{4h}\right)\right] + 0.65C_f\left(\frac{0.02\sqrt{\varepsilon_r}}{s/h} + 1 - {\varepsilon_r}^2\right) \tag{1.29}$$

The capacitance Cga can be modified from the capacitance of the corresponding coplanar strips, and expressed in terms of a ratio of two elliptic functions

$$C_{ga} = \varepsilon_o \frac{K(k')}{K(k)} \tag{1.30}$$

where $k = \frac{s/h}{\frac{s}{h}+2W/h}$ and $k' = \sqrt{1-k^2}$ and the ratio of the elliptic function is given at Eq. 1.27.

Table 1.2: Calculated characteristic impedance (Z_0), inductance (L), capacitance (C) and propagation time (T_p) per unit length (meter) of the 10, 25, 50 and 100 µm width of coupled microstrip

	10µm	25µm	50µm	100µm
Z_e (Ω)	124.7333	107.0005	92.1554	75.4664
Z_o (Ω)	26.7827	34.0744	38.8534	43.2461
L_e (µH/m)	0.6191	0.5562	0.5060	0.4454
L_o (µH/m)	0.1954	0.2505	0.2881	0.3326
C_e (pF/m)	33.7944	48.5814	59.5858	78.2087
C_o (pF/m)	272.4701	215.7515	190.8606	177.8571
T_{pe} (ns/m)	2.2415	2.1403	2.0262	1.8851
T_{po} (ns/m)	1.5246	1.5134	1.5003	1.4465

Using these equations, we calculated the characteristic impedance (Z_0), inductance per unit length (L) and capacitance (C) per unit length (meter) of the 10, 25, 50 and 100 μm width of coupled microstrip for even- and odd-mode as shown in Table 1.2. In these equations, we make five assumptions as follows: (a) all the signals propagate in quasi-TEM mode, (b) the geometry of the coupled line is symmetric, which means that the two traces have the same width and are on the same layer, (c) the thickness of the trace is negligible, (d) the return plane is much wider than the traces, and (e) there are no effects due to fields at the terminations. Comparing te results from Table 1.1 and 1.2, the total capacitances of the coupled microstrip lines are increased and it causes increases in the propagation time per unit length.

Coupled line analysis for this is quite different from the usual high frequency propagation. This is because the charge pulses are drifting on the electrodes rather than being transmitted. There are two effects to be considered here. The coupling effect on neighboring lines associated with propagation of a pulse down a reference line and coupling when there is simultaneous movement of charge pulses on neighboring lines due to interpolation of charges.

When a single pulse is travelling on a line, due to interline capacitance, a pulse of opposite polarity is created in the immediate neighboring line. This capacitive coupling between the lines dominates the backplane capacitance, which is the capacitance between the signal line and the ground as show in Fig. 1.24b. Since the capacitive coupling occurs on both neighboring sides of the reference line, the coupled signal noise is produced on both sides of the line and exhibits distributed signal strength. If there are many lines on either side of the reference line then coupling capacitances form a series of capacitances on both sides. This strip to strip capacitance aids in transferring the signal (when the strips are not connected or “floating”) or the signal noise (when the strip are connected to each other). For our case of study the signal noise travels through this capacitance to all the successive lines. When charges produced due to diffusion travel down the microstrip line, the coupling that occurs between the 2 signals on neighboring lines is an even mode coupling when they travel in the same direction and odd mode coupling when they travel in opposite direction. Even mode is when there is equal polarity on both the lines and odd mode, when they have opposite polarity. The even mode coupling produces an electrical

wall between the lines and the electrical lines of force recede from each other like an open circuit. For an odd mode there a magnetic wall and the electrical lines are coupled like a short circuit.

1.4.4 Slow factor calculation of meander-line on Metal-Insulator-Semiconductor

Multilayered structures that consist of metal, insulator and semiconductor layer (MIS), under certain conditions, are shown to support the Maxwell-Wagner polarization mechanism. The direct effect of this mechanism is the significant increase of the capacitance per unit length or the large slow-down of the propagation velocity in a frequency band [43-45]. Propagation of a slow-wave mode in MIS structure was first predicted by Guckel et al. [43], and it was experimentally confirmed by Hasegawa et al. [44]. The phase velocity of a slow-wave mode is less than that of a TEM (or quasi-TEM) mode by a factor which is called the slowing factor and is defined as the ratio of wavelength in free space to that of the waveguide, (λ_0/λ_g). The frequency range of the occurrence of slow-wave propagation is determined by a number of factors such as the substrate process, permittivity, conductivity and device geometry. Due to the smaller wavelength of the slow-wave mode, which is caused by the reduction of phase velocity, further miniaturization of passive components on semiconductor substrate that utilize slow-wave propagation has become possible. MIS transmission lines can be designed to potentially create a large slowing factor, but there are many engineering challenges associated with them such as low characteristic impedance and high losses [45].

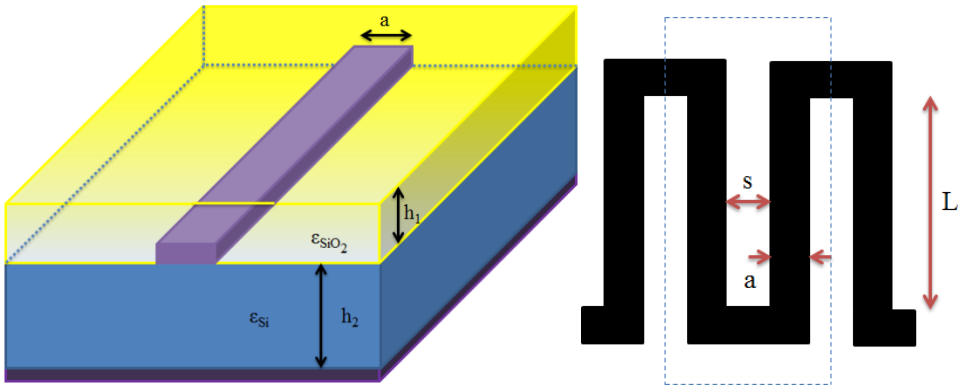


Fig. 1.25: MIS Transmission Line Structure and geometrical parameters (a) and meander line unit cell.

The effective permittivity of the metal-insulator-semiconductor (MIS) microstrip line on a substrate of Si-SiO_2, as shown in Fig. 1.25a, is studied. The transition from one propagation mode to the other is determined by several parameters, such as the frequency of the applied signal as well as the electric and magnetic characteristics of the materials involved. Since the dominance of each mode is strongly dictated by two parameters, the comparison of the substrate's thickness with the skin depth.

At intermediate frequencies and intermediate values of substrate conductivity, the skin depth is comparable to the thickness of the substrate; thus propagation is due to the slow-wave mode. This mode owes its existence to a thin space-charge layer which is formed at the insulator-semiconductor interface. Therefore, the electric field lines are greatly confined within the top metal layer and the insulator-semiconductor interface, introduces a large distributed capacitance in accordance with the small distance h_1 separating these two boundaries [46].

$$c_{ec} \cong \varepsilon_o \varepsilon_{overlay} \frac{a}{h_{overlay}} \tag{1.31}$$

Where a is the width of the microstrip. Additionally, the full penetration of magnetic fields implies an inductance which is related on the large thickness of the substrate thickness h_{semi}.

$$\mathrm{L} \cong \mu_o \frac{h_{semi}+h_{overlay}}{a} \tag{1.32}$$

In a perfect homogeneous dielectric, where the energy is guided through the entire substrate, the electric and magnetic fields penetrate to the same depth. By utilization of Eqs. 1.31 and 1.32 in Eq. 1.12, the propagation velocity of a wave through a microstrip line operating in the slow-wave mode region can be

$$v_p = \frac{1}{\sqrt{\mu_o \frac{h_{semi}+h_{overlay}}{h_{overlay}} \varepsilon_o \varepsilon_{overlay}}} = \frac{1}{\sqrt{\varepsilon_{ec}\mu_o}} = \frac{\omega}{\beta_{ec}} \tag{1.33}$$

where β_{ec} is now the phase constant of the slow-wave mode. Thus, the following equations are suggested to represent the effective complex permittivity (ε_{ec}) [44,46],

$$\varepsilon_{ec} \approx \frac{h_{semi}+h_{overlay}}{h_{overlay}}\ \varepsilon_{r,overlay} \tag{1.34}$$

where h_{semi} is the Si substrate's thickness, $h_{overlay}$ is the height of the overlay and $\varepsilon_{r,overlay}$ is the relative dielectric constant of the overlay and it can be calculated by using Eq. 1.10.

From the Eq. 1.21, the slowing factor, λ_0 / λ_g, can be expressed as

$$\frac{\lambda_0}{\lambda_g} = \sqrt{\varepsilon_{ec}} = 63.1650. \tag{1.35}$$

The propagation velocity is calculated by using Eq. 1.12 and the propagation time of the signal on the MIS microstrip delay line as

$$\tau_{MIS} = \frac{1}{v_p} = c\sqrt{\varepsilon_{ec}}. \tag{1.36}$$

For the 10, 25, 50 and 100 μm width meander line on MIS, the signal propagation times are calculated as shown in Table 1.3. The effective complex permittivity is increasing and it causes a several order-of-magnitude increase in the propagation time.

Table 1.3: Calculated characteristic slowing factor (λ_0 / λ_g), propagation velocity (v_p), and propagation time (T_p) per unit length (meter) of the 10, 25, 50 and 100 μm width of coupled microstrip in MIS structure.

	10μm	25μm	50μm	100μm
$\beta(=\lambda_0/\lambda_g)$	68.5042	65.0883	63.8254	63.1650
V_{pe} (m/s)	6.5124e+6	7.1783e+6	7.7326e+6	8.3983e+6
V_{po} (m/s)	9.5747e+6	1.0152+7	1.0443+7	1.0945+7
T_{pe}(ns/m)	153.5522	139.3085	129.3230	119.0723
T_{po} (ns/m)	104.4415	98.5046	95.7572	91.3682

1.4.5 Effect of the Microstrip Thickness

The thickness t is usually very small when the microstrip line is realized by conducting thin films; therefore, its effect may quite often be neglected. Nevertheless, its effect on the characteristic impedance and effective dielectric constant can be included [47].

$$Z_c(t) = \begin{cases} \frac{60}{\sqrt{\varepsilon_{re}}} ln\left\{\frac{8}{W_e(t)/h} + 0.25W_e(t)/h\right\} & if\ W/h \le 1 \\ \frac{120\pi}{\sqrt{\varepsilon_{re}}}\left\{\frac{W_e(t)}{h} + 1.393 + 0.667ln\left(\frac{W_e(t)}{h} + 1.44\right)\right\}^{-1} & if\ W/h \ge 1 \end{cases} \quad (1.37)$$

where

$$\frac{W_e(t)}{h} = \begin{cases} \frac{W}{h} + \frac{1.25}{\pi}\frac{t}{h}\left(1 + ln\frac{4\pi W}{t}\right) & if\ W/h \le 0.5\pi \\ \frac{W}{h} + \frac{1.25}{\pi}\frac{t}{h}\left(1 + ln\frac{2h}{t}\right) & if\ W/h \ge 0.5\pi \end{cases} \quad (1.38)$$

$$\varepsilon_{re(t)} = \varepsilon_{re} - \frac{\varepsilon_r - 1}{4.6}\frac{t/h}{\sqrt{W/h}} \quad (1.39)$$

In the above expressions, ε_{re} is the effective dielectric constant for t = 0. It can be observed that the effect of strip thickness on both the characteristic impedance and effective dielectric constant is insignificant for small values of t/h. However, the effect of strip thickness is significant for conductor loss of the microstrip line. Furthermore, conductor losses greatly exceed dielectric losses for most microstrip lines fabricated on low-loss substrates.

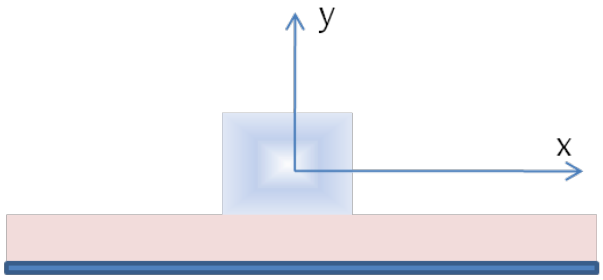


Fig. 1.26: At higher frequencies current crowds at the corners.

Skin depth is defined as the depth from the surface of the conductor where the current value falls to 1/e of that on the surface. The effects of skin depth on such a thick Al line cannot be ignored either. At higher frequencies the current flows through a thin layer on the outside surface of the microstrip conductors as shown in Fig. 1.26. Hence it can be assumed that there can be a maximum of 3 skin depths of conductor thickness from each surface for a uniform current distribution. Skin depth (δ_s) in the present case can be calculated to be

$$\delta_s = \sqrt{\frac{1}{\pi f \mu_0 \sigma}} = 2.6\mu\text{m} \tag{1.39}$$

where f is the frequency of the applied signal, μ_0 is the magnetic permeability of free space and σ is the conductivity of the substrate. The lateral dimension of 10 μm is close to this value (3δ) but there are 13 skin depths on the y direction of the conductor. Because of the charge accumulation towards the edges the effective resistance (whose reactance part is inductive) increases. The effect of this, which increases with line length, is to increase the internal inductance and resistance of the line. All calculations for the delay line detector's parameters are implemented by Matlab, and were included in APPENDIX A.

1.5 Objective of This Work

Recall that one can markedly enhance the directional resolution of a gamma-ray (or neutron) imager if the direction of the recoil electron (or product ion) can be measured, information that is lost in most existing devices. That direction can be inferred if the electron's ionization track is mapped, a task that requires fine position sensing and high charge sensitivity. This can be accomplished, at least crudely, using existing technology, but the devices are complex from an instrumentation standpoint and the charge collection time is lengthy. One of my goals is to thus accomplish charge cloud mapping but with fast information-collection and minimized readout intricacy. The complexity can be reduced by decreasing the number of channels required to collect the energy and positional information from each ionization event. For applications requiring high detection efficiency, such as border and portal monitoring, one prefers large detector areas, the deployment of which is cumbersome if each square centimeter of detector real-

estate requires a multitude of readout channels. Furthermore, if radiation detectors are to be deployed ubiquitously, then the cost is a non-negligible engineering concern, one means of controlling such is to minimize the cost of the underlying material. Despite its relatively low photoelectric cross-section, silicon can, in fact, succeed as a gamma-scatter camera if deployed in sufficient depth. If that depth takes the form of multiple detector slices, then one again desires minimal readout complexity per slice. We have therefore focused on the development of novel electrode partitioning geometries for which but a few readout pads are necessary.

We: (a) optimized the procedures for fabricating silicon-detectors with different electrode partitions, (b) built the analytical and numerical modeling tools for optimizing the detector and readout design, (c) used those tools to develop novel electrode configurations, and (d) continued the readout electronic development with the goal of minimizing the charge and temporal noise.

Our test-bed detectors are surface-barrier detectors, on which Schottky and ohmic contacts are deposited on high-resistivity silicon. Our fabrication procedures are sufficiently robust that low leakage-current devices can be made with high reproducibility. Furthermore, once an electrode patterned is designed, the fabrication requires only a few days to realize the detector. We thus have a smooth fabrication chain from which devices can be realized and compared to their predicted responses. Those predicted responses can be made with precision using the modeling tools developed. In Chapter II, we will discuss, theoretically, the methods by which the position of the ionization event can be inferred from the induced charge signal formed during carrier drift. Utilizing the shape of the induced current signal to determine the carrier's trajectory is the focus Chapters III and V, we will discuss a potential time-based method in which one or two readout contacts can be used to achieve fine lateral position sensing. Our diagnostic device under investigation will achieve, if successful, 20 μm or better lateral resolution across a sizable area using two readout channels. The physics governing the transformation between the charge-carrier motion and the induced current signal is thus well-understood. Precise predictions of the detector response are enabled by the development of the numerical modeling tools discussed in Chapter II. The radiation-solid interaction is modeled using Monte Carlo techniques via GEANT4, and the electric

field and weighting potentials in the detector environment are modeled with MAXWELL 3D as well as custom codes. The charge deposition from GEANT4 and the electric field information from MAXWELL is then input into custom C and Matlab codes to map each charge's motion from its point-of-creation to its collection at the electrodes or trapping at the surface. That motion results in induced charge (or current) signals at the electrodes. The preamplifier and other readout components are then modeled using PSpice.

The tools for accurately modeling the detector were therefore developed during the initial work period and put to good use in validating the behavior of the detector. In particular, an accurate model of the carrier's drift velocity variation with the electric field was needed to match the model to the measured alpha-particle response. Besides detector validation, the detector model will also be used to accelerate the rate at which different electrode geometries are investigated, and was used to develop the signal-processing routines which are necessary to intelligently interpret the pulse-shape information.

The degree of success in mapping the ionization events is determined by the uncertainties in the charge and temporal measurements. The experimental assessment of the time resolutions and noise properties of existing amplification components is made in Chapter V. Furthermore, the development, construction, and testing of a fast timing circuit is discussed. We thus have readout components with sub-nanosecond timing resolution and low noise so that the radiation events of interest can be accurately measured.

1.6 References

[1] G. F. Knoll, Radiation Detection and Measurement, 3rd ed, New York: John Wiley & Sons, Inc., 2000.

[2] M. Brigida, "A new Monte Carlo code for full simulation of silicon strip detectors", Nuclear Instruments and Methods in Physics Research A, vol. 533, pp. 322–343, 2004.

[3] R. T. Tung, "Schottkey Barrier and Ohmic Contact to Silicon", Contact to Semiconductors, Fundamentals and Technology, Ch. 3, Noyes Publication, 1993.

[4] J. Kemmer, "Fabrication of low noise silicon radiation detectors by the planar process," NIM 169, 1980.

[5]] J. Kemmer, "Improvement of detector fabrication by the planar process," NIM A 226, 1984.

[6] E. H. M. Heijne et al., "A silicon surface barrier microstrip detector designed for high energy physics," NIM 178, 1980.

[7] M. Caccia et al., "A Si strip detector with integrated coupling capacitors," NIM A 260, 1987.

[8] N. Bingefors et al., "The Delphi Microvertex detector," NIM A 328, 1993.

[9] C. Bozzi et al., "First-year experience with the BaBar silicon vertex tracker," NIM A 473, 2001.

[10] B. Sanghera et al., "Digital beta autoradiography using silicon microstrip detectors," NIM A 310, 1991.

[11] M. Overdick et al., "A 'Bioscope' system using double-sided silicon strip detectors and self-triggering read-out chips," NIM A 392, 1997.

[12] A. Czermak et al., "A new 2-dimensional high resolution Si detector for b-and g-radiography," NIM A 360, 1995.

[13] F. Arfelli, "Silicon detector for synchrotron radiation digital radiology," NIM A 353, 1994.

[14] P. Jalas et al., "New results with semiconductor drift chambers for X-ray spectroscopy," IEEE Trans. Nucl. Sci. Vol. 42, 1994.

[15] J. Kemmer and G. Lutz, "New detector concepts," NIM A 253, 1987.

[16] H. Soltau et al., “Fabrication, test and performance of very large X-ray CCDs designed for astrophysical applications,” NIM A 439, 2000.

[17] E. H. M. Heijne, “Semiconductor micropattern pixel detectors: a review of the beginnings,” NIM A 465, 2001.

[18] T. Mouthy, “Silicon pixel detector research and development,” NIM A 368, 1995.

[19] K. S. Kump, “Fast imaging of a 41cm a-Si flat panel detector for radiographic applications,” Proceedings of SPIE, Physics of Medical Imaging, Vol. 4320, 2001.

[20] Watanabe et al., “Development and evaluation of a portable amorphous silicon flat-panel x-ray detector,” SPIE Vol. 4320, 2001.

[21] D. L. Lee et al., “Imaging performance of a direct digital radiographic detector using Selenium and a thin-film-transistor array,” Computer Assisted Radiology, 1996.

[22] M. Yaffe and J. A. Rowlands, “X-ray detectors for digital radiography,” Phys Med. Biol. 42, 1997.

[23] J. M. Henry et al., “Solid state detectors for digital mammography,” SPIE Vol. 2432, 1995.

[24] M. Cuzin et al., “Preliminary characterization of a new hybrid structure with CdTe: X-ray imaging capabilities,” SPIE Vol. 2278, 1994.

[25] E. Samei et al., “DQE of direct and indirect digital radiography systems,” SPIE Vol. 4320, 2001.

[26] E. Bertolucci et al., “X-ray imaging using a pixel GaAs detector,” NIM A 362, 1995.

[27] B. E. Patt, “Development of a mercuric iodide detector array for medical imaging applications,” NIM A 366, 1995.

[28] C. Scheiber, “CdTe and CdZnTe detectors in nuclear medicine,” NIM A 448, 2000.

[29] R.F. Pierret, Semiconductor Fundamentals, 2nd Ed. Addison-Wesley, p. 65.1988.

[30] M. Jeong, D. A. Lawlor, and M. D. Hammig, “Optimization of the Position Resolution in Semiconductor Detectors”, 2007 IEEE Nuclear Science Symposium Conference Record, pp. 1456-1462.

[31] D.J. Griffiths, Introduction to Electrodynamics, 2nd Ed. p. 65, Prentice-Hall, 1989.

[32] K. Change, Microwave Solid-State Circuits and Applications, John Wiley & Sons, p. 56.1994.

[33] Bharathi Bhat and Shiban K. Koul, “Stripline-like Transmission Lines for Microwave Integrated Circuits”, Wiley Eastern Limited New-Delhi, 1980.

[34] E. H. Fooks and R.A. Zakarevicius, “Microwave Engineering using Microstrip Circuits”, Prentice Hall of Australia, 1990.

[35] Jia-Sheng Hong and M. J. Lancaster, “Microstrip Filters for RF/Microwave Applications,” John Wiley & Sons, Inc, 2001.

[36] H. R. Kaupp, “Characteristics of Microstrip Transmission Lines”, IEEE Trans., Vol EC-16, No. 2, April 1967.

[37] H. A. Atwater, “Tests of Microstrip Dispersion Formulas”, IEEE Trans., Vol. MTT-36, pp.619-621, March 1988.

[38] M. Kirschning and R. H. Jansen, “Accurate model for effective dielectric constant of microstrip and validity up to millimeter-wave frequencies”, Electronics Letters, Vol. 18, pp. 272-273, 18 March 1982.

[39] William J. Getsinger, “Microstrip Dispersion Model”, IEEE Trans., Vol. MTT-21, pp.34-39, January 1973.

[40] T. C. Edwards, “Foundations of Microstrip Circuit Design,” pp. 98-140, May 1987. Omar M. Ramahi, “Analysis of Conventional and Novel Delay line: A Numerical Study”, November, 2003.

[41] Terence C. Edwards and Roger P. Owens, “2-18 GHz Dispersion Measurements on 10-100 Microstrip Lines on Sapphire”, IEEE Trans., Vol. MTT-24, pp.506-513, August 1976.

[42] K. C. Gupta, Ramesh Garg, Inder Bahl, and Prakash Bhartia, Microstrip Lines and Slotlines, 2nd edition, Artech House, Norwood, MA, 1996.

[43] H. Guckel, P. A. Brennan, and I. Palocz, "A parallel-plate waveguide approach to microminiaturized, planar transmission lines for integrated circuits," IEEE Trans. Microwave Theory Tech., vol. 15, no. 8, pp. 468-476, Aug. 1967.

[44] H. Hasegawa, M. Furukawa, and H. Yanai, "Slow wave propagation along a microstrip line on Si-SiO_2 system,” Proc. IEEE, vol. 59, no. 2, pp. 297-299, Feb. 1971.

[45] D. Jager, "Slow-wave propagation along variable schottky-contact microstrip line," IEEE Trans. Microwave Theory Tech., vol. 24, no. 9, pp. 566-573, Sept. 1976.

[46] Michael Steer, Microwave and RF Design: A Systems Approach, Scitech Publishing, INC. 2009.

[47] I. J. Bahl and R. Garg, "Simple and accurate formulas for microstrip with finite strip thickness," Proc. IEEE, 65, pp. 1611–1612, 1977.

CHAPTER II

MODELING

2.1 Development of Numerical Modeling Tool

A charged particle- whether induced or directly incident- that traverses through a semiconductor creates electron-hole pairs along its trajectory. In general, both custom and commercial Monte Carlo codes were used to calculate the straggling functions that describe the distribution of energy loss inside the semiconductor detectors under study. For the results presented in this Chapter, we will generate simple deterministic charge distributions in order to better elucidate the effects of thermal diffusion and Coulomb drift. In order to predict the response of a silicon detector to ionizing gamma radiation in particular, a simulation of a detector was developed using the MAXWELL 3D electric field simulation package [1]. The charge carriers are produced by the primary electron as well as during the subsequent energy losses of the secondary electron-hole (e-h) pairs. The primary and hot-secondary carriers lose part of their kinetic energy by scattering and another part by ionization, and they are tracked until all particles have residual energies below an adjustable threshold, which is set to 1 keV when using GEANT4 [2], and when using a custom Matlab code, is set to the ionization energy, 3.62 eV [3]. Thus, the energy loss process simulation provides the spatial distribution of the energy deposition due to all relevant processes.

Maxwell 3D is then used to calculate the electric field and weighting potential inside the detector due to the applied bias, and custom Matlab codes are used to convert the charge motion into induced signals on the electrodes bounding the active volume. For circuit development purposes, the amplification and shaping of those signals is then simulated using PSpice circuit modeling tools, although for all of the results presented, the circuit effects simulated the (detector-integrated) induced charge or the detector's

induced current are not included. The implementation of these tools will be briefly discussed below.

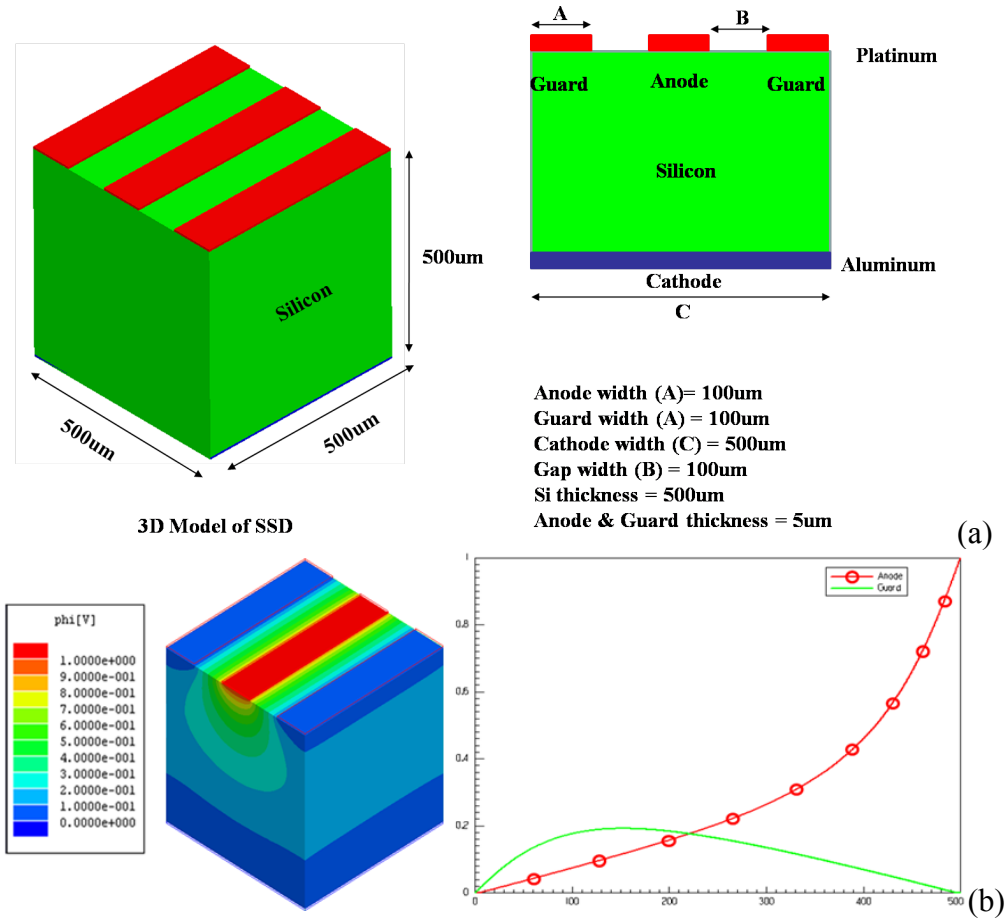


Fig. 2.1: (a) Detailed geometry for a diagnostic Si detector with a central strip anode and two grounded guard-strips. (b) Weighting potential over the total volume in 3D, and plotted through the center of detector.

For the sake of demonstrating the operation of the code, consider Fig. 2.1a, which shows a truncated Si detector with a central anode strip and two grounded guard strips, as modeled using Maxwell 3D. Fig. 2.1b shows the depth variation of the weighting potential for the aluminum anode and the 3D view of the total weighting potential values. For any given initial charge distribution - whether derived from the Bethe formula or by some other means- we can use the potential and field information from Maxwell 3D as inputs to the simulation code and track the motion of the electron-hole pairs.

2.2 Numerical Simulations given known Material Behavior

As discussed in Chapter 1, one reason that detailed information concerning the trapping characteristics at material interfaces is insufficiently predictive is that the assessment tools evaluate a composite signal from a surface over which the barrier characteristics can abruptly vary from point to point. The second reason is the high multiplicity of sources which can contribute to charge loss and trapping- dislocations, stress-induced states, impurities, and dangling bonds, for instance. We have therefore endeavored to develop probing tools for which the charge motion can be tracked as it encounters various material features. The modeling of the pure motion of the charges is topic of this sub-section.

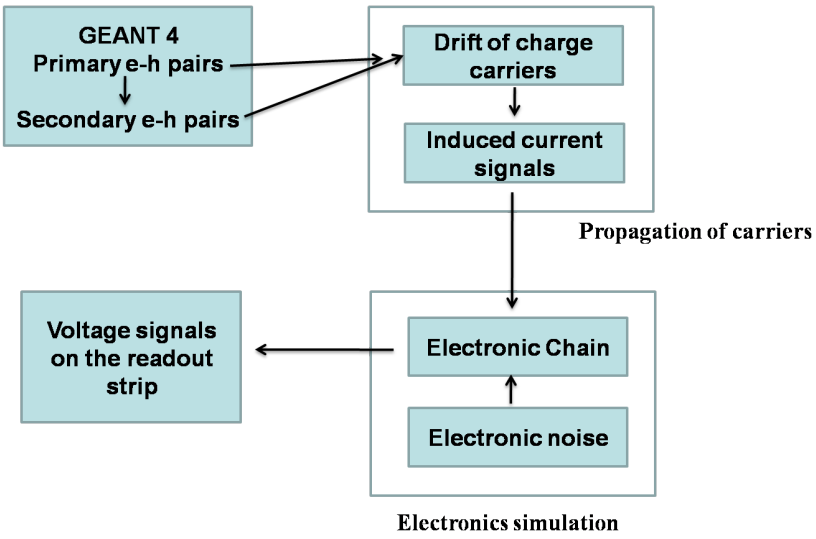


Fig. 2.2: General diagram of different simulation components used to model the detector.

The numerical and analytical modeling tools can expedite the development of different detector configurations and were put to use during our research to develop fine position-sensing schemes. If the behavior of the underlying material fails to conform with the model- as is typical with semiconductors undergoing various processes- then it is valuable to have a physics-based understanding of the factors which can cause the divergence. Nevertheless, for the development of silicon devices as a means by which that physics can be investigated, the models are of high utility. The diagram of the simulation tools used to convert the radiation-detector interaction into an electrical pulse

is shown in Fig. 2.2. A charged particle- whether induced or directly incident- which traverses through a silicon layer leaves an ionization track along its trajectory. The Monte Carlo Code GEANT4 is used to calculate the straggling functions that describe the distributions of energy loss inside our semiconductor detector. In order to predict the response of a silicon detector to ionizing gamma radiation in particular, a simulation of a detector was developed using the GEANT4 detector simulation toolkit and the MAXWELL 3D electric field simulation package.

The charge carriers are produced both in the primary collision and by the subsequent energy losses of the secondary electron-hole (e-h) pairs. All primary electrons and holes lose part of their kinetic energy by scattering and another part by ionization. The secondary charge carriers produced by the primary particle are followed via the same procedure until all particles have residual energies below the threshold. Thus, the GEANT4 simulation provides the spatial distribution of the energy deposition due to all relevant processes.

Maxwell 3D is then used to calculate the electric field and weighting potential inside the detector, and custom codes- one using Matlab, one using C- are then used to convert the charge motion into induced signals on the electrodes. The amplification and shaping of that signal is then simulated using PSpice circuit modeling tools. The implementation of these tools will be briefly discussed below.

2.2.1 Monte Carlo Modeling of the Radiation Impact Events

The following geometry, consistent with our experimental detectors, was used in GEANT4: a 500 μm thick wafer of silicon, 1 x 1 cm square, placed in the center of the world volume, which is filled with air. Furthermore, electrode strips of various sizes are placed on the top and bottom sides of the detector. No electric field is applied across the detector at this stage, since its effect on the ionizing radiation and secondaries is negligible. GEANT4 allows the user to specify which physics processes and particles to simulate. For our simulations to date, only electromagnetic processes were simulated, and only gamma rays and electrons/positrons were used. The standard electromagnetic physics processes in GEANT4 are valid only down to 1 keV, so the Penelope Low-Energy Electromagnetic Physics model was used instead. This model has been validated down to 250 eV for most processes (bremsstrahlung, photoelectric effect, Compton

scattering, pair production, Coulomb scattering, ionization, etc.). The secondary range cut was set to 1 nm. This means that when GEANT4 produces a secondary particle, it will continue to track it only if its expected range is greater than this cut. Otherwise, the particle is killed, and its energy is deposited at the position of its creation.

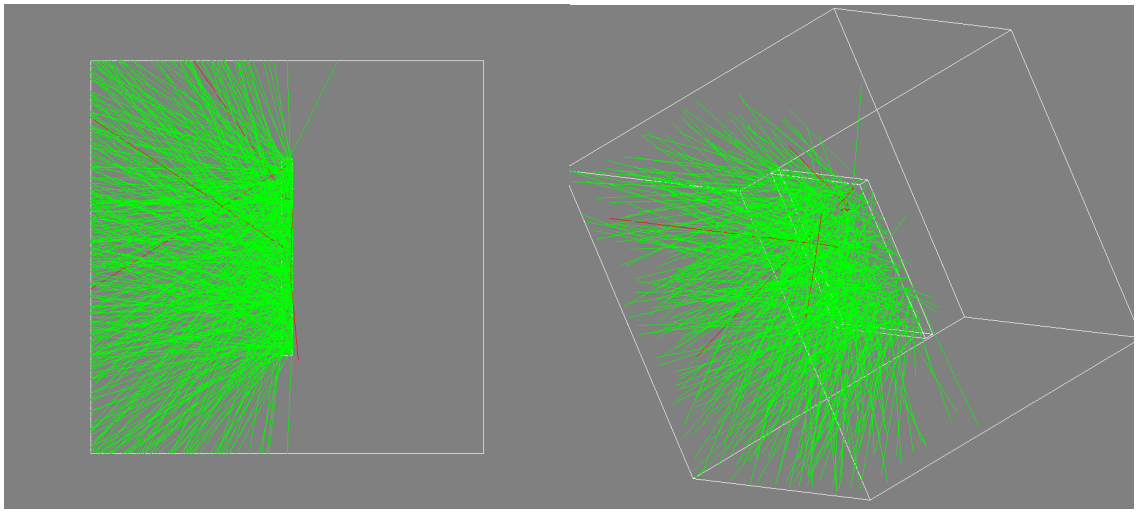


Fig. 2.3: Three-dimensional view of (500) primary gamma rays (green) uniformly incident on the detector face as well as a few fast secondary electrons (red).

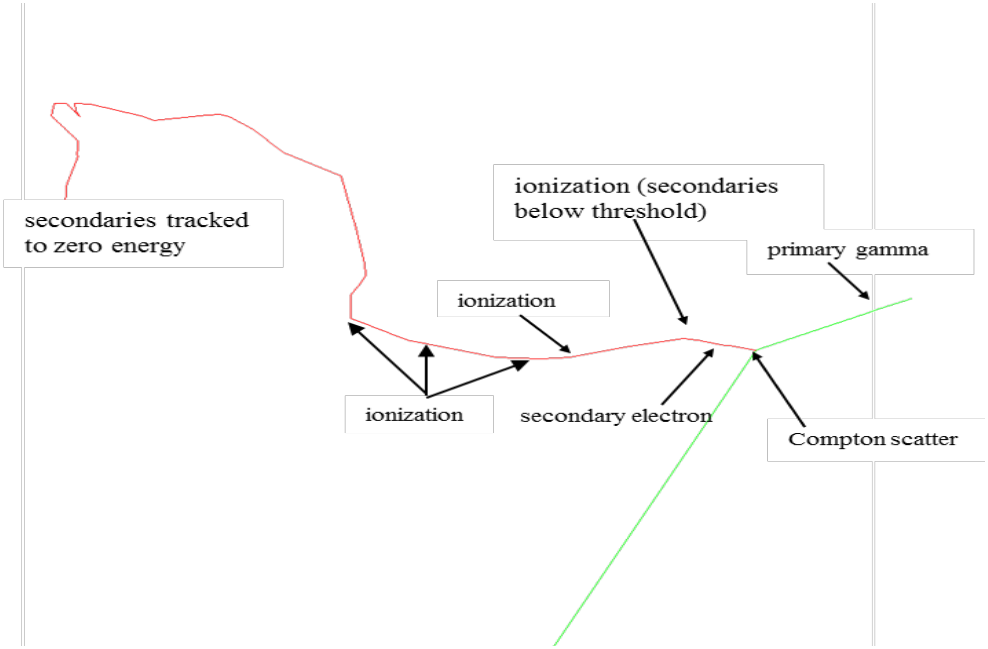


Fig. 2.4: Close-up view of Compton-scatter event, with processes and particles labeled.

The General Particle Source was used to generate primary gamma rays uniformly over a square a few microns above the top face of the detector. The momenta of the particles follow a cosine distribution which is uniform in 2π. Fig. 2.3 shows a collection of primary gamma rays (green) as well as a few secondary electrons (red). Figure 2.4 shows a close-up of one Compton-scatter event with the various processes and particles labeled. In the simulation, the silicon wafer is designated a "sensitive detector," which means that one can record hits within the volume. A hit is a detailed record of an interaction, and in this case includes the energy deposited by the particle and the location of the interaction. For example, in Fig. 2.4, a hit is recorded at each interaction point (Compton scatter and subsequent ionizations of the secondary electron). This information is written to an external file and used with the information from the MAXWELL electric field simulation to calculate the induced charge on the electrodes of the detector.

2.2.2 Numerical Modeling of the Electric Field Environment

As mentioned above, MAXWELL 3D is used to model the electric field and weighting potential for a specific electrode design. This has been done by dividing the detector volume into elementary cells and solving for the electric field via the scalar potential using Poisson's equation:

$$\nabla^2 \varphi = -\frac{\rho}{\varepsilon} \tag{2.1}$$

An example is shown in Fig. 2.5, which contains 36 strips electrodes on either side as well as a guard ring. By specifying the applied voltages to the electrodes, as well as boundary conditions, MAXWELL can solve for the electric field and potential and output the values at specified locations to an external file. An example plot of the electric potential and field vector are shown in Fig. 2.6, where 100 V was placed on one set of electrodes and all others (and the guard rings) were held at 0 V.

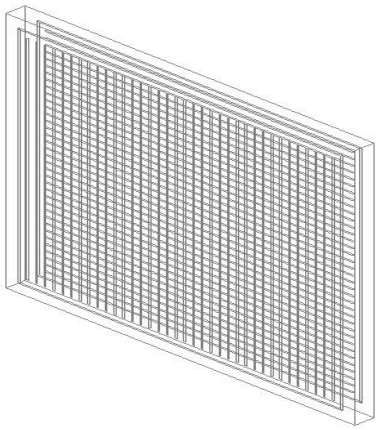


Fig. 2.5: Three-dimensional view of the electrode pattern used in MAXWELL.

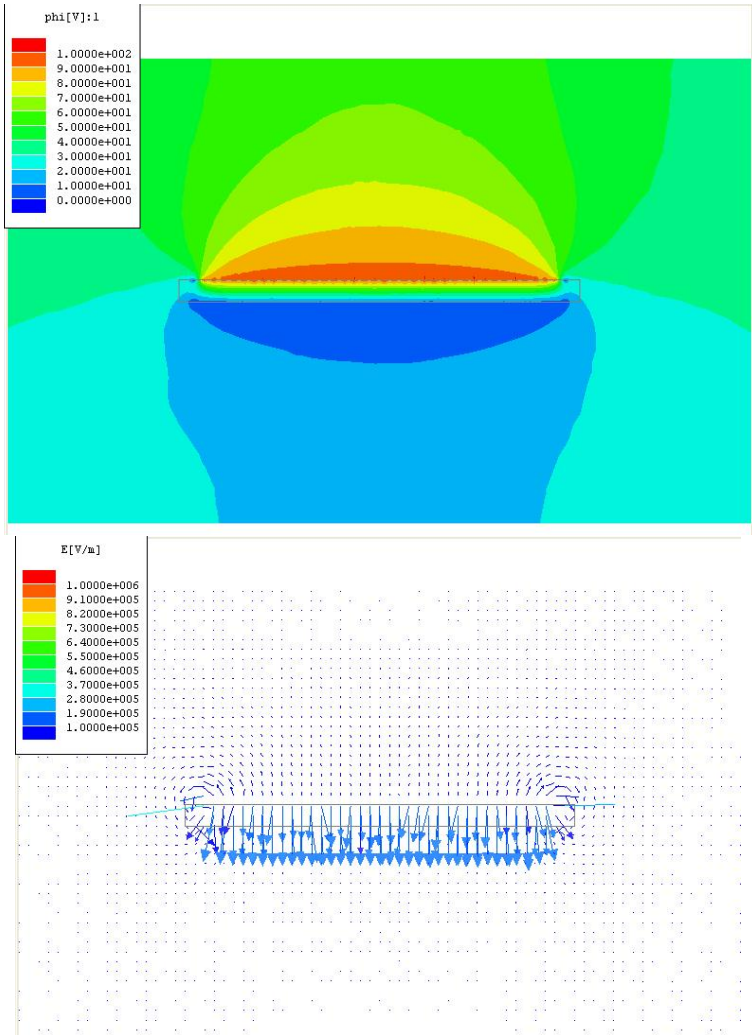


Fig. 2.6: As evaluated with MAXWELL, the electric potential (top) and vector field (bottom) for the silicon detector shown in Fig. 2.5.

The output of the GEANT4 and MAXWELL simulations can be combined to calculate the induced charge on the electrodes of the detector, using the Shockley-Ramo theorem. The MAXWELL program provides the weighting fields, as well as the electric field (which gives the drift velocity). Using the deposited energy information from GEANT4, one can then divide by the energy per electron-hole pair in silicon (w = 3.62 eV) to determine the number of carriers to start drifting at the position of the energy deposition.

2.2.3 Modeling of the Charge Carrier Transport and Signal Formation

The motion of the charge carriers can be described by the equation $\vec{v} = \mu\vec{E}$, where v is the drift velocity, E is the electric field, and μ is the mobility that depends on the electric field. During their drift, electrons and holes also diffuse by multiple collisions, and their simulated distribution follows a Gaussian law:

$$\frac{dN}{N} = \frac{1}{\sqrt{4\pi Dt}} exp\left(-\frac{r^2}{4Dt}\right) dr \tag{2.2}$$

where dN/N is the fraction of particles in the line element *dr* at a distance r from the production point and after a time t. The diffusion coefficient, D, is related to the mobility by the Einstein relation $D = \frac{kT}{e}\mu$, where T is the absolute temperature, k is Boltzmann's constant and e is the electron charge. If the mobility is assumed to be a constant, the root-mean-square of the distribution is given by $\sigma = \sqrt{2Dt}$.

In our simulation code the diffusion effect is taken into account during the solution of the equation of motion. After each step, a check is done to verify whether the carrier is still inside the detector volume: if the carrier is out of the detector volume, it will be stopped; otherwise a new step is performed.

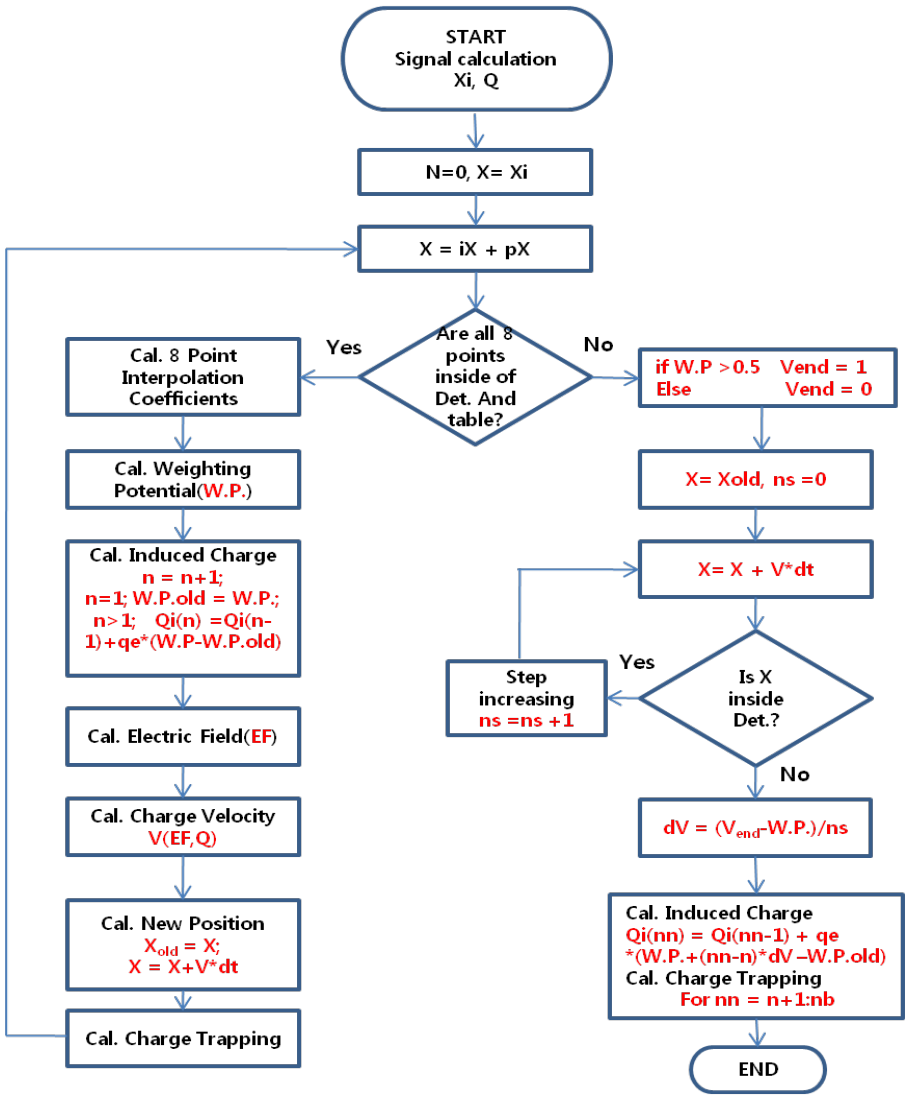


Fig. 2.7: The sequence diagram of our simulation code.

The induced currents are evaluated during the motion of the carriers, and the total currents are computed by adding the contributions from all carriers. A typical detector front-end consists of a preamplifier, which provides gain, followed by a shaper, which tailors the overall frequency response for optimum noise performance, and limits the pulse length to accommodate the signal rate. We have considered a standard front-end in which each induced current signal $i_k(t)$ is fed into the electronic chain formed by a preamplifier followed by a shaper. In this way we can simulate the voltage output event by event. Fig. 2.7 shows the sequence diagram of our simulation code.

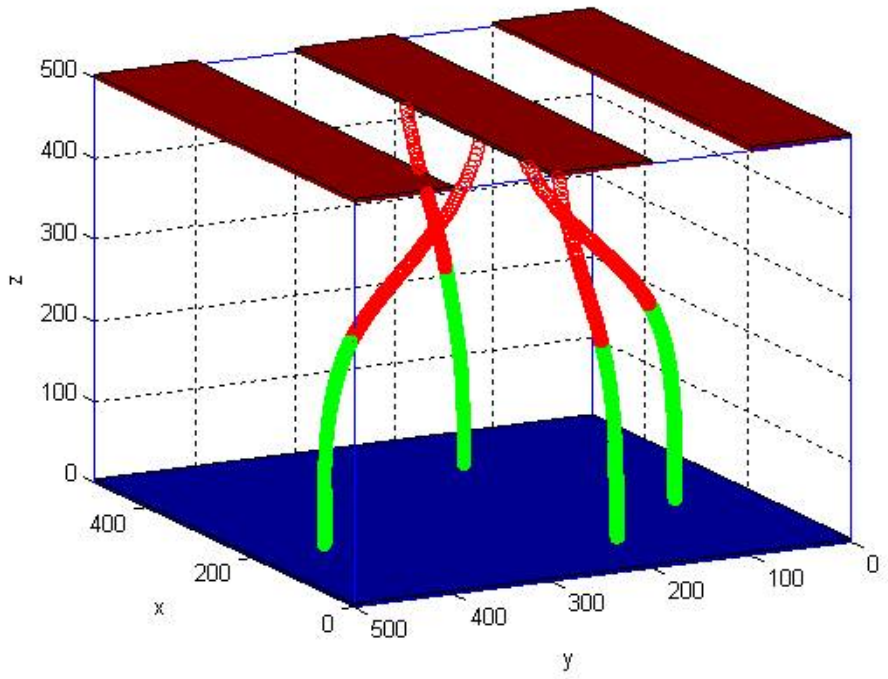


Fig. 2.8: Example picture of simulation result with four interaction positions, in which the anode plane is divided into a central strip and two guard strips.

For any given initial charge distribution therefore- whether derived from GEANT4 or by some other means- we can use the potential and field information from MAXWELL as inputs to the simulation code and track the motion of the electron-hole pairs. For instance, Fig. 2.8 shows the carrier tracks for four different electron-hole pairs, in which the (red) electrons drift toward the strip anode and the (green) holes are collected on the cathode. This information can then be combined with the field information to deduce the charge or current induced on the electrodes of interest, examples of which are shown in Figs. 2.9 and 2.10 for a planar device. The single-strip geometry of Fig. 2.8 is used to track carrier flight through different material regions, and fine lateral resolution can be achieved with small pitch widths; however, for investigating microscopic effects one must finely focus the exciting source; and further, in order to isolate the interaction in three-dimensions, small strips or pads must be used. Thus, we have endeavored to develop novel readout schemes discussed below.

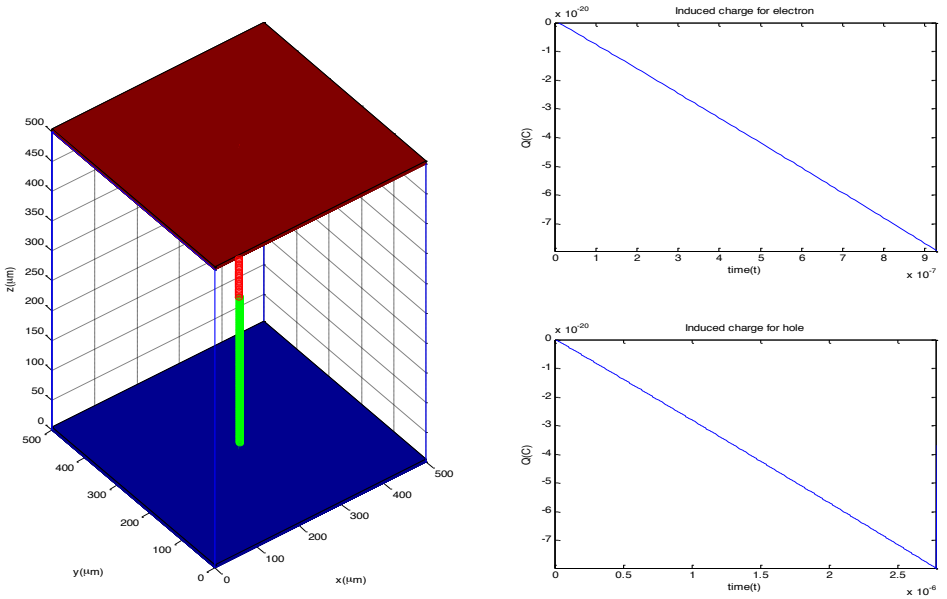


Fig. 2.9: Example picture of simulation result- including the induced charge signal- with single interaction position for which the electron (red) is collected at the anode and the hole (green) is collected at the cathode.

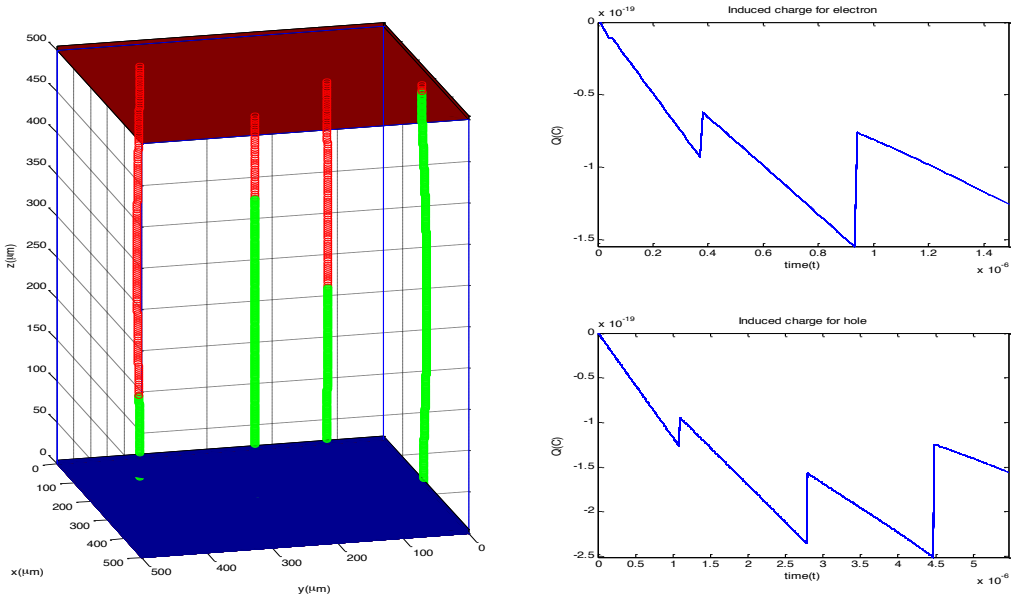


Fig. 2.10: Example picture of simulation result with four interaction positions. Note that the discontinuities in the induced charge plots are due to charge collection, and they therefore no longer participate in the signal formation.

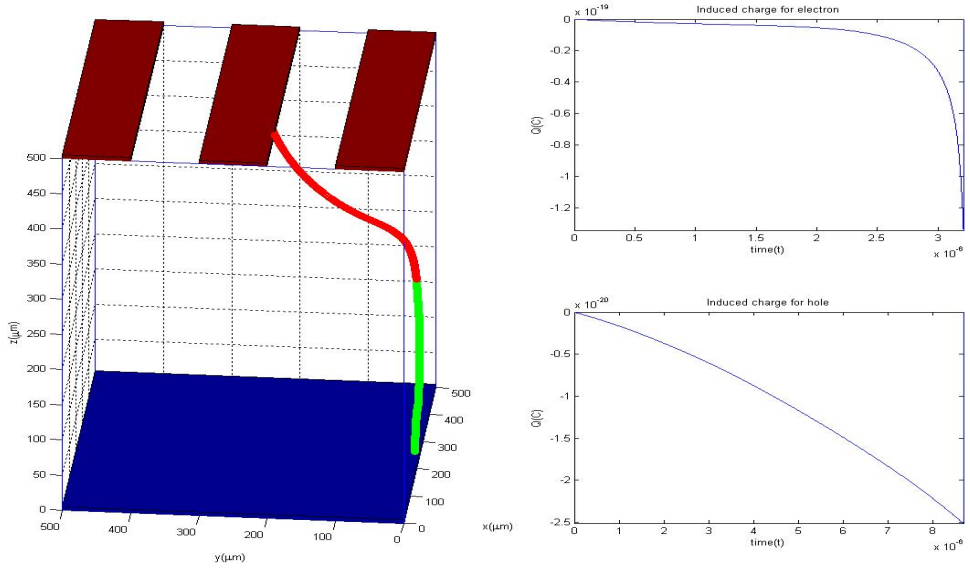


Fig. 2.11: Example of simulation result with an interaction near the edge but for which the electron (red) is collected at the anode and the hole (green) is collected at the cathode.

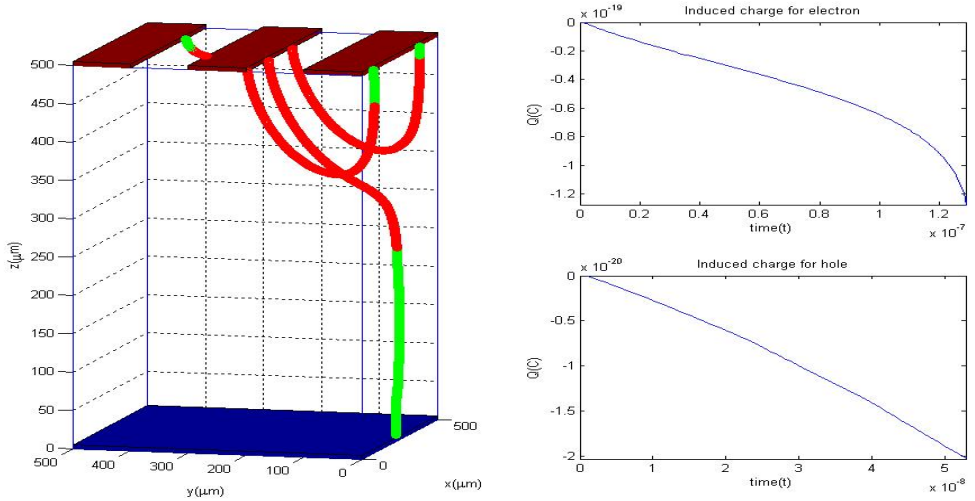


Fig. 2.12: Example picture of simulation result with four interaction positions. The induced charge plots are for the charge-pair creation near the guard ring in which the hole is collected at the guard ring.

For a partitioned device, which is of greater interest for charge tracking, Figs. 2.11 and 2.12 show examples of simulation results in which the charge trajectory and induced charge is calculated. Note in the following figures that the induced charge is plotted, which experimentally corresponds to the voltage developed across the detector

because the induced charge integrates across the capacitance of the detector. We are using both current-sensitive and voltage-sensitive preamplifiers. The shape of the current pulse is derived from the slope of the induced-charge pulses shown below.

In Fig. 2.12, three of the tracks show that if the interaction occurs sufficiently near the guard ring, then the hole is collected at the guard ring rather than at the cathode. The induced-charge or current pulse will then reflect this sequence, which doesn't impair the trajectory determination if one has mapped the field in the detector and uses that information to deduce the charge trajectory. However, the relationship between the pulse duration and depth at which the carrier was created is no longer monotonic.

One would prefer to physically eliminate the effect of those underlying discontinuities which trap charges; however, one can correct for trapping and delay in the transport by extracting the charge trajectories from the induced current signals. This method is used, in modified form, by one of our colleagues (Prof. Zhong He) in the context of CZT, but at the cost of high readout multiplicity. In the context of this and other research, we have endeavored to develop that capability but with a limited number of readouts, as described in the next section.

Table 2.1: Relative magnitude of drift force (due to 500 V acting on a planar detector that is 550 μm thick) compared with Coulomb force due to N charges concentrated at a point and separated from a single test charge a distance x away.

		$F_{drift}/F_{coulomb}$ X			
energy (keV)	N	1 nm	100 nm	1 micron	10 micron
10	2762	2.3E-07	2.3E-03	2.3E-01	2.3E+01
100	27624	2.3E-08	2.3E-04	2.3E-02	2.3E+00
1000	276243	2.3E-09	2.3E-05	2.3E-03	2.3E-01

The extraction of the recoil electron direction will ultimately be limited by those stochastic effects that significantly contribute to the charge motion; most notably, thermal diffusion. Coulomb drift can also be important for highly ionizing particles. Coulomb drift refers to the dynamic charge motion due to the electromagnetic forces of the space charge created during the radiation event. At small distances (< 10 μm), the electrostatic Coulomb force can, in principle, be much greater than the drift force. For instance,

suppose that we compare the drift force to the Coulomb force that acts on a test particle due to all of the other like charges concentrated at a point; thus, this is far and away the worst-case scenario. If incident energy E is converted into electrons (or holes, but not both), then we expect about E/w charges to be created. Assuming that they are at a point and separated from the test charge by the distances shown, one sees from Table 2.1 that for separations less than 1 micron, the drift force is smaller than the Coulomb force, and beyond a 10 micron separation the drift force begins to dominate the behavior.

Fortunately, the Coulomb effect is mitigated by two factors relative to the numbers implied in Table 2.1. First, the electrons are created alongside holes which dynamically shield the EM force during the initial drift phase. More importantly, the charges are less concentrated than that implied in Table 2.1, fast electrons losing about 1 keV per micron. Thus, a few hundred charges are within a micron of each other implying that their influence will not extend to those charges separated by 10 microns or more. Finally, the effect of Coulomb drift can, in principle, be removed via post-processing once the charge distribution is estimated, because the dynamics are not stochastic. Nevertheless, both Coulomb drift and diffusion, as explained below, can limit the position resolution if micron-scale sensing is achieved.

We modeled diffusion using two techniques, one found in the literature and based on the diffusion coefficient, the other based on the underlying physics in which the probability density function describing the random thermal motion is sampled to determine the random contribution to each charge's motion. As will be briefly shown, the effect of diffusion is significant, but its effect can be mitigated if the charges are drifting with alacrity.

2.3. Monte Carlo Modeling of the Energy Loss and the Induced Charge

The following geometry, consistent with our experimental detectors, was used in our simulation: a 500 μm thick wafer of silicon, 1 x 1 cm square, placed in the center of the world volume, which is filled with air. Furthermore, 5 μm thick planar electrodes composed of Al and Pt are placed on the top and bottom sides of the detector, respectively. During charge-pair *formation*, no electric field is applied across the detector since its temporal effect on the ionizing radiation and secondary by-products is negligible. For our simulations to date, only electromagnetic processes were simulated

and only gamma rays and electrons/positrons were used. As mentioned above, Maxwell 3D is used to model the electric field and weighting potential for a specific electrode design. This has been done by dividing the detector volume into elementary cells and solving for the electric field via the scalar potential using Poisson's equation Eq. 2.1. The output of a particle-creation code- whether deterministic or stochastic- and Maxwell 3D simulations can be combined to calculate the induced charge on the electrodes of the detector, employing the Shockley-Ramo theorem:

$$Q(t+1) = Q(t) + q_0(WP_{new} - WP_{old}) \tag{2.3}$$

where Q(t) is time-dependent induced charge, WP is weighting potential and q_0 can be expressed as n_0e where n_0 is the number of electron hole pairs and e is the electronic charge [4]. Based on the structures and the applied bias, the Maxwell program provides the weighting fields as well as the electric field, from which the drift velocity is derived. Using the deposited energy information, one can then divide by the energy per electron-hole pair in silicon (w = 3.62 eV) to determine the number of carriers to start drifting at the position of the energy deposition. For primary electrons, the specific energy loss is expressed with the Bethe formula:

$$-\frac{dE}{dx} = \frac{4\pi e^4 z^2}{m_0 v^2} NZ\left[\ln(\frac{2m_0 v^2}{I}) - \ln\left(1-\frac{v^2}{c^2}\right) - \frac{v^2}{c^2}\right] \tag{2.4}$$

where v is the primary particle's velocity, ze is its charge, e is the electronic charge, and N and Z are the number density and atomic number of the target atoms, respectively [3]. I is the average excitation energy of the target and is calculated as 13·Z [eV·Å] for the Si detector [5].

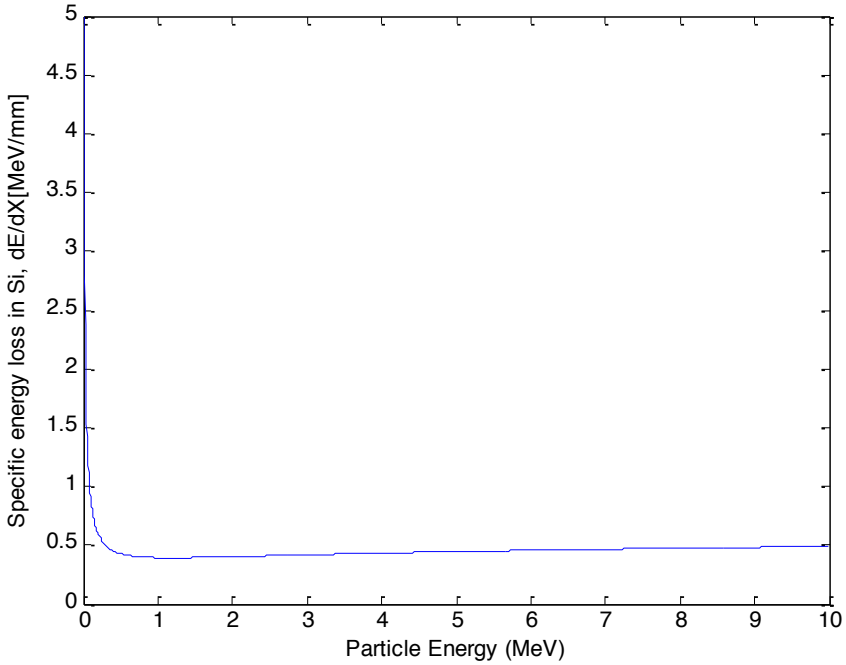
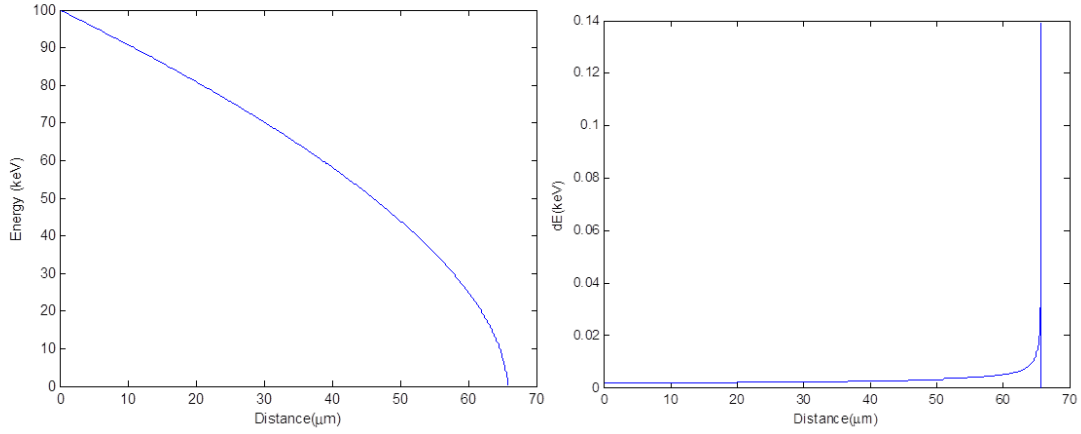


Fig. 2.13: In silicon, the specific energy loss of an electron as a function of its energy.



(a) (b)

Fig. 2.14: (a) The specific energy loss along the primary electron track for a 100 keV electron, as calculated with the Bethe formula. (b) Variation in the electron's energy along the track of an electron with initial energy 100 keV.

For the deterministic studies of radiation interactions in Figs. 2.13 and 2.14, Eq. 2.4 is used to derive the specific energy loss of the primary particle's energy in the Si detector. As shown in Fig. 2.13, the average specific energy loss is approximately 0.5 MeV/mm at high energies. Note that the radiative component of the energy loss, due principally to bremsstahlung radiation, is negligible compared to the ionization losses due to the low Z of Si.

Figs. 2.14a and 2.14b show the graphs of the specific energy loss and energy variation along the primary electron's track. For steps further along the track length, the primary particle's energy decreases more precipitously due to more effective momentum exchange with the charge cloud in the material.

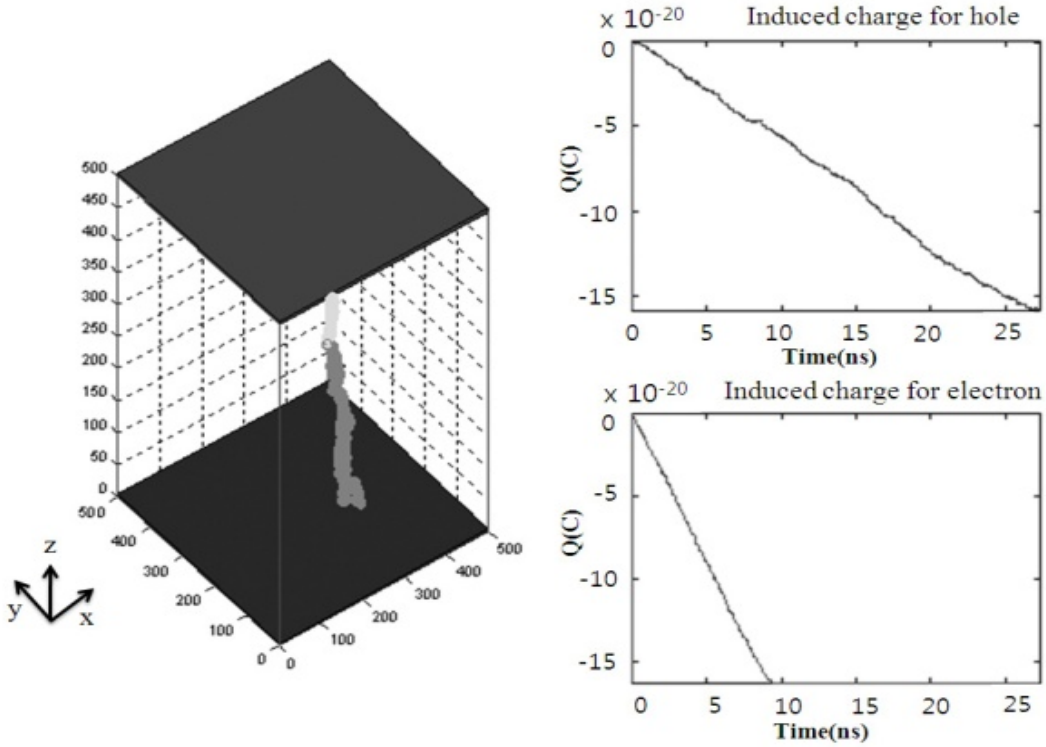
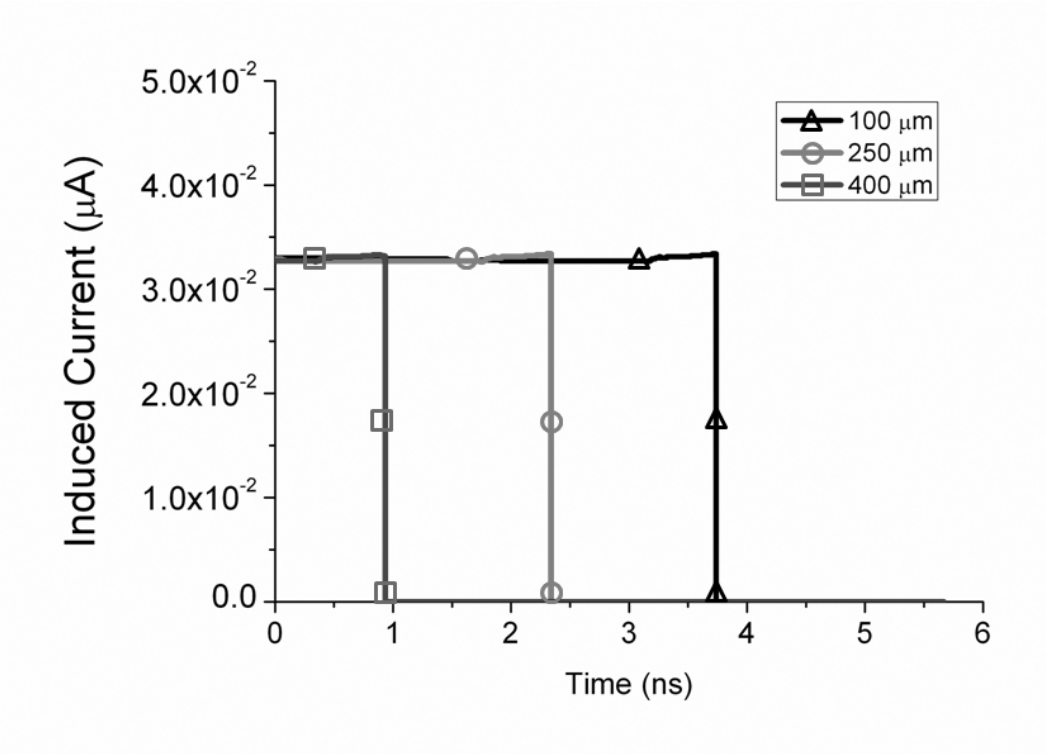
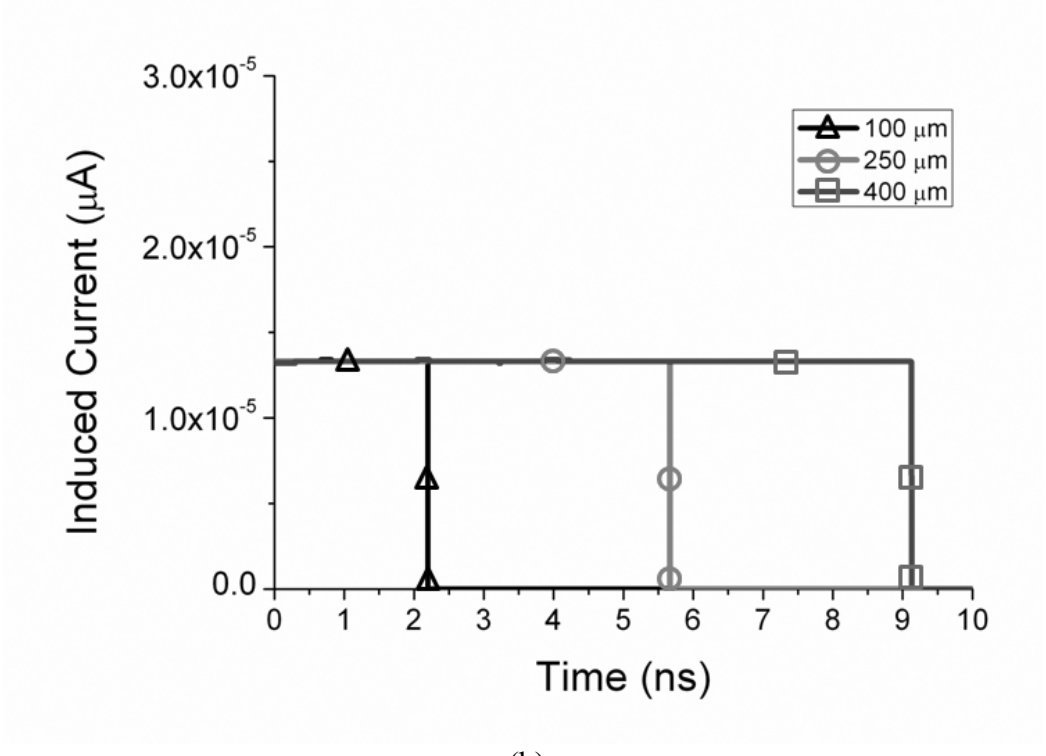


Fig. 2.15: Track of electrons (light) and holes (dark), including the effects of thermal diffusion, along with the induced charges on the anode due to their drift, when the bias voltage is 100 V.



(a)



(b)

Fig. 2.16: Induced current for a single electron (a) and a single hole (b) as the interaction position is varied y along the z-axis (from 100 μm to 400 μm) when the bias voltage is 500 V.

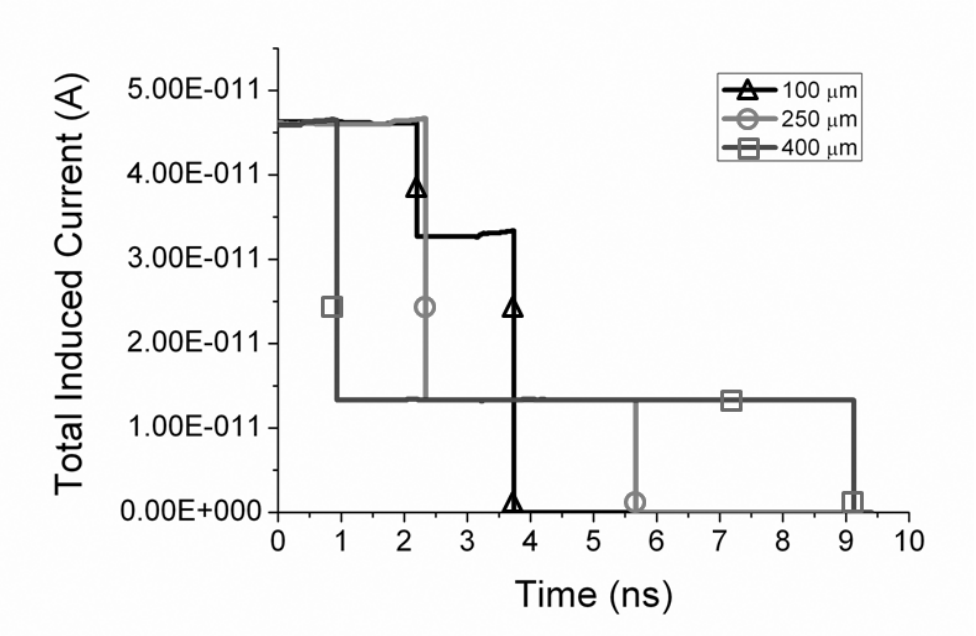


Fig. 2.17: Total induced anode current for a single e-h pair, as the interaction position is varied along the z-axis (from 100 μm to 400 μm) when the bias voltage is 500 V.

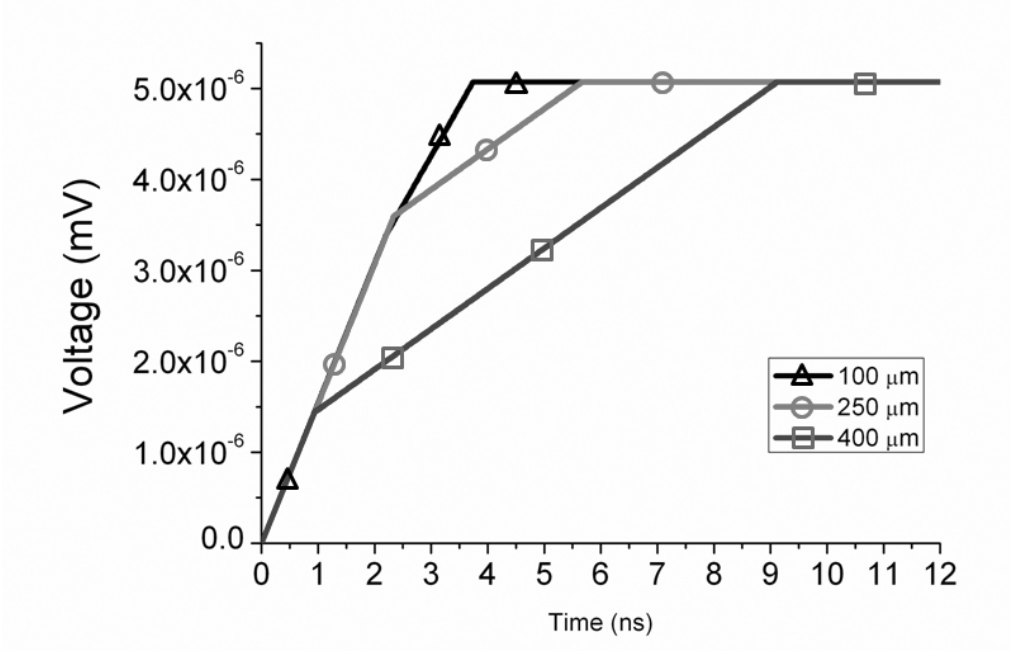


Fig. 2.18: Calculated voltage output from single e-h pair created at different depth positions due to the current when the bias voltage is 500 V.

Fig. 2.15 shows a typical numerical simulation result in which the carrier trajectories and induced charges are calculated for two e-h pairs, both created at the detector's midpoint. The induced-charge or induced-current pulse will reflect the carrier transport, and they can thus be used for position determination if one has mapped the field in the detector.

The shape of the current pulse is derived from the slope of the induced-charge pulses. Fig. 2.16 shows the variation in the induced current for an electron (a) and hole (b) pair, in which the depth (z-position) is varied (from 100 μm to 400 μm), and Fig. 2.17 shows the combined induced charge from both carriers when the bias voltage is 500V. In each case the electron and hole drift toward the anode and cathode respectively, and the charge and current waveforms provide a measure of the depth of interaction via the drift-time because of the known drift velocity as derived from the electric field, which is established during the design and application of bias. For this as well as more complex electrode geometries, the carriers are simply sampling different sections of the field.

Note that the variations in the pulses in Figs. 2.16 and 2.17 reflect the non-uniformity of the field in the cubic planar detector. In particular, the slight rise in the current signals toward the ends of the pulses reflects the concentration of the E-field lines nearer the electrodes. Note in Fig. 2.18 that the induced charge is plotted, which experimentally corresponds to the voltage integrated across the capacitance of the detector. For partitioned electrodes, the pulse duration reflects not only the depth at which the carrier was created it is also weakly dependent on the carrier's position lateral to the truncated collecting element, as detailed in [6]. If electrode strips are connected in a delay-line configuration, then the delayed arrival-time of the pulse at the end of the line can reflect the lateral position at which the charge is collected, as described in [7]. Regardless of the position-sensing method and track-reconstruction algorithm employed, a successful reconstruction of the electron's trajectory can substantially enhance the angular resolution of gamma-ray imaging instruments, as shown in [8].

2.4. Monte Carlo Modeling of the Thermal Diffusion

A fundamental physical limitation to mapping a carrier's trajectory is its thermal stochastic motion. One method of accounting for thermal diffusion is to introduce spread in the carrier's arrival position, which follows a Gaussian law. At low fields, the diffusion

coefficient D is related to mobility by the Einstein relation; furthermore, in cubic semiconductors, it is reduced to a scalar quantity [9]. At high fields however, D becomes a field-dependent tensor which describes the diffusion process parallel and perpendicular to the electric field, with respect to the crystallographic direction of electric field [8].

In our atomistic approach, the diffusion effect is taken into account during the solution of the equation of motion. The positional increment can be decomposed into a sum of two terms: $d\vec{r} = d\vec{r_E} + d\vec{r_D}$, where $d\vec{r_E}$ and $d\vec{r_D}$ are the positional increments due to the electric field and thermal diffusion, respectively. The stochastic contribution to the motion can be derived as follows. From the equipartition theorem, the mean thermal energy of the carrier is equal to $3k_bT/2$; thus, the mean thermal velocity can be expressed as:

$$\frac{1}{2} m_{eff} v^2 = \frac{3}{2} k_b T \qquad \Leftrightarrow \qquad v = \left[\frac{3k_b T}{m_{eff}} \right]^{\frac{1}{2}} \tag{2.5}$$

where k_b is Boltzmann's constant, T is temperature in Kelvin, and m_{eff} is the effective mass of particle, which depends on the crystallographic direction. When the energy distribution is given by the Boltzmann distribution $P(E) = \frac{1}{k_b T} e^{-\frac{E}{K_b T}}$ and the particle has a kinetic energy of $\frac{1}{2} m_{eff} v^2$, then the probability density function of the particle's velocity can be expressed as:

$$P(v) = P(E) \frac{dE}{dv} = P(E) \times \frac{1}{2} m_{eff} 2v \frac{dE}{dv} = \frac{m_{eff} v}{k_b T} e^{-\frac{m_{eff} v^2}{2k_b T}} \tag{2.6}$$

In order to find the random variable of velocity in terms of the uniformly distributed random variable ξ, we applied the probability density function of velocity (P(v)):

$$\xi = \int_{v=0}^{v} P(v) dv = \int_{v=0}^{v} \frac{m_{eff} v}{k_b T} e^{-\frac{m_{eff} v^2}{2k_b T}} dv = -e^{-\frac{m_{eff} v^2}{2k_b T}} + 1 \tag{2.7}$$

$$\ln|1 - \xi| = -\frac{m_{eff} v^2}{2k_b T} \qquad \Leftrightarrow \qquad v_{eff} = \left[\frac{2k_b T}{m_{eff}} \ln\left| \frac{1}{1 - \xi} \right| \right]^{\frac{1}{2}} \tag{2.8}$$

Thus, the increment in the position due to diffusion is given by $\vec{r} = v_{eff}\Delta t\hat{r}$, where the direction $\hat{r}$ can be found by sampling the direction, defined by polar angle, θ and azimuthal angle, Φ.

For the presented results, a one picosecond time increment, Δt, was chosen in order to fully capture the *deterministic* carrier motion, smaller increments simply resulting in a finer sampling of the behavior. If the scattering process is rotationally symmetric, then the random variable for the azimuthal angle Φ can be expressed in terms of ξ as:

$$\xi = \int_0^{\Phi} \frac{1}{2\pi} d\Phi = \frac{\Phi}{2\pi} \qquad \Leftrightarrow \qquad \Phi = 2\pi\xi \tag{2.9}$$

Defining μ as cosθ, μ also can be expressed in terms of ξ as μ = 2ξ-1 (-1≤ μ ≤1). To get the scattered direction vector u'(w_x', w_y', w_z'), we converted from polar coordinates to Cartesian coordinates, as follows, for the case of $w_z = \pm 1$,

$$w_x' = \sqrt{(1-\mu^2)}\cos\Phi$$

$$w_y' = \sqrt{(1-\mu^2)}\sin\Phi$$

$$w_z' = w_z\mu \tag{2.10}$$

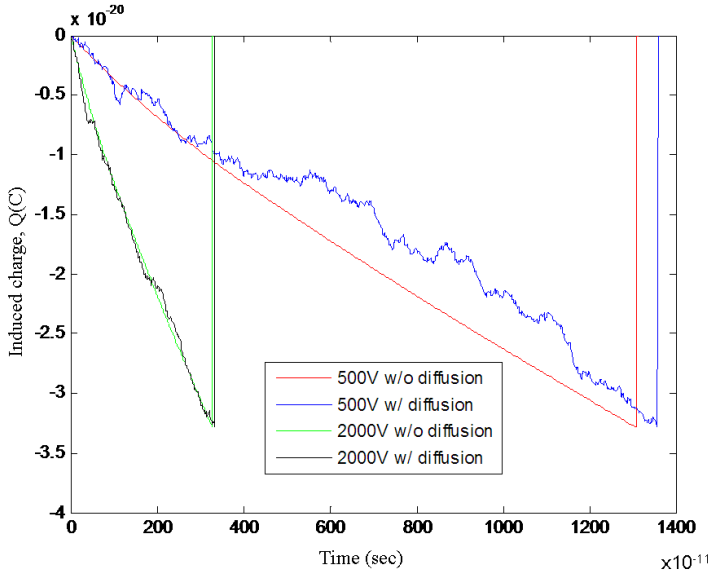


Fig. 2.19a: Plots of the induced charge Q due to the hole's drift with and without diffusion, and at two different biases, for the model detector configuration described above.

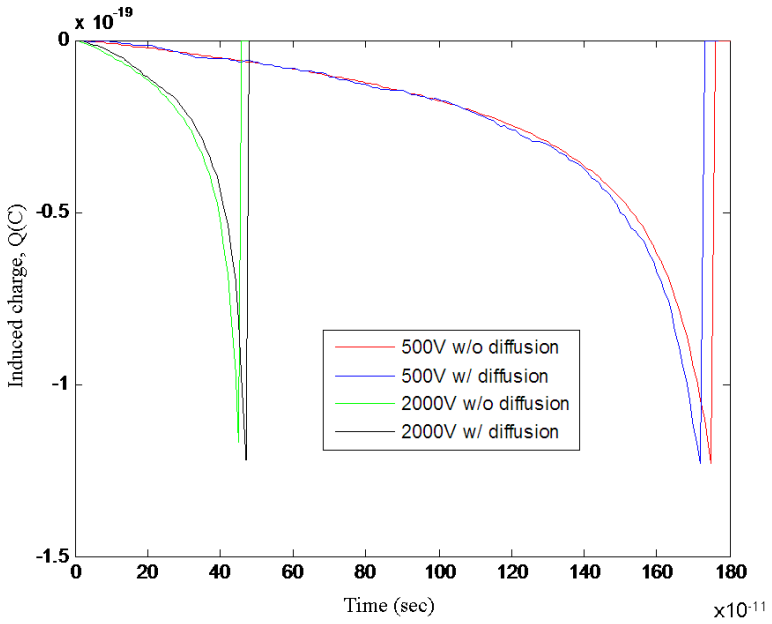


Fig. 2.19b: Plots of the induced charge Q due to the electron's drift with and without diffusion, and at two different biases, for the model detector configuration described above.

For this conductivity calculation, we applied an effective electron mass of 0.26 m_0 and a hole effective mass of 0.386 m_0, where m_0 is rest electron mass [10]. Fig. 2.19 shows graphs of induced charge due to the electron's and hole's drift, both with and without diffusion, and at two different bias voltages which affects the degree to which the drift velocity approaches saturation. Note that the geometry is a planar cathode and a strip anode, as reflected in the pulse shape; that is, for the electron motion, the change in the weighting potential is more precipitous near the partitioned electrode. At 500 V, the hole's velocity is nearly an order of magnitude below its saturated velocity, and the effect of diffusion is thus most pronounced (compare the blue and red lines in Fig. 2.19a).

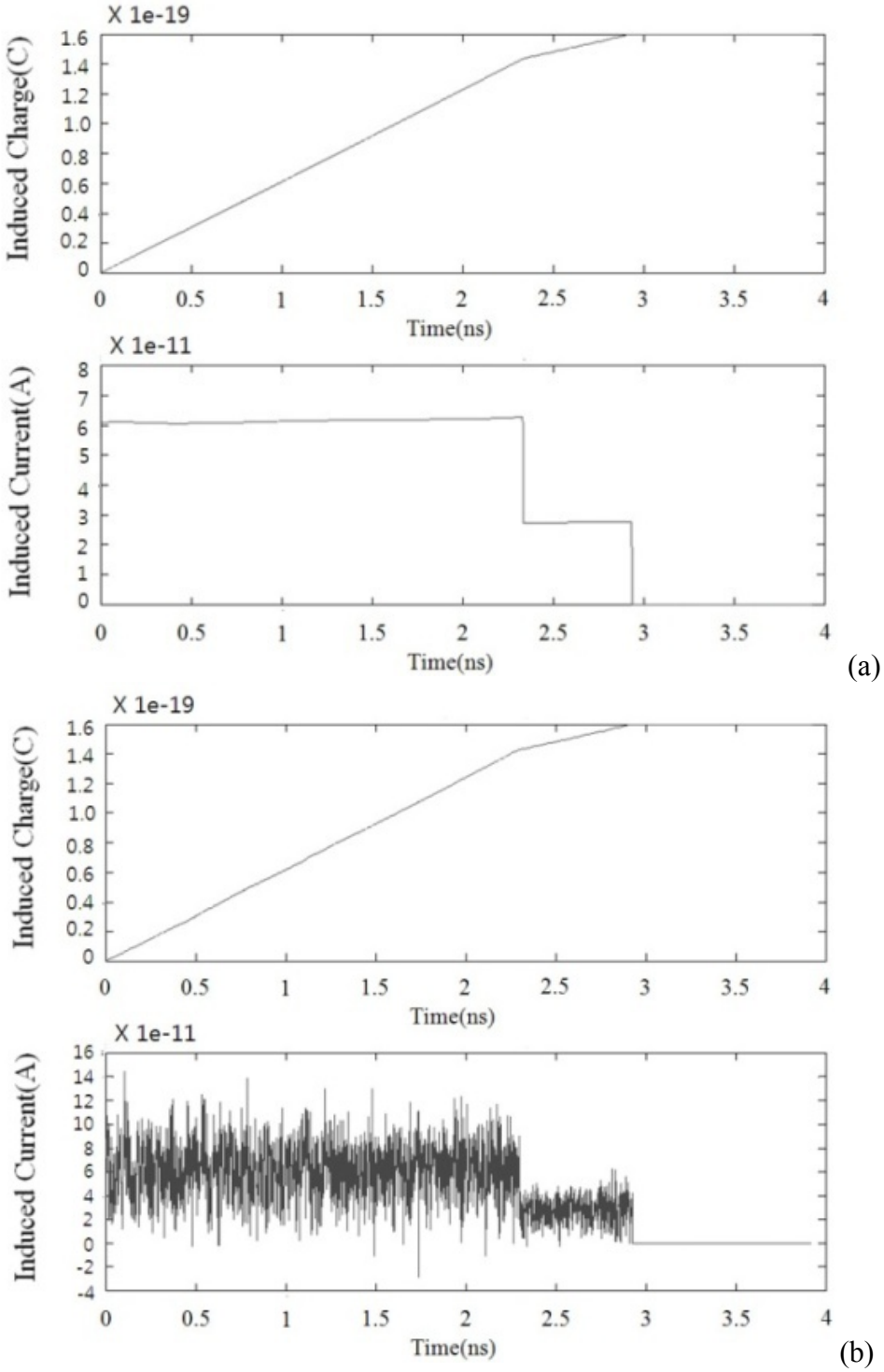
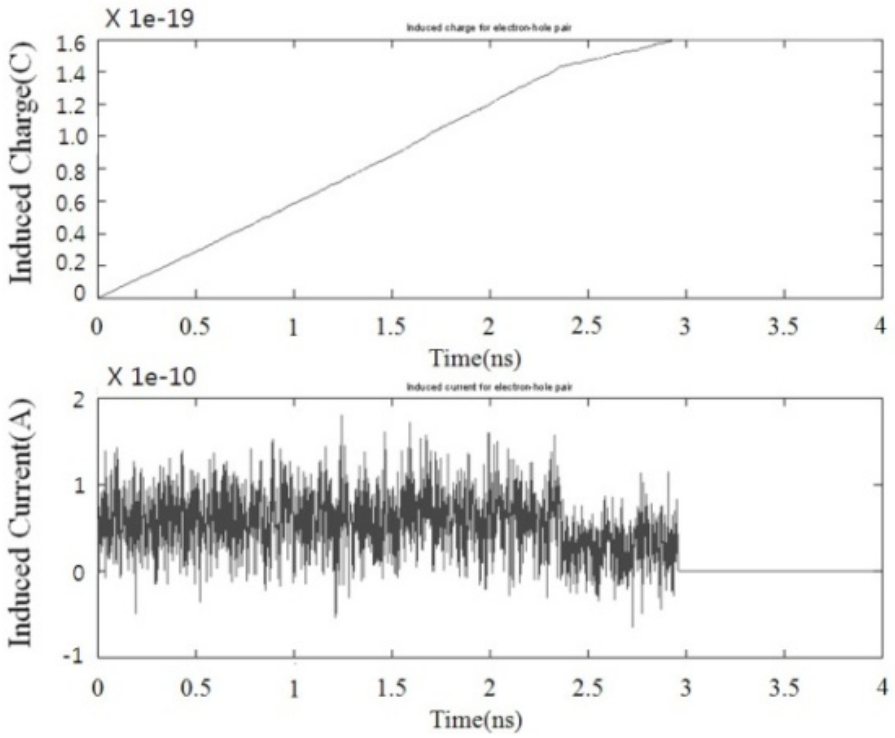


Fig. 2.20(a, b): Induced charge (top) and current (bottom) signals: (a) without diffusion, and (b) with diffusion using existing techniques when the bias voltage is 1000 V.



(c)

Fig. 2.20c: Induced charge (top) and current (bottom) signals with thermal diffusion with the described technique, when the bias voltage is 1000 V.

The random fluctuations can, in principle, obscure the amplitude variations due to charge collection; however, as shown in the graphs, the variation due to charge collection far exceeds that due to thermal jitter for even slow carriers. A more substantial consequence is in the charge collection time. Again at 500 V, the diffusion-influenced hole collection time can deviate from the pure-drift version by nearly half a nanosecond, which results in a miscalculation of the initial position by about 20 μm. As shown in the electron graphs, this deviation reduces to the desired sub-micron scale if the carrier's drift velocity is increased near saturation. Thus, in order to minimize the deleterious effects of diffusion, it is important to saturate the carrier velocities.

Figs. 2.20 shows graphs of induced charge and current due to a single electron and hole's motion for three different cases, in which the e-h pair is created at the detector's midpoint and the carrier velocities are saturated. In Fig. 2.20a, charge drift alone is modeled. To that deterministic motion, thermal diffusion is added in Figs. 2.20b and 2.20c, as modeled by two different methods. In Fig. 2.20b, diffusion is modeled

based on the existing diffusion-coefficient methods [1-2], and in Fig. 2.20c, the atomistic simulation of thermal jitter during the charge transport is validated. That is, the increment in the position due to diffusion, as outlined in Eqs. 2.8, 2.9, and 2.10 is used to calculate the new position, from which the change in the weighting potential and consequently, the induced charge, is calculated (see Eq. 2.3). The induced current, shown in the lower plots in Fig. 2.20, is then derived from the finite difference-derivative of the induced charge variation.

The thermal motion obviously complicates information extraction from the signal. Specifically, in Fig. 2.20a, the temporal position of the steps in the current pulse relative to the pulse start-time indicates- via two independent measures- the depth at which the charge pair is created. Furthermore, the size of the step reveals the number of charge carriers that are removed from the transport stream. For even this favorable case therefore, in which the carrier velocities are saturated, both the timing and size of each simulation step are obscured by the random motion. Means by which one can clarify the signal shape and thus extract the desired quantities are discussed in Sections III and IV.

2.5. Monte Carlo Modeling of the Coulomb Drift

Thermal diffusion certainly complicates the determination of the initial charge distribution, and moreover it places a fundamental limitation on the precision to which one can pinpoint the initial carrier distribution. Coulomb drift, the drift due to the electrostatic interactions of the cloud of charge carriers, can also complicate the electron tracking, but its effect can be removed because of its deterministic nature. For the sake of the Coulomb force calculation, we divide up the number of charges (or equivalently the energy) and spread them between the measurement points. That is, the Monte Carlo models will identify the energy loss at discrete points: $(\Delta r_1, E_1)$, $(\Delta r_2, E_2)\cdots(\Delta r_i, E_i)$, as implied in Fig. 2.21.

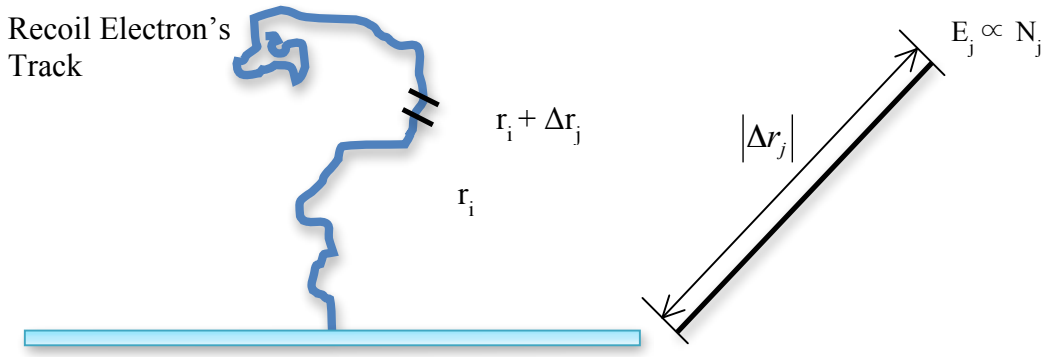


Fig. 2.21: A schematic of the track of the recoil electron (left) and an example position increment (right).

In the model, the E_i's are converted into electron-hole pairs by the Eq. 2.11. The number of electron-hole pairs created by the ionization energy of silicon is used as cut-off energy in this code.

$$N_{electrons} = N_{holes} = \frac{E_1}{w} \tag{2.11}$$

Ignoring the initial spatial separation between the electron and hole, we can spread the number of electrons (and holes) according to the energy loss formulas (greater number near end of track); however, for sufficiently small Δr_i, one can spread the charges uniformly between each energy-deposition step. Thus, the initial separation between electrons will be: $\Delta r_j / N_j$. For the results presented, we neglect the Coulomb force between electrons and holes until the charges separate, by more than 0.5 nm, in order to account for their discrete separation in the crystalline lattice. We thus modify the transport dynamics by assuming that each effect is independent:

$$m_{eff} a = \sum F = F_{drift} + F_{diffusion} + F_{Coulomb} \tag{2.12}$$

$$v = \sum v = v_{drift} + v_{diffusion} + v_{Coulomb} \tag{2.13}$$

where,

$$v_{drift}(\vec{r}) = \mu E(\vec{r}),$$

$$v_{diffusion} = \left[\frac{2k_bT}{m_{eff}}\ln\left|\frac{1}{1-\xi}\right|\right]^{\frac{1}{2}}, \text{ and}$$

$$v_{Coulomb}(x=\vec{r}_i) = \left\{\frac{q_i}{4\pi\varepsilon_0}\sum_{i\neq j}^{N}\frac{q_j(\vec{r}_i-\vec{r}_j)}{\left|\vec{r}_i-\vec{r}_j\right|^3}\right\}\times\frac{\Delta t}{m_{eff}} \tag{2.14}$$

As shown as Fig. 2.22, the electrons and holes are quickly separated by the Coulomb force, and its effect on the measured signal is therefore confined to very early part of the pulse. As shown in Figs. 2.20, in which the charge motion is mapped and the induced charge and current on the electrodes is deduced, diffusion can have a significant effect, particularly if the charges are drifted below their saturated velocities.

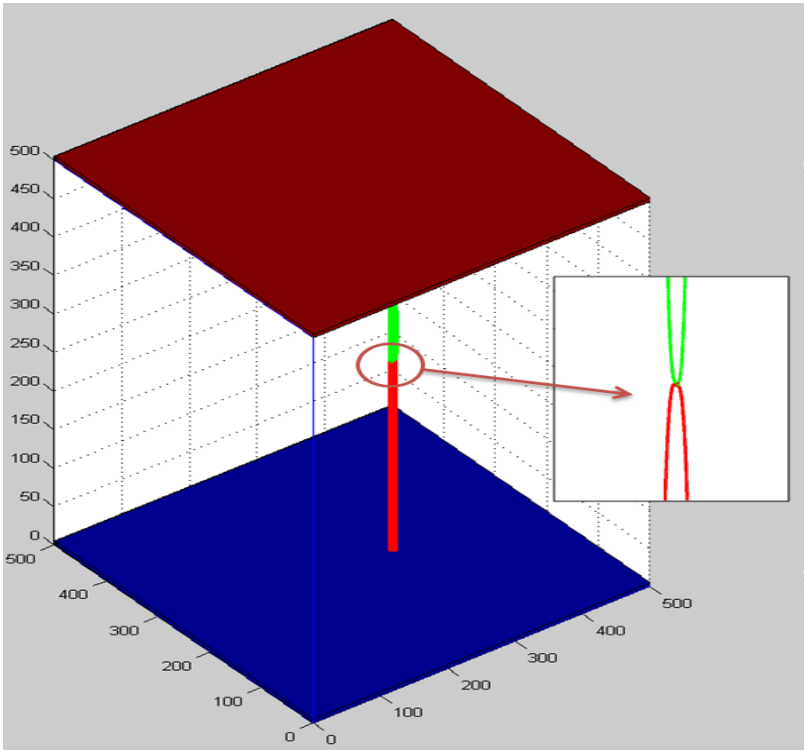


Fig. 2.22: For two electron-hole pairs initially at the midpoint of the detector, the divergence of their trajectories due to their mutual Coulomb force at small separations, as clarified in the magnified inset. The assumed applied bias is 100 V.

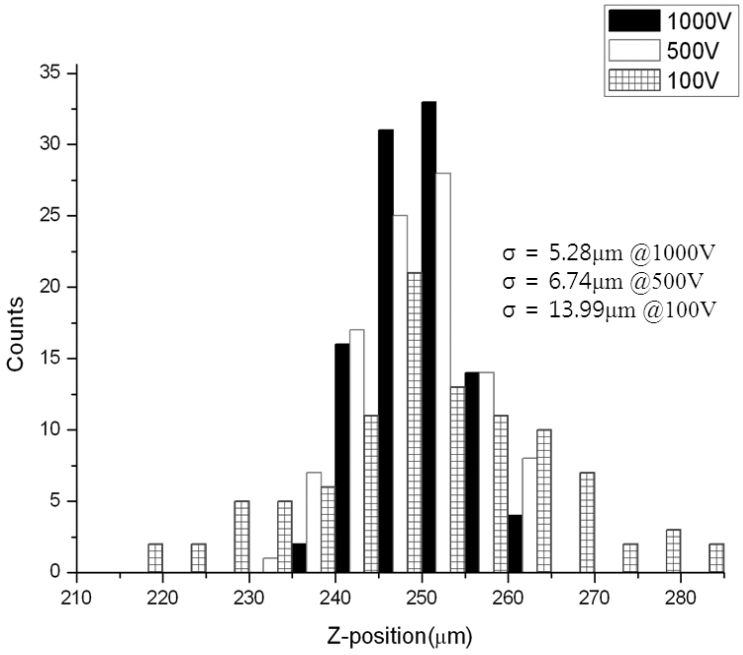


Fig. 2.23: Histogram of the measured depth-position and associated uncertainties, when the initial position of the carrier is 250 μm, for bias voltages of 1000 V, 500 V and 100 V.

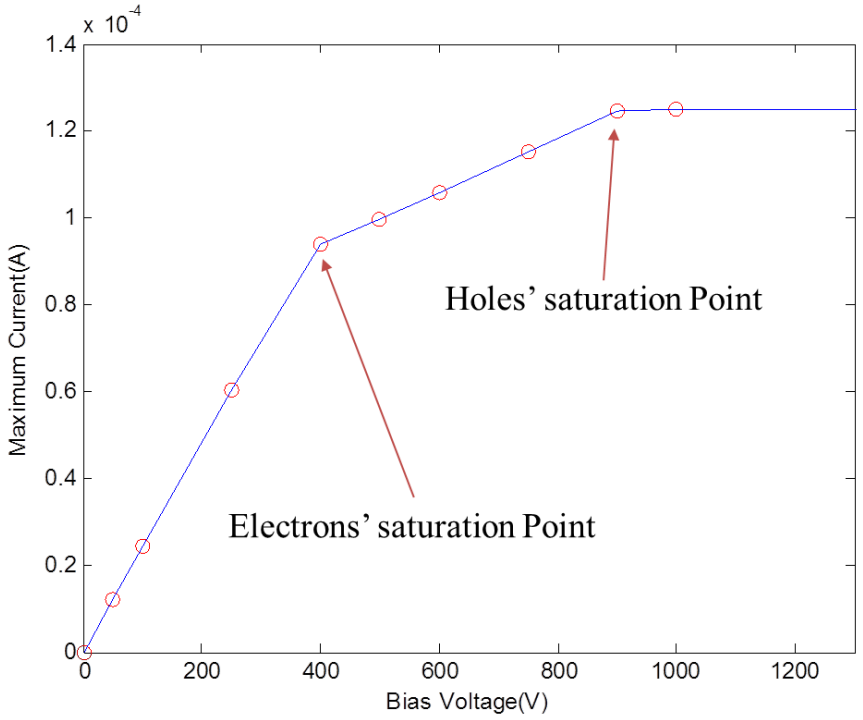


Fig. 2.24: Graph of maximum current with increasing bias voltage.

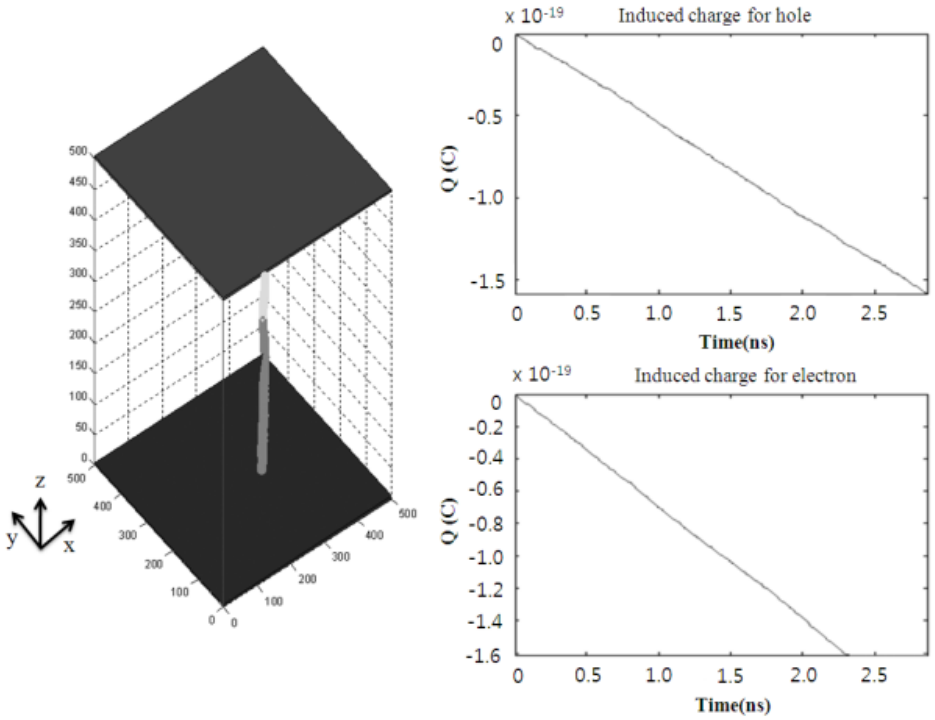
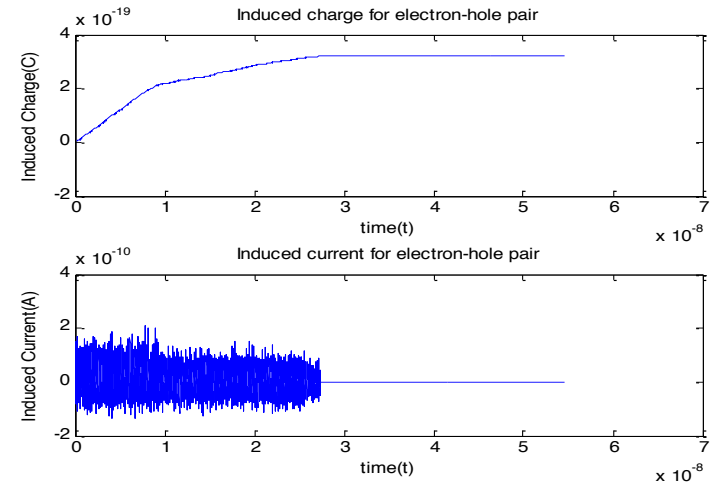


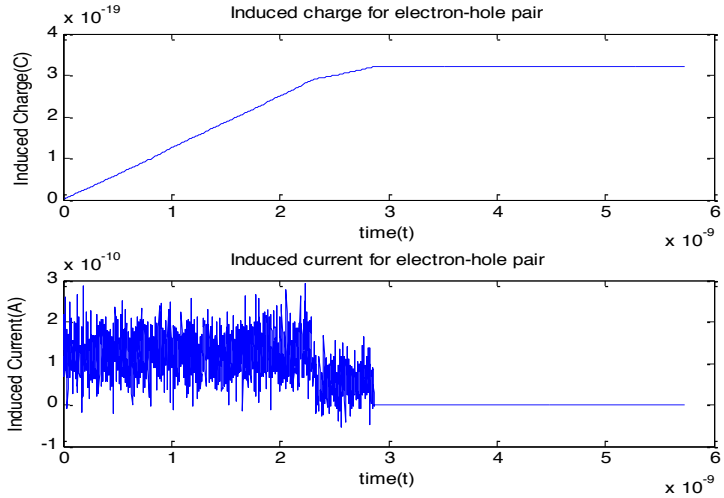
Fig. 2.25: Track of electrons and holes with thermal diffusion and induced charge due to their drift, when the bias voltage is 1000 V.

The random fluctuations can, in principle, obscure the amplitude variations due to charge collection; however, as shown in the graphs, the variation due to charge collection exceeds that due to thermal jitter for even slow carriers. A more substantial consequence is in the charge collection time, which can result in a miscalculation of the initial carrier position by 10's of micrometers as shown in Fig. 2.23, if they are drifted substantially below saturation. Fig. 2.24 shows the maximum current curve versus bias voltage, in which the saturation of the drift velocity is assumed to occur abruptly at the saturated electric field, and for lower fields, the drift velocity varies linearly with the electric field. The voltage for at which the electrons are saturated is approximately 410 V and the hole saturation voltage is 900V.

In order to clarify the effect of thermal diffusion as the bias voltage is varied, consider Fig. 2.25 in comparison to Fig. 2.15, in which the 1000 V bias results in the dominance of charge drift, relative to the thermal jitter.



(a)



(b)

Fig. 2.26: Total induced charge (top) and induced current (bottom) from two electron-hole pairs created at z = 250 μm, when the applied bias is (a) 100 V, and (b) 1000 V (i.e. carrier velocities are saturated).

Furthermore, to compare the effect of bias voltage on the impact of the existing thermal diffusion, we change the bias voltage from 100 V to 1000 V and obtain the current output graphs of Fig. 2.26. At higher biases, the total induced current has less jitter and one can therefore extract the drift time of the carriers more clearly; nevertheless, a clean extraction of that time remains difficult, unless one uses the comprehensive signal from multiple carriers to quench the magnitude of the thermal noise.

In Fig. 2.27, the variation in the current's signal-to-noise ratio (SNR) as a function of bias voltage is shown. The SNR can be acquired with the average power of the signal

(P_{signal}) and noise (P_{noise}) as shown in Eq. 2.15. For the single carrier case, the SNR saturates at approximately 20 dB near 900 V.

$$SNR(dB) = 10\log_{10}\left(\frac{P_{signal}}{P_{noise}}\right) \tag{2.15}$$

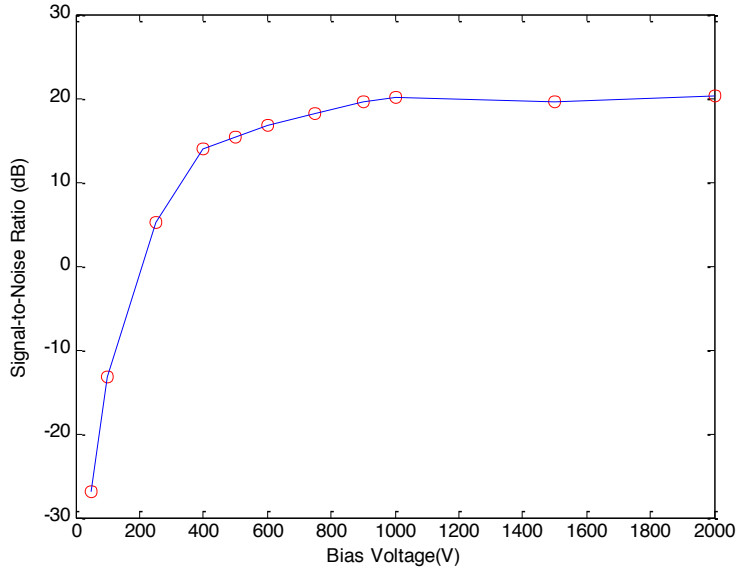


Fig. 2.27: The variation in the SNR as the bias voltage is varied.

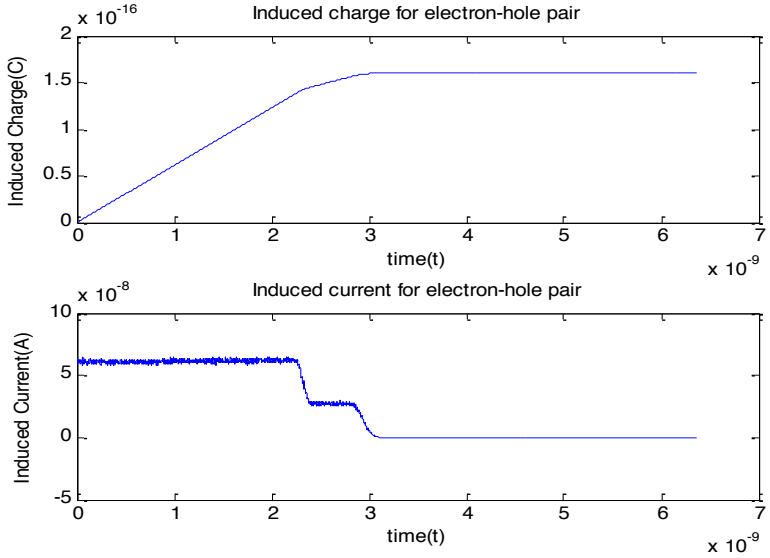


Fig. 2.28: Total induced charge (top) and induced current (bottom) when the number of electron-hole pairs is 1000 (initiated in a horizontal line) with thermal diffusion and 1000 V bias.

For multiple carriers, Fig. 2.27 shows the SNR variation. As clarified in the pulses of Fig 2.28 for which 1000 electron-hole pairs are created horizontally in the middle of the detector, the effect of thermal jitter is substantially reduced and the temporal width of the collection time provides a measure of the degree of random walk away from the drift plane of the carriers.

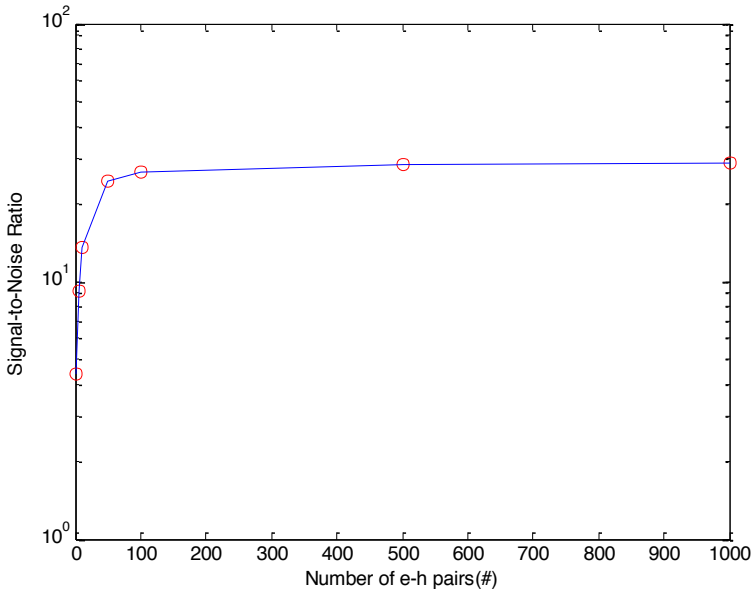


Fig. 2.29: Graph of the SNR variation with increasing the number of carriers with 500 V bias.

As shown in the plot of Fig. 2.29, the SNR increases rapidly as the number of carriers is increased from 1. For greater than 500 electron-hole pairs, the SNR was saturated at about 28 dB.

For the Coulomb force effect, we compare the total induced charge and current with and without the Coulomb force. During overall drift, the effect on the shape of the induced current pulse is negligible; however, the initial cloud repulsion can result in a large current spike at small times (sub-ps), as shown in Fig. 2.30. For standard measurement modalities, in which the induced charge is measured with a charge-sensitive amplifier, the Coulomb drift doesn't impact the measurement, regardless of the bias voltage.

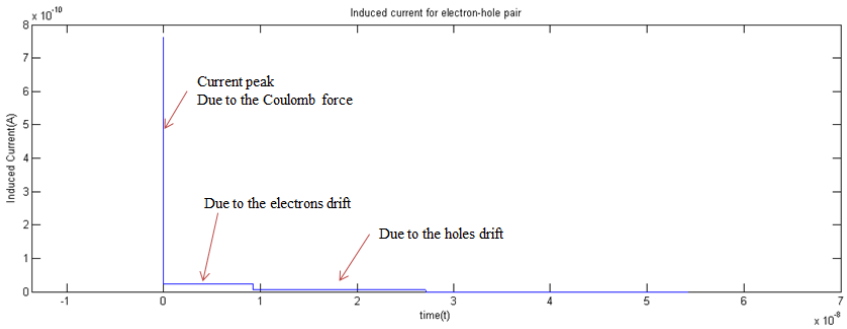
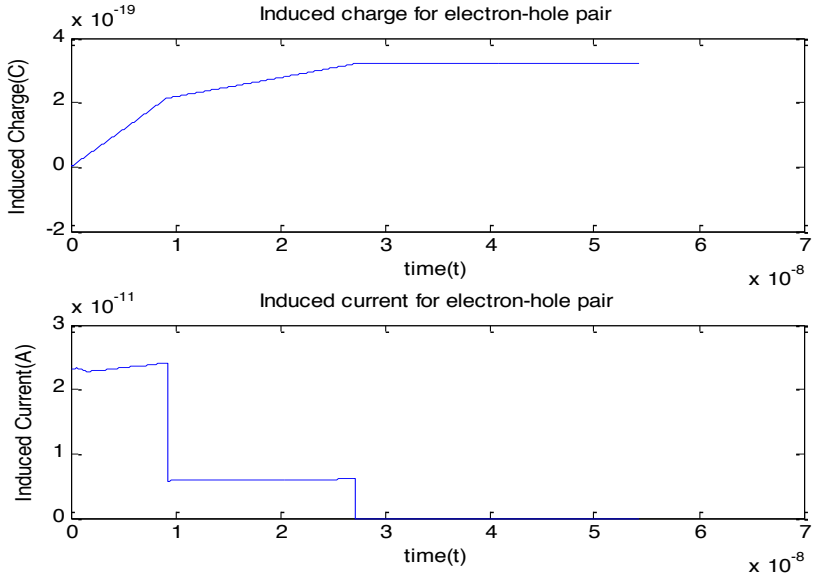
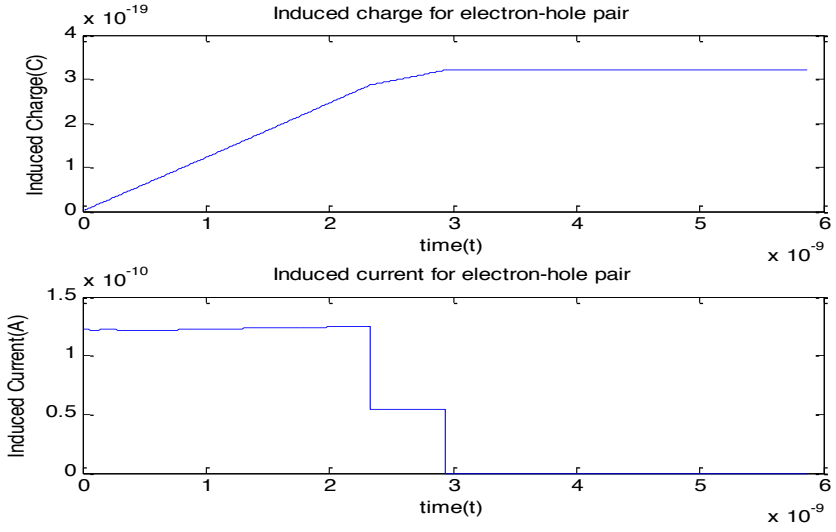


Fig. 2.30: Graph of the total induced current with Coulomb force.



(a)



(b)

Fig. 2.31: Total induced charge (top) and induced current (bottom) when the number of electron-hole pairs is 2 with Coulomb force and 100 V (a) and 1000 V (b) applied bias voltage.

As shown in Fig. 2.31, in which the pulses for two biases (100 V and 1000 V) are compared, the collection time variations are due solely to the change of velocity for both carriers. Nevertheless, at small carrier separations (< 10 μm), the electrostatic Coulomb force can be much greater than the drift force and thus impact the trajectory reconstruction. Fortunately, the effect of Coulomb drift can, in principle, be removed via post-processing once the charge distribution is estimated, because the dynamics are not stochastic. Nevertheless, both Coulomb drift and diffusion can limit the position resolution if micron-scale position-sensing is achieved.

Although Coulomb drift and thermal diffusion can potentially limit the resolution with which one determines the initial charge distribution in the detector, both effects can be counteracted such that micrometer-scale position sensing is possible. In particular, the timing and size of the transition in the current pulse during charge collection at the electrode provides an estimate of the quantity of charge produced at a given position. The effect of Coulomb drift on that estimate can be removed, in principle, by analyzing the pulse shape during the initial stages of the charge drift. Nevertheless, the simulated results imply that the space-charge effect does not significantly degrade the positional or charge determination.

Thermal diffusion can degrade the precision of that determination however, especially if only a handful of carriers are tracked in the detector. In order to mitigate its effect, the charge carriers should be drifted at their saturated velocities. In order for either effect to significantly limit the precision in the determination of the initial charge position, one must employ a low-noise readout that is sensitive to 100's of electrons at most, which we aim to achieve using on-chip transimpedance amplifiers [7].

2.6 Transmission line-based analysis and design of detector electrodes and integration of on-chip TIA

The primary structure of the detector is that of a meandered microstrip line, with a strip width and spacing of 10 μm each, in order to achieve good position sensitivity but with ease of microelectronic processing (smaller features requiring greater care). For radiation detectors based on silicon, one desires thick detectors- standard 550 μm wafers used in our case; however, under these geometrical conditions, a microstrip line presents considerable inductance which becomes more significant as we approach higher

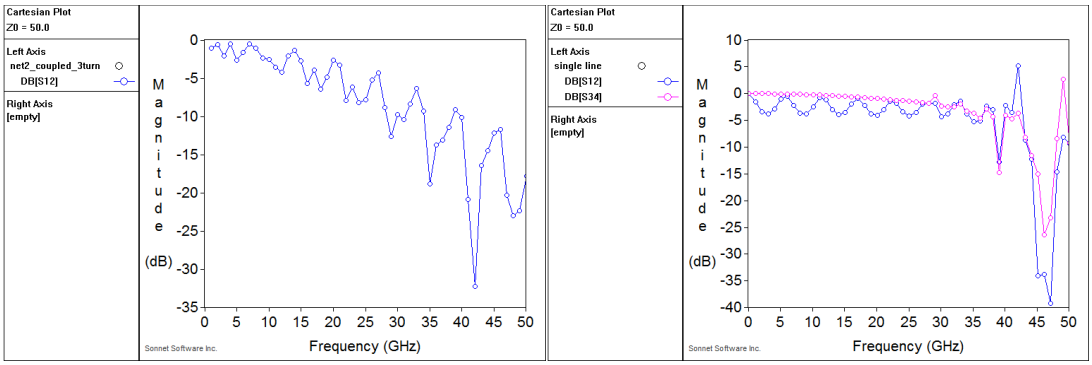
frequencies. In order to counteract this effect, an asymmetric strip-line structure was developed. Such structures were simulated using the Sonnet Simulation Software and considerable improvement was observed both in a single ended strip-line configuration and also in parallel and differentially coupled strip-line configurations.

The design further extends to accommodate a bare die-form transimpedance amplifier with differential output. The differential output will aid in noise reduction during the second stage amplification when the voltage amplifier will take in these differential signals at its input. Interconnects between the stages have been designed as 50 Ω coplanar waveguide (CPW) structures (the signal line guided by ground planes on either side of it). This is primarily done to reduce cross talk (the ground surrounding the signal nulls any external noise) and also to enable mating of these lines to an external SMA pin (the characteristic impedance of a CPW structure depends on the width of the line and the signal-ground spacing; thus, it is easier to maintain the 50 Ω while flaring).

2.6.1 Sonnet Simulation of microstrip and asymmetric strip line structures

The Sonnet EM analysis software package is a planar analysis tool which uses the method of moments to calculate the current distribution on a circuit metallization. The metallization is taken as a zero thickness layer sandwiched between the two dielectrics. The circuit is sub-sectioned and the overall electric field is calculated due to the subtended currents. For a lossless conductor, for which the tangential electric field component is zero, the current in each subsection is adjusted so that the tangential field on the surface of the conductor is nulled. Similarly, if there is conductor loss specified, then the tangential field in each subsection is made proportional to the current in each subsection, with metal surface resistivity being the proportionality constant [11].

For the present simulation, conductor loss (low frequency loss and skin-effect loss) for the aluminum metallization and silicon's dielectric loss were considered. The back metallization was also taken as aluminum for simulation purposes. The simulation showed that the impedance of the metal was rather high, at 126 Ω and that the line behaved similar to an LC filter with prominent resonant frequencies (harmonics of 6 GHz), as shown in Fig. 2.32a.



(a) (b)

Fig. 2.32: (a) Line resonances observed for 70 mm line with 5 turns (b) Resonances shifted and reduced for smaller line length of 15 mm. For a single turn line (S34) the resonances are shifted to higher frequencies compared to 5 turn line of the same length.

As the line length is decreased, or the number of turns is diminished, the resonances shift to higher frequencies, as shown in Fig. 2.32b. Conversely, as the line length is increased, the fundamental frequency shifts to lower values and the line showed more resonances, even for lower frequency signals.

Two main methods were investigated as means by which one can counteract the excess inductance of the line: a) adding compensation capacitances at nodes, and b) incorporating a dielectric overlay topped with a metallization, thus realizing an asymmetric stripline configuration.

For a low-loss line the characteristic impedance of a transmission line is given by [12],

$$Z_0 = \sqrt{L/C} \times \left[1 - j\left(\frac{R}{2\omega L} - \frac{G}{2\omega C}\right)\right] \tag{2.16}$$

For high resistivity silicon (ρ = 10 kΩ-cm), its conductance G is small. For microwave frequencies ωL » R and ωC » G, and for all practical purposes Z_0 is purely resistive.

$$Z_0 = \sqrt{(L/C)} \tag{2.17}$$

At an impedance of 50 Ω, a coaxial line presents the lowest possible loss for high power handling capabilities, and since our measurement components present 50 Ω, we seek to design the line and interconnects to match this impedance. At an impedance of Z_0 = 126 Ω, as present for our baseline design, the microstrip electrodes present an excess distributed inductance for a given capacitance. For the simulations, the input and output

ports were set to 50 Ω characteristic impedance, though for our semiconductor detectors, this is not entirely appropriate as the input generates charges and acts like a current source; however, assuming an output port impedance of 50 Ω is reasonable as this is what we are ideally trying to achieve in order to match with the rest of the circuitry.

2.6.2 Compensation using lumped capacitors

To compensate for the excess inductance a lumped capacitor in the form of a thicker turn on the corners of the meander was added, schematically shown in Fig. 2.33b. The graphs in Fig. 2.34 indicate that the compensation does reduce the resonances and supporting frequencies well up to 50 GHz with less than 3 dB loss at the resonant frequencies (comparing Figs. 2.34a to 2.34b); however, for lines longer than a 70 mm (Fig. 2.34c), the line behaves like a low pass filter, such as that shown schematically in Fig. 2.33a.

For a small length of line (when $L < \lambda/4$: 50 GHz corresponds to $\lambda = 6$ mm) the variation in voltage along the length of the line is small and it can be treated as a lumped inductive element; hence, the compensation with a capacitance is effective. However, when the line length is increased, it accommodates many wavelengths and the line starts to behave more like an LC filter [12]. Since our goal is to build detectors with only a limited restriction on the electrode length, the lumped capacitor compensation will, unfortunately, set the high frequency limit far below our region of interest.

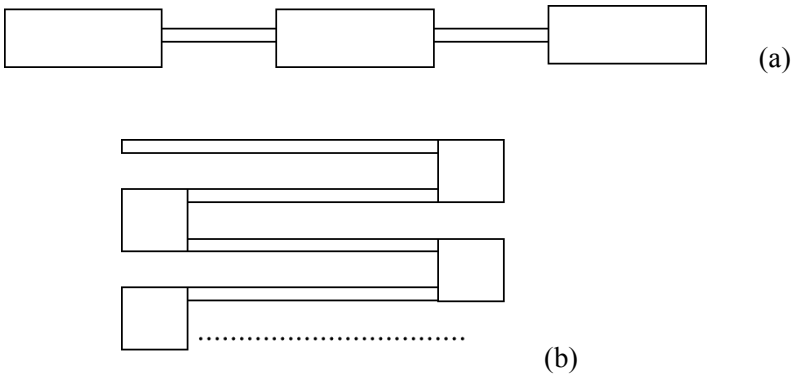


Fig. 2.33: (a) Typical LC Low pass microwave integrated circuit (MIC) filter and (b) meandered lines with compensation at the turns.

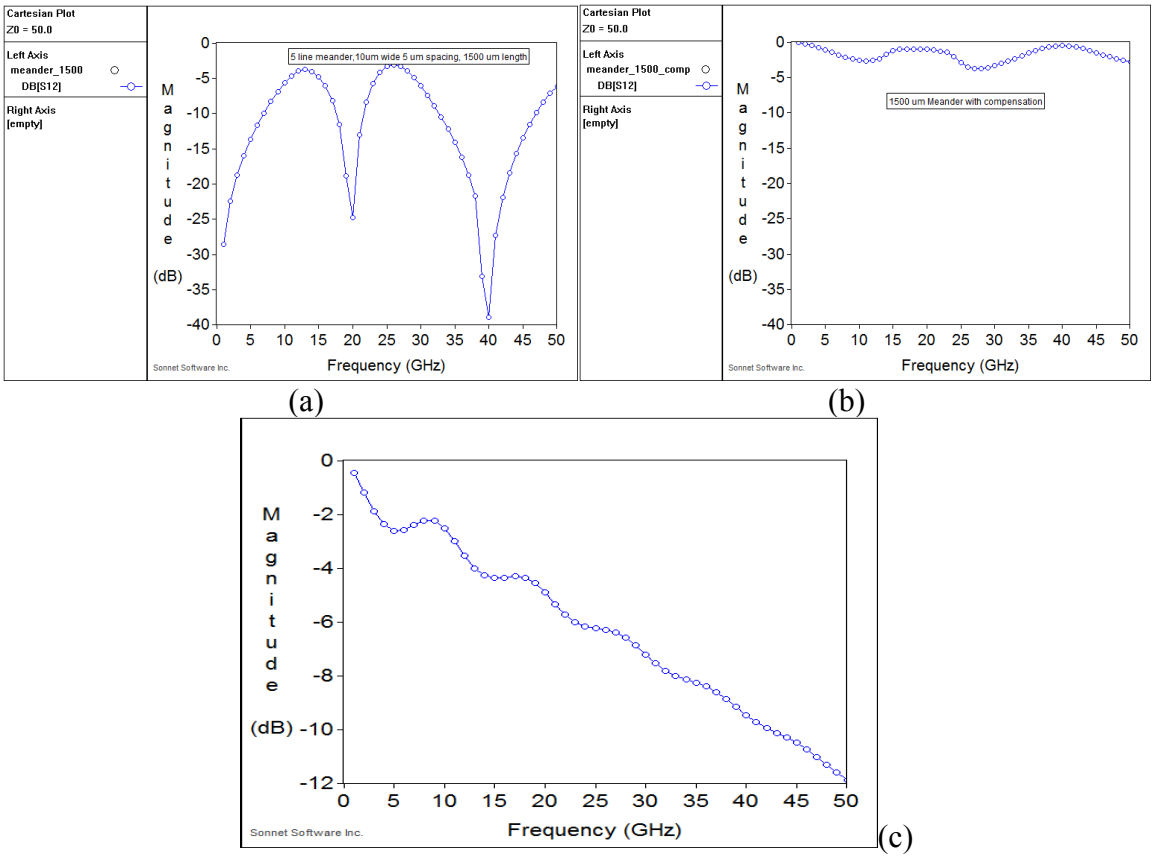


Fig. 2.34: With reduced strip width and spacing as 5 μm: (a) a highly resonating line at 1.5mm length, (b) when capacitive compensation on the turns is added, the resonances are reduced and the line is not as lossy, (c) when the line length is increased to 70 mm, a low pass filter response results.

2.6.3 Compensation using distributed capacitance

An alternative strategy is to form an asymmetric stripline configuration in order to introduce some additional capacitance and also to reduce the jitter due to coupling between the lines [13]. The design uses a polysilicon dielectric overlay on top of the microstrip line with a metallic ground on top. In such a configuration, the signal line is surrounded uniformly on all sides and the E-field lines avoid abrupt bends at material junctions, unlike in the case of a microstrip design.

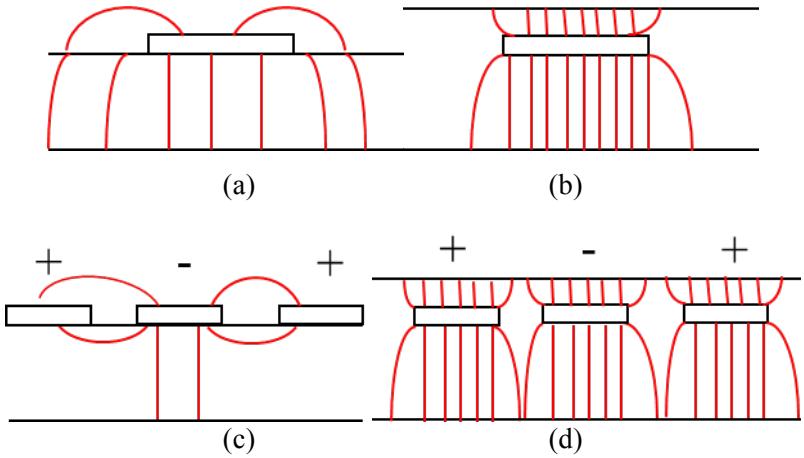


Fig. 2.35: Comparison between electric field lines on a microstrip and an asymmetric strip line [13] (a) Simple microstrip, electric lines of force uniformity lost at the junction between air and Si. (b) Asymmetric Strip line (c) Odd mode coupling between adjacent lines on a microstrip (d) Electric field lines deviated to top ground plane reducing the coupling.

For the 10 μm strip-width design, a 50 Ω line requires an overlay thickness of 13 μm, substantially smaller than the thickness of the detector bulk. The overlay thus allows us to decouple the signal propagation requirements from the detector requirements. The effect of the overlay on the 550 μm detector is thus to reduce the impedance slightly because of the uniformity introduced which can be adjusted by increasing the height on the thinner end (13 μm side) by a couple of units.

As shown in Figs. 2.35a and 2.35c, the microstrip structure supports odd mode coupling due to closely placed lines; however, if a ground is introduced close to the signal lines as shown in Figs. 2.35b and 2.35d, then the electric lines of force are diverted to the closer ground and the degree of coupling is reduced [14]. The coefficient of coupling K, follows a general rule of,

$$K = 1/\left(1+\left(\frac{s}{h}\right)^2\right)^2 \tag{2.18}$$

where s is pitch or center to center spacing between the lines and h is the distance of the line from the nearest reference plane [14]. Since we have very small spacing (s = 10 μm) compared to height of the substrate (h = 550 μm), bringing the ground plane closer

by adding a top silicon overlay of 13 μm will appreciably reduce the cross talk or coupling between the lines.

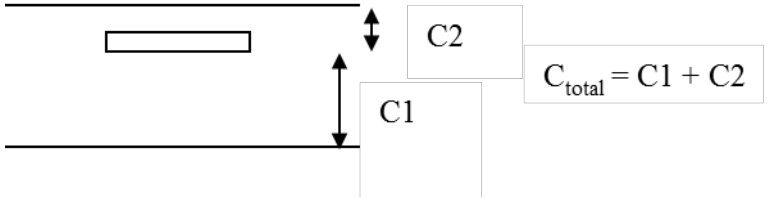


Fig. 2.36: Introducing distributed excess capacitance to compensate for the excess inductance and bring down the impedance of the line to 50 Ω.

An alternative method of examining the same structure is via the introduction of the compensation capacitance in a distributed manner. From Fig. 2.36, an excess capacitance of C2 is added to C1 (which the microstrip line alone presented), C2 balances the excess inductance thereby bringing down the characteristic impedance close to 50 Ω. Note that during a radiation event, both overlays may be depleted and thus, the entire volume can be active, although the pulse processing must account for transport in the entire E-field region.

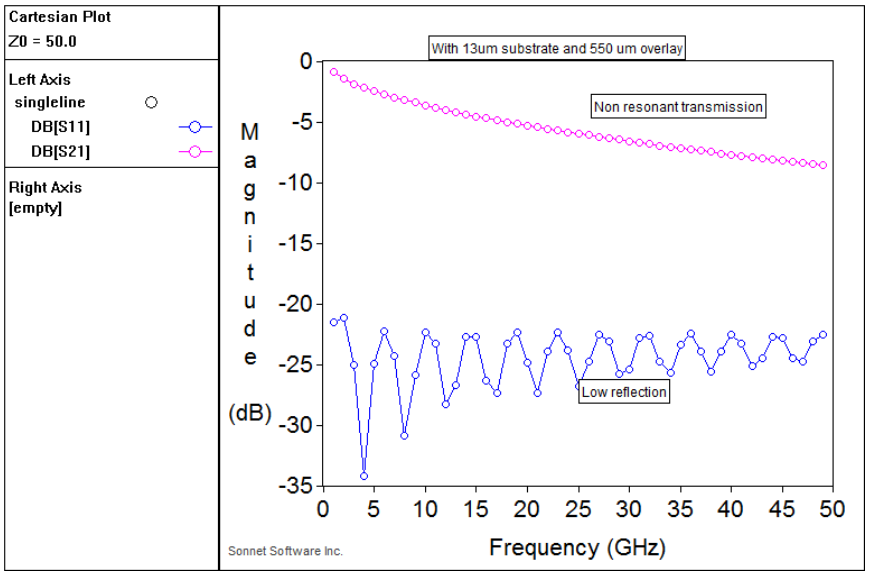


Fig. 2.37: Sonnet simulation of the asymmetric stripline structure. S11 (blue) is the reflection coefficient or return loss and S21 (red) is the Transmission Coefficient.

A Sonnet simulation of the asymmetric strip line shows satisfactory results, the results of which are displayed in Fig. 2.37. The return loss S11 was maintained below -20 dB for frequencies up to 50 GHz. The transmission coefficient S21 was non-resonant but did show up to -8 dB loss at 50 GHz which can be attributed to dielectric loss of the silicon substrate; nevertheless, the design of the detector as a 50 Ω transmission line structure is promising and meets the requirement for the TIA input.

2.7 Performance Validation and Signal-Shape Preservation for the Delay-Line Charge Tracker

Any long conductor carrying a high frequency pulse opposes the signal flow due to its self-inductance. The reactance presented by the inductance becomes more and more significant at higher frequencies (due to Lω). Whenever the return current line (the ground) is brought into close proximity to such a conductor the magnetic field lines cancel each other and the structure exhibits a lower inductance equal to its self-inductance minus the mutual inductance due to the ground. Since the meandering line on the surface of a 1 cm^2 detector gets longer with the decrease in the line width, its inductance only increases with the intended increase in position resolution. For timing measurements this would mean sacrificing the higher frequency components of the signal to line inductance. When a source is connected to a load via a “long” transmission line, the line's characteristic impedance dominates over the load impedance in determining the circuit behavior. In other words, an electrically “long” line acts as the principal component in the circuit, its own characteristics overshadowing the loads; for instance, consider a 50 Ω cable of infinite length. Regardless of the nature of the load that we connect to one end of this line, the source will only see 50 Ω of impedance because the line's infinite length prevents the signal from *ever reaching* the end where the load is connected. The most effective way to minimize the impact of transmission line length on circuit behavior is to match the line's characteristic impedance to the load impedance. If the load impedance is equal to the line impedance, then *any* signal source connected to the other end of the line will “see” the exact same impedance, and experience the same amount of current draw. In this condition of perfect impedance matching, line length only affects the amount of time delay from signal departure at the source to signal arrival at

the load. However, perfect matching of line and load impedances is not always practical or possible.

Taking the current design as an example, a small signal collected on a 5 μm wide meandering line that spans a 1 cm^2 area with 10 μm pitch on silicon would have to travel a length of anywhere from a few micrometers to 5 meters of length depending on the point of collection. If the signal needs to travel a few mm (say 5 mm) length of line (when the interaction happens on one edge and all that the signal travels through is the wire bond length before reaching the amplifier), transit time through 5 mm wire bond is only 36 ps, and let us assume that this is smaller than the time period of the highest frequency component of the pulse. In this case the signal only sees the input impedance of the amplifier and the characteristic impedance of the 5 mm line becomes unimportant. The oppositely directed signal has to traverse a 5 m length of meandering line before it sees the amplifier. In this case, the total transit time of the line is close to 36 ns (for ideal modeling) and during this time the highest frequency component can undergo several cycle changes before seeing the amplifier. In this case, the signal sees the line as a resistor of value equal to the characteristic impedance of the line and the transmission line properties therefore become important to this signal. However, if we design the lines as a 50 Ω line then the signal collected anywhere on the meandering line sees 50 Ω on both sides and is divided equally; furthermore, upon reaching a matched load the termination reflections can be avoided. The other advantage of having a matched line/load on both sides of the point of interaction is that the signals that reach the ends will be of similar strengths for a perfectly lossless line.

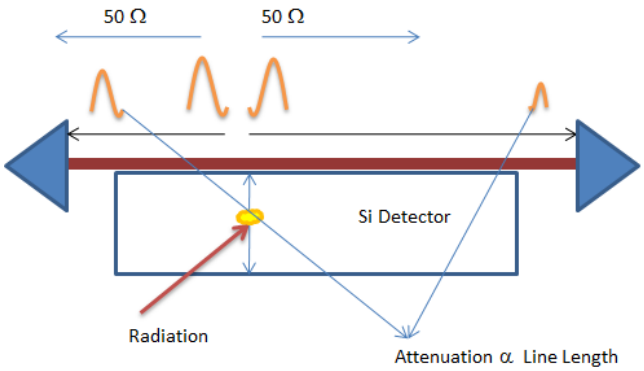


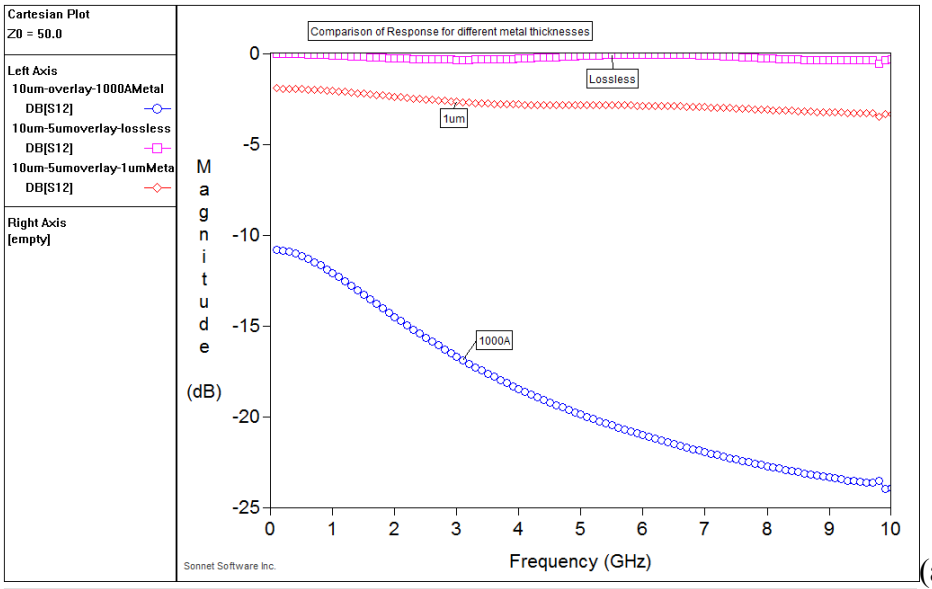
Fig. 2.38: For a matched Line, the signal splits equally but undergoes different amounts of attenuation depending on the conductor and dielectric losses of the line both of which are proportional to line length.

A lossy line on the other hand, will result in different signal strengths on the respective ends proportional to the line length. This will solely be due to the conductor and dielectric losses of the system, as illustrated in Fig. 2.38. Therefore, the line-design condition is that the lines have to be matched so that the signals, regardless of where they are created along the line, see similar impedance on both sides. In this case, they are split equally and are not subjected to reflections due to readout mismatch. The maximization of the signal-to-noise ratio (SNR) at the input of the amplifier is then dependent on how well we reduce the losses due to the limited conductivity of any metal as well as losses due to dielectrics.

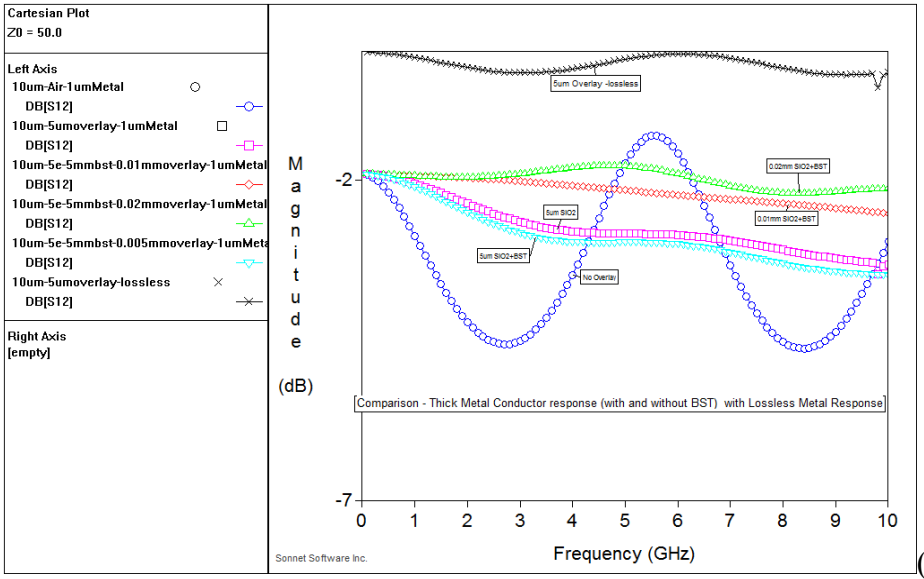
2.8 Minimize the Pulse Shaping via Different Transmission-Line Geometries: Frequency Domain

The work that has been done so far is based on the above concept, and we have tested and compared lines with and without impedance matching. As shown later, we have shown that using a transimpedance amplifier (TIA) could reproduce the instantaneous current signal that reaches the electrode. These signals were mostly formed close to the output pad of the detector and had to travel short line-lengths before seeing the amplifier. With lines now designed to match the amplifier input we could see better signal integrity with minimal reflections. Additionally, the detector bias of 450-500V could be reached, because of reduction in leakage current through the signal line, as discussed in Chapter 3. At 500 V bias, the signal rise time is approximately 175 ps, which tells us that the highest frequency component of the pulse is around 2 GHz, which informs the design goal to realize constant transmission response for all frequencies at and below 2 GHz. For the second generation design, a metallization thickness of 0.1 μm was sputtered on 2 μm, 5 μm and 10 μm lines. These have been fabricated with and without the silicon dioxide overlay. When simulated with Sonnet, the conductor loss for 0.1 μm is significantly higher than that with the 1 μm thick metal due to skin effect losses at lower frequencies. As shown in the frequency response plots of Fig. 2.39a, significant loss reduction is noticed when the metallization thickness is increased from 0.1μm (blue line) to 1 μm (red line). Being dependent on the frequency, the skin depth decreases by 10 times from 10 μm at 1GHz to 1 μm at 10 GHz, and the conductor losses dominate for thinner lines.

From Fig. 2.39b, we see that an unmatched line (blue), has the transmission coefficient (S12), that significantly oscillates within the frequency range. When the overlay is applied the values become more stable (pink) and we still see a dip at higher frequencies due to the dielectric loss that varies linearly with frequency. When a thin layer (50 nm) of BST (barium strontium titanate) is added to the overlay the dielectric loss further decreases at higher frequencies (red and green), a feature due to the 100 times lower loss tangent of the BST (10^{-4}), compared to SiO_2 which has loss tangent of 0.01.



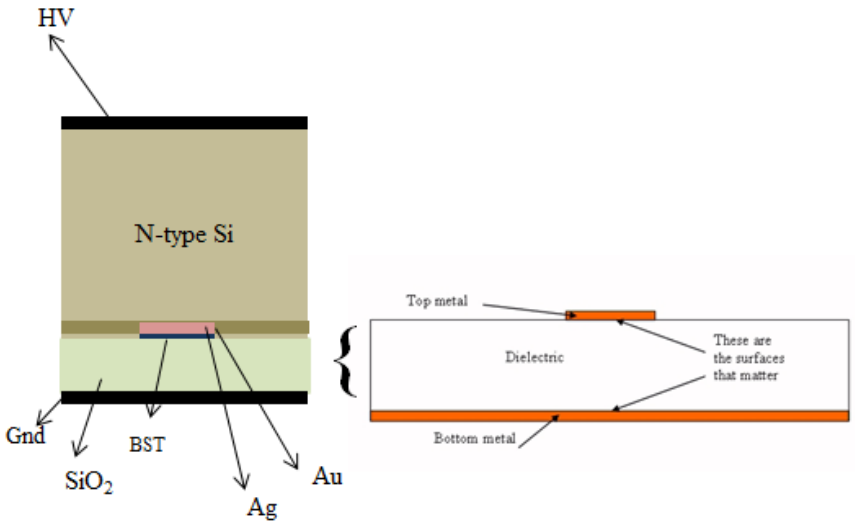
(a)



(b)

Fig. 2.39: Transmission coefficient (in dB) as: (a) the metal thickness is varied, and (b) for various overlay designs, including the use of a BST/SiO_2 sandwich overlay.

Further improvements in the metallization can be made with bimetal configurations, in which silver is used to conduct the high frequencies (higher conductivity) and gold remains to realize the Schottky barrier contact. As shown in Fig. 2.40 (inverted detector), the overlay with the ground metallization below is a microstrip transmission line structure.



Fig, 2.40: Higher Conductivity of Ag/Cu utilized for Signal Transmission.

The high frequency signals travel along the inner surface which can be chosen as a high conductivity metal like silver or copper. The gold or platinum films (which have lower conductivities) are essential for Schottky barrier formation and since they are on top of the higher conductivity metals (Ag/Cu) in the inverted detector configuration, they will not significantly impact the signal transmission. This fact justifies our use of Au/Pt for diode formation and Ag/Cu as an additional layer for transmission reasons.

2.9 Validation of the Effectiveness of the Higher Conductivity Electrode

Schottky barrier detectors were tested in which the metal contact consisted of a combination of high work-function Au for the interface metal and high-conductivity Ag for the current carrying medium. The sandwich layer would impact the leakage performance of the contact due to a balance of the Fermi levels anticipated and inter-diffusion of the species. An expectation, unfortunately confirmed via experiment in which the leakage current, was 10 or 100 times worse than a gold contact alone. We

therefore employ PIN structures to realize the electric field in the semiconductor and couple the charge motion to the circuit with an electrode designed solely for maximum pulse transmission. That is, we would prefer a simple Ag line with ~1 μm thickness, but the Ag doesn't adhere to the silicon, and the wire-bonds don't adhere to the Ag, as discussed previously; instead, it is desirable to use Au as the top layer for wire-bonding reasons. Thus, a 50 nm Ti interstitial layer is used to realize adhesion between the Si and the Ag (0.5 μm thick), and the Au (0.5 μm) adhesion to the Ag is realized with a Cr layer of thickness 50 nm.

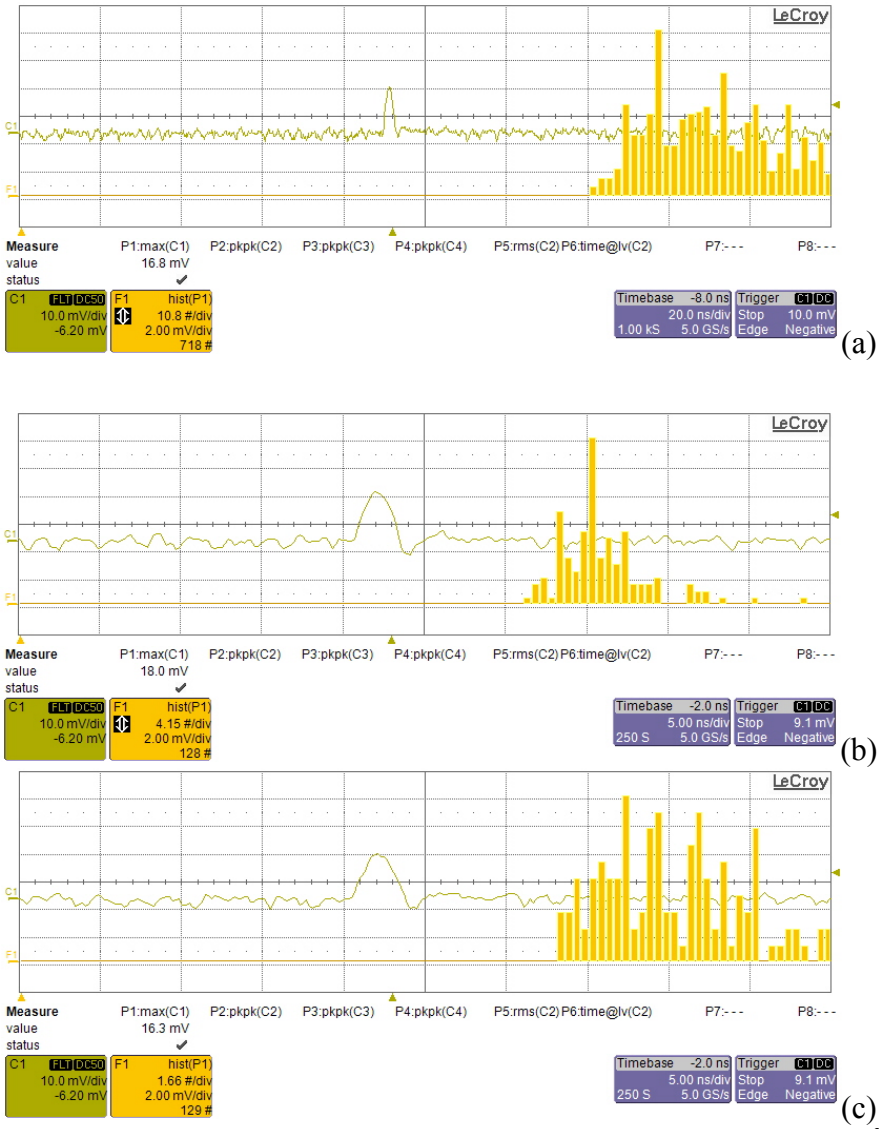
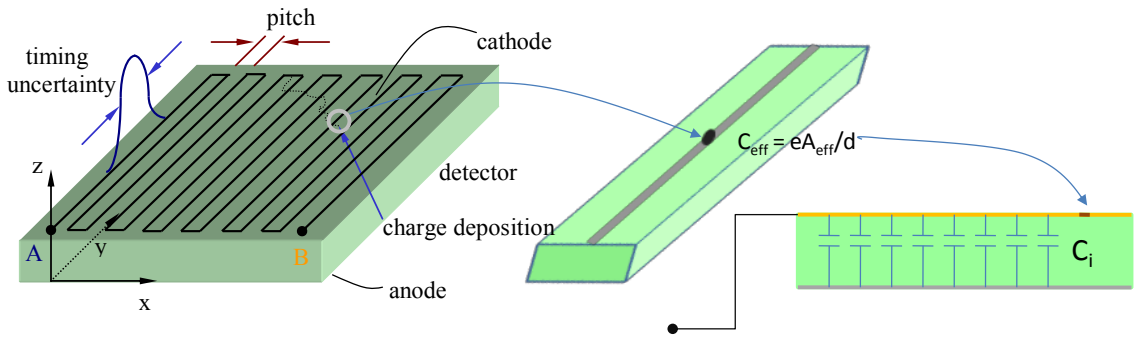


Fig. 2.41: (a) Near-end collection of pulses above 10 mV, on a line with 500 Å of Pt. (b) Far end collection of pulses above 10 mV, on a line with 500 Å of Pt. (c) Far-end collection of pulses above 10mV, on a line with 1 um of Ag.

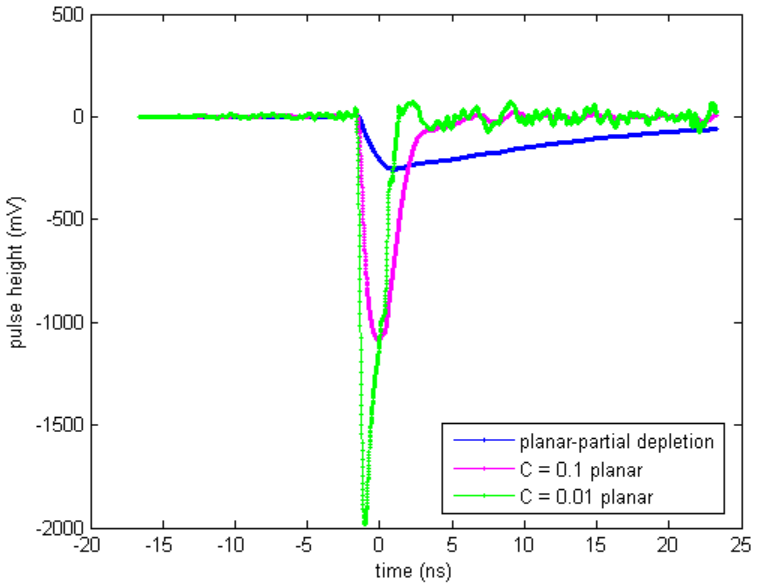
The empirical validation of the higher conductivity line is shown in Fig. 2.41. In Figs. 2.41a and b, the 30 minute height distribution of pulses above a certain threshold (10 mV in this case) is compared for the cases in which the alpha source is collimated at the end of the meander line nearer the readout circuit (Fig. 2.41a) versus that from a source collimated at the far end of the line (Fig. 2.41b). For interactions farther from the readout TIA the number of pulses above the threshold reduces and the height distribution shifts to lower amplitudes. As shown in Fig. 2.41c, in which the Ag electrode is employed with a 75 turn meander, the original near-end height distribution is largely recovered, which shows that the information is transmitting along the higher conductivity device. The extra effort required to process the high-conductivity electrode is thus worth it and has been used on various detectors, in which the thickness can be chosen to mitigate the conduction loss depending on the bias used and the highest frequency desired from the pulses.

2.10 Minimize the Pulse Shaping via Different Transmission-Line Geometries: Time Domain

The effect of various designs can be elucidated by looking at the time-domain response of the signals, which also clarifies the number of turns that can be employed before the pulses are significantly attenuated or shaped. First, consider Figs. 2.42 and 2.43, in which the shaping effect of an uncompensated line is illustrated in both simulations (Fig. 2.42) and measurements (Fig. 2.43). Recall that one of the main functions of the overlay is to convert the capacitive line into a resistive line. That is, if a strip-line is simply patterned on the surface of the detector as shown in Fig. 2.42a, then the capacitance that the signal experiences is given by the sum of the parallel capacitances (C_i) that the pulse sees as it travels to the readout end. Thus, interactions farther from the end of the line will be shaped by higher capacitance, as shown in Fig. 2.42b, in which a measured pulse (in green) is shaped by an RC circuit formed by the varying line capacitance and the TIA impedance (50 Ω).



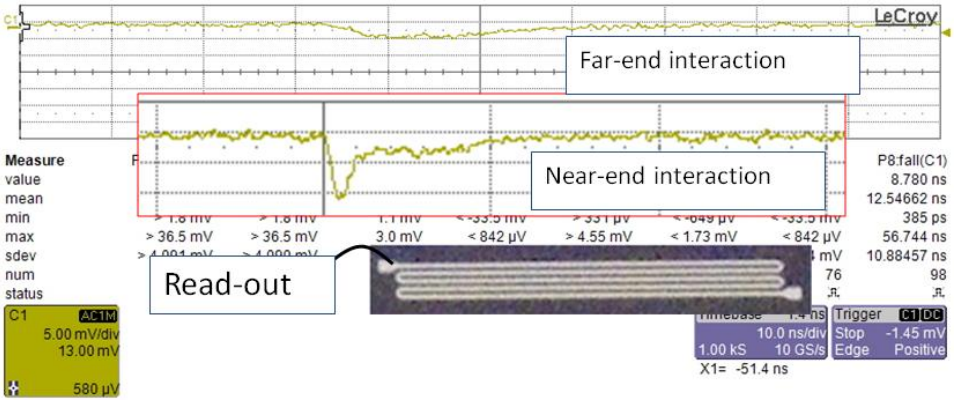
(a)



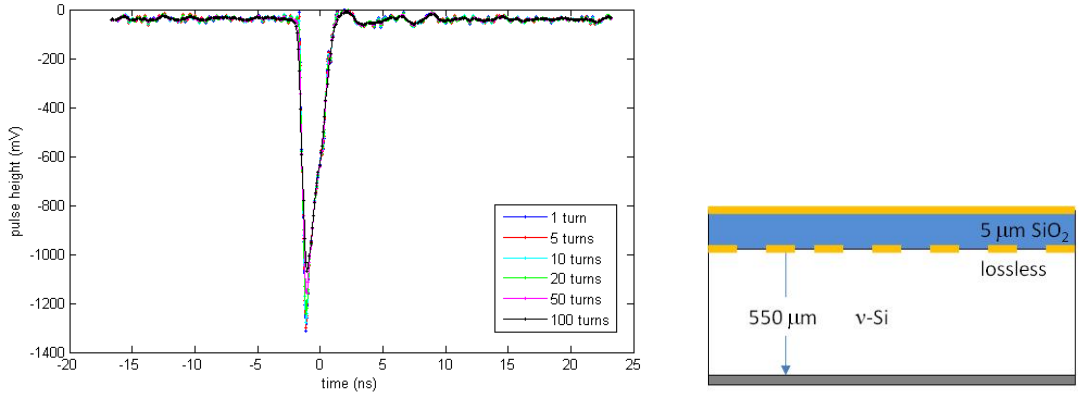
(b)

Fig, 2.42: (a) Illustration of the detector capacitance, which increases from a very low initial value, as governed by the area of induced charge, to a larger value, depending on the meander-line length. (b) The simulated shaping of a measured pulse, in green, by various detector capacitances.

For small line lengths, the capacitance has negligible effect on the pulse shape; however, if the capacitance is increased by 10 times (pink) or 100 times (blue), then the pulse is stretched. These simulated results are validated in Fig. 2.43, in which near-end interactions are minimally shaped (note similarity between pink pulse in Fig. 2.42b and near end interaction in Fig. 2.43) and the far-end pulses experience greater capacitance.



Fig, 2.43: For the meander line shown, the collimation of an alpha source on either the readout end or the far end of the line results in shaping due to increased line capacitance, as expected from Fig. 2.42.

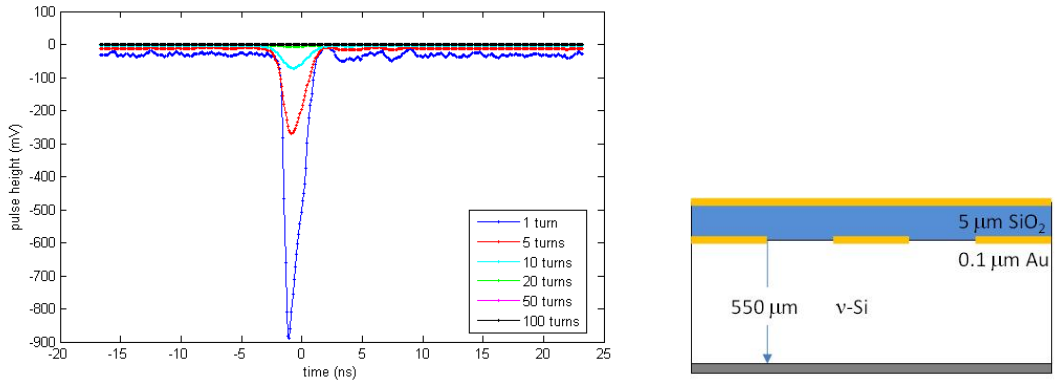


Fig, 2.44: For the lossless conductor, a 10 µm wide meander-line, and a 5 µm thick SiO_2 overlay, the resulting 50 µm structure results in minimal shaping and little attenuation.

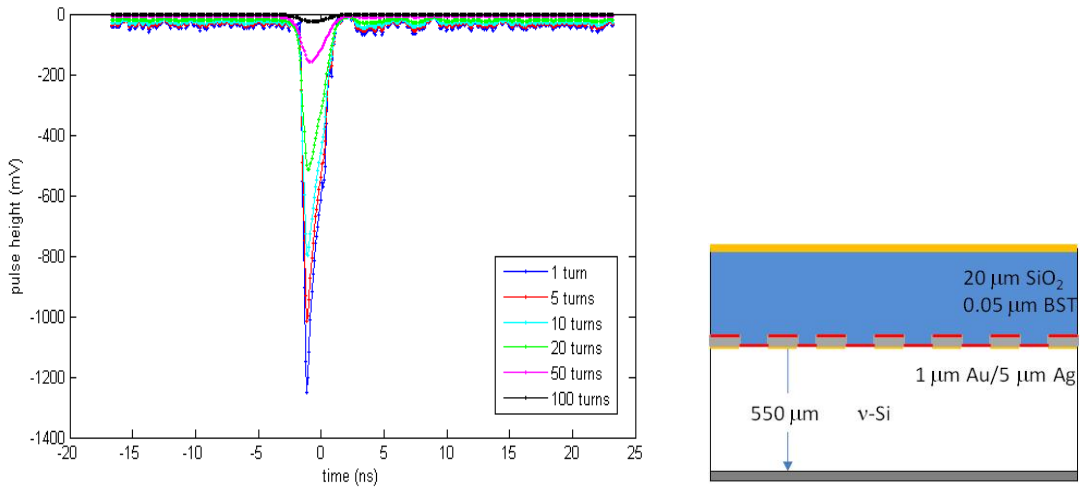
One must solve this capacitance issue therefore if the pulse shape is to survive propagation through a multitude of turns. As explained previously, the solution is to convert the simple strip-line into a *balanced* transmission-line via the addition of an overlay dielectric bounded by a ground plane. Note that the energy information is not lost during this shaping and it can therefore be tolerated if one is merely integrating the total current; however, if multiple interactions are coincident upon the detector, then the information is more extracted from the original induced signal.

The improvement in the shaping of the pulses is illustrated in Fig. 2.44, in which the transmission of a measured pulse through various line lengths is simulated with a custom Matlab code, in which the line response is derived from Sonnet. As shown in the

figure for a *lossless* conductor, the overlay 50 Ω design results in a signal that retains it shape with no apparent loss in fidelity even after passing through 100 sections, each 1 cm long. There is some slight attenuation (note different between blue and green peak values compared to black curve) due to dielectric loss, but this can be significantly reduced with the inclusion of BST.



Fig, 2.45: For a thin-film (0.1 μm) lossy conductor (Au) with a slightly unbalanced line (note the 100 μm width rather than 10 μm), the signal attenuated significantly within 10 or 20 turns.



Fig, 2.46: The addition of the 5 μm of Ag to the Au line reduces the conductor losses, and the BST reduces the degree of pulse shaping, compared with that shown in Fig.2.44.

The manner in which the *shaping* of Figs. 2.42 and 2.43 is avoided is therefore clear, and what remains is to minimize the *attenuation* due to conductor effects. As shown in Fig. 2.45, thin, poor conductivity conductors can significantly shape and attenuate the signal. We would prefer to propagate the pulses through hundreds of turns, at least, in order to minimize the readout complexity. As shown in Fig. 2.46, this is possible if: (a) thicker higher conductivity conductors are employed, and (b) a thin film of BST is added to the dielectric sandwich to minimize dielectric losses. As shown in the figure, a sizable signal is still present after 100 turns (each segment 1 cm long), and this is before full optimization of the design for the frequencies of interest.

The (partially) simulated results thus show that one can propagate the information along the lines as desired so that fine position-sensing is possible without enumerable readouts. This behavior is confirmed, as explained in the next section, in which the propagation properties of the lines with and without an overlay are compared.

2.11 Empirical Propagation Characteristics of Balanced Line

2.11.1 PECVD SiO_2 Overlay

The meandering lines were designed to widths 2, 5 and 10 μm, which were designed and processed on the same wafer. Since the overlay thickness was of common value for all the widths, only one of widths could be designed as a 50 Ω line (unique w/h ratio) per wafer. Hence, the 10 μm line over-layed with a 5 μm thickness of SiO_2 was the fabricated 50 Ω line for testing. The 2 μm (which needs 1 μm oxide) and 5 μm (which needs 2.5 μm oxide) lines, though overlayed with 5 μm thick oxide, were not considered for comparison as they do not present 50 Ω.

Clearly, from Fig. 2.47, the effect of the overlay is validated. With a fast-rise step input and narrow pulse input, the line without the overlay (unmatched in Fig. 2.47a) exhibited reflections in the form of ringing in the step pulse and distortion in the narrow pulse. The matched line however, outputs clean signals which look like the input signals (Fig. 2.47b). The step response of the overlayed line has a sharper fall time relative to the input step-pulse which is related to insufficient thickness (0.1 μm) of the lines. Because of increased skin depth losses associated with lower frequencies, they are subjected to

greater amount of attenuation and hence the chopping off of the long pulse; however, the rise time is maintained.

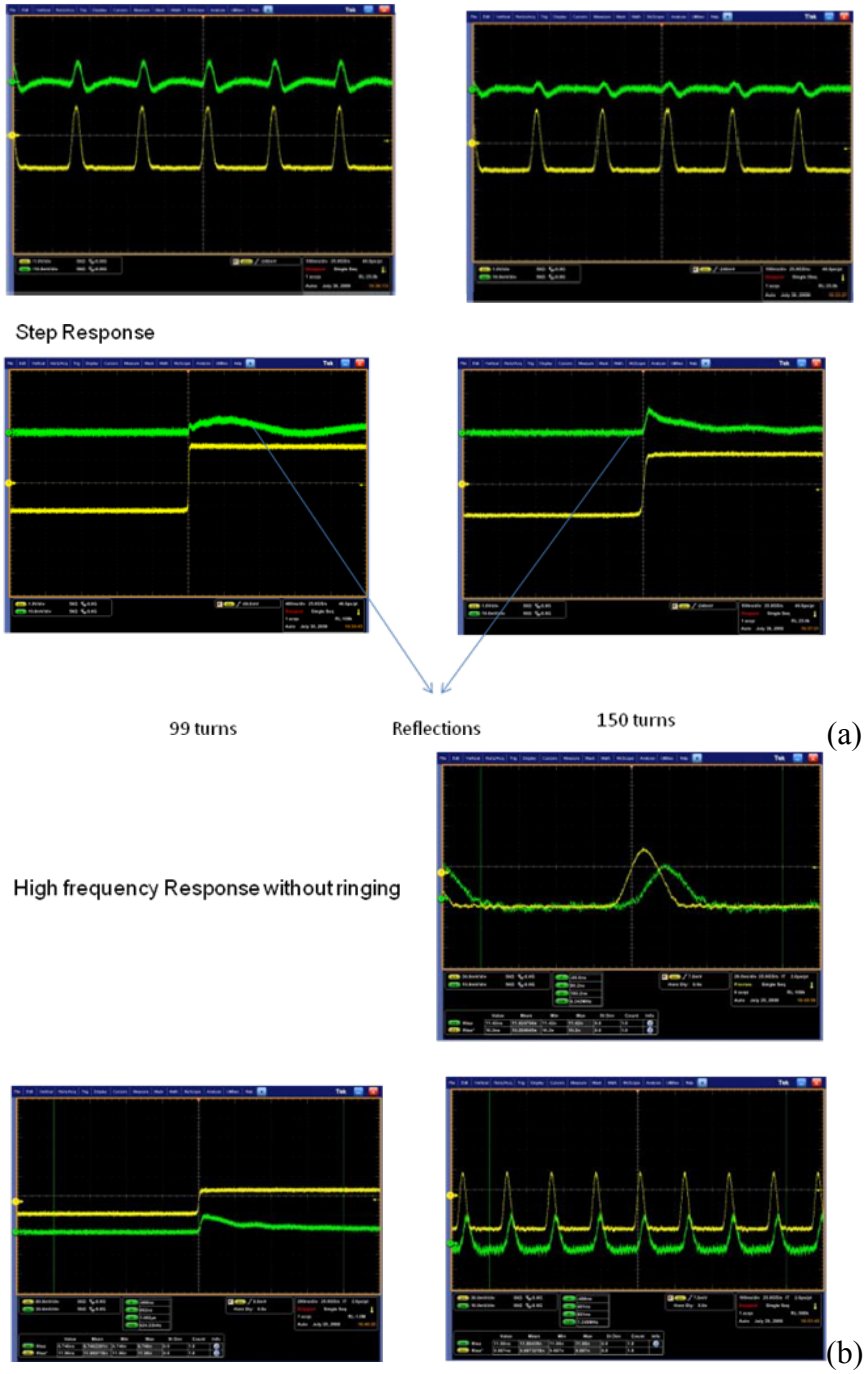


Fig.2.47: (a) 10 µm line without overlay showing reflections and pulse distortion. Note that yellow is the input and green is the output after propagation through 99 or 150 turns. (b) Lack of ringing or distortion due to presence of overlay.

Fig. 2.48, measured from the 5 µm (Fig. 2.48a) and 2 µm lines (Fig. 2.48b) *without* the overlay, exhibited greater reflection and distortion for the step pulse and narrow pulse inputs, relative to that measured from the 10 µm strip-lines (Fig. 2.47a), validating the fact that inductive and capacitive effects require compensation, particularly for smaller line widths, in order to maintain pulse integrity.

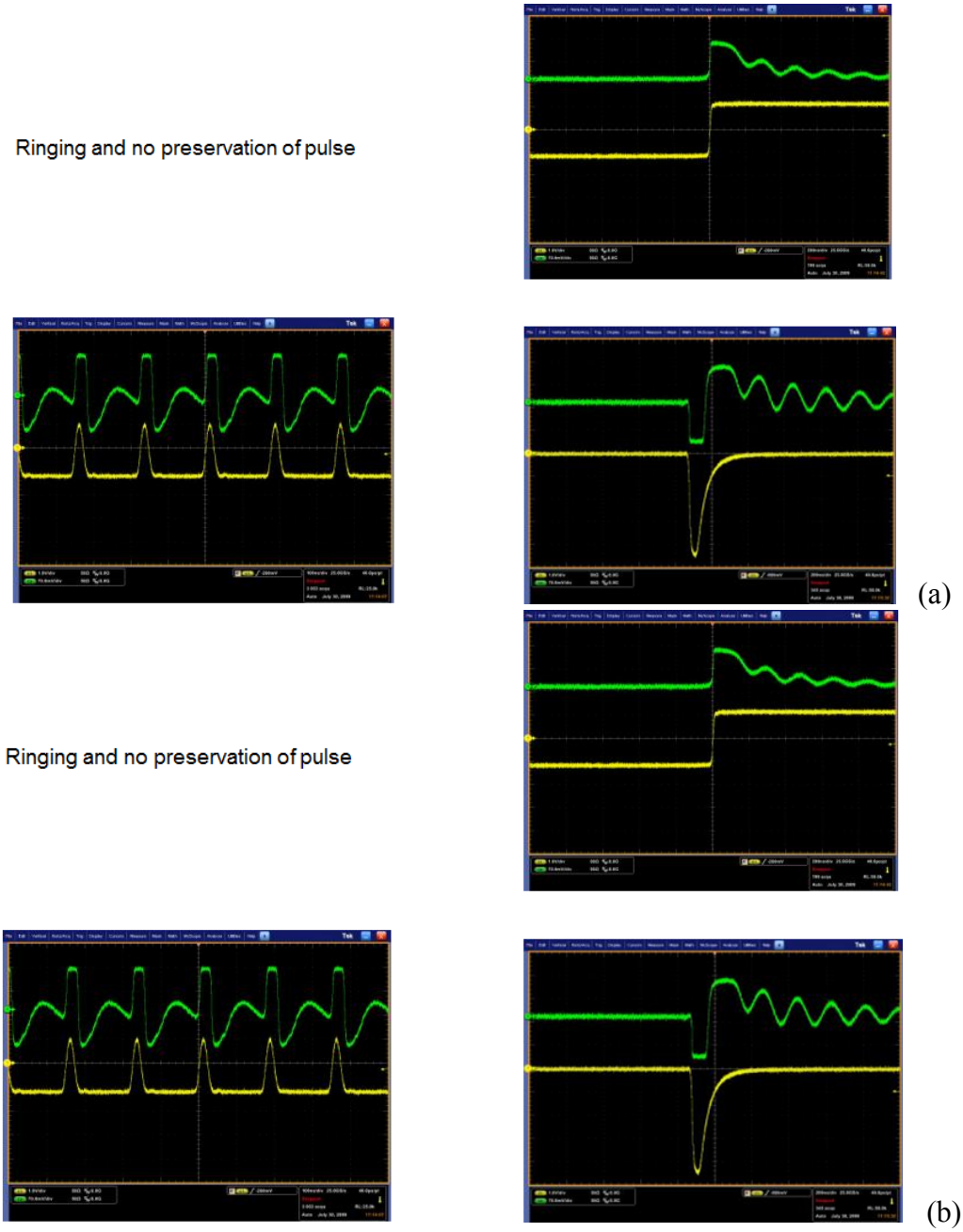


Fig. 2.48: Greater mismatch and ringing for: (a) 5 µm lines without overlay and (b) 2 µm lines without overlay

2.11.2 Plastic Block Overlay

Recall that one method by which one can process the dielectric overlay is to pattern the structure on a separate substrate and then mechanically or thermally couple the structure to the existing low-noise detector. We want to investigate this methodology for application to semiconductors with more exotic processing challenges than silicon; most particularly, the high Z, high band-gap semiconductors such as CZT and TlBr. During the research, we therefore tested the application of so-called block overlays and found that they produce the desired effect. As shown in Fig. 2.49, we created a matched line with a metalized polycarbonate sheet (i.e. adhesive conducting carbon tape connected to Saran-wrap) with a thickness of 10 μm. Since the permittivity of the polycarbonate ($\varepsilon_r = 3$) is similar to the SiO_2 ($\varepsilon_r = 3.9$) overlay that we have fabricated, the results are expected to be similar.

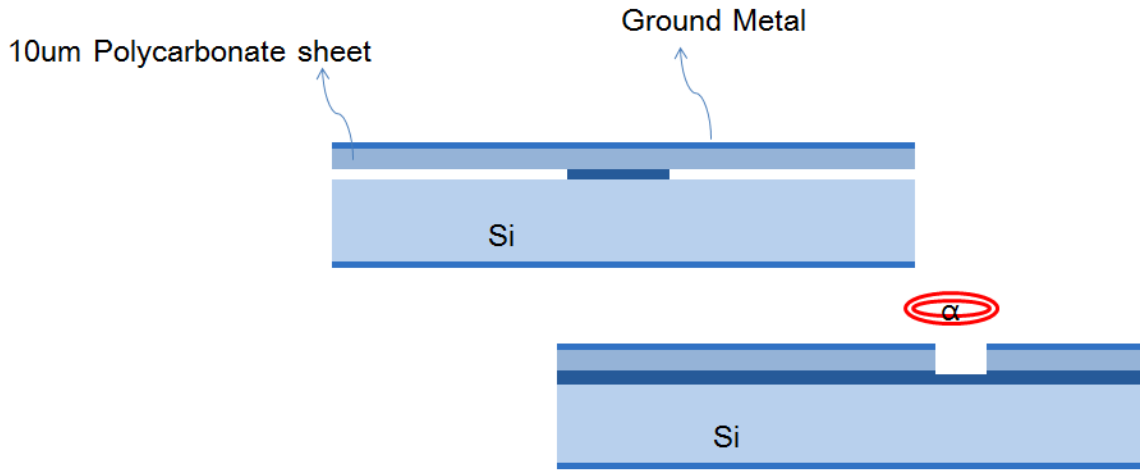


Fig 2.49: Arrangement to achieve 50 Ω line structure with a block polycarbonate overlay.

The 5 μm wide 75 turn meandering lines were used for the tests. The 10 μm polycarbonate metalized sheet, when placed on top of detector, converted the line to present an impedance of 50 Ω. In order to permit ion irradiation of the detector, we created vias in the overlay, as shown in Fig. 2.49.

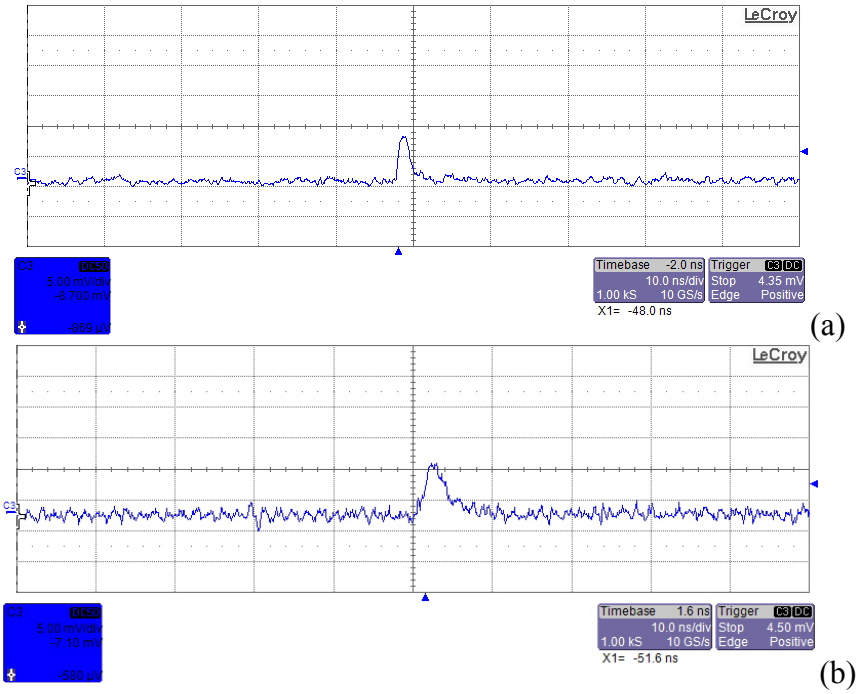


Fig. 2.50: (a) Typical signals from meandering line detectors without any kind of overlay. (b) Pulses from a block overlay made out of metalized polycarbonate sheet placed mechanically on top of the lines.

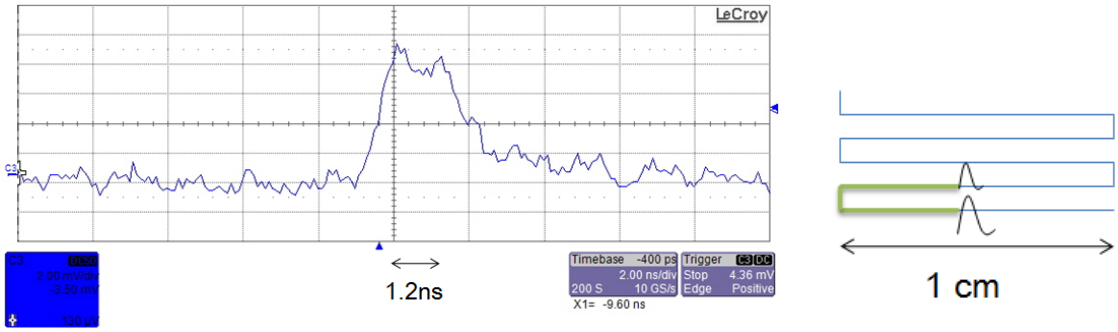


Fig. 2.51: Double humped pulsed with the polycarbonate overlay on top the meandering lines.

The effect of the overlay was an increase in the width of the pulses that propagated through the line with clear double humps close to the maxima (Fig. 2.50 & 2.51), the humps separated by 1.2 ns. This time difference is consistent with the signal propagation as *measured* with a time-to-amplitude converter, however, it is significantly slower than that modeled with Sonnet, which reflects that the bounding semiconductor participates in signal transmission, an effect not captured with the electromagnetic modeling software.

Nevertheless, we are able to see signals from both ends of the meandering line, as shown in Fig. 2.52, for two signals that are band-passed, and therefore shaped. Co-event signals were observed from both the ends, although for large turn multiplicities (order of 100), the signals can be small necessitating the use of downstream amplification and signal conditioning components.

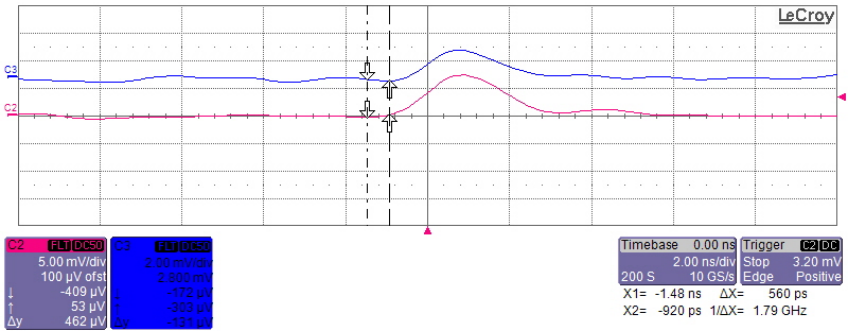


Fig. 2.52: Signals from the two ends of 5 µm, 7 turns meandering lines through the TIA, SFT-0100, when sourced somewhere in the middle of the length of the detector. (This may not be middle of length of the lines as the collimation was 300 µm for a 5 µm wide line.)

2.12 Group Velocity Control via Transmission-Line Design

As long as one avoids information overlap from a multitude of lines and depths, the track can therefore be reconstructed; however, the pulse information wants to travel at some significant fraction of the speed of light along relatively short sections for standard electrode geometries. We wished to therefore determine the transmission-line design which gave us adequate control over the propagation speed so that overlap between no more than two-legs occurred. Assuring adequate time separation can be achieved most simply by making longer lines; however, one can also employ slow-wave structures, for which there is an extensive body of literature. A well-known slow wave structure is a metal-insulator-semiconductor (MIS) based line (Fig. 2.53a), in which a thin layer of insulator separates the metal electrode from silicon. In a homogeneous dielectric material, with a uniform conductor, the electric and magnetic fields penetrate to the same depth, and the velocity of propagation equals $1/\sqrt{\varepsilon\mu}$. As indicated in Fig. 2.53b, in a MIS structure, the electric fields penetrate to a depth of h_1, and the magnetic fields penetrate down to h_2, disconnecting the homogeneous assumption.

(a) (b) (c)

Fig. 2.53: (a) MIS slow wave geometry. (b) Illustration of differing depths to which the electric and magnetic fields penetrate. (c) High permittivity dielectric in the overlay to slow down the propagation.

The resulting combination of large capacitance and large inductance slow the velocity of signal propagation ($v = 1/\sqrt{LC}$, for lumped transmission model). Unfortunately, the frequency dependency in such lines results in phase distortion, and from a detection standpoint, such an arrangement requires proximity sensing rather than charge collection.

We therefore modeled a high permittivity layer in between the SiO_2 and the metal, as shown in Fig. 2.53c, and simulated its response. We found that this helps to increase the group delay of the pulses without any significant group delay distortion. Perovskites are a group of ferroelectric crystalline ceramics with high permittivity which are typically used in varactor-like applications because their dynamic permittivity which can be varied with varying voltages. We explored two such compounds, Barium Strontium Titanate (BST) and Barium Titanate (BT). A 50 μm thick layer of BST (ε_r=300) could not raise the effective permittivity combined with SiO_2 as much as BT (ε_r=1000), as shown in Fig. 2.54. BT that is 50 μm thick shows an order of magnitude increase in group delay compared to MIS or 5 μm BST based lines, however, both improve the propagation characteristics and the BST is a slightly more mature material. We continue to consider both materials in our optimized high-permittivity designs. Note in Figs. 2.54a and b that the BST or BT are generally bonded to thick (500 μm for instance) SiO_2, which is no longer a thin film conducive to microelectronic processing. We thus employed glass and quartz wafers as the transmission line substrate upon which BST or BT was deposited. The block overlay is placed or bonded to the detector to improve the signal propagation characteristics.

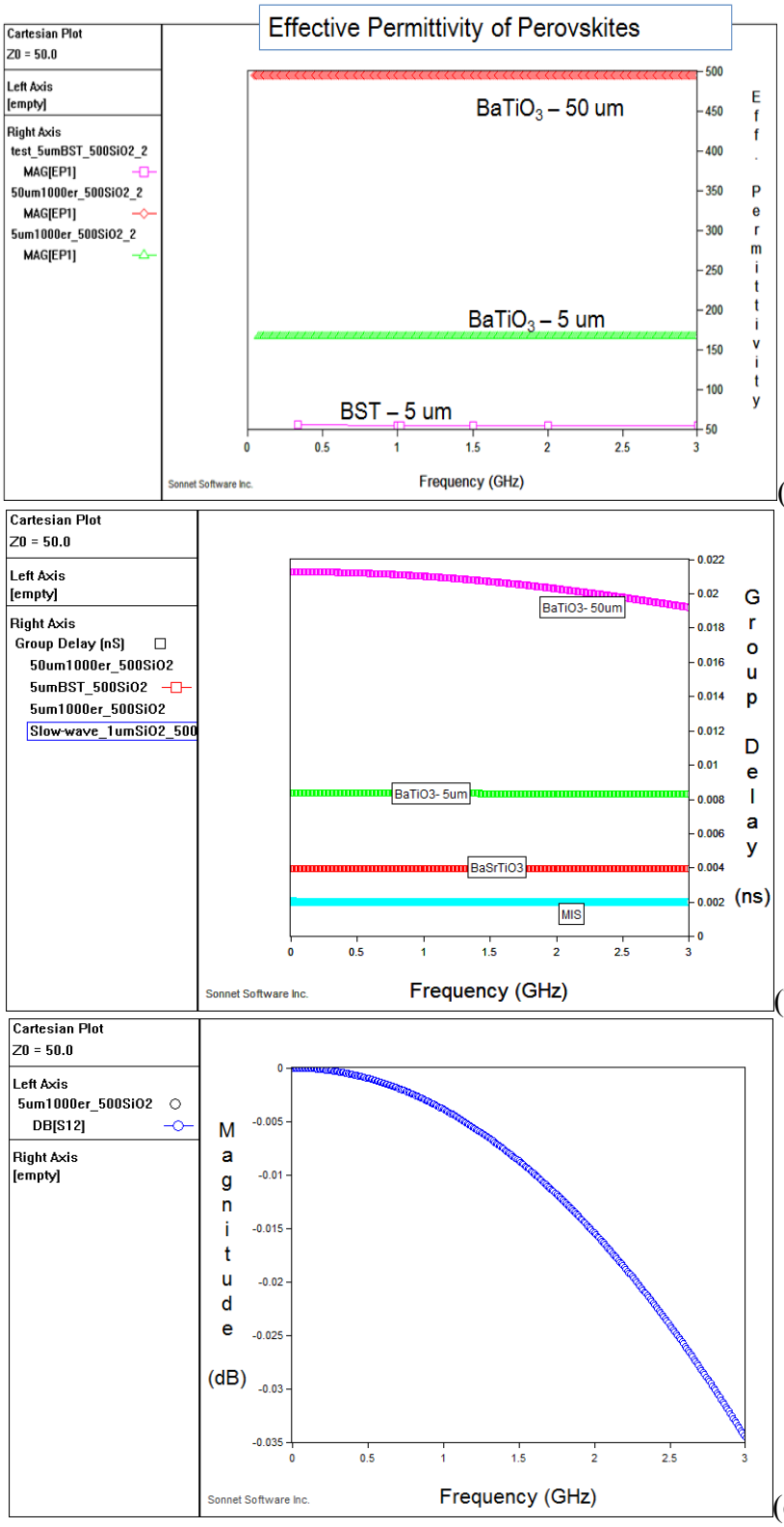


Fig. 2.54: Comparison of (a) effective permittivity, (b) group delay between BST, BT and SiO_2. (c) S21 of BT based line.

Note that we also wish to confirm that the loss properties of the BT are good, as indicated in the modeling of Fig. 2.54c, which shows that the transmission coefficient (S21) shows very little loss over the frequencies of interest.

The frequency response plots are transformed into spatial plots in Fig. 2.55 in order to clarify the impact of the propagation characteristics of the transmission line on the performance of the imager. The top two figures (simple strip-line, asymmetric strip-line with SiO_2 overlay) correspond to the two structures that we have realized experimentally to date, and the bottom two figures correspond to two potential improvements of the existing designs. Taking the simple strip-line (upper right figure) as an example, the figure shows the vertical design of the structure sitting atop the silicon detector on the left as well as a picture of the meander line on the right.

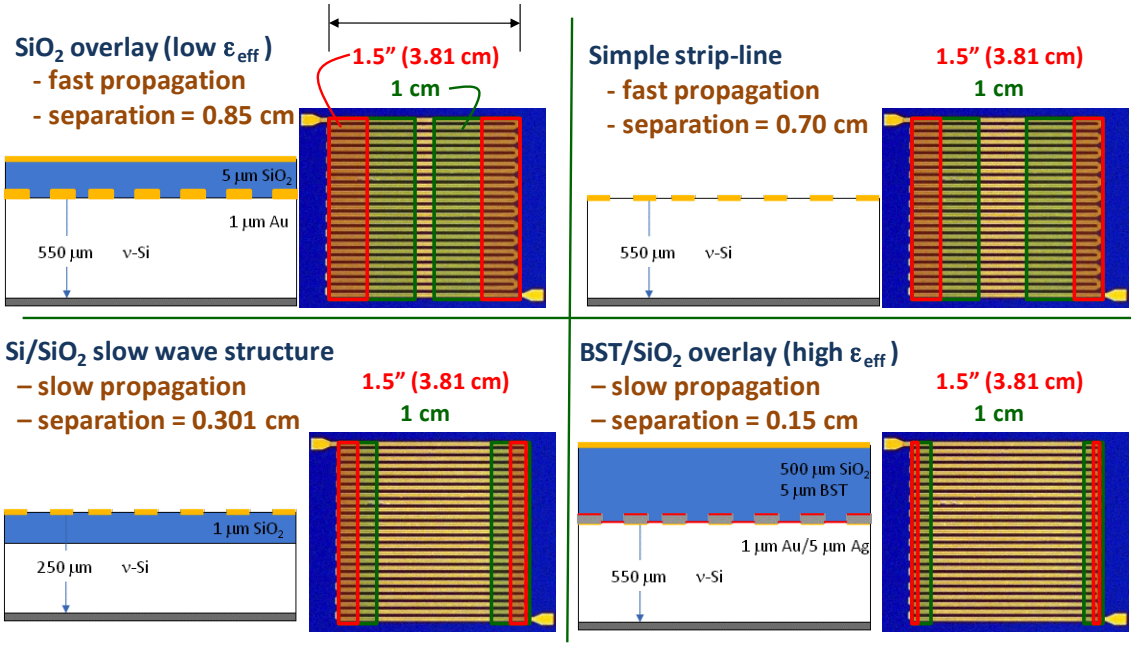


Fig. 2.55: The effect of the transmission line design on the area of the detector that will yield pulse overlap, assuming an overlap time of 380 ps (experimentally measured for an alphas particle with range 28 μm). The green rectangles indicate those regions close enough to the edge of a 1 cm line length detector to experience overlap, and the red regions show the overlap probability for a detector with 1.5" long legs.

We have processed different overall detector sizes but have mainly focused on 1 x 1 cm detectors and 1.5" (along the line) x 1 cm detectors, and they are therefore

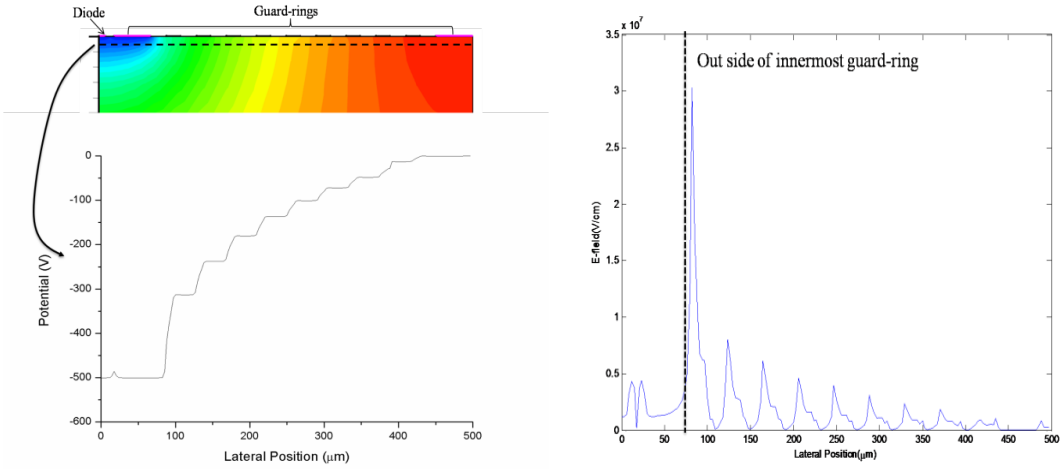
considered in the Fig. 2.55. The green boxes indicate that if the interaction is within that box, then it is close enough to the edge to result in overlap between neighboring lines. The red box indicates the probability of overlap for the longer 1.5" lines. As discussed in Chapter 3, for the straight tracks from ions, dual overlap can be cleanly separated during the analysis stage; however, for circuitous tracks such as those from electrons, we would like the open areas shown in Fig. 2.55 to be as large as possible.

For the asymmetric strip-line with SiO_2 overlay (upper left), the propagation speed is actually worse than the simple strip-line (but recall that it does effectively preserve the pulse shape by converting the capacitive line into a resistive line), in that there is a high probability of overlap for electron tracks. However, for the slow-wave structures modeled along the bottom row, most of the area is open and if the interaction is near the edge of the meander-pattern, then that can be corrected. Thus, for both preserving the pulse-shape and minimizing the propagation speed, the high permittivity dielectrics are optimal.

2.13 Semiconductor Device Simulation (Sentaurus Detector Modeling)

2.13.1 Metal field palates effects

For a 550 μm thick silicon detector, a bias of only 70 V is required to fully deplete the active area, as revealed through the current-pulse shape variation. That is, when the delay-line detector is fully depleted, the current signals scale with the detector bias (recall $I = qvE_w \sim V_{det\ bias}$); however, under partial depletion, the growth in the depletion depth compensates for the increases voltage and the current shows an approximately constant signal. In order to saturate the charge-carrier velocities, and hence maximize the *current* signal amplitude, an applied voltage near 1000 V is required, the application of which requires careful control of the electric field through geometric designs as well as optimal insulator selection. In order to find proper operational parameters for the fabrication of the delay-line detectors, we simulated with Sentarus TCAD [15] the effects of metal field plates (MFP's), floating multiple guard-ring structures, as well as biasing at the inner-most guard-ring. Fig. 2.56 shows the structure of the delay-line detector, in which every guard ring except the outermost has an inward (toward the detector area) metal field plate. The inward MFP can improve the leakage characteristics via the accumulation of electrons at the Si-SiO_2 interface due to positive oxide charge, but it degrades the breakdown characteristics.



(a) (b)

Fig. 2.56: Potential contour and its graph (a) and electric field graph (b) with 10 floating multiple guard-rings and bias (-500 V) at innermost guard-ring.

The serpentine delay-line has an outward overhanging metal field plate, which improves the breakdown characteristics since it can smooth the electric field at the right edge of the delay line where the field gradient is strongest. In order to minimize the leakage current and increase the breakdown voltage, we modeled various geometric designs for the over-layed delay-line detector, in which the diffusion barrier and electric field are formed by a p+ junction. The processing of the structure is discussed in Chapter III.

When the detector and innermost guard ring are biased to -500 V, and the outermost guard ring is biased to 0V, the electric potential 1 μm from the surface of the detector is shown in the graph of Fig. 2.56a (bottom), and the contour plot (top of Fig. 2.56a) shows its variation laterally (to 500 μm from the center of the detector) as well as in depth (to 20 μm from the anode surface). Compared with a detector absent a guard-ring structure, the gradient in the potential is reduced through its stepped division across the guard ring structure, the size of the voltage steps decreasing with larger number of rings. As shown in Fig. 2.56b, the guard ring structure successfully moves the strongest electric field region, and hence the strongest driver of leakage current, away from the active diode region to the outside of the inner guard-ring.

2.13.2. Inner guard-ring bias effects

If one compares the field with a device in which there is no bias at the inner guard-ring, then the highest electric field appeared at the edge of the active diode, as shown in Fig. 2.57, and the leakage current to the detector increases by almost three orders of magnitude, as shown in Fig. 2.58. Fig.2.59 shows the innermost guard-ring effects. Most surface leakage currents are collected at the inner guard-ring, so the total leakage current in the active area can be decreased by 10 times.

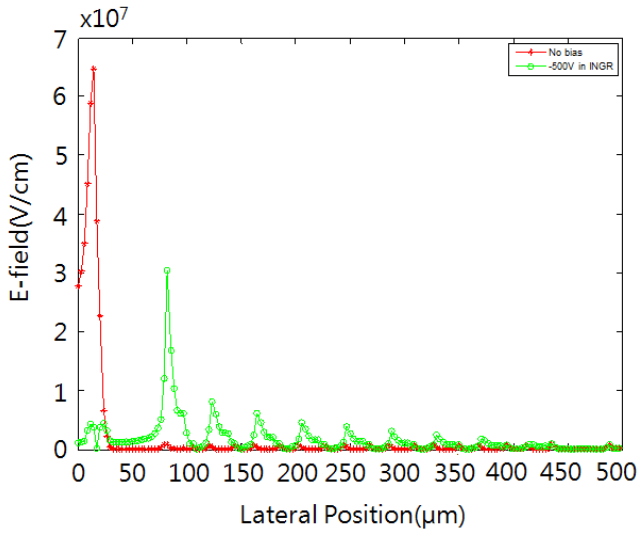


Fig.2.57: Electric field graph with (green) and without (red) bias voltage (-500 V) at inner most guard-ring.

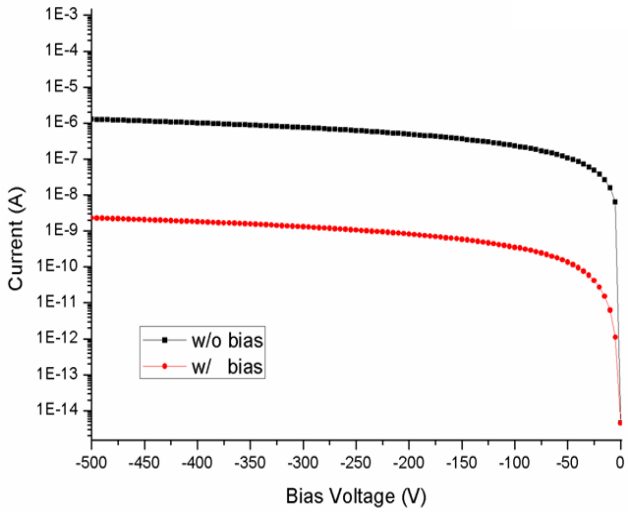


Fig.2.58: I-V characteristics with (red) and without (black) bias voltage (-500 V) at inner most guard-ring.

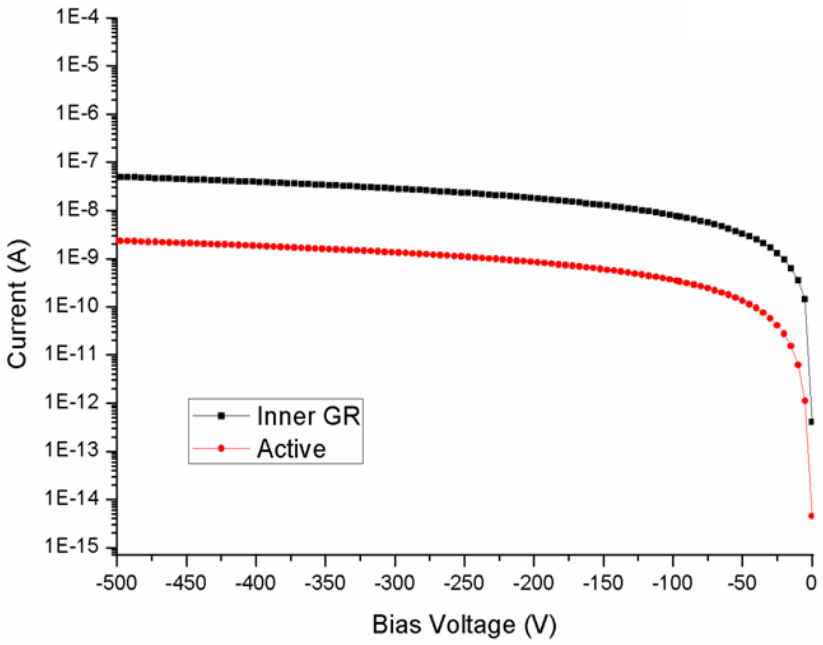


Fig.2.59: I-V characteristics comparison between active area of diode and innermost guard-ring with bias voltage (-500 V).

2.13.3 Passivation layer thickness effects

Beyond moving the high electric field away from the active region, one also endeavors to minimize its amplitude regardless of its position, in order to reduce the probability of soft or hard breakdown. As shown in Fig. 2.60, this reduction can be elicited by increasing the width of the metal field plate. As the MFP extent is increased from 3 μm to 5 μm, the magnitude of the maximum field decreases by more than a factor of 2, which implies that one can increase the applied detector bias commensurately. Although the surface leakage current may be the main source of noise, it can be collected by the inner guard ring, as shown in Fig. 2.61, in which 50 nA of leakage current is collected by the inner guard-ring and a current of only 1 nA current makes it to the detector.

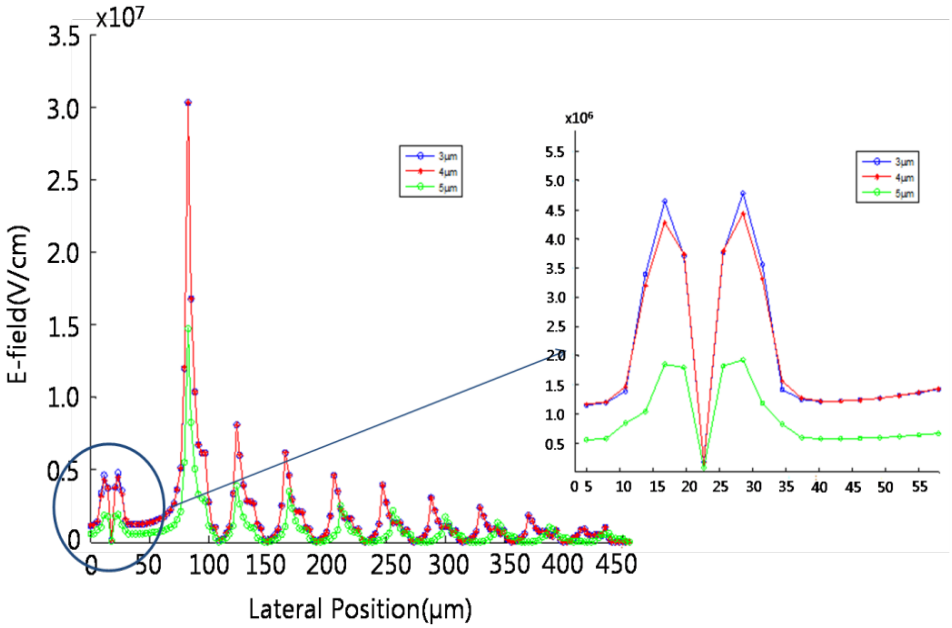


Fig.2.60: Electric field graph with varying the thickness of passivation SiO_2 layer from 3μm to 5μm with biasing (-500 V) at innermost guard-ring.

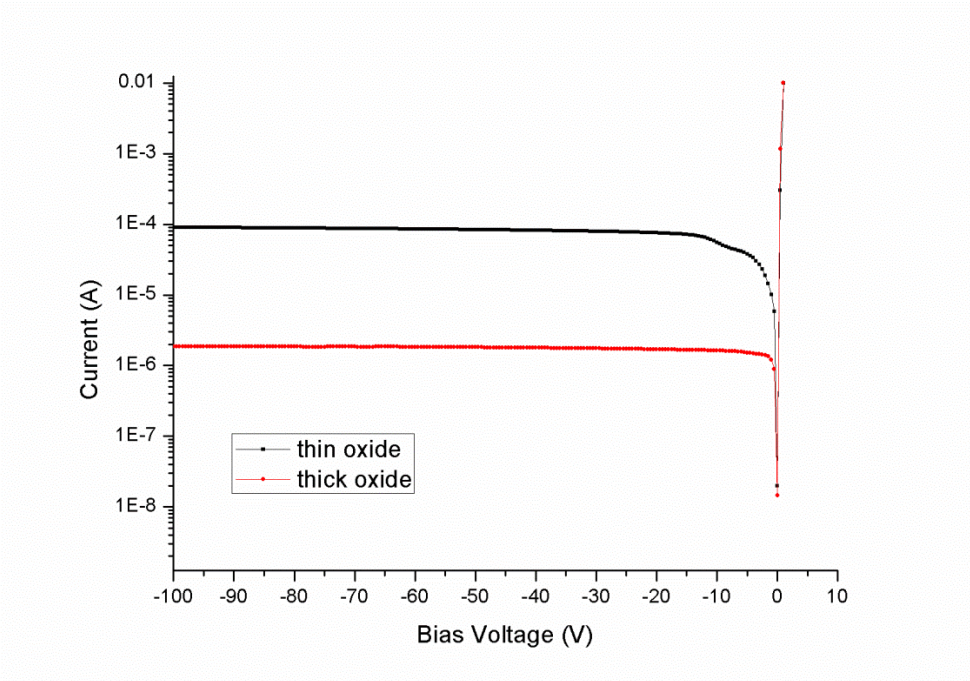


Fig.2.61: I-V characteristics comparison with thin oxide (50 Å, black) and thick oxide (5000 Å, red).

2.14 References

[1] MAXWELL 3D, Ansoft, Four Station Square, Suite 200, Pittsburgh, PA 15219, USA.

[2] GEANT4 website: http://geant4.web.cern.ch

[3] G. F. Knoll, Radiation Detection and Measurement, 3rd Ed, New York: John Wiley & Sons, Inc., 2000.

[4] Z. He, "Review of the Shockley-Ramo theorem and its application in semiconductor gamma ray detectors," Nuclear Instruments and Methods in Physics Research Section A, vol. 463, no. 1-2, pp. 250-267, 2001.

[5] L. C. Feldman and J. W. Mayer, Fundamentals of Surface and Thin Film Analysis, Elsevier Publishing Co., 1986.

[6] M. Jeong, D. A. Lawlor, and M. D. Hammig, "Optimization of the Position Resolution in Semiconductor Detectors", 2007 IEEE Nuclear Science Symposium Conference Record, pp. 1456-1462.

[7] M.D. Hammig, M. Jeong, D.K.Wehe, and S. Ramadoss, "Delay-line electrode partitioning as a means of simplified position sensing", IEEE Nuclear Science Symposium Conference Record, Nov. 2008.

[8] M.D. Hammig and R. Stevenson, "The degree of enhancement in a gamma-ray image gleaned from recoil-electron tracking", IEEE Nuclear Science Symposium Conference Record, Nov. 2008.

[9] M. Brigida, "A new Monte Carlo code for full simulation of silicon strip detectors", Nuclear Instruments and Methods in Physics Research A, vol. 533, pp. 322–343, 2004.

[10] K. Seeger, Semiconductor Physics, Springer, Wienna –New York, 1973.

[11] User's Guide, Sonnet Suites 11, p. 20-21.

[12] E.H. Fooks and R.A.Zakarevicius, Microwave Engineering using Micostrip Circuits, p.214.

[13] T.C. Edwards, *Foundations for Microstrip Circuit Design*, p.129 – p.151.

[14] Douglas Brooks, Crosstalk Coupling: Single ended vs. Differential (http://www.ultracad.com/mentor/crosstalk_coupling.pdf)

[15] Sentarus Device, Synopsys Inc., June 2006.

CHAPTER III

DETECTOR FABRICATION

Although we employed detectors with smaller partitions, our experimental development focused mainly on devices with 20 micrometer and 200 micrometer lateral resolutions. Those diagnostic detectors were used to find a more tolerable solution to the front-end readout circuit, in terms of its noise and timing performance, and to develop the fabrication methods for the overall detector assembly. As summarized in Chapter II, we accomplished the following: (a) optimized a fabrication method for the low leakage current detectors, (b) iterated through three different mask designs in the detector fabrication and produced a robust process, (c) fabricated mounts for the on-chip amplifiers, which were wire-bonded to the detector's meandering electrode, and (d) developed mechanical mounted procedures for the detectors and electronics on a printed circuit board

3.1 Detector Design and Fabrication

The performance of the detectors was inhibited by a high leakage current, measured in the 10's of μA range, although our goal, for charge sensitivity reasons and implementation concerns, is to yield detectors with leakage currents in the nA range or less. For the 3rd generation meander line detector, the leakage current was reduced by 10,000 times to the desired range, to less than 10 nA, which required multiple improvements in the processing, most related to controlling the surface leakage. First, an n+ cathode contact was used to block minority carrier motion, to lower the contact resistance, and to improve the ohmic properties of the aluminum contact. Second, a dry silicon dioxide was used to passivate the surface between the legs of the meander and across the open areas of the detector, which resulted in an order of magnitude

improvement in the leakage current. Third, sputtering was found to produce lower-leakage contacts, due mainly to the pre-deposition surface-cleaning plasma. Fourth, the surface state density and interband states were quenched by saturating the silicon dangling bonds by annealing the samples in a hydrogen atmosphere, which yields approximately 100 times improvement in the leakage characteristics. Finally, a guard ring structure, with a biased inner guard ring and six floating guard rings, was used to lower the leakage by another order of magnitude, to reach the desired noise goal. These extra steps slightly complicate the fabrication, but only extend the process time to 3 days rather than 1 day if one is simply making un-passivated metal-semiconductor detectors.

As summarized in Section 3.1 below, these passivation processes were applied to both Schottky-barrier detectors and PIN structures, the latter yielding slightly higher leakage and more extensive processing (2 weeks to fabricate rather than 3 days). However, the PIN structure has the advantage that the metal electrode structure, which bounds the detector does not participate in forming the electric field in the semiconductor bulk. Thus, one can choose the metal based on transmission-line design criteria alone. This is important since the length through which one can transmit the pulse information is limited principally by the conductive losses. The base semiconductor detector was thus realized with acceptable noise characteristics, but when implemented with a thick high-conductivity electrode and a dielectric overlay, the performance can be compromised.

3.1.1 Minimization of Charge Loss: Contact Characterization

In Chapter II, we discussed our development of the numerical simulation tools- GEANT4 for radiation interactions, MAXWELL for electric field modeling, custom codes for charge transport- that can be used to investigate different detector designs and validate experimental measurements. Nevertheless, the codes used to model radiation detectors do not include detailed material effects that can vary for different processes and over various time scales. For instance, the intrinsic stress of a metallic rectifying contact can change the bond lengths in the bounding semiconductor material and thus increase the surface-state density, the consequence of which is to increase the leakage current and reduce mechanical adhesion. Diffusion of oxygen through the metal can deteriorate the device's properties over time because of the formation of an intervening oxide, the presence of which reduces the barrier height and increases the probability of trapping at

the oxide-semiconductor interface. Finally, the details of the electrical contact can impact the device performance- a mechanical contact, for instance, can damage the interface, and increase the interband surface-state density; the pressure and heat of the wire-bond can alter the inter-diffusion characteristics of the contact. For instance, Fig. 3.1 shows the high sensitivity of the electrical properties of the contacts to their handling. Fig. 3.2 shows the performance degradation that can occur when wire bonding to a rectifying junction because of the damage caused by heating and pressure. The goal of surface process work is to develop fabrication recipes for which these effects are inconsequential, thus providing a path to robust contact formation.

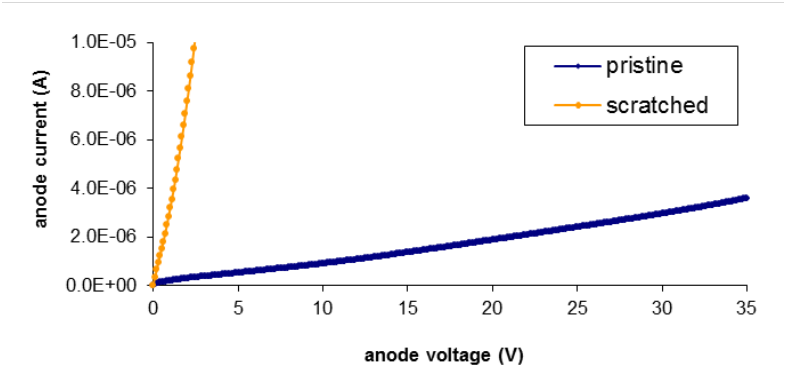
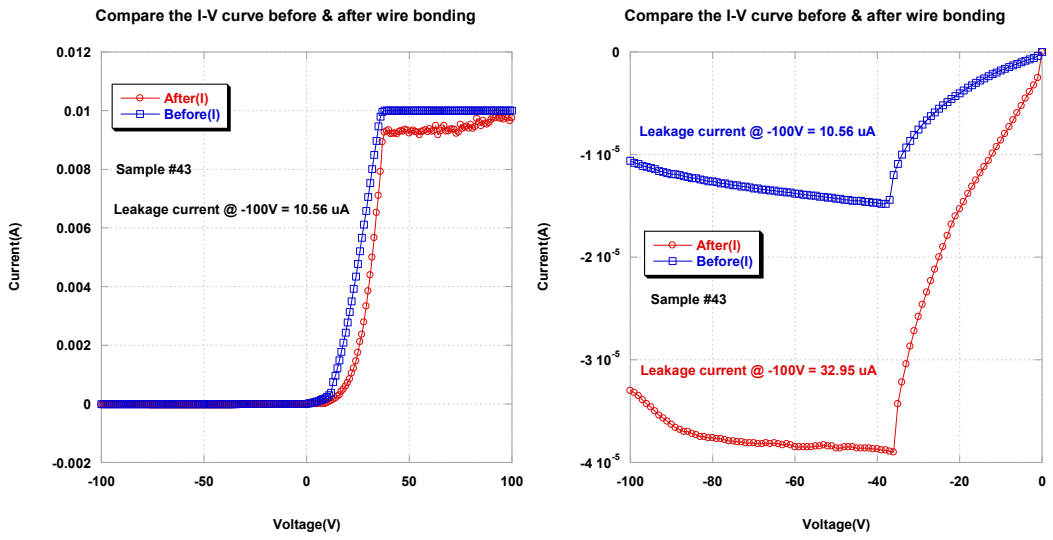


Fig. 3.1: Deterioration in the reverse-bias (shown in positive voltages) characteristic upon the creation of a conduction path during device coupling to a readout circuit.



(a) (b)

Fig. 3.2: Comparing the I-V curve before and after wire bonding (a) and its magnified graph at reverse bias region (b).

The solutions to these problems were developed. The intrinsic stress can be managed by interstitial low-stress layers, which also have the added advantage of lowering the fluidic diffusion rate. That is, a thin titanium layer can increase the surface adhesion of platinum contacts and a chrome layer commonly serves as an adhesion layer for gold. The interstitial layers have lower work functions than the contact layers and must therefore be made thin (10's of Angstroms) so that the diffusion barrier height does not decrease. As a specific example, platinum, with a work function of ~5.6 eV, as determined from IV-measurements and from consultations with the literature, has a diffusion barrier of approximately 0.9 eV [1,2,3]. If the titanium, with a work function of 4.3 eV begins to dominate the junction characteristics, then the diffusion barrier decreases and the leakage current increases. For thin interstitial layers however, the leakage current is preserved and therefore advantages of the mechanical improvements afforded by the added material can be taken.

The variation in the junction characteristics was also investigated as the nature of the electrical contact to the readout circuit was varied. The I-V characteristic indicates the underlying nature of the metal-semiconductor junction. That nature can vary markedly from point-to-point on a device that has undergone nominally identical processing [4]. That variation was sourced in alterations in the native oxide thickness and the degree of micro-scale heating during the evaporation procedure, the magnitude of which can be reduced by pre-deposition oxide etching (in HF or BHF) and the replacement of evaporation by sputtering.

One other consequence of the oxide growth, as discovered via experiments, was that the junction characteristics deteriorated over time. For instance, for Pt (rectifying) /Al (ohmic) detectors fabricated on n-type silicon, the typical leakage current is on the order of 3 μA, immediately after processing. After about a month however, that current would increase to ~10 μA, and upwards to 10's of μA's in subsequent months. In order to prevent edge leakage contributions, we passivated the surfaces, and tested the hypothesis that oxygen diffusion was impairing the intimacy of the Pt/Si contact, by heating the samples in a furnace. The long-term deterioration was replicated during these anneals. The surface state density (N_{ss}) was also determined, as described below, and exhibited an increase with time (or temperature), indicating that trap-assisted tunneling may be

contributing to the leakage current. Specifically, N_{ss} increased from 3.6 x 10^{16} eV^{-1} m^{-2} to 4.2 x 10^{16} eV^{-1} m^{-2}.

For compound semiconductors and composite materials, charge trapping can dominate the detector's performance, and we therefore investigated the alteration in the surface-state density at the interface. The charge trapping characteristics of defects can be measured using techniques such as deep-level transient spectroscopy (DLTS) [5]. Early in the research, we used the IV-characteristic to investigate the junctions; in particular, the surface-state density can be studied as various process parameters are altered. Specifically, one can model the metal-semiconductor junction with an interfacial layer using the thermionic emission theory model, where the emission current is expressed as:

$$I = aA^{**}T^2 \exp(-\chi^{0.5}d)\exp(-\frac{q\Phi_{bo}}{kT})\exp(\frac{q(V-IR)}{nkT})\left[1-\exp(-\frac{q(V-IR)}{kT})\right] \qquad (3.1)$$

where k is Boltzmann's constant, T is the temperature, n is the ideality factor, A^{**} is the effective Richardson contact, a is the diode area, Φ_{bo} is the zero-bias barrier height, and the $\exp(-\chi^{0.5}d)$ is the transmission coefficient through any interstitial layer [6]. For a metal-insulator-semiconductor diode with interface states N_{ss} in equilibrium with the semiconductor, the ideality factor can be related to N_{ss} as follows:

$$n(V) = 1 + \frac{\delta}{\varepsilon_i}\left[\frac{\varepsilon_s}{W_D} + qN_{ss}\right] \qquad (3.2)$$

Furthermore, the applied voltage can be related to the energy of the interface states, E_{ss}, with respect to the bottom of the conduction band, E_c as follows [7,8]:

$$E_c - E_{ss} = q(\Phi_e - V) \qquad (3.3)$$

Thus, using forward and reverse IV characteristics, like that shown in Figs. 3.3 and 3.4 in two different formats, one can evaluate the surface state density as a function of energy across the band-gap, if desired. That is, the linear portion of the forward I-V characteristic (< 1 V in Fig. 3.3) can be used to extract the ideality factor, n, and the zero-bias barrier height via the slope and intercept. The ideality factors, in the range of 1.08 to 1.75 for different samples, indicate the presence of an intervening oxide to various

degrees. At high forward bias, the response is dominated by the series resistance and in the intermediate regions, the presence of trap-assisted tunneling and R_s play a role.

The diagnostic detectors were separately investigated (over 100 in all), and those exhibiting the best performance (in terms of low leakage current) were mounted on a detector-PCB module, as shown in Fig. 3.5. The detector module allows one to separate the preamp design board from the detector board so that different circuit designs can be rapidly tested.

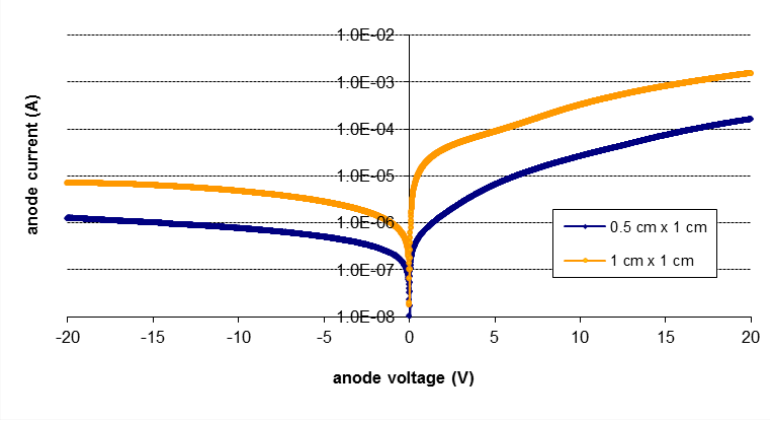


Fig. 3.3: Reverse and forward IV characteristics for Pt/n-Si/Al diodes for two different sizes (semi-log plot).

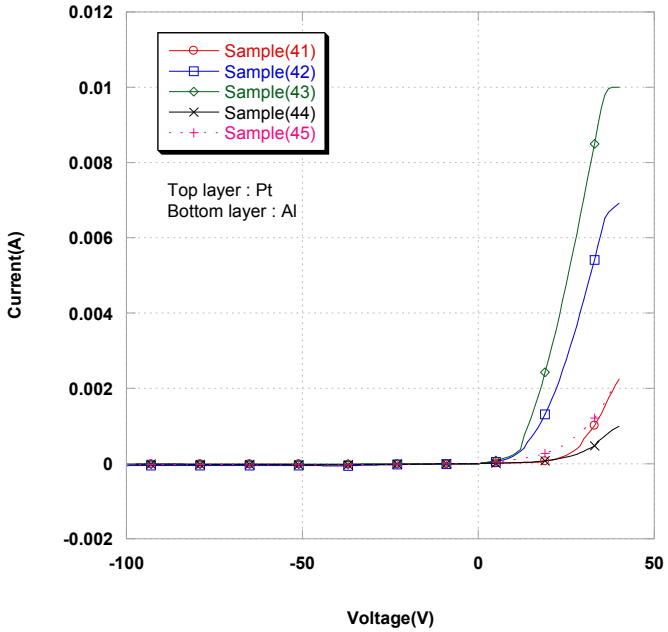


Fig. 3.4: For various detector samples, the reverse and forward IV characteristics for Pt/n-Si/Al diodes (linear plots).

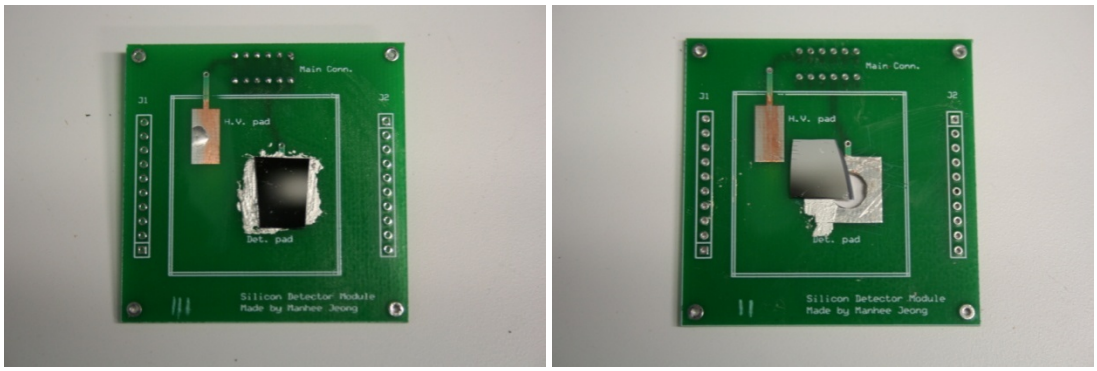


Fig. 3.5: Designed detector module PCB for wire bonding.

In summary therefore, electrical characterization measurements can indicate those factors, which lead to performance degradation, information that can be used to increase the uniformity of the junction fabrication. Those measurements were used to mitigate the electrical deterioration due to environmental interaction and readout connections, such that robust contacts were formed in Fig. 3.4.

3.1.2 Sources of Leakage Current

Fig. 3.6 schematically shows the dominant sources of leakage current in detectors, composed of I_{DR}, I_G, I_{INTER} and I_{SURF}. I_{DR} is the minority carrier diffusion current caused by the difference between the diffusion and the drift of the minority carriers. The E-field in a depletion region becomes stronger upon the application of reverse bias, which breaks the balance between the diffusion and drift of carriers. When minority carriers are drawn away by the E-field, the internal minority carriers come to replace them, the latter forming the diffusion current under reverse bias. $\mathbf{I_G}$ is the current due to the thermal generation of the electron-hole (e-h) pairs and is the bulk leakage current in a depletion region. This contribution increases with increases in the volume of the depletion region and can be decreased via cooling. For typical silicon devices, I_G is very low (recall that we modeled 10s of fA for ideal Pt Schottky diodes). $\mathbf{I_{INTER}}$ is the interface current at the Si/SiO_2 boundary, which depends on the interface states due to broken bonds, dangling bonds, impurities, and interface defects. The last source of leakage current is I_{SURF}, which is the outer surface current due to undesirables such as humidity and contaminations by fingerprints or other condensable vapors. I_{INTER} and I_{SURF} are passed through the active area if there is no guard ring present; however, a properly designed guard ring structure

can serve as a sink for these undesirable fluctuations in the charge number, as suggested in Fig. 3.7.

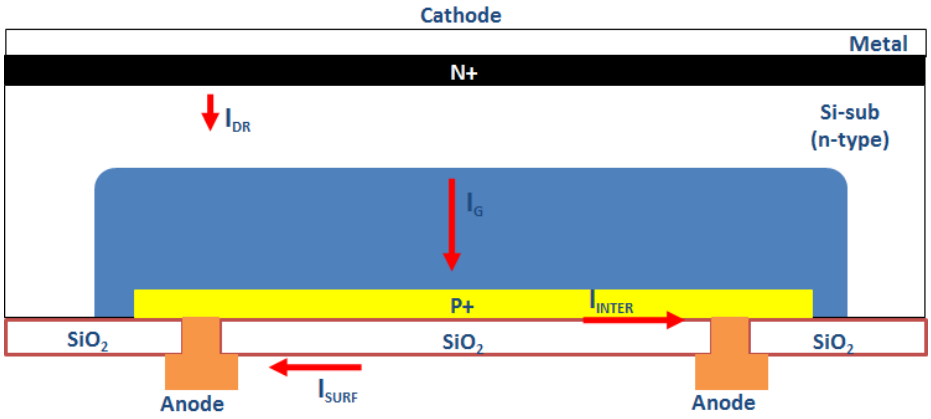


Fig. 3.6: Schematic diagram of contributions to leakage current.

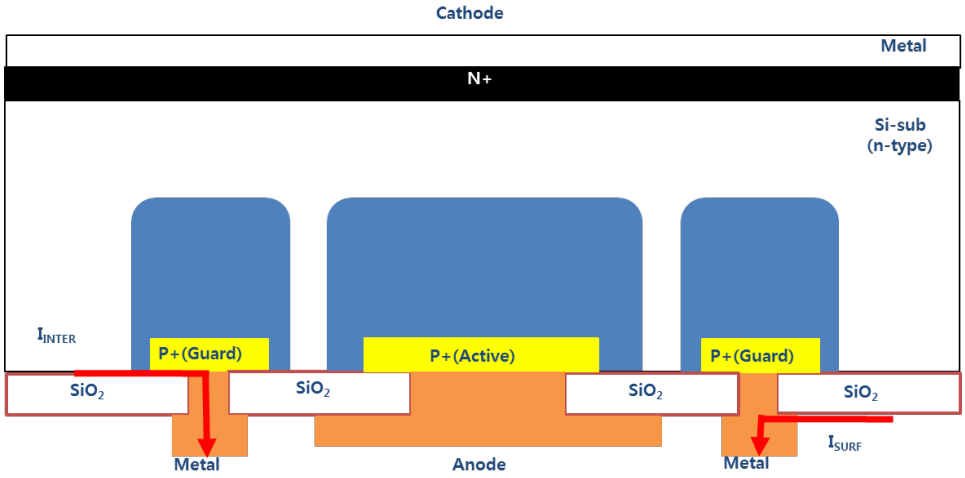


Fig. 3.7: The guard-ring effect demonstration.

3.1.3 Typical Mask Design

In order to simplify the coupling of the detectors into sockets, which can then be bonded to a PC board, we designed a mask for the second-generation devices, which has a detector active area of 8 x 8 mm^2. Across the face of the diagnostic mask shown in Fig. 3.8, are different size line widths (and gaps), a different number of guard rings (7 or 10), as well as a different multiplicity of lines. For instance, the lowest row has 30 meander-segments and the highest row can have up to 2000 turns. The “test pattern” shown in Fig. 3.8 is used to get the overall electrical characteristics of the detector.

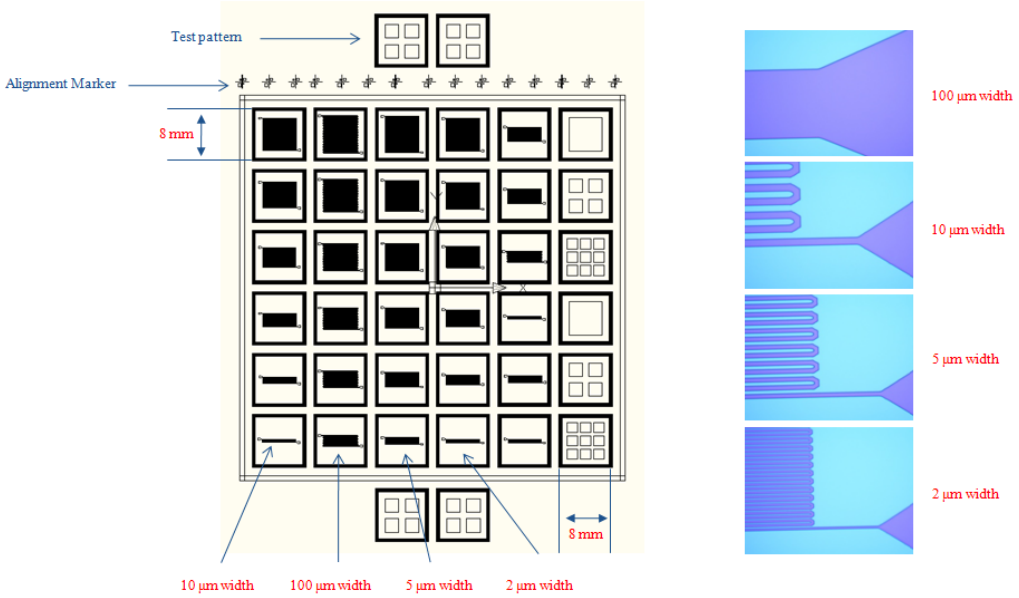
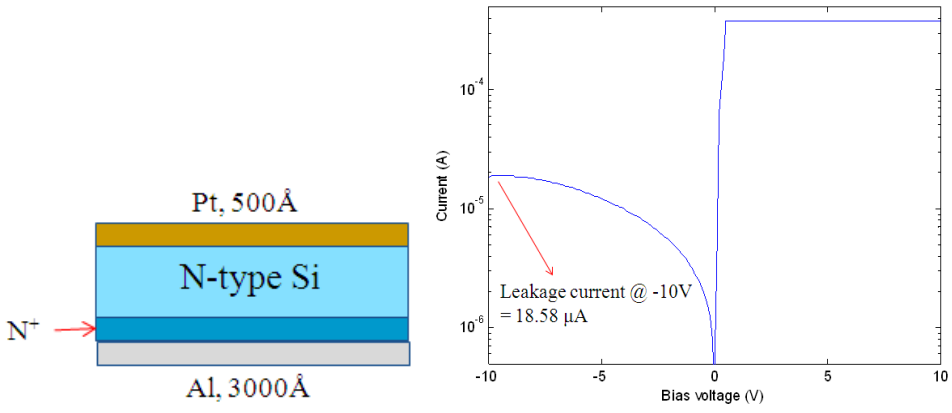


Fig. 3.8 (a) Designed mask which has a different size of line width from 2 μm to 100 μm and (b) their meander line samples after developing by photo-resistor.

3.1.4 N+ layer

In order to make a robust ohmic contact, we made an n+ contact, as shown in Fig. 3.9a with doping of phosphorous in a furnace with a temperature of 950 ^{o}C with N_2 gas. Planar detectors have no passivation region and guard-rings. These planar detectors have a leakage current of about 18 μA at -10 V bias voltage, as shown in Fig. 3.9b. Without the n+ contact, the leakage current at the ohmic contact portion is slightly increased due to the minority carrier motion. After making the n+ contact, however, the leakage current shows a good ohmic profile until 1V and then saturates.



(a) (b)

Fig. 3.9: Cross section of detector, which has an N^{+} (phosphorous) contact (a) its I-V curve (b).

3.1.5 Oxidation/passivation

Surface passivation prevents the deterioration of the electronic properties and reduces current-leakage paths of the semiconductor caused by moisture, ions, or other external contaminants. It also helps to prevent the recombination of e-h's at the edge of samples. With dry oxidation using a furnace with N_2 gas and a temperature of 1100 ^oC for 6 hours, we grew the SiO_2 layer with 2500 Å thickness. According to Fig. 3.10a, the oxidation reduces the surface leakage current by roughly 10 times to less than 1 μA.

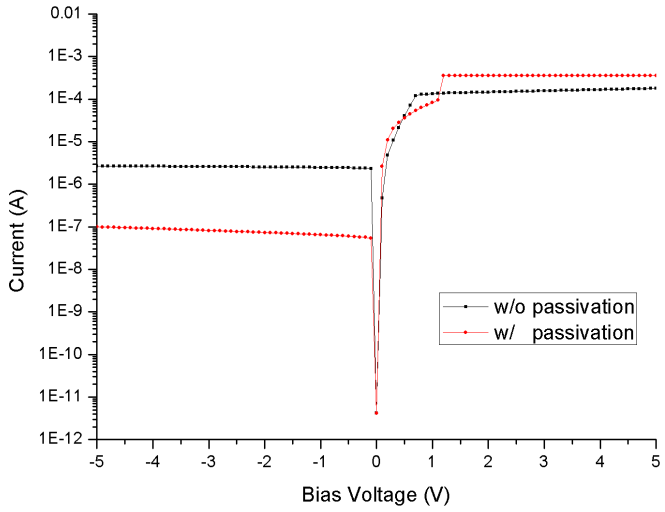


Fig. 3.10: I-V curve of comparison between surface passivated and non-passivated planar detector.

3.1.6 Schottky contact comparison between using the Evaporator and Sputter

Sputtering is a physical vapor deposition (PVD) process and was discovered in 1852 by Sir William Robert Grove. It was developed as a thin film deposition technique in the 1920s by Langmuir. The high-energy particles (Ar) strike a solid slab of high-purity target material (Au) and physically dislodge atoms. The advantage of sputtering is the ability to deposit and maintain complex alloys, the capability to deposit high-temperature and refractory metals, and the ability to deposit controlled, uniform films on large wafers (> 200 mm). It also has the capability of multichamber cluster tools to clean the wafer surface of contamination and native oxides before depositing metals (in situ sputter etch). Furthermore, it uses the sputter etch, which means that energetic argon and

reactant ions attracted to the surface due to wafer bias will bombard the surface and etch (dislodge) atoms. The sequence of sputtering is shown in Fig. 3.11.

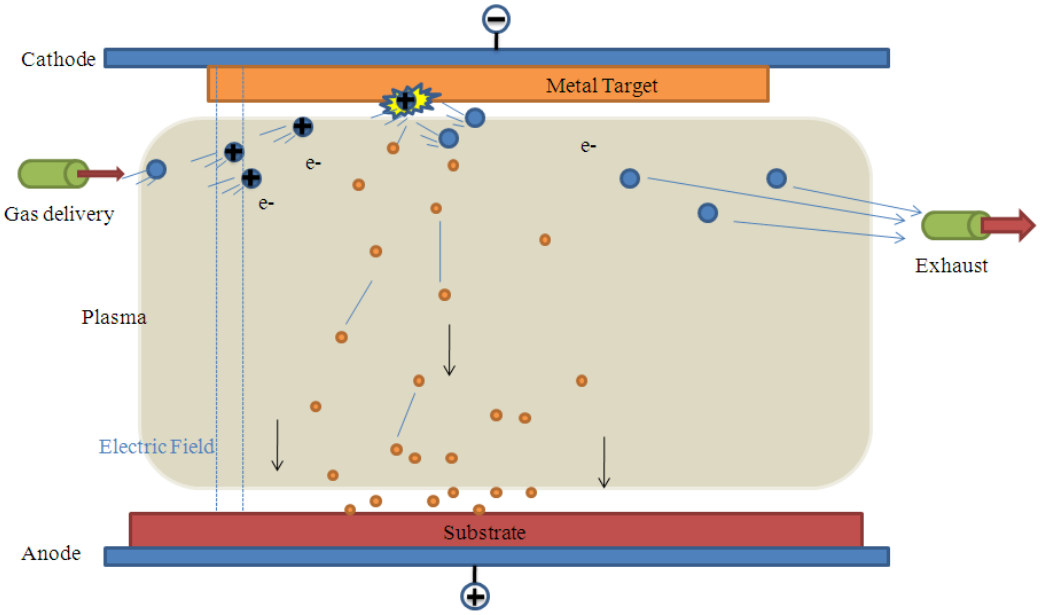


Fig. 3.11: Sputter process steps in the chamber.

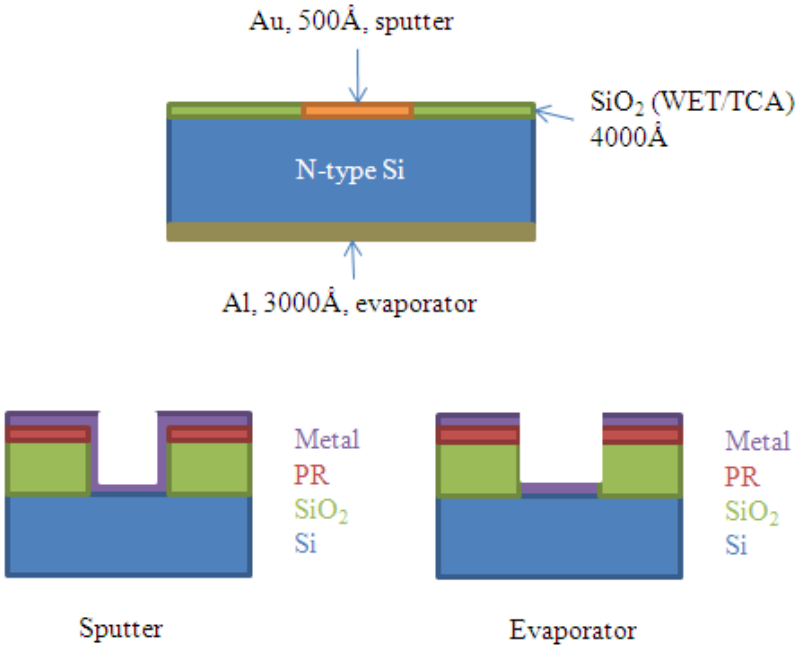
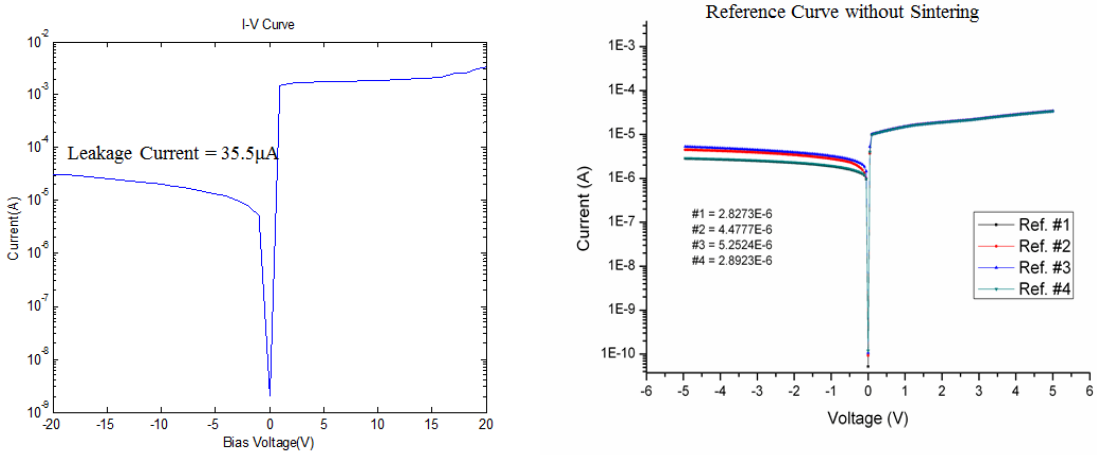


Fig. 3.12: The different method for deposition of metal between the sputter and evaporator.

At first, the electric fields create Ar^+ ions and the high-energy Ar^+ ions collide with a metal target. Then, the metallic atoms (Au) are dislodged (sputtered) from the target and metal atoms migrate toward the substrate (Si wafer). The excess matter is removed from the chamber by a vacuum pump. When using the sputterer, problems are associated in lifting off the deposited metal, as seen in Fig. 3.12.

In the evaporator case, the metal is deposited uniformly, making it easier to lift off the metal. However, in the sputter case, the metal is deposited isotropically, and therefore covers at the edge side of the photo-resistor (PR). As a result, we discovered the methods to easily lift off the metal, one being to harden the PR before the development. The other method is to use the LOA chemical to make the undercut at the photo-resist. Both methods facillate good lift off. There are 10's of differences in the leakage current between the evaporator and the sputter as shown in Fig. 3.13. The sputter etching helps to make a clean surface between the metal and bare silicon. Using the evaporator, the leakage current was 38 μA, but using the sputter, the leakage current of each sample distributed in 1 μA scale.



(a) (b)

Fig. 3.13: I-V curve with using the evaporator (a) and the sputter (b) with Au deposition.

3.1.7 Sintering with the forming gas

Tools we used for sintering/annealing are the rapid thermal process (RTP) and a furnace. The processes are similar, but there are differences in the composition of hydrogen. In the RTP case, the forming gas (N2:H2 = 95%:5%) with different time and temperature can be used. It is necessary to heat the interface to form a robust ohmic contact (low resistivity) between Al and Si and it fills the hydrogen into the weak dangling bond of Si to stabilize at the wafer surface. In the furnace case, the forming gas (N2:H2 = 90%:10%) can be used, but one cannot employ Pt since Pt is changed to Pt_2Si at temperatures above 200^{o}C. Fig. 3.14 shows the leakage current increasing due to the stress of the dicing, and Fig. 3.15 depicts the I-V curve comparison with increasing the temperature of RTP from100 ^{o}C to 350 ^{o}C.

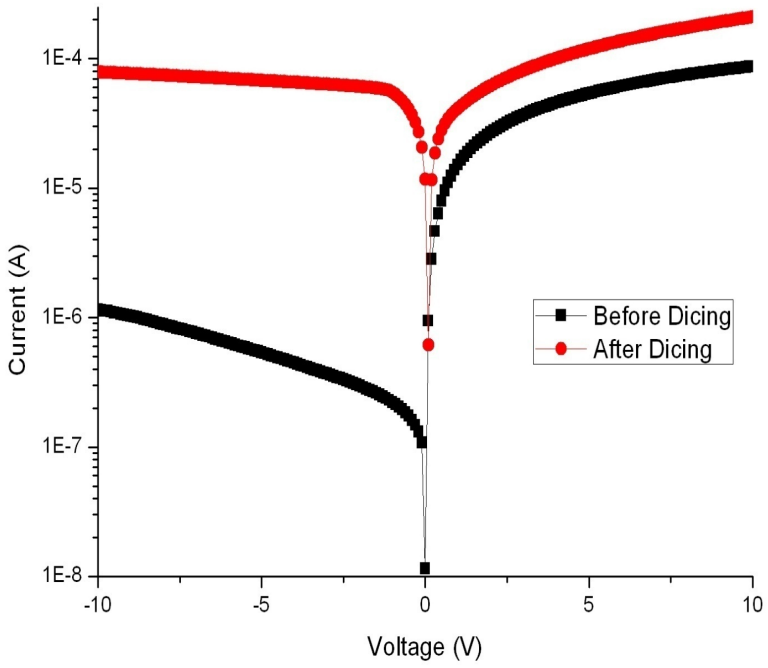


Fig. 3.14: I-V curve comparison before and after dicing the samples.

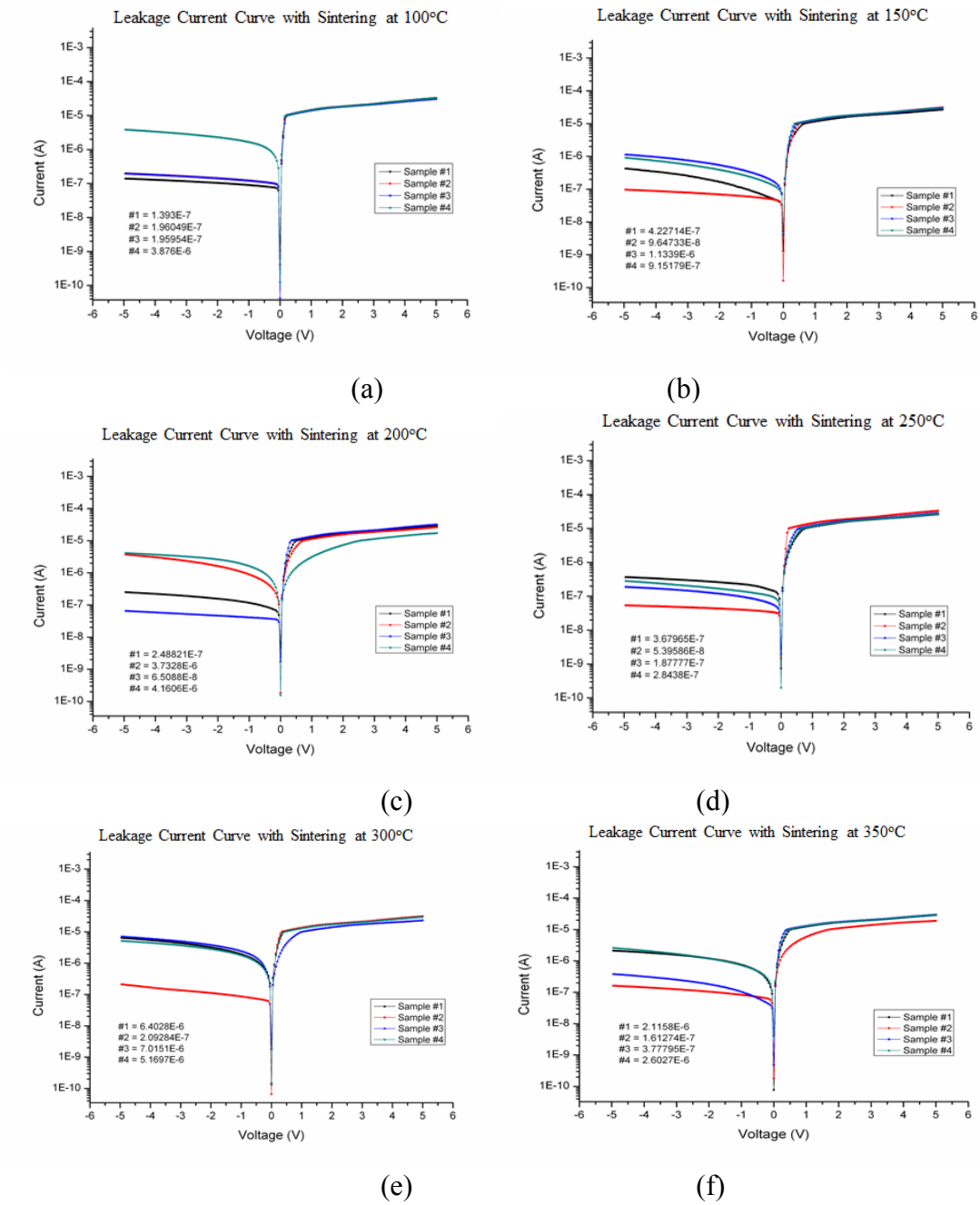
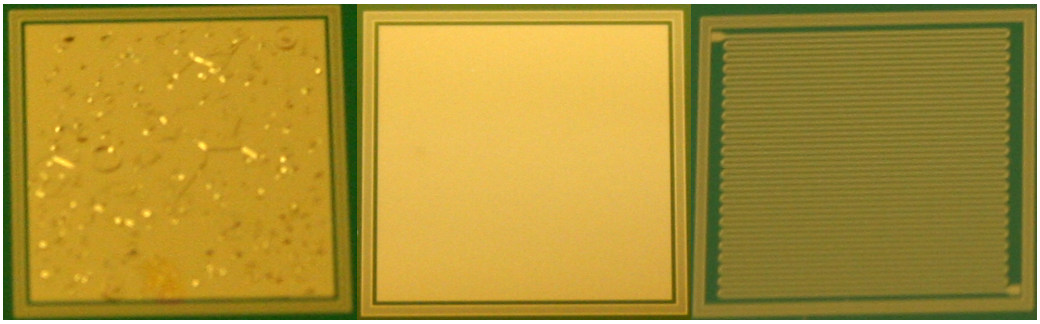


Fig. 3.15: I-V curve comparison with increasing the temperature of RTP from 100 ^{o}C (a) to 350 ^{o}C (f).

Furthermore, these sintering steps help to reduce the stress during the dicing as indicated in Fig. 3.16 indicate. During the dicing, stress on the edge side of the samples is formed, which increases the leakage current. However, the sintering/annealing helps to reduce this stress and with it, the leakage current.

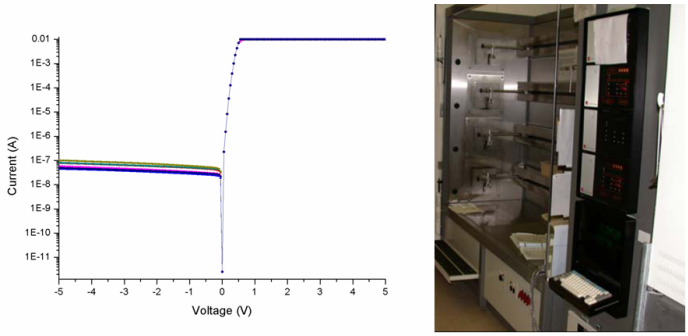
By increasing the temperature from 100 to 350 in 50 ^{o}C increments, sintering in the RTP was done for 10 min, as shown in Fig. 3.15. Without temperature increments, the leakage current was 10 times smaller. When the time with different temperatures is changed, increasing the time induces damage on the surface. Furthermore, in the case of the planar detector with an area of 1 cm^2, the surface stress is huge; it makes a bubble right after the lift-off step (Fig. 3.16a). However, after 30 min of sintering in the furnace at 200 ^{o}C, the surface stress is reduced and the bubbles are eliminated, as shown in Fig. 3.16b. By the way, the meander line detector has a small area and stress on the surface; it was the same both before and after the sintering as in Fig. 3.16c.



(a) (b) (c)

Fig. 3.16: Picture of the planar detector without curing of sintering (a) with curing of the sintering (b) and meander line detector (c).

Otherwise, when wafers are sintered in the furnace (Fig. 3.17b) the leakage current shows uniformity at each sample (Fig. 3.17a). Though the leakage currents are somewhat higher from 50 nA to 100 nA, the samples, having a low leakage current and the same properties, can be acquired.

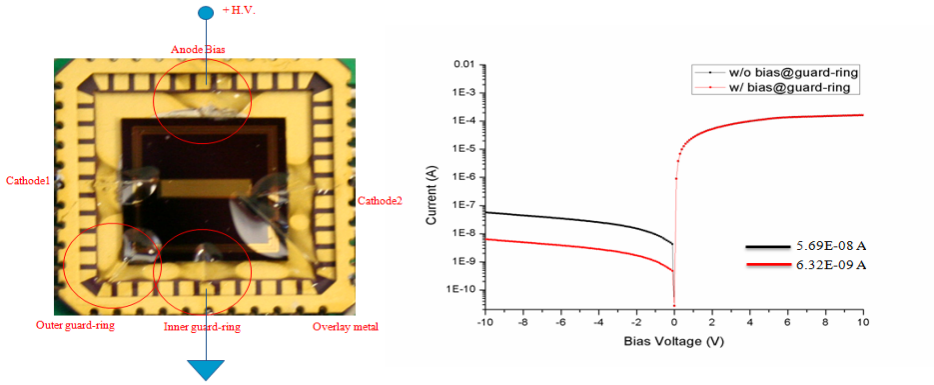


(a) (b)

Fig. 3.17: I-V curve of meander line detectors with sintering from A1 furnace with forming gas (10% H_2 and 90% N_2) (a) and picture of A1 furnace (b).

3.1.8 Guard Ring Biasing

Leakage current is a significant issue in detector systems, which cannot only add to the noise but also limit the extent to which we can bias the detector. To avoid overloading the high voltage source, leakage current reduction was further realized by designing and biasing the guard rings. For instance, Fig. 3.18a shows the improvement obtained just by biasing the inner guard ring at the same bias potential as the anode (top electrode). With improved passivation below the guard rings, further improvements are expected when we bias the outer guard ring at the same potential as the cathode (bottom electrode) bias.



(a) (b)

Fig. 3.18: (a) Picture of the packaged meander-line detector in the LCC socket, and (b) I-V curve comparison with and without bias at the inner guard-ring.

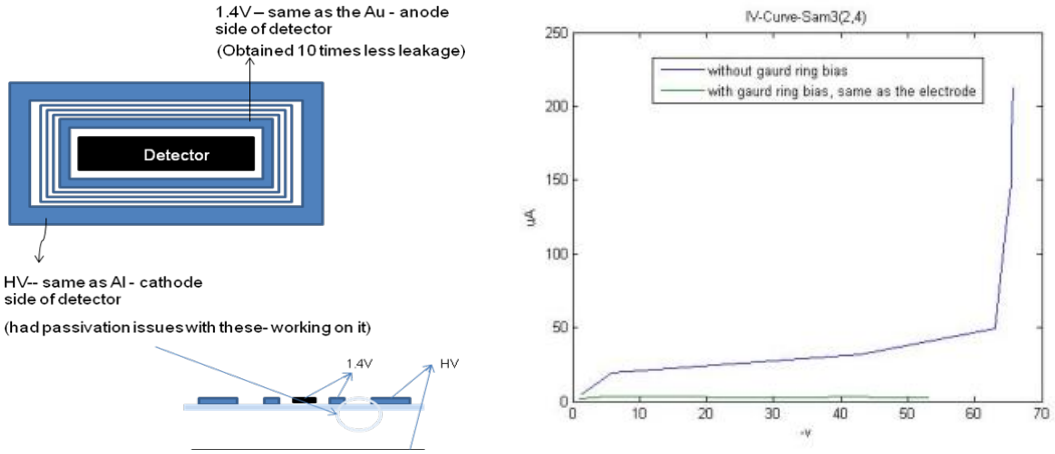


Fig. 3.19: Leakage current improvement with inner guard ring bias, as shown schematically on the left, resulting in the quenching of the reverse current bias (note reduction from blue curve to green curve).

Results similar to those shown in Fig. 3.19 showed that the guard ring was effective in reducing the leakage current to the desired nA range. For the results shown in Fig. 3.19, the inner guard ring is biased to that at the meander (0 V bias in this case) and the outer 6 or 9 guard rings were left floating. Fig. 3.18a shows the mounting of a meander detector with overlay (the gold rectangle in the middle) surrounded by the guard ring structure (the square on the outer edge of the detector), mounted in a socket for ease of mounting on a PC board. Fig. 3.18b compares the I-V curve derived with (red) and without (black) the guard ring biased, yielding a 6 nA leakage current in this case.

3.1.9 Overlay design and fabrication

For reducing the signal loss during its transmission through the meander line, the overlay detector was designed with a high dielectric material like SiO_2 or BST Barium-Strontium-Titanate (BST($Ba_{(x)}Sr_{(1-x)}TiO_3$)). Fig. 3.20 shows the processing steps to make an overlay detector. First, prepare the meander line detector, which has a low leakage current and make a photo-resist (PR) coating to prevent the damage on the passivated SiO_2 at the side of the meander line. Second, deposit 5 μm SiO_2 using the PECVD (Plasma enhanced chemical vapor deposition) at 200 ^oC. Third, with an e-beam evaporator, deposit the Cr/Au and then put a PR coating on it. Finally, remove the uninteresting regions with a wet etching and clean the PR. Fig. 3.21 shows the fabricated wafer with overlay detectors.

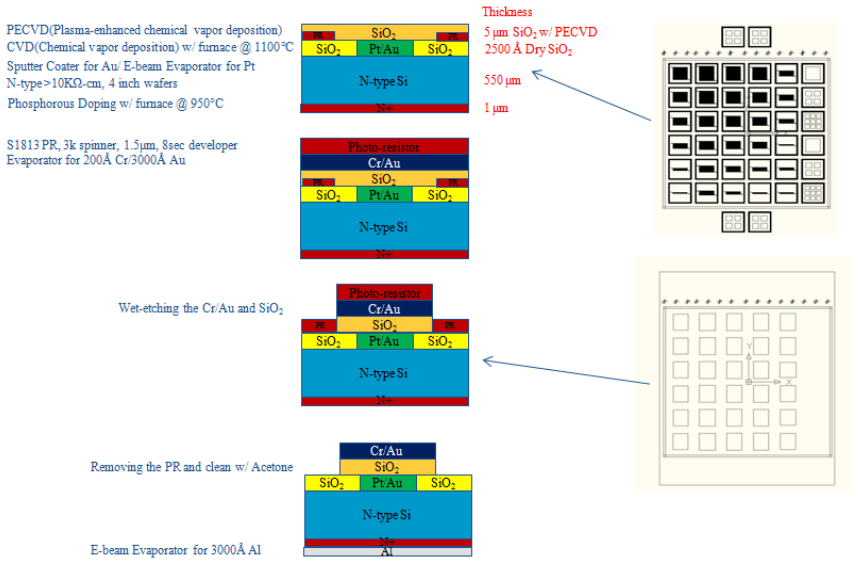
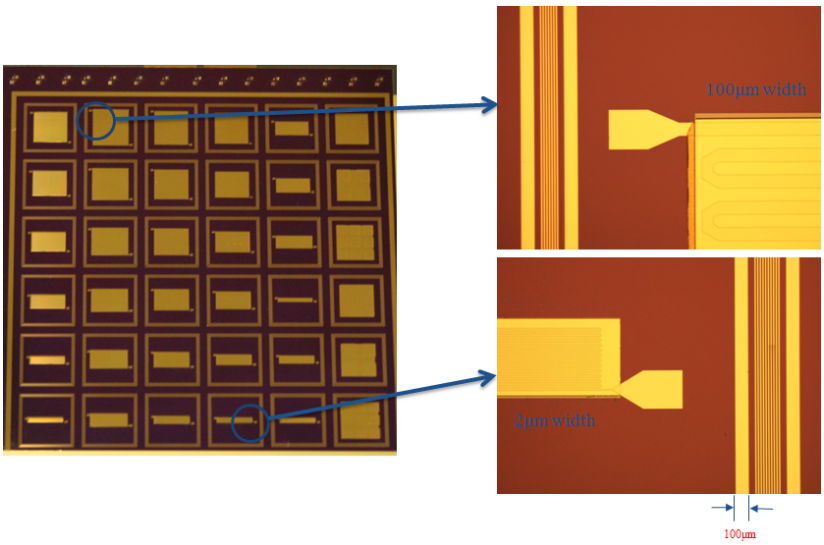


Fig. 3.20: Processing steps for overlay design.



(a) (b)

Fig. 3.21: Fabricated detector with new designed mask (a) and its zoomed in figures of 100 μm (up) and 2 μm (down) width meander lines (b).

3.2 Leakage Current Reduction Techniques

3.2.1 Gettering

Gettering means the transfer of unwanted impurities from the active area of the device to an unimportant region. The goal for this process is to maintain or enhance the starting minority carrier diffusion length. The gettering effect by phosphorus diffusion is used to improve the minority carrier diffusion length of polycrystalline silicon. In order to find and choose the proper material for gettering, I-V characteristics are compared with changing backside materials with 5000Å thick layers of low stress poly Si and amorphous Si (a-Si). In addition, I-V characteristics are also compared with and without 5000 Å thick passivation layers.

Fig. 3.22 shows the cross section of detectors and fabrication conditions. Fig. 3.23 shows the I-V characteristics with varying the materials for gettering. According to these results, lower leakage currents result when the a-Si and passivation layer are used. A-Si has a larger grain size than the low stress poly Si, so it collects minority carriers effectively and reduces the leakage current due to the drift of minority carriers. Furthermore, the passivation layer can block the leakage current path and mitigate the contamination via dust and humidity. Fig. 3.24 shows the summary of the I-V characteristics with different type's materials for gettering.

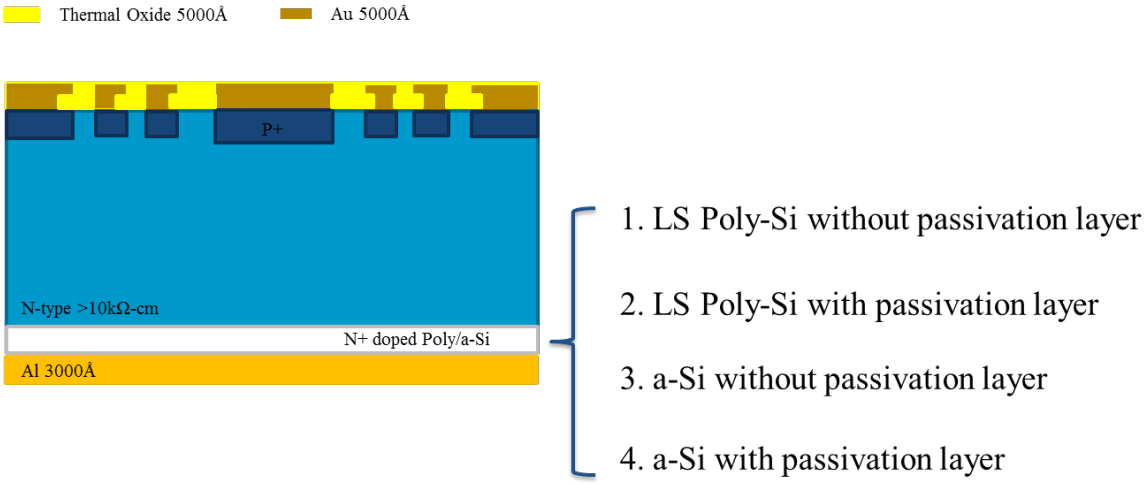


Fig. 3.22: Detector geometry with and without 5000 Å thermal oxide layer and different gettering materials (Low stress poly Si and a-Si).

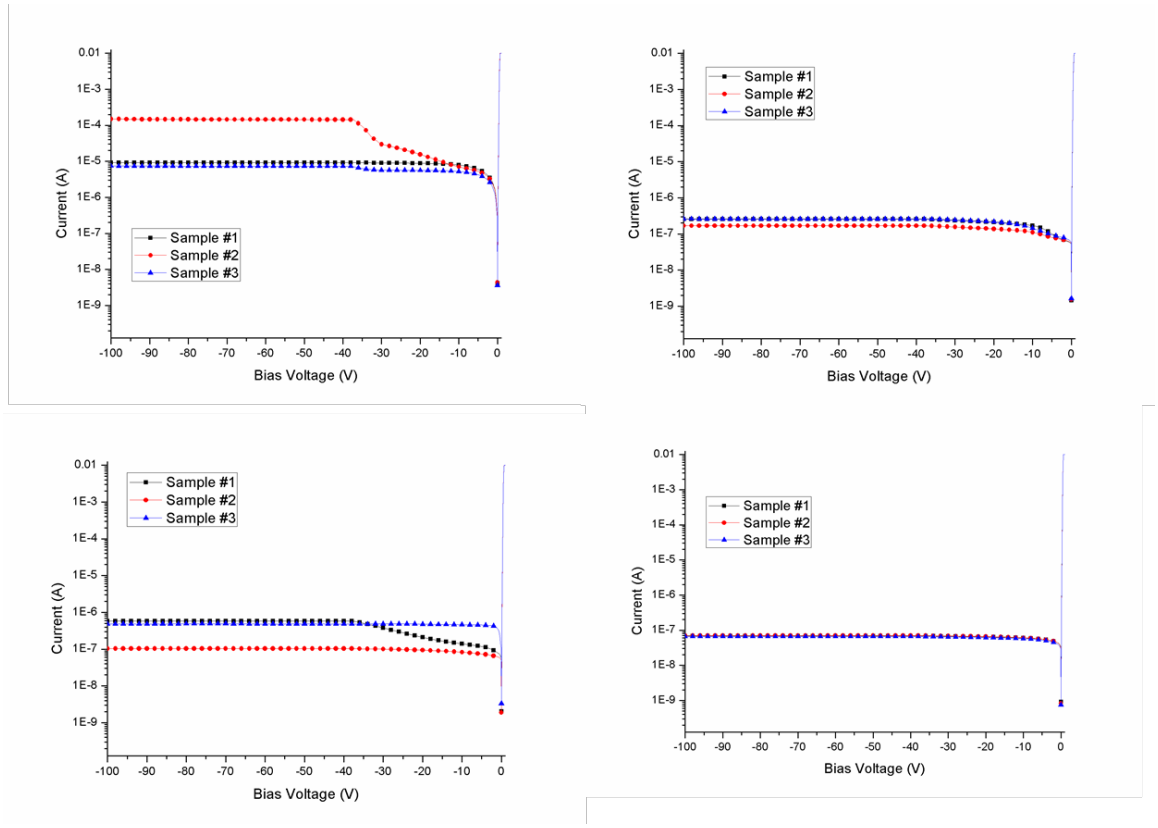


Fig. 3.23: I-V characteristics with different types of materials for gettering. Note that the left graphs are the no passivation layer with a LS poly Si (top) and a-Si (bottom) and the right graphs have a passivation layer and LS poly Si (top) and a-Si (bottom).

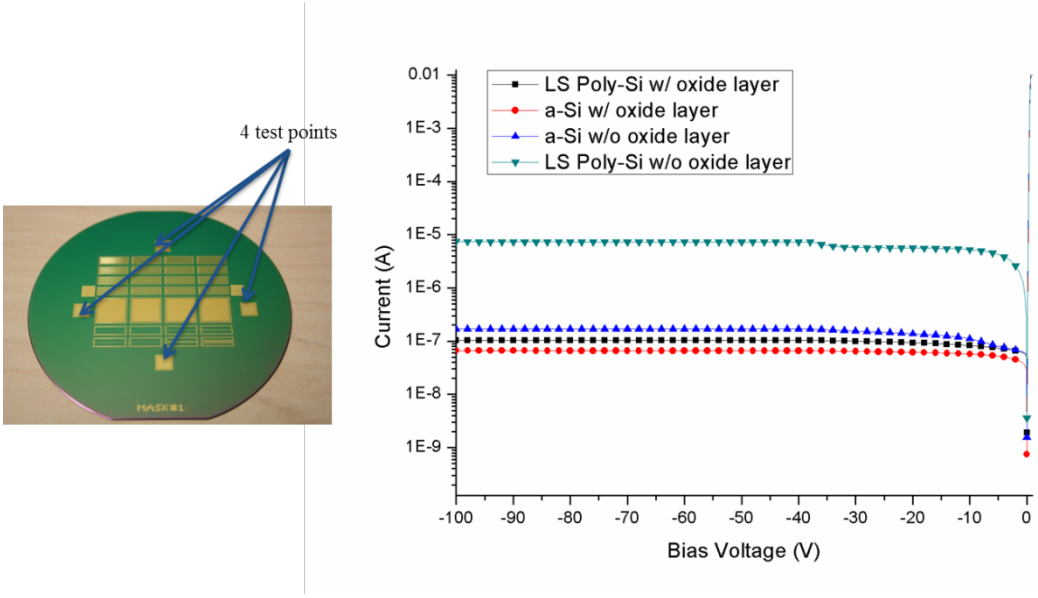


Fig. 3.24: I-V characteristics with different type of materials for gettering.

3.2.2 Surface passivation and annealing

Semiconductor detectors are in many cases used without protection against mechanical or chemical damage due to the environment where they have to operate. However, protection against chemical poisoning and environmental changes such as humidity is needed. Also, it needs mechanical protection of the soft platinum structures (meander line) during the sometimes elaborate mechanical assembly procedures such as wire bonding. In order to passivate the wafers, the wafer is covered with a robust electrically insulating layer.

Surface passivation prevents the deterioration of the electronic properties and reduces current-leakage paths of the semiconductor caused by moisture, ions, or other external contaminants. The hard SiO_2 layer will protect the silicon from scratches and processing damage, which might occur during backend final fabrication. Furthermore, passivation is also important for controlling the leakage current of junction devices and growing a stable gate oxide. But adequate oxide thickness is necessary to prevent electrical charging in a metal layer caused by a charge buildup in the silicon wafer surface, much like charge storage and breakdown characteristics of ordinary capacitors. This charging would lead to electrical shorting and other undesirable electrical effects. A thick layer of oxide that inhibits charge buildup from metal layers is the field oxide layer, which typically is between 2,500 Å to 15,000 Å thick. Silicon oxide also has a coefficient

of thermal expansion very similar to that of silicon. This expansion and contraction during the thermal excursions also prevents film stress from separating the oxide film from the silicon substrate.

Thermal oxide is grown by a chemical reaction between silicon and oxygen. Uniform oxide growth is achieved by exposing silicon to an elevated temperature in the presence of high-purity oxygen. The time and quality of this reaction varies, and is affected by the purity of the oxygen gas supplied to the silicon wafer surface and the reaction temperature. The typical temperature for the oxidation of the silicon during wafer fabrication is normally between 750 °C to 1100 °C.

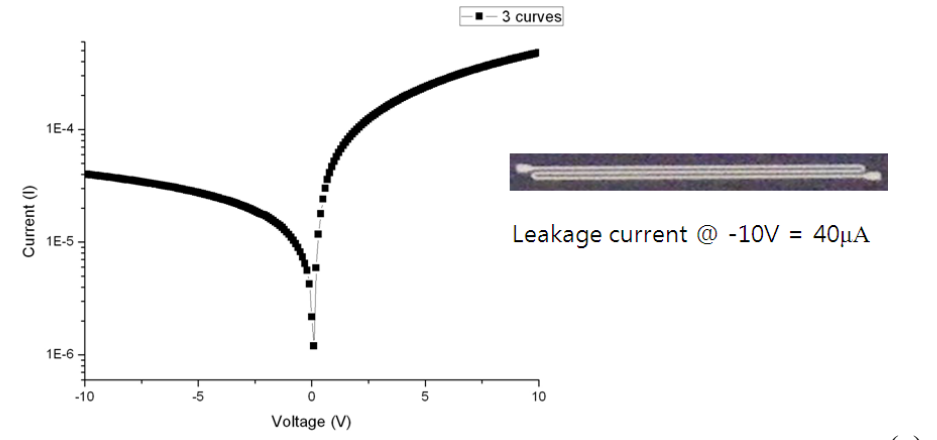
(a) (b) (c)

Fig.3.25: (a) Picture of the LPCVD at the University of Michigan, (b) passivated wafer with meander line pattern and (c) its cross section.

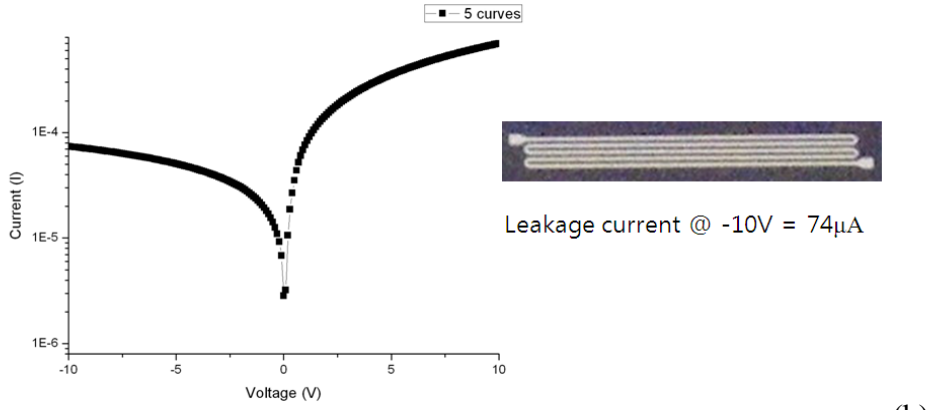
The Si meander line detectors were fabricated as follows: double side polished and (100) oriented Si substrates with n-type 10kohm-cm resistivity and 500 µm thick were oxidized at 400 °C in a dry O_2 ambient. The thickness of the thermal SiO_2 film was 1000 Å. A SiO_2 film was then deposited over the thermal SiO_2 using a high temperature, low-pressure chemical vapor deposition (LPCVD) equipment.

LPCVD is a chemical process used to produce high-purity, high-performance solid materials. The process is often used in the semiconductor industry to produce thin films. In a typical CVD process, the wafer is exposed to one or more volatile precursors, which react and/or decompose on the substrate surface to produce the desired deposit. The tool consists of a quartz vacuum process tube, surrounded by heating coils with three zone spike/profile temperature controls, a mechanical/wet vacuum pump, fore line with

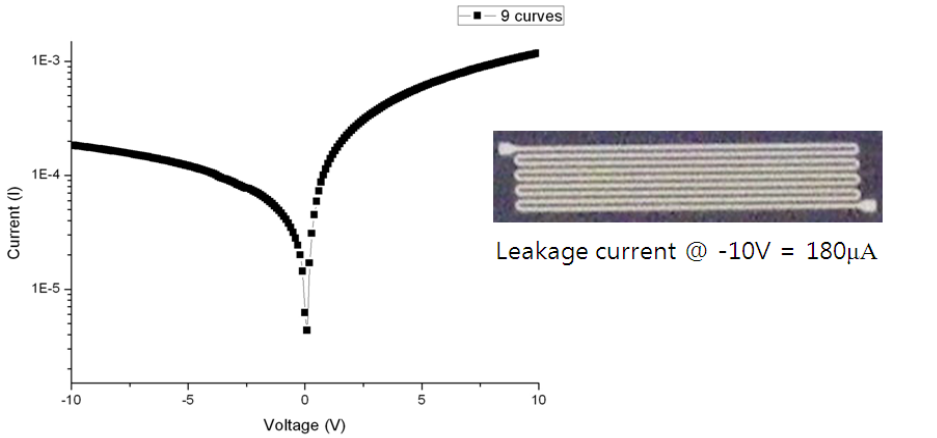
cold trap and multi-trap particulate filters, gas injection system, load and unload station, and MUX computer and tube control computers as shown in Fig. 3.25a.



(a)



(b)



(c)

Fig. 3.26: I-V curves for (a) 3 meander, (b) 5 meander and (c) 9 meander detector.

Before using the LPCVD, wafers should be cleaned via RCA/HF clean or Pre-Furnace clean. The purpose of the RCA clean is to remove organic contaminants (such as dust particles, grease or silica gel) from the wafer surface; then remove any oxide layer that may have built up; and finally, remove any ionic or heavy metal contaminants. This step is performed for CMOS compatible processes as discussed in APPENDIX B.

After the RCA clean, wafers are loaded at LPCVD with setup as 400 °C and 1000 Å thickness. The cross section is at Fig.3.25c. Then, we coated with a photo-resist of 1.5 μm thickness and made a pattern using a mask aligner on top of the photo-resist (PR). After making the pattern, we etched SiO_2 under the pattern area using a buffered HF (BHF) solution for 30 sec, which yields a passivated wafer, such as that shown in Fig. 3.25b. Upon dicing this wafer, we tested the I-V characteristics via the Keithley 4200-SCS (Semiconductor Characterization System). As shown in Fig. 3.26 and Fig. 3.27, the I-V curves of the meander line detectors were measured. An increasing number of meander turns results in higher leakage (Fig. 3.27), and the leakage current therefore scales with the size of the active area. In this case, the 4-inch wafer with 10 kΩ-cm resistivity was used, and the PIRAHNA clean and dip HF were done for removing the organic particles and dirty material. However, a BHF clean was not included, so the device possessed a SiO_2 layer, which is naturally grown at room temperature. Thus, this wafer has a bigger leakage current than that found in previous wafers.

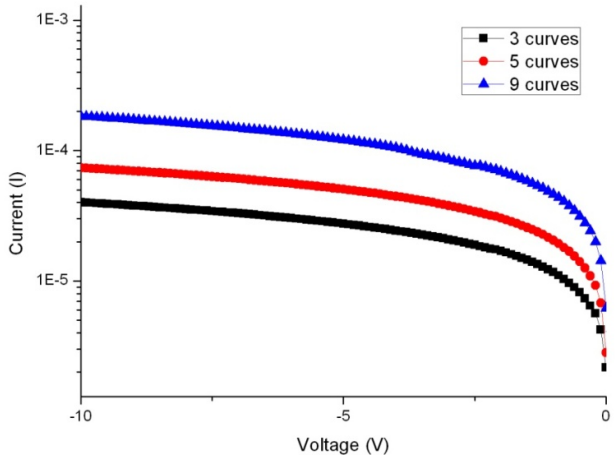


Fig. 3.27: Graph of the leakage current vs. reverse bias for different samples with 3, 5 and 9 meander.

If one wants to reduce the leakage current with passivation, one might do a BHF clean before the process. However, a robust meander line detector can be acquired through this passivation step, which can protect the wafer from humidity and mechanical weakness.

3.2.3 Metal Field Plates (MFP) and Multiple Guard-rings

Recall that metal-semiconductor contacts to interface with the detector were fabricated. However, one can use the metal-oxide-semiconductor (MOS) structure (a so-called "field plate" in the following) to enhance the noise design. For instance, if the MOS structure is operated in accumulation, the structure acts as a sink for the surface current components discussed in the beginning of Section 3.1. Furthermore, the field plate can be used to shape the electric field in order to reduce high-field regions and increase regions where soft-breakdown occurs.

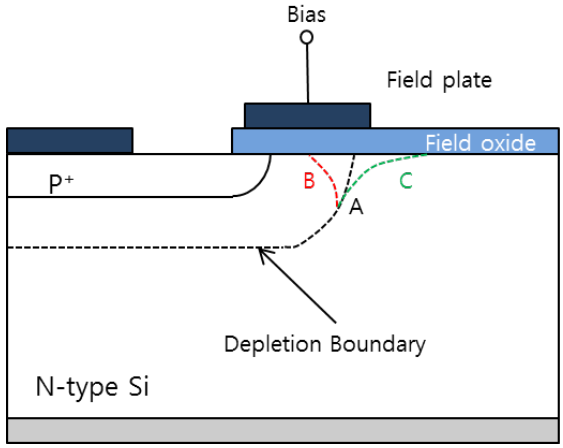


Fig. 3.28: The demonstration that the depletion boundary changes with varying the bias on the field plate.

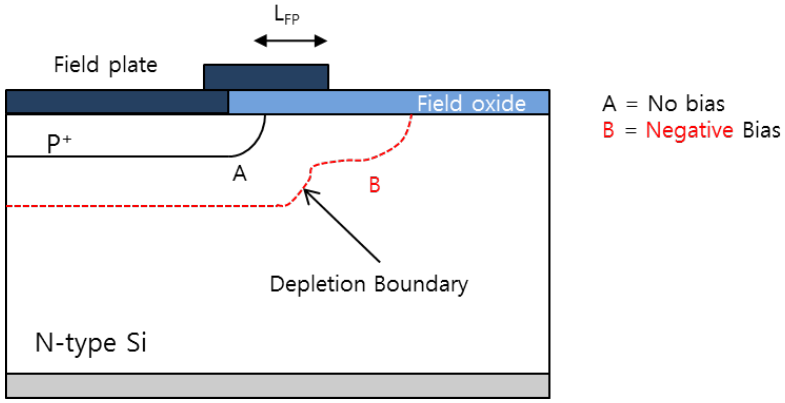


Fig. 3.29: Demonstration of the depletion boundary changing with varying the bias on an overhanged metal field plate.

As implied in Fig. 3.28, by varying the polarity of the bias voltage at a field plate the depletion boundary will change. For case B for instance, electrons are attracted to the surface and the electric field is concentrated, thus reducing the breakdown voltage. For case C, the bias repels electrons away from the surface and expands the E-field, and thereby increases the breakdown voltage. However, one does not need the additional complication of interdigitating MOS structures with depletion structures (MS or PN diodes) in order to accrue these benefits. One can use an over-hanging metal field plate as shown in Fig. 3.29, in which case B again shows the case in which the breakdown voltage is increased via the broadening of the electric field.

The leakage current benefits of guard-rings were discussed in Section 3.1.8, which includes sinking the surface leakage and isolating the active region. However, guard-rings can also improve the breakdown voltage of the detector by removing the high electric field regions. For instance, Fig. 3.30 shows the depletion region boundary being broadened by the presence of multiple floating guard rings. As the bias voltage is increased further, the depletion zone will continue to spread from the second guard ring, and the process will repeat for other guard-rings. The potential difference between any two guard-rings depends on the gap size, the doping density, the oxide charge, the field plate design, and the bias voltage, with more guard rings providing smaller abrupt jumps in the voltage. Fig. 3.31 is a schematic of the third generation of detector design considered in the research, in which the leakage current is controlled via multiple floating guard-rings and metal field plate structures. The inward metal field plates can improve the leakage characteristics, via the accumulation of electrons at the Si-SiO_2 interface due to positive oxide charge, but it degrades the breakdown characteristics. The outward metal field plates can improve the breakdown since it can disperse the electric field at the right edge of the delay line, where the electric field is strong, and keep the leakage current constant. The fabrication of the structure is discussed in Section 3.5, and the two transmission-line enhancements- the thick high-conductivity Ag electrode and the high permittivity (BST) dielectric overlay- are motivated in Chapter II.

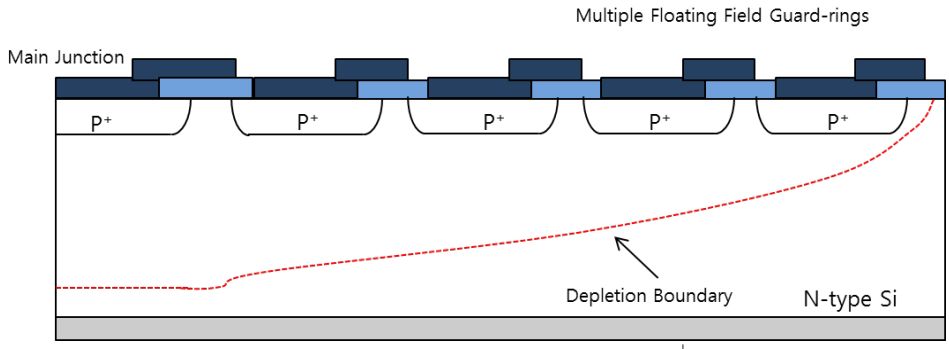


Fig. 3.30: Changing the depletion boundary due to floating P^+ field guard-rings.

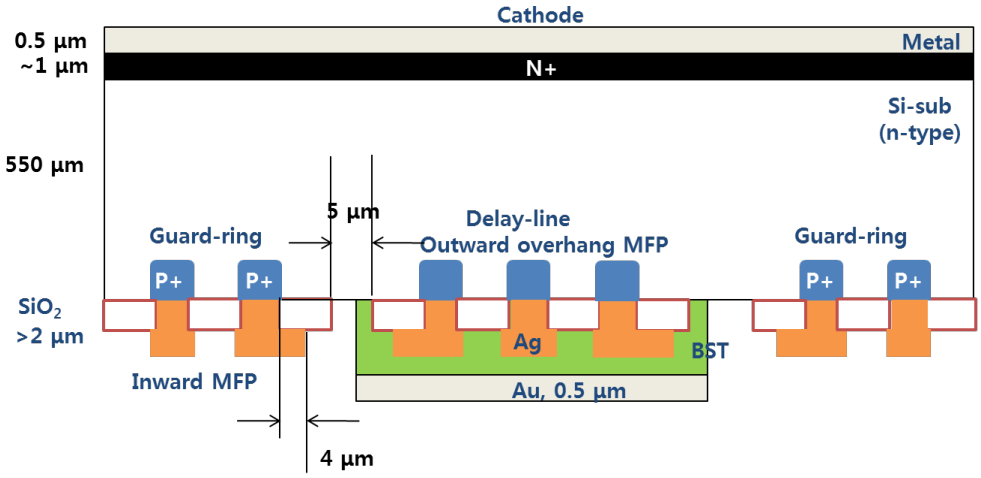


Fig. 3.31: Cross section of a 3rd generation designed meandering line detector with metal field plate and overlay.

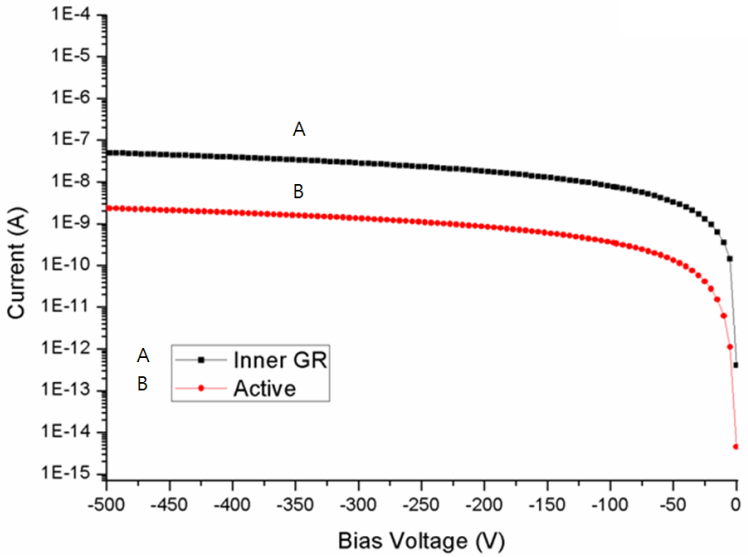


Fig. 3.32: Comparison between the I-V characteristics of the inner guard-ring (Inner GR) and the active diode.

Beyond moving the high electric field away from the active region, one also endeavors to minimize its amplitude regardless of its position, in order to reduce the probability of soft or hard breakdown [8]. This reduction can be elicited by increasing the width of the metal field plate. As the MFP extent is increased from 3 μm to 5 μm, the magnitude of the maximum field decreases by more than a factor of 2, which implies that one can increase the applied detector bias commensurately. Although the surface leakage current may be the main source of noise, it can be collected by the inner guard ring, as shown in Fig. 3.32, in which 50 nA of leakage current is collected by the inner guard-ring and a current of only 1 nA current makes it to the detector. In selecting an optimum guard-ring design, we also simulated the effect of altering the spacing between the structures. As shown in Fig. 3.33, the leakage current is reduced, approximately linearly with gap distance, as it is decreased from 40 μm to 10 μm, because of a more effective extension of the depletion region in the inter-gap region. The simulated results from Chapter II were buttressed with the experimental measurements, as follows. First, the breakdown voltage is increased with the addition of the MFP's.

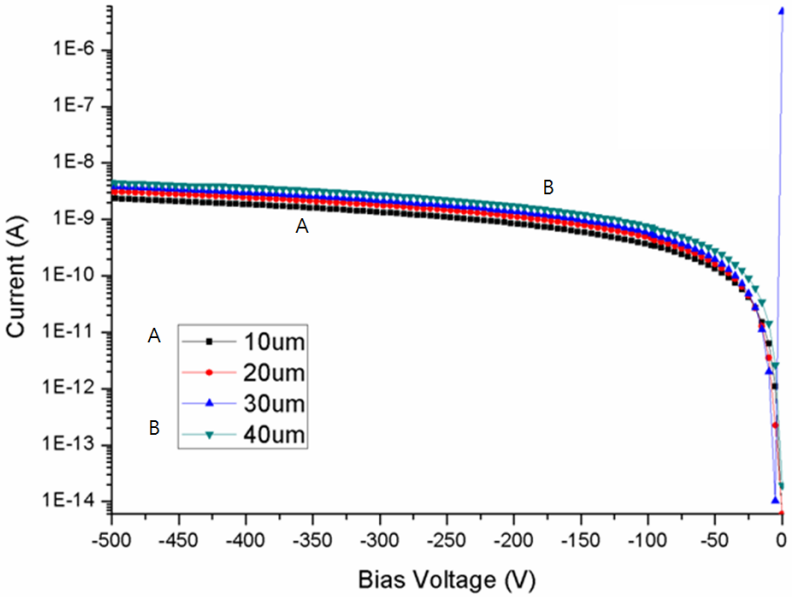
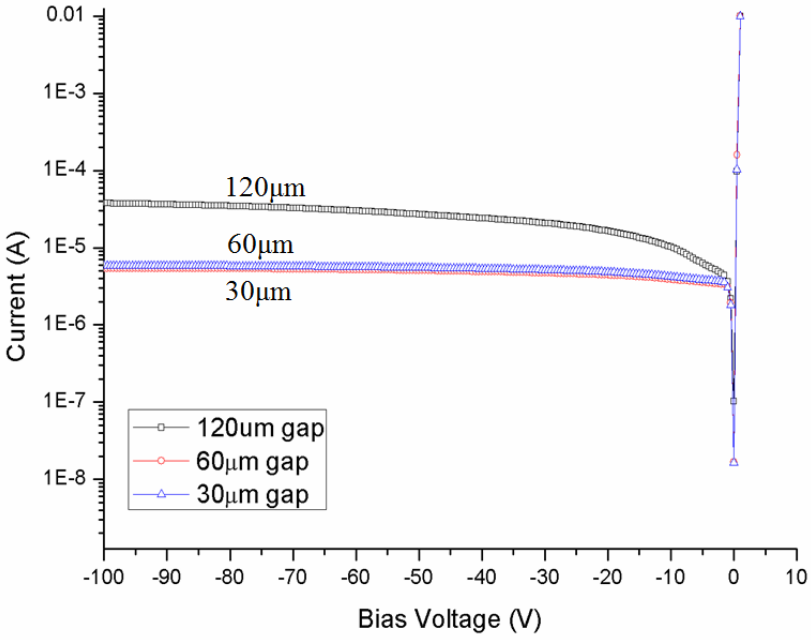
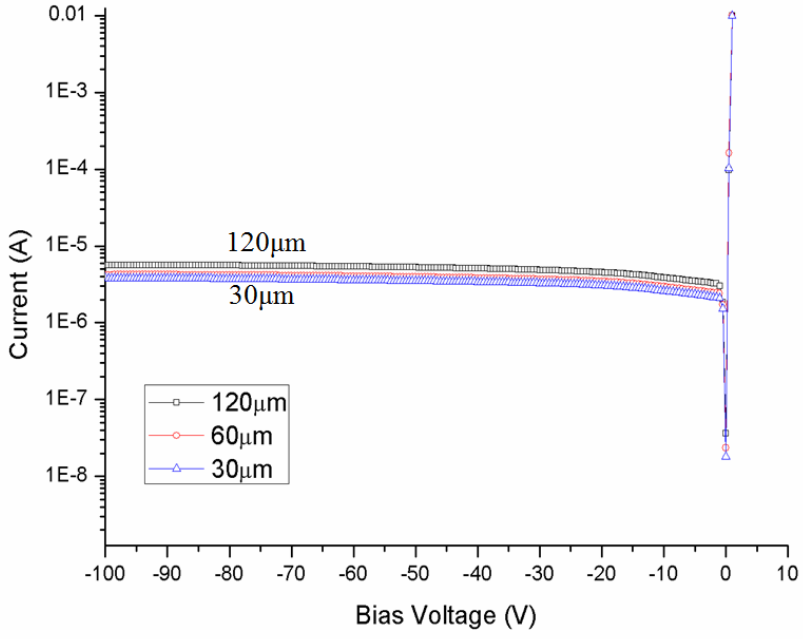


Fig. 3.33: Comparison between the I-V characteristics upon changing the gap size between the inner guard ring and the diode.



(a)



(b)

Fig. 3.34: Graph of the I-V comparison of detectors with different gap sizes between the active region and inner guard-ring (30, 60 and 120 µm, as indicated in the legend), when the gap between floating guard-rings is (a) 120 µm and (b) 60 µm.

Also validated in this research was the prediction that decreasing the gap between the active region and the inner guard ring decreases the leakage current. Specifically, Fig. 3.34 shows the I-V curves collected from detectors with different gap sizes, when the gap between the floating guard-rings is either 120 μm or 60 μm. Regardless of the inter-guard ring gap, the small gap decreases the leakage current with a similar moderate variation to that indicated in the simulations. Furthermore, the metal field plate structures successfully prevented soft breakdown, which is present in their absence.

3.3 The First Generation Detector Design and Fabrication

The first task of the first phase of the research was to understand the contact physics so that *robust* electrical contacts could be made to the detector. During those investigations, we started with p-type high-resistivity silicon, arranged as shown in Fig. 3.35a, and we produced a very low leakage-current design, as exemplified by the current-voltage (IV) characteristic of Fig. 3.35b. The resulting devices were sharply sensitive to gamma-ray interactions, as demonstrated with the 662 keV gamma-ray induced pulses shown in Fig. 3.35c.

Despite this sensitivity, the p-type devices lacked our most valued characteristic at the time, namely, a device with reproducible electrical characteristics. Recall that with the absence of extensive characterization runs, the devices were rectifying in only 30% of the cases. We therefore transitioned to n-type devices, which the calculations and models indicated would produce a more stable diffusion barrier, an expectation realized as 100 % of the platinum or gold-bounded devices were not only rectifying, but also proved to be suitable detectors for alpha particles. We therefore employed a stable Al-<n-Si>-Pt structure over the first half of the work, in order to develop the fine position-sensing methodology. However, the detector must again be made sensitive to small current and charge fluctuations, a goal compromised by the fluctuations in the high leakage currents (~1 μA to ~10 μA) that result from the n-type devices. Therefore, the practical task of optimizing the detector from the standpoint of the semiconductor physics in which the band-structure and electric field environment in the detector bulk is controlled, was done by exercising control over the process parameters, the geometry, and the electrical coupling to the readout circuit. In this section, some of the refinements to the simple

metal-semiconductor geometry, required to ensure robust behavior from the devices, are summarized.

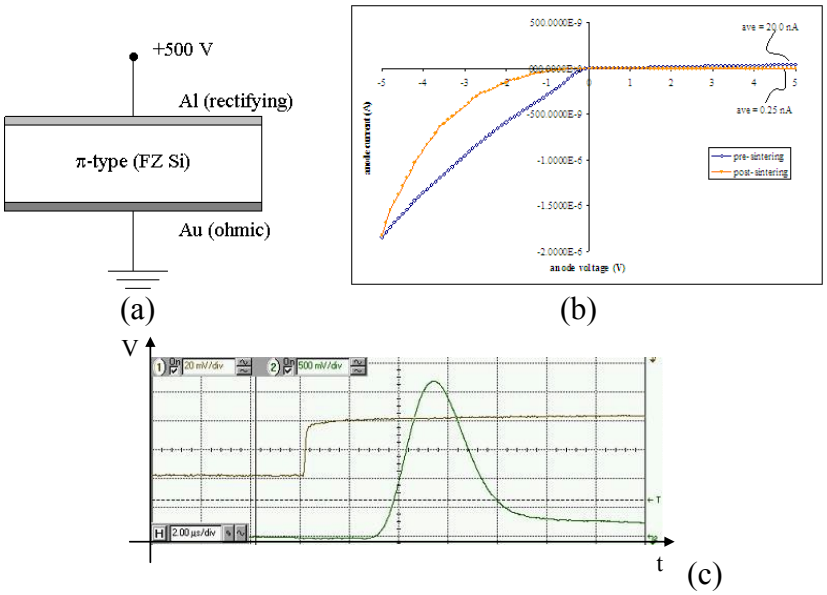


Fig. 3.35: (a) Detector configuration for p-type silicon with metal-semiconductor rectifying and ohmic contacts. (b) An example of IV characteristic, derived from a 1 x 1 cm planar detector, exhibiting nA scale leakage currents. (c) The resulting voltage signal derived from the depleted device, as measured with an Ortec 142A charge-sensitive preamplifier (in black) and a shaping amplifier (in green).

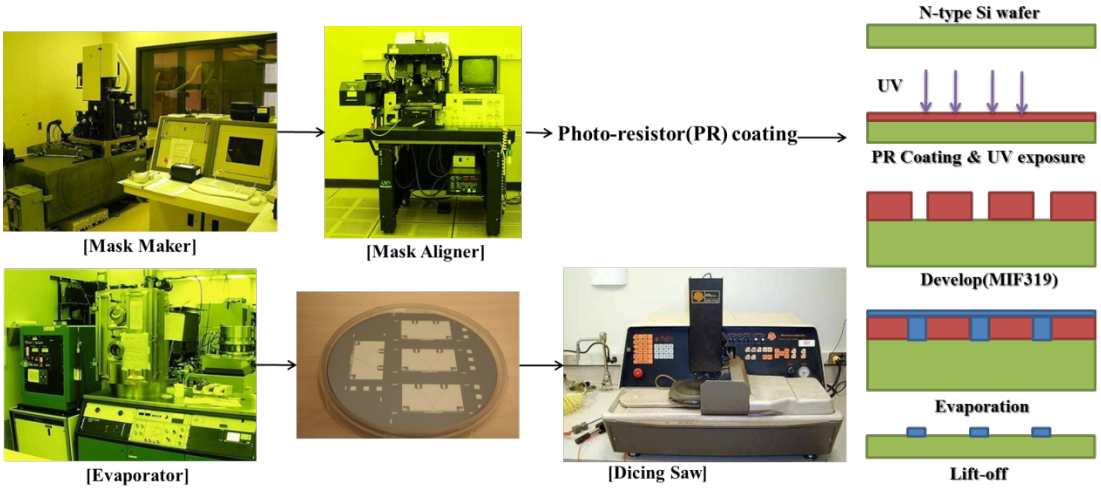
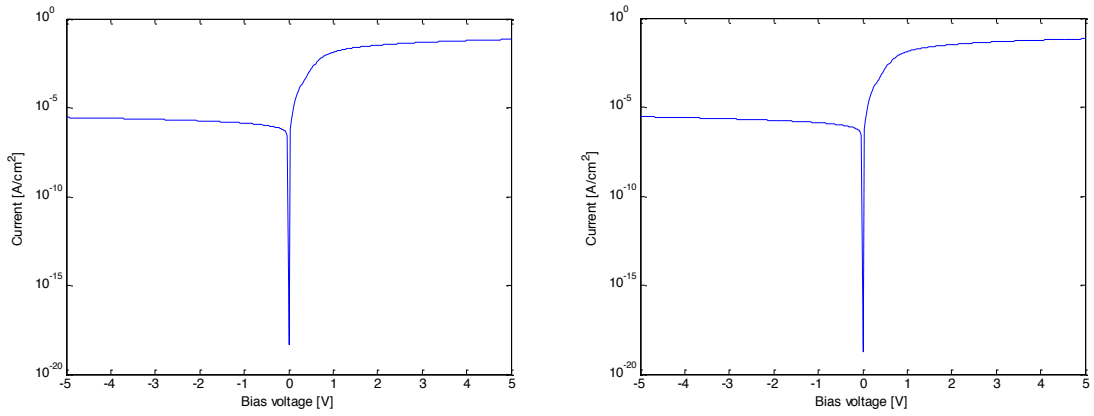


Fig. 3.36: (a) Wafer fabrication steps and equipment and (b) process profiles.

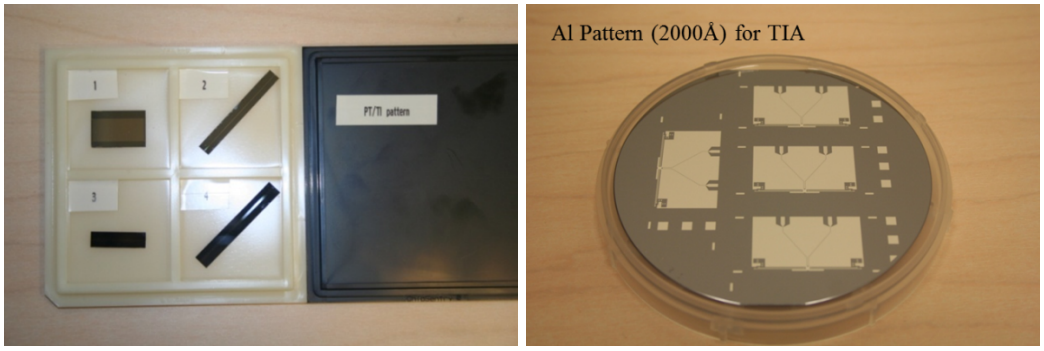
Regardless of the readout technique, the device fabrication methods are consistent, as outlined in Fig. 3.36. If one intends to measure the position of charge collection using amplitude-based techniques (cf. Chapter II), then one design the discontinuities on the ends of the meander line to attenuate the pulse amplitude need be designed. For time-based measurements, the compensation of the geometric changes to produce a balanced line is paramount. The main alteration is therefore in the original layout, drawn with different computer-aided design packages (AutoCAD and Mentor Graphics). However, the subsequent process steps do not change.

The first step is to transfer the computer-generated layout designs to chrome-patterned (glass) mask plates for use in the photolithographic processes. The bare high-resistivity (> 15 kΩ-cm) silicon wafers are cleaned of organic materials and debris with a PIRANHA clean (1:1 Hydrogen Peroxide (H_2O_2) and Sulfuric Acid (H_2SO_4) for 15 min). Before the metallic evaporation, the native oxide layer is removed with HF. The wafers are then patterned using standard photolithographic processes. In my case, photoresist (PR) 1813 (or 1827) is spun at 4 krpm for 30 sec resulting in photoresist thicknesses of 1.5 μm for 1813 and 3.5μm for 1827. The PR coated wafers are then UV exposed through the mask with the MA-6 mask aligner- for 6 sec for PR1813 and 13 sec for PR1827. The PR is then developed with MF319 for 1 min for 1813 and 1.5 min for 1827. An Enerjet evaporator is then used to deposit electrode metals Au, Pt or Al, which are subsequently lifted-off with 1112A for 1 day. The completed wafer is then diced with the help of a dicing saw, resulting in a completed device like those shown in Fig. 3.36.



(a) (b)

Fig. 3.37: Simulated I-V curves when the cathode is: (a) Pt and (b) Au.



(a) (b)

Fig. 3.38: Pictures of meander detector samples with a Pt/Ti rectifying contact (a) and pad for transimpedance amplifier (TIA) with Al contact (b).

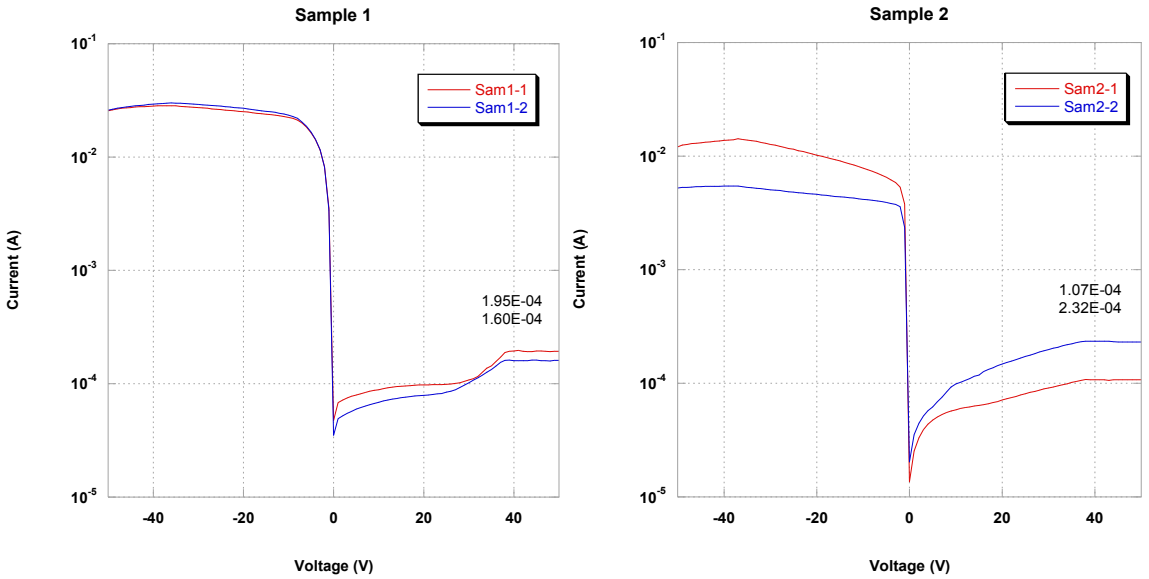
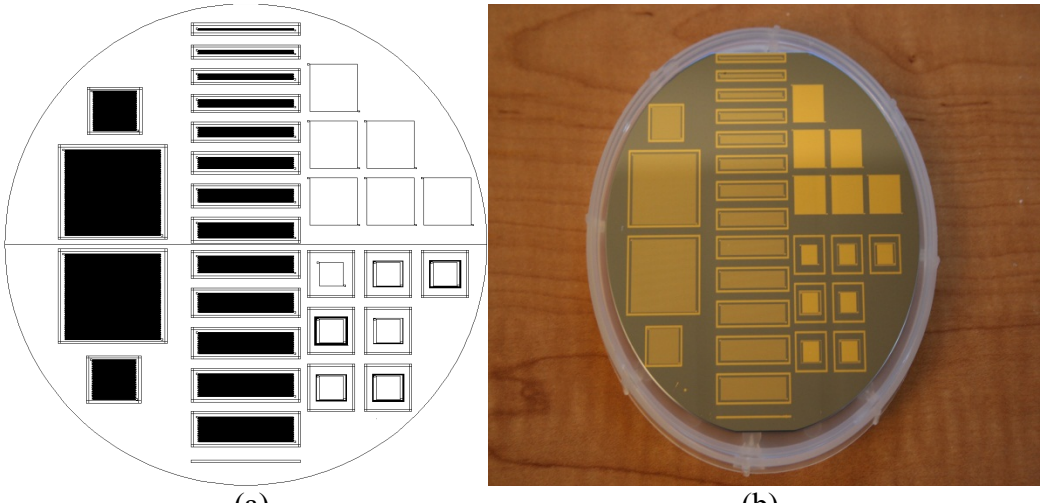


Fig. 3.39: Graphs of I-V characteristics for meander samples with a Pt/Ti patterned meander detector.

In order to validate the measured I-V characteristic data, the Sentaurus device simulation software was used, results of which are shown in Fig. 3.37. Fig. 3.37a shows the simulated I-V curve when the cathode is Pt (500 Å) and anode is Al (3000 Å). It has a leakage current of about 2.8 μA/cm^2, which is consistent with the value we measured from our 10 kΩ–cm batch of silicon wafers. However, about 10 times as much leakage current was measured for devices based on higher resistivity material (15 kΩ–cm). The detectors are shown in Fig. 3.38 and the resulting IV curves are shown in Fig. 3.39.

After seven microelectronic process iterations during a six-month period, the first-generation meander detector fabrication was made robust *for the Al-Si-Pt system*, which can be used to fully develop the charge tracking system. That is, detectors can be fabricated with reproducible behavior and patterned with partitioned electrodes on the aluminum side so that the delay-line design can be optimized, from which the lateral position-sensing can be practiced in one dimension. However, recall that measuring both lateral dimensions requires the fabrication of an orthogonal meander line on the backside of the detector. This, unfortunately, cannot be accomplished with platinum because of its poor adhesion to silicon. Therefore, some effort on optimizing the fabrication of a patternable rectifying contact was needed, as explained in the following.

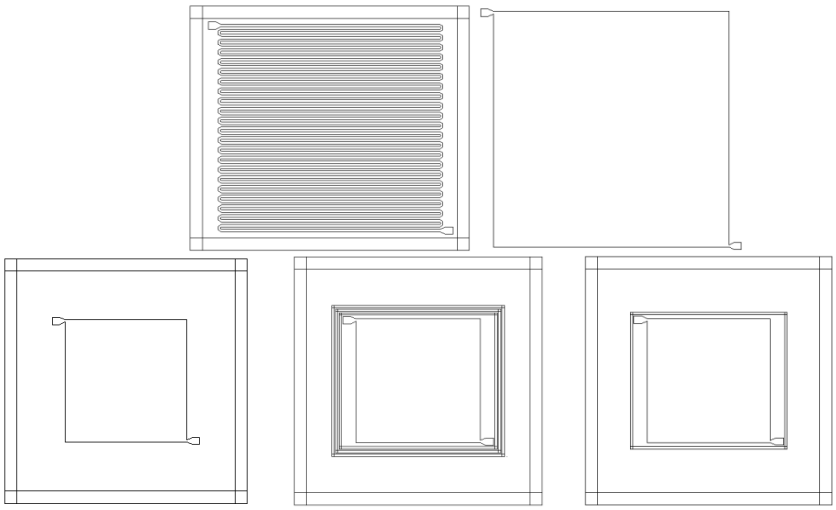
Principally, electronic measures such as the I-V curve to characterize various metal-semiconductor systems were used, using the Meander1 and TIA1 mask designs, examples of which are shown in Fig. 3.40. Standard photolithographic and lift-off techniques to pattern the wafers were used. For instance, Fig. 3.38 shows the I-V curves for Ti/Pt samples. Although they could be finely patterned, they all exhibited high leakage current, which implies that the Fermi level of the titanium participates in the contact-barrier formation. This can be avoided by making increasingly thin Ti adhesion layers (down to 3 nm). Even though the Ti/Pt devices have a high leakage current, they still provide rectifying contacts, which can be used to detect alpha particles, at least.



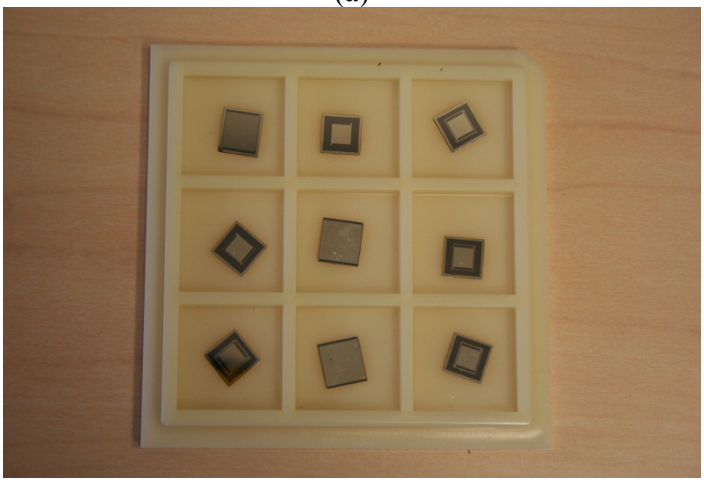
(a) (b)

Fig. 3.40 (a) Designed and (b) fabricated mask for n-type Si plane and meander detector with Cr/Au at pattern side and Al at bottom.

A robust gold-silicon-aluminum detector system also has been developed, but again, gold adhesion can become a problem, which is typically solved by using a chrome adhesion layer. Fig. 3.40a and Fig.3.41a show a mask design and Fig. 3.40b and 3.41b show the wafer and diced samples, which have various guard-ring and meander-line designs using the Cr/Au system fabricated on high-resistivity silicon.



(a)



(b)

Fig. 3.41: (a) Detailed design of meander and planar detectors with guard-rings and (b) various diagnostic detectors after fabrication and dicing.

The guard-ring implementations, examples of which are shown in Fig. 3.41a where a rectangular design with a 100 micron pitch and width were used, are employed to diminish the surface leakage current. Some representative IV-curve results are shown in

Fig. 3.42. As shown in Fig. 3.42, the leakage current is dependent on the sample size and the degree to which the surface is cleaned. Furthermore, Fig. 3.42d shows that the number of *floating* guard-rings was not critical, as the leakage current does not substantially vary between the devices. More importantly, the results show that as in the Ti/Pt sandwich case, the 6.6 nm Cr and 300 nm Au sandwich makes a high leakage contact, presumably because the barrier height is pulled down by the presence of the chrome. This conclusion is buttressed by a comparison with gold contacts, for which the leakage current is several orders of magnitude smaller than in the Cr/Au case, regardless of the chrome thickness. Again, radiation signals can be extracted from these detectors but a lower-noise material system was needed in order to increase the charge sensitivity.

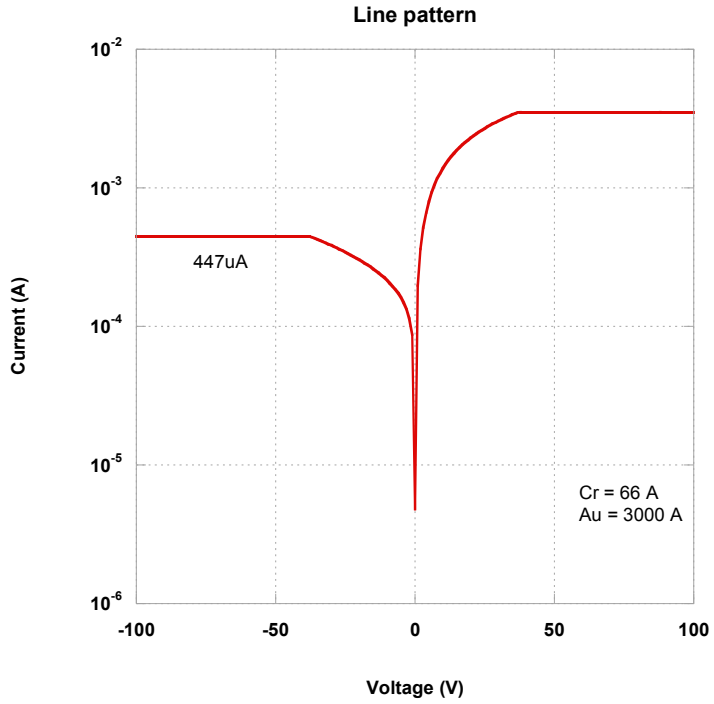


Fig. 3.42a: I-V curves of the Cr/Au-patterned detectors designed as a line.

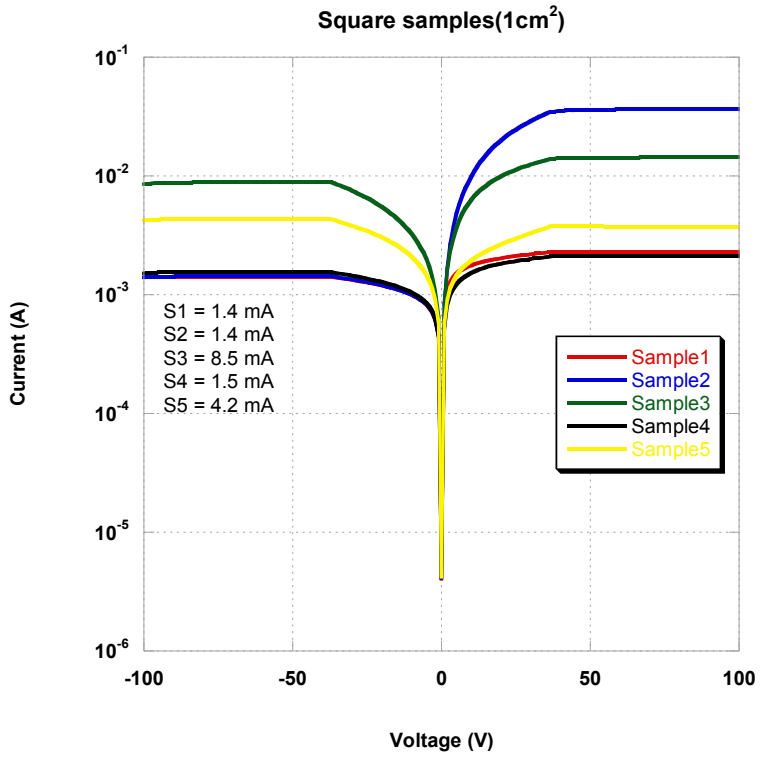


Fig. 3.42b: I-V curves of the Cr/Au-patterned detectors designed as a 1 cm^2 plane.

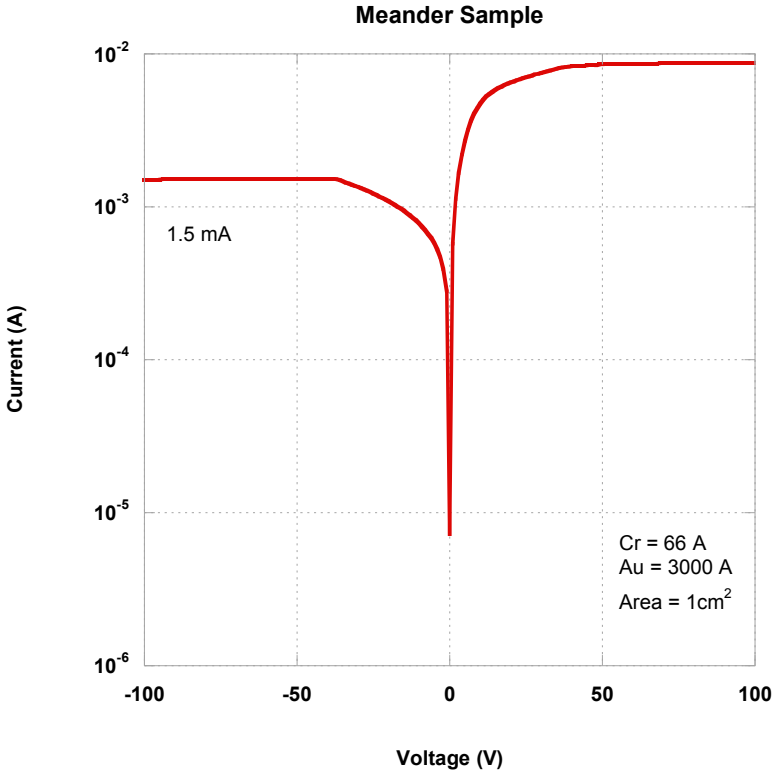


Fig. 3.42c: I-V curves of the Cr/Au-patterned detectors designed as a meandering line.

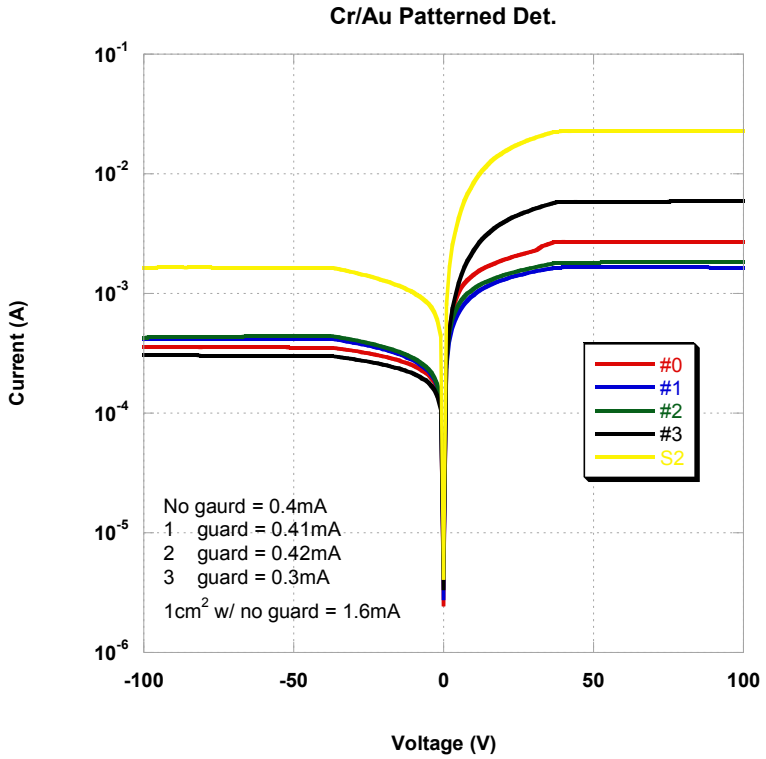


Fig. 3.42d: I-V curves of the Cr/Au-patterned detectors designed as planar detectors with various numbers of guard rings.

The depletion depth and the bulk capacitance of an n-type Si detector versus reverse bias with varying the area were simulated as Fig. 3.43-44. In order to make a 2nd generation detector, we used the high resistive (> 15 kΩ-cm) n-type Si wafer, which has a phosphorus dopant, 100 orientation, and 500 μm thickness. For a 500 μm thick n-type Si detector, which has 15 kΩ-cm resistivity, $8.85\text{x}10^{-14}$ permittivity [F/cm], and a relative dielectric constant 11.9, the calculated depletion voltage was 66.167 V [5]. The full depletion voltage (V_{dep}) was 66.198 V with using [9].

$$V_{dep} = \frac{q}{2\varepsilon\varepsilon_0}\left|N_{eff}\right|d^2 - V_{bi}, \tag{3.1}$$

$$V_{dep} = 4\left[\frac{\Omega\cdot cm}{\mu m^2}\right]\frac{d^2}{\rho_n} - V_{bi}, \quad \text{for n-type Si} \tag{3.2}$$

where, N_{eff} is effective dopping concentration ($2.8763\text{x}10^{11}$), d is thickness of wafer (500 μm) and V_{bi} is built-in voltage for Si detector (~0.5 V). The ~66 V predicted value compared well with the ~70 V measured.

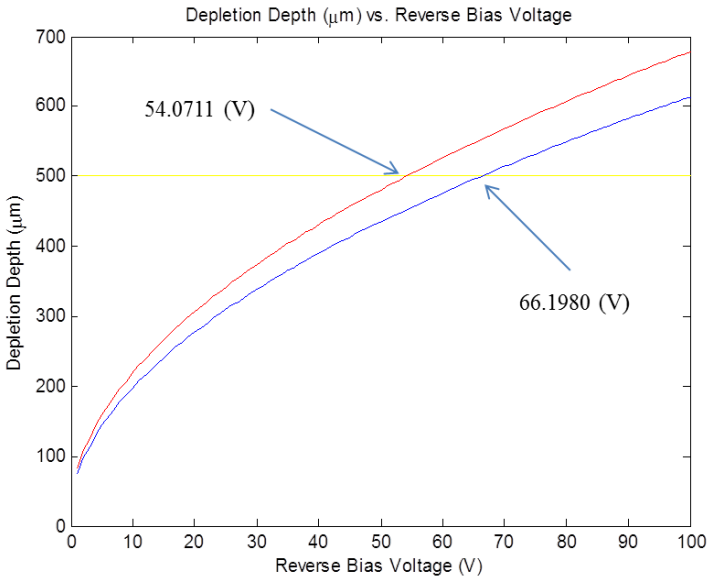


Fig. 3.43: Graph of depletion depth variation via reverse bias voltage for n-type Si detector.

Fig. 3.44 shows the simulation result for the depletion depth variation via reverse bias voltage for an n-type Si detector, and Fig. 3.44 and Fig. 3.45 show the bulk capacitance variation via reverse bias voltage for n-type Si detectors with different areas. The full depletion voltage is usually determined using the capacitance-voltage method.

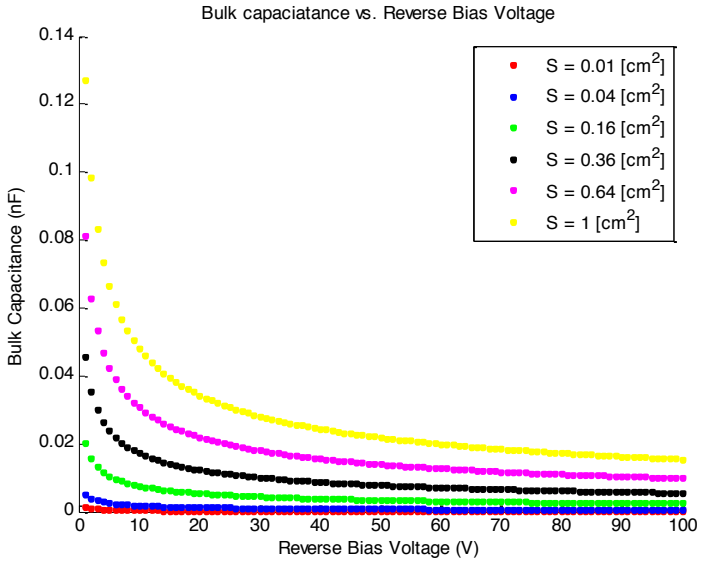


Fig. 3.44: Graph of bulk capacitance variation via reverse bias voltage for n-type Si detector with different areas.

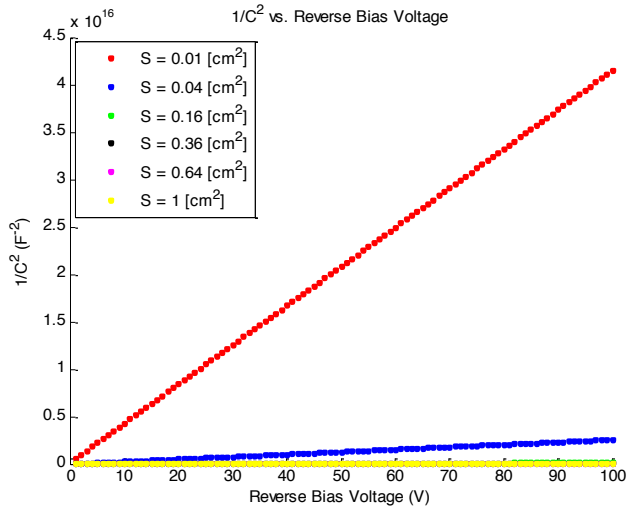


Fig. 3.45: Graph of $1/C^2$ via reverse bias voltage for n-type Si detector with different areas.

The boundaries of the depletion region act as the two plates of a planar capacitor. With increasing reverse bias, the distance between the planes increases and then the bulk capacitance decreases. The bulk capacitance is also dependent on the active area size. The equation to calculate for the bulk capacitance (Cbulk) is shown below.

$$C_{bulk}(V) = \frac{dQ}{dV} = S\sqrt{\frac{q\varepsilon\varepsilon_0|N_{eff}|}{2(V_{bi}+V)}}, \tag{3.3}$$

where, S is surface area, q is charge, $\varepsilon\varepsilon_0$ is dielectric constant, N_{eff} is effective dopping concentration, V_{bi} is built-in voltage and V is reverse bias.

3.4. The Second Generation Detector Design and Fabrication

With the patterned wafers, we made detector modules with two types of contacts between the metallic electrodes and the readout pins of the mounting sockets, and I-V curves were measured for various conditions, as shown in Fig. 3.46. Alterations in the mounting method and material did not impact the electrical performance. The practical consequence is one does not need wire-bonds during the electrical testing. For the planar detector modules, the leakage current was about 35μA, but in the meander-line modules, it was about 3.5 μA.

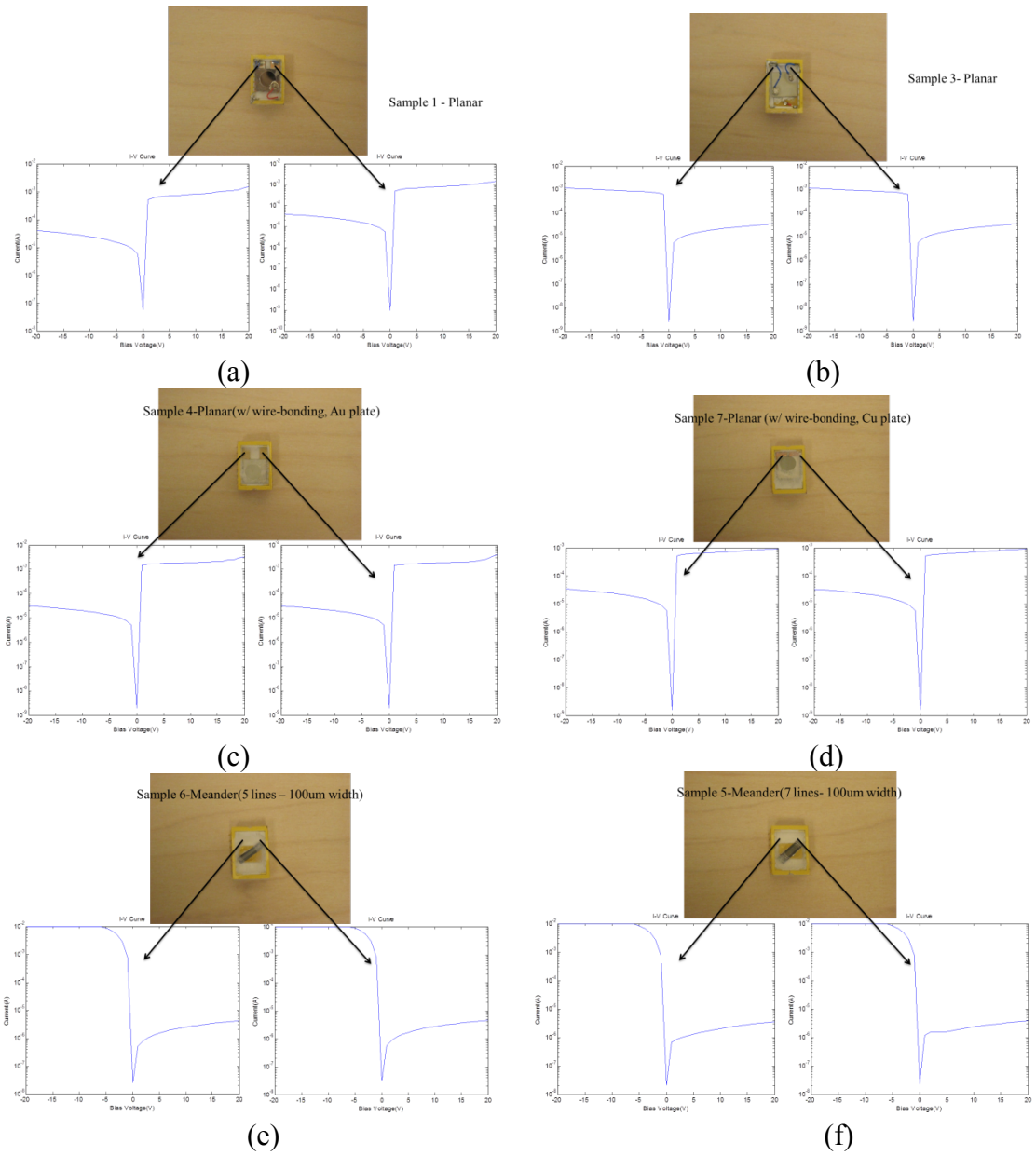


Fig. 3.46: Current-voltage relationships for (a,b) planar type Si detectors bonded with conductive epoxy, (c) a planar type detector bonded with Au wire, and a Au plate, (d) a planar type detector bonded with Au wire and a Cu plate, (e) 5 line meander detector module, and (f) 7 line meander detector.

To compare the Schottky contact and reduce the leakage current of the detector, we fabricated the P-I-N planar detector with several process steps, as shown in Figs. 3.47-49. These detectors also were passivated with a 2500 Å thick SiO_2 layer. The process steps are more complex than Schottky detectors, but each detector sample has a good uniformity in the sense of leakage current, as reflected in Fig. 3.50. Each sample

has a leakage current of about 100 nA. In previous results, a 1 μA scale leakage current was evidenced because an oxidation layer did not exist. However, it helped to reduce the leakage current by sintering steps with forming gas composed of 90 % nitrogen gas and 10 % hydrogen gas in the furnace at a temperature of 400 ^{o}C.

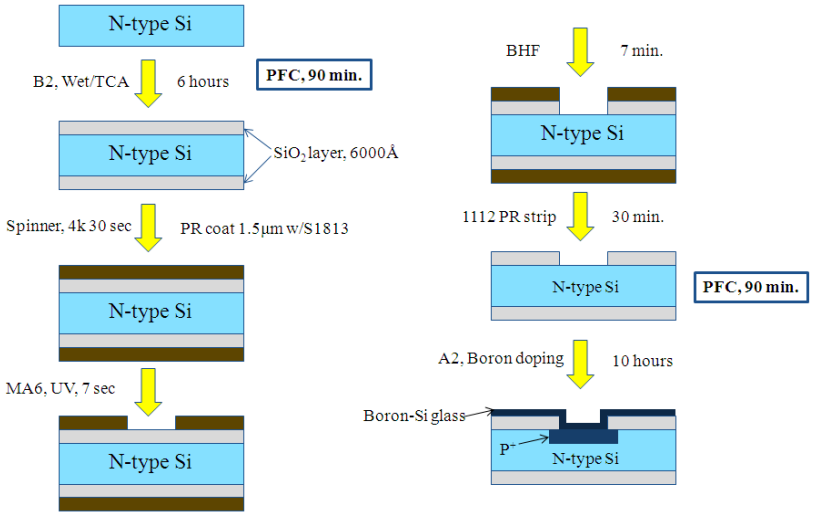


Fig. 3.47: Process steps for doping the P^{+} layer with Boron.

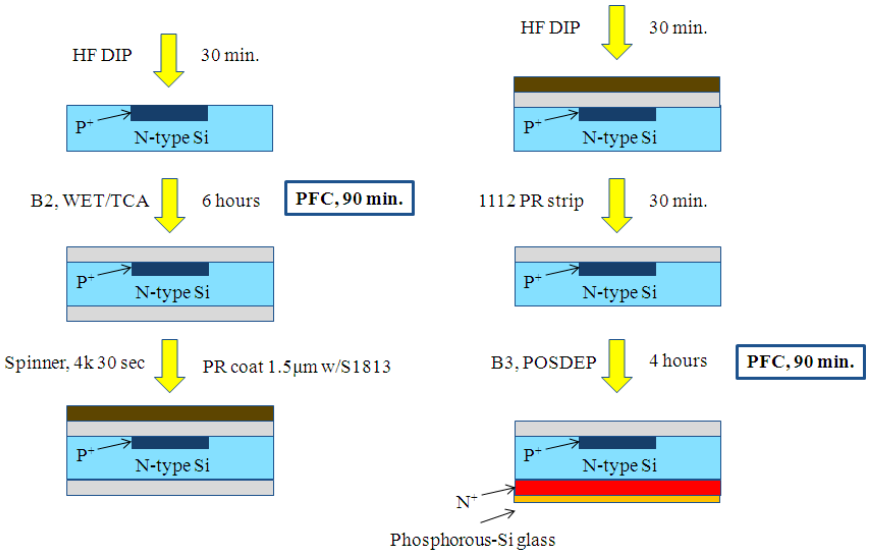


Fig. 3.48: Process steps for doping the N^{+} layer with Phosphorous.

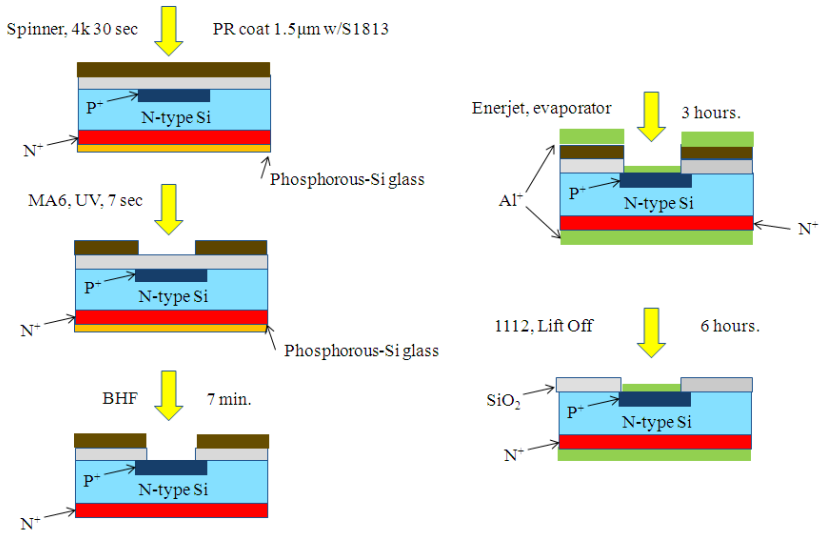


Fig. 3.49: Process steps for making the final P-I-N detectors.

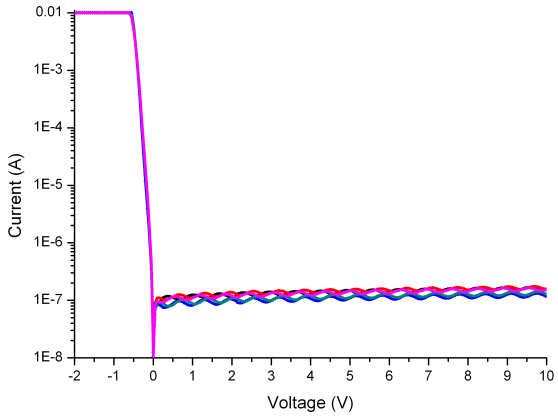


Fig. 3.50: I-V curve of the P-I-N planar detectors with different metals on P^+ layer and passivation.

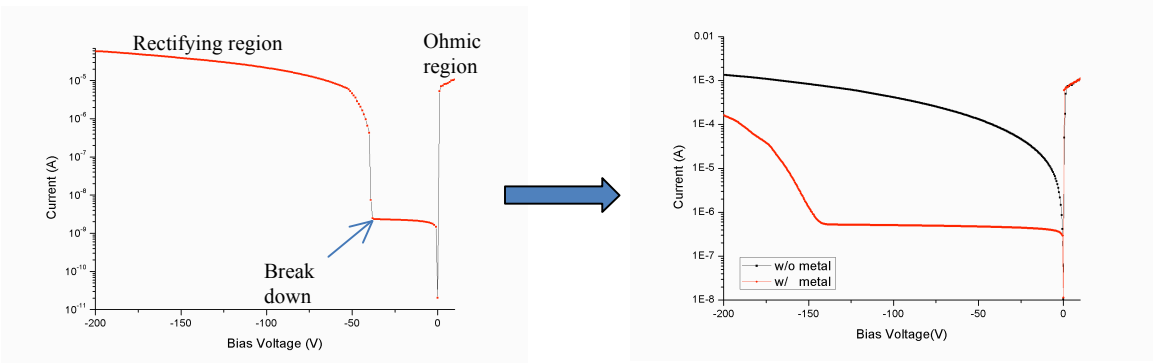


Fig. 3.51: Graph of I-V comparison of delay-line detectors which have a low breakdown voltage (left) and a high breakdown voltage (right).

Before the design of the PIN delay-line detectors, specific conditions should be considered as follows. First, for a fine position sensing (both in depth and laterally), a better signal to noise ratio (SNR) is needed. In order to obtain signals with good SNR, application of a high bias, up to 70 V to make a full depletion, is required. Moreover, a detector with low leakage current and high breakdown voltage are necessary as seen in Fig. 3.51 (right). For this purpose, the metal field plates (MFP) and multiple guard-ring structures were applied. Second, a P+ region with 1μm thick and width with low temperature boron doping (furnace temperature < 1000 ^{o}C) are essential, because other high temperature process, after boron doping process, should be excluded. Third, minimization of the conductor and dielectric losses with thick metal (> 1 μm) pattern with over-lay is needed. With these considerations, delay-line detectors were fabricated with following process steps (Fig.3.52):

1. Poly-Si and Phosphorous doping at back side
    a. Low temperature oxidation(LTO) with 450 ^{o}C LPCVD
    b. Photo-resist coating at front side
    c. Etching the back side LTO with 2 min. of BHF
    d. Removing the photo-resist
    e. Deposit the poly-Si with LPCVD
    f. Phosphorous doping at 975 ^{o}C (~1 μm depth) for 1hr.
2. Remove the LTO with HF and poly-Si at the top side with plasma etching (100 sec)
3. LTO with LPCVD ~5000 Å for mask (1) of boron layer
4. Photo-resist coating at both side & development at front side
5. Wet etching the LTO for the window of boron
6. Boron doping at 975 ^{o}C (< 1 μm depth) for 5 min.
7. Remove the LTO at both sides with HF
8. Wet thermal oxidation at 900 ^{o}C for 2000 Å thick and LTO for 5000 Å
9. Photo-resist coating and development with mask (2)
10. Pattern wet etching with 3.5 min. BHF for oxide layer
11. Photo-resist coating and development with mask (3)
12. Pattern wet etching with 0.5 min. BHF for metal deposition
13. Metal deposition and Lift-off with acetone for 2 days
14. Annealing at 150^{o}C with Rapid thermal process (RTP) for 10 min with H_2/N_2

foaming gas

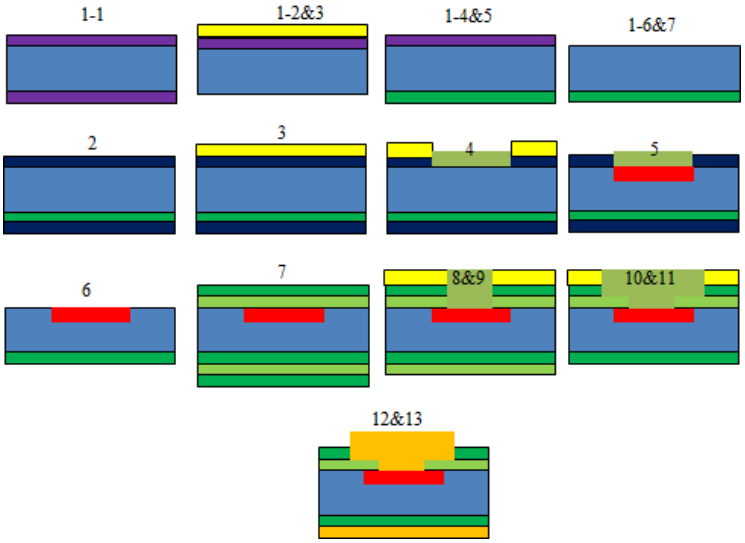
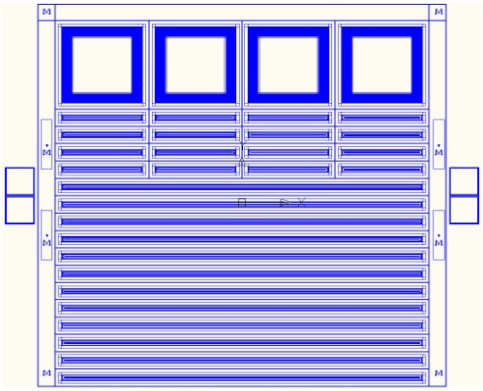
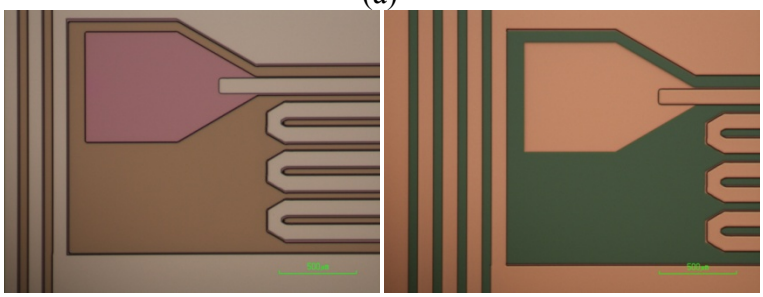


Fig. 3.52: The fabrication process used in this work.



(a)



(b) (c)

Fig. 3.53: (a) The figure of the mask for boron doping, (b) the delay-line detector area before the metal deposition, and (c) after the metal deposition and lift-off.

With prepared wafers having delay-line detectors, over-lay detectors were fabricated with the following process steps.

1. Prepare the wafer with delay-line detectors
2. Over-lay
   - ① Photo-resist coating & development for oxide layer
   - ② Oxidation with PECVD at 200 ^{o}C
   - ③ Metal deposition with Cr/Au
   - ④ Photo-resist coating and development for metal etching
   - ⑤ Oxide etch with plasma etching and O_2 ashing
   - ⑥ Remove the photo-resist

Fig. 3.54 shows the I-V characteristics of delay-line detectors with thin oxide (2000 Å) and thick oxide (5000 Å) for the passivation. The thick oxide can make a broadening electric field under the metal field plate, but it has a higher leakage current than thin oxide one.

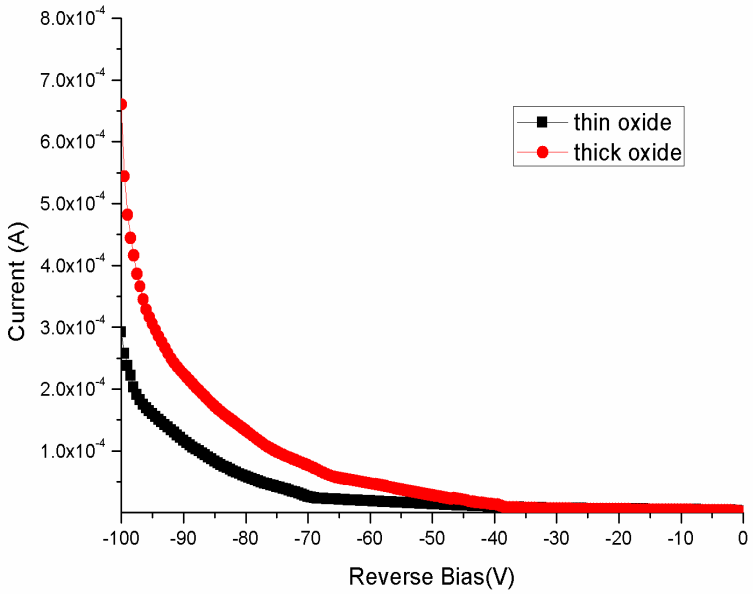


Fig. 3.54: The I-V characteristics of delay-line detectors with thin oxide for passivation.

Fig. 3.55 shows the sintering effect on the I-V characteristic. The sintering can help to mitigate the stress due to the vacant atom at the Si surface by filling the void with hydrogen. In this case, we used tungsten in order to match the contact between the top of the metal surface and probe tips. In addition, we also tested the effect of metal field plates as shown Fig. 3.56. Using metal field plates helps to increase the breakdown voltage up to 100 V.

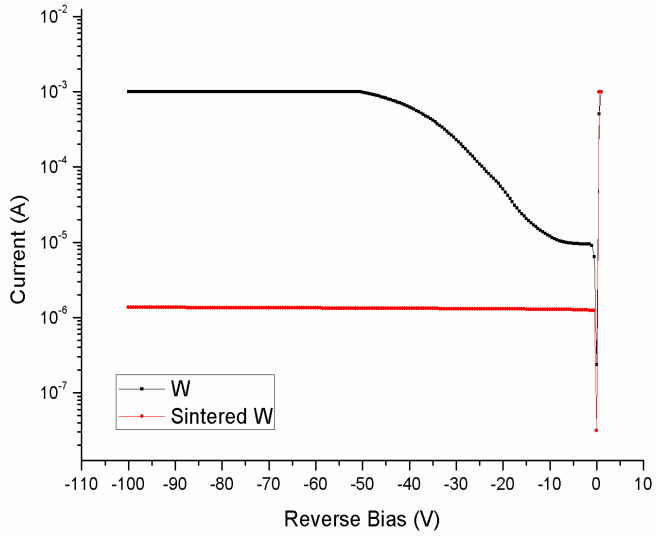


Fig. 3.55: The I-V characteristics of delay-line detectors with and without sintering.

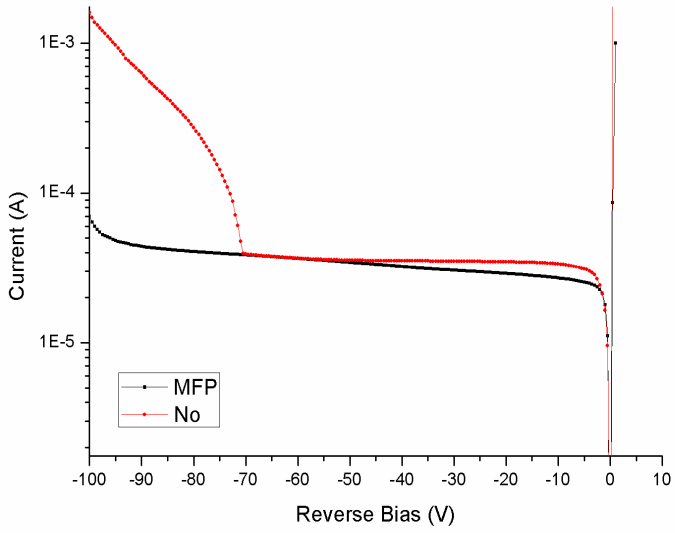
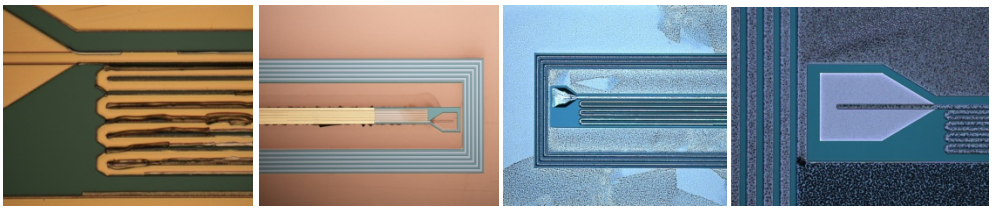


Fig. 3.56: The I-V characteristics of delay-line detectors with and without metal field plates (MFP).

However, we still have challenges regarding the low leakage current devices, which have a high breakdown voltage and over-lay as below:

A. Interface between SiO_2-Si decides the surface leakage current, so it needs a dry oxidation with a high temperature (1100 ^oC) which can change the doping profiles.
B. Thick SiO_2 layer for increasing the breakdown voltage causes increase of the leakage current.
C. Thick metal deposition with high conductive metal as Ag, Au or Cu causes difficulty in the lift-off due to the poor adhesiveness between Si and metals.
D. Wet etching is not good for the over-layer, so it needs a plasma etching for the oxide layer which can causes a hardening the photo-resist and a difficulty in lift-off step



(a) (b) (c) (d)

Fig. 3.57: (a) Pictures of damaged detectors during the fabrication about (a) lift-off of thick metal, (b) wet etching problems in the over-lay process, and (c, d) heat damages during the sintering.

Thus, to solve these challenges, the follow solutions are suggested.

A. Interface between SiO_2-Si reduces the surface leakage current with using a thin dry oxidation and thick LTO.
B. Thick metal deposition with high conductive metal as Ag, Au or Cu with a thick photo-resist (S1827 or SPR220 for 3 μm thickness)
C. Over-layer with a plasma etching for oxide with using the hot PRS2000 and acetone to remove the hardened PR.
D. Sintering with high temperature using by the RTP with 100 ^oC for 30 min.

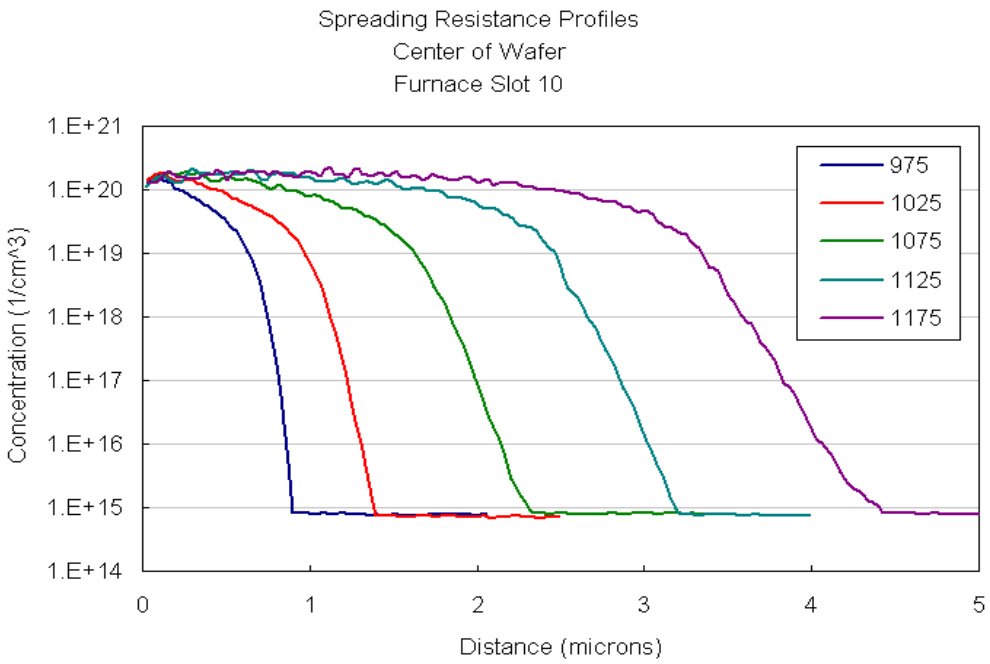


Fig. 3.58: The graph of doping profiles for A2 furnace with increasing the doping temperature in the University of Michigan Fabrication L, Ann Arbor, MI.

In Fig. 3.58, a doping profile of the A2 furnace, which was used for the boron doping to make a p+ region, is plotted. The p-doped regions require a narrow pattern as 1 μm width and gap, so these doping profiles should be considered for designing several masks. If it has a margin area like an overhang metal filed plate with 4μm, the temperature can be increased from 975 ^{o}C to 1125 ^{o}C. When the PIN detectors have a deep doping region, the detectors would have a lower leakage current. Because the thin doping region can make tunneling effects easily, it has a higher leakage current.

3.5 The Third Generation Detector Design and Fabrication

The third generation mask is designed, as shown in Fig. 3.61, having a detector active area 10 x 10 mm^2 and a total area, including multiple guard-rings, of 15 x 15 mm^2. Across the face of the diagnostic mask shown in Fig. 3.59, there are variety of meander designs with different size line widths (and gaps), and two types of planar detectors with 50 guard rings.

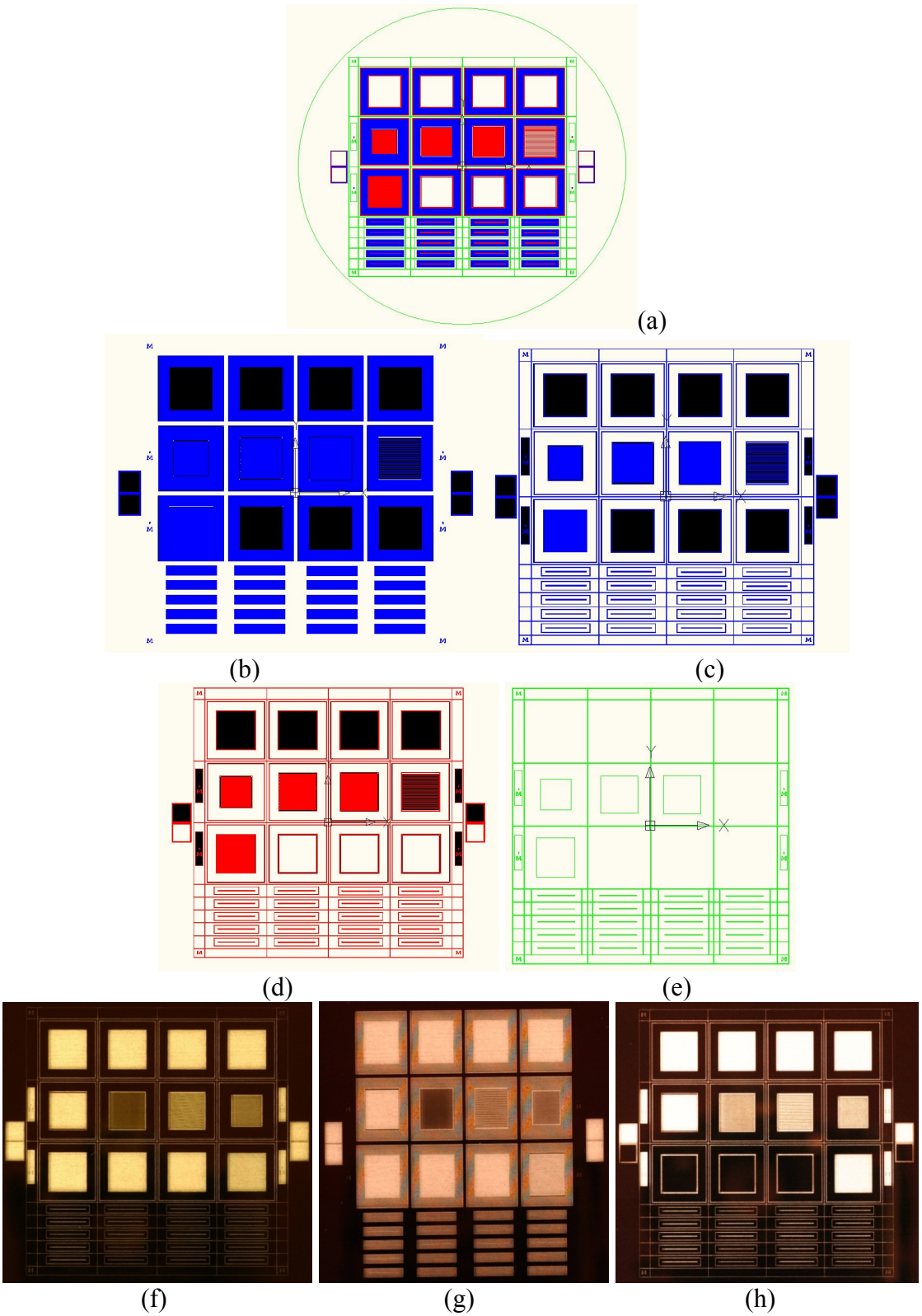


Fig. 3.59: Mask design for new meander and planar detectors. (a) Total mask layout, and masks for (b) boron doping, (c) oxide etching, (d) metal deposition and (e) overlay, and (f), (g), and (h) pictures of the fabricated masks, respectively.

As in the other mask designs, a test pattern (at the extreme right and left sides of Fig. 3.59a) is used to measure the overall electrical characteristics- typically extracted with I-V and C-V curves. Following all of the process steps, the resulting meander-line and planar detectors, are shown in Fig. 3.60.

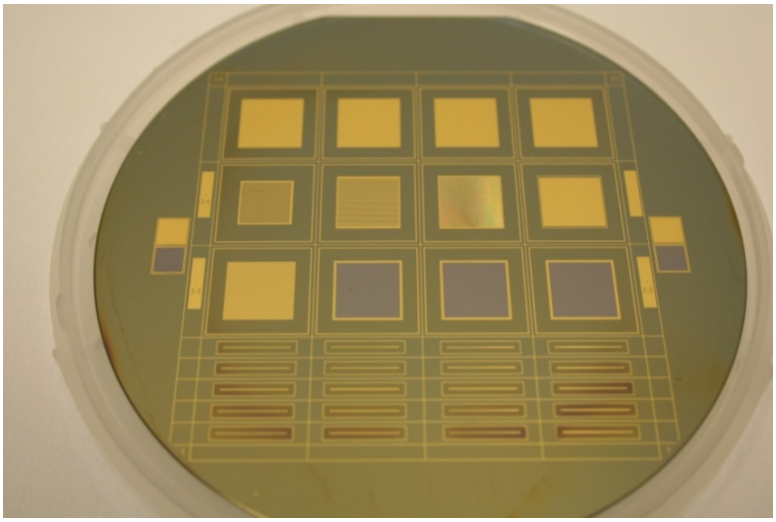


Fig. 3.60: Picture of the fabricated wafer with meander line and planar detectors, which has active area of 10 x 10 mm^2.

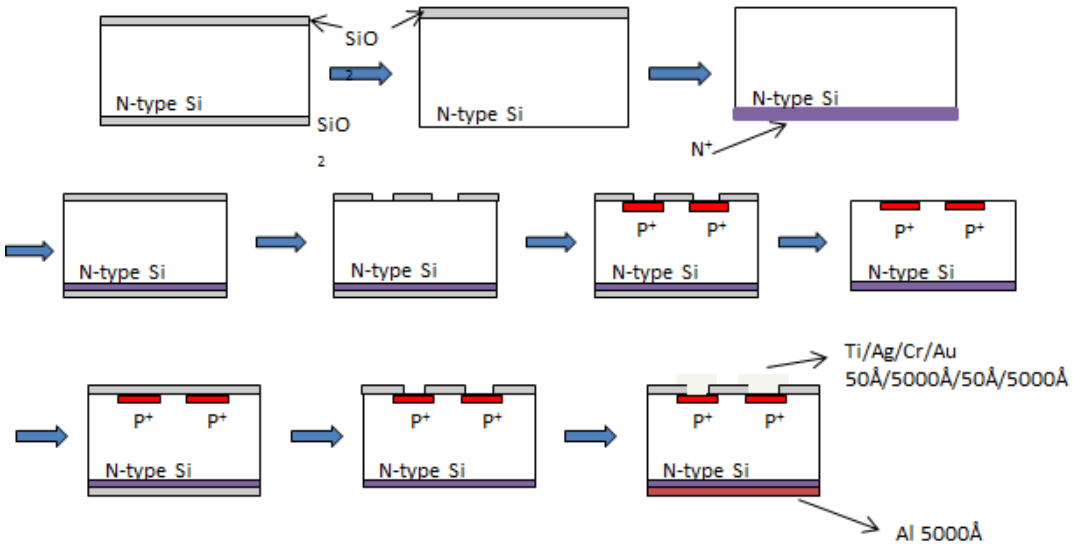


Fig. 3.61: Process steps for generation-3 meander line and planar detectors.

Fig. 3.61 illustrates the process steps used. First, we deposit a 1 μm or greater layer of thick wet SiO_2 with the oxidation furnace and strip one side to make an open

window for the phosphorous doping. After doping the phosphorous (> 1 μm depth) to make an n+ blocking layer for gettering, we deposit another SiO_2 layer with > 1 μm thick to make an open window for boron doping (at 975 ^oC for 5 min) on the opposite face, the patterned oxide being removed via wet etching (at a rate of approximately 1000 Å/min). Fig. 3.62 shows the fabricated 5 μm width meander line detector with innermost and outer guard-rings.

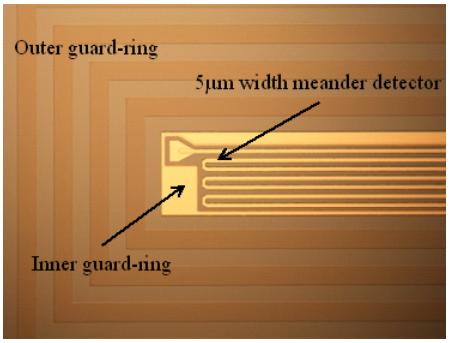


Fig. 3.62: 5 μm width meander line detector with innermost and outer guard-rings.

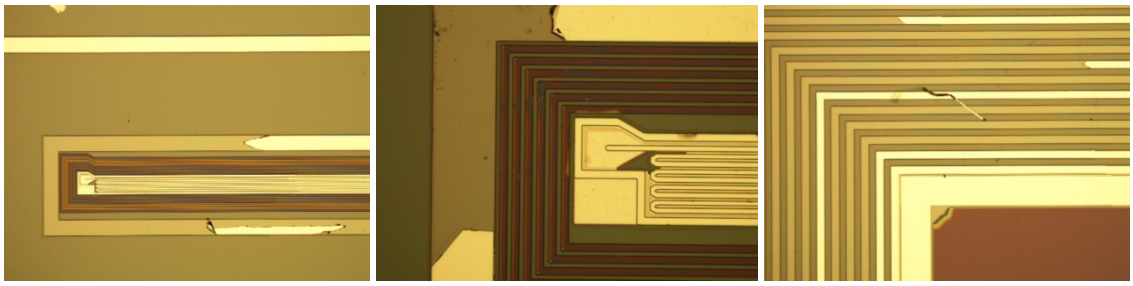


Fig. 3.63: Pictures of weak adhesiveness example between Si and thick silver.

As shown in Fig. 3.58, in which we characterized the doping profile of the furnace. One should use a low temperature of 975 ^oC (or lower) if desiring a shallow doping profile. After boron doping, we deposit the dry passivation oxide with a 2000 Å thickness and make a pattern for the T-shape metal cap, while removing the oxide on the opposite side at the same time. Finally, using the sputterer, we deposit the layered metal electrode with the following thicknesses: 50 Å Ti, 5000 Å Ag, 50 Å Cr and 5000 Å Au. Because of the adhesive weakness between the Si and Ag, we employed Ti as an adhesion metal. Furthermore, the adhesion layer Cr is required between Ag and Au to realize a good sandwich. The weakness of the adhesion between Si and Ag is illustrated

in Fig. 3.63. The fabricated detectors which have the two types of planar detectors and 2, 5, 10 and 100 μm width meander line detectors, are shown in Fig. 3.64.

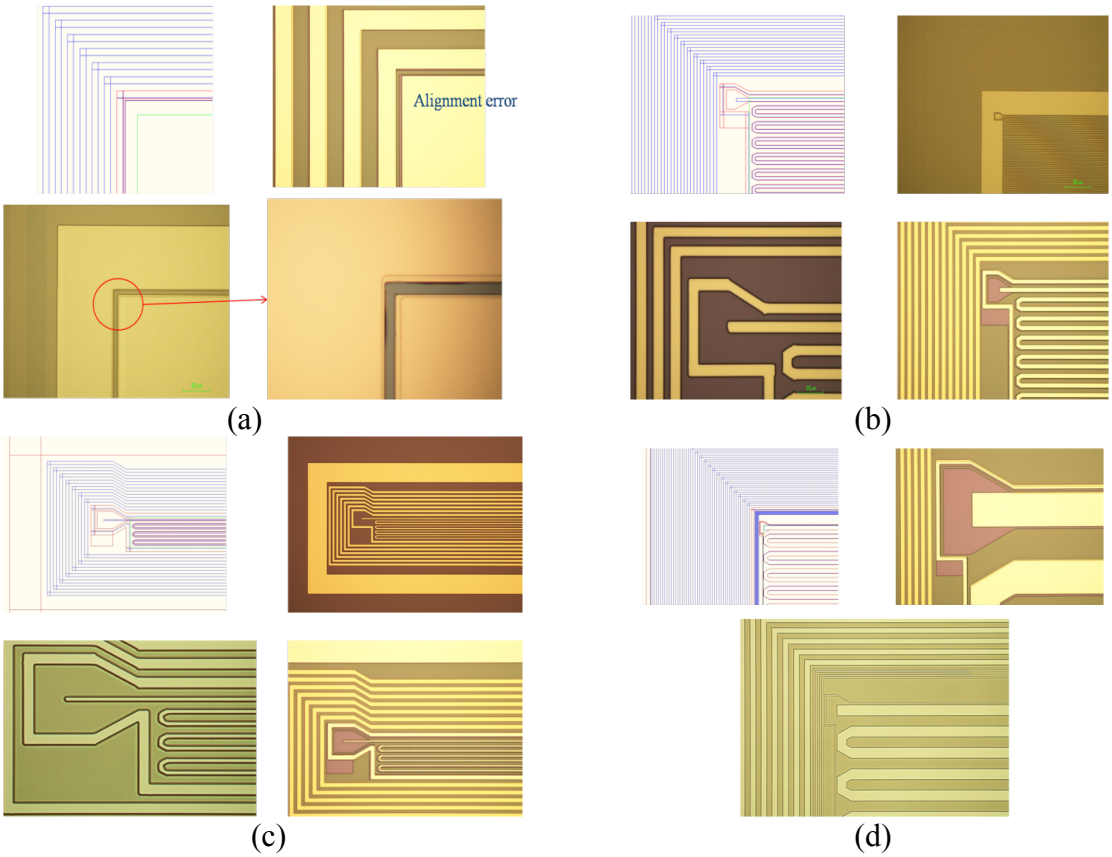
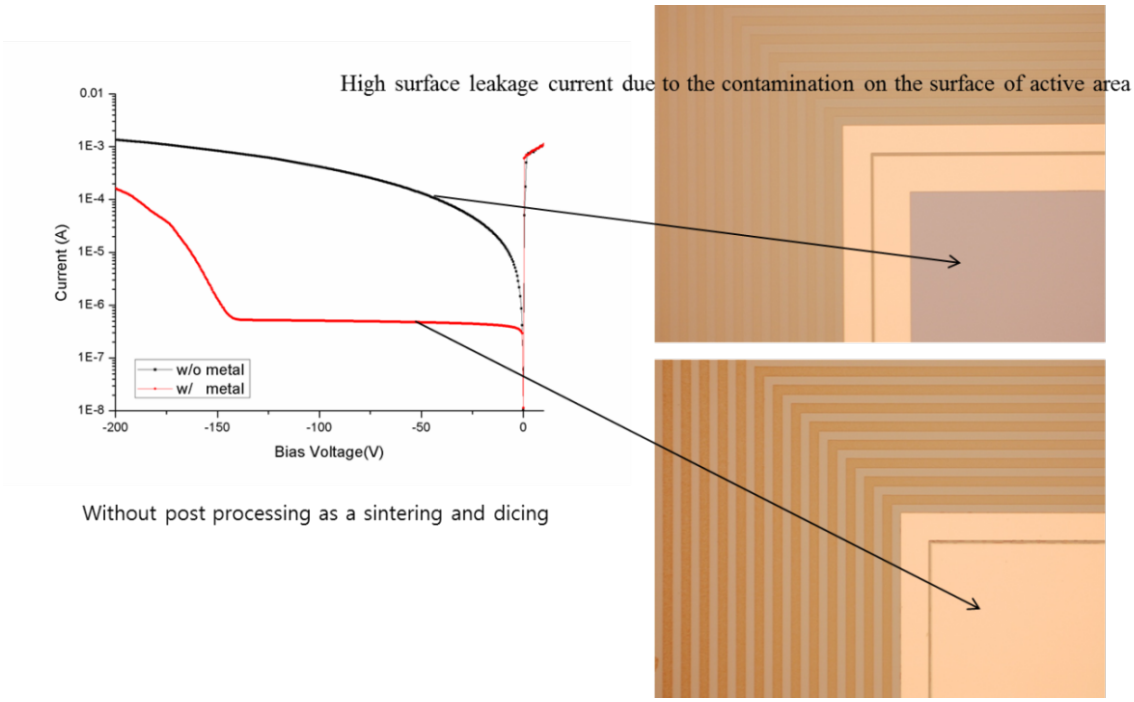


Fig. 3.64: Pictures of (a) the planar detector and meander line detectors with different pitches (b) 5 μm, (c) 2 μm and (d) 100 μm.

Fig. 3.65a shows the I-V curves measured from the two types of planar detectors as shown in Fig. 3.65b, in which the thick Ti/Ag/Cr/Au electrode is employed. When the metal contact is not employed, the surface leakage dominates the behavior, as shown in the black curve. In the red curve, the leakage current is shown to be superior to that realized previously, because the several hundred nano-ampere leakage is realized without post-processing steps that include H-annealing and guard ring bias. These steps together reduce the leakage current by a factor a factor of ~1000. Furthermore, the soft-breakdown point has been increased to -150 V and is expected to be enhanced further when the guard ring is biased.



(a) (b)

Fig. 3.65: I-V curve (a) for two type planar detectors, which have metal cover (b, top) and no metal cover (b, bottom).

Although the noise behavior will continue to be optimized, the underlying detectors have acceptable noise behavior so that both ions and electrons can be sensed. However, in order to realize the meander-line design, the low-noise detector must be coupled to an optimized transmission-line design, which includes not only the thick conductor but also a grounded dielectric with high permittivity.

3.6 The Fourth Generation Detector Design and Fabrication

In order to maximize the charge-sensitivity, and therefore increase the position and energy resolution of the detector, low surface-leakage detectors are fabricated with a PIN structure on a substrate of high-resistivity silicon (> 10 kΩ-cm). For the PIN detectors, the process sequence to fabricate the detectors, illustrated in Fig. 3.66, is as follows.

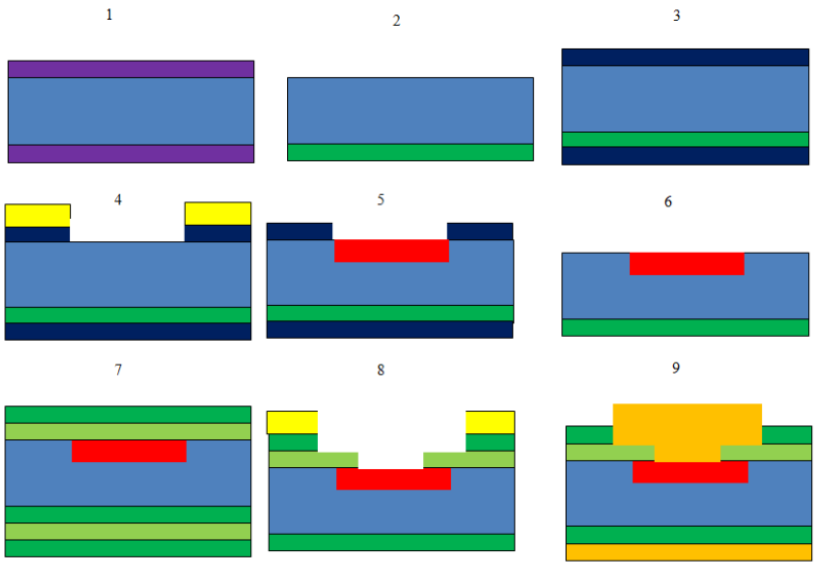


Fig. 3.66: A flow chart of the fabrication process for an n-type substrate with front side B doping with back side gettering.

An initial low temperature oxide (LTO) of 0.5 μm was grown at the temperature of 415 °C (Step 1 in Fig. 3.66). After removal, via wet-etching, of the oxide on the backside of the wafer and the deposition of polycrystalline silicon (Step 2), an n+ layer was formed via phosphorous doping with a diffusion furnace at the temperature of 975 °C (Step 3). In order to form a p+ active region, boron was diffused at 1050 °C for 20 minutes to make a shallow depth doping profile, after the removal of the initial oxide (Steps 4 to 6). Thick silver (Ag, 1 μm), used to minimize the signal's conductive losses, and a thin layer of gold (Au, 500 Å), used to increase the integrity of the wire-bond between the detector and the readout circuit, were sputtered to form the front-side metal anode. Aluminum (Al, 3000 Å) was sputtered on the backside of the detector to form the cathode (Steps 7 through 9). Finally, sintering with forming gases, which are comprised of 95% nitrogen and 5% hydrogen, was carried out to optimize the ohmic contacts and saturate the silicon dangling bonds. A post-processing annealing with a furnace at 200 ^{0}C for 10 minutes was also performed.

Aside from the detector's noise characteristics, another challenging problem is to minimize the loss and distortion in the signal as it propagates along the delay-line. The grounded insulated overlay structure not only minimizes the shaping of the signal, but the

use of dielectrics with low loss-tangent can reduce the degree of dielectric loss. Fig. 3.67 shows the mask design (left) and test patterns (right) for testing the charge sharing effects with 10μm and 100μm width meander lines.

With fabricated test device we measured I-V characteristics, as shown in Fig. 3.65. When the device does not have Metal field plates (MFPs) the leakage current at -10 V was quite low- 6 nA (over a 1 cm^2 area), but it had soft breakdown at -30 V. However, the breakdown voltage is increased with the application of the MFP's, as shown Fig. 3.68.

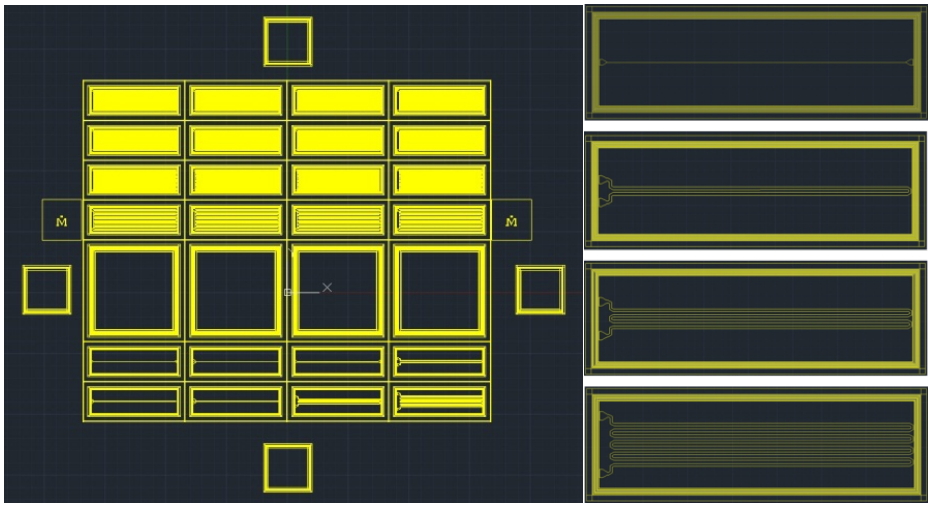


Fig. 3.67: Mask design (left) which has 10, 20, 25, 50 and 100 μm meander lines and planar. Right test patterns are for testing the charge sharing effects with 10 μm and 100 μm width meander lines.

For the five turn meander pattern with a 5 μm strip and gap width, the application of MFP's to the outer edge of the serpentine pattern as well as to the inner edges of the guard rings results in an increase of the soft breakdown voltage from 30 V to 200 V (the maximum voltage available on the test instrument- in experiments, the breakdown extends to 400 or 500 V). Furthermore, the boron diffusion depth was varied from 1μm to 5μm as Fig. 3.58 to see its effects. A furnace with solid boron target was used to deposit boron for the p+ contact. It can result in non-uniformity of the boron doping due to the distance variance between each target and wafer when a shallow doping depth is less than

1 μm. If the temperature is increased from 900 ^{o}C to 1100 ^{o}C, the boron diffusion length will be longer, and it can diffuse the boron uniformly and the I-V character is more stable.

Furthermore a passivation layer with SiO_2 is applied to protect the surface of the detector from dust and humidity to reduce the leakage current and to maintain its electrical properties. However, its electrical characteristics were changed after few weeks. Thus, this passivation layer is changed from SiO_2 to a SiO_2 and Si_3N_4 combination. With this combination, it can keep the device performance without losing its electrical property as shown at Fig. 3.69.

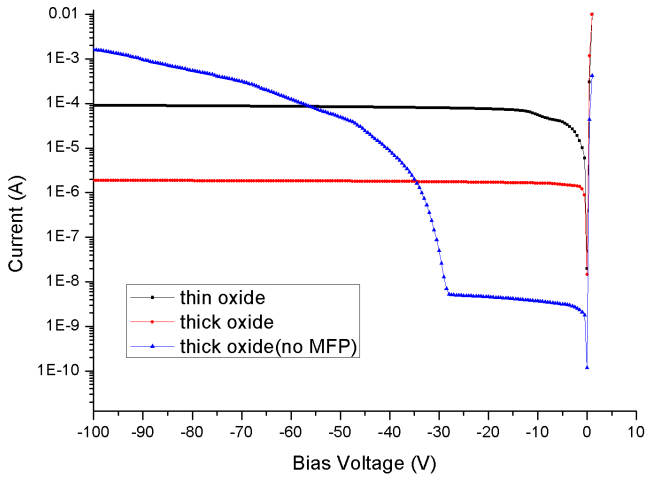
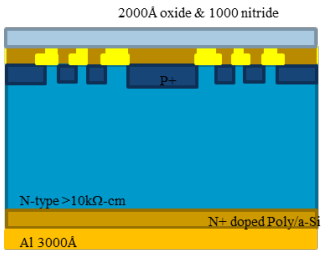


Fig. 3.68: I-V characteristics comparison of thin passivation layer (black), thick passivation layer (red) with metal field plates and without metal field plates (blue).



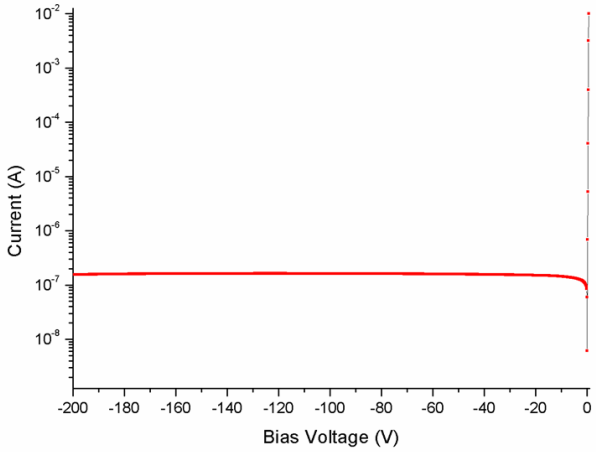


Fig. 3.69: I-V characteristics of fourth generation meander line detector with passivation layer of SiO_2 and Si_3N_4 combination.

3.7 The Fifth Generation Detector Design and Fabrication

Our clean room has been undergoing a continuous upgrade, one part of which is the replacement of the boron and phosphorous doping and oxidation furnaces with newer, cleaner units. The benefit of the new furnaces is better control over the process and lower impurity concentrations; however, they are different animals that require one to re-characterize the process in order to get similar device performance. Previously, we could achieve ~100 nA leakage current at -100 V (prior to guard-ring bias). If we can apply an effective gettering recipe, we can reduce the leakage current under that value and increase the breakdown voltage as well. Finding an equally effective or superior gettering recipe was the main challenge during the switch, but we were successful, with some (1 cm^2) devices exhibiting leakage current in the 100's of pA range. We will briefly describe the process used to realize good devices.

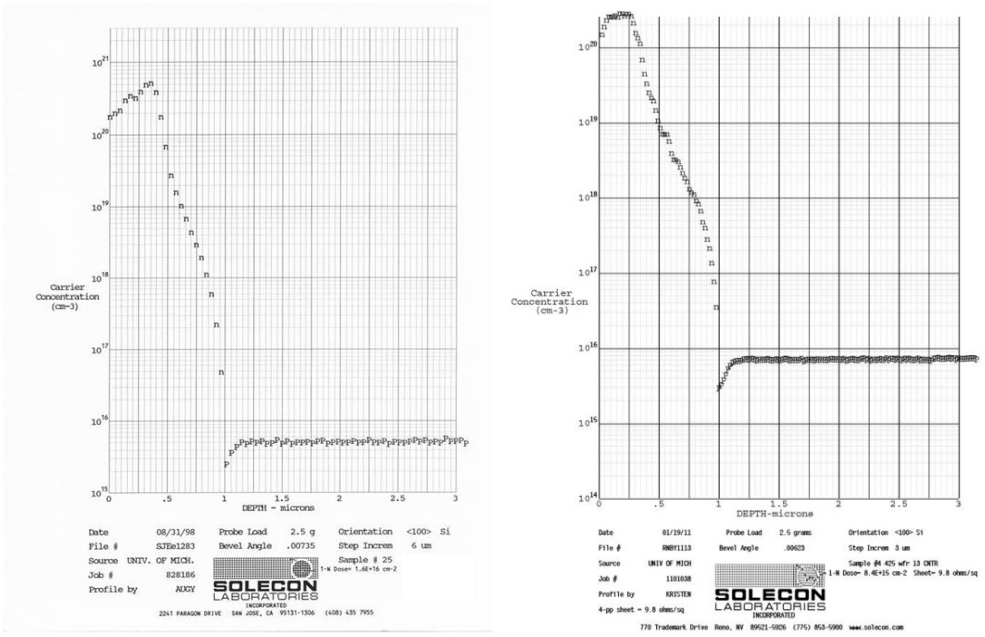
In the old furnace, we did a phosphoryl chloride (POCL) doping with run parameters as follows:

1. Temperature: 950°C
2. Settle time: 10 min.
3. Deposition time: 30 min. with 3 slpm of N_2, 150sccm of O_2 and 375sccm of N_2 and POCL
4. Drive time: 1 hour

With these parameters, we could get the POCL doping profile like that shown Fig. 3.70a. The maximum doping peak concentrate was 5 x 10^{20} (atoms/cm^3), and junction depth was ~0.95 μm when we used Si which had substrate doping level of 8 x 10^{16} cm^{-3}.

In the new furnace, we did a POCL doping with the following run parameters:

1. Temperature: 950 °C
2. Settle time: 10 min.
3. Deposition time: 20 min. with 5 slpm of N_2, 250 sccm of O_2 and 500 sccm of N_2 and POCL
4. Drive time: 25 min with 10 min. N_2 anneal.



(a) (b)

Fig. 3.70: Phosphorus doping furnace's reference profile from old (a) and new (b) furnace.

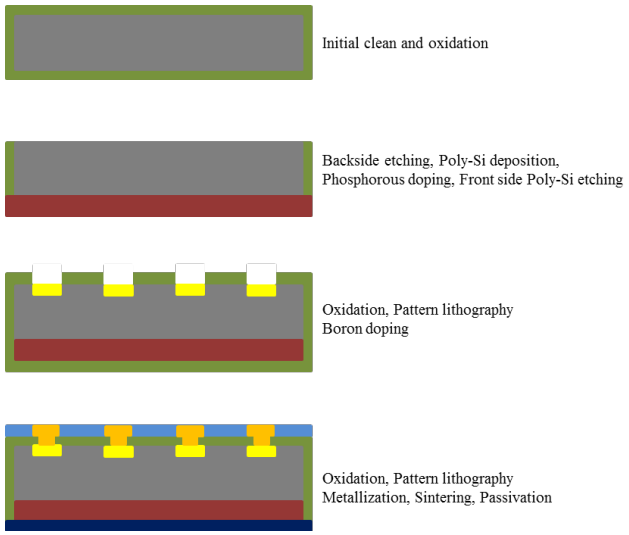


Fig. 3.71: Schematic of important aspects of the fabrication processes.

The new furnace's volume is larger than the old one and we therefore increased the gas quantities, such as those shown in run parameter #3. With these parameters, we could get the POCL doping profile shown in Fig. 70b. The maximum doping peak concentrate was $3x10^{20}$ (atoms/cm^3), and the junction depth was ~1 μm when we used Si which had

substrate doping level of 8 x 10^{16} cm^{-3}. In order to maximize the charge-sensitivity, and therefore increase the energy and position resolutions of the detector, low surface-leakage detectors are fabricated with a PIN structure on a substrate of high-resistivity (> 10 kΩ-cm) double side polished <100> n-type 550 μm-thick 4" wafer by using conventional processes, which include oxidation, furnace doping, photolithography, dry/wet etching and sputtering. For the PIN detectors, the process sequence to fabricate the detectors, illustrated in Fig. 3.71, is as follows.

At first, we prepared starting wafers via a pre-furnace clean, which includes an organic cleaning step, oxide etching, and an ionic clean. For the organic clean we used DI-H_2O + NH_4OH + H_2O_2. After this step, we removed the thin native oxide with 100:1 HF, and the ionic clean utilized DI-H_2O + HCL + H_2O_2, 10 min for each step. We then rinsed the detectors with an automatic SRD (spin/rinse/dryer). An initial thermal oxide of 0.5 μm was grown at a temperature of 1100 °C. After removal, via wet-etching, of the oxide on the backside of the wafer and the deposition of polycrystalline silicon (poly-Si, 0.5μm) with a CVD furnace at 585 °C for 100 min., an n+ layer was formed via phosphorous doping with a diffusion furnace at the temperature of 950 °C for 20 min diffusion time and 25 min drive-in time. For the protection of the poly-Si layer we performed a thermal oxidation with 3000 Å thick wet oxide. In order to form a p+ active region, boron was implanted at 1000 °C for 45 minutes to make a shallow depth doping profile, after the removal of the initial oxide with wet etching (buffered HF). For the pattering, we used S1813 photoresist (4k-rpm spin coat makes 1.5 μm thickness) which has good performance for the wet etching. Following each patterning process, the PIRANHA clean should be used in order to fully remove any organics- most topically, photoresist- from the wafer surfaces. Aluminum (Al, 3000 Å) was sputtered on the front and backside of the detector to form contacts. Finally, sintering with forming gases at the furnace on 425 °C for 15 min., which are comprised of 95% nitrogen and 5% hydrogen, was carried out to optimize the ohmic contacts and saturate the silicon dangling bonds.

For passivation as well as device protection (from humidity, handling, etc.) we deposited a thin oxide (2000 Å) and nitride (3000 Å) on the front side of the detector, and made a window just for the wire bonding pad side with plasma etching tool. Before the dicing, we coated wafers with photoresist, which can block the surface and keep it clean

from city water during dicing. Finally, we cleaned the diced detectors with IPA and acetone for 10 min. and anneal at 150 °C for 10 min. in order to dry the devices. Fig. 3.72 shows the I-V curves before and after dicing and annealing. Note that the main goal is to preserve the behavior of the devices achieved before dicing, but the extra annealing step yields about a factor of 10 improvements in the device leakage. Note that these values are all measured without guard ring bias. Fig. 3.73 shows the fabricated silicon PIN-type radiation detector. Most pertinently, the newer detectors have leakage currents that range from ~10 nA on the high end to 100 pA on the low end, the variation due to the non-uniform effectiveness of the getter layer. Thus, in our goal of minimizing both the detector and circuit noise in order to optimize the radiation-imaging performance, the silicon detectors do not constrain that performance.

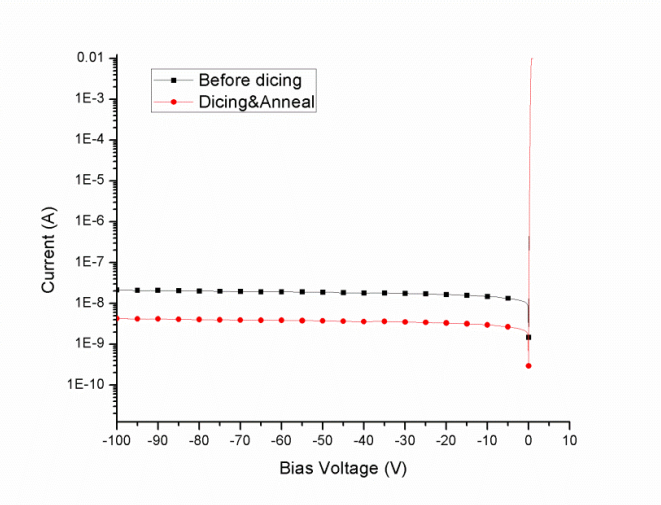


Fig. 3.72: I-V characteristics about the effects of dicing stress and annealing process.

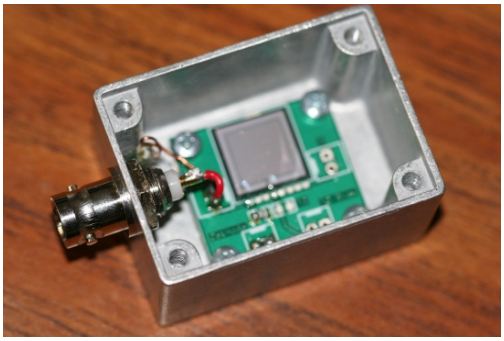
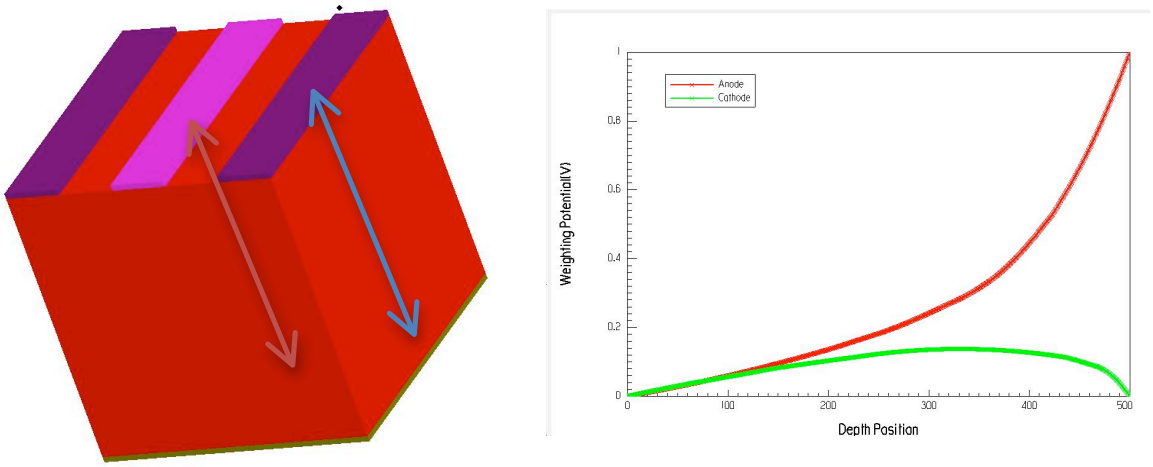


Fig. 3.73: Pictures of the fabricated silicon PIN-type radiation detector for low energy gamma spectroscopy. The active area was 1cm^2.

3.8 Delay-line CZT modeling and fabrication

In addition, the delay-line technique is applied onto the CZT as shown Fig. 3.76. Before the fabrication, we want to see the current signals and compare them with Si. Fig. 3.74 shows the geometry of CZT, plot of weighting potential in anode and outer guard ring and the table, which have physical characteristics of Si and CZT.

With this information, the simulation is run to see the current signals and can do a comparison. Fig 3.75 shows the simulation results with using the planar geometry for Si and CZT. The main difference of the signal profiles is due to the low hole mobility. So the signal can be more useful to discriminate the difference of time between the electron and hole drift.



Material	Atomic Number	Density (g/cm^3)	Band gap Eg (eV)	Electron mobility (cm^2/Vs)	Hole mobility (cm^2/Vs)
Si	14	2.33	1.12	1400	480
$Cd_{0.9}Zn_{0.1}Te$	48,30,52	5.78	1.572	1000	50

Fig. 3.74: The geometry of CZT (left top), plot of weighting potential in anode and outer guard ring (right top) and the table which have physical characteristics of Si and CZT.

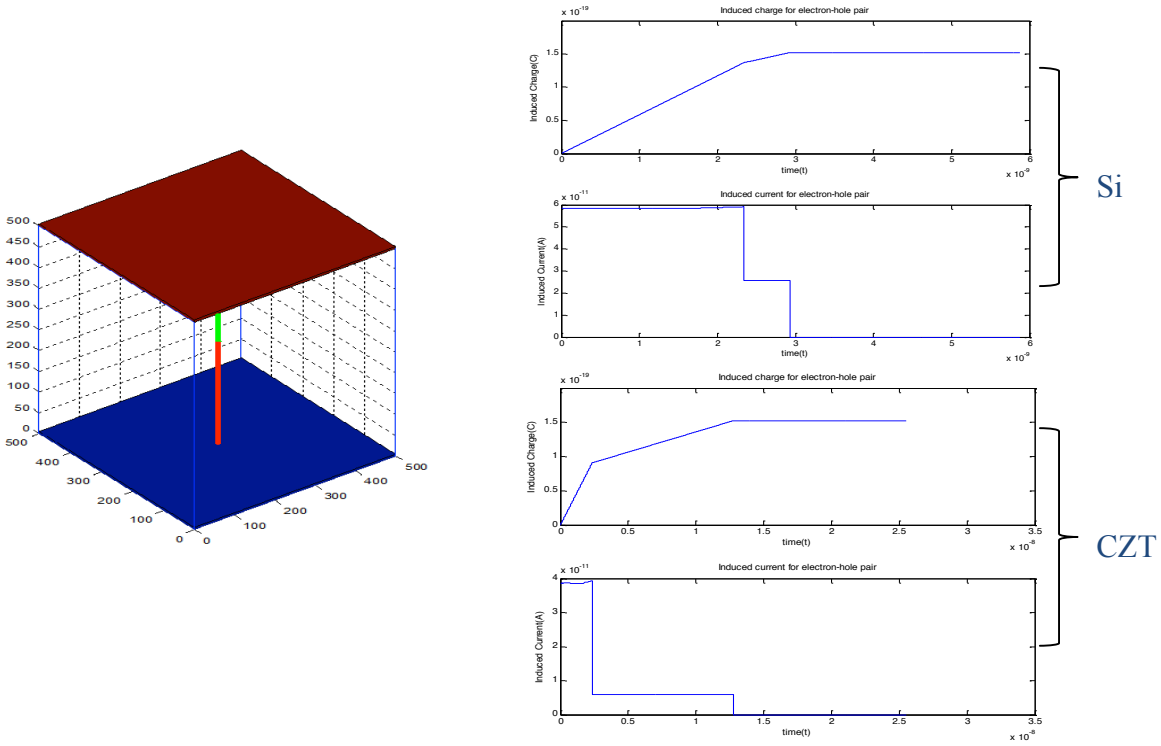


Fig. 3.75: The simulation results with using the planar geometry for Si and CZT.

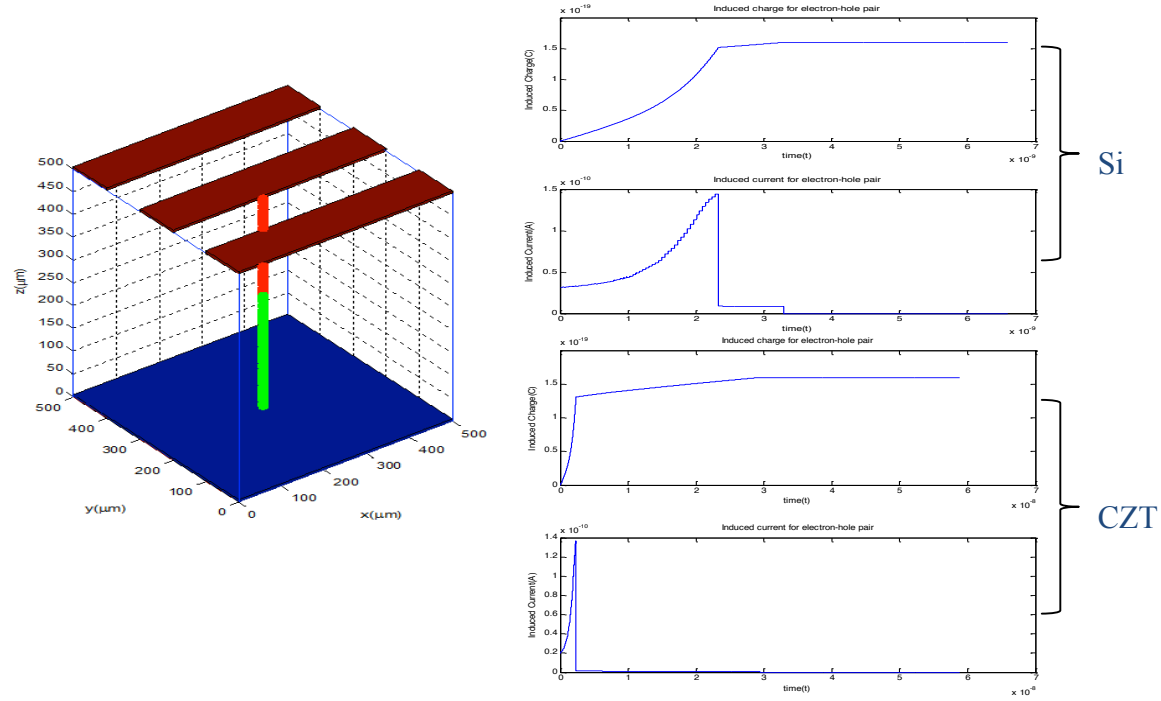


Fig. 3.76: The simulation results with using the strip geometry for Si and CZT.

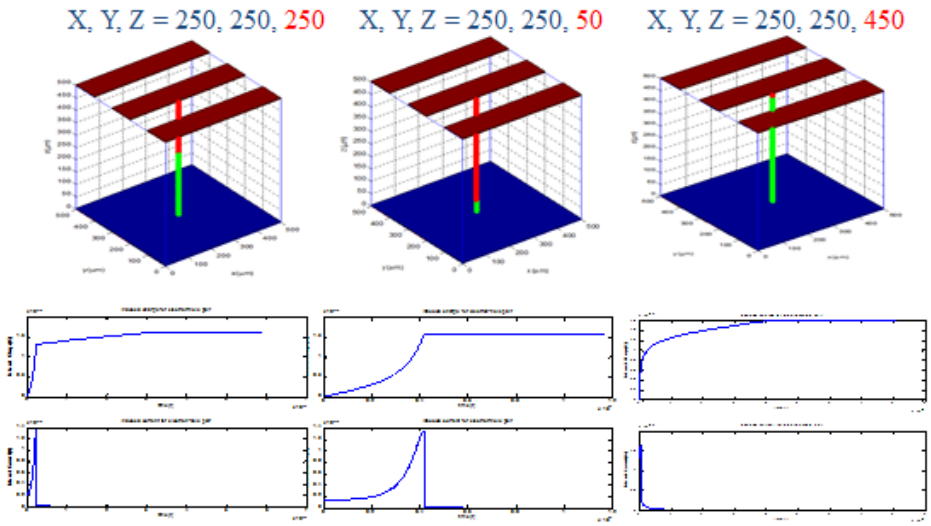


Fig. 3.77: The simulation comparison with changing the interaction position of z-axis from 50 μm to 450 μm.

Achieving low noise on CZT is much less challenging than on silicon because of the relatively large band-gap and less sensitivity to the surface properties. Furthermore, the poor hole transport that is the bane to the standard application of CZT for detection and imaging applications can be beneficial for the delay-line geometry because the hole collection would not participate in the signal formation and therefore complicate the extraction of the initial charge position and number information. Recall that the hole and electron signals provide two measures of these desired physical quantities and only one is required though the additional information can, in principle, be used to refine the estimate. We thus wish to implement the delay-line on CZT in order to test the ease with which one can apply the architecture to alternative semiconductor materials.

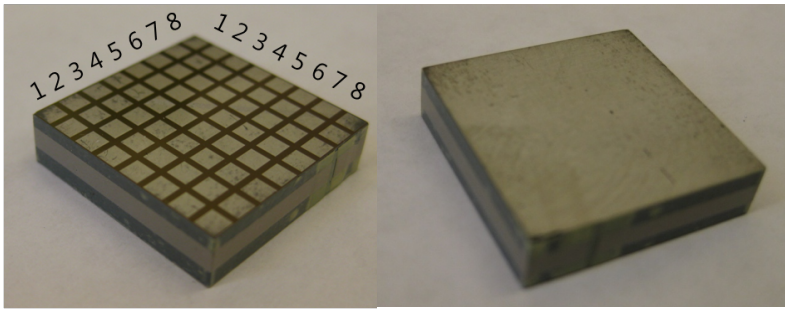
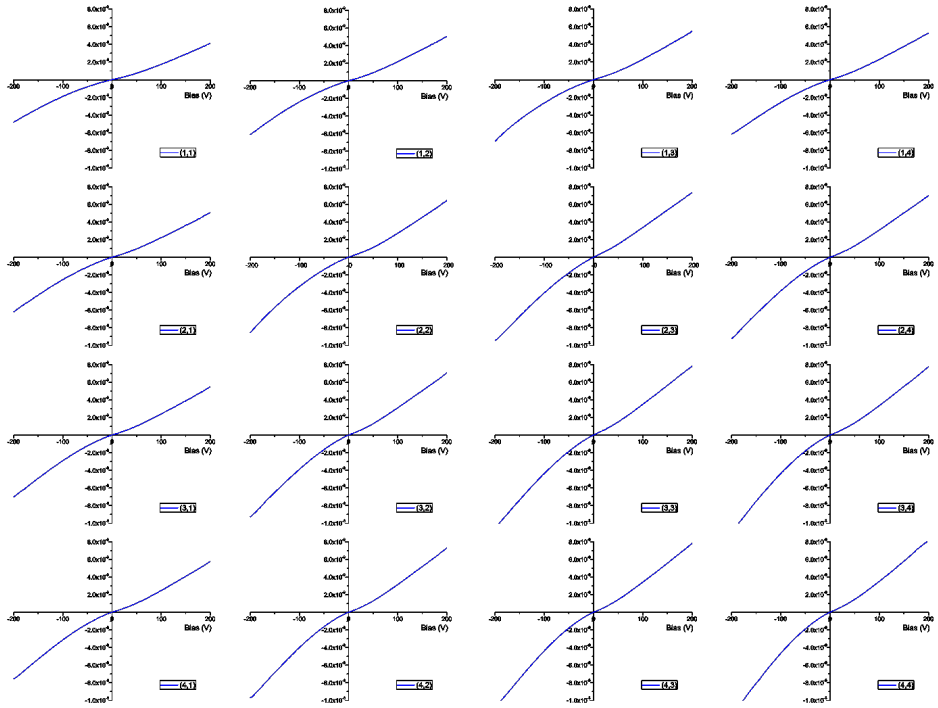


Fig. 3.78: Picture of the CZT detector, front side and back side.

To that end, we purchased several $Cd_{0.9}Zn_{0.1}Te$ (CZT) crystals (Serial Number H83717-2 B2-PL1) from ORBOTECH (bronze grade), and characterized the devices using the Keithley 4200-SCS device. As shown in Fig. 3.82, the detector, with guard-rings on the side walls, was bounded by pixellated indium contacts on the anode side and a planar cathode, processed on a 19.44 mm × 19.44 mm × 5 mm CZT crystal. As shown in the example grid of Fig. 3.79, IV curves from -200 V through 200 V and the capacitance value of each pixel were acquired for all of the 64 pixels. Referring to the face where guard ring strips are disconnected, (1,1) was assigned on the uppermost left pixel and pixel number was assigned as shown in Fig. 3.79.

Since the bulk resistivity dominates the current behavior, the IV curve does not show dominant rectifying behavior. We measured the noise behavior in terms of leakage current as well as the spectral response in order to map the behavior of the crystal and the contacts, so that upon reprocessing, comparisons with the baseline response can be made.



Pixel (1,1) – (4,4)

Fig. 3.79: Current-voltage (I-V) curves for example pixel of the CZT detector.

Leakage current at -200 V

	1	2	3	4	5	6	7	8
1	-4.78	-6.11	-6.96	-6.15	-6.20	-6.09	-6.32	-4.91
2	-6.19	-8.57	-9.43	-9.24	-8.41	-8.67	-7.80	-5.80
3	-7.04	-9.27	-10.43	-10.76	-9.68	-9.02	-9.17	-7.04
4	-7.55	-9.70	-10.40	-11.48	-10.41	-9.66	-9.24	-6.81
5	-7.63	-9.33	-9.85	-10.54	-10.53	-10.22	-8.63	-6.57
6	-8.31	-9.43	-10.00	-10.15	-10.98	-10.13	-9.03	-6.84
7	-7.95	-10.10	-10.29	-9.53	-9.76	-9.58	-8.84	-6.56
8	-6.59	-7.44	-8.03	-7.79	-7.84	-7.93	-7.46	-5.68

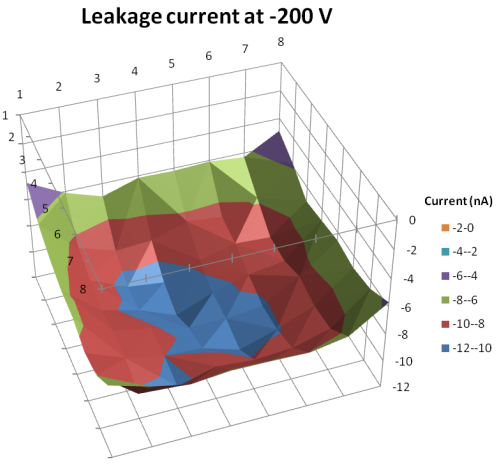


Fig. 3.80: Leakage current of each pixel on the CZT detector at -200 V bias applied on the cathode.

From Fig. 3.79, note that although the curves have the same general shape, the steepness of the curves did vary, and the measured leakage current value at 200 V on the anode side was somewhat smaller than that of -200 V bias applied, by factor of ~0.85. Leakage current values at -200 V and 200 V on the anode side was organized as a form of matrix and plotted, as shown in Figs. 3.80 and 3.81. Note that the leakage current was higher in the middle pixels compared to the outside and edge pixels because of the greater effective volume contributing to the current for the interior electrodes.

Leakage current at 200 V

	1	2	3	4	5	6	7	8
1	4.15	5.05	5.50	5.30	5.37	5.18	5.61	4.57
2	5.09	6.46	7.36	7.00	6.75	6.95	6.81	5.51
3	5.49	7.07	7.83	7.78	7.33	7.56	7.79	5.86
4	5.79	7.32	7.82	8.18	7.59	7.56	7.42	5.94
5	5.80	7.18	7.20	7.89	7.39	7.32	6.95	5.69
6	6.17	6.97	7.63	7.59	7.60	7.52	6.64	5.47
7	5.50	6.53	7.36	7.33	7.25	7.18	6.58	5.18
8	4.19	5.33	6.03	6.13	5.85	6.26	5.73	4.31

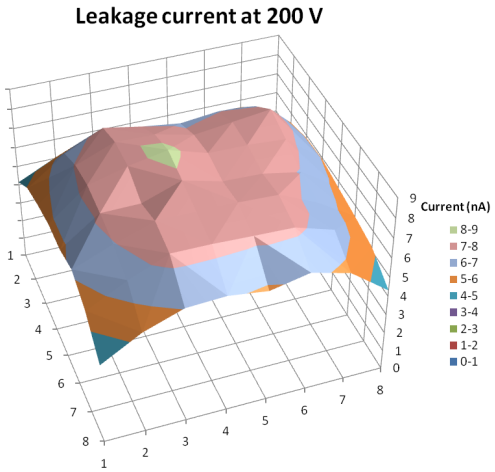


Fig. 3.81: Leakage current of each pixel on the CZT detector at 200 V bias applied on the anode.

In order to further quantify the performance of the CZT detector related to the measured parameters for each pixel above, the spectra derived from standard gamma sources (^{137}Cs, principally) and ^{241}Am, for all 64 pixels was measured. For the alpha particle measurements, the source was impinged upon the cathode side and the electron motion was thus across the bulk of the semiconductor. The movement of the peaks thus gives a measure of the charge trapping in the detector.

For the gamma-ray measurements, a 1000 V bias was applied, and each pixel was read out using the circuit, as illustrated in Fig. 3.82 and realized in Fig. 3.83. The Fig. 3.83 PCB board was designed to apply the same potential on all 64 pixels at the same time and extract the response from the pixel of interest, in which 64 switches were included to distinguish the collecting pixel from the uncollecting pixels. RC filters were also included in order to stabilize the bias, and the PCB board was attached to the anode pixel using a conductive epoxy.

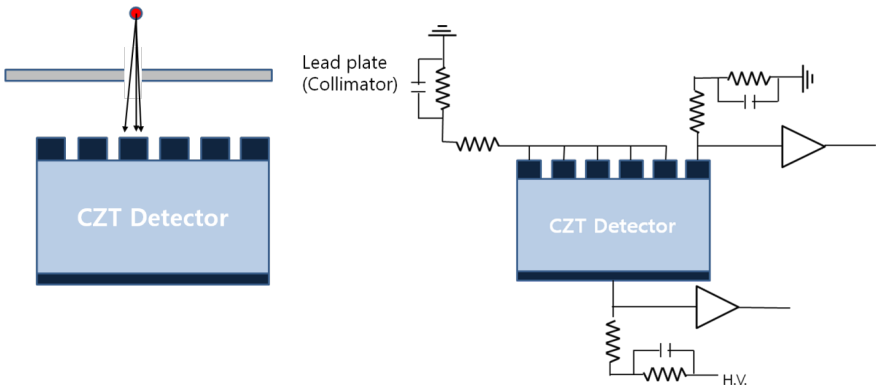


Fig. 3.82: Schematic of the setup used to acquire the pixel-by-pixel spectrum.

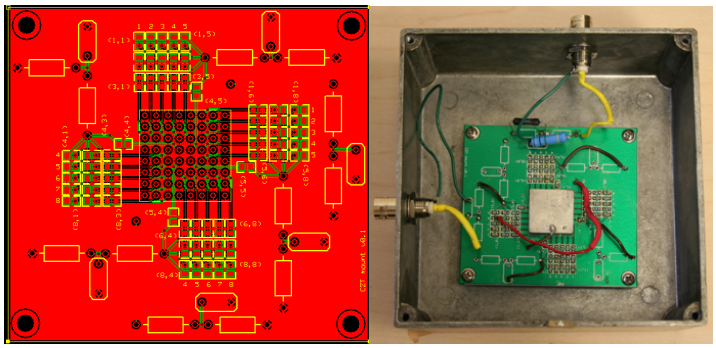


Fig. 3.83: PCB design for the CZT detector anode side, in which the central grid is bonded to the CZT pixel grid and the connections allow one to vary the biasing and readout across the pixel array.

Note that the single-polarity charge sensing is imperfect for the Fig. 3.83 design, in the sense that the weighting potential is not particularly sharp near the anode grid due to the large pixel size; thus, there is additional spread in the photopeak due to this non-uniform depth response as well as the variation due to the leakage current and increased electronic noise for the higher capacitance middle regions. The spectral responses to the impingement of 662 keV gamma-rays from ^{137}Cs are shown in Fig. 3.84.

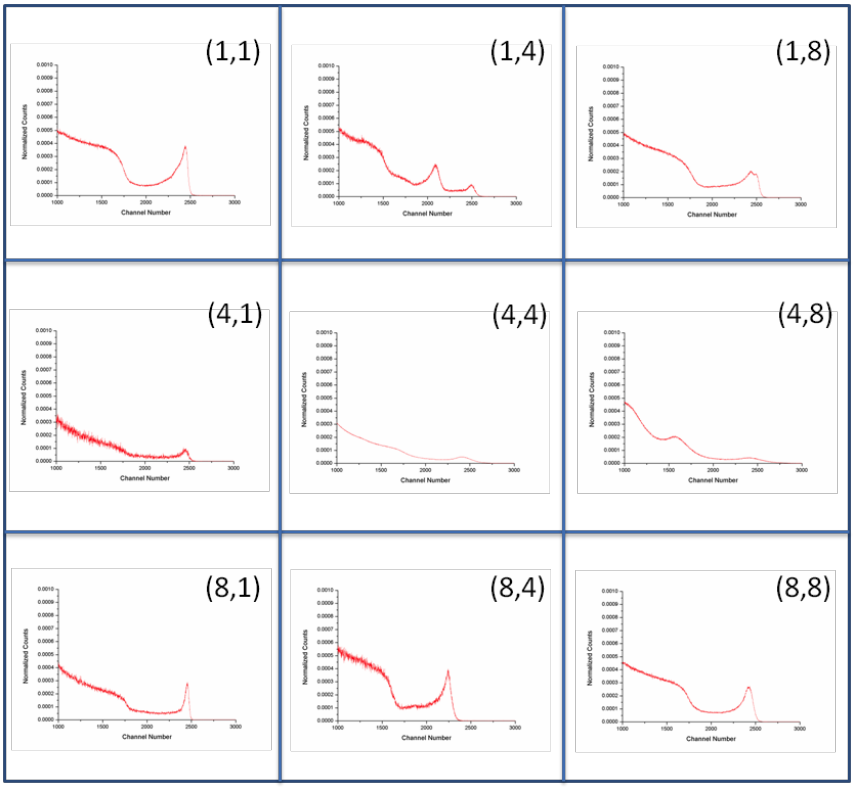
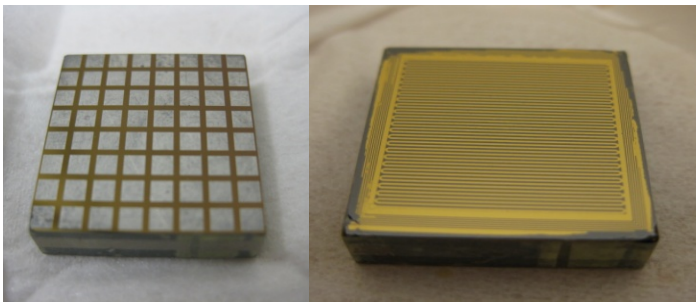


Fig. 3.84: Experimental setup and the acquired gamma-ray spectrum for Cs-137 source.



(a) (b)

Fig. 3.85: The CZT detector from ORBOTECH (a) the pixellated side and (b) Delay line electrode with guard ring structure pattern fabricated on the CZT crystal.

After this baseline testing, we diced the detector into quadrants, polished off the original contacts, and implemented our own contacts. We were able to produce detectors with lower noise than that provided by the commercial vendor, particularly with evaporated In (anode), evaporated Au (cathode) contacts and evaporated Au (anode), evaporated Zn (cathode) contacts. A meander on one of the 1 x 1 cm crystals are worked on processing as shown at Fig. 3.85.

With fabrication process steps like Fig. 3.83, a delay-line CZT detector with low leakage current can be fabricated. Fig. 3.85 shows the CZT detector from ORBOTECH (a) the pixellated side and (b) delay-line electrode with guard ring structure pattern fabricated on the CZT crystal with the process steps like Fig. 3.87. In addition, Fig. 3.86 shows the I-V characteristic curves of (a) original 8 × 8 pixel CZT detector and (b) reprocessed CZT crystals with different metal contacts and (c) delay line pattern on the CZT crystal.

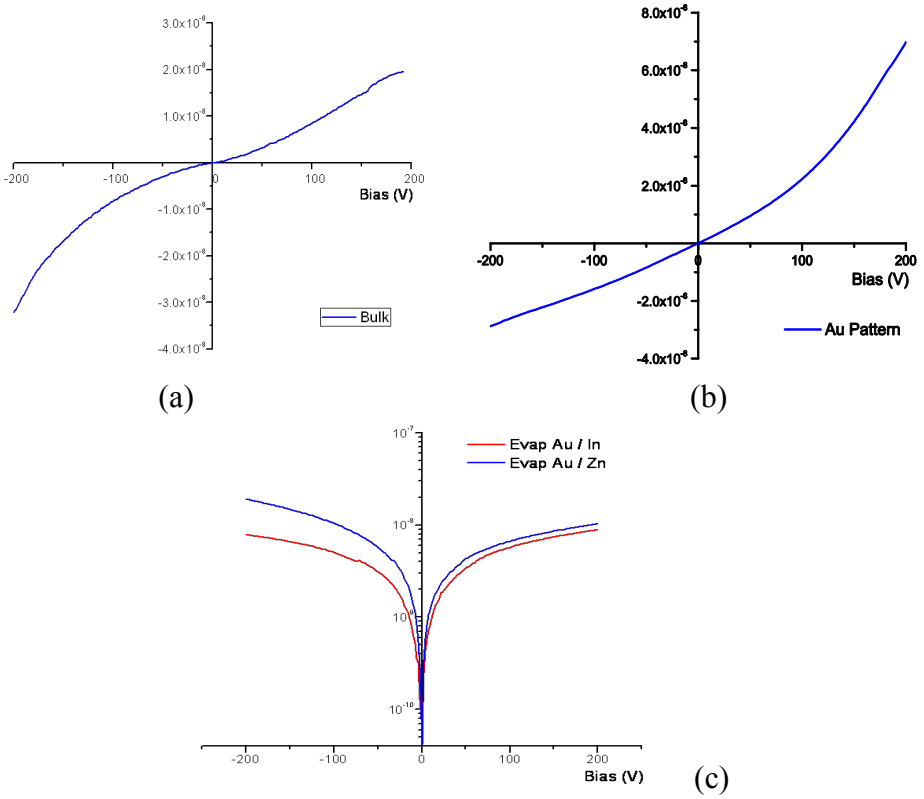


Fig. 3.86: I-V characteristic curves of (a) original 8 × 8 pixel CZT detector and (b) reprocessed CZT crystals with different metal contacts and (c) delay line pattern on the CZT crystal.

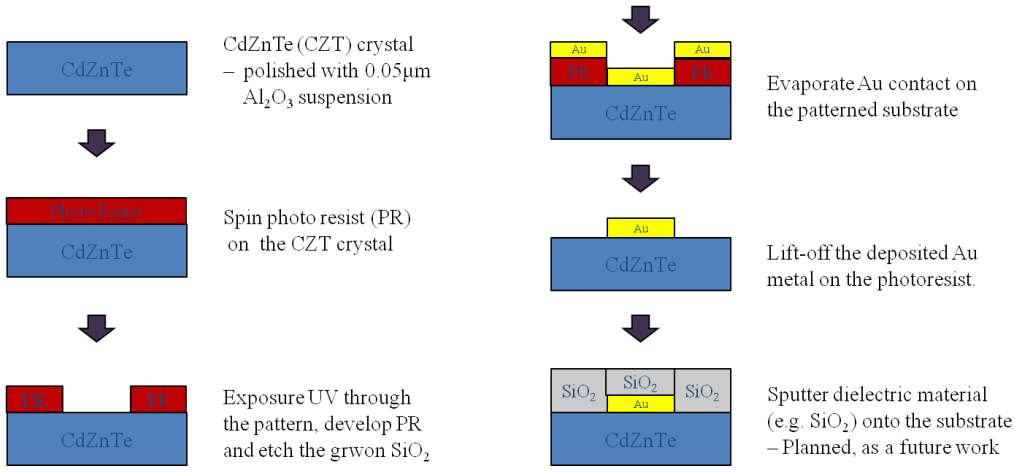


Fig. 3.87: Illustration of serpentine delay-line electrode fabrication process onto the CZT crystal.

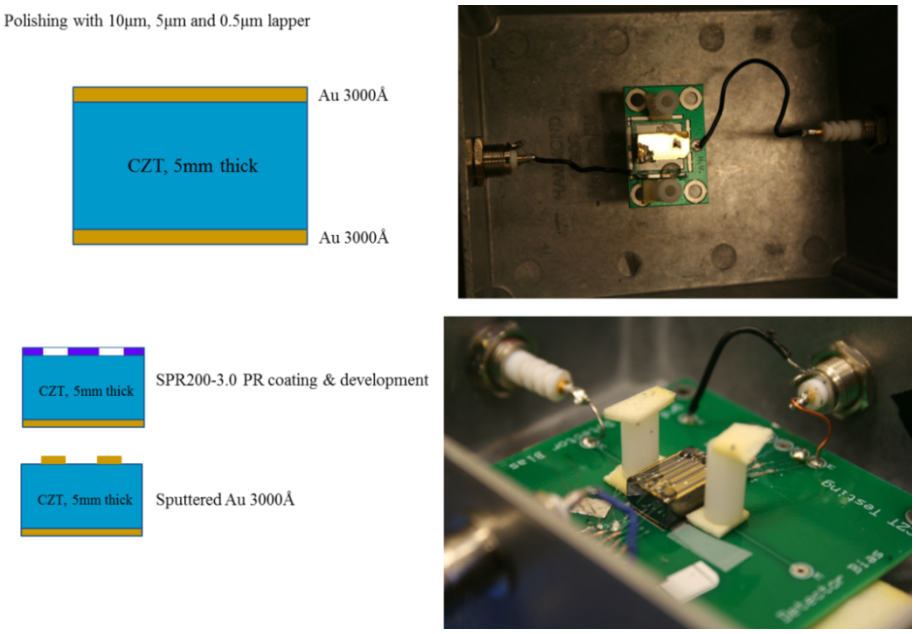


Fig. 3.88: CZT cross-section and fabricated CZT with planar and test meander line (1, 3, 5, 7 and 9 turns)

To validate the meander line for gamma rays, we fabricated CZT with different turn number of meander with 3000 Å thick gold, as shown in Fig. 3.88. We removed the coated material on the CZT with lapping and polishing, and deposited Au with the designed meander line pattern. During the photoresist (PR) coating, the thickness uniformity was poor and the development of PR was compromised. However, we found the proper balance between the PR thickness and the metal thickness for CZT and finally we successfully deposited meander line Au on CZT, as shown in Fig. 3.88. We used

SPR200-3.0 PR with spinning 4000 rpm, 6 seconds exposure time and 2 min. development time.

3.9 Production good performance CZT Detectors

Parallel to the silicon work, we fabricated both: (a) planar-type CZT with multiple floating guard-rings, and (b) meander-line CZT devices, such as those shown in Figs. 3.91 and 3.92 below. Before the process, we lapped and polished the crystals with 10 μm, 5 μm and 0.5 μm grain sizes of lapping and polishing papers. After polishing, we rinsed the bulk CZT with IPA and methanol for 5 min, after which we performed a Br-methanol (1%) etch for 1 minute. Fig. 3.89 shows microscopic pictures of the CZT surface after mechanical polishing and after Br-methanol etching, exhibiting the obvious improvement accrued from using chemical polishing.

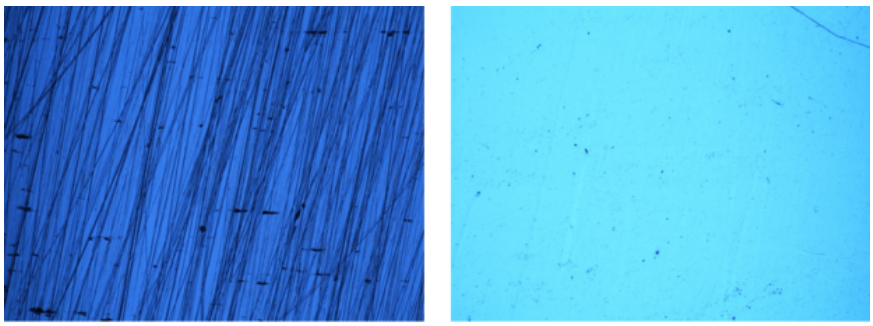


Fig. 3.89: Pictures of CZT surface after mechanical polishing (left) and Br-Methanol etching (right).

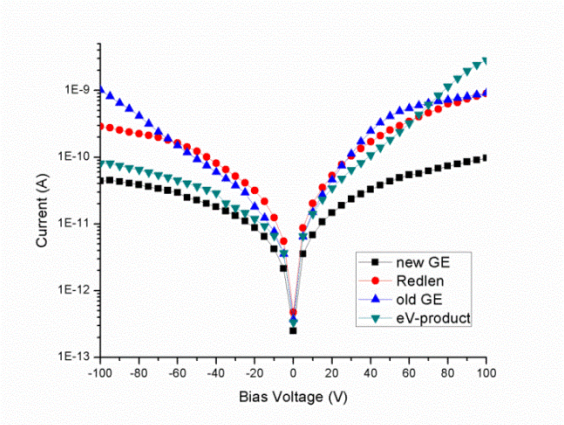


Fig. 3.90: Leakage current against reverse bias voltage with bare CZT from different venders.

Fig. 3.90 shows the I-V curves of the bulk CZT derived from our various samples, provided by either Redlen (via GE) or eV-product (via Orbitech) material. As shown in the plots, the leakage current improves as the defect level of the material is reduced. The mechanical grade CZT ("old GE") has the highest leakage, the spectroscopic grade Redlen material ("new GE") has the lowest leakage, and the eV-product material is intermediate between the two. In absolute terms, the leakage current is quite good because at -100 V, the value is ~35 pA for the nominal 5 mm thick sample, which is much lower than that produced by the samples upon the receipt from the vendors and about 10 – 1000 times smaller than that produced by silicon, which is 10 times thinner. Thus, the imperfections in the crystal are not preventing us from taking advantage of the wide band-gap of the material. The consequence of this low leakage is that the signals are substantially above the noise for both ions and gamma-rays, and one can therefore extract the charge information well, relative to the electronic noise. Of course, the poor hole transport and electron trapping still spreads the spectrum unless it is properly corrected.

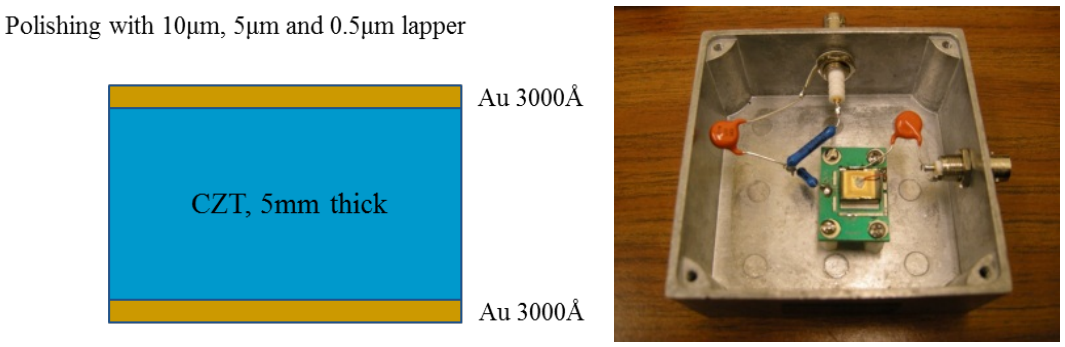


Fig. 3.91: Planar type CZT detector with multiple guard-rings.

With the polished CZT crystals, we deposited photoresist SPR-200 which has maximum thickness of 3 μm, which requires a 6 sec UV exposure time and 1 min. development time. Before the metallization, we performed an ashing run for removing the remaining photoresist. We then used an indium (In) deposition of 3000Å thick, which was patterned using lift-off, in which a 30 min. acetone dip was used for PR removal. N_2 annealing for 10 min. at 150 °C was used.

After metallization, we did a passivation with H_2O_2 and NH_4F (wt 1% each) for 10 min. to make an oxide on the CZT surface. Several different metallizations were investigated, including Au, In, Cr/Au, In/Ag, In/Au and Ti/Au, mainly to find the recipe that yielded the best adhesion, which facilitates the patterning of small features. When the evaporator was operated at the normal deposition rate (10 Å/s), the entire metal was taken off; however, when we reduced the deposition rate to 3 or 5 Å/second, finer grain structure was produced and adhesion between the CZT and the metal were good, and we can successfully pattern the meander, as the 20 and 25 μm width patterns of Fig. 3.92 show.

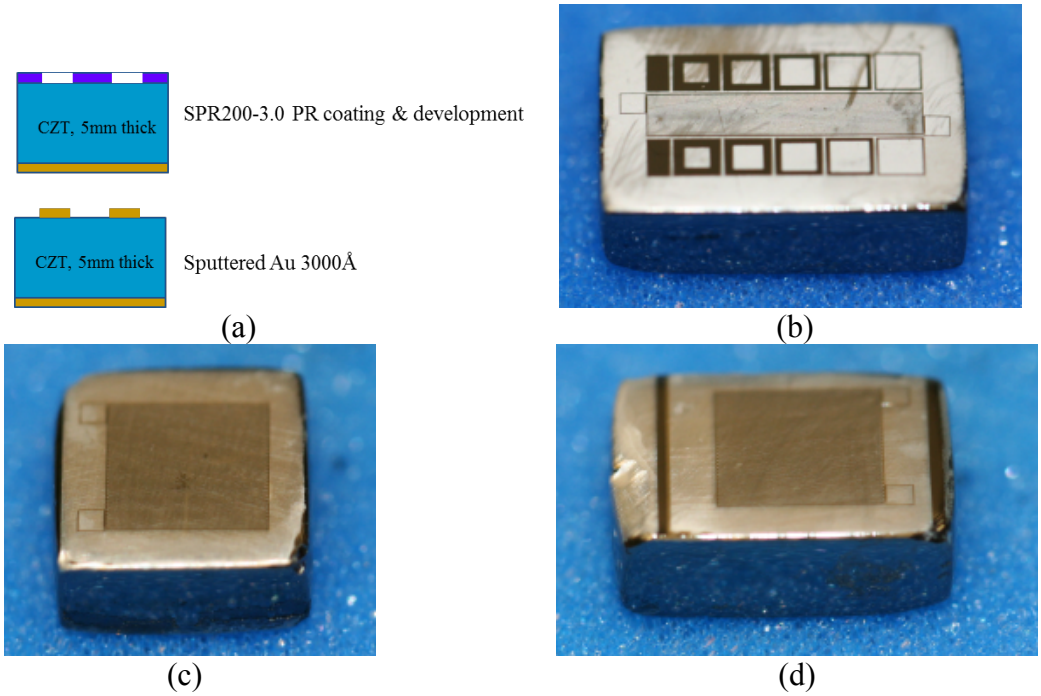


Fig. 3.92: (a) Simplified version of process used to fabricate meander-line CZT detectors. (b) Derived from spectroscopic-grade GE CZT, fabricated detector, which has test-pixel pattern. (c) 20 μm width meander, derived from eV-product CZT, and (d) 25 μm width meander detector from GE mechanical grade CZT.

3.10 References

[1] M. Jeong, D. A. Lawlor, and M. D. Hammig, “Optimization of the Position Resolution in Semiconductor Detectors”, IEEE Nuclear Science Symposium Conference Record, pp. 1456-1462, 2007.

[2] M.D. Hammig, M. Jeong, D.K.Wehe, and S. Ramadoss, “Delay-line electrode partitioning as a means of simplified position sensing”, IEEE Nuclear Science Symposium Conference Record, pp. 2493-2500, 2008.

[3] M. Jeong, S. Ramadoss, and M.D. Hammig, “Fabrication and Signal Readout of the Si-based Delay-Line Radiation Detector”, IEEE Nuclear Science Symposium Conference Record, pp. 2493-2500, 2009.

[4] Craig J.Scott, “Transmission Characteristics of Narrow Line-width Interconnections on Silicon Substrates”, VMIC Conference, pp. 426-428, 1989.

[5] G. F. Knoll, Radiation Detection and Measurement, 3rd ed, New York: John Wiley & Sons, Inc., 2000.

[6] Z. He, “Review of the Shockley-Ramo theorem and its application in semiconductor gamma ray detectors,” Nuclear Instruments and Methods in Physics Research Section A, vol. 463, pp. 250-267, 2001.

[7] Sentarus Device, Synopsys Inc., June 2006.

[8] P. A. Tove, "Methods of avoiding edge effects on semiconductor diodes," J. Phys. D: Appl. Phys., pp.517-536, 1982.

[9] Hulmuth Spieler, “Semiconductor Detector Systems,” Oxford, Aug. 25, 2005.

CHAPTER IV

TRACKING RECONSTRUCTION ALGORITHM

One factor that restricts the ubiquitous deployment of radiation imagers is the high cost of the existing instruments. One approach to overcoming this restriction is to develop more inexpensive detector materials; however, this will, in general require a new material modality, such as nanocrystalline semiconductors. An alternative approach is to make much greater use of the information that exists in the radiation field. For instance, if the imaging resolution can be improved by orders of magnitude, then smaller more economical detectors can be used. This research has been focused on enabling this improvement by including the directional information of the products of a radiation's interaction. However, we have not previously quantified the degree of improvement that can be achieved, and further, how that enhancement depends on the detector's characteristics, the most notable being its spatial resolution. In Chapter IV, the simulated results are summarized and we find that for even detectors of moderate spatial resolution, the gamma-ray direction can be highly localized.

Specifically, if the detector has 100 micron spatial resolution, then the direction of the recoil-electron can be confined to 10^0 or better, for energies greater than 300 keV, compared with a simple back-projection onto a 360^0 cone for Compton-scatter events. This same performance can be achieved with a 10 micron design down to a few 10's of keV recoil-electron energy; that is, over the entire energy range of interest. If the detector has 1 micron or better spatial resolution, then the Compton back-projected cone can be constricted to a 1^0 arc-segment or better, two or three orders of magnitude better than simple or filtered back-projected techniques. The quantitative studies indicate therefore that recoil-electron tracking is thus more effective than we initially anticipated in improving the imaging performance of gamma-cameras. That is, *a priori* one might have

had needed very high resolution (0.1 microns, say) in order to effectively determine the gamma-ray direction because of the circuitous nature of the electron tracks; however, we find that moderate resolution designs are sufficient for orders of magnitude improvement in the information utilization.

Our technological focus has been on tracking the products of radiation interaction in the semiconductor-detector volume; however, that information must be included in the gamma-ray (or neutron) image in application use. Therefore we spent effort in investigating the means of forming an image most efficiently. That is, given the usual deposited energy-position information, the software routines were developed to form the back-projected image. The recoil-product information also is included in order to clarify the gain that accrues from the methodology under development. The development of the gamma-ray imaging routines is described in this chapter.

4.1 The Reconstruction of the Initial Electron-Hole Track

The primary focus has been on the suitable extraction of the induced current (or induced charge) pulses that follow radiation interactions. As detailed in Chapters II and III, in this empirical work, the circuit components and their electrical coupling must be optimized in order to take best advantage of the amplification components. Given the successful extraction of those profiles, the algorithm was developed that will transform the current signals into spatial charge distributions. As shown in the example of Fig. 4.1, this manipulation is nontrivial.

In Fig. 4.1a, a meandering delay-line is patterned on the surface or surfaces of the silicon detector. When a gamma-ray or neutron interacts in the medium, via the creation and subsequent transport of an electron or an ion, respectively, an electron-hole track results, one example of which is shown in Fig. 4.1b for a 500 keV electron following a gamma-ray interaction (478 keV is the backscatter electron energy from a 662 keV gamma-ray Compton scatter, and is thus a typical energy that is expected to be observed both during development and subsequent deployment). As shown in Fig. 4.1b, the e-h track extends several millimeters in both lateral dimensions, and it nearly extends beyond the detector volume. The scale of the strip-pitch in the schematic Fig. 4.1a drawing thus corresponds to several millimeters, although the focus for our development detector is on 100 μm and 10 μm lines.

For a strip pitch of 20 μm (10 μm gap and 10 μm strip width), the pulses in Fig. 4.1c result, with some caveats. The current pulses in Fig. 4.1c are measured with 10 ps time steps, 500 strips compose the pattern, and the carriers are assumed to be saturated. For the sake of the exposition, we assume that the anode base is un-patterned, and we neglect the pinching of the electric field near the cathode. For small strip dimensions, the current pulses have a spike at the end of each current step corresponding to the drift and collection of each charge carrier. This effect doesn't modify the results discussed in this Section and is thus neglected.

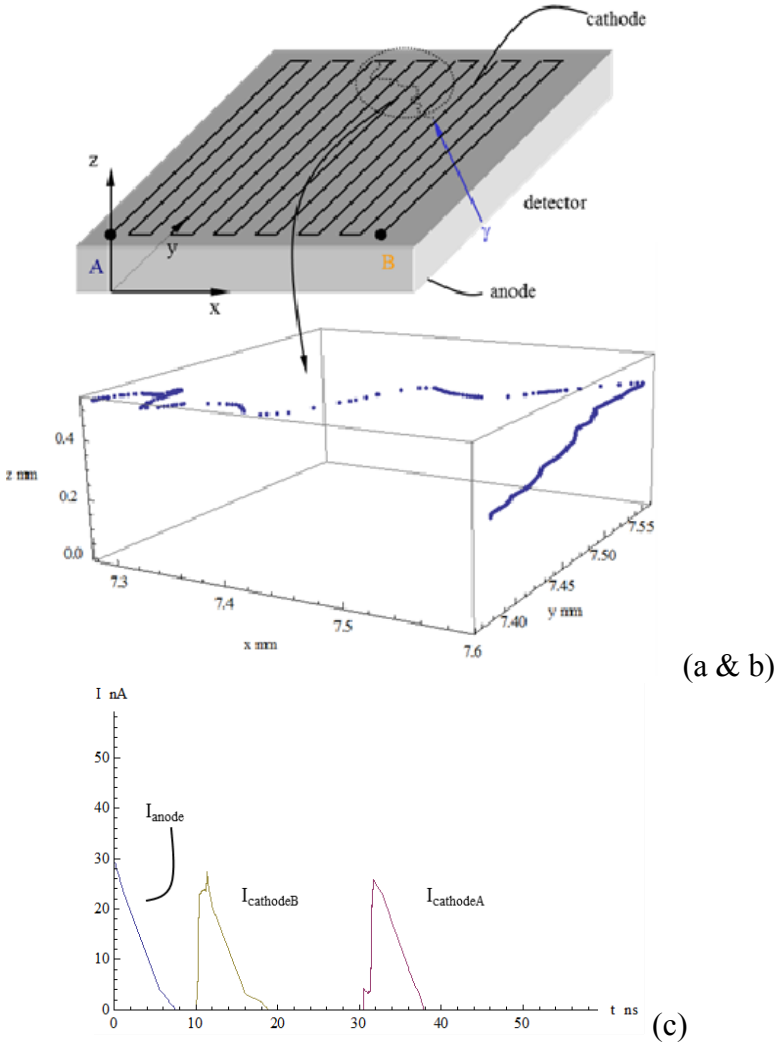


Fig. 4.1: (a) Schematic of delay-line detector (1 x 1 cm, 550 μm thick) with a gamma-ray interaction at ¾ of the detector's lateral length and width. (b) A typical meandering electron track in silicon, in this case, for a 500 keV electron created at the silicon surface and initially traveling normal to that surface. (c) The resulting induced current signals, for a planar anode and a delay-line cathode (with readouts at A and B).

One of challenges is to thus transform the current pulses in Fig. 4.1c into the e-h tracks of Fig. 4.1b. Once that track is mapped then the extraction of the recoil-electron direction can be achieved with marked improvements in the image quality. In this Section, the complicating effect of the delay-line structure on the current pulses will first be discussed, and secondly, the degree will be simulated to which the reconstructed track geometry is affected by variations in the detector parameters, such as the timing and positional resolutions.

4.1.1 Current-pulse Profiles through a Delay-Line Electrode

Recall that the current and/or charge induced on a conductor can be derived from the Shockley-Ramo Theorem. The weighting potentials and the saturated drift velocities are used to find the corresponding current pulses. Again, the sharp peak is neglected that occurs during hole collection at the partitioned cathode. For instance, consider Fig. 4.2, in which a single e-h pair is created at ¾ of the lateral width and length of the Fig. 4.1 detector, and adjacent to the cathode surface (i.e. at $z = 0$). As shown in Fig. 4.2a, a) only hole drifts, b) the width of any pulse provides a measure of the charge's depth, (c) the time delay between any two of the three pulses provides a measure of the position along the delay line, and (d) the current magnitude for these single charges is 29 pA, which prevents the use of some amplification components, if the sensing of a small number of carriers is desired.

Recall that the pass-band of the transmission line depends on the line's resistance, and is a function of frequency. If the signals that emerge from a wide-band line are high-passed, then only the sharp transitions will survive signal transit along the line, resulting in cathode pulses like those shown in Fig. 4.2b. As in the current pulses, the important timing information is preserved, and the collection position can be discerned. Recall that even if the line is made to be wide-band, the current characteristic provides a direct measure of the charge collected at a particular time, and thus, a particular position; therefore, one can take the derivative of the current pulse (or the double derivative of the charge pulse) in order to map the charge distribution. For instance, consider Fig. 4.3 in which the e-h pair of Fig. 4.2 is moved to the middle of the detector's thickness and both the electron and hole contribute to the pulse shape.

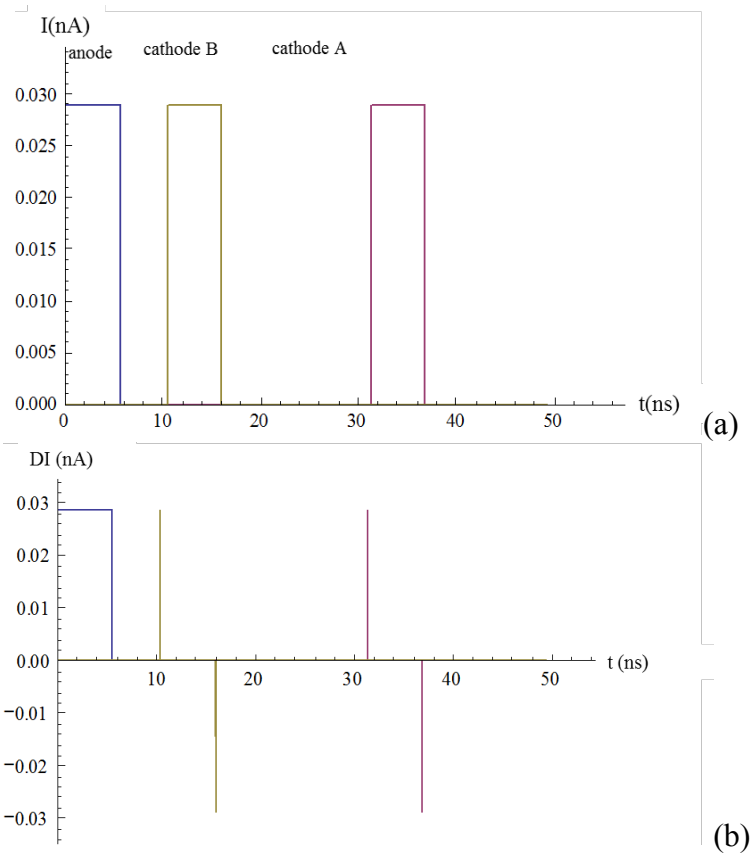


Fig. 4.2: (a) For a wide-band transmission line and amplifier, the current pulses from the planar anode and the ends of the delay-line cathode, for an e-h pair created adjacent to the cathode; thus, the current depends only on hole drift. Ten picosecond time steps are used. (b) The anode signal along with the change in the cathode signals, as can be produced by high-pass filtering.

The large positive current peak at the beginning of each pulse corresponds to the charge creation and charge acceleration stages, and as explained in previous Chapters, its width is sub-ps in general. The negative current excursions correspond to charge collection in the conducting electrodes, and those time widths are sub-ps as well (~0.2 ps for a saturated electron into aluminum), as estimated by the electron wave-packet penetration through the skin depth of the conductor. Furthermore, the bottom part of Fig. 4.3 shows that the electron and hole collection provide current peaks of different size because of their disparate drift velocities, and one can therefore identify the species collected by the transition strength, in principle.

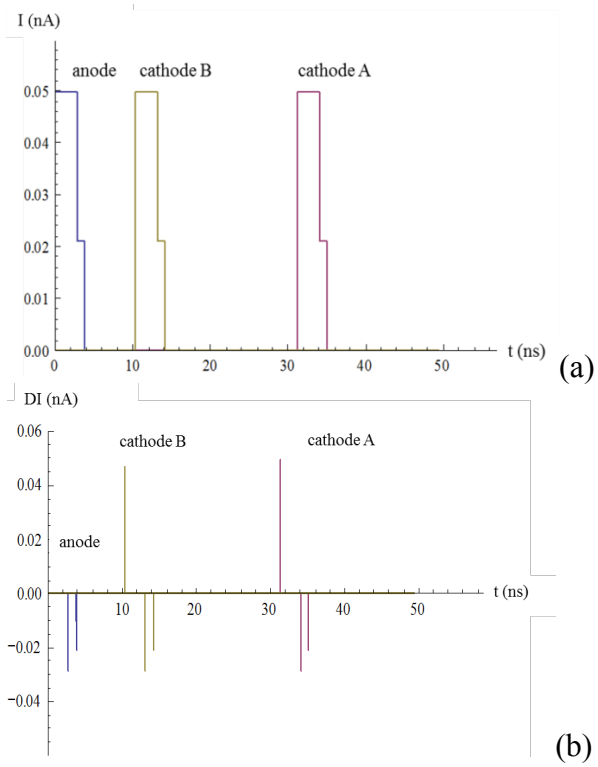


Fig. 4.3: Current (top) and current-slope profiles (bottom) for a single e-h pair created in the middle of the detector, in depth (t = 10 ps).

Note that the desired information is over-sampled in Fig. 4.3 by a factor of 6. That is, the depth of the e-h pair can be determined from the width of any one of the three signals, and further, from the width of one of the two charges' collection time. Similarly, the lateral position needs only the time difference between two signal-transitions that are sourced in the same material transition (for instance, the collection of the electron in the anode). The extra information can be used to improve the determination of the relevant position, and thus reduce the uncertainty by several times, but the added peaks can obscure transitions when a long continuous charge-train is transiting through the detector bulk.

As shown in Fig. 4.3 therefore, the collection of the charge on a given strip of the meandering electrode results in a copy of the underlying variation traveling to both ends of the line. This does not imply that the cathode A and B signals are always delayed copies of one another however. If the charge distribution is extended, like the Fig. 4.1 example, then the contribution from each strip has its own time stamp and the summation

signal thus varies between the readout pads. Consider the simple example of Fig. 4.4, in which three e-h pairs are placed a z = 0 (so only hole drift contributes), and at the following nonlinear lateral positions, where Lx and Ly are the detector's lateral dimensions:

1) (x,y) = (¾ Lx + 2 pitch, ¾ Lx + 2 pitch)
2) (x,y) = (¾ Lx + 4 pitch, ¾ Lx + 4 pitch)
3) (x,y) = (¾ Lx + 9 pitch, ¾ Lx + 9 pitch)

The cathode signals are approximate mirror images of each other in this case, a symmetry broken if the depth is varied such that the electron transport also contributes to the signal, as shown in Fig. 4.4 for z = thickness/2. As in Fig. 4.2, the electron collection events produce a larger current transition than the hole collection events. Furthermore, even though there are more pulses, one can still identify the symmetry in the patterns and identify which peaks are coincident and therefore determine the desired positions. For general distributions, the challenge is to identify the coincident peak patterns, whose relative placement depends on physical factors, such as the drift velocity and initial carrier depth relative to the strip length and signal propagation velocity.

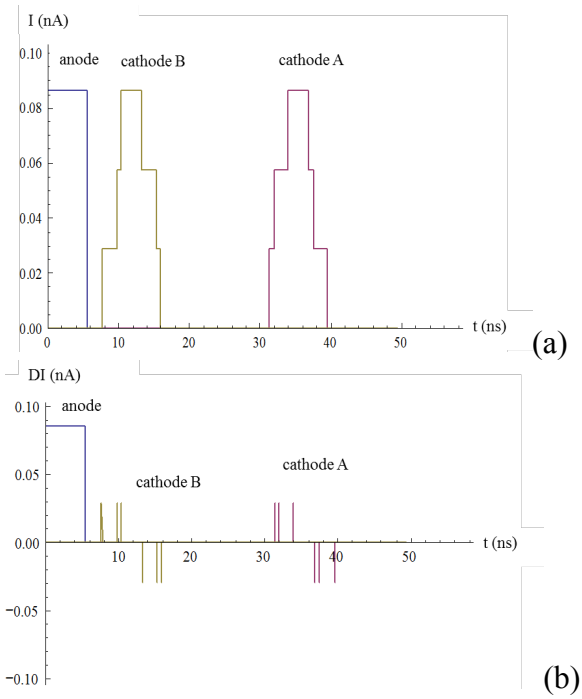


Fig. 4.4: Current (a) and current-slope profiles (b) for three e-h pairs created at z = 0 and at the lateral positions specified in the text (t = 10 ps).

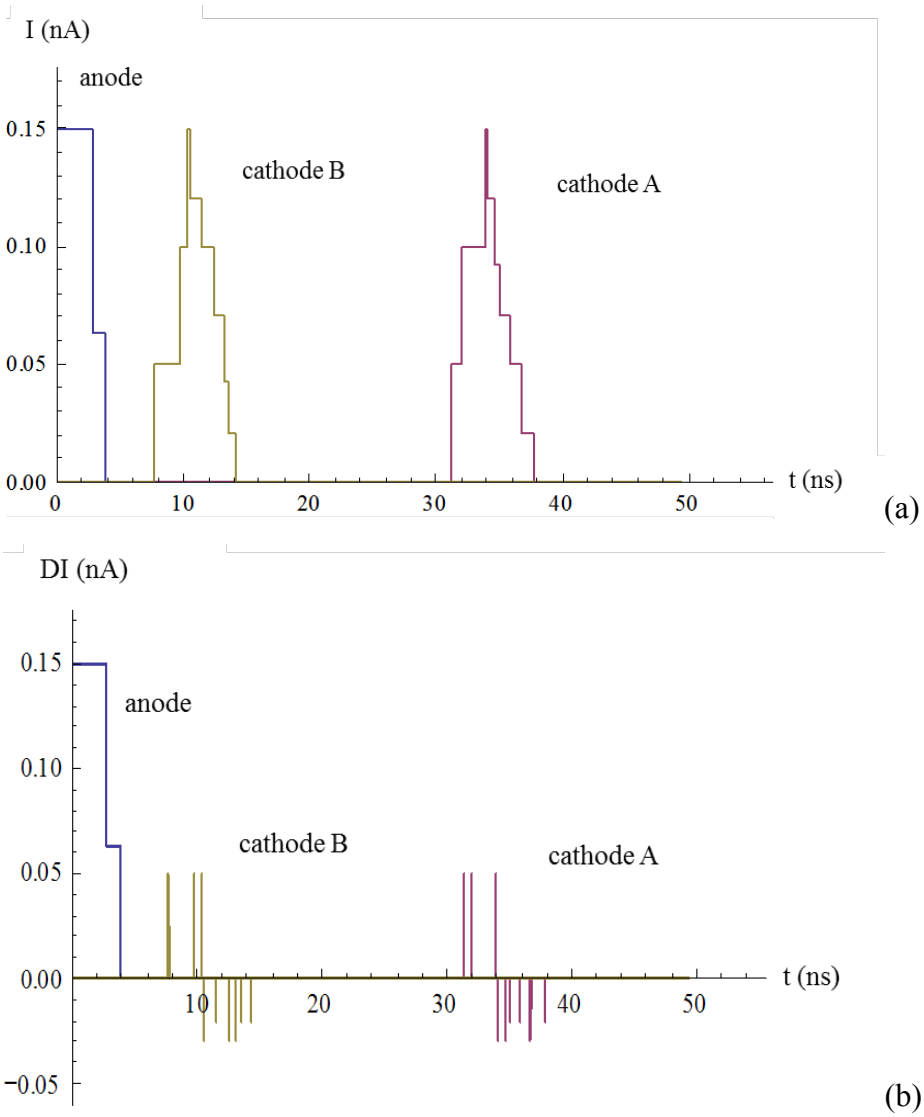


Fig. 4.5: Current (a) and current-slope profiles (b) for three e-h pairs created at z = thickness/2 and at the lateral positions specified in the text (t = 10 ps).

In general, the signals propagate 1000 times faster than the charges drift; thus, even though the carriers move through a smaller gap compared with the propagation length (~250 μm or so compared with a few meters), the induced signal from early collection events can coincide with the collection of later carriers along the delay-line. For realistic interactions, one can take advantage of the fact that the charge-pair distributions will be continuous, rather than random. Let us consider some e-h distributions simulated using GEANT4.

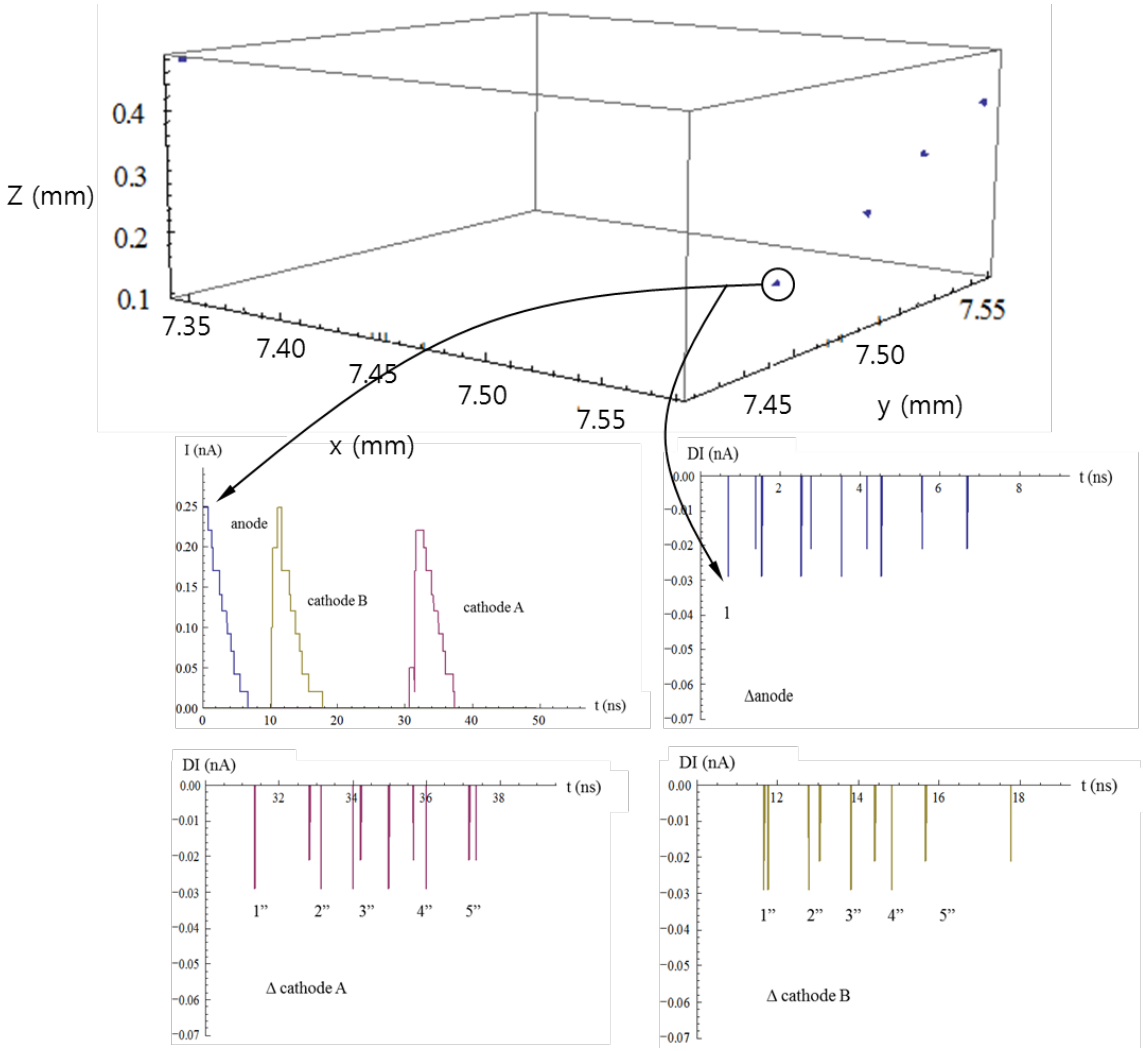


Fig. 4.6: (a) A 500 keV meandering electron track in silicon, created at the silicon surface and initially traveling normal to that surface, sampled at 100 μm depths. (b) Current profiles for the 5 created e-h pairs, as collected by the planar anode, and the meandering strip cathode (500 strips with width and gap of 10 μm). Current-slope profiles for the (c) anode, (d) cathode A, and (e) cathode B signals, sampled with 10 ps resolution.

In Fig. 4.6, the 500 keV electron track of Fig. 4.1 is sampled in 100 μm steps, and the 5 sampled interaction points (in blue) thus vary in some unpredictable manner. The discretely varying anode current in Fig. 4.5b, rather than the step pulses shown in the previous figures, is expected from electron and alpha particle interactions, because the extended charge depositions result in the gradual removal of carriers from the stream, rather than the abrupt removal of a large fraction of carriers.

The slope signals in Figs. 4.6c-e reveal the means by which one can extract the initial positions of the e-h pairs. First, consider Fig. 4.6c, which reveals 10 peaks arriving in two different sizes. The large transitions are due to electron collection, and hole collection produces the smaller peaks. Because the depths are sampled in even 100 μm increments, the peaks have equal separation. To highlight a particular carrier, consider the first peak, which corresponds to the electron collection of the circled charge pair. The anode collection time reveals the depth of charge creation, but the more difficult question is: which of the 5 electron peaks in either of the cathode signals corresponds to the first electron? In general, without any a priori knowledge, any of the five peaks in Figs. 4.6d and 27e (1'-5' or 1''-5'') could correspond to electron collection 1. For the assumed arrangement of Fig. 4.1 however, one can make use of the fact that:

$$t_{transit} = (tn' - t1) + (tn'' - t1), \qquad (4.4)$$

in order to find the appropriate n' and n''. As an example, if a carrier is collected directly adjacent to pad A, then the transit time of the cathode A pulse will be zero and that of the cathode B pulse will be $t_{transit}$, the total one-way transit time of the delay-line. If the carrier is collected at the midpoint of the line, then both signals require $t_{transit}/2$ to read their respective ends, but the total time between charge collection and the readout of both signals remains $t_{transit}$.

Table 4.1: Transit time table, as calculated from Eq. 4.4. The total transit time is equal to 41.73 ns for the delay-line modeled.

	n"				
n'	1	2	3	4	5
1	41.61	**41.73**	42.73	43.78	44.79
2	43.41	43.53	44.53	45.58	46.59
3	44.28	44.40	45.40	46.45	47.46
4	45.25	45.37	46.37	47.42	48.43
5	46.26	46.38	47.38	48.43	49.44

This constraint can be applied to the peaks in Fig. 4.6, as summarized in Table 4.1. Using the peak times and the known transit time (41.73 ns), we see from Table 1 that the corresponding peaks in the cathode pulses are 1', as measured at readout A, and 2'', as measured at readout B. Either time difference between anode peak 1 and the

corresponding cathode peak provides a measure of the position along the transmission line, and thus, the lateral position x.

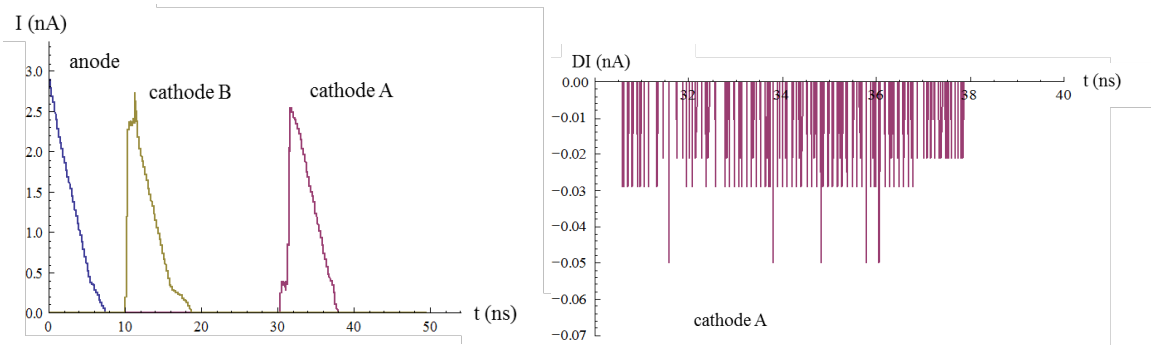


Fig. 4.7: Current profiles for the anode, cathode A, and cathode B signals, sampled with 10 ps resolution (a) and current-slop profiles for the cathode A signals (b).

Each of the additional peaks can be discerned in a similar manner, and the hole peaks provide an extra evaluation criteria to solidify the measurement. In the Fig. 4.6 example, the second anode electron peak corresponds to the 2' and 1'' peaks, and the remaining peaks have equal indices (i.e. anode 3×3' and 3'', 4×4' and 4'', and 5×5' and 5''). One can thus use this general technique to extract the e-h track for more finely sampled data, such as that shown in Fig. 4.7, in which the track is sampled at each micron. For the dense charge clouds typical of radiation interactions, the current-slope peaks can comprise not only steps corresponding to electrons and holes, but also the coincident collection of multiple carriers. These events slightly complicate the algorithm, but it can still successfully extract the charge tracks.

4.1.2 Impact of Imperfections on the Track Reconstruction

As discussed above, the positions of the e-h pairs can be reconstructed from the delay-line electrode geometry, even though the signal shapes are more complicated that than for simple electrode partitions. However, in a realized detector, the timing determination is uncertain, the signals are compounded by noise from thermal and electronic effects, and the geometry may be coarser than that assumed in Chapter II, although the latter is unlikely. In this subsection, the electronic noise is neglected, which is the current focus of our experimental work as discussed in previous section, and the

degree to which the track determination is impaired by imperfect timing and by geometric changes are gauged.

By design, the delay-line approach is robust in the face of relatively poor timing resolution. That is, readout with poor timing resolution (defined here at the $FWHM_t$ of the time spectrum) will result in a poorer determination of the interaction location along the transmission line length. However, the lateral increments along that length are truncated by the meandering design of the line. Specifically, even if the position is uncertain to within several centimeters along the line, the lateral position can be constrained to within a few strip pitches, typical in the 10s of microns range. Nevertheless, as clarified above, we prefer good timing resolution so that the charge centers can be identified.

If the strip pitch is widened, then a substantial deterioration in the track reconstruction can be expected, an expectation realized in Fig. 4.8, again for the 500 keV electron track in silicon. If the positional spacing is 10 μm or greater, or if one doesn't use interdigital pulse height analysis to refine the lateral position, then the discreteness of the grid will dominates as shown in the (10 μm, 100 ps) track reconstruction. Therefore, 1 μm or better spacing is required to map the track with sufficient clarity that the recoil direction can be determined in a substantial fraction of the radiation events of interest for Homeland Security applications.

Fine positional resolution is not enough however, as one can appreciate by closely examining the (1 μm, 1 ns) track reconstruction. With the data shown, one would fit a polynomial to the scatter of points at the beginning of the track; however, this would provide a wholly inaccurate measurement of the recoil electron direction because the electron meanders with must greater frequency than that sampled with the coarse timing resolution. Fortunately, the 100 ps timing does effectively reveal the relevant section of the track however. Thus, the target of 1 μm resolution with our already existing timing resolution is sufficient to extract information from e-h tracks originating from electrons with a few 100 keV energy. For lower energy electrons, extending into the x-ray regime and for shallow angle Compton scattering events, one might require sub-micron positioning, as shown in Fig. 4.9 for a 50 keV electron.

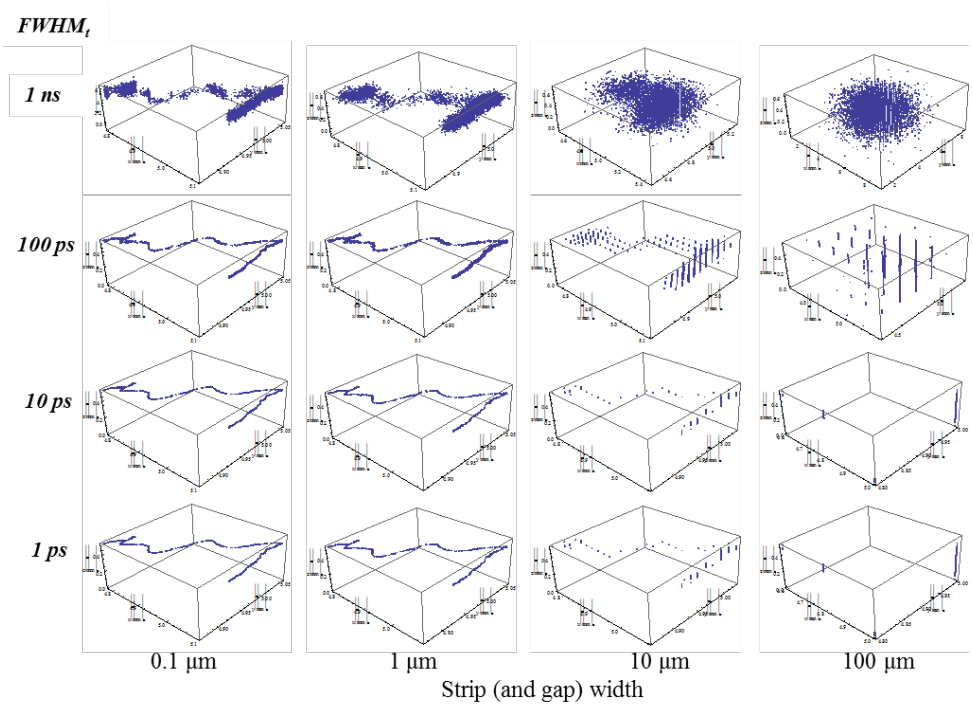


Fig. 4.8: Reconstruction of the (simulated) e-h track from a 500 keV electron, for various timing uncertainties (as measured at the FWHM) and meander-line geometries.

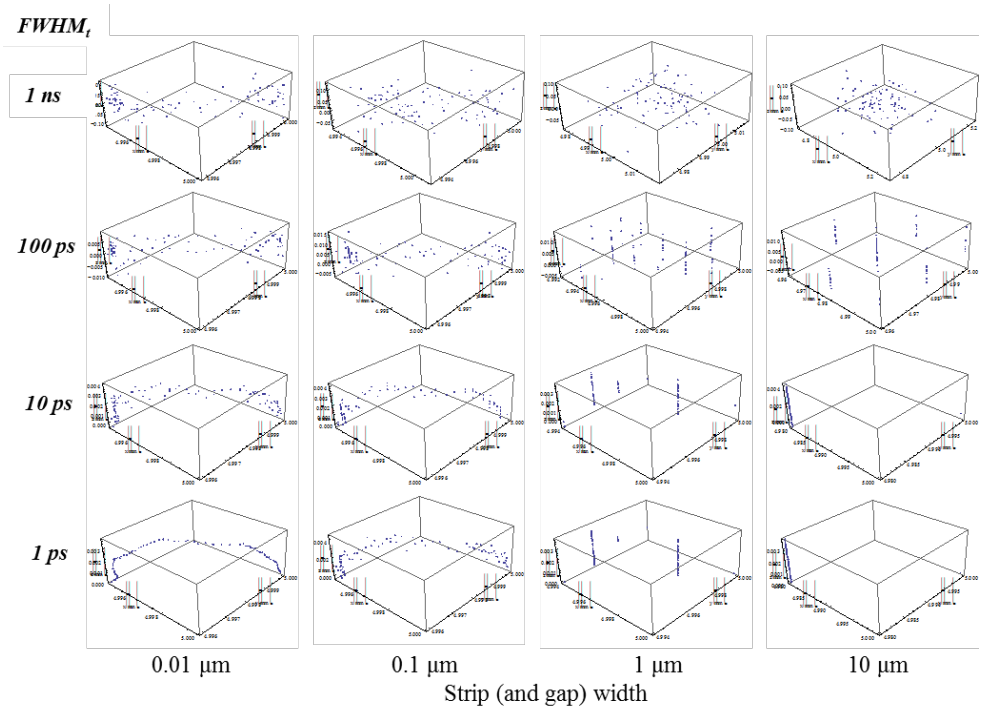


Fig. 4.9: Reconstruction of the (simulated) e-h track from a 50 keV electron, for various timing uncertainties (as measured at the FWHM) and meander-line geometries.

The underlying data is accurately represented by the track shown in the (0.01 μm, 1 ps) section. Note that for this particular track, the lateral extent is only 4 μm long; thus, one does not expect to be able to reconstruct its geometry with 1 μm or 10 μm spacing. For sub-micron grids however, no additional timing is required over that required for higher energy recoil products. Thus, if one wished to develop the detector (based on our or any other (CDD, CCD) approach) for x-ray applications, one should use electron-beam lithography for the device patterning.

Moving forward, the reconstruction algorithms were applied to our measured pulses, rather than their simulated counterparts, as done in Figs. 4.8 and 4.9. Extending our design to the 1 μm realm is straightforward from a fabrication standpoint, but the transmission line coupling increases, which slows the pulse propagation. The analysis has shown that the seemingly complicated pulses from the meandering electrode design can be converted into e-h trajectories, and one can successfully add recoil electron capability to the low and moderate energy gamma-ray regime if high performance instruments are developed. In particular, 1 μm or better positional resolution is desired and for our time-based design, 100 ps timing resolution or better is called-for.

4.1.3 Improvements in Angular Imaging via Recoil Charge-Carrier Tracking

The detector community has recognized that for gamma-ray imaging, the inclusion of recoil-electron information in the event reconstruction can substantially enhance the image resolution by constricting the extent of the back-projected probability cone. Similar improvements can accrue for neutron imaging devices if the ion-products are tracked. Despite this recognition, few have quantitatively studied the degree of improvement that can be achieved and how that improvement depends on the device characteristics.

The degree to which one can extract useful information from the electron track is limited because of its circuitous nature, examples of which are shown in Fig. 4.10, in which 400 keV electrons are impinged normally upon a silicon substrate.

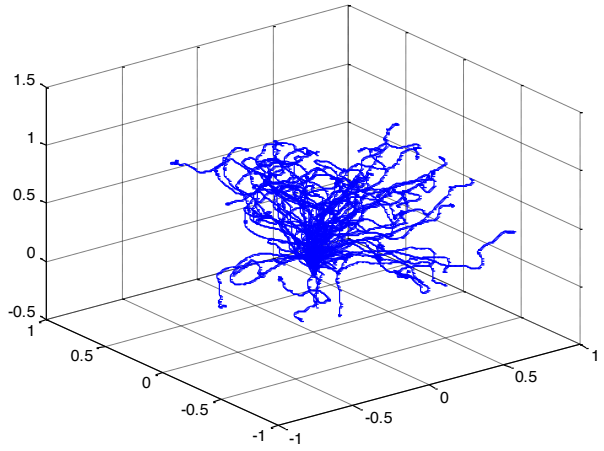
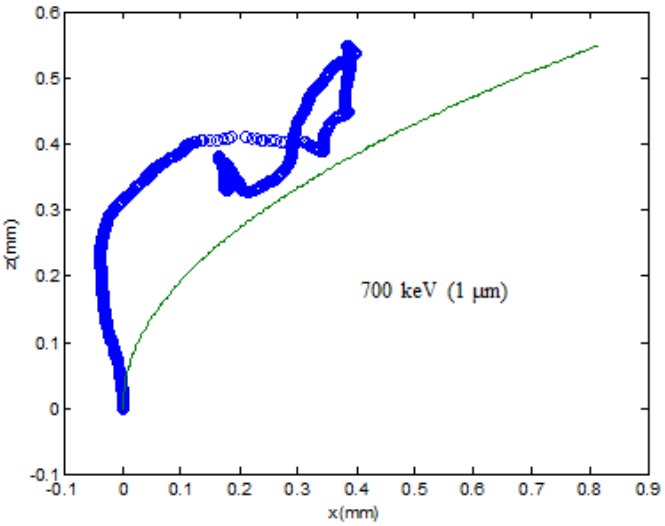


Fig. 4.10: Twenty-five trajectories for 400 keV electrons impinging upon a (functionally) semi-infinite slab of silicon, as derived using GEANT4.

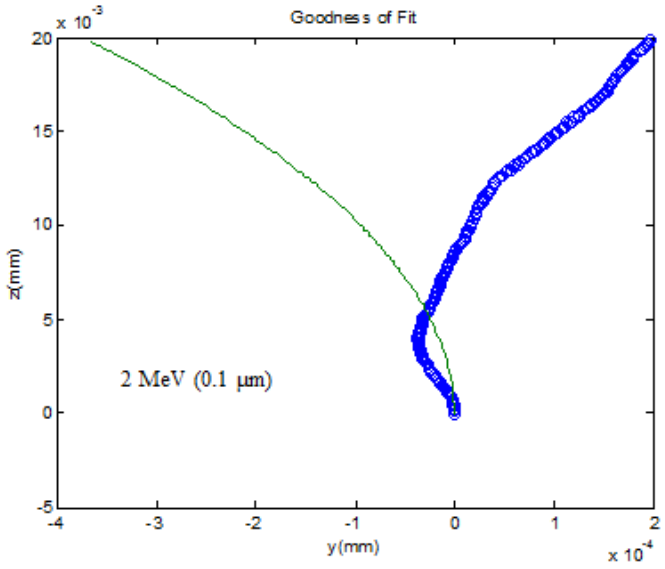
As suggested in Fig. 4.10, the initial directional information of the recoil electron is quickly lost during its momentum exchanges with the charge field in the active medium. If the useful part of the track is only 0.1 microns long say, then this places a very demanding constraint on the positional performance of the detector and requires a detector with sub-micron position sensing. However, if the directional information can be extracted from the first 100 microns of the track, then one can deploy relatively simple detector designs and achieve substantial imaging enhancements. Furthermore, the length of the electron tracks, and hence their useful fraction, will depend on the energy of the recoil electron. In order to limit the restrictions on the use of recoil-electron tracking, one would prefer to track small-energy recoils as well as their higher-energy counterparts. We have quantified these relationships using electron track information simulated in GEANT4 and with custom Matlab codes. In the functioning of the detector, one extracts the energy deposited- by counting the number of charge-carriers created- as well as the spatial distribution of the charge shower- using techniques that we have discussed at length previously. For the curving, twisting paths of the high-energy secondary electrons, various methods of extracting the *initial* direction of the primary secondary electron (either the photoelectron or the Compton scatter electron) were investigated. In particular, the electrode signals give us the path, as shown in Figs. 4.10 and 4.11, and we

intend to extract the recoil electron direction by fitting a curve to the initial part of the trajectory.

For instance, consider the electron trajectories shown in Fig. 4.11 along with the best polynomial fits. In Fig. 4.11a, a 2 MeV electron trajectory in silicon is sampled with 0.1 μm resolution, and in Fig. 4.11b, 1 μm resolution is used to map the path of a 700 keV electron. One can use alternative families of functions for the trajectory fitting, and as the results will show, one may not need to fit the trajectory at all in the statistical sense; however, we chose to investigate polynomial fitting because of its generality.



(a)



(b)

Fig. 4.11: Simulated electron trajectories (in blue) in silicon for (a) 2 MeV and (b) 700 keV electrons, as well as the best fit polynomials (in green).

Specifically, this code (see Appendix C) fits each order of polynomial (linear, quadratic, higher order polynomials) to each possible set of data points in the trajectory (3 or more data points for the linear fit, 4 or more data points for the quadratic fit, etc.) and then used statistical measures to find the best fit polynomial. The initial expectation was that the fast electron would travel more-or-less straight or at least in a continuous fashion until it underwent a large momentum exchange, and the best fit line would use all of the data points up to that kink in the trajectory, 20 points one time, 5 points another time, etc. The result did not meet initial expectations however, as shown in Fig. 4.12.

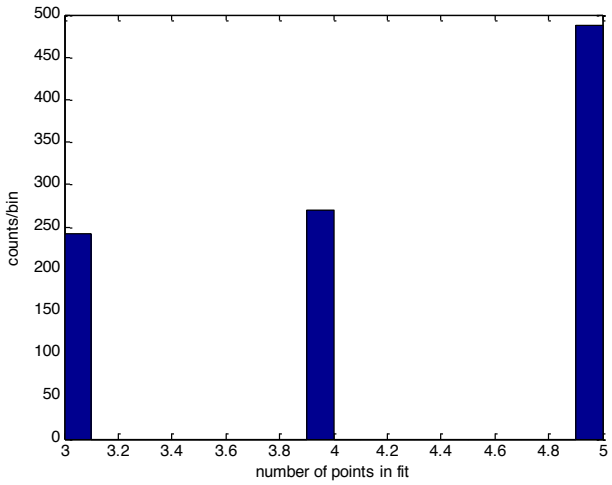


Fig. 4.12: Histogram of the number of points used in the best fit polynomial line, for 1000 paths of 400 keV electrons impinging silicon.

As shown in the Fig. 4.12 histogram, in which the number of points used in the best fit line are mapped, if we investigate only the first three orders of polynomials, the best fit line *always* used only 3 (for a linear fit), 4 (for a quadratic fit), or 5 points (for a 3rd order polynomial); that is, the polynomial order may change from trajectory to trajectory, but the best fit line always uses the minimum number of points.

In hindsight, this is understandable given the continuous nature of the data, and it suggests that one can simplify the recoil-electron determination by simply mapping the first two steps (as limited by the spatial resolution) and fitting a line to the points to measure the recoil electron direction.

Thus we employed a method that we can readily employ with our empirical data, as summarized in Fig. 4.13. In the simulation, we know the initial direction (i.e. 0^o) of the electron in the silicon, which then undergoes tortuous transport in the material. That trajectory is fit with the best fit polynomial, from which we estimate the initial electron direction (θ), using the equations shown in Fig. 4.13. Specifically, we extract the slopes (m_{xz} and m_{yz}) from the projections of the best fit polynomial and from them, derive the estimated electron direction. We would like the theta distribution to be a delta function at 0^o, but there is some spread in the distribution due to the finite resolution of the trajectory mapping. For a given electron energy, the distribution follows a one-sided Gaussian distribution such as that shown in Fig. 4.14. For an 800 keV electron as an example, 1000 electron tracks are used to derive the electron direction and form the distribution shown in the orange dots in Fig. 4.14, using a spatial resolution of 0.1 μm. For the sake of maximizing the utility of the study, what we really want is to know the functional form of the distribution so that the uncertainty due to recoil-electron energy and spatial resolution can be *calculated* and used in gamma-ray imaging algorithms, rather than having to track each electron in those studies, costing computation time. In principle, the form of the distribution can change depending on either variable; however, we discovered this variation is not significant in practice. First, the distribution in Fig. 4.14 follows the one-sided Gaussian form shown in Eq. 4.5 below and in green in the figure, in which the mean recoil angle, <θ> is equal to 0.0198^0. Note that the mean of the distribution is rather small and is, in fact, miniscule if one has sub-micron position resolution.

$$G(\theta) = \sqrt{\frac{2}{\pi}}\frac{1}{\theta}e^{-\frac{\theta^2}{2\bar{\theta}^2}} \tag{4.5}$$

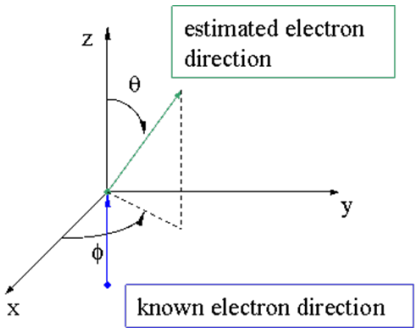


Fig. 4.13. The direction of the electron (θ) is estimated from the best-fit polynomial, via the slopes projected on the x,z and y,z planes. The distribution in θ then provides a measure of the certainty with which the recoil information can be utilized.

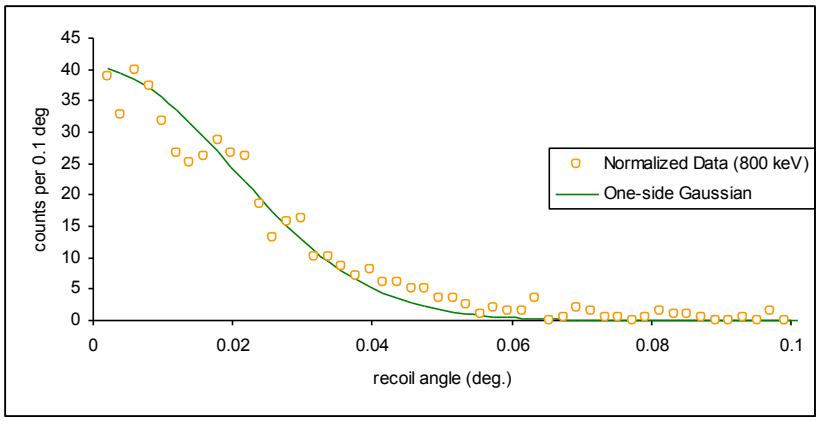


Fig. 4.14. Probability density distribution for the direction of the 800 keV electron (θ), as a function of energy, assuming 0.1 μm spatial resolution. The simulated data (in orange) and the one-sided Gaussian (in green), using the mean of the theta-deviation, are shown.

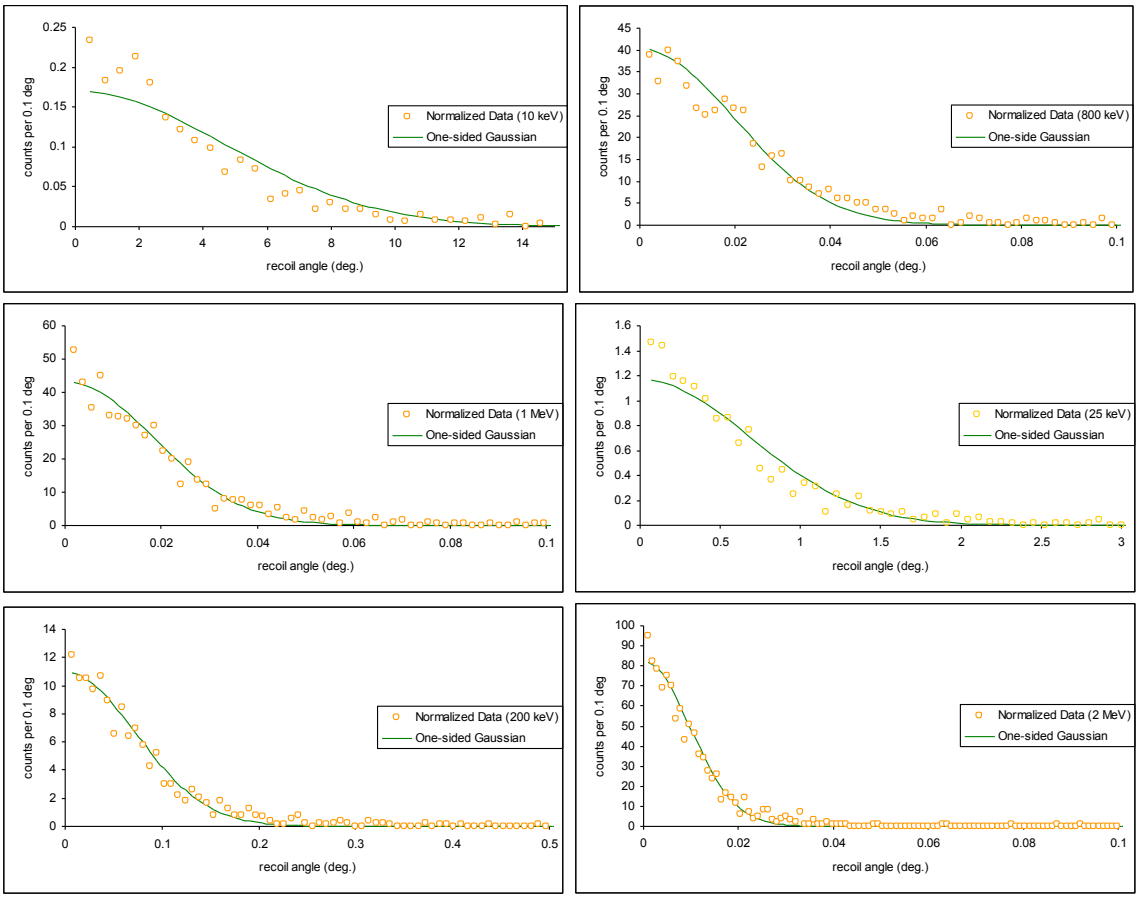


Fig. 4.15: Probability density distribution for the direction of the electron (θ), as a function of energy, assuming 0.1 μm spatial resolution. The simulated data (in orange) and the one-sided Gaussian (in green), using the mean of the theta-deviation, are shown.

In order to determine whether the Eq. 4.5 form varies, first with energy, we examined the data and the fit over various gamma-ray energies of interest, from 10 keV to 4 MeV. If the form doesn't vary, then one can simply calculate the mean of the theta-data and the entire distribution is determined, and this proves to be the case as shown in Fig. 4.15, in which the one-sided Gaussian, with the "empirical" mean, fits the data over the entire energy range. Note that the distribution deviates slightly at very low energies (< 20 keV), but the deviation is of negligible consequence in determining the uncertainty in the angular determination.

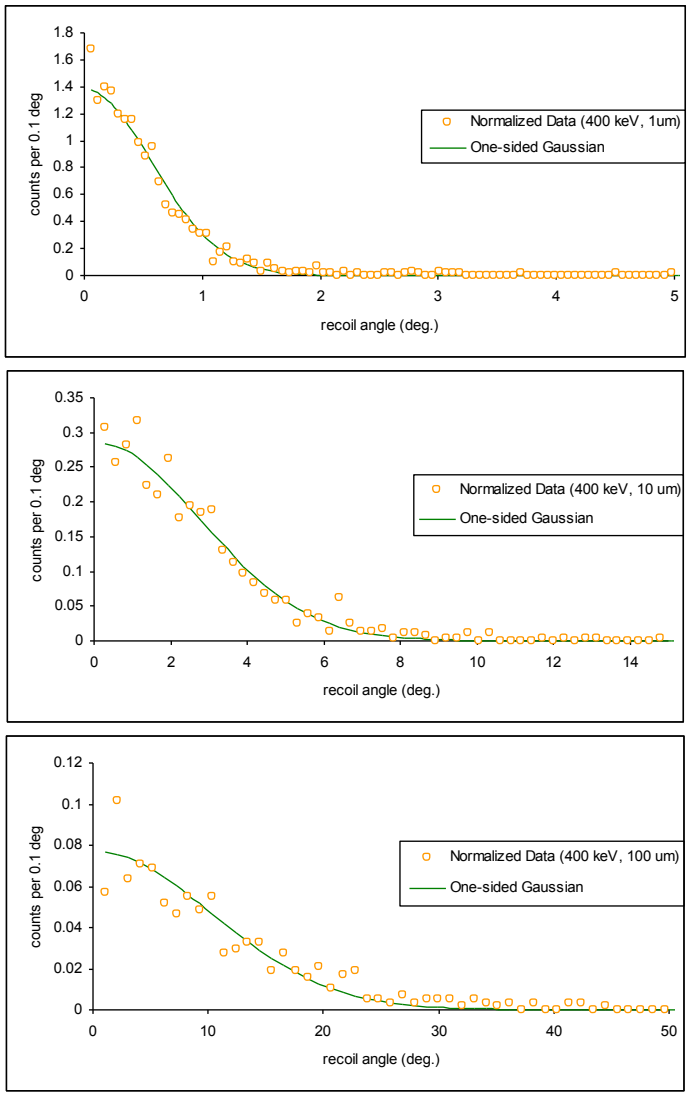
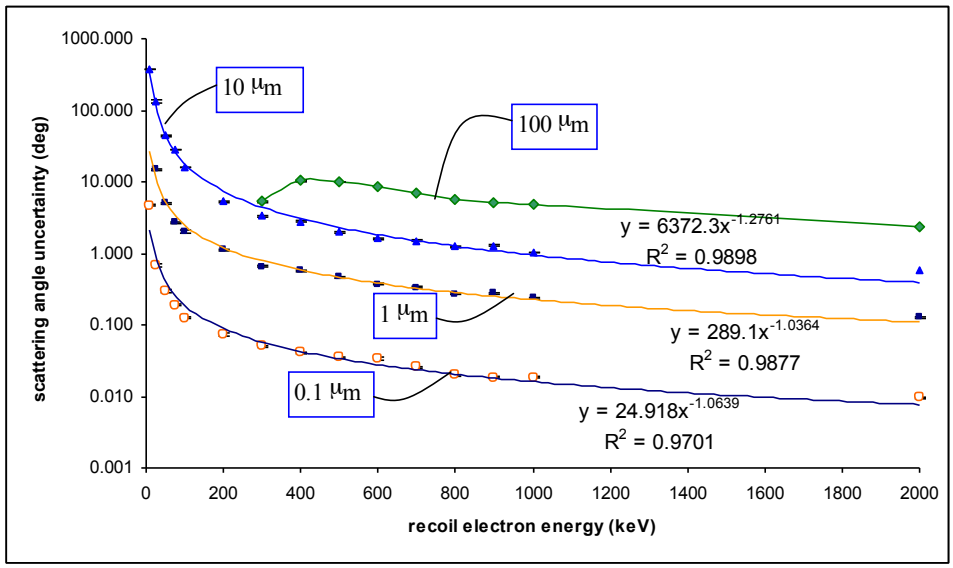
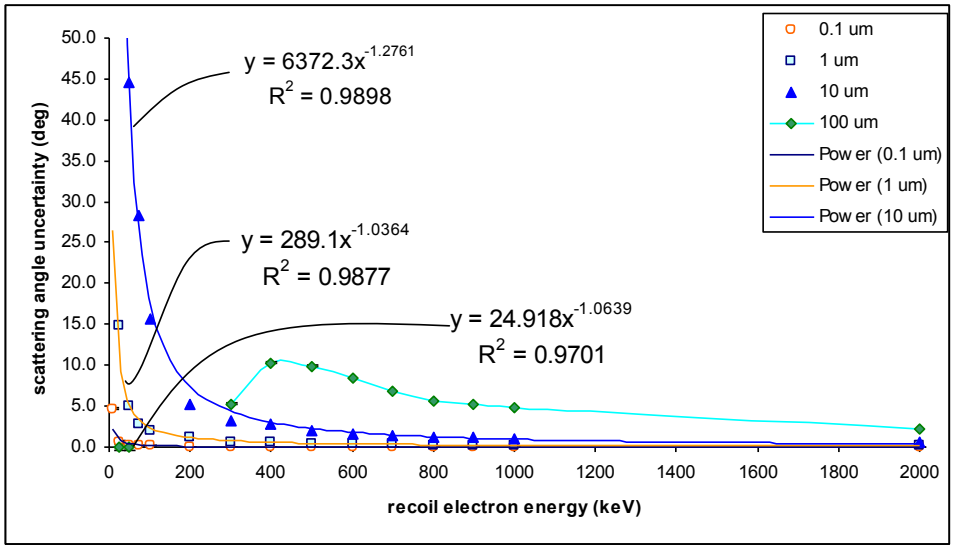


Fig. 4.16: Probability density distributions for the direction of the 400 keV electron (θ), as a function of spatial resolution. The simulated data (in orange) and the one-sided Gaussian (in green), using the mean of the theta-deviation, are shown.

We further determined that the probability density distribution (pdf) did not vary with the spatial resolution of the detector, as shown in Fig. 4.16, in which the 400 keV electron is tracked with three different spatial resolutions, 1 μm, 10 μm, and 100 μm. Note that the data shown in Fig. 4.16 provides a visual summary of the results over the entire energy range. Specifically, for 1 micron detector spatial resolution, the recoil-electron angle is determined to better than 1^0 ($\langle\theta\rangle = 0.6^0$), resulting in an angular determination improvement of 360 times compared with the use of simple back-projection.



(a)



(b)

Fig. 4.17: Semi-log (a) and linear (b) plots of the uncertainty in the recoil-electron angle as a function of recoil electron energy and parameterized by the detector spatial resolution.

At coarser resolutions, the angular determinations worsen, but the improvement that can be accrued from tracking the reaction products is still substantial. For instance, at 10 microns, the mean deviation in the recoil angle is approximately 3^0, and at 100 microns, the uncertainty in the recoil angle is 11^0. Thus, for the 400 keV electrons at least, electron tracking can improve the angular determination by 1 or 2 *orders of magnitude*, for even rather coarse electron-track mapping. The summary of the data over the entire energy range and for the spatial resolutions of interest is shown in Fig. 4.17. In Fig. 4.17, 1000 electron tracks at regular energies between 10 keV and 2 MeV are mapped with varying degrees of spatial resolution- 0.1 μm, 1 μm, 10 μm, 100 μm- and the pdf of the recoil angle and its mean deviation from the known true value (0^0) is derived. Those mean deviations are plotted in Fig. 4.17 as a function of energy on both semi-log (Fig. 4.17a) and linear scales (Fig. 4.17b). The data provides quantitative guidance regarding the spatial resolution that we should target when developing the enhanced-imaging instrument. As shown in the figure, 100 micron spatial resolution (green line in Fig. 4.17a) provides substantial imaging enhancement but only if the recoil electron energy exceeds 500 keV. The back-projected cone can be confined to a 100 or less arc-segment if a 100 micron tracking detector is deployed at these energies. Unfortunately, if one has only 100 micron resolution, then the limited length of the tracks restricts the application of the device to lower energies. For instance, at 300 keV, only 8 of the 1000 tracks were of sufficient length to allow any curve fitting (note the deviation in the data point from the overall trend due to this under-sampling). At smaller energies, the tracks are insufficiently long to allow one to gauge their direction. Since we would prefer wide applicability of the imager and further, we would prefer to enhance the imaging capability of even shallow-angle scattering events (for which the deposited energy is low), we seek better than 100 micron resolution. Nevertheless, one should keep in mind that for a high energy gamma-ray fields, for instance in nuclear resonance fluorescence or active neutron interrogation applications, on can deploy a 100 micron imaging to good effect.

At finer mappings, the track-length limitation still holds but it is less restrictive. For instance, if the recoil electron has energy of less than 20 keV, then the 10 micron resolution is too coarse to map the electron direction. Since the vast majority of (primary)

recoil electrons will exceed this threshold, then 10 micron spatial resolution can be used, and will result in a marked improvement in the angular reconstruction. That is, when back-projecting a cone (perhaps with modulated probably via filtered back-projection), then one can confine the cone to an arc-segment of angular extent of 1^0 to 10^0, depending on the recoil energy, as opposed to the 360^0 (modulated) back-projection using existing techniques. With even 10 micron spatial resolution then, one can significantly reduce the efficiency requirement of the imager because each event is so much more informative. At 1 micron or less resolution, then the electron direction can be pin-pointed precisely and further, the tracks of ions can be imaged as well, with application to neutron imagers.

Thus, if one is building a gamma-ray imager, then 10 micron spatial resolution is an acceptable target for substantial image enhancement and if a neutron-imaging device is envisioned, then 1 micron spatial resolution is preferred. The current generation of diagnostic devices employs both 20 micron and 200 micron spatial resolutions, and will therefore allow to empirically testing these simulated results, once their utility in optimizing the signal extraction is realized.

4.1.4 Maximum Likelihood Image Reconstruction for Compton Imaging

The image reconstruction for radiation-induced energy-deposition events can be considered as a gamma-ray source parameter estimation problem. In the literature of Compton imaging, the framework of maximum likelihood (ML) has been widely used in order to obtain estimates of parameters best describing the measured data. More rigorously, a ML estimate is defined as a set of parameters that generates the observed measurements with the maximum probability, that is

$$\hat{f}_{ML} = \arg\max_{f \geq 0} \Pr[g \mid f], \tag{4.6}$$

where $g = \{g_1, \cdots, g_I\}$ is a vector containing measurements and g_i denotes the number of counts measured as event i (=1,$\cdots$,I). The vector $f = \{f_1, \cdots, f_J\}$ described the source distribution and f_j denotes the intensity of the jth pixel of the source. Since each element of f typically stands for the intensity of the source image at the associated pixel location, it is reasonable to impose a positive constraint on the elements of $\hat{f}$ as in Eq. 4.6. By assuming that the gamma-ray imaging system that we use for image reconstruction is

well described as a linear system in the regime of energies that we are interested in, we are able to build up a relationship between a source distribution (input) and the corresponding measured data (output) as follows:

$$Tf = g, \tag{4.7}$$

where the matrix T is called a system response function, transforming the source distribution in the image space to the data in the measurement space. The elements, t_{ij}, in the system response function T can be thought of describing the probability that a photon or neutron emitted from the jth pixel of the source distribution creates the ith recorded event. In our application, the imaging space associated with the source distribution f is, in general, a 4π sphere. The ith event in the measurement space contains the information of the positions and deposited energies from interactions between incident quanta and the materials in detector. Fig. 4.18 illustrates the linear modeling of the gamma-ray imaging system represented in Eq. 4.7.

If we define a quantity using the elements of the system response function, called a sensitivity map as follows:

$$s_j = \sum_{i=1}^{I} t_{ij}, \tag{4.8}$$

which is the probability that a photon emitted from the jth pixel is detected as any event, then a spatial plot of S_j on the 4π sphere represents the sensitivity of the gamma-ray imaging system to photons from the jth pixel and it plays a key role in the update equation derived by the expectation-maximization (EM) algorithm as discussed below.

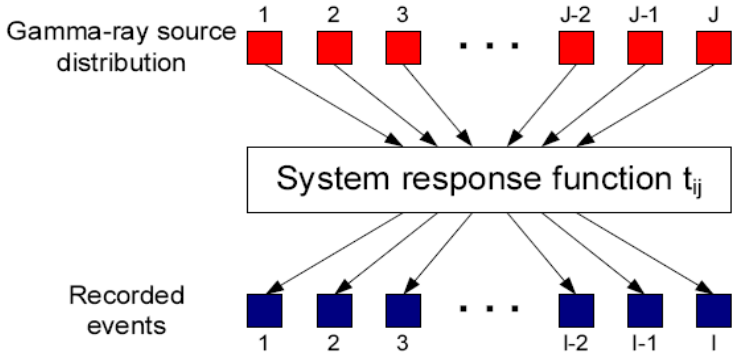


Fig. 4.18: A linear model for the gamma-ray imaging system.

4.1.4.1 Conventional ML-EM Algorithm

Under the preset time assumption- that is, collecting counts for pre-specified time duration- we accurately assume that the number of photons emitted from the j pixel obeys a Poisson random variable whose mean is given by f_j. Considering the definition of t_{ij}, then we can reasonably assume that the number of photons generated from the j th pixel and recorded as the ith event follows another Poisson random variable with mean $t_{ij}f_j$. Therefore, the total number of photons recorded as the ith event is a Poisson random variable with mean value of

$$\bar{g}_i = \sum_{j=1}^{J} t_{ij} f_j \quad . \tag{4.9}$$

Therefore, we have

$$\Pr[g_i \mid f] = \exp(-\bar{g}_i)\frac{(\bar{g}_i)^{g_i}}{g_i!} \quad , \tag{4.10}$$

for each event $i(=1,\cdots,J)$. Under the assumption of independent measurements (i.e. $g_1,\cdots,g_I$ are statistically independent), we are able to combine all probability mass functions (PMF) in Eq. 4.10, producing the following joint PMF parameterized by the source distribution in the measurement space:

$$\Pr[g \mid f] = \prod_{i=1}^{I} \Pr[g_i \mid f] = \prod_{i=1}^{I} \exp(-\bar{g}_i)\frac{(\bar{g}_i)^{g_i}}{g_i!} \quad . \tag{4.11}$$

By taking logarithm of Eq. 4.11, we obtain the log-likelihood function as follows:

$$L(g \mid f) = \log \Pr[g \mid f] = \sum_{i=1}^{I}\left[-\sum_{j=1}^{J} t_{ij} f_j + g_i \log\left(\sum_{j=1}^{J} t_{ij} f_j\right) - \log(g_i!)\right], \tag{4.12}$$

whose maximization produces a ML estimate of the source distribution f as in Eq. 4.6. Note that since the last term in Eq. 4.12 is independent of f, it can be ignored when the optimization is performed.

A popular method to maximize the given log-likelihood function in Eq. 4.12 is an EM algorithm, producing iterative equations. The main idea of the EM algorithm is to replace a difficult and/or numerically unstable direct maximization of Eq. 4.12 with a sequence of easier and/or stable maximizations of surrogate functions. Therefore, the EM algorithm consists of two steps; the expectation step (E-step), in which a surrogate function is identified, and the surrogate-function maximization step (M-step). The EM algorithm has guaranteed convergence properties [1] and yields remarkably simple iterations for updates in our case [2]. Following the standard procedure of the EM algorithm [1], one can arrive at the following equation for update:

$$\hat{f}_j^{k+1} = \frac{\hat{f}_j^k}{s_j} \sum_{i=1}^{I} \frac{g_i t_{ij}}{\sum_{j'=1}^{J} t_{ij'} \hat{f}_{j'}^k}, \tag{4.13}$$

where $\hat{f}_j^k$ denotes an estimate of f_j in the kth iteration and S_j is the jth pixel in the sensitivity map for j=1,⋯,J [2]. From a reasonably selected initial guess, by iterating Eq. 4.13, we can obtain a ML estimate of the source distribution, f.

4.1.4.2 List-Mode ML-EM Algorithm

Eq. 4.13 requires us to perform a summation over all (i) possible measurements. Even for two-pixel events, this requires a huge amount of computation. For example, in a single 3-D CdZnTe detector that our colleagues in the University of Michigan developed, if the depth is separated into 10 bins and the energy is represented by 1 keV per channel, for a system with dynamic range of 1 MeV, about 10^6 bins are necessary to store a single gamma-ray interaction. Therefore, for two-pixel events, the number of the needed bins is 10^{12}, which limits the practical use of the conventional EM algorithm.

To overcome this disadvantage and reduce the computational burden, an alternative method, the "list-mode ML approach" [7], was developed based on only the measured attributes of each event in a list. Since the number of measured events is typically far less than the total number of bins in practice, the measurement vector g is very sparse with elements of mostly zeros and ones. The list-mode ML-EM equation can be easily

obtained by replacing g_i in Eq. 4.13 with ones for the measured events [3]. As a result, the sum over all possible measurements becomes the sum over all measured events. Therefore, we have the following reduced form update equation:

$$\hat{f}_j^{k+1} = \frac{\hat{f}_j^k}{s_j} \sum_{i=1}^{M} \frac{t_{ij}}{\sum_{j'=1}^{J} t_{ij'} \hat{f}_{j'}^k} \tag{4.14}$$

in which the indices indicate the *j*th pixel in the *k*th iteration, and M denotes the number of measured events. The summation in Eq. 4.14 can be obtained using the small portion of the *j*th column of the system response function *T*, dramatically reducing the computational cost.

The stopping criterion of the ML-EM algorithm is a design parameter. For that purpose, the difference between two subsequent estimates of the source distribution, for instance,

$$D(k) = \sum_{j=1}^{J} | \hat{f}_j^{k+1} - \hat{f}_j^k | \tag{4.15}$$

is usually used. When $D(k)$ is less than a pre-specified value, the ML-EM algorithm is stopped and the estimate of *f* is obtained.

4.1.5 Angular Reconstruction via Compton Imaging

Our main focus has been on the theoretical and empirical development of the detection *modality* that will most simply allow us to map the recoil charge's transport with precision. Our *analytical* efforts to develop an optimum reconstruction technique will be summarized in this section. The method that one chooses for image reconstruction for Compton-scatter events depends, to some extent, on the envisioned application of the instrument. There is a trade-off to be made between improved image quality and processing time, as illustrated in Fig. 4.19. In general, the fastest angular-reconstruction techniques (simple back-projection) produce the coarsest image, while the sharpest image is based on statistical methods utilizing the entire data set, for which the image

optimization limits the overall speed of the system (e.g. Maximum Likelihood Expectation Maximization Techniques (ML)). To be specific, the ML algorithm can take over a minute to process the event sequences from 1000 radiation interactions (based on 662 keV imaging using a small CdZnTe crystal). The deterministic methods, based on either simple or filtered back-projection, can be done on an event-by-event basis, and are therefore real-time (assuming that the processing for the back-projection is slower than the event rate). As more counts are accumulated, the image increases in sharpness, but it never benefits from a comprehensive consideration of all events, which is achieved using more probabilistic techniques, such as the List-Mode MLEM algorithm.

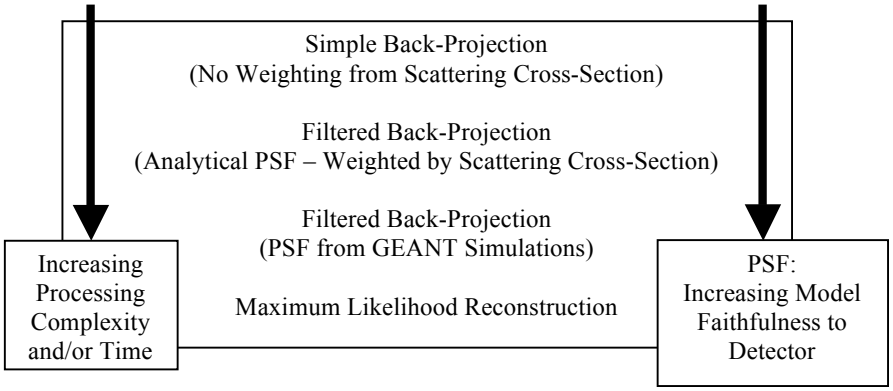


Fig. 4.19: Schematic of the image processing schemes and their relationship to processing time and image quality. Note that PSF stands for point spread function.

That is, the probabilistic methods use the knowledge of all of the events to condition the probabilities of each separate interaction during the back-projection. Thus, several or many iterations are needed to fully optimize the image reconstruction.

In particular for the special nuclear materials (SNM) detection application, the ML technique reduces the noise in the background, which is critical for isolating the target. However, as mentioned above, it is very processing intensive, requiring minutes of computation time for a limited number of counts, using a single standard desktop computer. Of course, one can deploy brute-force processing techniques in which a multitude of parallel processors are used. Our colleagues at the University of Michigan are using this technique, and we will piggy-back on their results if they are successful.

One can also anticipate faster computers in the future in concert. However, one should note that one advantage of tracking the recoil charge-products is that one can potentially eliminate this signal processing requirement, and thus remove one of the bottlenecks to the development of effective imaging instruments. For standard instruments, MLEM techniques play, for intermediate term applications, the role of providing a sharper image for coarser, faster techniques during the survey of the environment. Our initial focus has thus been on simple and filtered back-projection algorithms. As suggested in Fig. 4.19, simple back-projection calculates the scatter angle and smears the probability of incidence equally along the projected cone. This probability is not equal however, but depends on the non-uniformity of the Klein-Nishima scattering cross-section, as well as the disparate probabilities due to different path lengths for impinging gamma-rays. In fact, the detector characteristics can be best accounted for if one uses an accurate detector model, which we are building using GEANT4, as we discussed previously.

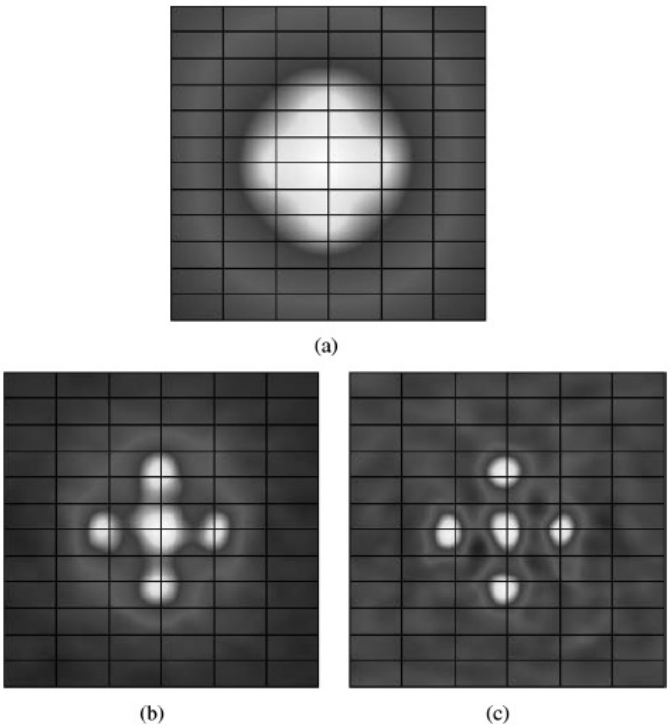


Fig. 4.20: Reconstructed images of simulated data using 5 Cs-137 point sources impinging a CdZnTe detector. (a) Simple back-projection. The five sources cannot be distinguished. (b) Filtered back-projection using theoretical PSF based on the scattering cross-section. The five sources can be distinguished. (c) Filtered back-projection based on the detector PSF, as derived from GEANT4, showing improved angular resolution.

Using a small (1.5 x 1.5 x 1 cm) CdZnTe detector, our colleagues have demonstrated the progressive gain in image quality at the cost of more sophisticated processing, as shown in Fig. 4.20 [4]. The simple back-projection algorithm fails to distinguish the five Cs-137 (E_γ = 662 keV) point sources, as shown in Fig. 4.20a. When the scattering cross-section is properly included, as in the filtered back-projection image of Fig. 4.20b, the sources are distinguished. Even further improvements in the angular resolution are gained if the detector effects are included, as demonstrated in the Fig. 4.20c image. Including the effect of the detector mainly consists of accounting for the non-uniform probability of gamma-ray impingement as limited by the transmission probability through the materials in the imager.

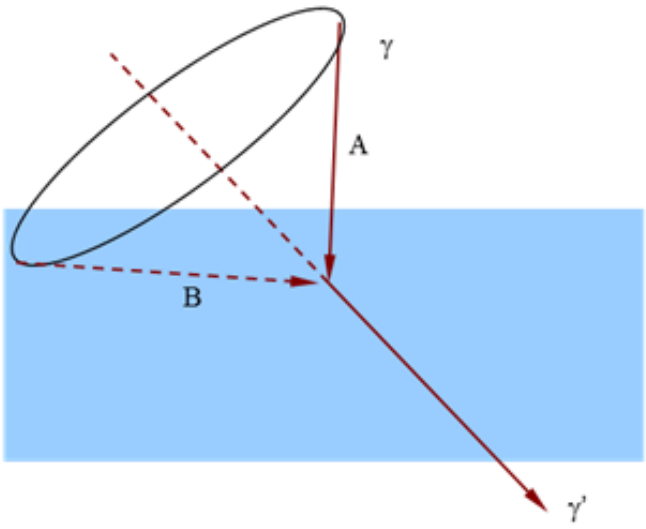


Fig. 4.21: Different back-projections reflect different path-lengths through which gamma-rays must traverse, and therefore different sequence probabilities; for instance, path A is more probable than path B because of the former's shorter traversal distance.

For instance, consider Fig. 4.21, in which the scatter cone is known. Those back-projections, Ω, nearer ray A are more likely than those near ray B, because the path length traversed by gamma-ray A is shorter (and thus traversed with higher probability) than gamma-ray B. If the attenuation in the gamma-ray flux through the detector is negligible, then we do not expect a significant detector effect (at least in this respect), and this is observed in Fig. 4.20, in which the differences between Figs. 4.20b and c are

noticeable but moderate. For a SNM imager, the unit may be designed such that a substantial proportion of the incident gamma-ray flux produces a useful count. Thus, the detector path-lengths are non-negligible and will be included in the image reconstruction, though at the cost of processing time because one must properly account for all of the various probabilities. Note that the detector point spread function will also vary with the incident gamma-ray angle unless the detector is spherically symmetric, which also complicates the reconstruction. It is therefore useful to employ simple back-projection and scattering-filtered back-projection for fast processing and for initial guesses for the MLEM algorithm.

4.1.6. Pulse width reduction due to higher bandwidth readout and matched readout

The signals were AC coupled using an ultra-broad band 0.1 μF coupling capacitors (ATC 545L) and read out with a higher bandwidth oscilloscope- 3.5 – 10 GHz bandwidth instead of the earlier utilization of a 2 GHz bandwidth with a 50 Ω input. The result, as implied in the typical pulses of Fig. 4.22, is that the signal rise time reduced by a factor of 10 (to the 100-200 ps range from 1 – 2 ns) and the pulse width reduced by a factor of 10 (to the 2 – 4 ns range), yielding a better representation of the transmitted signal.

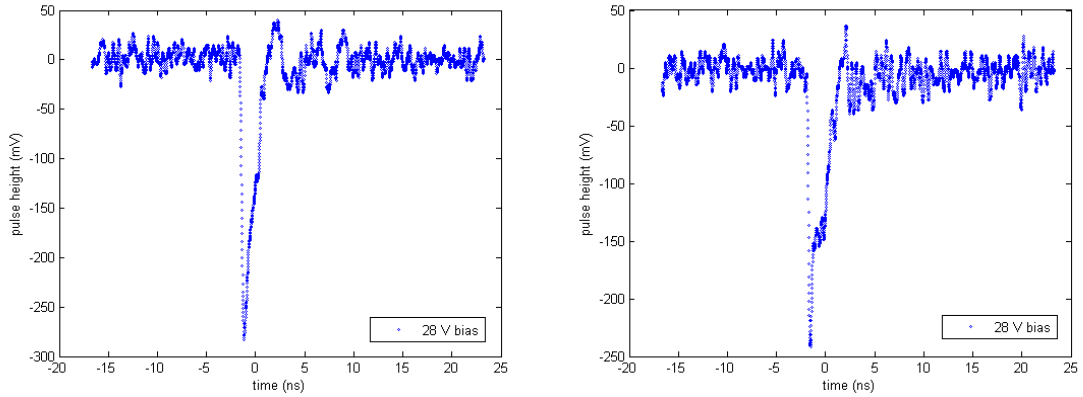


Fig. 4.22: Sharper, narrower pulse shapes reflecting the fact that a high bandwidth readout and proper termination are crucial for preserving the pulse shape.

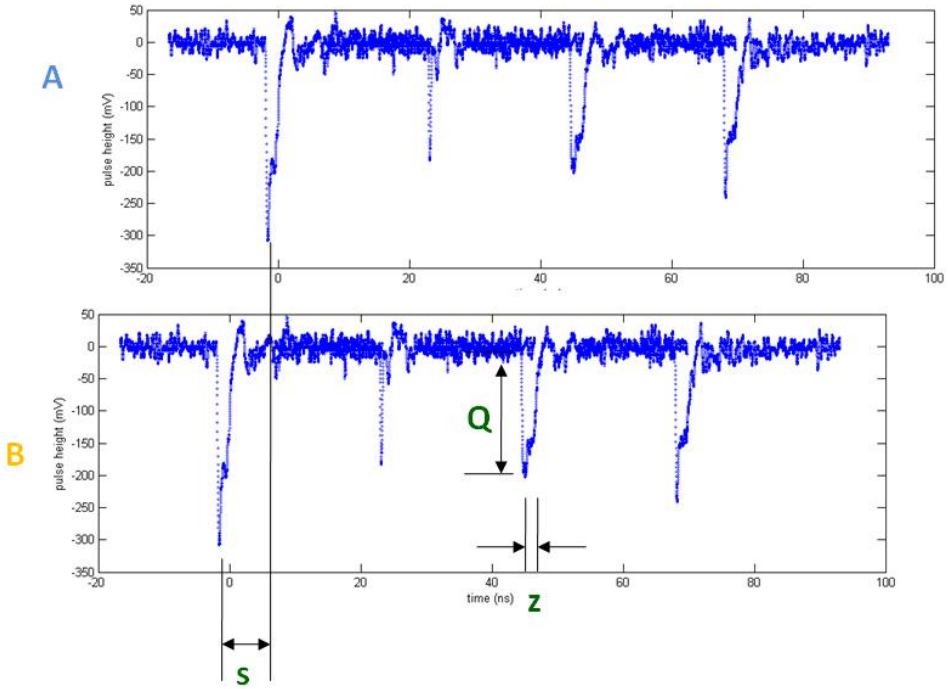


Fig. 4.23: Extraction of physical parameters, the charge Q under each strip, its distribution in depth z, and the position along the strip-lines, as measured from the two pulse trains measured on the ends, A and B, of the meander-line electrode.

The consequence of narrower pulses is that it becomes easier to extract the signals from multiple lines because the overlap will be reduced or eliminated altogether, depending on the segment length and the propagation speed of the line. As shown in Fig. 4.23, pulse separation would allow a simplified reconstruction algorithm, in which one can extract: (a) the charge deposited under each strip from the height of the pulse, (b) the depth of the charge distribution from the width of the pulse, and (c) the lateral position, along the meander line, from the time difference between the pulse trains at the two ends of the line.

4.1.7 Electron-hole Track Reconstruction

Recall that the charge number and position information is contained in the current *transitions* which provide a measure of the time and intensity of the electron or hole variation, relative to the time at which the e-h pair started its drift. One does not therefore require that the entire pulse width, a few nanoseconds wide and temporally

separated, but that only the sharp variations be distinct. For instance, a 28 μm alpha track was fully collected during a ~380 ps period, which establishes the time for which smaller temporal separations will result in overlapping information. If one has charge distributed over 3 μm, from an electron for instance, then the transitions must be separated by 38 ps, a leg-to-leg time difference that is easy to achieve for standard meander-detector configurations.

The operation of the algorithm is outlined in Fig. 4.24, operated on a chain of experiment pulses, in which the separation is changed in the code to simulate various degrees of overlap; for instance, the pulse separation in Fig. 4.24 corresponds to well-separated pulses. The pulses in the upper right figure, which correspond to charge collection on five neighboring lines, were experimentally measured from an alpha particle interaction, in which the low detector bias (28 V rather than 1000 V for full charge saturation), was chosen to yield pulses with a signal-to-noise ratio consistent with charge sharing between five lines of the meander.

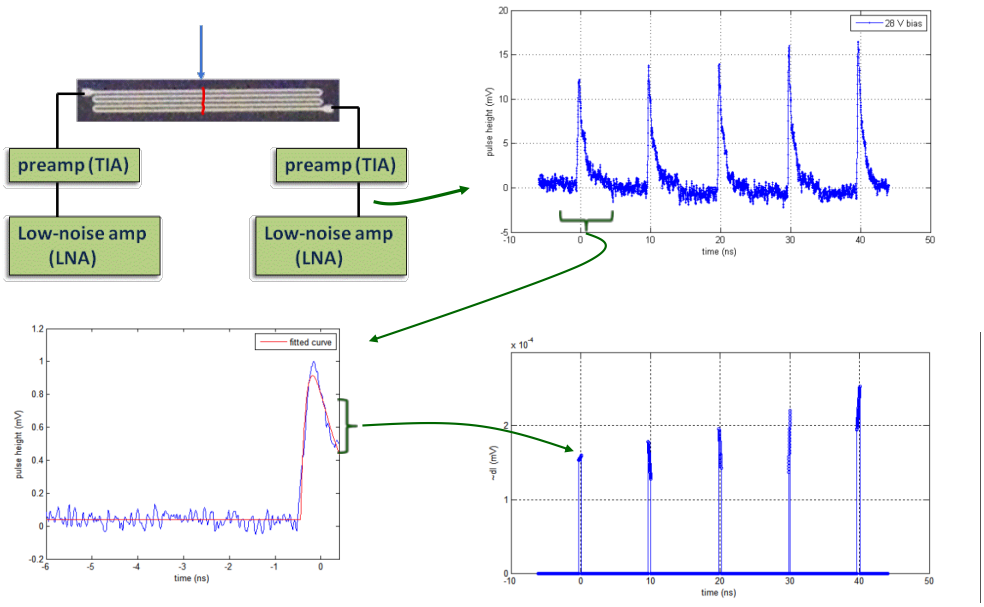


Fig. 4.24: Well-separated pulse train example, in which the separation is simulated, in order to demonstrate the principle behind the depth mapping algorithm. As shown in the lower half of the figure, the fitted curve to the loss portion of the pulse mapped to the depth dimension, as shown the lower right.

For each pulse, the depth distribution is extracted using the curve fitting methodology implied in the lower half of Fig. 4.24. Specifically, piecewise non-linear least squares fitting routines in Matlab are used to find the time at which the pulse starts (and hence the time at which the charge motion begins) as well as the time and intensity of the charge loss, which is then mapped to the depth coordinate, as shown in the lower right of the figure. Time windows are then used to assign the depth distribution to the leg of the meander, which then establishes its lateral coordinate. The mapping of the x, y, and z coordinates is shown in Fig. 4.25, which also includes the energy mapping, as follows. The absolute magnitude of the current-change signal reflects the charge which provides a measure of the energy deposited in the detector voxel, redder regions indicating greater energy loss.

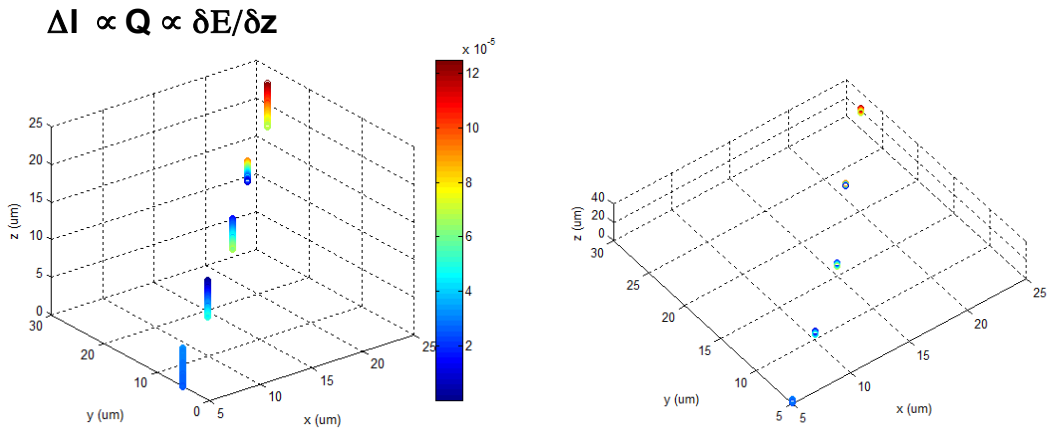


Fig. 4.25: Mid-point Interaction Track. Partially simulated ion track (from experimental pulses with simulated overlap), reflecting the spatial coordinates in three dimensions (5 μm resolution in the lateral dimensions and 0.75 μm resolution in the depth coordinate, z) and the relative energy deposited in the color mapping. The right-hand figure shows a view from the top which clarifies the lateral points at which the charge is assigned.

As shown on the right hand side of Fig. 4.24, the spatial points are mapped to the center of the strip lines, which are assumed to be separated by 5 μm. We could draw the colored pixels to reflect the positional uncertainty by drawing 5 x 5 μm square pixels; however, the information is being used to determine the direction of the recoil product, and the pertinent information are the positions shown. Even if a charge is created at (x,y) = (6 μm, 6 μm) for instance, it is collected on the line and assigned a position of (x,y) =

(5 μm, 5 μm), at least initially. The track in Fig. 4.24 is thus an accurate representation of the charge distribution and provides information that can enhance the imaging resolution of the neutron or gamma-ray device; however, if the interaction occurs sufficiently near the edge of the meander line, as shown in Fig. 4.25, then one can't avoid the overlap of pulses. The tracks in Fig. 4.26 show the consequence of this overlap.

First, on the left hand side of the figure, the five pulses shown in Fig. 4.24 are instead transformed into 3 pulses, two of which exhibit overlap. When the pulse train is applied to the first part of the imaging algorithm, the track emerges as shown on the right side of Fig. 4.26, yielding three distinct depth distributions rather than the five that are known to exist. Fortunately, this misidentification can be corrected in the code because the track exhibits two non-physical effects: (a) the lack of track continuity in the lateral dimension and (b) the lack of Bragg-curve energy depositions in the depth dimension. That is, the track is drawn near (x,y) = (5 μm, 5 μm) and (x,y) = (15 μm, 15 μm), but the distribution for the 10 μm line is missing from both dimensions, which indicated that the information is contained in an overlapping in the neighboring lines. The manner in which that overlapped information can be discerned from the expanded view of one of the depth distributions shown in the figure.

Specifically, the energy deposited is low (blue) then high (red), and then low again (blue), which indicates that the red portion is due to the addition, and overlap, of the two neighboring blue sections. One can then extract the original pulses by fitting the blue portions and assigning the deeper section (in this case) to the next line over. It is therefore rather straightforward to compensate for dual-track overlap in the code; however, a multitude of overlaps will muddy the distribution and lead to errors during the unfolding. For relatively straight tracks, like those from ions, we can design the detector so that only two pulses, at most, overlap, because note in the meander figure shown in Fig. 4.26, that for the third leg to produce an overlapping pulse, that contribution must travel all the way to the end of the line and back, a time that is generally too long to produce an overlapping problem.

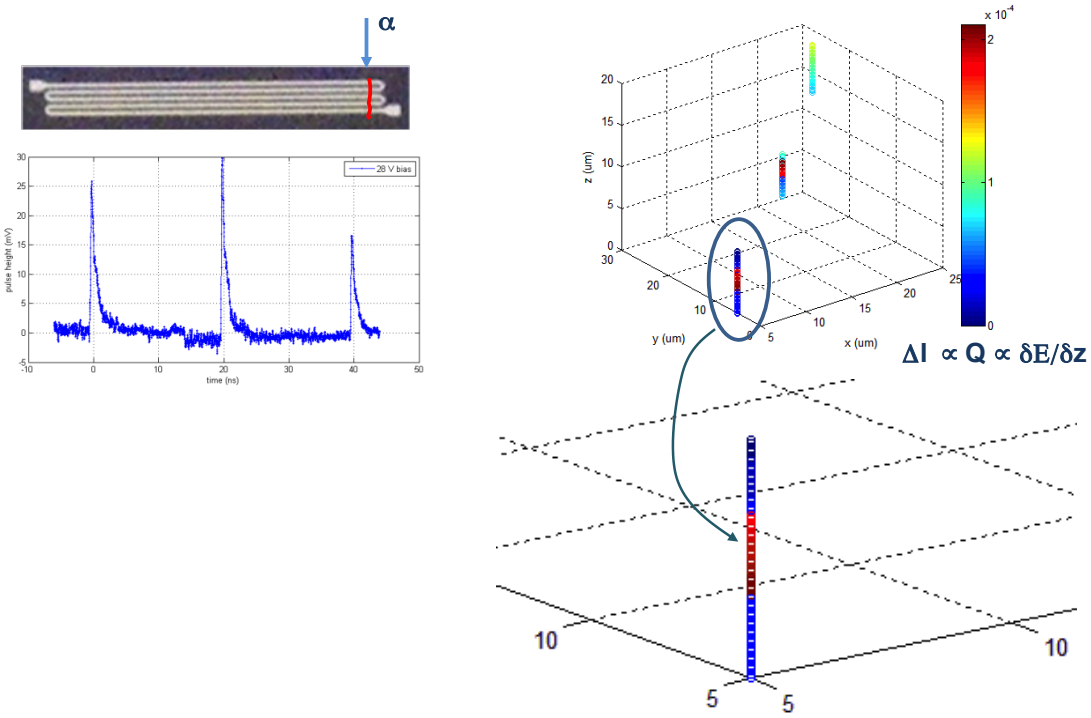
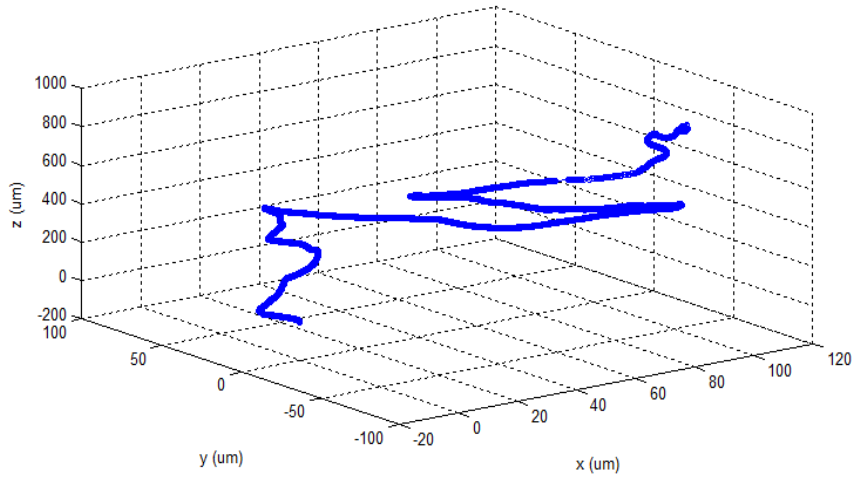
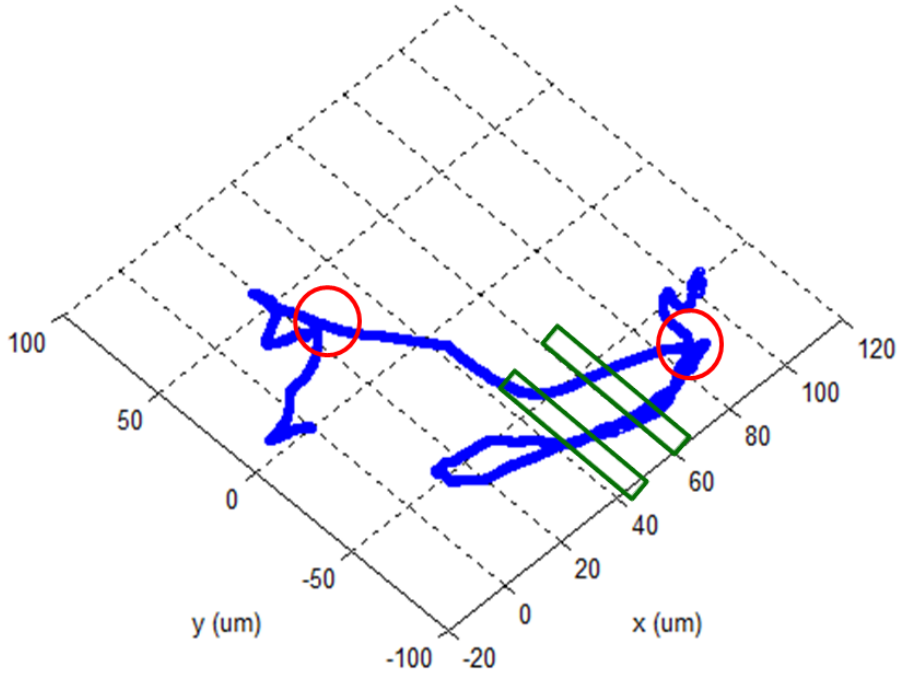


Fig. 4.26: Near-edge Interaction Track. Partially simulated ion track (from experimental pulses with simulated overlap), reflecting the spatial coordinates in three dimensions (5 μm resolution in the lateral dimensions and 0.75 μm resolution in the depth coordinate, z) and the relative energy deposited in the color mapping. The right-hand figure shows an expanded view of the depth distribution, reflecting the non Bragg-curve energy deposition.

For ion tracks therefore, the charge deposition is large and the tracks are straight, and although they are also short, our resolution is sufficiently fine that they can be reconstructed into discernible charge tracks. For electrons resulting from gamma-ray interactions, the tracks extend across a larger area of the detector and propagation speed requirements are therefore less demanding; however, the charge density is less and greater overlap can occur, as indicated for the example track in Fig. 4.27.



(a)



(b)

Fig. 4.27: (a) Typical electron track for a 500 keV electron, as simulated in GEANT. (b) Red circles show examples of those regions in which the charge is deposited in the same lateral coordinates and can therefore potentially overlap in the depth distribution. The green rectangles show that along a given lateral line, the back-tracking distribution can result in different parts of the track resulting in charge collection from a single line.

As shown in the red circles, the fast electron can roam through the detector bulk producing charge depositions at the same lateral coordinates but at different depths. Within the single 4 ns (maximum) that is required to drift the charges through the bulk, the electrons and holes from each of the separate packets will be collected. As long as their charge collections are not overlapping, then this depth overlap is not a problem, and the time over which a small charge pack will be collected is small. The challenge is therefore not to avoid overlapping, but to be sensitive enough to small charge numbers to be able to discern the transitions in the pulse. This fact motivates our experimental work because the improvement of the low-noise detectors and sensitive readout electronics directly impacts the gamma-ray imaging resolution.

The green rectangles in Fig. 4.27 show examples in which different parts of the electron track are collected on the same leg of the meander, but at different points along that leg. As for depth overlap, as long as the different charge distributions are not too close, they can be separated or their overlap can be corrected, but again charge sensitivity is the key to confidently extracting the relevant charge number variations. As long as these effects are corrected, then the track can be reconstructed as shown in Fig. 4.28, for two different lateral position resolutions. Note that what also mitigates the overlap concern is that the reconstruction does not have to be perfect for the entire track but only for the earliest part of the track, which provides the directional information desired. In general, we use the energy information from the whole track to determine which end was the start point and which was the end point, but then plan to focus the code on the early part of the track to maximize the precision of the reconstruction.

Note in Fig. 4.28 that the 10 μm resolution is insufficient to correctly identify the *initial* electron direction, which is the critical parameter that we are extracting from the distribution, because the electron track quickly kinks after entering the silicon. However, for this track, the 5 μm resolution design does provide an acceptable estimate of that initial direction. This example simply reflects the more detailed statistical analysis in which we discerned that for electrons of moderate energy; that is, for electrons in the Compton continuum in the 100 keV to 2 MeV range, resolution on the micrometer-scale is desired if the recoil-direction information is to be utilized from most of the interacting electrons.

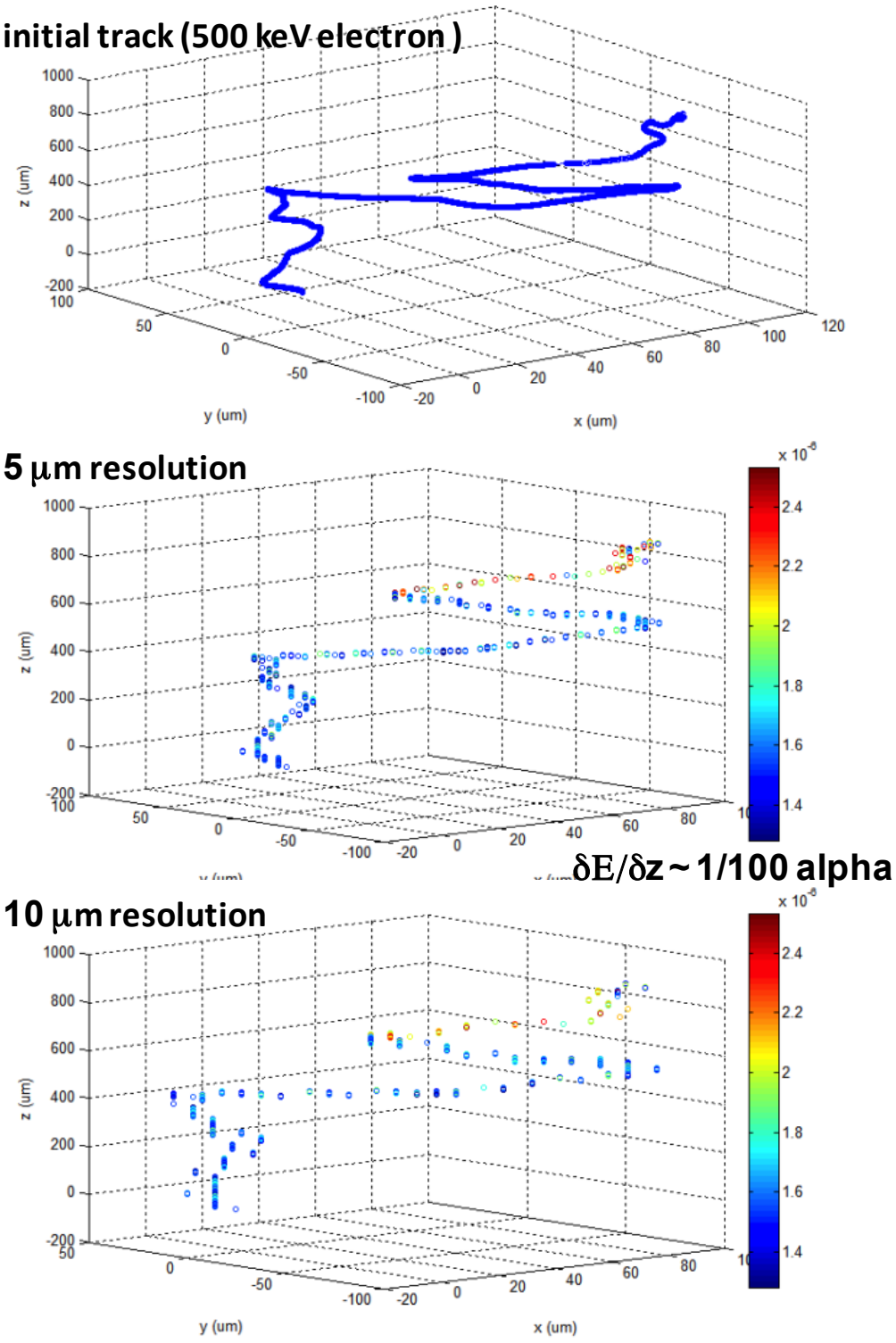


Fig. 4.28: Typical electron track for a 500 keV electron, and its mapping at two different resolutions one overlap effects are corrected. The depth distribution is mapped with pulses measured from alpha particle interactions and scaled to reflect gamma-ray depositions, and the initial track geometry is simulated in GEANT.

4.2 Angular Image Reconstruction in Advanced Compton Imaging

In this section, we review angular image reconstruction algorithms, including simple back-projection (SBP) and filtered back-projection (FBP), in advanced Compton imaging and quantify their degrees of enhancement provided by tracking the recoiled electron's direction. We have already reported some analysis results above. The investigation in this section is much more sophisticated than the previous one in the following five senses:

1) For SBP and FBP algorithms, the degree of improvement by electron-tracking is quantified in a more robust manner to statistical fluctuation by defining a new performance measure; distance-weighted root-mean-squared (DWRMS) error, which can capture not only the spread of point spread function (PSF) as full-width-half-maximum (FWHM) does, but also the RMS error of estimates. The DWRMS error allows us to avoid a Gaussian kernel fitting in order to measure FWHMs, thus reducing a potential source of model mismatch (in Gaussian kernel fitting) when reconstructed point-source images are sparse in a 4π sphere.
2) The FBP algorithm proposed by Parra [1] is modified in order to exploit the electron-tracking information in advanced Compton scattering imaging. Based on the modified FBP algorithm, the degree of enhancement provided by advanced Compton imaging is quantified and compared with that by SBP algorithm.
3) In SBP and FBP algorithms with electron-tracking ability, a different angular determination on a back-projected cone is assigned for each event. Note that, in our previous investigation of advanced Compton imaging, the directional determination of recoiled electrons was assumed 10^0 regardless of events and event-by-event variability was ignored. In this section, the segment on a back-projected cone is determined by the spatial resolution of a radiation detector and the energy of a recoiled electron for each event [2].
4) Instead of using a Fibonacci grid to plot the data (to ensure equal areas over the sphere), we display the reconstructed point-source images in the rectangular grid of (φ,θ), which is a more conventional way of displaying in the Compton scattering imaging literature [3,4,5,6]. To avoid well-known problems in mapping uniformly

distributed points on a unit sphere to the rectangular grid, we placed a point-source on the equator and simulated the data by MCNP5.

5) In order to obtain the better quality of reconstructed source images, we increase the resolution of the rectangular grid from 2^0 (previous code) to 0.25^0 (this section), requiring much more computational cost. We offset the increased computing cost by implementing SBP and FBP algorithms in Matlab, which provides much faster executions and more compatibility with C/C++ than Mathematica. (Note that the previous results were obtained and analyzed in Mathematica.) The Matlab implementation enables us to test a larger number of events (4000 data points) than the previous case (200 data points).

The remainder of this section is organized as follows. With electron-tracking information, we first qualitatively discuss the results obtained by the SBP algorithm and then consider the degree of enhancement provided by Parra's FBP algorithm. The performance evaluation of the modified Parra's FBP algorithm is conducted for three different finite orders of Legendre polynomials. Thirdly, we explain a newly defined performance measure; DWRMS error, and compare the conventional Compton imaging and advanced one quantitatively with it. Finally, we discuss some future work based on the presented results of advanced SBP and FBP algorithms.

4.2.1 Background on Back-Projection Based Algorithms

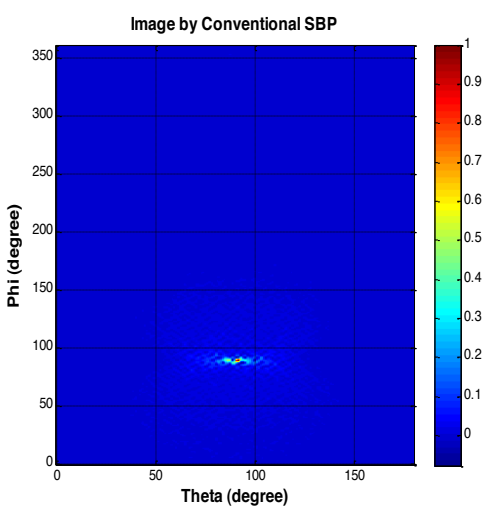
The SBP is probably the simplest and most intuitive algorithm to reconstruct an image of a radiation source distribution in Compton scattering imaging. In SBP, points on a back-projected cone are assumed uniformly probable and thus accumulating back-projected cones over all measured events produces a reconstructed source distribution of radiating gamma-ray photons. Since SBP uniformly weights all points on a back-projected cone, it equally treats all different gamma-ray traversal paths and thus typically generates a blurred and noise-contaminated point spread function (PSF). In FBP, for improving the quality of reconstructed images and avoiding unnecessary weighting on physically infeasible points on a back-projected cone, different weights based on the scattering cross-section are assigned to points along the boundary of the cone. Therefore, in general, FBP yields a sharper PSF with a smaller FWHM, but it requires more processing time than SBP. The imaging performance and execution time of FBP depends

on several factors, one of which is how to model an intrinsic system response, e.g. PSF. If an analytical PSF is used for FBP algorithm, it is usually requires less computational expensive than an empirical PSF, which can be obtained from nuclear particle interaction simulators such as MCNP and GEANT. Although FBP with an empirical PSF needs more processing time, the computation is done offline once and a look-up table is referenced for actual image reconstruction if the overall imager is reasonably approximated as a rotationally symmetric system. This relationship of SBP and FBP in terms of algorithm complexity, processing time, and reconstructed image quality is summarized in Fig. 4.29. Note that both SBP and FBP can be performed event-by-event, therefore real-time processing is possible, depending on the event-occurrence rate, i.e., arrival rate of impinging photons.

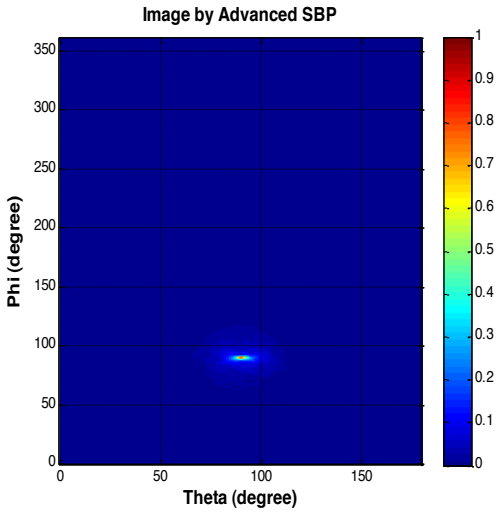
4.2.2 Simple Back-Projection with Electron-Tracking

As summarized above, the Matlab implementation allows not only the faster execution of the image reconstruction algorithms but it is also more compatible with other programming languages such as C/C++. It also provides more readiness for the radiation detection system integration since our other algorithms, for example, sequence reconstruction [3] and/or spectral deconvolution for chemical identification [10], have been already implemented by Matlab. In addition, much more external Matlab toolboxes and resources exist in the fields of signal processing and/or image processing than does Mathematica.

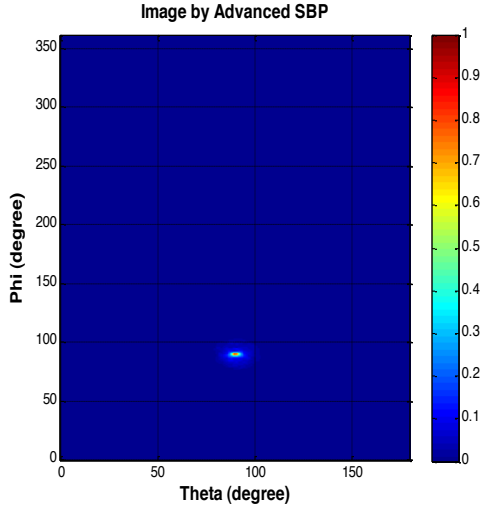
We apply SBP algorithms with and without electron-tracking ability to a dataset obtained from MCNP5, including 4000 two-interaction pairs. In order to synthesize the MCNP5 data, a true point-source, radiating 3 MeV gamma-rays, was assumed to be placed at $(\varphi,\theta) = (90^0, 90^0)$ on the equator, and a radiation detector was assumed to be $10m$ away from this radiation source. Note that at this energy, the performance is somewhat of a worse-case scenario since the gamma-rays are highly forward scattered and the angular uncertainty is therefore greater. Furthermore, the large number of events (4000) will tend to obscure the degree of improvement yielded by the extra direction information.



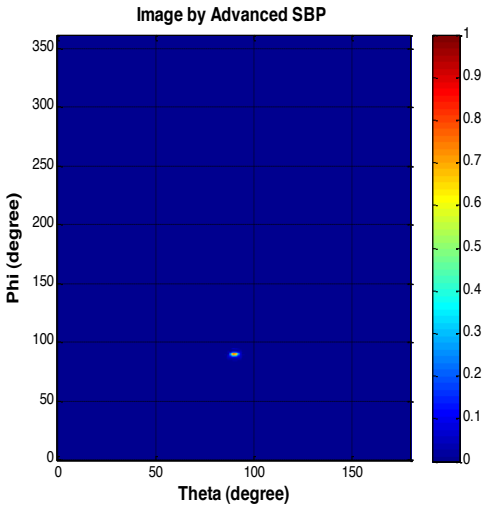
(a) Conventional SBP



(b) Advanced SBP w/ 10 μm



(c) Advanced SBP w/ 1 μm



(d) Advanced SBP w/ 0.1 μm

Fig. 4.29: The reconstructed point-source images by various SBP algorithms with different spatial resolutions of radiation detector for tracking recoiled electrons' paths: (a) by the conventional SBP without electron-tracking ability, (b) by the advanced SBP with electron-tracking and 10 μm spatial resolution detector, (c) by the advanced SBP with electron-tracking and 1 μm spatial resolution detector, (d) by the advanced SBP with electron-tracking and 0.1 μm spatial resolution detector. For all results by the conventional and advanced SBPs, the rectangular grid resolution was set to 0.25^0. Note that the true position of the point-source is picked by all SBP algorithms.

Two types of SBP algorithms (conventional and advanced ones) are implemented by Matlab, and were included in APPENDIX D. The reconstructed images by the

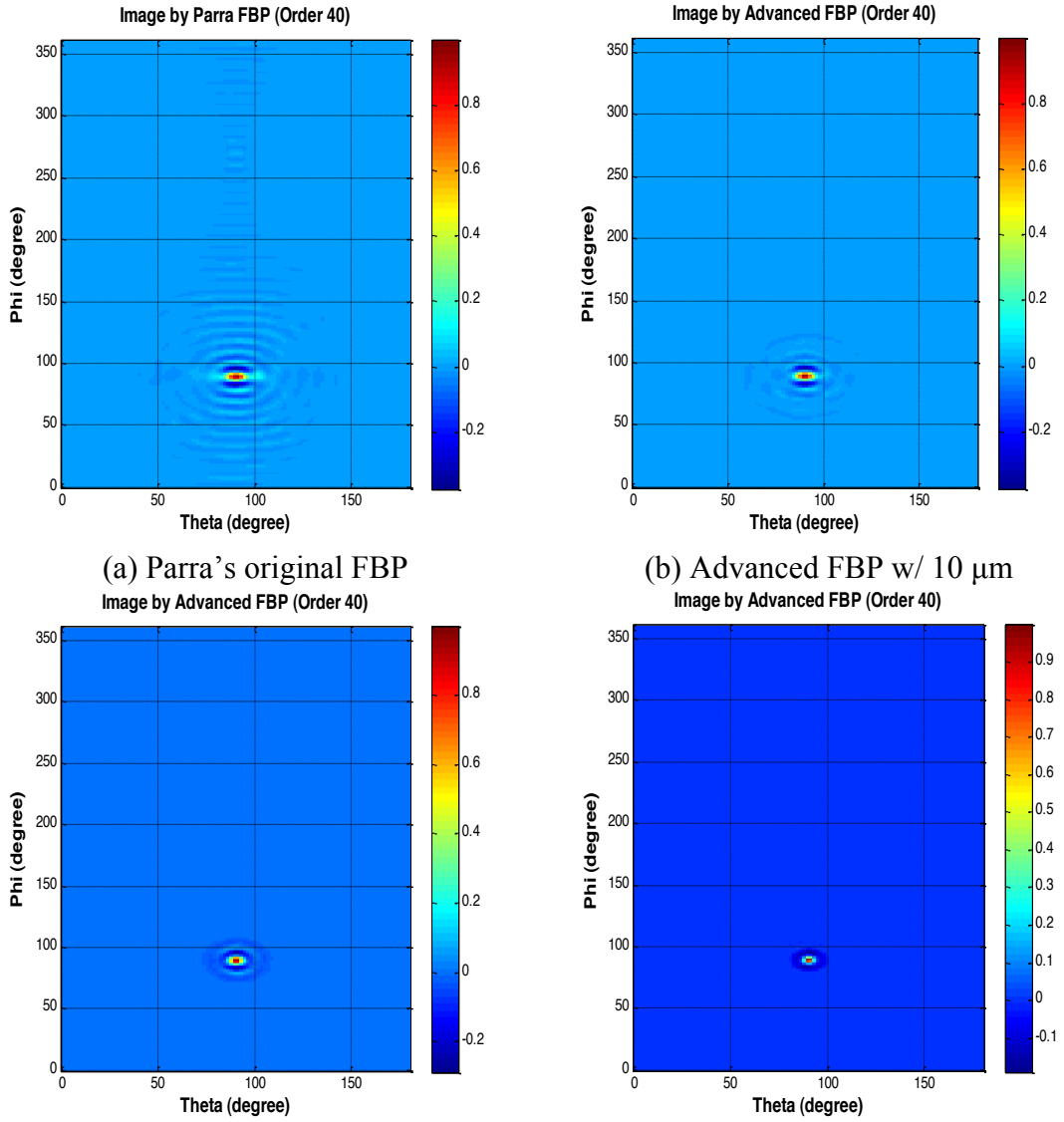
conventional and advanced SBP algorithms are shown on Fig. 4.29, where image intensities are normalized for the purpose of comparison. Three results are presented for the advanced SBP algorithm, depending on three different spatial resolutions (10 μm, 1 μm, 0.1 μm) of radiation detector, since different spatial resolutions of the radiation detector provide different angular uncertainty (in the sense of average) on back-projected cones. According to [2], on average, the finer the spatial resolution of a radiation detector is, the less angular uncertainty is achievable for a given energy of a recoiled electron. Due to the circuitous nature of recoiled electrons' paths [2], the spatial resolution of a radiation detector substantially affects the ability of determining the directions of impinging gamma-ray photons, i.e. the ability of confining the segments on back-projected cones for the image reconstruction of radiating sources.

As shown in Fig. 4.29, all SBP algorithms identified the accurate position of the true point-source located at $(\varphi, \theta) = (90^0, 90^0)$. From Fig. 4.29, we conclude that the advanced SBP algorithms with electron-tracking ability improve the performance of image reconstruction in terms of FWHM and the degree of enhancement increases as the spatial resolution of a radiation detector becomes finer for tracking recoiled electrons' paths. Since Fig. 4.30 shows only a qualitative comparison, the quantitative comparison with a newly defined performance measure, distance-weighted RMS error, will be provided in the latter section below.

4.2.3 Filtered Back-Projection with Electron-Tracking

We apply the Parra's FBP algorithm [1] and our modified version of it with electron-tracking to the same MCNP dataset analyzed in the previous section. The classical Parra's FBP and its advanced version exploiting electron-tracking information are both implemented by Matlab. According to [1], an infinite sum of spherical harmonic components constitutes the event response function (ERF) for each two-interaction event, which is practically infeasible to implement. From our experience with this MCNP dataset, ERFs based on Legendre polynomials up to the order of 40 provide good image reconstruction. Therefore we show the reconstructed images using only 40 Legendre polynomials in this section for qualitative comparison. For some more results based on other orders of Legendre polynomials, e.g. reconstructed images with 20 and/or 80 Legendre polynomials, the readers are referred to the DWRMS errors given in Table 4.2.

The Matlab codes for two types of FBP algorithms (Parra's original algorithm and our modified algorithm for tracking electrons' paths) were presented previously.



(a) Parra's original FBP

(b) Advanced FBP w/ 10 μm

(c) Advanced FBP w/1 μm

(d) Advanced FBP w/ 0.1 μm

Fig. 4.30: The reconstructed point-source images by various FBP algorithms with different spatial resolutions of radiation detector for tracking recoiled electrons' paths: (a) by the original Parra's FBP without electron-tracking ability, (b) by the advanced FBP with electron-tracking and 10 μm spatial resolution detector, (c) by the advanced FBP with electron-tracking and 1 μm spatial resolution detector, (d) by the advanced FBP with electron-tracking and 0.1 μm spatial resolution detector. For all results by the classical and advanced FBPs, the rectangular grid resolution was set to 0.25^0 and the order of Legendre polynomials was assumed to be 40. Note that the true position of the point-source is correctly detected by all FBP algorithms tested.

The reconstructed point-source images by Parra's original FBP algorithm and its advanced ones using 40 Legendre polynomials are shown on Fig. 4.30. In Fig. 4.30, all FBP algorithms identified the exact position of the true point-source, radiating 3MeV gamma-rays, located at $(\varphi,\theta)=(90^0,90^0)$. From Fig. 4.30, since subfigures in Fig. 4.30b- 4.30d by the advanced FBPs have less ripples in the side-lobes and narrower FWHMs in Gaussian kernel fitting than that in Fig. 4.30a by Parra's FBP, we conclude that the electron-tracking information provides the enhancement of image reconstruction in the qualitative assessment. The quantitative comparison using a newly defined performance measure, distance-weighted RMS error, will be given for various FBP algorithms in the latter section.

4.2.4 Distance-Weighted Root-Mean-Squared (DWRMS) Error

In this section, we define a new performance measure to compare the degree of improvement of several back-projection based image reconstruction algorithms provided by electron-tracking ability in advanced Compton scattering imaging. In general, the two most widely used classical performance measures in imaging processing and/or signal processing are probably root-mean-squared (RMS) error and FWHM. The RMS error can be defined as follows:

$$e_{RMS}=\sqrt{\frac{1}{N}\sum_{i=1}^{N}(f_i-\hat{f}_i)^2}, \tag{4.16}$$

where f_i is the true intensity value of a radiation source distribution at the ith pixel on a unit sphere and $\hat{f}_i$ is its estimate. N is the number of total pixels in the unit sphere considered. Note that the term inside the square root of Eq. 4.16 is an estimate of the sum of variance and squared bias of the estimate of image intensity. Therefore, the RMS error in Eq. 4.16 is used for measuring the variance and bias of the estimate of radiation source intensity. By assuming a one-dimensional Gaussian kernel, the FWHM is defined as follows:

$$FWHM=\sqrt{8\log 2}\sigma, \tag{4.17}$$

where σ controls the width of the one-dimensional Gaussian kernel. The FWHM indicates twice of the distance between the origin and the point where one-dimensional Gaussian kernel amplitude drops to the half of its maximum at the origin. In the case of two-dimensional image reconstruction, two different FWHMs are used for two

orthogonal directions. The FWHMs are typically used to measure the spread of PSF after image reconstruction is performed. So, the RMS error and FWHM capture two different aspects of image reconstruction algorithms in Compton scattering imaging.

To combine the benefits of the RMS error and FWHM into one measure, we here define a distance-weighted root-mean-squared (DWRMS) error as follows:

$$e_{DWRMS} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} d_i (f_i - \hat{f}_i)^2}, \tag{4.18}$$

where d_i denotes an Euclidean distance between f_i and $\hat{f}_i$ in the (φ, θ) plane. Note that the RMS error is a special case of the DWRMS error when $d_i = 1$ for all pixel locations. From the definition in Eq. 4.18, we notice that, if algorithm A provides a bigger FWHM than algorithm B, then it also produces a bigger DWRMS error than the algorithm B (Note that a point-source was used in the simulations of the previous sections in order to evaluate the SBP and FBP algorithms with and without electron-tracking ability.). Therefore, the DWRMS error can measure not only the variance and bias of intensity estimates but also the spread of PSF in reconstructed images in Compton scattering imaging. Note that computing the FWHM requires a Gaussian kernel fitting, whereas the DWRMS error does not impose any assumption on reconstructed images, providing more robust performance comparison when a radiation source distribution is sparse on a unit sphere than does FWHM. Thanks to its robustness and simplicity, we use the DWRMS error for the quantitative assessment of the improvement by advanced Compton scattering imaging.

4.2.5 Quantitative Assessment of the Enhancement by Advanced Compton Imaging

In this section, we present the quantitative assessments of the improvement obtained by tracking electrons' paths for the image reconstruction. With the same MCNP5 dataset reported in the previous two sections, we compute DWRMS errors for four different types of back-projection based algorithms (SBP verses FBP and with electron-tracking verses without electron-tracking). Their values of DWRMS errors are provided in Table 4.2 and 4.3.

Table 4.2: The DWRMS errors of reconstructed point-source images by several SBP algorithms with different spatial resolutions of radiation detector for tracking recoiled electrons' paths. For each column, the figures in the parentheses indicate the normalized DWRMS errors to that of the classical SBP, not exploiting the electron-tracking information.

	Rectangular grid Resolution (1.0^0)	Rectangular grid Resolution (0.5^0)	Rectangular grid Resolution (0.25^0)
No e-tracking	13.67 (100%)	19.56 (100%)	40.42 (100%)
With e-tracking $(10\mu m)$	6.07 (44%)	8.38 (43%)	16.60 (41%)
With e-tracking $(1\mu m)$	1.74 (13%)	2.30 (12%)	4.15 (10%)
With e-tracking $(0.1\mu m)$	0.84 (6%)	1.09 (5%)	1.57 (4%)

In Table 4.2, the DWRMS errors of the reconstructed images by various SBPs are shown for three different resolutions of the rectangular grid in the (φ,θ) plane. In Table 4.2, electron-tracking with a radiation detector whose spatial resolution is 10 μm provide about 58% reduction in terms of DWRMS error compared to the classical SBP algorithm, not exploiting electron-tracking ability. For finer spatial resolutions of radiation detector, 1 μm detector yields approximately 90% reduction, and 0.1 μm detector provides about 95% DWRMS error reduction compared to the conventional SBP algorithm. These reductions in terms of DWRMS error are consistent for three different rectangular grid resolutions tested in Table 4.2. Note that the finer rectangular grid resolution is used, the more execution time is required for implementing the SBP algorithm. In Table 4.3, the DWRMS errors of the reconstructed point-source images by Parra's FBPs [1] with and without electron-tracking ability are compared, where different orders of Legendre polynomials were used for the FBP algorithms. According to [1], an infinite sum of spherical harmonic components constitutes the event response function (ERF) for each two-interaction event, which is practically infeasible to implement. From our experience with the MCNP dataset, as shown in the figures on Fig. 4.31, Legendre polynomials up to the order of 40 provide good image reconstructions. So, we chose to test three different orders (n=20, 40, 80) of Legendre polynomials in this progress report. For more detailed discussion on selecting the order of Legendre polynomials for Parra's FBP algorithm, the

interested readers are referred to [8], arguing that more than 190 orders are necessary for image reconstruction. Note that the smaller order is used, the faster the Parra's FBP algorithm can be executed. In Table 4.3, electron-tracking with a radiation detector whose spatial resolution is 10 μm provided about 65% reduction in terms of DWRMS error compared to the original Parra's FBP algorithm without electron-tracking information. For finer spatial resolutions of radiation detector, 1 μm detector yields almost 90% DWRMS error reduction, and 0.1 μm detector provides about 95% DWRMS error reduction compared to Parra's classical FBP algorithm. These reductions in terms of DWRMS errors are fairly consistent for three different orders of Legendre polynomials tested in Table 4.3. For binning the rectangular grid in the (φ, θ) plane, the resolution was set to 0.25^0 for all scenarios in Table 4.3.

Table 4.3: The DWRMS errors of reconstructed point-source images by various FBP algorithms with different spatial resolutions of radiation detector and different Legendre polynomial orders up to 80. For all results by FBPs, the rectangular grid resolution was set to 0.25^0. For each column, the figures in the parentheses indicate the normalized DWRMS errors to that of Parra's classical FBP algorithm [1], not exploiting electron-tracking information.

	Legendre polynomial Order (n = 20)	Legendre polynomial Order (n = 40)	Legendre polynomial Order (n = 80)
No e-tracking	128.87 (100%)	63.20 (100%)	28.06 (100%)
With e-tracking $(10\mu m)$	45.67 (35%)	23.90 (38%)	9.32 (32%)
With e-tracking $(1\mu m)$	14.60 (11%)	9.14 (14%)	3.81 (13%)
With e-tracking $(0.1\mu m)$	5.08 (4%)	3.11 (5%)	2.42 (8%)

From the results shown in Table 4.2 and 4.3, we conclude that exploiting the electron-tracking information dramatically improves the quality of image reconstruction algorithms in terms of the newly defined DWRMS errors. Specifically, a radiation detector with a spatial resolution of 1 μm can provide approximately 90% reduction of DWRMS errors by using electron-tracking information for both SBP and FBP algorithms. Even with a radiation detector with coarse spatial resolutions, such as 10 μm, one can

achieve about 60% reduction of DWRMS errors by tracking recoiled electrons' paths and using their information in order to confine the segments on the back-projected cones in SBP and FBP.

4.3 References

[1] L. Parra, "Reconstruction of cone-beam projections from Compton scattered data", IEEE Trans. on Nuclear Science, Vol.47, No.4, pp.1543-1550, Aug. 2000.

[2] M.D. Hammig and R.C. Stevenson, "The degree of enhancement in a gamma-ray image gleaned from recoil-electron tracking", IEEE Nuclear Science Symposium Conference Record, N30-308, pp.2433-2439, Oct., 2008.

[3] Dan Xu, "Gamma-ray imaging and polarization measurement using 3-D position-sensitive CdZnTe detectors", PhD Thesis, Nuclear Engineering and Radiological Science Department, The University of Michigan, 2006.

[4] D. Xu and Z. He, "Filtered back-projection in 4-pi Compton imaging with a single 3D position sensitive CdZnTe detector", IEEE Transaction on Nuclear Science, Vol.53(5), pp.2787-2795, Oct., 2006.

[5] D. Xu and Z. He, "Gamma-ray energy-imaging integrated spectral deconvolution", Nuclear Instruments and Methods in Physics Research A, 574, pp.98-109, 2007.

[6] C.E. Lehner, Z. He, and F. Zhang, "4pi Compton imaging using a 3D position-sensitive CdZnTe detector via weighted list-mode maximum likelihood", IEEE Transaction on Nuclear Science, Vol.51(4), pp.1618-1624, 2004.

[7] H. Barrett, T. White, and L. Parra, "List-mode likelihood", Journal of the Optical Society of America, vol. A14, no. 11, pp. 2914-2923, Nov., 1997.

[8] T. Tomitani and M. Hirasawa, "Image reconstruction from limited angle Compton camera data", Phys. Med. Bio., Vol.47, pp.2129-2145, 2002.

[9] D.J. Lingenfelter, J.A. Fessler, and Z. He, "Sparsity regularization for image reconstruction with Poisson data", Proc. SPIE 7246, Computational Imaging VII, p.72460F, 2009.

[10] J. Noh, R.C. Stevenson, and M.D. Hammig, "Statistical methods for chemical compound identification from neutron-induced gamma-ray spectra", IEEE Nuclear Science Symposium Conference Record, Accepted for conference, 2009.

CHAPTER V

EXPERIMENTS PERFORMED AND DISCUSSION OF THE RESULTS

5.1 General Description: Evaluation of Proof-of-Concept

As discussed in previous chapters, one can use microelectronic fabrication techniques to construct geometries for which the pulse transit time is nonnegligible and further, provides a measure of the position-of-interaction. If the transmission line is properly designed, then the time difference between the pulses can then be used to measure the position at which the charge was collected. Furthermore, if the timing uncertainty is significantly less than the transit time along a single strip, then the position can be isolated to one strip of the pattern or better. The attractiveness of such a scheme is its relative simplicity since the readout of dozens or hundreds of separate strips can be replaced by the time-sensitive readout of only one or two contacts *per side* of the detector. Meander designs were fabricated with strip multiplicities ranging from 2 to over 3000. Recall that the lateral position orthogonal to that implied in the figure can be deduced by fabricating a similar structure on the detector backside, but with the strips in a perpendicular direction.

The signal can suffer loss and distortion, depending on both the nature of the signal and the strip-line design. In fact, simulations using the radio-frequency (RF) electromagnetic (EM) software Sonnet indicated that signal loss was expected, due to dielectric loss and impedance discontinuities at the ends of each line. Thus, we used this knowledge to test the signal propagation and ensure that unknown sources of loss and distortion were not present; in particular, to test the degree to which capacitive coupling took energy out of the signal.

5.2 Signal Readout with Preamplifiers and Cascade Amplifiers

5.2.1 Planar Detector Testing with High speed OP AMP

In order to: a) calibrate the timing system, and b) evaluate different preamp and amplifier designs, planar detectors were used, in configuration such as those shown in Figs. 5.1, 5.2 and 5.3.

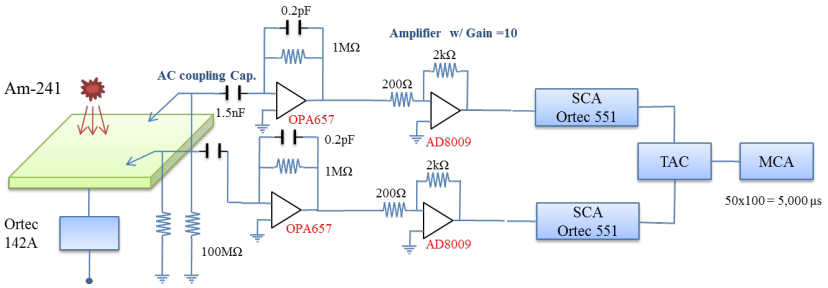


Fig. 5.1: Block diagram for the baseline fluctuation testing with two preamplifiers, amplifier, single channel analyzers and a TAC (JFET first stage with all bandwidths in the 1 -2 GHz range).

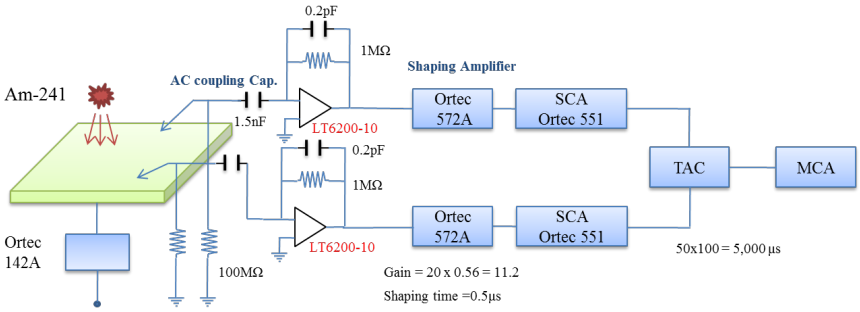
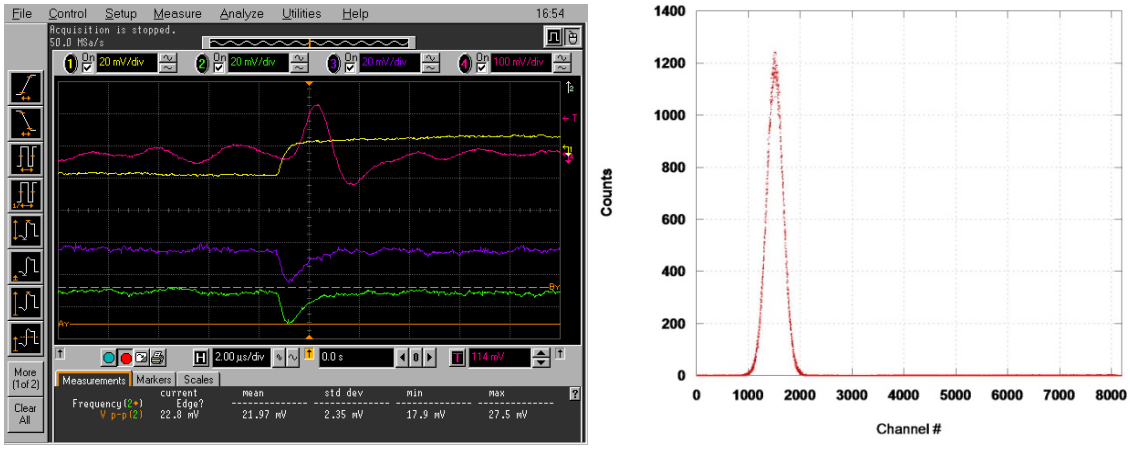


Fig. 5.2: Schematic of control experiment with different read-out architecture (two preamplifiers, shaping amplifiers, single channel analyzers and a TAC (preamps based on LT6000-10 with rail-to-rail input and with 1.6GHz bandwidth)).

(a) (b)

Fig. 5.3: Picture of the detector system (a) and two preamplifier circuits (b).

Examples of the signals extracted from the detector electrodes, for the Fig. 5.3 configuration, are shown in Fig. 5.4. The custom preamplifiers (outputs in purple and green) were designed with faster shaping than the commercial unit (output in yellow) due to the lower feedback capacitance (0.2 pF) and resistance (1 MΩ), although they exhibited unacceptably high noise. Nevertheless, their performance was sufficient to test the proof-of-concept (POC) device, Meander01. When the alpha-interaction is isolated from the noise, time spectra such as that shown on the right of Fig. 5.4b are produced.



(a) (b)

Fig. 5.4: Cathode signal from commercial preamp (Ortec 142A (yellow)), and anode signals from two custom preamplifiers (purple and green) and from shaping amplifier (red) (a), and baseline time spectrum from TAC (b).

Using the detector configuration in Fig. 5.2, the temporal response of the system was calibrated, as shown in Fig. 5.5, in which fast signals with moderate noise produced a FWHM of 63.2 ns. This rather large value is due to the small size of the signals relative to the noise fluctuations; nevertheless, it was sufficient to allow us to test the validity of the time difference method. As will be explained in Section 5.2.2, the achievement of 10s of picoseconds or better rather than 10s of nanoseconds in our next generation design was expected, but for proof-of-concept testing, the DIP-package custom op amps were acceptable, as will be detailed below.

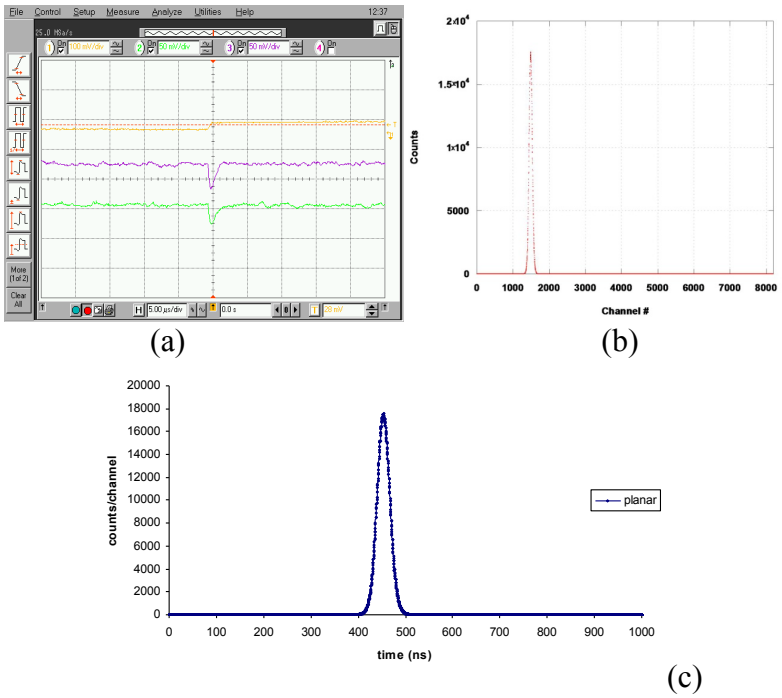


Fig. 5.5: Signal from Ortec 241A (yellow) and from preamplifiers (green and purple) (a), baseline time spectrum from TAC in channels (b), and same time spectrum in nanoseconds (c).

5.2.2 Meander Detector Testing with High speed OP AMP

In order to demonstrate that the signals measured from the two ends of a delay line could provide a measure of the position at which the charge was collected, two tests were used. First, the time difference between the signals should provide a measure of the position of interaction. Second, the pulses' amplitude-differences can be combined with their time-differences to gauge the device's operation.

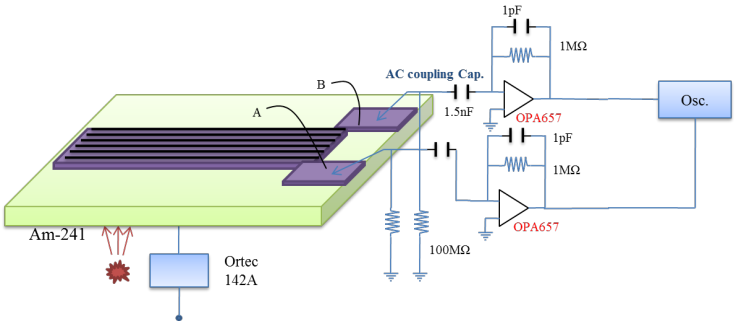


Fig. 5.6: Block diagram for the baseline fluctuation testing with two preamplifiers for a time difference fluctuation.

An electromagnetic-based velocity selector and fast-electron steering unit was constructed so that the position at which radiation impinges can be controlled. Furthermore, a collimated laser was used to develop the readout electronics, and such a device can be used to achieve micron-scale electron-hole injection. Nevertheless, for the proof-of-concept tests discussed below, the detector field with alpha particles has been flooded; thus, each interaction can occur randomly across the face of the meandering line.

With this fact in mind, consider a schematic of the experimental setup, as shown in Fig. 5.6, along with the strip design and its realization shown in Figs. 5.7 and 5.8, respectively. With various amplifier circuits and different strip multiplicities, the pulses were examined that emerge at the readout pads following each radiation interaction.

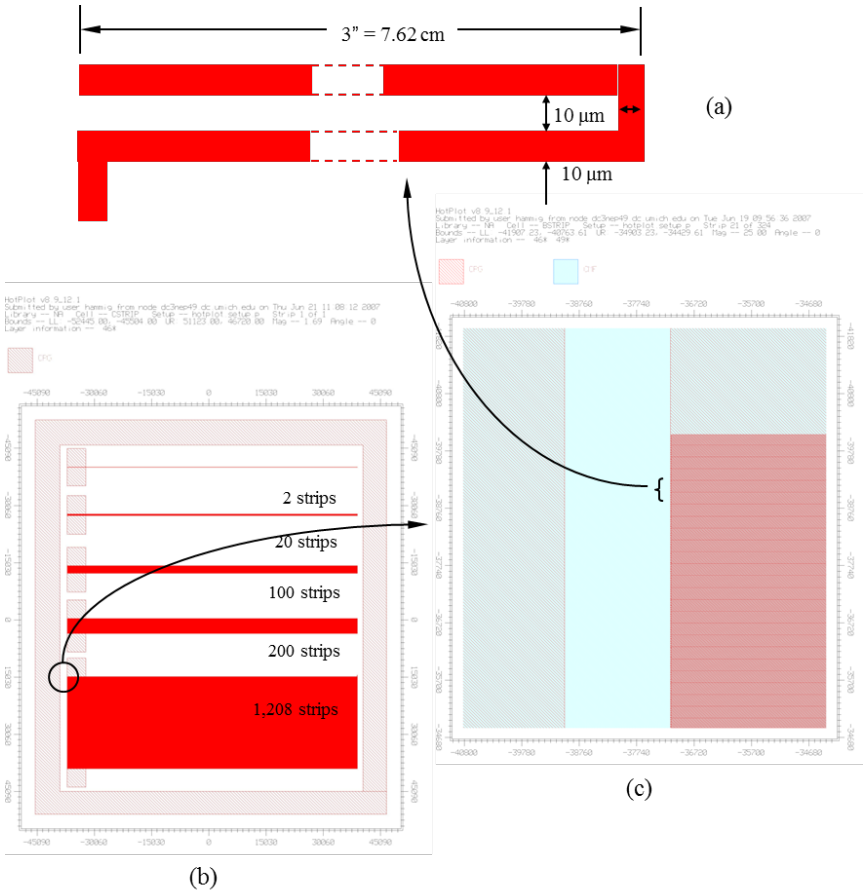


Fig. 5.7: Mask design for the diagnostic microstrip detector. (a) The geometry of the primitive, repeated in (b), contains designs consisting of 2, 20, 100, 200, and 1,208 strips.

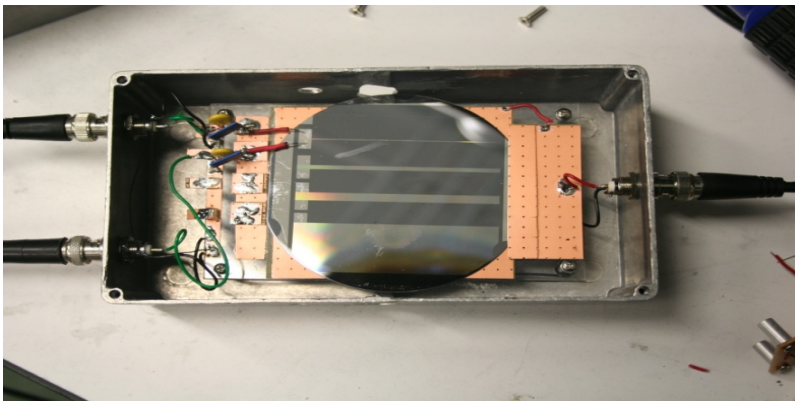


Fig. 5.8: Picture of the Meander01 detector system.

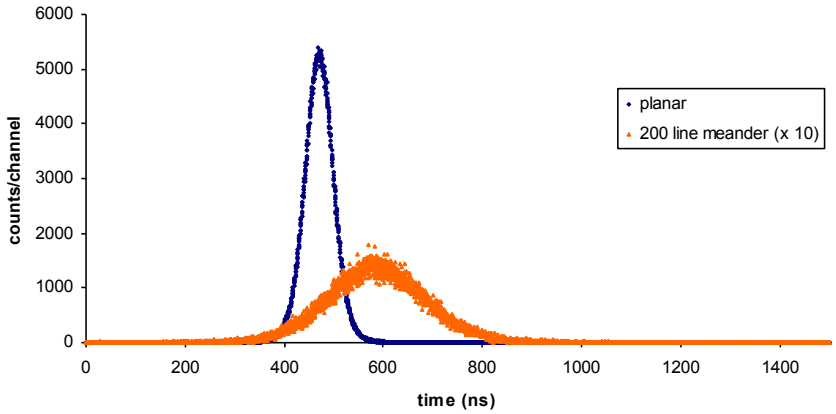


Fig. 5.9: Time spectra of the baseline-planar detector compared with the transit-time detector, demonstrating time differences consistent with those expected.

If radiation impacts nearer pad A (see Fig. 5.6), then that pulse appears sooner and, for the attenuating strip designed for Meander01, bigger than that measured from pad B. In order to test this expectation, both the pulse amplitudes and time differences via constant fraction timing were measured, for the 200 strip design, the results from which are shown in Figs. 5.9, 5.10, and 5.11.

In Fig. 5.9, the time spectra for the planar and meander design was shown, both using the Fig. 5.2 circuit. As shown in the baseline measurement, the time resolution is poor (347.3 ns FWHMt) but sufficient for determining the time differences inherent in the meander design. In particular, the overall time resolution of 925.4 ns implies an end-to-end *transit time* of 428 ns, consistent with a signal propagation velocity of 0.12 c, which is about what we expect. Note that because of the different geometries- the alpha

source was farther from the detector in the meander case and collimated to impact the strips- the coincident count rate was lower for the meander design. Further, the time shift shown in Fig. 5.9 is due to differences in the pulse shape between the cases.

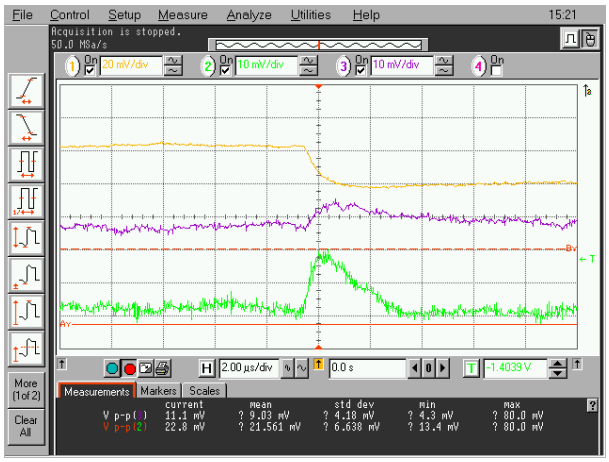
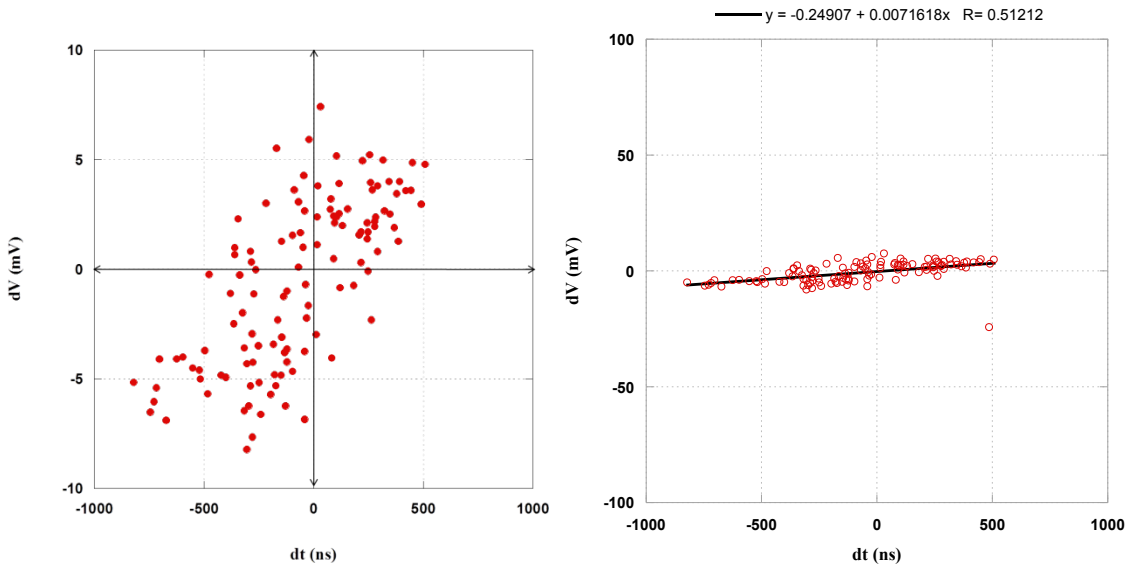


Fig. 5.10: Signal from Ortec 142A (yellow) and from preamplifiers (green and purple) from two readout pads of the meander. Note that the earliest pulse (in green) is larger and crisper than the delayed pulse (in purple).

The time spectra thus showed that the pulses contained the transit time information consistent with the geometry of the detector. In order to buttress this finding, the relationship between the pulse amplitude and the time difference was further examined, expecting those events that occur nearer the readout pad to produce a bigger response. For instance, consider Fig. 5.10, in which oscilloscope waveforms are displayed from the two readout pads and from the Pt cathode on the opposite detector face. The earlier signal (in green) is better defined and larger than that from the later signal (in purple), which is consistent with the expected pulse attenuation due to the imperfect impedance matching at the discontinuities in the first generation meander design. Note that for later generation detectors, we expect to largely eliminate this pulse shape attenuation.

A summary of the data is shown in Fig. 5.11, in which the difference in pulse amplitudes ($V_A - V_B$) is plotted versus the pulse time difference ($t_B - t_A$). We expect that if V_A is larger than V_B (positive ΔV in Fig. 5.11), then the V_A pulse will precede the V_B pulse and Δt will also be positive, which is what we observe. Note that the overall time differences are also in line with our transit time difference measurement of 428 ns, determined above.



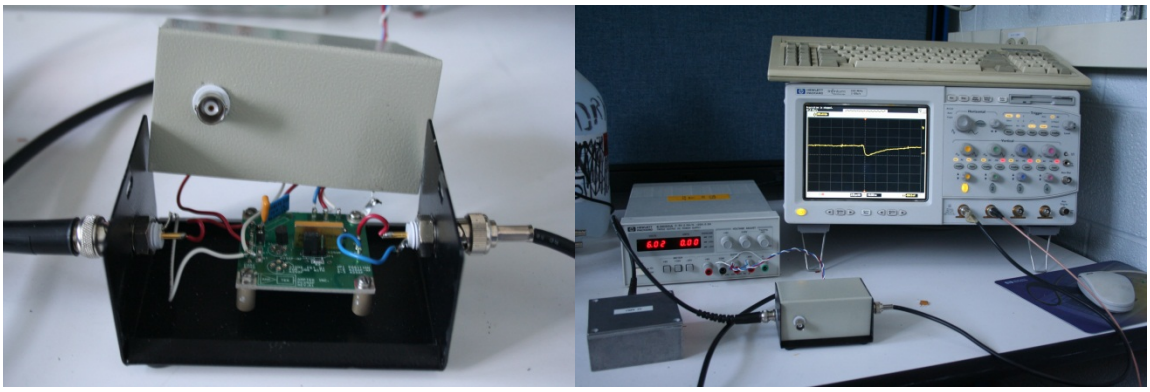
(a) (b)

Fig. 5.11: Plots of the voltage difference ($dV = V_A - V_B$) vs. the time difference ($dt = t_B - t_A$).

Our decision to overbuild the detector, in the form of long 3 inch strips thus enabled us to demonstrate the expected transit-time performance in the presence of an imperfect geometric design and relatively noisy and slow amplifier components. Our efforts to improve upon both areas are the subject of the next section, in which we expect orders-of-magnitude improvement in both timing resolution and noise mitigation.

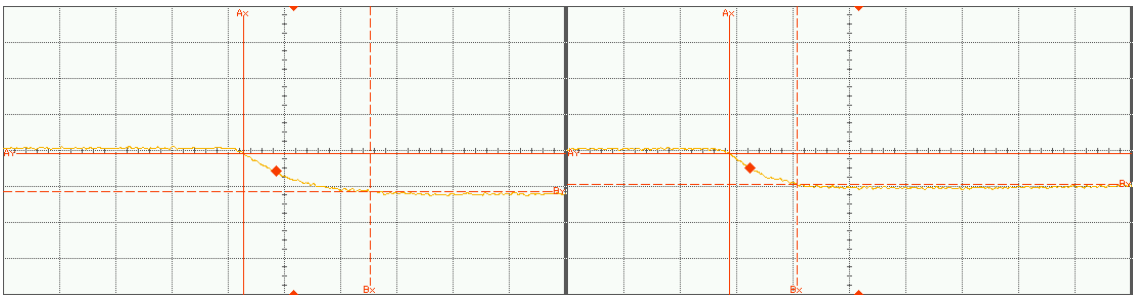
5.2.3 Depth-Sensing via Charge-Sensitive Amplifiers

Although the main focus is on optimizing the timing properties of the readout electronics using RF components, amplification technologies specifically designed for the low-noise readout of semiconductor detectors are also investigated. Specifically, in order to acquire the information for depth sensing using two steps of differentiation, readout circuits based on the A250 preamplifier from Amptek were made, which has ultra low-noise and a fast rise-time (4ns at 0 pF). Furthermore, it can perform positive or negative signal processing in contrast to the single-polarity Sirenza device described in the subsection above.



(a) (b)

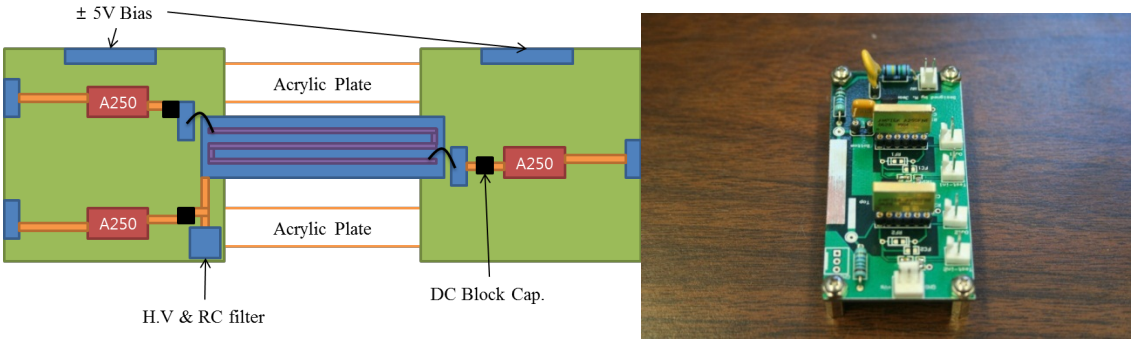
Fig. 5.12: Test board for A250NF (a) and the test environment (b).



(a) (b)

Fig. 5.13: Signals acquired with 0V bias (a) and 400 V bias (b), when the time setting was 200 ns/div and voltage setting was 200 mV/div.

In order to test the amplifier's performance, a test board was made with one A250NF and one planar detector having Au and Al contacts. Fig. 5.12a shows the test board for the A250NF and Fig. 5.12b shows the test environment. When the reverse bias voltage was 0 V, the rise-time was about 250 ns which reduced to 150 ns at 400 V reverse bias, as shown in Fig. 5.13. Most of the reduction is due to the decreased capacitance of the detector. Note that the SNR is again quite good for this device, but the die-form TIA performance of the previous subsection is more impressive because of its superior timing and noise performance, given the wider frequency band admitted by the device.



(a) (b)

Fig. 5.14: Layout of the meander readout using the high-end commercial charge-sensitive preamplifiers (a), and the assembled circuit (b).

In order to extract the time information from the signals, readout circuits were designed and made. Fig. 5.14 shows the diagram of the readout circuit for acquiring signals from two top and bottom sides of the detector, as well as the schematic, layout, and assembled circuit.

Thus two related but differing methods of extracting the induced signals from the delay-line detectors were investigated. The novel time-based technique was deemed to be the most promising, but its performance was compared with results using established readout techniques, like those described in this subsection. In both cases, the main hurdle that hampered the operation of the previous proof-of-concept devices has been overcome; specifically, an inadequate front-end readout stage particularly in terms of noise. Going forward, the electronic and mechanical implementation of the preamps will continue to be refined, and then applied to the delay-line devices.

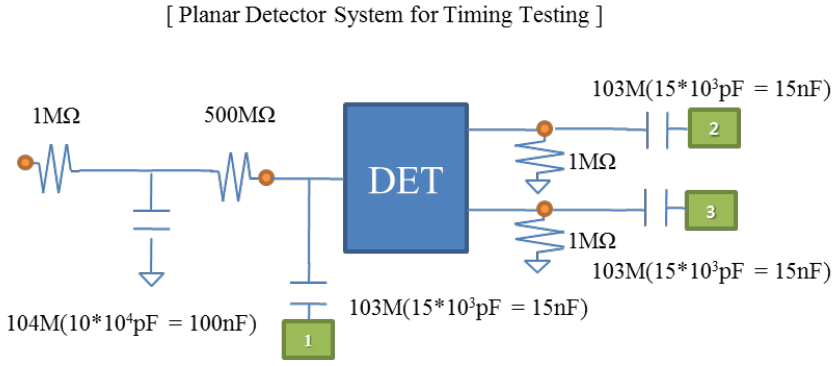
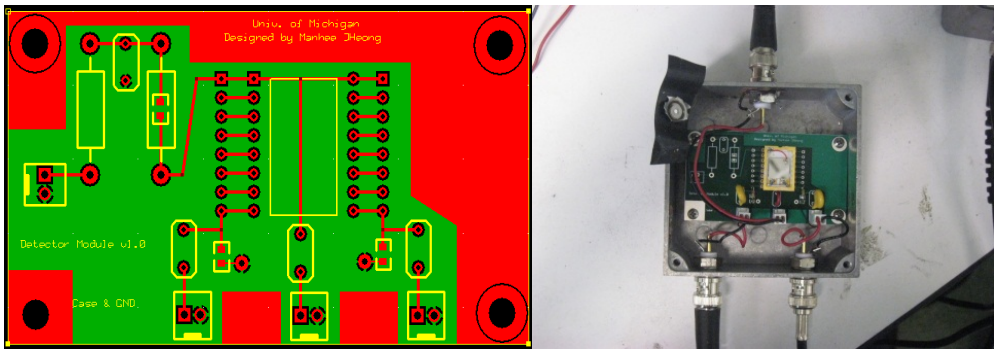


Fig. 5.15: The diagram of the detector module with three outputs from one cathode and two anodes.



(a) (b)

Fig. 5.16: The circuit schematic for the detector and three outputs (a), and its physical realization (b).

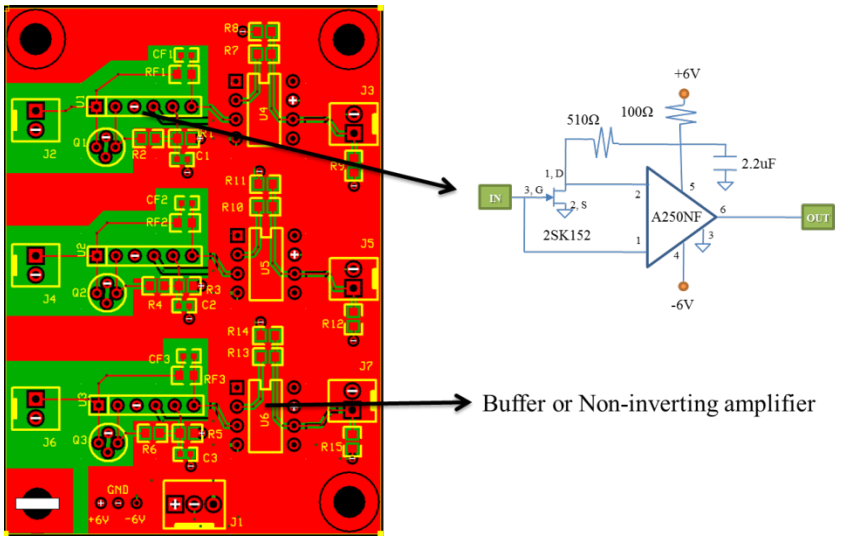


Fig. 5.17: The circuit layout for the preamplifier and its configuration.

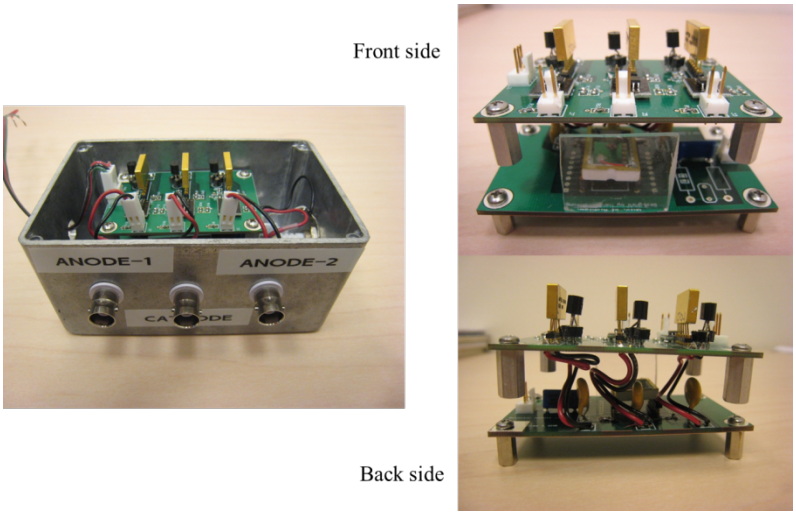


Fig. 5.18: The physical layout of the readout circuit module.

In order to measure the signal from each detector module, readout circuits for the detector and the preamplifier modules were designed, as shown in Figs. 5.15 and 5.16. In order to convert the induced charge into a voltage signal, the A250NF from AMPTEK Inc., a charge sensitive preamplifier was used, which is designed for low leakage-current semiconductor detectors. It possesses an external FET to increase the current gain, as shown in Fig. 5.17, which can be controlled by varying the drain-to-source current (I_{DS}). The external 1 kΩ resistor from pin 2 to pin 5 provides the FET with a drain current of 2.75 mA, and further reductions in the resistor increases I_{DS}, which results in a large current gain when the input current is small. In our detector however, the input current is sufficiently high that reductions in the external resistor value were required, typically to 510Ω.

Fig. 5.19: Photograph of system configuration (top left). The output signals from cathode (yellow), anode (green) and two SCAs (top right). Close-up view of the signals (bottom left). Zoomed-in signal of the anode (bottom right).

With the detector module and preamplifier circuits of Figs. 5.15-18, we acquire signals from one cathode and two anodes, such as the traces shown in Fig. 5.19. The rise time of the signals was about 500 ns for the planar devices and 250 ns for the meander-line detectors. In order to measure the time differences between the two sides of the anodes, we used timing single-channel analyzers (SCAs) and a time-to-amplitude converter, as shown in Fig. 5.20. The resulting time spectrum for a planar detector, and therefore a measure of the limiting time resolution of the readout electronics, is shown in Fig. 5.21, yielding a time resolution of 36 ns, as measured at the FWHM,. Thus, even though this system isn't designed for fast timing but rather good SNR, the time resolution is approximately a factor of 10 better than our results from the Section 5.2.2. Improvements of at least 100 times better than this for current-sensitive devices can be achieved.

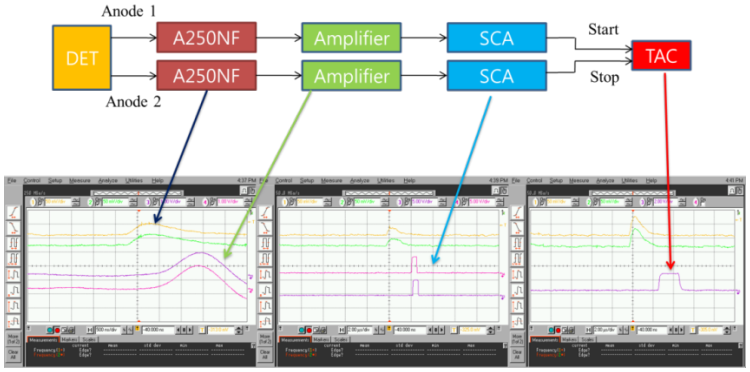


Fig. 5.20: Diagram of the readout system as well as representative signals.

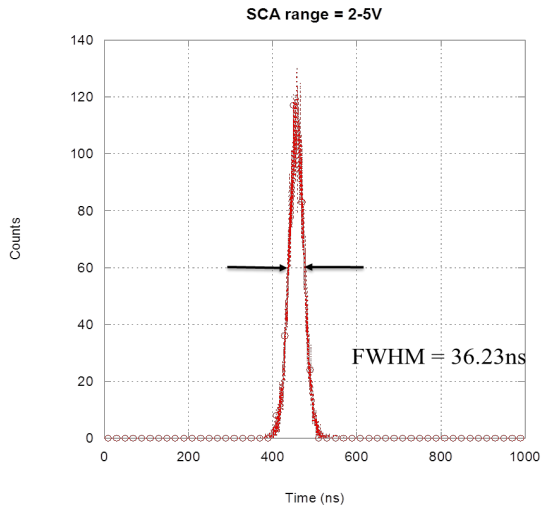
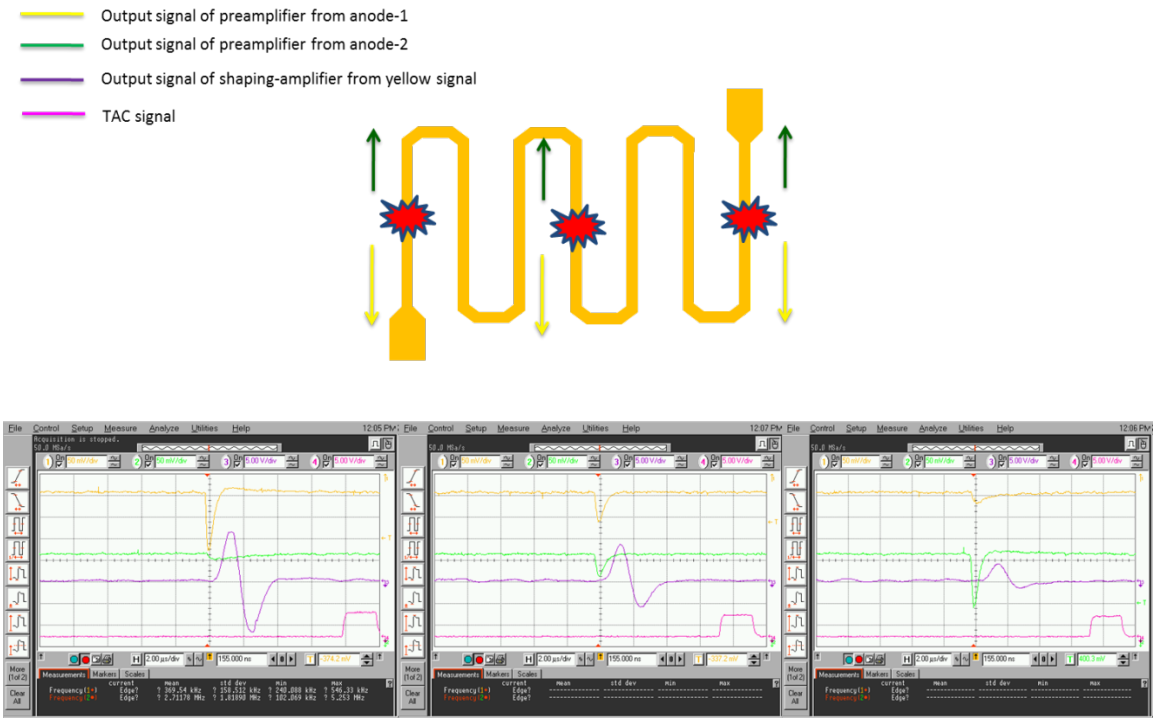


Fig. 5.21: Time spectrum of the planar detector system, yielding 36 ns FWHM time resolution.



(a) (b) (c)

Fig. 5.22: Signals from the 7-line meander detector (100 μm strip width and gap), when the interaction occurs under: (a) the left-mode strip, (b) the middle strip, and (c) the right-most strip.

The clean signals of Figs. 5.19 and 5.20 resulted in better timing resolution, as shown in Fig. 5.21, but they also allow us to measure the degree of signal attenuation along the meandering transmission line. In particular, the readout circuit was used to monitor the signals from the two ends of the meander-line anode and can infer the point of alpha-particle interaction by observing the size of the pulses, as shown in Fig. 5.22 for the 7-line meander with 100 μm strip width and gap. When the signal from one end is larger, that observed on the other end is proportionally smaller. Given the lossy design of the transmission line at the corners, one expects that those interactions under a given line will produce pulses of similar size and therefore, the energy spectrum should reflect the number of strips in the meander-line. This expectation is realized, as shown in Fig. 5.23, for the 5-line meander detector. Note that the peaks are not regularly spaced, nor are they uniform in size, which reflects both the non-uniform impingement of 5.5 MeV alpha particles on the device, as well as the non-linear attenuation along the line.

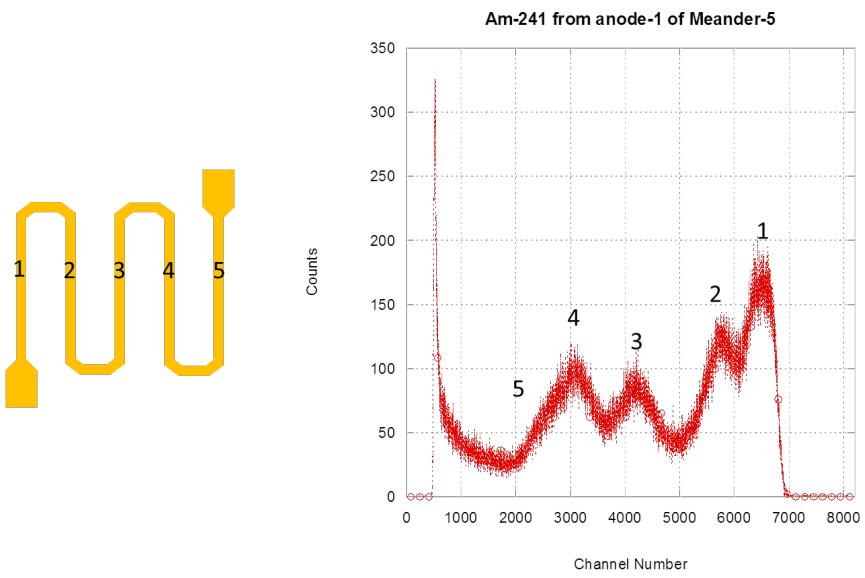


Fig. 5.23: Energy spectrum of 5-line meander detector system impinged by an Am-241 source, exhibiting 5 attenuation peaks.

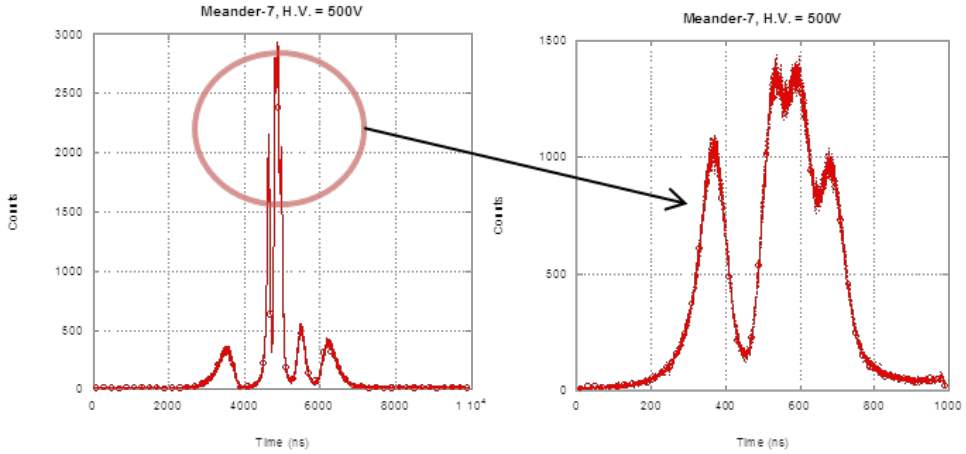


Fig. 5.24: Time spectrum of the 7-line meander detector system, with 7 peaks.

The energy spectra, such as those shown in Fig. 5.23, result in concomitant time spectra, because the amplitude differences map into amplitude-walk time differences, an example of which is shown in Fig. 5.24 for the 7-line meander. For the standard time spectra setup, the charge-sensitive configuration examined fails to provide the timing information necessary to identify the interaction location, but the amplitude does provide such information. One can design the line to attenuate the pulse with precision at the corners and then use the signals from the ends of the line to determine the strip on which

the charge is collected. However, given charge-sensitivity concerns, one does not want to sacrifice any of the signal amplitude, and from a practical standpoint, one may have an application in which 10,000 or more strips comprise the detector.

In these dense cases, implementing an amplitude-based, position-sensitive transmission line is not expected to be as sensitive as time-based methods, especially since one must separate the attenuation effect from the variations in the source energy. As shown in the simple 5-line example of Fig. 5.23, the five-lines are not cleanly distinguished in part because of the broadness of the thick alpha-particle source. For gamma-ray interactions, one must deconvolve partial energy transfers from attenuation effects. As we will show in the next section, these issues do not compromise the time-based methods and we will therefore pursue their implementation.

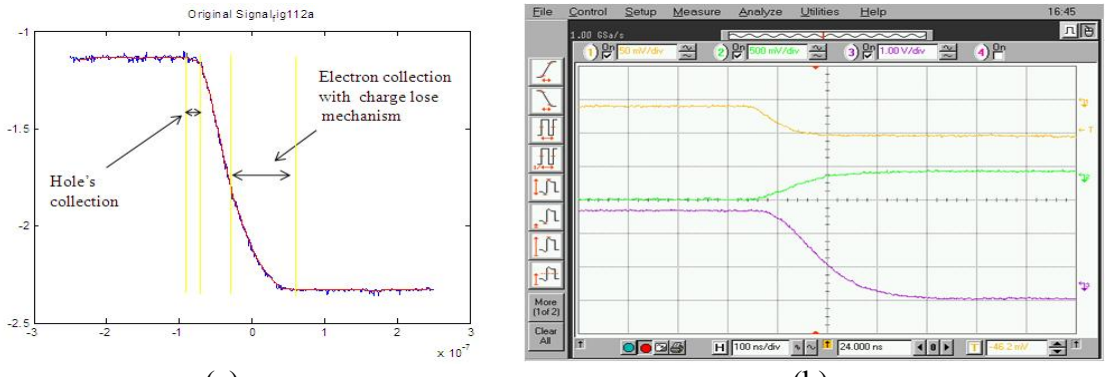
5.2.4 Depth Sensing via Analog Components and Digital Processing

Recall that one can measure the depth at which a charge-accumulation is created by either comparing the cathode and (partitioned) anode signals or by timing the charge drift through the known E-field geometry. The slope of the induced current signal can be used to extract the charge distribution because it measures alterations in the drift state and, in particular, the abrupt collection of the charges at the electrodes. Thus, the time of charge collection can be used to measure the depth from which the carriers originated, and the magnitude of the change measures the number of carriers that are removed from the drift stream.

The position resolution of such a scheme can be limited by either the timing or the charge resolution; however, the resolution loss due to timing can be mitigated by slowing the carriers near the surface, without , a substantial time penalty if thin "braking layers" are employed. For saturated electron or hole drift (~10^5 m/s), that is, for no E-field reductions, the second generation design (referred to above) is projected to have a depth resolution of 4.3 μm (if Δ_t = 33 ps) or better (sub-micron if Δ_t is in the ps range), if limited solely by timing. Thus, E-field reduction is not expected, via detector doping for instance, which is important if one wishes to extend the techniques to other materials of interest; in particular, if the detection methods are to be extended to advanced semiconductor materials for which dopant and oxidation chemistries are less well-established.

Therefore the use of time-based methods for extracting the depth profiles was investigated. In particular, both digital and analog signal-processing techniques were investigated in order to extract the initial charge distributions. Analog differentiation techniques can be costly from a SNR standpoint; however, for signal bandwidths greater than 10 GHz, digitization hardware can be costly. With somewhat slow signals (1-2 GHz bandwith), digital techniques were more effective; however, the purpose of the work was to establish the methodology of both so that they can be applied to the faster signals expected from the next-generation devices.

Note that the successful realization of a fast system may result in the position resolution being limited by the detector's charge resolution. For instance, if 3 ps timing resolution is achieved (1/10 the projected rise time of the system), then the time-limited position resolution is 0.4 μm. However, a fast electron will lose, in general, about 1 keV per micron in silicon; thus, about 276 electron-hole pairs per micron will be created. For instance, if the amplifier can only distinguish 400 electrons above the noise then its noise will limit the position performance. Note that for the 2nd generation TIA, the equivalent noise charge (ENC) of the preamp is less than a single electron for planar detector configurations and 2.7 electrons for the higher-capacitance meandering line; thus, the amplifier does not limit depth resolution via inadequate charge discrimination. The goal is to limit these systematic factors, so that physics limited effects- most notably, diffusion- will limit the ability to back-project the charge to its point of creation.



(a) (b)

Fig. 5.25: Acquired current signal with preamplifier (a) and its amplification (purple, G = 50) (b).

For signals from a *charge-sensitive* amplifier, the charge collection characteristics are most clearly revealed in its second derivative. Fig. 5.25a shows the preamplifier signal acquired following an alpha-particle event incident on the cathode, and Fig. 5.25b shows its amplification with gain of -10 (green) and 50 (purple). In Fig. 5.25a, the deviation of the straight line response at the beginning of the signal reveals, in part, the hole collection characteristics at the cathode, followed by the straight-line slope during electron drift, and finally, the deviation in the signal due to electron collection. In order to definitively assign these variations to the physical causes of interest, one must eliminate any circuit effects and, most particularly, the RC integration due to the detector capacitance and the feedback resistor, which can be accomplished with a small (0.1 pF) capacitance in series with the detector.

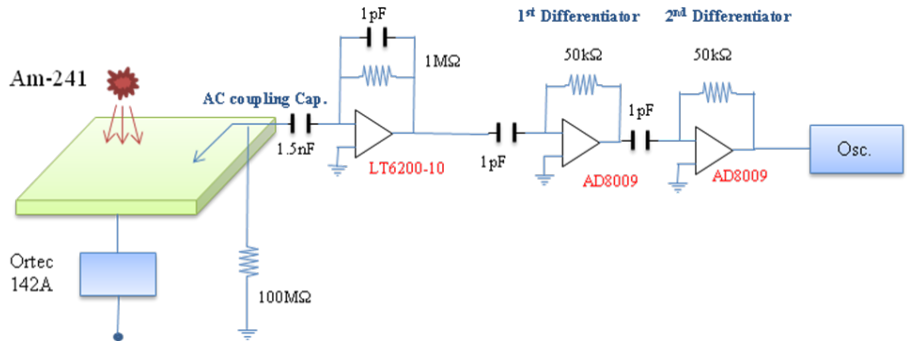
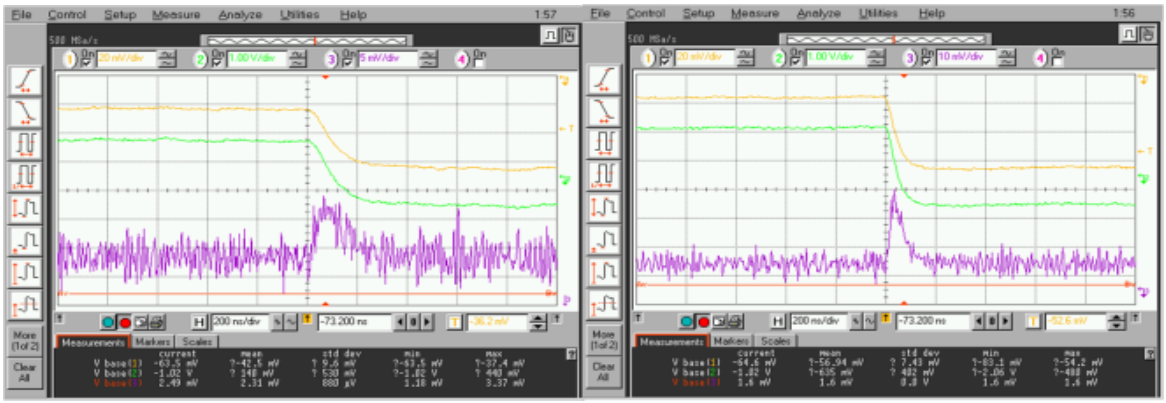


Fig. 5.26: Block diagram for testing the depth sensing with a two-stage differentiator and the planar type Si detector.

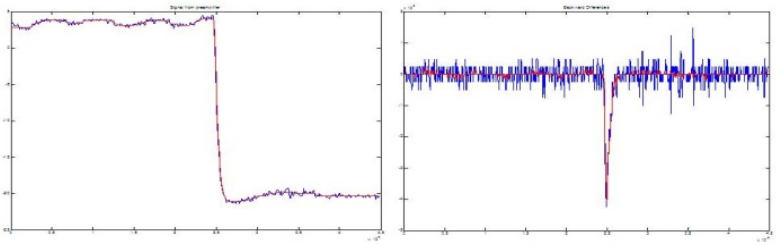


(a) (b)

Fig. 5.27: Acquired signal of preamplifier (yellow), amplified signal (green) and its (analog-circuit) differentiated signal (purple) with H.V. = 100 V (a), and with H.V. = 500 V (b).

The main purpose in the following was merely to determine the effectiveness, in terms of SNR, or to extract the charge distribution from the amplified signal, using either digital or analog techniques. Fig. 5.26 shows the block diagram for testing the signal processing with two stages of differentiation, using our planar silicon detectors.

Fig. 5.27 shows the amplified signal and its first derivative with different bias voltages, in which CR op amp differentiation was employed. In order to compare this analog conditioning with digital techniques, a differentiator using a MATLAB code was also applied. As shown in Fig. 5.28, the signal acquired from the preamplifier (blue) was either digitally filtered (red) (Fig. 5.28a) or used to form the first derivative (Fig. 5.28b in raw (blue) and filtered (red) forms). Using the filtered and differentiated signal, another differentiation provides a measure of the changes due to hole and electron collection if the circuit shaping effects are removed. As shown in Figs. 5.29 and 5.30, the initial Bragg-curve distribution of the charge cloud can be discerned for slow-drift conditions during electron collection (end of the pulse) and hole collection (beginning of pulse).



(a) (b)

Fig. 5.28: Signal of preamplifier (blue) and digitally filtered signal (red) (a). First derivation from preamplifier (blue) and filtered signal (red) with H.V. = 500 V (b).

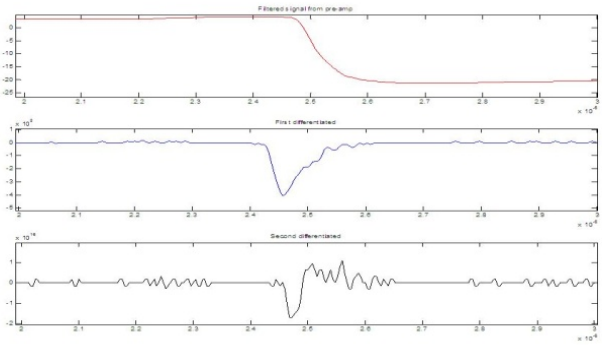


Fig. 5.29: Filtered signal from preamplifier (top), its first signal derivation (middle), and it is the second derivative (bottom), using digital techniques.

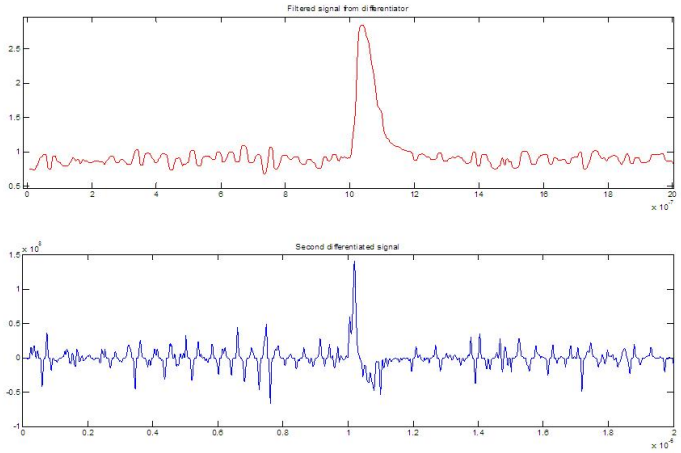


Fig. 5.30: Signal from differentiator (top) and its digitally-derived derivative (bottom), from which the initial charge distribution can be measured.

The conclusions are thus that *fine* depth sensing requires higher speeds and lower noise than that provided by commercial DIP-configuration amplifiers. The die-form transimpedance device and its relatives can provide that performance. Further, digital techniques were more effective in extracting the second derivative because the attenuated pulse due to analog differentiation resulted in poorer signal-to-noise. Nevertheless, for subsequent designs, high-speed analog components should be used so that most of the amplification and conditioning steps will be done on-chip. Digital processing may be necessary to facilitate the deconvolution of the characteristic response from the non-uniform electrode pattern.

5.2.5 Meander Detector Testing With High Bandwidth Transimpedance Amplifiers

For relatively feeble fast electrons (via gamma-rays), the sensitive preamplification of the signals from a semiconductor detector is critical. The silicon detector can produce a signal as small as a few microamps, and to detect such small currents an extremely low-loss line coupled to a very sensitive transimpedance stage is needed. Hence, the implementation was chosen with the Sirenza SFT-0100 trans-impedance amplifier which outputs a differential signal that can be further fed into a differential input voltage amplifier.

Aside from the sensitivity mentioned above, the chip has a DCDR (Duty Cycle Distortion Ratio) output pad, which compensates for the duty cycle distortion that can

result from the high input current seen by the amplifier. As the induced current is increased due to increases in the bias voltage, an automatic circuit is invoked to prevent distorted operation in the transistors of the TIA. The threshold input current, where the circuit turns on, can be adjusted by employing a shunt resistor to the ground with a value between 1 kΩ and 50 kΩ at the DCDR pin. The lower the value, the higher the input current threshold at which the circuit turns on. In order to optimize for maximum sensitivity and dynamic range, the I_{mon} current can be monitored while tuning this shunt resistor. The optimum sensitivity and dynamic range can be obtained by tuning the resistor value until only a few microamps of I_{mon} current are observed when there is no source present. Hence the chip preparation must be done in a controlled environment where the background influence is minimized. We observed a quiescent current of 28 μA which we reduced to 1 μA by using a shunt resistor of 40 kΩ at the DCDR pad.

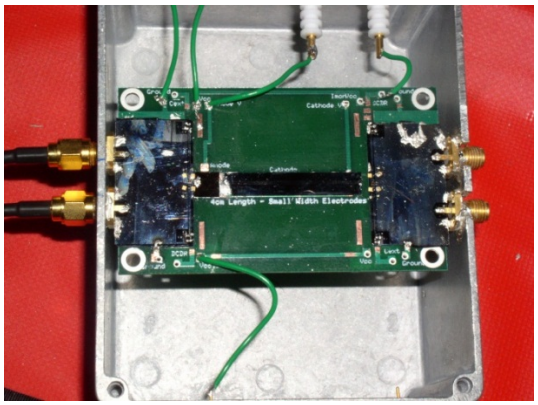


Fig. 5.31: The assembly of the meandered detector with TIAs on both ends of the electrode leading to SMA jacks.

Fig. 5.32: The assembly of the planar detector with the TIA.

A handful of detector fabrication and mounting runs were iterated in order to test the devices. More importantly, to understand the application and performance of the TIAs, in the process reducing sources of failure such as inadequate electrostatic discharge control and wire-bond cratering of the TIA package- 8 devices failed during these shakeout mounting runs. Note that the wire-bonding design is particularly critical because long bonds can droop and short-circuit, high currents can heat and cause failure of the bump or wedge bond, and wire-bonding to slippery platinum surfaces presents a challenge. In order to overcome these mounting challenges and to characterize the Sirenza TIAs in isolation of any partitioned-detector effects, a planar detector was used with a silicon-platinum Schottky junction, in which the platinum surface was directly wire-bonded with minimal bond length to the input pad of the TIA. The whole assembly is shown in Fig. 5.31 and again in Fig. 5.32. A custom-built printed circuit board (PCB) was used to give a base support and external biasing lines to the detector/TIA combination. Though the active radiation detector was directly placed on the PCB's biasing pad, the TIA was placed on a 550 μm silicon wafer which was patterned specifically to accommodate the TIA mounting requirements.

Fig. 5.31 shows the original assembly of the meandered detector and the TIA on both sides, outputting differential signals to the SMA pins on either side. By replacing the meander design with the planar detector, as shown in Fig. 5.32, the input conditions of the SFT-0100 were met, and the depletion region grew from the top surface, enabling probing with short-range alpha particles. The bias voltage to the detector was supplied to its bottom side through an SHV jack on to the PCB pad on which the detector was mounted. As the DC offset on the input pad of the TIA is around 1.45V, a more positive voltage of greater than 1.45V was applied to the bottom side (aluminum side) of the detector so that the reverse bias mode of the detector is enabled. The output lines that interfaced with the TIA output pads were designed as 50 Ω coplanar lines that mated with SMA tabbed jacks at the end. Single Layer Capacitors (SLCs) were used near the TIA for AC decoupling of the amplifier biasing voltage. The differential signals obtained from the TIA output pads are shown in Fig. 5.33. The signals were expectedly small (around 50 mV) and require a second stage of amplification, but the SNR is much better than that seen in the proof-of-concept devices.

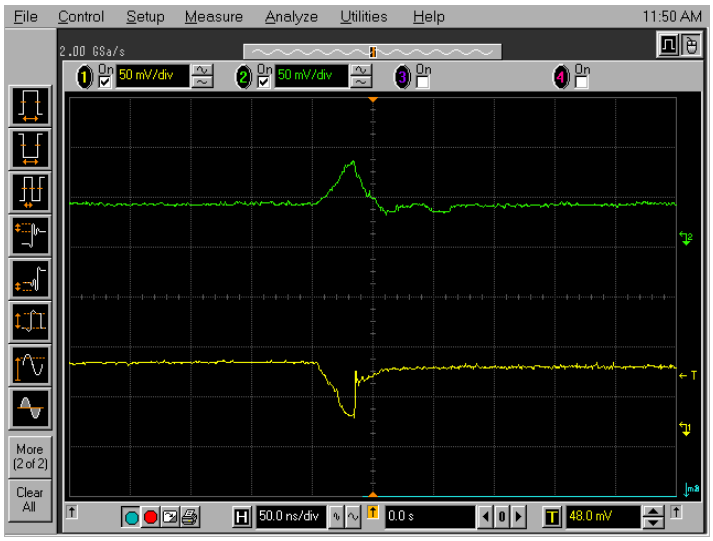


Fig. 5.33: Differential signal from planar detector/TIA assembly.

The somewhat small signal size is due to the limited transimpedance gain of the SFT-100 (approximately 2000 Ω), which when detecting a signal in the microampere range (say, 25 μA) outputs a voltage signal in the millivolt range. Fig. 5.33 waveforms also show a rise-time of around 25 ns, which is consistent with the charge drift through the detector, and which can be decreased with an increase in detector biasing voltage. The sharp fall time on the yellow curve is most likely due to intermittent mechanical contact at the SMA jack.

5.3 Position-sensing using a Current-Sensitive Transimpedance-based Preamplifier

The choice of preamplifier that follows a detector stage depends on the charge, current or voltage that we are interested in studying. The measured charge depends on the integration time of the capacitor and input resistance of the amplifier. Integration times larger than collection times of all charge carriers yield the full charge that was produced due to interaction. A smaller integration time yields the fraction of the charge that is momentarily integrated and under such a condition, there is instantaneous current in the circuit, which provides a better representation of the electrical carriers produced than that provided by the integrated charge-fraction since fully depleted silicon offers the minimum possible capacitance, and the only other way to decrease the integration time is to choose an amplifier with lower input resistance. A transimpedance amplifier has low input impedance and can convert the current signal to a proportional voltage signal. As

the instantaneous current on the detector electrodes rises sharply, one would prefer a proportional voltage signal, especially in timing measurements where our resolution in time delay depends on the fastest measurable component of the voltage pulse from the preamplifier. Because of the steep rise and fall times of a current pulse in a detector, the discretization of pulses in a high count environment becomes more feasible. This further reduces the additional step of dealing with piled up and masked pulses in the signal processing chain.

As explained in Chapter II, in which the reconstruction of low-charge-density electron tracks is discussed, the accuracy of the reconstruction is dependent on the sensitivity of the detector to radiation-induced drift electrons, which are in competition with the stochastic electrons flowing along the interfaces or thermally generated in the bulk. First and foremost, one desires that the current-sensing preamplifier be sensitive to a minimal number of electrons. Recall that an electron (or hole) drifted at saturation produces 29 pA of current, which can be sensed with various transimpedance amplifiers (TIAs). Amplifiers with fA sensitivity are available, but higher sensitivity usually comes at the cost of poorer bandwidth. The highest bandwidth TIA's can have sensitivities in the microamp range. One would like to use the TIA to reach or exceed the performance of the best (standard) charge-sensitive preamplifiers, in which ~100 electrons can be sensed. Once that amplifier sensitivity is reached, one must also quench the competition from the leakage current, which has made its quenching our continual focus until the charge sensitivity is limited by the amplification component. This quenching process, up until the charge sensitivity and the limitation by the amplification component, has been our focus.

5.3.1 Design Option

Transimpedance amplifiers have lower input impedances than charge-sensitive and voltage-sensitive devices, permitting their operation with higher capacitive sources. If one uses an operational amplifier for constructing the TIA, the input impedance at the inverting node is also low as it is at virtual ground. The design of such a TIA involves balancing the effect of detector capacitance and feedback resistance with a parallel feedback capacitance. On the other hand, microwave TIAs, which also have low input impedance, can be obtained in die form. The die-form amplifiers produce a faster

response by utilizing heterojunction bipolar transistor (HBT) monolithic configurations, which have a flat gain over their bandwidth. For the design discussion below, the two categories of TIAs examined were:

1) Die form TIAs – SFT 0100, Maxim 3724.

2) TIA designed using FET-input Operational Amplifier – OPA 657(FET based), OPA847

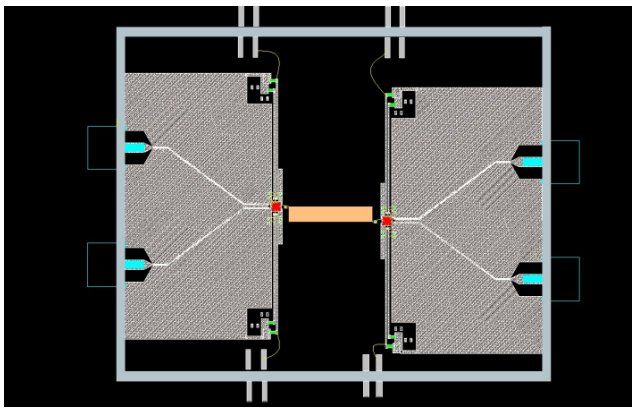


Fig. 5.34: The geometric layout of the detector (in orange), the readout amplifiers (in red), and the subsequent signal extraction (in blue).

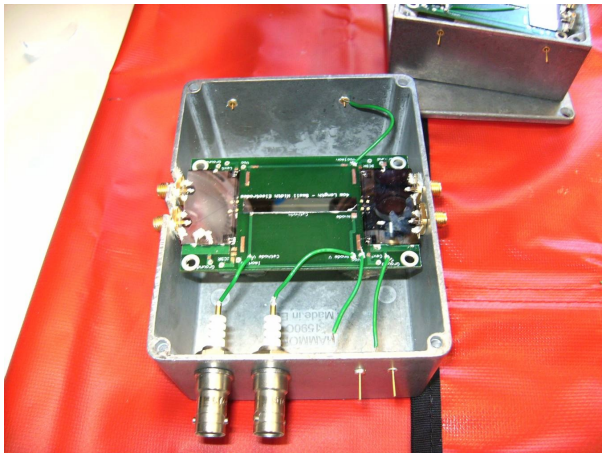


Fig. 5.35: The mechanical realization using a silicon surface barrier configuration.

As shown in Figs. 5.34 and 5.35, the design extends to accommodate a TIA with differential output, which aids in noise reduction during the second stage amplification when the voltage amplifier takes in the differential signals at its input. The interconnects

between the stages have been designed as 50 Ω coplanar waveguide (CPW) structures, in which the signal line is guided by ground planes on either side. This is primarily done to reduce cross talk- the ground surrounding the signal nulls any external noise- and also to enable mating of these lines to an external SMA pin, since the characteristic impedance of a CPW structure depends on the width of the line and the signal-ground spacing and it is thus easier to maintain the 50 Ω while flaring.

With an aim to accurately extract the current signal the Sirenza's SFT-0100, which is an InGaP based HBT transimpedance amplifier with a generous bandwidth of 10 GHz, was investigated. The advantages of using a MMIC microwave transimpedance amplifier is the uncompromised bandwidth, its ability to amplify the signals right out of the detector (being in a die form) and hence an increased sensitivity to the weak signals out of the detector. Though the gain of these TIAs is much less than an op-amp based TIA, the signals are much faster and subsequent stages of amplification can be included to provide the desired amplitude. The amplifiers provide a means by which the speed of the desired transitions can be judged, and inform the design for subsequent generations.

5.3.2 Design Considerations and Analysis

The main design criterion for a transimpedance amplifier is its current sensitivity. Input-referred RMS noise current determines the receiver sensitivity. A high sensitivity device can output error signals if the input noise exceeds the input referred noise of the amplifier, and a less sensitive amplifier can miss not only noise components but also some of the smaller signals. Working with high sensitivity TIAs, the external noise sources is struggled to limit, such as electromagnetic noise pickup, and thus used a bare-die TIA placed in close proximity to the detector. The input referred noise of the Maxim TIA is 0.325 μA at 2.1 GHz with 0.6 pF input capacitance. Since the diode detector we use has a minimum capacitance of 20 pF for a 1 cm^2 area, the input referred noise also increases to a few μA. The Sirenza SFT-0100, on the other hand, has spectral current noise density of 10pA/√Hz and a responsivity of 0.9 A/W, from which one can calculate the input referred noise to be 0.1 μA at 7 GHz, with a 50 Ω impedance as seen by the input of the amplifier, as outlined in Eq. 5.1.

$$\text{Noise Equivalent Power} = \frac{P_{signal}}{S/N} \tag{5.1}$$

$$= \text{Noise Current / Current Responsivity}$$

$P_{signal} = I_s^2 R$ where I_s is the signal current across the resistance of the detector and R is around 500 Ohms. Using this we can calculate the input referred noise (= I_s) by setting S/N = 1.

The dynamic range is the ratio of smallest detectable signal (which is limited by the noise floor at the amplifier input) and the largest signal (that would not cause any distorted operation of the transistors of the TIA). The Sirenza SFT-0100 offers a feature to compensate for the duty cycle distortion which is a result of high input current as seen by the amplifier. The threshold input current at which the protection circuit turns on can be adjusted using an external resistance, and thus gives better control when optimizing for maximum sensitivity and dynamic range.

5.3.3 Noise Considerations

Electronic noise determines the lowest detectable signal level and must therefore be minimized. In order to minimize the interference noise, a bare die-form TIA was placed close to the detector to extract the relatively low noise signal directly from the detector. The integrated detector and TIA combination was placed in a grounded die cast aluminum box. In another experimental setup (cf. Fig. 5.35), the detector was wire bonded to coplanar output lines (to minimize noise pickup) on a PCB board that interconnected to an opamp based transimpedance amplifier. In both the cases, we could achieve a better SNR than in the efforts based on simple op-amp designs; however, the noise inherent to the charge transport, such as fluctuations in carrier velocity (from thermal noise) or carrier number (from thermionic emissions) is still present. Since such noises are "white", their frequency spectrum is flat over the bandwidth and cannot be easily filtered out. For example, although the DC component of leakage current can easily be filtered before reaching the preamplifier, some fraction of its AC fluctuations occupy the signal's pass-band. Thus, from a detector-design standpoint, one desires to minimize the DC magnitude of the leakage current. As the signals are measured, the

degree to which one can separate the drift carrier signal from the thermal carriers is examined.

The input noise voltage to an amplifier does not depend on the detector capacitance whereas the signal voltage decreases with increasing capacitance, thus reducing the SNR. That can be avoided completely by converting a capacitive source to a resistive one. This can be done by treating the detector (meandered lines) more as a transmission line rather than a chunk of capacitance integrating the charges on its surface. Another advantage of making them resistive is the minimization of input-related noise. The input-related noise occurs at the input terminals and has its own source resistance provided by the amplifier's internals. This particular noise source resistance can be calculated as the ratio of input voltage noise divided by the input current noise. With regards to this noise component, the role of source and load are reversed; that is, the resistance of the attached signal source is the load resistance to the amplifier's input-related noise. Therefore this noise component is at minimum when the amplifier's input noise resistance is matched to its load, which is the resistance of the attached signal source. Hence, if one has detector which is matched to a 50 Ω input transimpedance amplifier, one not only decreases the input-related noise but also minimizes the loss of signal current and therefore achieves a better SNR.

The detector bias voltage can also be another source of noise into the system. A small voltage fluctuation on the bias supply of the silicon diode detector produces a current $I = C_D \Delta V/\Delta t$ which can reduce the sensitivity, where C_D is the detector capacitance. RC filtering is employed to remove the bias supply noise.

5.3.4 Board Design Considerations

For high frequency design, the choice of PCB board components, the interconnect dimensions, the mounting techniques, the conductive epoxy's resistivity, etc. become critical for low-loss signal transmission. To get the high performance from the optimized discrete components, passives must be scaled in size to fit the application. Smaller components have lower parasitic inductance and capacitance, a requirement for higher frequency operation; thus, surface mount devices were used all along the signal propagation path when working with packaged op-amp based TIAs and MIM single layer capacitors and resistors while mounting die-form high frequency TIAs. The

interconnections were all designed as 50 Ohm coplanar lines to minimize electromagnetic interference on the signal line. The vias on ground planes were widely used to avoid potential differences on the ground planes, and the SMAs and PCB mountable adapters were used for higher frequency handling.

5.3.5 Measurements with the Sirenza SFT-0100 TIAs

In the experiment, a tail pulse generator was first used with a series resistor to produce a current signal as an input in order to determine the minimally detected current from the detector. Second, an alpha source was used to create charge pulses in a planar detector, which was used as an input to the TIA. Comparing the signals from the two sources, the minimum detectable signal- which includes the limitations of measurement with the scope- was noted. Additionally, the fastest signal that can be seen without being limited by the preamplifier at various bias voltages was studied. This would determine the actual required bandwidth of the preamplifier at a particular value of bias voltage, which can be used in the reduction of high frequency noise by constraining the pass-band.

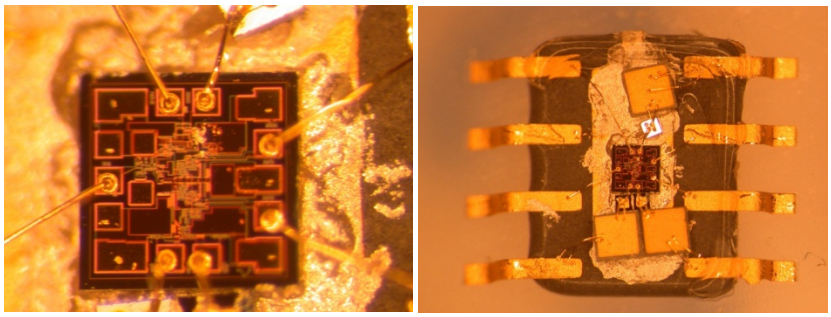
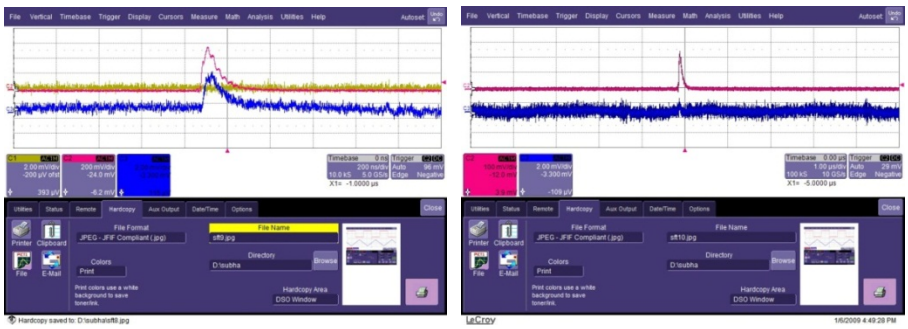


Fig. 5.36: Wire bonding the Sirenza SFT-0100 Transimpedance amplifier.



(a) (b)

Fig. 5.37: SFT-0100 TIA response to test input from tail pulse generator. (a) At 500 mV (corresponding current input of 125 μA) (b) Sensitivity degrades with input amplitude decreasing to 200 mV (corresponding current input of 50 μA).

When using a bare-die chip one must consider many factors with regard to its mounting and to ensure its safety from static electricity. As evidenced through the progressive knowledge obtained through each build of the device, a wire-bond encapsulant works well in protecting the chip not only from electrostatic damage but also from exposure to the incident alpha radiation. Fig. 5.36 shows the bonded SFT-0100, in which single layer capacitors were used for AC bypassing at the Vcc and also at the DCDR pads. Gold ball-bond was used on the die pads whereas aluminum wedge bonding was used to bond the platinum surface of the silicon detector.

The bare die was mounted to a PC board and tested. Fig. 5.37 shows the initial sensitivity of the device. With a 125 μA input, the TIA could reproduce an equivalent, though small voltage signal of 5 mV height, as shown in Fig. 5.37a. As the input current was decreased to 50 μA, the measurable signal also reduced as expected and is shown in Fig. 5.37b, in which the sensitivity of the oscilloscope is reached.

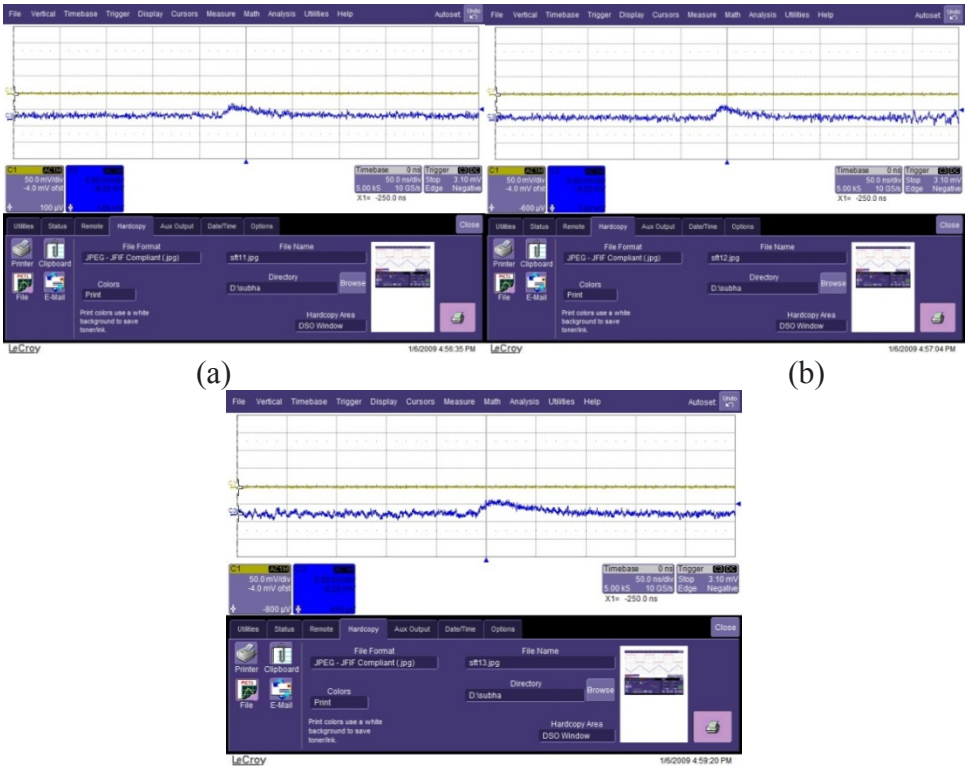


Fig. 5.38: SFT-0100 TIA response to alpha sourced planar detector biased at: (a) 25 V, (b) 33 V, and (c) 40 V.

The SFT-0100 has a photodiode monitoring pad, which not only reflects the average diode current, but it can also provide the means to manually adjust the DCDR threshold. With no input photocurrent applied to the TIA, the current monitor will give a finite quiescent current, measured in microamperes, which indicates that the automatic circuit is active. By employing a shunt resistor (1k-50k) on the CDCDR port to ground, this threshold can be reduced, setting the automatic circuit for maximum sensitivity and dynamic range. In order to read below a 50 μA input current, the sensitivity adjustment needs to be done at the DCDR port, but due to the instability associated with that measurement, the sensitivity could not be reduced until further design improvements are employed.

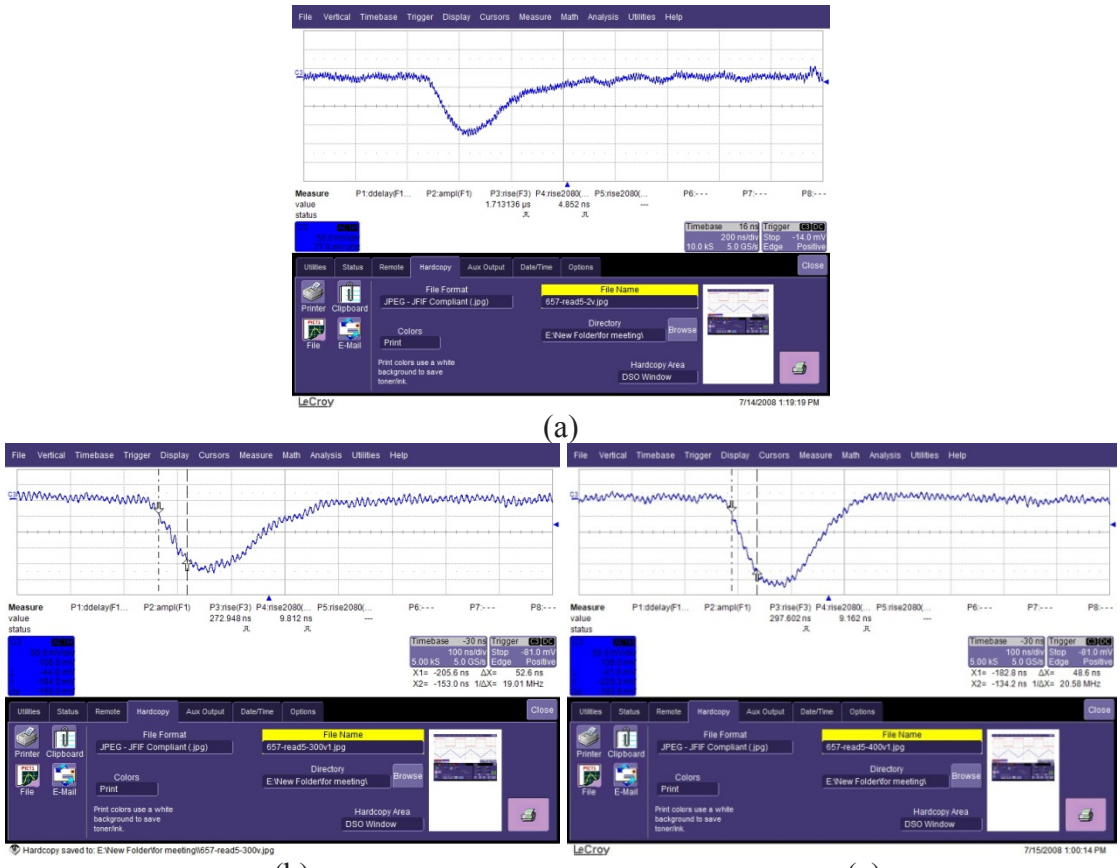
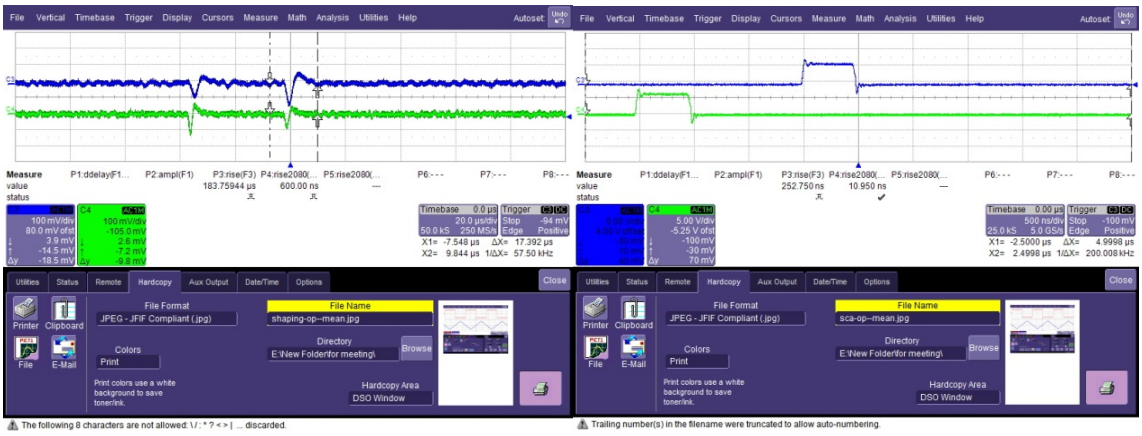


Fig. 5.39: OPA 657 based TIA output. Detector Biased at: (a) 2 V (rise time 90 ns, Amplitude of 100 mV), (b) 300 V (rise time 52 ns, Amplitude of 200 mV), and (c) 400 V (rise time 48 ns, Amplitude of 250 mV).

Fig. 5.38 shows the results obtained using the SFT-0100 when coupled to a planar detector impinged by alpha particles. At very low bias, the carriers move well below saturation and since the current amplitude is proportional to the carrier velocity, the SFT-0100 does not produce a large enough signal to see with a standard oscilloscope (or without additional amplification); however, at moderate biases, but about 2 orders of magnitude below the fields required for hole saturation, the signals are detectable, as shown in Fig. 5.39, in which the signals exhibit a rise time of approximately 8 ns. Recall that this rise time corresponds to the time to sweep the carriers out of the device and is consistent with our modeling, but further validation tests are needed to gauge whether there are some circuit limitations to this measurement. Further stages of amplification will allow us to monitor smaller signals, and Sirenza's broadband cascadable amplifier, the NBB-400 with 20dB broadband gain, were tested and 100 times improvement is expected in the amplitude with two stages of amplification.

With the 8 ns rise time from the Fig. 5.39 signals, we expect a marked improvement in the resolution with which we measure the time difference between the two signals from either ends of the meandered line, recalling that we achieved 160 ps timing resolution with OPA 657 TIA s which could output signals with 90 ns rise times.



(a) (b)

Fig. 5.40: Trace from the initial setup of 45 turn Meander through (a) the 657 TIA and (b) the 551 (SCA).

The initial measurements were made using a square centimeter planar detector and subsequent tests were performed using: (a) a 45 line meander detector (100 μm strip and gap), and (b) a 50 line meander, with 10 μm strip and gap.

As shown in the Fig. 5.40, the pulses become faster and their amplitude increases with increased detector bias voltage. As suggested in Fig. 40, the timing resolution gets lost when faster signals are passed through standard shaping amplifiers; for instance, with the Ortec 572 set with minimum shaping times, the time delay measurements on the TAC spectrum exhibited a low resolution; thus, faster pulse processing electronics are required until the timed-transitions in the pulses are extracted.

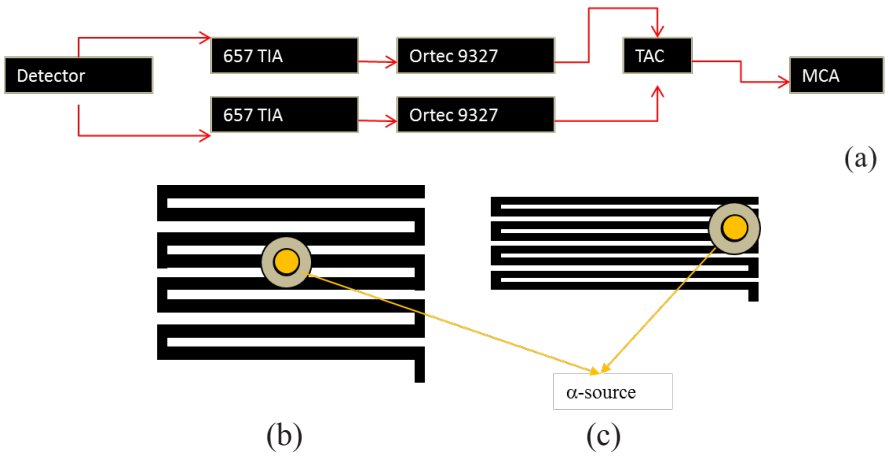


Fig. 5.41: Test setup for OPA 657 based TIA (a). Positioning the α-source at the top-center 100 um wide detector (Not to scale, in particular, the button alpha source extends across all of the lines in the vertical direction) (b). Positioning the α-source at the top-corner of 10 um wide detector (again not to scale) (c).

Fig. 5.41 shows the test setup for the OPA 657 based TIA. An Ortec 9327 amplifier and discriminator (1 GHz BW) were used for measuring the time delay via the TAC spectrum. The 5.5 MeV alpha source, collimated horizontally along the lines, as shown in Fig. 5.41, was placed using supports close to the detector. That is, the incident alphas could impact the detector through any of the electrode strips, but at only a given position along the length of the strips. For one measurement using the 100 μm strips, the source was placed at the midpoint of the strip-length of 1.5". For the 10 μm strip detector, the alpha source was placed closer to one end of the device (comprised of 1.5" long strips).

The resulting time spectra showed significant improvement from our previous results, yielding a time resolution of 160 to 170 ps, for either strip geometry, as shown in the next section. Note that this is achieved with a readout chip that is far from our best TIAs, the on-chip die-form Sirenza.

Thus, if needed, then a couple more magnitudes of improvement can be achieved in the timing resolution. A faster timing is always desired, but it may not significantly extend the capability of the detector as an imaging device of higher energy radiations. Nevertheless, we expect to implement the higher bandwidth, lower noise TIAs so that the position and charge sensing are limited only by physics-limitations, such as diffusion.

5.4 Lateral Position-Sensing based on Timing Variations

5.4.1 Single line charge positioning based on op-amp based TIA

As the instantaneous current on the detector electrodes rises sharply (cf. Fig. 5.39), one would prefer a proportional voltage signal, especially in timing measurements where the resolution in time-delay depends on the fastest measurable component of the voltage pulse from the preamplifier. A transimpedance amplifier (TIA) has low input impedance and can convert the current signal to a proportional voltage signal. For straightforward breadboard prototyping, we utilized the OPA657 operational amplifier, which employs a JFET input with a negligible current noise. The OPA657 is a high gain voltage–feedback amplifier with extremely high input impedance; hence, the input bias current is of the order of 2 pA, and most of the signal current takes the feedback path.

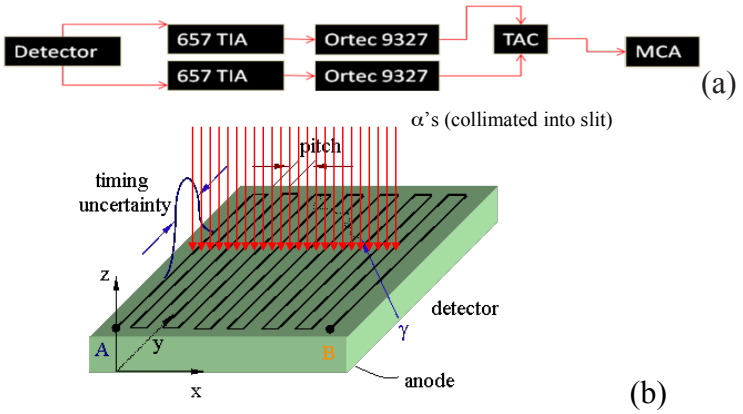


Fig. 5.42: Test setup for OPA 657 based TIA (a). Positioning the button α-source, such that the impingement extends across the x-axis of detector face, but along approximately only one y-coordinate (Not to scale) (b).

Fig. 5.42a shows the test setup for the OPA 657-based TIA. We used an Ortec 9327 amplifier and discriminator (1 GHz BW) for measuring the time delay between the pulses emerging from the two ends of the meander-line via the time-to-amplitude-converter (TAC) spectrum. The ^{241}Am alpha source, collimated horizontally along the lines, as shown in Fig. 5.42b, was placed using supports close to the detector. That is, the incident alphas could impact the detector through any of the electrode strips, but at only a given position along the length of the strips. Thus, we expect that a uniformly spaced comb-like TAC spectrum will result since the time-difference for interactions between neighboring strips should be temporally equidistant, an expectation realized in Figs. 5.43 and 5.44. For one measurement using the 100 μm strips, the source was placed at the midpoint of the strip-length of 1.5 inches. For the 10 μm strip detector, which was also comprised of 1.5 inch long strips, the alpha source was placed closer to one end of the device. The resulting time spectra exhibited a time resolution of approximately 170 ps for both 10 and 100 micrometer strip (and gap) geometries, which is sufficient for distinguishing the position of interaction along the transmission line. Specifically, the individual peaks in Figs. 5.43, 5.44, and 5.45 correspond to the charge collection on an individual strip, and one can therefore begin to process the pulses with downstream electronics so that electron-hole track reconstruction can commence, which is the subject of Chapter IV. Note that the equivalency of the peak separations in Figs. 5.43 and 5.44 reflects the fact that the strip lengths are equal (1.5") in the two cases.

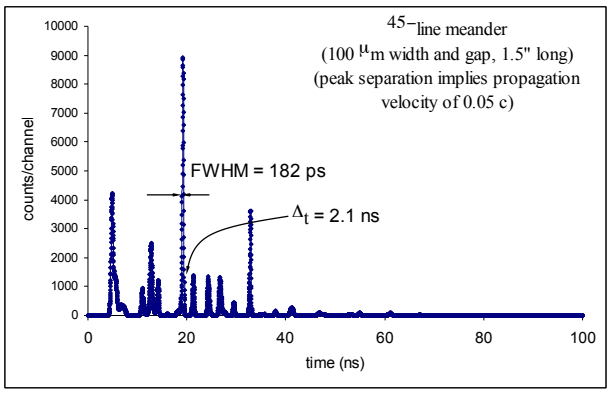


Fig. 5.43: TAC spectrum of 45-line device, with 100 μm strip width (and gap) and 1.5" strip length, exhibiting a FWHM timing uncertainty of 182 ps.

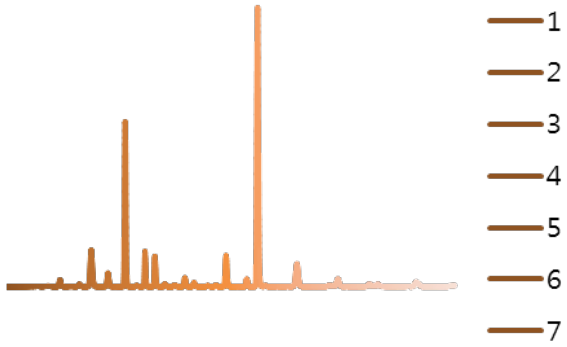


Fig. 5.44: TAC spectrum of 50-line device, with 10 μm strip width (and gap) and 1.5 inch strip length, exhibiting a FWHM timing uncertainty of 160 ps (the abscissa is the channel number and the ordinate is the counts/channel).

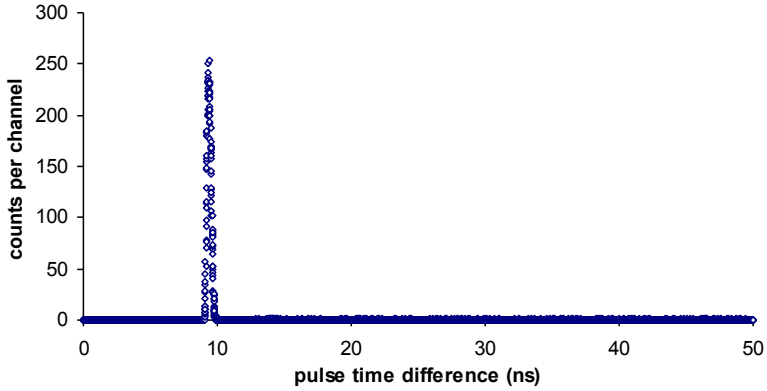


Fig. 5.45: For an alpha source collimated to a single 100 micrometer strip, the TAC spectrum derived from the 45-line, 100 micrometer-gap device.

If the alpha source is collimated to a single strip (using a perforated paper collimator), then the TAC spectrum simplifies to a single peak, like that shown in Fig. 5.45, for the 100 μm (strip and gap) design. The op-amp based TIA designs can thus deliver acceptable performance, but to achieve better results, we also investigated the use of bare die-form TIAs. The die-form amplifiers produce a faster response by utilizing hetero-junction bipolar transistor (HBT) monolithic configurations, which have a flat gain over their bandwidth.

5.5 Current Sensing as Means of Mitigating the Effects of Poor Charge

5.5.1 A Readout Solution: Initial Peak Current-Sensing

Recall that we don't simply collect the radiation residual products after creation and from their enumeration, derive the energy signal; instead, we measure the charge induced on the electrode by those in the bulk. Thus, just after charge-hole creation in the semiconductor bulk, such as that shown in Fig. 5.46a, the charges feel the effect of the local electric field and begin to drift. Immediately, a current, i, is induced on the bounding electrode in response to the changing electric field condition, that is equal to the charge q (which is proportional to energy), the drift velocity v, and the weighting field E_0, as shown in Eq. 5.2.

$$i = q\vec{v}\vec{E}_0 \tag{5.2}$$

For planar detectors, or for interactions far from some surface partitioning, the drift velocity is constant as is the weighting field, and the current magnitude is therefore proportional to the initial charge number. After the initial step in the current pulse, the subsequent charge shape reflects the charge transport. For instance, in Fig. 5.46b, the electric field increases throughout a single strip detector. Of greater concern to us is the case in which charges are lost, or delayed, during their motion through the bulk semiconductor. For instance, in Fig. 5.46c, we see a measured current signal derived from an alpha particle hitting the cathode (hole-collecting) side of a silicon detector. Since both the holes and electrons drift well in silicon, both contribute a substantial step in the current pulse. Regardless of the mobility details, the initial current peak reflects the initial charge number, which is subsequently diminished as: (a) the holes are first collected upon the top electrode, and (b) the electrons are collected on the anode after drifting through the entire 550 μm detector. If the electrons were trapped during transport, then the step following hole collection would exhibit a diminishing magnitude, but as long as we extract the initial current peak, our energy measurement is not affected. The method is therefore insensitive, in principle, to loss during transport, and it therefore opens the door to exotic semiconductors with charge trapping as well as poorly conducting heavily trapping materials (although polarization can impact the field formation). The current peak measurement method therefore works in principle.

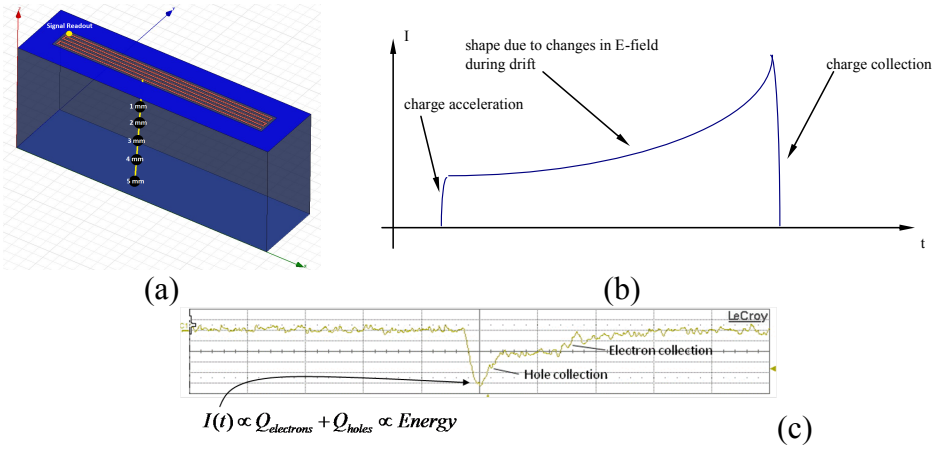


Fig. 5.46: Illustration of charge-track creation in the bulk (a). Schematic of the form of the induced current signal, in which the charges accelerate (in fractions of a ps) and subsequently drift through the field, and finally are collected at the electrode (in which the electric field is zero) (b). Measured signal derived from an alpha particle hit the cathode side of silicon (c).

5.5.2 Energy Spectra derived from Current Pulse

For either charge-sensitive or current-sensitive measurements, the fluctuations in the baseline can limit the signal-to-noise ratio (SNR), rather than the desired case in which the statistical counting associated with the charge-carrier creation limits the energy resolution. One of the advantages of integrating the charge in the detector is that the high-frequency noise is filtered as a natural part of the signal formation. In either case however, the detector leakage current, or the readout circuitry noise can limit the SNR and the resolution. The key to effectively utilizing current sensing is to therefore quench both the leakage current and the electronic noise in the circuit, so that the current peak magnitude is effectively extracted.

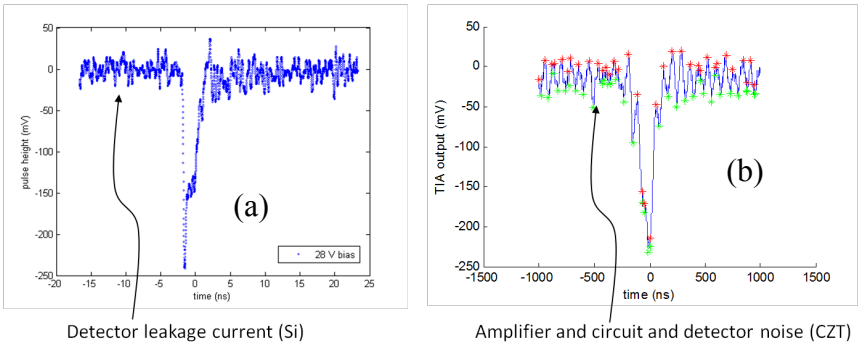


Fig. 5.47: Current pulse measured from Si, in which the high-frequency leakage current of the detector can modulate the current peak amplitude (a). Current pulse from CZT, in which the circuit lines are acting like an antenna for the current pulses traversing the circuit, resulting in a resonant noise (b).

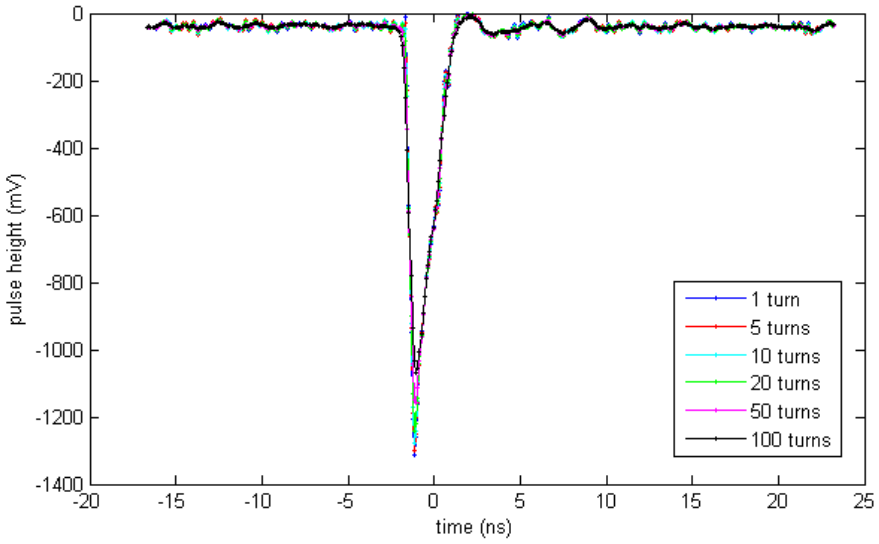


Fig. 5.48: Current pulse measured from Si, when fully biased to carrier saturated velocities, on a detector with low leakage current.

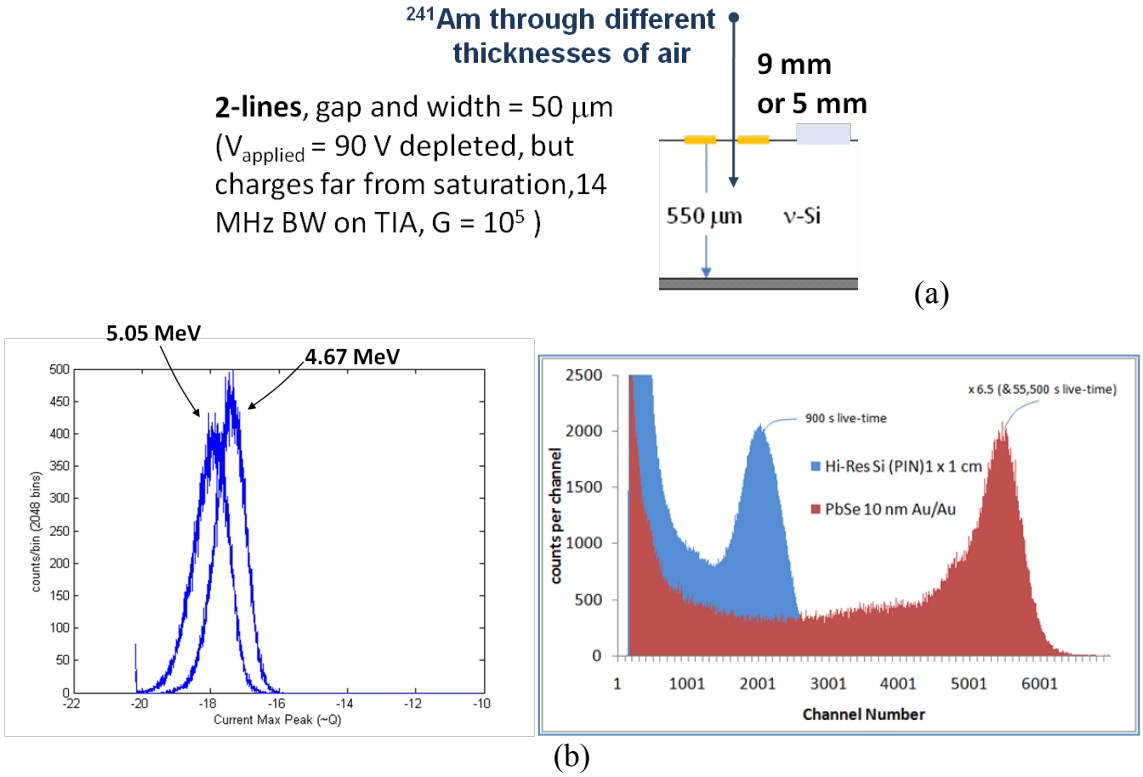


Fig. 5.49: Schematic of experimental setup (a). Pulse height distributions derived from the current peak, in which more negative currents correspond to higher energies (since the pulses have negative polarity) (b). The inset shows alpha particle spectra derived from a standard charge-sensitive setup (Ortec 142a preamp, 572 shaping amplifier), from either silicon (blue) or PbSe NC (maroon), exhibiting similar shape to that derived from the current pulse.

Fig. 5.47 shows two examples of current pulses, derived from Si (Fig. 5.47a) and CZT (Fig. 5.60b), in which the baseline fluctuations are substantial compared to the signal, due either to high leakage current- in the silicon case- or high frequency noise due to the circuit- in the CZT case. The elimination of these two sources of noise have been much of the practical focus of the work, in which the silicon noise was quenched, and lower-noise TIAs were implemented, as discussed in later sections. The goal of the work is to produce a pulse such as that shown in Fig. 5.48, which is measured from a silicon detector biased such that the carrier velocities are saturated (hence, the signal is maximized), and one for which the noise is quenched.

If one does utilize a detector with a stable baseline, then we showed that fine energy discrimination is possible. For instance, in Fig. 5.49, alpha particle spectra were derived from a 2-line (i.e. two legs of a meander-line pattern) silicon detector, in which the line width (and gap) was 50 μm. Since the range of the 5.5 MeV (maximum) alpha particle is only ~28 μm, the charge is generally collected on only one leg of the pattern. Fig. 5.49b shows that the spectra have the expected shape- as produced by the energy distribution from the thick source, which can be discerned by comparison with high resolution measurements made with Si and PbSe NC detectors using a standard charge-sensitive setup (see spectra in inset). The air gap is changed between the source and the detector (from 9 mm to 5 mm) in order to confirm that the spectra shifted in response to the differing alpha particle energies impacting the active region, and therefore found that the detector-readout system (FEMTO TIA with 14 MHz bandwidth and a gain of 105 V/A) can yield fine energy measurements. Note that higher energies correspond to leftward shifts in current since the signal polarity was negative (as shown in the pulses in Fig. 5.47 and 5.48). The fine energy resolution was further confirmed with gamma-ray measurements, in which a ^{60}Co source was used to impinge 1.173 and 1.332 MeV gamma upon the detector. Note that the silicon detector has a very low probability of photoelectric absorption at these energies (0.0044 %), and one therefore expects that the distribution will be dominated by the Compton scatter distribution (Compton scatter occurring 99.7 % of the time). This expectation was met, as shown in the distributions shown in Fig. 5.50. First, in Fig. 5.50 that for gamma-rays interacting upon the 50 μm

two-line, the range of the recoil electrons is large enough that the charge tracks will extend beyond the detector boundary for most of the interactions.

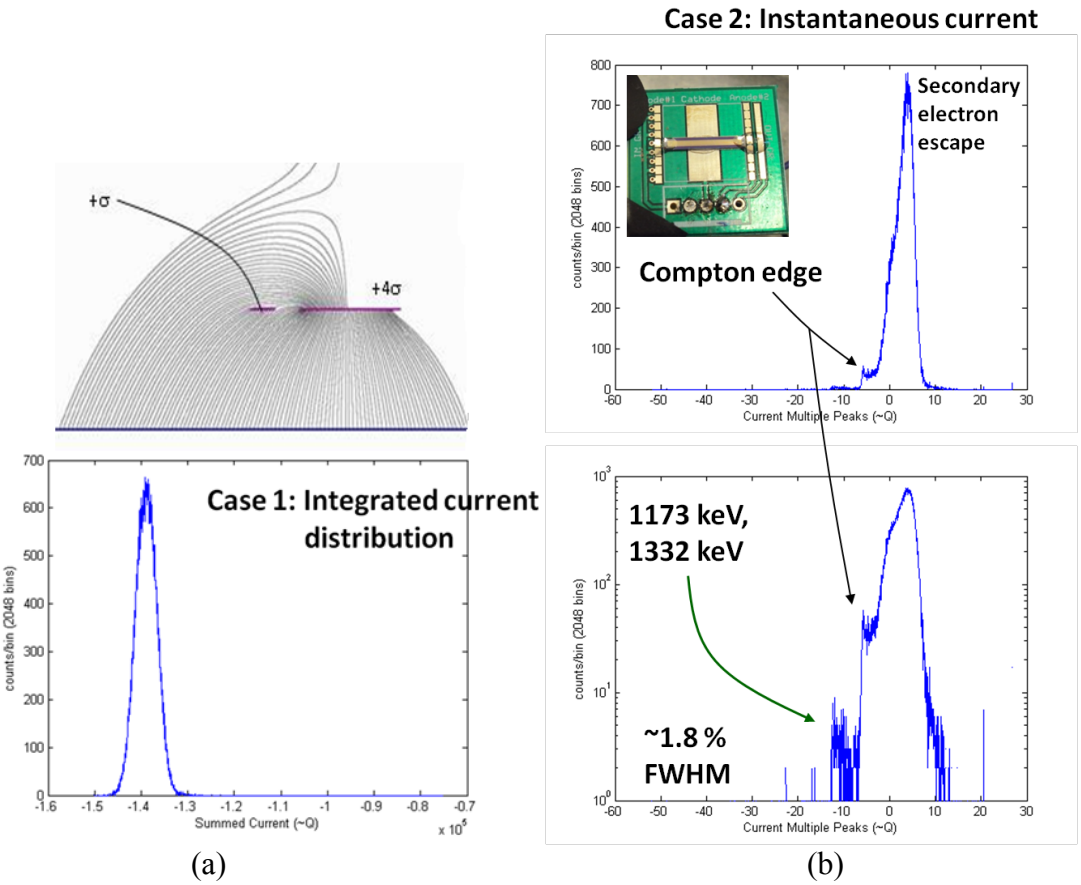


Fig. 5.50: Electric field lines for an example detector with two parallel strip-lines (top). The pulse height distribution derived from the integrated current signal (comparable to charge sensitive methods) (bottom) (a). In linear (top) and logarithmic (bottom) form, the pulse height distribution using current peaks (b).

The electric fields shown in Fig. 5.50a show that those charges that aren't collected on the electrode will generally be guided toward the silicon surface, where they will be trapped. Thus, although the detector uses silicon- for which trapping is negligible- the experiment simulates materials with substantial charge trapping. As shown in Fig. 5.50b, for which the charge-signal is used, as derived from the integrated current signal, the pulse height distribution does not reveal any energy information because of the trapping and the integration of the noise in the pulse. If we instead focus on the current peaks, as shown in Fig. 5.50b, then the expected energy distribution is produced, in which small full-energy peaks are accompanied by a large Compton continuum. There is

a substantial component at low energies as well due to the escape of the recoil electrons through the 550 μm silicon wafer. The work thus confirmed that if a low-noise detector was used with an optimized circuit, then extracting the energy from the current peak was feasible, even in the presence of charge trapping. The main limitation to the technique is thus not in the charge trapping, but rather in the leakage current noise which competes with the radiation-induced current pulses. However, the main focus is now on higher-Z wider band-gap materials, which are more naturally suited to gamma-ray detection.

5.6 Energy Spectrum Measurement with Low Leakage Current Si Detectors

5.6.1. Energy spectrum measurements from planar type detector

In order to measure the low energy gamma spectrum with a fabricated Si PIN-type planar radiation detector, three types of preamplifiers were used: 1) the Ortec(142A), 2) the AMTEK(A250NF) and 3) from Endicott Interconnect Technologies, Inc., the eV-5903. Fig. 5.51 shows the PCB layout for the eV-5093 preamplifier (left top), its picture (right top) and acquired signal (pink) and shaped signal (blue). Fig 5.52 also shows the PCB layout for the 2 of A250NF preamplifiers (left top), its picture (right top) and acquired signal (pink) and shaped signal (blue). Fig. 5.53 shows the PCB layout for the customized TIA (transimpedance amplifier, left), and its picture (right). Fig. 5.54 shows the block diagram of the detection system set-up with different kinds of preamplifiers (Case A: 142A, Ortec, B: A250NF, AMPTEK, C: eV-5093, EITI).

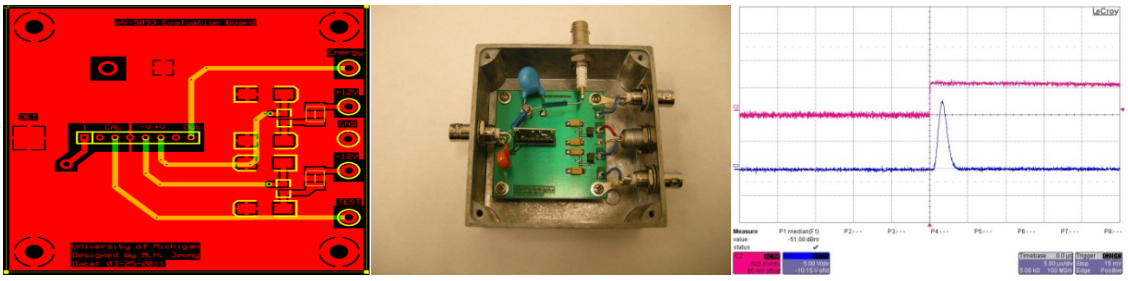


Fig. 5.51: The PCB layout for the eV-5093 preamplifier (left), its picture (middle) and acquired signal (right, pink) and shaped signal (right, blue).

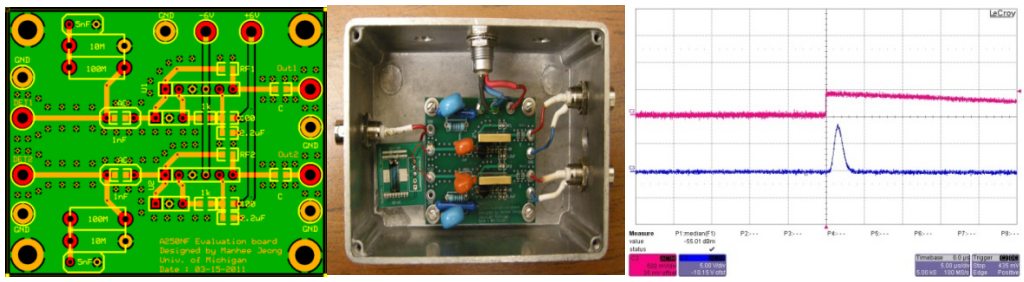


Fig. 5.52: The PCB layout for the 2 of A250NF preamplifiers (left), its picture (middle) and acquired signal (right, pink) and shaped signal (right, blue).

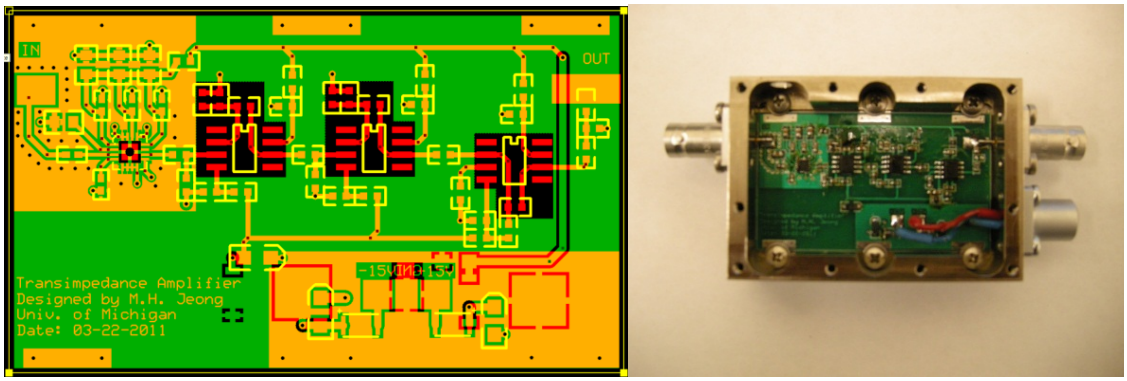


Fig. 5.53: The PCB layout for the customized TIA (transimpedance amplifier, left), and its picture (right).

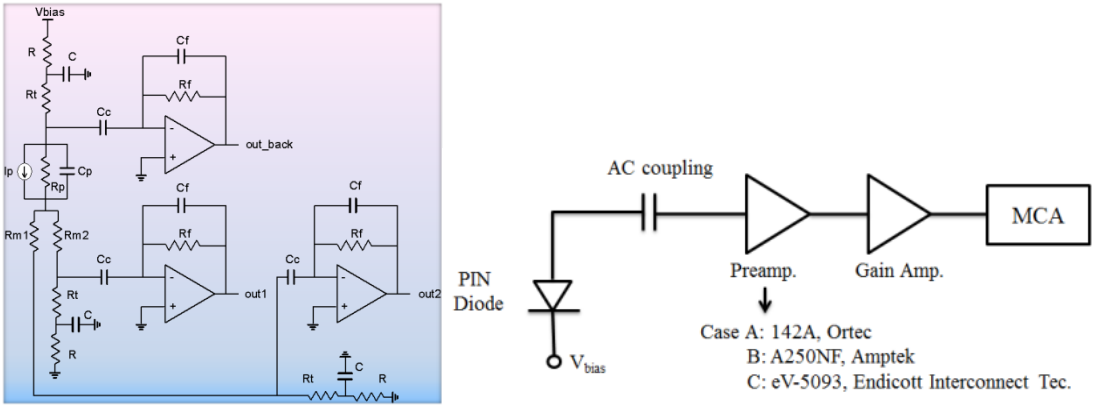


Fig. 5.54: Block diagram of the detection system set-up with different kinds of preamplifiers.

We used a passivation layer with SiO_2 to protect the surface of the detector from dust and humidity to reduce the leakage current and to maintain its electrical properties. However, its electrical characteristics were changed after a few weeks. Thus, we altered the passivation layer from SiO_2 to a SiO_2 and Si_3N_4 combination. With this combination,

we can maintain the device performance without compromising its electrical properties. With fabricated planar type detectors and a general configuration of the readout system, we measured alpha spectra (Am-241, 5.293MeV), as shown in Fig. 5.55.

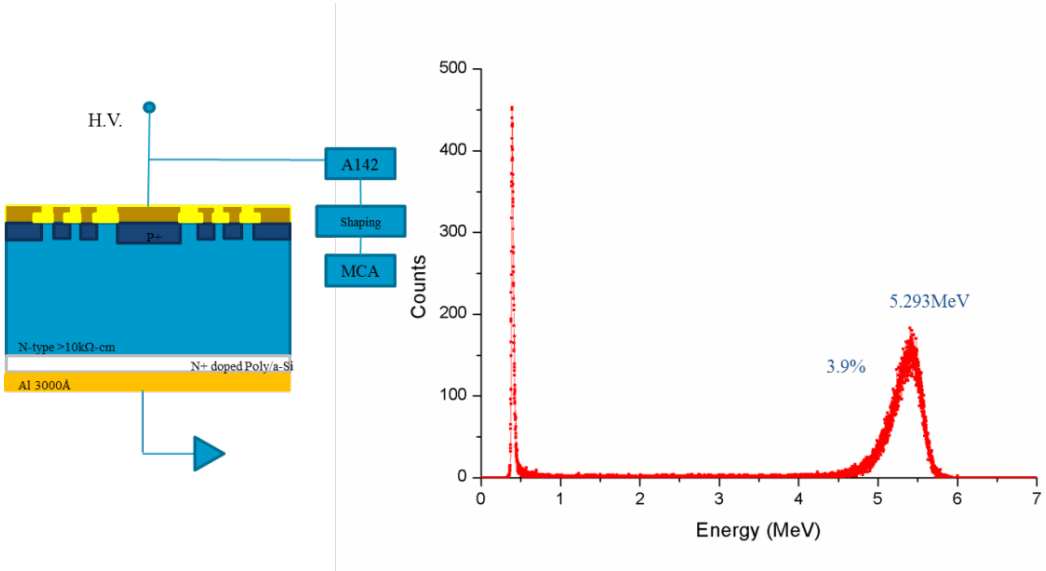


Fig. 5.55: Am-241 alpha spectrum and testing configuration.

With the Am-241 alpha source, we measured spectra as the distance was varied between the source and detector from 7 mm to 1 mm. To compare the mean energy, we ran an equivalent simulation using by SRIM, the results from which are shown in Fig. 5.56. The main peak shifts to the left side with increasing the gap between the source and detector as shown as Fig. 5.57.

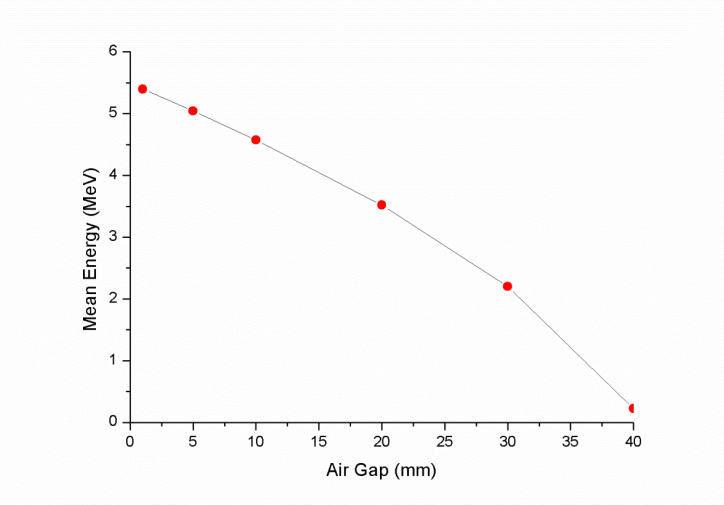


Fig. 5.56: Mean energy attenuation with varying the gap between the source and detector.

In addition, there were other peaks below the 100 keV, due to photon impingement. To verify the source of these peaks, we put a piece of paper between the source and detector to block the alpha particles along, the result from which is shown in Fig. 5.58. The gamma peak of 59.9 keV resulted, confirming that the feature corresponds to the gamma ray peak from Am-241. Furthermore, we increased the bias voltage from 5V to 100V and acquired alpha spectra, as shown in Fig. 5.59. In order to produce full depletion for the 550 μm thick Si detector, a bias of over 70V is required. As shown at Fig. 5.59, the main peak is saturated at the maximum energy of the alpha source. Below the 70V bias voltage, the main peak is shifted to lower amplitudes due to: principally, the increasing detector capacitance that accompanies partial depletion, and secondarily, larger dead layers, as the junction reaches less into the p+ top layer at lower biases.

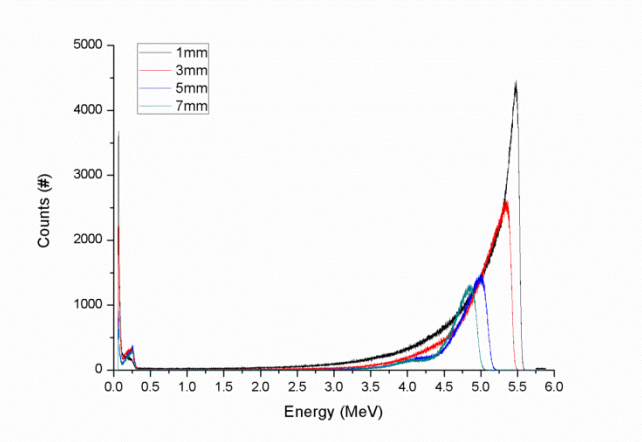


Fig. 5.57: Alpha spectrum with varying the gap between the source and detector.

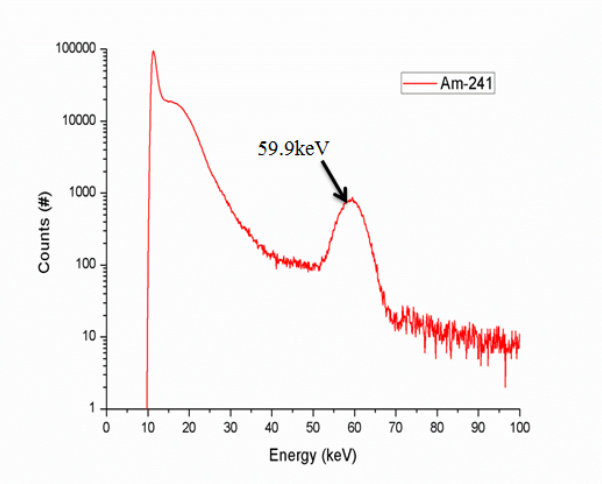


Fig. 5.58: Am-241 gamma spectrum.

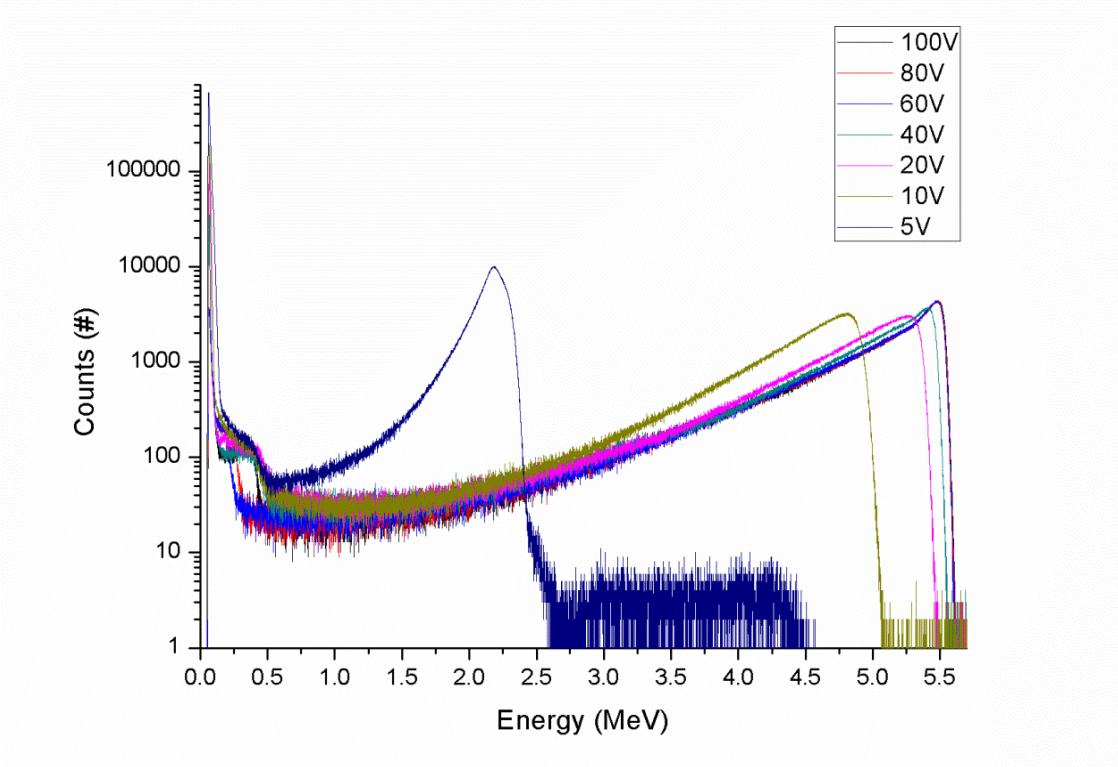


Fig. 5.59: Alpha spectrum with varying the bias voltage of the Si planar type detector.

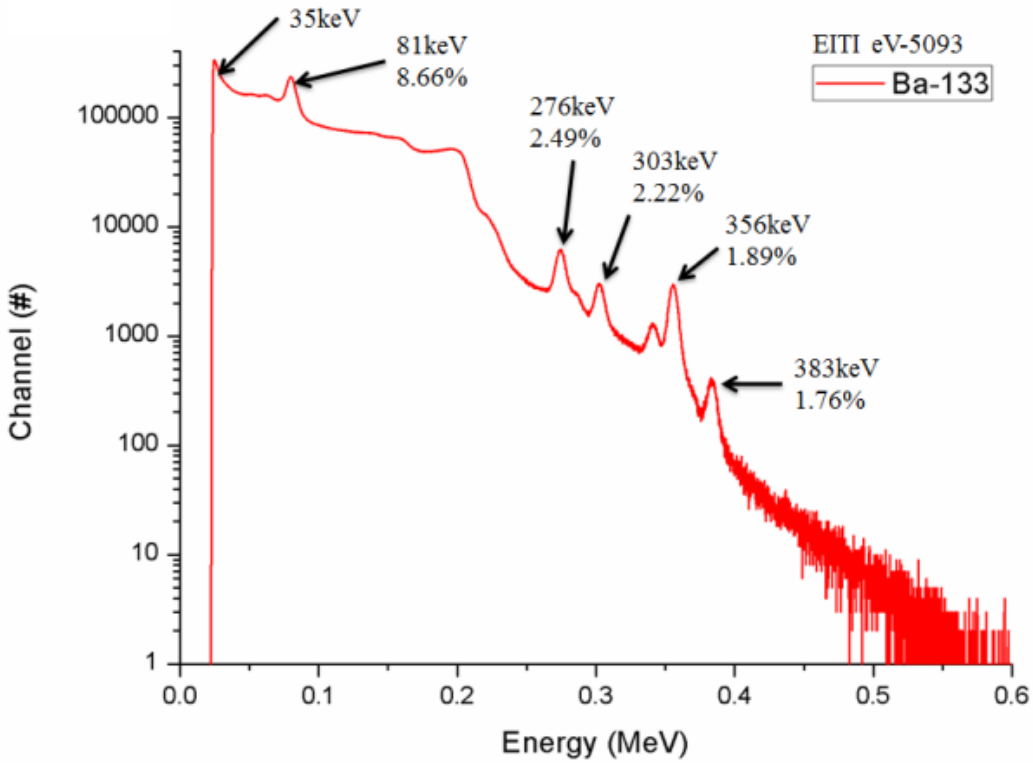


Fig. 5.60: Energy resolution of 1.89% of full width at half maximum (FWHM) were obtained for the 356 keV photopeak of Ba-133 with eV-5093.

With the measurements setup shown in Fig. 5.54, we measured low-energy gamma-ray spectra from the following sources: Ba-133, Co-57 and Cd-109, with main photopeak energies as follows: 122 keV (Co-57), 356 keV (Ba-133) and 88 keV (Cd-109). The gamma-rays following the electron–capture (EC) decay of Ba-133 at 35keV, 81keV, 276keV, 303keV, 356keV and 383keV have been observed as Fig. 5.60. The energy resolutions of the 356 keV photopeak are 2.29%, 2.25% and 1.89% for the three different preamplifiers described above (cases A, B and C); that is, the ev-5093 yields the sharpest features and the preamp introduces non-negligible noise.

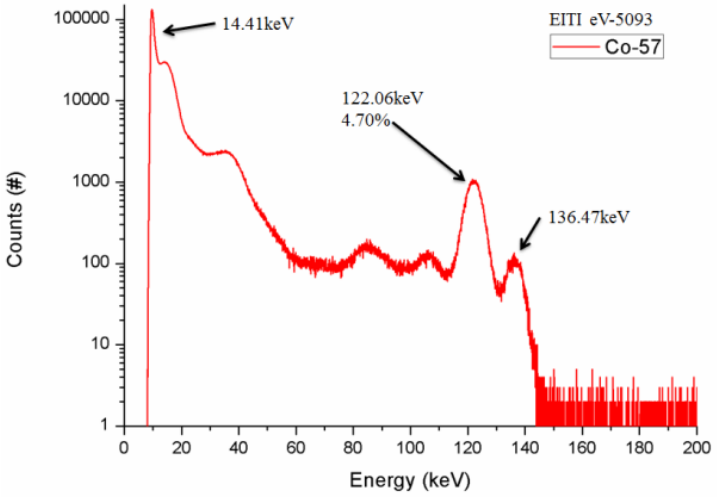


Fig. 5.61. Energy resolution of 4.7% of full width at half maximum (FWHM) were obtained for the 122 keV photopeak of Co-57 with eV-5093.

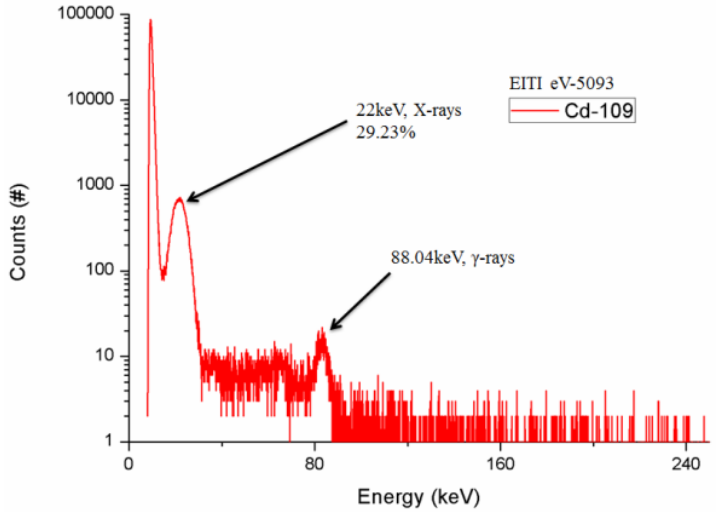


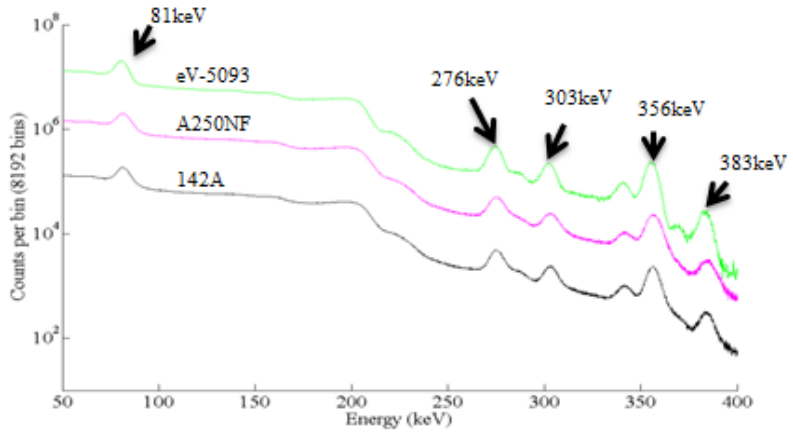
Fig. 5.62. Energy resolution of 29.23% of full width at half maximum (FWHM) were obtained for the 22 keV photopeak of Cd-109 with eV-5093.

The gamma-rays following the EC decay of Co-57 at 14keV, 122keV and 136keV have been observed, as shown in Fig. 5.61. The energy resolution of the 122 keV photopeak is 5.82%, 5.11% and 4.7% for three preamp cases (A, B and C), and the gamma-rays following the EC decay of Cd-109 at 88keV and X-rays at 22keV have been observed. The energy resolution of the 88 keV of gamma-ray is 6.6%, 5.9% and 5.6% for the case of A, B and C as shown as Fig. 5.62. Table 5.1 shows the test results with different low energy gamma sources.

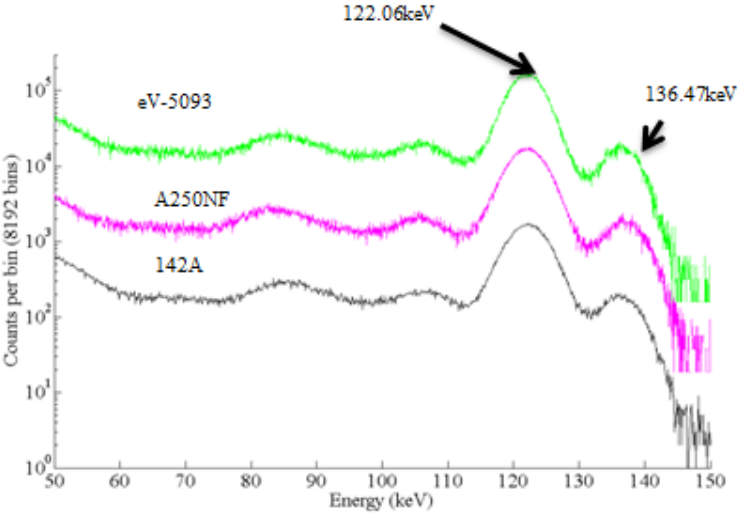
Table 5.1. Energy resolution test results with 3 type's preamplifiers and low energy gamma-ray sources under -100V revers bias and 0.5 μsec shaping times and 100 gain.

Preamplifiers	Source	Ba-133					Co-57		Cd-109	
	Energy(keV)	81	276	303	**356**	383	**122.06**	136.47	22, 25	**88.04**
Ortec 142A	H	326.89	1117.6	1232.07	1448.69	1562.33	1133.19	1268.29	226.04	819.72
	FWHM	30.76	31.44	29.32	33.28	30.34	66	46.67	61.15	54.04
	R(%)	9.409893	2.813171	2.379735	**2.297248**	1.941971	**5.824266**	3.6797578	27.05273	**6.592495**
Amptek 250NF	H	767.16	2642.45	2912.04	3431.8	3703.15	2440.79	2717.76	222.13	865.11
	FWHM	73.68	78.11	64.34	77.39	58.88	124.82	91.73	57.46	50.94
	R(%)	9.604255	2.955969	2.209448	**2.255085**	1.589998	**5.113918**	3.3752061	25.86774	**5.888269**
EITI eV-5093	H	705.58	2417.56	2666.95	3135.61	3382.35	2217.77	2476.29	201.75	777.92
	FWHM	61.08	60.27	59.3	59.24	59.53	104.45	96.95	48.98	43.81
	R(%)	8.656708	2.493009	2.223514	**1.889266**	1.760019	**4.7096859**	3.9151311	24.27757	**5.631684**

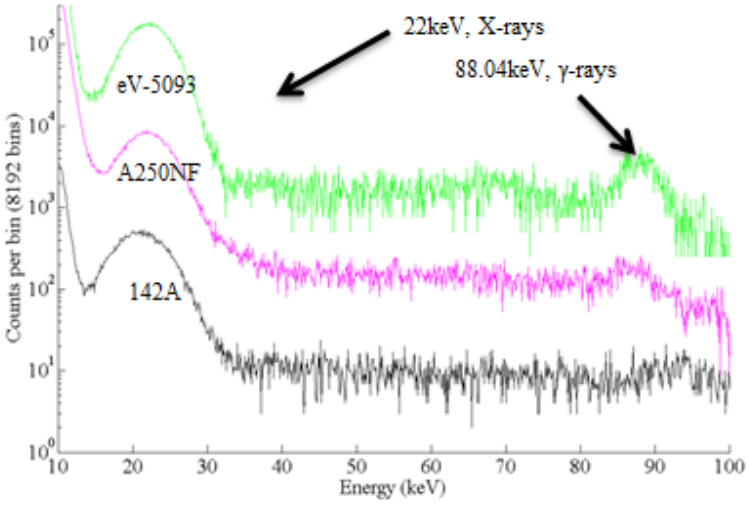
Figs. 5.63 show the spectrum comparison for the Ba-133, Co-57 and Cd-109 with 3 different types of preamplifiers. With a refined getter recipe, lower-leakage detectors are possible. Specifically, if the leakage of the 1 x 1 cm^2 detector is reduced from 150 nA (at 100 V), which existed for the Fig. 5.63 spectra, to 5 nA, then the peaks sharpen, as shown in Fig. 5.64, for the Ba-133 source.



(a)



(b)



(c)

Fig. 5.63: The gamma energy spectrum of Ba-133 (a), Co-57 (b) and Cd-109 (c) with different type of preamplifiers.

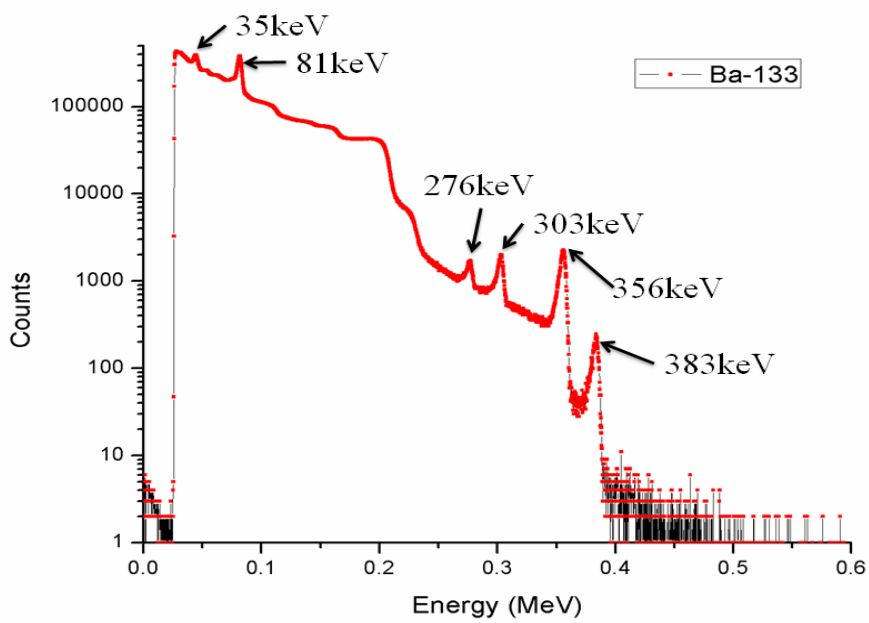


Fig. 5.64: The gamma energy spectrum of Ba-133 with low leakage Si meander line detector which has 5 nA at 100 V.

5.6.2. Position sensing with the meander line detector

Position sensing can be elicited via either timing or amplitude variations. As discussed in previous chapters, the main physical mechanism through which we ascertain the position via *amplitude*-based methods is through the phase-velocity dispersion and RC stretching of the pulse-shape. Recall that one can use an asymmetric strip-line to preserve the pulse shape, and one then extracts the position from the pulse timing and the energy from the steps in the signal, the advantage of which is that longer lines can be employed and more detailed information can be extracted.

However, if one wishes to simply use the detector as an energy and position measurement instrument, then one can eschew the overlay and permit pulse shaping. The degree of pulse shaping then provides a measure of the position-of-interaction along the transmission line. Practically speaking, the current step is elongated and thus diminished in amplitude as the pulse traverses the electrode, the consequence of which is that the *slope* in the charge-integrated signal is reduced as the interaction position increases. We measure that slope by taking the derivative of the preamp signal and use that as a measure of the position. In the following, we compare this phase-velocity distortion technique

with a timing method, in order to gauge the resolution with which the position can be extracted with the most simple transmission-line design.

Although this pulse shaping is the main mechanism through which we gauge the position, for longer lines, the resistance of the conductor can also play some role. As shown in Fig. 5.65a, the general circuit configuration consists of a photodetector and a transimpedance amplifier (TIA) which converts the current from the photodetector into a voltage at the output node of the TIA. Fig. 5.65b is an equivalent circuit modeled by the photodetector. The special gamma ray photodetector has a meandering line which is presented as two resistors R_{m1} and R_{m2}. Those resistors are variable due to a position where the gamma ray comes in the photodetector. Ideally, by measuring the current flowing through R_{m1} and R_{m2}, one can determine the gamma-ray interaction position.

Let V_x be the voltage across R_{m1} and R_{m2}.

$$\frac{V_x}{R_{m1}} + \frac{V_x}{R_{m2}} = I_P \tag{5.1}$$

I_p is a photocurrent generated by the incident gamma ray into the photodetector. If a current I_1 flows to TIA_1, then we can obtain the I_1 as follows:

$$\frac{V_x}{R_{m1}} = I_1 \tag{5.2}$$

Substituting Eq. 5.2 into Eq. 5.1,

$$I_1 = I_P \left(\frac{R_{m2}}{R_{m1}+R_{m2}}\right), \qquad I_2 = I_P \left(\frac{R_{m1}}{R_{m1}+R_{m2}}\right). \tag{5.3}$$

I_2 follows similarly. By comparing the amount of these current values, one can find the position of incident gamma ray in the. The resistance method is more applicable for a balanced transmission-line because its effect is negligible compared with the shaping inherent to a standard micro-strip line.

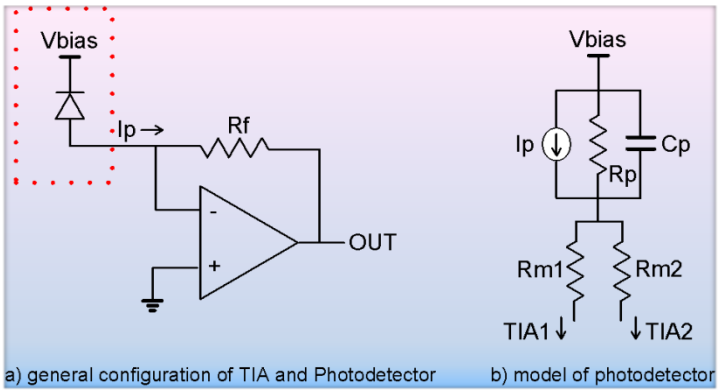


Fig. 5.65: The general configuration of TIA and photo detector (a) and the model of photodetector (b).

With 5nA leakage at 100V reverse bias voltage and the Am-241 alpha source collimated at one side of the 50 μm width meander lines, the meander line Si detector was used to acquire the time difference and amplitude difference at both ends, using the charge sensitive amplifiers as shown at Fig. 5.66. The system has 3 outputs- from preamplifiers for the both ends of meander line as well as the back side cathode. The shaped and amplified signal from the cathode was used as the trigger signal.

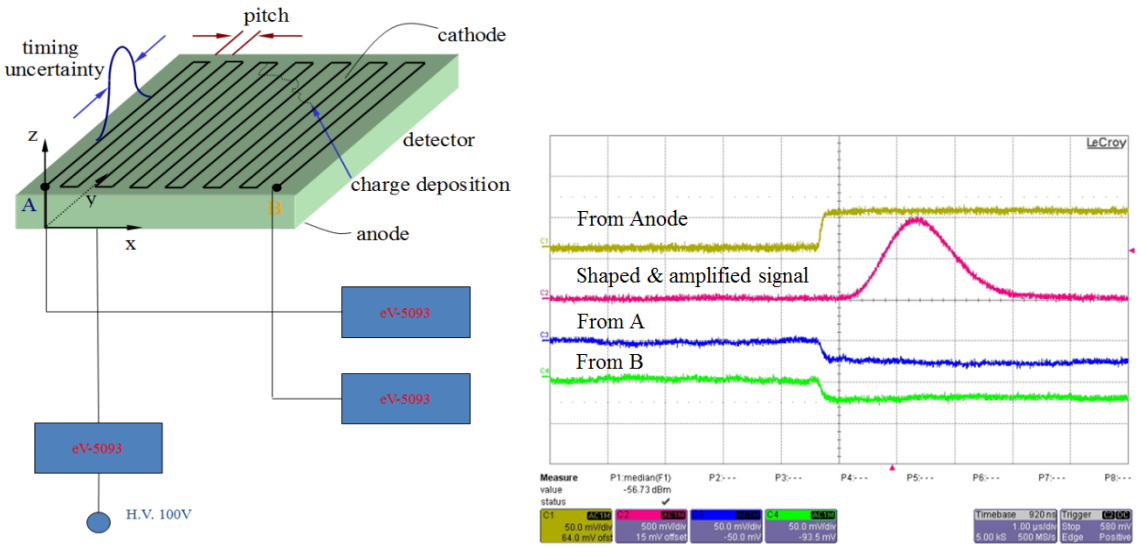


Fig. 5.66: The system configuration for the signal acquisition from anode and two ends of cathode with eV-5093 preamplifiers (left) and acquired signals (right).

For the signal acquisition, an oscilloscope (Lecroy) which has 2GHz bandwidth and 5Gs/s sampling rate was used. As shown in Fig. 5.67, outputs from the 3 preamplifiers (eV-5093) were acquired. The blue signal near the yellow signal (from the cathode) was calculated by signals of both meander ends (green and red). Fig 5.68 shows the digital shaped signals from ends of meander line (thick green and red), cathode signal (blue) and sum with opposite from both ends of meander (black). A digital shaping amplifier was designed using CR-RC^2 shaping and implemented in a Matlab code. In order to calibrate the position of interaction, a collimator made of copper tape was arranged as shown in Fig. 5.70. Specifically, the top edge of corners was exposed to the radiation; thus, the time or amplitude spectra should count the number of corners in the detector which is half of the number of lines in the meander-pattern. Fig. 5.71 shows examples of signals for the different interaction position at both meander ends. When the interactions occur at the extreme ends there are amplitude differences and time delays between the two signals, as shown in Fig. 5.71. In order to use the time difference between the signals at the meander ends for the position determination, the constant time fraction method is applied as shown in Fig. 5.72. With these acquired signals, amplitude differences between the meander ends are calculated for the position sensing as shown Fig. 5.73.

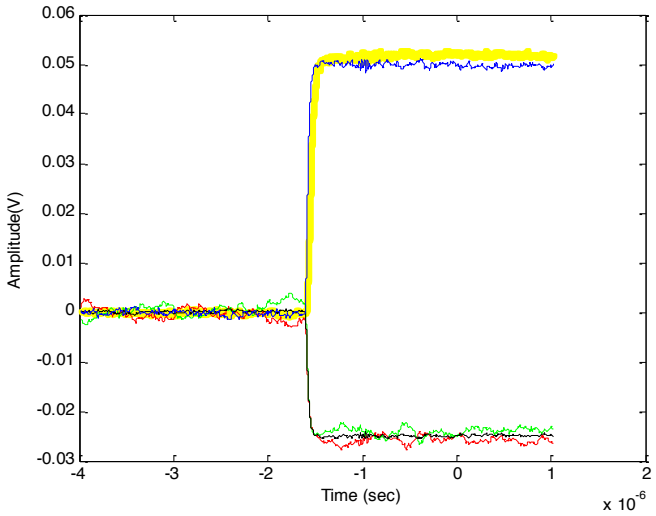


Fig. 5.67: Preamplifiers signals from both ends of meander line (green and red), their average signal (black), cathode signal (yellow) and sum with opposite from both ends of meander.

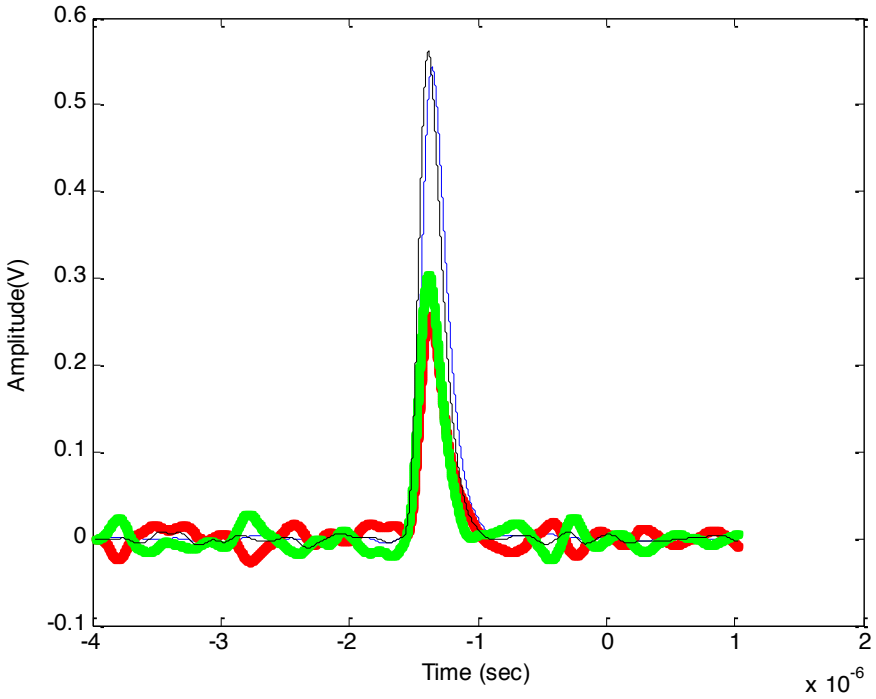


Fig. 5.68: Digital shaped signals from both ends of meander line (green and red), cathode signal (blue) and sum with opposite from both ends of meander (black).

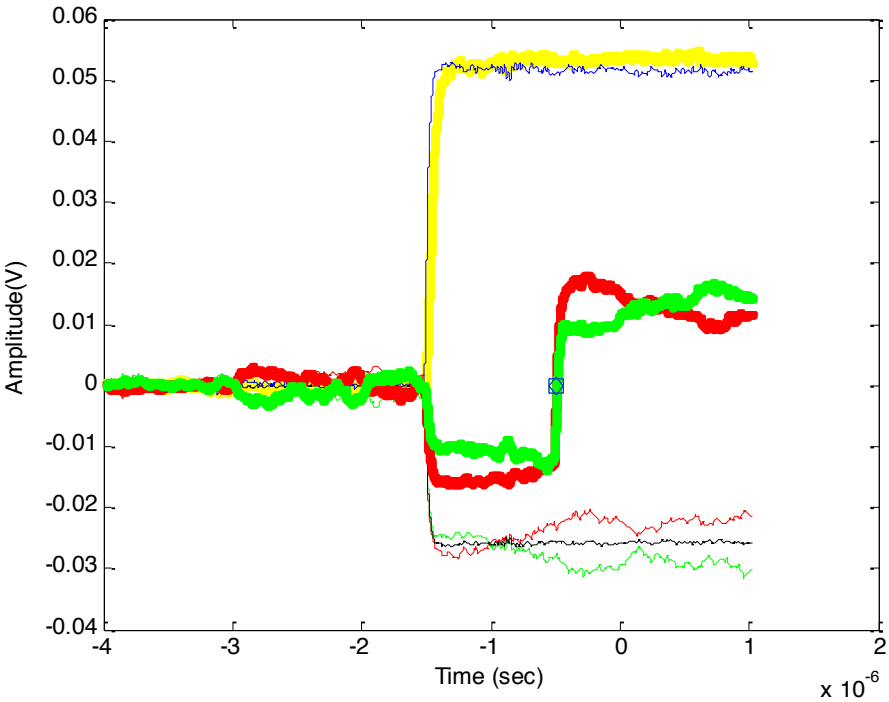


Fig. 5.69: Signals after applying the Constant Time Fraction (CFT) methods for both ends of meander (black). Note that small rectangular marks correspond to each t_0.

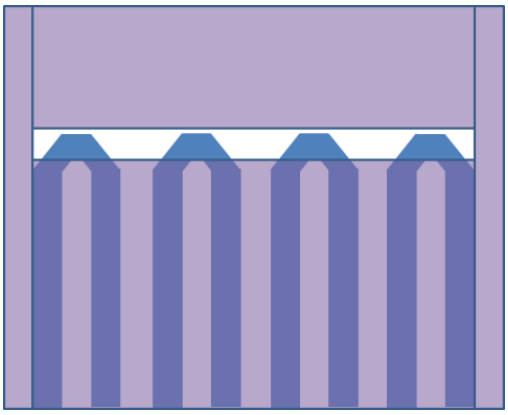


Fig. 5.70: The collimation at the each corner of 50 μm width and 8 lines meander. Note the blue line is the meander line and purple block is the copper collimator.

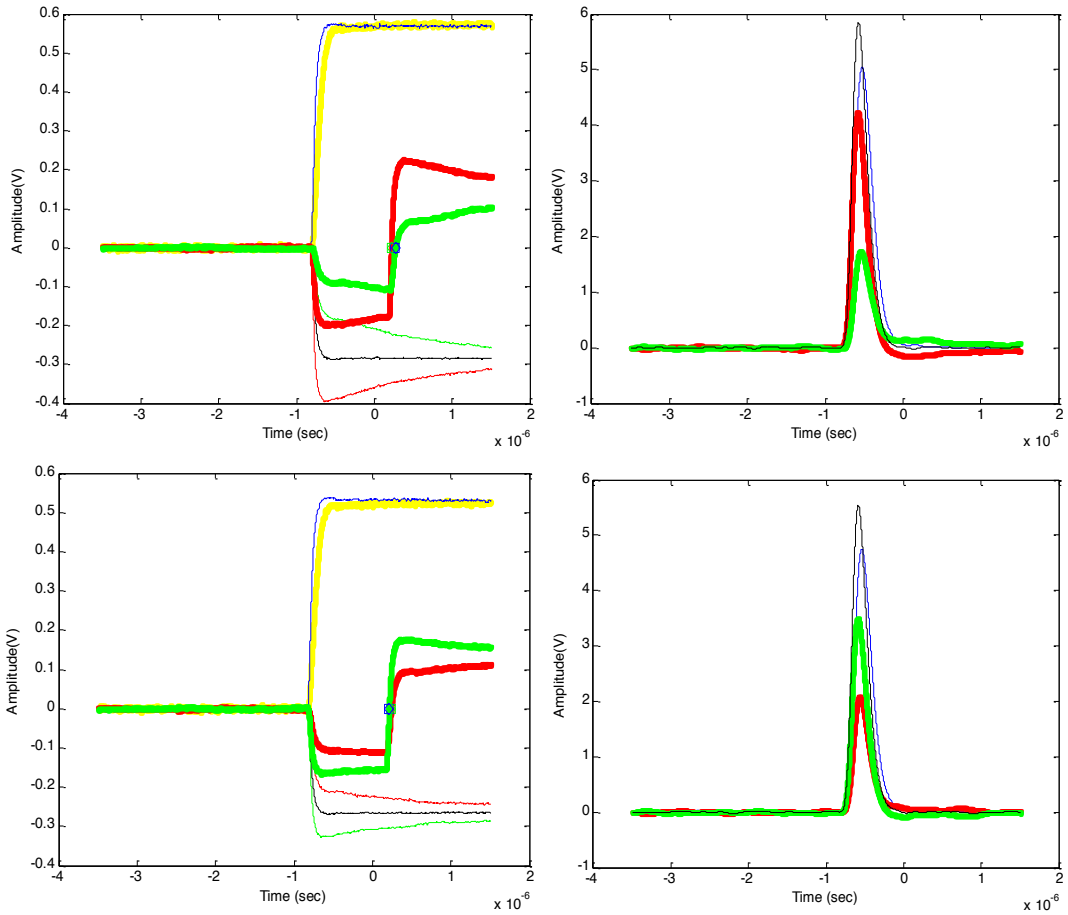


Fig. 5.71: Signals acquisitions when the interaction happens at left meander end (top) and when it happen the opposite site (bottom). Note that the pulses of right side are corresponding to their amplitude comparison.

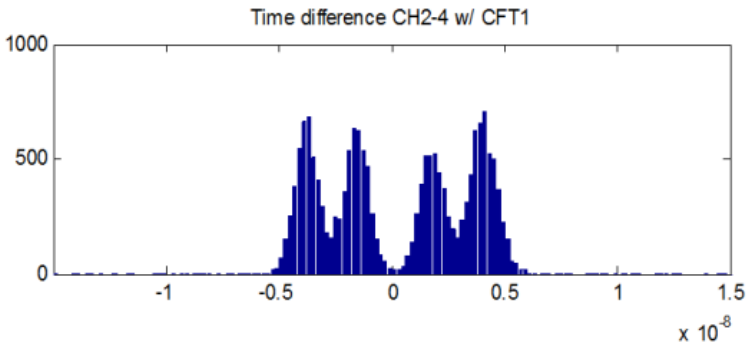


Fig. 5.72: Time difference histogram from the signals at both meander ends for the 50 μm width, 4 turns.

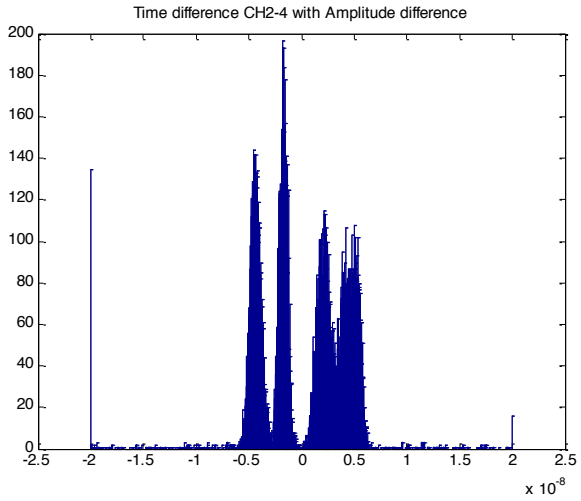


Fig. 5.73: Amplitude difference histogram from the signals at both meander ends for the 50 μm width, 4 turns.

The time spectra in Fig. 5.72 from the four-turn meander, indeed shows four peaks, in which the outer peaks are larger, in area, because of greater effective detection volume associated with the edge strips. Also note that the time difference between the two middle peaks is greater than the two outer features, which reflects the fact that signal propagation is more shaped by the increased capacitance and inductance of interior lines. The Fig. 5.73 amplitude distribution reflects the feature that the central peaks are generally sharper than the edge peaks because of superior SNR due to the higher signals amplitudes. That is, nearer the edge of meander line, the far pulse experiences a greater degree of pulse shaping, and the resulting signal is more adulterated by the electronic and thermal noise.

This trend is reflected in Fig. 5.75, in which the edge lines reach further down than the central line. That is, Fig. 5.74 shows the distribution of interaction positions corresponding to the time difference spectrum in Fig. 5.72, and Fig. 5.75 shows the alpha source image with collimation as shown at Fig. 5.70. The broadness in the horizontal direction is due to the line-collimation (rather than hole collimation), and the uncertainty in the vertical direction reflects the uncertainty in the position determination along a given line. According to the result of the time spectrum in Fig. 5.72, the transition time of the current between two lines was 2.032 ns, which means that the propagation speed is a rather slow 1.016 ns per 1 cm. However, this transition time is increased by increasing the number of turns because of increasing the inductance. Also, when the width of meander is decreased the transition time is increased.

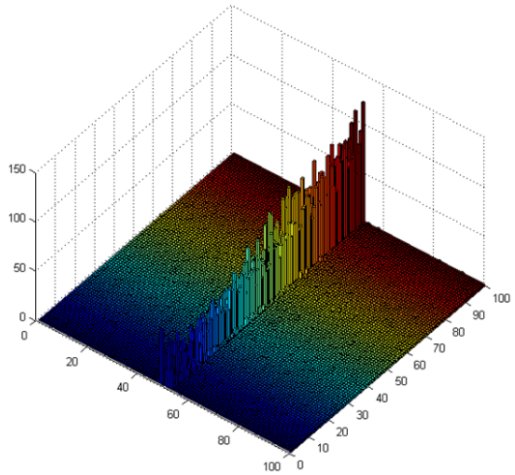


Fig. 5.74: Probability of the interaction position at the 50 μm width 4 turns meander line detector.

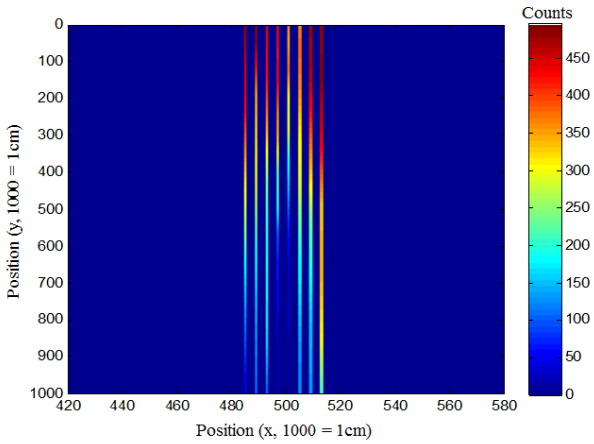


Fig. 5.75: Alpha source image with collimation like Fig. 5.70 for the 50 μm width 4 turns meander line detector.

For the purpose of evaluating the meander line detector, several detectors which have different number of turns were used, and time-difference and amplitude-difference spectra were measured, as shown in Figs 5.76 and 5.77. With these acquired data the alpha source collimated images were acquired as shown in Fig. 5.78. As shown in the figures, the total time difference and amplitude difference are increased by increasing the turn number. For the 4 turns, 6 turns and 12 turns, total time differences were 8 ns, 12 ns and 24 ns, respectively.

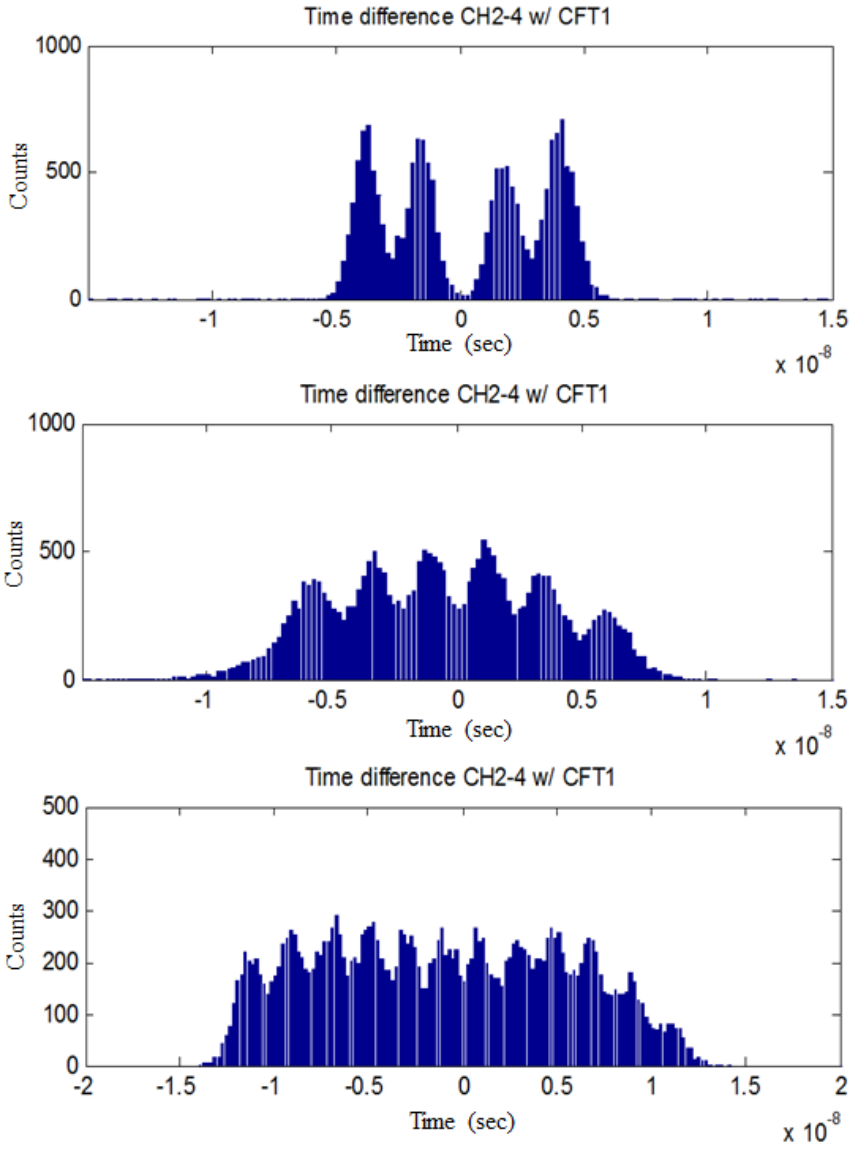


Fig. 5.76: Time difference histogram from the signals at both meander ends for the 50 μm width 4, 6 and 12 turn.

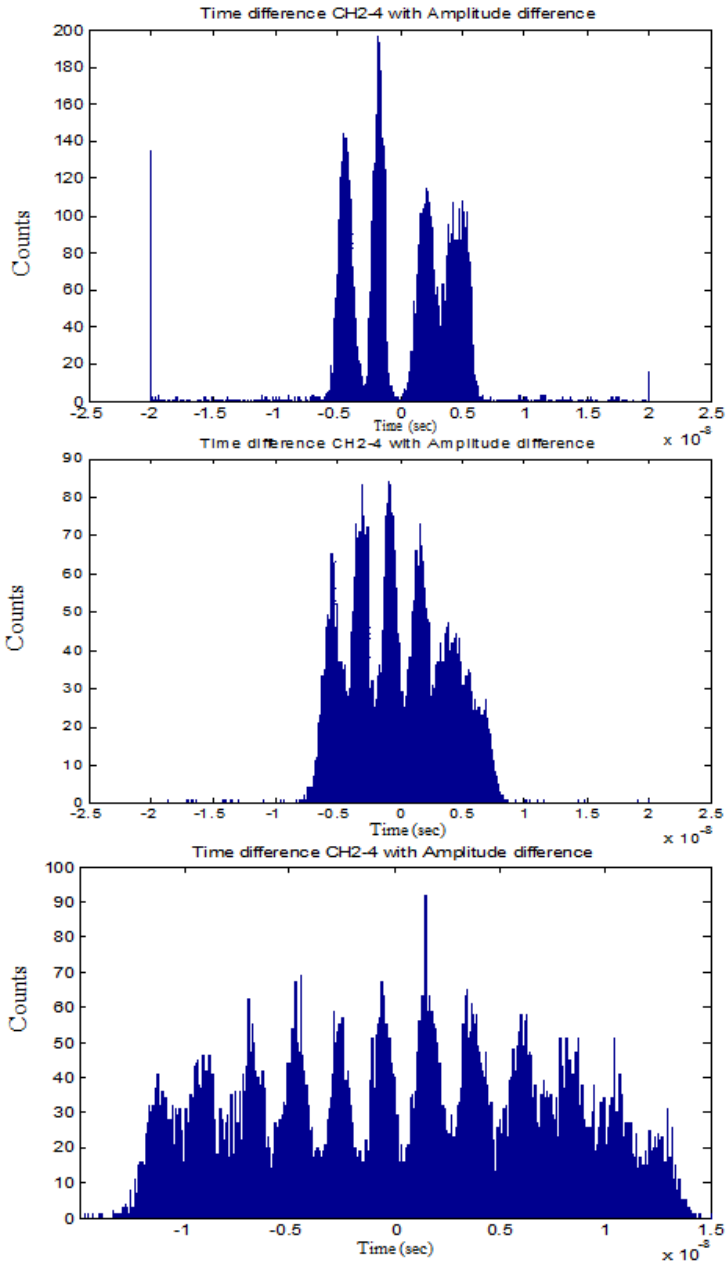


Fig. 5.77: Amplitude difference histogram from the signals at both meander ends which have 4, 6, 12 turns of the 50 μm width.

For the sake of further demonstrating position sensing with the meander line detector, the source was collimated to the upper half of the 24 line detector, resulting in the time spectrum of Fig. 5.79. With this information, the alpha source image with half collimation can be acquired, as shown in Fig. 5.80.

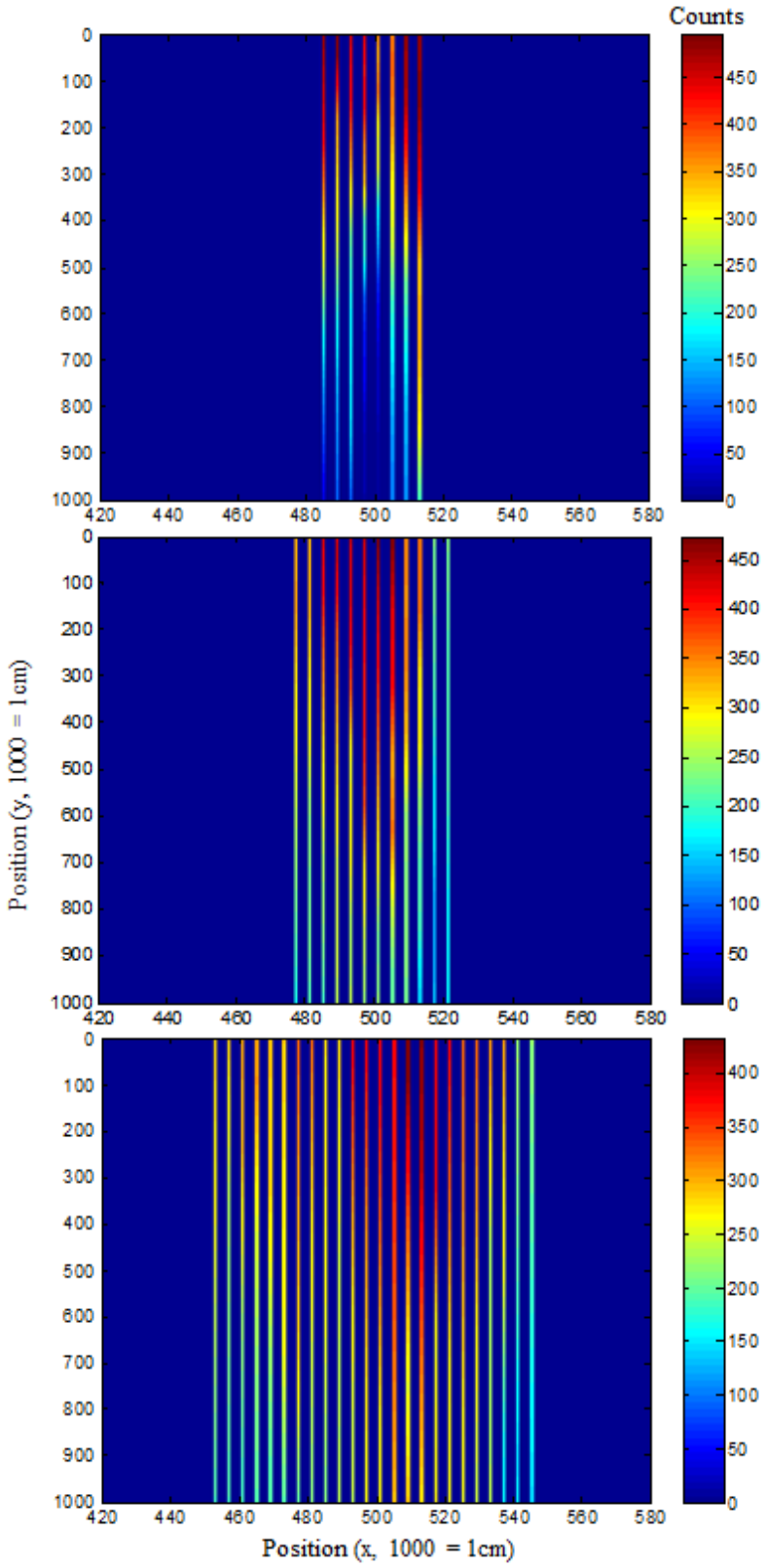


Fig. 5.78: Alpha source image based on amplitude variation with meander detectors which have 4, 6, 12 turns of 50 μm width.

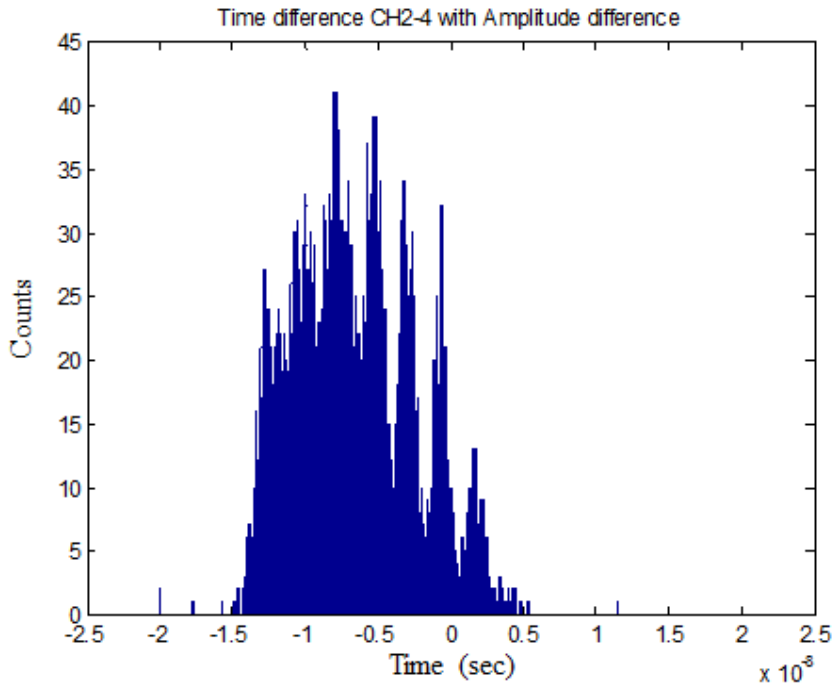


Fig. 5.79: Time difference histogram from the signals at both meander ends for the 50 μm width 12 turn with half collimation.

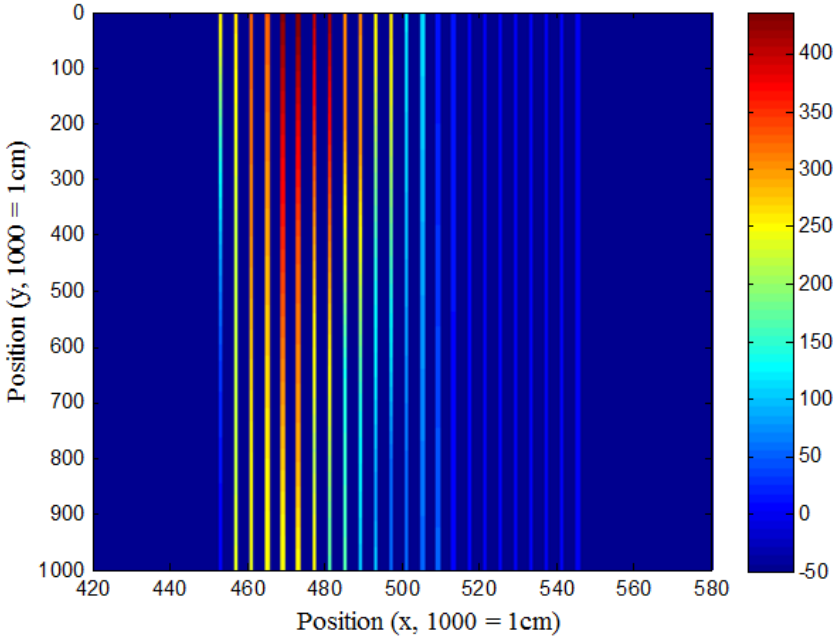
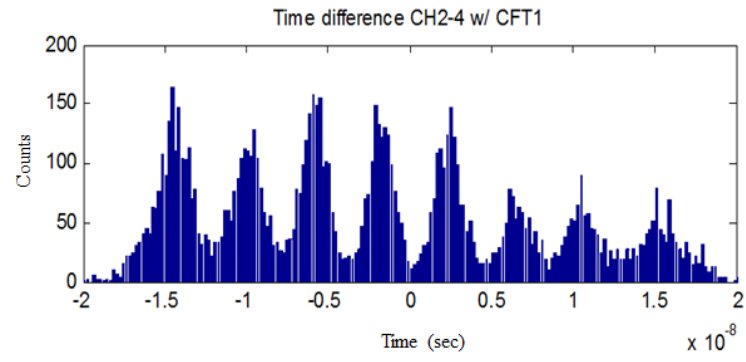
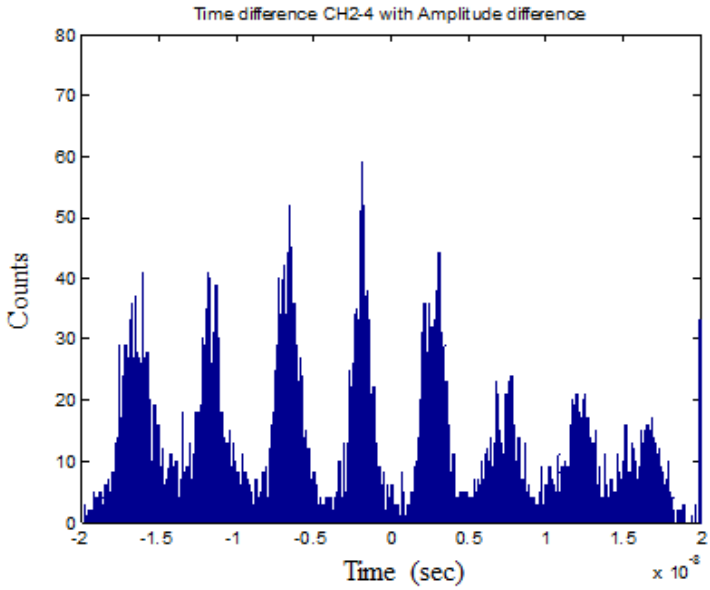


Fig. 5.80: Alpha source image of the meander detector which has 12 turns with half collimation.



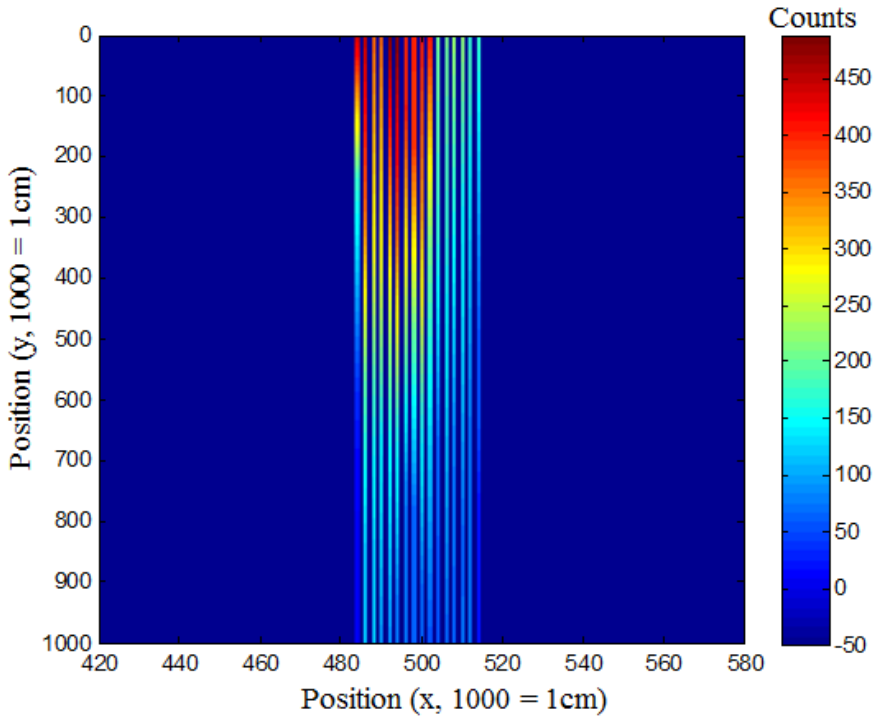
(a)



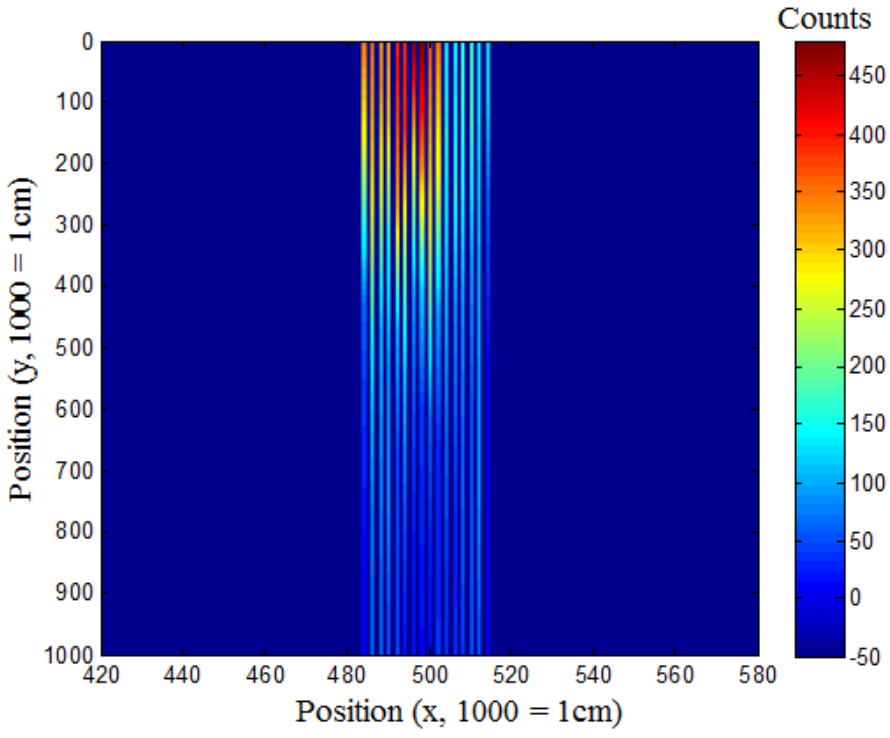
(b)

Fig. 5.81: Time difference histogram (a) and amplitude difference histogram (b) from the signals at meander's both ends for the 25 μm width 8 turn.

In order to validate the effect of varying the width of the meander line, 8 turns of the 25 μm width meander line detector was used. Fig. 5.81 shows the time difference histogram and amplitude difference histogram. The transit times between each line are from 4.08ns to 4.58 ns, which implies a propagation speed of ~2.04 ns per cm, which is almost double that measured for the 50 μm width meander (1.05ns). Favorably, smaller partitions result in lower propagation speeds and thus greater position clarity. With these acquired time difference and amplitude difference spectrum, the collimated alpha source images are acquired as shown at Fig. 5.82.



(a)



(b)

Fig. 5.82: Alpha source image of the meander detector which has 8 turns with using the time difference method (a) and amplitude difference method (b).

As shown Fig. 5.81, the position uncertainty due to timing is poorer than that realized when one employs the amplitude difference histogram. This is also reflected in the alpha source image, as shown in Fig. 5.82, in which the "reach-down" of the uncertainty in red is less for the amplitude-difference case. The spatial resolution is dependent on the meander's pitch size. Using either time-differences or amplitude differences, the position of interaction can be isolated to a single line as shown.

CHAPTER VI

CONCLUSION, SUMMARY AND FUTURE WORK

6.1 Conclusions

In conclusion, we have succeeded in developing the modeling tools for investigating different position-sensing schemes and used those tools to examine both amplitude-based and time-based methods, an assessment that indicates that fine position-sensing can be achieved with simpler readout designs than are conventionally deployed. This realization implies that one can deploy position-sensitive semiconductor detectors with only one or two contacts per side, which can increase the feasibility of deploying ubiquitous and inexpensive special nuclear materials (SNM) detecting technology. For the immediate research, which is focused on the materials-improvement issues, the further development of this capability allows one to more powerfully assess the behavior of charge-transport because the behavior over large areas can be rapidly assessed.

Thus, the existing electrical characterization tools have allowed us to develop suitably robust contacts and identify those general factors which can detrimentally trap charge carriers. The methods by which one can maximize the charge creation and minimize its inherent fluctuation have been previously studied in the context of optical photon detectors, and can, in principle, be applied for the fast charge-particles which result from nuclear radiation interactions. Finally, the numerical and analytical tools have been used to design a fine position-sensitive readout method for analyzing the charge motion within semiconductor-based detectors. We also have succeeded in fabricating robust contacts between the semiconductor and the readout electrodes and used those contacts to investigate both amplitude-based and time-based methods of extracting the lateral position of interaction with minimal hardware complexity. We have identified the limitations of each method and deduced that delay-line differential pulse sensing provides the opportunity for micron-scale position sensing but with a minimal number of readouts;

at best, only one readout per detector face is required, in principle. We have proven this concept for a relatively slow, noisy system and expect to complete the fabrication and testing of a high-speed design.

6.2 Summary

In Chapters I and II, we have described the delay-line radiation detector and its optimized fabrication. The semiconductor physics were simulated, the results from which guided the fabrication of the guard ring structure and the detector electrode, both of which included metal-field-plates. The measured improvement in the leakage current was confirmed with the fabricated devices, and the structures successfully suppressed soft-breakdown. Furthermore, we confirmed that: (1) minimizing the gap between the innermost guard-ring and the active diode region, and (2) extending the width of the metal field plate (within the constraints of the processing) optimized the noise properties. We also demonstrated that fabricating an asymmetric strip-line structure successfully minimizes the pulse shaping and increases the distance through which one can propagate the information of the deposited charge distribution.

In Chapter II, the focus of work was to optimize the detector fabrication recipes and develop the modeling tools required so we can rapidly test various electrode partitions and find the pattern or family of patterns that allows fine cloud-mapping but with minimal readout complexity. Our initial analyses indicated that amplitude-based techniques- in which the electric-field is made non-uniform across the entire detector can be used to achieve this goal. With the appropriate physics packages installed, GEANT4 provides an accurate gauge of the radiation response in the detector and MAXWELL and semiconductor modeling tools allows one to calculate the electric-field environment in the detector. The packaged and custom codes therefore provide an accurate model of the electrical signals formed on the detector electrodes, signals from which the original charge-cloud configuration is gleaned. The detector model is thus used to develop the signal processing chain from which the radiation-image is formed. With the successful realization of both n-type and p-type silicon surface-barrier devices which exhibit low leakage current, the electrical readout development has continued with the successful realization of both fast-timing circuits and current-sensitive amplifier chains. Therefore, the analyses indicate that recoil-electron tracking- and the concomitant enhanced

radiation imaging can be achieved. The principles have been analyzed and modeled, and complementary methods of realizing the cloud mapping goal are being investigated in numerical models. Although Coulomb drift and thermal diffusion can potentially limit the resolution with which one determines the initial charge distribution in the detector, both effects can be counteracted through readout techniques such that micrometer-scale position sensing is possible. In particular, the timing and size of the transition in the current pulse during charge collection at the electrode provides an estimate of the quantity of charge produced at a given position. The effect of Coulomb drift on that estimate can be removed, in principle, by analyzing the pulse shape during the initial stages of the charge drift. Nevertheless, the simulated results imply that the space-charge effect does not significantly degrade the positional or charge determination. Thermal diffusion can degrade the precision of that determination, however, especially if only a handful of carriers are tracked in the detector. In order to mitigate its effect, the charge carriers should be drifted at their saturated velocities. In order for either effect to significantly limit the precision in the determination of the initial charge position, one must employ a low-noise readout that is sensitive to hundreds of electrons at most, which one can achieve by using on-chip transimpedance amplifiers. Therefore, for detectors and readouts with high charge-sensitivity, the methodology described in Chapter II can be employed to quantify the degree to which the intrinsic uncertainties in the position reconstruction, due to either thermal diffusion or Coulomb drift, can impair that determination.

In Chapter III and V, we demonstrated that the metal field plates mitigate the leakage current and increase the breakdown voltage and inner guard-ring biasing helps to reduce the leakage current. In addition, thick oxides for the passivation helps to reduce the leakage current, and gettering can reduce the bulk leakage current and increase the breakdown voltage. Furthermore, short diffusion depths of the p+ doping can result in low leakage current at the cost of uniformity. A long diffusion depth can increase the leakage current, but it helps to smooth the electric field and increase the breakdown voltage. With this process we produced low leakage-current and high breakdown voltage device and acquired alpha spectra (Am-241) and gamma spectra (Ba-133, Co-57 and Cd-109). These techniques can also be applied to other materials, and we demonstrated the

methods by which one can robustly process the delay-line pattern on CZT with gold contacts. The successful quenching of the surface leakage has resulted in detectors that therefore yield noise performance that is limited by their inherent bulk properties, and those bulk properties have been improved, in the case of silicon, with optimized gettering recipes. Recall that the detectors can therefore be used to extract the charge information from both ion and gamma-ray interactions. Furthermore, standard charge-sensitive circuits yield high SNR pulses. The detectors and existing electronics can therefore be used to yield imaging instruments for neutron and gamma-rays, in the case of silicon, as we showed previously. For CZT, we would prefer to utilize current sensing to be able to clearly isolate the effects of the various charge-transport non-idealities, the full realization of which awaits the fabrication of the custom-designed TIA chip.

The degree to which one can improve the imaging (angular and energy) resolution for either neutrons or gamma-ray depends on the sensitivity with which one can extract the initial charge information, which can be muddied by detector and electronic noise, as well as by the inherent charge losses associated with transport through imperfect solids. We have successfully made both silicon and CZT detectors for which the detector noise is small compared with the electronic noise, and more importantly, small compared with the slight current fluctuations that accompany radiation interactions. In the case of silicon, the low noise detectors were enabled by a new generation of furnaces implemented in our clean room, with a refined gettering recipe which results in 10 – 1000 times improvement in the leakage current, compared with our detectors made earlier in the work. For the case of CZT, we have implemented surface passivation schemes that result in detectors with only 10's of pA of leakage near the nominal operating values. The detectors are therefore optimized for the imaging applications envisioned.

In Chapter IV, we performed the quantitative and qualitative performance comparisons of SBP and FBP algorithms with and without electron-tracking ability. For the robust quantitative assessment of different image reconstruction algorithms and for removing statistical fluctuations that might be caused by sparse radiation source distribution, we defined a new performance measure, "distance-weighted root-mean-squared (DWRMS) error", indicating not only the bias and variance of the estimates of image intensity but also the spread of PSF. Based on the DWRMS errors, four different

types of back-projection based algorithms (SBP versus FBP and with electron-tracking versus without electron-tracking) were compared and simulation data from MCNP5 were used to quantify the degree of improvement provided by tracking recoiled electrons' paths for SBP and Parra's FBP algorithms. Depending on the spatial resolution of the radiation detector used for tracking electrons' paths, the SBP and FBP with electron-tracking ability could reduce the DWRMS error from 60% to 95% compared to the conventional SBP and FBP algorithms, not exploiting directional information obtained from tracked electrons' paths. Therefore, we concluded that electron-tracking in advanced Compton scattering imaging could provide a substantial amount of improvement in terms of the bias and variance of image reconstruction, and the spread of PSF.

The extraction of the current pulses requires a very sensitive, low noise, high gain, high speed amplifier, all of which are not provided by any commercial chip. We therefore designed, at the transistor level, a transimpedance amplifier that has excellent current sensitivity (< 1 nA), high gain (> 1 MV/A), and high bandwidth (~ 2 GHz). We plan to tape-out the chip in a 130 nm CMOS process. Recall that the successful marriage of the chip with the detector has the following consequences. First, one can, in principle, use a planar CZT- with all of its transport inadequacies- and produce an excellent energy spectrum by concentrating on the initial part of the pulse. Second, one can more clearly map the charge trajectories if the sensitivity is improved beyond the current state-of-art, which simulations of the chip design indicate is possible.

Finally, we have developed either preliminary, mature, or final solutions for all aspects of the problem, and are putting the parts together to prove the concept for the envisioned neutron and gamma-ray imaging applications.

6.3 Recommended Future Work

We have further determined that on-chip amplifiers in the form of die-form components have the metrics to allow us to finely map the charge distributions; therefore, in situ microelectronic fabrication is not necessary, though it may still be pursued. Thus, the methods developed can be readily applied to other semiconductor materials because they rely on evaporated metal contacts for the electrode structures, rather those doped and insulating layers, for which the fabrication is more involved.

We began to develop the methodologies by which the charge distribution can be mapped, concentrating on the depth dimension primarily. Furthermore, following the successful testing of the second to fifth generation devices, as described in the Chapter V, we plan to apply those devices to full-scale 3D mapping of the initial charge distributions. The goal of such mapping is to determine the optimal method by which one can quickly and cleanly extract the information so that the radiation imaging resolution can be enhanced via inclusion of the recoil electron trajectory information.

Our focus will also shift to the application of the technology in which neutron imaging will be the initial imaging focus because of our existing capability of imaging ion tracks, followed by gamma-ray imaging tests once the charge sensitivity of the amplifiers and the noise of the detectors are optimized. Our commercial collaborator also expects to begin the commercialization of the detectors, now that stable low-leakage designs have been obtained. The development of commercial modules was completed during the work, in which we designed the module that can house a charged particle planar detector with floating guard rings, and we used an external company for their manufacture. In order to avoid the non-uniformity of the boron doping, we will use the ion implanter or new furnace with gas instead of using the furnace with solid target (uniformity is decided by length between source and wafer) and we will change the distance between the active area and the inner guard-ring to bias on the inner guard-ring and reduce the leakage current. We will also do a delay-line patterning on CZT and TlBr crystals with passivation layer and guard-rings. Finally, we will find a proper TIA (high bandwidth and gain) to sense the small current signal from our meander line devices.

We are also trying to make a customized TIA (transimpedance amplifier) with high-speed and wide bandwidth operational amplifiers. After verifying the performance of this TIA module, we will try to put together with meander-line Si, CZT and TlBr detectors to acquire current signals and verifying our concepts. Furthermore, in order to enhance the energy spectrum for the planar CZT or TlBr detector which has a poor energy resolution and energy loss due to the defects on CZT detector and the polarization on TlBr, we will apply this technique in order to get the initial energy amplitude which corresponds to the initial value of electron hole pairs.

Appendix A

Matlab code to design the meander line detector

```
clc; clear all; close all;
% Microstrip line
Parameters%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Physical structure
   c = 2.998e8;      % Speed of Light in free-space [m/s]
   f = 1;           %Frequency
   mu0 = 4e-7*pi;    % H/m magnetic constant, vacuum permeability
   sigma = 37.6676e6 ; % conductivity of Al (S/m)
   We = 100e-6;      % Width (meter)
   H = 500e-6;       % Substrate Height (meter)
   H1 = 0.5e-6;      % Overlay Height (meter)
   H2 = H+H1; S = We; % Gap size
   Er = 11.7;        % Relative Dielectric Constant of Si
   E0 = 8.854187817620e-12; % F/m
   Esio2 = 3.9;   % Sio2's relative permittivity
   N = 120*pi; U = We/H; U1 = S/H1; U2 = (H+H1);
   t = 0.5e-6;     % Metal thickness
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% For the single micro strip
% H1 < H  && H1 < We/2 condition
    if (We/H) >= 1/(2*pi)
       W = We + (H1/pi)*(1+log(2*H/H1));
   else  W = We + (H1/pi)*(1+log(4*pi*We/H1));
   end
   if (W/H) <= 1
      F = (1+12*H/W)^(-0.5) + 0.04*(1-W/H)^2;
   else  F = (1+12*H/W)^(-0.5);
   end
   E_re = 0.5*(Er+1) + 0.5*(Er -1)*F;   % Effective permittivity for static analysis
   E_re_t = E_re - ((E_re-1)/4.6)*(t/H)/sqrt(W/H);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Effective permittivity for static analysis
    if (W/H) <= 1
       E_re_Sio2 = ((Esio2+1)/2)+((Esio2-1)/2)*(1/sqrt(1+(10/U1))+0.04*(1-U1)^2);
    else E_re_Sio2 = ((Esio2+1)/2)+((Esio2-1)/2)*(1/sqrt(1+(10/U1)));
    end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Effetctive impedance, inductance and capacitance
   if (W/H) <= 1
```

```
        Z_om = (N/(2*pi*sqrt(E_re)))*log(8*H/W + 0.25*W/H);
    else
        Z_om = 1/((N/(2*pi*sqrt(E_re)))*(W/H) +1.393 + 0.667*log(W/H + 1.444));
    end
    C_s = (sqrt(E_re)/(c*Z_om))*1e12;   % pF/m
    L_s = (C_s*Z_om^2)/(1e6);   % uH/m
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%Effetctive impedance, inductance and capacitance with thickness effects
    if (W/H) <= 0.5*pi
        W_H = (W/H + (1.25/pi)*(t/H)*(1+log(4*pi*W/t)));
    else W_H = (W/H + (1.25/pi)*(t/H)*(1+log(2*H/t)));
    end
    if (W/H) <= 1
        Z_ct = (N/(2*pi*sqrt(E_re_t)))*log(8/W_H + 0.25*W_H);
    else Z_ct = 1/((N/(2*pi*sqrt(E_re_t)))*(W_H) +1.393 + 0.667*log(W_H + 1.444));
    end
    C_ct = (sqrt(E_re_t)/(c*Z_ct))*1e12;   % pF/m
    L_ct = (C_ct*Z_ct^2)/(1e6);   % uH/m
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Calculated Vp, Vg, Tp, Tg and Td   with effective thickness
    Vp_ct = c/sqrt(E_re_t);    % propagation velocity of single microstrip
    beta_ct = sqrt(E_re_t);   % propagation constant  of single microstrip
    Vg_ct = c*beta_ct;          % group velocity of single microstrip
    rate_ct = Vg_ct/c;
    Tg_ct = (1/Vg_ct)*1e9;      % group delay  of single microstrip
    Tp_ct = (1/Vp_ct)*1e9;      % propagation delay  of single microstrip
    Delta_s = sqrt(1/(pi*1e9*mu0*sigma))
    R_s = sqrt((2*pi*1e9*mu0)/(2*sigma))
    alpha_c = (0.686*R_s)/(Z_ct*W)
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Calculated Vp, Vg, Tp, Tg and Td
    Vp_s = c/sqrt(E_re);    % propagation velocity of single microstrip
    beta_s = sqrt(E_re);   % propagation constant  of single microstrip
    Vg_s = c*beta_s;          % group velocity of single microstrip
    rate_s = Vg_s/c;
    Tg_s = (1/Vg_s)*1e9;      % group delay  of single microstrip
    Tp_s = (1/Vp_s)*1e9;      % propagation delay  of single microstrip
    Td_s = 1.017*sqrt(0.475*E_re+0.67)/0.3048;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% For the coupled micro strip
A = exp(-0.1*exp(2.33-1.5*U)); Cp = (E0*Er*W)/H;  Cf = 0.5*(sqrt(E_re)/(c*Z_om) -
Cp);
%Cf1 = (Cf*sqrt(sqrt(Er/Ere_0)))/(1+(A/g)*tanh(10*g));
Cf1 = Cf/(1+(A/U)*tanh(10*U)); Ce = Cp + Cf + Cf1;
Cgd = (E0*Er/pi)*log(coth(0.25*pi*U)) + 0.65*Cf*(0.02*sqrt(Er)/U + 1 - (1/Er^2));
k = U/(U+2*U);
```

```
k1 = sqrt(1-k^2);
if (k^2 >= 0 && k^2 < (1/sqrt(2)))
   Kfactor_k = (1/pi)*log(2*(1+sqrt(k1))/(1-sqrt(k1)));
else Kfactor_k = pi/(log(2*(1+sqrt(k))/(1-sqrt(k))));
end
Cga = E0*Kfactor_k; Co = Cp + Cf + Cgd + Cga;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% For the air
Cp_a = E0*U; Cf_a = 0.5*(1/(c*Z_om) -Cp); Cf1_a = Cf/(1+(A/U)*tanh(10*U));
Ce_a = Cp_a + Cf_a+ Cf1_a; Cgd_a = (E0/pi)*log(coth(0.25*pi*U)) + 0.65*Cf*(0.02/U);
k2 = U/(U+2*U); k3 = sqrt(1-k2^2);
if (k2^2 >= 0 && k2^2 < (1/sqrt(2)))
   Kfactor_k2 = (1/pi)*log(2*(1+sqrt(k3))/(1-sqrt(k3)));
else Kfactor_k2 = pi/(log(2*(1+sqrt(k2))/(1-sqrt(k2))));
end
Cga_a = E0*Kfactor_k2; Co_a = Cp_a + Cf_a + Cgd_a + Cga_a;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
Z_ce = 1/(c*sqrt(Ce_a*Ce)); Z_co = 1/(c*sqrt(Co_a*Co));
Ee_re = Ce/Ce_a; Eo_re = Co/Co_a;
Ce_cs = (sqrt(Ee_re)/(c*Z_ce))*1e12;   % pF/m
Le_cs = (Ce_cs*Z_ce^2)/(1e6);   % uH/m
Co_cs = (sqrt(Eo_re)/(c*Z_co))*1e12;   % pF/m
Lo_cs = (Co_cs*Z_co^2)/(1e6);   % uH/m
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Calculated Vp, Vg, Tp, Tg and Td   for the coupled microstrip
   Vp_e_cs = c/sqrt(Ee_re);    % propagation velocity of single microstrip
   beta_e_cs = sqrt(Ee_re);   % propagation constant  of single microstrip
   Vg_e_cs = c*beta_e_cs;         % group velocity of single microstrip
   rate_e_cs = Vg_e_cs/c;
   Tg_e_cs = (1/Vg_e_cs)*1e9;      % group delay  of single microstrip
   Tp_e_cs = (1/Vp_e_cs)*1e9;      % propagation delay  of single microstrip
   Td_e_cs = 1.017*sqrt(0.475*Ee_re+0.67)/0.3048;
   Vp_o_cs = c/sqrt(Eo_re);    % propagation velocity of single microstrip
   beta_o_cs = sqrt(Eo_re);   % propagation constant  of single microstrip
   Vg_o_cs = c*beta_o_cs;         % group velocity of single microstrip
   rate_o_cs = Vg_o_cs/c;
   Tg_o_cs = (1/Vg_o_cs)*1e9;      % group delay  of single microstrip
   Tp_o_cs = (1/Vp_o_cs)*1e9;      % propagation delay  of single microstrip
   Td_o_cs = 1.017*sqrt(0.475*Eo_re+0.67)/0.3048;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%MIS slowing factor calculation
if (W/H2) <= 1
      E_re_mis = ((Esio2+Er)/2)+((Er-Esio2)/2)*(1/sqrt(1+(10/U2))+0.04*(1-U2)^2);
    else E_re_mis = ((Esio2+Er)/2)+((Er-Esio2)/2)*(1/sqrt(1+(10/U2)));
end
if (W/H2) <= 1
```

```
   Z_mis = (N/(2*pi*sqrt(E_re_mis)))*log(8*H2/W + 0.25*W/H2);
else Z_mis = 1/((N/(2*pi*sqrt(E_re_mis)))*(W/H2) +1.393 + 0.667*log(W/H2 + 1.444));
end
C_mis = (sqrt(E_re_mis)/(c*Z_mis))*1e12; % pF/m
L_mis = (C_s*Z_mis^2)/(1e6);   % uH/m
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%Calculated Vp, Vg, Tp, Tg and Td
   Vp_mis = c/sqrt(E_re_mis);    % propagation velocity of single microstrip
   beta_mis = sqrt(E_re_mis);   % propagation constant  of single microstrip
   Vg_mis = c*beta_mis;          % group velocity of single microstrip
   rate_mis = Vg_mis/c;
   Tg_mis = (1/Vg_mis)*1e9;      % group delay  of single microstrip
   Tp_mis = (1/Vp_mis)*1e9;      % propagation delay  of single microstrip
   Td_mis = 1.017*sqrt(0.475*E_re_mis+0.67)/0.3048;
```

Appendix B

The process steps of pre-furnace clean

a. Initial setup: First, fill the organic and ionic quartz tanks with DI water to one inch above the nitrogen bubbler tube (5L). Next, turn on the temperature controllers and set them to 95 °C and turn on the OT reset and heater for 30 min to warm up. Second, load wafers into the wafer carriers.

b. Organic clean: When the tank temperature reaches 92 °C in the RCA Organic tank, fill the beaker labeled Organic with Hydrogen Peroxide to the lowest black line (1L) on the label and pour into the tank. Next, fill the beaker labeled Organic with Ammonium Hydroxide to the lowest black line (1L) on the label and pour into the tank. Then, reset the temperature controller to 85 °C. Remove the handle and place it in and rinse it in the quench tank for 1 min. Leave wafers in this solution for 10 to 20 min. Remove the wafers, and place them in quench tank for two minutes.

c. HF Dip: Remove the wafers from the quench and place them in 10:1 HF tank (the 100:1 tank is for maintaining critical oxide thickness). After 30 sec, remove from the HF and place in the quench for 2 min. Have the handle facing downstream. Finally, for the Ionic clean, when the tank temperature reaches 92 °C in the RCA Ionic tank, fill the beaker labeled Organic with Hydrogen Peroxide to the lowest black line (850mL) on the label and pour into the tank. Next, fill the beaker labeled Organic with Hydrochloric Acid to the lowest black line (850 mL) on the label and pour into the tank. Reset the temperature controller to 85 °C. Lower the wafer carrier into the solution, using the top load wafer handle. Remove the handle, and place it in and rinse it in the quench tank for one minute. Leave the wafers in this solution for 10 to 20 min. Remove the wafers and place them in the quench tank for 5 min. Turn off both temperature controllers when finished cleaning. Start the cascade and N_2 bubbler 1 min prior to completing this step.

Appendix C

Matlab Code to Determine Initial Electron Direction in Silicon

```
%*********************************************************************
% Electron Tracking
% etrack04.m
% - determine the image enhancement that results from tracking the recoil
% electron, using GEANT4 data.
% - difference b.t. etrack04 and etrack02 is ability to adjust the
% position resolution during the data reading.  The underlying data
% is based on a resolution of 0.1 um
%*********************************************************************
clear all;
figure(1); clf(1); figure(2); clf(2); figure(3); clf(3);
figure(4); clf(4); figure(5); clf(5); figure(6); clf(6);
figure(7); clf(7); figure(8); clf(8); figure(9); clf(9);
figure(10); clf(10);
% read in the track info from the unedited file
clear file;
file = 'c:\matlabfiles\geant\700keV_100nm_1000.txt'; % 1MeV_1000b.txt'; %
fid = fopen(file);
clear file2;
file2 = 'c:\matlabfiles\geant\VAR_RES_out_700keV_1um_1000.txt'; %
1MeV_1000b.txt'; %
file3 = 'c:\matlabfiles\geant\VAR_RES_out_700keV_10um_1000.txt'; %
1MeV_1000b.txt'; %
file4 = 'c:\matlabfiles\geant\VAR_RES_out_700keV_100um_1000.txt'; %
1MeV_1000b.txt'; %
resultid1 = fopen(file2, 'w');
resultid10 = fopen(file3, 'w');
resultid100 = fopen(file4, 'w');
position_resolution1 = 1e-3;    % in mm
position_resolution10 = 1e-2;    % in mm
position_resolution100 = 1e-1;    % in mm
fprintf(resultid1,'Theta Data with position resolution %f um\n',
position_resolution1*1000);
fprintf(resultid1,'1 um ..... 10 um ..... 100 um info\n');
fprintf(resultid1,'theta (deg)   numpts(best)  fit_track_length (um)\n\n');
fprintf(resultid10,'Theta Data with position resolution %f um\n',
position_resolution10*1000);
fprintf(resultid10,'1 um ..... 10 um ..... 100 um info\n');
fprintf(resultid10,'theta (deg)   numpts(best)  fit_track_length (um)\n\n');
```

```
fprintf(resultid100,'Theta Data with position resolution %f um\n',
position_resolution100*1000);
fprintf(resultid100,'1 um..... 10 um ..... 100 um info\n');
fprintf(resultid100,'theta (deg)   numpts(best)  fit_track_length (um)\n\n');
stopstring = 'Parent';
particle1 = 1; particle10 = 1; particle100 = 1; theta_ave1 = 0; theta_ave10 = 0;
theta_ave100 = 0; track_length_ave1 = 0; track_length_ave10 = 0;
track_length_ave100 = 0;
while (~feof(fid))
  clear a; clear b; clear c; clear i;
  a = textscan(fid,'%s',1);
  if strcmp(a{1,1},stopstring)
   % see if particle track is base track (Parent ID = 0)
   b = textscan(fid,'%s %s %d',1);   % should be ID
   ID = b{1,3};
   if ID == 0
      % read in next three lines
      c = textscan(fid,'%[^P]');
      c = textscan(fid, '%s', 1);
      % should be at beginning of data line
      energy = 1000;     % some number, in MeV, greater than 0
       i = 0;            % counter which notes each track read
       i1 = 1; i10 = 1; i100 = 1;
       um_track1 = 0;     % 1um_track counter read data when = 0
       um_track10 = 0;    % 10um_track counter read data when = 0
       um_track100 = 0;   % 100um_track counter read data when = 0
       depth = -5e-5;     % in mm
       starting_depth = -5e-5;
       % end_depth = 1;
   while (energy > 0 & depth >= starting_depth) % & i < 1000) % & depth < end_depth)
    clear d;
    d = textscan(fid,'%d %f %f %f',1);
    if i > 0 % don't do anything for the first line (init line)
     clear depth;
     depth = d{1,4};
     if (depth >= starting_depth)
%         step(i,1) = d{1,1};
%         x(i,1) = d{1,2};
%         y(i,1) = d{1,3};
%         z(i,1) = d{1,4};
        if (um_track1 == 0)
         step1(i1,1) = d{1,1};  x1(i1,1) = d{1,2};   y1(i1,1) = d{1,3};
         z1(i1,1) = d{1,4};   i1 = i1 + 1;
        end % if 1um  read in the data at 1 um increments
        if (um_track10 == 0)
         step10(i10,1) = d{1,1};
```

```
      x10(i10,1) = d{1,2};  y10(i10,1) = d{1,3};  z10(i10,1) = d{1,4};
      i10 = i10 + 1;
     end % if 10um  read in the data at 1 um increments
     if (um_track100 == 0)
      step100(i100,1) = d{1,1};  x100(i100,1) = d{1,2};  y100(i100,1) = d{1,3};
      z100(i100,1) = d{1,4};  i100 = i100 + 1;
     end % if 100um  read in the data at 1 um increments
   end %if depth
  %    figure(1);
  um_track1 = um_track1 + 1;  um_track10 = um_track10 + 1;
  um_track100 = um_track100 + 1;
  if (um_track1 == 10) um_track1 = 0; end % read next data point at right pt.
  if (um_track10 == 100) um_track10 = 0; end
  if (um_track100 == 1000) um_track100 = 0; end
 end %if i > 0 so read in line
 e = textscan(fid,'%f %f %f %f %s %s',1);  energy = e{1,1};  i = i + 1;
end % while  (data for one track read into arrays)
num_points1 = length(step1);  num_points10 = length(step10);
num_points100 = length(step100);
% plot each track
figure(1);plot3(x1,y1,z1); grid on; hold on;
figure(4);plot(x1,z1); xlabel('x(mm)'); ylabel('z(mm)'); hold on;
figure(5); plot(y1,z1); hold on; xlabel('y(mm)'); ylabel('z(mm)');
% title('Electron Tracks'):
%*****************************************************************
% For each (x,y,z) track data, fit the curve with various
% orders of polynomials (1-3) in both (z,x) and (z,y) planes,
% find the best dual fit (this isn't necessarily the
% best method) and the number of data point used.  Then, find
% the scatter angle (theta) and azimuthal angle (phi) from the best
% fit and compare the scatter angle deviation from known input
% direcxtion (0 degrees).  Evaluate for different energies.
%*****************************************************************
best_fit1 = 1e20; % parameter with which to pick the best pts_fit
                  % points to make line fit
best_fit10 = 1e20; % parameter with which to pick the best pts_fit
                   % points to make line fit
best_fit100 = 1e20; % parameter with which to pick the best pts_fit
                    % points to make line fit
clear best_points1;  % number of point comprising the best fit
clear best_points10;  % number of point comprising the best fit
clear best_points100;  % number of point comprising the best fit
clear no_fit1; clear no_fit10; clear no_fit100;
fit1 = 0; fit10 = 0; fit100 = 0;   % set Boolean to true
% clear best_x;     % track step with the best x,z line fit
% clear best_y;    % track step with the best y,z line fit
```

```
for poly_order = 1:3
 if (num_points1 >= 5)
  fit1 = 1;
  for pts_fit = poly_order+2:num_points1
   % num_points or 20 for longer lines
   % fit line with 3 or more points (e.g.)
   % if N <= pts_fit   % just examine forward fit
    clear px; clear py; clear Sx; clear Sy;
    [px,Sx] = polyfit(z1(1:pts_fit), x1(1:pts_fit),poly_order); % line fit
    [py,Sy] = polyfit(z1(1:pts_fit), y1(1:pts_fit),poly_order); % line fit
    if abs(((Sx.normr)^2 + (Sy.normr)^2)^0.5) < best_fit1
     best_fit1 = abs(((Sx.normr)^2 + (Sy.normr)^2)^0.5);
     best_points1 = pts_fit;
     slope_x1 = px(poly_order); slope_y1 = py(poly_order);
     fit_x1 = polyval(px,z1); fit_y1 = polyval(py,z1);
    end    % if
  end % for pts_fit (cycling through all possible numbers)
 end % if num_points1 >= 5
if (num_points10 >= 5)
  fit10 = 1;
  for pts_fit = poly_order+2:num_points10
   % num_points or 20 for longer lines
   % fit line with 3 or more points (e.g.)
   % if N <= pts_fit   % just examine forward fit
    clear px; clear py; clear Sx; clear Sy;
    [px,Sx] = polyfit(z10(1:pts_fit), x10(1:pts_fit),poly_order); % line fit
    [py,Sy] = polyfit(z10(1:pts_fit), y10(1:pts_fit),poly_order); % line fit
    if abs(((Sx.normr)^2 + (Sy.normr)^2)^0.5) < best_fit10
     best_fit10 = abs(((Sx.normr)^2 + (Sy.normr)^2)^0.5);
     best_points10 = pts_fit;
     slope_x10 = px(poly_order); slope_y10 = py(poly_order);
     fit_x10 = polyval(px,z10); fit_y10 = polyval(py,z10);
    end    % if
   end % for pts_fit (cycling through all possible numbers)
 end % if num_points10 >= 5
 if (num_points100 >= 5)
  fit100 = 1;
  for pts_fit = poly_order+2:num_points100
   % num_points or 20 for longer lines
   % fit line with 3 or more points (e.g.)
   % if N <= pts_fit   % just examine forward fit
    clear px; clear py; clear Sx; clear Sy;
    [px,Sx] = polyfit(z100(1:pts_fit), x100(1:pts_fit),poly_order); % line fit
    [py,Sy] = polyfit(z100(1:pts_fit), y100(1:pts_fit),poly_order); % line fit
    if abs(((Sx.normr)^2 + (Sy.normr)^2)^0.5) < best_fit100
     best_fit100 = abs(((Sx.normr)^2 + (Sy.normr)^2)^0.5);
```

```
    best_points100 = pts_fit;
    slope_x100 = px(poly_order); slope_y100 = py(poly_order);
    fit_x100 = polyval(px,z100);fit_y100 = polyval(py,z100);
   end    % if
  end % for pts_fit (cycling through all possible numbers)
 end % for poly_order
end % if num_points100 >= 5
% best fit determined
% find the track length used in the line fitting
fit_tracklength1 = 0;
if fit1
  for j = 2:best_points1
   fit_tracklength1 = fit_tracklength1 + sqrt((x1(j,1)- x1(j-1,1))^2…
                          + (y1(j,1)- y1(j-1,1))^2 +(z1(j,1)- z1(j-1,1))^2);
  end % fit_tracklength calculated
  num_pts_in_fit1(particle1) = best_points1;
  phi1 = atan(slope_x1/slope_y1);  % in radians
  theta_rad1 = atan(1/(sin(phi1)*slope_y1));
  theta_deg1 = abs(180*theta_rad1/pi);
  theta_ave1 = theta_ave1 + 90 - theta_deg1;
 % sum up the thetas then divide by particle number at end
  track_length_ave1 = track_length_ave1 + fit_tracklength1;
  % again, sum then divide later
  theta1(particle1) = 90 - theta_deg1;
  fprintf(resultid1,'%f  %f  %f\n',90 - theta_deg1, best_points1, fit_tracklength1*1000);
  particle1  = particle1 + 1;
end % it fit
fit_tracklength10 = 0;
if fit10
 for j = 2:best_points10
  fit_tracklength10 = fit_tracklength10 + sqrt((x10(j,1)- x10(j-1,1))^2…
                          + (y10(j,1)- y10(j-1,1))^2 +(z10(j,1)- z10(j-1,1))^2);
 end % fit_tracklength calculated
 num_pts_in_fit10(particle10) = best_points10;
 phi10 = atan(slope_x10/slope_y10);  % in radians
 theta_rad10 = atan(1/(sin(phi10)*slope_y10));
 theta_deg10 = abs(180*theta_rad10/pi);
 theta_ave10 = theta_ave10 + 90 - theta_deg10;
 % sum up the thetas then divide by particle number at end
 track_length_ave10 = track_length_ave10 + fit_tracklength10;
 % again, sum then divide later
 theta10(particle10) = 90 - theta_deg10;
 fprintf(resultid10,'%f  %f  %f\n',90 - theta_deg10, best_points10,…
          fit_tracklength10*1000);
 particle10 = particle10 + 1;
end % 10 um fit
```

```
    fit_tracklength100 = 0;
    if fit100
     for j = 2:best_points100
      fit_tracklength100 = fit_tracklength100 + sqrt((x100(j,1)- x100(j-1,1))^2…
                          + (y100(j,1)- y100(j-1,1))^2 +(z100(j,1)- z100(j-1,1))^2);
     end % fit_tracklength calculated
     num_pts_in_fit100(particle100) = best_points100;
     phi100 = atan(slope_x100/slope_y100);  % in radians
     theta_rad100 = atan(1/(sin(phi100)*slope_y100));
     theta_deg100 = abs(180*theta_rad100/pi);
     theta_ave100 = theta_ave100 + 90 - theta_deg100;
     % sum up the thetas then divide by particle number at end
     track_length_ave100 = track_length_ave100 + fit_tracklength100;
     % again, sum then divide later
     theta100(particle100) = 90 - theta_deg100;
     fprintf(resultid100,'%f  %f  %f\n',90 - theta_deg100, best_points100,…
     fit_tracklength100*1000);
     particle100 = particle100 + 1;
    end % 100 um fit
    figure(7);hist(theta1, 100); xlabel('theta (deg)'); ylabel('counts/bin');
    figure(8);hist(num_pts_in_fit1, 20);xlabel('number of points in fit'); ylabel('counts/bin');
    figure(9);hist(num_pts_in_fit10, 20);xlabel('number of points in fit');
                ylabel('counts/bin');
    figure(10);hist(num_pts_in_fit100, 20);xlabel('number of points in fit');
                 ylabel('counts/bin');
    figure(2);plot(x1,z1,'o',fit_x1,z1,'-');xlabel('x(mm)');ylabel('z(mm)');%hold on;
    figure(3);plot(y1,z1,'o',fit_y1,z1,'-');xlabel('y(mm)');ylabel('z(mm)');%hold on;
               title('Goodness of Fit');
    figure(6);plot3(x1,y1,z1,fit_x1,fit_y1,z1);grid on;hold on;
   clear x1; clear y1; clear z1; clear step1; clear fit_x1; clear fit_y1;
   clear slope_x1; clear slope_y1; clear best_fit1; clear best_points1;
   clear theta_rad1;
   clear x10; clear y10; clear z10; clear step10; clear fit_x10; clear fit_y10;
   clear slope_x10; clear slope_y10; clear best_fit10; clear best_points10;
   clear theta_rad10;
   clear x100; clear y100; clear z100; clear step100; clear fit_x100; clear fit_y100;
   clear slope_x100; clear slope_y100; clear best_fit100; clear best_points100;
   clear theta_rad100;
  end % found ID 0
 end  % found Parent ID
end  % reached end of file
theta_ave1 = theta_ave1/particle1;
track_length_ave1 = track_length_ave1/particle1;
theta_ave10 = theta_ave10/particle10;
track_length_ave10 = track_length_ave10/particle10;
theta_ave100 = theta_ave100/particle100;
```

```
track_length_ave100 = track_length_ave100/particle100;
fprintf(resultid1, '1 um theta_ave = %f deg based on %f particles\n', theta_ave1,
particle1-1);
fprintf(resultid10, '10 um theta_ave = %f deg based on %f particles\n', theta_ave10,
particle10-1);
fprintf(resultid100, '100 um theta_ave = %f deg based on %f particles\n', theta_ave100,
particle100-1);
fprintf(resultid1, 'average track length used in 1 um fit = %f um\n',
track_length_ave1*1000);
fprintf(resultid10, 'average track length used in 10 um fit = %f um\n',
track_length_ave10*1000);
fprintf(resultid100, 'average track length used in 100 um fit = %f um\n',
track_length_ave100*1000); % multiply to convert mm to um
fclose(fid); fclose(resultid1); fclose(resultid10); fclose(resultid100); theta1
```

Appendix D

Image Reconstruction Algorithms for Compton Scattering Imaging

In this section, we present the Matlab code, implementing four different types of image reconstruction algorithms for Compton scattering imaging; SBP versus FBP and with electron-tracking ability versus without electron-tracking ability. The code enables users to select options for different image reconstruction algorithms, a spatial resolution of radiation detector for electron-tracking, the resolution of the rectangular grid, and the number of events considered for image reconstruction.

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
clear all;
CIRsetup; %set path
 %r1: first interaction position/E1: energy of first Compton%%%%%%%%%%%%%%
% scattered photon
%Data loading % r2: second interaction position/E2: energy of second Compton scattered
photon
filename = sprintf('ComptonPairData_1000cmSource_Dir001.%d.txt',0);
[x1 y1 z1 E1 x2 y2 z2 E2] = textread(filename, '%f %f %f %f %f %f %f %f', -1,
'delimiter', ' ');
%Basic quantities
E0 = 0.663;                               %source energy [MeV]
c = 299792458;                             %speed of light [m/sec]
me = 9.10938188*10^(-31);                   %mass of an electron
MeVperJoule = 1/(1.60218*10^(-19))/(10^6);  %conversion factor
Ee = me*c^2*MeVperJoule;                    %[MeV]
rot90mat = [1 0 0;0 0 1;0 -1 0];           %90 degree rotation about the x-axis (clockwise)
%Choosing the data size
while 1
   msg = sprintf('Input # of events to be tested among %d events:',length(x1));
   npt = input(msg); %npt: # of events to be considered
   if (npt > 0 && npt <= length(x1))
      break;
   end
   msg = sprintf('Wrong selection! Select # between 0 and %d',length(x1));
   disp(msg);
```

```
end
r1 = [x1(1:npt) y1(1:npt) z1(1:npt) E1(1:npt)];
r2 = [x2(1:npt) y2(1:npt) z2(1:npt) E2(1:npt)];
%energy   check%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
ind1 = find(E1 >= E0);      ind2 = find(E2 >= E0);
if (~isempty(ind1) || ~isempty(ind2)),  disp 'Illegal energy assigned';  end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%Choosing reconstruction parameters %
%90 deg rotation for Randy's MCNP data
while 1
   msgrot = sprintf('Do you want to rotate the data by 90 deg? (Y=1/N=0):');
   rot90sel = input(msgrot);
   if (rot90sel == 0 || rot90sel == 1)
      break;
   end
   disp('Wrong selection! Select 0 or 1');
end
%Simple BP or Filter BP%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
while 1
   msgBP = sprintf('Do you want Simple BP or Parra Filtered BP? (FBP=1/SBP=0):');
   BPsel = input(msgBP);
   if (BPsel == 0 || BPsel == 1)
      break;
   end
   disp('Wrong Selection! Select 0 or 1');
end
%Electron   tracking%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
while 1
   msgetrack = sprintf('Do you want to use electron tracking? (Y=1/N=0):');
   etracksel = input(msgetrack);
   if (etracksel == 0 || etracksel == 1)
      break;
   end
   disp('Wrong selection! Select 0 or 1');
end
%Detector  spatial  resolution%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
if etracksel == 1;
   while 1
      msgDspatresol = sprintf('Select detector spatial resolution (0.1um/1um/10um):');
      Dspatresol_str = input(msgDspatresol,'s');
      switch Dspatresol_str
         case '0.1'
            Dspatresol = 1;break;
         case '1'
            Dspatresol = 2;break;
         case '10'
```

```
            Dspatresol = 3;break;
        end
        disp('Wrong selection! Select among 0.1um/1um/10um');
    end
end
%Degree resolution on a 4pi sphere%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
while 1
    msgdegresol = sprintf('Select degree resolution on a 4pi sphere (0.25 deg/0.5 deg/…
                                1 deg/2 deg):');
    degresolsel = input(msgdegresol);
    switch degresolsel
        case 0.25
            resolR = 700;indstep = 25;break;
        case 0.5
            resolR = 324;indstep = 100;break;
        case 1
            resolR = 214;indstep = 100;break;
        case 2
            resolR = 104;indstep = 100;break;
    end
    disp('Wrong selection! Select 0.25 deg/0.5 deg/1 deg/2 deg');
end
%dE1: deposited energy by 1st interaction (by electron) %%%%%%%%%%%%%%%%%%
% Line projection % %dE2: deposited energy by 2nd interaction%%%%%%%%%%%%%%
dE1 = E0 - r1(:,4);
dE2 = r1(:,4) - r2(:,4);
%Scattering angle
KNfactor = 1-Ee*(1./r1(:,4) - 1./E0);
% - correction
indPErr = find(KNfactor > 1);       indNErr = find(KNfactor < -1);
if ~isempty(indPErr),   KNfactor(indPErr) = 1;  end
if ~isempty(indNErr),   KNfactor(indNErr) = -1; end
omeg = acos(KNfactor); %[0, pi]
%Line projection direction
rdiff = r1(:,(1:3))-r2(:,(1:3));
rdiffunit = rdiff./repmat(sqrt(dot(rdiff,rdiff,2)), [1 3]);
%90 deg rotation for Randy's MCNP data
if rot90sel == 1,   rdiffunit = transpose(rot90mat*transpose(rdiffunit));   end
%Spherical coordinate representation
cosphi = rdiffunit(:,1)./sqrt(1-rdiffunit(:,3).^2);
sinphi = rdiffunit(:,2)./sqrt(1-rdiffunit(:,3).^2);
OM2 = [acos(cosphi) acos(rdiffunit(:,3)) omeg];
indNsinphi = find(sinphi < 0);
OM2(indNsinphi,1) = 2*pi-acos(cosphi(indNsinphi));
%save('OM2.mat','OM2');
%hist(OM2(:,1),100);grid on;    %phi
```

```
%hist(OM2(:,2),100);grid on;    %theta
%hist(OM2(:,3),100);grid on;
%sanity check on a 4pi unit sphere
if 0
   [sx,sy,sz] = sphere(20);
   surf(sx,sy,sz,ones(size(sx)));axis equal;
   hold on;plot3(rdiffunit(:,1), rdiffunit(:,2), rdiffunit(:,3),'r.');grid on;
   hold off;title('Mapping on the Unit Sphere');
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Axes for cone generation %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
k0 = [0 0 1]; %north pole
if rot90sel == 1,   k0 = (rot90mat*k0')';   end
xp = repmat(k0,[npt,1]) - repmat(cos(omeg),[1 3]).*rdiffunit;
xpunit = xp./repmat(sqrt(dot(xp,xp,2)),[1 3]);
zpunit = rdiffunit;
yp = cross(zpunit,xpunit,2);
ypunit = yp./repmat(sqrt(dot(yp,yp,2)),[1 3]);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%Electron tracking angular uncertainty specification %
%formula from 2008 IEEE NSS-MIC paper %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
if etracksel == 1
   switch Dspatresol
      case 1 %0.1um
         AngUn = 24.918*(1000*dE1).^(-1.0639);  %in deg
      case 2 %1um
         AngUn = 289.1*(1000*dE1).^(-1.0364);
      case 3 %10um
         AngUn = 6372.3*(1000*dE1).^(-1.2761);
   end
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Rect grid generation % [#theta, #phi] in deg
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
phiR = linspace(0,360,2*resolR);
thetaR = linspace(0.0001,180,resolR);
[phiRR thetaRR] = meshgrid(phiR,thetaR);
del_phiR = phiR(2)-phiR(1);        del_thetaR = thetaR(2)-thetaR(1);
%Coordinates on a 4pi sphere
runitR = [cosd(phiRR(:)).*sind(thetaRR(:)) sind(phiRR(:)).*sind(thetaRR(:))…
            cosd(thetaRR(:))];
nptR = size(runitR,1);
%for tallying
scoreR= zeros(size(phiRR));        AscoreR = zeros(size(phiRR));
```

```
%Plot
if 0
   [sx,sy,sz] = sphere(20);
   surf(sx,sy,sz,ones(size(sx)));axis equal;
   hold on;plot3(runitR(:,1), runitR(:,2), runitR(:,3),'r.');grid on;
   hold off;title('Mapping on the Unit Sphere');
end
%Interpolated rect grid (Only for displaying purpose)
Ifactor = 2; %1 w/o interpolation
phiRI = linspace(0,360,Ifactor*2*resolR);
thetaRI = linspace(0.0001,180,Ifactor*resolR);
[phiRRI thetaRRI] = meshgrid(phiRI,thetaRI);
%Displaying setup
ndisplayevent = 5;
zcrossing1 = 0.1;
zcrossing2 = 0.2;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Source Image Reconstructions %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% 1. Simple back projection algorithm %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
if BPsel == 0
   %%%%%%%%%%%%%%%%%%%%
   % Cone generation % - Populating circle
   %%%%%%%%%%%%%%%%%%%%
   phi_del = degresolsel; %in deg
   phi_N = 360/phi_del;
   phi = (0:phi_N -1)'*phi_del*2*pi/360;
   %%%%%%%%%%%%%%%%%%%%%%%%%%%%
   % Image reconstruction algorithm  % Event-by-event
   %%%%%%%%%%%%%%%%%%%%%%%%%%%%
   tic
   for ii=1:npt
      %Cone data points
      rbp = repmat(cos(omeg(ii))*rdiffunit(ii,:), [phi_N 1]) ...
         + sin(omeg(ii))*(cos(phi)*xpunit(ii,:) + sin(phi)*ypunit(ii,:));
      ptXcart = cos(omeg(ii))*rdiffunit(ii,:) + sin(omeg(ii))*xpunit(ii,:);
      %Polar coordinate [theta,phi]
      thetaX = acosd(ptXcart(3));            phiX = acosd(ptXcart(1)/sqrt(1-ptXcart(3)^2));
      if (ptXcart(2) < 0),  phiX = 360 - phiX;    end

      %%%%%%%%%%%%%%%%%%%%%%%%
      % Conventional compton imaging % in Rect grid
      %%%%%%%%%%%%%%%%%%%%%%%%
      if etracksel == 0
         %Direct Rect grid tallying
```

```
        maxval = [];maxind = [];
        indxup  = ceil(size(rbp,1)/indstep);
        for st=1:indxup
           if st ~= indxup
              Crindwind = [indstep*(st-1)+1:indstep*st];
              rbp_runitR = rbp(Crindwind,:)*transpose(runitR);
           elseif st == indxup
              rbp_runitR = rbp((indstep*(st-1):end),:)*transpose(runitR);
           end
           [maxvaltmp maxindtmp] = max(rbp_runitR');
           maxval = [maxval maxvaltmp];
           maxind = [maxind maxindtmp];
        end
        %To consider multiple counts in the smae positions
        for ih=1:length(maxval)
           scoreR(maxind(ih)) = scoreR(maxind(ih)) + 1;
        end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%Advanced compton imaging by electron tracking              %
% - adjusted "arcind" depending on the angular uncertainty %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    elseif etracksel == 1
        %Angular uncertainty
        if (AngUn(ii)) < 360  %/sin(omeg(ii))
           arcind = ceil(AngUn(ii)/phi_del); %/sin(omeg(ii)));
           if rem(arcind,2) == 0,  arcind = arcind + 1;    end   %even # ind -> odd # ind
           Arbp = rbp([(1:(arcind+1)/2) (end-(arcind-1)/2+1:end)],:);
           %event-wise arc segmentation
        elseif (AngUn(ii)) >= 360   %/sin(omeg(ii))  %When angular uncertainty ~2pi
           Arbp = rbp;
        end
        %Direct Rect grid tallying
        Amaxval = [];Amaxind = [];
        indxup = ceil(size(Arbp,1)/indstep);
        for st=1:indxup
           if st ~= indxup
              Crindwind = [indstep*(st-1)+1:indstep*st];
              Arbp_runitR = Arbp(Crindwind,:)*transpose(runitR);
           elseif st == indxup && st ~= 1
              Arbp_runitR = Arbp((indstep*(st-1):end),:)*transpose(runitR);
           elseif st == indxup && st == 1
              Arbp_runitR = Arbp*transpose(runitR);
           end
           [Amaxvaltmp Amaxindtmp] = max(Arbp_runitR');
           Amaxval = [Amaxval Amaxvaltmp];
           Amaxind = [Amaxind Amaxindtmp];
```

```
    end
    %To consider multiple counts in the smae positions
    for ih=1:length(Amaxval)
        AscoreR(Amaxind(ih)) = AscoreR(Amaxind(ih)) + 1;
    end
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Looking for FWHM %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%Normalization
if etracksel == 0
    nscoreR = scoreR/max(scoreR(:));
elseif etracksel == 1
    nscoreR = AscoreR/max(AscoreR(:));
end
if 0
    %Finding zero-crossing pts
    thresh = 0;
    while 1
        zeroind = find(abs(nscoreR-zcrossing1) <= thresh);
        if ~isempty(zeroind)
            break;
        end
        thresh = thresh+0.001;
        if thresh > 1
            error('Threshold for finding zero-crossing pts is greater than 1!');
        end
    end
    distXzero = mean(sqrt((phiRR(zeroind)-phiX).^2…
                    + (thetaRR(zeroind)-thetaX).^2));
    GauCir = ((phiRR-phiX).^2 + (thetaRR-thetaX).^2 <= distXzero^2);
    %circle around pt X
    %Gaussian kernel fitting
    GauKersig = (0.01:0.05:15);
    for gi=1:length(GauKersig)
        GauKernel = exp(-1/(2*GauKersig(gi)^2)*((thetaRR-thetaX).^2…
                        + (phiRR-phiX).^2));
        GauKerdifsum(gi) = sum(sum(((GauKernel - nscoreR).*GauCir).^2));
    end
    [minGauKersig minindGauKersig] = min(GauKerdifsum);
    FWHM(ii) = sqrt(8*log(2))*GauKersig(minindGauKersig);
    FittedGauKernel = exp(-1/(2*GauKersig(minindGauKersig)^2)…
                            *((thetaRR-thetaX).^2 + (phiRR-phiX).^2));
    Fiterr = FittedGauKernel - nscoreR;
    GauCirFitbase = ((phiRR-phiX).^2 + (thetaRR-thetaX).^2…
                        <= (sqrt(2*log(10))*GauKersig(minindGauKersig))^2);
```

```
        rmserr(ii) = sqrt(sum(sum((Fiterr.*~GauCirFitbase).^2)));
    end
    %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    % Distance weighted RMS Error %
    %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    %Finding subscripts for max of nscoreR
    indmx = find(nscoreR == 1);
    if length(indmx) == 1
        [submxtheta submxphi] = ind2sub(size(nscoreR),indmx);
        %Distance weighted RMS error
        [subtheta subphi] = ndgrid((1:size(nscoreR,1)),(1:size(nscoreR,2)));
        theta_dist = abs(subtheta-submxtheta)*del_thetaR;
        phi_dist = abs(subphi-submxphi)*del_phiR;
        dwrmserr(ii) = sqrt(sum(sum(sqrt(theta_dist.^2 + phi_dist.^2).*(nscoreR.^2))));
    end
    %Displaying at every event when ii <= ndisplayevent, at every 100
    %events when ii > ndisplayevent
    if (ii <= ndisplayevent || (ii > ndisplayevent && rem(ii,100) == 0) || ii == npt)
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Displaying reconstructed images %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%Interpolation and plotting
    if etracksel == 0
        scoreRI = interp2(phiRR,thetaRR,scoreR,phiRRI,thetaRRI,'spline');
                figure(1+BPsel);imagesc(thetaRI,phiRI,abs(scoreRI'/max(scoreRI(:))));
                colorbar; grid on;axis xy;axis square; xlabel('Theta (degree)');
                ylabel('Phi (degree)');
                tit = sprintf('Reconstructed Image by Simple BP with the %dth…
                    event',ii);title(tit);
        elseif etracksel == 1
            AscoreRI = interp2(phiRR,thetaRR,AscoreR,phiRRI,thetaRRI,'spline');
            figure(1+BPsel);imagesc(thetaRI,phiRI,abs(AscoreRI'/max(AscoreRI(:))));
            colorbar;grid on;axis xy;axis square;
            xlabel('Theta (degree)');ylabel('Phi (degree)');
            tit = sprintf('Reconstructed Image by Advanced Simple BP with the %dth…
                event',ii);title(tit);
        end
        %subplot(122);imagesc(thetaR,phiR,FittedGauKernel');
        %grid on;colorbar;axis xy;axis square;
        %tit = sprintf('Fitted Gaussian Kernel with FWHM=%d deg',FWHM(ii));
        %title(tit);
        %status reporting
        if etracksel == 0
            msg = sprintf('Conventional Compton: %d events out of %d (deg
resol: %d)',ii,npt,degresolsel);
        elseif etracksel == 1
```

```
            msg = sprintf('Advanced Compton: %d events out of %d (deg
resol: %d)',ii,npt,degresolsel);
          end
            disp(msg);      %pause;
      end
%When non-accumulated images are needed
%       if etracksel == 1
%           AscoreR = zeros(size(AscoreR));
%       elseif etracksel == 0
%           scoreR = zeros(size(scoreR));
%       end
   end %event-by-event loop
   SBackPjtime = toc/60; %in min
   msg = sprintf('Simple Back Projection Time: %f mins',SBackPjtime);
   disp(msg);
   %Mesh plot
   if etracksel == 0
      figure(2+BPsel);mesh(thetaRI,phiRI,abs(scoreRI'/max(scoreRI(:))));colorbar;grid on;
      xlabel('Theta (degree)');ylabel('Phi (degree)');
      tit = sprintf('Reconstructed Image by Simple BP after %d events',ii);title(tit);
   elseif etracksel == 1
figure(2+BPsel);mesh(thetaRI,phiRI,abs(AscoreRI'/max(AscoreRI(:))));colorbar;grid on;
      xlabel('Theta (degree)');ylabel('Phi (degree)');
      tit = sprintf('Reconstructed Image by Advanced Simple BP after %d events',ii);
      title(tit);
   end
   if 0
      %FWHM plot
      figure(3+BPsel);stem(FWHM);grid on;
      xlabel('Number of events');ylabel('FWHM [deg]');
      outind = (abs(FWHM-mean(FWHM)) >= mean(FWHM));
      FWHMplot = FWHM(~logical(outind));
      tit = sprintf('FWHM [deg] vs Number of event with avg=%d',
                  mean(FWHMplot(20:end)));  title(tit);
   end
   %RMS error plot
   figure(4+BPsel);stem(dwrmserr);grid on;
   xlabel('Number of events');ylabel('dwRMS Error [AU]');
   tit = sprintf('RMS Error with avg=%d',mean(dwrmserr(20:end)));  title(tit);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% 2. Filtered back projection algorithm by Parra %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
elseif BPsel == 1
   %Loading Legendre coeffs for Parra's PSF
   load Hn_P.mat;         %upto 400th
   cos_thetaRR = cosd(thetaRR);       sin_thetaRR = sind(thetaRR);
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Image reconstruction algorithm by Parra % Event-by-event
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
PERF = zeros(size(phiRR));         Pnorder = 40; %Legendre order
adist = ceil(600/Pnorder); %in deg
tic
for evi=1:npt  %Event-by-event
 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
 % Creating Angular Uncertainty Mask % in deg
 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
 %Points of a circle (in Cartesian coordinate)
  ptXcart = cos(omeg(evi))*rdiffunit(evi,:) + sin(omeg(evi))*xpunit(evi,:);
  ptOcart = rdiffunit(evi,:);
  %In polar coordinate [theta,phi]
  thetaX = acosd(ptXcart(3));            phiX = acosd(ptXcart(1)/sqrt(1-ptXcart(3)^2));
  if (ptXcart(2) < 0),  phiX = 360 - phiX;    end
  thetaO = acosd(ptOcart(3));            phiO = acosd(ptOcart(1)/sqrt(1-ptOcart(3)^2));
  if (ptOcart(2) < 0),  phiO = 360 - phiO;    end
      if etracksel == 1
      %Angular uncertainty on a circle
      AngUnCir(evi) = AngUn(evi); %/sin(omeg(evi));
       if AngUnCir(evi) < 360     %When angular uncertainty is less than 2pi
                                  %segment a circle
      %Boundary points of a circle (in Cartesian coordinate)
         ptAcart = cos(omeg(evi))*rdiffunit(evi,:) ...
         + sin(omeg(evi))*(cosd(AngUnCir(evi)/2)*xpunit(evi,:) +
         sind(AngUnCir(evi)/2)*ypunit(evi,:));
         ptBcart = cos(omeg(evi))*rdiffunit(evi,:) ...
          + sin(omeg(evi))*(cosd(AngUnCir(evi)/2)*xpunit(evi,:) –
         sind(AngUnCir(evi)/2)*ypunit(evi,:));
      %In Polar coordinate [theta,phi]
         thetaA = acosd(ptAcart(3));   phiA = acosd(ptAcart(1)/sqrt(1-ptAcart(3)^2));
         if (ptAcart(2) < 0),  phiA = 360 - phiA;    end
         thetaB = acosd(ptBcart(3));   phiB = acosd(ptBcart(1)/sqrt(1-ptBcart(3)^2));
         if (ptBcart(2) < 0),  phiB = 360 - phiB;    end
         ptApol(evi,:) = [thetaA phiA];        ptBpol(evi,:) = [thetaB phiB];
         ptXpol(evi,:) = [thetaX phiX];        ptOpol(evi,:) = [thetaO phiO];
      %Angle XOA and XOB
         OA = [thetaA-thetaO phiA-phiO];     OB = [thetaB-thetaO phiB-phiO];
         OX = [thetaX-thetaO phiX-phiO];
         AngXOA = acosd(dot(OA,OX)/norm(OA)/norm(OX));
         AngXOB = acosd(dot(OB,OX)/norm(OB)/norm(OX));
         %Regions by line AO & BO
         deterAO_atX = (phiX-phiO-(phiA-phiO)/(thetaA-thetaO)*(thetaX-thetaO));
         deterBO_atX = (phiX-phiO-(phiB-phiO)/(thetaB-thetaO)*(thetaX-thetaO));
         %By line AO
```

```
  linemargin = 0;
  if (evi == 1 && Dspatresol == 1)
     linemargin = 5;
  end %deg margin correction
  if (deterAO_atX > 0)
     deterAO = (phiRR-phiO - (phiA-phiO)/(thetaA-thetaO)*(thetaRR-thetaO))
      >= -linemargin; %0;
  elseif  (deterAO_atX < 0)
     deterAO = (phiRR-phiO - (phiA-phiO)/(thetaA-thetaO)*(thetaRR-thetaO))
     <= linemargin; %0;
  end
  %By line BO
  if (deterBO_atX > 0)
     deterBO = (phiRR-phiO - (phiB-phiO)/(thetaB-thetaO)*(thetaRR-thetaO))
     >= -linemargin; %0;
  elseif (deterBO_atX < 0)
     deterBO = (phiRR-phiO - (phiB-phiO)/(thetaB-thetaO)*(thetaRR-thetaO))
     <= linemargin; %0;
  end
  %Circle by AOB
  CirAOB = ((phiRR-phiO).^2 + (thetaRR-thetaO).^2 <= (norm(OX)+adist)^2);
  %Angular uncertainty mask
  if (AngXOA+AngXOB <= 180 )   %AngUnCir(evi) < 180
     AngUnmask = double(deterAO & deterBO).*double(CirAOB);
  elseif (AngXOA+AngXOB > 180)
     AngUnmask = double(deterAO | deterBO).*double(CirAOB);
  end
  %Point indices
  [minthetaA indminthetaA] = min(abs(thetaR-thetaA));
  [minthetaB indminthetaB] = min(abs(thetaR-thetaB));
  [minthetaO indminthetaO] = min(abs(thetaR-thetaO));
  [minthetaX indminthetaX] = min(abs(thetaR-thetaX));
  [minphiA indminphiA] = min(abs(phiR-phiA));
  [minphiB indminphiB] = min(abs(phiR-phiB));
  [minphiO indminphiO] = min(abs(phiR-phiO));
   [minphiX indminphiX] = min(abs(phiR-phiX));
  indA = [indminthetaA indminphiA];     indB = [indminthetaB indminphiB];
  indO = [indminthetaO indminphiO];     indX = [indminthetaX indminphiX];
if 0%for sanity check
     %pts A,B,O,X
     AngUnmask(indA(1),indA(2)) = 2;    AngUnmask(indB(1),indB(2)) = 3;
     AngUnmask(indO(1),indO(2)) = 4;    AngUnmask(indX(1),indX(2)) = 5;
  end
elseif AngUnCir(evi) >= 360  %When angular uncertainty is more than 2pi
  %Circle by AOB
  ptOcart = rdiffunit(evi,:);   %center
```

```
        ptXcart = cos(omeg(evi))*rdiffunit(evi,:) + sin(omeg(evi))*xpunit(evi,:);
        thetaO = acosd(ptOcart(3));  phiO = acosd(ptOcart(1)/sqrt(1-ptOcart(3)^2));
        if (ptOcart(2) < 0),  phiO = 360 - phiO;    end
        thetaX = acosd(ptXcart(3));  phiX = acosd(ptXcart(1)/sqrt(1-ptXcart(3)^2));
        if (ptXcart(2) < 0),  phiX = 360 - phiX;    end
        OX = [thetaX-thetaO phiX-phiO];
        CirAOB = ((phiRR-phiO).^2 + (thetaRR-thetaO).^2 <= (norm(OX)+adist)^2);
        %Angular uncertainty mask
        AngUnmask = ones(size(phiRR)).*double(CirAOB);
    end
end    %end of loop for creating angular uncertainty mask
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Generation of Parra's ERF %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
arg1 = cos(OM2(evi,3));
arg2 = cos_thetaRR*cos(OM2(evi,2))+sin_thetaRR*sin(OM2(evi,2)).*…
        cos(phiRR*2*pi/360-OM2(evi,1));
%Legendre functions by Bonnet's recursion
%Initialization
Pn1(1) = 1;      Pn1(2) = arg1;
Pn2(:,:,1) = ones(size(arg2));  Pn2(:,:,2) = arg2;
%Main loop for Bonnet's recursion
for ll=2:Pnorder
  Pn1(ll+1) = ((2*ll+1)*arg1*Pn1(ll) - ll*Pn1(ll-1))/(ll+1);
  Pn2(:,:,ll+1) = ((2*ll+1)*arg2.*Pn2(:,:,ll) - ll*Pn2(:,:,ll-1))/(ll+1);
end
%Parra's ERF
Cn = (2*(0:Pnorder)+1).^2/(8*pi)./Hn_P(1:Pnorder+1);    %[1,Pnorder+1]
PERF_vol = permute(repmat(Cn'.*Pn1(:),[1 size(Pn2,1) size(Pn2,2)]),
                [2 3 1]).*Pn2; %[resolR,2*resolR,Pnorder+1]
if etracksel == 0
   PERF = PERF + sum(PERF_vol,3); %[resolR,2*resolR]
   %PERF = sum(PERF_vol,3);
elseif etracksel == 1
   PERF = PERF + sum(PERF_vol,3).*AngUnmask; %[resolR,2*resolR]
   %PERF = sum(PERF_vol,3).*AngUnmask;
   %PERFtmp = sum(PERF_vol,3); %for sanity check
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Looking for FWHM %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
nPERF = PERF/max(PERF(:));  %normalized Parra's ERF
if 0
   %Finding zero-crossing pts
   thresh = 0;
   while 1
```

```
        zeroind = find(abs(nPERF-zcrossing2) <= thresh);
        if ~isempty(zeroind)
            break;
        end
        thresh = thresh+0.001;
        if thresh > 1
            error('Threshold for finding zero-crossing pts is greater than 1!');
        end
    end
    distXzero = mean(sqrt((phiRR(zeroind)-phiX).^2 …
                    + (thetaRR(zeroind)-thetaX).^2));
    GauCir = ((phiRR-phiX).^2 + (thetaRR-thetaX).^2 <= distXzero^2);
    %circle around pt X
    %Gaussian kernel fitting
    GauKersig = (0.01:0.05:15);
    for gi=1:length(GauKersig)
        GauKernel = exp(-1/(2*GauKersig(gi)^2)*((thetaRR-thetaX).^2 +…
                        (phiRR-phiX).^2));
        GauKerdifsum(gi) = sum(sum(((GauKernel - nPERF).*GauCir).^2));
    end
    [minGauKer_fitdif minindGauKersig] = min(GauKerdifsum);
    FWHM(evi) = sqrt(8*log(2))*GauKersig(minindGauKersig);
    FittedGauKernel = exp(-1/(2*GauKersig(minindGauKersig)^2)…
                            *((thetaRR-thetaX).^2 + (phiRR-phiX).^2));
    Fiterr = FittedGauKernel - nPERF;
    GauCirFitbase = ((phiRR-phiX).^2 + (thetaRR-thetaX).^2 …
                        <= (sqrt(2*log(10))*GauKersig(minindGauKersig))^2);
    rmserr(evi) = sqrt(sum(sum((Fiterr.*~GauCirFitbase).^2)));
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Distance weighted RMS Error %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%Finding subscripts for max of nPERF
indmx = find(nPERF == 1);
[submxtheta submxphi] = ind2sub(size(nPERF),indmx);
%Distance weighted RMS error
[subtheta subphi] = ndgrid((1:size(nPERF,1)),(1:size(nPERF,2)));
theta_dist = abs(subtheta-submxtheta)*del_thetaR;
phi_dist = abs(subphi-submxphi)*del_phiR;
dwrmserr(evi) = sqrt(sum(sum(sqrt(theta_dist.^2 + phi_dist.^2).*(nPERF.^2))));
%Displaying at every event when ii <= ndisplayevent, at every 100
%events when ii > ndisplayevent
if (evi <= ndisplayevent || (evi > ndisplayevent && rem(evi,100) == 0) || evi == npt)
    %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    % Displaying reconstructed images %
```

```
      %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
      figure(2+BPsel);imagesc(thetaR,phiR,nPERF');grid on;colorbar;axis xy;
      axis square
      xlabel('Theta (degree)');ylabel('Phi (degree)');
      tit = sprintf('Reconstructed Image by Parra FBP with the %dth event',evi);
      title(tit);
      %status reporting
      msg = sprintf('Dealing with the %dth event',evi);  disp(msg); pause(0.5);
    end
  end %end of loop for event-by-event image recon
  PFBPtime = toc/60;
  msg = sprintf('Filtered Back-Proj time by Parra: %f mins for %d events',…
                PFBPtime,npt);
  disp(msg);
  figure(3+BPsel);mesh(thetaR,phiR,nPERF');grid on;colorbar;
  xlabel('Theta (degree)');ylabel('Phi (degree)');
  tit = sprintf('Reconstructed Image by Parra FBP after %d events',evi);
  title(tit);
  if 0
  %FWHM plot
  figure(4+BPsel);stem(FWHM);grid on;
  xlabel('Number of events');ylabel('FWHM [deg]');
  outind = (abs(FWHM-mean(FWHM)) >= mean(FWHM));
  FWHMplot = FWHM(~logical(outind));
  tit = sprintf('FWHM [deg] vs Number of event with avg=%d',…
        mean(FWHMplot(20:end)));
  title(tit);
  end
  %RMS error plot
  figure(5+BPsel);stem(dwrmserr);grid on;
  xlabel('Number of events');ylabel('dwRMS Error [AU]');
  tit = sprintf('dRMS Error with avg=%d',mean(dwrmserr(20:end)));
  title(tit);
end %loop for BP selection
```