Low Gain Avalanche Detectors for Particle Physics and Synchrotron Applications

Neil Moffat

Submitted in fulfilment of the requirements for the
Degree of Doctor of Philosophy

School of Physics & Astronomy
College of Science and Engineering
University of Glasgow

Abstract

Semiconductor detectors have a wide range of uses for particle physics and synchrotron applications. This thesis concentrates on the simulation, fabrication and characterisation of a new type of detector known as the low gain avalanche detectors (LGAD). The detector's characteristics are simulated via a full process simulation to obtain the required doping profiles which demonstrate the desired operational characteristics of high breakdown voltage (500 V) and a gain of 10 at 200 V reverse bias for low energy X–ray detection. The low gain avalanche detectors fabricated by Micron Semiconductor Ltd are presented. The doping profiles of the multiplication junctions were measured with Secondary ion mass spectrometry (SIMS) and reproduced by simulating the full fabrication process which enabled further development of the manufacturing process.

LGADs are interesting for high energy physics experiments due to their good timing performance. The need for such a detector is explained and results for 250 μm thick LGADs with a gain of 5 manufactured at Micron Semiconductor show comparable results to the other vendors, of 120 ps.

For low energy X–ray detection it is essential to operate at low noise levels. The aim of the project was to develop LGAD detectors with a highly segmented front side which would be compatible with the Timepix3 chip, which has an array of 256x256 pixels with a pixel pitch of 55 μm. However, when LGAD pixels are made to these size requirements their gain uniformity and fill factor are extremely degraded. Specific development for small pixel LGAD's was undertaken through simulation and possible structures have been identified to minimize this small pixel effect.

Acknowledgements

Declaration

I declare that except where explicit reference is made to the work of others, this dissertation is the result of my own work. This work has not been submitted for any other degree at the University of Glasgow or any other institution.

 Neil Moffat

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Semiconductor pixel detectors have been widely used in high energy physics experiments for precise tracking measurements, with a spatial resolution of a few micrometers [1]. In order to create an electron–hole pair in silicon an average energy of 3.6 eV must be deposited by an energetic particle or photon, this is around 10 times less than the energy required for a gaseous detector. Furthermore semiconductors have a greater stopping power than gas detectors, therefore they can be fabricated to be very thin with a fast signal response.

Silicon is used in high energy physics experiments as it has high efficiency, can be produced to be very thin and has fast charge carrier mobility, which results in a fast signal collection time. Silicon is the most widely used semiconductor detector and thus the most understood fabrication processes revolve around silicon. Particle tracks produced in the Large Hadron Collider (LHC) [2] can be resolved using 3-D tracking detectors, where the separation of the tracks is sufficient in time to determine each individual track. This is important for particle identification.

However the LHC is due to be upgraded by 2026 to the high luminosity LHC (HL–LHC), where the instantaneous luminosity is set to increase up to 7.5×10^{34} cm^{-2}s^{-1}. This corresponds to a 5 times increase in luminosity. This increase in luminosity will result in multiple events in the same region of the detector will occur very close in time. In order to better understand the interactions occurring, it would be useful to have a detector which is capable of tagging events in terms of the time of interaction. This will dramatically help with the reconstruction of the tracks. Therefore detectors are being developed which have both excellent time and spatial resolution.

The Low Gain Avalanche Detector (LGAD) has been developed to fulfil these require-ments. The technology is based on a silicon PIN diode. The LGAD has an extra doping layer underneath the junction implant which causes, when reverse biased, a high field region at the junction between these two implants. If the doping profiles are controlled precisely the electric field can be tailored to allow impact ionisation to take place. This

results in charge multiplication for negative charge carriers crossing this high field region. This design is similar to that of an Avalanche Photodiode (APD) but with lower gain. LGAD's tend to have a gain in the region of 5–20, compared with APDs which have a gain of 100's. High gain detectors tend to have large sensor noise and are difficult to segment. They are typically not run in full depletion mode. However LGAD's with a significantly lower gain that can be segmented and provide an increase in the slew rate with a gain of around 20 are the likely choice for the High Granularity Timing Detector (HGTD) [3]. If one makes an LGAD thin, this dramatically improves the timing performance due to the decrease in the rise time. Detectors with a good timing resolution produce signals that are large and fast. This technology has been shown to give a time resolution of around 30 ps [4].

Hybrid pixel detectors use a semiconductor material, which is segmented into an array of pixels, as the detection medium. Each pixel channel is connected to an individual low noise amplifier and signal processing electronics. Ionizing radiation creates free charge carriers in the semiconductor that drift, due an electric field gradient, towards the diode's electrodes, and in doing so induce a voltage. This is then converted to a charge by an integrating amplifier. The amount of charge collected is proportional to the energy of the incident radiation. The proportionality is given by the energy required to liberate an electron–hole pair. The signal is processed in the pixel to give either a count to register the pixel hit or the signal magnitude to allow spectroscopy. Both methods require a threshold to be applied to remove noise hits from the system. In modern hybrid pixel detectors the amplifier is followed directly by a discriminator, and counters are used to measure the time the signal is over this threshold to determine the signal magnitude. Such systems are already in use at the LHC and will be employed to benefit from the fast response of the LGAD detectors. An example of such a chip is the ALTIROC Asic [5], to be designed for the timing layer [6] of the ATLAS experiment.

For synchrotron applications there is desire to reduce the minimum energy of the X–rays detectable to a range of 1–5 keV. Both the Diamond Light Source and the Deutsches Elektronen–Synchrotron (DESY) have shown interest to improve their detection capabilities of X–ray photons in the tender energy window. The tender energy range is named as this because it lies between the typical *hard* (above 5 keV) and *soft* (below 1 keV) X–ray ranges. It is difficult to detect photons of this energy range due to the energy–dependant characteristic attenuation length of photons in silicon. As the energy of the photon reduces the attenuation length tends to decreases. At specific energies characteristic edges are found where attenuation length increases slightly, before decreasing again. At around 1–2 keV a distinct k-edge is found. As the attenuation length is higher at this energy, the percentage of energy transferred to the active region of the detector is greater.

There is a great motivation to pursue the study of electronic and atomic structure

of low-Z elements by probing their K absorption edges. It is especially interesting for biological studies since the K-edges of the main elements of interest, such as phosphorus (2.145 keV), sulfur (2.472 keV), chlorine (2.822 keV), potassium (3.608 keV) or calcium (4.0379 keV) which are all within the tender energy range [7]. It is thought that with the use of LGAD technology combined with hybrid pixel technology, it will be possible to increase the size of the initial signal at these energies to vastly improve the SNR and hence improve the image quality.

Hybrid pixel detectors are already used in single photon counting pixel detectors, which are used for a number of applications including, X-ray diffraction and X-ray crystallography. The discriminator threshold implies a minimum detectable signal.

One such group of pixel detector read-out chips for particle imaging and detection is the Medipix family of detectors. Several iterations of chips have produced a 256x256 array of 55 μm pitch pixels. The different iterations provide enhanced capabilities, starting with photon counting (Medipix Mode) to the latest technology, where the chips can simultaneously collect data on the deposited energy of a particle and provide time-of-arrival information. The minimum detectable signal, defined by the discriminator threshold has been reduced to an equivalent photon energy of 1.8 keV [8]. This is achievable with standard silicon pixel sensors using the Timepix3 readout chip [9].

Introducing an LGAD detector with fine segmentation to match the available readout electronics of this technology can provide improved detection capabilities. The internal gain of the LGAD sensor can allow sub-threshold detection of low energy photons by a factor defined by the gain. An internal gain of around 4 could provide enough electrons to obtain a meaningful output signal from a 1 keV X-ray interaction. Along with the fast response of the readout chip with a time resolution of 1.6 ns [9] it should be possible to have a detector with good time and high spatial resolution's able to detect sub-keV photons. The use of a well-known and established readout chip architecture should allow quick implementation of the technology for experimental and commercial applications in Synchrotron science.

Applications of this technology could vary widely across many synchrotron experiments which perform in the low energy x-ray range. Several low energy x-ray diffraction [10] and spectromicroscopy applications [11] could benefit from the use of LGAD detectors. Possible applications include the study of biological material, such a proteins, and new materials, such a graphene.

The aim of the work described in this thesis is to design a silicon detector with low gain, simulate its response using TCAD tools and fabricate and test prototype detectors of this type. The work has concentrated on the development of the simulation to produce a reliable fabrication process. The design was then altered to create pixellated structures compatible with the Timepix3 readout chip.

In this chapter, the basic principles of LGAD technology were introduced in terms of the electric field profile required to produce a detector with gain. Some of the possible applications of this technology are discussed, including for particle physics to allow the discrimination of tracks and for synchrotron applications, with the idea to benefit from existing readout electronics to produce a detector capable of detecting sub–threshold photons.

Chapter 2 gives an overview of the fundamental principles of semiconductor detectors, with a focus on the applications for detectors in physics experiments. The basic behaviour of semiconductor detectors are described in order to understand the principles behind producing a detector with gain. The basic fabrication process of LGAD detectors is introduced with comparison to a standard PIN diode, highlighting the importance of process control.

Chapter 3 gives an overview of the simulation package "Synopsys TCAD" [12] used in this work, giving details of the importance of using simulation for the development of detector technologies. The principles of semiconductor simulation are described with a focus on process simulation and the implementation of the fabrication steps used to create the required doping profiles. Electrical simulations are performed to provide results on electrical performance and internal gain for varying designs where the models used are described. The LGAD structure is discussed in detail to describe how simulations were performed and validated. The simulation process involved using 1D and 2D structures to first define a suitable multiplication region; followed by the design of the implants at the edge of the multiplication region capable of withstanding the high fields produced in this multiplication region.

Chapter 4 gives details of the fabricated devices produced for this project. Details for each fabrication run are provided with varying process parameters and for the different mask designs used. Devices fabricated varied in active pixel size from 55x55 μm^2 up to 5x5 mm^2. The devices were pixellated into arrays of 2x2 up to 256x256. The different sensor characterisation techniques are discussed in some detail, which include current–voltage (IV), capacitance–voltage (CV) and the Transient Current Technique (TCT). The results are presented in detail for a variety of sensors from each run. The final detectors presented give the required gain of 5–20 while withstanding reverse bias voltages of up to 500 V.

Chapter 5 introduces the concept of LGAD devices for applications requiring fast timing response. The sensor requirements needed for good timing resolution are described which include fast response and low noise as well as uniform signal response in time. The timing set–up developed is described with initial results from a 250 μm thick sensor presented a time resolution of around 150 ps, with details of the analysis procedure discussed.

Chapter 6 is devoted to the simulation of small pixel LGAD's where the goal is to use LGAD technology with the Timepix3 chip. The LGAD devices have shown a change in performance when the pixel size is reduced. This change results in the loss or reduction in

gain for pixels with a pitch of 55–220 μm. This reduction in gain is discussed and proposals for improving the performance of small pitch LGAD are presented. Due to their small size the LGAD's have a fill factor issue, defined as the area of the pixel with gain divided by the physical area of the pixel. As the size of the pixel reduces, the area of the gain region reduces in size compared with the size of the pixel. Current technology limits the fill factor to less than 30% for 55x55 μm^2 pixels, which is provided using a limited periphery region, resulting in a reduced breakdown voltage. Future work is expected to produce results from fabricated sensors of this pixel size to compare with simulated results.

Chapter 7 gives an overview of the results of this thesis. The general conclusions about the work are summarised with the main results discussed.

Chapter 2

Silicon Detectors

Silicon is one of many types of semiconductor material that is used for the detection of radiation. It is the most widely used and studied semiconductor material with some very important characteristics for use a detector in high energy physics (HEP) and synchrotron applications. Silicon is used as the semiconductor material in this project as the understanding of its electrical properties and fabrication technology is well developed. For a new technology it is important to have existing knowledge to build upon. Silicon is the most studied element of all and for fabrication purposes is the easiest to work with. It has a low ionization energy (3.6 eV on average) compared with other semiconductors, which is important for obtaining a good energy resolution. Silicon has high mobility charge carriers which enables fast charge collection, especially important for this work.

Furthermore for the R&D involved with this project it was important to have a large supply of available silicon wafers for use on a variety of fabrication test runs. Other semiconductor materials are available which have many advantages over silicon. Diamond detectors are used as they require less power, due to lower leakage current and have a lower radiation length which is good for tracking detectors. However they have a larger ionization energy which can reduce the collected signal. Silicon has advantages over other semiconductors for particle physics detectors such as: a modest bandgap (1.12 eV) which allows room temperature operation, unlike Germanium detectors with a bandgap of 0.66 eV. Silicon can be made defect free which reduces trapping and improves signal collection and therefore spectroscope performance, unlike Gallium Arsenide (GaAs) detectors which will always have defects. While GaAs is a well used semiconductor but has poor charge collection due to deep traps and is not radiation hard.

A summary of semiconductor physics is presented in the following section, along with some details on the fabrication process relevant to this work.

2.1   Introduction to silicon as a semiconductor

2.1.1   Crystal Structure

Silicon is an element of the IV–th group of the periodic table, characterised by four electrons in its outermost orbital shell (valence electrons). Through the formation of covalent bonds a diamond lattice structure is formed. The covalent bonds are created due to the sharing of electrons between atoms in the material. Pure silicon is too reactive to be found in nature. Instead various crystal growth techniques are used to produce mono–crystalline ingots, which is the basic material used to produce silicon detectors [13].

Silicon detectors are processed on thin wafers cut from the ingots of crystalline silicon. The lattice orientation of a silicon wafer is determined by the orientation of the seed used to grow the ingot. The crystal orientation is defined by the Miller indices of the cutting plane (wafer surface), denoted by (h,k,l). There are three possible wafer orientations used to produce silicon detectors, <100>, <110>, and <111> illustrated in figure 2.1. The difference between the three orientations can significantly change the electrical properties of the detectors. Typically, <111> and <100> silicon is used in high energy physics as they are more radiation hard compared with the <110> silicon [14].

The resistivity of a wafer depends on the crystal growth technique used. For ionising radiation sensors, the resistivity of the crystal should be large in order for full depletion to be reached at a lower bias voltage. Full depletion is the term used for when the semiconductor material has no majority free charge carriers within the bulk material. It should be noted that for this project <100> is exclusively used due to increased control of doping profiles, which is inherently important for the LGAD design. Resistivites of between 2–10 kΩ/cm are used, with a wafer thickness in the range of 200–300 μm.
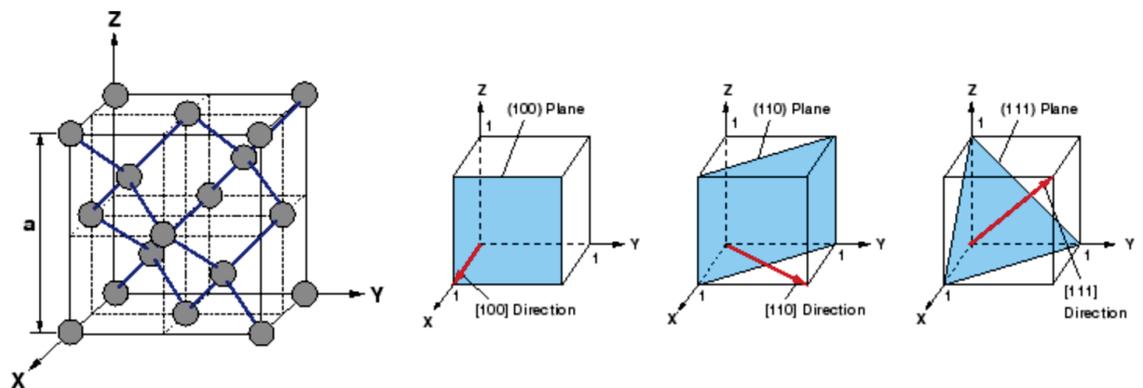
Figure 2.1: From left to right, diamond structure of the silicon crystalline lattice. (100), (110), and (111) lattice orientation planes (blue). The orientations <100>,<110>, and <111> refer to the perpendicular vector to the corresponding plane. [15]

2.2   Energy Bands

Crystals are characterised by some periodicity pattern in their atomic distribution. For an isolated atom, the electrons can have discrete energy levels. However when there are N, isolated atoms which are brought together, interactions occur which cause a shift in the energy levels due to forces of attraction and repulsion. This results in a continuous band of energy which extends over a few eV. In reality semiconductor band formation is slightly more complicated, for which the band formation of silicon will be described.

A silicon a atom contains 14 electrons. Of these 14, 10 are contained within two inner most energy shells. The remaining 4 are weakly bound in the outer shell and can be easily involved in chemical reactions. This outer shell has two sub–shells 3s and 3p. The 3s sub–shell has two allowed quantum states per atom, where it contains two electrons at T = 0 K. The 3p sub–shell has six allowed quantum states per atom, where the remaining two electrons are contained. When an N number of silicon atoms are brought together to form a solid, the atoms interact and create a continuous band as in the previously discussed case. However, because of the two sub–shells in the outer shell of silicon, two energy bands are formed, a low and a high. Both energy bands may contain 4 electrons per atom. At absolute zero the low energy (*Valence band*) will be full and all states in the high energy band (*Conduction band*) will be empty. The top of the valence band can be referred to as E_V and the bottom of the conduction band as E_C. The energy between the two bands $(E_C - E_V)$ is known as the band gap energy, E_g, which equals the energy required to break a bond to free an electron from the valence band to the conduction band, leaving a hole in the valence band [16]. The band gap energy is an inherent property of a material and dependent on temperature. The magnitude of this band gap energy defines the electrical properties of the material. The band gap energy can be used to describe the material in terms of its electrical and thermal properties. This is shown in figure 2.2, where it describes the difference in the energy band gap for a metal, insulator and a semiconductor.
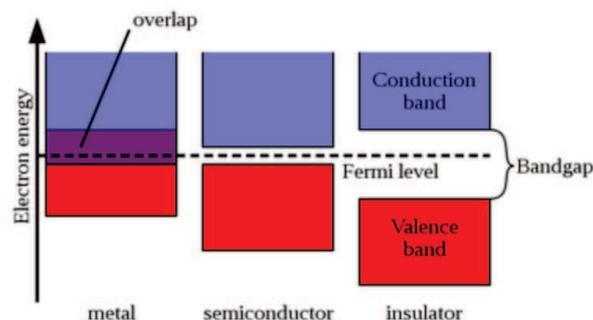
Figure 2.2: Simple diagram of the position of the valence and conduction bands for metal, insulator and semiconductor materials [17].

The materials can be described in the following way:

- **Insulator:** Large band gaps ($E_g \geq 2.5$ eV [18]) due to very tightly bound electrons between neighbouring atoms. At room temperature very few electrons are excited from the valence to the conduction band, resulting in very low electrical conductivity. Silicon dioxide is an insulator which is widely used in the fabrication of semiconductor devices as it does not conduct current.

- **Conductor:** Very small or non-existent energy gap ($E_g = 0$). A large number of electrons occupy the conduction band even at low temperatures. This results in a large number of freely moving electrons. Therefore, conductors can readily conduct current.

- **Semiconductor:** In between an insulator and a conductor. At $T = 0$ K no electrons occupy the conduction band, all the electrons are in the valence band. Which results in the semiconductor being a poor conductor at low temperatures. At room temperature some electrons are excited into the conduction band and thus, a moderate number of free charge carries are created in this material. A small applied electric field can easily move electrons resulting in a moderate current.

Silicon has an energy gap of 1.12 eV so is a semiconductor. Depending on its purity level semiconductors are classified as intrinsic or extrinsic semiconductors. All silicon used for particle or x–ray detection have some level of impurity so is classed as an extrinsic or doped semiconductor. Controlling this level of impurity is important for producing semiconductor devices.

2.3   Carrier Concentration

Semiconductor properties can be altered by introducing controlled amounts of impurities during the crystal growth process. This process is refereed to as "doping". If the impurities give electrons to the conduction band they are referred to as donors, and if they give holes to the valence band they are referred to as acceptors.

2.3.1   Intrinsic Semiconductors

A semiconductor is referred to as intrinsic if the concentration of impurities is lower than the concentration of thermally generated carriers. The thermally generated carrier concentration can be calculated by integrating over all the valence and conduction band states, taking the density of states and the probability of occupation into account using the Fermi–Dirac distribution [16]. The resulting hole (p) and electron (n) concentrations can be calculated as follows

$$p = N_v exp \frac{(E_v - E_F)}{kT}, \tag{2.1}$$

$$n = N_c exp \frac{(E_c - E_F)}{kT}, \tag{2.2}$$

where N_v and N_c are the density of the states in the valence and conduction bands respectively, k is the Boltzmann constant, T is the Kelvin temperature and E_F is the Fermi level. The Fermi level being the energy level in a semiconductor where the probability of finding an electron is half. The Fermi level of intrinsic silicon is in the middle of the band gap. This in when E_c - E_F = E_F - E_v, where E_c is the position of the bottom of the conduction band and E_v is the position of the top of the valence band. For intrinsic silicon we find that the carrier concentration of electrons and holes are equal, $n = p = n_i$, where n_i is the intrinsic carrier concentration.

2.3.2   Doped Semiconductors

Doping is the addition of controlled amounts of specific impurity atoms to alter the conductivity of the silicon substrate by increasing either the hole or electron free carrier concentration. The substrate is the term given for the bulk semiconductor material. By introducing small amounts of impurity atoms into a crystal lattice structure, the number of free electrons and holes can be altered. When an impurity element is introduced it will either donate an extra electron to the lattice or accept one depending on its outer electron shell number. For example if a silicon lattice is doped with boron, the boron atom replaces a silicon atom in the lattice structure. The boron atom has three valence electrons, therefore in order to create a strong covalent bond with the four neighbouring silicon atoms an additional electron is "accepted" and a positively charged "hole" is generated in the valence band. This creates a p–type semiconductor and boron is an *acceptor*. The other scenario is the introduction of a "donor" impurity, for example phosphorus. The phosphorus atom has five valence electrons and therefore when covalent bonds are formed with the four neighbouring silicon atoms a single loosely bound electron is "donated" to the conduction band. This creates an n–type semiconductor because of the addition of the negative charge carrier.

2.3.3   Carrier Transport and Multiplication

There is a variety of transport processes, which in combination control the working of a semiconductor device. Three important processes are *drift, diffusion* and *impact ionisation*.

Drift

Electrons in a semiconductor in thermal equilibrium move rapidly in all directions by thermal motion. The motion of these electrons can be attributed to random scattering from collisions with lattice atoms, impurity atoms and other scattering electrons. The net displacement of an electron, when no external field is applied, is zero. The average time between collisions is the *mean free time*, τ_c and the average distance between collisions is the *mean free path*.

When an external electric field, E, is a applied the electrons will experience a force equal to $-qE$ from the field and will be accelerated in the opposite direction to the field, where q is the charge of an electron. Each electron now has a net displacement due to the combination of thermal motion and *drift velocity* in the opposite direction to the field. The drift velocity, v_n can be calculated by equating the momentum of the electron and the impulse imparted on the electrons during the time between collisions. The momentum applied by the field is given by $-qE\tau_c$ and the momentum gained is $m_n v_n$. Where m_n is the effective mass of electrons [16]. Equating these terms gives

$$m_n v_n = -qE\tau_c. \tag{2.3}$$

Rearranging gives

$$v_n = -\left(\frac{q\tau_c}{m_n}\right)E. \tag{2.4}$$

The electron drift velocity in equation 2.4 is proportional to the applied electric field. The proportionality factor depends of the mean free time, the charge of an electron and the effective mass. This term is known as the *electron mobility*, μ_n. The mobility describes how strongly the electron motion is determine by the field. This can also be used to describe the hole motion in the valence band given by the hole mobility, μ_p. Electron and hole drift velocity can be expressed as

$$v_n = -\mu_n E, \qquad\qquad v_p = \mu_p E. \tag{2.5}$$

The drift velocity shows a linear increase with electric field strength at low fields. However at higher fields a non linearity appears due to increased scattering. The drift velocity saturates at high fields.

In the condition where there is a low impurity atom concentration the mobility is determined by the lattice scattering. Lattice scattering is produced by the thermal vibrations of the lattice atoms, which can occur at any temperature above zero kelvin. These vibrations cause disturbances in the lattice structure and enable energy to be passed between the carriers and the crystal. Lattice vibrations are increased with increasing temperature and therefore mobility decreases with increasing temperature. At high impurity atom concen-

tration, impurity scattering becomes a significant factor in the mobility. This is when a charge carrier passes by a dopant impurity and undergoes a coulomb interaction, losing energy. However at high temperatures the charge carriers move at a higher velocity with less chance of interaction and therefore this phenomenon is reduced. At values of over 200 ∘C the doping density has little effect on the mobility. At room temperature the mobility is limited by the ion impurity concentration, where the concentration of impurity ions starts to limit the mobility at doping densities of the order of 10^{16} cm^{-3}. For low impurity devices, such as those used in HEP the temperature dominates the mobility achievable.

Mobility is highest at low ion impurity and is limited by the lattice vibrations. It should be noted that the hole mobility is lower than that of the electron mobility. This is due mostly to the lower effective mass of electrons.

Diffusion

When there is a spatial variation in the carrier concentration within a semiconductor another current process is observed. The carriers can be initially disproportionately distributed so as to cause the carriers to be located in high and low concentrations respectively. The carriers tend to flow from an area of high concentration to an area of low concentration through a process called *diffusion current* to reach an equilibrium. The rate of carrier flow can be described as

$$F = -v_{th}l\frac{dn}{dx} \equiv -D_n\frac{dn}{dx},$$

(2.6)

where $D_n \equiv v_{th}l$ is called the *diffusivity*. The v_{th} is the thermal velocity of the electrons and l is the mean free path. In terms of electron flow, as each electron carriers a charge $-q$, the electron flow produces a current

$$J_n = -qF = -qD_n\frac{dn}{dx}.$$

(2.7)

Diffusion current is the result of random thermal motion of charge carriers within a concentration gradient. The current flows in the opposite direction of electrons and thus is positive. The diffusivity and the mobility can be related by the *Einstein relation* for electrons and holes respectively

$$D_n = (\frac{kT}{q})\mu_n \qquad\qquad D_p = (\frac{kT}{q})\mu_p.$$

(2.8)

When an electric field and concentration gradient are both present the resultant flow is due to both the drift and diffusion components given by:

$$J_n = q\mu_n nE + qD_n\frac{dn}{dx}\mu_n,$$

(2.9)

for electron current and given by

$$J_{\mathrm{p}} = q\mu_{\mathrm{p}}pE - qD_{\mathrm{p}}\frac{dp}{dx}\mu_{\mathrm{p}}, \tag{2.10}$$

for hole current. At high fields the terms $\mu_{\mathrm{n}}E$ and $\mu_{\mathrm{p}}E$ can be replaced with the saturation velocities v_{n} and v_{p} respectively [16].

Impact Ionisation

When the electric field in a semiconductor is increased above a certain value, the carriers begin to have sufficient kinetic energy to create electron–hole pairs by an *avalanche process*. If we imagine a free electron in the conduction band. If the electric field is high enough, when the electron is moving around the lattice it can gain enough kinetic energy to break a bond. This is because the electron can pass on a sufficient portion of its kinetic energy to the lattice. This results in ionising an electron from the valence band to the conduction band. This generates an electron–hole pair. Moreover, these two charge carriers now accelerate in the field and collide with the lattice, generating more electron–hole pairs. This process can continue and is called *impact ionisation*. The required kinetic energy for the ionisation process is given by [19]

$$E_0 = \frac{1}{2}mv_{\mathrm{s}}^2 = 1.5E_{\mathrm{g}}, \tag{2.11}$$

where m is the mass of the charge carrier, v_{s} is the saturation velocity and E_{g} is the minimum energy required to create an electron–hole pair. The actual energy required to generate an electron–hole pair in silicon is 3.2 E_{g} for electrons and 4.4 E_{g} for holes [16]. This process results in the multiplication of the carriers and a signal gain. The electron–hole pair generation rate, G, due to impact ionisation is given by [20]

$$G = \alpha_n n v_n + \alpha_p p v_p \tag{2.12}$$

where (n, p), (v_n, v_p) and (α_n, α_p) are the electron and hole density, velocity, and ionisation rate respectively. The ionisation rate strongly depends on the electric field (E) and may be defined as

$$\alpha = \frac{E}{E_{th}} \exp[-E_i/E] \tag{2.13}$$

where E is the high–field in the device, E_{th} is the high–field effective ionisation threshold energy (for silicon equal to 3.6 eV for electrons and 5.0 eV for holes) and E_i is the threshold field due to ionisation scattering [21].

Resistivity

A term which can be used to determine many attributes of a detector is the *resistivity*. Particle physics detectors tends to use high resistivity materials as they require detectors to be fully depleted during operation. The resistivity, ρ, of a semiconductor can be expressed as

$$\rho = \frac{1}{q(n\mu_n + p\mu_p)} \tag{2.14}$$

as shown by [16].

Higher resistivity substrates require a lower reverse bias voltage, compared with a lower resistivity substrate, in order to become fully deplete of free charge carriers. The resistivity is controlled by the doping concentration of the semiconductor. Typical values used in high energy physics are 1–10 kΩcm.

2.3.4 The p–n junction

This section aims to describe the basic behaviour of a semiconductor device by looking at the p–n junction. A p–n junction is created by having a semiconductor material with both p– and n–type regions in immediate contact. To help explain the operation of a p–n diode one can imagine the opposite sides of the junction to be at first isolated, and then brought into contact. A p–type material with an excess of holes is brought into contact with an n–type material with an excess of electrons. The charge carriers diffuse along the concentration gradient. That is the movement of electrons from the n–type to the p–type and vise-versa for the holes, as shown at the top of figure 2.3. As these charge carriers move across the junction they recombine. This produces a potential difference Φ_{bi} which prevents further charge flow. This is maintained by the static space charge built up around the junction as the donor and acceptor atoms are ionised. Equilibrium is reached when the junction's potential difference is large enough to prevent any net charge transfer across the junction. The value of this potential is called the built–in voltage (V_{bi}) and depends on the intrinsic carrier concentration, n_i, as well as the hole, p, and electron, n, concentrations respectively. This is shown by [20]

$$V_{bi} = \frac{kT}{e} \ln \frac{np}{n_i^2} \tag{2.15}$$

The space charge region around the *p–n* junction now has no free charge carriers and as such is called the depletion region as shown in figure 2.3. In reality a p–n junction is produced by introduction of impurity ions at the front and back sides of a silicon wafer, either p– or n–type. The process of this will be discussed in a following section. This produces an abrupt junction where the depletion region will extend further into the less

heavily doped side of the device. The depletion region can be calculated by first solving Poisson's equation given by

$$\frac{d^2\Psi}{dx^2} = -\frac{q}{\epsilon_s}(N_D - N_A + p - n).$$ (2.16)

Where Ψ is the electrostatic potential, ϵ_s is the permittivity of silicon, N_D is the donor concentration and N_A is the acceptor concentration. As the depletion region is free of charge carriers this can be simplified to

$$\frac{d^2\Psi}{dx^2} = +\frac{qN_A}{\epsilon_s} \quad for \quad -x_p \leq x < 0,$$ (2.17)

$$\frac{d^2\Psi}{dx^2} = -\frac{qN_D}{\epsilon_s} \quad for \quad 0 < x \geq x_n.$$ (2.18)

where x_n and x_p are the distances in which the depletion region has reached from the junction between the two regions. Thus, the total depletion width w is given by

$$w = x_p + x_n.$$ (2.19)

The electric field in the depletion region can be obtained by integration of equations 2.18 and 2.17, where the maximum field exists at x = 0 and can be expressed as

$$E_m = \frac{qN_D x_n}{\epsilon_s} = \frac{qN_A x_p}{\epsilon_s}$$ (2.20)

and shown in figure 2.3. One can then integrate again over the depletion region, which gives the total potential variation

$$V_{bi} = \frac{qN_A x_p^2}{2\epsilon_s} + \frac{qN_D x_n^2}{2\epsilon_s} = \frac{1}{2}E_m w.$$ (2.21)

This is the built-in voltage, which when rearranged can give a value for the depletion region width. However devices are generally run with an external potential, V, applied which effects the depletion width and is given by

$$w = \sqrt{\frac{2\epsilon_s(N_A + N_D)}{qN_A N_D}(V_{bi} - V)}.$$ (2.22)

If V is applied with the same sign as the built-in voltage (reverse bias) the width of the depletion region increases. This means applying a positive voltage to the n-type region and vice versa. If the voltage has the reverse polarity of the built-in voltage (forward bias), then the width of the depletion region shrinks [20]. Note that from equation 2.22, V is positive for a forward bias and negative for a reverse bias.

In forward bias the applied external voltage results in a reduction in the total electro-
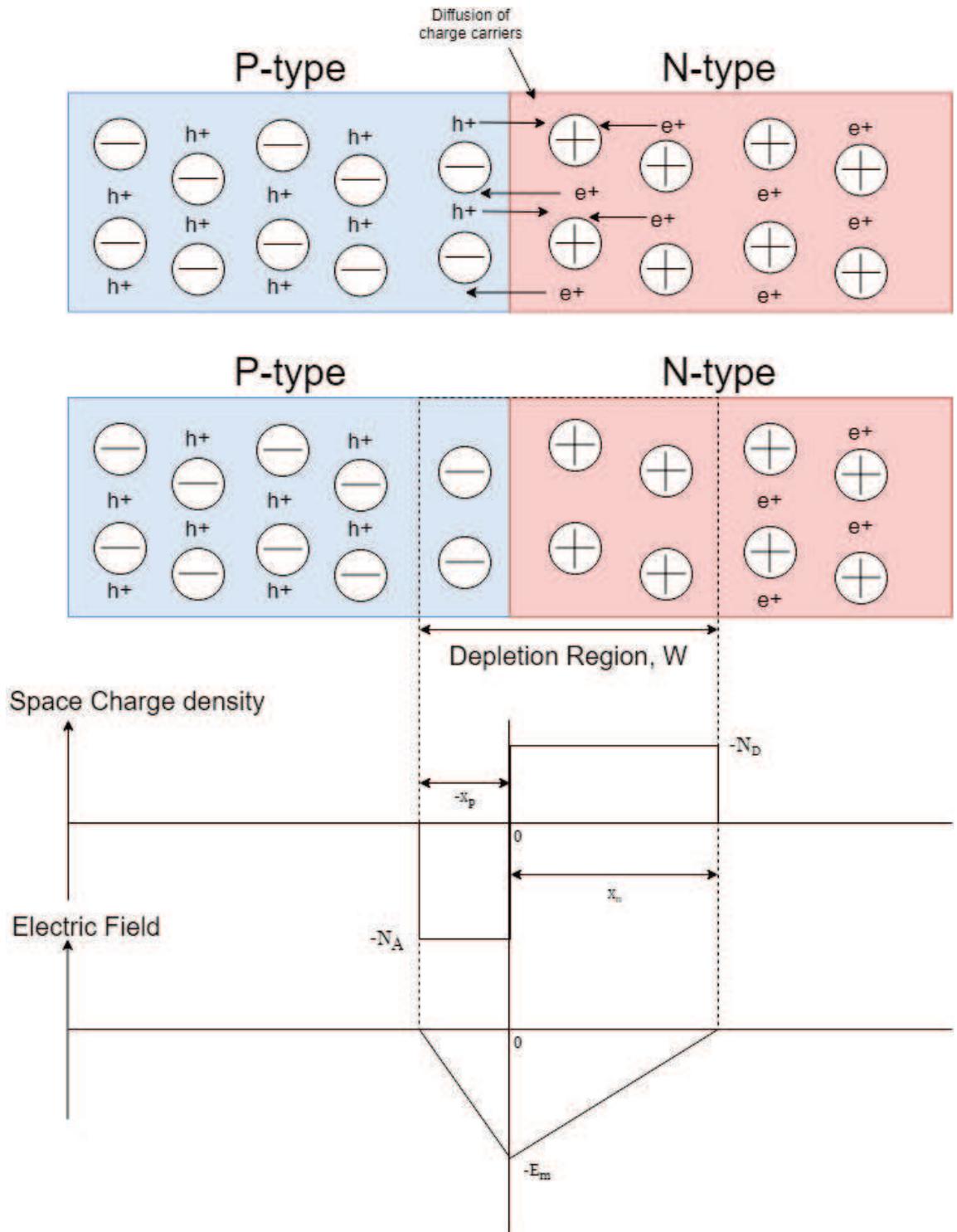
Figure 2.3: Simplified image of a PN junction in equilibrium showing the depletion region. The space charge density and electric field in the depletion region are shown.
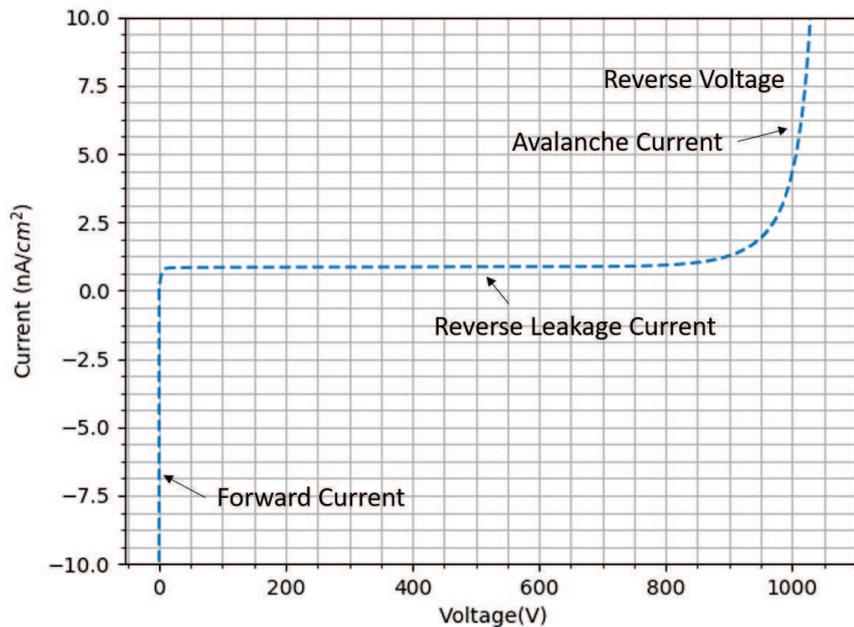
Figure 2.4: Ideal IV characteristics of a simulated p–n junction diode

static potential across the junction and thus the width of the depletion region decreases. Once the depletion region is zero a current can flow freely across the detector. The built–in voltage in silicon is around 0.7 V, depending on bulk doping concentration and tempera–ture [22]. This current grows exponentially with applied voltage above the built–in voltage.

However when a reverse bias is applied, the total electrostatic potential increases and the depletion width grows. In an ideal detector the current is low under reverse bias conditions until a critical point at which avalanche breakdown occurs, at an electric field of $\approx$ 300 kV/cm [23].

A typical IV characteristic for an ideal simulated p–n junction diode is shown in figure 2.4.

Full depletion is the term used for when the depletion region has been increased to the full thickness of the detector and be calculated using

$$V_{fd} = \frac{q(N_A + N_D)}{2\epsilon} w^2 \tag{2.23}$$

where V_{fd} is the full depletion voltage. High energy physics silicon particle detectors are always operated in reverse bias mode. Devices used in this project have full depletion voltages in the range of 60–120 V for 200 μm thick devices. This full depletion voltage can be measured experimentally by calculating the capacitance as a function of voltage. The capacitance of the depletion region will be given by

$$C = \frac{dQ}{dV},\tag{2.24}$$

where the charge is stored on either side of the depletion region. This is given by Q and –Q respectively and it can be shown that the capacitance per unit area depends on the width of the depletion region by $C = A\frac{\epsilon_s}{w_d}$. As the depletion region increases, the capacitance decreases and thus the full depletion voltage is calculated as the voltage at which the capacitance no longer decreases.

2.4   Principles of detector operation

A silicon detector is operated under reverse bias conditions. The depletion region may be only a few μm's or the full thickness of the detector depending on the structure of the detector. All devices in this work were operated fully depleted. This means that at all times there is an electric field across the full thickness of the detector. When photons or high energy particles traverse the detector they deposit some or all of their energy. This will create free electron–hole pairs, the required energy for this is $\approx$ 3.6 eV per electron–hole pair. As there is an electric field across the detector all the generated free electrons and holes drift in this field with their drift velocities v_n and v_p in the direction of the anode and cathode respectively. The transient current generated can be described by Ramo's theorem [24]. This is generally amplified and measured by external readout electronics. If all the energy of an incoming particle is deposited one can calculate the incident's particles energy, this is known as *spectroscopy*. Generally in high energy physics the particle track, the path which the particle takes, is desired as this allows the momentum of the charged particle to be determined from the reconstructed trajectory if the electric charge of the particle is known.

The basic structure used for this project was developed around a photodiode. The basic structure of which is shown in figure 2.5 showing the charge collection from a traversing particle. The process of fabrication is described in section 2.6. As the particle traverses the detection medium it generates electron–hole pairs uniformly as a function of depth for a *minimum ionising particle*, mip. Where a mip is a type of particle which loses close to the minimum amount of energy per unit distance while traversing a detector medium given by the Bethe–Bloch equation [25]. The electrons drift towards the n–type electrode and holes towards the p–type electrode inducing a measurable current signal. The integral of the current collected is the total charge collected, which for a detector with 100% charge collection efficiency is equal to the total charge introduced by the traversing particle.
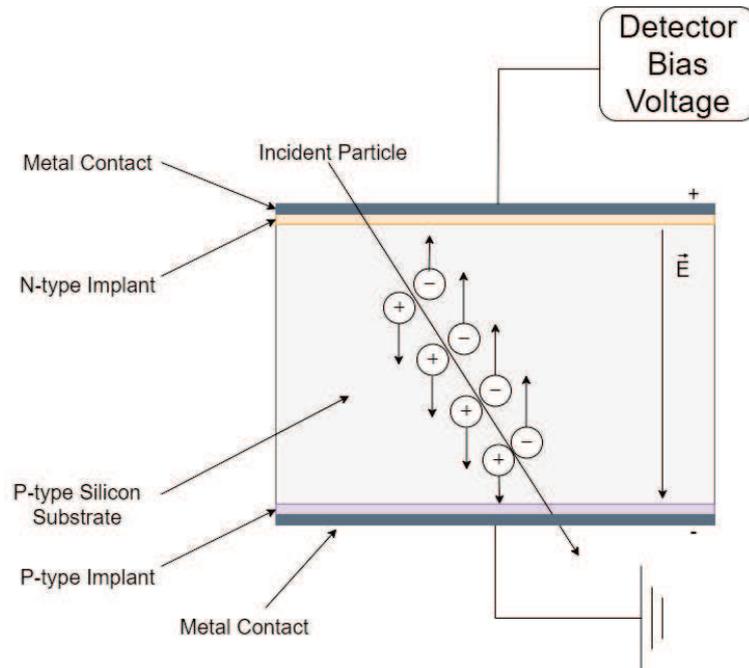
Figure 2.5: Structure of a simple photodiode, showing the charge collection from a traversing particle.

2.5 Introduction to Avalanche detectors

This section will give a brief introduction and review of avalanche detectors used in high energy physics and synchrotron applications.

2.5.1 Avalanche detectors

Charge multiplication is well understood in gases and solids and is based on the avalanche process, briefly described in section 2.3.3, whereby moving charges in a high field undergo impact ionisation producing a gain in the number of collected charges with respect to the initial generated charges. This has been used in semiconductors to produce silicon photosensors with internal multiplication such as avalanche photodiodes (APD) [26] with gain in the 100's , silicon photon multipliers (SiPM) [27] with a gain of the order 10^6 and Low Gain Avalanche Detectors (LGAD) with a gain of 10–20.

The Avalanche Photodiode (APD)

An APD is a p-n diode with internal gain due to a localised high electric field at the junction between p- and n-type doped silicon. The multiplication is in a proportional regime with typical values between 50 and 200. The temperature and voltage supply need to be highly controlled as the APD is operated close to the breakdown voltage. There are a few different structures used for APD's, however only the reach–through APD (RT-APD) structure will
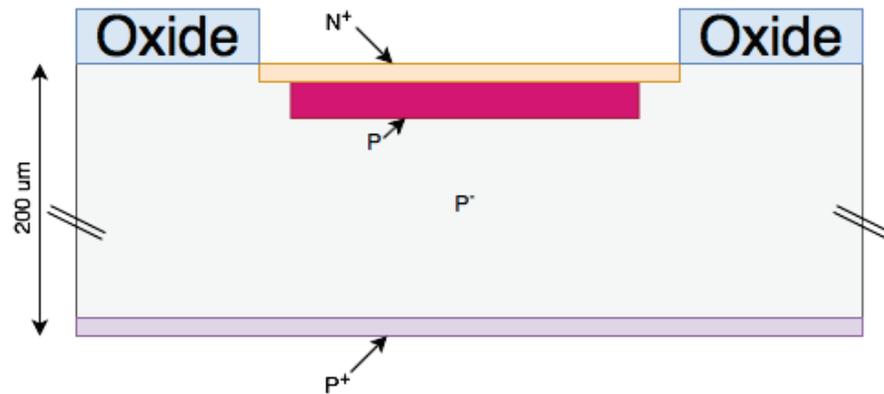
Figure 2.6: Schematic representation of a 200 μm thick RT–APD: X–ray photons generate electron–hole pairs within the fully depleted p⁻ active region. The n⁺/p junction generates a high field region where multiplication occurs.

be described here. The basic RT–APD structure corresponds to a $p^+/p^-/p/n^+$ [28], where p^+ indicates a region of comparatively high p–type doping and p⁻ indicates a region of comparatively low p–type, as shown in figure 2.6.

The typical electric field of such a device is shown in figure 2.7 where the peak electric field is at the n⁺/p junction and the active area is the majority of the bulk p–type substrate.

As electrons in silicon are more strongly ionizing than holes, due to their higher kinetic energy at a given electric field, they require a lower electric field for impact ionisation. These devices are run at a relatively low reverse bias voltage with a full depletion voltage of $\approx$ 100 V, [29]. As the devices are fully depleted this allows the device to be used with radiation entering through the p^+ contact. Electrons produced through the absorption of the radiation drift towards the high field region, where the multiplication occurs. The holes generated in this interaction drift across the full thickness of the detector to the p^+ contact.

These devices have been used in the detection of low to medium energy X–rays below 20 keV with good energy resolution [30], [28].

The Silicon Photo Multiplier (SiPM)

It is possible to run APDs in a proportional regime where the amplification is about 100 dependent on the reverse bias voltage. It is also possible to use an APD in the Geiger–Müller regime, above breakdown voltage. The internal amplification becomes huge and a photo–generated electron within the depletion region of the diode will cause it to breakdown catastrophically. This can be controlled using a "quenching" resistor to limit the breakdown current and return the diode to a stable state, ready for another photo–interaction. Devices operated in this regime are known as a "Silicon Photo Multiplier", (SiPM). It's important to note that the SiPM is run in a condition whereby both electrons and holes undergo impact ionisation in the high field region.
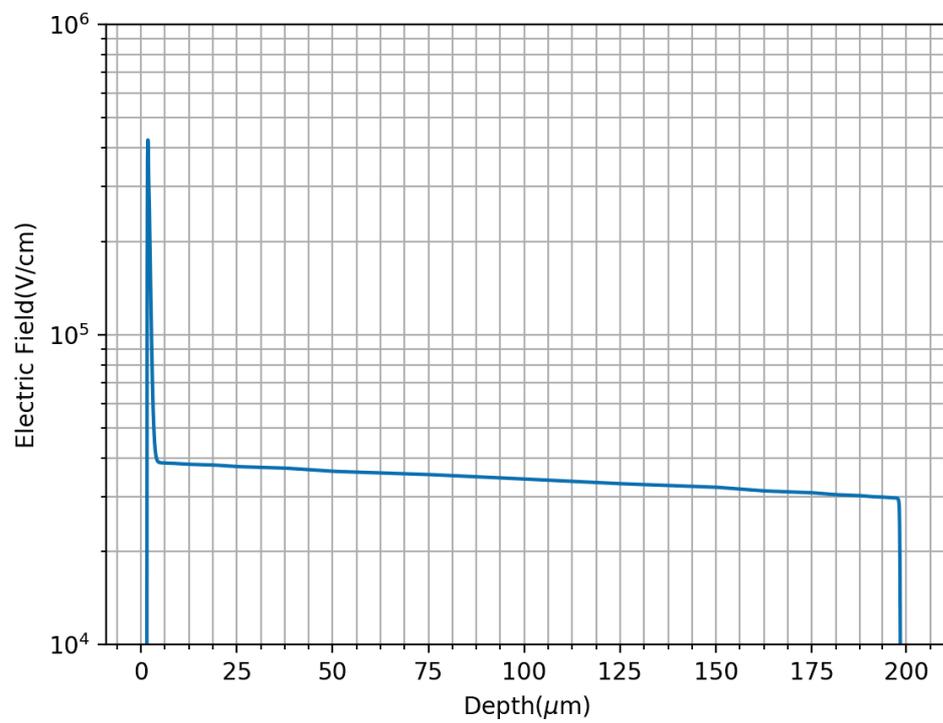
Figure 2.7: Simulated Electric field profile in an RT–APD at the centre of the main junction at a bias voltage beyond full depletion.
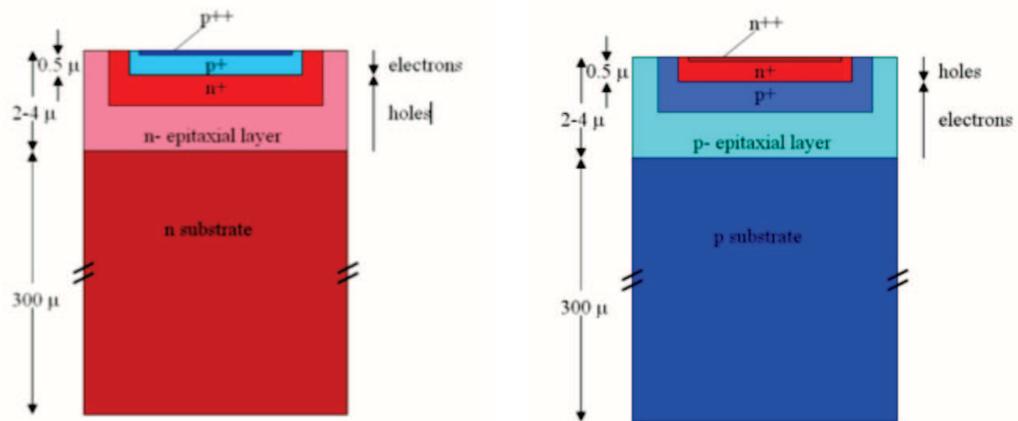
Figure 2.8: Basic view of a SiPM on both n– and p–type substrates. On the left is p–on–n structure and the right a n–on–p structure.

SiPM cannot be used in a full depletion mode as the probability of an early breakdown is larger with increasing bias voltage. The larger bias voltage increases the effective detector volume and increases the chance of a thermally generated electron–hole pairs in the bulk, causing a signal. For this reason only a thin depletion is possible, which has to be optimally fabricated depending on the application. SiPM can be optimised for different wavelengths of light, as the penetration depth of light depends on its wavelength. Typically a 300 μm low resistivity support wafer is used to grow a thin epitaxial layer (2 to 4 μm) of the same dopant type as the support wafer [31]. A heavily doped region is then created by diffusion of the same dopant type as the epitaxial layer and the support wafer, either n– or p–type. The p–n junction is then formed by a shallow implant of the opposite dopant, followed by an extremely thin but heavily doped layer of the same dopant. The structures for both p–on–n and n–on–p are shown in figure 2.8.

SiPM can be fabricated in array structures which make them useful for imaging purposes. These can be used to replace traditional PMT's in gamma cameras. Due to the structure of the SiPM they can be used for fast timing applications as the active area is only a few μm's, resulting in a fast response and a large signal. The timing resolution is of the order of 10's of ps [32]. The fill factor for standard SiPM is rather low with values in the region of 10-60% [33]. However recent improvements in technology has produced circular SiPM with a fill factor of 100%, for a pixel with a radius of 10 μm [34].

The Low Gain Avalanche Detector

A slight variation of the APD's structure leads to a device known as a Low Gain Avalanche Detector (LGAD) with a gain in the region of 10-20. The LGAD design is based on the modification of the doping profile, that is the control of the implants at the n$^+$/p junction.

The resultant doping profile, which has a large increase in doping concentration close to the junction, creates a large electric field when in reverse bias. The electric field profile is comparable to that of an APD, figure 2.7, where the electric field is more stringently controlled in the multiplication region. The implants need to be closely controlled in order to allow operation at a high voltage without breakdown, further details are described in chapter 3. The lower gain enables this technology to detect high energy charged particles and has been shown to work when finely segmented [35]. This allows the LGAD to be fabricated in pixel arrays and microstrip devices. The LGAD design enables the detection of sub–ns signals produced by minimum ionising particles, where the ionisation is produced uniformly as a function of depth in the detector. For an LGAD with a gain of 10, if the LGAD substrate is thinned by an order of magnitude compared to a standard PIN silicon detector the same signal will be produced from a minimum ionising particle by the two devices. The collection time will however be increased by an order of magnitude. The LGAD structure produces a fast rise time and fast silicon timing detectors with sub–ns rise times are therefore possible.

LGAD's invented at CNM [36] have set the benchmark for LGAD fabrication. Since then a number of groups worldwide including HPK, FBK [37], BNL [38] and Micron Semiconductor [39] have fabricated LGAD's for high energy physics applications. The research has been performed in the hope to use the LGAD technology for a ultra–fast silicon detector (UFSD). In order to achieve this the optimal parameters for such a device require the detector to be thin, ≈ 50 μm and have a gain of 20 or above.

The detectors have been tested at beam tests and have be shown to have a time resolution of 30 ps. Radiation hardness studies have shown the possibilities to use UFSD up to fluences of ≈ 1 x 10^{15} n_{eq} cm^{-2} [40]. At this time pads and small arrays of LGAD's are used to measure time resolution and radiation hardness information. Pads tested range from 1x1mm^2 – 5x5mm^2 in size, arranged inarrays of 2x2 and 5x5. It is the hope that the UFSD will be installed in the ATLAS upgrade as a 4-D tracking detector, which will be highly segmented and have a time resolution of ≈ 30 ps [41].

As well as high energy particle physics, LGAD's have been suggested for Synchrotron applications. The aim is to have a detector which has a small gain, sufficient enough to increase the collected charge from a low energy incident x-ray (< 1 keV) above the noise floor of the readout system.

Micron Semiconductor in collaboration with the University of Glasgow have pursued this challenge by creating a range of LGAD devices. The ultimate goal is to produce a hybrid pixel detector of highly segmented detectors with a pixel pitch of 55 μm of be readout using a Medipix/Timepix readout chip. The devices fabricated and tested will be discussed in this thesis, along with simulation results which have aided this process.

2.6   Semiconductor Fabrication

Understanding the semiconductor fabrication techniques used in the manufacturing of silicon detectors is essential when one is planning to design a new type of detector. The basic principles of fabrication will be described in this section with the main focus on the steps which can be controlled in the simulation software (Sentaurus) [12] which will be discussed in Chapter 3. There are a few semiconductor materials which can be used for physics experiments including CdT, CZT, GaAs and Si. For the purposes of this section only Silicon fabrication will be discussed as this is the substrate chosen for this technology. The reason this is chosen is due to basic demands of particle physics. There is the need for high resistivity substrates in order to achieve full depletion of a 200 to 300 μm thick detector with a reasonably low external bias voltage of below 300 V. Other demands require a low detector noise, which can be enabled by having very low bulk generation current resulting from the use of very pure, high lifetime material. Planar technology is used to manufacture silicon detectors with extremely low leakage currents (low noise). All fabrication is performed using 6 inch wafers of high resistivity silicon (2-10 kΩcm).

2.6.1   Thermal Processes

The capability of silicon to withstand high temperature processing gives it an advantage over other semiconductor substrates. Silicon wafer processing involves many high temperature (700 to 1200 ºC) procedures.

Oxidation

The native oxide of silicon, silicon dioxide (SiO_2), is a very stable and strong dielectric material that is routinely formed in high temperature processing. This oxide forms when bare silicon is exposed to the atmosphere, where it reacts with oxygen to form a thin layer (about 10 to 20 Å) of (SiO_2). This layer of native oxide grows until thick enough to stop further oxidation of the silicon at room temperature.

The fabrication process always begins with a very thorough cleaning of the wafers, which would usually include the use of Hydrofluoric acid (HF). HF is used as an etchant of silicon dioxide (SiO_2) in order to ensure removal of the smallest particulates. Other cleaning solutions include Hydrogen Peroxide, Sulphuric Acid, Ammonia and Hydrochloric Acid. Following cleaning, the wafer is ready for oxidation. Oxidation is performed in a high temperature furnace. The furnace used at Micron Semiconductor is a horizontal furnace, determined by the way the quartz tubes and heating elements are placed in the system. Many parts used in the furnace are made of fused quartz as it is a very stable material up to 1200 ºC. The wafers are placed in a jig and pushed slowly into the furnace for thermal processing. Normally the furnace runs at 500 ºC, once the wafers are placed inside the

furnace temperature is increased at a rate of about 10 ºC per minute. Oxidation is the process by which a layer of SiO_2 is grown thermally, in this case at a temperature of 1030 ºC. The furnace is flooded with oxygen which reacts with the bare silicon. Unlike the native oxide, this oxide can be grown to a thickness > 5000 Å. This is due to the oxygen molecules having a higher thermal energy at higher temperatures, meaning they move a lot faster. This means the oxygen atoms can diffuse across an existing oxide layer and react with silicon to form more SiO_2. The oxide is produced at the junction between the silicon and the oxide. Increasing the temperature will increase the rate at which the oxide is grown. The oxidation process takes several hours thus the need for batch processing of wafers. Oxidation can be performed by a *wet* or *dry* oxidation. Dry oxidation is performed in oxygen and can be expressed as

$$Si + O_2 \rightarrow SiO_2 \tag{2.25}$$

While wet oxidation is performed in water vapour and can be expressed as

$$Si + 2H_2O \rightarrow SiO_2 \tag{2.26}$$

During wet oxidation the flow rate of oxygen and hydrogen must be controlled in order to produce a high quality film. In general wet oxidation is used to grow the majority of the oxide as it is a faster process. This faster process is because at high temperatures H_2O can dissociate and form hydroxide (HO), which can diffuse in SiO_2 faster than O_2 can. However dry oxidation produces a higher quality oxide, so generally both methods are used. Small percentages of HCL are added to getter impurities and therefore improve the bulk properties (higher minority carrier life times).

There are many applications for the silicon dioxide once grown. It can be used as a mask for ion implantation, as most dopant atoms used in the semiconductor industry have a much lower diffusion rate in SiO_2 than they do in single–crystal silicon. Therefore, by etching windows in the masking oxide layer one can create isolated regions for ion implantation. For this process an oxide thickness of around 9000 Å was used. Another application of oxidation is the use of a screen oxide. A screen oxide is generally grown using dry oxidation, where a thin layer (100–200 Å) of high quality oxide is produced. This is grown for two main purposes. Ion implantation will be performed through this screen oxide where the ions scatter before entering the silicon substrate, this helps to minimize the channelling effect. The channelling effect is the process by which implanted ions penetrate deeply into the substrate by travelling through the gaps in the crystal structure between atoms. This can occur during the implantation step when the surface of the silicon substrate is orthogonal to the beam. This screen oxide also helps to minimise damage or contamination to the silicon by blocking any contaminants present in the ion chamber.

Thermal Annealing

Thermal annealing is a heating process in which the wafer undergoes physical or chemical change. During this process two different thermal annealing processes are used to achieve the required outcome.

Post–implantation annealing is performed as during an ion implantation process the high energy dopant ions can cause severe damage to the silicon crystal structure near the wafer surface. In order to meet device requirements the lattice structure must be repaired. This is done by thermal annealing to restore the lattice structure and activate the dopant ions. Post–implantation the dopant ions cause the lattice structure to be deformed. Once placed in a high temperature environment the atoms move quickly and rest at positions that have the lowest free energy, located at single–crystal lattices, hence restoring the lattice structure. The wafers are annealed in a furnace at around 900 °C for about 30 minutes in a nitrogen environment.

As the lattice is annealed the dopant ions can undergo diffusion into the lattice. The longer the annealing process, the deeper the diffusion of ions. This is also true if the temperature were to increase. For the multiplication implant stage of this process it is imperative that the dopant ions are diffused deeply into the substrate. Therefore a higher temperature is used (> 1000 °C) for several hours to achieve the required depth.

2.6.2   Photolithography

Photolithography is a patterning process that transfers a designed pattern from a mask to the photoresist on a wafers surface. This is one of the most important process steps during the fabrication as it determines the regions for various etching processes and therefore ion implantation. The photolithography process includes three major steps: photoresist coating, exposure and development. A photoresist is a photosensitive material used as a temporary coating on a wafer whereby a pattern can be produced. The photoresists used in semiconductor processing are sensitive to only UV light. There are two types of photoresist used, positive and negative. For the positive photoresists, the regions which are exposed to UV light become an aqueous–base–soluble material [42] in a photochemical reaction called photosolubilization. This can then be dissolved by the developer, leaving the unexposed parts remaining on the wafer surface. As for the negative resist, the regions which are exposed become cross–linked and polymerized due to a chemical reaction. They harden and remain on the wafer, while the unexposed parts can be dissolved by the developer. The positive Microposit S1811 photoresist was used for all lithography steps in this process. Figure 2.9 shows the two different kinds of photoresists and their pattern transfer mech–anism. The image produced on the positive photoresist is the same as that on the mask, and the image produced on the negative photoresist is the reversal of that on the mask.
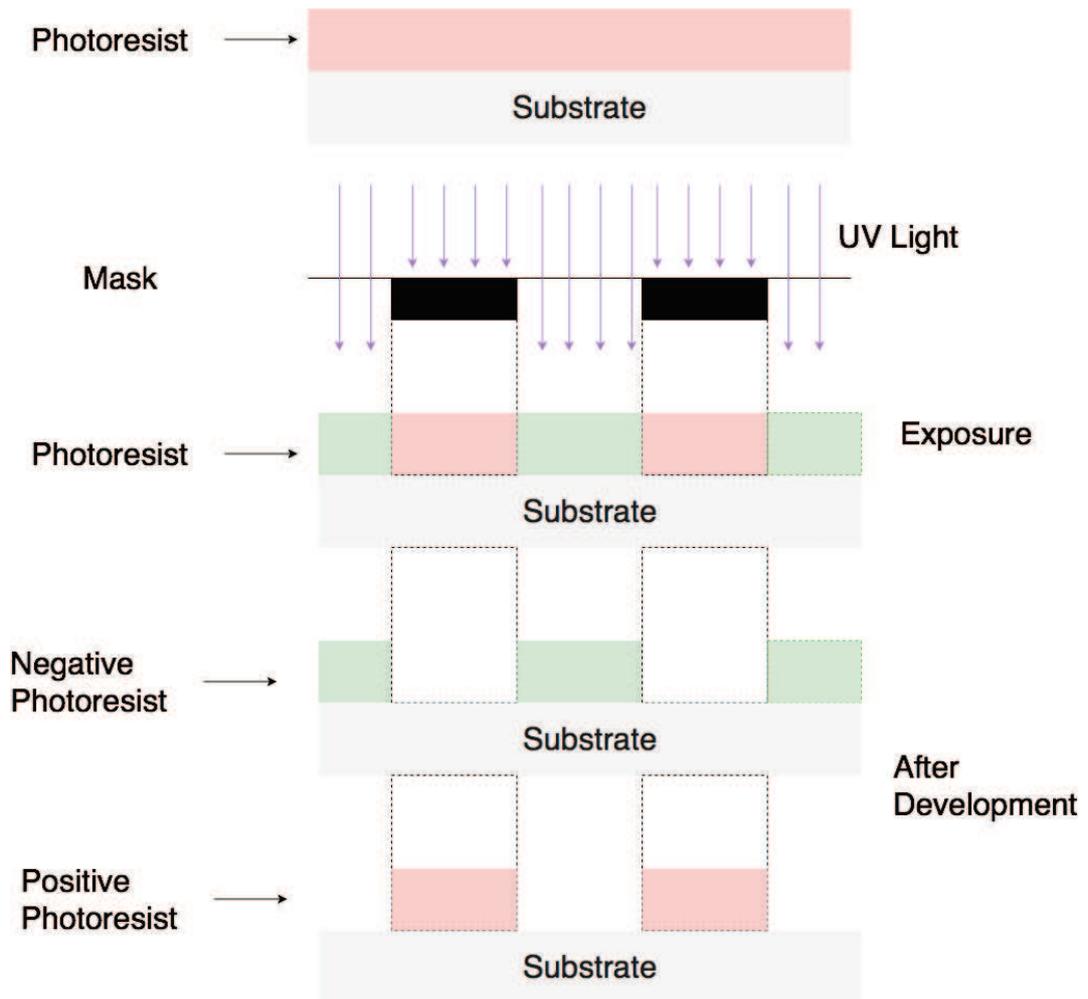
Figure 2.9: Patterning process with positive and negative photoresists.

Etching

Following a photolithography step, etching usually occurs. This is the process by which the oxide grown in the oxidation process is removed. The oxide will only be removed where there is no photoresist protecting it. This is a patterned etch, as discussed previously blanket etching can be used during the cleaning stages of the process. The wafers in this process undergo a wet etch, where the wafers are plunged in an etching solution where the time is precisely controlled. This is because the oxide reacts with the etching solution at a specific rate, depending on the concentration of the etchant, the etchant used and the temperature. Etchants are chosen to have an anisotropic etch rate, meaning the oxide is only etched in one direction. We want a high vertical etch rate compared with the etch rate in the lateral direction. The ideal case is where we have vertical profile where the oxide is masked by the photoresist. However in reality with a wet etch an isotropic etch occurs. The difference in etching is shown in figure 2.10. Other forms of etching are available which produce an anisotropic etch, important for small feature size. For our process this is

(a) Anisotropic Etch.
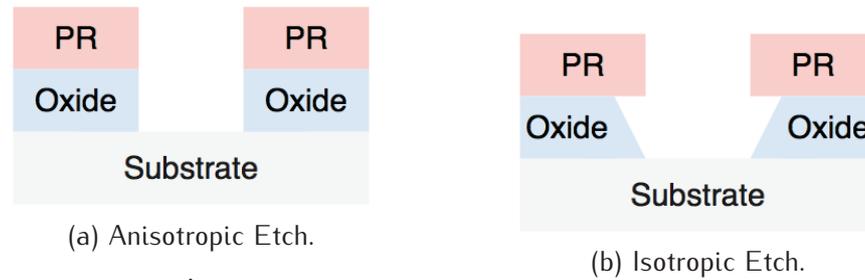
(b) Isotropic Etch.

Figure 2.10: Example of an isotropic and an anisotropic etch.

not necessary. It is also important to understand the selectivity of the etchant used, where the selectivity is the etch rate in one material compared with another. The etch rate must be much greater in the oxide than the photoresist.

2.6.3   Ion Implantation

A essential tool in any semiconductor fabrication process is an ion implanter. Ion implantation is necessary in order to control the conductivity of the semiconductor material by adding dopants. Silicon is doped with either n- or p-type dopants, in our process phosphorus and boron are used respectively. Ion implantation is a process by which dopant atoms are added into a semiconductor substrate by injection of an energetic ion beam. Unlike diffusion, ion implantation has far more control of the doping profile. The doping concentration and junction depth can be closely controlled. The doping concentration is controlled by the implantation time and ion beam current. The doping depth is controlled by the ion beam energy, the depth is the value at which the dopant concentration is greatest. Particularly for low gain avalanche detectors (LGAD) the control of the depth of the dopants is crucial in producing the correct gain value which is discussed further in chapter 3. As well as the uniformity of the ion implantation any small change in doping concentration or junction depth can have significant consequences for the device performance. A basic diagram of the process flow for ion implantation is shown in figure 2.11.

2.6.4   Metalisation

Metallisation is an additive process that deposits metal layers on a wafer surface. This is performed by sputtering in our process. Sputtering is where a wafer is placed inside a low pressure chamber, a metal target is bombarded with high energy ions (usually argon) dislodging atoms or molecules from the metal surface which are redeposited on the substrate surface to form a thin metal film ($\approx$ 1000 Å).

(a) Oxidation.

(b) Imaging of photoresist.

(c) Etching

(d) Ion Implantation

(e) Post Annealing
.
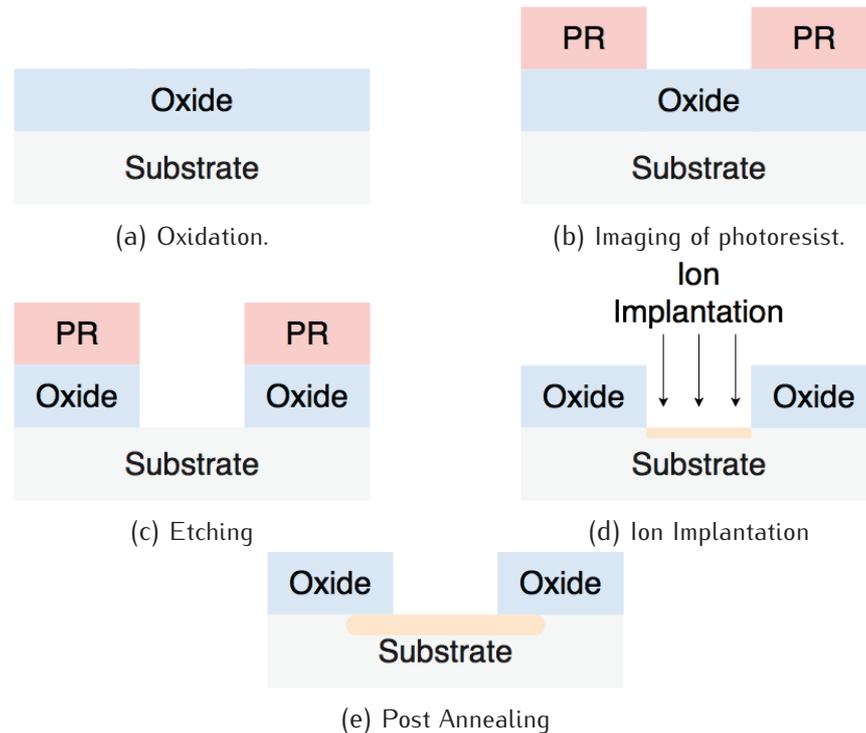
Figure 2.11: Process flow of a basic Ion implantation process

Lift–off

In order to make use of this, the wafer is first coated with photoresist and imaged as required. The wafer is then coated in metal, where some areas of the wafer will be coated in photoresist and metal and other areas will only have metal depending on how the wafer was imaged. The wafer is placed in a developer which attacks the photoresist. This lifts off all metal from the wafer which has photoresist between it and the silicon substrate leaving only the metal directly sputtered onto the wafer surface. This process is extremely important for creating appropriate contacts for applying a bias voltage. Metal designs can be produced which are very useful for device testing as shown in chapter 4 while using the transient current technique. Metal contacts are needed for making electrical contact to the device (wirebond pads and pixel interconnect pads) and reducing series resistance in long structures like strips. The metal used for all devices produced was aluminium. Figure 2.12 shows the basic fabrication process flow.

Etching

As well as lift–off, an etching technique can be used for metallisation. The metal is deposited on the surface of the substrate, where it is then coated with a photoresist. This photoresist is imaged and developed. The exposed metal is then etched away, leaving a patterned metal layer. This etching method is preferred for our process as this technique is better

(a) Imaging of photoresist prior to met‐allisation.

(b) Metalisation.
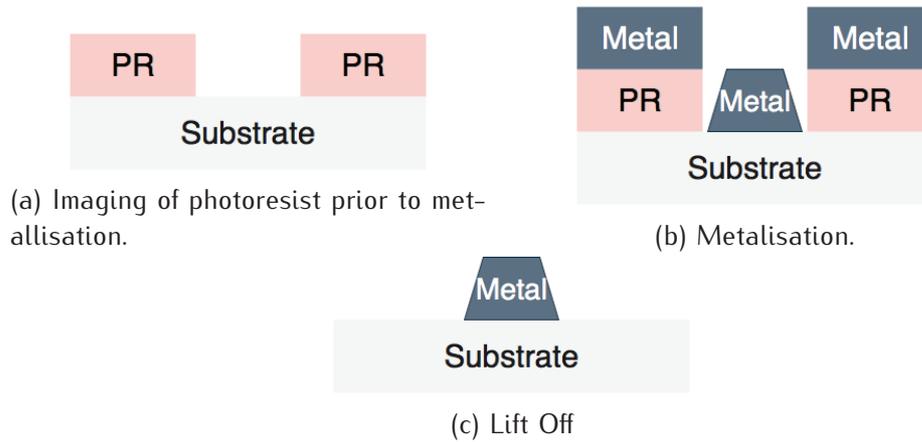
(c) Lift Off

Figure 2.12: Basic process flow for metallisation using lift off.

(a) Uniform deposition of metal.

(b) Photolithography.

(c) Metal etch and removal of photore‐sist
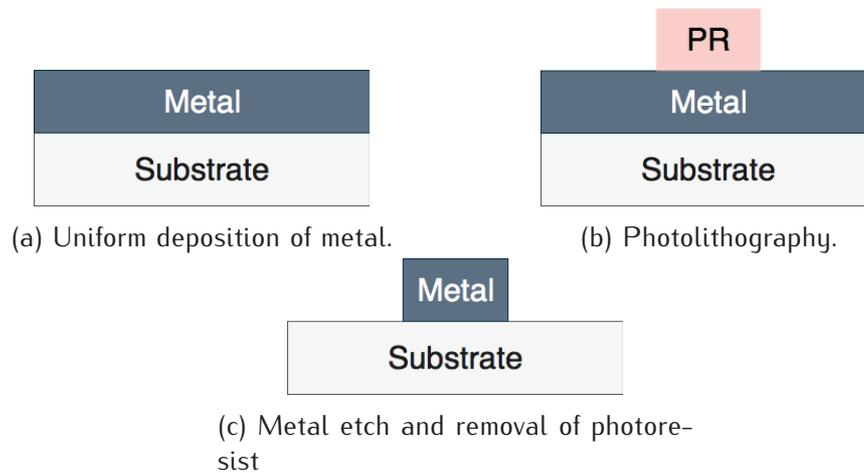
Figure 2.13: Basic process flow for metallisation using metal etching.

established at Micron Semiconductor. The process is shown in figure 2.13.

Following Metalisation one may require the use of a passivation layer. Silicon dioxide is deposited onto the surface of the wafer to protect the device from moisture and other contaminants.

2.6.5 LGAD fabrication process flow

For a basic p–n junction diode used in many applications within high energy physics the basic process steps are outlined below on a p–type substrate. Figure 2.14 shows the basic steps required to produce a working pad diode. Something not mentioned thus far is the use of a guard ring structure. Guards rings are used to reduce the field to the edge of the device in order to ensure the device does not undergo early breakdown [43]. Different approaches can be taken to produce guard rings, however this has not been pursued during this work. The standard guard ring structure used by Micron Semiconductor on the majority of their devices has been implemented in these devices as it has shown to produce devices

with a breakdown voltage of greater than 1000V. Extra steps are required when producing pixellated devices. These include the use of a p–stop or p–spray to isolate pixels. [44]. A p–stop is an isolated p–type implant (boron) between pixels, and a p–spray is a uniform p–type implant across the wafer. A p–stop requires an additional lithography step, therefore a dedicated mask must be produced. However, as p–spray is a uniform implant, no additional mask is needed. This is the simpler method and the one preferred for our process.

The process steps for a typical PIN diode are outlined below and shown in figure 2.14.

- Preliminary cleaning

- Oxidation at above 1000 ºC for $\approx$ 5 hours, to produce a field oxide with thickness $d_{ox} \approx 1 \ \mu m$.

- Formation of the front side ohmic n^+ contact:

    - Photolithography of n^+ region, this will form the active area and guard rings.

    - Oxide etch of the front side to produce windows for ion implantation and removal of the photoresist. This etch also removes all oxide from backside of wafer.

    - Ion implantation of phosphorus ions uniformly across the front of the wafer.

- Formation of the p^+ contact on the backside

    - Ion implantation of boron ions uniformly on the backside of the wafer.

- Further cleaning before the annealing stage.

- Wafer is annealed, activating the implanted dopants.

- Front and back side of the wafer is sputtered with $\approx$ 100 nm of aluminium.

- Photolithography, and aluminium etch on both front and back side of wafer.

    - It may not be necessary to image and etch the back side aluminium if a uniform metal covering is required.

When producing LGAD's a variation of this process is used as there is an additional implant stage. This additional implant stage comes before the junction implants and is responsible for the gain produced when the diode is run in reverse bias, explained further in chapter 3. The process steps for a typical LGAD device are outlined below and shown in figure 2.15.

- Preliminary cleaning

- Oxidation at above 1000 ºC for $\approx$ 3 hours, to produce an oxide with thickness $d_{ox} \approx$ 1 μm.

(a) P-type substrate

(b) Oxidation

(c) Photolithography

(d) Opening of Windows

(e) Ion implantation & annealing of junction implants

(f) Metalisation
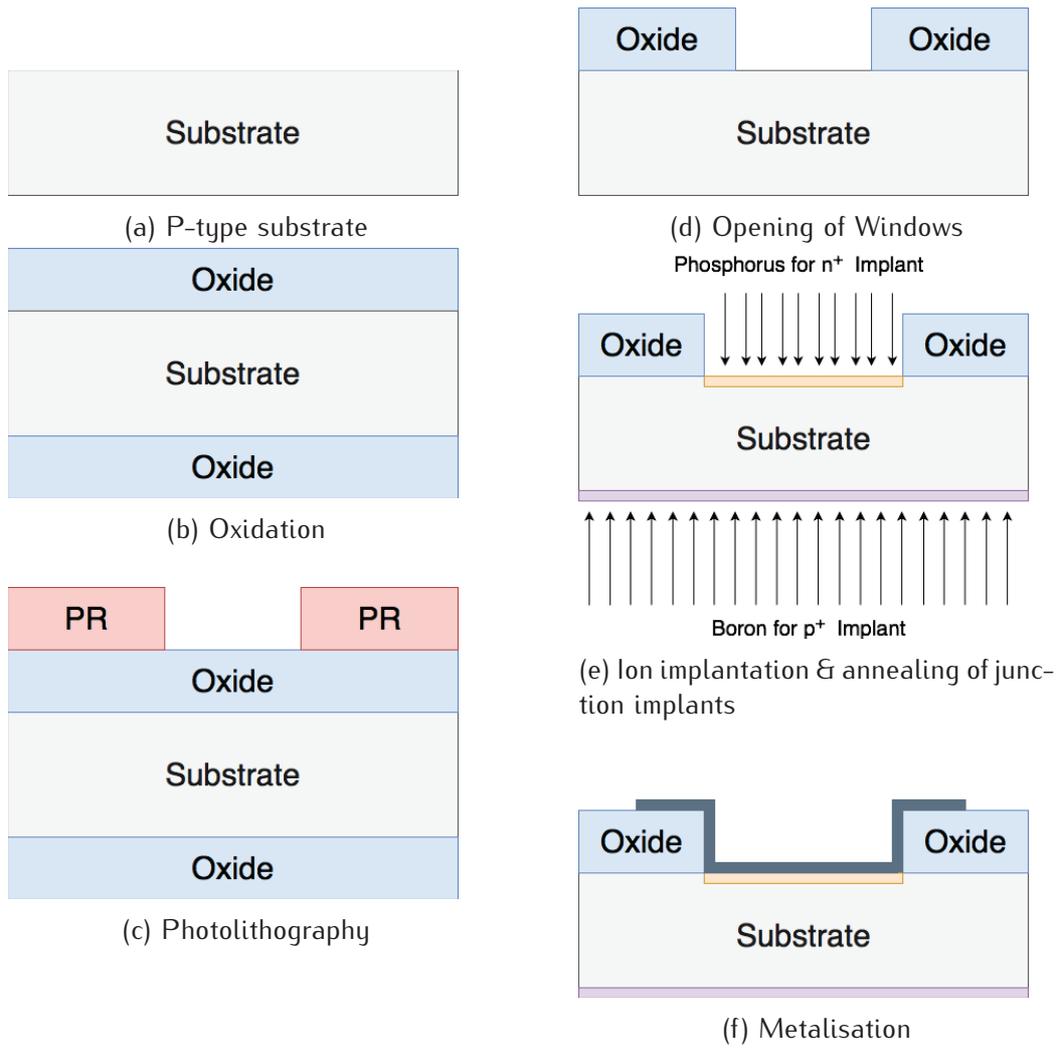
Figure 2.14: Basic process flow for the fabrication of a PIN diode.

- Integration of the Multiplication implant on front side of wafer.

  - Photolithography of the p^+ region, responsible for gain.
  - Oxide etch of the front side and removal of the photoresist. This etch also removes all oxide from backside of wafer.
  - Ion implantation of boron ions uniformly across the front of the wafer.
  - Oxide is stripped from wafer after ion implantation and thoroughly cleaned.
  - Wafer is then placed in a furnace for deep diffusion of boron ions for several hours at > 1000 ºC.

- A uniform p-spray is performed on the wafer and thoroughly cleaned.

- Oxidation at above 1000 ºC for ≈ 5 hours, to produce the field oxide with thickness $d_{ox} \approx 1$ μm.

- Formation of front side ohmic n^+ contact:

  - Photolithography of n^+ region, to form the active area and guard rings.
  - Oxide etch of the front side and removal of the photoresist. This etch also removes all oxide from backside of wafer.
  - Ion implantation of phosphorus ions uniformly across the front of the wafer.

- Formation of the p^+ contact on the backside

  - Ion implantation of boron ions uniformly on the backside of the wafer.

- Further cleaning before the annealing stage.

- Wafer is annealed, activating the implanted dopants.

- Front and back side of wafer is sputtered with ≈ 100 nm of aluminium.

- Photolithography, and aluminium etch on both front and back side of wafer.

  - It may not be necessary to image and etch the back side aluminium if a uniform metal covering is required.

(a) P-type substrate

(b) Oxidation

(c) Photolithography

(d) Opening of Windows

(e) Ion implantation & deep diffusion of multiplication implant

(f) P-type spray

(g) Second oxidation

(h) Opening of windows by photolithography

(i) Ion implantation & annealing of junction implants

(j) Metalisation
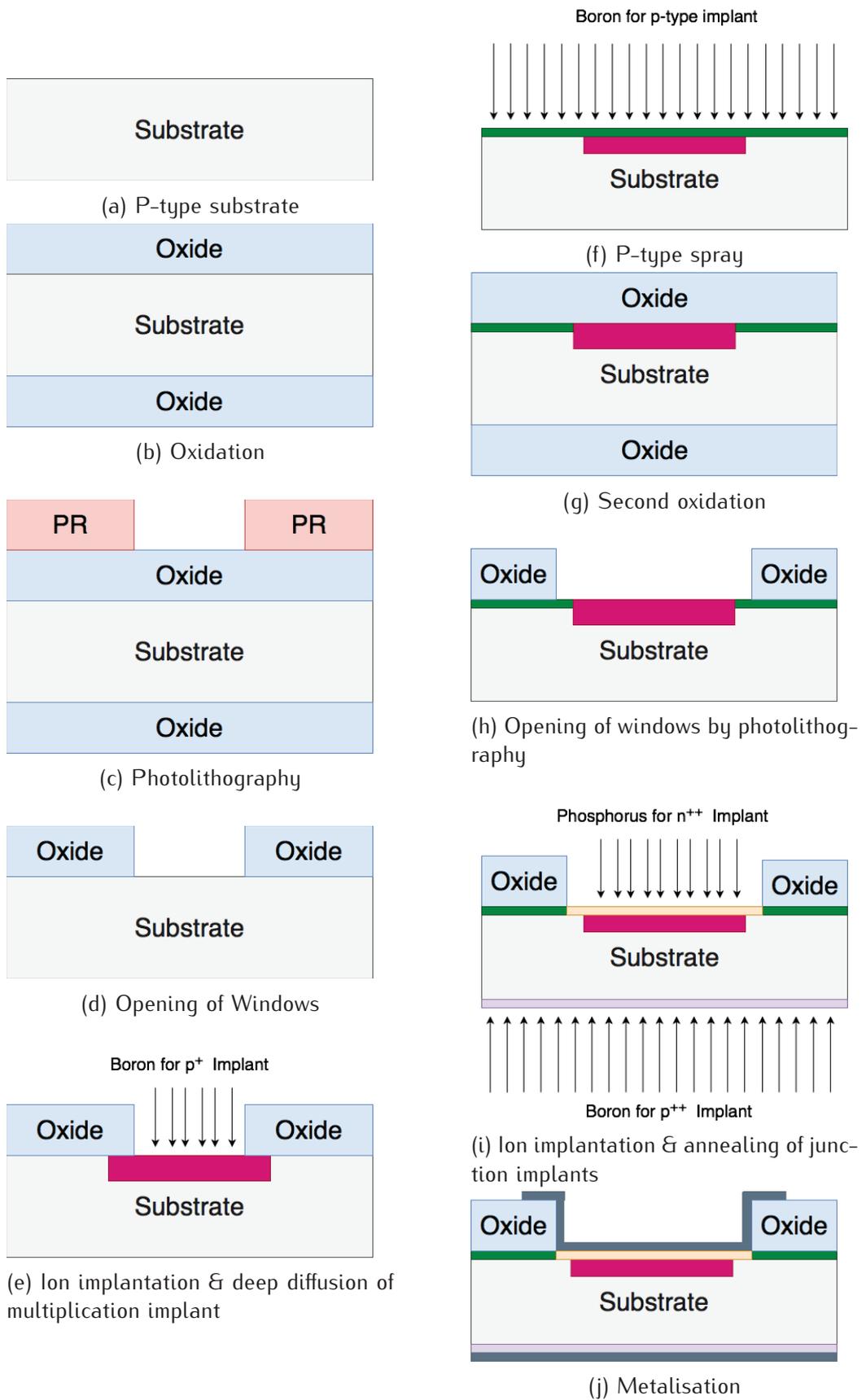
Figure 2.15: Basic process flow for the fabrication of an LGAD device.

Chapter 3

Simulations of Low Gain Avalanche Detectors

When developing a new detector technology one has to consider many factors including time, complexity and of course cost. Thus it's important for new technology to have computer aided programs to evaluate and optimise the design of a detector before fabrication. In a typical fabrication run only a few designs and process variations can be used given the cost and production timescales. Simulations can be extensively performed to define the final fabrication parameters needed in order to create a detector with the desired characteristics. Simulations can also be helpful for understanding the behaviour of a device.

This chapter gives an overview of the simulation package "Synopsys TCAD" [12] used to perform all of the simulated results. The principles of semiconductor simulation are described in sections 3.1.1 to 3.1.3. Where section 3.1.1 describes the implementation of the fabrication steps to produce the resultant doping profiles and the models used. This is known as process simulation. Section 3.1.3 describes the electrical simulations performed on the structures built in the process simulation, known as device simulation. Device simulation can be used to look at electric field profiles and depletion regions. Using transient simulation of charge the charge collection efficiencies can be investigated. A more complete description of the physics models related to both process and device simulation can be found in Appendix A and B respectively.

The LGAD structure will be looked at in detail describing how simulations were performed and validated in section 3.2. The simulation process involved using 1D and 2D structures to first define a suitable multiplication region doping profile, and then to determine an edge design which is capable of withstanding the high fields produced in this multiplication region.

A validation technique known as Secondary Ion Mass Spectroscopy (SIMS) [45] is used to evaluate the dopant profile of a fabricated device. Simulated and fabricated devices are compared and the results discussed in section 3.2.2.

3.1   Introduction to TCAD (Sentaurus)

Simulation packages in semiconductor device development are a commercial enterprise, researched for use with CMOS processing in evaluating electrical field calculations and Monte Carlo ion implantation. More recently they have been used by research facilities to help develop sensors for high energy physics applications. Current software is available in the form of Technology Computer Aided Design (TCAD) packages. SYNOPSYS Sentaurus [12] is the tool of choice for this research as the group already had good experience working with this tool. Synopsys manages to group Monte Carlo simulation, finite-element solutions algorithms, electrical field calculations and geometry profiles into one package. This allows for full 3D simulation of device processing and electrical behaviour making it possible to study carrier propagation through the detector and charge collection efficiency. However, due to the complex nature of the structure and the time required for simulation, 2D simulations were performed throughout this work. A 2D simulation is when the structure contains only information in the x and y axis. Where the y axis is along the surface of these devices and the x axis corresponds to the depth within the device. This can provide suitable information for creating diodes and simulating the movement of charge. However, effects seen only at the corners of devices cannot be modelled and thus can prove difficult when designing features at the edge of a junction.

In order to create the device structure and calculate electrical fields two simulation packages were used, Process simulation 3.1.1 and Device simulation 3.1.3.

3.1.1   Process Simulation

Sentaurus Process is the package used for process simulation. It uses a combination of diffusion equations, see Chapter 2 section 2.3.3, Monte Carlo Algorithms and extensive calibrated data libraries to provide predictive capabilities for silicon processing. The fabrication of a silicon detector has many technological process steps, where the main aim is to provide appropriate doping of the silicon substrate. Therefore it is extremely important to have extensive knowledge of the processing required. The use of photolithography, thermal diffusion and chemical reactions are all used within the Sentaurus package. The fabrication process is described in section 2.6 where the main steps are:

- Field oxidation of the substrate for the development of a mask layer to be used for doping purposes.

- Photolithography and etching to open up the appropriate oxide windows.

- Ion implantation to introduce dopant ions into the exposed silicon surface.

- Thermal diffusion for implant activation through high temperature exposure.

These processes are repeated as required for the final device structure.

Ion Implantation

Ion implantation is one of the most widely used and thus a very important processing step. It is this stage which the impurity atoms are introduced into the silicon substrate and are a major contribution to the final electrical performance of the device. Sentaurus Process uses a Monte Carlo (MC) method to compute the distribution of implanted ions and map the implantation damage. In order to calculate the penetration depth of implanted ions the MC method uses a statistical approach. This can be used to calculate the crystal damage and is based on the binary collision approximation [46]

The Monte Carlo engine used for this work is the Crystal-TRIM algorithm, which is based on the Transport of Ions in Matter (TRIM) code [46]. Monte Carlo implantation simulation has been developed for ion implantation into single–crystal materials and thus is ideal for simulation with silicon. Required for the simulation are explicit processing parameters, these include the implantation energy and dose as well as the orientation of the ion beam with respect to the wafer crystal axes.

During the implantation pseudoparticles are used to represent a fraction of the total dose required and are started from the surface of the substrate. The number of particles can be set however the default value of 1000 was used. More particles lead to more accurate results but use up large amounts of computing power, where less particles would result in less computing time but less accurate results. The particles are set on a trajectory determined by the initial conditions set. The final number of trajectories can be increased by allowing trajectory splitting. This artificially increases the number of trajectories allowed in regions of low trajectory density. If a pseudoparticle reaches an element with a low trajectory density, a split point is set, that means the initial particle is divided into two sub particles, which have half the statistical weight of the initial particle. These newly generated sub particles start under the same conditions as the initial particle. Trajectory splitting decreases the statistical noise and provides better modelling for the transition regions. In order to reduce the simulation time trajectory replication, by default, is used. This uses the fact that several regions in 2D structures have 1D topology, meaning simply copying the particles trajectory is possible and can provide accurate results. This method is much faster than calculating all trajectories.

As particles interact in the substrate they lose part of their energy due to elastic nu-clear scattering at target atoms and inelastic electronic iterations. Nuclear scattering is calculated using a Coulomb–screened pair potential (ZBL potential [47]). It is assumed the particles come to rest if their energy is < 15 eV, but above this value if the particle transfers > 15 eV to the target atom it can become displaced. By default only implanted ions are simulated and can cause interstitial formation in the silicon lattice. If the cascade–type

simulation is chosen the displaced atoms can in turn generate more displaced atoms, providing they have the required energy. This method provides the only way to correctly model the physical profiles of displaced atoms and thus is chosen for all simulations performed. Inelastic collisions are modelled using semi–empirical models to primarily investigate the channelling effects.

Diffusion Models

Following an implantation, a diffusion step is required to activate the implanted dopants and is used to drive implanted dopants deeper into the substrate. This activation process is important in the restructuring of the crystal lattice and is described using seven models in the Synopsis package:

- ChargedReact

- React

- ChargedPair

- Pair

- ChargeFermi

- Fermi

- Constant

Transport Models compute the particle flux, **J**, of dopant ions and are the main diffusion models. Pairing reactions can also be computed depending on the model chosen and the need. The models are described in detail in the Sentaurus Process User Guide [12], [48], therefore only the main idea for each model will be described here.

The *React* and *ChargedReact* diffusion models are the most advanced models used in Sentaurus Process and as such are the most complex. They solve up to three equations per dopant, a substitutional dopant, and up to two dopant–defect pairs and two sets of defect equations. They are also known as five–stream diffusion models. The *React* model is an uncharged version of the *ChargedReact* model, meaning the reaction rates are not charge state–dependent. The *ChargedReact* model is the most computationally expensive model.

The *Pair* and *ChargedPair* diffusion models uses the assumption that the dopant–defect pairs are in local equilibrium with the dopant and defect concentration. They also solve separate point–defect equations. Unlike the *React* models they solve one equation per dopant and two defect equations and hence are referred to as three–stream diffusion models. As with the *ChargedReact* model the *ChargedPair* model is charge dependent, allowing the pair coefficients to vary with charge state. These models are a good compromise for

accuracy and time but are not suitable for conditions where the ramp rates in the annealing process's are fast.

The *Fermi* and *ChargedFermi* models both work on the assumption that point defects and dopant–defect pairs are in equilibrium. The *ChargedFermi* model enables the diffusivity of every charge state to be set separately. These models are appropriate for long term diffusion's at high temperature.

The *Constant* diffusion model is used mainly for dopant diffusion in oxide as it assumes a constant diffusivity and no electric-field effect. The diffusion of ions is described by:

$$\frac{\partial C_A}{\partial t} = \Delta \bullet (D_{\text{star}} \Delta C_A{}^+) \tag{3.1}$$

where D_{star} is the intrinsic diffusivity of the impurity A and $C_A{}^+$ is the active portion of C_A, where C_A is the dopant. The diffusivity follows the Arrhenius law.

For all simulations performed the *React* model was used as the doping profiles are an essential part of the design and cannot be comprised for speed. For diffusion of the ions in silicon where there is an oxide, a doping dependent reaction is present and this is accounted for in the simulation and explained in section 3.1.1.

Oxidation

As described in Chapter 2, the process by which a masking layer is produced is through oxidation. This is achieved in Sentaurus Process by implementing an annealing step in the process. This annealing step is similar to the diffusion step, but is specified to be performed in an oxygen–rich atmosphere. This can be in either a *wet* or *dry* oxidation step by specifying either H_2O or O_2 in the *diffuse* command. The thickness of the oxide depends on the type of oxidation being performed and the temperature of oxidation, this makes it imperative to know the fabrication process in great detail.

If growing an oxide and there are impurity dopants in the silicon this can greatly effect the profile and concentration of these dopants. This is particularly important for the multiplication implant. This is due to the segregation of Boron in silicon oxide. When an oxide is grown on the silicon surface the boron is diffused further into the silicon substrate than into the oxide. The value for the diffusion coefficient of boron in silicon is around 10000 times greater than that of the diffusion coefficient in silicon dioxide [49]. This is very useful as the doping profile of boron for this technology is required to be very deep on the front side.

3.1.2 Meshing Strategy

With a finite element approach, calculations are performed in order to simulate the physical parameters of a device. In TCAD, semiconductor equations describing the movement of

charge via drift and diffusion process, described in chapter 2, section 2.3.3, as well as are solved at discrete nodes within a detector volume. The detector volume is represented by a network of these nodes and is referred to as the mesh.

In order to represent the detector effectively, the basic geometry of the device must be set with finite boundaries. To describe the overall detector each node of the mesh must contain three important variables. These are the electrostatic potential and the electron and hole concentrations, as all other variables can be calculated using these values. As each node is a finite distance from another node the region between them is split up in a series of elements. The value of the three variables mentioned is defined in each element as a linear interpolation of the surrounding nodes.

For accurate results one must produce a mesh, which is sufficiently small to be able to apply this linear interpolation. This is done by controlling the size of each element, thus the distance between each node is vital. For regions where the physical variables change abruptly, a finer mesh is required. However, if an extremely fine mesh is used for the full detector volume the computation time can be massive. If too coarse a mesh is used the calculations will not converge and the simulation will fail. Thus it is important to have a meshing strategy which is efficient in regions of minimal variations in physical parameters and still fine enough in regions of extensive variation. The number of nodes in the mesh is calculated and returned to the user after a re-mesh is performed.

When simulating the fabrication of a device, as in Sentaurus Process, it is necessary to redefine the mesh at various stages of the process as the device structure can change. Sentaurus Process has the ability to automatically generate meshes as the device structure changes. For all 2D simulations performed in this work the mesh was redefined after every geometrical operation. This includes etching, deposition and transformation operations. Transformation operations are used to copy or extend regions of the device which have a repeating pattern. The MGOALS library mesh engine was used to generate the mesh for all devices fabricated. Sentaurus Process has two meshing engines, the other being Sentaurus Mesh which is used for 3D simulations to reduce simulation time.

Before any process simulation is performed a general mesh is defined by the user as is the basic geometry of the device. For this device technology is was important to have fixed meshing placements. The mesh can be defined in the vertical and horizontal direction, where the vertical direction is into the silicon substrate and the horizontal direction is along the surface of the substrate. The vertical mesh is crucial for the multiplication region. The fixed meshing placements for a typical 200 μm device fabricated are described in table 3.1 where the node spacing's in the vertical direction range from 1 nm up to 10 μm. Depending on the size of the structure fabricated the horizontal node spacing's varied accordingly and had a finer mesh around the edge of any implants. Initial 1D simulations were performed to evaluate the doping profile and included a reflective boundary and as such a very

Position	Element Size
0 μm – 1 μm	1 nm – 10 nm
1 μm – 5 μm	10 nm – 400 nm
5 μm – 10 μm	400 nm – 1 μm
10 μm – 195 μm	1 μm – 10μm
195 μm – 199 μm	10 μm – 400 nm
199 μm – 200 μm	10 nm – 1 nm

Table 3.1: Fixed node spacing's for different regions of the simulated device.

narrow device of 10 μm could be simulated. In particular, for all devices simulated specific boundary conditions were enabled to ensure consistent results. The boundary conditions used were reflective. In reflective boundary condition, a reflected image with respect to the domain boundary is first constructed. Depending on the boundary condition specified on the other side, the composite structure is then either extended (if the other side is extended) or repeated (if the other side is also reflective). The added structure including the reflected image is removed after the implant is completed. As these were 2D simulations only the *Left* and *Right* boundaries had to be specified as reflective. This was done by using the commands; *pdbSet ImplantData LeftBoundary Reflect* and *pdbSet ImplantData RightBoundary Reflect* for the left and right boundaries respectively.

The meshing strategy for this is shown in figure 3.1, with a horizontal mesh of 0.5 μm. The vertical mesh is highly concentrated around the surface and the first few μm's of the structure as this is where the largest change in doping concentration is situated. The mesh should be such that a smooth doping profile is created with the minimum amount of meshing points.

To fully simulate the device and its electrical characteristics a series of device simulations are performed as described in the following section, 3.1.3. To prepare for this the mesh has to be redefined to enable an efficient simulation to be performed. The re-meshing is performed using Sentaurus Structure Editor where an adaptive mesh is introduced. Finite regions are selected at the surface of the structure where the most abrupt changes in doping concentration are found. It's essential to have a limited number of nodes for the device simulation and it's possible to keep this to a manageable level even for larger devices. The basic mesh for small devices can be used for larger devices as the mesh structure is almost identical. The bulk region of the detector can have a much larger element size due to the almost constant electric field profile in the bulk. The key areas for the mesh in an LGAD detector are the junction between the phosphorus and boron implants on the front side and the edge of these implants. These are the regions with the highest change in electric field. The mesh prepared for the 1D simulation is shown in figure 3.2.
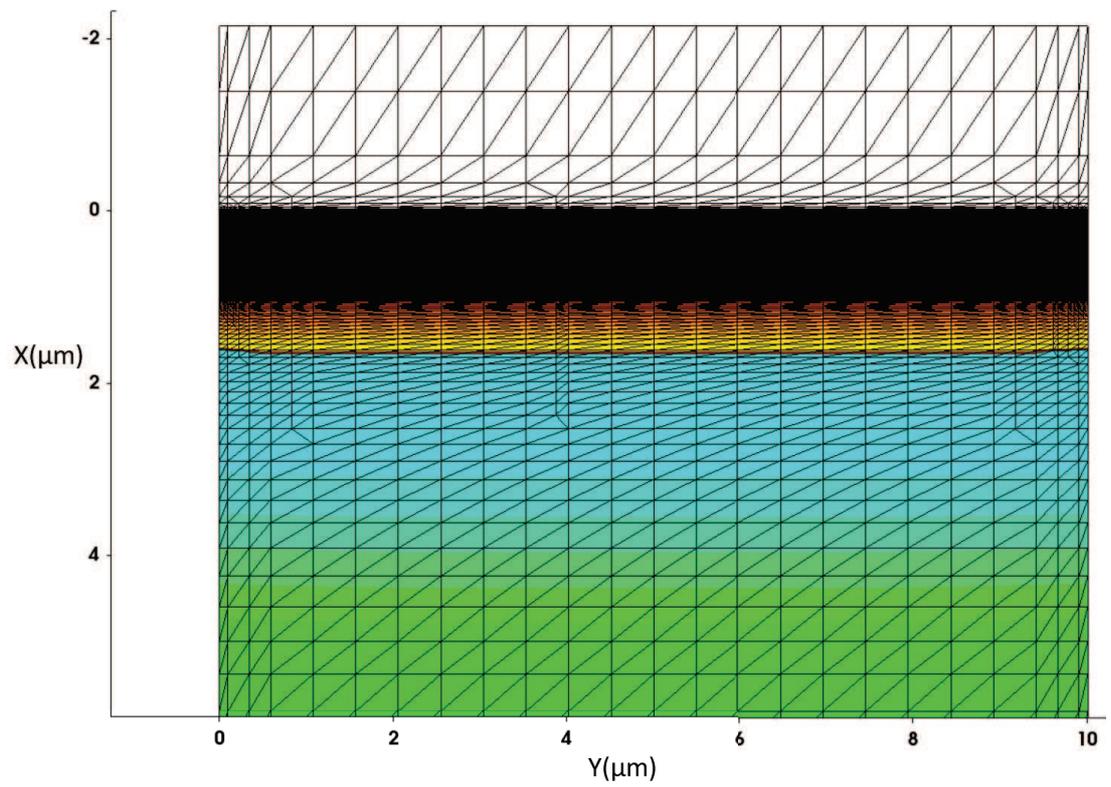
Figure 3.1: Fixed meshing strategy for 1D LGAD by using Sentaurus Process simulation. The colour highlights the doping concentration, with green having a lower concentration and red having the highest concentration.
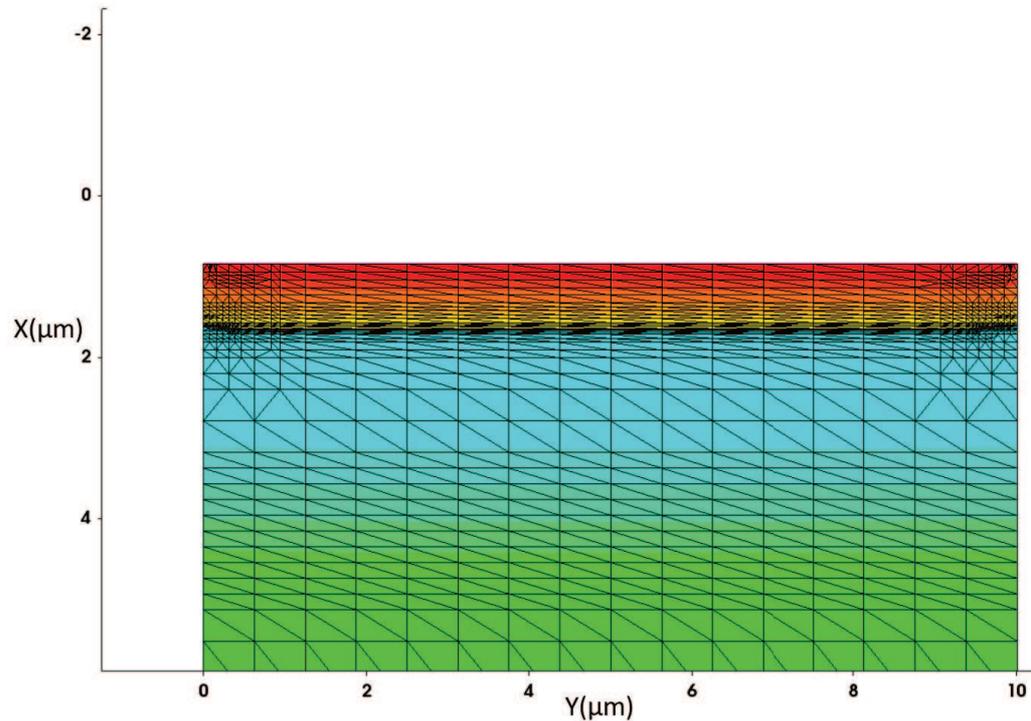
Figure 3.2: Re-meshed structure for a 1D device simulation. The number of meshing nodes is reduced compared with the sprocess structure to reduce the computational requirements.

3.1.3 Device Simulation

After creating a suitable mesh for the device, the Sentaurus Device tool can be used to run the simulation. The initial conditions as well as all physics models used are defined within a text file. Depending on the nature of the simulation the user can define the physics models required. This section will briefly describe the main models used in this work.

Physics Models

Following a full process simulation a device simulation is performed to simulate the electric field profile across the device. As this is a simulation of a photodetector, certain models have to be included.

The mobility of charge carriers is considered using doping concentration dependent and high field saturation dependent physical models. These are introduced by calling *DopingDep, eHighFieldsaturation* and *hHighFieldsaturation* respectively in the *Mobility* statement of the *Physics* section. The saturation model of electrons and holes are called independently as it may be useful to use different submodels to describe them. The default model for both is used and is an adaptation of the Canali Model [50] (based on Caughey–Thomas [51]) and uses the default driving field model GradQuasiFermi.

The saturation of charge carriers means that at higher and higher electric field they

Symbol	Parameter name	Electrons	Holes	Unit
a	a	0.426	0.243	V^{-1}
b	b	4.81×10^5	6.53×10^5	V/cm
c	c	3.05×10^{-4}	5.35×10^{-4}	K^{-1}
d	d	6.86×10^{-4}	5.67×10^{-4}	K^{-1}
γ	gamma	1	1	1
δ	delta	2	2	1

Table 3.2: Coefficients for the Okuto model.

will no longer increase in velocity but instead will reach a maximum, saturated velocity. The design for LGAD's is such that they should be used in saturation mode, as this helps create a uniform field in the bulk region. This is also essential for ultra fast timing detectors which will be discussed in chapter 5.

As the LGAD's are designed to have a confined high electric field region to cause impact ionisation, it is essential to include a model for this in the simulation. Sentaurus has six models to describe the threshold behaviour of this avalanche multiplication region. The avalanche multiplication is used by calling the keyword *Avalanche* in the *Recombination* of the *Physics* section with the model chosen by calling the appropriate keyword. For this work the *Okuto* [52] model was used as it best describes the generation of charge carriers at the required electric field, in the range of 10^5 Vcm^{-1} – 10^6 Vcm^{-1}. This uses an empirical model given by

$$\alpha\left(F_{\text{ava}}\right) = a \cdot \left(1 + c\left(T - T_0\right)\right) F_{\text{ava}}^{\gamma} \exp\left[-\left(\frac{b[1 + d\left(T - T_0\right)]}{F_{\text{ava}}}\right)^{\delta}\right], \qquad (3.2)$$

where F_{ava} is the driving force for impact ionization, T is the temperature and which by default uses the parameters for silicon and $T_0 = 300$ K. The parameters used are shown in table 3.2.

The Schenk model [53] was chosen to simulate band–to–band tunnelling, which is essential when dealing with high electric fields, as this best described the conditions under test. Importantly the model works up to an electric field of $\approx 8 \times 10^5$ V/cm and for doping levels at the anode and cathode greater than 1×10^{19} cm^{-3}. A key aspect of this simulation is to monitor the charge collected for different doping profile designs in the LGAD detector, further described in section 3.2. In order to achieve this one must generate extra electron–hole pairs in some region of the device. This is achieved by using the *Heavy Ion* model. When a heavy ion traverses a device structure it loses energy and creates a trail of electron–hole pairs. This process depends on the initial conditions, the type of ion, the energy of the ion, the angle of penetration and the relation between the lost energy and the number of charged pairs created. These can all be described in the Sentaurus *Heavy Ion* model. The generation rate by the heavy ion is used most commonly in transient

simulations described in 3.1.3. The initial time of Ion interaction with the structure can be set and the number of electron–hole pairs generated is added to the carrier densities in the simulation. This information can then be plotted as a function of time to determine the drift of the carriers. This information is useful as you can plot the electron and hole current as a function of time and determine the charge collected. This worked involved the comparison of simulated LGAD devices with no gain devices in order to calculate the gain. The gain defined as the ratio of the charge collected from both given the same simulation conditions.

Simulation Conditions

The most basic form of device simulation is the calculation of the device in a steady state, typically this is done by setting the electrodes to a specific voltage which do not change with time. As this is in a steady state the time–dependent terms in the semiconductor equations are zero. This simulation can be performed to find the electric field in a detector at a specific voltage, or to simulate a p–n junction and find out its characteristics at 0 V, for example the built–in voltage and space charge distributions.

The *quasi-stationary* mode of device simulation is used to evaluate the current–voltage characteristics of a device. The device is first solved in a steady state as mentioned previously, then the electrode voltages are changed by a small amount (of the order 0.1 mV) and the device is re–solved in this new steady state. This process is repeated many times up to the set voltage required by the user.

At each step an initial guess of the solution is performed by extrapolation from the previous solution to speed up the process. At the early stages of the simulation the step sizes tend to be small as there is an abrupt change in the electrical characteristics, however once the detector is depleted the current remains constant and the calculations to the solutions become much faster.

The final simulation performed is the *transient* simulation, previously mentioned. The device is simulated over time but at a set voltage. The steady state is found for the detector over a series of small time intervals. At each time interval the change in carrier concentration and electric potential is found for each node of the mesh, this is then used to calculate the following time interval. This is used to find the current signal produced in a detector after a heavy ion interaction.

An example of a simple device with no gain, with a thickness of 200 μm and a width of 100 μm was simulated and results described below. The detector simulated has a bulk resistivity of 5 kΩ The *quasi-stationary* mode produced the current–voltage measurement from 0–500 V and the results shown in figure 3.3a. The current increases rapidly at low voltages as the detector depletes. This current arises from thermally generated electron–hole pairs, where the rate of generation is very low. A capacitance–voltage measurement

was taken to calculate the full depletion voltage of the detector and was calculated to be around 90 V shown in figure 3.3b. Measurements for capacitance were taken at 10 kHz and at 5 V steps. The full depletion voltage can be calculated using equation 2.23 to give:

$$V_{fd} = \frac{\left(1.6 \times 10^{-19} \text{C}\right)\left(3 \times 10^{12} \text{ cm}^{-3}\right)}{(2)\left(11.7 \times 8.85 \times 10^{-14} \frac{C}{V \times cm}\right)} \times (.02 \text{ cm})^2 = 93 \text{ V} \tag{3.3}$$

Calculating this gives a full depletion voltage of 93 V. The device simulated had infinite boundary's meaning there is no edge effects which may cause electrical breakdown. There is no breakdown shown in this plot as expected from a PIN diode at this voltage range. It can also be useful to map the electric field profile with increase in voltage. The simulation can provide reliable field profiles which can indicate the depletion depth within the bulk of the material. In figure 3.4 the electric field is plotted as a function of depth for a voltage range of 0–150 V. The electric field extends into the bulk where the highest field is seen at the n^+ electrode. Once full depletion is reached the electric field within the whole of the detector increases, this would occur until the device breaks down.

Plots showing the electron current density (eCurrentDensity) and hole current density (hCurrentDensity) as a function of time are shown in figures 3.5 and 3.6 for a no gain device respectively. The charge is deposited in the centre of the device at a depth of 100 μm and 50 μm from the edge of the device. The amount of charge deposited is 1.282×10^{-5} pC. They show evolution of carriers through the silicon substrate, where they are collected at either electrode. It's clear from figure 3.6 that the movement of holes is much slower than that of electrons, due the difference in mobility and saturation velocity. The current collected at the top electrode (n^+ electrode) is plotted as a function of time and shown in figure 3.7 at a bias voltage of 500 V. The sharp edge indicates the arrival of the electrons to the electrode where they are collected almost instantly. However the tail of the curve shows the collection of the holes at the backside electrode, which is far slower. For fast detectors the total collection time is not as important but the leading edge, *rise time*, which is a main factor in the timing performance.

Simulation Flow

The simulation process as described previously follows the following pattern:

- A full device process simulation is performed with a set meshing strategy, which is defined at the beginning of the simulation. The is done using the *Sentaurus Process* tool

- After process simulation a re-meshing step is performed in order to best define the structure before any electrical characterisation. This is performed using adaptive
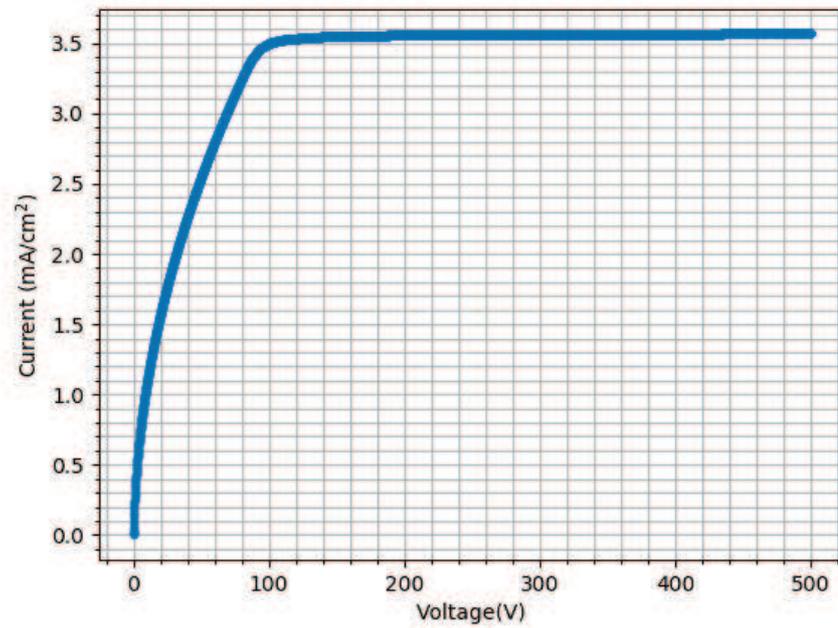
(a) IV simulation
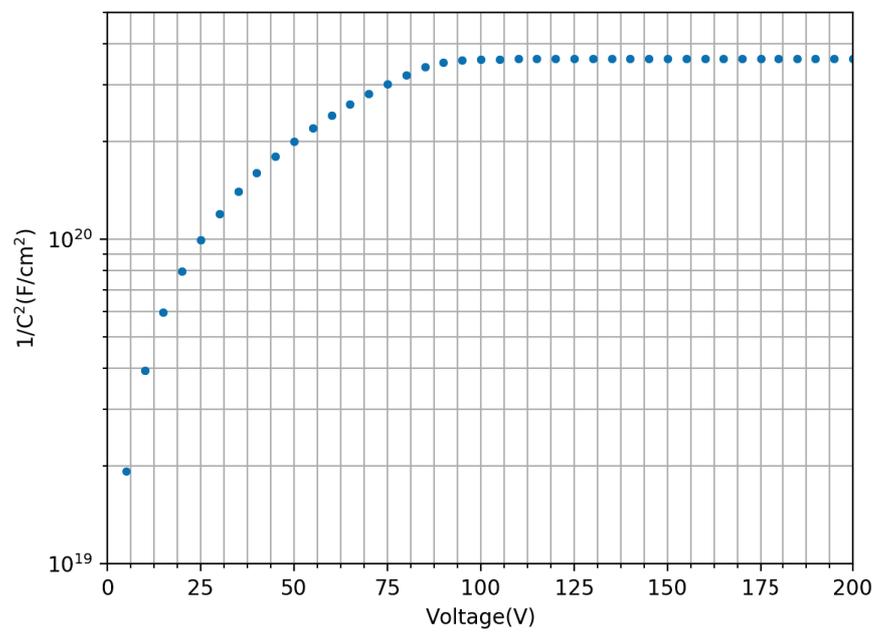
(b) CV Simulation

Figure 3.3: Example of a simulation for a PIN diode showing current–voltage and capacitance–voltage characteristics.
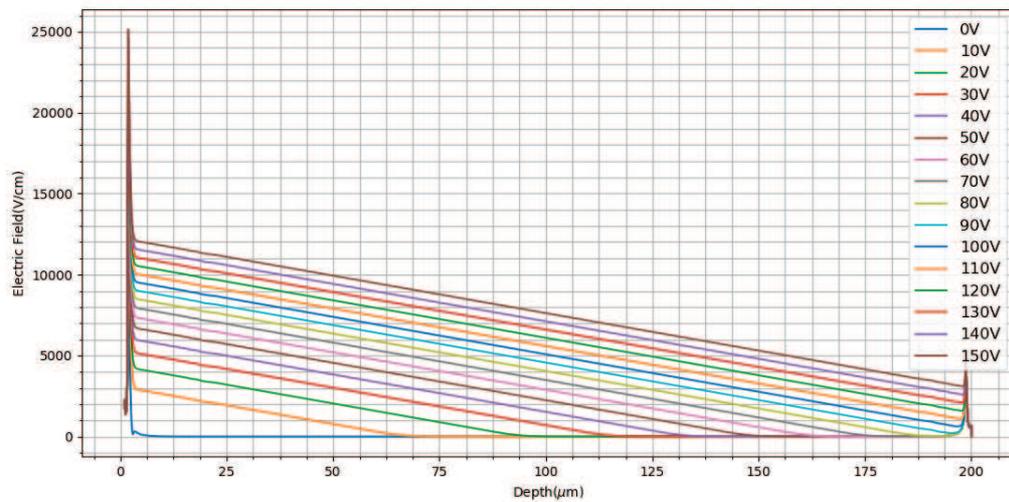
Figure 3.4: Electric field as a function of depth at a range of voltages for a PIN diode.

meshing techniques in the *Sentaurus Structure Editor* tool.

- The device undergoes a *quasi-stationary* simulation to characterise the electrical behaviour of the device using *Sentaurus sdevice*. From this the electric field profile is established at a range of voltage steps, the full depletion of the device can be calculated as well as the breakdown voltage.

- The device begins a *transient* simulation, again using *Sentaurus sdevice*, where a heavy ion interacts with the device structure. The charge collected can be visualised and calculated as a function of time.

3.2 LGAD Design

After some initial validation of the simulation package using a PIN diode structure. Where the current–voltage and capacitance–voltage measurements agreed with the simulations, the method was then applied to a new type of detector technology, the Low Gain Avalanche Diode (LGAD). Using the experience gained from simulation of a simple structure and using the literature available, [29], [36], simulation began to prepare a fabrication process to produce LGAD detectors with industry partner Micron Semiconductor Ltd[1] There was also some important collaboration with the inventors of this technology, CNM[2] in Barcelona, specifically for improving the breakdown capabilities of the Micron Semiconductor devices.

---

[1]Micron Semiconductor Ltd,1 Royal Buildings, Lancing, Business Park, West Sussex, BN15 8SJ, UK, http://www.micronsemiconductor.co.uk
[2]Centro Nacional de Microelectrònica de Barcelona, Campus UAB, Carrer dels Tillers, 08193 Cerdanyola, Barcelona, Espanya, http://www.cnm.es
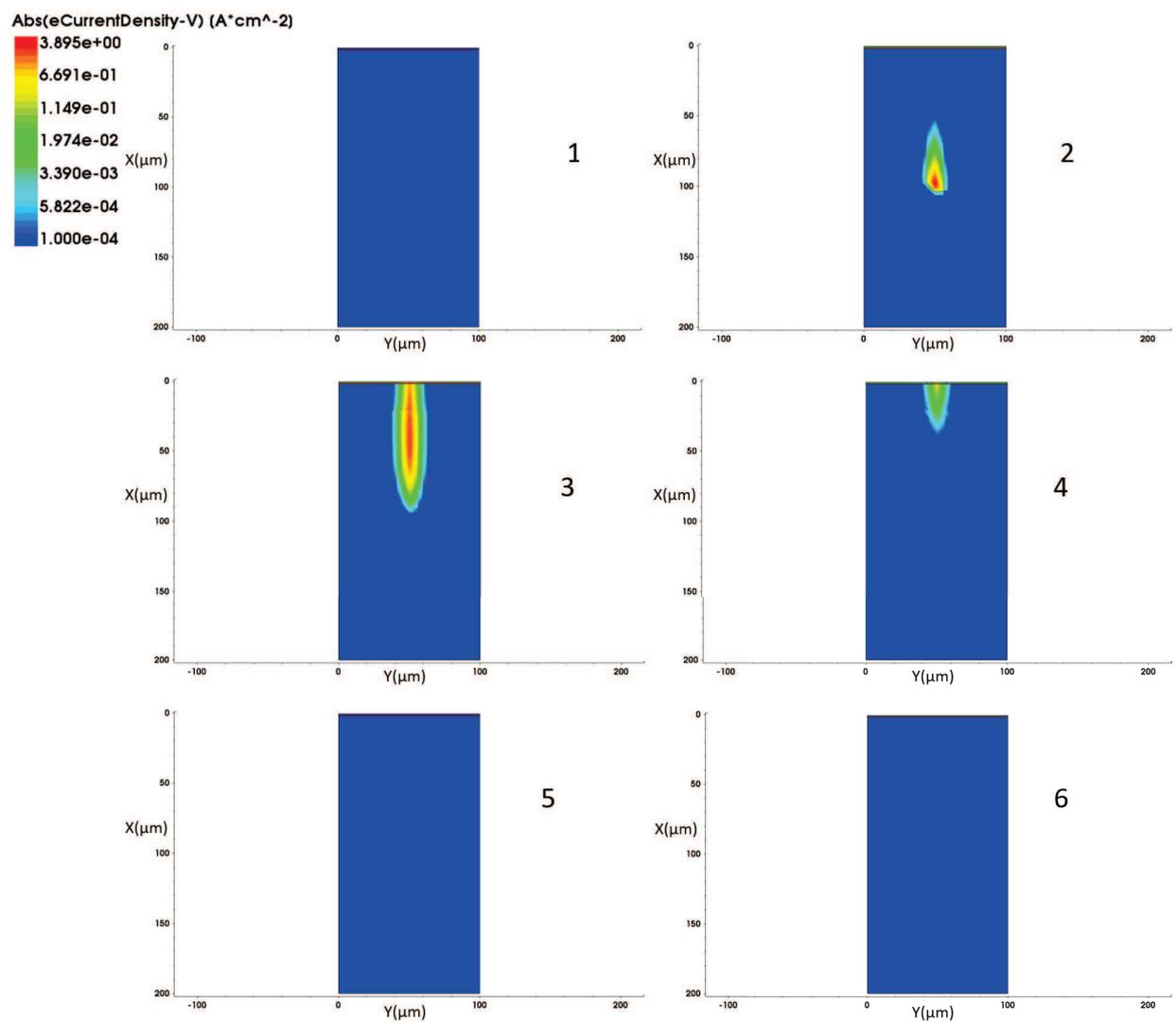
Figure 3.5: 2D plot of the electron current evolution in the simple diode after a heavy ion interaction for a bias voltage of 500 V. Each tile represents a different time slice in steps of 1 ns.
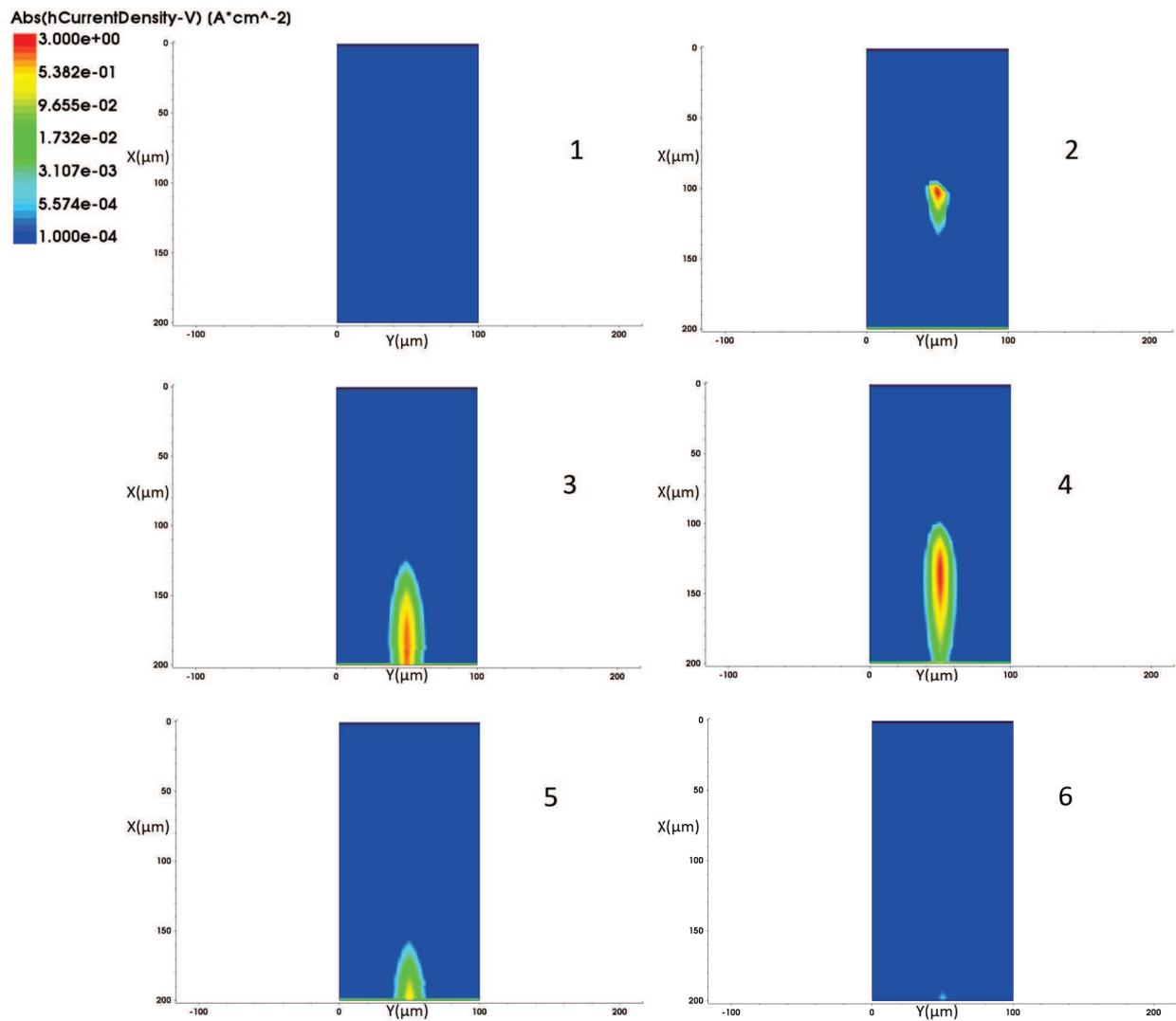
Figure 3.6: 2D plot of the hole current evolution in the simple diode after a heavy ion interaction for a bias voltage of 500 V. Each tile represents a different time slice in steps of 1 ns.
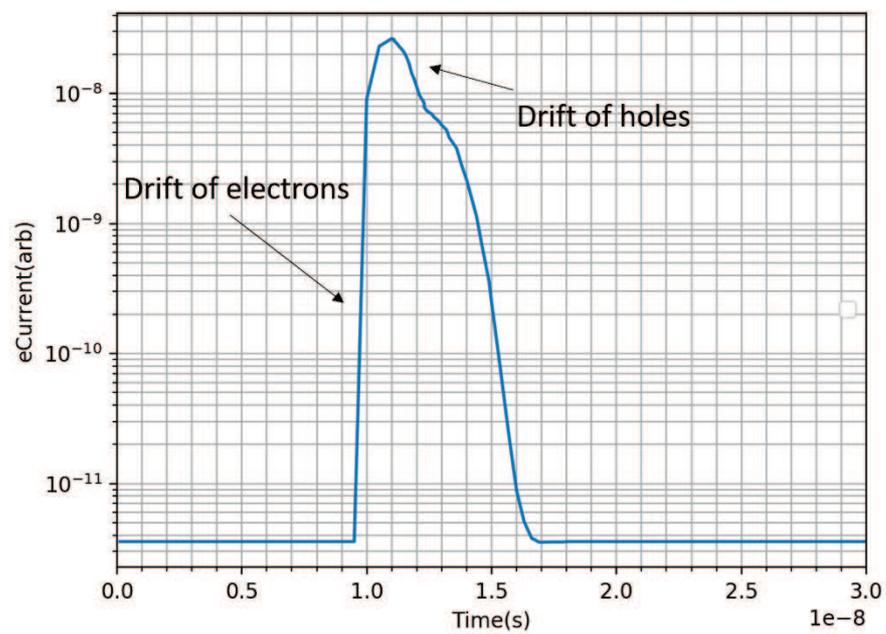
Figure 3.7: Current as a function of time after a heavy ion interaction in the centre of the device for a bias voltage of 500 V. The electrons are collected quickly at the top collecting electrode while the holes drift slowly through the device and are collected at the bottom collecting electrode.
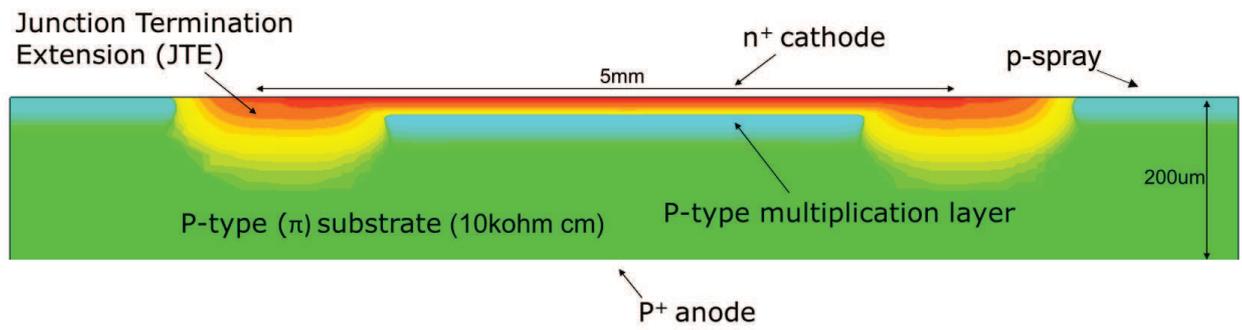
Figure 3.8: Schematic cross–section of the LGAD pad design. A deep p–type layer is located beneath the n^+ electrode to form the n^+/p/p$^-$ where the multiplication takes place. The colour highlights the doping concentration, with green having a lower concentration (of the order 5×10^{12}) and red having the highest concentration (of the order 1×10^{20}).

3.2.1   LGAD Concept

As mentioned in the previous chapter a new type of detector has been proposed which operates on the principle of avalanche multiplication by impact ionisation. This was first introduced to overcome the thermal and electronic noise induced by radiation damage. Having a region of charge amplification within the detector allows a signal to be produced which is much larger, also allowing for a lower gain amplifier in the readout electronics to be used. This behaviour is seen and used in photon detection by avalanche photodiodes (APD), however no information on the initial particle energy can be extracted. However if one were to design an avalanche diode with lower gain this information could be preserved.

The Low Gain Avalanche Diode (LGAD) detectors are based on an n in p diode with an extra multiplication layer, providing a desired gain of 5–20 at 200 V bias, with a breakdown voltage significantly higher than this of at least 400 V. The internal structure shown in the fabrication section, uses planer technology to produce a n^+/p/p$^-$/p$^+$ structure and is shown in figure 3.8. The figure is taken from a TCAD simulation showing the relative doping concentrations. It shows a highly doped n^+ cathode electrode with a moderately doped p–type region below, known as the multiplication implant. The p–type multiplication implant is diffused deeply into the bulk (≈ 4 μm) and has a peak doping concentration in the region of 1×10^{16} cm^{-3}. The n–type electrode has a peak doping concentration in the region of 1×10^{19} cm^{-3} and has a shallow profile into the bulk of ≈ 1 μm, typical for a PIN diode. The bulk material is high resistivity p–type silicon (≈ 5–10 kΩcm), corresponding to a doping concentration around 2×10^{19} cm^{-3}, with a p$^+$ anode electrode on the backside.

As a high resistivity p–type bulk is used a shallow uniform p–spray doping (order 1×10^{15} cm^{-3} and 4 μm deep) is implemented to isolate the cathodes. To reduce the magnitude of the electric field at the perimeter of each cathode an additional deep n^+ doping region (known as the Junction Terminating Extension, JTE) is present, discussed in more detail later.
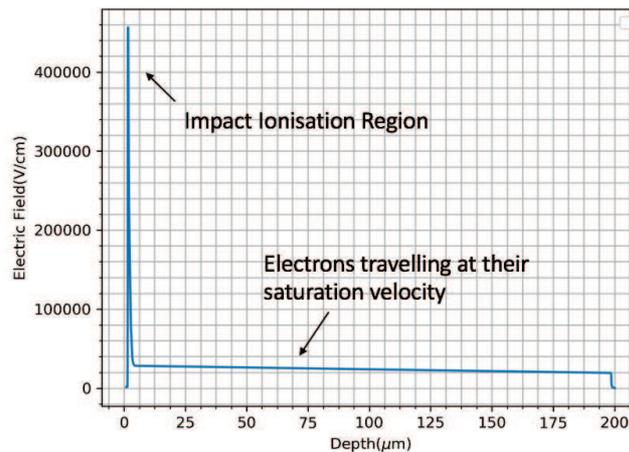
Figure 3.9: Typical electric field profile through device showing high electric field at junction between n+/p region. The device is 200 μm thick with a reverse bias of 500 V

As the detector is reverse biased, a high electric field region is present at the interface between the n^+ and p implants, an example of this is shown in figure 3.9 for a simulated LGAD device biased at 500 V. This field causes impact ionisation as electrons cross the high field region. The gain of the device is crucially dependent on the doping profile of the p-type multiplication layer. Simulations have been performed to accurately model the behaviour of LGAD detectors with various doping profiles.

LGAD Gain Structure

The LGAD structure, based on an APD, has to be designed very carefully in order to be used in the linear avalanche mode. The main driver of this design is to promote avalanche gain and to control it effectively. The aim is to produce a detector which has a gain with a small dependence on voltage, whereby the collected signal is proportional to that of the incoming photon or particle.

To control the gain of these devices, one must control the peak electric field at the junction between the n^+ and p implants. It has been shown that the electric field, thus the gain increases as a function of the boron implant dose. An example of the doping concentration through the n-type electrode and p-type multiplication implant is shown in figure 3.10, taken from a simulation. To understand the high-field performance of the device the concept of the effective p-type doping areal density, $Q_{effective}$ is introduced, shown in figure 3.10 and defined as the net p-type doping concentration integrated over the depth of the implant.

All simulations performed have undergone process and device simulation as described in section 3.1.3. The gain of the LGAD device was calculated by comparing the collected charge to that simulated for a standard PIN diode. For significant impact ionisation to
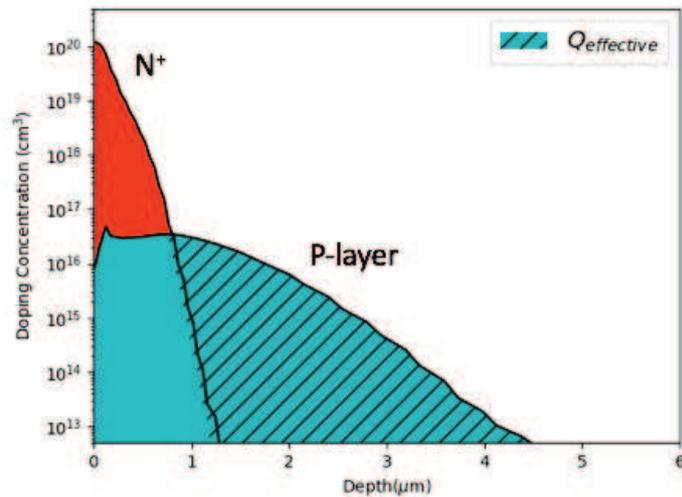
Figure 3.10: Example of the doping profile throughout the n$^+$/p region.

occur an electric field of $E_{crit} \approx 10^5$ V/cm is required. The E_{crit} is directly related to the doping concentration within the detector, as shown in [54]. The peak concentration in the lower doped side of the n+p junction is in the range of 1×10^{16} cm^{-3}. According to Baliga [55] for a doping concentration in this range the E_{crit} value is in the range of 4×10^5 V/cm, however lower E-fields will still result in some gain.

To investigate the effect of the p–type multiplication implant dose and doping profile on the device gain a series of simulations were performed for a range of boron implant dose concentrations and energies, while the phosphorus implant parameters were kept constant. A change in the p–type multiplication dopant level and distribution changes the $Q_{effective}$ of the device. The simulations demonstrated that the gain of the device is a function of $Q_{effective}$, shown in figure 3.11, for the LGAD under a reverse bias of 400 V. There is a strong correlation between gain and $Q_{effective}$ and a critical value of $Q_{effective}$, defined as $Q_{effective}^{crit}$, for which a gain of 1.1 is achieved, where $Q^{crit}_{effective} = 1.6 \times 10^{12}$ cm^{-2}. As the gain increases with increasing $Q_{effective}$ the breakdown voltage decreases due to higher fields in the device at lower bias voltages. In this case the breakdown voltage is caused by the high electric field in the multiplication region between the n+ implant and the p–type multiplication implant, as there are no edge effects present in the 1D simulated case. This simulation work agrees with the results shown in [54].

To understand the gain mechanism it is useful to look at the simulation plots of the electron and hole distribution with time, following a heavy ion interaction. Results have already been discussed with a simulation of a PIN diode in figures 3.5, 3.5 and 3.7. Similarly plots for the LGAD simulation are shown in figures 3.12 and 3.13 at a bias voltage of 500 V. The collection of electrons is very similar to the PIN diode case, the difference being the total amount of electrons collected is higher in the LGAD case. For
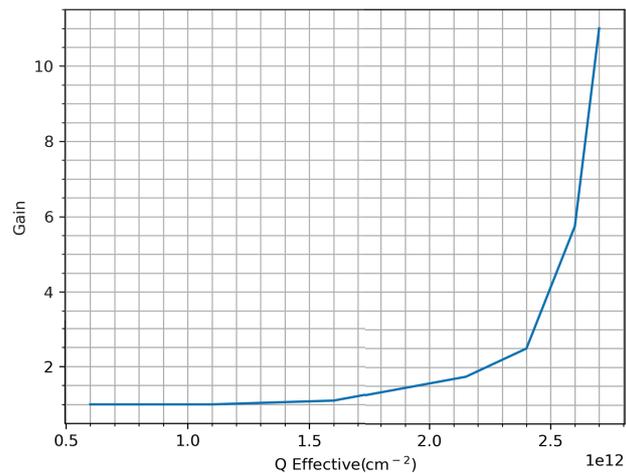
Figure 3.11: Simulated Gain as function of $Q_{effective}$ for a bias voltage of 400 V.

the collection of holes shown in figure 3.13 the gain mechanism is very clear. The electrons drift to the top electrode and e–h pairs are produced at the multiplication region. One can see clearly the generation of holes in this region and the drift of these holes to the backside electrode. The total amount of holes collected is significantly larger than that of the PIN diode clearly seen in this diagram. The charge collected over time can be plotted and is shown in figure 3.14. The initial drift of electrons is the same as with the PIN diode, however the onset of multiplication is clear where there is a large increase in collected electrons. The drift of holes is also evident and is also clearly much larger than with the PIN diode. This plot shows clearly the difference in drift velocity of the electrons and holes respectively.

The total charge collected can be calculated by integrating with respect to time and with comparison to the charge collected from the PIN diode, figure 3.7, the gain can be calculated. For this example the gain is calculated to be $\approx$ 13 at 500 V. Plotting the gain as a function of voltage produces the figure 3.15. It can be seen that the gain increases with increasing voltage. Calculating the $Q_{effective}$ gives a value of 2.76×10^{12} cm^{-2}, which is comparative with the plot in figure 3.11.

High Field Edge Effects

Having the desired doping profiles in the active region to produce the required gain, it was then important to simulate any edge effects of these devices, which can severally effect the performance of these devices. All simulations to this point have assumed infinite boundaries. 2D simulations were performed to investigate the breakdown voltage of the the device due to field distribution at the edge of the LGAD implants. There are a few design considerations for the optimum periphery of the LGAD detector.
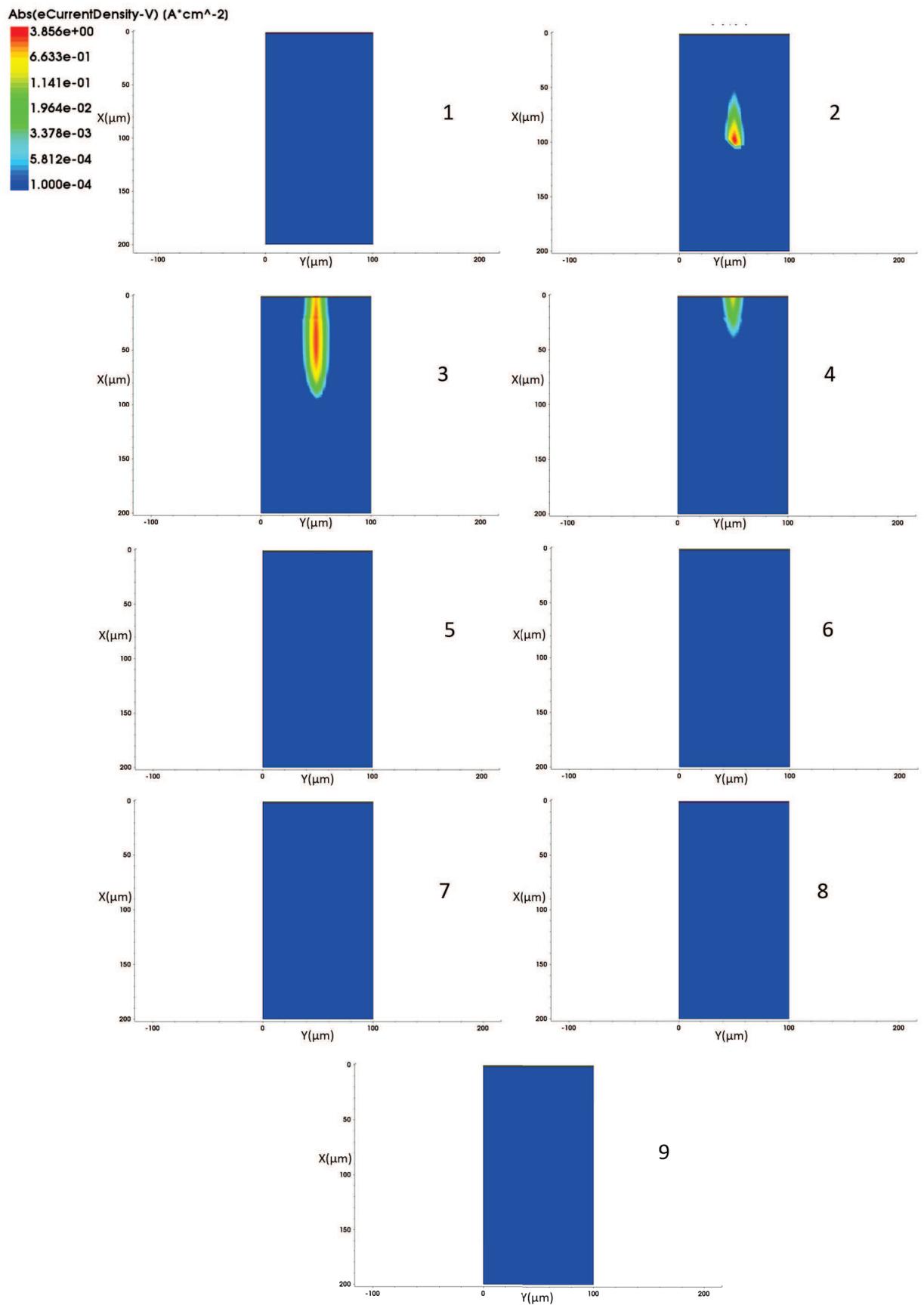
Figure 3.12: 2D plot of the electron current evolution in the LGAD after a heavy ion interaction for a bias voltage of 500 V. Each tile represents a different time slice in steps of 1 ns.
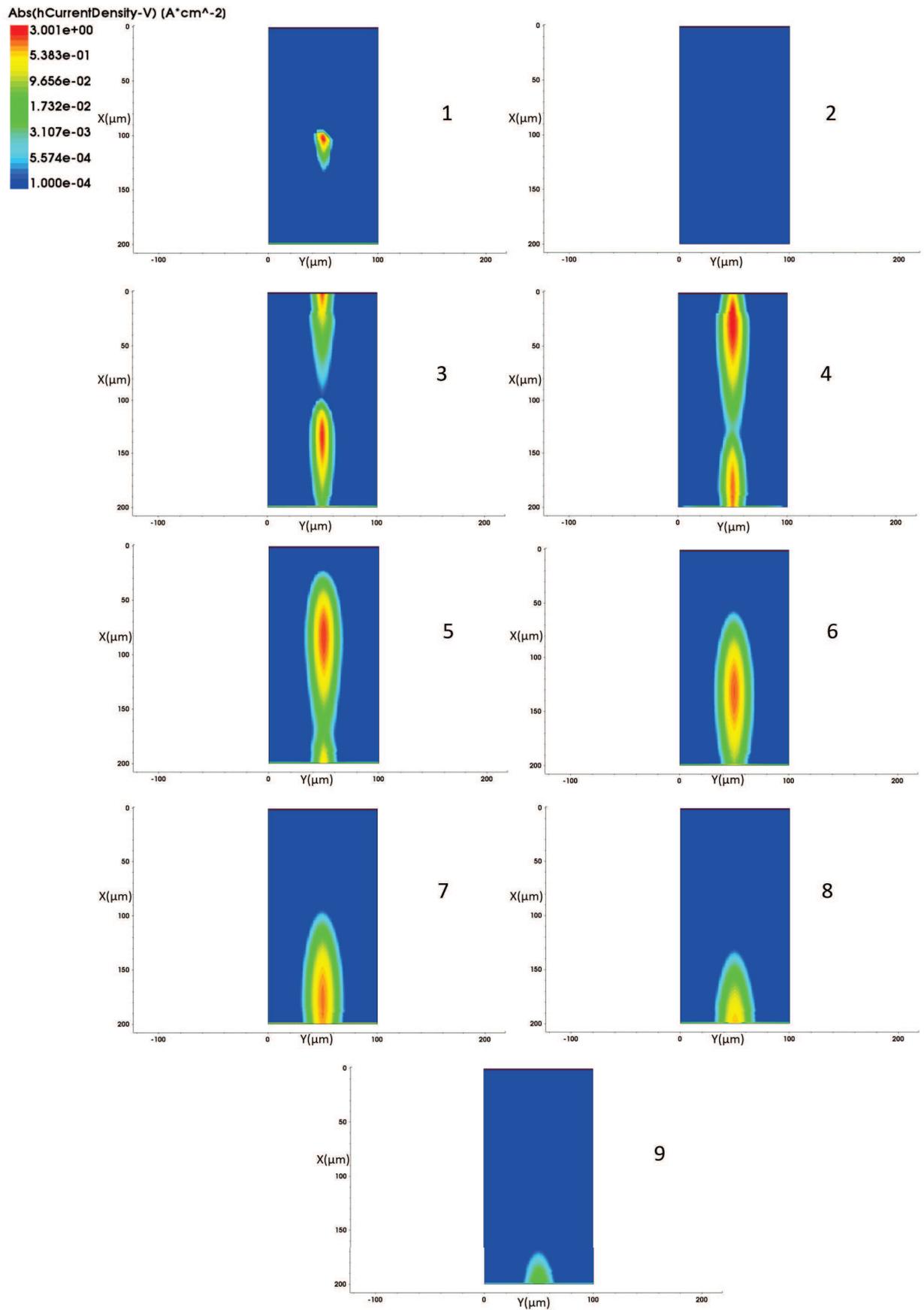
Figure 3.13: 2D plot of the hole current evolution in the LGAD after a heavy ion interaction for a bias voltage of 500 V. Each tile represents a different time slice in steps of 1 ns.
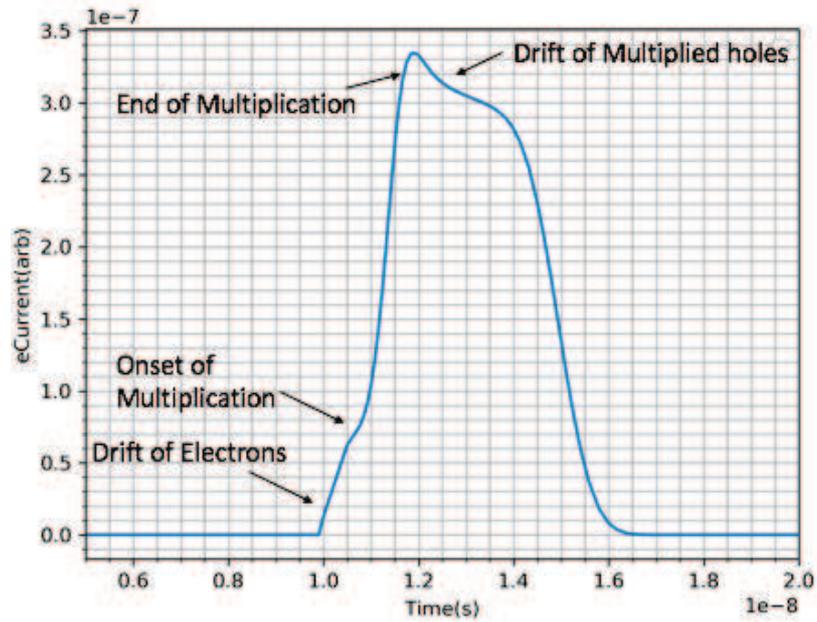
Figure 3.14: Electron current as a function of time after a heavy ion interaction. The final kink indicates the end of the drift of charge carriers, where the slope is due to diffusion of charge carriers.
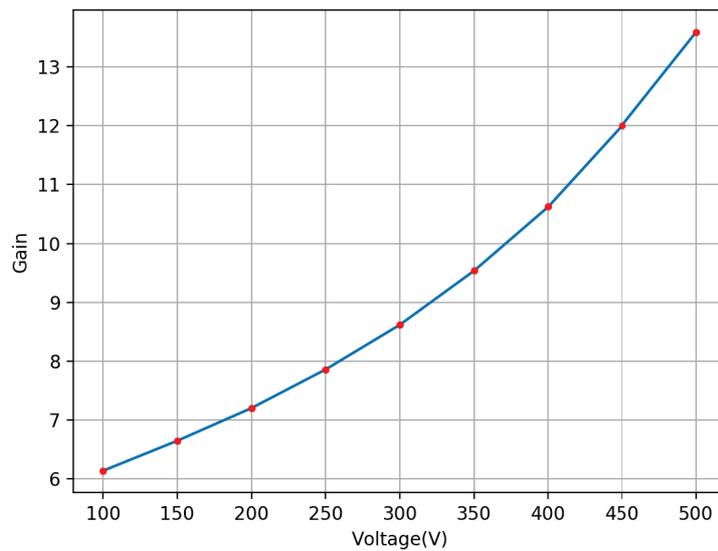
Figure 3.15: Simulated Gain as a function of Voltage

The process by which breakdown voltage occurs must first be considered. In the 1D simulation with infinite boundaries the breakdown voltage occurred in the multiplication region where the high electric field was designed to be. As the voltage increases in the LGAD this electric field increases to a critical value, where the junction breaks down. The critical electric field, E_C for breakdown has a dependence on the doping concentration in that region. Looking at the peak concentration in the lower doped side of the n^+p junction, we find that it's in the region of 1×10^{16} cm^{-3}. The critical electric field on the doping concentration, which is derived in [55], gives an E_C value in the range of 4×10^5 V/cm [56]. Due to the nature of avalanche gain it's key to control this field through accurate doping profile control.

In a PIN diode where there is no multiplication implant, the lower doping concentration at the junction between the bulk (p–type) and the cathode (n–type) is in the range of 1×10^{12} cm^{-3}, giving it a reduced E_C value of 1.5×10^5 V/cm. The edge region is therefore designed considering this information.

At the edge of the LGAD device the p–type multiplication implant can either align with the edge of the n^+ cathode or be smaller. No variation of this design has included having this larger than the cathode due to findings in [56]. Due to these findings the junction termination extension (JTE) has been introduced in LGAD detectors. This is an n–type implant which is deeply diffused into the bulk surrounding the active LGAD structure. It requires an extra mask level and to ensure deep diffusion it is the first step implanted in the production of LGAD's. The JTE is designed following research of power devices [57]. It allows the depletion region to spread over it which causes a spreading of the electric field, causing a decrease in the maximum electric field at the edge region.

The p–type multiplication implant must be smaller than the n^+ cathode due to the JTE, however this reduction in implant size reduces the active fraction of the device and leads to a spatial variation in the response of the device. Thus it's important to minimise this effect where possible. The edge region doping profile was simulated using process simulation and is shown in figure 3.8. The doping profile for each of the regions is shown in figure 3.16, where it's clear the JTE has a deeper profile with respect to the multiplication implant. The boundary conditions were reflective and only the last 30 μm of the n^+ cathode implant was simulated. The simulated p–type multiplication implant region was 20 μm wide, defined to be 10 μm inside the n–type implant. The region defined in the mask layer for the JTE implant was 10 μm wide and 10 μm from the edge of the multiplication implant. In order to ensure accurate final doping profiles a uniform p–spray implant was included and this extended 40 μm from the end of the JTE implant. Due to lateral diffusion of the dopants during thermal processing steps the exact implant regions can extend by up to 10 μm depending on the number of steps. The implant regions of the multiplication implant and the JTE were designed to diffuse together as this showed the best results.
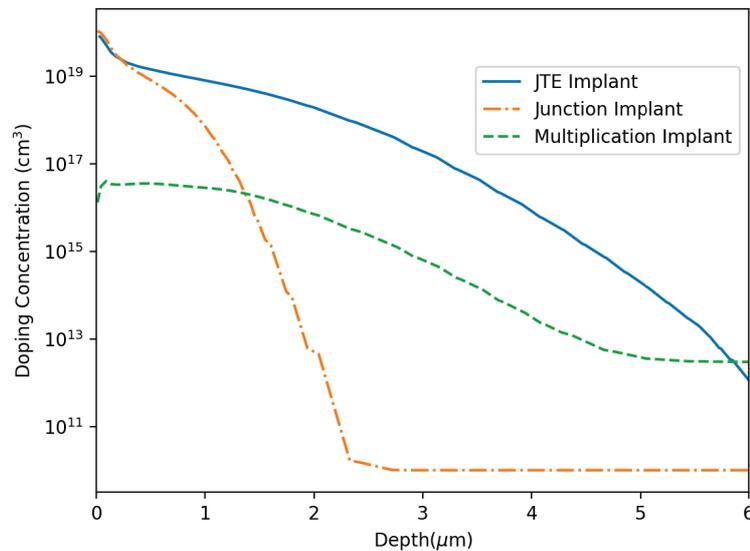
Figure 3.16: Simulated Doping profiles of the active area implants, where the JTE and junction profiles show the phosphorus concentration and the multiplication profile shows the boron concentration. The JTE implant profile is taken from the region across the JTE. The junction and multiplication implant profiles are taken from the regino across the junction. The caption is; Junction Implant: dot-dashed line, JTE implant: solid line and Multiplication Implant: dashed line

Another design considered was the overlap of the n^+ implant. The implant diffuses deeper into the bulk due to the lower doping concentration in that region, it was hoped this would have a similar effect as the JTE. It is a much easier method to use as it does not require an additional implant or implant mask. Figure 3.17 shows the maximum electric field at the edge of the LGAD structure for both the JTE and overlap designs at an applied bias of 200 V. The electric field at the interface between the n^+ and the p–type multiplication implant is approximately constant for both designs, at a distance of 0 to 15 μm. Following the edge of p–type multiplication implant there is a decrease in the maximum electric field, which then rises again as we reach the edge of the n^+ implant, at the interface of the p–spray and n^+ implant. For best results the peak electric field should be in the multiplication junction so that breakdown at the edge is avoided. Without the JTE the electric field at the edge of the n^+ cathode is larger than the electric field in the multiplication region and will result in breakdown taking place at this location.

The JTE is fabricated with the same dose as the n+ cathode implant, however it has a much longer diffusion time to obtain deeper dopant diffusion into the silicon. To reduce the edge electric field the JTE must be deeper than the p–type multiplication junction and wide enough to produce a gradual reduction in the potential to the p–spray implant. The peak concentration in the JTE is also a key factor in the effectiveness of the JTE for breakdown
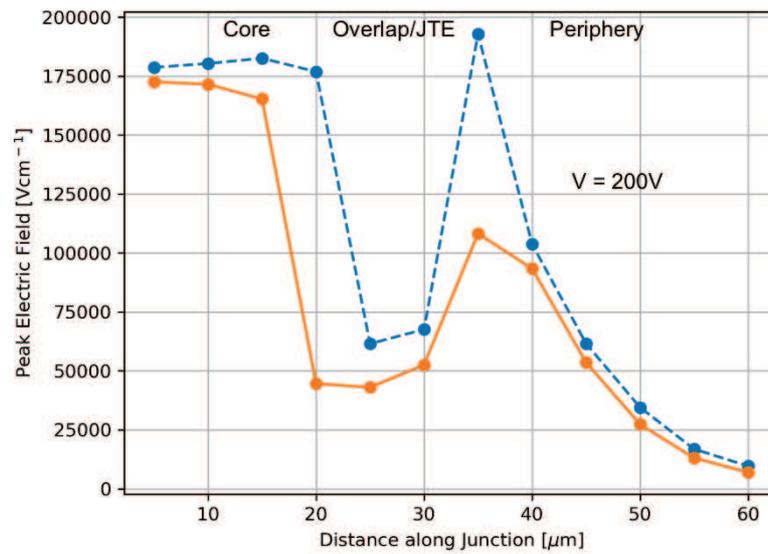
Figure 3.17: Simulated maximum electric field in the device for a bias of 200 V as a function of distance along the surface in the edge region of the n+ junction for a device with and without a JTE. The caption is, simulated without JTE but with overlap: dashed blue line, simulated with JTE: solid orange line.

performance, from [56], figure 3.17 shows the maximum electric field in the LGAD in the presence of a JTE. The maximum field in the edge region is reduced by a factor of two, for this design, and critically is less than the maximum field in the multiplication region.

An additional benefit of the JTE is the reduction of the leakage current in the detector. This is because without the JTE any thermally generated electrons will be multiplied in the multiplication region. As the JTE is deeper than an overlapping electrode it tends to collect more of these thermally generated electrons from the detector bulk and surface current. This means that they will not undergo avalanche multiplication.

The simulated current–voltage characteristics improves with the use of the JTE due to the lower edge electric field, shown in figure 3.17. The breakdown voltage of the device with a JTE is 780 V compared with 290 V without the JTE, shown in figure 3.18. Devices have been manufactured with and without a JTE for comparison with simulation and the results presented in chapter 4.

3.2.2 Secondary ion mass spectrometry (SIMS)

Secondary ion mass spectrometry (SIMS) is a technique used to analyse the composition of solid surfaces and thin films from a few angstroms up to tens of microns with a sensitivity of 10^{13} atoms/cm^3 [45]. The method works by sputtering the surface of the specimen with a focused primary ion beam , with an energy in range 0.5–20 keV. Typical ions used are O_2^+ and Cs^+. Analysing the ejected secondary ions by mass spectroscopy allows multi-element detection with high depth resolution. The SIMS method is a destructive technique
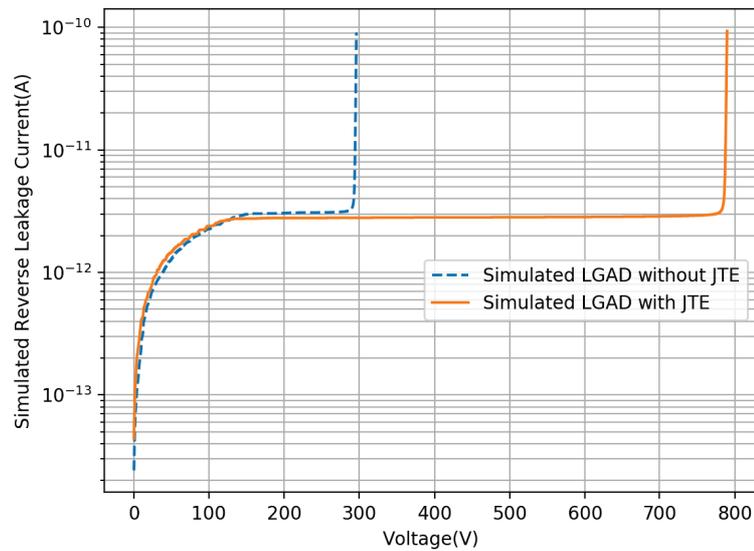
Figure 3.18: IV curves for simulated LGAD devices with and without a JTE. The caption is, simulated without JTE but with overlap: dashed blue line, simulated with JTE: solid orange line

which leaves a crater in the tested sample. This technique was used to monitor the doping profiles of both phosphorus and boron with a depth resolution of 1–5 nm [58], to compare with the simulated results. This is important to validate the fabrication process against the simulation model used. This method can also to be used to characterise the discrepancies from one run the next.

The SIMS measurements used in this work were all performed at EAG Laboratories [3]

Comparison of SIMS to simulation

Several process runs were performed for a small range of boron doses and implant energies based on the results from the simulations. These were also performed for a range of JTE implant doses and drive-in times. To obtain the required gain and sufficient breakdown voltage for these devices it was key to focus on the $Q_{effective}$ as described in section 3.2.1. In order to measure the doping profiles and hence the $Q_{effective}$ of the fabricated devices it was necessary to perform SIMS measurements on some of the fabricated devices for which the device characteristics are shown in chapter 4.

Figures 3.19b and 3.19a from one run show the comparison of the simulated doping profiles with the SIMS measurements. The simulation was tuned to the exact details of the fabrication process so as to allow the best comparison. Figure 3.19a shows the phosphorus doping profiles for the junction implant which shows the simulated profile matches very

---

[3]EAG Laboratories,14, Avenue du Docteur Maurice Grynfogel, ZAC Basso Cambo II, 31100 Toulouse, France , http://www.eag.com

well the fabricated device. At a depth range of 0-50 nm the SIMS result is quite different to that of the SIMS. The peak at 0 nm can be due to higher concentrations of ions present in the oxide. However it's difficult to explain the drop in concentration at 20 nm and thus is not present in the simulation. At around 800 nm there is a slight discrepancy between the results, however at this depth the error on the SIMS measurement becomes an issue for comparison.

Figure 3.19b shows the comparison of the multiplication implant (Boron). The general shape of the implant is very similar in both cases, there is however a few discrepancies in the profile. The peak implant dose at the surface is an order of magnitude higher in the SIMS measurement, however the two plots reproduce the same peak formed at a depth of $\approx$ 100 nm. The simulation shows a slightly higher peak at this depth. The simulated profile is slightly greater from this peak until $\approx$ 1000 nm depth, where after the SIMS measurements shows a greater doping concentration.

Calculating the $Q_{effective}$ for each of these profiles gives some indication of the likely gain to be produced. The simulation profiles produce a $Q_{effective}$ of 2.37×10^{12} cm^{-2}, while the SIMS measurement gives a value of 2.39×10^{12} cm^{-2}. The results of the gain of these devices are presented in chapter 4, where the SIMS results where taken from a detector from run 2. From figure 3.11 the gain for a fabricated device can be estimated to be in the range of 2-3 at a bias voltage of 400 V based on the simulation results

3.3 Conclusion

In this chapter the simulation process has been thoroughly described. The process simulation, which calculates the eventual doping profiles in the detector, has been described using the physics models and meshing strategy's available in Sentaurus TCAD. A brief description of the fabrication process used and the various ways to manipulate the doping profiles has been described. This includes altering implant doses, implant energies and thermal processing steps. The resultant doping profiles from simulation have been validated using the SIMS method for fabricated devices. A comparison has been made on the shape of these doping profiles as well as the $Q_{effective}$ where the results of gain will be shown in the following chapter.

Following process simulation a full device simulation was described and results of which presented. Device simulation was used to characterise the electric field profile in the device and calculate the resultant gain produced. It was shown that the total gain is related to the $Q_{effective}$, where there is a critical value of $\approx 1.6 \times 10^{12}$ cm^{-2} before any gain is present. This is consistent with all simulations performed. The gain was shown to increase as a function of voltage as the electric field in the gain region grows. The signal generated in an LGAD detector was shown, describing the drift of electrons, onset of multiplication and
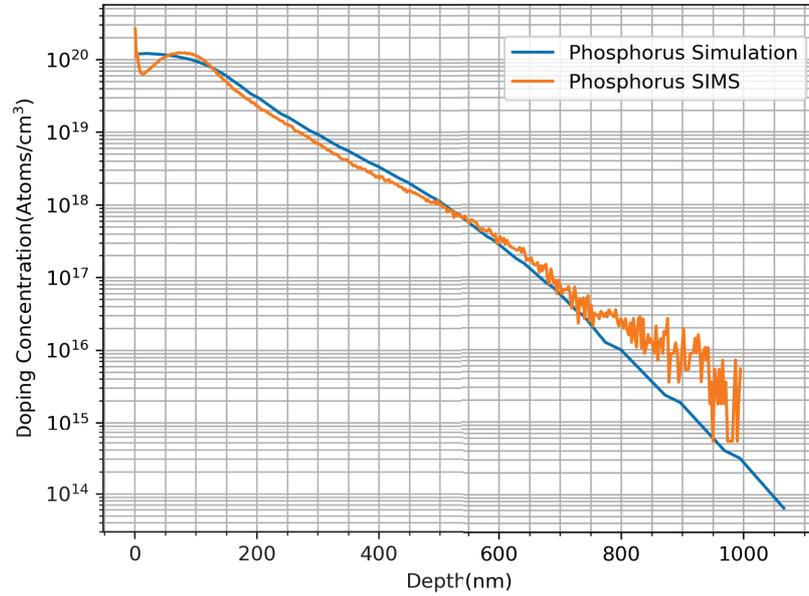
(a) Phosphorus doping profile.
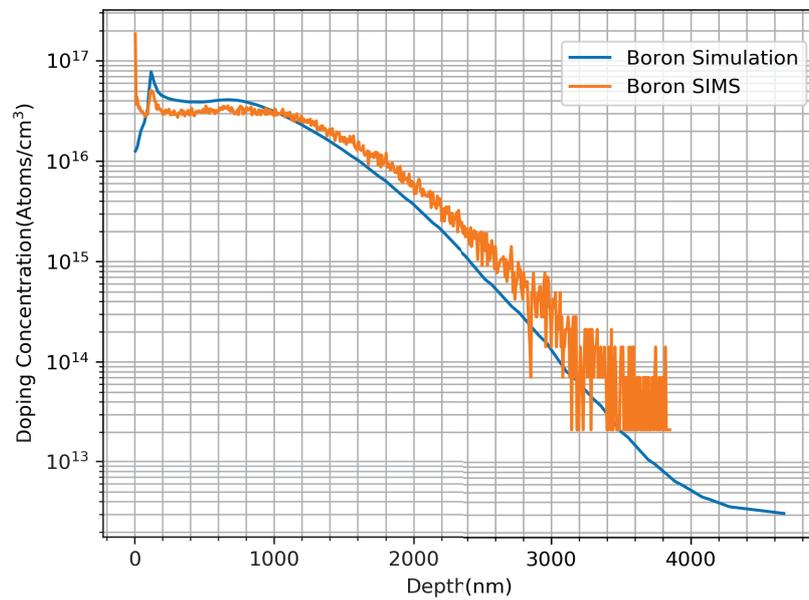
(b) Boron doping profile.

Figure 3.19: Comparison of dopant profiles from simulation and SIMS measurements

finally the drift of the generated holes.

Finally the high field edge effects were described for a typical LGAD detector, where a high electric field is present at the edge of the junction which causes a premature breakdown. Two solutions to this were described, the overlapping junction and the Junction Termination Extension (JTE). It was shown that the JTE was a better candidate for this as it better reduced the field at the edge and thus resultant breakdown voltage was greater when a JTE was used.

Chapter 4

Devices and Characterisation

This chapter discusses the devices fabricated and the characterisation techniques necessary to evaluate the performance of fabricated LGAD devices. The various techniques will be introduced and results shown as required. Section 4.2 discusses the electrical characterisation of devices to evaluate the full depletion voltage, leakage current and breakdown voltage. Section 4.3 describes the transient current technique used to evaluate the gain of the LGAD devices using a focused laser, as well as any gain variation across the detector. Finally section 4.5 discusses the method and results obtained using various radiation sources, including ^{90}Sr, Alpha Particles and X–rays.

4.1   Fabricated Devices

This section will give an overview of the fabricated devices discussed in this thesis. The aim is to provide an overview of the fabrication parameter difference between runs. Obviously not all parameters could be included for confidentiality reasons. Where appropriate relative doses have been given rather than actual values. More information on the fabrication process flow is shown in chapter 2. A total of five fabrication runs were performed throughout this project with various processing parameters and mask designs used.

Wafer runs 1–3 were fabricated using Mask set 1 (M1) and the relevant processing parameters shown in table A.2. An example of the wafer following fabrication is shown in figure 4.1, where the devices are situated around the outside of the wafer. Due to the design of the detectors and the limited space on this wafer a control wafer for each run had to be fabricated in order to test the gain of the devices. Results for run 3 are not shown in this thesis as all wafers fabricated produced no working devices. This was due to a process modification, where the multiplication implant was implanted at a different facility. The implanter was not the same as the one used at Micron which resulted in a slightly different dose that proved critical for this run.

Wafer run 4 was fabricated using a second mask design, Mask set 2 (M2) with relevant

Table 4.1: Fabrication runs 1–3, showing process parameters used

Wafer Run	Pad Size (mm)	Relative Boron dose	JTE	Relative Phosphorus Implant Energy
1	1x1, 2x2, 4x4, 5x5, 10x10	0.5, 1	None	1
2	1x1, 2x2, 4x4, 5x5, 10x10	1.3, 1.6, 1.9	None	1
3	1x1, 2x2, 4x4, 5x5, 10x10	1, 1.1, 1.2, 1.3	None	1
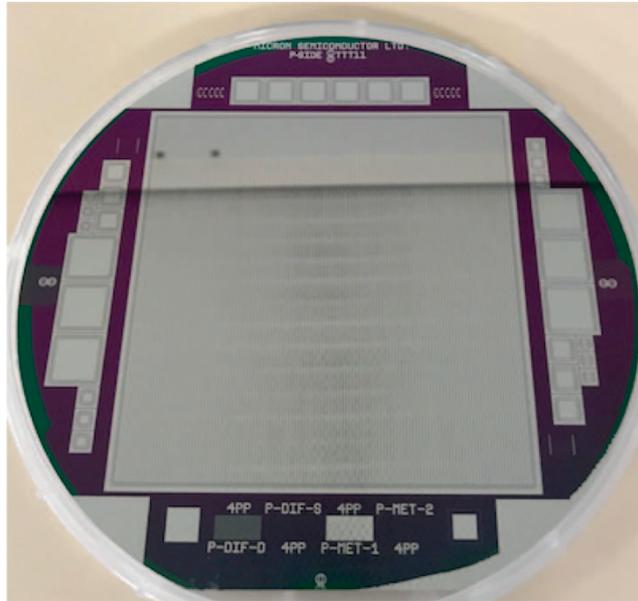
Figure 4.1: Example of a finished wafer from runs 1–3.

process parameters shown in table 4.2. The differences between M1 and M2 were the introduction of the JTE and control devices on the wafer. This meant no additional control wafer had to be fabricated. Eight wafers were fabricated for this run with four of each relative boron dose. Of these four wafers two of each were fabricated using a high and low JTE dose, the relative values are given in the table. For this run two junction implant energies were used to see if this would modify the gain as required. There are over 100 devices per wafer. Following IV measurements the lower energy implant provided no working detectors thus are not discussed in this thesis. An example of a wafer fabricated using this mask is shown in figure 4.2, where all devices are round the outer perimeter of the wafer.

Wafer run 5 used results obtained from run 4 to produce a new optimised mask design (M3) with a range of pad sizes and pixel arrays. The mask set M3, included designs compatible with bump bonding to the TimePix3 [59] readout chip and ALTIROC readout chip [5], for the high granularity timing detector (HGTD). There are over 50 device designs

Table 4.2: Fabrication run 4, showing process parameters used

Wafer Run	Pad Size (mm)	Relative Boron dose	JTE Dose	Relative Phosphorus Implant Energy
4	5x5	1.1, 1.3	1, 2.5	0.3, 1
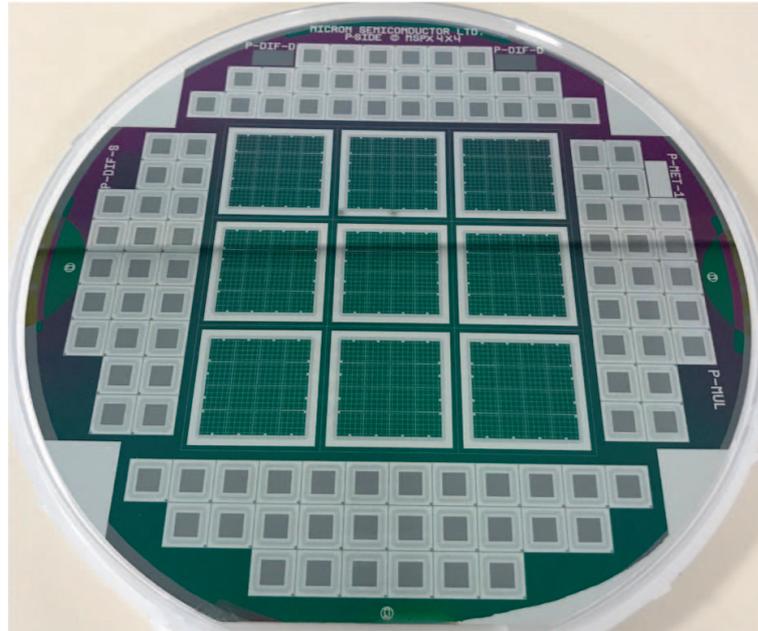
Figure 4.2: Example of a finished wafer from run 4.

Table 4.3: Mask set M3 devices fabricated.

Pixel Pitch	Arrays	Compatible with Bump Bonding
55 μm	256x256	Yes
110 μm	128x128	Yes
220 μm	1x1, 2x2, 3x3, 5x5, 64x64	Yes
0.5 mm	1x1, 2x2, 3x3, 5x5	No
1 mm	1x1, 2x2, 3x3, 5x5	Yes
1.3 mm	1x1, 2x2, 3x3, 5x5	Yes
2 mm	1x1, 2x2, 3x3, 5x5	No
5 mm	1x1, 2x2, 3x3	No

per wafer and the overview is shown in table 4.3, showing pixel sizes and arrays available. The devices with a pixel pitch of 55 μm, 110 μm and 220 μm (array of 64x64) are for use with the Timepix3 chips. Table 4.4 shows the process parameters used in this run. Figure 4.3 shows an example of a wafer fabricated from run 5. The entire wafer was dedicated to this project. The central region was mostly used for Timepix3 compatible detectors as the mask set was funded by Diamond Light Source[1] in Oxford who required these detectors.

---

[1]Diamond Light Source Ltd, Diamond House, Harwell Science and Innovation Campus, Didcot, Oxford-shire, OX11 0DE, https://www.diamond.ac.uk

Table 4.4: Fabrication run 5, showing process parameters used

Wafer Run	Relative Boron dose	JTE Dose	Phosphorus Implant Energy
5	1.1, 1.3	1, 2.5	1
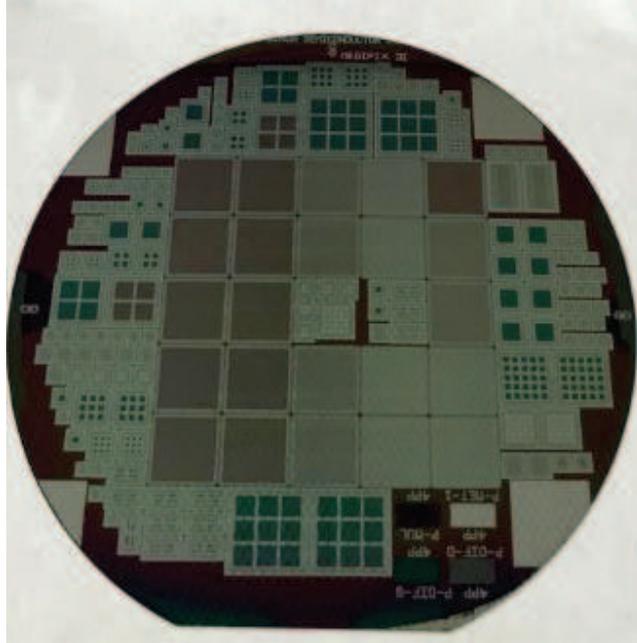
Figure 4.3: Example of a finished wafer from run 5.

4.2   Electrical Characterisation

After fabrication a series of standard visual inspection process quality checks are performed on the wafer. These checks are performed using a microscope to locate any possible defects on the surface of the wafer. It should be noted that during the fabrication run visual inspections are routinely performed to check for alignment issues and other defects. Possible problems that occur involve the etching of metal layers, either under or over etched. As well as alignment problems with the photoresist layer, including under or over developed. If these issues are not checked throughout the process the final result may have fundamental problems.

Electrical characterisation is the first step in any device characterisation study. This reduces down to two essential measurements. A bias voltage vs current measurement (IV) and a bias voltage vs capacitance (CV) measurement. These measurements determine the level of leakage current, breakdown voltage and full depletion voltage.

4.2.1   IV

For sensors in this work we require a low leakage current as the sensitivity of the detector is determined by the signal to noise ratio (SNR). The leakage current is the number of

randomly generated charges where no external excitation is provided. The total leakage current is the result of a collection of events; the diffusion current present due to the concentration gradient present in the semiconductor material, generation current caused by thermally generated electron–hole pairs, surface currents and high field effects. All leakage current introduces noise to the total system, where a high leakage current can cause significant noise reducing the device performance [43]. The relative contribution of the leakage current to the noise depends on the geometry of the device as capacitance also has a significant impact on the noise. One possible use of the LGAD detector is for low energy photon detection, thus any added noise provides a threshold constraint on the minimum energy detectable as it becomes impossible to distinguish between signal and noise.

The breakdown voltage is the potential value for which the electric field in the detector becomes so great that the structure can no longer withstand it. At such a point the sensor will become resistive. For our purposes we require a breakdown voltage of ≈ 500 V, where the working voltage will be in the range 200–400 V, but in general it's best to aim for a breakdown voltage as high as possible. LGAD devices have shown to perform with a breakdown voltage greater than 1000 V [36].

Setup for measurement of diode characteristics

In figure 4.4 the electrical setup for the IV measurements is shown. The samples for IV measurements were mounted on a Wentworth probe station in a clean room environment. The sensor was biased from the front side using a low impedance probe and grounded through the backside copper conductive chuck. The leakage current is measured through the same probe that applies the bias voltage, on the front side. Either a Keithley 2410 (1100 V, 1 A) or a Keithley 237(1100 V, 2 mA) were used as the bias source and current readout. The measurements were controlled by a computer over a GPIB bus using Labview and python scripts and written to text files. Data was analysed and plotted using a python script. Measurements of current were taken at 5 V steps from 0 to 1000 V.

IV results

As there was many runs over the course of this project the IV results from the latest two runs will be presented. Figure 4.5 shows a variety of pin diodes fabricated in run which have excellent IV characteristics. All devices presented show a breakdown voltage of greater than 1000 V and a leakage current of $\approx$ 10–100 nA/cm^2. This shows the fabrication process can be well controlled for no gain devices providing a stable platform to work from.

Run 4, as described in the previous chapter, was the fabrication of 5 mm detectors with various design splits per wafer. On each wafer, there was a variety of design variations which were included to establish the most optimal periphery design. The designs included
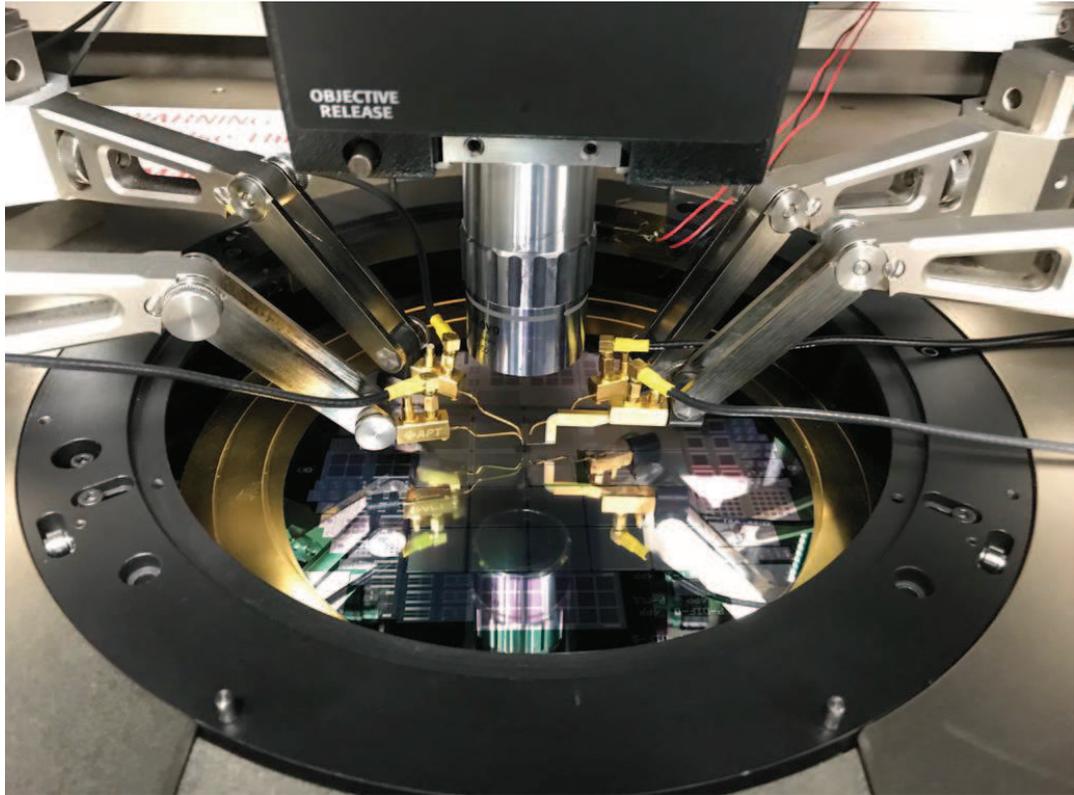
Figure 4.4: Probe station used for electrical measurements.

the addition of a Junction termination extension (JTE), an overlapping junction and various multiplication implant sizes relative to the junction implant size. Also included on the mask were devices which had an overlapping metal field plate, shown to reduce the field at the edge of junction [56]. These would be tested and compared with simulated results. Each wafer split had a variation of multiplication and JTE implant doses.

A typical device performance for all designs is shown in figure 4.7 taken from devices across an entire wafer. Devices with design B tended to have the characteristic of being faulty Detectors A, C, E and G have an overlapping metal field plate of 10 μm. Detectors B, D, F and H have an overlapping metal field plate of 25 μm. The detectors A and B have a multiplication implant which coincides with the inner position of JTE implant. C and D have a multiplication implant which is 10 μm smaller than the inner position of the JTE implant and E and F have a multiplication implant which is 500 μm smaller than the inner position of the JTE implant. Detectors G and H have no JTE implant and are used as controls devices, where the multiplication implant is the same size as detectors A and B. The designs are highlighted in table 4.5 and shown in figure 4.6.

In order to determine the best working detector design the factors considered are the leakage current and breakdown voltage. In general the detectors with the a larger overlap-ping field plate have a lower leakage current, < 100 nA/cm^2, than their equivalent detector by a factor of ≈ 10000. The high current devices may have a reasonably high breakdown
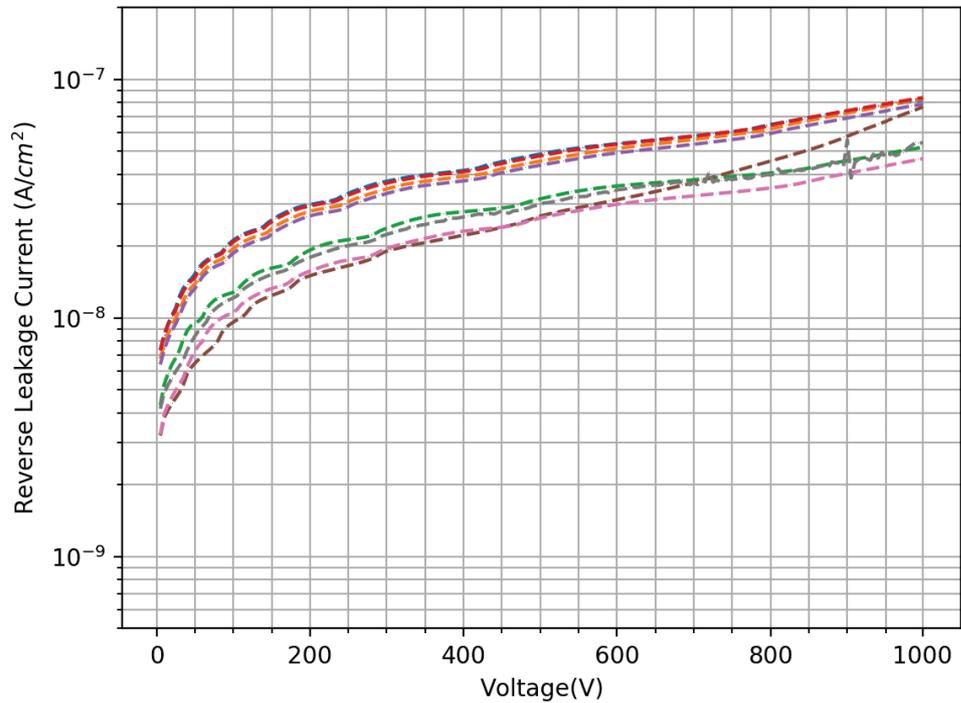
Figure 4.5: Experimental current vs voltage performance of PIN diode (no gain) detectors. Each colour represents an identical PIN diode from a wafer from run 4.

Table 4.5: Mask set M2 devices designed.

Device Letter	Metal Field Plate Overlap	Multiplication Implant size relative to JTE inner edge	JTE
A	10 μm	Coincides	Yes
B	25 μm	Coincides	Yes
C	10 μm	10 μm smaller	Yes
D	25 μm	10 μm smaller	Yes
E	10 μm	500 μm smaller	Yes
F	25 μm	500 μm smaller	Yes
G	10 μm	Same size as A	No
H	25 μm	Same size as B	No

Figure 4.6: Illustrations of device variations on mask set M2, as described in table 4.5.
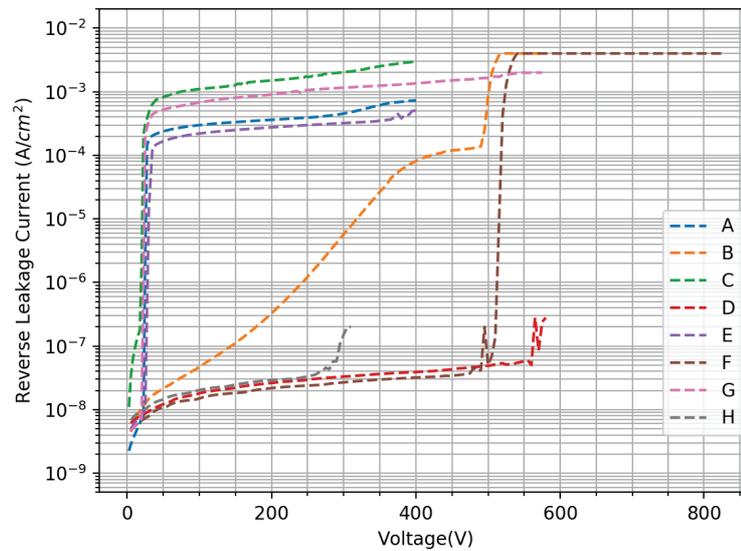
Figure 4.7: Experimental current vs voltage performance of LGAD detectors with a variety of periphery designs. All devices taken from one wafer fabricated using mask set M2. Device B in this plot shows clearly a defect in the device which was typical for B devices.

voltage but are not suitable as they have a large noise due to this high leakage current, > 100 μA/cm^2. Comparing devices with the same overlapping field plate design we find that that the breakdown voltage is highest for the devices with a 10 μm gap between the multiplication and JTE implants of greater than 500 V. For the gap of 500 μm we see similar results with a breakdown voltage of approximately 500 V. If there is no gap between the implants we see a high leakage current but a similar breakdown voltage. The device with no JTE also shows low leakage current but the breakdown voltage is significantly less at approximately 300 V, as expected from simulation. From these results devices with the best IV characteristics were then tested for gain using a variety of techniques discussed in the next section. These were the devices labelled D and F from this run.

For run 5 the results from run 4 were used to create a variety of detector designs as described in section 4.1, with a variety of pad sizes and pixel arrays. All detectors have a JTE and a 10 μm gap between the multiplication and the JTE implants. A variety of JTE widths with different designs were used to evaluate the breakdown performance on JTE width. As discussed there was a variety of device splits for this run, where results of the lower multiplication implant will be presented in this section.

Figure 4.8 shows the IV characteristics of 0.22x0.22 mm^2 and 1x1 mm^2 pads with varying JTE widths. The different JTE widths perform very similarly with the 1 mm pads breaking down at $\approx$ 400 V with a leakage current of 1–10 nA/mm^2. The 0.22 mm pads have a higher leakage current per unit area of 10–100 nA/mm^2 with a breakdown voltage in the region of 250–300 V.
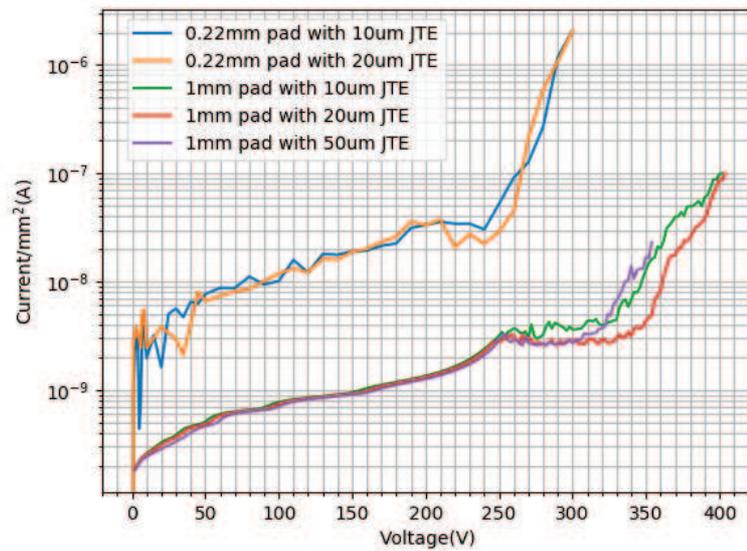
Figure 4.8: Experimental current vs voltage performance of LGAD pad detectors with a range of sizes and JTE widths.

Figure 4.9 shows the IV characteristics of several 2x2 arrays of LGAD detectors. For the 0.22 mm and 0.5 mm pitch arrays the leakage current is lower with the 10 μm JTE with respect to the 20 μm JTE with breakdown voltages of 400–500 V. The two 1 mm arrays have a very similar leakage current with the breakdown voltage being greater for the devices with a larger JTE, $\approx$ 350 V.

4.2.2   CV

To determine the full depletion voltage of these detectors a series of capacitance vs voltage (CV) measurements were performed. The samples were mounted on a Wentworth probe station in a clean room environment, same as for the IV measurements. The front contact was connected to the high potential and the backside to the low potential. The capacitance was measured with an LCR (Agilent 4284A 20 Hz – 1 MHz) meter in parallel mode at a frequency of 10 kHz. The maximum operation voltage by the LCR is 20 V thus an external power supply was used to supply the bias voltage via a special adapter to decouple the high voltage soruce from the LCR meter. Again these measurements were controlled by a Labview program and data stored in text files and plotted using a python script.

The full depletion voltage V_{fd} was extracted from a plot of $1/C^2$ against V by finding the point of interception of two lines fitted to the data before and after the kink in the plot. As well as the full depletion voltage the doping concentration of the bulk material can be calculated using equation 2.23
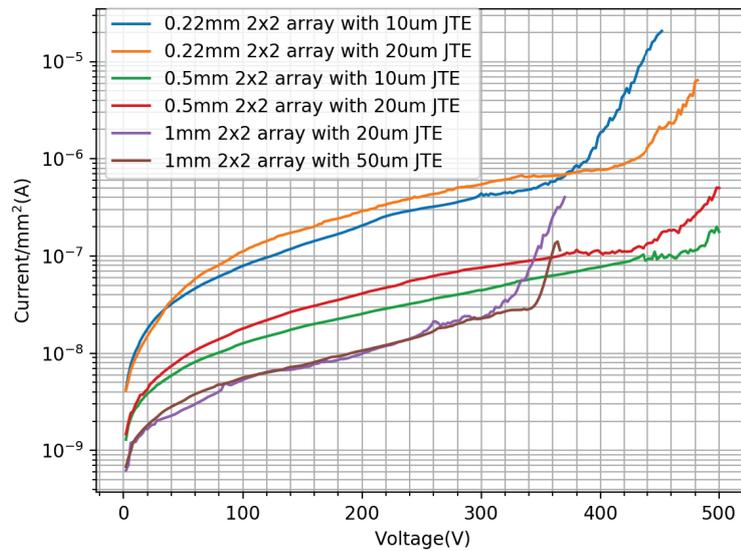
Figure 4.9: Experimental current vs voltage performance of 2x2 LGAD arrays with a range of sizes and JTE widths.

$$|N_{eff}| = \frac{2\epsilon V_{fd}}{qw^2}. \tag{4.1}$$

Substituting

$$C = \frac{\epsilon A}{w}$$

into equation 4.1 and rearranging we get

$$N_{eff} = \frac{2}{A^2 \epsilon \epsilon_0 q_0} \left(\frac{d\left(1/C^2\right)}{dV} \right)^{-1}. \tag{4.2}$$

This method was used to calculate the full depletion voltage and doping concentration of devices in figure 4.10. This figure shows the CV measurement of a PIN diode and an LGAD detector. The doping concentration is calculated using the slope of the curve. This is the same for both detectors as they are from the same wafer and is calculated to be 3.6×10^{12} Atom/cm^3, giving a resistivity of 3.7 k$\Omega \cdot$ cm. The full depletion voltage is calculated to be 108 V for the PIN diode and 138 V for the LGAD. The LGAD full depletion voltage is higher due to the presence of the multiplication implant. This has its own depletion voltage and can be determined from the plot as the first kink and is 30 V, the difference in the full depletion of the detectors.
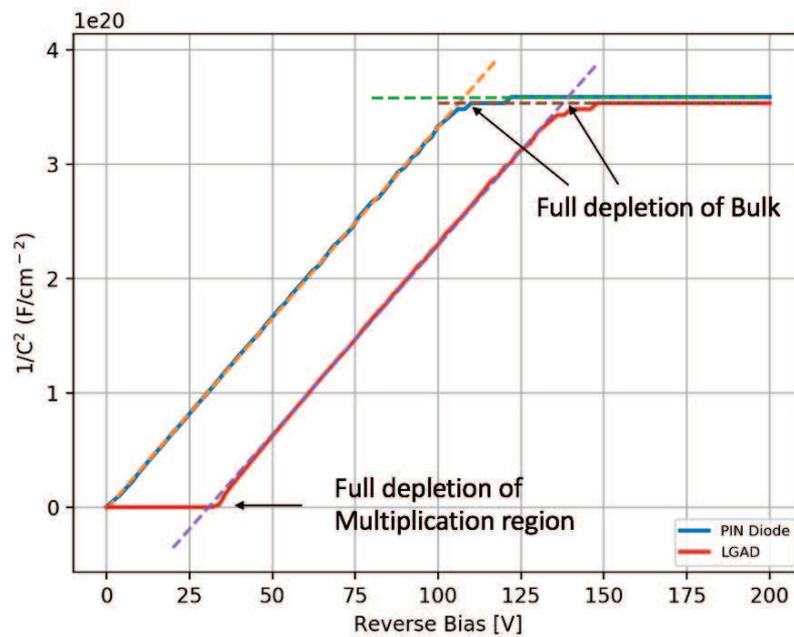
Figure 4.10: Experimental capacitance vs voltage measurement to calculate doping concen–tration and full depletion voltage. The dashed lines show the linear fits used to calculate the full depletion voltage.

4.3   Transient Current Technique, TCT

In order to understand the behaviour of detectors studied in this thesis the TCT set–up has been commissioned [60]. The basic concept of this set–up is to use a laser to generate e–h pairs within the silicon. The induced signal is generated and can be seen using an oscilloscope. The set–up is equipped with two lasers. A red laser of wavelength 640 nm and an infrared laser of wavelength 1060 nm. The red laser is used as it produces e–h pairs along the surface of the detector with a penetration depth of ≈ 5 μm. The infrared laser is used to imitate excitation by a charged particle traversing the detector, as charge is generated uniformly through the bulk of the detector. The lasers can be focused with the use of an optic lens. The position of the lens is controlled by a moving stage, the z-stage. There is also an x and y stage which control the position of the detector relative to the laser. This means it is possible to study charge variation across the detector, as well as to focus the laser. In order to increase the signal to noise ratio, the signal is passed through an amplifier before being readout on the oscilloscope. The reverse bias voltage is applied to the detector via a high voltage source. To prevent any high voltage spikes going to the detector a high voltage filter is connected between the voltage source and detector. Both the reverse bias and readout of the generated signal are connected to the front side of the detector. The backside is connected to ground. In order to separate the bias voltage and RF signal generated by the detector a bias T is used. The bias separates the voltage

from the RF signal, this allows the RF signal only to pass to the amplifier and onto the oscilloscope. To allow monitoring and control of the laser stability, the laser output is split 50:50 by a fiber optic beam splitter. Half of the signal goes to the detector, the other half goes to the beam monitor system. A cooling system (chiller and dry air generator) provides a temperature and humidity controlled environment for measuring the samples. A selection of humidity and temperature sensors are installed in the set–up to monitor any variation. The system is controlled by a PC (DAQ PC) that communicates via GPIB with the bias voltage supply and the oscilloscope. The PC also communicates via USB to the laser control unit which allows the user to set the laser parameters. The PC runs the DAQ (data acquisition) software in Labview supplied by Particulars [2] The basic design of the set–up is shown in figure 4.11.
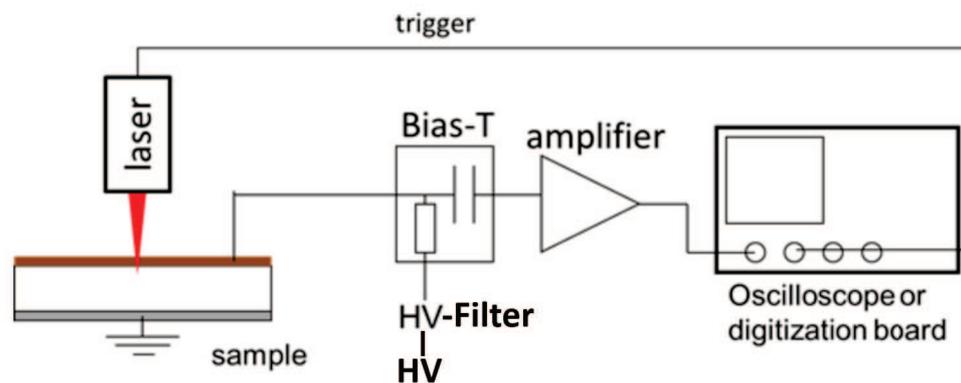
Figure 4.11: Sketch of the Basic TCT setup. In this setup the HV is positively biased as the backside is connected to ground.

4.3.1   Components of the setup

Most of the components of the setup are contained inside a metal box. This is to ensure a dark environment for measurements, preventing photo–currents induced on the detector by external light. It also provides a suitable space to provide a dry air atmosphere, which is required when measuring at low temperatures to avoid condensation on the device. The box and its components are shown in figure 4.12. The stages and laser are controlled remotely by the DAQ PC.

There are three stages used with this setup. Two are coupled to the detector and move as desired in the x and y axis. The other stage is coupled to the laser optics, z axis, is order to focus the laser beam. The movement of the stages can be controlled with a precision of 1/8 μm.

---

[2]Particulars d.o.o., Advanced Measurement Systems, Domžale, Slovenia, SI–1230 Domžale/Dragomelj 154/ Tel: +386 41 423 469, Fax: +386 1 477 3166, www.particulars.si
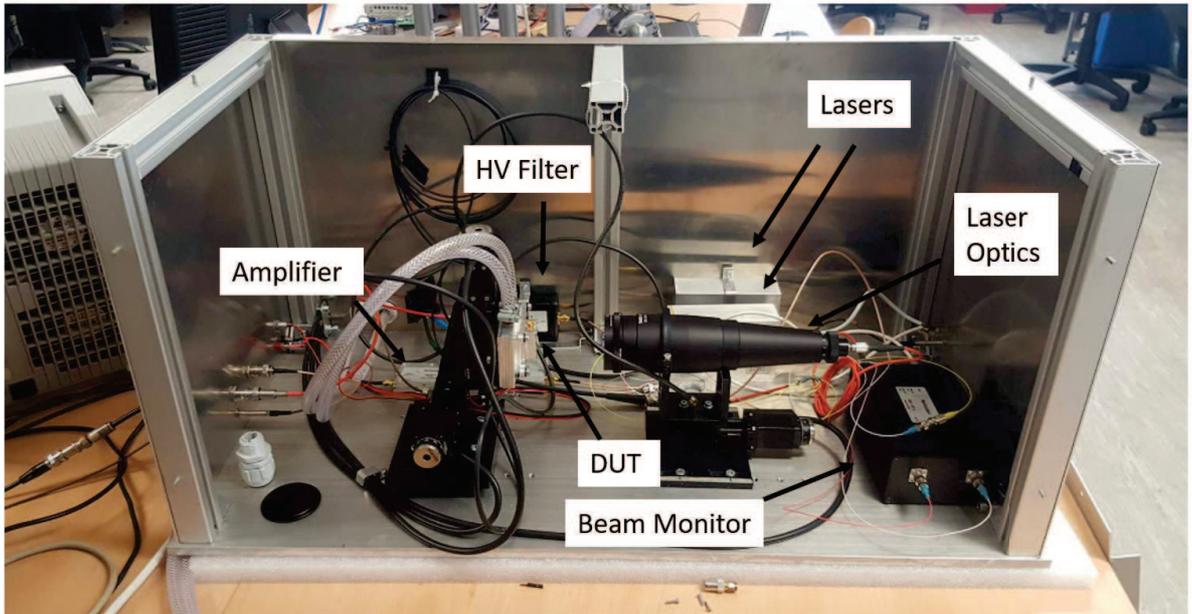
Figure 4.12: Components and setup of TCT system provided by Particulars.

The setup was provided with two lasers, a red (640 nm) and an infrared (1064 nm) as explained in the section 4.3. The laser us pulsed with a pulse frequency range of 50 Hz to 1 MHz. One can also vary the pulse duration from 350 ps to 4000 ps. There is a micro–controller in the control board of each laser which can be programmed via USB with a software called *LaserControl*. Using this program one can set the pulse frequency and duration. It is also possible to program a sequence of pulses however this function has not been used for the work. The laser system is equipped with laser optics. The optics contain a focusing lens and an optional 10 x optical filter to reduce the laser intensity. A trigger signal is sent from the laser to the oscilloscope to synchronize the oscilloscope data acquisition with the pulses generated and the signal collected in the diode under test.

A study in this thesis required the detectors to temperature controlled, in order to measure the temperature dependence, hence a cooling system is required. The TCT setup is provided with a Peltier Element, which is in direct contact with the Aluminium sample support mounted on the x–y stage. The Peltier element has a cold side and hot side. The cold side is in contact with the support, while the hot side is refrigerated using a chiller. In this work the chiller only was used to control the temperature of the sample. The temperature at the detector was measured using two resistance temperature detectors, connected either side of the detector.

An RF amplifier is used, with a 0.01 - 2000 MHz bandwidth, in order to increase the signal to noise ratio. The amplifier used has a gain of 53 dB. This is used in conjunction with a bias T, which can withstand DC voltages up to 2000 V.

Depending on the measurement two different oscilloscopes were used to record the signal. The first is a DRS4 board with a bandwidth of 700 MHz and a sample rate of

5 GSa/s. The second is a Agilent MS09404A with a bandwidth of 4 GHz and a sample rate of 20 GSa/s. The axis scales used in the measurements are mV for the voltage and ns for the time. The signal is triggered with an external trigger pulse from the laser controller via a lemo cable.

The high voltage source used (Keithley 2410) is a low noise DC power supply that provides up to 1100 V. It is able to read the leakage current and can be controlled by the DAQ PC via GPIB.

4.3.2 TCT Operational Parameters

As seen in section 4.3.1 , some components of the setup can be controlled by the DAQ PC. Therefore it is important to understand the parameters in order to achieve the optimized signal. The parameters are controlled by either the laser software (*LaserControl*) or with the *PSTCT* Labview program. The *LaserControl* program intuitively allows the user to change the following laser parameters:

- **Laser frequency:** The frequency of the laser pulses can be set from 50 Hz up to 1 MHz. This should not effect the waveform unless the time between pulses is lower than the charge collection time. The maximum frequency achievable by the software is however only 100 kHz. At this frequency it was noted that the waveform was effected, larger signal, than at lower frequencies. Thus a frequency of 1 kHz is used as standard with this setup.

- **Width:** The width parameter is related to the intensity and duration of the laser pulse. The length of the pulse is directly related to the intensity of the pulse. If the duration of the pulse is low, the intensity is also low and vice versa. The ideal situation is that of a delta pulse to simulate a particle hit. Thus, a compromise is needed between the laser intensity and the width parameter. This parameter is selected in the configuration file and goes from 0 – 100%. The width parameter controls the DAC laser voltage (the corresponding range is 0 – 2000 mV). The DAC voltage is proportional to the waveform amplitude, this obviously depends on the laser and device under test.

The DAQ software is designed to allow automatic scanning measurements to be made by selecting some parameters. These are as follow:

- **Detector Bias Voltage:** The bias voltage can be controlled to perform voltage scan measurements. By varying the bias voltage one can change the position of the depletion region within the detector and of course the electric field across the diode. The electric field is critical in the case of an LGAD device as this controls the gain of the detector, as well as the drift velocity of the charge carriers.

- **Detector Position:** The position of the detector relative to the laser can be controlled by moving the x and y stages. This can be programmed to run raster scans across a detector to look at charge collection efficiency for example.

- **Laser optics position:** The laser optics position can be controlled in order to focus the laser. The ideal scenario is to have as small a laser spot size as possible in order to study charge collected across pixel or strip detectors. The minimum spot size for the red laser is $< 8~\mu$m and for the infrared $< 11~\mu$m.

4.3.3   TCT signal and Analysis

The DAQ PSTCT *Particulars* Labview software enables the user to perform systematic studies of devices. Signals produced by a device and readout by an oscilloscope are saved into a file in a binary format. The data can then be analyzed using ROOT. *Particulars* provides a C++ framework from which one can do most of the basic analysis through ROOT macros. Further analysis can be achieved by altering these macros. The waveforms, generated by drifting charge, were induced by either an Infrared(IR) or red laser. The IR laser illuminated either the front side or back side of the detector in order to simulate a MIP-like particle. The red laser illuminated the backside of the detector to simulate a low energy x-ray. As the penetration depth of the red laser is much less than that of the IR laser the waveforms produced are very different.

Figure 4.13 shows an example of two waveforms generated by illumination of a red laser on the backside of the detector, one from an LGAD and the other from a PIN diode when biased at 300 V. The beginning of both waveforms is very similar, this is because both devices collect drifting electrons at this stage. Once the PIN diode waveform reaches its maximum it begins to fall. This fall is due to the electrons reaching the electrode. The holes generated should be collected almost straight away as they are generated very close to the electrode. However for the LGAD device there is a kink in the waveform. This is the point at which the drifting electrons cross the high electric field region close to the collecting electrode. At this stage the electrons undergo avalanche multiplication where additional e-h pairs are produced. The electrons produced at this stage are collected very quickly, hence the steep rising edge. The generated holes however, are collected much slower as they have to drift to the backside of the detector and the hole mobility and hence the hole velocity is lower than that of electrons.

In order to characterise the LGAD devices one must calculate its gain. This is done by comparing the charge collected with the charge collected by a PIN diode. To obtain the total charge collected (Q) one can simply integrate the current (I) over the charge collection time, $t_{col,e/h}$, for both holes and electrons for a given waveform:

$$Q = \int_0^{t_{col,e}} I_e(t)dt + \int_0^{t_{col,h}} I_h(t)dt \tag{4.3}$$
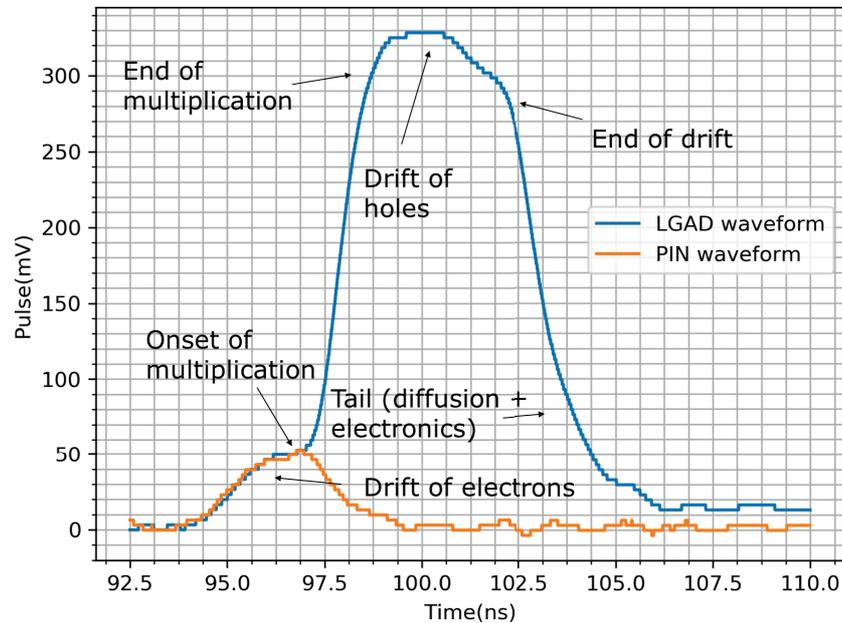
Figure 4.13: Example of backside red laser TCT waveforms from a PIN diode and an LGAD device.

.

The current cannot be directly read from the oscilloscope instead the voltage over a 50 Ω input resistance is recorded. As the current is proportional to the obtained voltage, Ohm's law (V = IR), one can use arbitrary units of mVns for the integrated charge. The waveform is integrated within the ROOT script.

4.3.4 Results

The setup as presented in section 4.3.1 allows the measurement of charge collection at different positions in the detector and at a range of voltages. In order to reduce noise each measured signal was the average of at least 100 waveforms. All measurements presented are the result of the laser illuminating the detector from either the front or backside at 90°to the surface. The laser width parameter was set as convenient for the measurement taken and results were taken at room temperature unless otherwise stated. Due to fluctuations in the laser intensity a beam monitor system was used and described in section here.

Beam Monitor

Laser fluctuations with time make it impossible to compare signals generated from one day to the next. If scanning a detector for relative changes in the charge collection over a small

time period this is not an issue. However for the majority of measurements made in this work it was imperative to compare charge collected over a period of days for a range of detectors. This is to characterise the gain of our LGAD's. Two approaches were used in order to obtain reliable results for gain measurements taken. The beam monitor approach, suitable for both IR and red lasers, and the signal fitting approach suitable for use with backside illumination by the red laser.

Beam Monitor System

The beam monitor system consisted of a parallel detector system with that of the device under test (DUT). The laser signal was split using a 50:50 fiber optic beam splitter, where 50% of the signal illuminated the device under test and 50% illuminated the beam monitor detector. The beam monitor detector was a 5 mm, low leakage current, high breakdown voltage PIN diode produced in run 4. The beam monitor diode was always biased to 200 V. The readout system for this beam monitor diode was an exact duplicate of the DUT readout system, with a bias T and amplifier. The HV supply was also the same.

A mount was made to hold the laser a well defined small distance away for the beam monitor diode in order to keep the focus the same for all measurement taken. The diode was illuminated from the back side due to the absence of metal, allowing light to pass easily to the detector bulk. The generated signal was passed to the oscilloscopes so that both this signal and the signal from the DUT could be seen simultaneously. The measurement required was then performed, where the signal for both devices was saved and analysis performed. The charge collected for the beam monitor diode was compared with the charge collected for the DUT to obtain a normalised charge. In order to calculate the gain of an LGAD detector this process was repeated for a both a PIN diode and an LGAD detector. The ratio of these measurements then gave the gain value. To determine the reliability of this process a long term study was performed to look at the charge collected for both the DUT and the beam monitor diode. The DUT in this case was a PIN diode. The charge collected for both was plotted as a function of time. The ratio of these charges were then plotted, to ensure they fluctuate at the same rate. The resultant plot should be a straight line as shown in figure 4.14.

Signal Fitting Approach

Due to the nature of the signal produced by backside illumination by a red laser it is possible to fit a signal produced by a PIN diode and an LGAD. The signal generated due to the drift of electrons from the backside to the front-side of the detector is identical for both detectors. The difference occurs near near the surface of the LGAD where avalanche multiplication occurs. Thus any change in the laser intensity between measurements can be addressed offline by fitting the signal from the PIN diode to the shape of the LGAD
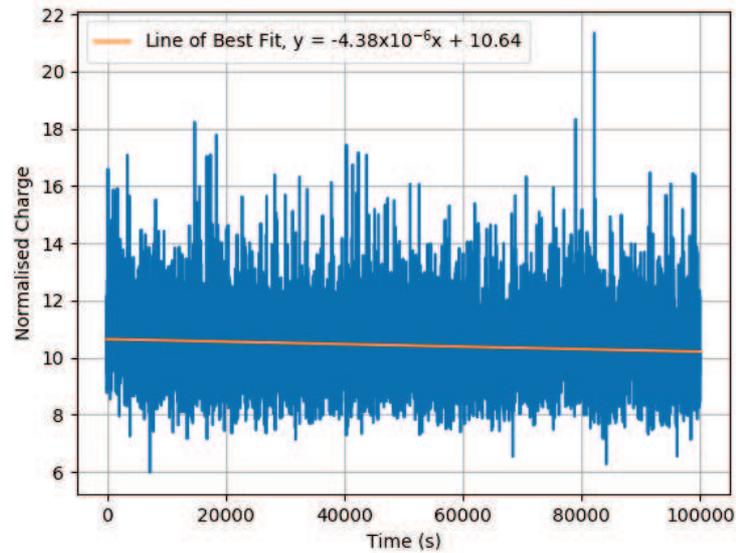
Figure 4.14: Normalised charge as a function of time from a PIN diode. Where the normalised charge is equal to the charge from the DUT divided by the beam monitor charge. This shows that normalised charge is stable with time as expected with a working beam monitor set-up.

.

signal. This has been done and the result shown in figure 4.13, where the PIN diode has been made to fit exactly the LGAD signal before the onset of multiplication. This is achieved by modifying the signal height in ROOT, no change to the time component is need as this is not effected by laser fluctuations.

An issue with this technique is the need for the manual alteration of signals which introduces errors. There is also a limit to this technique as it is only suitable for back–side illumination by a red laser, thus the beam monitor technique is the preferred option. However it can be useful if a beam monitor system is not available.

4.4   Charge collection studies using TCT

This section presents an overview of the results obtained using TCT to calculate the gain in a range of LGAD detectors. The results presented will be from various fabrication runs where the charge collected is measured as a function of voltage and compared with the charge collected from a PIN diode to calculate the gain as described in section 4.3. Unless stated the laser was focused and positioned in the centre of the active region of the detector. Both red and IR lasers were used in this work and results for each shown.
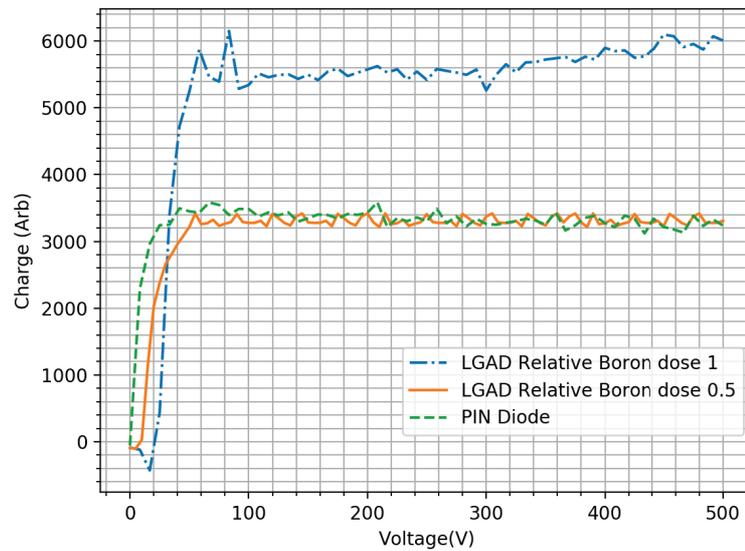
Figure 4.15: Charge as a function of voltage

4.4.1 Run 1

As described in the devices section, wafer run 1 had a range of pad sizes with two mul‐tiplication implant doses. Results for the two doses are plotted in figure 4.15, where the relative dose of 0.5 shows no gain. The device with a relative dose of 1 shows a small amount of gain, as the charge collected is greater than that of the PIN diode. Once calcu‐lated this gives a gain of up to 1.7. From simulation this dose was expected to give a gain value of between 1 and 2. The charge collected was induced by a red laser on the front side of the detector.

4.4.2 Run 2

From the results of run 1 it was clear the need to increase the multiplication implant dose in order to create high enough gain. Devices were fabricated using the same mask design but with relative doses of 1.3, 1.6 and 1.9 with respect to the standard doses from run 1. From initial current‐voltage measurements it was clear that the gain from the wafers with a dose of 1.6 and 1.9 was too high, as these had a breakdown voltage of $\approx$ 25 V. Therefore all TCT measurements were carried out on the devices from the wafer produced with a relative boron dose of 1.3. As with run 1, back side illumination by a red laser was used to induce charge and thus calculate gain with the signal fitting approach being used. Figure 4.16 shows the gain vs voltage results for four detectors. Two of which are 2x2 mm^2 and the other two are 1x1 mm^2 pad LGAD's. For each pad size the relative size of the multiplication implant is 75% or 80% of the the junction implant size. This has little
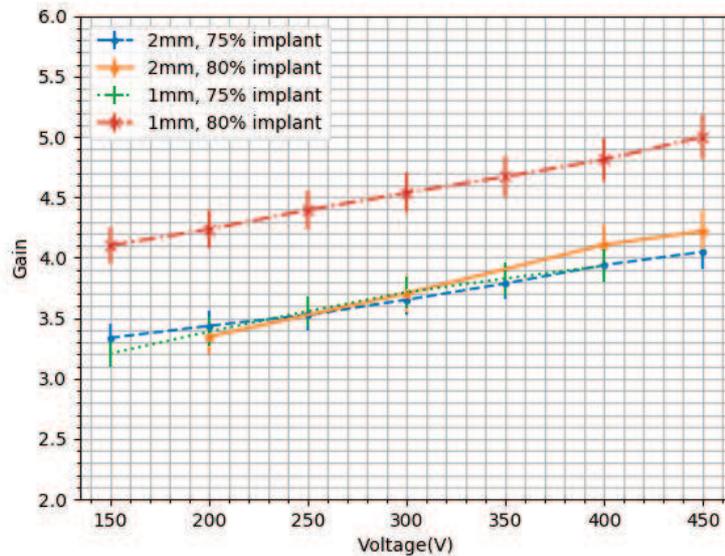
Figure 4.16: Gain as a function of voltage for run 2 LGAD's of 1x1 mm^2 and 2x2 mm^2 active area. These show the same gain within the normal process variation.

effect on the gain performance of the device as all devices show a gain in the region of 3–5. However the IV characteristics change significantly. The current per unit area against voltage is plotted in figure 4.17 and shows that the devices with a multiplication implant of 75% of the junction implant size have a larger leakage current than those with a relative size of 80% of around 10 times greater. However the breakdown voltage for all is around 400 V.

4.4.3 Run 4

The results from run 2 provided significant gain just below the required value of 5–10. It was at this stage that a new mask was designed in the hope to improve the electrical performance and to increase the gain of the detectors. This mask design used the simulation information to incorporate a JTE design around each pad detector. A variety of process splits were performed in order to evaluate the most efficient parameters for both multiplication implant dose and JTE dose. The IV characteristics have been described in section 4.2.1 for the lower gain devices of run 4. The multiplication implant doses used in this run have relative values of 1.1 and 1.3. All detectors have a pad size of 5 mm^2.

For this run a beam monitor system was used to help calculate the gain of the LGAD's using an IR laser which illuminated the front side of the detector. Various measurements were performed for different devices across the wafer. This section focuses on the higher boron dose of 1.3. Figure 4.18 shows a variety of LGAD pad diodes with the gain plotted as a function of voltage. These measurements were taken at 20°C. The gain increases as
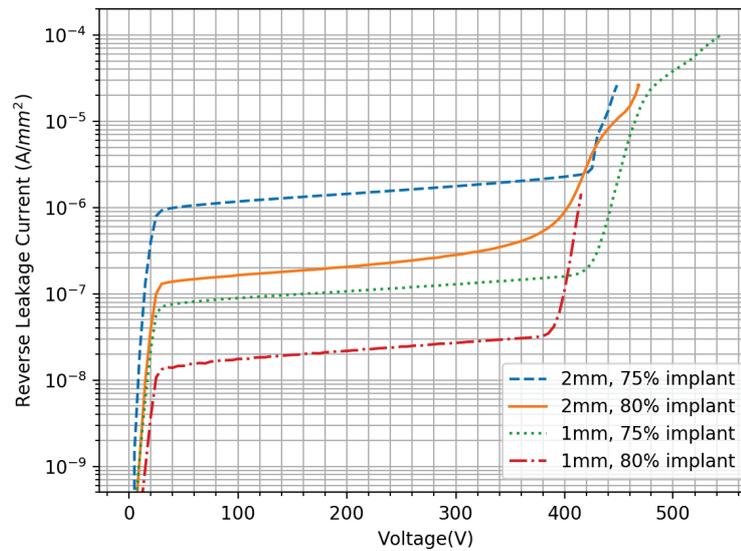
Figure 4.17: Current as a function of voltage for run 2 LGAD's of 1x1 mm^2 and 2x2 mm^2 active area.

a function of voltage as expected for all detectors, however there is a discrepancy in the gain from one device to the next. With an $\approx$ 10% variation at 300 V across these devices. The devices tested were from across the entire wafer. This highlights the need for a very controlled process, as this gain variation is most likely caused by small changes in the doping profile of the gain region.

Following gain measurements at 20° C a study was done to monitor the gain as a func-tion of temperature for three detectors. This was performed to see the effect of temperature on the gain and breakdown voltage as these devices are to be designed to be used in the ATLAS experiment which operates to a temperature down to –30° C. It has been shown that LGAD exhibit small variations in gain and breakdown voltage with temperature as described in [61]. This variation is predicted as the temperature changes the saturation ve-locities of charge carriers and the impact ionisation rates [62]. The decrease in temperature is predicted to have an increase in gain and a decrease in breakdown voltage.

The results of this study are shown in figures 4.19a, 4.19b and 4.19c. As mentioned in the section 4.3 it is possible to control the temperature of the DUT using a Peltier and cooling unit. For this study only the cooling unit was used. This provided cold water to a base plate up which the device was mounted. Temperature was monitored at either side of the device to get a reliable reading. To ensure no condensation built up on the device dry air was introduced into the box. The humidity was measured to ensure no danger to the device. Measurements of gain were taken in a voltage range of 150–350 V at temperatures of 15° C, 20° C and 25° C. For each voltage 100 measurements were taken, the standard deviation was calculated and plotted as the error.
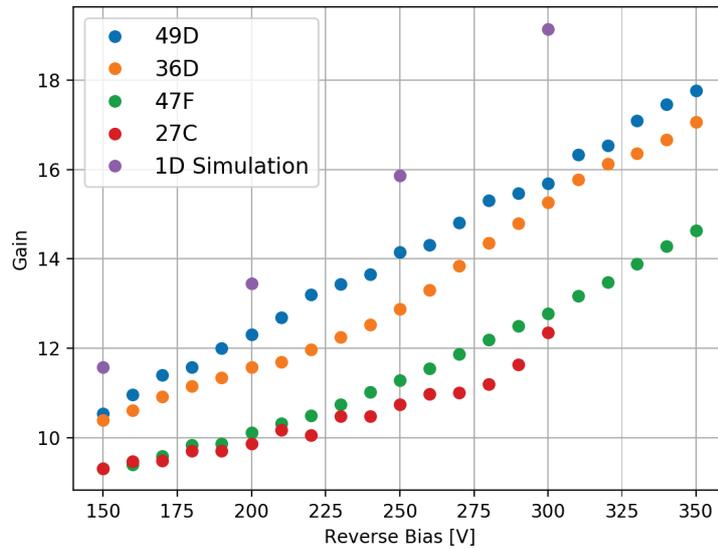
Figure 4.18: Gain as a function of voltage for a variety of devices from run 4, with a relative boron dose of 1.3. A 1D simulation of gain is also plotted for comparison with the fabricated devices.

From the results obtained there is little evidence that the temperature at this range has much effect on the gain performance of the device. Figures 4.19b and 4.19c seem to show the gain increasing with decreasing temperature, however the measurements obtained in 4.19a are not so convincing. For a clearer comparison it may be useful to test these at lower temperatures, to see a more obvious change in gain. It's clear that the leakage current is reduced with decreasing temperature by roughly 50% every 7° C. This becomes important when doing timing measurements described in chapter 5, where we want low noise detectors.

4.4.4   Run 5

The results from run 4 provided the optimised gain required for this work for pad devices with an active area of 5x5 mm^2. As the goal of this work was to produce pixellated LGAD detector for use with the Medipix readout chip, a mask was designed with multiple pad sizes and a variety of array's from 2x2 up to 256x256 as described in the devices section. The mask was designed using information gathered from run 4, in terms of multiplication implant size and JTE dose. A number of JTE widths were used to see the effect on electrical performance shown in section 4.2.1. The same boron doses were used with run 5 as for run 4.

Figure 4.20 shows the results of gain against voltage for a variety of 1x1 mm pad devices for relative boron doses of 1.1 and 1.3. All devices shown have a JTE width of 50 μm. The
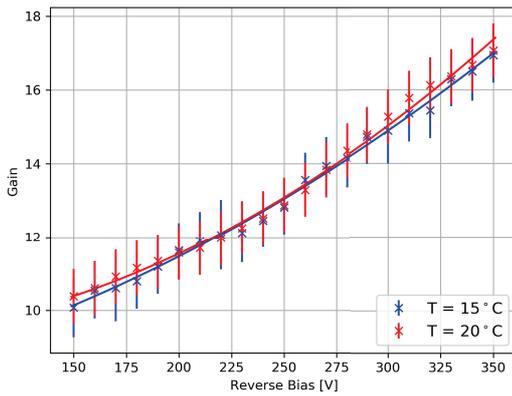
(a) Gain as a function of voltage and temperature for device 36D with a relative boron dose of 1.3
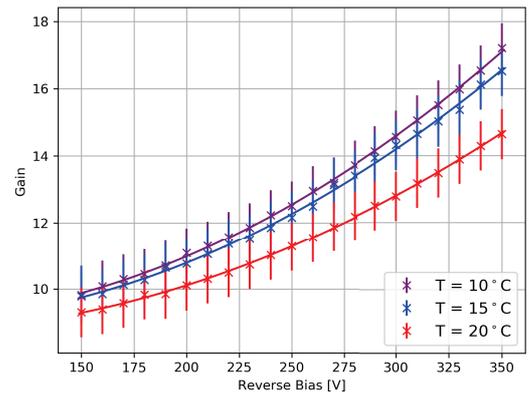
(b) Gain as a function of voltage and temperature for device 47F with a relative boron dose of 1.3
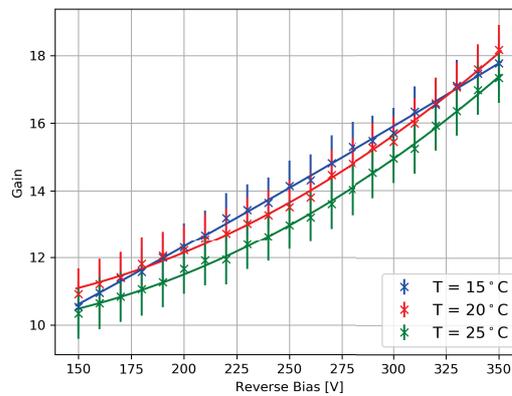
(c) Gain as a function of voltage and temperature for device 49D with a relative boron dose of 1.3

Figure 4.19: Temperature dependence of gain for LGAD detectors.
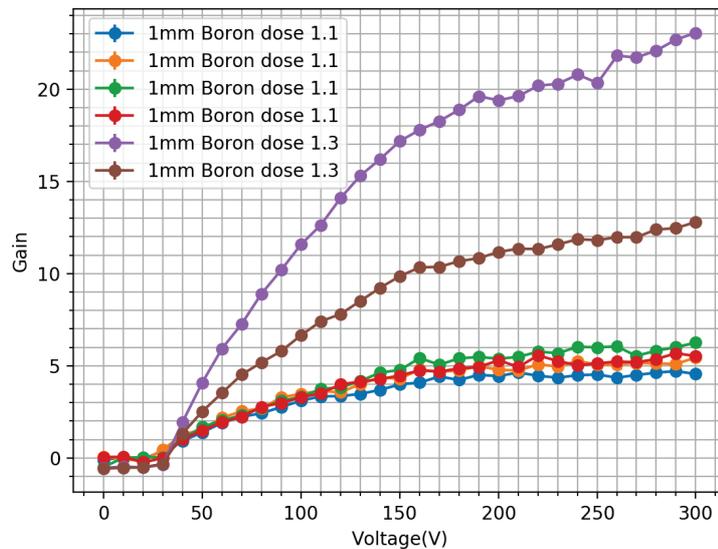
Figure 4.20: Gain as a function of voltage for a variety of 1 mm pad devices from run 5, with a relative boron dose of 1.1 and 1.3.

devices fabricated with relative boron dose of 1.1 provide a gain of ≈ 5 for all devices shown, with small a small error $< 10\%$. However the devices with a relative boron dose of 1.3 show a large variation in gain. One device has a gain of 12 at 300 V while the other has a gain of 23 at the same voltage. This is somewhat not unexpected as the IV results for these devices vary dramatically across the wafer. Where the yield for this wafer is very low. Comparing the results with run 4, we see that the gain is very similar for the relative boron dose of 1.3 of lower gain but obvious discrepancies if you compare the entire wafer. Hence all further measurements are performed using the wafers with a relative boron dose of 1.1.

Devices were fabricated with a range of pad sizes and gain results compared from one wafer with a relative boron dose of 1.1 with both higher and lower JTE doses. Figure 4.21 shows the gain as a function of voltage for pad LGAD's with an active area of 0.22x0.22 mm^2, 0.5x0.5 mm^2, 1x1 mm^2 and 2x2 mm^2. All devices have a JTE width of 20 μm. The devices with an active area of 0.5x0.5 mm^2 and 1x1mm^2 are from the wafer with the lower JTE dose and the 0.22x0.22 mm^2 and 2x2 mm^2 devices from the wafer with a higher JTE dose. The gain achieved by the 0.5x0.5 mm^2 and 1x1 mm^2 is very similar, however the gain for 2x2 mm^2 and 0.22x0.22 mm^2 varies quite a bit. This is the first indication that the size of pixel can have an effect on the gain. The gain can also vary from one wafer to the next due to small changes in doping concentration, as well as across the same wafer. The effect of the size of pixel on gain will be briefly discussed in the following sections of this chapter, but more in depth discussion is present in chapter 6.

Measurements were then taken for 2x2 mm^2 and 0.22x0.22 mm^2 pads with different JTE
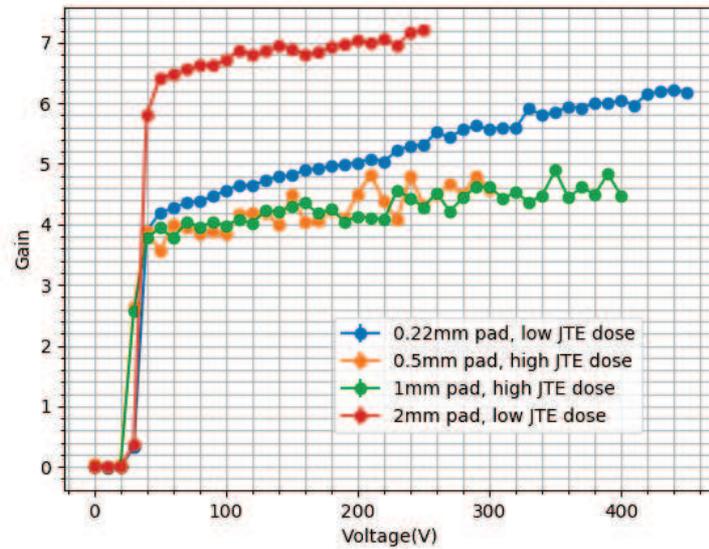
Figure 4.21: Gain as a function of voltage for a variety pad sizes from two wafers with relative boron dose of 1.1 from run 5.

widths, 10 μm and 20 μm. The results are shown in figure 4.22 where it's clear that there is an effect on the gain with JTE width at small pad size. The gain of the 1x1 mm pads is 5-7 for both JTE widths however this is not the case for 0.22x0.22 mm^2 pads. The pad with the 20 μm JTE has a lower gain with respect to the 10 μm by around 25%. This poses a problem when trying to fabricate small pixel LGAD's and this was investigated with the use of TCAD simulations. The results are presented in chapter 6, for small pixel LGADs.

First measurements of LGAD arrays using TCT were performed to characterize a 3x3 LGAD array of 0.5x0.5 mm^2 pixels. A raster scan was performed on the entire array to measure gain uniformity, shown in figure 4.23. The full array was wire bonded to the same readout channel, thus charge collected at a given laser position is the sum of the charge from all the pixels. Therefore no information can be given for charge sharing. Gain regions, in red, are surrounded by an aluminium electrode, where no charge is collected, in blue, due to the light being reflected by the electrode. Charge collected between pixels shows no gain as the electrons are collected through the JTE of the surrounding pixels. A reasonable gain uniformity is shown from pixel to pixel. The gain from the pixels is within a reasonable level to the pad detectors presented in figure 4.21, with the high dose JTE implant.
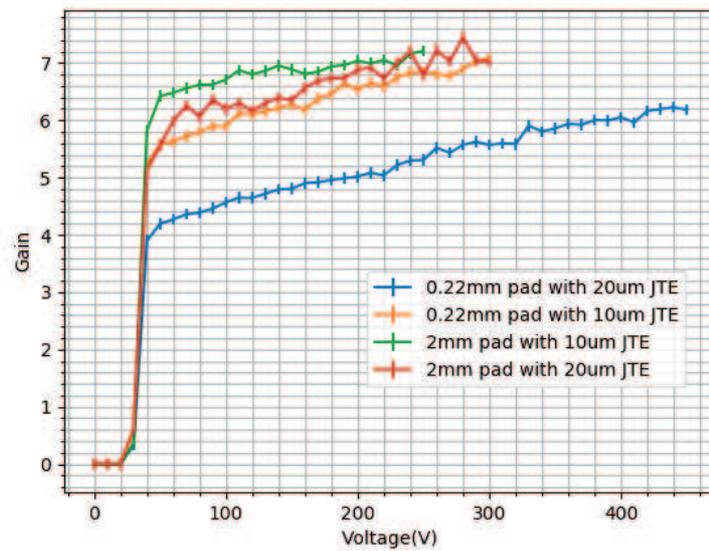
Figure 4.22: Gain as a function of voltage for LGAD pad sizes 0.22x0.22 mm^2 and 2x2 mm^2 from a wafer from run 5 with a relative boron dose of 1.1.
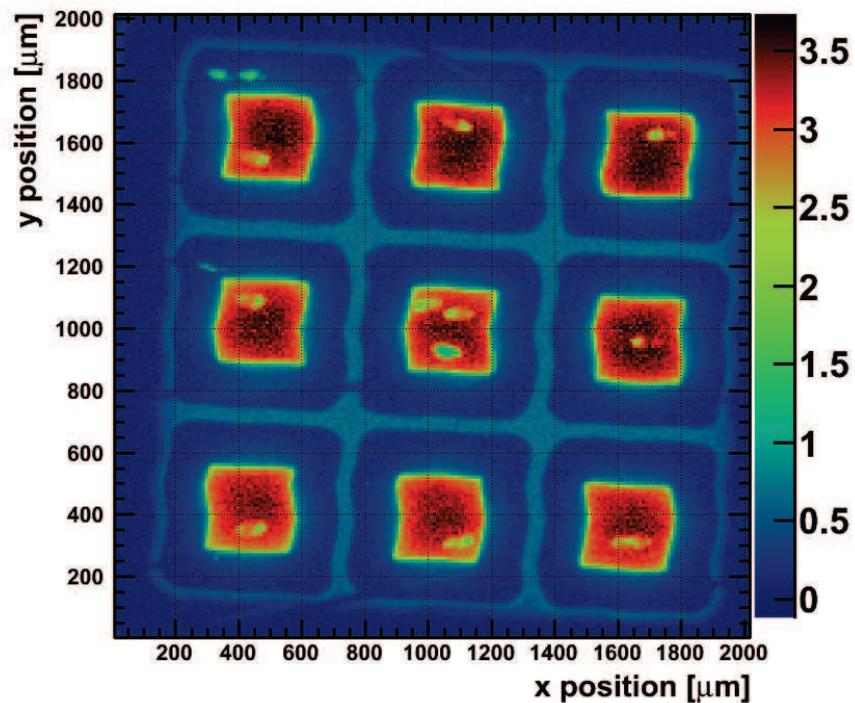
Figure 4.23: Raster scan of a 3x3 LGAD array of 0.5x0.5 mm^2 from a run 5 wafer with a relative boron dose of 1.1, showing gain as a function of laser position.

4.5 Characterisation using radiation sources

This section introduces the use of radiation sources as a characterisation tool for LGAD detectors. Alpha particles were used to measure the gain, similar to the TCT method. The technique and results will be described in section 4.5.1. Important for our work is the detection of X–rays, specifically at low energy. As these detectors have not been yet optimised for low energy X–ray detection a Variable X–ray source was used to determine the minimum detectable energy of these LGAD's, specifically from run 4. These results and setup are described in section 4.5.3

4.5.1 Alpha Particles

An alpha particle is a type of heavy charged particle which interacts via the Coulomb force with detector material causing ionisation and thus a charge deposition. It consists of two protons and two neutrons bound together, identical to a helium nucleus. The *specific energy loss* of the alpha particle with distance in a detector is described by the *Bethe* equation [25]

$$S \equiv -\frac{dE}{dx} = \frac{4\pi e^4 z^2}{m_e v^2} NZ \left[\ln \frac{2m_e v^2}{I} - \ln \left(1 - \frac{v^2}{c^2}\right) - \frac{v^2}{c^2} \right] \tag{4.4}$$

where v and $z{\cdot}e$ is the incident velocity and particle charge respectively. The electron rest mass is given by m_e, the target materials' number of atoms per unit volume given by N and atomic number is Z. The average ionisation of the detector material is given by I [63]. A plot of the specific energy loss with distance is known as the *Bragg curve*. The charge deposited has a specific depth at which the majority of the charge is lost. This distance is know as the penetration depth and varies with energy. For an alpha particle of energy 5 MeV the penetration depth is ≈ 4 μm in silicon. This makes it ideal for testing LGAD's as low energy X–rays have a similar penetration depth, but with less deposited charge.

Components of Setup Run 1

A similar approach to the TCT setup was used for the Alpha measurements. The same readout system was used, which consisted of a bias T and amplifier, for these measurements. The DUT was mounted inside a small vacuum chamber to reduce loss of energy by Alpha particles in air. The device was connected to the readout and HV supply via a SMA connector on the lid of the chamber to ensure a tight air seal. An Am^{241} source was used to produce the required alpha particles of 5 MeV. This was mounted on the backside of the detector to induce charge near the backside electrode. It was also mounted on the front–side for comparative reasons.

As the production of alpha particles is by a random process, the measurements were

(a) A PIN diode at 300V.

(b) An LGAD device at 300V.
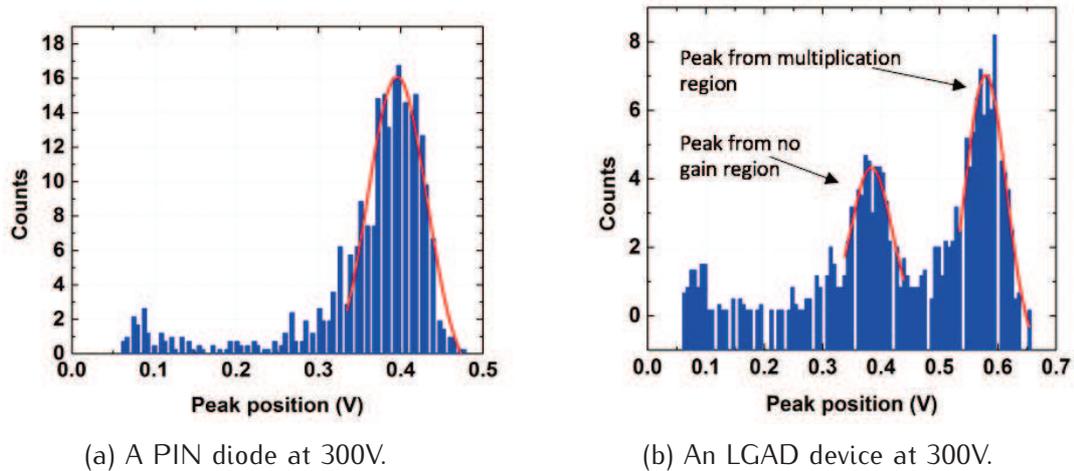
Figure 4.24: Numbers of counts as a function of peak position for a PIN and LGAD detector

taken over a period of hours to enable a sufficient number of counts to be collected. The oscilloscope was programmed via a Labview script to record a waveform every time the signal was triggered. The trigger was set to 10 mV above the noise level. A histogram of peak position was recorded and measurements were taken for a range of voltages.

Alpha Results Run 1

Alpha measurements were performed during the testing of run 1 detectors. An overview of the results will be presented here. Measurements were taken for both PIN diodes and LGAD devices. An example of the histogram of peak position for both detectors are shown in figure 4.24 . The PIN diode biased to 300 V in figure 4.24a shows the peak position at $\approx$ 0.4 V. However for the LGAD detector in figure 4.24b two peak positions are clearly shown. One of these peaks, the lower peak, coincides with the peak position of the PIN diode. The other peak has a larger value and the number of counts in this region is greater. In the LGAD the collection of generated charge carriers is through the multiplication region but also through the overlapping junction region. This is because the e-h pairs generated by the Alpha particle covers a large area depending on the incident particle. As the electrons drift to the collecting electrode, they do so perpendicular to their initial generation site. The number of counts in the no gain peak is lower because the relative size of this region is smaller with comparison to the gain region. One can calculate the gain of an LGAD detector using Alpha particles without having to compare results with a PIN diode. To be clear the 2 peaks occur from the interaction of different particles.

The gain was calculated for this device by Alpha illumination from both the front and backside. The peak position vs voltage is plotted in figure 4.25 , where the Planar detector is the PIN diode. This method shows a gain of 1.44 at 300 V, which is the same as
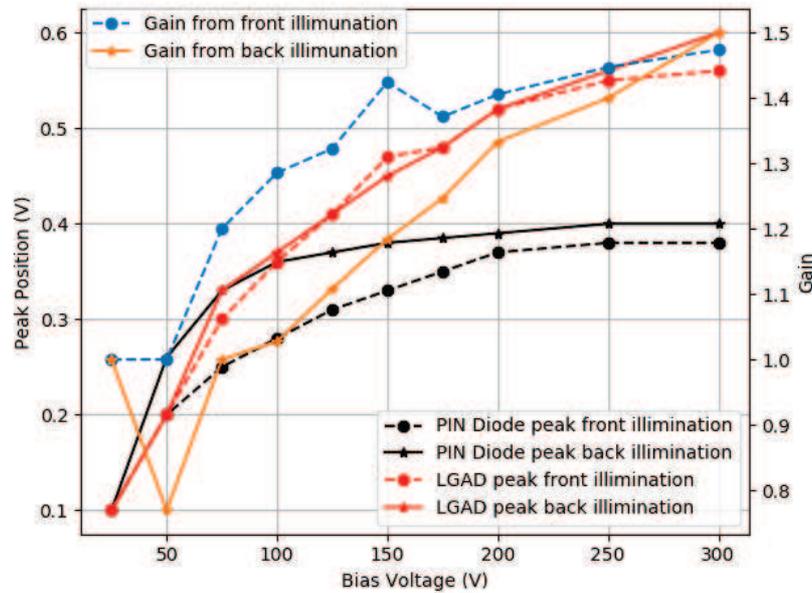
Figure 4.25: Peak position against voltage for a PIN diode and LGAD detector for front and backside illumination of alpha particles.

that obtained using the TCT method for the same detector. It can be seen that the peak position is slightly higher at 300 V for devices illuminated from the back than from the front. This could be due to the distance from the source. Due to mounting constraints, when illuminated from the front side the alpha source is slightly further from the detector with respect to the backside. Due to this a spreading in the generated charge carriers within the detector medium can be produced. The spreading can be such that the charge carriers are not collected, thus a small decrease in the signal generated. This decrease is the same for both PIN and LGAD detectors thus the gain calculated is the same.

Components of setup Run 4

A period of time during this PhD was spent at CNM in Barcelona, to characterise LGAD devices. Alpha measurements were performed using a different set up than the one used in Glasgow. The set up was of similar design however the signal's generated were fed through a multi channel analyser (MCA), thus only the peak information could be measured, not the entire waveform. The devices were mounted in an air tight box under vacuum, connected to a HV supply, where the signal generated was passed through an Amptek amplifier to the MCA. The source was placed above the device for all measurements. This setup was used for Beta measurements where the results are presented in section 4.5.2.
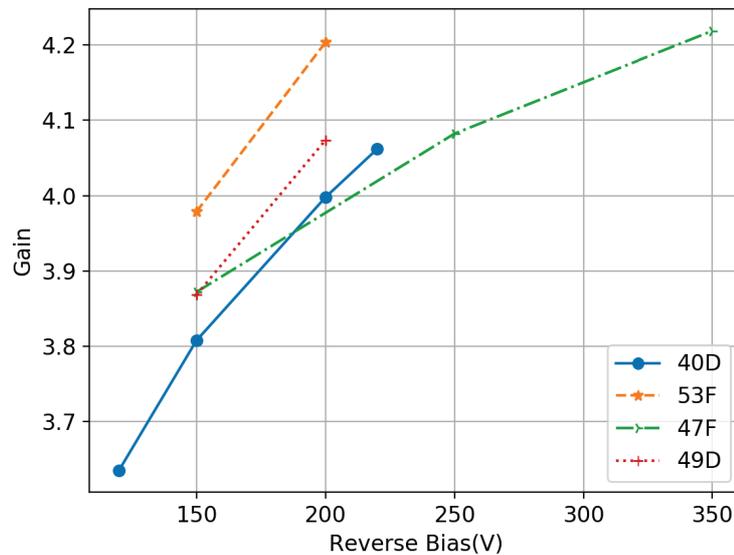
Figure 4.26: Gain against voltage for LGAD detectors illuminated by a Tri–alpha source.

Alpha Results Run 4

For these measurements a Tri–alpha source was used with isotopes Am^{241}, Pu^{239} and Cm^{244} where the central peak which had an energy ≈ 5 MeV. Due to the nature of the Tri–alpha emission three peaks were observed on the MCA, the central peak was used to calculate the gain. Again the gain was calculated by comparing the central peak of the LGAD to that of a PIN diode. Figure 4.26 shows the gain against voltage for a range of devices. The gain, appears to be lower than that calculated using the TCT method, which is ≈ 4 at 200 V for all devices, which calculated a gain of 10–12 at this voltage. The difference in gain could be attributed to the analysis method used. By the TCT method, the total charge collected was, while the Alpha measurements only measured the peak position. Typically the signal induced in the LGAD is longer and higher. The longer signal is due to the extra holes generated at the multiplication region. When using the peak position as a method for calculating gain, those extra holes can be missed, thus a lower gain is seen compared with TCT.

4.5.2 Beta

Strontium ($^{90}Sr_{52}$) source was used as a Beta source for characterisation of LGAD devices. A Beta particle, β , is a high-energy, high-speed electron or positron emitted by the radioactive decay of the Sr_{90} nucleus. The $^{90}Sr_{52}$ decays to Yttrium ($^{90}Y_{51}$), emitting a β^- particle. This then decays to Zirconium ($^{90}Zr_{50}$), releasing another β^- particle.

The energy of the source used in this work has a maximum energy of 2282 keV. This
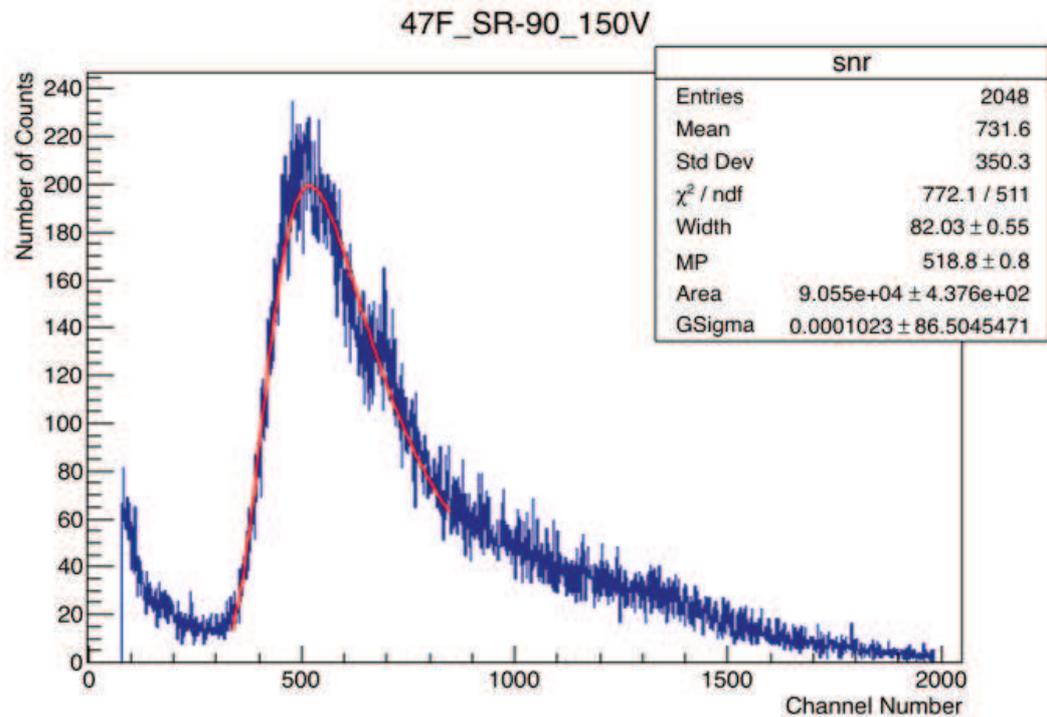
Figure 4.27: Histogram with fit of ^{90}Sr spectrum from an LGAD detector.

means the energy loss in the detection medium is dominated by ionisation. Energy loss at higher energies becomes dominated by *Bremsstrahlung radiation*. As the β^- can have a range of energies from 0–2282 keV a spectrum is plotted where the most probable value (MPV) is calculated. This is done by fitting the curve with a landau-Gaussian fit. The MPV is compared when calculating the gain of a detector. β^- particles and other relativistic particles are often referred to as *Minimum Ionising Particles* (MIPs), as they have a mean energy loss rate in the detector material which is close to or at the minimum charge deposited per unit area. In silicon this is value is 73 electron–hole pairs per μm [64] which corresponds to 21900 electron–hole pairs in a standard 300 μm thick detector.

The same setup was used for these measurements as with the Alpha measurements in section 4.5.1. In this experiment no external trigger was used. It was assumed that the total number of low energy electrons were low compared with the high energy ones.

Beta Results Run 4

As mentioned a 2282 keV ^{90}Sr$_{52}$ source was used to calculate the gain of LGAD detectors, the same detectors used with the Alpha measurements. Data was taken until 100000 acquisitions were made. Figure 4.27 shows the fitted histogram produced from such a data acquisition, showing the MPV to be 518 at 150 V. This method was used to calculate the MPV for various detectors at a range of voltages.

The MPV was compared with that of a PIN diode to calculate the gain. The gain
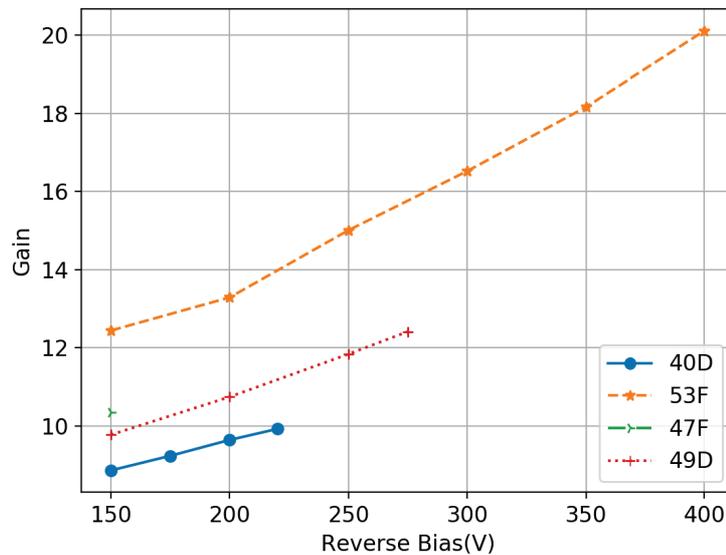
Figure 4.28: Gain against voltage for Beta measurements of LGAD detectors.

is plotted as a function of voltage in figure 4.28 . There is some difference in the gain obtained for each of the detectors. At 150 V the gain is in the range 9–13. Comparing device 49D with the results obtained from the TCT method at 250 V, we get a value for gain of $\approx$ 12 and $\approx$ 14 from Beta and TCT methods respectively. As we have seen from TCT measurements, figure 4.18, we have some gain variation from one measurements to the next with an estimated error of $\approx$ 10%, which allows these two measurements to be comparable. The Beta measurements don't incur the same error's but need a long time for each measurement to enable sufficient statistics for a reliable fit. Beta is the real use case so it is very positive to see that TCT agrees as TCT is easier to give position information.

4.5.3 X-ray

The final characterisation tool used in this work was performed using a variable X-ray source. The source was made by Amersham (code AMC.2084). The primary source is a 10 mCi ^{241}Am source. This source illuminates one of six metal foils, which fluoresce in the x-ray region from 8–44 keV, as shown in the table 4.6.

Components of X-ray Setup

The setup for these measurements was similar to the setup for alpha measurements in section 4.5.1. However for these measurements the vacuum chamber was not needed. The DUT was placed inside the TCT box connected to the same readout system. The variable X-ray source was placed at the backside of the detector with a thin metal layer between

Target	Energy (keV)	
	K alpha	K beta
Copper(Cu)	8.04	8.91
Rubidium(Rb)	13.37	14.97
Molybdenum(Mo)	17.44	19.63
Silver(Ag)	22.10	24.99
Barium(Ba)	32.06	36.55
Terbium(Tb)	44.23	50.65

Table 4.6: Variable X–ray source energies available.

the detector and source. This metal layer was to stop any alpha particles from the source impinging on the detector. A LabView script was run to collect data on the waveforms produced by X–rays and to plot peak position with number of counts. Using a python script a Gaussian fit was added to the plot to enable the peak position to be acquired. The resolution could then be calculated by the full width half maximum (FWHM) of this fit.

Results X–ray Run 4

The measurements for X–rays were taken at a range of voltages and for four energies ranging from 17.5 keV up to 44.5 keV. An example of the plot produced by the Labview script with a Gaussian fit is shown in figure 4.29 for an LGAD detector at 300 V with an X–ray of energy 44.5 keV. The detector used for these results if 40D, from run 4, which has a gain of $\approx$ 10 at 200 V. It was expected that a double peak would be seen due to the presence of a no gain signal as seen with the Alpha measurements. However due to the lower energy of the X–rays this was not the case as the no gain signal was below the threshold of the noise.

Figure 4.30 shows the peak position against X–ray energy for a range of voltages. As expected there is a linear trend as the energy is increased. The size of the peak also increases with increasing voltage, due to the increased gain. Figure 4.31 shows a similar plot but has peak position against voltage for a range of energies. From the errors on the plot it can be seen that the resolution tends to decrease with increasing voltage. This is plotted in figure 4.32, where the fwhm tends to increase as the peak position increases from data taken with the Tb (44.5 keV) source. This could be due to a less stable gain value at this voltage when approaching breakdown, as we have seen small changes in electric field can cause reasonable changes in the gain. The minimum detectable energy with this detector was 17.5 keV.
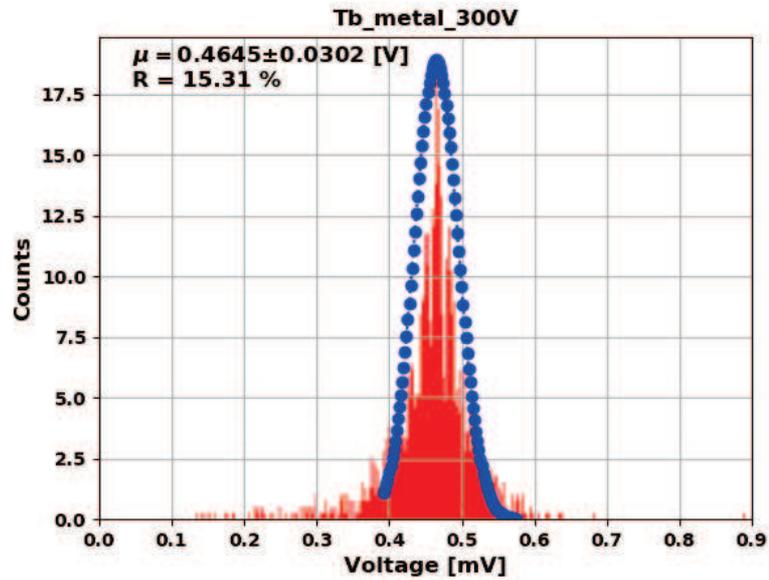
Figure 4.29: Number of counts against peak height for X–ray energy 44.5 keV. The blue curve shows the Gaussian fit of the data.
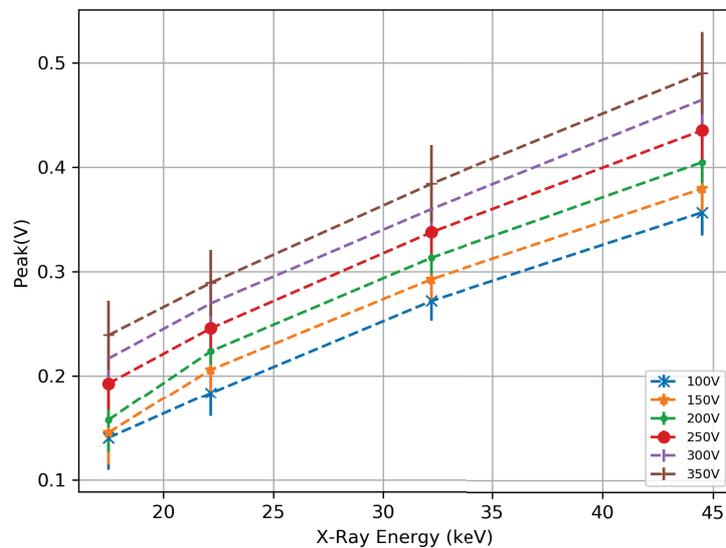
Figure 4.30: Pulse height against X–Ray energy for a range of voltages
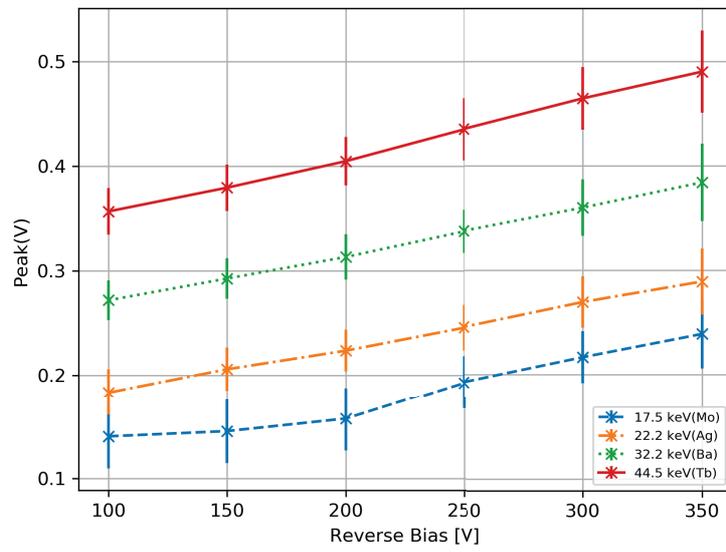
Figure 4.31: Pulse height against voltage for a range of energies
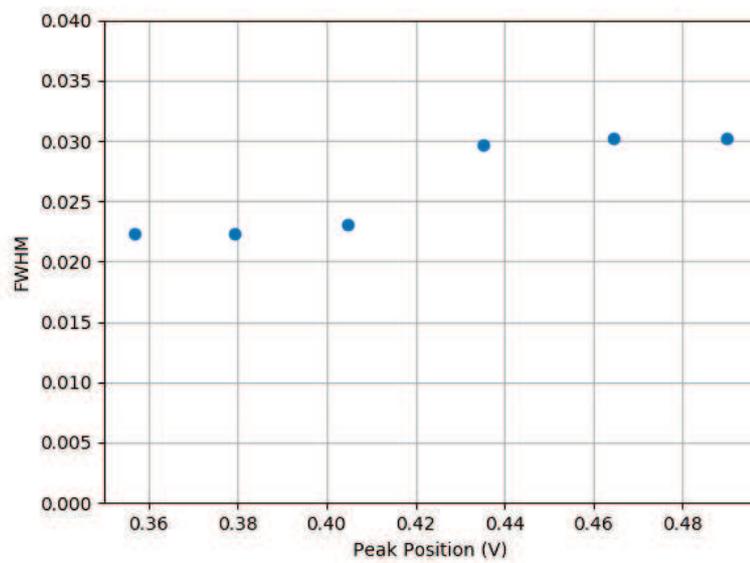
Figure 4.32: Pulse height against FWHM for an X-ray energy of 44.5 keV(Tb).

4.6 Conclusion

In this chapter the devices fabricated have been summarized, with details of the specific processing parameters used. In total there have been five fabrication runs of LGAD devices with a varying amount of success. The process parameters used for each run have been detailed where required, with some details held back due to confidentiality issues. The most important parameter for the multiplication implant has been detailed as a relative boron dose compared with the first dose used in run 1. Details of the different mask designs have been summarised to enable the comparison of design changes on device performance.

A summary of IV's and CV's measured for devices have been presented and where appropriate design differences discussed. The optimum peripheral region has been established through IV measurements for both breakdown voltage and leakage current. They were shown to improve with the addition of a Junction Termination Extension (JTE) as seen in the simulation, as well as with the addition of an extended field plate.

The gain was measured with a variety of techniques, TCT, Alpha TCT and Beta measurements. These have been compared where the gain for TCT and Beta measurements match within a 10% error, however the gain by Alpha TCT shows a significant reduction in gain. The fabrication process used to make devices in run 4 produced the desired gain of 5-20. This process was then used for the production of run 5, which also produced devices with the desired gain. Showing that the fabrication process can be reproducible, which is important for larger scale fabrication.

Finally X-ray measurements were made using a variable X-ray source to show the potential minimum detectable signal for these devices. The minimum energy detected was 17.5 keV for a 200 μm thick, 5 mm LGAD detector with a gain of 10 at 200 V. Due to the size of the detector the noise level was too high due to the high detector capacitance to measure any smaller signals. With a lower noise system it should be possible to detect < 5 keV signals, where the use of smaller pad sizes will reduce this further.

A potential issue found during these measurements was the reduction in the gain value for devices of pad size 0.22 μm when a larger JTE width was used. The gain was compared with a device with a JTE width of 10 and 20 μm, where the gain at 300 V was reduced by a factor of 25%. The issue is set to be more drastic when using pads of a smaller size and this is discussed in chapter 6 , through the use of simulations.

Chapter 5

Timing Detectors

This chapter will briefly discuss the requirements of fast timing detectors in high energy particle physics with relation to the upgrade of the Large Hadron Collider (LHC). The discussion will include the required time resolution and why Low Gain Avalanche Detectors (LGAD) may be a suitable candidate to provide this. Detectors fabricated at both Micron Semiconductor and CNM have been tested for timing capabilities using a dedicated set up developed by the author at the University of Glasgow. The results of these tests will be discussed along with the data acquisition and analysis strategy.

5.1   Timing Resolution Requirements

In many applications where the detection of charged particles is considered, the trajectories of such particles is required. In silicon tracking detectors, the trajectories can be reconstructed by analysis of charge deposition in multiple layers of detectors. Typically to measure this trajectory requires a detector with high 2D spatial resolution; such as that found in a pixel array. The segmentation concentration can vary, where the pixel pitch considered in this work varies from 55 μm up to 1.3 mm. It is also a given requirement that the total area covered by the detector is maximized, thus minimization of dead regions in detectors is preferred. Typically a particle detector will have a specific detection rate with which it can deal with the incoming particles. As technology has improved the rate capabilities of detectors has increased and has pushed the boundaries to more rapid measurements. This leads to significant increases in data collected which has improved the overall performance of the detector system. Silicon detectors are used in this field as they have been highly studied and the fabrication process is well understood through planar fabrication processes. The detectors have been developed to work in array structures of fine segmentation over large areas. Very high data rates are achievable with this type of detector where the readout chain is optimised. One limitation is the measurement of the arrival time of a particle which has traditionally been larger than 200 ps due to the charge

collection time in the sensor and the timing characteristics of the readout electronics. The trajectory taken by a particle at a typical high energy physics (HEP) collider experiment can be reconstructed by information obtained by a series of pixel detectors arranged at increasing distance from the origin of the track. Although there are many events occurring in a short time window, the density of events is relatively low as they occur in different locations. This means that the information obtained by traditional pixel detectors is sufficient to reconstruct every event.

However, as the LHC is to be upgraded in the coming years, such that the expected rate of events is to increase significantly by an approximate factor of 5 compared to the typical luminosities of LHC Run–2 [65]. This will mean that traditional sensors will not be fit for purpose as the track density will increase and miss–identification of pixels hits to the trajectory of the original particle will occur. The recent development of LGAD's shows the possibility of providing an improvement in the measurement of track arrival times. With the goal being the simultaneous high spatial resolution by fine segmentation and the capability of collecting data at a fast rate with good time resolution. Due to the nature of the time measurement, which requires a very short signal, the ability to deal with high data rates actually increases. With sufficient time resolution we can group particles that come from a well–defined primary vertex by their arrival time [66]. We will also be able to distinguish unique tracks which occur close in position but vary in arrival time [3]. The aim is to develop a detector and readout system that has a time resolution of around 10 ps.

5.1.1   Using Timing Information

In order to understand the need for timing information we can look at the information collected at the existing LHC. Typically charged particles are detected through a range of position sensitive measurements spaced over the particle's trajectory. As there are many particles to be detected, they are often detected within the same time frame. At the moment the density of events is low enough that 3D tracking information is sufficient to reconstruct each vertex as events do not occur in the same space.

However the increase in luminosity at the HL–LHC [65] will provide such an increase in the density of events that events will be overlapping in space and thus irresolvable. The present resolution for the separation of a vertex is 250–300 μm at the CMS and ATLAS detectors [67]. With the density of events expected, this will lead to 10–15% of vertexes composed of two events, leading to loss of events and precision of the reconstruction. Timing information is needed to improve this outlook, [66] shows the timing distribution of events expected form the HL–LHC has a distribution with an RMS of approximately 150 ps [66]. With the inclusion of a timing resolution of 30 ps one can divide the number of events into five distinct groups, each with a fifth of the total number of events, which allows separation in time the spatial overlapping events. Timing information is required to fulfill the potential

of increased interaction rate.

Interestingly it has been shown that it is not a requirement to have time information for each hit of the track, but to assign a time for each track. This allows the implementation of a single timing layer outwith the tracker volume which simplies the design.

5.2   Ultra-fast Silicon Detector

The need for good timing resolution has been briefly discussed in section 5.1. The implementation of this timing resolution requires a new detector technology. The choice for this detector is to build on existing detectors that are widely used and provide high spatial precision. As discussed in this thesis the LGAD detector has been investigated for its potential use as a timing detector as it uses planar technology, which is widely used to create pixellated detectors. The LGAD detector can be optimised for time measurements, thus the name ultra-fast silicon detector (UFSD) is given to this type of sensor.

The operation of an LGAD detector has been discussed previously in this thesis, where the control of the doping profiles is highly important to create the optimal electric field and hence the desired gain. The gain in LGAD's used in this work is initiated by drifting electrons toward the high field region. Using the knowledge that the multiplication mechanism for electrons occurs at a lower electric field value than for holes, the device can operate in a state outwith avalanche mode. Where only the electrons undergo multiplication, as the electric field is designed to be lower than that required for hole multiplication. This provides a reliable LGAD with a large bias voltage working range and a gain loosely dependent on the voltage. It's also important to note that this regime helps reduce the excess noise factor attributed to the multiplication process.

The internal multiplication process multiplies signal and noise, and in fact tends to increase the sensor noise more than the signal [66]. This has a negative effect on the signal to noise ratio (SNR) of the sensor. However if the sensor noise is kept low, by having low values of gain, the noise contribution of the detector will be dominated by the electronic noise and not the sensor noise. This produces a final SNR which is increased for devices with low gain.

5.2.1   Why Low Gain?

There has been many discussions on the need for low gain with respect to high gain devices. High gain silicon devices are available in the form of SiPM and APD, which are designed to detect single or a few photons respectively. A high gain is needed to ensure that charge from this single photon is detected above the noise. The high value of gain causes some negative features of the detector. This includes high sensor noise and the difficulty in segmenting such a device due to the high fields.
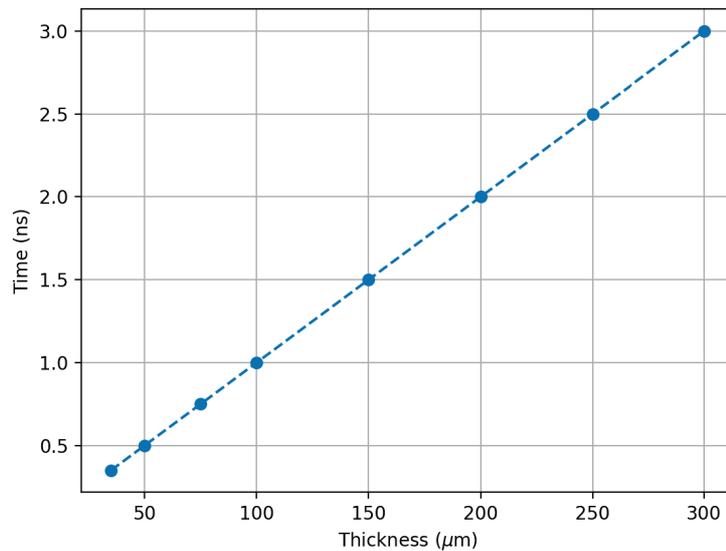
Figure 5.1: Electron drift time as a function of sensor thickness.

As we are discussing the timing resolution requirements for a particle collider we have an advantage of a larger initial signal generated by charged particles relative to single photons. This allows the use of a lower gain detector and hence the development of LGAD technology for fast timing applications. The idea of the UFSD is to develop thin LGAD sensors that have an optimised gain that is large enough to perform single particle time measurements.

The optimisation of the gain value is dependent on noise contributions. Excess leakage current is generated in high radiation environments which causes shot noise and heat. Therefore the sensors must be cooled sufficiently to reduce leakage current and control heat dissipated by the sensors.

5.2.2 Why Thin?

As mentioned in section 5.2.1 the UFSD will comprise of a thin LGAD sensor. To explain the reason for this we can look at the current signal generated by a MIP in an LGAD. The rise time of the signal is equal to the drift time of an electron travelling the full sensor thickness. The maximum current is dependent on the gain value. Thin sensors have a smaller rise time, compared to thicker detectors at a fixed gain value, which produces a better timing performance, discussed further in section 5.4.3. The expected time required for an electron to travel a range of distances is plotted in figure 5.1, assuming the electrons have achieved a saturation velocity of 1×10^7 cm s^{-1}.

A disadvantage of thin sensors when measuring MIPs is the reduction in signal size with reduction in sensor thickness. Therefore a high gain is required to generate large

enough signals to be measured by the dedicated read out electronics. They also have large values of capacitance which is detrimental to time resolution and noise. The sensor thickness has to be optimised along with the gain value to achieve the best time resolution.

5.3 Detectors optimised for time measurements

Fast time measurements rely on the speed of the readout electronics to measure a signal generated from the sensor material. They need to be designed to fulfill the potential of the sensor, more information on the design of such a readout is explained in [68] for the High Granularity Timing Detector (HGTD). However this thesis will concentrate on the optimisation and characterisation of the sensors.

Timing detectors have a limiting factor dependent on the signal shape. For fast timing it is required that the signal shape be constant and only scales with the amount of energy deposited. If the shape of the signal changes with varying particle positions then the time resolution will be negatively affected. Typically in a time–tagging detector the signal generated from the sensor is read out by a preamplifier, which shapes the signal. The output from the preamplifier is compared with a set threshold value V_{th} to find the time of arrival. The time of arrival, t_0, is the position in time when the signal crosses this set threshold value. Any changes in the shape of the signal close to this threshold value has an effect on the t_0 and thus effects the time resolution (σ_t).

One can group the effects that alter the time resolution into four categories [66]:

1) The total energy deposited by the particle that controls signal size ($\sigma_{\text{Time Walk}}$) and irregularities ($\sigma_{\text{Landau Noise}}$),

2) Changes in signal shape can be due to a non–uniform weighting field and variations in the drift velocities of the electrons ($\sigma_{\text{Distortion}}$),

3) The electronics used which cause noise and changes in the amplifier slew rate (σ_{Jitter}),

4) The effect of digitising the signal during readout, mostly limited by the time to digital converter (TDC) uncertainties (σ_{TDC}),

all of which will explained below.

The total time resolution can be expressed using as a sum of the terms mentioned and is shown as

$$\sigma_t^2 = \sigma_{\text{Time Walk}}^2 + \sigma_{\text{Landau Noise}}^2 + \sigma_{\text{Distortion}}^2 + \sigma_{\text{Jitter}}^2 + \sigma_{\text{TDC}}^2. \tag{5.1}$$

For this work the contribution in time resolution by the TDC binning is neglected as this does not contribute to the resolution of the sensors but of a readout chip which is not discussed further here.
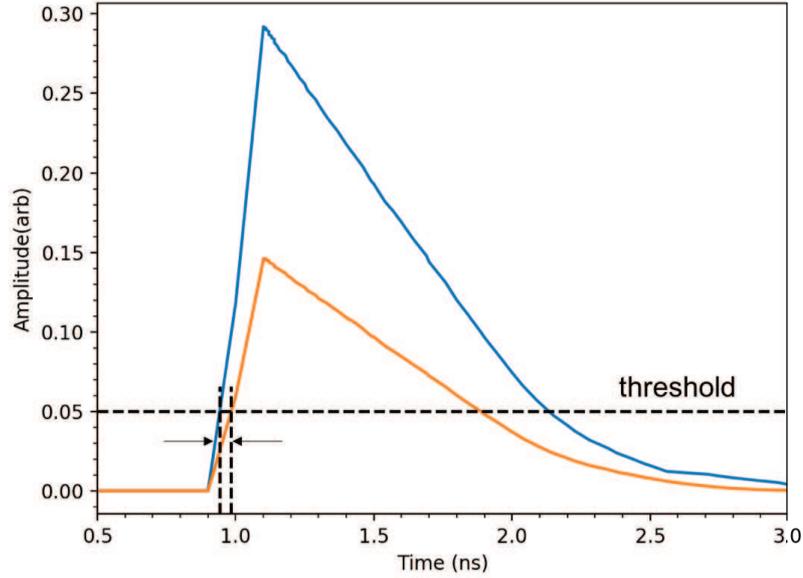
Figure 5.2: Simulated signals of different amplitudes cross the threshold value at different times. This causes a delay, t_d, on the firing of the discriminator. Larger signals fire first.

5.3.1   Time walk and Landau Fluctuations

Signal uniformity is limited due to the physical energy deposition by a traversing charged particle. The density of electron–holes pairs generated in the silicon along the particles path through the silicon changes from one event to the next. This is not only an issue for LGADs but for no–gain sensors also. Due to the variation in charge deposition for each event, this contributes to the change in the signal size which is the main cause of the time walk effect. This also produces irregularities in the signal, Landau noise.

Time Walk

Time Walk is the term used to describe the effect that larger signals cross a fixed threshold at an earlier point in time than smaller ones. In order to find a relationship to describe this, we assume a simple model of a linear signal, with amplitude S and rise time t_{rise}. The signal will cross the threshold value with a time delay of t_d, shown in figure 5.2. Geometrically a relationship is evaluated to be $\frac{t_d}{t_{rise}} = \frac{V_{th}}{S}$ where the delay time can be written as, $t_d = \frac{t_{rise} V_{th}}{S}$. Time walk is then expressed as equation 5.2 by the RMS of t_d

$$\sigma_{\text{Time Walk}} = [t_d]_{\text{RMS}} = \left[\frac{V_{th}}{S/t_{rise}}\right]_{\text{RMS}} \propto \left[\frac{N}{dV/dt}\right] \tag{5.2}$$

where the $\frac{dV}{dt}$ term is the slew rate, given by the ratio of the signal size to the rise time, $\frac{S}{t_{rise}}$. The threshold value is often dependent on the noise and is given as a multiple of the

system noise, N. Therefore Time Walk is minimised for detectors with a high slew rate and low noise.

5.3.2   Non–uniform weighting field and variations in the drift velocities

All particle detectors can have the shape of their current signal calculated by Ramo's theorem [24], which states that the current induced by a charge carrier is proportional to its drift velocity v_{drift}, the electron charge q and the weighting field E_w:

$$i(t) = -q\vec{v_{drift}} \cdot \vec{E_w} \tag{5.3}$$

Where equation 5.3 shows two main features by which the sensors should be designed around for accurate timing. The current signal depends on the drift velocity of the charged particles, thus good timing resolution depends on the ability to keep a constant drift velocity through the entire detector volume to minimise signal distortions. The easiest way to achieve this is to run the detector in a regime where the velocity of the charged particles is saturated. This occurs at a specific electric field value, which depends on the operating conditions. The electric field required for saturation velocity has a temperature dependence, where a lower temperature environment lowers the field necessary for saturation as well as achieving a higher saturation velocity. At room temperature the velocity saturates for an electric field of ≈ 30 kV cm^{-1} [69].

The weighting field E_w, should not vary as a function of the track position. This is an important issue when segmenting detectors into pixellated or strip detectors. As the implant width decreases the weighting field can alter around the pixel/strip implant edges and this contributes to a worsening of the time resolution. In general for a pixellated detector the pitch should be similar to the implant width and much greater than the thickness of the detector. This ensures the weighting field is concentrated underneath the pixel. For this work the weighting field contribution to the timing resolution can be ignored as only pad detectors have been tested, where the size of the pad is much greater than the thickness. This will become a problem for smaller pixellated devices. As this will be discussed in chapter 6 Nevertheless this information shows that the final detector should be as close as possible in geometry as a parallel plate capacitor with uniform electric and weighting fields.

5.3.3   Jitter

The time resolution depends on the jitter. The jitter is a term given for the time uncertainty around the firing of the comparator due to the presence of noise. This can be noise from the sensor or the electronics [66]. The jitter is directly proportional to the noise of the

system, N, and inversely proportional to the slew rate of the signal. If we assume that the slew rate is constant over the entire signal we can describe the jitter time resolution in terms of the slew rate and noise. This can be written

$$\sigma_{\text{Jitter}} = \frac{N}{\mathrm{d}V/\mathrm{d}t} \approx t_{\text{rise}} \bigg/ \left(\frac{S}{N} \right). \tag{5.4}$$

As previously shown, the slew rate is a function of the signal size and rise time. This equation is used to optimize the design of the electronics which requires low noise and a high slew rate. However large slew rates require a wide bandwidth, which increases the noise. Lowering the noise would require smaller slew rates. These competing elements must be balanced in the final design of such a detector. For the readout system used for the time measurements made in this work, the electronics used will be described where the trade-offs between noise and slew rate are discussed.

5.4   Measuring the Performance of LGAD's

Throughout this thesis the performance of the LGAD's has been discussed for various fabrication runs, where all detectors have been fabricated by Micron Semiconductor Ltd. Measurements of time resolution and gain of various detectors have been published in [36], [37] and [70] for devices fabricated at other facilities including CNM, FBK and HPK. These have been produced for the purpose of implementing this technology for the upgrade to the HL-LHC using the HGTD readout system [68].

This section will only concentrate on two detectors fabricated at Micron Semiconductor and tested at the University of Glasgow.

5.4.1   IV and Gain

Before undergoing time resolution measurements the devices were tested for IV and gain measurements. The set-up for these measurements has been previously described in chapter 4. Therefore only the results of these tests will be shown in this section. The devices tested were from a wafer in run 5. This wafer had a relative boron implant dose of 1.1 and the devices have an active area of 1x1 mm^2. The IV measurement for both devices is shown in figure 5.3a where the breakdown voltage is $\approx$ 200 V. The thickness of these detectors is 250 μm. The current is observed to level off at 200 V for both detectors for around 50 V before breaking down. During measurements of time resolution it was found that in this plateau region the devices were self triggering.

Measurements of gain were then performed using the TCT method described in chapter 4 for a range of voltages. The gain is plotted as a function of voltage in figure 5.3b for both devices. The gain for both devices at 200 V is $\approx$ 5. As shown in figure 5.3a, the
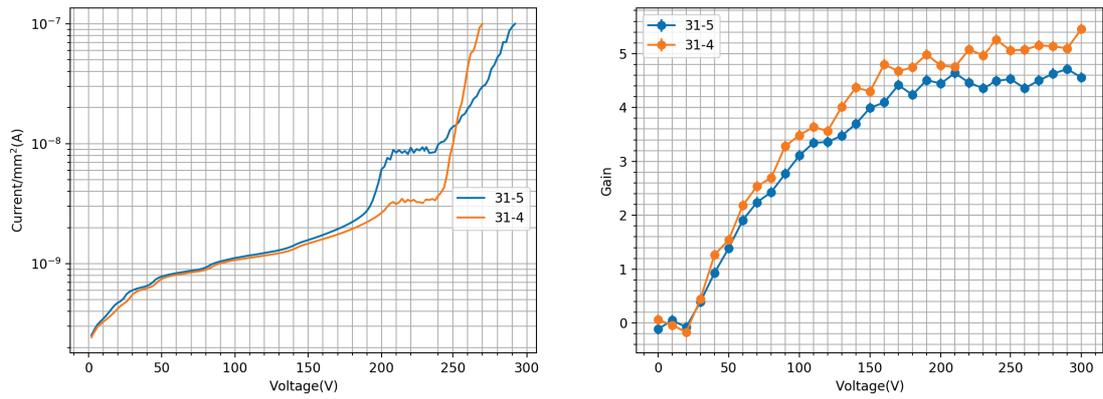
(a) IV measurements of 250 μm thick LGAD detectors with an active area of 1x1 mm^2 from a wafer in run 5.

(b) Gain measurements of 250 μm thick LGAD detectors with an active area of 1x1 mm^2 from a wafer in run 5.

Figure 5.3: Device Characterisation of the two detectors before time resolution is established

leakage current and the gain exhibit the same bias voltage dependence, as expected from the common charge multiplication mechanism. Above a bias of 200 V, the current shows a faster increase with the bias voltage than the gain, which might be due to the start of breakdown in the periphery of the pad, instead of the central bulk where the signal is generated.

5.4.2 Experimental Set-up

The time resolution measurements were conducted in a climate chamber located in a semi-clean room environment in order to accurately control the temperature and humidity during measurements. The experimental set-up will be described in the following sections with focus on the key elements required for good time resolution.

Mechanics

The general set-up consists of an Aluminium L-shaped support frame that supports the sensor test PCBs, as well as a radioactive source that is positioned on a mount fixed with screws to an aluminium breadboard optical plane. The layout of the structure is shown in figure 5.4, highlighting the necessary components. Due to the small size of the active area of the detectors precise alignment is required. As such it was found to be easiest to fix two PCB's to one support frame. The frames were designed in such a way so that the two detectors would perfectly align. The frames were designed at the University of Glasgow
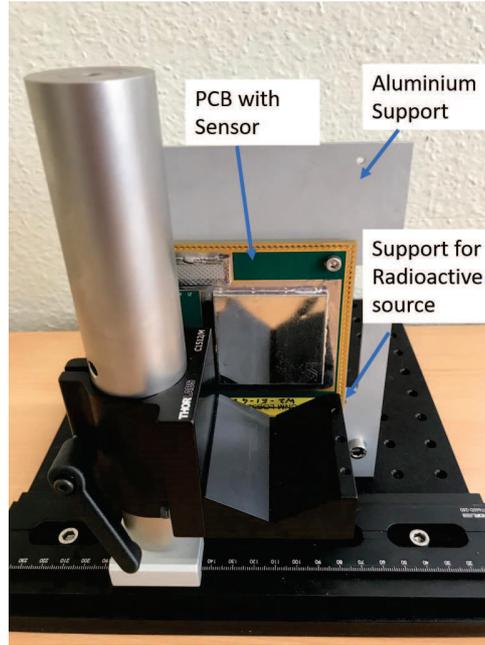
Figure 5.4: Mechanical set–up for timing measurements, showing position of sensors with respect to radioactive source.

using Solid–Works[1] and machined in house.

The aluminium breadboard base plate was purchased from Thorlabs to have an area of 250 mm x 300 mm. The thickness was chosen to be 12.7 mm to ensure stability of the set–up. Holes in the base plate were M6 with a spacing of 25 mm. The largest possible board was purchased to fit inside the climate chamber.

The support frames fixed to the base plate through specially designed holes. The holes on the frame were designed at a distance that matched the grid pattern of the base plate. This allowed movement of the frames across the board if needed. Thorlabs angle brackets were used to fix the frame in the desired location on the board using M6 screws.

Radioactive Source

A ^{90}Sr source is used for sensor laboratory testing. ^{90}Sr is typically a byproduct of nuclear fission and has a lifetime of 28.8 years. It undergoes β– decay into ^{90}Y with a decay energy of 0.546 MeV, which in turn undergoes β– decay with half–life of 64 hours and decay energy 2.28 MeV [71].

To support the radioactive source in the climate chamber components were purchased from Thorlabs to fit into the breadboard. A Dovetail rail was mounted on the board which allows the source to be moved along one axis. The source was mounted on a Post V–Clamp Mount which can rotate the source a full 360°about the x–axis. To fix the V–Clamp Mount

---

[1]Solid–Works is a solid modeling computer–aided design and computer–aided engineering computer program
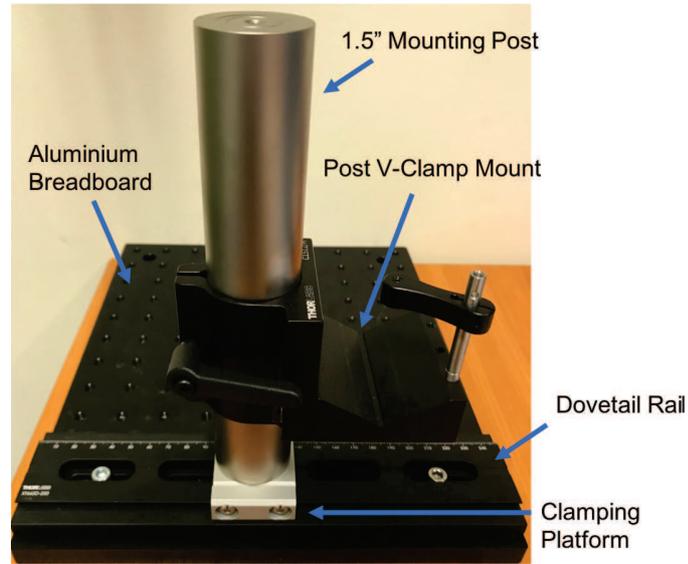
Figure 5.5: Mount used for radioactive source.

to the Dovetail rail a post is needed which, is in turn connected to a clamping platform. The Post V–Clamp mount allows the height of the source to be adjusted where the minimum and maximum height is determined by the length of the post used. The resultant mount is shown in figure 5.5, where it is connected to the aluminium breadboard.

A cylindrical source shield collimator was designed in Solid-Works to fit around the source. The collimator was manufactured using a 3D printer with a length of 100 mm and a collimator hole opening of 3 mm. The complete mounted source can be position wherever necessary, to be aligned with the detectors under test.

Readout Board

The readout board is based on a single transistor common emitter design and acts as an inverting trans–impedance amplifier [72]. An AC coupled silicon–germanium bipolar transistor is used as the first amplification stage. It has a transition frequency of 75 GHz. A gain of 29 dB is expected at a frequency of 1.9 GHz, where an integrated output noise of 260 μV is estimated. The board has a feedback loop which is designed for timing applications with small capacitance sensors, C3, inducing rise times which have a typical value around 800 ps, with a feedback resistor of 470 Ω, R3. The circuit diagram for the readout board is presented in figure 5.6.

Sensors are mounted to the boards by double sided conductive tape while the amplifier input is coupled to the front side metallisation layer via multiple wire bonds, to minimise inductance. A 1 MΩ resistor is attached between the input and ground which serves for detector biasing, followed by a pair of low forward resistance silicon pin diodes, D1/D2. The pair of diodes have a breakdown voltage of 50 V at 5 μA which protects the amplifier
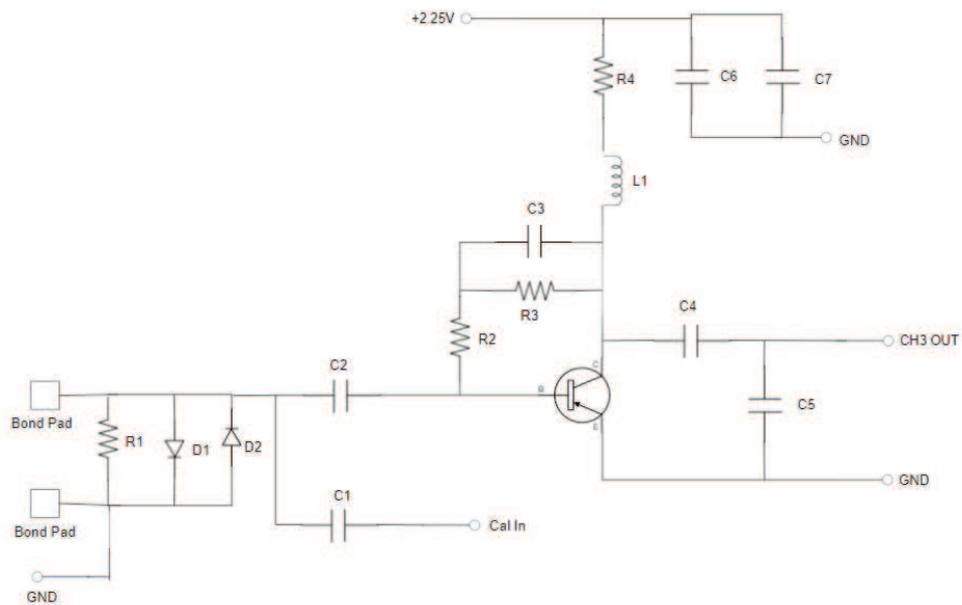
Figure 5.6: Circuit diagram of readout board [73].

Table 5.1: Electrical Characteristics of Board.

Supply Voltage	2.25 V
Expected supply current	15 mA
Amplifier gain	10
Expected noise (rms)	1.2 – 2.0 mV

input.

The overall trans–impedance within a 1.6 GHz bandwidth and terminated into 50 Ω, R4, can be estimated to be in the region of 470 with respect to a feedback resistor of 470 Ω. This was calculated assuming an external voltage amplifier of gain 10 is used.

The board has been designed to provide complete hermetic shielding on both sides up to a bandwidth of 3 GHz, where RC filtering is implemented on both the HV and low voltage input lines. Many high voltage decoupling capacitors, C6/C7, have been used to minimise parasitics and reduce inductance on the signal return path. A supply voltage of 2.25 V is required for the low voltage amplifier with an estimated working current of 15–17 mA. Powering is provided by a Thurlby Thandar Instruments (TTi, 32 V–3 A) power supply where grounding is assured via the signal readout cable at the DAQ level. The electrical characteristics of the on board amplifier are shown in table 5.1. A bias voltage is applied to both detectors by Keithley 2410's. One power supply for each detector.

Aluminium lids on both sides of the board had to be drilled to create a hole in order to allow β particles to pass through both detectors. The hole was then covered with thin aluminium foil , less than 50 μm thick, which was glued down by double sided conductive
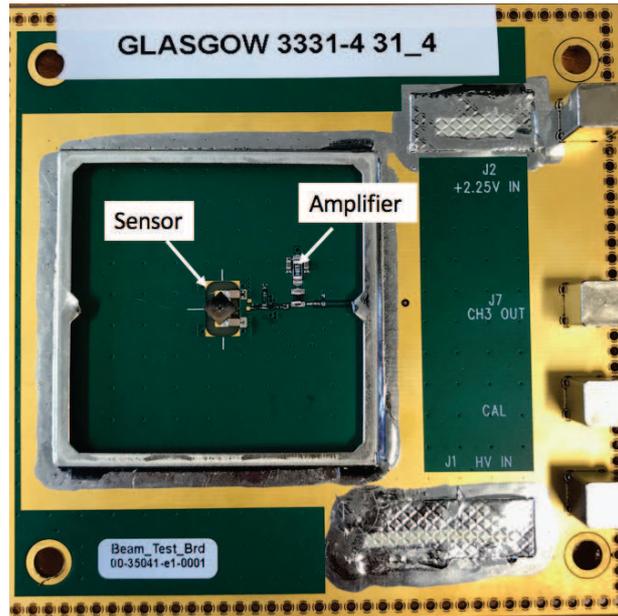
Figure 5.7: Test Board with mounted detector.

Table 5.2: Electrical Components of Evaluation Board.

Component	Value
A1	Gali-52(+)
C1	2400 pF
C2	1.2 - 2400 pF
C3	0.1 μF
R1	143 Ω, 0.75 W
R2	8.25 Ω, 0.25 W
CHK	Mini-Circuits TCCH-80+

copper tape in order to secure the hermetic shielding for noise reduction.

The Readout board with a mounted sensor is shown in figure 5.7 without a shield. The detector is biased from the backside where a wire bond is connected to the front side for reading out the signal. The guard ring is connected via a wire bond to ground.

Second Stage Amplifier

The second stage amplifier was purchased from Mini-Circuits. The amplifier used has part number TB-409-52+ (Gali-52(+)), which is based on a discrete integrated Gallium Phosphate broadband amplifier [74]. It has a dynamic range of up to 2 GHz with an input impedance of 50 Ω and an output power of 15.5 dBm. The amplifier is mounted on an evaluation board including a low voltage filter and the A/C coupling capacitors in a single metal frame. The layout is shown in figure 5.8, with the labelled components described by table 5.2. The electrical circuit for the evaluation board is shown in figure 5.9.
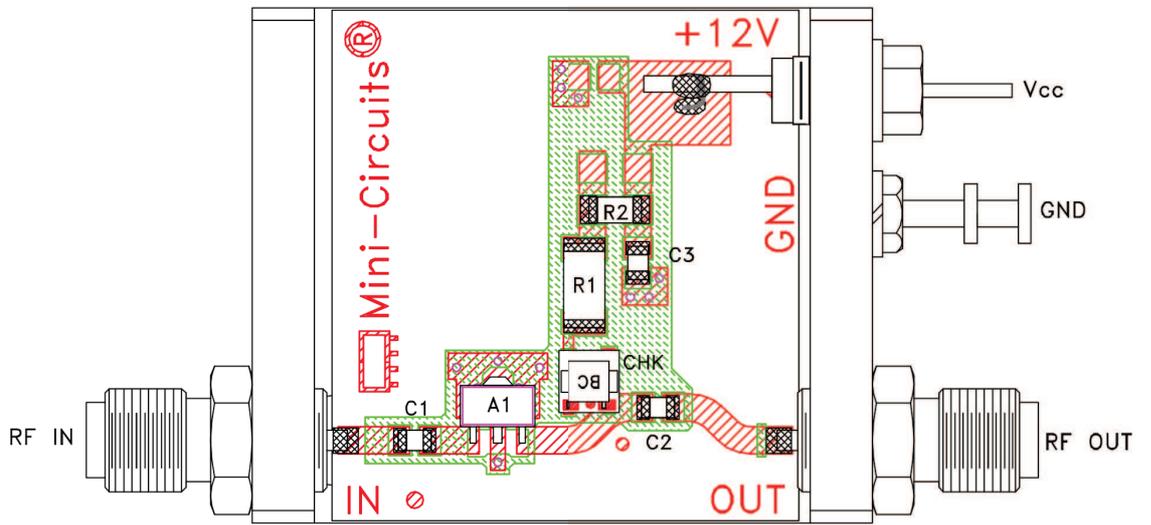
Figure 5.8: Evaluation board containing 2nd stage amplifier. [75]
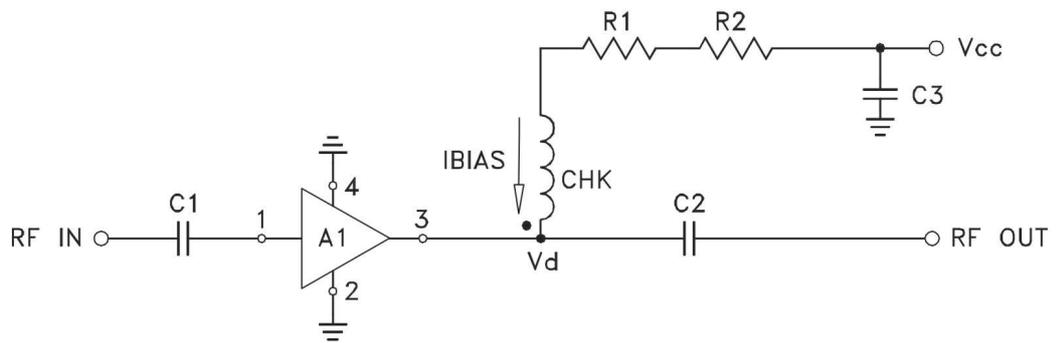
Figure 5.9: Electrical circuit for the 2nd stage amplifier. [75]

(a) Amplifier 1   (b) Amplifier 2
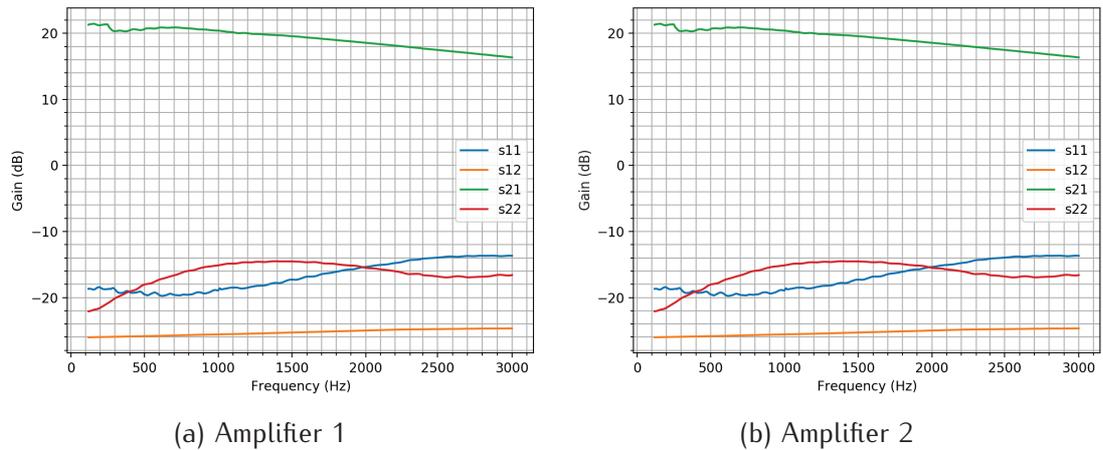
Figure 5.10: Calibration curves for both amplifiers used. S11 and S22 are the reflected signals induced from either side of the DUT. S21 corresponds to the transmitted signal at the output from a signal produced at the input. In this case it corresponds to the gain. S12 is the transmitted signal seen at the input from a signal produced at the output.

To prepare the board a set of banana connectors are soldered to the relevant terminals for low voltage powering. The supply voltage required is 12 V with a supply current of $\approx$ 60 mA. An aluminium cover was machined to cover the board and apply sufficient shielding for the amplifier. A layer of Kapton tape is used to cover the top side of the board to ensure no shorts.

The amplifier was chosen as it shows little gain dependence on temperature which is necessary for the tests performed. The evaluation boards were tested before any time resolution measurements were performed to ensure similar performance of operation over a range of frequencies. The measurements were taken using a Keysight PNA–L Network Analyzer from 0–3 Ghz. The results for both are shown in figures 5.10a and 5.10b showing the S curves [76]. The gain of the amplifier is given by the curve S21. Which for both is plotted in figure 5.11, and it can be seen that the are almost identical in performance, which is necessary for timing measurements. This meant that both these amplifiers could be used for two in series detectors.

Oscilloscope

Data was acquired using a 4 GHz – 8 bit vertical resolution MSO9404A Mixed Signal Oscilloscope by Keysight. It has a sampling rate of 20 GS/s which provides a time discretisation of 50 ps. A Labview script runs the DAQ for a set amount of triggered waveforms. The noise, defined as the rms of the amplifier output before the signal, was calculated for each waveform for the first 900 points as this was before any signal contribution. This was measured to be in the range 0.8 mV – 6.59 mV. The S/N ratio was calculated for each wave–
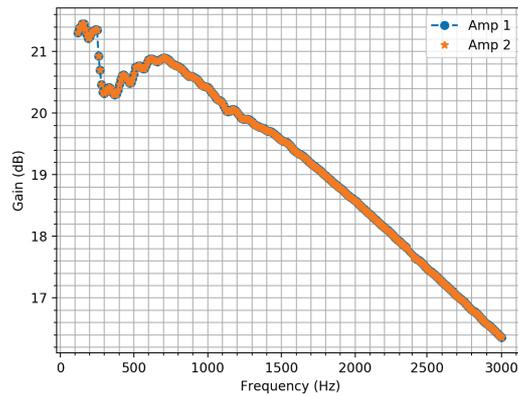
Figure 5.11: Gain as a function of frequency for two amplifiers.
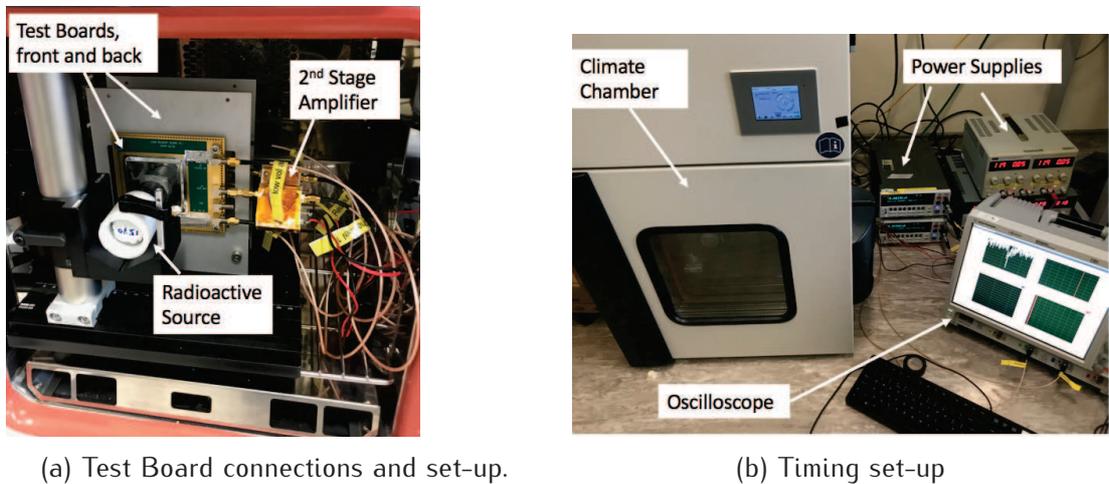
(a) Test Board connections and set–up.

(b) Timing set–up

Figure 5.12: Set–up for time resolution measurements.

form with a mean value used to calculate the signal time jitter. This can be approximated by $t_{rise} / \left(\frac{S}{N} \right)$. The average rise time of these detectors is ≈ 2.76 ns given by the time between 10% of the signal and 90% of the signal amplitude. Using the average $\left(\frac{S}{N} \right) \approx 22$ this gives a value for the time jitter of 124 ps. This gives a limit on the achievable time resolution with these detectors. The jitter was calculated for a device at –10° C biased to 180 V.

Figure 5.12 shows the complete set–up for the time resolution measurements, including the connections inside the climate chamber.

5.4.3   Time Resolution Results

Analysis procedure

The digital oscilloscope records the full waveform of the signal produced by the two de–tectors on separate channels for each event so complete information is saved for offline

analysis. The recorded data was analysed using the constant–fraction–discriminator (CFD) method. This was chosen as it is a procedure which can be used on chip and comparable to other results published. The oscilloscope can be programmed to trigger on a specific channel. It can also be programmed to trigger on two channels simultaneously. Events were selected by signals triggered on both detectors within a small time window to ensure coincidence. This allowed us to be confident that the signal on both channels was a result of the interaction on both detectors by a single β– particle. The trigger level was set to 10 mV, $\approx$ 5 mV above the noise for both channels. It is also possible to see non–gain events, signals from β– particle interactions in the perimeter of the device under the JTE, during these measurements. This was stopped by ensuring the pulse maximum fell within a window of 1 ns by the trigger settings.

Gain and Landau distribution

A Landau distribution of initial energy deposition is seen for traversing MIP's. The initial charge is then amplified by the internal gain of the detector and as a result the MPV varies with gain. The amplifier board used in this work contributes to the initial impulse response. This turns out to be a Gaussian curve. Therefore the final amplitude distributions has both Landau and Gaussian contributions. Figure 5.13 shows these distributions for a range of voltages (150 V, 160 V, 180 V and 190 V) with a Landau–Gauss convolution fit. As expected from the gain plots in section 5.4.1, the MPV does not increase in this bias range as the gain is stable at $\approx$ 5 in this region.

Timing resolution

The time of arrival of a β– particle in the sensor is determined using the constant fraction algorithm (CFD). This method involves measuring the time of arrival at a given threshold on the rising edge of the signal that is a certain fraction of the signal amplitude. The signal amplitude has been determined as the smallest value in the raw data set, as the signals are negative as shown in figure 5.14. The time of arrival at the chosen fraction value is determined as the first value in the data set below the requested value. This was performed for every waveform collected at a range of threshold values. For these measurements the CFD was performed with a constant fraction in the range of 10% – 90%. The time difference is defined as the time between the signal of one detector and the same point of the signal on the other detector. The time difference is then histogrammed, where the sigma of a Gaussian fit is declared as the time resolution. As this measurement was made using two identical detectors the time resolution of one detector can be calculated to be sigma divided by $\sqrt{2}$. This has been obtained for a range of bias voltages and temperatures. The voltage range is 140–190 V and the temperatures used are –10° C, –20° C and –30° C.

(a) 150V

(b) 160V

(c) 180V
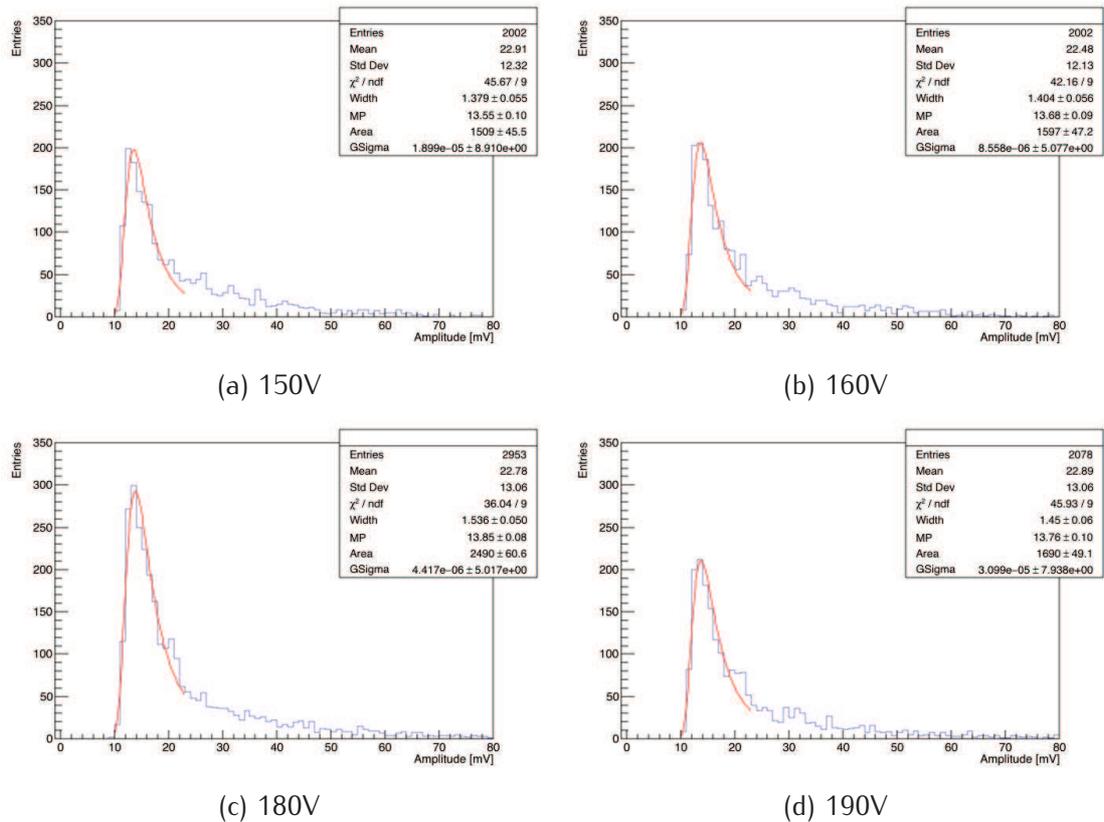
(d) 190V

Figure 5.13: Distribution of the signal amplitude for four bias voltages (150 V, 160 V, 180 V and 190 V), with a Landau-Gauss convolution fit. This fit is used to determine the most probable value. These plots are taken from a 1x1 mm^2 device from a wafer from run 5 with a relative boron dose of 1.1.
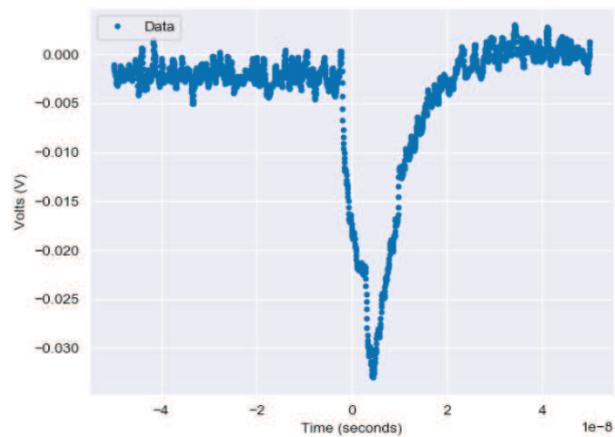
Figure 5.14: The full waveform of a typical signal observed.
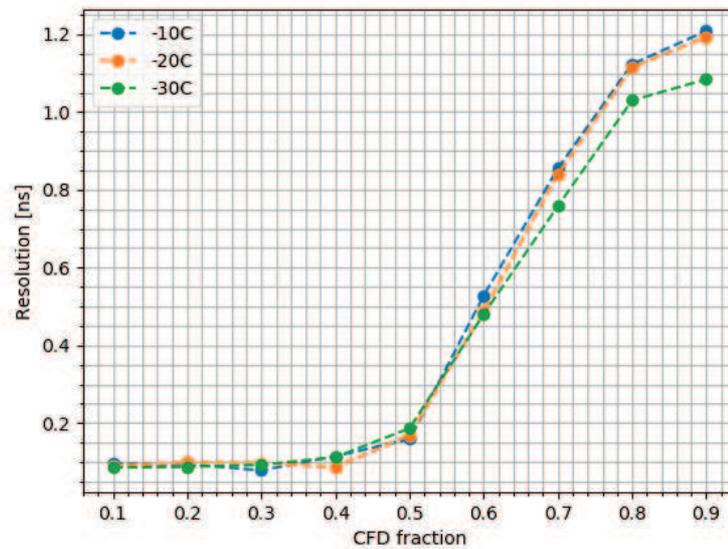
Figure 5.15: Time resolution as a function of CFD fraction at –10°C, –20°C and –30°C at 180V.

The time resolution was plotted as a function of CFD fraction and an example for one voltage measurement is shown in figure 5.15 for the range of temperatures. The lowest time resolution values are shown between 10% – 40% CFD fraction. Therefore a 40% CFD fraction was used for all further measurements. From this figure there appears to be a time resolution dependence on temperature. The time resolution was then plotted at a CFD of 40% for a range of voltages and temperatures. The results are shown in figures 5.16 and 5.17 showing the time resolution as a function of temperature and voltage respectively.

If we look at figure 5.16 it's clear that there is not a strong time dependence on temperature. The minimum time resolution recorded is $\approx$ 85 ps with the highest being $\approx$ 270 ps. The majority of measurements are in the range 85-130 ps. Figure 5.17 shows the same results but plotted as the time resolution as a function of voltage. Again there is no time resolution dependence on voltage. There are two reasons for this. The first being the achievable time resolution for this set-up and these devices are limited by the time jitter, which as previously mentioned is 124 ps. Also the detectors do not show any reasonable voltage dependence on the gain. Therefore, any possible advantages of increasing the time resolution by increasing the voltage are lost. Possibly at a higher voltage the gain could have be increased, providing the opportunity to improve the time resolution. However at higher voltages it was seen that both detectors showed signs of self triggering. This is the situation where a signal is generated by the detector outwith external sources. This shows the early signs of breakdown at the periphery of the detector, as no increase in leakage current is seen during this effect. Leakage current would increase rapidly if the device began to breakdown across the junction as the higher field would induce higher gain.
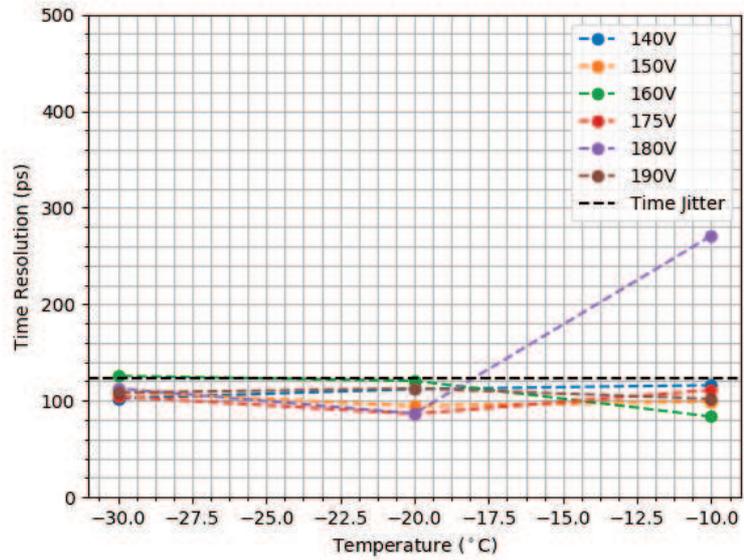
Figure 5.16: Time resolution as a function of temperature at a range of voltages.
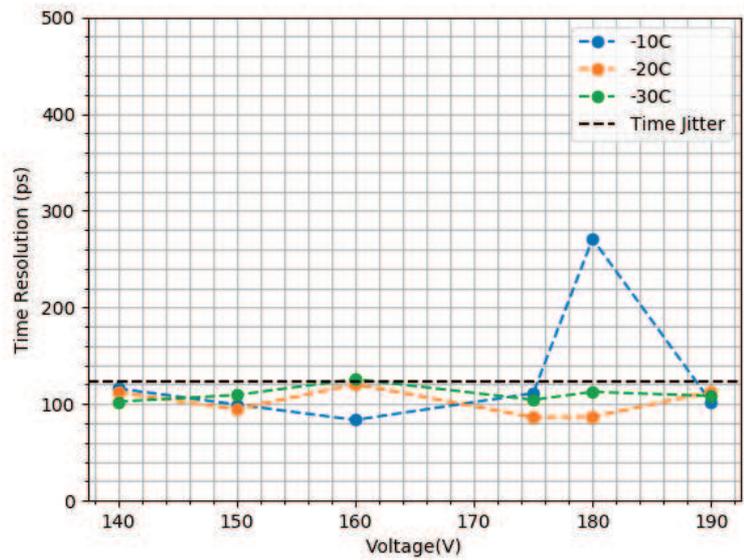
Figure 5.17: Time resolution as a function of voltage at –10° C, –20° C and –30° C.
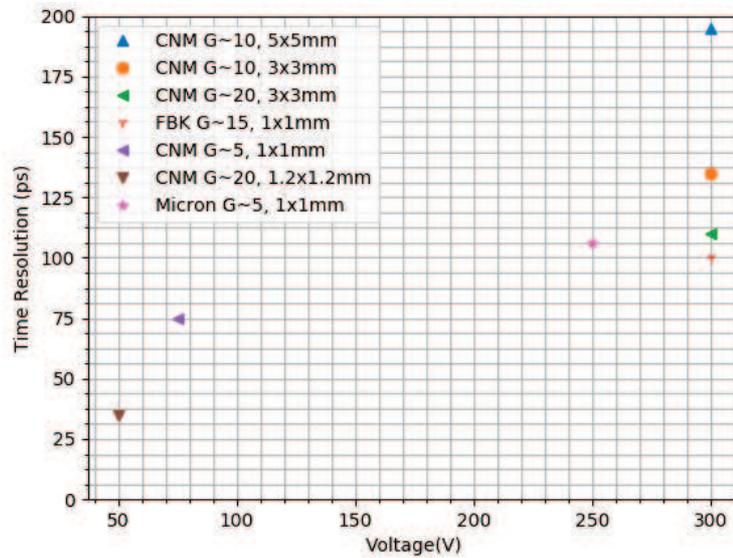
Figure 5.18: Comparison of time resolutions by various manufacturers for a range of LGAD detectors.

In order to compare with previously tested detectors from other manufacturers a summary of the timing studies performed is shown in figure 5.18. The results show timing resolution for a variety of pad sizes and thickness's. The results for the 300 μm thick sensors show the influence of gain and capacitance on the time resolution. Where smaller capacitance and higher gains yield to an improvement in performance. These results were taken from [66] where the data was taken at test beams at CERN. They indicate that using thinner detectors improves time performance as expected. The Micron sensor has been added to this figure and sits in a reasonable position in terms of its gain, thickness and size. Unfortunately there is a not a direct comparison but it can be inferred from the plot that the Micron sensor performs as well as the others respectively. This is promising and thus efforts have been made to produce thin, 50 μm thick Micron LGAD's. However these have not been fabricated in time to add any results to this thesis.

5.5   Summary and Outlook

For the 4D tracking of particles at the HL–LHC development of dedicated sensors has begun. This chapter has shown the work on the development of LGAD's at Micron Semiconductor as a possible candidate for the UFSD to be used in the HGTD. The most important effects on timing limitations have been discussed. We need detectors with a high field for saturated velocity and to minimise signal distortion. The size of the pixels should resemble a parallel plate capacitor which provides uniform electric and weighting fields. The collection of

electrons is preferred to holes, providing a faster rise time. Using thin sensors reduces the rise time, minimises the jitter and decreases the chance of charge trapping.

Two 250 μm thick LGADs from Micron Semiconductor have been evaluated for their timing performance. These devices have a gain of $\approx$ 5 at voltages up to 250 V with a leakage current less than 10 nA/mm^2. Measurements for timing resolution at a range of temperatures and voltages yielded a minimum time resolution of 85 ps. With comparison to other fabricated sensors from a variety of manufacturers this result seems reasonable. The time resolution did not have a voltage dependence as the gain was stable over the voltage range used. Also the jitter was calculated to be 124 ps which sets a limit on the achievable time resolution.

Next steps are to test recently fabricated 50 μm sensors from Micron Semiconductor with the hope of achieving a time resolution of less than 50 ps. Sensors with a variety of gain will be tested at this thickness to evaluate the gain dependence on the time resolution. Hopefully to be able to match the time resolution achieved by the CNM detectors. The sensors have also been fabricated in array structures where it is expected that the time resolution can be measured as a function of array size as this will be what is used in the final application.

Chapter 6

Small Pixels

This chapter will introduce the concept of producing small pixel LGAD's for use with the TimePix3 readout ASICs [59]. These readout ASICS typically have a pixel pitch of 55x55 μm in arrays of 256x256 pixels, producing an ASIC with greater than 65 k readout channels. Thus far these readout ASICs have been bump bonded to a variety of sensor substrate materials including Silicon and CZT with a range of thickness'. These sensors have been fabricated with a variety of pixel pitch sizes of which are 55 μm, 110 μm and 220 μm. For sensors with a pixel pitch of 110 μm every 4th channel on the readout ASIC is bump bonded to a single sensor pixel. For the 220 μm pixel design, every 16th channel on the readout ASIC is bump bonded to a single sensor pixel.

Issues with reducing the pixel size of LGAD's will be discussed with varying periphery designs to try and find the smallest pixel size reasonable for the technology available through TCAD simulation. Issues include the non-uniformity of gain across the pixel as well as having small fill factors. Designs will be discussed to push the limits of the existing technology to increase the fill factor and secure a better gain uniformity.

6.1   Hybrid Pixel Detectors

The development of hybrid pixel detectors with on–pixel electronics has been on going for several decades for experiments at the LHC. They have replaced strip detectors in the inner tracking layers of the vertex detectors as the track density was too high for strip sensors to correctly identify each track. Hybrid pixel sensors consist of a sensitive layer (Detector Material) bump bonded to a readout ASIC. Generally both the sensitive layer and readout layer are segmented in order to provide position resolution. An example of this formation is shown in figure 6.1. On the left side of the figure the entire structure is shown. The sensor is bonded to the readout with the front side pixellated. The back side of the sensors faces up and allows for a bias voltage to be applied. The front side is held to ground. Wire bonding pads are seen which connect the readout ASIC to the rest of the readout
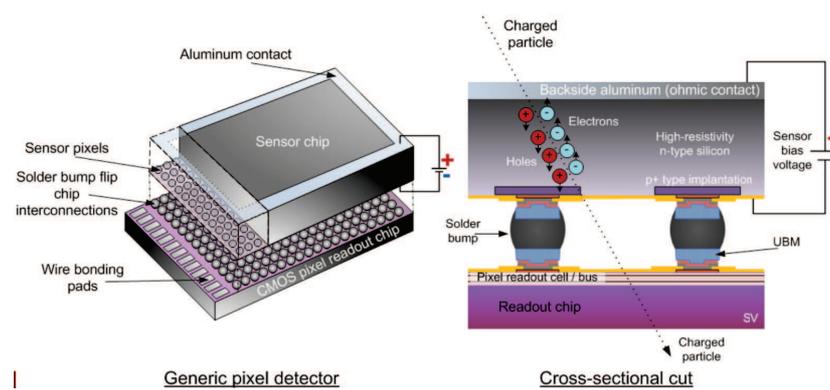
Figure 6.1: Example of a Hybrid Pixel Detector, with sensor bump bonded to Readout Asic. [77]

electronics. The right side of the figure shows how a charge is induced by an incoming particle and how the holes are collected at the segmented electrode. The holes generate a charge pulse which is collected by the readout ASIC via the solder bump. The solder bumps provide both electrical connection between the two layers and mechanical stability.

An advantage of this structure is the small input capacitance of each channel due to the small distance between the sensor and readout electronics. Typical values for this are 10–100 fF [78], with the input referred noise level of around 100 e$^-$ rms. The detector can be operated with a threshold of around 1000–2000 e$^-$ which essentially produces *noise free* hits. The Medipix Collaboration have produced detectors of varying complexity and will be briefly discussed in the following sections.

6.1.1   Medipix

The Medipix is a family of readout–ASICs for particle imaging and detection developed by the Medipix Collaboration. Thus far they have produced three ASICs, Medipix [78], Medipix2 [79] and Medipix3 [80]. These ASICs are photon counting ASICs where single hits for each pixel are processed on–pixel or within a neighborhood of pixels depending on the ASIC design. The number of counts per pixel is accumulated per pixel prior to a frame based readout, this requires the shutter being open while counting and closed at the end of each frame. This shutter time is predefined by the user.

The original Medipix detector was based on the Omega3/LHC1 [81] ASIC and submitted for fabrication in 1997. The ASIC uses CMOS technology with a minimum feature size of 1 μm. It is designed with a 64x64 readout array of channels with a pixel pitch of 170 μm. This ASIC had the ability to set a threshold for each individual pixel using a programmable 3–bit register. A global threshold is set using an external current source. The minimum usable threshold is $\approx$ 2000 e$^-$, which corresponds to a minimum photon energy of $\approx$ 7.2 keV This was the first ASIC used for large scale photon counting measurements and was used

in many medical imaging applications over a range of X-ray energies. With comparison to the technology at the time these ASICs provided a better signal to noise which offered an overall improvement in the quality of images produced for diagnosis. The ASIC was limited in size to one ASIC, as these could not be tiled together to produce a large array of detectors due to significant dead regions around the edge of the ASICs.

In 1999 the Medipix2 collaboration was formed with the aim of producing a larger area detector with finer segmentation. This was produced using 250 nm CMOS technology with a pixel pitch of 55 μm. An array of 256x256 pixels produces a detector with greater than 64000 readout channels [79]. Each cell of the readout ASIC has an 8–bit configuration register which allows for individual threshold adjustments as well as for masking noisy pixels. One main difference to the original Medipix ASIC is the ability to process both positive and negative input signals determined by the collected charge type. The minimum usable threshold is $\approx$ 1100 e$^-$, which corresponds to a minimum photon energy of $\approx$ 4 keV.

After the success of the Medipix 2 ASIC, the Medipix3 collaboration was formed in 2005 to develop the Medipix3 ASIC using 130 nm CMOS technology. The number of pixels and pixel size was kept the same as the previous ASIC with some improvements in performance, The ability to continuously read and write was implemented using two counters per 55 μm pixel, one which counts while the other reads out. Medipix3 implemented colour imaging and dead time free operation. The minimum operating threshold was reduced to 700 e$^-$, equivalent to a photon energy of $\approx$ 2.5 keV.

6.1.2 Timepix

Timepix [78] ASICs were developed in conjunction with the Medipix2 upgrade to extract more information from the data collected. The Timepix shares the same basic architecture as the Medipix2 ASIC for use with pre-existing readout systems and to build on already developed processing techniques. It has the same 256 x 256 pixel array with a pixel pitch of 55 μm. A modification of the original ASIC to include a 100 MHz internal clock provided the ability to track timing data by measuring the number of clock ticks from the moment the particle is detected until the shutter is closed, where the Medipix2 ASIC only counted the number of hits while the shutter was open. This gives the user the ability to measure the time of arrival (TOA) of the particle. This timing information can subsequently be used for measuring the detected photon energy. This is done by using the Time-over–Threshold (ToT) function, where the ToT is measured as the time the induced signal is above the threshold. The deposited energy related to the ToT and be calculated assuming the detector has been calibrated. The minimum operating threshold is $\approx$ 600 e$^-$, equivalent to a photon energy of $\approx$ 2.2 keV. The pixels on the ASICs can be programmed to be in three different modes. Medipix mode (counting mode) where the pixel simply counts the number of hits above a given threshold. Time of Arrival mode, where the number of clock ticks is

counted from the moment the discriminator is fired until the shutter is closed, giving time of arrival information with a resolution of 10 ns. Finally the Time–over–Threshold mode, where the number of clock ticks is counted while the given threshold is exceeded, providing the energy deposition of the particle or photon [8].

Although the Timepix ASIC has increased capabilities compared with the Medipix ASICs it has a number of important limitations. There is a limit on the time resolution in the measurement of the Time of Arrival (ToA). This is due to the number of clock ticks available from the moment the discriminator is fired to the shutter closing, where the maximum count is 11810, with a maximum time acquisition window of 11.810 ms. The time resolution is also limited by the slow clock frequency of 100 MHz In the Time–over–Threshold (ToT) mode if a second particle hits the same pixel while the shutter is open the resultant measurement will be an accumulation of both hits, producing an error in the measurement. Care is taken to use a shutter time of an appropriate value. If the hit rate is high then this becomes a huge problem as the shutter time needs to be less than the hit rate, but greater than the pulse duration. Each pixel is limited to using one of the modes available.

Therefore the Timepix3 ASIC [59] was established to mitigate these shortcomings as part of the Medipix3 upgrade using 130 nm CMOS technology. Timepix3 unlike the other ASICs, uses a data–driven architecture [9] to enable the readout and data acquisition to be simultaneous, keeping the pixels sensitive at all times. This enables the recording of ToA and ToT simultaneously in each pixel. To improve the time resolution capabilities of this ASIC the ToA information is recorded in a 14–bit register at an initial rate of 40 MHz, which can be refined to 640 MHz with the use of 4 further bits. This provides a time resolution of 1.5625 ns compared with the previous ASIC which has a 10 ns time resolution. As well as simultaneous measurements of ToA and ToT the ASIC has been designed to output data only from pixels which have seen an event, reducing the dead–time per pixel to 475 ns compared to 300 μs in Timepix [9]. To optimise the resources of the ASIC it has been split up into super pixel structures of 2 x 4 pixels. The pixels in each super pixel share the same readout logic which transport data to the End–of–Column. The minimum operating threshold was again reduced to $\approx$ 500 e⁻, equivalent to a photon with an energy of around 1.8 keV.

6.1.3 Velopix

The Velopix ASIC [82] is based on the design of the Timepix3 ASIC. One main difference between the ASICs is the data which is transmitted off the ASIC. The Velopix does not transmit any charge information off the ASIC but does use this information internally in each pixel [82]. The time resolution of this ASIC is degraded by a factor 16 compared with the Timepix3 ASIC to 25 ns. However it was designed to work in harsh radiation environments as required at the LHC. The ASIC has been designed to collect only negative

signal pulses from electron collection in the sensor. For the recent upgrade in 2018 these ASICs are capable of reading out every track provided by the CERN SPS H–6 beamline at maximum intensity. The ASIC can deal with a peak hit rate of 900 Mhits/s [82].

6.2   Why Small Pixels with Gain?

As discussed in the previous sections, the Medipix family of ASICs can be very useful for imaging and spectroscopy, with the Timepix ASIC useful for timing measurements and energy calibration. The evolution of the ASICs has seen the decrease in the overall electrical noise from the ASIC as well as the time resolution improvement of the Timepix.

Introducing a detector with an internal gain for each pixel can have many advantages. Firstly at the moment the minimum detectable signal is limited by the electronic noise and thus the threshold set on the ASIC. If the sensor used had an internal gain mechanism, assuming no increase in the electrical noise, the minimum detectable signal would decrease by an amount proportional on the gain. So if we have a noise floor of $\approx$ 500 e$^-$ for the Timepix3 ASIC, and a sensor with an internal gain of 10. The minimum detectable signal should now be of the order of 50 e$^-$ multiplied by a factor determined by the threshold, where the factor depends on the application. This factor can be very low for synchrotron applications but may be in the region of 5 to 10 for high energy physics experiments. This begins the search for sub–threshold signals. In reality there will be some noise associated with this gain mechanism therefore it is unlikely that the minimum detectable signal will scale directly with gain.

There are many manufacturers that produce LGAD's with a range of gain values. For the best time resolution it is said that an optimal gain of 20 is required [66] with the detector thickness of the order 50 μm. Using this value for gain and the Timepix3 ASIC we could potentially have the ability to detect single photons of less than 1 keV. To note there is a plan to upgrade the Timepix3 ASIC to the Timepix4 ASIC. This would have an estimated minimum threshold of < 300 e$^-$ with an increase in time resolution to values less than 200 ps.

The perfect detector would have high segmentation and high time resolution simultaneously which is not achievable at the moment. With the ability to detect sub–threshold signals it is hoped that LGAD detectors bonded to Timepix3 ASICs can detect < 1 keV X–rays at synchrotron sources. Working with the Medipix Collaboration, the Large Hadron Collider beauty (LHCb) experiment at CERN has produced a spin off ASIC known as the Velopix.

Looking toward the future upgrades the LHCb collaboration are working on options to deal with the increased expected luminosity at the HL–LHC. In order to determine the separate vertices, appropriate timing information is required with high segmentation. The

Velopix ASIC sits around 5 mm from the beam [83], therefore a higher hit rate is expected than with the HGTD for the ATLAS experiment. To determine the particle tracks they must be highly segmented , $< 55\text{x}55$ μm pitch, with time resolutions of the order 100 ps [83]. To achieve this the LGAD sensor is considered to provide this timing information as well as being segmentable. An upgrade of the Velopix ASIC would be required to match the timing performance of the sensor.

6.3   LGAD Segmentation

This section will discuss the limitations in the segmentation of LGAD sensors. These limitations include effective fill factor reduction with non uniformity in the gain. The segmentation limit will be discussed where issues involved with the gain mechanism when pixels are reduced down to a pitch of 55 μm. Devices have been fabricated in array structures with varying pixel size. With the minimum pixel size of 55 μm. Brief results of these tested arrays will be presented, where most results have been produced through TCAD simulations.

6.3.1   Fill Factor

In non–gain devices the total detector area is sensitive to incoming particles, thus a 100% pixel efficiency is possible. In the instance where charge is deposited between pixels, the charge collected is shared between these pixels, with the relative charge being used to calculate the position of the incident particle. If the particle hits the detector at an angle and has a track which crosses many pixels the angle of the track can be inferred by measuring the collected charge at each pixel. This is also determined from the rise time for each pixel, where if the particle crosses a pixel near the electrode the generated charge will be collected quickly. However if the track crosses a pixel close to the backside of the sensor it will take a longer time to collect the generated charges.

For devices with internal gain the fill factor is reduced. This is due to the presence of two regions, a gain region within the pixel core and a no gain region (pixel border) which surrounds the pixel. The gain region is defined as the area which has a gain of greater than a specified fraction of the maximum gain. Results for the fill factor in this thesis are given in terms of 50% and 90% of the maximum gain in the pixel. The fill factor is determined as

$$\text{Fill Factor} = \frac{\text{Gain Area}}{\text{Total Area}}.$$

The pixel border is a region where the gain=1, as the carriers generated in this area do not undergo multiplication as these are collected through the junction implant and do not pass through the high field (Gain) region. It is necessary to have this border region to host
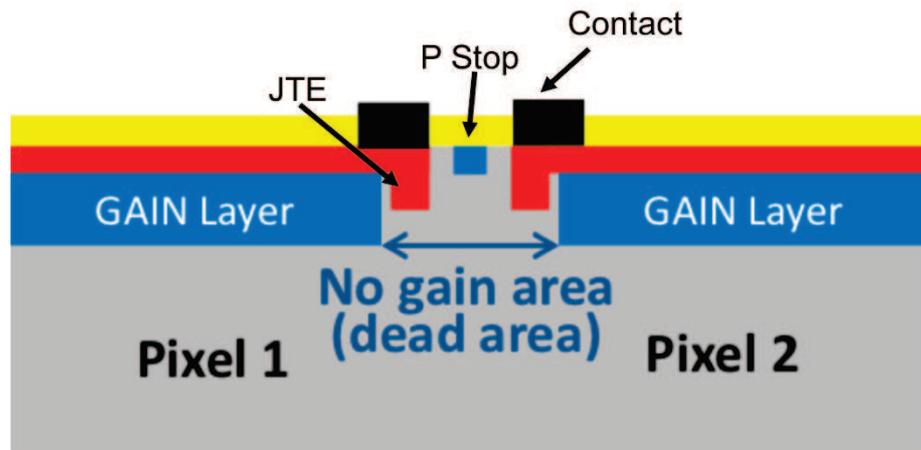
Figure 6.2: Fill factor reduction caused by Pixel Border dead region.

important structures such as the JTE, which is used to control the field within the pixel to provide a higher breakdown voltage. There must also be a region where the pixels can be isolated by either a p-stop implant or a p-spray. The pixel border width is a limiting factor on the possible fill factor. With a constant pixel border width, as the pixel size decreases the fill factor also decreases. The basic structure of two adjacent pixels is shown in figure 6.2, highlighting the non-gain region between the pixels.

For devices compatible with the Timepix3 ASIC an inter pixel spacing had to be established which could house all of the necessary device structures. TCAD simulation was performed to evaluate this. As previously mentioned the fill factor is dominated by this border therefore the minimum distance required was chosen. Simulations were performed using previously discussed doping profiles for the gain region and JTE structures. The simulations used two JTE widths as there was some evidence that this had an effect on the devices electrical performance. The widths chosen were 5 and 10 μm. The smallest JTE width used thus far on fabricated devices is 10 μm. From simulation it was found that 5 μm was the smallest achievable JTE size due to lateral diffusion as well as the minimum feature size possible through photolithography.

To obtain the effective fill factor for various size pixels, TCAD simulation was performed at the border of two pixels where each pixel and adjacent border was 40-50 μm. The device simulation was performed by applying a negative voltage to the backside and reading the current from each pixel separately. This was performed up to 300 V to obtain adequate electric fields. The nominal border is defined as the distance between the inner most part of the JTE implant and the edge of the pixel. The distance between the outer most part of the JTE and the edge of the pixel was chosen to be 9.5 μm. This means the nominal border is directly related to the size of the JTE. The fill factor as a function of pixel size is shown in figure 6.3a, where the nominal border is 14.5 μm and 19.5 μm for a pixel with a JTE width of 5 μm and 10 μm respectively. As expected the fill factor increases as the size
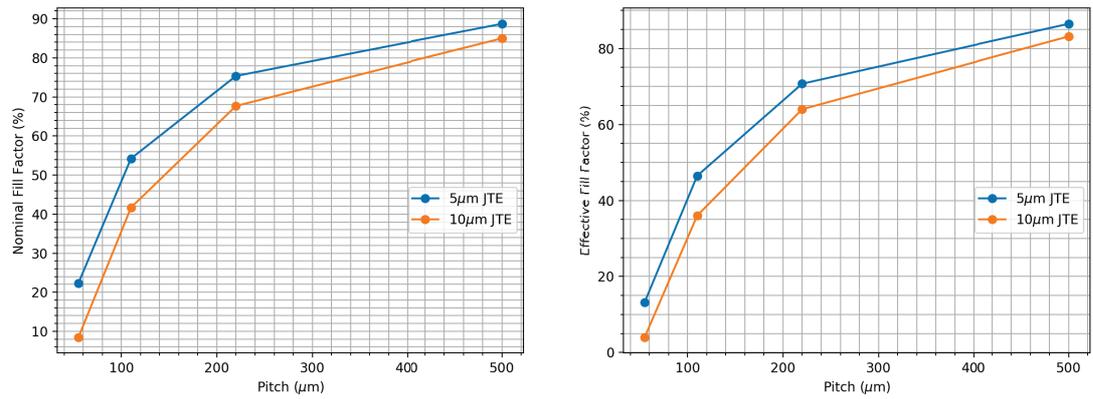
(a) Nominal Fill Factor, the Fill Factor de– fined by the geometry of the implants.

(b) Effective Fill Factor, the Fill Factor de– fined by the high Electric field region in the pixel as illustrated in figure 6.4a.

Figure 6.3: Fill Factor as a function of pixel size.

of the pixel increases. At a pixel pitch of 55 μm the nominal fill factor is $< 20\%$ for a JTE width of 5 μm, however this goes down to $< 10\%$ with a JTE width of 10 μm. The highest fill factor possible from these results is 89% for a JTE width of 5 μm and a pixel pitch of 500 μm.

However, it was found that this fill factor was actually less than anticipated due to the lateral diffusion of the JTE implant inwards, towards the gain region. This causes the gain region to shrink and an effective border is created. The effective border is calculated to be the distance from a region of high field (gain region) to the edge of the pixel. An example of the electric field profiles at the depth of the highest field across the pixels is shown in figure 6.4a at 300 V. The gain region is highlighted as the region with the high field (left of the image), this then drops off around the JTE. There is a small spike in electric field at the edges of the JTE, then a region of low electric field between the pixels. An identical pattern is shown for the second pixel. The designed distance between the pixels is 19 μm but was measured to be 12 μm, defined as the distance between the edge of the JTE of one pixel to the edge of the JTE of the other pixel. The electric field distribution in 2D is shown in figure 6.4b, which shows regions of Gain and no Gain. It should be noted that simulations were produced using a p–spray between the pixels as this is the process used in the fabrication of LGAD pixel sensors at Micron Semiconductor. It may be possible to reduce the pixel to pixel spacing by using a p–stop, a more defined region to isolate the pixels, however at this moment this method has not been explored through simulation as this would require a significant change in the fabrication process flow.

The Fill factor as a function of pixel size is shown in figure 6.3b, where the effective border is 17.5 μm and 22 μm for a pixel with a JTE width of 5 μm and 10 μm respectively. Figure 6.5 shows the nominal and effective dead border regions for one pixel. Plotted is
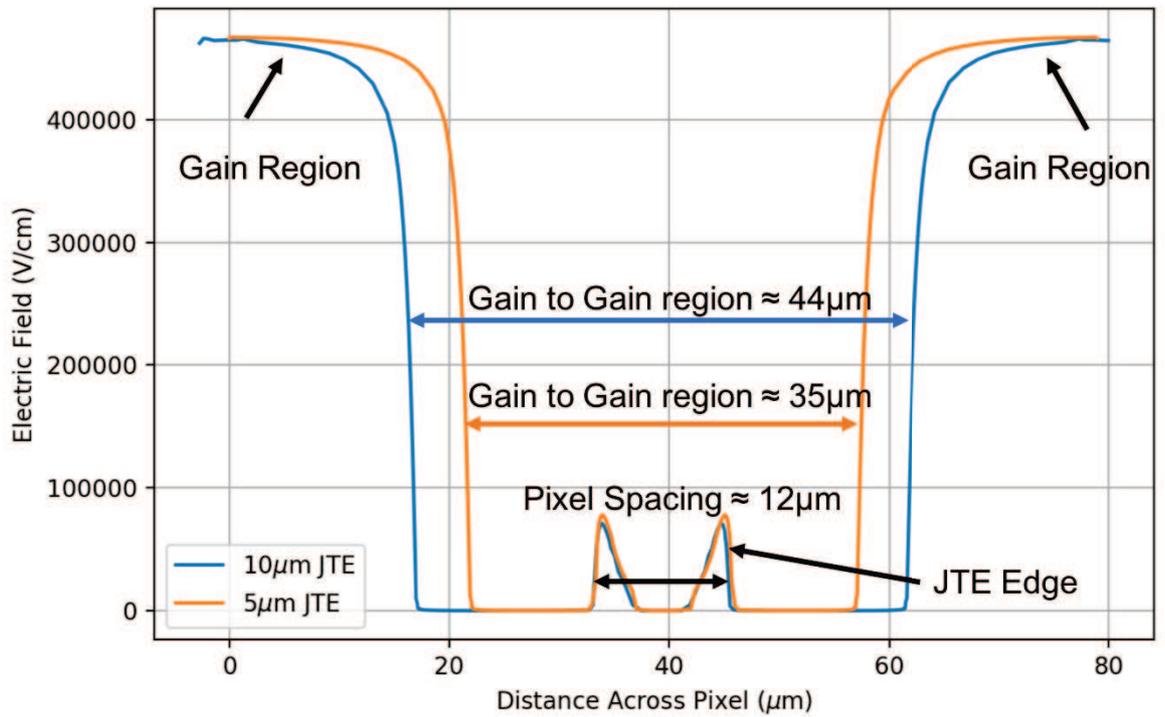
(a) Electric field profiles across two pixels at the region of highest field in the gain layer at 300V.
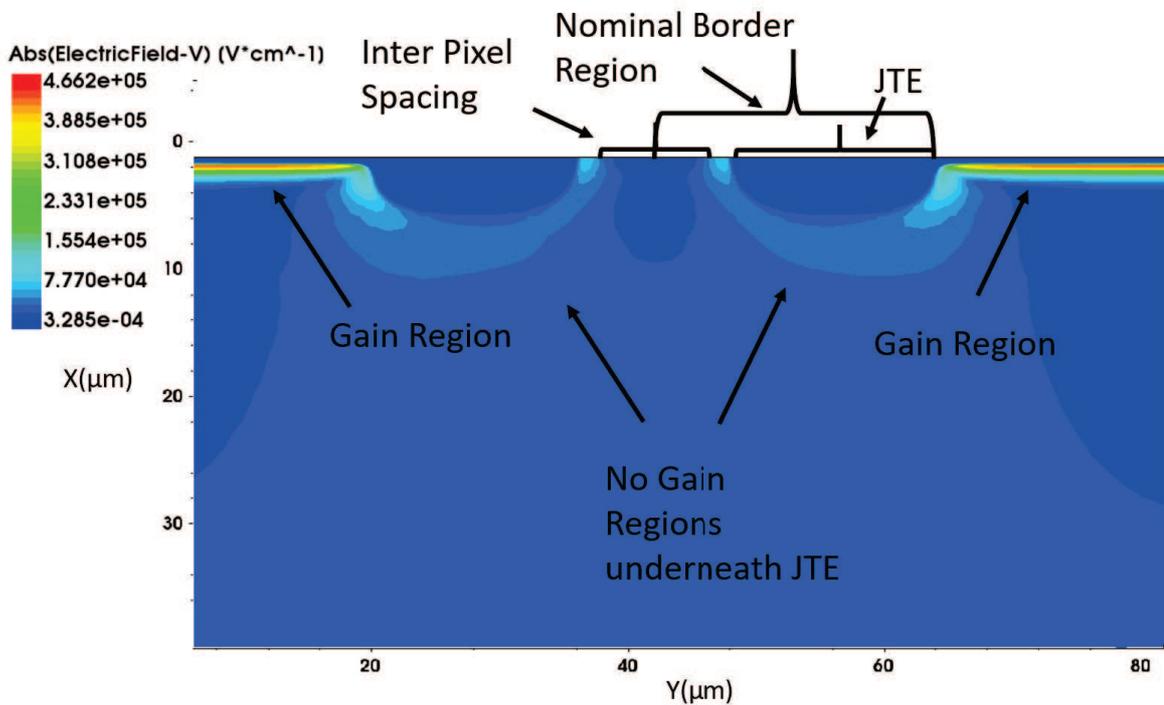
(b) Electric field profiles across two pixels with a JTE of 5 μm at a bias voltage of 300 V.

Figure 6.4: Electric field profiles at the border of two pixels with a JTE of 5 μm at a bias voltage of 300 V.
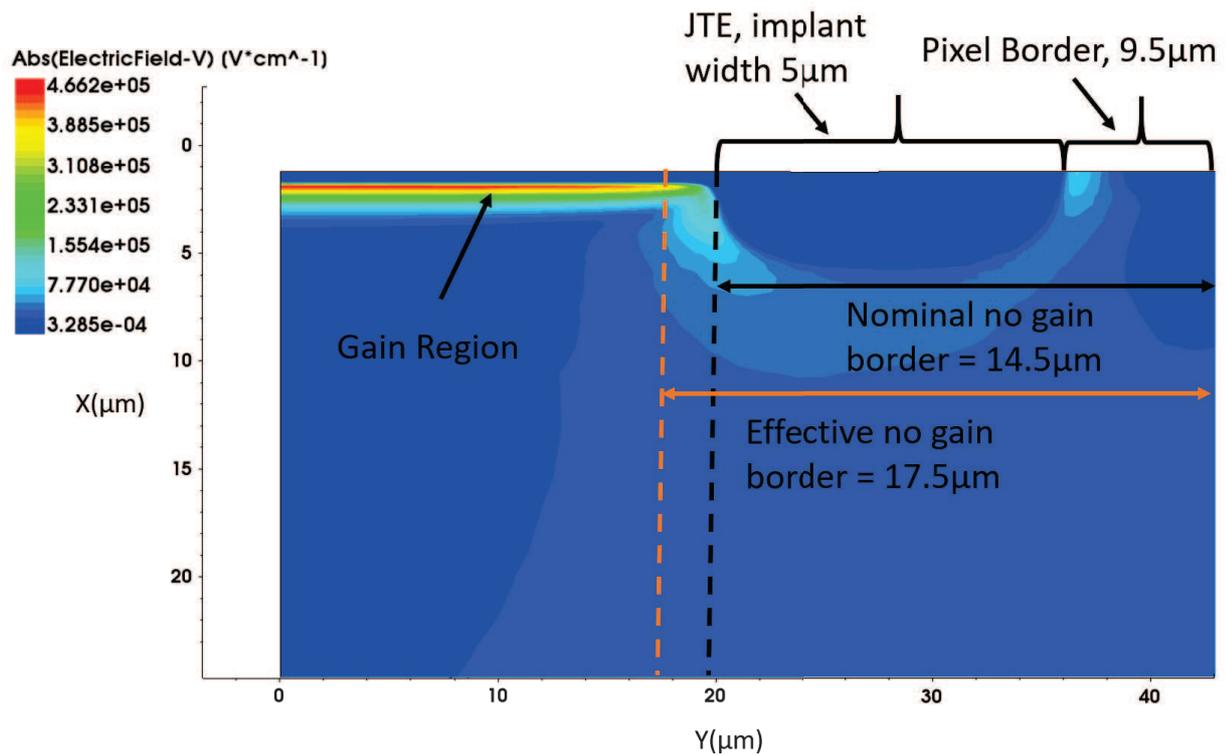
Figure 6.5: Border region of a pixel with a JTE of 5 μm at a bias voltage of 300 V

the electric field as a function of position across the pixel. The electric field is highest at the gain region. The area underneath the gain layer is the gain region, meaning any charge in this region will be multiplied. Any charge deposited outwith this region will not undergo multiplication as charge will be collected in the border region.

6.3.2   Gain

As it was shown the fill factor varies dramatically with pixel size for LGAD sensors, due to the non–gain region between the pixels. To fully understand the effect of the fill factor on the LGAD performance it was necessary to simulate the collection of electrons for a variety of pixel sizes and border structures. The structures were defined based on the pixel size requirements for the Timepix3 ASIC. Only a single pixel was simulated in this study to get an idea of the gain variation for different design types. Although this gives a good indication of how pixellated structures would perform.

The pixel pitch used in the simulation was 55 μm, 110 μm and 220 μm. The size of the total pixel implant, defined as the distance between the outer most region of the JTE at one side of the pixel to the other side. For an inter pixel spacing of 19 μm for all pixel pitches, this gives a total pixel size of 36 μm, 91 μm and 201 μm respectively. This study used two JTE widths of 5 and 10 μm, same as for the fill factor measurements. The multiplication

implant size was the size of the pixel minus the size of the two JTE's. For a JTE width of 5 μm, the multiplication implant size was 26 μm, 81 μm and 191 μm for the three pixel sizes respectively. For a JTE width of 10 μm, the multiplication implant size was 16 μm, 71 μm and 181 μm for the three pixel sizes respectively.

Electrical simulations were performed on the six device designs up to a bias voltage of 500 V. The detector thickness is 200 μm with the total width being 20 μm larger than the pixel size. Charge collection studies were performed using transient simulation in the TCAD software. Injection of charge was implemented using two strategies. The first strategy was to induce charge in a small region near the backside of the detector in the centre of the pixel, imitating the charge deposited by an Alpha particle or Red Light. The electrons must drift the full thickness of the detector before any multiplication takes place. If there is any gain mechanism as you can see holes drift from the front side to the backside of the detector after charge generation in the gain region. The second strategy was to induce charge evenly throughout the detector thickness, imitating a MIP, in the centre of the pixel. The gain for both methods should be the same for a large pad but small pixel effects may show differing results.

The gain was calculated for both methods by comparing the charge collected by each pixel for each method to a no gain pixel of the same dimensions at the same Bias Voltage. The charge collected does not change with increasing voltage for a no gain device above the full depletion voltage. Figure 6.6 shows the gain as a function of voltage for devices injected with an Alpha particle like charge. The device with a 55 μm pitch and JTE width of 10 μm shows no gain across the entire voltage range, however with a JTE width of 5 μm there is a small amount of gain at voltages $>$ 400 V up to a value for gain of $\approx$ 1.5. For the 110 μm pitch devices they show a high gain of $>$ 4 for voltages $>$ 100 V, up to a maximum gain of 10 at 500 V. Below 100 V there is a decrease in the gain. This effect is more prominent for the device with the larger, 10 μm, JTE implant. This is an indication that the JTE structures have an effect on the charge collection mechanism. Both 220 μm pixels show gain at all voltages with no drop in voltage at 50 V as seen with the smaller designs.

Figure 6.7 shows the gain as a function of voltage for devices injected with a MIP like particle. Unlike with Alpha particle injection, all devices show some gain. For the 55 μm pitch the gain is fairly low, around 1.3 over the full voltage range for a JTE width of 10 μm. However the gain increases above 300 V, similar to the Alpha measurements, for a JTE width of 5 μm up to a gain value of 2.5. Again the 110 μm pitch devices show high gains of $>$ 4 for voltages above 100 V up to values of 10 at 500 V. Lower gain values are seen for a 50 V bias, where the effect is magnified for JTE widths of 10 μm. This time the 220 μm pixels show some loss of gain at 50 V, this could be due to some distortion in the field that allows electrons to spread as they drift across the detector. This effect is more prominent
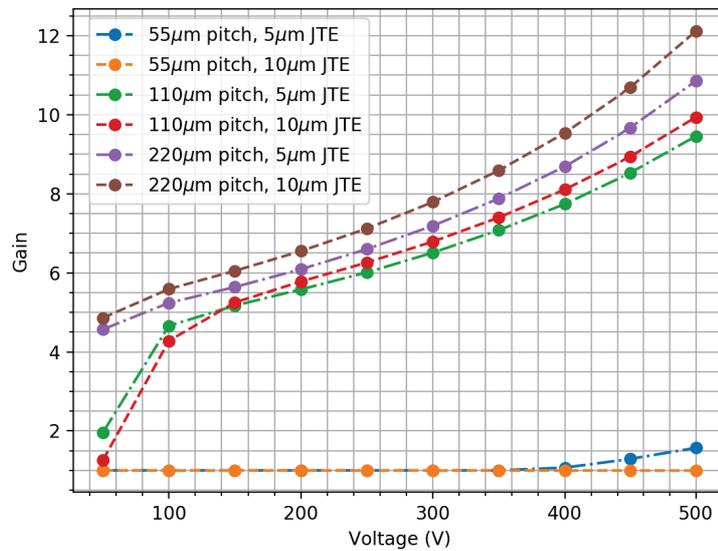
Figure 6.6: Simulated gain against voltage for devices injected with an Alpha Particle in the centre of the pixel at at depth of 100 μm.

for charge which is generated further from the collecting electrode.

It should be noted that total gain at 500 V for both sets of JTE's is slightly different. For a JTE width of 10 μm the gain is slightly higher with respect to the 5 μm JTE, this is due to the simulation process. The doping profiles for each device are created separately through a process of Monte Carlo simulations. Small random variations in the doping profiles are produced from one run to the next. Due to the sensitive nature of the gain mechanism, small changes in the doping profiles can alter the final gain. By operating devices at various voltages one can provide the same gain across many devices and pixels.

6.3.3 Charge Collection

To understand the loss of gain for small pixels it's important to simulate the path of charges after initial injection. An initial transient simulation will provide details of the total charge collected at a given electrode but will not show the full path of those charges from the generation site to the electrode. This full path can be simulated in TCAD as a series of snapshots in time where the electron and hole concentrations can be plotted as a function of position. This has been shown before in a previous chapter. Where for LGADs it's clear there is a gain mechanism when you see holes travelling from the front side to the backside of a detector after generation in the gain region.

The user has control over the time period of these snapshots to best show the reality of what is happening during charge collection. There are two cases to study here. The devices with a pitch of 55 μm show little or no gain at all voltage steps for both JTE widths. However
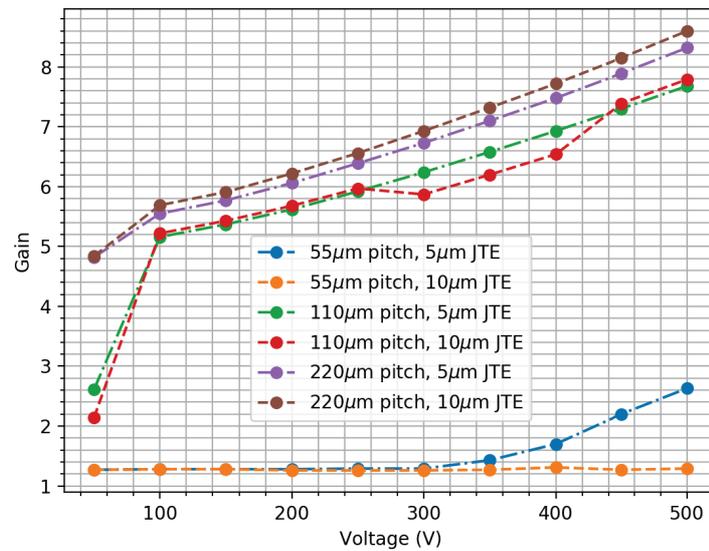
Figure 6.7: Simulated gain against voltage for devices injected with a MIP like particle in the centre of the pixel.

the devices with a pitch of 110 μm show no gain initially at low voltages but recover the gain at higher voltages. The first thing that one may consider to understand is the gain mechanism. This is the electric field required to produce avalanche multiplication. A plot of electric field as a function of depth in the device is shown in figure 6.8 for all devices at 150 V measured at the centre of the pixel, where a depth of 0 μm is at the surface of the detector. Both the devices with a pitch of 110 μm show the same peak electric field of $> 4 \times 10^5$ V/cm, high enough for impact ionisation and gain. However both the 55 μm pitch devices show significantly lower electric fields of $< 4 \times 10^5$ V/cm. The device with a JTE width of 10 μm has a peak electric field of 2.5×10^5 V/cm, almost half of the value for the same device with a pitch of 110 μm. This is the first indication why there may be a reduction or loss of gain for these devices at this pitch. The peak electric field is recovered for all devices at 500 V, however as gain is not present there must be some other process which causes the gain to disappear for the 55 μm pixels.

Looking at the 2D structures in the TCAD software the user can define what physical information is shown. The software also shows depletion regions as default. Using this information the electron current density was plotted as a function of position. An interesting feature appears below the multiplication implant at the centre of the pixel for both devices with a 55 μm pitch at all voltages. It appears that there is a region between the two JTE structures which has a collection of charge carriers which have not been removed during the reverse biasing of the device (un–depleted region), shown in figure 6.9a. The closer the structures are together the larger the effect. When plotted for electron current density we see a region of low electron density in the un–depleted region. This is because this region
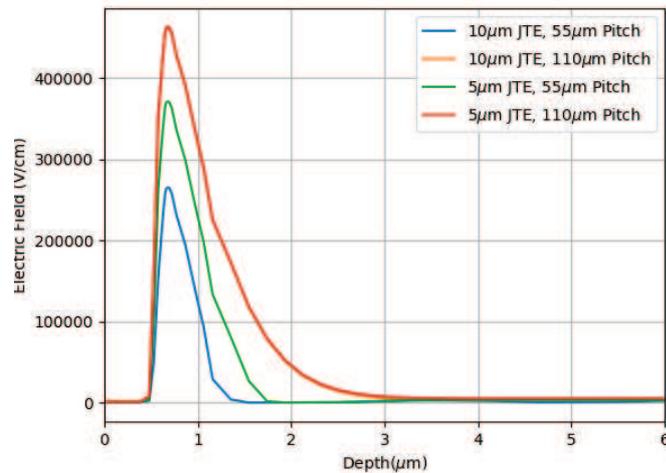
Figure 6.8: Electric field across the gain region in the centre of the pixel for various pixel designs at 150 V. The electric field lines for the 110 μm pitch overlap, shown in orange and red.

is p–type which has an majority charge carrier of holes. If a region is deplete of charge carriers there should be little or no variation in the charge carrier density. An example of this effect is shown in figure 6.9a for a 55 μm pitch device with a JTE width of 5 μm at 50 V.

If the voltage is increased the un–depleted region shrinks for both devices. However, even at 350 V as shown in figure 6.9b, the un–depleted region is still visible. This has an effect on the charge collection of electrons. Electrons will not pass through a region of un–depleted silicon as this has a low electric field in comparison to the JTE structures close by. This effect is shown in a series of images in figure 6.10 for the electron drift as a function of time for the region around the collecting electrode of a 55 μm pitch device with a JTE width of 5 μm at 300 V. These images are formed by transient simulation of an Alpha like particle near the backside of the detector, with a time delay of 0.5 ns per image.

This figure clearly shows the drift of electrons from the generation site towards the electrode, where a higher concentration of electrons is shown by a red/yellow colour. A low concentration is denoted by the blue region. At 500 ps after charge injection the first electrons have already arrived at the collecting electrode. The electrons drift vertically upwards towards the high field region, but at a depth of around 5 μm the electrons appear to move in almost a horizontal direction towards the JTE structures. No electrons are seen to pass through the high field region over the full collection period. Since all electrons are collected via the JTE region they do not undergo any avalanche multiplication and so the gain of this detector is 1. The same charge collected as a standard PIN diode of the same dimensions.

This effect is seen for both JTE widths at this pixel pitch. The electrons drift in this
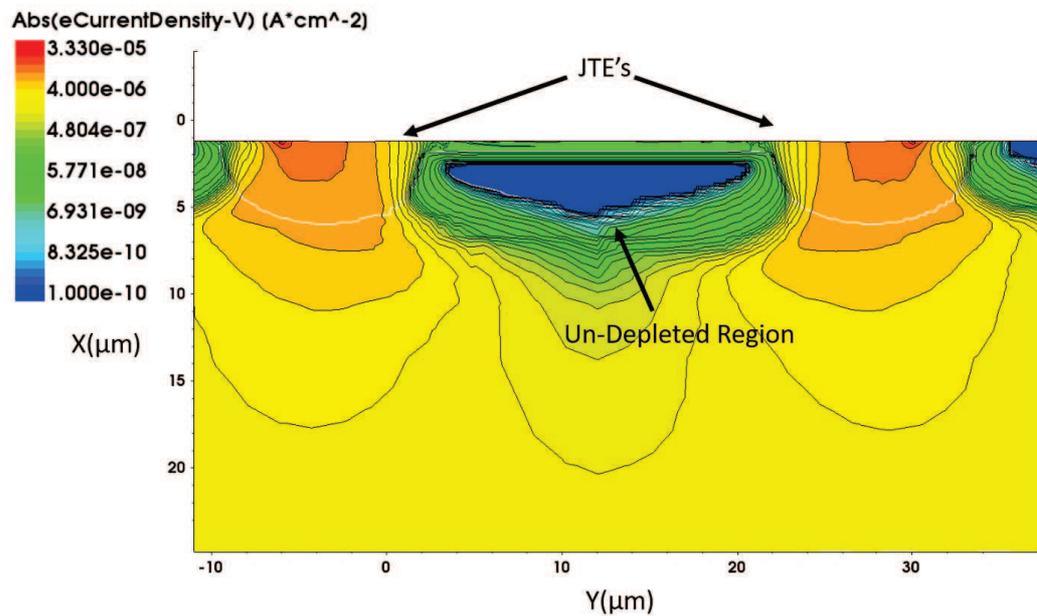
(a) Plot of Electron Density to show formation of an un–depleted region at a bias voltage of 50 V below the multiplication region in the centre of the pixel. The device shown has a 55 μm pitch with a JTE width of 5 μm.
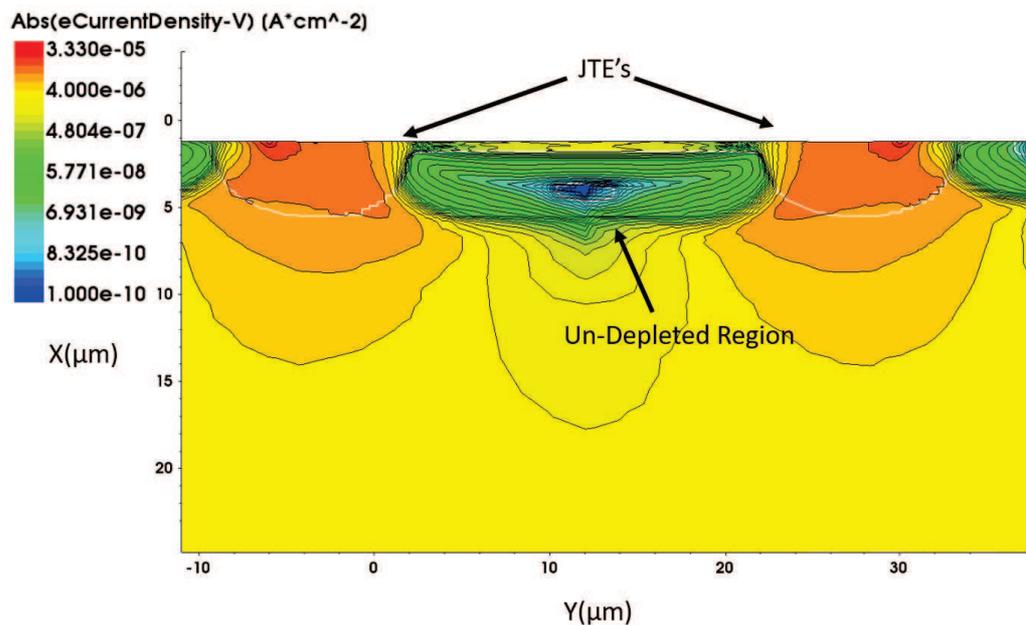
(b) Plot of Electron Density to show reduction of Un–depleted region in centre of a pixel at 350 V. The device shown has a 55 μm pitch with a JTE width of 5 μm.

Figure 6.9: Plots of Electron Density to show formation and reduction of Un–Depleted region between the JTE's by moving from 50 V to 350 V.
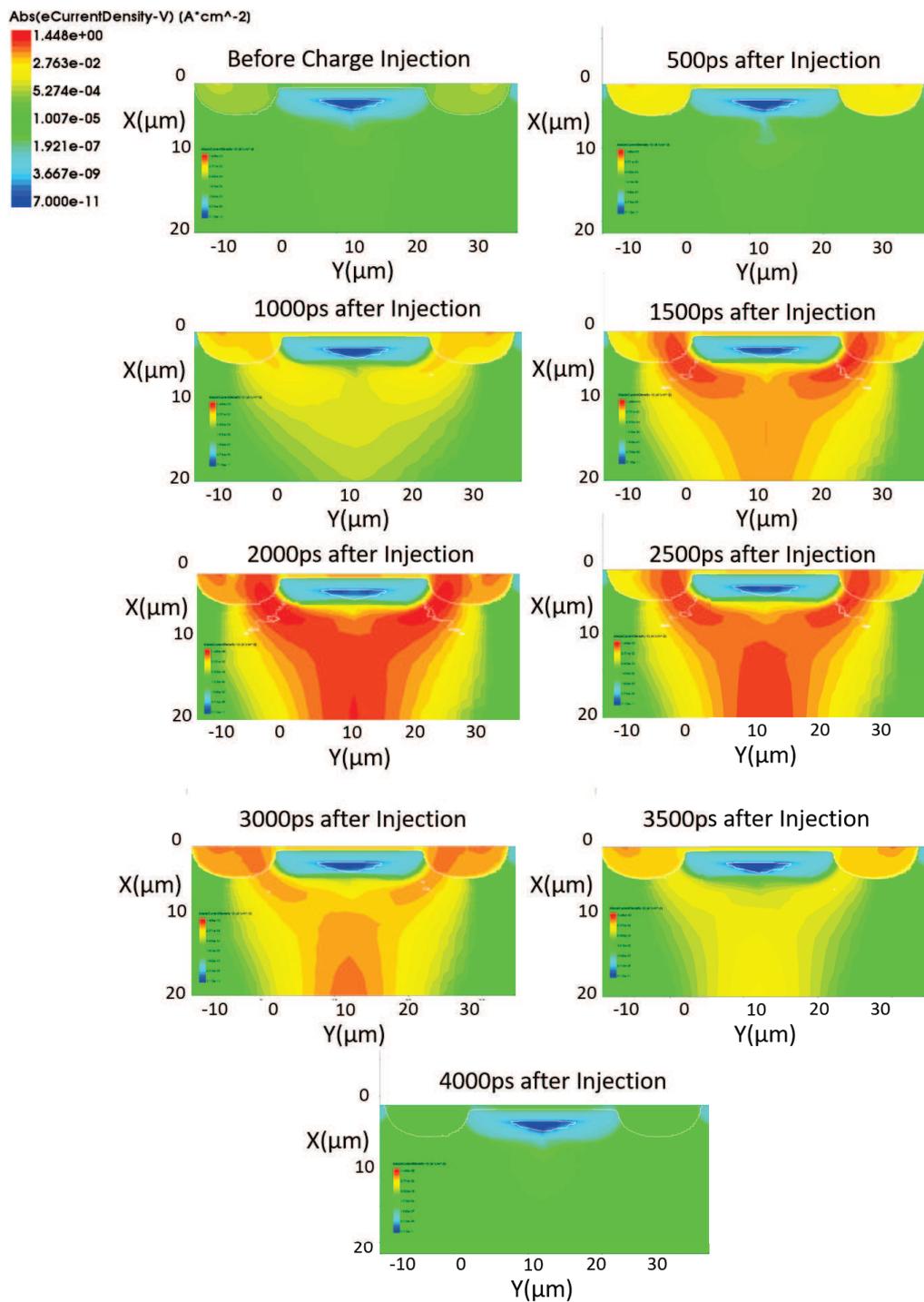
Figure 6.10: Electron charge density as a function of position for an extended time period after Injection of charge at 300V. The device shown has a 55 μm pitch with a JTE width of 5 μm. The device has a thickness of 200 μm but only the top 30 μm is shown.

way because of the electric field present around the electrode. The electric field is higher going towards the JTE structures than in the centre of the pixel, thus electrons favour this path. It was thought that increasing the voltage would mitigate this effect, however as the JTE size is close to the size of the pixel this effect dominates for all voltages tested for a JTE width of 10 μm. At 500 V some electrons are collected via the gain region for the device with a JTE of 5 μm producing a low value of gain.

This effect was then studied as we increased the pixel pitch to 110 μm. As we saw from previous simulations, the gain was recovered at increasing voltages. The width of the JTE did have an effect on the gain at low voltages. The electron current density was plotted as function of position to try and understand this, for a range of voltages and shown in figure 6.11 for a JTE width of 5 μm. At 50 V the electron current density underneath the pixel is still considerably lower than at the periphery of the pixel. However as we increase the voltage it's clear that there is a more uniform distribution of the electron density. At 100 V there appears to be a uniform distribution across the pixel, where the only regions of low electron density are at the very corners between the JTE and the gain region. This appears not to disappear even with increasing gain. If the charge is deposited in the centre of the pixel, above 200 V, all of the charge collected goes through the gain region and the expected gain is obtained.

The charge collection of electrons is shown as a function of voltage for both JTE widths for a pixel pitch of 110 μm in figures 6.12a and 6.12b respectively at 500 ps after charge injection. Looking at figure 6.12a for the 5 μm JTE we see that the electron paths change as the voltage is increased. At 50 V some of the electrons are attracted to the JTE structures, but the majority of charges flow vertically through the gain region. As we go up to 100 V the number of electrons attracted to the JTE structures has significantly decreased and by 150 V virtually none are seen to diverge from the gain region. This can be explained by the electric fields in the regions around the JTE. At low voltages these are higher than in the centre of the pixel so can attract electrons. At higher voltages, the electric field in the centre is higher and thus less electrons tend not to be attracted outwards, towards the JTE's.

Comparing to figure 6.12b, for the device with a 10 μm JTE we see very similar results. The differences being the contribution of electrons to the JTE at low voltages is greater than with the 5 μm JTE. The voltage required to overcome this effect is increased with increasing JTE width. Once up to 250 V all of the generated electrons flow through the gain region, hence the same gain for both devices at > 250 V.

Loss of gain at low voltages after Alpha Particle injection for the 110 μm pitch pixel can be explained by this flow of electrons. The number of electrons passing through the gain region is lower than that at higher voltages. Hence the gain is reduced, as this is defined as the amount of extra charge created after the initial particle collision.
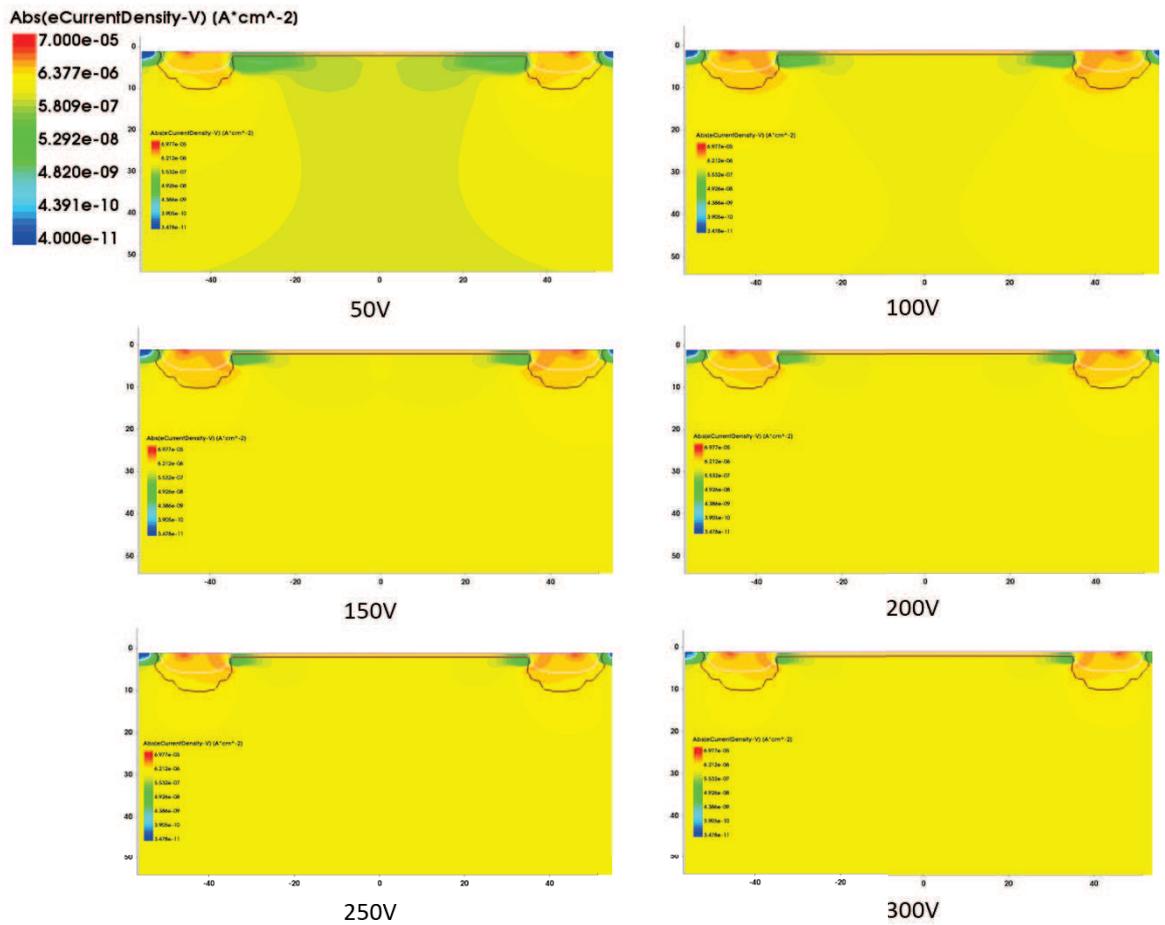
Figure 6.11: Electron current density as a function of position at a range of voltages with no injected charge. The device shown has a 110 μm pitch with a JTE width of 5 μm.
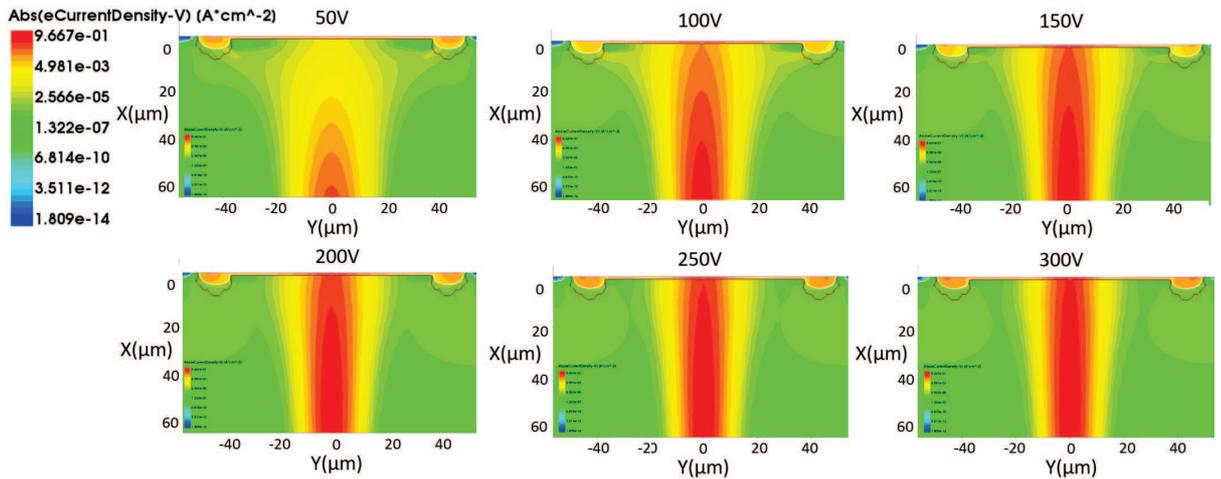
(a) Electron current density as a function of position for a range of voltages for a 110μm pixel with a JTE width of 5μm and 500 ps after charge injection.
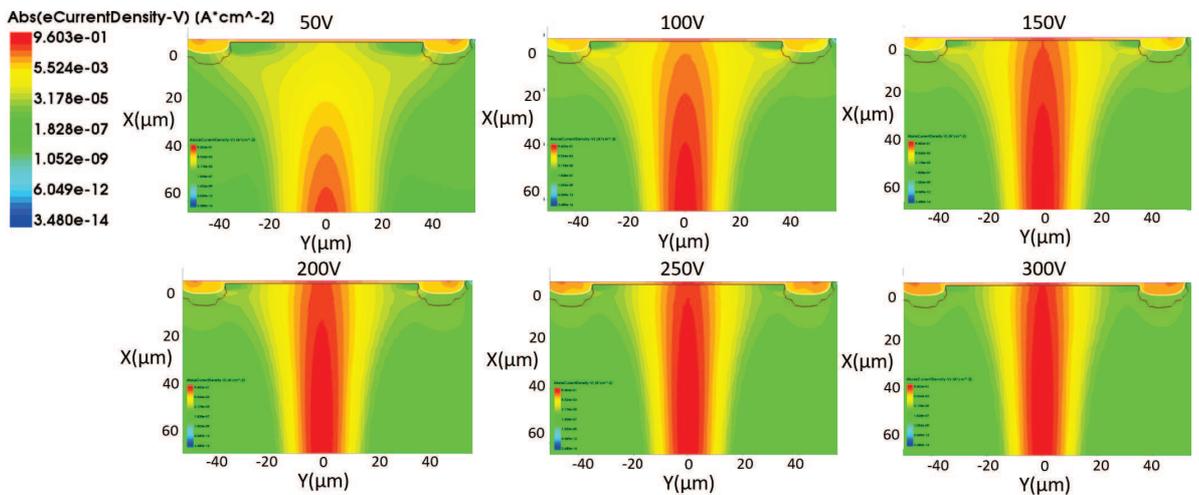
(b) Electron current density as a function of position for a range of voltages for a 110 μm pixel with a JTE width of 10 μm and 500 ps after charge injection

Figure 6.12: Electron current density as a function of position for a range of voltages for a 110 μm pixel with a JTE widths of 5 and 10 μm and 500 ps after charge injection.

Similar plots can be produced to investigate the charge collection after a MIP injection. The MIP was injected into the very centre of the pixel at the region of highest electric field. It was expected, after seeing the electron path following an Alpha particle injection, that the collection of electrons would be similar after a MIP interaction. In general this is true, where at the full voltage range all electrons generated deep in the bulk are collected via the JTE structures for a pixel pitch of 55 μm. However charges generated close to the surface are collected via the gain region. This results in the small amount of charge being multiplied, this corresponds to a low gain in the signal detected. The hole distribution at 500 ps after charge injection is shown in figure 6.13a for a voltage of 350 V and 55 μm pitch with a JTE of width 5 μm. It is clear that there is hole generation at the high field region as at this point in time any initial charges generated in this region would have drifted away. This can be seen in figure 6.13b which shows the same time stamp for a MIP injected into a non gain detector. The majority of generated holes from the initial injection of charge have been swept away from the front to the backside of the detector. There is an absence of generated holes from the multiplication region compared to that seen in the LGAD, figure 6.13a.

For the devices with a pitch of 110 μm a similar charge collection mechanism is seen for MIP injection as for Alpha. As with the Alpha injection of charge there is a lower value of gain at low voltages. This is due to the generated charge being collected at the JTE structures in the same process as described for the Alpha injection. However the gain loss is not so dramatic as not all charged deposited has the potential to flow through the JTE structures. The charge generated near the surface goes directly to the gain region and produces excess charge carriers. The same process seen in the 55 μm pitch design. There is a specific voltage for each JTE width that eliminates this behaviour and as seen before this voltage is dependent on sensor thickness. The larger thickness needing a higher voltage to fully deplete.

6.3.4   Gain Uniformity

The charge collection path of electrons in small pixels makes it difficult to produce a detector that has good uniformity with the present design of the periphery regions. It was shown that at high voltages the gain of the detector was as expected for charge deposited in the centre of the pixel for both Alpha and MIP like interactions. To understand the gain uniformity a simulation study was performed to look at the gain as a function of initial charge generation position. To simplify this only an Alpha particle interaction was studied. The study was performed for the 110 μm and 220 μm pixels for both JTE widths at 500 V, as it's important to understand the gain uniformity as a function of fill factor. The initial positions of charge interaction were separated by 10 μm steps, ranging from the edge to edge of each pixel either side of the centre of the pixel at a depth of 100 μm. Figure 6.14
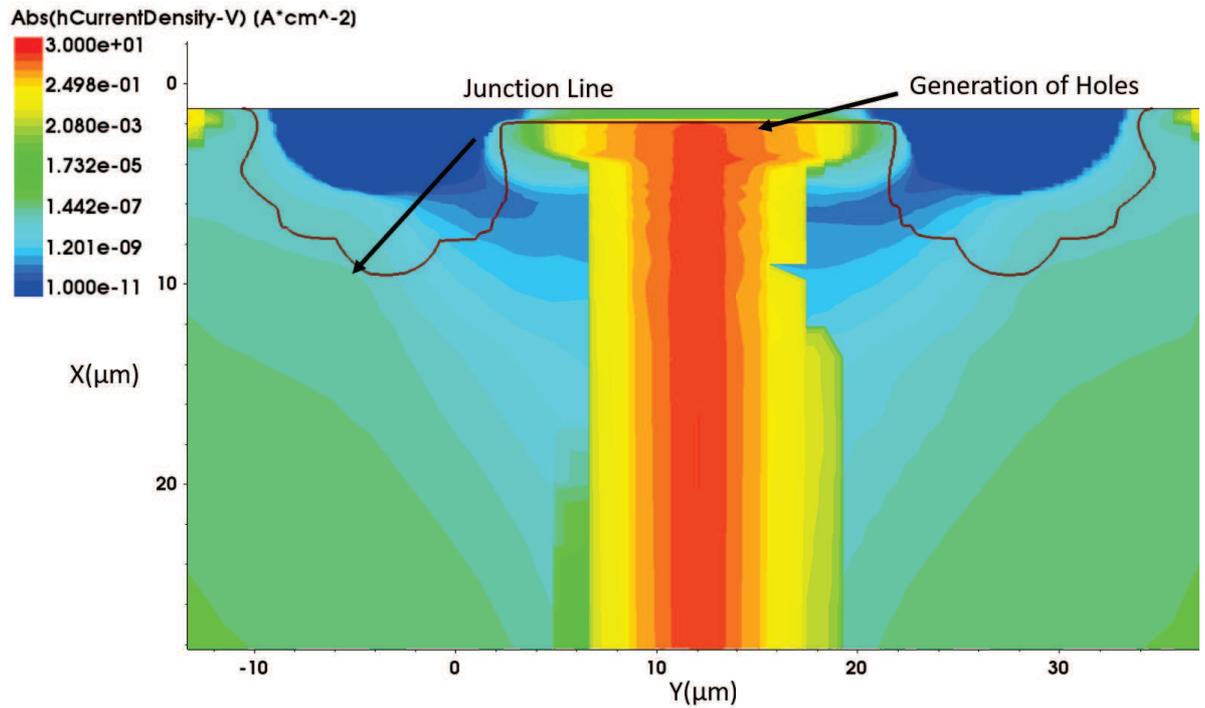
(a) Generation of holes in the gain region in the centre of the pixel at 350V, 500ps after MIP injection. Plot shows the effect on a single pixel where the centre of the image is the centre of the pixel
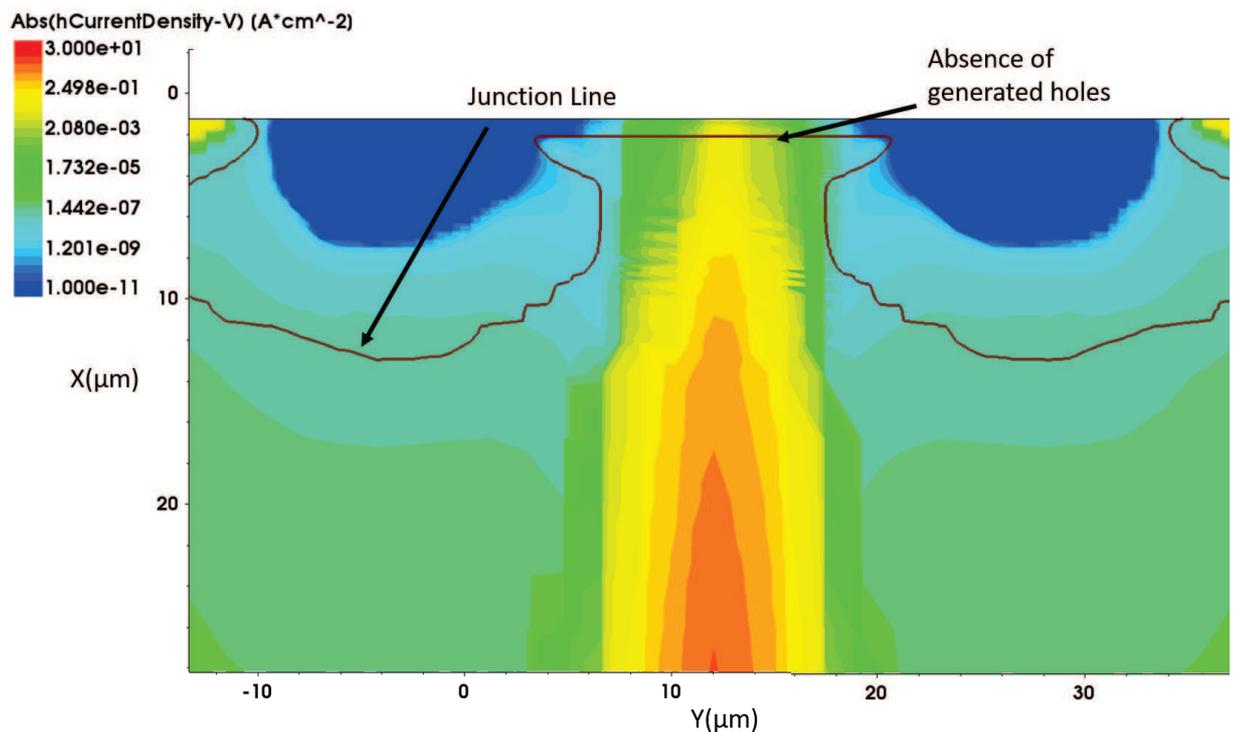
(b) Absence of holes at collecting electrode in a no gain sensor at 350V, 500ps after MIP injection. Plot shows the effect on a single pixel where the centre of the image is the centre of the pixel

Figure 6.13: Comparison of the hole current density for an LGAD and PIN diode at 350V, 500ps after MIP injection in the centre of the pixel. Both devices have a pixel pitch of 55 μm and a JTE width of 5 μm.
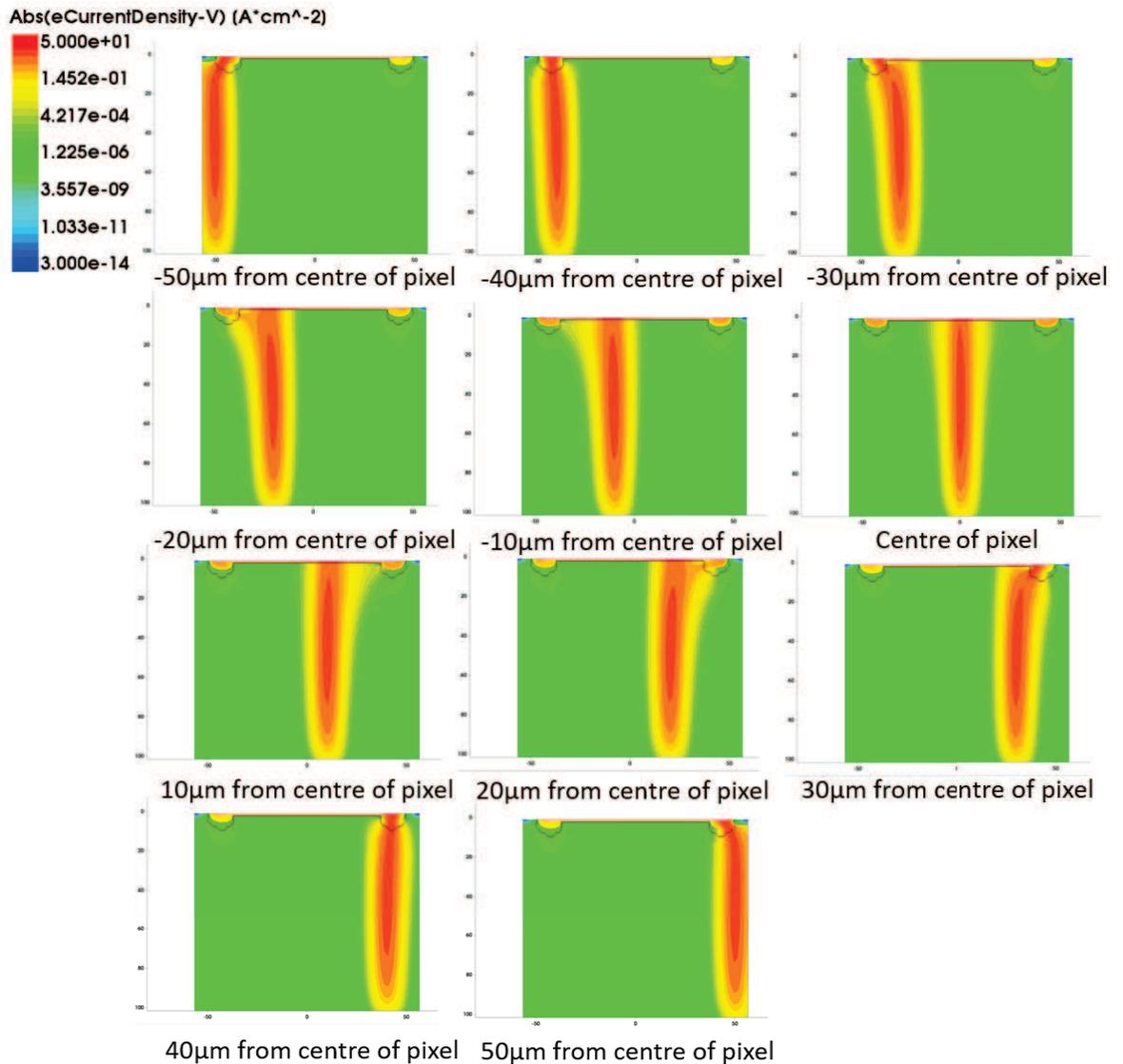
Figure 6.14: Electron charge density as a function of initial charge injection location for a pixel pitch of 110 μm and JTE width of 5 μm at a bias voltage of 500 V.

shows the electron charge density as a function of initial charge interaction position at 500 ps after the initial interaction for a 110 μm pixel with a JTE width of 5 μm. It's clear that the when the charge is deposited near the edges of the pixel that little or no electrons pass through the gain region. This is evident for charge injected at 40 and 50 μm from the centre. Charge deposited within 30 μm of the centre of the pixel starts to be collected via the gain region, where it is either all collected here or the electrons are split between the JTE and the gain region.

To understand the impact on the gain, the gain is plotted as a function of position relative to the centre of the pixel, shown in figure 6.15. At the edges of the pixel all charge is collected via the JTE structures and no gain is seen. As the charge injection location is moved closer to the centre of the pixel, the gain increases, as expected.

There is limited gain at around 20 μm from the centre of the 110 μm pixel for a JTE width of 10 μm and full gain at 10 μm either side of the centre. Using this information to calculate a new effective fill factor for the device with a JTE width of 10 μm, assuming an effective gain region at 90% of the peak height, we get a value of 2.9% fill factor which is much lower than expected. If we take 50% of the peak height we get a fill factor of 11.5%. For a JTE width of 5 μm there is gain at around 30 μm from the centre of the pixel. Where the peak gain is achieved in the region 15 μm either side of the pixel centre. Calculating using the same method for a JTE width of 5 μm we get a slightly increased effective fill factor of 6.8% and 17.7% for a peak height of 90% and 50% respectively.

The larger, 220 μm pixels show better gain uniformity due to their size. For the 220 μm pixel with a JTE width of 5 μm we see an effective fill factor of 31.3% and 47.4% for a peak height of 90% and 50% respectively. This slightly worsens for the 10 μm wide JTE giving values of 24.8% and 42% for the fill factor at a peak height of 90% and 50% respectively. With respect to the 110 μm pixels there is larger region of constant gain for these pixel which shows very good gain uniformity in the centre of the pixel. The 90% value gives us the gain at which the detector is designed for. In some applications a gain of 5 will suffice thus the 50% value is given. Using Timepix in the counting mode we can extract information on particles which have an energy that is five times lower than the minimum threshold over $\approx$ 18% of the 110 μm pixel. Assuming we have many hits this may be enough to create an image of reasonable quality. This does however, make it difficult to perform an energy calibration.

There is a transition region between the edge of the pixel and the centre where there is little or no gain. This transition region is an effect of the JTE width, and is reduced with smaller designs.

6.3.5   Improvements & Future Work

One way to improve gain uniformity is to increase the size of the pixel. As seen from the previous measurements this is quite effective going from 55 μm up to 110 μm. Increasing this up to 220 μm, compatible with Timepix ASICs, does indeed improve the uniformity but decreases the spatial resolution of the measurements. We saw that decreasing the size of the JTE's increased the gain uniformity by a factor of $\approx$ 2 . To increase the effective fill factor the contribution of the electrons drifting outwith the gain area must be reduced. One suggestion is to remove the JTE's from each pixel and have only an overlapping junction. Each pixel now would have a lower breakdown voltage, but would need a lower voltage to achieve the same gain value. This has been done for sensors fabricated by CNM for the HGTD module [84]. This design includes a single guard ring around the array of pixels.

This has been simulated where the total pixel size is 36 μm, 91 μm and 201 μm, allowing a pixel to pixel gap of 19 μm for the three pixel pitch's. The gain as a function of
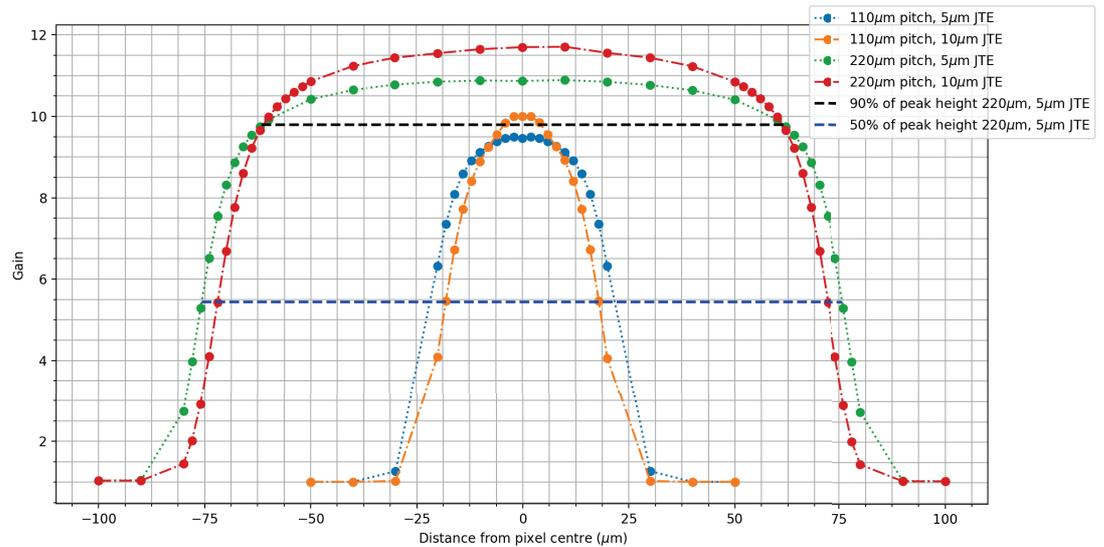
Figure 6.15: Gain as function of charge injection location relative to the centre of the pixel for various pixel designs.

achievable bias voltage is plotted in figure 6.17. Charge was injected via an Alpha particle like charge. Comparing the 110 μm results with the devices using a JTE in figure 6.6, we see that the effect on the gain at lower voltages is decreased. At 50 V the gain is around twice that of the gain produced from a detector with a JTE. This shows the reduced effect of charge collection via the non–gain edges of the detector. The 55 μm pixels show huge improvements with gain values of > 10 achievable, where the maximum gain for a device with a JTE was < 2. We see that the total gain reduces as a function of pixel size, this is not only due to small variations in the doping profile as seen for the different JTE widths.

The results for gain uniformity at 350 V are shown in figure 6.17, for charge injected along a pixel at different locations relative to the pixel centre. The effective fill factor for this design of a 55 μm pixel can be calculated to be 9.4% and 28.2% for at a peak height of 90% and 50% respectively. This is a significant improvement compared with the devices that included JTE structures. However the breakdown voltage of these devices is reduced to < 400V, where the breakdown voltage with a JTE is > 500V.

The simulation was then performed for a pixel of 110 μm pitch and no JTE. Again the breakdown voltage of this device was < 400V but the effective fill factor was significantly increased. At a bias voltage of 350 V the gain uniformity across the pixel is shown in figure 6.17. The gain for this device seems to be larger nearer the edge of the pixel. This is due to a slight increase in the field near the edge of the pixel, where this fields causes the early breakdown of the device. This could also be due to a finer mesh in this region. A finer mesh is required to handle the large change in field laterally at the edge of the pixel, partly due to the absence of the JTE. The effective fill factor for this design of pixel is
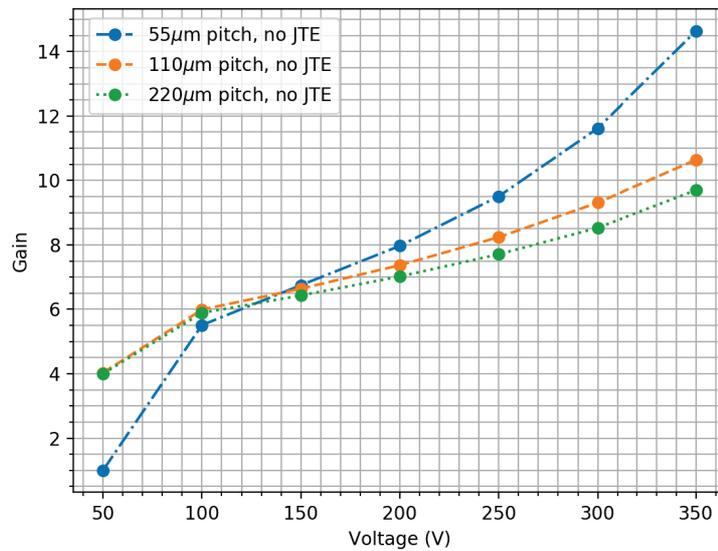
Figure 6.16: Gain as a function of voltage for devices without a JTE.

60.6% and 75.1% for at a peak height of 90% and 50% respectively, where the peak height was determined by the gain value at the centre of the pixel.

Looking at a pixel pitch of 220 μm and no JTE we see a significant increase in the effective fill factor and a reasonable breakdown voltage of $\approx$ 400 V. The gain uniformity is shown alongside the other pixel sizes in figure 6.17. A similar shape is seen as with the 110 μm pixel where the peak gain is near the edge of the device, again most likely as a result of a finer mesh in this region resulting in a slight variation of the electric field. The effective fill factor for this design of pixel is 78.4% and 88.7% for at a peak height of 90% and 50% respectively, where the peak height was determined by the gain value at the centre of the pixel.

It should be noted that a significant impact of the JTE is the reduction and total loss of gain in small pixel LGAD's. Removing the JTE width produces devices with gain across the entire pixel for three values of pitch. Gain is somewhat reduced at the edges but is still 4 at the edge of each device.

The fill factor is still quite low for the 55 μm pixels, especially for applications requiring high time resolution. As the signals generated will vary massively from one signal to the next, which limits the time resolution as described in chapter 5. Another suggestion is to modify the periphery of the pixel to reduce the non-gain region between pixels. At the moment there is a gap of 19 μm, which is essential for isolating the pixels, especially when JTE's are present. Without JTE's it may be possible to reduce the size of this gap to around 6 μm, assuming a reduced overlap to 5 μm. However for this to work a p–spray may not be effective due to the small gap, any diffusion of the junction implant into this gap may
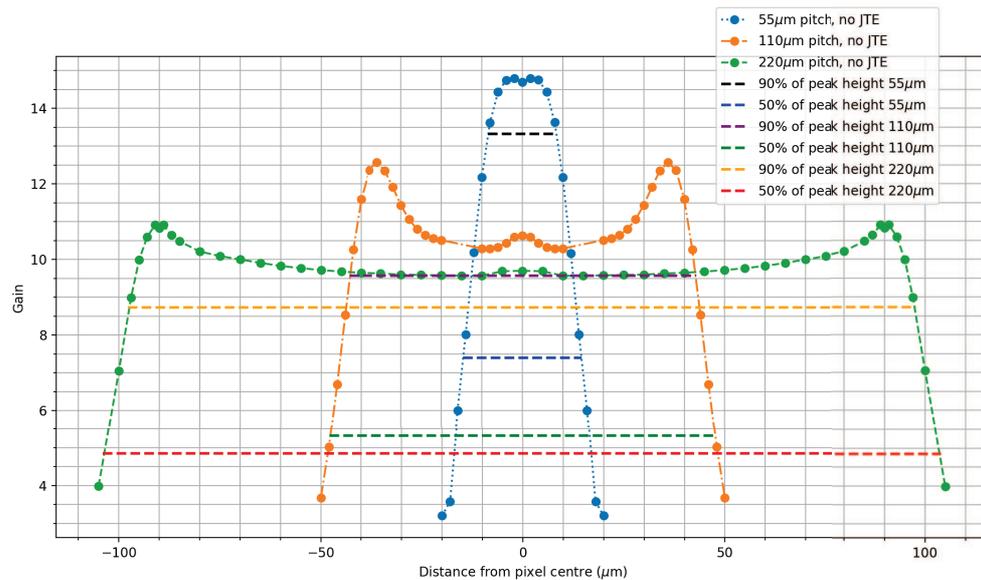
Figure 6.17: Gain as function of charge injection location relative to the centre of the pixel for devices without a JTE.

prevent the pixels from being isolated from one another. A p–stop would be required, which at the moment has not been used in our process, therefore no simulation of this design has been performed.

Another solution has been suggested by a group working with FBK in Torino [85]. This is the introduction of a trench at the junction between two pixels of the order of 1 μm width and around 10 μm deep. This technology has already been used to isolate SiPM pixel with a pitch of 5μm. This has been calculated to provide a non–gain border of 2–3 μm. For a pixel pitch of 55 μm this turns out to provide a fill factor of $> 84\%$ from initial calculations. Increasing the pixel pitch to 110 μm gives a fill factor of $> 90\%$. These results have been taken from a talk at an RD50 meeting proposing a project on High Density LGAD (HD–LGAD) [85].

To test the results from the simulated pixels with various pixel pitch a fabrication run was performed which included many Timepix3 compatible arrays. Devices of all pitches mentioned in this chapter have been bump bonded to Timepix3 ASICs and are awaiting tests. As of writing this thesis no data has been collected thus far for these devices. However these will be tested on completion of this thesis. The devices will undergo mea–surements at a Beam test in Diamond Light Source to evaluate the effective fill factor for all sizes with a JTE width of 10 μm. The 55 μm pitch Timepix arrays do not include JTE's as the devices with JTE were shown to have an effective fill factor of zero. Presented in this chapter was a pixel of a 55 μm pitch which produced gain and was the basis of this design.

6.4   Conclusion

This chapter has introduced the concept of highly segmented LGAD detector for sub-threshold detection with good time resolution using a hybrid pixel detector known as the Timepix ASIC. The evolution of the ASIC has been described to produce an ASIC with a time resolution of 1.56 ns, the Timepix3. The ASIC is segmented with a pixel pitch of 55 μm in an array of 256x256. Pixellated LGAD detectors have been developed at Micron Semiconductor that match this architecture and have been bump bonded to the Timepix3 ASIC.

Design and development of the pixel has been performed using TCAD tools to evaluate the gain uniformity and effective fill factor. Traditional designs of LGAD's include a Junction Termination Extension (JTE) that has a major impact on the fill factor for small pixels. The effect of reducing the size of the JTE was discussed in terms of the effective fill factor. Three pixel pitches were discussed as possible candidates for the final ASIC design with varying periphery structures. It was shown that any width of JTE effectively causes the 55 μm pitch pixels to have a fill factor of 0%, due to the charge collection via the JTE with charge avoiding the gain region. The fill factor was shown to increase for both the 110 μm and 220 μm pixels when a smaller JTE width is used. However the effective fill factor for the 110 μm pixel was limited to 18%. The effective fill factor achieved for the 220 μm pixel was 47%.

The larger pixels provide better gain uniformity but reduce the spatial resolution. Possible solutions to the pixel periphery design were briefly discussed to help increase the effective fill factor of the 55 μm pixels. Removing the JTE structures provided a 55 μm pixel with an effective fill factor of 28.2% but with a reduced breakdown voltage of < 400V. However increasing the pixel size to 220 μm provided a device with an effective fill factor of 88.7%. A Novel idea to include a trench to isolate neighbouring pixels which would allow a 55 μm pixel to have an effective fill factor of > 84% has been introduced.

Chapter 7

Conclusion & Future Work

7.1 Conclusion

In this thesis the simulation, fabrication and characterisation of LGAD detectors has been presented in detail. The basic principles of semiconductors as detectors was discussed in chapter 2 with further details on the requirements of detectors to achieve impact ionisation. This chapter also included the basic fabrication techniques used in commercial semicon-ductor fabrication facilities with details of the fabrication process variations used to achieve the required doping profiles for an LGAD. Further details of the fabrication process are described as part of chapter 3, where the technology was developed using the "Synop-sis" TCAD simulation package. Results from the characterisation of fabricated sensors are shown in chapters 4 and 5. Chapter 6 focused on the development of this technology into highly segmented pixel arrays where the challenges are described.

The simulations during this thesis proved vital in understanding the fabrication process required to produce a device with low gain. The process simulation used to calculate the eventual doping profiles of the detector was described using the physics models and meshing strategy's available in Sentaurus TCAD. It was important to understand the effect of altering the various processing parameters on the eventual doping profiles, such as implant doses, implant energies and thermal processing steps. The resultant doping profiles from simulation have been validated using the SIMS method for fabricated devices. A comparison has been made on the shape of these doping profiles as well as the $Q_{effective}$.

To calculate the electrical performance of these devices a full device simulation was performed to evaluate the electric field profile in the device and the resultant gain. The gain was shown to vary with voltage, as the peak electric field in the gain region grows. The $Q_{effective}$ was shown to be related to the gain with a critical value of $\approx 1.6 \times 10^{12}$ cm^{-2}. Periphery structures were evaluated to increase the breakdown voltage of these devices as high fields are present at the detector edge. The Junction Termination Extension (JTE), a deep n–type implant, was shown to improve the performance considerably and this was

therefore used in the fabricated designs.

The devices fabricated were summarised with details of specific processing parameters used that are critical for this type of detector. The devices were characterised using a variety of techniques to establish their electrical performance. IV measurements were used to determine the breakdown voltage of the devices, as well as the leakage current. With changes in the optimum peripheral region resulting in a varied set of results. The optimum choice of detector periphery was the combination of a JTE structure and overlapping field plate. Devices with the best electrical performance were then tested for gain using a variety of methods. The main method used was the Transient Current Technique (TCT), which induces charge in the detector by using a focused laser beam. The charge collected for a device without a gain region was used as a reference to calculate the gain. Alpha TCT was also used in a similar set-up but induced charge in the detector by alpha decay. The same analysis approach could be used by both TCT methods. Finally a ^{90}Sr source was used to calculate gain as the MPV would increase by the gain factor compared to a non-gain device. These measurements took much longer than the TCT methods. The gain for each run was presented with the final run producing a gain of 5-20 at voltages up to 500 V.

Initial measurements using an X-ray source established the minimum detectable energy of 17.5 keV for a 200 μm thick, 5 mm LGAD detector with a gain of 10 at 200 V.

Measurements for gain of pad sizes down to 220 μm were made, where the gain at this size showed a decrease relative to similar detectors of pad sizes > 0.5 mm.

Following successful fabrication of LGAD detectors with the required gain a dedicated timing set-up was produced at the University of Glasgow. The timing set-up was described where specific low noise readout boards were used to obtain the best achievable time resolution. Detectors fabricated at Micron Semiconductor were evaluated for their timing performance. These were detectors taken from run 5, which are 250 μm thick with an active area of 1x1 mm^2. These devices have a gain of $\approx$ 5 at voltages up to 250 V with a leakage current < 10 nA/mm^2. Measurements for timing resolution at a range of temperatures and voltages yielded a minimum time resolution of 120 ps. With comparison to other fabricated sensors from a variety of manufacturers this result seems reasonable. The time resolution did not have a voltage dependence as the gain was stable over the voltage range used. Also the jitter was calculated to be 124 ps which sets a limit on the achievable time resolution.

Fabrication is complete on a new batch of sensors which have a thickness of 50 μm, which are to be tested following the completion of this thesis. The results obtained should match those available by other vendors with a time resolution of less than 50 ps. Arrays of these devices have also been fabricated and measurements will be made to evaluate the performance of pixellated devices in comparison to pad detectors.

The final chapter introduced the concept of highly segmented LGAD detectors for sub-

threshold detection with good time resolution using a hybrid pixel detector known as the Timepix3 chip. The Timepix3 chip is part of well known Medipix family of detectors. The chip is segmented with a pixel pitch of 55 μm in an array of 256x256. The chip is bump bonded to a sensor with a matching array pattern. The sensor can have a variety of pixel sizes which can be connected to all or a fraction of the bump bonds. Typically these have pixel pitches of 55, 110 and 220 μm.

LGAD detectors of this pitch were studied for gain uniformity and effective fill factor using Sentaurus TCAD tools. It was shown that the Junction Termination Extension (JTE) has a major impact on the fill factor for small pixels. The effect of reducing the width of the JTE was discussed with the results showing increased fill factor with a reduction in JTE width. However all 55 μm pixels with a JTE showed no gain across the entire pixel. The maximum fill factor for the 110 μm and 220 μm pixels was limited to 18% and 47% respectively at 500 V. To increase the fill factor of the 55 μm pixels the JTE structures were removed and replaced with an overlapping junction design. This showed a decrease in breakdown voltage but at 350 V an effective fill factor of 28.2% was achieved. A dramatic improvement, also shown for the 110 μm pixel at 350 V with an effective fill factor of 75.1%. Further improvements need to be made and one such novel idea is to include a trench to isolate neighbouring pixels which would allow a 55 μm pixel to have an effective fill factor of > 84% as shown in the simulations.

7.2   Future Work

Thin LGAD detectors have been fabricated and will be tested for their timing resolution. Assuming comparable time resolutions with already established vendor detectors, detectors from Micron Semiconductor will undergo a radiation campaign to study the effect of radiation on the gain and resultant time resolution. This work must be prompt in order to establish the technology for use in the ATLAS timing detector at the HL–LHC.

Arrays of LGAD detectors will be tested for gain uniformity and time resolution with a variety of array and pixel sizes. A technique known as edge-TCT will be implemented into the existing TCT set–up to allow electric field profiles to be studied across the full width of a pixel as well as the pixel to pixel gap. PCB's have already been produced which will allow these measurements to occur quickly.

Work is already underway to produce Hybrid Pixel detectors which have a sensor with an internal gain. Pixellated LGAD detectors with pitches of 55, 110 and 220 μm have been bump bonded to Timepix3 readout chips. These will be fully characterised using the same methods as described in this work. A proposal is underway to get beam time at the Diamond Light Source in order to fully evaluate the effective fill factor and gain uniformity of these detectors. This will be work to compare with the simulation studies discussed in

this thesis. The Diamond Light Source uses a micro-focus X-ray beam which can be used to probe the charge collection to sub pixel resolutions, enabling the calculation of effective fill factor. A range of X-ray energies can be used which will determine the minimum detectable energy by this technology. The results should provide some understanding on the limits of the technology with the hope of detecting sub-threshold signals.

The use of the Timepix chip with LGAD technology is of great interest to the LHCb collaboration who wish to make a highly pixellated timing detector for their experiment. As we will be the first group to possess and test this technology we have a great advantage over other vendors. We will be able to provide great insight and knowledge on the capabilities of this technology, where design changes will most likely be required to fulfil the requirements.

Appendix A

Synopsis TCAD SProcess Models

All models used with the SProcess module are described where the details have been taken directly from the Synopsys Sentaurus TCAD manual [12].

A.1 Crystal–Trim Physical Model for Silicon

Crystal-TRIM simulation is based on the binary collision approximation (BCA), which represents the motion of ions in the target material as a set of binary collisions with the target atoms [46]. BCA is valid in a wide range of projectile energies, from approximately 100 eV to many MeV. It can, therefore, be employed over the whole range of energies of interest for ion implantation. For energies below approximately 100 eV, collective interactions may play an increasingly important role and BCA may become invalid. Nevertheless, the applications of Crystal-TRIM to ultra low–energy implants lead to results that are still sufficiently good compared with experimental data. At each collision, the projectile loses a part of its energy due to elastic nuclear scattering at target atoms and inelastic electronic iterations. The particles are assumed to come to rest if their energy is in the order of 15 eV.

A.1.1 Nuclear Collisions and Collision Cascades

Nuclear scattering is treated by classical mechanics using a Coulomb–screened pair potential (ZBL potential [86]). If the energy transfer to the target atoms exceeds the so–called displacement threshold (approximately 15 eV for silicon), the target atom can leave its site and become displaced (primary recoil). By default, only the trajectories of implanted ions are simulated. The number of vacancies and displaced target atoms produced at each collision is calculated approximately using the modified Kinchin–Pease formula. A full cascade–type of simulation is performed if the keyword cascades is used. The trajectories of energetic recoils are calculated in the same way as for the original ions. A primary

recoil with sufficiently high initial energy can generate more recoils (collision cascade). While both methods yield correct range profiles, only the full–cascade simulation produces physically correct profiles of vacancies and displaced atoms. However, a full–cascade simulation requires more computational time. In single–crystalline silicon, vacancies and recoils are often identified with the vacancies and interstitials responsible for transient–enhanced diffusion (TED) of dopants. The choice of a diffusion model determines whether the full–cascade mode of Crystal–TRIM must be applied.

A.1.2   Electronic Stopping

Electronic energy loss of the projectile is treated using semiempirical models. For crystalline target materials, the loss depends on the local electronic density in the environment of the projectile. Therefore, the use of a local approach is particularly important for investigations of channeling effects in single–crystalline substrates. A simplified local approach, the so-called modified Oen–Robinson formula [87], is used. The parameter *CEX1* describes the variation of the electron density for a projectile moving in the <110> direction of the crystal, while *CEX2* does the same for any other direction. The value for *CEX1* and *CEX2* are set in the parameter database and can be changed by using: *pdbSet <material> <dopant> CEX1 <n>* The value for *CEX1* should be close to 1 or at least within the range of 0.5 and 3. The default value for *CEX2* is 2.

A.2   Diffusion Models

During the fabrication process, dopants are introduced into the substrate with different concentration profiles. As processing proceeds through various thermal annealing cycles, the dopants diffuse and redistribute through the structure. The following effects contribute to dopant redistribution and can be modeled by Sentaurus Process:

- Dopant (de)activation

- Dopant–defect interaction,

- Chemical reactions at interfaces and in bulk materials,

- Material flow

- Moving material interfaces

- Internal electric fields

Table A.1: Solution names

Symbol	Boron	Arsenic	Phosphorus	Antimony	Indium
C_A	Boron	Arsenic	Phosphorus	Antimony	Indium
C_{AI}	BoronInt	ArsenicInt	PhosphorusInt	AntimonyInt	IndiumInt
C_{AV}	BoronVac	ArsenicVac	PhosphorusVac	AntimonyVac	IndiumVac
C_A^+	BActive	AsActive	PActive	SbActive	InActive

Table A.2: Point–defect names

Symbol	Interstitial	Vacancy
C_X	Int	Vac
C_X^*	EqInt	EqVac
$C_{X^0}^*$	IntNeutralStar	VacNeutralStar
C_{X^0}	IntNeutral	VacNeutral

A.2.1   Transport Models

Transport models compute the particle flux of dopants and are the core diffusion models solved by Sentaurus Process. In addition to particle flux, pairing reactions can be computed depending on the transport model selection. Transport models are usually used with one or more clustering or activation models available. The reaction or clustering models will not modify the dopant flux, but will compute terms to be added to from

$$\frac{\partial A^c}{\partial t} = -\nabla \cdot J_{A^e} + R_{A^2}^{trans} - R_{A^c}^{clus}$$

. The models used are described in detail below.

A.2.2   ChargedReact

The *ChargedReact* diffusion model is the most general transport model in Sentaurus Process. The model has an immobile substitutional dopant and up to two mobile charged dopant–defect pair species. Mobile charged point defects are also included in the model.

   The following reactions are considered:

$$A^z + I^i \leftrightarrow AI^{(z+j)} - (i-j)e \tag{A.1}$$

$$A^z + V^i \leftrightarrow AV^{(z+j)} - (i-j)e \tag{A.2}$$

$$I^i + V^j \leftrightarrow -(i+j)e \tag{A.3}$$

$$AI^{(z+i)} + V^j \leftrightarrow A^z - (i+j)e \tag{A.4}$$

$$AV^{(z+i)} + I^j \leftrightarrow A^z - (i+j)e \tag{A.5}$$

The differential equations that are solved in this model are:

$$\frac{\partial C_A}{\partial t} = -R_{AI} - R_{AV} + R_{AI,V} + R_{AV,I} - R_A^{c\ln s} \tag{A.6}$$

$$\frac{\partial C_{AI}}{\partial t} = -\nabla \cdot J_{AI} + R_{AI} - R_{AI,V} - R_{AI}^{clus} \tag{A.7}$$

$$\frac{\partial C_{AV}}{\partial t} = -\nabla \cdot J_{AV} + R_{AV} - R_{AV,I} - R_{AV}^{clus} \tag{A.8}$$

$$\frac{\partial C_I}{\partial t} = -\nabla \bullet J_I - R_{IV} - R_{AI} - R_{AV,I} - R_I^{clus} \tag{A.9}$$

$$\frac{\partial C_V}{\partial t} = -\nabla \cdot \mathbf{J}_V - R_{IV} - R_{AV} - R_{AI,V} - R_V^{clus} \tag{A.10}$$

where C_A is the concentration of substitutional (and assumed to be immobile) dopants, C_X is the concentration of 'free' defects of type X (either interstitials or vacancies), in other words, those defects not in clusters or pairs. The reaction rates of the different species (R) are defined later in this section.

Next, the flux of the mobile defect pair is considered. Working with

$$J_{A^c} = -d_{A^c} \left(\frac{n}{n_i} \right)^{-c} \nabla \left(A^c \left(\frac{n}{n_i} \right)^c \right) \tag{A.11}$$

for the charged pairs, the equation will be written in terms of the total concentration of pairs.

It is expected that the dopant–defect pairing reaction is in equilibrium, therefore, a set of constants for this pairing is defined:

$$C_{AX^{z+c}} = k_{AX^c} C_{A^z} C_{X^c} \tag{A.12}$$

where X is either I or V, z is the charge of the dopant A, and k_{AX^c} is the pairing coefficient for the pair AX, and is given by:

$$k_{AX^c} = k_{AX^c}^0 \exp \left(\frac{-k_{AX^c}^E}{k_B T} \right) \tag{A.13}$$

To set k_{AX^c}, use:

pdbSet <material> <dopant> <defect> ChargePair <c> <n> where <material> is a material name (see), <dopant> is one of the existing Sentaurus Process dopants, <defect> is either Interstitial or Vacancy, <c> is the charge state, and <n> is a Tcl expression that returns a number – it can be simply a number.

One commonly used Tcl procedure for setting parameters is *Arrhenius*. This procedure takes a prefactor and an energy as arguments and returns $prefactor \cdot exp\frac{-energy}{k_B T}$.

The flux of the pairs is computed from Eq. A.11:

$$J_{AX} = -\sum_c J_{AX^c} = -\sum_c D_{AX^c} \left(\frac{n}{n_i}\right)^{-c-z} \nabla \left(\frac{C_{AX}}{C_{X0}^* \sum_q k_{AX^q} k_X \left(\frac{n}{n_i}\right)^{-q}} \left(\frac{n}{n_i}\right)^z\right) \quad (A.14)$$

where C_{AX} is the total concentration of pairs that is the sum of the concentrations of pairs at every charge state and D_{AX^c} is an effective diffusivity of dopant point–defect pairs at charge state c and is related to the self diffusivity d_{AX^c} by:

$$D_{AX^c} = C_{X0}^* k_{AX^c} k_{X^c} d_{AX^c} = D_{AX^*}^0 \exp\left(\frac{-D_{AX^c}^E}{k_B T}\right) \quad (A.15)$$

where C_{X0}^* is the equilibrium concentration of the neutral defects and is related to the total equilibrium intrinsic concentration of defect X by:

$$C_{X0}^* = \frac{C_{X(intrinsic)}^*}{\sum_c k_{X^c}} \quad (A.16)$$

The quantities $C_{I(intrinsic)}^*$ and $C_{V(intrinsic)}^*$, which by default follow an Arrhenius law, can be changed by using the command:

pdbSet <material> <defect> Cstar <n> To set D_{AX^c}, use: pdbSet <material> <dopant> <defect> D <c> <n> A set of equilibrium–charging constants,k_{X^c}, for defect X is defined:

$$C_{X^c} = k_{X^L} C_{X0} \left(\frac{n}{n_i}\right)^{-c} \quad (A.17)$$

where X is either I or V, and k_{X^c} is the charging coefficient for the defect X and is given by:

$$k_{X^c} = k_{X^c}^0 \exp\left(\frac{-k_{X^c}^E}{k_B T}\right) \quad (A.18)$$

To set k_{X^c}, use:

pdbSet <material> <defect> ChargeStates <c> <n>

Similar to the pairs, the defect fluxes are computed from Eq. A.11:

$$J_X = -\sum_c J_{X^c} = -\frac{\sum_c k_{X^c} D_{X^c} \left(\frac{n}{n_i}\right)^{-c} C_X^*}{\sum_q k_{X^q} \left(\frac{n}{n_i}\right)^{-q}} \nabla \left(\frac{C_X}{C_X^*}\right) \tag{A.19}$$

where C_X is the total concentration of defects that is the sum of the concentrations of defect X at every charge state and D_{X^c} is the diffusivity of the defect X of charge state and is given by:

$$D_{X^c} = d_{X^c} = D_{X^c}^0 \exp\left(\frac{-D_X^E}{k_B T}\right) \tag{A.20}$$

To set D_{X^c}, use:

pdbSet <material> <defect> D <c> [Arrhenius <prefactor> <energy>] Now, the reaction rates can be written by considering Eqs. A.1 to A.5 and the general formula for the rate of all combinations of charge states:

$$A^a + B^b \leftrightarrow AB^c + (c-a-b)e R_{A^a,B^b,c} = k_{A,B^b,c}^f \left(C_{A^a} C_{B^b} - k_{A,B^b,c}^r C_{AB^c} \left(\frac{n}{n_i}\right)^{(c-a-b)}\right) \tag{A.21}$$

Therefore, summing all possible charge states gives:

$$R_{AX} \equiv -K_{AX}^{-f} \left(C_A C_X - \frac{C_{AX}}{\bar{K}_{AX}^r}\right) \tag{A.22}$$

$$R_{AI,V} \equiv \bar{K}_{AI,V}^f \left(C_{AI} C_V - \bar{K}_{AX}^r C_I^* C_V^* C_A\right) \tag{A.23}$$

$$R_{AV,I} \equiv \bar{K}_{AV,I}^f \left(C_{AV} C_I - \bar{K}_{AX}^r C_I^* C_V^* C_A\right) \tag{A.24}$$

where:

$$\bar{K}_{AX}^f \equiv \frac{\sum_i K_{f_x} KO_i k_X \left(\frac{n}{n_i}\right)^{-i}}{\sum_c k_{X^c} \left(\frac{n}{n_i}\right)^{-c}} \tag{A.25}$$

$$\bar{K}_{AX}^r \equiv \frac{\sum_i k_{AX^i} k_{X^i} \left(\frac{n}{n_i}\right)^{-i}}{\sum_c k_{X^c} \left(\frac{n}{n_i}\right)^{-c}} \tag{A.26}$$

$$\bar{K}_{AI,V}^f \equiv \frac{\sum_i \sum_j K_{f_{AI^i,Vj}} FT k_{AI^i} k_{Vj} \left(\frac{n}{n_i}\right)^{-(i+j)}}{\sum_c k_{AI^c} \left(\frac{n}{n_i}\right)^{-c} \sum_z k_{V^z} \left(\frac{n}{n_i}\right)^{-z}} \tag{A.27}$$

$$\bar{K}^f_{AV,I} \equiv \frac{\sum_i \sum_j K_{f,j,f} F T^k_{AV} k_N \left(\frac{n}{n_i}\right)^{-(i+j)}}{\sum_c k_{AV} \left(\frac{n}{n_i}\right)^{-c} \sum_z k_T \left(\frac{n}{n_i}\right)^{-z}} \tag{A.28}$$

where $\bar{K}^f_{AX}$ is the forward reaction rate for the kick–out mechanism, and X is either in–terstitial or vacancy, $\bar{K}^f_{AI,V}$ and $\bar{K}^{-f}_{AV,I}$ are forward reaction rates for the Frank–Turnbull mechanism and $\bar{K}^{-r}_{AX}$ is the equilibrium constant. The forward ($K_{f_X KO_{i,j}}$) kick–out reaction rates can be set by using the following commands:

pdbSet <material> <dopant> <defect> kfKickOut <c> <n> where c is the charge state.

For example:

pdbSet Silicon Boron Interstitial kfKickOut 1 [expr 4.0*3.14159*([Arrhenius 0.1 0.2]+[Arrhenius 0.3 0.4])] sets the forward reaction rate for $c = 1$ to $4\pi \left(0.1 \exp\left(\frac{-0.2}{kT}\right) + 0.3 \exp\left(\frac{-0.4}{kT}\right)\right)$. Similarly, the forward $\left(\bar{K}^f_{AI,V}, \bar{K}^f_{AV,I}\right)$ Frank–Turnbull reaction rates can be defined using the commands: *pdbSet <material> <dopant> <defect> kfFTM <i,j> <n>* The I–V re-combination reaction is given as:

$$R_{IV} = \bar{K}_{IV} \left(C_I C_V - C^*_I C^*_V\right) \tag{A.29}$$

where:

$$bar K_{IV} = \frac{C^*_{I(intrinsic)} C^*_{V(intrinsic)}}{C^*_I C^*_V \sum_z k_{I^z} \sum_z k_{V^z}} \sum_i \sum_j K_{I^i V^j} k_{I^i} k_{V^j} \left(\frac{n}{n_i}\right)^{-(i+j)} \tag{A.30}$$

The superscript '*' refers to the equilibrium concentration, and the subscripts I and V are for the interstitials and vacancies, respectively. The subscripts z,i,j are the charge states of the defects. $K_{I^i V^j}$ is the bulk recombination rate for interstitials and vacancies at the charge stated i and j, respectively. The bulk recombination rate $K_{I^i V^j}$ for each charged point defect can be set using the command:

pdbSet <mater> <defect> KbulkChargeStates <i,j> <n>

The equilibrium concentration of the unpaired point defect can be calculated by:

$$C^*_X = C^*_{X(\text{ intrinsic })} \frac{\sum_c k_{X^c} \left(\frac{n}{n_i}\right)^{-c}}{\sum_c k^s_{X^c}} \tag{A.31}$$

where $k^s_{X^c}$ is the scaled charging coefficient for the defect X and can be set by using:

pdbSet <material> <defect> ChargeStatesScale <c> <n>

$k^s_{X^c}$ is set to k_{X^c} as a default.

A.2.3   React

The *React* model is similar to the *ChargedReact* model, except that the reaction rates are not charge state–dependent and the electron concentration is computed directly from the net doping concentration. In addition, the Frank–Turnbull mechanism is not considered. The reactions considered are:

$$A + I \leftrightarrow AI \tag{A.32}$$

$$A + V \leftrightarrow AV \tag{A.33}$$

where A is the dopant, I is the interstitial, and V is the vacancy. The following set of differential equations represents the model:

$$\frac{\partial C_A}{\partial t} = -R_{AI} - R_{AV} - R_A^{clus} \tag{A.34}$$

$$\frac{\partial C_{AX}}{\partial t} = -\nabla \bullet J_{AX} + R_{AX} - R_{AX}^{clus} \tag{A.35}$$

$$\frac{\partial C_X}{\partial t} = -\nabla \cdot J_X - R_{IV} - R_{AX} - R_X^{clus} \tag{A.36}$$

where C_A is the concentration of substitutional (and assumed to be immobile) dopant and C_X is the concentration of 'free' defects of type X (either interstitials or vacancies), that is, those defects not in clusters or pairs.

Next, the flux of the mobile defect pair is considered. Working with Eq. A.11 for the charged pairs, the equation will be written in terms of the total concentration of pairs. In addition, this model assumes the charging constants for the pairs are the same as the charging constants for the defects. It is expected that the charging reaction is in equilibrium, therefore, a set of constants for this charging is defined:

$$C_{AX^{\varepsilon+c}} = k_{AX^c} C_{A^z} C_{X0} \left(\frac{n}{n_i}\right)^{-c} \tag{A.37}$$

However, in this model, the reaction rates are assumed to be independent of the charge state, so the pair charging constants are only needed for the flux of the pairs and are absorbed into the diffusivity of the pairs in this way:

$$J_{AX} = \frac{-\sum_c \left(D_{AX^c} \left(\frac{n}{n_i}\right)^{-c-z} \right)}{B_{AX}} \nabla \frac{C_{AX}}{C_X^*} \left(\frac{n}{n_i}\right)^z \tag{A.38}$$

where z is the charge state of dopant A, X is either interstitial or vacancy, and D_{AX^c} is the effective diffusivity of dopant point-defect pair at charge state c and is related to the bare

diffusivity, d_{AXc} by:

$$D_{AXc} = C_{X0}^* B_{AX} k_{Xc} d_{AXc} = D_{AXc}^0 \exp\left(\frac{-D_{AXc}^E}{k_B T}\right) \tag{A.39}$$

where k_{Xc} is a set of equilibrium charging constants for defect X defined by Eq. A.17 and Eq. A.18, and is the equilibrium concentration of the neutral defects defined by Eq. A.16.

To set D_{AXc}, use:

pdbSet <material> <dopant> <defect> D <c> <n>
where:

- <material> is a material name (see Material Specification on page 49).

- <dopant> is one of the existing Sentaurus Process dopants.

- <defect> is either Interstitial or Vacancy.

- <c> is the charge state.

- <n> is a Tcl expression that returns a number; it can be simply a number.

One commonly used Tcl procedure for setting parameters is *Arrhenius*. This procedure takes a prefactor and an energy as arguments and returns *prefactor* $\cdot exp\frac{-energy}{k_B T}$.

You can modify the entire array with the command (for example, arsenic–vacancy pairs):

pdbSet Si Arsenic Vac D 0 [Arrhenius 0.0 3.45] –1 [Arrhenius 12.8 4.05]

The defect flux is the same as the ChargedReact model and is given by Eq. A.19. The reaction can be written as:

$$R_{AX} \equiv K_{AXr}\left(C_A^+ C_X - \frac{C_{AX}}{B_{AX}}\right) \tag{A.40}$$

where X is either interstitial or vacancy, B_{AX} is the binding coefficient of defect X and dopant A, K_{AXR} is the rate constant for the chemical reaction, and C_A^+ is the active portion of C_A . The binding term between the defect and dopant also follows the Arrhenius law:

$$B_{AX} = B_{AX0} \exp\left(\frac{-B_{AXE}}{kT}\right) \tag{A.41}$$

The term can be changed with the command:

pdbSet <material> <dopant> <defect> Binding <n>
The chemical reaction term is expressed with:

$$K_r = K_{r0} \exp\left(\frac{-K_{rE}}{kT}\right) \tag{A.42}$$

and can be modified by using the command:

   pdbSet <material> <dopant> <defect> Krate <n>

   The defect recombination rate R_{IV} is the same as in the *ChargedReact* model and is given by Eq. A.29.

Appendix B

Synopsis TCAD SDevice Models

All physics models used with the SDevice module are described where the details have been taken directly from the Synopsys Sentaurus TCAD manual [12].

B.1   Mobility Models

This section presents the essential mobility models used in the SDevice simulations.

B.1.1   DopingDep

For any device material which has been doped a mobility model must be selected which takes this into account. This is done by choosing the doping-dependant mobility Model (DopingDep). Variations of this model can be used. However, the standard *Masetti Model* was used which was proposed by Masetti et al [88].

$$
\mu_{\text{dop}} = \mu_{\min 1} \exp\left(-\frac{P_{\text{c}}}{N_{\text{A},0} + N_{\text{D},0}}\right) + \frac{\mu_{\text{const}} - \mu_{\min 2}}{1 + \left(\left(N_{\text{A},0} + N_{\text{D},0}\right)/C_{\text{r}}\right)^{\alpha}} - \frac{\mu_1}{1 + \left(C_{\text{s}}/\left(N_{\text{A},0} + N_{\text{D},0}\right)\right)^{\beta}}
\tag{B.1}
$$

The reference mobilities μ_{min1}, μ_{min2}, and μ_1, the reference doping concentrations P_c, C_r, and C_s, and the exponents α and β are accessible in the parameter set DopingDep.

The corresponding values for silicon are given in table B.1.

B.1.2   HighFieldSaturation

In high electric fields, the carrier drift velocity is no longer proportional to the electric field, instead, the velocity saturates to a finite speed v_{sat}. Sentaurus Device supports different models for the description of this effect. The high-field saturation models comprise three submodels: the actual mobility model, the velocity saturation model, and the driving force model. With a some restrictions, these models can be freely combined. The actual mobility

Table B.1: Masetti model: Default coefficients

Symbol	Parameter name	Electrons	Holes	Unit
μ_{min1}	mumin1	52.2	44.9	cm^2/Vs
μ_{min2}	mumin2	52.2	0	cm^2/Vs
μ_1	mu1	43.4	29.0	cm^2/Vs
P_c	Pc	0	9.23×10^{16}	cm^{-3}
C_r	Cr	9.68×10^{16}	2.23×10^{17}	cm^{-3}
C_s	Cs	3.43×10^{20}	6.10×10^{20}	cm^{-3}
α	alpha	0.680	0.719	1
β	beta	2.0	2.0	1

model is selected by flags to *eHighFieldSaturation* or *hHighFieldSaturation*. The default Canali model is used.

The Canali model [50] originates from the Caughey–Thomas formula [51], but has temperature–dependent parameters, which were fitted up to 430 K by Canali et al. [50]:

$$\mu(F) = \frac{(\alpha + 1)\mu_{\text{low}}}{\alpha + \left[1 + \left(\frac{(\alpha+1)\mu_{\text{low}}F_{\text{hfs}}}{v_{\text{sat}}}\right)^{\beta}\right]^{1/\beta}} \tag{B.2}$$

where μ_{low} denotes the low–field mobility. The exponent β is temperature dependent according to:

$$\beta = \beta_0 \left(\frac{T}{300\text{K}}\right)^{\beta_{\text{exp}}} \tag{B.3}$$

The default velocity saturation model for silicon is used and is part of the Canali model, given by:

$$v_{\text{sat}} = v_{\text{sat},0}\left(\frac{300\text{K}}{T}\right)^{v_{\text{atatep}}} \tag{B.4}$$

The driving force model used is given by the flat *GradQuasiFermi*, given by:

$$F_{\text{hfs},n} = |\nabla \Phi_n| \tag{B.5}$$

B.2   Recombination Models

This section presents the essential recombination models used in the SDevice simulations.

B.2.1   Avalanche Generation

Electron–hole pair production due to avalanche generation (impact ionization) requires a certain threshold field strength and the possibility of acceleration, that is, wide space

Symbol	Parameter name	Electrons	Holes	Unit
a	a	0.426	0.243	V^{-1}
b	b	4.81×10^5	6.53×10^5	V/cm
c	c	3.05×10^{-4}	5.35×10^{-4}	K^{-1}
d	d	6.86×10^{-4}	5.67×10^{-4}	K^{-1}
γ	gamma	1	1	1
δ	delta	2	2	1

Table B.2: Coefficients for the Okuto model.

charge regions. If the width of a space charge region is greater than the mean free path between two ionizing impacts, charge multiplication occurs, which can cause electrical breakdown. The reciprocal of the mean free path is called the ionization coefficient α. With these coefficients for electrons and holes, the generation rate can be expressed as:

$$G_{\text{ii}} = \alpha_n n v_n + \alpha_p p v_p \tag{B.6}$$

There are multiple models in SDevice to choose. For this work the Okuto–Crowell model was chosen. Okuto and Crowell [52] suggested the empirical model:

$$\alpha(F_{\text{ava}}) = a \cdot \left(1 + c(T - T_0)\right) F_{\text{ava}}^{\gamma} \exp\left[-\left(\frac{b[1 + d(T - T_0)]}{F_{\text{ava}}}\right)^{\delta}\right], \tag{B.7}$$

where F_{ava} is the driving force for impact ionization, T is the temperature and which by default uses the parameters for silicon and $T_0 = 300$ K. These values are applicable to the range of electric field $10^5 \ V cm^{-1}$ to $10^6 \ V cm^{-1}$. The parameters used are shown in table B.2.

B.2.2   Band to Band tunneling

Phonon–assisted band-to-band tunneling cannot be neglected in steep p–n junctions (with a doping level of $1 \times 10^{19} \ cm^{-3}$ or more on both sides) or in high normal electric fields. Band-to-band tunneling is modeled using the expression given by Schenk et al [53]:

$$R_{\text{net}}^{\text{bb}} = A F^{7/2} \frac{\tilde{n}\tilde{p} - n_{\text{l,eff}}^2}{(\tilde{n} + n_{\text{ieff}})(\tilde{p} + n_{\text{ieff}})} \left[\frac{\left(F_C^{\mp}\right)^{-3/2} \exp\left(-\frac{F_C^{\mp}}{F}\right)}{\exp\left(\frac{\hbar\omega}{kT}\right) - 1} + \frac{\left(F_C^{\text{t}}\right)^{-3/2} \exp\left(-\frac{F_C^{\pm}}{F}\right)}{1 - \exp\left(-\frac{\hbar\omega}{kT}\right)}\right] \tag{B.8}$$

where $\tilde{n}$ and $\tilde{p}$ equal n and p for *DensityCorrection=None*, and are given by: Eq. $\tilde{n} = n\left(\frac{n_{\text{i,eff}}}{N_C}\right)^{\frac{\gamma_n |\nabla E_{F,n}|}{F}}$ for *DensityCorrection=Local*. The critical field strengths read:

$$F_C^+ = B\left(E_{\text{g,eff}} \pm \hbar\omega\right)^{3/2} \tag{B.9}$$

Symbol	Parameter name	Default value	Unit
A	A	8.977×10^{20}	$cm-1s-1V-3/2$
B	B	2.14667×10^{7}	$V/cm-1eV-3/2$
$\hbar\omega$	hbarOmega	18.6	meV

Table B.3: Coefficients for band–to–band tunneling (Schenk model)

The upper sign in Eq. B.6 refers to tunneling generation ($np < n_{i,eff}^2$) and the lower sign refers to recombination ($np > n_{i,eff}^2$). The quantity $\hbar\omega$ denotes the energy of the transverse acoustic phonon.

The parameters [53] are given in table B.3 and can be accessed in the parameter set *Band2BandTunneling*. The defaults were obtained assuming the field direction to be <111>.

Bibliography

[1] G. Charpak and F. Sauli, "High-resolution electronic particle detectors," *Annual Review of Nuclear and Particle Science*, vol. 34, no. 1, pp. 285–350, 1984.

[2] "Large hadron collider experiment." `https://home.cern/science/accelerators/large-hadron-collider`. Accessed: 2019-03-30.

[3] A. Collaboration, "Technical Proposal: A High-Granularity Timing Detector for the ATLAS Phase-II Upgrade," Tech. Rep. CERN-LHCC-2018-023. LHCC-P-012, CERN, Geneva, Jun 2018.

[4] N. Cartiglia, A. Staiano, V. Sola, R. Arcidiacono, R. Cirio, F. Cenna, M. Ferrero, V. Monaco, R. Mulargia, M. Obertino, F. Ravera, R. Sacchi, A. Bellora, S. Durando, M. Mandurrino, N. Minafra, V. Fadeyev, P. Freeman, Z. Galloway, E. Gkougkousis, H. Grabas, B. Gruey, C. Labitan, R. Losakul, Z. Luce, F. McKinney-Martinez, H.-W. Sadrozinski, A. Seiden, E. Spencer, M. Wilder, N. Woods, A. Zatserklyaniy, G. Pellegrini, S. Hidalgo, M. Carulla, D. Flores, A. Merlos, D. Quirion, V. Cindro, G. Kramberger, I. Mandić, M. Mikuž, and M. Zavrtanik, "Beam test results of a 16ps timing system based on ultra-fast silicon detectors," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 850, pp. 83 – 88, 2017.

[5] C. de La Taille, S. Callier, S. C. Di Lorenzo, N. Seguin-Moreau, P. Dinaucourt, G. Martin-Chassard, C. Agapopoulou, N. Makovec, L. Serin, and S. Simion, "ALTIROC0, a 20 pico-second time resolution ASIC for the ATLAS High Granularity Timing Detector (HGTD)," *PoS*, vol. TWEPP-17, p. 006, 2018.

[6] N. Cartiglia, R. Arcidiacono, M. Ferrero, M. Mandurrino, H.-F. Sadrozinski, V. Sola, A. Staiano, and A. Seiden, "Timing layers, 4- and 5-dimension tracking," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 924, pp. 350 – 354, 2019. 11th International Hiroshima Symposium on Development and Application of Semiconductor Tracking Detectors.

[7] J. Rinkel, D. Magalhães, F. Wagner, F. Meneau, and F. Cesar Vicentin, "Detective quantum efficiency for photon–counting hybrid pixel detectors in the tender X–ray domain: application to Medipix3RX," *Journal of Synchrotron Radiation*, vol. 23, pp. 206–213, Jan 2016.

[8] X. Llopart, R. Ballabriga, M. Campbell, L. Tlustos, and W. Wong, "Timepix, a 65k programmable pixel readout chip for arrival time, energy and/or photon counting measurements," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 581, no. 1, pp. 485 – 494, 2007. VCI 2007.

[9] T. Poikela, J. Plosila, T. Westerlund, M. Campbell, M. D. Gaspari, X. Llopart, V. Gromov, R. Kluit, M. van Beuzekom, F. Zappon, V. Zivkovic, C. Brezina, K. Desch, Y. Fu, and A. Kruth, "Timepix3: a 65k channel hybrid pixel readout chip with simultaneous ToA/ToT and sparse readout," *Journal of Instrumentation*, vol. 9, pp. C05013–C05013, may 2014.

[10] J. Fink, E. Schierle, E. Weschke, and J. Geck, "Resonant elastic soft x–ray scattering," *Reports on Progress in Physics*, vol. 76, p. 056502, apr 2013.

[11] A. P. Hitchcock, "Soft x–ray spectromicroscopy and ptychography," *Journal of Electron Spectroscopy and Related Phenomena*, vol. 200, pp. 49 – 63, 2015. Special Anniversary Issue: Volume 200.

[12] *Synopsys Inc. Synopsys TCAD manuals.* 2008.

[13] M. Moll, "Radiation damage in silicon particle detectors: microscopic defects and macroscopic properties," 1999. Presented on 30 Nov 1999.

[14] G. e. a. Lindstroem, "Radiation hard silicon detectors - developments by the rd48 (rose) collaboration," Jul 2001.

[15] "Diamond strucutre image taken from." `http://www.electrical4u.com/energy-brands-in-silicon`. Accessed: 2017-04-01.

[16] S. Sze, *Semiconductor devices, physics and technology.* Wiley, 1985.

[17] "Band diagram image taken from." `https://www.miniphysics.com/formation-of-bands.html`. Accessed: 2017-04-01.

[18] H. K. Charles Kittel, *Thermal Physics.* W.H Freeman, 1980.

[19] W. Maes, K. D. Meyer, and R. V. Overstraeten, "Impact ionization in silicon: A review and update," *Solid-State Electronics*, vol. 33, no. 6, pp. 705 – 718, 1990.

[20] S. Sze, "Physics of semiconductor devices (2nd edition),"

[21] K. K. Thornber, "Applications of scaling to problems in high–field electronic transport," *Journal of Applied Physics*, vol. 52, no. 1, pp. 279–290, 1981.

[22] A. Sinha, "Built–in voltage of a silicon p–n junction having a heavily doped region," *Journal of Applied Physics*, vol. 55, no. 2, pp. 460–462, 1984.

[23] M. D. Hanlon, "The Development of P–Type Silicon Detectors for High Radiation Regions of the LHC," 1998.

[24] S. Ramo, "Currents induced by electron motion," *Proceedings of the IRE*, vol. 27, pp. 584–585, Sep. 1939.

[25] H. B. und J. Ashkin, *Experimental Nuclear Physics.* J. Wiley, 1953.

[26] P. P. Webb and R. J. McIntyre, "Large area reach–through avalanche diodes for x–ray spectroscopy," *IEEE Transactions on Nuclear Science*, vol. 23, pp. 138–144, Feb 1976.

[27] A. Akindinov, A. Martemianov, P. Polozov, V. Golovin, and E. Grigoriev, "New results on mrs apds," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 387, no. 1, pp. 231 – 234, 1997. New Developments in Photodetection.

[28] P. Webb and A. R. Jones, "Large area reach–through avalanche diodes for radiation monitoring," *IEEE Transactions on Nuclear Science*, vol. 21, pp. 151–158, Feb 1974.

[29] P. Fernández-Martínez, I. A. S. Cortés, S. Hidalgo, D. Flores, G. Pellegrini, and M. A. Lozano, "Linear mode reach — through avalanche photodiodes for medium energy x–ray detection," *Proceedings of the 8th Spanish Conference on Electron Devices, CDE'2011*, pp. 1–4, 2011.

[30] J. Kataoka, T. Saito, Y. Kuramoto, T. Ikagawa, Y. Yatsu, J. Kotoku, M. Arimoto, N. Kawai, Y. Ishikawa, and N. Kawabata, "Recent progress of avalanche photodiodes in high-resolution x–rays and –rays detection," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 541, p. 398–404, Apr 2005.

[31] D. Renker and E. Lorenz, "Advances in solid state photon detectors," *Journal of Instrumentation*, vol. 4, pp. P04004–P04004, apr 2009.

[32] M. Kim, C. Avanzini, M. Bagliesi, G. Bigongiari, T. Lomtadze, P. Maestro, P. Marrocchesi, F. Morsani, and R. Zei, "Scintillation and cherenkov light detection with a 3 mm × 3 mm silicon photomultiplier," *Nuclear Physics B - Proceedings Supplements*,

vol. 197, no. 1, pp. 325 – 330, 2009. 11th Topical Seminar on Innovative Particle and Radiation Detectors (IPRD08).

[33] C. Piemonte, A. Ferri, A. Gola, T. Pro, N. Serra, A. Tarolli, and N. Zorzi, "Characterization of the First FBK High-Density Cell Silicon Photomultiplier Technology," Jan. 2013. Funded by the European Union. Part of the European Commission FP7 Open Access pilot.

[34] G. Zappalà, F. Acerbi, A. Ferri, A. Gola, G. Paternoster, V. Regazzoni, N. Zorzi, and C. Piemonte, "Study of the photo-detection efficiency of FBK high-density silicon photomultipliers," *Journal of Instrumentation*, vol. 11, pp. P11010–P11010, nov 2016.

[35] C. Allaire, J. Benitez, M. Bomben, G. Calderini, M. Carulla, E. Cavallaro, A. Falou, D. Flores, P. Freeman, Z. Galloway, and et al., "Beam test measurements of low gain avalanche detector single pads and arrays for the atlas high granularity timing detector," *Journal of Instrumentation*, vol. 13, p. P06017–P06017, Jun 2018.

[36] "Technology developments and first measurements of low gain avalanche detectors (lgad) for high energy physics applications," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 765, pp. 12 – 16, 2014. HSTD-9 2013 - Proceedings of the 9th International.

[37] G.-F. D. Betta, L. Pancheri, M. Boscardin, G. Paternoster, C. Piemonte, N. Cartiglia, F. Cenna, and M. Bruzzi, "Design and tcad simulation of double-sided pixelated low gain avalanche detectors," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 796, pp. 154 – 157, 2015. Proceedings of the 10th International Conference on Radiation Effects on Semiconductor Materials Detectors and Devices.

[38] G. Giacomini, W. Chen, F. Lanni, and A. Tricoli, "Development of a technology for the fabrication of low-gain avalanche diodes at bnl," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 934, p. 52–57, Aug 2019.

[39] N. Moffat, R. Bates, M. Bullough, L. Flores, D. Maneuski, L. Simon, N. Tartoni, F. Doherty, and J. Ashby, "Low gain avalanche detectors (LGAD) for particle physics and synchrotron applications," *Journal of Instrumentation*, vol. 13, pp. C03014–C03014, mar 2018.

[40] S. Otero Ugobono, M. Carulla, M. Centis Vignali, M. Fernández García, C. Gallrapp, S. Hidalgo Villena, I. Mateu, M. Moll, G. Pellegrini, and I. Vila, "Radiation tolerance

of proton–irradiated lgads," *IEEE Transactions on Nuclear Science*, vol. 65, pp. 1667–1675, Aug 2018.

[41] A. Collaboration, "Technical Proposal: A High–Granularity Timing Detector for the ATLAS Phase–II Upgrade," Tech. Rep. CERN-LHCC-2018-023. LHCC-P-012, CERN, Geneva, Jun 2018.

[42] H. Xiao, *Introduction to Semiconductor Manufacturing Technology, Second Edition*. SPIE, 2012.

[43] T. R. N. W. L.Rossi, P. Fischer, *Pixel Detectors: From Fundamentals to Applications*. Springer, 2006.

[44] T. R. N. W. L.Rossi, P. Fischer, *Pixel Detectors: From Fundamentals to Applications*. Springer, 2006.

[45] R. F. K. Herzog and F. P. Viehböck, "Ion source for mass spectrography," *Phys. Rev.*, vol. 76, pp. 855–856, Sep 1949.

[46] J. Biersack, "Basic physical aspects of high energy implantation," *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, vol. 35, no. 3, pp. 205 – 214, 1988.

[47] J. F. Ziegler and J. P. Biersack, *The Stopping and Range of Ions in Matter*, pp. 93–129. Boston, MA: Springer US, 1985.

[48] *Sentaurus Process User Guide G–2012.06*. 2008.

[49] P. Wilson, "The diffusion of boron in the si/sio2 system," *Solid–State Electronics*, vol. 15, no. 9, pp. 961 – 970, 1972.

[50] "Electron and hole drift velocity measurements in silicon and their empirical relation to electric field and temperature„" *IEEE Transactions on Electron Devices*, vol. 22, no. 11, p. 1045–1047, 1975.

[51] D. M. Caughey and R. E. Thomas, "Carrier mobilities in silicon empirically related to doping and field," *Proceedings of the IEEE*, vol. 55, pp. 2192–2193, Dec 1967.

[52] Y. Okuto and C. Crowell, "Threshold energy effect on avalanche breakdown voltage in semiconductor junctions," *Solid–State Electronics*, vol. 18, no. 2, pp. 161 – 168, 1975.

[53] A. Schenk, "Rigorous theory and simplified model of the band–to–band tunneling in silicon," *Solid–State Electronics*, vol. 36, no. 1, pp. 19 – 34, 1993.

[54] I. Cortés, P. Fernández-Martínez, D. Flores, S. Hidalgo, and J. Rebollo, "Gain estimation of rt–apd devices by means of tcad numerical simulations," in *Proceedings of the 8th Spanish Conference on Electron Devices, CDE'2011*, pp. 1–4, Feb 2011.

[55] B. J. Baliga, *Fundamentals of Power Semiconductor Devices.* Springer, 2008.

[56] P. Fernandez-Martinez, D. Flores, S. Hidalgo, V. Greco, A. Merlos, G. Pellegrini, and D. Quirion, "Design and Fabrication of an Optimum Peripheral Region for Low Gain Avalanche Detectors," *Nucl. Instrum. Meth.*, vol. A821, pp. 93–100, 2016.

[57] V. A. K. Temple, "Increased avalanche breakdown voltage and controlled surface electric fields using a junction termination extension (jte) technique," *IEEE Transactions on Electron Devices*, vol. 30, pp. 954–957, Aug 1983.

[58] V. Gkougkousis, *Background studies on the H→ZZ*→4l channel at LHC Run 1. Prospects of the bbH(→γγ) mode and studies for an improved pixel detector system for the ATLAS upgrade towards HL-LHC.* PhD thesis, Orsay, 2016-02-26.

[59] E. Frojdh, M. Campbell, M. D. Gaspari, S. Kulis, X. Llopart, T. Poikela, and L. Tlustos, "Timepix3: first measurements and characterization of a hybrid–pixel detector working in event driven mode," *Journal of Instrumentation*, vol. 10, pp. C01039–C01039, jan 2015.

[60] G. Kramberger, "Advanced Transient Current Technique Systems," *PoS*, vol. Vertex2014, p. 032, 2015.

[61] R. Mulargia, R. Arcidiacono, A. Bellora, M. Boscardin, N. Cartiglia, F. Cenna, R. Cirio, G. D. Betta, S. Durando, A. Fadavi, M. Ferrero, Z. Galloway, B. Gruey, P. Freeman, G. Kramberger, I. Mandic, V. Monaco, M. Obertino, L. Pancheri, G. Paternoster, F. Ravera, R. Sacchi, H. Sadrozinski, A. Seiden, V. Sola, N. Spencer, A. Staiano, M. Wilder, N. Woods, and A. Zatserklyaniy, "Temperature dependence of the response of ultra fast silicon detectors," *Journal of Instrumentation*, vol. 11, pp. C12013–C12013, dec 2016.

[62] C. R. Crowell and S. M. Sze, "Temperature dependence of avalanche multiplication in semiconductors," *Applied Physics Letters*, vol. 9, no. 6, pp. 242–244, 1966.

[63] G. F. Knoll, *Radiation detection and Measurement.* John Wisley and Sons, 1979.

[64] J. e. a. Beringer, "Review of particle physics," *Phys. Rev. D*, vol. 86, p. 010001, Jul 2012.

[65] G. Apollinari, I. Béjar Alonso, O. Brüning, M. Lamont, and L. Rossi, *High-Luminosity Large Hadron Collider (HL-LHC):Preliminary Design Report*. CERN Yellow Reports: Monographs, Geneva: CERN, 2015.

[66] H. F.-W. Sadrozinski, A. Seiden, and N. Cartiglia, "4d tracking with ultra-fast silicon detectors," *Reports on Progress in Physics*, vol. 81, p. 026101, dec 2017.

[67] "ATLAS Phase-II Upgrade Scoping Document," Tech. Rep. CERN-LHCC-2015-020. LHCC-G-166, CERN, Geneva, Sep 2015.

[68] A. Collaboration, "Technical Proposal: A High-Granularity Timing Detector for the ATLAS Phase-II Upgrade," Tech. Rep. CERN-LHCC-2018-023. LHCC-P-012, CERN, Geneva, Jun 2018.

[69] C. Jacoboni, C. Canali, G. Ottaviani, and A. A. Quaranta, "A review of some charge transport properties of silicon," *Solid-State Electronics*, vol. 20, no. 2, pp. 77 – 89, 1977.

[70] "Timing measurement on thin lgad tredi 2017." `https://indico.cern.ch/event/587631/contributions/2471694/attachments/1415772/2167581/Trento.pptx`. Accessed: 2019-09-30.

[71] S. G. J. H. C. S. S.C. Anspach, L.M. Cavallo, *Half Lives of Materials Used in the Preparation of Standard Reference Materials of Nineteen Radioactive Nuclides Issued by the National Bureau of Standards*.

[72] G. L. Patton, J. H. Comfort, B. S. Meyerson, E. F. Crabbe, G. J. Scilla, E. de Fresart, J. M. C. Stork, J. Y. . Sun, D. L. Harame, and J. N. Burghartz, "75-ghz f/sub t/ sige-base heterojunction bipolar transistors," *IEEE Electron Device Letters*, vol. 11, pp. 171–173, April 1990.

[73] "Readout board for timing." `https://twiki.cern.ch/twiki/pub/Main/UcscSingleChannel/proto_beam_test_board_V1.1C.pdf`. Accessed: 2019-09-30.

[74] "Gali-52+ galium phostpahte boradband amplifier." `https://www.minicircuits.com/WebStore/dashboard.html?model=GALI-52%2B`. Accessed: 2019-06-30.

[75] "Evaluation board with gali-52+ amplifer." `https://www.minicircuits.com/pcb/WTB-409-52+_P02.pdf`. Accessed: 2019-06-30.

[76] F. Caspers and P. Kowina, "RF Measurement Concepts," 2014.

[77] "Emerging fine-pitch bump bonding techniques." `https://indico.cern.ch/event/104203/contribution/0/attachments/14520/21112/Emerging_fine_pitch_bump_bonding_techniques_03-September-2010_-_Sami_VAEHAENEN.pptx`. Accessed: 2019-09-30.

[78] R. Ballabriga, M. Campbell, and X. Llopart, "Asic developments for radiation imaging applications: The medipix and timepix family," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 878, pp. 10 – 23, 2018. Radiation Imaging Techniques and Applications.

[79] X. Llopart, M. Campbell, R. Dinapoli, D. San Segundo, and E. Pernigotti, "Medipix2: A 64-k pixel readout chip with 55-/spl mu/m square elements working in single photon counting mode," *IEEE Transactions on Nuclear Science*, vol. 49, pp. 2279–2283, Oct 2002.

[80] R. Ballabriga, M. Campbell, E. Heijne, X. Llopart, L. Tlustos, and W. Wong, "Medipix3: A 64k pixel detector readout chip working in single photon counting mode with improved spectrometric performance," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 633, pp. S15 – S18, 2011. 11th International Workshop on Radiation Imaging Detectors (IWORID).

[81] E. H. H. et al, "Lhc1: A semiconductor pixel detector readout chip with internal, tunable delay providing a binary pattern of selected events," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 383, no. 1, pp. 55 – 63, 1996. Development and Application of Semiconductor Tracking Detectors.

[82] T. Poikela, R. Ballabriga, J. Buytaert, X. Llopart, W. Wong, M. Campbell, K. Wyllie, M. van Beuzekom, J. Schipper, S. Miryala, and V. Gromov, "The VeloPix ASIC," *Journal of Instrumentation*, vol. 12, pp. C01070–C01070, jan 2017.

[83] D. Dutta, "The LHCb VELO Upgrade," *PoS*, vol. VERTEX2018, p. 022, 2019.

[84] D. Lacour, "A high-granularity timing detector for the phase-II upgrade of the ATLAS calorimeter system: detector concept description and first beam test results," *Journal of Instrumentation*, vol. 13, pp. C02016–C02016, feb 2018.

[85] "First production of resistive ac-coupled silicon detectors (rsd) at fbk." `https://indico.cern.ch/event/812761/contributions/3459062/attachments/1860923/3059631/2019Mandurrino_34thRD50.pdf`. Accessed: 2019-09-30.

[86] J. Lindhard and M. Scharff, "Energy dissipation by ions in the kev region," *Phys. Rev.*, vol. 124, pp. 128–130, Oct 1961.

[87] O. S. Oen and M. T. Robinson, "Computer studies of the reflection of light ions from solids," *Nuclear Instruments and Methods*, vol. 132, pp. 647 – 653, 1976.

[88] G. Masetti, M. Severi, and S. Solmi, "Modeling of carrier mobility against carrier concentration in arsenic–, phosphorus–, and boron–doped silicon," *IEEE Transactions on Electron Devices*, vol. 30, pp. 764–769, July 1983.