Characterisation of Low Gain Avalanche Detectors for Large Area Precision Timing at Particle Colliders

Jonathan Franklin Mulvey

Thesis submitted for the degree of
Doctor of Philosophy

Particle Physics Group,
School of Physics and Astronomy,
University of Birmingham.

June 28, 2024

ABSTRACT

Low Gain Avalanche Detectors (LGADs) are the chosen technology for the timing layers which are to be placed in the forward region of the inner trackers in the AT-LAS and CMS experiments. Teledyne e2v (Te2v) is a silicon foundry known for producing large volume CCDs. This work aims to establish them as a future vendor of LGADs, particularly for High Energy Physics applications. Experimental set-ups and techniques have been successfully developed and refined in order to characterise the IV, CV, gain and timing properties of LGADs. Characterisation has been completed before and after irradiation with $27\,\mathrm{MeV}$ protons at the MC40 cyclotron in Birmingham. These results are reported here with key values compared to LGADs from alternative vendors. Te2v's LGADs achieve sub-$40\,\mathrm{ps}$ time resolution at a gain of ~ 20 or above. The time resolution also behaves similarly as a function of gain compared to other vendors. After irradiation, they can still achieve this time resolution below a fluence of $5.7 \times 10^{14}\,1\,\mathrm{MeV}\,\mathrm{n_{eq}/cm^2}$. At this fluence and above, the timing performance is limited to $50\,\mathrm{ps}$ at a gain of ~ 10. Gain non-uniformity across similar devices was observed in IV and gain measurements. The acceptor removal coefficient was measured to be $(9.7 \pm 0.5) \times 10^{-16}\,\mathrm{cm^2}$, $45\,\%$ larger than reported for Fondazione Bruno Kessler's (FBK) UFSD2 production where it is $6.7 \times 10^{-16}\,\mathrm{cm^2}$. Overall, this work has demonstrated that Te2v is capable of manufacturing LGADs with comparable performance to existing vendors. With further refinement of the gain layer implant energy and dose, Te2v would be a highly competitive supplier with potential for meeting the large orders anticipated for many future particle physics facilities.

DECLARATION OF AUTHORS CONTRIBUTION

I confirm that work presented here is my own unless expressed otherwise. While we operated under the Te2v collaboration, characterisation was primarily performed independently at the University of Oxford, the Rutherford Appleton Laboratory and the University of Birmingham.

Chapter 3 contains a discussion of the experimental set-ups and methods used in this work. I have made a contribution to the iteration and improvement of pre-established set-ups and methods. For measurements on the probe station and climate chamber, my contribution was the development of customised Python scripts which control the test procedure. These scripts are based on code initially developed by Matt Basso (University of Toronto), Tim Knight (University of Toronto), Will George (University of Birmingham) and Ioannis Kopsalis (University of Birmingham), with further alterations by Alasdair Winter (University of Birmingham). Gain measurements are performed using the TCT system from Particulars. I contributed to the design of the environmental chamber and the beam monitor calibration system. Again, my focus was on the development of custom scripts interfacing with the movable optical bench and the oscilloscope. The timing set-up was initially based on designs from colleagues at Ljubljana. However, I worked closely with one of our technicians (at the time), James Glover, to commission a new environmental chamber for cold measurements. We also worked together to redesign the entire set-up for more consistent cooling. I would also like to credit Daniel Thompson who suggested the method for measuring the time resolution of three LGADs without needing a reference device. Lastly, irradiations were performed by Amelia Hunter.

Chapter 4 contains the results which have all been acquired and analysed by myself. This is with the exception of the section on wafer probing. Wafer probing data was taken by two masters students: Marcus Madurai and Shivani Kapur. I provided insight and guidance throughout their masters project, including teaching

the operation of each set-up. However they performed the actual measurements and preliminary analysis, while I performed the final analysis.

Chapter 5 is the discussion of my work on a simulation package for predicting the Modulation Transfer Function (MTF) of CCDs. Where appropriate, the mathematics and analysis that supports the simulation is referenced. All of the code is written by myself, from scratch. Most of the challenges faced were solved by myself with the occasional input from my father, Kevin Mulvey (on the statistical analysis side), and from my industrial supervisor, Douglas Jordan (for the core concepts of CCDs, MTF theory and experimental methods).

ACKNOWLEDGEMENTS

Sometimes, one of the greatest motivators is to be told you cannot do something (be it by your peers, compiler or lab equipment). There is something innately human in the desire to prove that you can rise to the challenge, and with the submission of this thesis, I feel I have risen to one of the greatest academic challenges. It's no secret that I have always loved physics, desperately waiting to be old enough to be taught the "good stuff". I am still waiting, but in the meantime I've been lucky enough to join the University of Birmingham and be given the opportunity to be involved with some of the latest trends in semiconductor physics. I have thoroughly enjoyed every aspect of PhD life, but I have many people to thank for making it such a fantastic experience.

First and foremost, I would not be writing this thesis without my supervisors: Laura Gonella and Phil Allport. Their support has been immeasurable. They were always willing to discuss my (many) problems, offer solutions and even babysit experiments with me in the clean room. I always felt my ideas were given an equal weight in discussions. This is something I really appreciated and no doubt helped me to grow in confidence. I also want to thank Laura for not failing me in order to maintain the integrity of the PhD programme, despite continuous threats. Joking aside, it has genuinely been a pleasure to work with you both. I would like to thank my industrial supervisor Douglas Jordan. It has been a pleasure to work next to you (1 m to be exact). Even after my secondment in Chelmsford, you have been a fantastic supervisor, always willing to help.

Of course, it would also have been impossible to do anything in the last four years without support and advice from all the academics, postdocs, technical staff and fellow PhD students. I want to thank Ioannis Kopsalis for his help getting me up to speed with all of the equipment in the clean room in my first year. I also want to thank all of the technicians: Simon, James "the technician", Eve, Robbie and

Amelia for all of the *many* times they have stopped what they were doing to suit-up and perform a single wire-bond for me. I'd also like to thank James "the postdoc" specifically. He has been someone who I could rely on, either for quick discussion when I am panicking that my results are meaningless, or to fix my endless problems in the lab. You really have been invaluable. In general, I'd also like to thank my many colleagues, especially those I have had the pleasure to share an office with: Dan L, Gov, Patrick F, Dan T, Marcus, Ellie, Felicia, Niladri, Mihaela, Adrien, Adam, Eric, Marc, Julia, Josh, Gio, Harry, Ilyas, Stephen, Lex, Andy, Karol, Eleni, Alexandra, Ivan, Marc, Mark, Patrick K, Artem, Paul, Kevin, Steve, Nigel, Miriam, Dan J, Jeurgen, Panagiotis, Chris and Midas. They have all been fantastic colleagues and even better friends, a large part of the reason why this PhD has been such a great experience.

Lunch is usually a time of great excitement for students, especially those doing a PhD. But of course, no lunch is complete without an hour of sporcle. PhDs are hard, but without fail, lunch has genuinely been the highlight of some otherwise miserable days. I don't think I have ever laughed quite as hard as I have during those lunches, and for that I want to thank Robbie, Amelia and Alasdair for being fantastic company and support.

It might be difficult, but it's important that your PhD does not completely take over your life. Joining the Panthers Korfball Club was one of the best decisions I've made while in Birmingham and I'd like to thank everyone who I've met along the way. I have never left a session without a smile on my face and an unbearable urge to join everyone at the pub afterwards. Pool and Snooker has been in my blood since my undergraduate days, and SaPSoc has been the natural continuation. From the weekly sessions to the countless weekends away (but usually just to Stoke), I would not have kept my sanity if it wasn't for the friendship of so many of you. In particular I'd like to thank: Eliott, Ben, Paaras, Jat, Jon "Junior", Dan B, Henry, Max, James, Sam, Donald and Dan R.

Last but certainly not least, I would like to thank my family and close friends. I'd like to thank one of my closest friends, Ashley. We might not speak very often, but when we do, it feels like only minutes have passed by. You've always been there for me, willing to give me advice on any topic under the sun. Mum, Dad and Sarah, you have all been such wonderful support, not only over these last few years, but throughout my life. You've supported me financially and emotionally (and academically on occasions). You've encouraged me to keep going even when I don't want to hear it. No matter what time it is, I know I can always call and count on you, and that is something I will forever appreciate.
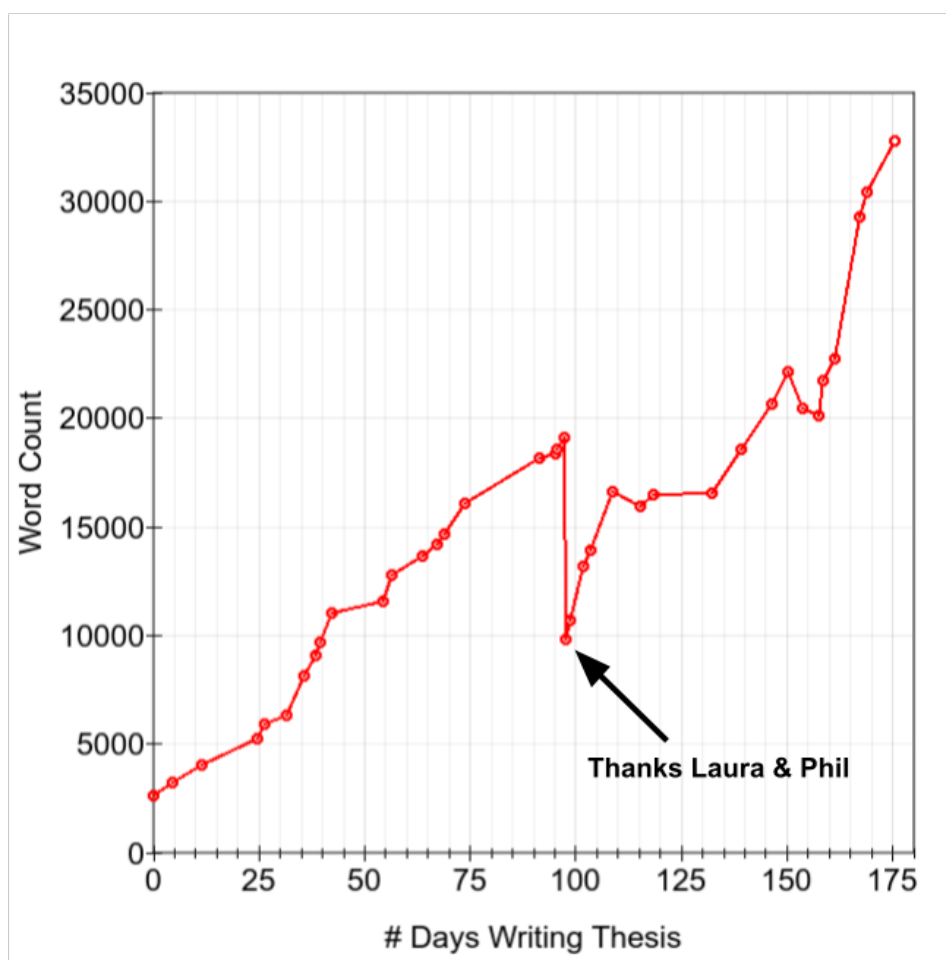
"It's worse than that, it's physics Jim"

The Firm

"Do. Or do not. There is no try"

(Not a physicist)

"wibbly-wobbly, timey-wimey"

LGAD Expert

Corporate needs you to find the differences between these two Santa Cruz Timing Boards. (One PCB was assembled by hand, by me. The other PCB was provided to us, professionally made.)

Thanks Laura & Phil

Contents

List of Tables

List of Figures

DEFINITIONS OF ACRONYMS

AC-LGAD ac-coupled LGAD

ATLAS A Toroidal LHC Apparatus

CCE charge collection efficiency

CCD charge coupled device

CFD constant fraction discriminator

CMS Compact Muon Solenoid

CSP charge sensitive preamplifier

CTT coincidence timing technique

DUT device under test

ESA European Space Agency

FBK Fondazione Bruno Kessler

FEM finite element method

FEMM Finite Element Method Magnetics

FFT fast fourier transform

FWHM full width at half maximum

GaAs gallium arsenide

HEP High-Energy Physics

HGTD High Granularity Timing Detector

HL-LHC High-Luminosity LHC

iLGAD inverse-LGAD

JTE junction termination edge

LGAD Low Gain Avalanche Detector

LHC Large Hadron Collider

LSST Large Synoptic Survey Telescope

mip minimum-ionising particle

MPV most-probable-value

MTF modulation transfer function

NASA National Aeronautics and Space Administration

ND neutral-density

NIEL non-ionising energy loss

op-amp operational-amplifier

PiN devices with no gain layer

PSF point spread function

RH relative humidity

RMS root mean square

SEB single event burnout

SIMS secondary ion mass spectrometry

SMU source measurement unit

SNR signal to noise ratio

Sr-90 strontium-90

TCAD technology computer-aided design

TCT transient current technique

TDC time-to-digital convertor

Te2v Teledyne e2v

ToT time-over-threshold

TVS transient-voltage-suppression

UFSD ultra-fast silicon detector

UKSA United Kingdom Space Agency

CHAPTER 1

INTRODUCTION

The Large Hadron Collider (LHC) is a proton-proton collider, commissioned in 2010 with a target centre-of-mass energy of 14 TeV [1]. The LHC has ushered in a new era for High-Energy Physics (HEP) which culminated in the discovery of the Higgs boson in 2012. It is expected that the LHC will be able answer many more questions about dark matter, supersymmetry or extra dimensions [1]. Since being commissioned, the LHC has gone through two long shutdowns for maintenance and improvements to the luminosity[1] as shown in Figure 1.1. The third long shutdown is for the LHC's high luminosity upgrade, and work is now expected to be completed by 2029. This phase of the LHC will see a factor of 10 increase in the instantaneous luminosity, up to $5 \times 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ [2]. The High-Luminosity LHC (HL-LHC) will see up to 200 events per 25 ns bunch crossing, which is an order of magnitude increase compared to the LHC's design specifications. An increased luminosity therefore results in a

---

[1]The instantaneous luminosity is a measure of the number of particles and how tightly packed they are in the colliding beams. When multiplied by the cross-section it gives the total interaction rate per second.
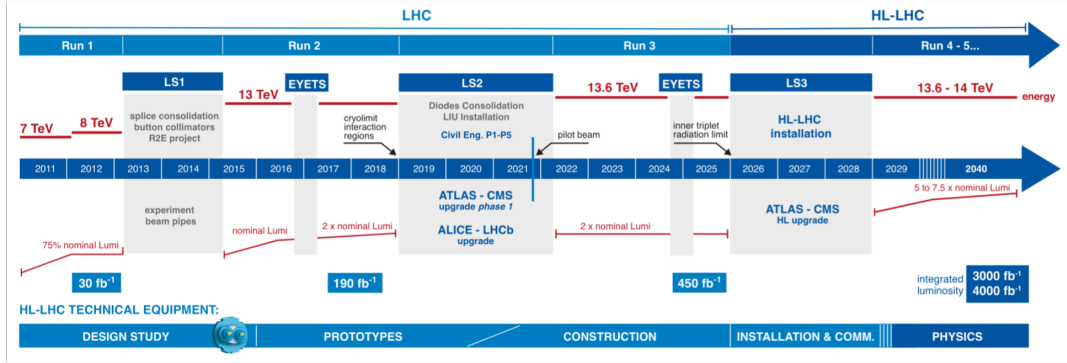
Figure 1.1: The planned maintenance and upgrades for the LHC as agreed in 2015. Taken from [4].

greater number of interactions (events) per unit time. This is vital for the future physics programme of the LHC. Following the discovery of the Higgs boson, a higher integrated luminosity will help with further characterisation of its properties to a much higher precision [3]. Beyond the Standard Model, searches for dark matter and long lived particles will become more common with previously statistically limited opportunities now becoming feasible [3].

The increase in instantaneous luminosity means an increase in the density of particle tracks emanating from the collision point at each bunch crossing. This is visualised in Figure 1.2. In particular, the track density in the direction parallel to the beam line is very high compared to the transverse case. This makes track reconstruction challenging, limiting the ability to associate tracks with the correct primary or secondary vertices. The ATLAS (A Toroidal LHC Apparatus) and CMS (Compact Muon Solenoid) experiments plan to install dedicated timing layers in the forward regions of the detector, just before the end-cap calorimeters [5, 6]. The High Granularity Timing Detector (HGTD) of the ATLAS experiment is shown in Figure 1.3. It provides two hits with fine timing information ($\sim 10\,\mathrm{ps}$) on tracks in the forward region of the detector (pseudorapidity between 2.5 and 4) where this is most useful due to the heavy congestion of tracks in the forward direction at colliders [5]. For further details on the spatial and temporal resolution, see [5, 6].

Low Gain Avalanche Detectors (LGADs) are a novel silicon sensor technology de-
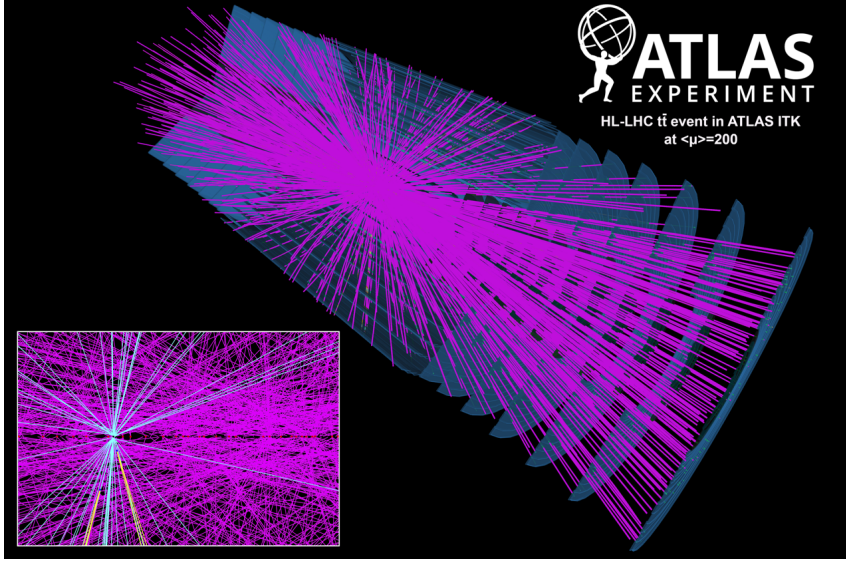
Figure 1.2: A simulated $t\bar{t}$ event. The HL-LHC will see a large increase in luminosity which will result in a large increase in unwanted scattering events known as pile-up. Tracks from the $t\bar{t}$ vertex are shown in cyan while tracks from secondary vertices are shown in yellow. The remaining (purple) tracks represent the intense amount of pile-up present in an interaction with 200 collisions per bunch crossing. Taken from [7].

signed to provide excellent temporal resolution with moderate spatial resolution ($\sim 300\,\mu$m). Low internal gain provides a controlled charge multiplication and allows them to achieve a 35 ps time resolution. They have been chosen as the primary detector technology in the timing disks in the ATLAS and CMS experiments, their first major application. Currently, LGADs are limited to a pixel size of $(1.3 \times 1.3)\,\text{mm}^2$, over 10 times the size of the pixels in the inner tracker. However, the HGTD is placed far enough away from primary vertices for this granularity to be sufficient to correctly assign the timing information to the correct track. This is the first step towards full large area 4D tracking at future colliders where timing information is added to hits in each tracking layer. For this however, the spatial resolution of LGADs will need much improvement, especially for tracking layers close to primary vertices. First discussed in 2012, with their first measurements reported in 2014, LGADs have seen significant improvements and alterations in the last twelve years [8, 9]. LGADs used in the HGTD will need to be able to cope with radiation fluences up to $1 \times 10^{16}\,1\,\text{MeV}\,\text{n}_{\text{eq}}/\text{cm}^2$. Thus characterising the acceptor removal in the internal gain layer (the main impact on LGAD performance after irradiation) is
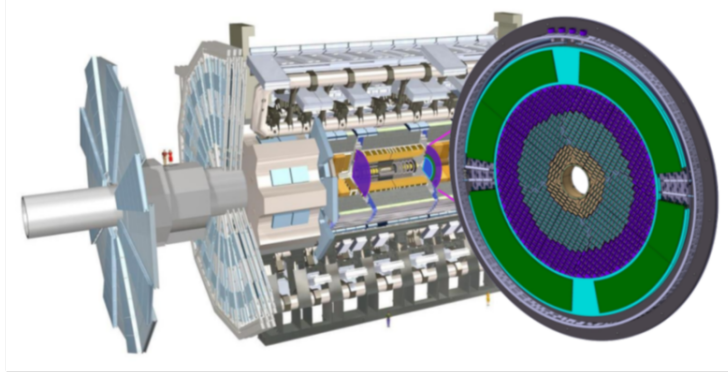
Figure 1.3: The ATLAS detector with the HGTD placed in the forward region of the inner tracker. Taken from [5].

crucial for assessing their performance under these high radiation conditions.

Various foundries have been developing LGADs in preparation for the HL-LHC [10, 11, 12]. Teledyne e2v (Te2v) is an established foundry but new to LGAD development. Established in 1947 as EEV, they have become a major producer of charge coupled device (CCD) imaging sensors for space, medical and defence applications [13]. The Te2v LGAD project, which is the subject of this thesis, is a collaboration between the University of Birmingham, University of Oxford, Rutherford Appleton Laboratory and the Open University, aimed at establishing Te2v as a potential future vendor for LGADs particularly in HEP applications. Te2v can provide a large production volume to HEP, well suited to equip areas of 10s to 100s of m^2 at future collider facilities. In general, the introduction of new vendors is important to ensure that detector development in HEP continues to innovate, driving down costs while driving up quality. In view of this, Te2v's first batch of LGADs finished fabrication in late 2020. This thesis is a discussion of the techniques, methods and set-ups which were developed at the University of Birmingham to characterise LGADs, presented in chapter 3. This is followed by a study of Te2v's LGADs including a study before and after irradiation. Key results will be presented in chapter 4 in the context of other vendors to show that Te2v is a suitable vendor for LGADs in the future.

Chapter 5 is a standalone chapter which documents work on a simulation package de-

signed to calculate the modulation transfer function (MTF) of CCDs manufactured by Te2v. The theoretical model is discussed including how it motivates decisions made for the simulation framework. Results from the simulation are compared to experimental results of Te2v's CCD250 detector. Based on this, current limitations and suggested improvements are then discussed.

CHAPTER 2

THEORY OF SILICON DETECTORS

In early 1940, Russell Ohl stood before his colleagues and a simple apparatus. A voltmeter was connected to either end of an inch-long black rod made of silicon with an impurity gradient along the rod. As Ohl pointed a flashlight towards the sample, the voltmeter flickered by almost half a volt. At the time, this was the most significant response to a light source anyone had seen from any photocell at the time. What Ohl had just stumbled across would later be called the pn junction [14].

Since then, the principle of the pn junction has been used to develop semiconductor detectors which play a crucial role in particle physics experiments. Finely segmented as pixels or strips, they can provide precision tracking of charged particles generated by colliders at these experiments. This chapter will explore the theory of silicon detectors and how their design can be modified into the LGAD, which is the primary focus of this work.

2.1   Semiconductor Theory

2.1.1   The Band-Gap Model

Silicon is a group-IV element with 4 valence electrons which can covalently bond with 4 other silicon atoms forming a face-centered diamond cubic lattice. Silicon is an intrinsic semiconductor with a resistivity of $230\,\mathrm{k\Omega cm}$ at room temperature [15].

Electrons in a solid can occupy various energy states. The band-gap model assumes that densely packed states can be treated as energy bands within which electrons can easily move between states. This is represented by the density of states, $N(E)$, shown in Figure 2.1. The highest occupied and lowest unoccupied bands are known as the conduction and valence bands respectively and are separated by the band gap energy, E_G. A consequence of the Pauli exclusion principle is that electrons tend to fill the lowest state which is not already occupied. Depending on the material in question, this leads to an empty, nearly empty or partially filled conduction band. This corresponds to an insulator, semiconductor and conductor[1] respectively. The band structure of a semiconductor is shown in Figure 2.1. The Fermi level, E_f, is defined as a hypothetical energy level which has a $50\,\%$ probability of being filled at any given time [15].

In silicon, the Fermi energy lies in the middle of the band gap which means the valence band is nearly filled. The band gap is $1.12\,\mathrm{eV}$ at room temperature and is small enough for thermally excited electrons to be promoted to the conduction band [15]. These electrons are free to move and conduct current, however they also leave behind a vacancy which other valence electrons can move into and fill. These vacancies are called holes and are treated as positive charge carriers which are also free to move. As free electrons and holes traverse around the material, they can meet and recombine. At a given temperature, T, thermal equilibrium is achieved once the recombination rate matches the rate of thermal electron-hole pair generation.

---

[1]A conductor can also have overlapping valance and conduction bands.
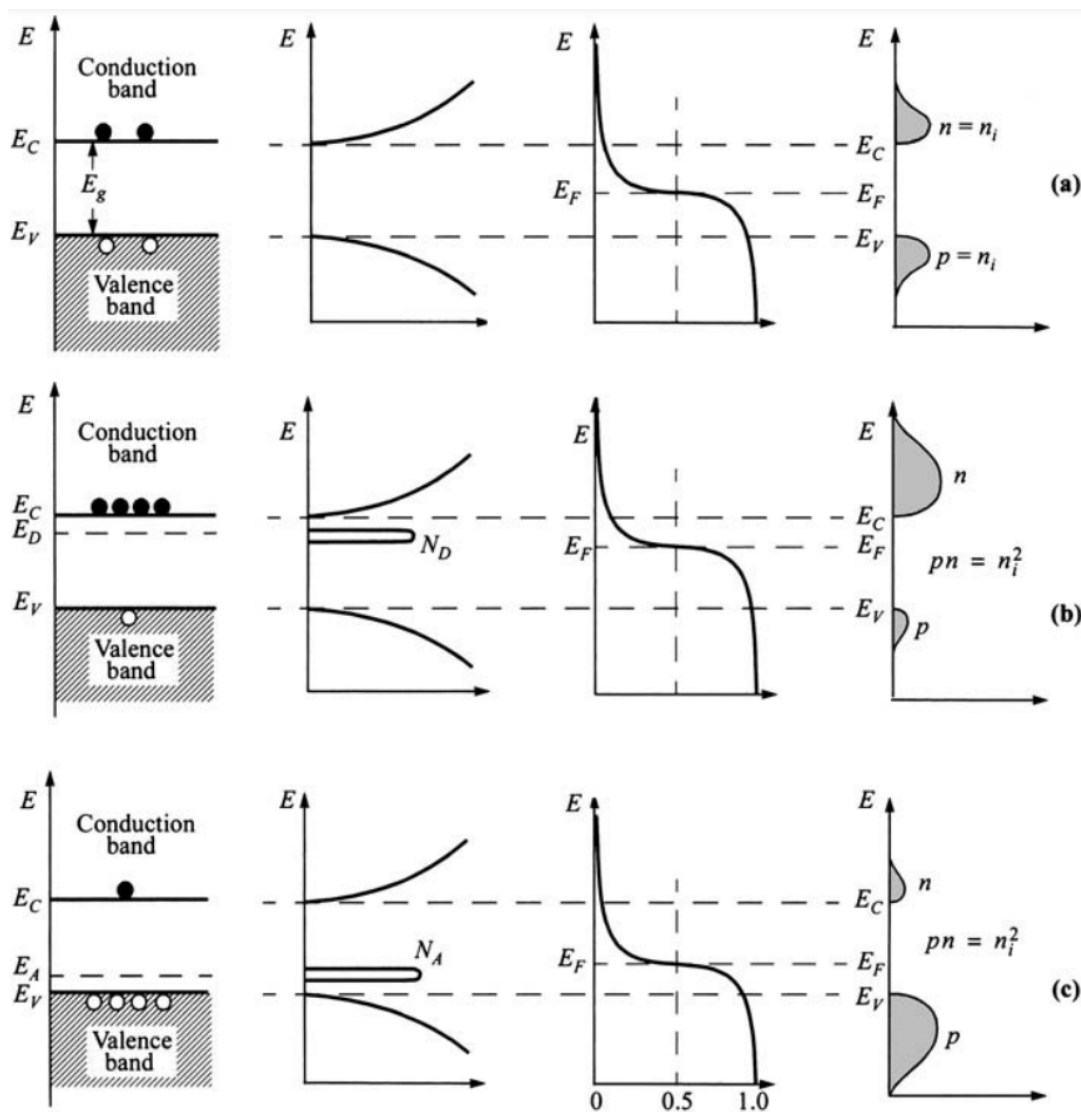
Figure 2.1: The band structure, density of states, Fermi-Dirac distribution and carrier concentrations are shown for (a) intrinsic (b) n-type and (c) p-type semiconductors. Taken from [15].

The consequence of this is the mass-action law,

$$n \cdot p = n_i^2 = const., \tag{2.1}$$

where n and p are the electron and hole densities respectively [15]. n_i is the intrinsic charge carrier density and has a temperature dependence given by,

$$n_i \propto T^{\frac{3}{2}} \cdot \exp\left(-\frac{E_G}{k_B T}\right), \tag{2.2}$$

where k_B is the Boltzmann constant. This dependence is not desirable as it means the resistivity is not well controlled while the number of charge carriers ($\sim 1.5 \times 10^{10}$ cm^{-3}) is also very small compared to the density of silicon atoms ($\sim 1 \times 10^{22}$ cm^{-3}). To avoid this silicon is purposely doped with impurities. This is discussed further in Section 2.1.1.2.

2.1.1.1   Direct & Indirect Semiconductors

Within the band structure of a semiconductor lattice, the maximum and minimum of the conduction and valance bands respectively can be displaced with respect to reciprocal ($\mathbf{k}$) space, as shown in Figure 2.2 [16]. The band gap comes from the energy difference between the minimum and maximum, in $\mathbf{k}$-space, of the conduction and valence bands respectively. However, these extremes do not occur at the same value of $\mathbf{k}$. Materials like this, such as silicon, are called indirect semiconductors. In order for a transition to occur from the valence to the conduction band, an additional momentum transfer to the lattice is also required. This means that a direct transition in silicon requires 3.6 eV. Transitions can still occur with an energy between 1.12 eV and 3.6 eV, albeit infrequently. In this case, the additional energy needs to come from the phonons in the lattice. The phonon energy is governed by the Bose-Einstein distribution which means that the probability of a indirect transition occurring is
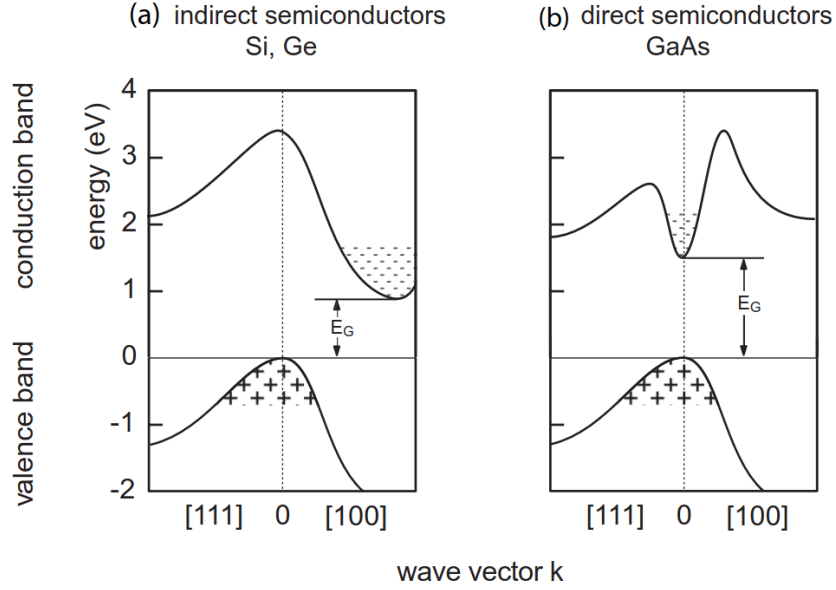
Figure 2.2: Example of a direct and indirect semiconductor. Transitions in an indirect semiconductors require energy in addition to the band gap energy. Taken from [16].

dependent on temperature [16].

2.1.1.2   Extrinsic (Doped) Silicon

Doping is a technique by which the conductivity of a semiconductor is increased in a controlled manner by introducing non-group-IV atoms [15]. Group-V elements, such as phosphorus, have five outer shell electrons and are known as donor atoms. When covalently bonded with silicon, only four can bond, leaving the fifth free to move and conduct. Silicon doped with donor atoms is known as n-type. Conversely, p-type silicon is created when a group-III element is introduced, such as boron. With its three outer shell electrons, boron is an acceptor atom and leaves behind a free hole once covalently bonded to silicon. These free charge carriers can be understood as new donor and acceptor energy levels which form close to the conductor and valence bands respectively, as shown in Figure 2.1 [15]. The energy difference between the new levels and their respective bands is on the order of 1×10^{-2} eV, which is easily overcome by thermal excitations [15].

The mass action law, $n_i^2 = np$, still holds for doped silicon however the charge carrier concentrations, n and p are no longer equal. In the case of n-type silicon, $n \gg p$ and hence electrons are known as the majority charge carriers while holes are the minority charge carriers. The reverse is true for p-type silicon. While the majority charge carrier concentrations change, doped silicon remains electrically neutral.

It can be shown that in n-type and p-type silicon respectively, the majority charge carrier concentrations are given by,

$$n_n \approx N_D \quad \text{and} \quad p_p \approx N_A \text{ [15]}.$$
(2.3)

Here, N_D and N_A is the number density of donor and acceptor atoms respectively. The corresponding minority charge carrier concentrations can be found with the mass action law.

Figure 2.1 shows the band structure and density of states, $N(E)$, for intrinsic, n-type and p-type silicon. The Fermi-Dirac distribution is also shown here and describes the average occupancy of a state at energy, E, as

$$F(E) = \frac{1}{1 + \exp\left(\frac{E - E_F}{k_B T}\right)} \text{ [15]}.$$
(2.4)

This distribution is dependent on the Fermi energy, E_F, which is modified by the new energy levels as shown in Figure 2.1.

2.1.2   The PN Junction At Equilibrium

A pn junction is formed when a single piece of silicon is doped with n- and p-type dopants in opposing halves as shown in Figure 2.3. This creates a strong concentration gradient of charge carriers which leads to a diffusion current, I_{diff} [15]. This means that electrons diffuse from n- to p-type and holes diffuse from p-
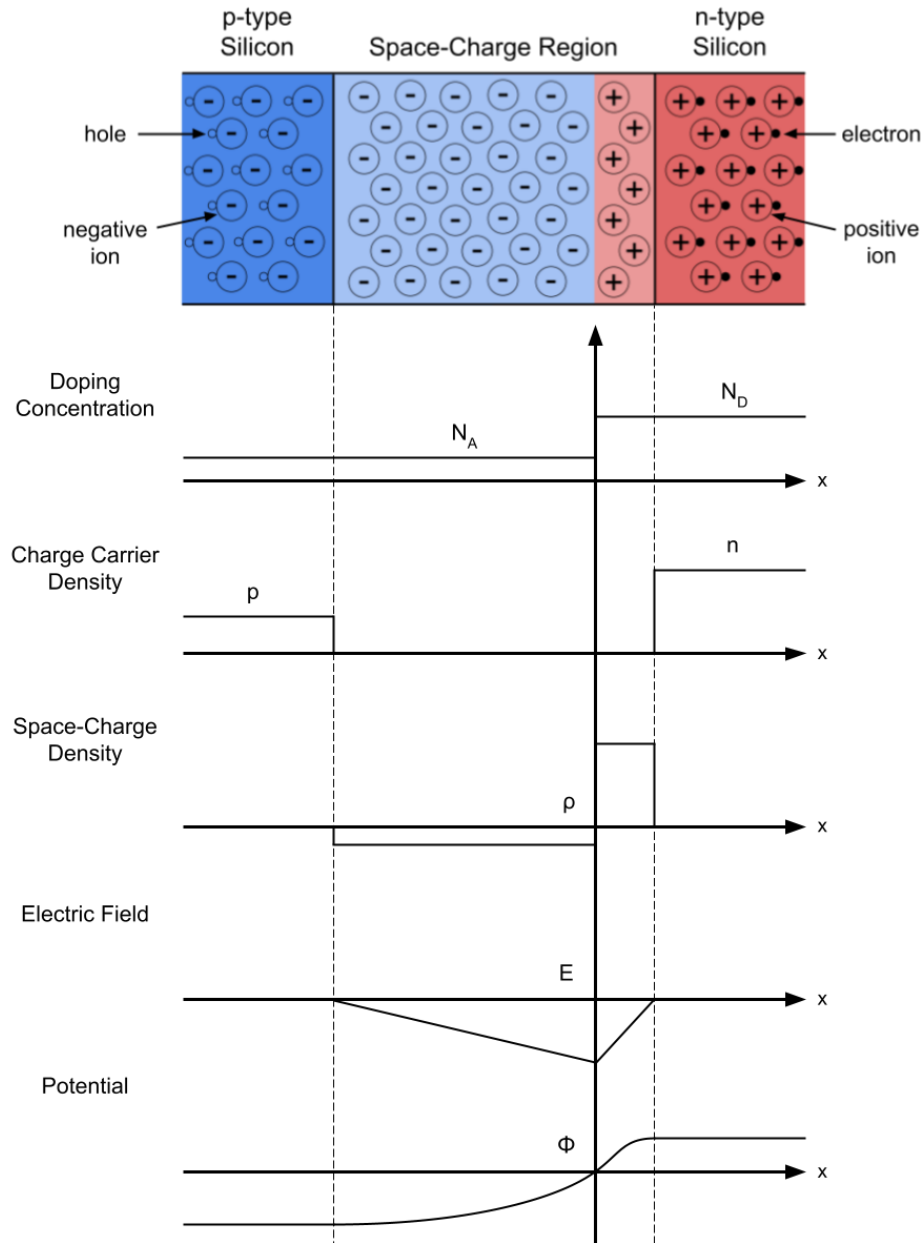
Figure 2.3: At thermal equilibrium, a space-charge region forms around the pn boundary. This region is depleted of charge carriers and has a potential difference between both ends, which creates an electric field across the junction. Adapted from [16].

to n-type. At the pn junction's boundary, electrons and holes meet and recombine. This creates a space-charge region around the boundary which is depleted of charge carriers, sometimes called the depletion region. Charge carriers which recombine leave behind ions which are fixed charges in the lattice structure. The space-charge region is assumed to have a homogeneous charge density, $\rho(x)$, given by,

$$\rho(x) = \begin{cases} -eN_A, & \text{-}x_p < x < 0, \\ +eN_D, & 0 < x < x_n, \\ 0, & elsewhere, \end{cases} \tag{2.5}$$

where e is the elementary charge and x_p and x_n are the thicknesses of the depletion region into the p- and n-type regions respectively [15]. The one-dimensional electric field, $\mathbf{E}$, generated by this charge density can be calculated with Poisson's equation,

$$\nabla \cdot \mathbf{E} = \rho/\epsilon, \tag{2.6}$$

where ϵ is the permittivity of silicon. Using the boundary condition $E(-x_p) = E(+x_n) = 0$. The electric field can be described by,

$$E(x) = \begin{cases} -\frac{eN_A}{\epsilon_r \epsilon_0}(x + x_p), & \text{-}x_p < x < 0, \\ +\frac{eN_D}{\epsilon_r \epsilon_0}(x - x_n), & 0 < x < x_n, \end{cases} \tag{2.7}$$

which peaks at the pn boundary where $x = 0$, reducing to the neutrality condition given as,

$$N_A x_p = N_D x_n \text{ [15]}. \tag{2.8}$$

This condition implies that the more weakly doped half of silicon will have a thicker depletion region. The electric field is the gradient of the electric potential, $\phi(x)$, over the boundary, which can be written as,
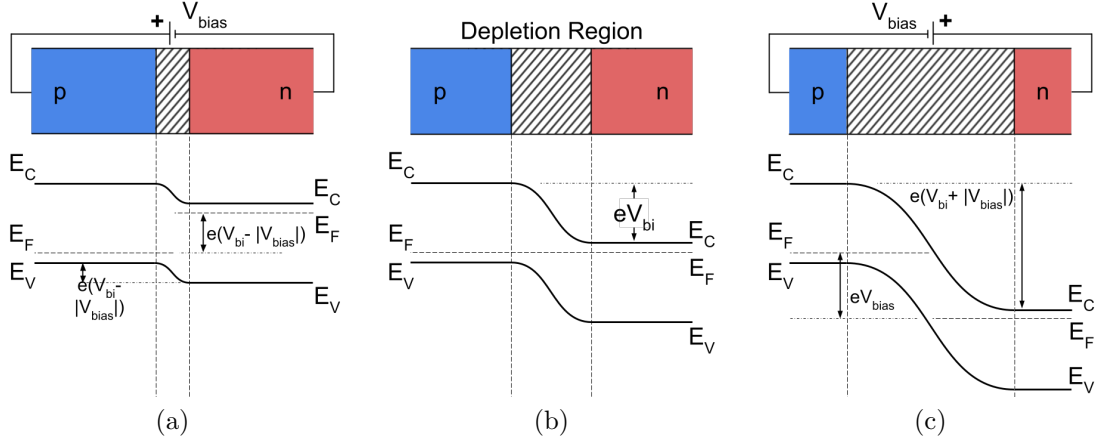
Figure 2.4: (a) Forward biased. Energy band bending is reduced and the depletion region shrinks. (b) Unbiased pn junction with equilibrium energy bands. (c) Reverse biased. Energy band bending is increased and the depletion region expands. Adapted from [16].

$$\phi(x) = \begin{cases} \phi(x_p) + \frac{eN_A}{2\epsilon_r\epsilon_0}(x + x_p)^2, & \text{-}x_p < x < 0, \\ \phi(x_n) - \frac{eN_D}{2\epsilon_r\epsilon_0}(x - x_n)^2, & 0 < x < x_n. \end{cases} \tag{2.9}$$

The built-in voltage, V_{bi}, is the electric potential difference between the ends of the space-charge region and is given as,

$$V_{bi} = \int_{-x_p}^{x_n} E(x)\, dx = \frac{e}{2\epsilon_r\epsilon_0} x_p^2 \frac{N_A}{N_D}(N_A + N_D) \; [15]. \tag{2.10}$$

This potential difference across the junction leads to the movement of minority charge carriers. This results in a drift current, I_{drift}, which opposes the diffusion current. The space-charge region reaches an equilibrium once there is no net current flow, such that $|I_{drift}| = |I_{diff}|$ [15]. At this stage the space-charge region is a fixed thickness and the built-in potential reaches a value of $\sim$0.6 V in the case of silicon.

2.1.3   Biased PN Junction

As will be discussed in subsequent sections, a thicker depletion region is desirable for a detector since more charge is collected. This is achieved by applying an external bias voltage, V_{bias}, across the junction. The effect of forward and reverse biasing on the depletion region and the band structure is shown in Figure 2.4.

A pn junction is reversed biased when a negative V_{bias} is applied to the p-type silicon with respect to the n-type side. This is in the same direction as the built in potential barrier and reduces the diffusion current, meaning that the system is no longer in equilibrium [15]. The flow of minority carriers, which are of very low density, is called leakage current. The external bias also attracts majority charge carriers away from the boundary which results in a thicker space-charge region. In contrast, in forward bias the barrier is decreased and the space-charge region shrinks. Figure 2.4 shows how the energy bands bend across the junction and the Fermi level is different on each side of the junction. The space-charge region can continue to grow with increasing reverse bias voltage until the entire volume is depleted. This is known as full depletion and is characterised by the full depletion voltage, V_{FD}. If the reverse bias voltage is raised further,

2.2   Particle Interactions With Matter

2.2.1   Photons

Photons, in the energy range of relevance to particle physics, usually interact through one of three primary mechanisms: photoelectric absorption, scattering or pair production [16]. The probability that each interaction occurs is governed by their cross-section which is dependent on the energy of the photon and the medium it traverses. The cross-section of each interaction in silicon is shown in Figure 2.5 as a function of energy. In this work, photons with an energy of a few $\sim$ eV up to a few
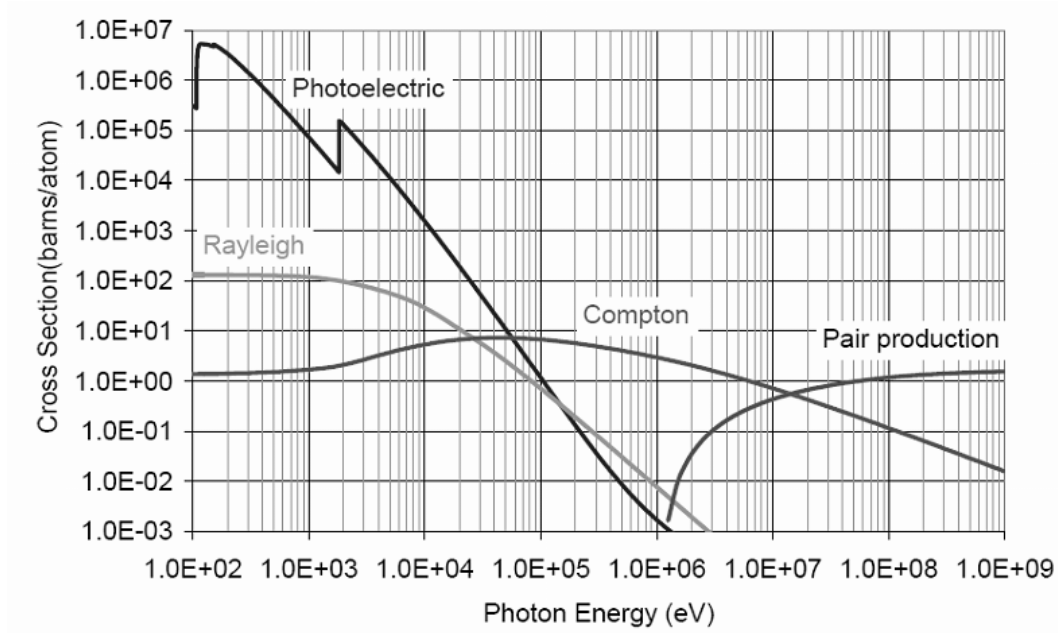
Figure 2.5: Photon interaction cross section as a function of photon energy. The photoelectric effect dominates for low energy photons, such as visible light. Taken from [17].

$\sim$ keV are used [16]. In these ranges, photoelectric absorption dominates.

Consider a beam of photons travelling through a material with mass density, ρ and atomic mass per mole, A. As the photons travels a distance, x, through the material, there is a chance that it will be absorbed. The probability is characterised by the absorption coefficient, ξ, which is defined as,

$$\xi = -\frac{1}{N}\frac{dN}{dx} = \rho\frac{N_A}{A}\sigma = n\sigma, \tag{2.11}$$

where N is the number of photons, $N_A = 6.022 \times 10^{23}$ is Avogadro's number, n is number density of the material and σ is the cross-section [16]. Equation 2.11 can be solved to show an exponential dependence on ξx. In a silicon crystal, when the photon energy exceeds the band-gap energy $(1.12\,\text{eV})$, photon absorption through electron-hole production can occur.

For photons with energies in the keV range, the dominant interaction is the pho-

toelectric effect [16]. The photon transfers all of its energy to an atom which then ejects an electron, leaving the atom ionised. This electron is free and unbound to the original material and has a surplus kinetic energy, $T = E_\gamma - E_B$. Here, E_γ is the energy of the photon which must be larger than the binding energy of the electron in the material, E_B, in order for this interaction to occur. The atom absorbs the recoil momentum, but the recoil energy is largely ignored due to the atoms relatively high mass.

The photoelectric cross-section is dependent on E_γ and Z, the proton number of the material. The cross-section rapidly decreases as E_γ increases while for a fixed energy it strongly increases with Z [16].

2.2.2 Charged Particles

Charged particles, in the energy range of relevance here, primarily interact through ionisation and excitation of atoms within a material. Any energy loss which is not from ionisation is known as non-ionising energy loss (NIEL) [16]. These are the dominant processes for low to medium energies. The average energy loss, dE, (stopping power) per path length, dx, for a particle with Lorentz factor γ and charge z can be described by the Bethe-Block formula:

$$-\left\langle \frac{dE}{dx} \right\rangle = K\frac{Z}{A}\rho\frac{z^2}{\beta^2}\left[\frac{1}{2}\ln\frac{2m_e c^2 \beta^2 \gamma^2 T_{max}}{I^2} - \beta^2 - \frac{\delta(\beta\gamma)}{2} - \frac{C(\beta\gamma, I)}{Z}\right], \quad (2.12)$$

where, $\beta = v/c$. Z and A are the atomic number and atomic mass of the material respectively [16]. c is the speed of light, ρ is the density of the target material and I is the mean excitation energy. δ is a density correction which applies to higher energies. C/Z is a shell correction for small β values. T_{max} is the maximum possible energy transfer. The constant K is defined as,
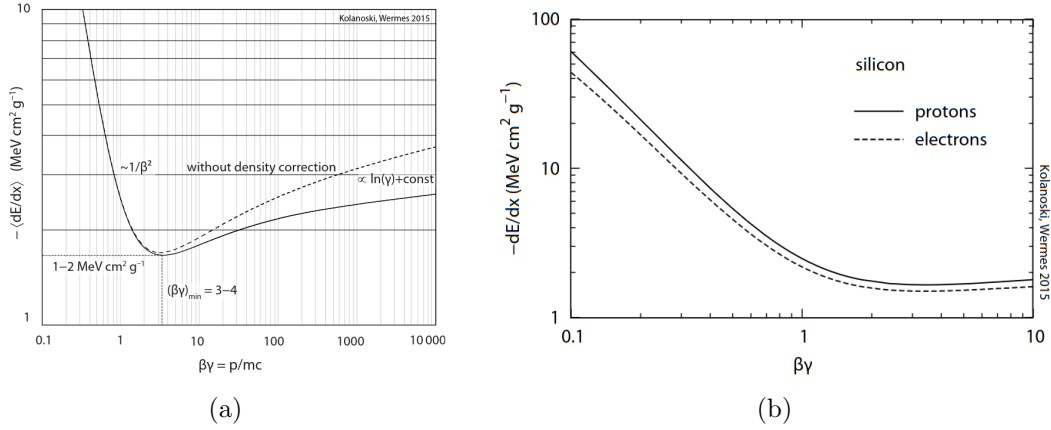
Figure 2.6: The Bethe-Bloch formula (a) as a function of a charged particle's momentum, for which $\beta\gamma$ is a measure and (b) showing the difference between protons and electrons in silicon, excluding bremsstrahlung. Taken from [16]

$$K = 4\pi N_A r_e^2 m_e c^2, \tag{2.13}$$

where m_e and r_e is the mass and classical radius of an electron respectively [16].

Figure 2.6 shows the Bethe-Bloch formula as a function of $\beta\gamma = p/mc$ for charged pions in silicon. For low energies, represented by the low $\beta\gamma$ region, the $1/\beta^2$ term dominates and results in a very large average energy loss per path length. A minimum forms at $\beta\gamma \sim 3$. Particles with an energy in this range and above are called minimum-ionising particles (mips) [16]. The average energy loss per unit path depends much more weakly on energy in this range, and thus it is often considered constant with respect to it. This is very helpful in particle physics tracking detectors since all highly relativistic charged particles can be considered approximately identical with respect to how they deposit energy within a detector.

While Equation 2.12 is valid for most heavy particles, positrons and electrons need to be treated slightly differently. Figure 2.6 shows how the ionisation losses from electrons differ from that of a proton in silicon. At the minimum, the electron's average energy loss is 10 % less than that of the proton and the difference increases either side of the minimum [16]. Electrons also lose energy from Bremsstrahlung
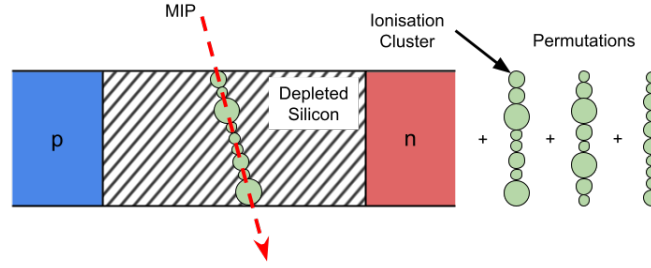
Figure 2.7: Visualisation of Landau fluctuations. Assuming there is one packet of energy deposition per μm, then for a $50\,\mu$m sensor there are 50 separate energy depositions. This gives 50! permutations leading to more Landau fluctuations. Adapted from [18]

which is when the electron interacts with the intense Coulomb field near the atomic nucleus, losing energy/momentum and producing a high energy photon [16]. This dominates energy loss by ionisation above the critical energy [16].

2.2.2.1   Landau Fluctuations

The Bethe-Block formula only describes the average energy loss. In reality, energy loss is stochastic and discrete [16]. This means that a number, N, of energy loss events (ionisation or excitation processes) occurs across a distance Δx. The total energy loss, ΔE can be described as the summation of these individual energy loss events, δE_n such that,

$$\Delta E = \sum_{n=1}^{N} \delta E_n \ [16].$$ (2.14)

The energy lost in individual events, E_n, also exhibits statistical fluctuations. This is well visualised in Figure 2.7. Combined with Equation 2.14, the net effect is called Landau fluctuations.

Assuming that the contributions of δE_n are independent from each other, then as N tends towards infinity, the distribution of ΔE tends towards a Gaussian as per the central limit theorem [16]. Therefore, the distribution has a Gaussian part, but

also contains a tail of high energy loss. This corresponds to hard collisions where much of the energy is transferred to an electron which is released as a delta ray [16].

This distribution is known as the Landau distribution and its exact form is parameterised with the ratio, κ, of mean energy loss, ξ, and the maximum energy loss, T_{max} [16]. For large κ, the distribution becomes a symmetric Gaussian. Conversely for small κ, it is very asymmetric. The distribution itself is shown in Figure 2.8. It is parameterised by λ which is defined as,

$$\lambda = \lambda(\Delta E_\omega, \xi) = \frac{\Delta E - \Delta E_\omega}{\xi} - 0.22278, \tag{2.15}$$

where ΔE_ω is the most probable energy loss, occurring at the maximum of the distribution where $\lambda = -0.222\,78$ [16]. The most probable energy loss tends to be a better metric than the average which is influenced by the long tail of the distribution.

The shape of the Landau is also directly affected by the thickness of silicon of which a particle traverses. As the thickness increases, the peak of the distribution shifts towards a higher energy loss per unit length. The width of the distribution also shrinks, representing a trend towards a Gaussian [19]. The detectors discussed in this work are thin enough ($\sim 50\,\mu$m) that the shape is predominantly Landau.

As will be discussed in Section 2.5.1, Landau fluctuations limit the time resolution of detectors. However, they also limit the spatial resolution through irregular charge distributions, making it harder to interpret non-perpendicular tracks [20].

2.3   Silicon Detectors

The principle of the reverse bias pn junction has been used to develop the silicon detectors now used for charged particle tracking in experiments like ATLAS and CMS [5, 6]. This section will explore the real-world sensor geometries as well as how a signal is formed and processed once a charged particle deposits energy in silicon.
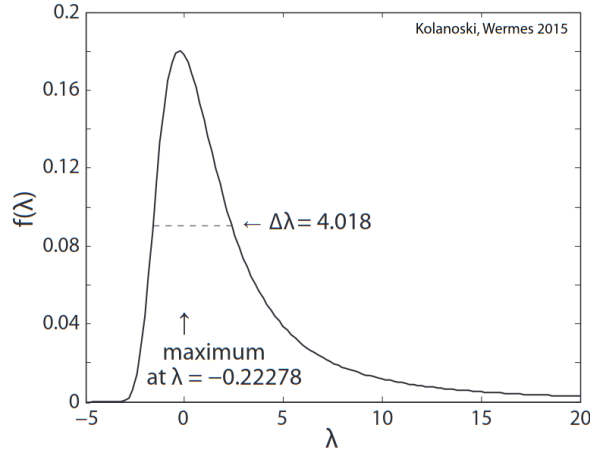
Figure 2.8: The generalised Landau distribution. Taken from [16]

2.3.1   Detector Structures

Semiconductor detectors operate in reverse bias in order to obtain a large depletion volume. This is necessary for sufficient charge to be collected from a mip, which deposits energy along its path. The most common sensor geometry is planar which is shown in Figure 2.9. The advantage of planar geometry is a uniform and strong electric field providing fast and efficient charge collection. For particle trackers, planar sensors are usually a p in n-bulk sensor, which means that the majority of the silicon is n-type, while a thin layer at the top is p-type [16]. The next generation of particle tracking detectors for the HL-LHC will be n in p-bulk to avoid the problem of n-type inverting to p-type (type inversion) from irradiation as discussed in Section 2.4.2.3.

For p-bulk detectors, in order to achieve a planar geometry, the doping concentration of the n-type needs to be significantly larger than that of the p-type. This reflects Equation 2.8, the neutrality condition. Typical values are shown in Table 2.1. In this example, the contribution of the n-type to the depletion depth can largely be ignored such that $x_p \gg x_n$. Hence, the depletion depth is often quoted as a total depletion depth, d, defined as,

$$d \approx x_p \approx \sqrt{\frac{2\epsilon_r \epsilon_0}{e} \frac{1}{N_D} (V_{bi} + V_{bias})} [16].$$

(2.16)

Parameter	Value
N_D	$10 \times 10^{19} \, \text{cm}^{-3}$
N_A	$2.3 \times 10^{12} \, \text{cm}^{-3}$
x_n	$4 \times 10^{-6} \, \mu\text{m}$
x_p	$\sim 20 \, \mu\text{m}$
V_{bi}	$\sim 0.4 \, \text{V}$ to $0.8 \, \text{V}$

Table 2.1: Values associated with a typical unbiased planar n in p-bulk sensor. Values are comparable to those of p in n-bulk sensors. Taken from [16]

Another advantage of planar geometry is that the thin region of n-type silicon can be structured in a specific way. These structures can be designed to segment the silicon into either pixels or strips, with a pitch between $50 \, \mu\text{m}$ and $200 \, \mu\text{m}$, as shown in Figure 2.9 [16]. Pixels and strips can be read-out individually, allowing the position of an impinging particle to be determined in 2D and 1D respectively. Very large pixels ($\sim 1 \, \text{mm}$) instead tend to be referred to as pads.

Additional structures can be implanted to optimise the detector. The p-stop is a p+ implant (opposite to the pixel/strip n-type) placed around pixels or between strips. It is required in order to ensure segmentation between neighbouring pixels in an array, which would otherwise short together through trapped positive charge at the silicon-oxide interface [18]. Trapped charge can create an inversion layer in p-type which will result in a conductive channel between n-implants [15]. Conversely with n-bulk, the positive trapped charge enhances the isolation between p-implants.

Figure 2.9 also shows a metal ring around the array of pixels/strips, called the guard ring. Metal contacts and an n-type implant help to define the boundary of the sensing area, shaping the field profile at the edge. If not floating, the guard ring also helps to sink any edge current coming, in particular, from any dicing induced damage.

At the bottom of each sensor in Figure 2.9 is a thin ($\sim 1 \, \mu\text{m}$) region of highly doped p-type silicon. This is called the ohmic contact and allows over-depletion. This is when the reverse bias voltage is increased past the point at which the bulk is fully depleted, increasing the electric field strength in the sensor [16].

(a) pn area diode.

(b) Area diode with guard ring.

(c) Pad or pixel detector.

(d) Strip detector.
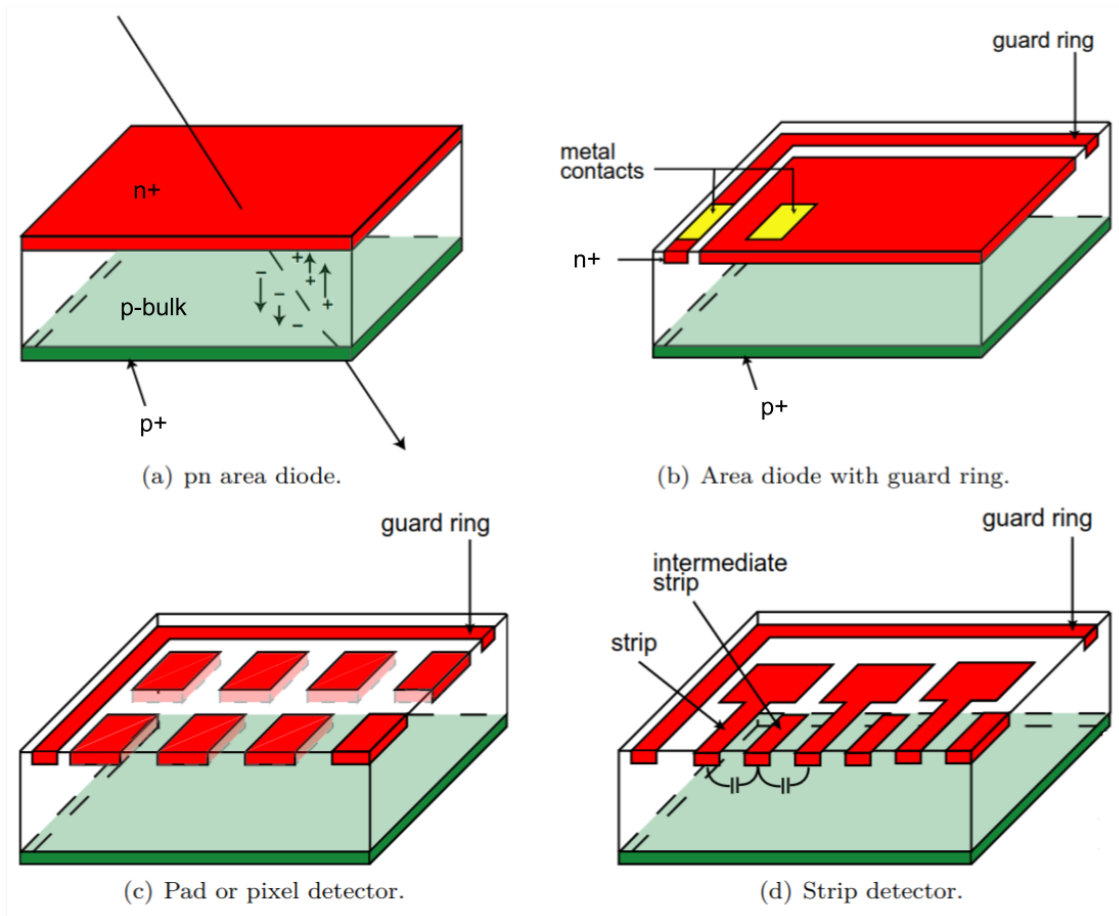
Figure 2.9: Example sensor geometries used for particle trackers. n in p-bulk are used in high-radiation environments such as the LHC. Adapted from [16].

2.3.2   Signal Formation

2.3.2.1   Shockley-Ramo's theorem

As discussed in the previous section, mips lose energy as they traverse through silicon. The energy loss results in energy deposited into the silicon. If the energy loss is through excitation, then a valence electron can be promoted to the conduction band, generating an electron-hole pair [16]. If the energy loss is through ionisation, an electron is released which typically also interacts within the silicon, resulting in further ionisation and electron-hole pair generation [16].

Once an electron-hole pair has been generated, it begins to separate under the electric field, producing a current. The Shockley-Ramo's equation describes the induced current, $i(t)$, as a function of time, t, as,

$$i(t) = -q\mathbf{v} \cdot \boldsymbol{\xi_W}, \tag{2.17}$$

where q is the charge, $\mathbf{v}$ is the drift velocity and $\boldsymbol{\xi_W}$ is the weighting field [18].

Electrons and holes travel in opposite directions but since their charges are opposite, their current contributions, $i_e(t)$ and $i_h(t)$ respectively, combine. The integral of their combined current is equal to the total charge generated, q, such that,

$$\int (i_e(t) + i_h(t))dt = q \text{ [18]}. \tag{2.18}$$

The relative contributions from each is dependent on the electrode geometry and applied electric field through the weighting field [18]. The weighting field is a coupling which describes how a charge within the sensor is seen by the read-out electrode [18]. For example, if $\xi_W = 0.5$ at a particular position, then a charge, q, is seen as $0.5q$ by the read-out electrode. The weighting field is calculated by holding all electrodes to ground, except for the collecting electrode. For the simple 1D case,

the weighting field has the same dependence upon the geometry of the electrodes as the electric field.

For large parallel electrodes with a separation d, the weighting field is constant and depends on $1/d$ [18]. This means that for a uniform initial distribution of charge, such as for a mip, the induced current signal depends only on $\mathbf{v}$ and d. This is shown in Figure 2.10 where large implants lead to a linear current profile over time. Holes and electrons contribute to the signal at different rates due to the difference in their mobility, but their integrated contributions are the same.

Now consider small electrodes with a larger pitch, such as for a strip sensor. The weighting field now has a dependence of $1/x$ where x is the distance from the electrode [18]. As a result, the signal is mostly generated when charges are very close to the electrode [18]. This means that electrons contribute more than holes overall. This is shown in Figure 2.10 where the area under the curve is larger for electrons than holes.

2.3.2.2   Traps

As charges drift through a silicon sensor, they can become trapped [21]. Traps are defects in the silicon lattice which create additional energy levels. They can arise during manufacturing where impurities or structural defects are introduced. Irradiation can also cause structural defects and therefore additional traps. If charges are trapped, they stop drifting, reducing their current contribution to zero [21]. This results in the loss of charge collection efficiency (CCE), which is defined as the charge collected as a percentage of the charge generated from the initial ionisation [18]. The probability that a charge is trapped is proportional to the distance travelled, and thus for N charges travelling a distance dl, the number of trapped carries is given as,

$$dN = -N \frac{1}{\lambda_{eff_{e,h}}} dl, \tag{2.19}$$
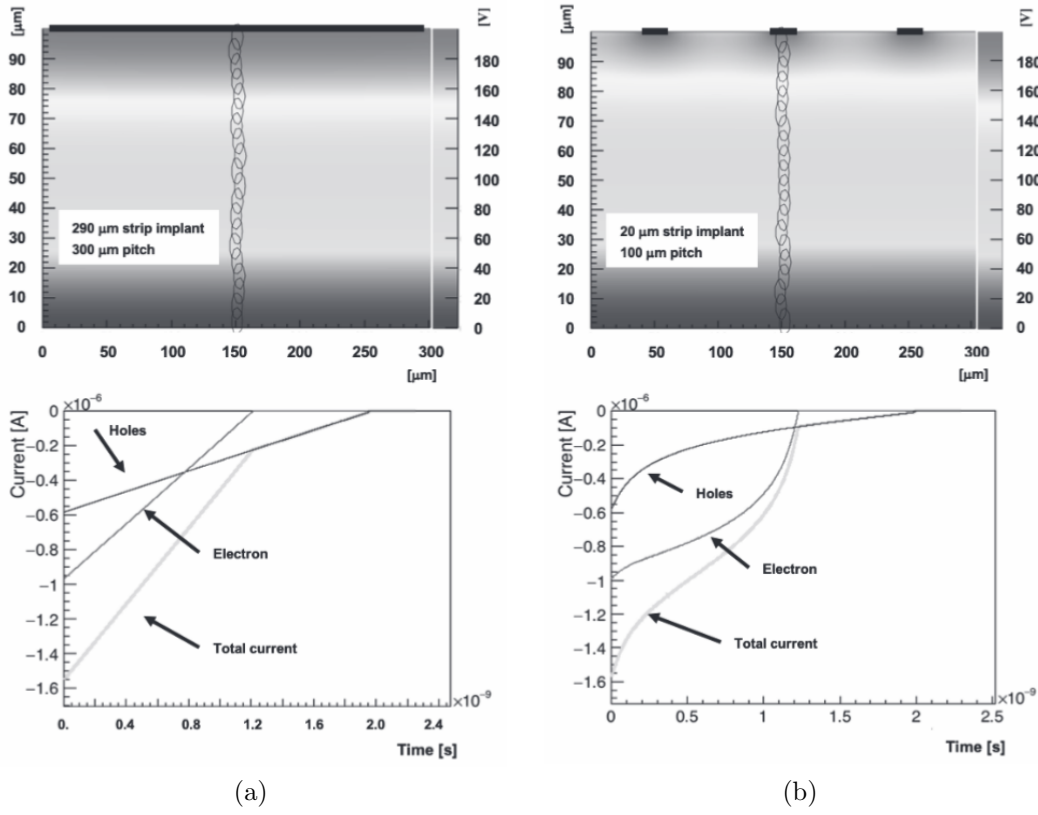
Figure 2.10: Weighting field and induced current in two configurations of a n-in-p silicon sensor: (a) Wide electrodes (b) Narrow electrodes. The overlapping ovals represent the initial discrete energy deposition of a mip as it traverses through the sensor. Taken from [18]
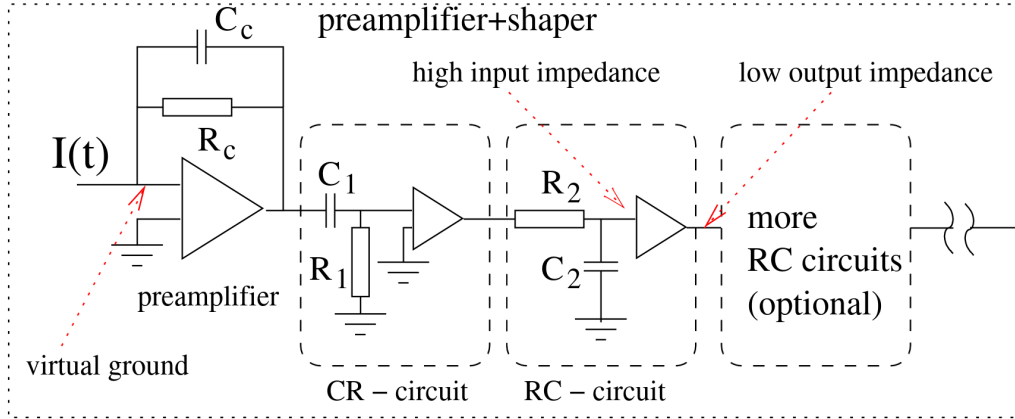
Figure 2.11: The full read-out chain of a silicon sensor. A charge sensitive preamplifier (CSP) converts the current into a voltage. This is followed by a CR circuit and number of RC circuits which shape the signal. Adapted from [21]

where $\lambda_{eff_{e,h}}$ is the effective carrier trapping distance [21]. Ideally, the trapping distance should be significantly larger than the thickness of the detector, keeping the CCE as high as possible.

2.3.3   Signal Processing

The induced current needs to be integrated to give the total collected charge and amplified so that it can be readout and recorded digitally. This is usually done with a charge sensitive preamplifier (CSP) and a shaping amplifier in a chain, similar to the arrangement shown in Figure 2.11, which produces a voltage proportional to the collected charge [22].

The equivalent circuit of a CSP is a resistor, R, and capacitor, C, in parallel. The output voltage, $U_{out}^{pre}(t)$, of the CSP can be shown to be,

$$U_{out}^{pre}(t) = \frac{1}{C_{pre}} e^{\frac{-t}{\tau_{pre}}} \int_{-\infty}^{t} I(t') e^{\frac{t'}{\tau_{pre}}} \, dt', \tag{2.20}$$

where $\tau_{pre} = R_c C_c$ is the time constant and $I(t')$ is the induced current as a function of time [21].
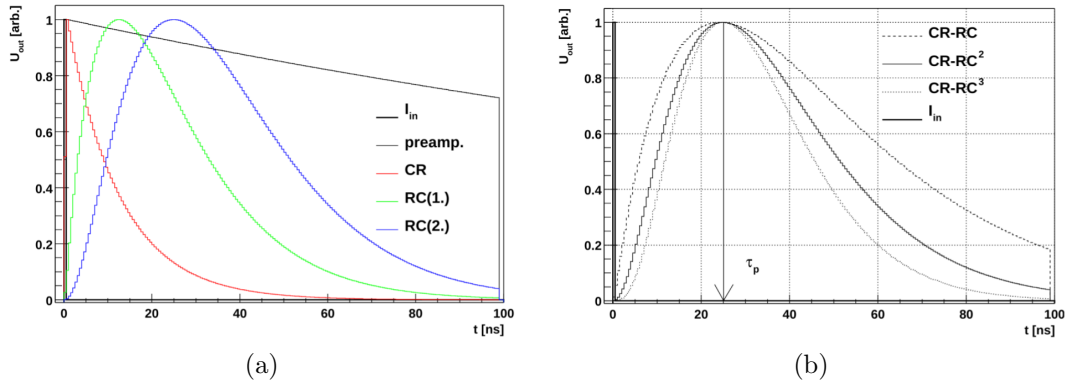
Figure 2.12: (a) The result of each stage of the read-out chain. (b) The effect of multiple RC circuits. Taken from [21]

After integration, the signal is shaped using an differentiator (CR) and a number of integrator (RC) circuits in combination with an operational-amplifier (op-amp). This shaping amplifier is designed to give a good signal to noise ratio (SNR) as well as minimising overlap between consecutive pulses [22]. The voltage response from a CSP and CR-RC^n chain, where n is the number of RC circuits in the chain, is shown in Figure 2.12.

The final voltage output from the amplifier chain can then be interrogated. There are multiple ways to do this, two of which are a discriminator or a readout stage. The discriminator checks whether the voltage is above a certain threshold and records the answer as a binary hit or no hit [22]. Alternatively a digitiser converts the entire waveform, $U_{out}^{RC}(t)$, to a binary string that can then be processed later [22]. A modern oscilloscope is an example of a digitiser.

2.3.4 Noise

Noise is considered to be any unwanted signal introduced into the system. It limits the ability to accurately measure the true signal due to an impinging particle. This is characterised by the SNR. There are various sources of noise which sum in quadrature for the total noise of the system [23]. Thermal noise arises from thermal fluctuations which lead to random electron velocity fluctuations. Low frequency, or

$1/f$, noise has a power spectrum inversely proportional to frequency [22]. This is due to the recombination of electron-hole pairs caused, for example, by traps in the bulk [22]. Finally, one of the more important sources is shot noise which is caused by statistical fluctuations in the integrated leakage current [22]. This form of noise increases after-irradiation and with internal gain (discussed in Section 2.5.3). Read-out electronics have their own source of noise which contributes in ways that also depend on the detector capacitance and the operating temperature. It is important to ensure that all sources of noise are minimised in order to maximise the SNR.

2.4   Radiation Damage in Silicon Detectors

Radiation degrades the performance of silicon sensors. At the HL-LHC, irradiation fluences of order $\sim 1 \times 10^{16}\,1\,\mathrm{MeV}\,\mathrm{n_{eq}/cm^2}$[2] are to be expected in the vertexing layers of the ATLAS and CMS [5, 6, 24]. Developing radiation-hard sensors requires an understanding of the mechanisms through which radiation damages silicon sensors.

There are two main types of radiation damage: surface and bulk. Surface damage is primarily a build-up of charge at silicon-oxide interfaces (or interfaces to other insulating layers) near the sensor surface [21]. Bulk damage is related to defects in the silicon lattice, which manifest as trapping centres and contribute mid-band-gap energy levels [21]. This section will summarise the origins and effects of bulk damage on sensor performance.

2.4.1   Generation of Defects

Bulk damage manifests in the form of defects [21]. Defects can take various forms, with the four most common simple defects shown in Figure 2.13. A vacancy (Figure 2.13a), labelled V, is a lattice with a missing silicon atom. An interstitial atom, labelled as I, is either a silicon or impurity atom (Figure 2.13b and Figure 2.13c

---

[2]Equivalent damage caused by $1\,\mathrm{MeV}$ neutrons of the same fluence.
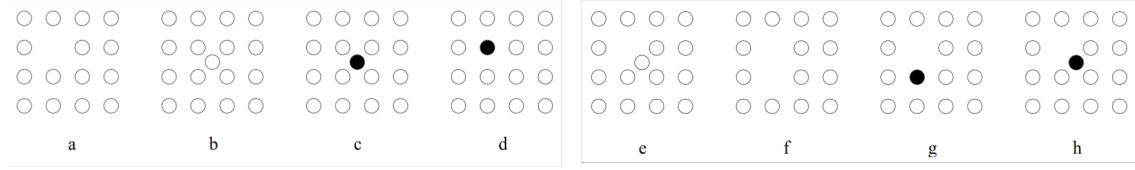
Figure 2.13: Simple defects: (a) vacancy, V, (b) interstitial silicon atom, (c) interstitial impurity atom, I, (d) substitutional impurity atom. Defect complexes: (e) close pair I-V (f) divacancy, V-V, (g) substitutional impurity and vacancy, VP complex, (h) interstitial impurity and vacancy, VO complex. Taken from [21]

respectively), which resides in the gaps between the regular silicon atoms. Lastly, there is the substitutional atom (Figure 2.13d), where an impurity fills a vacancy. Note that an interstitial usually accompanies a vacancy, known as a Frenkel pair. Some Frenkel pairs recombine while others separate, migrating around the lattice [21]. A combination of two or more simple defect types is known as a defect complex. An example of four complexes is shown in Figure 2.13e to Figure 2.13h.

Non-ionising energy lost to the lattice can produce simple defects if the deposited energy exceeds a threshold energy, $E_d \sim 25\,\text{eV}$[21]. These defects migrate and some, such as Frenkel pairs, recombine with no damage caused. Others will combine to form defect complexes, similar to the examples shown in Figure 2.13. If the energy deposited is significantly larger than $25\,\text{eV}$, the displaced atom can cause further ionisation and displacements, resulting in defect clusters [21].

The NIEL scaling hypothesis poses that the effects from different particles and energies can be scaled with respect to a reference NIEL value. The convention is to make a unit conversion of the fluence to $1\,\text{MeV}\,\text{n}_{\text{eq}}/\text{cm}^2$, which means equivalence to the damage caused by $1\,\text{MeV}$ neutrons [18]. Figure 2.14 shows the relative NIEL factor for neutrons, protons, pions and electrons as a function of energy. Note that low energy protons have a much higher NIEL factor than low energy neutrons, but they both converge for energies above $50\,\text{MeV}$.
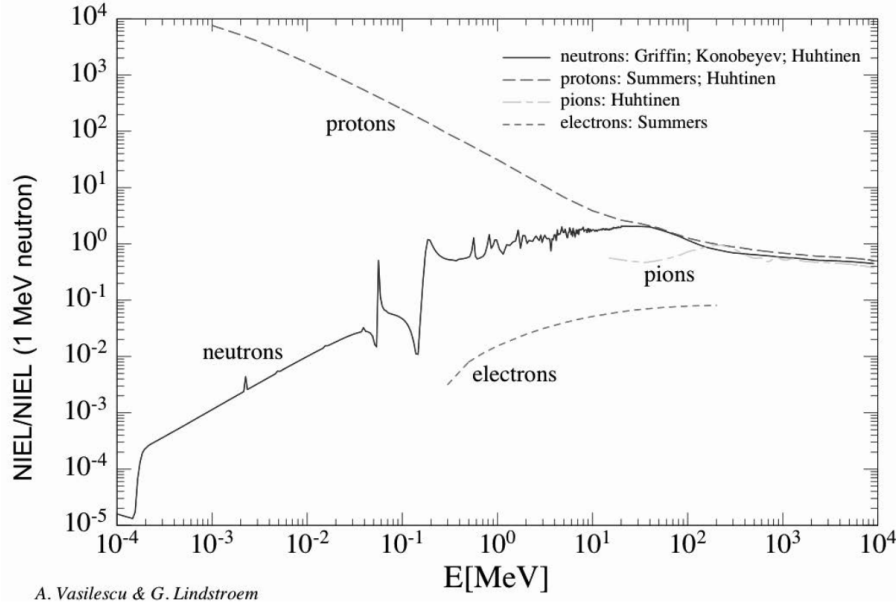
Figure 2.14: The relative NIEL factor for a selection of particles as a function of their energy. Taken from [18]

2.4.2   Impact on Silicon Detectors

2.4.2.1   Leakage Current

The first way a silicon sensor is affected by irradiation is the increase of leakage current, denoted as I_{bulk}. This is due to electron-hole pair generation centers in the middle of the band gap caused by defects. Pre-irradiation, the leakage current is given by,

$$I_{bulk} = Adq_0 \frac{n_i}{\tau_g}, \tag{2.21}$$

where A is the electrode area, d is the depletion depth, n_i is the intrinsic carrier density, q_0 is the electron charge and τ_g is the generation lifetime [18]. Post-irradiation, the change in leakage current, ΔI, is given as,

$$\Delta I = \alpha A d \Phi_{eq}, \tag{2.22}$$

where Φ_{eq} is the $1\,\mathrm{MeV}\,\mathrm{n_{eq}/cm^2}$ fluence and α is the current-related damage constant [18]. $I_{bulk}(\Phi_{eq})$ can be approximated to ΔI since $\Delta I \gg I_{bulk}$ for fluences above $1 \times 10^{12}\,1\,\mathrm{MeV}\,\mathrm{n_{eq}/cm^2}$ [18].

For sensors with internal gain, the bulk current is multiplied by the gain, G, and becomes a significant concern. However, current has a temperature, $T[K]$, dependence,

$$i(T) \propto T^2 \exp\left(\frac{1.2\ eV}{2\ k_B T}\right), \tag{2.23}$$

where k_B is Boltzmann's constant [18]. This temperature dependence can be exploited. For example, ATLAS and CMS intend to operate at $-30\,^{\circ}\mathrm{C}$ which results in two orders of magnitude drop in current compared to room temperature [5, 6]. At the same time, this helps to reduce the shot noise which increases due to high leakage current. For a given bandwidth interval, Δf, this is described by,

$$\sigma_{shot} = \sqrt{2q(I_{surface} + I_{bulk}G^2 F)\Delta f}, \tag{2.24}$$

where F is the excess noise factor and $I_{surface}$ is the surface leakage current [18].

2.4.2.2 Charge Collection Efficiency

Another consequence of irradiation is the generation of trapping centres as briefly discussed in Section 2.3.2.2. These deep defects trap charge carriers as they drift and stop them from contributing to the induced current signal [18]. This results in the loss of CCE [18]. The current signal, $i(t)$, induced by a charge moving in an irradiated sensor is given as,

$$i(t) = i(t)_{\Phi_{eq}=0} \cdot e^{-t/\tau_{eff}}, \tag{2.25}$$

which is an exponential function of time [18]. This is because the probability of trapping a charge carrier depends on the duration of the drift. τ_{eff} is the time constant and is defined as,

$$\frac{1}{\tau_{eff}} = \Phi_{eq}\beta_{e/h}, \tag{2.26}$$

where $\beta_{e/h}$ is the effective trapping damage constant [18]. If τ_{eff} is not significantly larger than the time it takes for electron-hole pairs to drift to their respective electrodes, from Equation 2.18, charge will be lost and thus the CCE reduced. Similarly, if the electronics shaping time is short compared to the drift time, signal will still be lost if τ_{eff} is less than that shaping time [18]. A correction to Equation 2.26 is required for fluences on the order of $1 \times 10^{15} \, 1 \, \text{MeV} \, \text{n}_{eq}/\text{cm}^2$, such that,

$$\tau_{eff} = 540 \cdot \Phi_{eq}^{-0.62} \, \text{ps}, \tag{2.27}$$

where Φ_{eq} is given in units of $1 \times 10^{15} \, 1 \, \text{MeV} \, \text{n}_{eq}/\text{cm}^2$ [18].

2.4.2.3   Doping Density

Radiation can also change the doping concentration of silicon through three different mechanisms. These are summarised by the effective doping density, N_{eff}, which is given by,

$$N_{eff} = N_{D_0}e^{-c_D\Phi_{eq}} - N_{A_0}e^{-c_A\Phi_{eq}} + g_{eff}\Phi_{eq}, \tag{2.28}$$

where N_{D_0/A_0} is the initial donor/acceptor concentration and $c_{D/A}$ is the donor/acceptor removal coefficient [18]. They represent an overall reduction in donor/acceptor concentration. $g_{eff} = 0.02 \, \text{cm}^{-1}$ is the coefficient for acceptor-like defect creation which represents an increase in the acceptor concentration [18]. For near intrinsic materi-

als, this can lead to type inversion from n-type to p-type.

2.4.2.4 Acceptor Removal

Acceptor removal is particularly important for sensors with internal gain, such as those studied in this work. This is because internal gain can be achieved through a highly doped region of acceptors. Irradiation fluences up to $1 \times 10^{16} \, 1 \, \mathrm{MeV} \, \mathrm{n_{eq}/cm^2}$ can be sufficient to remove most of the internal gain. However, secondary ion mass spectrometry (SIMS) measurements have shown that the boron atoms are merely deactivated rather than removed entirely [18]. The potential mechanisms and defects behind acceptor removal are discussed in [25]. One such mechanism begins with the Watkins replacement mechanism, which sees silicon intertitials, Si_i replace boron substitutionals, B_s [26]. This boron interstitial then interacts with other interstitial impurities such as oxygen, O_i, such that,

$$Si_i + B_s + O_i \rightarrow B_i + O_i \rightarrow B_iO_i, \tag{2.29}$$

forming B_iO_i complexes which replace an acceptor level with a donor level [18]. As touched upon in Equation 2.28, acceptor removal can be characterised as,

$$N_A(\Phi_{eq}) = N_A(0)e^{-c_A\Phi_{eq}}. \tag{2.30}$$

The acceptor removal coefficient, c_A can also be written as,

$$c_A = \frac{1}{\rho_0 \, eq} = \frac{\phi_{Si} \cdot k_{cap} \cdot N_{Int} \cdot \sigma_{Si}}{0.63 \cdot N_A(0)} \cdot \frac{1}{1 + (\frac{N_{Ao}}{N_A(0)})^{2/3}}, \tag{2.31}$$

where, $\rho_{Si} = 5 \times 10^{22} \, \mathrm{cm^{-3}}$ is the atomic density of silicon, $0.63 \cdot N_A(0)$ is the density of removed acceptors and σ_{Si} is the cross-section between the radiation and silicon [18]. N_{Int} is the number of defects created in the interaction. The capture coefficient,

k_{cap} is the probability for a defect to capture an acceptor which is dependent on the doping of the gain layer and the presence of other impurities such as carbon or oxygen [18]. The second term represents a correction required for low acceptor densities since the distance between defects and acceptors can be too large. Here, N_{Ao} is a fit parameter to be found through experimentation.

2.5   Ultra Fast Silicon Detectors

With the advent of the HL-LHC, high spatial resolution alone is no longer sufficient for regions of dense particle tracks. Instead, additional timing information is required to enable track reconstruction and separation primary vertices within a given bunch crossing. This section will discuss the different components which limit the time resolution of a detector and how they are optimised.

One design which achieves this is the LGAD. Alternative designs, such as ac-coupled LGADs (AC-LGADs) and inverse-LGADs (iLGADs), are aimed at improving the spatial resolution (currently $\sim 1300\,\mu$m pitch) to better match the specification of the current particle trackers ($< 100\,\mu$m pitch). Meanwhile, carbon-enriched LGADs and 3D sensors also approach the radiation-hard requirements of the HL-LHC. While only the basic design of an LGAD is the focus of this work, information on alternative designs can be found at [18, 27, 28].

2.5.1   Time Resolution

The total time resolution, σ_t of a silicon detector is made up from a contribution of multiple sources. These contributions can be treated as independent and be summed in quadrature such that,

$$\sigma_t^2 = \sigma_{jitter}^2 + \sigma_{Landau}^2 + \sigma_{timewalk}^2 + \sigma_{distortion}^2 + \sigma_{TDC}^2 \text{ [18]}. \tag{2.32}$$
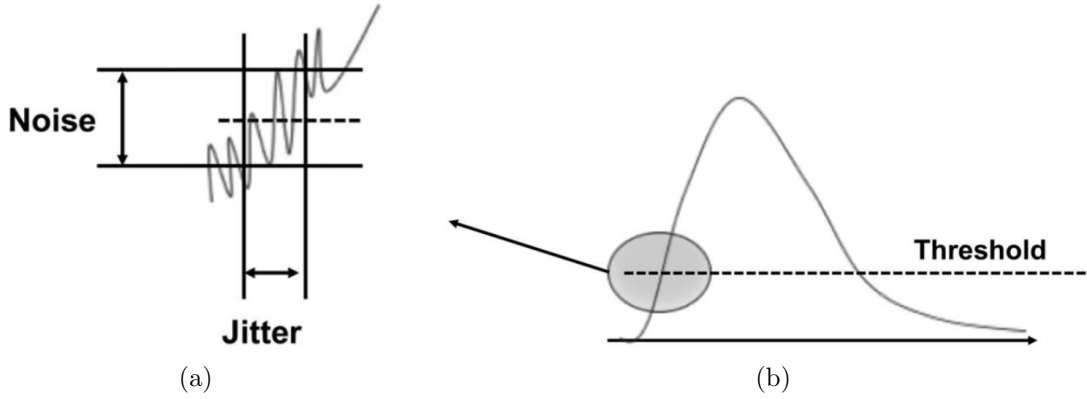
Figure 2.15: Jitter creates unwanted variation in the signal which can allow it to cross a threshold early or later than it normally would. Adapted from [18]

Jitter, σ_{jitter}, is when noise, N, causes a signal, S, to appear to arrive earlier or later than it actually does, as shown in Figure 2.15. The contribution to the time resolution can be characterised as,

$$\sigma_{jitter} = \frac{N}{dV/dt} \approx \frac{t_r}{S/N}, \tag{2.33}$$

where dV/dt is the slew rate and t_r is the rise time [18]. This approximation assumes that the slew rate can be approximated to S/t_r. In order to minimise σ_{jitter}, a sensor and its electronics should minimise the rise time, while maximising the SNR.

Landau noise and timewalk are the results of the discrete and stochastic nature of electron-hole pair generation, called Landau fluctuations. Landau noise, σ_{Landau} is the non-uniform current signals caused by irregular ionization from charged particles [18]. Figure 2.7 in Section 2.2.2.1 demonstrates the many possible permutations of energy deposition as a mip traverses through silicon. This leads to irregular contributions to the total signal which, similar to jitter, limit the time resolution. σ_{Landau} is a fundamental limit which can only be reduced with a smaller active thickness. For example, reducing from $300\,\mu$m to $50\,\mu$m reduces σ_{Landau} from $60\,$ps to $25\,$ps [18].

Time walk is the apparent change in time of arrival as a function of signal amplitude
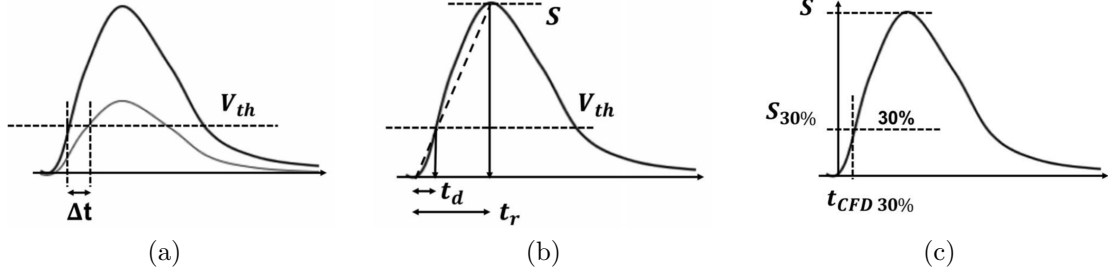
Figure 2.16: (a) Larger signals appear to arrive earlier than smaller signals. (b) A linear approximation can be used to estimate the arrival time. (c) A 30 % constant fraction discriminator (CFD) is used to estimate the time of arrival and minimise the effects of timewalk. Taken from [18]

for a given threshold [18]. Figure 2.16 demonstrates how a larger signal arrives earlier than a smaller signal. Using a linear approximation, the signal crosses a fixed threshold V_{th} with a time delay t_d, relative to an arbitrary zero time. Geometrically, one can write the ratios $t_d/t_r = V_{th}/S$ allowing t_d to be calculated. The time resolution contribution from time walk is defined as the root mean square (RMS) of t_d such that,

$$\sigma_{timewalk} = [t_d]_{RMS} = \left[\frac{V_{th}}{S/t_r} \right]_{RMS} \propto \left[\frac{N}{dV/dt} \right]_{RMS} [18]. \qquad (2.34)$$

The linear approximation allows the relation, $S/t_r = dV/dt$ to be used here. It is also common to express V_{th} as a multiple of the noise, N. In this work, a CFD is used to minimise the time walk contribution. The threshold, V_{th}, changes with each pulse and is a constant fraction of the signal amplitude, between 10 % and 30 %. This assumes the signal does not distort as the amplitude scales. Since timewalk can be almost completely corrected, it is usually ignored.

Distortion, $\sigma_{distortion}$, comes from changes in the current signal due to changes in the initial ionisation position within the sensor [18]. The Ramo-Shockley theorem states that $i \propto qv_d\xi_W$. To ensure uniformity in i, it is required that the drift velocity, v_d, and the weighting field, ξ_W, are as uniform as possible perpendicular to the implant. Planar geometry, discussed in Section 2.3.1, provides a uniform
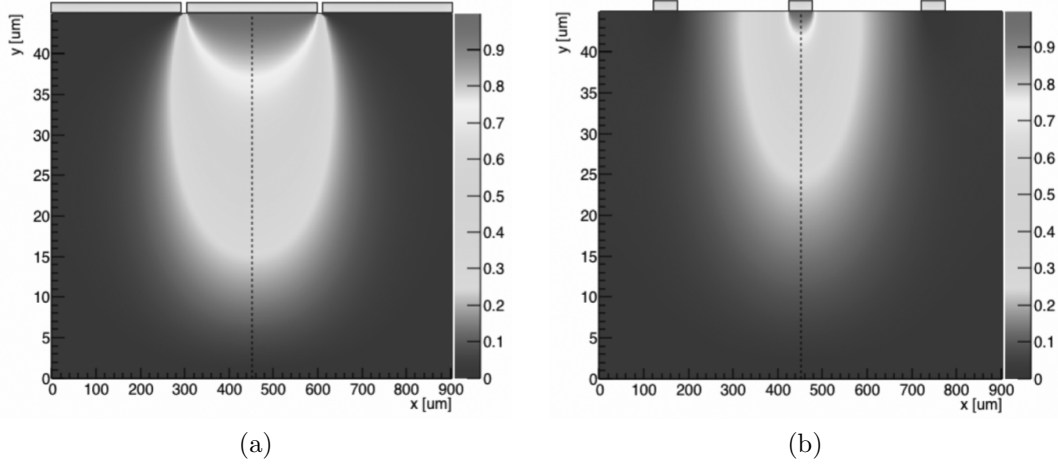
Figure 2.17: Weighting fields for a $300\,\mu$m pitch sensor with an electrode width of (a) $290\,\mu$m and (b) $50\,\mu$m. Taken from [18]

electric field and hence a uniform weighting field. This is shown in Figure 2.17 which shows the weighting field for a narrow and wide n-type region. The latter helps to ensure that the weighting field is uniform for as much of the pixel volume as possible. Saturating v_d also increases uniformity. For a sufficiently high bias voltage, the electric field strength can be in excess of $3\,\text{V}/\mu$m. In this limit, electron drift velocity saturates which naturally minimises variability in v_d. Note that the hole drift velocity only saturates at the limit of $10\,\text{V}/\mu$m which is only achieved at very high bias voltages. Uniformity in v_d and ξ_W means that $\sigma_{distortion}$ is small and therefore neglected.

A time-to-digital convertor (TDC) has a finite bin width, ΔT, which contributes to the time resolution as $\Delta T/\sqrt{12}$ [18]. Most TDCs used in HEP experiments have very fine binning, $25\,$ps for example, and thus their contribution, σ_{TDC}, to the time resolution can often be ignored.

2.5.2   Low Gain Avalanche Detector

The LGAD is a novel sensor design which aims to exploit Equation 2.32 and achieve a pre-irradiation time resolution of $\sim\!35\,$ps [18]. This corresponds to $25\,$ps from
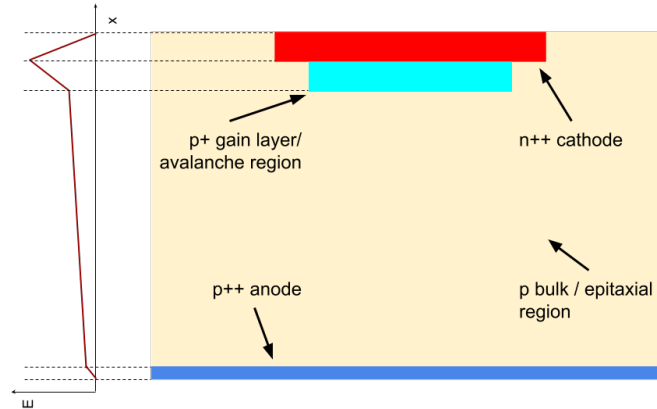
Figure 2.18: Simplest structure of an LGAD. The p+ gain layer creates a peak in the electric field which causes impact ionisation, leading to an intrinsic gain. Adapted from [29, 8]

the Landau noise floor which is minimised with a thin active thickness of typically $50\,\mu$m. The rest mostly comes from jitter which, while minimised, is impossible to eliminate completely. An issue with $50\,\mu$m thick sensors is that 6 times less charge is generated compared to $300\,\mu$m thick sensors, which leads to a small SNR. LGADs overcome this issue with intrinsic gain which increases the signal above the electronic noise floor [18]. A p+ region is implanted just below the n-type silicon as shown in Figure 2.18. This region, called the gain layer, is more highly doped than the p-bulk. As a result there is a very high electric field at the boundary between the n-type and p+ gain layer [18]. When electrons reach the gain layer, they are accelerated further and gain sufficient energy to cause impact ionisation, generating further electron-hole pairs. The exact value of gain is influenced by the bias voltage, V_{bias} across the sensor. Only with a sufficiently large V_{bias} is the electric field strength high enough for impact ionisation to occur [18]. As V_{bias} increases, so does the electric field strength, and thus more impact ionisation can occur, leading to higher gain.

The gain, G, of an LGAD can be defined as the ratio between the total charge injected, Q, and the total charge collected, Q', assuming a uniform weighting field. At the same time as the signal increases, additional noise is introduced due to variability in the random nature of the multiplication process. This emerges because the multiplication process is random and so is the variation associated in the value
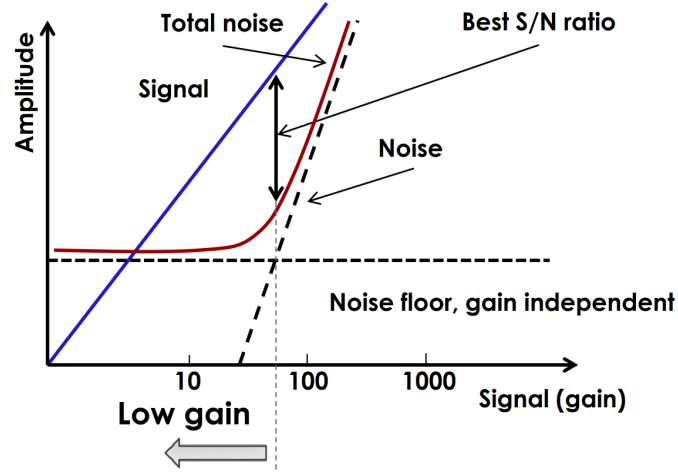
Figure 2.19: The noise floor from read-out electronics can be used to exploit an LGAD's gain. The noise from an LGAD increases more quickly than the signal, but until it is larger than the noise floor, the SNR can continue to rise with gain. Taken from [30]

of gain. The excess noise factor, F, characterises the increase in shot noise due to this variation and it can be shown that the SNR is proportional to $1/\sqrt{F}$ [18]. The excess noise factor is a function of G and α_p/α_n, where the latter is the ratio of the impact ionisation coefficient for holes and electrons respectively [18]. F can therefore be minimised by using an n-in-p configuration since at low gain, holes do not contribute to the multiplication process, minimising α_p. Nevertheless, if shot noise were the only source of noise, then the introduction of gain would result in a worse SNR.

In reality, the dominant noise at low gain is from the read-out electronics. This is a noise floor, independent of gain, which masks the shot noise [18]. This means that for increasing gain, the signal can be increased without the total noise increasing. This is clear in Figure 2.19 where the total noise remains unchanged until shot noise becomes the dominant source. There is a particular value of gain for which the SNR is maximised which, recalling Equation 2.33, minimises σ_{jitter} [18]. A typical gain to achieve this is on the order of 30 but it is important to note that the optimal gain is dependant on read-out noise. For worse performing electronics, the noise floor is larger and a larger gain is required before the shot noise becomes dominant again. Unfortunately, the SNR achieved at the optimal gain is also lower, thus σ_{jitter}
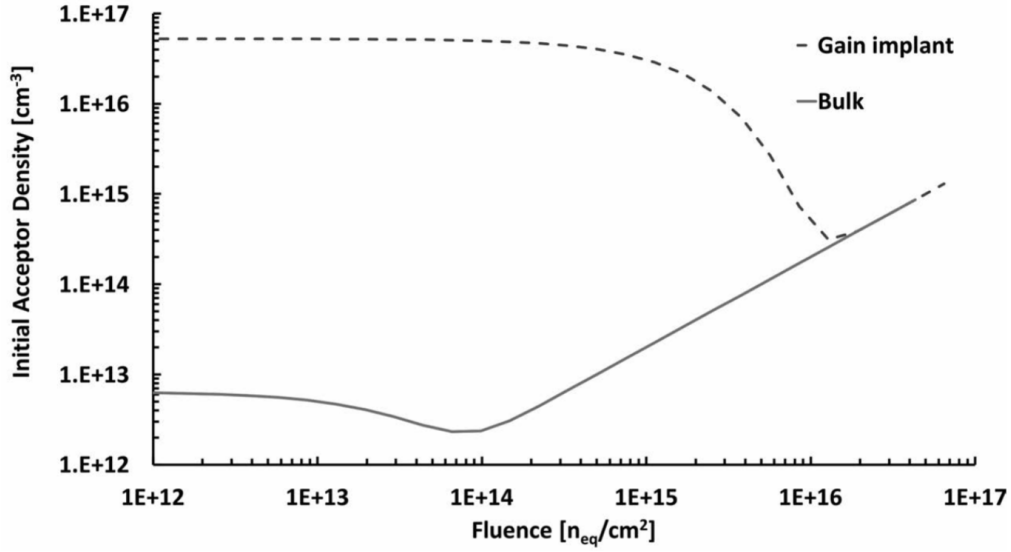
Figure 2.20: Acceptor density as a function of fluence. Note how the bulk and gain implant behave differently initially, but eventually converge for high fluences. Taken from [18]

remains larger.

2.5.3   Radiation Damage in LGADs

As discussed in Section 2.4.2.4, acceptor removal leads to a reduction of gain. The evolution of acceptor concentration is shown in Figure 2.20 which shows the difference between the bulk and gain implant. After a fluence of roughly $1 \times 10^{16} \, 1 \, \text{MeV} \, \text{n}_{\text{eq}}/\text{cm}^2$, the acceptor concentration of the bulk and gain implant is the same, and thus there is no more gain [18].

Acceptor removal leads to a high bias voltage requirement to achieve a similar value of gain. However, one needs to be cautious as the bias voltage approaches $12.7 \, \text{V}/\mu\text{m}$ [31]. Around this threshold, multiplication can occur in the bulk, known as bulk gain [18]. This can lead to single event burnout (SEB) where a single charged particle causes catastrophic and irreversible breakdown in the sensor [31]. In this work, a limit of $11 \, \text{V}/\mu\text{m}$ was adopted for added precaution.

It is important to remember that σ_{jitter} does not just depend on the SNR, but it

is also directly proportional to the rise time, t_r [18]. The rise time is the time it takes for electrons to reach the surface of the LGAD, which is dependent on their drift velocity. The drift velocity can be increased with a higher electric field caused by a high bias voltage. However, this is normally saturated very quickly. After irradiation, saturation takes place at a higher bias voltage [32, 33]. This means that after irradiation, an LGAD with low gain ($\sim$30) could have a better time resolution if the bias voltage required to achieve such a value of gain is higher.

DEVELOPMENT OF EXPERIMENTAL METHODS & CHARACTERISATION TECHNIQUES

Characterisation of LGADs is a necessary step in order to understand whether they meet the specification of the experiments they are designed for. Basic electrical characterisation: current-voltage (IV) and capacitance-voltage (CV), is common to all analogue silicon detector testing. LGADs can also be characterised by their most important operational properties: gain and time resolution. The theoretical motivation, techniques and analysis involved in characterising LGADs are discussed in this chapter. A large proportion of the work in this thesis has been the development and testing of these characterisation techniques and experimental methods.

Wafer #	Implant Energy	Implant Dose
19, **20**, 21	1.00	1.00
17, 18	1.07	1.00
15, **16**	0.92	1.05
12, 13, 14	1.00	1.05
9, 10, **11**	1.07	1.05
7, 8	1.15	1.05
4, 5, **6**	1.00	1.11
2, **3**, 24	1.07	1.11

Table 3.1: Wafers fabricated with one of eight variations of the gain layer implant energy and dose. These values have been normalised at the request of Te2v. Wafers tested at the University of Birmingham are in bold.

3.1 LGADs at Teledyne e2v

3.1.1 First Production Batch

The first production batch of 22 six inch wafers from Te2v finished fabrication in late 2020. This batch contained eight wafer variations, each with a different gain layer implant. The gain layer is varied by controlling the energy and dose of the boron atoms implanted into the region. Table 3.1 shows each type of wafer and their associated implant energy and dose which has been normalised with respect to the values for wafers 19, 20 and 21 at the request of Te2v for commercial sensitivity. Of the wafers produced, W20 (Wafer 20), W16, W11, W6 and W3 were selected for characterisation to represent a representative variation in implant energy and dose.

Each wafer, shown in Figure 3.1, consists of 42 repeating fields of which there are three different types. The repeating pattern allows for easy diamond saw dicing. The most common field is the test field which consists of LGADs and devices with no gain layer (PiN)[1] of various sizes and flavours. The second type of field is an array of LGADs each with a pad with dimensions $(1 \times 1)\,\mathrm{mm}^2$. The final field type

---

[1]Strictly speaking, these diodes without a gain layer are highly doped n-type in a very low doped p-type substrate with a high doped p-implant on the back to ensure an ohmic interface to the aluminium back contact. Because the the bulk p-type is of such low doping these devices are typically called PiN diodes, even though the substrate is not actually intrinsic.
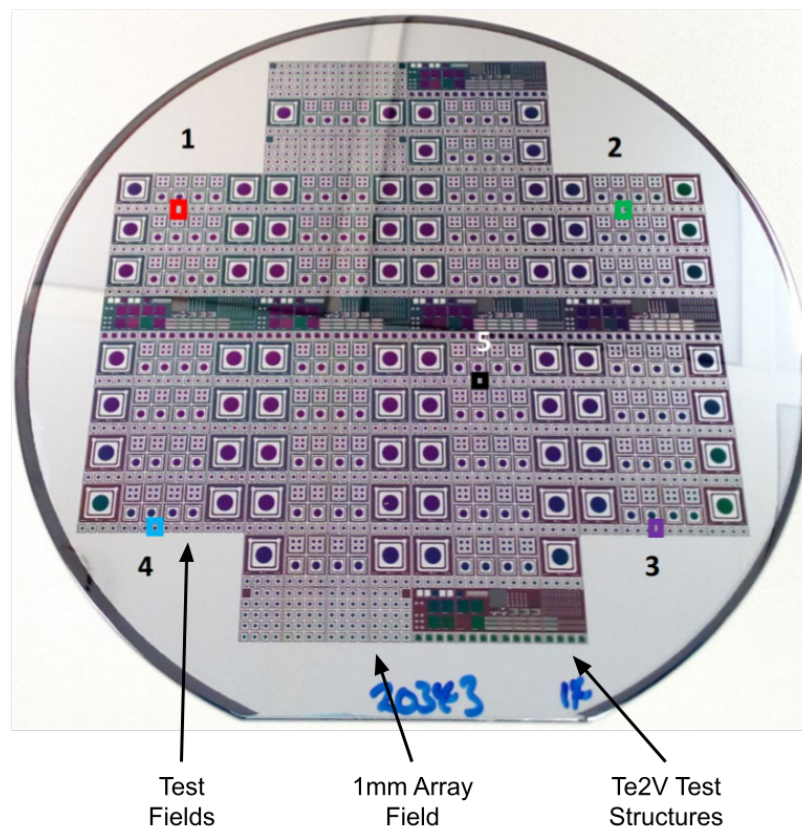
Figure 3.1: Wafer design for the first batch from Te2v. Each wafer consists of 42 fields of three different types. 22 wafers were fabricated in pairs/triplets of the same gain layer implant energy and dose. Indicated on the wafer are preferred test sites in order to test a variety of locations across the wafer.

consists of some simple test structures primarily used internally by Te2v. This work is concerned with the test field.
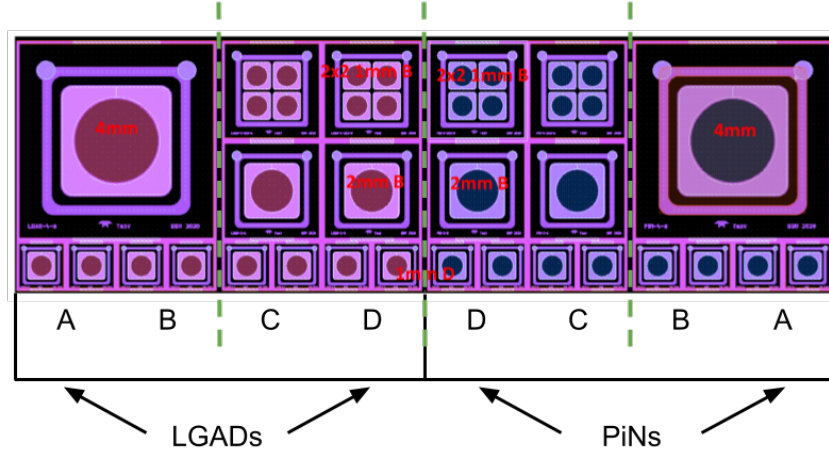
Figure 3.2: An example test field from the wafers in the first batch. Each test field contains LGADs and PiNs of various sizes and flavours. Each test field on a wafer is identical. Dotted lines represent the dicing lines. Taken from [34].

In each half of the test field, shown in Figure 3.2, there are: eight $(1 \times 1)\,\mathrm{mm}^2$ devices, two 2x2 $(1 \times 1)\,\mathrm{mm}^2$ arrays, two $(2 \times 2)\,\mathrm{mm}^2$ devices and one $(4 \times 4)\,\mathrm{mm}^2$ device [34]. This structure is mirrored down the centre of each field to give a set of LGADs and a set of PiNs (without the gain layer) [34]. In contrast to the $4\,\mathrm{mm}$ device, each of the other sizes have different flavours where properties, such as the distance from pad to guard ring, are varied [34]. For the $1\,\mathrm{mm}$ devices, there are four flavours with each coming as a pair [34]. The difference between these flavours was not investigated in this work, but was controlled and documented in the following chapter. The variety of test field locations means wafer uniformity can be investigated. Devices are usually identified by the wafer and field from which they come.

3.1.2 Te2v LGAD Design Details

The cross-section design of Te2v's LGADs is shown in Figure 3.3. In addition to the highly doped n-well and p+ gain layer, an n+ region called the junction termination edge (JTE) is implanted either side of the gain layer. Due to the way the edge of the gain layer is shaped, very high fields can form which can lead to early
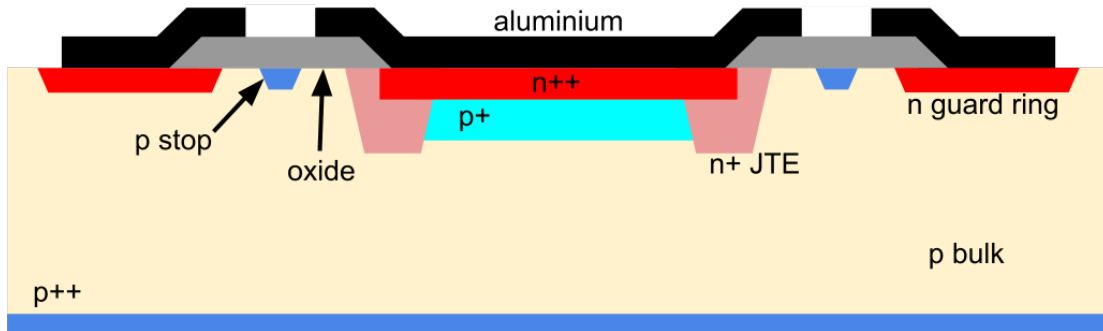
Figure 3.3: Sketch of the design used for the first batch of Te2v LGADs. Adapted from [35]

breakdown. The JTE shapes and limits these fields by creating a clean boundary at the edge of each pad, preventing this early breakdown [36]. This design also contains the p-stop and guard ring implants, discussed in Section 2.3.1 [37, 38]. All measurements are performed with the guard ring floating for simplicity, except for timing measurements where the guard ring is grounded to replicate the operating conditions in the final detector. Each $(1 \times 1)\,\mathrm{mm}^2$ LGAD has a roughly $800\,\mu\mathrm{m}$ diameter opening in the aluminium layer as shown in Figure 3.4. This aluminium layer is for electrical contact to the n-type region and the opening allows light to enter for gain measurements with a laser set-up.
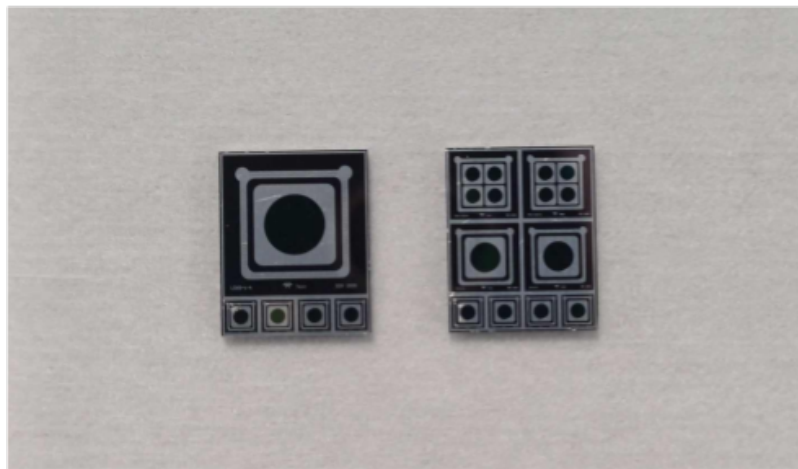
Figure 3.4: Diced LGAD devices. Note the circular openings in the metal layer and the guard ring surrounding each pad.

Wafer #	Dicing Method	Dicing Performed By
3	Laser	Scitech
11	Saw	Te2v
6, 16, 20	Saw	University of Birmingham

Table 3.2: Wafer dicing methods and where they were performed. Scitech was initially used for laser dicing. Low yield motivated a switch to saw dicing. This was split between Te2v and the University of Birmingham.

3.2 Device Processing

3.2.1 Dicing

In order to test individual devices, particularly with readout electronics, the wafers need to be diced into small segments. Each individual field is diced into roughly $(7 \times 9)\,\mathrm{mm}^2$ segments, as seen in Figure 3.4. There are two primary methods of dicing. The first is using a circular diamond saw which requires a regular segmentation pattern in the wafer design, such as that in Te2v's wafer. The second method allows for more complex dicing patterns by employing a high intensity IR laser which should also result in less edge damage compared to saw dicing. Table 3.2 shows which devices have been diced by which method and from where.

3.2.2 Thermal Annealing

Thermal annealing is the process of heating a semiconductor in order to reduce structural defects and activate dopants. Annealing is usually performed after implantation and is required to achieve full electric activation of the implanted impurities. It also helps to repair crystal damage due to the implantation as well as drive diffusion of the implanted impurities to give a smoother junction and reduced peak electric fields [15]. Dicing can also be a source of defects (or trapped moisture) at dicing boundaries which cause additional variability compared to measurements on a wafer. Peak electric fields can cause an early breakdown, and so the application of

Device Field ID	Fluence [$1\,\mathrm{MeV}\,\mathrm{n_{eq}/cm^2}$]
F11	5.6×10^{13}
F12	8.4×10^{13}
F13	1.0×10^{14}
F14	2.5×10^{14}
F08	4.0×10^{14}
F25	5.7×10^{14}
F28	8.3×10^{14}

Table 3.3: Seven 1 mm devices from wafer 3 were irradiated with increasing fluence. The field from which each devices comes from is used as their unique identifier.

another phase of thermal annealing was tested after dicing. In this work some devices were chosen to be annealed for 2 hrs at $150\,^{\circ}\mathrm{C}$ in order to try and improve the variability in their pre-irradiation IV curves after dicing [39]. This was performed at ambient conditions with $(50 \pm 10)\,\%\,\mathrm{RH}$.

3.2.3 Irradiation

As discussed in Section 2.4.2, understanding the effects of radiation damage is critically important, especially for LGADs which will be used at the HL-LHC. For this, seven devices were selected for complete characterisation before and after irradiation. These devices were all from W3, the highest combination of implant energy and dose. This is generally expected to have the greatest immunity to acceptor removal since the initial acceptor density is larger, following Equation 2.30 and Equation 2.31. This irradiation campaign also focused on the 1 mm ($(1 \times 1)\,\mathrm{mm^2}$) devices since the ATLAS and CMS experiments specify a pad size of 1.3 mm [5, 6]. Each device can be identified by its field location as shown in Table 3.3. The fluence is also shown, ranging from 5.3×10^{13} up to $8.3 \times 10^{14}\,1\,\mathrm{MeV}\,\mathrm{n_{eq}/cm^2}$. Irradiation was performed at the MC40 cyclotron at the University of Birmingham using 27 MeV protons.

3.3 Electrical Testing

This section focuses on the experimental setups designed for the characterisation of LGADs. The use of a probe station and climate chamber to perform IV and CV measurements will be discussed.

3.3.1 IV Measurements

The leakage current of the bulk region of a pre-irradiated silicon sensor, I_{bulk}, is discussed in Section 2.4.2.1 and is proportional to the depleted volume. Ideally, the leakage current should grow proportionally with the depleted volume while being as low as possible to minimise shot noise contributions to the readout. Hence, it is important to measure current as a function of reverse bias voltage (IV measurements) to ensure the current is minimised for the entire operating range. The voltage must be ramped slowly ($\sim 1\,\mathrm{V/s}$) in order to ensure the leakage current has reached equilibrium before being measured. The bias voltage is provided by a Keithley 6517b[2] which can also measure current of a few pA. The circuit diagram for an IV measurement is shown in Figure 3.5.

A Python script has been developed to control the ramping procedure in which the step size is 1 V for the first 100 V and then 2 V thereafter in order to provide sufficient precision while minimising execution time. An average of five measurements are taken after a 5 s wait at each voltage step. The script continues increasing the applied voltage until either the current or voltage exceed $200\,\mu\mathrm{A}$ or 650 V respectively. The current limit ensures that the full IV range is recorded, including the high current seen during the onset of breakdown. The voltage limit ensures that a destructive breakdown does not occur. Above this limit, a current threshold is insufficient to prevent destructive breakdown.

---

[2]https://www.tek.com/en/datasheet/6517b-electrometer-high-resistance-meter
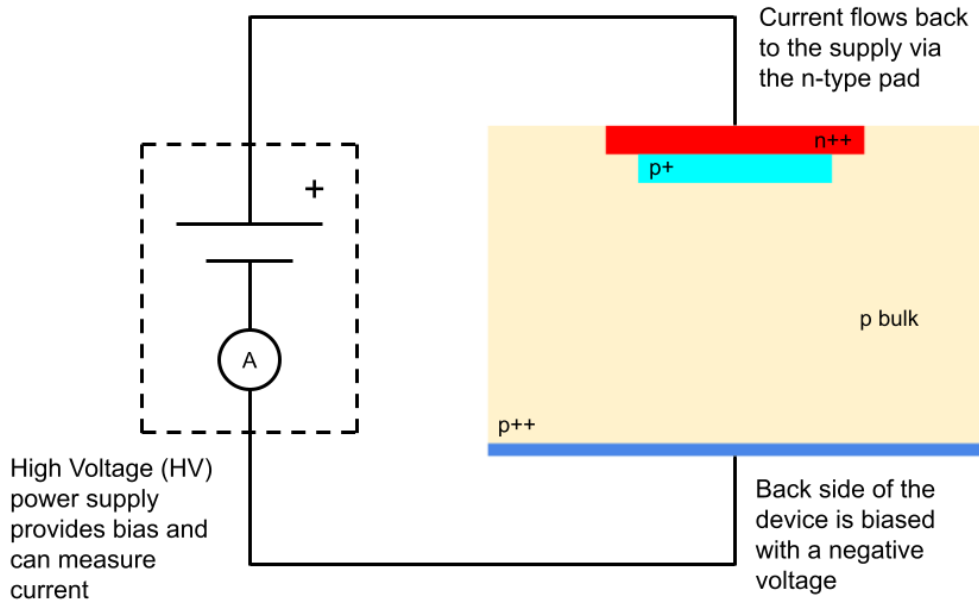
Figure 3.5: Circuit used to measure current as a function of reverse bias voltage. Bias is applied to the backside as a negative voltage.

3.3.2 IV Analysis

An important aspect of the current-voltage relationship is the breakdown voltage, V_{bk}. In this work, breakdown is defined where a very rapid increase in current happens as a function of voltage. In their recent LGAD market survey, ATLAS defined V_{bk} as the voltage for which the current (measured at $-30\,°C$) exceeds $200\,nA$ in a $(1.3 \times 1.3)\,mm^2$ device. This works well for pre-irradiated LGADs, although it is not strictly a definition of breakdown which can be defined as when the sensor enters a regime where the current is rapidly rising with time [15].

After irradiation, the leakage current increases and can exceed $200\,nA$ even at relatively low voltages. In this case, the threshold is increased to $10\,\mu A$. It should be noted that some post-irradiated devices will never exceed this threshold before the $650\,V$ ceiling.

The increase in leakage current, $i(T)$, post-irradiation is countered by operating at a lower temperature, such as $T = -20\,°C$, since,

$$i(T) = i(T_0) \left(\frac{T}{T_0}\right)^2 \exp\left(\frac{-1.2}{2k_B}\left(\frac{1}{T} - \frac{1}{T_0}\right)\right), \tag{3.1}$$

where $i(T_0)$ is the current at a different temperature, T_0. Correcting with this equation allows a direct comparison of the leakage current even when measurements are taken at different temperatures. A lower temperature also leads to a lower breakdown voltage. This is because the impact ionisation coefficient increases as temperature decreases, meaning the gain also increases for the same bias voltage [40]. Hence, the temperature dependence needs to be accounted for when comparing pre- and post-irradiation IV measurements.

3.3.3 CV Measurements

As the reverse bias voltage increases, the depletion depth also increases. This leads to a decrease in the capacitance, $C(V)$ as described by,

$$C(V) = \frac{\epsilon_r \epsilon_0 A}{d(V)}, \tag{3.2}$$

where A is the area of the sensor's pad. Here, ϵ_0 is the permittivity of free space and $\epsilon_r = 11.68$ is the relative permittivity of silicon. The change in capacitance is sensitive to the effective doping concentration, $N(d)$, which varies with depth as shown Equation 2.16. This means that $N(d)$ can be calculated from the change in capacitance as a function of voltage such that,

$$N(d) = \frac{2}{\frac{d}{dV}\left(C(V)^{-2}\right)} \cdot \frac{1}{\epsilon_r \epsilon_0 A^2}. \tag{3.3}$$

Measuring capacitance is not as straightforward as measuring current directly. Instead, it is assumed that the circuit can be modelled as a series LCR circuit, as show in Figure 3.6 [41]. A small AC voltage is applied across an inductor, a capacitor
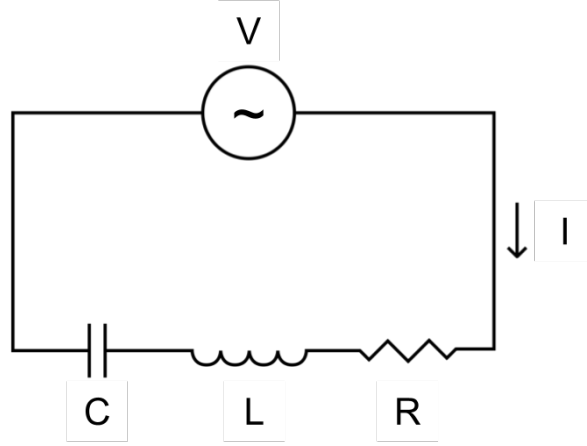
Figure 3.6: The standard LCR circuit used to study the effects of a capacitor, C, inductor, L, and resistor, R, on an AC voltage, V, driving and current, I.

and a resistor in series. To a very good approximation, $L \simeq 0$ for the full range of relevant frequencies, and therefore the impedance of the system can be written as,

$$Z = R - \frac{1}{\omega C}j = ze^{j\theta}, \tag{3.4}$$

where,

$$z = \sqrt{R^2 + (\frac{-1}{\omega C})^2}, \tag{3.5}$$

is the modulus of the impedance and,

$$\theta = \arctan(\frac{-1}{\omega CR}), \tag{3.6}$$

is the phase angle [41]. For an AC voltage signal, it can be shown that the capacitance and resistance can be calculated using

$$C = \frac{1}{\omega z}\sqrt{1 + \cot^2(\theta)}, \tag{3.7}$$
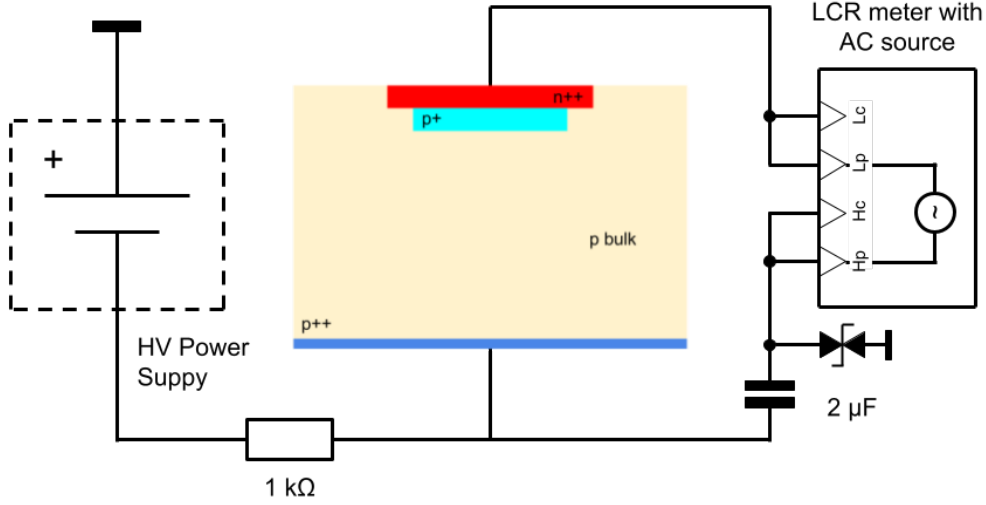
Figure 3.7: Circuit used to measure capacitance as a function of reverse bias voltage.

and

$$R = \frac{z}{\sqrt{1 + \tan^2(\theta)}}.$$ (3.8)

An LCR meter measures z and θ and calculates R and C. Either a Wayne Kerr 6500B[3] or Agilent E4980[4] were used for the measurements. The circuit diagram for a CV measurement is shown in Figure 3.7. A filtering capacitor is required to allow the AC signal from the LCR meter to pass, while blocking the high voltage bias from a source measurement unit (SMU) which could damage the meter. A bi-directional transient-voltage-suppression (TVS) diode is also used to add protection against high voltage. The resistor is required to prevent the AC signal seeing the SMU as a low-impedance path to ground. A coupling box is used to house these electronics. It includes a switch which allows a quick change from IV (Figure 3.5) to CV (Figure 3.7) measurements. The circuit diagram for this box is shown in Figure 3.8.

A python script was also developed for CV measurements. Since the transition from gain layer depletion to bulk depletion occurs over a short voltage range, a step size

---

[3]https://www.waynekerrtest.com/products$_detail.php?index_s = 4$
[4]https://www.keysight.com/gb/en/product/E4980A/precision-lcr-meter-20-hz-2-mhz.html
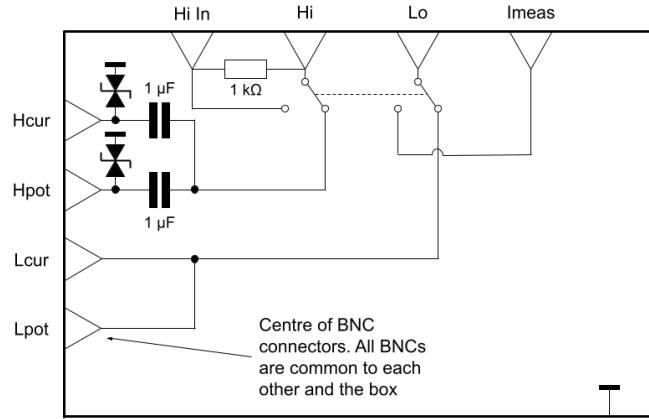
Figure 3.8: The coupling box which allows IV or CV measurements to be quickly selected. 'Hi' and 'Lo' refer to backside and pad of the device under test (DUT) respectively. 'Hi In' is the BNC connections for the HV power supply. 'Imeas' is the return BNC connection for measuring the current with the SMU

of 0.2 V is used for the first 50 V. From then on, a 2 V step is sufficient until 150 V is reached. Testing a range of frequencies at each voltage step ensures that any frequency dependence, as will be discussed in Section 3.3.5, can be monitored.

3.3.4   CV Analysis

As described in Equation 3.3, CV curves can be used to determine the doping concentration variations with depth. In the case of a PiN, the capacitance immediately falls before plateauing roughly 10 V later, as shown in Figure 3.9a. The PiN is said to have reached full depletion at a voltage, $V_{FD} = 10$ V. An LGAD takes a different route with a much slower fall in capacitance initially. This is caused by the gain layer which is first to deplete. Due to its high doping concentration, it takes up to 27 V to deplete fully in the example shown in Figure 3.9a. The dependence of this voltage, called V_{GLD}, on the doping concentration is crucial when comparing wafer flavours and the effects of irradiation. After the gain layer has depleted, the bulk can then deplete and full depletion occurs roughly 10 V later. Both depletion voltages can be extracted from a CV curve using two different methods.

The full depletion voltage is found by looking for a significant deviation in the value
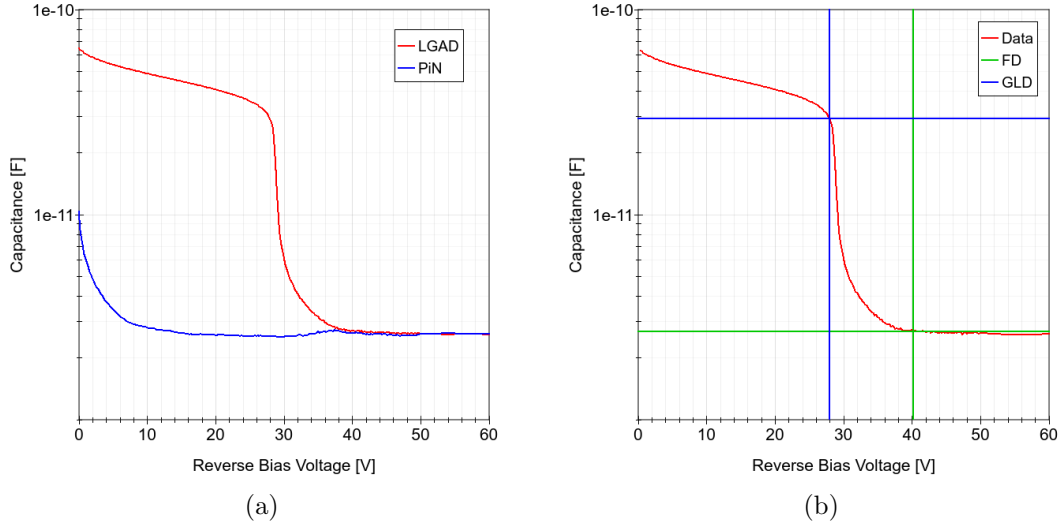
Figure 3.9: (a) Example CV curve for a 1 mm LGAD (from W3) and a PiN (from W2). Full depletion occurs at a much higher voltage for the LGAD since it must first deplete the gain layer. (b) The extracted depletion voltages for the LGAD.

of capacitance at full depletion. Once fully depleted, neither the active thickness nor the capacitance should change even if the voltage continues to increase, assuming there is no lateral depletion. Therefore an average and standard deviation can be calculated from everything measured above 50 V, chosen because no change in capacitance is observed beyond this point. The first voltage for which the measured capacitance no longer deviates from the average by more than five times the standard deviation is considered to be V_{FD}. This method and the result of using it is shown in Figure 3.10b and Figure 3.9b respectively, and demonstrates that the method tends to bias towards underestimating V_{FD}.

Figure 3.10a shows how the gain layer depletion voltage is extracted. The capacitance is differentiated twice with respect to voltage. Since the CV curve of an LGAD is primarily made up of 3 straight sections, the second derivative remains at 0. This is with the exception of the two turning points corresponding to V_{GLD} and V_{FD}. These turning points cause the two peaks seen in Figure 3.10a. Starting at the negative peak and working backwards, V_{GLD} is considered to be the first voltage at which the second derivative exceed the threshold $1 \times 10^{-11} \, \mathrm{FV}^{-2}$. The choice of threshold is arbitrary and was chosen as a value which best represents the turning

(a)   (b)
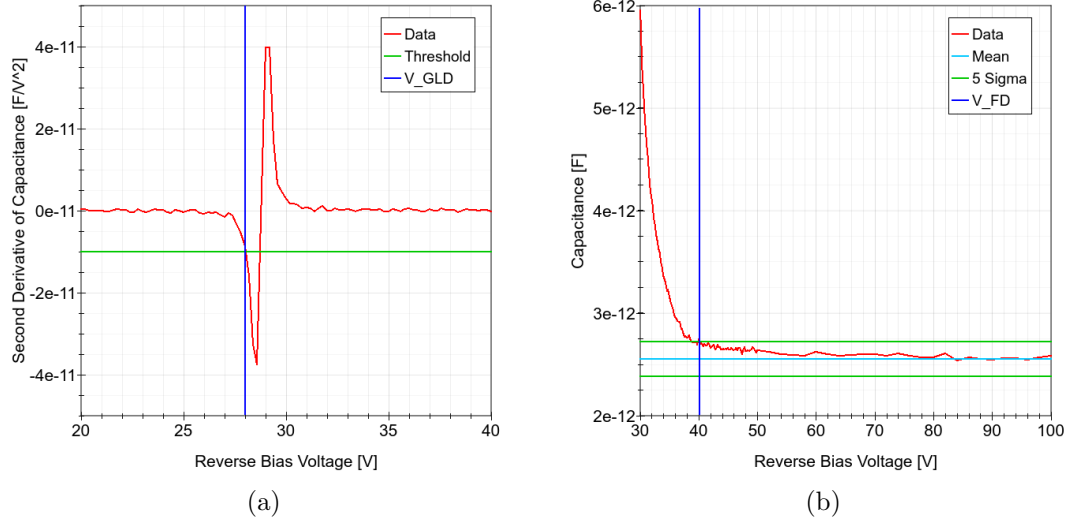
Figure 3.10: Visualisation of the extraction of (a) the gain layer depletion voltage
and (b) the full depletion voltage.

point that one might identify visually. Figure 3.9b shows the result of this method
on the example CV curve.

3.3.5   CV Frequency Dependence

The capacitance measured by the LCR meter depends on the measurement fre-
quency, ω, of the AC signal, as shown in Equation 3.7. This dependence is shown
in Figure 3.11 and indicates a region between 1 kHz and 20 kHz for which the ca-
pacitance does not change. This corresponds to Equation 3.2 which also has no
dependence on ω. After irradiation, the frequency dependence changes and only
frequencies below 2 kHz are valid, as shown in Figure 3.11. Choosing the right
measurement frequency is very important in order to calculate the correct value of
capacitance and depletion depth [42]. Contrary to Figure 3.11, it was found that
100 kHz was suitable for pre-irradiation measurements. Therefore, in this work,
CVs are measured at 100 kHz pre-irradiation and 1 kHz post-irradiation, unless oth-
erwise stated. 1 kHz was chosen after it became clear that 100 kHz was no longer
valid during post-irradiation studies. While the measured capacitance can be af-
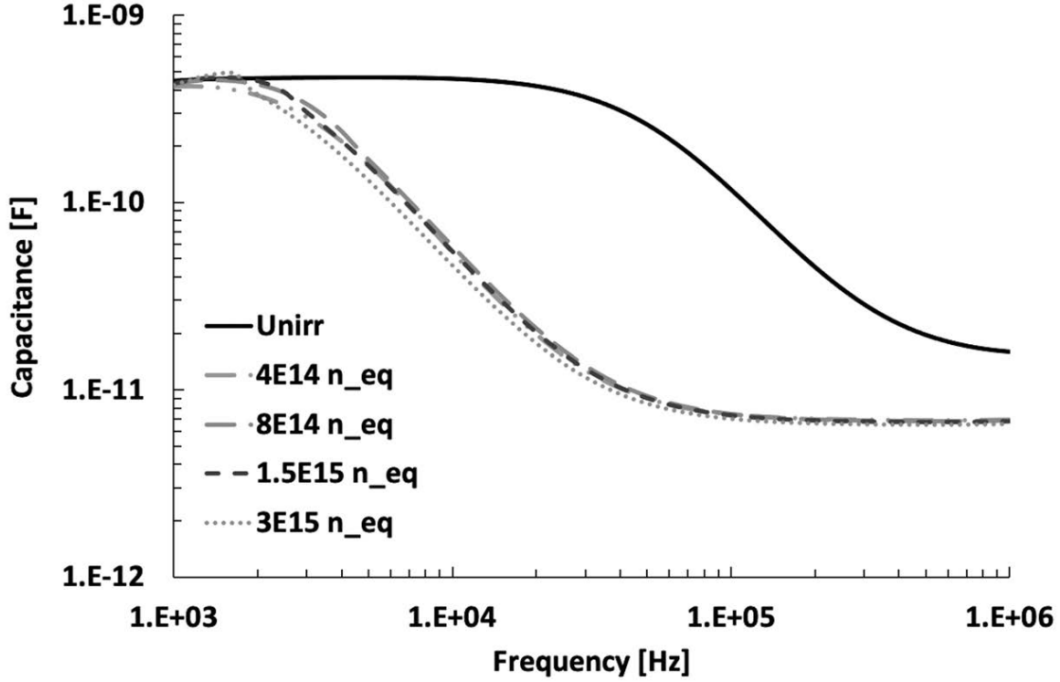
Figure 3.11: Partially depleted sensor biased at $-10\,$V. There is little dependence of C on frequency until very high values of frequency. Post-irradiated, the range of frequencies over which the capacitance is independent of frequency is significantly reduced. Taken from [18].

fected by frequency, the depletion voltages are independent. This is demonstrated in Figure 3.12 where one can see visually that V_{GLD} remains fairly consistent while the value of capacitance varies greatly. V_{FD} should also not vary, but it would appear that the capacitance's dependence on frequency is affecting when exactly full depletion appears to occur.

3.3.6   Automatic Probe Station

A probe station can be used to test wafers directly, as shown in Figure 3.13. The wafer is placed on a metal chuck which provides an electrical connection to the backside of the devices for biasing. Individual devices can then be quickly selected for testing by manually moving a needle probe into contact with the DUT. The chuck, which holds the wafer in place with a vacuum, can also be used with single diced devices, allowing for quick IV and CV measurements in order to check for any
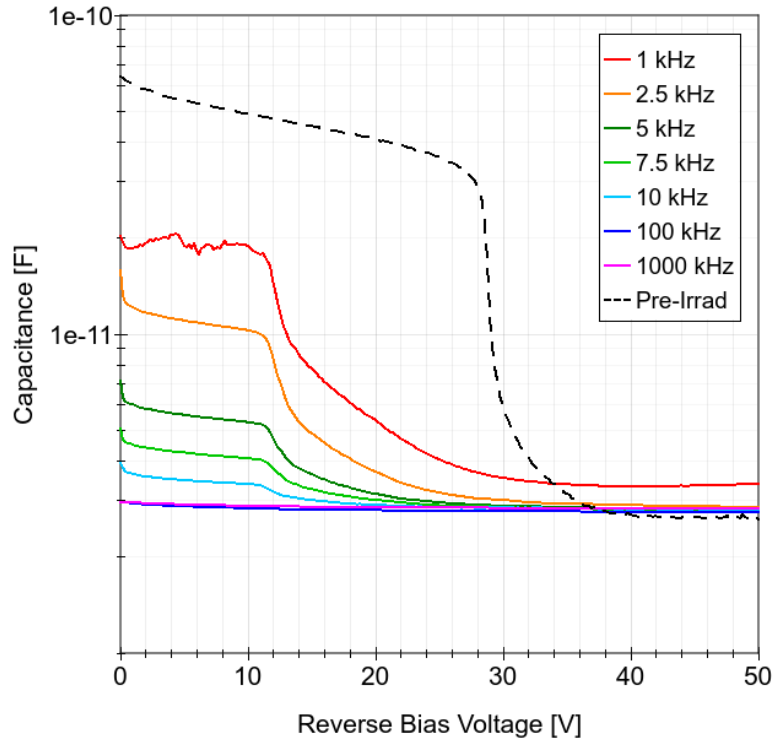
Figure 3.12: CV curve measured at different frequencies for an LGAD irradiated with a fluence of $8.3 \times 10^{14}\,1\,\text{MeV}\,\text{n}_{\text{eq}}/\text{cm}^2$. While the measured capacitance varies significantly, the gain layer depletion voltage remains unchanged. Full depletion is much harder to determine after irradiation (See Section 3.3.4).
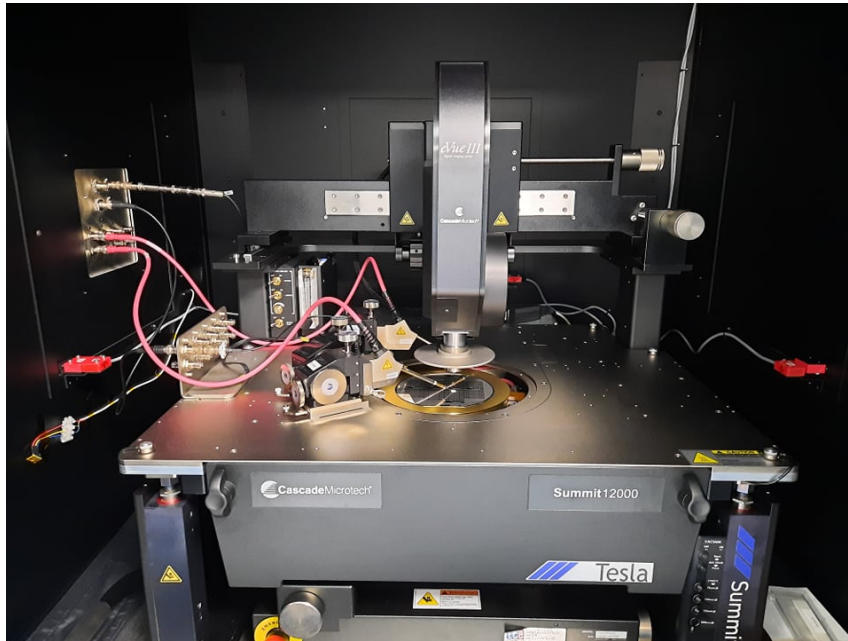
Figure 3.13: An automatic probe station allows wafers and diced devices to be quickly tested without the need for wire-bonding or gluing to a PCB.
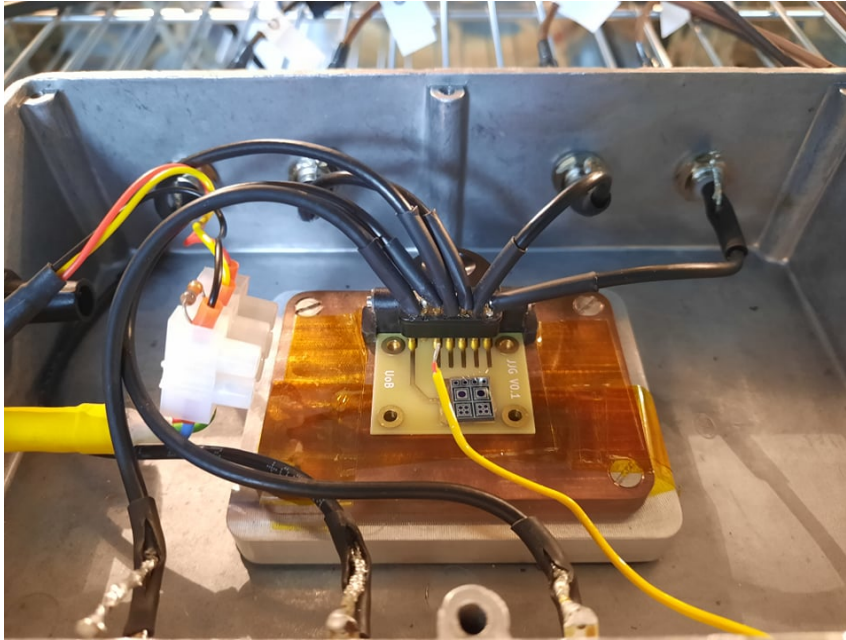
Figure 3.14: Aluminium box placed inside the climate chamber. The DUT is glued and wire-bonded to a small PCB which connects to external equipment via spring-loaded pins.

damage or changes after dicing. Bulkhead connectors allow the chuck and probes to be connected to the relevant SMU or LCR meter.

The probe station is light-tight, but not air-tight. Hence, any measurements are performed at room temperature. Environmental conditions are monitored for significant deviation, with humidity being less tightly controlled despite the ambient conditions being specified as $(50 \pm 10)\,\%\,\mathrm{RH}$ and controlled for a cleanroom environment.

3.3.7 Climate Chamber

A climate chamber is a light- and air-tight chamber in which the temperature can be lowered to $-20\,^{\circ}\mathrm{C}$ with a controlled relative humidity. This allows for IV and CV measurements to be performed after irradiation where a low temperature is required to maintain a low leakage current.

Within the climate chamber is an air-tight aluminium box which can be filled with

nitrogen, lowering the humidity to $<10\,\%\,$RH, as shown in Figure 3.14. The DUT is then placed inside after being glued via double-sided conductive tape to the PCB shown in Figure 3.15. The DUT can then be wire-bonded to the 5 pads allowing up to 5 different sensors to be tested.
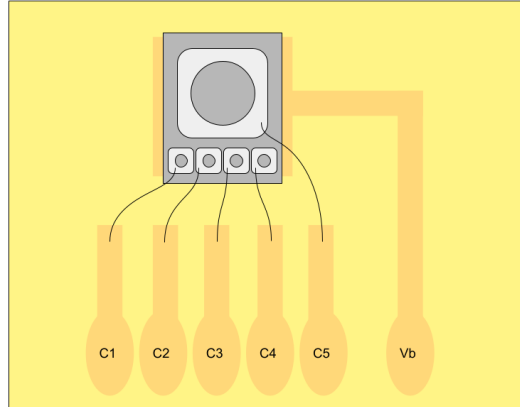
Figure 3.15: The PCB used to test a device in the climate chamber. It has 5 pads which can be wire-bonded to separate sensors on the same devices segment.

3.4   Laser Charge Injection

An LGAD's relationship between gain and bias voltage is different for every flavour of gain implant. As discussed in Section 2.5.2, a higher gain results in an improved time resolution up to an optimal value of SNR. Developing a method to measure the voltage-dependent gain is essential for understanding how this relationship evolves as LGADs are irradiated.

3.4.1   Set-up

A laser is an easy and reliable way to inject charge and measure the gain of an LGAD. This is done using a transient current technique (TCT) set-up shown on Figure 3.16 [21, 43]. A 1064 nm pulsed laser is directed into a telescope which focuses the beam onto the DUT below. The resulting voltage pulse is read out and the gain
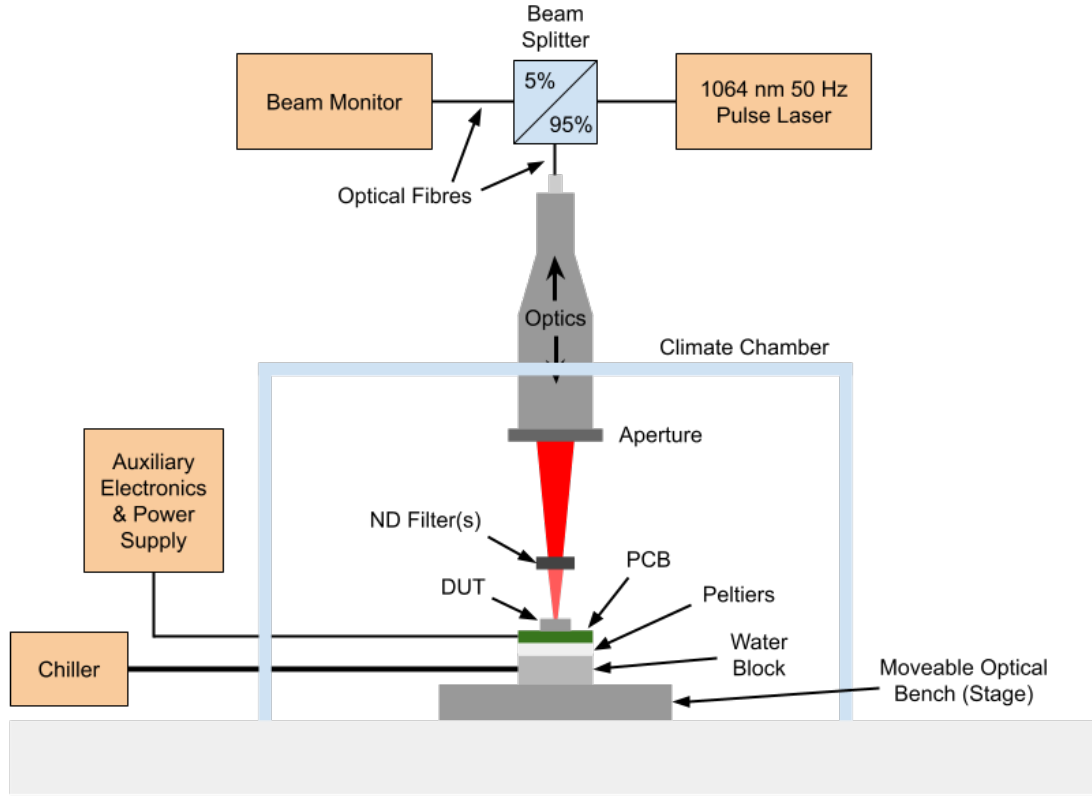
Figure 3.16: The TCT set-up used to measure the gain of an LGAD. A laser is split between a beam monitor and the optics required to focus the laser onto the DUT. The entire setup is enclosed in a light-tight aluminium box.

can be calculated by taking the ratio between the charge collected in an LGAD and in a PiN.

Dedicated software controls the laser and its settings. The pulse repetition rate is set to 50 Hz which is significantly longer period than the response time of the DUT and electronics. The intensity is controlled by the temporal width of the pulse which is smaller than the integration time of the DUT and electronics. The width is kept as high as possible so that it remains constant. The beam travels by optical fibre to a telescope which focuses the beam. The spot size can be further adjusted with an iris aperture at the end of the telescope. The spot size needs to be smaller than the opening in the LGAD's metalisation, but large enough to avoid the complication of gain suppression (discussed in Section 3.4.5). A $\sim 300\,\mu$m diameter is sufficient for this. The beam was focused onto the surface of the DUT using a red laser. It was

assumed that the divergence of the beam over the $50\,\mu$m depth does not result in the spot size exceeding the size of the DUT. However, the spot size also needs to be the same between associated LGAD and PiN measurements.
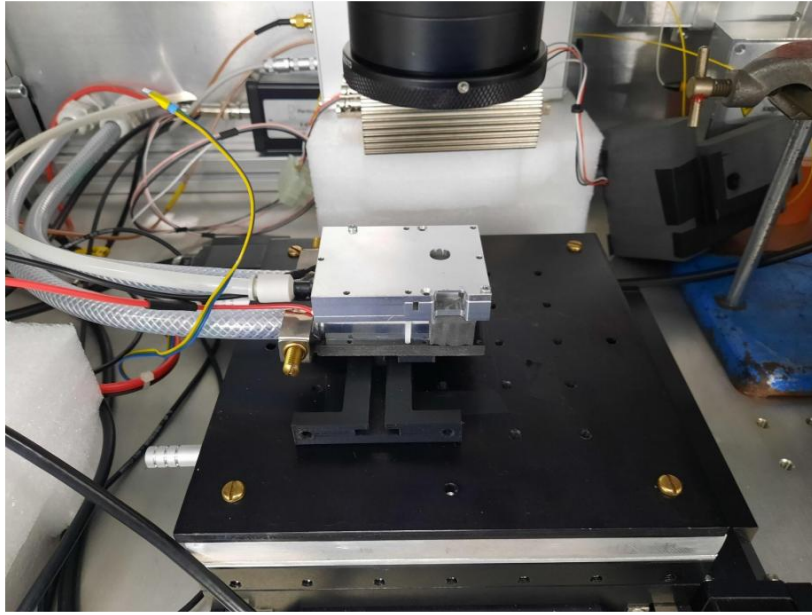
Figure 3.17: The DUT is attached on an optical bench which can be moved and aligned with the laser.

Below the optics is an optical bench on top of a movable stage, shown in Figure 3.17. A python script has been developed to operate the stage, moving with micron-level precision [43]. This allows the DUT to be quickly and accurately aligned with the laser spot.

LGADs can have a gain in excess of 100 at high bias voltages. Hence, the intensity of the laser must be kept low enough so that the readout electronics are not damaged. At the same time, a PiN does not have any gain and so requires a high intensity in order to see a signal above noise. In order to reduce the intensity by a well controlled factor for the LGAD, a neutral-density (ND) filter is used. The exact combination of filters is different for each LGAD and must be determined manually before a voltage sweep can be performed (but typically either ND0.5 or ND0.7 is used). Figure 3.18 shows the intensity recorded by an LGAD, held at a fixed voltage, as a function of ND number. The intensity has been normalised to that of ND0.0 which is the case
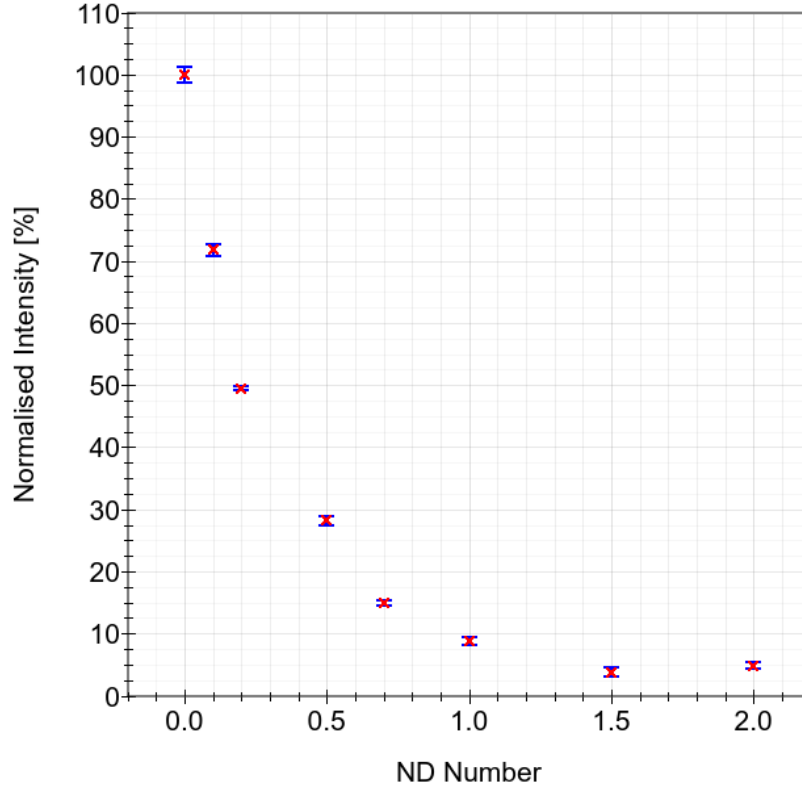
Figure 3.18: The laser intensity measured by an LGAD at a fixed voltage compared with varying ND filter. The intensity is normalised to ND0.0 which is when no filter is present. The error bars are in blue and show overlap between the intensity at ND1.5 and ND2.0 since the signal at these intensities is too small to measure accurately.

when no ND filter is present, as is always the case for a PiN measurement. This calibration is used for all gain measurements reported in this work.

3.4.2 Readout & Operation

The DUT is wire-bonded and glued with double-sided conductive tape to the PCB shown in Figure 3.19. Readout for a single device is achieved with a direct connection to an SMA connector which also provides the bias. Figure 3.20 shows the electronics required to bias and readout an LGAD in this set-up. Bias voltage from a Keithley 2410 is first passed through a HV filter. This is then passed into a bias-T which is connected to both the DUT and an amplifier. A bias-T uses an inductor and a

capacitor to block AC and DC respectively. This allows the DUT to be biased and read out via the same channel while isolating the amplifier from the high bias voltage. The inductor ensures that no additional AC is injected into the system. A single stage of amplification generates a voltage pulse which is probed by an oscilloscope. The amplifier is the AM-02 A 53 dB manufactured by Particulars, biased at 6.5 V [44]. The oscilloscope is triggered on the trigger signal of the laser.

Figure 3.19: The PCB used for TCT measurements. The board supports dual-channel readout where the resistors work with the parasitic capacitance built into the PCB to filter DC, similarly to the bias-T. However, it is primarily used in a simple single-channel readout which requires no electronics on the PCB.

A python script has been developed to control a full voltage sweep for each device. The voltage step varies as a function of voltage, starting at 50 V and ending at 2 V once the gain becomes particularly sensitive to voltage. Measurements at each voltage step are repeated 30 times to ensure their average has an acceptably small uncertainty. The script is set-up to detect significant voltage spikes in the signal. If this is detected, the measurement is taken again.
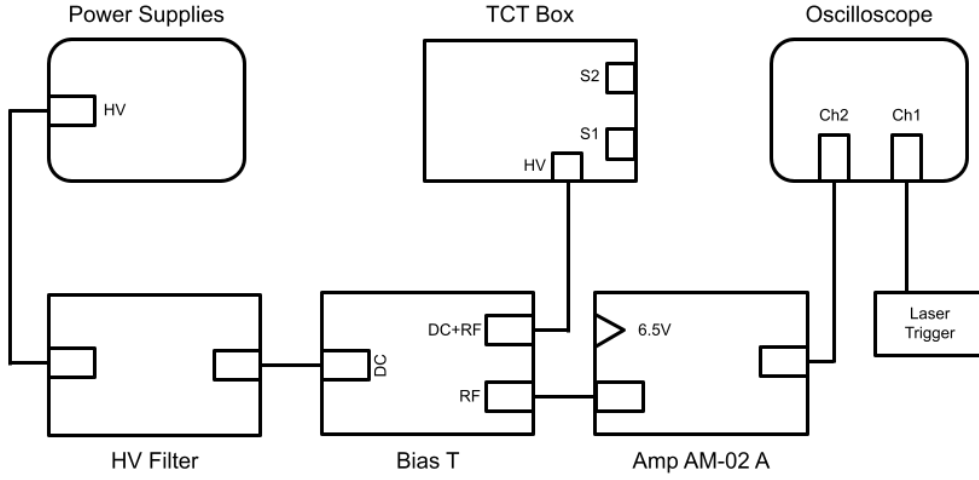
Figure 3.20: Block diagram of the electronics chain required to perform gain measurements in the TCT set-up. Lines between components represent SMA cables. The length of these cables should be optimised to minimise reflections of RF signals.

3.4.3 Cooling

The PCB is placed in an aluminium box shown in Figure 3.17. The box is held into place on the optical bench. Just below the box is a Peltier module which can cool the DUT to $-20\,°\text{C}$. Below this is a water block which is required to prevent the Peltier from overheating. The water block is supplied with $\sim2\,°\text{C}$ water from a chiller unit. The temperature is monitored with a PT100 on the PCB. The current supplied to the Peltier is manually tuned each time in order to ensure the correct temperature is achieved. The optics and stage are surrounded by custom transparent acrylic housing. This allows nitrogen to fill the housing which keeps the humidity below $10\,\%\,\text{RH}$, preventing ice from forming on the DUT at low temperatures.

3.4.4 Analysis

An oscilloscope records a waveform, $V(t)$ which is voltage as a function of time. The amplifier operates in current-mode meaning the integral of $V(t)$ is proportional to the charge collected, Q, such that,
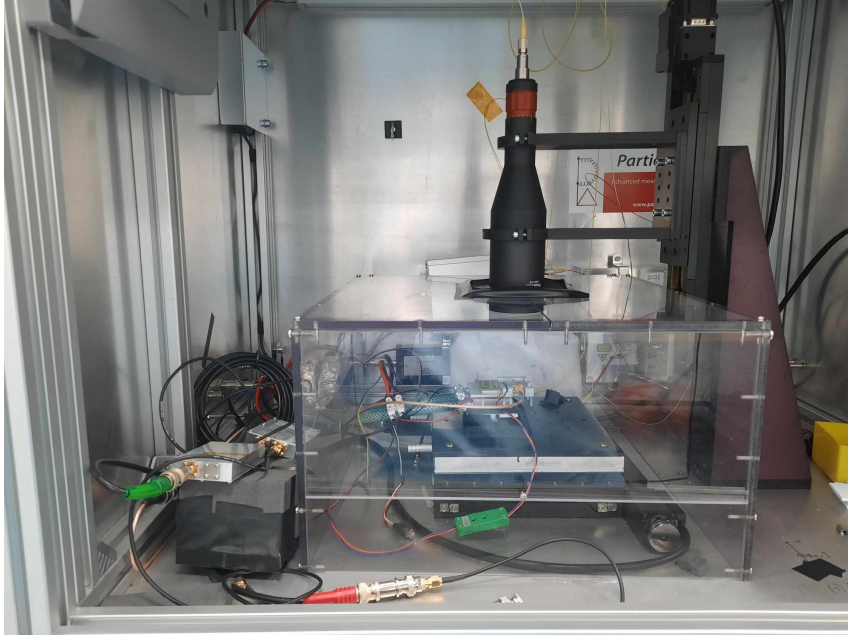
Figure 3.21: Surrounding the TCT system is a transparent acrylic box. A low humidity environment allows for the DUT to be cooled to $-20\,^{\circ}\mathrm{C}$, which is necessary post-irradiation.

$$S_{DUT} = \int_{t_0}^{t_0+t_w} V(t)dt \propto Q = qG, \tag{3.9}$$

where q is the charge generated by the incident light and G is the internal gain of the sensor [18]. A time window starting at t_0 of width t_w is used to integrate between. These times must be carefully selected to ensure that no reflections or excess noise are included. Reflections are usually minimised by cable length optimisation. In principle, integrating noise over a large enough window should cancel to zero, but in practice this is not always seen. Instead, a measurement must be made while the laser is not emitting light. Integrated over the same time window, this acts as a baseline which can be subtracted from the primary measurement. Figure 3.22 is an example of oscilloscope traces for an LGAD and a PiN with an encompassing time window.

The ratio of S between that of an LGAD, S_L, and a PiN, S_P, provides the relative gain. Assuming a PiN has a gain equal to one, the gain of the LGAD is hence given as,
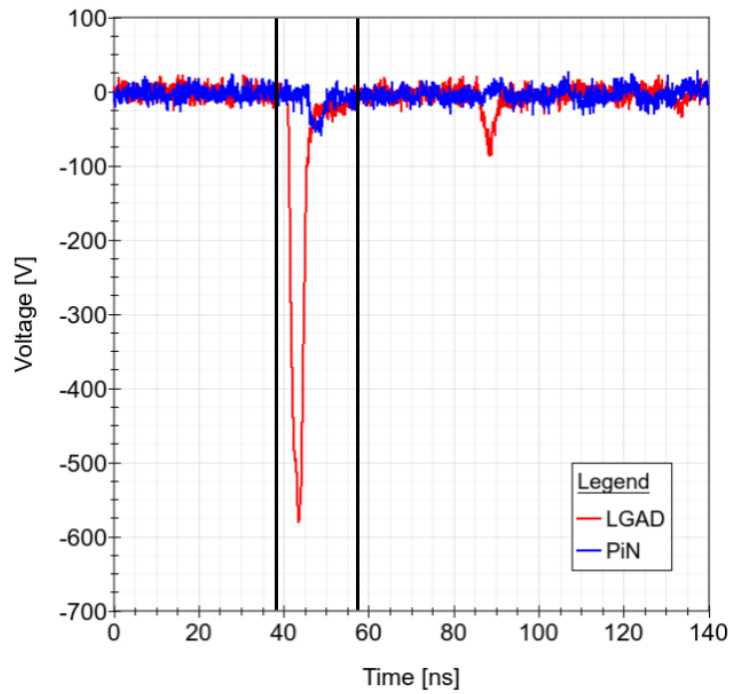
Figure 3.22: Example pulse from an LGAD and a PiN. A time window is carefully selected which encompasses the entire waveform for both the LGAD and the PiN. Note that care must be taken to avoid enclosing the reflected LGAD peak seen at ~90 ns.
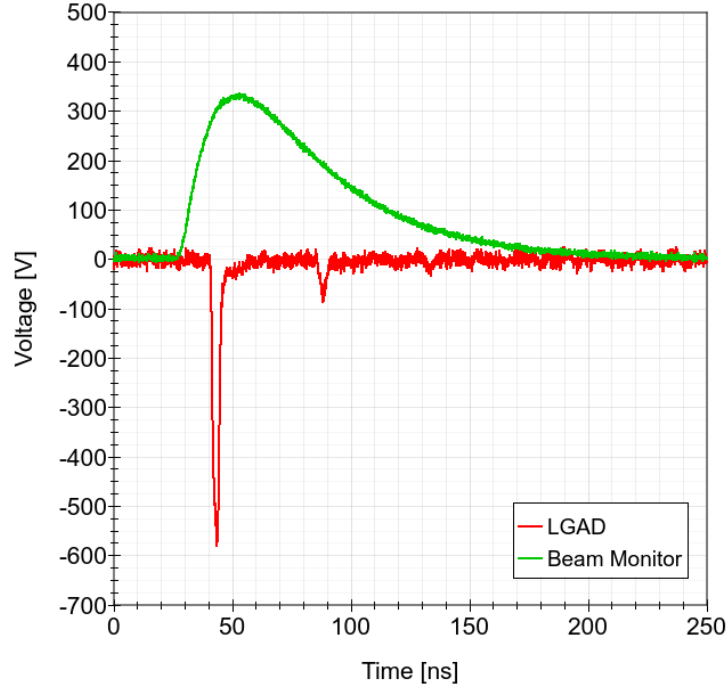
Figure 3.23: The pulse from an LGAD is shown with its associated beam monitor pulse. The decay time of the beam monitor signal is sufficiently large to make repeatable measurements for calibration.

$$\frac{G_L}{G_P} = G_L = \frac{S_L}{S_P}. \tag{3.10}$$

This relies on the assumption that the same amount of charge is injected into both the LGAD and the PiN. The intensity of the laser occasionally varies between subsequent measurements and this is corrected for using a beam monitor. The optical fibre cables has a $95\,\%/5\,\%$ splitter which sends a small fraction to the beam monitor for calibration as shown in Figure 3.23. The remaining $95\,\%$ is sent through the optics to the DUT. Similarly to the DUT, the integral of the beam monitor waveform is assumed to be proportional to the charge injected, such that,

$$S_{BM} \propto qG_{BM}, \tag{3.11}$$

where G_{BM} represents the gain of the beam monitor system. In order to correct for

a change in q, Equation 3.9 can be rewritten as

$$S'_{DUT}\frac{S'_{BM}}{S_{BM}} \propto qG_{DUT}, \tag{3.12}$$

where S'_{BM} is a reference value. When calculating gain, this cancels out and so it is often ignored entirely.

3.4.5   Charge Collection & Gain Suppression

If one has already calculated the gain of the sensor, then the charge collected can be calculated if one knows the type of impinging particles. For mips, an average of 80 electron-hole pairs are generated per μm [22]. Assuming they traverse perpendicularly through an LGAD with an active thickness of $50\,\mu$m, roughly 4000 electron-hole pairs and $0.64\,$fC of charge are generated. Since the charge collected is proportional to the gain of a sensor, the charge collected by an LGAD with a typical gain of 10 is $6.4\,$fC.

In this work, particles are assumed to travel perpendicular through a sensor. While this is valid during characterisation, it is not valid in general. Non-perpendicular paths result in a larger distance travelled and hence more charge is generated. This results in a larger SNR and eventually charge sharing between pads. The problem with perpendicular paths is an effect called gain suppression [45]. When charge is generated by a mip in a sensor, it is highly concentrated along the mip's path ($\sim 1\,\mu$m column diameter) and the generated electrons begin to screen the electric field. A reduction in the electric field strength means the avalanche is weaker and thus less charge is generated. Gain suppression only tends to occur at higher values of gain, upwards of 45 as shown in Figure 3.24. Below this limit, the difference between using a mip and a laser is negligible. This is an important effect and should be considered when reporting high values of gain with a laser as the source of charge.
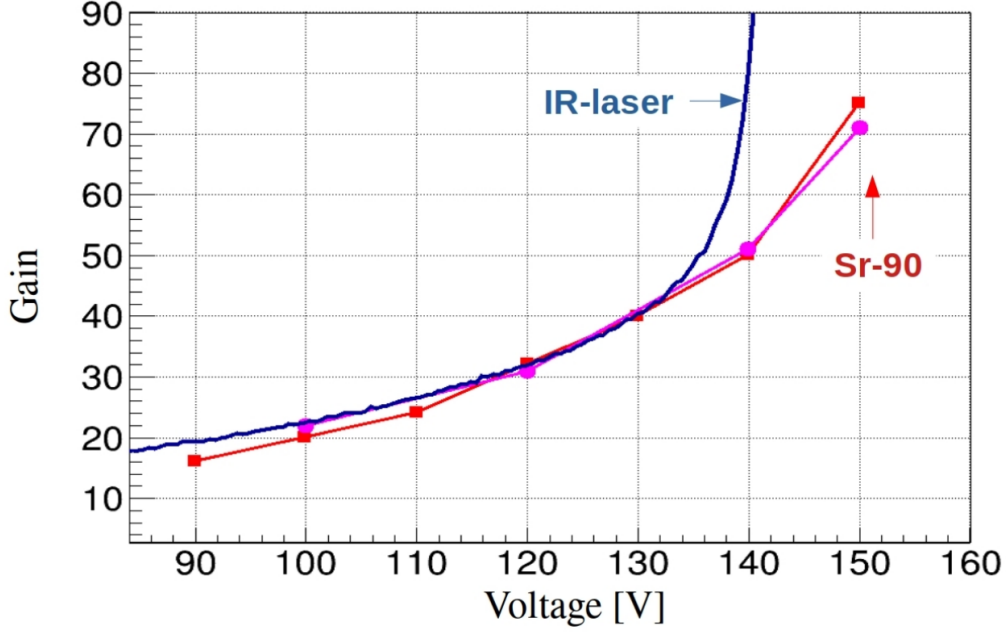
Figure 3.24: Gain as a function of voltage shown for two different charge injection methods: photons & mips. For low values of gain, the two methods are in very good agreement. As the gain reaches a value of ~45 and above, there is a large discrepancy due to gain suppression. Taken From [45]

3.5   Beta Source Coincidence Timing Technique

While a laser is ideal for injecting a consistent amount of charge, this is not the reality of what LGADs will see at the HL-LHC. As discussed in Section 2.2.2.1 and Section 2.5.1, the energy deposition of mips fluctuates along its path causing an additional contribution to the time resolution. Time resolution can be measured by placing two LGADs in coincidence with a source of mips, such as strontium-90 (Sr-90). Beta particles are emitted from the 150 MBq source and interact with both LGADs, as shown in Figure 3.25.

3.5.1   Set-up

The two LGADs are glued with double-sided tape to the PCB shown in Figure 3.26. They must be carefully aligned with a 1 mm diameter hole drilled through the main pad of the PCB. This hole allows beta particles to travel to both LGADs with
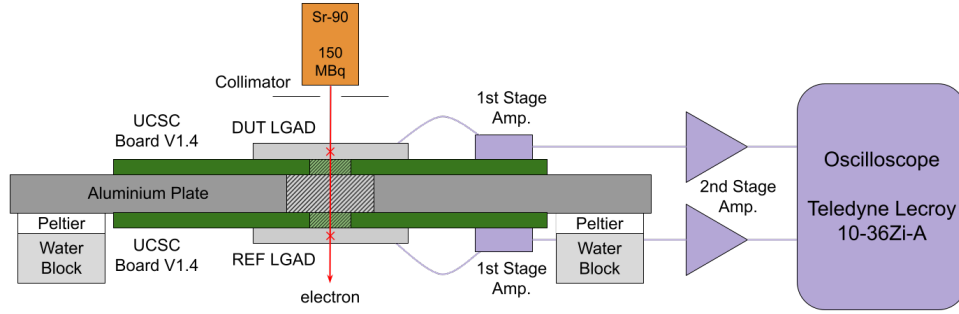
Figure 3.25: The experimental set-up used to measure the time resolution of a sensor by placing it in coincidence with a device of known time resolution.

minimal multiple Coulomb scattering [16].

These PCBs sandwich a 1 cm thick aluminium plate which has been milled in order to allow the PCB to lay flat, maximising contact with the thermal mass and without making electrical contact with any high voltage components. A 1 cm diameter hole is also required to allow beta particles to travel between the LGADs. Four threaded bars are attached to the aluminium plate and allow the PCBs to be secured and accurately aligned with each other. A kapton sheet is also sandwiched between the reference PCB and the aluminium plate to prevent a ground loop forming between the two boards. In standard operation the reference PCB is the bottom of the two, allowing the PCB on top, with the DUT, to be easily swapped.

3.5.2   Readout & Operation

This PCB shown in Figure 3.26 was developed by the University of Santa Cruz and uses a trans-impedance amplifier to integrate the current pulse while preserving timing information. This is the only case where the guard ring on the LGAD is grounded. While this can normally be ignored for other measurements for simplicity, the guard ring helps to shape the electric fields within the LGAD while replicating the operating conditions in the final detector. A second-stage amplifier, Particulars

Figure 3.26: Developed by the University of Santa Cruz, this PCB is used for time
resolution measurements [46].

AM-01 B 35 dB, follows the PCB to provide additional gain. Since two amplifiers
are required for the two LGADs they are both biased at 10 V, which ensures the gain
is saturated and thus the same between the two. The final waveform is recorded by
a Teledyne Lecroy 10-36Zi-A oscilloscope which is able to trigger on both LGADs
and saves the waveforms to a local hard drive.

A python script has been developed to automatically perform a voltage sweep. At
each voltage step, the script collects 2048 events. This provides sufficiently high
statistics, although a smaller number is acceptable in order to provide a faster mea-
surement. The exact list of voltages and number of events per step is determined
by hand. This list is checked with a short calibration script which can measure and
extrapolate how long the full sweep will take. For lower bias voltage the gain is
lower and thus some counts begin to be lost in noise. This reduces the count rate
as shown in Figure 3.27 and means that the total number of counts collected at this
voltage step should be reduced to allow a sensible data collection time.
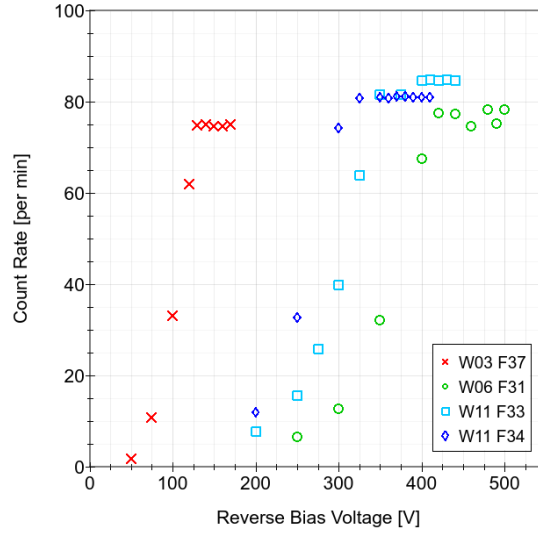
Figure 3.27: Example showing the count rate of an LGAD as a function of bias voltage. There is a clear linear increase before the count rate plateaus. Note how the curves are shifted for each sensor which is representative of their different gain characteristics.

3.5.3 Cooling

Two Peltier modules are placed either side of the aluminium plate to provide cooling down to $-20\,°C$. The advantage of the aluminium plate is its large thermal mass which uniformly cools both LGADs. This simplifies the procedure of the set-up by ensuring that both devices are always measured at the same temperature. Water blocks, cooled by a chiller unit, ensure the Peltier modules do not overheat and can achieve sub-zero temperatures. The entire set-up is placed inside a light- and air-tight container as shown in Figure 3.28. Nitrogen ensures the humidity is below $10\,\%\,$RH and ice cannot form anywhere within the container. To speed up the drying process, two pieces of ESD foam were added to minimise the volume of air which requires drying.

<center>(a)   (b)</center>
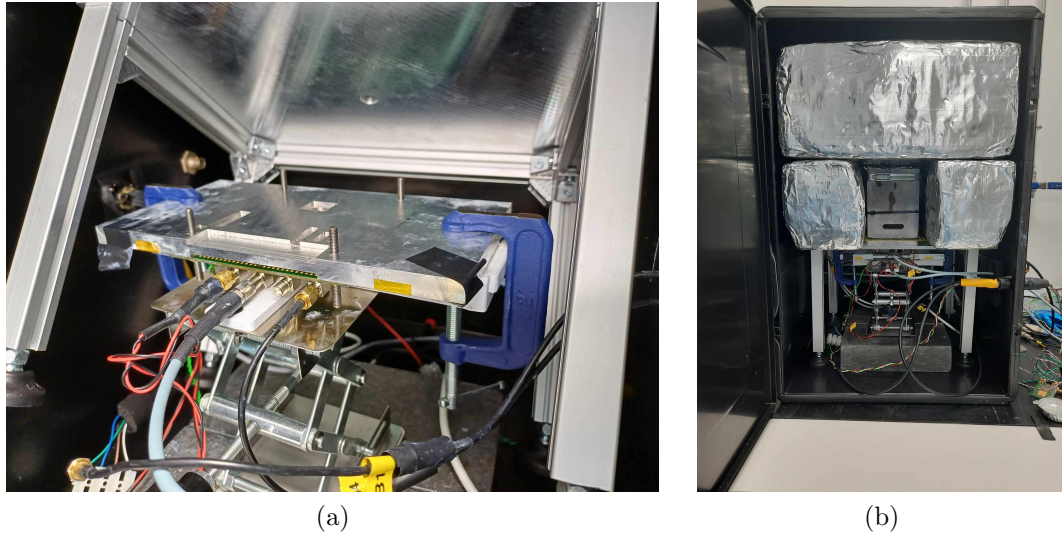
Figure 3.28: The setup is placed in a large black box which can be filled with nitrogen before cooling takes place. The box is also light-tight once closed.
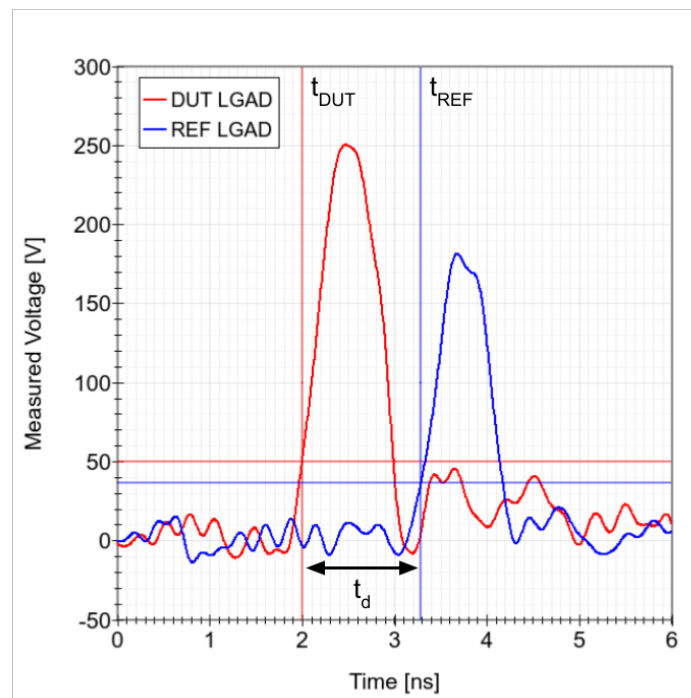
Figure 3.29: Example pulses from two LGADs in coincidence. They arrive with a particular time delay which can be measured.

3.5.4 Time of Arrival

The purpose of HEP-based LGADs is to provide precision timing information along-side good spatial resolution. This means that the time of arrival, t_{DUT} needs to be known with a sufficient resolution, σ_{DUT}. As discussed in Section 2.5.1, the time of arrival is determined using a constant fraction discriminator (CFD), as shown in Figure 3.29, which minimises the effects of time walk. In order to determine the optimal CFD fraction, a CFD sweep can be performed as shown in Figure 3.30 for two W3 devices at multiple voltages. The minimum for each curve lies between 10 % and 20 %. Ideally a smaller CFD fraction is preferred, but if too low, then the threshold can lie below the noise level seen in Figure 3.29. In this case, the measured arrival time will most likely be incorrect leading to a larger error and thus poor time resolution. A detailed study into this effect and how results at low gain are affected is beyond the scope of this work, but is important to consider when discussing time resolution results. To minimise this effect as well as possible while optimising the time resolution, a CFD fraction of 20 % was chosen for all subsequent measurements.

A measurement of t_{DUT} requires a reference point to have meaning, but radioactive decay is inherently random. Therefore, the reference time, t_{REF}, comes from the second reference LGAD. This LGAD is operated at a fixed voltage while the first is varied independently. Figure 3.29 shows the pulses from the two LGADs with an arbitrary time difference, t_d defined as,

$$t_d = t_{DUT} - t_{REF}. \tag{3.13}$$

Here, t_d is normally a function of the separation between the two LGADs. However in Figure 3.29, t_d is dominated by a difference in cable length running to the oscilloscope. A difference on the order of $0.5\,\text{m}$ results in a $1.7\,\text{ns}$ delay.

Subsequent measurements of t_d can be represented as a histogram as shown in
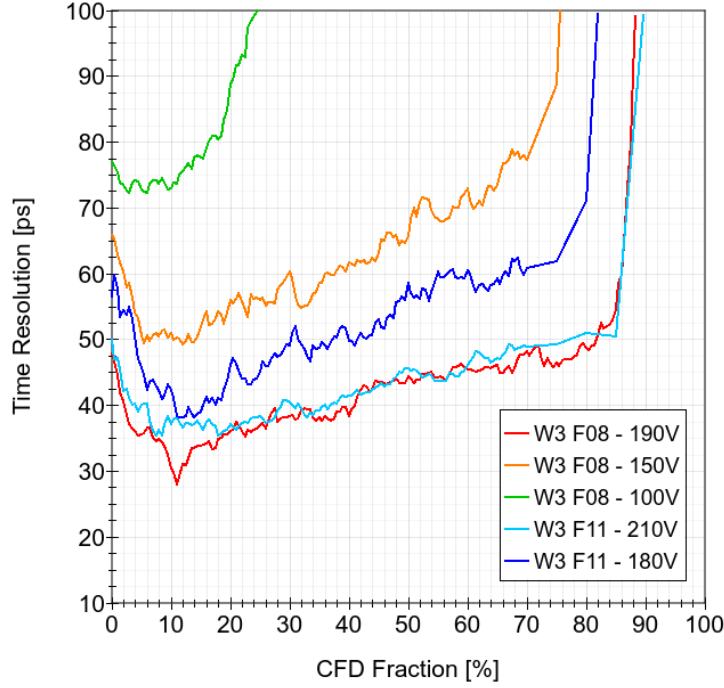
Figure 3.30: The calculated time resolution as the CFD fraction is varied. Two devices from wafer 3 are shown. Multiple voltages are also shown.

Figure 3.31. For a sufficient number of events, the distribution tends towards a Gaussian distribution with a standard deviation, σ_{MEAS}. The raw list of t_d can be fitted to a Gaussian using the "zfit" library for Python, extracting σ_{MEAS} [47]. The error associated with this, E_{MEAS}, can also be extracted.

3.5.5 Calculating Time Resolution

Since another LGAD is used for the reference time, its own time resolution, σ_{REF} contributes to σ_{MEAS} such that,

$$\sigma^2_{MEAS} = \sigma^2_{DUT} + \sigma^2_{REF}. \tag{3.14}$$

If an LGAD with known time resolution is used, then the calculation of σ_{DUT} in Equation 3.14 is trivial. If no reference LGAD is available, there are two approaches.
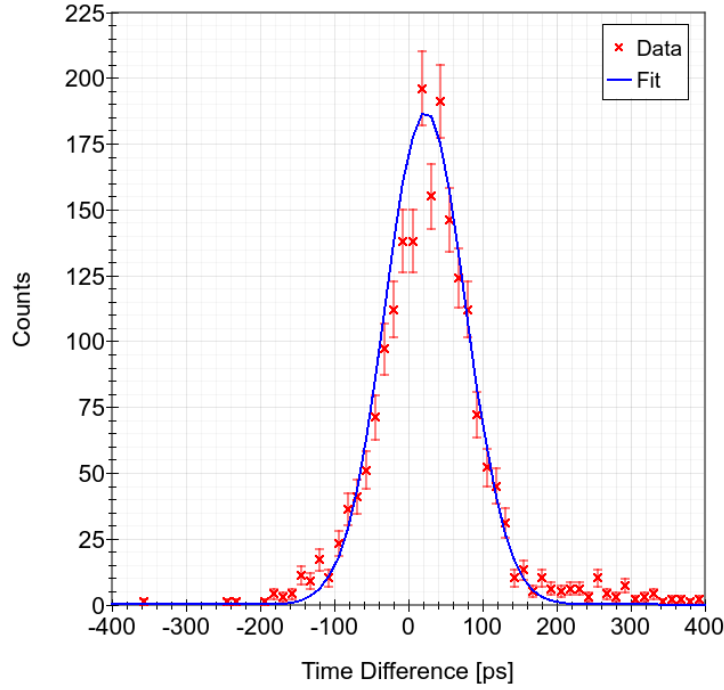
Figure 3.31: Histogram of time difference between two coincident LGADs. A Gaussian is fitted to the data and a standard deviation can be extracted and processed into a time resolution.

The first is to use identical LGADs and assume that they have an identical time resolution as a function of voltage. In this case, $\sigma_{REF} = \sigma_{DUT}$ and Equation 3.14 reduces to,

$$\sigma_{DUT} = \frac{\sigma_{MEAS}}{\sqrt{2}}. \tag{3.15}$$

However, this is not a reliable assumption. Even two LGADs which should be identical, tend to require a different voltage to achieve a similar value of gain and thus time resolution. This difference can be as large as $30\,\mathrm{V}$ and in this case makes the assumption in Equation 3.15 invalid. Instead, introducing a third LGAD with unknown time resolution allows the resolution of all three to be calculated. The three LGADs are denoted as A, B and C with time resolutions σ_A, σ_B and σ_C respectively. From Equation 3.14, the time resolution, measured pair wise, in all three possible pairings is given as,

$$\sigma_1^2 = \sigma_A^2 + \sigma_B^2, \tag{3.16}$$

$$\sigma_2^2 = \sigma_C^2 + \sigma_B^2, \tag{3.17}$$

$$\sigma_3^2 = \sigma_A^2 + \sigma_C^2. \tag{3.18}$$

Depending on how these equations are combined, each LGAD's time resolution can be extracted. For example,

$$\sigma_1^2 + \sigma_2^2 - \sigma_3^2 = 2\sigma_B^2. \tag{3.19}$$

Once these measurements have been performed once, one of these devices can then be selected to act as the reference device in Equation 3.14. Note that this reference value must be calculated separated for each operating temperature, including room temperature and $-20\,°\mathrm{C}$. This is because time resolution as a function of bias voltage changes with temperature. The time resolution of the three reference devices is shown in Table 3.4. LGAD A is chosen as the primary reference device, pre- and post-irradiation, since it has the smallest (best) time resolution and the associated error is similar between all three.

3.6   Cold Calibration For Gain and Timing Measurements

In order to ensure consistency between measurements, the temperatures of the DUT must be kept the same. To do this, the temperature needs to be monitored and this is done with a PT100 soldered to each PCB, for both set-ups, as shown in Figure 3.19 and Figure 3.26. Nominally, a PT100 has a resistance of 100 Ω at 0 $°C$

ID	Temp. [C]	Bias [V]	Time Res [ps]
A	+20	210	39.10 ± 0.58
B	+20	160	43.64 ± 0.52
C	+20	210	40.55 ± 0.55
A	-20	160	33.83 ± 0.54
B	-20	110	54.19 ± 0.34
C	-20	180	35.80 ± 0.51

Table 3.4: Three LGADs reference devices. Each has a specific bias voltage associated with the reference time resolution. Measurements were conducted at two different temperatures in preparation for characterisation before and after irradiation.

(a)   (b)
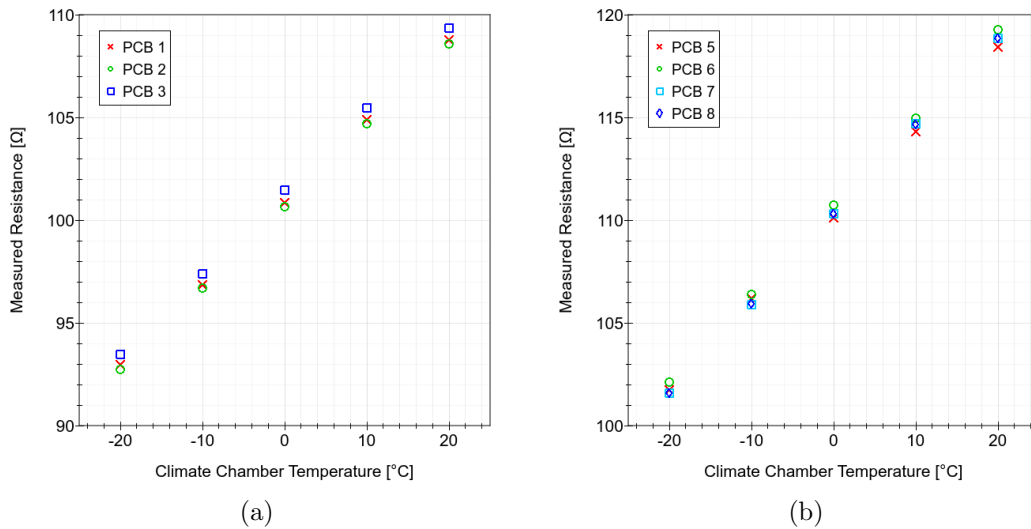
Figure 3.32: The PT100 on each PCB has been individually calibrated using the climate chamber. The PCBs are held at a selection of temperatures and the resistance is measured by their associated Arduino circuits. (a) TCT set-up (b) Timing set-up
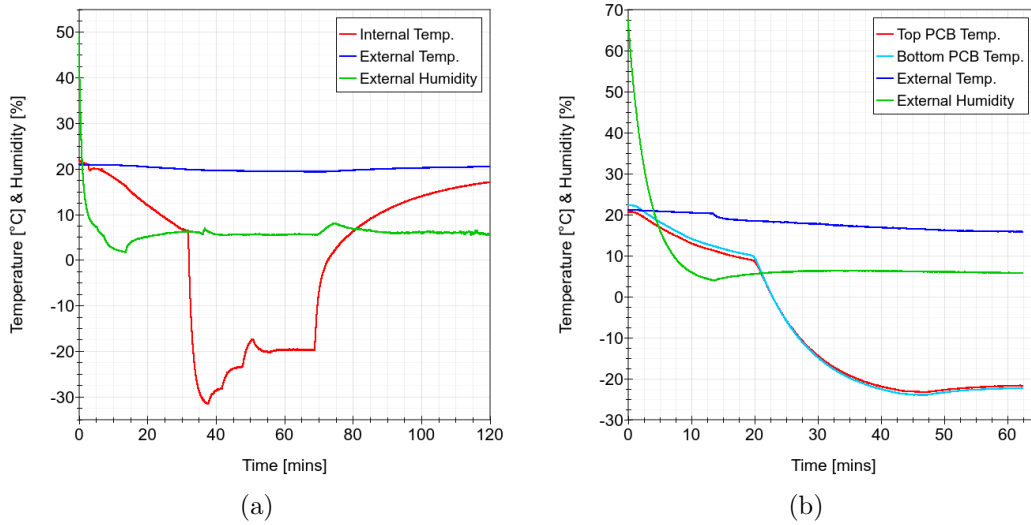
Figure 3.33: The cooling procedure for the (a) TCT set-up and the (b) timing set-up. Note that external refers to the conditions of the air surrounding the PCBs, while internal refers to the measurements on the PCB itself.

and varies linearly with temperature. However, manufacturing tolerances mean that this relationship is not perfect and a calibration is required. This was done using the climate chamber in which each PCB was tested at a range of temperatures from $+20\,^{\circ}C$ to $-30\,^{\circ}C$. Measured resistance as a function of temperature for each PCB is shown in Figure 3.32. From this, a linear fit can be extracted and used to convert the measured resistance into an accurate temperature in future measurements.

The experimental procedure for cooling both set-ups is shown in Figure 3.33. Firstly nitrogen is pumped into the test chamber. At the same time, the chiller is switched on and begins to cool to $2\,^{\circ}C$. Once the relative humidity is below $10\,\%\,\mathrm{RH}$ and the chiller has reached equilibrium, cooling via the Peltier modules can begin. Just before the 40 minute mark in Figure 3.33a, the current supplied to the Peltier modules is adjusted and the temperature is allowed to reach an equilibrium. This is repeated multiple times until the temperature reaches $-20\,^{\circ}\mathrm{C}$. At the 70 minute mark, the test has been successful and the chiller and Peltier modules are switched off. The DUT returns to room temperature after a further 50 minutes.

EXPERIMENTAL RESULTS

This chapter details the characterisation of Te2v's first LGAD batch. Wafer probing is first discussed for five wafer variants where IV and CV measurements have been performed on a selection of devices. After dicing, IV, CV, gain and timing measurements were then performed for four wafers. Gain measurements were preformed with laser while timing measurements were performed with mips. Finally, one wafer was selected for an irradiation campaign to investigate the acceptor removal effect. The five selected wafers, shown in Table 4.1, represent a wide range of energy and dose of the boron implant which provides gain.

4.1   Wafer Probing

Characterisation of the current and capacitance of all five wafers, shown in Table 4.1, is presented in this section. Three sizes, $1\,\text{mm}$, $2\,\text{mm}$ and $4\,\text{mm}$, of single pads are

Wafer #	Implant Energy	Implant Dose
20	1.00	1.00
11	1.07	1.05
16	0.92	1.05
6	1.00	1.11
3	1.07	1.11

Table 4.1: The five of Te2v's wafers selected for gain and timing measurements. The implant energy and dose have been normalised against W20.

(a)   (b)
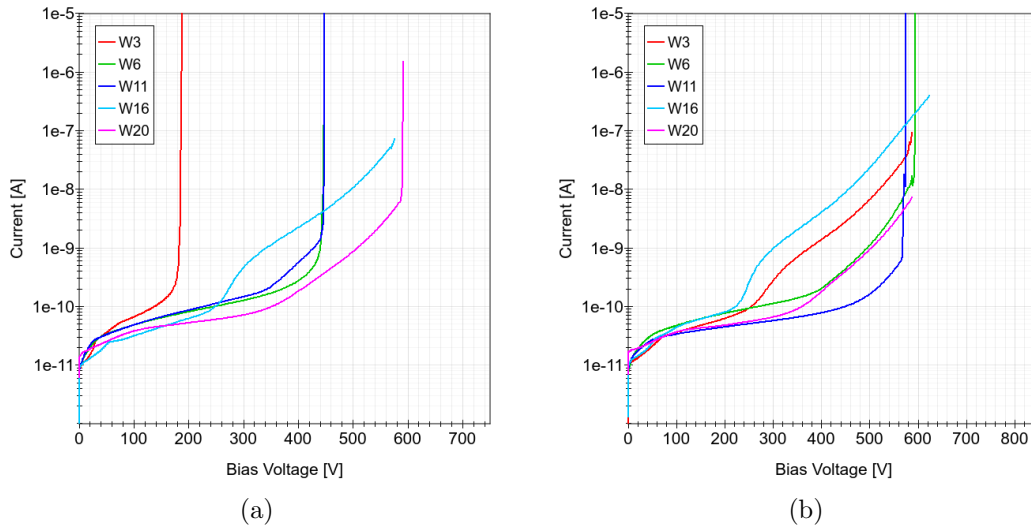
Figure 4.1: Example IV curves for one device from each wafer with a 1 mm pad size. (a) LGADs and (b) PiNs. The higher implant energy and dose, the lower the breakdown voltage. PiNs do not see this dependence. There is also some variability in the current, but this is not a concern.

tested. A sample of devices, up to five, for each combination of pad size and wafer type are tested. These devices are selected from different areas of each wafer as shown in Figure 3.1. Flavour D, as shown in Figure 3.2, is the focus of these tests. Not every combination of wafer or size, and sometimes only one device, could be tested. In each instance where these values are presented, the number of devices in the sample will be shown in brackets.
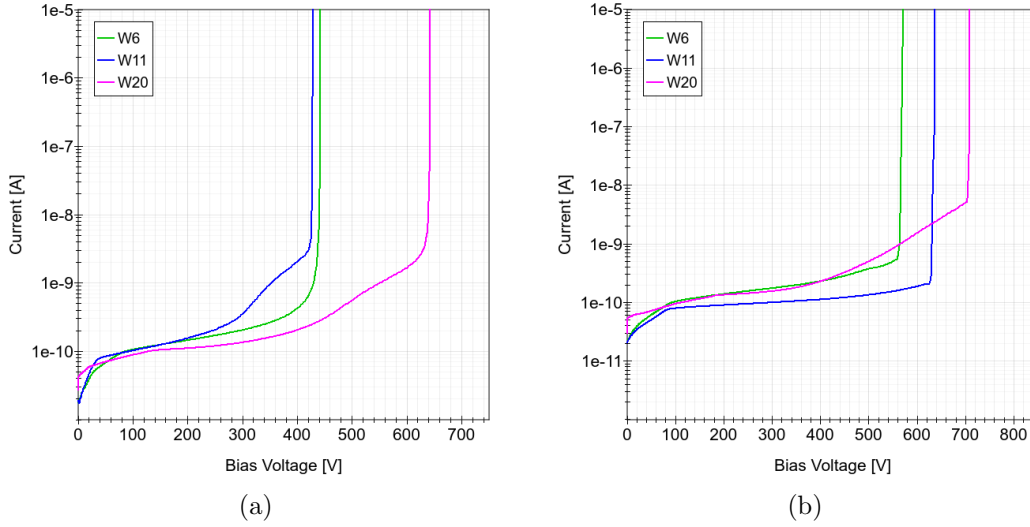
Figure 4.2: Example IV curves for one device from each wafer with a 2 mm pad size. (a) LGADs and (b) PiNs. The higher implant energy and dose, the lower the breakdown voltage. PiNs do not see this dependence, but there is some greater variability than in the 1 mm case.

4.1.1 IVs

The current as a function of bias voltage has been measured for all five wafers. Example IV curves for 1 mm LGADs are shown in Figure 4.1a. W3 (high implant energy and dose) clearly has the lowest breakdown voltage. W16 and W20 (low implant energy and dose) experience the highest breakdown voltage. While W6 and W11 (intermediate implant energy and dose) lie somewhere in the middle. It has been observed through testing that ambient humidity fluctuations can cause fluctuations in the current and this may explain the fluctuations seen here for LGADs and PiNs. Figure 4.1b shows the IV curves for 1 mm PiNs which should all be the same. Indeed they all experience breakdown at a similar voltage, however their current profiles vary greatly. From testing multiple devices, a current which exceeds $1\,\mu$A early (before breakdown) indicates a device is malfunctioning. Even with these variations in current profile, almost none of the 1 mm PiNs exceed such a current until breakdown.

IV curves for 2 mm LGADs are shown in Figure 4.2a. W3 and W16 were not
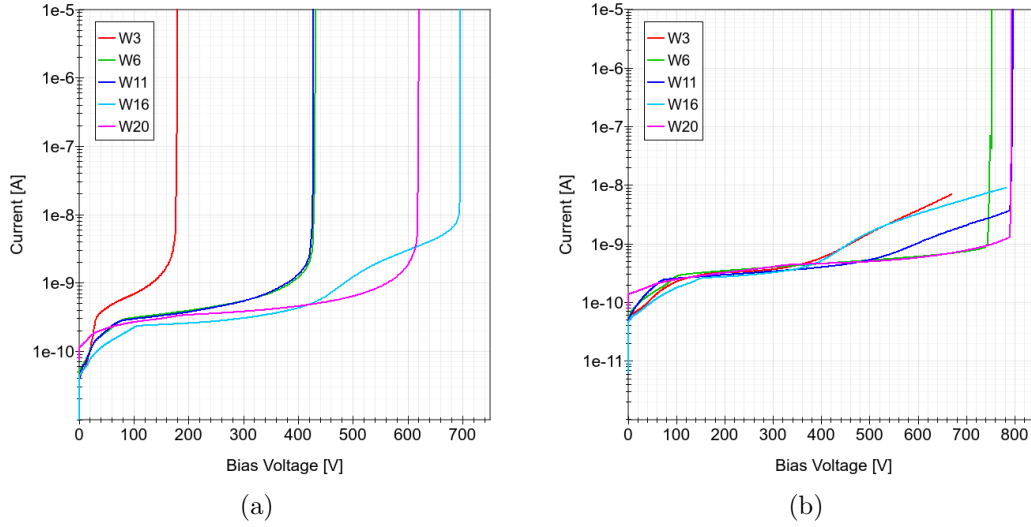
Figure 4.3: Example IV curves for one device from each wafer with a 4 mm pad size. (a) LGADs and (b) PiNs. The higher implant energy and dose, the lower the breakdown voltage. PiNs do not see this dependence.

tested here. In the context of this work, 1 mm is the focus since it is closer to the specification of ATLAS and CMS [5, 6]. For the remaining wafers, their breakdown voltage is similar to the 1 mm case. In the case of PiNs, shown in Figure 4.2b, there is much more variability in breakdown voltage, with some devices reaching a higher breakdown voltage than in the 1 mm case.

For all five wafers, 4 mm LGADs were tested and their IV curves are shown in Figure 4.3a. All wafers behave similarly to the 1 mm and 2 mm case. The exception is W16 (low implant energy and dose) which now experiences breakdown at a higher voltage. As Figure 4.3b for the 4 mm PiNs shows, irrespective of wafer number, in general the larger the PiN area, the higher the breakdown voltage.

Some of these devices, in Figure 4.3b for example, do not appear, visually, to reach breakdown and indeed numerically they do not reach the 200 nA definition. These are examples of devices which were tested with an older software. The voltage range had to be manually set and care had to be taken to ensure the final voltage did not result in a very high current such as 1 mA. Hence these voltage sweeps can stop early if the voltage range was set too low. In these cases, the last tested voltage is

Wafer # (Normalised	V_{bk} [V]		
Implant Energy & Dose)	1 mm	2 mm	4 mm
W03 (1.07 & 1.11)	192 ± 4 (2)	(0)	179 ± 5 (3)
W06 (1.00 & 1.11)	453 ± 11 (5)	461 ± 9 (5)	449 ± 8 (5)
W11 (1.07 & 1.05)	448 ± 6 (5)	437 ± 11 (5)	435 ± 4 (5)
W16 (0.92 & 1.05)	585 ± 5 (3)	(0)	689 ± 10 (3)
W20 (1.00 & 1.00)	580 ± 6 (5)	636 ± 11 (5)	625 ± 8 (5)

Table 4.2: The breakdown voltage, V_{bk} extracted for each LGAD. The average is shown for each combination of wafer type and pad size. As seen with the IV curves, the higher the implant energy and dose, the lower the breakdown voltage. There is no dependence on pad size, except for W16 and W20 (low implant energy and dose) where the 1 mm breakdown is lower.

taken to be V_{bk}. This is arguably not so unreasonable since each device was tested manually to find breakdown, however it does introduce and additional source of systematic errors which must be taken into account.

The extracted breakdown voltages are summarised in Table 4.2 for LGADs. For each combination of wafer and pad size, the average of V_{bk} is shown. The small number of devices tested is shown in brackets in Table 4.2. The trend in breakdown voltage is clear. W3 (high implant energy and dose) has the lowest breakdown voltage. W16 and W20 (low implant energy and dose) have the highest breakdown voltage. W6 and W11 (intermediate implant energy and dose) experience breakdown somewhere in the middle with overlapping standard deviations. For W3, W6 and W11, the change in pad size does not significantly affect the breakdown voltage. However, W16 and W20 show a difference between the breakdown of 1 mm and the 2 mm and 4 mm devices. This dependence on pad size is present for PiNs for all wafers as shown in Table 4.3. The LGAD results show that this trend disappears with the introduction of a significant gain layer, such as in W3, W6 and W11. This is likely because breakdown takes place in high-field region in the gain layer and is independent of the pad size. With a less doped gain layer, such as in W16, W20 and PiNs (which has no gain layer), breakdown is dominated by the high-field region between the n-well and the p-stop. As the pad size increases, so does the distance

| Wafer # (Normalised | Vbk [V] | | |
Implant Energy & Dose)	1 mm	2 mm	4 mm
W03 (1.07 & 1.11)	585 ± 1 (3)	(0)	685 ± 9 (3)
W06 (1.00 & 1.11)	597 (1)	576 (1)	755 (1)
W11 (1.07 & 1.05)	575 (1)	637 (1)	796 (1)
W16 (0.92 & 1.05)	593 ± 16 (3)	(0)	748 ± 36 (3)
W20 (1.00 & 1.00)	581 ± 7 (5)	693 ± 6 (5)	796 ± 5 (5)

Table 4.3: The breakdown voltage, V_{bk} extracted for each PiN. The average is shown for each combination of wafer type and pad size. There is no dependence on implant energy and dose as expected. However, there is now a dependence on pad size. The large variation in uncertainly is not expected and the reason is unclear.

between the n-well and the p-stop. For 1 mm devices, the distance is roughly $40\,\mu$m and doubles for each increase in device size. This increase in distance results in a lower electric field strength, and may lead to a higher breakdown voltage.

4.1.2   CVs

For these five wafers, the capacitance has also been measured as a function of voltage with a 100 kHz measurement frequency. Figure 4.4 shows example CV curves for all sizes of LGADs and PiNs. Wafers which experienced a higher breakdown voltage see an earlier gain layer depletion here. This is expected since both breakdown and depletion voltages depend on the implant energy and dose of the gain layer. W3 (high implant energy and dose) requires the highest voltage to deplete. W16 and W20 (low implant energy and dose) require the lowest voltage to deplete. W6 and W11 (intermediate implant energy and dose) deplete in the middle. It is also clear that the pad size influences the value of capacitance at each phase of depletion. This is expected since the capacitance of a parallel plate capacitor also depends on the area of the plates as shown in Equation 3.2. The same trend with pad size is clear for the PiNs.

These curves are characterised by extracting depletion voltages and the capacitance at which they occur. Table 4.4 shows the average of the gain layer depletion volt-
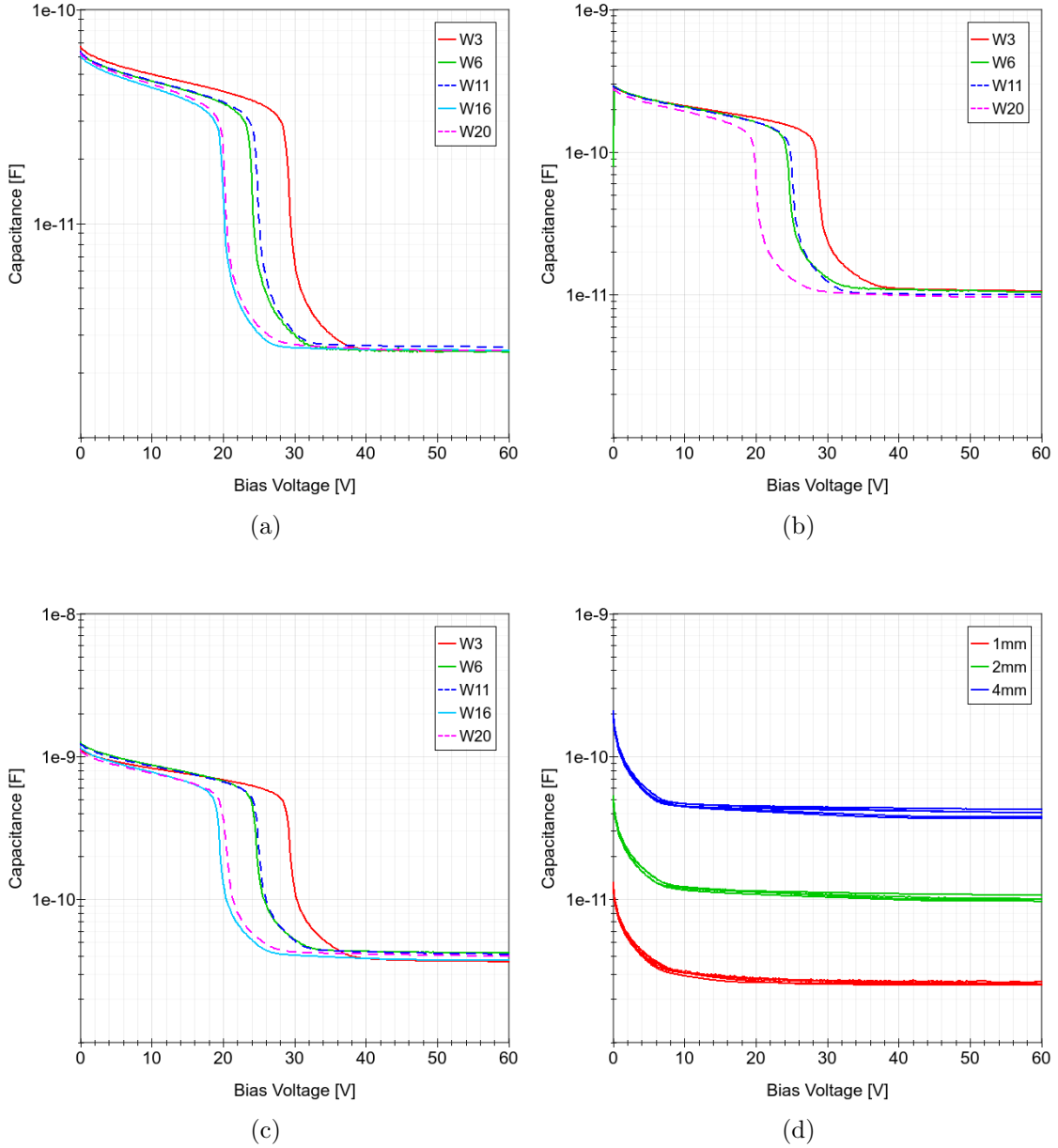
Figure 4.4: Example CV curves for (a) 1 mm LGADs (b) 2 mm LGADs (c) 4 mm LGADs and (d) all PiNs. All CV curves are very consistent and follow the trend that a higher implant energy and dose results in a higher depletion voltage. There is also a strong dependence on pad size. Both of these dependencies are expected

Wafer # (Normalised	V_{GLD} [V]		
Implant Energy & Dose)	1 mm	2 mm	4 mm
W03 (1.07 & 1.11)	28.2 ± 0.1 (3)	27.2 (1)	27.5 ± 0.1 (2)
W06 (1.00 & 1.11)	23.4 ± 0.4 (2)	22.6 ± 0.2 (3)	22.6 ± 0.2 (2)
W11 (1.07 & 1.05)	24.0 (2)	23.5 ± 0.3 (3)	23.3 ± 0.2 (4)
W16 (0.92 & 1.05)	18.7 ± 0.3 (3)	(0)	17.8 ± 0.1 (3)
W20 (1.00 & 1.00)	19.4 ± 0.2 (3)	18.6 (1)	18.5 ± 0.1 (3)

Table 4.4: The gain layer depletion voltage, V_{GLD} extracted for each LGAD. The average is shown for each combination of wafer type and pad size. The higher the implant energy and dose, the higher V_{GLD} is.

Wafer # (Normalised	V_{FD} [V]		
Implant Energy & Dose)	1 mm	2 mm	4 mm
W03 (1.07 & 1.11)	40.1 ± 0.3 (3)	35.6 (1)	45.2 ± 0.4 (2)
W06 (1.00 & 1.11)	37.3 ± 7.5 (2)	34.5 ± 2.4 (3)	31.4 ± 0.4 (2)
W11 (1.07 & 1.05)	35.5 ± 0.1 (2)	39.4 ± 2.4 (3)	36.2 ± 2.4 (4)
W16 (0.92 & 1.05)	32.1 ± 0.2 (3)	(0)	38.3 ± 2.7 (3)
W20 (1.00 & 1.00)	35.1 ± 2.7 (3)	40.2 (1)	36.3 ± 4.4 (3)

Table 4.5: The gain layer depletion voltage, V_{FD} extracted for each LGAD. The average is shown for each combination of wafer type and pad size. V_{FD} follows V_{GLD} and so sees the same dependence on implant energy and dose.

age, V_{GLD}, of LGADs for each wafer and pad size. Compared to V_{bk} the standard deviation on V_{GLD} is far smaller for all wafers, typically on the order of 0.3 V. There also appears to be no correlation between pad size and V_{GLD}, as expected. However, there is a trend across wafers where those with the highest combination of implant dose and energy have a larger V_{GLD}. This may be because a more highly doped gain layer, where a higher implant energy also gives a greater doping density at the gain layer as discussed in Section 4.2.3, requires a higher voltage to be depleted. W3 (high implant energy and dose) has the largest V_{GLD}. W6 and W11 (intermediate implant energy and dose) share a similar value of V_{GLD} with W11 being slightly higher. W16 and W20 (low implant energy and dose) have the smallest V_{GLD}. This is in line with the trend seen in V_{bk}, as expected.

The full depletion voltage, V_{FD} is shown in Table 4.5. The standard deviation varies

Depletion	Average Capacitance [pF]		
Stage	1 mm	2 mm	4 mm
0 V	60.4 ± 0.5 (13)	248.1 ± 1.7 (8)	1002.0 ± 12.3 (14)
(W03) V_{GLD}	29.1 ± 0.2 (3)	133.6 (1)	532.9 ± 0.1 (2)
(W06) V_{GLD}	28.0 ± 1.9 (2)	138.6 ± 0.4 (3)	585.0 ± 2.2 (2)
(W11) V_{GLD}	28.5 ± 0.1 (2)	133.1 ± 0.9 (3)	562.9 ± 4.0 (4)
(W16) V_{GLD}	29.2 ± 0.1 (3)	(0)	567.5 ± 2.3 (3)
(W20) V_{GLD}	29.7 ± 0.1 (3)	138.1 (1)	565.7 ± 1.0 (3)
V_{FD}	2.5 ± 0.1 (13)	10.9 ± 0.3 (8)	42.3 ± 1.0 (14)

Table 4.6: The average capacitance extracted at each phase of depletion. The average is shown for each pad size.

from about 0.3 V (W3) up to 10.6 V (W6) in the most extreme case. This reflects the difficulty in extracting V_{FD} from CV curves. Similarly to V_{GLD}, there is no strong correlation between pad size and V_{FD}, as expected. Any differences are likely to be due to the intrinsic difficulty in extracting V_{FD}. There appears to be a trend in terms of wafer type. One would expect that V_{GLD} and V_{FD} are separated by a constant, thus any trends in V_{GLD} are seen in V_{FD}. While this seems to be true for W3 (high implant energy and dose), the remaining wafers all have a comparable value of V_{FD}. It is not clear why the trend in V_{GLD} is not replicated in V_{FD} but it maybe be related to the large errors involved in extracting V_{FD}.

Table 4.6 shows the average capacitance at three key stages: initial depletion (0 V), gain layer depletion, full depletion. The capacitance at gain layer depletion was found to have no dependence on wafer type. This is not necessarily expected since a higher implant energy should lead to a deeper gain layer, resulting in a smaller capacitance when depleted. However, it may be an artefact of low statistics. The capacitance at 0 V and V_{FD} was found to have no dependence on the wafer type as expected and hence an average is shown here. There is a clear dependence on pad size, which is expected following Equation 3.2. The meaning of the capacitance can be better understood by using Equation 3.2 to calculate the depletion depth, d. This is shown in Table 4.7. Since the calculation takes the pad size into account, the depletion depth shows no dependence on the pad size. The depletion depth begins at

Depletion	Average Depletion Depth [μm]		
Stage	1 mm	2 mm	4 mm
0 V	1.72 ± 0.01 (13)	1.67 ± 0.01 (8)	1.66 ± 0.02 (14)
V_{GLD}	3.58 ± 0.04 (13)	3.05 ± 0.02 (8)	2.94 ± 0.02 (14)
V_{FD}	41.9 ± 2.5 (13)	38.17 ± 1.03 (8)	39.40 ± 0.88 (14)

Table 4.7: The depletion depth extracted at each phase of depletion. The average is shown for each pad size since no dependence on wafer type was found.

$\sim 1.7\,\mu$m before doubling after the gain layer depletes. At full depletion, the depth is roughly $40\,\mu$m. These values are generally expected, although the active thickness of these LGADs should be $50\,\mu$m. This is a significant difference but may arise due to lateral depletion. The guard ring was floating which means the bulk could deplete laterally as well as vertically. A small increase in the size of the pad (from 1.0 mm to 1.1 mm) would account for the discrepancy seen in Table 4.7. It should be noted that while no difference between wafers was found, this is not necessarily expected for the depth at gain layer depletion. A higher implant energy means the boron atoms can drive a deeper implant, but also a denser implant, as discussed in Section 4.2.3. However, these differences are likely to be much smaller than the statistical error from extracting the capacitance hence an average is taken.

For all of the relevant properties above, their respective values for PiNs are shown in Table 4.8. The full depletion voltage shows no correlation with pad size. However, the depletion voltage is not consistent with the difference between V_{GLD} and V_{FD}, measured for LGADs, as one might expect. This is inconsistent with the CV curves shown in Figure 4.4 where PiNs clearly deplete earlier than as reported in Table 4.8. This may indicate shortcomings or ambiguity in the extraction of V_{FD}. Lateral depletion might also create ambiguity in V_{FD}. Grounding the guard ring might mitigate this. The capacitance at 0 V and full depletion, C_i and C_{FD} respectively, show a linear dependence on the pad area similarly to LGADs. The depletion depths can again be calculated in each case and show no dependence on pad size. At full depletion, the depth is still expected to be the active thickness, $50\,\mu$m. Similarly to LGADs, the calculated depth is roughly $40\,\mu$m. This is also likely to be due

| CV | Average Value | | |
Parameter	1 mm	2 mm	4 mm
V_{FD} [V]	27.1 ± 2.3 (11)	33.0 ± 8.0 (4)	33.0 ± 3.6 (11)
C_i [pF]	5.0 ± 0.1 (11)	19.7 ± 0.4 (4)	74.7 ± 0.7 (11)
C_{FD} [pF]	2.68 ± 0.03 (11)	10.7 ± 0.6 (4)	41.1 ± 1.0 (11)
d_i [μm]	20.6 ± 0.2 (11)	21.1 ± 0.4 (4)	22.2 ± 0.2 (11)
d_{FD} [μm]	38.8 ± 0.5 (11)	39.1 ± 1.9 (4)	40.6 ± 0.9 (11)

Table 4.8: Various parameters extracted from the CV curve of a PiN. The average is shown for each pad size. Here, i represents the initial state of each parameter when at $0V$.

to lateral depletion. An increase in the pad size by $100\,\mu$m would be enough to cause this discrepancy. CV measurements while the guard ring is grounded would help to limit lateral depletion and confirm this explanation. More interestingly, the initial depth, d_i, is roughly $21\,\mu$m. This is much higher than for LGADs for which d_i is roughly $1.7\,\mu$m. This difference is caused by the gain layer which has a much higher doping concentration than the bulk region, meaning a smaller initial depletion thickness. The value of d_i for PiNs is in agreement with values predicted by theory in Table 2.1. A more detailed study could look to extract the effective doping concentration as a function of depth from CV curves, as described by Equation 3.2 and Equation 3.3.

4.1.3   Summary

A combination of five wafers and three pad sizes have been electrically characterised with their IV and CV curves analysed. Overall, it is clear that LGADs experience significant wafer non-uniformity when considering V_{bk}. Where there are sufficient statistics, the standard deviation was upwards of $20\,$V. This is also true for PiNs although as a percentage error there is slightly less variability. For all extracted values, CV curves seem to be much less sensitive to wafer non-uniformity than IV curves. W3 has the highest combination of gain layer implant energy and dose and this results in the lowest breakdown voltage and highest depletion voltages. W6 and

W11 follow with the second lowest V_{bk} and second highest depletion voltages. Their similarity is intriguing considering that they have a very different combination of implant energy and dose. The same is true for W16 and W20 which also exhibit similar breakdown and depletion voltages. A discussion of why W6 and W11 behave similarly can be found in the following section once gain measurements are introduced. As one of the worse performing wafers, W16 was not tested beyond this point.

4.2   Diced Devices

After a wafer has been diced, electrical tests need to be repeated to ensure no significant damage has been caused. Diced devices have been tested for gain and timing properties. This characterisation has been performed on a selection of 1 mm devices from four wafers shown in Table 4.1 excluding W16 to save time. The remaining wafers share implant energy or dose properties allowing their effects to be studied. In this section, a selection of devices from each wafer is tested. Flavour A, as shown in Figure 3.2, is the focus for these tests. This was changed from flavour D to increase the number of devices available for testing.

4.2.1   IVs

Current profiles for all four wafers have been measured and are shown in Figure 4.5. After dicing, some devices had an unusually low breakdown voltage, such as in Figure 4.5c. This, as well as the high variation generally seen for IV curves, was justification to apply a thermal annealing to each device after dicing. The details of this process are discussed in Section 3.2.2. IV curves measured after annealing are also shown in Figure 4.5.

The breakdown voltage, V_{bk}, for each IV curve has been extracted. The average and standard deviation of V_{bk} for each wafer before and after annealing is shown in
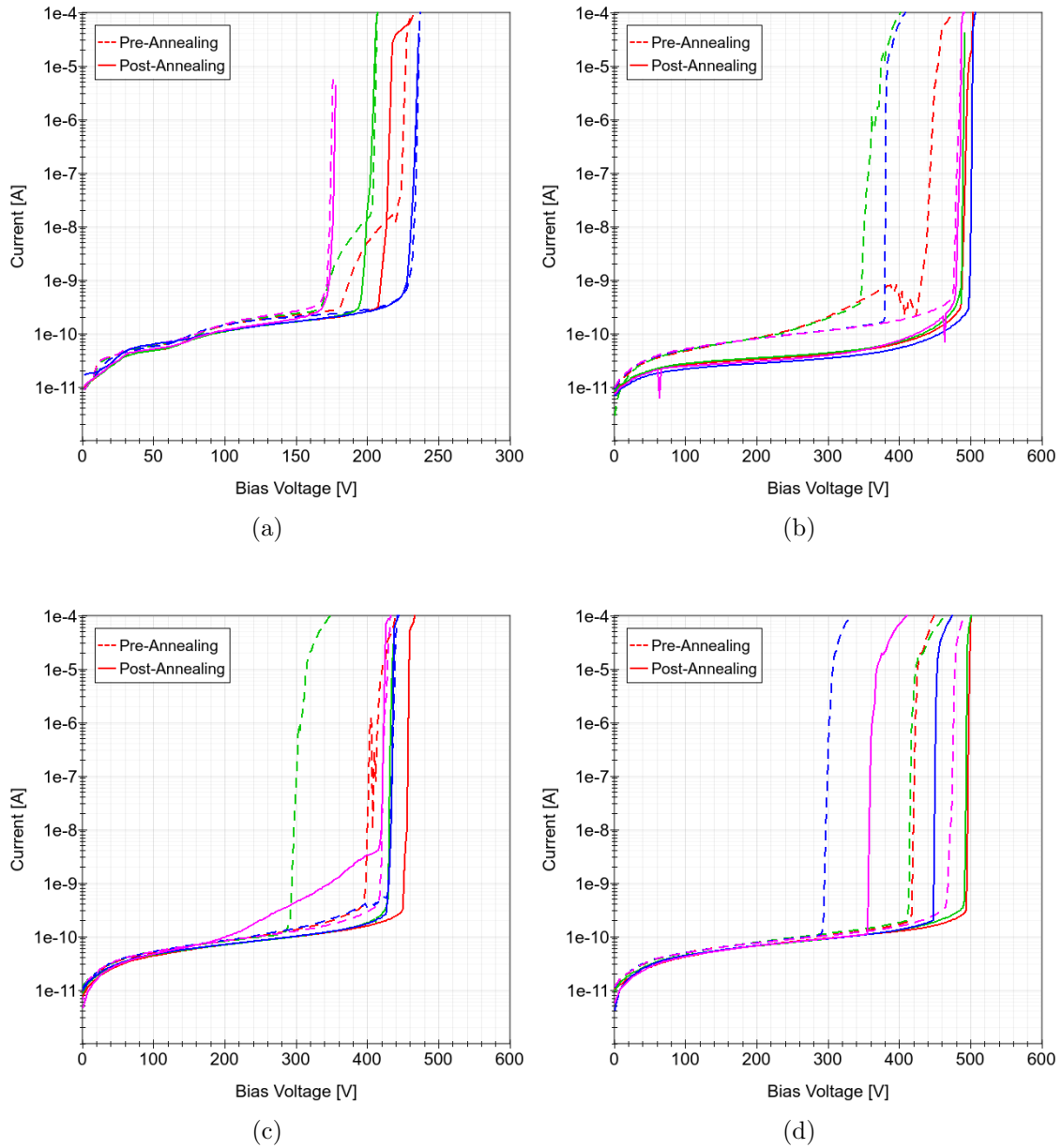
Figure 4.5: IV curves of a representative selection of 4 or 5 devices per wafer before and after thermal-annealing. (a) Wafer 3 (b) Wafer 6 (c) Wafer 11 (d) Wafer 20. Annealing shows a general improvement in breakdown voltage.

Wafer # (Normalised Implant Energy & Dose)	V_{bk} [V]		
	Pre-dicing	Post-Dicing	Post-Annealing
W03 (1.07 & 1.11)	192 ± 4 (2)	198 ± 7 (12)	200 ± 7 (12)
W06 (1.00 & 1.11)	453 ± 11 (5)	438 ± 11 (11)	471 ± 8 (11)
W11 (1.07 & 1.05)	448 ± 6 (5)	430 ± 10 (12)	440 ± 7 (12)
W20 (1.00 & 1.00)	580 ± 6 (5)	457 ± 15 (12)	482 ± 10 (12)

Table 4.9: For each device V_{bk} can be extracted. The average for each wafer is shown in three cases: Pre-dicing, Post-Dicing, Post-Annealing. The average breakdown voltage is improved after annealing. The standard deviation also improves.

Table 4.9. For all wafers, the average V_{bk} increases after annealing although only by a few volts. The standard deviation also decreases. Although these changes are small, they suggest that thermal annealing generally improves the performance of LGADs and reverses at least some of the damage caused by dicing. The average breakdown voltage for each wafer pre-dicing is also shown in Table 4.9. The standard deviation is lower in this case. While this suggests that dicing has caused this rise, it should be treated with caution. This is because the statistics on pre-dicing measurements is particularly poor in comparison to post-dicing results where more devices were tested. This is shown in Table 4.9 where the number of devices is shown in brackets. The same can be said for comparisons between laser and saw dicing. There appears to be no strong difference between each method, however a more detailed study is really required. Generally speaking, the average breakdown voltage pre-dicing is in agreement with post-dicing and post-annealing values. This is with the exception of W20 where the breakdown voltage has decreases by $100\,\text{V}$.

In addition to Table 4.9, the change in breakdown voltage after annealing on a case-by-case basis has also been calculated and is shown in Table 4.10. For all wafers the average change in V_{bk} is positive and for W6 and W20 it is also large. However, the standard deviation for each wafer is more than $100\,\%$ which means that some devices see a decrease in breakdown voltage. This is clear visually in Figure 4.5d where V_{bk} for one device is reduced by over $100\,\text{V}$. While thermal annealing seems to improve the IV performance of LGADs statistically, this does not mean that every

Wafer # (Normalised Implant Energy & Dose)	Change in V_{bk} [V]
W03 (1.07 & 1.11)	2 ± 2 (12)
W06 (1.00 & 1.11)	32 ± 14 (11)
W11 (1.07 & 1.05)	9 ± 11 (12)
W20 (1.00 & 1.00)	25 ± 16 (12)

Table 4.10: For each device, the difference in V_{bk} before and after thermal annealing can be calculated. The average for each wafer is shown here. While on average the breakdown voltage improves, some individual devices do see a decrease in breakdown voltage after annealing.

device sees an improvement.

It is also worth noting that most devices maintain a low leakage current before breakdown, rarely exceeding $1\,\mathrm{nA}$. Almost all the devices which have an abnormal current profile, such as those in Figure 4.5a, are rectified by the thermal annealing process.

4.2.2   CVs

Given that CV curves are generally very consistent, only a selection of devices and wafers had their CV curves measured. This was to confirm that the thermal annealing process does not cause any significant changes to the CV profiles. Figure 4.6 shows CV curves for W3 and W6. In addition, an example CV curve measured pre-annealing and pre-dicing is also shown. There is excellent agreement between the pre-dicing and post-annealing CV curves. The pre-annealing curve also is in agreement, but in the W3 case it does deviate in the region where the gain layer is starting to deplete. This deviation translates to a thicker initial depletion depth but this anomalous behaviour is rectified by thermal annealing.
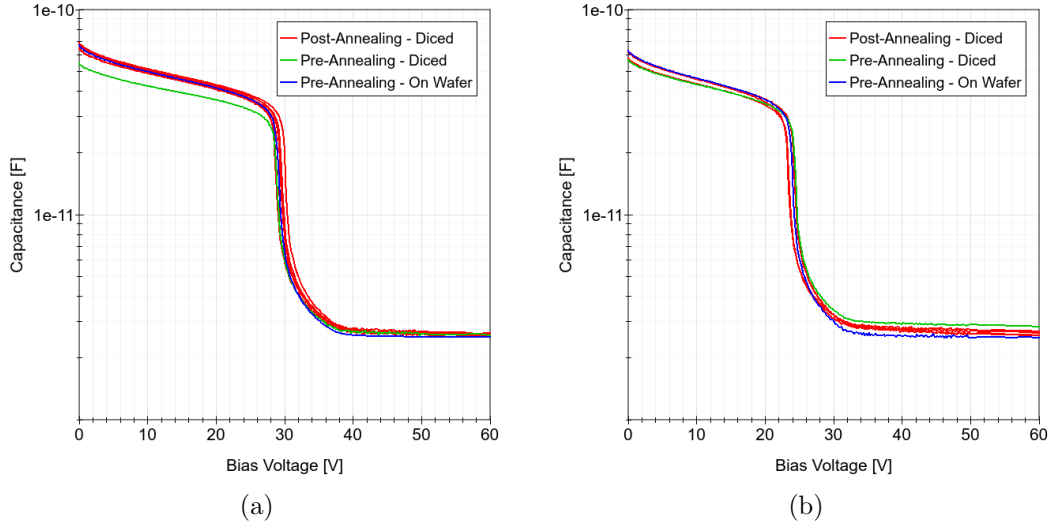
Figure 4.6: Example CV curves shown for three scenarios: pre-dicing, post-dicing, post-annealing. (a) Wafer 3 (b) Wafer 6. Annealing generally has no significant effect on capacitance, as expected

4.2.3   Gain

Figure 4.7b shows gain as a function of voltage for four devices from each wafer. It is clear that each wafer has significant non-uniformity shown by the way their gain depends on bias voltage. This is with the exception of W20 which barely exceeds a gain of 10 before reaching breakdown. For the other wafers, nearly all devices exceed a gain of 50 while most also exceed a gain of 100. There is some significant variation in the maximum achievable gain. This is not necessarily a breakdown limitation, but rather an experimental issue. Before the bias sweep takes place, the maximum bias voltage must be manually tested to ensure that the current draw and pulse size is not too large such that it might damage electronics. This also requires an optimal choice in ND filters (discussed in Section 3.4.1) to allow the larger voltage to be tested while preserving information at the lower voltages where the pulse height approaches the noise level. The achievable gain may be higher, but the need to avoid risk of damage to the electronics prevents pushing the operating voltage to the limit.

Considering the bias voltage variability, Figure 4.8 better characterises this. It shows
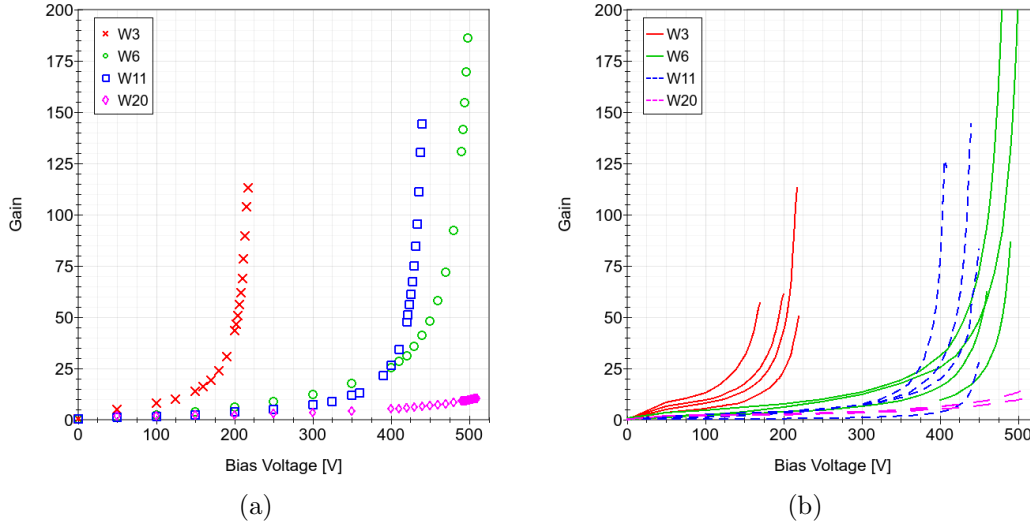
Figure 4.7: (a) Sample (one device per wafer) of gain curves shown to improve clarity and highlight the difference between wafers. (b) All gain curves shown to highlight variability. Higher implant energy and dose leads to high gain at a lower bias voltage, as expected.

the average bias voltage required to achieve a particular value of gain. W6 and W11 have a comparable profile for small values of gain, but for high gain (> 125) they begin to diverge, ending at $(489 \pm 10)\,\text{V}$ and $(422 \pm 16)\,\text{V}$ respectively. They also both begin with a significantly larger error on the order of $40\,\text{V}$. However this is largely a consequence of the initial plateau region in Figure 4.7 where small changes in gain require large alterations to the bias voltage. This variation is hardly seen with W3 which starts with a standard deviation of $8.6\,\text{V}$. The final data point also has a small standard deviation of $5.9\,\text{V}$ but this is because only two devices can achieve a gain of 60 and they happen to align quite well. For the rest of the curve, the standard deviation is fairly consistent and the bias voltage as a function of gain plateaus to a value of $(196 \pm 11)\,\text{V}$.

The similarities between W6 and W11 is interesting since they both share implant properties with W3 which achieves high gain at a significantly lower bias voltage than either W6 or W11. W6 shares its implant dose with W3, while W11 shares its implant energy with W3. Figure 4.7 suggests that a reduction in either the implant energy or dose, with respect to W3, increased the bias voltage required to achieve a
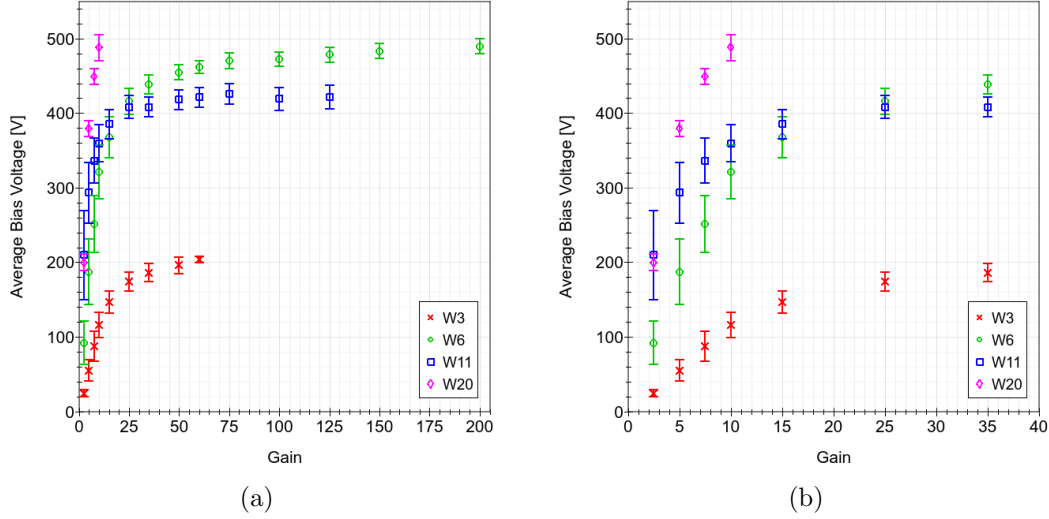
Figure 4.8: (a) The average bias voltage required to achieve a specific gain for each wafer. (b) Zoomed on the lower voltage region. There is a large variance between wafers in terms of the bias voltage required to achieve a specific value of gain.

high gain. Whilst one might expect the implant dose to be the main driver of the gain layer doping concentration, these results suggest the energy also plays a significant role. Figure 4.9 shows SIMS and technology computer-aided design (TCAD) results from a device from Wafer 2 (equivalent to W3) [48]. It shows significant overlap between the n and p-type regions. An increase in implant energy means the peak of the p-type region shifts and less of the region overlaps with the n-type region. While this does change the shape of the implant slightly, it also results in an increase in the doping concentration and hence a higher gain at a lower bias voltage. Further SIMS measurements are required to understand this effect in more detail.

4.2.4   Timing

All devices, including some additional devices for W11, have had their time resolution measured as a function of bias voltage using the Sr-90 set-up described in Section 3.5. These are shown in Figure 4.10. In terms of timing performance, W3 can reach sub-40 ps while W6 and W11 can reach sub-35 ps with some devices even below 30 ps. W20 has a considerably worse timing performance remaining above 50 ps.
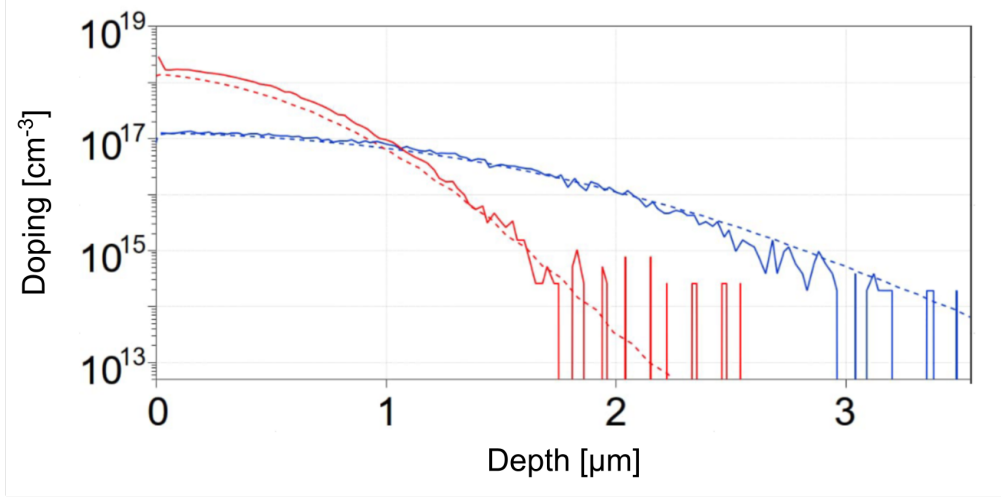
Figure 4.9: SIMS measurements (solid line) and TCAD simulations (dotted line) of the doping concentration in the n-well (red) and the p-type gain layer (blue). This is performed using wafer 2 (high energy and dose equivalent to W3). n- and p-type regions have considerable overlap. Taken from [48]
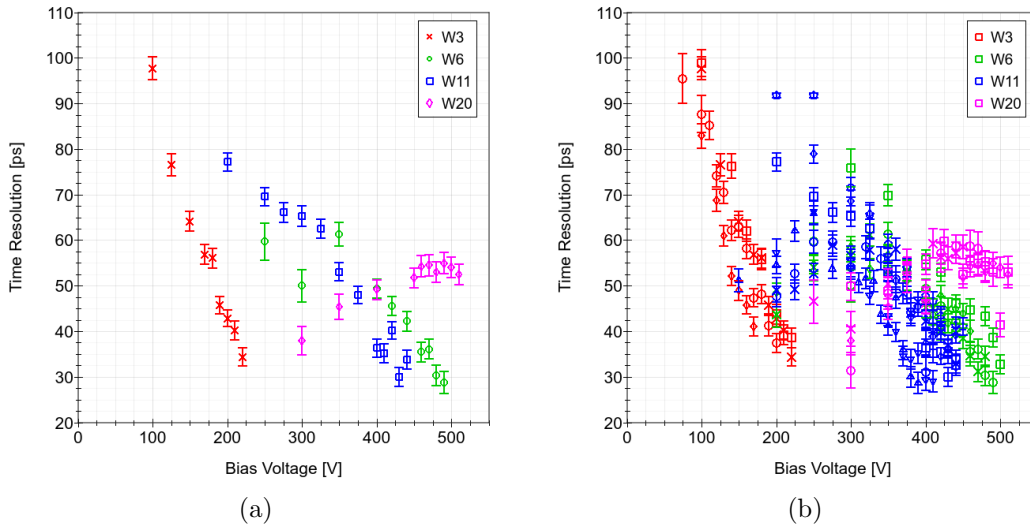
Figure 4.10: Time resolution as a function of bias voltage. (a) Sample (one device per wafer) shown to improve clarity and highlight the difference between wafers. (b) All devices shown to highlight variability. Higher implant energy and dose leads to a better time resolution at a lower bias voltage, as expected.

This aligns with its relatively low gain even at a bias voltage close to breakdown. Within each wafer there is significant variability with respect to the dependence on bias voltage. This relates to the variability in gain seen in Figure 4.7. This variability is characterised in Figure 4.11a where the average bias voltage required to achieve a particular time resolution is shown. Similarly to gain W11 requires a slightly lower bias voltage to achieve the same time resolution as W6, although the standard deviations overlap so this difference is not significant. Of course, W3 continues to require a much lower voltage for the same time resolution, reflecting its high gain at a lower bias. In the regions of good time resolution, the standard deviation is on the order of 18 ps. For lower values of time resolution, the standard deviation is also smaller. These deviations are modest, although W3 in particular seems to have the least variability. However, Figure 4.11b shows the average time resolution at a particular bias. This is important since in a real detector, all LGADs will be biased with the same voltage and thus any variability in the time resolution will impact the performance of the overall detector. In terms of standard deviation, this is quite large in general with W11 having a standard deviation on the order of 15 ps. W3 and W6 are slightly better, but not significantly. This variability is not ideal and again comes from the variation in gain profiles, which in turn depends on wafer uniformity.

4.2.5   Gain & Timing

Combining the time resolution and gain measured at a particular bias voltage is a good way to normalise against any variation in the gain profile, whether from across or within wafers. This is shown in Figure 4.12 where data for an LGAD manufactured by HPK, a more establish LGAD vendor, are also shown. The variation of time resolution as a function of gain is much less pronounced than as a function of voltage. This is true for all fields and wafers and is shown in Figure 4.13b where the average time resolution as a function of gain is shown. There is considerable overlap between all wafers implying that the relationship between time resolution

Figure 4.11: (a) Average bias voltage required to achieved a desired time resolution. (b) Average time resolution as a function of bias voltage. When all devices are biased at the same bias voltage (such as in the HL-LHC), the variance in the time resolution is large.
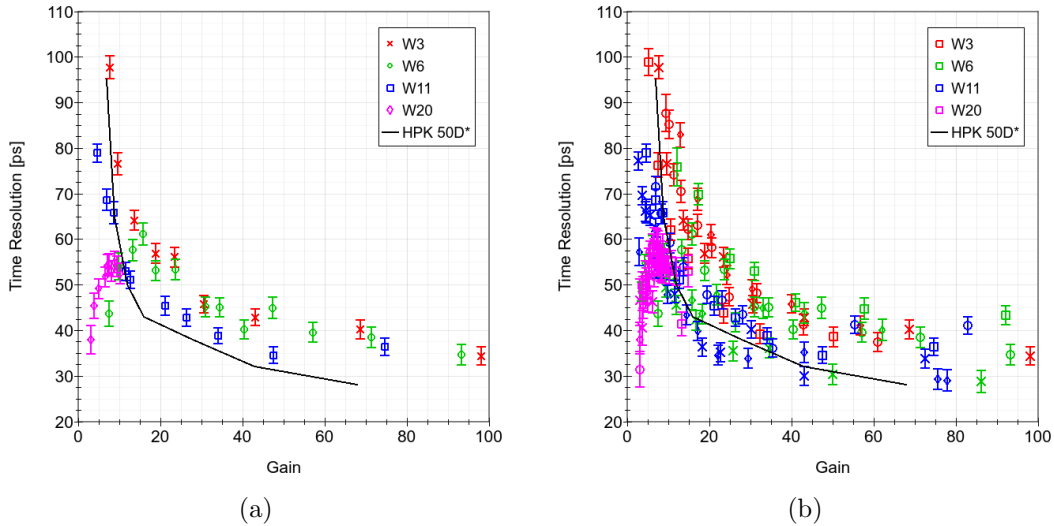
Figure 4.12: Time resolution as a function of gain. (a) Selection shown to improve clarity and highlight the difference between wafers. (b) All devices shown to highlight variability. The dependence on implant energy and dose is not visible in this case.
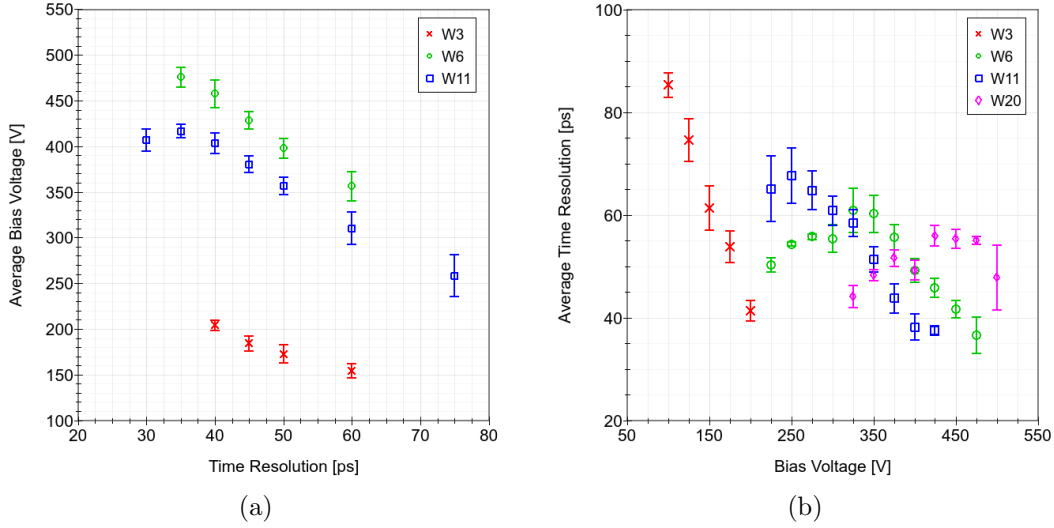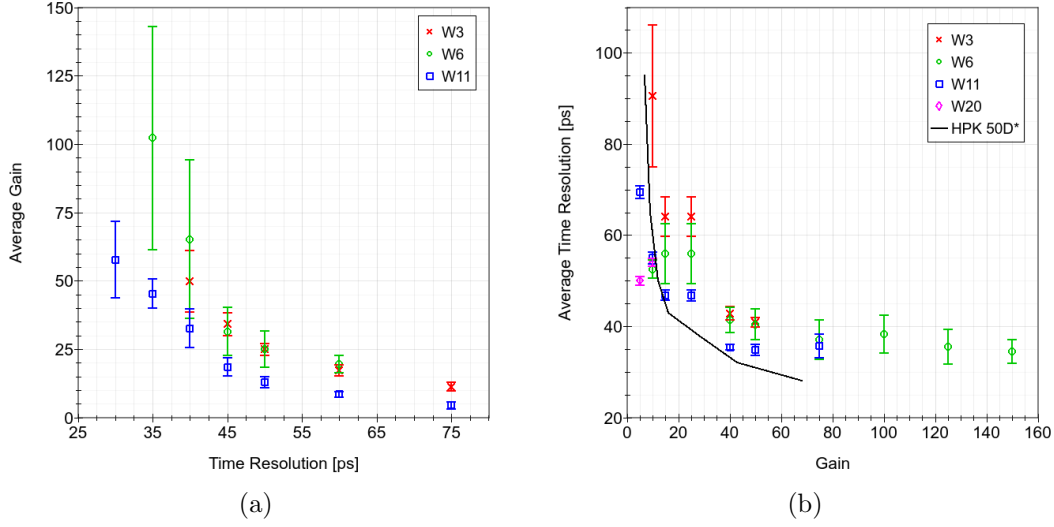
Figure 4.13: (a) Average gain required to achieved a desired time resolution. (b) Average time resolution as a function of gain. The lack of implant energy and dose dependence is highlighted here.

and gain is fairly universal regardless of the specific details of the gain layer. This is supported by the overlap with the data from HPK. The standard deviation for all wafers decreases as gain increases. At moderate to high gain, the standard deviation is about 5 ps for W6 and W11, whereas for W3 it is closer to 2 ps. This is an improvement on the variation as a function of bias voltage shown in Figure 4.11b. Note that W3 does see a very high standard deviation at a gain of 10. This is possibly a consequence of the low bias voltage required to achieve a gain of 10, meaning jitter is not minimised through a fast rise time. Figure 4.13a shows the average gain required for a particular time resolution and the standard deviation is shown to get much worse as the time resolution improves, especially for W6. Once more, this is because as the time resolution improves, it plateaus and the required gain is very sensitive to small changes in time resolution. Overall these plots show good agreement and a consistent time resolution dependence on gain, with no significant dependence on the type of wafer.
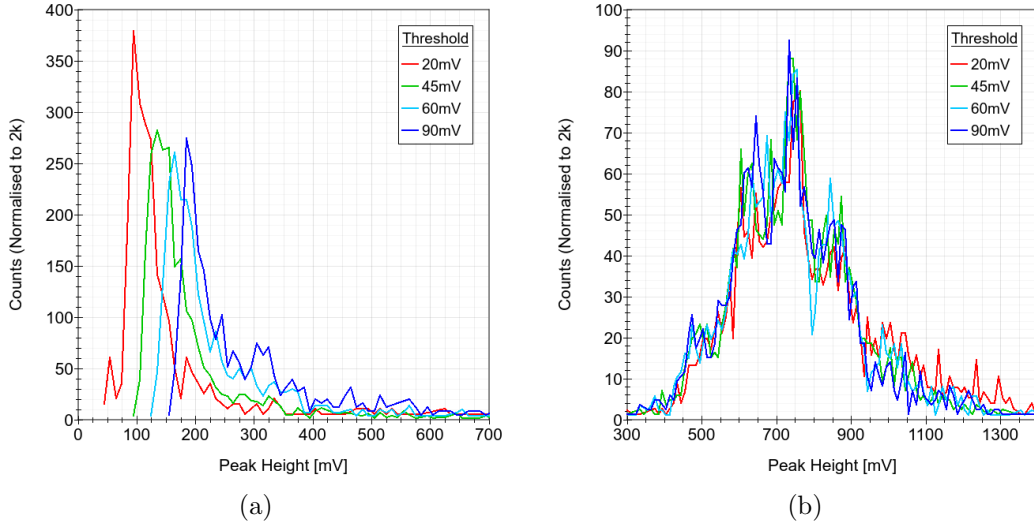
Figure 4.14: The peak height of the pulse from each coincidence event is plotted as a histogram for a 1 mm LGAD from W6 (intermediate implant energy and dose). The effect of changing the trigger threshold is shown in the case of (a) 200 V and (b) 430 V. A higher threshold results in the loss of low energy events, changing the shape of the distribution. This problem disappears at a high bias voltage.

4.2.6 Landau Cutting

Figure 4.10 and Figure 4.12 show a trend at low bias voltage, where there is low gain, in which the time resolution appears to improve. This is observed for W20, W11 and W6, but not for W3. This is not a real effect and is caused by the trigger threshold of the oscilloscope being set higher than some of the pulses generated by the LGAD. This happens when the threshold of the readout electronics is higher than lower energy events. Note that this is not the same as CFD. This threshold is to check whether an event has occurred at all and should be recorded. Figure 4.14a shows multiple histograms of signal peak heights captured with varying thresholds for a single device from wafer 6 biased at 200 V. As the threshold increases, the measured distribution is distorted as lower energy events are discarded. Figure 4.14b shows that at a bias voltage of 430 V, the threshold has no effect since it is significantly lower than even the smallest signals. However, the problem worsens at lower bias voltages where the gain and thus most-probable-value (MPV) of the peak height reduces. This is shown in Figure 4.15 where the time resolution as a function of bias
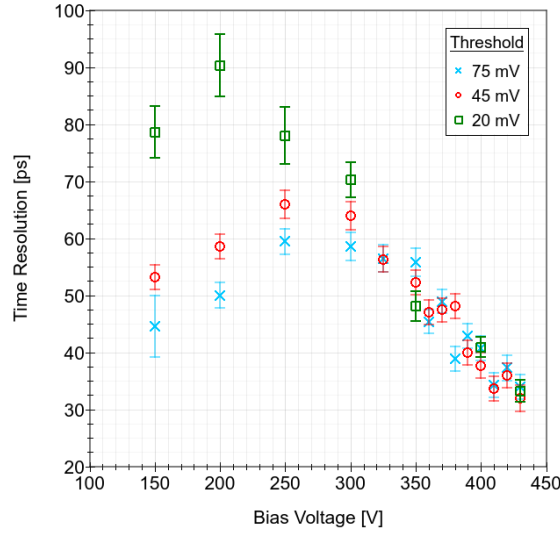
Figure 4.15: Time resolution versus bias voltage for W6 F37 (Intermediate implant energy and dose) with varying threshold voltage levels. A lower threshold results in less Landau cutting and hence a more realistic time resolution is measured.

voltage is shown for varying thresholds. For 350 V and higher the threshold shows a negligible difference. However, for lower bias voltages there is a strong dependence on the threshold where a higher threshold leads to an erroneously low value for the time resolution. This effect is minimised with a lower threshold which must be above the noise level. However, it cannot be eliminated completely, and as shown at 300 V in Figure 4.15. At 300 V there is a noticeable difference between the time resolution extracted with different trigger thresholds.

4.2.7   Summary

In this section, four wafers were diced and at least four devices from each were characterised, including gain and timing measurements. Thermal annealing was also performed on each device in an attempt to improve their breakdown performance after dicing. Although not necessarily true for every individual case, annealing successfully increased the average breakdown voltage while the standard deviation was reduced. Thermal annealing appears to have no significant effect on CV curves, as expected. There is significant gain non-uniformity in all wafers which also reflects

the non-uniformity seen with IV curves. Devices with a higher breakdown voltage achieve a high gain at a higher bias voltage. This is with the exception of W20 which does not achieve a high gain before reaching a test voltage ceiling. Time resolution also follows a similar trend where devices with a high breakdown voltage achieve a good time resolution at a high bias voltage. This is expected. As a function of gain, the non-uniformity across all devices is clear. Most importantly, each wafer sees large variability in the time resolution at a fixed value of bias voltage. Since all pads in a detector will be biased at the same voltage, this presents a problem since the time resolution of each individual pad cannot be accurately set, but instead will vary from pad to pad for the selected voltage that allows safe operation of the full array. Te2v's first batch of devices show a time resolution as a function of gain in line with that of more established vendors. Due to the high values of gain achieved at a low voltage, W3 was selected for an irradiation campaign presented in the following section.

Device ID	Fluence $[1\,\mathrm{MeV}\,\mathrm{n_{eq}/cm^2}]$
F11	5.6×10^{13}
F12	8.4×10^{13}
F13	1×10^{14}
F14	2.5×10^{14}
F08	4×10^{14}
F25	5.7×10^{14}
F28	8.3×10^{14}

Table 4.11: Seven LGADs, with identical specifications, were irradiated to a range of fluences. Taken from [49].

4.3   Irradiation Campaign

Radiation hardness is a major requirement for LGADs at the HL-LHC. The mechanisms through which irradiation damages an LGAD and affect the gain and timing performance are discussed in Section 2.5.3 and Section 2.4.2. In order to study these effects with Te2v devices, seven LGADs, with identical specifications, from W3 were selected to undergo complete characterisation before and after irradiation. Hence, all plots shown in this section are from devices from W3. Of all wafers tested so far, W3 has the best gain performance which should lead to a better resistance to acceptor removal. Flavour A, as shown in Figure 3.2, is still the focus for these tests. The fluences chosen ranged from $5.6 \times 10^{13}\,1\,\mathrm{MeV}\,\mathrm{n_{eq}/cm^2}$ up to $8.3 \times 10^{14}\,1\,\mathrm{MeV}\,\mathrm{n_{eq}/cm^2}$ and are shown in Table 4.11. In this section, the characterisation of these devices, before and after irradiation, is shown and discussed in comparison to other vendors. This section contains a more detailed analysis and discussion of data presented by the collaboration in [49].

4.3.1   Environmental Studies

As discussed previously, post-irradiation measurements must be performed cold in order to minimise leakage current. In order to better understand the results presented in this section, some brief studies of environmental conditions were conducted.
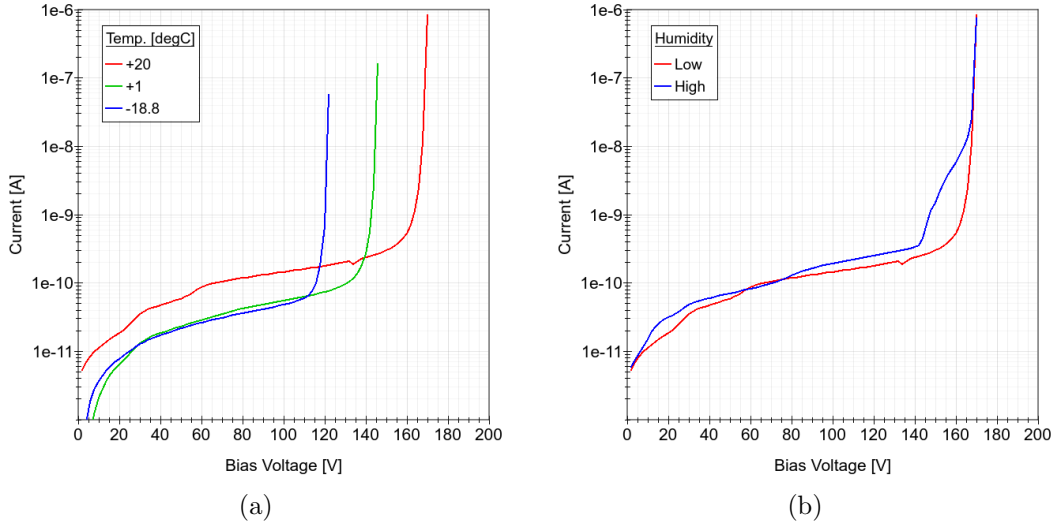
Figure 4.16: IV curves for an LGAD measured at different (a) temperatures and (b) relative humidities. A lower temperature results in a lower breakdown voltage. No humidity dependence is found for the breakdown voltage.

Figure 4.16a shows the IV curves for a device at three different temperatures. While there is a change in the leakage current, as expected, there is also a change in the breakdown voltage. A colder LGADs experiences an earlier breakdown since the gain at a given voltage is larger when the temperature is lower. This is because the impact ionisation coefficient increases as the temperature decreases [40]. Figure 4.16a suggests that the breakdown voltage has a dependence on temperature which approximates to linear for the temperature range of interest in these studies. This constitutes a difference of 24 V between each curve for a ∼20 °C difference.

The effect of humidity is shown in the case of high and low, ∼50 % RH and <10 % RH respectively, in Figure 4.16b. Generally speaking there is minimal difference. The leakage current is largely the same and the breakdown voltage is identical. However, there is an early increase in the leakage current in the case of high humidity. It has been observed in other Te2v LGADs that a very high humidity can result in an earlier breakdown voltage than usually measured by an IV sweep. This effect takes time to manifest and hence only becomes a problem for gain and timing measurements where devices are held at a high bias voltage for a long period of time. The solution is to keep the humidity below 10 % RH. All results presented beyond this point have
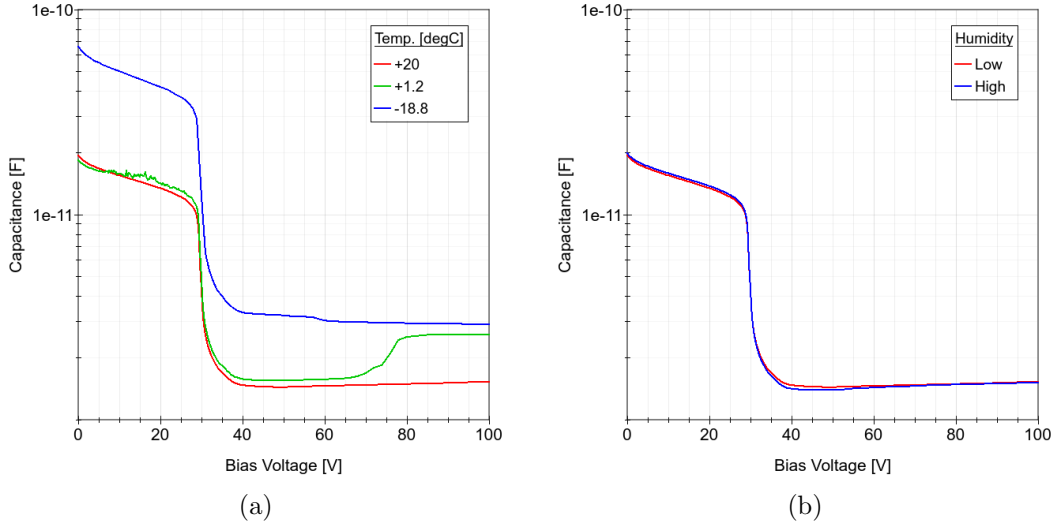
Figure 4.17: CV curves for an LGAD measured at different (a) temperatures and (b) relative humidities. Depletion voltages show no dependence on temperature or humidity

been measured in this way.

Similarly, CV curves have also been tested under different environmental conditions. Figure 4.17a shows the CV curves for the same three temperatures as before. In this case there is a difference in the measured capacitance when measured at $-18.8\,°\text{C}$. However, this is unfortunately caused by an incorrectly performed correction of the LCR meter at the start of the measurement, hence there is a constant offset across the voltage range. Of course, an incorrect correction does not explain the change in capacitance at $70\,\text{V}$ for the measurement at $1.2\,°\text{C}$. Nevertheless the key message from Figure 4.17a is that there is no temperature dependence on the gain layer and full depletion voltages. And indeed there is no dependence on relative humidity for either the depletion voltage or the capacitance as shown in Figure 4.17b.

As discussed previously the gain also has a temperature dependence. When measuring the gain with the laser system, the output of an LGAD is measured and the relative difference to a PiN measured at the same time is the gain. Since the PiN has no gain layer, it was naively assumed that it would have no temperature dependence and thus this reference measurement could be performed at room temperature to

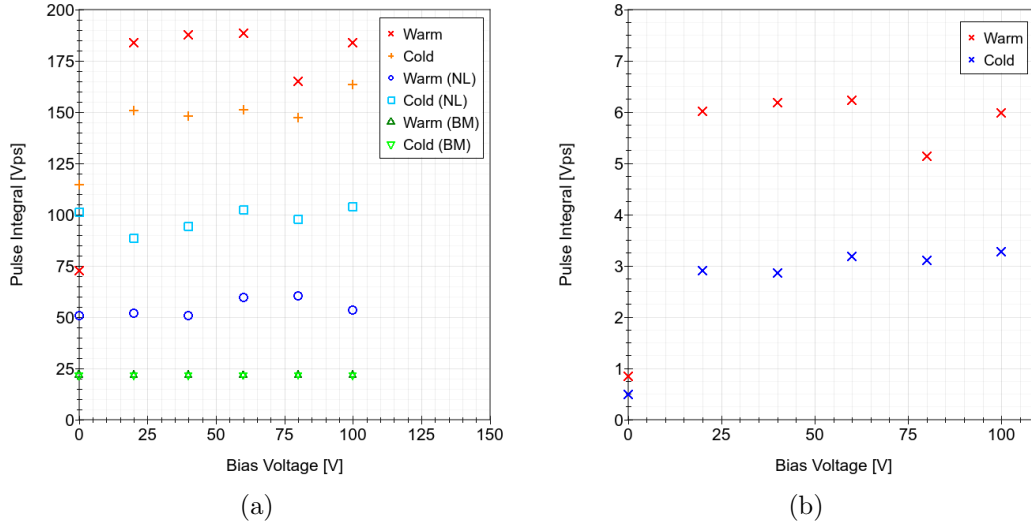Figure 4.18: The integrated pulse of a PiN as a function of voltage for two different temperatures: roughly $20\,°C$ and $-20\,°C$. (a) Uncorrected measurement show how each component (BM and NL defined in text) varies with temperature. (b) Corrected measurement as shown in Section 3.4.4. In this example, the integral of the beam monitor signal (both warm and cold) has been scaled down by a factor of 1000.

speed up the measurement process. This was quickly proven incorrect and hence a temperature sweep was preformed to better understand this dependence. This is shown in Figure 4.18. Figure 4.18a shows the pulse integral values of each step of a PiN measurement. Here, NL means a measurement performed without any injected charge from the laser and BM is the reference beam monitor which accounts for changes in the laser intensity between subsequent measurements. Figure 4.18b shows the pulse integral values after each correction (NL & BM) is applied as shown in Section 3.4.4. The BM measurements show no dependence on temperature, which is expected since the beam monitor is entirely isolated from the cooling system. The NL measurements show a strong dependence on temperature. This is a measure of the noise of the sensor and implies that the noise is reduced at a lower temperature, which is expected. Once the laser is introduced, the temperature dependence remains meaning the signal is reduced as the temperature lowers. Note that this is the opposite of an LGAD where the signal increases since the gain has a negative correlation with temperature. The reason for this temperature dependence is a combination of silicon being an indirect semiconductor and the use of an infrared

laser. As discussed in Section 2.1.1.1, silicon can absorb low energy photons through indirect transitions where additional energy from phonons in the lattice are required. This energy is temperature dependent and thus a lower temperature means the absorption coefficient is lower. Hence, fewer photons are absorbed in total, resulting in less charge being injected into the PiN [50, 51].

4.3.2 IVs

Current as a function of voltage (IV) has been measured for all seven devices before and after irradiation. IV curves measured before irradiation are shown in Figure 4.19a. Each device exhibits a similar current profile below $\sim$140 V where the current slowly rises from $\sim$10 pA to $\sim$200 pA. After this, the first of the devices goes into breakdown where the current rises rapidly over a small voltage range. There is then a large discrepancy characterised by a standard deviation of 20 V on the average breakdown which is 197 V. This large variance suggests poor uniformity between separate devices within the same wafer and is in agreement with variation discussed in Section 4.2.1.

Post-irradiation IV curves are shown in Figure 4.19b and exhibit a strong dependence on fluence. Generally the leakage current remains at $\sim$100 nA before breakdown, an increase of three orders of magnitude from pre-irradiation. The breakdown voltages have been extracted and are shown in Figure 4.20a as a function of fluence. Devices F25 and F28 are shown for completeness, but did not reach the 10 μA current limit before the 650 V maximum bias was reached. Although there appears to be a general linear trend, measurements at low fluences do not completely support this. This can be understood with Figure 4.20b, where the change in breakdown voltage is plotted instead. Ignoring the final two devices which did not undergo breakdown, a linear fit can be applied to these data. The fit reveals a 77 V increase in breakdown for an increase in fluence of 1×10^{14} 1 MeV n_{eq}/cm^2. This is in agreement with results from other vendors such as CNM and HPK [52, 53].
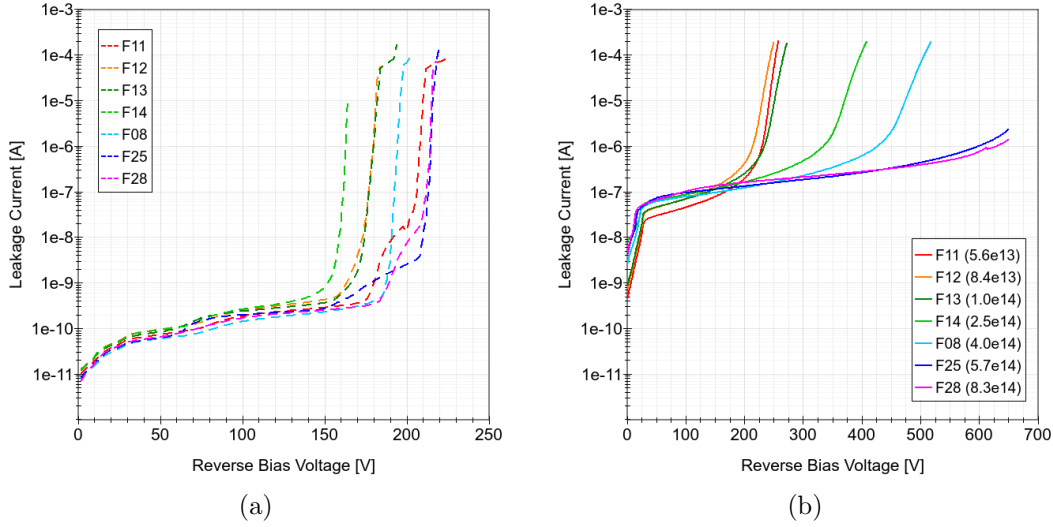
Figure 4.19: IV curves. (a) Pre-irradiation at room temperature and $\sim$50 % RH. (b) Post-irradiation at $-20\,°$C and <10 % RH. There is significant variability pre-irradiation. After irradiation, the change in breakdown voltage depends on the fluence, as expected

When calculating the breakdown voltage at different temperatures, the current has been corrected using equation Equation 3.1. As discussed earlier in Section 4.3.1, an LGAD's breakdown voltage is also dependent on temperature. However, the relationship is linear and hence the trend shown in Figure 4.20b is still valid.

4.3.3 CVs

Capacitance has also been measured for all seven devices as a function of bias voltage. Pre-irradiation CV curves measured at 100 kHz are shown in Figure 4.21a and show significantly lower variability compared to the IV curves in Figure 4.19a. These curves can be analysed to extract the depletion voltages using the methods described in Section 3.3.4. Pre-irradiation, each device should be identical meaning the average can be calculated and is shown in Table 4.12. Here, the three measurement frequencies are also shown and the extracted V_{GLD} from all three are in agreement. V_{FD} shows more variation as shown by the increased standard deviation as well as the difference in mean value between each frequency. This is due to the difficulty in
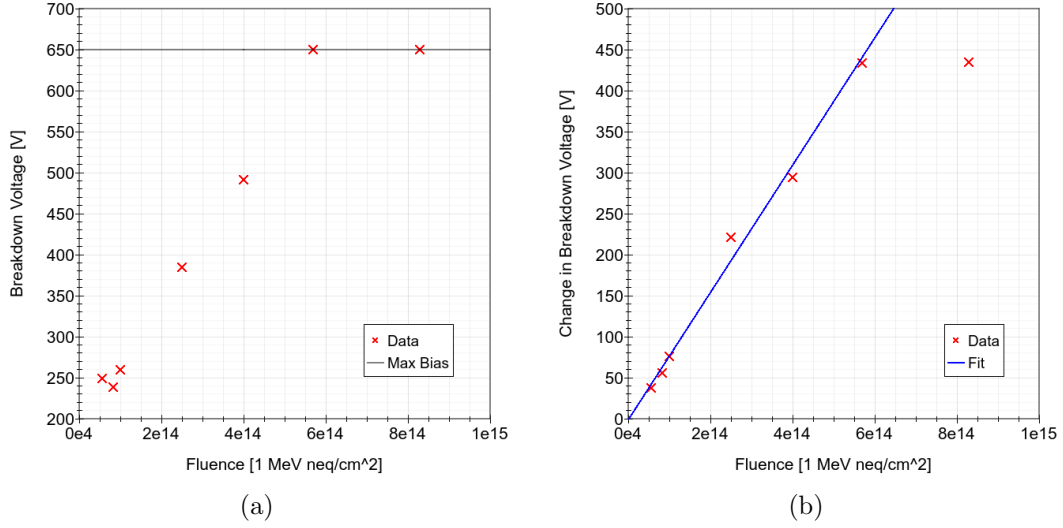
Figure 4.20: (a) Post-irradiation breakdown voltage as a function of fluence. (b) Change in breakdown voltage as a function of fluence. Note that for the final two fluences, full breakdown is not achieved at the maximum bias voltage of 650 V
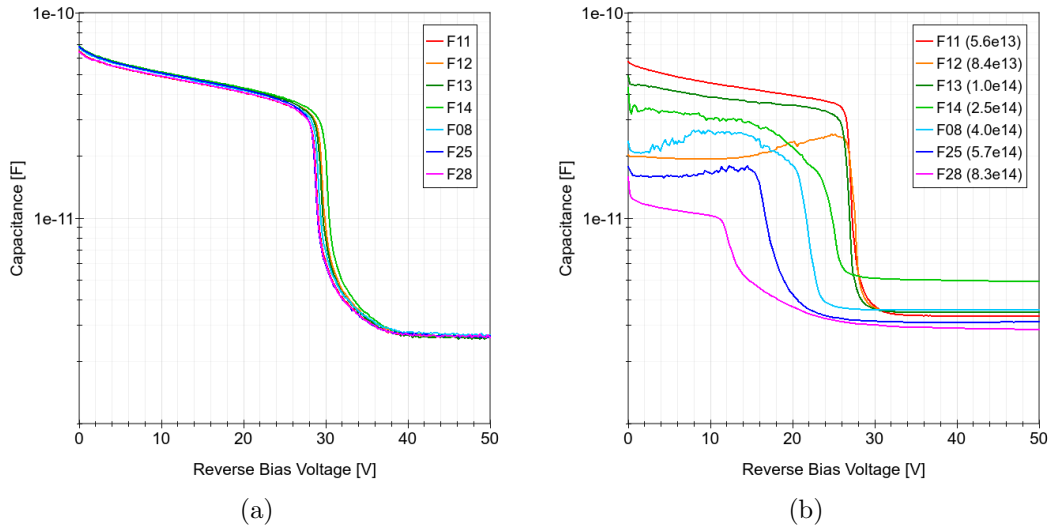
Figure 4.21: CV curves. (a) Pre-irradiation at room temperature and $\sim$50 % RH measured at 100 kHz. (b) Post-irradiation at $-20\,°$C and $<$10 % RH measured at 1 kHz. F28 is measured at 2.5 kHz due to an issue with the 1 kHz measurement. After irradiation, the depletion voltage is dependent on fluence.

Frequency [kHz]	V_{GLD} [V]	V_{FD} [V]
1	28.3 ± 0.2	40.2 ± 0.8
10	28.3 ± 0.2	41.2 ± 0.5
100	28.3 ± 0.2	40.2 ± 0.4

Table 4.12: Extracted depletion voltages for three different measurement frequencies pre-irradiation. No dependence on measurement frequency is found, as expected.

Frequency [kHz]	C_0 [pF]	C_{GLD} [pF]	C_{FD} [pF]
1	94.0 ± 7.2	33.4 ± 0.8	2.70 ± 0.02
10	66.0 ± 0.6	30.3 ± 0.2	2.70 ± 0.02
100	65.1 ± 0.7	30.1 ± 0.2	2.70 ± 0.02

Table 4.13: The capacitance, pre-irradiation, at each stage of depletion, including the initial capacitance with the bias held at $0\,\mathrm{V}$. $1\,\mathrm{kHz}$ shows a difference in capacitance compared to higher frequencies. This is not expected, however this may be due to the coupling box described in Figure 3.8 which is expected to cause issues at low frequency.

extracting V_{FD}, as previously discussed.

Table 4.13 shows the value of capacitance at each stage of depletion, including the initial capacitance when the bias voltage is held at $0\,\mathrm{V}$. From the standard deviation, it is clear that measurements at $1\,\mathrm{kHz}$ are not consistent with $10\,\mathrm{kHz}$ or $100\,\mathrm{kHz}$. Focusing on the $100\,\mathrm{kHz}$ measurements, the capacitance can be converted to a depletion depth using Equation 3.2. The initial depletion depth is calculated to be $(1.59 \pm 0.05)\,\mu\mathrm{m}$ and rises to $(3.44 \pm 0.05)\,\mu\mathrm{m}$ once the gain layer is depleted. At full depletion the depth is $(38.3 \pm 0.7)\,\mu\mathrm{m}$, which is significantly smaller than the specification's $50\,\mu\mathrm{m}$ active thickness. At $2.7\,\mathrm{pF}$, the full depletion depth is very sensitive to small systematic changes in the capacitance, such as parasitic effects from the probe station chuck. As discussed earlier, the measured capacitance is sensitive to lateral depletion which is likely to be the dominate cause of this discrepancy. These effects, however, should not affect the depletion voltages. The depletion depths are in agreement with those measured pre-dicing in Section 4.1.2.

Figure 4.21b shows the CV curves for these devices, post-irradiation. Similar to IV curves in Figure 4.19b, there is a strong influence from the irradiation fluence. As discussed in Section 3.3.5, the measured capacitance has a strong dependence on frequency after irradiation and makes conclusions hard to draw. However, this section also makes it clear that V_{GLD} is independent of measurement frequency and can hence be extracted. A reduction in V_{GLD} indicates acceptor removal and a reduction in gain as discussed in Section 2.5.3. The ratio between V_{GLD} before and after irradiation, is shown as a function of fluence in Figure 4.22. Equivalent data for proton-irradiated LGADs manufactured by FBK are also shown [11]. The acceptor removal coefficient, c_A, discussed in Section 2.4.2.4, can be extracted by fitting $V_{GLD}(\phi) = V_{GLD}(0) \cdot \exp(-c_A\phi)$ to these data [25]. c_A is found to be $(9.7 \pm 0.5) \times 10^{-16} \, \text{cm}^2$. This is 45% larger than in FBK's UFSD2 production where $c_A = 6.7 \times 10^{-16} \, \text{cm}^2$. Multiple factors such as the initial doping density, the doping profile and differences in the manufacturing processes can explain the difference seen in c_A for the devices in Figure 4.22 [49]. Subsequent batches from Te2v will require optimisation of these processes to improve c_A.
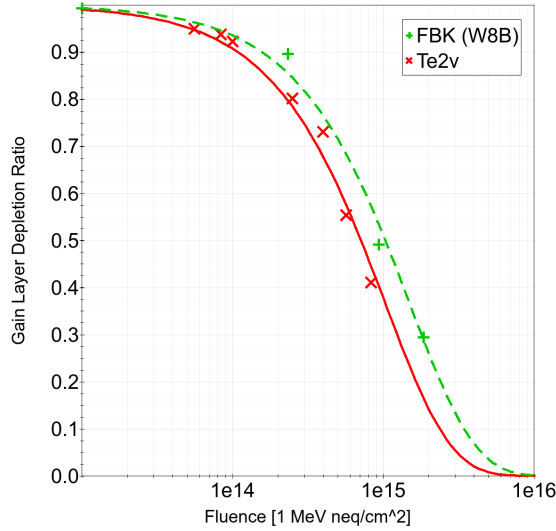

Figure 4.22: Gain layer depletion ratio as a function of fluence compared with proton irradiated devices from FBK. While similar, Te2v have a higher acceptor removal coefficient than FBK devices. [11]
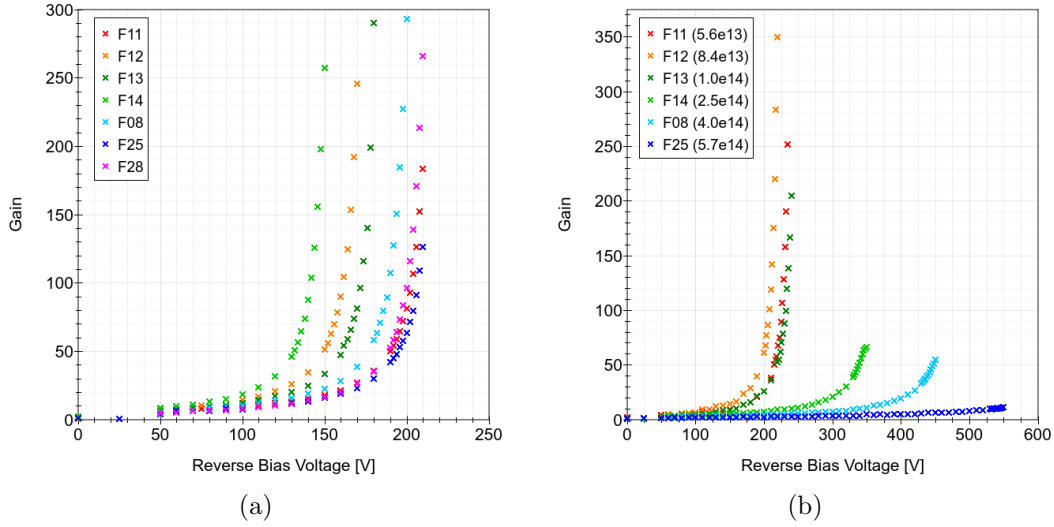
Figure 4.23: Gain measurements for W3. (a) Pre-irradiation gain measured at room temperature and humidity. (b) Post-irradiation gain measured at $-20\,°$C and $<10\,\%$ RH. There is significant variability in the bias voltage dependence of gain. This is in line with the breakdown voltage dependence.

4.3.4   Gain

Figure 4.23a shows gain as a function of bias voltage before irradiation. Each device can achieve a gain in excess of 125 with most devices reaching over 250. Similarly to the IV curves in Figure 4.19a, there is significant variability in terms of the voltage dependence. This is shown more clearly in Figure 4.24a where the average bias voltage required to achieve a particular value of gain is calculated. More importantly, the standard deviation very quickly plateaus to roughly 24 V about an average of $\sim$180 V. To try and correct for the variation in breakdown voltage, Figure 4.24b shows the gain achieved at a fraction of the breakdown voltage, V_{rel}. This suggests that either after correcting for non-uniformity, the shape of each gain curve is still different, or this correction is not valid.

After irradiation, the gain is measured again but at $-20\,°$C and $<10\,\%$ RH. This is shown in Figure 4.23b for all devices (except F28 which suffered from current instability when held at a high voltage for a long period of time) [49]. For fluences of $1.0 \times 10^{14}\,1\,$MeV n_{eq}/cm^2 and lower, a gain in excess of 200 is still achieved. The
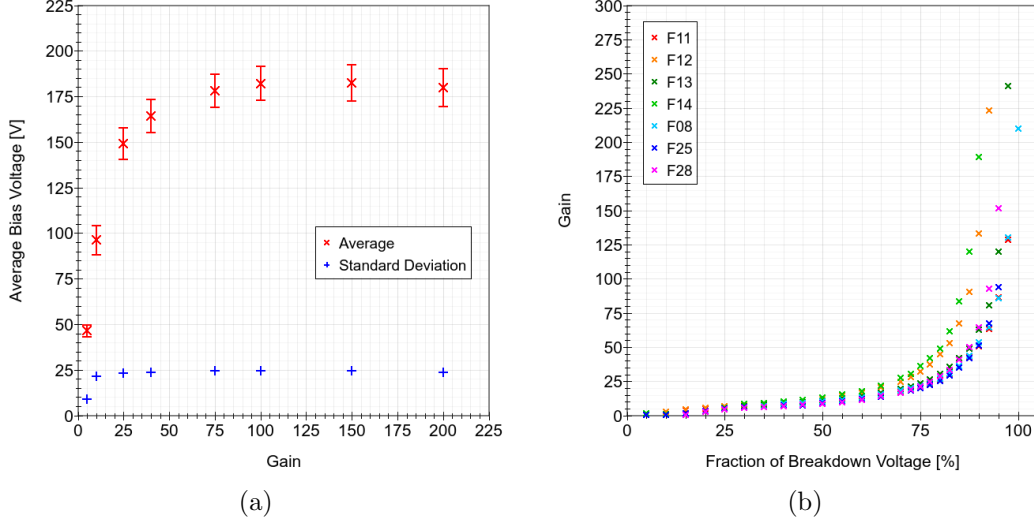
Figure 4.24: (a) Average bias voltage required to achieve a given value of gain pre-irradiation. (b) The pre-irradiation gain achieved at a fraction of the breakdown voltage.

bias voltage required to achieve a high gain has increased compared to pre-irradiation but only by $\sim$40 V. In contrast, at higher fluences a gain of 75 or 50 is more typical [49]. However a higher bias voltage, between 100 V and 300 V with respect to pre-irradiation, is required to achieve these values of gain. In the case of F25, this means that 550 V is required to achieve a gain of 10, compared to pre-irradiation where a gain of 125 could be achieved with only 210 V. This follows the trend expected from acceptor removal [49].

4.3.5   Timing

This section presents the time resolution (measured with mips) for each device as a function of bias voltage and then gain (measured with a laser). The pre-irradiated time resolution is shown as a function of voltage and gain in Figure 4.25a and Figure 4.25b respectively. It is clear that all devices can comfortably achieve a time resolution of 40 ps with some reaching 35 ps. Data for an LGAD manufactured by HPK is also shown in Figure 4.25b and seems to have a better time resolution as a function of gain [10]. This difference can come from various sources. For example,

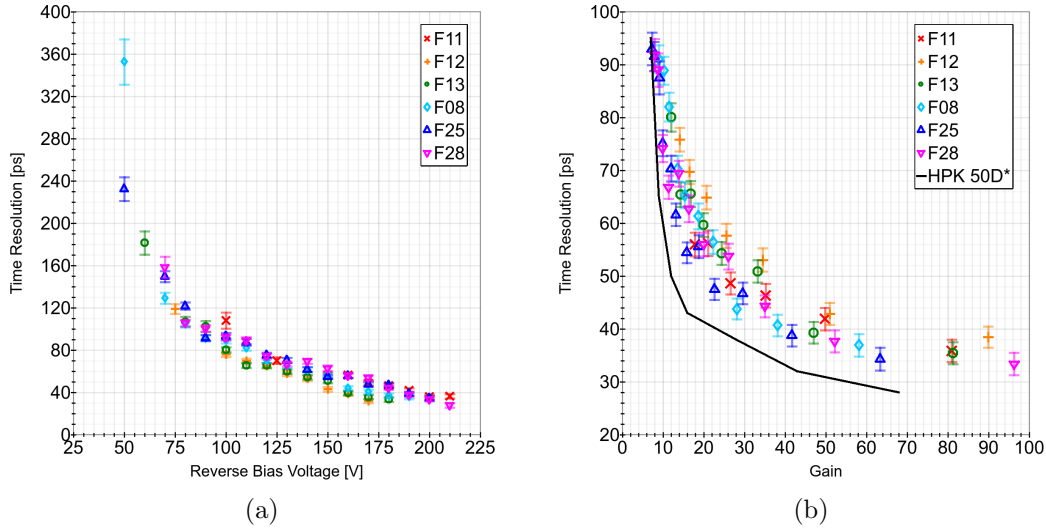Figure 4.25: Pre-irradiation time resolution as a function of (a) bias voltage and (b) gain. The dependence on gain is in line with other vendors.

an optimisation of the CFD combined with better readout electronics would help to reduce the jitter and improve overall time resolution. In addition, in this work the gain is measured using a laser, which does not cause the gain suppression seen with mips (the method used by [10]). This means gain (measured with a laser) is likely to be overestimated, thus causing a shift towards higher values of gain in the plot.

Each device shows a similar relationship with voltage and with gain. This is characterised by Figure 4.26. Figure 4.26a and Figure 4.26b shows the average time resolution achieved at specific values of bias voltage and gain respectively. As the gain increases, the standard deviation of time resolution begins to dramatically decrease from $17.8\,\text{ps}$ at a gain of 10 to $2.4\,\text{ps}$ at a gain of 50. This trend is expected, since the time resolution plateaus for a sufficiently large gain. In contrast, as bias voltage increases the standard deviation actually increases $4.6\,\text{ps}$ at $125\,\text{V}$ to $8.4\,\text{ps}$ at $170\,\text{V}$. This is a reflection of the variability seen in Figure 4.23a where the variability in gain grows at a high bias voltages, however it is also suppressed by the plateau region. Figure 4.26c and Figure 4.26d also tell a similar story where the average bias voltage and gain, respecitvely, required to achieve a specific value of time resolution are shown. In this case, as the time resolution decreases, the stan-
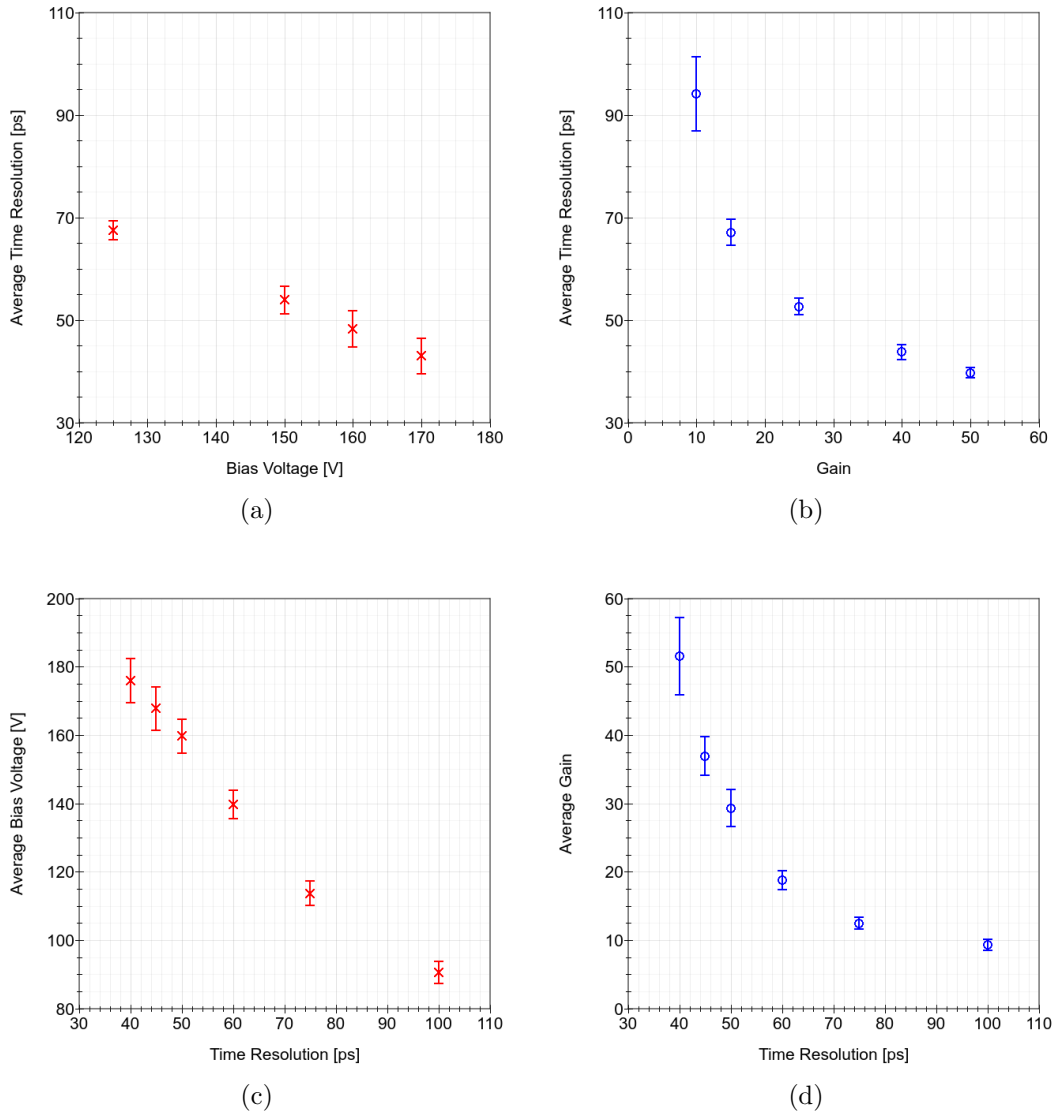
Figure 4.26: (a) The average pre-irradiation time resolution for a given value of bias voltage. (b) The average pre-irradiation time resolution for a given value of gain. (c) The average bias voltage required for a particular value of time resolution pre-irradiation. (d) The average gain required for a particular value of time resolution pre-irradiation.
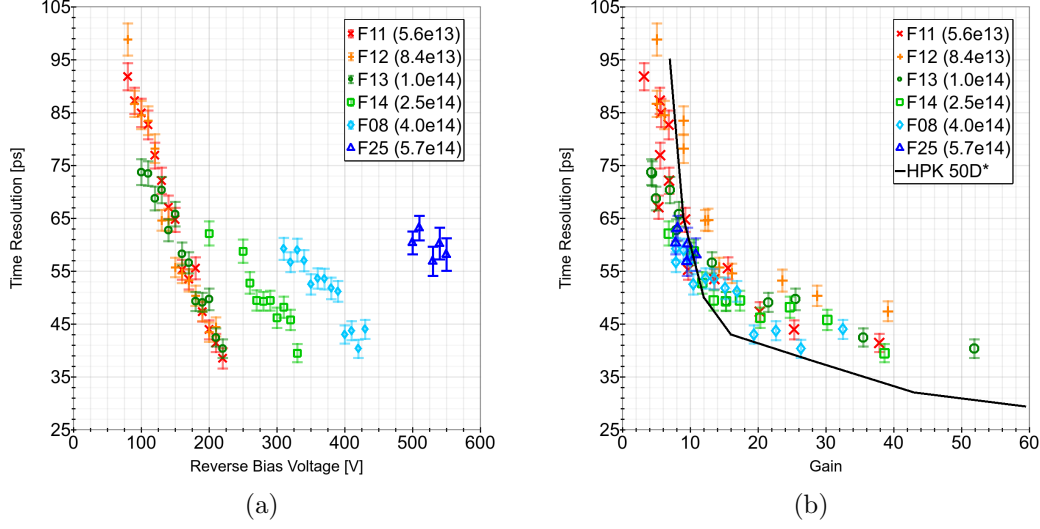
Figure 4.27: Post-irradiation time resolution as a function of (a) bias voltage and (b) gain. After irradiation, the dependence on gain is still in line with other vendors.

dard deviation on the required bias and gain increase. They start at 7.7 V and 2.2 respectively at a 100 ps time resolution, before increasing to 15.7 V and 13.8 at a 40 ps time resolution. This increase is caused by the plateau of the time resolution where the bias and gain is very sensitive to small differences in time resolution. The four plots in Figure 4.26 show that if one can achieve a consistent value of gain, then the variation in time resolution is minimised. However, in HEP experiments all sensors will be biased to the same value. In this case the variance is large and unavoidable.

Time resolution measured after irradiation is shown in Figure 4.27. As a function of bias voltage in Figure 4.27a, devices irradiated with a fluence of $4.0 \times 10^{14} \, 1 \, \text{MeV} \, \text{n}_{\text{eq}}/\text{cm}^2$ or less can achieve a sub-40 ps time resolution. The time resolution performance is only slightly worse compared to pre-irradiated results. The bias voltage dependence also changes with fluence in a similar way for gain in Figure 4.23b. As the fluence increases, a higher bias voltage is required to achieve sub-40 ps. For F25, the time resolution never reaches 55 ps before the maximum bias voltage of 550 V. Figure 4.27b shows that the time resolution has the same general gain dependence as with the pre-irradiated case. However, for particular

values of gain, the time resolution improves in the post-irradiation case. This difference is stronger for lower values of gain, as shown in Figure 4.28. There are three potential explanations for this improvement after irradiation.
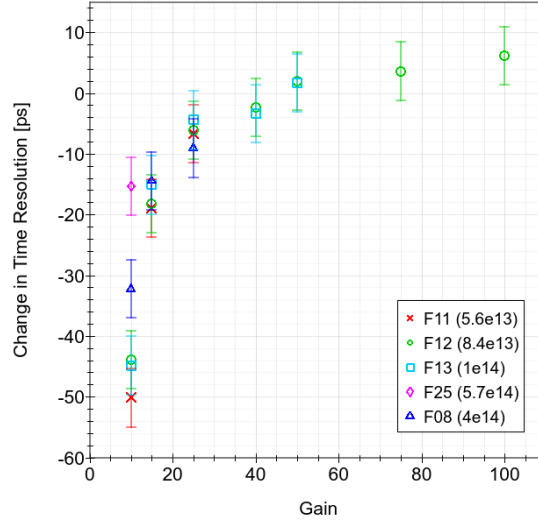

Figure 4.28: For each device where data are available, the change in time resolution, before and after irradiation, is plotted as a function of gain. As gain decreases, post-irradiation time resolution sees a great improvement in time resolution compared to pre-irradiation measurements. At high gain this difference is negligible.

Firstly, after irradiation the electron mobility no longer saturates at high electric fields [32, 33]. This means that as the bias voltage increases, so does the drift velocity. This shortens the rise time, which minimises the jitter component of the time resolution as shown in Equation 2.33. After a high fluence of irradiation, the same value of gain requires a higher bias voltage than pre-irradiation. This results in a decrease in rise time and hence an improvement in time resolution. The mobility also has a temperature dependence [32]. A lower temperature also results in a higher mobility and thus a lower rise time, improving the time resolution. Pre- and post-irradiation results are performed at $\sim 20\,^{\circ}\mathrm{C}$ and $-20\,^{\circ}\mathrm{C}$ respectively. This makes interpretation between pre- and post-irradiation challenging. The second effect is the introduction of trapping centres into the bulk region. If the trapping time constant is long enough, then the electron is lost and stops contributing to the signal. The trapping time at a fluence of $8 \times 10^{14}\,1\,\mathrm{MeV}\,\mathrm{n_{eq}/cm^2}$ is roughly $3\,\mathrm{ns}$ which is comparable to the signal duration [54]. Higher fluences result in a decreased

trapping time, making this effect more prominent. The further an electron has to travel, the more likely it is to fall into one of these traps. This disproportionately affects electrons further away from the n-type electrode. This makes the LGAD behave as though it is thinner which results in reduced Landau fluctuations and thus a better time resolution, as discussed in Section 2.5.1. Another effect that has been directly observed is Landau cutting. This has been discussed in Section 4.2.6 in the context of W20. Those which have been significantly affected by Landau cutting are not shown.

4.3.6 Summary

This section details how irradiation affects the IV, CV, gain and timing properties of LGADs. IVs and gain profiles see a bias voltage shift proportional to the irradiation fluence. As the fluence increases a very high gain is no longer achievable and eventually even a moderate gain is no longer achievable. Extracting the gain layer depletion voltage from CV curves before and after irradiation means that the acceptor removal coefficient can be calculated. While comparable with results from similar devices by FBK, there is still a significant (45 %) difference for Te2v devices. Improvement and optimisation in the design and manufacturing process will be required to bridge this gap. For example, increasing the gain implant energy and dose would be a simple way to improve the acceptor removal coefficient, which depends on the initial doping concentration. Timing performance after irradiation is in line with expectations. A higher fluence resulted in a higher bias voltage dependence. Taking the gain dependence into account, the dependence on fluence was removed. The timing performance as a function of gain seemed to improve compared to pre-irradiation due effects like Landau cutting, larger electron mobility and trapping centres. Overall, Te2v LGADs have been shown to have the expected characteristics and trends after irradiation.

CHAPTER 5

SIMULATION OF CCDS AT TELEDYNE E2V

5.1 Introduction

Part of the work of this thesis included a nine month secondment at Te2v. The topic of the work during this time was the simulation of CCD devices. This study is presented in this chapter.

Te2v is a well established foundry of CCDs for many applications including space imaging, medical imaging and defence among others [13]. They work with partners such as the European Space Agency (ESA), United Kingdom Space Agency (UKSA) and National Aeronautics and Space Administration (NASA) to deliver custom semi-conductor imaging sensors to meet their exact needs [13]. This means it is important for Te2v to be able to understand the capabilities of their current portfolio of detectors when corresponding with customers. By extension, it is also important that Te2v can understand the capabilities of future detectors, making informed decisions

before the first batch leaves the foundry.

This chapter is a discussion of the work completed at Te2v on a custom in-house simulation framework. It aims to provide an easy-to-use way to predict the behaviour of future detectors, specifically a property called the MTF. Simultaneously, it must be fast enough to keep up with changes and iterations of the design and specification from customers. This work is heavily inspired by the work of Daniel Weatherill's thesis where a similar simulation was developed [55].

5.2 Theory

5.2.1 Charged-Coupled Devices
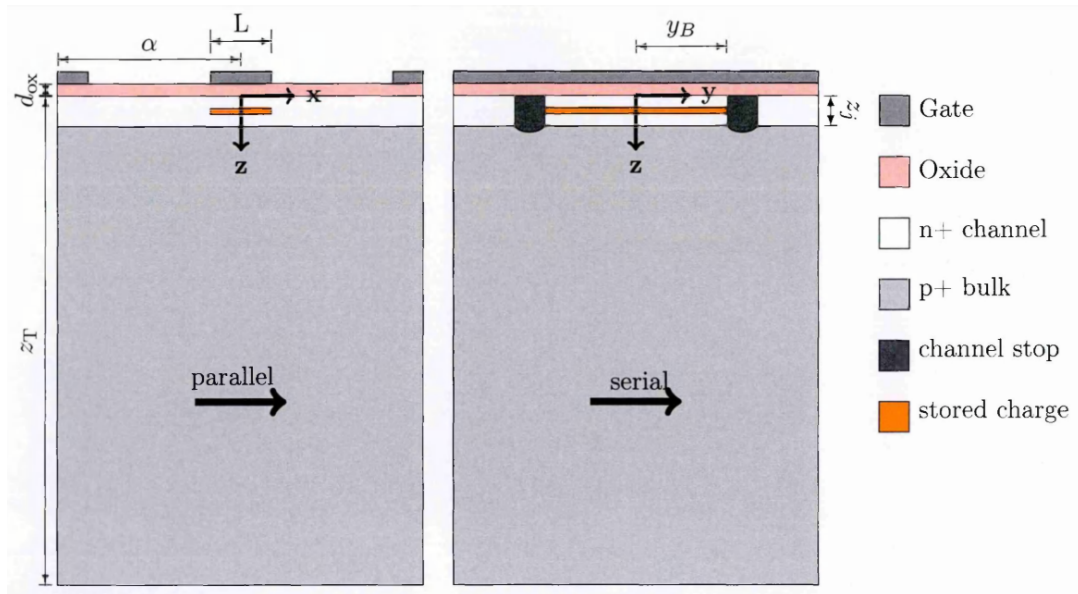
Figure 5.1: Schematic diagram of a CCD detector. The parallel (x) direction has a series of gates which form a potential well below the oxide layer. The serial (y) direction has p-stop (channel stop) implant strips which completes the segmentation of pixels. Taken from [55].

CCDs are a type of semiconductor detector which reads charge very differently to an LGAD. Rather than measuring an induced current as electron-hole pairs drift apart, the electrons are collected by a potential well beneath the electrode [15]. The key

difference is that the thin n-type region does not have a direct electric connection to the metal on the surface of the sensor, referred to as the gate in Figure 5.1. Instead there is a layer of silicon oxide which acts as an electrical insulator. When the drifting electrons reach the electrode, they fall into a minimum (well) in the potential field and remain there [15]. This potential well continues to collect charge until the detector is ready to enter readout mode. The longer this exposure time is, the more charge is collected which improves image contrast. The potential well can be manipulated by biasing neighbouring overlapping gates, as shown in Figure 5.1 [15]. This allows charge to be shifted from one pixel to another and eventually to the readout electronics where the charge collected can be integrated and the intensity of each pixel is recorded.
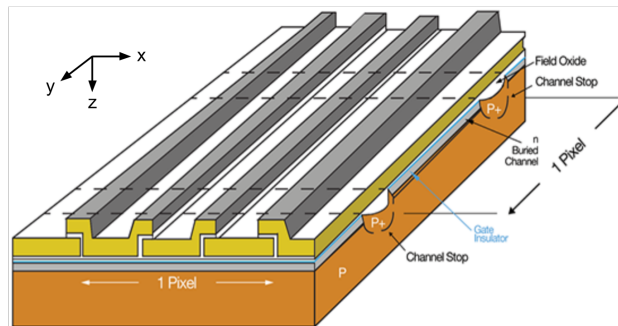
Figure 5.2: Segmentation is achieved in the x-axis with overlapping gates, only some of which are biased. In the y-axis, a channel stop implant strip achieves segmentation guiding charge to either side of the p implant. Taken from [56].

This work attempted to model CCD250, the sensor designed and fabricated by Te2v for the Large Synoptic Survey Telescope (LSST) project [57]. This sensor has a $10\,\mu$m pitch, the distance between the centres of neighbouring pixels. The pixels themselves have dimensions $(8 \times 7)\,\mu$m^2. It is a $100\,\mu$m thick, back-illuminated sensor. Parameters used for the simulation of CCD250 are shown in Table 5.1. All references to the sensor in this work refer to CCD250.

Symbol	Value	Description
α	$10\,\mu\text{m}$	Pixel pitch - Distance between centres of adjacent pixels (x & y).
L	$8\,\mu\text{m}$	Pixel width - Total width of a pixel when three adjacent gates are biased high.
z_j	$1\,\mu\text{m}$	Junction depth - Assuming an abrupt boundary.
z_T	$100\,\mu\text{m}$	Device thickness.
y_B	$3.5\,\mu\text{m}$	Channel stop position - Distance from centre of a pixel to the edge of the channel stop implant.
d_{ox}	$0.2\,\mu\text{m}$	Oxide thickness.
V_G	$10\,\text{V}$	Gate high voltage.
V_T	$0\,\text{V}$	Gate low voltage.
V_{BS}	$-70\,\text{V}$	Backside voltage.
ϵ_{si}	11.7	Relative permittivity of silicon.
ϵ_{ox}	3.4	Relative permittivity of silicon oxide.
T	$150\,\text{K}$	Operating Temperature.

Table 5.1: The key parameters of CCD250 used in the simulation. Adapted from [55].

5.2.2 Modulation Transfer Function

When taking an image, resolution and contrast are two important properties for determining the quality of the image [58]. The resolution, which is the ability of the sensor to resolve two nearby objects, is proportional to the size of an individual pixel. A smaller pixel improves the resolution. However, a larger pixel is able to collect more light while sacrificing resolution. This is important in space applications where a long exposure time is usually required in order to capture enough light to form an image with sufficient contrast. Contrast is a measure of the difference between the brightest and darkest parts of an image. The higher the contrast, the stronger the difference.

The modulation transfer function (MTF) is a method of characterising both the resolution and the contrast of an image allowing different sensors to be easily compared. More specifically it is the ratio of the contrast of the image, over the contrast of the object being imaged, measured in spatial frequency, k, space [59, 58]. Con-
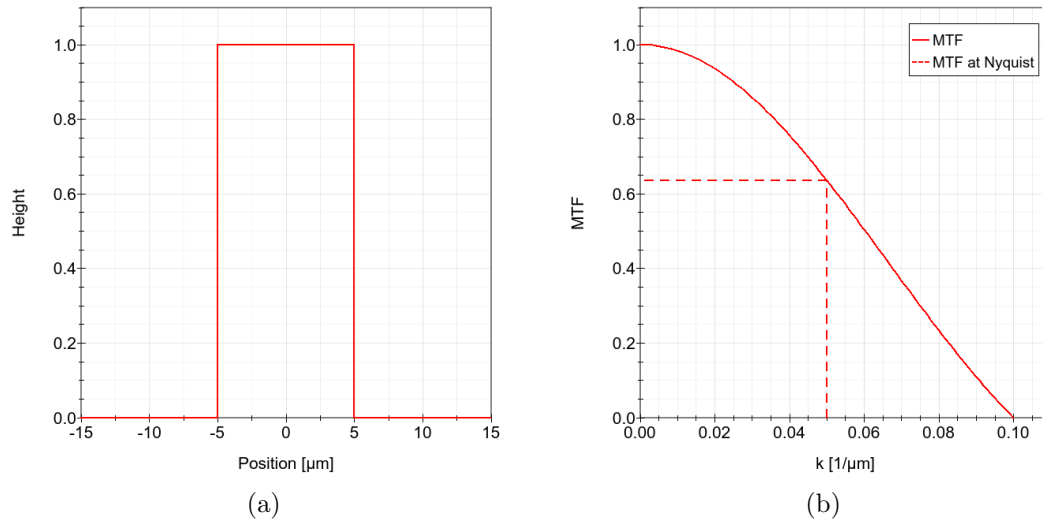
Figure 5.3: (a) A step function which represents a single pixel. (b) The FFT of a single pixel gives the MTF. The MTF at nyquist is often quoted and always equals 0.64 in the geometric case.

sider a single pixel in an array with a width $w = 10\,\mu$m, represented by the step function in Figure 5.3a. The fast fourier transform (FFT) of this step function is a complex number. The modulus is taken and normalised against the maximum as a function of k, as shown in Figure 5.3b. This is known as geometric MTF since only the geometry of the pixel contributes here. It is common to characterise this curve with MTF_{nyq}, the MTF at Nyquist. The Nyquist frequency is the maximum signal frequency that can be accurately sampled [60]. It is equal to half the sampling frequency. For a pixel array, the sampling frequency is $1/w$ and therefore the Nyquist frequency is $1/2w$.

Geometry is not the only factor in MTF. When charge is generated in a sensor, it diffuses and spreads as it travels to the electrodes. Assuming all the charge is generated in a single point, then the spread of this charge is characterised by the point spread function (PSF) [61, 58]. Assuming symmetry, this is approximately a Gaussian distribution with a standard deviation, σ, shown in Figure 5.4a, where $\sigma = 1\,\mu$m. To calculate the MTF in this new context, the step function must be convolved with the PSF as shown in Figure 5.4a [58]. The FFT can then be performed on this convolution to give an MTF which differs from the geometric
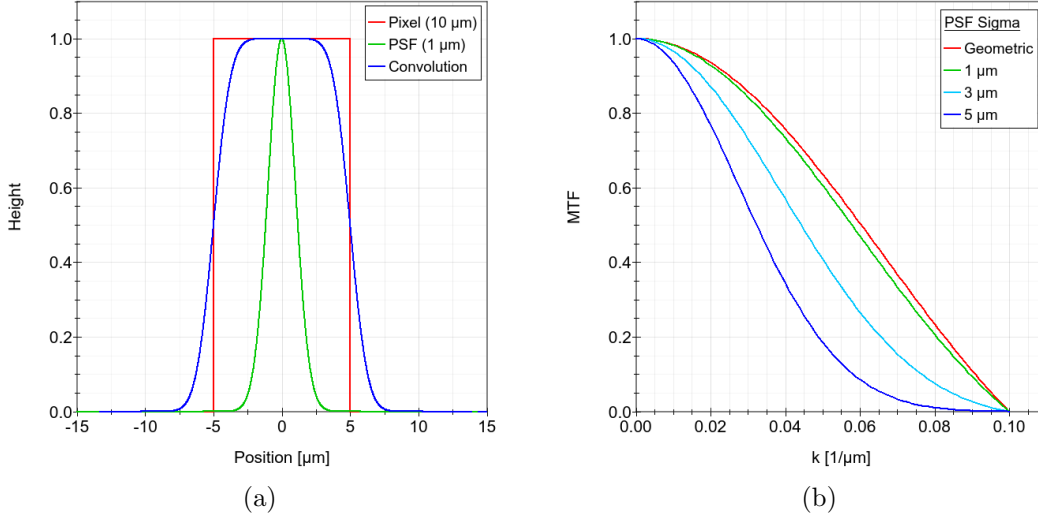
Figure 5.4: (a) The pixel's step function is convolved with a Gaussian PSF with $\sigma = 1\,\mu$m. (b) The FFT of this convolution is shown for a few values of σ and compared to the geometric case.

case, as shown in Figure 5.4b for varying σ. As before, MTF_{nyq} can be extracted and directly compared with the geometric case. This is shown in Figure 5.5a where the MTF for a selection of pixel widths is shown in the case of $\sigma = 3\,\mu$m. MTF_{nyq} is then extracted in each case and is shown in Figure 5.5b.

To be able to model the MTF of a CCD, one can instead model the PSF, measuring σ, from which the MTF can then be inferred. In order to do this one must understand how electrons move within silicon. This includes the potential fields and diffusion effects acting upon each electron which is discussed in the remainder of this section. The measurement of the PSF's σ is entirely contextual. It is affected by the structure of a CCD, particularly how well defined the pixel boundaries are. If σ is large enough and the pixel size is small enough, charge will be shared between neighbouring pixels [61]. The source of charge will also have an impact on the PSF. For example, different frequencies of visible light give a different value of σ [61]. While visible light illuminates the entire sensor area, X-rays manifest as single photon absorption events which also affects the σ measured. In this work, X-ray absorption is the focus and the simulation attempts to replicate experimental data performed with an Fe-55 source. More details on the experimental procedure is discussed alongside
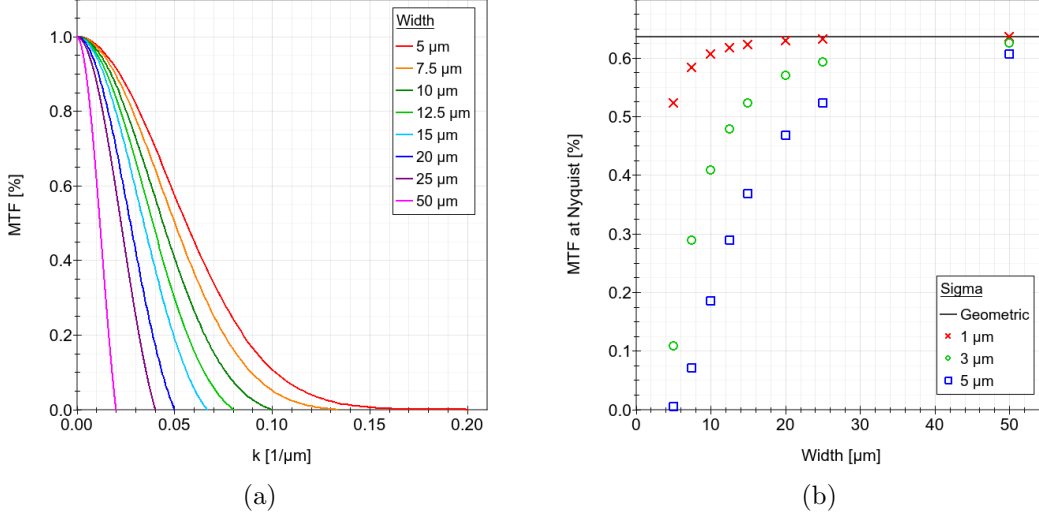
Figure 5.5: (a) Assuming $\sigma = 3\,\mu$m, the MTF is shown for varying pixel widths. (b) In each case, MTF_{nyq} is extracted and shown as a function of width. As the width increases, MTF_{nyq} tends towards the geometric case. This is repeated for different values of σ and shows that as σ increases, MTF_{nyq} decreases.

the simulation in Section 5.4

5.2.3   Potential Fields

To understand how charge moves within a sensor, one needs to start with the potential fields from the sensor itself. The potential field, $\Phi_{tot}(x, y, z)$, can be described as the summation of potentials from multiple sources such that,

$$\Phi_{tot}(x,y,z) = \Phi_H(x,z) + \Phi_B(x,z) + \Phi_D(x,y,z) + \Phi_C(x,y,z) + \Phi_E(x,y,z), \quad (5.1)$$

where x, y and z represent Cartesian coordinates within the detector as indicated in Figure 5.1 [55]. Each component can be calculated individually, and then the summation can be calculated as shown in Figure 5.6. This shows the potential well which electrons are guided into by the gradient in potential.
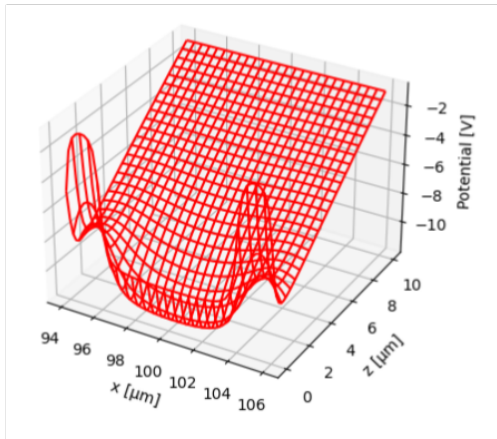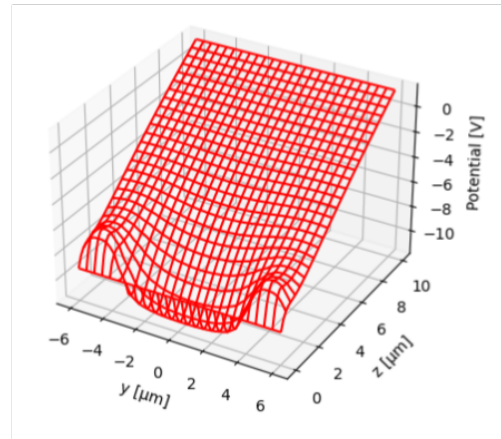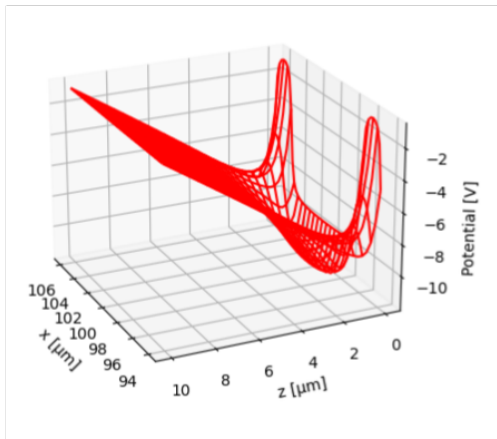
Figure 5.6: The total potential fields is shown as a function of (a & c) XZ and (b & d) YZ. There is a strong potential gradient which will cause electrons to drift towards a well defined potential well.

$\Phi_H(x, z)$ is known as the homogeneous potential and is responsible for a CCD's potential well [55]. As shown in Figure 5.1, a CCD has an array of metal gates above the oxide layer. These can be either be biased high at $V_G = 10\,\text{V}$ or held low at $V_T = 0\,\text{V}$. A single pixel has four gates, three of which are biased while charge is collected. By switching these gates on and off in a precise order, charge is shifted from one gate to another, and eventually to readout electronics. $\Phi_H(x, z)$ can be calculated analytically as,

$$\Phi_H(x, z) = V_T + \frac{V_G - V_T}{\pi} \cdot \left(tan^{-1}\left(\frac{z + d_{ox}}{x - L/2} \right) - tan^{-1}\left(\frac{z + d_{ox}}{x + L/2} \right) \right) \quad [55]. \quad (5.2)$$

For the values shown in Table 5.1, the homogeneous potential is shown in Figure 5.7 where the minimum in potential beneath the gate is clearly visible.

(a)   (b)

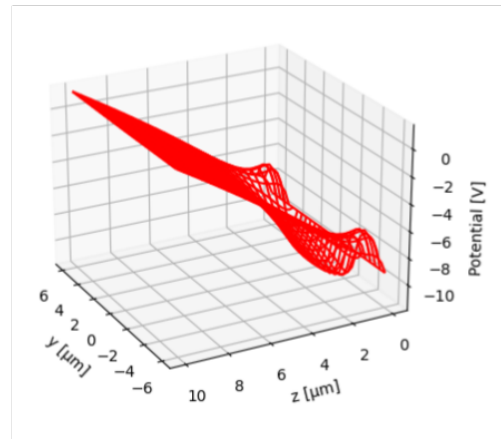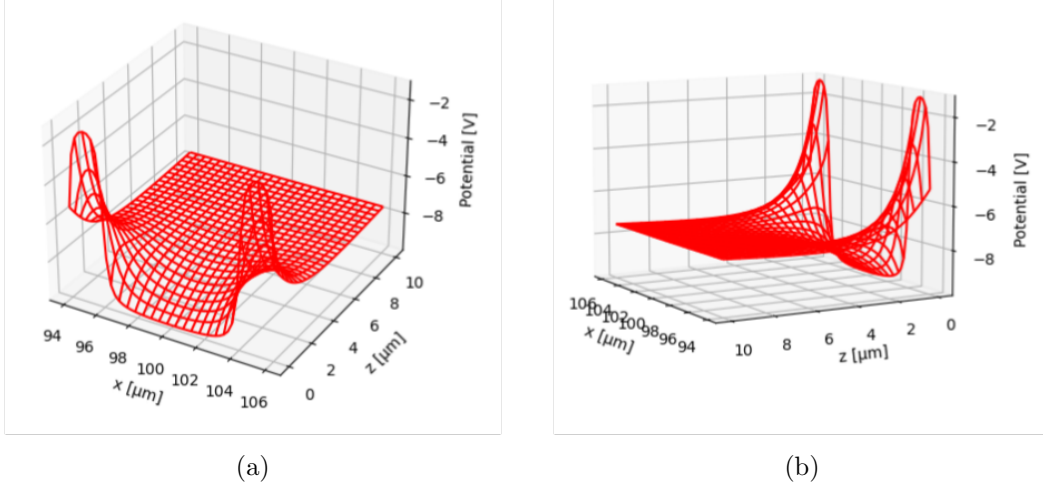Figure 5.7: The homogeneous potential, $\Phi_H(x, z)$, is derived from the overlapping gates. For a 4-phase CCD, three gates are biased at $10\,\text{V}$ while the fourth is held at $0\,\text{V}$. The gives a well in the potential for electrons to drift towards, defining the bounds of a pixel.

$\Phi_B(x, z)$ is called the backside potential and is from the potential difference applied between the front and back of the device [55]. For CCD250, $V_{BS} = -70\,\text{V}$ and is responsibly for full depleting the CCD and providing a signficant gradient in potential which causes the electrons to drift towards the potential well. It can also

be solved analytically such that,

$$\Phi_B(x,z) = \left(\frac{V_{BS} - \Phi_H(x, z_T)}{z_T + \frac{\epsilon_{si}}{\epsilon_{ox}}d_{ox}}\right) \cdot \begin{cases} \frac{\epsilon_{si}}{\epsilon_{ox}}(z + d_{ox}) & z < 0 \\ \left(z + \frac{\epsilon_{si}}{\epsilon_{ox}}d_{ox}\right) & z > 0. \end{cases} \tag{5.3}$$

$\Phi_E(x, y, z)$ is the electrostatic (Coulomb) potential and represents the potential from an electron, felt by other nearby electrons. The potential felt by an electron is,

$$\sum_{n}^{N-1} n \frac{1}{4\pi\epsilon_{si}\epsilon_0} \frac{q}{r_n} e^{-k_0 r_n}, \tag{5.4}$$

where ϵ_0 is the permittivity of free space, $q = 1.6 \times 10^{-19}$ C is the charge of an electron and r_n is the distance between the electrons [62]. N represents the total number of free electrons which can exert a potential on another. The exponential represents the Thomas–Fermi screening approximation where,

$$k_0 \equiv \sqrt{\frac{mq^2 k_F}{\epsilon_0 \pi^2 \hbar^2}}, \tag{5.5}$$

is the Thomas–Fermi screening wavevector [62]. k_F is the Fermi wavevector, $\hbar$ is the reduced Planck constant and m is the mass of an election. This is where the potential field is damped by the existence of mobile charge carriers such as those present in doped silicon [62]. This very quickly reduces the effect of the coulomb potential as the distance between electrons increases. While this can have an interesting effect during the initial movement of charge, the screening effect means that this potential can be largely ignored when considering the full thickness of the detector.

$\Phi_C(x, y, z)$ is the stored charge potential which originates from charge stored in the potential well [55]. As more charge enters the well, it begins to fill. There is a limit to the number of charges which the well can hold and over time it will become more likely that a charge is repelled and drifts into a neighbouring pixel. This potential acts to repel further charges from entering the well as it fills up. In this work, 5.9 keV

X-rays from an Fe-55 source are simulated. It is assumed that all pixels begin empty and hence $\Phi_C(x, y, z)$ can be ignored in this case.
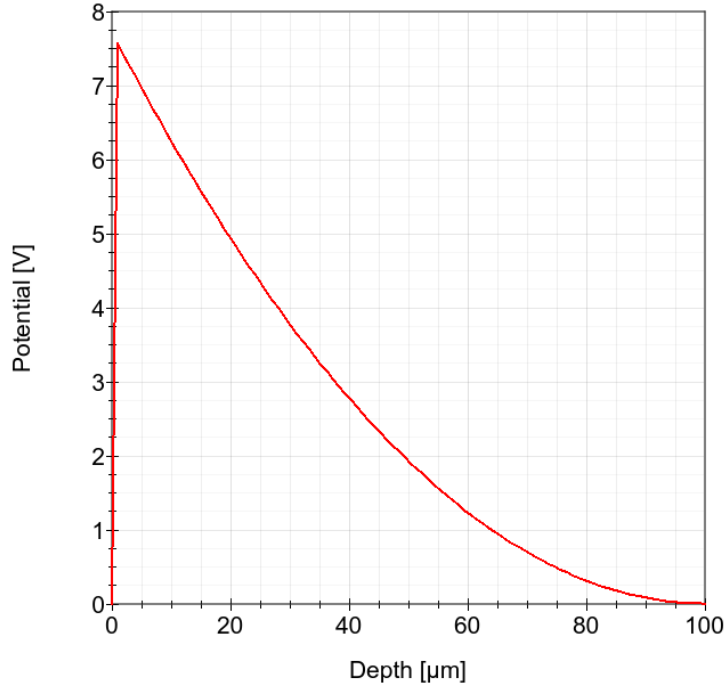
Figure 5.8: The particular potential calculated analytically using the doping concentrations of the n- and p-type silicon. An abrupt boundary is assumed and results in the peak in potential. This helps to define the potential well in the z-axis.

$\Phi_D(x, y, z)$ represents the potential due to the doping structure of silicon. This term encompasses other potentials, two of which are discussed by [55] in more detail. These are the particular potential and the charge stop potential. The particular potential acts as a barrier in the n-type region which complete the potential well in Figure 5.7. It can be calculated analytically by solving Poisson's equation for electrostatics and results in the curve in Figure 5.8. However, this assumes an abrupt junction in the doping concentration which is not an accurate assumption, as shown in Figure 5.9. The charge stop potential is used to model the p-stop implants which help to define the pixel boundaries in the y-axis. In [55], this potential is modelled in the same way as $\Phi_C(x, y, z)$. However, when trying to model the p-stop accurately, this potential became dominant and acted to repel electrons from the potential well, which is not physical.

Figure 5.9: The doping concentrations of the (a) n-well and (b) p-stop regions as a function of depth provided by Te2v. The values of doping concentration have been obfuscated at the request of Te2v. These curves have been approximated with segmented blocks of uniform doping. Note that for the n-well, the doping concentrati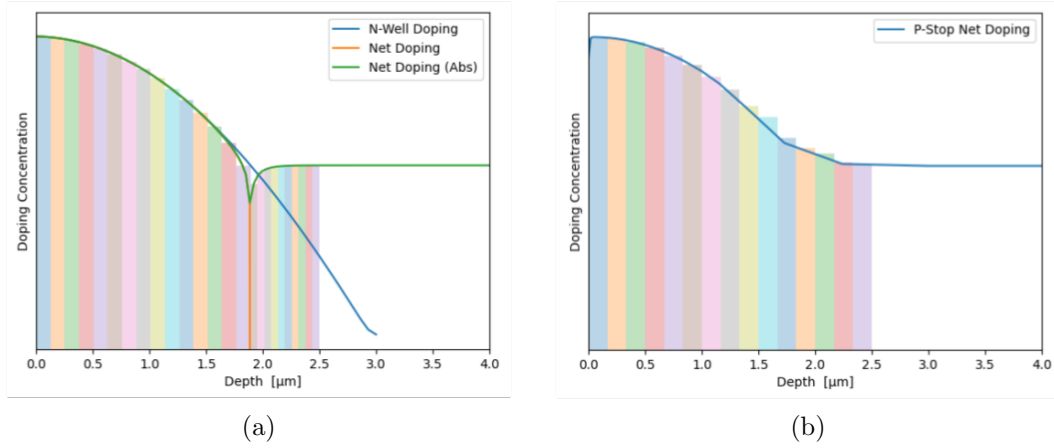on becomes negative after $1.89\,\mu$m (the junction depth) which represents the transition to the p-bulk. The approximation continues with a negative doping concentration (shown as positive here) to better represent the transition to p-bulk doping.

The doping profile in Figure 5.9 also shows the doping profile for the p-stop region. These profiles are the most accurate way to model $\Phi_D(x, y, z)$ for these devices, however they can only be modelled numerically. Therefore, the finite element method (FEM) is employed which allows differential equations to be solved numerically [63]. A program called Finite Element Method Magnetics (FEMM) is able to solve electrostatic problems given detailed structures [63]. This is shown in Figure 5.10 where a cross section of the CCD is shown in the $y - z$ plane. One can assume symmetry in the x-axis since only the homogeneous potential is required to define the pixel boundaries in the x-axis. In Figure 5.9, the doping profiles for both the n-well and p-stop are approximated with small segments of uniform doping. This segmented approximation is applied to FEMM and is shown in Figure 5.10. The potential field is calculated over a mesh which favours higher detail at the various boundaries. This is shown as a heat-map in Figure 5.10 and it is clear how the potential curves in around the pixel boundaries, driven by the p-stop.
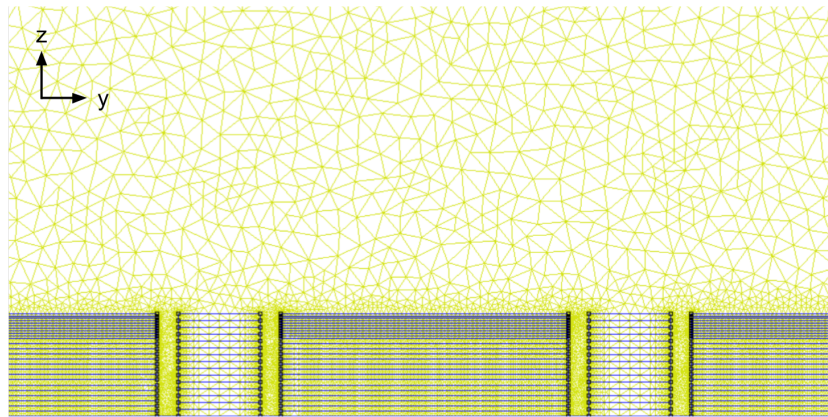
(a)
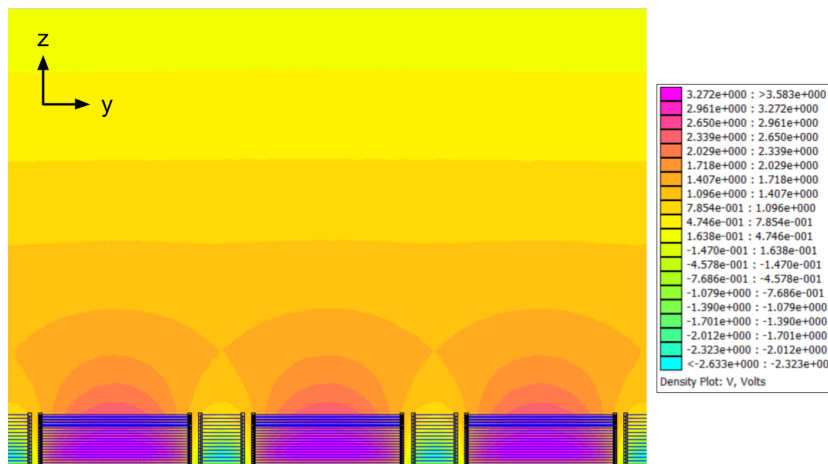
(b)

Figure 5.10: FEMM is a progamme which can solve complicated electrostatic problems, including simulation of the potential due to the n-well and p-stop implants. (a) The yellow mesh has a higher density in regions where the potential field is more complex and requires a more precise approximation. The doping profile is approximated as blocks of a fixed charge density which matches the doping concentration profiles provided by Te2v. (b) The potential field is shown as a heatmap and the gradient between neighbouring pixels is clearly shown, defining their boundaries to incoming electrons from the bulk.
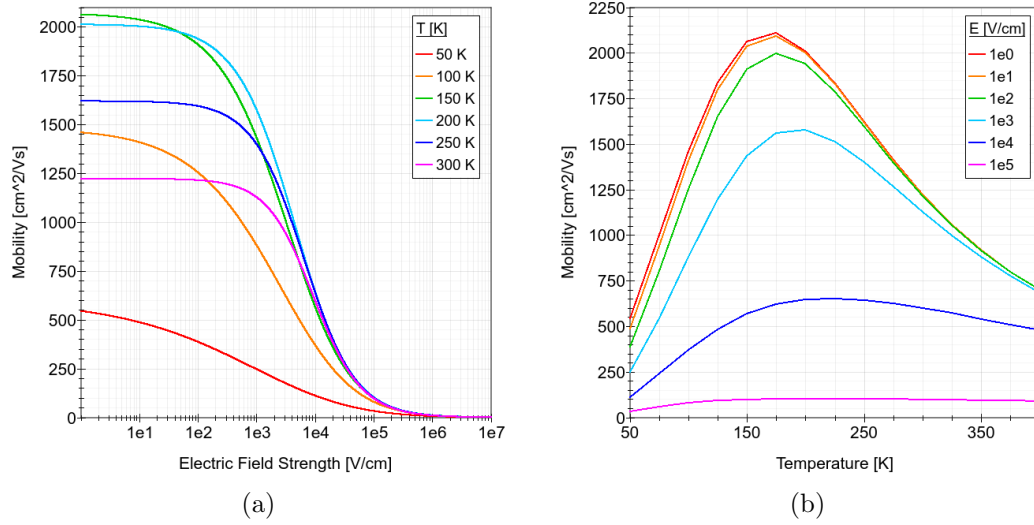
Figure 5.11: (a) Mobility decreases as the electric field strength increases. This is shown at multiple temperatures which show a more complicated dependence. (b) As a function of temperature, and more so for lower electric field strengths, the mobility peaks at roughly 175 K.

5.2.4 Mobility

Calculating the movement of electrons requires the field-dependent mobility, $\mu(\mathbf{E})$ to be correctly modelled. The electric field strength, $\mathbf{E}$, can be calculated from the potential field, discussed in Section 5.2.3 and Section 5.3.1. The mobility can be calculated analytically and the resultant dependence on the electric field strength, $\mathbf{E}$, is shown in Figure 5.11. There is minimal field dependence for low or high field strengths, however at 1×10^3 V/cm there is a strong dependence on $\mathbf{E}$. Equally there is a very strong dependence on temperature below 1×10^5 V/cm. Above this, the dependence disappears. For further details on the calculation of mobility, the reader is directed to [64].

5.2.5 Random Walk

When electrons are highly concentrated, they experience diffusion which aims to drive them apart, eliminating any concentration gradient. This diffusion is the

primary mechanism behind charge spreading and gives rise to the PSF, discussed in Section 5.2.2. Diffusion can be modelled by a random walk where the electron moves in a random direction with a magnitude proportional to the diffusion coefficient, D. The diffusion coefficient is given by the Einstein relation such that,

$$D = \alpha \frac{\mu(\mathbf{E})k_B T}{q}, \tag{5.6}$$

where k_B is the Boltzmann constant [15]. α is an additional diffusion factor. This allows for the diffusion to be tuned to see how a stronger diffusion influences the PSF. In all results presented, $\alpha = 40$. The reason why this factor is required is so far unclear. Since the mobility is dependent on the field strength, one could argue the diffusion coefficient needs to be split into parallel and perpendicular components which could be important since horizontal diffusion is the determinant for charge spreading. By definition, the perpendicular component is a zero-field region which would increase the mobility and hence the diffusion coefficient. The idea that diffusion has directionality has been discussed before, although the difference is a factor of two higher at best [65, 66, 67]. Nevertheless, understanding why an additional diffusion factor is required should be the focus of future work.

Over a total collection time, t_{col}, the total diffusion is approximately a Gaussian with a variance given as,

$$\sigma^2 = 2Dt_{col} \text{ [55]}. \tag{5.7}$$

While this is a good indication of the overall diffusion, it is insufficient when calculating σ for CCD250. This is because the intricacies and details of the potential fields are not taken into account as they become dominant close to the electrodes. For this a Monte Carlo simulation is required.

5.3 Simulation Framework

5.3.1 Drift & Diffusion

If the forces acting upon an electron are known, then its motion throughout the silicon device can be modelled and simulated. A Monte Carlo simulation allows multiple X-ray absorption events to be simulated independently and then analysed together. In the first version of this simulation, X-rays are assumed to be absorbed $30\,\mu$m into the backside of the CCD [68]. Future versions will need to take into account the distribution of absorption depths. For a $5.9\,$keV X-ray, 1600 electrons are generated in a flat distribution with a radius of $0.19\,\mu$m about the absorption site, a quick approximation of the initial charge cloud [69]. Each electron's motion needs to be simulated. They are largely independent except in the case where the electrostatic potential is considered. Although it is not considered in this first iteration, the simulation has been designed so that it maybe be considered in the future.

(a)   (b)
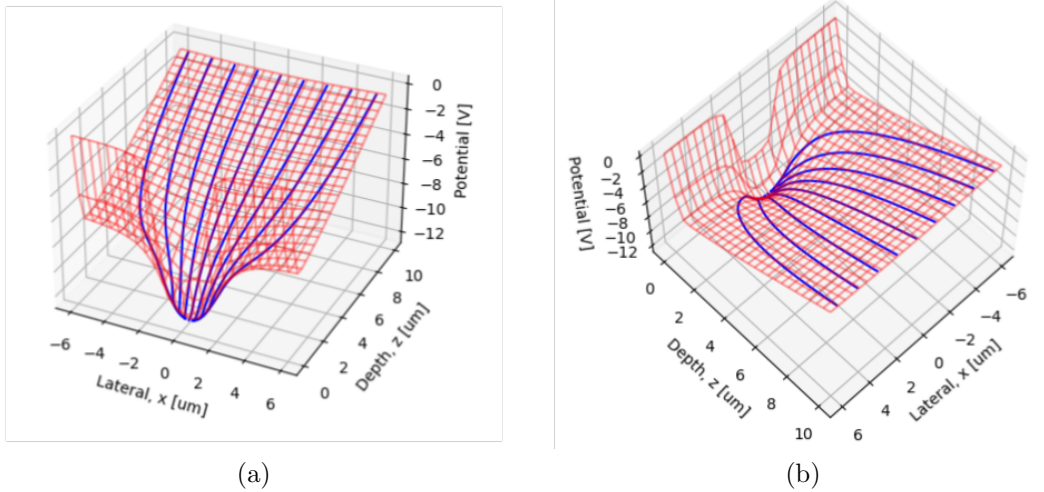
Figure 5.12: Nine electrons start at $50\,\mu$m at a range of x positions. They are simulated with drift only and predictably tend towards the potential well. Their paths are shown in blue

The simulation operates on a series of time steps, δt, for which $1\,$ps is chosen. This value ensures that the intricacies of the potential fields closer to the electrode

are preserved while keeping the simulation relatively fast. During each time step, potential fields cause the electron to drift. This is modelled as,

$$\begin{pmatrix} x_{i+1} \\ y_{i+1} \\ z_{i+1} \end{pmatrix} = \begin{pmatrix} x_i \\ y_i \\ z_i \end{pmatrix} - \mu(\mathbf{E})\delta t \begin{pmatrix} E_x \\ E_y \\ E_z \end{pmatrix} \ [55]. \tag{5.8}$$

Here, $\mathbf{E}$ is the electric field. The electric field is calculated as the gradient in potential such that,

$$\mathbf{E}(x,y,z) = -\nabla\Phi(x,y,z) \approx \begin{pmatrix} \frac{\Phi(x+\Delta x,y,z)-\Phi(x-\Delta x,y,z)}{2\Delta x} \\ \frac{\Phi(x,y+\Delta y,z)-\Phi(x,y-\Delta y,z)}{2\Delta y} \\ \frac{\Phi(x,y,z+\Delta z)-\Phi(x,y,z-\Delta z)}{2\Delta z} \end{pmatrix} \ [55]. \tag{5.9}$$

Here, the gradient is approximated numerically using a small perturbation, $\Delta x = \Delta y = \Delta z = 1\,\mathrm{nm}$, calculating the electric field in each axis independently. This perturbation is small enough to distinguish the intricacies of the potential field structure, without being unnecessarily small which would substantially increase the error associated with the approximation. At each time step, the total potential field, Φ_{tot} must be calculated six times (twice per axis to estimate the gradient) using the methods described in Section 5.2.3. Each electron is updated using Equation 5.8 before the next time step starts. This continues until an electron is deemed to have reached the n-well of a pixel. It is assumed that once an electron has reached the physical boundaries of the pixel, as defined by the gates and channel stop, it cannot leave and is treated as collected. The electron is not simulated after this point. The simulation continues until all electrons are collected. To illustrate the simulation, nine electrons have been simulated starting at various x positions along the backside. They are only under the influence of the potential fields and so drift, as expected, into just one pixel. Their paths are shown in Figure 5.12 where the y-axis has been ignored for simplicity.
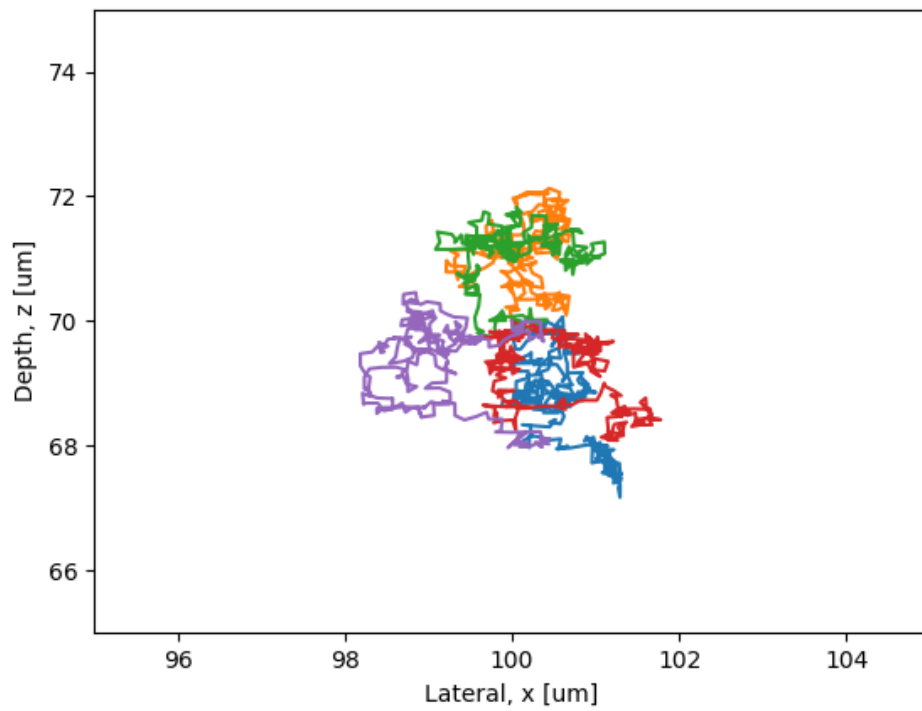
Figure 5.13: Five electrons, starting at $(100, 70)$ μm, are allowed to diffuse for a fixed amount of time. Their paths are shown and highlight the random and unpredictable nature of a random walk.

Thus far, the model only takes drift and not diffusion into account. As discussed in Section 5.2.5, variance in a particles position is proportional to the diffusion coefficient and time. This can be expanded and applied to each time step after drift has taken place. In this model,

$$
\begin{pmatrix} x_{i+1} \\ y_{i+1} \\ z_{i+1} \end{pmatrix} = \begin{pmatrix} x_i \\ y_i \\ z_i \end{pmatrix} + \sqrt{D\delta t} \begin{pmatrix} \sin\theta\sin\phi \\ \cos\theta\sin\phi \\ \cos\theta\cos\phi \end{pmatrix}, \tag{5.10}
$$

where θ and ϕ represent polar coordinates which are randomly chosen such that they are unique for each time step and each electron [55]. Figure 5.13 shows the diffusion component only for five different electrons, showing the random walk nature of diffusion over a fixed time.

5.3.2   Pixel Array

While an imaging sensor typically has millions of pixels arranged in a regular matrix, the full matrix does not need to be simulated here. A 7×7 array is ideal to minimised the simulation time and is shown in Figure 5.14a. The potential from additional pixels in the x-axis are calculated by calculating the homogeneous potential multiple times each with a different x displacement to represent the positions of each pixel. Additional pixels in the y-axis are taken into account by FEMM as shown in Figure 5.10. However, Figure 5.14a also shows that the potential at $z = 10\,\mu m$ curves towards the centre of the array. This is an edge effect which originates from the homogeneous potential, also shown in Figure 5.14, which does not have any termination in the x-axis. In Figure 5.14b, this effect has been minimised by expanding the array from 7 to 21 pixels. This is because the edge effects have dissipated once they reach the centre of the array. While more pixels would minimise this effect further, it also increases the computation time, since the contribution of the homogeneous potential from every pixel needs to be calculated. 21 pixels was chosen as

an ideal balance. This effect is also seen in the y-axis, but to a lesser extent. In this case, only 11 pixels are used since the calculation by FEMM takes much longer the more pixels are added. Overall this means a 21×11 array is simulated.
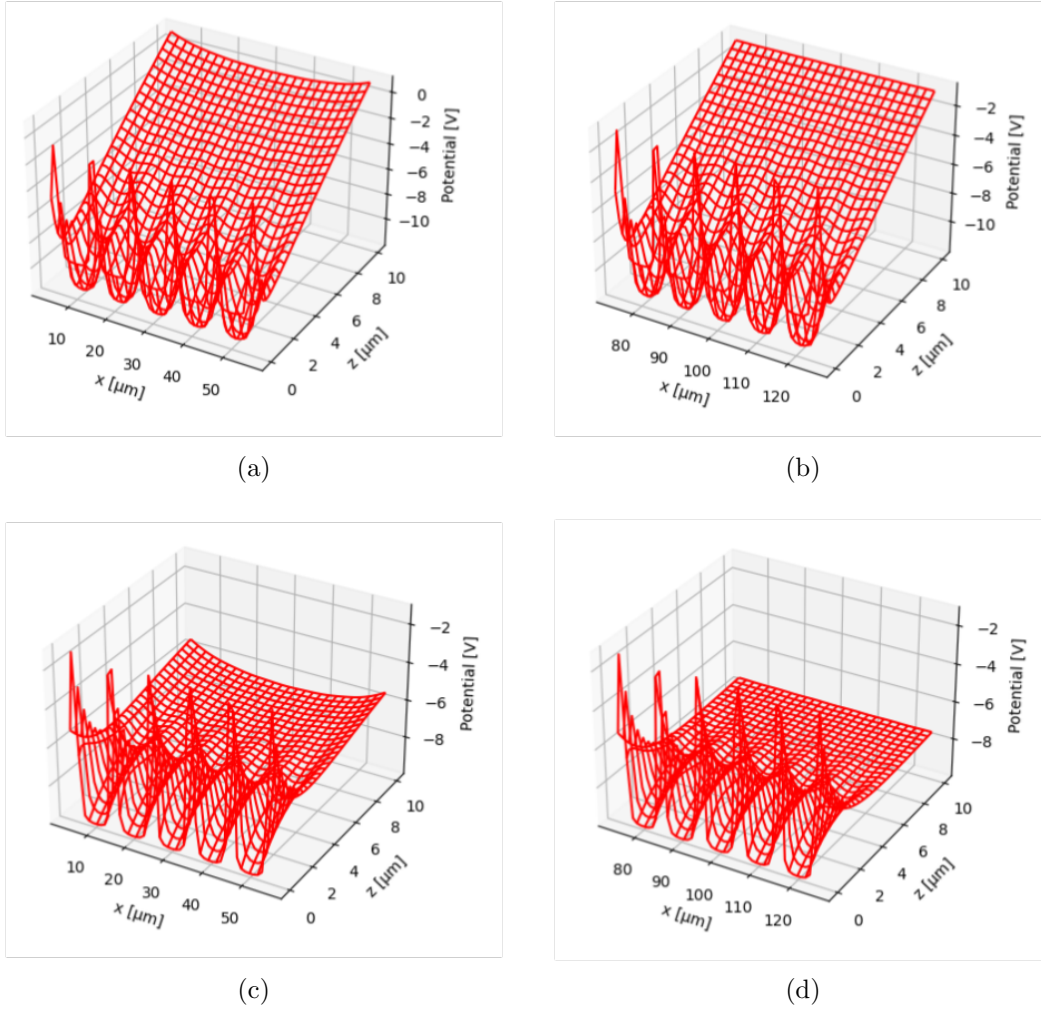
Figure 5.14:   The potential field of a pixel array is influenced by the number of pixels. Too few pixels and the edge effects dominate in the centre. This is shown in four cases: (a) Total Potential 7×7, (b) Total Potential 21×21, (c) Homogeneous Potential 7×7 and (d) Homogeneous Potential 21×21.

This curvature can be better characterised by simulation where both drift and diffusion is taken into account. 100 electrons are released from a particular x coordinate at a depth of 90 um and the number of electrons collected in each pixel is counted and recorded. This is repeated for a range of x values from the centre of the array to the centre of the next pixel ($10\,\mu$m pitch). The result is shown in Figure 5.15a where the black line represents the pixel boundary. If there was no curvature in the

potential field, then one would expect the cross over point of the two curves to be on the pixel boundary exactly. Since there is a curvature, the cross over point exists a few microns into the central pixel. This is a problem because it means that electrons are statistically more likely to fall into the pixels closer to the centre. Figure 5.15b shows the effect of increasing the number of pixels to 21. While it is still not perfect, the bias towards central pixels is reduced.

(a)   (b)
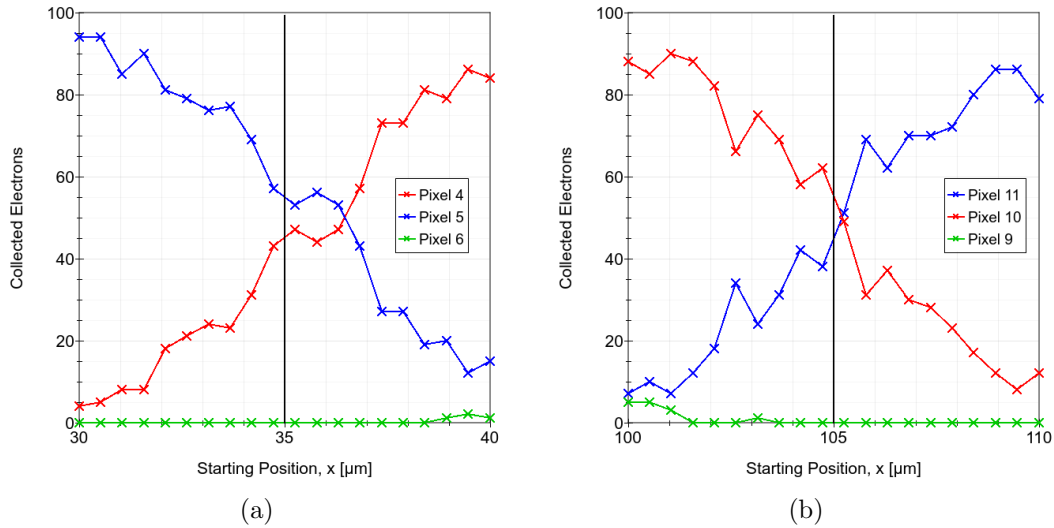
Figure 5.15: 100 electrons are deposited at a range of x positions and the pixel they are collected by is recorded. In an ideal world, with no edge effects, the cross over between two pixels is should be at the boundary. This is not true if too few pixels are simulated. As the number of pixels increases, the cross over point tends towards the boundary. This is shown in the case of (a) a 7×7 array and (b) a 21×21 array.

With the curvature problem minimised, Figure 5.16 shows the combination of drift and diffusion for three different starting locations, $70\,\mu$m from the potential well. This highlights the variation in path an electron can take from the same starting location, leading to different pixels from the one they started in. The black line represents the pixel boundaries. As before, the y-axis is ignored for simplicity.
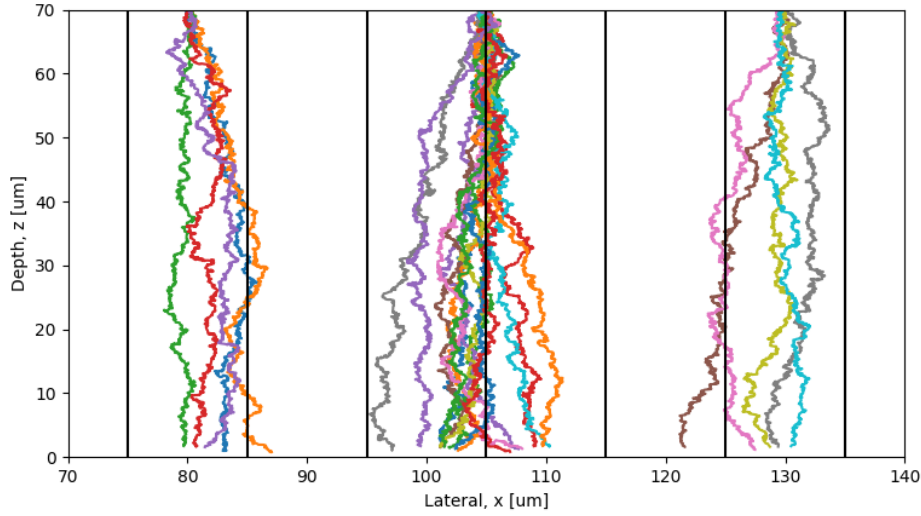
Figure 5.16: Drift and diffusion are combined to show a typical electron's path through a pixel array. The black lines represent the boundaries between pixels. Three locations are tested: two in the centre of a pixel and another at the boundary between two pixels. These highlight how the starting position of the electron influences its likely destination.

5.3.3 Fail Case

If electrons venture too far from the array, or take a significantly long time without being collected, they are recorded as a failed electron. Generally speaking, all electrons make their way to at least one pixel in the array. A common place where electrons can fail is at the top and back side of the sensor. When first set in motion close to the backside of the sensor, diffusion can be strong enough to move electrons to a value of z greater than the $100\,\mu$m thickness. Similarly, once electrons reach the n-well, they can sometimes move fast enough to jump over the n-well to a value of z lower than $0\,\mu$m. In both examples, the rudimentary solution has been to clamp the electrons to $0\,\mu$m $< z < 100\,\mu$m. So if an electron is moved above $100\,\mu$m in a time step, this is detected and the electron is moved back to $z = 100\,\mu$m. The first motivation for this is that it is unphysical for an electron to leave the silicon lattice and so any attempts to do so would be opposed. The second motivation is this method should not impact the results significantly. Electrons at the backside of the sensor are early in their lifetime and electrons at the n-well are most likely

already heading to one pixel or another.

5.4   Analysis Method & Preliminary Results

5.4.1   Experimental

In order to extract information about the PSF from simulation, one can look to how Te2v sensors are measured experimentally. CCD250 is first exposed with an Fe-55 source, producing $5.9\,$keV X-rays [61]. After a sufficient length of time, the source is removed and the resultant image is processed. Each pixel has an intensity associated with it which is proportional to the number of electrons collected during integration [61]. Assuming an X-ray is absorbed in the middle of a pixel, the distribution of electrons collected is shown in Figure 5.17 as a average from experimental data. This is an example of a single cluster where electrons are shared (in a Gaussian distribution) between neighbouring pixels. When analysing the entire sensor area, care must be taken to identify and discard double clusters (two overlapping Gaussians) as these will be too complicated to be worth analysing.

The PSF is responsible for the shape of a single cluster. This means that the standard deviation, σ, of the PSF can be extracted from one of these clusters [61]. To do so, it is assumed that the PSF is a one dimensional Gaussian with an amplitude, A, 2D position, $\mathbf{r}$, and standard deviation, σ. All three of these can be fitted to a single cluster hitmap, similar in shape to the average seen in Figure 5.17, collected by [61]. The Gaussian can be integrated between the bounds of each pixel. σ is found by minimising the squared difference between the calculated integral and the measured intensity of each pixel.

The extraction of σ is repeated many thousands of times for different clusters [61]. From this, a histogram can be formed as shown in Figure 5.18. This histogram peaks at $\sigma = 3.62\,\mu$m and the peak has a width at half maximum of roughly $0.59\,\mu$m. The
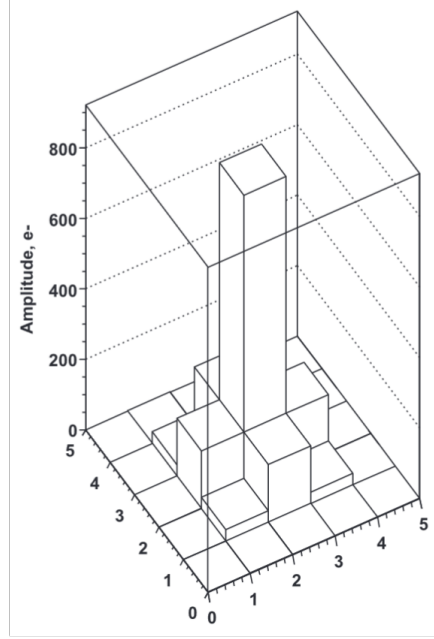
Figure 5.17: Assuming an X-ray is deposited in the centre of a pixel, this is the average expected number of electrons collected (during experimentation) by each pixel. Taken from [61].

asymmetry arises from the exponential decay in absorption of X-rays. A majority are absorbed within the first $20\,\mu$m, but some will make it much deeper.

5.4.2   Simulation

The same methodology described in the previous section can be applied to the simulation to extract the PSF's σ. However, symmetries can be exploited in order to simplify the methodology. Rather than exposing the entire sensing area, a small 5×5 sample of pixels in the center of the sensing area can be tested. Since only single event clusters are considered, each X-ray absorption event can be treated as independent by the simulation. Rather than simulating random X-ray absorption sites, a regular grid of test sites can be used. To minimise simulation time, a 10×10 grid covering one quarter of the central pixel of the 5×5 matrix is used. This exploits the mirror symmetry of a rectangular pixel and is shown in Figure 5.19. For each test point on the grid, the number of electrons collected in each pixel is recorded. The same methodology used in experimentation can be used here to extract σ.

Figure 5.18: From each cluster, the PSF's σ can be extracted. If this is repeated enough times, a histogram can be plotted, as shown here. It peaks at $3.62\,\mu$m which is the value used to extrapolate to the MTF_{nyq}. Taken from [61]

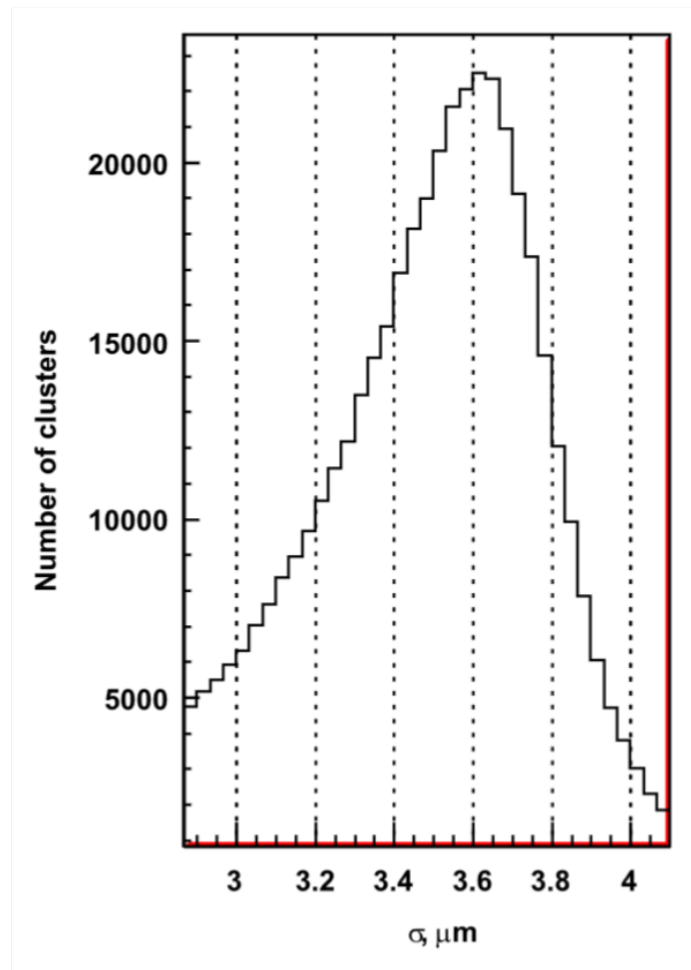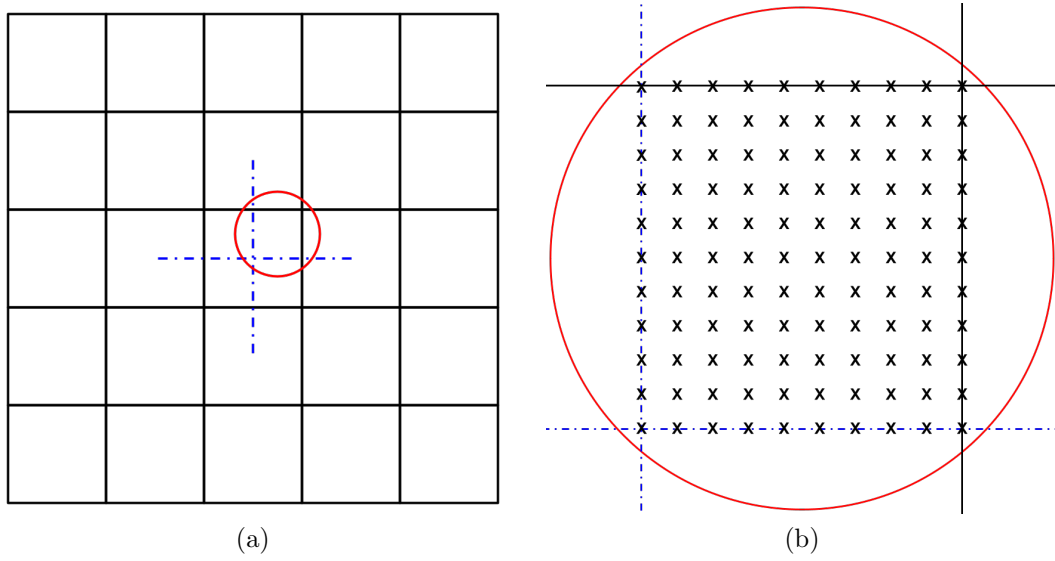Figure 5.19: (a) The number of electrons collected in a 5 × 5 sample of pixels is recorded in order to extract σ. (b) Exploiting symmetry, only one quarter of the central pixel needs to be tested. Within this, an array of 10 × 10 test positions can be used and σ can be calculated for each start position.

Ideally, all three parameters should still be fitted, as done in experimentation, but for simplicity (and to improve the reliability of extracting a sensible σ) A and $\mathbf{r}$ have been fixed so that only σ is varied and fitted to. This can be done since A and $\mathbf{r}$ are defined by the simulation. The simulation of the 10 × 10 test position grid and the extraction of σ has been repeated 20 times to provide a high statistic sample. This results in a heatmap of the average and standard deviation of σ at each test position, shown in Figure 5.20.

Variation in σ is generally minimised, as expected. However, as test positions tend closer to the corner of the pixel than the centre (located at $(25, 25)$ μm and $(20, 20)$ μm respectively), the average of σ decreases while the standard deviation of σ increases. This is also visualised in Figure 5.21 where the average and standard deviation of σ is shown as a function of distance from the centre of the pixel in the direction of the corner. It shows that the upper bound of σ remains unchanged, while the lower bounds decreases significantly, hence why the average of σ decreases while the standard deviation increases. This is caused by low statistics in the hitmap. Take the case where the test position is over the corner exactly.

All of the electrons are collected in the four pixels which surround the corner. In this scenario, it is only possible to calculate the upper limit of σ, since any value below this limit would produce exactly the same hitmap. For the corner test position specifically, the initial conditions of the fit are set slightly above this limit, and hence the algorithm finds a consistent value, keeping the standard deviation small. As soon as test positions deviate slightly from the corner, the electrons are no longer distributed evenly among the four neighbouring pixels and a very small number will start to be collected by other pixels. This means a lower bound can now be set on σ, but not very accurately since pixels which only collect a small number of electrons have a relatively large statistical error. This lower bound then increases as the test position moves away from the corner and the statistical error decreases.

(a)   (b)
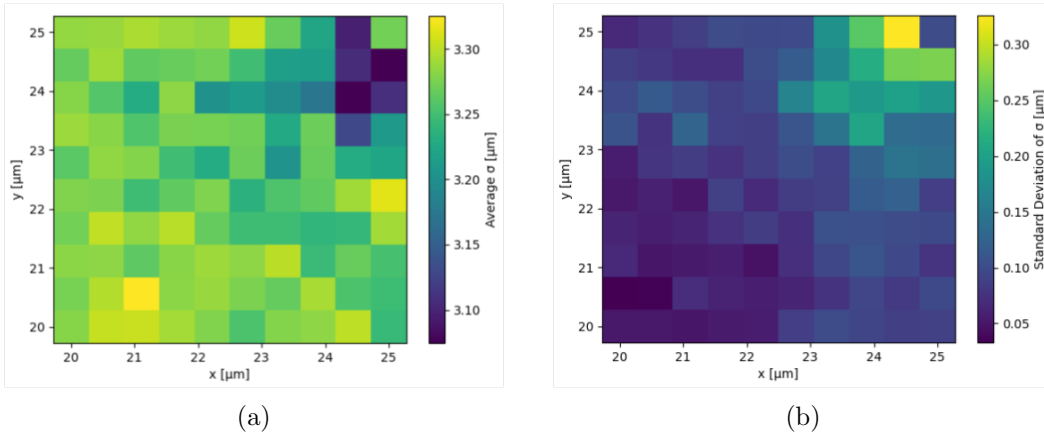
Figure 5.20: After 20 independent simulations of the 10×10 test grid, σ at each location, for each repeat, can be calculated. Over the 20 repeats, the (a) average and (b) standard deviation are shown. In the corner of the pixel, the average decreases while the standard deviation rises. This is a consequence of insufficient statistics when reconstructing σ in this region (discussed more thoroughly in the text).

Once the 10×10 grid of test positions has been simulated a histogram can be produced, similarly to the one shown in Figure 5.17. Experimentally this is done by taking a high statistics of single event clusters at random locations. It is assumed that the distribution of X-ray absorption locations is a flat random distribution over the entire sensing area. This is justified in the case of the 5×5 sample if it is assumed to be in the centre of the sensing area. Since a regular grid has been used by the simulation, a histogram can be produced from just those 100 samples.

Figure 5.21: Starting at the centre of the pixel and moving towards the corner, σ as a function of distance is shown. This shows the same trend seen in the heatmaps where there is insufficient statistics to calculate σ (discussed more thoroughly in the text).

To reduce the error in the histogram and allow a smaller bin width, one can either increase the number of samples in the test grid (up to 20×20, for example), or repeat the 10×10 test grid, as is done here, 20 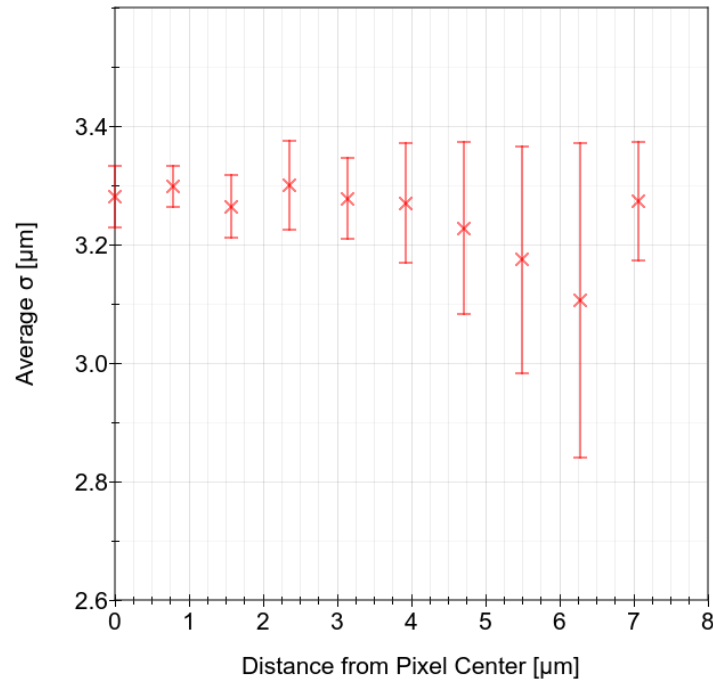times. Figure 5.22 shows a histogram of these 2000 samples. It also shows a histogram of the 100 samples from the averaged heatmap in Figure 5.20a. This second histogram shares the same peak as the first, but with a much small width, as expected from an average sample set. The peak at $\sigma = 3.27 \, \mu$m is similar to that found through experimentation, at $3.62 \, \mu$m. For reference, the method defined in Section 5.2.2 can be used to convert σ to MTF_{nyq}. For simulation and experimentation respectively, MTF_{nyq} is found to be 0.375 and 0.333. Unfortunately, these similarities do not have any meaning since the diffusion coefficient is arbitrarily increased by a factor of 40. However, since the peak is in a similar position, one can compare the widths of the peak. From simulation, the width at half maximum is roughly $0.18 \, \mu$m which is a little over three times smaller than for experimentation, at $0.59 \, \mu$m. The reason for this difference may be due to the depth at which X-rays are absorbed. In this particular simulation, all X-rays are absorbed $70 \, \mu$m. In reality, X-rays are absorbed at a range of depths. X-rays absorbed deeper have less time to diffuse whereas X-rays absorbed shallower have more time to diffuse, leading to a smaller and larger σ respectively. This variation will lead to an increase in the width at half maximum in Figure 5.22. The peak position will be dependent on the average absorption depth. The variation in depth also explains the asymmetry in Figure 5.18 (not seen as strongly in Figure 5.22) since X-rays absorbed shallower are limited by the thickness of the device, applying an upper limit to σ while there is no lower limit.

5.5   Conclusion & Outlook

The goal of this work was to develop custom software to predict a CCD's MTF via the PSF, given the specification of the device. The metric to test this against was CCD250 of the LSST project. After significant development and iterations to the simulation, it has been demonstrated that the simulation is able to simulate electron
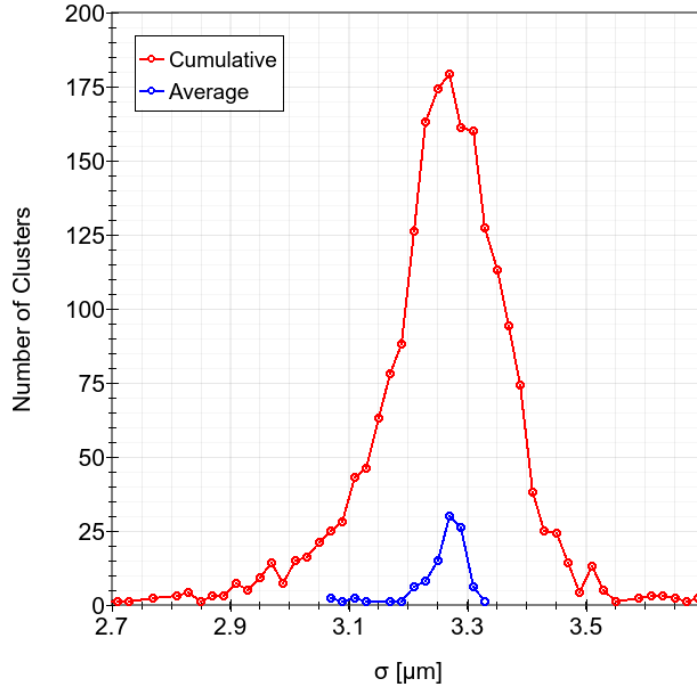
Figure 5.22: By sampling the interpolated heatmap, the same histogram of clusters, shown for experimental data, can be plotted for simulation. This peaks slightly lower at $3.27\,\mu$m and with a much narrower peak, however the general shape is similar.

motion through a silicon sensor and that the PSF's σ can be extracted through this process. The simulation can also replicate the experimental method used to measure CCD250 with Fe-55 X-rays. The peak value of σ is comparable between experimentation and simulation, $3.62\,\mu$m and $3.27\,\mu$m respectively. This means that the width of the peaks, which from experimentation is roughly 8.5 times larger than in simulation, can be compared and suggests further iterations are required. The similarity in the shape of the peaks suggests that the simulation is in some agreement with experimentation. However the artificial increase of the diffusion coefficient by a factor of 40 means that the simulation is not yet fully in agreement with experimentation. The reason for this factor is still unclear and reiterates the need for further investigation and improvement.

Time constraints meant that this project was halted before reaching a satisfying conclusion. The following are recommendations for future investigations should this project continue. Firstly, the distribution of X-ray absorption depths must be con-

sidered to understand to true distribution of the PSF's σ. This will increase the simulation time considerably, but it is necessary in order to reproduce the experimental data more accurately. Secondly, the factor applied to the diffusion coefficient must be understood. One approach would to attempt to simulate CCD250 with TCAD and the Allpix squared framework in order to make direct comparisons with this work [70, 71, 72]. This is a much more in-depth simulation package, but as a trusted source, it should provide insight into where the simulation presented here is going wrong. Finally, a series of commissioning tests should be performed to confirm the accuracy of this simulation. For example, investigating the dependency of X-ray absorption depth (following the implementation of the depth distribution). This should also include the testing of a separate device with different implant geometries, eliminating the risk that the agreement shown here is a coincidence.

The similarities between simulation and experimentation leave room for optimism. However, as state above, there are a number of areas that need to be explored before this project can be considered a success.

CHAPTER 6

CONCLUSION & OUTLOOK

The goal of this work was to characterise LGADs manufactured by Te2v in order
to establish them as a vendor for future particle physics applications. To achieve
this, test set-ups and experimental techniques were developed to investigate IV,
CV, gain and timing properties of five flavours of LGAD wafers. These properties
can be measured after irradiation which requires a consistent method of cooling to
preserve a low leakage current during characterisation. After initial commissioning
and subsequent adjustment, characterisation has been completed with a high success
rate. This work has demonstrated that these set-ups and techniques have matured
and are capable of testing a large volume of devices with good quality data. The
work put into developing this leaves the project in a strong position to further
characterise LGADs with Te2v should the project continue and further batches be
fabricated.

Te2v's first batch of $50\,\mu$m LGADs were produced in eight different wafer flavours
each with a unique combination of boron implant energy and dose for the gain layer.

Initial IV and CV characterisation took place on five wafers before dicing. These tests confirmed the expectation that the higher the combination of implant energy and dose, the lower the average breakdown voltage of devices from that wafer. This is in further agreement with the gain layer depletion voltage which increases for a higher combination of implant energy and dose. What also became clear was the doping non-uniformity within each wafer. The breakdown voltage varied by roughly 25 V in some cases which indicates a large variation in gain dependence on bias voltage.

Four wafers were diced and a sample of devices was tested for IV, CV, gain and timing. IV and CV measurements were performed before and after thermal annealing with the goal of improving the electrical performance by reducing damage due to dicing. On a case-by-case basis this was not always the case. However, the average breakdown voltage did see an increase while the variance fell, indicating a net improvement after thermal annealing. CV curves saw no significant changes, as expected. Gain measurements were in agreement with the trends found in IV and CV measurements. Three of the wafers tested could achieve a high gain of $\sim$100, but the average bias voltage required was lower for a higher combination of implant energy and dose. For the fourth wafer (low implant energy and dose), barely a gain of 10 could be achieved before the device entered breakdown. This is to be expected as is the large variance in gain, seen also with the IV curves. Timing measurements further agreed with gain measurements. A higher combination of implant energy and dose results in a good time resolution at a lower bias voltage. As a function of gain, the time resolution shows no dependence on implant energy and dose. It is very important here to consider the operation of LGADs in experiments like ATLAS and CMS. In these experiments, all LGADs will need to be biased at the same voltage. Under these conditions, there is considerable standard deviation on the time resolution, ranging from 4 ps up to 9 ps. This presents a challenge, further limiting the ability to set an accepted time resolution of Te2v devices.

The wafer with the highest combination of implant energy and dose was chosen for an irradiation campaign. Seven devices were irradiated with a fluence from

5.3×10^{13} up to 8.3×10^{14} $1\,\text{MeV}\,\text{n}_{\text{eq}}/\text{cm}^2$. IV, CV, gain and timing measurements were performed before and after irradiation allowing the effects of irradiation to be studied. IV measurements showed the expected trend where a higher fluence resulted in larger increase in breakdown voltage compared to the pre-irradiation case. CV measurements allowed the gain layer depletion voltage to be extracted. The ratio before and after irradiation allowed the acceptor removal coefficient to determined as $c_A = (9.7 \pm 0.5) \times 10^{-16}\,\text{cm}^2$. This is $45\,\%$ larger than for similar devices fabricated in FBK's UFSD2 production where $c_A = 6.7 \times 10^{-16}\,\text{cm}^2$. Gain measurements were in agreement with IV results. The higher the fluence, the larger the voltage required to achieve maximum gain. However, high fluences were unable to achieve as high of a maximum gain compared to the pre-irradiation case, falling below 10 at $5.7 \times 10^{14}\,1\,\text{MeV}\,\text{n}_{\text{eq}}/\text{cm}^2$. Timing measurements show a similar trend with fluence. A higher fluence resulted in a higher bias voltage required to achieve a good time resolution. However, at $5.7 \times 10^{14}\,1\,\text{MeV}\,\text{n}_{\text{eq}}/\text{cm}^2$, a time resolution below $55\,\text{ps}$ was not achieved at any voltage since the LGAD's gain could not exceed 10. As a function of gain, the time resolution's dependence on fluence was removed. However, it should be noted that for low values of gain the time resolution improved after irradiation. A few possible explanations are presented, but further experimentation is required to confirm where this effect originates. Lastly, the highest fluence, $8.3 \times 10^{14}\,1\,\text{MeV}\,\text{n}_{\text{eq}}/\text{cm}^2$ was not tested for gain and timing due to current instability issues. Since only one device was tested at each fluence, repeat measurements are needed to confirm if this problem is due to a high fluence or whether the device itself was faulty.

A further thread of the thesis was the development of a simulation framework to calculate the MTF of CCDs produced by Te2v, as these are their main commercial product. The modelling was found to only match the experimental performance of CCD250 with a large, nonphysical, scale factor of 40 applied to the diffusion coefficient. However, it has been found that using the low-field mobility for lateral diffusion in place of the field-dependent mobility in the drift direction results in a much better agreement between simulation and experimental results. Similarly, the

implementation of an exponential absorption depth profile means the new simulation also agrees with the asymmetry seen in experimental results. This subsequent work has been included as Appendix A.

In summary for LGADs produced by Te2v, the high variance seen in IV and gain measurements should be a major focus for future batches of LGADs from Te2v. In particular, reducing the variance in time resolution at a fixed bias voltage is crucial for improving LGAD performance. This includes their usefulness as a timing layer at future HEP experiments. With the high radiation requirements of the HL-LHC, radiation hardness should be the second major focus for Te2v. Optimisation of the gain layer, including exploring a higher implant energy and dose and the inclusion of carbon enrichment is required to bring Te2v in line with other vendors. With these in mind, Te2v, on their first attempt, have managed to fabricated LGADs with comparable properties to other vendors pre-irradiation and post-irradiation for low fluence values. They have achieved a sub-40 ps time resolution with a number of wafer variants. While their radiation hardness requires improvement, it is still comparable with vendors using a similar implant process. Finally, this work has established Teledyne e2v as a future vendor for any applications which wish to employ the Low Gain Avalanche Detector.

REFERENCES

[1] G. Apollinari et al., *High-Luminosity Large Hadron Collider (HL-LHC): Preliminary Design Report.* CERN Yellow Reports: Monographs. CERN, Geneva, 2015, 10.5170/CERN-2015-005.

[2] I. B. Alonso et al., *High-Luminosity Large Hadron Collider (HL-LHC): Technical design report.* CERN Yellow Reports: Monographs. CERN, Geneva, 2020, 10.23731/CYRM-2020-0010.

[3] M. Casado, *HL-LHC physics with the ATLAS detector*, Nuclear and Particle Physics Proceedings **309-311** (2020) 43–48.

[4] "The High-Luminosity LHC (HL-LHC)." `https://hilumilhc.web.cern.ch/article/ls3-schedule-change`.

[5] ATLAS collaboration, *Technical Design Report: A High-Granularity Timing Detector for the ATLAS Phase-II Upgrade*, Tech. Rep. CERN-LHCC-2020-007, ATLAS-TDR-031, CERN, Geneva, 2020.

[6] CMS collaboration, J. N. Butler and T. Tabarelli de Fatis, *A MIP Timing Detector for the CMS Phase-2 Upgrade*, .

[7] ATLAS Collaboration, "Simulated HL-LHC collision event in the ATLAS detector." 2019.

[8] G. Pellegrini et al., *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 **765** (2014) 12–16.

[9] H. Sadrozinski, *Exploring charge multiplication for fast timing with silicon sensors*, 20th RD50 Workshop (Bari), 2012.

[10] H. F. W. Sadrozinski et al., *4D tracking with ultra-fast silicon detectors*, Rept. Prog. Phys. **81** (2018) 026101, [`1704.08666`].

[11] M. Ferrero et al., *Radiation resistant LGAD design*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **919** (2019) 16–26.

[12] C. Grieco et al., *Overview of CNM LGAD results: boron Si-on-Si and epitaxial wafers*, Journal of Instrumentation **17** (Sept., 2022) C09021.

[13] "Teledyne e2v." `https://www.teledyne-e2v.com/en-us`.

[14] M. Riordan and L. Hoddeson, *Crystal fire: the invention, development and impact of the transistor*. W. W. Nortan & Company, first edition ed., 1997.

[15] S. Sze and K. Ng, *Physics of Semiconductor Devices*. John Wiley & Sons, 3rd ed., 2007.

[16] H. Kolanoski and N. Wermes, *Particle Detectors Fundamentals and Applications*. Oxford University Press, first edition ed., 2020.

[17] F. Berghmans et al., *An Introduction to Radiation Effects on Optical Components and Fiber Optic Sensors*, pp. 127–165. 12, 2007. 10.1007/978-1-4020-6952-9_6.

[18] M. Ferrero et al., *An Introduction to Ultra-Fast Silicon Detectors*. CRC Press, 1st ed., 2021, https://doi.org/10.1201/9781003131946.

[19] S. Meroli, D. Passeri and L. Servoli, *Energy loss measurement for charged particles in very thin silicon layers*, Journal of Instrumentation **6** (jun, 2011) P06013.

[20] R. Turchetta, *Spatial resolution of silicon microstrip detectors*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **335** (1993) 44–58.

[21] G. Kramberger, *Signal development in irradiated silicon detectors*. PhD thesis, Ljubljana U., 2001.

[22] H. Spieler, *Semiconductor Detector Systems*. Oxford University Press, 08, 2005, 10.1093/acprof:oso/9780198527848.001.0001.

[23] P. Allport, *Applications of silicon strip and pixel-based particle tracking detectors*, Nature Rev. Phys. **1** (2019) 567–576.

[24] A. F. Prieto and on behalf of the LHCb VELO collaboration, *The LHCb VELO Upgrade II: design and development of the readout electronics*, Journal of Instrumentation **19** (may, 2024) C05011.

[25] M. Moll, *Acceptor removal - Displacement damage effects involving the shallow acceptor doping of p-type silicon devices*, PoS **Vertex2019** (2020) 027.

[26] Watkins, G.D., ed. P. Baruch, *Radiation Damage in Semiconductors*. Dunod, Paris, 1965.

[27] S. M. Mazza et al., *Development of AC-LGADs for Large-Scale High-Precision Time and Position Measurements*, in *2021 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*, pp. 1–7, 2021. DOI.

[28] E. Currás et al., *Inverse Low Gain Avalanche Detectors (iLGADs) for precise tracking and timing applications*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **958** (2020) 162545.

[29] Z. Galloway et al., *Use of "LGAD" ultra-fast silicon detectors for time-resolved low-keV X-ray science*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **923** (2019) 5–7.

[30] N. Cartiglia, *From 3D to 4D tracking: The Impact of Avalanche-based Detectors in Particle Tracking*, 13th International 'Hiroshima' Symposium on the Development and Application of Semiconductor Tracking Detectors (HSTD13), Vancouver, Canada, 2023.

[31] G. Laštovička-Medin et al., *Studies of LGAD performance limitations, Single Event Burnout and Gain Suppression, with Femtosecond-Laser and Ion Beams*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **1041** (2022) 167388.

[32] R. Quay et al., *A Temperature Dependent Model for the Saturation Velocity in Semiconductor Materials*, Materials Science in Semiconductor Processing **3** (03, 2000) 149 – 155.

[33] A. Musiienko et al., *Deciphering the effect of traps on electronic charge transport properties of methylammonium lead tribromide perovskite*, Science Advances **6** (2020) eabb6393, [https://www.science.org/doi/pdf/10.1126/sciadv.abb6393].

[34] J. Mulvey et al., *Preliminary test results of LGADs from Teledyne e2v for the LHC's High-Luminosity upgrade*, Journal of Instrumentation **17** (oct, 2022) C10001.

[35] G. Giacomini, *Fabrication of Silicon Sensors Based on Low-Gain Avalanche Diodes*, Frontiers in Physics **9** (04, 2021) 618621.

[36] M. Carulla et al., *50µm thin Low Gain Avalanche Detectors (LGAD) for timing applications*, Nucl. Instrum. Meth. A **924** (2019) 373–379.

[37] Y. Iwata et al., *Optimal p-stop pattern for the N-side strip isolation of silicon microstrip detectors*, IEEE Transactions on Nuclear Science **45** (1998) 303–309.

[38] L. Evensen et al., *Guard ring design for high voltage operation of silicon detectors*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **337** (1993) 44–52.

[39] I. Mandić et al., *Gain recovery in heavily Irradiated Low Gain Avalanche Detectors by high temperature annealing*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **1055** (2023) 168553.

[40] R. Mulargia et al., *Temperature dependence of the response of ultra fast silicon detectors*, Journal of Instrumentation **11** (dec, 2016) C12013.

[41] A. J. Pointon and H. M. Howarth, *Alternating current theory*, pp. 23–47. Springer Netherlands, Dordrecht, 1991. 10.1007/978-94-011-3142-1_4.

[42] S. Mägdefessel et al., *Understanding the Frequency Dependence of Capacitance Measurements of Irradiated Silicon Detectors*, `2301.09371`.

[43] "particulars, advanced measurement systems: large scanning tct."
`https://www.particulars.si/products.php?prod=LargeScanTCT.html`.

[44] "Particulars Wide Band Current Amplifiers."
`https://www.particulars.si/downloads/ParticularsAmps-Manuals.pdf`, 2017.

[45] E. Currás et al., *Gain reduction mechanism observed in Low Gain Avalanche Diodes*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **1031** (May, 2022) 166530.

[46] E.-L. Gkougkousis, "Description and details of the UCSC single channel board design."
`https://twiki.cern.ch/twiki/bin/view/Main/UcscSingleChannel`, 2017.

[47] K. Onaru et al., *Study of time resolution of low-gain avalanche detectors*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **985** (2021) 164664.

[48] E. G. Villani et al., *Design and development of Low Gain Avalanche Detectors using Teledyne e2v process*, Journal of Instrumentation **19** (jan, 2024) C01038.

[49] J. Mulvey et al., *Timing and gain performance of Teledyne e2v's LGADs before and after irradiation*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment (2024) 169545.

[50] K. G. Svantesson and N. G. Nilsson, *Determination of the temperature dependence of the free carrier and interband absorption in silicon at 1.06μm*, Journal of Physics C: Solid State Physics **12** (sep, 1979) 3837.

[51] M. Zacharias and P. C. Kelires, *Temperature dependence of the optical properties of silicon nanocrystals*, *Physical Review B* **101** (June, 2020) .

[52] E. C. Rivera et al., *Gain layer degradation study after neutron and proton irradiations in Low Gain Avalanche Diodes*, *Journal of Instrumentation* **18** (oct, 2023) 10020.

[53] E. Currás et al., *Timing performance and gain degradation after irradiation with protons and neutrons of Low Gain Avalanche Diodes based on a shallow and broad multiplication layer in a float-zone 35μm and 50μm thick silicon substrate*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **1055** (2023) 168522.

[54] G. Kramberger, *Reasons for high charge collection efficiency of silicon detectors at HL-LHC fluences*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **924** (2019) 192–197.

[55] D. P. Weatherill, *Charge Collection in Silicon Imaging Sensors*. PhD thesis, 2016.

[56] Oxford Instruments Andor, "Low Dark Current Deep-Depletion (LDC-DD) CCD Technology." `https://andor.oxinst.com/learning/view/article/low-dark-current-deep-depletion-%28ldc-dd%29-technology`.

[57] "Vera C. Rubin Observatory." `https://www.lsst.org/`.

[58] H. H. Barrett and W. Swindell, *5 - Detectors*, in *Radiological Imaging* (H. H. Barrett and W. Swindell, eds.), pp. 192–290. Academic Press, San Diego, 1981. DOI.

[59] E. Gumbrell et al., *Characterizing the modulation transfer function for X-ray radiography in high energy density experiments*, *Review of Scientific Instruments* **89** (10, 2018) 10G118.

[60] P. Allisy-Roberts and J. Williams, *Chapter 5 - Digital radiography*, in *Farr's Physics for Medical Imaging (Second Edition)* (P. Allisy-Roberts and J. Williams, eds.), pp. 79–90. W.B. Saunders, second edition ed., 2008. DOI.

[61] I. Kotov et al., *X-ray analysis of fully depleted CCDs with small pixel size*, vol. 9154, p. 91540H, 07, 2014. DOI.

[62] N. W. Ashcroft and N. D. Mermin, *Solid State Physics*. Holt-Saunders, 1976.

[63] D. Meeker, "Finite Element Method Magnetics." `https://www.femm.info/wiki/HomePage`.

[64] D. Caughey and R. Thomas, *Carrier mobilities in silicon empirically related to doping and field*, *Proceedings of the IEEE* **55** (1967) 2192–2193.

[65] P. S. Cheung and C. J. Hearn, *The diffusion of electrons in semiconductors in high electric fields*, Journal of Physics C: Solid State Physics **5** (jul, 1972) 1563.

[66] D. J. Bartelink and G. Persky, *Diffusion of Electrons in Silicon Transverse to a High Electric Field*, Applied Physics Letters **16** (03, 1970) 191–194.

[67] R. Brunetti et al., *Diffusion coefficient of electrons in silicon*, Journal of Applied Physics **52** (12, 1981) 6713 – 6722.

[68] L. Fernandes et al., *Characterization of large area avalanche photodiodes in X-ray and VUV-light detection*, Journal of Instrumentation **2** (08, 2007) P08005.

[69] H. Tsunemi et al., *Measurement of the charge cloud shape produced by an X-ray photon inside the CCD using a mesh experiment*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **421** (1999) 90–98.

[70] Synopsys, Inc., "Synopsys TCAD." `https://www.synopsys.com/manufacturing/tcad.html`.

[71] The Allpix Squared Authors, "Allpix Squared: Semiconductor Detector Monte Carlo Simulation Framework." `https://allpix-squared.docs.cern.ch/`.

[72] S. Spannagel et al., *Allpix2: A modular simulation framework for silicon detectors*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **901** (2018) 164–172.

[73] D. Thompson. Private Communication, July 2024.

[74] "Pr/PrMCTools/src/DepositCreatorDetailed.cpp · run5 · LHCb / Rec · GitLab." `https://gitlab.cern.ch/lhcb/Rec/-/blob/run5/Pr/PrMCTools/src/DepositCreatorDetailed.cpp`, Sept., 2024.

[75] J.-P. Nougier, *Symmetry relations between the mobility and differential mobility-tensor components of cubic semiconductors in the hot carrier range*, Physica **62** (1972) 565–573.

[76] B. Hatanpaa, A. Choi, P. Cheng and A. Minnich, *Warm electron tensor in n-Si from first principles*, in *APS March Meeting Abstracts*, vol. 2022 of *APS Meeting Abstracts*, p. T48.002, Mar., 2022.

[77] E. Wajnryb, P. Szymczak and B. Cichocki, *Brownian dynamics: divergence of mobility tensor*, Physica A: Statistical Mechanics and its Applications **335** (2004) 339–358.

[78] R. C. Curby and D. K. Ferry, *Mobility of hot electrons in n-type inas*, Phys. Rev. B **3** (May, 1971) 3379–3389.

[79] "LHCb / Boole · GitLab." `https://gitlab.cern.ch/lhcb/Boole`, Aug., 2024.

IMPROVEMENTS TO THE SIMULATION OF TE2V CCDS

In the time after the initial submission of this thesis, prior to the viva, work was undertaken to improve upon the problems identified in Chapter 5. In particular, a possible reason behind the application of the scaling factor of 40 for the lateral diffusion during drift was investigated. Additionally, a more accurate simulation of the dependence on X-ray absorption depth was implemented.

A.1 Diffusion Scale Factor

In the original simulation, an additional scale factor, α, was added to the definition of the diffusion coefficient in Equation 5.6, with a value of $\alpha \approx 40$ being identified as necessary in order for the simulation results to match the experimental results, as shown in Figure 5.18 and Figure 5.22. This arbitrary scale factor can be removed through two key changes to the simulation.

The first change is the application of diffusion during each time step, initially shown in Equation 5.10. This diffusion step is different to the step used by Allpix2 where instead the change in each axis is a random number drawn from a Gaussian [71, 72]. This Gaussian has a standard deviation given by,

$$\sigma = \sqrt{2D\delta t},\eqno(\text{A.1})$$

which is more commonly used to predict diffusion over a longer timescale. By

comparing this with the RMS of $\sqrt{D\delta t} \cdot \sin(\theta)\sin(\phi)$ from Equation 5.10, a factor of eight difference is revealed. This means a scale factor of only $\alpha \approx 5$ is required after this improvement.

The second change is in the calculation of the diffusion coefficient and mobility. In Equation 5.6, the diffusion coefficient is given as proportional to the field-dependent mobility, $\mu(\mathbf{E})$. Discussion with colleagues within LHCb led to the suggestion to use the low-field mobility for lateral diffusion [73, 74]. Generally speaking, treating mobility as an anisotropic quantity is not new and intuitively, one might be able to justify its use [75, 76, 77, 78]. If the electric field is largely perpendicular to the electrodes, then the magnitude of the electric field component perpendicular to this is, by definition, zero. In the context of charge sharing, lateral diffusion is the dominant property. This change alone is not sufficient to account for all of the scale factor of $\alpha \approx 5$. For this, the mobility model also needs updating. The mobility model offered by [55] does not match the various models provided by Allpix2 [71, 72]. Therefore, the Jacoboni-Canali model was adopted instead. A comparison between the two models at two different temperatures (250 K and 150 K) is shown in Figure A.1. At 150 K (the operating temperature for LSST), the high- and low-field Jacoboni-Canali mobility differ by a factor of approximately five, eliminating the need for the scale factor α. A justification for the Jacoboni-Canali model comes from the scale factor used by LHCb of 2.25, where the Jacoboni-Canali mobility at 250 K (operating temperature of LHCb's tracker) predicts a factor of 2.35 [79].
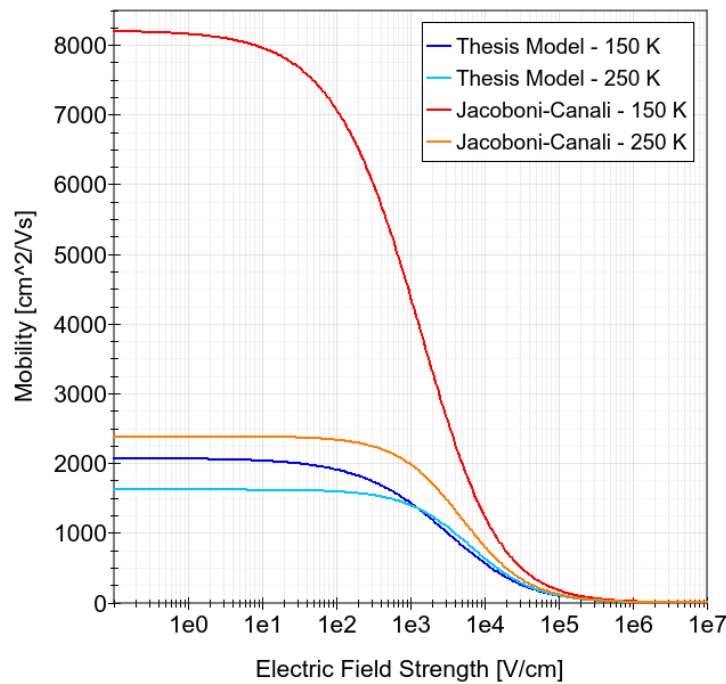

Figure A.1: The Jacoboni-Canali model is compared with the mobility model used initially in this work. They are compared at two temperatures: 150 K and 250 K.
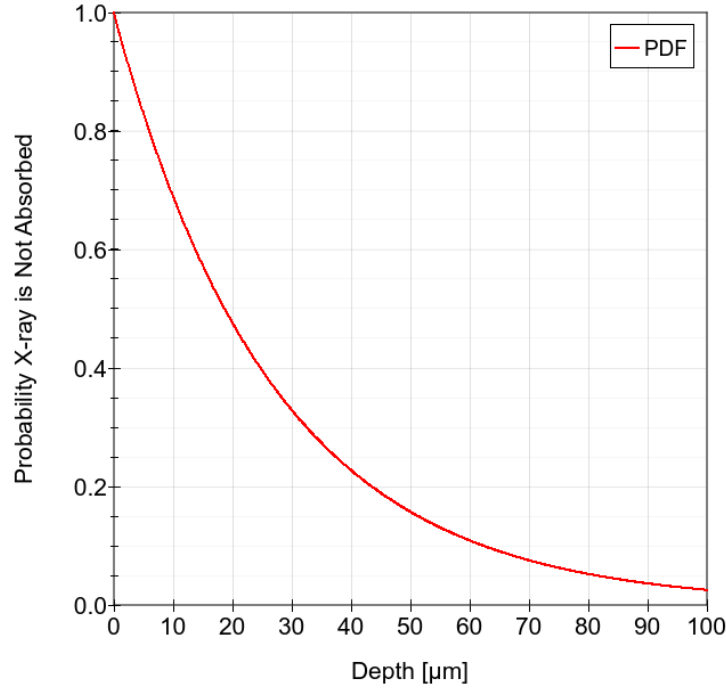
Figure A.2: The probability density function for X-ray absorption. In silicon, the absorption length is $28.8\,\mu$m for $5.9\,$keV X-rays [61].

A.2   X-ray Absorption Depth

A second major problem with the original simulation was that X-rays were assumed to absorb at a fixed depth. In reality, they can be absorbed anywhere within the sensor, with a heavy weighting towards the side of the sensor from which they enter. This dependency on depth is described by Equation 2.11 and the probability density function is shown in Figure A.2.

In order to accommodate different numbers of X-rays absorbed at different depths, the simulation now runs slightly differently. Firstly, the test grid shown in Figure 5.20 is still simulated and analysed, but at a range of depths from $0\,\mu$m (the electrodes) to $100\,\mu$m (where the X-rays enter). In order to correctly sample from these heatmaps, the simulation first chooses one position from the 10×10 grid. It then randomly selects a depth from the probability density function in Figure A.2. For a particular position in the 10×10 grid, the calculated PSF, σ, can be shown as a function of depth such as in Figure A.3. For a given depth, these curves can be interpolated to find a value of σ. This process is repeat many times to produce a histogram similar to the original simulation.

A comparison between the old simulation, the new simulation, and the experimental data from Figure 5.18 is shown in Figure A.4. This clearly shows that the new simulation's peak is only $0.4\,\mu$m from the experimental peak, but crucially there is
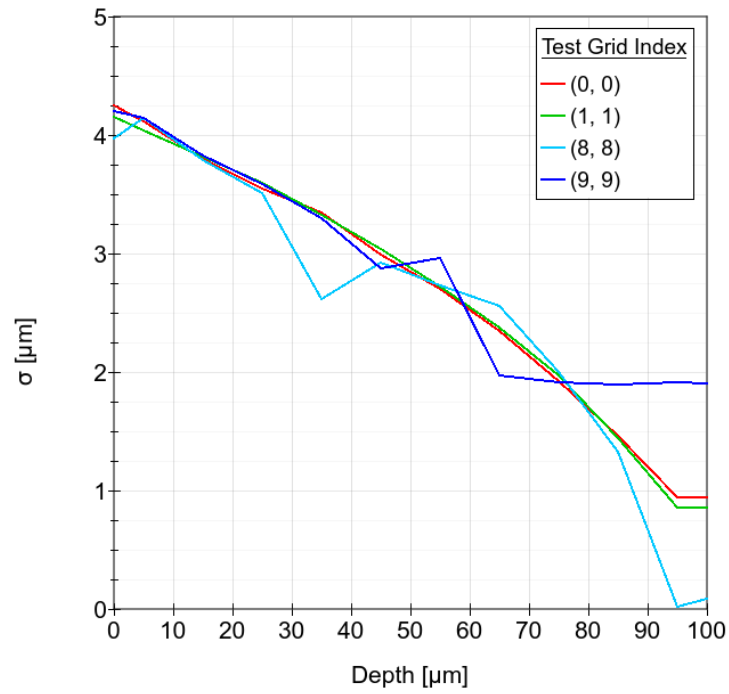
Figure A.3: For a sample of positions in the 10×10 test grid, the calculated PSF, σ, is shown as a function of absorption depth. The plateau region of the dark blue curve represents the inability to calculate σ in the corners of a pixel, as discussed in Section 5.4.2

no additional scaling factor for the diffusion coefficient ($\alpha = 1$). Figure A.4b shows the simulation histograms aligned with the peak of the experimental histogram. In this case, it is clear that the inclusion of the absorption depth distribution allows the new simulation to very closely match the shape of the experimental data. Of course, the new simulation does not perfectly match experimentation. However, further investigation would require much further fine-tuning of the simulation.
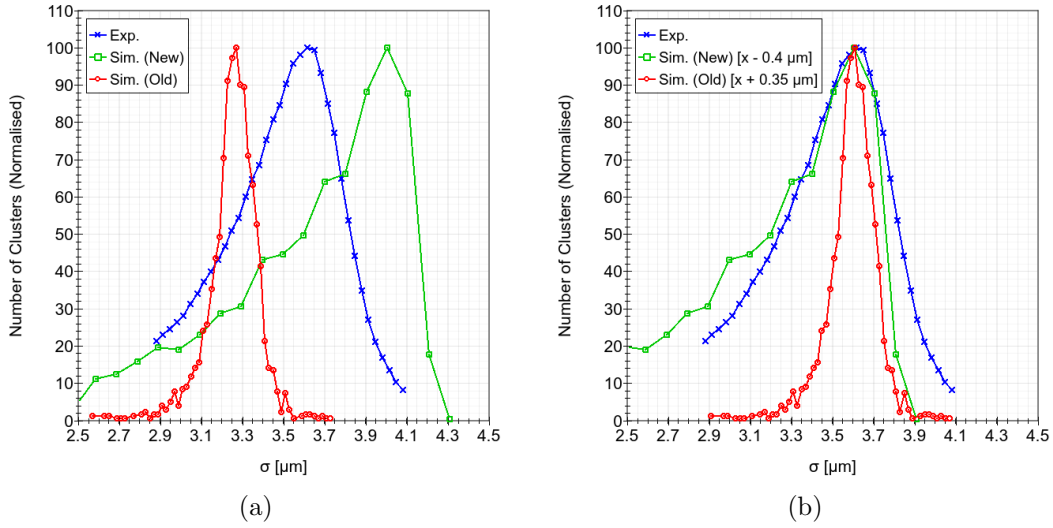

Figure A.4: (a) The new simulated cluster histogram is compared to the older simulation (including the diffusion scale factor of 40) and the experimental data. (b) Shows the simulation peaks aligned to the experimental peak, emphasising the shape of each histogram.

A.3   Conclusion

This supplementary study has successfully eliminated the need for an additional scale factor for diffusion as well as implementing a distribution of X-ray absorption depths. This has resulted in simulation data which is in much better agreement with experimental data than in the previous iteration of the simulation. While more work is required to fine-tune the simulation, confirming the validity of using the low-field mobility for lateral diffusion while using the much lower field-dependent mobility in the drift direction is a clear priority for future studies.