Hynds, Daniel Peter McFarlane (2014) *Resolution studies and performance evaluation of the LHCb VELO upgrade*. PhD thesis.

University of Glasgow
School of Physics and Astronomy

Resolution Studies and Performance Evaluation of the LHCb VELO Upgrade

Submitted in fulfilment of the requirements
for the degree of Doctor of Philosophy by
Daniel Peter McFarlane Hynds

Glasgow, September 2014

Abstract

The LHCb detector at CERN is scheduled to undergo an upgrade during the second long shutdown of the LHC. As part of this upgrade, the vertex detector (VELO) will be replaced with a new hybrid pixel detector, based on an evolution of the Timepix ASIC. The performance of this detector should improve upon that achieved by the current VELO, in addition to facilitating the complete detector readout at 40 MHz.

As part of the preparation for this upgrade, this thesis presents the results of studies carried out on the single hit resolution of silicon hybrid pixel detectors. The development of a particle beam telescope has been carried out to allow these studies, shown to operate with track rates in excess of 45 kHz and with a pointing resolution at the device under test of less than 2 µm. A wide range of sensor types, thicknesses and resistivities have then been tested under different operating conditions and the results presented, with single hit resolutions varying between 4 µm and 12 µm depending on the conditions and incident angle. The resistivity of the devices is observed to have a significant effect on the single hit resolution, with high resistivity substrates allowing operation at lower bias voltages. This facilitates increased charge sharing, and the corresponding improvement in resolution. At sufficiently large incident angles however, the resolution becomes independent of the electric field, being instead dominated by the sensor geometry and variations in the charge deposited along the track length. No significant differences were found between the various detector technologies (n-on-n, n-on-p and p-on-n) though a difference in performance is expected for low-voltage operation of higher resistivity samples. A simplified model of the physical processes contributing to the detector resolution has been constructed, shown to reasonably reproduce the observed resolution as a function of angle and bias voltage. This model is extrapolated to potential future directions in the design of pixel sensors, highlighting the differences between various technology choices.

The integration of the ATLAS FE-I4 ASIC into the telescope has been carried out, and the performance of an unirradiated planar silicon sensor was shown in order to verify this. Efficiency measurements show that the device is fully efficient in the angular range measured. The tracking performance of two irradiated sensors mounted on FE-I4 ASICs has been investigated, in addition to the mapping of collected charge over the pixel unit cell under various biasing conditions and at varying incident angles with respect to the incoming particles. For the sample irradiated to 2×10^{15} 1 MeV $\mathrm{n_{eq}}$ cm^{-2} the single hit resolution was 12.5 µm at perpendicular incidence, dropping to 8 µm at 22°. The sample irradiated to

4×10^{15} 1 MeV n_{eq} cm^{-2} was found to have a resolution of around 13.5 μm, which remained relatively insensitive to the incident track angle. The conclusions drawn suggest that the upgraded VELO detector will be able to overcome the difficult radiation environment if it is able to reach the high voltage operation required.

The implementation of these observations in the LHCb simulation environment has allowed some initial studies on the likely degradation of the detector performance to take place, showing that the high tracking efficiency (99.4 % for Long tracks) is likely to be maintained throughout the full lifetime of the upgrade. The impact parameter resolution was not observed to vary significantly. These studies have been carried out alongside simulations to gauge the expected compression that can be achieved in the data transmission of the VELOPix ASIC. Different designs of the front-end have been implemented, leading to the adoption of binary readout for the upgraded VELO. The uniformity of the pixel pitch across the detector has additionally been used to show the sensitivity of the system to multiple scattering, shown to be a credible tool with which to control the event reconstruction in the online LHCb trigger. A reduction of the number of VELO tracks passed to the forward reconstruction of almost 50 % has been shown, for the loss of only 5 % of tracks with momentum above 20 GeV/c. This could potentially replace the lifetime biasing cuts currently envisaged in the trigger.

Acknowledgements

There are many people who have made this work possible.

For my professional thanks I must first turn to Paula Collins, for her advice, instruction and continual request for just one more plot. Without her I believe that not only my own work but the progress of the VELO upgrade itself would be significantly lessened. Her drive and enthusiasm, along with her determination never to abandon hope, have been a strong positive influence in my years working with her. I must next turn to my supervisors, Chris Parkes, Lars Eklund and Paul Soler. All have performed admirably both in their understanding and patience, and have always been there to remind me of the painful unpleasant tasks that are nonetheless necessary at times (this thesis first and foremost). Their guidance has been greatly appreciated.

Further afield I would like to thank the various members of the VELO upgrade project, in particular the long-suffering Heinrich Schindler, with whom many a midnight hour passed under the watchful glare of the North Area lights, and many a kebab passed in the few blissful moments where everything worked as it should. To Martin van Beuzekom, Jan Buytaert and Raphael Dumps for their seemingly endless energy and boundless patience I express my most heartfelt thanks. To Matthew Reid and his apparently infinite knowledge of advanced programming tools (and for chef duties on behalf of the testbeam campaign), I am very thankful. Similarly to my ATLAS colleagues, in particular Jens Dopke and Christian Gallrapp, and their continual toil to follow the science regardless of its direction, I extend my warmest thanks.

And finally, to all of those in the LHCb collaboration which whom I have and the pleasure to work, both within the VELO upgrade group and beyond, you have my sincere gratitude. The collaboration has been a welcoming environment, understanding of new students and their relative inexperience, and it has been a profound pleasure to have taken part in this most wondrous achievement. This same is true of the PPE group in Glasgow, without whom I would not have had this opportunity. To the current PhD students: you are almost there. To the AIDA project, for their support and facility of testbeams, I will both always be thankful and never forgive them.

On a personal level this work could not have been carried out without the support of my friends and family, whom I hope can recognise who they are. To my parents, in whose debt I doubt I will ever escape, I thank them for the difficulties that they have put up with. To Paula I thank for her support and her smile.

There are many people who have made this work possible, and while it would be impossible to thank all of them there are at least a few whom deserve more than the words on these pages. To you, I hope that I have made my appreciation and gratification known outside of these brief words of thanks. I couldn't have done this without you.

Graciñas,

Daniel

March 2015

Declaration

The research results presented in this thesis are the product of my own work. Appropriate references are provided when results of third parties are mentioned. The research presented here was not submitted for another degree in any other department or university.

D. Hynds

Preface

This thesis presents the results of studies into the expected performance of the upgraded VELO detector, encompassing an extensive testbeam campaign to investigate the single hit resolution of hybrid pixel detectors as a function of the sensor design, the tracking performance after irradiation and the overall implications for the upgraded detector. This involved construction of a high rate beam telescope, the analysis of the detectors placed in the beam and simulations of the VELO upgrade both before and after irradiation.

An overview of the CERN accelerator complex and the LHCb experiment are presented in chapter 1, detailing the performance of the detector for a few key parameters important to the analyses taking place. Details of the plans to upgrade the detector at the end of the second LHC run are provided, with an emphasis on the upgrade of the vertex detector (VELO).

Chapter 2 presents an overview of the theoretical background of silicon detectors, describing the principles behind silicon as a radiation detector. The basic operating principles are summarised, along with the mechanisms for charge generation by energetic particles and the effects of radiation damage in silicon detectors.

In chapter 3 the LHCb VELO Timepix telescope is introduced and described in detail. Reconstructable track rates of 45.8 kHz are achieved, with a pointing resolution at the device under test of less than 2 μm. The timestamping efficiency is shown to be high, limited by the intrinsic Poisson timing characteristics of the particle beam. The work in this section involves collaboration amongst various individuals involved in the VELO upgrade project, and both the assembly and software were achieved with the help of many individuals. For my own contributions, I was the project co-ordinator who carried out the majority of the data taking along with H. Schindler from 2010-2013, assisted at times by members of the VELO upgrade team, which involved the installation and commissioning of a new fast readout system. I wrote the first version of the monitoring and control software, the slow control interface (with A. Leflat), parts of the analysis software (including a progressive fit-and-extrapolate pattern recognition with M. Reid) and the integration of external devices (with M. John). The evaluation of the pointing precision of the telescope was carried out by P. Collins and H. Schindler.

Chapter 4 contains the analysis of many different hybrid pixel assembles, all mounted on Timepix ASICs, and with a range of sensor designs including different thicknesses, substrate resistivity and bulk/implant types. The single hit resolution of all of the devices is measured for a wide range of incident track angles and bias voltages, and the trends between different designs determined. A simple model of the device resolution is constructed, after identifying the likely sources for the features found in the data. The alignment and application of eta corrections to the data was performed jointly between several members of the VELO upgrade group (C. Hombach, P. Rodriguez, Y. Ling, P. Tsopelas and myself), while the analysis of the results was carried out by the author.

In chapter 5 the integration of the ATLAS FE-I4 ASIC with the VELO Timepix telescope is described. This was carried out by the author, with the operation of the FE-I4 by the ATLAS IBL PPS group (in particular J. Dopke, C. Gallrapp, J. Jentzsch, T. Heim and M. Kocian). The performance of an unirradiated sensor mounted on an FE-I4 is shown, with measurements of the single hit efficiency and resolution. Two irradiated sensors mounted on FE-I4 ASICs are then presented, prepared by the MPI group in Munich (organised by A. Macchiolo). The charge collected across the pixel cell is shown, and the single hit resolution is observed to degrade significantly in comparison with unirradiated sensors. The active depth is observed to decrease, with the more heavily irradiated sample showing little variation with incident track angle. These analyses were carried out by the author.

Chapter 6 details the simulations which have been carried out with the full upgraded VELO detector. This builds upon the work of many people in the LHCb and VELO upgrade communities, including the material description of the detector and the initial description of the chip. The author has updated the description of the ASIC in the simulation environment and written the clustering algorithm used in all analyses presented in this section. A study was carried out by the author into the IP resolution observed for tracks using a linear fit in the VELO before the track momentum is known has been carried out, showing that a reweighting of the hit uncertainties can be used to improve the performance at this point. This study further shows that it is possible to selectively remove low-momentum tracks by observing differences in the track fit when the outer hits are given an increased weight, of use in the upgraded trigger where a decay-time biasing cut was envisaged. A literature review of the expected charge collected from sensors of different thicknesses as a function of radiation damage and applied bias voltage was carried out by the author with A. Morton, and the results implemented in the simulation framework. The expected performance of the upgraded VELO detector was then investigated by the author, showing no significant loss in performance over its lifetime.

Only two things are infinite,
the universe and human
stupidity, and I'm not sure
about the former.

If we knew what it was we
were doing, it would not be
called research, would it?

A. Einstein

Contents

1
Introduction

1.1 Physics at the LHC

1.1.1 The CERN Accelerator Complex

The European Organisation for Nuclear Research (CERN) [1] has been the home of European, and at many times worldwide particle physics since its establishment in 1954. Many important discoveries, including that of the W and Z bosons [2], and the establishment of the number of generations in the standard model, have been achieved through the accelerator infrastructure provided, continuing to the present day with the announcement of the discovery of a state compatible with the Standard Model Higgs boson [3].

Developed over the many years that CERN has existed, a long chain of accelerators now exist across the two main CERN sites. What were previously high energy colliders in their own right are now used in the injection chain to the current world leading machine, the Large Hadron Collider (LHC) [5]. The LHC is a 27 km long circular accelerator, with power provided to the beam by bunched Radio-Frequency (RF) cavities, and bending power via the more than 1200 superconducting dipole magnets. Primarily a proton-proton collider, the LHC is designed to deliver the highest ever man-made collision energy of 14 TeV.

The current accelerating structure at CERN is shown in figure 1.1. For the LHC chain, protons are produced by stripping the outer electrons from hydrogen molecules, supplied by a bottle at the start of the accelerator chain. These are then accelerated by a linear accelerator (LINAC 2) to 50 MeV, before insertion into a booster ring for acceleration to 1.4 GeV. This boost in energy allows a significant increase in the number of protons that can be accepted by the Proton Synchrotron (PS), which is the first large-scale accelerator on the path to the LHC. Once reaching 25 GeV in the PS, the protons are then transferred to the Super Proton Synchrotron (SPS), which accelerates them to up to 450 GeV and injects them into the LHC for their final boost to collision energies. The four collision points highlighted in figure 1.1 are the locations of the four largest experiments at the LHC: ATLAS, CMS, LHCb and ALICE [6–9].

An important installation that depends on this chain is the CERN North Area [10], containing several beam lines which are fed by the SPS. Directly in front of the experimental hall sit two target areas, T2 and T4, which can be used to convert the proton beam into a range of secondary beam types, or to simply attenuate the SPS beam intensity. The targets consist of multiple sheets of beryllium which produces pions and electrons on interaction with the incoming protons, and are followed by a

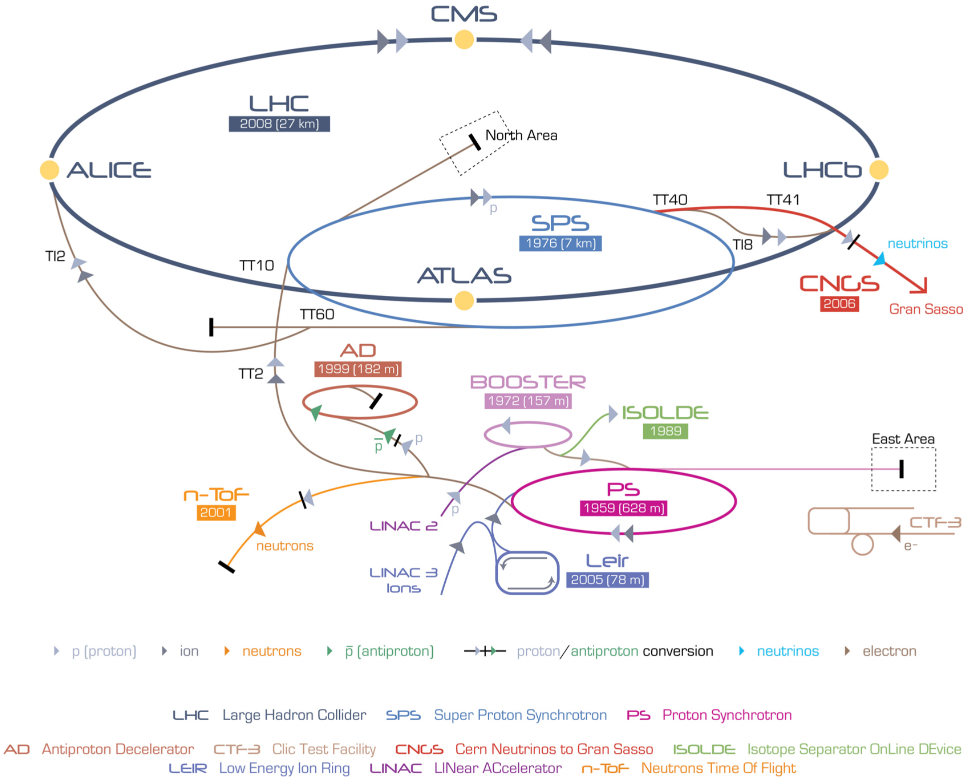

Figure 1.1: The CERN accelerator complex, figure taken from [4].

spectrometer that allows the selection of the secondary beam momentum. By increasing the amount of material in the target, or by including an optional filter after the start of the spectrometer, the electron contribution to the beam composition can be significantly decreased. This is advantageous as scattering within the experimental setup degrades the minimum possible resolution, and this is much larger for lighter particles.

1.1.2 The Standard Model

The Standard Model (SM) provides a description of the fundamental constituents of matter and the forces which act upon them, with the notable exception of gravity where the development of a Grand Unified Theory (GUT) has been a goal of the theoretical community for some time. While SM predictions of the behaviour of particles have been extremely accurate and self-consistent, there are noticeable areas of particle physics and cosmology that are not explicable within this framework. Experimental observations which point to the existence of dark energy and matter, or the non-zero mass of the various neutrinos are not described within the the SM, as is the imbalance of matter and antimatter in the visible universe. The different behaviour of matter and antimatter predicted by the SM cannot account for the evident asymmetry in annihilation which took place during the early formation of the universe. Answers to all of these questions are searched for across a wide range of fields, with particle physicists focussing on what model may lie beyond the SM.

The particle content of the SM is shown in figure 1.2 and contains two types of

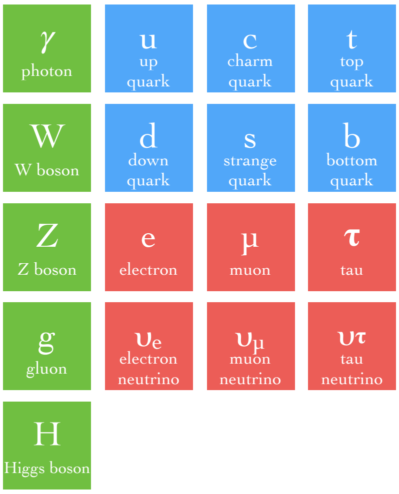

Figure 1.2: The particle content of the standard model, showing (green) the gauge bosons and the fermions, split into (blue) quarks and (red) leptons.

particles: fermions and gauge bosons. The latter particles consist of the force carriers within the SM, and include the photon (responsible for the electromagnetic interaction), the gluon (for the strong interaction) and the W and Z bosons (for the weak interaction). The Higgs boson, which arises in electroweak symmetry breaking, gives mass to particles through its interactions. The fermions are split into two categories and three generations. Each of the fermions interacts via the electromagnetic and weak forces, with the *leptons* and *quarks* differentiated by the strong interaction of the latter. Direct conversion between leptons and quarks is not possible (both lepton and quark number are conserved quantities in any interaction), and transitions within either group are mediated by the weak force.

The probability for any given interaction is related to the number of transitions required to take place. The processes are typically represented by Feynman diagrams [11], where the classification of the diagram is a general indication of its likelihood of taking place. Three such diagram types are shown in figure 1.3: those of tree, penguin and box processes. Tree-level diagrams can be used in direct searches for processes which are not described by the standard model, but are sensitive only to particles in the mass range within direct reach of the collider. Searches far beyond the interacting particle energies are best achieved by *indirect* searches. Typical indirect searches make use of processes which occur to first order through loop transitions (penguin or box diagrams). The observables associated to such processes, such as the branching ratio, are typically low and can be altered by the incorporation of any as-yet undiscovered particle in the loop process. Precise measurements, along with calculations from the SM, can be used to probe energy scales significantly higher than the particle energy, and detect physics beyond the SM.

Another search method for processes not described in the SM is by comparing the behaviour of decays where intermediate states are possible. Several neutral composite particles can *oscillate* into their anti-matter equivalent through box-diagram processes such as that shown for the B^0 meson in figure 1.4. Where both the B^0 and $\overline{B}^0$ can follow

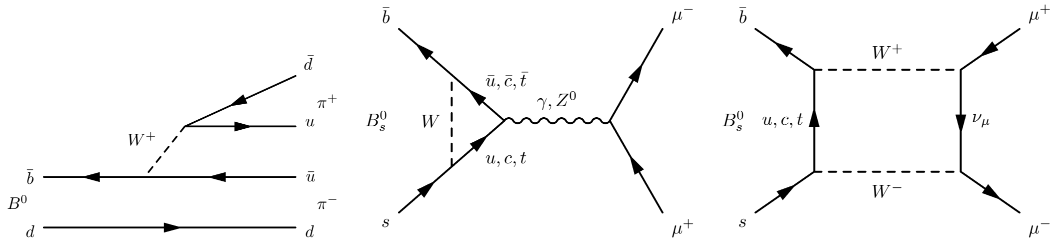

Figure 1.3: Examples of a (left) tree-, (centre) penguin- and (right) box-level decay processes. Figures taken from [12].

the same decay path the process may feature an intermediate step where the B^0 first oscillates into a $\overline{B}^0$ or vice versa. This process is not necessarily symmetric, as the SM describes the flavour and mass states as related by a rotation matrix (the CKM matrix), which contains complex phases allowing the interactions to proceed at different rates. Any unknown physics processes may increase or decrease these differences, making the measurement of asymmetries between the B^0 and $\overline{B}^0$ decays sensitive to physics beyond the standard model. This difference in behaviour is referred to as *Charge-Parity (CP) violation.*

1.1.3 Flavour Physics and LHCb

The LHC Beauty experiment (LHCb) is designed to study such heavy-flavour physics at the LHC [13], providing a complimentary range of measurements to those of the General Purpose Detectors (GPDs) in the search for physics beyond the SM. Specifically, LHCb studies b-physics and performs indirect searches for new particles or interactions by measuring the branching fraction of suppressed decays within the SM, and through the measurement of CP violation in a wide range of decays. This accounts for a large number of the analyses taking place on LHCb, but the much more varied physics programme includes searches for new resonance structures predicted by the quark model, searches for exotic particles and physics in the forward direction.

The physics goals of the experiment are the primary driving factor behind the detector design, giving LHCb a very focussed design remit. B-quarks are significantly heavier than their typical final decay products and have sufficiently long lifetimes as to travel from the interaction point before decay. This allows a trigger scheme based on

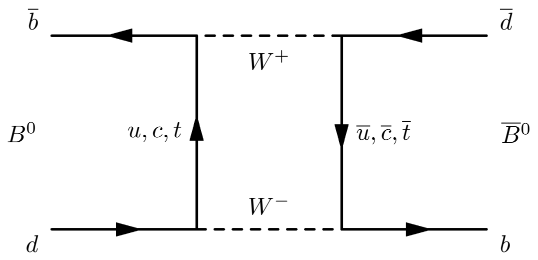

Figure 1.4: Mixing of the B^0 meson into its antimatter equivalent, $\overline{B}^0$. Figure taken from [12].

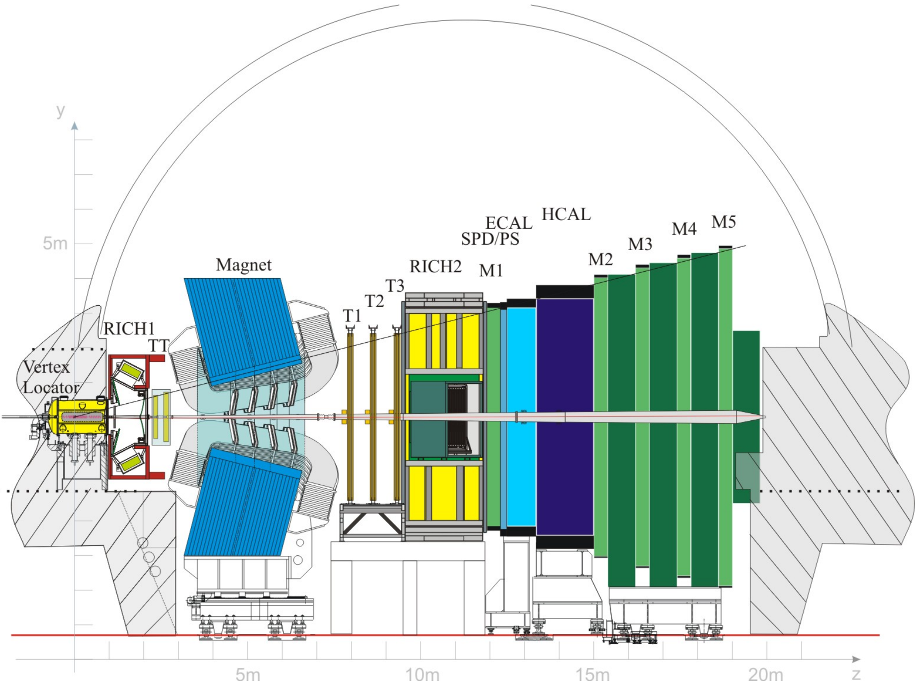

Figure 1.5: The LHCb Detector, showing the VErtex LOcator (VELO), the 2 Ring-Imaging Cherenkov Detectors (RICH1 and RICH2), the Tracker Turicensis (TT), the dipole magnet, the downstream tracking stations (T1, T2 and T3), the muon detectors (M1 to M5) and the calorimetry system (SPD, PS, ECAL and HCAL). Figure taken from [6].

large energy deposits in the calorimeters, combined with information about the flight distance of the particles produced. An extremely precise vertex detector is therefore required in order to resolve displaced vertices and identify long-lived particles. This allows not only the efficient separation of signal events from the large background, but provides the lifetime measurements for these particles to be studied, including resolving the oscillations which occur for neutral b-mesons such as the B^0 and B_s^0. The decay products of many of these states include both leptons (where muons are common final-state particles) and hadronic matter, where the identification of kaons and pions is important. For this reason LHCb was required to provide excellent Particle ID (PID), with strong separation for these two particles in particular. A series of muon chambers serves the identification of muons, which are further used in the first (hardware) level of triggering in order to maximise the yield of such events. In addition, a mass resolution sufficient to resolve neighbouring resonant structures was required, giving the layout of the detector shown in the next section.

1.2 The LHCb Detector

The LHCb detector is designed as a forward arm spectrometer, with a horizontal (vertical) angular coverage of 10 to 300 (250) mrad with respect to the incoming proton beams. This takes advantage of the $b\bar{b}$ pair production at the LHC, which is highly peaked in the forward and backwards directions. The detector covers roughly 2 % of the solid angle yet this acceptance covers $\sim$ 27 % of b quarks produced [14]. The final

LHCb layout, as installed in the LHC, is shown in figure 1.5. As described previously, the detectors can be separated into roughly two functional groups: those that provide tracking information and those that identify particle type.

1.2.1 Tracking Detectors

The tracking stations of LHCb consist of the VErtex LOcator (VELO) [15], the Silicon Tracker (ST) [16] comprising the Tracker Turicensis (TT) and Inner Tracker (IT), the Outer Tracker (OT) [17], and a single dipole magnet located after the TT [18]. These subsystems combine to allow the measurement of decay lengths of long lived particles, momentum information for charged final states, and reconstruction of the full decay by spatially matching the information from all subdetectors.

The VELO

The vertex detector for LHCb is designed to reconstruct the production and decay vertices of b- and c-quark hadrons, in order to accurately measure their decay times, their impact parameter (IP - the distance of closest approach) with respect to the primary vertex, and to provide tracking information upstream of the dipole magnet. For this reason, the detector is required to sit close to the interaction point, and to have high spatial resolution and minimal material.

To achieve such requirements, a silicon strip detector was constructed with R-ϕ geometry, consisting of planes of silicon arranged perpendicular to the beam axis. The design requirements for angular coverage and the minimum number of station hits per track (4) led to the optimisation of the z-layout shown in figure 1.6. In order to make the first measurement point as close as possible to the interaction point and to minimise the material content, the VELO was designed without a beam pipe to contain the LHC beam, and instead is separated from the LHC vacuum by a corrugated 300 um Radio-Frequency (RF) foil, which additionally shields the modules from beam-induced RF signals. This allows the innermost sensitive region to sit at a radius of 8.1 mm from the collision point. To provide sufficient aperture for the injection to the LHC, the detector was designed in two halves mounted on opposing sides of the beam upon a

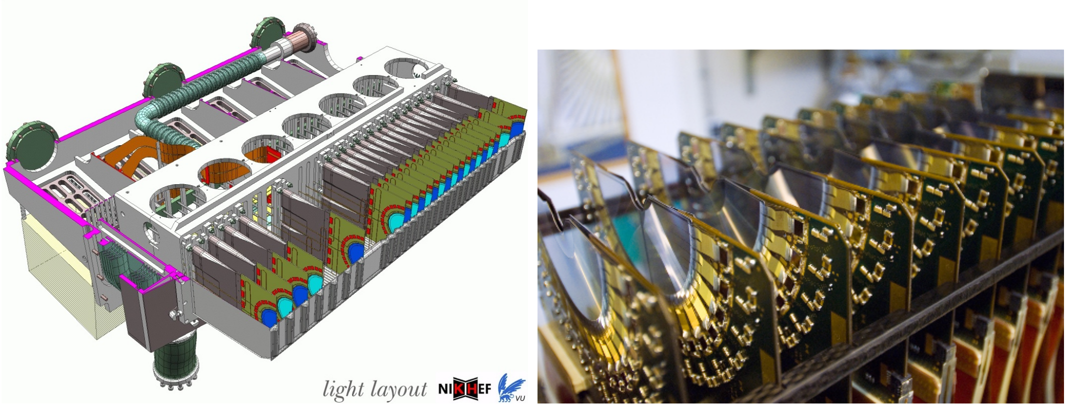

Figure 1.6: Left: Schematic of one full half of the VELO detector. Right: Close-up of one half of the VELO, focussing on the sensors.

precise mechanical stage such that both halves can be retracted by around 30 mm. This enables the VELO to remain open during beam injection, and to close symmetrically around the beam once stable beams have been created.

The composition of a single VELO module can be seen on the right hand side of figure 1.6 and consists of 2 silicon strip sensors mounted on opposing faces of the PCB containing the active components for the sensor readout. The sensors are produced with strip implants in either a radial or circular pattern, with each module containing one sensor of each design. This allows a single measurement point in R-ϕ space by each module. The readout chips can be seen around the sensor edge, mounted on the PCB and attached via a pitch adaptor to the metal readout lines patterned on the sensor. Thermal connectors are located at the top of the carbon fibre support, just before the PCB begins, and acts as heat sink which cools the PCB. The cooling system itself consists of a novel bi-phase CO_2 system, where the heat transfer proceeds by evaporative cooling.

The TT

The TT is the final subdetector before the dipole magnet, and as such is used to link VELO tracks with those produced in the tracking stations downstream of the magnet. The detector consists of 4 planes of silicon strip sensors, oriented at angles of 0°, +5°, -5° and 0° with respect to the vertical axis. The silicon sensors have a pitch of 183 µm and consist of p-doped implants on a 500 µm thick, n-type substrate.

The IT and OT

Downstream of the magnet there are 3 tracking stations, T1, T2 and T3, which contain both the Inner Tracker (IT) and Outer Tracker (OT). The IT sits closest to the beam pipe and each station contains four layers of silicon strip detectors, oriented at angles of 0°, +5°, -5° and 0° with respect to the vertical axis, as for the TT. Two sensor thicknesses are used: 320 µm for the sensors above and below the beam, and 410 µm for those to the left and right, all of which contain strips with a constant pitch of 198 µm. The bulk and implant types are the same as for the TT. As the particle flux falls off considerably with increasing radius from the beam, silicon was chosen for the inner region to provide greater granularity and improved hit resolution. For the total area of the downstream tracking however, the cost to cover the full acceptance with silicon (some 30 m^2) was prohibitively expensive, and a series of tracking planes consisting of straw tubes was constructed. As with the IT and TT, the straw tube planes of the OT are arranged in sets of four per station, with angles of +/- 5° on the central planes. The straw tubes are made up of 4.9 mm diameter cylindrical gas tubes, with a 25.4 µm thick gold-plated tungsten wire running through the centre. A gas mixture of 70 % Ar and 30 % CO_2 was chosen such that the ion drift time remains low (below 50 ns), but high enough to sufficiently resolve the drift time co-ordinate (roughly 200 µm).

The Dipole Magnet

LHCb contains a single dipole magnet which is used to provide a bending field for the measurement of charged particle momenta. Due to the cost and timescale of the original superconducting magnet proposed, a room temperature dipole was constructed

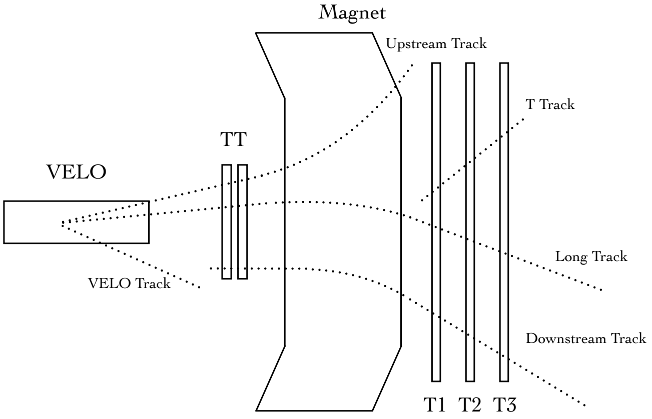

Figure 1.7: Cartoon illustrating the different track classifications within LHCb.

providing the required integrated field of 4 Tm. Other constraints that shaped the magnet design included the desire to have as high a field as possible between the VELO and TT, and to limit the residual field inside the RICH enclosures to less then 2 mT. Weighing 1600 tons and operating with a current of 5.85 kA, the field throughout the entire active volume has been mapped to high precision using hall probes, in order to achieve the required momentum resolution. Additionally, in order to investigate any observed asymmetries which could be due to the detector effects, the magnet polarity can be switched. Data taking during 2011 and 2012 was split approximately evenly between magnet orientations.

Tracking Performance Results

Tracks reconstructed within LHCb are split into different categories, depending on their geometry and which detectors are involved in their reconstruction (schematic in figure 1.7). Of primary importance are so-called *long* tracks, those which traverse all tracking detectors. Such tracks are of primary importance as they allow measurements of the impact parameter, decay time and momentum. This is more than just a useful nomenclature, as the resolution of the detector is different for each type of track. For instance, the lack of momentum information for VELO tracks degrades the precision of the track fit within the VELO, as the momentum measurement is used to account for material scattering throughout the detector.

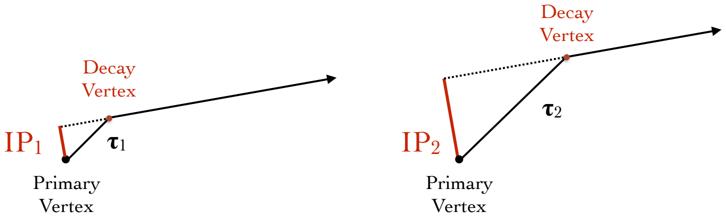

Figure 1.8: 1-dimensional measurement of the Impact Parameter (IP), and the dependence of track IP on particle decay time (τ).

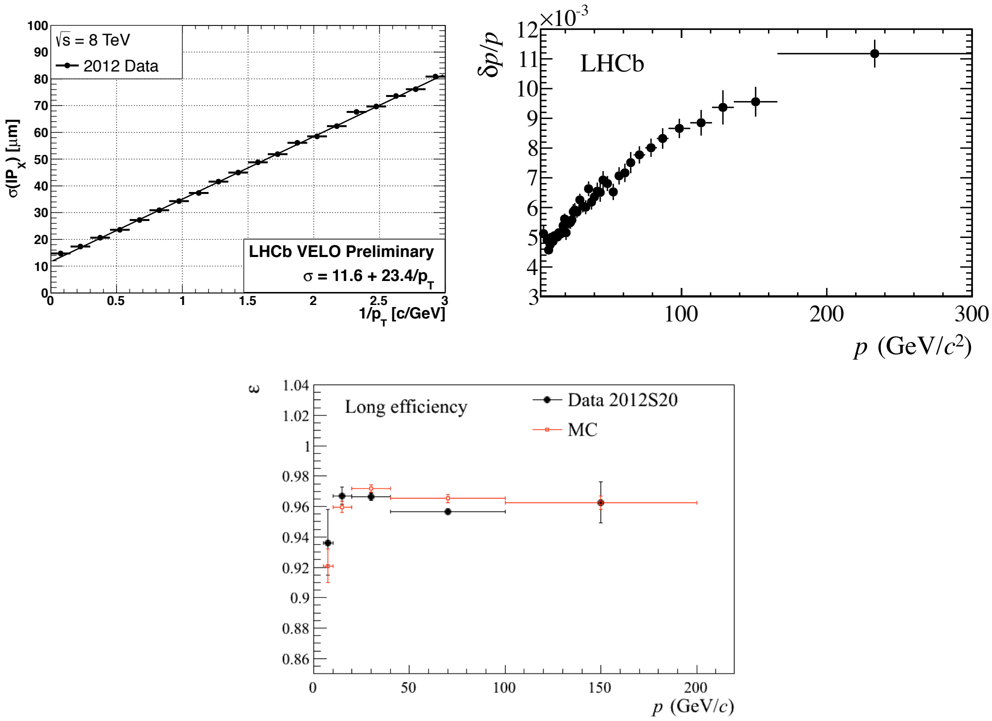

Figure 1.9: Tracking system performance plots. Left: Long tracking efficiency versus particle momentum. Centre: Impact Parameter (IP) resolution in the x-direction for Long tracks. Right: Momentum resolution versus momentum. Figures taken from [19].

The main performance criteria for the tracking systems are the track reconstruction efficiency, the momentum resolution and the track resolution parameters (consisting of the related single hit, vertex, and impact parameter (IP) resolutions). Of particular importance for the detector and the operation of the current trigger scheme is the IP resolution. The impact parameter for a given track is the distance of closest approach between the track and the primary (collision) vertex. Primary vertices are first reconstructed with a precision of 13 μm in the transverse plane and 71 μm in the beam direction. Tracks originating from the primary proton-proton interaction should have an IP compatible with zero, while large impact parameters are expected for long lived particles which decay within the detector. The average distance traversed by a b-quark at the LHC before decay, for instance, is of order 1 cm. A schematic of IP measurement, illustrating the dependence on particle decay time, is shown in figure 1.8. The bias introduced by using this parameter in the trigger will be discussed in later chapters.

Performance plots for the current detector can be seen in figure 1.9. The track reconstruction efficiency (long tracks) is around 96 %, and the momentum resolution is better than 1 % across the full momentum range. The IP measurement is extremely precise, and the accuracy in decay time measurements that the VELO provides allows LHCb to resolve oscillations between particles and their anti-matter equivalents.

1.2.2 Particle Identification

The second set of subdetectors are those which identify particle type, and consist of a Cherenkov detector placed in front of the magnet (RICH1), a second Cherenkov detector placed downstream behind the tracking stations (RICH2) [20], five stations of muon detector (M1 to M5) [21] and the calorimetry system, containing the Electromagnetic CALorimeter (ECAL), Hadronic CALorimeter (HCAL), Pre-Shower (PS) and Scintillating Pad Detector (SPD) [22].

The RICH Detectors

When relativistic particles traverse a medium where the local speed of light is lower than the particle velocity, the natural photon emission which occurs through electromagnetic interaction of the particle and the bulk becomes coherent. A cone of light is generated, with characteristic opening angle dependent on the particle type, velocity and refractive index of the material. Given the momentum measurement from the tracking detectors, a likelihood value for the identity of each track can be assigned. The separation power can be seen in figure 1.10 for particles passing through C_4F_{10}, the gas contained within RICH1. As can be seen, the separation power within a specific medium depends strongly on momentum, and is limited at low and high extremes. Particles must have a minimum momentum in order to start producing Cherenkov radiation, thus each medium has a detection threshold which is particle dependent. At high momenta, the opening angle saturates and the power to separate particle type is lost.

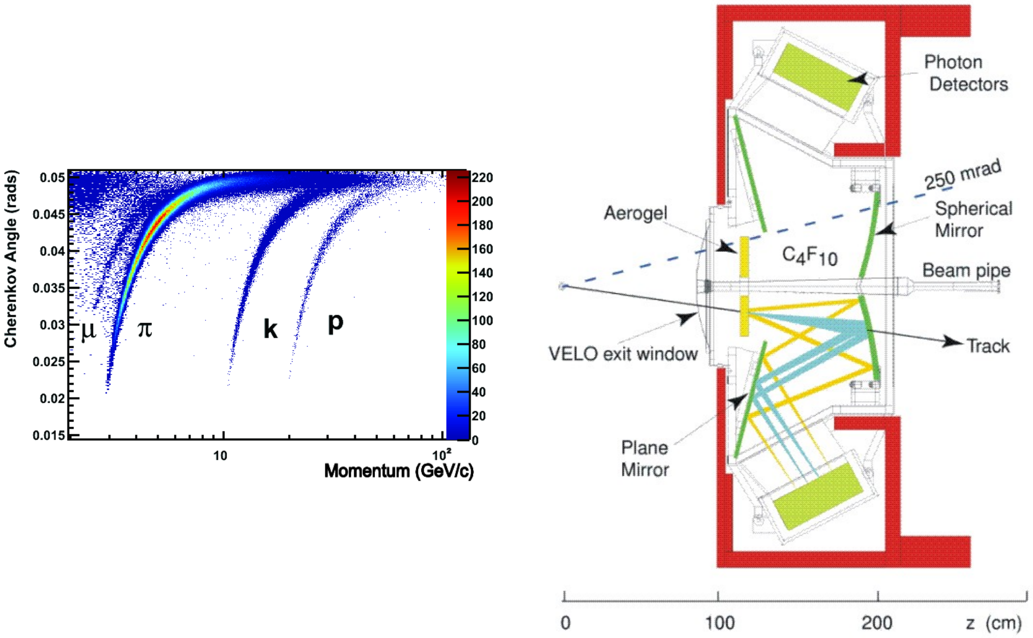

Figure 1.10: Left: Cherenkov angle versus momentum for several particle types, for a C_4F_{10} radiator. Right: Schematic of the RICH1 detector. Figures taken from [20].

As a large amount of the physics program at LHCb involves decays with pions, kaons and protons in the final states, the separation of these particles over the full momentum range was a strong requirement in the design of the experiment. To achieve accurate separation, two detectors were constructed: Ring Imaging Cherenkov detectors 1 and 2 (RICH1 and RICH2). RICH1 covers PID in the lower momentum region of 2 - 60 GeV/c, while RICH2 covers the region from 15 to over 100 GeV/c. The layout of both RICH detectors is similar, with differing focussing optics due to the different gas contents of each (C_4F_{10} for RICH1, CF_4 for RICH2). RICH1 contains, in addition to the C_4F_{10} gas, an Aerogel radiator which aids PID in the lower momentum region. A schematic of RICH1 can be seen in figure 1.10, and consists of an enclosure housing the gas, spherical mirrors which focus the Cherenkov light and transfer it to the focal plane, and the photon detectors mounted at the top and bottom of the detector. The photon detectors are Hybrid Photo Detectors (HPDs), which contain a single evacuated volume with a photocathode at the top and a silicon sensor bump bonded to the readout electronics at the bottom. An accelerating potential of $\sim$ 15 kV is placed across this volume in order to increase the detection efficiency of the generated photoelectron.

The Muon System

There are five muon stations in LHCb: one located directly in front of the calorimetry system (M1), and the others located at the back of the detector. The majority of the muon system is made up of Multi Wire Proportional Chambers (MWPCs), apart from M1 where the central region uses Gas Electron Multiplier (GEM) detectors. MWPCs are gas volumes containing thin wires held at high voltage. When particles interact with the gas mixture ($Ar/CO_2/CF_4$ in a ratio of 40:55:5) they create ionisation electrons which are attracted towards the wires: the resulting cascade provides the detection signal. GEMs work following this same concept, though consist of thin metal foils which contain an array of holes through which the cascade process is initiated. In LHCb, the GEMs detectors employ three GEM stages with the resulting signal collected directly by input pads on the readout PCB, and are used in the central region of the upstream station (M1) where there were concerns about the ageing of MWPCs.

The Calorimeter System

The calorimeter system in LHCb consists of several independent subdetectors: the Electromagnetic CALorimeter (ECAL), the Hadronic CALorimeter (HCAL), the Scintillating Pad Detector (SPD) and the Pre-Shower (PS).

The PS and SPD are used primarily in the online trigger, where they are used to distinguish electrons from photons (the SPD) and electrons from pions (the PS). Both detectors are built out of scintillating pads that use wavelength shifting (WLS) fibres to transmit the light to Multi Anode Photo Multipliers (MAPMTs), and are separated by a 15 mm thick lead converter. As photons do not produce scintillation light, a lack of signal in the SPD implies that the incoming particle was a photon. Misidentification from photon conversions does however occur, where the photon forms an electron-positron pair through material interactions upstream of the PS, though at low rate. The lead conversion layer liberates more energy from incoming electrons than it does from pions, and as such the energy deposited in the SPD can be used to distinguish between these two particles.

The ECAL and HCAL work on similar principles to the PS and SPD subdetectors. The ECAL contains 2 mm thick lead tiles interspersed with 4 mm of scintillator pads, which are read out by WLS fibres connected to phototubes. The HCAL is constructed using alternating iron and scintillator, with a total weight of 500 tons of iron to provide a depth of 5.6 interaction lengths. The effects of the PS/SPD (including lead convertor) are accounted for and allow the correct particle energy to be reconstructed.

PID Performance Results

Performance plots for particle identification are shown in figure 1.11. In general the PID efficiency varies between 90 % and 100 % for all particle types, varying with the acceptable rate of misidentification. The outcome of the various algorithms involved are normally presented as a log likelihood value, with cut values dependent on the analysis being carried out and the tolerable rate of mis-identification.

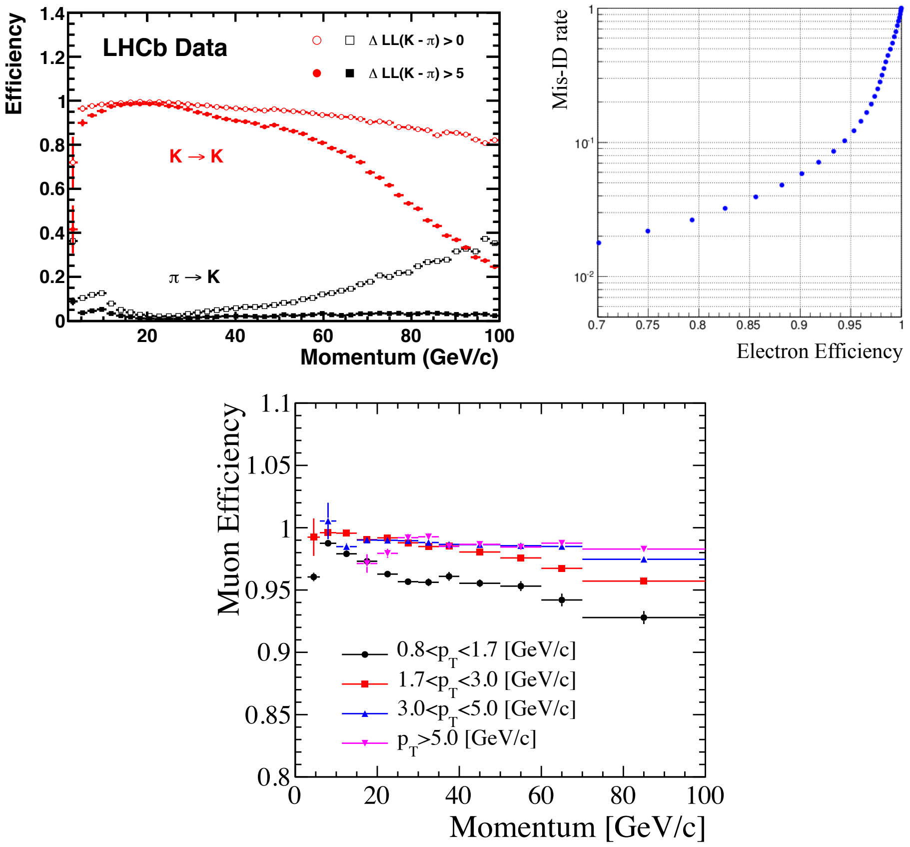

Figure 1.11: PID performance plots. Top left: Kaon efficiency and pion misidentification rate as a function of momentum. Top right: Electron efficiency versus misidentification rate. Bottom: Muon efficiency versus momentum in bins of transverse momentum. Figures are taken from [19].

1.2.3 Trigger

An important aspect in the operation of all detectors is the triggering strategy used to select interesting events. At the LHC, proton bunches are synchronised to a 40 MHz clock, giving a bunch spacing of integer multiples of 25 ns. During operation not all available bunch spacings are filled, and a total collision rate of up to 30 MHz is obtained in practice. The luminosity received is then the combination of beam area, number of particles and beam-beam overlap. As LHCb was designed to operate with a lower value of μ (the number of visible inelastic collisions per bunch crossing) than can be provided by the LHC, the relative overlap of the proton beams varies throughout a single fill (termed luminosity levelling) - allowing a constant luminosity to be delivered to the experiment. The nominal design values for LHCb are with $\mu = 0.7$ and an instantaneous luminosity of $2{\times}10^{32}$ cm^{-2} s^{-1}. For data taking during the first LHC run the bunch crossings were separated by 50 ns due to considerations for the accelerator, with 25 ns spacing expected for collisions in 2015 and beyond. Final data taking conditions of $\mu = 1.7$ and an instantaneous luminosity of $4{\times}10^{32}$ cm^{-2} s^{-1} were achieved during 2012 running, with a total integrated luminosity of roughly 3.5 fb^{-1} taken during the first run of the LHC.

Electronics at the LHC are designed around this concept of a 40 MHz bunch crossing frequency. On the detector front-end ASICs, events are sampled at 40 MHz and buffered locally for a fixed latency (160 bunch crossings), and can be read out using a dedicated trigger signal. The current trigger strategy, indicating the through rates at each stage, can be seen in figure 1.12. A hardware trigger (L0) takes information at 40 MHz from the calorimetry and muon systems, looking for large energy deposits or the presence of a moderately energetic muon. The L0 decision is distributed to the full range of subdetectors and the data for the selected bunch crossing is sent to the off-detector electronics. These consist of rack-mounted FPGA boards (the TELL1 [23]) which contain services such as slow control, clock distribution, data reception etc., and may perform some pre-processing of the incoming data (such as clustering). Once received,

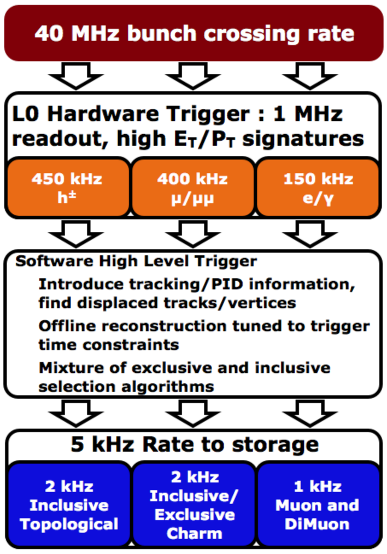

Figure 1.12: Trigger scheme for the current LHCb detector.

the data from these boards are combined over a switching network, and assembled as a single event for processing in the CPU farm.

Once inside the CPU farm, events are reconstructed and ran through several selective algorithms, split broadly into two sections: High Level Triggers 1 and 2 (HLT1 and HLT2) [24]. B-decays generally involve daughter particles originating from a displaced vertex and significant transverse momentum. This is used in HLT1, where tracks are selectively reconstructed throughout the full detector, with cuts on the IP and momentum. Those events which contain well-fitted tracks with sufficient momentum and displacement from the primary vertex are passed on to HLT2 for further processing, where there are several selection lines. For b-decays these include the n-body topological triggers, where events are accepted where they contain n tracks originating from the same displaced vertex, after a Boosted Decision Tree (BDT) multivariate cut.

As the LHC produces physics-quality beam for only around 30 % of its operation, the computing power of the CPU farm can be used to take advantage of the absence of beam. During 2012, 20 % of the L0 output was deferred to disk space on the CPU farm, to be processed during filling and machine down-time. This productive use of the computing resources allowed the lowering of some of the HLT1 threshold values for the reconstruction, and for a larger amount of data to be recorded to disk.

1.3 The LHCb Upgrade

As with most experiments, the diminishing returns for prolonged operation lead to the necessity of an upgrade, in order to pursue and extend the existing particle physics goals. LHCb is well placed to probe physics beyond the standard model both in the context of flavour physics (continuing with measurements of rare decays and CP violation searches) and more general physics in the forward direction. An upgrade to the detector is currently planned for the second long shut down of the LHC (LS2), expected at the end of 2018.

1.3.1 Trigger Upgrade

Both precision measurements and measurements of rare decays have in common the need for high statistics, and so the most obvious implication of an upgrade to LHCb is a significant increase in integrated luminosity. Before the upgrade of the detector an integrated luminosity of around 8 fb^{-1} will be accumulated, with a further 50 fb^{-1} to be taken during the upgrade. As mentioned above, this can only be achieved by increasing the overlap of the LHC beams, resulting in a higher number of proton-proton interactions in each bunch crossing. The expected instantaneous luminosity will therefore increase to around $2{\times}10^{33}$ cm^{-2} s^{-1}, with an average value of $\mu = 5.2$. Aside from any issues of increased granularity that might be required, the trigger becomes the most important limiting factor. While the current hardware L0 trigger is efficient at reducing the collision rate down to 1.1 MHz for input into the full reconstruction chain in the HLT, this will not be a valid approach in the upgrade due to the lack of discriminating power in the information available.

Importantly, the selective power for hadronic states in the current trigger scheme is very weak. While decays to muons can be easily distinguished due to the readout of the muon detectors for the L0 decision, the calorimeter information only allows for the

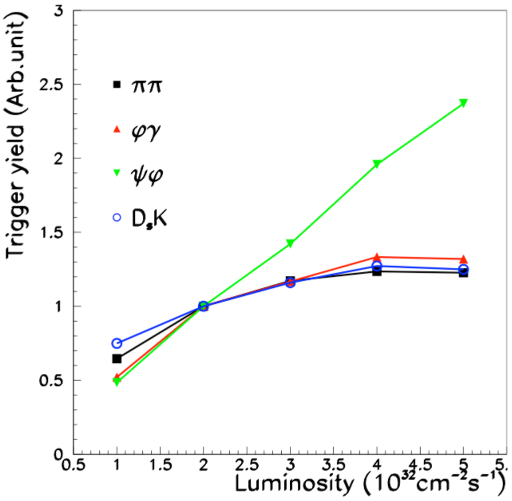

Figure 1.13: Trigger yields for several channels versus instantaneous luminosity, with the current LHCb trigger (current operational luminosity between 2 and $4{\times}10^{32}$ cm^{-2} s^{-1}).

selection of high energy objects, and in particular is unable to distinguish high energy hadrons produced in the primary vertex from those occurring in b-hadron decays. This leads to a loss in efficiency for hadronic-only final states with increasing instantaneous luminosity, illustrated quantitatively in figure 1.13. Simply increasing the luminosity at which the detector currently operates, or following the same strategy for the upgrade, would not significantly increase the yield of these events. This is why, while LHCb has been running significantly beyond its design luminosity during data taking in the first LHC run (due in part to the 50 ns bunch spacing), further increases necessitate a change in approach.

The trigger strategy for the upgrade is shown in figure 1.14 and is significant for the removal of all hardware triggering and the proposal to reconstruct each event in software. Such an approach is always advantageous (giving tracking information in the first decision stage), but due to the high collision rate for proton-proton collisions at the LHC rarely practical. Indeed, as will be discussed later, this requirement for the triggering is the most significant factor in the upgrade (requiring the replacement of almost all front end electronics). A more detailed description of the trigger is given with the subdetector changes below, and will be shown to successfully restore the scaling of trigger yield with luminosity.

1.3.2 Changes to Subdetectors

The VELO

Due to the higher occupancy and increased number of interactions per bunch crossing, the current VELO strip sensors must be replaced in the upgrade. In addition the current front end chip, the Beetle, has a maximum readout rate of 1.1 MHz, and is thus incompatible with the upgraded running conditions. Due to the heavy level of integration, this necessitates the replacement of the full VELO modules. After an internal review where two detector options were considered (a finer segmented strip

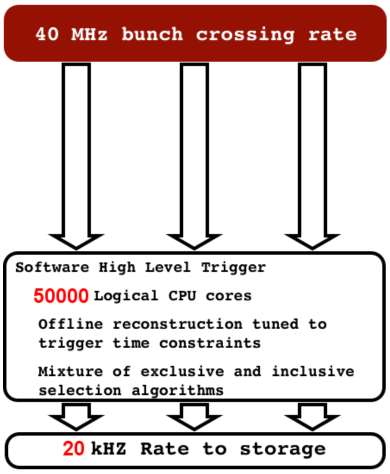

Figure 1.14: Trigger scheme for the upgraded LHCb detector.

detector or a pixel detector) it was decided to develop a pixel-based VELO [25]. Due to the different module geometries, a new RF-foil must also be produced.

The TT

As with the VELO, the TT modules must be completely replaced in order to increase their segmentation and readout functionality. The proposed Upgrade Tracker (UT) [26] will be a strip detector conceptually similar to the existing TT, read out using a new custom ASIC - the SALT chip. Following an idea which was originally intended to be used with the current detector, it has been proposed to use the small magnetic field that exists between the VELO and the UT in order to make a preliminary momentum measurement before track extrapolation through the magnet. This would allow a reduction in the search area for the downstream tracking, and allow faster reconstruction to be performed online for the trigger.

The Downstream Tracking stations

Due to the high occupancy in the OT even at current luminosity levels (in some cases surpassing 15 % with 50 ns bunch spacing), and the similar issues faced by the IT as for the VELO and TT, a substantial change in the downstream trackers is required in order to upgrade the experiment. Several proposals were put forward for this, all with a focus on reducing the occupancy through more granular detectors. An increase of the area covered by the silicon strips was suggested, in conjunction with the construction of shorter straw tube detectors and the upgrade of all front end electronics. A more novel idea was the construction of a scintillating fibre tracker, covering the full tracker region. After considering the expected performance and the practicalities involved (for example in the construction of new straw tubes, where the original construction site has been demolished) it was decided to proceed with a full Scintillating Fibre (SciFi) tracker.

As with the existing detector, there will be three downstream tracking stations. Each station will consist of four separate detecting layers, with orientations of 0°, +5°,

-5° and 0° with respect to the vertical axis, as with the UT. The detection layers will contain five or six stacked layers of fibres read out by Silicon Photo-Multipliers (SiPMs). Extensive studies have been carried out on both the performance and on the effects of radiation damage. For the latter in particular there has been detailed testing of the optical transmission properties of the fibres, and of the dark rate of the SiPMs after irradiation. Both have been found to be acceptable throughout the anticipated 50 fb^{-1} to be collected during the lifetime of the upgrade.

Expected Tracking Performance

Since the exact details of the detector designs are liable to change during the final optimisations currently underway for an installation during LS2, the performance of the upgraded tracking systems is also likely to vary in this time. Taking the performance from the Technical Design Reports (TDRs) some key performance plots are shown in figure 1.15. The upgraded tracking performance should deliver more precise IP and momentum measurements, in addition to maintaining the high tracking efficiency that has been achieved at present. Note that the momentum resolution shown in figure 1.15 does not include gains made by including UT hits in the track momentum measurement.

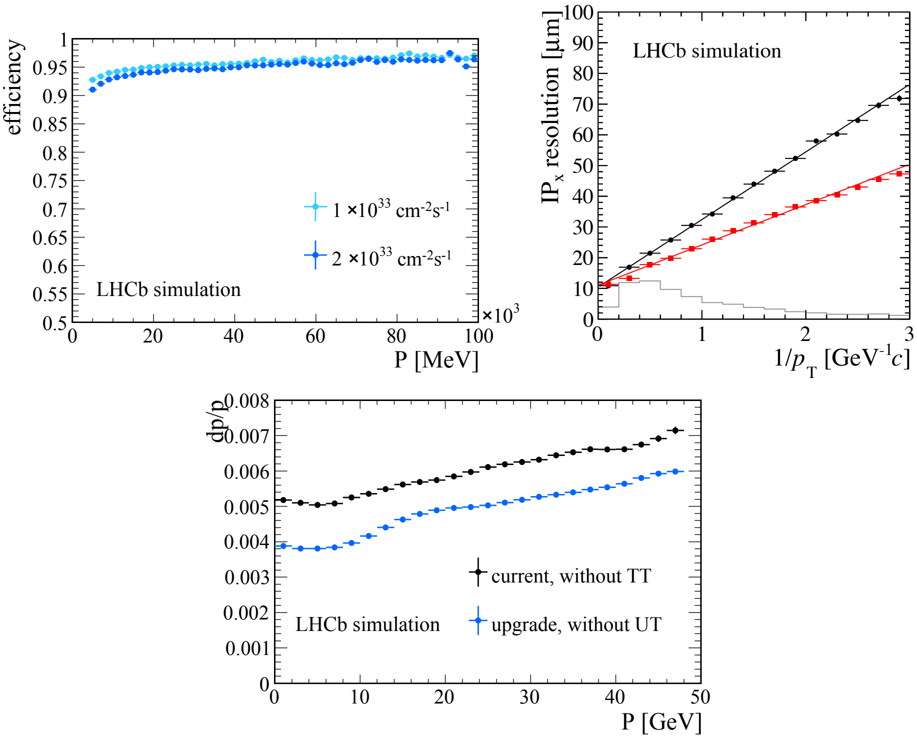

Figure 1.15: Upgrade tracking system performance plots. Top left: Long tracking efficiency versus particle momentum. Top right: Impact Parameter (IP) resolution in the x-direction for long tracks (current detector in black, upgraded VELO in red). Bottom: Momentum resolution versus momentum. Figures taken from [25, 26].

The RICH detectors

The changes to the RICH detectors are conceptually small, yet require a large amount of installation work to take place [27]. Indeed, several options were eventually discarded due to their installation times and the requirement that the work be completed during LS2. One such option was the creation of a combined RICH system with two gas volumes, focussing the light from each radiator onto different focal planes. Ultimately the changes to the detector will be limited to the replacement of the photon detectors and front end electronics, and a redesign of the RICH1 optics in order to reduce the occupancy on the focal plane. The current baseline is for Multi-anode Photo Multiplier Tubes (MaPMTs) to be used as the photon detector, read out by a custom ASIC which will contain the signal shaping, discrimination and further amplification stages.

The Calorimeters and Muon system

Both the calorimetry and muon systems require very little upgrading to cope with the conditions of the upgrade. For the calorimeters, new electronics to allow full data readout are under development, and will include some changes to the gain of the detector which will be reduced in order to enhance the lifetime of the MaPMTs. Due to the reduced role of the hardware trigger in the upgrade (see below) the information from the PS and SPD is not expected to play a part in the online trigger, and so these detectors will simply be removed. Similarly the muon detectors will be fitted with new off-detector electronics and the first station, M1 (before the calorimeters), will be removed. To reduce the particle rate on the inner section additional shielding will be put in place behind the hadronic calorimeter.

Expected PID Performance

A selection of simulated performance plots are shown in figure 1.16. Comparing with the existing detector, the RICH performance is expected to marginally improve across the whole momentum range, despite the more challenging track multiplicities. The performance of the muon detectors is however expected to degrade by around 5 % compared to present running conditions, mostly due to dead time in the muon chamber

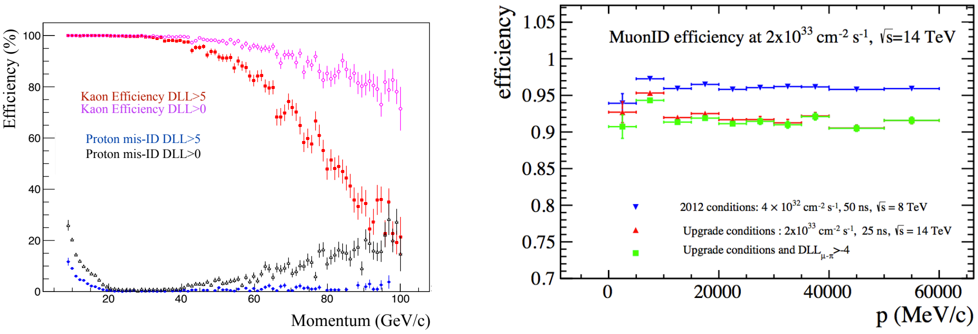

Figure 1.16: Upgraded PID performance plots. Left: Kaon efficiency and pion misidentification rate as a function of momentum. Right: Muon efficiency versus momentum. Figures taken from [27].

front ends, which are not being replaced. A possible upgrade to the muon system has been proposed, for a delayed installation in LS3, but the existing spare chambers and the upgrade of the off-detector electronics is considered sufficient to continue to provide high (order 93 %) efficiency in muon identification.

TORCH

Just as the aerogel in the current RICH1 detector was proposed to augment the PID performance in the low momentum region, a novel detector proposal has been put forward for the LHCb upgrade in order to improve significantly the performance for particles in the range 1 - 10 GeV/c [28]. The detector would use the time of flight to identify particle type, specifically using Time Of internally Reflected CHerenkov light (TORCH). TORCH is not tied strongly to the timescale of the upgrade, and will likely be installed in a subsequent long shut down (LS3).

Trigger Scheme

The upgraded trigger scheme, as mentioned, will be implemented fully within the CPU farm [29] and involve the removal of the current hardware L0 trigger. As with the current software trigger there will be two stages, roughly analogous to HLT1 and HLT2. Work on both triggers is still ongoing, but a rough plan for HLT1 will be for a stepped reconstruction chain very similar to the offline reconstruction, but with some mild cuts applied to the tracks after each processing stage. In this way the processing time can be reduced enough to fit within the available computing resources. With this information, candidates for the various inclusive and exclusive trigger lines can be selected, and written to disk as is done for the current detector. The output rate, and the implications for sharing of the output bandwidth, are currently under review and vary between 20 and 100 kHz event rates.

The amount of computing power expected at the time of the upgrade installation is also under review, and carries some uncertainty with it. The exact filling scheme, luminosity and ramping up of the experiment will all play a role in how much is needed and when, and so the move to full CPU processing for the trigger implies some risk. In order to have a fully functioning detector regardless of the exact size of the CPU-farm with the first LHC beams after LS2, a Low Level Trigger (LLT) has been proposed to

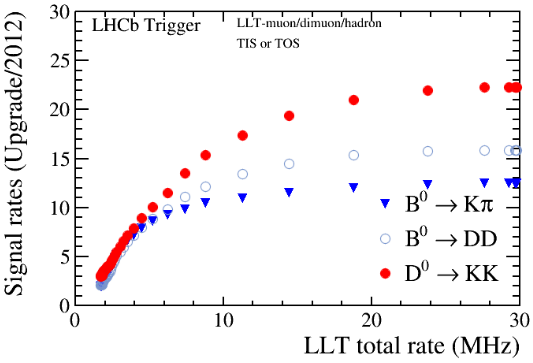

Figure 1.17: Signal yields for several physics channels as a function of LLT output rate.

throttle the event rate to the farm. This is an upgrade to some extent of the current L0 trigger, but with much looser requirements and a higher output bandwidth (being effectively removed once the online trigger farm can take the full event load). The gains expected for several signal channels can be seen in figure 1.17, and show that under upgrade conditions the trigger yield between what the L0 could output, and what the LLT can output (with an initial estimated output rate of 10 MHz) varies between a factor of 10 - 20.

1.4 The VELO Upgrade

The implications of the upgrade on the VELO are significant, and require the replacement of most of the existing detector. The move to 40 MHz readout necessitates a new ASIC, and the increased occupancy a more highly segmented sensitive area. This means the replacement of the modules, along with the electronics chain from the detector to the (new) off-detector DAQ boards. To fulfil these criteria, a pixel detector has been proposed based on an evolution of the existing Timepix ASIC [30].

1.4.1 Cooling and Module Design

In order to cool the modules to the -20°C that is required to prevent thermal runaway after heavy irradiation, and to control the annealing of the sensors, the VELO group have adopted the use of microchannel cooling for the upgraded detector. This consists of a series of tiny etched channels through the mechanical substrate through which the cooling fluid is directed, in order to deliver the cooling power directly to the heat source and replace the current method where a highly conductive spine allows the heat to flow out to the heat sink. Lab tests and simulations with the microchannel setup have shown extremely small temperature gradients across the proposed module

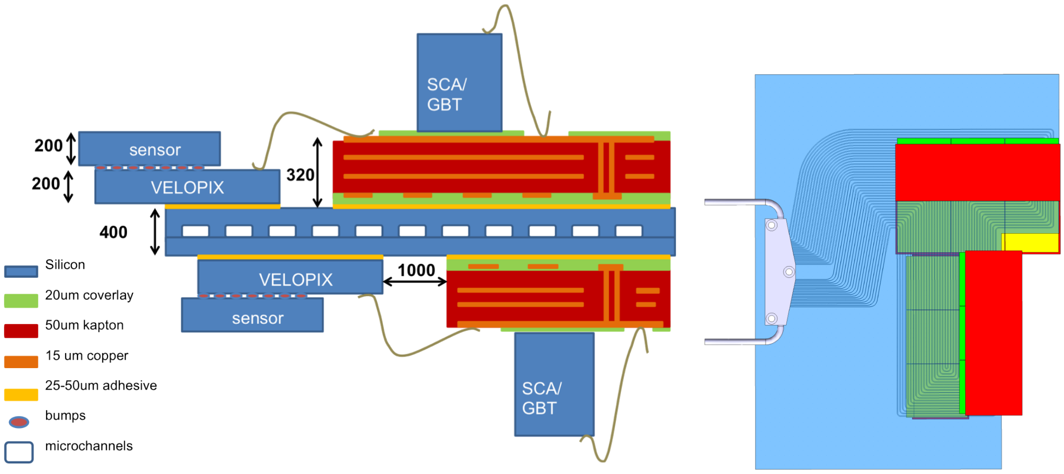

Figure 1.18: Left: Cross section of a VELO pixel module, highlighting the structure of the module including the microchannels which are etched into the silicon support and the overhang of the VELOPix ASIC and sensor. The colour indicates the material type. Right: Planar view, shown (red) top-mounted and (green) bottom-mounted sensor tiles. Figures taken from [25].

layout, with the most heavily irradiated silicon maintaining a temperature difference of approximately 6°C with respect to the cooling fluid. This was achieved using the expected power load at the end of lifetime, and shows that a sensor temperature of less than -20°C is feasible for the duration of the experiment.

In order to produce the cooling substrate, which will act as the mechanical support on which the module is constructed, a silicon wafer is first prepared by chemical etching of the coolant channels. Once these trenches have been produced, the wafer is then bonded to a second silicon wafer, which seals the channels. Typical dimensions for the conductive path are of the order 120 × 200 µm, with a total thickness of 400 µm for the complete substrate. On top of this substrate are mounted the VELOPix ASICs, and the kapton-mounted circuitry required to transmit the detector data to the edge of the vacuum tank. A planar view and cross section of the system are shown in figure 1.18 where the channels carrying the bi-phase CO_2 can be seen. Additionally, the figure shows the layout of the chips and sensors within the module - each sensor will cover three ASICs, and will be mounted on opposing sides of the module with a small overlap to eliminate inefficient regions. In order to reduce the amount of material close to the interaction point, it is anticipated to withdraw the cooling from the inner edge of the module, where the ASIC and sensor will then overhang.

1.4.2 VELOPix

One of the most challenging aspects of the upgrade will be the data collection and transmission from the front end. The VELOPix chip [31] will be a hybrid pixel chip with 256 × 256 pixels, of dimension 55 × 55 µm. The readout will be binary (without a measurement of the deposited charge) and use a data-driven readout method: hits registered will be immediately and independently transmitted off detector. This is necessary due to the huge number of hits and amount of data expected: for each colliding bunch, a mean of 32.8 tracks will traverse each module, with an average cluster size of 2.2. With a collision frequency of around 27 MHz this gives a mean output data rate of 36.8 Gbit s^{-1} per module, with large variation between the inner and outer chips due to the module orientation perpendicular to the beam. The inner chips will be required to transmit data at a mean rate of 10.2 Gbit s^{-1}, with a peak rate of around 15.1 Gbit s^{-1}.

In order to reduce the raw data rate coming off the chip, information is shared between several pixels to create a "Super Pixel" (SP). These are fixed regions covering 2 × 4 pixels where there is a central sharing of analogue circuitry along the row boundary, with the digital blocks placed on the outer regions. This allows some local buffering for events where the data from an individual pixel has not been sent out upon the arrival of a new hit, but more importantly allows the information sent off-chip to be compressed. The output packets (Super Pixel Packets - SPPs) contain the bunch crossing ID for the event when the hits were registered, the location of the super pixel within the chip, and an eight bit mapping which shows which pixels crossed threshold. A reduction of roughly 30 % in output data is achieved using this format.

Once the data from the chips has been transmitted it is assembled in the set of off-detector electronics boards, the TELL40 [32]. As the output from the VELOPix is data-driven, the hit packets arriving at the TELL40 have a variable latency, depending on the hit location and the link occupancy. The first job of the TELL40 is to time-order

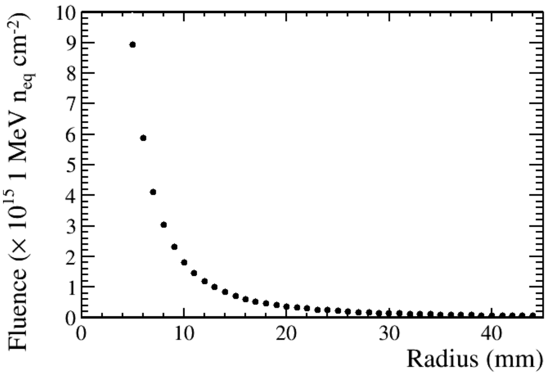

Figure 1.19: End of lifetime fluence across a single module close to the interaction point.

these data into events which can be passed on to the event-building farm, which gathers information from all of the subdetectors and passes them to the CPU farm for reconstruction. This is expensive in terms of FPGA resources, and does not leave much room for further pre-processing of the raw data before reconstruction. Further algorithms such as clustering are therefore expected to be carried out in the first software levels of the CPU farm.

While it is important that the TELL40 allows sufficient time for all of the data from a single event to be transmitted, there remains the issue of the front end response and the time at which each hit is registered. Timewalk is defined as the difference in response time for a circuit between an infinite charge injection and one close to threshold, and must be below the 25 ns bunch spacing in order for hits to be associated to the correct event. Current measurements with the prototype chip, TimePix3 [33], indicate that the time walk is below this level, but must be tested with sensors (as opposed to electrical injections by the front end) and after heavy irradiation.

1.4.3 Sensor Design

With sensor design for any experiment, the amount of inactive material around the sensor periphery is generally a concern, and indeed this is true for the VELO upgrade as well. However, the largest difficultly faced by the VELO is the high non-uniformity of the radiation damage across the sensor surface, and the consequences for running conditions. The variation of fluence at the end of lifetime for the VELO upgrade is shown in figure 1.19, and can be seen to vary by more than an order of magnitude across the module. In order to extract sufficient signal from the innermost region, bias voltages of up to 1000 V will be required [34]. As the current design calls for a three-chip sensor tile enclosed by a single guard ring structure, the regions of low fluence must be able to operate at these voltages without risk of electrical breakdown, placing considerable restrictions on the design of the sensor.

The decision to proceed with a three-chip sensor tile was intended to reduce dead regions between the separate ASICs. A spacing of 110 µm (2 pixels width) between chips has been proposed, with a subsequent elongation of the pixel implants in this region (shown in figure 1.20). This will give two rows of 110 µm pixels, on either the first or final row depending on the chip, with an inactive region surrounded the sensor

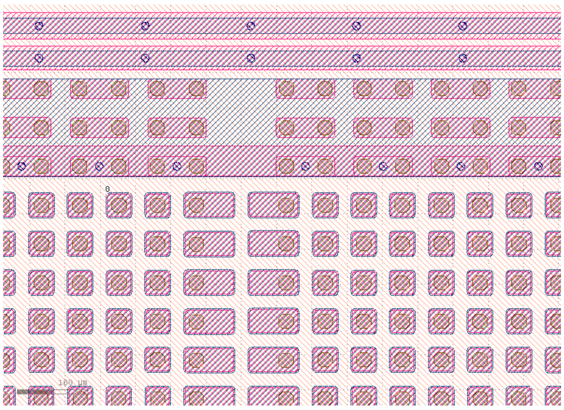

Figure 1.20: Pixel implants and bump bond pads, showing the elongated (110 µm long) structures at the boundary between two ASICs. Figure taken from [25].

tile of roughly 450 µm.

1.4.4 Infrastructure

While as much of the existing infrastructure will be reused for the upgrade, major components will by that time have been operating for around 10 years and thus there may be some replacement of voltage supplies etc. One of the major issues will be the replacement of the RF foil that separates the LHC vacuum from the module enclosure, and provides protection for beam-induced RF currents. The current foil was produced by metal pressing of a single massive sheet but it has proven difficult to control the foil thickness with this method. To produce the more complicated shape required by a square aperture, and to attempt to remove material from inside the LHCb acceptance, a new contraction method has been chosen which aims to provide a more uniform structure. This will involve the milling of a single block of Aluminium, removing 99 % of the material, and leaving only the 200 - 300 µm foil in the required shape. While a bold approach for a structure of such complexity and scale, prototyping of the RF foil has produced excellent results, with leak tests showing no detectable holes in the surface. Further thinning of the foil in the central region by means of chemical etching is also under investigation, and will aim to take the foil down to 100 - 200 µm in the region around the interaction point.

2

Theory

Starting with their first foray into particle physics at the NA11/NA32 [35] spectrometer at CERN in the early 1980s, silicon detectors have become a central component of tracking systems for modern particle physics experiments. Its predecessors, from the bubble chamber [36] to the Multi Wire Proportional Chamber (MWPC, invented 1968 [37]), were either limited in their detection and reconstruction time (requiring development of film or visible inspection of each event), or unable to achieve the resolution required for detailed investigations into the decay time of heavy particles. The ability to identify tracks at high rates on such a small scale, and observe decay points away from the interaction point (producing an array of combinatoric background), led to the adoption of silicon as the detector of choice where possible, and has enabled the measurement of decay times approaching the femtosecond range [38].

The development of VLSI (Very Large Scale Integration [39]) electronics, and more generally the advances made in silicon technology through the electronics industry, enabled the silicon detector to evolve and improve over time, leading to its widespread adoption by the particle physics community.

2.1 The Physics of Semiconductors

In isolation, an atom contains electrons held within discrete energy levels reflecting the rules of quantum mechanics through which they are governed. These levels are calculable for simple examples involving few bodies, typically only covering the solution of the Schrödinger equation for hydrogen atoms and ions of higher atomic number nuclei containing a single electron. The complexity involved in extrapolating this to more complex forms of matter is typically overcome by various approximations to states which can be more easily described. A powerful example of this is for a continuous medium, where the electronic properties of the bulk are to be studied. *Band theory* [40] calculates the single electron waveform for an infinite lattice of periodic structures of atoms or molecules, where the overlapping orbitals and interactions between atoms result in series of closely separated states which may be considered as single bands.

For single atoms of high atomic number, the inner-shell electrons sit in very deep energy wells, and these low energy states typically do not overlap between neighbouring atoms in a crystal lattice. The band produced from such states can usually be ignored, as they are unlikely to influence the electrical transport properties of the bulk material. Similarly there exist an almost infinite number of excited states, with increasingly high energies, which are not considered under normal conditions. The two most important

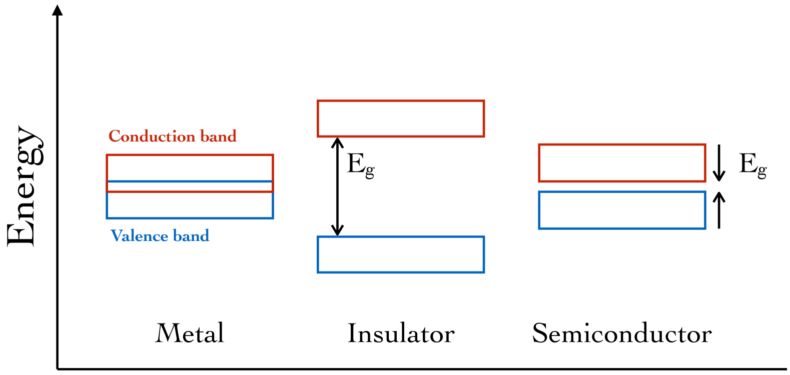

Figure 2.1: Band structure for conductors, insulators and semiconductors, illustrating the relative overlap of the conduction and valence bands.

bands used are the *valence* band and the *conduction* band. By observing the energy levels of these two bands (particularly their relative overlap), a given material is generally placed in one of three distinct categories: conductor, insulator or semiconductor (illustrated in figure 2.1). In conductors both bands overlap, allowing electrons to easily pass from the set of *filled* states in the valence band into some of the *empty* states in the conduction band. Once in the conduction band there are many empty states to which the electrons may move, and under the application of an electric field there is a significant flow of charge. For an insulator there is instead a wide energy gap (E_g) between the valence and conduction bands, far in excess of the thermal energy available at room temperature. The valence band remains full, with no transfer of states possible for the valence electrons, preventing the movement of charge under an applied field. Semiconductors sit in between these two extremes, with an energy gap between the valence and conduction bands low enough for there to be a supply of thermally-excited electrons occupying the conduction band. In this case there is a supply of charge carriers which will flow when an electric field is applied, both within the conduction band (where the electrons will pass between empty states) and within the valence band (where the *hole*, or empty state can be considered as a positively charged carrier).

2.1.1 Silicon

Silicon is a naturally occurring semiconductor, with an energy gap between the valence and conduction bands of around 1.12 eV at 300 K. In its single crystal form it has the same periodic structure as diamond: a face-centred cubic lattice, with a basis of two silicon atoms (one at (0,0,0) and one at $\left(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right)$). As indicated above, an electron from the valence band may be excited across the band gap into the conduction band, but for silicon this requires more energy than the band gap E_g. In the calculation of the band structure for any material, the momentum of the crystal lattice must be taken into account, and transitions between states with different momenta will require an interaction between the electron and the lattice. In silicon, the lowest energy state of the conduction band has a different momentum vector compared to the highest energy state in the valence band, and so a greater energy than E_g is needed in order to facilitate the momentum transfer to the crystal lattice. This results in an average electron-hole pair creation energy of 3.6 eV, and classifies silicon as an *indirect* semiconductor. Those materials where electron excitation occurs without a transfer to the lattice are known as *direct* semiconductors.

Carrier Density

The electron density n in the conduction band can be calculated using the Fermi-Dirac distribution function, which describes the probability for a state with energy E to be occupied at a given time. This probability $P(E)$ is given by

$$P(E) = \frac{1}{1 + e^{\left(\frac{E-E_F}{kT}\right)}} \tag{2.1}$$

where E_F is the Fermi level, k is Boltzmann's constant and T is the temperature. From the solution of the Schrödinger equation in a periodic potential, the density of energy states $N(E)$ can also be calculated,

$$N(E) = 4\pi \left(\frac{2m_n}{h^2}\right)^{\frac{3}{2}} \sqrt{E} \tag{2.2}$$

where h is Planck's constant and m_n is the effective electron mass. The electron density is then the integral of the occupied energy levels within the conduction band

$$n = \int_{E_c}^{E_{\text{top}}} P(E)N(E)\; dE \tag{2.3}$$

where E_c is the lowest energy state in the conduction band and E_{top} is the highest energy state. For values of $E - E_F$ much greater than kT (0.03 eV at room temperature) the Fermi-Dirac function can be approximated as

$$P(E) \approx e^{-\left(\frac{E-E_F}{kT}\right)} \tag{2.4}$$

and a final expression for the electron density in the conduction band can be reached

$$n = N_c\, e^{-\left(\frac{E_c-E_F}{kT}\right)}, \tag{2.5}$$

where

$$N_c = 2\left(\frac{2\pi m_n kT}{h^2}\right)^{\frac{3}{2}} \tag{2.6}$$

The expression for the density of holes in the valence band p can be similarly calculated, and found to be

$$p = N_v \; e^{-\left(\frac{E_F - E_v}{kT}\right)} \tag{2.7}$$

where

$$N_v = 2\left(\frac{2\pi m_p kT}{h^2}\right)^{\frac{3}{2}} \tag{2.8}$$

Analogously to before, m_p is the hole effective mass, and E_v is the highest energy state in the valence band.

Two concepts have been introduced here which play an important role in the electrical properties. The first is of the *effective mass* of the charge carriers. As the electrons in the conduction band are not free electrons but limited by the periodic boundaries of the crystal lattice, they appear to act as objects with a reduced mass. Holes are also considered to be real particles, which have opposing charge to the electron and their own effective mass. As charge can be carried by both, under an applied field there will be a current carried by electrons travelling through the conduction band but also by holes moving through the valence band. The second concept added above is the idea of the Fermi Level. In Fermi-Dirac statistics, the Fermi Level is the energy level where there is a 50 % probability of being occupied. Using the expressions above for the density of electrons and holes in a pure silicon semiconductor, and noting that these are equal, ie. $n = p = n_i$ (the intrinsic carrier density), then the Fermi Level can be calculated to be

$$E_F = \frac{E_c + E_v}{2} + \frac{3kT}{4}\ln\left(\frac{m_p}{m_n}\right) \tag{2.9}$$

At room temperature the second term is extremely small and the Fermi Level sits effectively halfway between the conduction and valence bands. Substituting values into the expression for the intrinsic carrier density similarly gives

$$n_i = \sqrt{np} = \sqrt{N_c N_v} e^{-\left(\frac{E_g}{2kT}\right)} \tag{2.10}$$

The intrinsic carrier density at room temperature is around 1.5×10^{10} cm^{-3}, compared with the atomic density of silicon which is roughly 10^{22} atoms cm^{-3}.

Doping

Given the large number of intrinsic carriers present at room temperature in pure silicon, any electrons or holes excited by ionising radiation will be difficult to distinguish from these thermally generated carriers. Two processes are carried out in order to overcome this: firstly, the silicon is *doped*, and secondly a *p-n junction* is created.

In the crystal structure described above, each silicon atom is covalently bonded to four other silicon atoms, completing the outer electron shell (the *p*-shell) and producing a stable structure. If a small number of impurity atoms with a differing number of valence electrons are introduced into the lattice, then the overall structure will remain the same while introducing a local excess or deficit of electrons. This is shown schematically in figure 2.2 for two common dopants, Phosphorus and Boron. The introduced impurity is generally either a group III or V element (as opposed to silicon, which is in group IV), and is said to be *n-type* if it results in an excess of electrons, and *p-type* if it results in a deficit.

When the doping is n-type, the excess electrons are not tightly bound to their parent atom, and are located in a state close to the conduction band (illustrated in figure 2.3). The energy difference between this band and the conduction band is dependent on the dopant used, but is typically close enough such that the majority of electrons have sufficient thermal energy to reach the conduction band. For this reason, the dopant is said to be an *electron donor*. The number of carriers introduced is approximately the same as the number of donor atoms, and given typical dopant densities in excess of 10^{14} atoms cm^{-3}, far exceeds the number of intrinsic carriers at room temperature. The electrons are thus the *majority carriers*, and are the dominant contributors to charge flow.

For p-type dopants, there is a local deficit of electrons around the impurity atoms. These holes can be occupied by electrons from the valence band, and this new set of empty states is typically only slightly higher in energy, allowing the majority of them to be filled by thermally-excited electrons at room temperature. The movement of electrons to the holes now present in the valence band give the apparent motion of a positively charged hole, allowing current to flow. In this case, the dopant is said to be an *electron acceptor*, and the majority carriers are the holes (analogously to n-type doped silicon). Both n-type and p-type doped silicon is referred to as an *extrinsic* semiconductors, as opposed to pure silicon where the charge carriers arise from the silicon itself.

In both cases, the addition of this new band leads to a change in the position of the Fermi Level, in order to reflect the higher electron occupancy of the conduction band (or increased hole occupancy of the valence band). For n-type doped silicon, the

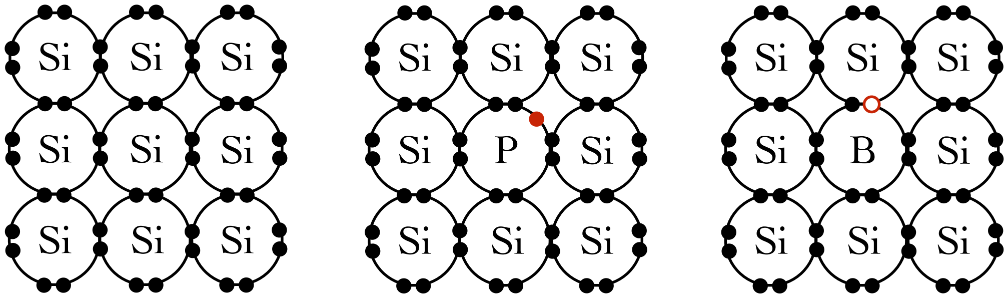

Figure 2.2: Electronic configuration of the valence electrons in pure (left), n-doped (centre) and p-doped (right) silicon.

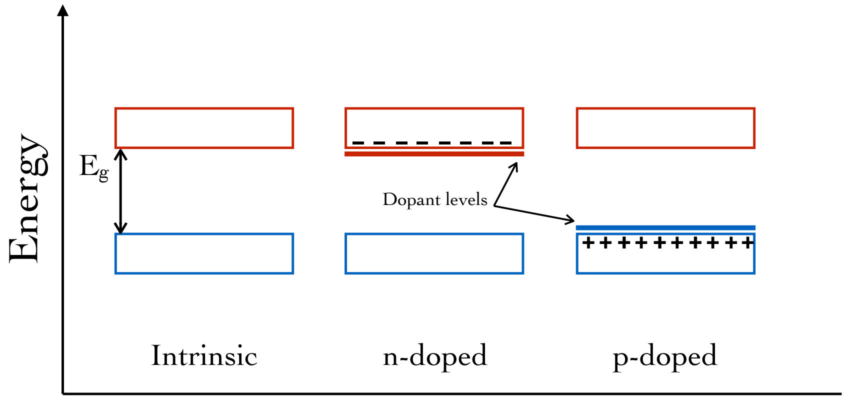

Figure 2.3: Band structure for pure (left), n-doped (centre) and p-doped (right) silicon. The carriers for doped silicon are additionally shown.

density of carriers in the conduction band (equation 2.5) is now equal to the n-type dopant concentration N_d

$$N_d = N_c \, e^{-\left(\frac{E_c - E_F}{kT}\right)} \tag{2.11}$$

giving a new expression for the Fermi Level of

$$E_F = E_c - kT \ln\left(\frac{N_c}{N_d}\right) \tag{2.12}$$

It can been from this that for the addition of n-type dopant, the Fermi Level is located closer to the conduction band. Analogously, for p-type doped silicon the Fermi Level is located closer to the valence band,

$$E_F = E_v + kT \ln\left(\frac{N_v}{N_a}\right) \tag{2.13}$$

where N_a is the p-type (acceptor) density.

Carrier Motion

The flow of charge within the semiconductor can be described by the current density J,

$$J = J_{\text{drift}} + J_{\text{diffusion}} \tag{2.14}$$

which contains two terms, one related to the random motion of the charge carriers $J_{\text{diffusion}}$, and the other related to their motion under an electric field J_{drift}. The latter term describes the collective motion of the charge carriers under an applied electric field. For electrons, this is

$$J_{\text{drift}} = -qnv_n \tag{2.15}$$

where q is the electron charge and the electron velocity is determined by conservation of momentum to be

$$v_n = \frac{1}{m_n} \int -qE \; dt = -\frac{qE\tau}{m_n} = -\mu_n E \tag{2.16}$$

The constant μ_n is the electron *mobility*, and τ is the mean free time before collision with the lattice. A similar expression for the motion of holes can be written, to give an overall drift current density of

$$J_{\text{drift}} = (qn\mu_n + qp\mu_p)E = \sigma E \tag{2.17}$$

This constant of proportionality is the *conductivity* of the silicon, with corresponding *resistivity* ρ defined as

$$\rho = \frac{1}{\sigma} = \frac{1}{qn\mu_n + qp\mu_p} \tag{2.18}$$

For doped silicon, where one of the carriers dominates, this expression can be simplified to

$$\rho = \frac{1}{qn\mu_n} \tag{2.19}$$

for n-type doping, and

$$\rho = \frac{1}{qp\mu_p} \tag{2.20}$$

for p-type.

The diffusive contribution to the current density, $J_{\text{diffusion}}$, arises from the motion of carriers from regions of higher to lower concentration, simply described by

$$J_{\text{diffusion}} = qD_n \nabla n - qD_p \nabla p \tag{2.21}$$

D_n and D_p are the electron and hole diffusion constants, related by the Einstein equation to the carrier motion

$$D = \frac{kT}{q}\mu \tag{2.22}$$

The overall expression for the current density can now be written,

$$J = \sigma E + qD_n \nabla n - qD_p \nabla p \tag{2.23}$$

The above derivation contains the mean free time τ, which is assumed to be a constant. This is due to the dominance of the thermal velocity in determining the mean free path. For modest values of applied field, the drift velocity remains far below the thermal velocity, and this along with the mean free path before collision with a lattice atom determines the mean free time. For larger fields, the drift velocity component will reduce the mean free time, and the drift velocity will be observed to saturate. Significant deviation for electrons occurs around 10^4 V cm^{-1}.

2.1.2 P-N Junctions

While the addition of charge carriers through doping might at first appear counterintuitive to the goal of making the signal of an interacting particle more easily detectable, it is a necessary step in producing the most fundamental device in modern electronics: the *p-n junction*. A p-n junction consists of two regions of silicon in contact with each other, one of which is contains p-type dopant, and the other n-type. This simple building block underpins modern transistors, and by extension the electronics industry.

The primary features of a p-n junction are the diffusion of carriers from one dopant region to another, and the response of this motion under an applied field. When brought into contact, the n-type region contains an excess of electrons in the conduction band, and the p-type region an excess of holes in the valence band. The empty conduction levels in the p-type region can accept electrons from the n-type, while electrons from the valence band of the n-type region can also occupy the empty valence states in the p-type. This sharp gradient of charge carriers gives the migration of both conduction electrons and valence holes into the oppositely-doped region.

Built-in Potential and Depletion Region

The removal of carriers from both sides of the junction results in the buildup of a static electric field. For n-type silicon, the loss of the mobile electron leaves an uncovered atomic charge, creating a positive field which acts to oppose the diffusive electron current. Similarly in the p-type side, the occupancy of holes by valence electrons from the n-type region causes a negative space charge buildup. This acts against further electron flow into the hole states, usually visualised as migration of holes. A steady-state situation is eventually reached where a non-zero electric field exists on both sides of the junction, leaving it devoid of free charge carriers. This is termed the *depletion region*.

The electric field that is created at the junction is dependent on the dopant concentrations on either side. An approximation of the space charge distribution is generally used, where the dopant atoms are assumed to be ionised fully over the depletion region. This is called the *full depletion approximation*, and gives the space charge and electric field profiles shown in figure 2.4. The width of the depletion region can be calculated by considering this accumulated charge. On both sides, the total charge must balance such that

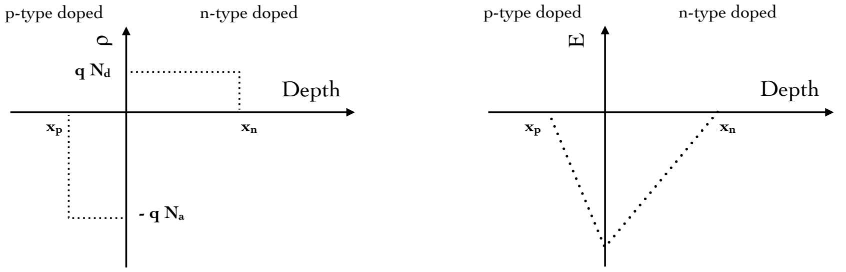

Figure 2.4: Built-in space charge (left) and electric field (right) distributions for a simple p-n junction.

$$qN_dx_n = qN_ax_p \tag{2.24}$$

where the depth of the depleted region in the n- and p-type sides is x_n and x_p respectively. At these points the electric field is zero. For the depletion region, the field can be calculated as a function of depth by use of Gauss' law

$$\frac{dE(x)}{dx} = \frac{\rho(x)}{\epsilon} \tag{2.25}$$

where $\rho(x)$ is the charge density. This can be written in terms of the dopant concentration in each side

$$\frac{dE(x)}{dx} = \frac{q}{\epsilon}(N_d(x) - N_a(x)) \tag{2.26}$$

where $N_d(x) = N_d$ in the n-type region, and 0 in the p-type (and vice-versa for N_a). The electric field in each region becomes

$$E(x) = -\frac{qN_a}{\epsilon}(x + x_p) \tag{2.27}$$

in the p-type region, and

$$E(x) = \frac{qN_d}{\epsilon}(x - x_n) \tag{2.28}$$

in the n-type. The naturally arising potential across the junction, V_{bi}, is the integral of this field, ie.

$$V_{bi} = -\int E(x)dx = \frac{qN_dx_n^2}{2\epsilon} + \frac{qN_ax_p^2}{2\epsilon} \tag{2.29}$$

Since the full depletion width x_d is simply the sum of the n-side and p-side depths, the total width of the depletion region can be described in terms of the built-in potential

$$x_d = \sqrt{\frac{2\epsilon V_{bi}}{q}\left(\frac{1}{N_a} + \frac{1}{N_d}\right)} \tag{2.30}$$

The evaluation of this potential can be performed in many ways, one of the most straightforward of which is to consider the change in mean electron energy across the junction. This is the difference in Fermi Level, and can be written using the previously derived expression for the Fermi Level in doped silicon as

$$\begin{aligned} qV_{bi} &= E_{F_{\text{n-side}}} - E_{F_{\text{p-side}}} \\ &= (E_c - E_v) - kT\ln\left(\frac{N_c}{N_d}\right) - kT\ln\left(\frac{N_v}{N_a}\right) \\ &= E_g - kT\ln\left(\frac{N_cN_v}{N_dN_a}\right) \end{aligned} \tag{2.31}$$

Using a modified form of the previous expression for the number of intrinsic carriers n_i

$$n_i^2 = N_cN_ve^{-\left(\frac{E_g}{kT}\right)} \tag{2.32}$$

we can replace the band-gap energy in the expression for the built-in potential

$$qV_{bi} = kT\ln\left(\frac{N_cN_v}{n_i^2}\right) - kT\ln\left(\frac{N_cN_v}{N_dN_a}\right) \tag{2.33}$$

$$\Rightarrow V_{bi} = \frac{kT}{q}\ln\left(\frac{N_dN_a}{n_i^2}\right) \tag{2.34}$$

These expressions now allow both the width of the depletion region, and the built-in potential to be calculated, and are entirely controlled by the doping of the two regions.

Forward and Reverse Biasing

The depletion region produced by the built-in potential can be be altered by the application of an externally generated potential difference, applied across the junction. This can act either in favour, or against, the already existing potential. If a circuit is constructed where the positive side is connected to the p-type region, and the negative side is connected to the n-type region, then the electric fields in both will be suppressed. This has the effect of reducing the width of the depletion region, and once this has been reduced completely there will be no barrier to charge flow across the junction. Further increasing the applied voltage will generate large currents, similar to the use of a single piece of n- or p-type doped semiconductor. Such a system is said to be *forward biased.*

The opposing situation, where the sign of the external potential matches the built-in potential (positive voltage to n-doped, negative to p-doped) will cause the depletion region to extend further from the junction, and the system is said to be *reverse biased.* No significant current will flow, although any thermally-generated current which is produced within the depletion region will be swept to the end of the depletion region. If the depletion width reaches the full extent of the p-n junction, the device is considered *depleted.* This is now the state at which incoming particles and radiation might be detected. As with the thermally-generated charge, any interaction of a particle in the p-n junction will produce free carriers, which will be transported through the external circuit and may be measured. This is the essence of silicon detector operation.

For reference, the width of the depletion region can be calculated analogously to the built-in depth, using

$$x_d = \sqrt{\frac{2\epsilon(V_{bi}+V)}{q}\left(\frac{1}{N_a}+\frac{1}{N_d}\right)} \tag{2.35}$$

where V is the applied voltage, and is positive for reverse biasing.

2.1.3 Charge Generation

It is necessary to investigate how ionising radiation interacts with matter in order to understand the signals it generates in a detector. For particles produced in high energy particle collisions which survive long enough to reach the vertex detector, there are broadly three categories in terms of their interactions with matter: charged particles, neutral hadrons and photons. Of these, neutral hadrons do not interact via the electromagnetic force and thus leave no signal in silicon detectors, and photons of the energies observed at the LHC predominantly convert into electron-positron pairs if they interact. Simply by virtue of their frequency, the most important decay products to measure are the charged particles which emanate from the interaction point.

There are several mechanisms through which a charged particle may transfer energy to intervening material [41]. Most of these interactions are electromagnetic, though nuclear interactions may occur with a much reduced probability. As such an interaction would likely prevent the particle from continuing on its path through the detector, and occur which such low frequency, they are not considered further. Of the possible electromagnetic interactions, for different ranges of incident particle energy different effects dominate. Electromagnetic interactions are primarily between the incoming particle and the atomic electrons which cover the greatest area of the atom, and the relative masses between the two interacting particles is an important factor in determining the collision dynamics. For this reason, electrons and positrons are generally considered separately, with their own distinct behaviour. The low mass of the electron has a significant effect on the range over which the various phenomena are dominant, in particular for photon radiation (discussed below).

For most particles considered, priority is thus for those which have a mass much greater than the electron (that is to say, all charged hadronic matter and heavy leptons). The two main energy loss mechanisms for the range $0.1 < \beta\gamma < 10^4$ (where β is the velocity as a fraction of light speed and γ is the Lorentz factor) are *electron scattering* and *radiative losses*. The first of these arises due to the energy transfer from the incoming particle to an electron held within the medium. In the case of silicon the energy transferred may be used to promote a valence electron into the conduction band, producing an electron-hole pair. This is the primary mechanism of signal generation in silicon detectors. Radiative effects occur only at very high energy, and are very strongly controlled by the mass of the particle. For light particles such as muons and pions this onset occurs beyond energies of several hundred GeV, and for heavier particles much higher still. Radiative energy losses come from the interaction of a moving particle with an electric field, which can result in several distinct events including photon emission (*Bremsstrahlung*), electron-positron pair creation and photo-nuclear interactions. As the most common mechanism, only energy transfer to electrons is considered further.

Mean Energy Loss of Particles in Matter

The mean energy loss $\left\langle \frac{dE}{dx} \right\rangle$ in the approximate range $0.1 < \beta\gamma < 10^4$ for charged particles in matter is described by the Bethe formula [42]

$$-\left\langle \frac{dE}{dx} \right\rangle = Kz^2 \frac{Z}{A} \frac{1}{\beta^2} \left[\frac{1}{2} \ln \frac{2m_e c^2 \beta^2 \gamma^2 T_{max}}{I^2} - \beta^2 - \frac{\delta(\beta\gamma)}{2} \right] \tag{2.36}$$

with variables defined in table 2.1. This expression comes about from considering the density of electrons through the absorbing medium, and the mean energy transferred through interaction with the incoming particle. A minimum energy is required to be transferred, implying a cutoff at the low end of the energy transfer spectrum. At the upper range, the maximum transferrable energy is limited by momentum conservation between the ingoing and outgoing states.

An example of the energy loss spectrum for muons passing through copper is shown in figure 2.5, where the distinct regions referred to above can be seen. The $\beta\gamma$ range from 0.1 to 2 - 3 is known as the *kinematic range*, where the energy loss varies as the inverse square of the velocity. There is then a transition into a slowly increasing range, where relativistic effects cause the extension of the particles transverse electric field. The junction between these two ranges, where $\beta\gamma$ = 2 - 3, is the point of minimal ionisation, and a particle with such characteristics is termed a *Minimum Ionising Particle* (MIP). Without corrections, the energy transfer is predicted to rise indefinitely due to relativistic effects, but in practice the ionisation of the medium tends to shield more distant electrons from interactions (usually termed the *density effect*). Thus the increase in transferred energy remains low until the onset of radiative contributions. For this reason, particles with $\beta\gamma > 2$ are often approximated as MIPs.

Symbol	Meaning
A	Atomic mass of absorber
m_e	Electron mass
N_A	Avogadro's number
r_e	Classical electron radius
K	$4\pi N_A r_e^2 m_e c^2$
z	Charge of incident particle (relative to e)
Z	Atomic number of absorber
c	Speed of light in vacuum
γ	Lorentz factor
β	Incident particle velocity (fraction of light speed)
T_{max}	Maximum possible energy transfer in a single collision
I	Mean excitation energy
$\delta(\beta\gamma)$	Density correction
x	Detector thickness in g cm^{-2}

Table 2.1: Variable definitions in the Bethe formula

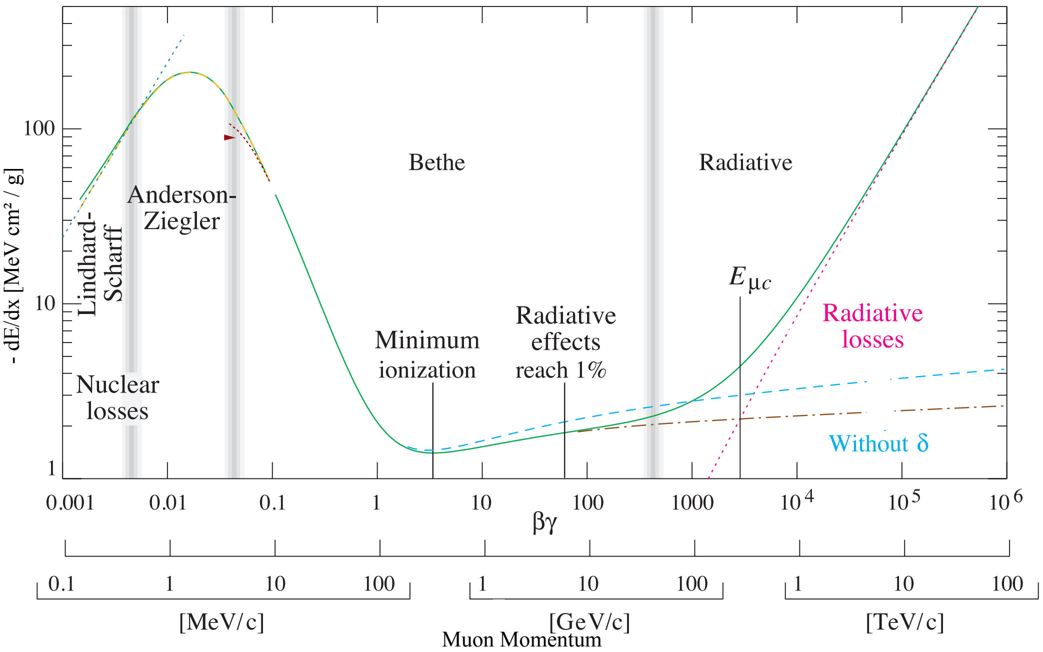

Figure 2.5: Mean energy loss spectrum for muons passing through copper.The dashed blue line indicates the rise in mean energy transfer if density effects were excluded. The correction is shown in the dashed brown line. Figure taken from [41].

Charge Spectra for Thin Materials

While the Bethe equation adequately describes the *mean* energy loss for a particle in the range $0.1 < \beta\gamma < 10^4$, the use of this in practical detector measurements is limited. The charge spectrum produced by interaction with a thin layer of material suffers disproportionally from high energy-transfer collisions, skewing the distribution from Gaussian. Many approaches have been applied in literature [43] in an attempt to adequately describe the energy spectrum for a given particle, momentum and absorber thickness, with the most common experimental use of a Landau distribution convoluted with a Gaussian. This is primarily due to the non-analytic nature of many of the other approaches, and the success of the Landau function in predicting the *most probable value* (MPV) which is characteristically taken as the value of interest (as opposed to the mean energy loss described above). The Landau-Vavilov distribution [44] describes the most probable energy loss Δ_p as

$$\Delta_p = \xi \left[\ln \frac{2m_e c^2 \beta^2 \gamma^2}{I} + \ln \frac{\xi}{I} + 0.200 - \beta^2 - \delta(\beta\gamma) \right] \tag{2.37}$$

where

$$\xi = \frac{K}{2} \left\langle \frac{Z}{A} \right\rangle \frac{x}{\beta^2} \tag{2.38}$$

using the definitions in table 2.1 and the mean value of Z/A for the material. An example spectrum is shown in figure 2.6. As can be seen, the high tail continues to

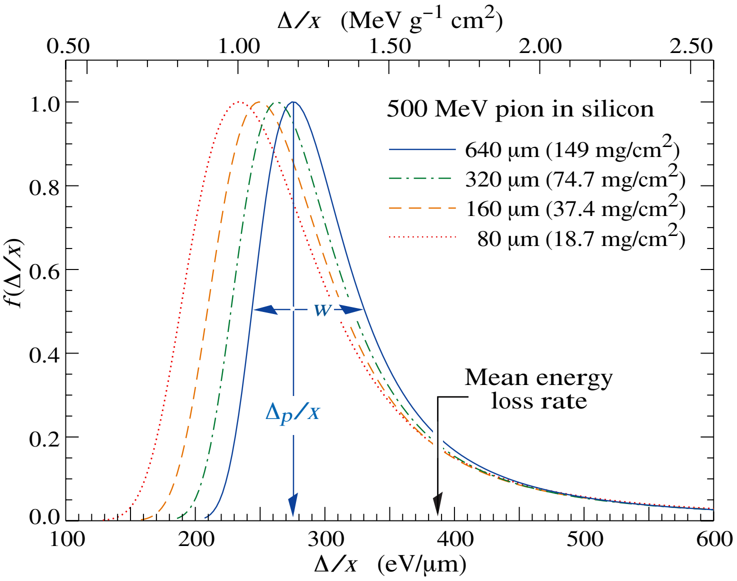

Figure 2.6: Energy loss spectrum for 500 MeV pions passing through silicon detectors of varying thickness, taken from [41].

significant energies which is what pulls the mean energy loss of the Bethe equation. The Landau function does not however accurately describe the width of the energy distribution in thin materials. In thicker materials, the spectrum width is given by 4ξ.

In addition to the overall shape of the energy spectrum, it is worth pointing out the existence of a particular type of event: *delta rays*. These are the electrons which are produced with energy significantly higher than the ionisation energy of the medium they are in, from the tail of the energy loss spectrum. They travel throughout the bulk, continuing to ionise further electrons, and following a straggling path (due to their equivalent mass). Such electrons can travel for many microns in silicon, and alter the spatial profile of the energy deposition.

Multiple Scattering

In addition to the transfer of energy in the interaction with material, the flight path of a particle will also be altered. Coulomb scattering is generally described using the formulation by Molière [45]. The angular distribution of the majority of tracks (>98 %) around the incident angle can be approximated as a Gaussian distribution with standard deviation θ, given by

$$\theta = \frac{13.6}{\beta cp} z \sqrt{\frac{x}{X_0}} \left[1 + 0.038 \ln \frac{x}{X_0} \right] \tag{2.39}$$

where $\frac{x}{X_0}$ is the thickness of the medium in terms of its radiation length (the mean distance over which a high energy electron will lose all but $\frac{1}{e}$ of its energy through bremsstrahlung). The distribution is observed to depend strongly on the particle momentum, and the extent of scattering reduces as the particle momentum increases.

This effect has a strong influence on the design of particle physics experiments, with the resolution in the inner region typically dominated by the scattering from material between the interaction point and tracking system, as well as between stations of the tracking system itself.

2.1.4 Detectors for Particle Physics

Silicon Sensors

Most silicon detectors are based on the use of p-n junctions as described above, though in a slightly more complicated form. In general, the detection medium will be a substrate of doped silicon, with an *implant* of high dopant concentration on one or both sides. When a bias voltage is applied to the substrate, the depletion region spreads from the bulk-implant interface to extend throughout the whole material. Charge generated through the bulk is collected at the implant (with the opposing charge carrier arriving at the back of the sensor), which is connected to the readout electronics. The implant can be either directly coupled (DC), such as for pixel detectors, or capacitively coupled (AC) as for most modern strip detectors. The physical shape of the implant further controls the spatial information provided about the event, with the two most common implementations being long parallel strip implants, giving the one-dimensional hit location, and a matrix of rectangular implants (pixels) giving two-dimensional information.

Given the negligible depth of the implant, the properties of the sensing medium will be controlled by the bulk doping type. As seen previously, the width of the depletion region with an applied voltage V is

$$x_d = \sqrt{\frac{2\epsilon(V_{bi}+V)}{q}\left(\frac{1}{N_a}+\frac{1}{N_d}\right)} \tag{2.40}$$

Taking the bulk doping to be N, and ignoring the built-in potential (typically < 1 V), this becomes

$$x_d = \sqrt{\frac{2\epsilon V}{qN}} \tag{2.41}$$

which means that in order to deplete the full substrate, a voltage V_{dep} must be applied, given by

$$V_{dep} = \frac{qNd^2}{2\epsilon} \tag{2.42}$$

for substrate thickness d. Using the previous relationship between the dopant concentration and the resistivity, this can be rewritten as

$$V_{dep} = \frac{d^2}{2\epsilon\mu\rho} \tag{2.43}$$

For a particle which traverses the full depth of the sensor, this is therefore the minimum bias voltage which must be applied in order to achieve complete charge collection. However, as full depletion is by definition the point at which the electric field extends the full depth of the sensor, the field strength is necessarily zero at one of the sensor surfaces. This implies that the charge collection time for charge deposited close to this surface will be quite long, given the small drift velocity imparted by the field. As such, detectors are typically operated over-depleted in order to ensure a rapid collection time. Typical sensor thicknesses are 100 - 300 μm, with the thinner end generally limited by the difficulties in handling the device and in collecting enough charge as to give a high signal to noise ratio.

The field created by the applied voltage extends isotropically from the implant-bulk interface for a homogeneous substrate. As the sensor dimensions are not infinite, the behaviour at the sensor edge is important. Sensor fabrication is performed on large wafers (typically 6" or 8"), and the individual devices are generally removed from the wafer either by mechanical dicing with a diamond-tipped saw or by laser etching. The sides of the sensor suffer considerable microscopic damage, and the cutting results in a rough, jagged edge containing many dangling bonds. If the electric field reaches this damaged region, there can be significant charge flow to the back side, giving a significant increase in the leakage current and physically damaging the sensor. To prevent this, structures are placed between the last operational pixel and the physical edge of the device, in order to step down the potential. Several doped rings are usually patterned onto the side of the sensor where the p-n junction (and thus highest field) lies, around the whole of the device. These *guard rings* help to dissipate the field before it reaches the damaged edge region. As the field is stepped down in this region, the guard rings normally represent an inactive border around the sensor, which is not sensitive to incoming charge. Much work is dedicated to the minimisation of this *dead region*, but modern particle physics experiments still operate with typically 400 - 500 μm or so distance between the final pixel and the cleaved edge.

Given the possibility of either p- or n-type doping, four possible sensor configurations exist for combinations of bulk and implant type. Historically, a combination of higher electron mobility and the availability from industry of high resistivity n-type wafers meant that n-type bulk detectors were most commonly used. For applications where radiation damage was not a strong consideration, a p-type implant was used to minimise processing costs. The use of a heavily doped n-type implant, embedded in an n-type bulk sensor, was used for particle physics experiments where the sensor was expected to be heavily irradiated. These n^+-on-n sensors (where n^+ denotes very high n-doping) operate by depositing a thin layer of p-type dopant on the back side of the sensor, creating a p-n junction at the opposite side from the implants. Such devices are *back-depleted*, and the electric field profile is reversed with respect to a p-on-n sensor.

P-on-n and n-on-p sensors are both *front-depleting* devices, with the main difference being the charge carrier collected by the implant. As charge generated in the bulk interacts with the electric field, the carriers are split and travel towards either the junction boundary or the back side of the sensor (which is metalised to allow the application of the bias voltage). Electronics connected to the implant allow the signal to be read out, with the differences in collected carrier mobility producing different signals. The depletion voltages and electric field profile vary with the bulk carrier.

It is possible to take advantage of both carrier types by producing implants on both sides of the sensor; such readout is termed *double-sided.* The p-n junction forms around the p-type implants, extending throughout the bulk until the n^+ implants on the back side. When a charged particle produces electron-hole pairs in the sensor the electrons will drift towards the n-type implant, with the holes going towards the p-type. This method of readout is advantageous for *strip detectors*, where the implants are patterned as long strips which provide 1-dimensional spatial information. Producing sensors with strips on both sides allows a 2D space point to be obtained, by introducing a stereo angle between the strip orientations on either side.

Shockley-Ramo Theorem

While it is conceptually more straightforward to consider the signal due to the carrier arrival at the implant, the signal is in fact generated before the charge is fully collected. The motion of electrons and holes generates an induced current in the implant, given by the change in electric field lines terminating there. The Shockley-Ramo theorem [46] states that this induced current is given by

$$I(t) = q\mathbf{v}(\mathbf{t}) \cdot \mathbf{E_w} \tag{2.44}$$

or, substituting the expression for the velocity v (given by the local electric field) from equation 2.16

$$I(t) = -q\mu \ \mathbf{E}(\mathbf{t}) \cdot \mathbf{E_w} \tag{2.45}$$

where q is the quantity of charge and E_w is the *weighting field*. This is the hypothetical field which would arise if there were no space charges, the charge collecting implant was held at 1 V and all other conducting surfaces in the system were held at ground. The generated current can be both negative and positive, with charge moving away from the implant (such as to a neighbouring implant) inducing a negative signal. This has the result that the time-integrated signal is equivalent to the total charge collected at the implant. As the shaping times in the electronics described below are large even with respect to the full collection time (before irradiation), the time-integrated approximation can be used.

Hybrid Pixel Detectors

One the main devices currently used for high-precision tracking in the inner regions of particle physics experiments is the hybrid pixel detector. This is shown schematically in figure 2.7, and consists of a doped silicon sensor with single-sided implant (as described above), DC coupled to a dedicated readout chip. The implants are patterned in a regular array, with a pitch which can be from hundreds of microns down to less than 50 µm. The chip and sensor are produced separately, decoupling the design process and allowing a great deal of flexibility for both. They are then joined together by the process of *bump-bonding.* This involves the deposition of solder balls on either the chip or sensor side, typically 10 - 15 µm or more in diameter. The two are then brought into physical contact and the assembly heated to the necessary temperature for the solder to melt and form a conductive bond between them. The material used differs between companies, and both lead-tin and indium bonding have been used in recent years by large-scale experiments.

Hybrid pixel detectors have increased in popularity with respect to several other technology types for many reasons. The use of high-resistivity substrates which can be operated with high bias voltages ensures both rapid signal detection and allows

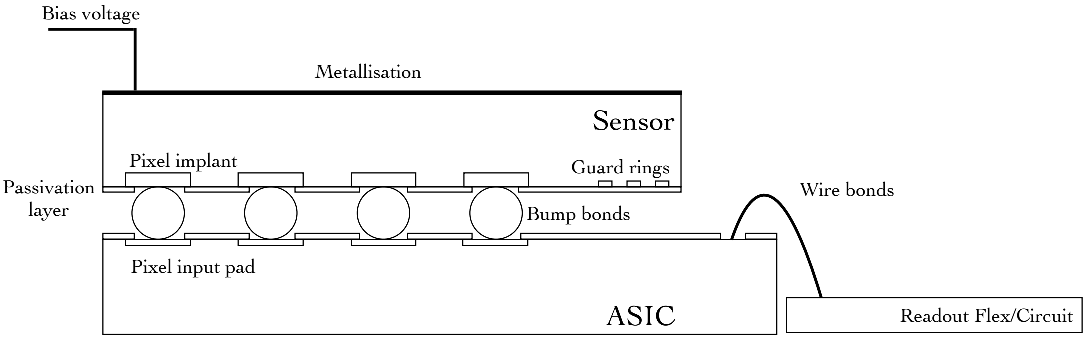

Figure 2.7: Schematic of a hybrid pixel detector.

the device to operate after extreme radiation damage. The granularity and 2D spatial information also gives significant advantages over strip detectors, and so are typically preferred for operation in high multiplicity environments. Even with a significant dead time due to the pulse shape of the electronics, the smaller pixel area means that the cell occupancy is substantially lower. Due to the variation in cost (with hybrid detectors more expensive), strip detectors are still widely used in tracking systems, but typically much further from the interaction point.

2.2 Irradiation Damage in Silicon Detectors

2.2.1 Radiation Damage

Interactions between incoming radiation and a silicon sensor can result in damage of two types, *ionising* and *non-ionising.* Ionising energy loss generally does not result in any permanent bulk effects, but can adversely affect electronics and lead to charge buildup in insulating regions. This is referred to as *surface damage*, and as it is primarily of concern to electronics design (rather than sensor design) is not considered further. The second type of energy loss involves damage to the crystal structure, where a small number of interactions result in the removal of atoms from their lattice positions. For most applications the effects of this are not observed, but for detectors operated inside collider experiments the extremely large flux of particles can damage the silicon sufficiently such as to degrade its performance. This *radiation damage* has been studied extensively within the particle physics community [47].

In order to remove a lattice atom from its natural position, an energy of greater than around 20 eV must be transferred to it during the (typically) electromagnetic interaction of the nucleus and impinging particle. This produces a single set of defects: a *vacancy* and *interstitial* defect. If the imparted energy is high enough, this displaced atom can continue to displace further atoms, forming a *cluster* of defects. Just as for electron-hole pairs these defects can recombine, returning the lattice to an ordered state. They can also interact with impurities within the silicon to create semi-stable or stable defects (Carbon and Oxygen are known to react in such a way). All of these interactions are highly temperature-dependent processes, and can be used to partially recover some of the detector performance. *Annealing* has two components however: the return of displaced atoms to now-vacant sites (*beneficial* or *short term annealing*), and the generation of more stable defects (*reverse annealing*).

Electrically, these defects can act as charge carrier donors or acceptors, with energy levels at intermediate energies within the band gap. This results in three important effects: a significant increase in leakage current; a change in the effective dopant concentration N_{eff}; and *charge trapping*. The first two effects are simply the result of these introduced energy levels, and in this regard is analogous (and supplementary) to the introduction of dopant atoms. The third effect involves deep energy levels located in the centre of the band gap: these states can accept generated charge carriers, such as those produced by an interacting particle, and require a reasonable amount of energy to escape. This results in the trapping of charge, which is slowly released again over time. This time constant varies with the defect type, but can be greater than the integration time of the front-end electronics.

A second consequence of this observed increase in the number of donors and acceptors is the change in effective dopant concentration, N_{eff}. This is used above in the calculation of the depletion voltage, and as such any changes introduced by radiation-damage defects will contribute towards this. Some data for these observations are presented below.

2.2.2 Physical Observables

In order to quantify the effects of differing types of incident radiation, a scaling mechanism is used. This is based on the hypothesis that the bulk damage is due to the energy transferred during displacement of an atom, and that the bulk effects vary linearly with the total energy imparted regardless of the spatial distribution. This is known as the *Non-Ionising Energy Loss* (NIEL) hypothesis. Each particle type can then be assigned a cross-section for energy transfer, based on its total energy, and the results compared to different species of radiation. The units used in literature are usually normalised to an equivalent fluence of 1 MeV neutrons (for a given surface area), the *1 MeV neutron equivalent* (1 MeV $\mathrm{n_{eq}}$ $\mathrm{cm^{-2}}$).

The change in leakage current per unit volume versus total fluence is shown in figure 2.8. The relationship shown is

$$\frac{\Delta I}{v} = \alpha \Phi_{eq} \tag{2.46}$$

where v is the volume, α is an experimentally measured constant and Φ_{eq} is the fluence in 1 MeV neutron equivalents. The relationship is observed to hold well over several orders of magnitude, and is an important tool in predicting the leakage currents for current and future detector systems. It is further observed to be very temperature dependent, with a normalisation factor proportional to the number of thermally-generated carriers (exponential with temperature).

The change in effective dopant concentration is shown in figure 2.9, along with the calculated depletion voltage which would be required for a 300 µm thick sensor. The defects produced by radiation damage are predominantly p-type. For p-type bulk sensors this has the effect of increasing the depletion voltage to several hundred or thousands of volts, particularly for the fluence levels expected for the inner tracking detectors at the LHC (which may reach fluences approaching 10^{16} 1 MeV n_{eq} cm^{-2}). Where the sensor bulk is n-type the introduction of p-type defects has the initial effect of reducing the depletion voltage, and eventually the majority carrier type is reversed, behaving as a p-type material. The point at which this crossover occurs is known as *type inversion.*

Studies of N_{eff} have shown that the behaviour can be described by three terms which describe the change with respect to initial dopant concentrations:

$$\Delta N_{eff} = N_c(\Phi_{eq}) + N_a(\Phi_{eq}, t) + N_y(\Phi_{eq}, t) \tag{2.47}$$

N_c is the *stable damage* term, which represents the stable and irrecoverable displacement of atoms, while N_a and N_y are the short-term and reverse annealing terms. The time shown in this equation is the *annealing time.* Expressions for each of these components can be written, with the stable damage term given by

$$N_c(\Phi_{eq}) = N_{c0}\left(1 - e^{-c\Phi_{eq}}\right) + g_c\Phi_{eq} \tag{2.48}$$

where N_{c0}, c and g_c are constants. The short term (recoverable) damage term is described by

$$N_a(\Phi_{eq}, t) = g_a\Phi_{eq}\, e^{-\frac{t}{\tau_a}} \tag{2.49}$$

which is in fact the sum of several related terms. Only the term with the longest decay constant τ_a is generally considered, with corresponding coefficient g_a. The final damage term is given by the long-term annealing, written as

$$N_y(\Phi_{eq}, t) = g_y\Phi_{eq}\left(1 - \frac{1}{1 + \frac{t}{\tau_y}}\right) \tag{2.50}$$

Again, τ_y and g_y represent the decay constant and constant of proportionality.

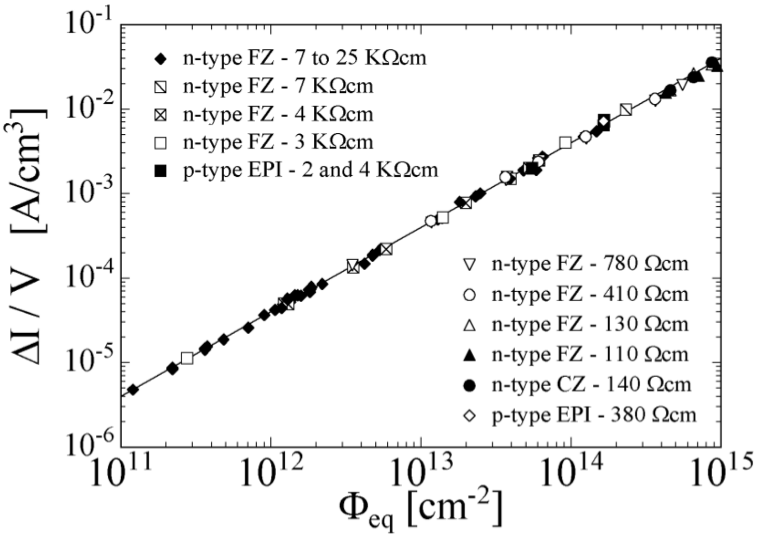

Figure 2.8: Variation in the leakage current per unit volume versus fluence, taken from [47]. Results are shown for a range of sensor thicknesses, resistivities and silicon type (FZ = Float Zone, CZ = Czochralski).

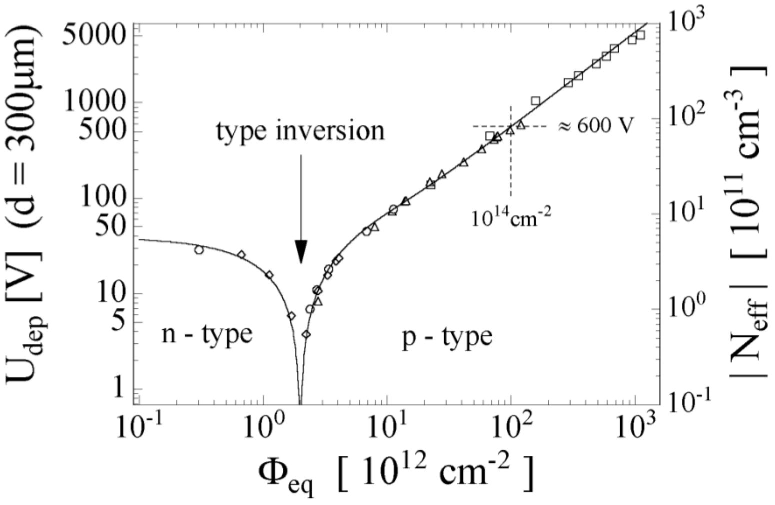

Figure 2.9: Effective dopant concentration versus fluence, taken from [48]. Also shown is the calculated full depletion voltage for a 300 µm thick sensor.

While these expressions allow the effective depletion voltage after irradiation to be calculated, this is only the point at which the applied electric field extends fully throughout the sensor. It does not describe the charge collection efficiency (CCE), which is affected by charge trapping (as mentioned above). Charge trapping gives an observed drop in the CCE after sufficiently high doses, and is due to charge carriers entering energy states produced by the radiation-induced defects. The probability of this occurring is analogous to the recombination of thermally-generated carriers, and depends on the mean free path. This can be reduced by further increasing the bias voltage beyond the effective depletion point, giving a steady rise in CCE after the plateau reached when the device is fully depleted.

The final effect of radiation damage that plays a significant role in detector design is *charge multiplication* [49]. The electric field configuration after heavy irradiation, and operated at large bias voltages (typically 1 kV or more) results in more collected charge than an equivalent sensor before irradiation. This is due to impact ionisation of the generated charge, and the resulting cascade has been observed to be proportional to the input charge (not reaching as far as Geiger mode). The noise rate does not change significantly in this regime, giving an enhanced signal to noise ratio, with further voltage increases ultimately limited by micro-discharges within the material.

It is worth highlighting again that all of these effects show a dependence on the contaminants found within the bulk silicon; in particular oxygen inclusion has been shown to significantly change the damage factors and result in an improved performance after equivalent fluence.

3

The Timepix Telescope

In order to study the precision of tracking detectors, and more generally for the characterisation of any type of detector, it is advantageous to use a probe where the incoming trajectory is well known. Several variations are employed for this function, depending on the study: collimated sources, micro-focussed x-ray beams and high energy particle beams. For the case of a non-focussed particle beam, a set of tracking stations are generally used in order to provide the probe information on a particle by particle basis. Such a system is generically called a telescope.

The LHCb VELO Timepix telescope [50] was designed to fulfil two functions. The first was the validation of the Timepix chip [30] for charged particle tracking, in anticipation of an upgraded VELO based on an evolution of the chip. The second was to provide LHCb with a testing setup for the particle beams provided at the SPS North Area. The telescope was required to operate at high rate and consistent with the known beam structure, and to accommodate testing of LHC-style devices (those synchronous to a 25 ns clock). The project was later incorporated into the AIDA project [51] and integration of external devices became an important consideration.

3.1 The Timepix ASIC

3.1.1 Chip Overview

The Timepix chip is a hybrid pixel chip designed and fabricated in 0.25 µm CMOS technology, containing a matrix of 256 rows × 256 columns with a pixel size of 55 × 55 µm. The pixel cell contains both analogue and digital circuitry, shown in detail in figure 3.1. The analogue side of the pixel contains a charge sensitive amplifier (CSA) with a tuneable discharge current I_{krum} and discriminator (Disc), with charge supplied either through the input pad or by means of an injection pulse through the capacitor *Ctest*. The front end can accept either positive or negative input charge, indicated by *Polarity*, and contains a 4-bit threshold adjustment.

Manufacturing imperfections in semiconductor electronics generally lead to small variations in performance both between devices fabricated on the same silicon wafer, and within the device itself. The global threshold set on the Timepix discriminator is the most significant example of this in relation to the chip operation. While a single value for the threshold is applied, local variations distort the true voltage observed on a pixel by pixel basis. This leads to a smearing of the threshold across the matrix, giving large levels of non-uniformity in the response. In order to combat this, a 4-bit local

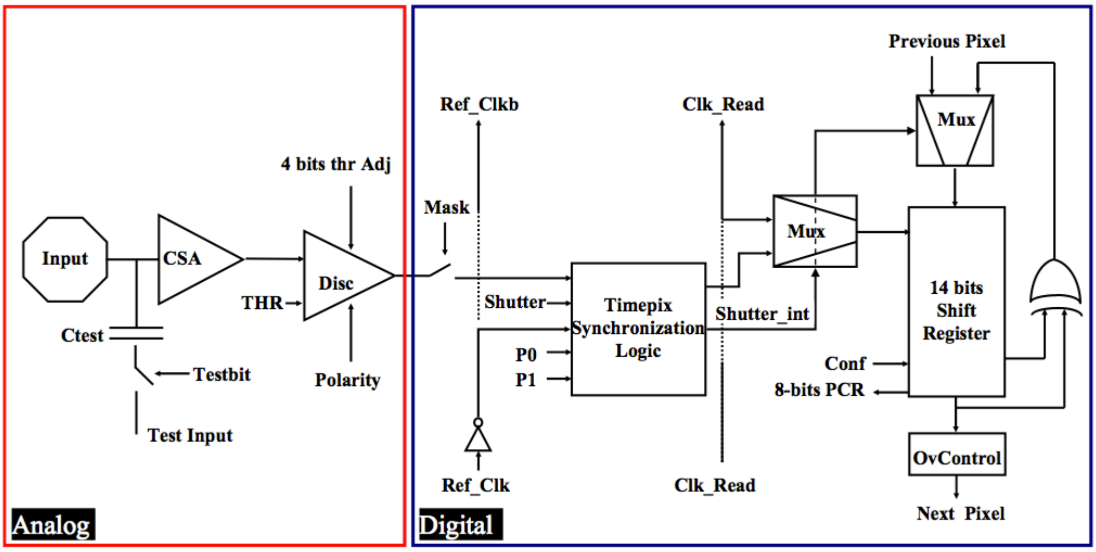

Figure 3.1: Schematic of the Timepix pixel cell, showing the analogue and digital circuitry. The analogue side contains a charge sensitive amplifier (CSA) and discriminator (Disc), while the digital side contains the shutter and counting logic (TSL) and a 14-bit shift register.

adjustment is applied to each pixel, to trim or augment the local value of the threshold voltage to match the global average. An *equalisation* of the front end is performed prior to use, in order to determine an optimal setting for the 4-bit adjustment on each pixel.

An illustration of the threshold equalisation steps, showing the dispersion at each step, is shown in figure 3.2. In the first step, the 4 adjustment bits are set to 0000. The global threshold DAC, *THL*, is scanned and the centre point of the pixel response is plotted (black in the figure). This is then repeated for an adjustment mask of 1111 (red in the figure). The range for each pixel is therefore known, and a linear relationship between the two end points is assumed. Each pixel is then assigned a bit-mask that places it as close as possible to the global mean threshold value. Some dispersion remains in the equalised chip (blue in the figure) but the response is significantly more uniform than for the raw device.

The digital side of the front end (figure 3.1) contains the Timepix Synchronisation Logic, which sets the operating mode and conditions for the counter incrementation. The 14-bit counter has a maximum useable range of 11,810 counts. The digital side also contains the readout logic, and a 14-bit shift register which transmits the counter value down the column during readout. The entire chip is controlled via a global shutter, which defines the active and inactive periods: while active, the whole pixel matrix is continuously sensitive. During readout, the digital side of the pixel does not respond to new events, and the chip remains insensitive. The most common readout systems use the serial LVDS link to transmit the data off-chip, giving dead times of roughly 9 ms, though a parallel CMOS bus readout is implemented with a dead time of 300 µs.

The chip can be operated in three distinct modes: Counting (Medipix), Time of Arrival (ToA) and Time over Threshold (ToT). In Medipix mode, the counter is incremented each time that the amplifier output exceeds the set threshold. This is uni-

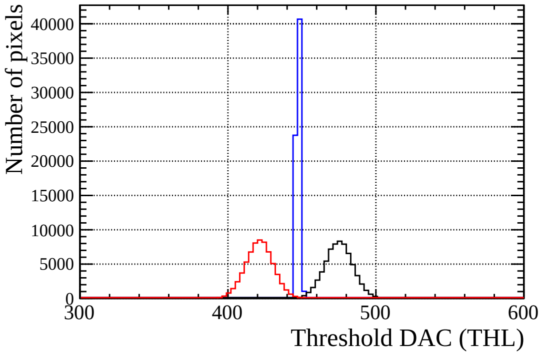

Figure 3.2: Spread of pixel responses during the front end equalisation, for mask 0000 (black), 1111 (red) and after equalisation (blue).

directional only, and it is the crossing from below- to over-threshold that causes the circuit to react; subsequent hits which occur before the discharge of the front end will not be reflected in the counter reading. Given the maximum frame rate of around 120 Hz (non-parallel readout) this allows the chip to register up to 1.4 MHits per pixel per second, and is typically used for imaging purposes. In ToA mode, the counter is again triggered by the amplifier output exceeding threshold, but in this case the pixel continues to count until the shutter closure, incrementing with each period of the clock. As the counter depth is 11,810, any hits which occur (11,810 × clock period) before the shutter closing will contain overflow values, and will no longer give a meaningful reading. The final mode of operation, ToT, allows for a measurement of the charge deposited in the pixel. As the front end is linearly discharged, a measurement of how long the amplifier output remains over threshold is approximately linearly proportional to the deposited charge, with the discharge current as the constant of proportionality. The counter, while the pixel is in ToT mode, is thus incremented with each clock period where the amplifier output is above threshold. As with Medipix mode, the pixel remains active for the full period that the shutter remains open, and therefore subsequent hits occurring during the same shutter will contribute to the ToT measurement.

In order to connect the Timepix to a silicon sensor, bump bonding is required. The pad size on the Timepix is 15 µm × 15 µm.

3.1.2 Charge Discrimination

As mentioned previously, the charge measurement on the Timepix chip is carried out by using a linear discharge current to relate the injected charge to the time that the CSA output spends over threshold. This is measured at discrete time intervals, with the period set by the measurement clock (either generated internally to the readout system or supplied externally). There are two primary sources of error in this measurement: the discretisation of the charge and the non-linearity of the circuit close to threshold. Due to the large range of the 14-bit counter, the former effect could be expected to be minimal. However, limitations to the clock frequency (less than 80 MHz) and a desire for the charge measurement time to be small compared to the shutter opening time, lead to discretisation errors of roughly 2 % (with the peak of a MIP occurring at around 100 ToT counts).

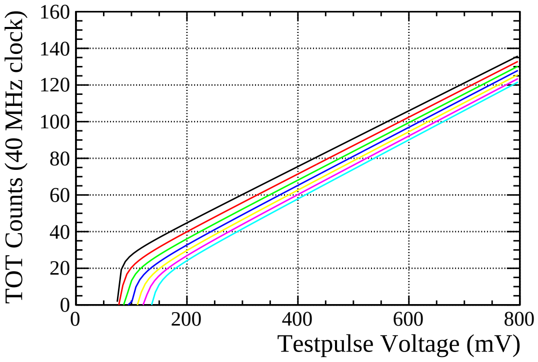

Figure 3.3: Calibration curves for a single chip operated at several threshold values. The lowest threshold shown is 1000 e$^-$, incrementing in steps of 500 e$^-$ each time.

For the non-linearity close to threshold, charge calibrations are used to reduce this as much as possible. The extent of this can be seen in figure 3.3, where the charge-to-ToT conversion is shown for several threshold values - the nominal operating threshold of the chips is 1000 e$^-$. The main chip parameters that control these effects are the clock frequency, the discharge current I_{krum}, the threshold and the preamplifier bias current (all of the parameters that alter the shape of the CSA output). The rise time of the CSA output is 100 - 150 ns, decreasing as more power is supplied.

3.1.3 Timing Effects

Another issue which is important for the chip operation is the timing information afforded by the ToA mode. In principle the primary limiting factor of the device response should be the clock granularity, however for charge deposits close to threshold there is an additional delay due to the response of the front end. Figure 3.4 shows the delay time on the front end response versus injected charge, for a threshold of 1000 e$^-$. The *timewalk* of the front end, defined as the difference in response time for an injected charge 1000 e$^-$ over threshold and an infinite charge injection, becomes particularly important for later generations of the chip. Note that the values in figure 3.4 are also relative to infinite charge, and not absolute response times (where the

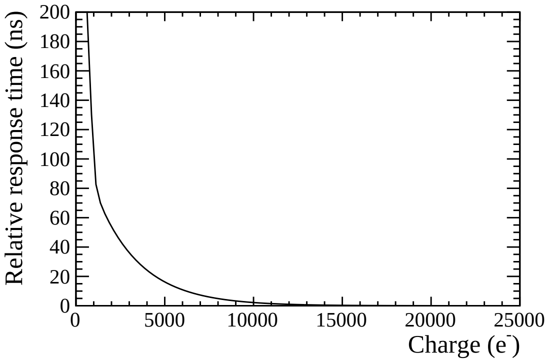

Figure 3.4: Response time relative to infinite injected charge.

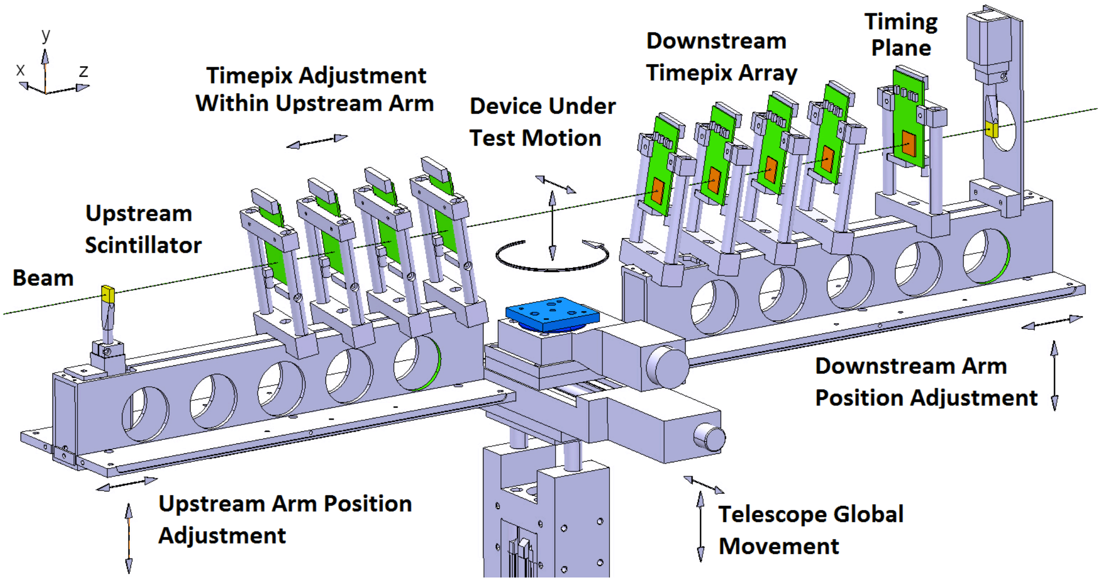

Figure 3.5: Layout of the Timepix telescope. Figure taken from [52].

response of the circuit to very high input charge still requires a fixed time to react).

3.2 Telescope Description

3.2.1 Mechanics

The telescope is composed of two separately mounted arms, which can be seen schematically in figure 3.5 (taken from [52]). The co-ordinate system is indicated, and is right-handed with the z-axis directed along the particle trajectory. On the upstream arm sit four Timepix planes operating in ToT mode and a single scintillator. On the downstream arm there are five Timepix planes, four of which operate in ToT mode with the other in ToA mode, and a second (inverted) scintillator. When mounted, the scintillators are aligned such that their active regions overlap within the telescope acceptance. All of the telescope planes operating in ToT mode are mounted with a 9° rotation around both the x and y axes, in order to increase the amount of charge sharing (approximately the angle at which all tracks must traverse at least two pixels). The ToA plane is mounted perpendicular to the beam, in order to reduce charge sharing and minimise the effects of low-charge-induced timewalk.

Many motion stages are employed to provide as flexible a system as possible for mounting devices and operation during beam time:

- Mounting rails allow manual z movement of the individual arms (with a maximum clearance of 46 cm).
- A linear screw system allows elevation of the individual arms of up to 20 cm.
- A global x-y table allows full setup to be remotely positioned within the beam.

- Remote operation of several motion stages is available for the Device Under Test (DUT), giving precision movement in the x and y directions and rotation around the y axis to an accuracy of a hundredth of a degree.

Both telescope arms are additionally housed within an enclosure that is flushed with dry air, to prevent the buildup of moisture and provide some modest cooling of the Timepix ASICs. Mounted on the side of these enclosures are the readout boards for the telescope. Low and high voltages are provided to the telescope from two racks, mounted in the beam area.

3.2.2 Data Taking

The principle behind the telescope layout is that the two arms of ToT detectors allow accurate track reconstruction, the ToA plane at the rear attaches a coarse timestamp (accuracy determined by the clock frequency), and that this timestamp can be associated to a hit from the coincident scintillators in order to give a significant improvement in resolution. In order to increase the total data taking rate, and largely due to the readout architecture of the chip (with a shutter based operation, and relatively long readout period), several tracks are accumulated during a single frame. The data streams for all devices are kept deliberately separate, in order to reduce the complexity of the setup and to more easily facilitate the integration of external devices.

The telescope planes are read out using a portable FPGA-based readout system, the RELAXd [53]. This allows the data taking periods to be synchronised using an externally generated shutter signal, based on the coincident scintillator signals. These are used to produce a beam present signal, which is vetoed by a busy from the RELAXd which persists during the system readout. When the readout is complete, a new shutter is generated, with this process repeating for as long as the beam present remains high. The total dead time during this readout period is roughly 8 ms, giving a frame rate of order 120 Hz (when the shutter opening time is small). As the granularity of the chips is very high, track rates far in excess of this frame rate can be written to disk. While the shutter is open, the number of tracks allowed to pass through the telescope is controlled by counting the number of coincident scintillator hits: once the required number of tracks have traversed the system, the shutter is closed and the chips read out. In addition (if timing information is required), a timer can be used to close the telescope after the maximum acceptable shutter length has been reached, in order to prevent overflow in the ToA plane. Table 3.1 shows the shutter length limits based on the clock frequency on the ToA, for those frequencies available by default in the

Clock Frequency (MHz)	Clock Period (ns)	Maximum Shutter Length (ms)
80	12.5	0.148
40	25	0.295
10	100	1.181
2.5	400	4.724

Table 3.1: Limitations on the shutter length for different clock frequencies.

readout hardware. Other values for the clock frequency are possible but must be supplied externally to the RELAXd. These two closing criteria allow complete control over both the occupancy and maximum shutter length.

A high precision TDC (time-to-digital convertor) built on VME is used to record the times of each shutter and both the raw and synchronous triggers. The system has a precision of around 1 ns.

3.2.3 Control Systems and Monitoring

As the DAQ of the various telescope components are kept separate, they each retain their individual control software and must be operated independently. This provided for some initial monitoring during data taking, though was restricted to a hitmap in the RELAXd DAQ software, buffer status for the TDC readout, and the values read back from the positioning, low-voltage and high-voltage systems. As the data rate of the system increased with the installation of RELAXd boards to replace the previous hardware, data management also became an issue.

In order to improve the operation and monitoring of the telescope, a run manager was developed based on the existing VELO monitoring GUI. This performed three main tasks: to write the run information into a log file on a run-by-run basis; to organise the raw data into run folders and make sure all of the data for each run was kept together (along with any preprocessing required); and to allow fast offline analysis of the data taken and display the results. As a full online monitoring system would have required integration with the DAQ software, it was decided to instead allow the run manager to launch instances of the offline software over a subset of the run data. Using this approach it was possible to have the full offline reconstruction performed within a minute or two of the first data being taken, and to quickly rectify any issues in the data taking.

3.3 Analysis Software

For the analysis of data taken with the Timepix telescope, a standalone reconstruction and analysis package has been written. As with the hardware design, this has been built with as few dependencies as possible, in order to simplify its use and make it transferable across platforms. The software is written in C++ and requires ROOT [54] and BOOST [55] C++ libraries.

3.3.1 Architecture and Data Flow

The conceptual design of the software is similar to that of the software framework used by both LHCb and the ATLAS collaboration (Gaudi [56]) though less complex. It revolves around the idea that the data can be split into individual events, over which sequential algorithms are applied. Information and objects can be deposited or retrieved from a central “Clipboard”, which is purged of content at the end of each event. Using a light framework, it is simple to add an individual algorithm into the program with very little overhead.

The software is effectively split into two separate operations. It first amalgamates the selection of raw data to produce an ordered tree, where the event format is fixed, and then acts as the reconstruction and analysis tool (where the output from the amalgamation acts as the input file). Each event in the testbeam software is a single Timepix frame, and the amalgamation contains the data decoders for all of the input files. As the data from each device is timestamped, and the number of frames should be identical for any given spill period (where the SPS spill lasts for 10 s, followed by 40 s without beam), it is straightforward to place the data from each device into a single event object. The case for external (non-Timepix) devices is subtly different, and will be discussed in section 3.5.

The analysis step sees several algorithms run sequentially, on a frame-by-frame basis, on the structured Tree. Each algorithm has access to the Clipboard, and will typically deposit there any output objects which are of use in subsequent steps (such as clusters to be picked up by the track reconstruction). A list of default global parameters is supplied via a single C-file, with all run specific information contained in a configuration file. This configuration file allows the user to specify which algorithms are run on the data, run-specific information such as the device ID of the DUT, and any specific requirements on the telescope reconstruction (number of planes required for each track, etc.). A typical analysis sequence would consist of: clustering, pattern recognition, track fitting, time stamping and some algorithm to analyse the DUT.

3.3.2 Track Reconstruction

The pattern recognition algorithm takes clusters from the telescope planes as input, which are produced by a separate algorithm and use the charge centre-of-gravity (COG) as the cluster position. The most upstream detector is used as the seeding plane, and a progressive fit and search is performed through the telescope. In the first instance the seed cluster on the first plane is extrapolated directly along the z axis to the next telescope plane. If a cluster is found within a given search window then it is added to this first cluster and a track is created between these two space points. This track is extrapolated to the next tracking plane, where a search is again carried out for clusters close to the projected intercept. If any are found, then the three points are fitted with a straight line and the extrapolation is continued until the required number of clusters are added to the track.

The search for clusters close to the projected track intercept is performed using a fast nearest-neighbour algorithm [57]. The search window can be defined manually, but is typically taken as the expected scattering due to material interactions (Multiple Scattering - MS) and takes into account the material content and distance between the tracking planes. The formula used for the scattering was defined in equation 2.39, and the estimated material content of a single plane is shown in table 3.2.

As expected, this method predominantly finds straight or shallow angled tracks. The maximum number of tracks that can be taken is also dependent on the occupancy, as track overlap becomes more frequent with an increasing number of tracks per frame. To reduce the number of ghost tracks created, or more to control the track quality where unrelated hits on the telescope may have been incorporated into tracks, a cut on the track χ^2 is performed.

	Material	Thickness (um)	x/X_0 (%)
Sensor	Si	300	0.32
Bump Bonds	Pb	0.6	0.01
ASIC	Si	700	0.75
PCB	Cu	100	0.70
	FR4	1300	0.82
Total			2.60

Table 3.2: Composition of an individual Timepix plane. The thickness used for the bump-bond contribution equates to 15 µm diameter bumps.

3.3.3 Time Stamping

Time stamping of the telescope tracks involves assigning a cluster on the ToA plane to a track (via the spatial pattern recognition), and further matching this coarse timing measurement to the high resolution information from the TDC. As the ToA is the time between the particle arrival and the shutter closing, the shutter length must be taken from the TDC measurement for each frame, and used to convert the ToA time into a true time stamp. A series of vetoes are used before the association of ToA to TDC times. An isolation veto on the ToA data is used to prevent cases where two hits occur close in time but where the first hit suffers from timewalk due to low charge deposited. This is usually taken as either 200 ns or three clock cycles (whichever is higher) and additionally protects against hits with very small timewalk but where the track passed close to the clock edge. Similarly, an isolation veto is applied to TDC hits, to ensure that where multiple hits are close together they are not misassigned to tracks. These conditions result in a clear association of the times, and allow the nanosecond timestamp to be attached to tracks.

As mentioned above, the limited counter size on the chip restricts the length of time for which the shutter can remain open, and thus there is a strong relation to be considered between beam intensity, clock granularity, shutter length and data rate. The clock must always be of sufficient granularity to distinguish the time between raw tracks, yet even with infinite counter depth will be limited by the varying timing response.

3.4 Performance

All planes in the telescope are equipped with 300 µm p-on-n silicon sensors, with a full depletion voltage of roughly 10 V, and are operated over-depleted (40 V). The beam momentum used is 180 GeV/c, consisting of mostly pions though with some sizeable muon contamination. All planes are clocked at a frequency of 40 MHz.

The performance of the telescope is characterised by two main criteria; the reconstruction efficiency and the precision of the track extrapolation at the DUT. The reconstruction efficiency is also tied to the overall track rate, due to the buildup of tracks during a single shutter.

3.4.1 Reconstruction

During the pattern recognition track, segments are continuously fitted and extrapolated in order to find hits on subsequent planes. The difference between the projected intercept point and the nearest cluster are shown in figure 3.6 for all of the telescope planes, with the telescope z positions shown in table 3.3 (note that there are no residuals for the first two planes, which are required to make the first fit and are the result of a simple extrapolation of the seed cluster along the z axis). During the extrapolation across the large gap between the two arms the distribution broadens significantly. This might at first be considered the result of scattering in the plane directly before the gap, yet the lateral displacement is significantly above the calculated $\sigma = 3.0\,\mu\text{m}$ (from equation 2.39). Considering the track at the point of extrapolation, an error in the track angle θ_x will come from the track gradient determination with (in this case) four position measurements. Projecting this error across the distance between tracking arms accounts for the observed performance in the first plane of the second arm. Once the track is complete and a linear fit is applied, the residual distributions improve significantly. These fitted (unbiased) residuals can be seen in figure 3.7 for the same set of tracks as figure 3.6; as expected there is no considerable difference in performance between the planes.

In order to avoid the observed impact of the prototrack gradient on the cluster search, a second algorithm was developed, which uses only an extrapolation along the z-axis. Each newly added cluster is directly projected along the z-axis onto the next telescope plane, and a search for the closest cluster is performed. This gives improvements upon the fit and extrapolate method - the error from one arm to the other is now dominated by the cluster position error on the last plane before the gap. This can be seen in figure 3.8. The error on the cluster search for the majority of planes is now equivalent to those in the final stages of the extrapolation method. Furthermore, the error in crossing the gap decreases from 33.1 µm to 22.3 µm.

The evaluation of the overall tracking efficiency of the telescope is non-trivial. Problems include the determination of the total overlap area of the individual tracking planes, their overlap with the scintillators positioned at both ends of the telescope, the scintillator single hit efficiency and timing effects due to hits arriving close to the

Chip ID	z position (mm)
C09-W0108	0
C10-W0108	23
J08-W0087	46
F11-W0108	69
D09-W0108	352
H03-W0092	375
G08-W0087	398
J03-W0089	421

Table 3.3: Z positions for the telescope planes shown in figures 3.6 and 3.7. The z-positions are recorded at the centre of the pixel matrix and do not play a large role in the track fit due to the relative insensitivity to translations along the track axis given by the limited angular range of tracks.

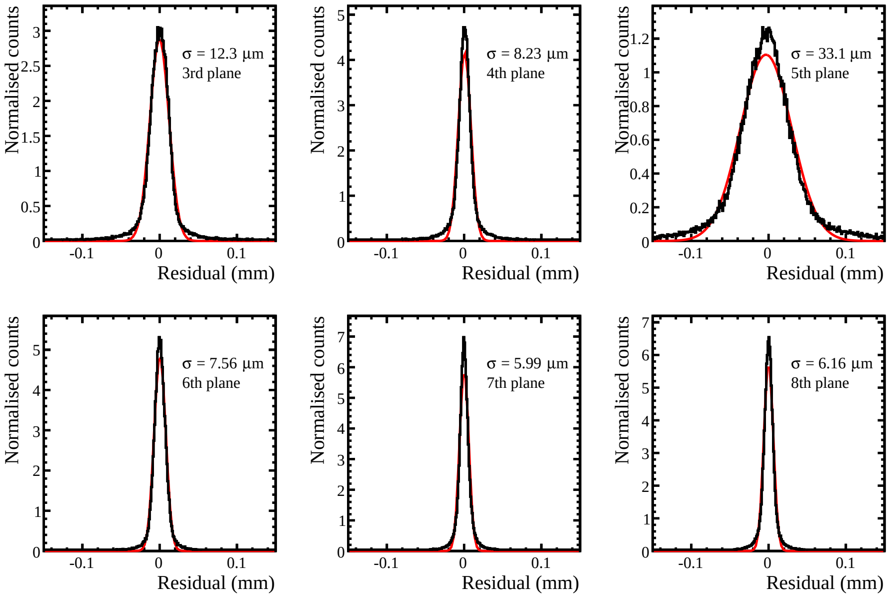

Figure 3.6: Projected residuals between the extrapolated intercept and closest cluster during track reconstruction, using the continual fit and extrapolation method. The plots are shown in order of z-position, with top left being upstream and bottom right being downstream.

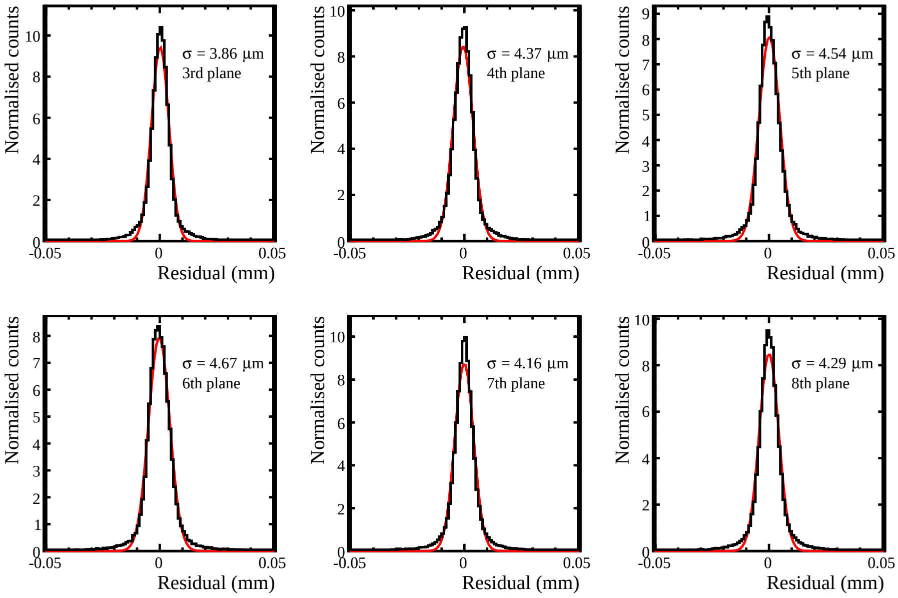

Figure 3.7: Final unbiased residuals for fitted tracks. The plots are shown in order of z-position, with top left being upstream and bottom right being downstream.

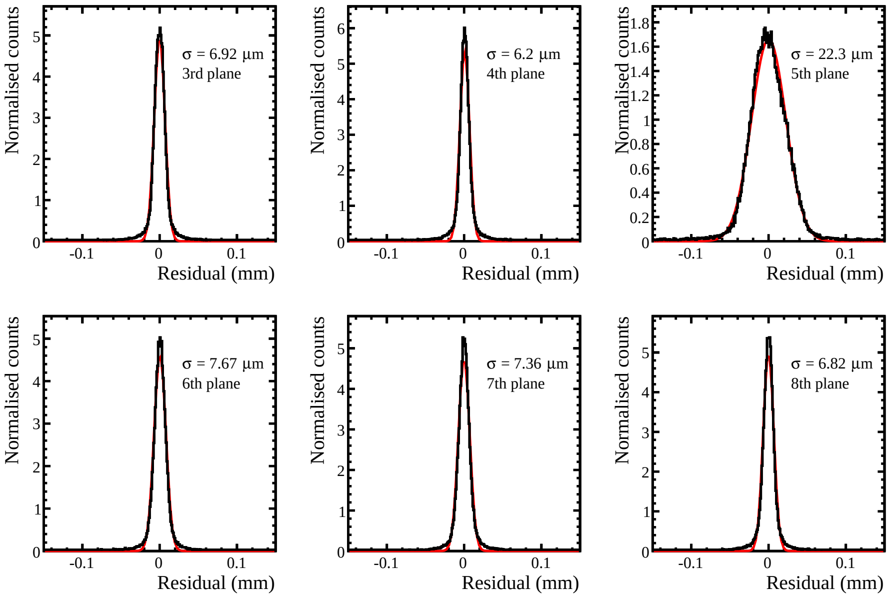

Figure 3.8: Projected residuals between the extrapolated intercept and closest cluster during track reconstruction, using the z-axis projection method. The plots are shown in order of z-position, with top left being upstream and bottom right being downstream.

shutter opening and closing. As the absolute efficiency is not in fact a consideration for the testing of devices, it is more useful to look at the absolute rate of tracks that can be written to disk. As described above, the frame rate with the RELAXd readout is around 120 Hz, and so the absolute track rate will just be the product of this with the number of tracks per frame. However, as the occupancy of the system increases the track reconstruction efficiency drops due to the total pixel occupancy of the chips, and this relationship diverges.

The reconstructed track rate as a function of the number of coincident scintillator hits is shown in figure 3.9. This is the raw number of tracks reconstructed, before the track is fitted and requiring only that there is a hit on every plane of the telescope (8 in total). The figure shows the results both for the continual fit and extrapolation method, and the direct z-axis projection. As expected from the increased error in the fitted method, fewer tracks are reconstructed. In both cases it is clear that there is an inefficiency in the tracking, and that this increases with the chip occupancy. This can be split into two components: a fixed inefficiency where the track scatters sharply and/or passes outside the telescope acceptance, and an inefficiency related to the occupancy of the pixel matrix. The efficiency ϵ then takes the form

$$\epsilon = 1 - \left(i_{fixed} + a \times \frac{4\ n_{tracks}}{256 \times 256} \right) \tag{3.1}$$

where i_{fixed} is the irreducible probability that a track will not be reconstructed, 256×256 is the pixel matrix size, $4\ n_{tracks}$ is the occupancy where the average cluster size is taken to be 4, and a is the constant of proportionality. Using this expression for the reconstruction efficiency versus number of tracks per frame (n_{tracks}), a fit was performed to the data. The fit results are shown in table 3.4, and give a reconstruction efficiency of

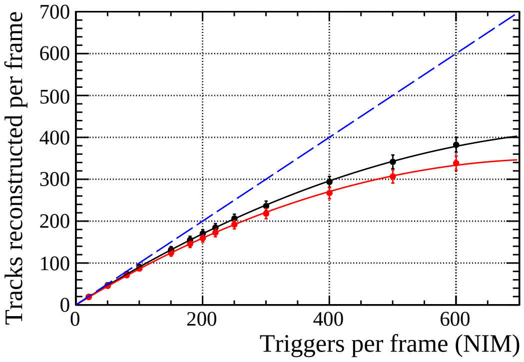

Figure 3.9: Number of reconstructed tracks versus the expected number from the shutter closing requirement, for the z-axis projection (black) and fit and extrapolate (red) methods. The fit results for equation 3.1 are shown, and the blue line represents the expected number of tracks.

$(95.79 - 0.05 \times n_{tracks})$ % for the z-axis projection method and $(91.79 - 0.06 \times n_{tracks})$ % for the fit and extrapolate method.

Figure 3.9 shows a maximum of 382 tracks reconstructed, taken with 600 tracks per frame set in the trigger logic. This gives a total data taking rate of 45.8 kHz of tracks versus the 2.3 kHz achieved at very low occupancy. Due to the corresponding increase in CPU time required to reconstruct the busier events, frames were generally taken with around 100 tracks each, for a mean track rate of 10.9 kHz. Given the typical spill length of 10 s in the NA Super Cycle this corresponded to 100k tracks each spill, and allowed a huge amount of data to be taken during beam periods.

	Fixed Inefficiency (%)	Occupancy constant (a)
Linear	4.21	8.93
Fitted	8.21	9.89

Table 3.4: Fit results for the fit and extrapolate and z-axis projection reconstruction methods (using equation 3.1). The fixed inefficiency term and the occupancy constant are shown.

The results shown for track reconstruction are obtained after alignment of the full set of telescope planes. An initial measurement of simple translational offsets in the x- and y-directions is obtained by comparing the global cluster positions of each plane with that of an arbitrarily chosen reference plane within the telescope. Once these translations are taken into account tracks can be produced, and a more detailed alignment can be performed by varying the alignment parameters for each plane and minimising the unbiased residuals with respect to the reconstructed track.

3.4.2 Pointing Resolution

The pointing precision of the telescope is estimated using a toy Monte-Carlo simulation [50]. The single hit resolution of the individual planes is left as a free parameter, and tracks are simulated with multiple scattering in each plane (using the material content

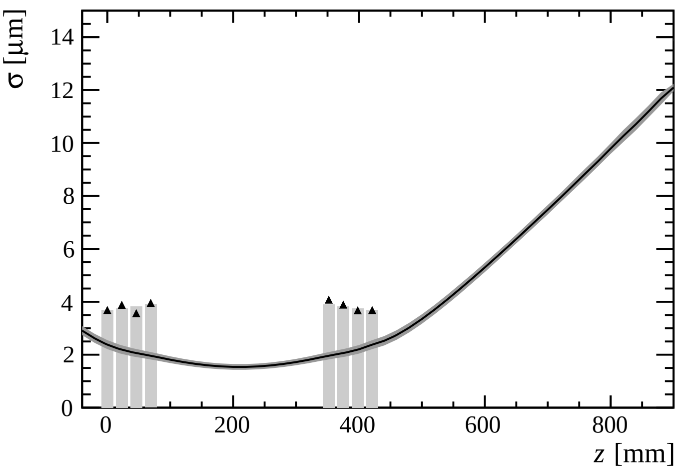

Figure 3.10: Simulated pointing resolution of the telescopes (fitted curve, with error). The telescope residuals measured are shown as points, and the fitted simulation residual values are given by the histogram bars.

in table 3.2). The distribution of telescope residuals obtained in the simulation can then be fitted to the experimental residuals to obtain the single hit resolution. The fitted simulation tracks are then used to determine the expected pointing resolution at a given point, by comparing the true and fitted track extrapolation. An example of this is shown in figure 3.10, where a resolution of 1.54 ± 0.1 μm was achieved. This value depends strongly on the conditions during data taking, in particular the distance between the telescope arms and the beam energy. This latter condition is the most significant, as the telescope resolution is limited by the high material content of the individual planes. Nonetheless, the typical resolution of the telescope is still below 2 μm during operations in the 180 GeV/c beam in the SPS North Area.

A simple step to improve the telescope performance would be the removal of material from the active overlap region. Thinned chips, and more critically removing the PCB from behind the chip on the chipboard, would allow significant improvements in the minimum achievable resolution, and in the performance with lower beam energies. This step is foreseen for the next evolution of the telescope, with the Timepix3 ASIC [33]. The resolution achieved has been invaluable in the detailed studies conducted for the VELO Upgrade.

3.4.3 Time Stamping

Inefficiencies in the time stamping can arise from four sources:

- Overflow of the ToA counter.
- Misalignment of the scintillators with the telescope.
- Tracks too close in time, resulting in ambiguity.
- Spatial tracking associates the wrong hit on the ToA plane to a track.

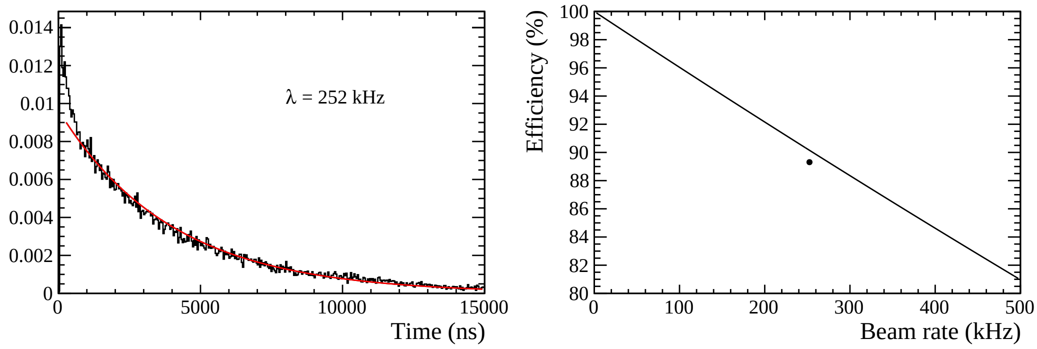

Figure 3.11: Left: Time between sequential hits in the scintillators. The beam rate is extracted from the exponential fit. Right: Expected time stamping efficiency versus raw beam rate (from equation 3.3), for a 200 ns isolation veto (line) and the measured efficiency of the telescope.

The first two inefficiencies can be effectively mitigated by control of the maximum shutter length, and by manual alignment of the scintillators before taking beam. The issue of track separation is always present however, and depends directly on the raw beam intensity. As the particle arrival in the telescope follows Poisson statistics, the probability that two tracks occur within δ_t ns follows an exponential relationship,

$$P\left(\delta_t\right) = 1 - e^{-\lambda\delta_t} \tag{3.2}$$

where the probability of two events occurring in time interval δ_t $\left(P\left(\delta_t\right)\right)$ is controlled by the raw beam intensity λ. Figure 3.11 shows the raw distribution of trigger separations for the telescope runs under consideration, and the extracted beam intensity. Taking the integral over the veto period (200 ns), and taking into account that both tracks would be vetoed, the expected efficiency ϵ should then be given by

$$\epsilon = 2 \times e^{-200\times10^{-9}\lambda} - 1 \tag{3.3}$$

for a raw beam intensity of λ. This is plotted in the right hand side of figure 3.11, along with the measured time stamping efficiency of the telescope during the low occupancy run (20 tracks per frame).

Once the scintillator events are isolated in time, and the telescope tracks are located in the overlap region between the telescope and the scintillators, the events are extremely clean. Figure 3.12 shows a typical event, along with the timing residuals for the Timepix ToA measurement. The maximum time difference is observed not to exceed 200 ns.

A priori, the efficiency is not expected to change significantly with the frame occupancy (as long as the shutter is not open long enough as to overflow the ToA counters). However, the spatial resolution of the ToA plane is less than for the other telescope planes, and any errors in the track reconstruction will result in an inability to associate the ToA time to a scintillator hit. The efficiency of the time stamping algorithm versus frame occupancy is detailed in table 3.5. No observable drop in efficiency is observed for track occupancies between 20 and 200 tracks per frame.

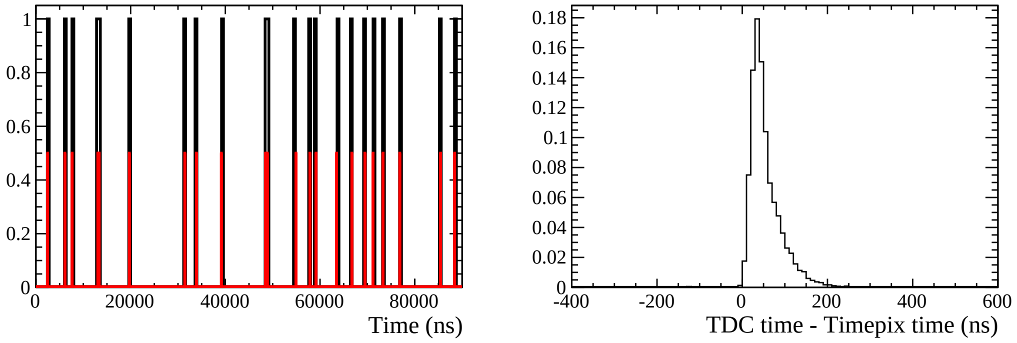

Figure 3.12: Left: Typical event showing the matching of TDC times (red) to Timepix ToA times (black). Right: Time difference between scintillator hits and track times, with only the isolation conditions applied.

Tracks per frame	Timestamping Efficiency (%)
20	89.31 ± 0.13
50	90.03 ± 0.09
100	89.81 ± 0.08
200	89.45 ± 0.28

Table 3.5: Time stamping efficiency for different frame occupancies, with a raw beam rate of 252 kHz. The expected efficiency is 90.16 %.

3.5 External Device Integration

An approach to the integration of non-Timepix devices was required both within the VELO upgrade project (due to the prototyping of strip detectors) and within the LHCb collaboration as a whole. Additionally, the involvement of LHCb in the AIDA European Project has led to the use of the telescope by users outside of the collaboration. As with the telescope itself, the concept of minimal hardware integration and communication between the systems was continued, in order to minimise the number of possible problems that could occur during beam time, and to limit the effort required before data taking.

The exact process for integrating with the telescope is dependent upon the readout architecture of the external front end. A typical case is the use of devices intended for the LHC - namely those which take a single trigger for each event, and may or may not require this to be synchronous to a 25 ns clock. In this case communication between the systems is limited to a single trigger line, which may or may not be synchronised, and a veto signal which may optionally be used to prevent overloading of the external DUT readout system. The DUT must provide a timestamp for the data read out by each trigger, with an accuracy of roughly 100 ns or better. No other interaction is required, and the data streams of both systems are combined offline by matching the timestamped telescope tracks with the data associated to that time from the DUT. Matching these together is a relatively straightforward process, given the uniqueness of the trigger pattern in time within a frame (making each telescope frame appear to have a “barcode”).

The matching of external device data may be performed within the existing software framework, or externally using custom code. Regardless of the route followed, once the code has been written the data can be reconstructed during beam time in order to have a fast data-quality check (typically 2 or 3 minutes delay).

The option requiring least integration into the Timepix telescope is evidently to work outside the dedicated software. In this case, the telescope software can be run by simply changing the configuration file, and will write out an ntuple of time stamped tracks which can be used for the analysis. Times attached to these tracks can be used to identify which spill and frame the track belongs to, and the time matching can be performed by the external group without any further interaction with the telescope.

Alternatively, the DUT data can be integrated into the amalgamation. This is performed by the addition of the relevant data decoder, and the time matching sequence between the DUT and TDC data. As this is added in the amalgamation stage of the data processing, new algorithms can be written into the software package to allow access and manipulation of the data during the analysis stage, and allows the user to take advantage of the existing framework (and algorithms such as the alignment).

To date, several devices have been integrated with the telescope, and users have been from both within the LHCb upgrade community (covering the TORCH [28], SciFi [26] and Calorimeter [27] upgrade groups) and external to the collaboration (ATLAS Planar Pixel Sensor and Diamond Beam Monitor groups, the CLiC vertex detector group, CERN RD50 [58] and the Medipix collaboration [59]).

4

Sensor Resolution Studies using the Timepix Front End Chip

The impact parameter resolution of the current VELO has been characterised extensively [60]. The contributing terms are: the radial distance of the first measured hit, the material content of the VELO and the single hit resolution. The first two terms are constrained by the geometry and module layout and are discussed in later sections. The final term, the single hit resolution, is of interest in both the sensor and ASIC design. Thicker sensors have higher charge deposits, improving both the signal-to-noise ratio of the detectors and the single hit resolution beyond that which is available from the device pitch, due to an increase in charge sharing. Electronics available at present are typically accurate enough to give negligible errors in the charge digitisation, but may have to sacrifice this in order to meet other operational requirements.

In general terms, the direction of the particle physics community in silicon detectors has moved in the opposite direction, leading to thinner active regions and minimising the material content. This has led to the continued use of large area strip detectors (beside cost considerations), in addition to developments in both sensor and ASIC thinning. Broadly speaking, the effects of multiple scattering at most experiments outweigh the considerations for improved single hit resolution, and so increasingly smaller pixel size has been employed with thinner substrates. At hadron colliders such as the LHC, radiation damage is also a strong design constraint, and results showing improved performance for thin sensors have been reported in literature. Furthermore, the increased electric field required in order to maximise charge collection efficiency after irradiation results in a much reduced charge spread, negating advantages that may have been present pre-irradiation.

The main challenge posed by the VELO Upgrade is the huge quantity of data that must be transmitted off-detector. However, the single hit resolution remains an important issue as the specifications for the upgrade are that the IP resolution of the existing detector should be maintained or improved upon. The fluence across the detector will be highly non-uniform, and with appropriate segmentation the sensors can be biased independently, allowing those at outer radius to retain their performance throughout the lifetime of the experiment. As input into the detector design, and to make comparisons with the model constructed in simulations, a campaign of device characterisation focussing on single hit resolutions has been carried out. Tests have been carried out on sensors with varying thickness, implant type and resistivity, and the observations compared with an analytic model of the expected behaviour. This model has been extrapolated to evaluate likely future directions for detectors for particle physics.

4.1 Sensor Results

4.1.1 Methodology

All of the data presented were collected with the Timepix telescope, described in detail in the previous section, on the H8 beam line of the SPS North Area using a beam of positive hadrons (primarily pions). Tracks reconstructed were required to have hits on each plane of the telescope, and a loose χ^2 cut was applied to improve the track quality (the DUT was at all times excluded from the track fit). For each device, a calibration of the ASIC was carried out using the built-in test pulse in order to convert the measured ToT counts into the physical charge. The resolution is then defined as the standard deviation of a Gaussian fit to the residual distribution between the charge-weighted cluster centre of gravity and the projected track intercept. This is calculated for three different scenarios: raw data, charge-calibrated data and eta-corrected data.

The raw resolution is calculated using Centre-of-Gravity (CoG) clustering on the DUT with the raw ToT values. As noted previously, this contains non-linearity for charge injections around threshold, and so is expected to distort the residual distribution. Charge calibration alleviates this problem, but as the charge shared between two pixels does not vary linearly with intercept position, this is still an incorrect assignment of the cluster centre. The third variant, eta-corrected data, uses a function (the eta-function) that parametrises the difference between the charge CoG and the true track intercept, due to the non-linear sharing of charge as the intercept position moves between one pixel and another. This is carried out using the projected track position from the telescope and the charge-calibrated cluster CoG, where the resulting distribution is fitted with an (arbitrary) 5-order polynomial. This is then applied as a correction factor in the clustering, allowing a more accurate measure of the true track position. The procedure followed is described in more detail in [50].

For all devices, the angular offset due to mounting is obtained by a quadratic fit to the cluster size fraction in the direction of rotation.

Source measurements were carried out on most of the sensors in order to determine their full depletion voltage. Alpha particles were used as the incident radiation, and both the cluster size and charge were plotted against the applied bias voltage. From these values and the known thickness, the resistivities were calculated. For several of the sensors it was not possible to perform lab tests due to a fault in the voltage regulators employed. As the communication between the Timepix chips and the RELAXd DAQ boards passes through SCSI (Small Computer System Interface) cables, and the chip receives powering from the RELAXd boards, it is possible for the communication to fail due to voltage drops through the cables. This resulted in an instability of the initial setup. To combat this problem, a small regulator board was designed to sit on the Timepix PCBs and to directly regulate a more powerful low voltage source. These boards were later found to suffer from a design error, in that the failure of the regulators resulted in the input voltage being passed directly to the chip. Such voltages exceeded the chip tolerance and led to the loss of several assemblies. The powering setup was later redesigned with a careful sequence of regulators with hard limits, in order to protect the system and to provide a more stable supply. Subsequently the system operated smoothly and without the communication loss observed with the prior setup.

4.1.2 Device Description

A complete list of the devices tested is shown in table 4.1, and contains samples with different bulk type, implant type, thickness and bulk resistivity. Sensors from several manufacturers were tested. In all cases the sensors were bump-bonded to Timepix ASICs using tin-lead solder, and mounted on the standard Timepix PCB (as used by the telescope planes). Unless stated otherwise, all of the detectors were operated with a threshold of 1000 electrons, after equalisation of the pixel matrix.

The samples obtained from VTT are active-edge sensors, designed to have minimal material between the end pixel and the physical edge of the device [61]. The manufacture of these devices involves doping of the sensor edge in order to prevent the extension of the electric field and the subsequent electrical breakdown. The samples received were successful in this respect (in that charge is observed for tracks which pass beyond the final pixel boundary), but suffered from breakdown voltages above but close to their depletion voltages, placing a hard limit on the upper bias voltage that could be applied. The leakage currents of these devices were also high at moderate bias voltages, in some cases surpassing tens of microamps before full depletion (though all devices were operated overdepleted). The sample produced by CNM with n-type bulk similarly suffered from high leakage currents, likely due to the tuning of the p-spray used to electrically isolate the pixel implants.

Possible effects of these high leakage currents on the results are expected to be negligible. The Timepix front-end contains leakage current compensation circuitry, which diverts the current-controlled supply I_{krum}. This is the same current that is

Device ID	Technology	Resistivity (kΩ cm)	Thickness (µm)	Depletion Voltage (V)	Manufacturer
C08-W0098	p-on-n	8	150	10	CNM
D04-W0015	p-on-n	32	300	10	CANBERRA
J08-W0171	n-on-p	5	100	20	VTT
E11-W0171	n-on-p	> 10	200	< 40	CNM
D07-W0160	n-on-n	5.5	150	14	VTT
H08-W0171	n-on-n	5.5	200	25	VTT
F08-W0171	n-on-n	5.5	200	25	VTT

Table 4.1: Summary of sensors tested. The resistivity values have been taken from a combination of the manufacturer and lab measurements of the depletion voltage (using the quoted thickness from the manufacturer). For device E11-W0171 no diode measurements were made by the manufacturer and the device was unavailable for further study. The upper limit on the depletion voltage is taken from the (lack of) change in charge spectrum between beam data taken at 40 V and 60 V.

used to discharge the pixel in order to make the ToT measurement. I_{krum} is divided into two flows, one of which goes towards compensating leakage current through the pixel. As typical values for I_{krum} are of the order 20 nA, the front-end can cope with a total leakage current through the sensor of order 0.5 mA. This is well above the level of currents that were observed in the sensors tested.

4.1.3 Resolution Measurements

Due to the range of sensors and the large number of variables that directly affect the detector resolution, it is not practical to produce devices that are identical in all other respects. Comparisons between devices therefore involve by necessity several possibly conflicting effects. In order to reasonably distinguish between these, the devices are presented under conditions intended to mitigate other influences as much as possible. In particular, the varying resistivities of the samples lead to very different electric field configurations, which has a direct effect on the diffusion of charge. For each of the comparisons made below, the electric field configurations of the samples are thus shown alongside the experimental results.

Raw, Charge Calibrated and Eta-corrected Data

To illustrate the effects of both charge calibrations and eta-corrections on the experimental results, data taken with a single detector (D04-W0015) and a single bias voltage (60 V) are shown in figure 4.1. The application of the charge calibration is observed to degrade slightly the resolution at angles close to perpendicular incidence, but to improve significantly after the further application of eta-corrections. This is due to two effects. The first is that, in particular for low angles, the charge shared between two pixels is not linearly related to the intercept position along the pixel cell. Thus for tracks passing though the sensor such that most of the charge is deposited in a single pixel, the small amount of charge deposited in any neighbours will disproportionately pull the cluster centre away from the track intercept. The eta-function at perpendicular incidence is shown in figure 4.1 to illustrate this.

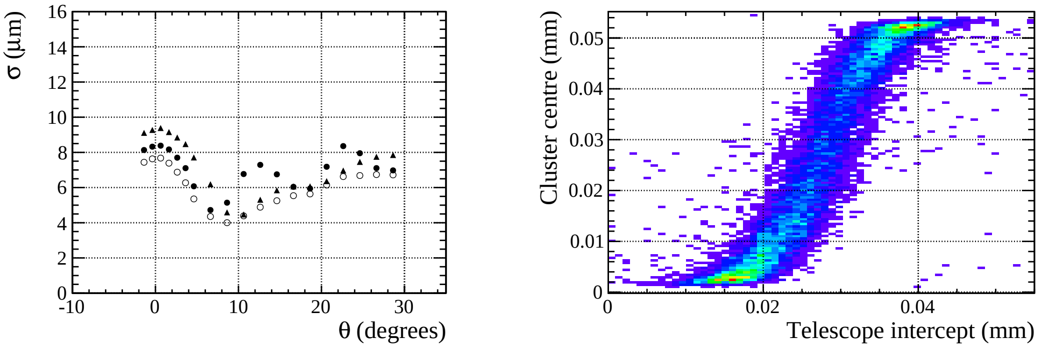

Figure 4.1: Left: Single hit resolution versus angle for raw (full circles), charge-calibrated (triangles) and eta-corrected (open circles) data. Right: Charge calibrated eta distribution at 0° (colour axis indicates the number of entries).

Without the application of charge calibrations, this effect is to some extent compensated by the response of the front-end electronics for low charge injections. As seen in the previous chapter, the non-linearity of the amplifier response is significant close to threshold, and so the weighting of the cluster centre in pixels with low charge is underestimated. A system with more linear charge response would then be expected to exhibit a worse initial resolution than that of the Timepix chip, but for this to be recoverable by the application of eta corrections.

Bias Voltage Dependence

Results for the device D04-W0015 are shown in figure 4.2. The sensor is made of n-bulk silicon, with p-doped implants (p-on-n), and is 300 μm thick, with a depletion voltage of 10 V. The cluster size in the direction of rotation is shown on the left hand plot of figure 4.2, and is interesting in that it shows that once the lateral path through the sensor exceeds the pitch (incident angles greater than 10°), the cluster size becomes independent of the applied bias voltage. Only at angles close to perpendicular incidence is the bias voltage a significant factor in determining the number of hit pixels. This is also observed to be the case for the single hit resolution, as shown in the right hand plot containing the charge calibrated eta-corrected resolution versus angle. Again, once beyond the angle at which tracks are geometrically required to pass through more than one pixel, the resolution of the device no longer appears to depend on the bias voltage.

The reason behind this feature is twofold. Firstly, the lateral spread of charge is facilitated both by the track passing laterally through the sensor, and through diffusion. As the incident angle increases, the contribution of charge diffusion decreases, notably in a non-symmetric fashion as charge deposited close to the pixel implants has a much smaller path length than that deposited at the back side of the sensor, and so will not diffuse over as great a distance. With sufficient track inclination the lateral extent of the charge cloud is therefore dominated by the lateral track length inside the sensor, and not by the diffusion of charge. Following this it would be expected that at sufficiently large angles the bias voltage should cease to have a significant impact on the net distribution of charge to neighbouring pixels.

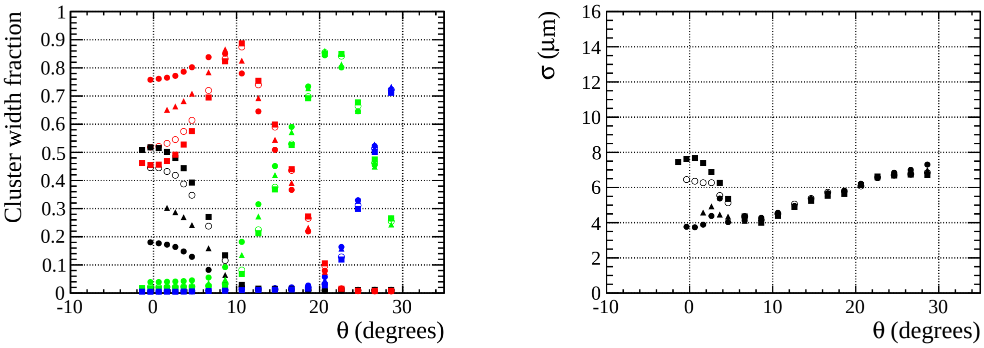

Figure 4.2: Results for device D04-W0015. Left: Fraction of n-pixel clusters versus angle (showing 1- (black), 2- (red) 3- (green) and 4- (blue) pixel clusters). Right: Single hit resolution (charge calibrated, eta-corrected) versus angle. Both plots show data for 10 (closed circles), 20 (triangles), 40 (open circles) and 60 (squares) V applied bias.

Secondly, the pixel pitch is large enough with respect to the charge cloud that a discrete change is observed between these two regimes (the first where the lateral diffusion dominates, the latter the track angle). If the pitch were small enough such that all clusters were multi-pixel even at perpendicular incidence, then a flatter resolution response would be expected; this is indeed the case as can be seen for the results in figure 4.2 at 10V. For this low field close to the depletion voltage (field distributions shown in figure 4.3), there is no significant transition in the resolution response where the incoming track guarantees at least two hit pixels (as the charge cloud size is large enough to reduce the contribution from single pixel clusters even at perpendicular incidence). At more moderate bias voltages however, there remains a significant fraction of single pixel clusters (over 50 % at 60 V). As the centre of gravity for these clusters is always the pixel centre, regardless of track intercept, they have a resolution given by

$$\sigma_{\text{single pixel clusters}} = f \times \frac{\text{pitch}}{\sqrt{12}} \tag{4.1}$$

where f is the fraction of single pixel clusters. This arises due to the self-selection of single pixel clusters close to the centre of the pixel and accounts for the shape of the resolution response at low angle, including the merging of the responses for all biases at the point where both the lateral track length dominates and the number of single pixel clusters is reduced to zero.

These results have several immediate implications. The most evident of these is the fact that with low electric field there is a significant amount of charge diffusion, and that this can significantly improve the spatial resolution of the detector. This may however have an impact on the timing of the system, where the response time may increase due to additional charge drift time or through timewalk effects in neighbouring pixels, which receive relatively low charge deposits. In addition, while the improvement in performance for this device is significant, it is likely that the gains for a different device would be sensitive to the resistivity of the sensor substrate. The difference in mean field between the data taken at 60 V and 10 V is around a factor of 4, and so it is likely that devices with lower resistivity, which require higher fields in order to fully deplete, would not show such significant gains even by operation at bias voltages close to depletion.

The second implication of the results is in suggesting forward directions in the

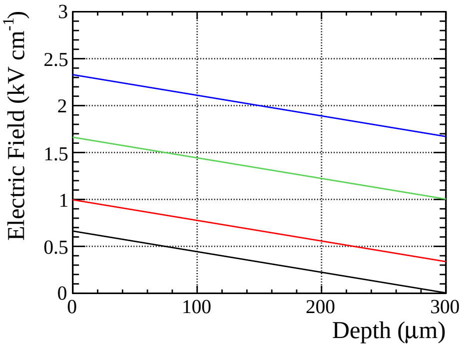

Figure 4.3: Electric field profiles for D04-W0015 at 10 (black), 20 (red), 40 (green) and 60 (blue) V. The depth is with respect to the pixel implant side of the sensor.

design of sensors and associated electronics. It becomes clear that at least until the pitch becomes comparable with the size of the charge cloud, it is advantageous to pursue smaller pixel sizes. In addition to allowing for potentially improved performance at low angles, smaller pixel size (but fixed thickness) would allow the field-independent regime to be achieved with lower incident angles, improving performance even at high voltage. In parallel, a move towards higher resistivity substrates would allow for flatter field profiles, where it might be expected that a detector with low depletion voltage operated in an over-depleted regime would outperform a lower resistivity detector operated just at depletion. Such a comparison is considered in detail below.

Threshold Dependency

The Timepix chip performs a digital measurement of the input charge during each acquisition. The threshold above which charge is measured is limited by the total noise of the system, and the acceptable noise occupancy for a given application. While unirradiated, the threshold is not a significant limitation on the single hit efficiency as the charge deposited in even a thin sensor is still higher than the operational threshold of most modern electronics. After irradiation this becomes more complicated, with the charge collected a function of both fluence and applicable bias voltage, and the threshold starts to become a limiting factor. The result of this has been to push front-end electronics towards lower thresholds, and to maximise the charge collected after irradiation. For resolution purposes, the most significant effect of non-zero threshold is the loss in resolution for hits where a neighbouring pixel does not register charge. This skews the charge centre of gravity, and cannot generally be recovered by the application of eta-corrections.

The results for a single detector operated at different thresholds are shown in figure 4.4. The sensor is 200 µm thick, with both n-type bulk and pixel implant (n-on-n), and a threshold of either 1000 or 2000 electrons. The cluster fraction plot on the left hand side shows some interesting trends: while it is expected that the average cluster size will be lower (with a higher fraction of single pixel clusters), it is notable that unlike the changes affected by varying the bias voltage, the shape of the two curves is identical,

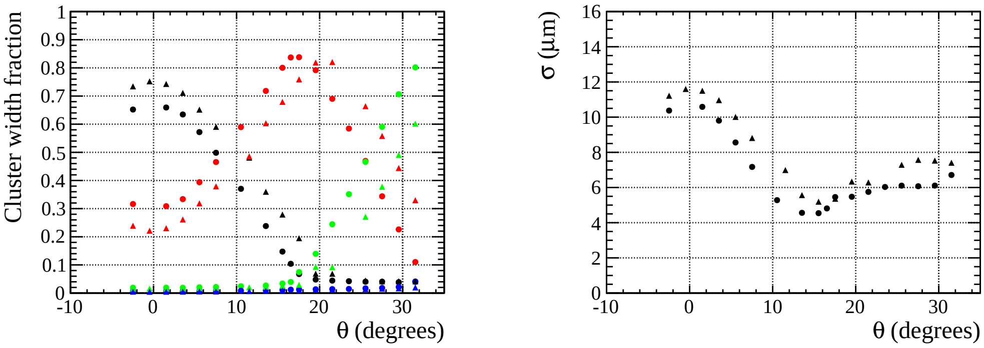

Figure 4.4: Results for device F08-W0171 at 60 V with 1000 (circles) and 2000 (triangles) electron threshold. Left: Fraction of n-pixel clusters versus angle (showing 1- (black), 2- (red) 3- (green) and 4- (blue) pixel clusters). Right: Single hit resolution (charge calibrated, eta-corrected) versus angle for 1000 (circles) and 2000 (triangles) electron threshold.

with a simple angular offset of $\sim 4^\circ$ between them. While the data taken at different voltages also showed that the device response becomes identical at a sufficiently large angle there does not appear to be such a relationship between the data by varying the threshold.

The resolution curves on the right hand plot of figure 4.4 also show this angular offset, in the same way as the cluster fraction distributions. As expected, the resolution is degraded in the device with higher threshold, but there is again no merging of the resolution response at higher angles. In addition, the minimum achievable resolution is worse, as is the resolution in the plateau beyond the geometric requirement for two pixel clusters. This suggests that a sizeable component of the detector resolution, in particular into the geometrically dominated region, comes from missing charge in pixels that remain under threshold.

N-on-n versus n-on-p

Electron collecting devices have become popular in high energy physics due to the higher mobility of electrons, and thus the shorter collection time and reduced trapping probability after irradiation. Two types of electron collecting sensors can be manufactured: those with p-type bulk and those with n-type. N-on-n sensors have the significant difference from both n-on-p and p-on-n devices of depleting from the back side of the sensor, rather than from the pixel implant. In order to deplete these devices, the back side is doped with p-type silicon, giving this reversal of the p-n junction location and causing the field to extend from the back side towards the pixel implants. The behaviour of such devices has been well studied, and was the technology of choice for the fabrication of the current VELO sensors.

One sensor of each bulk type is considered, with the operating conditions chosen to give similar charge collection profiles. In the absence of exact resistivity measurements this is taken from the cluster composition versus angle, where the detectors are compared where they have similar fractions of n-pixel clusters. The sensor on H08-W0171 has n-type bulk, with a depletion voltage of 25 V, while that of E11-W0171 has p-type bulk and depletes below 40 V; both sensors are 200 µm thick. Two sets of data are presented, that where both detectors were operated at 60 V and that where both were operated with 40 V (approximate fields shown in figure 4.5).

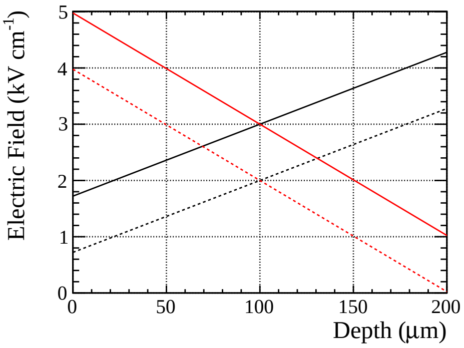

Figure 4.5: Electric field profiles for n-on-n H08-W0171 at 40 V (black dashed) and 60 V (black solid), and n-on-p E11-W0171 at 40 V (red dashed) and 60 V (red solid).

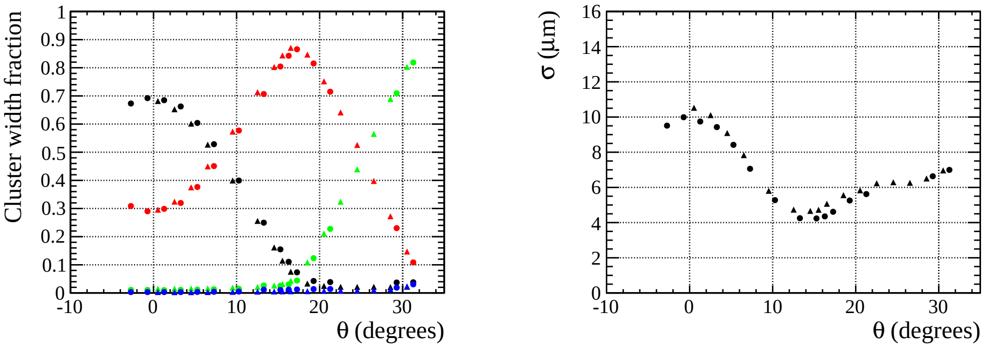

Figure 4.6: Results for n-on-n device H08-W0171 at 60 V (circles) and n-on-p E11-W0171 at 60 V (triangles). Left: Fraction of n-pixel clusters versus angle (showing 1- (black), 2- (red) 3- (green) and 4- (blue) pixel clusters). Right: Single hit resolution (charge calibrated, eta-corrected) versus angle.

The higher voltage results are shown in figure 4.6, and no significant variation between the devices is observed. In both cases charge diffusion is limited, with predominantly single pixel clusters (around 70 % at perpendicular incidence) accounting for the degraded spatial resolution. The minimum resolution is observed again to be around 4 μm, occurring around the geometrically predicted position of 15°.

The results at the lower voltage are more surprising, as shown in figure 4.7. While it would again appear that there is little variation in the cluster distributions between the two sensors, there is a noticeable difference in the single hit resolution, with the n-on-n sensor showing an improved precision of around 1 μm at perpendicular incidence. The source of this difference appears to come from the resolution obtained for single pixel clusters.

As described previously, the resolution observed by selecting only single pixel clusters can be approximated by equation 4.1. This is simply due to the relationship between the width of a square distribution to its Root Mean Square (RMS). Where there is no diffusion of charge, this would be 15.9 μm for the Timepix pitch of 55 μm. Where

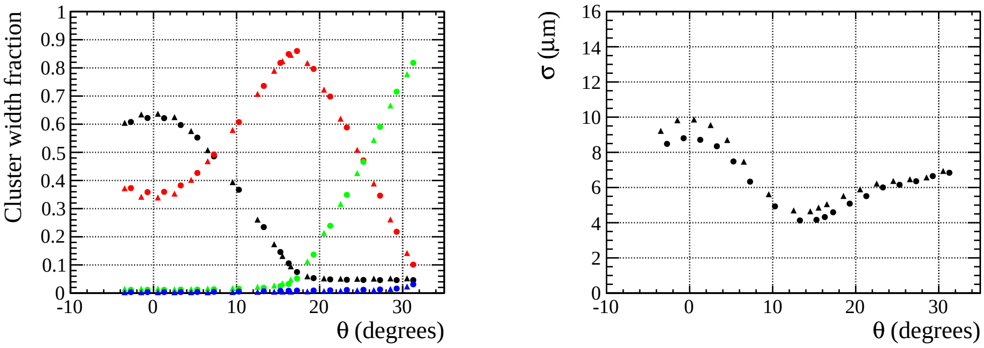

Figure 4.7: Results for n-on-n device H08-W0171 at 40 V (circles) and n-on-p E11-W0171 at 40 V (triangles). Left: Fraction of n-pixel clusters versus angle (showing 1- (black), 2- (red) 3- (green) and 4- (blue) pixel clusters). Right: Single hit resolution (charge calibrated, eta-corrected) versus angle.

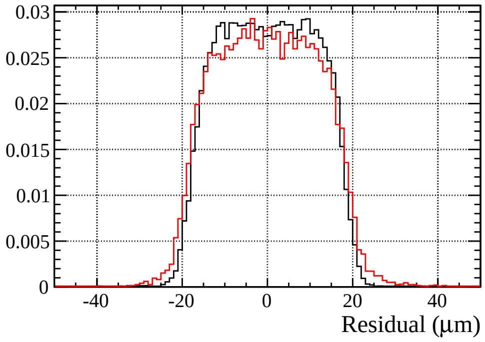

Figure 4.8: Single pixel residuals for n-on-n H08-W0171 (black) and n-on-p E11-W0171 (red) at 40 V and perpendicular incidence.

diffusion is present, the resolution of single pixel clusters is simply the region around the pixel implant where the charge cloud does not spread far enough to be collected by neighbouring pixels.

As the two devices compared here have operating conditions chosen such that the fraction of n-pixel clusters is approximately the same, it is reasonable to expect that the device resolution would therefore necessarily be the same. This is not strictly the case. The charge diffusion is controlled by both the depth of the charge deposit and the integrated field lying between the ionisation point and the pixel implant. While the integrated fields for both devices are similar over the full thickness, the field orientations throughout the bulk are reversed, meaning that charge in each scenario will have to pass through either the highest or lowest field in order to reach the pixel implants. Furthermore, the stochastic nature of the charge generation means that significant variations in the deposited charge exist along the track length.

For the n-on-n device the lower field region is located at the pixel implant side of the sensor. This means that charge deposits with smaller distances to travel will experience a lower electric field, and those where the field is highest will have a greater collection path. In both cases the field helps to counter the relationship that would be expected only considering the charge path: that diffusion increases sharply with depth. The combination of low-field, low-depth and high-field, high-depth acts to reduce the variation in diffusion with depth, minimising the effect of variations in the charge deposited along the path length.

This situation is reversed in the n-on-p device: charge deposited close to the pixel implants, which would not be expected to diffuse much, is also in the area of highest field, reducing the diffusion further. Charge at the back side of the sensor has the maximal conditions of low field and long collection distance. The response of the sensor, in particular whether a neighbouring pixel is hit or not, should be more susceptible to the location of any high charge deposits. For the same track position, those where a large deposit is created close to the pixel side of the sensor would produce (for that individual track) a single pixel cluster, while those where the deposit was close to the back side of the sensor are more likely to produce multi-pixel clusters.

The evidence of this is shown in figure 4.8, where the residual distribution for single pixel clusters can be seen at 40 V and perpendicular incidence for both devices. From the cluster fraction distribution in figure 4.7, it can be seen that the number of single pixel clusters at this angle is the same for both devices (within 2 %). The width of the distribution for the p-type bulk device appears wider however, with a pronounced tail in the fall off of single pixel clusters. Even at the pixel edge (27.5 µm) there are still hits where the charge does not spread to both pixels.

N-on-n versus p-on-n

Prior to the widespread availability of high resistivity p-type wafers, sensors were produced primarily with n-type bulk. The double-sided processing required to produce n-on-n sensors led to the use of p-on-n sensors for many applications, despite the lower carrier mobility of holes in silicon. As was seen previously in the comparison between n-on-n and n-on-p devices, where the carrier type is the same between two devices but the depletion side is reversed a small difference in resolution is seen for low bias voltages. For the devices shown below, the bulk type is the same for both devices, but both the depletion side and carrier type are different.

The detector D07-W0160 contains both n-type bulk and n-type implants, depleting from the back side. Device C08-W0098 has instead p-type pixel implants and collects holes. Both sensors are 150 µm thick, with depletion voltages around 10 V. The data shown are with both devices operated at 40 V. Significantly, the threshold for the p-on-n device is not the same as that of the n-on-n. Likely due to the same problems during the fabrication, which led to the significantly higher than expected leakage current for C08-W0098, the pixel matrix showed a large number of noisy pixels after equalisation and operation at 1000 electron threshold. The threshold was thus increased to 1500 electrons, and would be expected to have a significant effect (as seen from the previous sections).

The cluster width distributions for each device are shown in the left hand plot of figure 4.9. Two things are immediately noticeable: the difference in threshold and the apparent difference in diffusion. As seen in figure 4.4 for the device operated with different thresholds, the response of the cluster fraction plot remains roughly the same, but is offset in angle by several degrees. This appears to be the case in figure 4.9, where the data for the p-on-n device has a comparable gradient to that of the n-on-n, but

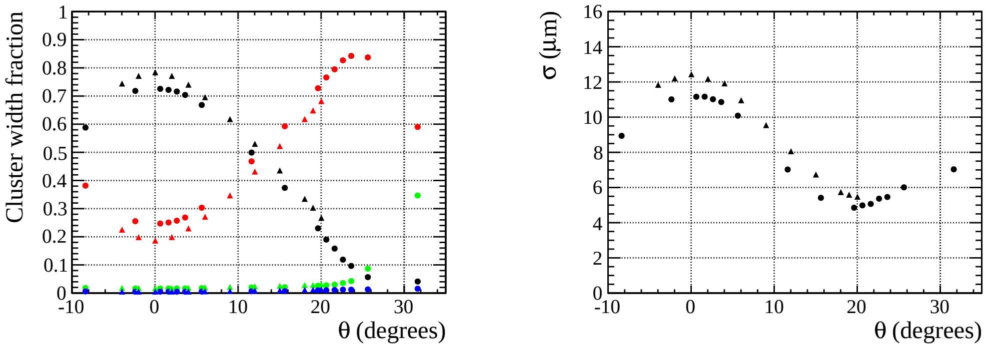

Figure 4.9: Results for n-on-n device D07-W0160 at 40 V (circles) and p-on-n C08-W0098 at 40 V (triangles). Left: Fraction of n-pixel clusters versus angle (showing 1- (black), 2- (red) 3- (green) and 4- (blue) pixel clusters). Right: Single hit resolution (charge calibrated, eta-corrected) versus angle.

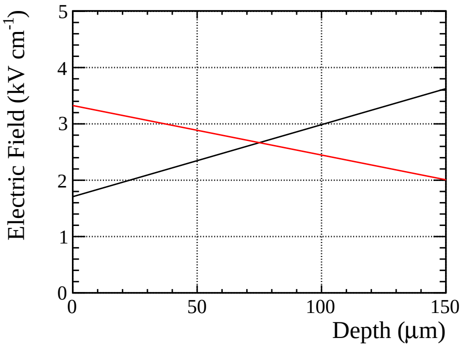

Figure 4.10: Electric field profiles for n-on-n D07-W0160 at 40 V (black) and p-on-n C08-W0098 at 40 V (red).

is roughly 2 - 3° offset. The change in gradient around 0° suggests that the level of diffusion is slightly different between the two devices as well, likely due to differences in the electric fields throughout the sensors (approximate fields shown in figure 4.10).

The resolution of each device is shown in the right hand plot of figure 4.9. It is not possible to definitively dismiss any differences due to the collected charge carrier type. From the difference in cluster fraction due to the higher threshold in the p-on-n device, the resolution curve should display a similar angular offset, which is indeed the case. Furthermore, the difference in depletion side has already been observed above to introduce a slight degradation in the resolution for angles below the geometric minimum. Both of these effects are sufficient to explain the 1 - 1.5 µm difference in resolution between the devices, and so any differences due to the carrier type must degrade the resolution by less than 1 µm.

Trends with Thickness

Now that the differences between the differing bulk and implant types have been investigated, the more significant trend in detector response can be considered. As the move towards less material in particle physics experiments progresses, both sensors and ASICs have been thinned in order to reduce the level of multiple scattering within tracking systems. Support structures, cooling and readout material are also following this direction, but the possible gains are steadily reducing as limits on the physical properties of the construction materials are reached. Novel structures and approaches, such as the microchannel substrate intended for use in NA62 and the LHCb VELO, are used where conventional techniques are prohibitively invasive. With sensors, the abilities in hybrid pixel detector construction are already below that required to progress from the current generation of experiments. Sensors with thicknesses down to 50 µm are already becoming available, though detector-scale productions at such extremes have not yet been carried out. Understanding the trends for detector performance as sensor technology progresses in this vein is important for the design considerations.

The results from three sensors are shown in figure 4.11: those for a 100 µm, 200 µm and 300 µm sensor, all operated at 60 V (with depletion voltages of 20 V, 25 V and 10[1] V respectively). The 200 µm sensor is back-depleted, while the other two detectors

---

[1] Due to the higher resistivity of the 300 µm sensor.

deplete from the pixel side of the sensor, though with differing bulk and implant types. The thresholds for all devices was 1000 electrons.

The main trends in the cluster size can be clearly seen in the left hand of figure 4.11. With decreasing thickness, the corresponding reduction in diffusion gives rise to a larger fraction of single pixel clusters at perpendicular incidence. For each device this falls off with angle, reaching zero once the track is required geometrically to pass through at least two pixels (with an offset due to the non-zero threshold). The resolution in this same region is entirely dominated by the resolution of these single pixel clusters, simply given by the standard deviation of the pixel region where charge sharing is not observed. In figure 4.11 it can be seen that at conditions where the cluster fraction is the same, the resolution is comparable, regardless of the device thickness. This is consistent with the previous observation that the doping type has little effect on the device performance.

Once the track angle is sufficient as to cross into the geometrically dominated region, the detector response follows the same trend for all devices. It is notable that the thinnest sensor has a worse minimum resolution than the other devices, which may be due to the relative size of the threshold with respect to the signal charge. The ratio of lateral pitch to thickness gives equivalent points between the different sensor thicknesses, and this pushes the bias-independent regime to higher angles with decreasing thickness. As a result, the uniformity of the device improves for thin sensors, though since the diffusion is significantly limited this makes the sensor response uniform at the cost of significantly degraded resolution with respect to thicker sensors operated at low bias voltage.

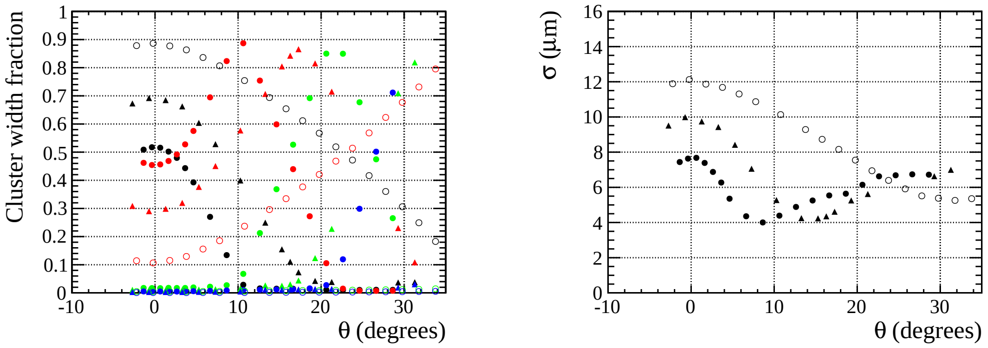

Figure 4.11: Results for 300 μm p-on-n device D04-W0015 (circles), 200 μm n-on-n H08-W0171 (triangles) and 100 μm n-on-p J08-W0171 (open circles), all at 60 V. Left: Fraction of n-pixel clusters versus angle (showing 1- (black), 2- (red) 3- (green) and 4- (blue) pixel clusters). Right: Single hit resolution (charge calibrated, eta-corrected) versus angle.

4.2 Analytic Model

In order to predict the performance of a given detector, Monte Carlo simulations are typically employed, with varying levels of detail in the description of the physical processes involved. The charge generation and transport as well as the production of secondary particles all have to be described, and many simulated events are used to generate the expected response. From the results shown above, it might be thought possible to instead derive an analytic expression that describes the detector resolution for a given set of conditions.

4.2.1 Observations in Data

There are several features observed in the data which suggest directions that should be followed. At angles close to perpendicular, there is a clear relationship between the resolution and the physical extent of the visible (above threshold) charge cloud. This appears to be dominated for most operating conditions by the fraction of single pixel clusters observed which, given the detector pitch, typically have a much coarser resolution. The extent of diffusion, controlled by the thickness and electric field strength, should be the primary source of error at such angles.

For large angles, it is clear that effects due to the electric field are not the dominating source of error in the position measurement. Further, there appear to be distinct differences between clusters of different size, facilitating a stepped loss of resolution at predictable angles based on the geometry of the pixel. The resolution in this region does not appear to be influenced by the electric field, and so other sources of noise must be considered. These might include threshold effects, variations in deposited charge, or the front-end electronics themselves.

4.2.2 Analytic Expression

In order to present a generic expression that is not strongly tied to front-end electronics, sensor geometry or data processing, an analytic expression is derived for a device using the true charge and without the application of any corrections. This contrasts with the results above where eta corrections have been implemented. A prediction of the detector resolution versus angle and bias voltage can be attempted by considering the various sources of error in the determination of the track intercept on a single plane. The experimental residual distribution can be thought of as the sum of several terms:

$$\sigma^2_{\text{single hit}} = \sigma^2_{\text{eta function}} + \sigma^2_{\text{landau fluctuations}} + \sigma^2_{\text{track extrapolation}} \tag{4.2}$$

The first of these terms comes from the clustering algorithm employed - used without knowledge of the track incident angle or detector dimensions - and is simply the difference in charge shared between pixels as a function of intercept position. The centre of gravity (CoG) determination of the cluster centre assumes a linear relationship between the track intercept and the sharing of charge, while this in reality is a more complicated function similar to the standard error function (such as that shown in figure 4.1). The second term comes from the fluctuations in deposited charge along the track length through the sensor, while the third is simply the track extrapolation error.

Charge Centre of Gravity

In order to calculate the term $\sigma_{\text{eta function}}$, some statements and assumptions about the sensor response must be made. Using the Einstein diffusion relation and the assumption of low field, where the carrier velocity, $v(z)$, is linearly proportional to the electric field E_z, the standard deviation $\sigma_{\text{diffusion}}$ for the lateral charge drift is

$$\sigma_{\text{diffusion}} = \sqrt{\frac{2k_bT}{q}\mu_c \int_0^z \frac{dz}{v(z)}} \tag{4.3}$$

for a discrete quantity of charge deposited distance z from the collecting electrode and with carrier mobility μ_c (where k_b is Boltzmann's constant, T is the temperature and q is the electric charge). The integral is evaluated differently for front- and back-depleting devices due to the differing field configurations, and for a front-depleted device with thickness d and field profile (for voltages V above the depletion voltage V_{dep}) of

$$E_z = \frac{qN}{\epsilon}(d-z) + \frac{V - V_{dep}}{d} \tag{4.4}$$

the integral inside equation (4.3) becomes

$$\int_0^z \frac{dz}{v(z)} = \frac{\epsilon}{qN\mu_c} \ln\left(\frac{d + d_0}{d + d_0 - z}\right) \tag{4.5}$$

where ϵ is the permittivity in the material, N is the dopant concentration and d_0 is an offset resulting from operating the sensor beyond full depletion (the distance beyond the sensor which would be required to reduce the electric field to zero), given by

$$d_0 = \frac{\epsilon(V - V_{dep})}{qNd} \tag{4.6}$$

For a back-depleted sensor this integral is instead

$$\int_0^z \frac{dz}{v(z)} = \frac{\epsilon}{qN\mu_c} \ln\left(\frac{d_0 + z}{d_0}\right) \tag{4.7}$$

Continuing for the moment with the front-depleted description, the final expression for the lateral diffusion then becomes

$$\sigma_{\text{diffusion}} = \sqrt{\frac{2k_bT}{q}\epsilon\rho\mu_b \ln\left(\frac{d + d_0}{d + d_0 - z}\right)} \tag{4.8}$$

where μ_b and ρ are the bulk mobility and resistivity respectively.

This expression for the charge diffusion notably does not contain any terms relating to the carrier mobility, and would therefore be the same for electron or hole collecting devices. The three terms which might be controlled are then the applied bias voltage, the resistivity and the bulk doping type. Figure 4.12 shows the calculated lateral charge diffusion versus depth, for the parameters listed for D04-W0015 (p-on-n, 300 µm thick with a resistivity of 32 kΩcm) and with the voltages used during data taking. Comparing this with the cluster size and resolution results at perpendicular incidence for D04-W0015, the calculated lateral diffusion seems to follow well the observed trend, in particular the relative changes between bias voltages.

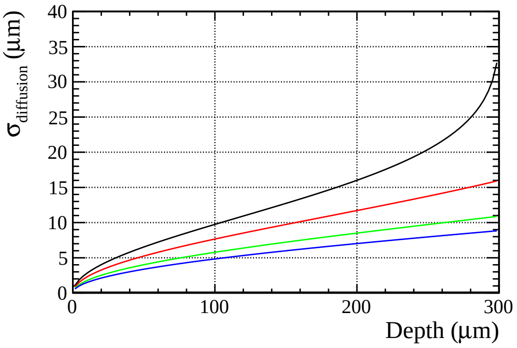

Figure 4.12: Diffusion parameter versus depth, calculated for device D04-W0015 at 10 (black), 20 (red), 40 (green) and 60 (blue) V.

Now that the diffusion of a discrete quantity of charge deposited at an arbitrary depth within the sensor is known, this must be extended to describe the collected charge profile. This is typically where the switch to a Monte Carlo simulation would be performed, in order to reproduce the stochastic nature of interactions with matter. Putting this to one side for the moment, in order to consider a general charge cloud or drift field that would arise from uniform and continuous charge deposits through the depth, the following lateral charge profile for a front-depleted sensor would be expected:

$$Q(x) = \frac{1}{\text{thickness}} \int_0^{\text{thickness}} \frac{Q_{total}}{\sigma_{\text{diffusion}}\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{\sigma^2_{\text{diffusion}}}} \, dz \tag{4.9}$$

where $Q(x)$ is the charge collected at lateral point x and μ is the lateral position of the charge deposit. This expression is simply the value of a Gaussian charge distribution (with mean μ) for an individual point at lateral distance x from the interaction, integrated throughout the full depth. Q_{total} is based on the Most Probable Value (MPV) of the charge distribution for the sensor thickness, and varies with the track angle θ such that

$$Q_{total}(\theta) = \frac{Q_{MPV}}{cos(\theta)} \tag{4.10}$$

For perpendicular incidence, μ is just the track entry position along the detector, but for angled tracks this varies as a function of the depth through the sensor such that

$$\mu(z) = z \times tan(\theta) \tag{4.11}$$

With this expression the charge arriving to the detecting plane of the sensor can be calculated, and is shown in figure 4.13 for the same conditions as figure 4.12 above. The left hand plot shows the lateral charge distribution as a function of bias voltage, while the right hand plot shows the calculated distributions for angled tracks at both low and high voltage. As would be expected once there is significant inclination of the track, the amount of charge registered to the left of the entry position ($x = 0$) falls off rapidly, while there is a significant plateau covering the lateral path of the track. This is in keeping with the extreme case of the track passing parallel to the sensor surface - a constant charge would be deposited in all pixels, under the above assumptions. The

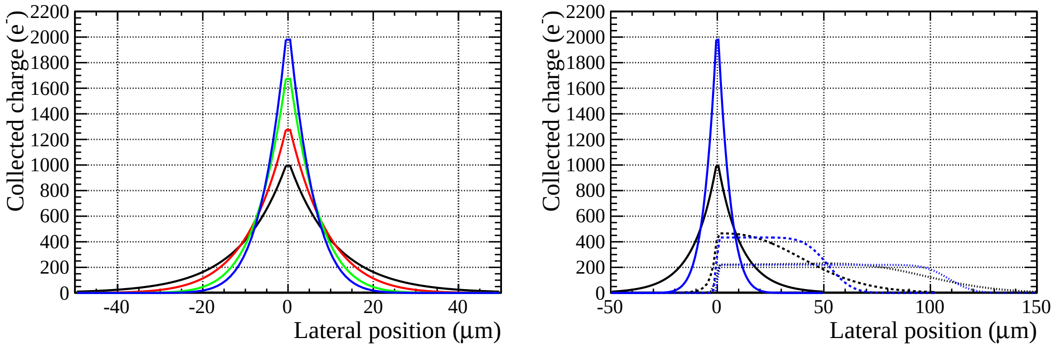

Figure 4.13: Lateral charge distributions calculated for device D04-W0015. Left: Lateral charge at 0° for 10 V (black), 20 V (red), 40 V (green) and 60 V (blue). Right: Lateral charge at 0° (plain), 10° (dashed) and 20° (dotted) degrees, for 10 V and 60 V.

variation of the charge collected for low and high fields is also observed to decrease with angle, as expected from the experimental results.

Now that the charge arriving to the pixel implant plane is known, the cluster centre of gravity can be calculated. The integral of the charge distribution over the physical pixel limits is used as a simplistic approximation, in order to avoid a full description of the electric field around the implant. The cluster barycentre is then given by the standard expression:

$$x_{cluster} = \frac{\sum x_i q_i}{\sum q_i} \tag{4.12}$$

where x_i is the pixel centre and the pixel charge q_i is given by:

$$q_i = \int_{x_{\text{low}}}^{x_{\text{high}}} Q(x)\, dx \tag{4.13}$$

with pixel lower and upper bounds of x_{low} and x_{high} respectively. Defining the average local residual as the difference between the true track intercept x (at the centre of the sensor) and the reconstructed cluster barycentre x_{cluster}

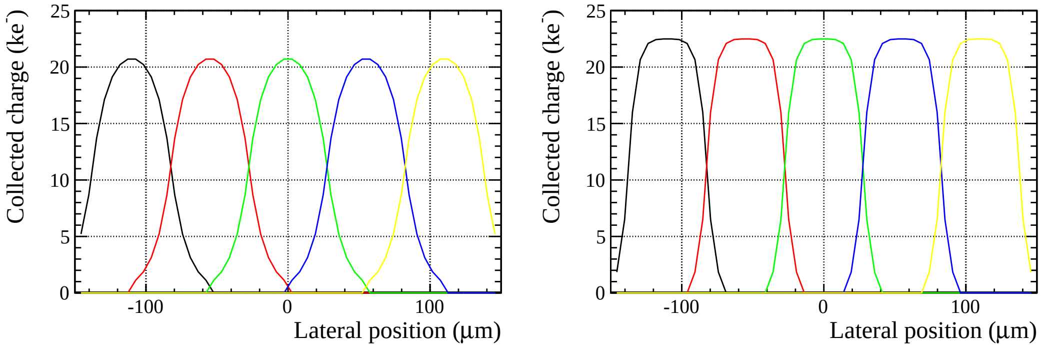

Figure 4.14: Pixel charge at 0° as a function of beam intercept position, for several pixels, at 10 V (left) and 60 V (right). Sensor parameters used to match device D04-W0115.

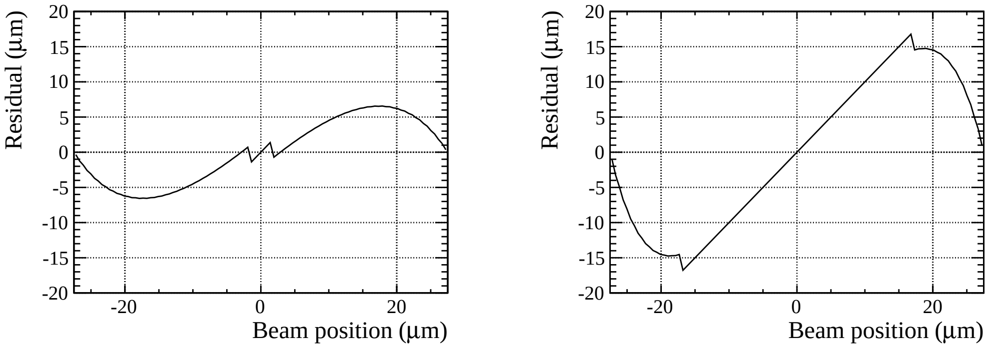

Figure 4.15: Average residual at 0° as a function of beam intercept position, at 10 V (left) and 60 V (right). Sensor parameters used to match device D04-W0115.

$$r(x) = x - x_{\text{cluster}} \tag{4.14}$$

the average residual versus intercept position can be plotted.

The calculation output at various stages is shown above. Figure 4.14 shows the response of several pixels as the beam travels laterally along the sensor surface. As before, these results are simulated using the parameters of D04-W0015. The residual distributions versus beam position over a single pixel cell are shown in figure 4.15. In both cases, plots are shown for low and high applied bias voltage, and highlight the differences in performance. In particular, the reduced overlap of the pixel responses for the data simulated at 60 V illustrates the prevalence of single pixel clusters, and the subsequent linear error in the residual function.

Extension to a 2-dimensional model

Using the above construction of the pixel response, it is notable that the charge cloud extends beyond the pixel boundary at perpendicular incidence (particularly for a bias voltage close to depletion). This suggests that the evaluation of the detector resolution at low field will not describe the loss of charge to pixels under threshold in the y-direction. To correct for this, the 2 dimensional charge cloud can be reconstructed by adding the corresponding component for diffusion in y. This is no different from the description shown above, with the exception of fixing the y-component of the incident angle to perpendicular. The residual distribution is still given by equation 4.12, but this time the charge q_i is given by:

$$q_i = \int_{x_{\text{low}}}^{x_{\text{high}}} Q(x)\ dx \int_{y_{low}}^{y_{high}} Q(y)\ dy \tag{4.15}$$

and the normalisation factor is $\sqrt{Q_{total}(\theta)}$ instead of $Q_{total}(\theta)$. The charge collection profile for a single pixel pixel at perpendicular incidence is shown in figure 4.16, again for both low and high voltages. The same plots are shown in figure 4.17 for 10° incidence.

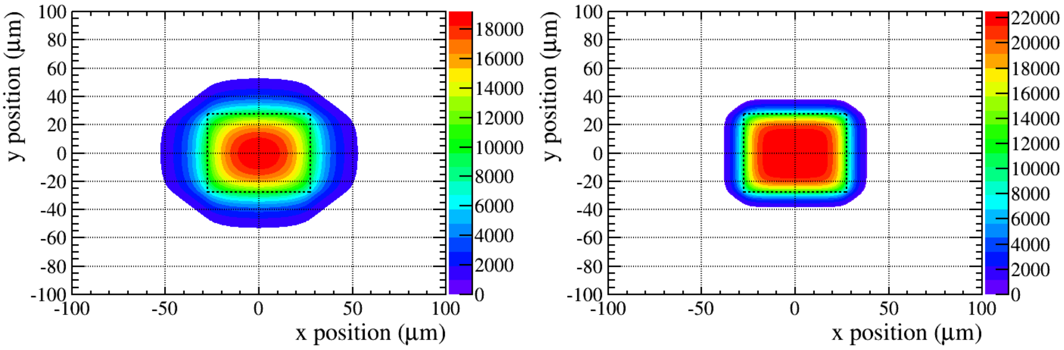

Figure 4.16: Pixel charge for a single pixel versus beam entry position at 10 (left) and 60 (right) V, both at perpendicular incidence. The dotted line shows the pixel boundary.

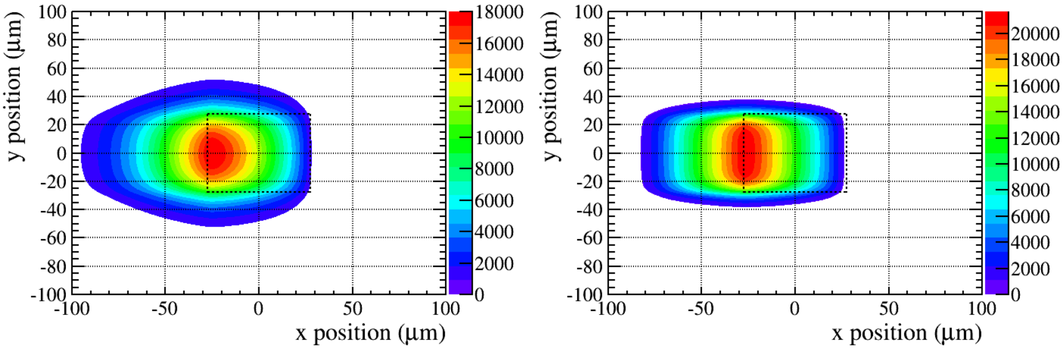

Figure 4.17: Pixel charge for a single pixel versus beam entry position at 10 (left) and 60 (right) V, both at 10° incidence. The dotted line shows the pixel boundary.

Charge fluctuation

Landau fluctuations in the deposited charge along the track length lead to an irreducible distortion of the charge cloud CoG. At angles close to perpendicular, multi-pixel clusters are purely the result of charge diffusion, and so any variations in the charge deposits are shared amongst all pixels. For angled tracks this is no longer the case, and variations between the charge deposited along the track are no longer shared strongly between pixels. In the experimental data this leads to a gradual loss of resolution with increasing angle. Since the charge is collected over a small number of discrete points (the pixels), the result of variations in the charge deposited will simply lead to an error in the measured charge for each pixel, dependent on the physical path length through it. The subsequent error on the reconstructed position can then be calculated separately for pixels of size n.

For clusters containing two pixels with centres x_1, x_2 and collecting charge q_1, q_2, the centre of gravity x_{CoG} is given by

$$x_{CoG} = \frac{q_1 x_1 + q_2 x_2}{q_1 + q_2} \tag{4.16}$$

Introducing an error on the charge of each pixel (σ_{q_i}), setting $x_1 = 0$ and $x_2 - x_1 = p$ (the pitch), an error on the reconstructed x-position $\sigma_{x_{CoG}}$ can be calculated by standard error propagation to be

$$\sigma^2_{x_{CoG}} = \frac{p^2 q_1^2}{(q_1 + q_2)^4}\,\sigma^2_{q_1} + \frac{p^2 q_2^2}{(q_1 + q_2)^4}\,\sigma^2_{q_2} \tag{4.17}$$

including the contribution from both pixels. Taking the average situation where the track intercept is such that there is an equal track length in each pixel (so that $\sigma_{q_1} = \sigma_{q_2} = \sigma_q$), the expression simplifies to

$$\sigma^2_{x_{CoG}} = \frac{p^2 \sigma_q^2 (q_1^2 + q_2^2)}{(q_1 + q_2)^4} \tag{4.18}$$

For the most pessimistic case where most of the charge is in one pixel, the approximation $q_1 \approx Q_{total}$ can be used to further simplify this expression to

$$\sigma^2_{x_{CoG}} = \frac{p^2 \sigma_q^2}{Q^2_{total}} \tag{4.19}$$

where the substitution $q_1 + q_2 = Q_{total}$ has also been used.

For three pixel clusters, the centre of gravity is given by

$$x_{CoG} = \frac{q_1 x_1 + q_2 x_2 + q_3 x_3}{q_1 + q_2 + q_3} \tag{4.20}$$

As the pixels can not now be considered identical, the standard error of each pixel is calculated separately. For pixel 1, this becomes

$$\sigma^2_{q_1}\left(\frac{\delta x_{CoG}}{\delta q_1}\right)^2 = \sigma^2_{q_1}\frac{p^2(q_2 + 2q_3)^2}{Q^4_{total}} \tag{4.21}$$

where $\delta x_{CoG}/\delta q$ is the partial derivate of the cluster centre of gravity with respect to the pixel charge. Again taking the pessimistic case where the excess charge is deposited furthest from the pixel, giving the approximation $q_3 \approx Q_{total}$, this can be simplified to

$$\sigma_{q_1}^2 \left(\frac{\delta x_{CoG}}{\delta q_1} \right)^2 = \sigma_{q_1}^2 \frac{4p^2}{Q_{total}^2} \tag{4.22}$$

Similar expressions can be found for the errors due to pixels 2 and 3, taking once more the overly pessimistic case where excess charge is deposited furthest from the pixel being considered

$$\sigma_{q_2}^2 \left(\frac{\delta x_{CoG}}{\delta q_2} \right)^2 = \sigma_{q_2}^2 \frac{p^2}{Q_{total}^2} \tag{4.23}$$

$$\sigma_{q_3}^2 \left(\frac{\delta x_{CoG}}{\delta q_3} \right)^2 = \sigma_{q_3}^2 \frac{4p^2}{Q_{total}^2} \tag{4.24}$$

to give an overall expression for the error of a 3-pixel cluster of

$$\sigma_{x_{CoG}}^2 = \sigma_{q_1}^2 \frac{4p^2}{Q_{total}^2} + \sigma_{q_2}^2 \frac{p^2}{Q_{total}^2} + \sigma_{q_3}^2 \frac{4p^2}{Q_{total}^2} \tag{4.25}$$

This approach has similarly been applied to higher order cluster sizes.

As at any given angle there will be a mixture of cluster sizes, depending on the track intercept along the pixel, it is necessary to combine the different errors which arise. This is performed using only the lateral path length through the sensor, as n-pixel clusters arising from charge sharing would not show the same inhomogeneity of charge (the extreme version being 2-pixel clusters at perpendicular incidence, where there is no variation in charge deposited as a function of the lateral position). The fraction of n-pixel clusters due solely to the track angle is calculated geometrically using the ratio of lateral distance to pitch, and the errors combined simply by summing the weighted errors

$$\sigma_{x_{\text{landau fluctuations}}} = \Sigma f_i \sigma_i \tag{4.26}$$

where σ_i is the calculated error for pixels of size i, and f_i is the fraction of i-pixel clusters at a given angle.

The final requirement for the calculation of these errors is the fluctuation on the charge measured in each pixel. There are many detailed calculations available in literature for the fluctuation in charge deposition along the path of a charged particle in matter, and many of these require tuned Monte Carlo simulations to determine the exact shape of the energy loss spectrum for a given material. In order to be able to write an analytic form for the purpose of this work, a few approximations have been made. Specifically, the shape of the energy loss spectrum has been taken from the Landau-Vavilov function, where the full width half maximum (FWHM) of the energy loss spectrum is given by 4ξ, where ξ is taken from equation 2.38. It is known that the Landau-Vavilov function underestimates the width of the energy loss spectrum for thin absorbers, including silicon detectors. For this reason, deliberately pessimistic approximations were used in the above calculations of the error due to a charge variation σ_{q_i} in each pixel. This is intended to compensate for the likely underestimation in the charge fluctuation. This was calculated using the geometric path length in each pixel, averaging over all intercept positions (so for a 2-pixel cluster the charge error in each pixel is calculated using half the total path length). For the purposes of comparison with data, a charge digitisation error was further added at the level of 1 % of the total charge deposited, to reflect a typical pixel charge of 40 ToT counts. This gives the total charge error on each pixel to be

$$\sigma_{q_i}^2 = \sigma_{\mathrm{Landa-Vavilov}}^2 + \sigma_{digitisation}^2 \tag{4.27}$$

4.2.3 Comparison with data

The above expressions have been used to calculate the expected resolution versus angle for different conditions that match the available data. There are no free parameters included in the analytic model that can be tuned, though given the level of knowledge about the substrate resistivities and absolute values of the threshold, etc., there is some scope for adjustment between the predicted and observed behaviour.

Figure 4.18 shows the output of the analytic model with the experimentally observed detector resolution for D04-W0015. The results are shown separately for each bias voltage for clarity. As with the experimental results above, the telescope pointing resolution is not subtracted, and for the analytic results a pointing precision of 3 µm is used (due to the fewer number of planes and higher material content during data-taking with D04-W0015).

The agreement between the model and data is reasonable, with resolution values obtained from the analytic expression generally much closer than 1 µm to the experimentally measured values. The shape of the response versus angle is also well described, though there appears to be an overestimation of the resolution given by the analytical expression for large angles. There is not enough experimental data to observe how this trend continues, but it is likely that at sufficiently large angle the Landau-Vavilov function will better describe the energy loss spectrum, and the assumptions made in the expressions for the pixel error will prove pessimistic.

The model can now be used to investigate the effects observed in data, taking advantage of the ability to keep all but one parameter fixed.

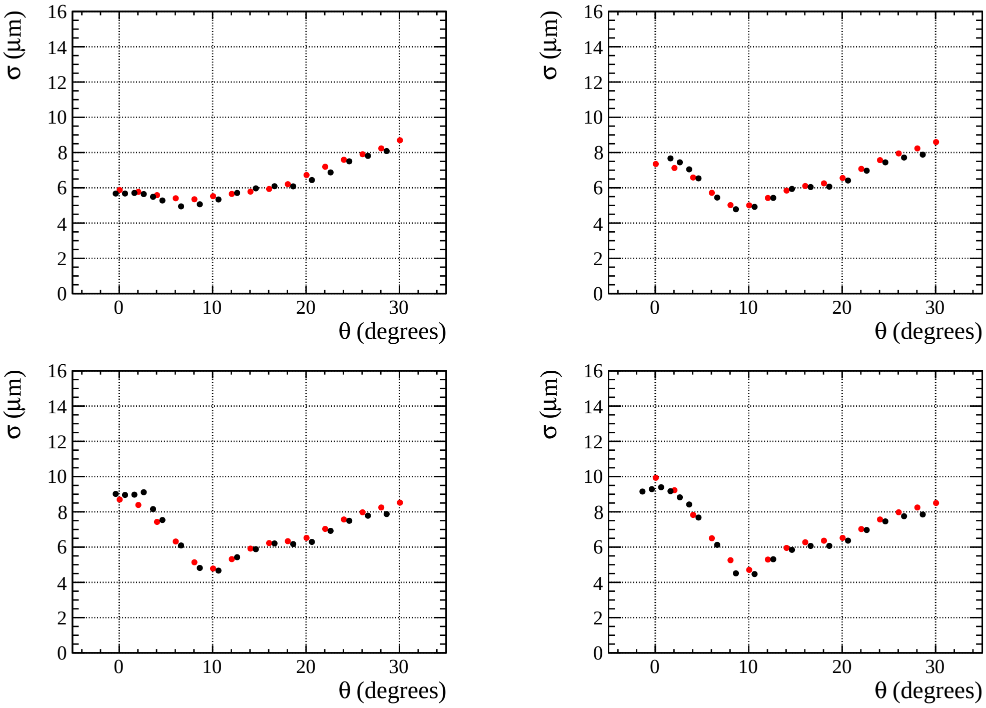

Figure 4.18: Measured (black) and predicted (red) charge-calibrated resolution versus angle for device D04-W0015, at 10 V (top left), 20 V (top right), 40 V (bottom left) and 60 V (bottom right).

4.3 Trends in Detectors Design

Given the range of combinations between the substrate type, resistivity, bias voltage, pixel size, etc. that can be varied in the detector design, it is useful to limit the scope of which sensor parameters are considered. In the following, the various parameters will be varied sequentially, following the direction which points towards improved spatial resolution and/or the more global trends in particle physics experiments (specifically the drive towards thinner devices). In order to decouple the effects of the measurement method, no error on the extrapolated position is added to the analytic model. Similarly, the calculations have been performed using the current front-end electronics unless otherwise specified, and where the technology exists to construct such a device at present. For this purpose the threshold has been kept at 1000 electrons, and the pitch at 55 μm, for all scenarios with the existing Timepix chip.

4.3.1 Interpretation of Experimental Results

The first parameters considered are those that were observed to have a significant effect in the experimental data, in order to confirm an understanding of the effects involved. An extrapolation towards future devices is then conducted.

Resistivity

For the samples measured in testbeams and presented above, one device was manufactured with a significantly higher resistivity substrate (D04-W0015). The performance of this device at bias voltages close to depletion is significantly better than that of any other device presented. That being said, the device thickness is also greater, which should significantly improve the lateral charge spread. Considering a device of such thickness, the resistivity (and therefore depletion voltage) may play an important role in determining the spatial resolution. In practice, there is a trend in particle physics for devices to be operated around 20 - 40 V above depletion, regardless of the absolute value of the depletion voltage. It can then be questioned if there is a significant gain to be made by choosing a high resistivity substrate and operating significantly over-depleted, rather than the typical lower resistivity employed, operated only mildly over-depleted.

Two hypothetical devices, similar to D04-W0015, are considered. Both are taken as having a thickness of 300 µm and being constructed out of p-on-n sensors, with pitch and threshold given by the existing Timepix electronics. The resistivities used are 4 kΩcm and 40 kΩcm, giving depletion voltages of 79 V and 7.9 V respectively. For the comparison of their performance, the calculations for the lower resistivity device are performed at depletion (80 V) and mild over-depletion (100 V), as might be used in the detector operation in practice. For the higher resistivity sensor, the performance at depletion (8 V) and significantly over-depletion (80 V) is shown. This allows a direct comparison of the devices under equal operating conditions, where the low-resistivity sample is only just depleted. The electric field profiles for these scenarios are shown in figure 4.19.

The calculated lateral diffusion parameter is shown in figure 4.20 for the scenarios described. The left hand plot illustrates the expected diffusion for the lower resistivity sample, showing that even at depletion the lateral diffusion is quite severely limited, with very little change in performance with operation at 20 V over depletion. The higher resistivity sample on the other hand shows significant lateral diffusion, in addition to a large variation with bias voltage. Even when operated at the same bias (80 V) as the lower resistivity sample, the device shows similar lateral charge spread despite being 70 V over-depleted.

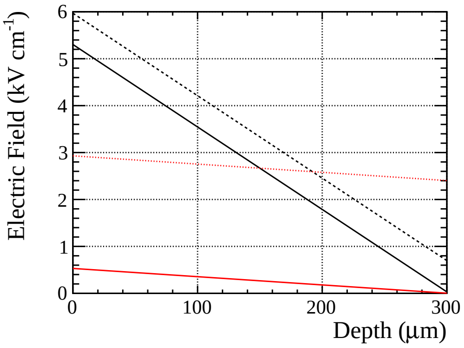

Figure 4.19: Electric field profiles for two hypothetical p-on-n sensors, with substrate resistivity 4 kΩcm (black) and 40 kΩcm (red). Fields are shown for depletion voltage (80 V and 8 V, solid line) and over-depletion (100 V and 80 V, dashed line).

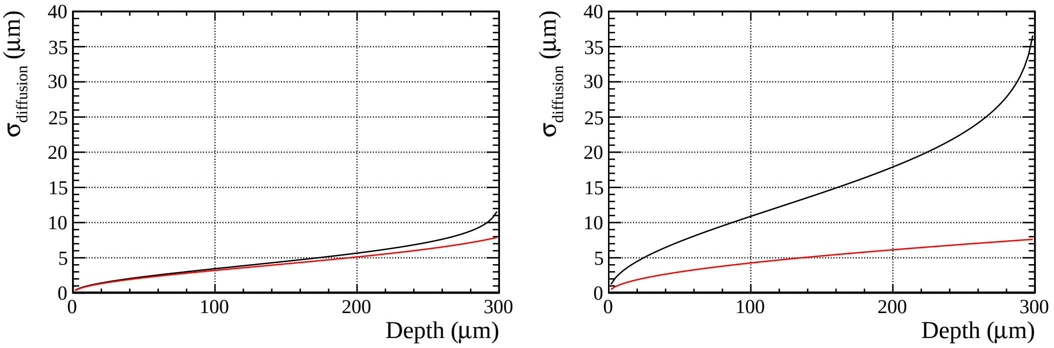

Figure 4.20: Diffusion parameter versus depth, calculated for devices with substrate resistivity 4 kΩcm (left) and 40 kΩcm (right) and operated at depletion (80 V and 8 V, black) and over-depleted (100 V and 80 V, red).

These differences in diffusion are reflected in the calculated resolution versus angle, shown in figure 4.21. On the left hand plot, the lower resistivity device can be seen to perform poorly even when operated at depletion, with no significant change with the application of higher bias. The higher resistivity device however, behaves much as D04-W0015, and shows a significant improvement in resolution for operation at depletion. The resolution remains approximately constant through the diffusion-dominated angular range, only degrading when it reaches the geometrically-dominated angles above 10°. When operated significantly over-depleted, at the same bias as the lower resistivity sample, some small (0.5 μm) difference is observed at perpendicular incidence for the high resistivity sensor. Despite this rather extreme operation (70 V over-depleted), it is clear that significant gains might be achieved as far as spatial resolution is concerned, in the drive towards higher resistivity substrates. Not only in the gains at equivalent voltage, but the prospects for lower operation voltage and the considerable improvement in resolution that can be achieved.

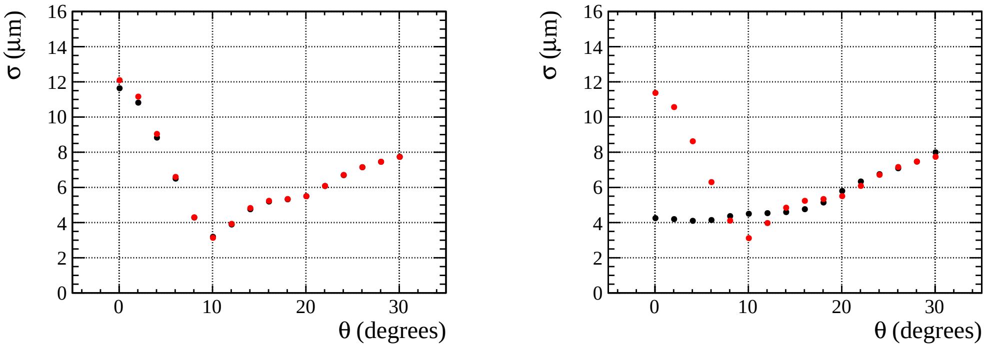

Figure 4.21: Calculated resolution versus angle for devices with substrate resistivity 4 kΩcm (left) and 40 kΩcm (right) and operated at depletion (80 V and 8 V, black) and over-depleted (100 V and 80 V, red).

Technology Type

Given the observations on the effects of substrate resistivity above, it is likely that for the samples tested in the testbeam that this effect dominated over the other contributions to the device resolution. The generally poor (8 - 12 μm) resolution at perpendicular incidence seems, to first order, to be attributable directly to the range of resistivities in the measured samples. Differences due to technology type would be expected to be more significant for higher resistivity samples, where the electric field is much lower and larger changes are observed with variations in the field. For this reason, and due to the significant gains in performance for higher resistivity substrates, the following calculations on bulk and implant type have been performed to match the device described in the previous section.

In equation 4.8 there are no terms related to the carrier type, only the bulk mobility, resistivity and depletion depth. These terms dictate the electric field (and therefore diffusion) throughout the sensor, and so for devices produced with different bulk types the only change expected would be from the ratio of (resistivity × mobility). Comparing an n-on-p to p-on-n sensor, the differences brought about by the change in bulk mobility could therefore be mitigated by increasing the substrate resistivity by a factor of $\mu_{\text{hole}}/\mu_{\text{electron}}$. In all other respects the devices are expected (from the model presented) to behave in the same fashion, and so such a comparison is not considered further.

For an n-type bulk sensor with n-type implants, this is a very different scenario. The reversal of the field direction, as mentioned previously, augments the diffusion of charge deposited close to the pixel implants (where the collection path is low), while depressing the diffusion of charge close to the back-side. Two sensors are again hypothesised for comparison, the p-on-n sensor described in the previous section and an n-on-n device with equivalent thickness (300 μm) and resistivity (40 kΩcm). As the location of the depletion zone is directly in front of the pixel implants, and a failure to fully deplete the sensor will result in a non-depleted region reducing the charge collection, three field configurations are used for the comparisons: depletion (8 V), slightly over-depleted (20 V) and significantly over-depleted (80 V). The electric field profiles are displayed in figure 4.22.

As before, the calculated lateral diffusion parameters for both devices can be seen in

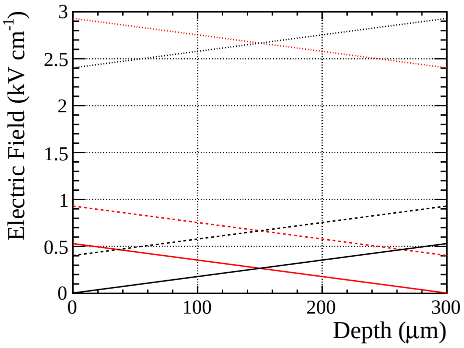

Figure 4.22: Electric field profiles for two hypothetical sensors, with sensor type n-on-n (black) and p-on-n (red). Fields are shown for depletion voltage (8 V, solid line), mild over-depletion (20 V, dashed line) and significant over-depletion (80 V, dotted line).

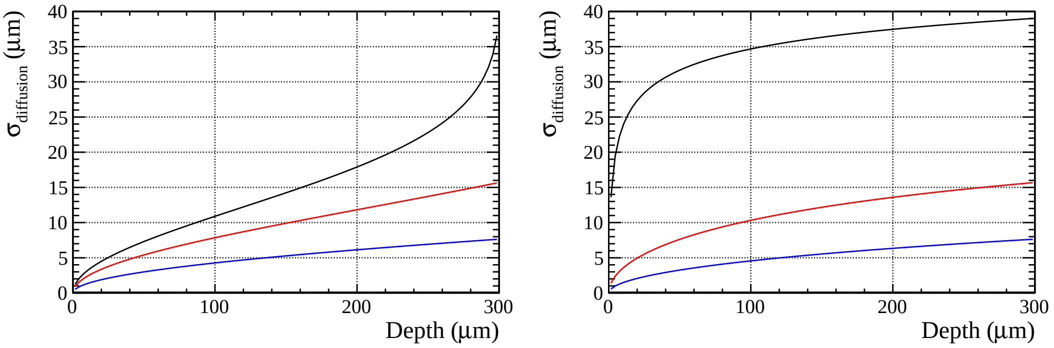

Figure 4.23: Diffusion parameter versus depth, calculated for an n-on-p (left) and n-on-n (right) sensor, operated at depletion (8V, black) mildly over-depleted (20 V, red) and significantly over-depleted (80 V, blue).

figure 4.23. The n-on-p device (left hand plot) is the same as that presented previously, with the addition of results at 20 V. The right hand plot shows the expected diffusion for the n-on-n device, and for operation at depletion shows a significantly larger increase in diffusion, across the full sensor depth. This is not surprising considering that all charge in this case must pass through the region of virtually zero field close to the pixel implants. The more realistic (and practical) scenarios are for 20 V and 80 V. For the former, the difference between the two devices is less marked, though in the region closest to the pixel implants the n-on-n sensor still shows a distinct increase in diffusion. This reduces for the higher applied bias, where both sensors appear to behave roughly the same.

The calculated resolution for both devices again follows the expected diffusion (and can be seen in figure 4.24). For significant over-depletion there is no real difference between the two devices, while for the realistic case of mild over-depletion the n-on-n sensor shows a distinct improvement of around 2 μm at perpendicular incidence. The predicted performance at depletion shows a drastic improvement for the n-on-n sensor, but this should be treated with more caution as the model contains no description of the electric field focussing around the pixel implants. While it might be expected to be

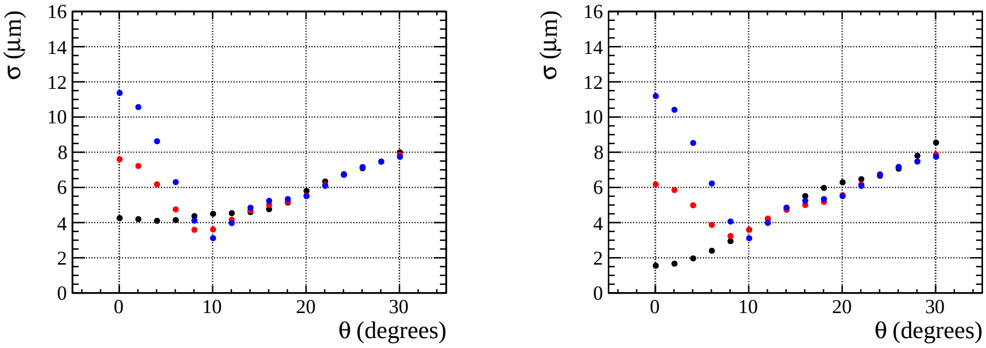

Figure 4.24: Calculated resolution versus angle for an n-on-p (left) and n-on-n (right) sensor, operated at depletion (8 V, black) mildly over-depleted (20 V, red) and significantly over-depleted (80 V, blue).

broadly correct, the results for mild over-depletion should be taken as a more realistic demonstration of the gains possible.

4.3.2 Future Directions

Thickness

Current sensor-thinning technology has been able to produce sensor thicknesses down to 50 µm. As the trend for reduced material in particle physics experiments is in this direction, the question of how to improve the detector resolution is considered. From the previous section, the advantages of n-on-n sensors in encouraging diffusion for charge deposited close to the pixels would place this as the technology of choice.

Three n-on-n sensors are considered, with equal resistivity (40 kΩcm), readout electronics and thicknesses of 300 µm, 100 µm and 50 µm, leading to depletion voltages of 8 V, 0.9 V and 0.2 V respectively. Each device is compared for three biasing conditions: at depletion, mild over-depletion and significant over-depletion. As the electric field in an n-on-n sensor extends from the junction at the back side, the voltage used in the calculation of the device performance is varied in order to keep the field approximately constant with respect to the region in front of the pixel implants. The fields for each device are shown on the left hand plot of figure 4.25. All sensors are operated at or slightly above their depletion voltages (8 V for the 300 µm sensor, 1 V for the 100 µm and 0.25 V for the 50 µm), with the small differences showing larger deviations as the sensor thickness reduces. For the mild over-depleted scenario, voltages of 20 V, 5 V and 2.25 V are used, while for the higher field voltages of 80 V, 25 V and 12 V are used.

The corresponding calculation of the lateral drift parameter for each device is shown on the right hand side of figure 4.25. While the drift parameter for the thinnest sensor is still large at depletion, it is worth noting that with no change in the electronics this will mean a signal to threshold ratio of just over 3.5. Significant charge diffusion is likely to result in possible inefficiencies, with a majority of single pixel clusters despite this lateral diffusion. This is confirmed in the calculated resolution plots shown in

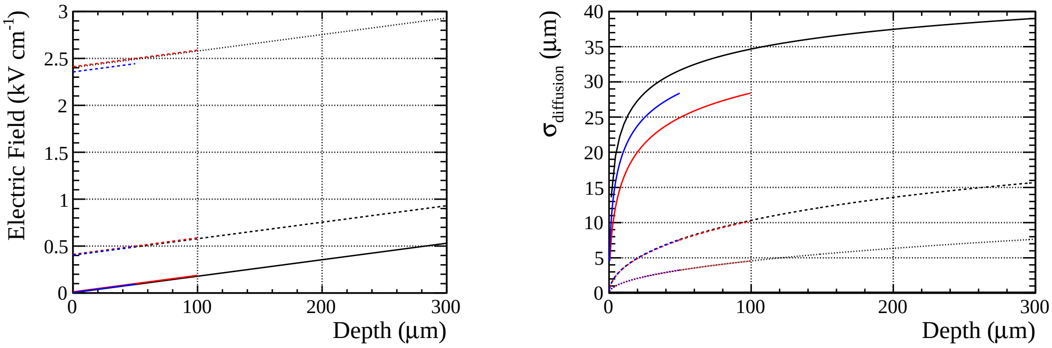

Figure 4.25: Left: Electric field profiles for three hypothetical n-on-n sensors, of thickness 300 µm (black), 100 µm (red) and 50 µm (blue). Fields are shown for depletion voltage (8 V, 1 V and 0.25 V, solid line), mild over-depletion (20 V, 5 V and 2.25 V, dashed line) and significant over-depletion (80 V, 25 V and 12 V, dotted line). Right: Calculated diffusion parameter versus depth.

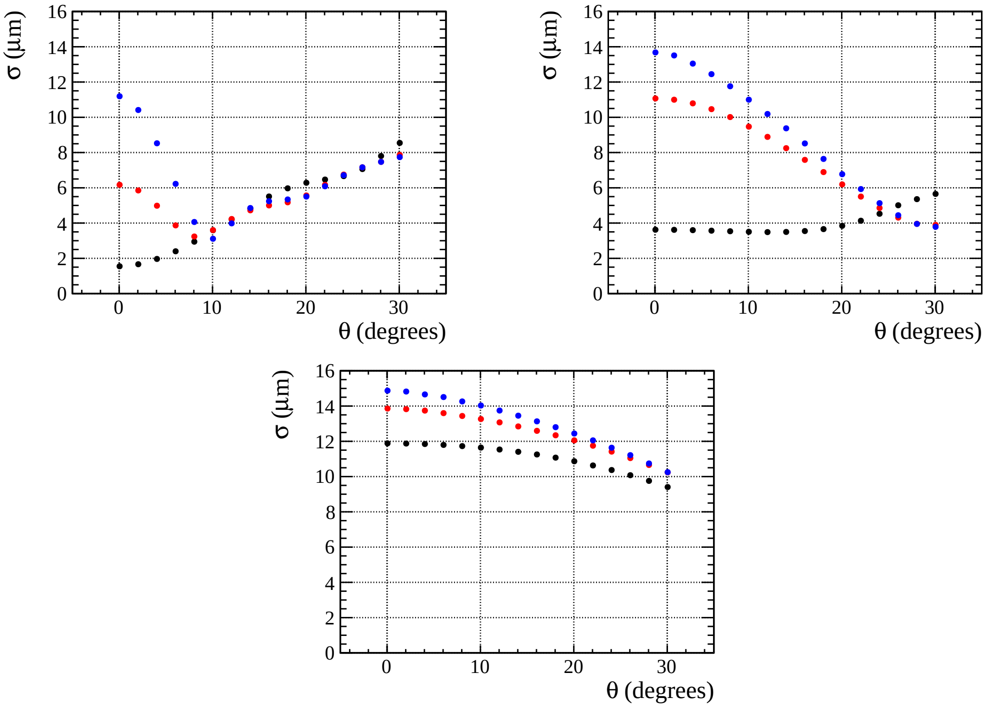

Figure 4.26: Calculated resolution versus angle for three n-on-n sensors, of thickness 300 μm (top left), 100 μm (top right) and 50 μm (bottom), with applied bias around depletion voltage (8 V, 1 V and 0.25 V, black), mildly over-depleted (20 V, 5 V and 2.25 V, red) and significantly over-depleted (80 V, 25 V and 12 V, blue).

figure 4.26. What is interesting about the variation in response between the different devices is the performance of the 100 μm thick sensor close to depletion. While thick sensors have the general advantage of multi-pixel clusters (in addition to better signal-to-noise ratio), and the subsequent improvement in resolution which this facilitates, the degradation due to Landau fluctuations in the geometrically-dominated region begins at 10°. Thinner sensors switch into this regime at much higher angles (around 29° for a 100 μm thick sensor with 55 μm pitch), giving a flatter response and improved resolution for many angles, as long as sufficiently low biases are applied. The extreme switch in performance observed going to the 50 μm sensor suggests that significant changes to the front end electronics are required in order to take full advantage of such thin substrates.

Pixel Size

As discussed, the general reduction of charge sharing and loss of resolution when equipped with thinner sensors has led to the use of smaller pitch electronics, in order to recover some of these losses. The two important changes which are generally introduced with subsequent iterations of pixel chips for particle physics, in addition to more complicated functionality or increased output bandwidth, are improved granularity and lower threshold. Both are important if there is to be a complete shift to sensors below 100 μm thickness.

Using the parameters listed in the previous section for the 50 μm sensor (including

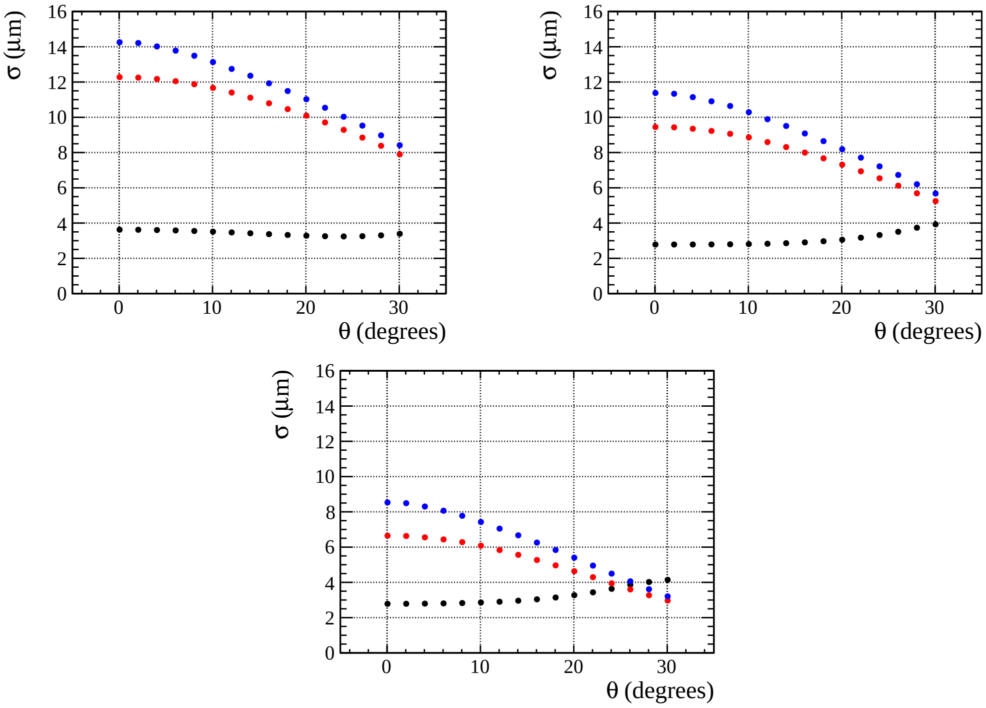

Figure 4.27: Calculated resolution versus angle for three 50 μm n-on-n sensors with pixel pitch of 55 μm (top left), 45 μm (top right) and 35 μm (bottom), with applied bias around depletion voltage (0.25 V, black), mildly over-depleted (2.25 V, red) and significantly over-depleted (12 V, blue).

bias voltages), three devices featuring updates with respect to the current Timepix chip are considered. Each has a reduced threshold of 500 electrons, and a pixel pitch of 55 μm, 45 μm or 35 μm. The calculated resolution versus angle for all three can be seen in figure 4.27, for 0.25 V, 2.25 V and 12 V. The difference in resolution from simply lowering the global threshold to 500 electrons (top left plot, compared with that of figure 4.26) is significant, particularly close to depletion. The further gains due to reducing the pixel pitch are better than would be expected from simple scaling, and a flat response of roughly 3 μm is observed in the best case.

4.4 Summary

The single hit resolution for a variety of silicon sensors mounted on Timepix ASICs have been measured, and the differences between varying thickness, substrate resistivity, implant/bulk type and operating conditions have been investigated. The bias voltage is observed to have a strong effect on the single hit resolution for high resistivity substrates, where operation close to their depletion voltages can significantly improve the spatial resolution (from 8 μm to 4 μm in the device shown). At sufficiently large angles the bias voltage is however found not to play a role in determining the device performance, once the lateral track path exceeds the pixel pitch. Variations due to the bulk/implant type and depletion side have been shown to be small, with the low resistivities of the samples tested dominating.

To complement the experimental measurements, an analytic model has been constructed to describe the single hit resolution. This has been observed to match well the experimental data, and is used to compare different sensor designs in order to determine the individual contributions of several effects. The model has also been used to make some statements on the likely direction of travel for future sensor development.

5

Sensor Resolution Studies Using the FE-I4 Front End Chip

Despite the success and progress made in sensor testing with the Timepix ASIC, the Timepix does not fulfil the full requirements for prototyping the VELO upgrade. Ignoring changes required to the ASIC operation, which require the design and production of new ASICs (the Timepix3, followed by the final VELOPix), the most important limitation to sensor testing posed with the Timepix ASIC is radiation tolerance. The maximum fluence in the VELO is found at the stations around the interaction point, and is radially distributed as shown in figure 5.1 (for the end of detector lifetime). The inner region of the chip will suffer a total dose of around 8 - 9 × 10^{15} 1 MeV n_{eq} cm^{-2} , giving surface damage equivalent to about 370 Mrad [62]. The existing Timepix is radiation tolerant to around 200 krad [63], with the result that sensors cannot be irradiated to such levels without destroying the ASIC. Bump bonding of the sensor after irradiation is possible, but the temperatures involved would cause a significant amount of reverse annealing, limiting the usefulness of any such tests.

To allow initial tests to be carried out on the sensor performance after irradiation to these levels, a radiation-hard ASIC is required. A collaboration between LHCb and the ATLAS Inner B-Layer (IBL) Planar Pixel Sensor (PPS) group was established, due to the ATLAS development of the Front-End IBM 4 (FE-I4) chip [64] to be installed during the first long shutdown of the LHC. This chip has been used to test heavily irradiated sensors [65] for use in the IBL, and while the chip dimensions and functionality do not

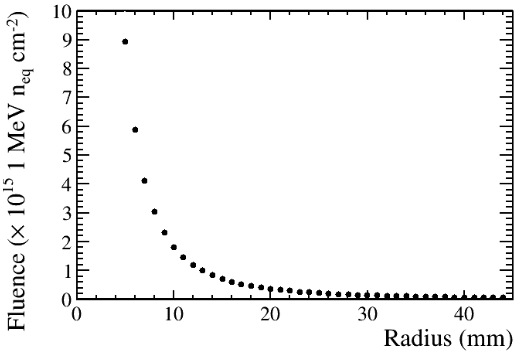

Figure 5.1: End of lifetime fluence across a single VELO detector module close to the interaction point.

match those required by the VELO, it allowed the testing of irradiated sensors which could be fed back into the VELO upgrade design.

5.1 The FE-I4

5.1.1 The FE-I4 ASIC

The FE-I4 is a hybrid pixel ASIC designed in 130 nm CMOS process, the same proposed technology for the Timepix3 and VELOPix ASICs. It contains an array of 80 columns and 336 rows with a pixel dimension of $250 \times 50\,\mu\mathrm{m}$, giving it a comparable pitch to the Timepix in the row direction and a total area of 20.0×16.8 mm^2. The deposited charge is measured by the Time Over Threshold method as for the Timepix, with the result encoded in a 4-bit counter. The pixels contain both analogue and digital functionality, and the full ASIC is expected to cope with a trigger rate of around 75 kHz (the ATLAS level 1 trigger acceptance rate [66]), with raw hit rates of around 400 Mhits cm^{-2} expected. Two versions of the FE-I4 have been produced: a prototype version FE-I4 A, and the final production ASIC FE-I4 B.

A schematic of the analogue front-end of the FE-I4 A is shown in figure 5.2. The pixel cell contains a two-stage amplification process, designed to allow for lower power consumption, faster rise times and increased protection against sensor leakage current after heavy irradiation [64]. The ASIC is arranged into 40 double-columns, where the digital regions in the two columns are neighbouring such that information may be shared between them, allowing local buffering of event data. If a pixel discriminator crosses threshold, the ToT reading and time of the event are stored locally for each 2 × 2 pixel region, enabled by the neighbouring digital regions of these four pixels. If a trigger is sent to the chip during the data latency period, all data matching the trigger time are sent out.

Similarly to the Timepix ASIC, a threshold tuning is performed to remove channel-to-channel variation and reduce the overall threshold dispersion. The FE-I4 performs this by a two-stage iterative process. The chip is tuned to an overall threshold, obtained by an injection of charge directly to the front-end with the value of the required

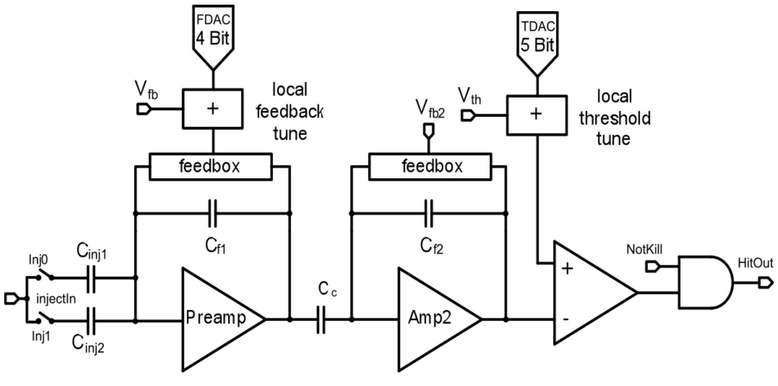

Figure 5.2: Schematic of the analogue section of the FE-I4 A pixel ASIC. Figure taken from [64].

threshold, and varying the local 5-bit tuning DAC (TDAC). The gain is then controlled by injecting a (higher) fixed quantity of charge and varying the discharge current so as to observe this at a specified number of ToT counts. This is again varied locally at the pixel level, with a 4-bit DAC (FDAC). When quoting the threshold tuning both values are given, so a chip might be said to have been used with a 1 ke$^-$ threshold tuned to 10 ke$^-$ at 10 ToT.

In the prototype version of the chip, the FE-I4 A, the capacitances used for charge injection during the threshold tuning are estimated from a comparison with source data (typically a β-emitter). This gives a poor estimate of the true capacitance, and significant variation between chips has been observed. This is rectified in the production version of the chip (FE-I4 B) with the introduction of dedicated circuitry to measure the injection capacitances. Where prototype chips have been employed, the data have been kept in ToT counts, with the tuning shown to provide an estimate of the charge collected. This is expected to vary by at least 20 %.

5.1.2 FE-I4 DAQ

To read out the FE-I4, a system based on the Reconfigurable Cluster Element (RCE [67]) built by SLAC [68] was used. This system is split into a back-end readout unit (the RCE) and a front-end High Speed Input/Output (HSIO) which can be placed near the device under test. While the back-end processes data into recorded events and allows for online monitoring of the DAQ, the HSIO is self-sustained, feeding triggers into the DUT, reading back the data and converting it into the RCE input format. The RCE allows control of the DUT in order to equalise and configure the FE-I4.

The local buffering of data within the 2 × 2 pixel regions on the chip are designed to allow time for a global first level trigger decision about the event to be taken, as it is not possible to push the complete event data off-detector. This is suited to the LHC bunch structure, where events are collided with an integer spacing of 25 ns. For asynchronous events, where the FE-I4 clock cannot be tuned to give a fixed phase with respect to events, it is possible that hits may occur in the 25 ns preceding or succeeding the expected time window. To alleviate this, the FE-I4 DAQ can request multiple time bins to be read out for each trigger received. The timestamp at the pixel level, running with a 40 MHz clock, is referred to as the Bunch Crossing ID (BCID).

5.1.3 Sample Description

The results presented were performed on three detector assemblies, all constructed using the prototype FE-I4 A ASIC. An unirradiated sample was used to evaluate both the tracking performance of the FE-I4 pre-irradiation (in particular observing the 4-bit ToT charge measurement) and the integration of the device with the Timepix telescope. Two additional samples, both irradiated to fluence levels of interest to the VELO upgrade, were also tested, with a focus on the single hit resolution and charge collection properties. At this time the expected thickness of sensor to be used in the upgraded VELO was 150 µm, which has since been reconsidered and is now expected to be 200 µm. For this reason the irradiated sensors measured are thinner than the current VELO proposal.

The unirradiated FE-I4 is bump-bonded using lead-tin solder to a 200 µm thick silicon sensor, with p-type bulk and n-type implant. The depletion voltage is 50 V, and the device was operated with a bias voltage of -70 V. The two irradiated sensors were both produced at the semiconductor laboratory of the Max-Planck Institute for Physics in Munich (HLL) [69], with the bump-bonding performed at the Fraunhofer Institute (IZM [70]) in Berlin using lead-free (indium) solder. They are fabricated using 2 kΩ cm p-type wafers thinned to a thickness of 150 µm, with n-type implants. Irradiation of both devices occurred at different locations, with one sample irradiated to 2×10^{15} 1 MeV $\mathrm{n_{eq}}\ \mathrm{cm}^{-2}$ with 25 MeV protons at the Compact Cyclotron in Karlsruhe (KIT) [71] and the other irradiated to 4×10^{15} 1 MeV $\mathrm{n_{eq}}\ \mathrm{cm}^{-2}$ with 800 MeV protons at LANCSE [72] in Los Alamos. As this work was carried out before the decision was taken to reduce the inner radius of the upgraded VELO detector, this corresponded to 25 % and 50 % of the expected end-of-lifetime fluence in the hottest regions.

5.2 Device Integration

The telescope operation with the FE-I4 remains unchanged in most respects, although the requirement of timing information on the tracks (not necessary for the analysis of a Timepix sensor) acts as a limit on the maximum achievable event rate. As described in previous sections, the time-tagging requirements determine the maximum shutter length for a given beam intensity. If the beam rate is low enough such that the FE-I4 can absorb the raw trigger rate of the beam, then there are no changes with respect to Timepix-only running. If on the other hand the raw trigger rate exceeds that which can be handled by the FE-I4, then a correspondingly smaller fraction of events will be recorded. The telescope shutter must be closed once the occupancy or shutter-length requirements have been met, regardless of the number of FE-I4 accepted triggers.

For the FE-I4, the coincident scintillator triggers are gated with the telescope shutter and passed directly to the RCE. In return, a busy signal can be (optionally) used. This was found necessary due to the high instantaneous rate of the beam at the SPS North Area, and these two lines of communication remain the only hardware-level integration of the two systems. The busy signal was additionally used on the telescope side to veto any raw scintillator triggers that were not passed to the FE-I4: with hindsight, this requirement should be removed (discussed in section 5.3.1).

The offline integration is performed by matching the trigger times within each shutter from the TDC and FE-I4. By looking at the time difference between triggers, rather than the absolute times, delays and other problems such as dropped triggers/frames are avoided. The reconstruction sequence pairs the TDC time with a ToA cluster located on the track, attaching data from the FE-I4 in the process. The analysis of the FE-I4 DUT can then proceed.

Due to the asynchronous nature of the incoming beam, a number of BCIDs were read out for each trigger sent to the FE-I4. As the data rate coming out of the chip increases with the number of BCID read out, this was limited to 4. A latency scan was performed in order to account for the difference in relative propagation times between the systems.

5.3 Unirradiated Performance Studies

A quick integration test was performed with the unirradiated n-on-p detector, operating at 70 V. The device was rotated around the short pixel axis, increasing the path length in the 250 µm column direction. Data was taken for some representative shallow angles and additionally for tracks almost parallel to the sensor surface (due to the use case in the barrel-shaped IBL). The ToA was clocked at a frequency of 2.5 MHz in order to extend the shutter opening time and maximise useful data collection, giving the ToA precision of 400 ns for this data set. The general detector performance was investigated by the cluster size, resolution and single hit efficiency. In particular, the efficiency measurement was intended to highlight any issues with the device integration. Using an over-depleted unirradiated silicon sensor was deliberate for this reason, as the efficiency should be 100 %.

5.3.1 Efficiency and Integration Results

After analysing the data taken, it is immediately apparent that there is an issue in the integration of the FE-I4 with the telescope. Before considering the FE-I4, the association of tracks with scintillator hits is performed, using the ToA and TDC times respectively. The residual distribution for associated events, after the isolation vetoes described previously have been applied, is shown in figure 5.3. It is noticeable that the peak is distinctly asymmetric, and by a much greater degree than can be explained by timewalk or any phase differences with respect to the ToA clock. The efficiency of the FE-I4 is observed to change markedly depending on any cut in this distribution, yet still lies below 100 % when a window of 400 ns (the clock period) is used.

The source of this error is the veto accepted by the telescope while the FE-I4 is unable to accept new triggers. While this is used to prevent data loss in the FE-I4, for this data set the veto was also applied to the TDC input, such that only triggers that were transmitted to the FE-I4 are recorded. When this occurs and the veto remains high (around 1 µs), a second track can pass through the telescope and scintillators. As no timing information is used in the track reconstruction, both tracks can be potentially reconstructed, and an ambiguity arises regarding which one to assign to the TDC time. This is usually solved by the isolation veto: the two tracks would have times close to each other and so both would be discarded from further analysis. The problem arises

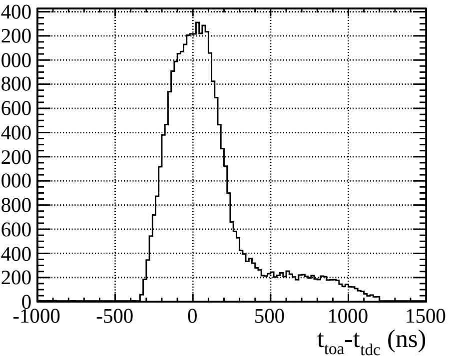

Figure 5.3: Timing residuals between the ToA plane and TDC for associated events.

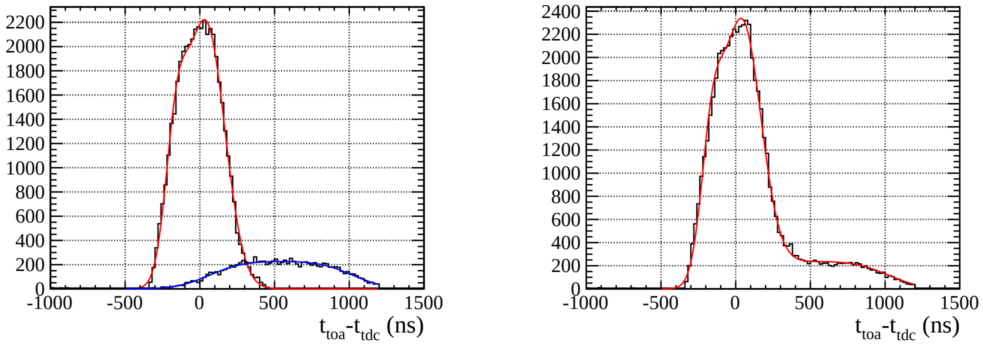

Figure 5.4: Left: Individual fits for signal (red) and background (blue) in the timing residual distributions for where an associated FE-I4 cluster was or was not observed. Right: Combined fit of the timing residual distribution to extract the purity.

when the first track has not been reconstructed. When this happens, there appears to be an isolated TDC time, along with a time-isolated track. These are subsequently matched, and explain both the direction and asymmetry of the tail in the residual distribution. The TDC always belongs to the earlier of the two tracks which have passed, while the second track arrival will be some time period over the next 1 µs while the busy signal remains high.

In order to measure the efficiency of the FE-I4 without retaking this data, an estimate had to be performed of the *purity* of each data set, and a fit was performed to the timing residual spectrum in order to measure the quantity of such mis-associated tracks. The shape of both the “signal” and “background” is described analytically as a top-hat function with Gaussian edges, namely the sum of two error functions. The signal additionally may have contributions from tracks arriving close to the clock edge or suffering from timewalk, and so a further Gaussian was added to the signal distribution. In order to fit the signal and background parameters, the data taken at perpendicular incidence was used. The residual distributions for tracks where an FE-I4 cluster *was* observed were fitted as signal events, while those where no corresponding cluster was found were fitted as background. These two individual fits are shown on the left hand plot of figure 5.4. These parameters were then fixed, so that the original distribution could be fitted allowing only the relative amount of each to vary (right hand plot). The fit was normalised by the number of tracks used for the efficiency measurement.

For all subsequent runs, the raw efficiency could thus be combined with a measured purity value of the sample, and the “true” efficiency measured. The results for moderate rotation are shown in figure 5.5, where the efficiency can be observed to be around 100 % in all cases as expected. For subsequent data-taking periods the veto from the RCE was not used to prevent triggers being sent to the TDC, mitigating this issue. The high timing information of the RCE (roughly 1 ns) allows accurate matching of TDC and FE-I4 data even if the TDC data contains a larger number of triggers, as the time differences between accepted triggers during a frame can still be accurately measured.

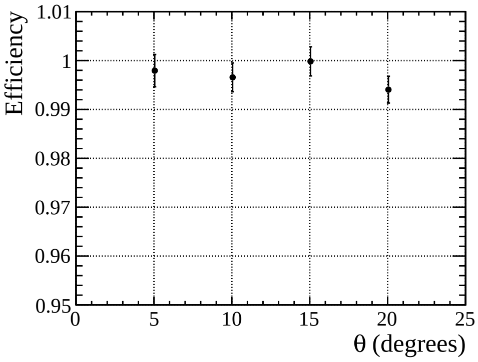

Figure 5.5: Measured single hit efficiency versus incident angle.

5.3.2 Tracking Performance

The cluster width in the direction of rotation and the single hit resolution for this device can be seen in figure 5.6. The resolution is shown in the 250 µm direction, where the binary resolution is 72 µm, and the rotation axis is parallel to the short pixel edge. At perpendicular incidence the response is better than binary as expected, given the 4-bit ToT information available, but the sensor does not exhibit significant charge sharing. Both the extreme pitch, compared with typical diffusion lengths of some microns or tens of microns, combined with the thin substrate both act against the direction required for better precision. In this case the gains, from purely the single hit resolution perspective, of the 4-bit charge information are not likely to compensate for the increased bandwidth required to transmit the data off-detector.

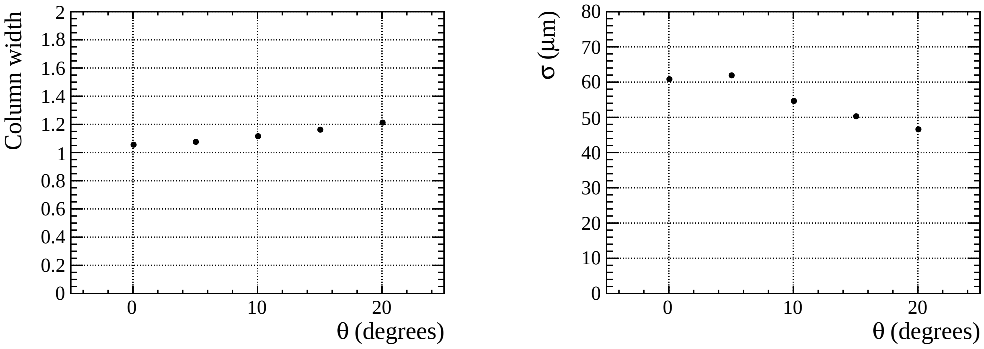

Figure 5.6: Left: Cluster size versus angle in the direction of rotation. Right: Single hit resolution versus incident track angle.

5.4 Irradiated Performance Studies

Both of the irradiated sensors were operated in the testbeam with dry ice cooling to prevent annealing and allow the application of high voltage. The detectors were mounted in the rotation axis inside a thermally-insulated box with two compartments, one containing the mounted assembly and the other the dry ice. The cooling power was transferred using high-conductivity copper braid, affixed to a thin aluminium layer mounted directly under the ASIC. Accurate temperature measurements were not available, but as the sublimation temperature of dry ice is -78.5 °C and a temperature sensor close to the ASIC registered around -50 °C, the sensor temperature was likely between -40 °C and -50 °C.

5.4.1 Sensor Properties

After irradiation and before testing in the SPS North Area, the devices were both characterised by the ATLAS PPS group in Munich. Leakage current measurements were carried out in a dry environment at -50 °C, and tests with a ^{90}Sr source were performed. The exact annealing of both samples is not accurately known, due in part to the sample transport and cooling during the irradiation. Tuning of the front-end for both samples was carried out for the lab measurements, using a threshold of 1 ke$^-$ for the lower irradiated sample and 2 ke$^-$ electrons for the higher. Both were tuned such that the charge to ToT conversion was 10 ToT at 10 ke$^-$.

The leakage current measurements, taken at a temperature of -50°C, are shown in figure 5.7. Predictions for the leakage current before annealing are 2.8 µA for the sample irradiated to 2×10^{15} 1 MeV n_{eq} cm^{-2} and 5.6 µA for the sample irradiated to 4×10^{15} 1 MeV n_{eq} cm^{-2} , both under the assumption of full depletion (where approximate values of effective depletion voltage from the Hamburg model are ~ 1.5 kV and ~ 3 kV respectively). For the lower irradiated sample, the IV curve is observed to reach a plateau just before 200 V, after which there is a slow linear rise in the current measured. This is likely due to the increase in field throughout the sample, which reduces the drift time of the charge carriers. Since the trapping probability is proportional to the mean free time, faster charge collection reduces the total amount of trapped charge, giving an increase in observed leakage current. For the higher irradiated

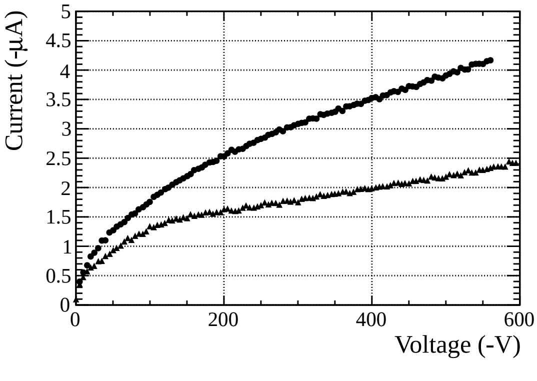

Figure 5.7: Measured IV curves taken at -50°C for 2×10^{15} (triangles) and 4×10^{15} (circles) 1 MeV n_{eq} cm^{-2} irradiated samples.

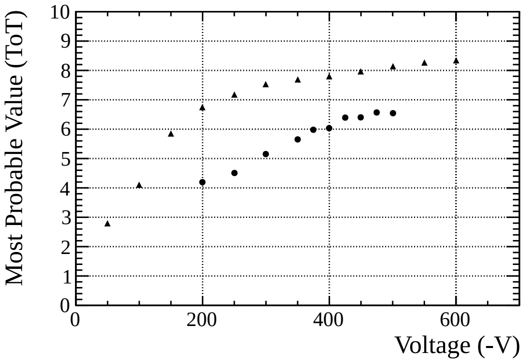

Figure 5.8: Measured charge collection curves taken with ^{90}Sr for 2×10^{15} (triangles) and 4×10^{15} (circles) 1 MeV n_{eq} cm^{-2} irradiated samples.

sample the whole IV curve is much less distinct. In addition to showing higher overall current, the onset of a plateau is not cleanly displayed, although the increase in current does appear to become linear after roughly 400 V.

To determine the operating points of the samples for the testbeam measurements, source measurements were taken in the lab using a ^{90}Sr source. The MPV of a convoluted Landau-Gaussian fit to the observed spectra is plotted in figure 5.8 versus bias voltage. As with the leakage current measurements, there appears to be a turning point in the charge collected with the lower irradiated sample, with a crossover point around 200 V. After this there is a marginal increase in the measured charge as the field strength within the sample reduces trapping. For the higher fluence assembly a saturation in detected charge appears just after 400 V, with some gradual increase shown for the bias points above this voltage. For operation in the testbeam, an operational voltage of 400 V was chosen for both devices.

5.4.2 Tracking Performance

Both samples were mounted such that the rotation axis aligned with the long pixel edge, allowing the track angle in the 50 μm direction to be varied. Tuning of the front end was performed again in the testbeam setup, with a threshold of 1.4 ke$^-$ and a charge to ToT conversion of 5 ToT = 8 ke$^-$ for the sample irradiated to 2×10^{15} 1 MeV n_{eq} cm^{-2} . The higher fluence sample was tuned with a threshold of 1.2 ke$^-$ and a conversion of 6 ToT = 6 ke$^-$, but was operated at a threshold of 1.9 ke$^-$ during data taking due to the large number of noise hits.

Bias Scan

Using the reconstructed tracks from the telescope, it is possible to accurately map the CCE across the unit pixel cell in order to better account for the loss of charge close to the pixel boundary. Tracks were reconstructed through the telescope and, after alignment of the devices under test (DUTs), the intercept position of the track with the device was calculated. Splitting each pixel into smaller regions and averaging over all pixels, the charge spectra for each sub-pixel region could be reconstructed based on the intercept position and the charge observed for the resulting cluster. This was

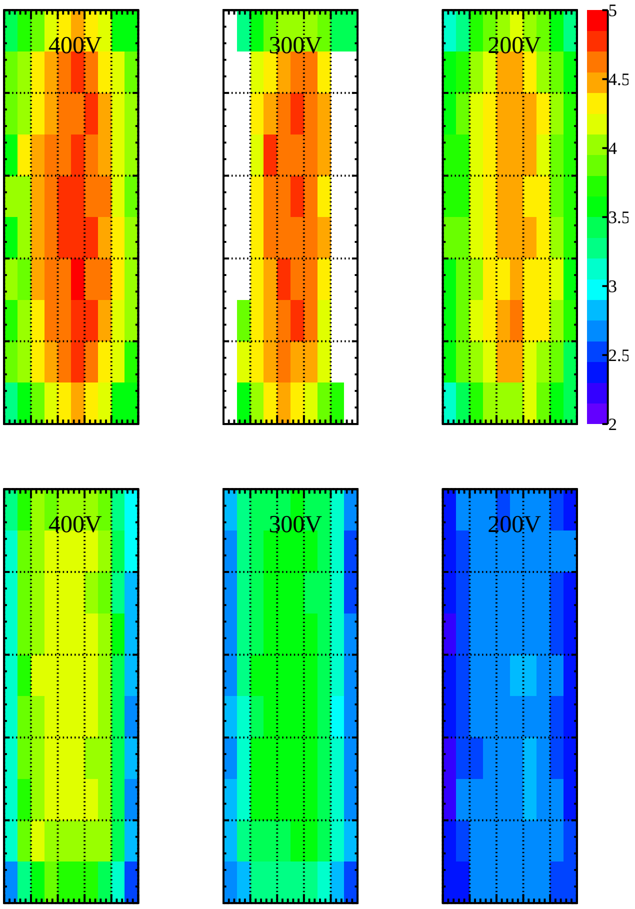

Figure 5.9: Charge collection maps as a function of bias for the 2×10^{15} (top) and 4×10^{15} (bottom) 1 MeV n_{eq} cm^{-2} samples at perpendicular incidence, for single pixel clusters. Each plot is over the unit cell and the colour axis displays the Landau MPV in ToT counts.

performed for perpendicular incidence at a range of bias voltages, to supplement the lab measurements taken previously.

The results for the two samples are presented in figure 5.9, showing the MPV of a convoluted gaussian-landau fit in each sub-pixel region. The plots have been made using only single pixel clusters, and show several features aside the increase in collected charge with higher bias voltage. As would be expected, the CCE across the unit pixel cell is observed to be very non-uniform, due to charge loss to neighbouring pixels which remain under threshold. However, this effect is significantly reduced in the more heavily irradiated sample, suggesting that the active depth of the sensor is less than that of the lower fluence sample. This is in keeping with the increased effective dopant concentration due to higher defect concentration.

As mentioned previously, with the prototype version of the FE-I4 exact charge calibration is not possible, and large errors exist in the ToT to charge conversion. Nonetheless, taking the approximate values of charge measured in both samples, and comparing with values observed in literature [34], it is reasonable to suggest collected

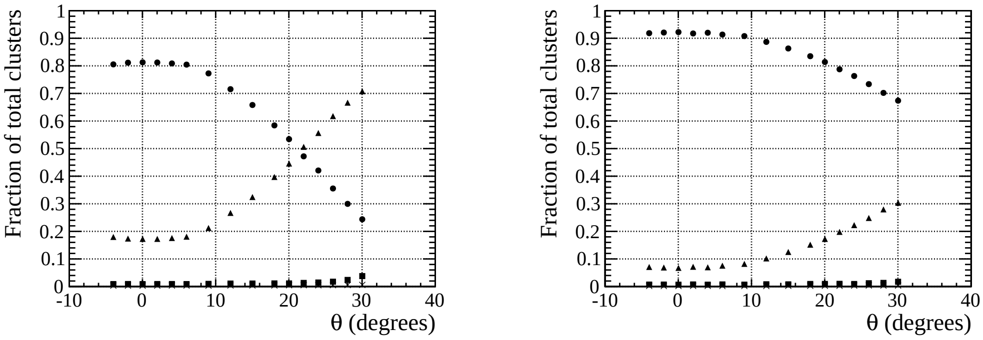

Figure 5.10: Fraction of n-pixel clusters as a function of angle for (left) 2×10^{15} and (right) 4×10^{15} 1 MeV n_{eq} cm^{-2} irradiated samples. The 1-pixel cluster fraction is represented by circles, 2-pixel by triangles, 3-pixel by squares and 4-pixel by crosses.

charge of around 7.5 ke^{-} for the sample irradiated to 2×10^{15} 1 MeV n_{eq} cm^{-2} , and around 5 ke^{-} for the higher fluence sample.

Charge Sharing

The observations from the charge collection maps suggest that the higher irradiated sample collects charge from a reduced volume compared with the lower fluence sample, in particular the extent of uniformity across the pixel cell and the reduced total charge. This should be reflected in the response of the device under rotation, where the behaviour may differ from that of unirradiated samples. It has been shown in previous sections that the cluster fraction in the direction of rotation becomes dominated by geometry at angles where the lateral path length is greater than the pixel pitch.

The fraction of n-pixel clusters in the direction of rotation is shown for both samples in figure 5.10. On the left hand plot the lower fluence sample is shown, where already the onset of multi-pixel clusters occurs at higher angle than would be expected were charge to be collected from the full bulk. Both the high threshold with respect to the total charge collected and the drop in charge collection efficiency versus depth will drive the detector response towards smaller clusters. This is much more prominent

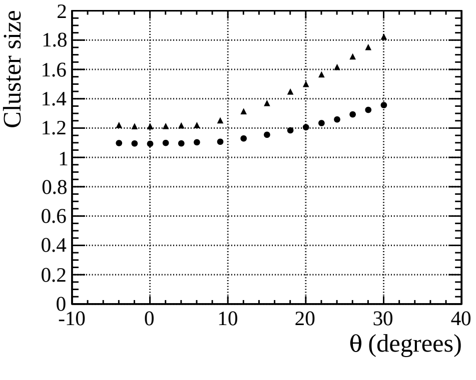

Figure 5.11: Average cluster size in the direction of rotation for 2×10^{15} (triangles) and 4×10^{15} (circles) 1 MeV n_{eq} cm^{-2} irradiated samples.

in the sample irradiated to 4 × 10^{15} 1 MeV n_{eq} cm^{-2} , where the response is flatter, suggesting only a very thin active region of the device. At the extreme shown (30°) the lateral path length through the sensor is roughly 85 µm, but almost 70 % of hits still result in charge deposits on a single pixel.

For completeness, the average cluster size versus incident angle is shown in figure 5.11 for both samples. The relative differences between the samples can be seen clearly, with only a factor of 2 in the fluence received by each. At 30° rotation the more heavily irradiated sample has an increase of just under 30 % in average cluster size with respect to perpendicular incidence, while for the the less irradiated sample an increase of 50 % is observed.

Charge Collection Efficiency

The angular distribution of the cluster size is of course not the only way in which the charge distribution within the sensor can be probed. Following the method for the bias scans performed with telescope data (section 5.4.2), the charge collection efficiency

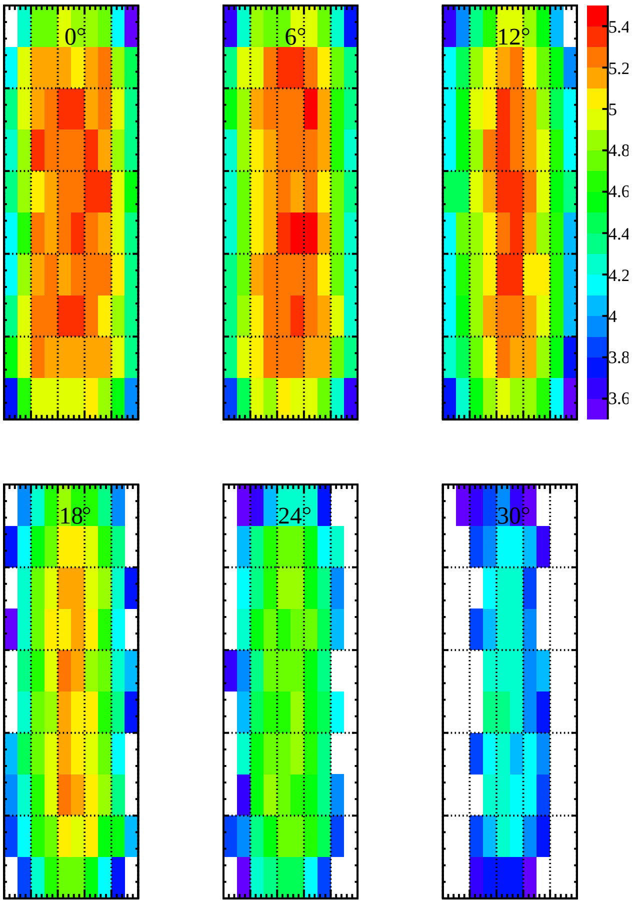

Figure 5.12: Charge collection maps for the 2 × 10^{15} 1 MeV n_{eq} cm^{-2} irradiated sample, for single pixel clusters. Each plot is over the unit cell and the colour axis displays the Landau MPV in ToT counts.

across the unit pixel cell can be mapped under rotation. As before, this is performed separately for single- and two-pixel clusters (where statistics permit).

The charge collection maps for the lower fluence sample can be seen in figures 5.12 and 5.13 for single- and two-pixel clusters respectively. The charge collected follows the same response as the cluster fraction shown above, and the effects of neighbouring pixels remaining under threshold can be seen by the drop in charge collection over the central region of the pixel. The drop is perhaps not so marked as it would be if charge were being collected from the full volume of the sensor, without which the lateral path is effectively shorter by the same fraction. Similarly for the charge collection map for 2-pixel clusters, at 30° it might be expected that charge is lost to neighbouring pixels on either side, when in fact an increase in charge collected is observed. This is likely due to the shorter active lateral path, which prevents charge from being lost until significantly higher angles, combined with the increased path length close to the implant-side which would give an increase in charge for the same depth.

The results for the 4×10^{15} 1 MeV $\mathrm{n_{eq}}$ $\mathrm{cm^{-2}}$ sample are shown in figure 5.14,

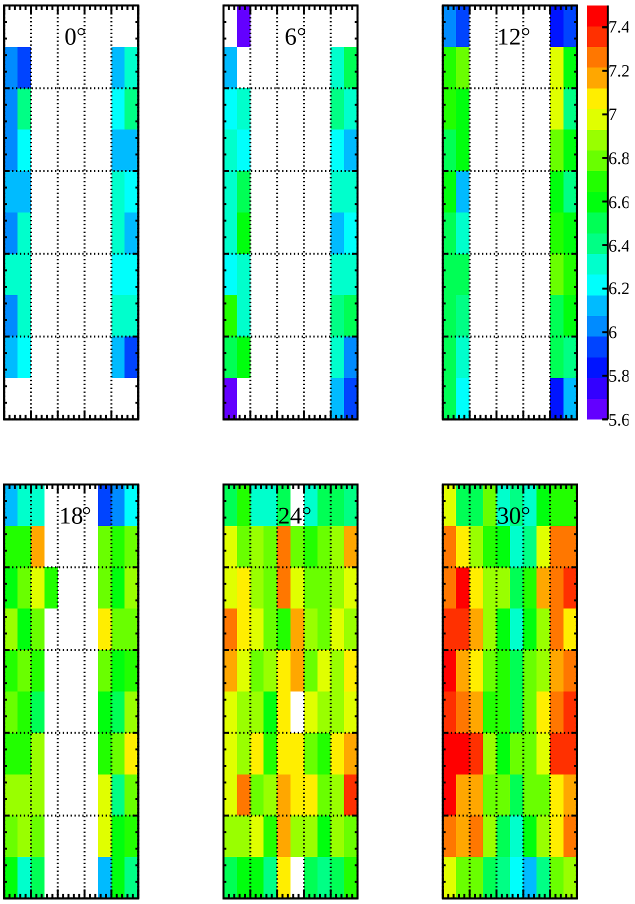

Figure 5.13: Charge collection maps for the 2×10^{15} 1 MeV $\mathrm{n_{eq}}$ $\mathrm{cm^{-2}}$ irradiated sample, for 2 pixel clusters. Each plot is over the unit cell and the colour axis displays the Landau MPV in ToT counts.

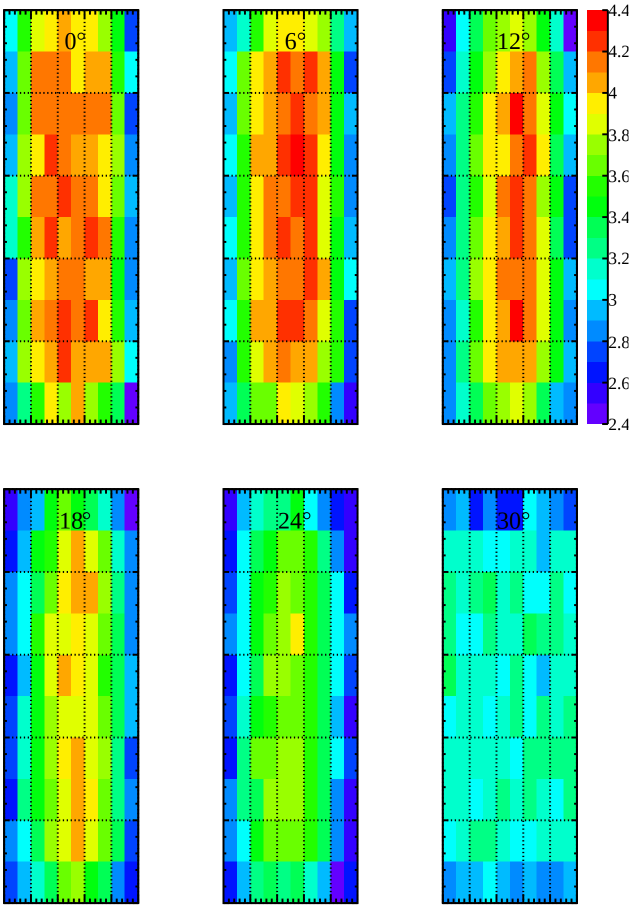

Figure 5.14: Charge collection map for the 4×10^{15} 1 MeV n_{eq} cm^{-2} irradiated sample, for single pixel clusters. Each plot is over the unit cell and the colour axis displays the Landau MPV in ToT counts.

showing charge collection maps for different angles of only single pixel clusters. The response becomes very uniform at larger incident angles, showing losses due to the neighbouring pixel threshold. The large uniformity at perpendicular incidence again indicates the limited extent of charge diffusion, while the active lateral path length at high angle does not appear to spread charge over more than two pixels.

Single Hit Resolution

The resolution of both devices versus angle is shown in figure 5.15. The resolution is again taken as the sigma of a Gaussian fit to the residuals, subtracting the telescope pointing precision in quadrature, and is shown alongside the resolutions for single- and two-pixel clusters individually. For the 50 µm pixel size the expected binary resolution is 14.4 µm, and it is noticeable that the performance of both samples is better than this simplistic expectation. For the sample irradiated to 2×10^{15} 1 MeV n_{eq} cm^{-2} the response improves significantly with angle, reaching an optimal resolution of 8.5 µm at 24°. This is driven by the change in cluster fraction, and the improved resolution offered by two-pixel clusters. As with unirradiated detectors, the reduced area giving rise to single pixel clusters means the average distance between the track intercept and the pixel centre is smaller, giving improving resolution versus angle for these clusters. Conversely, two pixel clusters have improved resolution for shallow angles, before charge loss to neighbouring pixels and uncertainties in the charge deposition (from landau fluctuations and poor charge resolution) degrade their performance.

The results of the 4 × 10^{15} 1 MeV n_{eq} cm^{-2} sensor can be seen on the right hand side of figure 5.15. As with all previous measurements, the resolution remains virtually unchanged over the full angular range.

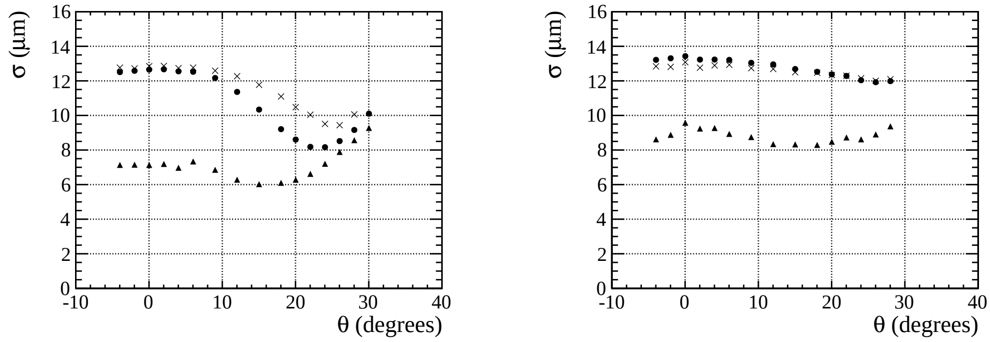

Figure 5.15: Resolution versus angle for the 2 (left) and 4 (right) × 10^{15} 1 MeV n_{eq} cm^{-2} irradiated samples, showing resolution for single- (crosses), two- (triangles) and any- (circles) pixel clusters.

5.4.3 Implications for the VELO Upgrade

The results shown for both the charge collection and tracking efficiency of the irradiated sensors are encouraging for the prospects of the upgraded VELO detector. As shown in figure 5.1, the damage across a single module falls sharply with radius which, while giving a higher fluence at the very tip of the module, results in a relatively moderate dose across most of the detector. High voltage operation is vital in order to extract sufficient charge from the sensor, but despite the non-uniformity in the collection of charge across the pixel matrix, sufficient signal is observed even in the pixel corners (see figure 5.9), suggesting that the detector efficiency should remain high throughout its lifetime. The reduction in cluster size, and the resulting degradation of the single hit resolution (figures 5.10 and 5.15) will affect the detector performance, although the extent of those changes require simulation of the complete environment (discussed in section 6.4).

5.5 Summary

Integration results for the FE-I4 ASIC with the Timepix telescope have been presented, showing some representative performance plots for an unirradiated planar silicon sensor analysed within the analysis software framework. The single hit efficiency has been measured versus the track incident angle, with the device fully efficient as expected. The single hit resolution in the long (250 µm) pixel direction has been shown under rotation, with a minimum resolution of 47 µm at 20° incidence.

These tests have been used to optimise data taking conditions in order to allow the analysis of FE-I4 assemblies irradiated to fluence levels comparable to the upgraded VELO detector (2 and 4 $\times$ 10^{15} 1 MeV n_{eq} cm^{-2}). The analysis of these devices has been presented, with particular importance on the charge collection across the unit pixel cell and the single hit resolution achieved. Performance close to binary has been observed for the more heavily irradiated sample, with an apparent reduction in the depth from which charge is collected. All of the results indicate that the upgraded VELO detector should be able to operate efficiently until its end of life, though with the loss of single hit resolution in the inner regions.

6

Simulated Performance of the Upgraded VELO

Long before construction begins on an experiment, detailed simulations are carried out in order to verify the detector design and ensure that the predicted performance meets that required by the experiment. Monte Carlo (MC) simulations are used extensively in particle physics for this purpose, typically in parallel with lab testing of the individual detector components. The complexity of the MC description usually evolves ahead or in step with the hardware design, and by the time of construction will contain a full description of the hardware involved, complete with electronics effects and simulation of the readout chain. For the upgrade of LHCb, these simulations have been carried out individually for each of the sub-detectors proposing significant changes, and have been combined to allow different configurations of proposed systems to operate together. The VELO is a largely self-consistent sub-detector, with nothing lying between it and the collision point and interacting with the rest of the detector solely through the use of the measured momentum in correcting the track fit. Given the precision of the momentum determination expected (see figure 1.15), the true particle momenta from the generated event can be used without unduly influencing the results.

6.1 Overview of the Simulation Environment

6.1.1 Simulation Chain

The simulation chain in LHCb comprises three distinct steps: event generation and propagation; digitisation and readout; event reconstruction. These are carried out by the software packages *Gauss*, *Boole* and *Brunel* respectively and are built on Gaudi, combining algorithms written in C++ with python configurations and operating on an event-by-event basis.

The two components of Gauss are very much decoupled, and utilise shared software packages common to many other particle physics experiments. First, the proton-proton interaction is simulated using *Pythia*, which determines for each collision the relevant decay products. Pythia contains both theory and models for a wide range of high energy physics phenomena, describing the possible states produced and their decay products and properties (angular distribution, daughter momenta, etc.). This provides the particle physics event which will be observed in the detector. The second stage is to propagate all of the particles produced, and simulate their interaction with the matter through which they will pass. This will be the sum of beam pipes, mechanical structures, detecting elements, etc. that exist in the simulated detector, and describes only the direct result of the particle interaction with matter. For a silicon detector, this means that the creation of electron-hole pairs is described, but not their motion due to the presence of an electric field. Effects due to the interactions on the propagated particles are also considered, such as multiple scattering or the conversion of photons into electron-positron pairs. This step is run using the *Geant4* software package.

Once the particles produced have been propagated outside of the detector volume, the result is a series of interaction products in the various systems. For the VELO, this is the charge deposited at each stepping point throughout the sensitive region of the silicon. *Boole* is used to describe the detection of this charge, along with the processing steps carried out by the detector electronics. For the upgraded VELO this comprises the VELOPix ASIC and the TELL40 DAQ board. The electric field and sensor parameters (resistivity, electric field) are used to determine the diffusion of discrete charge deposits, and these are simply summed over each pixel and digitised using an approximation for the expected VELOPix response.

The output of *Boole* mimics that of the output from the off-detector electronics which is passed to the computing farm for reconstruction, written in so-called *Raw Banks.* The final stage of the simulation chain (*Brunel*) takes these data banks (which may be from real data in the case of the existing detector) and reconstructs the collision and decay products. Track reconstruction algorithms are performed for several of the tracking detectors in isolation, and also using track seeds extrapolated between them. The primary and secondary vertices are reconstructed, particle IDs are assigned, and the variables required for performing an analysis on the particle physics event are calculated. These may all be output to a separate file type, such that the reconstruction is not required to be rerun for each separate analysis that considers the same event.

6.1.2 VeloPix Detector Description

Several aspects of the upgraded VELO detector require descriptions in the simulation environment in order to represent the different detector properties. The material physically present is an important part of this, as the detector performance will involve components due to multiple scattering. The accurate mapping of active silicon volumes is also necessary for the generation of charge by passing particles, in addition to the subsequent reconstruction of the event. Besides the physical layout, the sensor properties and electronics description will play a role in both of these areas, and all of these must be accessible throughout the full MC simulation chain.

In some cases below, simulations were carried out before certain technology decisions were taken (the most recent are those found in the Technical Design Report (TDR) [25]). Where this occurs the difference from the TDR and rationale are given, along with any consequences for the design.

Material Description

An initial description of the material content exists for the VELO upgrade, where the detector is abbreviated as *VP* (VELO Pixel). As the module design is still under development this is likely to change significantly over the coming years, but for present studies it is sufficiently detailed to gauge the expected performance. The physical layout is written in XML (Extensible Markup Language) and is shown in figure 6.1. Individual stations are modelled with a 400 µm thick silicon substrate (the microchannel plate), with a hybrid structure on top to represent the readout circuitry. This has been taken as a uniform layer of 450 µm thick kapton, with 30 µm of copper. In order to reduce the amount of material present before the first measured point, the cooling substrate has been retracted by 5 mm from the inner edge. There are 4 silicon sensor tiles of 200 µm thickness, mounted on top of 200 µm thick ASICs on opposing sides of the module. To prevent gaps in the coverage, these overlap by 2 pixel widths (110 µm). The module control chip (marked GBT in the figure) is represented by a 1 mm thick silicon block whose dimensions matches those of the VELOPix, while the final mechanical items are the cooling connector (Aluminium, 2 mm thick on one side, 1 mm thick on the other) and the support bars (carbon-fibre, 6 mm outer diameter and 1 mm wall thickness).

The RF foil is visible in both schematics and is described by a 250 µm thick corrugated aluminium structure, closely approximating the current design. Both halves of the VP detector are contained within vacuum tanks, with the RF foil forming the wall which faces the other half (and the beam). When closed, the closest pixel on the modules is 5.1 mm from the beam position. The RF foil sits 3.5 mm from the beam.

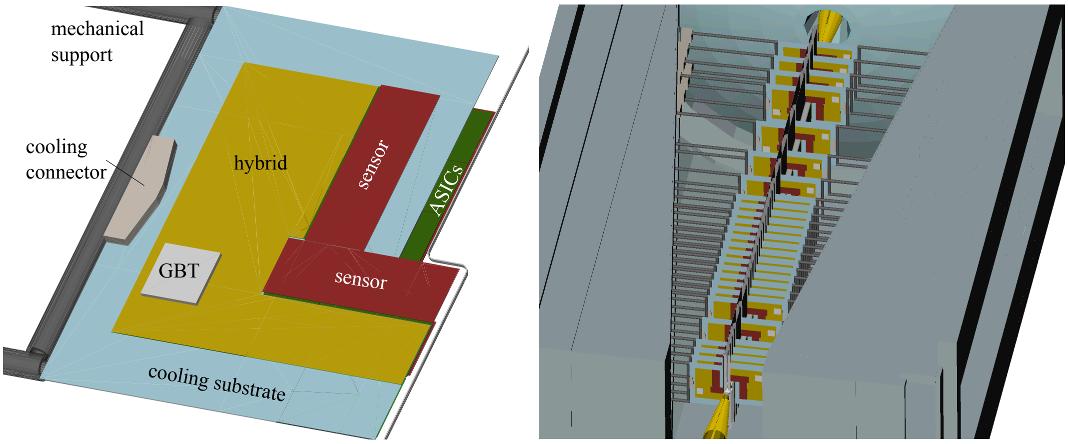

Figure 6.1: Left: Material description of a single VP module. Right: The full VP detector geometry (top surface removed for clarity). Figures taken from [25].

ASIC and Sensor Description

The description of the active components of the VP detector are at present an approximation based on the current ASIC architecture, which will require updating as the ASIC design evolves. The sensor, as stated above, is 200 µm thick and the discrete charge deposits generated by Geant4 are diffused according to a Gaussian distribution with

$$\sigma = \sqrt{\frac{2k_BTd}{q}}\sqrt{z} \tag{6.1}$$

where k_B is Boltzmann's constant, T is the temperature, d is the sensor thickness, q is the charge of the electron and z is the distance between the charge deposit and the collecting implant on the sensor. This approximation is used at present when it is expected that the bias voltage will be far greater than the depletion voltage, allowing a constant electric field to be assumed. This may be subject to change, but will be driven by the timing characteristics of the ASIC (where the charge collection time may affect the timewalk) and the rate at which the voltage must change to counter the effects of radiation damage.

The ASIC has been implemented using a ToT charge discrimination, though with the number of bits set to 1 after the decision to proceed with a binary chip. The time spent over threshold (t) is calculated using

$$t = \gamma(q - q_{th}) \tag{6.2}$$

where q and q_{th} are the charge and threshold charge respectively and γ is the discharge rate. For historic reasons, the default discharge rate is set such that 15,000 electrons would register a time over threshold of 250 ns. The clock frequency of the chip is assumed to be 40 MHz, giving a 25 ns sampling period.

6.1.3 Reconstruction

Track reconstruction in the VELO is biased towards tracks produced in the downstream direction, ie. tracks where the particle lies within the LHCb acceptance. Tracks outside of this range are still of use in constraining the position of the primary vertex, but are not essential and play no further part in the tracking. The first step in the tracking is to decode the raw data banks and cluster the pixel hits observed for the event. This is performed using a seeded eight-way flood fill algorithm, taking advantage of common imaging tools.

Once clusters are built, their global positions and uncertainties (pitch / $\sqrt{12}$) are fed into the pattern recognition. Starting at the downstream modules (those closest to the rest of the LHCb detector), clusters on neighbouring same-side modules are combined if they fulfil the conditions

$$\left|\frac{\delta x}{\delta z}\right| < 0.4, \quad \left|\frac{\delta y}{\delta z}\right| < 0.4 \tag{6.3}$$

This “seed” track is then extrapolated upstream, and the closest cluster within a given window on the module is added to the track. If no clusters are found, the opposite-side module is checked, followed by the next same-side module. If no clusters are found after 3 stations (comprising the same- and opposite-side module) then the search is abandoned.

For tracks with 3 clusters, a cut on the track χ^2 is applied, and all clusters are required to be unused by another track. For tracks with more than 3 clusters, at least 50 % of the clusters are required to be unused, and no χ^2 cut is applied. In all cases, the tracks are fitted with a fixed-momentum Kalman filter and saved for use in further tracking algorithms.

6.2 Detector Performance

For the VELO upgrade TDR, samples have been prepared in order to benchmark the expected detector performance and guide the technology decisions taken to date. The following simulations were carried out using a mean number of visible interactions per bunch crossing $\mu = 5$, which would be observed for operation at a luminosity of 2×10^{33} cm^{-2} s^{-1}. The samples were generated using the decay $B^0 \to K^{*0}\mu^+\mu^-$, in order to provide a harder momentum spectrum and reconstruct a decay vertex featuring many tracks. The beam energy used is 7 TeV, for a collision energy of 14 TeV and with the nominal 25 ns bunch spacing expected.

6.2.1 General Properties

The response of the simulated detector for different track properties (incident angle, momentum, etc.) does not provide a full picture of the expected performance without also showing the distribution of these properties observed in LHC collisions. Some of the general track distributions are shown here such that a more complete picture can be obtained.

The angular distributions of reconstructed tracks from the generated $B^0 \to K^{*0}\mu^+\mu^-$ events are shown in figure 6.2, where the angles displayed are the pseudorapidity η (defined as $-\ln\left(\tan\left(\frac{\theta}{2}\right)\right)$), the polar angle ($\phi$) and the incident angle relative to the beam axis (θ). Given the orientation of the silicon sensors with respect to the beam axis, this is also the track angle relative to the sensor normal direction. The track distribution in terms of the components of θ are shown in figure 6.3, where θ_x is the

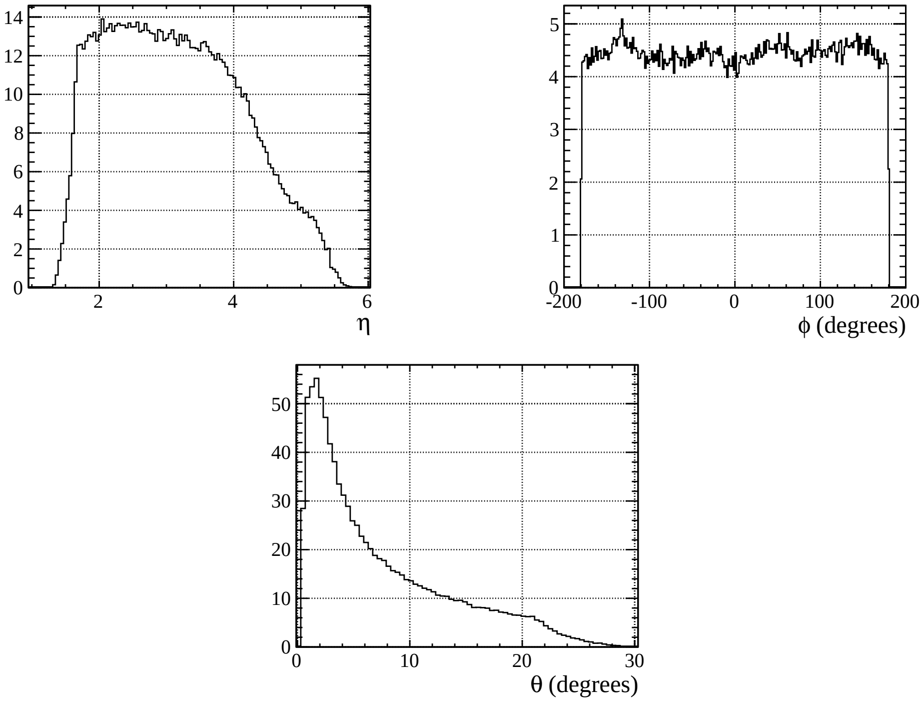

Figure 6.2: Distribution of all reconstructed tracks in (top left) η, (top right) ϕ and θ.

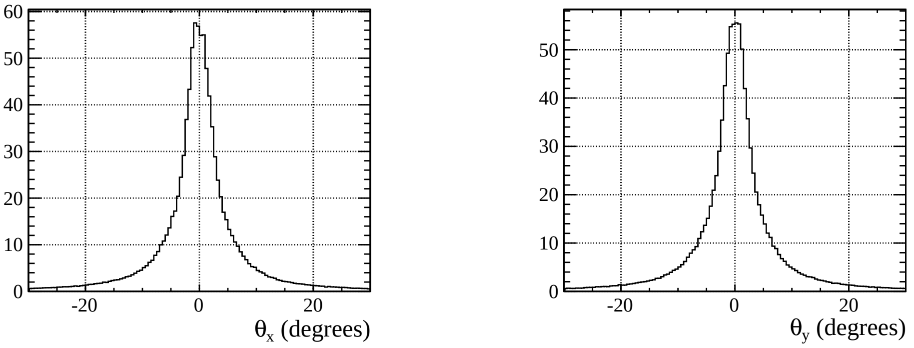

Figure 6.3: (Left) x- and (right) y-component of the track angle with respect to the silicon sensor surface.

component in the silicon sensor x-direction. There are several noteworthy features in these distributions. First, the cut on track angle applied during the reconstruction can be clearly seen in the distribution of θ, where the cut of 0.4 radians corresponds to approximately 22°. Furthermore, as expected by the orientation of the VP with respect to the beam axis, it can be seen from the distributions of $\theta_{x,y}$ that most incident hits occur at relatively small angles with respect to the normal of the sensor surface. Without a very high resistivity substrate and operation close to bias voltage, it can be expected that the resolution will be dominated by single pixel clusters, and therefore the pixel pitch. Taking into consideration the high bias voltage required after irradiation, the pitch should be the driving force behind the single hit precision.

The track momentum (p) and transverse momentum (p_{T}) have a strong effect on the detector resolution, and the scattering term shown below for the IP resolution will be seen to be large compared with the effects of the single hit resolution. The total and transverse momentum distributions for all reconstructed tracks are shown in figure 6.4. The observed structure in the transverse momentum distribution around 1 GeV/c is likely due to the high transverse momentum inherent in the B^0-decay daughters, above the underlying event from the proton-proton collisions.

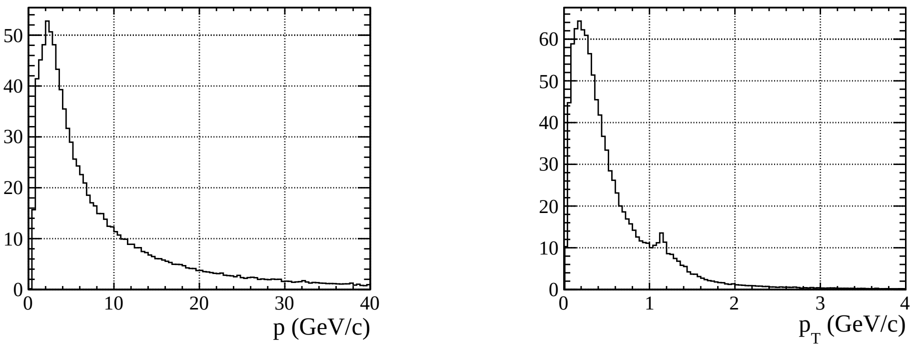

Figure 6.4: (Left) Total and (right) transverse momentum distributions for all reconstructed tracks.

6.2.2 Single Hit Resolution

From the plots above, the predominant incident angle for tracks within the LHCb acceptance is normal to the sensor surface. As for the samples measured in testbeam, the primary factor controlling the resolution is therefore the extent of charge sharing, which is deliberately simplified in the detector simulation at present. The high bias voltages which may be applied for timing considerations (and will definitely be required after radiation damage) will likely render moot any fine-tuning of the sensor parameters which might be carried out at the present time, when even the sensor resistivity is not yet known.

The extent of charge sharing can be seen in the two plots of figure 6.5, where the cluster size and fractions in the direction of rotation are respectively shown versus the incident track angle. In order to provide sufficient statistics but prevent the influence of inclined tracks in the perpendicular direction, the angular component θ_y is constrained to +/- 2°. The response is similar to that of the Timepix devices measured in testbeam, and the geometric shape of the cluster fraction plot behaves as expected of a 200 μm thick sensor.

The single hit resolution is shown in figure 6.6, displaying both the absolute measurement error versus overall track angle and the component of both in the x-direction (aligned with the pixel cell). The resolution was calculated from the RMS of the cluster residual distribution, taking the difference between the reconstructed cluster centre and the MC truth information of the hit position (in the centre of the sensor). As many of the tracks reconstructed in the VP detector contain only 3 or 4 hits, the biased or unbiased residual between the track intercept and cluster centre was not used, in order to decouple the fitted track error from the single hit resolution. In the component of the resolution aligned with the pixel cell (right hand side plot), the minimum in resolution is observed to begin earlier than for the Timepix data measured previously. This is due to the binary readout employed in the simulation of the VELOPix, which results in a minimum residual when the area of the pixel giving rise to 2-pixel clusters is equal to that of single-pixel clusters. Thus the minimum point appears at roughly half of the expected angle.

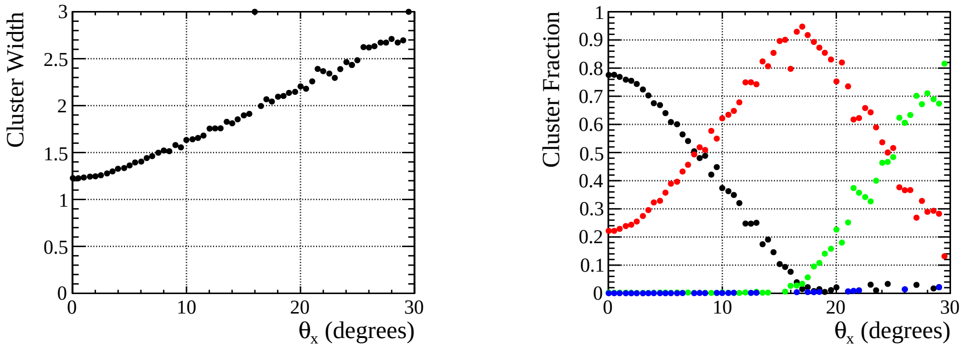

Figure 6.5: Left: Cluster width in the direction of rotation θ_x. Right: Fraction of n-pixel wide clusters versus track incident angle, for (black) 1-, (red) 2-, (green) 3- and (blue) 4-pixel wide clusters. For both plots $|\theta_y| < 2°$.

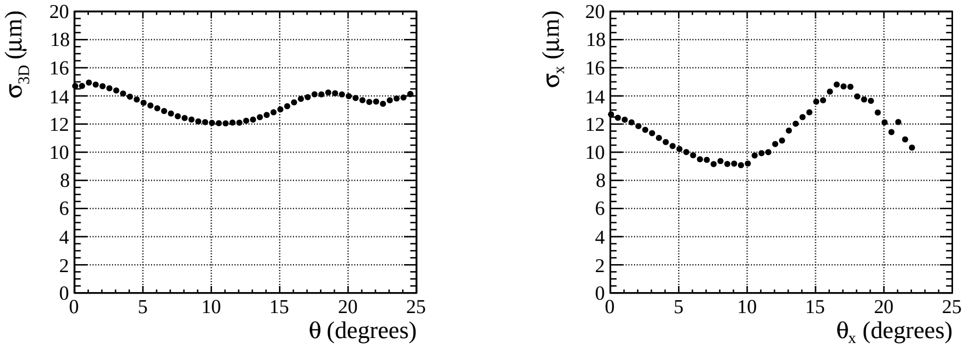

Figure 6.6: Left: 2D resolution versus absolute track angle with respect to the sensor surface normal. Right: Resolution in the x-direction versus angular component in this direction. As above, $|\theta_y| < 2°$.

6.2.3 Impact Parameter Resolution

The IP resolution of both the present and upgraded VELO is one of the most important figures of merit for the detector performance. The present detector has been studied extensively [60], and an approximately linear relationship between the IP resolution and the inverse track momentum transverse to the beam direction has been established (where deviation at high momentum occurs due to the relative contributions of scattering and single hit precision). The relationships for the 3D IP resolution and the component in the x-direction of the global LHCb reference frame are shown in figure 6.7. The results are shown for tracks reconstructed within the upgraded VELO (without considering other sub-detectors downstream) and fitted with a Kalman filter which assumes a fixed particle momentum. This step in the reconstruction is important due to the effects of extrapolating the track to the downstream tracking stations, and more importantly for the online trigger where a cut on IP is performed before a measurement of the track momentum is made. In both plots, the performance significantly exceeds that of the existing VELO at low momentum, while for high momentum the resolution is approximately the same (a comparison of the two detectors is given in figure 1.15). This is due to the coarser single hit resolution of the VP detector, which

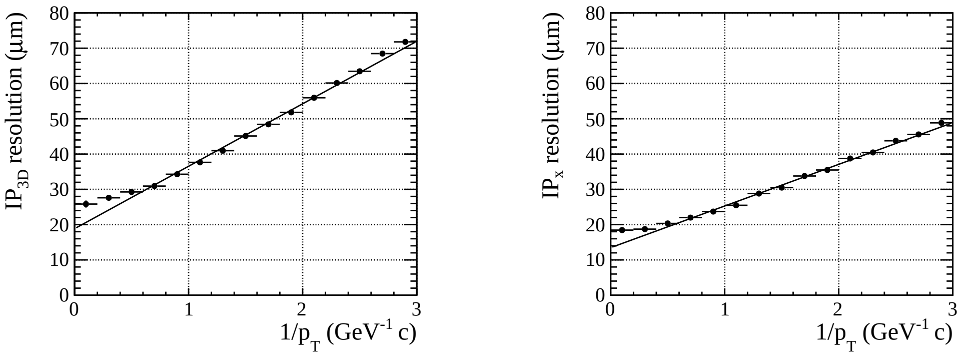

Figure 6.7: (Left) 3D and (Right) x-component of the Impact Parameter resolution versus inverse track momentum transverse to the beam direction. Both plots are for VELO tracks, using a track fit with fixed-momentum Kalman filter.

controls the IP resolution at infinite momentum, but is compensated for by the location of the VP closer to the interaction point than the current detector.

6.2.4 Reconstruction Efficiency

The second important figure of merit is the absolute track reconstruction efficiency for a range of track types observed in the LHCb acceptance. In all cases the efficiency is defined as the number of reconstructed tracks with respect to the number of *reconstructable*: those within the LHCb acceptance and which leave a minimum number of hits in the relevant detectors. This is defined along with the definition of the track type (shown in figure 1.7), as:

- VELO tracks: Hits from this particle exist on a minimum of 3 different VP modules
- Long tracks: Hits from this particle exist on a minimum of 3 different VP modules and have at least one “x” and one “stereo” hit in each of the three downstream track seeding stations

The reconstruction efficiencies for both types of tracks, in addition to those with a minimum momentum of 5 GeV/c and those produced by the decay of b-hadrons, are shown in table 6.1 along with the equivalent performance for the current VELO. The efficiency is observed to be extremely high for all track types, and outperforms the existing VELO not only for the upgrade running conditions, but also for those at which the VELO is currently running. The efficiency is extremely flat over the full LHCb acceptance, described in more detail in [25].

Track Type	VP Reconstruction Efficiency (%)	Current VELO Reconstruction Efficiency (%)
VELO, $p > 5$ GeV/c	98.9	92.7
Long	99.4	93.7
Long, $p > 5$ GeV/c	99.6	95.7
b-hadron daughters	99.6	95.4
b-hadron daughters, $p > 5$ GeV/c	99.8	96.6
Ghost rate	2.5	25.0

Table 6.1: Track reconstruction efficiency for the existing VELO and for the VP detector, with $\mu = 5$ corresponding to $\mathcal{L} = 2 \times 10^{33}$ cm^{-2} s^{-1}. Note that this instantaneous luminosity is far beyond the design requirements for the existing VELO.

6.3 Simulation Studies

The following studies were performed during the preparations for both the internal technical review meeting, where the decision to proceed with an upgraded VELO based on a pixel detector with microchannel cooling was taken, and the subsequent technical design report. At those times some of the detector parameters had not yet been fixed, or would later be changed for various reasons. In each case below, the relevant changes are highlighted, and the result of the study on the detector design discussed.

6.3.1 Track Fit

The present VELO detector is constructed using back-to-back silicon strip sensors, with strips of variable pitch orientated in either the Φ or radial direction. When tracks are reconstructed within the detector, they are fitted before being stored for use in seeding tracks to the downstream tracking stations. This fit is a simple linear least-squares fit, using the strip pitch at each of the measurement points to give an error of pitch/$\sqrt{12}$. Once *Long* tracks have been created, the full track is refitted using the measured momentum with a Kalman filter which accounts for the material content and subsequent scattering within the detector.

The same approach was used in the initial simulation of the VP detector, and the IP resolution was observed to be significantly worse for VELO tracks than for Long tracks (see figure 6.8), specifically for tracks with lower transverse momentum. This feature is not observed in the current detector, and was found to be due to the linear fit performed on VELO tracks. Given the orientation of the VELO with respect to the beam axis and tracks produced at the interaction point, successive hits on any given track generally occur with increasing radial distance from the beam axis. In the current VELO, where the strip pitch gradually increases with larger radius, this provides a heavier weighting for the first measurement points on a given track with respect to those further downstream in the detector. This is not the case for a pixel-based VELO.

The VP detector has a uniform pixel pitch, giving a constant hit error for all measurement points. The result of this on the track fit for VELO tracks is that downstream measurements, after having passed through material and suffered from scattering, are

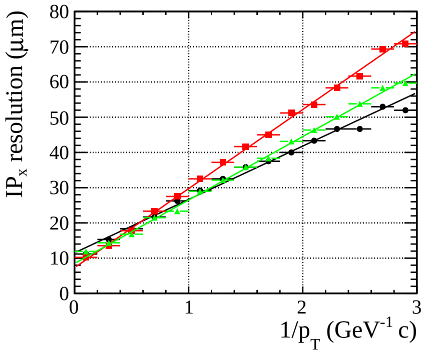

Figure 6.8: IP resolution in the global x-direction for (black) Long tracks, fitted with a Kalman filter after measurement of their momentum, (red) VELO tracks with a linear fit and uniform hit error and (green) VELO tracks with a linear fit, reweighted with hit errors proportional to the hit radius.

given a stronger weighting than they should be assigned, pulling the fitted track away from the true position. Figure 6.8 shows the IP resolution for (black line) Long tracks, where the momentum has been measured and the track refitted with a Kalman filter, and (red line) VELO tracks with a uniform hit weight.

A solution for this problem is to construct artificial weights for the hits measured in the VP and to use these in the linear track fit. Such an approach is shown (green line) in figure 6.8. The hit error was assigned to be

$$\sigma = 2 \times 55\,\mu\mathrm{m} \times \frac{\mathrm{r}}{45\ \mathrm{mm}} \tag{6.4}$$

where r is the hit radius, the 55 µm comes from the pixel pitch and the 45 mm allows the weight at the outer radius to equal twice the pixel pitch. At small radii the hit error is approximately pitch/$\sqrt{12}$.

While not replicating the full Kalman fit performed on Long tracks, this reweighing improves the IP resolution for low transverse momentum tracks. This can be taken advantage of in the online trigger, as detailed below, in discriminating low-momentum tracks. For the VP detector, this approach has been superseded by a fixed-momentum Kalman filter, which further improves the performance with respect to the final tracks produced.

6.3.2 Charge Discrimination

The initial design for VELOPix was that each pixel would record the deposited charge using the same Time over Threshold mode as the existing Timepix ASIC. Given the large hit rates and subsequent data volume that must be transmitted from the innermost ASICs in the detector, the ToT counter was envisaged to be only 4 bits long. This was intended to both aid in the monitoring of radiation damage within the detector and to improve the single hit resolution with respect to binary readout. After more detailed design work had been carried out on the VELOPix, the availability of space on the pixel came into question. Simulations were carried out to identify any gains for continuing with a binary ASIC design.

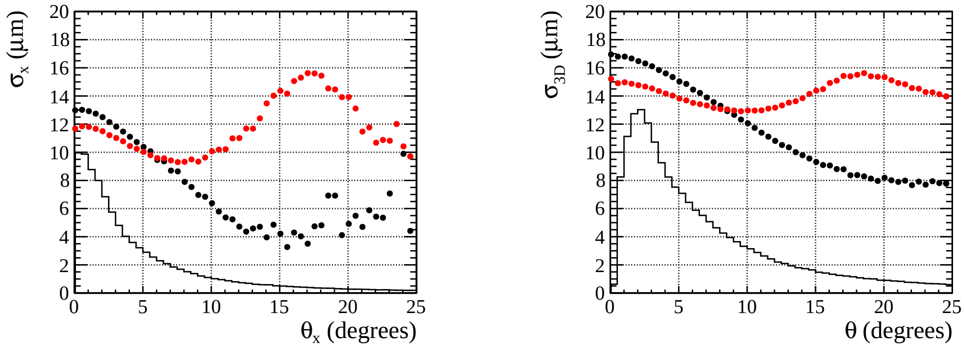

Figure 6.9: Left: Singe hit resolution in the pixel x-direction, versus incident angle. Right: 3D hit resolution versus track angle. Results are shown for (black) a 4-bit ToT charge measurement and (red) a binary ASIC. The angular distribution of tracks in each case is also shown.

Two plots are shown in figure 6.9: the single hit resolution in the pixel x-direction and the 3D hit error versus incident track angle. The simulations are run for ASICs with a 4-bit ToT (with a discharge current such that 10 ToT = 15,000 e^-) and with binary readout, both with a threshold of 1000 electrons. There are stark differences between the response in either case. The residuals where the pixel charge is measured behave in a similar manner to the current Timepix ASIC, namely a gradual improvement in resolution from perpendicular incidence to $\tan^{-1}$(pitch/thickness). The resolution is controlled in this region by the extent of charge sharing, where figure 6.5 shows a single pixel fraction of 80 % for perpendicular tracks. For a binary readout ASIC, the resolution behaves differently. The reconstructed position of n-pixel clusters is always the same, regardless of the intercept position of the track. So all 1-pixel clusters have a reconstructed position in the pixel centre, all 2-pixel clusters at the pixel edge, etc. The residual distribution for each of these is then a top hat function, with width depending on the angle and charge sharing. Where most tracks deposit charge in two pixels, the single-pixel cluster resolution will be very precise, due to the self-selection of tracks that pass close to the pixel centre. The angle of minimum resolution will no longer be that at which all tracks cover two pixels, but where half of the clusters produced are 2-pixel. At this point the residual distributions for single and 2-pixel clusters are approximately the same. Beyond this point, the single pixel residuals will continue to narrow, and the 2-pixel residuals will increase, leading to a local maximum in the resolution as seen in figure 6.9.

The effect of the single hit resolution on the parameter of relevance to analysis, the IP resolution, is however much less. The angular distribution of tracks shown in figure 6.9 highlight that most tracks are incident at shallow angle, where the difference between binary and analogue (at least 4-bit digitised) readout is minimal. In addition, for lower momentum the IP resolution becomes more and more dominated by multiple scattering and less dependent on the individual measurement errors.

The IP resolution plots for both cases can be seen in figure 6.10. In the left hand plot, VELO tracks (reweighted with the radial error shown in equation 6.4) are shown for both (black) 4-bit ToT and (red) binary readout, and the difference in performance can be seen to be minimal. For comparison, the IP resolution using the generated Monte

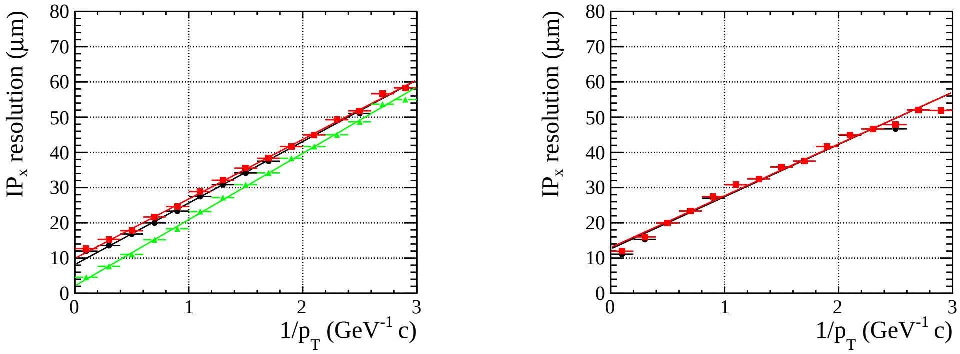

Figure 6.10: IP resolution versus inverse transverse momentum for (left) VELO tracks and (right) Long tracks. The results shown are for (black) 4-bit ToT and (red) binary readout and (green) the generated Monte Carlo hit positions. The Monte Carlo information is shown only for VELO tracks, due to the way this information is accessed in the software

Carlo hit location is also shown (green), giving the performance of an infinitely precise detector. At high momentum the difference observed between this and the other results is approximately the single hit resolution close to perpendicular as expected, while for low momentum tracks the response is almost identical. When the momentum of the tracks is measured and the full Kalman fit is performed (right hand plot), the difference between 4-bit ToT and binary readout is virtually zero. With no gain to the detector performance in terms of parameters relevant to analyses, the decision was taken to design the VELOPix as a binary ASIC.

6.3.3 Data Compression

The most formidable challenge for the VELO upgrade is the volume of data which the innermost ASICs are required to transmit. In order to have a fully software-based trigger, the detector must read out each 25 ns bunch crossing, which given the average number of tracks per event will lead to data rates of tens of Gigabit per second for the inner ASICs. In an attempt to reduce the total output bandwidth, the VELOPix chip is designed to share information between groups of pixels, so-called *Super Pixels* (SPs). This is achieved by designing the chip with a series of grouped columns which act as a single block. The analogue pixel circuitry may lie on the outside of this structure, with the location of the digital circuitry in the centre allowing the sharing of information between pixels. Buffering of the data can thus take place in this larger SP, with a reduction in transmitted data due to shared bunch crossing ID and location addressing.

Before the decision to adopt a binary readout was taken, it was envisaged that VELOPix would output a 4-bit measurement of the pixel charge. In this regime, a 4 × 4 region of pixels was planned for the SP structure. In addition, during these studies the expected sensor thickness was 150 μm, as opposed to the presently proposed 200 μm. Given the size of the pixel matrix (256 × 256) and the expected latency in data transmission to the off-detector electronics, the data packet required to read out a single pixel is 32 bits: a 12-bit bunch ID, a 16-bit pixel address and a 4-bit ToT. By combining the pixel region into super pixels, this can be reduced, using the data format shown in figure 6.11. In this scenario the bunch ID remains the same (12 bits), the SP

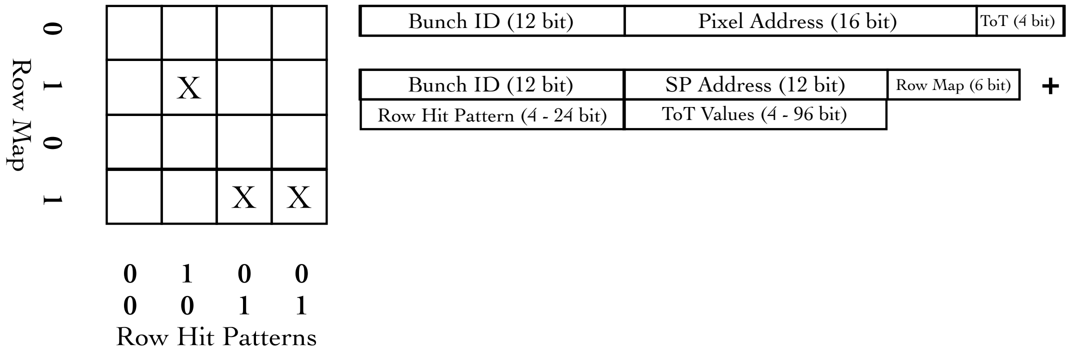

Figure 6.11: Output data package for (top) single pixel and (bottom) super pixel architectures. The pixel address encoding for super pixels is additionally shown.

address is transmitted (12 bits) and the individual hit information is packaged using a row map (6 bits, indicating which rows contain hit pixels) and a hit pattern for each row (4 bits for each row, giving the status of each pixel - hit or empty). The 6 bits in the row map allow further communication between super pixels, where the first two rows of the succeeding super pixel may be read out with the SP above.

Using this format, the information transmitted for single pixel hits increases from 32 to 38 bits, but for higher numbers of hit pixels reduces significantly. Where 2 hits occur within a super pixel a saving of 30 % is achieved, while for 3 hit pixels this increases to 44 %. The fraction of n-pixel clusters and average cluster size as a function of radius are shown in figure 6.12 for the assumed threshold of 1000 electrons. At low radius (the highest occupancy region) around 35 - 40 % of clusters contain more than a single pixel, giving potential benefits for data reduction with the SP architecture.

Each station in the VP detector contains two modules, mounted on either side of the beam axis. Within a module the natural unit size is the ladder (or sensor), which covers 3 ASICs. As the modules in a station are related to each other by a 180° rotation, it is sufficient to number the ASICs 0 - 11 and show results for these without reference to the detector half. The number scheme and arrangement within a module are shown in figure 6.13, where the ASIC number is given by (3 × ladder number) + ASIC position on ladder. The columns are aligned with the pixel y-direction, such that the periphery of the 3 ASICs on a ladder are aligned. The sensor ladders are also mounted on opposing sides of the module (ladders 0 and 3 on the front, 1 and 2 on the back) such that there is sufficient space for the readout traces.

As would be expected from the orientation of the module with respect to the beam, the highest hit rates are observed on ASIC 2 of ladder 0, and ASIC 0 of ladder 2. Since the hit rate falls off exponentially, it is easier to visualise the relative occupancy across the module in terms of the column number. The 4 × 4 pixel SP structure gives 64 readout columns per ASIC, and a unique column number on each module can be obtained by offsetting this with the ASIC number. The relative occupancy according to this can be seen in figure 6.14, where the single pixel occupancy within the 4 × 4 region is shown (black line), in addition to the readout packet occupancy (red line). A variation in hit pixel rate of roughly a factor of 7 is observed between the most and least active regions.

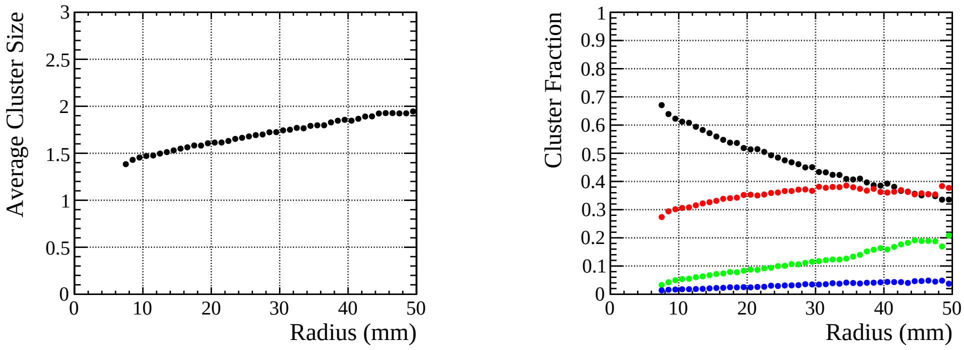

Figure 6.12: Left: Average cluster size as a function of radius. Right: Cluster fractions as a function of radius for (black) 1-, (red) 2-, (green) 3- and (blue) 4-pixel clusters. An operating threshold of 1000 electrons has been used.

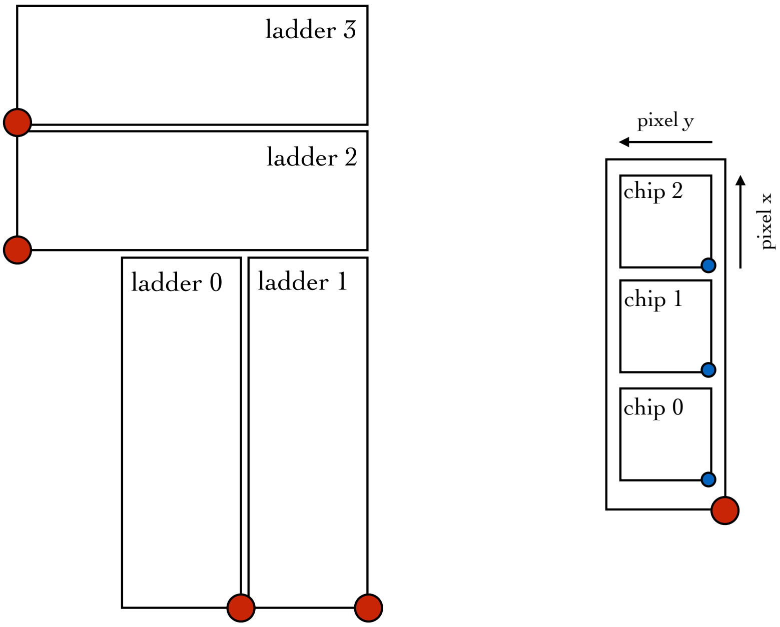

Figure 6.13: Reference frame for a VP module. The red and blue dots indicate the reference point of ladders and ASICs respectively. The ASIC columns (rows) are aligned in the pixel y- (x-) direction.

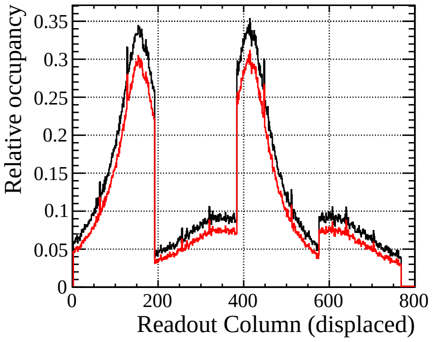

Figure 6.14: Relative occupancy for each (displaced) column on a module, showing (black) the single pixel and (red) data packet occupancy.

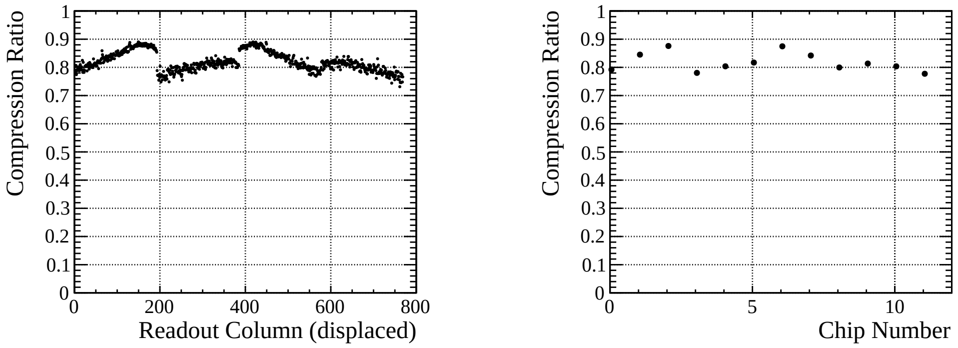

Figure 6.15: Relative data rates showing the gains of implementing a super pixel structure, as a function of (left) column and (right) ASIC number.

The compression ratio, defined as the data volume output using the super pixel format as opposed to the raw data rate expected, is shown as a function of column and ASIC number in figure 6.15. On both the column and chip level the gain varies between 10 % and 25 %. As expected, for the outermost ASICs the gain is more significant, as tracks which pass through these regions must be inclined to a greater extent and therefore produce clusters of greater size. However the occupancy in these regions is almost an order of magnitude lower than the inner ASICs, where the gain is most critical. Gains for these ASICs are just over 10 %.

The gain obtained from formatting the raw data on-chip is further impacted upon by the number of "split" clusters which occur: where a cluster contains more than one pixel but where the pixels lie within the boundary of more than one super pixel. As mentioned above, functionality to allow the super pixel to pick up hits in the two following two rows have been proposed, but this is not possible in the column direction. A significant number (between 10 % and 25 % - see figure 6.16) of clusters occur in such a configuration, with a variable gain or loss rate depending on the exact number of pixels hit in each super pixel. In this configuration the mean data rate for the hottest ASICs was expected to be 12.4 Gbit s^{-1}.

While the simulations show a modest gain in data compression for the implementation of a 4 × 4 super pixel structure, further gains might be expected given the sub-

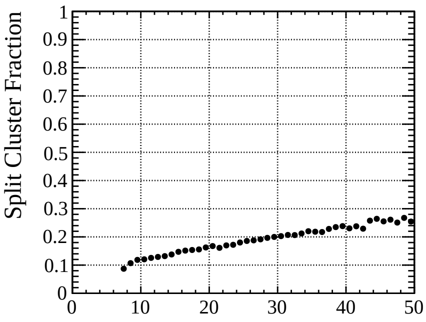

Figure 6.16: Fraction of split clusters, where pixels from a multi-pixel cluster are located within multiple super pixel boundaries.

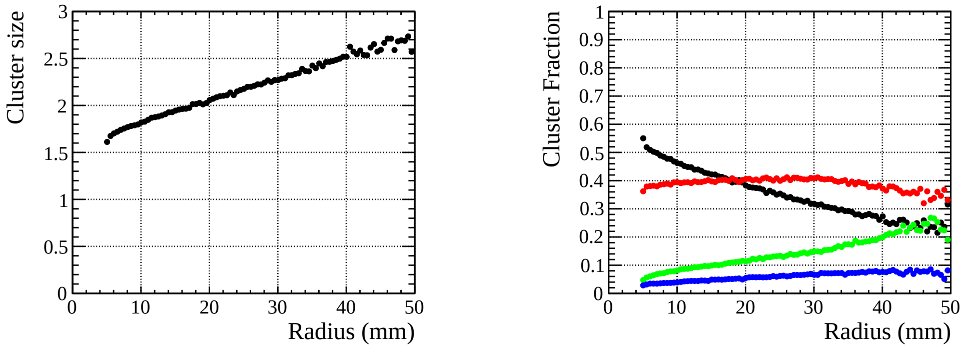

Figure 6.17: Left: Average cluster size as a function of radius. Right: Cluster fractions as a function of radius for (black) 1-, (red) 2-, (green) 3- and (blue) 4-pixel clusters. An operating threshold of 1000 electrons has been used.

sequent decision to construct the upgraded VELO with 200 µm thick silicon sensors. The increase in cluster size, detailed in figure 6.17, will increase the compression ratio given the drop of more than 10 % in the number of single pixel clusters. Similarly the decision to use a binary ASIC, both for the transmission of single pixel information and for SP packets, will change the total data transmitted.

Subsequent to this work, the design of the VELOPix super pixel was reconsidered, in particular taking into account the crossover of digital and analogue regions of the pixel. It was decided to continue with a SP structure of 2 × 4 pixels, such that the digital circuitry is touching between neighbouring columns. Each double column will be read out separately, and estimates for the expected mean data rate reach 10.2 Gbit s^{-1} for the innermost ASICs (with a peak rate of 15.1 Gbit s^{-1}).

6.4 Radiation Damage at the VELO Upgrade

As input to the technical design document of the VP, a study was carried out to investigate the robustness of the detector performance to the expected radiation damage during the total lifetime. The projected integrated luminosity is assumed at present to be 50 fb^{-1}, with an expected hardware upper limit on the applied voltage of 1 kV.

6.4.1 VELO Environment

The expected fluence to be received both by the current and upgraded VELO detectors has been estimated in detail from simulation, and for the current detector the predictions overwhelmingly agree with the measured changes in bulk leakage current [62]. The dose expected for the upgrade can be described in terms of the total integrated luminosity, and is dependent on both the radial position from the beam, as well as the z-position of the sensitive volume (with higher radiation damage concentrated close to the interaction point). The fluence Φ_{eq} (in 1 MeV n_{eq} cm^{-2}) is described for a given z-position by

$$\Phi_{eq} = A\mathcal{L} \times R^{-k} \tag{6.5}$$

Polynomial term	Coefficient A_i ($\times\ 10^{13}$)	Coefficient k_i
z^0	3.63	2.30
z^1	2.72×10^{-4}	1.22×10^{-4}
z^2	-3.30×10^{-6}	-5.11×10^{-6}
z^3	-3.43×10^{-9}	6.22×10^{-9}
z^4	6.36×10^{-12}	2.79×10^{-11}
z^5	1.94×10^{-15}	-6.87×10^{-14}
z^6	1.63×10^{-18}	4.18×10^{-17}

Table 6.2: Polynomial fit for A and k, the constants used in the calculation of the fluence, versus station z position, where $A = \Sigma A_i z^i$ and $k = \Sigma k_i z^i$ for a given z. Fit taken from [62].

where the constants A and k vary as a function of z, $\mathcal{L}$ is the integrated luminosity and R is the radius in cm. Both A and k are taken from fits to the predicted fluence versus station z-position, and are described by a polynomial fit with coefficients A_i and k_i given in table 6.2. This has been incorporated into the simulation of the full VP detector in order to describe the fluence at all points in the detector for a given point during it's lifetime.

The final state of the detector ($\mathcal{L}$ = 50 fb^{-1}) is shown in two plots in figure 6.18. On the right hand plot, the fluence at each simulated hit point has been calculated such that the histogram represents the relative occupancy of the various regions. As might be expected with the sharp decay of the fluence with radial distance from the beam, relatively few hits occur in the most damaged regions (though these are evidently the most active parts of the detector). Averaging over all hits, around 27 % occur in regions where the fluence received exceeds 2×10^{15} 1 MeV n_{eq} cm^{-2} , with less than 5 % occurring in regions at a level of 5×10^{15} 1 MeV n_{eq} cm^{-2} or more. In the right hand plot, the relative occupancy of the detector is shown expanded in the z-direction, where the intensity at each z-Φ_{eq} point indicates the frequency that such a position is hit. This highlights the variation in maximum fluence observed on the different modules throughout the VP detector, but these high-fluence regions are again shown to have relatively low occupancy when compared with the total number of hits. The large area of detector receiving only a modest dose will have benefits in terms of the performance

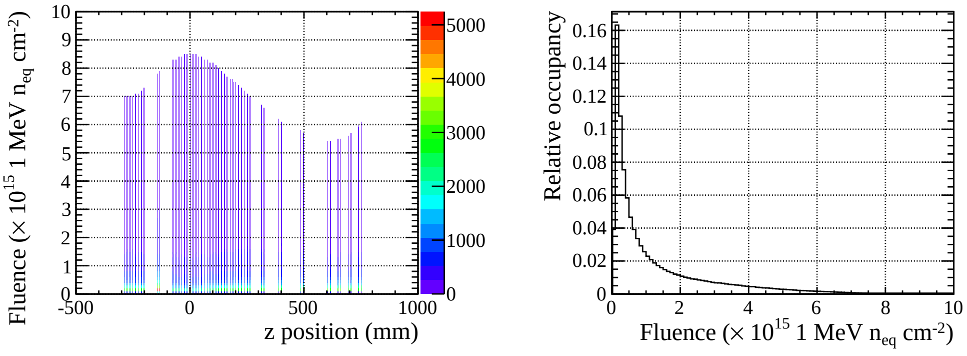

Figure 6.18: Left: (Arbitrary) Number of hits at the various fluence values on each station at the end of lifetime. Right: Relative hit occupancy in terms of the local radiation dose at end of lifetime (normalised to unity).

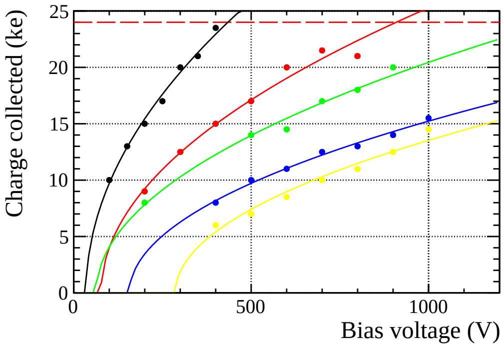

Figure 6.19: Fitted data (taken from [73]), showing the charge collected from a 300 µm sensor at different bias voltages (Most Probable Value) for samples irradiated to (black) 0.1, (red) 0.2, (green) 0.5, (blue) 1 and (yellow) 3 × 10^{15} 1 MeV n_{eq}. The dashed line represents the MPV of the deposited charge spectrum.

towards the end of its life.

6.4.2 Parametrisation of CCE

Given the fluence on each module at any point during the detector lifetime, it is necessary to describe the response of the silicon sensors, in particular the total amount of charge collected. For simplicity and due to the unknown high voltage distribution which may be implemented, it is assumed that all sensors must be operated at the same bias voltage. For the modelling of radiation damage effects, while there are a number of efforts to model physically the charge trapping and motion within irradiated sensors, for these global robustness tests an effective parametrisation from results shown in literature was used. The charge collected from sensors of different thicknesses was observed as a function of bias voltage and fluence, and given both the shape of the responses and the fact that at sufficient fluence the depletion voltage increases to many kilovolts, the response was fitted as

$$Q = a \times \sqrt{V + b} \tag{6.6}$$

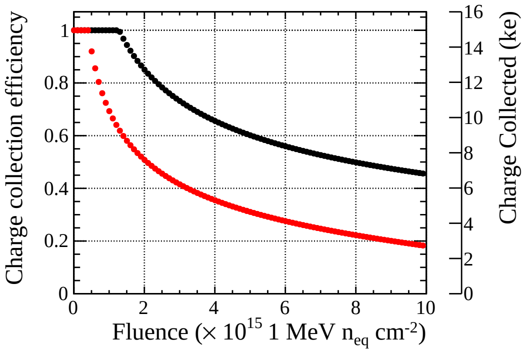

Figure 6.20: Estimated charge collection efficiency and absolute charge collected for a 200 µm thick sensor versus fluence, operated at (black) 1000 V and (red) 500 V.

where a and b are constants that depend on both the fluence and the sensor thickness, V is the applied bias voltage and Q is the collected charge in ke$^-$. An example of the fitted data (taken from [73]) is shown in figure 6.19, for a 300 µm sensor at a range of fluences up to 3 × 10^{15} 1 MeV $\mathrm{n_{eq}}$ $\mathrm{cm^{-2}}$. After fitting these data for each thickness the relationship between the integrated dose and the fit constants was determined. An estimate was then made for the case of a 200 µm sensor, which produces a response intermediate between those sensor thicknesses for which data was available. The result of this can be seen in figure 6.20, where both the CCE and absolute charge value are shown versus fluence, for an applied bias voltage of (black) 1000 V and (red) 500 V. The constants a and b for the determination of the total charge were estimated as

$$a = 2.2 \times 10^4\ \Phi_{eq}^{-0.306} \tag{6.7}$$

$$b = -0.136\ \Phi_{eq}^{0.228} + 191.6 \tag{6.8}$$

6.4.3 Implementation and Results

The calculations of local fluence and expected charge collection were implemented in the digitisation stage of the LHCb simulation framework. Instead of a direct scaling of the charge to match the expected CCE, the depth from which charge is collected was modified. This is a simplistic approximation of the effects observed in irradiated sensors, where charge trapping and the increase in N_{eft} result in charge collection predominantly from a thin region close to the depletion boundary. It allows both the reduction in lateral diffusion and the shift in residual distributions to be replicated in simulation (whereas normally the sensor measurement assumes a hit occurs half way through the sensor, this is not the case when charge is collected only close to the surface).

To confirm the implementation and observe the expected charge versus radius, simulations were first performed with the 4-bit ToT VELOPix description. The MPV of a convoluted Landau-Vavilov-Gaussian fit to the charge spectrum is shown in figure 6.21, for different points throughout the detector lifetime. Results are shown for an

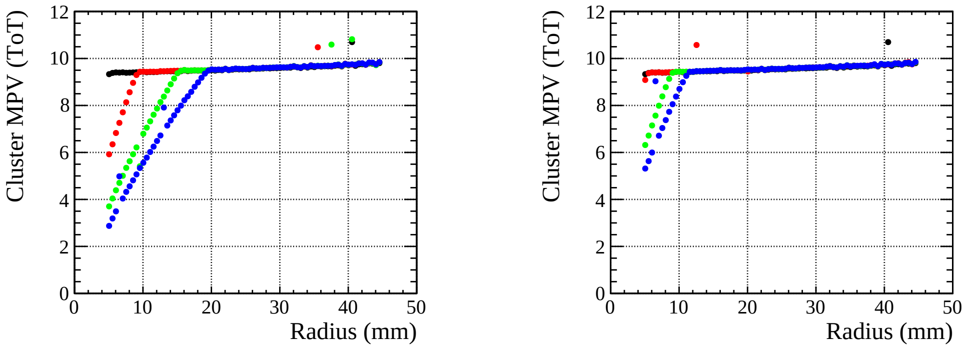

Figure 6.21: Collected charge versus radius, for (black) the unirradiated detector and for (red) 10, (green) 30 and (blue) 50 $\mathrm{fb^{-1}}$ of collected data. Left: Operation at 500 V, Right: Operation at 1000 V.

operation voltage of (left hand plot) 500 V and (right hand plot) 1000 V. In the latter case, even at the end of detector lifetime the collected charge is around 5 ke$^-$ for the innermost radius, well above the expected threshold of 1000 electrons. For 500 V the observed charge in the inner region becomes much closer to threshold, highlighting the importance of high voltage operation for the upgraded VELO detector.

For all subsequent plots, the ASIC description is that of the TDR (a binary ASIC with 1000 electron threshold). Given measurements of irradiated sensors showing a lateral diffusion greater than might be expected for an applied voltage of several hundred volts, the voltage used in the lateral drift calculation has been fixed to 100 V for all irradiated samples. This may have some effect on the single hit resolution, but is not expected to impact unduly on the impact parameter resolution, as shown from the relative insensitivity of moving from a 4-bit ToT to binary readout.

The single hit resolution versus angle is shown in figure 6.22, along with the distribution of track angles. Again the results are shown for both 500 V and 1000 V, where the improved charge collection efficiency at 1000 V leads to the smaller change in resolution. In either case, since the track angles are predominately perpendicular to the sensor surface and the resolution there is dominated by the prevalence of single pixel clusters, there are no significant changes in performance. A change from 12 μm to 15 μm resolution at the end of lifetime is not a considerable setback, as shown in the impact parameter resolution plots in figure 6.23. With the caveats noted above, in terms of the absolute charge collected and lateral diffusion observed, the degradation of the the detector performance over its full lifetime does not appear to be severe, and critically is further reduced by increasing the applied voltage. For both scenarios however there appears to be no change in the tracking performance, with the efficiency for track reconstruction of Long tracks virtually unchanged throughout the full lifetime (figure 6.24). This is only to be expected given the redundancy in number of stations, and the relatively high signal to threshold ratio even after irradiation.

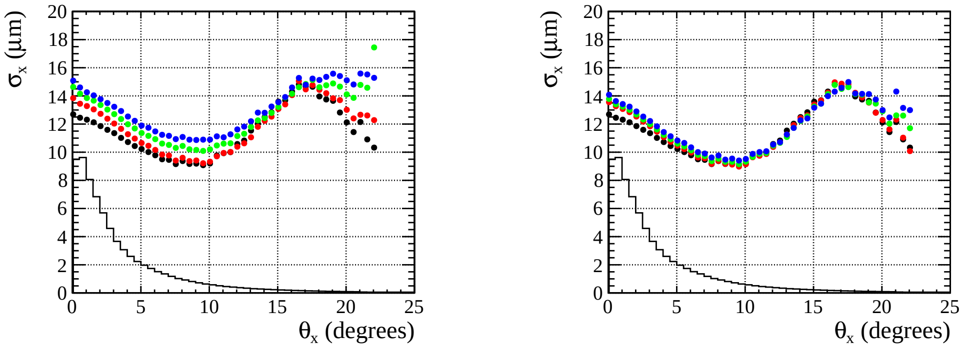

Figure 6.22: Single hit resolution versus track angle in the direction of rotation, for (black) the unirradiated detector at 60 V and after an integrated luminosity of (red) 10, (green) 30 and (blue) 50 fb^{-1}. The irradiated simulations are performed with an applied bias voltage of (left) 500 V and (right) 1000 V.

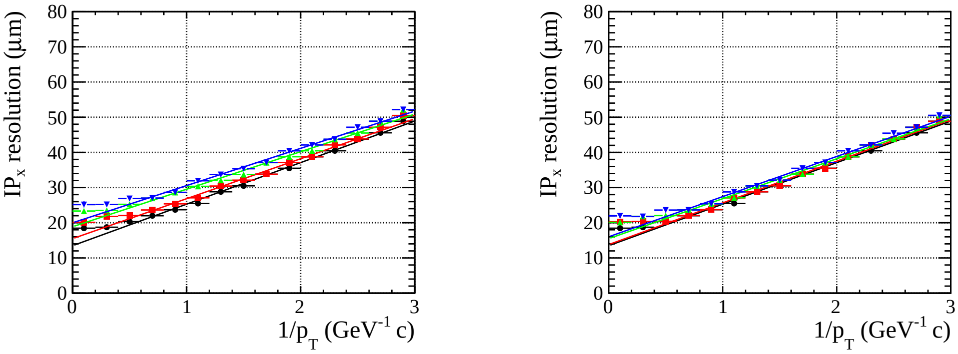

Figure 6.23: IP resolution for Long tracks versus inverse transverse momentum, for (black) the unirradiated detector at 60 V and after an integrated luminosity of (red) 10, (green) 30 and (blue) 50 fb^{-1}. The irradiated simulations are performed with an applied bias voltage of (left) 500 V and (right) 1000 V.

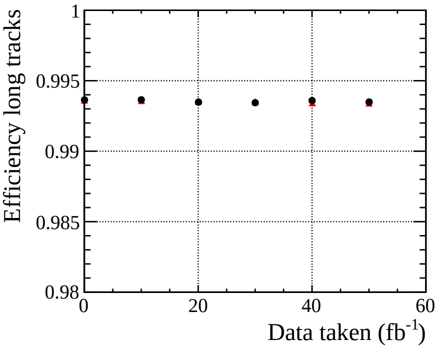

Figure 6.24: Reconstruction efficiency for Long tracks versus integrated luminosity, for (black) 1000 V and (red) 500 V operation.

6.5 Effects of a Pixel Detector on the LHCb Trigger Scheme

The driving force behind each of the sub-detector upgrades is primarily the design of the LHCb trigger: the requirement to read out the full detector for each 25 ns bunch crossing and make a decision about which events should be retained for offline analysis. The saturation of the current trigger architecture with increasing luminosity (for hadronic final-states) limits what gains may be achieved by simply increasing the detector granularity. But it is also true that the exact trigger scheme employed will depend on the performance of each sub-detector, and that any technology decisions taken will have an impact upon the online event selection. This is the case for the VP design, which as indicated above differs from the current VELO strip detector in the way which it treats track hits during the linear fit performed during the track reconstruction stage.

6.5.1 Multiple Scattering in the VELO

The effects of multiple scattering were shown above in figure 6.8, where the uniform weighting of hit points in the detector has the effect of distorting the track fit where multiple scattering occurs. As this is more drastic for lower momenta, the impact parameter resolution degrades for these tracks while high momentum tracks are relatively unchanged. This has been corrected above with the use of a radially-dependent hit error, which provides an IP resolution closer to the full Kalman-fitted track after the momentum is measured via the downstream tracking.

This principle might also be used to instead *degrade* the IP resolution of low momentum tracks, pulling the fitted track further away from the true trajectory. This is illustrated in figure 6.25, where the IP resolution is shown for the cases above (linear fit, Kalman fit and radial hit errors) along with another track fit where the errors applied to each hit are extremely small (5 µm). The IP resolution relative to the radial assignment of hit errors is almost a factor of 2 for tracks with transverse momentum of 300 MeV/c. This opens up the possibility of identifying low-momentum tracks purely within the VELO, before the downstream tracking has been performed.

With a view to identifying low-momentum tracks in the VP detector, the track

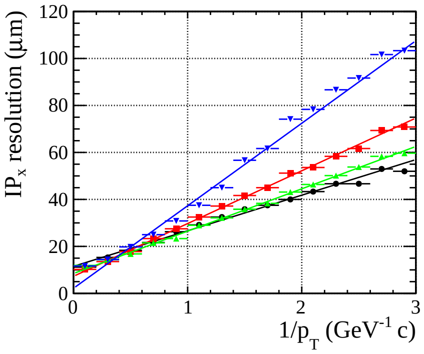

Figure 6.25: IP resolution in the global x-direction for (black) Long and (red) VELO tracks. Linearly fitted VELO tracks with differing hit errors are also shown, where (green) the error varies with the hit radius and (blue) where the hit is given a stronger (uniform) weight.

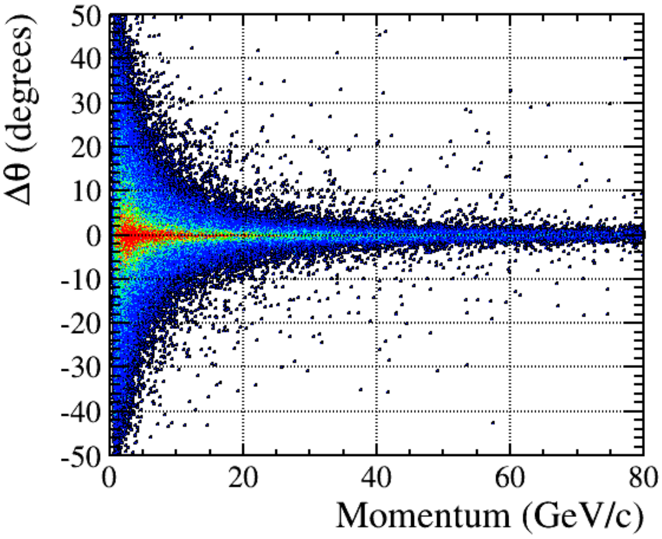

Figure 6.26: Change in track opening angle when the track is refitted using different hit weights, showing the dependence on track momentum.

fit is performed twice: once with hit errors which provide an IP resolution similar to the Kalman filter, and once with hit errors which increase the pull of the scattered downstream hits. The difference between the two fitted tracks can then be used to identify those produced by low-momentum particles. For proof of concept, this has been taken as the track opening angle with respect to the LHC beam (θ). The change in this angle versus particle momentum is shown in figure 6.26, where the potential discriminating power can be seen clearly.

6.5.2 Trigger Requirements

The use of multiple scattering in the VP to identify low momentum tracks is of interest due to the current reconstruction scheme expected for the online trigger. Uncertainties over the exact computing resources that will be available at the point of installation have led to the proposal of a temporary Low Level Trigger (LLT) which will function similar to the existing L0 hardware trigger, and reduce the event rate going into the CPU farm. The output of the LLT would be scaled back as the size of the farm increases, until the full 40 MHz bunch crossing rate is processed by the HLT. As for the present software trigger, this would be split into two parts, so-called HLT1 and HLT2, which are run sequentially. HLT1 performs a partial reconstruction of the event, selecting tracks based on their impact parameter and momentum, while HLT2 performs the full reconstruction and builds candidate objects which are stored according to numerous selection algorithms.

Current trigger studies relating to the upgrade have focussed on the performance of the HLT1 algorithm *Hlt1TrackAllL0*. This trigger line accounted for 75 % of the total HLT1 bandwidth in 2012, and is the main trigger line for analyses which do not contain muon final states. The algorithm follows the steps illustrated in figure 6.27, with the cut values for each step shown in table 6.3. First, the full VELO tracking is performed and

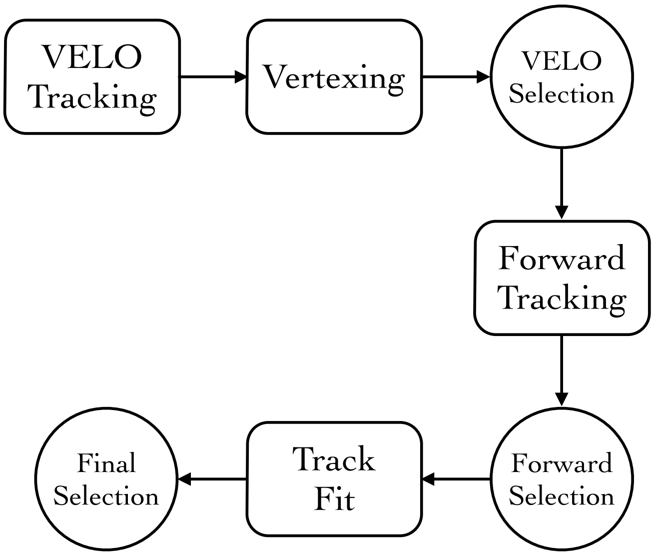

Figure 6.27: Flow diagram of the HLT1 algorithm Hlt1TrackAllL0.

the primary vertices are reconstructed. A cut is then applied to the impact parameter of each track, in order to reduce the number of tracks passed to the forward tracking algorithm, which matches VELO tracks with hits in the downstream stations. This is the main timing constraint on the trigger processing, and dominates the HLT1 timing (and thus maximum bandwidth for given CPU resources or the required size of the CPU farm). It is also this cut that introduces a lifetime bias in the triggered data, as tracks originating from the primary vertex which have a low decay time will not be passed to the forward tracking. Replacing this with a cut on the track scattering would enable the removal of low momentum tracks (where a cut of 10 GeV/c is applied after the forward tracking) without introducing a decay time bias, if the number of tracks which are accepted is comparable to that with the existing cuts.

Quantity	Cut value
VELO Selection	
IP	> 0.1 mm
VELO hits	> 3
Forward Selection	
p	> 10 GeV/c
p_T	> 1.7 GeV/c
T hits	> 10
Final Selection	
Track χ^2	< 1.5
IP χ^2	> 16

Table 6.3: Cut values for the selections performed by the Hlt1TrackAllL0 trigger line.

6.5.3 Low Momentum Track Suppression

In order to suppress low momentum tracks in the Hlt1TrackAllL0 algorithm, a simple cut on the change in track angle between the two fitting methods ($\Delta\theta$) has been chosen, with an optimisation carried out separately for tracks with different numbers of measurement points. This is intended to show the feasibility of such an approach, though fits containing additional variables may prove more discriminating. The tuning of the cut value is determined in each case by specifying the required number of tracks to be retained above a given momentum. For initial tests the momentum used was 20 GeV/c, and the relevant figures of merit for the trigger have been evaluated for a range of retention rates.

Figure 6.28 shows the track retention rate as a function of the cut on $\Delta\theta$, for both the total number of tracks (black) and the number of tracks with momentum above 20 GeV/c (red). The right hand plot shows (black) the raw momentum distribution and (red) the track momentum after a cut corresponding to 90 % signal retention is applied (ie. where 90 % of tracks with momentum greater than 20 GeV/c are kept). Low momentum tracks can be seen to have been suppressed by a significant amount, with only around 40 % of VELO tracks surviving the cut. The number of VELO tracks which are retained, as a function of the signal retention, can be seen in figure 6.29. The

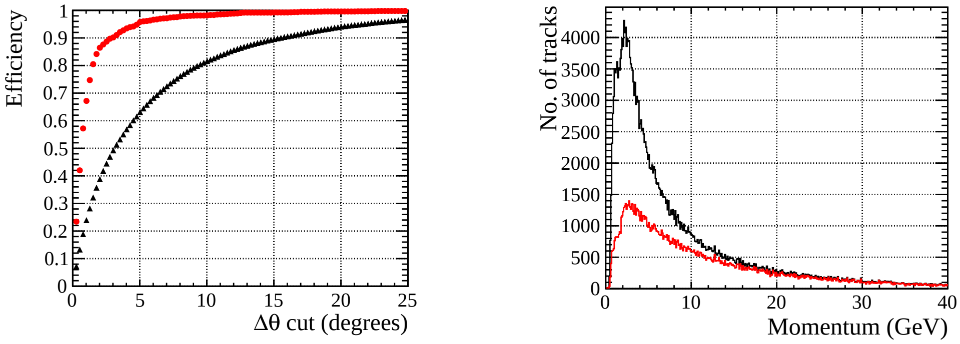

Figure 6.28: Left: Track retention for (black) all VELO tracks and (red) tracks with $p > 20$ GeV/c. Right: Initial momentum spectrum (black) and track momenta after the scattering cut (red).

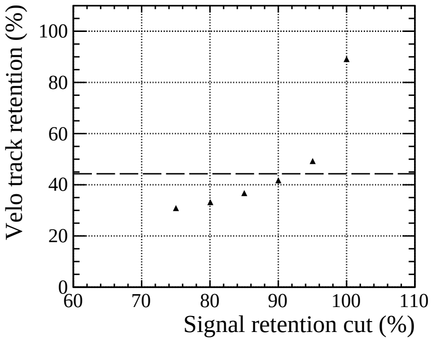

Figure 6.29: VELO track retention for different signal retention cuts. The horizontal line indicates the VELO track retention of the currently envisaged IP cut.

horizontal dashed line indicates the number of VELO tracks retained with an IP cut of 0.1 mm, as envisaged for the upgrade. It can be seen that a signal retention cut of 90 - 95 % produces a similar result.

Two important performance criteria are shown in figure 6.30; the average time taken to process each event and the number of events which are retained after all of the selection cuts in HLT1 are applied. The timing in particular is required to be similar to the current method using an IP cut. The results are shown for both (black) minimum bias events and (red) those simulated with $B^0 \to K^{*0}\mu^+\mu^-$. For the total event retention the minimum bias retention with a cut on track scattering is similar to the use of an IP cut. For a signal retention of 95 % the yield of $B^0 \to K^{*0}\mu^+\mu^-$ events is also the same, although in this case the timing for minimum bias events increases by around 30 %. As a rough indication though, the scattering method does appear to have similar discriminating power and give approximately the same results as using a cut on the impact parameter, without an explicit lifetime cut. To confirm this, the efficiency of these cuts as a function of the B^0 decay time are plotted in figure 6.31, where the (black) IP cut shows a sharp drop for decay times less than 1 ps. The efficiency is much less biased for the case of both (red) 95 % and (green) 90 % signal retention with the scattering cut applied.

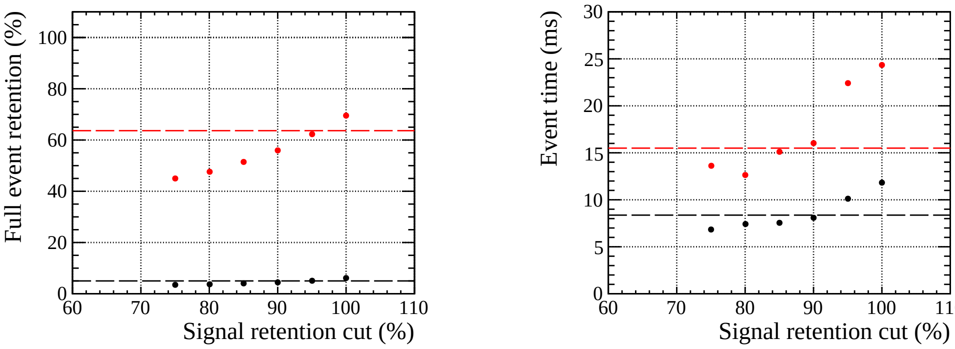

Figure 6.30: Left: Event retention for different cuts on the track scattering. Right: Processing time versus scattering angle cut. Both plots contain results for (black) minimum bias and (red) $B^0 \to K^{*0}\mu^+\mu^-$ events. The horizontal line indicates the performance of the currently envisaged IP cut.

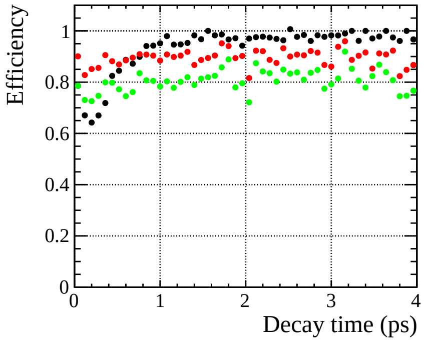

Figure 6.31: Decay time acceptance for $B^0 \to K^{*0}\mu^+\mu^-$ decays using (black) a cut on the IP of 100 µm, and a scattering cut with (red) 95 % and (green) 90 % signal retention.

All of the above results indicate that low momentum tracks can be preferentially discarded at the end of the VELO tracking sequence by observing the amount of scattering in the detector. The number of signal tracks lost for a representative decay channel ($B^0 \to K^{*0}\mu^+\mu^-$) by such an approach does not exceed the number currently lost by the application of a cut on IP, while the CPU time required remains comparable. The replacement of this decay time biasing cut in the trigger may be of use for analyses involving either lifetime measurements or time-dependent decays. Subsequent to this work, the forward tracking algorithm for the detector has been rewritten, taking into account the technology choice (scintillating fibres) for the downstream tracking stations. It now appears feasible, for the projected size of the computing farm available during the upgrade, to perform the full reconstruction chain on all tracks. This would effectively remove the need to perform a selection cut after the VELO tracking stage, though this will be examined during commissioning of the upgraded detector.

6.6 Summary

The performance of the upgraded VELO detector has been presented, paying attention to the influence of the single hit resolution on the IP determination. Given the generally low track inclination with respect to the sensor surfaces and the dominant effect of single pixel clusters, the hit resolution is not observed to change significantly for the majority of tracks when the 4-bit charge measurement is replaced by a binary ASIC. The IP resolution in both cases has been found to be the same, and this has been adopted as part of the VELOPix design. Studies with super pixel structures have been studied to gauge the possible gains in compressing pixel information on-detector, and compression ratios of greater than 20 % have been achieved.

Given the proximity of the upgraded VELO to the LHC beam, the effects of radiation damage on its performance have been investigated. Charge loss is expected to result in a signal of approximately 6000 electrons in the innermost region, if operated at 1000 V. Nonetheless, no significant change in either the track efficiency or the IP resolution have been observed.

A method to preferentially remove low-momentum tracks in the selection performed before the forward tracking has also been shown. This involves measuring the change in track angle when the outer hits on a track are given more or less weighting in the linear track fit. Rejection of ~ 50 % of low momentum VELO tracks has been achieved, for a loss of only 5 % of tracks with momentum greater than 20 GeV/c. This has been shown to comply with the CPU requirements for the online trigger, and reduce the decay time bias in the trigger selection. Further improvements using a multivariate discriminator are likely but have not as yet been performed.

7

Conclusions

This thesis presented an overview of the LHCb detector and its search for unknown contributions to physical observables at the LHC from physics beyond the standard model. Each of the sub-detectors were described, and the rationale behind the upgrade of the experiment was laid out. The importance of the move to a fully software-based trigger was highlighted, and the implications for each of the existing sub-detectors was discussed. For the VELO, the replacement of the existing silicon strip detector with a new hybrid pixel detector was shown, with emphasis on the difficulties in the design of such a system. The challenges posed by the data that the innermost ASICs are required to transmit, the damage inflicted by operation in such a high fluence environment, and the design of a module incorporating microchannels for efficient cooling were all discussed.

In chapter 2, an overview of the physical principles behind radiation detection with silicon was shown, starting from a description of the band structure that leads to the behaviour of semiconductors. The extra energy levels introduced by the presence of group III or V impurity atoms, close to either the conduction or valence band, was shown to lead to the basic detecting element of a silicon detector: a p-n junction. The behaviour of charge in this region and the establishment of a built-in depletion region was described, along with the behaviour of this region under the application of an external bias voltage. The main mechanisms for energy transfer during the passage of charged particles through silicon was shown and the most probable energy loss given in the form of the Landau-Vavilov formula. Finally, the damage mechanisms for non-ionising energy losses in silicon detectors was briefly detailed, with the corresponding changes in physical attributes for the silicon such as effective dopant concentration, leakage current and charge trapping.

Chapter 3 described the construction and performance evaluation of a particle beam telescope based on the Timepix ASIC. Two pattern recognition algorithms were presented, based on either a simple z-extrapolation of hits from one plane to the next, or the continual fitting of a track during construction and using the extrapolated track intercept to search for new clusters. This latter method was found to suffer from the relatively poor constraint of the track over the distance between the two telescope arms, leading to a track reconstruction efficiency of $(91.79 - 0.06 \times n_{tracks})$ %. This is in comparison to the simple z-axis extrapolation which was found to have a reconstruction efficiency of $(95.79 - 0.05 \times n_{tracks})$ %. A pointing precision at the device under test of less than $2\,\mu$m was achieved. The method of attaching a high precision timestamp, through the association of coincident scintillator triggers with the spatially reconstructed timing plane clusters, achieved an efficiency of 89.31 ± 0.13 %, consistent with the calculated Poisson separation of particles in time for the given beam intensity. The integration of external devices, specifically considering the case of a generic LHC-style readout, was discussed.

In chapter 4 the testbeam results for a variety of unirradiated pixel sensors, mounted on Timepix ASICs, are presented. The samples vary in their thickness, substrate resistivity, implant/bulk type and operating conditions, and in each case the single hit resolution versus angle is investigated. In the first such comparison, a single device operated under different applied bias voltages is observed to show a strong dependence on the electric field strength, leading to a resolution of between 4 µm and 10 µm at perpendicular incidence. This is attributable to the increased diffusion of charge and corresponding reduction in the number of single pixel clusters. The resolution is observed to become independent of the applied field once the later track length is comparable to the pixel pitch. The effects of increasing the discriminator threshold are subsequently shown for a single device, operated in otherwise identical conditions. The resolution in both cases behaves identically with respect to the incident track angle, with a simple angular offset for the data taken with increased threshold (due to the loss of the charge to neighbouring pixels). A series of comparisons are then made between technology types, specifically p-on-n versus n-on-n and n-on-p versus n-on-n. In all cases the devices are found to behave approximately the same, with the exception of the low bias data where an n-on-n device was found to have a marginally improved resolution compared with n-on-p.

After the analysis of the testbeam data is presented, an analytic expression is constructed to describe the effects observed in data. The expressions for lateral charge diffusion are derived, and a contributing term to the resolution due to charge diffusion is defined, based only on the integrated diffusion through the sensor depth and the most probable size of the expected charge deposits. Two more resolution terms were defined, namely that due to variation in the charge deposited along the track length, and the track extrapolation error. A comparison with the testbeam data showed this model to represent the data well, and a series of hypothetical devices are simulated in order to isolate the contributions from different sensor type and resistivity. Some comments upon the future direction of silicon sensors are provided, examining the move to thinner devices operated with lower thresholds and increasing pixel granularity.

Chapter 5 shows the results of integration tests performed with the ATLAS FE-I4 ASIC inside the Timepix telescope. An issue with the veto applied to the coincident scintillator signals recorded by the TDC was observed to result in the mis-association of tracks with TDC times, resulting in an apparent efficiency drop for the device under test. This was understood and rectified for subsequent data-taking, and the analysis of the original device was able to continue by assigning a purity to the collected data based on the telescope timing residuals. After correcting for this, the assembly was observed to be fully efficient. Subsequently, two irradiated silicon detectors mounted on FE-I4 ASICs at fluences of 2 and 4×10^{15} 1 MeV $\mathrm{n_{eq}}$ cm^{-2} were tested in order to serve as input to the design of the upgraded VELO. The charge collection across the pixel unit cell was mapped for both devices at a range of bias voltages and under rotation to angles representative of those found in the VELO. For the lower-irradiated device, the single hit resolution was observed to reach 8 µm at a track angle of 22°, while the more heavily damaged sample remained relatively static around 13 µm. The limited response of the 4×10^{15} 1 MeV n_{eq} cm^{-2} suggested that the active depth of the sensor was less than the sample irradiated to 2×10^{15} 1 MeV n_{eq} cm^{-2} , in addition to the lower charge collected by the device.

For the final chapter, simulations performed within the LHCb software framework are presented, beginning with the characteristic distribution of tracks both spatially within the detector and in terms of their momentum. The single hit resolution expected for the upgraded VELO, as envisaged at the time of the Technical Design Report, is shown alongside the cluster size, fraction and impact parameter resolution. The 3D resolution on the single hit position is found to vary between 12 µm and 15 µm depending on the track opening angle, while the 3D resolution on the impact parameter varied between 25 µm and 70 µm depending on the transverse track momentum. Efficiencies for the reconstruction of different track types are compared between the existing and proposed detectors operating at the expected instantaneous luminosity of the upgrade, and a significant improvement is achieved through the replacement of the current VELO.

A series of simulation studies follow this, first considering the effects of the linear track fit performed before the extrapolation of tracks to the downstream tracking stations. The uniform pixel pitch (and corresponding hit uncertainty) is found to degrade the impact parameter resolution for this process, due to the effects of multiple scattering and the need to reduce the contribution of hits that have scattered to the track fit. An initial reweighting of the pixel hits based on their radial position is found to improve the impact parameter resolution by around 10 µm for low momentum tracks. The effects of varying the charge discrimination offered by the VELOPix ASIC are then considered. The difference in impact parameter resolution found between a 4-bit charge measurement and binary readout were observed to be negligible, leading to the binary option adopted as the baseline for the ASIC design. The digital architecture is also applied to the simulated data in order to investigate possible gains in the quantity of transmitted data for the implementation of a super pixel structure. A 4 × 4 super pixel with a 150 µm silicon sensor is found to give compression ratios of 12 - 22 % depending on the ASIC location.

Further to the studies on the performance of irradiated silicon sensors, a parametrisation from literature is presented of the expected charge collection as a function of bias and received fluence. This has been implemented in the simulation of the full detector in order to make some initial predictions of the detector performance throughout its lifetime. The tracking efficiency for Long tracks is observed to remain unchanged at 99.4 %, and the degradation of the impact parameter resolution is deemed not to be significant. This relies on the application of 1000 V at the end of the detector lifespan.

Finally, a method for preferentially removing low-momentum tracks within the upgraded VELO detector is proposed. This involves the removal of tracks where heavier weighting of the outer hits significantly shifts the linearly fitted track angle, and is intended to be used before the downstream tracking algorithm has been performed and a measurement of the track momentum made. A rejection of around 50 % of all VELO tracks has been achieved with a loss of only 5 % of tracks with momentum greater than 20 GeV/c. This cut could be used as a replacement for the current cut on the track impact parameter, virtually eliminating the decay time bias that such a cut naturally introduces. Further improvements with a multivariate discriminator are expected.

The planned upgrade of the LHCb detector during the second long shutdown of the

LHC continues to move forwards, and detailed designs for each of the sub-detectors exist and are being fine-tuned as prototypes are constructed. The planned installation date, while challenging, is expected to be sufficient for the full detector upgrade in time for operations in LHC Run 3. Until that point, the construction of the sub-detectors will continue in tandem with searches for the as-yet unobserved physics beyond the standard model. The possibilities opened by the upgrade of the detector are wide-ranging, providing some hope that the next steps for particle physics may yet come from the LHC.

References

[1] `www.cern.ch`. [Cited page 1].

[2] UA1 Collaboration, *Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at* $\sqrt{s} = 540$*-GeV*, Phys. Lett. **B122** (1983) 103. [Cited page 1].

[3] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Physics Letters B **716** (2012), no. 1 1 . [Cited page 1].

[4] `http://te-dep-epc.web.cern.ch/te-dep-epc/machines/general.stm`. [Cited page 2].

[5] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008). [Cited page 1].

[6] LHCb Collaboration, *LHCb : Technical Proposal.* CERN-LHCC/98-004. 1998. [Cited pages 1 and 5].

[7] ATLAS Collaboration, *ATLAS: technical proposal for a general-purpose pp experiment at the Large Hadron Collider at CERN.* CERN-LHCC/94-43. 1994. [Cited page 1].

[8] CMS Collaboration, *Technical Proposal.* CERN-LHCC/94-38. 1994. [Cited page 1].

[9] ALICE Collaboration, *ALICE: Technical proposal for a Large Ion collider Experiment at the CERN LHC.* CERN-LHCC/95-71. 1995. [Cited page 1].

[10] `www.sba.web.cern.ch`. [Cited page 1].

[11] R. P. Feynman, *The theory of positrons*, Phys. Rev. **76** (1949) 749. [Cited page 3].

[12] `www.physik.uzh.ch/∼che/FeynDiag`. [Cited page 4].

[13] LHCb Collaboration, *Road map for selected key measurements from LHCb*, CERN-LHCb-PUB-2009-029. Comments: 379 pages. [Cited page 4].

[14] LHCb Collaboration, *Measurement of* $\sigma(pp \to b\bar{b}X)$ *at* $\sqrt{s}$*=7 TeV in the forward region*, Phys. Lett. B **694** (2010) 209. [Cited page 5].

[15] LHCb collaboration, *LHCb VELO: Technical Design Report*, CERN-LHCC-2001-011. [Cited page 6].

[16] LHCb collaboration, *LHCb inner tracker: Technical Design Report*, CERN-LHCC-2002-029. [Cited page 6].

[17] LHCb collaboration, *LHCb outer tracker: Technical Design Report*, CERN-LHCC-2001-024. [Cited page 6].

[18] LHCb collaboration, *LHCb magnet: Technical Design Report.* Technical Design Report LHCb. CERN, Geneva, 2000. [Cited page 6].

[19] LHCb collaboration, *LHCb Detector Performance. For submission to: IJMPA*, . [Cited pages 9 and 12].

[20] LHCb collaboration, *LHCb RICH: Technical Design Report.* Technical Design Report LHCb. CERN, Geneva, 2000. [Cited page 10].

[21] LHCb collaboration, *LHCb muon system: Technical Design Report.* Technical Design Report LHCb. CERN, Geneva, 2001. [Cited page 10].

[22] LHCb collaboration, *LHCb calorimeters: Technical Design Report.* Technical Design Report LHCb. CERN, Geneva, 2000. [Cited page 10].

[23] G. Haefeli *et al.*, *The LHCb DAQ interface board TELL1*, Nucl. Instrum. Meth. **A560** (2006) 494. [Cited page 13].

[24] LHCb HLT project, J. Albrecht, V. Gligorov, G. Raven, and S. Tolk, *Performance of the LHCb High Level Trigger in 2012*, J. Phys. Conf. Ser. **513** (2014) 012001, `arXiv:1310.8544`. [Cited page 14].

[25] LHCb Collaboration, *LHCb VELO Upgrade Technical Design Report*, Tech. Rep. CERN-LHCC-2013-021, CERN, Geneva, Nov, 2013. [Cited pages 16, 17, 20, 23, 110, 111, and 117].

[26] LHCb Collaboration, *LHCb Tracker Upgrade Technical Design Report*, Tech. Rep. CERN-LHCC-2014-001. LHCB-TDR-015, CERN, Geneva, Feb, 2014. [Cited pages 16, 17, and 61].

[27] LHCb Collaboration, *LHCb PID Upgrade Technical Design Report*, Tech. Rep. CERN-LHCC-2013-022, CERN, Geneva, Nov, 2013. [Cited pages 18 and 61].

[28] *TORCH: A Large-Area Detector for Precision Time-of-Flight Measurements at LHCb*, Physics Procedia **37** (2012) 626 . [Cited pages 19 and 61].

[29] LHCb, Collaboration, *LHCb Trigger and Online Upgrade Technical Design Report*, Tech. Rep. CERN-LHCC-2014-016, CERN, Geneva, May, 2014. [Cited page 19].

[30] *Timepix, a 65k programmable pixel readout chip for arrival time, energy and/or photon counting measurements*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **581** (2007), no. 1–2 485 . [Cited pages 20 and 45].

[31] T. Poikela *et al.*, *Architectural modeling of pixel readout chips Velopix and Timepix3*, JINST **7** (2012) C01093. [Cited page 21].

[32] F. Alessio and R. Jacobsson, *System-level Specifications of the Timing and Fast Control system for the LHCb Upgrade*, . [Cited page 21].

[33] V. Gromov *et al.*, *Development and applications of the Timepix3 readout chip*, PoS **VERTEX2011** (2011) 046. [Cited pages 22 and 58].

[34] A. Affolder, P. Allport, and G. Casse, *Charge collection efficiencies of planar silicon detectors after reactor neutron and proton doses up to 1.6 × 10^{16} n_{eq} cm^{-2}*, Nucl. Instrum. Meth. **A612** (2010) 470. [Cited pages 22 and 102].

[35] B. D. Hyams *et al.*, *A silicon counter telescope to study short-lived particles in high-energy hadronic interactions*, Nucl. Instrum. Meth. **205** (1982) 99. [Cited page 24].

[36] D. A. Glaser, *Some effects of ionizing radiation on the formation of bubbles in liquids*, Phys. Rev. **87** (1952) 665. [Cited page 24].

[37] G. Charpak, R. Bouclier, T. Bressani, J. Favier, and C. Zupancic, *The use of multiwire proportional counters to select and localize charged particles*, Nucl. Instrum. Meth. **62** (1968) 262. [Cited page 24].

[38] C. Langenbruch, *Measurement of the B_s^0 mixing phase in the decay $B_s^0 \to J/\psi\phi$ with the LHCb experiment*, . [Cited page 24].

[39] S. Kleinfelder, *VLSI FOR PHYSICS*, . [Cited page 24].

[40] H. Spieler, *Semiconductor detector systems*, Ser. Semicond. Sci. Tech. **12** (2005) 1. [Cited page 24].

[41] H. Bichsel, D. Groom, and S. Klein, *Passage of particles through matter*, . [Cited pages 34, 36, and 37].

[42] H. Bethe, *Theory of the passage of fast corpuscular rays through matter*, Annalen Phys. **5** (1930) 325. [Cited page 35].

[43] H. Bichsel, *Straggling in Thin Silicon Detectors*, Rev. Mod. Phys. **60** (1988) 663. [Cited page 36].

[44] L. Landau, *On the energy loss of fast particles by ionization*, J. Phys. (USSR) **8** (1944) 201. [Cited page 36].

[45] G. Moliere, *Theory of the scattering of fast charged particles. 2. Repeated and multiple scattering*, Z. Naturforsch. **A3** (1948) 78. [Cited page 37].

[46] S. Ramo, *Currents induced by electron motion*, Proc. Ire. **27** (1939) 584. [Cited page 40].

[47] M. Moll, *Radiation damage in silicon particle detectors: Microscopic defects and macroscopic properties.* PhD thesis, 1999. [Cited pages 41 and 43].

[48] R. Wunstorf, *A Systematic investigation of the radiation hardness of silicon detectors for high-energy physics experiments.* PhD thesis, 1992. [Cited page 44].

[49] J. Lange, J. Becker, E. Fretwurst, R. Klanner, and G. Lindstrom, *Properties of a radiation-induced charge multiplication region in epitaxial silicon diodes*, Nucl. Instrum. Meth. **A622** (2010) 49, `arXiv:1007.4735`. [Cited page 44].

[50] K. Akiba *et al.*, *Charged Particle Tracking with the Timepix ASIC*, Nucl. Instrum. Meth. **A661** (2012) 31, `arXiv:1103.2739`. [Cited pages 45, 57, and 63].

[51] `http://aida.web.cern.ch`. [Cited page 45].

[52] K. Akiba *et al.*, *The Timepix Telescope for High Performance Particle Tracking*, Nucl. Instrum. Meth. **723** (2013) 47 , `arXiv:1304.5175`. [Cited page 49].

[53] *A gigabit per second read-out system for medipix quads*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **633, Supplement 1** (2011), no. 0 S22 . [Cited page 50].

[54] `http://root.cern.ch`. [Cited page 51].

[55] `http://www.boost.org`. [Cited page 51].

[56] G. Barrand *et al.*, *GAUDI - A software architecture and framework for building HEP data processing applications*, Comput. Phys. Commun. **140** (2001) 45. [Cited page 51].

[57] M. B. Kennel, *KDTREE 2: Fortran 95 and C++ software to efficiently search for near neighbors in a multi-dimensional Euclidean space*, ArXiv Physics e-prints (2004) `arXiv:physics/0408067`. [Cited page 52].

[58] `http://rd50.web.cern.ch`. [Cited page 61].

[59] `http://medipix.web.cern.ch`. [Cited page 61].

[60] M. T. Alexander, *Constraints on mixing and CP-violation in the neutral charmed meson system at LHCb.* PhD thesis, University of Glasgow, 2012. [Cited pages 62 and 116].

[61] X. Wu, J. Kalliopuska, S. Eränen, and T. Virolainen, *Recent advances in processing and characterization of edgeless detectors*, Journal of Instrumentation **7** (2012), no. 02 C02001. [Cited page 64].

[62] P. Collins *et al.*, *The LHCb VELO upgrade*, Nucl. Instrum. Meth. **A636** (2011) S185. [Cited pages 93, 125, and 126].

[63] X. Llopart Cudie, C. Frojdh, and M. Campbell, *Design and Characterization of 64K Pixels Chips Working in Single Photon Processing Mode.* PhD thesis, Mid Sweden U., Sundsvall, Sundsvall, 2007, Presented on 25 May 2007. [Cited page 93].

[64] M. Karagounis *et al.*, *Development of the ATLAS FE-I4 pixel readout IC for b-layer upgrade and Super-LHC*, Proceedings, Topical Workshop, TWEPP-08, Naxos, Greece, 15-19 September 2008 (2008) 70. [Cited pages 93 and 94].

[65] S. Terzo *et al.*, *Heavily Irradiated N-in-p Thin Planar Pixel Sensors with and without Active Edges*, JINST **9** (2014) C05023. [Cited page 93].

[66] W. Buttinger, *The ATLAS Level-1 Trigger System*, ATL-DAQ-PROC-2012-024 (2012). [Cited page 94].

[67] ATLAS Collaboration, J. Jentzsch, *Quality assurance and functionality tests on electrical components during the ATLAS IBL production*, JINST **8** (2013) C02048. [Cited page 95].

[68] `www6.slac.stanford.edu`. [Cited page 95].

[69] `www.mpp.mpg.de`. [Cited page 96].

[70] `www.izm.fraunhofer.de`. [Cited page 96].

[71] `www.kit.edu`. [Cited page 96].

[72] `lansce.lanl.gov/`. [Cited page 96].

[73] G. Casse, A. Affolder, P. Allport, and M. Wormald, *Measurements of charge collection efficiency with microstrip detectors made on various substrates after irradiations with neutrons and protons with different energies*, PoS **VERTEX2008** (2008) 036. [Cited pages 127 and 128].