Characterisation of advanced silicon pixel detector prototypes for the Inner Tracker of the ATLAS experiment for the High Luminosity LHC

Abhishek Sharma
St Cross College, Oxford





Thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy at the University of Oxford

Michaelmas Term, 2019

Abstract

In preparation of the high luminosity running of the Large Hadron Collider at CERN, the ATLAS experiment will undergo a major overhaul of its mechanical structures and sensor technologies, including their services and readout system. A novel monolithic active pixel sensor based on 180 nm technology is investigated as a possible solution for layer 4 of the ATLAS Inner Tracker's (ITk) pixel detector.

Three such chip designs are assessed through extensive laboratory measurements as well as test-beam campaigns in both unirradiated and several after doses of irradiation. Analyses focus on their suitability for operation at the LHC in the 1×10^{15} n_{eq}/cm^2 fluence, and 70 Mrad regime. Results show that chips possessing optimised wafer substructures in pixel corners allow for high efficiency up to the target irradiation doses. In order to perform additional measurements to show correlations between the different pixel design geometries and their effects on charge sharing, a new edge-transient current technique facility was built. Results from these measurements identified charge-sharing fractions for a range of pixel and electrode dimensions.

In an effort to qualify all mechanical structure (stave) prototypes being considered for the ITk, a thermal characterisation facility was developed to identify the thermal figure of merit of all such designs being considered and to assess their suitability for the different sub-detector layers they are targeting. Results obtained for all designs include the dependency of stave orientations as well as convective air-flow on the thermal figure of merit. A thermal cycling campaign was also conducted on five prototypes for failure testing. Several guidelines are proposed to enhance future characterisations of this kind.

In the context of the development of 1:1 scale demonstrator models housing fully mounted sensors and thermal modules, cable saver boards were designed to allow interfacing between delicate flex cables emanating from the sensors and laboratory instrumentation. The different board designs are described along with details on their integration.

Following the developments of the cable saver boards, contributions in the integration of the FELIX readout system towards its integration for the ITk are presented. These include: the redesign of a compact readout board enabling communication with up to 16 FEI4b ASICs via optical links; chip configuration and scan timing measurements; and improvements to the different configuration scans needed for FEI4b chips.

Acknowledgements

Firstly, I would like to thank the University of Oxford Mathematical, Physical & Life Sciences (MPLS) Division, the CERN Doctoral Programme, Prof. Daniela Bortoletto and Dr. Christoph Rembser and my college Senior Tutor Dr. Joanna Ashbourn for granting me the funding that has enabled my enrolment and participation towards this doctorate.

The submission of this thesis has been made possible by the careful and consistent support of my academic supervisor Prof. Daniela Bortoletto to whom I am incredibly grateful for facilitating every stage of my DPhil. I am also very thankful to her for providing an abundance of discerning feedback and her invaluable advice throughout its course both towards my academic output as well as my career development.

For a duration of over 1000 consecutive days at work, the daily mentorship of my CERN supervisor Dr. Carlos Solans Sanchez has fostered the single most productive and rewarding period of my career thus far. Spanning a multitude of new technical skills, from meticulous programming, to PCB & CAD design, 3D printing and wide-ranging instrumentation from first principles, this has consistently been executed with devotion towards the growth of his students and ultimately, with the inevitable result of cultivating a thriving cross-disciplinary team.

I am also very grateful to both my supervisors for mentoring and encouraging me in the direction of new opportunities, including attending multiple training workshops and conferences.

Through the daily teamwork, the CERN-based ATLAS ITk Pixel group provided a wealth of opportunities for knowledge transfer. This hardware development team, comprising over the course of this DPhil of Heinz Pernegger, Valerio Dao, Enrico Schioppa, Susanne Kuehn, Petra Riedler, Florian Dachs, Lingxin Meng, Sascha Dungs, Ignacio Asensi, Florian Haslbeck, Patrick Freeman, Branislav Ristic, Christian Riegel and Karola Dette provided a vast and rich aggregate experience from which I benefited greatly.

I equally would like to thank the chip designers team who were the source of crucial information throughout this DPhil: Walter Snoeys, Thanushan Kugathasan, Roberto Cardella, Ivan Berdalovic, Leyre Flores, Francesco Piro and Magdalena Munker.

Working in close collaboration with the Oxford-based Pixel CMOS team formed of Richard Plackett, Kirk Arndt, Luigi Vigani, Kaloyan Metodiev and Maria Mironova had a very positive impact and considerably facilitated the advancement of my work.

My colleagues from external institutes whose collaborations ranged from CMOS R&D, FELIX readout, Thermal Cycling as well as rehearsals, I would like to thank Bojan Hiti, Laura Franconi and Joern Grosse-Knetter and Ram Krishna Sharma.

Very significant contributions were made towards the results presented in this thesis by the CERN technicians Valery Akhnazarov, Jan Mladek and Evgueni Danilevich, who

machined, assembled and repaired innumerable components. Crucial assistance towards developments in mechanics were provided by the CERN EP-DT & Gas Group team members Andrea Catinaccio, Danilo Giugni, Diego Alvarez Feito, Neil Dixon, Christophe Bault, Lukasz Zwalinski and Piotr Dziurdzia. The CERN DSF wire-bonding & SMD lab technicians Florentina Manolescu, Ian Mcgill, Alan Honma, Sylvain Kaufmann, Raphael Berberat, Stephane Excoffier and Nicole Wauquier, provided careful yet rapid sample wire-bonding, PCB assembly and round the clock support for any sample preparation needs.

Finally, I am eternally grateful for the consistent and unwavering belief in me of my wife Pooja and the support of our families, that has propelled this endeavour towards its fruition.

Contents

1 Introduction **1**
- 1.1 The Standard Model . . . 1
- 1.2 The Large Hadron Collider . . . 1
  - 1.2.1 ATLAS . . . 2
- 1.3 Upgrades of the accelerator and the ATLAS detector . . . 10

2 Monolithic Active Pixel Sensors **16**
- 2.1 Interaction of particles with matter . . . 16
- 2.2 Silicon and the P-N junction . . . 18
- 2.3 Pixel Detectors . . . 23
- 2.4 Radiation Damage Effects and Irradiation Sites . . . 23
- 2.5 The TowerJazz Designs . . . 27
  - 2.5.1 The Investigator Prototype . . . 27
  - 2.5.2 MALTA . . . 31
  - 2.5.3 mini-MALTA . . . 32
- 2.6 Conclusions . . . 34

3 Sensor Characterisations **36**
- 3.1 Laboratory Measurements . . . 36
  - 3.1.1 Edge Transient Current Technique . . . 36
  - 3.1.2 I-V Measurements . . . 47
  - 3.1.3 Fe-55 Source Measurements . . . 49
  - 3.1.4 MALTA . . . 51
  - 3.1.5 mini-MALTA . . . 53
- 3.2 Testbeam Measurements . . . 56
  - 3.2.1 MALTA . . . 56
  - 3.2.2 MALTA Telescope . . . 59
  - 3.2.3 mini-MALTA . . . 59
- 3.3 Conclusions . . . 65

4 Contributions to the ITk Outer Barrel Demonstrator Programme **66**
- 4.1 Thermal Figure of Merit measurements for Prototype Staves . . . 66
  - 4.1.1 The TFM Concept . . . 67
  - 4.1.2 TFM Setup Description . . . 69
  - 4.1.3 NTC Calibration . . . 70
  - 4.1.4 The VolTRON Temperature Readout System . . . 71

4.1.5 Measurement Method . . . 78
4.1.6 Heat Load Optimisation . . . 79
4.1.7 Performance Results . . . 81
4.1.8 Further IR Analysis . . . 85
4.1.9 Thermal Cycling . . . 86
4.1.10 Conclusions and Outlook . . . 89
4.2 Cable Saver Boards for The Demonstrator Programme . . . 92
4.2.1 The Demonstrator Programme . . . 92
4.2.2 Cable Saver Boards . . . 94
4.2.3 Summary . . . 98
4.3 The FELIX Readout System . . . 99
4.3.1 Upgrades of the ATLAS DAQ System . . . 99
4.3.2 ITk FELIX Integration . . . 101
4.3.3 Outlook and next steps . . . 108

5 Summary and Outlook **110**

Bibliography **115**

Glossary

- AC: Alternating Current
- ADC: Analogue to Digital Converter
- ALICE: A Large Ion Collider Experiment
- ALPIDE: Alice Pixel Detector
- APC: Angled Physical Contact
- API: Application Programme Interface
- ASIC: Application-Specific Integrated Circuit
- ATLAS: A Toroidal LHC AparatuS
- BC: Bunch Crossing
- BCID: Bunch Crossing Identifier
- BER: Bit Error Rate
- CAD: Computer Aided Design
- CERN: European Organisation for Nuclear Research. In french: Conseil Européen pour la Recherche Nucléaire
- CLK: Clock
- CMOS: Complementary Metal-Oxide Semiconductor
- CPU: Central Processing Unit
- CSB: Cable Saver Board
- CSC: Cathode Strip Chambers
- DAC: Digital to Analogue Conversion
- DAQ: Data Acquisition
- DBM: Diamond Beam Monitor
- DCS: Detector Control System

- DESY: Deutsches Elektronen-Synchrotron
- DIN: Data Input
- DMA: Direct Memory Access
- DOUT: Digital Output
- DRS: Domino Ring System
- DSF: Department Silicon Facility
- DUT: Device Under Test
- ELSA: Elektronen-Stretcher Anlage
- ETCT: Edge-Transient Current Technique
- EYETS: End of Year Technical Stop
- FE: Front-End
- FEA: Finite Element Analysis
- FEC: Forward Error Correction
- FELIX: Front-End Link Exchange
- FIFO: First-In First-Out
- FPGA: Field Programmable Gate Array
- FW: Firmware
- FWHM: Full Width Half Max
- GBL: General Broken Lines
- GBT: GigaBit Tranceiver
- GUI: Graphical User Interface
- HEP: High Energy Physics
- HW: Hardware
- IBL: Inner Barrel Layer
- ID: Inner Detector
- IP: Interaction Point
- IR: Infra-Red
- ITk: ATLAS Phase 2 Inner Tracker

- LINAC: Linear Accelerator
- LVDS: Low-Voltage Differential Signalling
- MIP: Minimum Ionising Particle
- MOS: Metal Oxide Semiconductor
- MOSFET: Metal Oxide Semiconductor Field Effect Transistor
- NTC: Negative Temperature Coefficient
- PCB: Printed Circuit Board
- PCI: Peripheral Component Interconnect
- PMOS: Positive channel MOS
- PSB: Proton Synchrotron Booster
- PSPP: Pixel Serial Powering Protection
- RMS: Root Mean Square
- ROD: Readout Driver
- ROS: Readout System
- RPC: Resistive Plate Chamber
- SCT: Silicon Tracker
- SEU: Single Event Upset
- SFP: Small Form-factor Pluggable
- SMA: Sub-Miniature version A
- SPS: Super Proton Synchrotron
- SWROD: Software Readout Driver
- TCAD: Technology Computer Aided Design
- TFM: Thermal Figure of Merit
- TID: Total Ionising Dose
- TIM: Thermal Interface Material
- TOT: Time Over Threshold
- TRT: Transition Radiation Tracker
- VLDB: Versatile Link Demonstrator Board

Chapter 1

Introduction

The ATLAS experiment located at the Large Hadron Collider (LHC) at CERN will undergo a major upgrade known as Phase 2 to exploit the full physics potential of the High Luminosity LHC. These hardware and software upgrades will be necessary to cope with the higher radiation environments and particle hit rates that come with higher luminosity. A description of the ATLAS experiment, its forthcoming upgrade and finally the respective hardware specifications required is provided in this chapter.

1.1 The Standard Model

The Standard Model of particle physics is an accurate framework which predicts the behaviour of fundamental particles and their interactions. Fig. 1.1 shows the list of the fundamental particles in the Standard Model.

These particles can be classified firstly as fermions, following Fermi-Dirac statistics or bosons, exchange particles enabling the four fundamental forces. Fermions can be further subdivided into quarks and leptons. The gauge carriers of the forces are: the photon, the $W^{\pm}$ and Z^{o}, and the gluons for the electromagnetic, weak and strong interaction respectively. The Higgs boson provides a mechanism for particles to acquire mass and was discovered at the LHC in 2012 [2, 3]. It should be noted that the Standard Model does not include a description of gravity. Similarly, it presents shortfalls in the description of the matter-anti-matter asymmetry in the universe and it does not provide a dark matter candidate.

1.2 The Large Hadron Collider

The key beam parameters of the LHC [4] are listed in Table 1.1 while Fig. 1.2 shows the full CERN accelerator complex. Collisions occur at four interaction regions around the LHC and they are recorded by the experiments located at those points: ALICE [5], ATLAS [6], CMS [7] and LHCb [8]. The four experiments can be subdivided in two groups. ATLAS and CMS are multi-purpose detectors which measure particles propagating across a wide range of

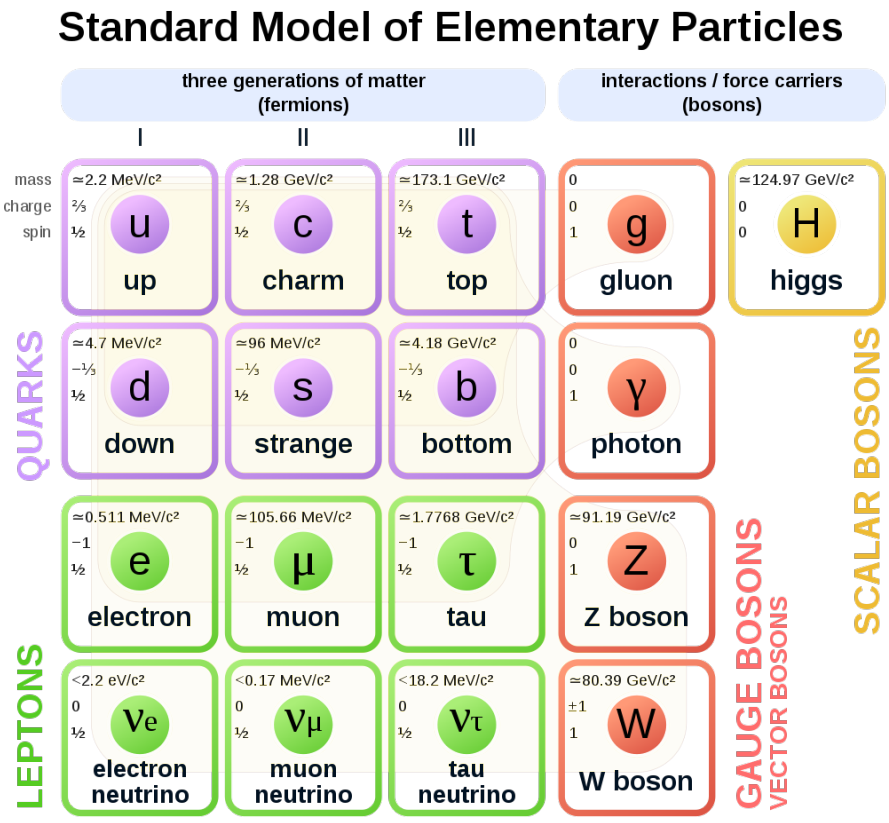


Figure 1.1: The Standard Model of Particle Physics [1].

spatial directions, except along the beam axis. They are designed to search for new physics with high transverse momentum signatures, for example the search for new particles which are predicted by theories such as of supersymmetry and to gather more statistics to study the Higgs mechanism and other Standard Model processes. They are designed to receive more instantaneous luminosity from the LHC than ALICE and LHCb. The total recorded luminosity after Run 2 has been 162 fb^{-1} for CMS and 158 fb^{-1} for ATLAS. ALICE and LHCb have recorded 67 pb^{-1} for ALICE and 6.7 fb^{-1} for LHCb [9].

ALICE focusses on the study of the quark-gluon plasma to understand the origin of confinement, a mechanism that binds quarks and gluons in composite particles, and the generation of the mass of those composite particles.

LHCb is a forward detector which is designed to make precision measurements, especially on b-quarks and their counterpart the anti-b quarks. The study of slight differences in the decay channels between those two particles can help understand the matter-antimatter asymmetry observed in the universe.

1.2.1 ATLAS

The LHC provides a particle collision energy and luminosity that extends the region of discoverable physics far beyond every other preceding experiment. The ATLAS detector

Injection energy	450 GeV
Design beam energy	7000 GeV
Number of particles per bunch	1.15×10^{11}
Number of bunches per fill	2808
Nominal luminosity	10^{34} cm^{-2}s^{-1}
Inelastic cross section	60 mb
Total cross section	100 mb
Revolution frequency	11.245 kHz
Bunch frequency	40.08 MHz
Circumference length	26.66 km
Radius	4.24 km
Number of dipole magnets	1232
Number of quadrupole magnets	392
Nominal magnetic field strength	8.33 T

Table 1.1: Design LHC beam parameters.

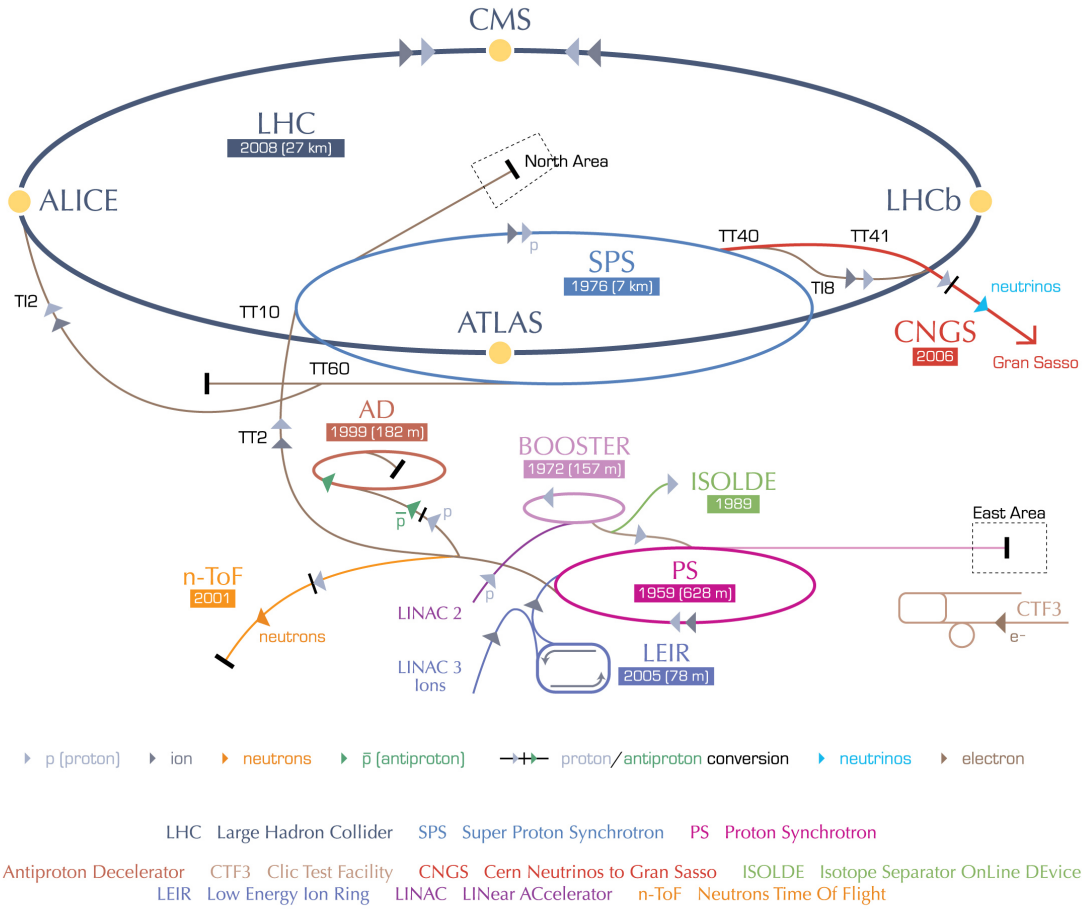


Figure 1.2: The CERN Accelerator Complex [10].

was designed and optimised to maximise the discovery potential for new physics beyond the standard model as well as to perform high precision measurements of established physics. ATLAS is the largest detector at the LHC with a diameter of 25 m, a length of 46 m and a weight of roughly 7000 tonnes [6]. Four different subsystems provide the data for the event

reconstruction. Those subsystems are: The Inner Detector, Electromagnetic Calorimeter, Hadronic Calorimeter and the Muon Chambers. Momemtum measurements are possible due to a large magnetic field permeating through the experiment. This magnet system is divided into a solenoid magnet providing a 2 T field at 2.4 m radii and 3 toroidal magnets. One of these is in the barrel region, made out of 8 superconducting air-core toroidal magnets that give ATLAS its characteristic shape averaging 0.8 T, and two are in the end-cap (forward) regions averaging 1.3 T. These different magnet sections are shown on Fig. 1.4 along with a summary of their key properties on Table 1.2.

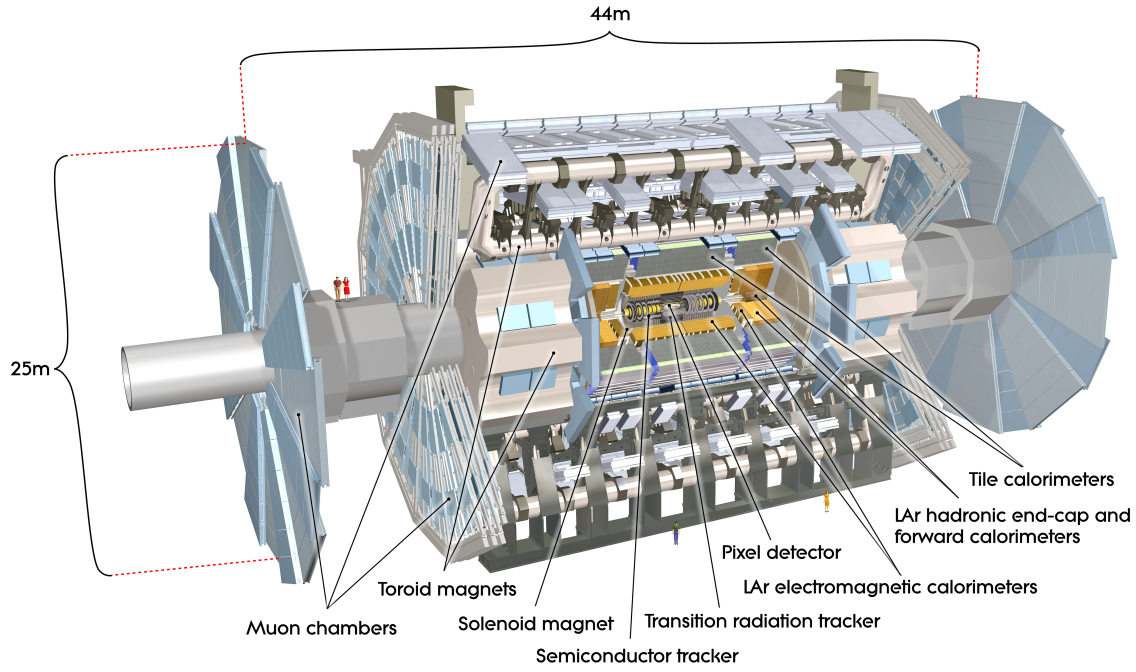


Figure 1.3: The ATLAS Experiment's detectors and subdetectors for Phase 1. [11]

The tracking system provides efficient track reconstruction up to high luminosities to enable the measurement of high transverse momentum leptons, identification of electron, τ-lepton and charged particles. To achieve the best performance the tracking system should cover a large region in pseudorapidity (η) and the full azimuthal angle ϕ. The coordinate system used in ATLAS is a right-handed system using cylindrical coordinates r and ϕ (azimuthal angle around the z-axis) in the transverse plane, its centre point being located at the interaction point in the middle of the detector and the z-axis along the beam pipe. For the x- and y-axis the LHC is taken as frame of reference, where the x-axis points towards the centre of the LHC and the y-axis points from the ring towards the surface. The pseudorapidity is defined as: $\eta = -\ln\left(\tan\left(\frac{\theta}{2}\right)\right)$ with θ being the polar angle with respect to the beam direction. The calorimeter system is optimised for electron and photon identification as well as for precise measurements of jets and missing transverse energy, while the muon system provides high-precision measurements of muon tracks.

A sketch of the detector is shown in Fig. 1.3. The Inner Detector forms a cylinder with a length of 7 m and a radius of 1.15 m and consists of three different systems, each optimised for different requirements due to their distance from the interaction region. Closest to the

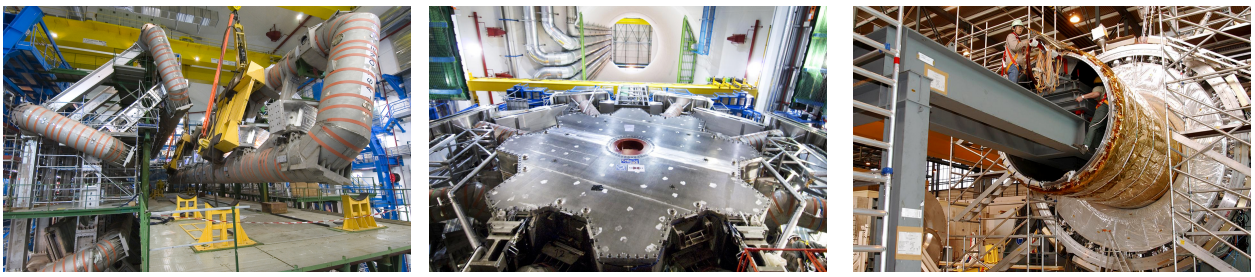


Figure 1.4: The Barrel Toroid, End-Cap Toroid and Central Solenoid Magnets housed inside the ATLAS experiment. [12]

Parameter	Barrel Toroid	End-Cap Toroid	Central Solenoid Magnet
Length [m]	25.3	5.0	5.3
Outer diameter [m]	20.1	10.7	2.4
Coils	8	8	-
Mass [t]	830	240 (each)	5
Superconductor magnetic field [T]	4	4	2
Stored energy [MJ]	-	250 (each)	389
Superconducting wire [km]	100	-	9
Conductor (Al/NbTi/Cu) [km]	56	13	-
Nominal current [kA]	20.5	20.5	7.73
Working point temperature [K]	4.7	4.7	-

Table 1.2: Properties of the different magnet system components housed inside the ATLAS experiment. [13]

interaction point are two silicon-based systems, namely the pixel detector and the SCT detector. The original pixel detector was made out of 4 barrel layers between 50.5 and 120.5 mm in radius, and 3 disks in the end-cap region. The barrel has 1744 modules with 46080 read-out channels each, and a pixel size of 50×400 μm, with a total of 80 million channels, covering an area of 1.7 m^2. The 15 kW power consumption is subject to -25°C by an evaporative cooling plant that uses C_3F_8. An additional pixel barrel layer known as the IBL [14], shown on Fig. 1.5, was added in 2015 at a radius of 30 mm made out of 224 modules with 50×250 μm pixel size. This new layer uses the FEI4 [15] front-end chip that will be used in the demonstrator for the Phase 2 tracker known as the ITk, described in detail in section 4.2.1. The size of the modules for the ITk are based on the 2×2 cm sized modules of the IBL. These sensors are 3D and planar, which will be default designs for ITk.

The SCT is an assembly of 4 barrels placed between 299 and 514 mm from the beam line with a length of 1492 mm and is made of 2112 modules cooled to -7°C as well as 18 disks. It is operated with a high voltage between 350 and 450 V to compensate for inefficiencies due to radiation damage. It consists of silicon strip modules with 80 μm pitch where each module contains two sensor layers rotated with respect to each other by a 40 mrad stereo angle to increase the space resolution. Outside to the SCT a continuous straw-tube detector called TRT is located as the outermost part of the tracking system. It consists of straw tubes with a diameter of 4 mm, holding a thin wire in their centre. The tubes are filled with Xenon gas that gets ionised when traversed by charged particles. The TRT can be used to distinguish between electrons and hadrons by measuring their transition radiation. For

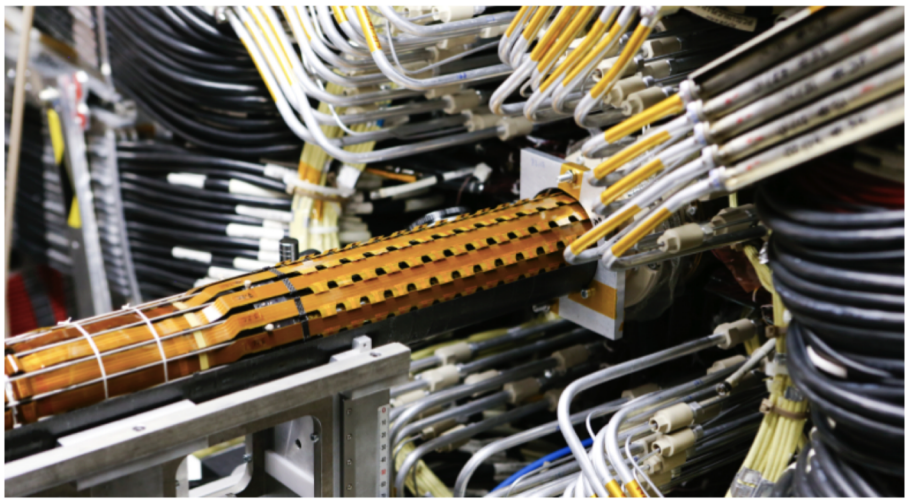


Figure 1.5: The Insertable Barrel Layer (IBL) added to the ATLAS Pixel Tracker in 2015 [14].

instance, electrons would radiate a larger number of photos when traversing the gas than would pions, leading to a greater ionisation of the gas. In turn, a greater charge would be measured on the gold plated wire.

The tracking system is surrounded by a solenoid magnet with a field intensity of 2 T that bends the trajectories of charged particles and provides the measurement of particles' momenta. Schematics of all three subsystems are shown in Fig. 1.6.

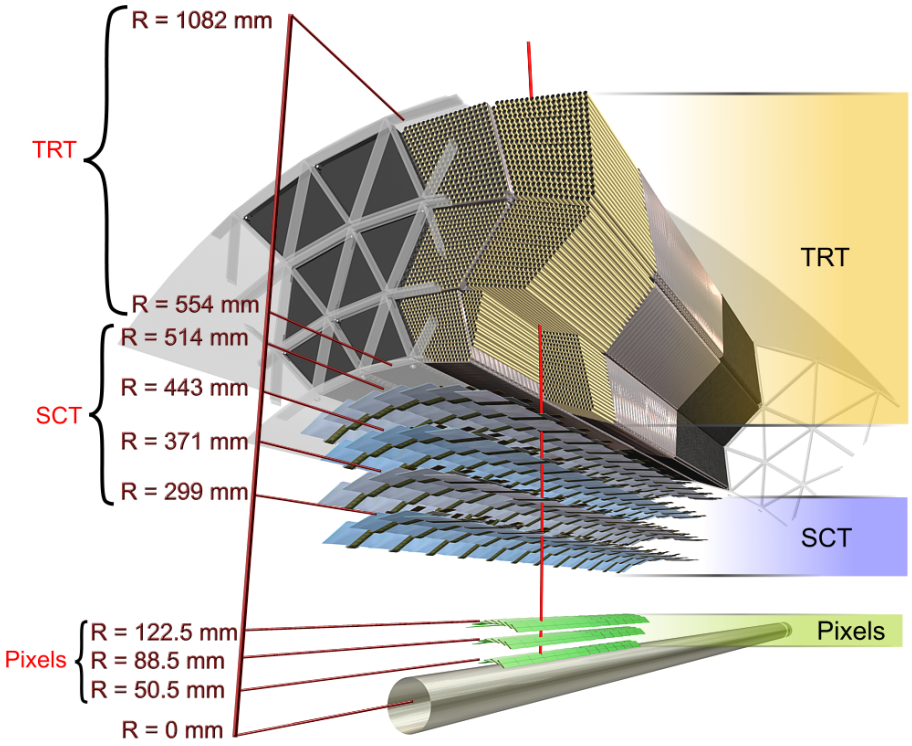


Figure 1.6: The ATLAS Experiment's Tracker sub-detector structure before LS1 [16].

In the electromagnetic and hadronic calorimeter, particles are stopped to measure their energy. For the electromagnetic calorimetry a non-compensated sampling calorimeter is used

which is built out of passive absorbers and active detection material. In ATLAS the absorber material is lead and the detection material is liquid argon (LAr). It is divided into a long barrel of 5.56 m and two extended barrels of 2.91 m at an outer radius of 4.25 m. The calorimeter is contained in a cryostat (at 80 K), is 6.8 m in length with an outer radius of 2.25 m and covers the pseudorapidity range of $|\eta| < 3.2$. The hadronic calorimeter also uses liquid-argon for the forward region ($2.47 < \eta < 4.8$) with copper as passive medium. The central region ($-2.4 < \eta < 2.4$) uses a non-compensated sampling calorimeter made of plastic scintillating tiles as active medium and iron plates as absorber. The resolution of the LAr is described by a=10%/$\sqrt{E}$, b=170 MeV/E and c=0.7%, where a is the stochastic term, b is the electronic noise and c includes to detector effects. To minimise the stochastic term (a=45/$\sqrt{E}$%) the tiles are placed perpendicular to the beam direction, and the detector effects term is c=3% due to a precise Cs-137 calibration system [17]. In general terms the resolution of the calorimeter can be expressed by Eq. 1.1.

$$\frac{\sigma}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c \tag{1.1}$$

Nearly all particles are stopped inside the calorimeter system with two exceptions. Due to their small interaction probability with matter, muons and neutrinos pass through the calorimeter system. While the only possibility for neutrino detection is the calculation of missing energy, muons can be measured in dedicated muon systems. The muon spectrometer is the outermost system of the ATLAS detector and located at an outer radius of about 11 m. A toroidal magnetic field bends the muon trajectories to allow momentum measurements. Four different systems are used for the muon detection: TGC, RPC, MDT and CSC. Fig. 1.7 shows the location of each system.

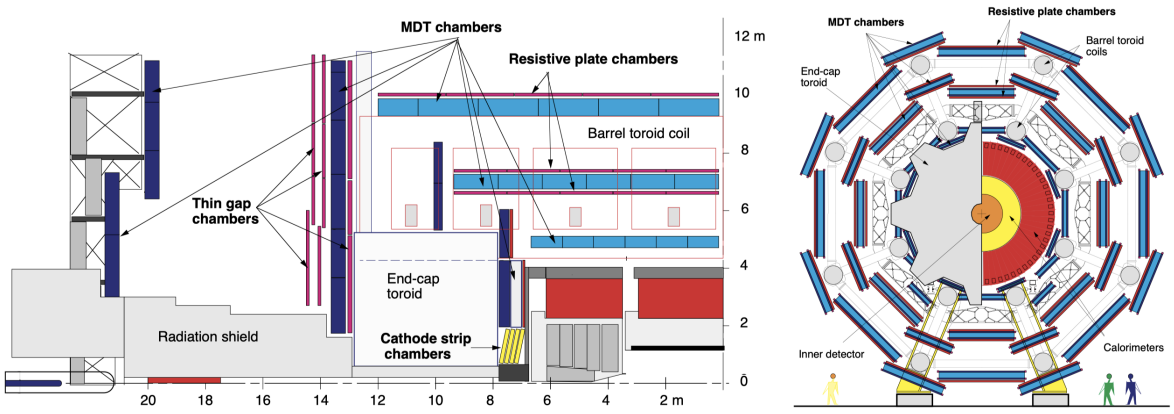


Figure 1.7: Side view of one quadrant of the muon spectrometer (left) and transverse view of the spectrometer (right) [18].

TGC are Thin Gap Chambers that are used in the muon system in the forward region of ATLAS. They consist of a plane of closely spaced wires at high voltage between resistive cathode planes. The TGC delivers signals within 25 ns with an effective rate of 30 to 500 Hz cm^{-2}. The gas used inside the TGC chambers is pressurised n-Pentane with a high voltage range of 1 to 10 kV. The Cathode Strip Chambers (CSC) within the muon system cover the forward region $2.0 < \eta < 2.7$, making use of CF_4 and argon gas. They are housed

on two wheels with 16 sectors within each wheel. Each chamber has 4 precision (η) layers with 192 strips in each layer, with pitch 5.31 and 5.36 mm for big and small chambers, and 4 transverse (ϕ) layers with 48 strips in each. The signal is sampled every 50 ns. The barrel houses, within a psudorapidity of $\pm$ 2, the RPC and MDT, where the RPC (or Resistive Plate Chambers) provide the trigger and the MDT (Monitor Drift Tube) the readout. The gas mixture of the RPC is $C_2H_2F_4$ at 94.7%, C_4H_{10} at 5% and SF_6 at 0.3 %. Chambers in the barrel region are located between and on the eight coils of the superconducting barrel toroid magnet, while the end-cap chambers are in front and behind the two end-cap toroid magnets. The muon chambers consist of 8 octants subdivided into small and large sections arranged into 3 concentric barrels at 7.5 and 10 m radius. These last two systems provide the sector logic whereby coincidence triggers are produced up to a rate of 100 kHz. They have a high voltage operation range of 9.6 kV. Signal formation in the RPCs is smaller than 15 ns. The MDT uses argon and CO_2 with its chambers being built on top of all RPCs. It has a track position resolution of 40 μm and $\Delta p_T = 2$ to 10 % for transverse momenta between 6 GeV and 1 TeV over a pseudo-rapidity range of $\eta < 2.7$.

ATLAS showed an exceptional performance throughout Run 2 where it recorded more than 90% of the delivered luminosity. The evolution of the luminosity over the year and the data recorded by ATLAS is shown in Fig. 1.8 and the average number of collisions per bunch crossing was 30, as shown on Fig. 1.9.

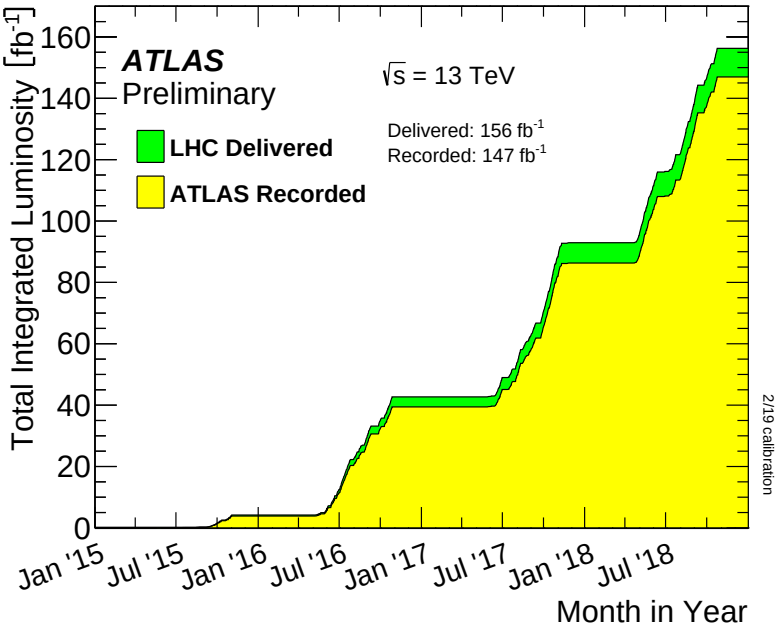


Figure 1.8: Integrated delivered and recorded luminosity by the ATLAS experiment in Run 2. [19]

Many physics searches at the LHC experiments require precise information about hadron-jets containing b-quarks, so called b-jets. Tracking does not only give information about the trajectories of particles produced in collisions and their subsequent decay products but also information about a particle's charge and momentum. Information on impact parameter resolution and distinguishing primary and secondary vertices is used as inputs to so called b-tagging algorithms which are able to identify b-jets. Identifying events that include high

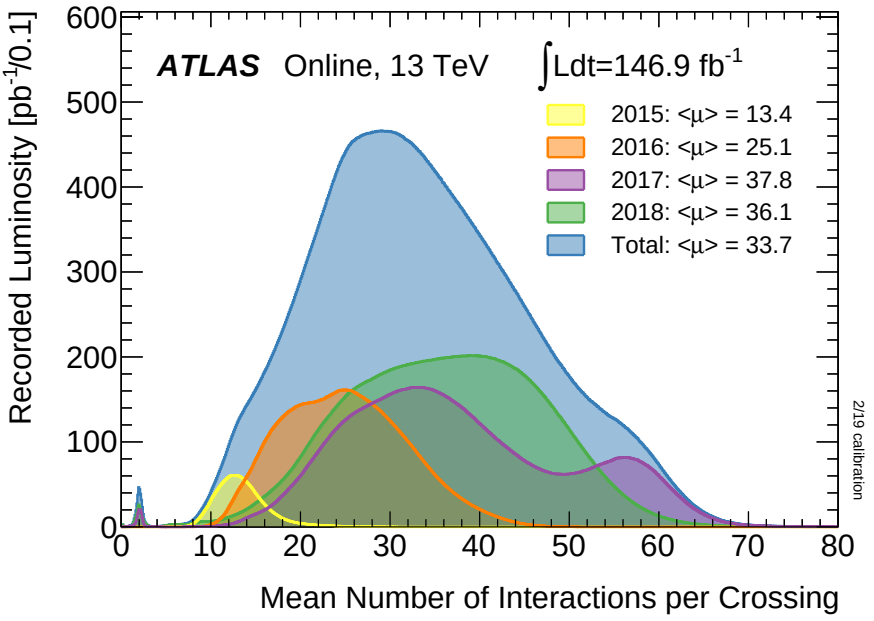


Figure 1.9: Number of interactions per crossing over the different years of operation of the LHC. [6]

p_T b-jets in their final state is essential for analyses that aim for measurements of the top-quark, studies of the Higgs boson decays involving b-quarks and also searches for new physics signatures. An improved tracking performance enhances the quality of the b-tagging and is therefore directly beneficial for those analyses. Eq. 1.2 [20] details the effect of material thickness on tracking momentum resolution where β is v/c of the particle, B the magnetic field, L the distance to the detector perpendicular to the beam axis, X_0 the radiation length of the material and X_{tot} the total thickness of the detector material. A CMOS based pixel detector can be extremely thin and yield an improved momentum resolution.

$$\frac{\Delta p_T}{p_T} \approx \frac{0.0136 \text{ GeV/c}}{0.3\beta B[\text{T}]\text{L}[\text{m}]}\sqrt{\frac{X_{tot}}{X_0 sin\theta}} \tag{1.2}$$

An additional key parameter in detector tracking performance is the impact parameter, known as z_0 and d_0 in the longitudinal and transverse directions respectively. The geometrical definition of these parameters is shown on Fig. 1.10. Optimising the resolutions of the impact parameters play a key role in measuring displaced vertices. Their respective resolutions are described by Eq. 1.3 and 1.4 [20].

$$\Delta d_0 \approx \frac{0.0136 \text{ GeV/c}}{\beta p_T} r_0 \sqrt{\frac{d}{X_0 sin\theta}} \sqrt{1 + \frac{1}{2}\left(\frac{r_0}{L_0}\right) + \frac{N}{4}\left(\frac{r_0}{L_0}\right)^2} \tag{1.3}$$

$$\Delta z_0 \approx \frac{0.0136 \text{ GeV/c}}{\beta p_T} \frac{r_0}{sin\theta} \sqrt{\frac{d}{X_0 sin\theta}} \tag{1.4}$$

where N represents the number of detector layers and the remaining parameters are

defined in Fig 1.10.

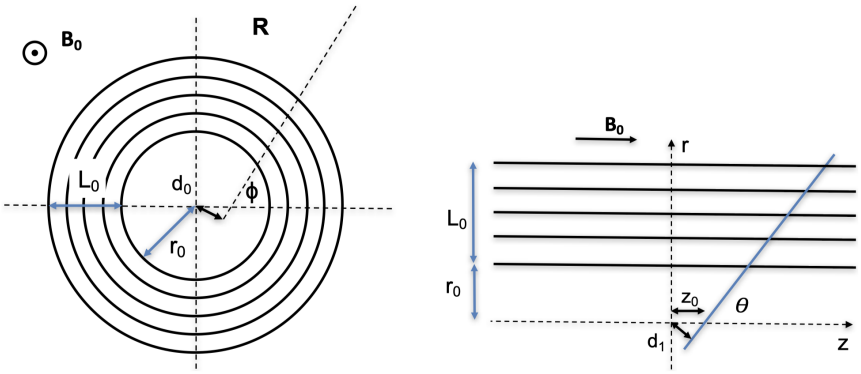


Figure 1.10: Impact parameters d_0 and z_0 in the transverse (left) and longitudinal (right) directions respectively, where ϕ is the azimuthal angle, θ the angle of the particle trajectory in the z-r plane, B_0 the magnetic field, L_0 the distance between the first and last plane of the detector and r_0 the distance between the beam axis and the first layer of the detector. [20]

1.3 Upgrades of the accelerator and the ATLAS detector

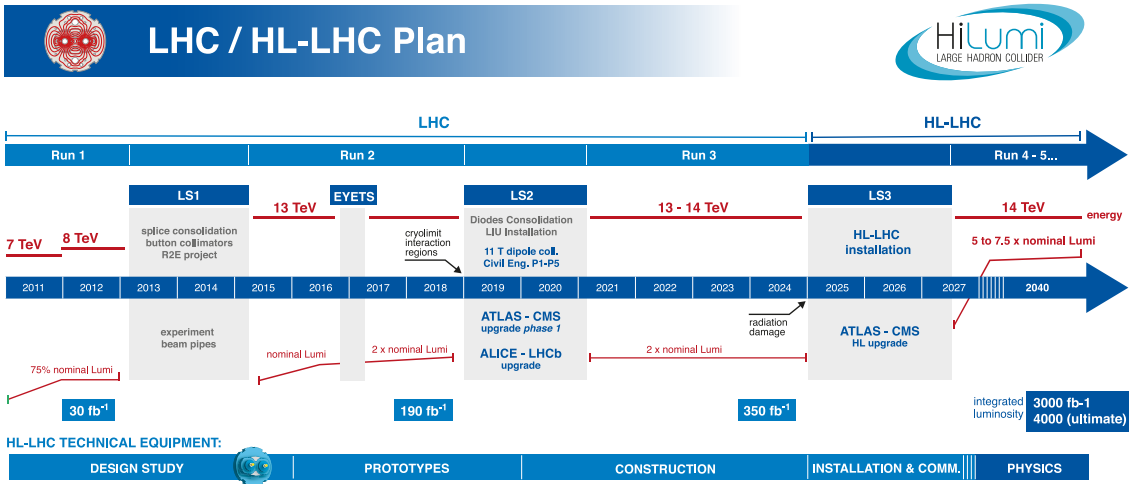


Figure 1.11: Timeline of key development stages for the LHC (and future HL-LHC). [21]

The LHC started operating in 2008, and the experiments started collecting data in what is known as Run 1 that spanned from 2010-2012. Up to 34 pb^{-1} where collected by ATLAS during this period. The timeline of the LHC's various developmental stages of commissioning and operation, including transition towards high luminosity, are summarised on Fig. 1.11 and described in greater detail in this section.

Phase 0: The first long shutdown (LS1) of the LHC occurred from February 2013 to July 2014 to reinforce the LHC's magnet inter-connects. This was necessary to operate the LHC at its design luminosity of 10^{34} cm^{-2}s^{-1} with each beam having an energy of 7 TeV. During

LS1 the addition of the ATLAS pixel detector's IBL (Insertable Barrel Layer) also took place. This shutdown was followed by the Run 2, which lasted from 2015 to 2017 collecting up to 150 fb^{-1} of data at a centre of mass energy of 13 TeV.

EYETS: The Extended Year End Technical Stop in 2016 saw the replacement of the laser drivers by new service quarter panels of the rest of the pixel detector and the installation and commissioning of the DBM (Diamond Beam Monitor).

Phase 1: At the time of writing, the LHC is undergoing a major upgrade, Phase 1, to sustain its luminosity at approximately 2.2×10^{34} $\text{cm}^{-2}\text{s}^{-1}$ during a long shutdown from 2018 to 2020 (LS2), during which period two upgrades are planned. The injector system for the Proton Synchrotron Booster (PSB) will be the LINAC4, replacing the LINAC2. The PSB beam intensity will be doubled enabling a greater LHC luminosity. The luminosity will be further increased through upgrades to the LHC collimation system which will focus the beams at the interaction points. A delivery of 500 fb^{-1} is expected from the LHC over the course of Run 3 [22]. In order to ensure the robust performance of the ATLAS detector during this period of greater luminosity, significant upgrades specifically targeting the ATLAS Trigger system will take place. ATLAS will experience 55 to 80 pile-up events per bunch crossing at an instantaneous luminosity of 1.9×10^{34} $\text{cm}^{-2}\text{s}^{-1}$. Reducing the event rate by altering the threshold on the p_T would result in a loss of physics signal efficiency. In order to avoid this, background events, namely, jets misidentified as electrons, will need to be reduced. This error originates in the calorimeter with fake muons in the spectrometer. An increase in the granularity of the calorimeter readout will therefore also take place. New calorimeter readout boards for the electromagnetic and forward calorimeters are needed to achieve this goal. The Phase 1 FELIX system will be deployed to address these issues. The system is described in detail in Chapter 4, where the preparation towards Phase 2 will be emphasised. The fake muon background is addressed by introducing a new small wheel [23] device and new MicroMegas tracking detector technology in the forward muon spectrometer.

Phase 2: The LHC will undergo yet another upgrade to exploit fully its physics potential over the course of a shutdown after 2022 (LS3), in order to deliver the required luminosity of 3000 fb^{-1}. This upgrade is scheduled to take place from 2024 to 2026 after which the LHC luminosity will be increased to between 5 to 7×10^{34} $\text{cm}^{-2}\text{s}^{-1}$. In order to achieve this operational target, the relevant technologies are undergoing intensive R&D. These will encompass upgrades to the injector chain with crab cavities as well as to niobium and tin (Nb_3Sn) alloy focus quadrupole magnets. The LHC experiences a reduction in instantaneous luminosity over the course of a fill. The maximum instantaneous luminosity is thus above the average luminosity. Luminosity levelling methods, intended to compensate this effect to alleviate the effects on experiments are also being studied and prototyped for Run 3.

The 5 to 7-fold increase in luminosity with respect to the LHC design is expected to lead to around $140 - 200$ instantaneous interactions per bunch crossing. The ensuing higher radiation damage and occupancy are aspects the next generation detector will have to take into consideration. Parts of the present tracker system would likely experience close to 100% occupancy along with radiation induced catastrophic failure in such an environment. The

higher luminosity will also equate to reaching the lifetime of the current pixel and strips tracker systems much sooner. Manifestation of detector ageing will appear in the form of increased depletion voltages, exceeding the operational 600 V high voltage limit, as shown on Fig. 1.12. A comprehensive upgrade of the inner tracking system will thus take place through LS3 in 2024 to achieve improved vertex resolution and track reconstruction performance in ATLAS over the Phase 2 period.

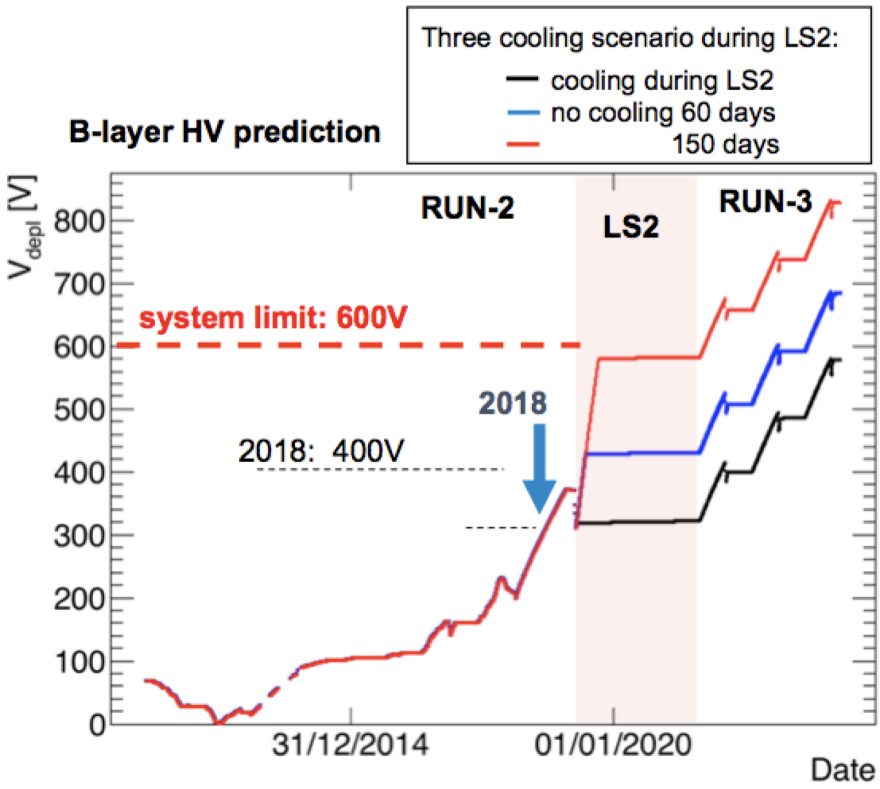


Figure 1.12: Gradual increase of the depletion HV necessary to maintain full depletion of the sensors caused by TID effects, only partially mitigated by cooling the sensors. [21]

The new inner tracker design under consideration will be formed entirely of silicon, comprising of five inner layers forming the pixel detector at close proximity to the beam pipe to ensure improved pattern recognition as well as vertex measurements. Beyond the pixel detector will be four strip layers. The different geometrical layouts being considered are aiming to allow improved tracking performance for pile-up events of up to 200. Subsequently, an upgrade of the readout system will also occur to account for the much greater data rates expected in conjunction with the higher occupancies.

3000 fb^{-1} is expected to be collected during HL-LHC running which, for the pixel layer closest to the interaction point, will correspond to a total ionising dose (TID) of 7.7 MGy as well as a 1 MeV neutron equivalent fluence of 1.4×10^{16} cm^{-2}. This will be significantly greater than the levels the current detector can withstand. On the opposite end, the outermost pixel barrel layer will be exposed to a TID of 0.9 MGy and a 1 MeV neutron equivalent fluence of 1.7×10^{15} cm^{-2}. As a result of the order of magnitude reduction in radiation environment from the inner to outer layers, the technologies considered for the different layers target very different designs. Fig. 1.13 shows the expected high radiation environment the

extended ITk region would be exposed to during the HL-LHC phase. The expected fluence at different z locations don't always decay along r in an $1/r^2$ fashion but instead display minor variations. These can be attributed to neutrons produced from nuclear collisions in the inner layers.

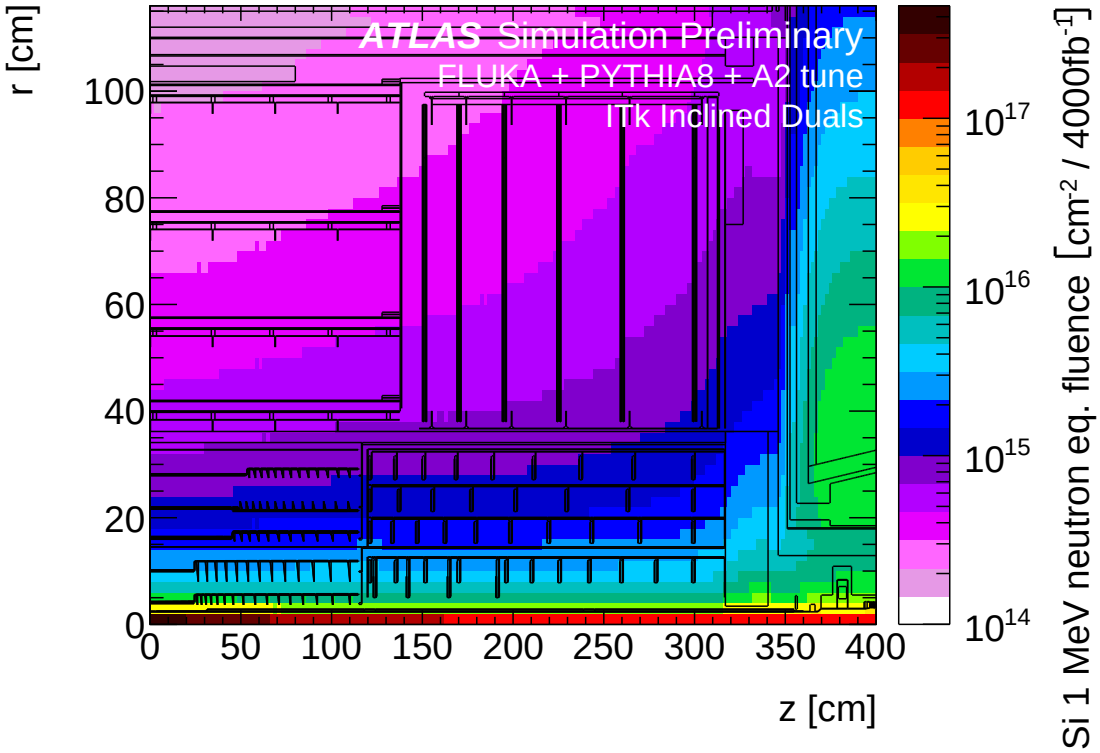


Figure 1.13: Radiation environment within the extended ITk region. [24]

The ITk pixel detector readout electronics is designed to provide adequate bandwidth for a 1 MHz new Level-0 trigger, with the option to upgrade to 4 MHz for the track based trigger system. The trigger algorithm exploits regions-of-interest to estimate the p_T of the tracks, which is then used in the Level-1 trigger algorithm following the new proposed trigger rate, with the only restriction of a minimum bandwidth. Resolving multiple pile-up vertices in the high luminosity environment and assigning the high p_T jets, tracks and secondary vertices to the vertex of particular interest is the main task of the new pixel detector. In turn, this requirement dictates the granularity of the detector at different radii. Commensurate to the above requirements, a pixel detector layout is proposed, by splitting the pixel detector in two parts. Each of these parts uses distinctive technologies known as planar or 3D sensors. For the inner-most layers 3D sensors have been shown to provide an exceptionally high radiation tolerance due to their stronger electric fields across a shorter distance coupled with their reduced carrier drift distances as shown on Fig. 1.14. For a detailed overview of 3D sensors, the reader is referred to [25].

Planar sensors will be used for equipping the outer layers, as shown in the proposed layout for the ITk in Fig. 1.15. The process that forms the hybrid pixel detector exploits a strategy involving the sensing element and the readout to be designed in two distinct silicon parts. These are then connected to each other through a process called bump bonding and a

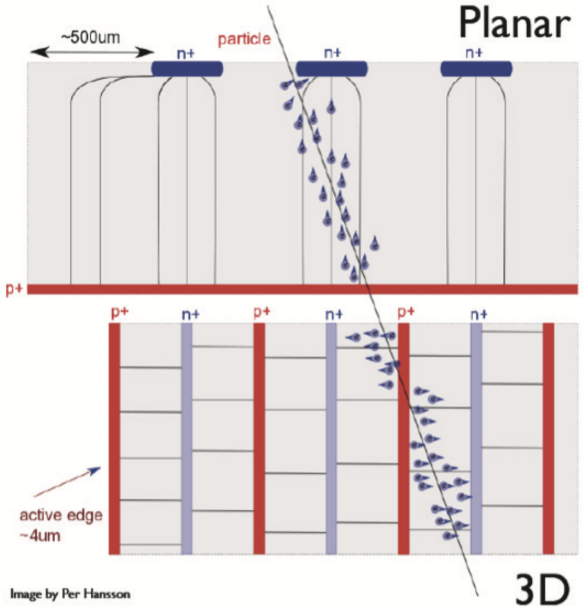


Figure 1.14: Schematic of a planar (top) and 3D (bottom) sensor design geometry. [26]

hybrid pixel detector is thus formed. This is a very expensive process, in addition reducing the granularity of the sensor because of the required added gaps for these bump pads as well as eventual cumulative thickness.

Another approach offering several benefits exploits industrial CMOS processes with the option to deplete the bulk material for the sensor production. One possibility is to have an initial amplification stage in a simpler CMOS which is then read out by a more intricate chip.

The monolithic sensor approach, detailed in Chapter 2 is a second, interesting possibility. Here both the amplification and the additional readout logic are realised on the sensor itself. This approach offers the possibility of exploiting industrial grade silicon manufacture for LHC and future HEP experiments.

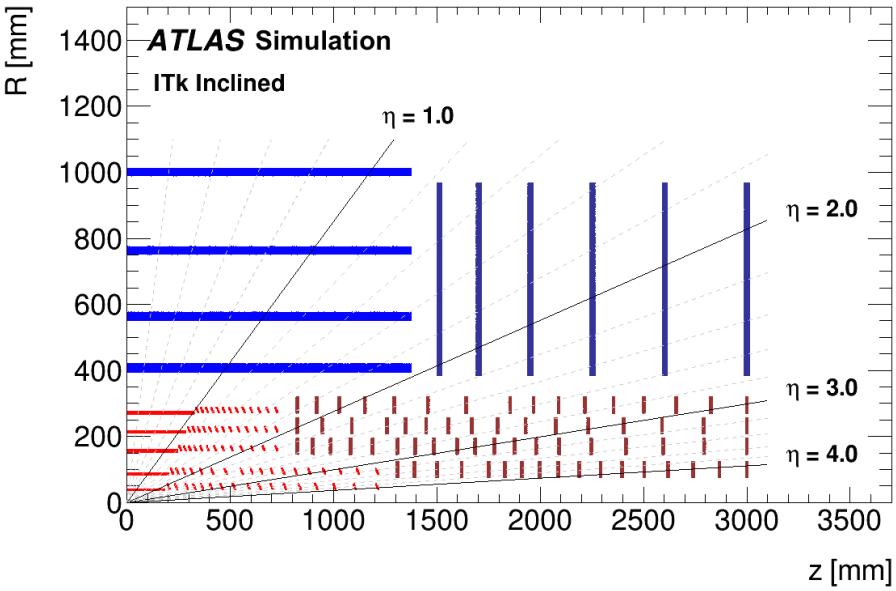


Figure 1.15: Proposed layout for the ATLAS Phase 2 Inner Tracker showing tilted sensor arrangement with strip sensors in blue and pixel sensors in red. [27]

Chapter 2

Monolithic Active Pixel Sensors

Monolithic Active Pixel Sensors (or MAPS) are an attractive technology for HEP tracking environments due to their highly integrated circuitry and simplified fabrication. This chapter outlines the basic building blocks that comprise such a sensor, starting with the basic P-N junction and finishing with a comparison of MAPS and hybrid sensors. For an exhaustive overview of these topics the reader is referred to [25].

2.1 Interaction of particles with matter

The detection of a particle by any detector inherently requires some form of interaction between that detector and the incident particle. Detectors rely on the energy loss experienced by a particle traversing the sensitive region of a sensor to generate a signal. For charged particles, the prevailing interaction mechanisms involved are inelastic collisions with atomic electrons leading to energy loss through Coulomb interactions, and elastic collisions with the nuclei leading to damage of the crystal lattice (see section 2.4). Particle interaction with matter is approximated by the Bethe-Bloch formula shown in Eq.2.1.

$$-\frac{dE}{dx} = 4\pi N_A r_e^2 m_e c^2 z^2 \frac{Z}{A}\frac{1}{\beta^2} - \left[\frac{1}{2}\ln\left(\frac{2m_e c^2\beta^2\gamma^2 T_{max}}{I^2}\right) - \beta^2 - \frac{\delta(\gamma)}{2}\right] \tag{2.1}$$

where,

- r_e: classical electron orbital radius (2.817×10^{-13} cm)
- m_e: electron rest mass (9.11×10^{-31} kg)
- c: speed of light (2.998×10^{8} ms^{-1})
- N_A: Avogadro's number (6.022×10^{23} mol^{-1})
- Z: atomic number of absorbing material
- z: charge of incident particle in units of e

- A: atomic weight of absorbing material
- β: v/c of the incident particle
- γ: Lorentz factor ($\gamma = \frac{1}{\sqrt{1-v^2/c^2}}$)
- T_{max}: maximum transferable kinetic energy from incident particle to atomic electrons, in units of eV.
- I: mean excitation energy of the material, 173 eV for silicon.

The approximated energy loss as observed by a muon travelling across copper is shown in Fig. 2.1.

When considering the energy loss of particles, a specific energy regime is of particular interest, namely that close to the minimal ionisation energy producible by an incident particle through the sensor material. This corresponds to the range $1 < \beta\gamma < 10$ on Fig. 2.1.

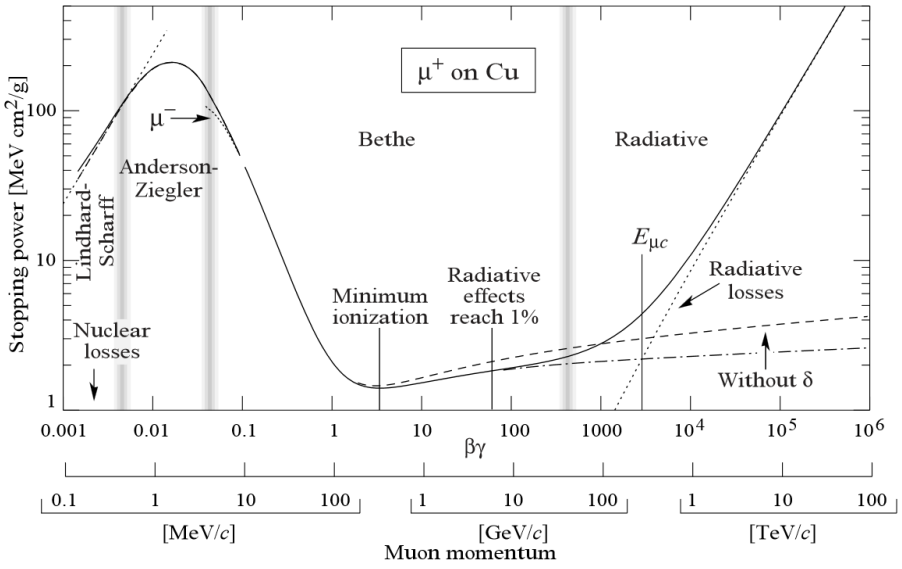


Figure 2.1: Average energy loss approximated by the Bethe-Bloch formula for a muon across a copper medium. [28]

Such particles are referred to as Minimum Ionising Particles (MIPs) and their importance is related to the fact that they lose a minimal amount of kinetic energy as they traverse the silicon sensor. The energy deposited by particles traversing the silicon layer is probabilistic in nature and is best represented by a Landau distribution, shown on Fig. 2.2 for different silicon thicknesses. The largest value of this distribution is referred to as the Most Probable Value in terms of deposited energy. The rarer higher energy deposition events, generating what are known as δ electrons, lead to the distinct long tail in this distribution.

When considering photon interactions with matter, three main mechanisms should be considered:

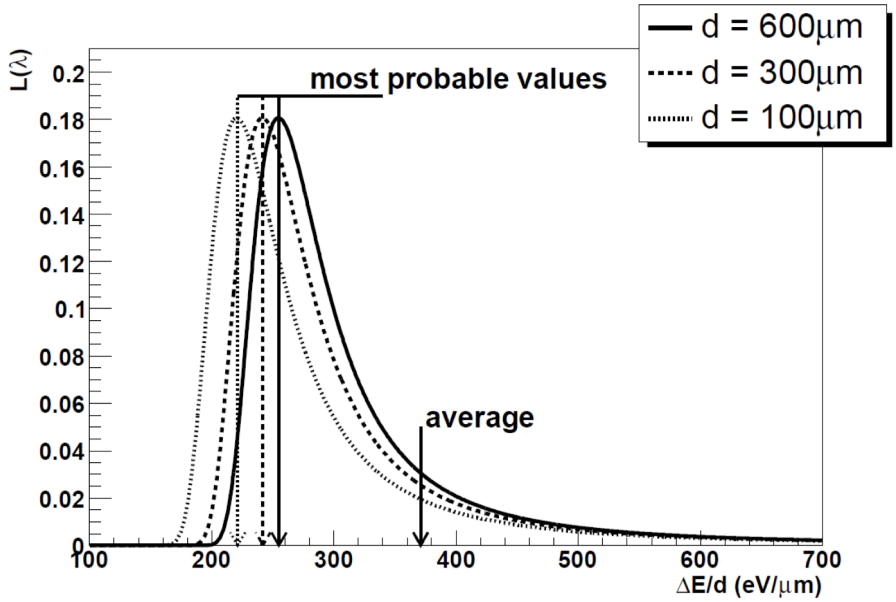


Figure 2.2: MIP energy loss distribution in silicon for various thickness d, normalised to d, taken from [29]

- **Photoelectric Effect**:
  This occurs when the energy of the incident photon is just above the ionisation energy of an atom, following which an electron is expelled from that parent atom.

- **Compton Scattering**:
  When the incident photon has an energy much larger than the atom's ionisation energy, an electron can be ejected taking energy from the scattered photon, leading to a increase in its wavelength.

- **Pair Production**:
  In cases where the energy of the incident photon is superior to twice the rest mass of an electron, the production of an electron-positron pair can occur.

For the study detailed in section 3.1.1, in the case where infra-red light is shone on silicon, the primary mode of interaction of light is through photoelectric effect. As the beam of photons is shone on the material and penetrates deeper through it, the energy of the photons remains the same but the beam intensity reduces exponentially, $I(x) = I_0(x) \cdot e^{-\mu x}$, where x is the distance traversed within the material and μ is its absorption coefficient.

2.2 Silicon and the P-N junction

Silicon is a very cost-effective, highly studied material with well measured physical properties and widely used in the electronics industry. It is particularly interesting for applications in particle physics trackers due to its relatively low electron-hole pair creation energy of 3.6 eV, compared to gas ionisation chambers requiring upwards of 20 eV for ionisation to occur. This

low electron-hole pair creation energy provides the possibility for incremental pair production per small energy deposition steps, hence providing greater energy resolution. In addition, its high density reduces the scattered secondary particles travelling substantial perpendicular distance along the sensor, as a result, providing excellent position resolution. Its high density also yields a substantial energy loss of around 300 eV/μm, which translates to around 80 *e-h* pairs created for every micron of silicon traversed.

In terms of electrical behaviour, materials can be categorised as conductors, insulators or semi-conductors, depending on their level of intrinsic conductivity. These differences in conductive behaviour are governed by the electronic structures of elements at their atomic level. The band-gap theory can be used to understand these variations in conductivity for different elements. As shown on Fig. 2.3, the band-gap theory describes two distinct regions of discrete electronic states where electrons can reside within an atom: the valence band and the conduction band (where electrons are free to move to neighbouring atoms). The energy difference between these two levels is known as the band gap energy. Metals have no band gap, allowing their electrons to seamlessly move between atoms within the metallic lattice. Insulators on the other hand have a significant band gap, requiring energy superior to that available through thermal excitation to promote an electron to the conduction band. Finally, semi-conductors have an intermediate level band gap, the Fermi-level which can be altered to vary their conductivity, as needed, through a process called doping.

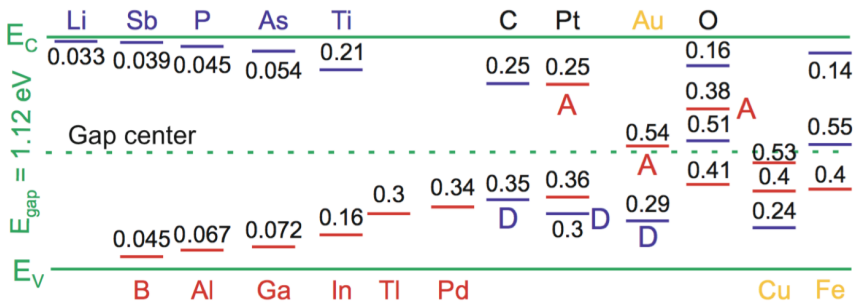


Figure 2.3: Band gap energy in eV for different elements. [30]

At room temperature and pressure, silicon forms a regular diamond lattice. Being a group IV element, it can be doped with group III or V elements to alter its dielectric properties such as resistivity (inversely proportional to conductivity). Doping involves implanting a given number of silicon atoms with the desired dopant atoms. Boron is a group III element, having three valence electrons and hence known as an acceptor. It is commonly used as a dopant to produce an electron vacancy or hole in the silicon lattice. This leads to the boron doped region of the silicon material as being *p-type*, with holes being the majority carrier. Similarly, Phosphorous is commonly used as a group V element, having five valence electrons and thus acts as a donor. It provides an additional free electron and thus will produce an *n-type* region in the silicon, with the electron being the majority carrier. In both cases, the carriers diffuse freely through the silicon lattice. The silicon lattice-level electronic configurations of different doping types is shown on Fig. 2.4.

In addition, the probability of occupation of individual energy states can be described by the temperature dependent Fermi function:

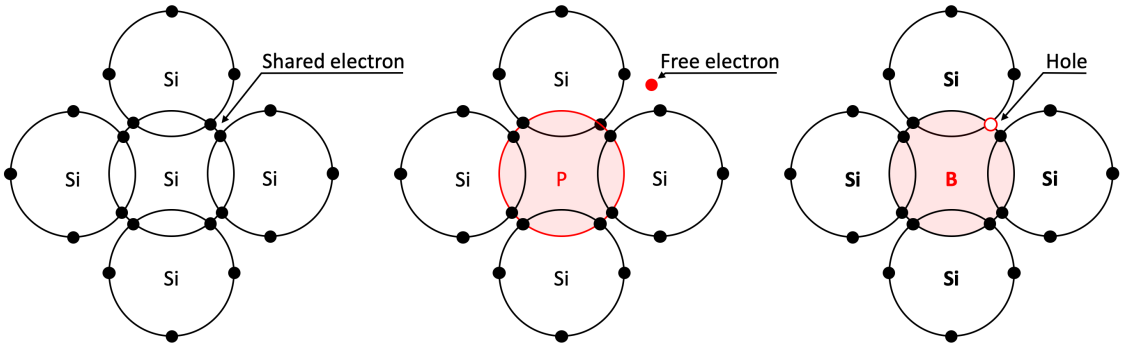


Figure 2.4: Electronic structure of intrinsic (left), n-doped (centre) and p-doped (right) silicon lattices.

$$f(E,T) = \frac{1}{e^{(E-E_f)/k_B T} + 1} \tag{2.2}$$

Here, k_B is the Boltzmann constant and E_f is the Fermi level, typically close to midway between the valence and conduction bands. For undoped silicon, the number of free charge carriers, n, can then be determined by the product of the density of states in the conduction band with the probability of the occupation of those states as evaluated by the Fermi function:

$$n = \int_0^\infty N(E) f(E,T) dE \tag{2.3}$$

If doped silicon is being considered, then the number of free charge carriers will be exactly equivalent to the concentration of the dopant element.

The ability of dopants to affect the resistivity of an intrinsic semiconductor through the release of free charge carriers and subsequently affecting the band-gap energy is a crucial tool in the design of custom silicon tracking detectors as it controls the electric field gradient across the cross-section of a silicon sample. The resistivity, ρ, is dependent on the mobility of each charge carrier, which is in turn dependent on the electric field applied across the silicon sample.

$$\rho = \frac{1}{c} = \frac{1}{e(\mu_e n + \mu_h p)} \approx \frac{1}{e(\mu_d N_d)} \tag{2.4}$$

Here c is conductivity, μ_e and μ_h are the mobilities and n and p the densities, respectively of electrons and holes. In the event of a much greater number of charge carriers of a particular type, this equation can be simplified as shown in Eq. 2.4 where μ_d and N_d are the mobility and density of the majority dopant.

If a *p-* and *n-type* silicon material are placed in contact, the respective mobile carriers become attracted their oppositely charged counterparts. This leads to the movement of these electrons and holes, as shown in Fig. 2.5, at which point some at the interface recombine causing a depletion region to form.

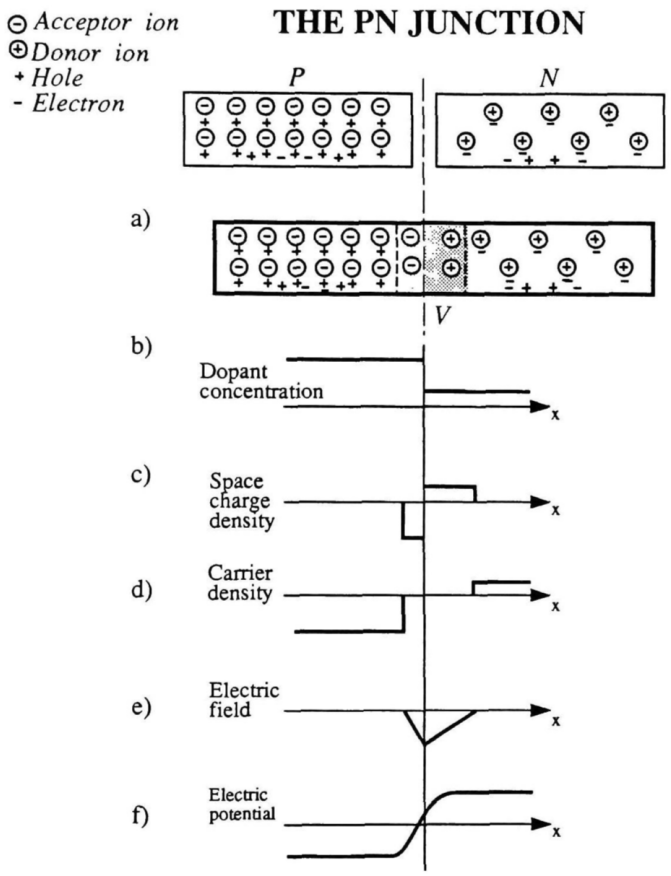


Figure 2.5: P-N junction before and after physical contact, illustrating the different electrical characteristics of such a configuration. [31]

This diffusion of the electrons onto the holes in the neighbouring silicon volume, or recombination, continues until an equilibrium is reached. This forms a space-charge region also known as a depletion region. The electric field across this volume is known as the built-in voltage (V_{bi}), described by:

$$V_{bi} = \frac{kT}{e} ln \left(\frac{N_d N_a}{n_i^2} \right) \tag{2.5}$$

Here, as before, N_d and N_a are the respective densities of donors and acceptors and n_i^2 is the product of the number of free carriers of each type, n_a and n_d for acceptors and donors respectively. In the case of intrinsic silicon these are equal, as follows: $n_a = n_d$ hence $n_a \cdot n_d = n_i^2$. V_{bi} values can typically reach up to 1 V in the context of the CMOS designs described within this thesis.

Since immobile ions of opposing charge are present on either side of the depletion region, this volume acts much like a charged capacitor of value C per unit area:

$$C = \epsilon_0 \epsilon_{si} / W = \sqrt{e \epsilon_0 \epsilon_{si} N_{eff} / 2V} \tag{2.6}$$

with the voltage V being the sum of built-in voltage and any additional external biasing

applied, N_{eff} the effective dopant density and ϵ_{si} the permittivity in silicon.

"Forward biasing" means applying a voltage where the positive node is connected to the p-region and negative node to the n-region. The positive node repels the holes and the negative node repels the electrons, causing them to be pushed towards the small intrinsic depletion region and as a result, reducing its width. "Backward biasing" has the opposite effect and leads to increased depletion layer width. The growth of the depletion layer, and in other words, the removal of free carriers in the depleted region forms the basis of a silicon tracking detector.

In the event of energy deposition within the depleted region, electrons and holes produced drift to their respective electrodes and induce a potential change. The shape of this signal provides critical information on the energy deposition and its position in the depleted volume. It is desirable to ensure that the silicon bulk is fully depleted and devoid of free charge carriers as energy deposition within under-depleted volumes would increase the likelihood of recombination of the e-h pairs produced before.

The depth of the depleted volume can be provided as a function of the donor and acceptor densities N_a and N_d respectively, along with the built-in voltage V_{bi} already present across the P-N junction, and finally the material's permittivity, ϵ, as shown below:

$$d = \sqrt{(V + V_{bi})\frac{2\epsilon}{e}\left(\frac{N_A + N_D}{N_A N_D}\right)} \approx \sqrt{V\frac{2\epsilon}{eN_{eff}}} \tag{2.7}$$

In most practical cases, one type of dopant (such as acceptors) strongly outnumbers the other, in which case, Eq. 2.7 can be simplified as shown, where N_{eff} is known as the effective dopant density and is equivalent to $N_A - N_D$. Finally, rearranging this equation isolating voltage and setting the depth to the physical thickness of the silicon bulk, provides what is known as the depletion voltage, or the potential necessary to fully deplete the thickness of the bulk (d_t), as described by Eq. 2.8. For a silicon detector's sensitive region to be an effective charge collecting volume, achieving a full depletion of the bulk is preferred.

$$V_d = \frac{d_t^2}{2\epsilon\mu\rho} \tag{2.8}$$

2.3 Pixel Detectors

A silicon sensor is at its very fundamental level a reversely biased p-n junction. As elaborated in section 2.2 such a biased junction, say, a diode, would already comprise of a particle sensing element, namely its depleted volume. In order to transition from a simple diode to a sensor for detector applications, spatial coordinates of the particle trajectory need to be recorded. In order to achieve this, historically, silicon sensors were designed in what are known as strip geometries. As shown on Fig. 2.6, these provide only a 1-dimensional spatial reference point when a particle deposits energy.

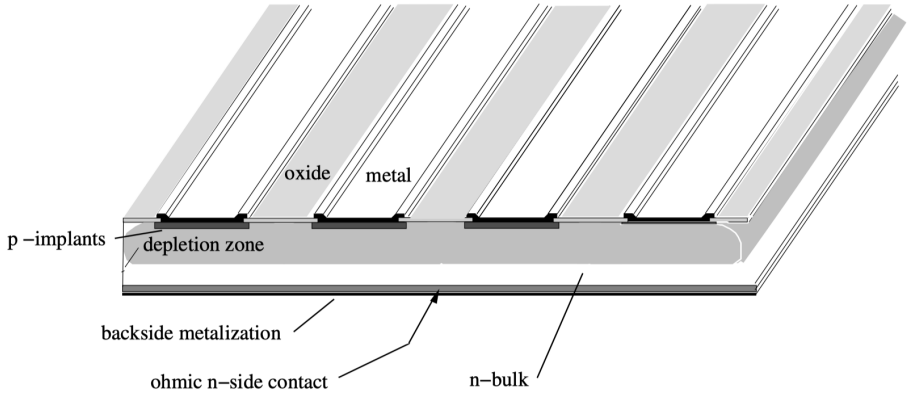


Figure 2.6: Example geometry of a strip detector [25].

Using a pair of such strip detectors with a stereo angle (as described in section 1.2.1), the location at which a particle traverses the detector is measured. Higher sensor granularity can be achieved if the sensing element is subdivided into pixels, with individual sensing elements as small as 20 μm. Fig. 2.7 shows the smallest readout chain used to form a pixel sensor, from its sensing element to its readout electronics via a circular soldering connections known as bump bonds.

The pixellation is necessary to provide a spatial measurement of the trajectory of an traversing particle. As a result, the structure of pixel sensors can involve a depleted region, capable of producing a significant amount of e-h pairs that induce a signal in the electrodes (sometimes achieved instead with the implementation of a float-zone based design [25], and a thin enough substrate layer, to reduce the dead material in the experiment). Pixel detectors are built as a composition of sensor and read-out chip (hybrid), or as a single device on silicon, implementing the read-out circuitry in the sensor design (monolithic). The purpose of this thesis is to characterise the recent developments in monolithic pixel detectors.

2.4 Radiation Damage Effects and Irradiation Sites

Massive particles can displace a silicon atom from the lattice bulk resulting in a lattice vacancy and a hole in the silicon also called a Frenkel pair. If the **p**rimary **k**nock-on **a**tom (PKA) has enough energy it will create ionisation or further atom displacement. An agglomeration of this type of damage is referred to as a cluster. The typical range of a PKA

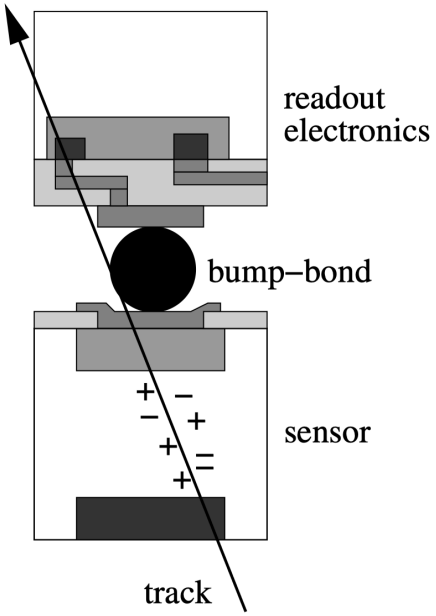


Figure 2.7: Full vertical assembly of a hybrid pixel sensor [25].

is 100 nm. The kinetic energy needed to create a Frenkel pair or cluster by neutral massive and charged particles is listed in Table 2.1.

	Frenkel Pair	Cluster
Neutrons	185 eV	35 keV
Electrons	255 eV	8 MeV

Table 2.1: Kinetic energy required for Frenkel pair and cluster production by neutrons or electrons.

Most important mechanisms that induce a defect on the silicon are:

- **Generation**:
  Defects with energy levels close to the middle of the band gap act as generation centres facilitating the creation of *e-h* pairs.
- **Recombination**:
  Simultaneous capture of an *e-h* pair leading to reduction of free carriers.
- **Compensation**:
  Process through which doping levels are affected by acceptors or donors.
- **Trapping**:
  Process in which lattice defects, known as traps, capture a charge carrier. This is particularly harmful when the time of release of the charge is long as it can compromise the ability of a detector to collect charge within a specified time window, known as in-time charge collection efficiency.

The average number of electrons in a single state can be calculated with Fermi-Dirac statistics in thermal equilibrium. The Shockley-Read-Hall model [32] in turn allows the

calculation of the statistics of recombination of holes and electrons and their lifetimes wherein recombination occurs through the mechanism of trapping.

The effects at macroscopic level are:

- Change in the reverse current: It is characterised by $\Delta I(\Phi) = \alpha\ \Phi\ V$, where α is the current damage parameter (typically 4×10^{-17} A/cm [33]), Φ the equivalent fluence and V the volume of the semiconductor [33].
- Depletion Voltage: Ionised donors create positive charge and acceptors create negative charge, affecting the depletion voltage.
- Type Inversion: In the case of, say, an n-type silicon sensor, irradiation results in the occurrence of negative space charge which nullifies the initial positive space charge of the n-type silicon. With increasing particle fluence, the net space charge decreases to near pure silicon levels. This level along the irradiation is when the space charge sign changes from positive to negative [34].
- Reduction in charge collection efficiency: caused by the macroscopic effects of trapping [35]. $\tau_{eff} = \frac{1}{\beta\ \phi}$, where τ is the recombination time, β is the damage parameter with units of cm^2/s and ϕ is the effective fluence.

The mechanisms to recover from radiation (annealing) are the following:

- **Migration**:
  This is a temperature dependent effect whereby the defects move through the lattice until they become trapped.
- **Complex Formation**:
  The case where a defect reacts with other impurities in the lattice forming a complex defect structure.
- **Dissociation**:
  The simplification of complex defects into individual damage effects.

In order to assess the radiation hardness tolerances of the different sensor designs, the following irradiation facilities were made use of to expose the sensors to predefined fluences in order to simulate the life-cycle exposure that they would experience:

- Neutron irradiations:
  - TRIGA nuclear reactor in Ljubljana [36] providing 5 MeV neutrons with a nuclear flux of up to 10^{12} $\mathrm{n/cm}^2/\mathrm{s}$ [37].
- Proton irradiations:
  - Proton-Synchrotron Facility at CERN delivering 25 GeV/c protons.

  - Birmingham MC40 Cyclotron Facility [38] 23 MeV/c protons.

  The ability to distinctly probe the effects of low and high energy protons provides benefits in silicon damage studies. This is due to the fact that low energy protons, typically interact through elastic collisions, causing ionisations of the silicon lattice, which in turn can lead to Single Event Upset (SEU) events. Such events can be the cause of bit flips and as a result, impact the Bit Error Rate (BER). High energy protons on the other hand can lead to nuclear reactions and therefore bulk damage.

- X-ray irradiations:

  - X-RAD iR160 biological X-ray machine at CERN, producing a maximum energy of 160 keV.
  - Custom built X-ray irradiation chamber at the University of Glasgow, with 50 keV.

  Both facilities provide a large volume in which to house the sample allowing irradiation rate adjustments to be made by placing the sample further from the X-ray tube.

2.5 The TowerJazz Designs

As part of the Phase 1 upgrade of the LHC (2019-2020), the ALICE experiment developed an Inner Tracking System (ITS) comprising of 7 concentric layers of high resolution silicon pixel sensors through collaboration with the TowerJazz [39] silicon foundry, who use 180 nm CMOS technology. Fig. 2.8 (left) shows the cross section of their sensor design. This design was incorporated into the R&D program of the ATLAS ITk. In order to cope with the higher radiation levels in ATLAS and to ensure full depletion of the sensor, an additional step was introduced in the manufacturing process, namely, a horizontal n-implant which is shown in Fig. 2.8 (right). This would allow the uniform vertical growth of the depletion layer and facilitate a full depletion of the epitaxial layer and avoid any undepleted sections remaining in the pixel edges.

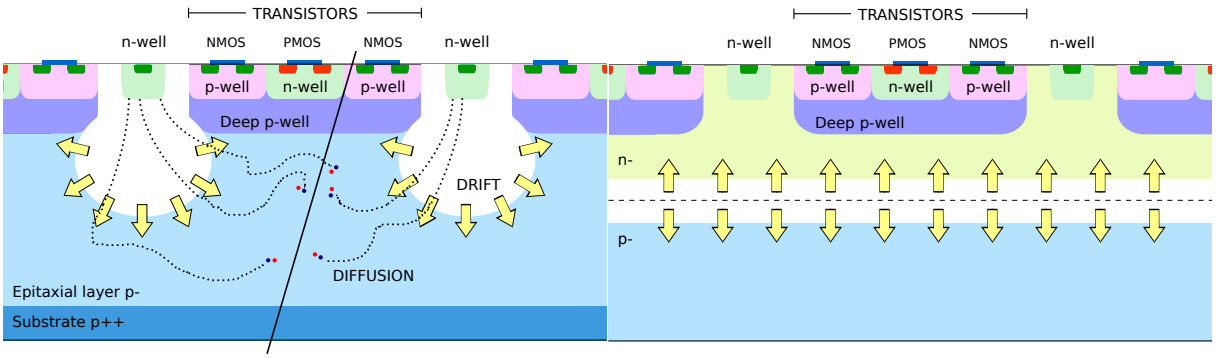


Figure 2.8: Wafer cross sections of the TowerJazz pixel design from the ALICE ITS R&D (left) and ATLAS ITk design (right). The arrows show the direction of growth of the depletion volume for both designs.

2.5.1 The Investigator Prototype

The Investigator I chip, also inherited from the ALICE ITS programme, is a fully analogue device, hence allowing for the direct study of the pulse formed during energy deposition, without it being altered by any amplification or other processing. It is designed to characterise pixel geometries varying in size from 20 μm^2 to 50 μm^2 and to study their effects on charge collection efficiencies as well as charge sharing. Within the ATLAS ITk R&D, two additional versions of the Investigator chip were designed, named: Investigator IBis and Investigator II. The Investigator I consists of 134 structures known as mini-matrices, each in turn comprising of 10 by 10 individual pixels. The outermost pixels in each mini-matrix are referred to as dummy pixels since they cannot be readout even when they are fully biased. Their presence is purely to create a gap between every mini-matrix as well as to conceal any edge-effects in order to avoid interferences from neighbouring mini-matrices during measurements. The Investigator chip houses a mirrored set of these mini-matrices, as shown in Fig. 2.9, with the axis of symmetry passing horizontally across the centre of the chip and with the top half being used to read the analogue signals from pixels via ultrasonic wedge wire-bonds and where the bottom half was intended for pixel-level charge injection.

The bottom half of the chip though was found to possess a design flaw making that section unusable in the Investigator I and was subsequently corrected in the IBis version.

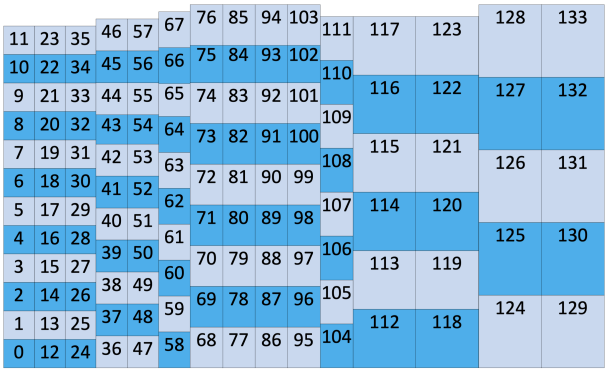


Figure 2.9: Layout of the 134 mini-matrices in the TowerJazz Investigator I chip, varying in square pixel widths from 20 μm to 50 μm.

The geometrical pixel parameters that vary between mini-matrices are shown in Fig. 2.10. This includes changes in the total pixel size, collection electrode size, spacing between the electrode and the neighbouring p-well as well as p-well depths. An additional parameter that varies is the reset method which is either a "diode reset" or an "active reset". Fig. 2.11 shows the circuitry of the two different types of resets available on the Investigator chips.

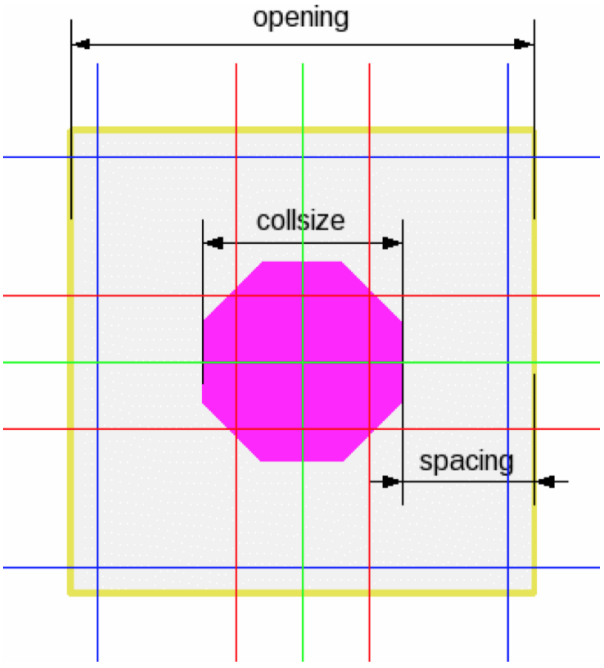


Figure 2.10: Geometrical pixel parameters varied across the different mini-matrices on the Investigator chips.

The basic circuit involved in the signal propagation once induced at the electrode level is shown on Fig. 2.12. After being reset by a dedicated signal (reset signal), three source followers propagate the output signal to the periphery of the chip, each powered by a certain current labelled IFOL (source follower current). Fig. 2.13 shows an image of the full

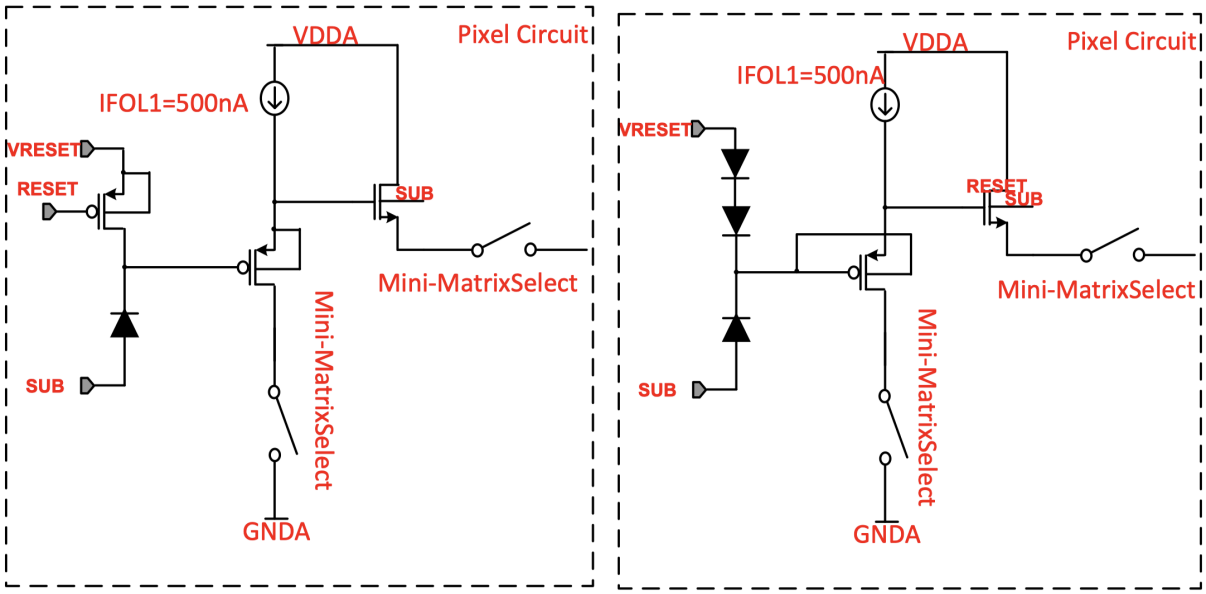


Figure 2.11: PMOS reset mechanism (left) and diode reset (right) circuitry on the Investigator chips.

Investigator I chip along with a close-up of the mini-matrices and wire-bond pads. A specific mini-matrix can be selected for readout by encoding its corresponding binary address onto the chip's respective 9 bit addressing pads.

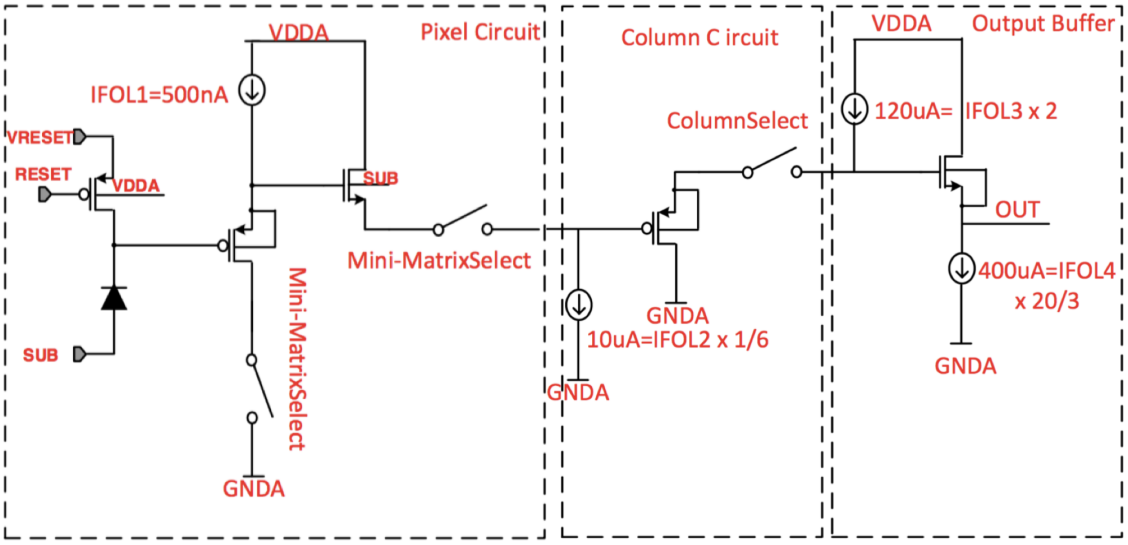


Figure 2.12: Basic signal propagation circuit as found on the Investigator chips.

The changes implemented in version IBis are summarised below:

- Ability to separately bias the p-well and substrate (both being shorted together in version I).
- Fixed charge injection circuit (pixel pulsing).
- Fixed diode reset mechanism.

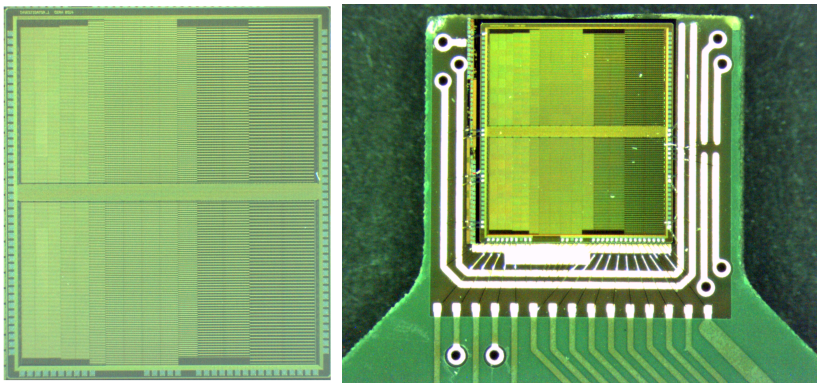


Figure 2.13: A diced and unbonded Investigator I chip (left) and the same chip mounted on its carrier board and connected with ultra-sonic wedge wire-bonds (right).

Additionally, the Investigator II implements the following changes, on top of the Investigator IBis improvements:

- Faster signal propagation to chip periphery through the use of faster source followers (requiring new IFOL currents as a result).
- New set of mini-matrices (pixel sizes and geometries).

As part of this study, a dedicated carrier board was designed to house the Investigator chips, which underwent numerous revisions, adding new features such as an on-board temperature probe (NTC), expansion of readout channels, independent p-well and substrate biasing capabilities, grounding isolations, as well as a jumper-based mini-matrix selection mechanism. Fig. 2.14 shows the first and last versions of the carrier boards. In addition, a PCB shield was also designed to allow the mini-matrix selection to be done through software via the use of an Arduino Mega 2560 board [40].

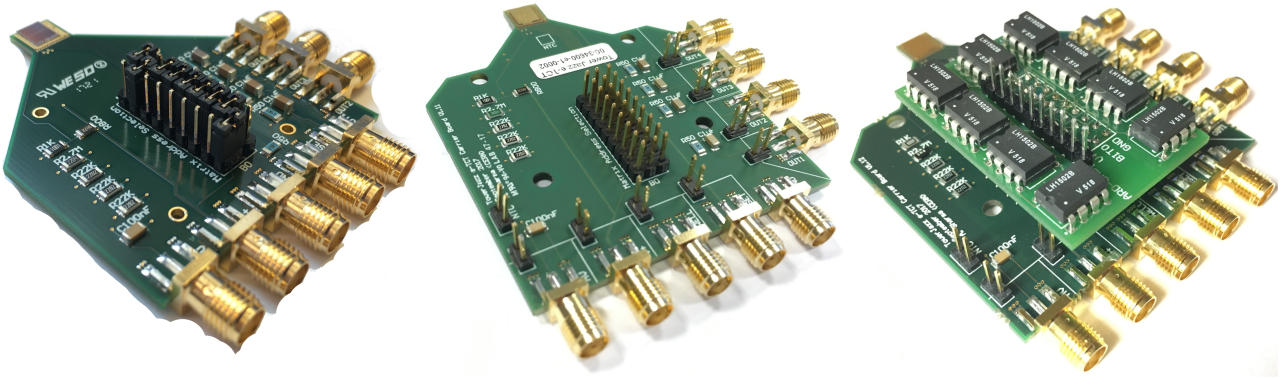


Figure 2.14: Selected carrier board iterations for the Investigator chips, showing version 1.0 (left) and version 1.11 (centre) and with remote mini-matrix selection shield (right). Version 1.11 is backwards compatible with every type of Investigator chip.

2.5.2 MALTA

The MALTA [41] sensor (Monolithic from ALICE to ATLAS), shown in Fig. 2.15, is a 22×20 mm^2 full size demonstrator comprising of 8 distinct vertical sectors with different pixel flavours. The chip's matrix architecture is fully clock-less, leading to a low power consumption. Charge deposition information is extracted from the signal's propagation delay along the respective double column until it reaches the chip periphery. This is also referred to as the signal time walk. The pixel size is 36.4×36.4 μm^2, with the sensor being populated with 512×512 pixels in total. Octagonal collection electrodes vary in diameter between 2 to 3 μm, reducing the input capacitance. A separation between this electrode and the nearby electronics of 3.4 to 4 μm is incorporated into the pixel design to minimise cross talk. The power consumption is of the order of 1 μW/pixel for the analogue power (75 mW/cm^2) and 10 mW/cm^2 for the digital power which makes it suitable for layer 4 in ATLAS ITk pixel detector, given its power budget of 700 mW/cm^2 (detailed in section 4.1).

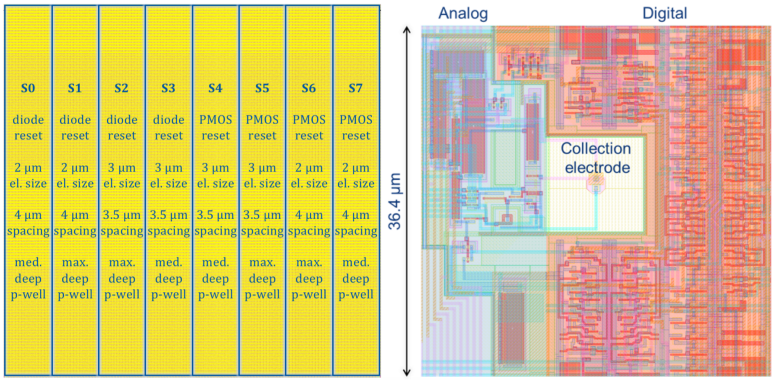


Figure 2.15: Representation of the 8 different vertical sectors that form the MALTA chip (left) and a close-up view of a single pixel showing the separation between analog and digital circuitry (right).

The generation and propagation of a hit signal begins with a single pixel transistor that is charged with the energy deposited on the sensor that acts as a diode. The charge is collected by the collection electrode which in turn causes the the transistor gate to be opened and discharged. The current from that transistor is discriminated by a discriminator that has an adjustable threshold. This subsequently produces a hit signal with a pulse width adjustable between 0.5 ns and 2 ns transmitted asynchronously on a dedicated bus, with one hit emanating per group. Here, a group is defined as an 8×2 pixel structure. There are 32 groups in a double column. The hit propagates along the bus which has 16 lines, with one connecting to every alternating group. There are hence 2 buses, one for every odd and even group respectively along a double column. The reason for using two buses is to avoid cross talk and signal merging. Upon reaching the bottom of the bus, the hit signal reaches the periphery of the chip. At this location two buses from every double column are received. The electronics located at the periphery assign bits to the hit to identify its group type (odd or even), the double column as well as the pixel from which the hit originated. There are hence 33 bits per bus, comprising of the reference (1 bit), the pixel pattern (16 bits), the group address (5 bits) and finally the double column address (9 bits). Before being propagated outside the chip, the hit signal goes through 11 successive levels of merging in

order to produce a single output from the 256 double columns. This process adds a further 2 bits to the output word. Onto this word are appended the additional bits corresponding to the Bunch Cross ID (BCID), which is obtained from a timestamp of a 40 MHz clock that is provided to the periphery of the chip. The Field Programmable Gate Array (FPGA) adds an event counter which is provided by the Level 1 Trigger ID (L1ID) which corresponds to the trigger signal. For the purposes of the results shown within this study, the BCID was replaced with the event counter from the FPGA during data-taking.

As opposed to a typical sampling of the signal at 4 GHz rate, up to 8 copies of the signal are sampled with the rising and falling edges of a 320 MHz clock. For a 1 ns pulse, the time resolution is of the order of 390 ps, where 38 individual signals are sampled. The spread between these readings is of under 1 ns and the efficiency greater than 99%. No temperature dependence (between 10°C to 40°C) has been observed. Timing corrections of the order of 250 ps, known as taps, for each of the 38 signals are applied in order to align the rising edge of each signal. A tap calibration measurement result is shown on Fig. 2.16, where 2000 pulses (each 2 ns long) are propagated along all 38 signal lines and their time of arrival is correspondingly adjusted using taps for a synchronised readout.

A dedicated carrier board was also designed for MALTA with two flavours: a single chip and double chip. The single chip board was used for characterisations, and the double chip to validate chip to chip communication.

The MALTA chips were produced on a 300 μm thick wafer with a selection of samples being thinned down to 150 and 100 μm at the time of writing. Nevertheless, all results shown within this thesis refer to measurements conducted on 300 μm thick sensors.

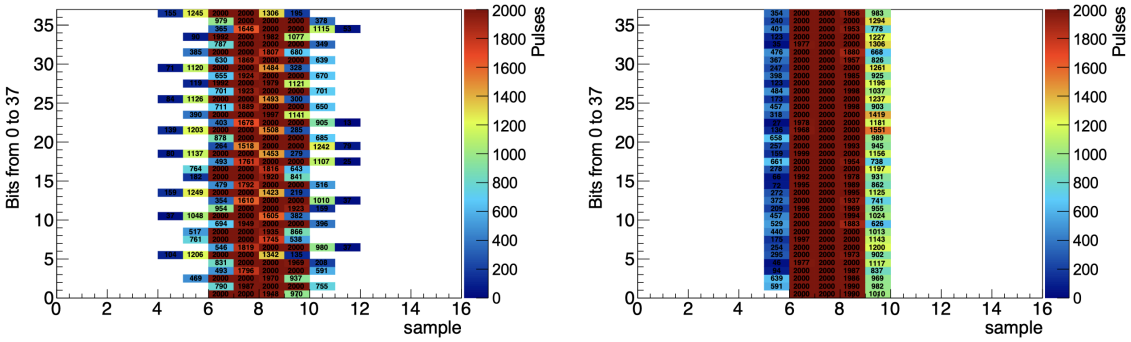


Figure 2.16: Synchronised readout of 38 signal lines reading 2000 pulses each, where the pulse duration is of 2 ns, with applied taps adjusting the time of arrival accordingly.

2.5.3 mini-MALTA

The mini-MALTA is the latest iteration of the TowerJazz designs and comprises of a 64×16 matrix of pixels sized 36.4×36.4 μm^2. This includes 8 sectors with variations in the analogue front-end design, reset mechanism and wafer cross-sectional processing. There are 8×2 special pixels in the top corners to monitor the analogue signals on the collection electrode and front-end output for different sectors. A total then of 64 rows and 16 columns form the 8 double column read-out. A fast asynchronous pulse is formed of 22 bits with synchronisation

and fine time provided at a rate of 640 MHz and BCID at 40 MHz. Every hit, encoded in 22 bits, is transmitted to a short term memory where it is timed in with a 640 MHz clock, where additional 24 bits are added. Then it is input into a FIFO (First In, First Out data buffering memory) 64 events deep at 200 MHz. The FIFO can be read out by a Low Voltage Differential Signal (LVDS) driver at 1.2 Gb/s that uses an 8b/10b encoding, or through a slow read-out path at 40 Mb/s without encoding. The LVDS driver is the same as in the ALPIDE chip developed for the ALICE ITS upgrade [42]. Fig. 2.17 shows the layout of this sensor.

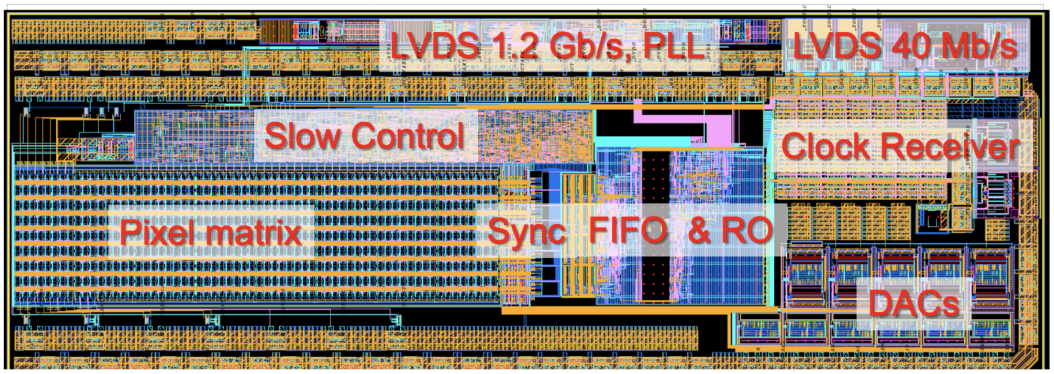


Figure 2.17: Layout of the different components that form the mini-Malta chip.

The analog pixel of mini-MALTA is the same as that of the one for MALTA. Enlarged transistors have been implemented for M6 and M3, shown on Fig. 2.18, in sectors S4, S5, S6, S7 to address the long tail in the noise distribution, observed in MALTA. As shown on Fig. 2.18, sectors S1 and S5 implement an extra deep p-well implant that requires an additional mask. Sectors S3 and S7 implement a gap in the n-implant. Both options try to optimise the electric field at the edges of the pixel, as shown on Fig. 2.19. Sectors S0, S1, S3, S4, S5, S7 implement a diode reset, sectors S2 and S6 implement a PMOS reset, which was identified as being faulty. As a result these two sectors were disabled for all measurement results subsequently shown in Chapter 3. Extensive simulations [43] show these options will improve the charge collection efficiency. The read-out of mini-MALTA is a standard serialiser running at 40 MHz for the slow data output or at 1.2 Gbps for the fast data output.

The mini-MALTA carrier board was also designed on this occasion to be compatible with a larger family of FPGAs, introducing voltage level shifters in order to accommodate this.

The choice of the TowerJazz foundry for these designs was driven by their ability to control design geometry and doping profiles from the feedback provided by measurements conducted, in part, within this thesis but also for practical reasons beyond the technology. With a significant development history between CERN and TowerJazz, numerous software and firmware modules have been created to accelerate the design of the next generation of sensors. For instance, the mini-MALTA sensor's design phase was shortened considerably through the availability of a custom firmware module provided directly by the foundry. Opting for an alternate foundry using a technology other than 180 nm could prove beneficial but would cause a delay in the development as its distinct design processes of the transistor technologies and bespoke tools would need to be understood and adopted.

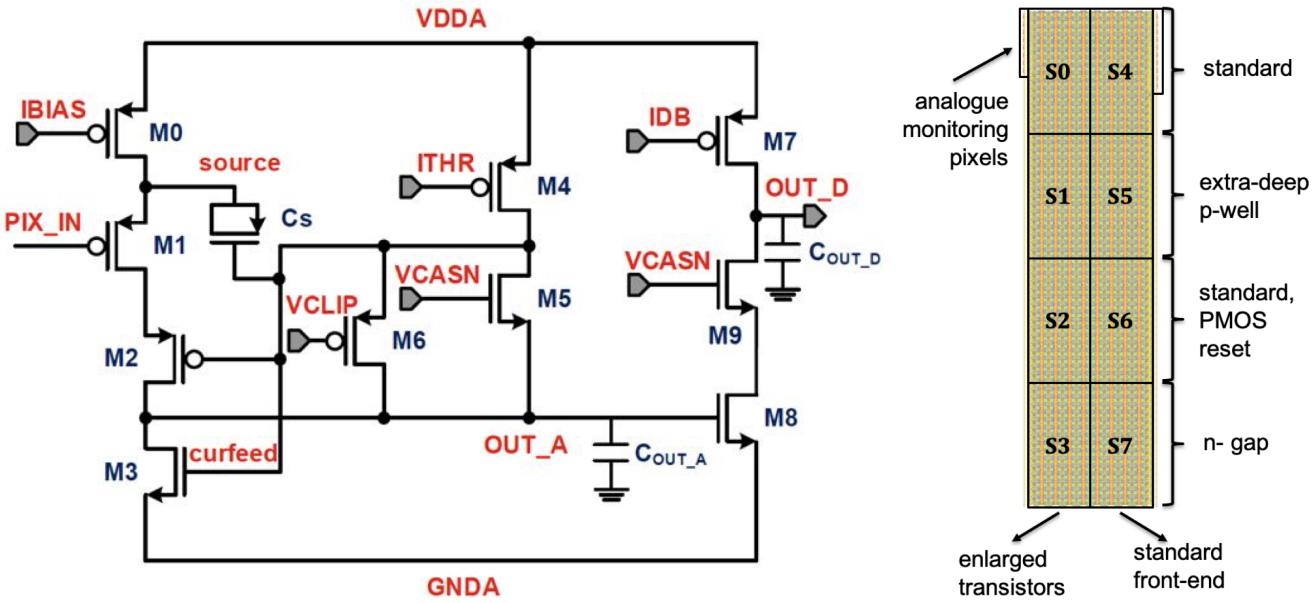


Figure 2.18: Analogue front-end circuitry of the mini-MALTA chip (left) and the layout of the different sectors (right) that form the mini-Malta chip, with pixel 0,0 being in the upper left corner.

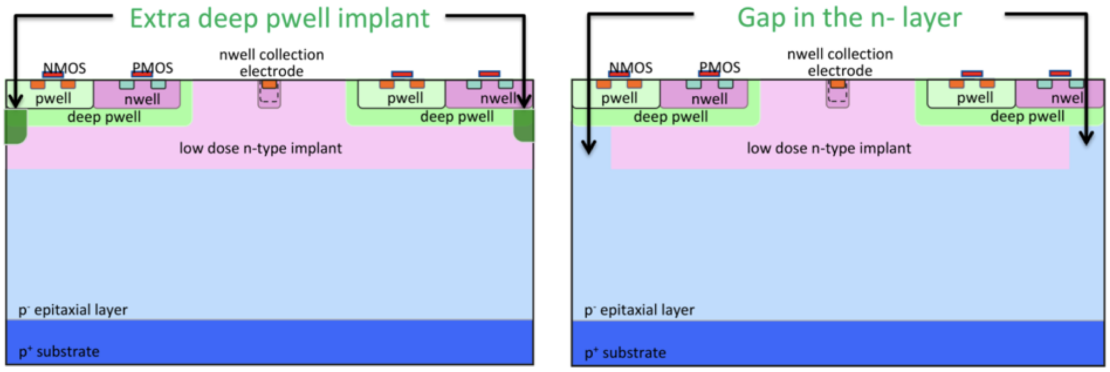


Figure 2.19: Cross-section of the mini-MALTA sensor showing the the extra-deep p-well implants added (left) and the additional gap in the n-layer (right) in order to obtain an optimised E-field configuration for charge collection at the pixel corners.

2.6 Conclusions

Three classes of chip design have been produced towards the consideration of monolithic silicon detectors for layer 4 of the ATLAS ITk, namely, the Investigator chips inherited from the ALICE R&D, the MALTA full sized demonstrator sensor and the smaller mini-MALTA targeting specific pixel design improvements.

Through this process, bespoke carrier boards were produced to interface with these chips. In the case of the Investigators, a highly flexible design was conceived in order to improve on pre-existing boards used by the collaboration. The enhanced board allowed for the complete omission of wire-bonding for mini-matrix addressing, including support for remote addressing, whilst significantly improving signal integrity and being optimised for E-TCT

studies. All PCBs designed towards these studies have been archived onto official CERN repositories within the public domain in order to facilitate future developments [44].

The MALTA readout architecture has targeted 1 Gbps sampling of the output bit-stream of the chip. This relatively high data rate has been made possible through the use of an asynchronous readout scheme bypassing the need for a distributed clock over the active matrix. This design has been achieved whilst maintaining a very low power dissipation of 75 mW/cm^2 and 10 mW/cm^2 for the analogue and digital power respectively. This places the heat load of this chip at 10% of the 700 mW/cm^2 allotted for final sensor operation (assuming a hybrid design), providing a significant margin towards cooling considerations, discussed in Chapter 4.

The extensive design R&D of these chips has enabled a thorough understanding of the TowerJazz 180 nm technology. Numerous subtle design variations are now possible including changes in wafer masks, transistor sizes, pixel sub-structure variations as well as the understanding of the company's memory structure offerings.

Chapter 3

Sensor Characterisations

The TowerJazz Investigators, MALTA and mini-MALTA chips have all undergone extensive characterisations in order to understand their performance and optimise their design for radiation hardness. In order to achieve this, in the first instance, all chips underwent laboratory testing to validate their correct operation and calibration measurements along with more specific assessments identifying charge sharing in the case of the Investigator chip. Following these laboratory-based studies, efficiency measurements were conducted at the various test-beam facilities mentioned in section 2.4. This chapter details the numerous assessments conducted on all of these chips along with the different test-beam campaigns to study the MALTA and mini-MALTA samples.

3.1 Laboratory Measurements

Lab measurements were conducted to characterise the different prototypes. Namely, the charge collection efficiency through E-TCT (section 3.1.1), punch-through through current-voltage characteristics measurements known as I-V measurements (section 3.1.2), the gain with radioactive sources (section 3.1.3), and finally the threshold (section 3.1.4) through dedicated scans.

3.1.1 Edge Transient Current Technique

Laser Transient Current Technique (TCT) is a method used to characterise the signal response of silicon sensors through the deposition of photons onto the particle sensing region (epitaxial layer) using a highly focussed laser beam. The focal point of the beam is used to scan the volume of the sensor to obtain qualitative assessments of its charge collection efficiency, charge sharing fraction and signal timing measurements for various depletion voltage configurations before and after different irradiations. This method is referred to as Edge-TCT in cases where the laser light is oriented to be incident on the edge of a sample, necessary when a metalisation layer is present on the sample surface or to study the depletion depths.

The wavelengths commonly used for characterising silicon sensors are that of red (660 nm) and infra-red (1064 nm) light. These are selected for their practical absorption lengths in silicon, as shown on Fig. 3.1 as well as their commercial availability. The results presented in this study were obtained using the 1064 nm infra-red laser light.

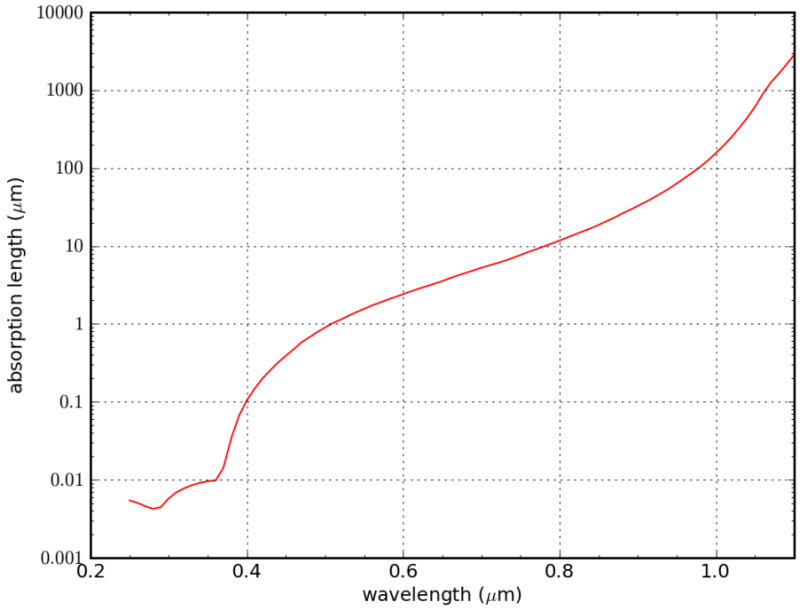


Figure 3.1: Absorption length as a function of wavelength in silicon at room temperature. [45]

The beam spot size is typically ensured to be a fraction of the size of the smallest pixel structure located on the sensor. In the case of the Investigator pixel sensors, this corresponds to < 20 μm and the target laser beam spot size is thus of the order of < 5 μm. The beam spot measurement is described below.

The expected beam spot diameter (W_0) and a Rayleigh length (Z_R) are given by the following expressions:

$$W_0 = \frac{\lambda f}{\pi W} \qquad Z_R = \frac{W_0^2 \cdot \pi}{\lambda} \tag{3.1}$$

where λ is the wavelength of the laser system, f is the iris filter aperture and W is the diameter of the beam before the focusing lens. These lead to an expected $W_0 \approx 5$ μm and $Z_R = 300$ μm. These different parameters are illustrated geometrically on Fig. 3.2.

Experimental Setup

The experimental setup is described in three stages: the optical chain, auxiliary equipment and the readout chain.

The optical chain begins with the powering and laser triggering being provided by a PicoQuant PDL 800-B single channel laser driver at frequencies ranging from 5 to 80 MHz with a pulse width of 50 ps. This in connected to a PicoQuant LDH-P-C-660 diode laser

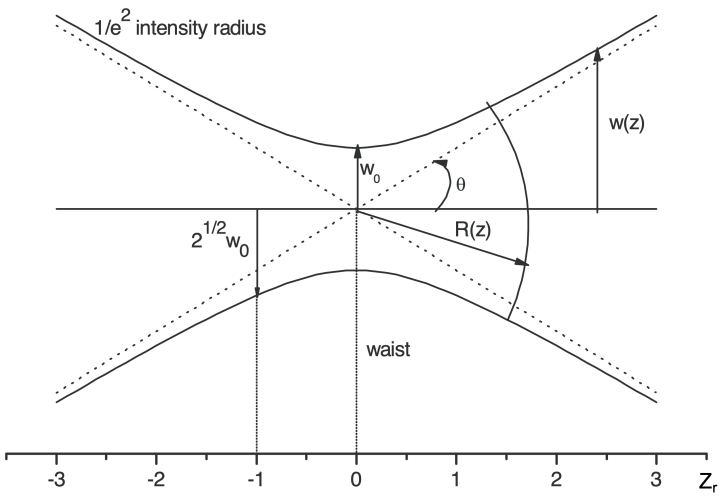


Figure 3.2: Laser beam width as a function of position in units of Raleigh Range z_R.

head which emits near infra-red light of 1064 nm wavelength. The laser light produced is then transmitted to the optical assembly inside a large copper Faraday cage via a 2 m long TN1064R2A1B Angled Physical Contact (APC) optic fiber. The optical assembly consists of: a micro-focus collimator, diameter adapter, a 15× Galilean beam expander, a plano-convex lens and a mechanical iris, as shown on Fig. 3.3.

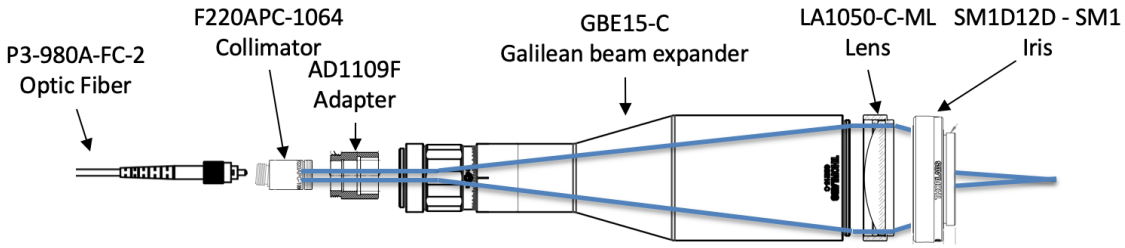


Figure 3.3: Optical assembly used to achieve a beam spot size of $< 5\ \mu m$.

The rationale for initially expanding the beam in order to then obtain a small spot size comes from the need to counter the effects of spherical aberrations, which is the dominant source of aberrations in the optical system. A qualitative representation of the different contributions of diffraction and spherical aberrations is illustrated in Fig.3.4.

The focussed beam of light is then shone directly onto the Device Under Test (DUT). The DUT, resting on its carrier board (see section 2.5.1), is held in place on a custom mechanical support made of plexiglass and copper. This support assembly is anchored on two perpendicular UTS100PP Newport linear stages. This allows for the DUT to be swept across a 2D plane perpendicular to the laser beam. The optical assembly is held in place on a large vertical IMS100V Newport linear stage allowing for the focal plane of the focussed light exiting the optical assembly to be shifted relative to the DUT. All three linear stages are controlled via a Newport ESP301 controller.

The DUT, linear stages and entire optical assembly are housed inside a 3 m^3 light-tight copper Faraday cage. The copper section of the mechanical support allows for silicon oil-based cooling to be circulated via a Huber chiller [47], which can bring the temperature of

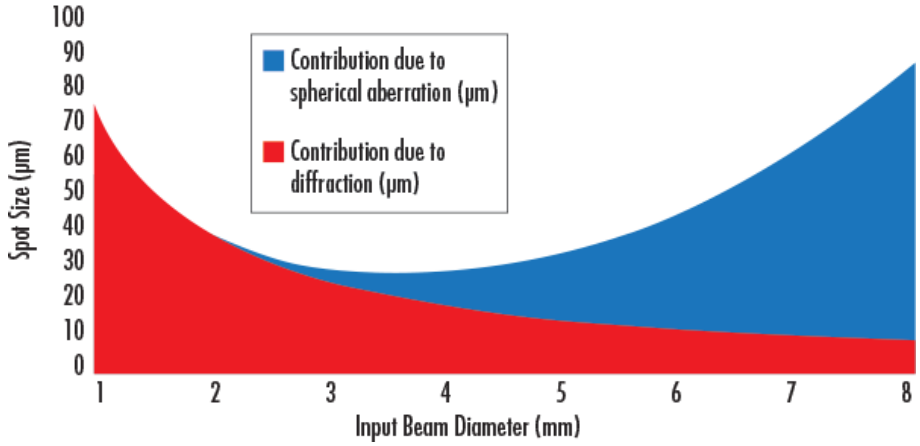


Figure 3.4: Qualitative representation of the distinct contributions of spherical aberrations and diffraction towards the final beam spot size. [46]

the DUT to around -15°C in the case of irradiated sensors (see section 2.4).

The output signal readout comprises of exiting the respective carrier board trace via an SMA (SubMiniature version A) connector into a Cividec 40 dB amplifier [48], which is then split between an oscilloscope for real-time monitoring and a DRS [49] digital oscilloscope for data acquisition. All signals throughout the readout chain are propagated via SMA cables to benefit from their enhanced shielding design.

Once fully operational, the triggering scheme, shown on Fig. 3.6, involves providing a negative-logic reset pulse to the chip (resetting on the pulse's low voltage level), followed by a laser trigger signal (triggering on the falling edge). The laser pulses instantly (∼ns) after receiving the trigger signal and a delay of a few microseconds is placed between the chip reset pulse and the laser trigger in order to reduce interference pick-up on the signal line. The signal generated from the hit can then be observed ∼200 ns after the laser trigger. A sample output signal from a single pixel is shown on Fig. 3.7, comprising of a sharp increase from drifting charges being collected rapidly followed by a slower signal contribution coming from charges undergoing diffusion.

The coordinate system used in this study is shown on Fig. 3.8, where the laser beam propagates in the direction of the Y-axis and the sensor thickness with the direction from substrate to collection electrode is along the Z-axis. Neighbouring pixels in the same focal plane are hence along the X-axis.

In order to ensure the focussed beam spot is under the 5 μm target, a First Sensor PC10-Q-TO [50] test diode shown on Fig 3.9 with enhanced sensitivity to 1060 nm wavelength was used to conduct what is known as a knife-edge measurement. This involves depleting the diode through reverse bias and collecting charge generated by the laser as it is gradually swept across the edge of the diode. This should yield a sigmoid curve with no signal generation when the laser is outside the active area of the diode and a gradual increase as it enters the diode active area. Once the laser beam spot is fully located within the diode's active area, the signal amplitude then plateaus. The slope of this S-curve can then be measured to infer the size of the laser beam spot. A larger (defocussed) beam would yield a low slope, whereas

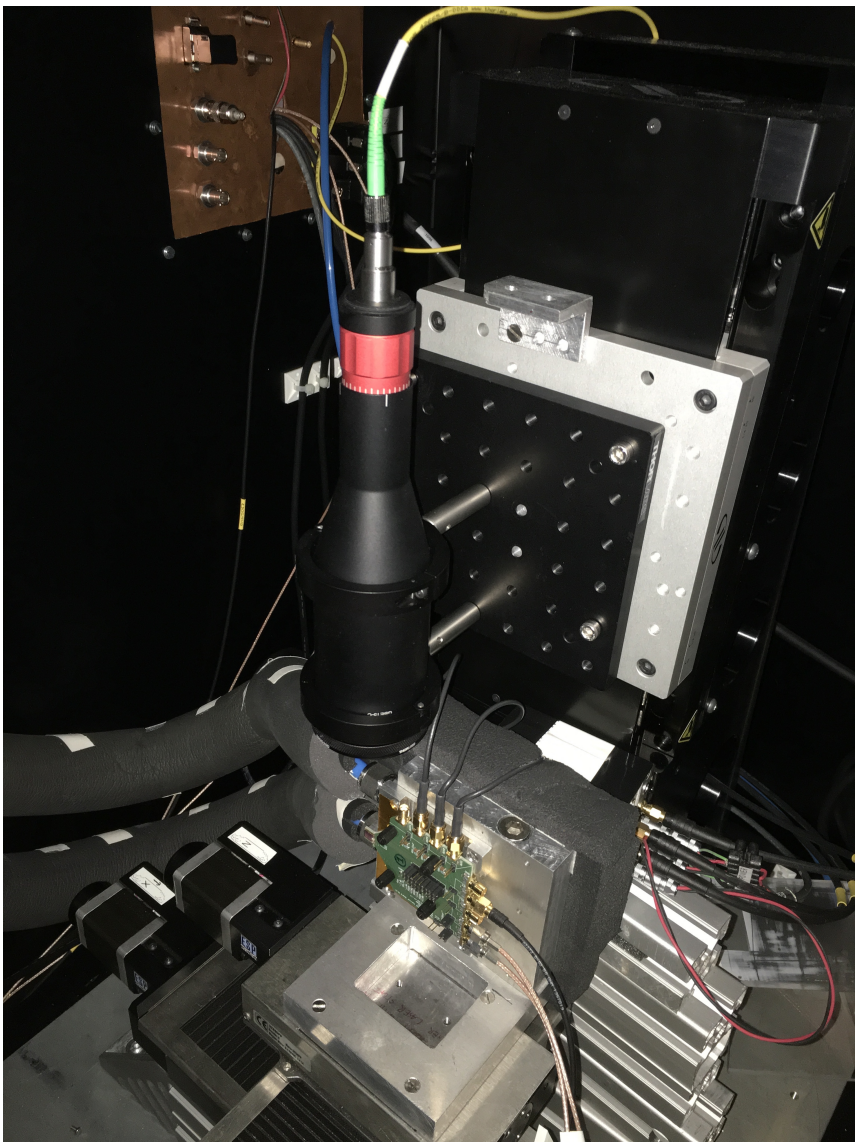


Figure 3.5: Linear stages and optical assembly housed inside a copper Faraday cage.

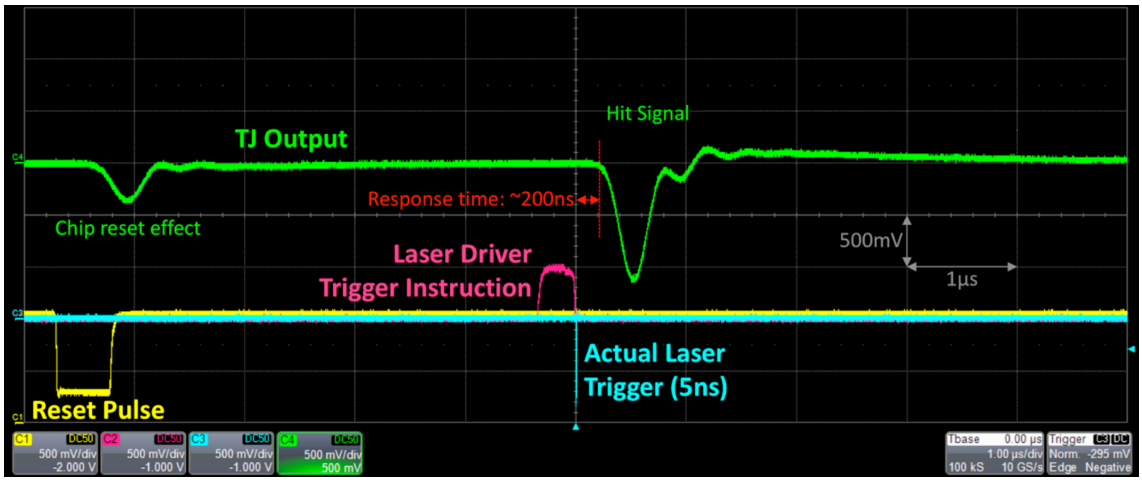


Figure 3.6: Triggering scheme for all Investigator chips undergoing E-TCT measurements.

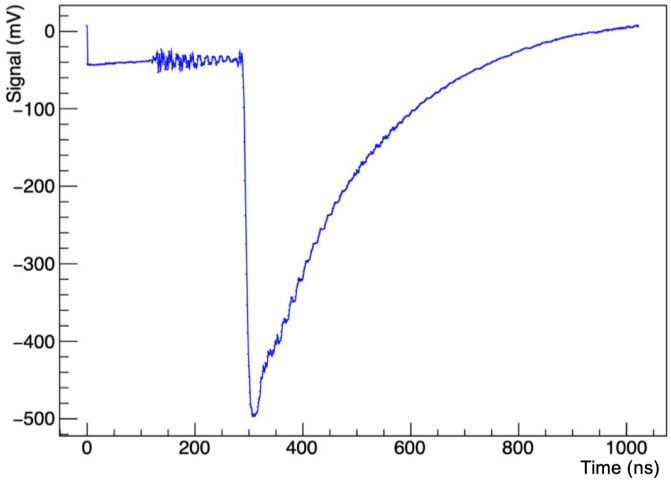


Figure 3.7: Sample signal produced on an Investigator I's 50 μm pixel.

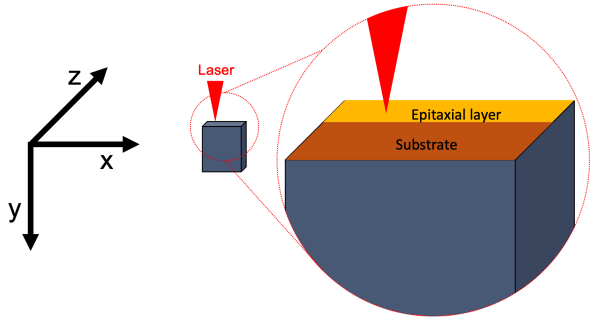


Figure 3.8: Definition of the coordinate system used in the E-TCT study.

a highly focussed narrow beam would lead to a sharp rise in the signal collected as the laser crosses over onto the active region of the diode. Such a knife-edge measurement is shown on Fig. 3.10. The measurement is repeated by varying the distance of the focal plane relative to the test diode's surface in order to obtain the smallest possible beam spot size.

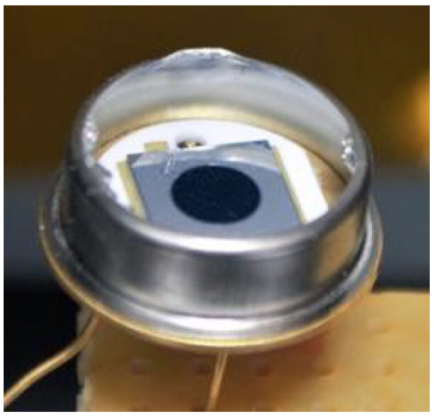


Figure 3.9: Image of the First Sensor PC10-Q-TO diode used for knife-edge measurements.

An additional parameter that can be optimised to further reduce the laser beam spot size is the diameter of the aperture on the iris mounted on the exit lens of the Galilean beam expander. This allows one to circularly truncate any aberrations on the beam shape and

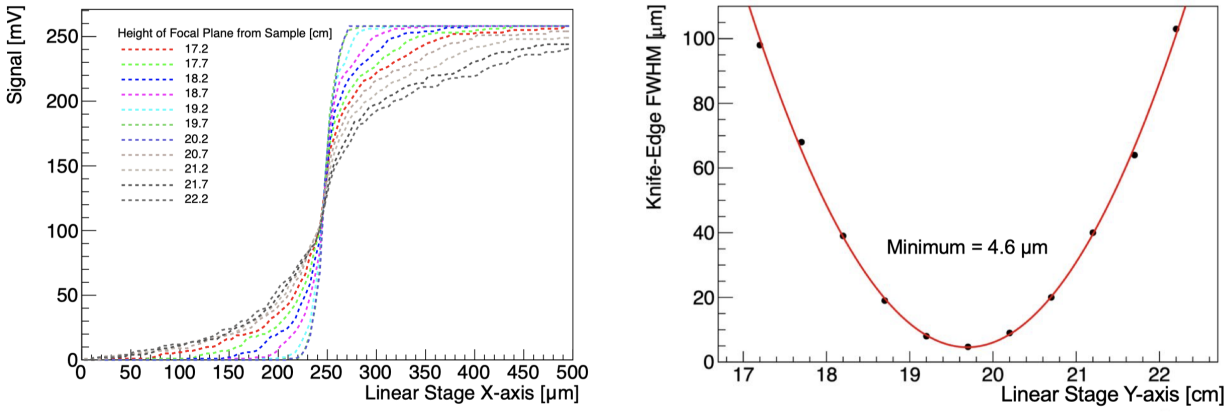


Figure 3.10: Multiple knife-edge measurements showing S-curves of varying slopes (left) and the subsequent minimum beam spot waist obtained through the FWHM of the curves.

produce a cleaner and finer beam spot size. This parameter must be adjusted with care as too large an aperture would lead to a needlessly unfocussed beam yet too narrow an opening would drastically reduce the beam intensity reaching the DUT. Fig. 3.11 (right) shows the effect on the S-curve's slope as the iris aperture is varied.

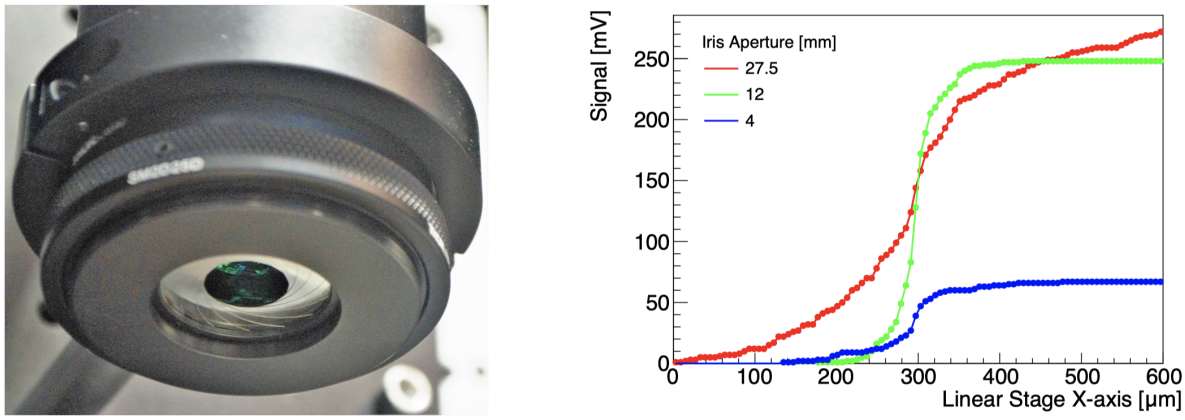


Figure 3.11: Image of the laser iris as well as the direct effect upon changing its aperture on the slope of the knife-edge S-curves.

Sensor Polishing

An important stage in the preparation of samples towards E-TCT measurements is the assessment of their edge roughness, typically conducted visually. In the case of severe scratches from the rough dicing of the samples from their original wafer, sensor polishing must be undertaken in order to ensure a uniform smooth edge to allow for homogeneous penetration of the laser light into the silicon as it is swept across its edge surface. This can be achieved following a four step process [51]:

1. **Clamping**: the sample securely between two teflon bars, as shown on Fig. 3.12.
2. **Lapping**: using 3 μm lapping film to remove coarse aberrations on the silicon edge's surface.

3. **Polishing**: fine polishing using a 1/10 μm grade diamond paste to remove finer imperfections.

4. **Cleansing**: finally, all paste and silicon residue must be carefully removed using isopropanol.

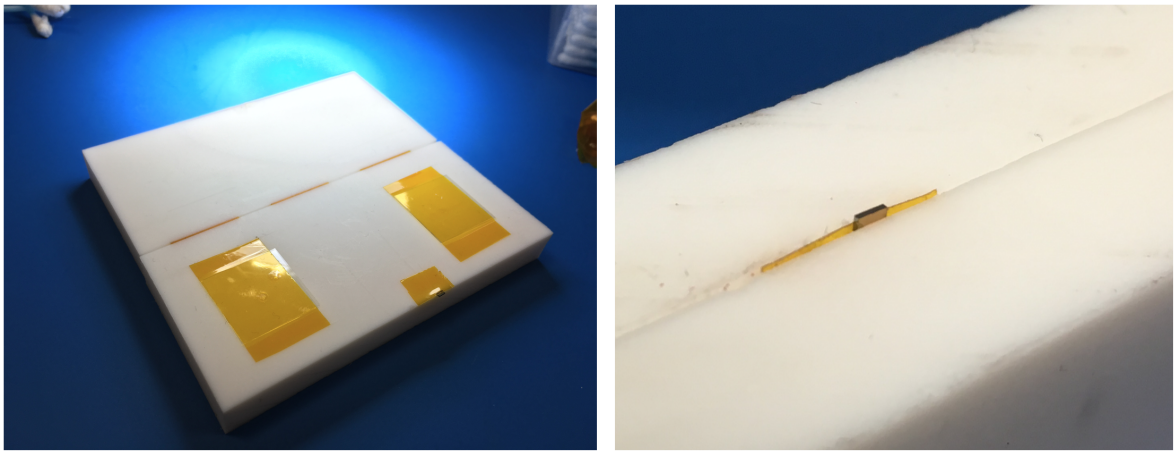


Figure 3.12: Teflon bar based clamp to secure silicon sensors for edge polishing.

Fig. 3.13 shows the sensor polishing steps that were optimised using a dummy KC53A [52] sensor.

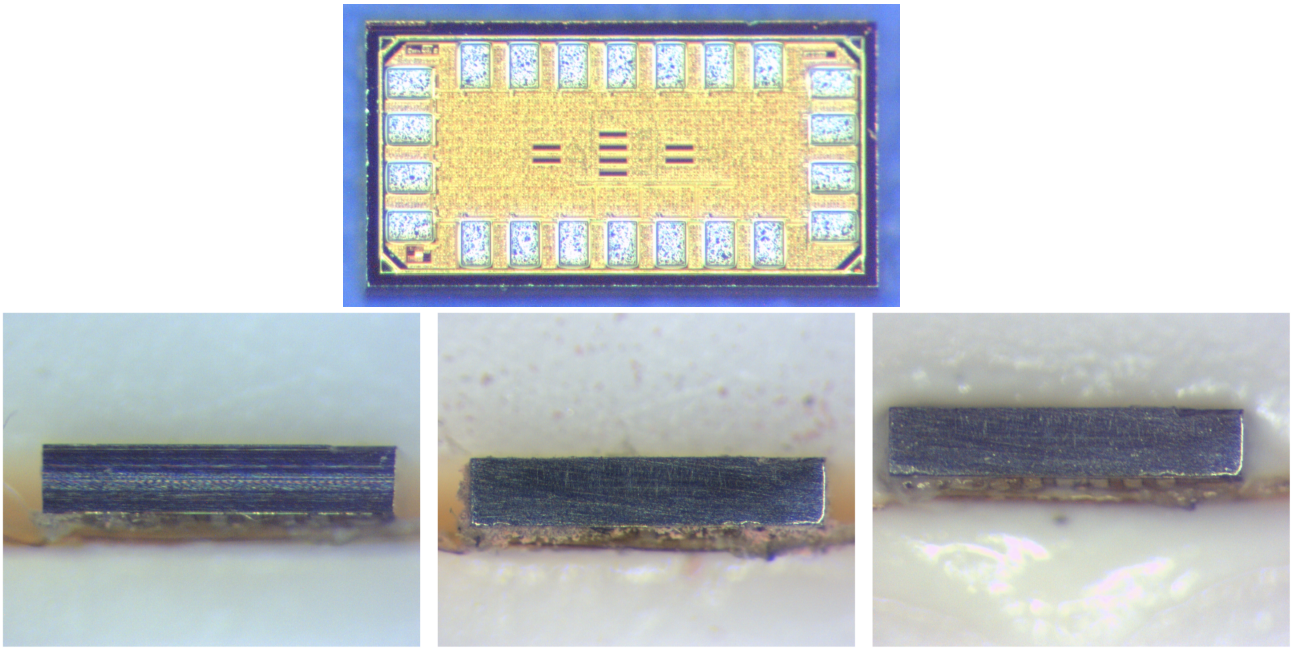


Figure 3.13: KC53A test sample (top), with its original edge (left) undergoing lapping (centre) and fine polishing (right), yielding a smooth uniform surface.

Measurements

All E-TCT measurements begin with an intentionally unfocussed beam at high intensity, in order to locate the position of the DUT (given its typically small size of $< 50\ \mu m^2$). Once a low resolution (large step size) surface sweep has been conducted in the XZ plane which yields a signal, a focus scan is performed. This involves a three step procedure, where first: an XZ scan is performed for different Y-coordinates, as shown in Fig. 3.14 (left). Secondly, a Z-projection is taken for each 2D scan along the centre of the DUT from which a peak of varying width will be obtained, shown on Fig. 3.14 (right). Finally, the width of these distributions is then plotted for every Y-coordinate yielding a parabola, where its minimum pertains to the location of the focal plane coinciding with the DUT shown on Fig. 3.15.

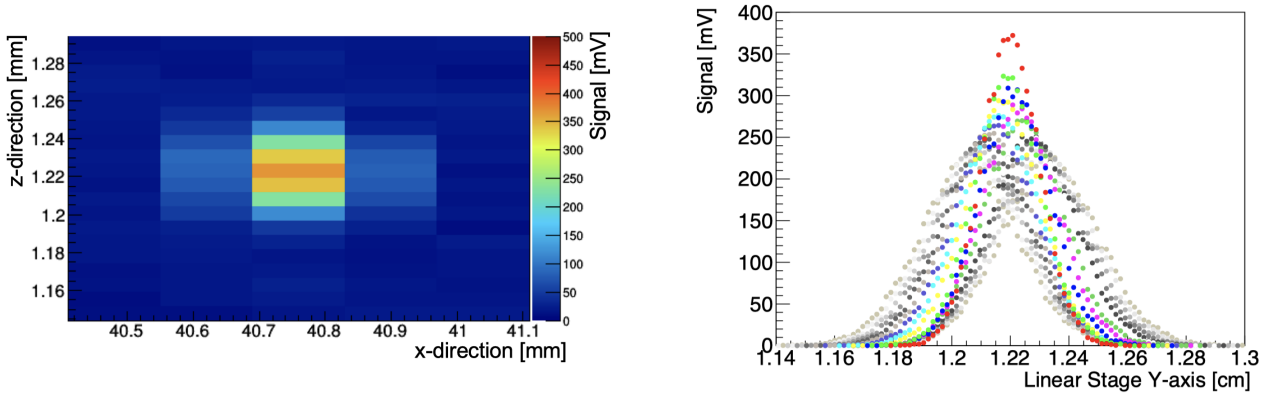


Figure 3.14: 2D surface sweep of the edge of a sensor as the first step of a focus finding scan and its corresponding Z-projections along the centre of the DUT yielding peaks of varying widths with the narrowest one corresponding to the focal plane coinciding with the DUT.

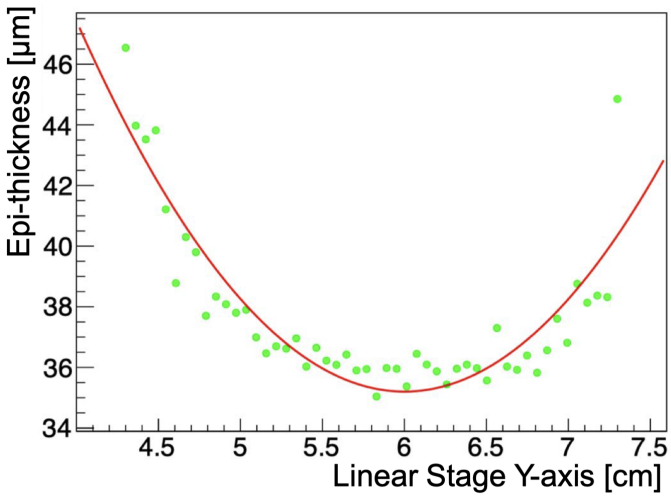


Figure 3.15: Parabola formed from the distribution of the peak widths yielding the epi-width as its minimum.

Upon identifying the correct focal plane coordinates, the Y-coordinate is frozen, and a fine surface scan can be conducted. Signal response can then be observed with the same scan at varying substrate voltages, also shown on Fig. 3.16.

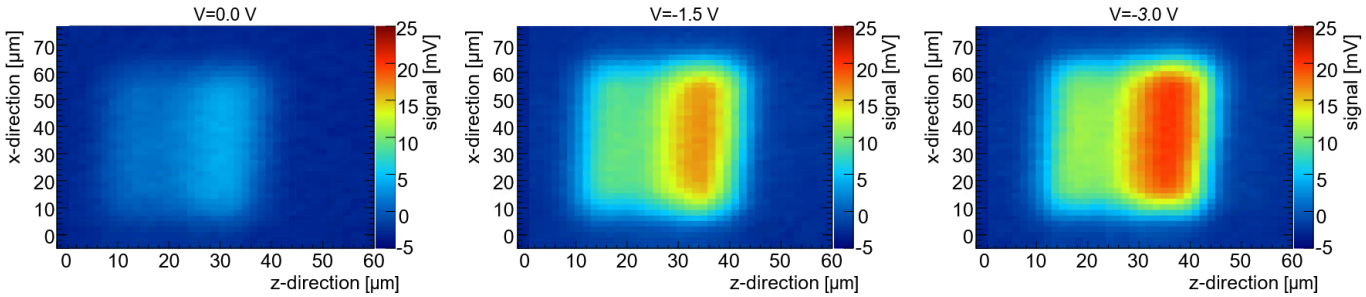


Figure 3.16: Surface scan at 0, -1.5 and -3 V substrate biasing on an Investigator I, showing the increased signal amplitude on a single 50×50 μm^2 pixel.

Further analysis can then be performed using the 2D scans obtained over the chip's wafer cross-section at the correct focus. This involves:

1. **Depletion depth** measurement using a Z-projection. By repeatedly scanning the laser along Z at various depletion, the depletion depth can be estimated from the "full width half max" (FWHM) of that projection as shown in Fig. 3.17 as measured on mini-matrix 129 on a single 50 μm pixel. Fig. 3.18 shows the effect of varying the substrate voltage on the projection.

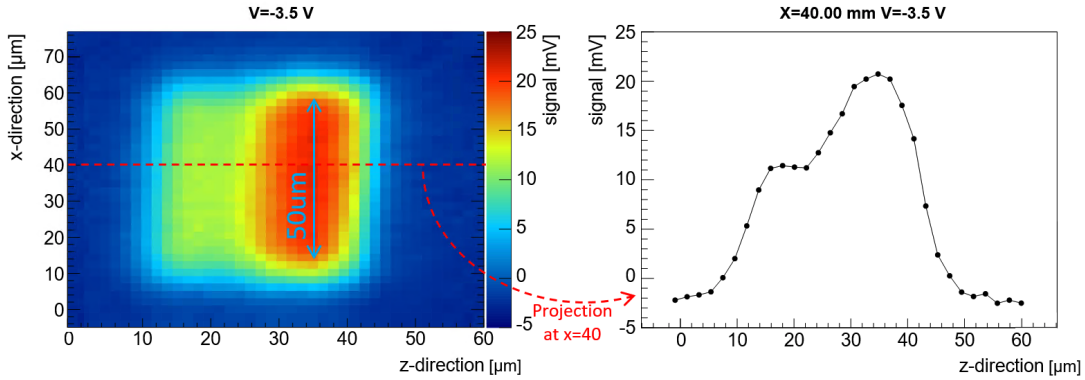


Figure 3.17: Z-projection taken from a 2D surface scan producing a signal peak.

2. **Resistivity** measurement of the epitaxial layer. Using the depletion depth data as the function of the voltage, the resistivity can be extracted by applying a fit using the relationship described in Eq. 2.8. Though this method of obtaining resistivity values is theoretically feasible, the measurement results obtained were provided to the foundry which in turn declared them to be unreliable as they displayed large variations with the known resistivity. A more robust (though destructive) method of identifying the resistivity of such a sample is Spread Resistance Profiling, which involves diagonally cutting the wafer and conducting a direct resistance measurement using a probe along the depth of the wafer.

3. **Charge sharing** performance. This can be assessed by considering the output of the neighbouring pixels along a row perpendicular to the laser. Since the output of up to

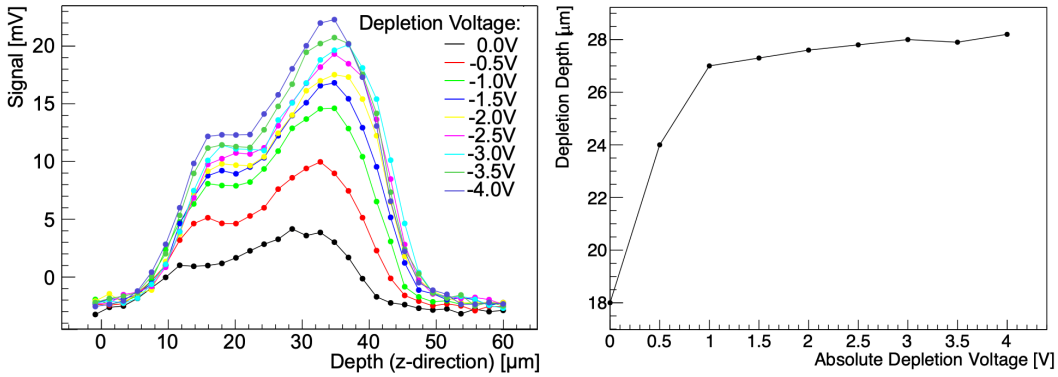


Figure 3.18: A rapid increase in the signal generation can be observed as the substrate voltage is gradually increased (left) along with its corresponding expected increase in depletion depth obtained from the FWHM of the projection data along Z at the different substrate voltages.

three pixels can be read out simultaneously on individual channels, by taking an X-projection of each channel, the contribution of the charge collected from neighbouring pixels (charge-sharing) can be identified. This is shown on Fig. 3.19.

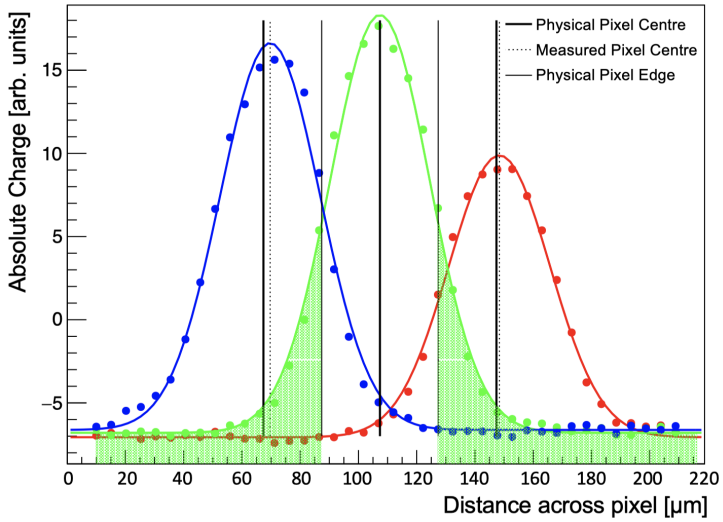


Figure 3.19: Charge sharing analysis conducted on mini-matrix 123 using three consecutive 40 μm pixels individually read out. The green area marks the contribution to the central pixels signal coming from charge collected from the neighbouring pixel volumes. Dots represent the charge collected at each location, to which are applied Gaussian fits for each pixel.

A general trend in charge sharing fraction was observed across the different mini-matrices measured whereby the smaller the overall pixel size the higher the charge sharing fraction, as expected. The size of the collection electrode though did not have a significant effect on this fraction, as summarised in Table 3.1, with the error extracted from the Gaussian fits applied on each pixel output. It must be noted that a variation in the peak heights was also observed in Fig. 3.19 and subsequently attributed to unmatched track lengths of each pixel output on the carrier board PCB shown on Fig. 2.14. This variation is also the cause for the asymmetry in the charge sharing results when comparing the left and right sides. The qualitative trend observed across different pixel types still holds nevertheless.

Pixel Size [μm]	Electrode Size [μm]	Right Side % ($\pm$ 0.1)	Left Side % ($\pm$ 0.1)
40	20	20.4	14.9
40	15	20.3	14.5
40	10	20.2	15.3
40	4	22.7	26.6

Table 3.1: Summary of charge sharing results across a selection of mini-matrices on the Investigator I chip.

3.1.2 I-V Measurements

Current-Voltage (I-V) measurements are a common way to characterise silicon sensors as they can measure breakdown voltage and leakage current which are sensitive to process problems. In the case of the Investigator IBis chip, such a measurement was of particular interest due to the observation of a failure point on TCAD simulations [43] performed using the original Investigator I chip design. This showed what is referred to as a punch-through current being observed in the case of a large difference between the substrate and p-well biases overcoming the isolation layer normally formed at the horizontal n-implant in the epitaxial layer as shown on Fig. 2.8 (right).

In order to specifically measure the punch-through current rather than the total leakage current, the wiring of the I-V measurement setup must be such that the reference ground is applied to the p-well. The full wiring scheme is shown on Fig. 3.20. In order to physically connect the different terminals onto the chip, a probe station with micron-sized needles is used as shown in Fig. 3.21.

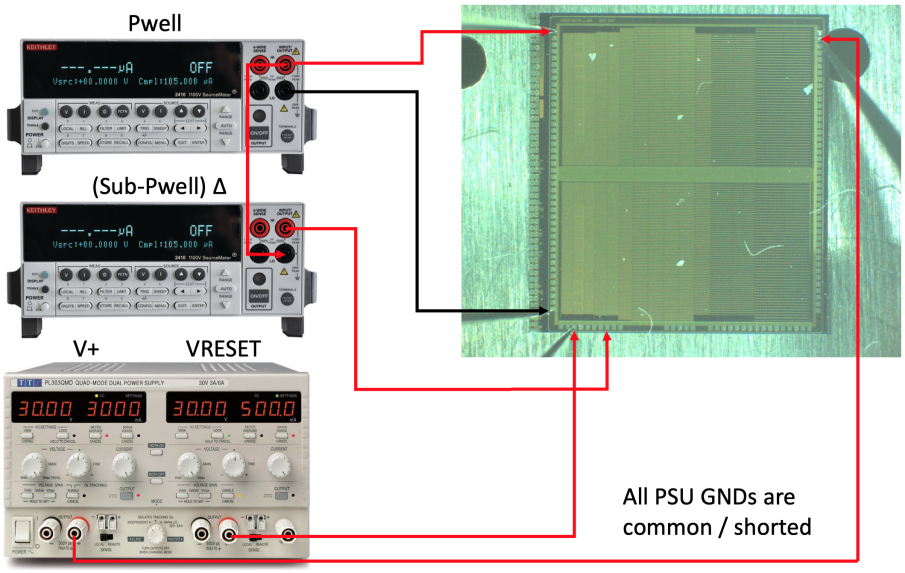


Figure 3.20: Wiring schematic for punch-through current measurements on the TowerJazz Investigator IBis.

The measurements were conducted on several samples of Investigator IBis originating from wafer 4 (high doping) and wafer 12 (very high doping) each. Fig. 3.22 shows the punch-through current measured on the two wafers. The punch-through was observed for

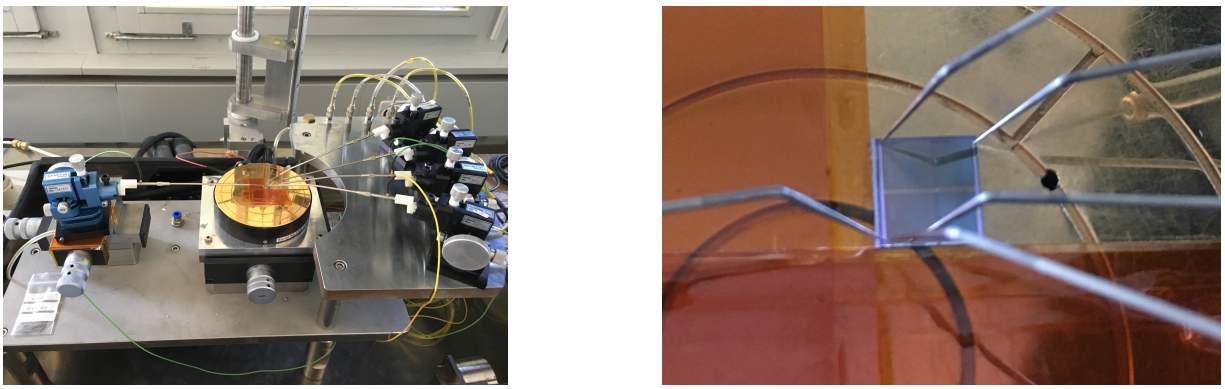


Figure 3.21: Probe station setup for punch-through I-V measurements (left) with close-up of probe needle connections (right).

the wafer 4 samples while for the higher doped wafer 12 displayed no increased current at 15 V. These results show that an increased doping level in of the horizontal n-implant provides a greater degree of isolation at the p-well-to-n-implant and n-implant-to-epitaxial layer junctions. As a result, a higher p-well potential would be needed to overcome these junctions, and the maximum operating voltage range of the p-well still prevents any punch-through current to occur. This has also lead to a preference towards higher doped designs to benefit from the superior punch-through current prevention. The outlying error bars on Fig. 3.22 are likely caused by poor contact between the probe needles and the bond pads of the chip under test.

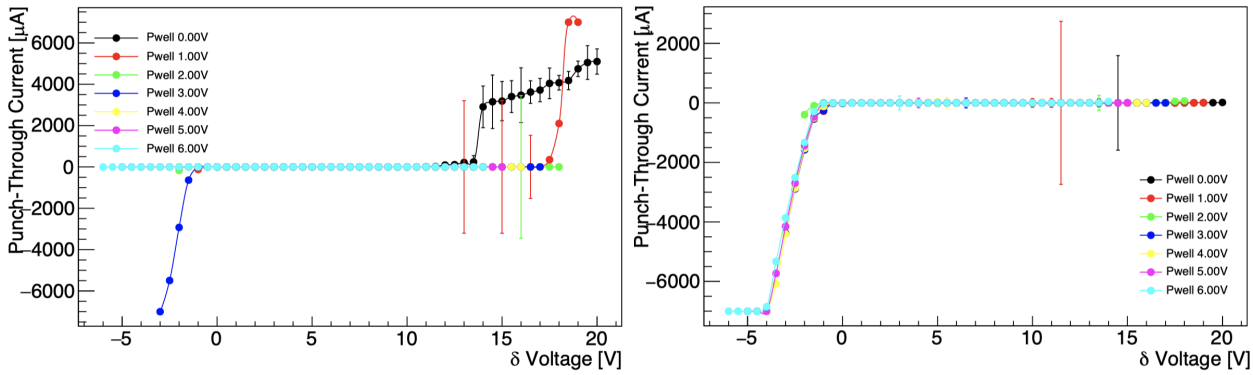


Figure 3.22: Punch-through current measurements on wafer 4 (left) and wafer 12 (right). A current spike is observed clearly on the lower doped wafer 4 samples.

3.1.3 Fe-55 Source Measurements

Characterisations of sensors using radioactive sources provides numerous insights as their decays are well known and in some cases can provide clean monochromatic sources of energy deposition, as is the case with Fe-55 which has a half life of 2.7 years. For example the Fe-55 decay into Mn-55 through electron capture is regularly used to reliably calibrate devices. The resulting lowered electronic reconfiguration then leads to the emission of Auger electrons and X-rays. The three main X-ray photon energies produced are k_α at 5.89875 keV with an emission probability of 16.2% and 5.88765 keV at 8.2%, appearing monochromatic, and k_β at 6.49045 keV with an occurrence of 2.85% [53]. k_α corresponds to electron transition to the innermost k shell from the second l shell. Similarly, k_β corresponds to electron transition to the k shell from the higher m shell.

Such characterisations were systematically conducted on an Investigator IBis sample. In order to achieve this, a new support was developed as shown on Fig. 3.23, allowing for a Fe-55 source to be held <3 mm above the surface of the chip to maximise hit rates as well as making use of the linear motion stages to sweep across the different mini-matrices on the sample. Fig. 3.24 shows a sample pulse produced by the iron source. This signal originally produced as a negative pulse has been inverted for this study to produce the energy spectrum matching general convention increasing along the x-axis.

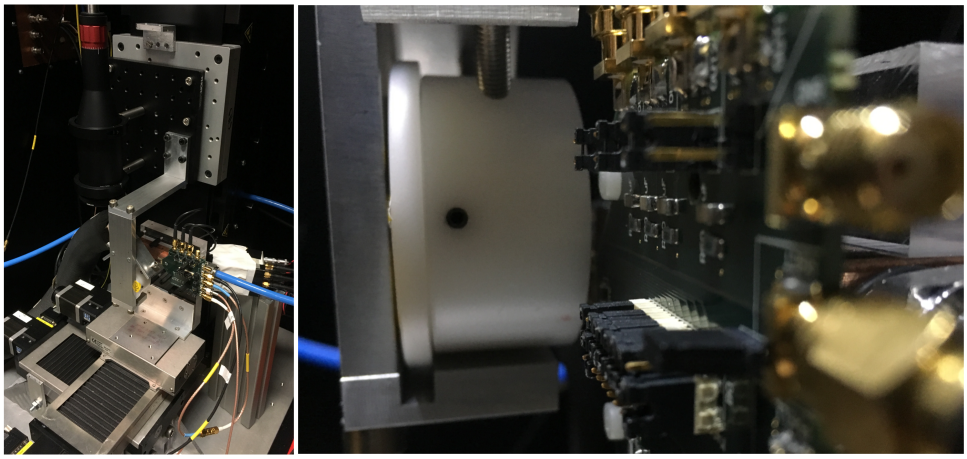


Figure 3.23: Fe-55 support holding source <3 mm away from Investigator IBis surface, iteratively moving across the different mini-matrices, where each is selected automatically via the mini-matrix selector shield.

100k waveforms were acquired each for different p-well and substrate configurations. Each signal's maximum amplitude was then extracted to produce an energy spectrum, as shown on Fig. 3.25. This is shown for varying substrate voltages whilst maintaining the same p-well level. The higher amplitude k_α and lower k_β are clearly visible and can be observed to shift towards a higher signal amplitude with increasing substrate voltage, as expected.

Variations in baseline noise were also quantified, as shown on Fig. 3.26, where a double Gaussian structure was observed. Nevertheless, biasing was shown to have no effect on the baseline noise.

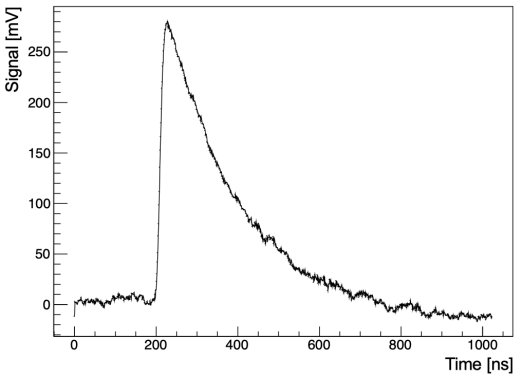


Figure 3.24: Sample signal produced by an Fe-55 source on an Investigator IBis' 28 μm pixel. The originally negative signal has been inverted for the subsequent analyses.

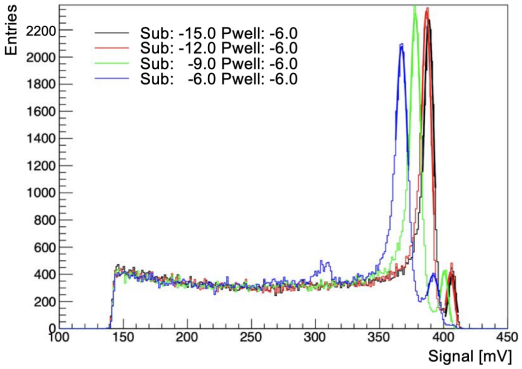


Figure 3.25: Signal response on an Investigator 1Bis chip to Fe-55. The sensor was biased to -15 V and the p-well bias was fixed to -6 V.

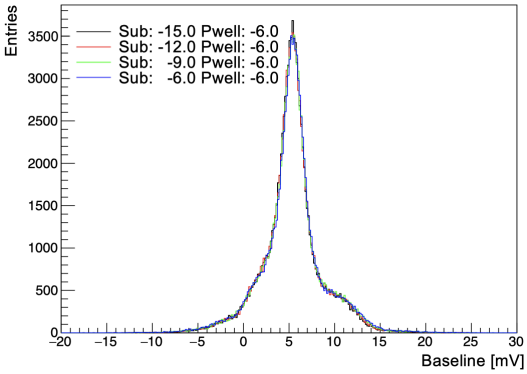


Figure 3.26: Baseline noise showing no dependence on biasing configurations, as observed on an Investigator IBis' 28 μm pixel.

The energy difference between both peaks for a given spectrum can be used to deduce the gain for a given mini-matrix. This is shown on Fig. 3.27 where the gain is calculated as described by Eq. 3.2. A clear decrease in the gain can be observed as a function of decreasing substrate bias of the order of 0.1 mV/e$^-$ when biasing the substrate from -6 to -15 V.

$$\text{Gain}[\text{mV/e}^-] = \frac{\Delta\text{V}(\text{k}_\alpha, \text{k}_\beta)\ [\text{mV}]}{(6.49 - 5.89)\ [\text{keV}]\ /\ 3.6\ [\text{eV}]} \tag{3.2}$$

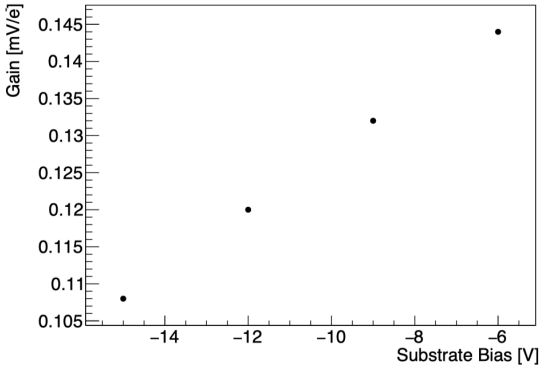


Figure 3.27: Gain as a function of bias voltage on an Investigator 1Bis.

The signal rise times can be calculated using the delay between the signal crossing 10% up to 90% of the maximum amplitude (also known as the "10-90" rule). The fall times can be calculated in a similar manner. The distributions of the rise and fall times are summarised on Fig. 3.28, as observed on mini-matrix 75. The rise times show a consistently narrow time distribution expected from their drift signal contributions. The fall times though display a broader distribution at lower signal amplitudes due to the larger influence of noise.

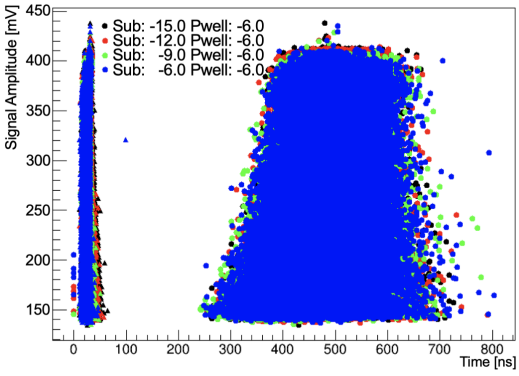


Figure 3.28: Timing measurements showing the rise (triangles) and fall (circles) times for different p-well and substrate configurations as observed on mini-matrix 75.

3.1.4 MALTA

MALTA comprises of 8 analogue monitoring pixels. These bypass the digital circuitry altogether and provide a direct access to the analogue output. This makes it possible to perform timing as well as gain measurements using radioactive sources. Fig. 3.29 was produced using a Sr-90 and Fe-55 source. The timing measurements reveal that the charge collection time and front-end timing are within the 25 ns requirement. The energy resolution performance shows some degradation after proton irradiation to 10^{15} n_{eq}, but still allows for the observation of the two distinct k_α and k_β peaks produced by the Fe-55 source.

Similarly, the gain can be estimated from the amplitude difference between these k_α and k_β peaks. This is performed by obtaining the ratio of the amplitude difference between both peaks and the difference between the k_α and k_β energies. The variation in distance between

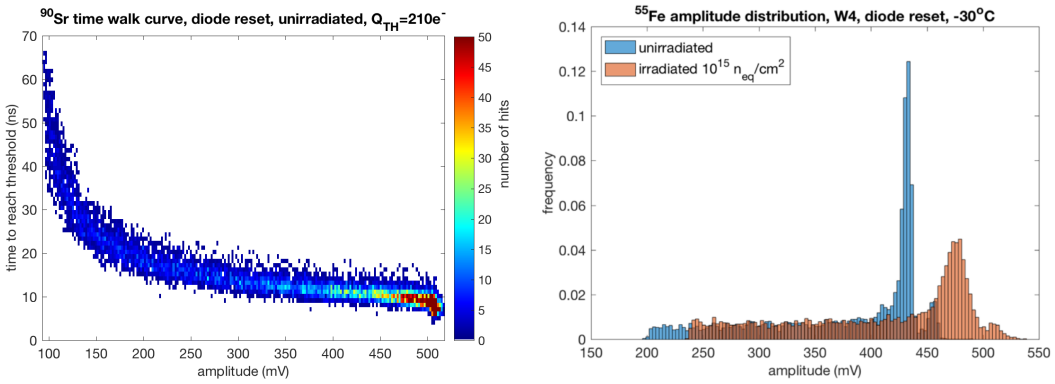


Figure 3.29: Sr-90 source measurements, showing timing results (left) and energy resolution performance of the sensor (right) before and after irradiation using Fe-55.

both these peaks with varying substrate voltage is shown on Fig. 3.30. The variations in these values as a function of substrate voltage are tabulated in Table 3.2 along with their respective measurement uncertainty of ±0.002 mV/e.

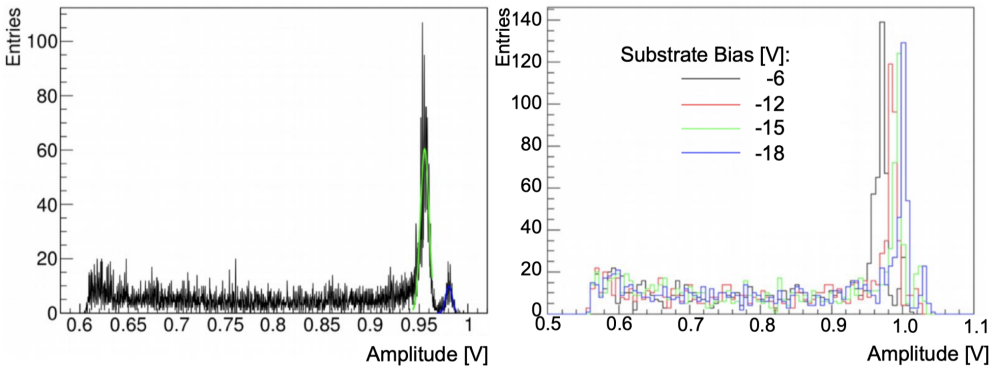


Figure 3.30: (Left) Fe-55 spectrum obtained with the MALTA chip. The fits to the k_α and k_β peaks are also shown. (Right) The variation in the amplitude difference between both peaks as a function of the substrate voltage.

SUB (V)	Gain±0.002 (mV/e)
-6	0.146
-12	0.148
-15	0.137
-18	0.136

Table 3.2: Gain measurements on an unirradiated MALTA sensor obtained using its response to Fe-55 for varying substrate voltages.

Another characterisation parameter assessed for this sensor is the threshold dispersion. This involves injecting a known amount of charge into a pixel, known as pixel pulsing, a given number of times and repeating this procedure for different injected charge amplitudes. The result is a turn-on curve (or S-curve), as is shown on Fig. 3.31. Determining its inflection point, being the point at which the pixel fires half of the time, corresponds to the threshold of that specific pixel. Repeating this process for a large number of pixels determines the

distribution of the individual pixel thresholds across the chip, known as the threshold dispersion. In the case of the MALTA sensor, this threshold dispersion had a design specification of 20 electrons but was measured to be twice as large, as shown in Fig. 3.31 (right).

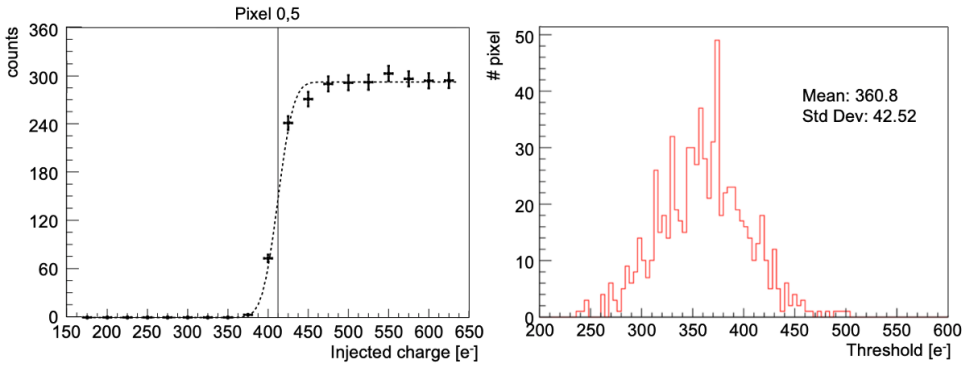


Figure 3.31: S-curve (turn-on curve) produced by pulsing pixels multiple times at varying threshold levels (left) along with the distribution of thresholds for over 700 pulsed pixels (right).

Threshold dispersion measurements were repeated on a 70 Mrad X-ray irradiated sample. After irradiation, a 50% increase in the threshold dispersion was observed. The noise, obtained by the slope of the threshold measurement S-curves (where a steeper slope would correspond to a lower noise level), is also observed to increase post X-ray irradiation. Fig. 3.32 shows the threshold dispersion measurement of an unirradiated sample (top-left), along with the increase in threshold dispersion by over 50% after 70 Mrad of irradiation (bottom left). The corresponding increased noise effect post irradiation is shown on the bottom right. Despite these significant changes after irradiation, the circuit was observed to be functional after a TID level of 70 Mrad.

3.1.5 mini-MALTA

As shown on Fig. 2.18 (right), the mini-MALTA sensor was designed with sectors forming its left half housing enlarged transistors. These enlarged transistors resulted in lower noise and as a result, the ability to reach lower thresholds. This can be observed on Fig. 3.33 for two sectors and Fig. 3.34 (right) showing the stark difference in threshold over the entire chip.

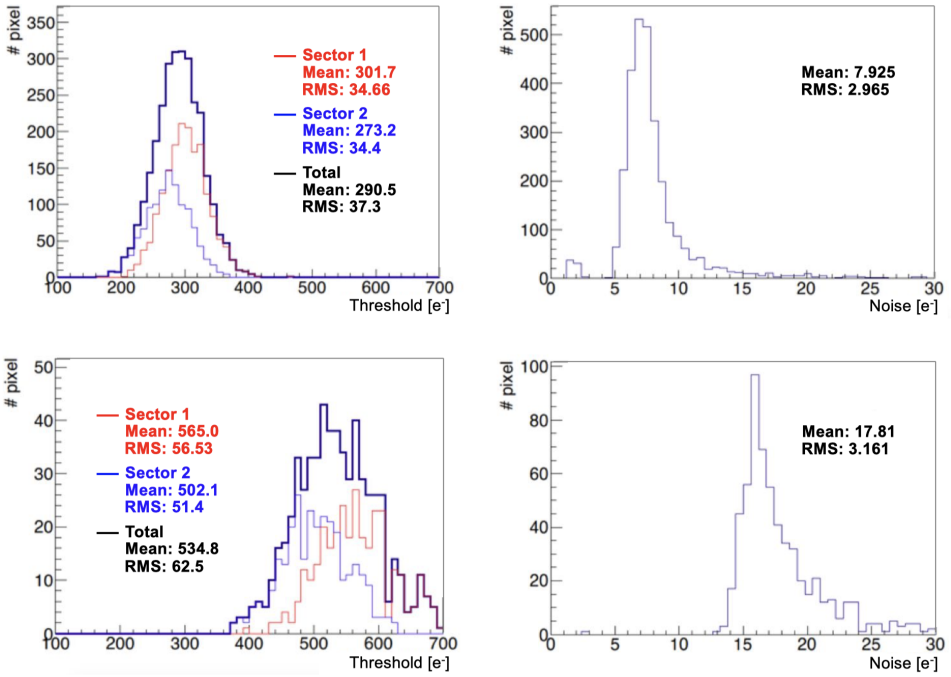


Figure 3.32: Threshold dispersion measurement (left) for sectors 1 (red) and 2 (blue) along with the corresponding noise distributions (right) for an unirradiated (top) and 70 Mrad x-ray irradiated MALTA sensor (bottom).

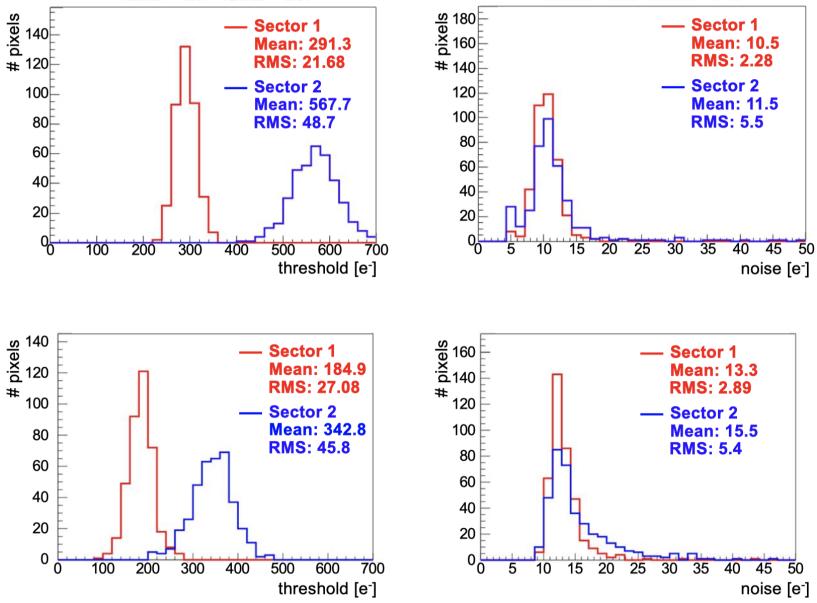


Figure 3.33: Threshold (left) and noise (right) measurements on an unirradiated (top) and 1×10^{15} n_{eq}/cm^2 neutron irradiated (bottom) mini-MALTA sample, showing sector 1 in red and sector 2 in blue.

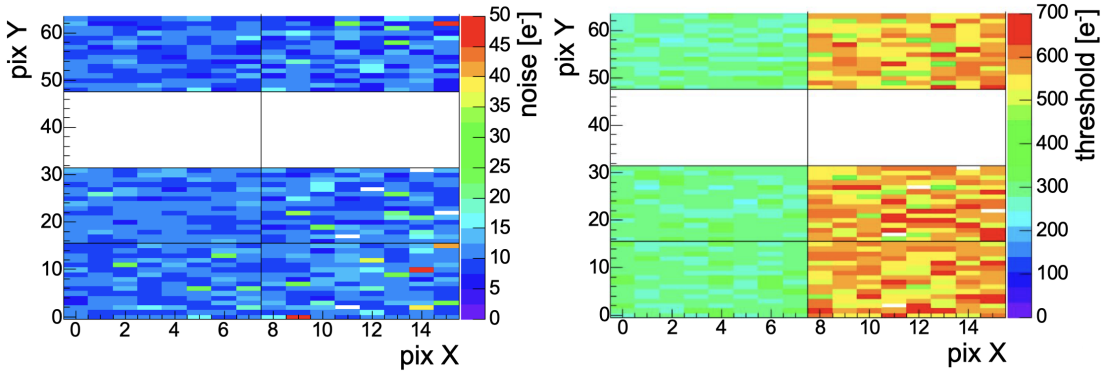


Figure 3.34: Noise and threshold measurements on an unirradiated mini-MALTA sample.

3.2 Testbeam Measurements

3.2.1 MALTA

A test-beam campaign was conducted on the MALTA chip at CERN's Super Proton Synchrotron (SPS) facility providing 180 GeV/c pions. The telescope assembly used to gather the necessary track reconstruction data comprised of 6 MIMOSA [54] planes with an overall resolution of under 2 μm [54]. Irradiated samples were kept at -20°C by a silicon-oil refrigeration system. The samples tested included unirradiated devices as well as ones that were irradiated to 5×10^{14} n_{eq}/cm^2, 1×10^{15} n_{eq}/cm^2 with neutrons and to 80 Mrad with X-rays.

The cluster size, or number of surrounding pixels that register a hit upon a particle track traversing the MALTA sensor can be obtained through track reconstruction conducted using the Proteus framework [55]. Tracks are reconstructed by fitting a straight line. The fit procedure minimises the distance between the track and the cluster centre across planes. The cluster size measurement was assessed for varying thresholds for both unirradiated as well as neutron irradiated (5×10^{14} n_{eq}/cm^2) chips. Fig. 3.35 compares the cluster size for sector 2 of the unirradiated and was as neutron irradiated sample for decreasing threshold levels. A clear increase in the cluster size can be seen when the threshold is decreased from approximately 600 to 250 electrons, as would be expected.

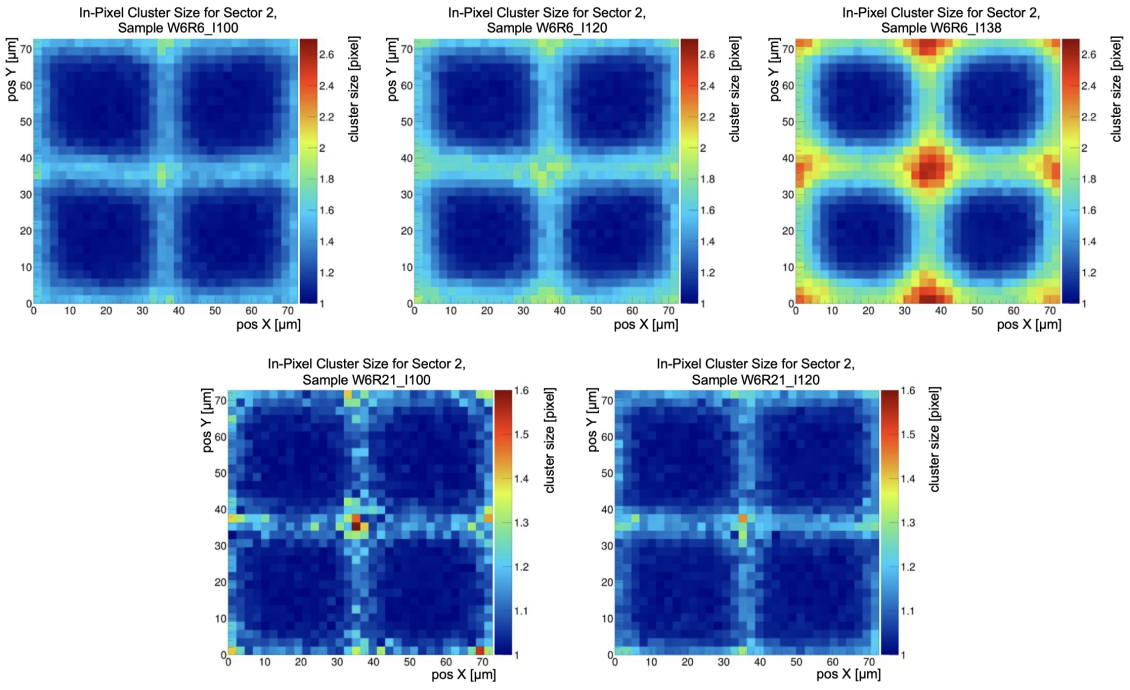


Figure 3.35: Variation in cluster size for an unirradiated MALTA sample (top row) and a neutron irradiated 5×10^{14} n_{eq}/cm^2 sample (bottom row) at decreasing thresholds, from approximately 600 to 250 electrons, from left to right. Note that the lowest threshold of 200 electrons could not be obtained for the irradiated sample.

The efficiency is calculated as the ratio between the number of hits on a pixel and the number of hits expected. Given the excellent 2 μm resolution of the telescope, in-pixel

efficiency could also be probed. This is shown on Fig. 3.36 where the corners of the pixels are revealed to be low efficiency regions after $1{\times}10^{15}$ n_{eq}/cm^2 irradiation.

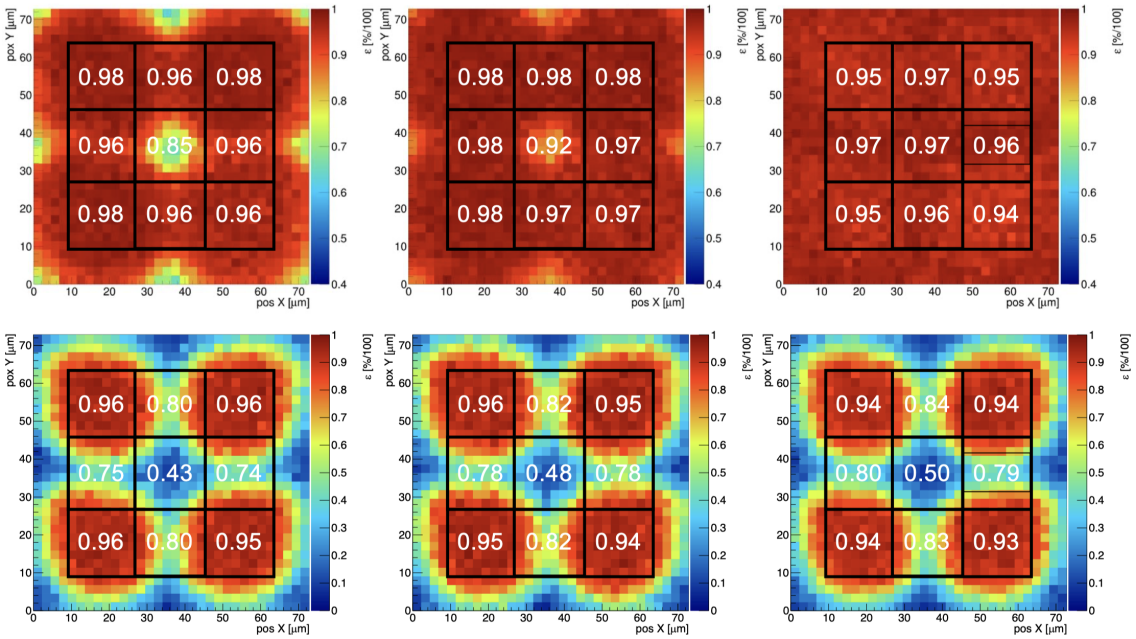


Figure 3.36: In-pixel efficiency on an unirradiated sample (top row) and a neutron irradiated 1×10^{15} n_{eq}/cm^2 sample (bottom row) at decreasing thresholds (approximately 600 to 250 electrons), showing an increase in inter-pixel efficiency as well as decrease pixel centre efficiency.

Though the efficiency at the centre of the pixel is expected to increase with decreasing threshold, the opposite effect is observed in these samples. This efficiency reduction at the pixel centre arises partly due to issues at the masking of noisy pixels and merging level in the cases of simultaneous hits. Additionally the pixel centres have a higher probability of collecting the total charge (above threshold) from energy deposition leading to these regions being more sensitive to noise hits. Similarly, inter-pixel regions have lesser sensitivity to noise hits as the likelihood is greater that the charge deposited between pixels will be collected by two or more pixels, reducing the chances of a noise hit causing the pixel hit to be classified as a fake hit. This behaviour is observed consistently across all samples, even the unirradiated ones, as shown on Fig. 3.36.

The shape of the deep p-well extension in the pixel design which is shown in Fig. 3.37 (right) has an effect on the contour of the efficiency as seen in Fig. 3.37 where sector 3 incorporates the so called "medium deep p-well" and sector 2 the "maximum deep p-well". Their respective layer cut-outs are also shown. This difference between p-well shapes becomes increasingly visible on irradiated chips at high threshold.

Figure 3.38 summarises the efficiency trends observed in the test-beam campaigns for the different sectors, as described in Fig. 2.15. The error of the efficiencies are obtained through binomial error calculation, $\Delta\epsilon = \sqrt{\epsilon(1-\epsilon)/n_{tracks}}$, where ϵ is the efficiency and n_{tracks} the number of tracks. In the non-irradiated case, the efficiency at the four pixel centres is always above 96%, and in the inter-pixel region varies from 80 to 95% with decreasing pixel

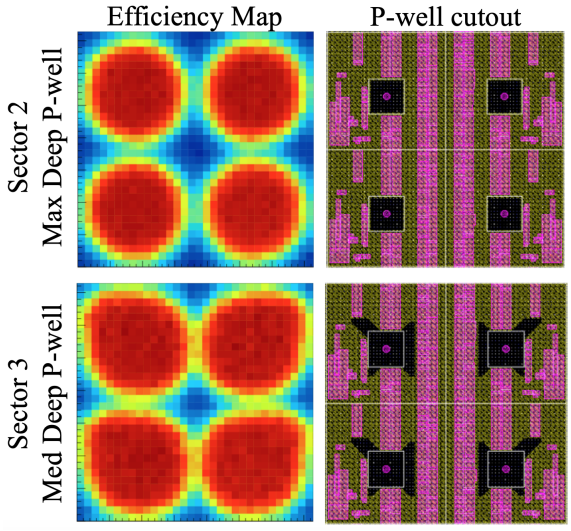


Figure 3.37: A view of the effect of the p-well cut-out shape (right) on the subsequent contours of the pixel efficiency maps (left), for sectors 2 (top) and 3 (bottom).

threshold. In the irradiated case, the efficiency in the inter-pixel region drops from 80 to 30% in the worst case, and raises up to 50% at the lowest threshold achievable of approximately 500 e^-.

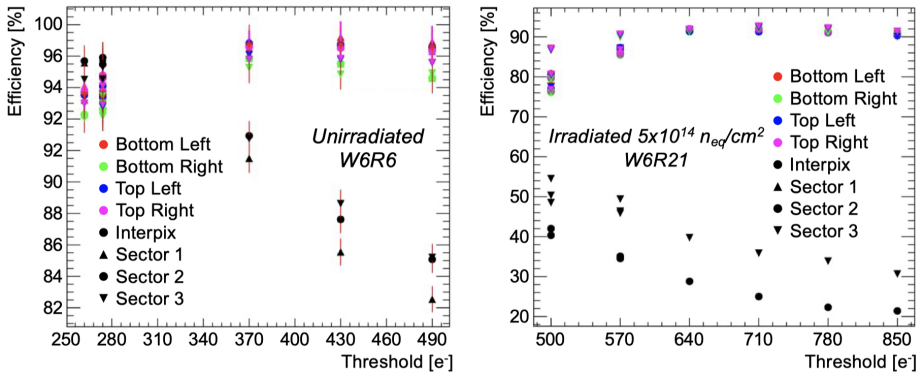


Figure 3.38: Efficiency variation as a function of threshold for an unirradiated (left) MALTA sample and after 5×10^{14} n_{eq}/cm^2 neutron irradiation (right).

The MALTA sensor is a full demonstrator monolithic CMOS sensor produced in 180 nm TowerJazz technology. It uses a novel asynchronous readout with no clock being distributed over the matrix to minimise power consumption. Its design integrates a small charge collection electrode for low capacitance and low analogue power consumption. Measurements have been performed on unirradiated and neutron irradiated samples (5×10^{14} $n_{eq}/cm^2$and 1×10^{15} n_{eq}/cm^2). This lead to the observation that as the merging logic is faulty and as a result permanently off, it causes some hits to be lost due to erroneous merging of those true hits with noise hits.

Analogue tests showed over 95% of hits being registered within the 25 ns in-time efficiency requirement. This was also measured at low noise levels.

Most issues observed in the test-beam results are explained by the degraded detection

efficiency, especially in the pixel corners due to masking and merging errors as well as the higher than expected threshold spread as compared to simulations. The next design submission of the chip should hence also include adjustment bits for greater threshold settings control.

All other issues that have been identified through this study should be addressed in the next iteration of the sensor design with very promising performance for the Phase 2 upgrade of ATLAS.

3.2.2 MALTA Telescope

A full-fledged telescope was developed which makes use of MALTA sensors for each of its telescope planes, shown on Fig. 3.39. It comprised of 6 MALTA chip-based planes, 6 FPGAs for their readout and centralised powering. The configuration of the planes is described in Fig. 3.40. Simulations yield a 3.9 μm resolution with 180 GeV/c pions, before applying any optimisations [56]. With an electron beam having a higher propensity for scattering than pions, a resolution of 13 μm was obtained. Fig. 3.41 shows the telescope used at the SPS and ELSA facilities, where in the latter, the telescope planes had to be brought closer to each other in order to compensate for the scattering effects predicted in an Allpix [57] simulation conducted for the telescope, and summarised in Table. 3.3.

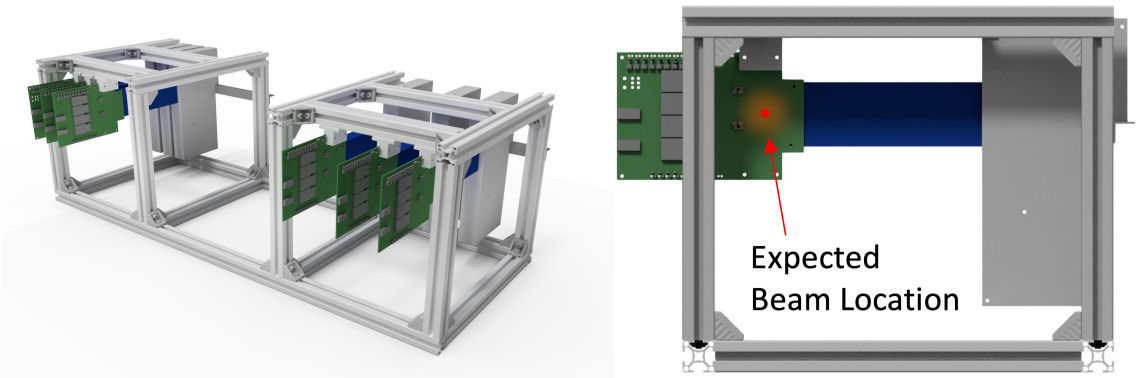


Figure 3.39: CAD representation of the MALTA telescope.

3.2.3 mini-MALTA

At the ELSA facility, a 2.5 GeV electron beam was provided. The MALTA Telescope was used and the tracks were reconstructed using the General Broken Line (GBL) [58] reconstruction algorithm. This accounts for the substantial scattering of electrons as they pass through the hundreds of microns of MALTA sensor silicon layers, and provides a greater resolution as a result. The expected number of hits are themselves interpolated from the planes from reconstructed tracks. Tracks are selected by applying a cut on the χ^2, defined in Eq. 3.3, where O_i are the observed values, E_i the expected values and i the number of entries.

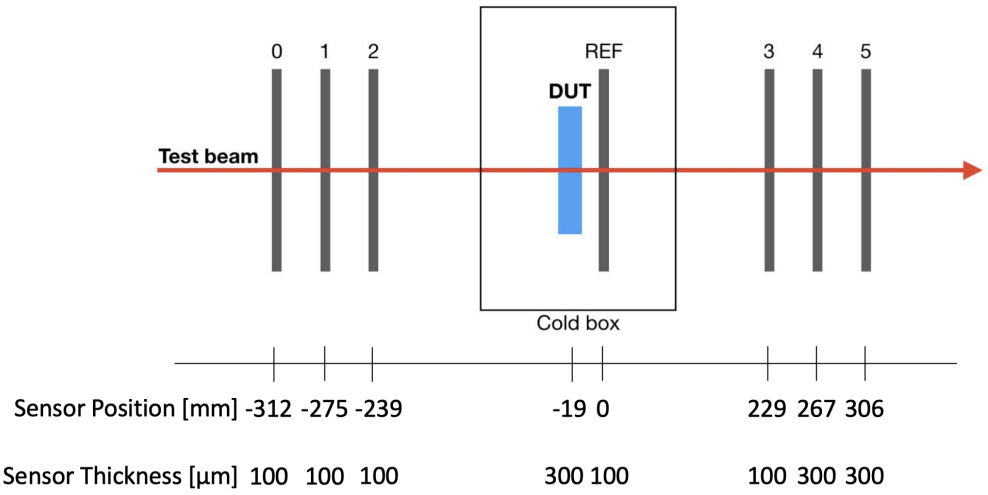


Figure 3.40: The plane configurations used to assemble the MALTA telescope along with the respective sensor thicknesses.

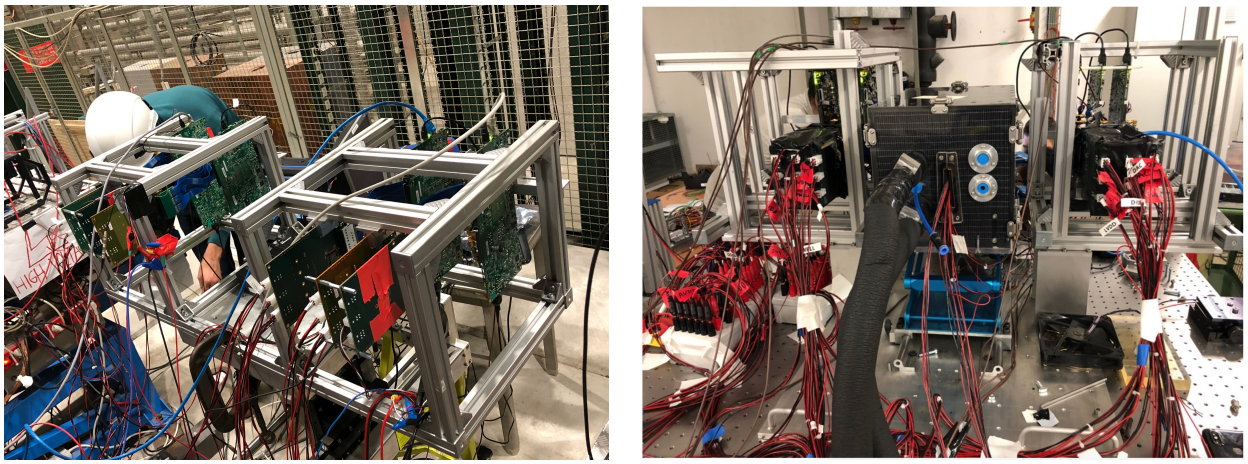


Figure 3.41: The MALTA telescope deployed at the SPS (left) and ELSA (right) test-beam facilities.

	Pions 180 GeV	Electrons 3 GeV		
Plane #	Distances [cm]	Distances [cm]	Distances [cm]	Distances [cm]
1	0	0	0	N/A
2	7	7	2	0
3	15	15	4	2
DUT	39	39	6	4
4	65	65	8	6
5	79	79	10	8
6	85	85	12	N/A
Residual	4 μm	87 μm	16 μm	17 μm

Table 3.3: Allpix [57] simulated residuals of the MATLA telescope plane configurations for 180 GeV pions and 3 GeV electrons.

$$\chi^2 = \sum_{n=1}^{i} \frac{(O_i - E_i)^2}{E_i} \tag{3.3}$$

In this case the χ^2 cut is applied where $\chi^2 < 5$. A fiducial region was created on the mini-MALTA which included the active area of each sector, discarding the periphery, and excluded a single pixel-wide area around noisy pixels that could not be masked. The tracks are interpolated into the mini-MALTA plane yielding an expected position of the hit on the chip. This expected position of the hit is matched with a cluster within a 100 μm radius of the remaining pixels. A cluster is defined as the number of immediately neighbouring pixels that all register a hit for a given event. The efficiency is then calculated as the ratio of the observed number of hits and the expected of number of hits, which is obtained for each bin of a size equivalent to the telescope resolution. Such efficiency plots are shown on Fig. 3.42 which shows a progressive deterioration in the efficiency post irradiation, but nevertheless maintaining a >95% efficiency in sector 1. In every case, the extra deep p-well sector yields the highest efficiency.

Fig. 3.43 shows the efficiency versus threshold on neutron irradiated mini-MALTA samples at 1×10^{15} 1 MeV n_{eq}/cm^2. All sectors housing enlarged transistors yield a clearly superior efficiency as well as the added ability to reach much lower thresholds (due to their lower noise). The additional pixel wafer substructures of an extra deep p-well and the n-gap feature also clearly produce an efficiency meeting the layer 4 CMOS specification of >97% after irradiation.

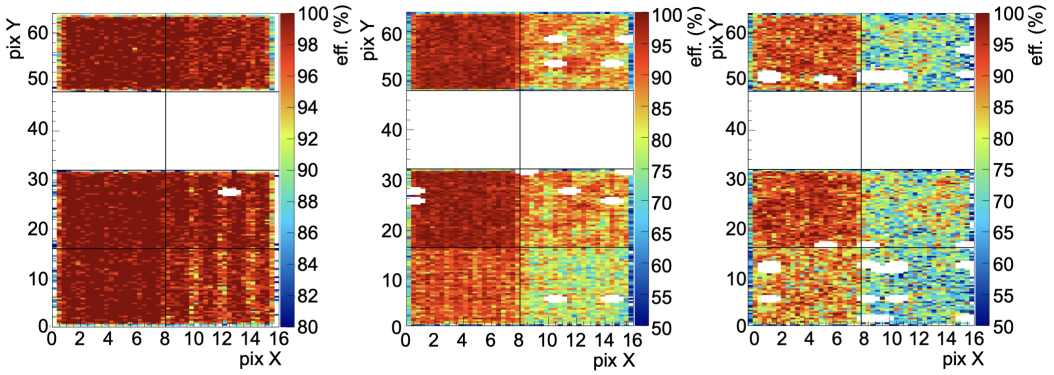


Figure 3.42: Efficiency plots for an unirradiated (left), 1×10^{15} 1 MeV n_{eq}/cm^2 (centre) and 2×10^{15} 1 MeV n_{eq}/cm^2 proton irradiated samples.

To obtain very high (<2 μm) resolution pixel response maps of the different sectors of the mini-MALTA chip, a test-beam campaign at the Diamond Light Source Facility was undertaken [59], shown on Fig. 3.44. This involved mounting a mini-MALTA sample on a X-Y-Z linear stage with 500 nm position precision, perpendicular to a (X-ray) photon beam of 8 keV with a beam waist at the focal distance of 2 μm. 8 keV was chosen as it is closest to the expected energy deposition of a MIP in the mini-MALTA substrate with 25 μm epi-taxial thickness. Hits were collected over a time window of 1 second, and the stage moved every 3 seconds by 2 μm, which resulted in 50 M hits per scan.

The number of hits at each position of the beam for each pixel were measured and represent a map of adjacent pixels with the sum of all the hits at each position. The relative response of each pixel could hence be observed, as shown on Fig. 3.45. Variations in the hit entries from one pixel to the next arise from the non-uniform threshold set on all pixels. Variations within a pixel, though arise from non-uniformities within the electric field map.

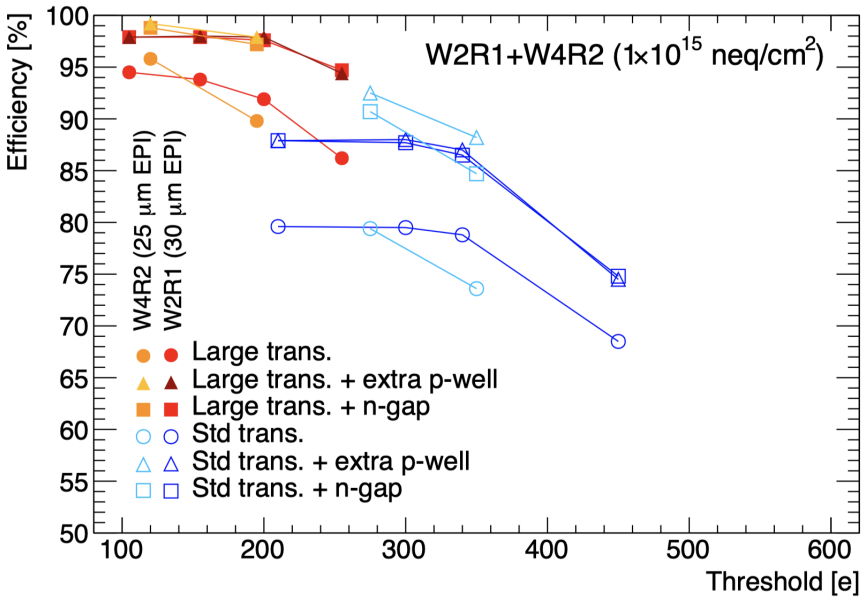


Figure 3.43: Efficiency versus threshold on neutron irradiated mini-MALTA samples at 1×10^{15} 1 MeV n_{eq}/cm^2. Different sensor regions are presented: standard MALTA-like (circles), modified with extra deep p-well (triangles) and modified with extra n-gap layer (rectangles). Results are also shown for sensor regions with standard (open markers) and enlarged (full markers) transistors, as well as for sensors with different epitaxial layer thicknesses: 25 μm (orange or light blue symbols) and 30 μm (red or dark blue symbols).

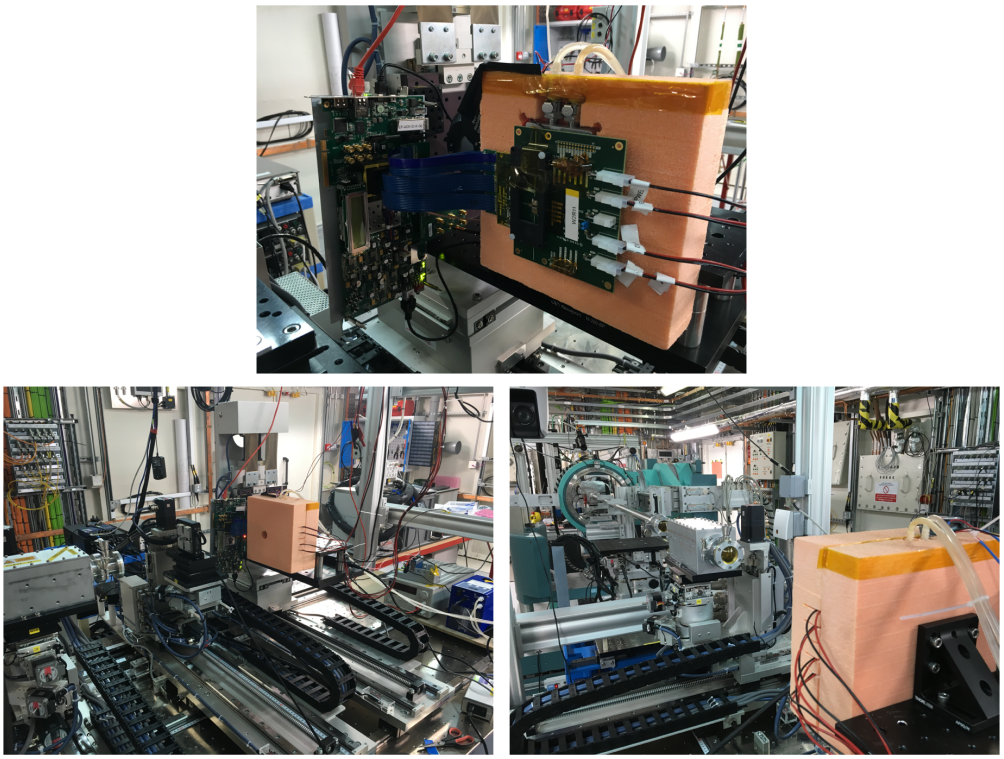


Figure 3.44: mini-MALTA chip mounted onto beam-line 16 support structures during test-beam campaign.

New design iterations of this sensor are expected to include threshold adjustment bits to allow the threshold configuration compensating for dispersions across the pixels as opposed to setting a global threshold.

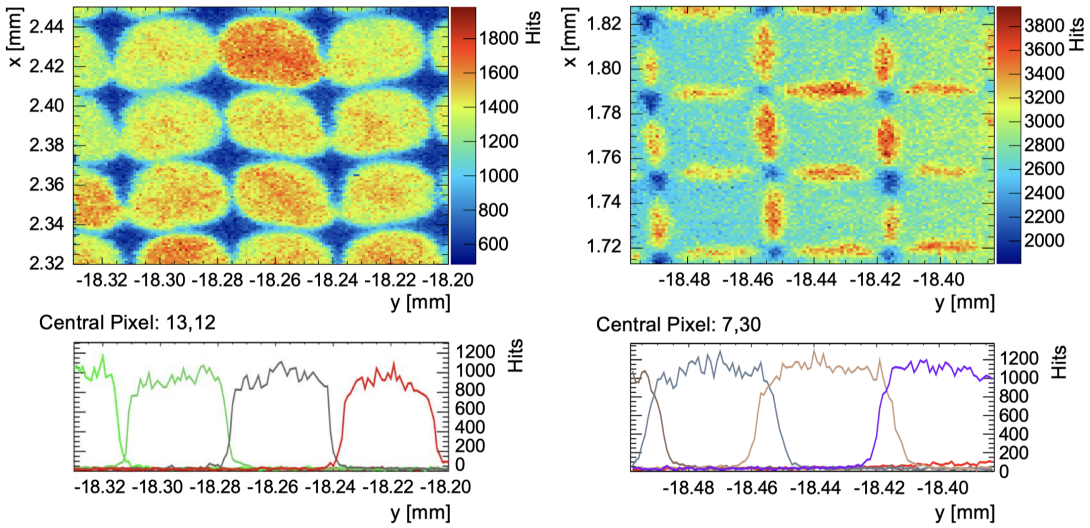


Figure 3.45: Pixel response maps through the measurement of the observed hits in a 1 second window of irradiation of 8 keV photons on the MALTA-like sector (left) and the extra deep p-well sector of the mini-MALTA (right).

Further analysis was conducted to identify the relative pixel response to the X-ray photons. In order achieve this the hit map data shown on Fig. 3.45 was reprocessed with the following procedure. As shown on Fig. 3.46, a Gaussian filter was applied to the image and the in-pixel coordinates with the largest value were identified. In a small area around this point, the average hit number was calculated, as shown by the green square. This average was used to normalise the hits in all neighbouring whole pixels, yielding the removal of pixel to pixel variations. The physical size of the pixels being 36.4×36.4 μm^2, the average normalised number of hits was calculated within this area shown as a red square in Fig. 3.46. The relative photon pixel response is defined as this average value. The pixel response uncertainty is in turn defined as the standard deviation of the number of hits in the green square forming the normalisation area.

In terms of this relative pixel response to the X-ray photons, the outline of the p-well cut-out is clearly observed as well the double column symmetry structure of the chip. For the pixel response as a function of dose rate the continuous n$^-$ layer showed near complete charge collection at pixel centres but poor charge collection in some pixel corners. The charge collection was degraded by over 10% after neutron irradiation to 1×10^{15} n$_{eq}$/cm^2. The extra deep p-well and n-gap sectors on the other hand displayed a improved performance with the full pixel area collecting charge uniformly before irradiation, even in pixel corners, and retained the full charge collection capability even after irradiation. In line with the pixel response, the charge sharing was observed to decrease for the continuous n$^-$ layer after irradiation as expected, given the further deteriorated charge collection efficiency at the pixel corners. The extra deep p-well and n-gap sectors showed no change in charge sharing as a function of dose rate. For the pixel response as a function of biasing voltage, it was observed that the pixel response reduces in the case where the sample is biased as high as 20 V, instead

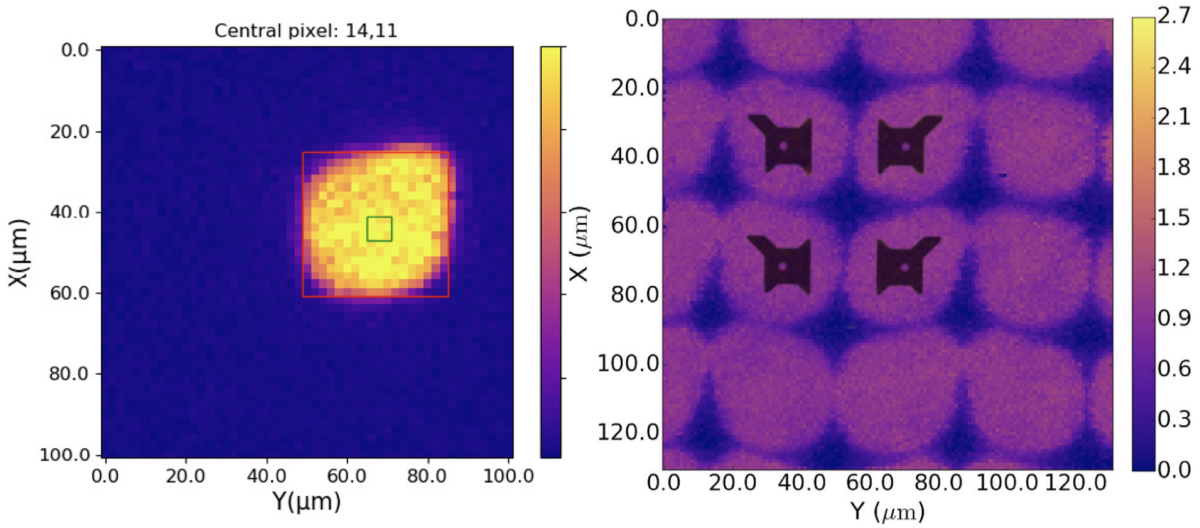


Figure 3.46: Relative photon response of a single pixel, showing the pixel centre area in green to which hits are normalised and in red the physical area of the pixel (left). Pixel response of several pixels with the deep p-well cutout overlayed on 4 pixels (right).

of the nominal 6 V. Through corroborating TCAD simulations [60], it was identified that the higher biasing voltage leads to a longer drift path for the charges in the pixel corners below the deep p-well. As these charges drift vertically along the substrate until reaching the deep p-well itself, at which point the drift direction is effectively horizontal and much slower as the E-field strength is independent of the biasing in that direction, leading to an increased likelihood for the charges to get trapped. This was observed across all sectors of the unirradiated sample, with an approximately 5% reduction in pixel response when biasing from 6 V to 20 V.

3.3 Conclusions

The latest generation of the MALTA family of pixel detectors designed in the CMOS 180 nm technology by TowerJazz shows full efficiency after 1×10^{15} n_{eq}/cm^2. General trends observed include that the 25 μm epitaxial layer samples were more efficient than 30 μm ones. The charge collection efficiency's uniformity had variations of less than 5% across the pixels. The mini-MALTA chip could be operated at a substrate bias of up to 20 V before breakdown, which is useful to compensate for efficiency losses post-irradiation. The extra deep p-well and n-gap processing in the pixel designs were validated as effective substructures to correct for pixel corner charge collection efficiency losses. Timing measurements for MALTA are expected to be conducted over the course of 2020.

A new version of the chip, named MALTA 2, will be produced in 2020. Its dimensions will be constrained by a smaller area available on the wafer reticle and will thus be approximately half the size of the current MALTA (2×1 cm^2). This lack of area will also lead to compromises in the layout of the next version. As the PMOS reset was shown to provide no appreciable advantage, only the diode reset will be implemented, albeit with two distinct sectors. The first will feature a FE with enlarged transistors, larger even than those present in mini-MALTA, and the second with what is known as a cascoded version of the FE, which involves increasing the resistance of the FE transistors. Both these features are expected to provide a higher gain and hence the ability to reach yet lower thresholds. As with its predecessor, MALTA 2 will also incorporate chip-to-chip communication capability. Due to the reduced area, the previously two sets of 40 pads on either side of the chip available for this feature will be reduced to only one set, resulting in reduced chip-to-chip orientation options, but retaining the full communication features. MALTA 2 will be produced in various process flavours, where each will also comprise of a variation of pixel geometries and processes, varying the extra deep p-well, n-gap and continuous n-minus layer masks. It will also feature a redesigned slow-control in an attempt to reduce threshold scan times as well as optimise chip configuration.

This technology is of special interest to HEP experiments due to its reduced requirements in power consumption, cooling and material budget. This technology is part of the Strategic R&D Programme on Technologies for Future Experiments of CERN's EP department [61].

Chapter 4

Contributions to the ITk Outer Barrel Demonstrator Programme

In the context of the ATLAS Phase 2 upgrade towards the HL-LHC, beyond the sensor technology development covered in Chapter 3, is the need for a thorough characterisation of the supporting structures holding these sensors in place. These structures also provide support for the many sub-detector services such as cooling, powering and data cabling. As a result, they have to undergo extensive R&D to ensure that they are radiation hard, mechanically robust enough to endure the stresses of supporting the sensors and services as well as displaying beneficial thermal characteristics allowing for the efficient cooling of the sensors. Section 4.1 delves into the systematic thermal characterisation measurements developed and performed in order to assess the viability of mechanical support designs.

A major stage in the R&D towards the new sub-detector involves the construction of a fully functional set of demonstrator 1:1 scale models that mimic the production and assembly techniques foreseen to be used for the actual ITk detector. A key element is the cable saver board, which serves in connecting the delicate flex cables originating from the sensor towards the more robust instrumentation cabling. Section 4.2 describes the design and production effort towards this R&D for the ITk.

In addition, readout system developments have also been necessary to cope with the much larger number of channels and hit rates expected for the ITk. Section 4.3 summarises the contribution made in the qualification of the FELIX readout system and ITk software integration.

4.1 Thermal Figure of Merit measurements for Prototype Staves

Staves are support structures, analogous to scaffoldings, which hold in place the sensor modules, readout and power cables as well as cooling tubes. The assessment of a stave design rests on three key inputs: finite element analysis simulations, mechanical stress tests as well as the characterisation of its thermal properties. This section focuses on thermal characteri-

sations of numerous candidate stave prototype designs, submitted by an international group of over 10 ATLAS institutes.

4.1.1 The TFM Concept

The necessity for qualifying the thermal behaviour of a stave prototype arises from the fact that one of its intended purposes is to keep the silicon sensors cold. This is a particularly important factor for silicon sensors as over the course of irradiation their leakage current increases. This leakage current is in turn proportional to temperature. It can be approximated that the leakage current doubles every 7°C. As a result, the power dissipated by silicon sensors is equivalent to $P = IV$, where I is the leakage current and V is the bias voltage. It is therefore imperative to ensure that sensors remain cold in order to diminish the likelihood of thermal runaway by dissipating the heat produced by the sensors. Keeping the sensor cool also provides the added benefit of reducing thermally induced noise as well as preventing the development of secondary defects after long periods of annealing, known as reverse annealing

The key thermal parameter sought after to qualify each stave sample is known as the Thermal Figure of Merit (TFM). The nature of this parameter can be explained, in its simplest form, using Fig. 4.1 where a block of material has its top surface heated at 2 W reaching 30°C and bottom surface cooled to 0°C. The temperature experienced by the top surface will depend on the effectiveness with which the 2 W heat is dissipated downwards through the block. This will in turn depend on its geometry and thermal properties. The TFM in this case is defined as:

$$\mathrm{TFM} = \frac{T_{hot} - T_{cold}}{Power/_{area}} \tag{4.1}$$

Where T_{hot} is the temperature of the hot surface (30°C), T_{cold} is the cold temperature directly under the probed location on the hot surface (0°C) and the heat flux, power/unit area, is in this case 2 W/cm^2.

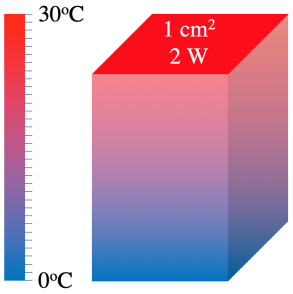


Figure 4.1: Schematic of a block of material with heated top surface and cooled bottom surface, displaying a temperature gradient inversely proportional to its TFM, as measured on the top surface.

In the context of stave characterisations, this can be represented by Fig. 4.2, showing an example of an actual sample known as an I-Beam stave, described in detail in section 4.1.7.

Here the sample has a CO_2 coolant-filled pipe passing through the entirety of its length and the surface intended to have sensors glued onto it is attached with 350 μm thick silicon tiles to simulate the operation of a silicon sensor. Finally, Negative Temperature Coefficient (NTC) thermistors are placed along the surface of the silicon tiles (that act as heaters) as well as at the inlet and outlet of the CO_2 pipe. Computing the TFM for a given location on the stave sample in this scenario is done in two steps. First, T_{hot} is provided by the corresponding NTC. T_{cold} is then obtained by linearly interpolating between the CO_2 pipe inlet and outlet NTCs and identifying the value along the pipe length matching the position of the selected NTC. The power supplied to the silicon heaters is transferred via copper braids attached on one end to the heaters and on the other directly to a power supply. The area in the TFM Eq. 4.1 corresponds to the total area of the powered silicon heaters.

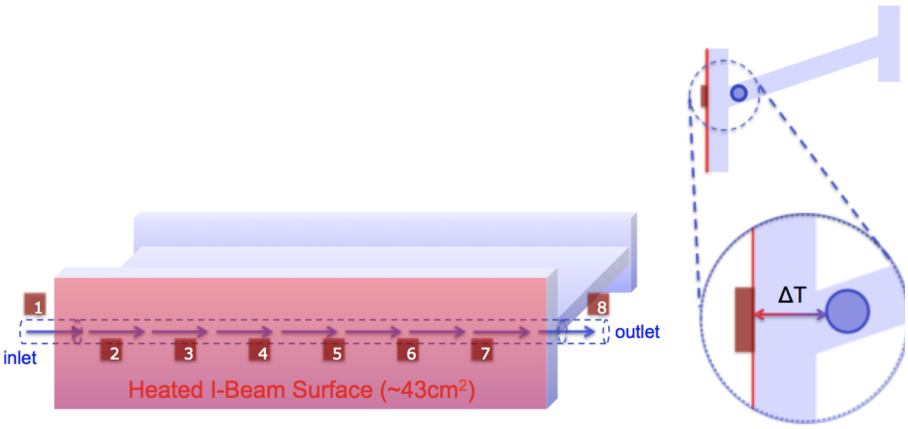


Figure 4.2: Diagram of the TFM concept in the context of a stave support cooled via an internal pipe.

The TFM hence provides a quantitative value for the ease with which heat can be transported away from the sensors towards the cooling fluid. Fig. 4.3 shows an image of the I-Beam stave sample installed in the setup.

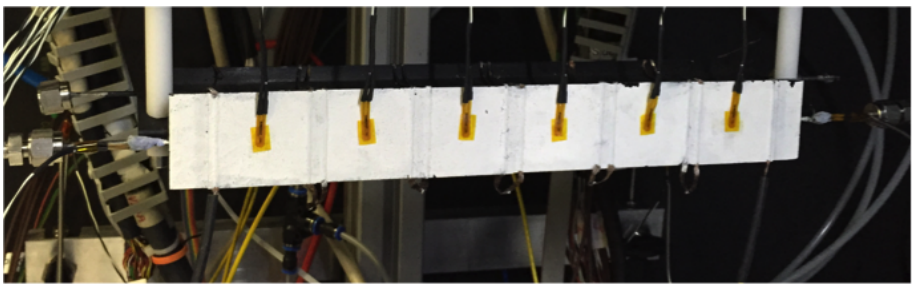


Figure 4.3: Image of the version 1 Berkeley I-Beam sample (wide side view) with 6 quad heater cells attached, each with an NTC glued at its centre.

Fig. 4.4 shows the expected TFM curves to be produced upon supplying an incrementally higher power across the heated surface of a given stave. The significantly higher TFM values between 0.1 Wcm^{-2} to around 0.5 Wcm^{-2} are due to heat pick-up from the ambient air. As the heat flux is increased, the contribution of the radiation surpasses that of convection and as a result, a more representative value emerges.

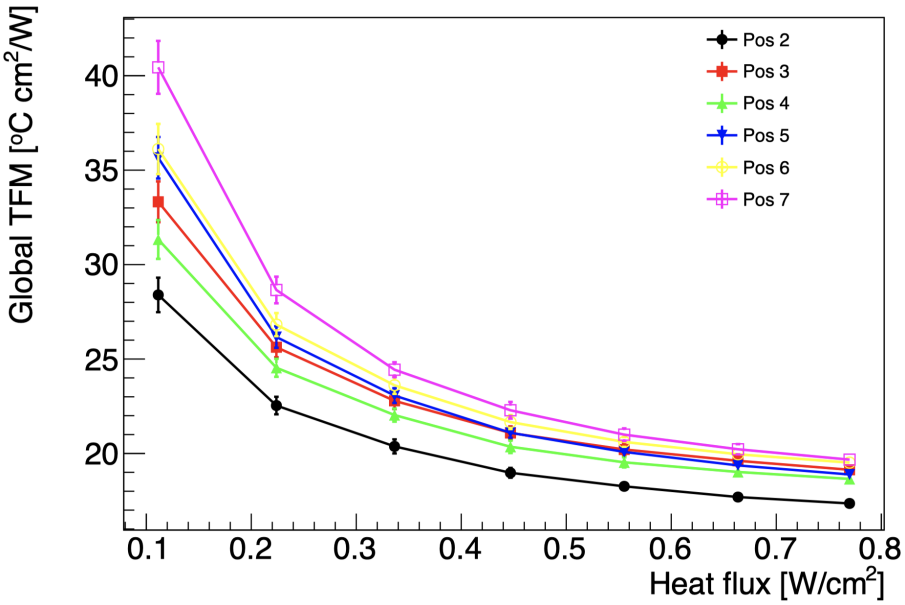


Figure 4.4: A sample TFM result showing the different curves produced by individual NTC readings at varying heat fluxes.

4.1.2 TFM Setup Description

The different hardware components that form the experimental setup are shown in Fig. 4.5. Each component is described following the direction of the CO_2 flow. The entire measurement system resides within an accessible aluminium enclosure containing all hardware components and the device under test (DUT). This enclosure's internal walls are fully padded with 10 cm thick ArmaFlex [62] to secure hermeticity and thermal insulation. The only component located outside this chamber is the TRACI v.1 [63] evaporative CO_2 cooling recycling system. This device provides the necessary CO_2 in its bi-phase state to provide temperatures as low as $-30°C$. For practical purposes, a 3-point valve is situated at the entrance of the climate chamber in order to prevent the CO_2 in the full cooling circuit being evacuated during the disassembly or replacement of stave samples. Following the valve is a coriolis mass flowmeter [64] which provides values through the measurement of oscillatory vibrations generated by fluid flow via a U-shaped pipe crossing the flowmeter. These vibration are proportional to the mass flow crossing the device. At this stage, a restriction valve is located which can be used to adjust the aperture of the pipe at that location, having the resulting effect of altering the pressure of the CO_2 . A cartridge heater [65] is then installed which can supply heat into the CO_2 to ensure that it is in a boiling state prior to entering the DUT. Keeping the CO_2 in a bi-phase state is critical to its cooling performance. Finally, the stave sample under test is then placed to undergo testing, detailed in section 4.1.7. Additional hardware components include Pressure-Temperature (PT100) thermocouples [66] at the inlet and outlet of the stave sample. These probes are installed so as to make a direct contact temperature measurement of the CO_2 . Section 4.1.4 describes in greater detail the custom temperature readout developed for this setup. And lastly, a triaxial pipe, composed of three concentric pipes with insulation separating the inner most and outer most pipes, is installed at the inlet and outlet of the CO_2 pipes in the immediate vicinity of the TRACI cooling

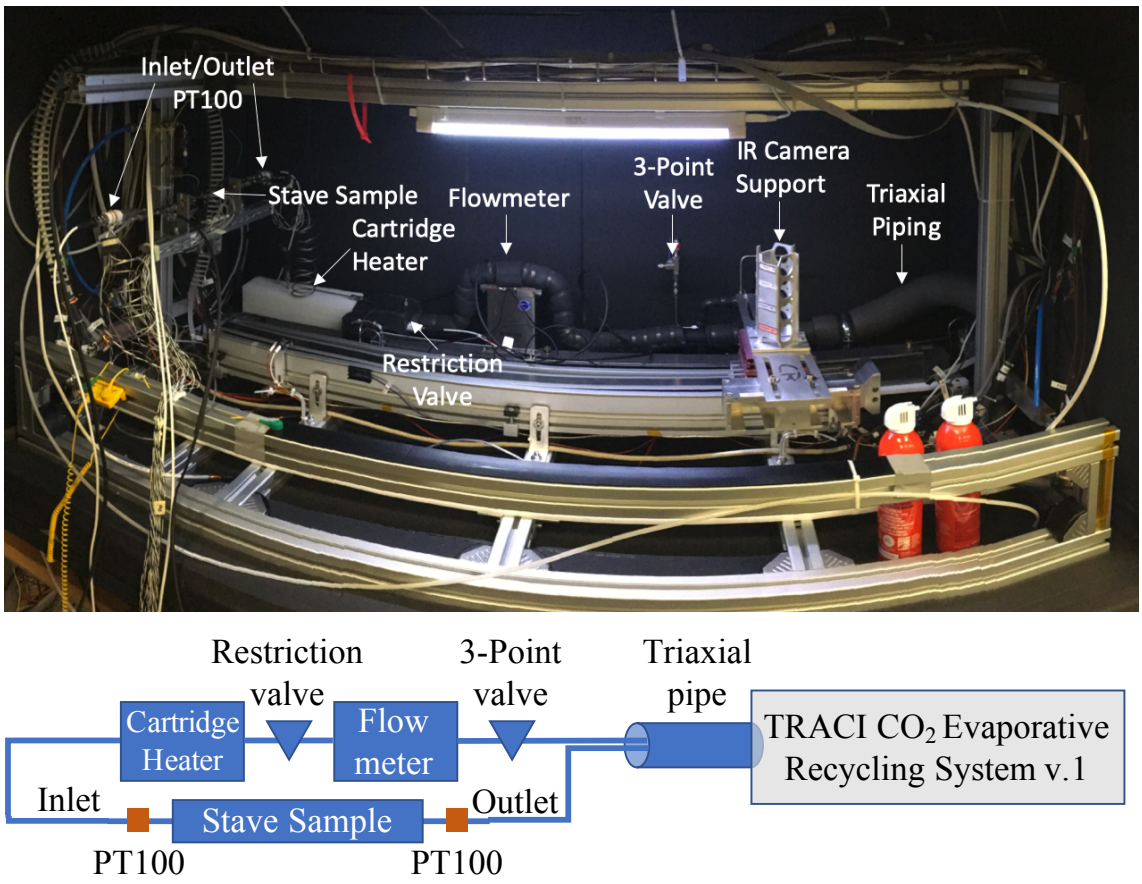


Figure 4.5: Image of the TFM measurement apparatus comprising of: a CO_2 cooling unit (TRACI), triaxial pipe, 3-point valve, a flowmeter, restriction valve, a cartridge heater and PT100s at inlet and outlet relative to the stave sample. All of these components reside inside a 2 m^3 insulated aluminium enclosure. The bottom diagram shows a schematic of the CO_2 cooling circuit.

system. This provides a significant protection against thermal pickup.

4.1.3 NTC Calibration

The NTCs used in this study underwent a calibration procedure whereby 8 probes were secured onto a copper block along with a reference PT100 sensor as shown on Fig. 4.6. The copper block was in turn submerged inside a glycol bath initially cooled to $-25°C$ and allowed to reach room temperature over time. The resistance evolution of all 9 probes was monitored over this duration, and the NTCs yielded a measured β value, shown on Fig. 4.7 of 4192, differing from its data sheet value of 4092. Here, the β value is defined as follows:

$$\beta = \frac{ln\left(\frac{R_{T1}}{R_{T2}}\right)}{\frac{1}{T_1} - \frac{1}{T_2}} \tag{4.2}$$

where R_{T1} = Resistance at Temperature 1, R_{T2} = Resistance at Temperature 2, T1 =

Temperature 1 (^{o}K) and T2 = Temperature 2 (^{o}K). β is the parameter that represents the behaviour of a negative thermistor or NTC. An estimate of the impact of a 1% error in this value was identified as being < 0.6°C.



Figure 4.6: 8 NTCs and a reference PT100 sensor secured within a copper block using thermal paste before submersion inside a glycol bath.

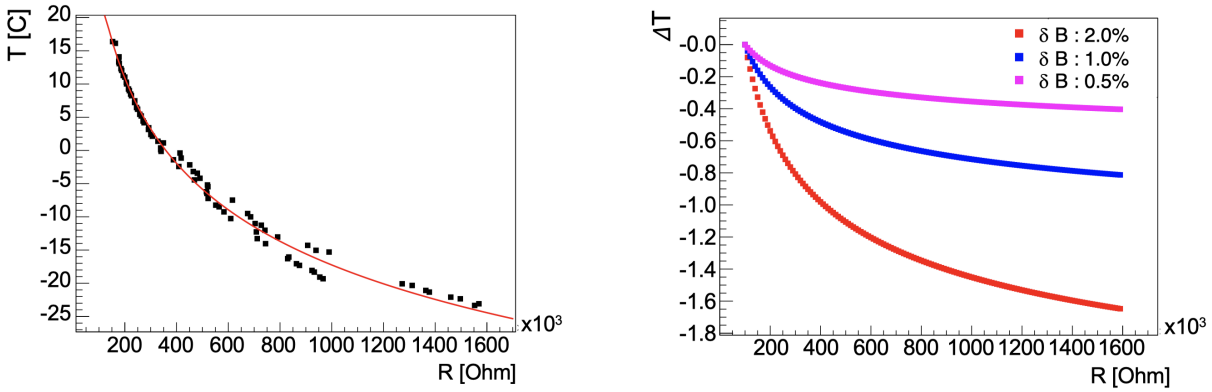


Figure 4.7: β value extracted from the resistance measurements of 8 NTCs (left) along with a projection of the impact of a 1% error in this estimate being < 0.6°C (right).

4.1.4 The VolTRON Temperature Readout System

The dimensions of the stave samples tested varied from single heater cells to upwards of 12 cells. Given that around five NTC sensors per cell were needed to obtain meaningful temperature measurements over the surface area of each cell, this required the development of a custom NTC readout system to cope with all expected channels. This lead to the creation of the VolTRON readout system.

The VolTRON (Voltage to Temperature ReadOut iNstrument) is a temperature and voltage readout device for the thermal characterisation of ITk Stave prototypes. It has a

small footprint of 125×125×40 mm^3, with 16 input channels and a USB control interface, that also serves as the 5 V power source. VolTRON measures the resistance in the range of 2 kΩ to 2 MΩ of the input channels and transmits the values with a configurable precision and latency through the USB. The software, written in Python, allows monitoring and read-out of the data with a sampling rate of 1 Hz or lower, with conversion into temperature in the range of -25°C to 25°C. A picture of VolTRON is shown in Fig. 4.8.

VolTRON was developed to address the need for a replacement of the obsolete resistivity measurement system, the Cirris Touch1 [67], previously inherited for this task. It provides measurements with equal or better performance, and brings several advantages such as a user-friendly programming language, small footprint and cost effectiveness.

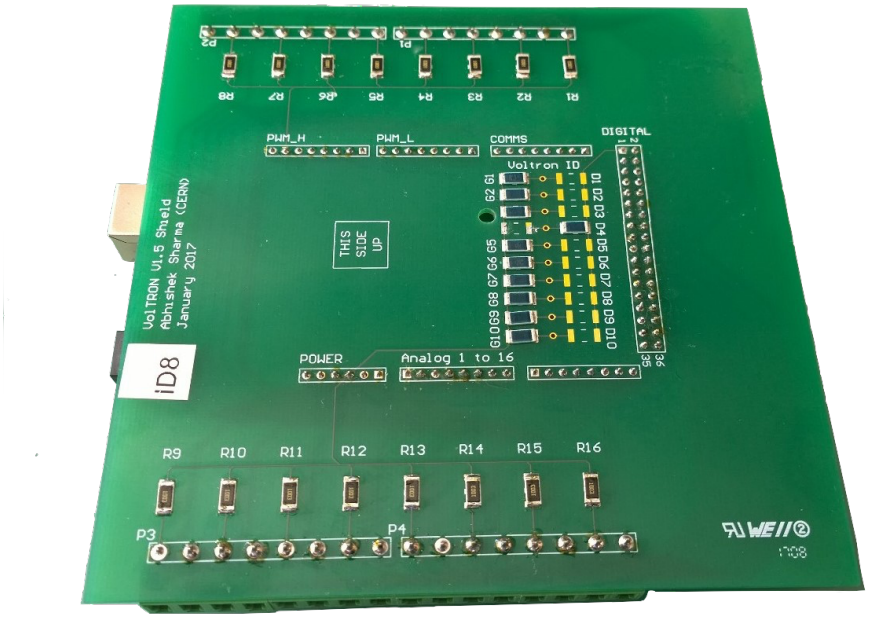


Figure 4.8: Picture of VolTRON PCB Shield Version 1.5.

The Arduino Mega [40] board, see Fig. 4.9, is the motherboard of VolTRON. It is based on the ATmega1280 micro-controller and comprises of 16 analogue input pins, 54 digital input and output pins, a 16 MHz clock and a USB connection. Each analogue input is connected to a 10-bit ADC with very high impedance (∼ MΩ), which provides very little uncertainty on the voltage measurement. Each of the read-out channels is independently powered and read-out during the measurement to avoid sensor heating due to current flow. The input reference voltage used for the measurement is 5 V, and the reference voltage for the analogue input can be 2.56 V as well as 5 V. Both voltages have an uncertainty of 0.1%. The minimum uncertainty associated with the voltage reading is thus 0.1%.

A custom designed VolTRON shield PCB provides the input connectors and cable routing to the motherboard, along with reference 100 ± 0.1 kΩ resistors and board identification resistors. Each channel of the VolTRON is equipped with a voltage divider, which is a passive linear circuit that produces an output voltage that is a fraction of its input voltage. An example of a voltage divider, as seen in Fig. 4.10, is two resistors connected in series, with

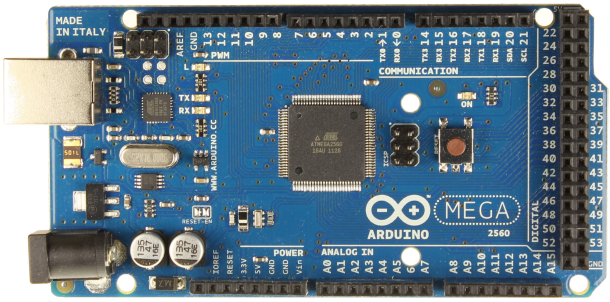


Figure 4.9: Arduino Mega 2560 revision 3 board [40].

the input voltage applied across the two resistors. The output voltage then emanates from between both resistors. The voltage output is proportional to the input voltage as shown in Eq. 4.3 where each readout channel has the reference resistor placed in position $R2$, and the NTC in position $R1$.

$$V_{out} = V_{in} \cdot \left(\frac{R_2}{R_2 + R_1} \right) \tag{4.3}$$

Each VolTRON unit is placed in a support structure shown on Fig. 4.11 that provides stability and eases the cable routing.

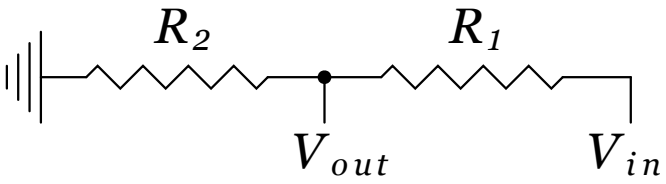


Figure 4.10: Circuit diagram of a resistive voltage divider.



Figure 4.11: VolTRON tower accommodating up to 7 readout units in parallel, totalling 112 NTC channels.

As discussed earlier, the reference resistance and the NTC have both a value of 100 kΩ at 25°C. Therefore, using Eq. 4.3, if the input voltage is 5 V, the voltage at the output of the voltage divider at 25°C is 2.5 V.

In order to increase the precision of the measurements below 25°C, the reference voltage for the ADC can be set to 2.56 V. The smallest temperature that can be measured by the 10-bit ADC is for a resistance of the NTC that is 1022 times bigger than the reference resistance, approximately 100 MΩ, which corresponds to temperatures below −50°C, which is shown in the following:

$$V\left(R_1 = 1022 \cdot R_2\right) = \frac{R_2}{R_2 + 1022 \cdot R_2} V_{in} = \frac{1}{1023} V_{in}$$

However, the typical resistance for the selected NTCs at −25°Cis around 2 MΩ, which means that the minimum voltage is approximately 0.23 V, as shown in the following:

$$V\left(T = -25°\text{C}, R_1 = 20 \cdot R_2\right) = \frac{R_2}{R_2 + 20 \cdot R_2} V_{in} = \frac{1}{21} V_{in}$$

Thus a reference voltage of 2.56 V would provide a higher voltage measurement resolution.

The Arduino firmware is programmed in a proprietary language that resembles Java, and the control software is written in Python. At power up the Arduino board enters a set-up routine followed by an infinite loop of a measurement routine. Individual channels can be enabled or disabled depending on the needs of the given sample being measured. Each unit has a hard-coded identifier encoded with a binary address using a sequence of 0 Ω resistors, allowing the parallelisation of the units. The control software in Python is based on pyserial [68] and ROOT [69] libraries. A class definition has been developed to interface each unit and a command line application connects to the serial port and displays real-time a ROOT-based plot of the temperature as a function of time as shown in Fig. 4.12.

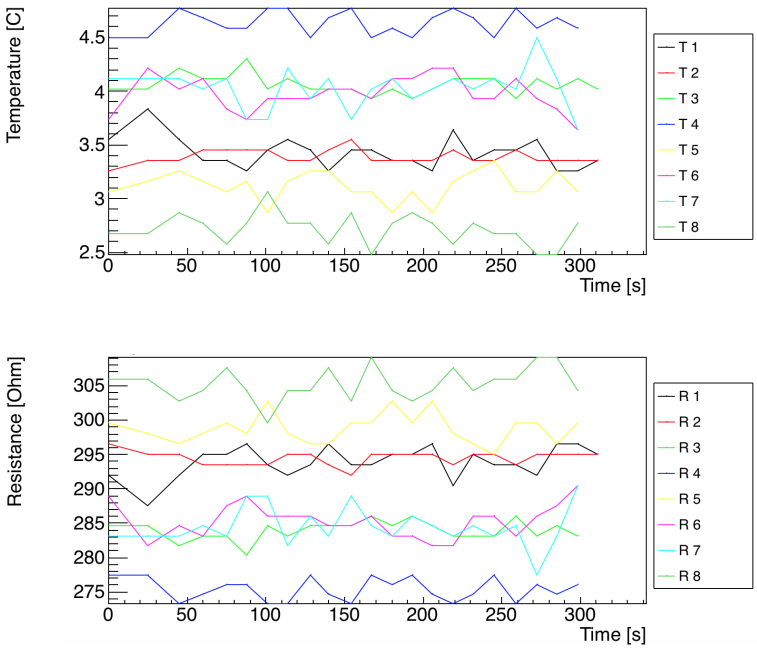


Figure 4.12: Temperature and resistance plot produced in realtime during the measurements by the control software.

The input channels have to be inter-calibrated for each board to provide better precision measurements. Three high precision resistors of values, 100 kΩ, 500 kΩ and 1 MΩ, with a precision of 0.1% were placed in each of the channels and the response of the system was then measured. Fig. 4.13 shows the measured values as a function of the reference values, and Table 4.1 shows the slope and independent term (y=mx+n) of a linear fit. The evaluated non-linearity as the value of the quadratic term to a second degree polynomial was identified as being smaller than 10^{-7} and the uncertainty of the measurements below 0.1%.

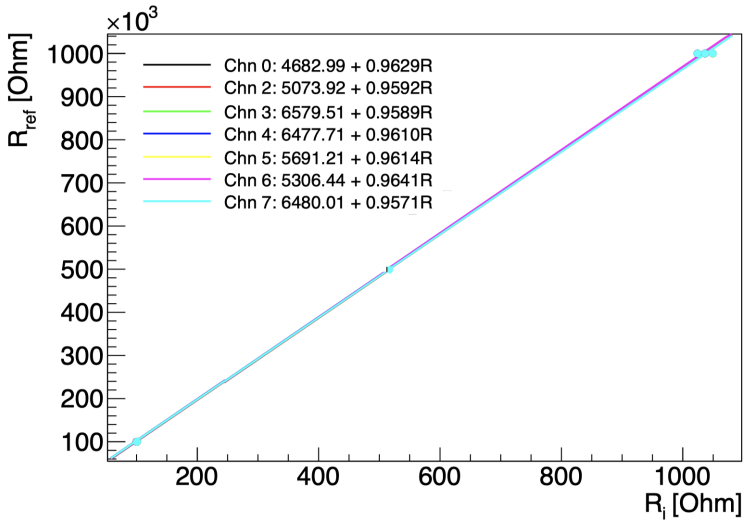


Figure 4.13: Measured values as a function of reference resistance.

Chn	m	$\varepsilon(m)$	n	$\varepsilon(n)$
0	0.96285	0.00146	4682.99	988.18
2	0.95918	0.00091	5073.92	639.07
3	0.95892	0.00161	6579.51	1079.19
4	0.96097	0.00135	6477.71	899.77
5	0.96141	0.00141	5691.21	930.00
6	0.96406	0.00137	5306.44	899.11
7	0.95710	0.00182	6480.01	1223.91

Table 4.1: Absolute calibration constants per channel.

As mentioned earlier, one of the two resistors in the voltage divider is known very accurately ($R_{ref} = 100 \pm 0.1$ kΩ) and the other is the NTC thermistor ($R_{NTC} = 100 \pm 1$ kΩ at 25°C) that varies in resistance as a function of temperature following the Steinhart-Hart Eq. 4.4.

$$\frac{1}{T} = \frac{1}{\beta} ln \left(\frac{R}{R_0} \right) - \frac{1}{T_0} \tag{4.4}$$

This is typically expressed as an equation of a single parameter β, when the parameters of the former are expressed as $a = (1/T_0) - (1/\beta)ln(R_0)$, $b = 1/\beta$, and $c = 0$. The reason for using this single parameter formula is that NTC thermistor manufacturers provide a

measurement of β at a temperature of 25°C. The NTCs used in VolTRON have a β value of $4092 \pm 1\%$.

The temperature measurement is then completely determined from the voltage reading by substituting the resistance (R) in Eq. 4.4 by the measured resistance of the NTC (R1) of Eq. 4.3. Fig. 4.14 shows the temperature as a function of the voltage and the uncertainty associated with each temperature value. As can be seen, the designed circuit has an uncertainty below 0.1°C in the range of interest from −25°C to 25°C.

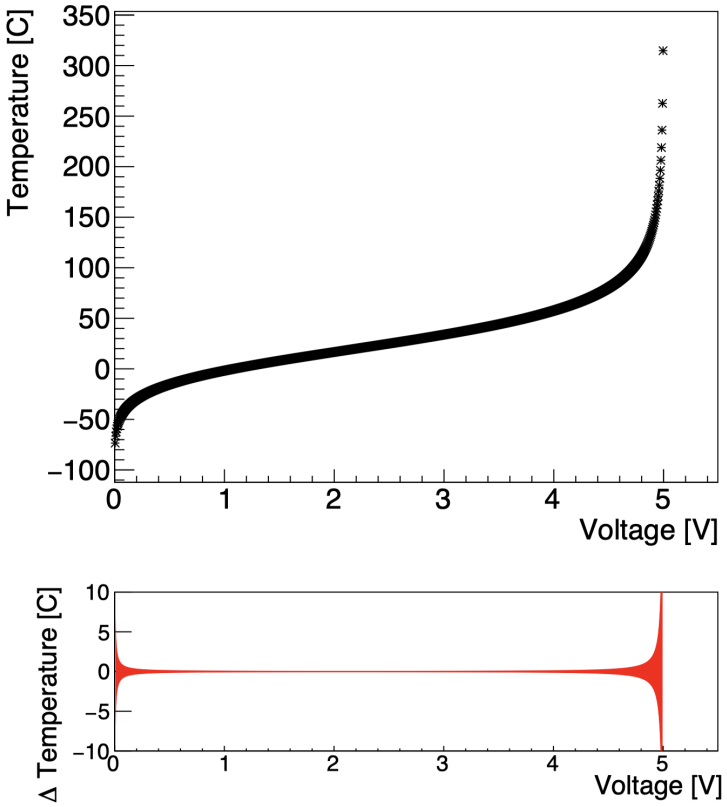


Figure 4.14: NTC temperature versus ADC voltage

The input voltage uncertainty is 0.1%, equivalent to 0.003 V, and the uncertainty of the 10-bit ADC is one bit, or 1/1023 V=0.001 V, given the precision of the reference resistor is 0.1%, and the NTC has a precision of 1%, the expected uncertainty in the voltage measurement is calculated using the error propagation approximation shown in Eq. 4.5. Since the input voltage is fixed, this uncertainty depends on the value of the reference and NTC resistance. Fig. 4.15 shows the uncertainty in the voltage measurement as a function of these two values. The systematic uncertainty associated to the voltage measurement is up to three times the uncertainty of the input voltage and decreases with resistance of the NTC. The uncertainty associated with the reference resistance is of the order of the input voltage uncertainty as shown in Eq. 4.4. Thus the output voltage uncertainty is $\Delta V \leq 0.5\%$. The associated temperature uncertainty is smaller than 0.5°C.

$$\begin{aligned}\varepsilon(V)^2 &= \left(\frac{\partial V}{\partial R_2}\right)^2 \varepsilon(R_2)^2 + \left(\frac{\partial V}{\partial R_1}\right)^2 \varepsilon(R_1)^2 + \left(\frac{\partial V}{\partial V_{in}}\right)^2 \varepsilon(V_{in})^2 \\ &= \left(\frac{R_2}{(R_2+R_1)^2} V_{in}\right)^2 \varepsilon(R_2)^2 + \left(-\frac{R_2}{(R_2+R_1)^2} V_{in}\right)^2 \varepsilon(R_1)^2 + \left(\frac{R_2}{R_2+R_1}\right)^2 \varepsilon(V_{in})^2 \end{aligned} \tag{4.5}$$

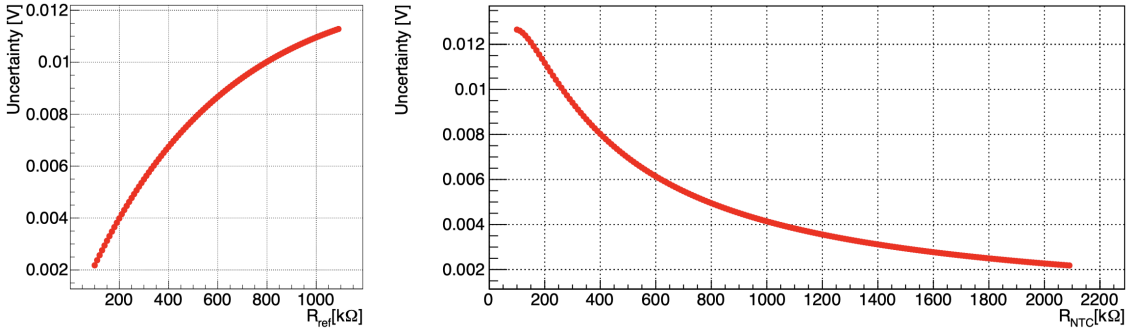


Figure 4.15: Voltage uncertainty vs reference resistance (left) and NTC resistance value (right).

In summary, VolTRON is a temperature readout system designed for the thermal characterisation of stave prototypes for the ITk that has undergone numerous development iterations, as shown on Fig. 4.16.

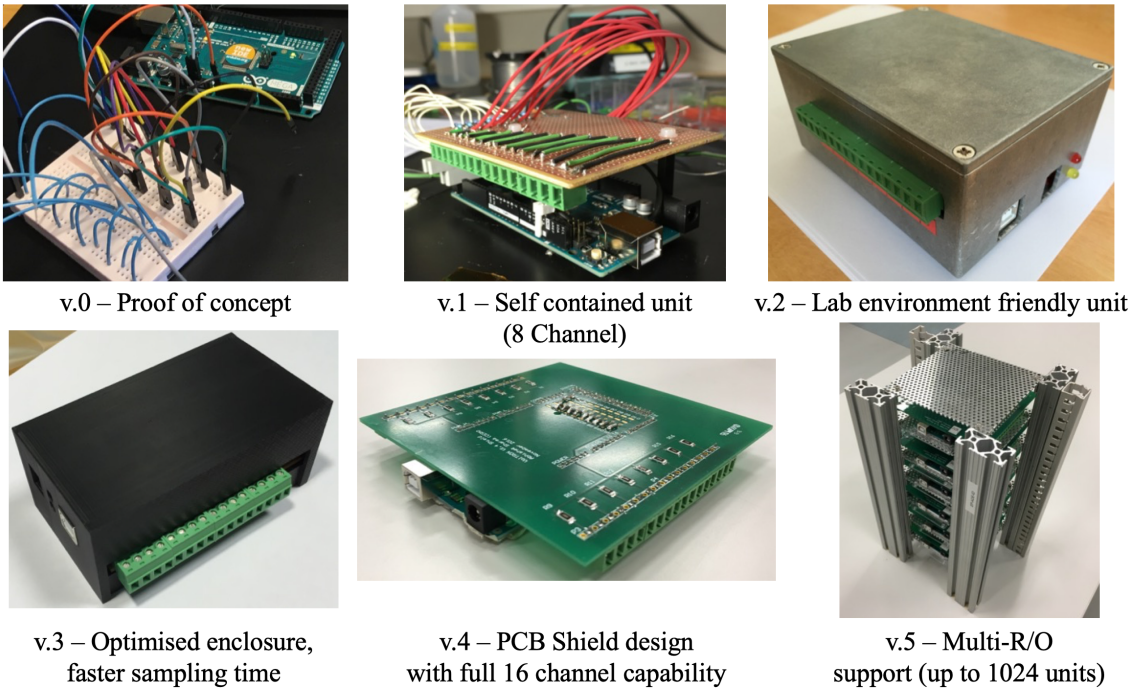


Figure 4.16: The complete development cycle of VolTRON, showing each different design iteration chronologically.

It makes use of a voltage divider circuit to quantify changes in the resistance of NTCs. It provides greater physics performance as compared to its predecessor, and has a small form

factor. It is entirely Python-based and its hardware and software are fully understood from the ground up.

Improvements at the level of system reliability and understanding of the systematic uncertainties associated with the system, which are below 0.5°C, provide great confidence in the measurements performed with the system within the temperature range of −25°C to 25°C.

4.1.5 Measurement Method

The measurement sequence begins with the installation of the stave sample into the climate chamber, for which a bespoke mechanical support is prepared. Below are the key stages in the measurement method, summarised at the bottom in the flowchart shown on Fig. 4.18:

- The sample is placed at ϕ=0, where ϕ is the angle between the vertical plane and the plane perpendicular to the silicon heaters, as shown in Fig. 4.17.

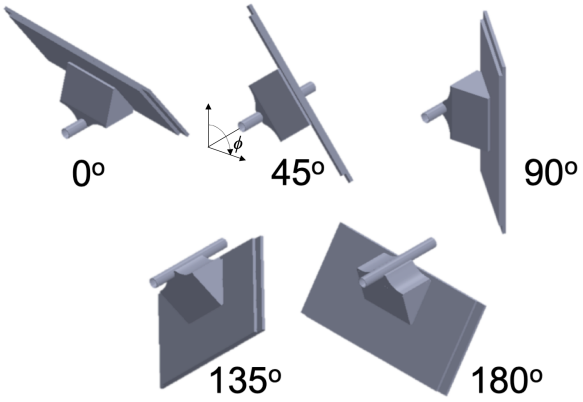


Figure 4.17: Visualisation of the different ϕ angle orientations the stave samples are secured in.

- Dry air valves are then opened to flush the entire climate chamber with dry air in order to drop the humidity, and hence the dew point, to −60°C to prevent any condensation and resulting ice build-up following the cool-down of the sample.

- The TRACI cooling system is then enabled with a temperature set-point of −20°C. Given the substantial thermal load across the circuit (from piping, sample, flowmeter, etc) this can take up to 4 hours. The CO_2 flow-rate is set to approximately 3 g/min.

- The Data Acquisition (DAQ) routine then starts, saving 100 temperature readings for every NTC connected to the VolTRON readout system at a rate of 0.2 Hz. The median values are output serially to the connected computer. This continues for a predefined time window, typically set to 10 minutes.

- The silicon heaters are supplied with an iteratively higher target power, ranging from 0 Wcm^{-2} to 0.7 Wcm^{-2} in steps of 0.1 Wcm^{-2}. Each step is followed by a 5 minute pause to allow for thermalisation at the sample level, before data-taking is resumed.
- The sequence thus far is repeated but with the dry air flow valve closed, preventing any air flow within the chamber. This is done to simulate extreme cases when convective air flows within the volume of the detector. After each data-taking step, the routine is paused to allow the dry air to be replenished within the chamber for 15 minutes, followed by an additional 5 minutes to allow the air to settle before resuming data-taking at the next target heat-flux step.
- Finally, the entire sequence is repeated with the sample incrementally tilted by 45^o, until the angles 0^o, 45^o, 90^o and 135^o are tested.

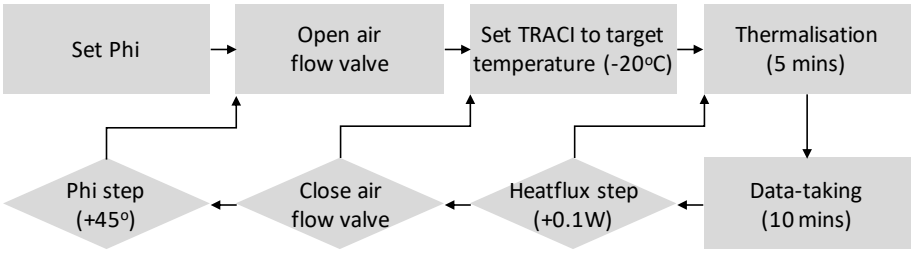


Figure 4.18: Flowchart descibing the TFM measurement DAQ routine.

4.1.6 Heat Load Optimisation

The 300 μm silicon heater cells are provided power via a set of copper braids formed of individual copper wires. The ends of these braids are placed in contact with the heater cells through a high temperature soldering process, shown on Fig. 4.19.

Current flow, and as a result heat generation across these heater cells was observed to be non-uniform and highly dependent on the geometrical configuration of the copper braids. A series of qualitative infra-red (IR) imagery tests were hence performed to observe the effects of these different configurations. Fig. 4.20 shows the increasing heat dissipation uniformity achieved with a higher number of contact points across the silicon heater.

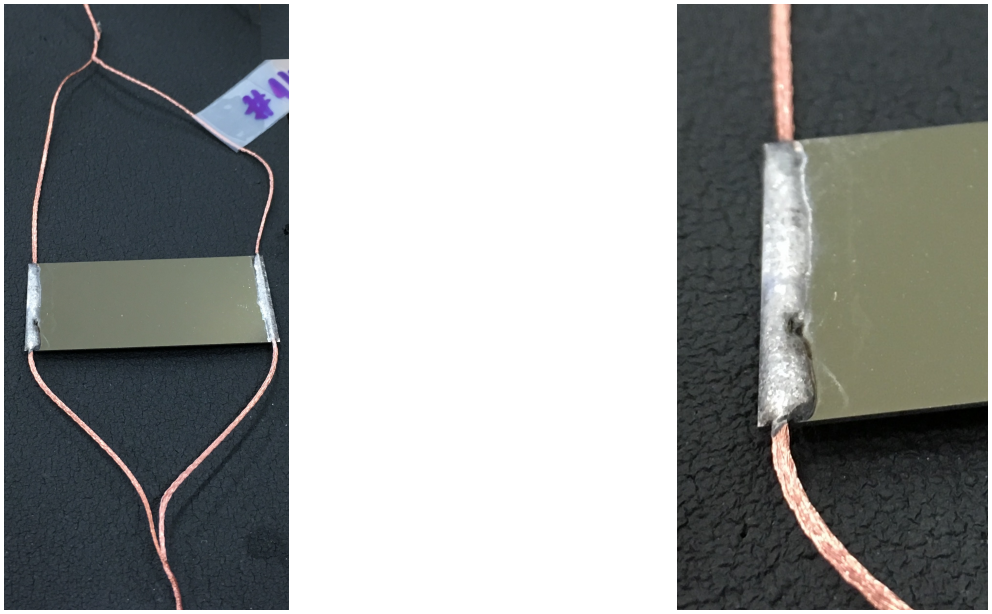


Figure 4.19: Copper Braids used to supply power to a silicon dual heater cell.

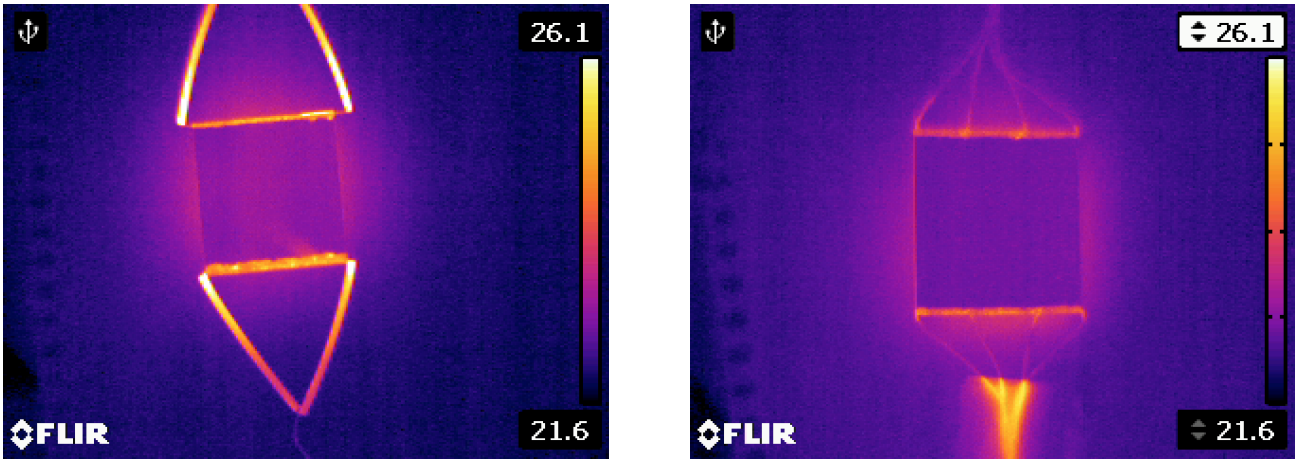


Figure 4.20: 2-point versus 4-point copper braid contact. The 2-point configuratin shows a slightly higher temperature area at the heater centre, whereas the 4-point connections provide a fully uniform response.

4.1.7 Performance Results

In this section, the different stave prototypes measured are introduced and their respective results are shown only in the case where outlying behaviours were observed. A full summary of the measurement results are detailed at the end of this section.

Stave samples were designed as a single cell, or triplets comprising of 3 heater cells. They targeted specific layers and sections in the ITk pixel stave geometry and as a result were either single heater, duals or quad module-sized cells.

SLIM Flat Barrel [70] Designed by the University of Geneva, this design originates from the Stiff Longeron for ITk Modules (SLIM) family of stave designs. These have been prepared using a stainless steel pipe with an outer diameter of 3.2 mm. Samples have been prepared in the formats of a single flat, single dual flat, single dual tilted, single quad flat and triplet quad flat cells. The triplet sample held its 3 cells in place on a stave design known as a Longeron, which is a thin hollow carbon-fibre tube designed to maximise mechanical stability whilst minimising its contribution to the material budget, shown on Fig. 4.21.

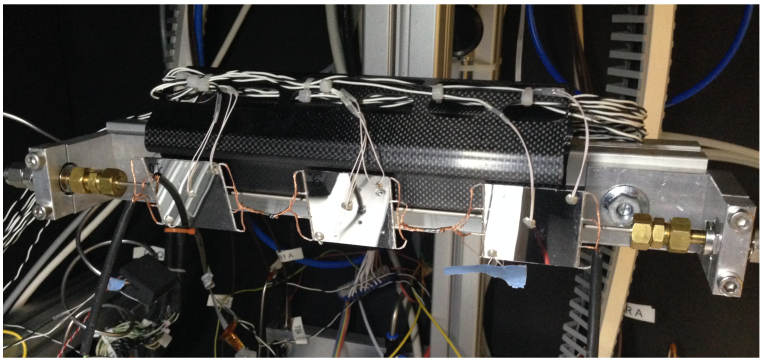


Figure 4.21: A quad flat barrel triplet stavelet secured onto a Longeron support.

These designs target layers 0 and 2. Fig. 4.22 shows the TFM curves produced at varying heat fluxes for a tilted single cell.

I-Beam [71] This design developed by the Lawrence Berkeley National Laboratory (LBNL), resembles a slanted I-beam in cross-section. Here, the design strategy involves the bulk of the beam being made of a densely packed highly thermally conductive foam covered by a carbon fibre outer layer and joined using a custom thermal interface material (TIM). This stave design inherently accommodates quad and dual sensors simultaneously on its opposite faces. In either case the cells are always in the flat configuration, not tilted. It houses two distinct cooling pipes each optimised for either the larger quad modules or the dual section. The first prototype was observed to have failed due to the absence of a thermal interface required to be applied between the foam and the cooling pipes. It was found to be absent on the cooling pipe for the quad side of the beam. This led to thermal runaway followed by the delamination of the quad heater cells, as shown on Fig. 4.23.

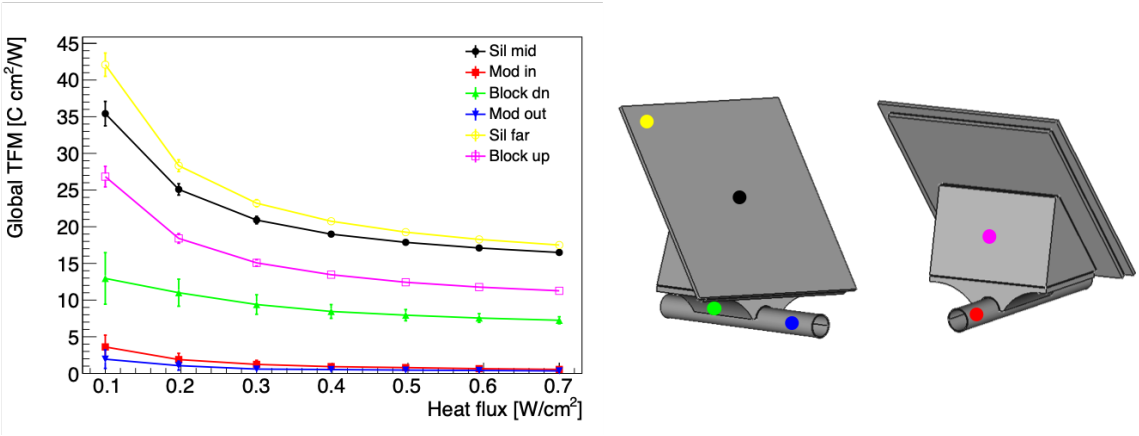


Figure 4.22: TFM versus heat flux for a SLIM tilted single cell (left), shown for various locations across the sample surface (right).

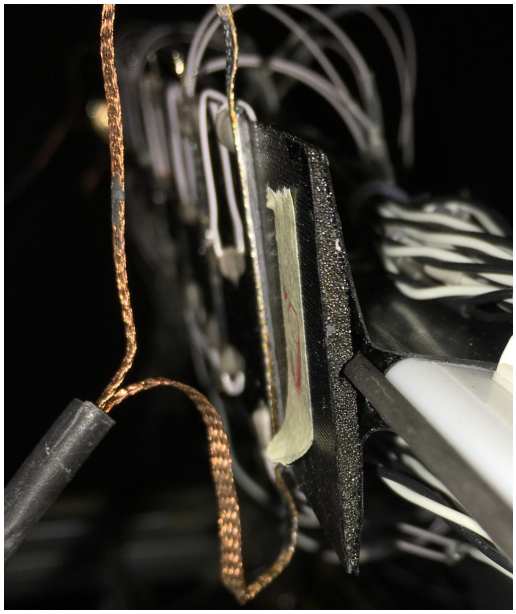


Figure 4.23: Absence of thermal interface material along the cooling pipe on an I-Beam sample leading to thermal runaway.

ALPINE [72] A novel "mountain"-shaped approach was taken by the Physics Laboratory of Annecy (LAPP) which combined the dense foam packing strategy employed in the I-Beam and expanded it to the tilted cell geometry. It also allowed for two cooling pipe contact points along the cell, and thus displayed excellent cooling performance, particularly for a tilted geometry.

Wuppertal [73] A series of staves designed by the University of Wuppertal, offered an alternative to the I-Beam's proposal for flat dual and quad cells. The design strategy revolved around minimising the material as well as volume of the stave, making it the most compact proposed prototype tested in this study. Design iterations focussed on the contact surface area between the stave and the pipe as well as changes in the pipe's material itself, testing both stainless steel as well as titanium, along with the cross-sectional shape of the stave, as shown on Fig. 4.24

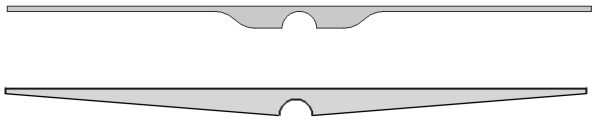


Figure 4.24: Geometry optimisations made between versions 2 (top) and 3 (bottom) to improve thermal performance on the Wuppertal stavelets.

The Manchester End-Cap [74] Three iterations of this design, prepared by the University of Manchester, have been constructed at the time of writing. Its heaters were placed on both surfaces of the stave and offset by half a heater length in order to maximise hermeticity with the fewest cell numbers. For the purposes of TFM tests, these samples housed only 5 cells, with 3 on one face and 2 on the other. The first design exhibited major cooling failures and as a result, thermal runaway, as shown on Fig. 4.25. This problem was identified as originating from a poor thermal conductivity of the chosen foam of which its body was predominantly composed of. The second iteration proved much more reliable with a corrected foam selection. The third design has as yet not been made available for testing.

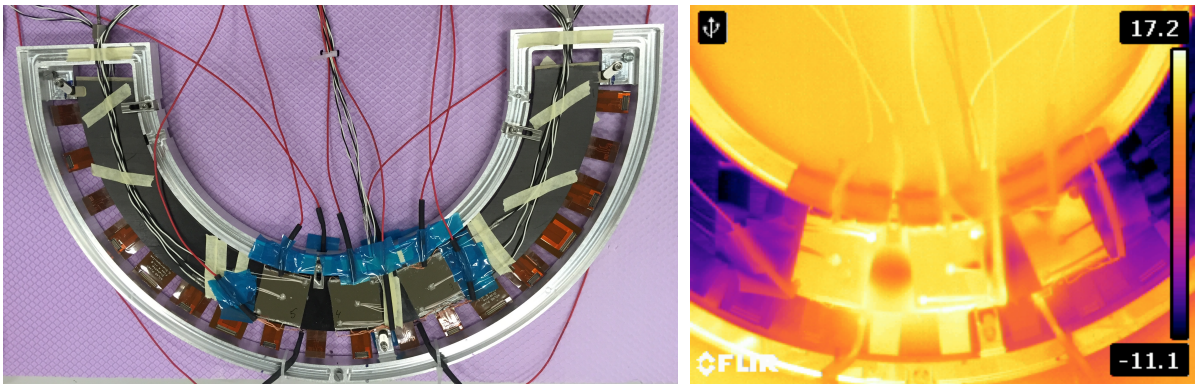


Figure 4.25: The first Manchester endcap ring prototype (left) and an IR image showing thermal runaway due to cooling failure (right).

The Berkeley End-Cap Double-Rings [75] LBNL also produced a double ring end-cap stave design targeting layers 0 and 2. This design populated sensors on both sides of the rings which accommodated quad modules in the outer ring and single modules within the inner ring. Similarly to the Manchester half-ring, the cells were also positioned in an alternating fashion so as to maximise sensing hermeticity.

Summary

Table 4.2 lists the TFM specifications for each layer and display variations commensurate with the radiation levels expected at each layer along with the number of cooling lines. It must be noted that these specifications undergo significant re-evaluations over the course of the ITk R&D.

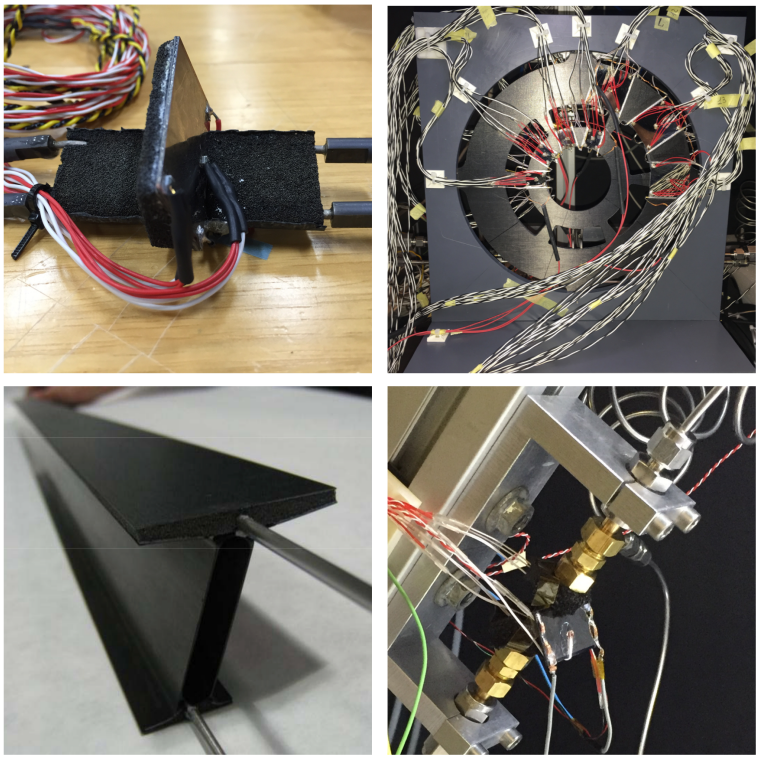


Figure 4.26: A selection of stave samples that were prepared for characterisation (clockwise): The LAPP Alpine (top-left), Berkeley Rings, SLIM titled dual and the Berkeley I-Beam (v.1) samples.

Layer	TFM K·cm^2/W
0	15.53
1	24.70
2	20.53
3	23.54
4	26.45

Table 4.2: TFM specification for the different ITk layers [76].

The summary of results obtained of the TFM for the different samples is shown on Table 4.3. The statistical uncertainty quoted is the root mean square (RMS) of the distribution of the TFM at 0.7 W/cm^2 with the cooling pipe in horizontal position and the heaters at ϕ = 90^o. The systematic uncertainty was identified as being $\pm$0.5 K·cm^2/W. This table also indicates the status of the stave samples relative to the specification as to whether or not they have been met, are partially met, in the case of only some layers, or fail to meet the required TFM values for the target layer completely. A clear improvement can be observed across all stave designs as they are iterated from their first versions to the next. The results obtained in this study were able to provide insight towards failure points across the surface of the samples through the many NTC locations that probed their geometries.

Sample	Type	Target Layer	# samples	TFM K·cm²/W	Status
I-beam	V1 Narrow Flange (dual heater)	0,2	1 (6 heaters)	11.8 ± 0.9 (stat)	Pass
I-beam	V1 Wide Flange (quad heater)	1-4	1 (6 heaters)	17.5 ± 0.6 (stat)	Pass
I-beam	V2 Narrow Flange (dual heater)	0,2	1 (6 heaters)	11.1 ± 1.5 (stat)	Pass
I-beam	V2 Wide Flange (quad heater)	1-4	1 (4 heaters)	25.9 ± 0.6 (stat)	Partial
Alpine	Mountain	0-4	1	13.3 ± 1.1 (stat)	Pass
SLIM	Flat Cell (quad heater)	1-4	3	18.0 ± 0.5 (stat)	Pass
SLIM	Tilted Cell (dual heater)	1-4	1	17.1 ± 0.6 (stat)	Pass
SLIM	Flat Cell (dual heater)	0	1	11.0 ± 0.5 (stat)	Pass
SLIM	Tilted Cell (single heater)	0	1	18.1 ± 0.1 (stat)	Fail
SLIM	Triplet Flat Cell (quad heater)	2	1 (3 heaters)	17.6 ± 1.1 (stat)	Pass
Wuppertal	V1 Ti Pipe Flat Cell (dual heater)	0,1	1 (3 heaters)	11.8 ± 3.5 (stat)	Pass
Wuppertal	V2 Stainless Steel Pipe, Flat Cell (dual heater)	0,1	1 (3 heaters)	8.7 ± 1.4 (stat)	Pass
Wuppertal	V2 Stainless Steel Pipe, Flat Cell (quad heater)	2-4	1 (3 heaters)	24.9 ± 3.2 (stat)	Partial
Wuppertal	V3 V-Shape Flat Cell (quad heater)	2-4	1 (3 heaters)	11.6 ± 2.1 (stat)	Pass
Manchester	V1 End-Cap (quad heater)	2	1 (5 heaters)	45.1 ± 10.2 (stat)	Fail
Manchester	V2 End-Cap (quad heater)	2	1 (5 heaters)	22.1 ± 3.8 (stat)	Partial
LBNL	V1 End-Cap (quad heater)	1	1 (6 heaters)	17.8 ± 1.5 (stat)	Pass
LBNL	V1 End-Cap (single heater)	0	1 (8 heaters)	4.4 ± 1.4 (stat)	Pass

Table 4.3: Summary of results for all stave samples TFM performance.

4.1.8 Further IR Analysis

In an attempt to verify the results, an alternative analysis method was established involving temperature measurements using a FLIR-E60 infra-red camera. The analysis comprised of taking infra-red images at every heat flux step, examples of which are shown in Fig. 4.27, and extracting the hottest point on each individual heater cell. This was done by defining rectangular areas in the pixel mapping of each IR image overlapping with each silicon heater, scanning for the hottest pixel, and matching its value to the temperature scale displayed at the edge of each IR image. This scale itself provided a maximum and minimum value which was input in the analysis manually, and a linear fit along this vertical scale was applied to interpolate the corresponding matching value of the hottest pixels for each heater cell. With this information, an alternative TFM measurement could be obtained, a comparison of which is shown on Fig. 4.28.

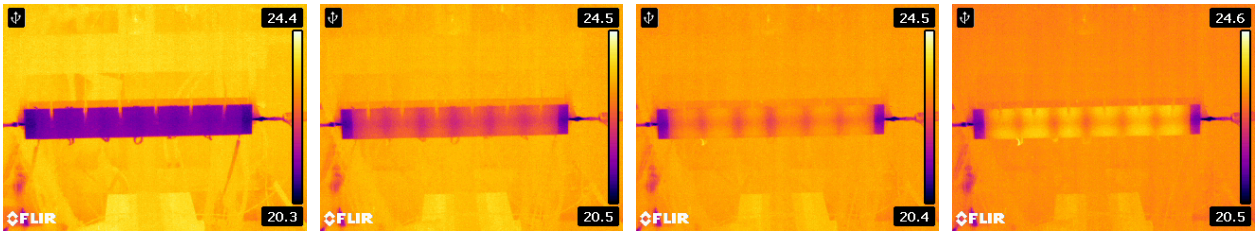


Figure 4.27: Example IR-images of the first Berkeley I-Beam protytpe at 0.0 $W\text{cm}^{-2}$, 0.1 $W\text{cm}^{-2}$, 0.2 $W\text{cm}^{-2}$ and 0.3 $W\text{cm}^{-2}$ (from left to right).

This method proved to be a very useful cross-check for early measurements with samples such as the Berkeley I-Beam (1^{st} prototype), where a matt white outer coating was applied, shown previously on Fig. 4.3. Such coatings were eventually omitted as part of the stave sample assemblies in order to not interfere with auxiliary mechanical characterisations, such as bending and warp tests. As a result, this method of analysis was discontinued as the bare surface of the silicon heaters exhibited a very high level of reflectivity, hindering an accurate

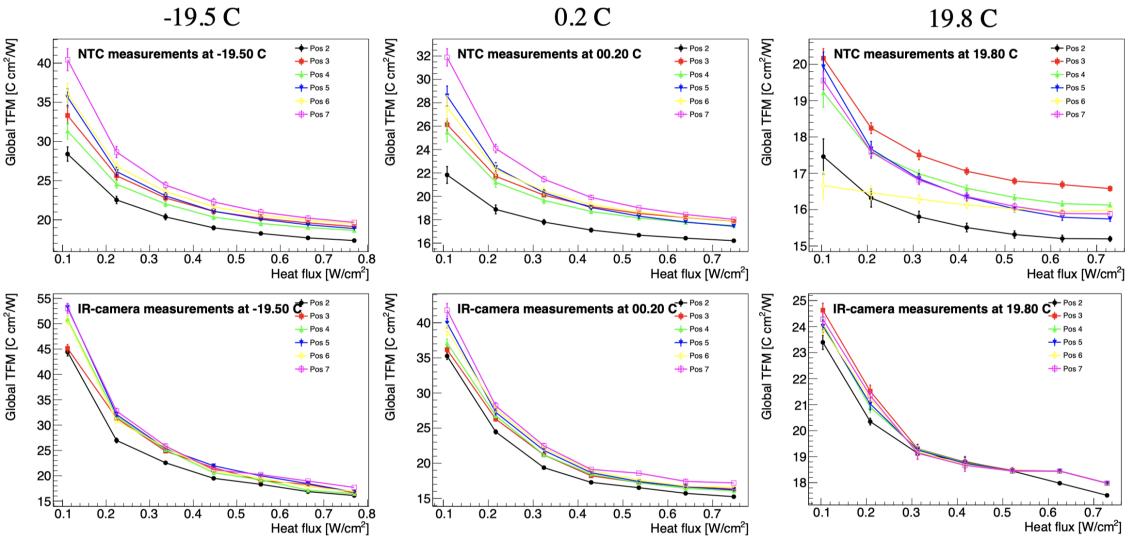


Figure 4.28: Comparative study between standard NTC-based TFM measurements and the IR-image approach. Results are shown for TFM measurements on the Berkeley I-Beam's first prototype sample at −19.5°C, 0.2°C and 19.8°C.

temperature measurement via IR imagery.



Figure 4.29: FLIR E60 infra-red camera (left) used for this analysis, shown imaging a 2×4 cm^2 dual SLIM prototype (right).

4.1.9 Thermal Cycling

An important aspect to establish that a design meets qualification is its ability to sustain repeated thermal shocks in the case of catastrophic failures in the cooling system. As mentioned earlier, typical coolant temperatures are expected to be around −30°C. Nevertheless, in the case of a leak in the stave's CO_2 piping, the rapid pressure release (expected to be just below 60 bars), would lead to a sudden temperature reduction, or thermal shock, followed by solid-ice build-up around the point of failure. In such an event, the stave must be designed

to withstand substantial temperature ranges well below the operational range, going from as low as −60°C to 25°C expected during technical stops.

To simulate such a failure, a selection of stave samples underwent thermal shock testing at the Goettingen University Temperature Shock facility, where they were subjected to fluctuations within these temperature extremes inside a "Voetsch Industrietechnik VT 7012 S2" thermal shock instrument. Table 4.4 lists the samples used for this study. This instrument comprises of 2 chambers, one kept hot and one cold located below the heated one, and a trolley that can slide between both chambers in under 5 seconds.

Stave	Type
Wuppertal	4×4 triplet, "40/2", V-shape
Wuppertal	2×4 triplet, "20/2", V-shape
SLIM	Flat Barrel 2×4 single
SLIM Longeron	Flat Barrel 4×4 triplet
I-Beam	4×4 6-cell & 2×4 6-cell, optimised

Table 4.4: List of stave samples that underwent thermal cycling.

The thermal cycling parameters chosen for this study involved pre-cooling the bottom chamber to −65°C and preheating the top chamber to 35°C. These values overshoot by 5°C the target temperatures to account for the thermal inertia of the samples. The samples were to remain at each temperature extreme over a dwell time of 3 minutes, for a total of 100 cycles.

During this process the NTC cabling was also kept inside the chambers for practical reasons. A temperature probe was secured in physical contact with the cooling pipe of a Wuppertal stavelet using thermal grease (Waemeleitpaste GS 0001/A Losimol), as shown on Fig. 4.30.



Figure 4.30: Thermal cycling instrument (left) with oscillating trolley filled with stave samples (middle) and temperature probe contact against a Wuppertal cooling pipe using thermal grease.

Fig. 4.31 shows the temperature readings of the probe in contact with the Wuppertal stave pipe along with dedicated probes to the respective chambers reading the ambient temperatures. The first 5 cycles on this plot pertain to tests conducted in order to optimise the characterisation parameters. The thermalisation at cold temperature is observed to take approximately 68 seconds, whereas heating thermalises slower at 100 seconds.

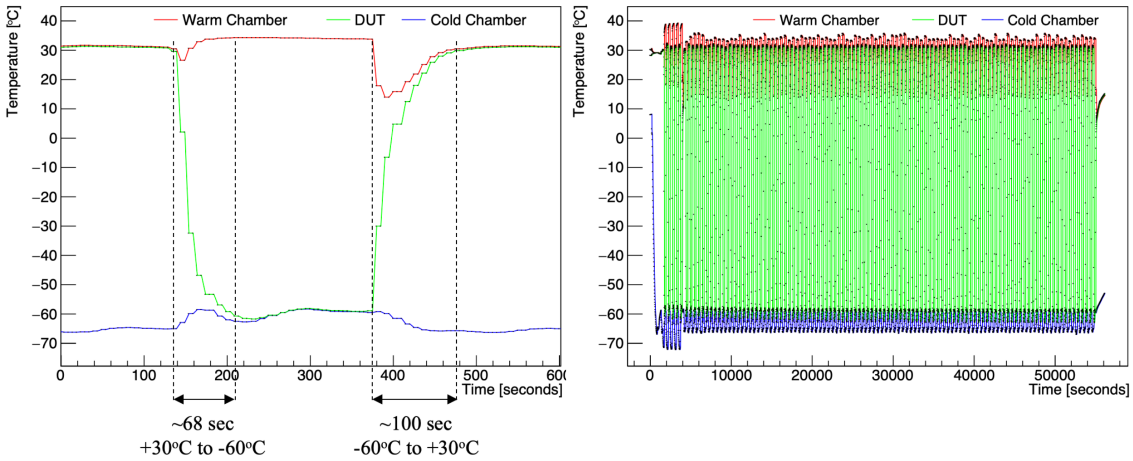


Figure 4.31: Temperature readings of the warm, cold chambers and the DUT for a single cycle (left) indicating the thermalisation times and all cyles (right).

After 100 cycles, the samples displayed no visible signs of damage except for the Wuppertal 2×2 triplet stave where all 3 heater cells fully detached from the main body, shown on Fig. 4.32. This has now been attributed to non-uniform application of the thermal interface material (TIM) used for this prototype.

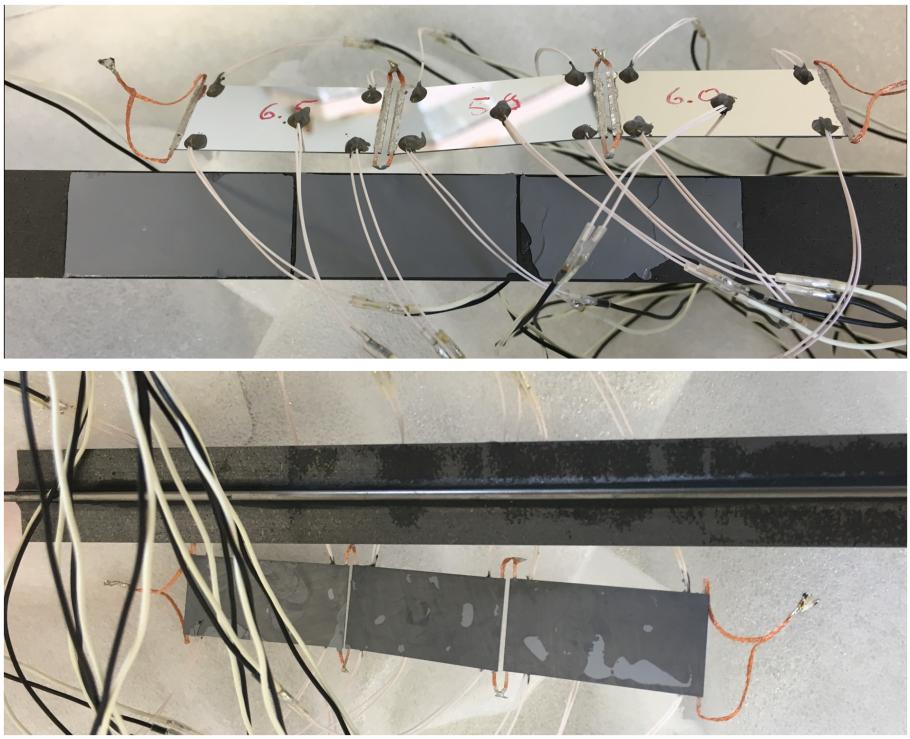


Figure 4.32: Wuppertal 2×4 triplet stave failure after 100 thermal cycles showing the full detachment of all 3 heater cells.

TFM measurements were repeated on the SLIM single cell sample after these thermal stress tests and the results are summarised on Fig. 4.33, showing a clear increase in the TFM post-cycling albeit still well within the specifications, validating the design for this particular sample.

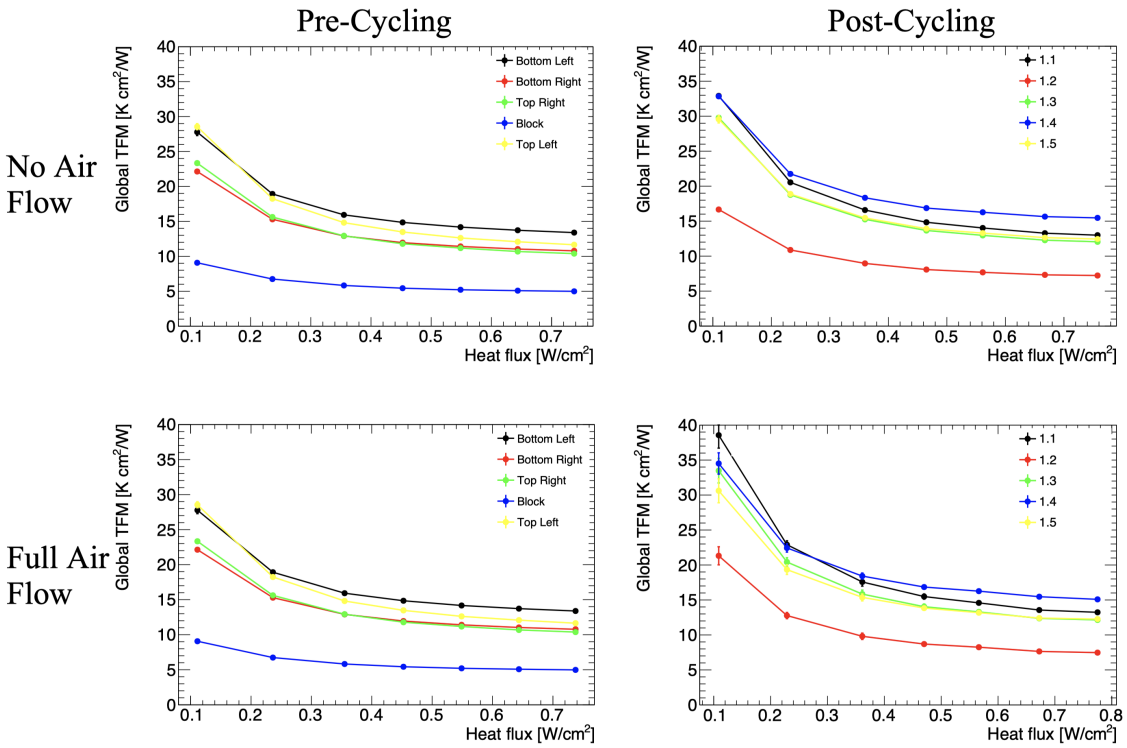


Figure 4.33: Performance results of the SLIM single cell sample before (left) and after (right) thermal cycling. Measurement results are shown for both scenarios of without any air flow (top) and full air flow (bottom), simulating worst case convection.

4.1.10 Conclusions and Outlook

This study characterised the full set of prototype stave designs produced for consideration towards the ITk tilted and flat layout support geometries. A comprehensive TFM measurement facility was constructed from the ground up including a custom scalable readout system. All results were summarised and submitted to the mechanics community of the ITk as a key engineering input. As of the date of writing, a tilted layout geometry has been selected for the design of the ITk using, in part, the assessments presented in this study.

A practical framework for further characterisation involving failure testing through thermal shock has also been clearly established, laying the path for systematic studies of the selected designs.

Over the course of this study, some areas of improvements for future such characterisations have been identified:

- Thermal pick up at the level of the stave sample is expected and is part of the qualification. Nevertheless, thermal pick up at the setup level, through its piping and other hardware components must be avoided and can be achieved to a greater extent by implementing the following changes to the setup:
  - Minimising the cooling pipe length by placing the CO_2 cooling unit as close as possible to the stave sample.

  - Maximising tri-axial piping coverage along the cooling circuit in order to prevent thermal pick up in the CO_2 and reduce the likelihood of boiling the CO_2 at the start of the cooling circuit.
  - Opting for a CO_2 cooling unit which incorporates a mass flowmeter within it, omitting the need for an external flowmeter (currently located just before the sample CO_2 inlet) and its substantial thermal inertia to contribute to the thermal load on the system.
  - Drastically improving thermal insulation by collapsing the entire climate chamber onto a small volume only marginally larger than the stave sample and sandwiching the entire assembly (other than the CO_2 cooling plant) in between two thick planes of ArmaFlex insulation foam. This would albeit result in the loss of convective effects around the stave.

- The CO_2 vapour quality measurement could be conducted dynamically in real-time by reading pressure and temperature sensor values at the inlet and outlet of the sample and extracting pressure-enthalpy data live from a repository known as REFPROP [77].

- With a known value of the CO_2 vapour quality, a given short stave sample could be simulated as being located along a specific length along the full stave dimension. This could be achieved by altering the CO_2 vapour quality which is expected to increase as it passes along a full stave length. Failure tests could hence be conducted at the worst possible cooling scenario along a stave length, namely near its outlet.

- Some institutes have opted to build their own TFM measurement setup replicating similar aspects of the work done in this study. These new facilities are now allowing for the cross checking of some of these results. In order to compare the performance and systematics in the measurements between the different setups, a reference stave sample would be greatly beneficial. Such a reference stave design has been proposed within the scope of this study and would comprise of a simple low thermal conductivity aluminium block. Its low thermal performance would lead to high temperature differences across its surface and thus easily differentiable TFM values across the easily machinable aluminium block. A preliminary FEA simulation, shown on Fig. 4.34, has been conducted on a $1.5\times4\times8$ cm^3 block which could accommodate two quad heater cells and with a pipe opening of 3.2 mm to mimic the cooling diameter expected to be used.

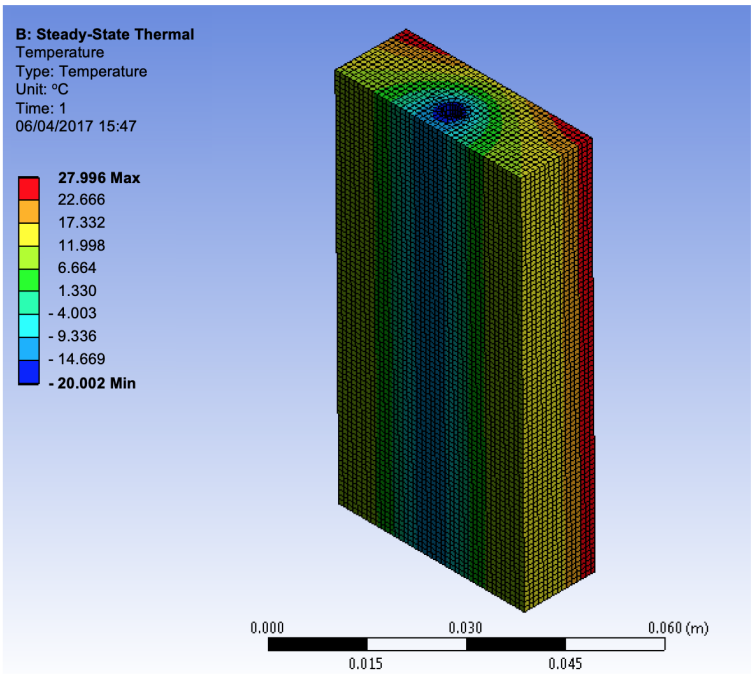


Figure 4.34: Aluminium reference stave with cooling pipe opening askew relative to the stave cross-section centre. FEA analysis of the block's cooling pipe lining surface applied a temperature of -20°C and its heater surface being applied a heat flux of 0.7 Wcm^{-2}, showing as expected a significant ΔT in order to facilitate TFM measurements.

4.2 Cable Saver Boards for The Demonstrator Programme

4.2.1 The Demonstrator Programme

In April 2017, the ITk Pixel Outer Barrel Demonstrator Programme was conceived with the aim of producing a series of detector section mock-ups (known as demonstrators), both semi-operational and fully operational, intended to evaluate their designs against the relevant Phase 2 engineering specifications.

Key stages of this programme including the different demonstrator types involved are listed below:

1. One dual and one quad heater with heater flexes loaded on individual cells.
2. One quad module with flex loaded on cell.
3. Small integration prototype.
4. Small electrical structure (single 300 mm long cooling line loaded with 7 × quad electrical modules).
5. Long thermal prototype (single 1.6 m long cooling line fully loaded with heaters).
6. Thermo-mechanical prototype (functional Longeron with 4 × 1.6 m cooling lines).
7. Full Demonstrator System Test Facility with CO_2 cooling, Detector Control System (DCS) and Serial Powering (SP).

In this context, a demonstrator comprises of a 1:1 scale stave (section or full-length) designed using materials, production assembly methods and tools, closely matching those foreseen to be used for the final construction of the ITk outer barrel stave. Figure 4.35 shows the different components to be included when developing a demonstrator. In the case of thermal staves, silicon heater cells with PT100 temperature sensors on its four corners and one centrally placed are used, or, FEI4 chips in the case of electrical staves.

One of these components is the Cable Saver Board. Its purpose is to transition the delicate custom-made power and data flex cables exiting the staves to the more robust general-purpose laboratory connectors routing the stave data and power to their relevant instruments for testing. Table 4.5 summarises all the different CSBs needed for the various stave prototypes where the module positions are described in section 1.3 and where X+ is in the direction of ϕ (described in section 4.1.5) and X− in the opposite direction.

This section presents a description of the cable saver boards.

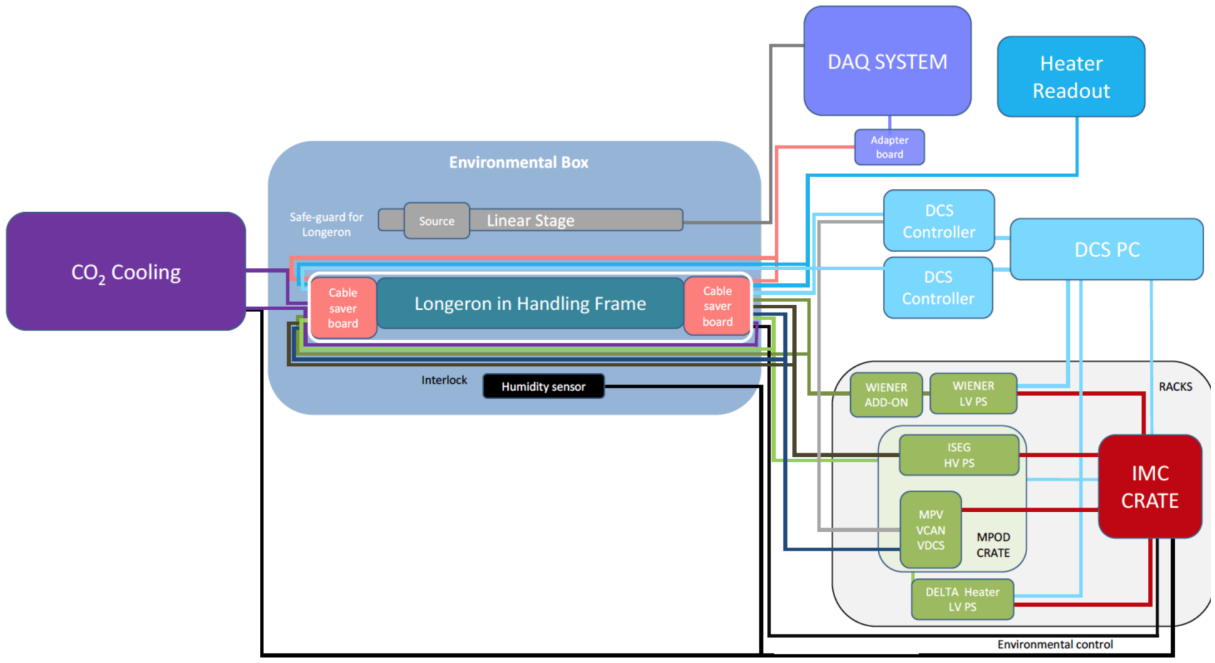


Figure 4.35: System setup overview for a generic stave demonstrator, with the Cable Saver Boards (CSB) shown in pink.

Cable Saver Board	Demonstrator	Location	Quantity
Thermal (short)	Short Integration	Side A X+ L3	1
Electrical	Short Electrical	Side A X+ L3	4
Thermal (long)	Full Thermal	Side A X+ L3	1
Thermal (long)	Full Thermal	Side C X+ L3	1
Electrical	Full Demonstrator (Electrical)	Side A X+ L3	1
Electrical	Full Demonstrator (Electrical)	Side C X+ L3	1
Thermal (long)	Full Demonstrator (Thermal)	Side A X− L3	1
Thermal (long)	Full Demonstrator (Thermal)	Side C X− L3	1
Total	-	-	11

Table 4.5: Summary of CSBs designed for the entire demonstrator programme.

4.2.2 Cable Saver Boards

Short Integration Prototype

The short integration prototype has been specifically designed to assess the following features:

- Integration tests including heater on cell and cell on Longeron.
- Flex bending
- Validating connector pin-out mappings for Thermal Prototypes.
- Commissioning readout Thermal Prototypes.
- Initial Thermal Figure of Merit (TFM) tests.

This prototype is comprised of 5 dual and 4 quad silicon heaters, with 5 PT100 sensors on each heater cell, mimicking a section of the barrel region's layer 3 stave dimensions. Figure 4.36 shows a view of its CAD model and Figure 4.37 shows a view of its CSB's layout and integration with the stave prototype.

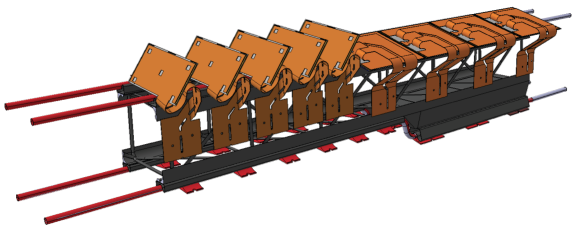


Figure 4.36: Short Integration Prototype CAD view.

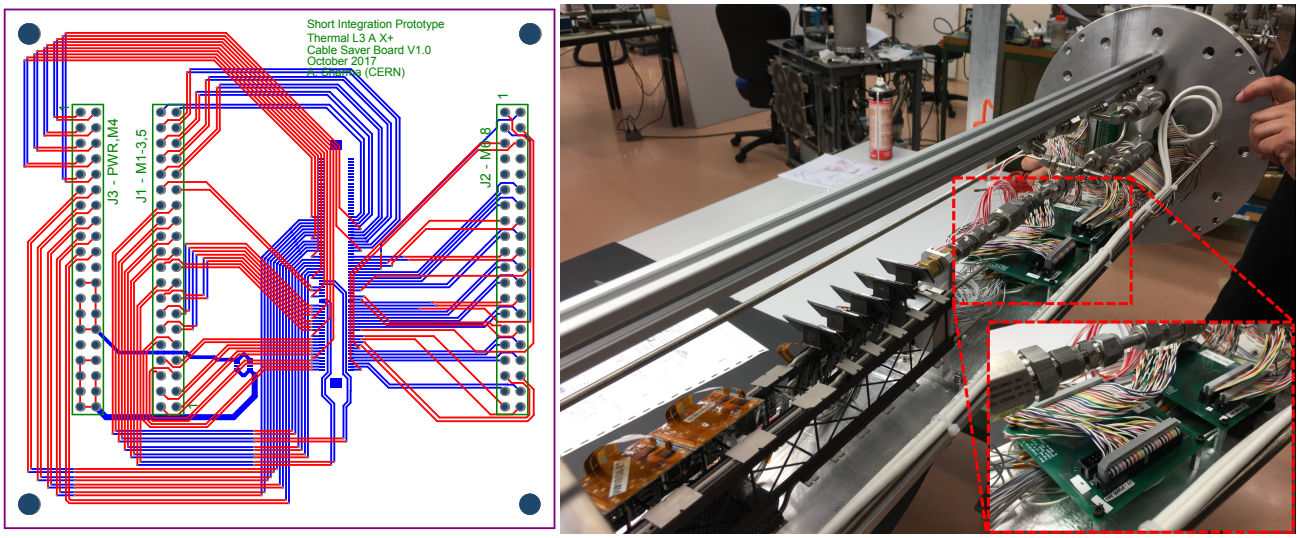


Figure 4.37: Short Integration Prototype CSB Layout (left) and assembly view (right).

Long Thermal Prototype

The purpose of the Long Thermal Prototype stave involves assessments of the Thermal Figure of Merit (see section 4.1) across the full cooling line, as well as examining manufacturing variability and performing CO_2 vapour quality measurements. The prototype is built of 44 heater cells with 5 PT100 sensors on each cell (all four corners and one central).

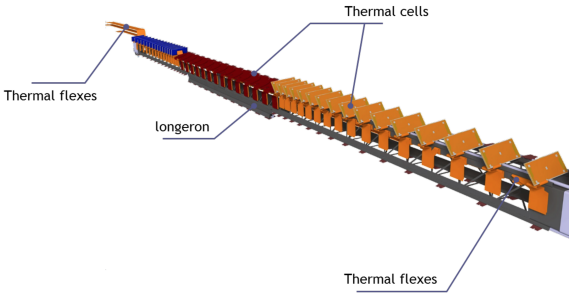


Figure 4.38: Long Thermal Prototype CAD view.

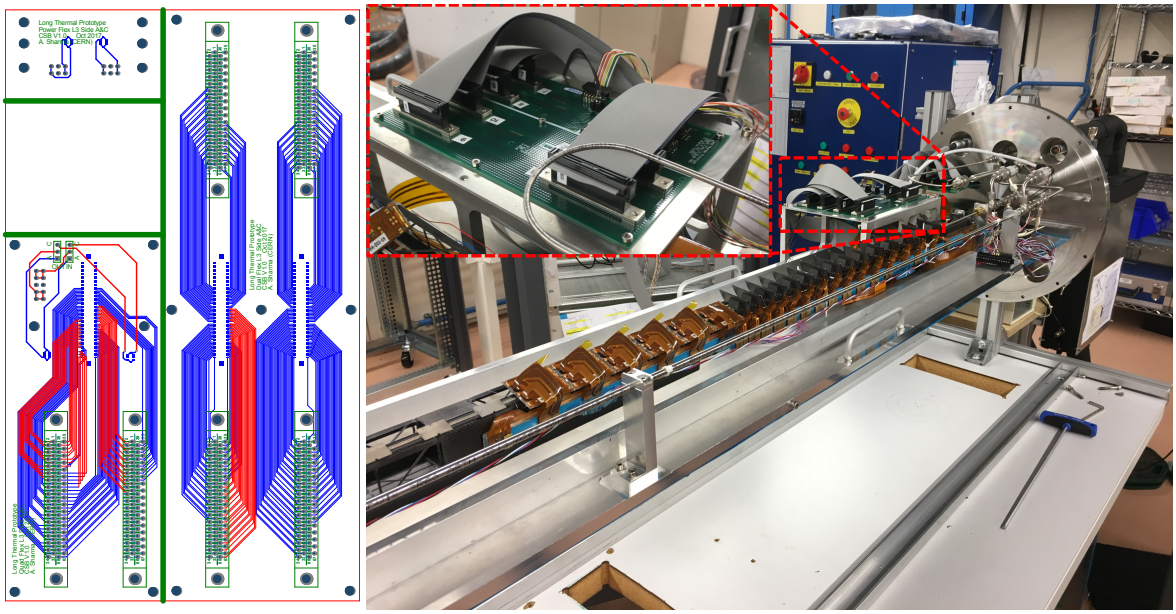


Figure 4.39: Long Thermal Prototype CSB layout (left) and assembly views (right).

Short Electrical Prototype

The Short Electrical Prototype stave's aim is to assess electrical functionality on a small stave section prior to proceeding with longer demonstrator productions. Tests unertaken on this smaller structure include interlock tests, electrical contacts, DCS, rack, power supplies, Pixel Serial Powering Protection (PSPP [78]) chip tests and finally radioactive source tests. This prototype also mimics layer 3 barrel stave dimensions and comprises of 7 quad modules mounted on a stave section.

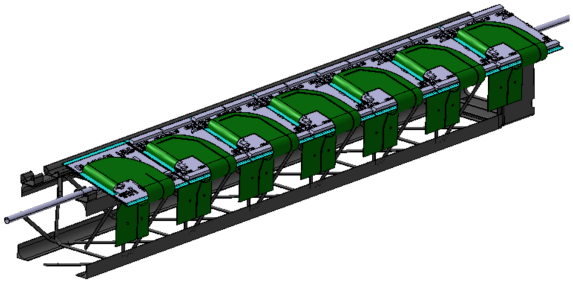


Figure 4.40: CAD model view of the Short Electrical Prototype.

Multiple versions of the electrical CSB were produced to cater to the symmetrical nature of the prototype stave. One of the features under test in the Short Electrical and Full Demonstrator Staves is the performance of various serial power chain configurations. The electrical CSB has hence been designed with the ability to allow the user to select between 1 or 2 low voltage powering chains as well as the option of providing high voltage through 1, 2, 3 or 4 distinct channels.

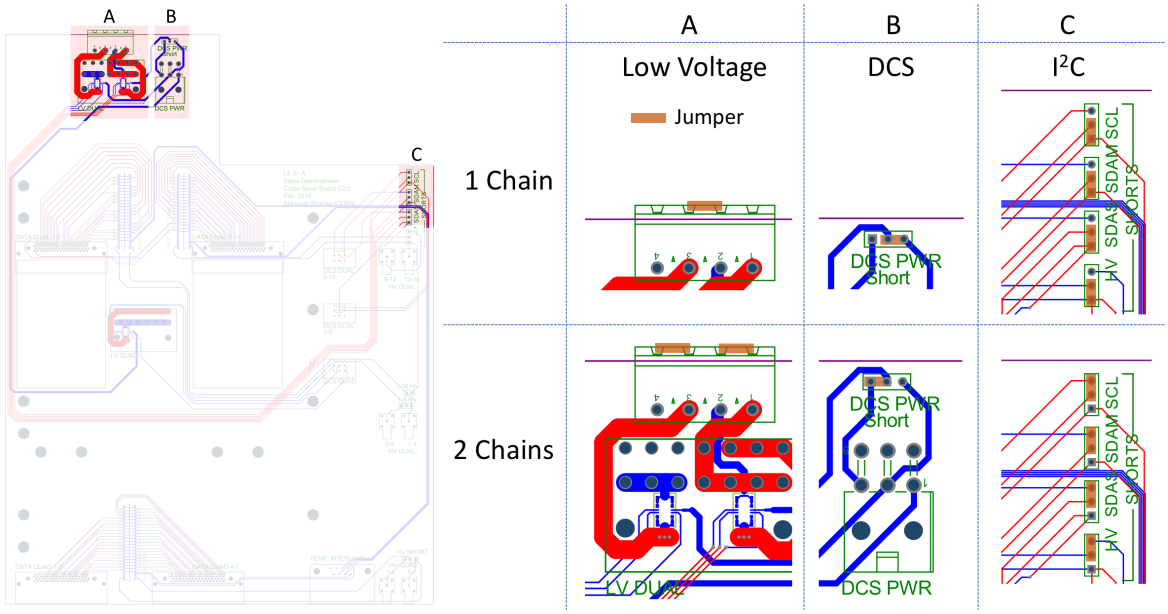


Figure 4.41: Electrical CSB jumper configurations for low voltage and DCS power chains.

The dimensions of the electrical CSBs were constrained by the positions of the power and data flex cable connectors as well as the shape of the handling frame of the Final Demonstrator. Data transfer rates are expected to be up to 320 Mbps. As a result, all differential data tracks were designed to have a track pair separation no greater than 100 μm and neighbouring track separation of larger than 1.2 mm. Track length differences between given differential pairs was constrained to be within 15 cm. This low limit has been set as a contingency to take into account the total track length differences for a given differential pair along the entire data chain (from FEI4 module to readout) across all connectors and cables. The impedance targeted was of 100 Ω.

A particular data connector (VHDCI) pin-out scheme was established so as to maximise

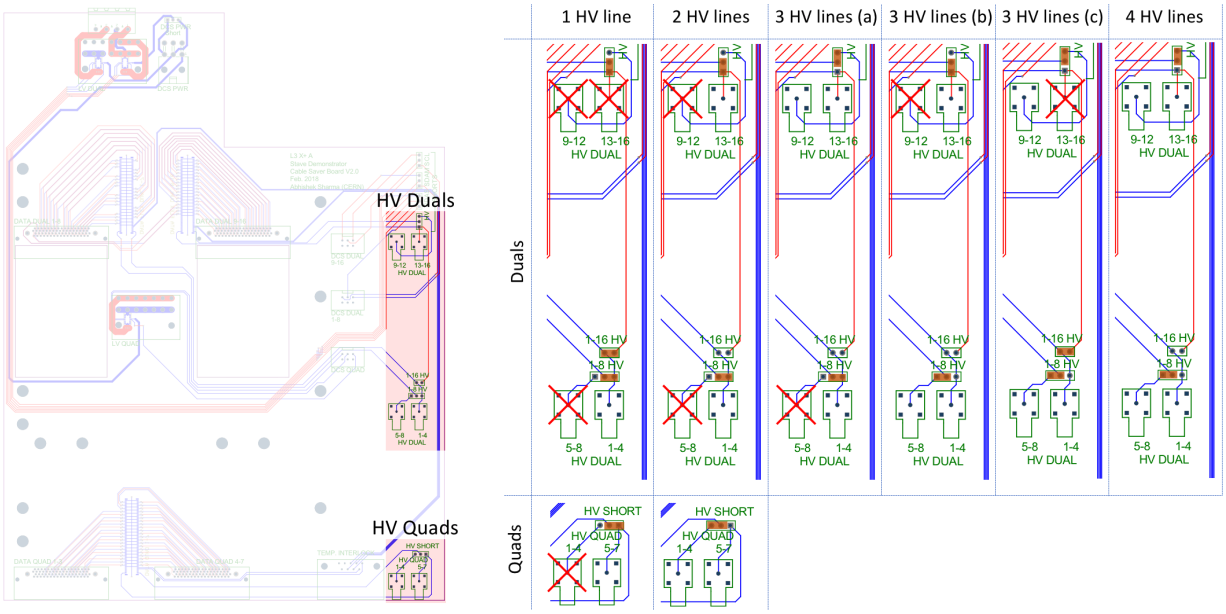


Figure 4.42: Electrical CSB jumper configurations for high voltage lines, including indications on which corresponding SMC connectors to leave disconnected.

modularity at the readout level. Consequently the VHDCI data cables can be connected to the relevant readout instruments independently of whether the cables pertain to dual modules or quad modules.

Full Demonstrator

The Full Demonstrator combines a full layer 3 cooling line populated with 30 dual and 14 quad heater cells, with a neighbouring layer 3 cooling line comprised of up to 32 dual and 14 quad modules. Its research aims include examining:

- Serial powering capabilities across 16 modules.
- The capability to cool two full cooling line.
- Full performance evaluation of electrical (dual and quad) modules, including source tests.
- Variability in the production of modules (5 institutes).
- Feasibility of 6 serial power chains running together.
- Readout tests for all Front-Ends.
- Integrating such a large prototype as well as its handling.
- Commissioning of the system setup.

The same CSB models used for the Long Thermal and Short Electrical prototype staves are made use of for the corresponding sections of the Full Demonstrator.

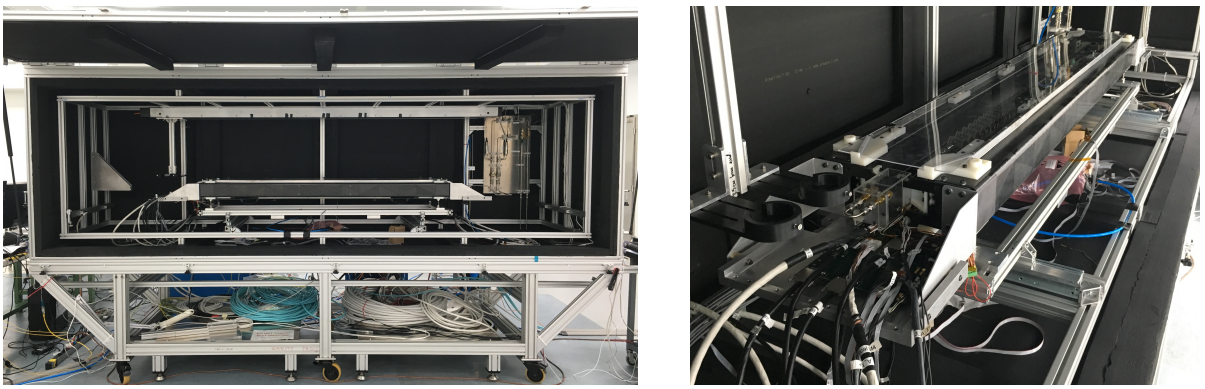


Figure 4.43: Final Demonstrator CSB integrated assembly view inside its enclosure.

4.2.3 Summary

The development of CSBs for every demonstrator within the ATLAS ITk Outer Barrel Demonstrator Programme was done in order to facilitate the R&D of the integration and testing being undertaken to qualify the electrical and thermal sensors under consideration. In order to achieve this, PCB trace connection patterns had to be identified for all CSBs so as to minimise the number of layers required to achieve the necessary flex to connector pin mappings all whilst abiding by the stringent data speed and differential signal requirements of the data connections. This also included special considerations for the high current powering lines as well as the mechanical constraints that were placed by the larger assemblies surrounding the CSBs, such as the handling frames and cooling lines. The completion of this task has allowed for extensive electrical characterisations and full system tests to be performed on all ITk stave demonstrators.

An additional benefit of the CSB development has been to make more accessible the powering and data connectivity needed to qualify the ITk readout system. This is discussed in greater detail in section 4.3.

4.3 The FELIX Readout System

Beyond the sensor and mechanical developments being undertaken towards the Phase 1 and 2 upgrades of the LHC, these luminosity upgrades will also have repercussions on the number of electronic channels required to read out the detector. For this reason, the whole experiment will replace its back-end electronics with the new Front-End Link-eXchange (FELIX) system [79].

One of the sub-detectors that will make significant use of FELIX, due to its large number of electronics channels, will be the ITk pixel [27]. This section provides an overview of the efforts undertaken to demonstrate a full tuning sequence of the FEI4 [15]. This is the Front-End used in the very first ITk pixel demonstrator.

4.3.1 Upgrades of the ATLAS DAQ System

The first element within an ATLAS detector to receive data was known as a Read-Out Driver (ROD). The next, is the Read-Out System (ROS). The link between both these units has been a completely custom developed high speed S-link [80]. The rest of the data flow goes through commercial off the shelf elements (or COTS). In the next iteration of the ATLAS DAQ system, the output of the detector will avoid the custom link in favour of high speed links called GBT links. The element that will receive these links is called FELIX. The connection to the data flow downstream will still involve COTS.

In the past the ROD had among its functionalities, formatting and monitoring as well as including a busy generation and Timing Trigger Control (TTC) synchronisation. Formatting and monitoring will now be handled by FELIX, such that when formatting is needed, the data will be propagated to a data handling unit called a Software ROD (SWROD). The SWROD will be a new element in the ATLAS DAQ system, providing formatting and monitoring, and sending it to the COTS. The shift from custom to commercial links will allow for the use of common units which are dramatically cost efficient. A diagram of the Trigger and Data Acquisition (TDAQ) architecture for the Phase 2 upgrade is shown in Fig. 4.44.

The FELIX-based readout is expected to be introduced by the Phase 1 upgrade on numerous ATLAS sub-detectors. This will reduce the number of sub-detector specific components. Towards Phase 2, the remainder of the ATLAS sub-detectors will replace their ROD and ROS elements by a FELIX readout and SWROD system.

FELIX supports different protocols for data and commmand links. The FULL mode which uses the full 3.3 Gbps payload bandwidth of the link is designed to transfer packets that are multiples of 32-bit words, with no maximum packet size over a 9.6 Gb/s link with a payload of 7.68 Gb/s due to 8b/10b encoding. The GigaBit Tranceiver (GBT) mode [81] transmits 120-bit packets over a 4.8 Gb/s speed link, which includes a 4-bit header, a 4-bit slow control word, 80 bits of user data (3.2 Gb/s), and a 32-bit Forward Error Correction (FEC) block. The 80 bits of the user data can then be further split into E-links. These are a variable width logical link on top of the GBT protocol that makes it possible to separate different streams on the same physical link. The GBTX 130 nm CMOS ASIC [81] is designed

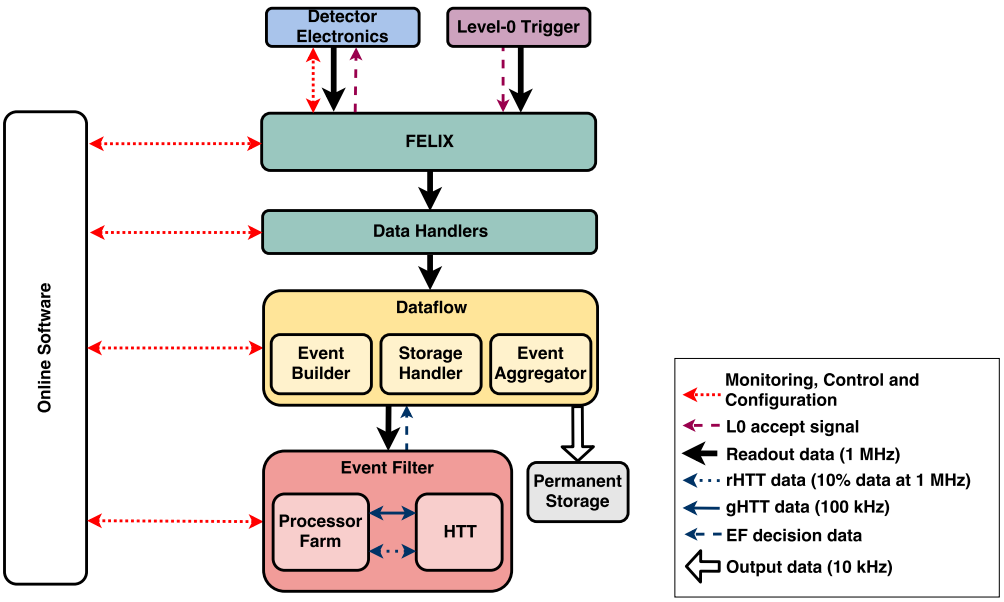


Figure 4.44: TDAQ architecture for Phase 2 upgrades [79].

for front-end electronics and has been qualified to be radiation hard.

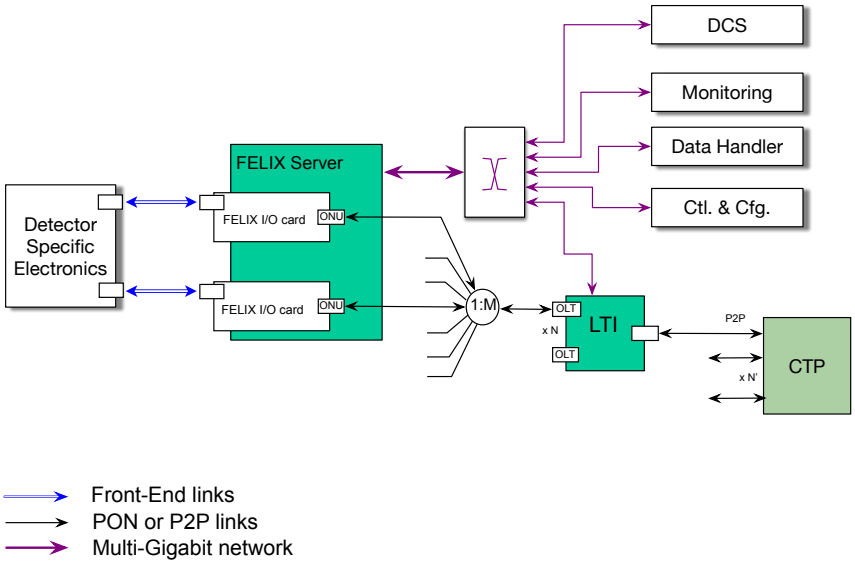


Figure 4.45: FELIX structure for Phase 2 upgrades [79].

The FELIX structure is shown in Fig. 4.45. It is based on custom input-output boards residing within a server with a high bandwidth network connection. There are up to 48 links on each FELIX card that can handle 48 links in GBT mode or up to 24 links in case of FULL mode in the Phase 1 system. For Phase 2, it is foreseen to handle up to 24 Low Power GBT (lpGBT) links, currently at the early development stage.

The FELIX group develops the complete software stack for Scientific Linux 6 and CERN

CentOS 7 operating systems. This includes the development of the kernel drivers, low level application program interface (API), the command line interface tools, as well as the FELIX core application. This core application provides a network interface communication layer called NETIO for the high level connectivity to the SWROD, or the ITk calibration software. The application runs two threads per FELIX input-output card for the uplink, since there is one process per direct memory access (DMA) and 2 DMA per Peripheral Component Interconnect (PCI) endpoint. The downlink has independent threads for slow control, monitoring, and busy.

4.3.2 ITk FELIX Integration

The integration of the FELIX detector interface with the ITk demonstrators has been carried out in small scale systems that represent a full readout chain of the detector, from the FEI4 detector modules, up to the ITk calibration software. As depicted in Fig. 4.46, the FEI4 front-end modules accept command signals at 40 Mb/s, and transmit data at 160 Mb/s. The data from up to 20 individual FEI4 front-ends is aggregated into a single GBT link at 3.2 Gb/s via a GBTx ASIC which provides 20 E-links per optical link. The optical links are plugged into a FELIX input-output card which is controlled by the FELIX core application, and interfaced to the ITk software through the NETIO protocol.

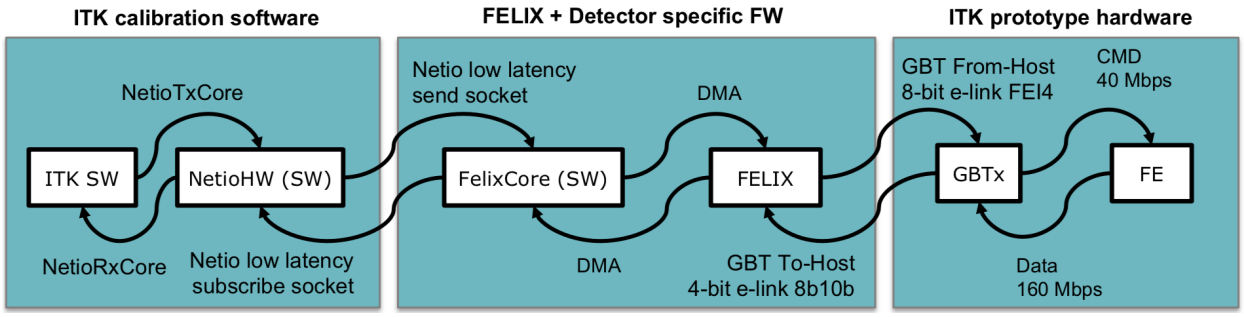


Figure 4.46: Flow of data and commands submitted to and from a module through a scan on ITk front-end prototypes

The FELIX firmware is running a sub-detector specific module in the uplink in order to communicate with front-ends that are AC coupled. This requires the use of a 'Manchester'-encoding which is defined as each bit being either low then high, or high then low, for equal amounts of time in the communication. This provides the main advantage of making the signal synchronised and self-clocking, minimising the error rate. The demonstrator requires 8 GBT links. Several boards have been designed for the integration of FELIX with the FEI4:

1. Small adapter from module shown on Fig. 4.47 allowing for the inclusion of individual modules.

2. VHDCI to 16×HDMI-A, Fig. 4.48, which can cope with up to 4 quad modules or 16 individual front-ends, with jumper configurations enabling master/slave functionalities.

3. VHDCI-GBTx-SFP board shown on Fig. 4.49. This board, based on the Versatile Link Demonstration Board [82] was developed to drastically condense the connectivity required to transduce command, clock and output data signals from the CSB (see section 4.2) via a GBTx ASIC to an optical signal through a Small Form Factor (SFP) transceiver. This board can accommodate data transfer for up to 16 individual front-ends modules or 8 dual modules. It also houses two FEAST [83] DC/DC converter chip modules providing a clean source of 1.5 V and 3 V minimising any rippling effects from power sources, given the fragile power level acceptance of the GBTx chip.

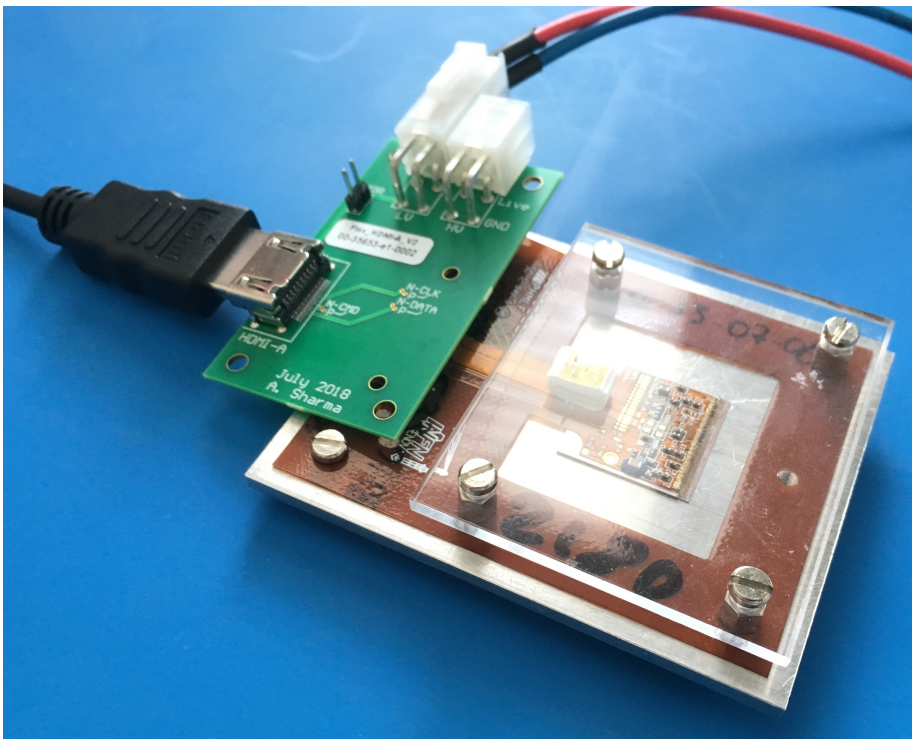


Figure 4.47: Flex to HDMI-A adapter board for data, low voltage and high voltage propagation for a single front-end.

The demonstrator software is based on the implementation of what is known as the FEI4 'tuning sequence' using FELIX. This tuning is specific to the FEI4 and is an iterative process that involves many sub-processes known as 'scans'. Each scan is a computing intensive process that handles a large data volume. For example, one of the goals of the tuning is to define the threshold at which the pixel will produce a hit, which should be the same for all pixels in the chip, in a manner similar to that defined for the turn-on curve (or S-curve) in 2.5.2. This is done with two different scans, one that sets the threshold globally for the whole front-end, and one that adjusts the threshold per pixel. The requirement for the first one is that the mean value of the threshold distribution matches the expectation, the second scan requires that all pixels have a similar threshold distribution. It is possible that the mean value of the threshold distribution changes after the per pixel threshold scan, thus requiring the execution of a second global threshold scan.

The ITk software required for calibration runs allow the implementation of several front-ends like the FEI4, or the RD53 [84], which is the next pixel front-end prototype for ITk.

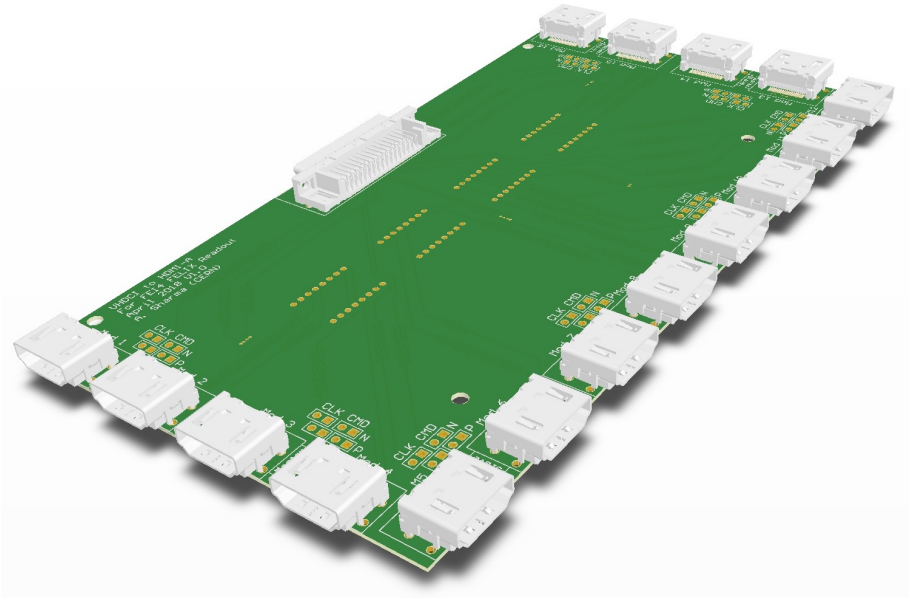


Figure 4.48: VHDCI to 16×HDMI-A board allowing for data, low voltage and high voltage to be propagated to 16 individual front-ends or 4 quad modules.

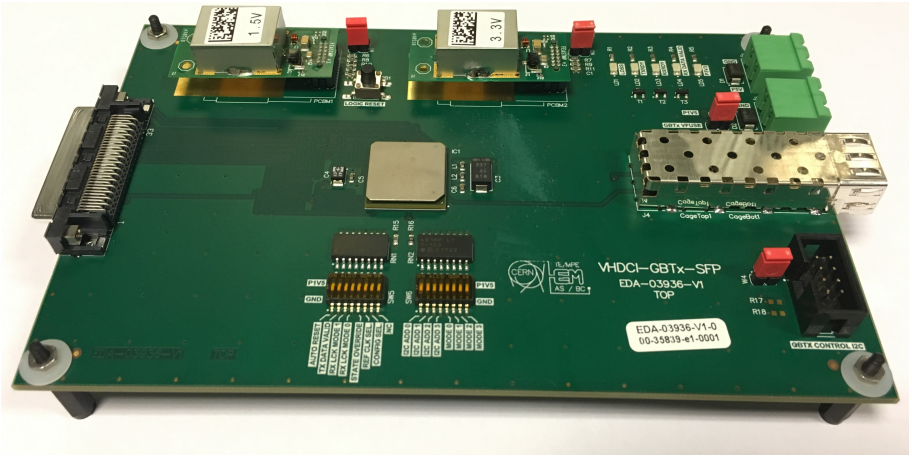


Figure 4.49: The VHDCI-GBTx-SFP board.

The components of the software required for a calibration run are the scan, that steers the calibration in a series of nested loops over the parameter of the chip, and gathers the data, a decoder that processes the raw data into event data, a histogrammer that fills histograms from event data, and the analysis that accumulates the results from the histograms and provides feedback to the scan in several cases. One of the characteristics of the current implementation is the fact that the communication with each front-end at each step of the scan is done in series.

A FELIX-based tuning sequence has been successfully implemented, with the performance measured as a function of the number of front-ends, with up to two GBT links at a time. The scan time measured as user, real or system time, shows a clear linear dependence with the number of front-end chips for both the digital and analogue scans, as shown in Figures 4.51. These measurements were conducted on modules kept in an air cooled and

light-tight dark box shown on Fig. 4.50 for the duration of the scans. The CPU load as a function of number of front-ends shows a steep increase for lower number of chips, and a high value for relatively few front-end chips. A full tuning sequence has been measured to take place in under 15 minutes on a quad. Complementary measurements on the full scale demonstrator will provide valuable insight to the current implementation. As mentioned before, at each step of the scan the settings of each front-end are addressed individually. However, FELIX allows for addressing each front-end independently and in an asynchronous way. Since the tuning of each front-end is a completely independent operation, it means that the tuning of each front-end could be implemented in a fully parallel manner, where each front-end would run in an independent process. In conclusion, the linear dependency with the number of chips is expected, and the optimal scaling would be the one where there is no dependency on the number of front-ends in the scan. This parallelisation of front-end scans is currently under development using the many modules available on the demonstrator stave.

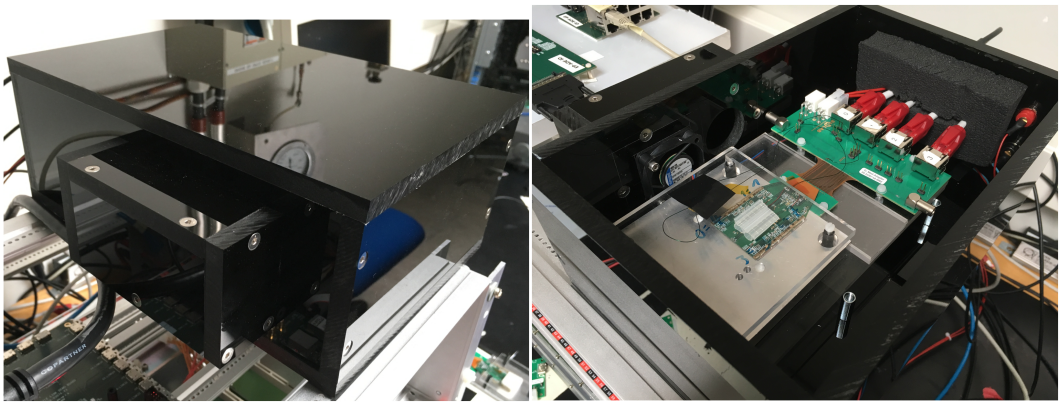


Figure 4.50: Light-tight and air-cooled dark box housing up to two quad modules at a time.

The following scans are needed to tune the FEI4. Each scan is designed to find the optimal value for a given FEI4 register. These registers can be global in nature or specific to each pixel. The pixel specific scans are inherently more computing intensive.

- **Digital**: A digital scan is the simplest scan that can be performed on the FEI4. It injects a signal in the digital circuit of the pixel, and checks for the corresponding hit in the data. Because the output data bandwidth of the FEI4 is limited, the injection is done in a few pixels at a time (known as masks) and in groups of double column (known as DCs). The analysis of the digital scan checks that every single pixel has the required number of hits. If there is a discrepancy in the number of hits in a given set of pixels, it is possible to generate a mask to ignore those pixels in subsequent scans.

- **Analogue**: Injects a known charge in the analogue pixel in groups defined by masks and double columns. The injections are looped over these masks, double columns triggers and number of data acquisitions. An occupancy map is then generated. In the analysis we check for the occupancy at each level, identifying defects in the digital and analogue parts respectively, such as extremely high noise.

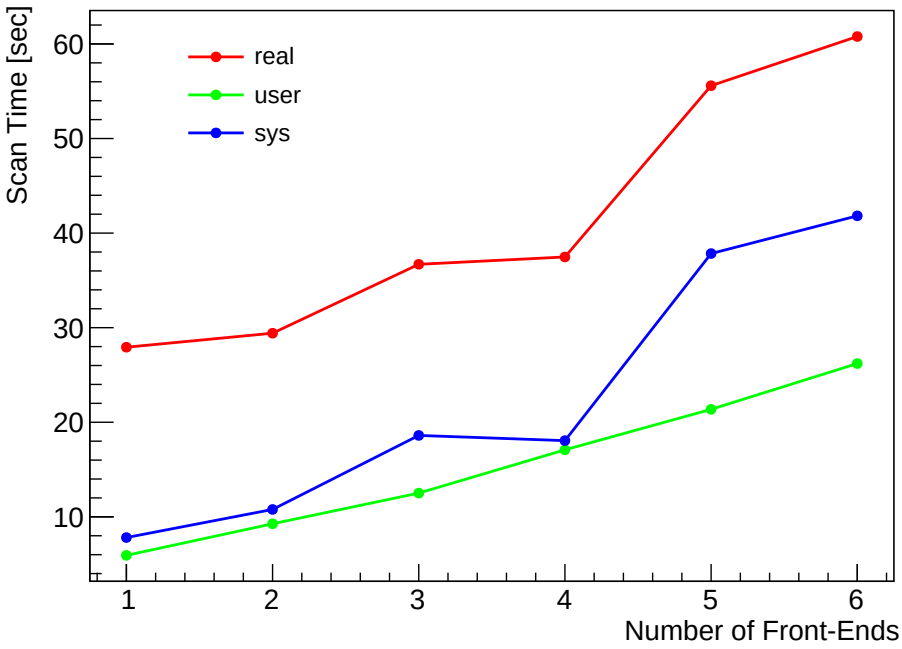


Figure 4.51: Scan time versus number of front-ends conducted on 2 FEI4 quad modules, where "real" pertains to the total time taken to run the scans including any auxiliary processes and "sys" and "user" times are specific to the scan processes.

- **Global threshold DAC calibration**: Tunes the global threshold registers (known as Vthin_ Coarse and Vthin_ Fine) of the FEI4 in order to target a new mean value for the threshold in the chip. This is done by evaluating the mean occupancy of the chip at the target charge. This global threshold register has 16 bits, 8 bits for Vthin_ Coarse and 8 bits for the Vthin_ Fine. The possible values for each one is from 0 to 255. The initial value for Vthin_ Coarse is 0, and for Vthin_ Fine is 50, unless the retuning flag has been set, in which case their pre-existing values will be read from a configuration file. The analysis for this scan provides the next global value for the threshold. The first part of the scan sets the global threshold value, after which the scan loops over the different masks of the chip and the double columns. The charge selected by the user, typically 3000 e^-, is then injected and sends the triggers to the chip and the data is subsequently read out for each chip. The mean occupancy per chip is accumulated at each step of the threshold loop and evaluated by the analysis. If the occupancy of the chip is larger than 50%, the value of the Vthin_ Coarse and Vthin_ Fine is lowered and if the occupancy of the chip is lower than 50% the value of the Vthin_ Coarse and Vthin_ Fine is raised. The algorithm converges when the occupancy of the chip is between 49% and 51% of the 98% of the unmasked pixels. The scan finishes when 98% of the unmasked pixels have been tuned. This scan does not mask pixels at the end and requires high voltage to be applied on the front-ends.

- **Pixel threshold DAC tune**: Similarly to the previous scan, this tunes each pixel threshold register (known as TDAC_ FB) of the FEI4 in order to reduce the threshold dispersion across the chip. This is done by evaluating the occupancy of the pixel at the target charge. This pixel threshold register has 5 bits, with possible values ranging

from 0 to 31. The initial value for every pixel is 16 unless the retuning flag has been set, in which case it will be read from the configuration file. The analysis provides the next value for the next iteration. At the beginning of the scan the threshold value for each pixel is set, the different masks of the chip are looped over, as well as the double columns, a charge is injected and triggers are sent to the chip. Data is then read out for each chip. The occupancy per pixel is accumulated at each step of the threshold loop and evaluated by the analysis. In cases where the occupancy of the pixel is larger than 50% the value of the TDAC_ FB is lowered and if the occupancy of the pixel is lower than 50% the value of the TDAC_ FB is raised. The algorithm converges when the occupancy of the pixel is between 49% and 51% of the 98% of the unmasked pixels. The new value of the pixel threshold register is passed onto the next iteration of the loop. The scan finishes when 98% of the unmasked pixels have been tuned. This scan also does not mask pixels at the end and requires HV to be applied across the front-end.

- **Global preamplifier DAC calibration**: Tunes the global preamplifier register (known as PrmpVbpf) of the FEI4 in order to target a new mean value for the Time Over Threshold (TOT) in the chip. This is done by evaluating the mean TOT of chip at the target charge. This global threshold register has 8 bits, with possible values ranging from 0 to 255. The initial value of the register is 100, unless the retuning flag has been set, in which case they it be read from the configuration file. The scan sets the global threshold, then loops over the different masks of the chip, as well as over the double columns. Charge is injected and triggers are sent to the chip. As before, the data for each chip is then read out. The TOT per chip is accumulated at each step of the threshold loop and evaluated by the analysis. If the average TOT of the chip is larger than 50%, the value of PrmpVbpf is lowered and if the average TOT of the chip is lower than 50%, the value of PrmpVbpf is raised. The algorithm converges when the average TOT of the chip is within 5% of the target TOT for the 98% of unmasked pixels. The new value of the PrmpVbpf register is passed onto the next iteration of the loop. The scan finishes when 98% of the unmasked pixels have been tuned. This scan does not mask pixels and requires HV to be applied.

- **Pixel feedback current (FDAC_ FB) tune**: Similarly to the previous scan, this tunes the feedback current register (FDAC_ FB) per pixel of the FEI4 in order to reduce the TOT dispersion across the chip. This is done by evaluating the mean TOT of the pixel at the target charge. This pixel threshold register has 4 bits, with possible values from 0 to 15. The initial value for every pixel is 8 unless the retuning flag has been set, in which case it will be read from the configuration file. The analysis provides the next value for each pixel to the ongoing loop which sets the FDAC_ FB value for each pixel, looping over the different masks of the chip and double columns. As before charge is injected and the triggers are sent to the chip after which the data is readout for each chip. The TOT per pixel is accumulated at each step of the FDAC_ FB loop and evaluated by the analysis. If the TOT of the pixel is above the target TOT, the value of the FDAC_ FB is lowered. If the TOT of the pixel is lower than the target TOT, the value of the FDAC_ FB is raised. The algorithm converges when the TOT of the pixel is within 5% of the target TOT. The scan finishes when 98% of the unmasked

pixels have been tuned. This scan does not mask pixels at the end and requires HV to be applied.

These tuning scans can be performed via a dedicated GUI, shown on Fig. 4.52 developed to allow for systematic performance checks for available front-end modules.



Figure 4.52: Scan mananger GUI allowing for systematic performance checks on FEI4 modules.

Fig. 4.53 and 4.54 show the occupancy maps produced after digital and analogue scans respectively.

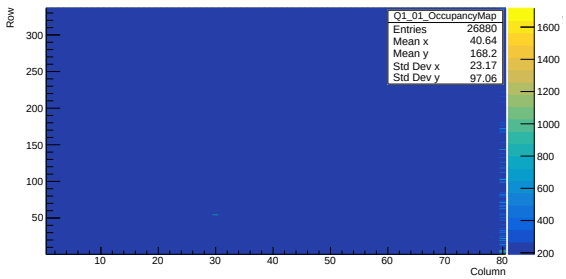


Figure 4.53: Occupancy map produced after a digital scan.

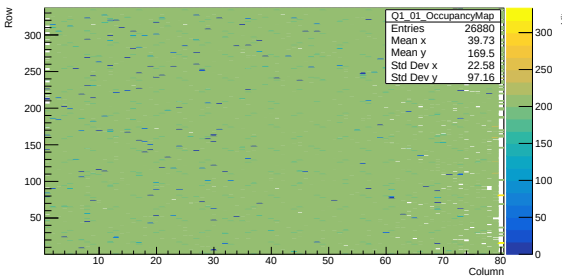


Figure 4.54: Occupancy map produced after a analogue scan.

The tuning sequence of the FEI4 requires the successful completion of these scans. In order to verify the actual threshold per pixel, a threshold scan is required. Much like for the MALTA chip in section 2.5.2, a charge of gradually increasing magnitude is injected onto individual pixels following a pattern governed by masks and double column loops. Naturally, a charge below the threshold will not produce any hits, and a charge above threshold will always produce hits. The threshold is defined as the value at which the pixel fires 50% of the times. This response can be fitted by a sigmoid function, parameters of which yield the

threshold and the noise. Fitting a sigmoid to this S-shaped curve yields the threshold at the inflection point of this fit. Sample S-shaped curves for 3 pixels are shown on Fig. 4.55.

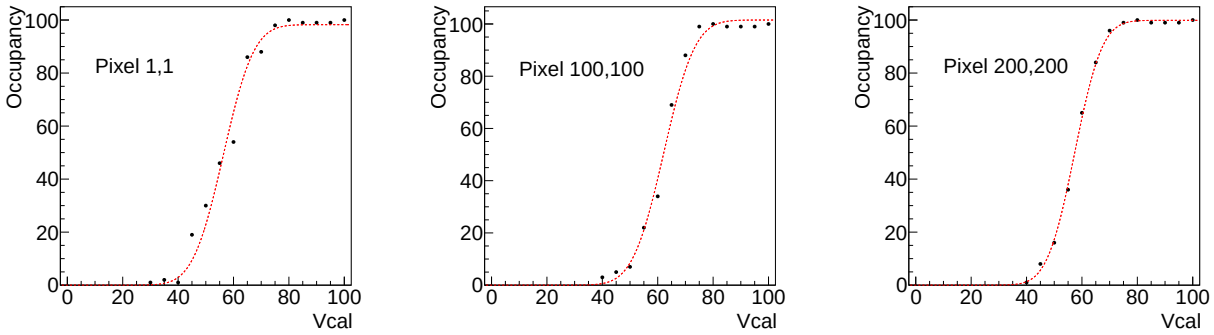


Figure 4.55: Example of S-curves produced for 3 specific pixels during a threshold scan, where V_{cal} is proportional to the injected charge, that is 3000 e^-.

These thresholds are tuned, either globally or per pixel, by their respective scans, and yield a shift in the mean of the threshold distribution in the case of a global tuning, or a narrowing of the threshold dispersion in the case of pixel-level tuning. Such a sequence is shown on Fig. 4.56, Fig. 4.57, where changes in the width of the distribution are negligible but the mean is reduced slightly and Fig. 4.58 where the largest contribution in the optimisation of the threshold dispersion is in the significant narrowing of its width.

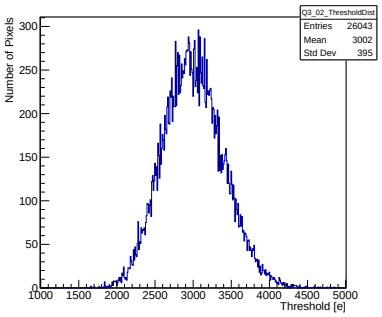


Figure 4.56: Threshold distribution for all unmasked pixels on a single module before any tuning.

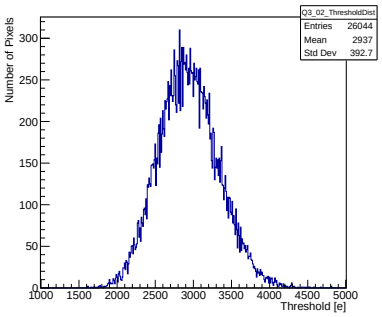


Figure 4.57: Threshold distribution for all unmasked pixels after performing a global threshold tune.

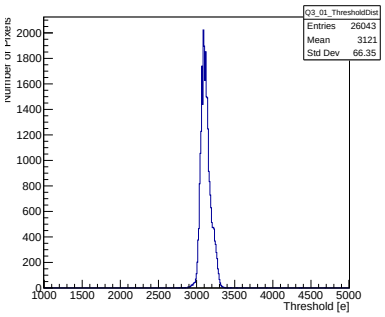


Figure 4.58: Threshold distribution for all unmasked pixels after performing a pixel threshold tune.

A detailed step by step iteration of the pixel-level tuning can be seen on Fig. 4.59, where in 5 steps, the occupancy of the pixels is optimised around a mean occupancy of 100.

4.3.3 Outlook and next steps

The FELIX detector interface for ATLAS will be installed in staged fashion, first into few sub-detectors in the Phase 1 upgrade of ATLAS by 2021, and then into all sub-detectors by the Phase 2 upgrade in 2026. The current implementation achieves the required Phase 1 performance with emulators, and the integration with ITk pixel is progressing well.

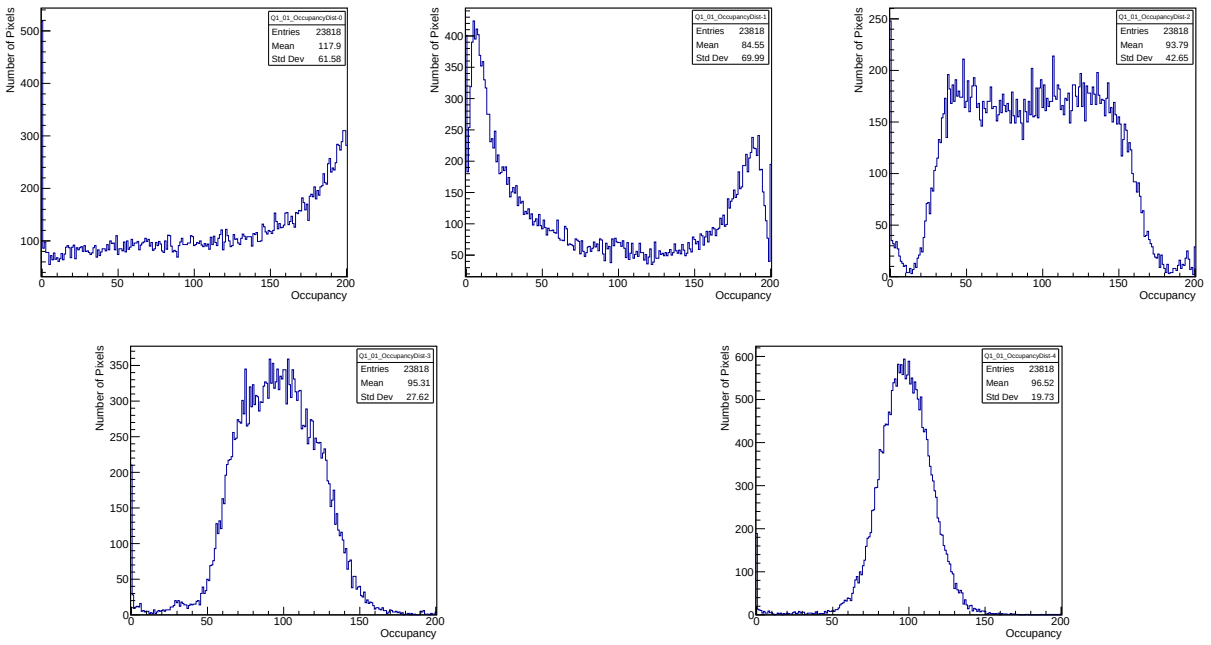


Figure 4.59: Iterative pixel threshold tuning sequence showing a gradually narrower occupancy distribution reaching a mean of 100.

The integration with the next pixel detector prototype chip for the ATLAS ITk pixel, the RD53A read-out chip, has already started. It requires an intermediate PILUP [85] board for data formatting, and allows the read-out through FELIX in FULL-mode. The plan is to develop a full sized demonstrator by the end of 2019 based on the RD53A chip. This will require a 20 fold increase in the data bandwidth to FELIX, thus requiring the installation of a 100 Gb/s data network that is planned for the surface integration site at CERN. This will allow further integration steps with Phase 1 FELIX and SWROD components.

Chapter 5

Summary and Outlook

Over the course of the work conducted towards this thesis, numerous functional areas of the future ITk pixel detector were discussed. This offered an outstanding opportunity to make contributions over a broad spectrum of technologies and scientific instrumentation. A key objective of this thesis was to evaluate the suitability of novel monolithic CMOS sensor designs. To this end numerous characterisations, including the development of a new Edge-TCT facility, a multitude of laboratory measurements and test-beam campaigns were conducted. In a second stage, the stave prototypes onto which such sensors would be held in place underwent thermal characterisations. In order to achieve this, a brand new centralised Thermal Figure of Merit characterisation facility was created.

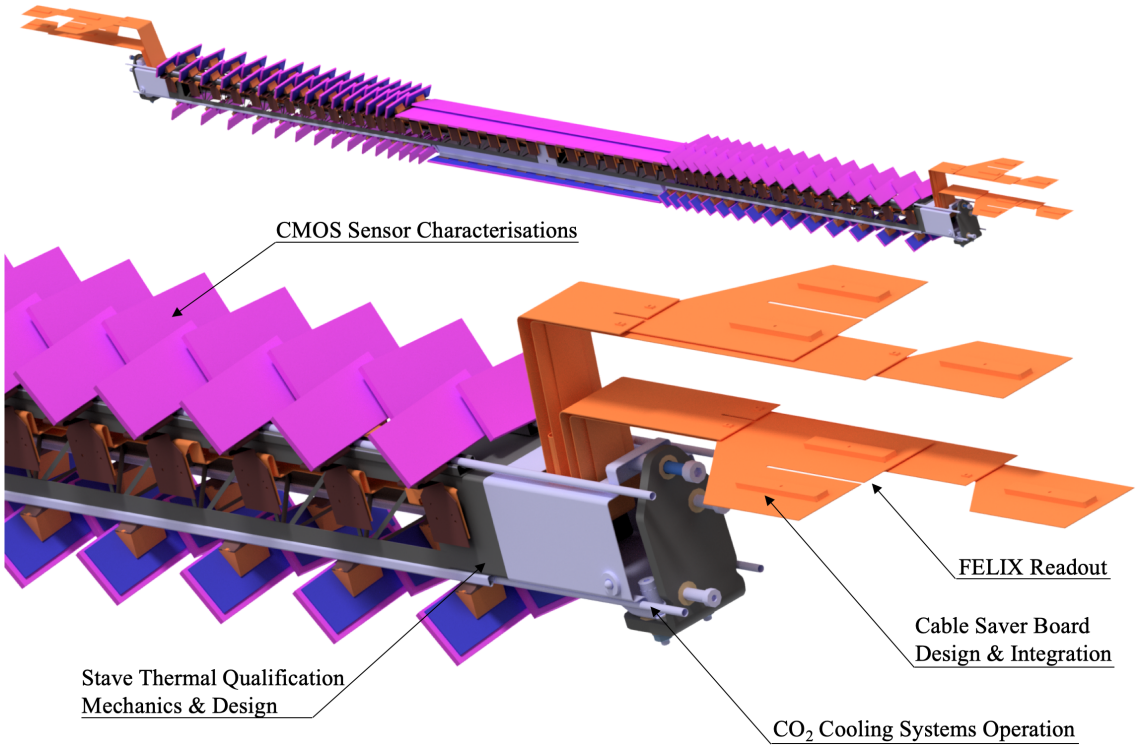


Figure 5.1: CAD representation of a full demonstrator stave, showing the different research domains covered within this thesis.

Additionally, in an effort towards the full integration of demonstrator staves replicating the anticipated ITk pixel design using the technology presently available, a series of custom cable saver boards were developed. Finally, in order to handle the data and the configuration of future front-ends, contributions were made towards the development of a robust FELIX system capable of meeting the specifications required for the Phase 2 upgrade. Fig. 5.1 shows a representation of the closest integrated design at the time of writing of a fully assembled 2-layer pixel stave along with the visual overview of the different domains covered within this thesis. This section distils the key findings, observations and achievements of this thesis.

As part of the sensor characterisation portion of this study, identifying a feasible CMOS design solution for layer 4 of the ITk pixel detector lead to the assessment of three distinct sensor prototypes. These were produced in 180 nm TowerJazz technology integrating a small charge collection electrode for low capacitance including greatly reduced power consumption. The first was a demonstrator sized 2×2 cm^2 chip dubbed MALTA. This chip underwent comprehensive laboratory qualification to evaluate its mode of operation and initial baseline responses to radioactive sources as well as optimisations in the configuration of thresholds and noise levels. One of the features of this design is that there is no threshold setting per pixel. The threshold dispersion was hence identified and observed to be twice that intended by design and this was further increased by a 50% spread after X-ray irradiations. Nevertheless, MALTA was observed to be a functioning chip up to 70 MRad TID and 1×10^{15} n_{eq}/cm^2 fluence irradiations still configurable with reliable slow control communication. Analogue tests showed over 95% of hits being registered within the 25 ns in-time efficiency requirement. This was also measured at low noise levels. Following test-beam campaigns with electrons at the ELSA facility and pions at SPS, resolutions for each type of beam of 14 μm and 4 μm were obtained respectively. It was also observed that the merging logic on-board the chip did not work as expected and hence had to be permanently disabled. Some hits were as a result being lost due to erroneous merging of those true hits with noise hits. Efficiency measurements also clearly outlined pixel corners as being poor charge collection regions post-irradiation. Nevertheless, benefiting from the timing capabilities and full charge collection of unirradiated MALTA chips, a 6 plane telescope was developed allowing for the continued characterisation of future sensors throughout the LHC's two-year long shut down during which the SPS beam lines will not be available for test-beam campaigns. The resolution of this telescope, particularly in the case of an electron beam could be substantially improved through the process of wafer thinning down to 50 μm or less, greatly reducing particle scattering. Additionally the next iteration of the sensors should include threshold adjustment bits to effectively eliminate the currently broad threshold dispersion.

Remedying these shortfalls was a key incentive in the development of a second sensor, the mini-MALTA. This sensor focussed on implementing corrective pixel substructures within the wafer thickness, namely extra-deep p-wells and n-gap designs in order to draw any previously uncollected charge in the pixel corners towards the collection electrodes. The effectiveness of these design strategies were specifically probed through a high resolution efficiency measurement using a 2 μm wide X-ray beam at the Diamond Light Source facility. The results showed that pixel flavours incorporating these wafer substructures indeed compensated for the initial efficiency loss and lead to a high efficiency of >99.7% even for samples irradiated to 1×10^{15} n_{eq}/cm^2. The MALTA readout architecture has aimed towards

a 1 Gbps sampling of the output bit-stream. An asynchronous readout scheme which bypasses the requirement of a distributed clock over the active matrix has made the readout of this relatively high data rate possible. As a result of forgoing the clock distribution, a very low power dissipation of 75 mW/cm^2 and 10 mW/cm^2 for the analogue and digital power respectively was achieved. This drastically reduced dissipation amounts to around 10% of the 700 mW/cm^2 budgeted for final sensor operation, providing a substantial safety margin particularly towards cooling constraints. Hybrid sensor designs will be implemented for the outer layer of the ITk in 2024-26, but these CMOS design candidates will be considered for integration in layer 4 during the 2030's EYETS during which time the pixel detector's inner layers will be replaced.

Finally, a series of chips inherited from the ALICE ITS R&D, known as Investigators, were used to probe pure analogue output of a large matrix comprising of pixels with varying geometries. In order to identify the charge sharing properties of these chips, a new laser E-TCT setup was commissioned, using which trends were identified, in particular that wider charge collection electrodes reduce charge sharing fractions. This is likely due to a more uniform vertical electric field across the epitaxial layer. I-V measurements on those very samples provided insight into the so-called punch-through effect, whereby large potential differences applied across the p-well and substrate could potentially lead to discharges. Punch-through was observed for low-doping samples whereas higher doped wafer samples showed no signs of the punch-through effect at every operational voltage configuration. Gain measurements were conducted on Investigator IBis samples using Fe-55, showing a near-linear decrease in gain with increasing depletion voltage.

Through this work multiple custom carrier boards were developed to house each chip and allow its powering and communication. In the case of the Investigators, a very flexible design optimised for E-TCT was produced to enhance pre-existing designs allowing for a high degree of automation in pixel matrix selection, improved signal integrity and the complete omission for the need of wire-bonding for matrix addressing. For every PCB, all design files have been made available on a repository within the public domain in order to facilitate future and cross-institute developments.

The extensive R&D of these chips has lead to a firm understanding of the TowerJazz 180 nm production chain. Numerous subtle design variations are now possible including changes in wafer masks, transistor sizes, pixel sub-structure variations as well as the understanding of the company's memory structure offerings. This technology has shown a very promising performance for the Phase 2 upgrade of ATLAS and beyond, with considerably reduced material budget. This technology is also part of the Strategic R&D Programme on Technologies for Future Experiments of CERN's EP department. Though a CMOS candidate sensor will not be used for the outermost layer of the pixel detector at the start of Run 3, a technical stop is expected around 2030 during which time the inner most layers will require replacement. This may present an opportunity to reconsider CMOS sensors as a possibility for layer 4. In addition, the MALTA sensor design presents numerous applications beyond that of the HL-LHC, such as being used in imaging devices in space given its low power consumption and subsequently noise levels. Another application being investigated is the idea of creating micro-via channels in the substrate of the MALTA sensor so as to use it

for radioactive contaminants in, for example, the recirculated cooling water in nuclear power plants.

The second part of this study focussed on the assessment of the thermal characteristics of a full set of prototype stave designs produced for consideration towards the ITk tilted and flat layout support geometries. In order to achieve this, a centralised Thermal Figure of Merit measurement facility was constructed from the ground up to provide measurements under the same conditions. This included the development of a custom scalable readout system, dubbed VolTRON. All performance results were summarised and submitted to the ITk mechanics community forming a key engineering input. The findings included the identification of a small number of severely failing designs, namely leading to thermal runaway. These were subsequently associated to under-performing thermal foam forming the bulk of the stave prototypes. Thermal convection effects along with the impact of bi-phase CO_2 bubbles affecting differing surfaces of the staves depending on their orientation were also established.

Beyond the baseline performance measurements, a framework for further characterisation involving failure testing against thermal shocks was also established, laying the path for future systematic studies of the selected designs. Initial thermal shock tests additionally revealed the need for uniformity in the application process of thermal interfaces, particularly in cases where stave designs employed a compacted powder-like iso-graphite, failing which full delamination was likely.

As a result of the cumulative experience gathered in the execution of such thermal characterisations, a large number of aspects have been identified that should be taken into consideration for such studies. In future, firstly, the most significant parameter to control is the thermal pick of the setup, not solely that of the stave sample. This is a particularly problematic issue that can lead to the premature boiling of CO_2 through insufficiently insulated as well as needlessly long piping. In order to fully minimise thermal pick up within the setup, the CO_2 cooling unit must be kept as close as possible to the DUT with minimal pipe lengths, all whilst maximising tri-axial piping coverage along the cooling circuit. Physical contact with hardware components other than piping results in unintended thermal pickup due to the significant thermal inertia of components such as flowmeters. Therefore, a CO_2 cooling unit must be chosen which already incorporates a flowmeter in order to reduce the thermal load on the system. A drastic improvement in thermal insulation can be achieved by reducing the measurement volume. This can be accomplished by sandwiching the full assembly (excluding the cooling plant) between thick layers of ArmaFlex insulation foam. The gain in control of the CO_2 vapour state would outweigh the loss of information regarding convective air flow around the sample due to the restricted volume. CO_2 vapour quality measurements could be made in realtime by combining dynamic readings from temperature and pressure probes at the inlet and outlet of the sample as well as that of the mass flowmeter. These data could then be processed within REFPROP, a software used by the CERN Gas Group which matches the data to reference pressure-enthalpy tables in order to estimate the gas' vapour quality. Fine control of the vapour quality can allow new assessments such as subjecting a 30 cm short stave sample to the CO_2 vapour quality expected towards the outlet of a full length stave. This could allow worst case scenario testing, as the CO_2 vapour is expected to comprise of nearly entirely gas as it reaches the stave outlet,

having gradually picked up increasing amounts of heat. Finally, given the importance of understanding the thermal behaviour of staves, an increasing number of institutes have now begun to develop independent setups to characterise the samples in-house post-production, replicating similar aspects of the work done in this study, including in cases the adoption of the VolTRON readout system. In order to conduct cross-check measurements between these different setups, a reference stave would be highly beneficial. This could comprise of a simple rectangular block of low thermal conductivity aluminium with a 3 mm stainless steel pipe passing across its length, with an offset from its central axis in order to amplify temperature differences across the thickness of the block. This reference stave could provide insights into systematic uncertainty variations from one setup to the next.

Additionally, contributions were made within the scope of the ITk Demonstrator Programme which aims to produce full scale functional replicas of fully mounted and serviced staves. Towards this end, every cable saver board necessary for interfacing between delicate flex cables and more robust instrumentation cabling was designed, assembled and integrated. This aimed to facilitate the R&D of the wider powering and electrical integration and subsequent testing necessary to qualify these demonstrators. This comprised of multiple unique PCBs catering to 5 distinct demonstrators, ranging from pure thermal as well as fully electrical module testing staves. In the case of electrical modules these CSBs were optimised for data rates of 320 MB/s differential signals, all whilst maintaining 100 Ω impedance and catering to the high current loads expected from the low voltage power supply lines. These CSBs also offer the feature of serial powering testing. The completion of this task has allowed for extensive electrical characterisations and full system tests to be performed on all ITk stave demonstrators.

Finally, the FELIX readout system to be integrated ATLAS-wide by 2024 is undergoing substantial development in order to improve its level of interface for the ITk. To facilitate this development, a new readout board was designed which can connect between up to 16 electrical modules from an electrical demonstrator and route these data and configuration signals to an optical link to directly interface with the FELIX system. This board has been fully qualified and is now in use as a standard link in the readout testing chain. ASIC configuration and scan timing measurements were also conducted which lead to the confirmation that parallel scanning is a necessity due to the cumulative scan time increase for serial scans. The next steps will involve updating this chain so as to interface with the next generation of readout chips known as the RD53A. By the end of 2020, it is expected that a full sized demonstrator based on this new ASIC will be constructed. This will require a 20 fold increase in the data bandwidth to FELIX, thus requiring the installation of a 100 Gb/s data network that is planned for the surface integration site at CERN.

The work conducted within this thesis has been part of a wider R&D phase of the ATLAS ITk pixel. Each activity has been motivated by the fact that further developments are needed in those areas in order to meet the specifications required for satisfactory operation in the forthcoming era of the HL-LHC. The conclusions from this study have lead to optimisations in radiation hard sensor technology, mechanics prototyping, integration and read-out.

Bibliography

[1] Wikimedia Commons. "File:Standard Model of Elementary Particles.svg — Wikimedia Commons, the free media repository" (2019). Accessed 12-July-2019. https://commons.wikimedia.org/w/index.php?title=File:Standard_Model_of_Elementary_Particles.svg&oldid=341423479

[2] G. Aad et al. *Physics Letters B*, **716**, 1 (2012) 1 – 29. ISSN 0370-2693. http://www.sciencedirect.com/science/article/pii/S037026931200857X

[3] S. et al. Chatrchyan. *Physics Letters B*, **716**, 1 (2012) 30–61. ISSN 0370-2693. http://dx.doi.org/10.1016/j.physletb.2012.08.021

[4] Lyndon Evans et al. *Journal of Instrumentation*, **3**, 08 (2008) S08001–S08001.

[5] The ALICE Collaboration. *Journal of Instrumentation*, **3**, 08 (2008) S08002–S08002.

[6] The ATLAS Collaboration. *JINST*, **3** (2008) S08003.

[7] The CMS Collaboration. *JINST*, **3** (2008) S08004.

[8] The LHCb Collaboration. *JINST*, **3** (2008) S08005.

[9] "Accelerator Statistics" (Accessed: 04 July 2019). http://acc-stats.web.cern.ch/acc-stats/lhc/overview-panel

[10] Esma Mobs. General Photo. https://cds.cern.ch/record/2197559

[11] Joao Pequenao. "Computer generated image of the whole ATLAS detector" (2008). https://cds.cern.ch/record/1095924

[12] Claudia Marcelloni. "View of the ATLAS Cavern with the EndCap Magnets in place" (2007). https://cds.cern.ch/record/1058286

[13] "ATLAS Magnet System" (Accessed: 12 July 2019). https://atlas.cern/discover/detector/magnet-system

[14] A. La Rosa. *Journal of Instrumentation*, **11**, 12 (2016) C12036–C12036.

[15] M. Garcia-Sciveres et al. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, **636**, 1, Supplement (2011) S155 – S159. 7th International Hiroshima Symposium on the Development and Application of Semiconductor Tracking Detectors.

[16] Joao Pequenao. "Computer generated image of the ATLAS inner detector" (2008). http://cds.cern.ch/record/1095926

[17] O Solovyanov et al. *Journal of Physics: Conference Series*, **119**, 2 (2008) 022012.

[18] "ATLAS muon spectrometer: Technical Design Report". Technical Design Report ATLAS. CERN, Geneva (1997). https://cds.cern.ch/record/331068

[19] "ATLAS, Run 2 Luminosity Public Results" (Accessed: 25 July 2019). https://twiki.cern.ch/twiki/bin/view/AtlasPublic/

[20] Zbynek Drasal et al. *Nucl. Instrum. Meth.*, **A910** (2018) 127–132.

[21] Apollinari G. et al. "High-Luminosity Large Hadron Collider (HL-LHC): Technical Design Report V. 0.1". CERN Yellow Reports: Monographs. CERN, Geneva (2017). http://cds.cern.ch/record/2284929

[22] Frederick Bordry, Michael Benedikt, Oliver Bruning, John Jowett, Lucio Rossi, Daniel Schulte, Steinar Stapnes & Frank Zimmermann. "Machine Parameters and Projected Luminosity Performance of Proposed Future Colliders at CERN". Technical Report arXiv:1810.13022, CERN, Geneva (2018). * Temporary entry *. http://cds.cern.ch/record/2645151

[23] T. Kawamoto et al. "New Small Wheel Technical Design Report". Technical Report CERN-LHCC-2013-006. ATLAS-TDR-020 (2013). ATLAS New Small Wheel Technical Design Report. https://cds.cern.ch/record/1552862

[24] "ATLAS, Run 2 Radiation Simulation Public Results" (Accessed: 25 July 2019). https://twiki.cern.ch/twiki/bin/view/AtlasPublic/

[25] Leonardo Rossi et al. "Pixel detectors: from fundamentals to applications". Particle Acceleration and Detection. Springer, Berlin (2006). https://cds.cern.ch/record/976471

[26] P. Grenier. *Physics Procedia*, **37** (2012) 874 – 881. ISSN 1875-3892. Proceedings of the 2nd International Conference on Technology and Instrumentation in Particle Physics (TIPP 2011). http://www.sciencedirect.com/science/article/pii/S1875389212017865

[27] ATLAS Collaboration. "Technical Design Report for the ATLAS Inner Tracker Pixel Detector". Technical Report CERN-LHCC-2017-021. ATLAS-TDR-030, CERN, Geneva (2017). https://cds.cern.ch/record/2285585

[28] J. et al. Beringer. *Phys. Rev. D*, **86** (2012) 010001.

[29] Malte Backhaus. "High bandwidth pixel detector modules for the ATLAS Insertable B-Layer". Ph.D. thesis (2014). Presented 30 Oct 2014. http://cds.cern.ch/record/1696901

[30] Christian Riegel. "Performance tests of depleted CMOS sensors for application at HL-LHC". Ph.D. thesis (2018). Presented 01 Oct 2018. http://cds.cern.ch/record/2671033

[31] Anna Peisert. "Silicon microstrip detectors", pages 1–79. https://www.worldscientific.com/doi/abs/10.1142/9789814360333_0001

[32] W. Shockley & W. T. Read. *Phys. Rev.*, **87** (1952) 835–842. https://link.aps.org/doi/10.1103/PhysRev.87.835

[33] Michael Moll. "Radiation damage in silicon particle detectors: Microscopic defects and macroscopic properties". Ph.D. thesis, Hamburg U. (1999). https://inspirehep.net/literature/513308

[34] M. Moll. *IEEE Transactions on Nuclear Science*, **65**, 8 (2018) 1561–1582. ISSN 1558-1578.

[35] Gregor Kramberger. "Signal development in irradiated silicon detectors". Ph.D. thesis (2001). Presented 21 Jun 2001. http://cds.cern.ch/record/1390490

[36] Smodis B. et al. *The JSI TRIGA Mark II Reactor, Slovenia. International Atomic Energy Agency (IAEA): IAEA.*

[37] et al. Chiesa Davide. *The European Physical Journal Plus*, **135**, 4. ISSN 2190-5444. http://dx.doi.org/10.1140/epjp/s13360-020-00334-7

[38] Carl Wheldon David Parker. *Nuclear Physics News*, **Vertex2016** (2018) 15–20.

[39] "TowerJazz" (Accessed: 04 June 2019). http://TowerJazz.com

[40] "Arduino" (Accessed: 04 June 2019). https://store.arduino.cc/mega-2560-r3

[41] I. Berdalovic et al. *Journal of Instrumentation*, **13**, 01 (2018) C01023–C01023. https://doi.org/10.1088/1748-0221/13/01/c01023

[42] M. Šuljić. *Journal of Instrumentation*, **11**, 11 (2016) C11025–C11025.

[43] M. Munker et al. *Journal of Instrumentation*, **14**, 05 (2019) C05013.

[44] "CERN, Engineering & Equipment Data Management Service (EDMS)" (Accessed: 24 July 2019). https://edms.cern.ch/project/CERN-0000181566

[45] Martin Green & Mark J. Keevers. *Progress in Photovoltaics: Research and Applications*, **3** (1995) 189 – 192.

[46] "Edmund Optics: Advantages of Using Beam Expanders" (Accessed: 25 August 2019). https://www.edmundoptics.com/resources/application-notes/lasers/advantages-of-using-beam-expanders/

[47] "Huber Unichiller 007w-H" (Accessed: 04 June 2019). https://www.huber-online.com

[48] "Cividec Instrumentation" (Accessed: 04 June 2019). https://cividec.at

[49] Stefan Ritt. *NIM A*, **623** (2009) 486–488.

[50] "First Sensor AG" (Accessed: 04 June 2019). https://www.first-sensor.com

[51] Sven Wonsak. "Sensor Preparation" (2015). 1st TCT Workshop at DESY.

[52] Ettore Zaffaroni et al. *PoS*, **Vertex2016** (2017) 063.

[53] Nuclear Energy Agency. "Joint Evaluated Fission and Fusion File". www.oecd-nea.org

[54] J. Baudot et al. "First test results Of MIMOSA-26, a fast CMOS sensor with integrated zero suppression and digitized output". In "2009 IEEE Nuclear Science Symposium Conference Record (NSS/MIC)", pages 1169–1173 (2009).

[55] Kiehn Moritz, et al. "Proteus beam telescope reconstruction" (2019). https://zenodo.org/record/2586736#.Xqh5jS2B1hE

[56] M. Dyndal et al. *Journal of Instrumentation*, **15**, 02 (2020) P02005–P02005. https://doi.org/10.1088/1748-0221/15/02/p02005

[57] S. et al. Spannagel. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, **901** (2018) 164–172. ISSN 0168-9002. http://dx.doi.org/10.1016/j.nima.2018.06.020

[58] Claus Kleinwort. *Nucl. Instrum. Meth.*, **A673** (2012) 107–110.

[59] M. Mironova et al. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, **956** (2020) 163381. ISSN 0168-9002. http://www.sciencedirect.com/science/article/pii/S0168900219316195

[60] M. Munker et al. *Journal of Instrumentation*, **14**, 05 (2019) C05013–C05013. https://doi.org/10.1088/1748-02211405c05013

[61] "M. Aleksa et al.". "Strategic R&D Programme on Technologies for Future Experiments". Technical Report CERN-OPEN-2018-006, CERN, Geneva (2018).

[62] "ArmaFlex Insulation" (Accessed: 24 June 2019). https://corporate.armacell.com/en/

[63] "IBL CO2 cooling system operation and maintenance manual" (Accessed: 30 August 2019). https://edms.cern.ch/document/1501950/2

[64] "Rheonik Mass flow meter model RHM015TAP1PM0M0N1N" (Accessed: 24 June 2019). https://www.rheonik.com/

[65] "FIREROD Cartridge Heater" (Accessed: 24 June 2019). https://www.watlow.com

[66] TDK. "NTC thermistors for temperature measurement. Glass-encapsulated sensors, standard type. Series B57550G1" (2014). https://en.tdk.eu/inf/50/db/ntc_13/NTC_Glass_enc_sensors_G550_G1550.pdf.

[67] Cirris. "Touch1". https://www.cirris.com/61-product-information/91-touch-1

[68] "Python serial port library". https://github.com/pyserial/pyserial

[69] "ROOT webpage". https://root.cern.ch

[70] A. Sharma C. Solans. "TFM of SLIM Flat Barrel Quad Size Prototypes". Technical Report 1763878 v.1 (2017). https://edms.cern.ch/document/1763878/1

[71] A. Sharma C. Solans. "TFM of Short I-Beam Baseline prototypes". Technical Report 1743103 v.2 (2017). https://edms.cern.ch/document/1743103/2

[72] A. Sharma C. Solans. "TFM of Alpine cell prototype". Technical Report 1742900 v.1 (2016). https://edms.cern.ch/document/1742900/1

[73] A. Sharma C. Solans. "TFM of Wuppertal stave prototypes". Technical Report 1743239 v.3 (2017). https://edms.cern.ch/document/1743239/3

[74] A. Sharma C. Solans. "TFM of EndCap Half Ring prototype". Technical Report 1742929 v.2 (2018). https://edms.cern.ch/document/1742929/2

[75] A. Sharma C. Solans. "TFM of the Berkley Inner Rings Prototype 0". Technical Report 2026763 v.1 (2018). https://edms.cern.ch/document/2026763/1

[76] D. Giugni. "ITk Pixel Local Support Design Specifications". Technical Report 1534572 v.4 (2018). https://edms.cern.ch/document/1534572/4

[77] "NIST Reference Fluid Thermodynamic and Transport Properties Database (REFPROP)" (Accessed: 27 August 2019). https://www.nist.gov/srd/refprop

[78] Lehmann Niklaus et al. "Prototype Chip for a Control System in a Serial Powered Pixel Detector at the ATLAS Phase II Upgrade". page 026 (2018).

[79] ATLAS Collaboration. "Technical Design Report for the Phase-II Upgrade of the ATLAS TDAQ System". Technical Report CERN-LHCC-2017-020. ATLAS-TDR-029, CERN, Geneva (2017). https://cds.cern.ch/record/2285584

[80] "CERN S-LINK homepage". http://hsi.web.cern.ch/HSI/s-link/

[81] P. Moreira et al. https://cds.cern.ch/record/1091474

[82] Martín Lesma et al. *Journal of Instrumentation*, **12** (2017) C02020–C02020.

[83] Orlandi Stefano et al. "Optimization of Shielded PCB Air-Core Toroids for High-Efficiency DC–DC Converters". volume 26, pages 2073 – 2080 (2009).

[84] Maurice Garcia-Sciveres. "The RD53A Integrated Circuit". Technical Report CERN-RD53-PUB-17-001, CERN, Geneva (2017). https://cds.cern.ch/record/2287593

[85] A Gabrielli et al. *Proceedings Of Science*, **321**. Sixth Annual Conference on Large Hadron Collider Physics (LHCP2018).