UNIVERSITY
OF OSLO

Milou van Rijnbach

Radiation Hard Multi-chip Sensor to Module Integration for Monolithic and Hybrid Pixel Detectors

Thesis submitted for the degree of Philosophiae Doctor

Department of Physics
Faculty of Mathematics and Natural Sciences

The European Organization for Nuclear Research (CERN)
ATLAS Pixel Group

2023

© Milou van Rijnbach, 2023

Series of dissertations submitted to the
Faculty of Mathematics and Natural Sciences, University of Oslo
No. 2699

ISSN 1501-7710

All rights reserved. No part of this publication may be
reproduced or transmitted, in any form or by any means, without permission.

Cover: UiO.
Print production: Graphic center, University of Oslo.

Preface

This thesis is submitted in partial fulfillment of the requirements for the degree of *Philosophiae Doctor* at the University of Oslo. The research presented here was conducted at the University of Oslo and at CERN, under the supervision of professor Heidi Sandaker, Heinz Pernegger, and Petra Riedler.

The thesis is a collection of three papers, presented in chronological order of writing. The papers are preceded by an introductory chapter that relates them to each other and provides background information and motivation for the work. The author of this thesis is the first author of all three papers.

Milou van Rijnbach
Geneva, November 2023

Acknowledgements

As the writing of my thesis comes to an end, it is inevitable that I am writing these acknowledgements from a slightly hurried place. I apologise in advance if I have forgotten to mention somebody, I thank you all for being part of my journey.

Firstly, I would like to express my sincere gratitude to my university supervisor, Heidi Sandaker, whose support and comforting words have been invaluable throughout my PhD journey. Her guidance during the final phase of thesis writing was particularly appreciated and she helped me in gaining confidence as a researcher. I'm immensely grateful to my CERN supervisors, Heinz Pernegger and Petra Riedler, for granting me the opportunity to pursue my PhD at CERN. I would like to thank Heinz for welcoming me in the ATLAS pixel group, for his guidance, support and the fruitful discussion we have had. I would like to thank Petra for letting me be part of the EP R&D work package on Module Development. I value our conversations and I'm grateful to Petra for teaching me how to place my research topics within a broader context. I would like to thank her for the continuous support she provided, especially during challenging times. I would like to thank Ole Rohne for helping me in my first year of my PhD, especially in setting up and defining my research and university-related work. I would like to thank him, Stefan Guindon, and Abhishek Sharma for their valuable comments during my thesis writing process.

This thesis is a culmination of extensive collaborative work, building on the contributions of numerous colleagues. I owe a lot of gratitude to those who laid the foundation for my research and I acknowledge their invaluable contributions. During my PhD journey, my colleagues played a vital role, consistently offering me their support and assistance. In this context, I would like to express my gratitude to Andrea Gabrielli for taking on the role of my daily supervisor during my first year of the PhD. I would like to thank him for the incredible patience he exhibited, especially as I had a significant learning curve to overcome in the topic of particle detectors. I would like to express my sincere appreciation to Abhishek Sharma for his support and guidance in my work for the work package, throughout my ATLAS qualification task, and for always making time to address my questions. I want to extend a special appreciation to Carlos Solans for his continuous help and, particularly, his support in making our research results relevant to a broader audience. I would like to express my gratitude to my colleagues at both CERN and beyond, Julian, Ignacio, Leyre, Dominik, Florian H., Florian D., Katharina, Justus, Giuliano, Vlad, Martin, with whom I had a pleasure to work with, from whom I learned a lot and who made work an enjoyable place. A special thanks to Leyre, for being my "anti-social" companion

and introducing me to the addictive habit of doing sports during lunch time.

Lastly, I owe a special thanks to Valerio Dao. While we may have exchanged jokes about the length of my acknowledgments dedicated to him, it is only fair to say that my gratitude towards him goes beyond a few sentences. I would like to thank Valerio for taking me on the MALTA journey with him, learning me all the intricacies of data analysis, hardware, particle physics and more (usually at quite a high pace). I want to express my appreciation for his ability to transform challenging situations into positive ones and for consistently lifting my spirits when I felt lost in my research. Valerio has been an invaluable source of knowledge (and gossip), and the depth of learning I've gained from him during my PhD has exceeded my expectations.

I would like to thank my dear friends and roommates, Sarah, Vaiva, and Marine, for making me always feel at home in Geneva. A special thanks goes to my dear friend Sinah, who made my time in Geneva especially wonderful.

As I am writing these last pages of my thesis back home, in the Netherlands, it is even more evident for me that the efforts of the past years have been fruitful due to the support I have received from my family and friends here. I would like to give a special thanks to Sophie, Deline, Esha, Marjolijn, Jane, Chanan, Gregory, Manon, Jolien and all others who have been part of this journey with me. A special thanks goes to my very dear friends Nora and Pleun, for always showing interest in my research, their immense loving words and encouragements when it was needed. I am so happy to call you my best friends.

A very special thanks goes to Hicham. His unconditional love, support, encouragement, and the (more than) occasional pep-talks have allowed me complete this thesis while maintaining a reasonable level of mental sensibility. He has stayed by my side during moments of desperation, dinners while I was checking my Skype messages, and many evenings and weekends of working. He has also always been the first person to celebrate any victory, small or big. I could not have wished for a better support. Thank you for everything.

Finally, a thanks to my family, to whom I owe everything. My parents, Pieter Jan and Desirée, and my siblings, Justus and Sophie, have been my biggest supporters during my whole life and have always believed in me. Three years ago (at the same spot as where I am sitting now) they encouraged me to embark on the PhD journey, whereas I had big doubts if I could even manage to do it. Finalizing the final words of my thesis at home, with them around me, has been a full circle moment.

Abstract

Within the scope of the high luminosity upgrade of the LHC, the ATLAS experiment is undergoing a significant upgrade with the development of the Inner Tracker (ITk). MALTA, a Depleted Monolithic Active Pixel Sensors (DMAPS) in 180 nm Tower Semiconductor CMOS imaging technology, was developed with the aim of exploring its suitability in ATLAS ITk and other high-energy physics experiments. This work centers around the latest prototype of the MALTA family, MALTA2. The results in this work show that MALTA2 exhibits a notable reduction in RTS (Random Telegraph Signal) noise, which is attributed to the implementation of a cascode front-end and by increasing the size of specific transistors. Compared to its predecessor, MALTA, at similar operating threshold (340 e^-), the results show an improvement in the tails of the noise distribution for MALTA2, with a decrease of the standard deviation of approximately 40%. This work shows that non-irradiated MALTA2 samples on Czochralski substrates can achieve efficiencies of 99% and an average cluster size of 2 pixels at low threshold settings (150 e^-). In these conditions, a timing resolution of σ_t=1.7 ns can be obtained, where more than 98% of the hits are collected within 25 ns. This study demonstrates that MALTA2 sensors, fabricated on Czochralski substrates and subjected to a backside metallisation post-processing procedure, can operate at high fluence levels ($>3\times10^{15}$ n_{eq}/cm^2). Superior performance at these fluence levels is found on samples with very high doping of the n^- layer. At an operating threshold of 110 e^-, these samples achieve a maximum efficiency of 98% and an average cluster size of 1.7 pixels. In these conditions an RMS of the time difference distribution equal to 6.3 ns can be obtained, with 95% of the clusters being collected within a 25 ns time frame. This work reviews the architecture and performance of the MALTA telescope which features a spatial resolution of σ_s=4.1±0.2 µm, based on the linear regression approach, and a track timing resolution of σ_t=2.1 ns. This research presents the testing outcomes of the first generation multi-chip module assembly featuring the MALTA chip. The results demonstrate successful chip-to-chip data transmission through CMOS drives without appreciable distortion of the signal. Versatile coating and embedding layers for multi-chip modules are explored to ensure durability, dependability, and resilience in high-energy experiments. This work demonstrates that the encapsulant Sylgard 186 can safeguard wire-bonds against potential mechanical damage, catering to both monolithic and hybrid pixel detectors, and that the usage does not affect the discriminator's pixel threshold. Additionally, the work reveals that Parylene N is an effective coating layer for safeguarding hybrid modules against high voltage. Moreover, preliminary groundwork is performed to assess the application of Parylene N as a protective barrier against humidity and accelerated aging for both monolithic and hybrid detectors.

Sammendrag

I forbindelse med oppgraderingen av LHC til høy luminositet gjennomgår ATLAS-eksperimentet en betydelig oppgradering av sin Inner Tracker (ITk). MALTA, en DMAPS-sensor (Depleted Monolithic Active Pixel Sensors) laget i 180 nm Tower Semiconductor CMOS-teknologi, har blitt utviklet med sikte på å utforske dens egnethet i ATLAS ITk og andre høyenergifysikkeksperimenter. Arbeidet presentert i denne avhandlingen er sentrert rundt den nyeste prototypen i MALTA-familien, MALTA2. Resultatene viser at MALTA2 har en betydelig reduksjon i RTS-støy (Random Telegraph Signal), som skyldes implementeringen av en kaskode-frontend og en økt størrelse på spesifikke transistorer. Hvis man sammenligner med forløperen, MALTA, ved tilsvarende signalterskelverdi (340 e^-), viser resultatene en forbedring i halene av støyfordelingen og en reduksjon i standardavviket på omtrent 40%. Dette arbeidet viser at ubestrålte MALTA2 sensorer på Czochralski-substrater kan oppnå en effektivitet på 99% og en gjennomsnittlig pikselklyngestørrelse på 2 piksler ved lave signalterskelverdier (150 e^-). Under disse forholdene kan man oppnå en tidsoppløsning på σ_t=1,7 ns, der mer enn 98% av pikseltreffene samles inn i løpet av 25 ns. Denne studien viser at MALTA2-sensorer produsert på Czochralski-substrater og med en baksidemetallisering etter fabrikering kan fungere ved høye bestrålingsnivåer ($>3\times10^{15}$ n_{eq}/cm^2). Overlegen ytelse ved disse bestrålingssnivåene er funnet for prøver med svært høy doping av n^--laget. Ved en signalterskel på 110 e^- oppnås en maksimal effektivitet på 98% og en gjennomsnittlig pikselklyngestørrelse på 1,7 piksler. Under disse forholdene kan man oppnå en RMS for tidsdifferansefordelingen på 6,3 ns, og 95% av klyngene samles inn innenfor en tidsramme på 25 ns. Dette arbeidet gjennomgår også arkitekturen og ytelsen til MALTA-teleskopet, som har en romlig oppløsning på σ_s=4,1±0,2 µm, ved bruk på lineær regresjon, og en tidsoppløsning på σ_t=2,1 ns. Denne forskningen presenterer testresultatene fra den første generasjonen av multibrikkemoduler med MALTA detektorer. Resultatene viser at dataoverføring fra brikke til brikke med hjelp av CMOS teknologi er vellykket og uten nevneverdig forvrengning av signalet. Coating- og innkapslingslag for flerbrikkemoduler er utforsket for å sikre holdbarhet, pålitelighet og robusthet i høyenergi-eksperimenter. Dette arbeidet viser at innkapslingsmaterialet Sylgard 186 kan beskytte wire-bonds mot potensielle mekaniske skader, og at de er egnet for både monolittiske og hybride piksel-detektorer, og at bruken ikke påvirker diskriminatorens signalterskel. I tillegg viser arbeidet at Parylene N er en effektiv coating for å beskytte hybridmoduler mot høy spenning. Videre er det utført et grunnarbeid for å kunne studere anvendelsen av Parylene N som en beskyttende barriere mot fuktighet og akselerert aldring for både monolittiske og hybride detektorer.

Contents

Preface **i**

Acknowledgements **iii**

Abstract **v**

Sammendrag **vii**

Contents **ix**

List of Figures **xi**

1 Introduction **1**

2 High Energy Physics at CERN **3**
2.1 The Large Hadron Collider 3
2.2 The High Luminosity LHC 6
2.3 ATLAS Inner Tracker Upgrade 7
2.4 Conclusion 9

3 Silicon Pixel Detectors **11**
3.1 Fabrication of Silicon Sensors 11
3.2 Front-End Electronics of Pixel Detectors 13
3.3 Radiation Damage in Silicon Detectors 19
3.4 Hybrid Pixel Detectors 23
3.5 Monolithic Pixel Detectors 27
3.6 Conclusion 29

4 State of the Art: MALTA Monolithic CMOS Sensor **31**
4.1 MALTA Process Technology 32
4.2 MALTA Pixel 34
4.3 Mini-MALTA Demonstrator 39
4.4 MALTA on Czochralski Substrates 42
4.5 MALTA2 46
4.6 Conclusion 50

5 Silicon Pixel Detector Modules **53**
5.1 Interconnection Technologies 53
5.2 Modularization of the MALTA Sensor 58
5.3 Coating and Embedding Studies for Pixel Detectors 59

5.4 Characterization of the MALTA multi-chip module 74
5.5 Conclusion 76

6 Conclusion and Outlook 79

7 Introduction to Papers 83

Papers 86

I Radiation hardness and timing performance in MALTA monolithic pixel sensors in TowerJazz 180 nm 87

II Performance of the MALTA Telescope 97

III Radiation Hardness of MALTA2 Monolithic CMOS Sensors on Czochralski Substrates 111

Bibliography 129

List of Figures

2.1 Schematic image of the overall layout of the LHC and its relation to the city of Geneva (grey) and the surrounding region. The Super Proton Synchrotron (SPS) accelerator (indicated by the smaller ring) sends the beam to the LHC. Image is obtained from Ref.[19]. 4

2.2 Cumulative luminosity per day delivered to ATLAS during stable beams and for high energy *pp* collisions. Image obtained from Ref.[28]. 5

2.3 Schematic image of the layout of the ITk described in the Strip TDR referred to as ITk inclined. Only one quadrant and only active detector elements are shown. The four outermost layers (blue) compose the Strip detector, the five innermost layers (red) compose the pixel detector. The horizontal (z) axis is the axis along the beam line with zero being the IP. The vertical (R) axis is the radius measured from the interaction region. The outer radius is set by the inner radius of the barrel cryostat that houses the solenoid and the electromagnetic calorimeter. Image is obtained from Ref.[44]. 8

3.1 Probability density function f(δ/x) for 500 MeV pions in silicon of different thicknesses (640, 320, 160, and 80 µm), normalised to 1 at the most probable value (MPV) Δp/x. Image is obtained from Ref.[64]. 15

3.2 Schematic block diagram of a typical front-end processing chain in pixel detectors for HEP applications. Image is obtained from Ref.[72]. 17

3.3 Consequences of the creation of deep energy levels to the operation of semiconductor detectors. Process (a) shows the existence of of shallow charged defects, process (b) illustrates that defects can trap and detrap free carriers and (c) shows that defects act as generation-recombination centers. Electrons and holes are denoted by *e* and *h*. Image is obtained from Ref.[20]. 21

3.4 Change of the full depletion voltage (U_{dep}) of a 300 μm thick silicon sensor and its absolute effective doping (N_{eff}) versus the normalized fluence (ϕ_{eff}), immediately after irradiation. Image is obtained from Ref.[78]. 22

3.5 Cross section of an exemplary individual hybrid pixel cell. The bottom layer represents the sensing layer, or the passive pixel sensor (light blue), and the top layer represents the read-out chip (dark blue). The two layers are connected through a ball-like feature (grey), or bump-bond. The specifications of the CMOS circuitry of the readout chip are not shown, but they are represented by the alternating orange and blue blocks. Image is obtained from Ref.[80]. 24

3.6 Picture of a dummy ATLAS ITk pixel module consisting of the module flex (dark green base layer), wire-bonds (located on left and right side of the module), passive components, data connector (white feature in the center of the module) and power connector (bottom center). 26

3.7 Top: Cross section of an exemplary MAPS with charge collection in an epitaxial layer (light-blue) mainly by diffusion. The n-well (orange) acts as the charge collection node. The other n-wells are shielded by the p-well (dark-blue). Bottom: Cross section of an exemplary DMAPS. Multiple wells (orange and blue) on high resistive substrate allow complete embedding of the CMOS electronics in the charge collecting deep n-well (light orange). Images are obtained from Ref.[80]. 28

4.1 Cross section of the Tower Semiconductor 180 nm CMOS imaging technology fabricated on a high resistivity p-type epitaxial layer. Left images shows the standard process, the right image shows the standard modified process (STD) where an additional n^- layer is introduced. Indicated are the n-well collection electrode, p-well and n-wells hosting the NMOS and PMOS, respectively, and the deep p-well. Images are not drawn to scale and are adapted from Ref.[64]. 33

4.2 Layout of a pixel in the MALTA chip with a pixel pitch of 36.4 µm. The analogue and digital part are well separated and shielded from each other and the collection electrode to avoid crosstalk. Image is obtained from Ref.[64]. 34

4.3 Layout of the analog front-end circuit of the MALTA sensor. Image is obtained from Ref.[64]. 35

4.4 Layout and main building blocks of the full MALTA chip. The pixel matrix is divided into 8 sectors (S0-S7) with different sensor and front-end designs. The power pads are positioned at the bottom, left and right side of the matrix. The image indicates the position of the digital periphery, DACs, and LVDS output. Image is obtained from Ref.[64]. 36

4.5 Organisation of a double column in the MALTA digital readout architecture. Each pixel hit generates a signal on its corresponding line of the pixel bus that is transmitted asynchronously at the time of the discriminator output. The pixel bus is shared by alternating pixel groups, divided in blue or red groups. All of the blue groups connect to same line. Image is obtained from Ref.[64]. 37

4.6 In-pixel 2D efficiency map projected over a 2×2 pixel matrix for non-irradiated (left) and irradiated at 5×10^{14} 1 MeV n_{eq}/cm^2 (right) STD MALTA samples measured with a 180 GeV proton beam. Results are shown for sector 3 of the MALTA matrix, corresponding to a maximum extension of the deep p-well. The chips are operated at -15 V bias voltage and -20°C. Operational threshold corresponds to 250 (left) and 400 (right) electrons. Images are obtained from Ref.[90]. 39

4.7 Layout of the mini-MALTA pixel matrix with 8 pixel sub-groups (S0-S7) of 8×16 pixels. The left half of the pixel matrix features enlarged transistors, the right half features the standard MALTA front-end. The bottom row of sub-groups features the STD pixel flavour, the second row features the XDPW flavour, the third row implements different reset mechanisms (standard or PMOS), and the fourth row features the NGAP flavour. Image is adapted from Ref.[88]. 40

4.8 Cross section of the process modifications for the MALTA sensor: NGAP flavour with low dose n^- layer removed at the edge of the pixel (left), and XDPW flavour with extra deep p-well at the edge of the pixel (right). Images are not drawn to scale and are adapted from Ref.[88]. 41

4.9 2D efficiency map for a non-irradiated (left) and irradiated at 1×15 1 MeV n_{eq}/cm^2 (right) mini-MALTA sample measured with a 2.5 GeV electron beam. The pixel matrix is divided into sub-groups featuring different process modifications and desings. The left half of the pixel matrix features enlarged transistors, the right half features the standard MALTA front-end. The bottom row of sub-groups features the STD pixel flavour, the second row features the XDPW flavour, the third row implements different reset mechanisms (no data available) and the fourth row features the NGAP flavour. The empty bins represent dead pixels. The chips were operated at -6 V SUB voltage and -20°C, and were tuned to low threshold. Images are obtained from Ref.[88]. . . . 42

4.10 Efficiency versus threshold mean for neutron irradiated mini-MALTA samples at 1×10^{15} 1 MeV n_{eq}/cm^2 (top) and 2×10^{15} 1 MeV n_{eq}/cm^2 (bottom) measured with a 2.5 GeV electron beam. The chips were operated at -6 V SUB voltage and -20°C. Different sensor regions are presented: STD (circles), XDPW (triangles) and NGAP (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). Images are obtained from Ref.[88]. 43

4.11 Sensor efficiency and cluster size for samples irradiated at 1×10^{15} 1 MeV n_{eq}/cm^2 on Cz substrate with different pixel flavour (STD, NGAP, and XDPW). Left figures shows the efficiency (top) and cluster size (bottom) as a function of substrate voltage with constant threshold across the samples. Right figures show the efficiency (top) and cluster size (bottom) as a function of threshold with constant substrate voltage of -50 V. Images are obtained from Ref.[11]. 44

4.12 Top image: A schematic sketch of the MALTA beam telescope where the six tracking planes (P1 - P6), the devices under tests (DUTs), and the scintillator are indicated. Bottom image: Top view of the MALTA telescope. Equipment mounted on the main stage are inside yellow lines, cold box and DUTs positioned on DUT stage are inside the red lines. 45

4.13 Layout of the MALTA2 matrix featuring 224×512 pixels. The power pads are positioned at the bottom, left and right side of the matrix. The image indicates the position of the slow control, DACs, LVDS output, and merging and arbitration logic. Image is obtained from Ref.[16]. 47

4.14 Layout of the MALTA2 cascode analog front-end. Image is obtained from Ref.[16]. 47

4.15 Threshold (left) and noise (right) distributions of a MALTA (black) and MALTA2 (red) sample. Both samples are non-irradiated, fabricated on an epitaxial layer, NGAP, 300 µm thick, high doping n^- layer, and operated at -6 V SUB bias. Images are obtained from Ref.[98]. 48

4.16 Threshold (left) and noise (right) distributions of non-irradiated (black) and neutron irradiated (red) at 3×10^{15} n_{eq}/cm^2 MALTA2 samples. Both samples are fabricated on an epitaxial layer, XDPW, 100µm thick, low doping n^- layer and operated at -6 V SUB bias. The threshold and noise scans are obtained for the lowest threshold configuration achievable in that chip. Images are obtained from Ref.[98]. 49

5.1 Image of an assembled and wire-bonded MALTA quad-chip board at the Bondlab at CERN. The wire bonding was performed through aluminium wedge wire-bonding. The bonding tool hovers above the four MALTA chips, arranged side by side on the PCB board. 55
5.2 Electrical connectivity tests of (fused silica) test structures bonded on a 50 µm flexible PCB mounted on a glass substrate. The test structures are flip-chip bonded to the PCB, using ACF from Dexerials [106] as the interconnection technology. 56
5.3 SEM image of copper nanowires grown on two squared MALTA pads (88×88 µm). The diameter of the nanowires is approximately 100 µm, with a length of several micrometers. 57
5.4 Image of fully assembled and functional MALTA quad-chip board module using ultrasonic aluminium wedge wire bonding. The three (secondary) left MALTA chips forward data to the primary chip (right) where all data is finally is read out. For the third secondary MALTA chip (far-most left), the presence of wire-bonds are indicated with the blue dashed line. 59
5.5 Two stages of encapsulation of the wire-bonds of RD53A modules. Top: Module prior to encapsulation. The blue dashed boxes indicate the areas with wire-bonds, also located on the right side of the module. Bottom: Image of a fully encapsulated module. All wire-bonds are encapsulated with Sylgard 186, the transparent glue-like epoxy. The blue dashed boxes indicate the encapsulated area, also present on the right side of the module. No evidence of encapsulant leakage beyond the module's boundaries is apparent, signifying the success of the procedure. 62
5.6 The ITk Pixel Outer Barrel RD53A demonstrator features the installation of 18 RD53A modules onto the longeron. These modules are effectively protected by the application of the encapsulant Sylgard 186, a transparent adhesive substance that is carefully deposited along the outer edges of the modules (indicated by the blue arrow), specifically targeting the wire-bond areas. The Sylgard 186 encapsulation ensures the safety and integrity of the wire-bonds on the modules. 64
5.7 Image of MALTA dual-chip module with Sylgard 186 encapsulation of the wire-bonds. The wirebonds are located at the bottom, left and right side of the single MALTA chips. 65
5.8 Threshold distributions of four (S0-S4) sectors of one MALTA chip in the MALTA dual chip module. The quoted mean (µ) threshold value is obtained by using a Gaussian fit to the core of the threshold distributions separately, from which the dispersion (sigma) is extracted. After Sylgard encapsulation (right), no appreciable difference in the threshold can be observed. As only a sub-set of pixels are scanned, threshold and sigma values are rounded. 65

5.9 Left: Image of an unsuccessful Parylene coated silicon dummy after cross-cut test, where a lot of flaking of the Parylene can be observed. Right: Image of a successful Parylene coated silicon dummy after cross-cut test. Clean cut lines are visible due to the use of a primer (Silane) prior to coating. 68

5.10 IV curve indicating the electrical characteristics of a Parylene coated high-voltage test structure that has undergone X-ray irradiation equivalent to 660 Mrad. The black curve represents a single scan conducted immediately after irradiation. The red curve illustrates the same scan (100 times), but after subjecting the kit to 100 cycles of thermal cycling (10x -45°C,+45°C + 1x -55°C,+ 60°C). The green curve reflects the IV measurements (100 times) following 25 cycles of thermal cycling between -55°C,+ 60°C. The blue curve represents the last IV measurements taken after 25 cycles of thermal cycling under the same conditions. The final IV curve was only performed 1.5 times due to a significant increase in leakage current, indicating a breakdown of the sensor. All IV measurements covered the voltage range of 0-900 V. . . . 69

5.11 Before (top) and after (bottom) Parylene-coated wirebond samples containing 3 samples (red, green, blue) with 200 short (3 mm) and 200 long (10 mm) 25µm Al wire bonds. The mean (µ) pull strength and their standard deviation (σ) are indicated, quantifying the mechanical strength of the deposited Parylene. After Parylene coating, the mean pull strength shifts for long wire bonds from approximately 6.5 g to 9.5 g. For short wire bonds, the shift is more dramatic, as the mean pull strength goes from approximately 11 to 23.5 g. 70

5.12 Three stages of Parylene coating of a RD53A module. Top left: Module prior to masking and Parylene coating. Top right: Module after masking and prior to Parylene coating. The features that are masked are indicated by the blue arrows. Blue dicing tape is located on the pick-up points (circular features) and backside of the module. Black cap protects the power connector with a silicon glue. Bottom: Module after Parylene coating and demasking. A small degree of delamination (flaking) around data connector can be observed. 71

5.13 X-ray scans of four front ends of an ITk pixel module (ID: CERNPixQ28–HPk-HPk150 quad Module) before (top four) and after (bottom four) Parylene coating. No difference in the results of the scan can be observed after coating, indicating that the masking, Parylene coating, demasking, and other related tests have been successful and did not harm the module. These measurements were performed by colleagues from the ATLAS ITk module testing group at CERN. 73

5.14 Time of arrival of the leading hit in the cluster with respect to a scintillator reference for the first, second, third and fourth chip of a MALTA quad module. The timing distributions include a scintillator (~0.5 ns) and a sampling ($3.125/\sqrt{12}$=0.9 ns) jitter and were obtained by illuminating the four chips separately. For every chip, the peak value of the timing distribution is quoted and the sigma is extracted by fitting a Gaussian to the core of the distribution. 75
5.15 2D Efficiency plot of the full matrix of four chips of the MALTA quad-chip board, represented by the secondary chips 1 (bottom left), 2 (top right), and 3 (top left) and the primary chip (bottom right). 77

Chapter 1

Introduction

The pursuit of high luminosity is a crucial endeavor in particle physics, enabling scientists to explore uncharted territories and unveil new insights into the fundamental building blocks of our universe. CERN [1], or the European Organization for Nuclear Research, as a global leader in particle physics research, is embarking on the High Luminosity upgrade [2] of their Large Hadron Collider (LHC) [3] to enhance the capabilities of its experiments and facilitate groundbreaking discoveries. The ATLAS experiment [4], situated at the LHC, is an essential component of the quest for scientific breakthroughs. The ongoing ITk upgrade [5] of the ATLAS detector plays a vital role in ensuring its readiness to handle the unprecedented luminosities foreseen at the High Luminosity LHC (HL-LHC). One of the foremost challenges faced for future collider experiments is the development of radiation-hard silicon pixel detectors. These detectors must withstand the intense radiation environment and maintain their functionality and performance over extended periods. Radiation hardness, one of the main topics in Chapter 3, is a key requirement to ensure accurate and reliable data collection in the presence of high particle fluxes and harsh radiation conditions [2].

MALTA [6], a Depleted Monolithic Active Pixel Sensor (DMAPS) [7] fabricated in Tower Semiconductor 180 nm CMOS imaging technology [8], was initially developed for potential use in the ATLAS experiment at the HL-LHC upgrade and possible integration in other future high-energy physics experiments, such as the Upstream Tracker upgrade of LHCb [9]. Therefore it was designed to fulfill stringent requirements such as radiation hardness, high hit-rate capability, and low mass, further discussed in Chapter 4. Although it did not ultimately find integration within the ATLAS ITk, the MALTA chip remains highly relevant for other future collider experiments. Consequently, extensive studies have been conducted on boosting its performance and modularizing the MALTA chip to achieve a large area and lightweight multi-chip module, discussed in greater in detail in Chapter 5.

The core of this thesis centers around three papers. They are presented in chronological order of writing, that encompass a substantial portion of the research efforts that were made for MALTA in this thesis. The first (Chapter I) and third paper (Chapter III) primarily focus on test beam results obtained using the MALTA and MALTA2 sensors for radiation hardness studies. The first paper specifically compares samples fabricated on high-resistivity epitaxial silicon with Czochralski substrates before and after neutron irradiation. It also presents initial findings from MALTA2, highlighting the improvements in the front-end. The third paper is entirely dedicated to demonstrating the radiation

hardness of MALTA2 on Czochralski substrates. It emphasizes the performance in terms of timing resolution and efficiency before and after neutron irradiation, with a special focus on the process modifications that were implemented to achieve radiation hardness $>3\times10^{15}$ 1 MeV n_{eq}/cm^2. The second paper (Chapter II) provides a comprehensive overview of the architecture and performance of the MALTA telescope, that was utilized in the test beam campaigns conducted between 2021 and 2023 at the SPS North Area. This paper highlights how the MALTA telescope leverages the best qualities of the MALTA sensor: a full prototype (large area), high granularity, self-triggering capability, and excellent spatial and timing resolution. The collaborative efforts within the MALTA working group have resulted in the publication of numerous proceedings and papers in peer-reviewed journals, including Ref.[10–17], of which the author of this thesis is a co-author.

Throughout this thesis, a comprehensive exploration of the aforementioned topics will be presented, covering the theoretical background, advancements, results obtained during the work performed for this thesis, challenges, and potential solutions within each domain. The research conducted aims to contribute to the broader scientific community's understanding of high-energy physics, radiation hardness, and the development of advanced silicon pixel detectors for future collider experiments. By addressing the complexities of the High Luminosity upgrade, the ITk upgrade of ATLAS, and the unique characteristics of hybrid and monolithic detectors, this thesis aims to provide valuable insights and pave the way for further advancements in the field.

Chapter 2

High Energy Physics at CERN

The LHC was built with the ambitious motivation of pushing the boundaries of the energy frontier in particle physics, enabling scientists to explore the fundamental building blocks of the universe at unprecedented scales. By studying the Standard Model (SM) [18], further discussed in section 2.1.2, of particle physics at the TeV scale, the LHC seeks to unlock deeper insights into the nature of matter and the fundamental forces that govern the cosmos. CERN is one of the world's leading particle physics research centers. It is located in Geneva, Switzerland, and is home to the LHC, which is the world's largest and most powerful particle accelerator. At its core, CERN's mission is to study the fundamental properties of matter and the forces that govern it. This involves accelerating subatomic particles to incredibly high speeds and colliding them into one another [1].

This chapter provides an introduction to the LHC, outlining its purpose and the physics motivations behind its construction. We also delve into the rationale for enhancing the LHC's luminosity to achieve the HL-LHC. Additionally, we focus on the upgrade of one of the four major experiments conducted at the LHC, namely ATLAS, exploring the reasons and objectives behind this upgrade. The content of this chapter is primarily based on references [2, 5, 19, 20], if not stated otherwise.

2.1 The Large Hadron Collider

The LHC consists of a circular tunnel that is 26.7 km in circumference and lies beneath the French-Swiss border, illustrated in Figure 2.1. The tunnel houses two parallel beams of protons (pp) that travel in ultrahigh vacuum in opposite directions and are guided by superconducting magnets. The beams are accelerated to nearly the speed of light by a series of radio-frequency cavities [1]. The LHC is equipped with four large detectors positioned at different locations around the ring. Two of these detectors, ATLAS and CMS [21], are general-purpose detectors as they investigate the largest range of physics possible. The other two, LHCb [22] and ALICE [23], are specialized detectors; LHCb specializes in precise measurements of the quark flavor sector [24], while ALICE is optimized for studying heavy ion physics [25] and the properties of the Quark-Gluon Plasma (QGP) [26]. Once the particles are accelerated to their desired energy levels, they are made to collide at various points around the accelerator, where the large detectors are used to capture and analyze the particles and their resulting decay products [19].

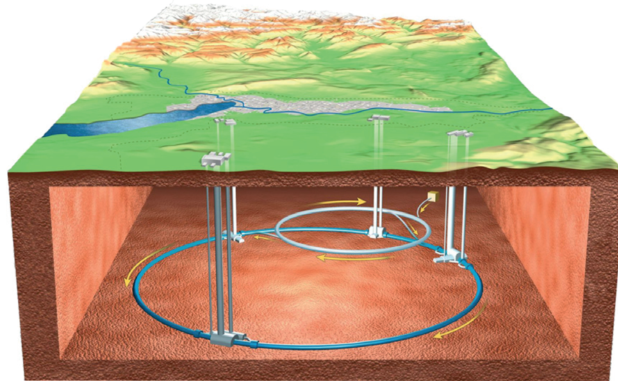


Figure 2.1: Schematic image of the overall layout of the LHC and its relation to the city of Geneva (grey) and the surrounding region. The Super Proton Synchrotron (SPS) accelerator (indicated by the smaller ring) sends the beam to the LHC. Image is obtained from Ref.[19].

2.1.1 Luminosity

Each proton beam in the LHC consists of approximately 2800 bunches, with each bunch containing around 10^{11} protons. Each beam circulates the LHC ring 11245 times per second. Due to the high revolution frequency and the large number of bunches, there are approximately 40 million bunch crossings per second at the center of each experiment. This, coupled with the high number of protons per bunch and small beam sizes at the Interaction Points (IP), the place where the bunch gets squeezed down to 16×16 µm for collisions to take place, leads to exceptionally high collision rates between protons. A critical parameter for a collider is the luminosity, which represents the number of collisions per unit area per second at a specific IP [3]. Luminosity directly correlates to the collision rate observed by the experiment whereas the integrated luminosity refers to the total number of potential collisions per unit area over a given time period. When multiplied by the cross section, the integrated luminosity provides the total number of collisions observed by a particular experiment during that time [19]. In 2018 a record annual integrated luminosity of 65 fb^{-1} was achieved, indicated in Figure 2.2 where the cumulative luminosity for a multi year period is shown for the ATLAS detector. In the ongoing Run 3 period from 2022 to 2025, the LHC aims to further increase the total integrated luminosity, with a current goal of reaching 350 fb^{-1} by the end of Run 3 [27], surpassing the initial LHC target of about 300 fb^{-1}.

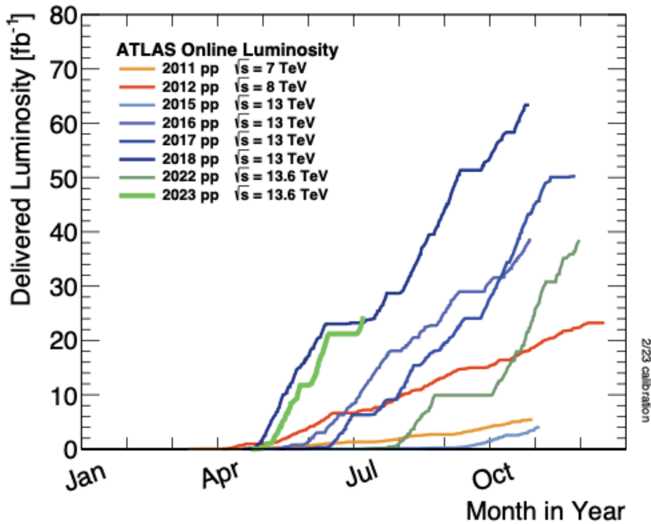


Figure 2.2: Cumulative luminosity per day delivered to ATLAS during stable beams and for high energy *pp* collisions. Image obtained from Ref.[28].

2.1.2 The Standard Model

The collisions that occur at the LHC serve as experimental tests for the SM of particle physics, which represents our current understanding of the fundamental constituents of nature (elementary particles) and their interactions with one another. The SM of particle physics is a comprehensive theoretical framework that explains the properties and interactions of all known elementary particles. It unifies the electromagnetic and weak forces into a single electroweak interaction at energies above a few hundred GeV. At higher energies the strong force and electroweak meet. The SM does not account for gravitational interactions [18]. While the SM has demonstrated remarkable success in describing the observed phenomena and making accurate predictions, such as the discovery of the Higgs boson by the ATLAS and CMS collaborations in the year 2012 at the LHC [29], there are several phenomena it fails to explain. One of the unresolved aspects is the existence of dark matter, which has been inferred through gravitational effects in astronomical observations. Dark Matter (DM) does not emit light, making it undetectable by telescopes searching for electromagnetic emissions. Additionally, the SM does not provide an explanation for the matter-antimatter asymmetry in the Universe. Despite the expectation that the Big Bang should have produced equal amounts of matter and antimatter, our Universe appears to be predominantly composed of matter [30].

Another fundamental question is whether a unified theory exists that encompasses all fundamental forces (strong, electromagnetic, weak, and gravitational). Such a theory would need to incorporate the SM and its extensively tested predictions accumulated over the past six decades. The pursuit of a unified theory is driven by the desire to understand the connections between these forces and their underlying fundamental principles. The theoretical challenge, known as the hierarchy problem, arises from the significant disparity in strength between the gravitational force and the other fundamental forces (by a factor of 10^{17}). This disparity remains an unsolved puzzle within the current understanding of physics [31][19].

2.2 The High Luminosity LHC

To fully exploit the physics potential of the LHC and to address the unresolved challenges within the SM, CERN established the HL-HLC project in late 2010. The project's targets include achieving a peak luminosity of 5×10^{34} cm^{-2}s^{-2} through a levelling and attaining an integrated luminosity of 250 fb^{-1} per year, with a goal of accumulating around 3000 fb^{-1} over a span of approximately 12 years after the upgrade. This integrated luminosity is roughly ten times higher than what the LHC was initially predicted to achieve in its original configuration [2]. The HL-LHC introduces a new configuration that incorporates several groundbreaking innovations. Key advancements include the utilization of 11-12 T superconducting magnets, compact superconducting cavities for precise beam rotation control, novel technologies for beam collimation, and the implementation of 100-meter-long high-power superconducting links with minimal energy dissipation. Achieving these innovations requires extensive research and development efforts conducted on a global scale over several years [2].

The HL-LHC upgrade, applicable to nearly all major LHC experiments, encompasses a broad range of physics objectives. By increasing the number of collisions, new possibilities for observing rare processes and particles will emerge. The primary driving force behind these goals is the significant boost in integrated luminosity. The major physics objectives of HL-LHC fall into five categories: improved measurements of the Standard Model, searches for physics beyond the SM (BSM) [32], exploration of flavor physics involving heavy quarks and leptons, investigation of the properties of the Higgs boson, and studies of QCD matter at high density and temperature. Understanding the Higgs boson and its connection to electroweak symmetry breaking remains a primary focus. In the realm of flavor physics, LHCb, ATLAS, and CMS will collectively test the unitarity of the Cabibbo-Kobayashi-Maskawa matrix [33], while ATLAS and CMS will specifically measure properties of the top quark, the fermion with the highest known mass and Yukawa coupling [34]. HL-LHC will also contribute to the understanding of Parton Distribution Functions (PDFs) [35] by measuring various Standard Model processes involving jets, top quarks, photons, and electroweak gauge bosons in their final states. The production of jets and

photons in heavy-ion collisions will serve as probes for QCD perturbation theory [36]. The high-energy collisions of HL-LHC also hold the potential to detect BSM phenomena, such as baryogenesis [37], DM, answers to the flavor problem, neutrino masses, and insights into the strong CP problem [38].

2.3 ATLAS Inner Tracker Upgrade

The ATLAS detector is one of the two general-purpose particle detectors at the LHC at CERN, which is designed to study the collisions of high-energy proton beams. The ATLAS detector has been instrumental in the aforementioned discovery of the Higgs boson and in the search for BSM particle physics. The increase in luminosity for the HL-LHC comes at the cost of introducing significant challenges to the detector technologies, particularly at the IP. To address these challenges, the ATLAS experiment has adopted a staged approach for upgrading various components such as the calorimeters [39], muon spectrometer [40], and the trigger and data acquisition (DAQ) systems [41]. As an intermediate step, dedicated upgrades were implemented during Long Shutdown 1 (LS1), a period of planned maintenance and consolidation of the entire LHC complex. During LS3, the current Inner Detector (ID) [42] will be completely replaced by a new all-silicon Inner Tracker (ITk) [5]. The ITk is designed to deliver performance comparable to or even surpassing the present ATLAS tracker, while operating under the more demanding conditions of high luminosity data-taking [5][43].

The design of the ITk detector is centered around an all-silicon tracker, comprising a pixel subsystem surrounded by a strip subsystem, within a 2 T solenoidal magnetic field. In comparison to the current tracker, the ITk incorporates new technologies across most of its components. The pixel system comprises five barrel layers and endcap rings. Refer to Figure 2.3 for a schematic depiction of the ITk Inclined layout, which corresponds to the configuration detailed in Ref.[44]. This layout pertains to the arrangement where the sensors are inclined in the forward section of the barrel layers. The two innermost pixel layers are designed for replacement after accumulating around 2000 fb^{-1} of data, while the three outermost layers will remain operational until the end of the experiment. The strip system encompasses four barrel layers and six end-cap disks, extending up to a maximum radius of 1 m. Compared to the present ATLAS tracker, the ITk system is significantly more complex, with a tenfold increase in strip channels and a sixty-fold increase in pixel channels. The ITk incorporates approximately five times the number of modules compared to the current tracker, with the surface area of each module being three times larger for the strip part and seven times larger for the pixel part. Despite the increased surface area and complexity of, among others, the pixel module assembly and integration (see Chapter 3.4), the ITk achieves a reduction in material compared to the present system. Various advanced solutions have contributed to this material reduction, including improved services utilizing CO_2 cooling and serial powering, advanced materials for local supports, and thinner modules and inclined sections have also

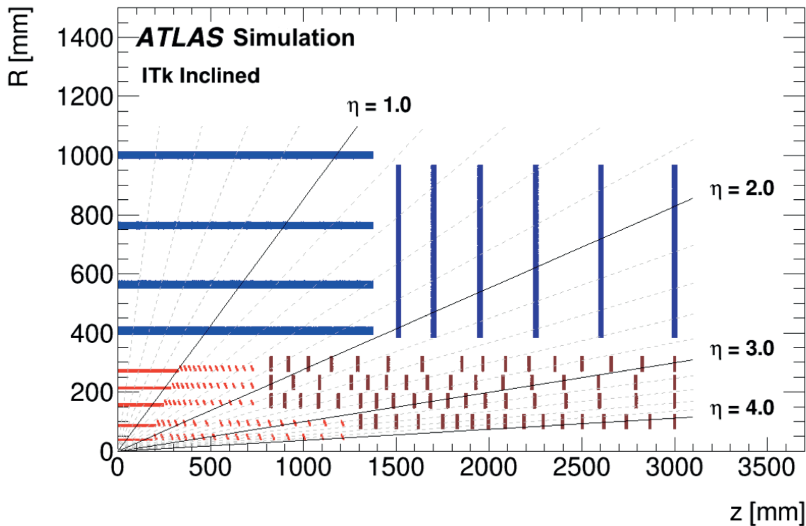


Figure 2.3: Schematic image of the layout of the ITk described in the Strip TDR referred to as ITk inclined. Only one quadrant and only active detector elements are shown. The four outermost layers (blue) compose the Strip detector, the five innermost layers (red) compose the pixel detector. The horizontal (z) axis is the axis along the beam line with zero being the IP. The vertical (R) axis is the radius measured from the interaction region. The outer radius is set by the inner radius of the barrel cryostat that houses the solenoid and the electromagnetic calorimeter. Image is obtained from Ref.[44].

been employed [5]. The rationale behind minimizing material within a particle detector will be elaborated upon in more detail in Chapter 3. A list of the main requirements for the outermost pixel layer of the ATLAS ITk are listed in Table 2.1. Further explanation of the performance requirements listed in this table will be provided in the subsequent chapters.

The tracking performance [45], the effectiveness and accuracy of the tracking system in reconstructing the paths of charged particles produced in the IP, of the ITk benefits not only from the material reduction. It also benefits from enhanced granularity, improved hermeticity, and an increased number of expected hits-on-track [5]. Redundancy plays a crucial role, not only to address potential module failures over time but also to reduce the reconstruction combinatorics. The layout of the ITk detector is designed such that it can identify charged particles with high efficiency and purity and that it can measure their proprieties with extremely high precision. The ITk is necessary to have the best possible performance of object reconstruction for HL-LHC conditions. These include, but are not limited to, photon conversion reconstruction studies, flavour tagging

performance, pile-up jet rejection and E_T^{miss} performance, electron and muon reconstruction and identification, and study of τ identification using the ITk [5].

Main performance requirement ATLAS ITk	
Detection efficiency [%]	>97
Timing resolution [ns]	25
Particle rate [MHz/mm^2]	1
Radiation tolerance (NIEL) [1 MeV n_{eq}/cm^2]	10^{15}
Radiation tolerance (TID) [Mrad]	50
Power consumption [mW/cm^2]	<500
Material budget [% of x/X_0]	<2

Table 2.1: Main performance requirements for the outermost pixel layer of the ATLAS ITk [5]

2.4 Conclusion

In conclusion, this chapter has provided a brief overview of several key aspects in the field of particle physics, which will serve as the framework for this thesis. The discussion started with the LHC at CERN, which collides protons at extremely high energies in order to unveil the fundamental building blocks of the universe. The SM has served as the theoretical framework that encapsulates our current understanding of elementary particles and their interactions. While the SM has been remarkably successful in explaining a wide range of phenomena, there are still unresolved issues, such as the nature of DM and the matter-antimatter asymmetry. To tackle these unresolved questions and further explore the frontiers of particle physics, CERN launched the HL-LHC project. The HL-LHC aims to significantly increase the number of collisions, luminosity, and integrated luminosity, allowing for more precise measurements and the potential discovery of BSM phenomena. Within the scope of the HL-LHC project, the ATLAS experiment is undergoing a significant upgrade with the development of the ITk. The ITk, based on an all-silicon tracker, will feature improved performance and capabilities to handle the challenging conditions of the HL-LHC. It incorporates advanced technologies to enhance tracking performance and maintain efficient data collection. These advancements will allow us to push the boundaries of knowledge and continue to deepen our understanding of the fundamental laws that govern our universe.

Chapter 3

Silicon Pixel Detectors

Semiconductor detectors have become an integral tool in high-energy physics experiments, enabling groundbreaking discoveries and precise measurements in recent years, as discussed in the previous Chapter (2. Thanks to the rapid progress of the semiconductor industry, there have been constant technological advancements leading to the development of better devices [20]. Among them, silicon-based detectors have gained widespread use in numerous experimental setups of the LHC, such as the LHCb Vertex Locator (VELO) [46] and the CMS pixel detector [47], and are the primary focus of this thesis. Within the realm of silicon detectors, numerous subdivisions exist, ranging from depleted to non-depleted, hybrid to monolithic, planar versus 3D, and more [48]. While this thesis cannot cover all these variations, it is essential to acknowledge their vast diversity, each type possessing unique characteristics tailored to specific target applications. Currently, strip [49] and pixel detectors [48] are the two mainstream types of silicon detectors. In the case of the former, the sensitive area is segmented in only one dimension, while pixel detectors provide two-dimensional division and are usually rectangular. This chapter provides an outline of the concepts of hybrid and monolithic silicon pixel detectors, emphasizing design characteristics relevant to this thesis. Nonetheless, it will be evident that certain topics covered, like signal formation, extend beyond pixel detectors and are applicable to a broader range of detectors as well. We start with a description of the main fabrication techniques of silicon wafers, followed by an overview of signal formation and typical front-end components of silicon sensors. As solid-state detectors consist of a crystalline medium, their performance heavily relies on lattice properties, such as the bandgap, carrier mobility, and lifetime, which can be affected by radiation exposure. The main types of radiation damage and its implication on the performance on detectors will be described and will form the theoretical basis for the main results of this thesis (presented in Chapter I, II, III). Finally, the two main pixel detector technologies are introduced, accompanied by corresponding examples that serve as the main devices under investigation in this thesis's research. The content of this chapter is primarily based on Ref.[5, 20, 48, 50, 51], if not stated otherwise.

3.1 Fabrication of Silicon Sensors

Silicon is extracted from quartzite, a relatively pure form of sand (SiO_2). Through various processing steps a high pure polycrystalline silicon can be obtained, also known as Electronic Grade Silicon (EGS), which can be used as the starting material for growing large single crystals using one of two main techniques: Float Zone (FZ) [52] and Czochralski (Cz) [53]. In the FZ method,

a polycrystalline rod of ultrapure EGS is melted using a RF heating coil to generate a localized molten zone within the rod. This molten zone serves as the starting point for the crystal ingot to grow. The growth is initiated using a seed crystal placed at one end of the rod. In the Cz growth process, silicon is melted in a crucible made of SiO_2 with the desired concentration of dopants, such as boron (B), phosphorus (P), or Arsenic (As). A slowly rotating seed crystal (with known crystal orientation) is pulled from the surface of the liquid, resulting in a solidified monocrystalline ingot. As the crucible slightly dissolves during the growth, oxygen will be always present in Cz silicon, i.e. around 5-20 ppma [50] [48].

As the FZ does not use a silica crucible, the oxygen content is much less compared to the Cz method (10^{16} cm^{-3}, compared to 10^{18} cm^{-3} for Cz). This also holds for the elimination of any metal contaminants from the crucible. However, as the oxygen mechanically strengthens the silicon, Cz wafers are usually mechanically stronger than FZ wafers. Until recently, silicon detectors were typically produced using the FZ technique, due to its precise control of impurity and dopant concentrations [20]. In the context of developing radiation hard silicon detectors (see also section 3.3), it was found that the impurity content of the used silicon had a strong impact on the observed radiation damage and motivated to perform studies on materials with different oxygen concentrations [54]. This lead to the exploration of the enrichment of float zone silicon using oxygen diffusion (DOFZ) [55] and more recently, detectors have been processed using Cz and epitaxially grown silicon in light of developing radiation hard silicon detectors [48]. The latter will become apparent through the work performed in this thesis, specifically in the publications presented in Chapters I, II, and III.

Epitaxy is a form of thin-film deposition on a crystalline substrate. The deposited layer registers the crystalline information from the substrate. In order to do so properly, the crystal lattices of the film and the substrate must be identical or closely matching. Epitaxial wafers offer extreme purity: carbon and oxygen are practically absent in epitaxial layers. However, the epitaxial layers are not defect-free, as the stacking faults created in epitaxial growth are the largest yield limiters. Epitaxial deposition is very reproducible, for both resistivity and thickness [50]. As will be discussed in great detail in Section 3.5, Monolithic Active Pixel Sensors (MAPS) [56], use commercial CMOS imaging technology to produce a monolithic pixel detector in which pixel sensor and electronics circuitry form one entity. In this technology, the sensing volume is an epitaxially grown layer grown on top of an often lower quality substrate [57]. As MAPS can profit from the mature CMOS imaging technology, it uses standardized foundry fabrication processes. These processes for CMOS transistors are separated in front-end and back-end processes. The front-end is characterized by oxidation, diffusion and ion implantation. The CMOS process is further characterized by the creation of the wells, isolation, gates, contacts, metallization and passivation. Back-end processes create wiring to interconnect the transistors to each other [50].

Within the context of the European Strategy for Particle Physics [58], a strong focus will be put on detector fabrication techniques, more specifically on the close relationships that can be formed with industrial partners. By meticulously controlling the wafer fabrication process, including factors such as crystal growth, doping profiles, and surface quality, the desired characteristics of the silicon material can be achieved. The silicon wafers then serve as a reliable foundation for the subsequent implementation of front-end electronics, ensuring optimal signal amplification, readout, and data processing capabilities within the pixel detector. Therefore, a seamless connection between the fabrication techniques of silicon wafers and the design and integration of front-end electronics is vital for the overall performance and functionality of silicon pixel detectors.

3.2 Front-End Electronics of Pixel Detectors

The basic detection mechanism of silicon detectors is the generation and movement of mobile charges (electrons and holes) in a silicon p-n junction [57]. When an energetic particle, such as a photon or a charged particle, interacts with the sensor material of a pixel detector, it creates electron-hole pairs. The number of ionised charges depends on the energy loss of the traversing particle within the material, which is described in section 3.2.1. The average number of electron-hole pairs (N) generated by a constant amount of absorbed energy can be obtained by dividing the energy (E) by the average energy needed to produce an electron-hole pair (w):

$$N = \frac{E}{w} \tag{3.1}$$

In silicon, w=3.6 eV, which is more than three times the bandgap (1.12 eV). This excess energy, beyond the bandgap, is used to generate phonons. The process of creating e-h pairs and generating phonons is subject to fluctuations and uncertainties. As a result, the number of e-h pairs generated in response to incident radiation can vary with F, the so-called Fano factor [59]:

$$< \Delta N^2 >= FN = F\frac{E}{w} \tag{3.2}$$

The Fano factor is in the order of 0.1 for most semiconductors and provides the ultimate limit of energy resolution in semiconductor detectors.

3.2.1 Energy Loss of Charged Particles

The detection of particles is based on their interaction with detector material, leading to the deposition of a fraction or all of their energy. Different mechanisms are responsible for the energy transfer of ionising radiation. Charged particles passing through matter experience Coulomb interactions with the medium.

These interactions can be either elastic, causing deflection without energy transfer, or inelastic, resulting in energy transfer to the medium or radiation emission, such as Bremsstrahlung, where radiation produced by the deceleration of electrons when deflected by the nuclei [60]. Moderately relativistic charged particles, excluding electrons, primarily lose energy in matter through ionization and atomic excitation processes [61]. The rate at which energy is lost, known as the stopping power, is described by the Bethe-Bloch equation [62]:

$$-\frac{dE}{dx} = Kz^2\frac{Z}{A}\frac{1}{\beta^2}\left[\frac{1}{2}ln\frac{2m_e c^2\beta^2\gamma^2 T_{max}}{I^2} - \beta^2 - \frac{\delta}{2}\right] \tag{3.3}$$

where Z, A and I are the atomic number, mass and the mean excitation energy of the medium, respectively. The charge of the incoming particle is denoted by z and β as well as γ are related to the particle velocity via β=v/c and $\gamma = 1/\sqrt{1-\beta^2}$. The electron mass is given by m_e and c denotes the speed of light. The maximum energy transfer in a single collision is represented by T_{max}. In the low $\beta\gamma$«3 range, particles exhibit high ionization and deposit a significant amount of energy in the material. Around $\beta\gamma \sim$3, there is a global minimum in energy loss, resulting in relatively low energy deposition. Particles in this regime are known as Minimum Ionizing Particles (MIPs). At $\beta\gamma$»3 values, radiative processes cause a gradual increase in energy loss. The mean energy loss per unit absorber thickness, as given by Equation 3.3, exhibits statistical fluctuations because of the stochastic nature of energy losses. This distribution of energy loss is described by the Landau distribution [63]. This probability density function, denoted as f(Δ/x), is depicted in Figure 3.1 for 500 MeV pions in silicon with varying thicknesses (640, 320, 160, and 80 µm). The normalization is set to 1 at the most probable value (MPV) Δp/x. The most probable energy loss is typically less than the mean energy loss predicted by the Bethe-Bloch equation due to the long tail in the distribution, which represents a few high-loss events. Additionally, the Most Probable Value (MPV) decreases as silicon thickness decreases. For very thin layers, the energy loss distributions deviate from the classical Landau function, necessitating the use of other models.

Energy loss in electrons and positrons differs from that in heavy particles due to factors like kinematics, spin, and their interaction with the electrons they ionize. At low energies, electrons mainly lose energy through ionization, but the mean excitation energy of the medium requires modification. At higher energies, typically above a critical energy of a few tens of MeV in most materials, bremsstrahlung becomes the dominant energy loss mechanism. The radiation length (X_0) describes the mean distance over which a high-energy electron loses all but 1/e of its energy through bremsstrahlung. When a charged particle traverses a medium, it undergoes multiple small-angle scatters, primarily due to the Coulomb interaction with the nuclei in the medium. The scattering angle, as the particle exits the material after numerous interactions, roughly approximately a Gaussian distribution with a Root-Mean-Square (RMS) [65, 66]:

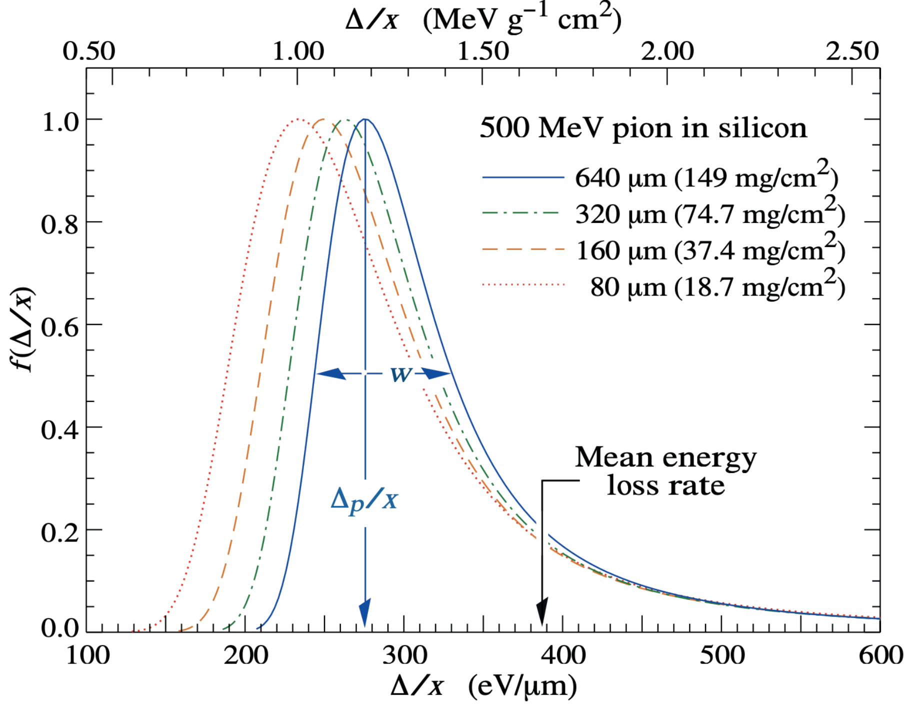


Figure 3.1: Probability density function f(δ/x) for 500 MeV pions in silicon of different thicknesses (640, 320, 160, and 80 µm), normalised to 1 at the most probable value (MPV) Δp/x. Image is obtained from Ref.[64].

$$\theta_{plane}^{RMS} = \frac{13.6MeV}{\beta pc} z \sqrt{\frac{x}{X_0}} \left[1 + 0.038 ln \frac{x}{X_0} \right] \tag{3.4}$$

where (x/X_0) represents the material content in terms of radiation lengths, p the particle momentum, z represents the charge, and βc the velocity. The radiation length is an inherent property of the material, denoting the average distance over which the energy of a high-energy electron diminishes to a fraction of 1/e of its initial energy. This multiple scattering will have an impact on the position resolution of a silicon detector composed of multiple layers. The interaction of photons with matter, including the processes such as the photoelectric effect, Compton scattering and pair production, is not discussed in this work. More information on this matter can be found in, for instance, Ref.[20].

3.2.2 Signal Formation

In nearly all silicon particle detectors, the fundamental component of the sensor is a reverse-biased p-n junction. At the boundary between the n-type and p-type materials, majority carriers from one side diffuse to the other and combine with the majority carriers, leading to a region depleted of free carriers. The presence of space charge in this depletion region results in the development of an electric field across the junction [51]. In order for the generated electron-hole pairs to be collected and to contribute to a signal, an external electric field is applied. The instantaneous current induced (I_{ind}) can be expressed following the Shockley-Ramo theorem [67] [68]

$$I_{ind} = q\vec{E_w}\vec{v} \tag{3.5}$$

where the infinitesimal motion of a charge (q) with velocity ($\vec{v}$) can be calculated from the weighting field ($\vec{E_w}$). The weighting field is obtained by applying a unit potential to the electrode under consideration and zero potential to all other electrodes. The carrier velocity is subject to constant scattering processes with lattice vibrations (phonons), impurities, surfaces or other imperfections, which is expressed in the carriers mobility. The mobility, related to the carrier velocity, depends on various macroscopic parameters such as the temperature, doping concentration or the electric field. Phenomenological mobility models, such as Masetti [69] or Canali [70], are used to express a carrier's mobility, however they are typically valid for a restricted parameter space [51].

The induced charge, Q_{ind} or the integral of the induced current, is given by

$$Q_{ind} = \int_{t_0}^{t_1} I_{ind}dt = q[\phi_w(\vec{x_1}) - \phi_w(\vec{x_0})] \tag{3.6}$$

which relates to the differences in the weighting potential (ϕ_w) between two positions ($\vec{x_0}$ and $\vec{x_1}$) and their respective time (t_0 and t_1). The electric field accelerates the charge carriers towards the readout electrodes through a process called drift. As the electron-hole pairs drift towards the pixel electrodes, they also undergo diffusion. Diffusion is driven by a concentration gradient and thermal energy causes the carriers to move in random directions. The diffusion constant, or diffusivity, is linked to the mobility by the Einstein relation [71], with the collection time being inversely proportional to the carrier mobility. Diffusion causes the spreading of charge carriers in space, leading to a widening of the charge cloud. Note that the weighting field and electrical field are distinctly different, as the electrical field determines the trajectory and velocity of a charge, whereas the weighting field only depends on the geometry and determines how charge motion couples to a specific electrode [51].

3.2.3 Front-end Electronics

The front-end electronics of silicon pixel detectors are responsible for the signal processing of the induced current. A schematic block diagram of a typical front-end processing chain in a pixel detector is shown in Figure 3.2.

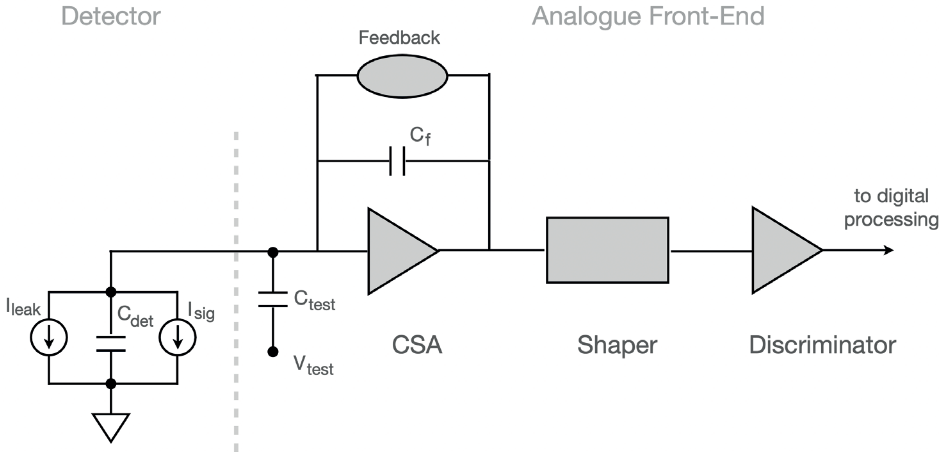


Figure 3.2: Schematic block diagram of a typical front-end processing chain in pixel detectors for HEP applications. Image is obtained from Ref.[72].

The sensor converts the energy deposited by a particle or photon to an electrical signal. The common challenge in semiconductor detector readout is achieving low-noise measurements of the signal charge while operating under stringent constraints such as high-speed operation, low power consumption, limited space, and often encountering high radiation levels. Low noise is crucial for accurately measuring small signals with a low detection threshold and a high Signal to Noise Ratio (SNR) is advantageous for improved timing measurements. The precise characteristics of the signal and noise are contingent upon the specific design of the sensor and the accompanying readout circuitry, commonly employed circuitry for silicon detectors often incorporates an amplifier and a shaper [48], discussed in more detail below.

3.2.3.1 CSA

The charge-sensitive amplifier (CSA) comprises an inverting amplifying circuit designed to generate an output voltage that is ideally proportional to the input. It features a feedback capacitor, C_f, which acts as a controllable charge-to-voltage factor. It is a controllable parameter of the electronics and not an intrinsic property of the sensor. To bring the circuit into its operational state, a high-resistance feedback path is necessary in the loop. This feedback path serves

to discharge the signal charge stored in C_f and restore the output voltage to its baseline level [20].

3.2.3.2 Shaper

The pulse shaper plays a vital role in semiconductor detector systems by enhancing the SNR. Sensors typically employ a shaper after the amplification stage to suppress high-and low-frequency noise. The main objective of the shaper is to tailor the frequency response to favor the signal while attenuating the noise, as the frequency spectra of the signal and noise differ. This filtering process improves the overall quality of the signal and allows for better discrimination between the desired signal and unwanted noise components [48]. The shaper also shortens the pulse duration to prevent signal pile-up. A simple pulse shaper consists first of high-pass filter, or differentitator, that sets the duration of the pulse by introducing a decay time constant. The second filter, a low-pass filter, or integrator, increases the rise time to limit the noise bandwidth [51].

3.2.3.3 Discriminator

A discriminator evaluates the output of the shaper by comparing it to a configurable threshold. Signals below the threshold are disregarded to eliminate low-amplitude noise contributions, while signals surpassing the threshold are forwarded to the digital logic [48].

3.2.3.4 Electronics Noise, Power Consumption, and the Q/C ratio

Noise is an inherent characteristic of all measurement procedures discussed in this thesis. While noise can arise from various stages in the measurement process, the sensors employed in this study are primarily influenced by electrical noise. Specific examples of the main noise sources are thermal noise, 1/f noise, and shot noise.

Thermal noise arises from the random movement (Brownian motion) of charge carriers and is inherently linked to temperature [73]. It is shown in Ref.[74] that when the thermal noise of the input transistor is the dominant factor in a system, which is often the case, the analog power consumption (P_{analogue}) required to achieve a certain SNR can be expressed as

$$P_{analogue} \propto \left(\frac{Q}{C}\right)^{-m} \text{with } 2 \leq m \leq 4 \tag{3.7}$$

with an inversely proportional relation to the charge Q and directly proportional relation to the capacitance C. Therefore the analogue power consumption can be significantly reduced by minimising the capacitance [48].

At the output of the shaper, thermal and 1/f noise contributions are proportional to the capacitance. 1/f noise encompasses various noise sources characterized by a frequency-dependent noise spectrum that decreases as the frequency increases. Recombination and generation of charge carriers, along with the trapping of carriers due to impurities, give rise to a power spectrum that is frequency-dependent. Random Telegraph Signal (RTS) noise can be considered as 1/f noise with a low frequency in MOS transistors [73, 75]. This type of noise arises from the trapping and detrapping of single or multiple electrons by defects at the Si/SiO_2 interface or within the oxide layer. This process results in the occurrence of single or multi-level switching events in the drain current. As it is associated with real-time electron switching events in the drain current or voltage, it is able to probe active traps individually. This effect is particularly pronounced in small devices, such as heavily down-scaled MOS transistors with reduced channel length and width [76].

Contrary, white noise (thermal noise) is characterized by a consistent power spectral density across all frequencies [73]. Shot noise is a consequence of the discrete nature of charge carriers and typically emerges at potential barriers. In silicon sensors, this noise arises due to the statistical generation of individual electron-hole pairs, leading to current fluctuations. As a result, shot noise can become apparent in leakage current.

The robustness and resilience of front-end electronics in silicon pixel detectors play a critical role in mitigating the detrimental effects of radiation damage. As these detectors are exposed to high-energy particles, they are prone to various radiation-induced phenomena that can degrade their performance over time. To counteract this, the design and implementation of front-end electronics incorporate radiation-hardened techniques, such as utilizing radiation-tolerant materials and optimizing circuit layouts. These measures aim to minimize the impact of radiation damage on the functionality of the detectors and maintain their desired performance characteristics under irradiation.

3.3 Radiation Damage in Silicon Detectors

Detectors used in high-energy physics experiments, such as those installed in the HL-LHC, operate in environments with high particle flux. As discussed in Chapter 2, the HL-LHC project anticipates an increase in both instantaneous and total integrated luminosity, leading to more demanding performance requirements [5]. Semiconductor devices in these detectors are typically affected by two primary radiation damage mechanisms: displacement damage and ionization damage. In the context of this work, both mechanisms are relevant and must be studied. Ionization damage primarily impacts the front-end and readout electronics, causing damage to surface oxide layers and the $Si\text{-}SiO_2$ interface. This damage is quantified using a parameter known as total ionizing dose (TID), typically measured in units of rad. However, since the characteristics of the depleted regions in radiation-hard diodes primarily depend on bulk properties,

displacement damage is the more critical damage mechanism addressed in this study. Furthermore, commercial CMOS technologies that are resilient against ionization damage are readily available and extensively studied. In contrast, bulk damage in these technologies is not yet well understood and requires further investigation.

3.3.1 Displacement Damage

Displacement damage refers to the process where an incident particle or photon can dislodge a silicon atom from its lattice site and hereby create defect clusters. To compare the damage caused by different types of particles with varying energies, we use a quantity called non-ionizing energy loss (NIEL). NIEL quantifies the damage induced by a specific particle fluence (time-integrated flux of radiation or particle stream per unit area) relative to the damage caused by a fluence of 1 MeV neutrons. Therefore, displacement damage is often expressed in terms of neutron equivalent fluence, typically denoted as n_{eq}/cm^2. This allows for a standardized measure of damage across various particle types and energies [77][20]. Consequently, the damage produced by different particles (neutrons, protons, pions) or particles with varying energy should be scalable through their NIEL [54]. Displacement damage can be expressed in three different ways: formation of energy mid-gap states, charge trapping through states close to the band edges, and change in doping characteristics [51].

Figure 3.3 schematically illustrates these three effects. The first (a) process illustrated in Figure 3.3 shows how shallow levels can act as radiation induced doping [51]. The second (b) way displacement damage manifests is through trapping of carriers at a deep level, where a carrier can recombine and be lost for detection, or at a shallow level, where a carrier is temporarily captured at a defect centre and is later emitted to its band. Finally, the formation of states close to the middle of the enegery gap (mid-gap states) (c), can assist electrons to travel from the valence band to the conduction band, which can lead to the increase of leakage current. As mid-gap states serve as generation-recombination centers, which can influence the movement of charge carriers within the material. The phenomena of Figure 3.3 are described in more detail below. For an extensive review on radiation-induced effects on silicon, please refer to [51] or [20].

3.3.2 Leakage Current

When energy levels exist near the middle of the bandgap, they can result in an increase in the generation current. This occurs because the thermally generated electron-hole pairs are separated by the electric field before they have a chance to recombine. The total leakage current (I) of a device, that is produced by the defect levels through the emission of electrons and holes, can be expressed as

$$I = q_0 w A \sum_{defects} G_t \tag{3.8}$$

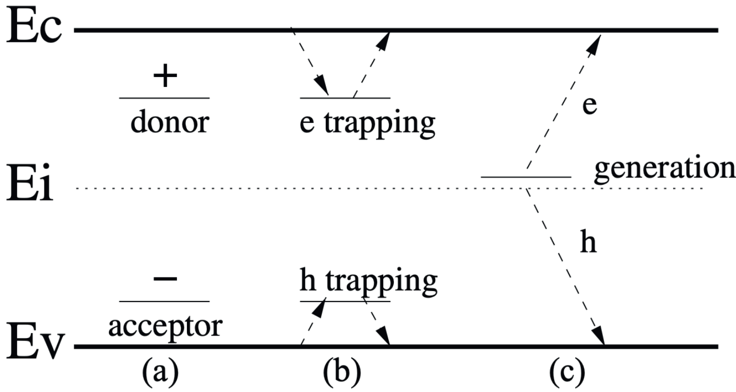


Figure 3.3: Consequences of the creation of deep energy levels to the operation of semiconductor detectors. Process (a) shows the existence of of shallow charged defects, process (b) illustrates that defects can trap and detrap free carriers and (c) shows that defects act as generation-recombination centers. Electrons and holes are denoted by e and h. Image is obtained from Ref.[20].

with q_0 the elementary charge, w the depletion width, area A, and the generation rate of a single defect type (G_t). Experimentally determination of the leakage current is challenging, as the magnitude can be influenced in various ways, such as fluence and temperature [54][48].

3.3.3 Doping Concentration

Some of the defects that are created through energetic radiation can be charged, which in turn changes the effective doping concentration and the electric field. These changes can shift the depletion voltage, which might require higher (or lower) operation voltages in order to deplete the same volume [20]. The dependence between fluence and the effective doping concentration and the depletion voltage is shown in Figure 3.4.
Beginning with n-doped material, the doping concentration gradually decreases until it reaches a specific fluence range of $\phi(2\text{–}5)\times10^{12}$ cm^{-2}, where the space charge nearly disappears. Subsequent irradiation causes an increase in the absolute effective doping concentration (N_{eff}), primarily driven by the presence of acceptor-like defects with a negative space charge. This shift in behavior resembles that of p-material, resulting in a type inversion or, more precisely, a reversal of the space charge sign. Consequently, the pn-junction relocates from the p$^+$-side of the sensor to the n$^+$-side, leading to the expansion of the space charge region from that point onwards [48]. The Hamburg model [79], described as

$$N_{\text{eff}} = N_{\text{eff},\phi=0} - [N_C(\phi) + N_a(\phi, T_a, t) + N_Y(\phi, T_a, t)] \tag{3.9}$$

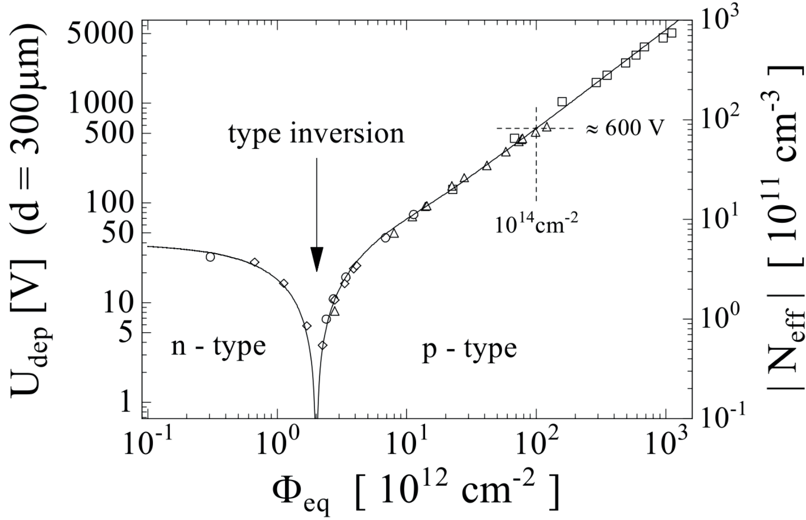


Figure 3.4: Change of the full depletion voltage (U_{dep}) of a 300 μm thick silicon sensor and its absolute effective doping (N_{eff}) versus the normalized fluence (ϕ_{eff}), immediately after irradiation. Image is obtained from Ref.[78].

parametrises the change of parameters, such as effective doping (N_{eff}), as a function of fluence and temperature history. The model is remarkably accurate up to fluneces of approximately 10^{14} cm^{-2} [20]. The term $N_C(\phi)$ is the fluence dependence of the effective doping and is only dependent on the fluence (ϕ) [48]. Overall, four components contribute to the change in space charge: donor removal and build-up of stable charge (N_C), beneficial annealing (N_a), and anti-annealing (N_Y). The term N_C is only dependent on fluence and not on temperature and is therefore referred to as stable damage [51].

3.3.4 Charge Trapping

When defect levels capture drifting charge carriers that are generated by ionizing particles in the space charge region, it is referred to as trapping and it is one of the limiting factor for high-fluence applications. Trapping centers refer to the location where the charge gets trapped. The process of trapping is characterized by the effective trapping time, $\tau_{\text{eff,e}}$ for electrons or $\tau_{\text{eff,h}}$ for holes, as

$$\frac{1}{\tau_{\text{eff,e}}} = \sum_{\text{defects}} c_{(n,t)}(1 - f_t)N_t \tag{3.10}$$

$$\frac{1}{\tau_{\text{eff,h}}} = \sum_{\text{defects}} c_{(p,t)} f_t N_t \tag{3.11}$$

with c_n and c_p the capture coefficients for electrons and holes, respectively, f_t the defect occupancy with electrons, and N_t the concentration of defect types [54]. Subsequently, when trapping centers are present, a change in carrier concentration is observed. In circumstances where the defect concentration increases, due to increasing particle fluence, a larger amount of charge carriers are trapped. As the charge is trapped into a defect and not released within the signal collection time of the sensor, the signal is reduced, leading to a decrease of the charge collection efficiency. Beyond a certain fluence level, typically encountered at the LHC, the dominant factor affecting efficient operation is the process of charge trapping. The impact of trapping on charge collection becomes evident in fully depleted detectors, where the deterioration of induced charge solely results from trapping. Ref. [20] demonstrated that at a fluence of approximately 10^{15} cm^{-2}, only around half of the charge observed in a non-irradiated detector is measured at the voltage threshold. However, the induced charge increases again for bias voltages surpassing the full depletion voltage. This increase can be attributed to a higher electric field, which reduces the drift time and consequently mitigates the influence of charge trapping.

While it is challenging to mitigate radiation damage to individual devices within a specific technology, numerous techniques can be employed to minimize the impact of radiation damage on an entire system. The objective of radiation-hard design is not merely to achieve a system that remains unchanged under irradiation, but rather to ensure that the system maintains its required performance characteristics throughout its lifespan. In the following section, we will examine two distinct types of silicon pixel detectors: the hybrid and monolithic detector technology. We will emphasize the key advantages and disadvantages of these designs and also discuss an example of each technology that demonstrates radiation-hard properties.

3.4 Hybrid Pixel Detectors

Hybrid pixel detectors are a type of radiation detector used in various scientific applications, such as high-energy physics and X-ray imaging [20]. They consist of a sensor layer, usually made of silicon, and a readout chip, which are connected using bump bonding technology, illustrated in Figure 3.5. The sensor layer, denoted in the image as the passive pixel sensor (light blue), is composed of an array of small pixels, typically with a size of a few tens of micrometers, which are designed to detect charged particles or photons. The readout chip, located on top of the sensor in Figure 3.5 (dark blue), is designed to amplify and digitize the signals produced by the sensor pixels, and to perform various other functions, such as triggering, data acquisition, and communication with external systems [20]. A so-called bump-bond (grey) connects the two layers. This connection is established through two metal contacts situated on opposite sides of the layer. Hybrid pixel technology has established itself as the preferred pixel technology for the detectors in high-energy physics and high-rate fixed

target experiments, primarily due to their proven performance in demanding high-radiation environments, among other advantages [48, 57].

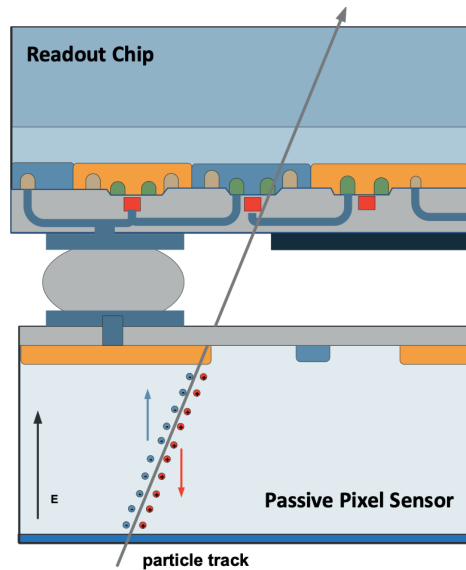


Figure 3.5: Cross section of an exemplary individual hybrid pixel cell. The bottom layer represents the sensing layer, or the passive pixel sensor (light blue), and the top layer represents the read-out chip (dark blue). The two layers are connected through a ball-like feature (grey), or bump-bond. The specifications of the CMOS circuitry of the readout chip are not shown, but they are represented by the alternating orange and blue blocks. Image is obtained from Ref.[80].

Hybrid pixel detectors offer several advantages over other types of detectors, such as small pixel dimensions, high readout speed, and low noise. The small size of the pixels allows for precise localization (σ_s=9.4 for CMS with pixel pitch 100 µm [81]) of the interaction point of a particle or photon, which is important for imaging and tracking applications. The fast readout speed, typically in the range of a few microseconds or less, allows for high event rates and reduces the likelihood of pileup, which is the overlap of signals from multiple interactions. The low noise, achieved through careful design and optimization of the readout electronics, results in high signal-to-noise ratios and improves the detector's sensitivity.

However, the hybrid technology also has notable disadvantages. Hybrid pixels constitute a relatively large material budget, which is distributed among the various module components. As will become evident in the next subsection (3.4.1), the production of hybrid modules is highly complex, due to processes such as the bump-bonding and flip-chipping. These processes consist of a large number of production steps and require specialized equipment and skilled personnel, implying usually high cost. Despite these limitations, hybrid pixel detectors remain a popular choice in high-energy physics experiments and other fields that require high-performance detectors. Ongoing research and development efforts are aimed at overcoming the limitations and further improving the performance of hybrid pixel detectors [57]. Further discussion of this topic will be presented in Chapter 5 as a component of the research conducted within this thesis.

3.4.1 ATLAS ITk pixel module

The ATLAS ITk Pixel Detector, introduced in Chapter 2, requires very particular developments of the sensor technology, including stringent demands on, among others, hit efficiency, power dissipation, and radiation hardness (also detailed in Chapter 2 in Table 2.1). The ITk pixel module is the basic mechanical and electrical unit of the ITk Pixel Detector using the hybrid pixel detector technology. It serves as the main electrical interface to the read-out, external services, and thermal interface to the local supports. The ITk pixel module consists of a bare module (silicon sensor bump-bonded to front-end chips), the module flex (flexible PCB that provides electrical connection to the module and is glued to the backside of the sensors), DCS or Pixel Serial Powering chip (ensures reliability of the entire serial powering chain, monitoring and control), and the aggregator chip (multiplexing of multi-chip modules into a number of high speed electrical data cables) [5]. An exemplary ITk pixel dummy module is shown in Figure 3.6. Given that a part of the work presented in this thesis relates to the bare module, additional details about the fabrication process can be found below.

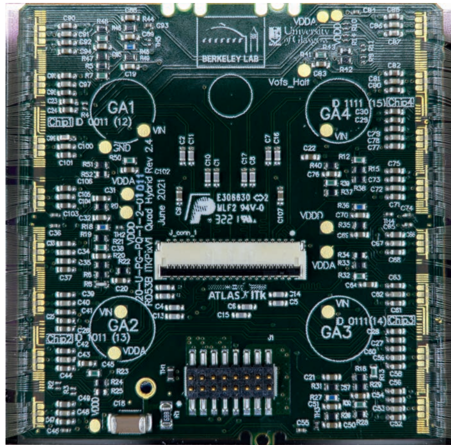


Figure 3.6: Picture of a dummy ATLAS ITk pixel module consisting of the module flex (dark green base layer), wire-bonds (located on left and right side of the module), passive components, data connector (white feature in the center of the module) and power connector (bottom center).

In order to fabricate the bare ITk pixel module a process called hybridization is required. During this process, the sensor and front-end chips are connected together (hybridized), through a time consuming post-processing procedure. It is the most expensive part of the pixel module construction and consists of four main stages: bump (solder or indium) deposition on the front-end chip wafer, sensor wafer processing, wafer thinning & dicing, and finally face-to-face joining of the parts, i.e. flip-chip bonding. The process has shown to be technological challenging, due to the very thin pixel assemblies (bare module ∼150 µm) and high bump bond density. Currently the industry can handle bump pitches in the order of 25-50 µm. Moreover, the intricacy of the process presents challenges regarding production rate and volume, given the notable variability in consistency observed among hybridized modules [5].

The front-end chip is responsible for integrating the charge in the sensor generated by the incident particles, the amplification and digitization of this signal, and finally sending the hit information to the Data Acquisition (DAQ) system. It is produced in 65 nm CMOS technology on 300 mm diameter wafers. The front-end chip is 20 mm wide by 21 mm tall die with 153600 pixels, with a pixel input bump pitch of $50{\times}50$ μm. Chapter 5 will discuss in more detail the requirements and processes that are involved in the high-voltage protection of ITk pixel modules through the use of protective coating layers.

3.5 Monolithic Pixel Detectors

In MAPS, sensor and readout electronics are integrated on the same silicon chip. In this technology, part of the substrate is used as detector material. It exploits the use of the mature commercial CMOS imaging technology and can therefore be very cost effective. As MAPS integrate the pixel sensor array and readout electronics on a single chip, the final detector can become more compact and is associated with smaller material budget. Due to the fact that for MAPS the integrated circuit technology is produced on a single wafer, the fabrication is scalable, significantly easier, and in turn cheaper, compared to hybrid pixel detectors. As these sensors have potential for high radiation tolerance, they are expected to play an important role in future particle physics experiments [58][20].

In an exemplary NMOS transistor MAPS (top image of Figure 3.7), the n-well serves as the collecting electrode, and all transistors are situated within the p-wells. In a region adjacent to the n-well, charge carriers will be depleted, while signal electrons are collected through drift. However, the primary sensitive volume, the p-type epitaxial layer, remains field-free, resulting in charge collection mainly through diffusion. This diffusion process is inherently slow and leads to significant charge spreading into neighboring cells [20]. To address the challenge of slow charge collection by diffusion, which makes the sensor susceptible to bulk radiation damage, DMAPS are under development, bottom image of Figure 3.7. DMAPS are fabricated on substrates with resistivity ranging from 100 Ωcm to a few kΩcm and operate with depletion depths typically ranging from 50 to 200 μm [57].

Two approaches have been pursued in the field of DMAPS: the large collection electrode and the small collection electrode design. In the large collection electrode design the readout circuitry is placed into the collection electrode. Inside the electrode, the electronics are well-shielded from the rest of the sensor, which enables the application of a high bias voltage that leads to a sizeable depletion region and offers a more uniform electric field, resulting in shorter drift distances and enhanced radiation tolerance. The size of the collection electrode is driven by the necessity to house the required on-pixel electronics. However, the larger size has the disadvantage of larger capacitance, around 100 fF per pixel, along with additional well-to-well capacitance, leading to increased noise, reduced speed, higher power consumption, and potential cross-talk between the sensor and digital electronics [20]. In the small collection-electrode design, as holds for the MALTA chip (to be discussed in Chapter 4), the readout electronics are placed on deep well structures that are separated from the collection electrode. As a result, the size of the electrode can be minimised, offering lower capacitance, improved noise and speed at lower power. However, the electric field in the sensor is not uniform, leading to increased sensitivity to radiation damage.

In the field of DMAPS, both the large and small collection electrode design have

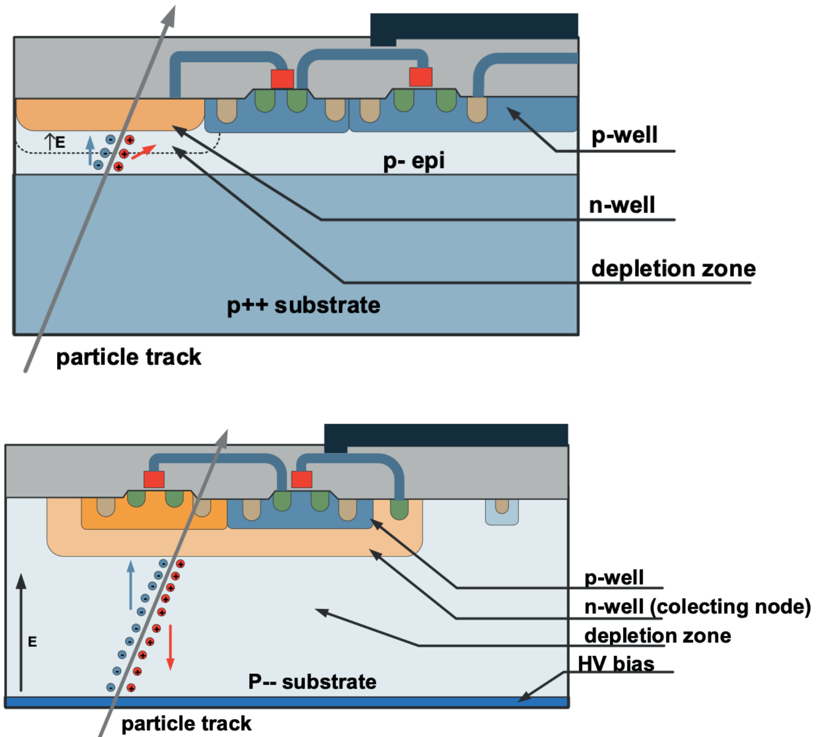


Figure 3.7: Top: Cross section of an exemplary MAPS with charge collection in an epitaxial layer (light-blue) mainly by diffusion. The n-well (orange) acts as the charge collection node. The other n-wells are shielded by the p-well (dark-blue). Bottom: Cross section of an exemplary DMAPS. Multiple wells (orange and blue) on high resistive substrate allow complete embedding of the CMOS electronics in the charge collecting deep n-well (light orange). Images are obtained from Ref.[80].

been fabricated by various foundries using technologies such as 150 nm, 180 nm, and 350 nm. These sensors have demonstrated resilience even after irradiation with hadrons at fluencies surpassing a few 10^{15} cm^{-2}. Concluding the choice between DMAPS with small collection electrodes and those with large collection electrodes depends on the specific requirements of the application. If high spatial resolution is critical, DMAPS with small collection electrodes may be preferred. If high signal-to-noise ratio and charge collection efficiency are more important, DMAPS with large collection electrodes may be the better choice. [20].

3.5.1 MALTA: Monolithic ALICE to ATLAS

Prototypes of DMAPS, such as [80] and [82], have been successfully developed using the Tower Semiconductor 180 nm CMOS imaging process. These prototypes were developed in order to assess their suitability for the Phase-II upgrade of ATLAS for the HL-LHC, as well as for future HEP experiments. MALTA was born with the emphasis on fulfilling requirements on radiation hardness, specifically up to a Total Ionizing Dose (TID) of 100 Mrad and a Non-Ionizing Energy Loss (NIEL) of $> 1\times10^{15}$ 1 MeV n_{eq}/cm^2 [83]. MALTA is fabricated on both high-resistivity epitaxial layers produced on low resistivity substrate and on thick high-resistivity Cz substrates. In order to ensure that ionization charge is not trapped in the non-depleted part of the sensing volume, various process modifications on the sensor level were introduced by Ref.[84]. The original MALTA chip features a small collection electrode design ($\sim$2 μm diameter) which allows for small pixel size, low capacitance, and low power consumption. As the electrode is separated from circuitry, as will be shown in more detail in Chapter 4, the small pixel size can offer high spatial resolution. The matrix consists of 512×512 pixels (2×2 cm) with a pixel pitch of 36.4 µm and features an asynchronous readout architecture (further discussed in Chapter 4.2.1. Chapter 4 will discuss in extensive detail the timeline, development, and performance of MALTA throughout the years. In particular, it will demonstrate how a combination of process modifications and front-end design can effectively counteract the significant impact of radiation-induced damage, while still maintaining excellent performance. It is the main subject of the work performed in this thesis and of the publications presented in Chapters I, II, and III.

3.6 Conclusion

This chapter has provided an overview of silicon pixel detectors, covering various aspects crucial to their functionality and performance. Understanding the intricacies of wafer production is essential for optimizing the quality and characteristics of the sensors. The front-end electronics of sensors play a vital role in signal amplification, readout, and data processing. The significance of efficient front-end electronics design has been emphasized, as it directly impacts the detector's sensitivity, noise levels, and overall performance. Moreover, the chapter addressed the issue of radiation damage in silicon pixel detectors. Radiation-induced effects pose significant challenges in maintaining the detectors' long-term stability and functionality. The discussion shed light on the various types of radiation damage, their underlying mechanisms, and the strategies employed to mitigate their adverse impacts. Finally, the chapter explored two prominent types of pixel detectors: hybrid and monolithic. The examples of the ITk Pixel and MALTA served as illustrative cases, demonstrating the practical application of hybrid and monolithic pixel detectors, respectively.

Chapter 4

State of the Art: MALTA Monolithic CMOS Sensor

The utilization of MAPS in commercial CMOS technologies marked a significant milestone in various experiments such as the STAR experiment [85] and the Inner Tracking System (ITS) [82], an upgrade of the ALICE experiment, with the development of the ALPIDE monolithic active pixel sensor [86]. The ALPIDE sensor, based on Tower Semiconductor 180 nm CMOS imaging technology [8], served as the foundation for the development of the Monolithic ALice To Atlas (MALTA) sensor [6]. The primary goal of MALTA was to explore its applicability in the Phase-II upgrade of ATLAS for the High Luminosity LHC and other high-energy physics experiments. Consequently, the detector was required to satisfy certain criteria, which are detailed in Table 4.1 [87].

Requirements of MALTA	
Radiation hardness (NIEL)	$>10^{15}$ 1 MeV n_{eq}/cm^2
Radiation hardness (TID)	100 Mrad
Hit rate capability	>100 MHz/cm^2
Response time	25 ns (LHC bunch-crossing)

Table 4.1: Overview of the main requirements of MALTA in terms of radiation harndess (NIEL and TID), hit rate capability, and response time.

To meet these requirements, MALTA underwent several design and process modifications over the years, evolving from MALTA to the mini-MALTA demonstrator [88], MALTA on Cz substrates [11], and finally the latest generation, MALTA2 [16]. For a chronological overview of the MALTA development stages, please consult Table 4.2. The specifications of each version, such as front-end type and pixel flavour, will be elaborated on in the following sections. A large contribution of this thesis originated from the characterization and testing of MALTA on Cz substrates. The outcomes of these efforts are detailed in the publications presented in Chapters I and II. The trajectory of research extended to MALTA2, detailed in the publication presented in Chapter III. Given the interconnected nature of the work across these versions, which had commenced prior to this thesis, this chapter aims to present an overview of the evolution of MALTA over the years. It will highlight the baseline process technology, the motivations behind the most significant changes, and performance highlights of each design will be presented. Given that a majority of the results presnted in the following paragraphs stem from a period preceding the work undertaken in this thesis, the primary sources referred to are [11, 64, 88–90], if not stated otherwise.

Overview of the MALTA-family					
Version	**Year**	**Size**	**Front-end**	**Substrate**	**Pixel flavour**
MALTA	2018	2×2 cm	Standard	Epitaxial	STD
mini-MALTA	2019	1.7×0.5 mm	Standard or Enlarged transistor and cascoded	Epitaxial	STD, NGAP, XDPW
MALTA Cz	2019	2×2 cm	Standard	Epitaxial and Cz	STD, NGAP, XDPW
MALTA2	2021	2×1 cm	Enlarged transistor and cascoded	Epitaxial and Cz	STD, NGAP, XDPW

Table 4.2: Overview of the MALTA-family. For each version its respective year of availability for testing is indicated. Furthermore, the summary includes details on the size, front-end type, substrate material, and the range of available pixel flavours.

4.1 MALTA Process Technology

The left image of Figure 4.1 illustrates a cross-sectional view the Tower Semiconductor 180 nm CMOS imaging technology process, referred to as the "standard process technology". In the center of the pixel lies the small, octagonal shaped n-well collection electrode (diameter of approximately 2 µm), which collects the charge generated within the sensor. The small collection electrode allows to achieve a high SNR and to reduce analog power consumption, resulting in a low (femto-Farad) sensor capacitance. Due to the low sensor capacitance, the sensor junction is, consequently, also small in size. The wells that host the in-pixel electronics are positioned to the left and right of the collection electrode in the image, with a p-well for the NMOS transistors and an n-well for PMOS transistors. To minimize lateral capacitance between the electrode and the wells, a separation of several microns exists. A deep p-well envelops the n-wells of the PMOS transistors, ensuring that there is no competition with the collection electrode in charge collection. This design choice facilitates the incorporation of full CMOS and consequently allows for more complex readout circuitry within each pixel. As the foundry offers the flexibility of utilizing various starting materials for the sensor, a high-resistivity (>1 kΩcm) p-type epitaxial layer is used, in light of the requirements imposed on radiation hardness (Table 4.1), to enhance the depletion region surrounding the collection electrode. The design shown in Figure 4.1 represents a pixel fabricated on an epitaxial layer with a thickness ranging from 25 to 30 µm. To increase the depletion depth even further and to reduce the sensor capacitance, a reverse bias of up to 6 V is applied on the p-type substrate. As the NMOS transistors see the same reverse bias applied

to the substrate, the reverse bias is limited to the breakdown of the source/drain junctions of the NMOS transistors. Overall this process technology offers several advantages: it eliminates the need for a dedicated backside implantation and allows to achieve high resistivity necessary for depletion. However, the thickness of the epitaxial layer imposes limitations on the depletion thickness [89].

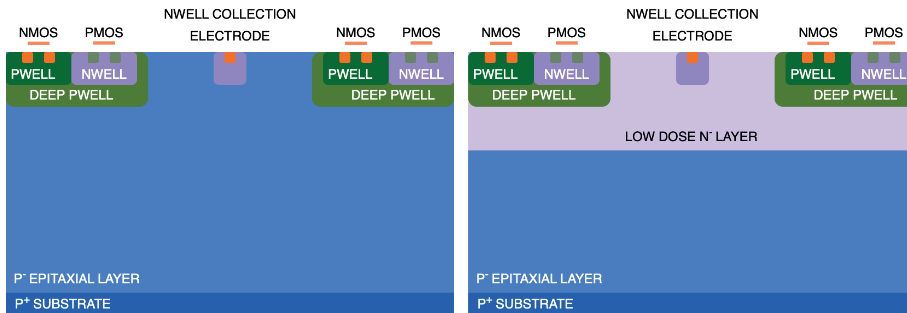


Figure 4.1: Cross section of the Tower Semiconductor 180 nm CMOS imaging technology fabricated on a high resistivity p-type epitaxial layer. Left images shows the standard process, the right image shows the standard modified process (STD) where an additional n^- layer is introduced. Indicated are the n-well collection electrode, p-well and n-wells hosting the NMOS and PMOS, respectively, and the deep p-well. Images are not drawn to scale and are adapted from Ref.[64].

For the standard process technology, left image of Figure 4.1, even with a high resistivity epitaxial layer and maximum sensor bias, the depleted region is severely limited. Additionally, TCAD simulations by Munker et al. [84] have revealed the presence of an electric field minimum at the corner of the pixel. This introduces a complex electric field configuration that limits the speed of charge collection, especially for charges originating from the pixel corner. The impact of diffusion and the electric field minimum on performance has been extensively discussed in Ref.[91]. Although the small collection electrode used in the standard process technology can offer excellent spatial precision as it allows for smaller pixel size, it suffers from slow charge collection (few times tens of ns), resulting in reduced radiation hardness, timing, and detection efficiency. To address the limitations imposed by the small sensor junction on the depletion depth, a modification was implemented. In order to enable complete lateral depletion, a deep low dose n^- layer was added, as depicted in the right image of Figure 4.1. This modification effectively isolates the CMOS circuitry from the backside, removing the constraint on the bias voltage, as discussed previously. Consequently, the sensor can be operated at larger reverse bias voltages [89]. This modified version is referred to as the "standard modified process technology", or simply STD process in this work.

4.2 MALTA Pixel

In 2018 the first full-scale (2×2 cm) MALTA featuring the standard modified process technology (MALTA STD) was developed. MALTA comprises a 512×512 pixel matrix, with a squared pixel of pitch 36.4 µm, illustrated in Figure 4.2. The front-end circuit occupies an area of approximately 160 µm^2 and is placed to the left of the small collection electrode with other analog circuitry. The remaining part is occupied by the digital circuit. The analogue and digital regions of the pixel are also well separated and shielded from each other with metals to avoid any cross-talk. For this reason, the two regions use different power domains for their supply voltage [16, 64].

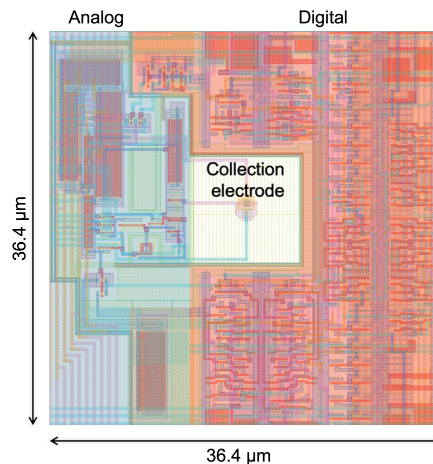


Figure 4.2: Layout of a pixel in the MALTA chip with a pixel pitch of 36.4 µm. The analogue and digital part are well separated and shielded from each other and the collection electrode to avoid crosstalk. Image is obtained from Ref.[64].

The use of a small collection electrode allows to achieve a high Q/C ratio at the circuit input by minimizing the input capacitance. For a sensitive layer thickness of 25 µm, a most probable ionization charge of approximately 1500 e^- is expected. If an ionization charge of 63 electron-hole pairs per µm path length and a total electrode capacitance of 5 fF is assumed, this corresponds to a voltage step of around 50 mV. This voltage step provides the opportunity to employ an open-loop voltage amplifier as the initial amplification stage, contrary to the conventional charge-sensitive amplifier setup with a feedback capacitor (see also section 3.2.3). This not only saves space but also simplifies the circuit. The collection electrode input voltage undergoes a reset after a particle hit. The front-end amplifier output is connected to a discriminator, generating the digital signal for a hit pixel. The discriminator threshold is globally set for the entire sensor [88]. Figure 4.3 illustrates the analog front-end circuit.

Figure 4.3: Layout of the analog front-end circuit of the MALTA sensor. Image is obtained from Ref.[64].

The full MALTA pixel matrix is subdivided into eight sectors, each featuring different sensor and front-end pre-amplifier configurations, see Figure 4.4. These sectors (S0-S7) deviate in terms of collection electrode size, ranging from 2 to 3 µm in diameter, as well as the gap between the electrode and the adjacent deep p-well, varying between 3.5 and 4 µm. Another notable distinction involves the extension of the deep p-well within the pixel. Precisely, as the deep p-well is strictly necessary only beneath the n-wells of PMOS transistors, half of the sectors adopt a medium (med) deep p-well layout. Conversely, the other sectors employ a conventional maximum (max) deep p-well layout, where all transistors, whether NMOS or PMOS, possess a deep p-well beneath them. In terms of front-end design, a distinction is made on the circuit employed for resetting the collection electrode's voltage. This is achieved using either a diode or a PMOS transistor [64]. Figure 4.4 indicates the position of the digital periphery, digital-to-analogue converters (DACs), and the low-voltage differential signal (LVDS) output. MALTA incorporates an asynchronous readout mechanism, where the hit information is transmitted directly from the pixel to the chip periphery through 37 parallel LVDS output signals that describe the hit address. Further details regarding this readout mechanism are discussed in the following section.

4.2.1 Asynchronous Readout

The MALTA chip uses an asynchronous digital readout architecture which eliminates the need to propagate a clock signal to the pixel matrix. This in turn

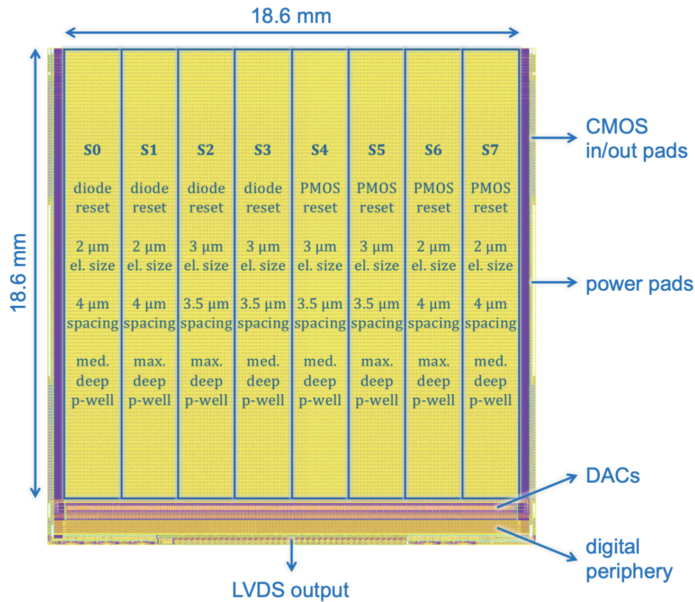


Figure 4.4: Layout and main building blocks of the full MALTA chip. The pixel matrix is divided into 8 sectors (S0-S7) with different sensor and front-end designs. The power pads are positioned at the bottom, left and right side of the matrix. The image indicates the position of the digital periphery, DACs, and LVDS output. Image is obtained from Ref.[64].

minimizes analog-digital cross-talk and allows for low per consumption, i.e. 10 mW/cm^2 at 100 MHz/cm^2 and 70 mW/cm^2 analog power. The pixels of the MALTA matrix are organised in double columns and within a double column in groups of 2×8 pixels. Alternating sets of 16 pixels (or pixel group) are connected to two output buses per double column, schematically presented in Figure 4.5. When a pixel within a group detects a hit exceeding the charge threshold of the pixel discriminator, a reference pulse is generated, which is appended to the pixel and group address, respectively 16-bit and 5-bit. The hits are distributed in two parallel 22-bit wide busses, one for even groups and the other for odd groups. If two pixels within one set react simultaneously, the two corresponding lines are activated and only one word is transmitted on the bus. If pixels within one set receive a sufficiently different amount of charge and react one after the other, two words are transmitted sequentially over the bus, guaranteeing sufficient separation of the pulses on the bus for proper transmission. Data is transmitted almost instantaneously, and is therefore available at the chip periphery only a few nanoseconds after the hit took place. The readout of each double column is completely independent of the others, so that multiple columns can transmit data at the same time. At the chip periphery, the hit information is merged and

finally provided on a 40-bit wide parallel LVDS output port, which includes the full pixel address and timing information. During testing, the MALTA chip is read out with KC705 Evaluation Board with a Kintex Field Programmable Gate Array (FPGA) [92], which asynchronously oversamples the 40 output signals and performs further processing of the hit data [83]. For more information on the asynchronous readout of MALTA, please consult Ref. [64].

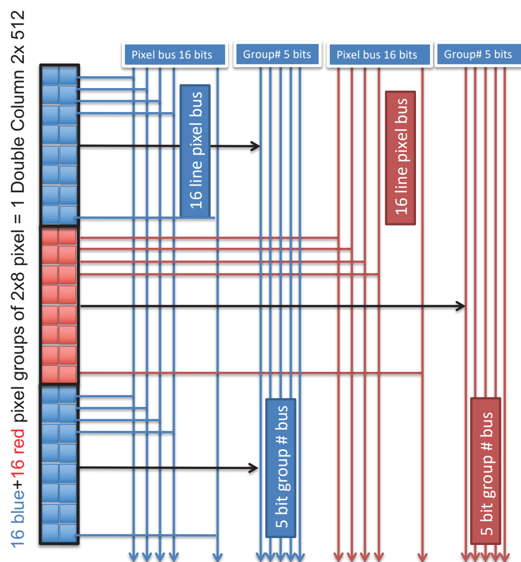


Figure 4.5: Organisation of a double column in the MALTA digital readout architecture. Each pixel hit generates a signal on its corresponding line of the pixel bus that is transmitted asynchronously at the time of the discriminator output. The pixel bus is shared by alternating pixel groups, divided in blue or red groups. All of the blue groups connect to same line. Image is obtained from Ref.[64].

Depending on the final deployment location of MALTA, specific features, such as asynchronous readout, will gain greater significance. In the case of a trigger-less detector [93], like the Upstream Tracker upgrade of LHCb [9], a data-driven solution proves highly valuable. The asynchronous data transmission from pixel to the end of the column enables the use of tracking information for a trigger decision at a later time. Furthermore, this clock-less readout scheme contributes to low power consumption, as mentioned previously. Given that the power consumption for MALTA is mainly tied to the number of hits and the readout to the periphery, there was no observed change in power dissipation for the MALTA chips after irradiation.

For a trigger-based detector, a synchronous readout scheme may be more appropriate. Monopix2 [94], a monolithic prototype, is equipped with a synchronous readout. This prototype has undergone development in two DMAPS lines, LF-Monopix and TJ-Monopix, fabricated using 150 nm LFoundry and 180 nm Tower Semiconductor CMOS imaging technology, respectively. While featuring distinct pixel layouts, front-end implementations, and biasing schemes, they share the same readout architecture. The employed readout is a fully synchronous column-drain architecture [95], resembling the scheme used in the FE-I3 and FE-I4 [96] chips in the current ATLAS pixel detector and IBL [97], respectively. It includes the Time of Arrival (ToA) and the total length, or time over threshold (ToT), of the pulse, which are determined by the leading and trailing edges of the discriminator output, utilizing a 40 MHz clock. The advantages of ToT information include improved timing and spatial resolution, features that MALTA lacks due to space constraints in the matrix. A drawback of this readout scheme is its limitation in operating at a high hit rate, as the frequency is limited to 40 MHz. Additionally, issues such as analog-digital cross-talk and high power consumption have been reported in Ref. [94]. In conclusion, the choice between an asynchronous or synchronous readout depends on the application and accompanying requirements, considering factors like hit rate, resolution, and power consumption.

4.2.2 Performance of the STD MALTA Sensor

Figure 4.6 displays the in-pixel (edge between two pixels) efficiency maps projected onto a 2×2 pixel matrix. This is shown for two MALTA STD samples on an epitaxial substrate, both before (left) and after (right) irradiation, at an operating threshold of 250 and 400 e^-, respectively. The hit detection efficiency is calculated as the number of matched clusters on the Device Under Test (DUT) over the total number of reconstructed tracks. A matched cluster is found by associating hit clusters on the DUT to a track, which should be found within 100 µm. Due to the very large statistics, a small statistical error is recorded for the hit efficiency.

From Figure 4.6, it can be observed that prior to irradiation, the efficiency is uniformly distributed across the pixels, with an efficiency response of approximately 97%. However, after irradiation, a noticeable degradation of the efficiency is observed, primarily originating from the pixel corners where the efficiency falls down to 30%. This is motivated by the fact that the lateral electric field cannot sufficiently push the deposited charge towards the small central electrode. The results presented in Figure 4.6, as discussed in Ref.[90], highlight the impact of the electric field minimum on the loss of efficiency due to the long drift path. These findings, along with other observations, have emphasized the need to mitigate the effects of the electric field minimum on performance. This mitigation is done with further optimization the sensor layout, as will be described in the next section (4.3).

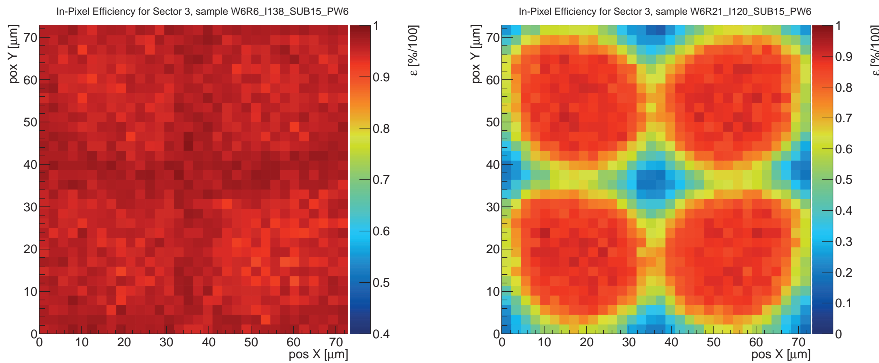


Figure 4.6: In-pixel 2D efficiency map projected over a 2×2 pixel matrix for non-irradiated (left) and irradiated at 5×10^{14} 1 MeV n_{eq}/cm^2 (right) STD MALTA samples measured with a 180 GeV proton beam. Results are shown for sector 3 of the MALTA matrix, corresponding to a maximum extension of the deep p-well. The chips are operated at -15 V bias voltage and -20°C. Operational threshold corresponds to 250 (left) and 400 (right) electrons. Images are obtained from Ref.[90].

4.3 Mini-MALTA Demonstrator

In 2019, the mini-MALTA demonstrator was introduced, comprising of a 16×64 pixel matrix with the same pixel geometry of MALTA, i.e. square with pixel pitch of 36.4 µm. Its total size measures 1.7×5 mm. It was designed as a testing vehicle to evaluate various front-end and process designs. Consequently, the detector was partitioned into eight distinct pixel sub-groups labeled as S0-S7, as depicted in Figure 4.7. The sub-groups vary in terms of analog front-end design (enlarged transistors or standard front-end), reset mechanisms (standard diode or PMOS transistor), and pixel flavors (STD, NGAP, or XDPW). These distinctions are further discussed below. In the mini-MALTA demonstrator a new slow control was implemented based on a shift-register, elaborated on in Ref.[14].

4.3.1 Sensor and Front-End Modifications

One of the most important changes that was featured in the mini-MALTA demonstrator were the different pixel flavours. As was shown in Figure 4.6, the efficiency response, especially in the pixel corners, was severely compromised after neutron irradiation. This motivated Munker et al. in Ref.[84] to perform dedicated TCAD simulations in order to determine how the geometry of the sensor could be modified such that charge collection after irradiation would be improved. The simulations that were performed indicated that in the default MALTA pixel layout, the lateral electrical field is not able to sufficiently push

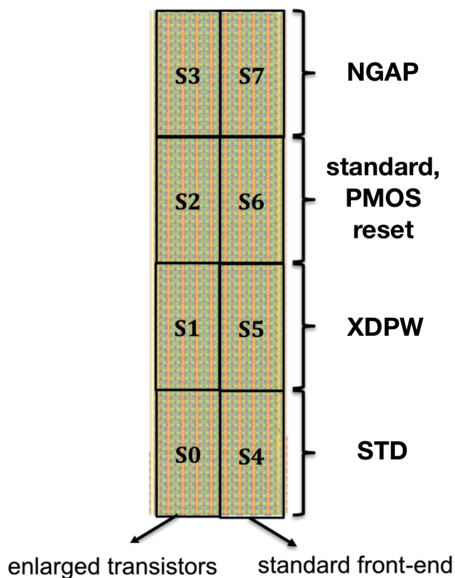


Figure 4.7: Layout of the mini-MALTA pixel matrix with 8 pixel sub-groups (S0-S7) of 8×16 pixels. The left half of the pixel matrix features enlarged transistors, the right half features the standard MALTA front-end. The bottom row of sub-groups features the STD pixel flavour, the second row features the XDPW flavour, the third row implements different reset mechanisms (standard or PMOS), and the fourth row features the NGAP flavour. Image is adapted from Ref.[88].

the deposited charge towards the small collection electrode. For this reason the process technology was further modified by adding a gap in the low dose n^- layer (left image of Figure 4.8) or adding an extra deep p-type implant (right image of Figure 4.8). We refer to these configurations as NGAP and XDPW [84], respectively. The purpose of these modifications is to improve the charge collection at the pixel edges and corners through the creation of a stronger lateral field, which focuses the ionization charge towards the collection electrode. Initial measurements of the MALTA circuit revealed a significant contribution of RTS noise which prevented the sensor to be operated at low threshold settings. This was attributed to the size of the M3 transistor in Figure 4.3. Consequently, in specific sectors of the mini-MALTA matrix, denoted as the "enlarged transistor" in Figure 4.3, the size of the M3 transistor was increased. The study conducted by Dyndal et al., as referenced in [88], demonstrated that enlarging the transistor size substantially mitigated RTS noise, both pre and post-irradiation. Moreover, this modification resulted in a notably increased gain and a reduced charge threshold under the same settings.

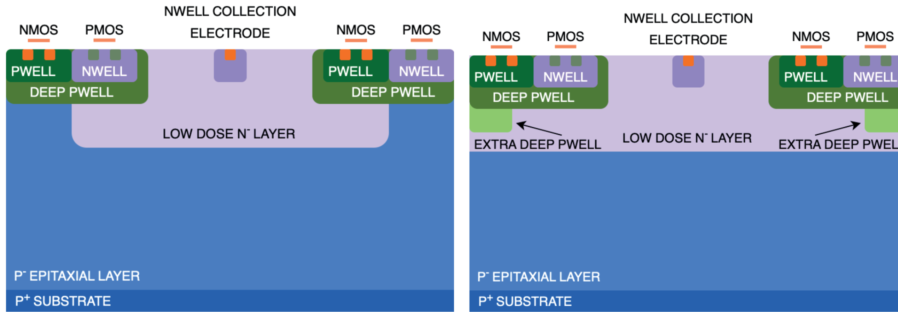


Figure 4.8: Cross section of the process modifications for the MALTA sensor: NGAP flavour with low dose n^- layer removed at the edge of the pixel (left), and XDPW flavour with extra deep p-well at the edge of the pixel (right). Images are not drawn to scale and are adapted from Ref.[88].

4.3.2 Performance Before and After Displacement Damage

In Figure 4.9 a 2D efficiency map for a non-irradiated (left) and an irradiated at 1×15 1 MeV n_{eq}/cm^2 (right) mini-MALTA samples are shown, measured with a 2.5 GeV electron beam. The arrangement of the full matrix shown in Figure 4.9 corresponds to the layout depicted in Figure 4.7. The different sub-groups of the matrix are visible: STD (bottom row of each chip), XDPW (second row) and NGAP (top row). No results are shown for the sub-groups featuring the different reset mechanisms (standard or PMOS). The empty bins represent dead pixels. Results are also shown for sensor regions with standard (right side of each chip) and enlarged (left side) transistors. From the left plot, it can be observed that prior to irradiation uniform efficiency response above 98% can be achieved for all sub-groups.

After irradiation (right), the effect of the process modifications on performance are visible. There is an approximately ∼13% improvement due to the enlarged transistor size and a ∼6% improvement from process modification, with a similar improvement from XDPW and NGAP. This was attributed to the fact that the enlarged transistors allow for higher gain, smaller gain spread and reduced RTS noise [88]. Figure 4.10 shows the efficiency response versus operating threshold for neutron irradiated mini-MALTA samples at 1×10^{15} 1 MeV n_{eq}/cm^2 (top) and 2×15 1 MeV n_{eq}/cm^2 (bottom) measured with a 2.5 GeV electron beam operated at -6 V bias voltage. In both figures, different regions of the chip with different sensor flavours are presented, i.e. STD (circles), XDPW (triangles), and NGAP (rectangles). The open markers represent regions with the standard transistor size whereas the full markers represent the enlarged transistors. The orange or light blue symbols represent the 25 µm thick sample and the red or dark blue symbols represent the 30 µm thick sample.

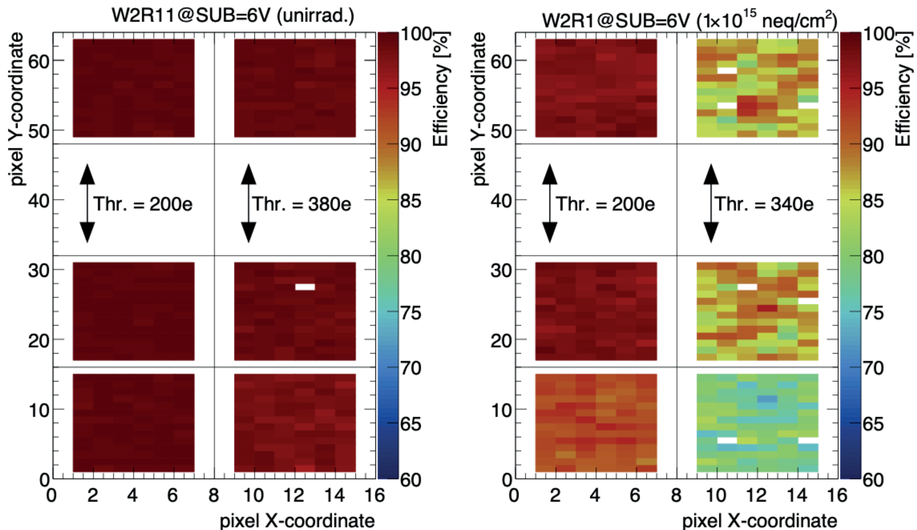


Figure 4.9: 2D efficiency map for a non-irradiated (left) and irradiated at 1×15 1 MeV n_{eq}/cm^2 (right) mini-MALTA sample measured with a 2.5 GeV electron beam. The pixel matrix is divided into sub-groups featuring different process modifications and desings. The left half of the pixel matrix features enlarged transistors, the right half features the standard MALTA front-end. The bottom row of sub-groups features the STD pixel flavour, the second row features the XDPW flavour, the third row implements different reset mechanisms (no data available) and the fourth row features the NGAP flavour. The empty bins represent dead pixels. The chips were operated at -6 V SUB voltage and -20°C, and were tuned to low threshold. Images are obtained from Ref.[88].

For samples irradiated at 1×10^{15} 1 MeV n_{eq}/cm^2, the efficiency response lies well above 97% when the operating threshold is below 200 electrons for the sensors with enlarged transistors. For samples irradiated to 2×10^{15} 1 MeV n_{eq}/cm^2 with larger transistors, an efficiency larger than 90% can be achieved [88].

4.4 MALTA on Czochralski Substrates

In order to generate a larger active sensor volume compared to the MALTA sensors fabricated on the high-resistivity epitaxial layer, MALTA sensors were fabricated on high-resistivity (3-4 kΩcm) Cz substrates. The substrate thickness varies between 100 and 300 µm. The production of CMOS sensors with small collection electrodes and the implant geometries introduced before (STD, XDPW, NGAP) on high-resistivity Cz substrates brings together the benefits of small electrode sensors and thicker detection layers. This combination ensures the retention of low pixel capacitance, resulting in low noise and low power

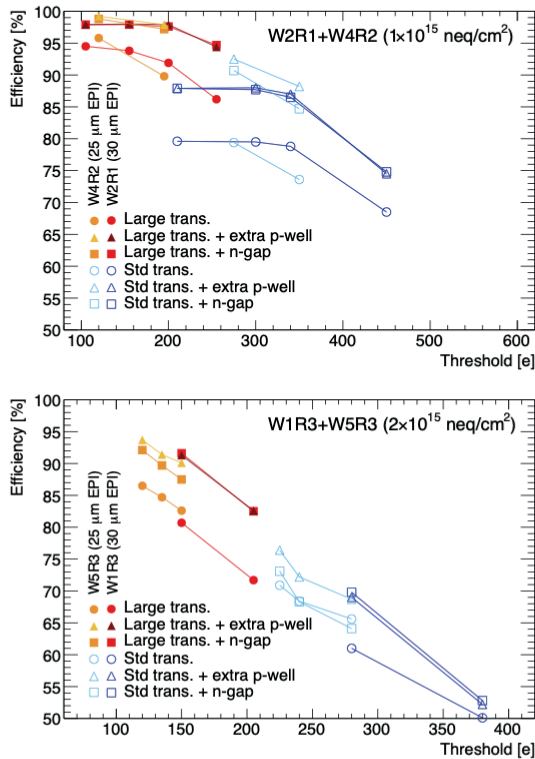


Figure 4.10: Efficiency versus threshold mean for neutron irradiated mini-MALTA samples at $1{\times}10^{15}$ 1 MeV n_{eq}/cm^2 (top) and $2{\times}10^{15}$ 1 MeV n_{eq}/cm^2 (bottom) measured with a 2.5 GeV electron beam. The chips were operated at -6 V SUB voltage and -20°C. Different sensor regions are presented: STD (circles), XDPW (triangles) and NGAP (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). Images are obtained from Ref.[88].

performance. Simultaneously, the signal amplitude is greatly enhanced due to the larger ionization charge present in thicker depleted sensors [11].

4.4.1 Performance of MALTA-Cz after Displacement Damage

To evaluate the influence on the overall performance of samples featuring diverse pixel designs and fabricated on a Cz substrate, the efficiency of the sensors was measured at medium operating threshold (approximately 300 e^-). Figure 4.11 illustrates the efficiencies (top row) and cluster sizes (bottom row) of Cz sensors irradiated at $1{\times}10^{15}$ 1 MeV n_{eq}/cm^2 for the STD, NGAP, and XDPW process

modifications. The S2 and S3 labels indicate that the data were obtained from MALTA sector 2 and 3 (Figure 4.4). In the left column, the efficiency (cluster size) is shown as a function of substrate voltage while maintaining a constant threshold. It is observed that NGAP and XDPW sensors exhibit improved corner efficiency. This behavior aligns with the qualitative predictions from TCAD simulations [11]. The right column demonstrates that sensors with these additional modifications, i.e. NGAP and XDPW, offer a wider operational range for threshold settings while maintaining higher efficiency and cluster size compared to STD Cz sensors. The improvements in corner efficiency, previously observed in epitaxial sensors, also qualitatively apply to sensors manufactured on Cz substrates. The publication presented in Chapter I, "Radiation Hardness and Timing Performance in MALTA Monolithic Pixel Sensors in Tower Semiconductor 180 nm" is dedicated to showcase some of the major advantages non-irradiated and irradiated MALTA-Cz samples have over MALTA on epitaxial substrate.

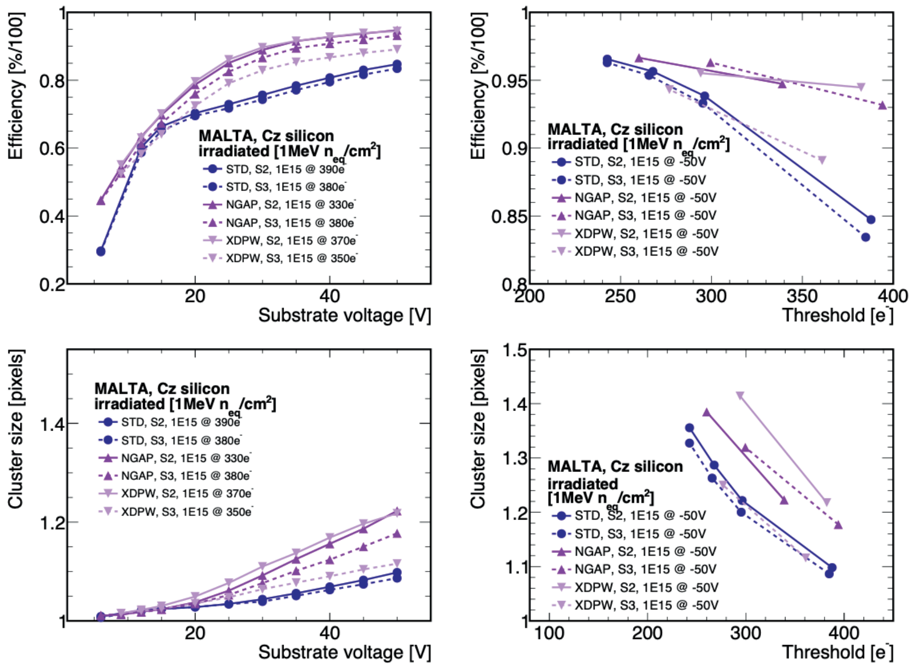


Figure 4.11: Sensor efficiency and cluster size for samples irradiated at 1×10^{15} 1 MeV n_{eq}/cm^2 on Cz substrate with different pixel flavour (STD, NGAP, and XDPW). Left figures shows the efficiency (top) and cluster size (bottom) as a function of substrate voltage with constant threshold across the samples. Right figures show the efficiency (top) and cluster size (bottom) as a function of threshold with constant substrate voltage of -50 V. Images are obtained from Ref.[11].

4.4.2 Applications of MALTA-Cz

Test beam measurements provide an opportunity to thoroughly characterize silicon sensors in a manner that goes beyond what can be achieved in a laboratory setting. To ensure an unbiased evaluation of the sensor response, an external reference is necessary. In the case of charged particle measurements, a beam telescope composed of several well-characterized sensors serves as this external reference by reconstructing the paths of the incoming particles. By comparing the performance of the DUT against the accurately defined trajectories of the beam particles, it becomes possible to conduct detailed studies that go beyond the inherent resolution of the DUT [48]. In 2021, a custom telescope with six MALTA tracking planes was build in the North Area of the SPS at CERN, illustrated in Figure 4.12.

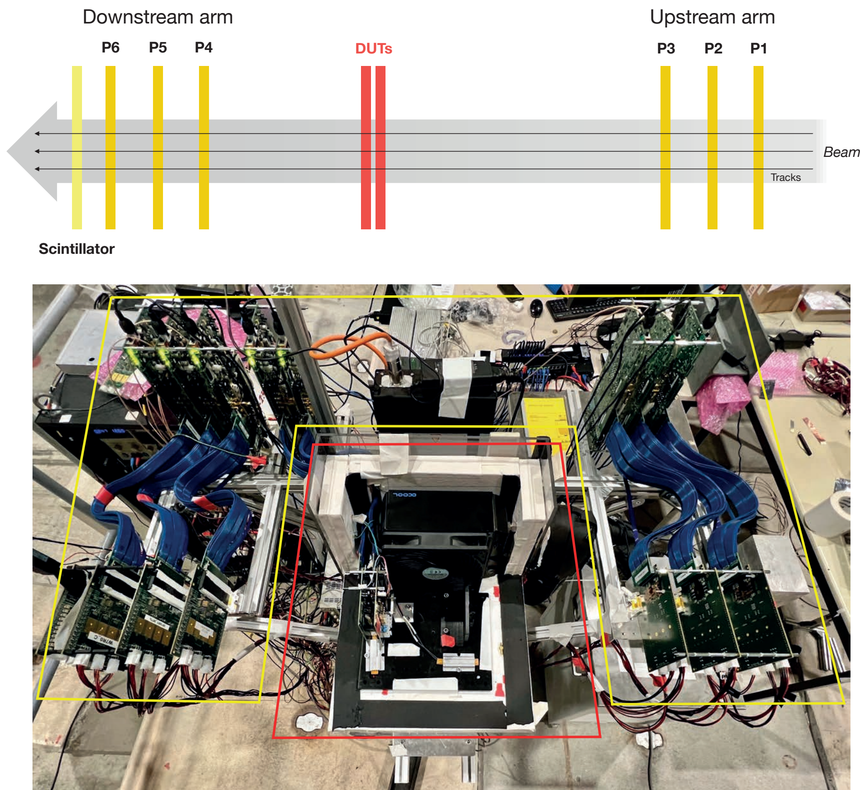


Figure 4.12: Top image: A schematic sketch of the MALTA beam telescope where the six tracking planes (P1 - P6), the devices under tests (DUTs), and the scintillator are indicated. Bottom image: Top view of the MALTA telescope. Equipment mounted on the main stage are inside yellow lines, cold box and DUTs positioned on DUT stage are inside the red lines.

The publication presented in Chapter II, "Performance of the MALTA telescope", is dedicated to the MALTA telescope and constitute a substantial portion of the research showcased in this thesis. This publication presents, for the first time, a comprehensive summary of the architecture and performance of the MALTA telescope during the test beam campaigns held at the SPS North Area in both 2021 and 2022. The MALTA telescope features, among others, a dedicated custom-built Trigger Logic Unit (TLU) with the possibility to trigger on up to four telescope planes, a scintillator for precise timing reference, and in-chip Region Of Interest (ROI) capabilities in order to serve both large and small DUTs ($< 2\times2$ cm^2). The MALTA telescope can operate at a maximum rate of 50 kHz, which is limited by the external readout.

The track time of the MALTA telescope, i.e. the averaged time of arrival of the fastest hit in the cluster for the six tracking planes, equates to σ_t=2.1 ns. The excellent time performance allows to separate individual track contributions, even in intense beam settings (6×10^6 particles/spill). In order to capitalise on the larger cluster size of the MALTA-Cz sensors at large bias voltage, compared to sensors fabricated on epitaxial layers, two MALTA-Cz STD samples were positioned closest to the DUT. This has allowed to improve the spatial resolution significantly. The measured resolution based on the linear regression approach of the full telescope was estimated at $\sigma_s = 4.1 \pm 0.2$ µm. Comparison with simulations, where the cluster size is not accounted for, have highlighted how the cluster size affects the spatial resolution improving the predictions by about 10% for various tested configurations. The motivation behind this was driven by the fact that the cluster position determination in MALTA relies on geometrical averaging. Consequently, when the cluster size exceeds 2, the (straight) particle hit is positioned closer to the pixel edge rather than the pixel center. The decision to incorporate only two MALTA-Cz STD tracking planes is driven by the restricted availability of this particular sample type. These results are discussed in more detail in the publication presented in Chapter II.

4.5 MALTA2

MALTA2 is the second full prototype of the MALTA family and takes center stage in the work presented throughout this thesis. It is approximately half of the size of the original MALTA, as its matrix counts 224 $\times$ 512 pixels (10$\times$20 mm), illustrated in Figure 4.13. MALTA2 inherited the signature asynchronous readout of its predecessors and implemented, similar to the mini-MALTA demonstrator, a shift-register as the slow-control protocol. Figure 4.14 illustrates the MALTA2 analog front-end layout, the design details of which are thoroughly discussed in Ref.[16].

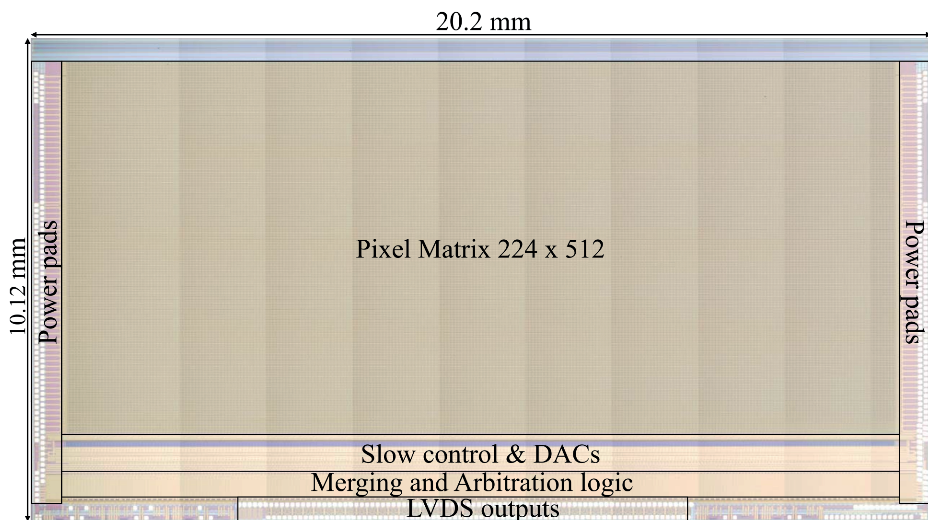


Figure 4.13: Layout of the MALTA2 matrix featuring 224×512 pixels. The power pads are positioned at the bottom, left and right side of the matrix. The image indicates the position of the slow control, DACs, LVDS output, and merging and arbitration logic. Image is obtained from Ref.[16].

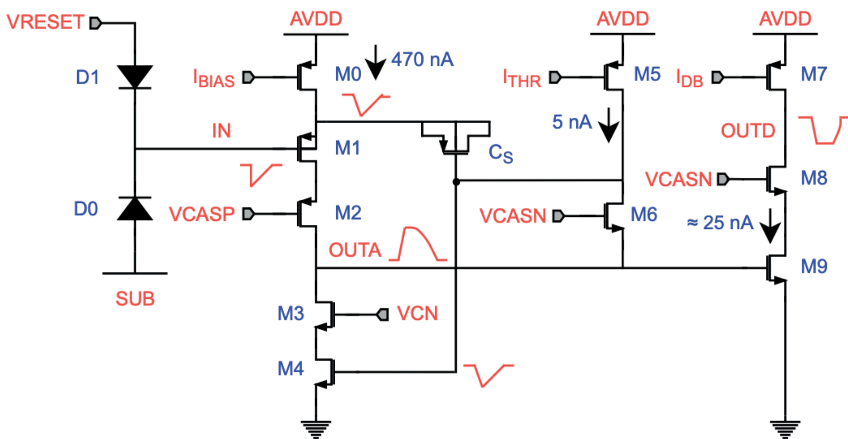


Figure 4.14: Layout of the MALTA2 cascode analog front-end. Image is obtained from Ref.[16].

One primary enhancement in the MALTA2 front-end involves the cascoding of the M4 transistor to achieve a substantial transconductance for optimal timing performance. Nevertheless, pushing the aspect ratio too high poses a risk of compromising on the output parasitic capacitance, detrimentally affecting gain and speed. Consequently, cascoding the transistor, among other advantages, serves to isolate it from the output node, affording greater flexibility in its sizing. In Ref. [16], Piro et al. highlighted that, for RTS noise, the transistors M1 and M4 play a crucial role, requiring an iterative sizing process. To mitigate RTS noise, the gate area of the M1 transistor was increased; however, this led to a larger effective sensor capacitance. The chosen gate size (0.18 μm^2) represents a compromise between capacitance penalties and noise considerations. As for the M4 transistor, a larger gate area (2.4 μm^2) was selected. This decision was influenced by its larger noise transfer function to the output node and since RTS noise is typically more pronounced in NMOS transistors. These modifications have allowed to increase the gain and decrease the noise, which is illustrated in Figure 4.15.

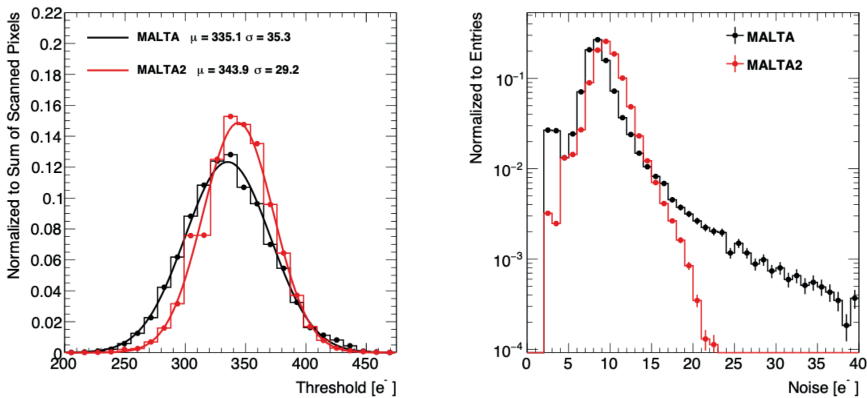


Figure 4.15: Threshold (left) and noise (right) distributions of a MALTA (black) and MALTA2 (red) sample. Both samples are non-irradiated, fabricated on an epitaxial layer, NGAP, 300 µm thick, high doping n^- layer, and operated at -6 V SUB bias. Images are obtained from Ref.[98].

Figure 4.15 showcases the threshold and noise distributions for a MALTA (black) and MALTA2 (red) sensor at the same operating conditions. The mean value of the Gaussian fit for the MALTA and MALTA2 threshold distributions are similar, μ=335.1 and μ=343.9 electrons, respectively, with an approximate 10% dispersion from the mean. The RMS values of the noise distributions lie at 3.5 and 2.25 for MALTA and MALTA2, respectively. The noise distribution shows a strong improvement for MALTA2, as the tail of the distribution is much less significant [98].

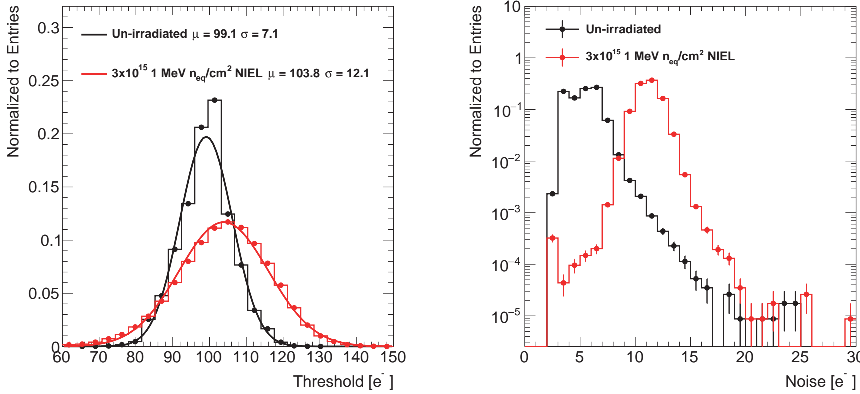


Figure 4.16: Threshold (left) and noise (right) distributions of non-irradiated (black) and neutron irradiated (red) at $3{\times}10^{15}$ n_{eq}/cm^2 MALTA2 samples. Both samples are fabricated on an epitaxial layer, XDPW, 100µm thick, low doping n^- layer and operated at -6 V SUB bias. The threshold and noise scans are obtained for the lowest threshold configuration achievable in that chip. Images are obtained from Ref.[98].

Figure 4.16 displays the threshold (left) and noise (right) distributions, similar to Figure 4.15, but illustrated for two MALTA2 sensors. The black curve corresponds to a non-irradiated sample, while the red curve represents a sample irradiated to $3{\times}10^{15}$ neq/cm^2. Threshold and noise scans were conducted for the lowest threshold configuration achievable for each respective sample. Despite exposure to a high irradiation dose, the sensor remains operational at low thresholds, exhibiting a threshold dispersion close to 10%. The noise distribution exhibits elevated noise at $3{\times}10^{15}$ neq/cm^2, attributed to the increased leakage current resulting from radiation damage. No significant non-Gaussian component is observed in the noise distribution, indicating no major contributions from Random Telegraph Signals (RTS) to the total noise.

These outcomes mark the initial characterization findings for MALTA2, a process in which the author of this thesis actively participated. It paved the way for subsequent test beam measurements of a large amount of MALTA2 samples, of which a selection is highlighted in Table 4.3. This summary underscores the extensive quantity of samples subjected to testing, which is a consequence of the large number of distinct operating parameters for the MALTA(2) sensor. These parameters include factors such as thickness, substrate type, and the doping level of the n^- layer. Of particular significance has been the latter, the n^- layer's doping level, which has been a central focus of attention within this research. A more thorough exploration of this aspect will be undertaken in the publication outlined in Chapter III, titled "Radiation Hardness of MALTA2 on

Czochralski Substrates." This paper reviews the performance of MALTA2 sensors on Cz substrates before and after irradiation. It showcases that by implementing various process modifications, MALTA2 has exceeded its initial performance target requirement (Table 4.1), i.e. radiation hardness $>10^{15}$ 1 MeV n_{eq}/cm^2 NIEL.

Overview of MALTA2 samples tested

Substrate	Pixel flavour	Total thickness [µm]	Doping level of n^- layer	Fluence level [$\times10^{15}$ n_{eq}/cm^2]
Epitaxial	NGAP	300	High	0, 1, 2, 3
Epitaxial	XDPW	100	Low	0, 1, 2, 3
Cz	XDPW	100	High	0, 1, 2, 3
Cz	NGAP	300	High	0, 1, 2, 3
Cz	NGAP	100	High	0, 1, 2, 3
Cz	XDPW	100	Very high	0, 1, 2, 3

Table 4.3: Overview of MALTA2 samples tested in the SPS test beam campaign from 2021 - 2023. The samples differ in substrate type (epitaxial or Cz), pixel flavour (NGAP or XDPW), total thickness (100 or 300 µm), doping level of n^- layer (low, high, and very high) and the fluence level.

Ultimately, MALTA and other monolithic CMOS contenders were not selected as the sensor technology for the ATLAS ITk upgrade. A review conducted by a CMOS committee in march 2019 determined that, among various considerations, CMOS monolithic sensors could only be a viable choice if a nearly fully functional chip and plug-compatible modules were readily accessible. Additionally, opting for a CMOS solution should not introduce substantial risks or delays to the established pixel upgrade baseline. Considering these factors and other organizational considerations, the review committee concluded that the implementation of hybrid modules would be the preferred approach for the pixel upgrade in the ATLAS project.

4.6 Conclusion

MALTA was developed with the aim of exploring its suitability for the Phase-II upgrade of ATLAS in the HL-LHC and other high-energy physics experiments. To meet the rigorous demands of these experiments, MALTA underwent numerous iterations of design and process modifications. The evolution of MALTA began with the standard process of 180 nm Tower Semiconductor CMOS imaging technology, featuring a small collection electrode. To enhance performance after irradiation, a process modification involving an additional n^- layer was introduced (STD). This enabled full lateral depletion and expanded the operational window of the detector. Upon observing a degradation of efficiency originating from the pixel corners after irradiation, two additional process

modifications, NGAP and XDPW, were implemented. These modifications were incorporated into the mini-MALTA demonstrator, along with the enlargement of selected transistors and a cascoded stage at the input branch. The NGAP and XDPW process modifications effectively mitigated the impact of the electric field minimum on charge collection, resulting in improved lateral field and charge collection in the corner regions of the pixel. Furthermore, in the case of MALTA on Cz substrates, an increase in the depleted region was observed with higher substrate voltage. This led to a significant improvement in efficiency when operating at higher voltages. These iterative developments ultimately led to the latest generation of the detector, MALTA2. The continuous improvements and adaptations of the detector design have paved the way for enhanced radiation tolerance and improved performance. As the field of high-energy physics progresses, the experiences and lessons learned from MALTA will undoubtedly contribute to the development of future detectors, as will be discussed in the concluding chapter of this thesis (Chapter 6).

Chapter 5

Silicon Pixel Detector Modules

The next generation of particle trackers will need to cover a significantly larger area than current ones. The current ATLAS Pixel Inner Detector covers an area of approximately 1.7 m^2. However, the proposed ATLAS ITk aims to cover a much larger surface area, with the pixel barrels alone accounting for 8.3 m^2. Reducing the dead areas, i.e. non-sensitive areas, is a critical concern for the development of large-area detectors. Large-area pixel detectors offer several advantages, including a reduced material budget due to fewer connections for data transmission and power, as well as lower assembly costs, as fewer modules can cover the same area. Therefore, efforts are focused on minimizing dead areas to maximize the active sensing area of the detectors [90]. The development and construction of hybrid and CMOS pixel detector modules plays a vital role in advancing our future collider experiments. In this study, a CMOS module refers to a flex (flexible PCB) and one or more chips. A hybrid module consists of multiple front-end chips bump-bonded to a single sensor die [90] and will specifically refer to the ITk pixel module presented in Chapter 3.4.1.

The work presented in this chapter highlights the possibilities for creating a large-area, lightweight module concept tailored for monolithic silicon sensors [90, 99]. It presents various research activities related to enabling technologies and the characterization of detector modules. The exploration of enabling technologies, including interconnection technologies, post-processing, packaging, coating, and embedding, is of large importance in the development of cutting-edge detector modules [100]. A comprehensive study has been conducted utilizing both hybrid (ITk pixel) and monolithic (MALTA) sensors, examining diverse aspects of such a module. First, this section will outline the endeavors undertaken prior to this thesis and by collaborative peers, such as Ref.[17, 101], in exploring diverse interconnection technologies for detector modules. Next, we will introduce an illustrative CMOS pixel detector module, initially presented in Ref. [90], which has served as the foundation for testing studies conducted within this thesis. Following that, the research that has been performed for this thesis on coating and embedding layers will be presented, encompassing both hybrid and monolithic pixel detector modules. Finally, test beam results of the CMOS detector module will be presented, for which the author has been actively involved in.

5.1 Interconnection Technologies

In the field of pixel detector modules, the choice of the interconnection technology plays a crucial role in ensuring efficient data transmission. The primary objectives in selecting an interconnect in this study are facilitating chip-to-chip data

transmission without distortion of the signal, maximizing the module's active area, enhancing mechanical robustness, and ensuring that the chosen interconnect can be produced through a streamlined manufacturing process. This section delves into the exploration of three distinct interconnection technologies employed in the realization of a multi-chip module utilizing the MALTA sensor. At the time of writing, two interconnection technologies (ACF and nanowires) were subjects of investigation by colleagues; however, they had not yet been incorporated into a testable device. Consequently, the work presented later in this chapter will center on modules employing wire-bonds as the chosen interconnect method.

5.1.1 Wire-Bonds

Wire-bonding is a commonly used technique in the field of microelectronics for connecting integrated circuits (ICs) or other electronic components to the substrate or packaging. It involves creating electrical connections between the IC and the package by using fine wires made of materials such as gold, aluminum, or copper. The wire-bonding technique is a solid phase welding process which uses thin wire and a combination of heat, pressure and/or ultrasonic energy. A bonding machine employs a bonding tool, often a fine-tipped capillary, which is maneuvered to establish contact between the bond pad on the IC and the corresponding pad on the substrate or package. Depending on the bonding agent, three major wire bonding processes can be performed: thermocompression, ultrasonic, and thermosonic. Once the bond is formed, the bonding tool creates a loop in the wire and then trims it to the desired length, ensuring proper electrical connection and tension [102]. Wire-bonding has found extensive use in silicon pixel detector technology [100] and has been crucial in various applications, such as medical imaging and high-energy physics research. The wire-bonds provide low-resistance electrical connections between the pixel detectors and the readout electronics, ensuring efficient signal transfer without significant losses. Furthermore, wire-bonding enables fast and reliable signal transmission between the pixel detectors and the readout circuitry, allowing for high-speed data acquisition [102]. Due to its ease of replacement or repair when necessary, the interconnection method has been established as highly reliable, adding to its appeal, particularly in the context of prototyping. This is depicted in Figure 5.1, showcasing the successful assembly and bonding of four MALTA sensors into a functional (monolithic) module, the MALTA multi-chip module, by the Bondlab at CERN. Given that the testing and characterization of this module have significantly contributed to this thesis, discussion regarding it will follow in Chapter 5.2 and 5.4.

However, it is worth noting that wire-bonding does have some limitations, such as the parasitic capacitance and inductance which can affect the performance of high-frequency applications [102]. Due to their fragile nature, wire-bonds are susceptible to damage, such as mechanical stress or thermal cycling. This in turn makes large scale integration and dense packaging challenging, especially when no additional protection is provided, further discussed in section 5.3. The

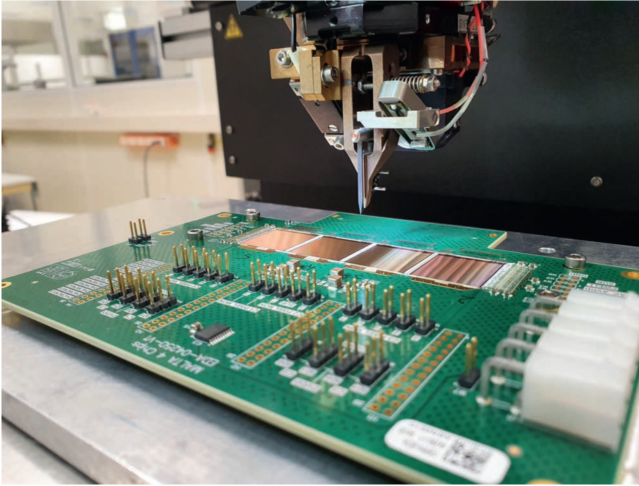


Figure 5.1: Image of an assembled and wire-bonded MALTA quad-chip board at the Bondlab at CERN. The wire bonding was performed through aluminium wedge wire-bonding. The bonding tool hovers above the four MALTA chips, arranged side by side on the PCB board.

wire-bonding process poses a significant challenge for the monolithic MALTA sensor, specifically due to the small pad size and the dense arrangement of pads (>700), located at the bottom, left and right side of the matrix. Moreover, the bond length between chips, presently approximately 1.9 mm within the MALTA multi-chip module, imposes a constraint on how closely the individual chips can be interconnected side by side. Consequently, this restriction hinders the reduction of the dead area, signifying the absence of detecting material, within the module.

5.1.2 Conductive Adhesives

Anisotropic Conductive Film (ACF) is an industry-standard interconnection technology utilized in the production of LCD screens [103]. In contrast to the bump-bonding process discussed in Chapter 3.4, the solder bumps are now replaced by conductive micro-particles embedded in an epoxy film. This technology involves establishing an electro-mechanical connection between the sensor and the ASIC through thermocompression of the ACF using a flip-chip device bonder [104]. The necessary pixel (MALTA) pad topology is achieved through Electroless Nickel Immersion Gold (ENIG) [105], which acts as a layer to elevate the ACF particles and prevent the formation of an oxide layer [101]. Figure 5.2 shows a 50 µm flexible PCB on a glass substrate with two (fused silica) test structures bonded by ACF from Dexerials [106].

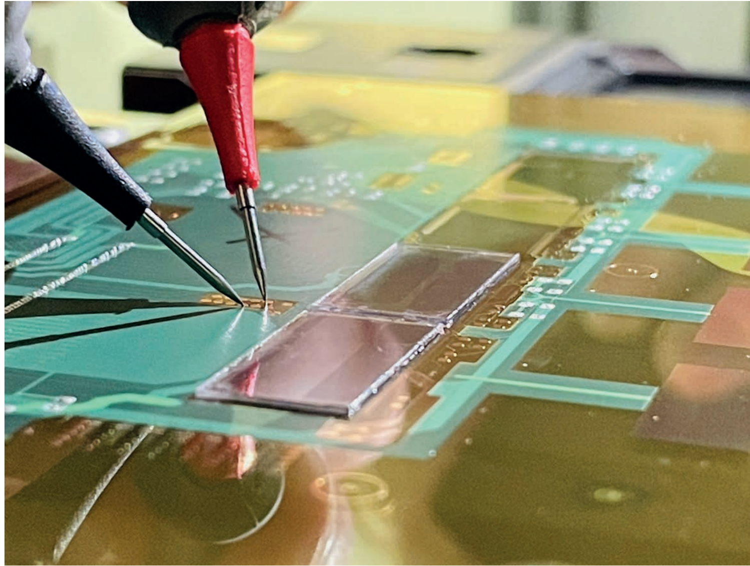


Figure 5.2: Electrical connectivity tests of (fused silica) test structures bonded on a 50 µm flexible PCB mounted on a glass substrate. The test structures are flip-chip bonded to the PCB, using ACF from Dexerials [106] as the interconnection technology.

ACF technology offers several significant advantages, making it an attractive choice for various applications, particularly in research and development (R&D) settings. One of the primary benefits of ACF is its cost-effectiveness. Compared to other interconnection technologies, ACF provides a more economical solution without compromising performance. Another advantage of ACF is its simplicity and lack of dependency on sophisticated tools. Unlike some complex interconnection methods, ACF does not require specialized equipment and is a maskless and in-house assembly technology. Furthermore, ACF proves highly advantageous for prototyping purposes in R&D. Its ease of use and simplicity allow for quick prototyping and testing. However, one of the primary limitations of ACF technology is the lack of proven radiation hardness and longevity in extreme conditions, such as those encountered in high-energy physics experiments or extremely cold environments. Pioneers adopting this technology have not yet performed comprehensive testing on these aspects. Consequently, it is uncertain how well ACF would perform or maintain its functionality under such demanding circumstances. Further research and testing are required to determine the feasibility and durability of ACF in these specific conditions [101].

5.1.3 Nanowires

The use of nanowires is the final interconnection technology explored for the modularization of the MALTA chip, described by Weick et al. in Ref.[17]. For this bonding process, a titanium layer is deposited on the aluminum MALTA pads together with a gold finish to prevent oxidation. This serves as a base for a copper seed layer on which nanowires, with a diameter of approximately 100 µm and a length of several micrometers, are grown. The subsequent bonding process is referred to as the glueing process, where the nanowires are only grown on one side of the device. A non-conductive glue serves as an underfill which allows for reduced pressure and heat and enhanced mechanical stability during the flip-chip process. Figure 5.3 shows a scanning electron microscope (SEM) image of two MALTA pads on which copper nanowires are grown.

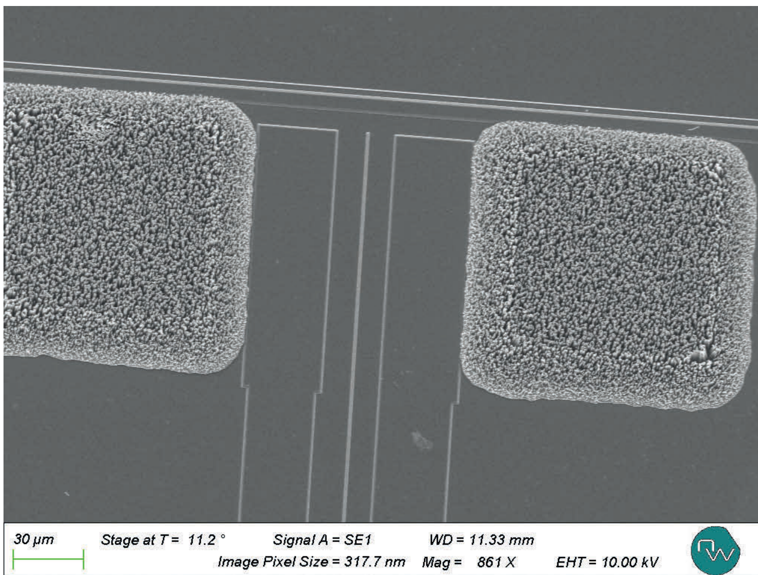


Figure 5.3: SEM image of copper nanowires grown on two squared MALTA pads (88×88 µm). The diameter of the nanowires is approximately 100 µm, with a length of several micrometers.

Using nanowires as an interconnection technology offers several significant advantages, particularly in terms of electrical performance. One of the key benefits is the capability to achieve extremely low resistance through a cold welding process. This process allows the nanowires to establish strong, reliable connections with minimal resistance, ensuring efficient and high-speed transmission of electrical signals. Additionally, nanowire interconnections exhibit low parasitic inductance and capacitance. This characteristic enables improved signal quality, reduced latency, and enhanced overall system performance. Furthermore, nanowire interconnections can be applied either on the targeted carrier or directly on the

chip, whether at the chip level or wafer level. This versatility allows for flexible integration into different device architectures and manufacturing processes.

One of the primary drawbacks of employing nanowires as an interconnection technology is the inherent complexity of the process. Implementing nanowires involves two lithography processes, adding an additional layer of intricacy to the fabrication procedure. This complexity increases the overall manufacturing time and may require specialized expertise to ensure accurate and reliable results. Another significant downside is the cost associated with prototyping using nanowires. Since the fabrication process is performed on a wafer level, it incurs higher expenses compared to other interconnection techniques. Prototyping using nanowires requires the use of expensive equipment and materials, further contributing to the elevated costs.

5.2 Modularization of the MALTA Sensor

The MALTA chip features CMOS transceiver blocks positioned at the left and right edges of the chip, along with most of the power supply pads, also shown in Figure 4.4 in Chapter 4. This design enables the assembly of modules by arranging individual chips side-by-side. Each chip can be configured to receive data from a CMOS transceiver on one side, merge the data collected from its own matrix, and transmit the combined data stream to a CMOS transceiver on the other side. Ultimately the data stream is transmitted to the LVDS output located at the bottom.

Towards the development of a large area, lightweight monolithic module with MALTA sensors, several multi-chip modules have been built and characterized. This included the so-called dual chip board, with two MALTA sensors, and the quad-chip board, with four MALTA sensors. A dedicated carrier board, see Figure 5.4, was designed for the assembly of these modules. The interconnection technology between the chips and to the PCB were realized using ultrasonic aluminium wedge wire bonding. The carrier board features a custom pad layout that allows the chips to be bonded consecutively, where first a minimal bonding sequence is followed in order to validate the readout and configuration of the respective chip. This procedure allows for trouble-shooting in case any complications are encountered during the highly complex wire-bonding process, as a MALTA quad-chip board features more than 2500 wire-bonds.

Similarly to the single MALTA detector, the carrier board is interfaced with the Xilinx KC705 Evaluation Board with a Kintex-7 FPGA via FPGA Mezzanine Card (FMC) connector [107]. It interfaces the slow control and reference bit of each individual chip as well as the LVDS readout of the primary chip. This LVDS output of the primary chip is responsible for the readout of the entire module, as all secondary chips forward their data from left to right (following the geometry

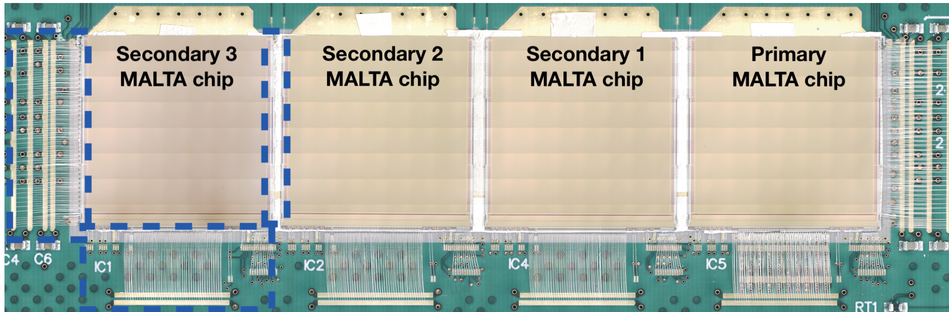


Figure 5.4: Image of fully assembled and functional MALTA quad-chip board module using ultrasonic aluminium wedge wire bonding. The three (secondary) left MALTA chips forward data to the primary chip (right) where all data is finally is read out. For the third secondary MALTA chip (far-most left), the presence of wire-bonds are indicated with the blue dashed line.

presented in Figure 5.4) via their CMOS transceivers. A primary objective in the development of the MALTA multi-chip module is to showcase the achievement of seamless data transmission across all four chips, ensuring the preservation and integrity of the data without any distortion or loss. The outcomes of the tests conducted to validate this critical aspect of the module are detailed in Section 5.4.

Wire-bonds offer a reliable and cost-effective solution for connecting the various components of pixel detectors, as has been demonstrated for the MALTA multi-chip module. As the wire bonds are very delicate, they require for large scale integration a protective housing. The next section explores this vital aspect: protecting and embedding pixel detectors with coating and/or protective layers. These layers not only shield the module from external factors but also enhance their performance and longevity. These studies will be demonstrated on both hybrid pixel detectors as monolithic multi-chip modules.

5.3 Coating and Embedding Studies for Pixel Detectors

In the pursuit of advancing pixel detector technology, this section focuses on an essential aspect of this development: coating and embedding studies. In this chapter, the work conducted for this thesis on diverse protective coating and embedding methods for both hybrid and monolithic modules is presented. The modules discussed earlier, namely the ITk pixel module (Chapter 3.4.1) and the MALTA multi-chip module (Chapter 5.2), serve as the subjects of investigation in this endeavor. By delving into these studies, we aim to enhance the reliability, functionality, and longevity of these detector components in various scientific and technological applications.

5.3.1 Embedding Studies for Wire-Bonded Pixel Detectors

The encapsulation, or embedding, of wire-bonds can protect them from possible mechanical damage, corrosion, or other deteriorating effects [108, 109]. As both the ITk Pixel and the MALTA multi-chip module discussed in this thesis use wire-bonds as the interconnection technology, they face similar challenges and demand careful consideration on the protection of their wire-bonds during their integration in large scale systems. The goal of encapsulation is to completely wrap the wire-bonds with an encapsulant, i.e. a glue-like substance, for full coverage within a defined area. The process of encapsulation should be achieved with a high degree of precision, as the encapsulant should not leak beyond the boundaries of the module. The utilization of encapsulants to safeguard wire-bonds has also been investigated by other LHC experiments, including CMS for their pixel detector upgrade [110].

5.3.1.1 Hybrid Technologies

The ATLAS ITk Collaboration conducted studies, prior to this studies, to explore an appropriate encapsulant for safeguarding the wire-bonds located on the ITk modules [111, 112]. The desired properties for the encapsulant are listed in Table 5.1, following the requirements of the ATLAS ITk Pixel detector discussed in Chapter 2.

Requirements for Encapsulation	
Linear CTE [ppm/°C]	< 25
Expansion/shrinkage upon moisture and temperature [%]	<0.5
Filler particle size [µm]	<25
Curing time at room temperature [hours]	<72
Electrical resistivity [Ωcm]	$>10^{14}$
Glass transition temperature [°C]	>60
Viscosity [cps]	1 - 7×10^4
Radiation Hardness (NIEL) [1 MeV n_{eq}/cm^2]	$>10^{15}$
Radiation Hardness (TID) [Mrad]	100
Thermal cycling [range in °C]	-45, +40 (10x), -55, +60 (1x)

Table 5.1: Overview of the main requirements for an encapsulant to protect the wire-bonds for Outer Barrel ITk Pixel modules of the ATLAS ITk Pixel detector.

Several encapsulants, such as EP30TC [113], Vitralit 1605 [114], Sylgard 170 [115], and Dymax9001 [116], were under consideration by Kobayashi et al. [111, 112]. The objective of these investigations was to evaluate the performance of the encapsulants following exposure to irradiation and thermal cycling. For the irradiation studies, diverse samples containing encapsulants underwent proton irradiation, reaching up to 1.6×10^{16} n_{eq}/cm^2 (3.5 MGy and 10 MGy), as well as gamma irradiation up to 15 MGy. The outcomes of these investigations

unveiled a wide range of observations, including alterations in the color and shape of the encapsulant (EP30TC), disconnection of bump-bonds (Sylgard 170), solidification of the encapsulant (Vitralit), and thermal shrinkage along with disconnected wire-bonds observed across all encapsulants. In summary, the results from visual inspections and bump disconnection studies suggested that the available encapsulants did not meet the desired performance criteria, as discussed in Ref. [112, 117]. Consequently, it was proposed that an alternative protective mechanism should be employed for wire-bond protection in each system, tailored to its specific configuration. For further information, please consult [5] and [117].

However, for the ITk Pixel Outer Barrel RD53A demonstrator [118] the decision to utilize an encapsulant for wire-bond protection was reevaluated [119]. The measurements and tests conducted in this context constituted a substantial effort conducted for this thesis. The Outer Barrel demonstrator is a large-scale testing vehicle for integration and system-related aspects pertaining to the forthcoming outer barrel system within the ATLAS ITk pixel detector [120]. Notably, the encapsulant employed for the demonstrator need not adhere to radiation hardness, given that the demonstrator itself will not be subjected to this specific environments. The presence of the encapsulant ensures the protection of wire-bonds from potential damage and safeguards them during the installation process onto a longeron. The longeron serves as a flat, carbon-based support structure located in the outer barrel region of the ITk Pixel detector, formerly known as the stave in the Inner Detector [5]. For more information regarding the design specifications of the Outer Barrel Demonstrator, please refer to Ref. [118, 120].

The chosen encapsulant for the modules assembled in the Outer Barrel Demonstrator was Sylgard 186 [121], a two-component silicone. The selection of Sylgard 186 was motivated by its flexible rubber properties, ability to withstand mechanical shocks, and user-friendliness. The two-part encapsulant is mixed, poured into a syringe, and connected to an air-powered fluid dispenser. The syringe, held in a three-axis stage, controls the shape of the encapsulant deposited on the wire-bonded module based on the dispensed volume and the rate of motion. The modules used in this study featured the RD53A [122] readout chip. The usage of these so-called RD53A modules served various purposes, including initial design validations, testing prototype for module assembly and Parylene masking (refer to section 5.3.2), and served as the primary modules for integration into the Outer Barrel Demonstrator. Figure 5.5 shows an RD53A module prior and post encapsulation. The dashed blue line indicates the position of wire-bonds on one side of the module. The bottom image illustrates the presence of a glue-like epoxy, Sylgard 186, post-encapsulation. This endeavor was considered successful, as no wire-bonds suffered damage during the process, and there was no observable indication of encapsulant leakage beyond the module's boundaries.

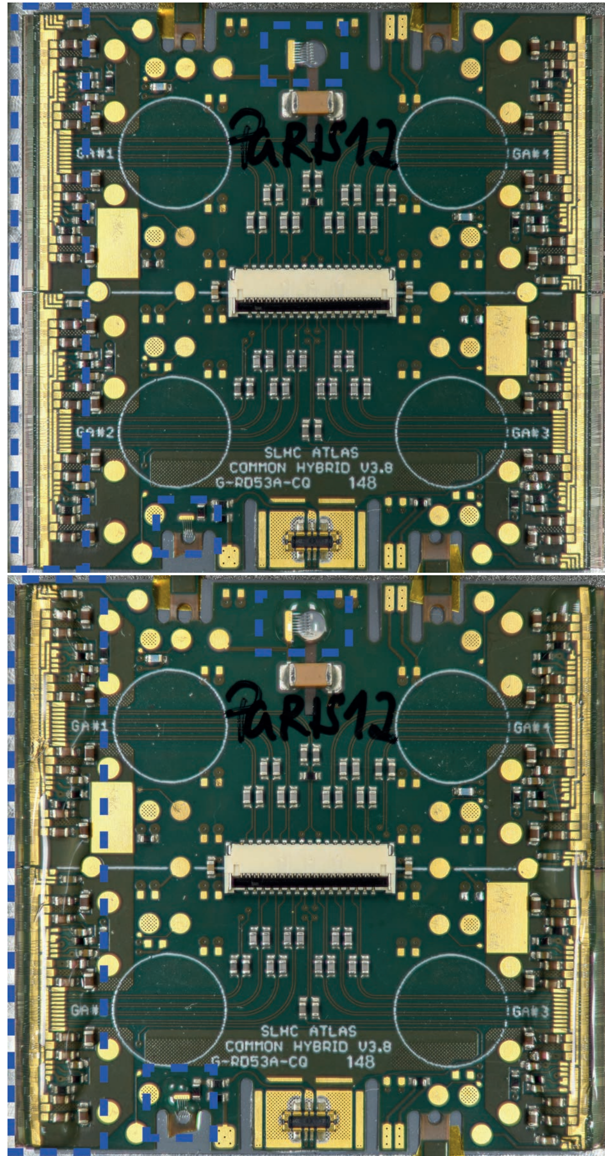


Figure 5.5: Two stages of encapsulation of the wire-bonds of RD53A modules. Top: Module prior to encapsulation. The blue dashed boxes indicate the areas with wire-bonds, also located on the right side of the module. Bottom: Image of a fully encapsulated module. All wire-bonds are encapsulated with Sylgard 186, the transparent glue-like epoxy. The blue dashed boxes indicate the encapsulated area, also present on the right side of the module. No evidence of encapsulant leakage beyond the module's boundaries is apparent, signifying the success of the procedure.

Once the Sylgard encapsulant has thoroughly cured, the subsequent phase involves mounting the modules onto the longeron of the Outer Barrel Demonstrator, as illustrated in Figure 5.6. In the image, the blue arrows indicate the wire-bond positions (following the same symmetry as depicted in Figure 5.5) along with the encapsulant. The image displays the arrangement of 18 RD53A modules, organized into two separate serial powering chains. Among them, six RD53A modules are interconnected in series, while the remaining 12 modules are collectively powered. Notably, the orange pigtail flexes seen in the image are positioned to traverse perpendicular to the wire-bonds. The encapsulant serves to shield the wire-bonds from potential harm stemming from the flex pigtails and safeguards them during the installation process onto the longeron.

5.3.1.2 Monolithic Technologies

The knowledge and experience that was gained from the encapsulation studies for the ITk Demonstrator formed an important starting point to explore the use of an encapsulant for the MALTA multi-chip module. These investigations were undertaken within this thesis to determine whether encapsulation could viably safeguard the wire-bonds of monolithic modules without introducing any adverse impacts on module performance. Since individual MALTA chips and MALTA modules are handled regularly during test beam campaigns, providing supplementary protection for the wire-bonds is considered beneficial. The MALTA dual chip module, a monolithic module with two MALTA sensors, was used for these studies. As the wire-bonds of the module are of height $\sim$800 µm, a thick layer of the encapsulation material was required, as can be seen in Figure 5.7.

To ensure that the encapsulation of wire bonds does not impact the module's performance, a threshold scan was conducted both before and after the module's encapsulation process, as depicted in Figure 5.8. The figure illustrates the threshold distribution of one MALTA chip in the MALTA dual chip module, specifically for sectors 0, 1, 2, and 3 (see Figure 4.4 in Chapter 4). As explained in Chapter 4, the front-end per sector differs in MALTA and therefore the threshold dispersion among different sectors observed in Figure 5.8 aligns with the expectations. The mean threshold value of the distribution and its sigma (σ) is determined by fitting a Gaussian function to the core of the distribution. As seen in the figure, the encapsulation of wire bonds using Sylgard did not induce any noticeable changes to the discriminator's pixel threshold, which is in line with the expectations from LHC experiments such as ATLAS ITk [119] and CMS [110]. In the realm of module development, upcoming efforts aimed at creating large area, lightweight modules are centering their focus on the adoption of a flip-chip bonding technique with ACF or nanowires as the interconnection technology. This approach deliberately avoids the utilization of wire-bonds for interconnection purposes. Therefore, the exploration of encapsulant usage has not been pursued further.

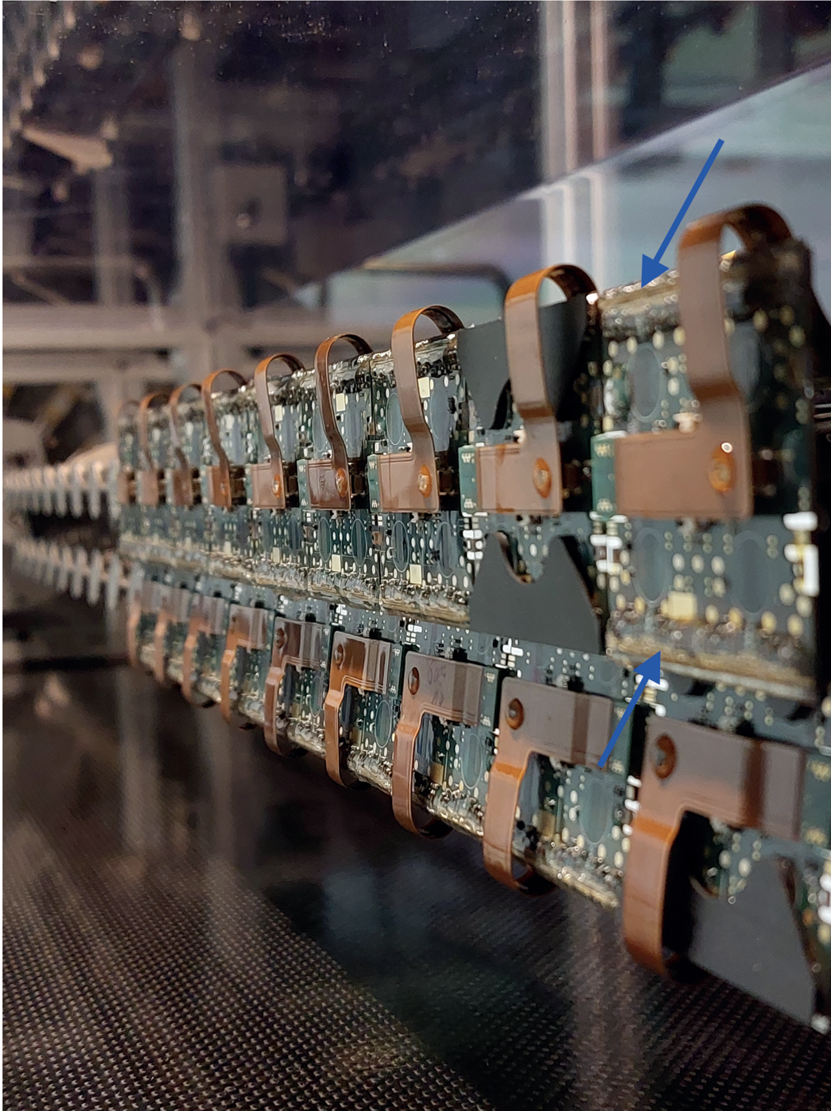


Figure 5.6: The ITk Pixel Outer Barrel RD53A demonstrator features the installation of 18 RD53A modules onto the longeron. These modules are effectively protected by the application of the encapsulant Sylgard 186, a transparent adhesive substance that is carefully deposited along the outer edges of the modules (indicated by the blue arrow), specifically targeting the wire-bond areas. The Sylgard 186 encapsulation ensures the safety and integrity of the wire-bonds on the modules.

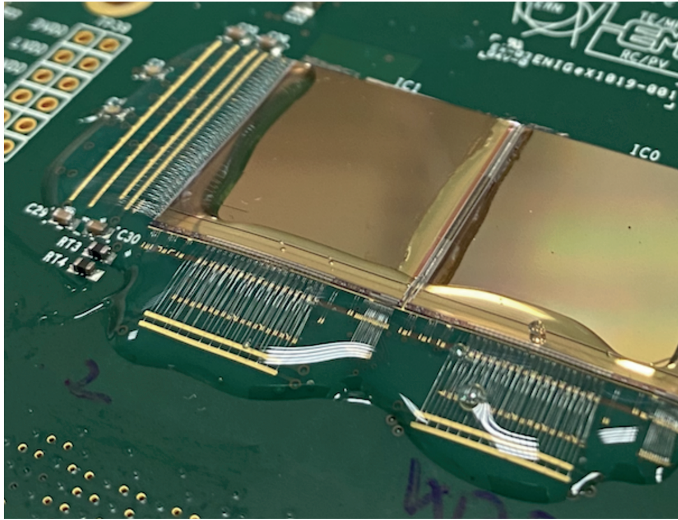


Figure 5.7: Image of MALTA dual-chip module with Sylgard 186 encapsulation of the wire-bonds. The wirebonds are located at the bottom, left and right side of the single MALTA chips.

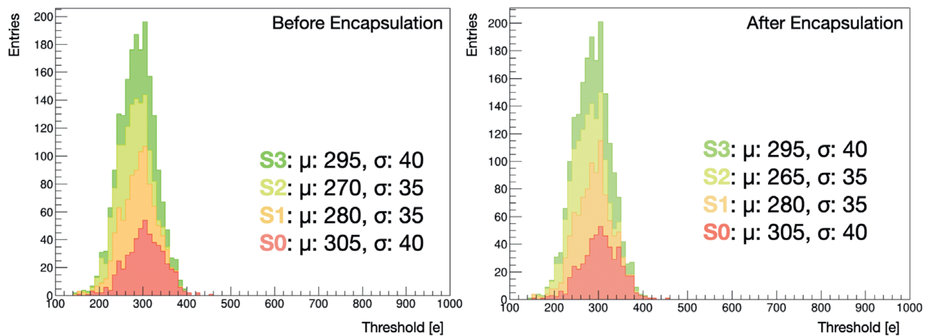


Figure 5.8: Threshold distributions of four (S0-S4) sectors of one MALTA chip in the MALTA dual chip module. The quoted mean (μ) threshold value is obtained by using a Gaussian fit to the core of the threshold distributions separately, from which the dispersion (sigma) is extracted. After Sylgard encapsulation (right), no appreciable difference in the threshold can be observed. As only a sub-set of pixels are scanned, threshold and sigma values are rounded.

5.3.2 Coating Studies for Pixel Detectors

As became evident in the previous section, embedding layers, such as Sylgard 186, offer protection to wire-bonds but are not yet developed such that they can survive the demanding environments of high energy physics experiments. This has motivated the search for a coating layer that can provide protection against environmental corrosion. More specifically for hybrid pixel detectors additional spark protection between the sensor and front-end is required. Parylene N [123] was identified as a highly suitable candidate for this purpose, as will be discussed in more detail below. Parylene is the generic name for a family of polymers referred to as the poly(para-xylylene) derivatives. In industry there are various grades of Parylene used, such as Parylene N, Parylene C, and Parylene HT [123]. It is deposited through Chemical Vapor Deposition (CVD) with a thickness that can vary between 1 µm to several mm. Due to its dielectric strength, i.e. up to a few hundred MV/m, it is ideal for protecting small spaces, such as between a sensor and ASIC, as well as individual wire-bonds [50].

5.3.2.1 Hybrid Technologies

The usage of Parylene coating for ITk pixel modules serves multiple purposes [124]. Firstly, it effectively prevents high voltage discharge from occurring at the sensor's edge to the ASIC ground pads, a phenomenon often observed during operations at high bias voltages of the sensor (800 V) [5]. Additionally, the coating provides protection against environmental and handling damage, while also enhancing the strength of wire-bonds and bump-bonds within the modules [125]. In the past, both Parylene N and Parylene C have been utilized for coating purposes within ATLAS and LHCb [126]. However, due to the chlorine content in Parylene C, a potential risk exists that the halogen is released onto the substrate, leading to corrosion. To mitigate this concern, the ATLAS ITk collaboration has opted for Parylene N as the preferred material for coating. The desired thickness for the Parylene coating is set at 7$\pm$2 µm to meet the aforementioned high voltage requirements [5].

To qualify as a vendor for the deposition of Parylene on a significant number of pixel modules, it is mandatory to undergo a series of qualification tests [127]. These tests have been established to guarantee uniform evaluation of all potential vendors. The aim is to assess them based on standardized criteria and ensure alignment with the specifications outlined by ATLAS ITk and are listed in Table 5.2. Every qualification item will be discussed further below and as these tests are a part of the work carried out in this thesis, several test results will be showcased. The qualification tests are not required to be performed on actual pixel modules; instead, they can be performed using glass or wire-bonded silicon samples, depending on the specific test being conducted. If the results of any of the qualifying tests are deemed unsatisfactory, the vendor has the option to adjust or optimize their process settings accordingly.

Specification for Parylene Coating for ITk			
Qualification	**Requirement**	**Test before coating**	**Test after coating**
Thickness	7 ±2 µm	Substrate flatness and thickness	Thickness
Adhesion	ISO2409 Class 0 or 1	Clean surface with isopropyl alcohol and compressed air	Cross-cut test
Dielectric strength	No	IV scan up to 200 V	IV scan up to 900 V (100 cycles)
Wire-bond pull strength	>10 g after coating	Pull test on non-coated long wires (4 mm) and short wires (1 mm)	Pull test on coated long wires (4 mm) and short wires (1 mm)

Table 5.2: Requirements tests for vendor qualification for Parylene coating. The type of qualification, its requirement and the qualifying tests are described.

The first test involves measuring the thickness of the sample before and after coating, ensuring the coated layer falls within the specified range of 7 ±2 µm. This measurement is typically carried out using a profilometer with an accuracy of ≤2 µm. Next, the adhesion of the Parylene to the substrate is assessed using a cross-cutter multi-blade, which scratches the surface of the Parylene. The adherence of the Parylene to the substrate is evaluated based on the ISO2049 standard specifications and should fall between 0 and 1. Figure 5.9 presents two silicon samples coated with Parylene, after scratch tests. In the left image, the sample shows significant amount of flaking of the Parylene after the scratch test, whereas the right sample exhibits cleaner cuts. The right image of Figure 5.9 meets the requirements listed in Table 5.2, as the degree of flaking is not visible. It has been suggested that the vendor responsible for the coating of the left sample did not utilize a Silane primer prior to the coating process, which has been identified as a crucial step in preventing the flaking of Parylene after coating as it allows for better adherence of the coating layer.

As soon as the thickness of the Parylene adheres to the specifications of Table 5.2, the dielectric strength can be measured, as these two parameters are interrelated. The purpose of this dielectric strength test is to recreate the structure of the pixel sensor and readout chip and evaluate the dielectric strength of this configuration following the application of a known thickness of Parylene by a vendor or institute. To conduct these tests, a dummy module is used by using metal or silicon sensors (diodes or single pixel sensors) along with a suitable spacer that mimics the bump height of a module or an actual single chip module. Before the coating process, an external bias voltage of 200 V is applied for one cycle. After coating, an external bias voltage is applied in the range 0-900 V at a rate of 5V/s measuring the IV characteristics for hundred cycles. Furthermore, in this work, the test structure underwent X-ray irradiation (up to 660 Mrad) after

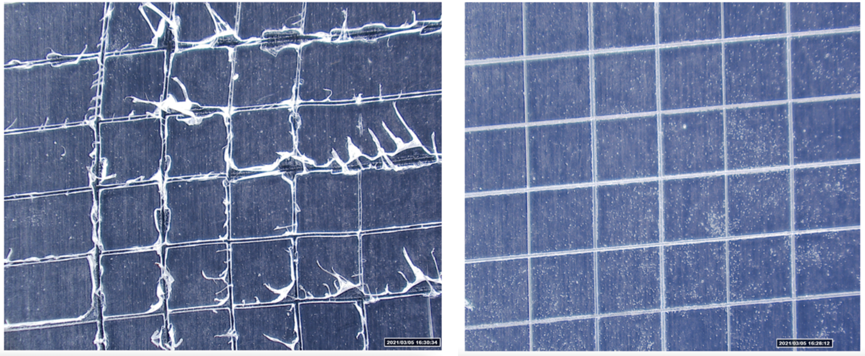


Figure 5.9: Left: Image of an unsuccessful Parylene coated silicon dummy after cross-cut test, where a lot of flaking of the Parylene can be observed. Right: Image of a successful Parylene coated silicon dummy after cross-cut test. Clean cut lines are visible due to the use of a primer (Silane) prior to coating.

the coating process. Subsequently, various IV measurements were performed after sequential thermal cycling. Figure 5.10 illustrates an IV curve (0-900 V) after four different thermal cycling steps: prior to thermal cycling (black), after 10x -45°C,+45°C + 1x -55°C,+ 60°C (red), after 25x -55°C,+ 60°C (green), and after 25x -55°C,+ 60°C (blue). The red and green curves represent 100 cycles of the IV curve, the black curve shows only one cycle, and the blue curve shows 1.5 cycles, as the leakage current increased significantly, indicating a breakdown of the sensor. This is primarily explained by the radiation damage to the SiO_2-Si interface of the diodes causing high leakage currents. Due to the extended duration of the x-ray irradiation campaign, various test structures were distributed among multiple institutes, and therefore only one sample is shown in this work.

Finally, the last qualification measurement is the wire-bond pull test, shown in Figure 5.11. Three Parylene coated wire-bonded samples containing 200 short (3mm) and 200 long (10mm length) 25μ aluminum wire bonds are used for this test. The aim is to quantify the mechanical strength of the deposited Parylene. Half of the wire-bond should be pulled before coating and the remaining half after coating. The mean pull strength (µ) and their standard deviation (σ) are recorded. After Parylene coating, the mean pull strength shifts for long wire bonds from approximately 6.5 g to 9.5 g. For short wire bonds, the shift is more dramatic, as the mean pull strength goes from approximately 11 g to 23.5 g. As the requirements (Table 5.2) state that after coating the mean pull length should be >10 g, this test is deemed successful.

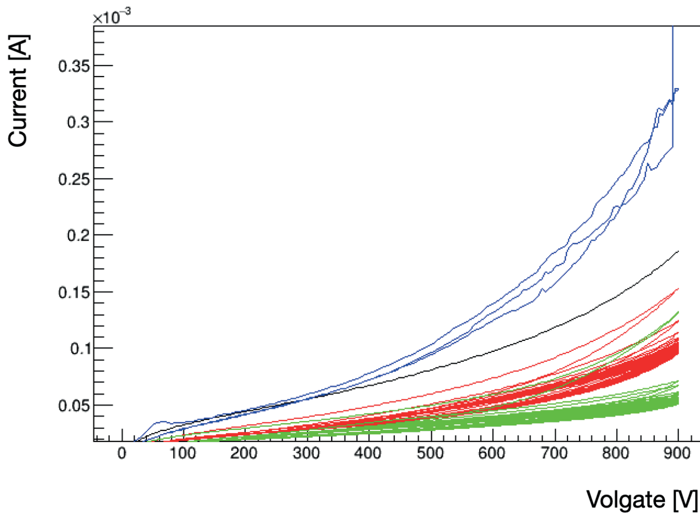


Figure 5.10: IV curve indicating the electrical characteristics of a Parylene coated high-voltage test structure that has undergone X-ray irradiation equivalent to 660 Mrad. The black curve represents a single scan conducted immediately after irradiation. The red curve illustrates the same scan (100 times), but after subjecting the kit to 100 cycles of thermal cycling (10x -45°C,+45°C + 1x -55°C,+ 60°C). The green curve reflects the IV measurements (100 times) following 25 cycles of thermal cycling between -55°C,+ 60°C. The blue curve represents the last IV measurements taken after 25 cycles of thermal cycling under the same conditions. The final IV curve was only performed 1.5 times due to a significant increase in leakage current, indicating a breakdown of the sensor. All IV measurements covered the voltage range of 0-900 V.

Prior to coating by a qualified vendor, the pixel modules are masked, as data connectors, power connectors, and specific areas of the module flex should not be covered with Parylene [125]. While certain vendors provide internal masking services, the delicate nature of pixel modules and the high associated costs have led to the decision of performing Parylene coating masking at module building institutes before sending the modules to the coating vendor. Within the ITk collaboration, various masking procedures for Parylene coating of ITk Pixel modules have been developed in order to make them reproducible amongst all module building groups, i.e. no reduction in the yield of the module production. The masking procedure was developed initially within the RD53A program and was later adopted to the ITk pixel module. The three stages of Parylene coating (prior to masking, post masking, post demasking) are visualised in Figure 5.12.

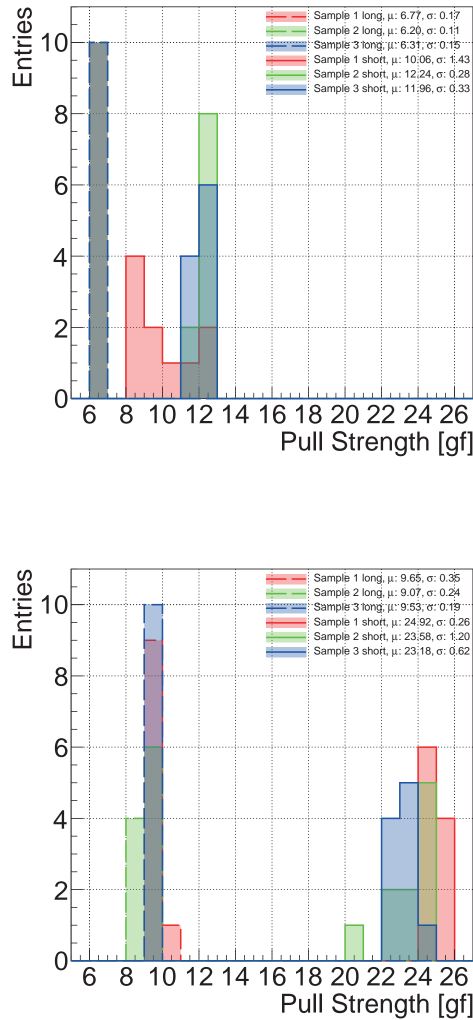


Figure 5.11: Before (top) and after (bottom) Parylene-coated wirebond samples containing 3 samples (red, green, blue) with 200 short (3 mm) and 200 long (10 mm) 25µm Al wire bonds. The mean (µ) pull strength and their standard deviation (σ) are indicated, quantifying the mechanical strength of the deposited Parylene. After Parylene coating, the mean pull strength shifts for long wire bonds from approximately 6.5 g to 9.5 g. For short wire bonds, the shift is more dramatic, as the mean pull strength goes from approximately 11 to 23.5 g.

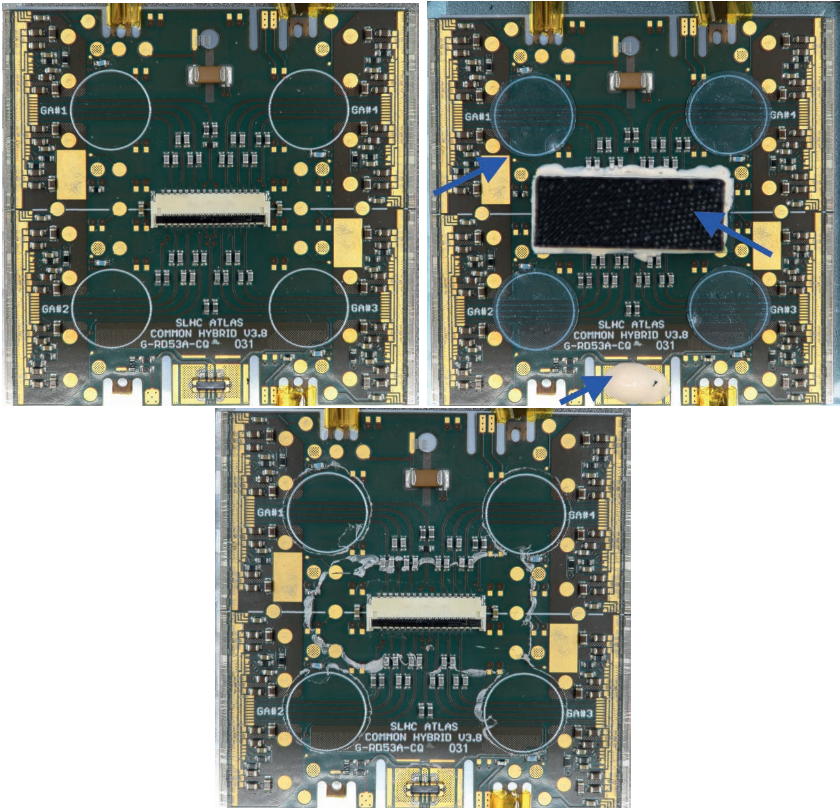


Figure 5.12: Three stages of Parylene coating of a RD53A module. Top left: Module prior to masking and Parylene coating. Top right: Module after masking and prior to Parylene coating. The features that are masked are indicated by the blue arrows. Blue dicing tape is located on the pick-up points (circular features) and backside of the module. Black cap protects the power connector with a silicon glue. Bottom: Module after Parylene coating and demasking. A small degree of delamination (flaking) around data connector can be observed.

The masking procedure combines the usage of dicing tape, silicone glue and 3D printed caps, or masks, to protect selected components and areas of the modules during the Parylene coating. Dicing tape is applied on the bare silicon back side of the flex and on the pickup points located on the flex (circular features). 3D printed covers (black) protect data and power connectors and are fixated onto the flex by a silicone glue (white). Samples are placed back in the carrier after masking and are packaged in such a way that there is minimal intervention from the vendor during coating. Upon reception of the module, after Parylene coating by a qualified vendor, the modules are demasked. The removal of the dicing tape is carried out with tweezers after light scoring of the coated tape edge to give minimal delamination when the mask is lifted. The silicone can be removed by using a scalpel and 3D components are successively removed with a pair of tweezers. It should be noted that upon removal of the mask, the degree of delamination of the Parylene should be minimal. For that reason, as mentioned previously, the use of a primer prior to coating has been found crucial.

During the final phase of the Parylene coating process, the modules undergo electrical testing to assess their performance. These measurements encompass various tests, such as bump disconnection, threshold scans, analog scans, digital scans, and IV measurements, among others. These tests are conducted both before and after the coating process to identify any potential anomalies that may have occurred during coating. Figure 5.13 shows the results of an illustrative X-ray scan before (top four front-ends) and after (bottom four front-ends) Parylene coating. These results indicate that this specific module for this specific hybridization vendor, Hamamatsu Photonics K (HPk) [128], survived the Parylene coating stage and all intermediate steps involved (masking/demasking), and the module was not harmed. These measurements were performed by colleagues from the ATLAS ITk module testing group at CERN.

5.3.2.2 Monolithic Technologies

To explore the suitability of Parylene N as a protective coating for monolithic technologies, its effectiveness against humidity, mechanical damage, and other environmental factors is of great interest. The ideal scenario would involve using Parylene as the sole protective layer, covering the entire multi-chip MALTA module, including the wire-bonds. However, due to the considerable height of the wire-bonds in the MALTA multi-chip modules, a significantly thick Parylene coating layer is necessary. In communication with a vendor responsible for Parylene coating, SCS [123], deposition of Parylene N with a thickness beyond 40 µm is impractical and cost-prohibitive, as it would involve an extensive deposition process lasting approximately one month. Moreover, the mechanical properties of Parylene N undergo a transition towards brittleness when the coating exceeds a thickness of 100 µm [50]. For these reasons, after the module has been coated with Parylene, the wire-bonds can be subsequently encapsulated with Sylgard 186 (as explained in the previous section). This approach allows for an effective protective measure while maintaining the desired mechanical

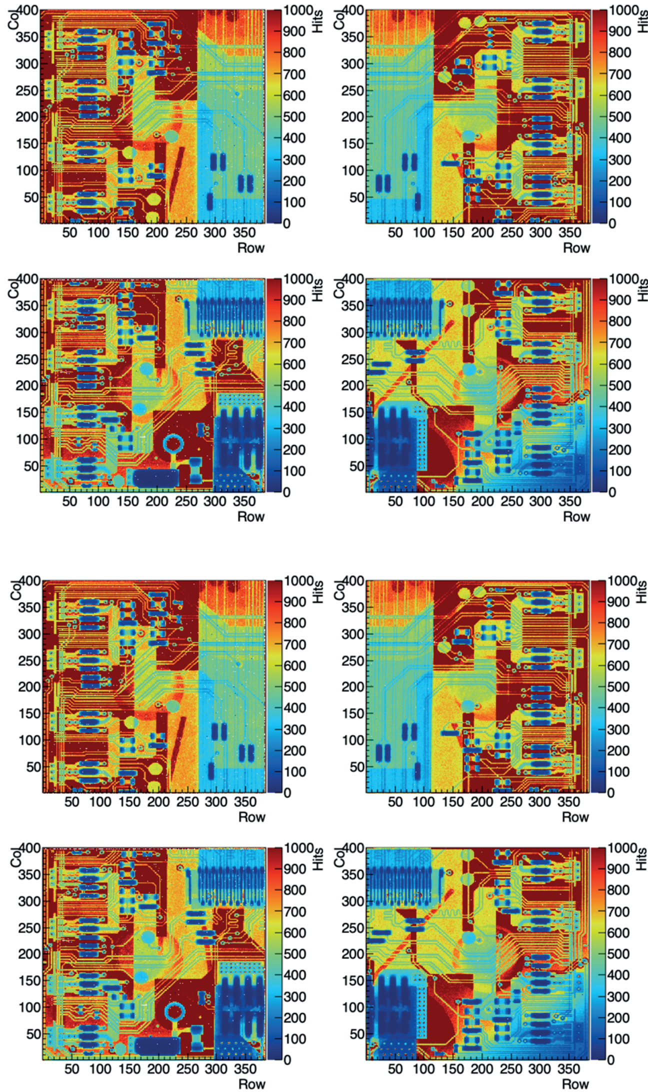


Figure 5.13: X-ray scans of four front ends of an ITk pixel module (ID: CERNPixQ28–HPk-HPk150 quad Module) before (top four) and after (bottom four) Parylene coating. No difference in the results of the scan can be observed after coating, indicating that the masking, Parylene coating, demasking, and other related tests have been successful and did not harm the module. These measurements were performed by colleagues from the ATLAS ITk module testing group at CERN.

and cost considerations. Preparatory work has been initiated to investigate the potential of Parylene N as a protective measure for modules exposed to environments characterized by high humidity levels or subjected to accelerated aging testing.

In this section, we delved into the crucial studies on coating and embedding techniques for pixel detector modules. These investigations explored methods to enhance the performance and durability of these modules, ensuring their reliability in challenging environments. Now, in this final section, we shift our focus towards the practical implementation and characterization of a monolithic pixel module. By subjecting the module to testing and analysis, we aim to gain valuable insights into its operational capabilities, its potential for efficient data transfer, and overall performance.

5.4 Characterization of the MALTA multi-chip module

In a multi-chip module, a crucial aspect is the seamless transmission of data between chips, ensuring minimal data loss or distortion. Additionally, the module should allow for simultaneous readout of all the chips integrated within it. To evaluate these key features effectively, a test beam set-up is commonly employed. This configuration involves utilizing a high-rate particle beam to assess the module's performance in terms of data transfer and chip readout.

5.4.1 Test Beam Set-Up

Two test beam campaigns were conducted at CERN's SPS in 2021 and 2022 to test the MALTA quad-chip board. The MALTA telescope, described briefly in Chapter 4 and extensively in the publication presented in Chapter II, was utilized with the quad-chip board as the Device Under Test (DUT), while six MALTA tracking planes served as a reference. To enable sequential illumination of all four chips on the module, the DUT was mounted on a moving linear stage and subjected to a 180 GeV pion beam. The MALTA quad-chip board underwent characterization under two different settings. In the first setting, all chips on the module except one were disabled (masked) to facilitate single chip characterization. The linear stage was positioned such to ensure that the beam only illuminated the enabled chip. In the second setting, all four chips were enabled, and the linear stage was adjusted to move the beam across all four chips. Special attention was given to the gap (dead area) between neighboring chips when the beam illuminated that region. The following section will present the results obtained from the single chip characterization in the first setting, as the author of this thesis was strongly involved in the test beam characterization and data analysis.

5.4.2 Single Chip Characterization

The MALTA quad-chip module that was tested in the test beam set-up consists of four MALTA samples on an epitaxial layer with thicknesses of 100 μm for the secondary chips and 300 μm for the primary chip. The operating threshold for all four chips corresponds to approximately 400 electrons.

The timing performance of each individual chip in the module is shown in Figure 5.14. It shows the time of arrival of the leading hit in the cluster with respect to a scintillator reference for the first, second, third, and fourth chip of the MALTA quad module. The timing distributions include a scintillator (~ 0.5 ns) and a sampling ($3.125/\sqrt{12}$=0.9 ns) jitter. The quoted sigma and peak value are extracted by fitting a Gaussian to the core of the distribution. There is a consistent 8 ns delay for each gap between the chips (limited by the wire-bond length) that is attributed to the CMOS transceivers. Figure 5.14 demonstrate that chip-to-chip data transfer is successful without distortion of the timing information.

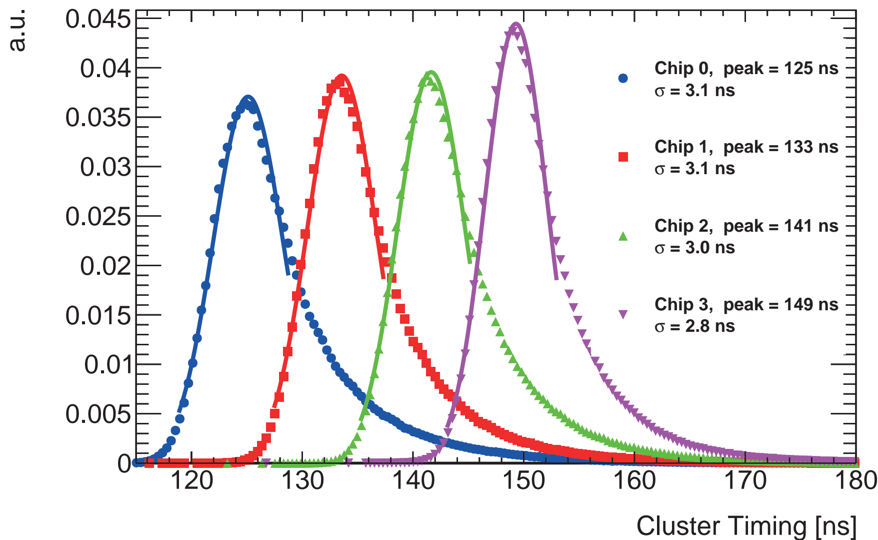


Figure 5.14: Time of arrival of the leading hit in the cluster with respect to a scintillator reference for the first, second, third and fourth chip of a MALTA quad module. The timing distributions include a scintillator ($\sim$0.5 ns) and a sampling ($3.125/\sqrt{12}$=0.9 ns) jitter and were obtained by illuminating the four chips separately. For every chip, the peak value of the timing distribution is quoted and the sigma is extracted by fitting a Gaussian to the core of the distribution.

Figure 5.15 depicts the efficiency response of all four chips on the MALTA quad-chip board. Simarily for MALTA (Chapter 4, the hit detection efficiency is calculated as the number of matched clusters on the Device Under Test (DUT) over the total number of reconstructed tracks. A matched cluster is found by associating hit clusters on the DUT to a track, which should be found within 100 µm. Due to the very large statistics, a small statistical error is recorded for the hit efficiency. It is shown that all chips exhibit an efficiency response surpassing 97%, meeting the quality standards established by ATLAS ITk (see Table 2.1 in Chapter 2). Notably, the secondary chips (top row and bottom left) demonstrate a remarkably similar efficiency response, with any minor fluctuations attributed to slight variations in operating thresholds among the chips. Comparatively, the primary chip demonstrates an approximately 0.5% lower efficiency response in relation to the secondary chips, which can be attributed to the different thickness of the primary chip. These findings align with the results presented for individual MALTA chips in Chapter 4, providing consistent outcomes across different chip configurations.

The proposed modularization of MALTA is specifically tailored to the current functionalities of the chip, including the CMOS transceiver blocks and the readout architecture. However, as developments of MALTA are ongoing, alternative modularization techniques could be further explored, such as the stitching technique [129]. This technique offers the capability to manufacture devices significantly larger than the dimensions of the design reticle. This is achieved by custom designing geometries into reticle sub-frames, enabling the creation of large chips with diagonals approaching the wafer diameter. This approach is, for instance, explored for the ITS3 monolithic pixel sensor of the ALICE collaboration [130]. The current periphery of MALTA, with its extensive 512 columns readout through 40 LVDS drivers, is not suitable for the stitching technique. However, in a prospective MALTA design featuring a serialized output [131], this approach could be considered as a potential modularization strategy.

5.5 Conclusion

In conclusion, this chapter has provided a comprehensive exploration of various aspects related to silicon pixel detector modules. We have examined different interconnection technologies, including wire-bonds, ACF, and nanowires, highlighting their potential for module development. Specifically, our focus has been on the modularization of the MALTA sensor, where wire-bonds were utilized as the interconnection method. Through coating and embedding studies, with respectively Parylene N and Sylgard 186, we have gained valuable insights into enhancing the performance and reliability of both hybrid and monolithic pixel detector technologies. Furthermore, the characterization of the MALTA multi-chip module has been a central focus, where rigorous testing and analysis

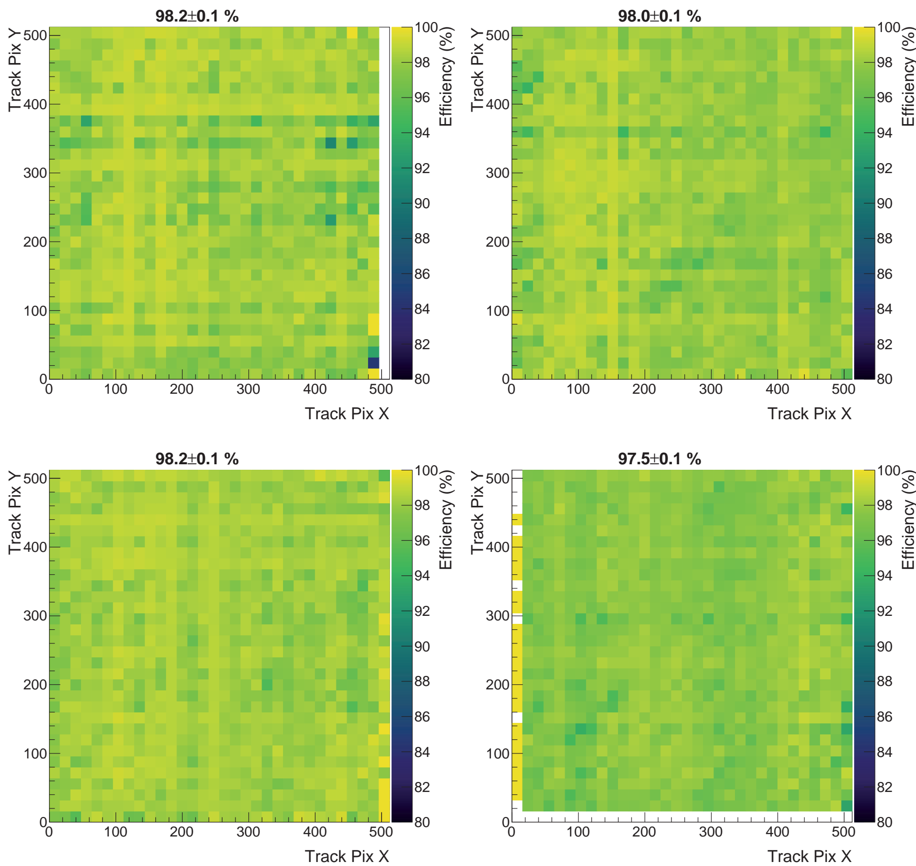


Figure 5.15: 2D Efficiency plot of the full matrix of four chips of the MALTA quad-chip board, represented by the secondary chips 1 (bottom left), 2 (top right), and 3 (top left) and the primary chip (bottom right).

have allowed us to evaluate its operational capacity and efficiency. These modules has demonstrated successful chip-to-chip communication and demonstrated that there is no appreciable data loss of the signal across the CMOS drivers. These findings contribute to our understanding of the module's capabilities and pave the way for future advancements in pixel detector technology.

Overall, the topics covered in this chapter demonstrate the intricate nature of silicon pixel detector modules and their potential impact in future collider experiments. By exploring versatile interconnection methods, modularization techniques, and coating and embedding studies for both hybrid and monolithic technologies, a solid foundation for future research, design improvements, and applications in particle physics experiments, medical imaging, and beyond is

made.

Chapter 6

Conclusion and Outlook

Within the scope of the HL-LHC project, the ATLAS experiment is undergoing a significant upgrade with the development of the Inner Tracker (ITk). The ITk, based on an all-silicon tracker, will feature improved performance and capabilities to handle the challenging conditions of the HL-LHC. MALTA, a DMAPS technology featuring a small collection electrode fabricated in Tower Semiconductor 180 nm CMOS imaging sensor technology, was developed with the aim of exploring its suitability for the Phase-II upgrade of ATLAS in the High Luminosity LHC and other high-energy physics experiments. In order to fulfill the stringent requirements on radiation hardness, high granularity, fast response time, and superior radiation tolerance, the MALTA detector underwent multiple design and process modifications.

This work has centered around the latest prototype of the MALTA family, a Depleted Monolithic Active Pixel Sensors (DMAPS) in 180 nm Tower Semiconductor CMOS imaging technology, MALTA2. By implementing a cascode front-end and by increasing the size of specific transistors, as detailed by Piro et al. in Ref. [16], the MALTA2 sensor has shown a notable reduction in RTS (Random Telegraph Signal) noise. The results presented in this work have shown results on its improved performance compared to its predecessor, MALTA, in terms of decreased RTS noise while maintaining a similar operating threshold (340 e^-). This improvement is particularly evident in the tails of the noise distribution, where the standard deviation has been observed to decrease by approximately 40%. As a result of the reduced front-end RTS noise, the MALTA2 sensor can now be effectively operated at a threshold as low as 150 e^- for non-irradiated samples. My research has demonstrated that MALTA2 sensors, produced on Czochralski substrates as suggested by Pernegger et al. in Ref. [11], and subjected to a backside metallisation post-processing procedure, exhibit the capability to function effectively even at high fluence levels ($>3\times10^{15}$), at an operating threshold as low as 90 e^-. It has been demonstrated in this work that non-irradiated MALTA2 samples on Czochralski substrates can achieve efficiencies of 99% and an average cluster size of 2 pixels at low threshold settings (150 e^-). In these conditions, a timing resolution of σ_t=1.7 ns can be obtained, where more than 98% of the hits are collected within 25 ns. Also it has been shown that superior performance at the highest irradiation dose (3×10^{15} 1 MeV n_{eq}/cm^2) was found on samples with very high doping of the n^- layer, initially discussed for the mini-MALTA demonstrator by Dyndal et al. in Ref. [88]. At an operating threshold of 110 e^-, these samples achieved a maximum efficiency of 98% and an average cluster size of 1.7 pixels. These conditions also yielded an RMS of the time difference distribution equal to 6.3 ns, with 95% of the clusters

being collected within a 25 ns time frame. The collaborative efforts within the MALTA working group have resulted in the publication of numerous proceedings and papers in peer-reviewed journals, including Ref.[10–16], of which the author of this thesis is a co-author.

Collaborating with a small team of physicists and engineers, we successfully constructed a custom 6-plane telescope utilizing the MALTA chip. In 2022, I undertook the position of the test beam coordinator, bearing the principal duty of supervising all facets pertaining to the MALTA telescope and its affiliated operations. In this work the architecture and performance of the MALTA telescope has been reviewed and it has been shown that the MALTA telescope provides a spatial resolution of σ_s=4.1±0.2 µm, based on the linear regression approach, and a track timing resolution of σ_t=2.1 ns.

This research has presented the testing outcomes of the first generation multi-chip module assembly featuring the MALTA chip. The results have demonstrated successful chip-to-chip data transmission through CMOS drives without appreciable distortion of the signal. The results have formed a stepping stone towards the realisation of a large area, lightweight monolithic module. Within the context of module development, the exploration of versatile coating and embedding layers has been a central focus to ensure durability, dependability, and resilience. This study has successfully demonstrated the viability of Sylgard 186 as an encapsulant that safeguards wirebonds against potential mechanical damage, catering to both monolithic and hybrid pixel detectors, and does not affect the discriminator's pixel threshold. Additionally, the investigation reveals Parylene N as an effective coating layer for safeguarding hybrid modules against high voltage. I was assigned the responsibility of driving the development and validation of the Parylene coating process for the ITk pixel modules on behalf of the CERN team. The efforts in this realm positioned CERN as one of the pioneering institutions within the project, achieving qualification for this critical step in module assembly. Moreover, preliminary groundwork has been conducted to assess the application of Parylene N as a protective barrier against humidity and accelerated aging for both flip-chip bonded monolithic and hybrid detectors.

Returning to the introductory paragraph of this chapter, it was stated that MALTA was developed with the intention to be integrated into the Phase-II upgrade of ATLAS in the HL-LHC and other high-energy physics experiments. It has since become evident that the ATLAS Inner Tracker (ITk) has opted for the hybrid ITk Pixel module, and MALTA will not be integrated into the HL-LHC of ATLAS. The decision to adopt the hybrid technology was made due to the limited readiness of MALTA at that time. Nonetheless, MALTA still holds relevance in the context of other high-energy physics experiments that require specifications such as radiation hardness, small pixel size, and high timing efficiency. Two of these experiments are briefly highlighted below.

MALTA is being considered as a potential candidate for the LHCb Upgrade

II [9], which involves the upgrade of the Upstream Tracker (UT) with MAPS technology. The UT project demands a MAPS solution that can meet specific requirements, including radiation hardness (NIEL) within the range of 0.5-3×10^{15} 1 MeV n_{eq}/cm^2, a timing resolution of <3 ns, pixel dimensions of $<50\times150$ µm, and high-rate capability [9, 87]. As demonstrated in the work presented in this thesis, MALTA exhibits the capacity to fulfill all of these prerequisites. Furthermore, the interest in MALTA for the LHCb UT stems from its unique capability to output all raw data without the need for a trigger. This feature opens up the possibility of constructing a secondary trigger system based on cluster information received from the tracker, effectively establishing a track-based trigger system [9]. This aspect adds to the appeal of MALTA as a candidate for the LHCb Upgrade II.

The high-luminosity high-energy Electron-Ion Collider (EIC) [132], to be constructed at Brookhaven National Laboratory (BNL) [133], will study several fundamental questions in the high energy and nuclear physics fields. Therefore, the EIC project necessitates a silicon vertex and tracking detector characterized by high granularity and a low material budget. Such a detector is vital for delivering precise measurements of primary and displaced vertices, along with accurate track reconstruction that boasts exceptional momentum and spatial resolutions [132, 134]. To meet these demanding requirements for the vertex and tracking detectors, the EIC project has explored various technologies, including the utilization of MAPS. These requirements encompass specifications such as achieving a spatial resolution of <5 µm and maintaining a material budget below 0.8% of a radiation length in the barrel (endcap). Of particular note is the special interest expressed in Reference [134] regarding MALTA2 for its excellent timing resolution (<2 ns), as well as the threshold over noise ratio (>10) even after irradiation at a rate of 3×10^{15} 1 MeV n_{eq}/cm^2. These features make MALTA2 a compelling candidate for deployment in the EIC project.

While the work in this thesis allowed to demonstrate the radiation hardness of the MALTA2 sensor above 3×10^{15} 1 MeV n_{eq}/cm^2, further work will try to push its performance at higher fluences. Especially the use of MALTA2 (or future generation) with the very high doping of the n-implant and backside metallisation for fluences $>10^{16}$ 1 MeV n_{eq}/cm^2 will be of large interest. Furthermore, the foreseen MALTA3 demonstrator [131] has the intention of bridging the gap between the asynchronous pixel matrix and the synchronous DAQ. Its readout architecture will serve as a baseline for the full-scale prototype, MALTA3, with focus on timing performance. The synchronization memory will be upgraded to allow clock speeds of up to 1.28 GHz, with the goal of achieving a sub-nanosecond on-chip timing resolution.

Chapter 7

Introduction to Papers

The following pages present a collection of papers showcasing the results achieved during the various stages of the MALTA development. These papers are arranged in chronological order of writing. Paper I and II have been published in peer-reviewed journals, namely the Journal of Instrumentation and the European Physics Journal C, respectively. Paper III is available on arXiv and has been submitted to the European Physics Journal C.

The first and third paper primarily focus on test beam results obtained using the MALTA and MALTA2 sensors for radiation hardness studies. The first paper specifically compares samples on high-resistivity epitaxial silicon with Czochralski substrates before and after neutron irradiation. It also presents the first results obtained with MALTA2, highlighting the improvements in the front-end. The third paper is entirely dedicated to demonstrating the radiation hardness of MALTA2 on Czochralski substrates. It emphasizes the performance in terms of timing resolution and efficiency before and after neutron irradiation, with a special focus on process modifications that were implemented to achieve radiation hardness greater than $3{\times}10^{15}$ 1 MeV n_{eq}/cm^2. The second paper provides a comprehensive overview of the architecture and performance of the MALTA telescope, that was utilized in the test beam campaigns conducted between 2021 and 2023 at the SPS North Area. This paper highlights how the MALTA telescope leverages the best qualities of the MALTA sensor: a full prototype (large area), high granularity, self-triggering capability, and excellent spatial and timing resolution.

In addition to being the first author of the aforementioned publications, the author of this thesis has played an active role in obtaining the corresponding results. Apart from being involved in lab-scale characterization of samples and participating in test beam installation, data-taking shifts, and data-analysis, the author served as the test beam coordinator during the 2022 test beam period. The collaborative efforts within the MALTA working group have resulted in the publication of numerous proceedings and papers in peer-reviewed journals, including Ref.[10–17], of which the author of this thesis is a co-author.

Papers

Paper I

Radiation hardness and timing performance in MALTA monolithic pixel sensors in TowerJazz 180 nm

M. van Rijnbach et al.

Published in *Journal of Instrumentation*, April 2022, Volume 17, Number 4, Pages C04034 Publisher IOP Publishing DOI: 10.1000/1748-0221/17/04/C0403461.

I

PUBLISHED BY IOP PUBLISHING FOR SISSA MEDIALAB

RECEIVED: *October 25, 2021*
REVISED: *January 19, 2022*
ACCEPTED: *February 25, 2022*
PUBLISHED: *April 29, 2022*

TOPICAL WORKSHOP ON ELECTRONICS FOR PARTICLE PHYSICS 2021
20–24 SEPTEMBER, 2021
ONLINE

Radiation hardness and timing performance in MALTA monolithic pixel sensors in TowerJazz 180 nm

M. van Rijnbach,[a,b,*] P. Allport,[c] I. Asensi,[d] I. Berdalovic,[e] D. Bortoletto,[f] C. Buttar,[g] R. Cardella,[h] F. Dachs,[a] V. Dao,[a] H. Denizli,[i] D. Dobrijevic,[h] M. Dyndal,[a] L. Flores,[f] P. Freeman,[c] A. Gabrielli,[a] L. Gonella,[c] M. LeBlanc,[a] K. Oyulmaz,[i] H. Pernegger,[a] F. Piro,[a,j] P. Riedler,[a] H. Sandaker,[b] C. Solans,[a] W. Snoeys,[a] T. Suligoj,[e] J. Torres[d] and S. Worm[k]

[a] *CERN Experimental Physics Department, Meyrin, Switzerland*
[b] *Department of Physics, University of Oslo, Oslo, Norway*
[c] *Department of Physics and Astronomy, University of Birmingham, Birmingham, United Kingdom*
[d] *Department of Electronics Engineering, University of Valencia, Valencia, Spain*
[e] *Faculty of Electrical Engineering and Computing, University of Zagreb, Zagreb, Croatia*
[f] *Department of Particle Physics, University of Oxford, Oxford, United Kingdom*
[g] *Faculty of Physics and Astronomy, University of Glasgow, Glasgow, United Kingdom*
[h] *Department of Nuclear and Particle Physics, University of Geneva, Geneva, Switzerland*
[i] *Department of Physics, Bolu Abant Izzet Baysal University, Bolu, Turkey*
[j] *Advanced Quantum Architecture Laboratory (AQUA), Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland*
[k] *Deutsches Elektronen-Synchrotron (DESY), Zeuthen, Germany*

E-mail: milou.van.rijnbach@cern.ch

ABSTRACT: The MALTA family of depleted monolithic pixel sensors produced in TowerJazz 180 nm CMOS technology target radiation hard applications for the HL-LHC and beyond. Several process modifications and front-end improvements have resulted in radiation hardness $>10^{15}$ 1 MeV n_{eq}/cm^2 and time resolution below 2 ns, with uniform charge collection and efficiency across the pixel of size $36.4 \times 36.4\ \mu m^2$ with small collection electrode. This contribution will

*Corresponding author.

© 2022 The Author(s). Published by IOP Publishing Ltd on behalf of Sissa Medialab. Original content from this work may be used under the terms of the Creative Commons Attribution 4.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

https://doi.org/10.1088/1748-0221/17/04/C04034

2022 JINST 17 C04034

present the comparison of samples produced on high-resistivity epitaxial silicon with Czochralski substrates, before and after neutron irradiation, and results from MALTA2 with a new cascoded front-end flavour that further reduces the RTS noise.

Keywords: Radiation-hard detectors; Materials for solid-state detectors

2022 JINST 17 C04034

Contents

1 Introduction **1**

2 MALTA-Czochralski **2**

2.1 Non-irradiated samples 2

2.2 Irradiated samples 3

2.3 Timing performance 3

3 MALTA2 **4**

4 Conclusion **5**

1 Introduction

Future collider experiments demand their detector technologies to withstand the harsh environments they will operate in, such as large radiation dose, high hit rate with high granularity, and fast response time [1]. One of the most interesting detector technologies that can address these requirements is the monolithic pixel sensor with small collection electrode due to its large signal-over-noise ratio. Compared to the more largely used hybrid pixel sensors with bump-bonding technology, monolithic technologies offer advantages such as reduced production effort, reduced costs, and material usage due to the fact that readout electronics and sensor are integrated in the same silicon wafer [2, 3].

MALTA is a Depleted Monolithic Active Pixel Sensor (DMAPS) with small collection electrode, fabricated in TowerJazz 180 nm CMOS imaging technology. The MALTA matrix consists of 512 × 512 pixels with a pixel pitch of 36.4 µm and an octagonal shaped collection electrode with a diameter of 2 µm. The asynchronous readout avoids the distribution of high frequency clock signals across the matrix to reduce power consumption (10 mW/cm^2 digital and 70 mW/cm^2 analog power) and minimises analog-digital crosstalk [4]. The pixel design and sensor processing of MALTA have been developed such that it can withstand a large radiation dose while maintaining the advantages of using a small collection electrode with very low capacitance, i.e. minimising noise and low power dissipation. In order to achieve high detection efficiency after irradiation, the pixel design is optimised such that it can achieve excellent charge collection and electrical field configuration. This is done by using three different implant designs: the standard modified (STD) introduces a low-dose n-type blanket implant across the full pixel matrix, a design with a gap in this n-blanket implant (N-GAP), and a design with an additional deep p-well implant (XDPW) [5]. These pixel flavours are produced on high-resistivity epitaxial substrates (30 µm depletion layer) and on thick high-resistivity p-type Czochralski (Cz) substrates. Introducing the same sensor design on high-resistivity (3–4 kOhm) Cz substrates enables the combination of the advantages

2022 JINST 17 C04034

of small electrode CMOS sensors with those of a thick (100 µm) detection layer. Here, the low capacitance is maintained while the signal amplitude is increased due to the thicker depleted sensor layer.

2 MALTA-Czochralski

In the following sections results will be presented on MALTA sensors produced on epitaixal and Cz substrates before and after neutron irradiation. The data presented were recorded at the DESY test beam facility, using a 3–4 GeV electron beam with a telescope consisting of three MALTA epitaxial tracking planes. The non-irradiated devices under test (DUT) were operated at room temperature, whereas the irradiated DUTs were operated in a cold box at −14 °C.

2.1 Non-irradiated samples

Figure 1 illustrates the difference in efficiency response as a function of substrate bias for the non-irradiated epitaxial and Cz sensors. The hit detection efficiency is defined as the fraction of clusters on the DUT matched to telescope tracks over the total number of tracks. The DUT hit is matched to a track if the distance between the track interpolation position and the centre position of the cluster is smaller than 100 µm.

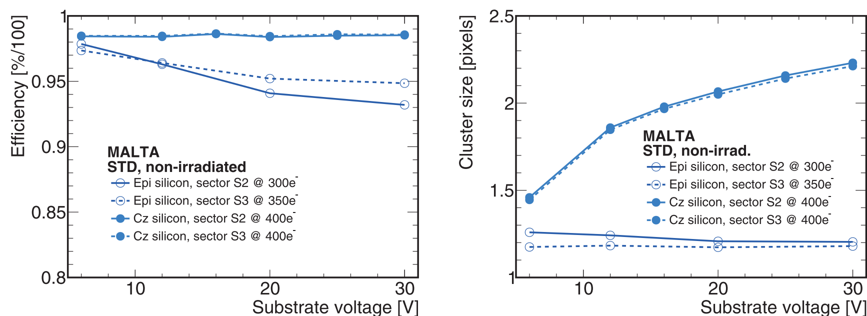


Figure 1. Left: efficiency of non-irradiated STD MALTA samples on epitaxial and Cz substrate versus substrate bias. Right: average cluster size of non-irradiated STD MALTA samples on epitaxial and Cz substrate versus substrate bias. Indicated are the respective discriminator thresholds (electrons) and the corresponding sector of the pixel matrix (S2 or S3). The sectors differ in the extension of the deep p-well, respectively for S2 and S3 medium and maximum [4].

Whereas the epitaxial sensors show decreasing efficiency at increasing bias voltage due to the increasing leakage current, the Cz samples show a flat dependency versus substrate voltage. Figure 1 also illustrates the relationship between cluster size and bias voltage. The cluster size, the number of adjacent pixels firing after a charged particle transverses the detector, for epitaxial sensors does not increase with substrate voltage. However, for a STD Cz sample the cluster size increases up to 2.2 pixels at 30 V due to the larger depletion depth in the thick substrate and therefore more charge is collected.

2022 JINST 17 C04034

2.2 Irradiated samples

Figure 2 shows the efficiency response of N-GAP epitaxial and Cz sensors as a function of substrate bias after neutron irradiation to 1×10^{15} 1 MeV n_{eq}/cm^2 and 2×10^{15} 1 MeV n_{eq}/cm^2. This figure illustrates the motivation to move to Czochralski sensors after irradiation, since the efficiency of Cz samples increases substantially with substrate voltage. As the bias voltage increases, the depleted region in the high resistivity substrate increases and with it its signal. After neutron irradiation the epitaxial sensors achieve a maximum efficiency at −12V, whereafter the efficiency declines.

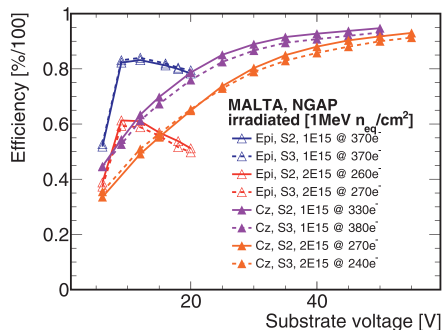


Figure 2. Efficiency of MALTA samples irradiated to 10^{15} 1 MeV n_{eq}/cm^2 and 2×10^{15} 1 MeV n_{eq}/cm^2 on NGAP epitaxial and Cz silicon versus substrate bias. Indicated are the respective discriminator thresholds (electrons) and the corresponding sector of the pixel matrix (S2 or S3). The sectors differ in the extension of the deep p-well, respectively for S2 and S3 medium and maximum [4].

2.3 Timing performance

With the bunch-crossing clock of the LHC being 25 ns, sensors and readout architecture have to be sufficiently fast in order to match the hits with the corresponding bunch-crossing. Furthermore, sub-nanosecond level resolution is required to contribute to jet sub-structure reconstruction [1]. Figure 3 illustrates that non-irradiated MALTA Cz STD is capable of fulfilling these requirements. Timing measurements were performed with a trigger scintillator for timing reference and a dedicated CERN developed ASIC, the Pico-TDC [6], to measure the difference in time between the fastest MALTA signal and the scintillator. The timing distribution is measured as a function of substrate voltage and the 50%-, 68%-, and 95%-integral of the time difference identify the core of the distribution and its tails. These results show that at higher substrate voltage, the signal is faster and has higher amplitude. From the integral of the timing distribution it can be concluded that there will be no pile-up of bunch-crossings, since all data can be read-out within 25 ns. Furthermore, figure 3 indicates that at a bias voltage >15 V 50% of the hits have arrived and are tagged within 2 ns for the full-size sensor. The ability to read out within 2 ns and preserve the timing information makes this technology interesting for physics analysis.

2022 JINST 17 C04034

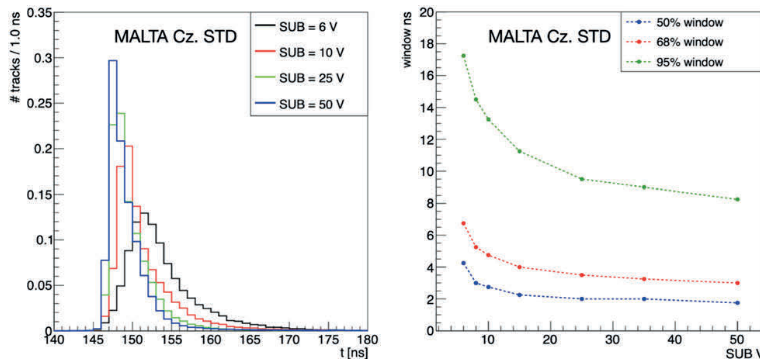


Figure 3. Left: difference in time of the fastest hit of the cluster (matched with the track in the DUT) and the time of the hit in the scintillator versus substrate bias. Right: integral of the difference in time distribution (left) versus substrate bias. Timing measurements were done for a non-irradiated MALTA Cz. STD sample and with low energy electrons from Sr-90 β-decay.

3 MALTA2

MALTA2 is the second prototype of the family of DMAPS in TowerJazz 180 nm technology. The matrix consists of 224 × 512 pixels, with a pixel size of 36.4 × 36.4 μm^2, and an active area of 18.33 mm^2. MALTA2 is fabricated on both epitaxial and Cz substrates and consists of the same pixel flavours as MALTA (STD, N-GAP and XDPW). MALTA2 has two front-end designs implemented: the standard MALTA and a cascode design as implemented for the MiniMALTA demonstrator [7]. Furthermore, the size of selected transistors are increased to reduce the Random Telegraph Signal (RTS) noise. This benefit was also shown in the MiniMALTA demonstrator, where the larger transistor size significantly decreased the RTS noise, both before and after irradiation [7]. Similar to the MALTA design, the readout of the MALTA2 chip is fully asynchronous. The benefit of the new front-end design for MALTA2 is illustrated by figure 4, which presents a threshold and a noise scan for a non-irradiated Epi N-GAP MALTA and MALTA2. Both designs show similar threshold dispersion, approximately 10% of the mean. However, the MALTA2 sensor shows significantly less RTS, indicated by the smaller tails in the noise scan, compared to the MALTA sensor at the same threshold (~350 electrons) and same bias voltage (−6 V). The reduction of the RTS from the front-end opens the possibility to operate at lower thresholds and therefore reach higher efficiencies.

2022 JINST 17 C04034

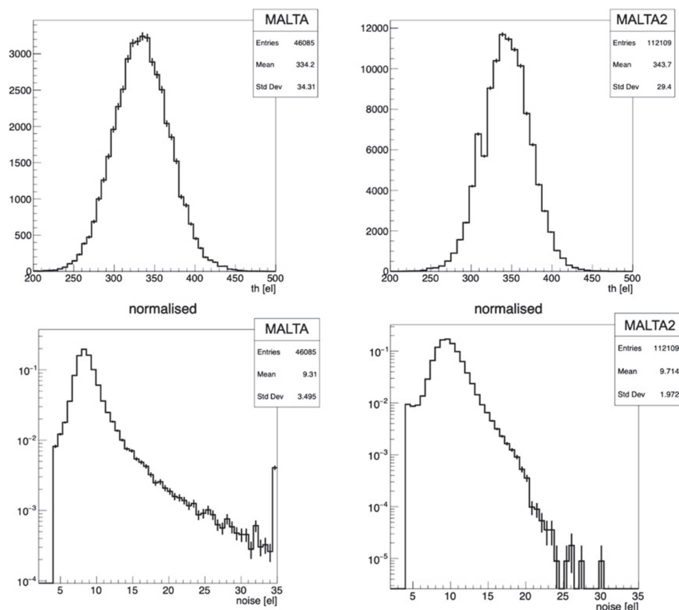


Figure 4. Threshold scan (top) and noise scan on logarithmic scale (bottom) of MALTA (left) and MALTA2 (right) (Epi, NGAP, 300 μm thick, high doping of n-blanket) for respectively 46085 and 112109 pixels, at −6 V SUB bias and −6 V PWELL bias.

Test beam at the Super Proton Synchrotron (SPS) at CERN is currently ongoing with the goal to demonstrate MALTA2 performance in terms of radiation hardness (>10^{15} 1 MeV n_{eq}/cm^2) and timing performance. The custom telescope dedicated to demonstrate the performance of MALTA2 contains 6 MALTA tracking planes and a cold box which can host up to 2 DUTs. Preliminary results on new generation irradiated MALTA2 samples demonstrate uniform efficiency response over a wide range of threshold.

4 Conclusion

The combination of the pixel design and sensor processing of MALTA Cz. has provided excellent results for a full-size monolithic CMOS sensor with small collection electrode. MALTA2 should confirm radiation hardness >10^{15} 1 MeV n_{eq}/cm^2 and excellent timing performance.

Acknowledgments

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement Number 101004761 (AIDAinnova), and Number 654168 (IJS, Ljubljana, Slovenia). Furthermore it has been supported by the Marie Sklodowska-Curie

2022 JINST 17 C04034

Innovative Training Network of the European Commission Horizon 2020 Programme under Contract Number 675587 (STREAM). The measurements leading to these results have been performed at the TestBeam Facility at DESY Hamburg (Germany), a member of the Helmholtz Association (HGF), and at the E3 beam-line at the electron accelerator ELSA operated by the university of Bonn in Nordrhein-Westfalen, Germany.

References

[1] CERN Experimental Physics Department, *Strategic R&D Programme on Technologies for Future Experiments*, CERN-OPEN-2018-006 (2018).

[2] H. Pernegger et al., *First tests of a novel radiation hard CMOS sensor process for depleted monolithic active pixel sensors*, 2017 *JINST* **12** P06008.

[3] W. Snoeys et al., *A process modification for CMOS monolithic active pixel sensors for enhanced depletion, timing performance and radiation tolerance*, *Nucl. Instrum. Meth. A* **871** (2017) 90.

[4] R. Cardella et al., *MALTA: an asynchronous readout CMOS monolithic pixel detector for the ATLAS High-Luminosity upgrade*, 2019 *JINST* **14** C06019.

[5] M. Munker et al., *Simulations of CMOS pixel sensors with a small collection electrode, improved for a faster charge collection and increased radiation tolerance*, 2019 *JINST* **14** C05013 [arXiv:1903.10190].

[6] L. Perktold and J. Christiansen, *A multichannel time-to-digital converter ASIC with better than 3 ps RMS time resolution*, 2014 *JINST* **9** C01060.

[7] M. Dyndal et al., *Mini-MALTA: radiation hard pixel designs for small-electrode monolithic CMOS sensors for the High Luminosity LHC*, 2020 *JINST* **15** P02005.

2022 JINST 17 C04034

Paper II

Performance of the MALTA Telescope

M. van Rijnbach et al.

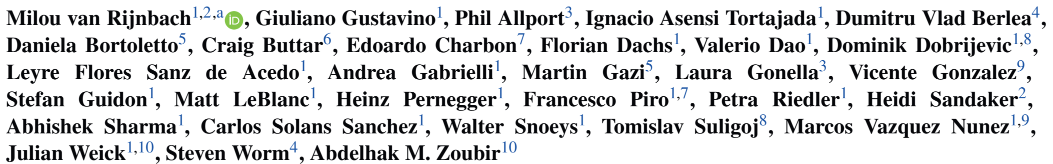
Published in *The European Physical Journal C*, July 2023, Volume 83, Number 7, Pages 581 Publisher Springer DOI: 10.1140/epjc/s10052-023-11760-z.

Eur. Phys. J. C (2023) 83:581
https://doi.org/10.1140/epjc/s10052-023-11760-z

THE EUROPEAN PHYSICAL JOURNAL C

Regular Article - Experimental Physics

Performance of the MALTA telescope


Milou van Rijnbach[1,2,a], **Giuliano Gustavino**[1], **Phil Allport**[3], **Ignacio Asensi Tortajada**[1], **Dumitru Vlad Berlea**[4], **Daniela Bortoletto**[5], **Craig Buttar**[6], **Edoardo Charbon**[7], **Florian Dachs**[1], **Valerio Dao**[1], **Dominik Dobrijevic**[1,8], **Leyre Flores Sanz de Acedo**[1], **Andrea Gabrielli**[1], **Martin Gazi**[5], **Laura Gonella**[3], **Vicente Gonzalez**[9], **Stefan Guidon**[1], **Matt LeBlanc**[1], **Heinz Pernegger**[1], **Francesco Piro**[1,7], **Petra Riedler**[1], **Heidi Sandaker**[2], **Abhishek Sharma**[1], **Carlos Solans Sanchez**[1], **Walter Snoeys**[1], **Tomislav Suligoj**[8], **Marcos Vazquez Nunez**[1,9], **Julian Weick**[1,10], **Steven Worm**[4], **Abdelhak M. Zoubir**[10]

[1] CERN, Geneva, Switzerland
[2] University of Oslo, Oslo, Norway
[3] University of Birmingham, Birmingham, UK
[4] DESY, Zeuthen, Germany
[5] University of Oxford, Oxford, UK
[6] University of Glasgow, Glasgow, UK
[7] EPFL, Lausanne, Switzerland
[8] University of Zagreb, Zagreb, Croatia
[9] Universitat de València, Valencia, Spain
[10] Technische Universität Darmstadt, Darmstadt, Germany

Received: 24 April 2023 / Accepted: 26 June 2023
© The Author(s) 2023

Abstract MALTA is part of the Depleted Monolithic Active Pixel sensors designed in Tower 180 nm CMOS imaging technology. A custom telescope with six MALTA planes has been developed for test beam campaigns at SPS, CERN, with the ability to host several devices under test. The telescope system has a dedicated custom readout, online monitoring integrated into DAQ with realtime hit map, time distribution and event hit multiplicity. It hosts a dedicated fully configurable trigger system enabling to trigger on coincidence between telescope planes and timing reference from a scintillator. The excellent time resolution performance allows for fast track reconstruction, due to the possibility to retain a low hit multiplicity per event which reduces the combinatorics. This paper reviews the architecture of the system and its performance during the 2021 and 2022 test beam campaign at the SPS North Area.

1 Introduction

Beam telescopes are tracking detector systems used for the characterization of pixel detector prototypes for a wide range of applications. They allow for studies that are beyond feasible in a laboratory, by analysing the sensor's response to ionising particles, mimicking the operation conditions in real tracking detectors. Particles passing through the telescope are reconstructed and used in the characterisation of a given Device Under Test (DUT).

MALTA is a Depleted Monolithic Active Pixel Sensor (DMPAS) designed in Tower 180 nm CMOS imaging technology. A custom telescope has been developed for test beam campaigns at the Super Proton Synchroton (SPS) at CERN using up to six MALTA tracking planes, a scintillator for precise timing reference, and the ability to host several DUTs. The telescope system has a dedicated readout, a fully configurable trigger system, an excellent time resolution, and a low hit multiplicity per event that reduces the combinatorics, which in turn allows for fast track reconstruction. Furthermore, there is the possibility for in-chip Region Of Interest (ROI) implementation, serving both large and small prototypes. The MALTA telescope is permanently installed at the H6 beamline (see Fig. 1) in the North Area (NA), one of the secondary beam areas of SPS. The high-energy and high-resolution H6 beam line can transport mixed hadron beams within the range of 10–205 GeV/c to the respective experiments, with a beam intensity ranging between 2×10^5 and 4.5×10^6 particles per spill.

This paper is dedicated to reviewing the architecture, operation, and performance of the MALTA telescope. First the MALTA sensor will be introduced, which will elaborate on

[a] e-mail: milou.van.rijnbach@cern.ch (corresponding author)

Published online: 08 July 2023

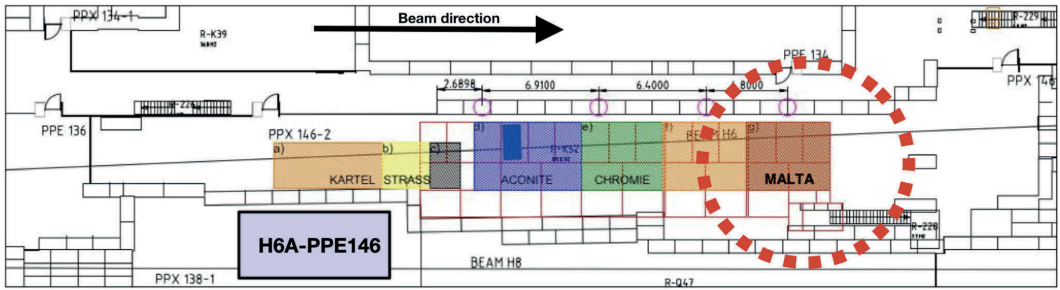


Fig. 1 A schematic image of the H6A/PPE146 beam line at the Super Proton Synchrotron (SPS) at CERN. The experimental areas, including the permanent installation of the MALTA telescope downstream of the beam line, are indicated

the motivation for a MALTA based telescope and specifications on the tracking planes. Second, the main components of the telescope will be addressed, including the mechanical structure, trigger logic, and the data acquisition. Furthermore, the reconstruction and offline analysis framework will be discussed including the results on the spatial telescope resolution, followed by discussion on the timing resolution of the MALTA telescope. Finally, examples of DUT integration will be provided, focusing on results obtained with the MALTA2 sensor during the 2021 and 2022 test beam campaign and integration of various ATLAS R&D prototypes within the custom Trigger Logic Unit (TLU).

2 The MALTA sensor

MALTA is a DMAPS with small collection electrode, fabricated in Tower 180 nm CMOS imaging technology. The sensor was originally developed for its use in the Phase-II upgrade of ATLAS for the High Luminosity LHC [1] and for other future collider experiments, the latter being still relevant for possible applications of MALTA. The large interest in monolithic pixel sensors has been driven by the possibility to minimize the material budget, reduce the production effort, and lowering the costs as the readout electronics and sensor are fabricated in the same silicon wafer. More specifically, the MALTA pixel sensor offers excellent radiation hardness up to about 100 Mrad Total Ionizing Dose (TID) and greater than 1×10^{15} 1 MeV n_{eq}/cm^2 in Non-Ionizing Energy Loss (NIEL) with a fast charge collection [2,3], as will be explained further below.

The MALTA matrix consists of 512×512 square pixels with a pixel pitch of 36.4 μm. The size of the small, octagonal-shaped collection electrode (diameter of 2 μm) results in a small capacitance, which consequently minimizes noise and allows for low power dissipation (10 mW/cm^2 digital and 70 mW/cm^2 analog power). The asynchronous readout transmits the hit information directly from chip to periphery and consequently through 37 parallel output signals (2 ns output signal length). The asynchronous readout avoids the distribution of a high frequency clock signals across the matrix, which minimizes analog-digital cross-talk. Pixels are organised in groups of 2×8 and hits from a pixel are sent to a reference pulse generator which is common within the group. A reference pulse is generated, which is appended to the pixel and group address, respectively 16-bit and 5-bit. The hits are distributed in two parallel 22-bit wide busses, one for even groups and the other for odd groups [4]. The benefit of this distinction lies in the fact that adjacent groups cannot share the same bus, reducing cross-talk on the hit address bus. Due to the area constraints of the chip, no Time-over-Threshold (ToT) information is directly available from the chip.

Several studies [4–6] have reported on the performance of irradiated DMAPS with small collection electrode, specifically on the sensors with the standard modified process modification (STD), where a low dose n-type blanket implant is introduced across the full pixel matrix. The studies showed that the detection efficiency for these type of sensors is compromised in the pixel corners. Therefore, the pixel design and sensor processing has been modified such that these deficiencies could be overcome while maintaining the advantages of a sensor with a small collection electrode. Through dedicated TCAD simulations [7] it was found that two different process modifications could improve the charge collection and increase the radiation tolerance. These include: a design with a gap in the n-blanket implant (NGAP) and a design with an additional deep p-well implant (XDPW). These pixel flavours are produced on high-resistivity epitaxial (Epi) substrates (approximately 30 μm sensing layer on a 70–270 μm substrate) and on thick p-type Czochralski (Cz) substrates. Introducing the same sensor design on high-resistivity (3–4 kOhm) Cz substrates enables the combination of the advantages of small electrode CMOS sensors with those of a thick (100–300 μm) detection layer. Here the low capacitance is

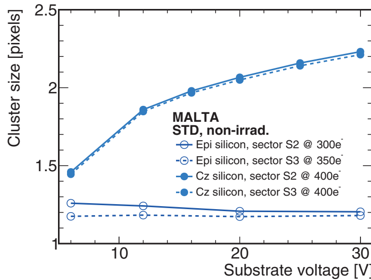


Fig. 2 Average cluster size of non-irradiated STD MALTA samples on epitaxial and Cz substrate versus reverse substrate bias. Indicated are the respective discriminator thresholds (electrons) and the corresponding sector of the pixel matrix (S2 or S3). The sectors differ in the extension of the deep p-well, with medium and maximum extension for sectors S2 and S3 respectively [4]

maintained while the signal amplitude is increased due to the thicker depleted sensor layer [8].

2.1 MALTA telescope planes

It has been shown in previous test beam campaigns [8] that the substrate type and process modification flavour (i.e. STD, NGAP or XDPW) have great influence on the cluster size, defined in Sect. 4. As can be observed in Fig. 2, the cluster size for Epi sensors does not increase with substrate voltage. However, for a STD Cz sample the cluster size increases up to 2.2 pixels at 30 V reverse substrate bias due to the larger depletion depth in the thick substrate and therefore more charge is collected. This insight has been extremely valuable in designing the MALTA telescope and choosing the appropriate tracking planes, especially in order to improve its spatial resolution.

Table 1 displays some of the key parameters of the MALTA planes that are implemented in the telescope. This includes the substrate type (Epi or Cz), the flavour (STD, NGAP, or XDPW), the thickness of the sensing layer, the bias voltage at which they are operated, and their relative position with respect to Plane 1. The position of the MALTA Cz STD planes (Plane 3 and 4) have been chosen such that they are positioned before and after the DUT(s) and are operated at their maximum bias voltage in order to guarantee the generation of large clusters.[1] The benefits of these operating conditions on the performance of the telescope will become more evident in Sect. 5.

[1] The criterium for the maximum bias voltage is chosen such that the leakage current is minimised.

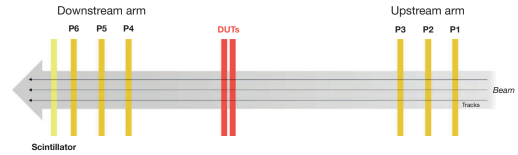


Fig. 3 A schematic sketch of the MALTA beam telescope where the six tracking planes (P1–P6), the devices under tests (DUTs), and the scintillator are indicated

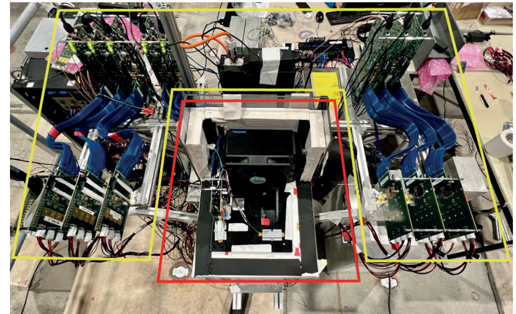


Fig. 4 Top view of the MALTA telescope. Equipment mounted on the main stage are inside yellow lines, cold box and DUTs positioned on DUT stage are inside the red lines

3 Components of the telescope

3.1 Mechanical architecture

A layout sketch of the MALTA beam telescope is shown in Fig. 3. It consists of two arms placed around one or more DUTs, where each arm houses three reference planes spaced 8 cm from each other. The two arms are placed 78 cm from each other to accommodate the coldbox (explained further below).

The DUTs and telescope arms are held by a metal structure that is supported by two moving stages each, a STANDA micro-position and an ISEL linear stage, which can be independently positioned in both perpendicular and horizontal direction with respect to the beam. The stage that moves the arms of the telescope planes has a range of 75 cm, whereas the range of the stage holding the DUTs is 15 cm (in horizontal and perpendicular direction). Furthermore, a rotational stage allows to variate the incident angle between the beam and the DUTs. The position of the planes and the DUTs is controlled by dedicated software packages and can therefore conveniently be operated remotely.

The telescope relies on a JULABO FP50-HE air-cooled chiller along with a Julabo H5 thermal fluid. This cooling system is capable of reaching temperatures as low as –50°C and it is employed to cool irradiated DUTs through convection. The DUTs can be installed inside a coldbox ($46 \times 29 \times 39$ cm)

Table 1 Overview of the main specifications of the MALTA telescope planes. For every telescope plane (1–6) the substrate type, pixel flavour, thickness, and operating voltage have been indicated. The position of the planes with respect to Plane 1 are listed

MALTA telescope planes						
Plane Specifications	1	2	3	4	5	6
Substrate type	Epi	Cz	Cz	Cz	Cz	Epi
Sensor flavour	STD	NGAP	STD	STD	NGAP	STD
Total thickness [µm]	100	100	300	300	100	300
Operation voltage [V]	− 6	− 6	− 30	− 30	− 6	− 6
Distance [cm]	0	8	16	94	102	110

located on top of their linear stage, as seen in Fig. 4. The box is completely encapsulated with a 4 cm thick expanded polystyrene layer, which insulates, in addition to the polymethyl methacrylate walls of 5 mm, the interior. The chiller is connected to a heat exchanger and a fan to circulate the air inside the box. The usage of an external dry air source reduces the humidity in order to avoid water condensation and ice formation inside the box. A temperature, humidity, and dew-point sensor allows for remote monitoring of the climate inside the coldbox. Through a small slit in the coldbox, the dry air pipes, power, communication, and temperature-monitoring cables are drawn in. If no cooling is required, the lid of the coldbox can be removed and the operation can proceed at ambient temperature and humidity.

Every tracking plane is connected to a Xilinx Kintex (KC-705) or Virtex (VC-707) commercial FPGA evaluation board using a flexible HPC FMC cable. The FPGAs are in turn connected to the TLU through coaxial cables, further described in Sect. 3.2. A custom framework allows the remote control of the power supplies, including TTIs (PL303QMD-P) and Keithleys (2410). This has allowed to separately control and monitor the power domains of the tracking planes and DUTs, i.e. the power supplied to the substrate (SUB), PWELL, and the analog and digital voltage domains of the chip (AVDD and DVDD). A network-connected Power Distribution Unit (APC AP7921B) with multiple outputs distributes the power for several parts of the telescope including the PC, linear stages, and private network switches. This allows the system to be remotely rebooted independently from the PC and the private network.

3.2 Trigger logic unit

The predecessor of the current MALTA TLU was based on Nuclear Instrumentation Modules (NIM), requiring manual configuration of the trigger logic, which made the process a complex and time-consuming task. However, FPGAs can be used to build a more versatile TLU for test beam campaigns, additionally reducing the cost and weight. The current TLU of the MALTA telescope is based on a Kintex-7 KC705 evaluation board, which is used to process the combination logic and provide online monitoring. The TLU is interfaced using

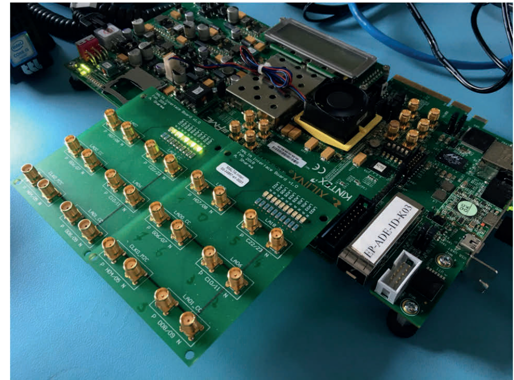


Fig. 5 SMA to FMC converter cards interface with TLU Kintex-7

SMA cables to the telescope planes and scintillator through two custom SMA-to-FMC converter cards [12] for input and output signals, as seen in Fig. 5. The Gigabit Ethernet port uses the IPbus protocol [13] for readout communications, control, and configuration. A micro-USB port is used for firmware programming of the FPGA through JTAG. The connections between the telescope and the TLU are shown in the Fig. 6.

Figure 7 shows the process that triggers the recording of events in the MALTA telescope. When a particle is detected, each MALTA plane produces two signals. The first signal is a fast signal (or hitOR), showed in red, that is sent without any processing to the TLU. The second signal is a full time stamping signal (shown in blue) which, after a defined internal delay, is sent out of the FPGA to be recorded by the standard readout. Depending on the telescope configuration, up to 4 fast MALTA signals are connected to the TLU input connectors. These signals are processed and then addressed in a combination logic to produce a combined signal. The scintillator produces a faster signal, also in red, that allows the TLU for precise timing measurements. The coincidence of MALTA planes in the combined signal has a time resolution of several ns but the fast signal of the scintillator is used as a timing reference, hence ensuring a precision in the trigger signal of a fraction of a ns. The trigger signal or Level

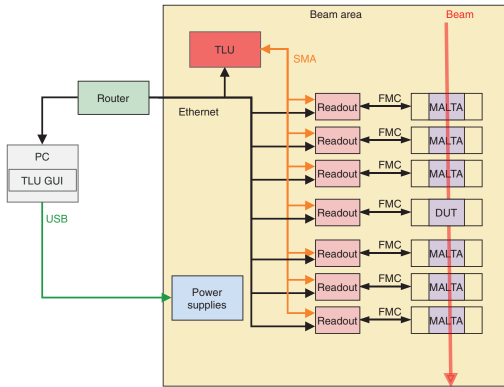


Fig. 6 Diagram of the connections between the MALTA telescope and the TLU

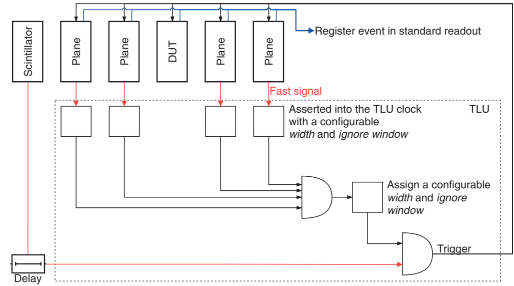


Fig. 7 Diagram of the signal processing from fast signals to trigger and data recording. The dashed black line represents the processes that occur inside the hardware of the TLU

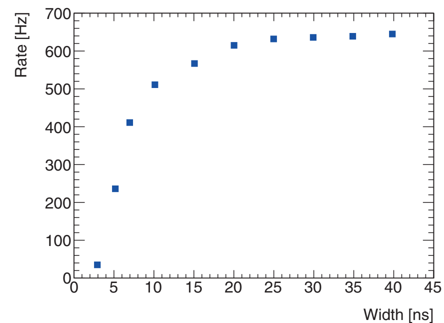


Fig. 8 Trigger rate as a function of the length assigned to the fast signals to the combination logic. Measurement was done with low energy electrons from ^{90}Sr β-decay

1 Accept (L1A) is sent from the TLU to each MALTA plane (FPGA) to trigger the recording of the full time stamping signal by the standard readout.

The FPGA runs a firmware divided in 3 main modules: input, coincidence logic, and output. A 320 MHz clock is generated from the FPGA internal clock for signal processing modules and logic. In the input module, as seen in Fig. 7, the asynchronous input fast signals are captured by a signal processing block into the internal clock. Each signal is transformed into a standard logic for the subsequent processing including the stretching to a programmable length and the implementation of a possible veto window to avoid too close signals. The module contains a 32-bit counter to monitor the input rate of each channel.

The coincidence module combines the selected individual channels in an AND gate. The width of the signal from the previous step, i.e. before the coincidence, acts as a coincidence window in the combination step. This window is necessary due to the nature of the signal from the MALTA planes on which the arrival time of the hits is proportional to the charge deposition. As such, the input signals are spread by typically 5 to 15 ns. This is observed in Fig. 8, where the L1A rate as a function of the stretched window width of the signals for three telescope planes is shown. This measurement was performed with a ^{90}Sr source, which explains why the observed rate is lower than in test beam conditions. The larger the time window of the MALTA planes is, the greater the opportunity is for the coincidence logic to form a L1A trigger. A saturation effect can be observed after approximately 25 ns as hit signals are ignored during the long stretched processed signals.

The output processing module is similar to the input processing, though it allows to regulate the output signal. The capability to control the output signal length is important to interface the TLU to the devices receiving the trigger, while the veto is used to implement a maximum trigger rate. While individual chips and the TLU can theoretically perform at very high rate, the output rate of the TLU is limited to 50 kHz due to limitations of the FIFO size in the FPGA that interfaces the individual planes.

3.3 Data aquisition

The DAQ process of the TLU is implemented inside the FPGA as a Finite State Machine (FSM) that is controlled remotely via IPbus. A Command Line Interface (CLI) and Graphical User Interface (GUI) are available for this, the latter shown in Fig. 9. The baseline is a C++ class responsible of communication with the FPGA. The upper panels allow the selection of trigger planes, maximum allowed rate (veto), the output signal width, L1A and connection settings. The middle panel allows the user to start and stop a data-taking run. The lower panel is used to monitor the number of triggers of the respective planes and L1A.

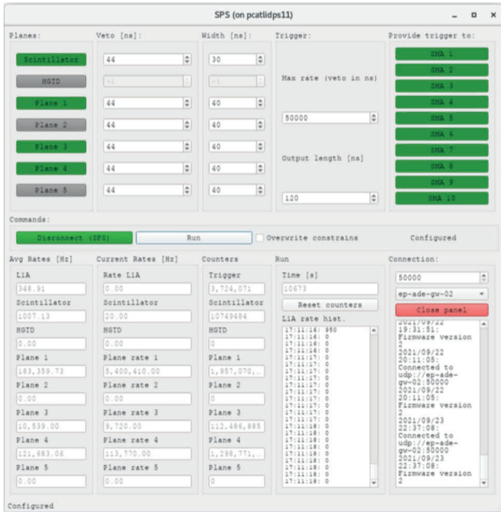


Fig. 9 Screenshot of the in-time trigger logic unit GUI during data-taking. The GUI allows the user to start and stop a run, select on which plane (or scintillator) to trigger on, the maximum rate that is allowed, the output length of the trigger signal, and displays in-time rate for every plane

Figure 10 shows the in-time data acquisition window during data-taking. Every column represents a tracking plane (Plane 1–6), while the last two planes represent the two DUTs (Plane 7 and 8). The top row shows the hit map of each plane, which allows additionally to visualise a ROI. The middle row shows the number of hits as a function of time of arrival since the L1A signal with respect to the trigger scintillator. The scale of the plot (500 ns) represents the readout window after the L1A. The bottom row shows the hit multiplicity histogram, i.e. the number of hits as a function of the number of pixel per event. The realtime monitoring plots allow for quick feedback on and fast assessment of the DUT performance.

4 Reconstruction and offline analysis

Precise knowledge of the positions of the detector planes is required to guarantee a high track quality and correct association with DUT clusters. The MALTA telescope is mechanically built with as much precision as possible. However, as robust mechanical frameworks have a limited spatial accuracy when setting the positions of the planes, offline measurements of their deviation can be accounted for during the software alignment. For the alignment, track reconstruction, and the offline analysis of test beam data the software package Proteus is used [9]. Proteus software takes raw data in the form of hits per event and groups them into clusters, where a cluster is defined as the geometrical average of the adjoining hit positions. It can also provide track reconstruction by finding tracks from clusters on the tracking planes and it calculates the optimal track parameters on selected planes. Finally, Proteus provides the user with output data for further offline analysis. The sequence of alignment, reconstruction, and analysis with Proteus are discussed in more detail below. Using these modular steps, the performance of the telescope planes is discussed in the final section of this chapter.

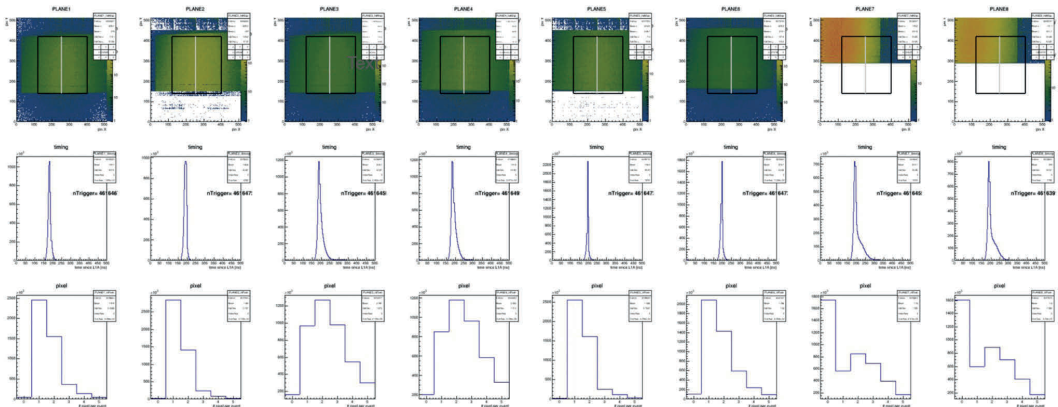


Fig. 10 Screenshot of the in-time telescope data acquisition window during data-taking. From left to right the six tracking planes are indicated, with the last two columns representing the two DUTs. The top plot shows the hit-map, the middle plot shows the time of arrival with respect to the trigger, and the bottom shows the hit multiplicity

Table 2 Overview of the characterisation of the MALTA telescope planes. For every telescope plane (plane 1–6) the average efficiency, average number of hits per event within a 500 ns acquisition window, and average cluster size are listed. The error for the average efficiency is expressed as the statistical uncertainty. The error on the latter two parameters (average number of hits per event and cluster size) is expressed as the Root Mean Square (RMS) of the relative distributions

Characterisation of telescope planes						
Plane	1	2	3	4	5	6
Average efficiency [%]	97.4 ± 0.1	96.9 ± 0.1	99.2 ± 0.1	99.0 ± 0.1	95.0 ± 0.1	94.0 ± 0.1
Average number of hits per event	1.7 ± 0.9	1.8 ± 1.0	2.2 ± 1.1	2.5 ± 1.2	1.7 ± 0.9	1.6 ± 0.9
Average cluster size [pixels]	1.1 ± 0.3	1.1 ± 0.4	1.7 ± 0.8	1.9 ± 0.9	1.1 ± 0.3	1.1 ± 0.4

4.1 Alignment

The alignment step accounts for possible misalignment from the nominal telescope description. This nominal telescope position uses Plane 1 (upstream of the beam) as the origin of a global coordinate system to which the other planes downstream are referenced towards. In total, three rotational and three translational degrees of freedom (around x-, y- and z-axis) per tracking plane are considered. During the first step of the alignment, i.e. the coarse alignment, the hit correlation distribution between consecutive planes is calculated. Assuming a parallel beam, the two translational coordinates (x-and y-axis) perpendicular to the beam are inferred by the correlation offset. The next alignment step, i.e. fine alignment algorithm, uses the unbiased residuals to iteratively perform a χ^2-minimization. This fine alignment is motivated by the fact that the track position is influenced by the position of the given plane being aligned. The fine alignment runs until the χ^2 is minimal and convergence is achieved, where any calculated shift is used in the initial geometry.

4.2 Track fitting

The alignment step is completed when the positions of the telescope planes are well-defined, after which the tracks can be reconstructed. By using a seed cluster, the tracking algorithm searches for clusters on the consecutive telescope planes. Depending on the beam type, an angle in which this search is performed can be chosen in order to compensate for potential scattering. In case multiple clusters are found within the region the algorithm searches, the track search will split and continue in both directions. Only tracks that contain the largest number of associated clusters are kept, after which a χ^2-cut is applied to filter out tracks of bad quality, for instance due to multiple scattering or nuclear interaction with the telescope planes. The $\chi^2/\text{NDF}<10$ requirement is applied to the tracks, which allows for some buffer during automated analysis for unexpected misalignment without significantly affecting the performance in terms of efficiency nor timing resolution.

4.3 DUT cluster matching

Finally, hit clusters on the DUT need to be associated to a track. In order to do this, the fitted track is extrapolated to the DUT and the matched clusters are restricted to a user-defined distance. The minimum value of this distance is dependent on the spatial resolution of the DUT. In our studies this distance is set to 2.5 times the size of the pixel pitch (36.4 μm), i.e. 80 μm. When the matched clusters on the DUT are found, it is possible to define the DUT efficiency as the number of tracks with a matched cluster on the DUT over the total number of reconstructed tracks.

4.4 Characterisation of telescope planes

The telescope planes were characterised during test beam measurements through dedicated runs where one plane at a time was selected as the DUT. Due to the flexibility of the TLU, other planes except the DUT could be selected as triggering planes. As a result the average efficiency and cluster size of these planes could be measured. The results of these measurements are presented in Table 2. Additionally, in this table the average number of hits per event (within a 500 ns acquisition window) is shown, a variable that can be monitored during real-time data-taking (Fig. 10). The results are in line with the quality conditions of MALTA. Telescope planes 3 and 4 have larger efficiency, cluster size, and hit multiplicity, as they are Cz STD sensors which allow for larger depletion depths at higher bias voltages.

5 Spatial resolution

To evaluate the spatial telescope resolution first an analytical estimate [11] that takes into account the planes' position, the radiation length and intrinsic resolution of the sensors and the beam energy are used. For sufficiently high momentum of the incident particles the latter parameter does not have an impact on the resolution. This is the case of the following estimation where the charged hadron beam used at SPS has an energy greater than $\mathcal{O}(100)$ GeV. The intrinsic resolution

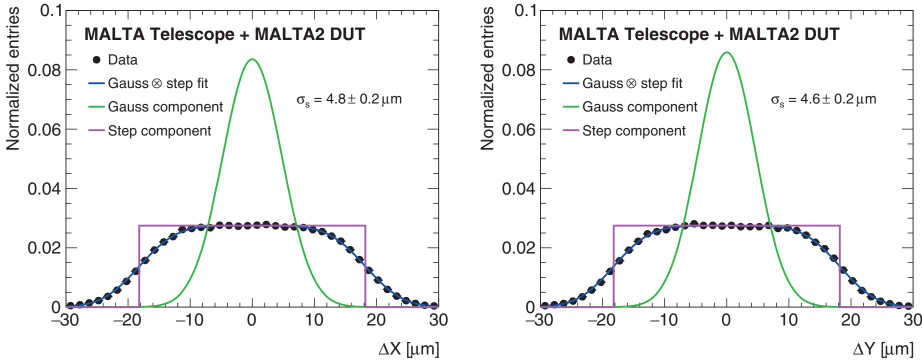


Fig. 11 Distribution of the residuals on the X (left) and Y (right) directions between the fastest sensor hit and the track intercept. The fit result of a convolution of a Gaussian with a two-sided step distribution is shown (blue) together with the Gaussian (green) and step function (magenta) components

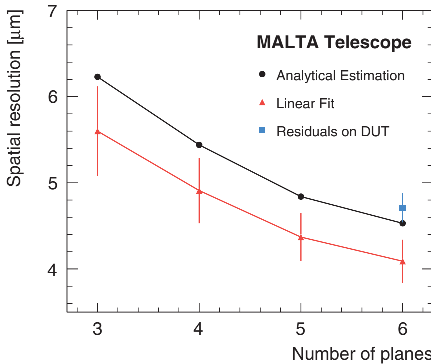


Fig. 12 Telescope resolution as a function of the number of planes considered. The analytical estimation (black) is compared with the measurements based on the linear fit (red). The total resolution is also extracted from the fit of the convolution of a Gaussian with a two-sided step function on the distribution of the residuals on a DUT (blue)

of the sensor is set to $\sigma_{int} = 10.5$ μm assuming a uniform distribution on the pixel pitch. The radiation length of the sensor is about 0.1% for the planes with samples 100 μm thick and 0.3% for those of 300 μm.

Tracks are reconstructed with a linear fit of the clusters position based on the χ^2 minimisation method. This provides the measurement of the several tracks parameters and the projection of the intercepts on the DUT. The uncertainty of the latter gives an estimation of the spatial resolution. Only events where exactly one track is reconstructed fulfilling the aforementioned fit quality criteria are selected. The RMS of the estimated resolutions is assigned as the uncertainty of the central value. The General Broken Lines (GBL) algorithm [10] was also tested to account for Coulomb scattering due to the material interaction of the beam with the telescope planes and the DUTs. Results are found to be compatible with the linear fit regression.

An alternative approach adopted to evaluate the tracking resolution is based on the distance from the track intercept and the fastest hit of the closest reconstructed cluster of a DUT. At first, a residual mis-alignment correction is applied to data by subtracting from each calculated distance in the pixels its average obtained in the corresponding position in the chip matrix. The track spatial resolution of the telescope can be described by a Gaussian function whereas the probability of hitting the sensor is assumed uniform along the pixel pitch. Hence, the convolution of a Gaussian and a step function is used to perform a fit on the residual on the two transversal coordinates X and Y of the DUT, as shown in Fig. 11. The resulting spatial resolution is extracted from the width of the Gaussian distribution.

The analytical description of the spatial resolution on the DUT position as a function of the number of telescope planes is shown in Fig. 12 together with the estimation given by the track parameters obtained from the linear fit. The set of planes chosen in the several configurations provides the best measured resolution for the fixed subset of samples. These always include the planes closer to the DUT, having a cluster size greater than one, hence improving the spatial resolution. The comparison highlights how the cluster size, not taken into account in the simulation, affects the spatial resolution improving the predictions by about 10% for all the tested configurations. The measured spatial resolution based on the linear regression approach of the full telescope is $\sigma_s = 4.1 \pm 0.2$ μm. The average spatial resolution extracted from the Gaussian fit of the two transversal components is also shown and corresponds to $\sigma_s = 4.7 \pm 0.2$ μm. This latter approach provides a larger estimate if compared with the analytical description and the linear regression measurements. This is

attributed to the effects of inhomogeneity and time resolution at the edges of the DUT that are not taken into account in the fit and inflate the width of the Gaussian function.

6 Timing resolution

The timing performance of the MALTA tracking planes has been measured extensively during the past test beam campaigns. In order to describe the timing response of the telescope planes, several intrinsic and extrinsic effects need to be accounted for. One of these systematic intrinsic effects is the signal propagation, i.e. the time required to reach the periphery along the column direction in the pixel matrix. This effect consists of a contribution of the pulse propagation to the pixel and a contribution from the hit propagation to the periphery. As this effect has a linear behaviour, a dedicated propagation correction is applied and integrated over the whole pixel matrix. An extrinsic effect that is taken into account is the jitter originating from the scintillator, approximately 0.5 ns, and from the FPGA. The FPGA latches the trigger with the 320 MHz clock, which adds a 3.125 ns per-event jitter. Therefore, the contribution of the FPGA to the timing distribution is approximated at $3.125/\sqrt{12} \sim 0.9$ ns. It should be pointed out that due to the fact that the trigger is internally latched within the FPGA, a lot of flexibility and adaptability to any trigger source is gained within the set-up.

Figure 13 shows the track timing, where the time of arrival of the fastest hit in the cluster with respect to the scintillator for the six tracking planes is averaged to a single timing distribution. The RMS of the timing distribution of all six telescope planes were found to be compatible. The standard deviation extracted from a Gaussian fit performed on the track timing distribution provides an estimation of the timing resolution of $\sigma_t = 2.1$ ns. The excellent timing performance allows to separate the individual contributions of the track, even in high rate beam settings ($> 4 \times 10^6$ particles per spill).

7 Device under test integration

Ultimately, a telescope is dedicated to the characterisation of R&D prototypes. As upcoming colliders and physics experiments will become more demanding, the next generation detectors will need to be able to operate at high rate and fluence conditions. The flexibility of the MALTA telescope has allowed to test a wide variety of detector prototypes during the test beam campaigns of 2021 and 2022, which provided relevant insights and results for future LHC experiments and upgrade programs. This section will highlight results that were achieved with two different types of DUTs and their respective integration in the MALTA telescope.

7.1 MALTA2: latest radiation-hard full scale MALTA sensor

The next prototype of the MALTA family, MALTA2, had a dedicated test beam campaign in the MALTA telescope starting in May 2021 until November 2021 and did an equal lengthy campaign in 2022. During this test beam campaign, the radiation hardness and timing performance of the second generation sensors was the focus point. Furthermore, post-processed MALTA2 samples were extensively tested.

Analogous to its small-scale demonstrator, the Mini-MALTA [14], MALTA2 has two front-end designs implemented: the standard MALTA design and a cascode design. The size of selected transistors was increased to reduce the Random Telegraph Signal (RTS) noise, which opens up the possibility to operate at lower threshold and therefore reach higher efficiencies [15,16]. The matrix of MALTA2 consists of 224 × 512 pixels, with equal pixel size of the original MALTA (36.4 μm). Effectively, the active area of the chip corresponds to ~ 18.33 mm^2. Similar to MALTA, this second prototype is fabricated on both Epi and Cz substrate and both the NGAP and XDPW pixel flavors are implemented.

The benefits of testing the MALTA2 sensor in the MALTA telescope become evident from Fig. 14. This figure shows a projection of the difference between the time of arrival of the leading hit in a pixel cluster and the average arrival time of signals over the entire chip projected on a 2 × 2 pixel matrix. As one can observe from the plot, a difference of 2–3 ns is observed between signals originated from the pixel center and the signals from the corners [17]. Here, the excellent resolution of the telescope allows for fine observation of these effects between pixels and charge sharing are useful insights for future simulation work. These results have prevailed that the MALTA telescope has allowed for precise hit timing for DUTs of the order of $\mathcal{O}(ns)$.

7.2 Carbon-enriched low gain avalanche detectors for the ATLAS high granularity timing detector

Low Gain Avalanche Detectors (LGAD) have been extensively studied within the context of the High Granularity Timing Detector in the ATLAS experiment, which will be added during the Phase-II upgrade of the High-Luminosity Phase of the LHC [18]. The LGAD sensors are required to achieve at a maximum fluence of 2.5×10^{15} 1 MeV n_{eq}/cm^2, among other requirements, a timing resolution of 50 ps per hit, a hit efficiency of 97% and a collected charge > 4 fC at the start of their operation. During the summer of 2021, a dedicated test beam campaign with the MALTA telescope was performed using irradiated (up to 2.5×10^{15} 1 MeV n_{eq}/cm^2) LGAD sensors with a carbon enriched gain layer in order to improve the radiation hardness [19]. The MALTA telescope was used to track the incident charged particles and

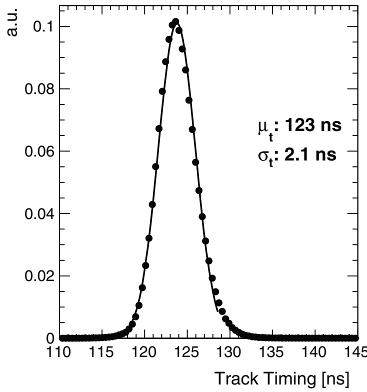


Fig. 13 Averaged timing distribution of the six MALTA tracking planes with respect to the scintillator reference. The quoted standard deviation and mean result from fitting a Gaussian to the core of the distribution. Measurements were done with a 180 GeV hadron beam at SPS at CERN in 2022

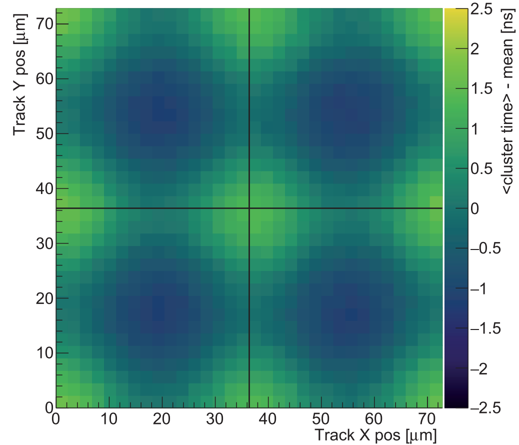


Fig. 14 In-pixel timing projected over a 2 × 2 pixel matrix for a MALTA2 Epi, XDPW, 100 μm thick, low doping of n-blanket, at -6 V SUB bias and -6 V PWELL bias. Colour scale indicates the difference in timing of the leading hit in the cluster and the average timing over the entire matrix. Threshold corresponds to 130 electrons. Measurements were done with a 180 GeV hadron beam at SPS at CERN in 2021

provide the position of the incoming particle in the frame of each DUT. Below, the integration of LGAD sensors into the MALTA telescope is described.

The LGAD sensors, manufactured by Fondazione Bruno Kessler (FBK) and Institute of High Energy Physics (IHEP), are mounted on a custom readout board with integrated amplification stages to enhance their signal. A four-channels oscilloscope is used to sample the waveform from the DUTs and initiates its DAQ by the trigger system. The firmware of the TLU is modified such that it only uses the accept state from the oscilloscope, where only events triggered and accepted are recorded. In order to trigger on a particle, the second MALTA tracking plane is used in coincidence with the scintillator placed behind tracking Plane 6. If the two sensors record a signal, the TLU records the telescope data from all six planes and the waveforms from both the DUT and a second LGAD that is used as a timing reference (timing resolution of ∼ 55 ps). Finally, the track reconstruction is performed with the software package Proteus, as described in Sect. 4. The tracks that are reconstructed are extrapolated onto the plane of the DUT, taking into account multiple scattering by using the GBL algorithm (discussed in Sect. 5).

One of the key parameters that was measured during the test beam campaign was the time resolution of the LGAD sensors. The time resolution is extracted by subtracting the time of arrival of the DUT with the second LGAD installed for time reference. Figure 15 shows the time resolution for sensors irradiated at a fluence of 1.5 and 2.5 × 10^{15} 1 MeV n_{eq}/cm^2 as a function of the bias voltage. As can be observed in this plot, the time resolution improves with bias voltage and a time resolution of 40 ps for sensors FBK−2.5 (at 550 V) and 43 ps for IHEP−2.5 sensors (at 450 V) is achieved. Figure 16 shows that as the bias voltage increases, efficiencies up to approximately 99% can be achieved for sensors irradiated to 1.5 × 10^{15} 1 MeV n_{eq}/cm^2. For sensors irradiated to higher fluences, efficiencies up to ∼ 95% can be obtained.

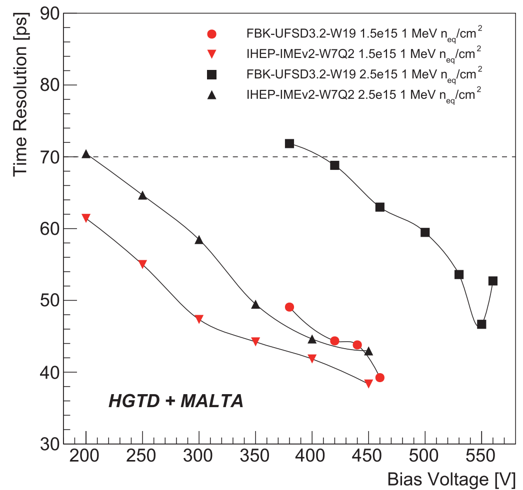


Fig. 15 Time resolution as a function of bas voltage for the HGTD test beam campaign in the MALTA telescope. The dashed lines corresponds to the minimum requirements for the future HGTD. The results are shown for different single-pad sensors (FBK and IHEP) irradiated at 1.5 and 2.5 × 10^{15} 1 MeV n_{eq}/cm^2

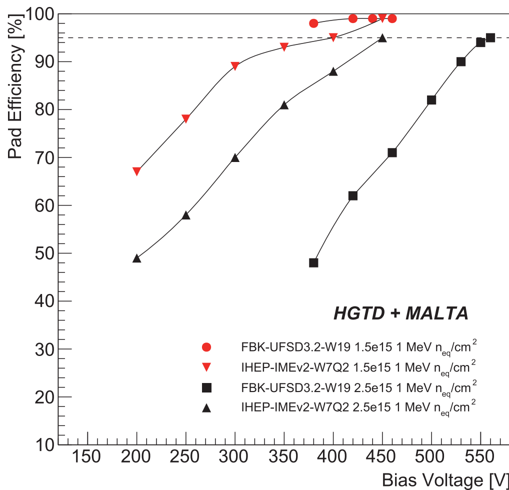


Fig. 16 Efficiency as a function of bias voltage for the HGTD test beam campaign in the MALTA telescope.The dashed lines corresponds to the minimum requirements for the future HGTD. The results are shown for different single-pad sensors (FBK and IHEP) irradiated at 1.5 and 2.5×10^{15} 1 MeV n_{eq}/cm^2

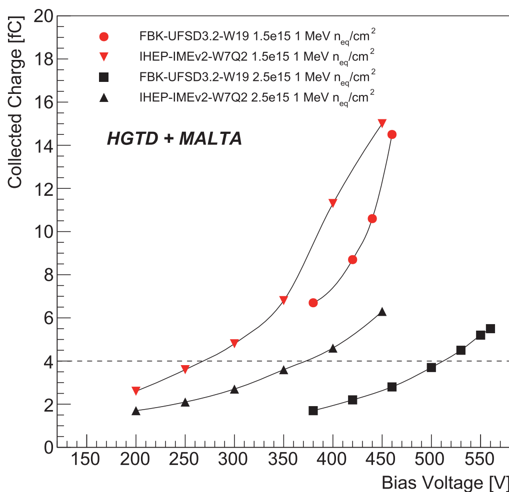


Fig. 17 Collected charge as a function of bias voltage for the HGTD test beam campaign in the MALTA telescope. The dashed lines corresponds to the minimum requirements for the future HGTD. The results are shown for different single-pad sensors (FBK and IHEP) irradiated at 1.5 and 2.5×10^{15} 1 MeV n_{eq}/cm^2

Finally, Fig. 17 shows that at a fluence of 1.5×10^{15} 1 MeV n_{eq}/cm^2, the collected charge is above the minimum required amount of 4 fC. At higher fluence, the collected charge at the same bias voltage is less, though the HGTD requirements are still fulfilled at higher bias voltage. These results confirm the feasibility of an LGAD-based timing detector for HL-LHC. More extensive analysis and description of this test beam campaign can be found in Ref. [19].

8 Conclusion

Since 2021, the MALTA telescope has been extensively used for the characterization of R&D prototypes during test beams in the H6 beamline of the North Area at CERN. The MALTA telescope is based on six monolithic MALTA tracking planes, which are fabricated in Tower 180 CMOS imaging technology. The MALTA telescope exploits the best qualities of the MALTA sensor: a full prototype (large area), high granularity, self-triggering, and good spatial and timing resolution. This allowed to build a telescope that offers a high degree of flexibility and versatility, and allows to test a wide variety of prototypes that target different requirements. The MALTA telescope provides a spatial resolution of $\sigma_s = 4.1 \pm 0.2$ µm, based on the linear regression approach, and a track timing resolution of $\sigma_t = 2.1$ ns. Continuous detector R&D activities will be a crucial ingredient in order to achieve ultimate performance at future colliders and therefore the continuous development and improvement of test beam telescopes is essential.

Acknowledgements This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement number 101004761 (AIDAinnova), and number 654168 (IJS, Ljubljana, Slovenia). Furthermore it has been supported by the Marie Sklodowska-Curie Innovative Training Network of the European Commission Horizon 2020 Programme under contract number 675587 (STREAM).

Data Availability Statement This manuscript has no associated data or the data will not be deposited. [Authors' comment: All data generated during this study are contained in this published article.]

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.
Funded by SCOAP[3]. SCOAP[3] supports the goals of the International Year of Basic Sciences for Sustainable Development.

References

1. ATLAS Collaboration et al., Technical design report for the ATLAS inner tracker strip detector, ATL-TDR-025 (Section 15.1) (2017)
2. H. Pernegger et al., First tests of a novel radiation hard CMOS sensor process for depleted monolithic active pixel sensors. JINST **12**, P06008 (2017)
3. W. Snoeys et al., A process modification for CMOS monolithic active pixel sensors for enhanced depletion, timing performance and radiation tolerance. Nucl. Instrum. Methods A **871**, 90–96 (2017)
4. R. Cardella et al., MALTA: an asynchronous readout CMOS monolithic pixel detector for the ATLAS High-Luminosity upgrade. JINST **14**, C06019 (2019)
5. I. Caicedo et al., The Monopix chips: depleted monolithic active pixel sensors with a column-drain read-out architecture for the ATLAS Inner Tracker upgrade. JINST **14**, C06006 (2019)
6. E.J. Schioppa et al., Measurement results of the MALTA monolithic pixel detector. Nucl. Instrum. Methods A **958**, 162404 (2020)
7. M. Munker et al., Simulations of CMOS pixel sensors with a small collection electrode, improved for a faster charge collection and increased radiation tolerance. JINST **14**, C05013 (2019). arXiv:1903.10190
8. H. Pernegger et al., MALTA-Cz: A radiation hard full-size monolithic CMOS sensor with small electrodes on high-resistivity Czochralski substrate, arXiv preprint. arXiv:2301.03912 (2023)
9. M. Kiehn et al., Proteus beam telescope reconstruction (2019). https://doi.org/10.5281/zenodo.2586736
10. C. Kleinwort et al., General broken Lines as advanced track fitting method, NIM 673 (2012). https://doi.org/10.1016/j.nima.2012.01.024
11. Telescope Optimizer. https://mmager.web.cern.ch/telescope/tracking.html
12. A. Sharma, MALTA FMC Board Version 4.0, Tech Rep. CERN EDMS Number 2142554. https://edms.cern.ch/document/2142554/
13. C. Ghabrous Larrea et al., MIPbus: a flexible Ethernet-based control system for xTCA hardware. JINSThttps://doi.org/10.1088/1748-0221/10/02/C02019
14. M. Dyndal et al., Mini-MALTA: radiation hard pixel designs for small-electrode monolithic CMOS sensors for the High Luminosity LHC. JINST **15**, P02005 (2020)
15. F. Piro et al., A 1-μW radiation-hard front-end in a 0.18-μm CMOS process for the MALTA2 monolithic sensor. IEEE Trans. Nucl. Sci. **69** (2022)
16. M. van Rijnbach et al., Radiation hardness and timing performance in MALTA monolithic pixel sensors in TowerJazz 180 nm. JINST **17**, C04034 (2022)
17. G. Gustavino et al., A timing performance of radiation hard MALTA monolithic pixel sensors. JINST **18**(03), C03011 (2022)
18. ATLAS Collaboration et al., Technical Design Report: A High-Granularity Timing Detector for the ATLAS Phase-II Upgrade CERN-LHCC-2020-007, ATLAS-TDR-031 (CERN, 2020). https://cds.cern.ch/record/2719855
19. S. Ali et al., Performance in beam tests of carbon-enriched irradiated Low Gain Avalanche Detectors for the ATLAS High Granularity Timing Detector, arXiv preprint. arXiv:2303.07728 (2023)

Paper III

Radiation Hardness of MALTA2 Monolithic CMOS Sensors on Czochralski Substrates

M. van Rijnbach et al.

Submitted for publication. arXiv: 2308.13231.