VRIJE UNIVERSITEIT BRUSSEL

FACULTEIT WETENSCHAPPEN EN
BIO-INGENIEURSWETENSCHAPPEN DEPARTEMENT
NATUURKUNDE

The Upgraded Outer Tracker for the CMS Detector at the High Luminosity LHC, and Search for Composite Standard Model Dark Matter with CMS at the LHC

JARNE DE CLERCQ

PROMOTER
PROF. DR. STEVEN LOWETTE

Proefschrift ingediend met het oog op het behalen van de academische graad van Doctor in de Wetenschappen

January 2020

Doctoral examination commission:

Prof. Dr. Krijn De Vries (Vrije Universiteit Brussel, *secretary*)
Prof. Dr. Jorgen D'Hondt (Vrije Universiteit Brussel, *chair*)
Prof. Dr. Steven Lowette (Vrije Universiteit Brussel, *supervisor*)
Prof. Dr. Nadia Pastrone (INFN, Torino)
Prof. Dr. ir. Gerd Vandersteen (Vrije Universiteit Brussel)
Prof. Dr. Pascal Vanlaer (Université Libre de Bruxelles)

Cover illustration: Patrick De Clercq

Printed by
Crazy Copy Center Productions
VUB Pleinlaan 2, 1050 Brussel
Tel : +32 2 629 33 44
crazycopy@vub.ac.be
www.crazycopy.be
ISBN: 9789493079571
NUR: 924, 926, 987

Contents

Summary

The Standard Model of particle physics describes the interactions between particles at the smallest known scale. The Standard Model has passed many tests with flying colours, but several observations, such as the need for dark matter, do not find an explanation in this framework. The Large Hadron Collider (LHC) at CERN is an excellent tool to study the Standard Model and to try and find answers to the open questions. The LHC collides protons resulting in a center of mass energy of 13 TeV, the highest energies ever attained at a particle collider. This enables physicists to study matter at an energy scale which is normally not accessible and which reflects the energy density in our Universe a fraction of a second after the Big Bang.

At one of the four interaction points along the LHC, the Compact Muon Solenoid (CMS) experiment is located. This general purpose experiment is designed to give an as good as possible measurement of the kinematic properties of the particles produced in a collision. CMS does this by using several subdetectors, one of which is the tracking system. The CMS tracker is based on silicon technology and consists of a pixel and a strip detector, designed to reconstruct the tracks of charged particles. The strip detector has been taking data since the start of the operation of the LHC in 2010 and will continue to do so until 2024. During the 2025-2027 Long Shutdown of the LHC, the accelerator complex will be upgraded to start the High Luminosity LHC (HL-LHC) phase at the end of 2027, which will provide the experiments with higher luminosities by increasing the number of proton-proton collisions per bunch crossing (pileup). In order to keep tracking performance at pre-HL-LHC levels in this harsher HL-LHC environment, the strip tracker will be replaced by an Outer Tracker, consisting of pixel-strip (PS) and strip-strip (2S) modules. These modules have local track reconstruction logic. The output of this logic is used as input to the L1 trigger system.

This thesis consists of a hardware part, focussing on the testing of Outer Tracker prototypes, and a search for new physics using the data collected by the CMS experiment.

The Upgraded Outer Tracker for the CMS Detector at the High Luminosity LHC:

Over the past years, Outer Tracker prototype ASICs and modules have become available. These require a dedicated test bench for characterization. This thesis discusses in detail the Phase-2 Outer Tracker and the structure and scope of the firmware project (μDTC) set up to read out ASIC and module prototypes of this new detector. The test bench is designed to test both 2S and PS modules and components. Several tests, to which the author contributed, are presented, going from bench-top testing of single ASICs, through radiation hardness testing of these chips and operating module prototypes in test beams.

Search for Composite Standard Model Dark Matter with CMS at the LHC:

The Sexaquark (S), composed of uuddss quarks, is a hypothetical particle that was proposed to be stable and a potential dark matter candidate. $\bar{\mathrm{S}}$ particles could be produced in the proton-proton collision and could subsequently annihilate on a neutron in the beampipe or detector material. This annihilation could result in a K^0_S and $\bar{\Lambda}^0$ which in turn can decay to charged products which are reconstructable with the CMS

tracker. This is the signal used in this thesis to look for the $\bar{\mathrm{S}}$ in the CMS 2016 dataset by reconstructing the $\bar{\mathrm{S}}$ kinematic properties and its annihilation vertex. The difficulties of reconstructing such a low momentum, displaced and off-pointing signature with the default CMS reconstruction algorithms will be studied and a first-ever limit on the $[\sigma(\mathrm{p}+\mathrm{p}\rightarrow\bar{\mathrm{S}})\times\sigma(\bar{\mathrm{S}}+\mathrm{n}\rightarrow\mathrm{K}^{0}_{\mathrm{S}}+\bar{\Lambda}^{0})]$ cross section will be presented.

Samenvatting

Het Standaardmodel van de deeltjesfysica beschrijft de interactie tussen deeltjes op de kleinste schaal. Dit model heeft reeds vele testen doorstaan. Toch vinden verschillende observaties - zoals de nood aan donkere materie - geen verklaring in dit framework. De Large Hadron Collider (LHC) aan het CERN is een excellent platform om het Standaardmodel te bestuderen en om antwoorden te vinden op onbeantwoorde vragen. De LHC versnelt protonen om ze nadien te laten botsen bij een massamiddelpuntsenergie van 13 TeV. Dit is de hoogste energie ooit bereikt in een laboratorium en daardoor stelt de LHC fysici in staat om materie te bestuderen bij een energieschaal die normaal niet toegankelijk is en die de energiedichtheid van ons universum reflecteert een fractie van een seconde na de Big Bang.

Aan één van de vier interactiepunten van de LHC bevindt zich het Compact Muon Solenoid (CMS) experiment. Dit experiment is ontworpen om een zo goed mogelijke meting te maken van de kinematische eigenschappen van de deeltjes die geproduceerd worden bij een botsing. CMS doet dit door gebruik te maken van verschillende subdetectoren. Een van deze subdetectoren is de sporendetector. De CMS-sporendetector is gebaseerd op siliciumtechnologie en bestaat uit een pixel- en een stripdetector. Deze detectoren zijn ontworpen om het pad van geladen deeltjes te reconstrueren. De huidige stripdetector is in gebruik sinds het begin van de LHC-exploitatie in 2010 en zal verder operationeel zijn tot 2024. Tijdens de Long Shutdown in 2025-2027 zal het versnellercomplex geüpgraded worden om de hogeluminositeitsfase (HL-LHC) van de LHC aan te vatten eind 2027. De HL-LHC zal de experimenten voorzien van hogere luminositeit door het aantal proton-proton botsingen per bundelkruising te verhogen. Om de performantie van de sporenreconstructie te garanderen in deze moeilijke omstandigheden zal de stripdetector vervangen worden door een nieuwe *Outer Tracker*. Deze subdetector zal bestaan uit twee soorten detectormodules: pixel-strip (PS) en strip-strip (2S) modules. Deze modules hebben logica om lokaal sporen te reconstrueren. De output van deze logica wordt gebruikt als input voor het L1-trigger-systeem.

Deze thesis bestaat uit een hardware-gedeelte, dat focust op het testen van Outer Tracker-prototypes, en een zoektocht naar nieuwe fysica, gebruikmakend van data verzameld door het CMS-experiment:

De Upgrade van de Buitenste Sporendetector van de CMS-detector voor de Hoge Luminositeit LHC:

De prototype ASICs en modules voor de Outer Tracker vereisen een specifieke testopstelling voor karakterisatie. Deze thesis beschrijft in detail de Outer Tracker en het doel en de structuur van het firmware-project (μDTC) dat werd opgezet om prototypes van ASICs en modules van deze nieuwe detector uit te lezen. De testopstelling is ontworpen om zowel 2S als PS modules en componenten te testen. Verschillende testen, waaraan de auteur heeft bijgedragen, worden voorgesteld: testen van standalone chips, testen van chips onder straling en de kwalificatie van prototype modules in test beams.

Zoektocht naar Samengestelde Standaardmodel Donkere Materie met CMS aan de LHC:

Het Sexaquark-deeltje (S), met uuddss quark-inhoud, is een hypothetisch deeltje dat stabiel kan zijn. Mits de juiste eigenschappen is het deeltje ook een donkerematerie-kandidaat. $\bar{\mathrm{S}}$ deeltjes kunnen geproduceerd worden bij proton-proton botsingen en kunnen nadien annihileren op neutronen in de bundelpijp of detectormateriaal. Deze annihilatiereactie kan aanleiding geven tot de vorming van een $\mathrm{K}_{\mathrm{S}}^{0}$ en een $\bar{\Lambda}^{0}$ die kunnen vervallen naar geladen dochterdeeltjes. Deze geladen deeltjes kunnen gereconstrueerd worden met behulp van de CMS-sporendetector. Dit is het signaal dat gebruikt wordt in deze thesis om te zoeken naar $\bar{\mathrm{S}}$ in de CMS 2016 dataset door de kinematische eigenschappen en de annihilatievertex van de $\bar{\mathrm{S}}$ te reconstrueren. De moeilijkheid van het reconstrueren van een dergelijk signaal met de standaard CMS-reconstructiealgoritmes wordt onderzocht en een limiet op $[\sigma(\mathrm{p}+\mathrm{p}\rightarrow\bar{\mathrm{S}})\times\sigma(\bar{\mathrm{S}}+\mathrm{n}\rightarrow\mathrm{K}_{\mathrm{S}}^{0}+\bar{\Lambda}^{0})]$ wordt afgeleid.

Author's Contribution

The author contributed to the development and use of the test system for CMS Phase-2 Outer Tracker ASICs and module prototypes and was the main developer of a first search for Sexaquarks at the CMS experiment. The contributions of the author are specified at the end of the chapters where the work is described and are summarised briefly below.

The work on the test system consisted of development, testing and maintenance of firmware blocks. The author then contributed to the testing of single chips, ASIC irradiation tests, integration tests and test beams with prototype modules and analysed the results of these tests. This work resulted in the following, two of which are still to be published, papers:

- CMS Tracker Collaboration, "Test beam demonstration of silicon microstrip modules with transverse momentum discrimination for the future CMS tracking detector", JINST 13 (2018) no.03, P03003, (2018-03-06), DOI: 10.1088/1748-0221/13/03/P03003
- CMS Tracker Collaboration, "Beam Test Performance of Prototype Silicon Detectors for the Outer Tracker for the Phase-2 Upgrade of CMS", CERN-CMS-NOTE-2019-006, *submitted to JINST*
- CMS Tracker Collaboration, "Performance of the Prototype CBC3-based Outer Tracker Modules for the Phase II Upgrade of CMS before and after neutron irradiation", *in preparation*

and following proceedings:

- T. Gadek et al., "Quality Control Considerations for the Development of the Front End Hybrid Circuits for the CMS Outer Tracker Upgrade", TWEPP2017, POS proceeding, Volume 313, DOI: `https://doi.org/10.22323/1.313.0061`
- D. Ceresa et al., "Characterization of the MPA prototype, a 65 nm pixel readout ASIC with on-chip quick transverse momentum discrimination capabilities", TWEPP 2018, POS proceeding, Volume 343, DOI: `https://doi.org/10.22323/1.343.0166`
- A. Caratelli et al., "Characterization of the first prototype of the Silicon-Strip readout ASIC (SSA) for the CMS Outer-Tracker phase-2 upgrade", TWEPP2018, POS proceeding, Volume 343, DOI: `https://doi.org/10.22323/1.343.0159`
- A. Caratelli et al.,"Low-power SEE hardening techniques and error rate evaluation in 65nm readout ASICs", TWEPP2019, POS proceeding, *submitted*
- B. Nodari et al., "First results of CIC data aggregation ASIC for the future CMS tracker", TWEPP2019, POS proceeding, *submitted*
- M. Kovacs et al., "A High Throughput Production Scale Front-End Hybrid Test System for the CMS Phase-2 Tracker Upgrade", TWEPP2019, POS proceeding, *submitted*

- J. De Clercq et al., "OT-μDTC, a test bench for testing CMS Outer Tracker Phase-2 module prototypes", EPS-HEP2019, POS proceeding, *submitted*

The work on this test system was recognized by the CMS collaboration by awarding a 2018 CMS detector award to the author for *"Significant contributions to the baseline DAQ for the Phase-2 Outer Tracker, and for µDTC firmware design and implementation"*.

The author was the main developer for the Sexaquark search. Extracting a limit on $[\sigma(\mathrm{pp} \rightarrow \bar{\mathrm{S}}) \times \sigma(\bar{\mathrm{S}} + \mathrm{n} \rightarrow \mathrm{K}^0_{\mathrm{S}} + \bar{\Lambda}^0)]$ required setting up the signal reconstruction algorithm using dedicated vertexing and kinematic fitting techniques, implementing the simulation chain for this unique signal in the CMS simulation framework, including a non-standard encoding of its interaction using the GEANT toolkit, evaluating systematic uncertainties, investigating the background and using multivariate discrimination techniques to suppress it. Throughout this, several obstacles were tackled in searching for the Sexaquark signal with the CMS detector. The observations from these in-depth studies are key to guide the next Sexaquark search towards a more competitive result.

"In a distant and second-hand set of dimensions, in an astral plane that was never meant to fly, the curling star-mists waver and part... See... Great A'Tuin the turtle comes, swimming slowly through the interstellar gulf, hydrogen frost on his ponderous limbs, his huge and ancient shell pocked with meteor craters. Through sea-sized eyes that are crusted with rheum and asteroid dust He stares fixedly at the Destination. In a brain bigger than a city, with geological slowness, He thinks only of the Weight. Most of the weight is of course accounted for by Berilia, Tubul, Great T'Phon and Jerakeen, the four giant elephants upon whose broad and startanned shoulders the disc of the World rests, garlanded by the long waterfall at its vast circumference and domed by the baby-blue vault of Heaven. Astropsychology has been, as yet, unable to establish what they think about. The Great Turtle was a mere hypothesis until the day the small and secretive kingdom of Krull, whose rim-most mountains project out over the Rimfall, built a gantry and pulley arrangement at the tip of the most precipitous crag and lowered several observers over the Edge in a quartzwindowed brass vessel to peer through the mist veils. The early astrozoologists, hauled back from their long dangle by enormous teams of slaves, were able to bring back much information about the shape and nature of A'Tuin and the elephants but this did not resolve fundamental questions about the nature and purpose of the universe. For example, what was Atuin's actual sex? This vital question, said the Astrozoologists with mounting authority, would not be answered until a larger and more powerful gantry was constructed for a deep-space vessel. In the meantime they could only speculate about the revealed cosmos. There was, for example, the theory that A'Tuin had come from nowhere and would continue at a uniform crawl, or steady gait, into nowhere, for all time. This theory was popular among academics. An alternative, favoured by those of a religious persuasion, was that A'Tuin was crawling from the Birthplace to the Time of Mating, as were all the stars in the sky which were, obviously, also carried by giant turtles. When they arrived they would briefly and passionately mate, for the first and only time, and from that fiery union new turtles would be born to carry a new pattern of worlds. This was known as the Big Bang hypothesis."

Terry Pratchett
The Colour of Magic

Part I

Setting the Stage

Chapter 1

Preamble

If you do not measure, you do not know and the Universe might as well be existing of giant tortoises...

The world as pictured in the quote at the start of this thesis from Terry Pratchet's novel "The Colour of Magic" is perhaps not physically exactly correct (at least not for our Universe), but it does exemplify the need for fundamental research, which would otherwise make the professions of astrozoologist and astropsychologist not completely redundant.

The Large Hadron Collider (LHC) at CERN (European Organization for Nuclear Research) together with the CMS (Compact Muon Solenoid) experiment allow physicists to study, in a laboratory environment, the laws of physics at a new energy frontier. At the LHC two beams of protons are accelerated in opposite direction and are made to collide at four interaction points along the accelerator ring. At one of these interaction points the CMS experiment is located. By detecting the particles which are created during a collision, already known processes can be studied in more detail or signatures of undiscovered particles might be found. The discovery potential of the LHC was already proven by experimentally confirming the existence of the Higgs boson.

The Higgs particle and its properties fit within current accuracy in the Standard Model of particle physics. Astronomical and cosmological observations, as well as theoretical considerations, hint however at physics beyond the Standard Model. The data taken by the CMS experiment already allowed for putting tight limits on several beyond the Standard Model models, but other searches as well as high precision Standard Model measurements could benefit from a larger dataset. Therefore, the LHC will run at a higher luminosity from 2026 onwards[1]. This requires the experiments, such as CMS, to upgrade their detectors. One of the foreseen upgrades is a full replacement of the CMS Tracker.

Performing better Standard Model measurements or experimentally investigating physics beyond the Standard Model requires a long process to be followed. From a bird's eye view this process consists of following steps: theoretical predictions, experiment design often including detector R&D and prototyping, experiment construction, detector operation, data analysis and hopefully at the end discovery. This thesis essentially splits up into and touches upon two parts of the aforementioned process applied to the CMS experiment:

- Detector prototyping: data acquisition development for, and testing of, CMS Phase-2 Outer Tracker prototype ASICs and module prototypes.
- Data analysis: search for the Sexaquark particle, a Standard Model composite dark matter candidate of which the anti-particle could initiate a signal in the CMS tracker.

[1] This thesis, with the exception of the Summary, adopts the pre-December 2019 LHC schedule with Long Shutdown 3 planned from 2024 until 2026. According to the latest schedule of the LHC, the Long Shutdown 3 is planned from 2025 until 2027.

The above two parts are stand-alone research subjects with as commonality the CMS experiment and more specifically the CMS tracker subdetector. This thesis is therefore built up as follows: the first part is a general introduction to the thesis and contains useful introductory information for both research topics. It gives an introduction to particle physics, detector techniques for particle physics experiments and the CMS experiment. More details are provided on the outstanding problem of dark matter, silicon based detectors and on the CMS tracker subdetector as these topics are particularly relevant for the rest of the work. A good understanding of the full CMS detector is essential to understand the role of the tracker subdetector in the full experiment and the reason for the upgrade of this subdetector. This upgrade is the topic of the second part of the thesis which starts with an introduction on the CMS Phase-2 Outer Tracker upgrade and proceeds with the tests performed on several prototypes. The third, and last part of this thesis, introduces the Sexaquark particle and discusses the search for this particle using 2016 RunG and RunH data from the CMS experiment.

Chapter 2

Introduction to Particle Physics

2.1 Introduction

Elementary particle physics tries to explain the nature and interaction of matter at the smallest scales. Decades of experiments have contributed to the development of the *Standard Model of particle physics* which classifies the elementary particles and describes the fundamental forces which act between them. This chapter gives an overview of the Standard Model (section 2.2), an introduction to its theoretical basis (section 2.3) and a description of one of its shortcomings: explaining the nature of *dark matter* (section 2.4).

2.2 The Standard Model of particle physics

Figure 2.1 shows the building blocks of the Standard Model of particle physics. As far as we know, all known matter is built up of fermions which are particles with half-integer spin. The fermions are divided in leptons and quarks and both of these particle types are arranged into three generations which have increasing masses. Particles belonging to higher generations are heavier and unstable and when they are produced they eventually decay back to particles from the first generation. Each of the fermions has an anti-matter counterpart which has all the charges reversed.

The fermions interact through the exchange of particles with integer spin, the bosons. Each one of the bosons can be linked to a fundamental force. The most familiar boson is the photon (γ) which mediates the electromagnetic force. The $W^{\pm}$ and Z bosons mediate the weak force and are apparent for example in nuclear decays. The gluon (g) is linked to the strong force, which e.g. binds quarks in nuclei and keeps protons and neutrons together in the atomic nucleus. The fourth fundamental force, gravity, has not yet been described in the Standard Model.

There are six different *flavours* of quarks in the Standard Model: up, down, charm, strange, top and bottom. Quarks have a *colour* charge and they cannot exist isolated, a property of the strong interaction called confinement, which only allows colour neutral states to propagate freely. A whole spectrum of bound states of quarks, so-called hadrons, has been discovered. Hadrons built up from an even number of *valence quarks*[1] (e.g. K^0_S) are referred to as mesons whilst particles consisting of an odd number of valence quarks (e.g. Λ^0) are referred to as baryons. Next to a colour charge, the quarks have a weak charge (*flavour*) and a fractional electric charge which makes them also interact through the electroweak interaction.

Leptons on the other hand are not subject to the strong interaction. Leptons come in pairs: each charged lepton is associated with a neutrino. Neutrinos are massless in the Standard Model. They also do not have an electric charge and therefore only interact

[1] The quarks and anti-quarks which give rise to the quantum numbers of the hadron.

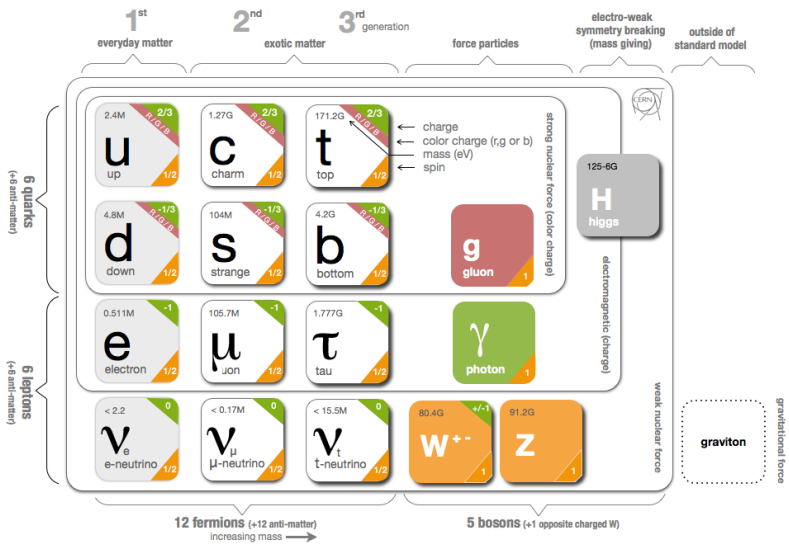


FIGURE 2.1: Elementary particles of the Standard Model [1].

through the weak interaction which makes them notoriously difficult to detect experimentally.

Fermions and massive bosons get their mass through interaction with the Brout-Englert-Higgs (BEH) field [2, 3], or in short Higgs field, of which the Higgs bosons are the excitations. The BEH mechanism will be discussed in more detail in section 2.3. The interaction between the fundamental particles is described using a Quantum Field Theory (QFT), which is the subject of the next section.

2.3 The mathematical framework of the Standard Model

The interactions between elementary particles happen at the smallest distance scales with energies typically higher or comparable to the particle's mass. Using a QFT approach, both quantum mechanics and relativity are included to describe the particle interactions. The solutions of the relativistic Schrödinger equations in the QFT are fields and the quantization of these fields are linked to particles in the Standard Model. An extensive discussion on QFT and the Standard Model can be found for example in Ref. [4].

The QFT description starts with a Lagrangian ($\mathcal{L}$), describing the field, from which the equation of motion can be extracted by minimising the action ($\mathcal{S}$) defined as:

$$\mathcal{S} = \int \mathcal{L}(x)\, d^4x, \tag{2.1}$$

where x is the space-time coordinate. The Standard Model Lagrangian comprises specific terms for each of the fundamental interactions.

The Lagrangian of a free fermion contains a kinetic and a mass term:

$$\mathcal{L}_{Dirac} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi, \tag{2.2}$$

where the Einstein convention is adopted of summation over repeated indices, m is the fermion's mass, γ^μ are the Dirac matrices ($\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}$ with $g^{\mu\nu}$ the Minkowski metric) and ψ and $\bar{\psi}$ ($= \psi^\dagger\gamma^0$) represent the fermion and anti-fermion fields respectively. The modulus $|\psi|^2$ as well as the Lagrangian are required[2] to be invariant under local phase transformations:

$$\psi \rightarrow \psi' = U(x)\psi = e^{i\vec{\alpha}(x)\cdot\frac{\vec{\tau}}{2}}\psi, \tag{2.3}$$

with $\vec{\tau}$ the generators of a Lie group and $\vec{\alpha}(x)$ rotation parameters. To get the first term in the Lagrangian in Equation (2.2) invariant under this gauge symmetry, the derivative can be replaced by its covariant form:

$$D_\mu = \partial_\mu + ig\frac{\vec{\tau}}{2}\cdot\vec{A}_\mu, \tag{2.4}$$

where $\vec{A}_\mu$ is a vector field which couples to the fermions with a coupling strength g and which transforms according to $A_\mu \rightarrow A'_\mu = A_\mu - \partial_\mu\alpha$. Combining Equation (2.2) and (2.4) gives a new Lagrangian:

$$\mathcal{L}_{Dirac} = i\bar{\psi}\gamma^\mu D_\mu\psi - m\bar{\psi}\psi, \tag{2.5}$$

which is indeed invariant ($\mathcal{L}_\psi = \mathcal{L}_{\psi'}$) under the local gauge symmetry (Equation (2.3)). The above principle naturally gives rise to a new field ($\vec{A}_\mu$) which is the gauge field associated to the gauge transformation which was performed. This field allows information to be propagated from one matter field to another.

The same principle can be applied using other symmetry groups than the one used in Equation (2.3) and can be extended to the full Standard Model Lagrangian which is invariant under the symmetry group[3] G_{SM}:

$$G_{SM} = SU(3) \times SU(2) \times U(1). \tag{2.6}$$

$SU(3)$ represents the symmetry of the theory describing the strong force, quantum chromodynamics (QCD) with coupling strength g_s, whilst $SU(2) \times U(1)$ describes the symmetry of the electroweak theory which combines the weak and the electromagnetic interaction. The covariant derivative leaving the full Standard Model Lagrangian invariant under the transformation G_{SM} is:

$$D_\mu = \partial_\mu + ig_s\frac{\lambda_a}{2}G^a_\mu + ig\frac{\sigma_i}{2}W^i_\mu + ig'\frac{Y}{2}B_\mu, \tag{2.7}$$

where the second term represents the strong force and its $SU(3)$ symmetry, with the Gell-Man matrices λ_a (a = 1...8) as generators. It introduces eight gluon fields of which the massless gluons are excitations. The third and fourth term are associated to the electroweak interaction. The electroweak $SU(2) \times U(1)$ symmetry group, with generators σ_i (i=1...3), the Pauli matrices, and Y the hypercharge, introduces 3 gauge fields W^α_μ and one gauge field B_μ. The electroweak coupling strengths g and g' are linked through the weak

[2]The argument for introducing this requirement is that if there is no way for fields to communicate through space and time, a local change in their phase should not change the physics of the problem.

[3]$SU(n)$ are the groups of $n \times n$ unitary matrices of determinant 1. The dimension of a group is given by n^2 for $U(n)$ groups and n^2-1 for $SU(n)$.

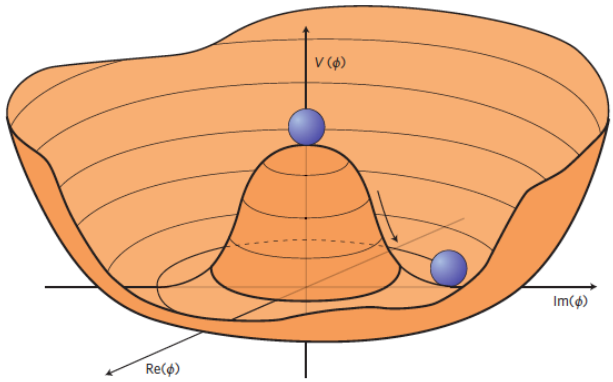


FIGURE 2.2: Mexican hat potential [5].

mixing angle

$$\tan\theta_W = \frac{g'}{g} \tag{2.8}$$

and the physically observed bosons W_μ^-, W_μ^+, Z_μ and A_μ (the photon) are formed by mixing the W_μ^i fields, associated to the $SU(2)$ symmetry, and the B_μ field, associated to the $U(1)$ symmetry:

$$W_\mu^\pm = \sqrt{\frac{1}{2}}\left(W_\mu^1 \mp iW_\mu^2\right) \tag{2.9}$$

$$Z_\mu = W_\mu^3 \cos\theta_W - B_\mu \sin\theta_W \tag{2.10}$$

$$A_\mu = W_\mu^3 \sin\theta_W + B_\mu \cos\theta_W. \tag{2.11}$$

The above theory is however unable to assign a non-zero mass to the gauge bosons without breaking the gauge invariance. The mechanism introduced to explain the particle masses without breaking the gauge invariance of the Standard Model Lagrangian is the Brout-Englert-Higgs mechanism. This mechanism introduces a complex scalar doublet (ϕ) which has a non-zero vacuum expectation value v. The Higgs field Lagrangian is:

$$\mathcal{L}_H = \left(D^\mu\phi\right)^\dagger \left(D_\mu\phi\right) - V(\phi), \tag{2.12}$$

with the potential $V(\phi)$:

$$V(\phi) = \mu^2\phi^\dagger\phi - \lambda\left(\phi^\dagger\phi\right)^2, \tag{2.13}$$

where λ represents the self coupling strength and μ is a mass parameter. If $\mu^2 < 0$ and $\lambda < 0$ this potential is known as the *Mexican-hat potential* which is depicted in Figure 2.2.

The BEH mechanism allows for the three observable massive weak gauge bosons to acquire a mass whilst the photon can remain massless. The excitations of the scalar field ϕ introduced here are the Higgs bosons. Fermions acquire mass by the addition of *Yukawa* terms, of the form $g_y\bar{\psi}\phi\psi$, to the Standard Model Lagrangian where g_y represents the coupling strength of the scalar field to the fermion.

The Standard Model as introduced in the previous section and its mathematical basis depicted in this section describe high energy physics experiments to high accuracy.

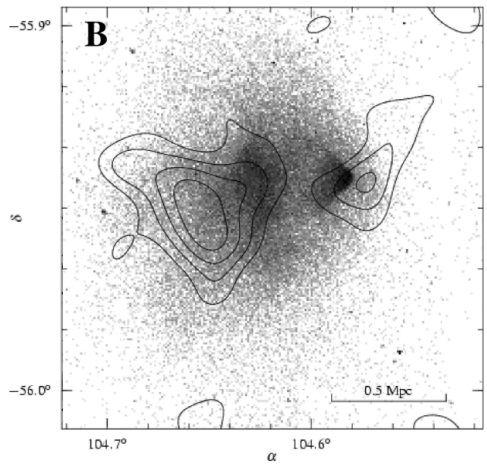


FIGURE 2.3: Image of 1E0657-558 [7]. The black contours are the mass contours extracted from weak lensing. Each contour represents a change in the surface mass density of 2.8×10^8 $M_\odot/kpc^2$. The grey-scale is the X-ray image of the hot gas obtained from the Chandra Observatory. The so-called bullet cluster is visible on the right in the grey-scale.

Experimental observations however suggest that extensions of the Standard Model are needed as the Standard Model for example does not explain the difference in the electroweak and gravitational scale (*hierarchy problem*), the neutrino masses, the matter-antimatter asymmetry in the Universe nor does it predict the existence of dark matter. The latter will be discussed in more detail in the next section as the dark matter problem is relevant for the search described in Chapter 10.

2.4 Dark matter

2.4.1 Introduction

The nature of dark matter, together with e.g. baryon asymmetry, the hierarchy problem and the neutrino masses is one of the main mysteries in particle physics. All of these do not find a place in the Standard Model. Dark matter does not appear to interact via the electromagnetic force and thus would not absorb, reflect or emit any light, hence its name. Until today, only gravitational evidence for the existence of dark matter is available and this through astronomical and cosmological observations. Measurements show that our Universe is built up of roughly 5% known matter, $\approx$ 27% dark matter and $\approx$ 68% dark energy [6].

2.4.2 Evidence for dark matter

A particularly instructive example of evidence for the existence of dark matter is 1E 0657-558 [7], also known as *the bullet cluster*, which is shown in Figure 2.3. In this figure the grey-scale shows the mass distribution derived from X-ray measurements of hot gas,

which represents the majority of the baryonic matter in the cluster. It reveals the shape of a large galaxy cluster on the left and a smaller cluster which demonstrates a shock-wave on the right. The right cluster obtained this specific shape (a bullet) by passing through the large cluster. Using the effect of gravitational lensing, the total mass distribution of both clusters can be extracted. The obtained mass distributions are shown as the contours in Figure 2.3. Comparing the distribution of the hot gas with the total mass distribution shows a clear offset. This is a smoking gun for the existence of dark matter: the gas in the clusters is slowed down due to frictional electromagnetic forces. The dark matter on the other hand is not affected by this friction and travels further on its path, in this way splitting off from the gas. Similar observations were done in other colliding galaxy clusters such as MACS J0025.4-1222 [8].

These colliding clusters are historically not the first evidence for dark matter. It was F. Zwicky whom in 1933 measured the velocity of galaxies in clusters and found that the visible mass in the clusters could not deliver the gravitational pull to keep the galaxies on their orbit [9]. He therefore conjectured that additional matter needs to be present in the cluster to deliver the extra mass. This type of matter he called *dark matter*. Similar observations can be done on a smaller scale by looking at stars within a galaxy.

Besides colliding galaxy clusters and galaxy rotation curves, evidence for dark matter was also found in the Cosmic Microwave Background (CMB) [10, 11]. The CMB is a remnant from the photon decoupling which happened when the Universe was about 379,000 years old. The decoupling took place when the energy density after the Big Bang dropped below the energy threshold for the photons to ionize hydrogen and maintain a hydrogen plasma. Hydrogen atoms could now stay combined and photons could move freely. These photons can still be observed today, although as less energetic ones due to the expansion of the Universe. This relic radiation, having a thermal black body spectrum, can be observed using radio telescopes. Small perturbations in the CMB, measured over the full sky, reflect, amongst others, information on the small anisotropies in matter density at the time of decoupling.

One of the most important results of the study of the CMB are the measurements of the mass density of several types of matter in our Universe. The Planck collaboration for example quotes [6] a dark matter density and baryon density of respectively $\Omega_c h^2 = 0.120 \pm 0.001$ and $\Omega_b h^2 = 0.0224 \pm 0.0001$. Here $\Omega_X = \rho_X / \rho_{crit}$ for a component X with ρ_{crit} the critical mass density which means that the total mass density $\Omega_{tot} = 1$ corresponds to a flat Universe and h is the Hubble constant in units of 100 km/(s.Mpc).

The observation of old galaxies at a redshift of ≈ 10 together with the small density perturbations seen in the CMB also require extra matter, besides the visible component, in order for these old galaxies to could have formed [10]. Without dark matter the density perturbations would not start to grow into galaxies at such an early stage. Furthermore this dark matter should be *cold* which means that it should have been non-relativistic at the start of galaxy formation. This implies for example that neutrinos cannot be the dark matter: neutrinos were kept in equilibrium with photons and electrons through $\gamma \leftrightarrow e^+ + e^- \leftrightarrow \nu_i + \bar{\nu}_i$. This reaction stopped at around 3 MeV, leaving behind a relic relativistic population of neutrinos. An upper bound on the contribution of light neutrinos to the dark matter can be extracted from analysis of structure formation: $\Omega_\nu h^2 \leq 0.0062$ (at 95% CL) [10].

2.4.3 Dark matter candidates

Many candidates for dark matter have already been proposed. Some of them are hard to detect Standard Model particles, but most are completely new exotic particles. The constraints on dark matter candidates are numerous. M. Taoso, G. Bertone and A. Masiero summarise in Ref. [12] a 10-point test for a dark matter candidate:

1. Does it match the appropriate relic density? Non-relativistic dark matter, first being in thermodynamic equilibrium with the plasma in the early Universe freezes-out when its interaction rate drops below the expansion rate of the Universe. The remaining relic density today can be calculated for a given dark matter candidate and should be within the bounds of the limits set by CMB measurements.

2. Is it cold? This constraint was already explained in the previous section. Only a small part of the dark matter content of the Universe is allowed to be hot dark matter in order not to jeopardize the early onset of galaxy formation.

3. Is it neutral? Most dark matter candidates have to be neutral to explain the fact that they do not interact through the electromagnetic interaction. Only in specific models the dark matter candidate is allowed to have an electric charge. These models will not be discussed here.

4. Is it consistent with Big Bang nucleosynthesis? A set of Boltzmann equations can explain the measured abundances of light nuclei in the context of Big Bang nucleosynthesis (BBN) to very high accuracy. A dark matter candidate should therefore not introduce stress on the BBN model which is described very well by Standard Model processes. E.g. introducing dark matter which in decays would generate photons which in turn could alter the formation process of light nuclei would be a direct constraint on this type of dark matter.

5. Does it leave stellar evolution unchanged? Dark matter candidates can be produced in the center of stars and if the candidates only interact weakly they can furthermore leave the star. This would represent an energy loss channel which could possibly alter the evolution of the star. Several measurements, for example the measurement of the neutrino flux, neutrino energy and the duration of supernova SN1987A put limits on these additional channels.

6. Is it compatible with constraints on self-interactions? Multiple observations put constraints on the self-interaction cross section of dark matter particles. One example of these measurements are the colliding clusters described in the previous section from which an upper limit on the dark matter self-interaction can be extracted. Another observation is the *core-cusp* problem, which is the discrepancy seen in material distribution between the cuspy dark matter haloes predicted in cold dark matter N-body simulations and the actual observed distribution in e.g. dwarf galaxies. Self-interacting dark matter could alleviate this problem by smoothing the cuspy profile. Recent measurements [13] show however that the core-cusp problem can also be resolved by taking into account the effect of bursty star formation which can kinematically heat up dark matter at the center of galaxies through fluctuations in the gravitational potential.

7. Is it consistent with direct DM searches? As will be discussed in section 2.4.4, direct detection of dark matter through detection of nuclear recoils has put several constraints on the dark matter interaction cross section versus mass of the dark matter particle.

8. Is it compatible with gamma-ray constraints? As will be discussed in section 2.4.4, another way to look for dark matter candidates is by detection of their annihilation products. A good candidate product to look for are high energy photons. For mass ranges of the dark matter candidate in the GeV-TeV region this results in γ-rays. Non-observation of clear evidence of dark matter annihilation can put several constraints on dark matter candidates.

9. Is it compatible with other astrophysical bounds? Besides photons, also neutrinos can be formed directly or indirectly in annihilation of the dark matter particles. These neutrinos can subsequently be detected in high-energy neutrino telescopes through the detection of Cherenkov light from secondary muons produced when the neutrino interacts with the material. Other messengers such as positrons, antiprotons or X-rays could also give more information and put further limits on dark matter models.

10. Can it be probed experimentally? This is not a fundamental requirement for a good dark matter candidate, but it is a requirement to affirm the hypothesis.

A few candidates for dark matter were already mentioned above. Two other interesting examples of possible candidates are:

- Massive astrophysical compact halo objects (MACHOs): these are a natural solution to the dark matter problem as they do not require introducing any exotic particle. MACHOs are astronomical bodies composed of normal baryonic matter with typical masses of 0.001–0.1 solar masses. These objects emit only little or no radiation at all. Examples of MACHOs are (primordial) black holes and neutron stars. Major sky surveys have been set up to discover these objects through microlensing[4] without positive result [11].

- Weakly interacting massive particles (WIMPs): WIMPs are particles with masses roughly between 10 GeV/c^2 and a few TeV/c^2 and with cross sections comparable to the weak interaction. WIMPs freeze out once the rate of reactions which change WIMPs to Standard Model particles and vice versa becomes smaller than the Hubble expansion rate of the Universe. It is rather remarkable that when WIMPs are requested to explain the energy density of dark matter one naturally arrives at an interaction cross section which is in the order of magnitude range of the weak interaction. WIMPs therefore form a preferred solution to the dark matter problem. An example of a WIMP is the lightest superparticle in supersymmetric models with R-parity conservation [11].

It is of course not required that dark matter consists out of a single type of particle. Therefore e.g. combinations of the different possibilities listed above are also possible.

[4] If the object acting as a lens has a small mass (compared to normal lensing where galaxies or clusters are the lensing object) it is not possible to distinguish multiple images of the lensed object. An amplification of the light can however be measured. MACHO events can be distinguished from variable stars using the fact that the amplification is the same for red and blue light.

Furthermore it is possible that dark matter is actually not required to explain the observations discussed in section 2.4.2. These observations rely on our current understanding of physics which might not be applicable at cosmological scales. Alternative theories can be postulated which rely on modifications to the laws of gravity to explain for example the rotation curves of stars in galaxies. Indeed, there is a valid basis for such theories as the acceleration of stars in the outer radius of galaxies is extremely small. This is a regime in which Newtonian dynamics has never been tested. It proves difficult however to reconcile all observations, in particular the CMB measurements, in such theories.

2.4.4 Experimental searches for the nature of dark matter

Besides specific experiments designed to detect a certain type of dark matter, such as experiments looking for MACHOs, most dark matter experiments can be categorised into one of the following three types:

- Direct detection: the dark matter particles present in our galaxy are expected to interact with Standard Model particles through e.g. nuclear recoil interactions. The recoil nuclei can be recorded for example by detecting emitted scintillation light, induced charge or phonons.
- Indirect detection: dark matter particles can annihilate on their anti-matter partner and can result in e.g. γ-rays or neutrinos which can be detected with dedicated experiments. Regions with high mass density, such as the galactic center or the center of the Earth and the Sun are particularly interesting to study. Detection of γ-rays is most conveniently done in space based observatories because γ-rays do not penetrate the atmosphere (e^+e^- pair production) to reach ground based telescopes. Indirect detection of γ-rays is however possible using ground based telescopes which are sensitive to the Cherenkov light and secondary particles in a cosmic shower initiated by a γ-ray. As discussed in the previous section, also neutrinos, positrons, anti-protons, anti-nuclei, X-rays and measurements of radio emission can be used to either do independent searches or to complement measurements in any other channel.
- Production at particle colliders: dark matter particles can also be produced at collider experiments through the collision of two Standard Model particles. The produced dark matter particle could pass through the detector without leaving a signal, but it can be produced in association with other Standard Model particles. In a hermetic detector, the presence of the dark matter particle can then be inferred from the momentum imbalance in the total transverse[5] momentum.

2.5 Summary

The Standard Model of particle physics and its mathematical foundations which can be found in Quantum Field Theory describe particle physics phenomena to high accuracy. The known matter around us is built up of fermions which interact through the exchange of bosons. The mass of the bosons and fermions can be explained by interaction with the Higgs field. Astronomical, cosmological and theoretical considerations point at the existence of physics beyond the Standard Model. One of the outstanding problems is the

[5]This is the plane transverse to the beam direction.

nature of dark matter. The evidence for dark matter is eclectic: colliding galaxies, rotation curves of stars in galaxies, CMB measurements,... The fact that dark matter is needed to explain these observations also implies that a dark matter model is constrained from many angles. Particle accelerators, such as the LHC, are a good platform to test several of these theories through the potential production of dark matter particles in high energy collisions.

Chapter 3

Detector Techniques in Particle Physics

3.1 Introduction

A good understanding of particle interactions with matter is required to both understand particle detection techniques and the radiation damage which these interactions can induce on detector components. An overview of particle interactions with matter is therefore given in section 3.2. The actual detection of charged particles will then be discussed in section 3.3 and the usage of silicon sensors for charged particle tracking detectors is described in section 3.4, which includes a brief summary on the effects of radiation on silicon sensors. An introduction to the basics on integrated circuits is provided in section 3.5 and more detail is given on readout ASICs (Application-Specific Integrated Circuits) for silicon tracking detectors in section 3.6. The effects of radiation on ICs (Integrated Circuits) is discussed in section 3.7.

3.2 Particle interactions with matter

Particles can interact in many different ways with matter. These interactions form the basis of particle detection. The particle interaction with matter depends on the type of particle which interacts. For the following discussion four particle categories are identified:

- Heavy charged particles
- Electrons
- Photons
- Heavy neutral particles

where the adjective *heavy* is used to distinguish other charged particles and neutral particles from electrons and photons respectively. The following sections will give a brief overview [14] of all four categories, going a bit more into depth on heavy charged particles as these are of importance for the rest of the work.

3.2.1 Heavy charged particle interactions with matter

Heavy charged particles interact through the Coulomb force with the orbital electrons from the material. As a result of these interactions they excite orbital electrons or electron-ion pairs are formed. In several particle detectors the latter forms the basis for the detector response. In interactions where a lot of energy is transferred to the orbital electron, the orbital electron can gain enough energy to also create secondary ionization. In those cases the electrons are referred to as *δ-rays*. Interactions of the incoming particle with

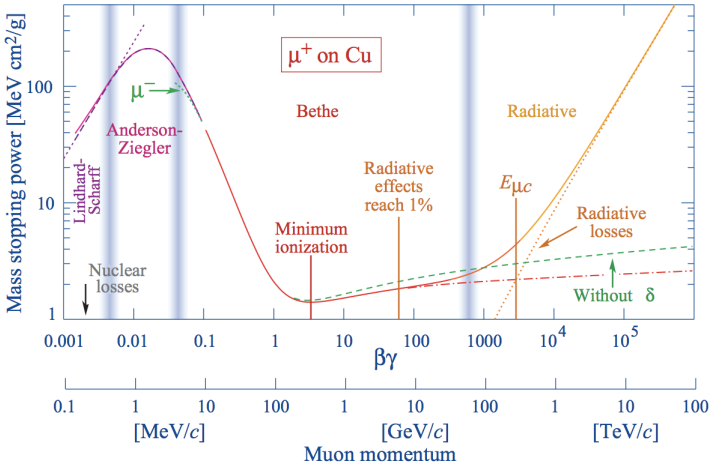


FIGURE 3.1: LET divided by the material density for positive muons in copper as a function of the muon momentum and $\beta\gamma$ ($=p/(Mc)$). In the central region, the Bethe-Bloch approximation is valid. For lower ($\beta\gamma \lessapprox 0.1$) and higher ($1000 \lessapprox \beta\gamma$) momenta other effects need to be taken into account [10].

the nuclei, such as Rutherford scattering, are also possible but are less probable than ionization interactions and are normally not exploited in radiation detectors.

The Linear Energy Transfer (LET):

$$LET = -\frac{dE}{dx} \tag{3.1}$$

describes the energy loss (dE) over a path length (dx). The LET for a muon is depicted in Figure 3.1 for nine orders of magnitude of the muon's energy.

The average LET for particles with $0.1 \lessapprox \beta\gamma \lessapprox 1000$ (with $\beta = v/c$ and $\gamma = 1/\sqrt{1-\beta^2}$) can be calculated using the relativistic Bethe-Bloch formula [14]:

$$-\left\langle \frac{dE}{dx} \right\rangle = \frac{4\pi}{m_e c^2} \cdot \frac{nz^2}{\beta^2} \cdot \left(\frac{e^2}{4\pi\varepsilon_0} \right)^2 \cdot \left[\ln\left(\frac{2m_e c^2 \beta^2}{I \cdot (1-\beta^2)} \right) - \beta^2 \right], \tag{3.2}$$

with

$$n = \frac{N_A \cdot Z \cdot \rho}{A \cdot M_u}, \tag{3.3}$$

where v is the speed of the particle, z its charge, E its energy, n, I and ρ respectively the electron density, the mean ionization potential and density of the material, c the speed of light, ϵ_0 the vacuum permittivity and e and m_e the charge and the mass of the electron, Z and A respectively the atomic number and the relative atomic mass of the material, N_A the Avogadro number and M_u the molar mass constant. Important to note is that the LET scales quadratically with the charge of the incident particle and increases with the electron density of the absorber material. The LET reaches a minimum around $\beta\gamma = 3$. Particles with this energy are called Minimum Ionizing Particles (MIPs). At low energies ($\beta\gamma \lessapprox 0.1$) the above equation is not valid any more and correction terms have to be

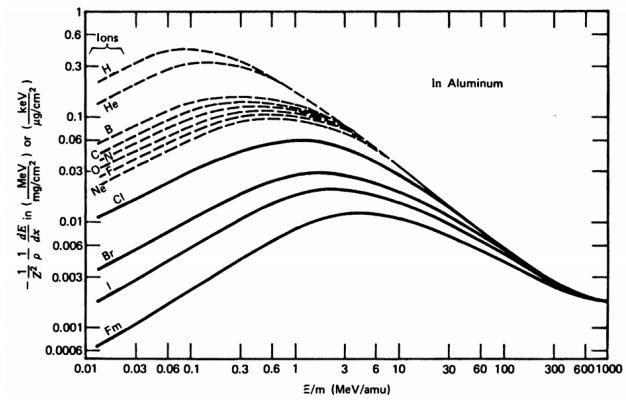


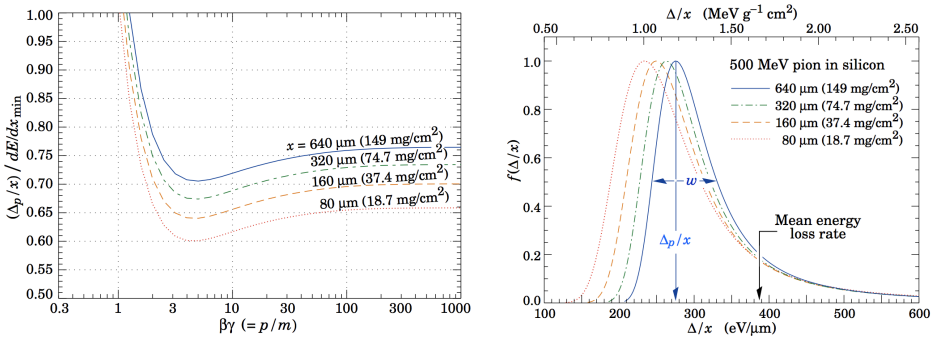
FIGURE 3.2: The energy loss for several heavy ions passing through aluminium. ρ is the density of aluminium, Z the ion's atomic number, and m is the mass of the ion expressed in atomic mass units [14].

added [10] to take into account atomic binding of electrons and the fact that the incident particle starts to carry atomic electrons with it. The latter reduces the incoming particle's effective charge and thus LET. The larger the charge of the incoming particle the more important this effect becomes and it becomes apparent at higher energies as shown in Figure 3.2. At high energies ($1000 \lessapprox \beta\gamma$) radiative losses, e.g. due to bremsstrahlung, need to be taken into account.

It is interesting to look at a few values and distributions for one specific material: silicon. Silicon is used as detection material, and is furthermore the main component of electronics. A good understanding of the energy deposition in silicon is thus important to understand both detection mechanisms and radiation damage.

For high energy physics the tracking of charged particles with silicon detectors relies on the use of thin layers of silicon in which the particle deposits energy due to ionization, leaving behind electron-hole (e^-h^+) pairs. The average mean energy loss of a MIP in silicon is 388 eV/μm but corrections have to be taken into account for thin layers from which part of the secondaries can escape the volume [15]. These correction terms together with the evolution of the LET with the particle's energy are illustrated in Figure 3.3a. For a silicon sensor with a typical thickness of ≈ 300 μm this gives a most probable value of 76 e^-h^+/μm for the number of electron-hole pairs formed along the path of the MIP (3.6 eV is required to generate an e^-h^+ pair in silicon). The distribution of the deposited energy by a MIP in silicon is illustrated in Figure 3.3b and shows the typical skewed Landau distribution where the upper tails are due to δ-rays.

Low energy heavy ions are an important particle type to consider for the effects they can have on the operation of ICs (section 3.7.3). Low energy heavy ions are produced in material by an incoming particle, mostly neutrons, protons or pions, undergoing a nuclear interaction and producing a shower of particles and a nuclear recoil. The nuclear recoil is a heavy ion and typically has an energy below 10 MeV. Figure 3.4a shows the LET versus energy for a silicon and a magnesium recoil and shows that a silicon recoil with an energy of 10 MeV deposits $\mathcal{O}(2\ \text{MeV}/\mu\text{m})$, which is far above the $\mathcal{O}(100\ \text{eV}/\mu\text{m})$ range for MIPs. As a result of this high LET the range of these heavy ions is short as shown in Figure 3.4b, where the range is the penetration depth of the particle in the material before coming to rest. Figure 3.4b shows that silicon ions with an energy below 10 MeV have a



(a) (b)

FIGURE 3.3: LEFT: Most probable energy loss in silicon layers of different thickness, scaled to the mean energy loss at minimum ionization (388 eV/μm) [10]. RIGHT: Normalized distribution of the energy deposit per μm for a 500 MeV/c pion in silicon layers of different thicknesses [10].

range of less than 10 μm in silicon.

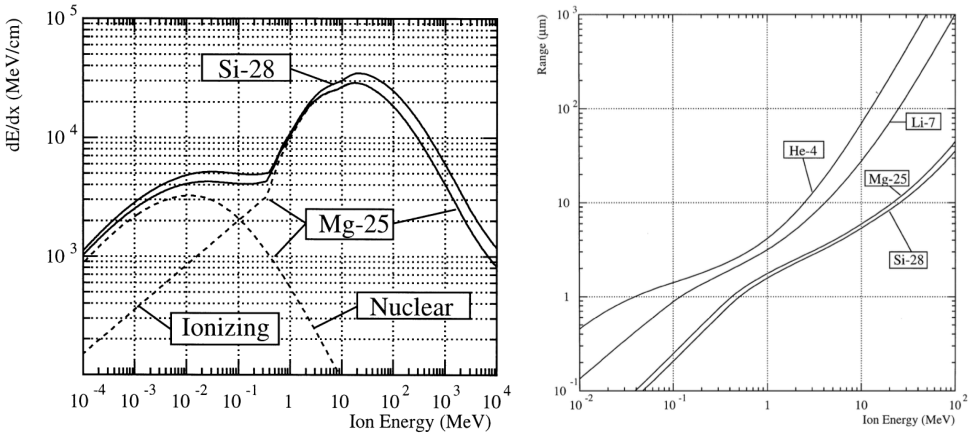
A last important type of heavy charged particles to consider are muons. Muons do not interact through the strong interaction, they have a much higher mass than electrons which makes them lose only a small fraction of their energy in ionization and radiative energy losses are furthermore suppressed due to their high mass. As a result high energy muons can have ranges of many meters through dense material.

3.2.2 Electron interactions with matter

Due to the low mass of the electron compared to the particles discussed in the previous section the electron requires some special attention. The electron interacts with orbital electrons when it enters the material and therefore large fractions of its energy can be lost and its direction largely changed by single encounters. Besides interactions with orbital electrons the impinging electrons can also lose energy through interactions with nuclei and by radiative processes such as bremsstrahlung. For heavy charged particles, radiative corrections are only of importance at the highest energies ($\beta\gamma \approx 1000$). For high energy electrons (E $> \mathcal{O}(10$ MeV$)$) however, radiative processes are the main reason for energy loss. The mean distance over which a high-energy electron loses all but 1/e of its energy by bremsstrahlung is referred to as the *radiation length* (X_0). For detection techniques, ionizing interactions of electrons play an important role. In silicon, for electrons with γ = 3.3 (p = 1.6 MeV/c), the electron loses 348 eV/μm due to ionization, which is close to the value for heavy MIPs [10].

3.2.3 Photon interactions with matter

Interaction of photons with material are dominated by three mechanisms: photoelectric absorption, Compton scattering and pair production [14]. In all of these mechanisms the photon's energy is conveyed (partially) to electrons. For high energy photons pair production is by far the most important mechanism for energy loss. The *radiation length* for high energy photons is defined by: 7/9 of the mean free path for pair production.



(a) (b)

FIGURE 3.4: LEFT: LET for silicon and magnesium ions in silicon [16]. RIGHT: Ranges of some ion types in silicon [16].

3.2.4 Heavy neutral particle interactions with matter

The interaction of heavy neutral particles with matter requires the incoming particle to interact with the nucleus of the absorbing material [14]. The secondary radiation produced when a neutral particle interacts will mostly consist of heavy charged particles, such as heavy ions, whose material interactions were discussed in section 3.2.1. Inelastic scattering of the neutral particle with the nucleus on the other hand can elevate the nucleus to a higher energy state from which it decays by emitting a γ-ray. At low energies, neutrons can induce (n,γ), (n,α), (n,p) and (n,fission) reactions.

3.3 Charged particle detector techniques

The type of detector to use for a certain application largely depends on the type of particle which needs to be detected and the properties of the particle which need to be determined. Experiments can be designed to detect neutrinos, neutral heavy particles, charged particles, photons or a combination of any of these. Throughout the years many different particle detectors have been designed and constructed to detect specific properties of any of the above particle types.

Table 3.1 summarises a few important properties of a selection of charged particle detectors. The values listed in this table are only order of magnitude approximations and of course depend on the exact detector implementation [10]. All the detectors listed in Table 3.1 share the same basic detection principle: a sensitive volume is present which the incoming particles locally ionize. This sensitive volume can be a gas, liquid or solid. The liberated charge is collected or induces a charge on a nearby electrode which is connected to the readout electronics[1].

Table 3.1 shows that detectors based on silicon have very good spatial and time resolution. This type of detector will be discussed in more detail in the next sections.

[1]With the exception of the bubble chambers and emulsion techniques where recording of the charged particle's track is performed with cameras or microscopy.

Detector type	Intrinsic spatial resolution	Time resolution	Dead Time
Single gap resistive plate chambers	$\lesssim 10$ mm	1 ns	-
Scintillation tracker	≈ 100 μm	100 ps/(refraction index)	10 ns
Micro-pattern gas detectors	30-40 μm	< 10 ns	10-100 ns
Silicon strip (with digital readout)	strip pitch/$\sqrt{(12)}$	few ns	$\lesssim 50$ ns
Silicon pixel	$\lesssim 10$ μm	few ns	$\lesssim 50$ ns
Bubble chamber	10-150 μm	1 ms	50 ms
Emulsion	1 μm	-	-

TABLE 3.1: Typical properties of charged particle detectors [10].

3.4 Silicon based detectors for tracking

3.4.1 Introduction

Silicon, with the correct doping concentration, is a very convenient detector medium. A few reasons why silicon is such a useful material are [15]:

- Only 3.6 eV of energy, compared to e.g. 30 eV in gas, is required to generate an electron-hole pair in silicon. Together with an increased mass density compared to gas, the same particle generates more charge over a given path length which allows silicon detector modules to be compact. Important to note here is that in most configurations silicon detectors for tracking purposes do not make use of charge multiplication as opposed to many gaseous detectors.
- All commonplace electronics are based on silicon technology. As a result companies are available which have the necessary equipment for large scale production of wafers with high density μm-size features resulting in a fine spatial granularity of detector channels.
- With the appropriate design, silicon sensors can be made radiation hard which makes them an excellent detector material to be used in high radiation environments e.g. close to the interaction point in hadron colliders.

In the following section, basics of silicon as a detector material will be introduced. In section 3.4.3 the layout and operating principle of a position sensitive silicon sensor will be described followed by a brief discussion on radiation damage in silicon sensors in section 3.4.4.

3.4.2 Silicon doping and the pn-junction

The electrical potential created by atomic nuclei in solids limits the electrons to only occupy certain energy levels. These levels are separated by gaps in which no available energy levels are available for the electrons [15]. The discrete states of allowed energy levels are very dense and can therefore be interpreted as energy *bands*. In insulators and semiconductors two energy bands can be identified: the valence and the conduction band. The *valence band* comprises the energy levels in which electrons are normally present at 0 Kelvin. The *conduction band* comprises the lowest range of free electronic states. The valence and conduction band are separated by the *bandgap*. Materials with a bandgap larger than approximately 4 eV are insulators, materials with a bandgap in the range of

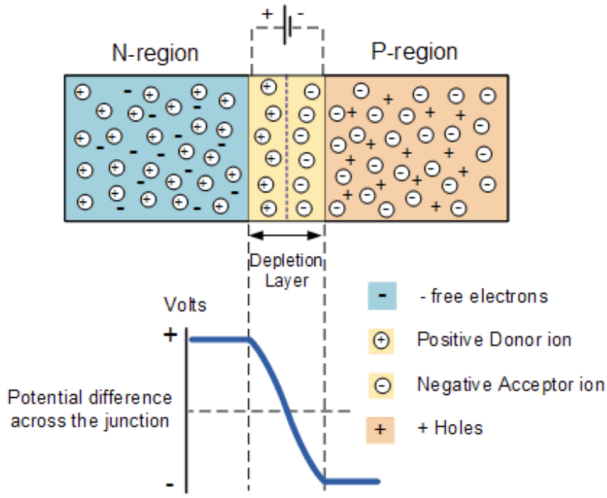


FIGURE 3.5: The pn-junction [17].

≈ 1 to ≈ 4 eV are semiconductors and materials without a bandgap are metals. The probability of finding an electron in a certain energy state is given by Fermi-Dirac statistics. In semiconductors, as opposed to insulators, electrons can occupy the conduction band at room temperature.

Silicon is by far the most used semiconductor material for electronics applications and also in semiconductor based detectors it is often used. At room temperature, pure silicon has about 1 x 10^{11} free charge carriers per cm^3. As a first step in making silicon useful as a detector material, or any kind of electronic application, other elements need to be added to it. This process is referred to as *doping*. The four valence electrons of silicon form bonds with four neighbouring silicon atoms. If for example phosphorus, with five valence electrons, is added to silicon, the extra electron increases the overall conductivity of the substrate. When elements like boron, with three valence electrons, are added, the conductivity in terms of holes increases. These types of doped silicon, or semiconductor in general, are referred to as *n-doped* and *p-doped* respectively.

The influence of doping can be expressed in terms of the band model: the doping introduces energy levels inside the silicon bandgap. n-type semiconductors have additional levels, so-called *donor levels*, close to the conduction band whereas p-type semiconductors have *acceptor levels* close to the valence band.

To come to a piece of silicon apt for detecting charged particles, n and p-type silicon are put together to remove the thermally excited charge carriers which exceed by far the number of charge carriers created by an ionizing particle. Figure 3.5 shows such a pn interface, the *pn-junction* or *diode*, after reaching equilibrium: the e^- (h^+) from the n-type (p-type) material have diffused to the p-type (n-type) material, leaving behind fixed positively (negatively) charged ions. This layer with fixed space charge is referred to as the *depletion layer* and results in an electric field which counteracts the diffusion and in the end generates the equilibrium condition as shown in Figure 3.5. This depletion layer can be extended in size by applying a potential difference between the n and p sides: if a positive voltage is applied on the n-side with respect to the p-side then both n and p-side will be evacuated from free charge carriers. In this *reverse-biasing* scheme, the depletion

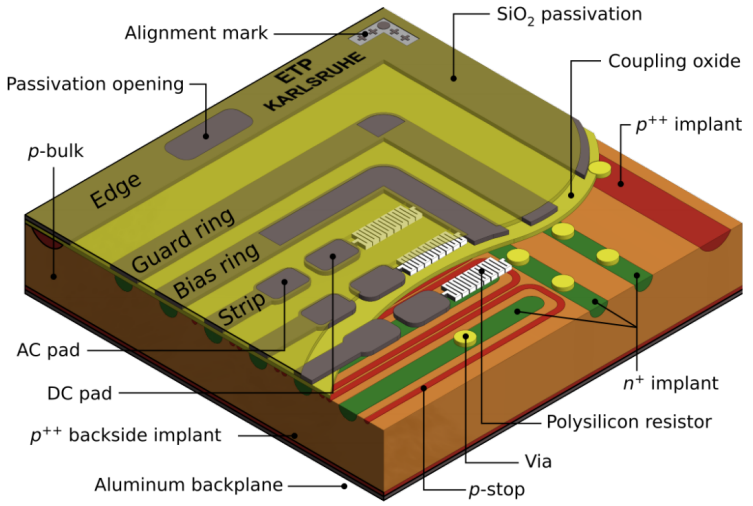


FIGURE 3.6: Layout of an n-in-p AC coupled silicon strip sensor [18].

layer, if the *biasing voltage* is high enough, extends across the full device and the volume is depleted from free charge carriers. In this way the doped silicon material becomes an excellent active volume for capturing charge deposited by ionizing particles.

Due to constantly produced $e^- h^+$ pairs the depletion region is never completely evacuated from free charge carriers and this results in a constant current through the device. This current is referred to as leakage current and needs to be minimised to e.g. not stress the power supply system and reduce the heat generated by the sensor.

3.4.3 Layout and operation of a position sensitive silicon sensor

Figure 3.6 shows the typical layout of an *AC-coupled* position sensitive silicon strip sensor for the *n-in-p* configuration. The main components are the *p-bulk* (shown in orange in Figure 3.6) and the n^+*-implants* (the green structures in Figure 3.6) forming a diode structure [15]. The diode is reverse biased by applying a high potential difference between the aluminium backplane (negative in this case and at the bottom of the volume in Figure 3.6) and the bias ring (positive in this case). The aluminium *DC pads* then receive this positive voltage through the polysilicon resistors[2] (shown in white in Figure 3.6) which are connected through a *via* with the n^+-implants whilst for the rest of the surface the n^+-implants and surface structures (such as the strips) are separated from each other by a SiO_2 layer. This biasing scheme results in the depletion of the p-bulk from free charge carriers. Due to the applied potential, highly mobile electrons are also drained from the SiO_2 leaving behind a slightly polarised oxide layer which attracts electrons from the bulk which concentrate at the Si-SiO_2 layer and can form a conductive path between the n^+-implants. This conductive path can affect the spatial resolution of the sensor by shortening the strips together or even shielding them from the bulk. In order to resolve this, local p^+-implants, so-called *p-stop* structures (shown in red in Figure 3.6), are introduced

[2]These resistors are needed to ensure a high ohmic path between the DC pad and the bias ring. A low ohmic path would result in all the induced charge going to the bias ring rather than to the readout electronics.

between the n^+-implants. These implants capture electrons and this results in a local negative space charge which repels other electrons and ensures the inter-strip isolation.

When a charged particle passes through the sensor most of the charge will be deposited in the p-bulk. The electrons will start to drift to the positive n^+-implants whilst holes follow an opposite path towards the backplane. The drifting electrons induce a current on the n^+-implants which are capacitively coupled to aluminium readout strips (*Strips* in Figure 3.6). A very homogeneous, well defined and preferentially thin SiO_2 layer is thus required between the implants and the strips to guarantee a uniform capacitance along and for all strips. The *AC pad* is then connected to a readout chip through wire bonds. This *AC coupling* scheme, where the capacitor is implemented on the sensor, has the advantage that the leakage current does not need to be drained by the readout chip, as opposed to the *DC coupling* scheme where the aluminium pad is directly connected to the implants and the readout chip thus needs to cope with a free bulk charge flow. In the DC coupling scheme the capacitor is then implemented in the readout chip.

The fact that these n^+-implants are localised (in the example in Figure 3.6 as strips) give the spatial resolution to the sensor. Other segmentations of the implants, such as pixels, can also be used. In that case the connection to the readout chip is performed by *bump bonding* the pixel to the readout chip's front-end.

The pixel dimensions or the strip *pitch*, defined as the distance between strip centres, is a very important design parameter of the sensor. It basically drives the spatial resolution of the sensor. The strip number where the signal is measured tells where the particle passed the sensor, when applying corrections for Lorentz drift of the e^-h^+ pair when the sensor is placed in a magnetic field. In case the signal spreads over more than one strip and the readout electronics provide information on the pulse height of the signal, the center of gravity of the signal can be calculated. When the readout electronics adopt a binary readout then the variance on the position resolution of a single strip cluster is given by [15]:

$$< \Delta x^2 >= \frac{1}{p} \int_{-p/2}^{p/2} x^2 dx = \frac{p^2}{12}, \tag{3.4}$$

where p is the strip pitch and x is the dimension orthogonal to the strip direction.

3.4.4 Radiation damage in silicon sensors

Operating sensors in a radiation environment can greatly affect their performance. In general, the radiation damage introduced in silicon sensors can be split up in *surface damage* and *bulk damage* [15].

Surface damage is most important for AC-coupled sensors as it manifests at the SiO_2 layer or Si-SiO_2 interface and the fundamental principle is comparable to the *total ionizing dose* effect on MOSFETs described in section 3.7.2: ionizing particles generate charge in the SiO_2 layer, the electrons are drained by the strip electrode and the holes start to drift to the Si-SiO_2 interface where they accumulate and attract electrons formed in the bulk. This results in a negative space charge which can result in shorts between the strips. These can be alleviated using p-stop techniques, which were already discussed in section 3.4.3.

Bulk damage is largely the result of non-ionizing interactions which affect the lattice structure of the silicon. Non-ionizing energy loss (NIEL) happens when a particle interacts with the atomic bodies via the electromagnetic or strong force and displaces the atomic body, like this creating a Frenkel pair: a vacancy and an interstitial. Also more

complex defects can be created. A commonly used unit for NIEL is the 1 MeV neutron equivalent: the non-ionizing radiation damage introduced by radiation of different type and energy can in first order be normalized to the radiation damage induced by 1 MeV neutrons. This is referred to as the NIEL hypothesis [19] which allows to compare NIEL irradiation tests performed with different particle types and/or energies.

The changes to the silicon lattice structure introduced by non-ionizing interactions result in additional energy levels in the silicon band gap which manifest as:

- Higher leakage current due to extra energy levels in the middle of the bandgap which facilitate generation and recombination of free charge carriers.
- Change in depletion voltage due to extra energy levels close to the valence and conduction band which effectively change the doping concentration.
- Lower charge collection efficiency due to defects capturing free charge carriers.

The bulk damage induced by radiation can change over time, even when the sensor is not irradiated: displaced atoms are able to move from their displaced position and the defect can either be cured or the movement can introduce new defects. This effect is referred to as *annealing* and has both positive and negative consequences. A positive implication is that the leakage current of a sensor, when no longer irradiated, typically goes down over time and this effect accelerates with temperature as the mobility of displaced atoms increases. As annealing changes the concentration of bulk defects it affects the effective doping concentration, which in turn modifies the depletion voltage and the charge collection efficiency of the sensor. The effect of annealing on the depletion voltage and the charge collection efficiency is positive at first (*beneficial annealing*), but over time *reverse annealing* takes over and results in a higher depletion voltage and lower charge collection efficiency. This reverse annealing can be slowed down by cooling the sensors to sub-zero temperatures [15].

The induced surface damage is typically not fully reversible and the recovery depends for example on the oxide quality. Furthermore emission of captured charge carriers from existing defects is suppressed due to the large band gap in the oxide. Annealing of the surface damage only sets in at high temperatures (around 100-400 °C).

3.5 Integrated circuits

An IC (Integrated Circuit) or a *chip* is a small, usually silicon, wafer which can hold millions of micro scale structures such as resistors, capacitors, transistors,... for use in analogue or digital applications. The main building blocks of integrated circuits are transistors which come in two main families: *bipolar junction transistors* and *field-effect transistors* (FETs). The latter have become the preferred type for use in *very large-scale integrated* (VLSI) circuits thanks to their small dimensions and the possibility they offer to create circuits with low power consumption [20]. The two most important types of FETs are *junction-gate FETs* and *insulated-gate FETs*. The latter are also more widely known as *MOSFETs* (metal–oxide–semiconductor FETs) and form the preferred type of FETs usually found in VLSI circuits due to e.g. easier fabrication, higher input resistance and smaller leakage current than junction-gate FETs. Throughout the years, channel lengths of MOS devices have substantially decreased. This *scaling* of the devices nowadays culminates in the 5 nm technology node being commercially available.

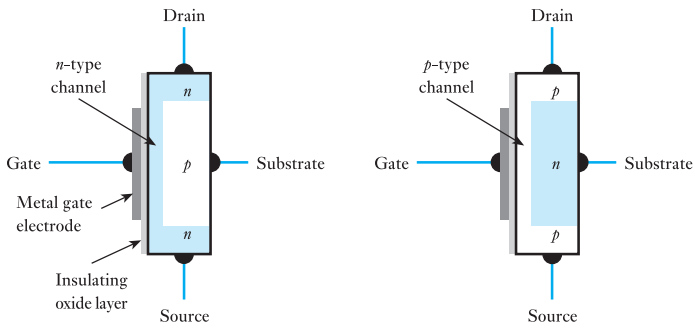


FIGURE 3.7: n-type channel (LEFT) and p-type channel (RIGHT) MOSFET [20]. The figures illustrate a depletion MOSFET or an enhancement MOSFET in the ON state.

Figure 3.7 illustrates the layout of an *n-type channel* (*NMOS*) and *p-type channel* (*PMOS*) MOSFET. The metal *gate* terminal of the transistor is separated from the *body* by an insulating layer, in this way creating the metal-oxide-semiconductor (MOS) sandwich. The voltage on the gate regulates the conductivity of the semiconductor body between the *drain* and *source*. In the NMOS configuration a positive voltage (with respect to the body) at the gate will attract electrons and make the n-type channel wider resulting in a higher conductivity between the n-type source and drain. By applying a negative voltage the channel becomes thinner and conductivity decreases. Other types of MOSFETs also exist. These *enhancement MOSFETs* do not have a manufactured n-type channel and can only operate in *enhancement mode* [20]. Enhancement MOSFETs are the transistor implementation used in *complementary MOS* (CMOS) devices.

NMOS and PMOS devices can be combined within a single circuit to form CMOS devices. These devices have the advantage that in a static configuration one of the transistors is always in the OFF state, which greatly reduces power consumption as the only current which flows is the one drawn by the load. Only during switching both transistors are conducting which results in a small burst of current from supply to ground.

CMOS transistors form the basis of nearly all electronic systems. They can be used as amplifiers in a huge variety of applications and as switches, logical gates, memory cells, etc. in digital systems. Figure 3.8a for example shows how NMOS and PMOS transistors are used to form a logic NAND gate. The implementation as shown in this figure, without a sampling clock, is referred to as *static* logic. Figure 3.8b shows the NAND gate implemented in *dynamic* logic, where the signal is evaluated only during a certain part of the clock cycle and where the output is stored on the capacitor X at the output.

3.6 Readout chips for silicon detectors

For particle physics experiments ICs are extremely interesting as they can provide an on-detector signal amplification, first-level data reduction and/or data buffering. The small dimensions of an IC and their low power consumption make them also suitable for use in tracking detectors. The ICs developed as on-detector electronics are mostly designed to meet one specific application and are therefore referred to as *Application-Specific Integrated Circuits* or *ASICs*.

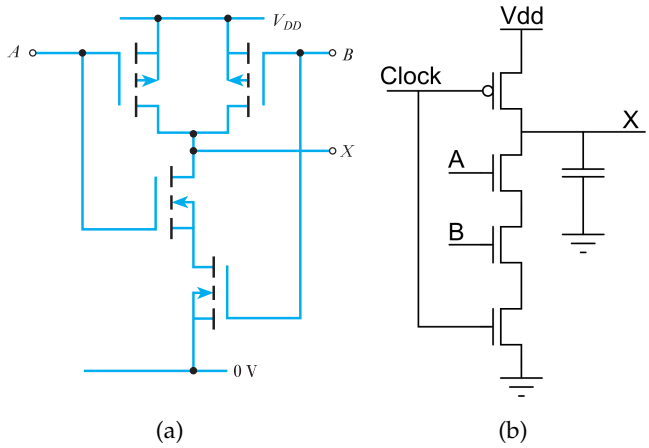


FIGURE 3.8: CMOS logic implementation of a static (LEFT) and dynamic (RIGHT) NAND gate [20].

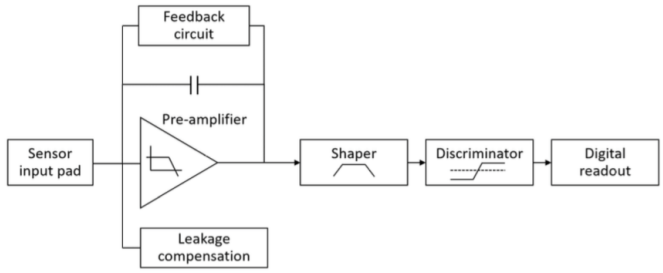


FIGURE 3.9: Typical front-end of an ASIC for the readout of a silicon sensor [21].

A typical ASIC for readout of a silicon sensor for the CMS Phase-2 tracker (Chapter 5) consists of an analogue front-end and digital back-end. In the past, for example for the Phase-0 strip tracker (section 4.3.2), fully analogue chips have been developed as well.

A schematic of a typical analogue front-end and the interface to a digital back-end is shown in Figure 3.9. It consists of a charge sensitive pre-amplifier stage with active feedback which basically operates as a charge integrator converting the input charge into a voltage. At this stage also leakage current compensation is implemented if the sensor is designed using the DC process (section 3.4.3). The next stage is a shaper circuit which filters low and high frequency noise and shapes the signal to a given rise time and return to baseline. The shaper is followed by a discriminator with a configurable threshold. The output of this discriminator is then sampled and used in the digital readout. More advanced front-end designs can be used if additional information is required such as more accurate timing information on the signal, a multi-threshold readout,...

Thanks to the scaling of features on ICs more advanced logic can be implemented in the digital part of the chips without increasing its area and without jeopardizing the power budget. This allows to already perform data reduction on the ASIC resulting in a smaller bandwidth to the back-end and/or to implement intelligence which aids the operation of the full system (see section 5.5). ASICs for silicon sensor readout are

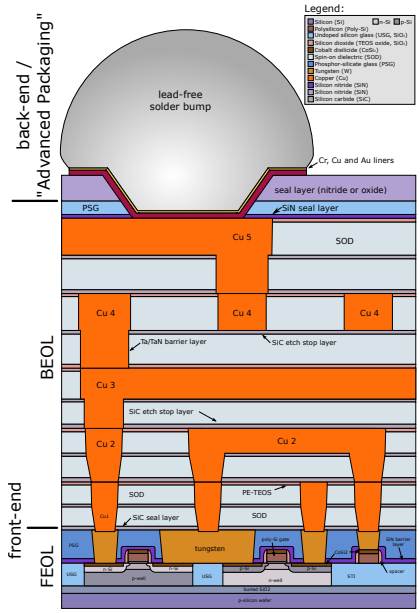


FIGURE 3.10: Typical layer structure of an integrated circuit [22].

typically produced using a standard planar process leading to a layer structure shown in Figure 3.10 where the transistors are in the first μm levels at the surface. Most of the volume is occupied by the routing of the interconnections and powering.

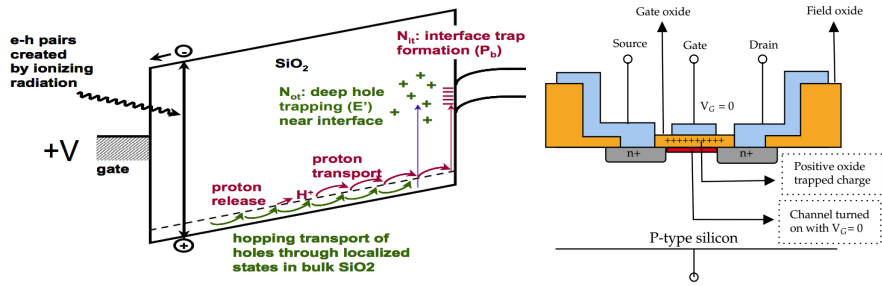
3.7 Radiation effects in integrated circuits

3.7.1 Introduction

During the design phase for the LHC experiments ASIC development for accelerator based particle physics experiments entered a new era. Due to the intense radiation environment generated by this high luminosity hadron collider the chips needed to be qualified for operation in these extreme conditions. Also now, during design for the HL-LHC, the electronics need to be hardened against and tested for several types of radiation induced effects.

Radiation effects in ICs can be very roughly categorized into two types: cumulative and transient radiation effects. Cumulative effects gradually change the performance of the device whilst transient effects are a sudden upset of the device as a result of radiation. Both of these effects can cause malfunctioning of the device. Radiation dose is expressed in units of Gray (= 1 J/kg) or rad (1 rad = 0.01 Gy) which represents the absorbed ionisation energy. For transient effects the particle flux for different particle types and energies is the unit of interest.

In the following sections both total ionising dose (TID), a cumulative effect, and single-event effects (SEE), a transient phenomenon, will be described. Displacement damage due to non-ionizing energy loss is of less importance in MOS devices and is therefore not discussed [23].



(a) (b)

FIGURE 3.11: LEFT: The result of ionising radiation in the gate oxide of a transistor. The holes can be transported to the Si-SiO_2 interface where they can be trapped by defects. The holes can interact with defects containing hydrogen. The liberated proton can introduce dangling silicon bonds at the interface [24]. RIGHT: Extreme result of TID in an NMOS transistor: the localised positive charge in the gate oxide can keep the channel turned on without application of a gate voltage.

3.7.2 Total ionizing dose

3.7.2.1 Introduction

A good understanding of TID effects [24] is important for devices which need to operate in a high radiation environment such as inside CMS. TID effects are also of importance in space applications although it is important to note that most electronics in earth satellites (in low earth orbit) receive a TID(SiO_2) of $\mathcal{O}$(50 Gy/year) [25], whereas the most inner levels of the CMS detector are expected to aggregate up to 12 MGy at the HL-LHC. In the next sections the physical mechanism behind TID effects will be discussed at transistor level (section 3.7.2.2) after which TID hardening techniques (section 3.7.2.3) and the methodology behind TID testing (section 3.7.2.4) will be introduced.

3.7.2.2 Total ionizing dose mechanism

When an ionising particle passes through a MOS transistor it can generate electron-hole pairs in the SiO_2 layer. About 17 eV of energy is required to create an electron-hole pair in SiO_2. Figure 3.11a illustrates the effect of ionising radiation in a MOSFET: the electrons produced by the passage of the ionizing particle have high mobility and are evacuated via the positive voltage on the gate or recombine with holes. The holes which do not recombine are forced away from the gate, but have a lower mobility than the electrons. The hole transport towards the Si-SiO_2 interface happens on localised oxygen vacancies until they reach the Si-SiO_2 interface where they can be trapped by defects. These defects are introduced in the production process and are e.g. caused by the lattice mismatches at the Si-SiO_2 interface. The trapped holes result in a layer of fixed positive charge, which, as illustrated in the transversal cross section of an NMOS transistor in Figure 3.11b, can in extreme cases result in the transistor being on without an applied gate voltage. The latter is an extreme case but exemplifies the influence of TID on the operation of a transistor.

Besides the direct generation of a localised positive charge near the Si-SiO_2 interface the ionizing radiation can also indirectly introduce defects: the holes on their way to the

$Si\text{-}SiO_2$ interface can also interact with defects containing hydrogen. This reaction liberates a proton which by diffusion or drift can move to the interface where it can remove hydrogen which passivates a dangling bond of a silicon atom. The remaining silicon dangling bonds can be seen as interface traps.

The device can actually recover from the dose it received. This is referred to as the *annealing process*. Tests have shown that whereas the positive charge buildup anneals, the induced traps do not [26].

In highly scaled devices, such as 65 nm CMOS technology, the effect of trapped charge in the gate oxide is typically negligible due to the small size of this layer. Other technology dependent TID effects have however been seen for these technologies and are discussed in Ref. [27]. Decreased drive currents[3] are reported at ultra high doses (> 1 MGy) and were traced back to certain areas in the transistor layout which are likely rich in defects and hydrogen making them ideal sites for interface traps and radiation-induced charge buildup. These areas are located in the source and drain regions. The radiation damage increases the electrical resistance of these regions.

3.7.2.3 Total ionizing dose hardening techniques

In general the TID effect becomes less important with decreasing technology feature size as the volume of SiO_2 decreases. Besides this, also active measures can be taken to harden the integrated circuit which mostly implies designing a special layout of the transistor to minimise the leakage current. These layouts typically require more area and will not be discussed here in detail. More information can be found for example in Ref. [28].

3.7.2.4 Total ionizing dose testing

Testing the device's TID sensitivity is mostly done by accelerated irradiation using X-rays or ^{60}Co γ-irradiation[4] where the end of life dose of the device can be reached in a few hours or days. The type of ionising radiation is important as the efficiency with which non-recombining holes are generated in the oxide is dependent on the radiation type and energy as illustrated in Figure 3.12. Electric field values, in for example 65 nm devices, are close to 5 MV/cm and at these values the generation of non-recombining holes with ^{60}Co irradiation and X-rays are respectively $\approx 100\%$ and $\approx 80\%$. Other particle types have a fairly low efficiency which makes testing practically less interesting. Typically the radiation field in which the device is eventually deployed does not only consist of photons and the testing with ^{60}Co and X-ray irradiation is therefore a conservative approach.

Also the state of the device when it is irradiated is important for TID evaluation. It is recommended to irradiate the device in a mode as it would be operated in the final application as the switching of the transistors influences the final effect of the irradiation in a non-trivial way.

Temperature during irradiation has a strong influence on the eventual radiation damage caused by TID. In general the radiation effects are more severe at higher temperatures [26]. Therefore testing the device at higher temperature than the eventual operation temperature is a conservative approach.

[3] In CMOS technology, drive current is defined as the drain current when the gate and drain are connected to the supply voltage whilst source and bulk are grounded.

[4] The decay of ^{60}Co gives rise to two γ peaks at 1.17 MeV and 1.33 MeV.

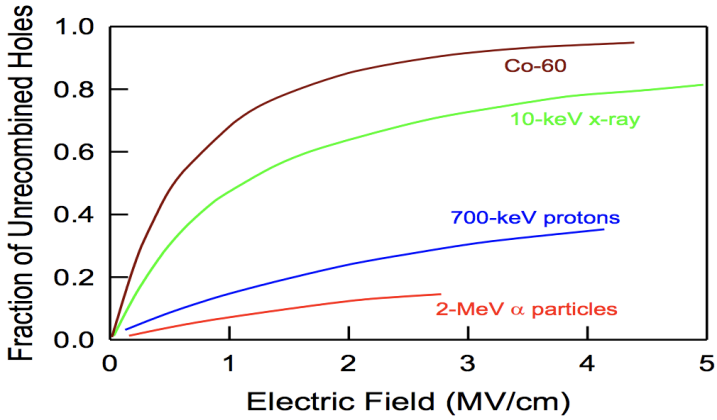


FIGURE 3.12: Efficiency for different types of radiation to generate electron-hole pairs, which do not recombine, in a transistor gate oxide as function of the electric field strength [24].

3.7.3 Single-event effects

3.7.3.1 Introduction

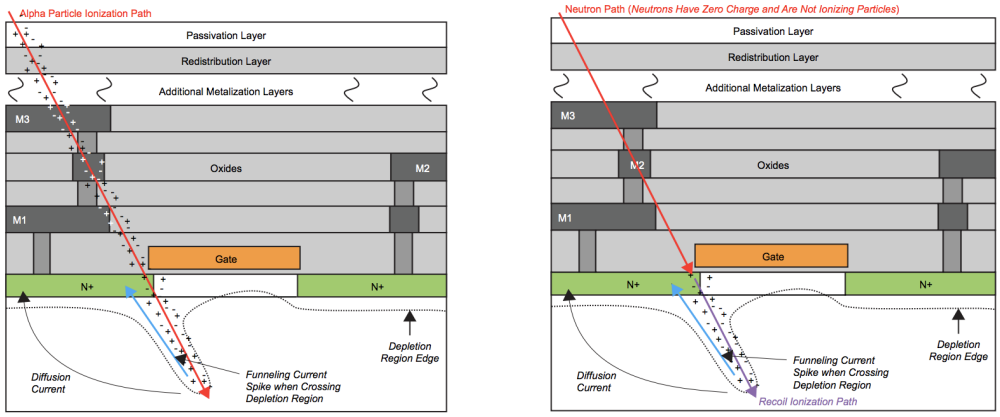
Next to the TID effect also single-event effects (SEE) need to be taken into account as radiation effects in ICs. These single-event effects are a sudden change in the operation of the device due to the passage of an ionizing particle. SEEs can be divided into two categories: soft and hard errors. Soft errors do not result in hardware damage and are correctable. Hard errors on the other hand can destroy the device. A lot of study has already been done on SEEs and SEE mitigation. The topic is relevant for particle physics experiments in high radiation environments, but also for aeronautics and space applications and is gaining more relevance in large scale computing farms, even at sea level.

Soft SEE can be categorized in Single-Event Transients (SET), Single-Event Upsets (SEU) and Single-Event Latchup (SEL). SETs are glitches in the circuitry caused by passage of an ionizing particle. These glitches show up as a temporary voltage spike. SEUs change the state of a storage element and SETs can thus result in SEUs when they affect such a sensitive node. A single ionizing particle can upset a single bit (single bit upset) but also multi-bit upsets (MBU) are possible. In SELs the ionising particle can create a high current state which might require a power reset of the device for it to go back to normal operation or the high current state can harm the device in which case it is catalogued as a hard error.

In the next sections the mechanism explaining the formation of an SEU will be discussed more in detail (section 3.7.3.2) followed by a brief overview of some SEU hardening techniques (section 3.7.3.3) and an introduction on how to test the SEU sensitivity of an IC (section 3.7.3.4).

3.7.3.2 Single-event upset mechanism

Figures 3.13a and 3.13b show an SEU triggered by a charged and neutral particle respectively. In case of the charged particle, the primary particle itself generates the electron-hole pairs along its track. In case of the neutral particle, the particle has to interact with



(a) (b)

FIGURE 3.13: Illustration of an SEU induced by a charged (LEFT) and neutral (RIGHT) particle [30].

the material and create charged secondaries which result in an ionizing track. The formation of an SEU can then be divided into three steps:

1. An ionising particle passes through the material to form free charge carriers.
2. The free charge carriers move by drift and diffusion whilst also recombining.
3. The free charge carriers are collected on a sensitive node (a node whose electrical potential can be modified by internal injection or collection of electrical charges) where they may trigger an SEU if the deposited charge in a given amount of time is large enough.

If the ionizing track crosses close to a pn-junction or a drain-gate-source region then the electric field, normally limited to the depletion region, extends further into the bulk silicon. In case of NMOS (PMOS) devices the electrons (holes) can be acquired by the NMOS (PMOS) drain. This electron or hole current is referred to as the funneling current [29]. Basically all ionising particles can generate SEUs, but only particles which give rise to a LET which is higher than a certain, device-dependent, threshold might generate SEUs. Therefore low energy heavy ions are of interest here. Low energy heavy ions giving rise to an SEU need to be produced in the silicon itself to generate an upset as the range of these ions is below 10 μm for a silicon nuclei with energy lower than 10 MeV as shown in Figure 3.4b.

Whether the free charge carriers created by the ionizing particle result in a minimum amount of charge on a sensitive node to trigger an SEU also depends on the charge collection mechanism. The charge collection is the result of drift (such as the funneling current) on one hand and diffusion of the free charge carriers which are left when the electric field restores on the other hand. During their travel to the sensitive nodes the electrons/holes can also recombine. The free charges have to be collected on the sensitive node in a time which is shorter than the response time of the circuit. The charge collection mechanism is a stochastic process and depends on charge trapping probability, doping levels, holes mobility,... and is therefore a very involved mechanism to describe for a complex device.

If enough charge is deposited on a sensitive node it can affect the operation of the logic where this node is integrated. An example of an SEU happening in an SRAM cell is illustrated in Figure 3.14, where the SEU results in a bit flip.

3.7.3.3 Single-event upset hardening techniques

Several mitigation techniques can be applied to make circuits less sensitive to SEUs. These techniques can be separated in physical hardening and logical hardening. Below, a few concepts of both types of hardening will be described. It is not intended here to give a full overview of hardening techniques as the existing hardening techniques are numerous and can differ a lot given the technology and application.

In order to mitigate SEUs in combinatorial logic it is advised not to use dynamic cells, where the information is stored on high-impedance nodes, but rather use static combinatorial logic where an active circuit that provides current and restores the correct value is always present [32].

In the same way DRAM memory is more susceptible to SEUs than SRAM. Besides using physical hardening in memory also logical hardening can be implemented in the form of forward error encoding to be able to correct bit upsets by SEUs.

Triplication is a common design methodology to make digital logic more SEU hard. A control state register in the IC might for example be triplicated and with a certain refresh frequency the values in the three registers can be compared with each other using a voter. This voter selects the majority value as the true value and rewrites the configuration if any of the registers does not match the other two. Like this a multi-bit upset in two different registers is required to generate an error. In order to minimise the probability for multi-bit upsets the physical separation between triplicated registers should be large enough (e.g. 15 μm for 65 nm technology).

During the design phase of the IC a lot of effort has to be put into designing and optimizing the triplicated logic on the chip and evaluating the overhead (increased chip size and power consumption, more complex routing,...) introduced by this extra logic. It is important to realise that SEUs in the control logic of the chip, which might introduce dead time, operation in an unknown state or destruction of the device are much more severe than single-event upsets resulting in bit errors in the output data of the chip. Most of the time a few upset bits in the output data stream can be tolerated.

The mitigation techniques implemented for two front-end chips for the CMS Phase-2 Outer Tracker are discussed in more detail in section 5.5.4.

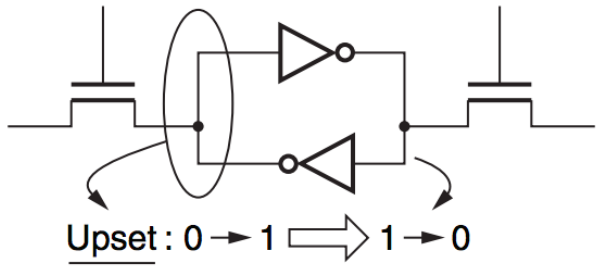


FIGURE 3.14: Illustration of how an SEU can affect the bit stored in an SRAM cell [31].

3.7.3.4 Single-event upset testing

Characterising the IC's sensitivity to SEUs and other SEE effects is typically done at particle accelerators such as cyclotrons. The die can be irradiated with different kinds of ionizing particles. Typically low energy protons or heavy ions are used for their large LET. During the testing the irradiation is changed so that the LET changes. This is achieved by changing the particle energy, type or incident angle. In the end the SEU cross section:

$$\sigma = \frac{\#SEU}{\phi}, \tag{3.5}$$

needs to be determined for different LET. ϕ in Equation (3.5) is the particle fluence on the device. In theory the SEU cross section versus LET would show a step function behaviour, as a minimal LET is required to generate the critical charge on a node. In reality this step function is smoothed and can be represented by a cumulative Weibull distribution function [16]:

$$\sigma = \sigma_0 \left\{ 1 - exp\left[-\left(\frac{E_{dep} - E_0}{W} \right)^s \right] \right\}, \tag{3.6}$$

with W and s as shape parameters and with E_{dep} the energy deposited by the ionizing particle in a sensitive volume. The smoothing can be explained by the variation in the charge collection efficiency across the sensitive area and by junctions with different sensitivities. The critical value (E_0) then represents the minimal energy deposit to upset the most sensitive part of the design. MBUs, more probable for higher LETs, can increase the saturation threshold (σ_0). Once the SEU cross section as a function of LET is known the bit error rate due to SEU for a particular radiation environment can be predicted by convoluting the SEU cross section as function of LET with the probability distribution to have a certain energy deposit in a certain sensitive volume for a given particle spectrum and multiplying this with the particle flux. This will be explained in more detail and applied in section 8.4.

3.8 Summary

Understanding the particle interactions with matter is indispensable for describing both particle detection mechanisms and the radiation damage which these particles can induce in detector components. The interaction of a particle with material strongly depends on the particle's properties such as its mass, energy and charge. In silicon sensors, the signal generation happens in the bulk of the sensor material which is depleted from free charge carriers by applying a bias voltage over the pn-junction. Charged particles passing the sensor will deposit energy, which results in the production of e^-h^+ pairs ($\approx 76\, e^-h^+/\mu m$ for a 300 μm thick sensor) in the silicon bulk and their charge flow induces a signal on the readout strips which are connected to the readout ASICs.

For silicon sensors, both radiation damage induced by non-ionizing and ionizing interactions are important. The former is linked to bulk damage, whereas the latter is linked to surface damage in the sensor. Bulk damage mainly results in higher leakage current, change in depletion voltage and a lower charge collection efficiency. Surface damage on the other hand can affect for example the interstrip resistivity. The surface damage mechanism in sensors is comparable to the TID effect introduced by radiation on ICs. ASICs in the CMS Outer Tracker will aggregate a dose up to 600 kGy [33] and without careful

design choices this could have severe effects on the operation of the ASICs. The TID effect is greatly dependent on the exact technology which is used for the ASIC and testing the ASIC's TID sensitivity by accelerated irradiation with ionizing radiation is imperative in the ASIC's validation procedure. Besides TID effects also SEUs are a consequence of operating ASICs in a radiation field. Highly ionizing particles, such as low energy heavy ions, induce these effects as they can deposit a large charge in the sensitive volume around a sensitive node. SEUs can cause malfunctioning of the device if the device is not properly designed. Design techniques such as triplication are therefore critical to mitigate SEU effects. The mitigation strategy can be evaluated by testing for SEE and measuring the SEU rate at heavy ion facilities.

Chapter 4

The LHC and the CMS Experiment

4.1 Introduction

Accelerator based high energy physics experiments have proven to be indispensable as discovery avenues in particle physics and for testing of the Standard Model. The most powerful accelerator available today is the Large Hadron Collider (LHC) which allows physicist to do lab experiments at the energy frontier. This chapter will give an overview of the LHC and some of the experiments based at the LHC (section 4.2), goes into more detail on the Compact Muon Solenoid experiment (section 4.3), sets the stage for the next chapter by introducing the High Luminosity LHC and its physics motivation (section 4.4) and the upgrade plans of CMS for the HL-LHC phase (section 4.5).

4.2 The LHC and the LHC experiments

The Large Hadron Collider is the most powerful man made particle accelerator ever constructed. It is housed in a 26.7 km long subterranean tunnel and forms the final stage for particle acceleration at the CERN accelerator complex. The accelerator complex is illustrated in Figure 4.1 where the particle acceleration starts at the LINAC, after which the booster, PS and SPS follow. The LHC itself consists of accelerating structures as well as superconducting magnets to focus the particles, keep them on track and squeeze them closer together right before the interaction point.

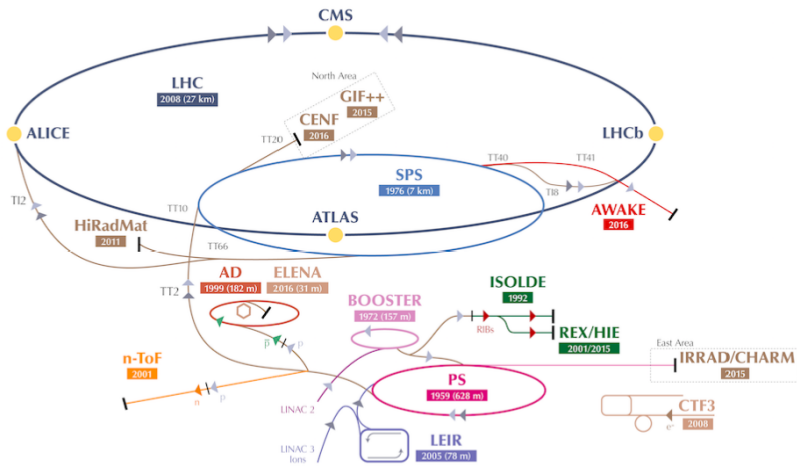
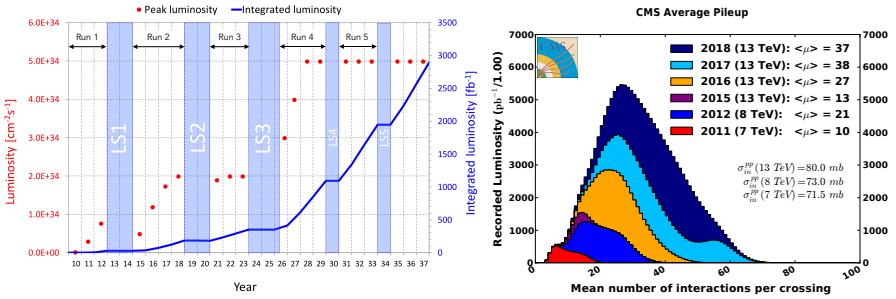


FIGURE 4.1: Illustration of the CERN accelerator complex [34].

The LHC can accelerate both protons and heavy ions. The protons and/or heavy ions are accelerated in opposite directions in two separate accelerator tubes. They are then made to collide at four interaction points along the LHC ring where detectors are located with the goal to reconstruct the particles which are formed during a high energy collision. The protons or heavy ions are accelerated in bunches and the LHC delivers 40 million bunch crossings per second. About 2800 proton bunches travel around the LHC, each bunch containing of the order of 10^{11} protons [35].

The four main experiments at the LHC are ALICE (A Large Ion Collider Experiment), ATLAS (A Toroidal LHC ApparatuS), CMS (Compact Muon Solenoid) and LHCb (Large Hadron Collider beauty). ATLAS and CMS are general purpose detectors which were designed with the goal to discover the Higgs boson but also to be fit for a multitude of other physics analysis. ALICE is specifically designed to analyse the quark-gluon plasma formed by heavy ion collision and LHCb is specifically designed to study the properties of the B mesons.

Figure 4.2a shows the schedule of the LHC: in 2010 operation started, followed by a running period at 7 TeV in 2011 and 8 TeV in 2012. The period from 2011 until 2012 is referred to as *Run 1*. Run 1 directly confirmed the physics potential of the LHC with the Higgs boson discovery by ATLAS and CMS [36, 37]. *Run 2* was started in 2015 after a 2 year Long Shutdown where the accelerator and experiments were consolidated. The centre of mass energy during Run 2, which lasted until 2018, was 13 TeV.



(a) (b)

FIGURE 4.2: LEFT: Acquired and projected integrated and instantaneous luminosities at the LHC [38]. RIGHT: Pileup as measured by the CMS experiment over the period from 2011 until 2018 [39].

Figure 4.2a also shows the acquired and projected integrated and instantaneous luminosities for the ATLAS and CMS experiment. The instantaneous luminosity (L) is a beam dependent parameter and is defined as the number of collisions (N) in a certain time (t) normalised to the interaction cross section (σ):

$$L = \frac{1}{\sigma}\frac{dN}{dt}. \tag{4.1}$$

The unit of instantaneous luminosity is $\mathrm{cm}^{-2}\mathrm{s}^{-1}$, but also $\mathrm{b}^{-1}\mathrm{s}^{-1}$ (1 b = 10^{-24} cm^2, 1 mb = 10^{-27} cm^2) is used. The integral of the instantaneous luminosity over time is simply referred to as *integrated luminosity*:

$$L_{int} = \int L\, dt. \tag{4.2}$$

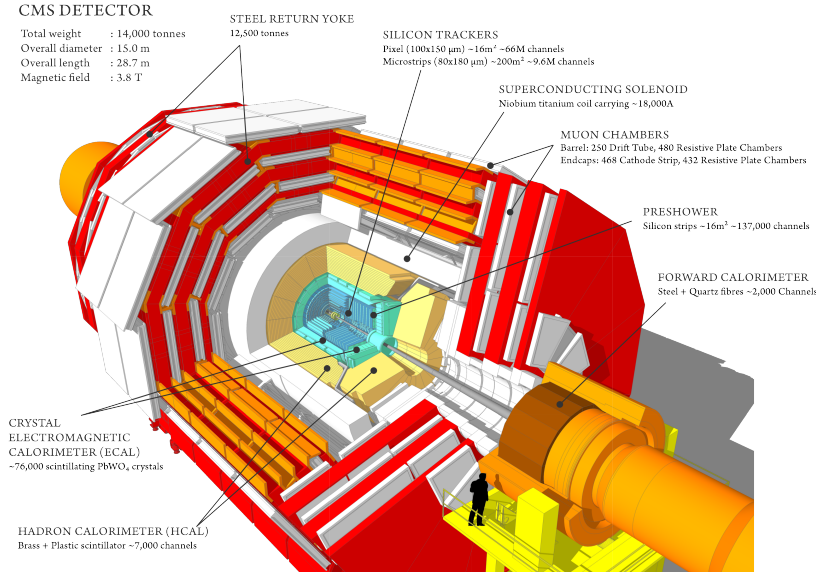


FIGURE 4.3: Layout of the CMS detector [40].

At the end of 2018, about 192 fb^{-1} [39] of proton-proton (pp) collisions were delivered to CMS. The gradual increase in instantaneous luminosity is a direct consequence of improvements in the beam quality and intensity which results in a higher number of proton-proton collisions happening during one bunch crossing (BX). The latter number is referred to as *pileup* (PU). The distribution of the pileup measured at CMS is shown, for each running year, in Figure 4.2b, from which it is clear that the LHC surpassed its original design value of $<PU> \approx 25$. This larger pileup extends the physics reach of the experiments due to the higher probability of occurrence of an interesting collision, but on the other hand forms a challenge as most of the PU interactions are soft and in general less interesting to test new physics models. The detectors should therefore be able to disentangle every proton-proton collision. A significant increase in luminosity is foreseen by 2026 as a result of an upgrade of various parts of the accelerator complex. This is an important milestone in the LHC schedule and is referred to as the *High Luminosity LHC* (HL-LHC). As a result of this upgrade, many of the detectors at the LHC will need to be upgraded to cope with the increased pileup. The average expected pileup at HL-LHC is expected to be 150, or even 200 if the ultimate design prospects are achieved [33]. The preparation for HL-LHC started already in 2019 during Long Shutdown 2. After this Long Shutdown, another 3 year run at 14 TeV is scheduled, followed by a 3rd Long Shutdown from 2024 to mid 2026 where the main upgrades for the HL-LHC will be put in place. After this 3rd Long Shutdown the HL-LHC will provide collisions to the experiments for another 10 to 15 years. Section 4.4 will give more information on the HL-LHC.

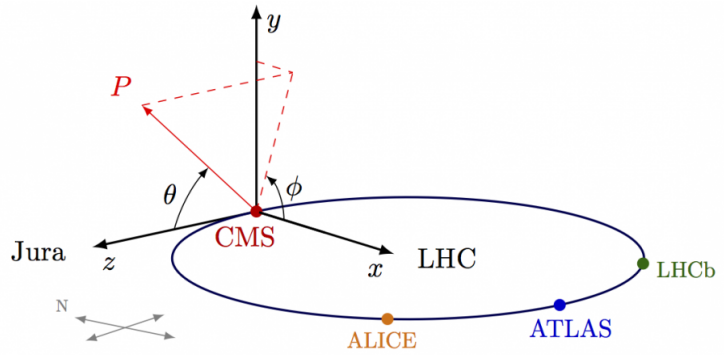


FIGURE 4.4: Definition of the coordinate system adopted by CMS [41].

4.3 The CMS experiment

4.3.1 Introduction

Particles can decay to other less massive particles following conservation laws such as energy conservation, momentum conservation, baryon number conservation, etc. Massive particles can be created in the reverse process where high energy particles are collided. At the LHC's interaction points experiments were build to study the particles formed during such high energy collisions. The experiment, located at LHC P5 in Cessy, France, is the CMS experiment.

Figure 4.3 shows the structure of the CMS detector with at its centre the interaction point. Figure 4.4 shows the definition of the coordinate system: the incoming particles travel along the z-axis and collide at $z \approx 0$, the azimuthal angle ϕ is measured in the plane transverse (*transverse plane*) to the z-axis and the polar angle θ is measured from the z-axis. CMS is almost symmetrical for positive and negative z coordinates.

Going from the interaction point outward, the particles formed in the collision can pass several subdetector systems. These subdetectors all have a specific design, optimised in such a way that combining the data from each subdetector gives the best possible characterisation of the particles. Figure 4.5 shows a longitudinal quadrant of the CMS experiment where the different subdetectors can be distinguished. Going radially outward from the interaction point, a particle not stopped by the material, like a high energy muon, would encounter: the pixel silicon tracker, the strip silicon tracker, the electromagnetic calorimeter, the hadron calorimeter and the muon system. In between the hadron calorimeter and the muon system sits a superconducting solenoid generating a 3.8 T magnetic field in which charged particles are bend. A measurement of the bend can be used to extract the momentum of the particle. Figure 4.5 shows, for five different types of particles, that each particle has a distinct signature in CMS. This forms the basis of the particle identification.

Figure 4.5 also shows the forward coverage of the different subdetectors. Important to note here is that the CMS detector was designed pre-HL-LHC to have coverage in all subdetectors up to $|\eta| = 2.5$ where η is the pseudorapidity defined as:

$$\eta = -\ln\left(\tan\left(\frac{\theta}{2}\right)\right). \tag{4.3}$$

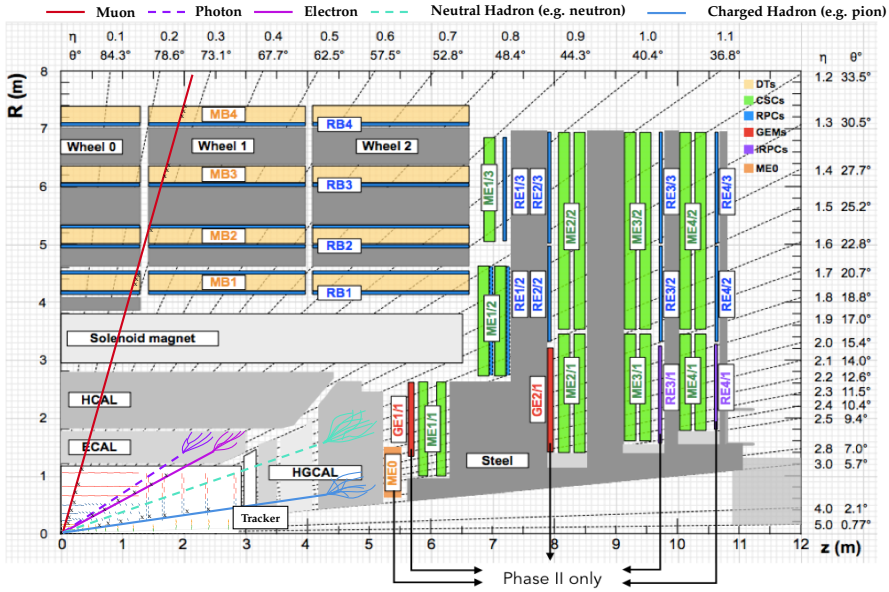


FIGURE 4.5: Quadrant of the CMS experiment (adapted from [42]). The particles collide at $R \approx z \approx 0$. The passage of some particle types through the different detector layers are illustrated. The muon wheels (RE3/1, RE4/1, GE1/1, GE2/1, ME0), which will only be present for the Phase-2 upgrade (see section 4.5) are indicated. The location of the Phase-2 HGCAL is also indicated.

Forward calorimeters with higher forward acceptance than $|\eta| = 2.5$ are also present (section 4.3.3).

Within the tracker acceptance the charged particle density per η unit is approximately constant. This, together with the fact that the particle density in general reduces when going radially away from the interaction point has several consequences:

- In order to keep the signal occupancy in a given sensor channel low enough the density of readout channels has to scale with the particle density: the closer to the interaction point the smaller the dimensions of single channels. This complicates the design of the detectors sitting close to the interaction region and often an on-detector data reduction has to happen to be able to pass the data to the readout system.
- In general, the closer to the interaction region the more radiation the detector and readout electronics have to cope with.
- In the opposite direction, the further away from the detector the larger the detectors need to be to cover the solid angle at that radius. This has implications on the type of detectors which can be used. Also financial aspects play an important role here.

In the next sections a description of each CMS subdetector system will be provided and the basic particle detection mechanism for each system will be introduced. Special attention is paid to the tracker system as this system is of particular importance for the

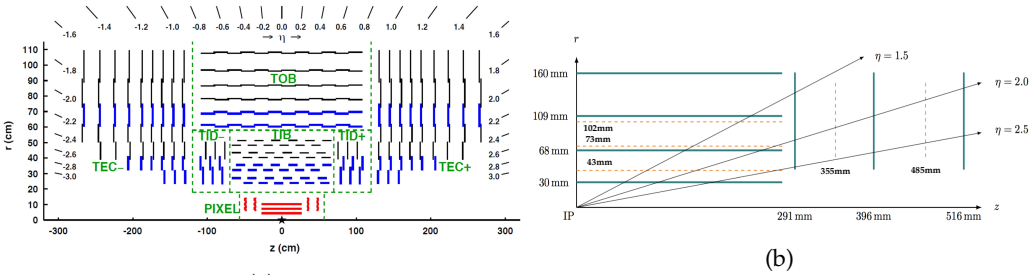
rest of the work. The system deciding which events to keep for offline storage, the CMS trigger system, will be introduced in section 4.3.5.

4.3.2 The tracker system and track reconstruction

4.3.2.1 The tracker detector system

The tracking system at CMS performs precise reconstruction of charged particle trajectories inside the 3.8 T magnetic field. It does this using silicon sensor technology to perform the particle detection as discussed in section 3.4. The tracker system plays an important direct role in electron, muon, hadron, tau, jet and primary vertex reconstruction and contributes to vetoes for neutral particles. The CMS tracker is split up in two subdetectors: the pixel and strip tracker.

In the 2016 extended year-end technical stop the first pixel detector, the Phase-0 pixel detector, was replaced by the Phase-1 pixel detector. This replacement was needed in order to guarantee an efficient running for the remainder of Run 2 and for Run 3. Without this upgrade significant loss in tracking performance due to radiation damage or due to limited buffering capabilities in the readout chip were expected. Figure 4.6a and 4.6b respectively show a longitudinal cut of the Phase-0 full tracker system and a quadrant of both pixel detectors.



(a) (b)

FIGURE 4.6: LEFT: CMS Phase-0 pixel and strip tracker geometry [43]. The modules indicated in blue are the stereo-modules. RIGHT: One quarter of the CMS Phase-1 pixel detector. The dashed orange lines represent the Phase-0 geometry [44].

The Phase-0 pixel detector has 66 million 100 μm $\times$ 150 μm and 285 μm thick pixels spread over 3 barrel layers, at radii of 4.3, 7.3 and 10.2 cm, and 2 endcap disks at both sides at a longitudinal distance of 35.5 and 48.5 cm from the interaction point. The pixels are read out using the PSI46 ASIC.

The Phase-1 pixel detector consists of 4 barrel layers and 3 endcap disks at both z sides summing up to a total of 124 million pixels. Besides the increased hit coverage the Phase-1 pixel detector is also lighter than the Phase-0 pixel detector due to several mechanical improvements and relocation of services. The sensor design is identical to the Phase-0 sensors, but the readout chip is improved with increased buffer stages and an increased output bandwidth: the PSIDIG chip was used for all layers except for the innermost layer where an improved high rate version, the PROC600 was used. The operation of all three chips is similar: each chip is bump bonded to 52$\times$80 pixels, the chip performs zero-suppression and the full pulse height information of a hit channel is passed to the

readout boards. The main difference between the PSI46 on one hand and the PSIDIG and PROC600 chip on the other hand is that the PSIDIG and PROC600 digitise the pulse height using an ADC (analogue to digital converter) and send this information out digitised whereas an analogue output is used in case of the PSI46 [15]. In both the Phase-0 and Phase-1 pixel detector, the modules are oriented to provide the best resolution in the $r\phi$ bending plane.

The strip tracker consists of a barrel region (TIB and TOB) where the modules are parallel to the beampipe and endcaps (TIDs and TECs) where the modules are perpendicular to the beampipe. Each line in the strip tracker region in Figure 4.6a represent a detector module. The modules are oriented in such a way that the strip direction provides the best resolution in the $r\phi$ plane. An example of such a strip tracker module is shown in Figure 4.7a.

In total, the strip tracker has 9.3 million channels on 320 μm or 500 μm thick silicon sensors with a strip pitch varying from 80 to 184 μm. The layers indicated in blue in Figure 4.6a are instrumented with a second module which is mounted back-to-back to another module at a stereo angle of 100 mrad. These *stereo-modules* allow to measure the hit position along the strip direction. The strip sensors are read out by the APV analogue chip [45], each APV receiving the signal of 128 channels.

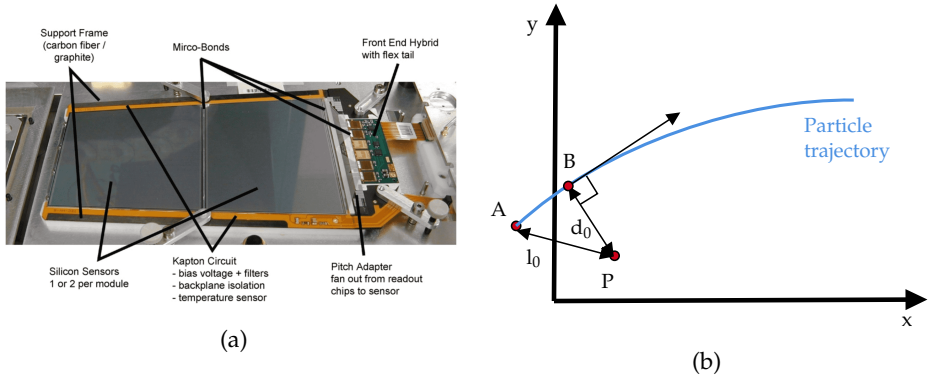


FIGURE 4.7: LEFT: An example of a CMS Phase-0 strip tracker module [46]. RIGHT: The definition of the d_0 parameter and l_0 parameter of a track originating from point A. The point P is the reference point and could for example be the beamspot or a primary vertex. The point B is the reference point of the track and is the point of closest approach in xy of the track to P.

4.3.2.2 Charged particle track reconstruction

The combination of pixel and strip information allows for very precise charged particle track reconstruction, referred to as tracking, which basically boils down to a complex exercise of connecting the hits in several tracker layers. Tracking starts with the reconstruction of hits, followed by generation of seeds, track finding, track fitting and track selection. These different steps will be discussed briefly below. More information can be found in Ref. [47].

The pixel detector ASIC performs zero-suppression with an adjustable threshold per pixel. In the offline reconstruction, clusters are formed from this zero-suppressed data.

Iteration	Seeding layers	p_T (GeV/c)	d_0 (cm)	$\lvert d_z \rvert$
0	Pixel triplets	>0.8	<0.2	<3 σ
1	Mixed pairs with vertex	>0.6	<0.2	<0.2 cm*
2	Pixel triplets	>0.075	<0.2	<3.3 σ
3	Mixed triplets	>0.35	<1.2	<10 cm
4	TIB 1+2 & TID/TEC ring 1+2	>0.5	<2.0	<10 cm
5	TOB 1+2 & TEC ring 5	>0.6	<5.0	<30 cm

TABLE 4.1: Configurations used during the track seeding. For each of the six iterative tracking steps the detector layers used to seed the tracks are listed, as well as the requirements on the minimum p_T, maximum transverse (d_0) and longitudinal (d_z) impact parameter with respect to the beamspot. *Mixed pairs with vertex* are seeds that use two hits and a third point given by the location of a pixel vertex. *Mixed triplets* are seeds produced from three hits formed from a combination of pixel hits and matched strip hits with at least one pixel hit. σ is the Gaussian standard deviation corresponding to the length of the beamspot along the z-direction. The asterisk indicates that the longitudinal impact parameter is calculated relative to a pixel vertex instead of the centre of the beamspot.

The exact position of the cluster on the sensor is extracted using a fast algorithm during track seeding and track finding. For the final track fit a more precise algorithm is used.

For the strip detector, the zero-suppression happens in off-detector electronics. Clusters are built offline and the cluster position is determined from the charge-weighted average of the strip positions. At this stage corrections are performed for the Lorentz drift and for inefficiently collected signal at the sensor's backplane in the thickest sensors caused by the finite integration time of the signal in the APV.

Once the clusters are reconstructed the tracking takes into account differences between actual and assumed location of the sensing elements as extracted from alignment procedures.

The actual CMS track reconstruction is based on Kalman filters [48, 49, 50] and happens in several iterations. This iterative tracking allows for pattern recognition and track fitting to happen within the same framework. During the first iterations the tracks which are easiest to find (e.g. large transverse momentum, small displacement) are reconstructed and after each iteration the hits associated to tracks reconstructed in this step are removed, making subsequent iterations easier. In the last iterations, lower p_T tracks and tracks with larger displacement are reconstructed. A single iteration consists of four steps:

1. Seed generation: a first estimate of the 5 track parameters is made using either three 3D hits[1] or two 3D hits and a constraint on the production vertex of the track such as the beamspot or primary vertices[2]. Only 3D hits are used during the seed generation to improve the quality of the seeds and the speed of the seeding algorithm. 3D hits are obtained from pixel hits or matched strip hits in the stereo-modules. The seeding works from the inside out and for every iteration different *seeding layers* are used with different constraints on kinematic parameters of the extracted seeds. Examples of constraints are a minimal p_T and their consistency with originating from

[1] A 3D hit is a hit that provides a 3D position measurement.

[2] The beamspot and primary vertices needed in this step are reconstructed with a special reconstruction algorithm which reconstructs pixel tracks and vertices.

the pp interaction region. The requirements, as given in the last public document [47] on the CMS tracking, are listed in Table 4.1 together with the detector layers used for seeding in each iteration. The required number of pixel hits in the seeding layers decreases as the iterative procedure advances. The *seeding layers* range from *pixel triplets*, where the three seeds are only allowed to consist of pixel hits, in the first iteration to *strip pairs*, where seeds consist of a pair of matched hits in the layers/rings of the strip detector which are equipped with stereo-modules. The latter seeds do not require any pixel hits and are therefore useful to find tracks that do not leave hits in the pixel detector or are produced outside its volume.

2. Track finding: the seed is used as input to a Kalman filter to generate the first coarse estimate of the track parameters. Subsequently, compatible detector modules, lying along the extrapolated track[3], are searched for. The hit locations within the module are refined using the impact parameter of the already reconstructed track and compatible hits are selected using a χ^2 test. With the addition of each hit, the track parameters and covariance matrices are updated by the Kalman filter. By using a Kalman filter, stochastic effects, such as multiple scattering can be taken into account. Furthermore, thanks to the Kalman filter, the track finding always carries along the full information available up to a certain point and thus gives the best possible prediction for the next point. Hits are added until a stop condition (e.g. the track going out of the tracker volume or too many missing hits) is fulfilled. The track finding searches for outward going tracks, but when the stop condition is reached an inward search is also started which can add hits which were for example not included in the seed when the seeding was done using matched hits in the stereo modules.

3. Track fitting: once all hits are assigned to a track, the track is refitted using a Kalman filter. This fitting stage gets rid of biases due to constraints (e.g. beamspot) applied during seeding, updates hit position information, takes into account the inhomogeneous magnetic field and looks for outlier hits.

4. Track selection: quality requirements are used to reduce the track fake rate. The criteria are the minimum number of layers in which the particle generates a hit, the χ^2 of the track fit and significances of several of the track parameters. A quality flag is extracted from the track selection and high purity tracks are flagged.

The tracking performance for some important tracking parameters is shown in Figure 4.8. Figures 4.8a and 4.8b respectively show the tracking efficiency as function of p_{T} and radial displacement of the track creation vertex of a subset of the tracks as defined in the figure labels. The efficiency is defined as the number of matched reconstructed tracks divided by the number of simulated tracks. A high track reconstruction efficiency down to low values of the transverse momentum can be observed and an $\approx 35\%$ tracking efficiency is obtained at radii of the track production vertex as high as 45 cm. Figure 4.8c and 4.8d show respectively the resolution of the transverse momentum and transverse impact parameter (d_0). The definition of the impact parameter is illustrated in Figure 4.7b. The p_{T} resolution is better than 4% over the full p_{T} range and for the impact parameter, resolutions below 150 μm can be attained for the most central tracks [51].

[3]Either an analytical extrapolation to the next detector layer is used when no multiple scattering is assumed or a more extended and time consuming extrapolation is performed which does include multiple scattering and energy loss due to the particle crossing material.

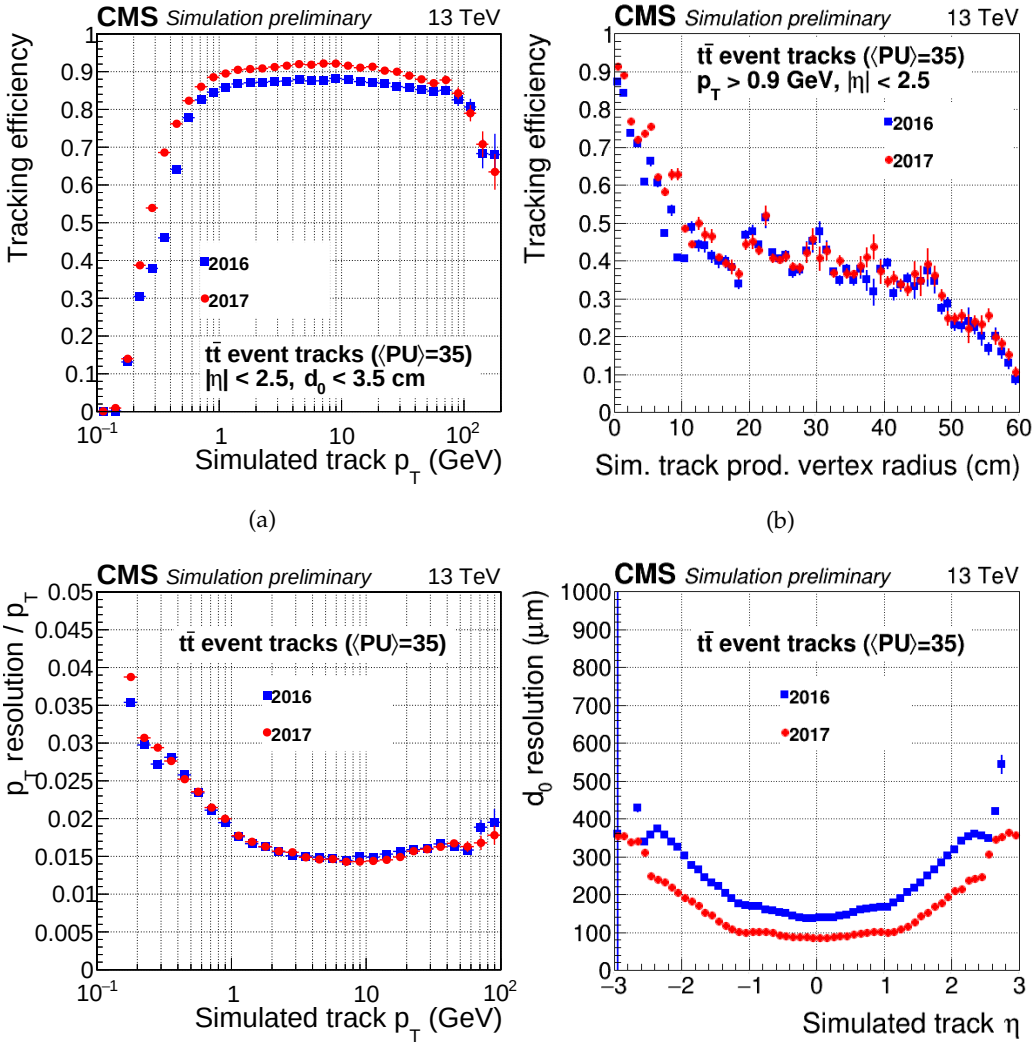


FIGURE 4.8: TOP: Tracking efficiency versus p_T (LEFT) and displacement (RIGHT) for both the 2016 and 2017 configuration of the tracker [51]. BOTTOM: Resolution of the transverse momentum (LEFT) and transverse impact parameter (RIGHT) for both the 2016 and 2017 configuration of the tracker [51].

4.3.3 The calorimeter system

As the name suggests the calorimeter system is designed to measure the energy of the particles. In contrast to the tracking system, which is designed as light as possible not to change the particle's track the calorimeters need to fully stop the particle inside their volume in order to measure the particle's energy. The CMS calorimeter system consists of the electromagnetic (ECAL) and hadronic calorimeter (HCAL).

The electromagnetic calorimeter [52] is specifically designed to measure the energy of

photons and electrons. The ECAL is build up of 75,948 lead tungstate ($PbWO_4$) crystals which are very dense and form a good scintillating material with fast timing characteristics and an emission peak at 425 nm which is in a good range for photon detectors. The ECAL is organised in a barrel (EB) and two endcap (EC) sections. The crystals used in the barrel (endcap) have a length of 23 (22) cm and a frontal surface of $\approx 2.2 \times 2.2$ cm^2 (2.68×2.68 cm^2), resulting in a $\Delta\eta \times \Delta\phi$ coverage of 0.0175×0.0175 in the barrel while in the endcap it varies from 0.0175×0.0175 to 0.05×0.05.

The scintillation light is detected by two avalanche photodiodes per crystal in the barrel and by one vacuum photo-triode per crystal in the endcaps. Amplified and digitised pulse information are stored in a pipeline, which sits on a dedicated front-end readout card, for 256 bunch crossings awaiting readout triggered by an L1A (Level1-accept) trigger (see section 4.3.5). The front-end readout card receives data corresponding to 5×5 crystals and performs transverse energy sums at BX (bunch crossing) frequency[4]. These transverse energy sums are used in the L1 (Level1) trigger system (see section 4.3.5) and in trigger terminology such a detector unit is referred to as a *tower*.

The hadron calorimeter [53] is also build up of barrel (HB) and two endcap (HE) sections. Most of the HCAL is located inside the magnet, but there is also a part outside the magnet (outer HCAL (HO)) to increase the coverage for high energy hadrons and a forward (HCAL forward (HF)) calorimeter ($2.9 < \eta < 4.1$) is also present. HB and HE are sampling calorimeters made of brass absorber material interleaved with 17 layers of scintillator as active material. Each scintillator is read out by a wavelength shifting fibre. The wavelength shifting fibres from HB and HE as well as the ones connected to the scintillators which constitute HO are read out by photodiodes.

HF is build of steel and quartz fibres, with two different lengths, parallel to the beam direction. The fibres conduct the Cherenkov light, created in their volume by the passage of charged shower particles, out of the HF where the light is read out using PMTs (photo multiplier tube). The $\Delta\eta \times \Delta\phi$ segmentation of HCAL decreases from 0.087×0.087 to 0.175×0.35. The signal from photodiodes or PMTs is integrated and digitised in a custom made chip which sits on the detector.

HCAL also provides input data to the L1 trigger system from which transverse energy sums are calculated. The L1 information from ECAL and HCAL are then combined before input to the *Global Trigger* system (see section 4.3.5).

4.3.4 The muon system

The muon system [54] is split up into a barrel and two endcap sections. The system consists of three types of detectors: drift tubes (DTs) in the barrel region, cathode strip chambers (CSCs) in the endcap region and resistive plate chambers (RPCs) in both the endcap and the barrel region. The entire muon system has an area of $\approx$ 25,000 m^2. A detailed view of the configuration is shown in Figure 4.5. CMS uses the fact that muons are deeply penetrating particles by placing the muon detectors at different stages inside an iron return yoke which guides the magnetic field lines. This return yoke allows CMS to do performant muon reconstruction up to high muon momenta whilst still maintaining a compact design.

All three CMS muon detectors are based on the localised ionisation of a gas volume inside the detector chamber due to the passage of an ionising particle. The DTs and CSCs

[4]The term *BX* will be used in this work not only to refer to an actual crossing of bunches at the LHC, but also as a unit of time, being the time between bunch crossings.

provide very high spatial resolution. The spatial resolution per chamber was measured to be 80 to 120 μm in the DTs and 40 to 150 μm in the CSCs [54].

The DTs feature a gold-plated stainless-steel anode wire operating at 3600 V at the center of a 42 $\times$ 13 mm^2 cell. Each cell has a length equal to the length or width of the chamber, which range from $\approx$ 2.5 m up to $\approx$ 4.1 m. The sides of the cell are electrodes which shape the electric field on two sides. The electrodes also function as readout strip. One chamber consists of 2$\times$4 such layers: four layers with the wire parallel to the beam line and 4 layers with the wire transverse to the beamline. The chamber is filled with an Ar and CO_2 gas mixture.

The muon rates in the endcap regions are higher and therefore a faster detector is required. CSCs, having a shorter drift path than DTs are appropriate for this. Each CSC consists of 6 layers. Each layer consists of a gas volume with an anode wire (along ϕ) operated at 2.9 kV or 3.6 kV and perpendicular to the wire cathode strips (along r). CSCs range in length from 1.7 to 3.4 m in the radial direction and span either 10 or 20° in ϕ. A single CSC layer contains 80 cathode strips with a width between 2.2 and 4.7 mrad in ϕ. The orthogonal anode wires have a diameter of 30 or 50 μm, are separated by 2.5 up to 3.16 mm and are arranged in groups of 5 to 16 wires. The chambers are filled with a mixture of CO_2, Ar and CF_4.

Besides the DTs and the CSCs which provide a fine spatial resolution a muon detector needs to be present with an $\mathcal{O}(1\text{ ns})$ time resolution. This is achieved by introducing RPCs in the system which allow for almost unambiguous identification of the correct bunch crossing and the time coincidence of track segments from the different muon subsystems. The RPCs used at CMS have two 2 mm gas gaps, with at both sides a 2 mm thick resistive Bakelite plate coated with a graphite layer to which a voltage of 9.6 kV is applied. The gas which is used is a mixture of $C_2H_2F_4$, i-C_4H_{10} and SF_6 to which water vapour is added to obtain a relative humidity of 40-50%. The readout strips on which the charge of an ionizing event is induced are placed between the two gaps and lie along η.

With the combination of high spatial resolution DTs and CSCs on one hand and high time resolution RPCs on the other hand CMS features two complementary inputs for the trigger system. The large number of layers per DT or CSC chamber is used to construct track segments within the chambers. These track segments can be used to set sharp transverse momentum cuts up to 100 GeV/c. For higher momenta the information from all muon systems needs to be combined.

4.3.5 The trigger system

A back-of-the-envelope calculation shows that the collision rate at CMS is of that order that it is technically impossible to store the result of each collision to disk: at a bunch crossing rate of 40 MHz and a 2 MB event size a storage capacity of 800 TB/s would be required.

To cope with this, only high energy head-on interactions, which are the ones likely to reveal new physics, should be saved to disk for offline analysis. CMS does this reduction in data rate in two stages: the Level 1 (L1) and high-level (HLT) trigger, which bring back the trigger rates to respectively a maximum of $\approx$ 100 kHz and $\approx$ 1 kHz.

The L1 trigger system is implemented in hardware and receives input from the calorimeter and muon detectors. The information which is used as input to the L1 trigger is typically a subset of the full-event information from which already a first fast reconstruction can be performed into potentially interesting objects such as high energy muons, electrons or photons, τ leptons, missing energy or jets. This reconstruction has to be done

in a very limited fixed time budget during which the full-event information is stored in the on-detector electronics pipelines. The time in which the decision is made to read out the full-event information is referred to as the *L1 latency*. For CMS this is a time window of ≈ 4 μs which also includes the latency of taking the data out of the detector to the counting room and sending the L1A back to the on-detector electronics. The L1 latency should thus be shorter than the depth of the pipelines in the front-end electronics and should also be fixed so that the front-end electronics on reception of an L1A signal, sent over the TTC (trigger, timing and control) link, can respond by outputting the full data stored at a given location in the pipeline.

As briefly mentioned in the previous sections, the input which is used for the L1 trigger system are the energy deposits from over 8000 ECAL and HCAL towers, track segments formed by the front-end electronics on the DTs and CSCs, and the hit information from the RPC system. Through several stages of first combining the calorimeter, RPC, CSC and DT information in dedicated subdetector trigger systems, all information is combined in the global trigger where a decision is made whether or not to reject the event. A schematic illustration of the system is shown in Figure 4.9.

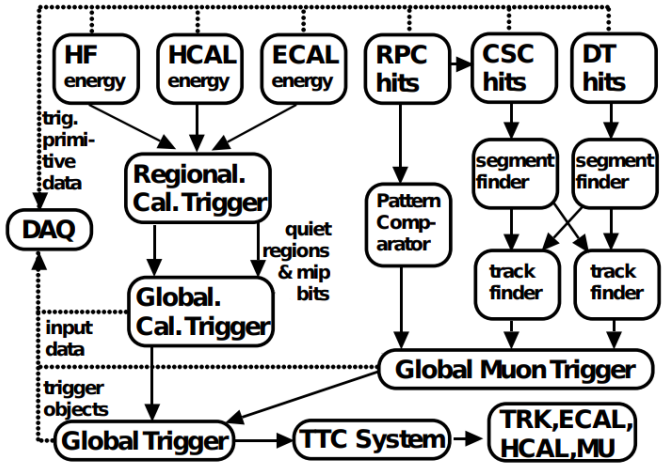


FIGURE 4.9: Illustration of the L1 trigger system in CMS [55].

The full-event information is subsequently used in the HLT, which implements several software algorithms running on a farm of commercially available computers to further improve the purity of the physics objects. Tracking for example becomes also possible at the HLT. HLT algorithms can be very flexible which allows for dedicated triggers to be developed for specific signals. In the end the HLT makes the final decision on which events will go for offline storage. First the data is stored locally on disk and then later transferred to Tier-0 computing centres where another offline reconstruction happens and where the data is stored permanently [55].

4.4 The HL-LHC upgrade

In order to maintain scientific progress and exploit the full LHC potential the LHC will be upgraded to the HL-LHC [38] which will provide the experiments with an increased instantaneous luminosity, a factor 2.5 higher than during the nominal operation in 2018.

The goal of the HL-LHC is to deliver to the experiments an integrated luminosity of 250 fb^{-1}/year, compared to for example the $\approx$ 68 fb^{-1} delivered to CMS in 2018. This will be obtained by increasing the pileup, from nominal mean values around 37 for the 2018 run to values of 150 at the HL-LHC. This will result in a peak luminosity of 5×10^{34} $cm^{-2}s^{-1}$ and an end of life total integrated luminosity of 3000 fb^{-1}. In the ultimate performance scenario the HL-LHC could deliver a pileup of 200 resulting in an instantaneous luminosity of 7-7.5$\times10^{34}$ $cm^{-2}s^{-1}$ and an integrated luminosity of up to 300-350 fb^{-1} per year which, by the end of life of HL-LHC could result in a total integrated luminosity of 4000 fb^{-1}.

All this will be achieved technically by several upgrades. Most of these upgrades are related to making existing systems robust against longer operation in an increased radiation field, preparing the accelerator optics for higher beam intensities, upgrading cryogenic systems,... Other upgrades are new concepts specifically engineered for the HL-LHC. Two of the latter are listed here:

- In contrast to the LHC, where the instantaneous luminosity continuously drops during a run due to the decrease of particle density in the bunches, the HL-LHC will make use of *luminosity levelling* where in the first part of the run the luminosity will be kept constant by changing the focussing of the beam at the interaction points. Like this the maximum workable pileup for the experiments can be maintained for a longer time.

- A new optics scheme will allow for a reduction of the β^* of the LHC beam. This requires even stronger quadrupole magnets with a larger aperture. In this configuration the flattened bunches collide under an angle which would result in a decrease in luminosity simply due to geometrical reasons. Specific radio-frequency cavities, so-called crab cavities, are being designed which use electrical fields to rotate the bunches right before collision, so that their geometrical overlap is maximised right at the interaction point.

Due to the increase in luminosity the experiments will be more sensitive to rare and statistically limited Standard Model signals [56]. An example of such a Standard Model measurement is the measurement of the Higgs self coupling which can be probed in events where Higgs bosons are produced in pairs. This measurement needs high statistics as the cross section of these events is about a factor of 1000 smaller than for single Higgs boson production. Furthermore the determination of the Higgs boson coupling to muons, τ, b and t quarks and W and Z bosons will also benefit from larger dataset.

Also beyond the standard model searches will reach higher significance with larger datasets. Due to increased statistics the searches will be able to probe smaller couplings with increased significance. Furthermore, rare channels with low production cross section or small branching ratios will become available.

4.5 The CMS Phase-2 upgrade

CMS subdetector upgrades and replacements will be performed in order to guarantee good quality physics data taking during the HL-LHC phase. Furthermore, several subdetector systems will be extended and new detectors such as the MIP Timing Detector (MTD) and Gas Electron Multipliers (GEMs) will be installed. Below, the major upgrades to the CMS subdetectors will be described. This upgraded version of CMS is referred to as *CMS Phase-2*.

A major requirement, which drives a substantial part of the planned upgrades, is maintaining the performance of the L1 trigger system. Therefore an increase in L1 latency to 12.5 μs and an increased L1A rate to 750 kHz is required. This puts direct requirements on the on-detector electronics as the buffer depths and the output bandwidth need to increase. In addition the new Phase-2 Outer Tracker will be designed in such a way that tracker information can be used in the L1 trigger system. Tracking information is required to keep the L1A trigger rate acceptably low with sharp turn-on curves (see Figure 5.31). Also other subdetectors will be upgraded in order to provide better input to the L1 trigger system.

The CMS Phase-1 pixel detector will be replaced by a new pixel detector [33] for operation at the HL-LHC. This Phase-2 pixel detector will feature 25×100 or 50×50 μm^2 (a factor 6 smaller than the Phase-0 and Phase-1 pixels) 100-150 μm thick (compared to 285 μm for Phase-0 and Phase-1) pixels. The pixel sensors are integrated into a pixel module where the pixels of a single pixel region are bump bonded to an RD53 chip. The module design has to withstand total doses up to 12 MGy, hadron fluences up to 2.3×10^{16} n_{eq}/cm^2 and hit rates up to 3 GHz/cm^2. Despite these unprecedented hit rates, the increased granularity will still result in a per pixel occupancy around the per mille level.

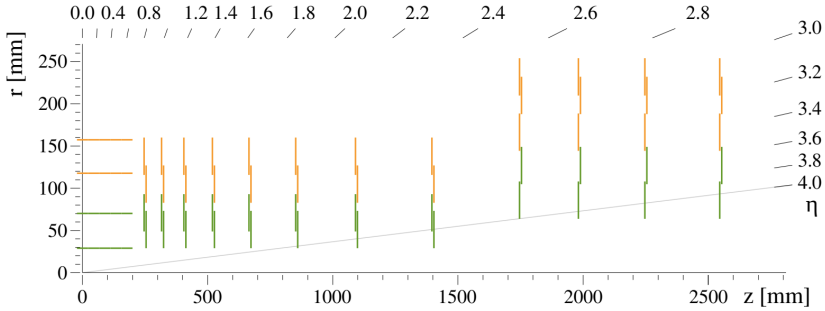


FIGURE 4.10: Layout of the Phase-2 pixel detector. Green and yellow lines respectively represent modules containing two and four readout chips [33].

Comparing the layout of the Phase-0/1 pixel detector, shown in Figure 4.6b, to the layout of the Phase-2 pixel detector, shown in Figure 4.10, immediately shows that the Phase-2 detector features a substantial extension of the pixel system. An extensive expansion in z is foreseen: going from 3 to 12 disks at each z end, like this covering up to $|\eta| \approx 4$ and $z = \pm 2.5$m.

For the CMS Phase-2 upgrade the complete strip tracker will be replaced with a new Outer Tracker. This upgrade will be discussed in detail in Chapter 5.

The EB on-detector and back-end electronics will be upgraded [57] to accommodate the requirement for a higher latency and L1A rate. The upgraded front-end electronics will allow a finer time resolution ($\approx$ 30 ps for photons with E $>$ 30 GeV) which will also mitigate the noise problems currently affecting the performance. The operating temperature of the EB will also be reduced from 18°C to 9°C to further reduce the increase in electronic noise expected by irradiation. The system will furthermore provide input to the L1 trigger with higher granularity: instead of using the 5×5 crystals as a unit for the trigger primitives the single crystal information will be made available to the L1 system.

The HB detector is refurbished [57] already in Long Shutdown 2 where the photodiodes will be replaced by SiPMs (Silicon PhotoMultipliers). The higher photo-detection

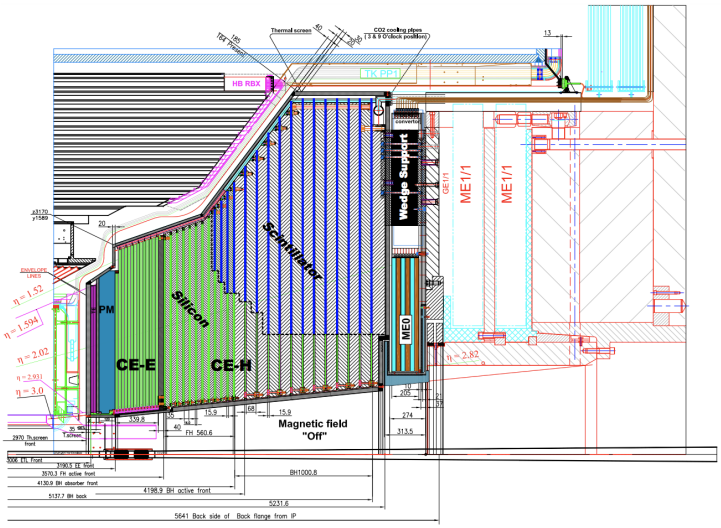


FIGURE 4.11: Layout of the HGCAL endcap calorimeter system for the CMS Phase-2 upgrade [58]. Both the electromagnetic (CE-E) and hadronic calorimeter (CE-H) are shown. The colour codes show the technology used: silicon sensors for the innermost part and scintillators for the more outward part. The full system will be operated at -30°C.

efficiency of the SiPMs with respect to the currently used photodiodes will compensate the loss in scintillator signal due to radiation. Also, to manage the higher rates an upgrade to the back-end electronics will be performed.

A complete replacement of the endcap calorimeter system is foreseen for Phase-2 [58]. Both the current EE and the HE will suffer too much radiation damage to provide good physics data at HL-LHC with the current detector. Therefore a new calorimeter will be installed, the so-called HGCAL (high granularity calorimeter). A detailed layout of the HGCAL system is shown in Figure 4.11. Hexagonal silicon sensors with a width of 163 mm will constitute the active material for the innermost part of the upgraded detector whilst the part of the HGCAL at higher z and smaller $|\eta|$, subject to lower radiation levels, will be equipped with plastic scintillator tiles read out by SiPMs. The longitudinal structure of the calorimeter is fairly complex and is a stacking of different layers of cooling, sampling and readout. In total there will be 28 and 24 longitudinal samplings for the electromagnetic and the hadronic part respectively. The silicon cell size is 0.5 to 1 cm^2 and the plastic scintillators have a segmentation of 4 to 30 cm^2. Together with the high resolution timing information (30-40 ps for hadrons with $p_T > 5$ GeV/c) that the silicon detectors bring, a highly segmented calorimeter like this will allow for accurate shower reconstruction, pileup rejection and vertex identification.

The current design of the CMS muon detectors were tested at HL-LHC conditions and were still found to be adequately functioning, despite the higher hit rates and possible ageing effects. As a result no major replacements of detectors have to happen, but electronics upgrades and extension of the muon system are planned and summarised briefly below [59].

The on-detector electronics of the DTs will be replaced by more performant ones which will allow higher latency and trigger rates up to 750 kHz. Radiation hardness

of some of the electronic components will be increased. Furthermore, most of the logic which generates input for the L1 trigger system will be moved from the on-detector to the back-end electronics so it becomes easily accessible and more advanced triggering algorithms can be implemented.

The CSCs, located in the endcaps, cope with even higher radiation levels than the DTs. Parts of the on-detector electronics will be replaced to make them more radiation hard. Higher speed output optical links will be implemented to handle the higher L1A rate and chamber occupancies and buffer sizes will be increased to allow for larger latencies. The back-end DAQ (Data AcQuisition) board will also be upgraded to cope with the higher rates.

The RPC system's off-detector *link system*, which synchronises and compresses the data from the on-detector electronics and forwards it to the L1 trigger system, will be replaced. This upgrade will result in a higher time resolution for input to the L1 trigger. Besides the upgraded electronics, the RPC system will be extended towards higher η (from $|\eta| < 1.8$ to $|\eta| < 2.4$), which is a challenging region for muon reconstruction due to high backgrounds and where until now only CSCs provide information to the L1 trigger. An improved RPC (iRPC) design has been developed using High Pressure Laminate as electrodes and a smaller gap and electrode thickness compared to the current RPCs. The on-detector electronics for these iRPCs are also custom tuned to these new chamber's properties and allow to read out the strips at both ends which gives access to the time difference between the hits so that a higher spatial resolution can be attained. In addition to the iRPCs, GEM (Gas Electron Multiplier) detectors will be installed in the forward region. The location of the iRPC and GEM chambers is illustrated in Figure 4.5.

A completely new subdetector, the MIP timing detector [60], is also foreseen to be installed for the HL-LHC operation. The basic idea behind this detector is to exploit the fact that not all pp interactions during a single BX happen at exactly the same time: there is a time window with an RMS spread of approximately 180-200 ps wherein the pileup interactions happen. It is the goal of the MTD to measure timing of hits from MIPs with a time resolution of 30 ps. Like this the beamspot can be sliced in consecutive exposures of 30 ps, which would bring down the maximum pileup of 200 to values close to the current mean pileup. This will greatly ease the assignment of tracks and energy deposits to the correct primary vertex. The upgraded EB and HGCAL also provide timing information, but do not provide this information for MIPs. The MTD will be implemented as a dedicated detector layer in between tracker and calorimeters consisting of a barrel ($|\eta| < 1.5$) part and two endcap (up to $|\eta| = 3$) parts. The barrel layer will be equipped with crystal scintillators (12×12 mm^2) read out by SiPMs whilst LGAD (Low-Gain-Avalanche-Diodes) silicon sensors will be used in the endcaps. The baseline is that the MTD will be read out on reception of an L1A. Full read out of the MTD for use in the L1 trigger is not feasible due to bandwidth constraints, but it is considered that MTD regions of special interest could be pre-triggered by a so-called *Level 0* request from the track, calorimeter or muon trigger. In this scheme timing information could be used in a later step of the L1 decision.

The event size going to the HLT for the full CMS Phase-2 detector is 7.4 MB [61], compared to 2 MB during Run 2. At a trigger rate of 750 kHz this requires an event builder that can maintain a throughput of 44 Tb/s in order for the HLT to reduce the offline storage rate from 750 kHz to 7.5 kHz in the end resulting in 5.3 PB of stored data per day of operation.

4.6 Summary

CMS is a general purpose detector located at one of the four interaction points along the LHC ring. CMS consists of several subdetector systems, sitting in a magnetic field, which provide an as complete as possible picture of an event. One of these subdetectors is the tracking system, which is positioned closest to the interaction point and which is designed to measure the tracks of charged particles. This detector has to be light in order to reduce multiple scattering. The CMS tracker consists of several layers of silicon modules, of which the hit information is used by the tracking software to reconstruct the tracks. The subsequent calorimeter and muon system are respectively designed to measure the particle's energy and to reconstruct tracks from high energy muons. The latter two subdetectors provide input to the L1 trigger system which decides in a short and fixed time whether the event is interesting. If the event is considered interesting, the full-event data is requested from the on-detector electronics and used in the HLT system which decides if the event is kept for offline storage. This two-level trigger system reduces the event rate from 40 MHz to $\approx$ 1 kHz.

To maintain scientific progress, the full LHC potential will be exploited by the HL-LHC upgrade. Several improvements will be made to the accelerator complex which will allow LHC to deliver a significantly higher luminosity to the experiments. This has severe implications on the operation of CMS as it will have to cope with an even harsher radiation environment and higher pileup. A major challenge will be to guarantee sharp trigger turn-on curves so that low enough trigger rates can be maintained.

To assure performant operation of CMS throughout the full HL-LHC running, CMS is preparing to perform several upgrades: on-detector electronics will be improved to allow for a larger L1 latency and trigger rate, subdetector systems will be replaced and even completely new subdetectors will be installed to help mitigate the harsh pileup conditions. One of the major upgrades is the replacement of the full strip tracker system. This upgrade is the topic of the next chapter.

Part II

The Upgraded Outer Tracker for the CMS Detector at the High Luminosity LHC

Chapter 5

The CMS Phase-2 Outer Tracker

5.1 Introduction

The current CMS strip tracker, designed to operate at an average pileup of 25 and up to an integrated luminosity of 500 fb^{-1} [62] is coping well with the above design pileup values of Run 2. The strip tracker is however suffering from performance degradation due to radiation damage which will in the end result in a tracker which cannot be operated any more above 1000 fb^{-1}. The main reason for malfunctioning is an increase in the leakage current and depletion voltage of the sensors. As a result for example, the efficiency for reconstructing tracks with transverse momentum larger than 0.9 GeV/c in $t\bar{t}$ events would drop from $\approx$ 90% to $\approx$ 50% at $|\eta| = 1.5$ when going from a tracker with no ageing at a pileup of 50 to an aged tracker at 1000 fb^{-1} operated at a pileup of 140. While the efficiency for track reconstruction drops, the single track fake rate for this specific case would go up from $\approx$ 5% to $\approx$ 45%. Furthermore, the current front-end electronics do not allow triggering at 750 kHz and the maximum latency budget is only 4 μs [63].

As a result the current strip tracker needs to be replaced with a new *Outer Tracker* (OT) during Long Shutdown 3. The goal of this CMS Phase-2 tracker is to provide accurate tracking performance (see section 5.10), equal or better than the current tracker, in the harsh high pileup environment which the HL-LHC will provide. The Phase-2 tracker has to cope with higher particle fluxes and higher radiation levels than the current tracker and furthermore has to provide input to the CMS L1 trigger system. The full tracker is designed to be operated at a temperature of -30°C to reduce radiation damage.

To provide a better tracking performance the material budget of the new tracker system will be greatly reduced compared to the current tracker. The different contributions

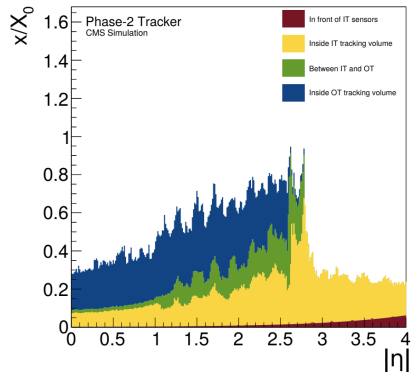
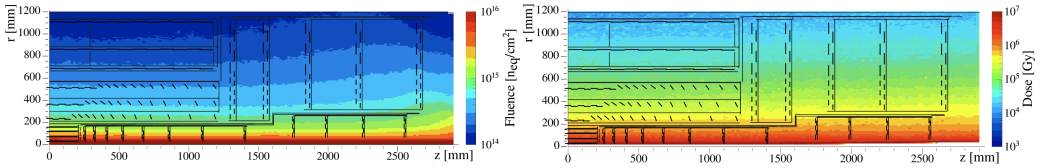


FIGURE 5.1: Simulated material budget of the Phase-2 tracking systems in units of radiation length [33].



(a) (b)

FIGURE 5.2: Expected integrated particle fluence (LEFT) and total integrated dose (RIGHT) for the CMS tracker environment after 3000 fb^{-1} of pp collisions at 14 TeV centre of mass energy [33].

to the material budget, extracted from simulation, are shown in function of η in Figure 5.1. Both track reconstruction and calorimeter performance will greatly benefit from this reduction in material budget. An important factor in reducing the material budget is the new geometry of the Outer Tracker, described in section 5.3.

The requirement for tracking information for the L1 trigger requires a unique OT module design with on-module transverse momentum measurement capabilities. These so-called p_T *modules* will be described in sections 5.2. The p_T modules will be able to discriminate between particles with a p_T lower and higher than a given threshold. The nominal threshold p_T is 2 GeV/c and in this baseline configuration only track primitives, so called *stubs*, from particles passing the 2 GeV/c threshold are passed on to the detector back-end, at BX frequency, to be used in the L1 track finder. Like this the fact that only 3% of the tracks at 14 TeV pp collisions have a p_T larger than 2 GeV/c is exploited to keep the Outer Tracker L1 bandwidth within limits [64]. The full Outer Tracker comprises $\approx$ 200M channels which is a substantial increase compared to the $\approx$ 10M channels in the Phase-0 strip tracker.

The expected integrated particle fluence and total integrated dose for the Outer Tracker environment at 3000 fb^{-1} HL-LHC operation are shown in Figure 5.2a and 5.2b respectively. The upgraded Outer Tracker is designed to be fully efficient up to 3000 fb^{-1}, with a 50% margin, without the need for maintenance interventions. The radiation levels in the Outer Tracker volume, going up to 10^{15} n_{eq}/cm^2 integrated fluence and 600 kGy integrated dose, form a particular challenge for the operation of the silicon sensors and the front-end electronics. The sensors and the front-end electronics will be the subject of section 5.4 and 5.5 respectively. The on-module electronics will be discussed in detail, as a good understanding of their operation and how they are implemented on the module (section 5.6) is important for the rest of this work.

How the p_T modules are planned to be built and qualified is briefly described in section 5.7. A basic understanding of the assembly procedure and qualification is essential to frame the work described in later chapters. How the DAQ and the L1 track trigger of the final system will look is introduced in section 5.8 and section 5.9. Some expected performance results of the Phase-2 tracker are discussed in section 5.10.

5.2 p_T module concept and design

The requirement for an L1 track trigger necessitates a module design with on-module track reconstruction capabilities. This will be accomplished using the p_T *module design*. The basic idea is to use two parallel silicon sensors, separated by a few mm, in one single

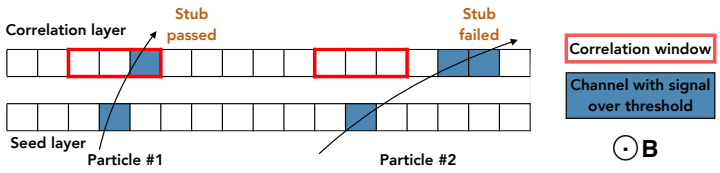


FIGURE 5.3: Two particles generating a hit in both the seed and correlation layer. Particle #1 generates a hit in the correlation layer within the correlation window (indicated by the red squares) of which the center is defined by the hit in the seed layer. This hit pair is recognised as a stub. The hit from particle #2 in the correlation layer falls outside the correlation window and thus does not result in a stub.

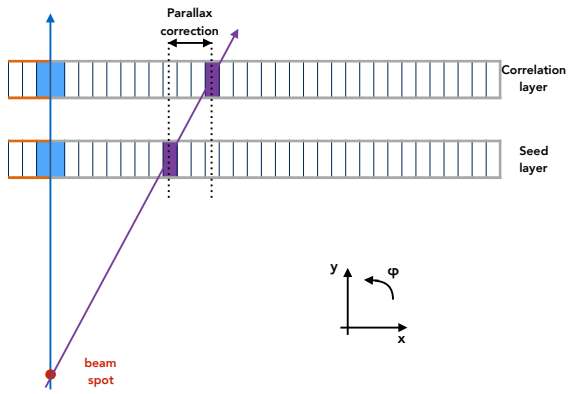


FIGURE 5.4: Both module types implement a parallax correction in the front-end chips, as explained in the text.

module. The signal routing on the modules is such that the front-end ASICs receive the data from both the top and bottom sensor and can thus correlate hits from both layers. This concept is illustrated in Figure 5.3 where two charged particles pass through both sensor layers and bend in the magnetic field: the ASIC records a hit in one of the two layers chosen to be the *seed layer* and then looks for a corresponding hit in the *correlation layer* within a programmable *correlation window*. If a hit is found within the correlation window, a *stub* is formed and the chip will output the stub location and bend[1]. The size of this correlation window directly maps to a certain p_T threshold for the stubs. These stubs serve as input to the L1 track trigger system from which the track finder reconstructs high p_T tracks. A programmable correlation window is required to set a uniform p_T threshold for the full Outer Tracker volume.

The p_T module design, with planar sensors, inherently suffers from parallax errors as explained in Figure 5.4: two particles, with infinite p_T originating from the beamspot, are shown passing through the two sensors of a module. The stubs which are formed should have a bend value of zero, but only particles passing the module at a location where the

[1]Given some requirements, which are described in section 5.5, are satisfied.

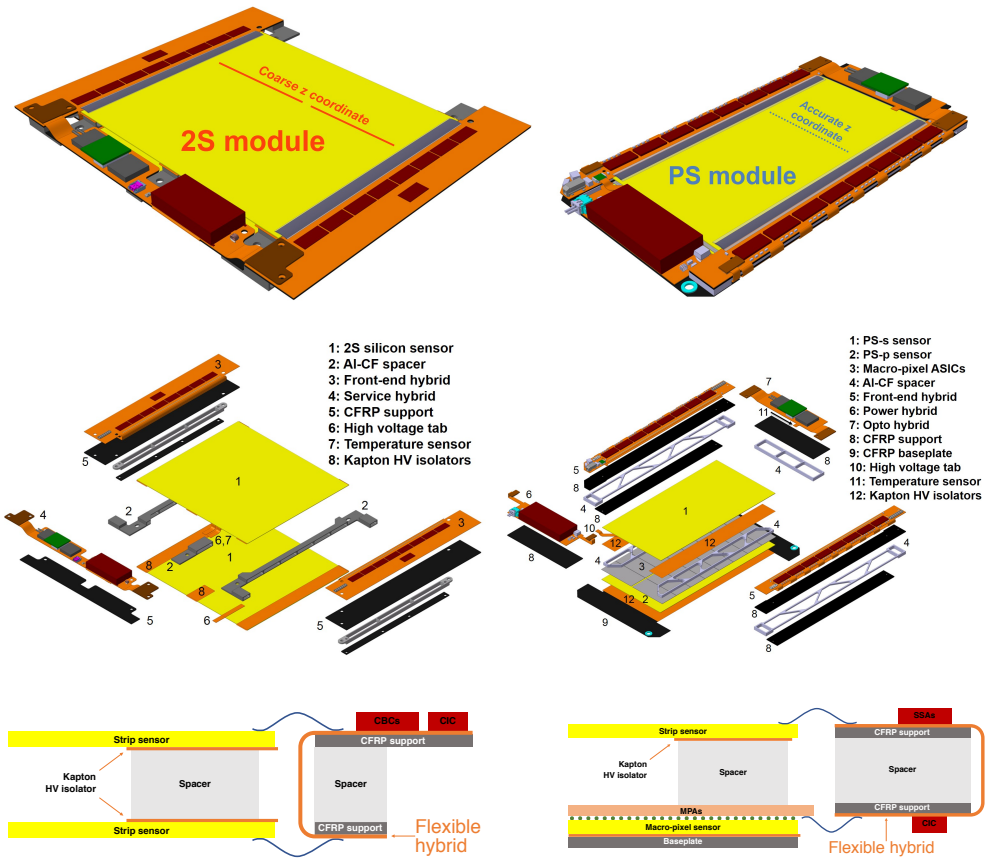


FIGURE 5.5: Assembled and exploded views of the 2S (TOP and MIDDLE LEFT) and PS module (TOP and MIDDLE RIGHT) showing the different components of both p_{T} module designs. The bottom row shows a transversal cut through the module and shows how the front-end ASICs are connected to the sensors [33].

normal vector on the sensor plane is pointing to the beamspot will give rise to stubs with zero bend. To recover from parallax errors for particles not hitting at this specific location, the sensors can be artificially shifted with respect to each other using parallax correcting logic in the SSA and CBC (see section 5.5).

The Phase-2 Outer Tracker tracker will adopt two major module designs: the pixel-strip (PS) module consisting of a macro-pixel sensor and a strip sensor and the strip-strip (2S) module consisting of two identical strip sensors. The PS modules are designed to be operated in the innermost parts of the Outer Tracker, hence their higher granularity, whereas the 2S modules will be used in the outer parts. The 2S modules come in 2 variants with 1.8 or 4.0 mm sensor spacing and the PS modules come in 3 variants with 1.6, 2.6 or 4.0 mm sensor spacing. The different values for the sensor separation arise from the requirement to have a uniform p_{T}-cut across the full Outer Tracker volume, but still minimize the number of module variants.

5.2.1 The PS module

The PS module is shown on the right in Figure 5.5. It shows the PS-pixel (PS-p) sensor and the PS-strip (PS-s) sensor. Both sensors have an active area of $\approx 5 \times 10$ cm^2. The 32×960 macro-pixels of the PS-p sensor are bump bonded to a total of 2×8 Macro Pixel ASICs (MPA[2]) together forming a so-called Macro Pixel Sub Assembly (MaPSA). Each macro-pixel has a size of 1.5 mm $\times$ 100 μm.

The 2×960 2.5 cm long PS-s strips are wire bonded to 2×8 Short Strip ASICs (SSA) which sit alongside both sides of the sensor on the *front-end hybrids* (FEH) together with the CIC (Concentrator Integrated Circuit) chips. The strip pitch of 100 μm matches the width of the macro pixels. The neighbouring SSA chips on a hybrid are interconnected in order to pass hit data from one chip to the other, thus making the sensor continuous for stub finding.

Spacing of the sensor is done with Al-CF spacers with a thin kapton layer on both sides to which the sensors are glued. Different spacer thicknesses will be used for the different sensor separations and the thickness of the FEHs will be matched accordingly.

How the interconnection of chips is achieved in the module is shown in the transversal cut at the bottom of Figure 5.5: the SSA, receiving the sensor signal over wire bonds, sends data to the MPA chip over the hybrid and wire bonds. The MPA receives the signals from the pixels, forms the stubs by correlating them with the SSA clusters, formats the full-event data package and sends the data over another set of wire bonds to the CIC chip. The full module data path will be discussed in section 5.6.

Both front-end hybrids are connected to two additional hybrids. One hybrid, the *power hybrid*, carries a DCDC converter providing the correct LV (low voltage) power levels for the front-end ASICs and also facilitates the HV (high voltage) connection. The other hybrid, the *opto hybrid*, carries the lpGBT (low-power GigaBit Transceiver) and VTRx+ (Versatile Link Plus) chip for serialisation and conversion from the electrical to the optical domain.

The cooling of this module will be done using a cooled baseplate to which the module is glued.

5.2.2 The 2S module

The 2S modules, shown on the left in Figure 5.5, consist of two identical strip sensors (10×10 cm^2) with 2×1016 5 cm long strips on each sensor. The strips, with a pitch of 90 μm, are wire bonded to the CBCs (CMS Binary Chips) which sit on two front-end hybrids along the sensor. In total 16 CBC chips, 8 on each FEH, are used to read out the full 2S sensor. Each CBC chip receives data from 127 top and bottom channels by passing the signal over the flexible hybrid as illustrated at the bottom left in Figure 5.5. The CBCs also receive data from their neighbours in order to make the module uniform for stub finding. The CBC data is then sent to the CIC chip which sends the data to the *service hybrid* (SEH), which carries all the components for LV and HV powering and data treatment. The components (DCDC converter, lpGBT, VTRx+) on this service hybrid are the same as used in the PS module.

As in the design of the PS module, the sensors are separated using an Al-CF spacer with kapton isolator. Different spacer thicknesses will be used for the different sensor separations and the thickness of the FEHs will be matched accordingly.

[2]The front-end chips of both PS and 2S modules are discussed in more detail in section 5.5.

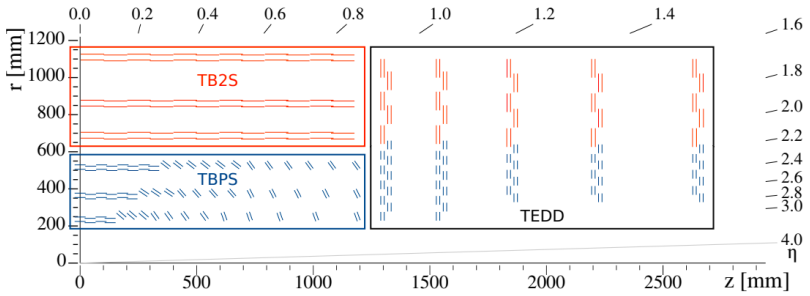


FIGURE 5.6: Layout of one quarter of the Phase-2 Outer Tracker. The blue and red lines represent respectively the PS and 2S modules [33].

The cooling of this module is obtained through five mounting points at the sides of the sensor spacers.

5.3 Outer Tracker geometry

How the 7680 2S and 5616 PS modules are arranged to form the Outer Tracker is illustrated with a cross section view in Figure 5.6. The Outer Tracker consists of a barrel region and two endcaps. Three distinct parts are defined:

- TBPS: the barrel region equipped with 3 layers of PS modules. A tilted module geometry is used for modules at larger z. The tilting of the modules reduces the total number of modules required for the Outer Tracker and also eliminates stub inefficiencies introduced by particles which pass module or sensor edges.
- TB2S: the barrel region equipped with 3 layers of 2S modules.
- TEDD: the endcap regions consisting of 5 double discs carrying both PS and 2S modules.

The different modules, with different sensor spacings, will be distributed across the Outer Tracker as shown in Figure 5.7. The correlation window configuration resulting in a uniform p_{T} cut of 2 GeV/c is also illustrated in this figure.

In both the barrel and endcap regions the layout was carefully designed to mitigate dead areas by letting modules overlap by slightly staggering them on different surfaces within the layer. This guarantees hermetic coverage of each layer. The modules are arranged in *plank* (TBPS) and *ladder* (TB2S) structures parallel to the beam as shown in Figure 5.8a. The orientation of the strips and macro-pixels are also illustrated in this figure. In the endcaps, illustrated in Figure 5.8b, the modules are arranged in double discs where each disk has modules on both sides.

5.4 Silicon sensors

After an extensive R&D [65] project, n-in-p sensors were found to be more radiation hard than p-in-n sensors. As a result, n-in-p sensors, with electron readout, and a thickness of 320 μm will be used to equip the p_{T} modules. P-spray and p-stop structures will ensure sufficient inter-strip isolation [65].

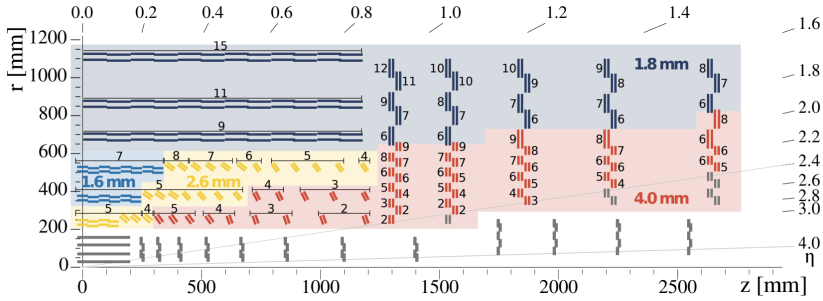


FIGURE 5.7: The 2S modules are designed in two variants: with 1.8 mm and 4.0 mm sensor spacing. The region of the tracker where these types will be placed are indicated respectively in dark blue and red. The PS modules are designed in three variants: 1.6 mm, 2.6 mm and 4.0 mm sensor spacing. These are indicated respectively in light blue, yellow and red. The numbers close to the modules indicate the acceptance windows which will be used to obtain a p_T cut > 2 GeV/c [33].

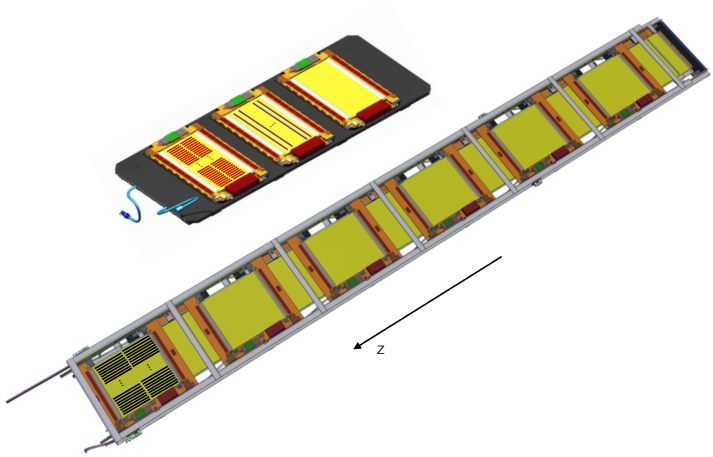
The basic characteristics of the 2S and PS sensors were already given in section 5.2.1 and 5.2.2. For the 2S and PS-s sensors a bias ring which surrounds the active area distributes the ground potential to each strip, the sensors are designed to be biased up to a maximum of -800 V from the sensor's backside and the strips are AC-coupled to the front-end electronics. The PS-p sensor is DC-coupled to the MPA front-end. Using a common bias grid and punch-through structures each macro pixel is grounded and the sensor is biased from the backside.

Tests have shown that signals of about 12 ke$^-$ (9 ke$^-$) can be extracted after irradiation to 2 (1) times the nominal end of life fluence for the 2S (PS) modules and this for sensors with an active thickness of 300 μm at a bias voltage of -750 V. This is well above the 1 ke$^-$, 0.7 ke$^-$ and 0.2 ke$^-$ RMS noise expected for the 2S, PS-s and PS-p sensors respectively [33]. The increase in leakage current due to irradiation damage was evaluated and the extra power was found to be within cooling power limits.

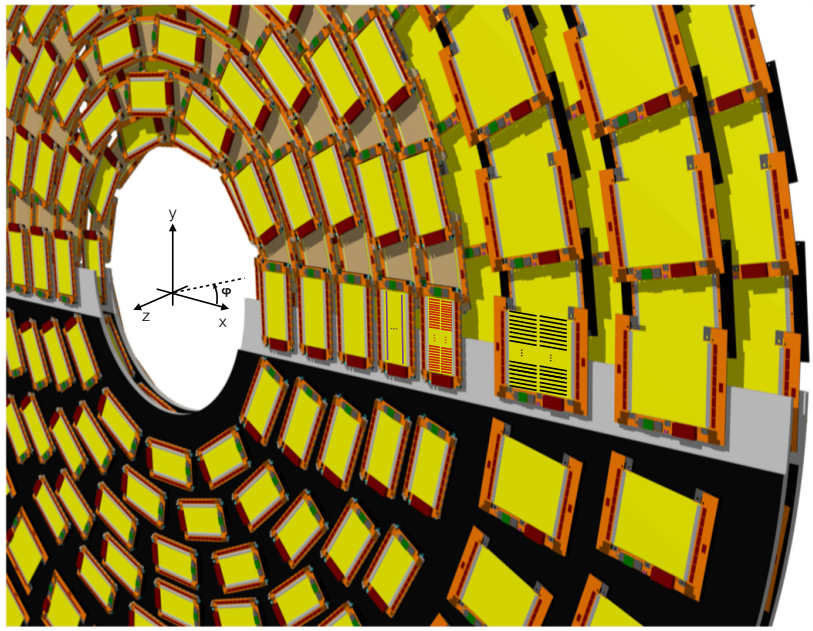
5.5 p_T module ASICs

The two different p_T module designs and the unique requirement for these modules to generate stubs implies the need for dedicated readout chips for each module type. The PS modules, residing in a higher occupancy regime than the 2S modules, consist of a pixelated and a strip sensor. This already requires two different readout chips for the PS module. The SSA chip receives the signal from the strip sensor and sends the data towards the MPA, which receives the data from the pixel matrix and from the SSA to form stubs and full-event data packages. The 2S modules, consisting of two identical strip sensors, are read out by CBC chips, which also build the stubs. The CIC chip is a common chip for use in 2S and PS p_T modules. It receives the data from 8 MPA or 8 CBC chips, buffers and forwards the full-event data and selects the stub data. Both data streams are then sent to the lpGBT chip for transmission to the back-end.

In the following sections the main working principles of SSA, MPA, CBC and CIC, relevant for the rest of this document, will be discussed. Some features are not discussed and much more information on the MPA, SSA, CBC and CIC chips can be found in the respective manuals Refs. [66, 67, 68, 69].



(a)



(b)

FIGURE 5.8: TOP: Example of the planks and ladders of PS and 2S modules from which the barrel part of the tracker will be build. BOTTOM: PS and 2S modules placed on the TEDD double discs. In both figures the orientation of the 2S strips are shown by the black lines. The purple lines give the orientation of the macro pixel boundaries in the coarse dimension and the red lines give the orientation of the fine dimension of the macro pixels and PS-s strips (adapted from [33]).

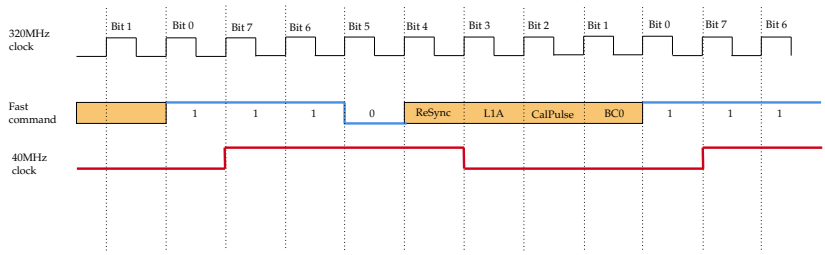


FIGURE 5.9: Fast command encoding: MPA, CBC, SSA and CIC receive a 320 MHz clock with which the fast command line is sampled. From the sync pattern of the fast command the chips extract the 40 MHz clock used for most internal logic on the ASICs.

5.5.1 Commonalities between the SSA, MPA, CBC and CIC ASICs

From the user point of view the OT ASICs share many commonalities:

- All chips are provided with a clock from the lpGBT chip or from the back-end if the lpGBT is not in the chain yet.
- The format of the commands to send instructions to the front-end chips at BX frequency, the so-called fast commands (or t1), is the same for each chip. The different flavours of chips respond however, although to a small extent, differently to the fast commands.
- The I^2C protocol is used to do the slow[3] control of the chip.
- The data output runs at 320 Mbps (except 640 Mbps for certain configurations of CIC2): there is a single line for the full-event data and 5 (MPA and CBC), 8 (SSA) and 5 or 6 (CIC) lines for the stub data. The exact data format amongst chip types differs and will be discussed in sections 5.5.2, 5.5.3, 5.5.5 and 5.5.6. All output lines are fully digital. As a note for the reader: in the rest of the work the SSA data which is made available at BX frequency will also be referred to as *stub data*. Physically these data do not represent stubs, they encode cluster positions.

The commonalities naturally follow from the fact that these chips will be implemented on the same module, data transfer between chip flavours is required, chips need to be running on the same clock and are required to act on the same fast commands. Another natural result is that the test bench for these devices can be very similar. This test bench will be described in Chapter 6 and results obtained with this test bench are discussed in Chapters 7, 8 and 9.

The fast commands are transmitted over one differential line, running at 320 MHz, from the lpGBT (final module) or directly from the back-end (during prototyping) to the MPA, SSA, CBC and CIC chip. The fast command encoding consists of 8 bits: three header bits ('110'), four payload bits and one trailer bit ('1'). The format is illustrated in Figure 5.9 and the definitions of the payloads are shown in Table 5.1. This table shows

[3]In contrast to the fast command operation which allows communication to the chip at 40 MHz.

Fast command (t1)	Bit 7 (MSB)	Bit 6	Bit 5	Bit 4	Bit 3	Bit 2	Bit 1	Bit 0 (LSB)
ReSync (Fast Reset)	1	1	0	1	0	0	0	1
L1A Trigger	1	1	0	0	1	0	0	1
Test Pulse Trigger (CalPulse)	1	1	0	0	0	1	0	1
BC0 (Orbit Reset)	1	1	0	0	0	0	1	1
BC0 & ReSync	1	1	0	1	0	0	1	1
BC0 & L1A Trigger	1	1	0	0	1	0	1	1
BC0 & Test Pulse Trigger	1	1	0	0	0	1	1	1

TABLE 5.1: Definitions of the fast command. The header and trailer are indicated in blue. All seven allowed fast command payloads are shown. The table also states some commonly used other terminology for the fast commands (t1).

all the allowed combinations of fast commands which can be sent to the chips. The action taken by each chip on reception of a fast command can be slightly different[4], but in general the purpose of the fast commands is as follows:

- ReSync: resets the FIFOs and the state machines, and except for the CBC the bunch crossing counter and the L1 counter are also reset.
- L1A trigger (or L1A): the SSA, MPA and CBC chip will respond by outputting the full-event data present at a certain location, corresponding to the latency setting, in the memory to the output FIFO.
- CalPulse: triggers the discharge of the test pulse capacitor in the chip's front-end. This feature is not relevant for the CIC. For the MPA and SSA there is also the possibility to inject data into the digital logic.
- BC0: resets the bunch crossing counter and the L1 counter, except for the CBC where only the L1 counter is reset.

From the header of the fast command the chips also extract the 40 MHz clock, as shown in Figure 5.9, which is used for most internal chip logic.

The I^2C is implemented according to the Philips Semiconductors I^2C v2.1 specifications with a general call address within a chip family and a maximal I^2C clock frequency of 1 MHz. In the final module there will be two I^2C buses, one for each hybrid. The lpGBT chip will be the master of both buses. One bus connects to 8 MPAs, 8 SSAs and 1 CIC chip in the PS case and 8 CBCs and 1 CIC in the 2S case. Each chip on the bus has a specific address which allows to address the chips separately. The I^2C protocol is a wide spread protocol and will therefore not be discussed in detail here. The full slow control specifications for the OT module ASICs can be found in Ref. [70]. One thing to note however is that the MPA, SSA and CIC implement 16-bit register addressing whereas the CBC implements 8-bit register addressing. The CBC makes use of a *paging* system to cope with the large amount of registers.

The CBC chip is implemented in 130 nm CMOS technology. The MPA, SSA and CIC1 are implemented in 65 nm CMOS technology. The CBC operates at a nominal voltage of 1.2 V. The MPA and SSA chip will be operated at 1 V for the digital power supply and 1.2 V for the analogue and the I/O (Input/Output). The CIC will operate at two digital

[4]For the MPA and SSA the interpretation of the fast commands also depends on the mode (synchronous or asynchronous) the chip is configured in (see section 5.5). The L1A marks the start of the counting for the ripple counters, whereas the BC0 marks the end. The sending of a BC0 together with a ReSync starts the read out of the counters, the combination of a BC0 and L1A resets all the counters.

voltages: 1 V in the PS mode and 1.2 V in the 2S mode. The I/O of the CIC is powered with 1.2 V in both cases.

The data pads with which the back-end interfaces are all according to the differential SLVS (Scalable Low-Voltage Signaling) standards. Also the clock and fast command line are implemented as differential SLVS. The I^2C and hard reset pad are implemented in CMOS standard. Of course there are many more pads on the chips and the full specification list can be found in the chip manuals.

5.5.2 The SSA chip

5.5.2.1 Introduction

Each SSA chip [67] is designed to: read out 120 strips of a PS-s sensor using a specifically designed front-end, digitize the sensor signal, provide full-event data on reception of a trigger and synchronous centroid data at 40 MHz to the MPA chip. In section 5.5.2.2 the front-end of the SSA chip will briefly be introduced. In section 5.5.2.3 the transition to the digital domain is explained and the digital part is then discussed in sections 5.5.2.4 and 5.5.2.5. Figure 5.10 shows the top-level architecture of the SSA chip.

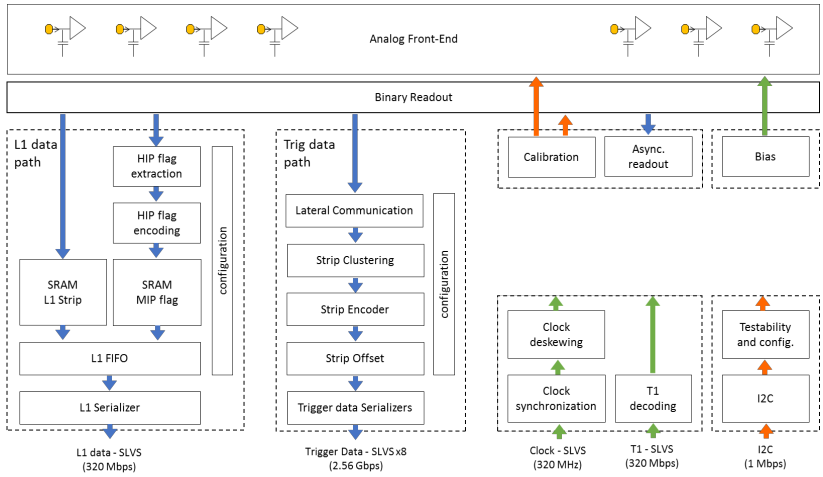


FIGURE 5.10: Top-level architecture of the SSA chip [67].

5.5.2.2 SSA front-end

120 separate front-ends are present in the SSA and are AC-coupled to one of the 120 PS-s sensor strips. Each front-end consists of a pre-amplifier, a booster and two discriminator stages. Two discriminator stages are present to provide the possibility to set a *HIP threshold* next to the detection threshold. A 5-bit trimming DAC (digital to analogue converter) is present to compensate for the channel-to-channel discriminator offset. Each channel input is connected to a 52 fF calibration capacitor. The amplitude of the calibration pulse generated by this circuit can be controlled by an 8-bit DAC. An edge sensitive circuit forms the transition to the digital logic of the chip and is described in the next section.

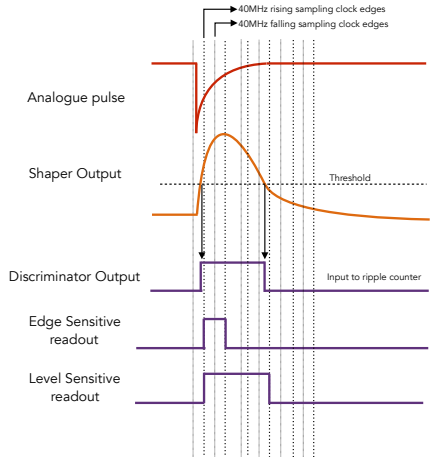


FIGURE 5.11: Illustration of the sampling logic of the SSA and MPA.

5.5.2.3 SSA sampling logic

The SSA uses a clock signal, the so-called sampling clock, to sample the incoming signal. The phase of the sampling clock with respect to the chip's 40 MHz internal clock can be configured over the full 25 ns range with a resolution of 200 ps using an internal DLL (delay-locked loop) and a coarse shift, based on the 320 MHz clock.

The discriminator output is then sampled in three different ways, as illustrated in Figure 5.11:

- Ripple counter: this asynchronous counter counts the number of pulses from the comparator stage during a user controlled window. A counter is present for each channel. This mode is especially useful for calibration purposes. These 15-bit counters can be accessed over I^2C or the SSA can be configured to output them on the stub data lines. In the latter case the 8 LSBs (Least Significant Bits) are transmitted on the first stub data line whilst the 7 MSBs (Most Significant Bits) are transmitted on the second stub data line. The counters are transmitted every 8th BX.

- Edge sensitive: in this mode comparator transitions which are non synchronous with the sampling clock are registered which guarantees no blind cycles. In this mode the output pulse has a fixed length of 1 BX.

- Level sensitive: the signal is sampled at the rising edge of the sampling clock. In case the discriminator output is longer than a single acquisition clock cycle the output of this sampling stage will also last multiple clock cycles. The maximum duration can however be configured and the output is forced to zero after the set maximum number of clock cycles.

The two latter modes or a logic OR or XOR of the two is used to feed into the digital logic (see section 5.5.2.4 and 5.5.2.5) and signifies the time of arrival of the analogue pulse with respect to the acquisition clock. The sampling setting is configurable for each front-end, but usually all front-ends are configured in the same mode. The masking of channels is also implemented at this level: a switch per channel is implemented to disconnect the front-end from the sampling logic.

Line 0	$C1_{Bit6}$	$C1_{Bit5}$	$C1_{Bit4}$	$C1_{Bit3}$	$C1_{Bit2}$	$C1_{Bit1}$	$C1_{Bit0}$	$C1_{Bit0.5}$
Line 1	$C2_{Bit6}$	$C2_{Bit5}$	$C2_{Bit4}$	$C2_{Bit3}$	$C2_{Bit2}$	$C2_{Bit1}$	$C2_{Bit0}$	$C2_{Bit0.5}$
Line 2	$C3_{Bit6}$	$C3_{Bit5}$	$C3_{Bit4}$	$C3_{Bit3}$	$C3_{Bit2}$	$C3_{Bit1}$	$C3_{Bit0}$	$C3_{Bit0.5}$
Line 3	$C4_{Bit6}$	$C4_{Bit5}$	$C4_{Bit4}$	$C4_{Bit3}$	$C4_{Bit2}$	$C4_{Bit1}$	$C4_{Bit0}$	$C4_{Bit0.5}$
Line 4	$C5_{Bit6}$	$C5_{Bit5}$	$C5_{Bit4}$	$C5_{Bit3}$	$C5_{Bit2}$	$C5_{Bit1}$	$C5_{Bit0}$	$C5_{Bit0.5}$
Line 5	$C6_{Bit6}$	$C6_{Bit5}$	$C6_{Bit4}$	$C6_{Bit3}$	$C6_{Bit2}$	$C6_{Bit1}$	$C6_{Bit0}$	$C6_{Bit0.5}$
Line 6	$C7_{Bit6}$	$C7_{Bit5}$	$C7_{Bit4}$	$C7_{Bit3}$	$C7_{Bit2}$	$C7_{Bit1}$	$C7_{Bit0}$	$C7_{Bit0.5}$
Line 7	$C8_{Bit6}$	$C8_{Bit5}$	$C8_{Bit4}$	$C8_{Bit3}$	$C8_{Bit2}$	$C8_{Bit1}$	$C8_{Bit0}$	$C8_{Bit0.5}$

(a)

Left Lateral	Hit_7	Hit_6	Hit_5	Hit_4	Hit_3	Hit_2	Hit_1	Hit_0
Right Lateral	Hit_{119}	Hit_{118}	Hit_{117}	Hit_{116}	Hit_{115}	Hit_{114}	Hit_{113}	Hit_{112}

(b)

FIGURE 5.12: LEFT: Stub data format as output by the SSA chip. RIGHT: Lateral data format as output by the SSA chip.

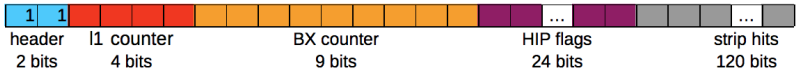


FIGURE 5.13: The full-event data as output by the SSA chip.

For testing purposes a digital pattern can be injected in the digital logic. This pattern is configurable per channel, can be enabled or disabled per channel, is configurable for both low and high threshold, is 8 bits long and is injected in the digital logic on reception of a test pulse trigger.

5.5.2.4 SSA stub data path and format

The SSA has two data paths: the stub data path and the full-event data path. Both data paths share the same same front-end and the same sampling logic. The stub data path performs clusterisation of hits in consecutive strips. The centroids formed in this clusterisation and which pass the configurable cluster width cut are sent to the MPA stub logic. Up to eight clusters per BX can be sent to the MPA. The format of the stub data is shown in Figure 5.12a. To guarantee continuity of the cluster data across SSA chips, lateral inter-chip communication is implemented between the SSAs: the hit information from the eight edge channels are sent to the adjacent SSA chip at every BX over a single SLVS line. Also logic is implemented to do parallax corrections[5] as introduced in section 5.2. There is also the possibility to configure a repetitive eight-bit sequence to be output on the stub data lines. The pattern is configurable for each of the eight stub lines separately. The latter feature is very useful for debugging and alignment purposes at the back-end.

5.5.2.5 SSA full-event data path and format

The full-event data path is used to output the full PS-s sensor image to the MPA chip. The full sensor image, for both threshold settings, is stored in the SSA chip for 512 BXs defining the maximum for the L1 latency. Figure 5.13 shows the output format of the SSA full-event frame, which the SSA outputs on reception of an L1A trigger signal. The image for the lower threshold is output without any sparsification. The sensor image for the higher threshold passes a clustering logic and is output sparsified as 24 HIP (highly ionizing particle) flags where an active HIP flag indicates whether the higher threshold

[5]To do this the SSA is split up in six regions: 4 regions covering the 120 strips and 2 regions for the inter-strip communication.

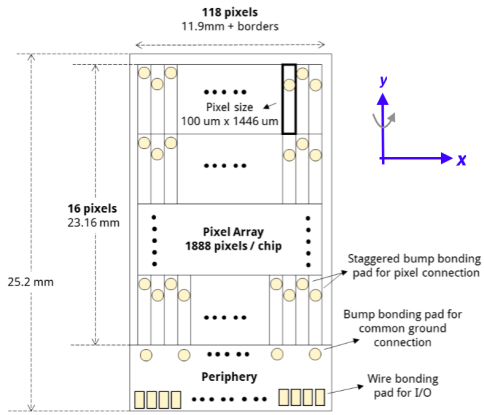


FIGURE 5.14: The dimensions of the MPA chip and the pixel array [66]. The coordinate system shown here is not the one used in CMS and is included here for later reference in section 8.5.

was passed for at least one strip within the cluster. It takes 18 bunch crossings to output the full-event data package. 16 consecutive triggers can however be sent thanks to the presence of a FIFO which can store up to 16 full-event packages. The minimal time between the end of the package end the start of the next package is 2 clock cycles.

5.5.3 The MPA chip

5.5.3.1 Introduction

The MPA chip [66] is designed to read out 1888 (118×16) macro-pixels from a PS-p sensor. The layout of a single MPA chip is shown in Figure 5.14 which shows the dimensions of the chip and macro-pixel matrix. The signal from the pixels is processed by the MPA's custom designed front-ends and digitised in the MPA's sampling logic.

Besides the input from the pixel matrix, the MPA receives input data from the SSA chip. On reception of a trigger, the MPA outputs the sparsified full-event data from the pixel matrix and the sparsified full-event data from the SSA. Besides this asynchronous full-event data stream, the MPA chip also provides stub information at BX rate. Stubs are formed in the dedicated MPA stub logic by correlating clusters in the pixel matrix with the centroid data received from the SSA.

In section 5.5.3.2, the front-end of the MPA chip will briefly be introduced. In section 5.5.3.3 the transition to the digital logic is discussed. The digital logic is then described in sections 5.5.3.4 and 5.5.3.5. Figure 5.15 shows the top-level architecture of the MPA chip.

5.5.3.2 MPA front-end

Each macro-pixel has its own front-end and is DC-coupled to it. Compared to the SSA, the MPA front-end has an extra stage at the start which provides leakage current compensation up to 200 nA for the macro-pixels. The remainder of the front-end acts as an amplifier, integrator and threshold stage. The common threshold is set using an 8-bit *threshold* DAC, whilst a second *trimming* DAC, present for each front-end, is used to equalize the

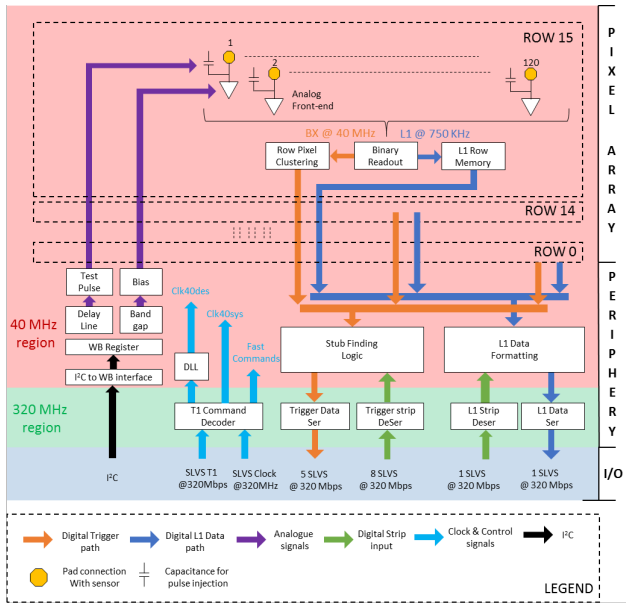


FIGURE 5.15: Top-level architecture of the MPA [66].

offset spread of the discriminators. A calibration circuit, using 20 fF calibration capacitors, is also implemented in the front-end. Each front-end has its own capacitor and the connection to the front-end can be enabled on a pixel-by-pixel basis. An injected charge of 0 to 9 fC in steps of 0.035 fC can be applied.

5.5.3.3 MPA sampling logic

The MPA sampling logic delivers the input signal to both the MPA stub and full-event logic. As described in section 5.5.1 the MPA ASIC, as is the case for the other front-end ASICs also, generates its 40 MHz clock from the fast command pattern. The sampling clock is derived from this clock and can be shifted over the full BX period with a resolution of 200 ps using the combination of a phase-shifter, which allows shifting the clock in steps of 3.125 ns, and a DLL. The three sampling modes present for the SSA, and described in section 5.5.2.3, are also implemented in the MPA. In the MPA the asynchronous mode is implemented using a 15-bit ripple counter for each pixel. As is the case for the SSA, this counter can be read out over I^2C or over the stub data lines. In the latter case, the MPA sends one counter every 8 BXs in the MPA stub package and the bits in the frame reserved for transmission of the first stub are replaced with the counter data:

$$\text{stub1}[14:0] = \text{Ripple counter}[7:0] + \text{Ripple counter}[14:8], \tag{5.1}$$

where the + sign represents a concatenation.

The same testing feature as in the SSA is implemented in the MPA: configurable 8-bit patterns, this time on a pixel-by-pixel basis, can be injected in the MPA digital logic. Furthermore the MPA can be programmed to output a repetitive 8-bit pattern on the full-event and stub data line. This is very useful for alignment of the CIC or the back-end.

Masking of pixels is also done at this level: a switch allows disconnecting the pixel front-end from the sampling logic.

5.5.3.4 MPA stub data path and format

The hit data from the sampling logic is clusterised within each row, this is the *row pixel clustering* as indicated in Figure 5.15. The position of up to 8 pixel clusters, which pass the cluster width cut, are encoded and the MPA stub finding logic correlates these clusters with the strip sensor centroids at each bunch crossing. The clock edge together with the phase at which the MPA samples the incoming SSA centroid data can be selected manually. This allows for sampling outside of the metastable region and for word alignment of the data.

The correlation between pixel and strip data is performed for pixel and strip clusters within a configurable window representing the maximum bend in the $r\phi$-plane as represented in Figure 5.16a. The maximum separation, defining the p_{T} cut, between top and bottom hit is 15.5 strip pitches. The p_{T} cut will in general be put at lower bend values as shown in Figure 5.7. For this reason and to minimise the bandwidth, a configurable bend look-up table was implemented in the MPA to map the 5-bit bend information onto a 3-bit value.

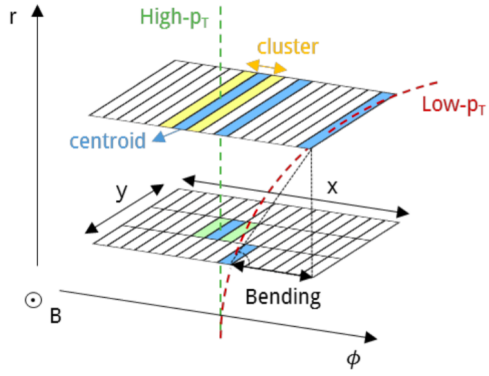
The seed cluster for the stub, either being the pixel or the strip cluster, is configurable. This is the so-called *layer swapping*. Also for testing purposes the MPA can be configured to either output the strip or pixel clusters on the stub data lines.

The output stub data format is shown in Figure 5.16b and spreads over 2 BXs. A single stub data package can contain data from two consecutive BXs and the package contains a synchronisation bit, which is always 1 for the first part of the package and 0 for the last part of the package, the number of stubs which belong to the first BX, the address of the stubs (both ϕ and z or r coordinate) and the bends.

5.5.3.5 MPA full-event data path and format

The digitised hit information from the pixel matrix is stored without any data reduction in the MPA memories. There are 16 memories on the MPA, each storing the full-event data of one row. The full sensor image is stored in the MPA memories for 512 BXs which defines the maximum L1 latency. When the MPA receives an L1A trigger the data in the MPA memory for the configured latency is sent to the sparsification logic together with the full-event data which is received from the SSA chip[6]. The sparsification results in a data format as shown in Figure 5.17. The frame starts with a 19-bit header, which is unique and cannot be found elsewhere in the full-event data, followed by 2 error bits, an L1 ID and the number of strip and pixel clusters present in the package. The actual payload encodes the widths and the locations of the first strip or pixel in the clusters. For the strip clusters also the HIP flag is passed on and for the pixel clusters the z or r coordinate. A cluster has a maximum width of 8 strips or pixels. If the cluster is larger, two clusters will be formed. The maximum number of clusters which can be sent in the frame is 31 pixel clusters and 24 strip clusters. The entire frame ends with a trailer bit. This sparsified frame is passed to an 8 BX deep FIFO from which it is output from the MPA on a single differential line operating at 320 MHz.

[6] As for the stub data lines the clock edge together with the phase at which the SSA full-event data is sampled is configurable.



(a)

$B5^0$	$Z5^3$	$Z5^2$	$Z5^1$	$Z5^0$	BX N+1
$S5^2$	$S5^1$	$S5^0$	$B5^2$	$B5^1$	
$S5^7$	$S5^6$	$S5^5$	$S5^4$	$S5^3$	
$B4^0$	$Z4^3$	$Z4^2$	$Z4^1$	$Z4^0$	
$S4^2$	$S4^1$	$S4^0$	$B4^2$	$B4^1$	
$S4^7$	$S4^6$	$S4^5$	$S4^4$	$S4^3$	
$B3^0$	$Z3^3$	$Z3^2$	$Z3^1$	$Z3^0$	
S_2	$S3^1$	$S3^0$	$B3^2$	$B3^1$	
$S3^6$	$S3^5$	$S3^4$	$S3^3$	$S3^2$	BX N
$Z2^3$	$Z2^2$	$Z2^1$	$Z2^0$	$S3^7$	
$S2^1$	$S2^0$	$B2^2$	$B2^1$	$B2^0$	
$S2^6$	$S2^5$	$S2^4$	$S2^3$	$S2^2$	
$Z1^3$	$Z1^2$	$Z1^1$	$Z1^0$	$S2^7$	
$S1^1$	$S1^0$	$B1^2$	$B1^1$	$B1^0$	
$S1^6$	$S1^5$	$S1^4$	$S1^3$	$S1^2$	
S_1	N^2	N^1	N^0	$S1^7$	
Line 0	Line 1	Line 2	Line 3	Line 4	

S_x: sync bit (S_1 = 1, S_2 = 0)
N^x: bit y of number of stubs in BX N
SX^y: bit y of stub address of stub X
BX^y: bit y of bend code of stub X
ZX^y: bit y of z coordinate of stub X

(b)

FIGURE 5.16: LEFT: Example of hit correlation between the pixel and strip sensor resulting in a stub if the bend is smaller than a configurable window [66]. RIGHT: stub data format as output by the MPA chip. A single stub data package spans 2 BXs and can contain data from two consecutive BXs.

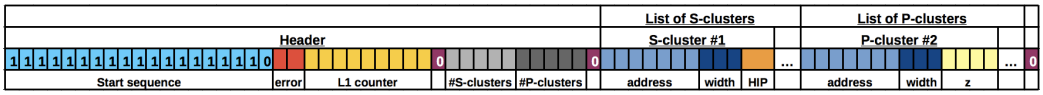


FIGURE 5.17: The full-event data as output by the MPA chip.

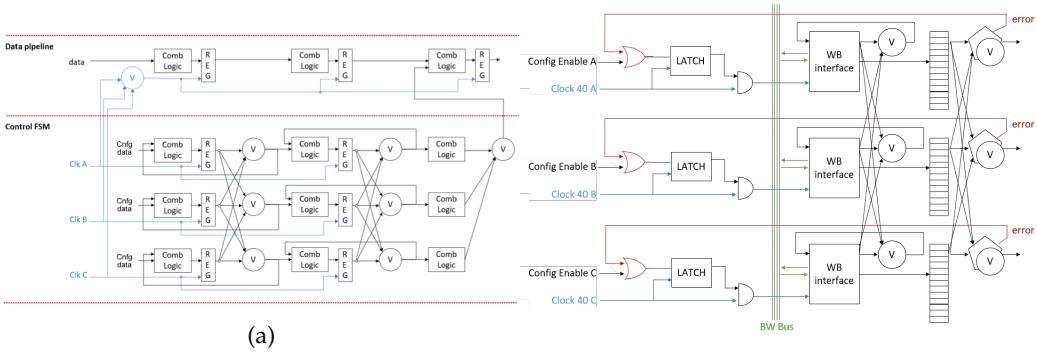


FIGURE 5.18: The TMR techniques implemented for the MPA and SSA chip's control (LEFT) and configuration (RIGHT) logic. On the left figure the input at the top referred to as *data* can for example be (processed) hit or stub data. The input, *cnfg data*, to the triplicated part is the content from configuration registers or the output from other state machines [21].

5.5.4 MPA and SSA single-event effect tolerant design strategy

As discussed in section 3.7.3 SEE hardening techniques have to be implemented in ASICs designed to operate in radiation environments to mitigate the effect of SEE on the operation of the chip. In this section, the SEE hardening techniques used in the MPA and SSA chip will be described in more detail.

The SSA and MPA design does not make use of dynamic cells where the data is stored on high-impedance nodes. All the combinatorial logic in the ASICs makes use of static logic. For what concerns the memories on the chips no error correction algorithm is implemented. The upset rate due to SEU in the memories can be evaluated during testing and the results can be used in MC simulations to understand whether this error rate has a significant impact on the track reconstruction.

The most important logic which needs to be SEU immune is the control and configuration logic in the chips as SEUs happening in this logic could introduce dead-times or damage to the ASICs. In order to increase the immunity against SEUs, Triple Module Redundancy (TMR) techniques are implemented. The TMR logic implemented for the control of the data path is illustrated in Figure 5.18a, which shows that the data path is not triplicated, but that the control path and the clock tree are triplicated. The lower part of the figure shows the control state machines, which are triplicated, and after each stage a voting mechanism (V), itself also triplicated, delivering the input to the next stage. A feedback after every voter refreshes the state variables periodically.

The TMR technique described above for the control state machines requires an active clock for refreshing the state variables and therefore has fairly high power consumption. This approach would not be suitable as TMR for the configuration registers in the ASICs (e.g. 1128 8-bit registers for the SSA) as it would not fit into the tight power budget (250 mW per MPA+SSA pair). Therefore a clock gated TMR method is implemented which is illustrated in Figure 5.18b. Here the error detection logic checks amongst the triplicated registers if there is an upset bit and if an upset is detected then the clock is activated, the voting system becomes operational and the upset is cancelled. This lowers

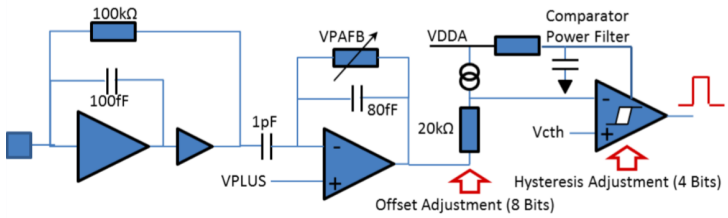


FIGURE 5.19: CBC simplified front-end schematic [68].

the error signal which activated the clock. In this way the clock is deactivated until a next upset happens [21].

5.5.5 The CBC chip

5.5.5.1 Introduction

The CBC chip [68] is designed to read out 127 strips from the top 2S sensor and 127 channels from the bottom 2S sensor. The signal from the strips is processed in the CBC's custom designed front-ends and digitised in the sampling stage. The CBC correlates the hits in the top and the bottom sensor and forms stubs. It outputs this stub data every BX. On reception of an L1A trigger, the CBC outputs the full unsparsified event information. In section 5.5.5.2 the CBC front-end will briefly be introduced. In section 5.5.5.3 the sampling logic is discussed and the digital logic is the topic of section 5.5.5.4 and 5.5.5.5. Figure 5.19 and 5.20 show the CBC's front-end and top-level architecture respectively.

The CBC version which will be described in the following sections is CBC3. The CBC2 chip, being the predecessor of the CBC3, has many similarities with the CBC3 version. In section 5.5.5.6 a few differences between CBC2 and CBC3 will be pointed out.

5.5.5.2 CBC front-end

Each CBC input channel has its own front-end stage and the front-end is specifically designed to read out AC-coupled n-in-p type silicon sensors. As shown in Figure 5.19, the CBC's front-end consists of a pre-amplifier, post-amplifier and comparator stage. The input charge from the strip is read out by the pre-amplifier and integrated on the 100 fF capacitor. This capacitor is discharged through the resistive feedback network, which uses a 100 kΩ resistor, into the post-amplifier. The pre-amplifier and post-amplifier are designed to result in a combined gain of 50 mV/fC. The shaping of the input signal is designed to give a rise time of less than 20 ns and a return to baseline within 50 ns. The post-amp's offset can be adjusted on a front-end to front-end basis using an 8-bit DAC. This setting allows for correction of channel-to-channel mismatches in the amplifiers and the comparator stage due to process variations. The last part in the front-end is the comparator stage, with configurable hysteresis, which outputs a fixed amplitude signal for the time that the input signal crosses a configurable threshold. This threshold setting is the same for each front-end and represents the actual threshold at which the chip is operated. The configurable hysteresis logic is added to the comparator stage to make the comparator stage less sensitive to small input variations: when the input is lower than a chosen threshold, the output is low, whereas when the input is above a different higher

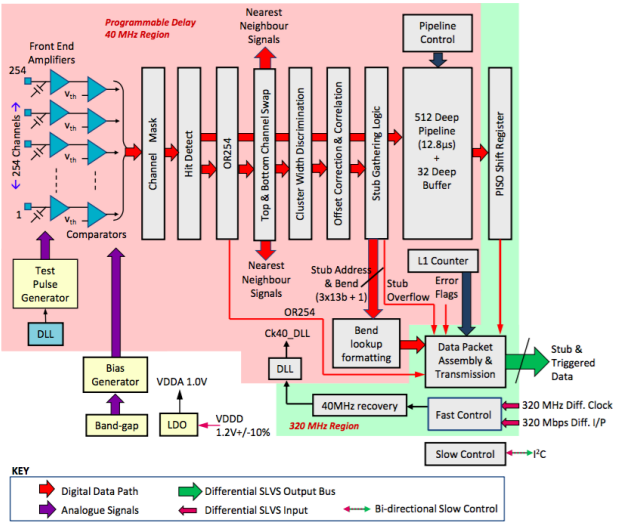


FIGURE 5.20: CBC top-level schematic. The *Hit Detect* logic is equivalent to the sampling logic [68].

chosen threshold the output is high. When the input is between the two levels the output retains its value.

The analogue front-ends furthermore comprise a test pulse capacitor. This 20 fF $\pm$ 25% (3σ) capacitor can be discharged into the chip's pre-amplifier and mimics the charge injected into the front-end from a charged particle depositing energy in the sensor. The test pulse is enabled for all channels in the same test group. There are 7 such test groups containing 32 channels and 1 test group containing 30 channels. The injection is triggered by the CBC receiving a test pulse trigger and the fine timing of the discharge can be controlled by configuring a dedicated DLL circuit which allows a time resolution of 1 ns. This DLL circuit uses as its timing reference the output of the DLL for the sampling logic (see section 5.5.5.3). All the capacitors in the same test group will receive a voltage step from a dedicated voltage step generator which uses an 8 bit DAC to set the amplitude of the applied voltage step. The DAC uses a 1.1 V supply as its reference so the test pulse can be selected in steps of 0.086 fC up to a value of 22 fC. Which channels belong to which test group is shown in Table 5.2. This distribution is chosen carefully: consecutive top and bottom channels are in the same group so that stubs can be generated and the stub logic tested. After sending a test pulse trigger the voltage step generator has to be reinitialised which results in a maximum rate of 100 kHz at which the test pulse trigger can be sent.

5.5.5.3 CBC sampling logic

The CBC uses a DLL to provide a clock, phase-shifted with respect to the internal 40 MHz clock, to sample the output of the comparator. The internal 40 MHz clock is derived from the 320 MHz clock and the fast command pattern. The resolution of the sampling phase is 1 ns. As for the MPA and SSA chip different modes of sampling are implemented in the CBC chip. These modes are illustrated in Figure 5.21:

Channel group	Stimulated channels
0	242, 241, 226, 225, 210, 209, 194, 193, 178, 177, 162, 161, 146, 145, 130, 129, 114, 113, 98, 97, 82, 81, 66, 65, 50, 49, 34, 33, 18, 17, 2, 1
1	244, 243, 228, 227, 212, 211, 196, 195, 180, 179, 164, 163, 148, 147, 132, 131, 116, 115, 100, 99, 84, 83, 68, 67, 52, 51, 36, 35, 20, 19, 4, 3
2	246, 245, 230, 229, 214, 213, 198, 197, 182, 181, 166, 165, 150, 149, 134, 133, 118, 117, 102, 101, 86, 85, 70, 69, 54, 53, 38, 37, 22, 21, 6, 5
3	248, 247, 232, 231, 216, 215, 200, 199, 184, 183, 168, 167, 152, 151, 136, 135, 120, 119, 104, 103, 88, 87, 72, 71, 56, 55, 40, 39, 24, 23, 8, 7
4	250, 249, 234, 233, 218, 217, 202, 201, 186, 185, 170, 169, 154, 153, 138, 137, 122, 121, 106, 105, 90, 89, 74, 73, 58, 57, 42, 41, 26, 25, 10, 9
5	252, 251, 236, 235, 220, 219, 204, 203, 188, 187, 172, 171, 156, 155, 140, 139, 124, 123, 108, 107, 92, 91, 76, 75, 60, 59, 44, 43, 28, 27, 12, 11
6	254, 253, 238, 237, 222, 221, 206, 205, 190, 189, 174, 173, 158, 157, 142, 141, 126, 125, 110, 109, 94, 93, 78, 77, 62, 61, 46, 45, 30, 29, 14, 13
7	Top Dummy, 240, 239, 224, 223, 208, 207, 192, 191, 176, 175, 160, 159, 144, 143, 128, 127, 112, 111, 96, 95, 80, 79, 64, 63, 48, 47, 32, 31, 16, 15, Bottom Dummy

TABLE 5.2: Mapping of the CBC test pulse groups to the channel numbers.

- Fixed pulse width mode: in this case the output of the comparator is latched. As in this mode the input to the latch is the direct output of the comparator this mode is able to detect comparator pulses which are non-overlapping with the sampling clock. Hits in consecutive clock cycles in the same channel will only be captured if the signal returns below the comparator threshold for each hit. The output signal in this mode has a fixed pulse duration of 25 ns for each detected comparator output transition.

- Sampled mode: in this mode the sampling of the comparator output is performed synchronously. For the sampling logic to therefore output a hit, the comparator output needs to be high at the time of sampling. Comparator outputs being high for more than a single BX, either due to consecutive hits or due to single hits generating a signal above the threshold lasting more than one clock cycle, will result in the output of the sampling logic to be high also for several clock cycles. The output will only return low again on the clock cycle succeeding the comparator output's return to zero. In this mode comparator outputs which are high only in between the sampling clock edges will not result in an output from the hit detect logic.

- Logical OR mode: the logical OR of the fixed pulse width and sampled mode.

- HIP suppressed mode: in this mode either the output of the sampled mode or the Logical OR mode is used, but extra logic forces the output of the hit detect logic low in this way limiting the maximum number of consecutive clock cycles where the output of the sampling logic can be high. This extra logic can be used to eliminate the effect of highly ionizing particles generating an active output of the sampling logic for several clock cycles.

The output of the sampling logic is used as input to the digital part of the chip. There are two types of digital logic on the chip: full-event logic and stub logic. Which output from the sampling stage is used can be configured for the full-event and stub logic separately. The channel masking is implemented right before the sampling logic: there is the possibility for each channel to disconnect the output from the front-end to the sampling logic as shown in Figure 5.20.

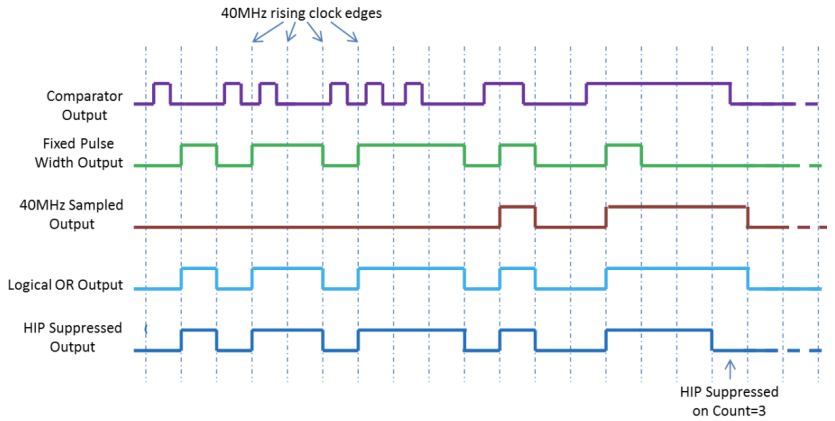
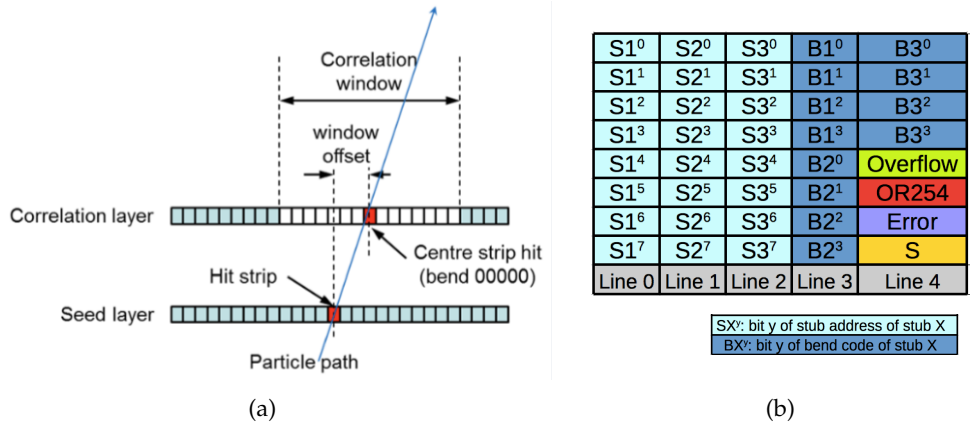


FIGURE 5.21: CBC sampling modes [68].

5.5.5.4 CBC stub data path and format

The CBC receives the sampled data from the top and the bottom channel and is therefore able to correlate clusters in both sensors to generate stubs. Inter-chip communication guarantees that this correlation can be performed across chips. The stub data path is represented in Figure 5.20. The full sampled sensor image is passed into a logical OR network which produces an active high output (*OR254*) if any of the channels are hit. This bit is passed along in the stub data output package. Subsequently the hit data from the full sensor is passed to the *layer swapping logic* which allows for swapping of the top and bottom channels so that the user can choose which layer to use as seed layer. The *cluster width discrimination* then filters out clusters which have a width larger than a programmable value[7]. Clusters surviving this filter are passed to the *offset correction and cluster correlation logic* which allows for programming an offset of the correlation layer with respect to the seed layer to correct for parallax errors. Four separate regions per CBC chip are defined for which an offset of the *correlation window* can be programmed. The offset is programmable up to ± 3 channels with half strip resolution. The CBC then tries, for every cluster in the seed layer, to find a cluster in the correlation layer, within a programmable window, as close as possible to the center of the seed cluster's location. The maximum size of this window is ± 7 channels from the center of the seed cluster and the window is configurable with a resolution of half a channel. If the CBC finds a correlation between a seed cluster and a correlation cluster within this window, a stub is formed. This window therefore defines the p_{T} cut. The definition of the correlation window and the offset are shown in Figure 5.22a. When a stub is found the location of the seed cluster is decoded as the location of the center of the cluster in the *stub gathering logic*. Half strip resolution is thus possible for clusters of size larger than 1. Also the offset between the cluster in the seed and correlation layer is calculated and this represents the bend. Given the maximum correlation window of ± 7 strips and a half strip resolution for the bend, the bend can be represented by a 5-bit number. The 5-bit bend information is however only used internally on the chip. In order to reduce the output bandwidth of the chip, the 5-bit bend is translated to a 4-bit bend code in the configurable bend LUT. The CBC outputs the location and the bend of up to 3 stubs per BX using 5 lines running

[7] With a maximum of 4 strips.



(a) (b)

FIGURE 5.22: LEFT: Definition of the CBC's correlation window and window offset [68]. The correlation window defines the p_T cut and the layer offset is implemented to perform the parallax correction explained in Figure 5.4. This figure shows a hit in the seed layer being matched to a hit in the correlation layer. The resulting bend is zero (this is before the bend LUT) due to the fact that the window offset is set to 3 channels. RIGHT: Output format of the CBC stub data. This package covers a single BX.

at 320 Mbps as illustrated in Figure 5.22b. The stub data is not ordered and in case there are more than 3 stubs in a BX the stubs with lowest address are output. In this case, the CBC reports in the stub data package that more than 3 stubs were found by enabling the stub overflow bit. The output data frame also contains a synchronisation bit which is always high. This bit can be used by the CIC or the back-end to check the alignment of the incoming data. Furthermore, the package contains the OR254 bit and an error flag which is the logical OR of the full-event FIFO full flag and the latency error flag (see section 5.5.5.5).

5.5.5.5 CBC full-event data path and format

For every bunch crossing the output from the sampling logic for each channel is written into the CBC's memory. The location of the memory to which the data is written is given by the write pointer which is a 9-bit counter. A second pointer, the trigger pointer, is used to access the data at the correct latency: the pointer is driven by the trigger counter which does not start counting until a given configurable latency after the write counter starts

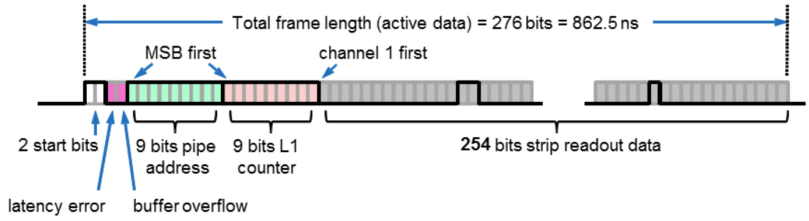


FIGURE 5.23: The CBC full-event data format [68].

counting. On reception of an L1A trigger, the CBC chip accesses the memory location to which the trigger pointer points. The maximum L1 latency, defined by the depth of the memory, is the same as for the MPA and SSA chips and is 512 BXs. The full-event data from the memory is passed on to an output FIFO with a depth of 32 from which it is serialised, foreseen of a header, and output over one differential line at 320 Mbps towards the back-end. The output full-event data frame is shown in Figure 5.23. The full-event frame starts with a header of two bits, followed by a field for two error bits, representing the latency error and the buffer overflow error. The latency error is raised when there is discrepancy between the trigger pointer and the write pointer on one end and the programmed latency on the other end. The buffer overflow error bit represents a FIFO full condition of the 32 deep output FIFO. After the error bits follows a 9-bit pipeline counter and a 9-bit L1 counter which increments every time an L1A trigger is received. This is followed by 254 bits which represent the sensor image. Including 3 BX of dead time between two consecutive full-event data packages, the time required to output a single full-event data package is 950 ns. This allows for a maximum sustainable full-event data rate of 1 MHz. The first bit of the full-event package at the output of the CBC is in sync with the output of the sync bit in the stub data package.

5.5.5.6 The CBC2 chip

From the user's point of view the CBC3 and CBC2 chip are fairly similar. There are however a few important differences, which are summarised below:

- A few tweaks have been made to the CBC's front-end when going to the CBC3 version. The CBC2 has the possibility to read out signals with both polarities as at the time of the CBC2 design it was not decided yet whether n-bulk or p-bulk sensors were going to be used.
- In the CBC2's full-event frame no L1 counter is implemented.
- The stub data information in the CBC2 is very basic: there is a single 40 Mbps line per CBC which flags the presence of a stub in that BX. No stub location or bend information is output.

5.5.6 The CIC chip

5.5.6.1 Introduction

In this section the CIC1 chip will be introduced. The CIC1 is a prototype and the final version will be the CIC2, discussed briefly in section 5.5.6.5. Most of the functionalities of the final CIC are however already implemented in the CIC1 design, so from a user point of view both versions are fairly similar. Due to the familiarity between both versions, the CIC1 and CIC2 will often be referred to as *CIC* in this work when the discussion is applicable for both.

The CIC ASIC [69] receives full-event and stub data from eight MPA chips in the PS module and eight CBC chips in the 2S module. The task of the CIC chip is to group the data from a module and to reduce the amount of data output by a module. The CIC does this by selecting the stub data with the lowest bend and by sparsifying the full-event data in case of the CBC. The CIC is a fully digital chip and is not a readout chip as the ASICs which were described until now. The CIC can therefore be seen more as a hub where the data gets concentrated: 48 differential input lines running at 320 Mbps

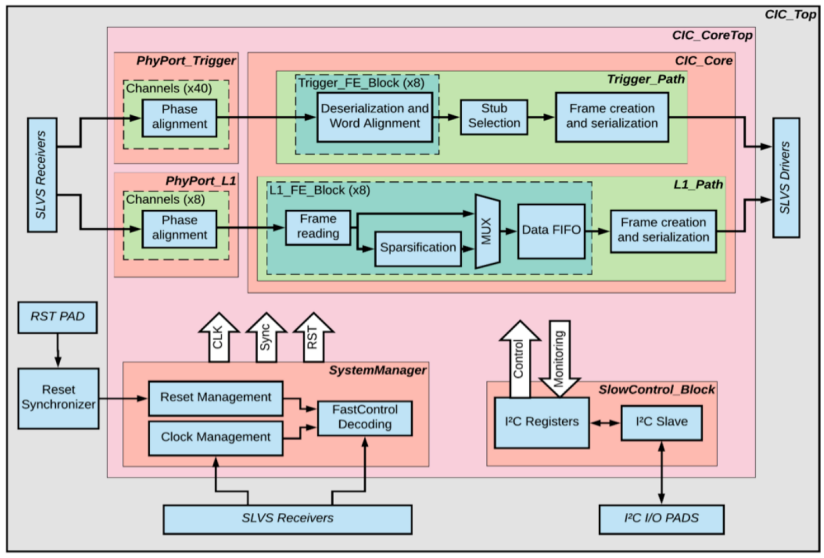


FIGURE 5.24: Top-level representation of the CIC1 [69].

are concentrated to 6 or 7 differential lines[8] at 320 Mbps. This results in a reduction in bandwidth with a factor of ≈ 8. For the CIC2 there will also be the possibility to output the data at 640 Mbps. This mode will be used for PS modules in the highest occupancy regions where the higher bandwidth is required to limit the stub losses.

From the user point of view the CIC is therefore quite transparent as to first order the chip only changes the data format. The only caveat to this is that the CIC also drops data:

- Stub data is dropped when the number of stubs received from the front-end chips is larger than the number of stubs which can be output by the CIC. The stubs which are dropped are the ones with the largest bend.
- Full-event data is dropped when the L1A trigger rate is too high and the FIFOs on the CIC get filled up.
- Clusters can be dropped in the sparsified full-event readout if the cluster occupancy is too high.

In the next sections the full-event and stub data path in the CIC chip will be discussed, but first the *phy ports*, common to both the full-event and stub path, will be described.

5.5.6.2 Phy ports

As shown in the CIC top-level schematic in Figure 5.24, the first logic the input data encounters are the CIC's phy ports. These blocks phase align the incoming data with respect to the CIC's internal 320 MHz clock. This is required, because all the ASICs on the module run on the same clock, directly provided by the lpGBT (section 5.5.7), so the CIC knows the exact sampling frequency, but it does not know the sampling phase. A phase alignment mechanism is thus required to sample the bit at the optimal phase, i.e. away from the bit transitions.

[8]Depending on the CIC mode.

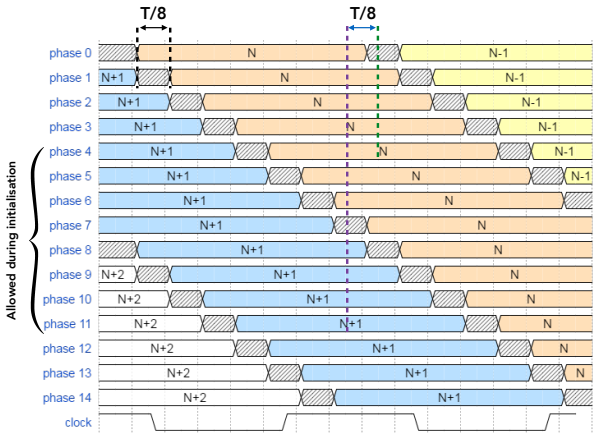


FIGURE 5.25: Illustration of the output of the delay line in the phase aligner block. During initialisation only tap values between 4 and 11 are allowed. The purple and green line indicate the optimal sampling phase for tap values of 4 and 11, which are equivalent modulo 1 tap value [71].

The CIC implements the phase aligner blocks developed for the lpGBT for which the input lines are grouped into *phy ports*. Each phy port receives 4 input data lines which need to operate at the same frequency. The phase can be different for each line within the phy port. In the CIC case the input data lines all operate at the same frequency, but the optimal sampling phase might be different for each line due to different timing between front-end ASICs or due to routing delays on the hybrid. The CIC phy ports allow to select the sampling phase (*taps*) with a 391 ps (= 3.125 ns/8) resolution when running at 320 MHz. Figure 5.25 illustrates the phase tuning mechanism: the phase alignment block delays each incoming data line with 15 different delay values. Each delay line is shifted by $1/8^{\text{th}}$ of the clock period with respect to the other. During the initialisation procedure, during which the optimal taps are automatically extracted, only tap values between 4 and 11, which actually cover a full clock cycle, are allowed.

To find the optimal alignment setting the CIC needs to be fed with a repetitive data pattern. After starting the CIC phase tuning procedure the chip reports for each of the 48 input lines if the phase tuning succeeded. When the CIC sits on a hybrid this alignment should in principle be done only once as the delays between the different lines are fixed. Once the delays are known the optimal tap values can be programmed over I^2C. More details on the exact alignment procedure can be found in section 9.2.3.

The CIC can be configured to bypass the phy ports and the full CIC core logic: the inputs to one phy port can directly be forwarded to the CIC output lines. Which phy port is under investigation can be configured. In this way, access to the raw input data from the front-end chips is provided. This is a very useful debugging feature which allows to trace back issues to the front-end chips.

5.5.6.3 CIC stub data path and formats

The CIC input stub data formats are described in section 5.5.3.4 and 5.5.5.4. Depending on the connected front-end chip, stub data packages arrive at a rate of 40 MHz (CBC)

Front-end chip	CBC (2S)				MPA (PS)			
Bend info	Bend info present		Bend info not present		Bend info present		Bend info not present	
#Output lines	5	6	5	6	5	6	5	6
Width of a single stub (bits)	18		14		21		18	
Max #stubs/boxcar	16	19	20	25	13	16	16	19
Total stub payload/boxcar (bits)	288	342	280	350	273	336	288	342

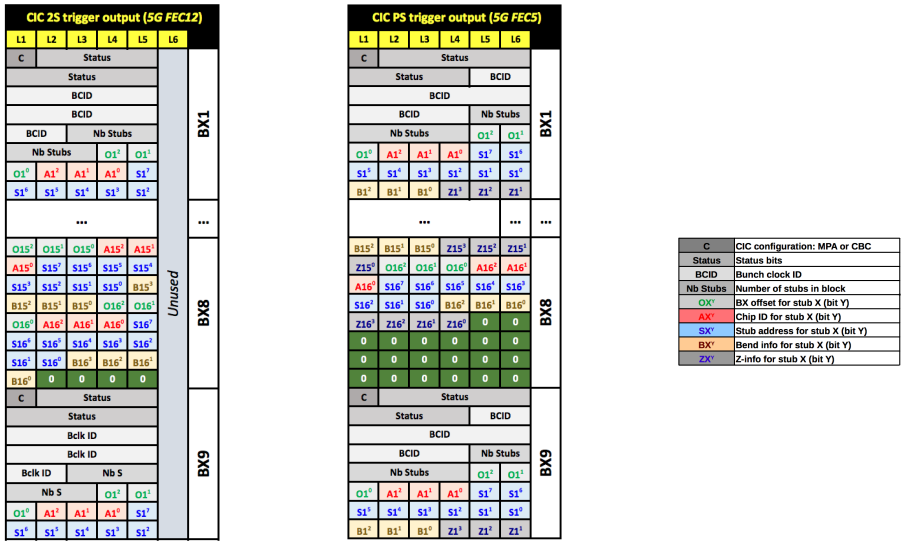
TABLE 5.3: 320 MHz stub data configurations for the CIC1.

or 20 MHz (MPA). In order to choose the correct, out of eight, bit as the first bit of a *word* the CIC needs to be fed with word alignment patterns on each of the five stub data lines. The CIC reports through an I^2C register if the word alignment succeeded. The word alignment pattern can be different for each of the five lines. This alignment can correct a time skew up to 9.375 ns between the lines of different front-end chips. The word alignment patterns are also used to perform the BX0 alignment (see later). More details on the word alignment procedure can be found in section 9.2.3. As is the case for the phase alignment, the word alignment should in principle also only be done once when the CIC sits on the hybrid. Once the word alignment is achieved, the alignment block still uses the sync bits present in both the MPA and CBC stub data stream to check the validity of the word alignment.

There are 8 (one per front-end chip) stub reconstruction blocks present in the CIC which output 5 (MPA) or 6 (CBC) stubs per block at a 20 MHz rate to the stub selection logic where stubs from 8 consecutive BXs are gathered and sorted on bend value. The selected stubs are then passed to the *frame creation and serialization* where the header is attached and the stub data is put on the output lines. The exact output format depends on the CIC configuration. All the possible CIC1 configurations are shown in Table 5.3. Besides configuring the CIC for the connected front-end chip type, also the number of lines on which the CIC outputs the stub data can be chosen. This can be 5 (so-called *FEC12* mode[9]) or 6 (so-called *FEC5* mode) lines. Furthermore, there is a configuration to select whether the bend info should be transmitted in the payload. Using more lines and/or not using the bend information in the output data frame increases the number of stubs which can be transmitted by the CIC. The bend information is not required in the final operation: the L1 trigger system will most likely not use the stub bend information to reconstruct the tracks. The presence of the bend information can however be used to validate the stub sorting logic on the CIC.

Two examples of the stub data format produced by the CIC are given in Figure 5.26. A frame of data spans 8 consecutive bunch crossings. This frame is referred to as a *boxcar* in CMS DAQ terminology. In order for the CIC to select the correct first BX of data from the front-end chips to start aggregating the stub packages, a *BX0 alignment* needs to be performed. This alignment procedure will also be explained in more detail in section 9.2.3. The status of the BX0 alignment can be read through an I^2C register. In the final operation, the BX0 alignment is important to make sure that each CIC throughout the OT takes the same BX as BX0. Each stub boxcar starts with a header, where the *CIC conf* bit specifies whether the CIC is configured for PS or 2S mode (0 for CBC, 1 for MPA). There are 9 status bits in the header, one for the CIC and one for each of the front-end chips. The CIC error bit is set on a CIC stub overflow or if a wrong sync bit was received from the front-end chips. The front-end chip statuses in the CBC case are a logical OR of the error bits in the eight CBC stub packages and is also high if a wrong sync bit is received. In

[9]When using 6 lines for stub data from CIC to lpGBT, the lpGBT cannot be operated in the FEC12 mode any more, but has to be operated in FEC5 mode, hence the naming (see section 5.5.7).



(a) (b)

FIGURE 5.26: Examples of the stub data format as output by the CIC chip [69]. Two modes are shown: the CIC-CBC, with bend info and 5 lines configuration (LEFT) and the CIC-MPA, with bend info and 6 lines configuration (MIDDLE).

the MPA case they only signify whether the sync bit is wrong. In the header, the CIC also attaches a bunch crossing ID and the number of stubs present in the payload is reported. The payload represents a stub as if it would be coming from the front-end chip: each stub has an address, a bend and for the MPA also the z or r coordinate is reported. The CIC attaches to each payload the chip ID from which the stub was received and the timing offset of the stub within the CIC stub boxcar.

5.5.6.4 CIC full-event data path and formats

The *frame reading block* in the CIC looks for the header of the front-end chip's full-event data frame (as described in section 5.5.3.5 and 5.5.5.5) for a configurable time after the reception of an L1A trigger. A time-out error is raised if the full-event header was not detected within the programmable time window. The received full-event data for the MPA, which is sparsified, is directly stored in the data FIFO. There is one FIFO for each chip. For the CBC there are two options: the full unsparsified frame is stored in the FIFO or the CIC performs clusterisation on the CBC full-event frame and the sparsified frame is stored in the FIFO.

The FIFO is the same for both 2S and PS. Events are only written into the FIFO if the internal L1 counter on the CIC matches the L1 counter in the received full-event frame and if the trailer bit is correct. If this is not the case an empty frame with the error bits high is produced and written into the FIFO. The width of the FIFO defines the maximal occupancy of clusters which can be saved per chip. The depth of the FIFO directly relates to the maximum short term trigger rate the CIC can cope with. For the CIC1 a FIFO with a depth of 10 and a width of 550 bits was implemented for each of the eight front-end chips. The FIFO width allows for the full CBC full-event frame to be stored in case of the

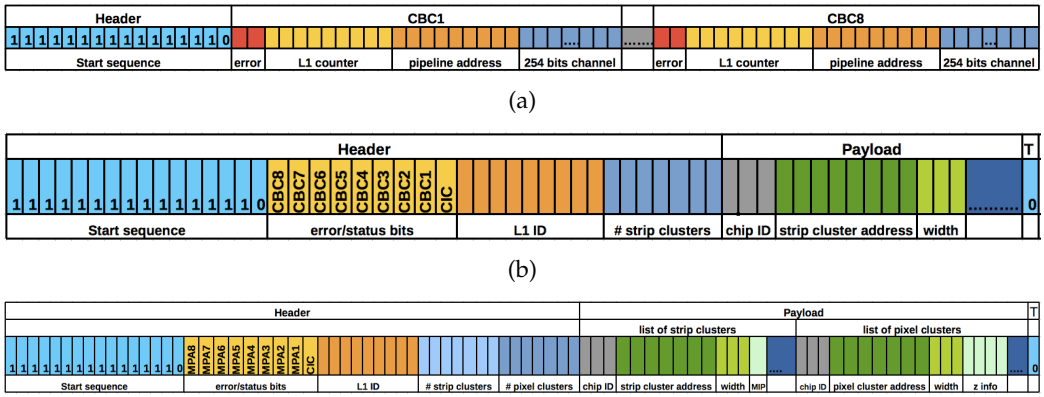


FIGURE 5.27: The full-event data frame format output by the CIC chip: TOP: CIC-2S in unsparsified mode; MIDDLE: CIC-2S in sparsified mode; BOTTOM: CIC-PS mode.

CBC unsparsified mode and allows for events with 17 pixel clusters and 17 strip clusters to be stored in case of the PS mode.

The *frame creation block* reads from the FIFOs and forms the full-event data package which represents the sensor image of half a module. The output formats are shown in Figure 5.27. The full-event data is serialised and put on a single line out of the CIC. The maximum number of clusters in the CIC package is 127 [33]. This occupancy is however not sustainable by the CIC bandwidth. The bandwidth allows a continuous occupancy of $\approx$ 25 clusters per CIC and per L1A trigger when the CIC is operated in the 320 MHz mode.

5.5.6.5 The CIC2 chip

An improved version of the CIC1 chip has recently been developed and includes several changes. The most important changes are the triplication of the logic and the addition of the configuration where the CIC is operated in the 640 MHz mode. This mode will be used for the PS modules in the regions with highest stub occupancy. For these modules higher bandwidths are required as can be seen from Figure 5.28 where the stub rates per module (so 2 CICs) are shown per BX at an average pileup of 200. In the 640 MHz mode with stub data on 6 lines, each CIC boxcar, which integrates data of 8 consecutive BXs, can contain up to 35 (bend info included) or 40 (no bend info included) stubs in the data stream.

5.5.7 The optical link

The lpGBT chip [71] in cooperation with the VTRx+ is designed to provide a high speed optical link in radiation environments. The lpGBT is a multi-purpose radiation hard chip to be used by several LHC experiments. A typical use case of the lpGBT is shown in Figure 5.29 where a single bidirectional link is used between the FPGA in the counting room and the on-detector electronics. The optical link embeds all configuration, control and data information into a single downlink (towards the detector) and single uplink

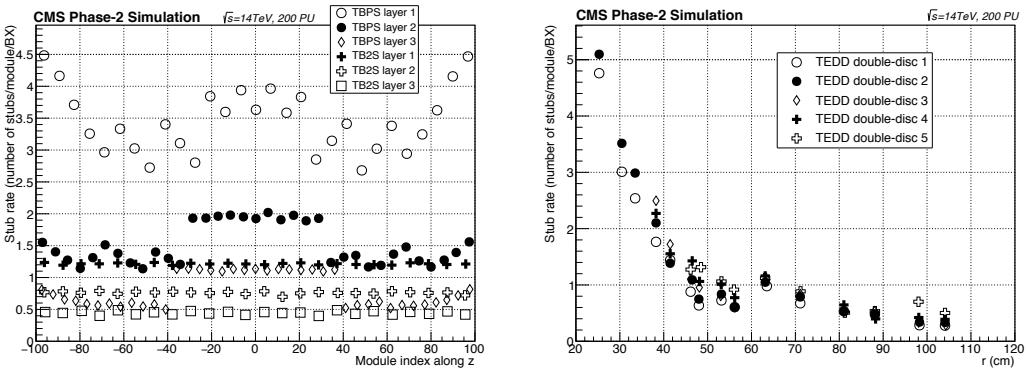


FIGURE 5.28: Expected stub rates per BX per module for an average pileup of 200 for the different parts of the Outer Tracker [33].

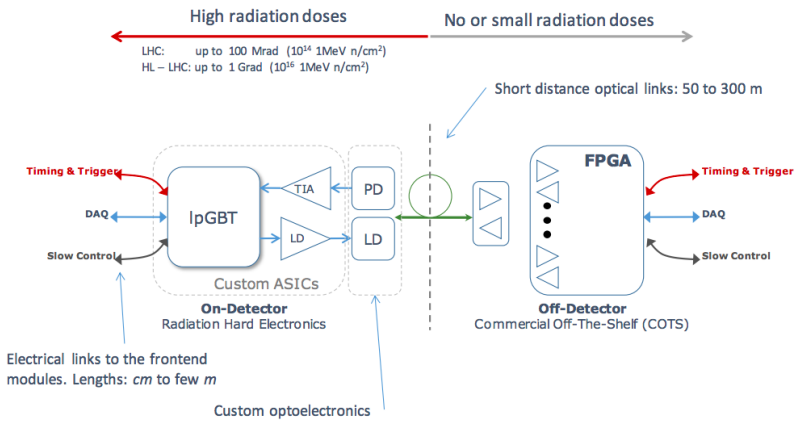


FIGURE 5.29: Optical link using the lpGBT [73]: custom ASICs and optoelectronics are developed for the high radiation environment. Commercial off-the-shelf components can be used to instrument the back-end.

(from the detector) package with forward error correction (FEC). There are 2 FEC modes available in the lpGBT: FEC12 and FEC5. FEC12 and FEC5 can respectively recover 12 and 5 consecutive wrong bits. The optical link is received by the VTRx+ [72] containing receivers and transmitters to do the conversion from the optical to the electrical domain and vice-versa.

The lpGBT performs the task of phase aligning, serialising and encoding the incoming data into a custom designed data package. The data is received over so-called *eLinks*. Furthermore it decodes and deserialises the downlink package and extracts from it a configurable clock for internal logic and further distribution. The lpGBT is the clock source for the other ASICs on the module and the incoming data to the lpGBT is synchronous and has a fixed and stable phase relationship to the lpGBT internal clock which in turn is locked to the LHC clock. The data rate of the eLinks is configurable (160, 320, 640, 1280 Mbps). The configurations used in the p_{T} modules will be the 320 Mbps and 640 Mpbs mode.

The lpGBT provides a bandwidth of 2.56 Gb/s for the downlink and 5.12 or 10.24 Gb/s for the uplink. The bandwidth available for uplink detector data depends on the operating mode and forward error encoding which is used: 3.84 Gb/s and 4.48 Gb/s for 5.12 Gb/s mode for FEC12 and FEC5 respectively and 7.68 Gb/s and 8.96 Gb/s for the 10.24 Gb/s mode for FEC12 and FEC5 respectively.

The decoding and encoding of the lpGBT frames in the back-end happens in the lpGBT-FPGA block implemented on the FPGA sitting on the DAQ card in the counting room. This firmware block is provided by the lpGBT team and the combination of lpGBT and lpGBT-FPGA essentially makes the optical link transparent to the user.

5.6 p_T module data path

Figures 5.30a and 5.30b show respectively the main lines which are important for control and data transmission on the 2S and PS modules. Note that not all connections are shown here. Power lines, ground lines,... are not included for clarity reasons. A good understanding of the control and chip interconnections is important in the development of a test bench to read out module prototypes and final modules, which is the topic of Chapter 6.

Figure 5.30a shows the interconnections on a 2S module where physically three main blocks can be identified: the two front-end hybrids carrying the CBCs and the SEH which, from a data transmission and control point of view, carries the lpGBT and VTRx+ chips. In case of the PS module the inter-chip connectivity (Figure 5.30b) is more complicated than the 2S case due to the fact that the MPAs, physically located on the MaPSA, have to receive data from the SSAs which sit on a separate front-end hybrid and the MPAs in turn have to send data back to the CIC on the FEH.

For both the 2S and PS module there is a single fast command line which goes from the lpGBT to each CBC, SSA and CIC chip. The 320 MHz clock is generated by the lpGBT and there is a separate clock line for each front-end hybrid. In the PS module the MPA chips receive the clock and the fast commands buffered and forwarded by the SSA chip. For the PS module there is also the possibility to run the CIC in the 640 MHz mode, therefore a separate clock line from the lpGBT to the CICs is present. Besides the fast commands and the clock the lpGBT also controls the reset lines. There is a dedicated reset line for each chip flavour. The configuration of the front-end chips and the CIC happens over I^2C. For the 2S module there is one I^2C bus for each FEH and for the PS module there is one I^2C bus per left/right side of the module. Other I^2C buses are routed from the lpGBT to the DCDC converter and the VTRx+.

The data flow was already explained in the sections dedicated to the front-end ASICs. To summarise: in case of the 2S module each CBC receives input from 254 sensor channels and sends 5 lines of stub data and 1 line of full-event data at 320 Mbps to the CIC. In the PS module, the SSAs receive data from 120 sensor channels and send 8 lines of stub data and 1 line of full-event data at 320 Mbps to the MPA. In addition, the MPA acquires the data from 1920 macro pixels. Each MPA sends 5 lines of stub data and 1 line of full-event data at 320 Mbps to the CIC.

In both modules, the CIC thus receives data on 48 lines at 320 Mbps and concentrates the data on 5 or 6 stub data lines and 1 full-event data line and sends it to the lpGBT at 320 Mbps. In the PS case there is the possibility to also run the CIC in the 640 MHz mode which essentially doubles the bandwidth from each side of the module. The lpGBT thus receives 2 lines of full-event data and 10 or 12 lines of stub data, either at 320 Mbps

or 640 Mbps. The lpGBT serialises this data and transfers it to the VTRx+ chip which translates it to the optical domain using a bandwidth of 5.12 or 10.24 Gbps. The downlink of 2.56 Gbps facilitates the control and configuration of the full module.

Figures 5.30a and 5.30b show the interconnections in a finalised module. However, during R&D and also still during production, sub-assemblies need to be tested. For the 2S case for example the FEHs need to be tested during production as stand-alone objects before they are assembled into a module. This requires electrical communication to the CIC directly. During even earlier prototyping stages the CIC is not yet present on the 2S FEH and thus direct electrical communication with each of the 8 CBC chips is required. For the PS case, testing of sub-assemblies is complicated by the fact that the MPA and SSA physically sit on different substrates. During production the PS front-end hybrids also need to be tested stand-alone. At this stage no MPA chip is in the system. Therefore the testing of PS-FEHs is rather complex as all the wire bond pads which facilitate connection to the MPA will need to be made available temporarily to the test-system. Like this, the testing of the PS-FEH boils down to testing of stand-alone SSA chips and testing of stand-alone CIC chips: the SSAs and CIC sit on the same physical object, but no MPA is present, which would normally form the interconnection between the SSA and the CIC. The MaPSA assemblies themselves also need stand-alone testing. Each MPA on the MaPSA will be tested and it is thus a requirement to also stimulate the MPA with SSA-like input to test the full MPA's functionality.

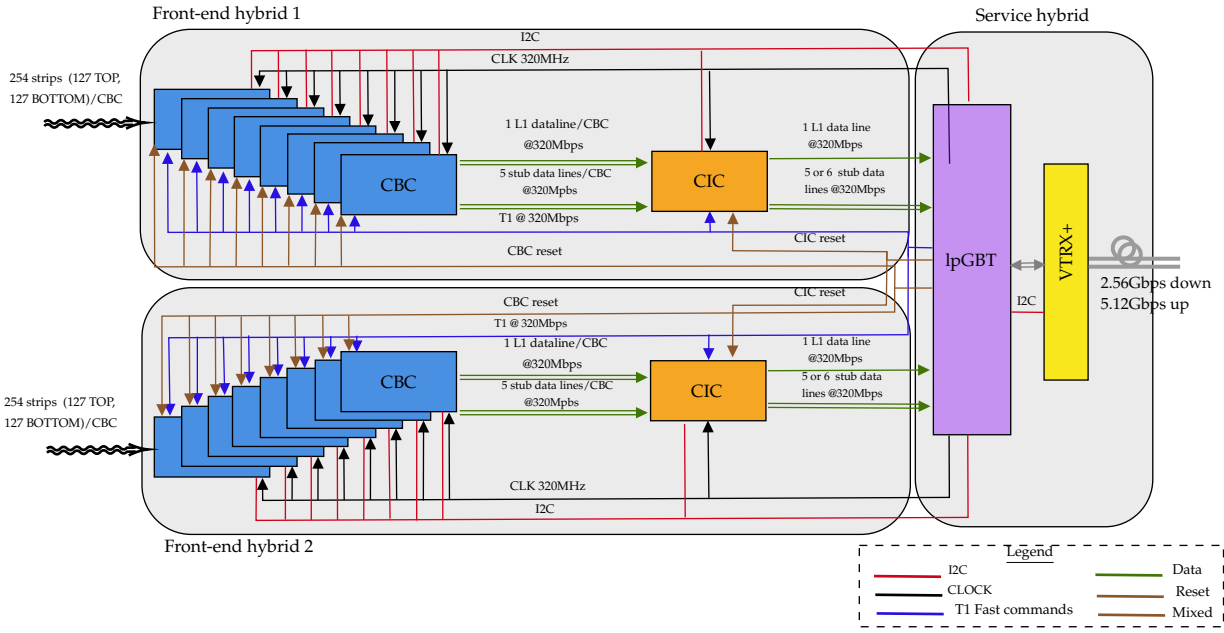
The above advocates for a flexible test bench which can be configured to read out different hardware configurations. This test bench will be the topic of Chapter 6.

5.7 Module assembly and testing during production

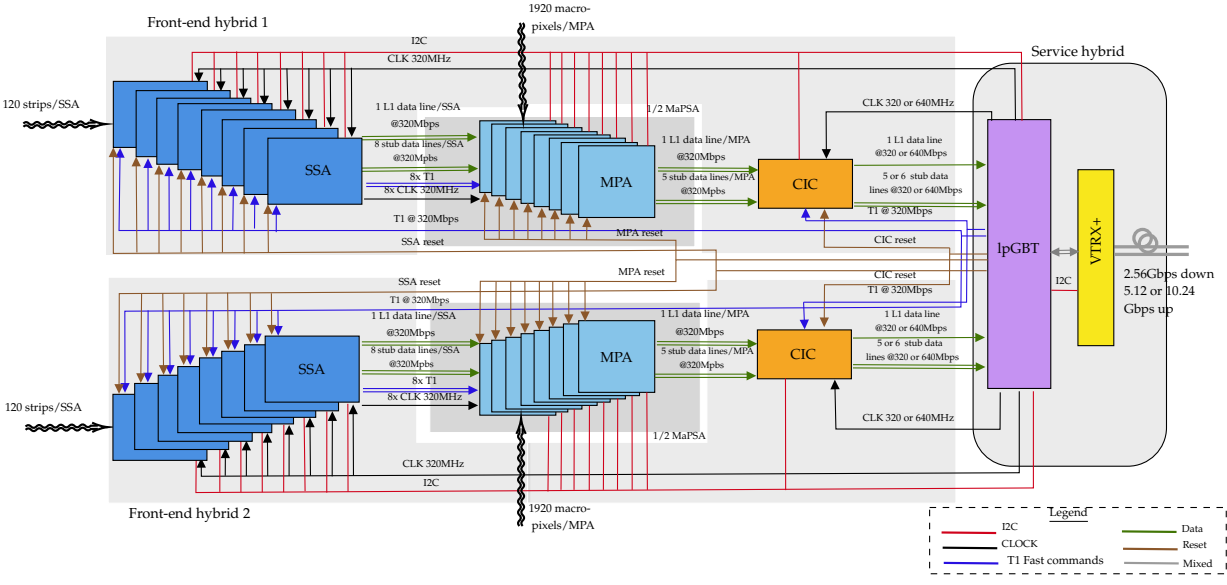
The assembly of the p_{T} modules will be done largely manually. The main ingredients which have to be put together to a full module are: two sensors, glued onto the spacers with kapton foils, two front-end hybrids and one or two service hybrids. The assembly procedure consists of gluing the different components together with high mechanical precision using jigs, followed by electrically connecting the front-end hybrids to the sensors using wire bonds.

Before the components are assembled into a module they have to be qualified as replacing any of the components listed above in an assembled module is not possible. The sensors, front-end and service hybrids will therefore be tested as stand-alone objects before integration into a module.

Dedicated test benches are being developed to electrically test the stand-alone hybrids both at the manufacturer and at the assembly centres. These test benches are focussed on testing these sub-assemblies and are very different from the readout of a final module. A nice example here is the PS-hybrid: the test system will have to test the CIC chip which is on the hybrid as well as the 8 SSA chips without the MPA chips in the data path. Also tests in cold are foreseen to make sure the hybrids can be turned on and function at tracker operational temperature. Once a module is finalised it will be tested during thermal cycling, the so-called *burn-in* test. During this test the module will be connected to the test system by the optical link.



(a)



(b)

FIGURE 5.30: Illustration of the DAQ-related inter-chip connections in a 2S (TOP) and PS (BOTTOM) module.

5.8 Data, trigger and control system

In the final DAQ system the back-end Data, Trigger and Control (DTC) card will be a custom designed ATCA board using Kintex Ultrascale FPGAs. Each DTC board will control, configure and receive data over optical fibres and multi-channel optoelectronic transceivers from 72 modules. The data stream consists of three flows: the full-event data transmitted upon an L1A trigger (DAQ steam), stub data received at BX rate (TRIG stream) and the TTC&CTRL stream which includes clocking, fast commands and the slow control and configuration. The downlink and uplink have a bandwidth of respectively 2.56 Gb/s and 5.12 or 10.24 Gb/s for each of the 72 modules. The current prototype architecture for the DTC is an ATCA card with two high-end FPGAs and a system on chip for monitoring, control and calibration procedures.

The DAQ stream is forwarded by the DTC to the CMS central DAQ system with an estimated event size of 1.15 MB[10] and the stub data stream is forwarded to the L1 track finding system.

5.9 L1 track finding

The L1 track finder is loaded with the task to make tracks from the stub data provided by the p_T modules. The system is designed to perform tracking for particles with $p_T > 2\,\mathrm{GeV/c}$. The track finder has to construct these tracks out of an average of 15k stubs per BX and has to do this in 5 μs to fit in the overall latency budget of 12.5 μs. The L1 track finder will be implemented using FPGAs and four steps can be distinguished in the track finding process: data organization, pattern recognition, track fitting and duplicate removal.

5.10 Phase-2 tracker expected performance

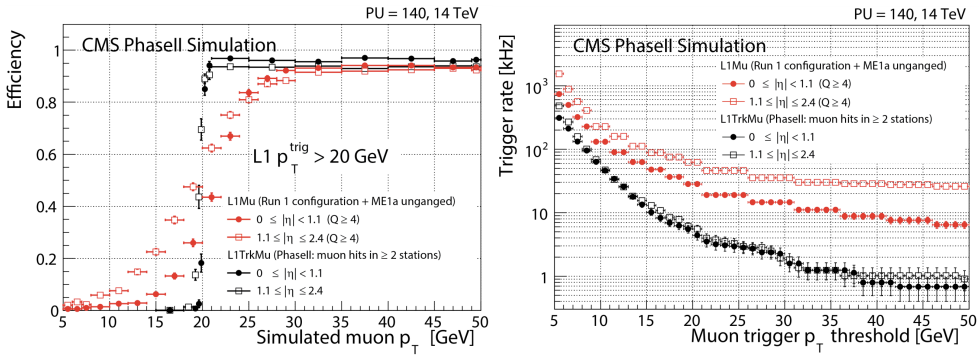
Tracking performance will greatly benefit from the increased granularity and the smaller material budget of the upgraded tracking system. To give an example: the material budget, expressed in radiation lengths, within the full tracker volume will decrease with a factor ≈ 2 at $|\eta| = 1.5$ and therefore reduce the multiple scattering of the charged particles.

Extensive simulations were run to validate the performance of both L1 and offline tracking using both the upgraded pixel and Outer Tracker detector. The simulations include an accurate description of the sensor and electronics response, detector granularity and layout and take into account effects due to pileup to which optimised tracking algorithms were applied.

The simulations of the L1 tracking show sharp turn-on curves and efficiencies saturating at approximately 98% for muons with a p_T larger than ≈ 4 GeV/c over the full η range in an environment with an average of 200 pileup collisions. A similar environment results in L1 tracking efficiencies for electrons saturating at 90% [33].

The importance of the addition of the track trigger to the global trigger at HL-LHC is exemplified in Figure 5.31a and 5.31b where a case study is shown for a single muon trigger at pileup of 140 simultaneous collisions. The difference between an L1 system with only input from the muon chambers on one hand and the combination of tracker and muon chamber information on the other hand is shown. It is clear from Figure 5.31a

[10] The estimated full event size for CMS Phase-2 is 7.4 MB [61].



(a) (b)

FIGURE 5.31: Result of the L1 track trigger input to the L1 trigger system for the example of a single muon trigger with 20 GeV/c p_T threshold. The turn-on curve (LEFT) of this trigger is much sharper in the case where muon system and tracker information is combined with respect to an L1 trigger system which only receives input from the muon system. This results in a decreased L1A trigger rate (RIGHT) [63].

that the addition of tracker information results in much sharper turn-on curves at the requested p_T which as a result greatly reduces the L1A trigger rate, as shown in Figure 5.31b, due to the decreased probability of misidentifying low momentum muons as high momentum muons. A decrease in L1A rate with a factor of 10 is obtained for this case [63].

Besides substantially decreasing the L1A trigger rate, the inclusion of tracking at L1 opens up the opportunity for new online event selection algorithms at HL-LHC. In Ref. [74] for example, the possibility of developing a dedicated trigger for displaced tracks is discussed, which could be used to target physics signals with tracks being produced a few mm from the beamspot.

The efficiency of the offline tracking at Phase-2 is shown in Figure 5.32a and 5.32b with respect to p_T and η respectively. Both results are obtained for tracks from $t\bar{t}$ simulated events with average PU 140 and 200. Figure 5.32a shows a tracking efficiency higher than 90% for tracks with a p_T higher than 1 GeV/c. Also in η efficiencies above 90% are attained almost for the full coverage.

The resolution of the offline tracking is illustrated in Figure 5.32c and 5.32d showing the p_T resolution and the resolution on the transverse impact parameter (d_0) for both the Phase-1 and the Phase-2 tracker. For both quantities, the Phase-2 tracker shows an improved performance over the full η range with transverse impact parameter resolution going below 10 μm for the most central regions.

5.11 Summary

The performance of the current CMS strip tracker is degrading due to radiation damage. This, together with the increased pileup foreseen for HL-LHC requires the strip tracker to be replaced by a new detector. This Phase-2 Outer Tracker is designed to operate at high pileup conditions and will also allow to use tracker information in the L1 trigger system.

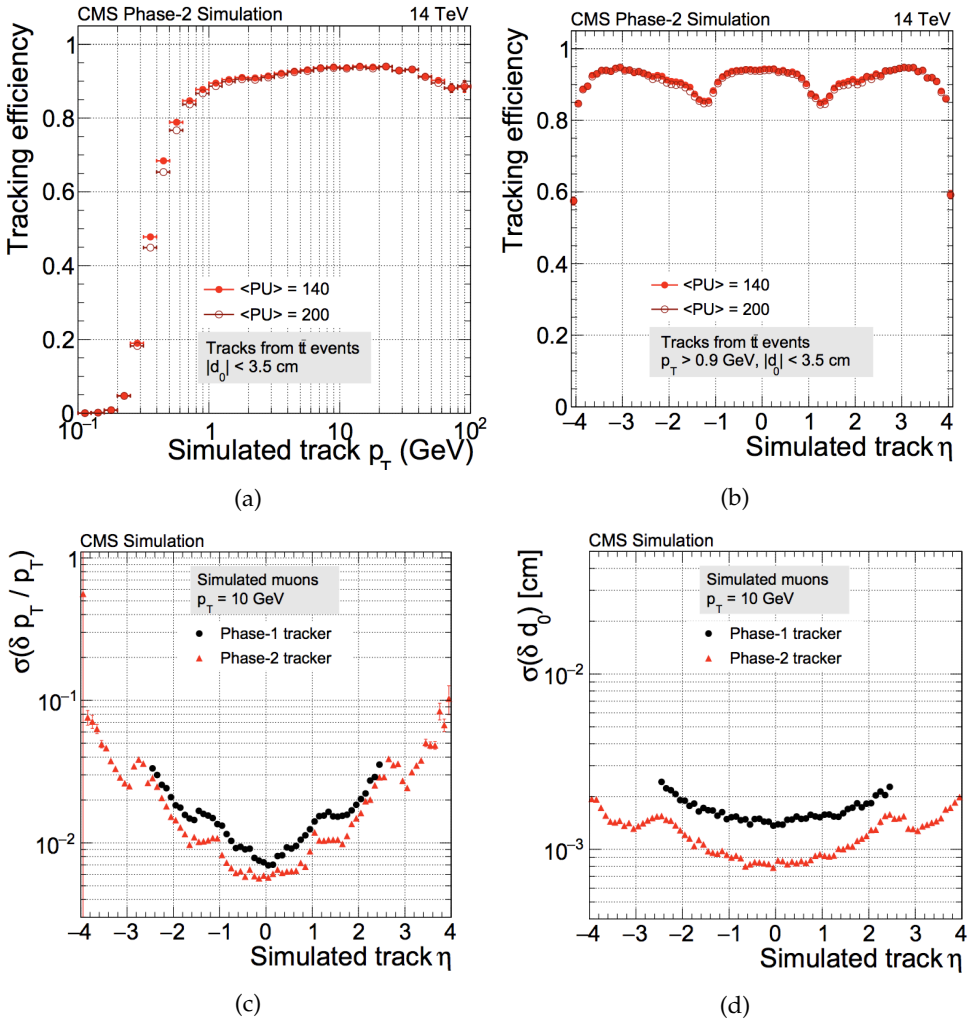


FIGURE 5.32: TOP: Track reconstruction efficiency as function of p_{T} (LEFT) and η (RIGTH) in $t\bar{t}$ events at a pileup of 140 and 200 for tracks produced with a transverse impact parameter smaller than 3.5 cm. For the η dependence tracks with $p_{\mathrm{T}} \leq 0.9\,\mathrm{GeV/c}$ were discarded. BOTTOM: Comparison between the Phase-1 and Phase-2 tracker for the resolution on the transverse momentum (LEFT) and transverse impact parameter (RIGHT) for muons with a transverse momentum of 10 GeV/c [33].

To accomplish the latter, the CMS Outer Tracker will adopt a unique p_T module design which will allow for on-module generation of high p_T track primitives (*stubs*) which can be sent from the module to the back-end at BX rate. Two types of p_T modules will be used to equip the Outer Tracker volume. The different silicon sensor designs used in these modules require three specific readout ASICs to process the analogue signal, form the stubs and store the full-event information. A concentrator ASIC will be used to group and to select the highest p_T stubs and to process and buffer the full-event information per module side. In the end this results in two data streams to the back-end: the stub data and the full-event data. During prototyping and construction, single chips, chip assemblies and prototype modules need to be tested. These devices all need a test bench to validate their operation and a good understanding of the ASICs' functionalities and the data and control paths on the module is indispensable when designing such a test system.

Chapter 6

Firmware Development for Phase-2 Tracker Prototypes

6.1 Introduction

In order to qualify the CMS Phase-2 Outer Tracker module prototypes, a test system was set up which covers the readout of all possible DUTs (Device Under Test). This test system will be one of the main test benches for testing during R&D and module production. Therefore, this system has to cover a fairly large number of devices, starting from bench-top testing of single chips to multi-module testing in beam or during burn-in and this for both 2S and PS flavours. The readout system consists of a firmware part, the so-called μDTC, running on an FC7 [75] AMC (Advanced Mezzanine Card) card and the Ph2 ACF (Phase-2 Acquisition and Control Framework) C++ software package which controls the firmware to run the calibration routines and data taking. A top level view of the readout framework is shown in Figure 6.1: at the lowest level the hardware prototypes which need to be configured, controlled and read out are located. They communicate over a dedicated link to the FC7 readout board. This link can be electrical or optical depending on the type of prototype which is connected. The FC7 runs the hardware specific μDTC firmware which makes use of the system firmware provided by the FC7 team. This system firmware implements, amongst others, the firmware blocks for communication with the FC7 board peripherals and an IPbus core (see section 6.4.2) for communication to the back-end. The back-end is a Linux machine running the μHAL library which is a low level library to operate the IPbus communication over the machine's Ethernet port. DUT specific and firmware specific settings and routines are coded into the Ph2 ACF's *Middleware layer* which makes the *User Routines layer* as agnostic as possible of changes in firmware or chip versions. Examples of user routines are latency scans and tuning of the front-ends of the ASICs.

This chapter will describe in more detail the layer connecting the DUT with the software: the μDTC system, with as its backbone the FC7 card carrying a Kintex7 FPGA. This chapter contains an introduction on FPGAs (section 6.2), a word on the scope of the μDTC project (section 6.3), an overview of the hardware platform (section 6.4) and a detailed description of the firmware blocks designed for the μDTC (section 6.5).

6.2 FPGAs

Field Programmable Gate Arrays (FPGAs) are chips consisting of a matrix of configurable logic blocks (CLBs). Each CLB contains the same elementary logic (LUTs, Digital Signal Processing (DSP) units, multiplexers, flip-flops). The interconnection between the CLBs and connection to other FPGA resources is performed using a reprogrammable switching

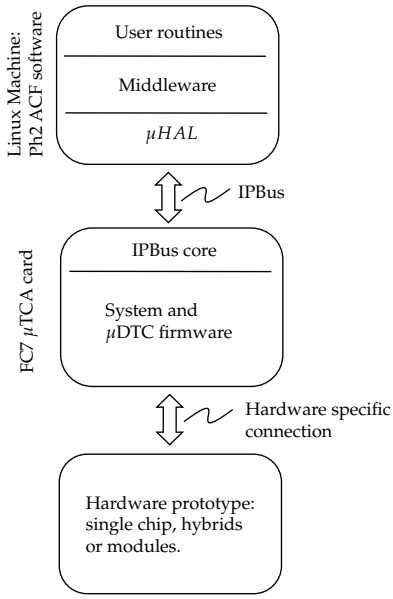


FIGURE 6.1: Top level view of the readout system for Phase-2 Outer Tracker prototypes.

matrix which is a vast network of wires and switches to interconnect different parts of logic.

Next to the CLBs, dedicated hardware elements (e.g. block memory, ISERDESes (input serializer/deserializer), OSERDESes (output serializer/deserializer), clocking resources (e.g. MMCM (mixed-mode clock manager)), dedicated I/Os,...) are integrated on the die. A design like this makes FPGAs apt for a multitude of applications with the main advantage over ASICs being the fact that FPGAs can be reprogrammed.

The architecture of a Xilinx 7 [76] series FPGA (the FPGA on the FC7 board is of this family) is used here as an example to have a look at the typical configuration of a CLB slice. Each CLB of a 7 series FPGA contains two so-called *slices* as shown in Figure 6.2. The layout of the slice itself is shown in Figure 6.3. Each slice contains four 6-input LUTs which can be used to programme any 6-input combinatorial logic function. Besides the LUTs, the slices contain several multiplexers for efficient utilization of the resources, dedicated hardware carry logic to perform arithmetic functions (DSP) and flip-flops to implement synchronous operations. Approximately one-third of the slices on a Xilinx7 series FPGA allow the LUTs in these slices to be used as distributed 64-bit RAM.

Designing so-called firmware code for FPGAs most commonly proceeds through hardware description languages such as VHDL or Verilog. Software packages such as Vivado are used to simulate the design and to generate the bitstream which is eventually uploaded to the FPGA.

6.3 Scope of the μDTC test-bench

The μDTC has to support the control and readout of several devices based on the ASICs described in section 5.5. The control and configuration protocols of these chips are fairly

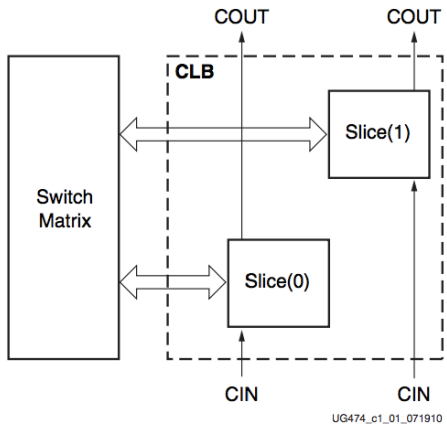


FIGURE 6.2: Slices within a CLB connected to the switching matrix [76].

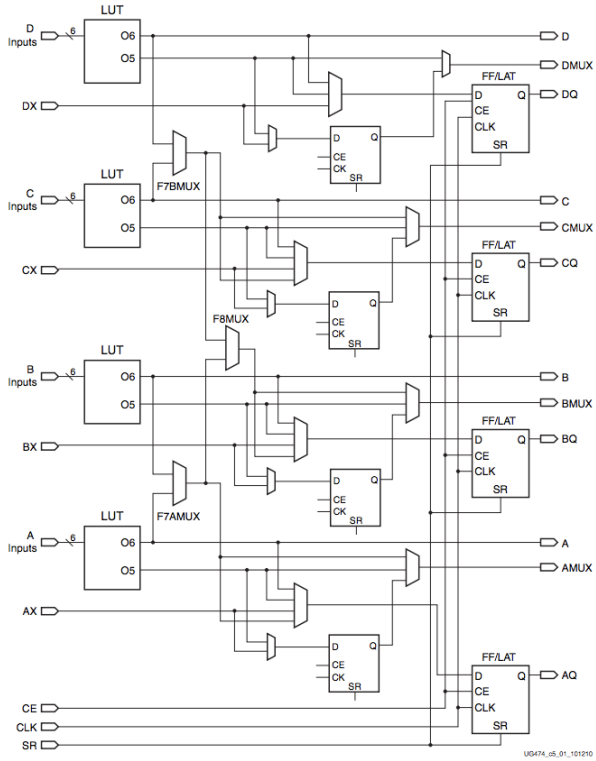


FIGURE 6.3: Simplified configuration of a Xilinx 7 series slice [76].

similar which advocates for the use of a single firmware project which implements the logic for all possible DUTs. The main difference between the ASICs is the output data format. This can however be caught in a single firmware layer, the so-called Physical Abstraction Layer (see section 6.5), which delivers the data to the next firmware layer in a format which is chip independent. Some of the prototype devices which were developed and which the μDTC supports are described in section 7.2.2 and 8.2.4 for the 2S and PS case respectively. Besides the variety in DUTs, the number of DUTs connected to the μDTC is also changing from set-up to set-up. Therefore the number of chips per hybrid and the number of hybrids itself is a parameter in the firmware configuration. As long as the readout of the DUTs is electrical the number of devices which can be connected to an FC7 is limited by the limited I/O (section 6.4) capabilities of the FC7. Once the communication towards the front-end is done optically the number of modules that can be connected to a single FC7 will increase.

Many of the firmware blocks written for the μDTC are R&D specific. In the final DTC system the variety of devices to read out is limited and as a result the number of required firmware images is small compared to the number during prototyping. During prototyping however, the number of devices is large, mainly due to the different ASICs which need to be interfaced with. The μDTC project's time line is therefore largely dictated by the availability of new hardware. This hardware is getting more and more mature and is reaching the final architecture. Therefore the μDTC also evolves to a configuration to read out a final module which basically means receiving data from two CIC chips over an optical link, which is the same configuration as will be read out by the DTC system. Some strategies or firmware blocks might therefore be transitioned from the μDTC to the DTC.

Besides for R&D, the μDTC will also be used during module production where it will be operated to do the acceptance testing of the front-end hybrids, the single module testing and the burn-in of the modules.

6.4 Hardware platform

The hardware platform chosen to operate the test bench for CMS Phase-2 tracker prototypes and testing during module production is based on the μTCA (Micro Telecommunications Computing Architecture) hardware environment [77]. A specific DAQ and control card, the so-called FC7 [75], engineered according to the AMC specification, forms the backbone of the testing. In this section the μTCA standard, the IPbus protocol, the FC7 AMC card and a typical test set-up will be introduced.

6.4.1 The μTCA standard

The CMS DTC system will follow the ATCA (Advanced Telecommunications Computing Architecture) standard for the implementation of the DTC card. The TCA standard is a common standard for data acquisition applications in telecommunication. For the test systems a downscaled version of the ATCA, the μTCA standard, is used. The choice for this standard with respect to the ATCA standard is driven by price, development effort, throughput and form factor considerations: ATCA cards are much more expensive than the μTCA cards, already developed μTCA cards, such as the FC7 card, can be used and the maximal number of modules which the μDTC will have to read is of the order of eight. Using an ATCA card to do this would be overkill and the smaller form factor of

the μTCA cards (73.8 $\times$ 181.5 mm^2 for single width or 148.8 $\times$ 181.5 mm^2 for double width) with respect to the ATCA card (280 $\times$322 mm^2) makes them much more practical for use during R&D and module production.

Multiple μTCA blades can be operated in μTCA crates (see for example Figure 6.9b) which also house powering modules and a shelf manager, the so-called MCH (MicroTCA Carrier Hub). This MCH controls the powering, cooling, hot swap and failure scenarios of the AMCs in the crate. The shelf manager and the AMC cards are interconnected over the μTCA crate's backplane. The MCH is also used as data hub to the AMCs. MCHs implementing different kinds of protocols (Gigabit Ethernet, PCI Express,...) are available. External systems can connect to the MCH via e.g. front panel Ethernet ports.

Besides using the FC7 μTCA card in a crate, the card can also be used in a bench top configuration, as illustrated in Figure 6.5. In this case the backplane is simply replaced by a dedicated PCB (Printed Circuit Board) which enables connection to the back-side adapter on the FC7.

6.4.2 The IPbus protocol

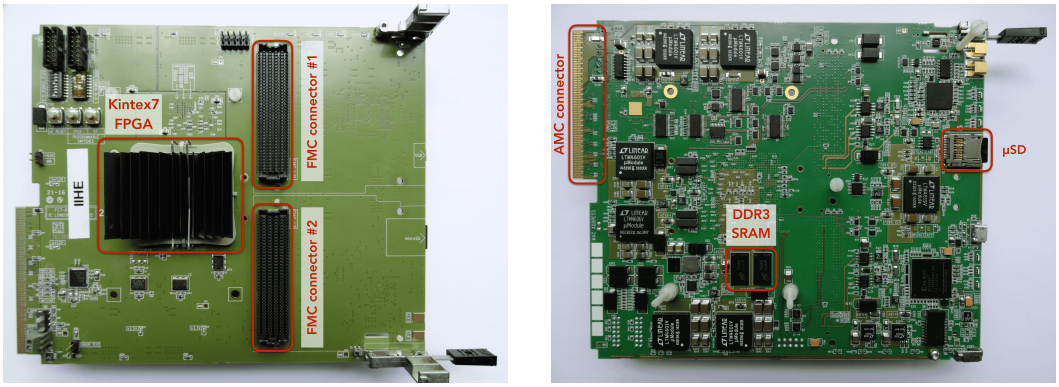
The IPbus protocol [78] is a reliable and high performance control link specifically designed for use in particle physics systems as a control link for μTCA and ATCA based systems. The IPbus suite consists of a software interface (μHAL) and a firmware block which can be adopted by users. The IPbus protocol operates over a bus with 32-bit addressing, which is more than sufficient for most applications, and 32-bit data transfer. The IPbus packet is packed into a UDP, IP and Ethernet packet. The IPbus packet itself consists of a header, followed by a number of IPbus transactions, which each consist of a header and a body. The header specifies the request (read or write) and the size of the request, specifying the depth of the transaction. FIFO read and write are also implemented where the address stays the same. On the target side the IPbus protocol is decoded by a master which forwards commands to IPbus slaves which each cover a certain address range and can handle the decoded IPbus requests the way they like (configure a setting, start a state machine,...) and respond with an acknowledge and data in case of a read.

Most efficient transfer of data happens if there are many IPbus transactions in one package, as this limits the relative contribution of the headers to the transferred data. Through a combination of packet IDs, a protocol between client and target to retrieve information on the target and resends of lost packages the IPbus protocol ensures reliable transactions.

The IPbus throughput, defined as the amount of user data transferred or received per unit of time, was measured to be around 0.5 Gb/s [78] for 1-to-1 block transfers with blocks containing a payload larger than 1 MB. This is for the case where the software controls a single hardware device.

6.4.3 The FC7 AMC card

The FC7 AMC card, shown in Figure 6.4, is a custom designed DAQ and system control card for applications in the CMS collaboration. It is the successor of the CMS GLIB AMC card and was designed initially for usage in for example the CMS Pixel Front-End Driver which performs the DAQ of the CMS Phase-1 Pixel detector. The central component on the FC7 is a Kintex7 XC7K420T FPGA which is programmable by the user to implement the experiment specific DAQ and control logic. Some of the available resources on this specific FPGA are summarised in Table 6.1. The FC7 also hosts a 4 Gb (128M$\times$32b) DDR3



(a) (b)

FIGURE 6.4: The FC7 AMC card with some important components highlighted on both the top (LEFT) and bottom side (RIGHT).

RAM memory and 2 FMC (FPGA mezzanine card) connectors which can be used to interface to the front-end. The FMCs are of the Low Pin Count type, each delivering 34 differential I/Os, but in total also 20 (12+8) high speed serial lanes are made available. The high speed lanes to the front-end allow running high bandwidth protocols, such as the GBT protocol (see section 5.5.7), for communication to the front-end. The communication to the back-end typically happens over IPbus or, for the data, over dedicated pins on the AMC connector which connect over the back-plane to another dedicated AMC card such as an AMC13. This last option was not used yet for the μDTC system and will therefore not be discussed here.

Slices	6-input LUTs	Max distributed RAM (Kb)	Shift register (Kb)	Flip-flops	Max block RAM (Kb)
65,150	260,600	5,938	2,969	521,200	30,060

TABLE 6.1: The main resources of the Xilinx Kintex7 XC7K420T [79].

6.4.4 Example of a typical μDTC set-up

Figure 6.5 shows a typical configuration of the electrical readout of a piece of prototype electronics. Here the example of a 2CBC3 hybrid is shown, but the components are typical for each set-up: the device under test connects to a custom made interface card, which typically provides power regulation and hosts level translators to do the conversion to/from SLVS from/to LVDS (Low-Voltage Differential Signaling) signals. This interface board connects to the FMC connectors on the FC7 through an FMC card. The FC7 is shown here in the bench-top configuration where the AMC connector plugs into a dedicated PCB which replaces the basic functionalities of the back-plane. The main functions of the PCB in this case are the power distribution and making available the Ethernet port which allows IPbus communication.

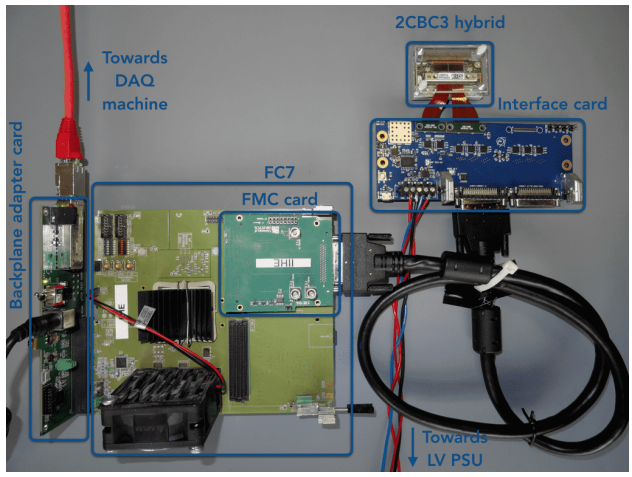


FIGURE 6.5: Example of a typical μDTC set-up for electronics testing. The DUT, in this case a 2CBC3 hybrid, is connected to a dedicated interface card. A custom made FMC card plugs into the FC7's FMC connector. The FC7 is operated here in the bench top mode where the AMC connector plugs into an adapter card which provides the powering to the FC7 and makes the Ethernet connection available. The Ethernet cable goes to the DAQ PC which communicates to the FPGA using the IPbus protocol.

6.5 μDTC firmware structure

6.5.1 Introduction

Several application specific firmware blocks, so-called user blocks, are added to the system firmware provided by the FC7 team. These firmware blocks form the μDTC firmware. A block diagram representing the structure of the μDTC is shown in Figure 6.6. Each block has its own function and given a well defined interface between the blocks, changes to block X will minimally affect block Y. The μDTC user code heavily relies on the IPbus interface to the back-end. An IPbus core is provided in the system firmware. This core interfaces with a μDTC specific firmware block which interfaces to each of the user core firmware blocks in this way enabling the control, configuration and status control of the entire user firmware.

Slow control configuration of the DUTs are also initialised over IPbus. The IPbus commands for slow control go into a FIFO, waiting for the previous command to have been executed, in the Command Processor Core and are passed on to the Physical Interface Abstraction Layer, hereafter referred to as the Phy layer, where they are translated to a DUT specific protocol and routed to the DUT which is addressed.

The Fast Command Core contains several state machines to generate dedicated fast command sequences, with BX timing resolution. Furthermore it can receive for example trigger and clock from a DIO5 (see section 6.5.5) or from the AMC13 core and forwards these to the Phy layer where they are passed along on the payload of the fast command line according to the format expected by the front-end chip.

Also for the data the μDTC needs to be transparent to the device to which it is connected, to have the blocks downstream of the Phy layer generic across DUTs. As for fast

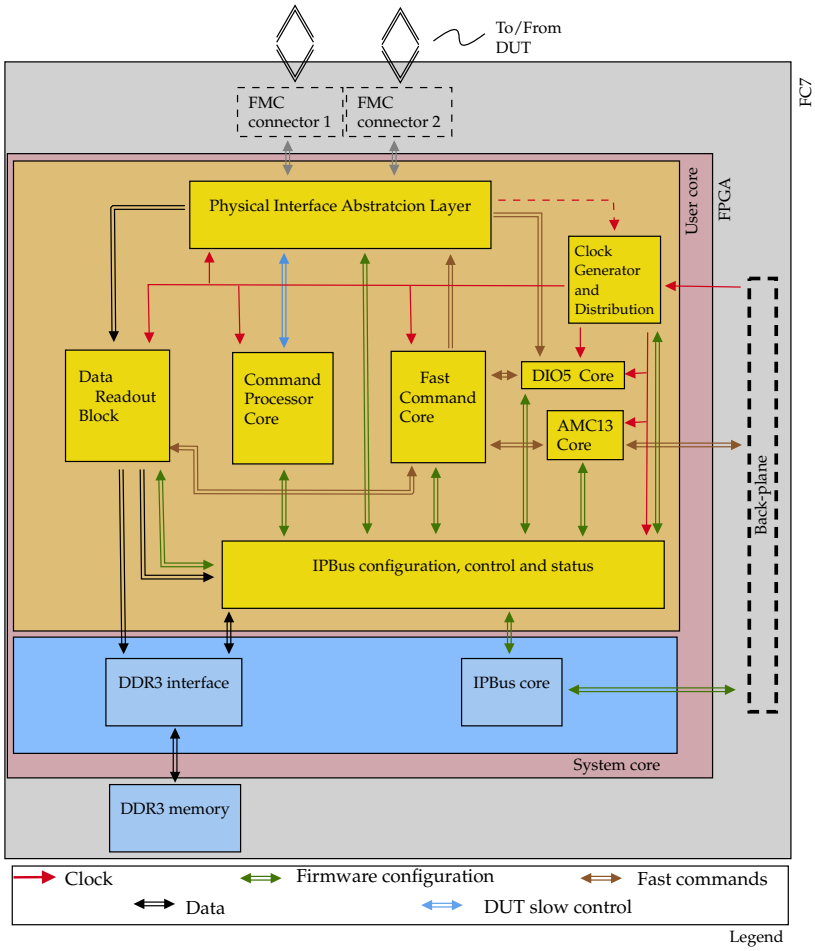


FIGURE 6.6: Top level view of the μDTC firmware showing only the most relevant blocks.

and slow control communication, the changes in data format due to a change in hardware type are mostly caught in the Phy layer. In this layer, the data from the front-end is received and formatted to a more generic form. A more detailed description of the Phy Layer is given in section 6.5.2. The formatted data from the Phy Layer is sent to the Data Readout Block (section 6.5.3) where the data is buffered and awaiting readout over IPbus or is sent to the DDR3 memory on the FC7.

All the blocks are clocked from the Clock Generator and Distribution block which uses the FC7's oscillator or an external clock as a reference. Most of the firmware runs at a frequency of 40 MHz. There are some exceptions where higher clock frequencies are used to save on logic utilisation.

The following sections give a bird's eye view on the μDTC firmware. Some specific blocks, which are dedicated for specialised tests will be described in later chapters where the testing is also explained.

6.5.2 The Physical Interface Abstraction Layer

The Physical Interface Abstraction Layer forms the interface to the front-end. It also abstracts the incoming and outgoing data formats for different front-ends in order to minimise the changes which need to be done on the downstream part of the firmware. The Physical Interface Abstraction Layer interfaces the fast commands, the slow control commands and the data from the front-end with the other blocks of the firmware. More information on these data streams can be found in the following sections.

6.5.2.1 Fast commands

For the fast command, the Physical Interface Abstraction Layer forms the transition from the 40 MHz regime to the 320 MHz regime using an OSERDES. The 110XXXX1 pattern is continuously sent whilst a 4-bit bus at 40 MHz is received from the Fast Command Block which holds the fast commands to be transmitted to the front-end. The phase, with respect to the 320 MHz clock, of the fast command going to the chip is configurable.

6.5.2.2 Slow control

The slow control part of the Phy layer is illustrated in Figure 6.7. The input slow control command from the Command Processor Block to the Physical Interface Abstraction Layer specifies: whether a broadcast to all chips is requested, the hybrid id (4 bits), the chip id (5 bits), the page (only for CBC chips), read/write bit, the register address (up to 2 bytes to support 16-bit addressing for the MPA and SSA) and the data to be sent (multiple bytes to support the sequential read/write).

The hybrid id is used in the multiplexer which forwards the slow control command to the correct hybrid's master and waits for an acknowledge from this master signifying the acknowledge from the front-end chip. When a command is forwarded to a hybrid, the multiplexer requires an acknowledge within a specified time and if not received it generates a time-out error. This time-out error is propagated to the Command Processor Block.

For each hybrid the wrapper handles for example the register paging system which is used in the CBC chips. It also translates the chip id (5 bits) to the chip address (7 bits). The exact translation can be configured over IPbus. Furthermore it handles the sending of the correct chip address in case a broadcast to all chips is requested.

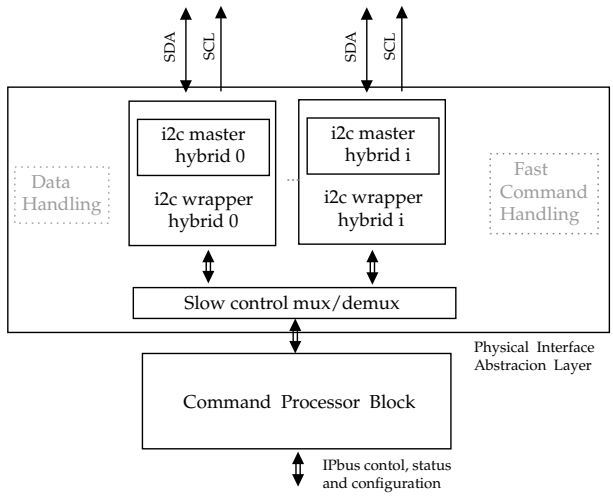


FIGURE 6.7: Layout of the slow control command handling in the Phy layer.

The wrapper then in turn interfaces with the real I^2C master. Here an I^2C master state machine is encoded. It drives both the clock (CLK) and the data (SDA) line of the I^2C. The I^2C clock frequency is configurable and the nominal and maximum frequency is 1 MHz.

The response from the slave is then sent back to the Command Processor Block. It passes the data read from the I^2C register in case of an I^2C read, or just an acknowledge in case of a write. Also as discussed above this bus is used to propagate an error to the Command Processor Block.

6.5.2.3 Phase tuning of the incoming data

The phase of the incoming data with respect to the FPGA's 320 MHz sampling clock is unknown and the correct sampling phase needs to be extracted from the incoming data itself. This is performed by inputting a known repetitive pattern (e.g. the CBC3 sync bit line) from the front-ends to the FPGA and delaying the input line using the IDELAY resources in the FPGA which allows for a resolution of 78 ps. When scanning the IDELAY range the bit transitions can be extracted and the sampling in the metastable regime can be avoided. The IDELAY setting can be extracted for all incoming data lines as long as a predefined repetitive pattern can be configured on this line. Besides the phase also the correct 8-bit word can be extracted given an adequate pattern which allows determining the first bit in the byte. This is called the word alignment and is performed by changing the *bitslip* setting in the ISERDES.

6.5.2.4 Front-end data acquisition

For every data line[1] coming to the FPGA an ISERDES is used to translate a single data line at 320 MHz to an 8-bit wide bus at 40 MHz. In this way the internal logic on the FPGA can run at 40 MHz which makes timing closure of the FW design much easier.

The handling of the full-event and stub data happens in separate logic and there are as many full-event and stub decoders in the firmware as there are front-end chips from

[1]Except for CBC2, there the data lines operate at 40 MHz instead of 320 MHz.

which the FPGA receives the data. The full-event and stub data formats which the μDTC has to decode were already introduced in Figures 5.12a, 5.12b, 5.13, 5.16b, 5.17, 5.22b, 5.23, 5.26 and 5.27.

The full-event data acquisition across chips of different flavour is very similar: the full-event data decoder block receives 8-bit words, which are properly phase and word aligned, from the ISERDES instance and looks for the full-event header in the data stream. When the header is found the full-event frame is aggregated in a state machine and built into a multiple of 128-bit words which are sent to the Data Readout Block. In case of the CBC and SSA chip the length of the full-event data package is fixed and thus the size of the data package to the readout block is also fixed. In case of the MPA and CIC chip the full-event data is sparsified and the full-event data decoder interprets the header data, containing the number of strip and pixel clusters which will follow in the payload, to know how the rest of the aggregation should proceed.

The stub data is synchronous, so the state machine which aggregates the stub data runs continuously. In case of the CBC (SSA), the 5 (8) bytes of stub data are parsed into a package of 40 (64) bits and passed to the back-end. In case of the MPA, the stub data package is spread over two 40 MHz clock cycles and can contain data from two BXs. In order to do the event building, the MPA stub data has to be *boxcar unpacked*: the header of the MPA stub package is interpreted, the 2 BX long package is buffered and then the first clock cycle after the buffering the data attributed to the first of the two BX is output to the Data Readout Block followed in the next clock cycle by the data attributed to the second BX. For the CIC also the stub data package contains data from multiple BXs. How the boxcar unpacking is achieved in the case of the CIC is explained in more detail in section 9.2.4.2.

6.5.3 Data Readout Block

The Data Readout Block receives two data streams from the Phy layer: an asynchronous full-event data stream and a synchronous stub data stream.

In the end, the μDTC should build events, which means that it should match stub and full-event data which belong to the same event. To do this, two parameters need to be tuned: the L1 trigger latency and the stub latency. The L1 trigger latency is a setting on the front-end chip which programmes the depth of the on-chip memory for the full-event data. The setting is dependent on the latency between the reception of the data in the chip's full-event data memory and the arrival of a trigger over the fast command line and is thus directly dependent on the latency introduced by the trigger system (speed of trigger logic, length of cables, etc). All the full-event frames which are received from the Phy layer are written into a dedicated FIFO for the full-event data.

In the Data Readout Block the stub data is passed through a shift register with a programmable depth. In this way the stub data is delayed in the firmware in order to be able to match it with the full-event data. Only the stub data at a certain depth (referred to as the stub latency), with respect to an L1 trigger, in the shift register will be written to the stub readout FIFO.

Both the stub and L1 latency parameters can be obtained by scanning both latency ranges for data. When both L1 and stub latency are correctly set the stub data will match the full-event data and an event containing both types of data can be built.

There is one full-event FIFO and one stub FIFO for each chip from which the firmware receives data. Besides these data FIFOs also FIFOs are present for saving counters such as FC7 L1 counters, FC7 BX counter, TDC values,...

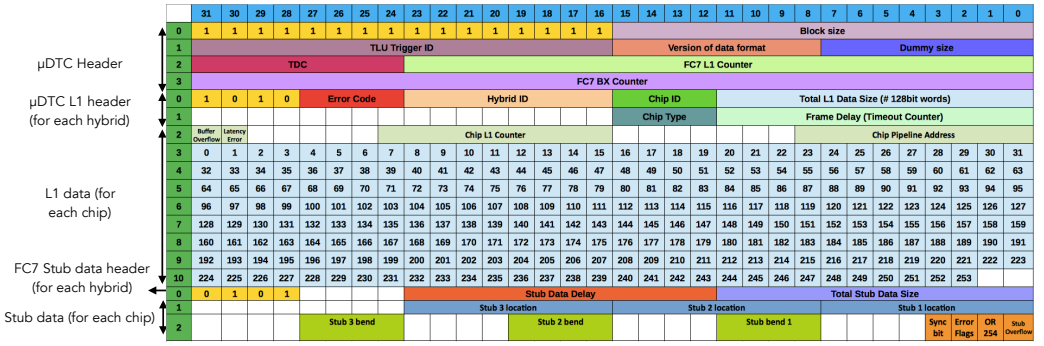


FIGURE 6.8: Example of an event format as sent from the μDTC over IPbus to the DAQ machine. The example shows the format for a single CBC3 chip. The length of the package scales with the number of hybrids and number of chips.

The output from all FIFOs are assembled and for a single event are foreseen of an event header and for each hybrid data with a hybrid header. This event data package is packed into a FIFO or in the DDR3 memory on the FC7. Which memory is used is selected at compilation time of the firmware. When a specified amount of events has been written to either the DDR3 or the FIFO the firmware raises a flag which is polled over IPbus by the software. When the software sees the flag is raised, the data readout over IPbus to the back-end is started. An example format of how the data is eventually transmitted over IPbus is shown in Figure 6.8.

6.5.4 Command Processor Block

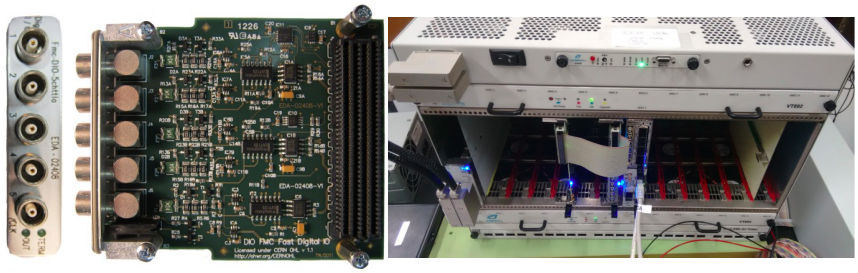
The Command Processor Block stores and interprets the slow control commands sent over IPbus and passes them on to the Phy layer where the slow control command is translated to the front-end chip specific protocol, typically I^2C.

The Command Processor Block contains some intelligence to send, based on specific flags in the IPbus frame, specific requests to the front-end. There is for example the possibility to use a flag in the IPbus request package to apply the same slow control command to all the front-end chips on all the hybrids connected to the FC7.

6.5.5 Fast Command Block

The Fast Command Block generates the payload for the fast command. The generation of the payload happens at 40 MHz, the translation to the 320 MHz as required in the fast command protocol of CBC, MPA, SSA and CIC happens in the Phy layer. The Fast Command Block muxes the fast commands from different sources. Examples of these sources are:

- State machine to send triggers at a user defined frequency.
- State machine to send test pulse triggers and triggers with a user defined spacing and frequency.
- Triggering on stubs: the stub data payload is checked, for example for the presence of a stub or for the active high of the OR254 bit in the CBC3 stub data package. Like this a module can be set in a self triggering mode.



(a) (b)

FIGURE 6.9: LEFT: The DIO5 FMC, very useful for its LEMO connectors which can be used to interface easily with the FPGA logic. RIGHT: Two interconnected FC7s, one running the emulator firmware and the other the DAQ firmware, to test the first features of the μDTC firmware. The FC7s are operated in a μTCA crate which can host up to 12 AMC cards. Centrally in the crate two half-width AMCs are visible: the MCH receiving the Ethernet cable at the bottom and the AMC13 at the top. The power modules for the crate are located at the far left of the crate.

- External signals delivered through a dedicated FMC mezzanine such as a DIO5, shown in Figure 6.9a which provides LEMO inputs.

The user can select which of these sources to use.

6.5.6 Firmware testing

To commission the DAQ chain of a new DUT, emulators of chips, running on FPGA, were written to be able to test the firmware without having the actual chips or assemblies in hand. This is very useful as first functionality of the DAQ chain can be proven, but furthermore the FW, being an important part of the test bench of a new device, should already be verified so that testing of the new device can be done with an already validated DAQ chain.

These emulators are very rough approximations of the real device: the full stub and full-event logic is not implemented. The emulator typically has an I^2C slave and some registers implemented to test the slow control, outputs a configurable (over I^2C) pattern on the stub lines and responds to a trigger by outputting a configurable (over I^2C) full-event frame. The emulators can be configured to run on the same FPGA as the μDTC or can be configured to run on a separate FPGA. In the former case the emulator's I/O lines are routed internally to the Phy block and as a result some parts of the logic, such as the ISERDESes, are bypassed. In case the emulator runs on an external FPGA, the set-up resembles very closely the system which is aimed for: the DAQ FC7 receives data over the FMC cards which would be used during testing of the new device. In this way the full DAQ firmware is tested: from I/O pin mapping, through phase and word tuning, ISERDESes, Phy layer and up to the back-end firmware blocks. In this case the external emulator runs on a 40 MHz clock which it extracts from the fast command line, as the real ASICs do. This is required to keep the emulator and the DAQ synchronous whilst not having to change the DAQ FW. The external emulator also incorporates OSERDES instances which serialise the data from the 40 MHz to the 320 MHz regime.

6.6 Summary

The μDTC test bench, based on the FC7 AMC card, is used to qualify prototype ASICs and modules for the Phase-2 Outer Tracker and will also be used for qualification of components and modules during module construction. The μDTC is used to configure, control and acquire data from the DUT. The firmware's block structure allows most of the firmware to be left unchanged when going to a different flavour of DUT. Most of the changes in data format are caught in the μDTC's Physical Interface Abstraction Layer. Validation of the test bench is required before actual testing with the real hardware can start. This is mostly done by developing emulations of the DUT which can run on a second FC7.

6.7 Author's contribution

The author was one of the main contributors to the OT μDTC firmware since the start of the development. Experience had to be gained on how to structure the code and a significant effort was made to make the code generic for different devices. The author mainly contributed to the development of the Phy Layer in the firmware and to testing of the firmware's operation by using emulators developed by the author. The testing also requires a thorough understanding of the front-ends and the other firmware blocks and requires some dedicated software. The author also contributed to the latter.

Once the firmware was tested with emulators and with a first real device, the test bench was made available to the community. The author also provided support for the use of the test-bench, most importantly during test beam activities.

Chapter 7

2S Prototype Testing

7.1 Introduction

CBC chips were the first Phase-2 Outer Tracker readout chips to become available. As a result, the first goal of the μDTC project was to be able to read out and control devices based on these ASICs. This chapter will briefly discuss the specificities of this μDTC flavour and its use cases (section 7.2). In section 7.3, a study of cross-talk between CBC channels is discussed which exemplifies the usability of the μDTC for bench-top testing.

7.2 Specific firmware development for 2S based systems

7.2.1 Introduction

μDTC firmware blocks developed for the readout of CBC based devices are fairly straightforward: no boxcar unpacking of the stub data needs to be performed and the full-event data package has a fixed length. Furthermore the μDTC test bench was not used to perform the initial single chip qualification. The initial testing of the CBC chips was already performed with another set-up and during the design of the μDTC for CBC readout it could be trusted upon that e.g. the data formats were exactly the ones provided in the documentation. The μDTC was used during later stages of the development to test and qualify the CBC based prototypes which are listed in the next section and to perform test beams with CBC based prototype modules.

7.2.2 Use cases

The μDTC was used to test the following CBC-based devices:

- 2CBC2 hybrid and 2CBC2 mini-module.
- 8CBC2 hybrid and 8CBC2 full size module.
- PS-MCK FEH hybrid: a hybrid with the form factor of a PS front-end hybrid, but with two CBC2 chips as active components.
- CBC3 stand-alone: used for initial verification of the CBC3 μDTC.
- 2CBC3 hybrid and 2CBC3 mini-module.
- 8CBC3 hybrid.

These devices are illustrated in Figure 7.1.

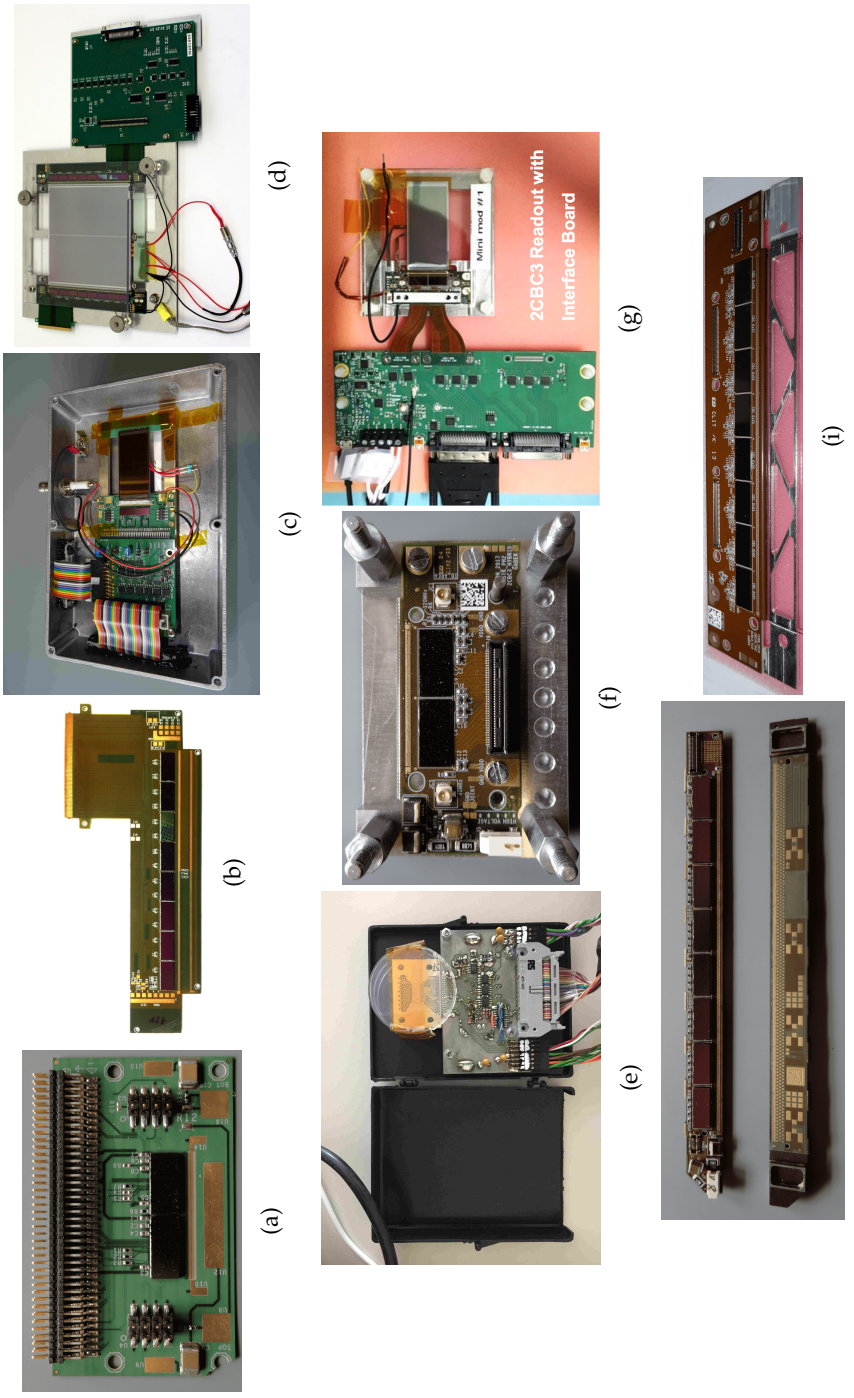


FIGURE 7.1: (a) 2CBC2 hybrid. (b) 8CBC2 hybrid. (c) 2CBC2 mini-module. (d) Full-size module with prototype 2S sensors and 8CBC2 hybrids, one of which is connected to an interface board. (e) Single CBC3 chip on carrier card connected to an interface board. (f) 2CBC3 hybrid. (g) 2CBC3 mini-module connected to its interface card. (h) PS-MCK hybrid. (i) 8CBC3 hybrid.

Several test beams with 2CBC3 mini-modules were already conducted which use the μDTC firmware for control, data acquisition and interface with the telescope and trigger system. An example result of such a test beam, conducted at the Fermilab Test Beam Facility in December 2017, is shown in Figure 7.2. This figure shows the stub detection efficiency when rotating a 2S mini-module, with 1.8 mm sensor separation, in the beam. The rotation axis is along the strip direction. The stub efficiency is measured by correlating the stub locations with tracks reconstructed with the telescope in the beam. The figure shows the stub efficiency for three different settings of the correlation window fitted with a sigmoid function. The larger the correlation window the more the device has to be rotated before the stub efficiency starts to drop. The stub efficiency reaches a plateau around 99%.

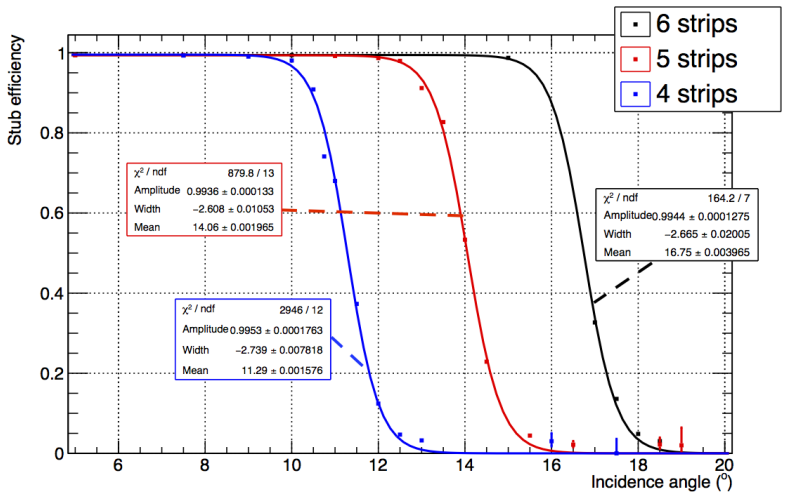


FIGURE 7.2: Stub efficiency as measured on a 2CBC3 mini-module for three different settings of the stub correlation window as given in the legend. The window size mentioned here is the search window size in one direction with respect to the hit in the seed layer [80].

7.3 CBC3 test pulse injection testing

7.3.1 Introduction

In this section the usage of the test pulse injection mechanism of the CBC3 chip is described to qualify the front-end of the CBC3 chip. Using the test pulse injection mechanism, the gain of the front-ends can be extracted. However, when measuring the gain in this way the measurement is limited by the fact that the capacitance of the test pulse capacitor is only know up to 10% accuracy which results in a 10% uncertainty on injected charge. Besides the gain, also the pulse shape can be investigated. This allows to study the shaping of the input signal and for example a measurement of the time walk. Before the gain and the time walk measurements are described, the cross-talk between channels is investigated in more detail.

Delay resolution	1 ns steps at 40 MHz Master Clock
Delay range	25 ns at 40 MHz Master Clock
Delay variation	$\pm$ 100 ps in 25 ns or $<\pm$ 5 ps in 1 ns
Charge step resolution	1 test pulse LSB = 0.086 fC = 536.7 e^-
Charge step dynamic range	255 test pulse LSB = 22 fC = 137.3 ke^-
Charge step variance due to calibration capacitor variation	0.0086 fC = 53.7 e^- ($\pm$10%)
Test groups	32 (except for the last test group which has 30 instead of 32 channels) channels are pulsed at the same time, resulting in 8 test groups for the full chip.
Connectivity	1 test capacitor per channel

TABLE 7.1: Summary of the main properties of the CBC test pulse injection mechanism [68].

7.3.2 Procedure

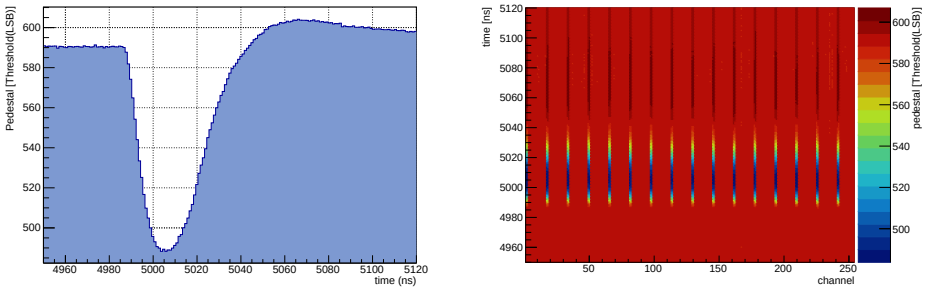
The reconstruction of a full pulse shape with the CBC3 digital readout is fairly involved. The procedure described here uses the CBC3's internal test pulse injection mechanism, described in section 5.5.5.2 and summarised in Table 7.1, and is as follows:

1. Set-up:
   (a) Tune the front-ends of the chips so that a common threshold can be applied to all channels.
   (b) Configure the firmware to send test pulse trigger fast commands followed by L1A triggers and find the correct L1 latency setting for the chip. In this specific case the L1 latency is known as it is the time between the sending of the test pulse trigger and the L1A trigger, which is a setting of the firmware state machine.
2. Run:
   (a) Make an S-curve for all the CBC3 channels whilst injecting charge. Making an S-curve consists of scanning the threshold range: at each threshold setting a number of triggers has to be sent in order to extract the average occupancy, with sufficient statistical precision, from the full-event data.
   (b) Extract the pedestals from the S-curve and save the pedestals of these S-curves for all the channels. The pedestals are defined as the threshold for which the occupancy is 50%.
   (c) Delay the test pulse using the on chip DLL which has a resolution of 1 ns and a range of 25 ns. If it is required to go beyond this range, change the L1 latency to make a 25 ns step.
   (d) Start over at 2.(a).

In the end this two dimensional scan in threshold and time will result in the reconstructed test pulses for all the channels if the pedestal, deduced from the S-curve, is plotted as function of time.

7.3.3 Results

Figure 7.3a shows an example of a reconstructed pulse shape, when injecting a charge representing a MIP (30 test pulse LSB = 30 $\times$ 0.086 fC = 2.58 fC = 30 $\times$ 536.7 e^- = 16.1 ke^-)



(a) (b)

FIGURE 7.3: LEFT: Example of a reconstructed test pulse for an injected charge corresponding to a MIP in a 220 μm sensor. RIGHT: Example of reconstructed test pulses across one CBC3 chip for an injected charge corresponding to a MIP in a 220 μm sensor.

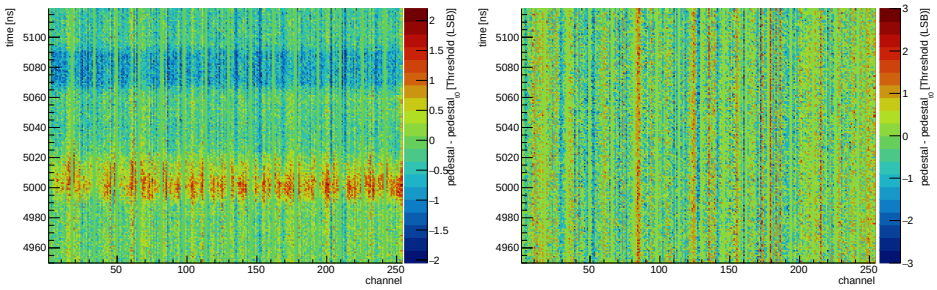
for sensors with a thickness of $\approx$ 220 μm (extracted from Figure 3.3a). The time, shown on the x-axis, reflects the L1 latency setting at which these tests were run and corresponds to a latency setting around 200 BXs. The figure furthermore shows that the rise time of the signal is less than 20 ns, which agrees with the front-end specifications.

Figure 7.3b shows the same measurement, but now across channels, with the colour code representing the amplitude of the reconstructed pulse. The channels receiving the test pulses stand out. Note that there are 32 of these channels as there are 32 channels in one test group. The other channels do not receive any directly injected charge, but it is worthwhile to investigate if any activity, and thus cross-talk, is visible on these channels.

Figure 7.4a shows the change, with respect to a time where no test pulse is induced, in pedestals and this only for the channels in which no charge is injected directly. The channels in which charge is injected are masked in this figure to make the channels not receiving the test pulse stand out. The figure shows that there is some signal injected in the non-pulsed channels around the time (around 5000 ns) when the test pulse is injected. Looking in more detail at Figure 7.4a reveals an extra feature: around 5080 ns there is also a minimum of the induced signals, reflecting an undershoot. When the hysteresis (see section 5.5.5.2) setting on the chip is turned on, the induced signal is masked. This can be seen in Figure 7.4b where an injected charge representing a MIP was injected (same as Figure 7.4a) but no induced signal is visible.

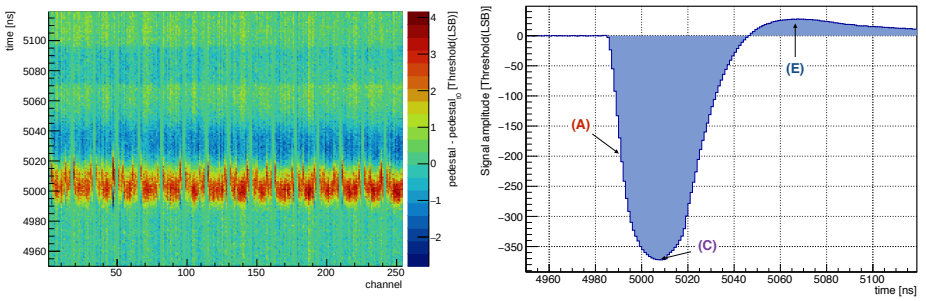
Figure 7.5a shows the same behaviour as Figure 7.4a, but this time the test pulse amplitude is set to a charge which corresponds to 4 times the charge of a MIP in a 220 μm sensor. The induced signals become even more clear. Figure 7.5b shows an example of a reconstructed test pulse when a charge corresponding to 4 times the charge deposited by a MIP was injected. Figure 7.6a shows the signal on a channel which does not receive this test pulse directly.

Figure 7.4a and 7.5a show that the induced signal does not peak at the same time for all the channels. To quantify the strength of the induced signal the maximum induced signal over the entire time range over which was scanned is used. The times at which these maxima are reached are shown in Figure 7.6b which shows clearly that the maximum signal amplitude on the channels which do not receive the test pulse directly is



(a) (b)

FIGURE 7.4: LEFT: Example of induced signals on the channels not receiving the test pulse directly (for an injected charge corresponding to a MIP in a 220 μm sensor) with the hysteresis setting disabled. RIGHT: Example of induced signals on the channels not receiving the test pulse directly (for an injected charge corresponding to a MIP in a 220 μm sensor), with the hysteresis setting on the chip enabled (set to maximum).

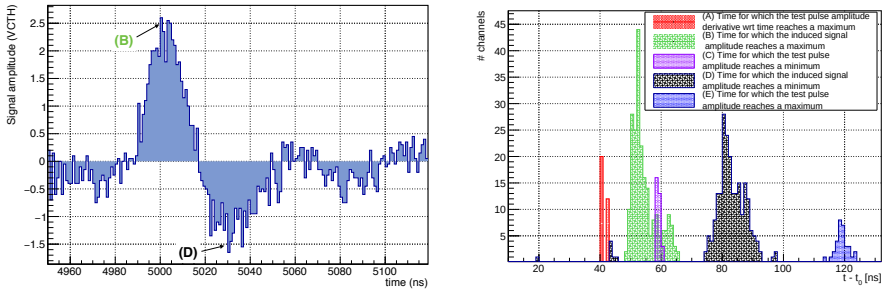


(a) (b)

FIGURE 7.5: LEFT: Example of induced signals on the channels not receiving the test pulse directly (for an injected charge corresponding to 4 times the charge of a MIP in a 220 μm sensor) with hysteresis setting disabled. RIGHT: Example of reconstructed test pulse with the hysteresis setting disabled for an injected charge corresponding to 4 times the charge of a MIP in a 220 μm sensor.

(A)	Time at which the derivative of the test pulse amplitude with respect to time reaches the maximum.
(B)	Time at which the induced signal amplitude reaches a maximum.
(C)	Time at which the test pulse amplitude reaches a minimum.
(D)	Time at which the induced signal amplitude reaches a minimum.
(E)	Time at which the test pulse amplitude reaches a maximum.

TABLE 7.2: Definitions of the moments in time used in Figure 7.6b.

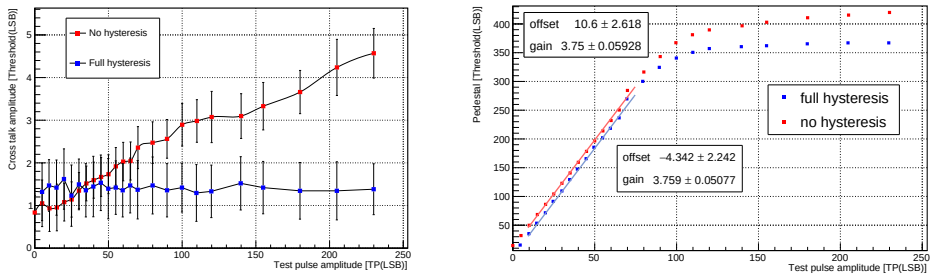


(a) (b)

FIGURE 7.6: LEFT: Example of induced signals on a channel not receiving the test pulse directly (for an injected charge corresponding to 4 times the charge of a MIP in a 220 μm sensor) with hysteresis disabled. RIGHT: Timing of test pulse signals and of the induced signals.

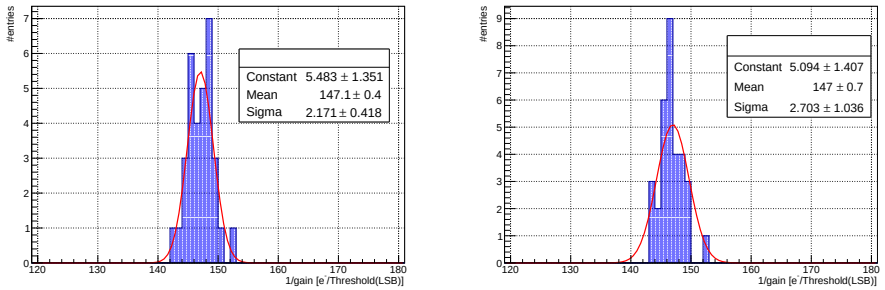
indeed correlated with the time of test pulse injection. The figure shows some other specific moments in time for the different channels. The definitions of these specific moments in time are illustrated in Figure 7.5b and 7.6a and summarised in Table 7.2. The red distribution in Figure 7.6b shows the distribution of the times where the reconstructed test pulse has the largest negative slope. This is followed by the green distribution which shows the time when the induced signal is peaking. The purple distribution gives the time distribution for which the test pulse reaches the minimum. The black and the blue distributions give the times of the overshoot of respectively the induced signal and the injected signal.

Figure 7.7a shows the strength of the induced signal for different amplitudes of the test pulse. It shows that the hysteresis setting masks the effect of induced signal and comparison with Figure 7.7b, which shows the actual measured amplitude of the test pulse, learns that the induced signal is of the order of 1% of the injected charge. This is a rather small effect and it can be concluded that this cross-talk behaviour can be tolerated. In the tracker collaboration a set-up was developed to inject external signals in the front-ends by using photodiodes connected to the wire bond pads on the hybrid [81]. The photodiodes are illuminated with light pulses from LEDs driven by a pattern generator on FPGA. In this way, up to 48 CBC channels can be stimulated at the same time, but more interestingly for cross-talk studies is the fact that single channels can be pulsed. The latter is not possible when using the internal test pulse mechanism. With this set-up also cross-talk of the order of 1% is measured on neighbouring channels.



(a) (b)

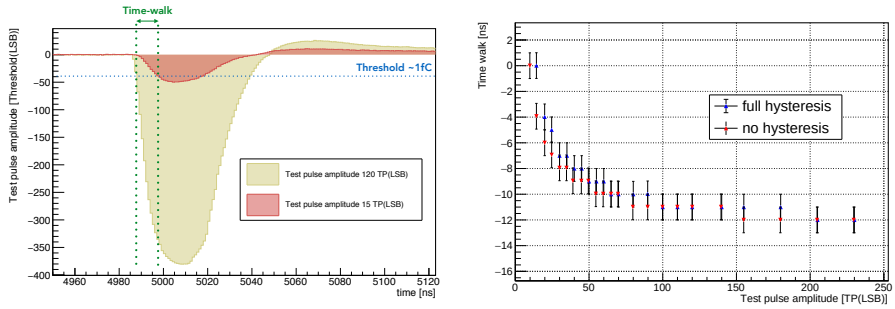
FIGURE 7.7: LEFT: Cross-talk averaged over all channels which do not receive a test pulse directly. The cross-talk is shown as a function of the test pulse amplitude and for two hysteresis settings of the chip. The error bars reflect the spread in the measured cross-talk amplitude amongst all channels not receiving the test pulse directly. RIGHT: Gain for a single channel for two front-end hysteresis settings.



(a) (b)

FIGURE 7.8: LEFT: Distribution of the inverse of the gain with the hysteresis enabled. RIGHT: Distribution of the inverse of the gain with the hysteresis disabled.

An example of a gain measurement for a single channel, extracted using the internal test pulse mechanism, is shown in Figure 7.7b both for a hysteresis off and on setting. The extracted gain value is equivalent for both hysteresis settings as expected. From test pulse amplitude $\approx$ 80 LSB onwards the response of the front-end is not linear any more. If a line is fitted to the data points corresponding to a test pulse with an amplitude smaller than 80 LSB, then we can extract the gain of the CBC3 front-end. The distribution of the inverse of the gain for all 32 channels in the test group are shown in Figure 7.8a and Figure 7.8b with respectively the hysteresis setting enabled and disabled. In both cases the spread on the extracted gain values can be attributed to the spread on the actual gain, but also on the spread of the injected charge, which essentially is linked to the spread in capacitance of the test pulse capacitors. The spread on the latter is a result of the manufacturing process and the spread is therefore expected to be the largest between chips. The gain measurement was also performed for a single test group in 3 other chips. The resulting values are summarised in Table 7.3.



(a) (b)

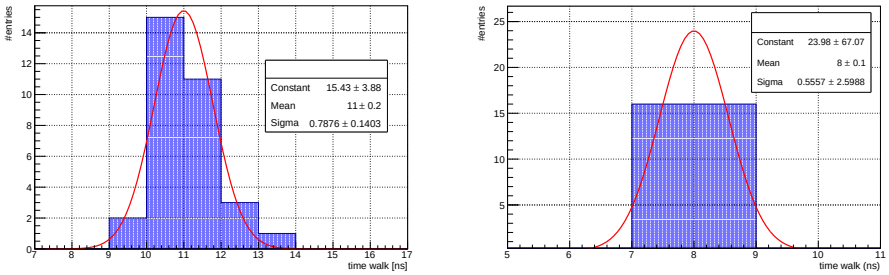
FIGURE 7.9: LEFT: Reconstructed test pulses for injected charges of 1.25 and 10 fC as reference for the time walk measurement. RIGHT: Comparator firing time for different settings of test pulse amplitude with the hysteresis enabled and disabled. The reference time is chosen to be the comparator (at a 1 fC threshold) firing time when a 1.25 fC signal in the full hysteresis setting is applied.

The manufacturer quotes a maximum spread of 10% (1σ) on the capacitor values, which limits the use of the internal test pulse mechanism to extract an accurate measurement of the absolute gain. In order to do a more accurate measurement an external, well calibrated signal could be injected in the front-end.

Hybrid	CBC	Mean inverse gain no hysteresis (e^-/Threshold(LSB))	Spread inverse gain no hysteresis (e^-/Threshold(LSB))	Mean inverse gain full hysteresis (e^-/Threshold(LSB))	Spread inverse gain full hysteresis (e^-/Threshold(LSB))
1	0	146.7±0.7	2.70±0.77	149.0±0.4	1.82±0.40
	1	147.0±0.7	2.70±1.04	147.1±0.4	2.17±0.42
2	0	154.2±0.2	3.19±0.78	151.9±0.7	2.91±0.77
	1	152.4±0.7	2.69±0.71	152.4±0.4	2.07±0.36

TABLE 7.3: Gain measurements performed on four CBC3 chips from two 2CBC3 hybrids. The gain measurement was performed for hysteresis on and hysteresis off setting.

From the reconstructed test pulse also the time walk can be extracted. The time walk is defined here as the difference in firing time of the comparator for an injected signal of 1.25 fC ($\approx$ 15 test pulse LSB) and 10 fC ($\approx$ 120 test pulse LSB) when the comparator threshold is set at 1 fC ($\approx$ 42 Threshold LSB). An example of reconstructed test pulses for these two values of injected charge are shown in Figure 7.9a where also the definition of time walk is illustrated. The time walk values for both settings of hysteresis are shown in Figure 7.10a and 7.10b. For both settings the measured time walk is well within the specified maximum of 16 ns [68]. To have a more global view on the time walk the comparator firing time is shown in Figure 7.9b for a range of test pulse amplitudes and for both hysteresis on and off setting. The reference time in this figure is chosen to be the comparator firing time when a 1.25 fC signal in the full hysteresis setting is applied with a 1 fC threshold setting. The figure shows that, the higher the amplitude, the faster the signals goes over threshold. Also, as expected, disabling the hysteresis makes the signal go over threshold faster.



(a) (b)

FIGURE 7.10: LEFT: Time walk for all the channels in the same test group with full hysteresis (LEFT) and no hysteresis (RIGHT).

7.4 Summary

The μDTC firmware blocks developed for readout and control of CBC chips can be used to test several devices based on these ASICs. The configurability and modularity of the μDTC is very convenient at this point and the readout of CBC based objects was actually a first milestone in the development of the μDTC. Bench-top testing of the CBC3 chip, using the on-chip test pulse injection mechanism, revealed a 1% cross-talk between the channels. The μDTC also allows for testing of module prototypes in test beams where e.g. the stub reconstruction logic can be exercised.

7.5 Author's contribution

The CBC2 and CBC3 based devices were the first devices to be read out and controlled by the μDTC. The testing with real devices allowed for tuning and improvement of several blocks of the firmware. The author contributed to having the μDTC configurable and operational to read out the multitude of CBC based devices which are shown in Figure 7.1. This required an intense collaboration with the colleagues developing specific tests and the system was used in real-life conditions during several test beams in which the author also participated and of which an important result is shown in Figure 7.2. The bench-top measurements, which are discussed in this chapter, were carried out by the author.

Chapter 8

PS Prototype Testing

8.1 Introduction

The qualification of the full PS module design requires the validation of each single component, starting with single chips. The specificities included in the μDTC for testing PS module components will be described in section 8.2. These developments were then used to do bench-top testing of the first MPA and SSA chips. This testing and an SEU test performed on the MPA and SSA will be the topic of section 8.3 and 8.4 respectively. Furthermore, a prototype MaPSA was tested in beam. The prototype, referred to as single-MaPSA, consists of a single MPA chip bonded to a baby PS-p sensor. This test beam will be discussed in section 8.5.

8.2 Specific firmware development for PS based systems

8.2.1 Introduction

In contrast to the CBC3 chips, where a VMEbus based readout was used by the chip developers to qualify the first chips, the μDTC system was used to do the qualification of the very first MPA and SSA chips. Therefore a more conservative approach was used in the development of the PS-μDTC firmware in order to minimise the chance of introducing bugs in the test system. A *debug* version of the firmware was developed and is briefly described in the next section. Section 8.2.3 describes the more DAQ oriented PS-μDTC firmware.

8.2.2 MPA/SSA debug firmware and PS specific test features

A *debug* version of the μDTC firmware was developed which scopes the data lines coming from the MPA and SSA a given time after a fast command was sent to the chips. The data on the lines are stored in a FIFO and can be read out over software. 5 μs of full-event data is stored after an L1A signal is sent to the chip and up to 500 μs of stub data is stored after a test pulse trigger is sent. The time window of 500 μs allows to also capture the output of the counters in the asynchronous mode.

In this debug mode, finding and parsing the actual data package needs to be performed in software, which implies a lot of overhead, but also minimises the complexity of the test system. The only intelligence present in the firmware to capture the data are the functionalities listed below, which were already tested in the CBC3 case:

- selecting the correct phase of the incoming data with respect to the FPGA's internal clock
- deserialising the 320 Mbps lines to 8$\times$40 Mbps.

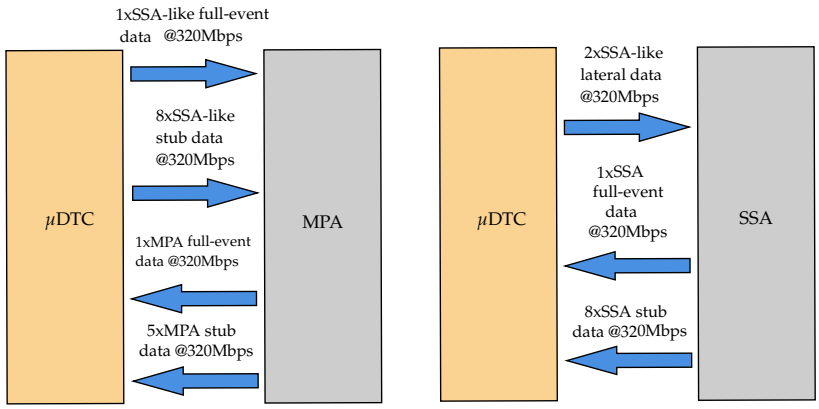


FIGURE 8.1: The data streams for stand-alone testing of the MPA (LEFT) and SSA (RIGHT). Only the full-event and stub data is shown, the fast command, slow control and clocking is not shown in this figure.

As the MPA and SSA chips are designed to operate together, but will at first be tested as stand-alone objects, the input data to the chips needs to be provided by the μDTC. Functionality is implemented on the μDTC to send SSA-like full-event and stub data to the MPA and to send lateral data to the SSA. The data streams are illustrated in Figure 8.1. The sending of this emulated data needs to be synchronised in time and therefore the transmission of this emulated data is triggered by certain fast commands. The content of the data can be configured over software. In order to make sure that the emulated data arrives at the chip at the correct time the sending of the data can be delayed in steps of 25 ns after the reception of the fast command which triggers the output of the data. Also the phase of this data, with respect to the clock sent to the chips, can be configured over software. This phase-shift was implemented in the firmware by clocking the OSERDES (see section 6.2) with a phase-shifted 320 MHz clock generated by a separate MMCM. The phase of this clock is configurable at runtime with a resolution of ≈ 18 ps.

The same I^2C master is used as in the case of the CBC testing, with some minor modifications such as 16-bit instead of 8-bit register addresses. The logic for sending the fast commands is the same as in the CBC3 case.

As the test system had to be fully operational when the chips became available, it needed to be tested without actual DUT. Therefore a basic emulation of an MPA chip was coded to run on FPGA. The firmware block responsible for the MPA emulation resides in the same μDTC firmware project. The output lines of the emulator block, which are the full-event and stub data lines, can be connected internally to the Phy layer to test the firmware blocks, for receiving MPA data, in simulation.

The MPA emulation block can also be compiled stand-alone. In that case a set-up with 2 FC7s can be used: one FC7 running the DAQ and the other FC7 running the emulator. The data sent back and forth between the two FC7s has the same format as in the case of a real chip. The emulated chip receives the fast commands and the 320 MHz clock from which it generates a 40 MHz clock for its internal logic. It acts on the fast commands by sending MPA-like stub and full-event data. Also an I^2C slave is present and is used to test the I^2C communication and to let the user change the content of the emulated stub and full-event data. This set-up allowed the testing of the new firmware blocks and the software provided to run the tests. An emulation of the SSA chip was not prepared as all

the relevant features of the firmware could be tested with the MPA emulation.

Operating the test system using the debug blocks to acquire the data introduces a large overhead in data transmission to the back-end. This slows down the data taking and needs to be optimised by building events into a data package similar to the one illustrated in Figure 6.8, rather than reading out the raw data. A *DAQ firmware*, similar to the CBC3 firmware, was implemented and is discussed in the following section.

8.2.3 MPA/SSA DAQ firmware

Adding the MPA and SSA chips to the μDTC firmware exemplifies the usefulness of the Physical Interface Abstraction Layer. The parsing of the data formats is performed in this layer.

The MPA full-event data, illustrated in Figure 5.17, is of non-fixed length and is sparsified. The maximum number of S-clusters and P-clusters which can be sent is 24 and 31 respectively. The manual of the MPA specifies that after the transmission of a full-event data package there are at least three 40 MHz clock cycles where the full-event data is 0. As the data is sparsified and of non-fixed length, the state machine processing the full-event data in the Physical Interface Abstraction Layer layer has to interpret the header information which specifies how many strip and pixel clusters are to follow in the payload and tells how to split the payload data into S-clusters and P-clusters.

The MPA stub data also requires some special attention compared to the CBC3 and SSA stub data. For the MPA, as can be seen in Figure 5.16b, the stub data is spread over 2 BXs. In order to attain a fixed stub latency the Phy layer foresees an 80-bit wide bus to the Data Readout Block: one 25 ns clock cycle carries the data from BX n, and the subsequent 25 ns clock cycle carries data from BX n+1. This is the boxcar unpacking of the MPA stub data.

The time it takes the full-event data package to travel through the Phy layer is a function of the length of the data package, so it does not arrive a fixed time after the trigger signal in the Data Readout Block. On the other hand, there is the stub data package which is boxcar unpacked in the Phy layer and from then on is synchronous again. Like this the stub data can still be delayed in a similar way as was done in the CBC3 case and only the stub data with a certain time offset (stub latency) with respect to the L1A trigger is written to the Data Readout Block. In this way the event building for the MPA case is guaranteed.

The MPA emulator, described in section 8.2.2 was also used to validate the above event building. In order to test this, a self triggering on stubs was implemented. The test routine is as follows:

1. Send a test pulse trigger.
2. The emulator will respond on the test pulse trigger by outputting non-zero stub data.
3. The firmware notices that there is non-zero stub data and when configured correctly will generate a trigger.
4. The emulator responds to this trigger with full-event data.
5. One can now check that for different stub data patterns (only stub data in the even BX, only stub data in the odd BX, stub data in both even and odd BX (this will result in 2 consecutive triggers)) the stub latency is indeed fixed and that it is possible to set the L1 and stub latency to one fixed value in order to read the data.

8.2.4 Use cases

The following list gives an overview for which devices and tests the PS-μDTC firmware was used and the pictures in Figure 8.2 show the hardware:

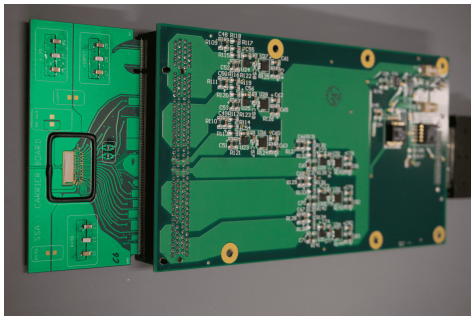
- MPA stand-alone: initial verification (see section 8.3.1), wafer probing and TID and SEU testing. The latter required specific firmware development (see section 8.4.3) and this firmware was also used to characterise the power consumption of the MPA at high trigger rates.
- SSA stand-alone: initial verification (see section 8.3.2), wafer probing and TID and SEU testing. As for the MPA also special firmware was designed (see section 8.4.3) to run the SSA SEU test and this firmware was also used to characterise the power consumption of the SSA at high trigger rates.
- Single-MaPSA assembly: tested in beam (see section 8.5).
- 2SSA mini-module: the first operation of SSA chips with a sensor.
- MPA-SSA board: first implementation of an MPA and SSA chip on a common carrier. After stand-alone testing of the single MPA and SSA chip this board was developed as proof that the MPA and SSA chip can talk to each other. From the firmware point of view the system is similar to reading out a stand-alone MPA chip.

The PS-μDTC firmware will also be used to test PS hybrids. A dedicated test system is being developed to test PS front-end hybrids as stand-alone objects. This test is required before assembly of the hybrid into a module in order to guarantee the functionality of the hybrid and thus the full functionality of an assembled module. The test system is based on plug-in cards which carry a single PS hybrid. These plug-in cards are operated in a crate system, housing several plug-in cards[1], and the μDTC is connected to the backplane of the crate. Each plug-in card makes the I/Os of the SSAs available by temporarily connecting, with spring-loaded sockets, to the wire-bond pads. Also the I/Os to the CIC are made available. In the end the system is limited to simultaneous access to only two SSAs due to the limited number of lines available as input to the FC7. A multiplexing scheme on the plug-in card is used to select 2 SSA chips on a specific hybrid. As a result it appears to the firmware that 2 SSA chips are connected and the same firmware version (modulo a new pin assignment) can be used as for the 2SSA mini-module. The CIC chip on the PS-hybrid can be tested with a similar test firmware as the one described in section 9.2.

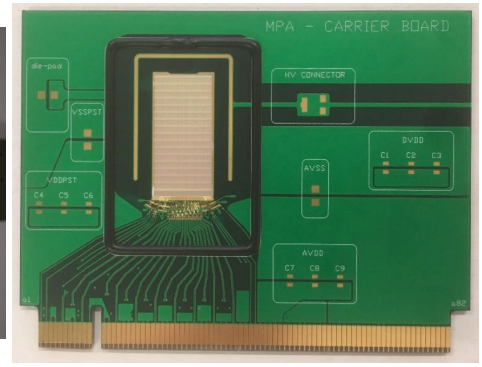
8.3 Single chip testing

The debug firmware allows for testing of all the digital features on the MPA and SSA chips. Also the functionality of the front-end can be assessed by taking S-curves both with and without the test pulse injection. The below sections list the functionalities which were tested. Most of these tests were performed short after the first prototypes of the MPA and SSA became available in January 2018.

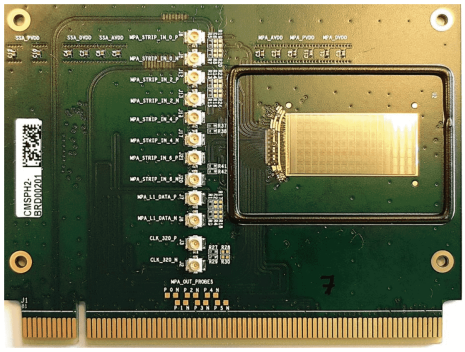
[1] Selection of the plug-in card, and thus hybrid, under test is performed using a multiplexing system on the backplane.



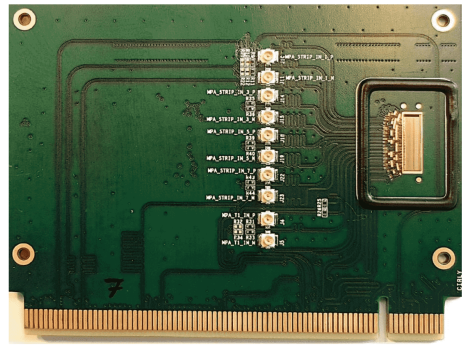
(a)



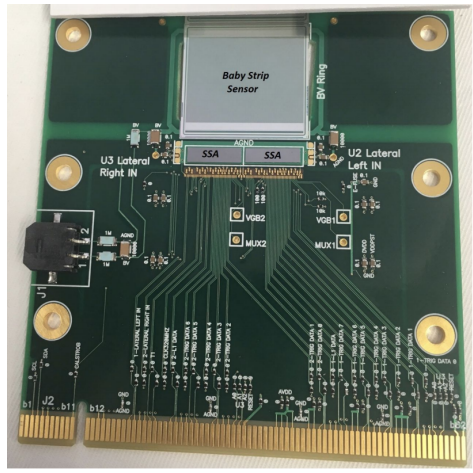
(b)



(c)



(d)



(e)

FIGURE 8.2: TOP LEFT: SSA on carrier card connected to the interface card. TOP RIGHT: MPA on carrier card. MIDDLE LEFT: MPA side of the MPA+SSA intercommunication card. MIDDLE RIGHT: SSA side of the MPA+SSA intercommunication card. BOTTOM: 2SSA mini-module.

8.3.1 MPA

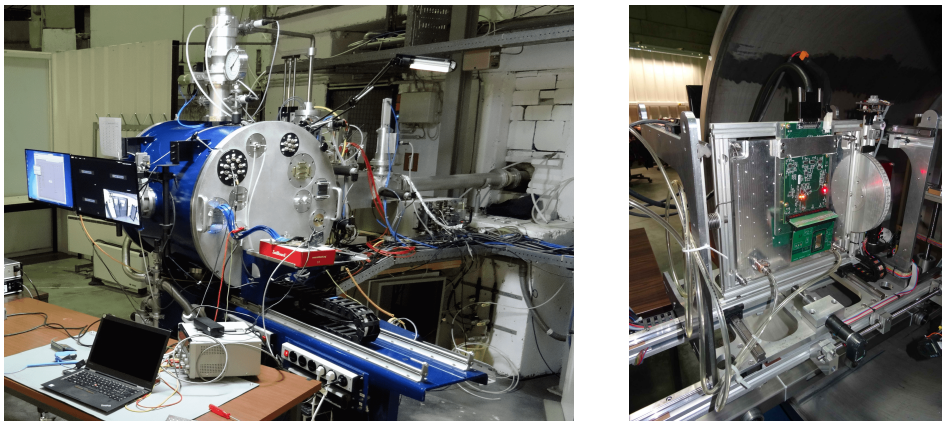
The following functionalities were tested and validated to be working on the MPA chip:

- I^2C communication
- Alignment of the output lines
- Pixel masking
- Stub logic: correlation window, cluster cut
- Strip-strip mode: the MPA sends out the strip centroids on the stub data lines
- Pixel-pixel mode: the MPA sends out the pixel centroids on the stub data lines
- Hit detect logic
- Asynchronous mode
- Digital test pulse injection
- Analogue test pulse injection
- Digital pattern generation
- Full-event frame and full-event memory
- Verification of the analogue front-end (e.g. gain measurements using the test pulse injection)

8.3.2 SSA

The following functionalities were tested and validated to be working on the SSA chip:

- I^2C communication
- Alignment of the output lines
- Strip masking
- Stub logic: cluster cut
- Hit detect logic
- Asynchronous mode
- Lateral communication
- HIP flags
- Digital test pulse injection
- Analogue test pulse injection
- Digital pattern generation
- Full-event frame and full-event memory
- Verification of the analogue front-end (e.g. gain measurements using the test pulse injection)



(a) (b)

FIGURE 8.3: LEFT: The irradiation chamber at the UCL HIF facility. RIGHT: MPA and interface card mounted on a cooling plate and rotation stage inside the HIF irradiation chamber.

8.4 MPA and SSA single-event upset testing

8.4.1 Introduction

In section 5.5.4, the SEU tolerant design strategy of the MPA and SSA chip was introduced. The following sections describe how the SEU sensitivity of the MPA and SSA chip was tested in order to determine the SEU related upset rate in the HL-LHC environment. The tests described here were performed in October 2018.

8.4.2 Heavy ion facility

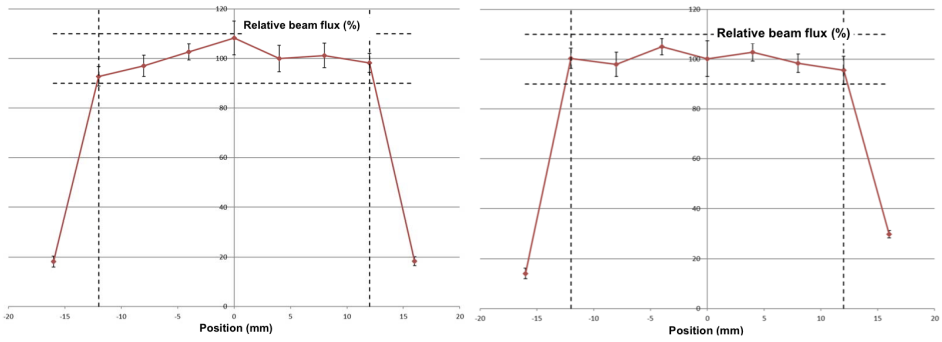
For the MPA and SSA SEU testing, the UCL (Université Catholique de Louvain) CRC (Cyclotron Resource Centre) HIF (Heavy Ion Facility), shown in Figure 8.3a, was used. At the HIF, heavy ions are accelerated by a cyclotron and are delivered into a vacuum chamber where the chip can be placed on a rotation stage. The available ions and their properties are listed in Table 8.1.

The beam is uniform over a size of $\approx 2.4 \times 2.4$ cm^2 and has a maximum fluence of 15k particles /cm^2/s. The beam energy and uniformity is calibrated before the irradiation. Figure 8.4a and 8.4b show the uniformity of the beam in two orthogonal directions. An SEU rate measurement is also performed with a reference device and the rate was found to be uniform within 10% of the mean value over 2×2 cm^2.

Figure 8.3b shows the MPA chip and interface card mounted on the cooling plate, which is attached to the rotation stage, in the vacuum chamber. The VHDCI cable was connected to an FMC feed-through which on the other side connects to the FC7 using FMC cables. The set-up is almost the same for the SSA chip as the interface card can also host an SSA carrier card. As a result only the carrier card and the firmware need to be changed to test the SSA chip.

Ion	M/Q	Energy on device [MeV]	Range on device [μm]	LET on device [MeV/mg/cm^2]
$^{13}C^{4+}$	3.25	131	269.3	1.3
$^{22}Ne^{7+}$	3.14	238	202.0	3.3
$^{27}Al^{8+}$	3.37	250	131.2	5.7
$^{40}Ar^{12+}$	3.33	379	120.5	10.0
$^{53}Cr^{16+}$	3.31	513	107.6	16.0
$^{58}Ni^{18+}$	3.22	582	100.5	20.4
$^{84}Kr^{25+}$	3.35	769	94.2	32.4
$^{103}Rh^{31+}$	3.32	972	88.7	45.8
$^{124}Xe^{35+}$	3.45	995	73.1	62.5

TABLE 8.1: Properties of the ions (with mass M and charge Q) available at the HIF facility. The range is defined in silicon. Values as provided by the facility.



(a) (b)

FIGURE 8.4: Uniformity of the HIF beam in two orthogonal directions.

8.4.3 Test set-up

In order to catch every possible SEU induced error on the data stream and thus extract the best value for the SEU cross sections, ideally all full-event and stub data coming out of the chip has to be checked within the time slot available at the irradiation facility. This requires online checks of both the stub and full-event data. The stub data are produced by the chip every BX, whilst the L1A trigger can be sent up to 1 MHz. Checking the stub data every BX and the full-event package at a rate of 1 MHz has to happen on the FPGA as the IPbus bandwidth is too limited to send all this data to the back-end, where a comparison in software would be possible. Furthermore a large amount of storage would be needed. To give an example: storing the full data stream output by a single SSA chip triggered at 1 MHz for a single second would require $\approx$ 2.7 Gb of storage.

The MPA and SSA chips can be configured to output full-event and stub data frames with a fixed payload. This is achieved by using the digital test pulse injection mechanism which inputs digital patterns into the digital part of the chips each time a test pulse trigger is received (see section 5.5.2 and 5.5.3). In this way the full digital logic on the chip is used which is the SEU sensitive part of the ASIC.

Three state machines were added to the firmware, each dedicated for online checking of the incoming data:

- Full-event data check for the MPA and SSA

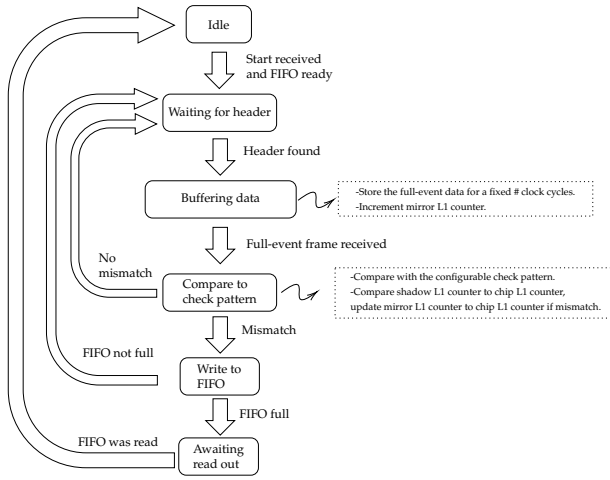


FIGURE 8.5: Diagram representing the dedicated state machine written for running the MPA and SSA SEU testing for checking the full-event data stream. During all the states it is checked if a stop was sent by the user.

- Stub data check for the MPA
- Stub data check for the SSA.

These state machines are illustrated in Figures 8.5 and 8.6.

The state machine for testing the full-event data stream during the SEU test, illustrated in Figure 8.5, looks for the header of the full-event frame. Once the header is found it starts to buffer the data on the full-event line for a fixed amount of clock cycles. For the MPA this was set to a time which matches the time needed to output a full-event package containing 8 strip clusters and 8 pixel clusters. Transmitting packages containing more strip or pixel clusters would not be possible at a rate of 1 MHz. In the SSA case the full-event frame has a fixed length. The payload from the buffered full-event packages is subsequently compared to a pattern which is configurable over IPbus and the L1 counter is compared to a mirror counter which increments on the FPGA every time a full-event header was found. In case of a mismatch between the check pattern and the payload or a mismatch between the chip L1 counter and the mirror L1 counter, the full-event package together with the μDTC's 32-bit rolling BX counter is written to a FIFO which can be read out after stopping the test. This BX counter (BX resolution over a range of ≈ 108 s) gives a time reference of when the SEU happened. This is used to see if any consecutive errors due to SEU could be observed. The test is stopped when the FIFO is almost full or when the test is stopped by the user. This last option would typically be used to start running the test with a new pattern.

Similar state machines, described in Figure 8.6a and 8.6b, were added to test the stub data streams. This data is delivered by the chips synchronously at 40 MHz and all data needs to be checked for occurrences of SEUs. A start, sent from software, starts the state machine which checks the incoming raw stub data against a configurable check pattern. It was chosen to send the test pulse trigger to the chip every second clock cycle. In this configuration, for the SSA chip, there will be a BX with data followed by a BX where the stub lines are all 0. In order to find the BX which carries the data the test pulse triggering is preceded by a fast reset of the chip. The state machine on the firmware waits for this fast

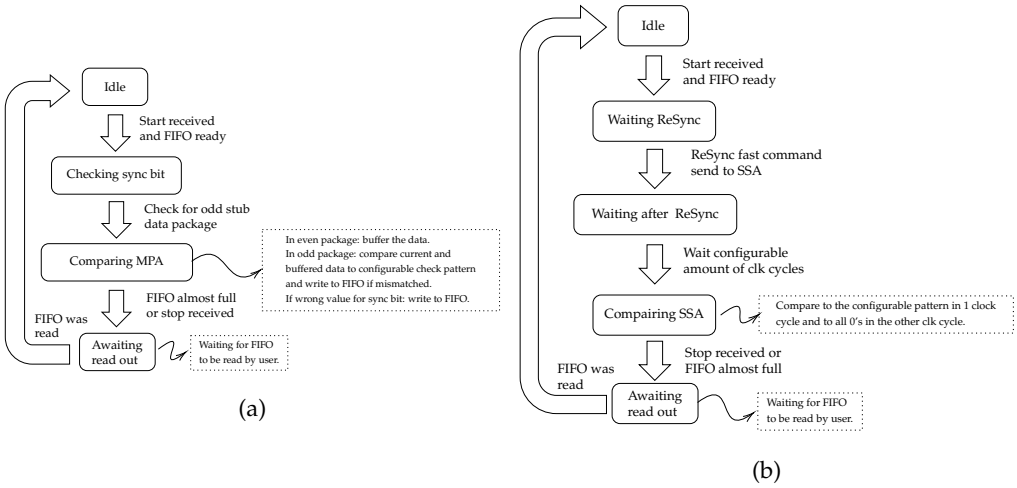


FIGURE 8.6: Diagrams representing the dedicated state machines written for running the MPA and SSA SEU testing for the MPA stub data (LEFT) and SSA stub data (RIGHT). During all the states it is checked if a stop was sent by the user.

reset being sent and then waits a configurable number of clock cycles before interpreting the data. By tuning this waiting time manually, it is possible to find the BX which carries data.

For the MPA this way of sending the test pulse trigger can still result in data on the stub lines in both odd and even BXs as the MPA sends stub data in packages spanning 2 BXs. The MPA has the advantage of the presence of the sync bit to synchronise the comparison to the check pattern. During the BX where the sync bit is high (referred to as the *even* BX) the data is buffered on the FPGA and during the next BX (referred to as the *odd* BX) the buffered data together with the current data on the stub lines is compared to the check pattern and the data together with the BX counter is written to a FIFO in case of mismatch. Furthermore, if the sync bit is not present an error is raised and the data is written to the FIFO.

In order to test the complete digital logic against SEUs also the data inputs to the ASICs have to be stimulated. This means that during the MPA testing, the MPA needs to receive SSA-like full-event and stub data from the FPGA. In a similar fashion, the SSA needs to receive lateral SSA-like input data. These input data streams have to be synchronised with the data generation on chip. This input data to the chip is therefore also triggered by the L1A trigger (MPA full-event data input) and the test pulse trigger (MPA stub data input and SSA lateral data input). The exact tuning can be done manually. As discussed before, it is possible to both delay the data in steps of 25 ns and to shift its phase. In this configuration it is possible to choose the arrival time of the input data at the chip carefully in order for it to result in data in the same data package as the digital test pulse injection. The L1 counter in the SSA full-event data package going to the MPA has to be synchronised with the L1 counter on the MPA as well. This is required because the MPA does a check for the alignment between these two counters. In case of mismatch, an error bit is raised in the full-event frame of the MPA.

In order to maximise the amount of data checked during the SEU test, the full-event

Ion	Angle 1 (0°)	Angle 2 (30°)	Effective LET angle 1 [MeV/mg/cm^2]	Effective LET angle 2 [MeV/mg/cm^2]
$^{13}C^{4+}$	0	30	1.3	1.5
$^{27}Al^{8+}$	0	30	5.7	6.6
$^{40}Ar^{12+}$	0	30	10	11.5
$^{58}Ni^{18+}$	0	30	20.4	23.6
$^{84}Kr^{25+}$	0	30	32.4	37.4
$^{103}Rh^{31+}$	0	30	45.8	52.9
$^{124}Xe^{35+}$	0	30	62.5	71.2

TABLE 8.2: SEU irradiation programme for the MPA chip.

and stub data routines were run in parallel. The test is performed in the following sequence:

1. Configure the sending of SSA-like full-event and stub data to the MPA or lateral data to the SSA: configure the correct timing and configure the payload in order to get a defined output on the data lines.

2. Configure the digital test pulse injection to create a specific pattern on the stub and full-event data lines.

3. Configure the check patterns for the full-event and stub data. These check patterns have to be encoded manually by the user.

4. Configure the L1A triggering and test pulse triggering state machine in the firmware in order to create a fixed L1A trigger with 1 MHz frequency and a fixed test pulse trigger with 20 MHz frequency.

5. Start the state machine which compares the full-event data to the check pattern. The state machine is waiting for the first header.

6. Start the sending of L1A triggers and test pulse triggers for an undefined time, preceded by a single ReSync. The state machine which does the full-event comparison becomes active as headers start to arrive.

7. Start the state machine which does the stub comparison. Stub data is already present, so the state machine will be able to find the correct BX.

This routine was run without beam for several hours. No errors were seen for several hours of dry running.

During the SEU testing, a certain input pattern was run for $\approx$ 30 seconds after which the I^2C settings of the chip were checked for upset configuration registers. In case of error in the I^2C settings these errors were logged and the configuration rewritten. After that the next pattern is configured. This test was done at two different L1A trigger latencies in order to check the effect of SEUs in the memory. These runs were done for different types of ions and at two different angles: 0° and 30°. The MPA and SSA chip were irradiated under the conditions shown in Table 8.2 and 8.3 respectively. Each irradiation with a specific ion and under a specific angle lasted for $\approx$ 20 minutes.

8.4.4 SEU cross section calculation

In order to come to a bit error rate due to SEU in the CMS HL-LHC environment the results of the irradiation with heavy ions needs to be ported to the HL-LHC radiation environment. In order to do this the computational method described in Ref. [16] is used. In

Ion	Angle 1 (0°)	Angle 2 (30°)	Effective LET angle 1 [MeV/mg/cm^2]	Effective LET angle 2 [MeV/mg/cm^2]
$^{13}C^{4+}$	0	-	1.3	-
$^{22}Ne^{7+}$	0	-	3.3	-
$^{27}Al^{8+}$	0	-	5.7	-
$^{40}Ar^{12+}$	0	30	10	11.5
$^{58}Ni^{18+}$	0	-	20.4	-
$^{84}Kr^{25+}$	0	30	32.4	37.4
$^{103}Rh^{31+}$	0	30	45.8	52.9
$^{124}Xe^{35+}$	0	30	62.5	71.2

TABLE 8.3: SEU irradiation programme for the SSA chip.

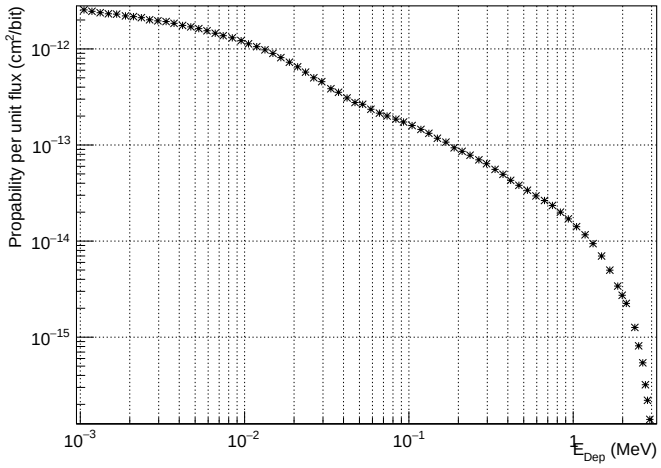


FIGURE 8.7: Probability to have within a volume of $1\times1\times1$ μm^3 of silicon an ionizing deposition greater or equal to the indicated E_{Dep}. The incoming particles which give rise to these energy depositions are representative for the charged hadron spectrum with E > 20 MeV for proton-proton collision at 13 TeV in the Outer Tracker environment [16].

Ref. [16] it is shown that a Weibull fit to experimental heavy-ion SEU data can be used to quantify the sensitivity of the circuit for SEUs. Probability distributions representing the probability to have an energy deposition within a certain volume triggered by a particle of a certain type and energy are given in Ref. [16]. These probabilities can be calculated for different particles and a given energy spectrum as shown in Figure 8.7. Here the example is given for hadrons with E > 20 MeV emanating from proton-proton collisions at 13 TeV for the CMS Outer Tracker environment specifically.

Only E > 20 MeV hadrons are considered, because Ref. [16] concludes that other types of particles and lower energy neutrons have an SEU cross section which is small compared to hadrons with E > 20 MeV and furthermore their flux is at least an order of magnitude less in the Outer Tracker environment. These two factors combine to a negligible contribution to SEU for particles other than E > 20 MeV hadrons.

8.4.5 Results

In the next two paragraphs the results of the SEU testing will be discussed. The results are divided in two sections: particular upsets seen in the configuration and control of the ASICs in section 8.4.5.1 and the calculation of the SEU cross sections based on the bit upsets seen in the data in section 8.4.5.2.

8.4.5.1 SEU triggered configuration and control upsets

For the MPA chip only two configuration errors, not corrected by the triplication logic, were observed during the testing. To see an upset in the MPA configuration, a multi-bit upset is required where two out of the three registers get upset and as a result the wrong value is forced also on the third one. These multi-bit single-event upsets were traced back to happen on a register where the spacing of the triplicated registers was smaller than 15 μm. Normally the design tool ensures that the spacing is larger than 15 μm to decrease the SEU sensitivity, however successive place and route steps sometimes move the registers closer again. This will be fixed for the production of the next version of the MPA chip. Single bit upsets are counted in the *SEU counter* which counts the number of corrections that need to be applied to the chip's configuration. For design reasons not all the MPA registers are connected to the SEU counter. For the MPA 3×1181 flip-flops and for the SSA 3×256 flip-flops[2] from the configuration logic are connected. The actual values for the SEU counter are given in Appendix A where it can be seen that, depending on the type of ion which is used, the SEU counter rate can go up to ≈ 2 Hz.

For the SSA, in 3 out of 128 runs the strip configuration registers got reset and in one case also the periphery configuration was reset. This was traced back to a problem in the triplication of the reset tree. In the next version of the SSA a constraint will be implemented in the design to force triplication of the reset tree and also the physical reset pad will be triplicated. Dedicated SEU tests of the corrected versions of the MPA and SSA will be performed to extract a final upset rate for the configuration registers.

8.4.5.2 SEU cross section and bit upset rates on data paths

By counting the number of SEU (# SEU) in a given time slot of length t, the SEU cross section σ can be calculated as:

$$\sigma = \frac{\#\text{SEU}}{\phi.t}, \tag{8.1}$$

where ϕ is the heavy ion flux.

Figure 8.8 and 8.9 show the SEU cross section for the different kinds of logic at different LETs for the MPA and SSA chips respectively. The data are fitted with a Weibull distribution:

$$\sigma = \sigma_0 \left\{ 1 - \exp\left[-\left(\frac{E_{Dep} - E_0}{W} \right)^s \right] \right\}, \tag{8.2}$$

where σ_0 is the saturation value of the SEU cross section and E_0 is the SEU threshold energy. W and s are shape parameters. To calculate the deposited energy from the LET, the following formula is applied:

$$E_{Dep} = LET \times d \times \rho_{Si} = LET \times 0.2328 \, \frac{\text{mg}}{\text{cm}^2}, \tag{8.3}$$

[2]The factor 3 is for the triplication.

where ρ_{Si} (= 2.32 g/cm^3) is the density of silicon and d is the sensitive depth which was here assumed to be 1 μm.

Figure 8.8 and 8.9 show for both chips the SEU cross sections for the stub data path and the full-event data path for two latency settings. For the MPA (SSA) low latency corresponds to a latency setting of 40 BX (100 BX) and high latency corresponds to a latency setting of 500 BX (500 BX). Due to the low statistics for the SSA data at low latency, a fit with a Weibull distribution is hard and therefore not shown. It can be seen however that the data points for the low L1 latency of the SSA measurements are all below the data for high latency. The high latency value for the error rate can thus be used as an upper limit for the SSA low L1 latency setting. Anyway, the high latency setting is a more representative value for the application in the CMS experiment. Also for the MPA chip the dependency on the latency setting is clear: when the chip is configured for a higher latency the data resides longer in the memory and therefore the probability for upset data being present at the output is larger for the high latency setting. Important to note for the full-event data is that the L1A rate was 1 MHz. The SEU cross section which is calculated below, and as a result the upset rates, are a function of the L1A trigger rate. 1 MHz is the upper limit for the L1A trigger rate and as such the cross section and upset rates extracted here are the worst expected. In order to evaluate the bit upset rate, due to SEU, coming off a module, this SEU rate should be downscaled to a L1A rate which is sustainable by the CIC chip. This will be done later in this section.

Comparing the SEU cross sections of the MPA and SSA chip, it can be concluded that the MPA is more SEU sensitive than the SSA chip. This is a direct consequence of the higher number of logic cells on the MPA.

In Figure 8.8 and 8.9, the uncertainty on the cross section values is attributed to the Poisson uncertainty on the counts, the uncertainty on the fluence value (taken to be 100 particles/s/cm^2) and the uncertainty on the duration of the test (taken to be 1 s).

In order to extract the SEU cross sections for the CMS Outer Tracker environment at 13 TeV proton-proton collisions the Weibull fits need to be convoluted with the probabilities in Figure 8.7:

$$\Gamma = \sum_{E_i} \frac{P_i \Delta\sigma_i \sigma_0}{A} = \sum_{E_i} \frac{P_i(\sigma_{i+1} - \sigma_i)}{A}. \tag{8.4}$$

Here Γ is the SEU cross section under specific conditions, P_i are the probabilities represented in Figure 8.7, A is the sensitive area which is assumed here to be 1×1 μm^2, according to the recommendations from Ref. [16], and the sum runs over the energy range in Figure 8.7.

The results of these calculations are shown in Table 8.4. The relative uncertainties (originating from the uncertainty on the fit) on these values are typically of the order of 10%. All these values should however be interpreted as an order of magnitude calculation due to the approximations used in the computational method. For both chips and for the different logic the SEU cross section can be estimated to be around 10^{-10} cm^2.

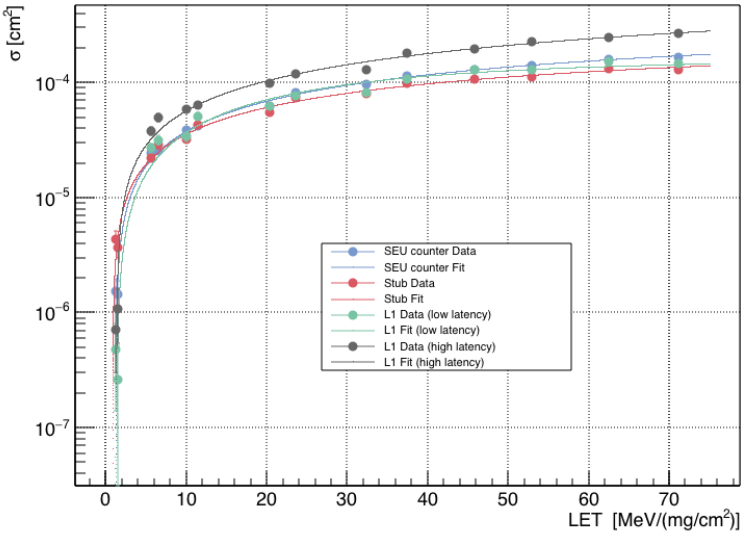


FIGURE 8.8: SEU cross sections at different LET for the MPA chip

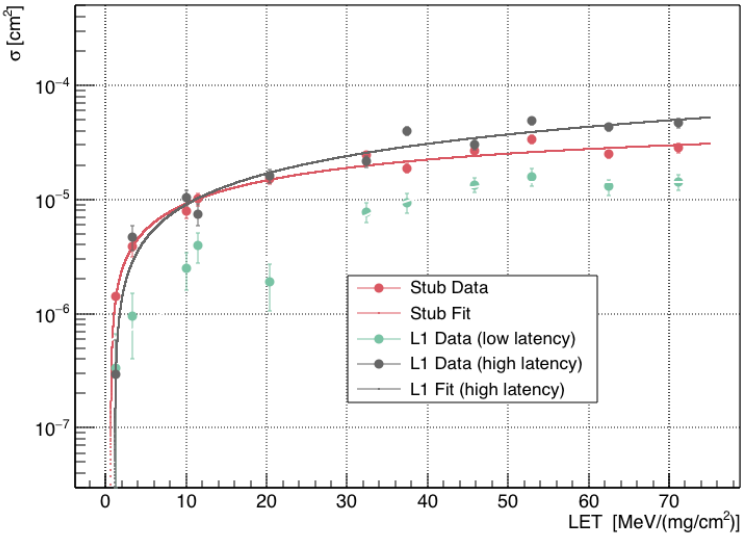


FIGURE 8.9: SEU cross sections at different LET for the SSA chip

Chip	SEU counter (cm^2)	Stub data (cm^2)	Full-event data, low latency (cm^2)	Full-event data, high latency (cm^2)
MPA	6.77×10^{-11}	7.92×10^{-11}	5.59×10^{-11}	9.55×10^{-11}
SSA	-	2.74×10^{-11}	1.39×10^{-11}	1.39×10^{-11}

TABLE 8.4: SEU cross sections for the different types of logic in the MPA and SSA chip, extrapolated to the CMS Outer Tracker environment at 13 TeV pp collisions. The cross section on the full-event data is dependent on the L1A trigger rate and is shown for a rate of 1 MHz, which is the rate at which the SEU test was performed. The SEU counter data for the SSA chip is not usable and therefore not shown. The value for the high latency setting is used as an upper limit for the low latency setting in the SSA case.

The bit upset rates are calculated by multiplying these SEU cross sections with the flux of hadrons with E > 20 MeV in the HL-LHC conditions. The hadron fluxes are extracted from the CMS FLUKA simulation v3.7.2.0 [82]. The flux for hadrons with E > 20 MeV at an HL-LHC instantaneous luminosity of 75000 $\mu b^{-1}s^{-1}$ (= 7.5×10^{34} $cm^{-2}s^{-1}$) is shown in Figure 8.10 for the region where PS modules will be located. This instantaneous luminosity value is the estimated peak luminosity of the HL-LHC.

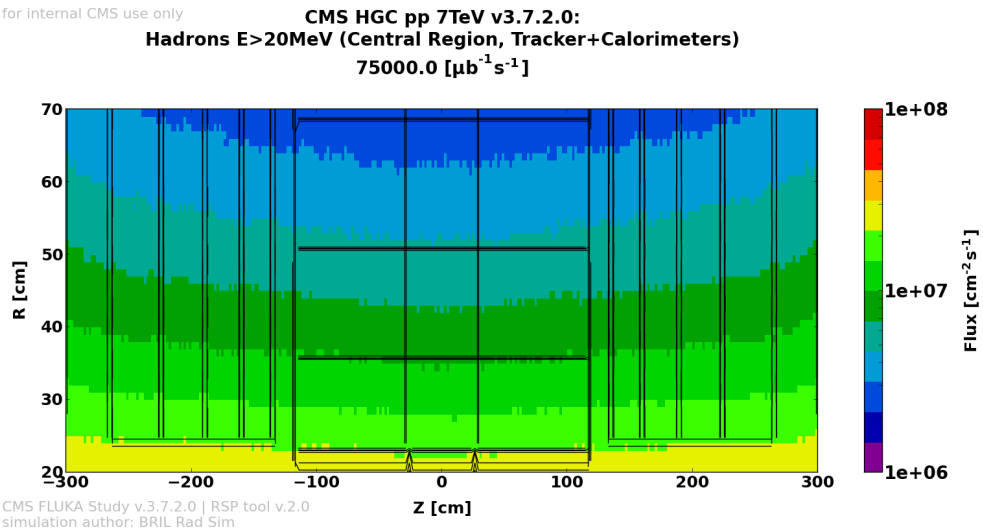


FIGURE 8.10: Flux of E > 20 MeV hadrons in the region of the CMS detector where PS modules will be placed. This is extracted for an HL-LHC instantaneous luminosity of 75000 $\mu b^{-1}s^{-1}$ [83].

Figure 8.11 shows the distribution of the fluxes which the PS modules will receive in the conditions mentioned above. Outer Tracker tkLayout[3] version 614 was used to extract the PS module locations. The median value of 6.4×10^6 $cm^{-2}s^{-1}$ is taken here as a representative value for all the PS modules. Using an uncertainty of 20%, quoted on the FLUKA simulations, we can use the flux value of 7.7×10^6 $cm^{-2}s^{-1}$ for a conservative estimate of the upset rates. The results of this calculation are shown in Table 8.5. These are the estimated upset rates due to SEU for stand-alone MPA and stand-alone SSA chips.

[3]tkLayout is the 3D modelling and performance estimation software tool used to optimise the layout of the CMS Phase-2 tracker [84].

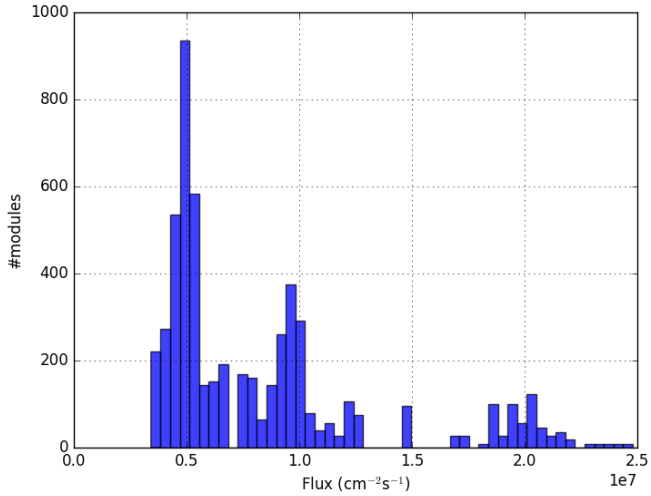


FIGURE 8.11: Flux of hadrons with E > 20 MeV for the PS modules at an instantaneous luminosity of 75000 $\mu b^{-1}s^{-1}$.

Chip	SEU counter ($s^{-1}chip^{-1}$)	Stub data ($s^{-1}chip^{-1}$)	Full-event data, low latency ($s^{-1}chip^{-1}$)	Full-event data, high latency ($s^{-1}chip^{-1}$)
MPA	4.34×10^{-4}	5.07×10^{-4}	3.58×10^{-4}	6.11×10^{-4}
SSA	/	1.75×10^{-4}	8.88×10^{-5}	8.88×10^{-5}

TABLE 8.5: SEU upset rates for the different types of logic in the MPA and SSA chip extrapolated to the CMS Outer Tracker environment at 13 TeV pp collisions at an instantaneous luminosity of 75000 $\mu b^{-1}s^{-1}$ for a representative PS module. The SEU rate on the full-event data is dependent on the L1A trigger rate. The value used for these upset rates is a L1A rate of 1 MHz which is the same frequency at which the test was performed. The value for the high latency setting is used as an upper limit for the low latency setting in the SSA case.

In reality these chips will not be used as stand-alone objects. The SSA data will be passed through the MPA and the data from multiple MPA chips is concentrated by the CIC chip. Therefore these single chip error rates need to be translated to the error rate for a module. To make an estimation of the module SEU rate the following is assumed:

- The same particle does not generate SEUs in both the MPA and SSA chip.
- The CIC is insensitive to SEUs.
- The MPA treats upset data from the SSA as non-upset data. Similarly the CIC treats upset data from the MPA as non-upset data.
- A strip/pixel hit occupancy and L1A rate such that the CIC bandwidth is at the saturation threshold.

The first point can be assumed because the MPAs and the SSAs do not physically overlap in a module and the chance of 2 SEUs arising from the same particle is very small. The

CIC SEU sensitivity is not known and therefore not used in this calculation. The last point is considered such that the SEU rate is actually maximal, with the CIC in a state where it is just able to pass all the stub and full-event data which it receives.

The SEU rate on the full-event data should be rescaled for a trigger rate which the CIC can just maintain. Taking into account the fact that the tested full-event data packages contain 8 strip clusters and 8 pixel clusters this corresponds to a CIC package with a length of 297 bits, whereas $40 \times 8 = 320$ bits are available between two L1A triggers at 1 MHz. It is thus required to scale up the measured SEU rate with a factor of 1.08 [4]. This is for the case where a single MPA chip would be connected to the CIC. In the configuration where 8 MPA chips are connected to the CIC and all 8 produce full-event data with the same occupancy of strip and pixel clusters as considered before, then one full-event package has a length of 2033 bits, which takes the CIC ≈ 255 BX to output. If we want a continuous output of full-event data from the CIC chip in this configuration we need to reduce the L1A rate to 0.157 MHz, which means lowering the L1A rate with a factor of 6.4 compared to the L1A rate during the SEU test. For this last configuration we thus have to scale the SEU rate with a factor of $8/6.4 = 1.25$ which is close to the factor of 1.08 in the case of the connection of a single chip. The larger factor in the case where 8 MPA chips are connected originates from the fact that the header contributes relatively less to the number of bits which are transmitted in case data is received from 8 MPAs. These two examples are just to show that this scaling factor is close to unity and for this order of magnitude calculation can indeed be assumed to be unity[5].

The full-event data produced by the MPA chip will also contain the SEU induced errors in the SSA full-event data. Therefore the probability of an SEU in the MPA full-event data stream is the sum of the two upset rates in Table 8.5 where we use the high latency value as the most conservative and also most representative value for the operation in CMS. This results in an upset rate of $\approx 7 \times 10^{-4}$ $s^{-1}MPA^{-1}$ which, given the reasoning above, is also a good approximation for the upset rate on the full-event data output by the CIC chip. In one module, data from two CIC chips is combined and furthermore there is the possibility to run the CIC in the 640 MHz mode. Both these considerations will essentially double the SEU rate from a module compared to a single CIC at 320 MHz. Taking these two items into account the upper limit[6] on the median SEU rate for a single module can be estimated to be $\approx 3 \times 10^{-3}$ upsets/s/module on the full-event data path.

For the SEU rate estimation on the module stub data stream, a similar reasoning can be applied, but it should be considered that in order to generate a fake stub due to an SEU on the MPA a hit should already be present in the stub search window. This hit can be a real physical hit, a noise hit or can also be generated by an SEU. To take into account all details, full Monte Carlo simulations should be performed. As an upper limit, we can however sum the upset rates on the stub logic of both chips. The CIC, in the configuration where it outputs the data without bend at 640 Mbps on 6 lines, can output 40 stubs in 8 bunch crossings. In 8 bunch crossings a single MPA chip can output maximally 20 stubs. The upset rate on the stub logic for a single MPA chip thus has to be multiplied by a factor of 2. This results in an upset rate of 1.3×10^{-3} $s^{-1}CIC^{-1}$ and in an upset rate of 2.7×10^{-3} $s^{-1}module^{-1}$. This value is again an upper limit which was calculated for a maximum instantaneous luminosity and a maximum bandwidth of the CIC.

[4] In reality there will be a dead-time between the CIC full-event packages which will reduce this factor.

[5] In the end, the MPA and SSA upset rates will be used as input to large scale MC simulations to understand the effect of so-called *front-end inefficiencies* on the full system.

[6] Maximum instantaneous luminosity, close to maximal latency and maximum readout rate were considered.

Given the fact that there will be 5616 PS modules in the Phase-2 CMS Outer Tracker the above upset rates result in ≈ 15 upset bits in the stub data stream every second for all the PS modules, this is ≈ 1 bit every 2.6×10^6 collisions. For the full-event data the upset rate is ≈ 17 upset bits every second at a L1A rate of 1 MHz. This corresponds to an upset approximately every 60,000th triggered event.

The measured single chip SEU induced upset rates should be used in Monte Carlo simulations of the full tracker to evaluate their effect on the track trigger and tracking. The above estimation learns however that the upset rates due to SEU induced upsets in MPA and SSA can most likely be tolerated.

8.5 Single-MaPSA test beam

8.5.1 Introduction

A baby PS-p sensor bump bonded to an MPA chip, a so-called single-MaPSA (Macro Pixel Sub Assembly), was tested in beam at the CERN H6 North test beam area in April 2018. It was the first time such an assembly was tested in beam. Different configurations (depletion voltages, thresholds, beam incident angles) were tested. A telescope was present in the beam area and was used to do the track reconstruction. The following sections will give more information on the beam area, the set-up, the tested assembly and the results.

8.5.2 Beam area

8.5.2.1 Beam parameters

The CERN H6 North area is provided with 400 GeV protons from the SPS from which a secondary beam of hadrons is extracted for the test beam areas. Typically charged pions are delivered during an extraction. During the single-MaPSA test beam the momentum of the charged pions was 120 GeV/c. An extraction lasts for ≈ 4 seconds and occurs about every minute. During the extraction the beam can be considered to be continuous, there is no specific timing of the particles within a single extraction.

8.5.2.2 Telescope

The AIDA telescope, shown in Figure 8.12a, present in the beam area, was used to do the track reconstruction. This telescope consists of 6 planes of MIMOSA26 chips, a chip developed for the ILC project. 1 MIMOSA26 chip has an active area of $\approx 10.6 \times 21.2\,\text{mm}^2$ and contains 576×1152 pixels with a pitch of 18.4 μm. The chips are read out with a rolling shutter at a rate of ≈ 8.68 kHz, which is slow compared to the 40 MHz rate of the DUT DAQ. As a result the MIMOSA26 chips integrate many tracks during a full readout cycle. To be able to unravel this pileup and to match the tracks to hits on the DUT an extra timing reference is included in the telescope: the FE-I4 plane. The FE-I4 plane, developed for the ATLAS tracking system, has the required time resolution of 25 ns and has an active area of $16.8 \times 20.0\,\text{mm}^2$ (336 rows $\times$ 80 columns). The telescope is part of the EUTELESCOPE project and the EUDAQ software is available to do the track reconstruction.

The passage of particles through the telescope is indicated by the coincidence signal coming from 2x2 (2 downstream and 2 upstream of the telescope) scintillators. This

(a) (b)

FIGURE 8.12: LEFT: AIDA telescope at the CERN H6 beam line. RIGHT: single-MaPSA assembly with interface board mounted on the rotation stage at the center of the AIDA telescope.

coincidence signal is distributed using a trigger logic unit (TLU) [85] to all the detector systems: the 6 MIMOSA26 planes, the FE-I4 reference plane and the DAQ of the DUT itself. Also a trigger ID is sent to the subsystems. This ID can be used to match the events in the offline analysis.

8.5.3 Single-MaPSA

The single-MaPSA assembly consists of a PS-p baby sensor bump bonded to an MPA chip. The sensor has a width of 12 mm, a length of 25 mm and a thickness of 200 μm. The pixels themselves are 100 μm wide and 1467 μm long. Note that the sensor pixels are slightly longer than the MPA pixels which have a length of 1446 μm (see Figure 5.14). The spatial resolution is obviously defined by the sensor dimensions. Figure 5.14 also explains the coordinate system used for this analysis: the x and y-axis are respectively along the fine and coarse pixel dimensions. The incident angle α is the angle between the normal vector of the sensor plane and the particle direction. During the angular scan the single-MaPSA was rotated around the y-axis. A CV curve of the sensor showed a depletion voltage of -150 V.

The pixel layout, shown in Figure 8.13, shows the boundary of 4 pixels. The bias rail supplies the ground to the n^+ *dots* of the punch-through structure. When the sensor is biased from the backside (using a negative, with respect to ground, high voltage) a depletion region starts to grow starting from the dot. As soon as this depletion region reaches the pixel implant, the pixel implant will be set to the same potential as the dot (which is ground in this case). This is the so-called *punch-through* effect. The p-stop ring around the dot prevents a short between the dot and the main pixel implant. This short would result in all the charge, generated by an ionizing particle, to be drained by the bias rail rather then collected by the aluminium pads which are on top of the pixel implant.

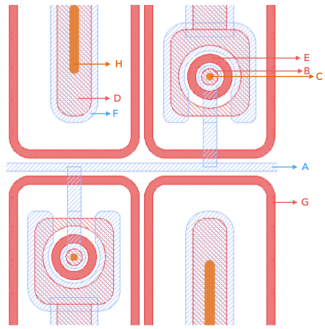


FIGURE 8.13: Layout of the single-MaPSA pixel sensor matrix [18]. The bias structure (A) distributes the ground to the *dots* (B) of the punch-throughs using a via connection (C). These dots are a small n^+ implant with the same donor concentration as the pixel implant. They are separated from the main pixel implant (D) using p-stop isolation (E) around the punch-through structure. The charge is collected on the Al pad (F) on top of the pixel implant. The pixel implant is connected to the aluminium pad using a via (H) (DC process). The pixels are physically separated using p-stop isolation (G).

8.5.4 Single-MaPSA set-up

The single-MaPSA assembly was placed at the center of the telescope on a rotation stage as shown in Figure 8.12b. The MPA was read out with the same hardware as used for the single chip testing: the MPA, wire bonded to a dedicated carrier card was connected to the same interface board, as used for the bench top tests. The interface board regulates the power to the chip and does the level translation to and from the back-end (μDTC system), which is connected with a 1 m VHDCI cable. The trigger signal was provided to the VHDCI-FMC, which hosts a LEMO connector. The μDTC exerts back pressure to the TLU. This back pressure is raised from software whenever a run is not ongoing. The back pressure can also be raised from the firmware when the data buffers on the FC7 are almost full. In this case the back pressure signal is raised and the triggers to all the subsystems are suppressed.

The clock that drives the MPA chip is generated on the FC7 and is uncorrelated with the arrival time of the particles. Therefore the arrival of the trigger signal from the TLU is not always happening at the same phase with respect to the clock, and thus the sampling time, on the chip. A time to digital converter (TDC) with a resolution of 3.125 ns was implemented in the firmware to give a better resolution on the arrival time of the trigger signal from the scintillators. The TDC value is passed on to the back-end in the event's data package.

8.5.5 MPA data streams

The MPA chip was operated in two modes during this test beam: the asynchronous and the synchronous mode (see section 5.5.3). In case of the asynchronous mode the shutter is opened for some time during which the ripple counter for each pixel counts the number of times a signal above threshold was detected. After closing the shutter the counters are read out over the stub lines.

The disadvantage of the asynchronous mode is that the counters on the MPA are self triggered and there is no common trigger signal available which can be sent to the telescope. As a result for these runs no matching to telescope tracks is available. The asynchronous mode has the advantage that data can be collected at high rate contrary to the synchronous mode where the DAQ rate is limited by the rate at which the telescope planes can be read out. When running in asynchronous mode the counts are normalised to the counts which were received by a scintillator downstream of the telescope.

The synchronous readout is the default operation mode of the MPA chip and in this mode the full-event information is passed to the back-end on reception of a trigger. Being a single sensor assembly, the single-MaPSA does not have the ability to form stubs. The MPA can however be run in so-called pixel-pixel mode where the stub lines do not carry the stub information but are used to carry the location of pixel clusters. The decoding, as shown in Figure 5.16b, is the same as for the stubs with the bend being always zero. In this way there is a full-event and a cluster data stream, which are independent, but should carry the same information, with the limitation that the cluster data is bandwidth limited to 5 clusters/2 BXs. The data stream on the stub lines does not give us extra information, on top of the full-event data, but the consistency with respect to the full-event data stream can be checked and furthermore provides feedback whether the event building in the μDTC is working properly.

8.5.6 Commissioning of the single-MaPSA

Before data taking can start the chip needs to be commissioned. This includes recalibrating the MPA chip's bias block DAC settings to the nominal settings, tuning the front-ends of each pixel in order to be able to apply a common threshold and finding the optimal latency setting. These commissioning steps were performed when the assembly was set-up in the test beam area.

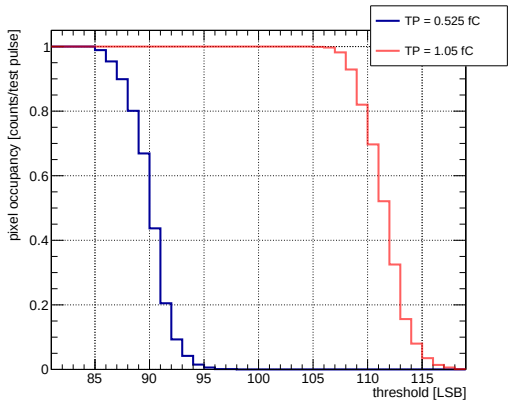
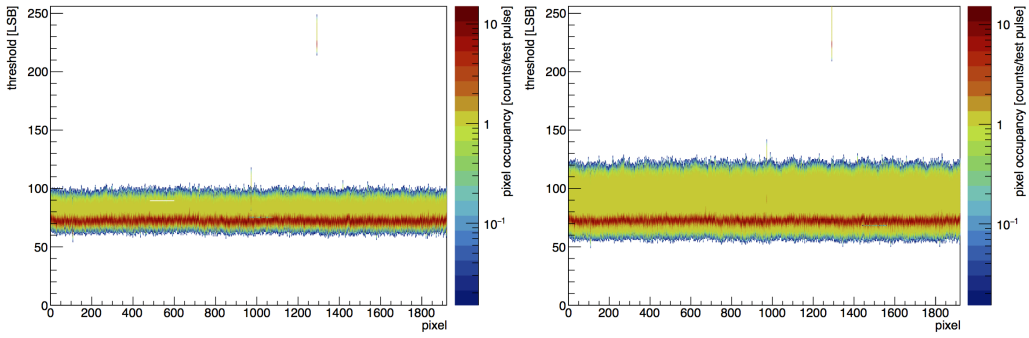


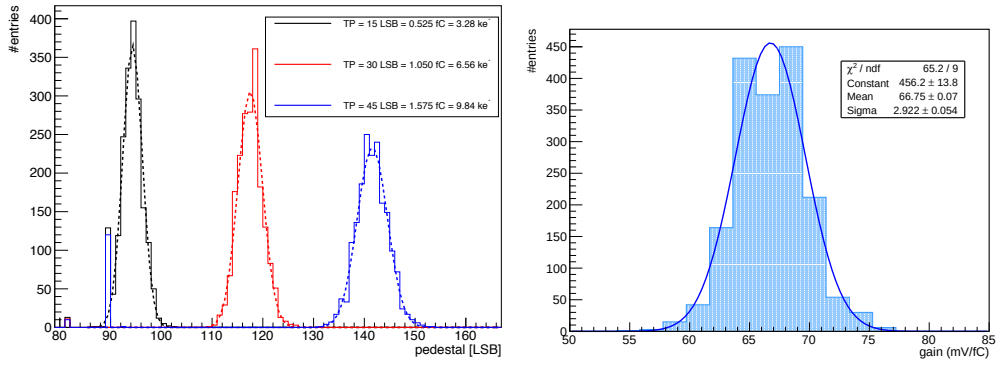
FIGURE 8.14: Example of an S-curve extracted for a single pixel when pulsed with test pulses of different amplitude.

The optimal bias block DAC settings were extracted by measuring the output of the analogue multiplexer (AMUX) with a multimeter, scanning the full DAC range and finding the setting for each DAC which gives the output on the AMUX as close as possible to the default settings. These default settings are the ones reported in the MPA manual [66].



(a) (b)

FIGURE 8.15: S-curves after calibration for a calibration pulse of 0.525 fC (LEFT) and 1.05 fC (RIGHT).



(a) (b)

FIGURE 8.16: LEFT: Pedestal distribution for three different test pulse amplitude settings. RIGHT: Gain distribution showing the gain calculated for each pixel.

These default values are extracted from the chip simulation and are provided by the chip designers.

After the calibration of the DAC settings the tuning of the pixel front-ends can be done. This calibration was performed with the sensor biased to -200 V. The tuning procedure *trims* the front-ends of each pixel in order to get the threshold for the comparator for each pixel identical. The results can be checked by injecting test pulses using the calibration capacitors in the front-ends of the chip. By performing a threshold scan and measuring the occupancy, a result as the one shown in Figure 8.14 is extracted. The figure shows an *S-curve* measured for a single pixel on the MaPSA for two different settings of the test pulse amplitude (0.525 fC $\pm$ 20% and 1.050 fC $\pm$ 20%, the 20% uncertainty is given here by the process variations in the chip production). The S-curves as measured for all the pixels are illustrated in Figures 8.15a and 8.15b. These calibration results were obtained by running the MPA in the asynchronous mode and injecting test pulses. The typical shape of the curve is explained as follows, going from high to low threshold:

Run #	Assembly	Type	Threshold (LSB)	Bias voltage (V)	Incident angle α (°)	MPA Synchronous/ Asynchronous	Telescope data available?	Sampling mode
1	1	Threshold scan	70, 72, ..., 250	-200	0	Asynchronous	No	Ripple counter
2	2	Threshold scan	70, 71,..., 251	-200	0	Asynchronous	No	Ripple counter
3	1	Latency scan	90, 110, 250	-200	0	Synchronous	No	Level
4	1	Latency scan	100, 120, ..., 240	-200	0	Synchronous	No	Edge
5	1	Angular scan	110	-200	-15, -10, -5, 0, 10, 20, 30, 40	Synchronous	No	Level
6	1	Bias scan	110	-200,-150, -100, -50	0	Synchronous	No	Level
7	1	Threshold scan	90, 110, 150	-200	0	Synchronous	Yes	Level
8	1	Bias scan	110	-200, -150, -100	0	Synchronous	Yes	Level
9	2	Sampling modes	110	-200	0	Synchronous	Yes	Level and Edge

TABLE 8.6: List of scans performed during the single-MaPSA test beam.

at high threshold the comparator never switches, so no counts are registered. Then at lower threshold (e.g. threshold $\approx$ 100 for Figure 8.15a) the comparator starts to switch and 1 count per test pulse is registered. The middle (50% occupancy) of the S-curve is referred to as the *pedestal*. At even lower thresholds the threshold is low enough so that the comparator also triggers on noise (e.g. threshold $\approx$ 70 for Figure 8.15a). For even lower thresholds the pixel occupancy per test pulse drops again to zero as there will be no more comparator transitions because the noise will always be above threshold.

The distribution of the pedestals for all pixels are shown for three different test pulse settings in Figure 8.16a and the distributions of the gains, as calculated for each pixel is shown in Figure 8.16b. An average gain of 66.75 mV/fC and a baseline[7] of 106 mV are extracted.

After calibration of the DACs and tuning of the front-ends a threshold can be chosen above the noise in order to do the latency scan. Two latency values need to be extracted to set up the system for data taking: the L1 latency and the stub latency. It has to be kept in mind that the *stub* data in this specific configuration are actually pixel clusters. The results from the latency scans will be discussed in section 8.5.7.1

After calibration of the DACs, the tuning of the offsets and the latency scans, the system is ready for data taking.

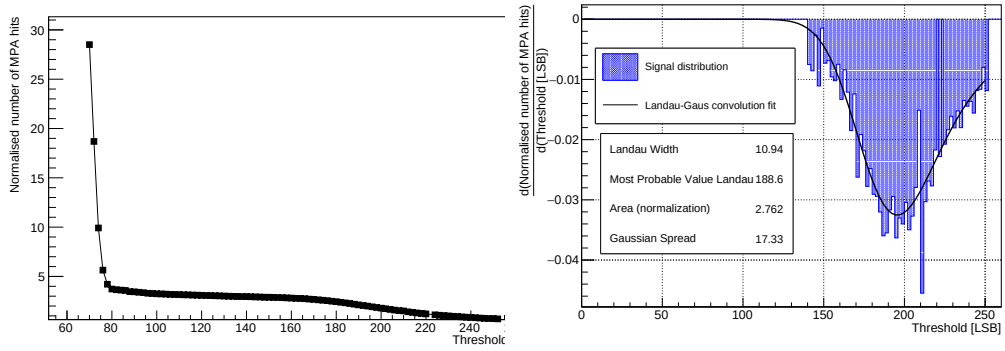
8.5.7 Results

During the single-MaPSA test beam, two assemblies were tested. The scans which were performed are summarised in Table 8.6 and will be discussed in the next sections.

[7] Pedestal at zero injected charge.

8.5.7.1 Assembly #1 and #2: DUT only

A fine threshold scan in asynchronous mode at perpendicular incidence and a bias voltage of -200 V was performed for both assembly #1 and assembly #2. The results are shown in Figure 8.17 and 8.18 respectively. The integrated counts on the ripple counter from the full MPA, normalised to the number of hits recorded by a downstream scintillator[8] are shown in Figure 8.17a and 8.18a. The derivative of these distributions are shown in Figure 8.17b and 8.18b respectively and show the signal distribution. The signal distribution in the signal region is fitted with a convoluted Gaussian and Landau.



(a) (b)

FIGURE 8.17: LEFT: Number of hits recorded on the ripple counter of all the MPA pixels normalised to the number of hits recorded by a downstream scintillator in function of the threshold. RIGHT: The derivative with respect to threshold of Figure 8.17a gives the signal distribution. The signal distribution is fitted with a convolution of a Landau and a Gaussian. The result for values lower than 135 LSB are not shown. These results are for assembly #1.

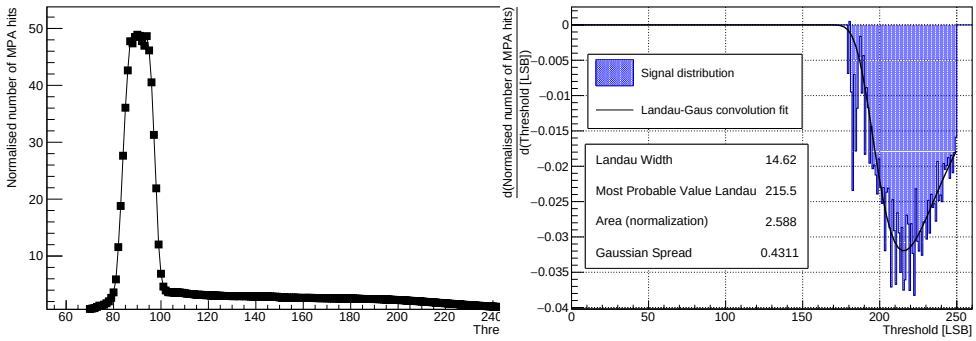
For assembly #1 this results in a most probable value for the signal of 188.6 threshold LSB. Together with the baseline at 72 threshold LSB, a most probable signal of ≈ 14 ke$^-$ (70 e$^-$h$^+$/μm, from Figure 3.3a for a 200 μm thick sensor) and the fact that 1 threshold LSB corresponds to 1.456 mV [66] a gain of 75.7 mV/fC can be extracted which is close to the 66.75 mV/fC extracted using the test pulse measurement. The discrepancy can be explained by the uncertainty of the injected charge when using the test pulse and the unknown charge collection efficiency of the sensor.

For assembly #2 a similar calculation, using a most probable value for the signal of 215.5 LSB and a baseline at 97 LSB, leads to a gain of 76.6 mV/fC. The gain values, calculated with the data from the beam particles, is coherent when comparing the two assemblies.

For the threshold settings which were used during this test beam the physical threshold values obtained with the gain values calculated above are given in Table 8.7.

Figure 8.19a and 8.19b show the L1 latency scan performed for assembly #1 in the level and edge sampling mode (see Figure 5.11) respectively. For both modes it can be seen that the higher the threshold, the smaller becomes the latency window to operate in. For the level sensitive mode a single latency value can be chosen which will result in

[8]The disconnected pixels (see later) were masked for the analysis.



(a) (b)

FIGURE 8.18: LEFT: Number of hits recorded on the ripple counter of all the MPA pixels normalised to the number of hits recorded by a downstream scintillator in function of the threshold. RIGHT: The derivative with respect to threshold of Figure 8.18a gives the signal distribution. The signal distribution is fitted with a convolution of a Landau and a Gaussian. The result for values lower than 175 LSB are not shown. These results are for assembly #2.

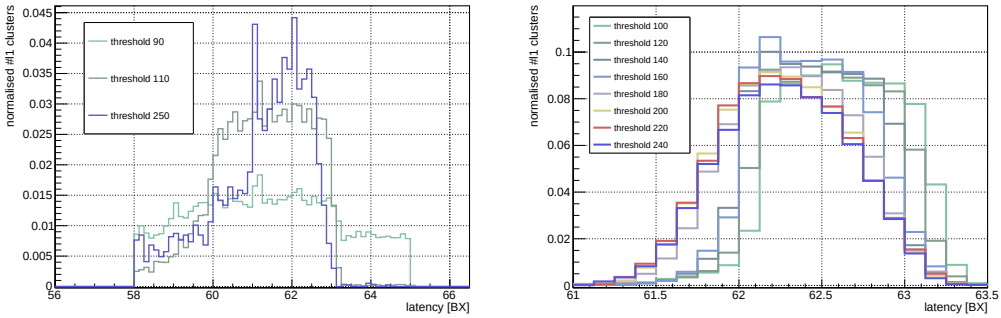
Threshold (LSB)	Assembly #1 Physical threshold	Assembly #2 Physical threshold
90	0.346 fC = 2.15 ke$^-$	/
110	0.731 fC = 4.56 ke$^-$	0.247 fC = 1.54 ke$^-$
150	1.500 fC = 9.36 ke$^-$	/

TABLE 8.7: Threshold settings as used during the test beam for both assemblies.

a maximal efficiency for the full threshold range shown in Figure 8.19a. The latency scan in the level sensitive mode also shows that the time over threshold of the comparator output can span several clock cycles, this represents the relatively slow return to baseline of the shaper in the MPA front-end which is of the order of 6 BXs. For the edge sensitive mode the shift in the optimal latency with threshold is more pronounced. Both figures show that the lower the threshold, the earlier (larger latencies) becomes the time over threshold. Most of the runs were taken in the level sensitive mode, so that a tuning of the latency setting is not required when changing the threshold.

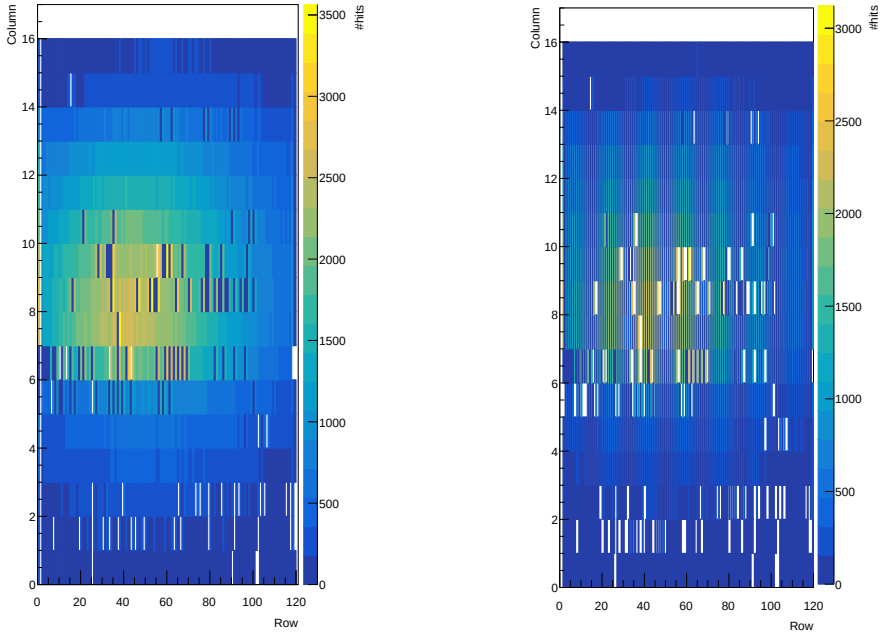
Figure 8.20a shows the 2 dimensional beam profile on assembly #1 reconstructed from the full-event data after setting the correct L1 latency. The data was taken at an incident angle of 0° and a threshold setting of 110 LSB. The dead pixels, due to pixels not properly bump bonded to the MPA front-ends, are clearly visible in the 2D hit map. A similar beam profile can be reconstructed, after proper tuning of the stub latency, using the cluster data as shown in Figure 8.20b. The latter figure shows the half-strip resolution which can be attained for two strip clusters.

Figure 8.21a shows the mean cluster size obtained during the angular scan (-200 V bias voltage and threshold of 110 LSB) of assembly #1 in the synchronous run. A fit with $1/\cos(\alpha)$ is shown. The rotation axis of the single-MaPSA assembly lies along the coarse pixel dimension. The increase in cluster size with non-perpendicular incidence



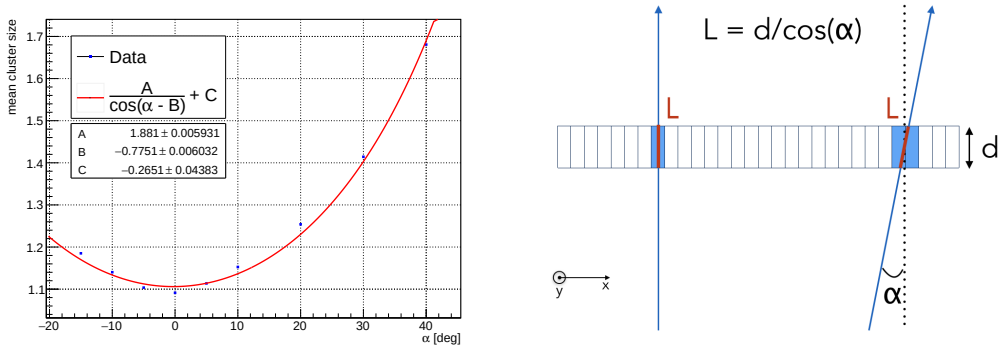
(a) (b)

FIGURE 8.19: L1 latency scan for assembly #1 in the level (LEFT) and edge (RIGHT) sampling mode.



(a) (b)

FIGURE 8.20: Two-dimensional beam profile on assembly #1 from full-event data (LEFT) and stub data (RIGHT).



(a) (b)

FIGURE 8.21: LEFT: Mean cluster size distribution for single cluster events on the single-MaPSA versus incident angle. The distribution follows a $1/\cos(\alpha)$ behaviour, representing the increase in charge sharing between pixels due to the particle crossing more of the active volume when the incidence is not perpendicular. RIGHT: Explanation of the charge sharing with rotation: d is the thickness of the sensor, α the incident angle and L the path length of the particle through the sensor.

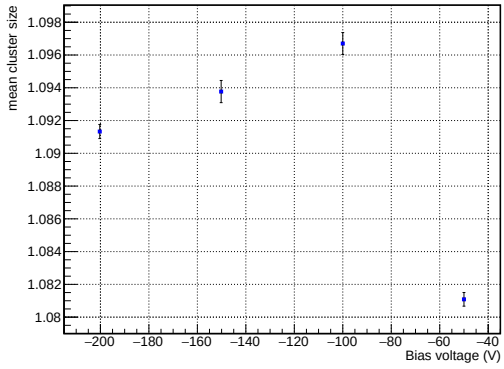


FIGURE 8.22: Mean cluster size distribution for single cluster events on the single-MaPSA versus bias voltage for assembly #1.

represents the charge sharing due to the particles travelling through more of the active sensor volume as illustrated in Figure 8.21b. For this result only events with a single reconstructed cluster were used. This was done to circumvent the pollution to the cluster size distribution which the MPA introduces by splitting clusters of size larger than 8 into two separate clusters.

In Figure 8.22 the mean cluster size on assembly #1 is shown as a function of bias voltage when the threshold was set to 110 LSB and with perpendicular incidence of the particles. Again only single cluster events were used. At low voltages (-50 V in Figure 8.22) the sensor is not fully depleted yet and the charge collection efficiency of the sensor is not optimal, which reduces the probability of getting cluster sizes larger than 1. At higher absolute voltages the drop in mean cluster size can be attributed to the decrease in diffusion due to the higher electric field strengths.

8.5.7.2 Assembly #1: combined results of threshold scans

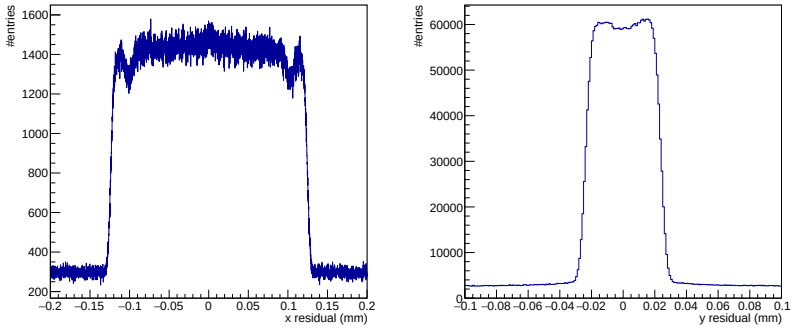
In this section some qualitative results are shown which combine the results obtained for the three runs at different threshold settings (90, 110, 150 LSB) for assembly #1. This combination is done to increase the statistics, or are presented combined, because the results are independent of the threshold setting. The quantitative results are given in the following sections, where the three threshold settings are treated separately.

As a sanity check for the track reconstruction with the telescope, the x and y residuals of the reference plane, which in this set-up is the FE-I4 plane, are shown in Figure 8.23. The x (y) residual is defined as the x (y) distance between the location of the extrapolated telescope track on the DUT plane and the location of the best matching hit on the DUT. Figure 8.23 indeed shows the FE-I4 pixel dimensions which are 250×50 μm^2 and also show the reduced number of entries at the end of both sides of the pixel where the punch-throughs are located. Based on these residual plots a cut is made on the tracks for further analysis: tracks with an absolute residual on the reference plane which is larger than 0.11 mm in x or 0.02 mm in y are not considered for the rest of the analysis.

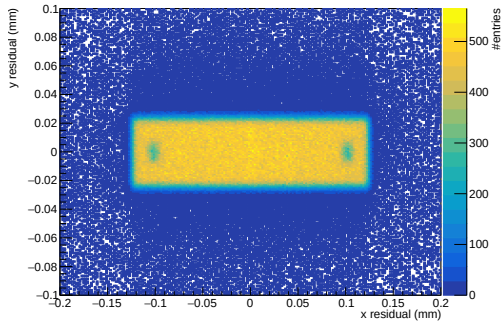
Using these selected tracks, also the residuals on the DUT can be calculated and are shown in Figure 8.24. These residual values represent the pixel dimension of the single-MaPSA, which is 100 μm $\times$ 1467 μm, convoluted with the resolution of the telescope and the DUT. The residual in x, which is the fine pixel dimension, also shows the contribution of clusters with size larger than 2, which give an increased spatial resolution.

A DUT hit is only considered to be matched to a telescope track if the absolute residual in x is smaller than 0.07 mm and smaller than 0.9 mm in y. Only hits which pass this criterion contribute to the numerator of the DUT hit efficiency. DUT hit efficiency maps, based on this definition, are shown in Figure 8.25a which shows that there are dead pixels in the assembly and that the telescope does not cover the full single-MaPSA assembly in the y direction.

The dead pixels are due to a suboptimal assembly procedure of the PS-p baby sensor to the MPA chip. Later assemblies with improved assembly procedures are able to obtain connection of the full pixel matrix to the MPA chip. In further analysis and efficiency calculations these dead pixels are masked and also a fiducial region, masking the edges of the single-MaPSA assembly is defined if not stated otherwise. These two masking steps result in the efficiency map shown in Figure 8.25b. This efficiency map shows inefficiencies along the bias rails for every row of pixels, but more importantly shows a high pixel efficiency over the full studied assembly.

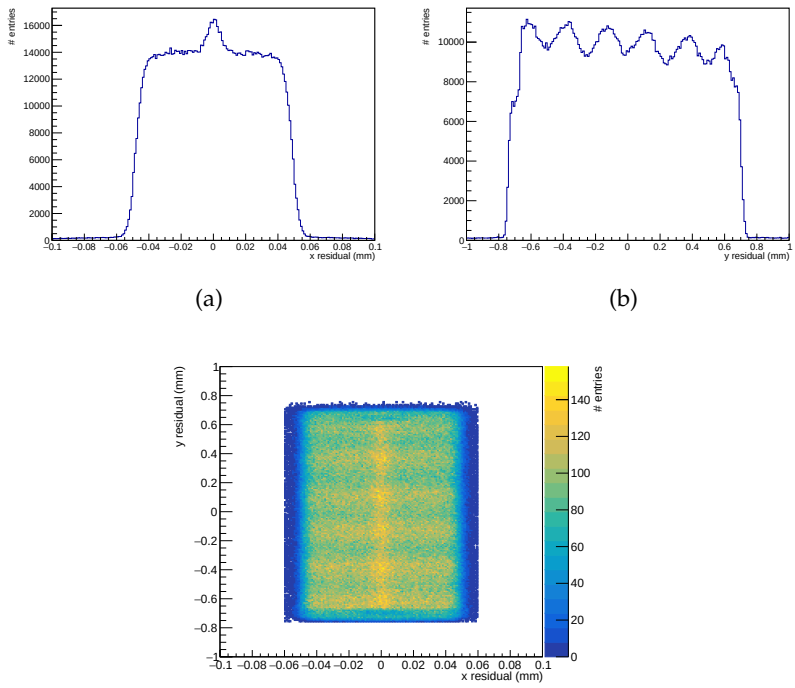


(a) (b)



(c)

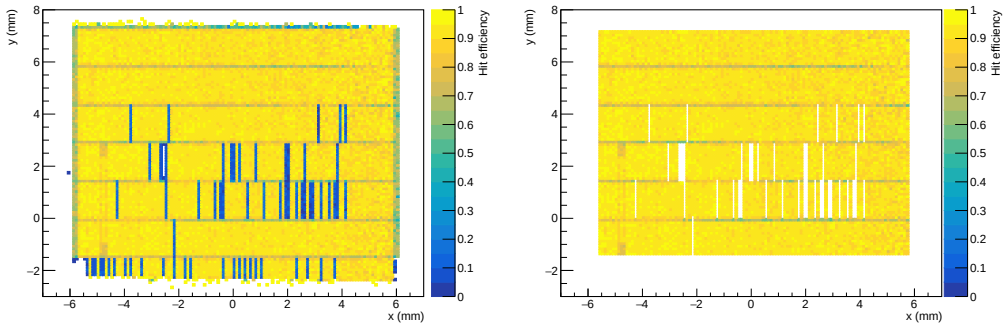
FIGURE 8.23: Residual in x (TOP LEFT), y (TOP RIGHT) and xy (BOTTOM) of the FE-I4 reference plane.



(a) (b)

(c)

FIGURE 8.24: Residual in x (TOP LEFT), y (TOP RIGHT) and xy (BOTTOM) of the single-MaPSA.



(a) (b)

FIGURE 8.25: xy efficiency map of the first single-MaPSA assembly showing the dead pixels (LEFT) and with masked dead pixels and a fiducial region applied (RIGHT).

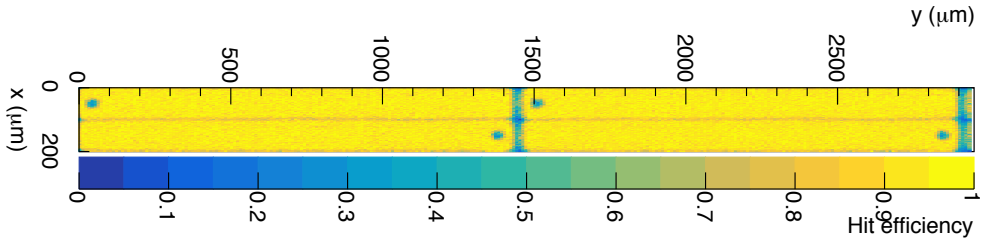


FIGURE 8.26: Inpixel efficiency for 4 pixels. This is obtained by folding the data from the studied part of the first single-MaPSA assembly to increase the statistics inside the pixels.

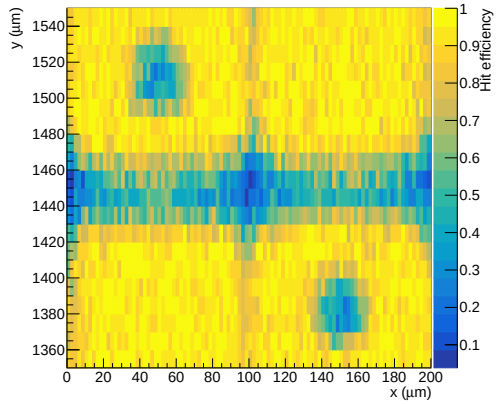
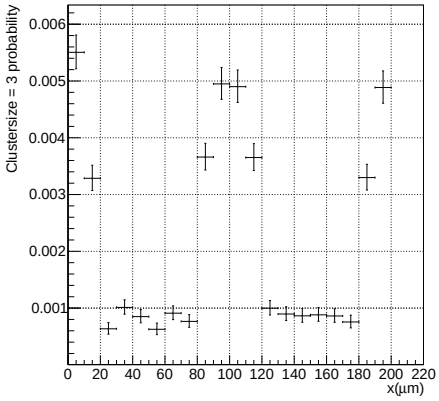


FIGURE 8.27: Zoom on the bias rail and punch-through region for the data shown in Figure 8.26.

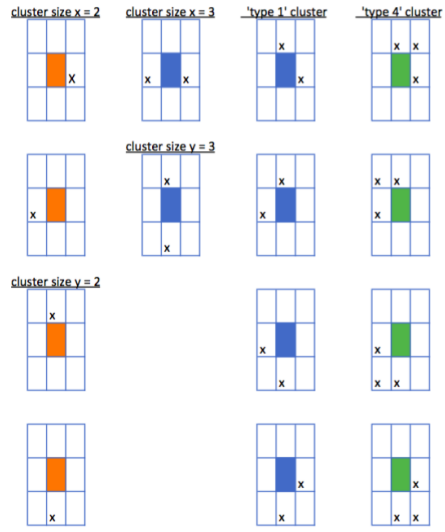
To better study the inefficiencies due to the bias structure, so-called *inpixel efficiency maps* can be extracted by overlapping the data from different pixels on the same set of 4 pixels. This set of 4 pixels is the unit of the layout of the pixel matrix, as represented in Figure 8.13. An inpixel efficiency map is shown in Figure 8.26 with a zoom on the bias and punch-through region in Figure 8.27.

Besides the efficiencies, also the cluster size distributions can be studied. Probability distributions for clusters of size 1 and 2 will be discussed later (see section 8.5.7.3). The focus here will be on clusters with sizes larger than 2. The study of this type of clusters requires the combination from the data of all three threshold settings due to the low probability for this type of clusters to occur. Figure 8.28a shows the probability for cluster size 3 as function of the hit location of the particle. It can be seen that when the particle hits close to the pixel borders the probability of creating a cluster of size 3 increases. It must be noted however that the probability of clusters of size 3 is very low. The only size 3 clusters which were included in Figure 8.28a are clusters which have 3 hits in consecutive pixels along the fine pixel dimension.

Cluster distributions can also be studied along the coarse pixel dimension. Here the overall probability of clusters spanning multiple pixels is smaller due to the larger pixel



(a)



(b)

FIGURE 8.28: LEFT: Probability of cluster size 3 along the fine pixel dimension combining data from the three threshold settings used for the first assembly. RIGHT: Definition of cluster types where the coloured pixel represents the central pixel which is hit and the pixel filled with an 'x' the neighbouring pixel(s) which is (are) hit. The two leftmost columns show the standard 2 and 3 pixel clusters along x and y, simply for completeness. The two right columns introduce the new cluster types *type 1* and *type 4*.

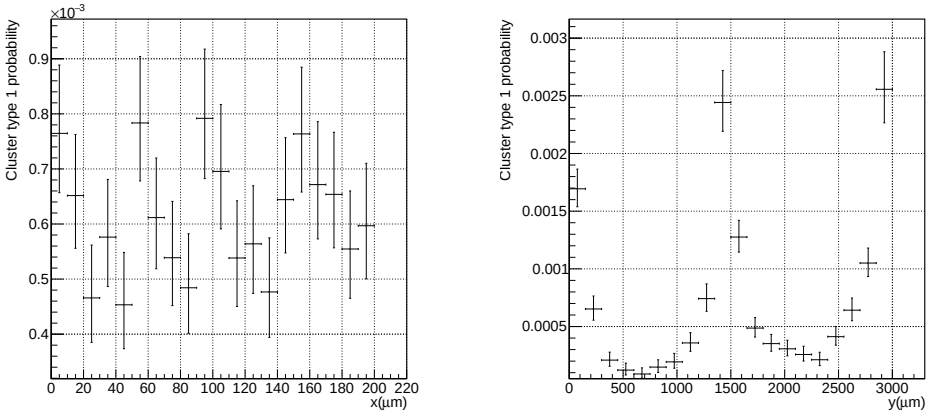
size in this dimension. Furthermore the MPA does not perform clusterisation along y, so clusters in this dimension will not add to an improved spatial resolution, unless reconstructed offline. The probability for cluster sizes 1 and 2 in function of y can be found in section 8.5.7.3. The probability for clusters spanning 3 consecutive pixels in the y direction was found to be zero.

Two other cluster types were also studied. These will be referred to as *type 1* and *type 4* and are defined in Figure 8.28b[9]. A non-zero probability for these clusters was found and their probability as function of x and y are shown in Figure 8.29 and 8.30 for type 1 and type 4 respectively. It must be noted that the probability for these cluster types is very low, around the per mille level, but nevertheless a higher probability for these clusters is observed along the pixel boundaries in y.

8.5.7.3 Assembly #1: threshold scan

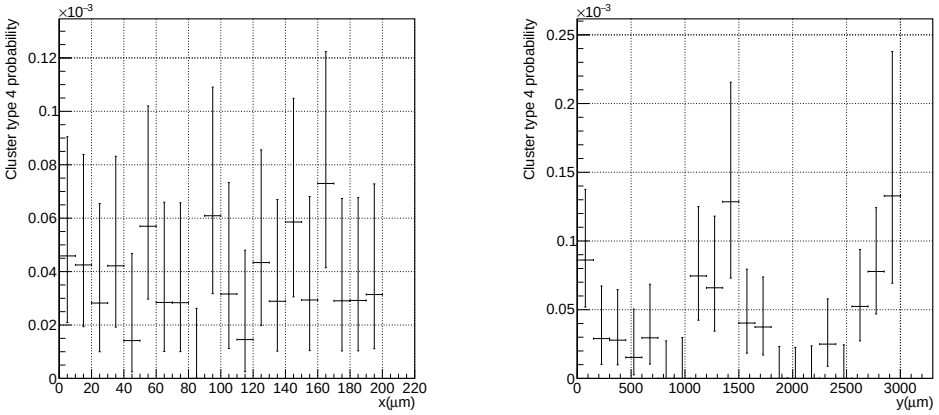
After a more qualitative discussion in the previous section, the exact efficiency values for the different threshold settings will be discussed in more detail in this section by projecting the 2D efficiency maps along x and y. This will be done for the three threshold settings with which assembly #1 was operated. Figure 8.31 shows the TDC dependence of the overall efficiency. It shows that for the highest threshold settings the efficiency

[9]Type 4 clusters contain type 1 clusters, but no entry is made in the numerator of the probability for type 1 clusters if a type 4 cluster was found.



(a) (b)

FIGURE 8.29: Probability of clusters of type 1 (defined in Figure 8.28b) along the fine (LEFT) and coarse (RIGHT) pixel dimension obtained by combining data from the three threshold settings used for the first assembly.



(a) (b)

FIGURE 8.30: Probability of clusters of type 4 (defined in Figure 8.28b) along the fine (LEFT) and coarse (RIGHT) pixel dimension obtained by combining data from the three threshold settings used for the first assembly.

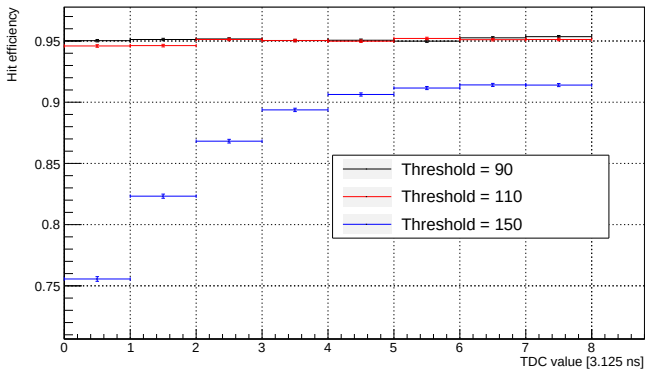
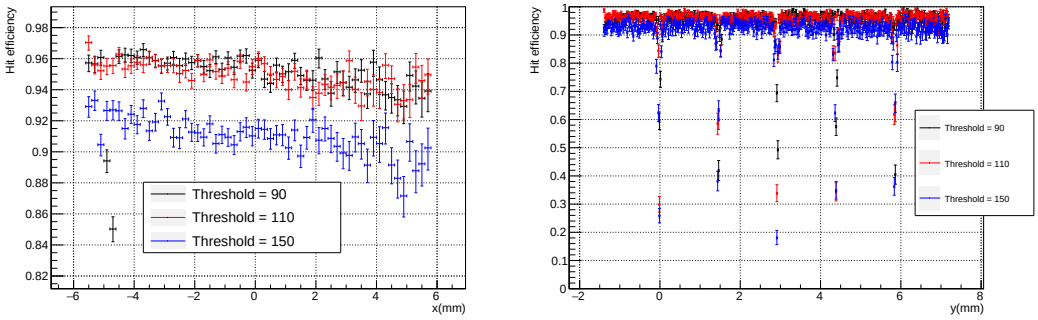


FIGURE 8.31: TDC dependence of the efficiency for the three tested threshold settings with the first single-MaPSA assembly.



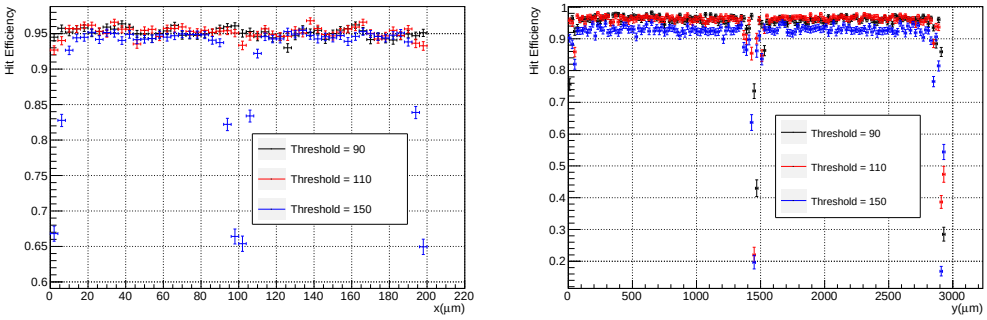
(a) (b)

FIGURE 8.32: Efficiency along the fine (x) and coarse (y) pixel dimensions for the three threshold settings used during the test beam for the first single-MaPSA assembly.

strongly depends on the TDC phase. Therefore for the rest of the results in this section only events which end up in TDC bins 6 and 7 will be used for the efficiency calculations.

Projections of the efficiency along x and y for the three studied threshold settings can be found in Figure 8.32. Figure 8.32a shows, for the lowest thresholds an efficiency higher than 94%. Also a decrease in efficiency with increasing values of x is visible. The reason for this dependency is not understood and was also observed for the second assembly (see section 8.5.7.5). Figure 8.32b shows that the efficiency drops to as low as 20% at the bias rails, even for the lowest threshold setting. The decrease in efficiency due to the bias rail can be recovered by tilting the assembly so that the particles do not enter the silicon exactly perpendicular.

The inpixel efficiencies can be found in Figure 8.33 for both pixel dimensions. Figure 8.33a shows that also along x there are inefficiencies along the pixel boundaries. These inefficiencies can partly be attributed to charge sharing between pixels. For the two lowest thresholds these inefficiencies are rather small ($\approx$ 2% compared to the central regions).



(a) (b)

FIGURE 8.33: Inpixel efficiency along the fine (x) and coarse (y) pixel dimensions for the three threshold settings used during the test beam for the first single-MaPSA assembly.

The influence of the punch-through on the efficiency is illustrated in Figure 8.34 where the 4 pixels are folded on top of each other to 1 pixel. For Figure 8.34a only tracks were used which were found to pass within 5 μm from the center of the punch-through in y. A similar cut was done for Figure 8.34b where only tracks were selected which pass within 5 μm from the center of the punch-through region in x. Figure 8.34a shows a dip in the central region for the two highest threshold settings, where the efficiency drops as low as 70% for the x values which coincide with the punch-through region. A similar behaviour can be seen in Figure 8.34b where for the y values lower than 100 μm the efficiency is low, this first due to the pixel boundary and then the punch-through region.

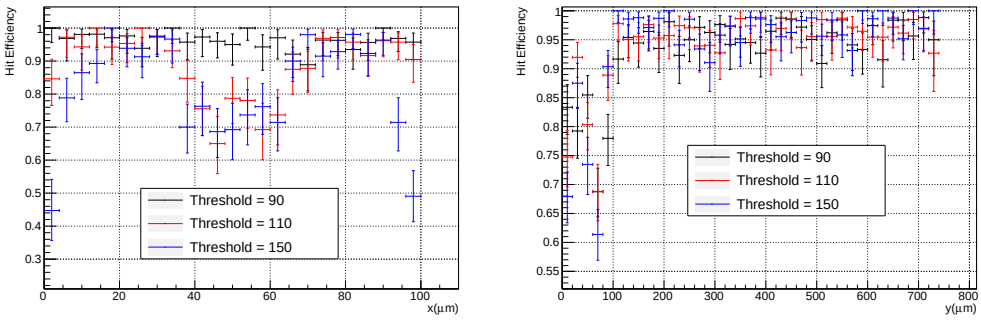
Figure 8.35 shows the efficiency averaged over the full sensor for three separate fiducial areas:

- case 0: No masking was applied.
- case 1: Dead pixels were masked and a fiducial region was defined in order to exclude the border of the assembly (Figure 8.25b).
- case 2: On top of the one in the bullet above, an extra fiducial region inside the pixel is defined in order to not include the inefficiencies along the bias rails, pixel edges and the punch-through structures. This fiducial region excludes inpixel hits with a location smaller than 20 μm, between 80 μm and 120 μm and larger than 180 μm in x and smaller than 100 μm, between 1300 μm and 1600 μm and larger than 2800 μm in y.

Figure 8.35 shows that by defining the fiducial region as in *case 2*, the inefficiencies seen at operation with a threshold of 150 LSB can largely be recovered. For thresholds of 90 and 110 LSB the efficiencies for case 2 are larger than 96%.

Different sensor designs were tested in subsequent test beams with designs which reduce the inefficiencies around the punch-throughs and bias rails. Details on these studies can be found in Ref. [18].

The probability for clusters of size 1 and size 2 along the fine and the coarse pixel dimension at different thresholds are shown in Figure 8.36 and 8.37 respectively. The figures show the expected behaviour: the probability for clusters of size 2 is smaller when



(a) (b)

FIGURE 8.34: Inpixel efficiency along the fine (x) and coarse (y) pixel dimensions, specifically for particles passing the punch-through region in y (LEFT) and in x (RIGHT). This is shown for the three threshold settings used during the test beam for the first single-MaPSA assembly.

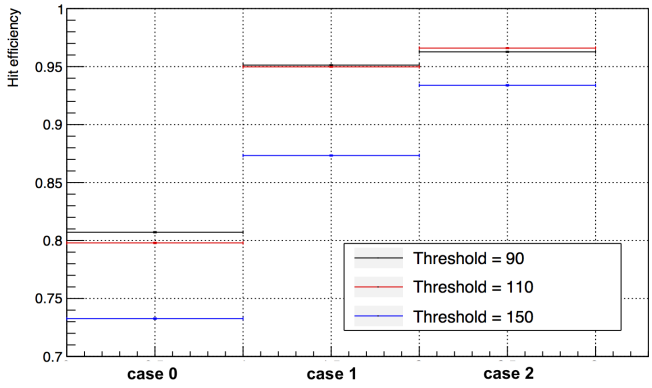
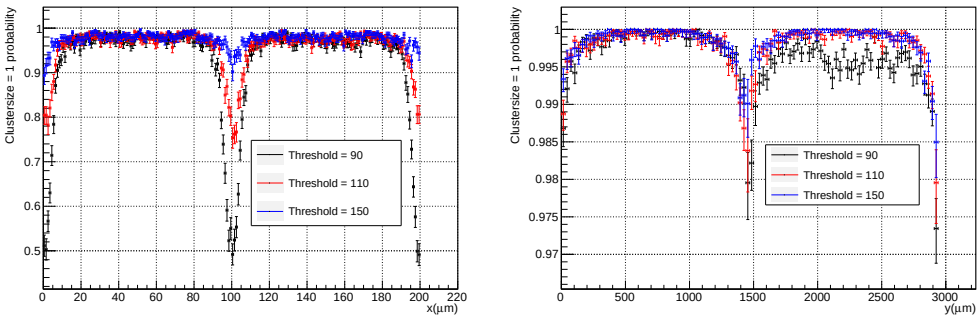
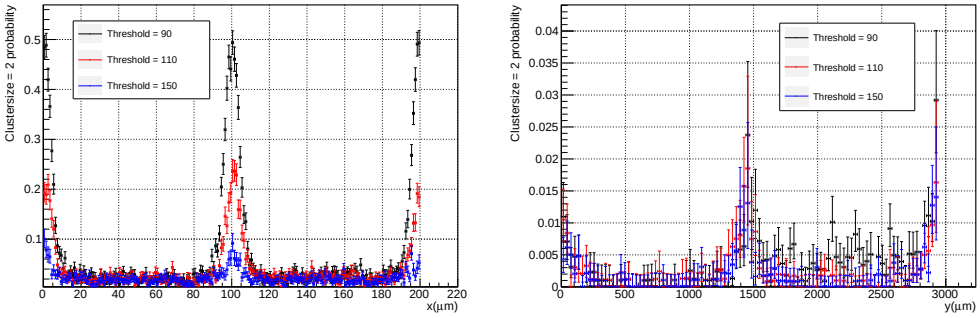


FIGURE 8.35: Average efficiency values for the first assembly for three definitions of the fiducial region during the threshold scan of the first single-MaPSA assembly. The definitions of the fiducial regions are explained in the text.



(a) (b)

FIGURE 8.36: Inpixel probability for clusters of size 1 along the fine (x) and coarse (y) pixel dimensions for the three threshold settings used during the test beam with the first single-MaPSA assembly.



(a) (b)

FIGURE 8.37: Inpixel probability for clusters of size 2 along the fine (x) and coarse (y) pixel dimensions for the three threshold settings used during the test beam with the first single-MaPSA assembly.

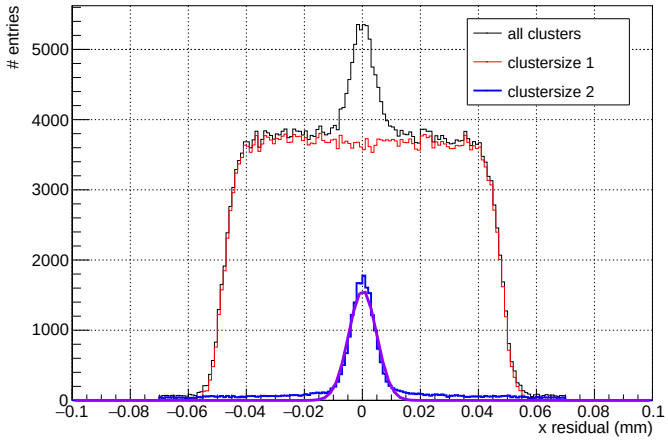


FIGURE 8.38: Residuals for clusters of size 1 and of size 2 for the first single-MaPSA assembly at a threshold of 90 LSB. Clusters of size 2 exhibit a smaller residual which shows the higher resolution which can be obtained with this type of clusters. The purple line in the figure shows a Gaussian fit to data for clusters of size 2.

the threshold is increased. For the lowest threshold setting the probability for clusters of size 2 along the fine pixel dimension can reach 50% near the pixel boundary.

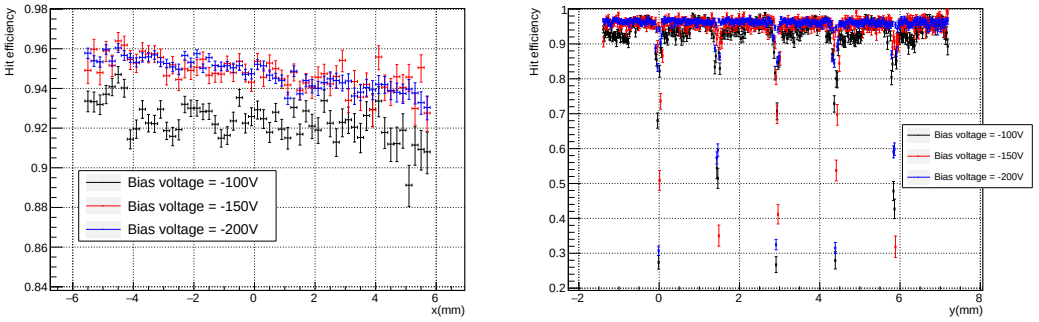
Figure 8.38 shows the x residuals for the run taken at a threshold of 90 LSB unfolded for cluster sizes 1 and 2. The events with cluster sizes of 2 show a better spatial resolution. A Gaussian fit to the distribution of cluster sizes of 2 gives a standard deviation of 27.5 μm. Both the resolution of the single-MaPSA and the track pointing resolution from the telescope contribute to this resolution. Taking a telescope track pointing resolution of 2 μm [86], a resolution of 27.4 μm can be extracted for the single-MaPSA, which is close to the 28.9 μm resolution obtained using a pixel pitch of 100 μm in Equation (3.4).

8.5.7.4 Assembly #1: bias scan

Figure 8.39 shows the x and y dependence of the efficiency over the full studied part of the single-MaPSA assembly when biasing the sensor to -100, -150 and -200 V and operating at a threshold of 110 LSB. Again the decrease in efficiency with increasing x values can be seen. Both figures show that the efficiency is lower when the sensor bias was only set to -100 V, which is below depletion of the sensor.

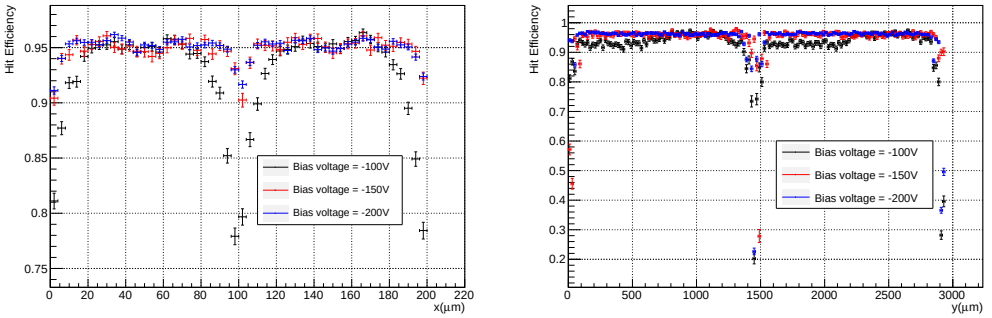
Figure 8.40 shows the inpixel efficiency values for the three tested bias voltages. Figure 8.40a shows that the lower efficiency in the -100 V case can be attributed to inefficiencies along the pixel boundaries which are larger in case of the -100 V bias setting. Along the y direction (Figure 8.40b) the -100 V bias setting shows that the inefficiencies are most pronounced in one half of the pixel: the efficiency is lower for the range in y from $\approx$ 100 μm to $\approx$ 700 μm and then again from $\approx$ 1600 μm to $\approx$ 2200 μm. This behaviour is not yet understood.

Similar to Figure 8.34, Figure 8.41 shows the contribution of the punch-through region to the inefficiencies. Along the fine pixel direction (Figure 8.41a) a clear drop in the center



(a) (b)

FIGURE 8.39: Efficiency along the fine (x) and coarse (y) pixel dimensions for the three bias settings used during the test beam for the first single-MaPSA assembly.



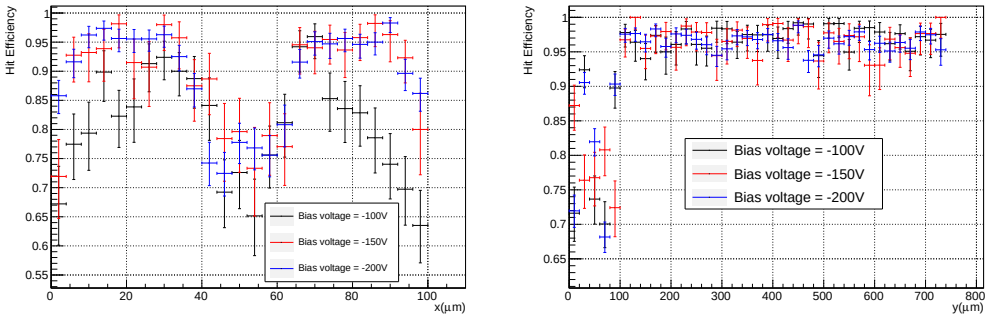
(a) (b)

FIGURE 8.40: Inpixel efficiency along the fine (x) and coarse (y) pixel dimensions for the three bias settings used during the test beam for the first single-MaPSA assembly.

of the pixel is visible for all three bias settings. What is apparent is that for the lowest voltage setting the inefficiencies along the pixel boundaries are more significant compared to the other two bias settings. In Figure 8.41b the contribution of the pixel boundary and the punch-through region to the inefficiency is visible in the first ≈ 100 μm.

Figure 8.42 clearly shows the fact that these inefficiencies along the pixel boundaries are more pronounced in the setting with the lowest bias. The figure shows the average efficiency of the single-MaPSA assembly for the three fiducial regions which were used in Figure 8.35 and described in section 8.5.7.3. When going from fiducial region *case 1* to *case 2*, which excludes the pixel boundaries, we recover the efficiency for the lowest bias voltage setting to values which are comparable to the other two bias voltage settings.

For completeness, the probability for clusters of size 1 and 2 along both pixel dimensions are given in Figure 8.43 and 8.44 respectively. No strong dependence on the bias setting can be observed for the cluster size distributions.



(a) (b)

FIGURE 8.41: Inpixel efficiency along the fine (x) and coarse (y) pixel dimensions, specifically for particles passing the punch-through region in y (LEFT) and in x (RIGHT). This is shown for the three bias settings used during the test beam for the first single-MaPSA assembly.

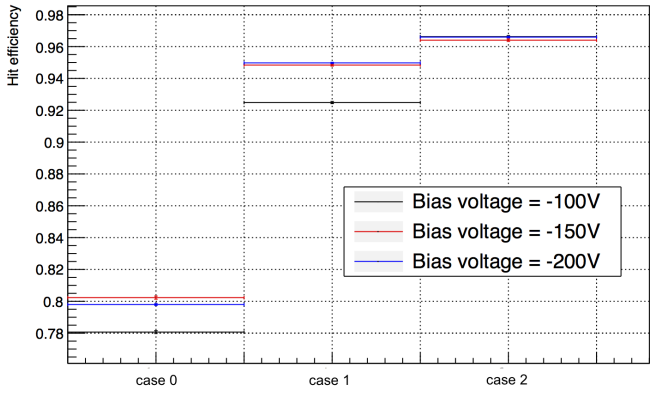
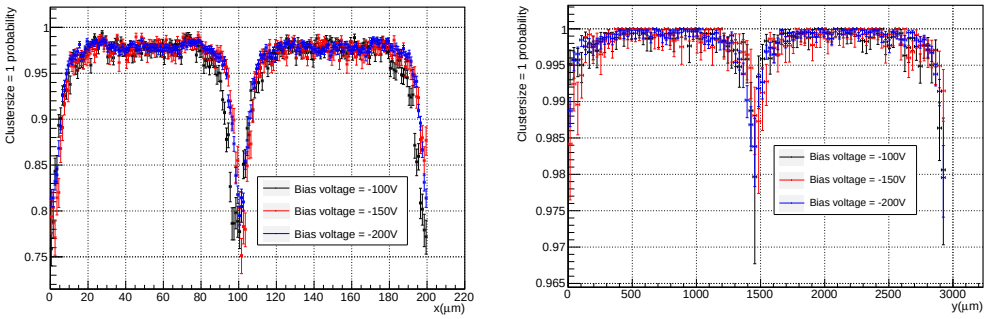
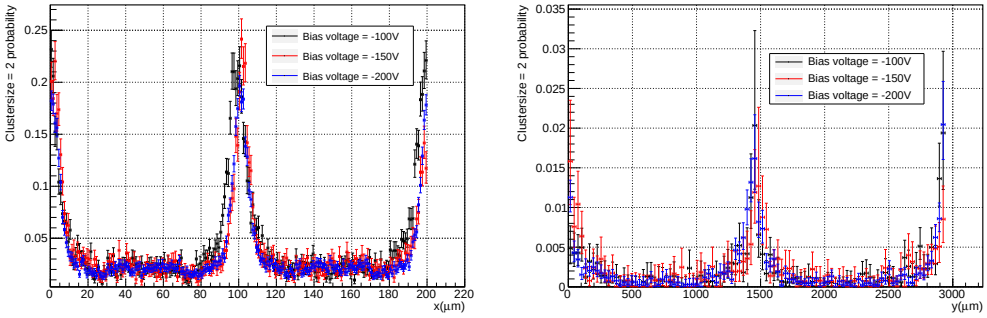


FIGURE 8.42: Average efficiency values for three different bias voltages and three definitions of the fiducial region for the first single-MaPSA assembly. The definitions of the fiducial regions are explained in the text.



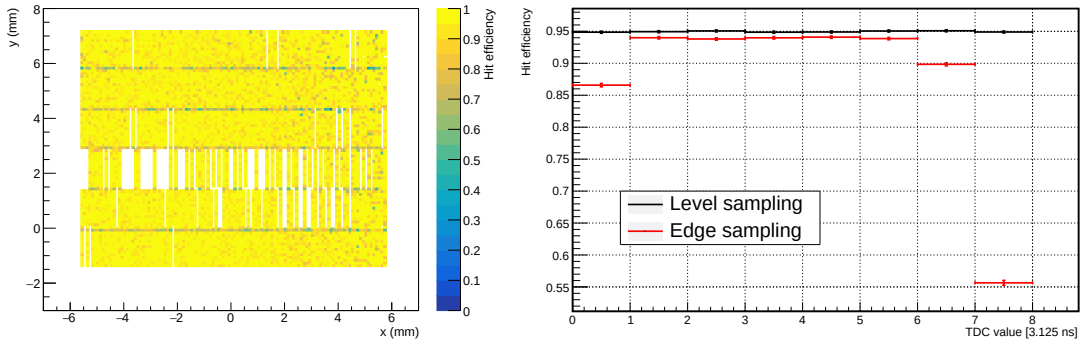
(a) (b)

FIGURE 8.43: Inpixel probability for clusters of size 1 along the fine (x) and coarse (y) pixel dimensions for the three bias settings used during the test beam for the first single-MaPSA assembly.



(a) (b)

FIGURE 8.44: Inpixel probability for clusters of size 2 along the fine (x) and coarse (y) pixel dimensions for the three bias settings used during the test beam for the first single-MaPSA assembly.



(a) (b)

FIGURE 8.45: LEFT: xy efficiency map for the second single-MaPSA assembly. A fiducial region is defined as for the first assembly and dead pixels are masked. RIGHT: Efficiency dependence on the TDC value for the two sampling modes for the second assembly operated at a threshold of 110 LSB.

8.5.7.5 Assembly #2

Figure 8.45a shows the xy global efficiency map for the second assembly obtained at a threshold of 110 LSB. Dead pixels were masked and a fiducial area to mask the edges, similar to Figure 8.25b was applied. There are more dead pixels in this assembly compared to assembly #1 therefore assembly #1 was more extensively tested.

Figure 8.45b shows that in edge sensitive sampling mode there is a strong dependence of the efficiency on the TDC value. For the edge sensitive sampling mode the chip needs to be accurately timed in. There is a plateau of 5 TDC values (which corresponds to $\approx$ 15.5 ns) where the efficiency is highest. The following results are obtained for this efficiency plateau.

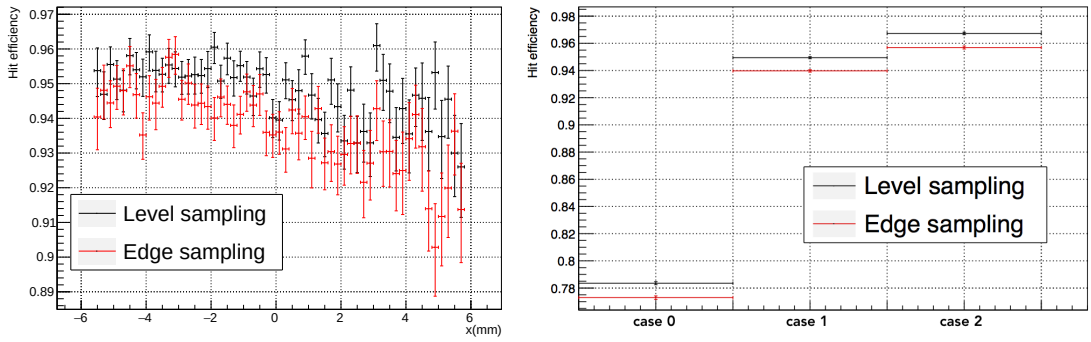
Figure 8.46a shows the global efficiency along the fine pixel dimension for two different sampling modes. A slightly smaller efficiency is observed in the edge sensitive sampling mode. This figure furthermore shows a similar trend as Figure 8.32a: the efficiency drops when moving to higher values of x.

Figure 8.35 and 8.46b give very consistent values for the efficiencies when comparing the data for a threshold equal to 90 and 110 LSB in Figure 8.35 with the data for the level sensitive sampling mode in Figure 8.46b which was obtained at a threshold of 110 LSB. Both measurements were made at a bias voltage of -200 V, at perpendicular incidence in level sensitive sampling mode, but for two different assemblies.

8.6 Summary

The testing of the MPA and SSA chips was performed with the μDTC test bench. Bench top testing was performed to validate both the front-end and digital logic of both chips.

The test system was also used to validate the SEU hard logic which was implemented on these chips. The SEU test was performed by placing the MPA and SSA in a heavy ion beam and measuring the upset rates in the configuration logic, the full-event data path



(a) (b)

FIGURE 8.46: LEFT: Efficiency along the fine pixel dimension for the two sampling modes in which the second assembly was tested. RIGHT: Average efficiency for the two sampling modes in which the second assembly was tested for three definitions of the fiducial region. The definitions of the fiducial regions are given in the text.

and the stub data path. Using the computational method described in Ref. [16], the measurements were used to compute the expected SEU induced upset rates at the HL-LHC. An order of magnitude estimate is obtained for the expected upset rates: $\mathcal{O}(5 \times 10^{-4})$ upsets/s/chip for both the stub and full-event data stream in a single MPA and $\mathcal{O}(10^{-4})$ upsets/s/chip for both the stub and full-event data stream in a single SSA. For the full PS part of the Outer Tracker, the upset rate due to SEUs in the MPA or SSA is then estimated as 1 bit every 2.6×10^6 collisions in the stub data stream and 1 upset bit every 60,000th triggered full-event data package.

An MPA chip was equipped for a first time with a baby PS-p sensor to form a single-MaPSA assembly. This assembly was tested in a particle beam which allowed for qualification of the sensor, ASIC and DAQ. The results presented in this chapter show, amongst others, efficiency measurements which demonstrate that the major inefficiencies are due to the presence of bias and punch-through structures on the sensor. These inefficiencies can be alleviated in the final operation by tilting the modules so that the majority of the particles do not have a perpendicular incidence. Further testing, with new sensor layouts, has also been performed [18].

The μDTC test bench, exercised here to perform the stand-alone testing of the MPA and the SSA, can also be used in the future to test more advanced objects such as PS front-end hybrids and MPAs integrated on a MaPSA.

8.7 Author's contribution

The author intensively contributed to the firmware blocks (e.g. generation of input data for the ASICs) which were specifically developed for initial testing of the MPA and SSA ASICs. This firmware was tested using an emulator developed by the author. The author then worked together with the chip developers to do the qualification of the first MPA

and SSA ASICs. The μDTC allowed for testing of most of the MPA's and SSA's functionality in a bench-top set-up as described in section 8.3. The specific firmware blocks for the MPA and SSA SEU test were largely developed by the author and the author participated in running the tests and eventually performed the analysis of the data to estimate the expected SEU induced upset rate in MPA and SSA data streams in the HL-LHC environment. For operation during test beams, a more DAQ-apt firmware was required and the author contributed to this by developing the MPA and SSA specific firmware blocks in the μDTC's Phy Layer. Several test beams with single-MaPSA devices have been operated in the meantime and the author participated to the first of these. In this test beam, the sensor, ASIC and readout system were all operated for a first time in a beam. The analysis of the test beam data, performed by the author, was presented in section 8.5.

Chapter 9

CIC1 Related Testing

9.1 Introduction

After having discussed tests for the CBC, MPA and SSA chip in the previous two chapters, the testing of the CIC will be the topic of this chapter. The CIC is different from the front-end chips as it is not a readout chip, but can be seen rather as a *data hub*. This chapter will discuss results on testing of the CIC as a stand-alone object (section 9.2) and a first-ever test of the CIC with real front-ends (section 9.3). The stand-alone tests start with the validation of the CIC readout blocks developed for the μDTC. The readout blocks and the validation are thoroughly described in section 9.2.1-9.2.6. These tests were performed short after the first availability of CIC1 chips (February 2019).

In the final modules the output data of two CIC ASICs will go to an lpGBT chip and the data from the lpGBT will then be decoded in the lpGBT-FPGA firmware block. In this way the lpGBT-FPGA makes the GBT link essentially transparent to the user. As a consequence the firmware blocks developed for the electrical readout of the CIC can also be used for the optical readout. The electrical readout test of the CIC is therefore close to an optical readout of a p_{T} module, what concerns the OT specific firmware blocks which need to be developed. This configuration is therefore also representative to what the DTC will need to deal with. A good understanding of the FPGA resource utilization for handling CIC data can therefore be relevant for the design of the DTC. Section 9.2.7 shows how the framework, set up for testing the readout blocks, is used in a TID (total ionizing dose) test of the CIC and section 9.3 then describes the operation of the CIC with 8CBC3.1 chips.

9.2 CIC stand-alone testing

9.2.1 Introduction

The CIC chip is designed to receive data from either 8 MPA chips or 8 CBC chips. Very roughly, as a reminder, the CIC performs the following tasks:

- Phase and word alignment of the incoming data.
- Buffering full-event data from the eight front-end chips and outputting the full-event data over 1 serial line.
- Selecting stub data based on bend and outputting this data over 5 or 6 serial lines.

Before actually connecting the CIC to real front-ends, it needs to be tested as a stand-alone object to validate its operation and the correctness of the μDTC decoder blocks developed for handling CIC data. Having a test bench where a CIC can be tested stand-alone has a few major advantages:

- One has full control over the input data as this data can be stored on FPGA and *played* to the CIC chip. Furthermore the input can be changed on a BX-by-BX basis, as opposed to a configuration where the input data is coming from real CBCs or MPAs where the test patterns can only be configured by changing the CBC or MPA configuration over I^2C, which does not allow for BX-by-BX control.
- The CIC chip can be put on a dedicated carrier card, therefore eliminating the overhead of testing the CIC chip on a complex carrier such as a front-end hybrid.

A test bench like this furthermore allows for doing CIC specific studies such as irradiation tests. Also during construction of the Outer Tracker, CICs will still need to be tested as stand-alone objects. Indeed, on the PS front-end hybrids (as described in section 8.2.4) the CIC can be seen as a stand-alone object as it is not connected yet to the MPA ASICs.

Compared to the MPA, SSA and CBC chip, the CIC testing is different as the CIC is a mostly digital chip which actually makes data input from an FPGA to the CIC straightforward. Logic blocks were developed for the μDTC Phy layer to decode both the stub data and the full-event data in order to build an event in the Data Readout Block. The end goal of the CIC stand-alone testing, described in section 9.2.2 until 9.2.6, is to verify these firmware blocks and is very straightforward: the data which is played to the CIC chip should also be the data which is output by the CIC decoding blocks in firmware, modulo the data which the CIC is supposed to drop. The tests described in these sections are devoted to testing the CIC stub data decoding blocks on the firmware rather than focussing on testing the functionalities of the CIC1 chip. Implicitly the CIC1 is of course benchmarked against its validation model.

To summarise, in order to perform these tests the following ingredients are required:

- A reference model (section 9.2.2): the input data to this reference model comes from an emulated CBC chip or from the model of the MPA + SSA chip. This data functions as input to the CIC1 model of which the output is used to compare to the output of the CIC data decoder blocks in the μDTC.
- Front-end data player (section 9.2.3): a firmware block which stores data from the reference model and plays it to the CIC chip.
- A full CIC1 hardware set-up (section 9.2.6).
- CIC FW decoder blocks for full-event and stub data (section 9.2.4):
  - Full-event data: the full-event data needs to be aggregated and stored for readout.
  - Stub data: the CIC aggregates stubs from 8 consecutive BXs and orders the stubs on bend. As a result the data is not ordered in time when it goes out of the CIC. In order to match the full-event with the correct stub data the firmware block needs to reorder the stub data so it is ordered in time.
- A DQM (data quality monitoring) tool (section 9.2.5) to compare the output from the CIC decoder to the output of the CIC model.

Results from the testing are discussed in section 9.2.6 and how this framework was used to evaluate the CIC's TID sensitivity is described in section 9.2.7.

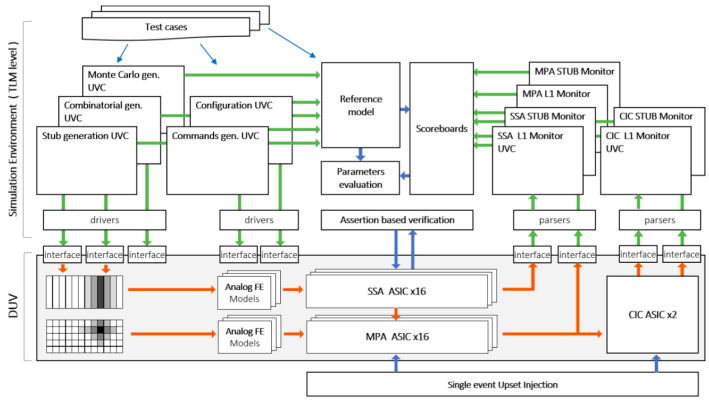


FIGURE 9.1: Schematic of the front-end chip validation framework [21].

9.2.2 Reference model

A framework for the validation of the CBC+CIC1 and SSA+MPA+CIC1 systems is developed at CERN [21] and illustrated in Figure 9.1 for the PS case. The main purposes of this framework are: evaluate design choices for the ASICs, show that each chip's RTL[1] (register-transfer level) implementation is flawless, validate the interfacing of different ASIC types and investigate data losses and bandwidth limitations of the system.

The validation framework can be stimulated by two different types of generators: random generated stimuli or stimuli extracted from physics MC simulations of the Outer Tracker. The signal injection happens at the front-end level where a 32-bit signal represents the charge injected from the pixel or the strip to the front-end. The validation framework can thus also be used to study the effect of time of arrival, time walk and other front-end related causes of inefficiency. The random generator stimuli can be set up to generate runs with different configurable parameters. Examples of these parameters are: hit occupancy, hit distribution across the sensor, transverse momentum distribution,... Another generator takes care of handling the generation of clock and fast command signals. The hit stimuli are injected in the chip front-ends on one hand and into a reference model on the other hand. The reference model is a high level description of the CMS Outer Tracker readout functionalities and does not implement any inefficiencies. In the end the output of the reference model and the full ASIC chain are compared in the scoreboard which can be used to investigate inefficiencies in the ASIC chain.

What is important for the CIC stand-alone testing is the digital part of the simulation framework which is used to provide reference data. The validation framework offers the data streams pre- and post-CIC for both stub and full-event data. These are the two data streams which are required for the testing of the CIC data decoder blocks in firmware: the input data to the CIC which has to be played from the front-end data player and the output data of the CIC to check that what the firmware decoder blocks deliver, corresponds to the output data from the CIC model.

[1]Abstraction for the modelling of digital logic. The modelling is performed in terms of the flow of digital signals and the logic operations performed on these signals.

9.2.3 Front-end data player

The CBC or MPA data from the reference model needs to be loaded on the FC7 card and played to the CIC chip. The data needs to be stored on-board as online playing of data from the DAQ machine to the FC7 is not possible due to the limited bandwidth between the FC7 and the DAQ computer: the data input to the CIC is 8(chips)×6(data-lines/chip)× 320Mbps + 1(t1 line)× 160 Mbps, which results in a total of 15.52 Gbps, and would require a dedicated link from DAQ machine to FC7.

The FPGA which plays the data to the CIC also receives the data back from the CIC. This implies that the DAQ firmware blocks need to be present on the same FPGA as the front-end data player block. Given the fact that the DAQ uses the external DDR3 memory on the FC7 board, the front-end data player has to use the internal FPGA BRAM resources to store the data which is to be played. In this way the data playing interferes as little as possible with the DAQ firmware blocks and represents a test as closest as possible to the final configuration.

Each entry in the BRAM contains one BX of data:

- 4 bits of fast command payload (the header and trailer is sent by default)
- 8 bits (1 word) × 5 (5 stub lines/front-end chip) × 8 (8 front-end chips) = 320 bits of stub data
- 8 bits (1 word) × 1 (1 full-event line/front-end chip) × 8 (8 front-end chips) = 64 bits of full-event data

summing up to 388 bits of data to be sent every BX. The fast command needs to be sent by the data player as proper timing is mandatory in this application.

Before sending the actual data to the CIC chip the CIC needs to phase and word align each line of incoming data. This is referred to as the *CIC start-up sequence*. In order to perform a test with the stand-alone CIC which is representative of the scenario with real front-ends the front-end data player mimics this start-up sequence as close as possible. For this reason the state machine shown in Figure 9.2, where the specific case of the front-end data player emulating CBC chips is shown, is adopted.

A full overview of the start-up sequence can be found in Appendix B. There are basically five different main stages in this synchronisation exercise to set up the CIC and DAQ for data taking:

1. Locking of the fast command: the CIC can be configured to sample the fast command on the rising or falling edge of the 320 MHz clock.
2. Phase alignment of all the CIC data input lines: a '...1010...' repetitive pattern needs to be sent on each CIC data input line in order for the CIC to find, for each input line, the correct phase at which to sample. When a start is sent from software the front-end data player starts to send these phase alignment patterns. The state machine goes to the next state when the software reads from the dedicated CIC I^2C register that the phase alignment has locked on all the data lines driven with a phase alignment pattern. In reality the MPA can be configured in a mode where the data on the lines is the output of a configurable 8-bit deep shift register. Therefore for the CIC phase alignment, when MPAs are connected, the phase alignment pattern is a periodic pattern of '10101010' on all lines. In the CBC case such a shift register is not present and the chip therefore has to be put in a special configuration to generate

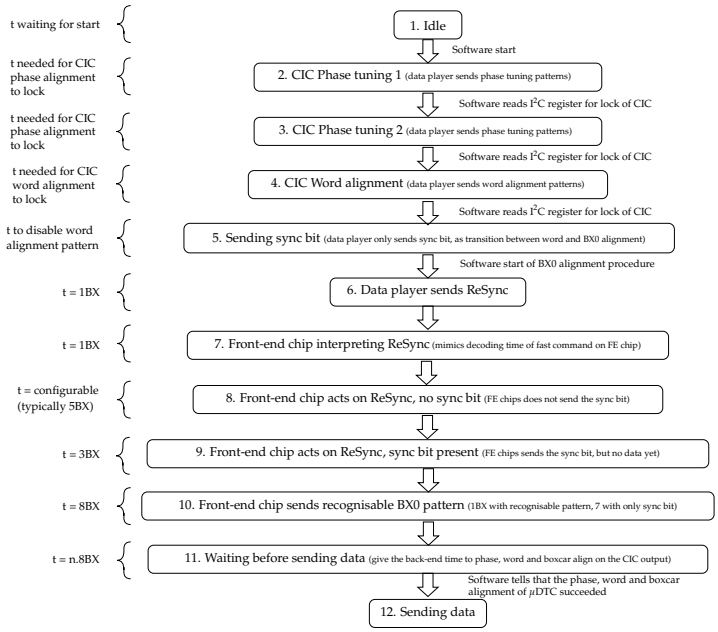


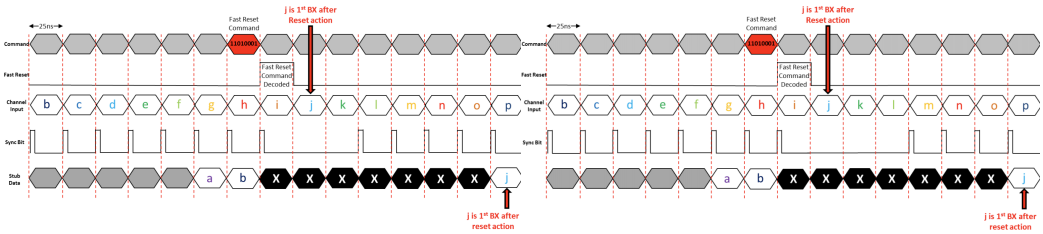
FIGURE 9.2: State diagram of the front-end data player for testing of the CIC chip for the case where the data player emulates the configuration in which CBCs would be connected to the CIC.

a fixed pattern, usable for phase alignment, on the output lines. The phase alignment patterns chosen in case CBCs are connected are summarised in Table 9.1. The phase alignment in case of the CBC proceeds in 2 separate steps. First stub line 1, 2 and 4 are phase aligned after which line 3 and 5 are phase aligned[2]. These patterns are mimicked in the data player where both a CBC and an MPA configuration are available. It has to be noted here that the data player does not mimic exactly the phase tuning of the full-event data line in the CBC mode. The data player generates a repetitive '...101010...' pattern on the full-event data lines, but it is not possible to generate this pattern with the real CBC. How this is performed with the real CBC chip is discussed in section 9.3.

3. Word alignment of the CIC stub data input lines: after locking of the phase alignment is obtained the word alignment patterns are sent. The CIC uses these patterns to find the correct consecutive 8-bit patterns. The MPA word alignment pattern can again be generated by the shift register and is the same for all lines: '1101 0000'. The alignment pattern for the CBC is different for each line and the respective patterns are also given in Table 9.1[3]. The front-end data player again mimics these real life configurations. After the software reads from the CIC that the word alignment has locked, the software tells the front-end data player state machine to proceed to the next state.

[2]Where line 5 is the line carrying the synchronisation bit.

[3]Note that these are not the only patterns which can be used for CBC word alignment. Which word alignment pattern the CIC has to look for is configurable for each stub data line, but the patterns are the same for each front-end chip.



(a) (b)

FIGURE 9.3: ReSync behaviour of the CBC3.1 [68]: after reception of a ReSync the CBC chip takes 1 clock cycle to decode this command to which it responds by not outputting the sync bit for 2 clock cycles for low DLL settings (LEFT) and 3 clock cycles for high DLL settings (RIGHT). BX j is the first BX where the data from the CBC is valid after the ReSync.

4. BX0 alignment: after a ReSync is sent the CIC starts to look for a pattern[4] on one of the stub lines from one of the front-end chips[5]. In this way the CIC can find the BX0, which is the first BX after a ReSync where the CIC will interpret the data from the front-end chips as valid data. This BX0 value is deterministic. For the CBC we know for example that after a ReSync is decoded on the CBC, the valid data will appear 7 clock cycles later on the CBC output (see Figure 9.3). However, due to additional delays on the hybrid or delays introduced by the CIC phase and word alignment, the timing of the first valid data after a ReSync might not be the expected one. The only way to confirm this is by letting the front-end chip output a certain pattern in clock cycle j (see Figure 9.3). Once this extra delay is known for a given hybrid, this should not change during operation. For the CBC the generation of this pattern can be done with the test pulse logic (for more info see section 9.3). When running the start-up sequence from the data player, also a BX0 alignment is performed and the behaviour from the front-end chips is mimicked as close as possible. As illustrated in Figure 9.3, the CBC for example does not send the sync bit for 2 or 3 clock cycles, depending on the chip's DLL setting, after reception of a ReSync. This behaviour is emulated also in the data player.

5. Alignment of the back-end: the CIC can be configured to send a repetitive pattern on the output data lines. The pattern is '1110 1010' every 8th BX followed by 7 BXs of '1010 1010'. Using this pattern the back-end can find: the correct phase to sample the data, the beginning of a word and also the beginning of a stub boxcar.

9.2.4 CIC decoder blocks

9.2.4.1 Full-event data

The full-event data processing in the firmware is fairly straightforward. Three different data streams must be handled:

- CBC unsparsified data

[4]The pattern can be configured by the user.
[5]Which chip and which line is configurable.

	stub line 1	stub line 2	stub line 3	stub line 4	stub line 5 (with sync bit)
phase alignment 1	0b0101 0101 location 1 = 0d85	0b1010 1010 location 2 = 0d170	0b0000 0000 location 3 = 0d00	0b1010 1010 bend code 2 = 0d10 bend code 1 = 0d10	0b1010 0000 error bit = 0 OR254 bit = 1 overflow bit = 0 bend code 3 = 0d00
phase alignment 2	0b0011 0010 location 1 = 0d50	0b0110 0100 location 2 = 0d100	0b1010 1010 location 3 = 0d170	0b1010 1010 bend code 2 = 0d10 bend code 1 = 0d10	0b1010 1010 error bit =0 OR254 bit = 1 overflow bit = 0 bend code 3 = 0d10
word alignment	0b0111 1010 location 1 = 0d122	0b1011 1100 location 2 = 0d188	0b1101 0100 location 3 = 0d212	0b0011 0001 bend code 2 = 0d03 bend code 1 = 0d01	0b1010 0001 error bit = 0 OR254 bit = 1 overflow bit = 0 bend code 3 = 0d01
BX0 alignment	0b0001 0010 location 1 = 0d18	0b0010 0010 location 2 = 0d34	0b0100 0010 location 3 = 0d66	0b0010 0010 bend code 2 = 0d02 bend code 1 = 0d02	0b1000 0010 error bit = 0 OR254 bit = 0 overflow bit = 0 bend code 3 = 0d02

TABLE 9.1: Table summarising the phase and word alignment patterns for CBC+CIC. The highlighted cells show in which state the CIC is locking on which line. For the BX0 alignment, in the configuration shown here, data is generated on all five lines so that the user can choose which of the five lines to use. The actual payload is also shown in decimal format and decoded to show that these configurations are indeed possible with the CBC: the addresses on the stub lines increment with line number, bends can be programmed by offsetting the masked channels, the OR254 bit will be high when enabled due to the presence of data, the overflow bit is low because in all configurations exactly 3 stubs are programmed and the error bit is zero. In case the latter one would cause trouble a fast reset can clear it. To program the CBC in these states the values of the register settings are given in Appendix C. It should be noted here that the above table gives one of the configurations to run the CBC+CIC start-up sequence. Other configurations are possible.

- CBC sparsified data
- MPA sparsified data

The handling of the full-event data stream is very similar to how the data is handled in case of MPA readout as explained in section 6.5.2.4 and briefly explained again here: a state machine looks for the full-event data header and when it is found transitions to a state where it buffers the data. In case of the sparsified stream the information in the header is interpreted and the state machine handles the acquisition of the remainder of the package accordingly after which it sends the data out to the Data Readout Block over a similar bus as described in section 6.5.2.4.

9.2.4.2 Stub data

As explained in section 5.5.6.3 the CIC aggregates the stub data from the front-end chips over 8 BXs and selects from these stubs the ones with the lowest bend to send to the back-end. Handling CIC stub data is therefore fairly different from handling stub data from the readout chips:

- The firmware needs to buffer the 8 BXs long CIC boxcar before being able to process it.
- The firmware needs to resort the payload on the offset value, available for each stub in the payload, which encodes the time offset of the stub within the 8 BX long boxcar.

For the latter item a bitonic sorting network was implemented which sorts the stubs within one boxcar based on the offset value. The concept of a bitonic sorting network is illustrated in Figure 9.4. This sorting algorithm is a commonly used technique for sorting data in parallel: in Figure 9.4 every arrow represents a comparator stage which simply makes a decision on which of the two inputs is largest. The largest input goes to the end of the arrow. By pipelining several of these blocks, the data is sorted after a given amount of clock cycles. The number of clock cycles (n_{stages}) required to do the sorting is dependent on the number of inputs n (which here is equal to the number of stubs to sort) and is given by:

$$n_{\text{stages}} = \log_2(n)\frac{\log_2(n)+1}{2}. \tag{9.1}$$

At the output of this block the stubs will be sorted on the offset value. In addition it is required to pass the stub data on to the readout block separated by BX. If the full CIC stub package would be input to the bitonic sorter, the output package would still contain data from all 8 BXs. So somehow the now sorted data needs to be output BX-by-BX. This sequence of resorting the data and separating it on BX is referred to as *boxcar unpacking*.

Besides boxcar unpacking the data, the data also needs to go to the readout block right aligned[6]. All of this can be accomplished in two ways, both based on the bitonic sorting: the first one will be referred to as the *hot-bit-map + bitonic-sorter* and the second one as the *bitonic-sorter+shifter* approach. These approaches are described in the next two sections.

[6] Aligned towards the LSB end of the bus.

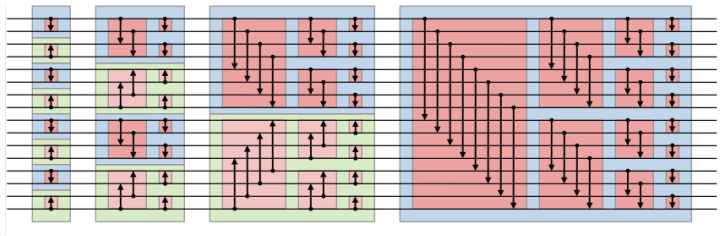


FIGURE 9.4: Illustration of a bitonic sorting network with 16 inputs. The data will reside in each stage for 1 clock cycle, after which it is passed on to the next stage (in this example there are 10 stages). The arrows represent comparator stages. The coloured blocks represent entities which can be implemented in separate logic blocks [87].

9.2.4.2.1 Hot-bit-map + bitonic-sorter approach

It is possible to get the required boxcar unpacking and right alignment by feeding a single CIC stub package in 8 consecutive clock cycles to the same sorting network. In clock cycle 0 the data input to the sorter are the stubs with offset value 0, in clock cycle 1 the stubs with offset value 1 and so on until the stubs with offset value 7. At the output of the bitonic sorter there will be 8 consecutive clock cycles of right aligned stub data, which are ordered in time. This approach requires of course some additional logic to create the *hot-bit-map* which decides on which stubs will be input to the sorter at a given clock cycle. Next to the additional logic for creating the hot-bit map, this approach also occupies the logic of the bitonic sorter for 8 consecutive clock cycles, which reduces the number of CICs which could be handled by the same sorter. Assuming that the sorter stage could run at 240 MHz this would allow for sharing of the sorter network over 6 CICs.

9.2.4.2.2 Bitonic-sorter + shifter approach

Instead of inputting the data to the bitonic sorter BX-by-BX by using the hot-bit-map approach as described in the previous paragraph, it would also be possible to feed the full boxcar to the bitonic sorter and to get on output the package sorted on offset values. The drawback of this approach is that the data does not come out separated in BX, neither does it come out right aligned. It is thus required to add an additional stage after the bitonic sorting. This stage, referred to as the *shifter* stage, is a pipelined approach for outputting the data separated on BX and right aligned. The shifting operation is illustrated in Figure 9.5. The clock cycle = -1 represents an empty shifter network with a stub package with 16 stubs, sorted on offset value by the bitonic sorter, coming in. The colours make it possible to follow the different packages throughout the different clock cycles and the numbers represent the offset value. Stubs with offset value 0 are output in the shifter stage = 0, stubs with offset value 1 are output in stage = 1, and so on. In this way the shifter stage can handle a continuous stream of stub packages.

This approach requires a lot of multiplexing resources on the FPGA as for every level in the shifter the number of stubs to output can range from 0 to the maximal number of stubs. The advantage compared to simply doing a similar approach on the raw unsorted boxcar is explained by an example: when looking at the yellow stub package in for example clock cycle 1, two extreme cases are possible: either the payload does not contain any stub with offset 1 or it contains stubs which all have offset 1. The example in Figure 9.5

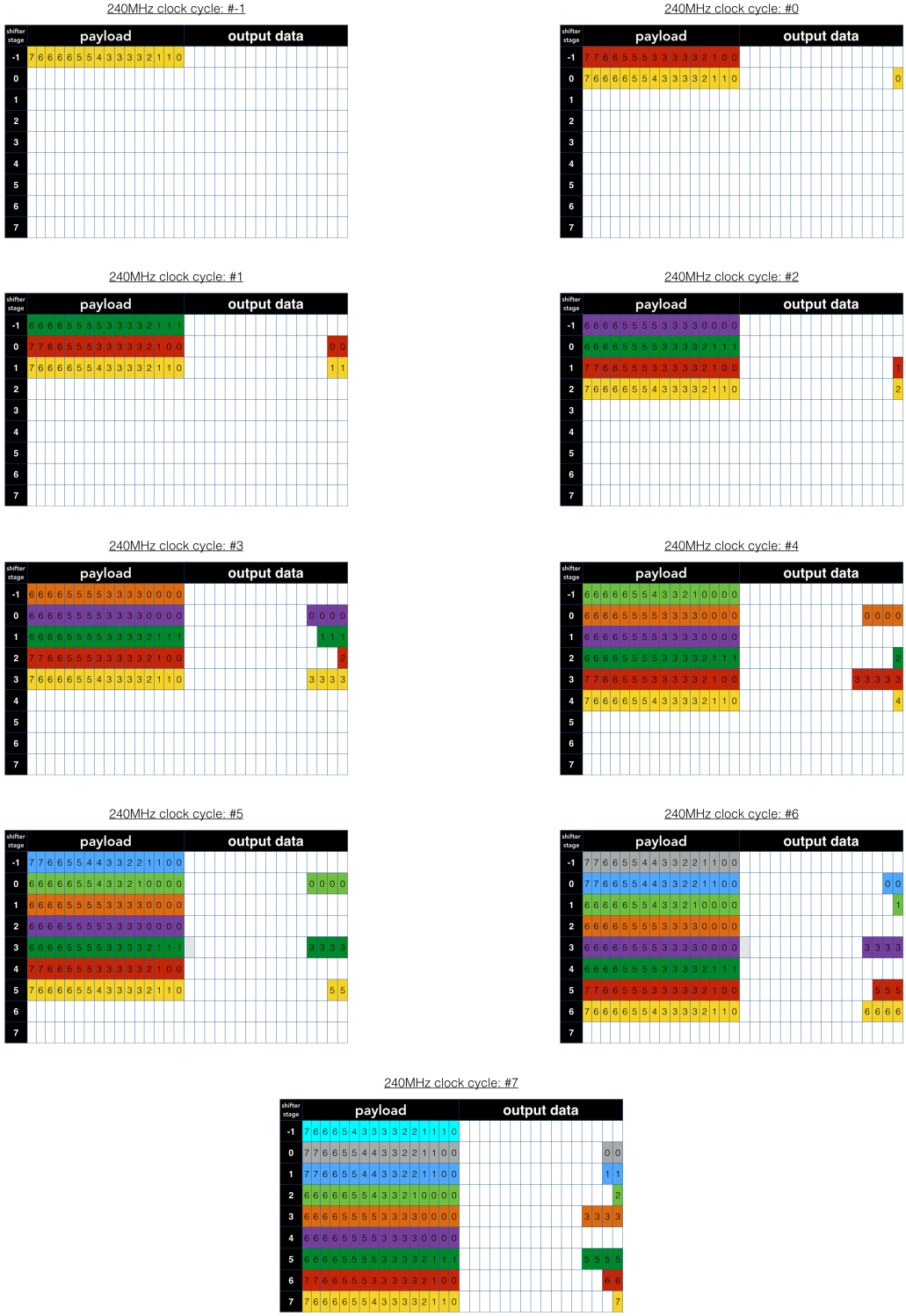


FIGURE 9.5: Representation of the shifter stage for 9 consecutive clock cycles. The colour code represents the different CICs and the numbers in the *payload* and *output data* represent the offset value of the stub within the boxcar.

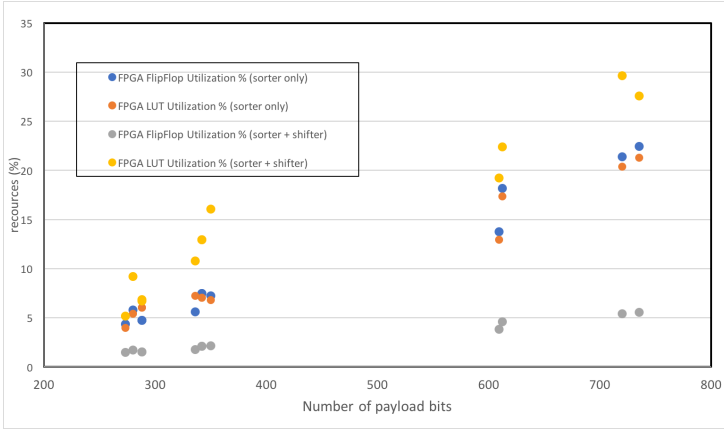


FIGURE 9.6: Resource utilization of two approaches of the CIC stub decoder block for a Kintex7 XC7K420T FPGA. The data points of the 2S with bend info and PS case without bend info for both the FEC12 and FEC5 320 Mbps mode overlap as the number of payload bits, and as a result the resource utilization, is identical.

shows an intermediate case where there are two stubs with offset 1. In firmware the logic has to be present anyway for outputting all 16 stubs, but the multiplexing is limited by the fact that the rightmost entry on the *output data* side can receive 16 stubs, whilst the one next to it can only receive data from 15 stubs and so on. This reduces the logic utilisation. Results on the logic utilisation can be found in section 9.2.4.2.3. The advantage of this approach is that the logic for the sorting and the shifting can be shared over many more CICs than in the *Hot-bit-map + bitonic-sorter approach* approach. When again assuming that both stages can run at 240 MHz this logic could be shared over 48 CICs.

9.2.4.2.3 Logic utilisation of the CIC stub decoder

The final goal of the μDTC is to be able to read out 8 modules with a single FC7. It is therefore relevant to further investigate the resource utilisation of this block, also in view of using this stub decoder block in the final DTC.

The flip-flop and LUT resource utilization for the two approaches for stub data decoding is shown in Figure 9.6 for the Kintex7 XC7K420T FPGA. The resource estimate is shown for the eight possible CIC1 configurations (see Figure 5.3) and also for the four 640 MHz modes which will be available in CIC2. The resource utilization is given as function of the total number of stub payload bits in the CIC boxcar for the example of reading out 48 CIC chips which would require 8 *Hot-bit-map + bitonic-sorter* blocks and 1 *bitonic-sorter + shifter* block, assuming all blocks run at 240 MHz.

It can be seen that the *Bitonic-sorter + shifter* approach has a higher LUT utilization for all the configurations. The flip-flop utilization is however much reduced in the *Bitonic-sorter + shifter* approach. The latter can be assigned to the fact that the *Hot-bit-map + bitonic-sorter* approach requires 8 such bitonic sorters, compared to only one for the *Bitonic-sorter + shifter* approach.

To conclude, the two approaches have the following advantages and disadvantages:

- Given that both stages can be clocked at 240 MHz, a single *Bitonic-sorter + shifter* approach can be used for 48 CICs. A single *Hot-bit-map + bitonic-sorter* approach can be used for 6 CICs. In case less than 48 CICs would be connected to the DAQ, resources would not be used in the *Bitonic-sorter + shifter* approach. In that sense the *Hot-bit-map + bitonic-sorter* is more suitable for scaling down.
- In the case of 48 CICs the *Hot-bit-map + bitonic-sorter* approach has a lower LUT utilization, but a higher flip-flop utilization. What is most attractive must be decided by looking at the overall resource utilization of the full μDTC. It should be noted that these flip-flops and LUTs are contained in the same slice on the FPGA, so in case one wants to exploit the smaller flip-flop utilization of the *Bitonic-sorter + shifter* approach one has to remember that slice resources would have to be shared between different blocks, which would complicate the routing for the compilation tool.

What is not included yet in this estimate is the fan-in (from the ISERDESes) and fan-out (to the Data Readout Blocks) of the data to and from the sorting networks. Similar resources will however be needed in both approaches to accomplish this.

For now the *Hot-bit-map + bitonic-sorter* approach was implemented in the μDTC as this is the most applicable for reading out the limited number of modules that will be connected to a single FC7.

9.2.4.2.4 Test bench

Given the fact that the stub decoder block is a completely new and rather complicated block, this block was carefully reviewed in simulation. Data from the CIC1 reference model was fed to the CIC1 stub decoder block using a Vivado test bench and the output data was checked against the input data in software. In this way the decoder block was verified before testing with real hardware.

9.2.5 Data quality monitoring

A dedicated stand-alone and small scale DQM was set up to mainly compare the data output of the CIC1 validation model against the output from the μDTC CIC data decoder blocks. As a starting point the input and output CIC1 data, delivered by the validation model, were compared to get acquainted with the data streams. Assessing the input data to the CIC also gives a better feeling on how the CIC is stressed (trigger rate, stub occupancy,...). Investigating the data streams unveiled three problems in the ASIC designs:

- When the CIC1 full-event output FIFO is full, due to a high trigger rate, the FIFO is not properly reset for the subsequent events, therefore subsequent full-event data packages are lost. This will be fixed in CIC2.
- The front-end error bits are not correctly timed in into the CIC1 stub boxcar: the CBC error bits are reported in the previous CIC stub boxcar's header if stubs have a time offset of 0 or 1 within the boxcar. This will be fixed in CIC2. For an offset starting at 2 the front-end status bits are reported in the same boxcar as the actual stub.
- The MPA full-event frame extends, with meaningless data, beyond the trailer bit. This is due to an error in the MPA RTL model and will be fixed for the next submission of the chip. The full-event frame itself is however not affected by this bug.

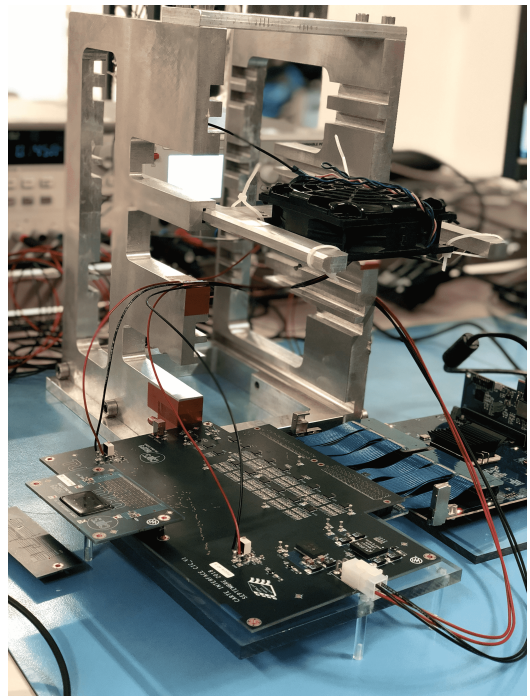


FIGURE 9.7: Test bench of the stand-alone CIC chip. The CIC1 mounted on its carrier card, on the left in the figure, is connected through a single FMC connector to the CIC interface board which connects to the 2 FMC connectors on the FC7.

9.2.6 Data processing test results

The CIC1 stand-alone test set-up is shown in Figure 9.7. It shows the FC7 and the CIC on a dedicated carrier card connected to an interface board for level translation and power regulation.

For the purpose of testing the stub data path in the CIC1 and the FW decoder blocks the data output from the firmware stub decoder blocks is stored in the DDR3 memory on the FC7 for a full test run. This approach has the advantage that the full continuous data stream is available: every BX of data can be accessed. This is not the case in the standard μDTC readout scheme where the stub data is available only on a sample basis, where the sampling rate is equal to the L1A trigger rate. This continuous test speeds up the testing and more importantly is a requirement, as in the final system the full continuous stub data stream will be used in the L1 trigger system.

In the end the goal of the testing is very straightforward: what comes out of the CIC decoder blocks should be matching what comes out of the CIC1 validation framework. In this way the CIC1 is benchmarked against its RTL model and the FW decoder blocks are validated. Figure 9.8 shows the FW blocks which are needed for CIC readout and emphasises the ones which are used during the testing of the stub data flow: the slow control blocks to configure the CIC and to read its status, the data player to feed the CIC with a fast command and a total of 48 stub and full-event data lines at 320 Mbps, the phase and word tuning to align the incoming CIC data, the decoder blocks to do the boxcar unpacking of the stub data, the DDR3 packer to store the stub data in the DDR3

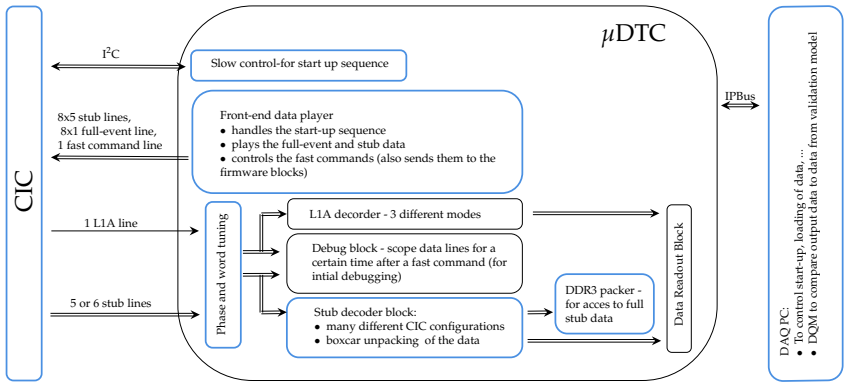


FIGURE 9.8: Illustration of the FW blocks used for the CIC stand-alone testing. The blocks used for the testing of the stub data flow are highlighted.

memory and a software DQM tool to do offline checks on the data.

The objective of the test is rather simple and as a result the visualisation of the results is elementary: all the fields in the CIC boxcar (see Figure 5.26) can be compared to the output data from the validation framework. The parameters which thus need to be checked are: the configuration bit, the front-end and CIC status bits, the stub front-end index, address, bend and in case of the PS data also the z/r coordinate. For one typical run these parameters are compared for $\mathcal{O}$(1k) BXs. The output of the DQM is shown in the scatter plots in Figure 9.9 for the MPA-CIC1 configuration using 5 lines for stub data with bend information included. The plots are simply added to give some visual feedback, and are a tool for debugging, but the exact matching is also checked in the software code. Note that checking the consistency between validation and output of the decoders on a BX-by-BX basis implicitly checks the correctness of the boxcar unpacking of the data based on the stub offset field within the CIC boxcar. This consistency check was performed for all eight 320 MHz CIC1+decoder configurations and 100% matching between validation model on one hand and CIC1+decoder blocks on the other hand was attained.

9.2.7 Irradiation testing

9.2.7.1 EP-ESE irradiation system

The X-ray irradiation system, ObeliX, at the CERN EP-ESE was used to conduct a TID test of the CIC1 chip. The X-ray tube uses a tungsten target to produce the X-rays. Dose rates of 90 kGy(SiO_2)/hour can be attained. Dose calibrations are done with a pin diode and a systematic uncertainty smaller than 30% is expected. This irradiation test was performed in June 2019.

9.2.7.2 Test set-up

The CIC1 was tested with X-ray irradiation up to a total ionising dose (TID) of 2.1 MGy at a dose rate of 88 kGy(SiO_2)/hour. This TID is sufficient for qualification of the CIC1 for HL-LHC conditions where, for an integrated luminosity of 4500 fb^{-1}, a TID of 1 MGy is expected for the PS modules in the most hostile radiation environment. Figure 9.10

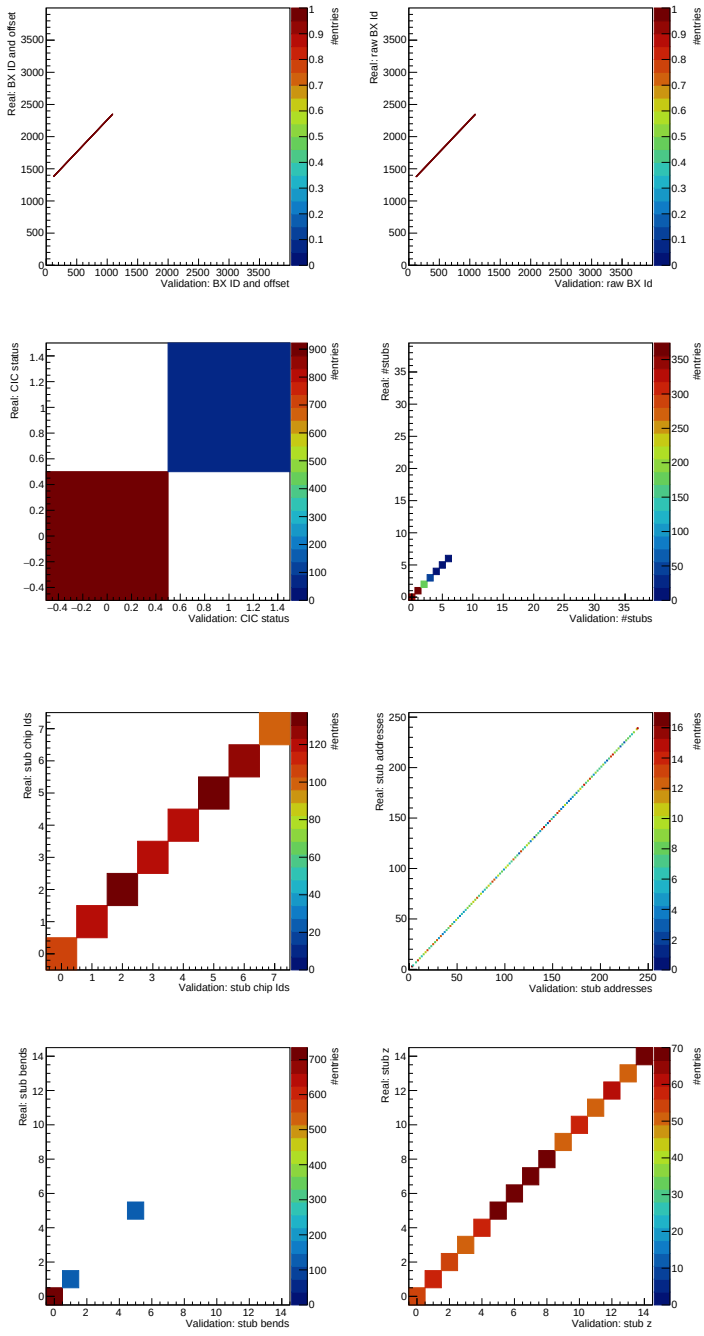


FIGURE 9.9: Visualisation of the consistency between the data from the CIC1 validation model and the data obtained from the CIC and firmware decoder blocks when both were stimulated with the same input data. The test case shown here is the MPA-CIC1, 5 lines, with bend data configuration. TOP LEFT: Correlation between the BX Ids. The firmware takes the BX counter from the CIC (*raw BX* counter, incrementing by eight) and makes it a counter with BX resolution. TOP RIGHT: The raw BXId. SECOND ROW LEFT: correlation of the CIC status bit to the validation model which confirms proper treatment of CIC stub data package header in the firmware. The rest of the figures show the correlation of the payload, from which it can be concluded that the boxcar unpacking on the firmware is properly implemented.

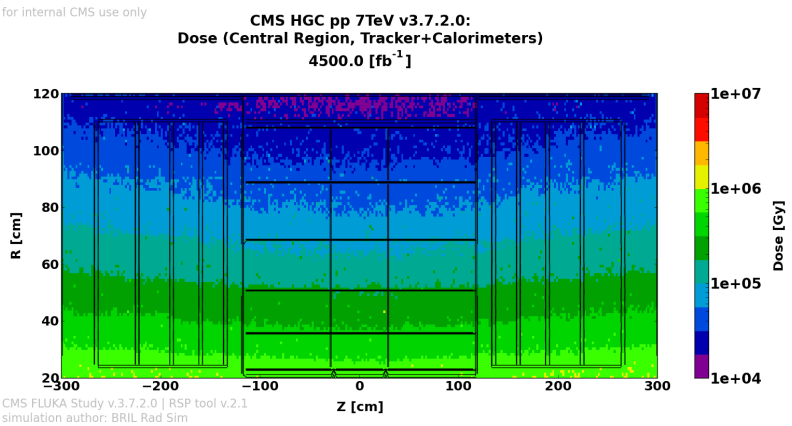


FIGURE 9.10: Integrated dose for the Outer Tracker environment at an integrated luminosity of 4500 fb^{-1} [83].

shows the total expected integrated fluence for the Outer Tracker environment for an integrated luminosity of 4500 fb^{-1}.

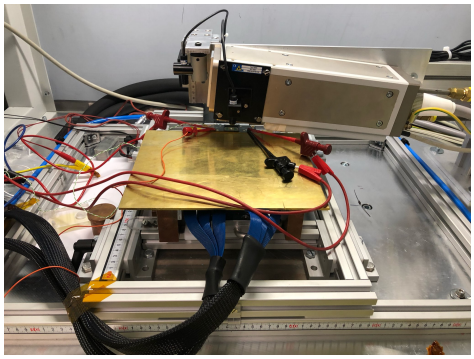
The TID set-up is shown in Figure 9.11a where the interface card is sitting below the shielding. At the far back of the image, under the X-ray tube, the CIC1 can be distinguished with its protecting raisin removed.

The uniformity of the beam, as measured just before the irradiation, is shown in Figure 9.11b where the location of the CIC1 is indicated by the grey rectangle. The dose rate is very uniform over the CIC1 area with a maximum fluctuation of -4.7% from the maximum value. Together with the systematic uncertainty on the absolute dose, which is guaranteed to be smaller than 30%, and the uncertainty on the HL-LHC dose rates from FLUKA simulations, which is below 8%, the TID value of 2.1 MGy has a sufficient safety margin with respect to the 1 MGy of maximal TID expected for the final operation.

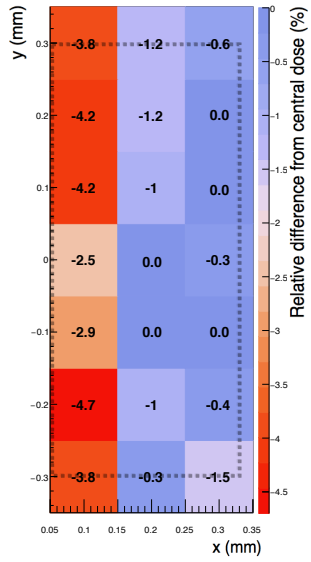
During the irradiation, the same test loop was repeated continuously where the CIC start-up sequence (including phase alignment, word alignment and BX0 alignment) was run as if CBCs were connected after which the CIC1 was stimulated with $\approx$ 1k BXs of stub data. The CIC1 was configured in the CBC-CIC mode outputting stub data on 6 lines with bend information included. For each run the phase alignment, word alignment and BX0 alignment results and the CIC1 output stub data were saved for offline analysis. A single loop of initialisation and data playing takes about 65 seconds, so during the irradiation the loop was run about 1300 times. The currents consumed by the CIC1 core and CIC1 periphery were also monitored continuously. During irradiation the temperature on the CIC was $\approx 40^{\circ}$C.

9.2.7.3 Results

The current consumption with respect to aggregated dose is shown in Figures 9.12a and 9.12b. These show respectively the currents drawn by the core logic (I_{core}) of the CIC1 and the periphery (I_{PST}). The currents were measured in each loop at two specific times: during the phase alignment and during the streaming of the stub data. It can be seen that the CIC1 uses 22 mA (1.8 mA) more current in the phase alignment (stub streaming) stage on I_{core} (I_{PST}). The relative change in current, during stub streaming, is shown with respect to the aggregated dose in Figure 9.13a for both I_{core} and I_{PST}.

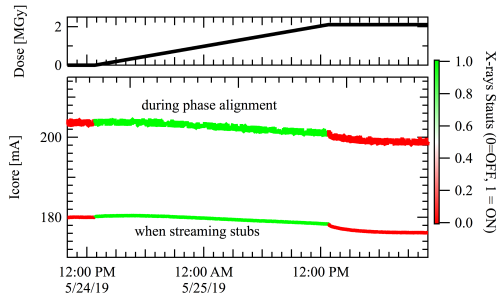


(a)

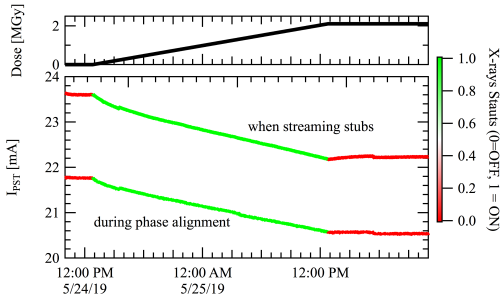


(b)

FIGURE 9.11: LEFT: The CIC1 under the X-ray source in the TID set-up. RIGHT: Dose rate uniformity in the X-ray machine. The CIC1 borders are indicated with the grey rectangle.



(a)



(b)

FIGURE 9.12: Currents on I_{core} (LEFT) and I_{PST} (RIGHT) pre, during and post irradiation for two specific moments during the test routine: during phase alignment and during stub data streaming.

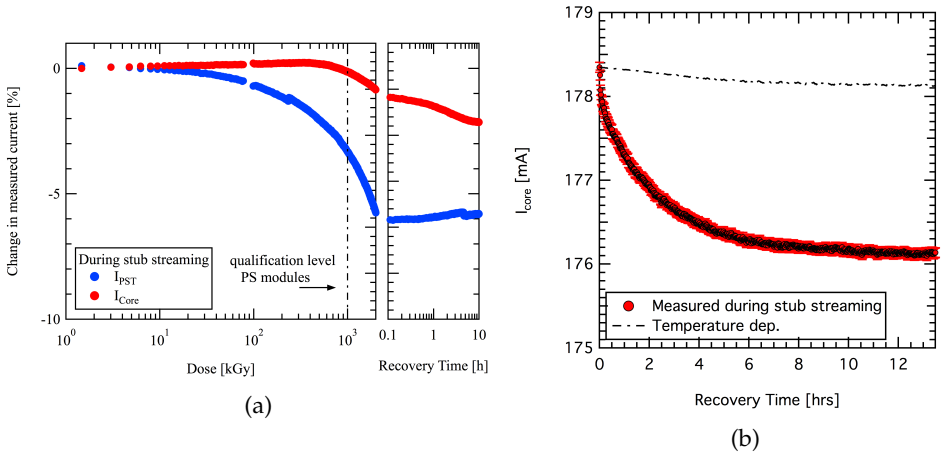


FIGURE 9.13: LEFT: Relative current changes on both channels during irradiation and after irradiation. RIGHT: Current on I_{core} during the annealing period. A stable value is reached and the change in current due to temperature is negligible.

I_{core} slightly rises at the start of the irradiation and starts dropping around 400 kGy to reach a value $\approx$ 1.5 mA below the pre-irradiation value. Also after irradiation I_{core} keeps going down, but reaches a stable value $\approx$ 2 mA below the value when the irradiation was stopped. The post irradiation behaviour of I_{core} is shown in more detail in Figure 9.13b, from which it is clear that the current reaches a stable value. The expected change in current due to temperature changes is also shown in this figure and is negligible. I_{PST} drops throughout the irradiation with $\approx$ 1.5 mA and is stable after irradiation. According to Ref. [27] a decrease in drive strength is indeed expected.

The results of the optimal taps found during the phase alignment stage are shown in Figure 9.14 where in the first figure the dose rate is illustrated. The timespan during which the CIC1 was irradiated is shown between the vertical delimiters. The shifts in optimal tap of 1 unit are also seen without irradiation. The seemingly large jumps between tap values of 4 and 11 are actually equivalent to changes in tap values of 1 unit, as explained in Figure 5.25.

The phase alignment, as well as the word and BX0 alignment, always reported to have locked during the irradiation. Furthermore, the data for each run was compared against the validation framework: two different sets of data were played to the CIC1 and no upset in the data was recorded.

9.3 8CBC3 + CIC1 testing

9.3.1 Introduction

A *CIC mezzanine* was developed which plugs into an 8CBC3 front-end hybrid to mimic a final 2S front-end hybrid (see Figure 9.15). From the DAQ point of view this is equivalent to a final 2S hybrid. In this case the front-end hybrid can be read out electrically over the connector which would connect to the SEH. This is also the configuration in which 2S front-end hybrids will be tested during production: the front-end hybrid will be tested

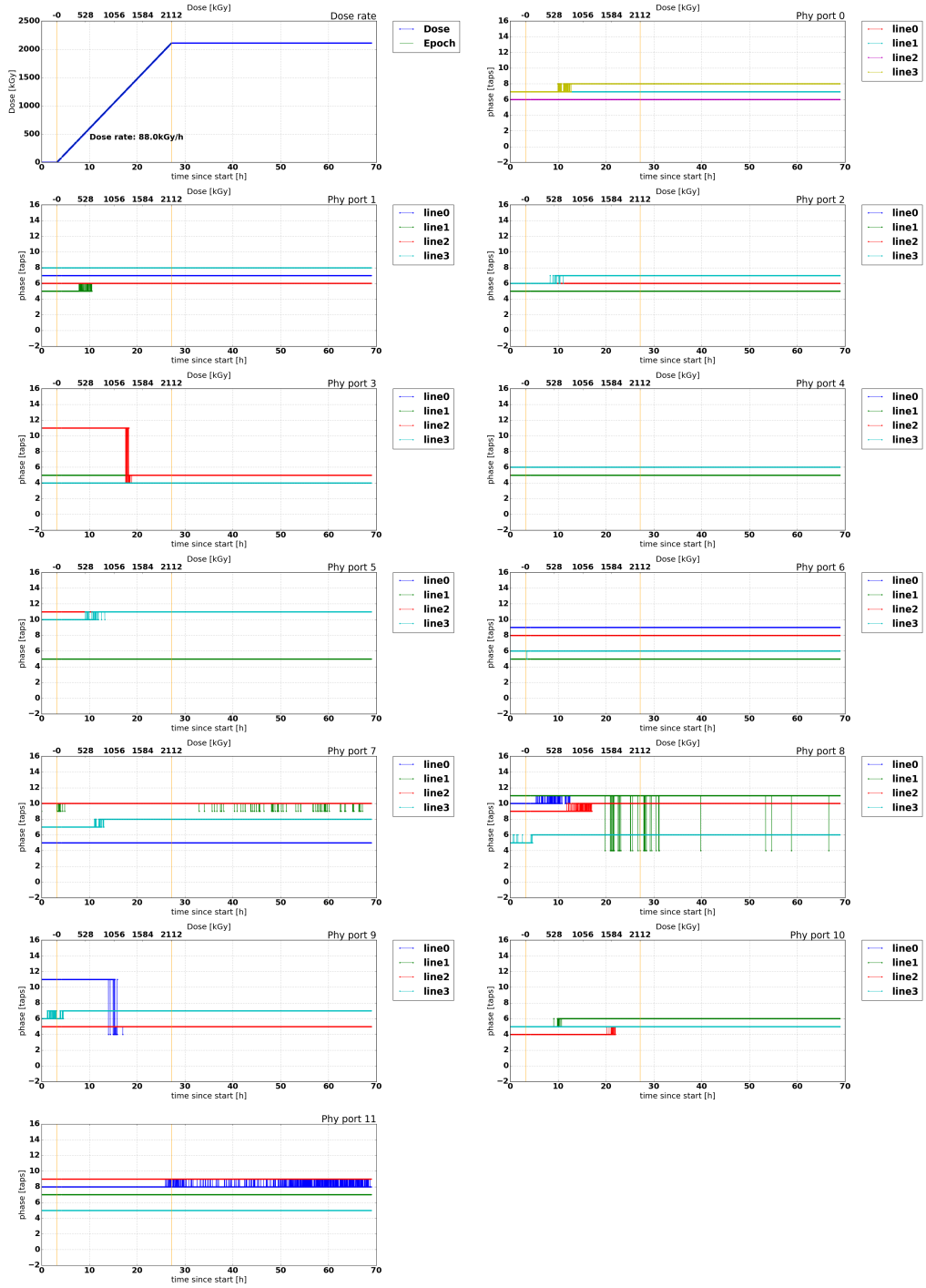


FIGURE 9.14: Optimal tap values found by the CIC phase alignment procedure during TID testing. The results are separated by Phy port. In the first figure the integrated dose over time is shown.

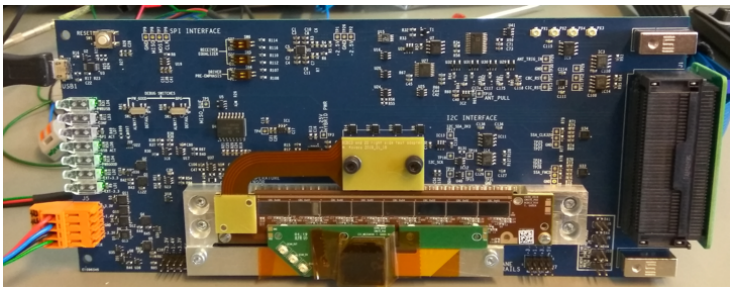


FIGURE 9.15: A 2S FEH carrying 8 CBC3.1 chips to which a CIC mezzanine card is connected to mimic a final 2S front-end hybrid. The CIC1 output data is routed back to the front-end hybrid and goes out of the hybrid over the connector which in a module would connect to the service hybrid, but here connects to a dedicated interface card.

electrically before implementation into a module. A dedicated interface card was developed to run the hybrid in this specific configuration. This interface card can be used to test single hybrids, in which case it is connected with a single FMC cable to a single FMC slot on the FC7, but the card is also designed to operate in a crate system in which many hybrids, mounted on interface cards, are plugged into a backplane. In the latter case a multiplexing scheme permits to address the hybrids one by one.

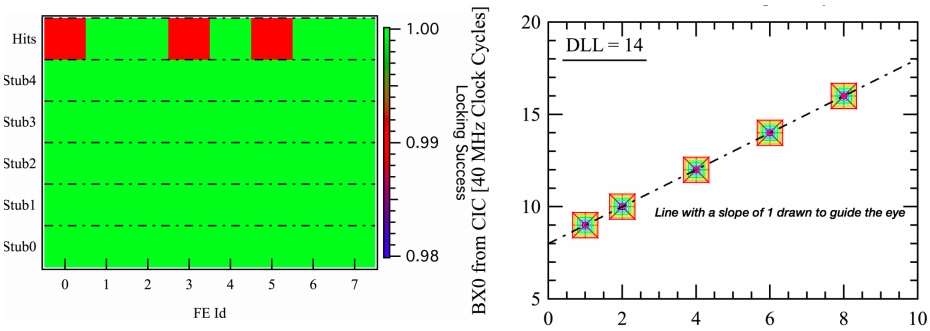
The scope of the 8CBC3+CIC1 testing is to guarantee that CBC3 and CIC1 chips are compatible. The start-up sequence will be tested which will assure that CBCs can be configured in such a way to produce the required patterns for alignment with the CIC. After running the start-up sequence successfully, the CBCs can be configured to produce stub and full-event data by masking dedicated channels or by using the test pulse injection mechanism, and the μDTC data decoder blocks can be exercised. During all these tests, the connectivity of all the lines on the hybrids to the chip I/Os are checked implicitly. The first tests of 8CBC3+CIC1 were conducted in April 2019.

9.3.2 Running 8CBC3 + CIC1

During the testing of this prototype 2S front-end hybrid there are a few important things to keep in mind, which are straightforward, but are easily forgotten when one is used to test chips of the same flavour:

- There is a single fast command line per hybrid. Each fast command which is sent is received and decoded by both chip flavours.
- There are two reset lines on the hybrid: one connected to the CIC1 chip and one connected to the CBC chips.

Before actual data taking with this object can start the CIC start-up sequence needs to be performed. In contrast to the stand-alone CIC testing there is no direct access to the CIC input lines and the dedicated patterns for phase, word and BX0 alignment need to be produced by the CBC chips. The below steps explain how this is done. These are the specific steps which need to be performed to do the phase, word and BX0 alignment when using the CBC chips connected to the CIC. The rest of the start-up sequence, as explained in detail in Appendix B, is still applicable. The exact configuration of the CBC chips to generate the patterns as described below are presented in Appendix C. The below steps



(a) (b)

FIGURE 9.16: LEFT: Result from the phase tuning sequence repeated for 100 times on the same hybrid. Only percent level of failures are seen where the phase tuning does not manage to lock. This happens on the full-event data line. RIGHT: BX0 alignment procedure run for all the 40 stub lines for different delays between the ReSync and the test pulse trigger. The marker colours represent the different chips.

are run on a hybrid where the CBC chips are calibrated so that a common threshold can be applied.

1. Phase tuning: the phase tuning happens in 3 steps for all eight chips in parallel:

   (a) Phase tuning of the stub lines 1, 2 and 4 with the patterns as described in Table 9.1. These patterns need to be sent from the CBC chips continuously. To do this the CBCs are put to a low threshold so that they are always firing on noise, but a large part of the channels are masked in order to only generate two stubs, one in location 42.5 and one in location 85, each with a bend such that the output code defined in the bend LUT is '1010' for both stubs.

   (b) Phase tuning of stub lines 3 and 5 with the patterns as described in Table 9.1. The principle is very similar to the first phase tuning step, it is just required to generate a stub on location 85 and two stubs with lower address. The third stub's bend also needs to result in a bend code of 1010 and the OR254 bit needs to be enabled in the stub output frame in order to get the full '10101010' pattern on the 5th stub line.

   (c) Phase tuning of the full-event data lines: in order to do this, the threshold is again set to a low value and every top channel is masked, while leaving the bottom channels unmasked. In this way the payload in the CBCs full-event data package will be 254 bits of '...1010101....' The CBC chips are then triggered at an L1A frequency of 1 MHz in order to generate this full-event phase tuning pattern with the highest possible repetition rate. This step thus requires to configure the μDTC to generate consecutive L1A triggers at 1 MHz.

   Statistics was gathered on a single hybrid for running the phase tuning. The procedure was run for 100 times and as can be seen from Figure 9.16a it is very stable except the percent level failure on the full-event data line.

2. Word alignment: word alignment of all stub lines with the patterns as described in Table 9.1 can be done by putting the threshold of the CBC chips low and generating stubs by unmasking three top and three bottom channels and masking all other channels.

3. BX0 alignment: in this case specific timing of a pre-defined pattern on the stub data lines is required with respect to the ReSync. This is different from the previous patterns which needed to be generated for the phase and word alignment. In the two latter cases the patterns needed to be repetitive whilst for the BX0 alignment the pattern needs to be there at a given time. This timing constraint requires the use of the test pulse feature on the CBCs: the CBC is configured to generate a test pulse on reception of a test pulse trigger and all channels are masked except two consecutive channels which are in the configured test pulse group. This will result in a stub pattern on the first stub data line which encodes the stub location. This pattern can be loaded to the CIC and the CIC will be looking for this pattern after receiving a ReSync. A dedicated state machine in the μDTC can be configured to send a ReSync directly followed by a test pulse trigger in the next BX. During operation, the BX0 alignment procedure will only be done on one of the 40 stub input lines from the front-ends. Which line the CIC uses is configurable. If it is desired to do the BX0 alignment on a different stub line, then the same mechanism with the test pulse trigger can be used to generate more stubs by unmasking more than 2 channels. If the BX0 alignment needs to be carried out on the line carrying the bend information, this can be achieved by using the parallax correction logic in the CBCs which will artificially generate a bend even though the test pulse triggers are injected in consecutive channels. One important remark is that if the BX0 alignment needs to be performed on the line which carries the sync bit then the OR254 bit which is present in this package needs to be disabled as it was discovered that this OR254 bit is transmitted two stub packages before the actual stub data package it belongs to. The BX0 alignment was tested on a hybrid and the procedure was run for each of the 40 stub data lines. The measured raw data output from the CBCs during the BX0 alignment is shown in Table 9.2. Results from the BX0 alignment are shown in Figure 9.16b where the BX0 value reported by the CIC1 is shown versus the delay, controllable from the firmware, between the ReSync and the test pulse trigger. It can be concluded that the BX0 value reported for each line is the same and that the actual values correlate with the programmed time between ReSync and test pulse trigger.

The above results validate the start-up sequence of the CBC3 and CIC1. Together with the tests performed with the stand-alone CIC1 it can be concluded that the CIC1 and CBC3 can operate together and it guarantees furthermore the availability of a working DAQ FW for the CBC3+CIC1 configuration.

9.4 Summary

Firmware blocks to handle the CIC stub and full-event data were developed and benchmarked against the CIC validation framework. The stub decoding logic on the FPGA is resource heavy and two approaches, both based on a bitonic sorting network, have been investigated. The framework, set up to test the decoding blocks, was then also used to test the CIC's TID sensitivity. No problems were observed in the CIC operation up to

	line 1	line 2	line 3	line 4	line 5
	00000000	00000000	00000000	00000000	10000000
	00000000	00000000	00000000	00000000	10000000
	00000000	00000000	00000000	00000000	10000000
ReSync received	00000000	00000000	00000000	00000000	10000000
ReSync decoded/ Test pulse received	00000000	00000000	00000000	00000000	10000000
	00000000	00000000	00000000	00000000	00000000
	00000000	00000000	00000000	00000000	00000000
	00000000	00000000	00000000	00000000	00000000
	00000000	00000000	00000000	00000000	10000000
	00000000	00000000	00000000	00000000	10000000
	00000000	00000000	00000000	00000000	10000000
	00000000	00000000	00000000	00000000	10000000
	00000000	00000000	00000000	00000000	10000000
Data on stub lines	00010010	00100010	01000010	00100010	10000010

TABLE 9.2: The raw output lines from one CBC chip when a ReSync is received measured with the μDTC: each row represent a 40 MHz clock cycle and each column a stub line (*line 5* carries the sync bit). The CBC takes one clock cycle to act on the fast commands and acts on the ReSync by not outputting the sync bit for 3 BXs. Data, as a result of the test pulse trigger, is visible in the last line. It are these data patterns which can be used for BX0 alignment.

doses which surpass the expected HL-LHC dose. The current consumption on both I_{core} and I_{PST} dropped with respectively 3.5 mA and 1.5 mA due to the irradiation. Such a drop is expected as described in Ref. [27] and the TID test showed that this does not have an influence on data integrity.

A first-ever operation of the CIC with real front-end chips was demonstrated. The CIC was connected to eight CBC3.1 chips, sitting on a front-end hybrid, and the start-up sequence for the CIC was carried out using data generated by the CBC chips. This start-up sequence is not trivial and required a good understanding of the operation of both CBC and CIC. The fact that these ASICs can operate together is an important milestone in proving the feasibility of the data-path on the p_{T} modules. At the moment of writing, the full 2S module data path, including 2 front-end hybrids and a prototype service hybrid with a GBT instead of lpGBT, has been successfully operated through the optical connection.

9.5 Author's contribution

The author contributed to the follow-up of the implementation of the CIC data decoder blocks in the Phy layer. The author performed an evaluation on the resource utilization of the implemented block and compared it to a second approach as described in section 9.2.4.2, where this second approach was coded by the author. The author then set up a system to validate the implemented firmware blocks which entailed using the data produced by the CIC reference model (section 9.2.2), developing the front-end data player which mimics as closely as possible the operation of real devices (section 9.2.3) and obtaining and evaluating the test results for all possible CIC configurations (section 9.2.5

and 9.2.6). The author then set up this system to be used in the TID testing of the CIC and contributed to analysing the results of this test. The author used the experience gained with the front-end data player and the stand alone CIC set-up to contribute to the running of the initialization sequence of the CIC with CBCs, which makes a 2S front-end hybrid ready for data taking.

Part III

Search for Composite Standard Model Dark Matter with CMS at the LHC

Chapter 10

Sexaquarks

10.1 Introduction

G. Farrar introduces in Ref. [88] the possibility that a compact state of uuddss quarks could be deeply bound and have low enough mass to be stable or essentially stable[1]. This particle is referred to as the *Sexaquark* (S). Particles consisting of more quarks than the commonly known baryons are allowed by QCD (quantum chromodynamics) and particles with four and five quarks have indeed been experimentally discovered as discussed in section 10.2. The quark content of the S reflects the hypothesis of the H-dibaryon, which was proposed by R. Jaffe [89] in 1976. The H-dibaryon proposal and negative searches for this particle will be discussed in section 10.3, followed by a description of the S (section 10.4) and the possibility for it being a dark matter particle (section 10.5). Farrar proposes two discovery channels in Ref. [88] which are explained in section 10.6. One of the discovery channels is looking for $\bar{\mathrm{S}}$ which are produced in pp-collisions at the LHC and which can subsequently annihilate on material. The products of this annihilation reaction can be charged particles which can be reconstructed.

Such a search, using the CMS tracker detector, is introduced in more detail in section 10.7. Section 10.8 and 10.9 describe which datasets are used for this search and how the signal is reconstructed. Section 10.10 explains the signal simulation which was set up to outline the search strategy and investigate the signal reconstruction efficiency. Results obtained with this simulation are presented in section 10.11. How the background suppression is performed is described in section 10.12. The evaluation of the systematic uncertainties is the topic of section 10.13 and finally the results and a summary and outlook can be found in section 10.14 and 10.15 respectively.

10.2 Exotic hadrons

Hadrons which do not consist of two or three valence quarks fall outside the conventional quark model and are referred to as exotic hadrons. The Belle experiment at the KEKB $\mathrm{e^+e^-}$ collider was the first experiment to observe such an exotic state in 2003 [90]. This excited state is referred to as X(3872) and was found in the decay of $\mathrm{B^\pm}$ mesons.

The Belle result marked the onset of the discovery of many exotic hadron states, by several experiments, which proceeds until today reaching to the order of 50 claimed discoveries. Most of these states are considered to be tetraquark states of the $\mathrm{q\bar{q}Q\bar{Q}}$ type where q represent a light (u or d) and heavy (c, s or b) quark respectively. Pentaquark states were recently observed by the LHCb collaboration [91]: $\Lambda_\mathrm{b}^0(\mathrm{udb}) \rightarrow \mathrm{J/\Psi(c\bar{c})p(uud)K^-(\bar{u}s)}$ gives indirect evidence for the existence of two pentaquark states ($\mathrm{P_c^+}(4380)$ and $\mathrm{P_c^+}(4450)$) with quark content $\mathrm{c\bar{c}uud}$.

[1] Lifetime larger than the lifetime of the Universe.

In many cases the quark content and the exact configuration of the quarks in these exotic hadrons is still an open question. Quarks can be organised in tightly bound states, or a tetraquark or pentaquark can be built like a molecule of two loosely bound mesons or a loosely bound meson and baryon respectively. Other, even more exotic multi-quark states are also allowed by QCD.

10.3 The H-dibaryon

In 1976 R. Jaffe [89] proposed the existence of a light dibaryon (H) with quark content uuddss as a bound state of two Λ^0s (uds). Jaffe found that the potential of the Λ^0s can be such that they are lightly attractive which would make this dibaryon's mass lie below the threshold, $2 \times m_{\Lambda^0}$ (= 2231 MeV/c^2), for strong decay. Jaffe calculated a mass of 2150 MeV/c^2 for the H. As a result this H-dibaryon would decay weakly to e.g. $\Lambda^0 p\pi$. Several experiments were set up in the decades after this prediction to look for the H. Many of these experiments were located at Brookhaven National Laboratory [92]. Also other experiments at KEKB (Belle) [93], SPS (NA 3) [94], FNAL [95] and LHC (ALICE) [96] started looking for the H-dibaryon or were probing its parameter space implicitly.

Several experiments with different lengths of the *decay region* were designed to investigate production of Hs in a target sitting in a kaon, pion or proton beam with subsequent decays (e.g. to $\Lambda^0 p\pi^-$) of the potentially produced H in the extracted neutral beam. These experiments were sensitive to several orders of magnitude of lifetimes of the H with masses around the $\Lambda^0\Lambda^0$ threshold. No evidence for the existence of an H-dibaryon with a typical weak lifetime was seen. Some of the experiments operating with this principle also had to restrict their search range to a mass > 2 GeV/c^2 to eliminate the neutron background [92].

At Belle, the production of the H+X could also be foreseen in Υ decay, very similar to the observed Υ decay mode to anti-deuterons + X. No evidence for this decay (with H$\rightarrow \Lambda^0 p\pi$ or H$\rightarrow \Lambda^0\Lambda^0$ (the latter above the $m_{\Lambda^0\Lambda^0}$ threshold)) mode of the Υ was seen. This search was restricted however to a fairly small mass region ($\mathcal{O}$(10 MeV/c^2)) around the $\Lambda^0\Lambda^0$ threshold.

Stable Hs living in a nuclear potential of a host nucleus were also looked for. At KEK for example an event was recorded in nuclear emulsion which is linked to the formation and decay of a ${}^{6}_{\Lambda^0\Lambda^0}$He nucleus[2]. This double strange hypernucleus is formed by placing a target in a K^- beam. Ξ^- particles can be formed in the target which, when they come to rest can interact with a proton in the nucleus and form a Λ^0 pair, like this giving rise to a double strange hypernucleus. By studying the kinematics of the recorded event in detail the binding energy of the $\Lambda^0\Lambda^0$ system could be calculated as $BE_{\Lambda^0\Lambda^0} = 7.13$ MeV which shows this system is lightly attractive. A lower mass limit for the H could be extracted as: $m_H > 2 \times m_{\Lambda^0}$ - $BE_{\Lambda^0\Lambda^0} = 2223.7$ MeV/c^2 (at 90% CL) [97], which excluded Jaffe's original prediction. Other counter-experiments (e.g. Ref. [98]) also claim the observation of double strange hypernuclei through detection of a charged pion in each of the reactions where $\Delta s = +1$.

The observation of unstable double strange hypernuclei can be interpreted as evidence against the existence of a strong-interaction stable H where the binding force resembles the nuclear force. There are however claims against this conclusion, as presented in Ref. [99] where it is proven that the coupling strength needed for the H to bind to light

[2] A ^{6}He nucleus with 2 neutrons replaced by 2 Λ^0s

nuclei would be unphysically large and that therefore current limits extracted from double strange hypernuclei are not relevant to extract conclusions on the existence of the H.

Besides a fairly loosely bound H-dibaryon, for which no experimental evidence exists, the possibility that the uuddss quark combination is actually not a loosely bound state of two baryons, but a strongly bound state of six quarks, as in the Sexaquark model, is still viable. This idea will be discussed in the next section.

10.4 Properties of the Sexaquark

Considerations regarding the stability of the S can be summarised as follows for different values of the S mass [100]:

- $m_p + m_n - m_e - 2 \times BE \approx 1860$ MeV/c^2 $< m_S < 2 \times (m_p + m_e) \approx 1878$ MeV/c^2: where BE is the binding energy of nucleons in the nucleus and m_p, m_n and m_e are respectively the mass of the proton, neutron and electron. The lower limit is needed to guarantee nuclear stability against the formation of an S in a nucleus. If the S mass also lies below the stated upper limit then it would be stable against decay to two protons. For the latter to hold, 16% binding energy, compared to two Λ^0s, is required.

- $2 \times (m_p + m_e) \approx 1878$ MeV/c^2 $< m_S < m_p + m_e + m_{\Lambda^0} \approx 2055$ MeV/c^2: if the S lies within this mass range it can decay through a double weak process and for $m_S < 1890$ MeV/c^2 its lifetime is larger than the age of the Universe.

- $m_p + m_e + m_{\Lambda^0} \approx 2055$ MeV/c^2 $< m_S < 2 \times m_{\Lambda^0} \approx 2232$ MeV/c^2: the S can decay through a single weak interaction and it would not be stable.

- $2 \times m_{\Lambda^0} \approx 2232$ MeV/c^2 $< m_S$: the S can undergo a strong decay to two Λ^0s and it would thus not be stable.

Therefore for nuclei and Sexaquarks to be stable on a cosmological timescale it is required that the S mass lies in the 1860-1890 MeV/c^2 range.

Lattice QCD could render an exact calculation of the S mass. Most recent calculations [101] are however still far from a realistic six-quark state but they show the existence of a deeply bound state for the uuddss combination. In literature several calculations of the S mass can be found. Results range from 1130 MeV/c^2 [102] to 2240 MeV/c^2 [103]. It would therefore be beneficial if an experimental search for the S could probe this full mass range.

An important consideration [88] is that one could end up with a high mass for the S when naively using the constituent quark model. Indeed, if one uses the effective masses of the quarks as in the constituent quark model one arrives at 4×300 MeV/c^2 + 2×450 MeV/c^2 = 2100 MeV/c^2. A mass for which the S cannot be stable. Other hadron masses e.g. $\pi^{0,\pm}$ can also not be predicted using this model. So it is fair to say that the intuition from the constituent quark model is not a good guideline for predicting masses.

In Refs. [99, 104] it is estimated that the limits on the H-dibaryon obtained from double strange hypernuclei decays (see section 10.3) are not applicable for a strongly bound S as the S has isospin/flavour symmetry and would therefore not exchange pions. This gives it a much smaller physical size compared to other baryons where pions account for the long range of the nucleon-nucleon interaction. The pions extend the size of the baryons

beyond their Compton radius. If the S radius is smaller than 0.24-0.41 fm, it is calculated in Ref. [88] that the formation time of the S in a strangeness = -2 hypernucleus is larger than $\approx 10^{-9}$ s due to the small wavefunction overlap between the S and two baryons. The S would therefore not have the time to get formed in these hypernuclei. This would make limits from double strange hyperon decays inapplicable. It should be noted however that the arguments in Ref. [99, 104] are approximations, as first principle approaches for calculating the binding energy of standard nucleons, let alone Sexaquarks, in nuclei do not exist.

Whether the small wavefunction overlap between two baryons and an S could also make nucleons in nuclei insensitive to decay to S on the timescale of the Universe could be investigated by e.g. Super-Kamiokande in the $\{n + n\} \rightarrow S + \pi^0$ channel, where $\{\}$ denotes that these neutrons belong to the same nucleus.

Assuming that the S-nucleon interaction goes through the exchange of an f_0 allows for a crude estimation [105] of the physical size of the S: given that the typical size of a nucleon (r_N) is 0.9 fm and 0.2 fm is its Compton wavelength (λ_N), the formula:

$$r_N \approx \lambda_N + \frac{1}{2}\lambda_\pi \tag{10.1}$$

can be put forward with λ_π = 1.4 fm the Compton wavelength of the pion. Given that the S does not couple to pions but to the f_0 gives:

$$r_S \approx \lambda_S + (0 \rightarrow \frac{1}{2})\lambda_{f_0} = 0.1 \rightarrow 0.3 \text{ fm}, \tag{10.2}$$

when using λ_S = 0.1 fm and λ_{f_0} = 0.4 fm. This is a compact object compared to the more common baryons.

10.5 Sexaquarks as a dark matter candidate

Besides being an interesting new type of matter, the S can also be a good dark matter candidate. It should therefore essentially pass all the requirements which were discussed in section 2.4.3.

The study of the S could benefit from knowledge on several S interactions: S-nucleon, S-S and S-$\bar{S}$ scattering; all accompanied with a certain interaction cross section which is velocity dependent. In the literature [88, 99, 105, 106, 107], several estimates and limits on these cross sections can be found. It is clear that their is no unambiguous result, therefore exact values will be avoided below. Some findings will be discussed which will show that evaluating limits for these cross sections is not so straightforward. This will give a feeling that with a proper cross section the dark matter under the form of S particles might have evaded detection.

With the correct S-nucleon cross section the S could have evaded detection in direct detection experiments as the S would thermalize on its way through the atmosphere. The energy it would consequently deposit in direct detection experiments could be below the detection threshold [107].

In Ref. [106] it is shown that the thermalisation could even happen on interstellar gas. This thermalisation would only occur at relatively small radii from the galactic center due to the gas density being larger there. This could even make balloon based direct detection searches inapplicable.

The interaction of the S with the baryonic matter would result in an extra source of heating for the baryonic matter. This extra heating could explain the star formation quenching problem of non-interacting cold dark matter models which do not explain the star formation rate coherently by only including supernova and active galactic nuclei as sources of heating [107].

In Ref. [107] the implications of self-interacting dark matter are discussed. Making the dark matter candidates interact results in heat transfer from dynamic parts of the dark matter halo to less dynamic parts which can result in a smoothing of the halo. If σ_{DM-DM}/M is in the range of $8 \times 10^{-25} - 1 \times 10^{-23}$ $cm^2GeV^{-1}c^2$ self-interacting dark matter (with dark matter-dark matter interaction cross section σ_{DM-DM} and mass of the candidate M) would be a natural solution for e.g. the core-cusp problem. In Ref. [107] it is however shown that the most interesting part of the mass range ($\mathcal{O}(GeV/c^2)$) is excluded by direct detection experiments for the most natural interaction cross section of σ_{DM-p} (dark matter proton interaction cross section) $\approx 10^{-24}$ cm^2. It should however be noted that these are limits on the σ_{DM-p} cross section, but it is hard to imagine a process which could make σ_{DM-p} lie several orders of magnitude from σ_{DM-DM} for a Standard Model dark matter particle such as the Sexaquark.

As explained in section 10.4, the small probability of S to nucleon binding would also explain why the S would not affect Big Bang nucleosynthesis and the S would furthermore not get accumulated in stars hence not affecting stellar evolution [107].

One of the main requirements for a dark matter candidate is for it to be able to explain the relic abundance. In Ref. [105] it is shown that Sexaquarks in the mass range of 1870-1880 MeV/c^2 can quite naturally explain $\Omega_{DM}/\Omega_b \approx 5.3$. The production of the S is supposed to happen at the QCD phase transition where the Universe goes from a quark gluon plasma (QGP) to hadrons. This transition happens when the temperature drops from 170 MeV to 140 MeV. The timescale of the interactions which form the hadrons ($\mathcal{O}(10^{-23}$ s)) is much smaller than the duration of the phase transition (the lifetime of the Universe at this stage is given by: $7.3 \times 10^{-5}(\frac{100\text{ MeV}}{T})^2$ s). Equilibrium equations can thus be used to estimate the abundance of the S at the end of the QCD phase transition. The Sexaquarks produced at this stage are supposed to be maintained until today [105] as the expected small size of the S makes it insensitive to break-up reactions such as $\pi + S \rightarrow \Sigma + \Lambda^0$ or $K^+ + S \rightarrow p + \Lambda^0$. The results for the S relic abundance obtained in Ref. [105] are however in thus far unresolved contrast with the results from Ref. [100] where, also using equilibrium statistics during the QCD phase transition, the relic abundance of Sexaquarks is calculated, after freeze-out, to be $\mathcal{O}(10^{-12})$ the fraction of baryons.

It is thus still debated if the S could explain all or even a significant fraction of the dark matter. Also the actual form in which the S would contribute to the dark matter is under debate, as besides being particulate dark matter it could also be contained in primordial black holes [105]. Nevertheless, the S remains an interesting type of matter which deserves a dedicated search.

10.6 Sexaquark search channels

In Ref. [88] G. Farrar points out two possible discovery channels for the S: one based on production of S in Υ decays and the other based on production of $\bar{S}$ at hadron colliders and subsequent interaction of the $\bar{S}$ with beampipe or detector material. Both channels will be discussed below.

10.6.1 Search channels at B-factories

Very similar to the Belle search discussed in section 10.3 for the H-dibaryon, the S could be looked for in the decay of the Υ. Both the Belle and BABAR experiments would be well suited for this. The Belle search for the H-dibaryon does not exclude the existence of the S as the search window in Ref. [93] was only $\mathcal{O}(10\ \mathrm{MeV/c^2})$ around the $\Lambda^0\Lambda^0$ threshold. The proposed discovery channel for the S at B-factories would be:

$$\Upsilon \rightarrow \mathrm{S}\bar{\Lambda}^0\bar{\Lambda}^0 + \pi\mathrm{s\ and/or}\ \gamma\mathrm{s}, \tag{10.3}$$

for which all the decay products, except the S, can be reconstructed. The S can therefore be searched for in the invariant missing-mass spectrum. The decay channel to $\bar{\mathrm{S}}\ \Lambda^0\Lambda^0\mathrm{X}$ is of course also a possibility.

In Ref. [108] results are discussed of such a search performed by the BABAR collaboration where the mass range $0 < m_{\mathrm{S}} < 2.05\ \mathrm{GeV/c^2}$ is probed using the $\mathrm{S}\bar{\Lambda}^0\bar{\Lambda}^0$ channel. For the full mass range, limits could be set for the branching fraction of $\Upsilon(2S)$ and $\Upsilon(3S)$ to the $\mathrm{S}\bar{\Lambda}^0\bar{\Lambda}^0$ channel of (1.2-1.4) $\times\ 10^{-7}$ (at 90 % CL) which is a branching fraction two orders of magnitude below the branching fraction for decay to the channels containing an anti-deuteron [10].

10.6.2 Search channels at hadron colliders

The S could also be produced in hadron collisions. In order to find this needle-in-the-haystack a large dataset is required. This automatically leads to a search at the LHC. The S and $\bar{\mathrm{S}}$ production rate can be estimated [88] to be of the order of 10^{-4}-10^{-6} relative to the production rate of pions. As a result S and $\bar{\mathrm{S}}$ could be routinely produced at the LHC.

The proposed methodology is to look for $\bar{\mathrm{S}}$ produced in the proton-proton collision. The $\bar{\mathrm{S}}$ could subsequently annihilate on nucleons in the beampipe or detector material, which could lead to any of the below signals:

$$\begin{aligned} \bar{\mathrm{S}}(\bar{\mathrm{u}}\bar{\mathrm{u}}\bar{\mathrm{d}}\bar{\mathrm{d}}\bar{\mathrm{s}}\bar{\mathrm{s}}) + \mathrm{p}(\mathrm{uud}) &\rightarrow \bar{\Lambda}^0(\bar{\mathrm{u}}\bar{\mathrm{d}}\bar{\mathrm{s}}) + \mathrm{K}^+(\mathrm{u}\bar{\mathrm{s}}) \\ \bar{\Lambda}^0(\bar{\mathrm{u}}\bar{\mathrm{d}}\bar{\mathrm{s}}) &\rightarrow \bar{\mathrm{p}}(\bar{\mathrm{u}}\bar{\mathrm{u}}\bar{\mathrm{d}}) + \pi^+(\mathrm{u}\bar{\mathrm{d}}) \end{aligned} \tag{10.4}$$

$$\begin{aligned} \bar{\mathrm{S}}(\bar{\mathrm{u}}\bar{\mathrm{u}}\bar{\mathrm{d}}\bar{\mathrm{d}}\bar{\mathrm{s}}\bar{\mathrm{s}}) + \mathrm{n}(\mathrm{udd}) &\rightarrow \bar{\Xi}_0^+(\bar{\mathrm{d}}\bar{\mathrm{s}}\bar{\mathrm{s}}) + \pi^-(\bar{\mathrm{u}}\mathrm{d}) \\ \bar{\Xi}_0^+(\bar{\mathrm{d}}\bar{\mathrm{s}}\bar{\mathrm{s}}) &\rightarrow \bar{\Lambda}^0(\bar{\mathrm{u}}\bar{\mathrm{d}}\bar{\mathrm{s}}) + \pi^+(\mathrm{u}\bar{\mathrm{d}}) \\ \bar{\Lambda}^0(\bar{\mathrm{u}}\bar{\mathrm{d}}\bar{\mathrm{s}}) &\rightarrow \bar{\mathrm{p}}(\bar{\mathrm{u}}\bar{\mathrm{u}}\bar{\mathrm{d}}) + \pi^+(\mathrm{u}\bar{\mathrm{d}}) \end{aligned} \tag{10.5}$$

$$\begin{aligned} \bar{\mathrm{S}}(\bar{\mathrm{u}}\bar{\mathrm{u}}\bar{\mathrm{d}}\bar{\mathrm{d}}\bar{\mathrm{s}}\bar{\mathrm{s}}) + \mathrm{n}(\mathrm{udd}) &\rightarrow \bar{\Lambda}^0(\bar{\mathrm{u}}\bar{\mathrm{d}}\bar{\mathrm{s}}) + \mathrm{K}_S^0\left(\frac{\mathrm{d}\bar{\mathrm{s}} - \bar{\mathrm{d}}\mathrm{s}}{\sqrt{2}}\right) \\ \bar{\Lambda}^0(\bar{\mathrm{u}}\bar{\mathrm{d}}\bar{\mathrm{s}}) &\rightarrow \bar{\mathrm{p}}(\bar{\mathrm{u}}\bar{\mathrm{u}}\bar{\mathrm{d}}) + \pi^+(\mathrm{u}\bar{\mathrm{d}}) \\ \mathrm{K}_S^0\frac{(\mathrm{d}\bar{\mathrm{s}} - \bar{\mathrm{d}}\mathrm{s})}{\sqrt{2}} &\rightarrow \pi^+(\mathrm{u}\bar{\mathrm{d}}) + \pi^-(\bar{\mathrm{u}}\mathrm{d}) \end{aligned} \tag{10.6}$$

In the above reactions all the final state particles can be reconstructed and therefore the invariant mass of the $\bar{\mathrm{S}}$-nucleon system as well as the interaction vertex of the $\bar{\mathrm{S}}$ with the nucleon can be extracted.

The formed $\bar{\mathrm{S}}$ will have a transverse momentum comparable to other hadrons formed at LHC pp-collisions, which is $\mathcal{O}$(GeV/c). This implies that the final state products will also have low transverse momenta. This, together with the fact that the annihilation cross section of the $\bar{\mathrm{S}}$ with material is low, and the fact that final state products are produced significantly displaced from the luminous region makes this a challenging signal to look for.

10.7 Adopted search strategy

The next sections describe a first-ever search for the Sexaquark particle at the LHC. The signal which is considered in this search is the one described in Equation (10.6) and the CMS detector is used to look for $\bar{\mathrm{S}}$ annihilating on neutrons in the CMS beampipe. In case the search is negative, an upper limit can be placed on the product of the $\bar{\mathrm{S}}$ production cross section ($\sigma(\mathrm{pp} \to \bar{\mathrm{S}})$) and the cross section for $\bar{\mathrm{S}}$ annihilation on a neutron resulting in a $\mathrm{K}^0_{\mathrm{S}}$ and an $\bar{\Lambda}^0$ ($\sigma(\bar{\mathrm{S}} + \mathrm{n} \to \mathrm{K}_S + \bar{\Lambda}^0)$). This will be performed by applying the following formula:

$$[\sigma(\mathrm{pp} \to \bar{\mathrm{S}}) \times \sigma(\bar{\mathrm{S}} + \mathrm{n} \to \mathrm{K}^0_s + \bar{\Lambda}^0)]^{\mathrm{UL}} = \frac{N^{\mathrm{UL}}}{\frac{N_{\mathrm{ev}} \times <\mathrm{PU}> \times (1-C)}{\sigma_{\mathrm{pp}}} \times \epsilon_{\mathrm{reco}} \times \epsilon_{\mathrm{fid}} \times \epsilon_{\mathrm{BkgCuts}} \times N_{\mathrm{n}}}. \quad (10.7)$$

Here N^{UL} is the upper limit on the number of signal events after all selection cuts, N_{ev} the total number of pp interactions in the considered dataset, <PU> the average pileup in this dataset, C the overlap rate between datasets, σ_{pp} the proton-proton inelastic cross section, ϵ_{reco} the $\bar{\mathrm{S}}$ reconstruction efficiency, ϵ_{fid} the efficiency for the signal to be contained in a fiducial region in which systematic uncertainties are controlled, $\epsilon_{\mathrm{BkgCuts}}$ the efficiency with which the $\bar{\mathrm{S}}$ candidates survive the background cuts and N_{n} the neutron column density of the beampipe. In the sections which follow all these terms will be evaluated.

Due to the low momenta of the final state particles, no dedicated trigger can be used and the strategy adopted here is to look for the $\bar{\mathrm{S}}$ signal in all the available datasets. The $\bar{\mathrm{S}}$ could as well be formed in pileup interactions and all pp collisions within an event are therefore potentially interesting. The datasets which are used in this search are discussed and listed in section 10.8. Data in CMS is categorised in *primary datasets*. Each primary dataset has a list of triggers which are linked to it. In this way a certain dataset contains a certain type of physics. Triggers linked to different primary datasets can however trigger on the same event and as a result the same event might end up in different primary datasets leading to overlap between datasets. This overlap rate was measured (see section 10.8) and used to calculate the number of unique pp interactions in the datasets.

The decay of the $\mathrm{K}^0_{\mathrm{S}}$ and $\bar{\Lambda}^0$ can happen within the CMS tracker volume as the c.τ of the $\mathrm{K}^0_{\mathrm{S}}$ and $\bar{\Lambda}^0$ are respectively $\approx$ 2.7 cm and $\approx$ 7.9 cm. The overall branching fraction of the decay of the $\mathrm{K}^0_{\mathrm{S}}$ and $\bar{\Lambda}^0$ to the final state considered in this search is 44.2% (69.2% for $\mathrm{K}^0_{\mathrm{S}} \to \pi^+ \pi^-$ and 63.9% for $\bar{\Lambda}^0 \to \pi^+ \bar{\mathrm{p}}$) [10]. These charged final state particles can

be reconstructed with the CMS tracker and actually the K^0_S, Λ^0 and $\bar{\Lambda}^0$, referred to as V^0 *particles*, are part of the standard reconstruction in CMS [109]. This so-called V^0 *fitting* had to be adapted however to fit the kinematics of the V^0s from the signal as described in section 10.9.

To outline the analysis and evaluate the reconstruction efficiency and the background rejection, a dedicated signal simulation is required. The generation of the $\bar{S}$ is performed with the EPOS-LHC [110] event generator and the $\bar{S}$-neutron interaction is implemented in GEANT4 [111] which is part of the standard CMS simulation software. More details on the signal simulation and how it was used to estimate the reconstruction efficiency can be found in sections 10.10 and 10.11 respectively.

To evaluate the background, reconstructed S candidates are used. S, as opposed to $\bar{S}$, will not annihilate on neutrons and reconstructed S candidates can therefore be used as a background control sample. The reconstructed S only differ from the $\bar{S}$ by the charge of the decay products of the $\Lambda^0/\bar{\Lambda}^0$. To achieve a performant signal to background discrimination, a boosted decision tree (BDT) is used. Section 10.12 explains in detail how the background rejection is performed and shows what is the source of this background.

Several sources of systematic uncertainties are described in section 10.13. Taking these into account, an upper limit on the product of cross sections is then evaluated in section 10.14.

As can be expected from a first-ever search, the analysis described below is not optimised yet in every possible way. Several fundamental, although sometimes non-anticipated nor intuitive, insights and difficulties for this specific signal will be discussed. The outlook (section 10.15) will then review the possible improvements for this analysis.

10.8 Datasets

Due to the low momenta of the final state particles no dedicated trigger can be used for this signal. For this reason, the signal is searched for in all existing primary datasets and the pileup interactions are also considered. This is a unique feature of the analysis: to our knowledge no analysis so far has used all recorded pp collisions in a beyond the Standard Model search.

The data considered here is the data collected in 2016 by CMS. In early 2016 runs (until RunF) the CMS tracker was coping with a problem introduced by the smaller inter-bunch crossing time, higher peak luminosity and the change in detector temperature with respect to Run 1. This resulted in a saturation effect of the APV chip's pre-amplifier which led to decreased hit efficiencies severely impacting tracking performance. No simulation was available which accurately describes this so-called *APV saturation effect* and therefore it was not possible to evaluate a robust systematic uncertainty on tracking efficiencies (see section 10.13) for early 2016 data. The search was thus limited to the second half of the 2016 data taking.

The primary datasets which were analysed are listed in Table 10.1, where the number of events for each run and dataset are shown together with the average number of *valid* primary vertices (PVs) per event. *Valid* PVs are defined here and in the following sections by requesting the number of degrees of freedom of the vertex to be larger than 4, its $|z|$ location to be smaller than 24 cm and its radial displacement to be smaller than 2 cm.

The datasets listed in Table 10.1 contain a total number of $\approx 2.01 \times 10^9$ recorded events (N_{ev}) with on average 18.2 reconstructed valid primary vertices per event. The conversion from valid reconstructed PVs to the true number of inelastic interactions in the event

Dataset	Number of events	Average number of valid PVs
/BTagCSV/Run2016G-07Aug17-v1/AOD	100,014,471	18.1
/BTagCSV/Run2016H-07Aug17-v1/AOD	64,209,667	19.1
/BTagMu/Run2016G-07Aug17-v1/AOD	27,342,838	16.8
/BTagMu/Run2016H-07Aug17-v1/AOD	28,329,985	17.6
/Charmonium/Run2016G-07Aug17-v1/AOD	61,398,862	17.2
/Charmonium/Run2016H-07Aug17-v1/AOD	70,124,218	18.1
/DisplacedJet/Run2016G-07Aug17-v1/AOD	18,006,002	17.6
/DisplacedJet/Run2016H-07Aug17-v1/AOD	19,433,403	19.1
/DoubleEG/Run2016G-07Aug17-v1/AOD	76,538,905	17.7
/DoubleEG/Run2016H-07Aug17-v1/AOD	82,569,062	18.9
/DoubleMuon/Run2016G-07Aug17-v1/AOD	44,182,144	18.2
/DoubleMuonLowMass/Run2016G-07Aug17-v1/AOD	44,835,919	17.4
/DoubleMuonLowMass/Run2016H-07Aug17-v1/AOD	52,461,380	18.3
/HTMHT/Run2016G-07Aug17-v1/AOD	30,133,219	18.2
/HTMHT/Run2016H-07Aug17-v1/AOD	33,060,229	19.9
/JetHT/Run2016G-07Aug17-v1/AOD	118,152,949	17.2
/JetHT/Run2016H-07Aug17-v1/AOD	120,597,407	18.3
/MET/Run2016G-07Aug17-v1/AOD	25,726,953	19.6
/MET/Run2016H-07Aug17-v1/AOD	37,033,146	22.2
/MuOnia/Run2016G-07Aug17-v2/AOD	34,955,119	16.7
/MuOnia/Run2016H-07Aug17-v1/AOD	37,575,978	17.8
/MuonEG/Run2016G-07Aug17-v1/AOD	33,573,015	18.1
/MuonEG/Run2016H-07Aug17-v1/AOD	28,812,785	18.7
/SingleElectron/Run2016G-07Aug17-v1/AOD	152,098,617	17.5
/SingleElectron/Run2016H-07Aug17-v1/AOD	127,093,812	18.7
/SingleMuon/Run2016G-07Aug17-v1/AOD	147,945,745	17.6
/SingleMuon/Run2016H-07Aug17-v1/AOD	171,137,991	18.7
/SinglePhoton/Run2016G-07Aug17-v1/AOD	31,884,899	17.2
/SinglePhoton/Run2016H-07Aug17-v1/AOD	33,674,739	18.1
/Tau/Run2016G-07Aug17-v1/AOD	78,733,260	17.7
/Tau/Run2016H-07Aug17-v1/AOD	75,409,819	19.0
	Total = 2,007,046,538	**Average, weighed by number of events = 18.2**

TABLE 10.1: Number of events and valid primary vertices in the primary datasets which were analysed.

is extracted from simulation studies [51] of which the result is shown in Figure 10.1. For an average value of 18.2 reconstructed valid primary vertices per event this results in 21.4 interactions per event. The true number of inelastic pp interactions stored in the datasets listed in Table 10.1 is thus $\approx 43.0 \times 10^9$.

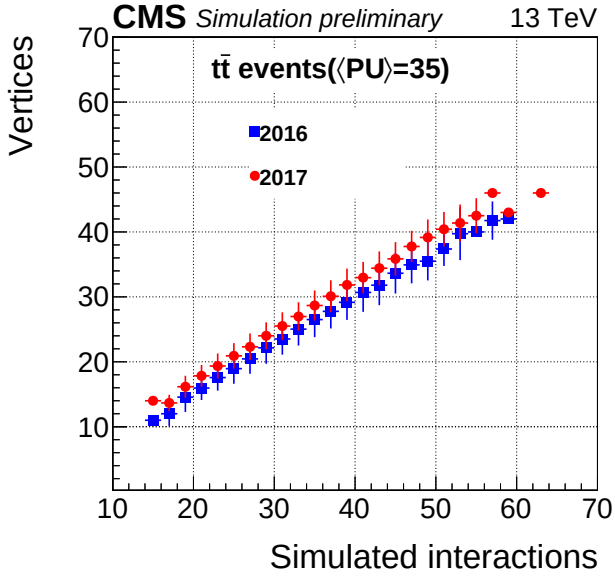


FIGURE 10.1: Number of reconstructed vertices as function of the number of simulated interactions [51].

Different triggers, mapped to different primary datasets, might fire on the same event and thus the primary datasets contain duplicate events. This overlap factor was determined by measuring the sum of the rates of all triggers which feed into the primary datasets in Table 10.1 on one hand and the total combined HLT rate for these triggers on the other hand. The difference in rates gives the overlap factor. This study was performed on special datasets (*HLTPhysics*) which have no HLT selection applied. An overlap of 12.6±0.5% and 13.1±0.7% between the datasets listed in Table 10.1 were obtained for RunG and RunH respectively. By reweighing these overlap factors with the number of events in the runs this leads to an overall overlap factor (C) of 12.84±0.61%.

The total number of unique pp collisions available in the datasets listed in Table 10.1 is thus <PU> $\times (1-C) \times N_{\text{ev}} = 21.4 \times (1-0.1284) \times 2.01 \times 10^9 = 37.4 \times 10^9$. Using a cross section of 70 mb for 13 TeV inelastic pp collisions this number of events corresponds to an integrated luminosity of 0.535 pb^{-1}.

10.9 Physics objects and signal reconstruction

To reconstruct the $\bar{\text{S}}$ signal, track pairs are formed and checked whether they could be originating from V^0 particle decays. A common vertex is fitted to these reconstructed K^0_S and $\Lambda^0/\bar{\Lambda}^0$ and these form the $\text{S}/\bar{\text{S}}$ candidates. These two steps will be explained in more detail below.

K^0_S, Λ^0 and $\bar{\Lambda}^0$ reconstruction is part of the standard CMS reconstruction [109]. This standard V^0 reconstruction was modified slightly to be able to reconstruct the $\bar{S}$ signal, but the standard algorithm is maintained: the V^0 reconstruction module takes all reconstructed tracks as input but only uses tracks which pass the following quality criteria: the significance of their transverse impact parameter should be larger than 2, the normalised[3] χ^2 should be smaller than 10, the track should have at least 7 hits and a p_T larger than 0.35 GeV/c.

The vertexing of track pairs to a V^0 candidate uses oppositely charged tracks which satisfy the requirements in the previous paragraph and which have a point of closest approach in $r\phi$ not larger than 2 cm. This point of closest approach should lie within the tracker volume and the scalar product of the 3D track momentum vectors at this point of closest approach should not be negative. The invariant mass of the track pair is then calculated under the assumption that both tracks are pion tracks and track pairs with an invariant mass larger than 0.6 GeV/c^2 are excluded.

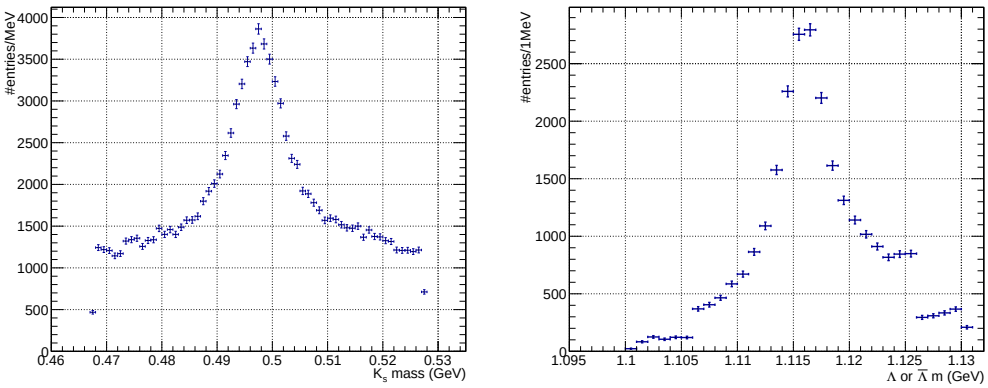
Track pairs which fulfil the above requirements then serve as input to a Kalman vertex fit [112]. For track pairs for which the vertex-fit succeeds the reconstruction proceeds. The normalised χ^2 of the vertex-fit is required not to be larger than 15 and the significance of the transverse distance of the vertex with respect to the beamspot is required to be larger or equal to 10. The refitted tracks from the Kalman vertex-fit are subsequently used to calculate the energy of the V^0 using the three hypotheses whether the V^0 is a K^0_S, Λ^0 or $\bar{\Lambda}^0$. The distinction between Λ^0 and $\bar{\Lambda}^0$ is made by looking at the charge of the track with the highest momentum, which is taken to be the (anti-)proton track. As a last constraint the candidate V^0s are required to pass mass cuts constraining them within $\pm$0.070 GeV/c^2 ($\pm$0.050 GeV/c^2) from 0.498 GeV/c^2 (1.116 GeV/c^2) for the K^0_S (Λ^0 and $\bar{\Lambda}^0$).

In contrast to the standard V^0 reconstruction the one modified for the $\bar{S}$ reconstruction does not request the V^0 candidate to be pointing to the luminous region. This cut is applied in the standard V^0 reconstruction module to eliminate background from random combinations of tracks. However the K^0_S and $\bar{\Lambda}^0$ resulting from the $\bar{S}$ annihilation are not necessarily pointing to the luminous region. This means that the reconstructed K^0_S and $\bar{\Lambda}^0$ present in the standard CMS collections cannot be used as input to the $\bar{S}$ reconstruction, but that the V^0 reconstruction has to be rerun on the full dataset. The latter is a computationally heavy operation. For this adapted reconstruction, no requirement is placed on the impact parameter of the V^0s and the mass constraints are tightened in order to decrease the background. Mass constraints of $\pm$ 0.03 GeV/c^2 ($\pm$ 0.015 GeV/c^2) around the masses stated in the previous paragraph are used for the K^0_S (Λ^0 and $\bar{\Lambda}^0$).

Mass distributions of K^0_S, Λ^0 and $\bar{\Lambda}^0$, reconstructed with this algorithm, are shown in Figure 10.2. These V^0s were reconstructed in *SingleMuon_Run2016H-07Aug17* data. The K^0_S invariant mass distribution shows a significant background contribution. This background can be rejected by requiring the K^0_S to point to the luminous region, as will be illustrated in section 10.13, but as explained before, accepting non-pointing K^0_S is essential for reconstructing the $\bar{S}$ signal.

The output K^0_S, Λ^0 and $\bar{\Lambda}^0$ of this *adapted V0Fitter* are used as input to the S and $\bar{S}$ reconstruction. This reconstruction makes use of kinematic fitting. Kinematic fitting [113] is a technique commonly applied for signal reconstruction. This technique allows to add physical constraints to the fitting problem. These additional constraints result in smaller uncertainties on the fitted parameters. For signal reconstruction using tracks, examples of constraints are the invariant mass of combinations of tracks or constraining

[3] χ^2 divided by the number of degrees of freedom.



(a) (b)

FIGURE 10.2: K^0_S, Λ^0 and $\bar{\Lambda}^0$ masses as reconstructed with the *adapted* V0Fitter in *SingleMuon_Run2016H-07Aug17* data.

multiple tracks to a common vertex. The constraints are applied by the use of Lagrange multipliers.

In the $S/\bar{S}$ reconstruction K^0_S and $\Lambda^0/\bar{\Lambda}^0$ are kinematically refitted to their world average mass [10] and the V^0 pairs are fitted to a common vertex using a Kalman vertex-fit. This vertex would in the signal case represent the annihilation vertex of the $\bar{S}$. The vertex of the two V^0s is required to have a normalised χ^2 not larger than 10.

10.10 Signal simulation

10.10.1 Introduction

The following sections describe the signal simulation which was set up to shape the analysis and to extract the $\bar{S}$ reconstruction efficiency. The signal simulation uses a special tuned version of the EPOS-LHC minimum bias event generator which implements a production process for S and $\bar{S}$ particles. The generated events are then input to the CMS simulation software where a dedicated $\bar{S}$ interaction with matter was implemented in the GEANT4 framework (section 10.10.3). The GEANT4 software also performs the simulation of the particle interactions with matter for all the particles in the event. Consequently, the default CMS software is used to simulate the detector response and apply the digitisation of the signal, superimpose pileup on the signal event and run the reconstruction algorithms. The simulation is very important here to understand the full topology of the signal, which due to the material interaction, is fairly uncommon. The simulation shows that the neutron momentum in the target nucleus actually has a non-negligible influence on the reconstruction of the $\bar{S}$ kinematical parameters (section 10.10.4) and gives an idea about the acceptance of the signal (section 10.10.5). The kinematics of the signal V^0s and final state particles are discussed in section 10.10.6, which sets the stage for investigating the reconstruction efficiencies for the signal as described in section 10.11. In section 10.10.2 the kinematics of $\bar{S}$ with different masses as obtained with EPOS-LHC are shown. In the successive sections the focus is on $\bar{S}$ with a mass of 1.8 GeV/c^2.

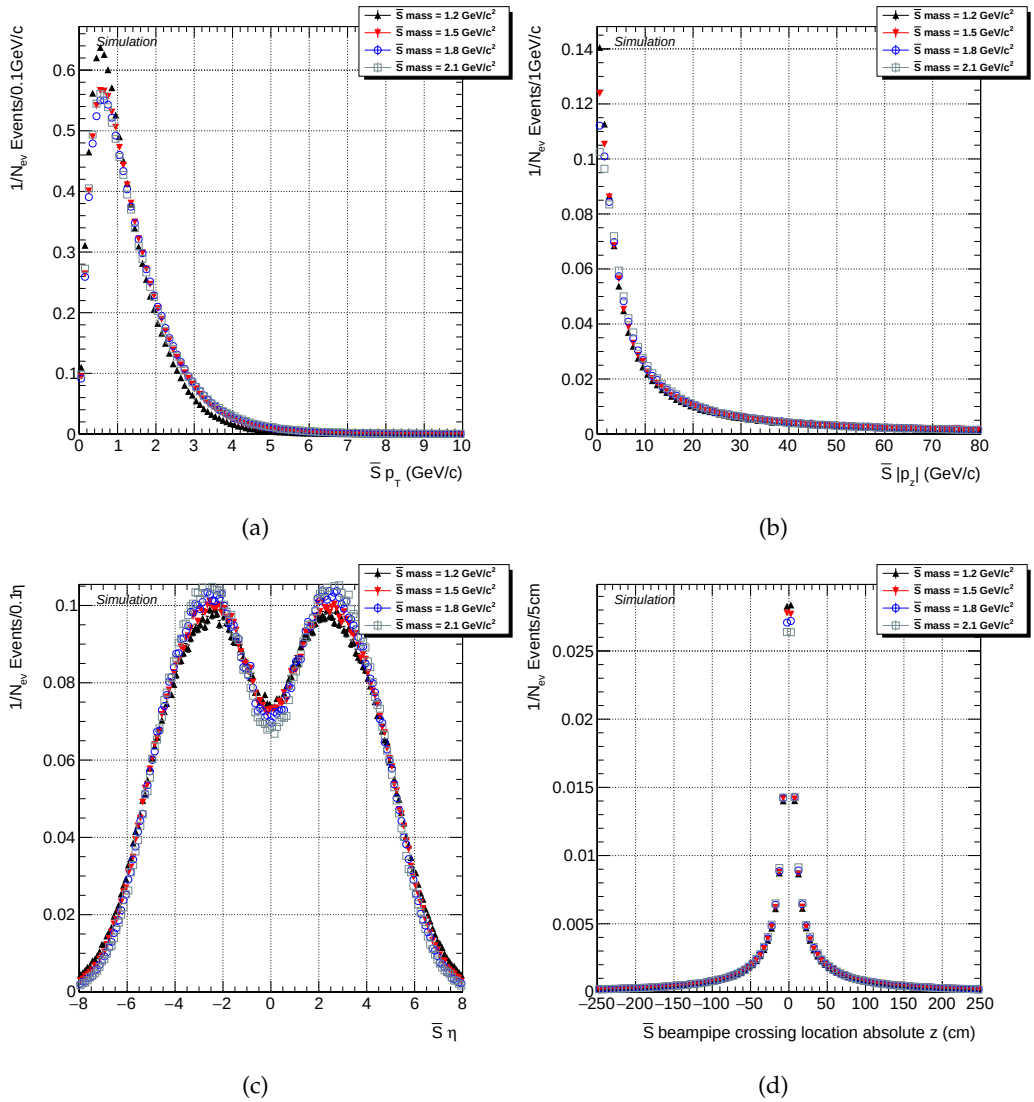


FIGURE 10.3: Main kinematics of the $\bar{S}$ at generator level.

10.10.2 $\bar{S}$ kinematics

The starting point for the signal simulation is a special tuned[4] version of EPOS-LHC [110] which includes the production of S and $\bar{S}$ in the pp-collision. EPOS-LHC is mostly used as event generator for heavy-ion collisions, but as Ref. [110] shows it is also an excellent tool to describe strangeness production in LHC pp collisions.

The p_T and p_z distribution of the generated $\bar{S}$ are shown in Figure 10.3 for $\bar{S}$ masses of 1.2, 1.5, 1.8 and 2.1 GeV/c^2. These distributions show the typical $\mathcal{O}$(GeV/c) range for the p_T of hadrons produced at LHC collisions. The final state particles will therefore have low transverse momenta as well, which will complicate their reconstruction. Figure 10.3 furthermore shows the pseudorapidity distribution of the $\bar{S}$, which shows that the majority of the $\bar{S}$ are produced forward. The last figure in Figure 10.3 shows the z location[5] of where the $\bar{S}$ would cross the CMS beampipe using a beampipe radius of 2.21 cm (see section 10.10.3).

The distribution of the z location of the valid primary vertices and the number of valid primary vertices per event are different in simulation and data. Since an accurate description of these variables is important for this analysis a *reweighing* procedure is applied to reweigh the simulation to data. In this procedure, an event with a primary vertex which is more comparable to a primary vertex in data is given a larger weight.

Figure 10.4 illustrates the result of this two-dimensional reweighing procedure which corrects for the different z locations of the PVs as well as the number of primary vertices in the event. The figures show that the reweighed MC (Monte Carlo) distributions indeed coincide with the distributions for data. The distributions which were used here for data are extracted from all the events used in this search (listed in Table 10.1) in order to have an as representative as possible reweighing of simulation to data. This event-by-event reweighing of the simulation is always applied to obtain the results which are shown hereafter.

For the next sections only the case of $\bar{S}$ candidates with a mass of 1.8 GeV/c^2 was considered. The simulated sample includes a total of $\approx 1.3 \times 10^8$ $\bar{S}$ in $\approx 3.4 \times 10^7$ events.

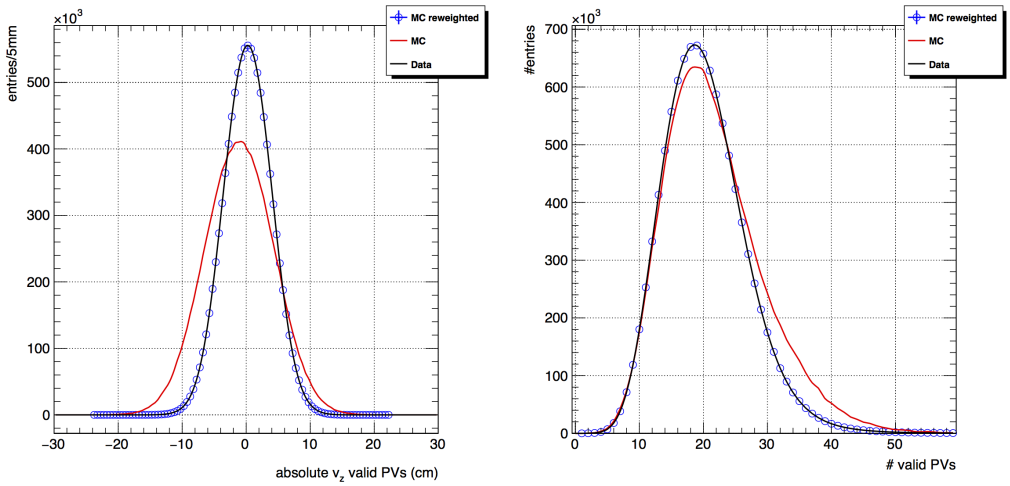
10.10.3 $\bar{S}$-neutron interaction

The actual interaction of the $\bar{S}$ with a neutron was implemented specifically for this signal in GEANT4 [111]. The $\bar{S}$ was added as a new particle to the GEANT framework and its neutron interaction was implemented as a hadronic interaction and a two body particle decay of the $\bar{S}$-neutron system to a K^0_S and an $\bar{\Lambda}^0$. No angular preference due to potential spin effects were included in this reaction. The annihilation reaction leaves behind a nucleus with a neutron less.

The interaction cross section of the $\bar{S}$ with material is very small. To simulate this, the interaction cross section could indeed be put very low in GEANT which would guarantee that the $\bar{S}$ has a uniform interaction probability along its passage through the material. This approach would unavoidably lead to many $\bar{S}$ escaping the beampipe volume without interacting. To guarantee an economic simulation whilst maintaining a low interaction cross section and thus a uniform interaction probability, a *looping* mechanism, which iterates the transport of the $\bar{S}$ through the sensitive volume, was implemented. This was

[4] Private version provided by the EPOS-LHC author which has S and $\bar{S}$ production included.

[5] Coordinates are referred to in some figures as *absolute* in order to distinguish from coordinates calculated with respect to other reference points (e.g. beamspot or center of the beampipe) than the absolute origin.

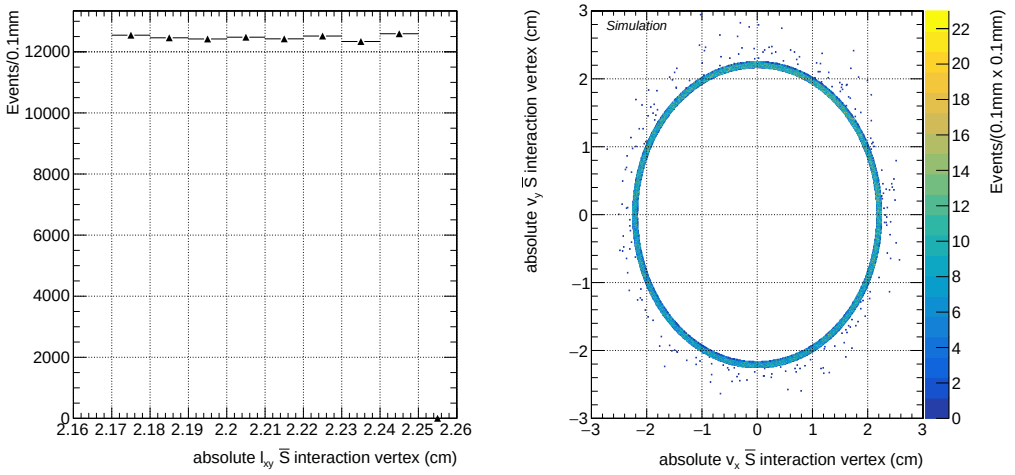


(a) (b)

FIGURE 10.4: z locations of the valid primary vertices (LEFT) and the number of valid primary vertices per event (RIGHT) in data and simulation. A reweighing procedure is applied to reweigh the MC simulation to the data, as described in the text.

done by *killing* (GEANT terminology) the $\bar{S}$ which go outside a defined volume and creating a secondary, as an exact copy of the killed $\bar{S}$, at its original creation vertex (*cv*). When the interaction cross section is put low enough, the $\bar{S}$ can do several iterations through the beampipe volume before interacting. In the end, this results in a flat interaction probability per crossed unit of material volume for $\bar{S}$ with different pseudorapidities. In this search we focus on the beampipe to find $\bar{S}$ annihilation vertices. The CMS beampipe is made of beryllium, has a thickness of 0.8 mm with an inner radius of 2.17 cm and is a cylinder up to $|z|$ = 145.7 cm after which it adopts a conical shape [114]. In the simulation the $\bar{S}$ is killed and a copy is created at its original creation vertex when it moves radially further than 2.5 cm. This value was chosen to include the outer radius of the beampipe at 2.25 cm and does not yet include the first layer of the pixel detector which is located at a radius of 4.3 cm. Furthermore, $\bar{S}$ which are produced very forward and would only interact at high $|z|$ are killed when the $\bar{S}$'s z coordinate becomes larger than 160 cm. It will be shown later (section 10.10.5) that $\bar{S}$ which interact with material at these high $|z|$ values are anyway not reconstructable.

The result of this iterative transport is shown in Figure 10.5. The left figure shows that this approach results in a uniform annihilation probability throughout the beampipe. The beampipe structure which shows up on the right hand side in Figure 10.5 is indeed the expected one with an inner radius of 2.17 cm and an outer radius of 2.25 cm. Some interactions outside of the beampipe radius are also visible on the right hand side in Figure 10.5. These interactions happen on air molecules. Interactions within the beampipe do not occur due to the beampipe-vacuum. In simulation, the beampipe-axis is centred at (x = 0 cm, y = 0 cm), but in reality the beampipe center (*bpc*) is offset and was found to be lying at (x = 0.124 cm, y = 0.027 cm) [115]. This difference between data and simulation will be of importance later when a search window in the lateral direction is placed to find



(a) (b)

FIGURE 10.5: Coordinates of the $\bar{S}$ interaction vertices at generator level.

$\bar{S}$ particles.

Figure 10.6 and 10.7 show two simulated events in which the $\bar{S}$ got reconstructed (see section 10.11). Figure 10.6 shows an event where the $\bar{S}$ has a fairly high pseudorapidity whereas Figure 10.7 shows a more central event.

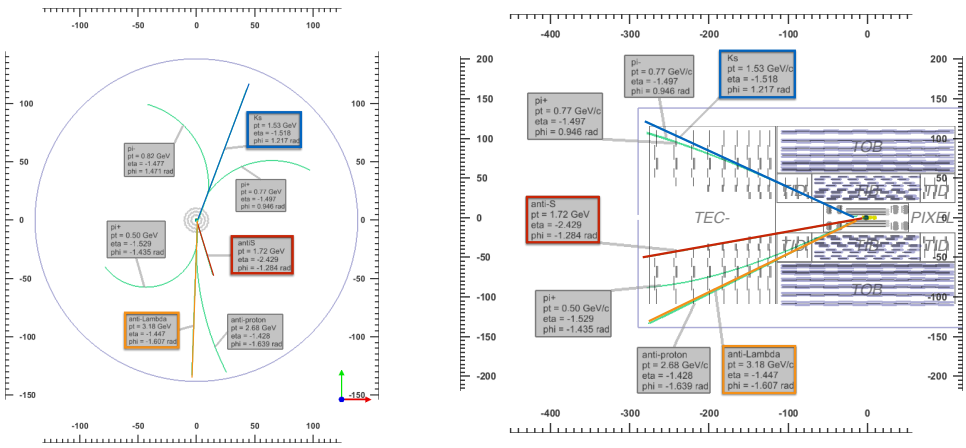
The iterative transport guarantees that all $\bar{S}$ will interact with the beampipe. In reality however, not all $\bar{S}$ have the same interaction probability. The interaction probability of the $\bar{S}$ is a function of the $\bar{S}$ pseudorapidity as $\bar{S}$ with larger $|\eta|$ will have a larger path length through the beampipe. An additional event-by-event weight parameter is therefore introduced and defined as $1/\sin\theta_{\bar{S}}$. The weight parameter is shown as function of η of the $\bar{S}$ and z location of the $\bar{S}$ interaction vertex in Figure 10.8.

From now on all results which are shown have this weight parameter applied. The reweighed distribution of the $\bar{S}$ transverse and longitudinal momentum, pseudorapidity and z location of its interaction vertex are shown in Figure 10.9[6]. Comparison with Figure 10.3 shows that the $\bar{S}$ signal is significantly enhanced in the forward region and the $\bar{S}$ average p_T is significantly reduced.

10.10.4 Influence of neutron Fermi momentum

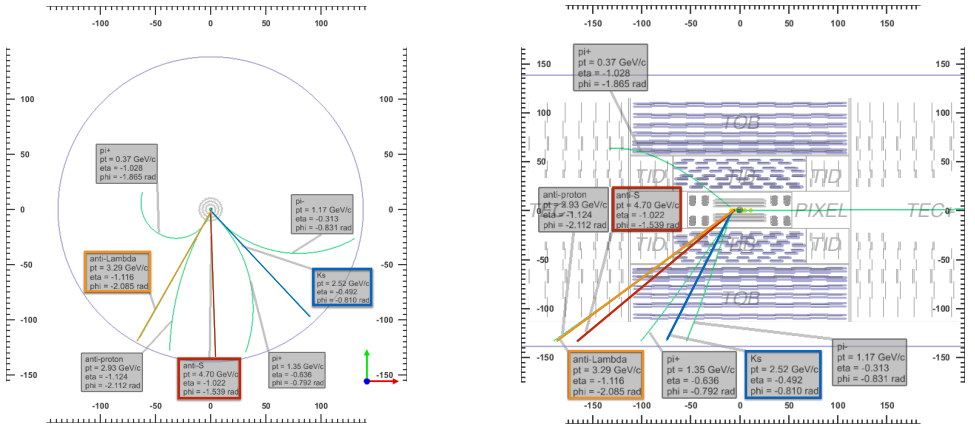
The neutron on which the $\bar{S}$ annihilates is not stationary but is moving around in the nucleus with a certain Fermi momentum of the order of 200 MeV/c [116]. The fact that this neutron is not at rest spoils the accuracy with which the $\bar{S}$ properties, such as its invariant mass and its direction of propagation, can be reconstructed. The exact interaction vertex position of the $\bar{S}$ can however still be extracted.

[6]The z location of the interaction vertex was calculated using the pseudorapidity of the $\bar{S}$ and a radius of the beampipe of 2.21 cm, as in Figure 10.3.



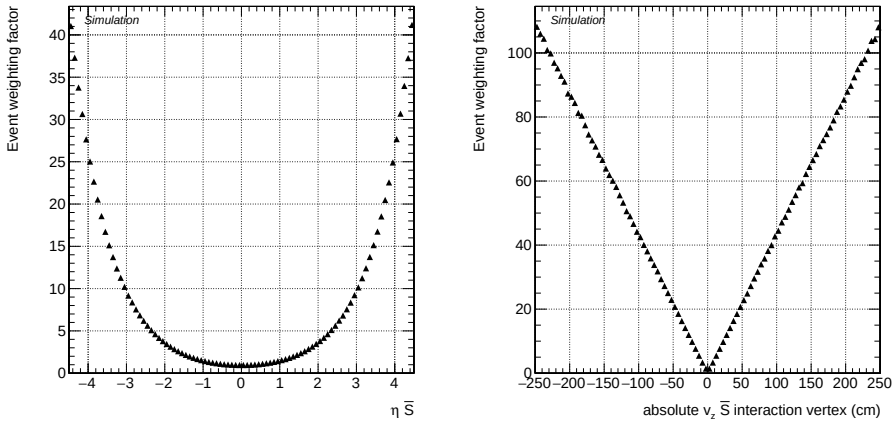
(a) (b)

FIGURE 10.6: Event display in the $r\phi$ (LEFT) and the rz-plane (RIGHT) of a simulated event where the $\bar{S}$ got reconstructed. At the center of the $r\phi$ plane the pixel barrel is illustrated and the outer circle represents the tracker barrel. The paths of the simulated particles are shown here.



(a) (b)

FIGURE 10.7: Event display in the $r\phi$ (LEFT) and the rz-plane (RIGHT) of a simulated event where the $\bar{S}$ got reconstructed. At the center of the $r\phi$ plane the pixel barrel is illustrated and the outer circle represents the tracker barrel. The paths of the simulated particles are shown here.



(a) (b)

FIGURE 10.8: In the simulation all $\bar{S}$ interact in the beampipe due to the iterative transport. In reality however the interaction probability scales with the path length through the beampipe and an event-by-event weight factor has to be applied. This weight factor is defined as $1/\sin(\theta_{\bar{S}})$ and shown here in function of η and the z location of the interaction vertex for generated $\bar{S}$.

The momentum distribution of nucleons in a ^{9}Be nucleus, as provided by the framework described in [117], was implemented for the $\bar{S}$-neutron interaction. The implemented neutron momentum distribution is shown in Figure 10.10a. Beryllium was used here as this is the material the beampipe is made of.

The effect of the neutron momentum on the invariant mass calculation of the $\bar{S}$ is rather large and illustrated in Figure 10.10b[7], where the correlation between $m_{\bar{S},\text{obs}}$ and the size of the vectorial sum of momenta of the K^0_S and $\bar{\Lambda}^0$ are shown. $m_{\bar{S},\text{obs}}$ is the $\bar{S}$ invariant mass calculated under the assumption that the target neutron is at rest. For low momenta of the K^0_S-$\bar{\Lambda}^0$ system, the invariant mass is indeed calculated to be $\approx 1.8\,\text{GeV}/c^2$, as expected. With increasing momenta however, the spread on the invariant mass becomes significant and even negative invariant masses can be obtained. Negative invariant masses are defined for events where the sum of the energies squared of the final state particles is smaller than the vectorial sum of the final state particles squared. $m_{\bar{S},\text{obs}}$ is shown here with respect to the sum of momenta of the K^0_S-$\bar{\Lambda}^0$ system as this vector can be reconstructed and is the best proxy for the actual momentum of the $\bar{S}$, which cannot be directly reconstructed. This is illustrated in Figure 10.11, where the influence of the neutron momentum on the reconstruction of the actual $\bar{S}$ direction of propagation is shown in the $\Delta R(=\sqrt{\Delta\phi^2+\Delta\eta^2})$ variable. This reduces the accuracy with which the $\bar{S}$ can be reconstructed as pointing to the luminous region.

The counter-intuitive result, where the deviation of $m_{\bar{S},\text{obs}}$ from the true $\bar{S}$ mass scales with the momentum of the incoming $\bar{S}$, can also be shown analytically.
For an $\bar{S}$ with energy $E_{\bar{S}}$ and momentum $\vec{p}_{\bar{S}}$, annihilating on a neutron (n) at rest with mass m_n where a K^0_S and $\bar{\Lambda}^0$ are formed with respective energies $E_{K^0_S}$ and $E_{\bar{\Lambda}^0}$ and momenta $\vec{p}_{K^0_S}$ and $\vec{p}_{\bar{\Lambda}^0}$, the following is valid:

[7] The figures shown in this section are for reconstructable $\bar{S}$ events as defined in section 10.10.5.

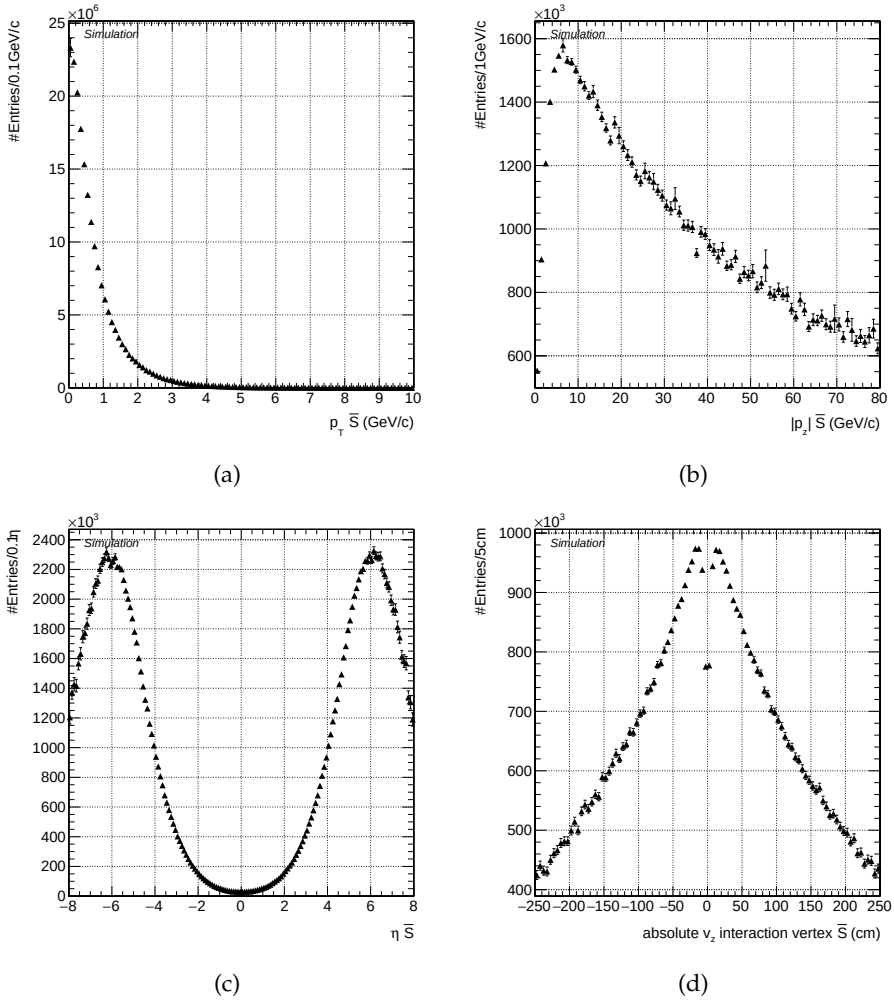
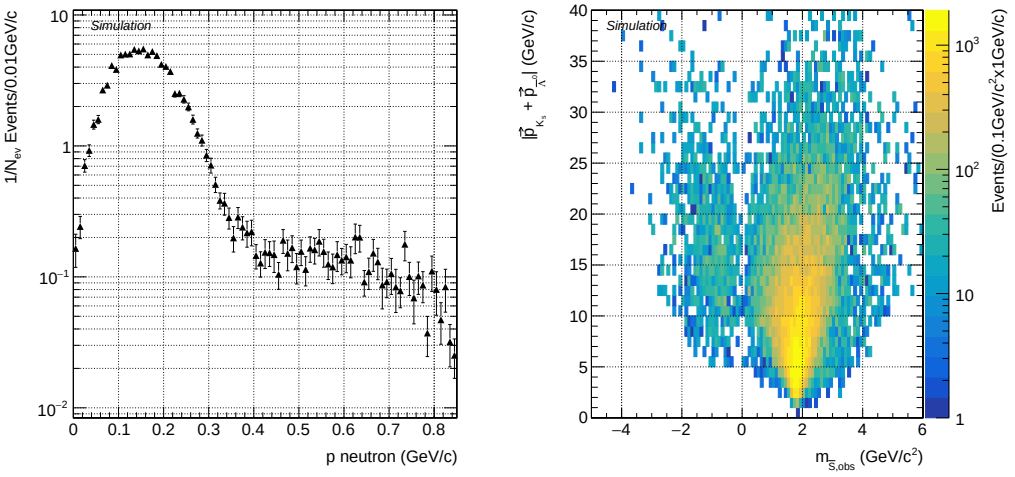


FIGURE 10.9: The reweighed transverse and longitudinal momenta of the $\bar{S}$, pseudorapidity of the $\bar{S}$ and z location of its interaction vertex at generator level.



(a) (b)

FIGURE 10.10: LEFT: The momentum distribution of a neutron [117] in beryllium. RIGHT: The influence of the neutron momentum on the invariant mass reconstruction of the $\bar{S}$ extracted at generator level in reconstructable events.

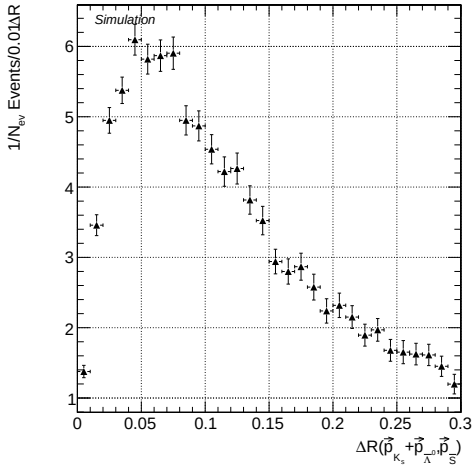


FIGURE 10.11: The influence of the target neutron momentum on the direction of propagation of the $\bar{S}$, extracted at generator level in reconstructable events.

$$(E_{\bar{S}}, \vec{p}_{\bar{S}}) + (m_n, 0) = \left(E_{\bar{\Lambda}^0} + E_{K_S^0}, \vec{p}_{\bar{\Lambda}^0} + \vec{p}_{K_S^0}\right), \tag{10.8}$$

which results in an invariant mass of the $\bar{S}$:

$$m_{\bar{S},obs} = \sqrt{\left(E_{\bar{\Lambda}^0} + E_{K_S^0}\right)^2 + m_n^2 - 2m_n\left(E_{\bar{\Lambda}^0} + E_{K_S^0}\right) - (\vec{p}_{\bar{\Lambda}^0} + \vec{p}_{K_S^0})^2}. \tag{10.9}$$

In reality, the neutron is not at rest and is moving with a momentum $\vec{p}_n$ in the nucleus. Equation (10.8) then becomes:

$$(E_{\bar{S}}, \vec{p}_{\bar{S}}) + (E_n, \vec{p}_n) = \left(E_{\bar{\Lambda}^0} + E_{K_S^0}, \vec{p}_{\bar{\Lambda}^0} + \vec{p}_{K_S^0}\right), \tag{10.10}$$

which leads to an invariant mass of the $\bar{S}$:

$$m_{\bar{S}} = \sqrt{\left(E_{\bar{\Lambda}^0} + E_{K_S^0}\right)^2 + m_n^2 - 2E_n\left(E_{\bar{\Lambda}^0} + E_{K_S^0}\right) - (\vec{p}_{\bar{\Lambda}^0} + \vec{p}_{K_S^0})^2 + 2\vec{p}_n(\vec{p}_{\bar{\Lambda}^0} + \vec{p}_{K_S^0})}. \tag{10.11}$$

If we assume that the momenta of the K_S^0 and $\bar{\Lambda}^0$ are collinear and substitute $m^2_{\bar{S},obs}$ from Equation (10.9) in Equation (10.11), we get:

$$\begin{aligned} m^2_{\bar{S},obs} &= m^2_{\bar{S}} + 2\left(E_n - m_n\right)\left(E_{\bar{\Lambda}^0} + E_{K_S^0}\right) - 2p_n\left(p_{\bar{\Lambda}^0} + p_{K_S^0}\right) \\ &= m^2_{\bar{S}} + 2\left(E_n - m_n\right)\left(E_{\bar{S}} + E_n\right) - 2p_n\left(p_{\bar{S}} + p_n\right) \end{aligned}, \tag{10.12}$$

which shows that $(m^2_{\bar{S},obs} - m^2_{\bar{S}})$ is a function of the $\bar{S}$ momentum.

The fact that the neutron momentum smears the invariant mass peak of the $\bar{S}$ vouches for a zero background analysis as it will not be possible to extract a narrow mass-peak.

10.10.5 Reconstructability

When discussing the reconstruction efficiency for the signal, it is interesting to try to factorise the efficiency losses due to detector acceptance on the one hand and detector inefficiencies and algorithmic limitations on the other hand. Three factors are therefore identified which contribute to the signal reconstruction efficiency:

1. Detector acceptance: all four final state particles should fall in the tracker acceptance. This requirement can be evaluated by counting the number of hits that the particle would generate in the tracker modules. A number of hits equal or larger than 7 is required to evaluate this acceptance criteria. This condition is used as it is the same number of hits that is required for tracks entering the V^0 reconstruction module (section 10.9).

2. Branching fraction: as stated in section 10.7 the branching fraction for $K_S^0 \to \pi^+\,\pi^-$ is 69.2% and for $\bar{\Lambda}^0 \to \pi^+\,\bar{p}$ is 63.9%. The combined branching fraction of K_S^0 and $\bar{\Lambda}^0$ to charged particles is thus 44.2%. $K_S^0 \to \pi^0\,\pi^0$ and $\bar{\Lambda}^0 \to n\pi^0$ complement

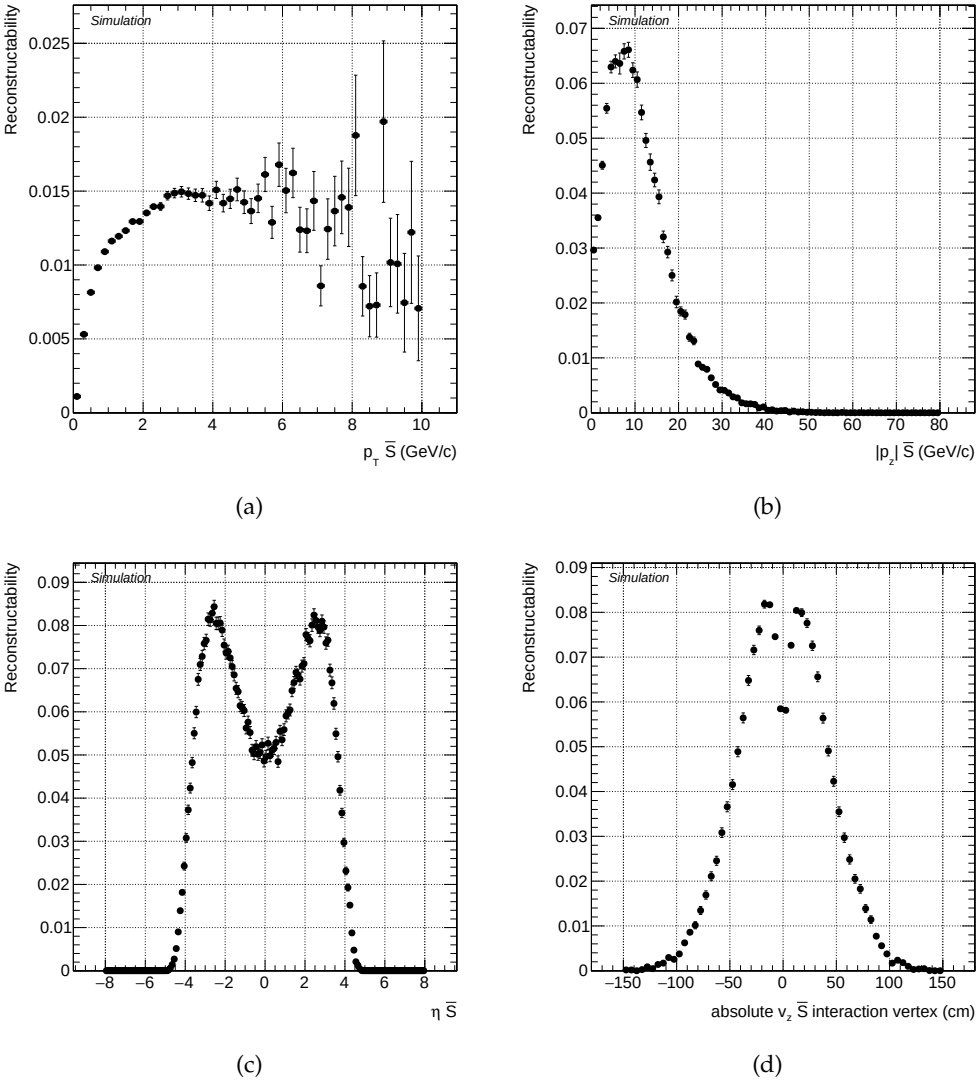


FIGURE 10.12: *Reconstructability* of generated $\bar{S}$ as function of the $\bar{S}$ transverse momentum, longitudinal momentum, pseudorapidity and z coordinate of the interaction vertex.

the K^0_S and $\bar{\Lambda}^0$ decay modes used in this analysis almost completely with respective branching fractions of 30.7% and 35.8%. Other charged final states are thus negligible.

3. Detection efficiency: it is not guaranteed that an $\bar{S}$ event which satisfies the above two requirements will be reconstructed. These reconstruction inefficiencies will be the subject of section 10.11.

From now on, the acceptance requirement, as in point 1 above, and the branching fraction, as in point 2 above, will be combined and an $\bar{S}$ fulfilling these requirements will be referred to as *reconstructable*.

Figure 10.12 shows the *reconstructability* of the $\bar{S}$, i.e. the probability for an $\bar{S}$ to be reconstructable, with respect to the $\bar{S}$ transverse momentum, longitudinal momentum, pseudorapidity and z coordinate of its interaction vertex. It shows that the reconstructability of $\bar{S}$ vanishes for $\bar{S}$ with $|\eta| \gtrapprox 4.5$ and $|z| \gtrapprox 140$ cm. Also a minimal transverse momentum is required in order for the $\bar{S}$ to be reconstructable. The overall reconstructability was found to be 0.75% which shows that acceptance criteria are indeed an important inefficiency factor. The results shown in the rest of this section are for reconstructable $\bar{S}$.

10.10.6 V^0 and final state particles kinematics

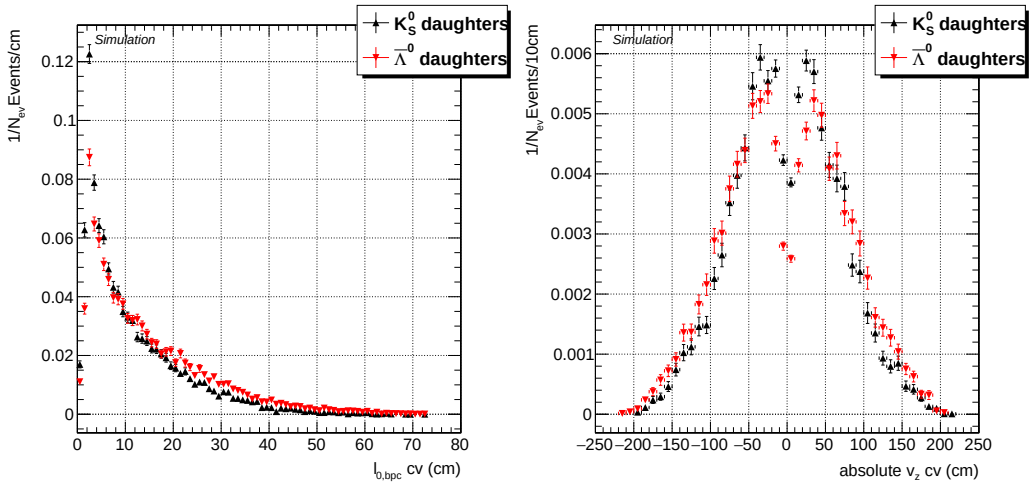
Before studying actual signal reconstruction efficiencies, it is instructive to look at the kinematics of the final state particles in reconstructable events. In this section it will become clear that the $\bar{S}$ annihilation produces final state particles which have a very unusual topology compared to final states in typical CMS analyses.

The transverse and longitudinal displacement of the K^0_S and $\bar{\Lambda}^0$ decay vertices are shown in Figure 10.13 and in two dimensions in Figures 10.14 and 10.15 for the K^0_S and $\bar{\Lambda}^0$ respectively. These figures basically represent a convolution of the $\bar{S}$ interaction vertices with the decay lengths of the K^0_S and the $\bar{\Lambda}^0$.

The actual momentum distributions of the K^0_S and $\bar{\Lambda}^0$ and their respective daughters are shown in Figure 10.16 and 10.17 respectively. Important to note here is the low p_T of the final state particles and especially the low momentum of the soft pion in the $\bar{\Lambda}^0$ decay. For the π^+ in the $\bar{\Lambda}^0$ decay, the transverse and longitudinal momentum are also shown in Figure 10.18 as function of the pseudorapidity of the particle. It shows that the majority of the π^+ are produced forward.

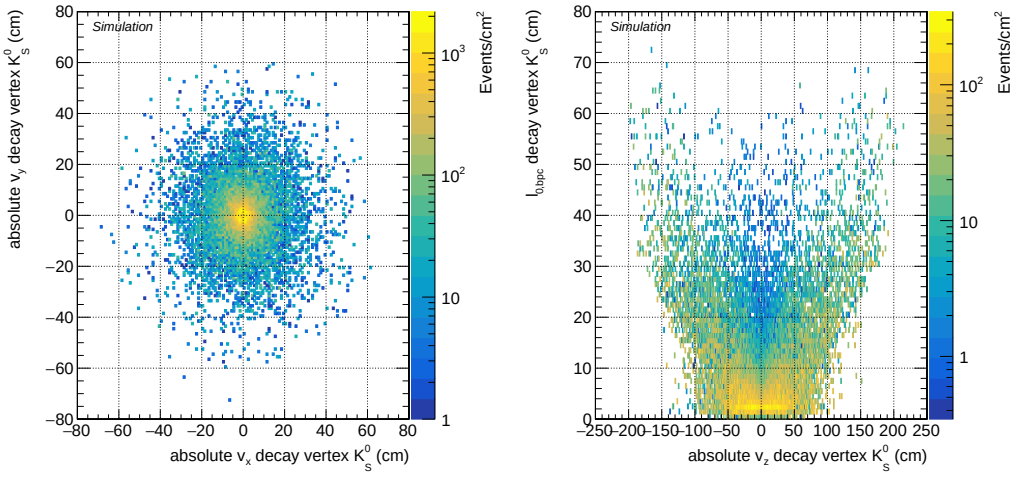
Besides being produced displaced and having low momenta, the final state particles are also not necessarily pointing to the luminous region as shown in Figures 10.19 where the transverse (d_0) and longitudinal (d_z) impact parameters with respect to the beamspot (bs) are shown calculated at the creation vertex of the final state particles. The sign of d_0 is given by the scalar product of the track momentum vector and the vector connecting the beamspot and the track creation vertex.

The above discussion shows that the final state particles have low momenta, are produced mainly at high $|z|$ and they do not necessarily point to the luminous region. These facts imply that the final state particles do not necessarily benefit from the tracker layout which was optimised to have particles, created at the primary vertex with a significant p_T pass as many tracker and seeding layers as possible.



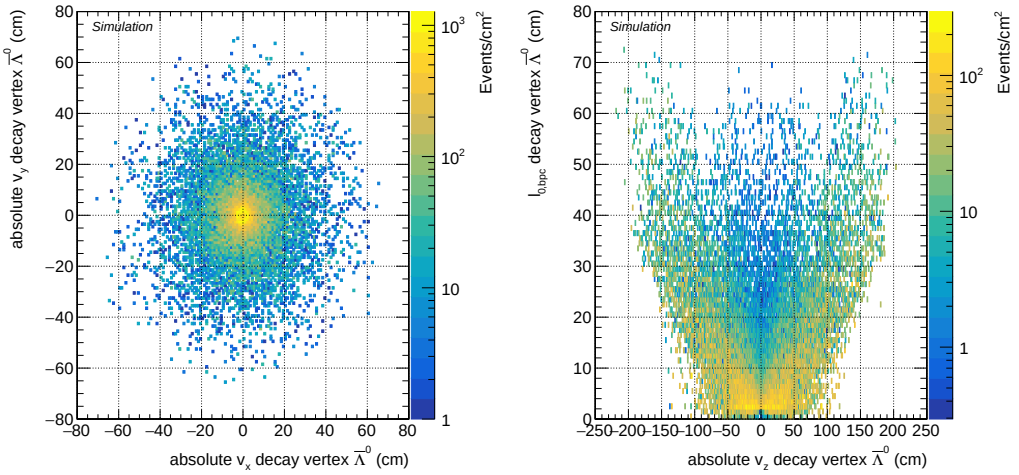
(a) (b)

FIGURE 10.13: Displacement of the decay vertices of the V^0s (creation vertices (cv) of the final state particles) with respect to the beam pipe center (bpc) and $z = 0$ respectively. This is obtained at generator level in reconstructable events.



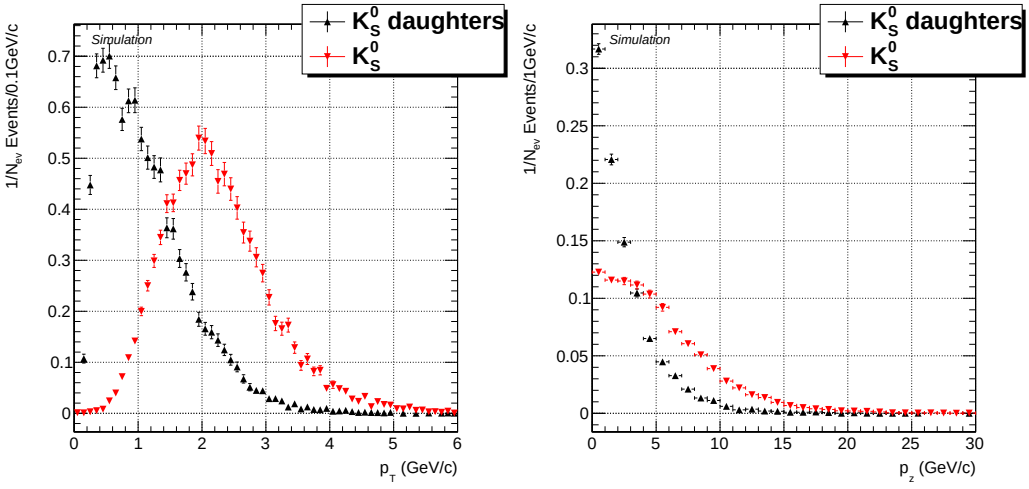
(a) (b)

FIGURE 10.14: Decay vertices of the K^0_S (creation vertices of the final state particles) at generator level in reconstructable events.



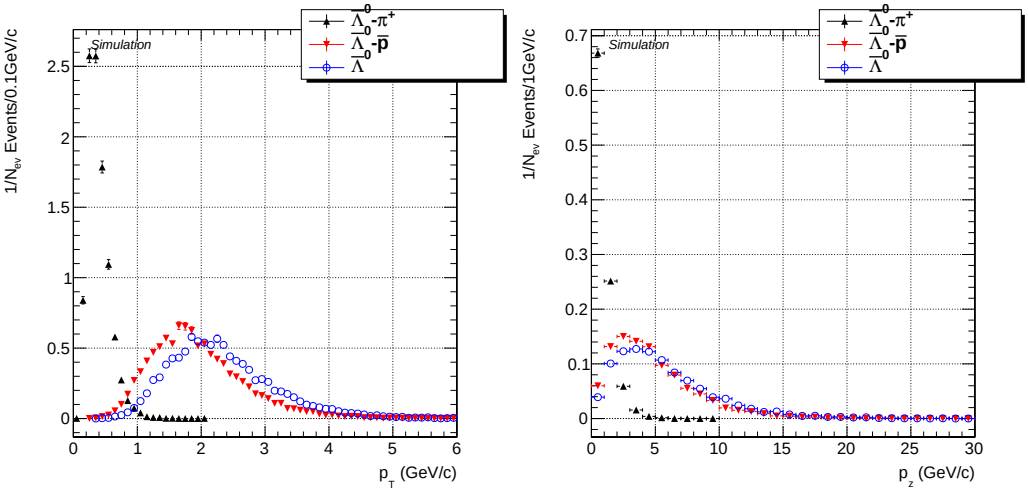
(a) (b)

FIGURE 10.15: Decay vertices of the $\bar{\Lambda}^0$ (creation vertices of the final state particles) at generator level in reconstructable events.



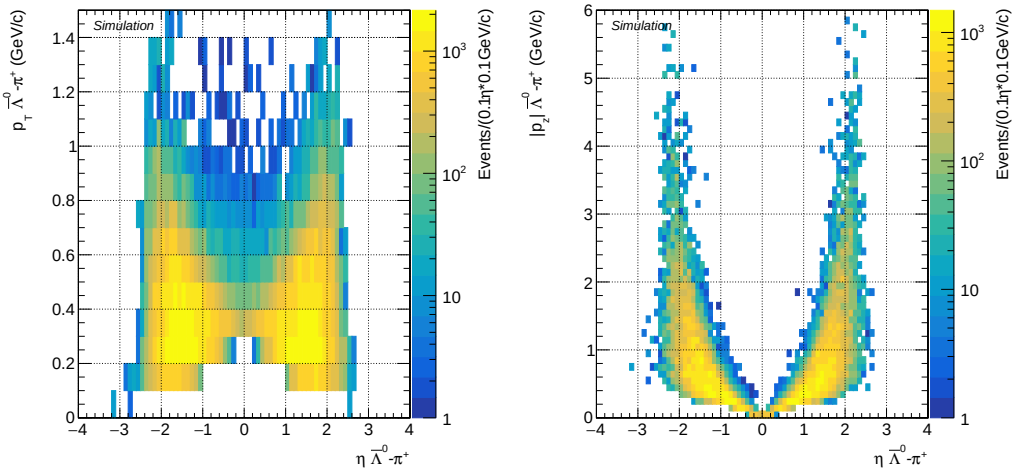
(a) (b)

FIGURE 10.16: Momenta of the K^0_S and its daughters at generator level in reconstructable events.



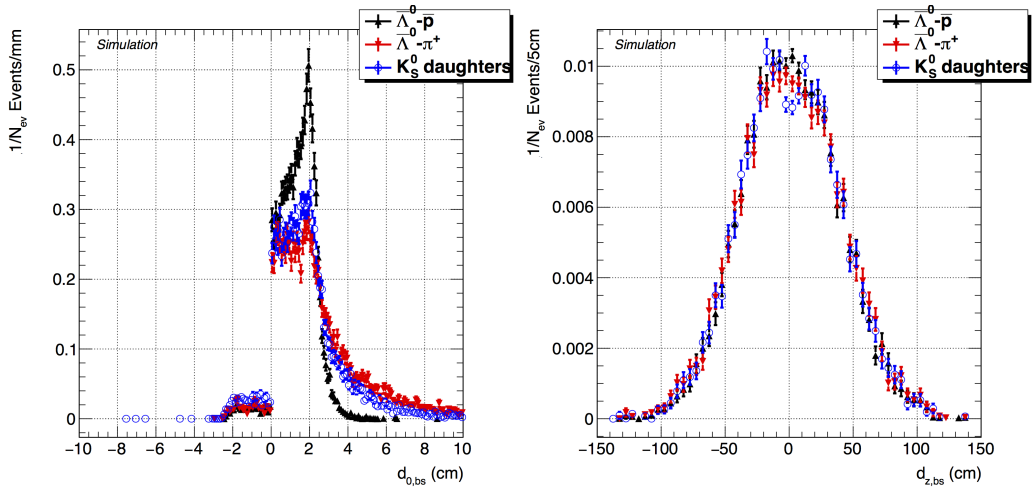
(a) (b)

FIGURE 10.17: Momenta of the $\bar{\Lambda}^0$ and its daughters at generator level in reconstructable events.



(a) (b)

FIGURE 10.18: Momenta of the π^+ produced in the $\bar{\Lambda}^0$ decay as function of the π^+ pseudorapidity at generator level in reconstructable events.



(a) (b)

FIGURE 10.19: Impact parameters of the final state particles at generator level in reconstructable events.

Particle	Matching criteria
Final state particles	Hit matching: $> 75\%$ of the reconstructed track hits should be matched to the simulated track
K^0_S	$\Delta L_{\text{xyz,dv}} < 2$ cm and $\Delta R(\vec{p}_{\text{SIM}}, \vec{p}_{\text{RECO}}) < 0.03$
$\bar{\Lambda}^0$	$\Delta L_{\text{xyz,dv}} < 3$ cm and $\Delta R(\vec{p}_{\text{SIM}}, \vec{p}_{\text{RECO}}) < 0.03$
$\bar{\text{S}}$	All final state particles and V^0s matched, $\Delta R(\vec{p}_{\text{SIM},\bar{\text{S}}}, \vec{p}_{\text{RECO},\bar{\text{S}}}) < 0.5$ and $\Delta L_{\text{xyz,iv}} < 2$ cm

TABLE 10.2: Matching criteria between simulated and reconstructed particles in an $\bar{\text{S}}$ event.

10.11 Signal reconstruction efficiency

In this section, inefficiencies in $\bar{\text{S}}$ reconstruction in events which pass the reconstructability requirement, defined in the previous section, are investigated using the signal simulation. The reconstruction efficiency of all particles in the event will be evaluated by matching simulated (*SIM*) particles with reconstructed (*RECO*) particles. The matching criteria between simulated and reconstructed particles are given in Table 10.2. The matching for charged particles is performed based on tracker hits[8]. For the V^0 particles the distance between the simulated and reconstructed decay vertex ($\Delta L_{\text{xyz,dv}}$) together with the angular separation between simulated and generated momentum direction (ΔR) is used for matching. For the $\bar{\text{S}}$, on top of requirements on $\Delta R(\vec{p}_{\text{SIM},\bar{\text{S}}}, \vec{p}_{\text{RECO},\bar{\text{S}}})$ and the distance between the simulated and reconstructed interaction vertex ($\Delta L_{\text{xyz,iv}}$), also matching for both the V^0s and all tracks in the event is required.

[8] A ΔR matching is not possible here due to the signal tracks being displaced. As a result the reference points of the SIM and RECO track are not the same.

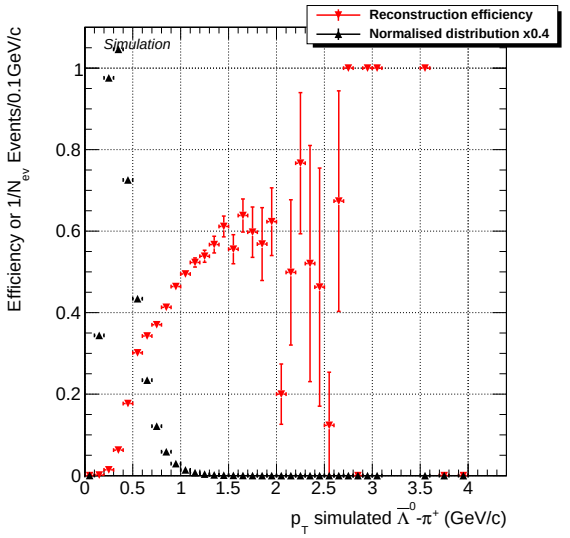


FIGURE 10.20: Tracking efficiency for the π^+ in the $\bar{\Lambda}^0$ decay in reconstructable events as function of the particle's transverse momentum.

These matching criteria can now be used to evaluate reconstruction efficiencies in function of different kinematic variables. The kinematic variables which are of interest are the displacement of the particle, its transverse momentum and its pseudorapidity. The reconstruction efficiencies with respect to these variables can be found for all particles in an $\bar{S}$ event in Appendix D. Some of the most interesting results are shown and discussed below. It should be mentioned that these results are not always straightforward to interpret as the efficiency for a certain kinematic variable always integrates the inefficiencies in the other variables as well.

The softest among all final state particles is the π^+ in the $\bar{\Lambda}^0$ decay. The momentum of the π^+ in reconstructable events is shown together with the reconstruction efficiency in Figure 10.20. It is immediately clear that a large fraction of the reconstructable π^+ will not be reconstructed due to the low transverse momentum of this particle. Good to remember here is that particles with a transverse momentum of 0.25 GeV/c will travel on a trajectory with a bending radius of only ≈ 22 cm.

Another interesting result is the reconstruction efficiency with respect to the transverse displacement of the track's creation vertex. This is shown for the $\bar{p}$ in the $\bar{\Lambda}^0$ decay on the top left in Figure 10.21. The efficiency is rather high until ≈ 65 cm after which it steeply falls. This can be explained by the fact that no more track seeding is possible at these large distances. Indeed, the last tracker modules which are used to extract track seeds are located at a radius of ≈ 70 cm as discussed in section 4.3.2. The decreased reconstruction efficiency for tracks created at ≈ 35 cm lateral displacement is explained by the fact that for a radius between ≈ 40 cm and ≈ 60 cm no stereo-modules are present. The figure on the top right in Figure 10.21 shows that the tracking is most efficient at small $|z|$, where particles traverse the pixel barrel layers. Most of the particles are however produced at higher $|z|$ (≈ 40 cm) which is where the tracking efficiency already starts to decrease sharply. The figure at the bottom of Figure 10.21 illustrates the above conclusions in a more qualitative way.

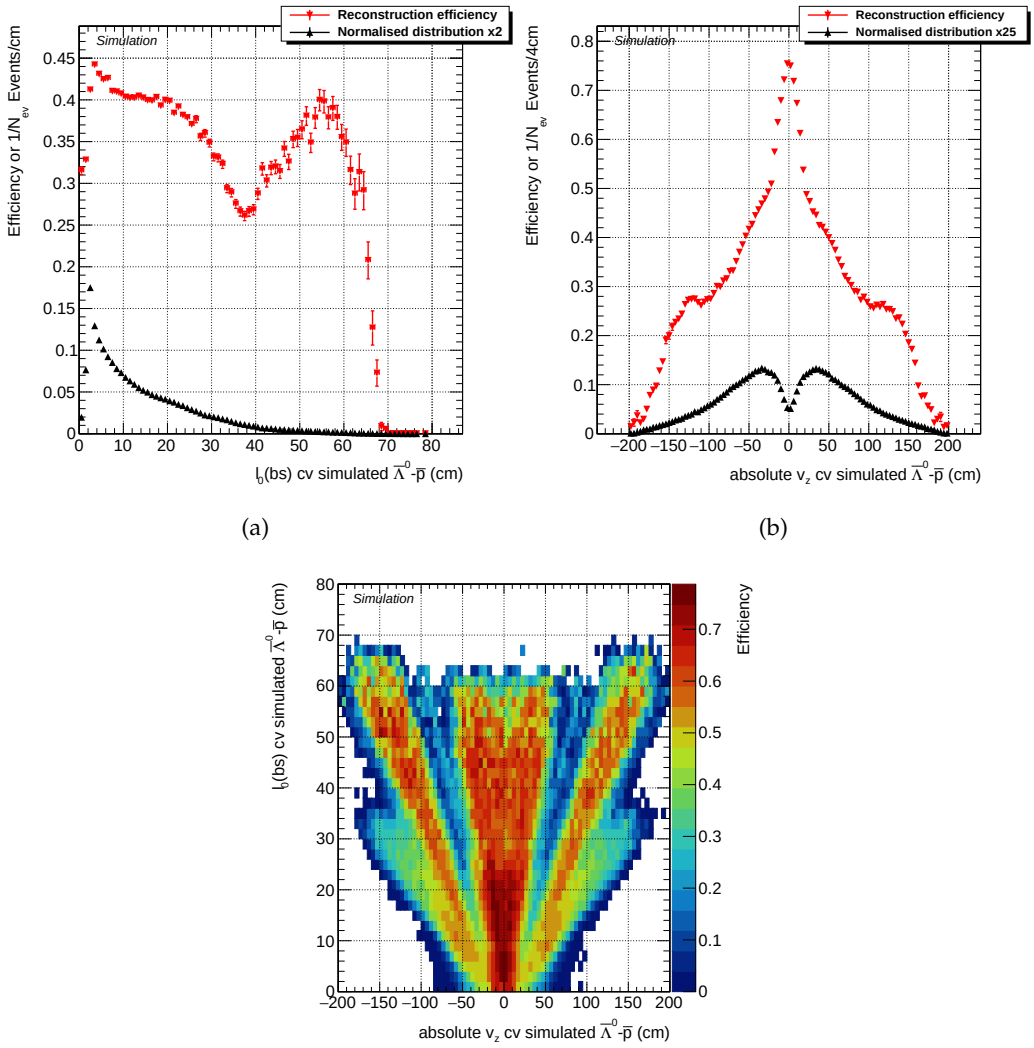
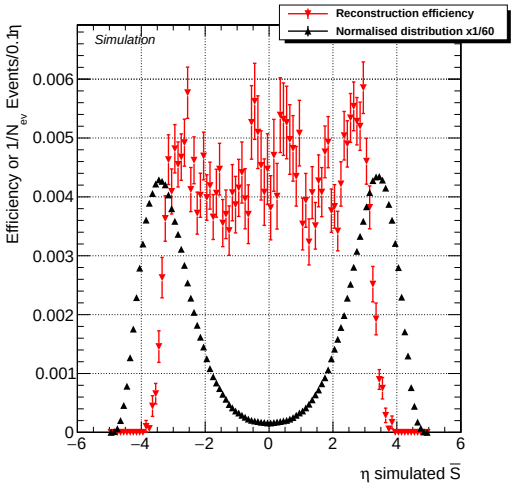
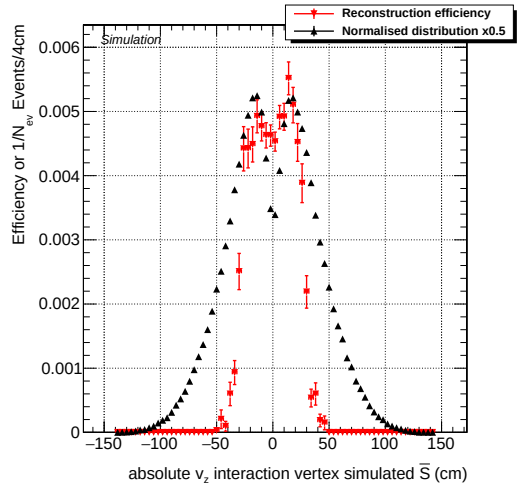


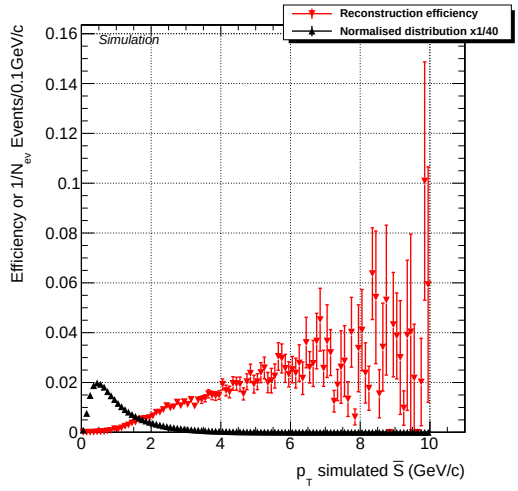
FIGURE 10.21: Tracking efficiency for the $\bar{p}$ in the $\bar{\Lambda}^0$ decay in reconstructable events as function of the displacement of the particle's creation vertex (cv). The actual distribution of the variables is also shown for the two figures in the top row.



(a)



(b)



(c)

FIGURE 10.22: Reconstruction efficiency for the $\bar{S}$ in reconstructable events. The actual distribution of the variables is also shown.

$\bar{S}$ reconstructability	0.751%
K^0_S-π^+ reconstruction efficiency in reconstructable events	31.1%
K^0_S-π^- reconstruction efficiency in reconstructable events	31.3%
$\bar{\Lambda}^0$-π^+ reconstruction efficiency in reconstructable events	12.9%
$\bar{\Lambda}^0$-$\bar{p}$ reconstruction efficiency in reconstructable events	39.2%
K^0_S reconstruction efficiency in reconstructable events	6.54%
$\bar{\Lambda}^0$ reconstruction efficiency in reconstructable events	2.78%
$\bar{S}$ reconstruction efficiency in reconstructable events	0.194%
Overall $\bar{S}$ reconstruction efficiency (ϵ_{reco})	0.0014%

TABLE 10.3: Reconstructability and reconstruction efficiencies for the track, V^0 and $\bar{S}$ reconstruction.

The reconstruction efficiencies for the V^0 particles, shown in Figure D.5 and D.6 for the $\bar{\Lambda}^0$ and the K^0_S respectively, show similar features as at the tracking level. Figure 10.22 shows the eventual reconstruction efficiency for the $\bar{S}$. It is clear from both the first and second figure that a significant part of the $\bar{S}$ is produced too forward and thus interacts at too high $|z|$ to get reconstructed. The third figure shows that, as expected, $\bar{S}$ with a significant transverse momentum have the highest reconstruction efficiency.

The breakdown of the overall reconstruction efficiency into reconstructability and reconstruction efficiencies of reconstructable objects is given in Table 10.3. The overall reconstruction efficiency of the $\bar{S}$, of 0.0014%, is thus very low and implies that, although a large dataset is available to look for the $\bar{S}$, the discovery potential is significantly reduced by this inefficient reconstruction.

Important for the background suppression is the accuracy with which the $\bar{S}$ interaction vertex can be reconstructed. The location of reconstructed interaction vertices are shown in the top row in Figure 10.23. The reconstruction indeed gives us back the beampipe radius and Figure 10.23 furthermore shows that no $\bar{S}$ are reconstructed which interact at $|z|$ larger than 50 cm. The actual accuracy of the reconstruction is shown on the bottom row of Figure 10.23 where Monte Carlo truth is compared to reconstructed objects. $\mathcal{O}$(1 mm) accuracy can be attained on both the transverse and longitudinal displacement.

Figure 10.24 shows the distribution of the estimated invariant mass of the $\bar{S}$. It shows that also $\bar{S}$ with negative invariant mass can be reconstructed, and shows the large spread on the invariant mass introduced by the neutron momentum.

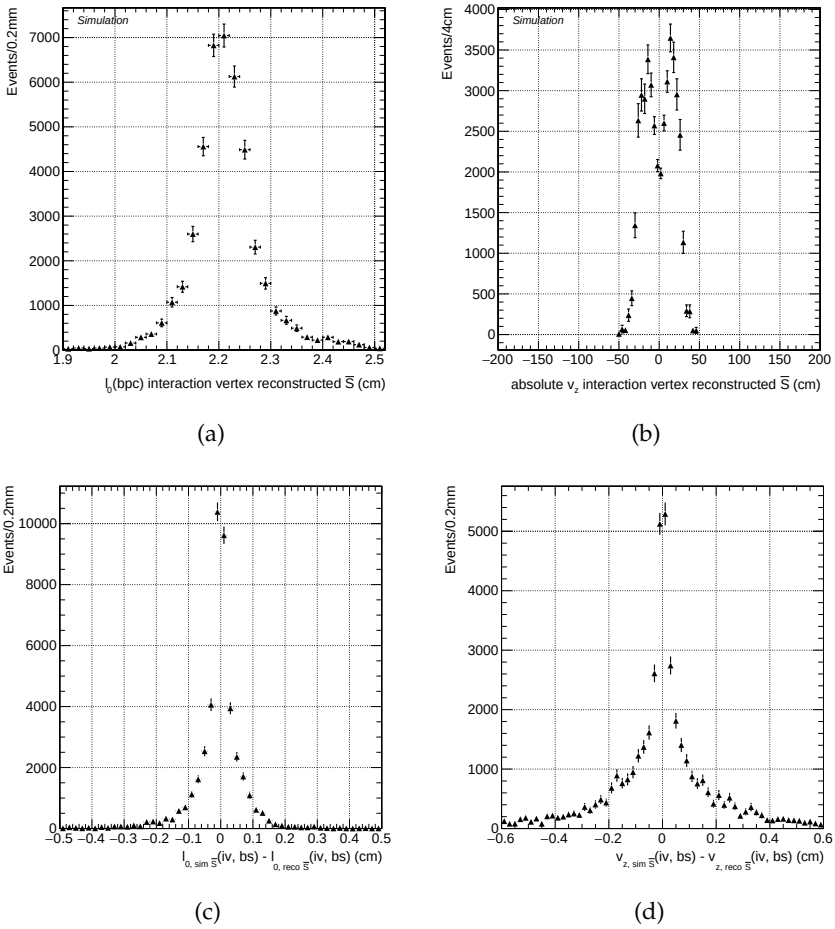


FIGURE 10.23: Location of the $\bar{S}$ reconstructed interaction vertices (TOP) and their reconstruction accuracy (BOTTOM).

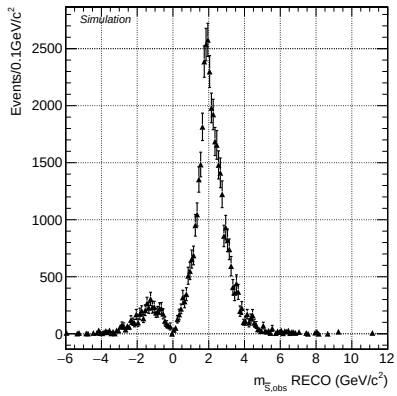


FIGURE 10.24: Invariant mass of the reconstructed $\bar{S}$ calculated under the assumption that the target neutron is at rest.

10.12 Background rejection

The background for the $\bar{S}$ signal is evaluated by looking at reconstructed S particles. This collection does not contain signal since S, as opposed to $\bar{S}$, do not annihilate in the beampipe. Throughout this section, five background samples are compared to each other:

- S reconstructed in the *DYJets* MC sample, which was also used for the systematics study on tracking efficiencies (see section 10.13).
- $\bar{S}$ reconstructed in the *DYJets* MC sample, which was also used for the systematics study on tracking efficiencies (see section 10.13).
- Reconstructed S from *SingleMuon_Run2016H*. This is the default background sample.
- Reconstructed $\bar{S}$ from *SingleMuon_Run2016H* where it is guaranteed that this is background by placing a maximum on the value of the BDT classifier.
- $\bar{S}$ which are reconstructed across events: $\bar{\Lambda}^0$ are fitted to a common vertex with K_S^0 from the previous event to form $\bar{S}$ candidates.

The first four background samples are compared to validate that the S background candidates, reconstructed in data, form a good reference for the $\bar{S}$ background in data. This comparison will be made for all the kinematic variables shown in this section. The last background sample is used to show that the background is due to random combinations of K_S^0 and $\Lambda^0/\bar{\Lambda}^0$ which form $S/\bar{S}$ candidates. The fact that the background is indeed *combinatorial* will be discussed later in this section, but it is relevant to already be aware of this fact when interpreting the background distributions shown in this section.

The background suppression is achieved by first introducing three background rejection cuts which show an obvious difference between the signal and the background. To further eliminate the background, a BDT is used. An overview of the introduced cuts which will be introduced below can be found in the cut flow table (Table 10.4) for the signal simulation and the background extracted from data[9]. In what follows, reconstructed S and $\bar{S}$ candidates are investigated that have the parameters of their final state particles within a fiducial region, which will be introduced in section 10.13, for which tracking efficiencies are well described in simulation.

The first background rejection cut is applied to remove the hard component in the event by requesting the absolute angular difference in ϕ of the K_S^0 and the $\bar{\Lambda}^0$ or Λ^0 ($|\Delta\phi(K_S^0, \bar{\Lambda}^0/\Lambda^0)|$) to be larger than 0.5. The distribution of this variable is shown on the left in Figure 10.25 for events which have their final state particles in the fiducial region as defined in section 10.13. The $|\Delta\phi(K_S^0, \bar{\Lambda}^0/\Lambda^0)| > 0.5$ has an efficiency of 97.9% for signal and 48.9% for background.

To further suppress the background, the fact that the $\bar{S}$ reconstructed interaction vertex location should be compatible with the beampipe position can be exploited. It should be noted here that this distance is calculated differently in simulation and data as the center of the beampipe is at a different location for both. For simulation this distance is calculated with respect to the absolute origin, whereas for data this distance is calculated with respect to ($x = 0.124$ cm, $y = 0.027$ cm) which is the measured beampipe center [115].

[9] No check for duplicate events due to overlaps in primary datasets was performed when constructing this table.

	Signal MC: # RECO $\bar{S}$ (weighed)	All considered datasets: #RECO S
Reconstructed S or $\bar{S}$ candidates	100,000	29,524,230
Fiducial region for systematic uncertainties	58,339	12,653,661
$\lvert\Delta\phi(K^0_S, \bar{\Lambda}^0/\Lambda^0)\rvert > 0.5$	57,007	6,188,455
2.02 cm $< l_{0,\mathrm{bpc}}(\mathrm{iv}) <$ 2.40 cm	55,111	1,258,606
$0 \leq d_{0,\mathrm{bs}}(\bar{S}/S)/l_{0,\mathrm{bs}}(\bar{S}/S) < 0.5$	55,053	402,043
BDT classifier value > 0.35	28,784	7

TABLE 10.4: Cut flow table for signal $\bar{S}$ reconstructed in signal simulation (arbitrarily normalised) and S reconstructed in all considered datasets. For the latter no check for duplicate events due to overlaps in primary datasets was performed when constructing this table.

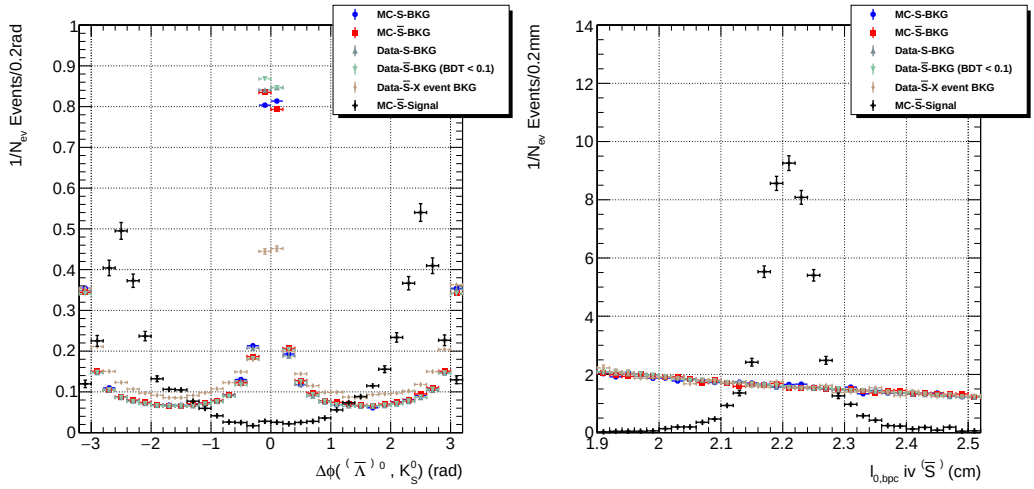
The distribution of the location of the lateral location of the reconstructed interaction vertices are shown on the right in Figure 10.25 for events which have their final state particles in the fiducial region as defined in section 10.13 and where the $|\Delta\phi(K^0_S, \bar{\Lambda}^0/\Lambda^0)| > 0.5$ cut has been applied. A cut is introduced which requires the events to have this distance in the range from 2.02 to 2.40 cm. This cut has an efficiency of 96.6% for signal and 20.3% for background after applying the $|\Delta\phi(K^0_S, \bar{\Lambda}^0/\Lambda^0)| > 0.5$ cut.

On top of these two cuts a third background rejection cut can be defined based on the transverse impact parameter of the S or $\bar{S}$ normalised to its transverse displacement, both calculated with respect to the beamspot: $d_{0,\mathrm{bs}}(\bar{S}/S)/l_{0,\mathrm{bs}}(\bar{S}/S)$. The distribution of this variable is shown in Figure 10.26 for S and $\bar{S}$ with final state particles in the fiducial region defined in section 10.13, $|\Delta\phi(K^0_S, \bar{\Lambda}^0/\Lambda^0)| > 0.5$ and 2.02 cm $< l_{0,\mathrm{bpc}}(\mathrm{iv}) <$ 2.40 cm. It shows that for signal only a fraction of the $\bar{S}$ is reconstructed with $d_{0,\mathrm{bs}}(\bar{S}/S)/l_{0,\mathrm{bs}}(\bar{S}/S) < 0$ or $d_{0,\mathrm{bs}}(\bar{S}/S)/l_{0,\mathrm{bs}}(\bar{S}/S) > 0.5$. We therefore introduce the cut: $0 \leq d_{0,\mathrm{bs}}(\bar{S}/S)/l_{0,\mathrm{bs}}(\bar{S}/S) < 0.5$. After applying the other cuts defined before this requirement is 99.8% efficient for signal and 31.9% efficient for background.

To further discriminate signal and background a boosted decision tree is used. A BDT is a multivariate analyses machine learning technique commonly applied for classification problems [118]. In particle physics, BDTs are usually applied to discriminate signal from background. These can be separated from each other in the multi-dimensional parameter space set up by the kinematic variables used as input to the BDT. A subset of the signal and the background data is used to *train* the BDT and to extract a single *classifier*.

The BDT exploits the differences between signal and background listed below. The distribution of these variables, which are used as input to the BDT, are given in Appendix E for $\bar{S}$ or S of which the final state particles lie within the fiducial region as defined in section 10.13, with $|\Delta\phi(K^0_S, \bar{\Lambda}^0/\Lambda^0)| > 0.5$, 2.02 cm $< l_{0,\mathrm{bpc}}(\mathrm{iv}) <$ 2.40 cm and $0 \leq d_{0,\mathrm{bs}}(\bar{S}/S)/l_{0,\mathrm{bs}}(\bar{S}/S) < 0.5$.
The BDT input parameters are:

- Location of the reconstructed interaction vertex of the S or $\bar{S}$ (Figures E.1).
- Angular correlation between the S ($\bar{S}$), K^0_S and Λ^0 ($\bar{\Lambda}^0$) (Figures E.2, E.3, E.4).
- η distributions of the K^0_S and $\bar{S}$ (S) (Figures E.5): the signal has a more forward topology than the background.



(a) (b)

FIGURE 10.25: LEFT: Angular difference in ϕ of the K^0_S and $\bar{\Lambda}^0$ from S or $\bar{S}$ candidates for $\bar{S}$ signal and five background samples for events which have all their final state particles in the fiducial region as defined in section 10.13. RIGHT: Transverse distance between the beampipe center and the reconstructed $\bar{S}$ or S candidate's interaction vertex for MC signal and five background samples. Only events are shown where the final state particles lie within the fiducial region as defined in section 10.13 and which have $|\Delta\phi(K^0_S, \bar{\Lambda}^0/\Lambda^0)| > 0.5$.

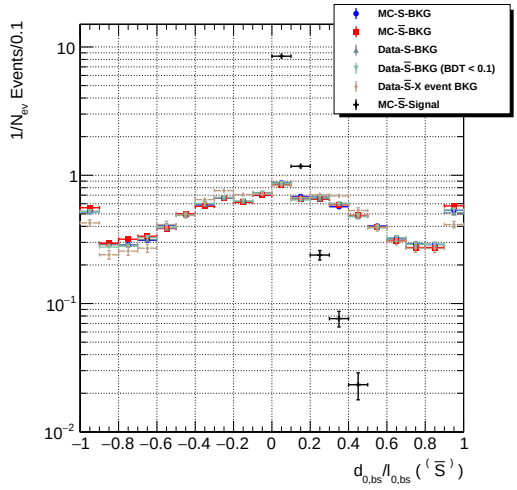
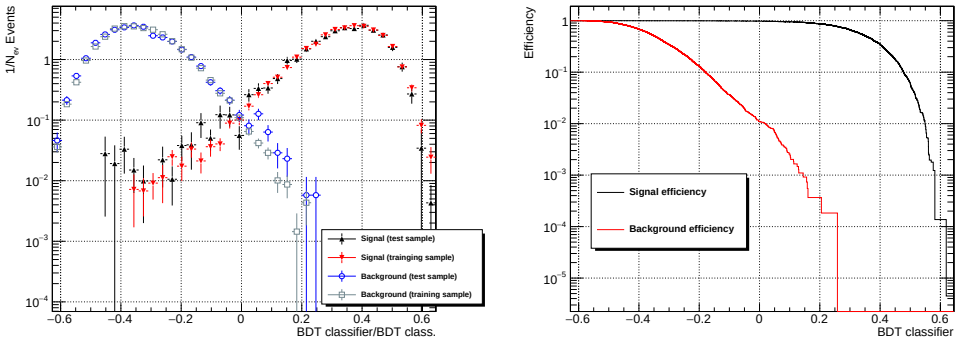


FIGURE 10.26: Longitudinal impact parameter of the $S/\bar{S}$ calculated with respect to the beampipe center divided by the transverse displacement of the $S/\bar{S}$ interaction vertex with respect to the beamspot for MC signal and five background samples. Only events are used where the final state particles lie within the fiducial region as defined in section 10.13 and which have $|\Delta\phi(K^0_S, \bar{\Lambda}^0/\Lambda^0)| > 0.5$ and 2.02 cm $< l_{0,\mathrm{bpc}}(\mathrm{iv}) <$ 2.40 cm.



(a) (b)

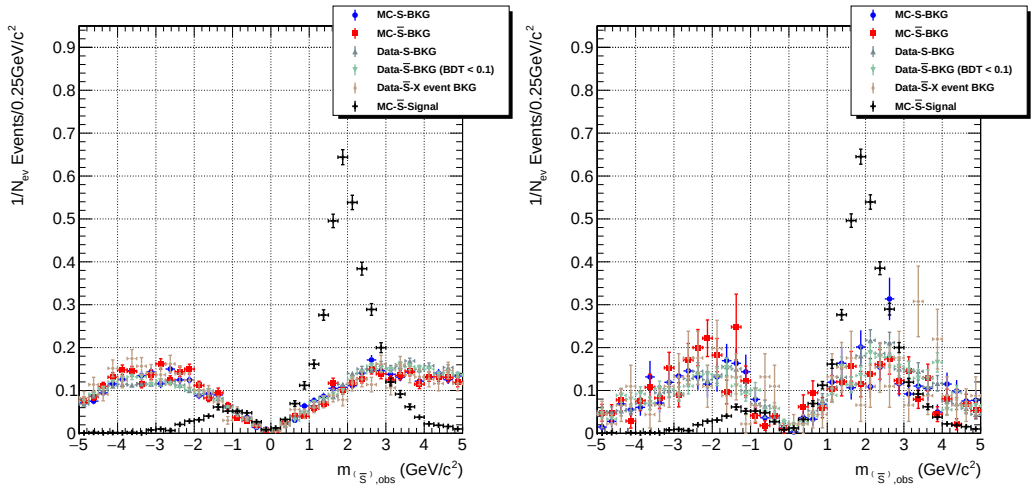
FIGURE 10.27: LEFT: BDT response for the background and signal training and testing sample. RIGHT: Data and background efficiency with respect to the BDT classifier.

- Pointing of the S ($\bar{S}$), K^0_S and Λ^0 ($\bar{\Lambda}^0$) to the primary vertex (Figures E.6 and Figures E.7): the reconstructed signal $\bar{S}$ point to the luminous region whereas the K^0_S and $\bar{\Lambda}^0$ originating from an $\bar{S}$ +n annihilation on average do not. The reconstructed background S are less probable to point back to the luminous region as they can be the result of randomly vertexed K^0_S and Λ^0s. The background K^0_S and Λ^0s in their turn do point back to the luminous region, the main origin of K^0_S and Λ^0. This topology can be exploited by calculating the transverse impact parameters of the K^0_S, $\Lambda^0/\bar{\Lambda}^0$ and $S/\bar{S}$ with respect to the beamspot and their longitudinal impact parameter with respect to the valid primary vertex that minimises this longitudinal impact parameter.
- The transverse momentum of the K^0_S (Figure E.8 left): the K^0_S from signal have a harder spectrum than the background.
- The transverse displacement of the Λ or $\bar{\Lambda}^0$ decay vertex (Figure E.8 right).
- The fit quality of the S or $\bar{S}$ vertex (Figure E.9): vertices of background S, if from random overlaps, are expected to have a larger χ^2/ndof.

Figures E.1-E.9 indeed show that in all kinematic variables the S background extracted from data follows the MC background samples. This vouches for the use of the S background sample extracted from data as reference for the $\bar{S}$ background. The use of the S sample from data as reference for the $\bar{S}$ background allows for a high statistics sample for training of the BDT.

The resulting BDT classifier is shown for both signal and background in Figure 10.27a. For training a signal MC sample of $\approx$ 12,600 (sum of weights $\approx$ 53,000) reconstructed $\bar{S}$ was used and for background $\approx$ 27,000 S reconstructed in the *SingleMuon_Run2016H* dataset. The signal and background efficiencies as function of the BDT classifier are shown in Figure 10.27b as obtained with the BDT testing sample.

As an extra validation, the variables used as input to the BDT are evaluated for background events of which the BDT classifier value is in the tail of the distribution towards the signal region. These distributions are given in Appendix F and again show that there



(a) (b)

FIGURE 10.28: Invariant mass distribution of S and $\bar{S}$ reconstructed in the five background samples and the signal MC sample for particles which have their final state particles within the fiducial region as defined in section 10.13, with $|\Delta\phi(K^0_S, \bar{\Lambda}^0/\Lambda^0)| > 0.5$, 2.02 cm $< l_{0,\mathrm{bpc}}(\mathrm{iv}) <$ 2.40 cm and $0 \leq d_{0,\mathrm{bs}}(\bar{S}/S)/l_{0,\mathrm{bs}}(\bar{S}/S) < 0.5$. In the right figure the BDT classifier value is furthermore requested to be $\geq$ -0.15.

is no strong difference between the data background sample and the MC background samples.

The invariant mass of the reconstructed S and $\bar{S}$ is explicitly not used in the BDT in order to keep this as a free parameter in the search. The invariant mass distributions are shown in Figure 10.28. The left figure shows the invariant mass distribution for $\bar{S}$ or S of which the final state particles lie within the fiducial region as defined in section 10.13, with $|\Delta\phi(K^0_S, \bar{\Lambda}^0/\Lambda^0)| > 0.5$, 2.02 cm $< l_{0,\mathrm{bpc}}(\mathrm{iv}) <$ 2.40 cm and $0 \leq d_{0,\mathrm{bs}}(\bar{S}/S)/l_{0,\mathrm{bs}}(\bar{S}/S) < 0.5$. In the right figure the BDT classifier value is furthermore requested to be $\geq$ -0.15.

Figure 10.25 and 10.26 and the figures in Appendix E and Appendix F show that the $\bar{S}$ which are reconstructed across events indeed show the same behaviour as the other four background samples. The discrepancy between the $\bar{S}$ which are reconstructed across events and the other background samples on the left figure in Figure 10.25 is due to the $S/\bar{S}$ candidates being reconstructed in the hard event. These $S/\bar{S}$ candidates have $\Delta\phi(K^0_S, \bar{\Lambda}^0/\Lambda^0) \approx 0$. When constructing $S/\bar{S}$ across events this type of S and $\bar{S}$ candidates disappear as hard events are not correlated from one event to the other. Indeed, after applying the $|\Delta\phi(K^0_S, \bar{\Lambda}^0/\Lambda^0)| > 0.5$ requirement, the cross-event background sample coincides with the other background samples as shown in Figure 10.29. The fact that this cross-event sample follows the other background samples shows that the background in this search is purely combinatorial and provides robustness to the background estimate.

Figure 10.30 shows the distribution of the BDT classifier applied to the S reconstructed in all datasets which are considered here. It shows that the background in the different datasets is comparable and that the background in the *SingleMuon_Run2016H*

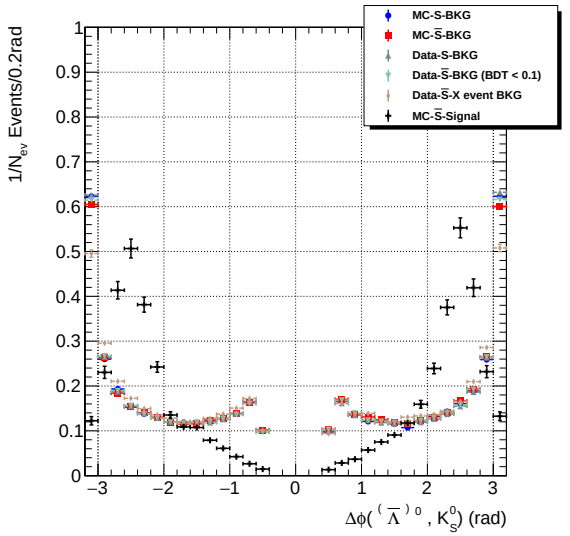


FIGURE 10.29: Angular difference in ϕ of the K^0_S and $\bar{\Lambda}^0$ from S or $\bar{S}$ candidates for MC $\bar{S}$ signal and five background samples for events which have all their final state particles in the fiducial region as defined in section 10.13 and with the $|\Delta\phi(K^0_S, \bar{\Lambda}^0/\Lambda^0)| > 0.5$ cut applied.

dataset, used for BDT training, is thus representative for the other runs. Again, also the S and $\bar{S}$ background samples from simulation are shown in the last figure in Figure 10.30, and demonstrate that the BDT classifier for these samples has the same behaviour as the background extracted from data. This similarity between background samples is shown also in Figure 10.31, where the distribution of the BDT variable for all five background samples is shown for events which have their final state particles within the fiducial region as defined in section 10.13 and which have $|\Delta\phi(K^0_S, \bar{\Lambda}^0/\Lambda^0)| > 0.5$, 2.02 cm $< l_{0,\mathrm{bpc}}(\mathrm{iv}) <$ 2.40 cm and $0 \leq d_{0,\mathrm{bs}}(\bar{S}/S)/l_{0,\mathrm{bs}}(\bar{S}/S) < 0.5$.

10.13 Systematic uncertainties

Several sources of systematic uncertainty are considered. The uncertainty on the signal reconstruction efficiency is dominated by uncertainties on single-track reconstruction efficiencies which feed into the overall $\bar{S}$ reconstruction efficiency, as discussed in section 10.13.1. Furthermore, the systematic uncertainty on the number of pp collisions is evaluated in section 10.13.2 and the systematic uncertainty which accounts for the fact that S are used as background reference for the $\bar{S}$ is evaluated in section 10.13.3.

10.13.1 Signal reconstruction efficiency

Systematic uncertainties need to be evaluated on the reconstructed objects used for the signal reconstruction in order to account for detector effects not well described in simulation, which could affect the value estimated for the $\bar{S}$ reconstruction efficiency. The only objects used as input to the signal reconstruction are tracker tracks.

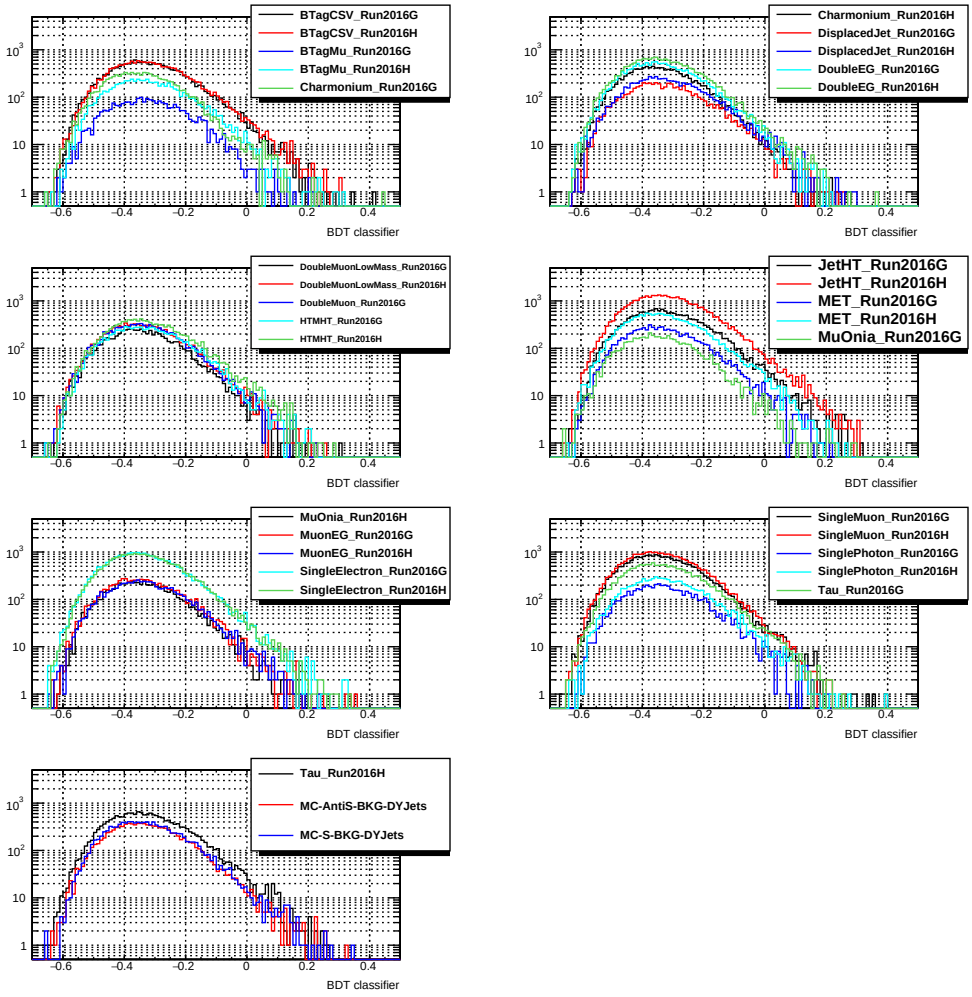


FIGURE 10.30: Distribution of the BDT parameter for the S background in all the data samples which are used and for the background as extracted from simulation.

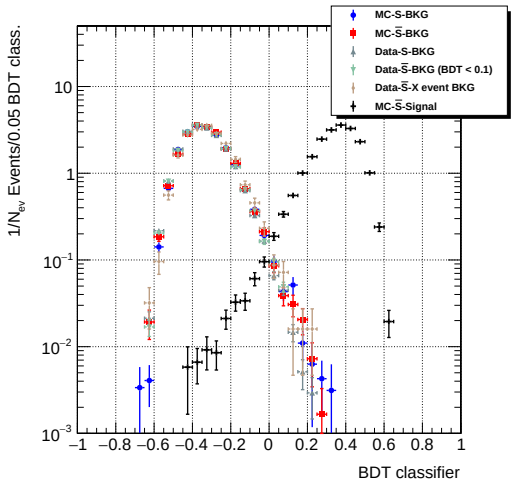


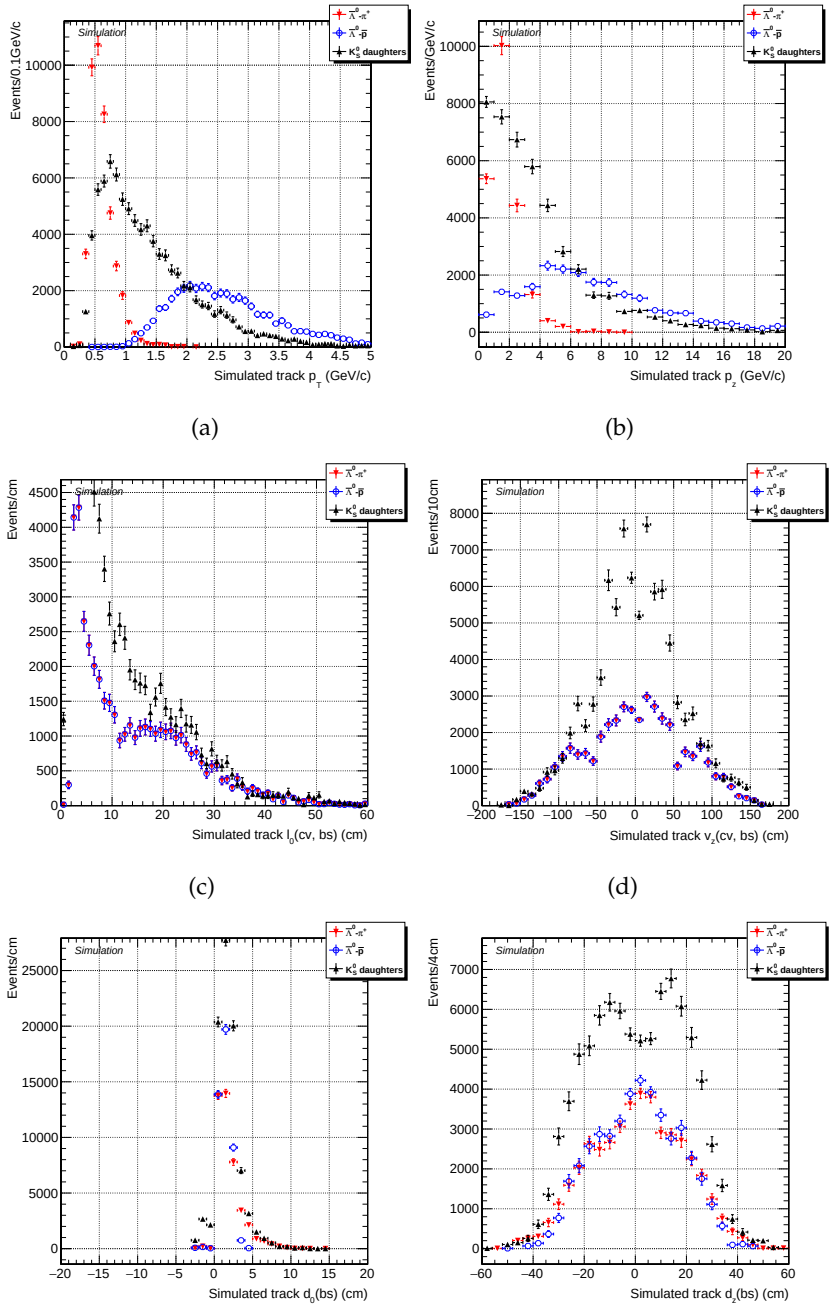
FIGURE 10.31: Distribution of the BDT classifier for MC $\bar{\mathrm{S}}$ signal and the five background samples for events which have their final state particles within the fiducial region as defined in section 10.13 and which have $|\Delta\phi(\mathrm{K}^0_\mathrm{S}, \bar{\Lambda}^0/\Lambda^0)| > 0.5$, 2.02 cm $< l_{0,\mathrm{bpc}}(\mathrm{iv}) <$ 2.40 cm and $0 \leq d_{0,\mathrm{bs}}(\bar{\mathrm{S}}/\mathrm{S})/l_{0,\mathrm{bs}}(\bar{\mathrm{S}}/\mathrm{S}) < 0.5$.

Figure 10.32 shows the distribution of six kinematic variables used here to parametrise the tracks associated to final state particles in reconstructed $\bar{\mathrm{S}}$ events:

- Transverse (p_T) and longitudinal (p_z) momentum.
- Transverse (l_0) and longitudinal (v_z) displacement of the track's creation vertex with respect to the beamspot and the absolute origin respectively.
- The impact parameters d_0 and d_z in the transverse and longitudinal plane respectively. The impact parameter is evaluated at the track's creation vertex and calculated relative to the beamspot for d_0 and *closest* valid PV for d_z, where the *closest* valid PV is the valid primary vertex amongst all valid primary vertices minimising the d_z.

These six parameters reflect what is specific for the signal tracks: low momenta, large displaced production and not necessarily pointing to the luminous region. These distributions are important as they show the parameter space for which systematic uncertainties on the track reconstruction efficiency has to be well controlled.

The strategy adopted here is to use a data control sample consisting of a track collection which has similar properties, and for which the production process is known to be well described at the generator level in simulation. In this way data-to-MC discrepancies are restricted to differences in tracking reconstruction efficiencies. Such a sample is provided by the tracks produced in the decay of K^0_S produced in the *underlying event* (UE): in events where the product of the hard collision is found along one axis, the underlying event can be studied in the two azimuthal *transversal* regions defined in Figure 10.33 where the $\Delta\phi$ angle is calculated relative to the direction of the leading object. K^0_S production in the UE is indeed known to be well described by simulation for certain Pythia tunes, as shown in Refs. [119] and [120].



(a) (b)

(c) (d)

(e) (f)

FIGURE 10.32: Distribution of kinematic variables, used to parametrise the tracks, for the final state particles in reconstructed $\bar{\Sigma}$ events.

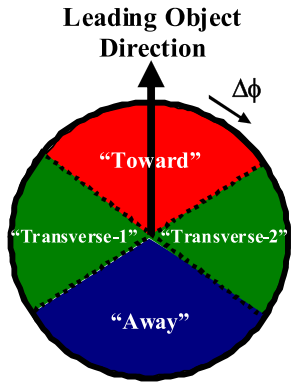


FIGURE 10.33: Definition of the azimuthal regions with respect to the leading object in the event [121]. The transversal regions are indicated in green.

The leading object in Figure 10.33 can be a Z boson. We use the 2016 RunG and RunH *DoubleMuon* dataset to reconstruct events with Drell-Yan (DY) Z production and will compare the tracks from K^0_S decay products in the UE to the same track collection obtained from *DYJets* MC simulation. The latter MC sample was produced with Pythia8, TuneCUETP8M1-amcatnloFXFX. This approach has the advantage that in DY Z production the leading object and thus the UE are straightforwardly defined. Considering K^0_S produced in the UE provides a sample of K^0_S, and thus decay products, which are kinematically similar for data and simulation.

For the Z reconstruction, only *TightMuons* are used with a $p_T > 20$ GeV/c, $|\eta| < 2.4$ and a longitudinal impact parameter within 0.01 cm from the z location of the first PV. Furthermore, the muons need to satisfy the isolation requirement as in Equation (10.13), where a cone of $\Delta R = 0.4$ is built around the muons to compute the flux of particle flow candidates. In Equation (10.13), I_{ch} is the sum of the p_T of the charged hadrons in the cone, I_{nh} the sum of the transverse energy (E_T) of the neutral hadrons in the cone, I_g the sum of the E_T of the photons in the cone and I_{chPU} the sum of the p_T of charged particles in the cone from particles not originating from the primary vertex.

$$I^{\mu}_{rel} = \frac{1}{p_T}(I_{ch} + \max(I_{nh} + I_g - 0.5 \cdot I_{chPU}, 0)) < 0.15. \tag{10.13}$$

Event selection is performed by requiring both data and MC simulated events to pass one of the below triggers:

- *HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ*
- *HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_DZ*.

Z candidates are then formed from oppositely charged muons of which the invariant mass has to lie within $\pm$ 7.5 GeV/c^2 from the known Z mass [10]. Events where more than 2 *LooseMuons* are found are vetoed. The underlying event is defined here by looking at activity in the *transversal* region with respect to the Z: the two regions which have an azimuthal angular difference larger than $\pi/3$ from both the Z and back-to-back direction of the Z. Only K^0_S in this transversal region are considered. Furthermore it is requested that no jet activity is seen in this transversal region by requiring no jets with energy larger

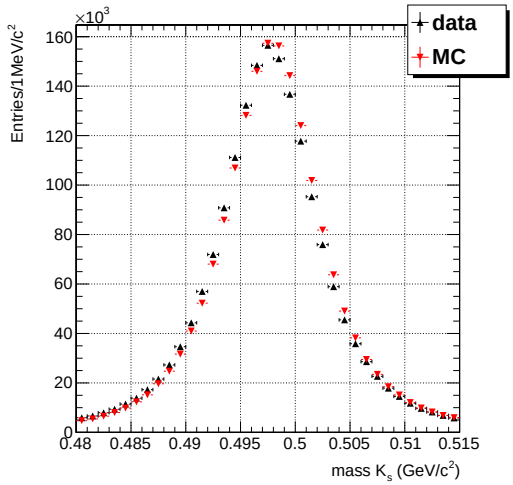


FIGURE 10.34: Mass of the K^0_S which were used in the systematic uncertainty study.

than 30 GeV to have more than a $\pi/4$ difference in ϕ from the Z or back-to-back of the Z. A tighter angular constraint is used for the latter compared to the definition of the transversal region, to take into account possible spill-over of the jet in the transversal region. K^0_S tagged as produced in the PU interactions are also considered, and are required to have a longitudinal impact parameter $|d_z|$ larger than 1 cm with respect to the primary vertex where the Z got reconstructed.

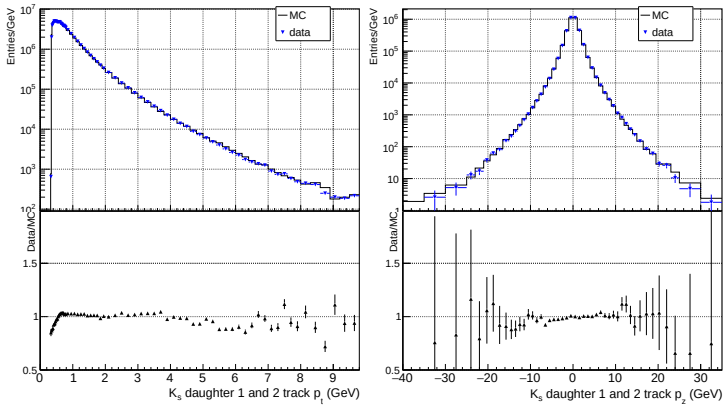
With the above selections, a collection of K^0_S, either produced in the underlying event or in the pileup interactions, is obtained. The K^0_S in these collections were reconstructed with the *adapted* V^0 reconstruction algorithm as discussed in section 10.9 and are required to have $|\eta| < 2$, $0 < d_{0,\text{bs}} < 0.1$ cm and an absolute longitudinal impact parameter with the closest (based on minimising the d_z) matching valid PV to be smaller than 0.2 cm. Furthermore the MC distributions are reweighed for the difference in z location of the closest PVs in data versus simulation.

The mass of K^0_S candidates reconstructed in data and simulation are shown in Figure 10.34. For these K^0_S, the six track parameters of the K^0_S daughters are studied. The results are shown in Figure 10.35 and overall a good agreement can be seen between data and simulation.

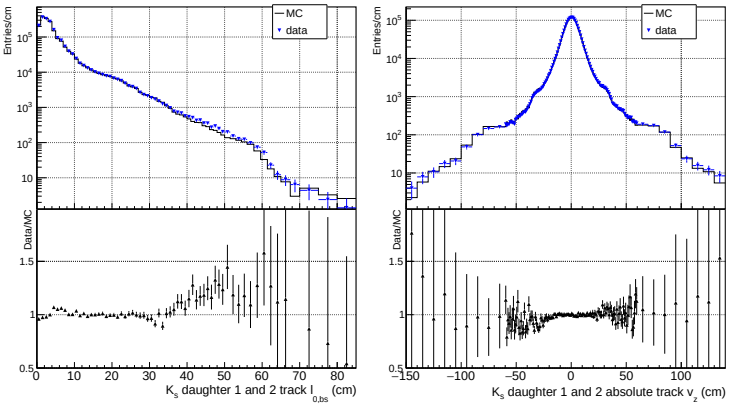
The data to MC ratios in Figure 10.35 are then used to estimate a systematic uncertainty on the reconstruction efficiency of the $\bar{S}$. This is performed by calculating a *correction factor* (C_k) for each reconstructed signal $\bar{S}$ in the simulation:

$$C_k = \prod_{j=0}^{3} \left(\prod_{i=0}^{5} \left(\frac{\text{Data}}{\text{MC}} \right)_i \right)_j . \tag{10.14}$$

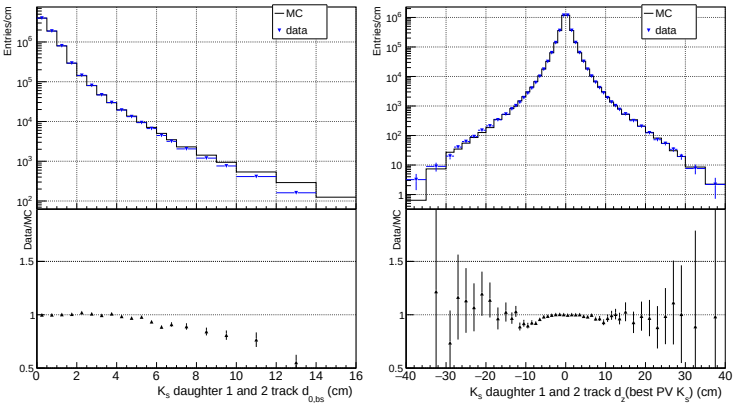
Here j represents the 4 final state particles and i the six kinematic parameters used to parametrise a track. Equation (10.14) translates the data-to-MC discrepancies of reconstruction efficiencies at the track level to data-to-MC discrepancies on the reconstruction efficiencies at the $\bar{S}$ level, but requires all individual Data/MC ratios to be uncorrelated.



(a) (b)



(c) (d)



(e) (f)

FIGURE 10.35: Data-MC ratios for several kinematic track parameters used in the systematic uncertainty study on tracking efficiencies.

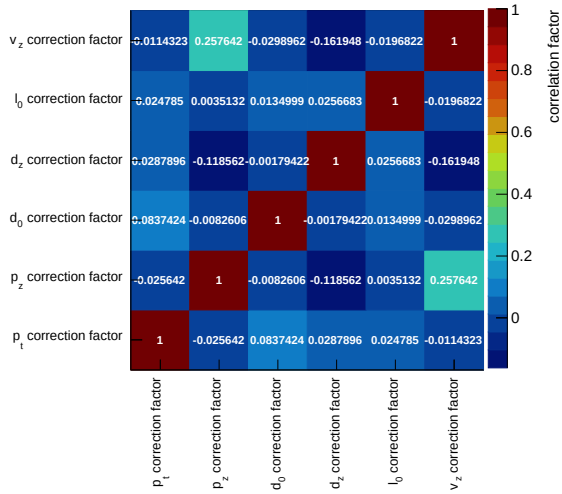


FIGURE 10.36: Correlation factors between the correction parameters of the kinematic variables of the tracks of reconstructed $\bar{S}$.

Residual correlations may lead to double counting and an overestimate of the systematic uncertainty. The correlation factors between the correction factors for all six kinematic variables are shown in Figure 10.36 and teach us that the correction factors can, to first order, indeed be treated uncorrelated.

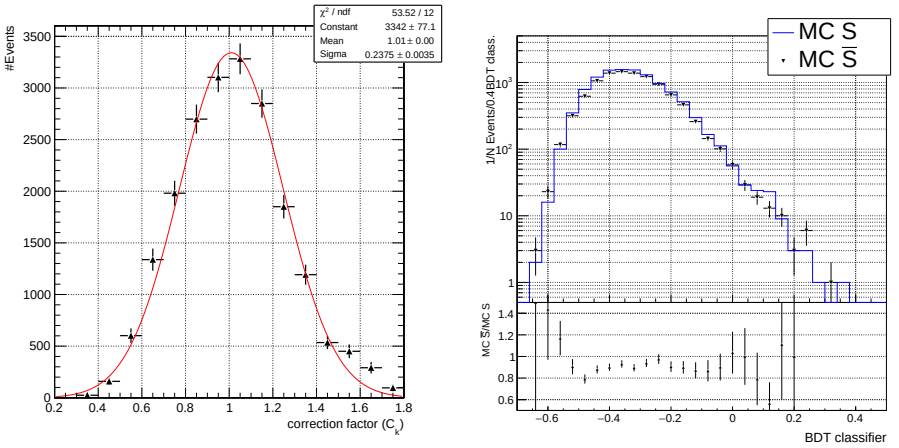
A fiducial region for which the Data/MC ratios shown in Figure 10.35 are within 0.8 to 1.2 is used to evaluate the correction factors. The fiducial region is defined as:

- $p_T \in$ [0.33 GeV/c, $+\infty$[
- $p_z \in$ [-22 GeV/c, 22 GeV/c]
- $l_0 \in$ [0 cm, 44.5 cm]
- $v_z \in$ [-125 cm, 125 cm]
- $d_0 \in$ [0 cm, 9.5 cm]
- $d_z \in$ [-27 cm, 27 cm]

Defining this fiducial region results in a loss in acceptance and thus also contributes to the overall signal reconstruction efficiency. The signal efficiency for this fiducial region is 58.3%.

The $\bar{S}$ overall correction factors as calculated from Equation (10.14) for reconstructed $\bar{S}$ with final state particles in the fiducial region are shown in Figure 10.37a. The width of the fitted distribution is taken here as the systematic uncertainty introduced due to tracking efficiency and is $\approx 24\%$.

Their is however one caveat to this approach: the fact that tracks from K^0_S are used implies that the K^0_S already needs to be reconstructed, which in turn implies that a second track needs to be reconstructed. This second track is correlated with the track under study and therefore the results in Figure 10.35 can only be interpreted as *single track ratios* if the



(a) (b)

FIGURE 10.37: LEFT: Distribution of the $\bar{S}$ correction factor as calculated from Equation (10.14). RIGHT: The ratio of $\bar{S}$ to S reconstructed in MC simulation is used to estimate a systematic uncertainty for the fact that S are used as background reference sample for the $\bar{S}$ background.

correlation in reconstruction efficiencies between the track under study and the second track are the same in data and simulation. The latter is not necessarily true and implies that there is some double counting of reconstruction efficiencies in the results shown in Figure 10.35 resulting in a too large, and thus conservative, estimate of the systematic uncertainty.

10.13.2 Number of pp collisions

A systematic uncertainty on the true number of pp-collisions can be extracted from the result shown in Figure 10.1. Here the size of the error bars is used as an estimate for the conversion from number of reconstructed primary vertices to true number of interactions which results in a systematic uncertainty of 4.8% on the average number of primary vertices per event in the full dataset considered here (Table 10.1). The uncertainty on the value for the event overlap, as discussed in section 10.8, is purely statistical and is negligible with respect to the 4.8% systematic uncertainty on the true number of pp-collisions.

10.13.3 S as background reference

A systematic uncertainty for the use of S as reference for the $\bar{S}$ background is extracted from simulation. Figure 10.37b shows the ratio of $\bar{S}$ to S reconstructed in the *DYJets* MC simulation sample. No obvious dependence on BDT classifier value is observed for this ratio. During limit setting (section 10.14), the statistical uncertainty on the ratio of the number of $\bar{S}$ to S events with a BDT classifier value larger than 0.2 will be used as systematic uncertainty for the use of S as reference for the $\bar{S}$ background. For this specific cut on the BDT classifier value this systematic uncertainty is 53%.

10.14 Limit setting and results

10.14.1 Limit setting introduction

If no significant $\bar{S}$ signal excess is observed, an upper limit can still be placed on $[\sigma(pp \rightarrow \bar{S}) \times \sigma(\bar{S} + n \rightarrow K_S^0 + \bar{\Lambda}^0)]$, referred to in the following loosely as the *signal cross section*. This limit setting is performed by comparing the background only hypothesis with the signal plus background hypothesis for a certain discriminating variable. A *test statistic* is built using this variable and the distribution of this test statistic is obtained by generating pseudo-experiments for the signal plus background and background only hypothesis, or, in case the number of background events is large, by using the asymptotic approximation [122].

In order to then obtain an upper limit, at a certain confidence level (CL), on the signal cross section, the CL_s method [123] is used. The LHC test statistic [124], which is used here, is defined as:

$$q_\mu = -2 \ln \frac{\mathcal{L}\left(\text{data}\,|\mu, \hat{\theta}_\mu\right)}{\mathcal{L}(\text{data}\,|\hat{\mu}, \hat{\theta})}, \text{with constraint} : 0 \leq \hat{\mu} \leq \mu, \tag{10.15}$$

with $\mathcal{L}$ the likelihood, μ the signal strength, $\hat{\theta}$ representing the full suite of nuisance parameters, $\hat{\theta}_\mu$ the maximum likelihood estimator of θ for a given μ and (pseudo) data, and the pair of parameter estimators $\hat{\mu}$ and $\hat{\theta}$ correspond to the global maximum of the likelihood. The expected number of events in a certain bin i with s_i signal events and b_i background events is given by: $\mu s_i + b_i$. With this definition, $\mu = 1$ corresponds to the signal plus background hypothesis with the nominal cross section and $\mu = 0$ to the background only hypothesis.

Two p-values p_{s+b} and p_b are now defined for an observation q_μ^{obs}:

$$p_{s+b} = \int_{q_\mu^{\text{obs}}}^{\infty} f\left(q_\mu|\mu, \hat{\theta}_\mu^{\text{obs}}\right) dq_\mu, \tag{10.16}$$

$$1 - p_b = \int_{q_{\mu=0}^{\text{obs}}}^{\infty} f\left(q_\mu|\mu = 0, \hat{\theta}_{\mu=0}^{\text{obs}}\right) dq_\mu, \tag{10.17}$$

with $f\left(q_\mu|\mu, \hat{\theta}_\mu^{\text{obs}}\right)$ and $f\left(q_\mu|\mu = 0, \hat{\theta}_{\mu=0}^{\text{obs}}\right)$ respectively the probability density function of the test statistic q_μ for the signal plus background hypothesis and background only hypothesis as constructed from the pseudo-experiments with the nuisance parameters, best describing the experimentally observed data, fixed to $\hat{\theta}_\mu^{\text{obs}}$ (signal+background) and $\hat{\theta}_{\mu=0}^{\text{obs}}$ (background-only). The signal is then excluded at 1 - α confidence level if

$$CL_s = \frac{p_{s+b}}{1 - p_b} \leq \alpha. \tag{10.18}$$

When extracting an upper limit at 95% CL the smallest signal strength is searched for which has $CL_s \leq 0.05$. The signal cross section corresponding to this signal strength is then quoted as the upper limit on the signal cross section.

To accomplish the limit evaluation the *Higgs Combine Tool* [125] is used which implements the above method, and is the common approach in CMS for limit evaluation.

10.14.2 Results

For this analysis the expected background rate is extracted from the reconstructed S in the data samples under consideration. However, due to the fact that we are dealing here with pp collisions a larger number of Λ^0s are expected compared to $\bar{\Lambda}^0$ s and as a result more S background is expected compared to $\bar{\text{S}}$ background. This is shown in Figure 10.38 where the distribution of the BDT classifier for three collections is shown:

- S background reconstructed in all datasets under investigation.
- $\bar{\text{S}}$ reconstructed in all datasets under investigation. The BDT classifier value is requested to be < 0.1 to avoid the signal region.
- The result from a 10% unblinding[10] of $\bar{\text{S}}$ signal. This distribution is shown scaled with a factor 10.

For all these collections, duplicate S and $\bar{\text{S}}$ candidates, due to overlaps in primary datasets, were explicitly removed and only events are used which have their final state particles in the fiducial region as defined in section 10.13 and which have $|\Delta\phi(\text{K}^0_\text{S}, \bar{\Lambda}^0/\Lambda^0)| > 0.5$, 2.02cm $< l_{0,\text{bpc}}(\text{iv}) <$ 2.40 cm and $0 \leq d_{0,\text{bs}}(\bar{\text{S}}/\text{S})/l_{0,\text{bs}}(\bar{\text{S}}/\text{S}) < 0.5$ (as introduced in section 10.12).

Comparing the first two collections indeed shows that less $\bar{\text{S}}$ background is observed compared to the S background, and the extrapolation factor from $\bar{\text{S}}$ to S background is obtained by evaluating the ratio of the $\bar{\text{S}}$ to S yield for BDT classifier values from -0.1 to 0.1. This range was chosen because it is towards the signal side, which is where this ratio needs to be evaluated, but is for BDT classifier values for which no signal is expected to show up yet. Figure 10.38 shows this ratio, and the fit[11] which is shown gives an extrapolation factor of 0.903 ± 0.011.

In order to evaluate the signal strength, three nuisance parameters on the signal and three nuisance parameters on the background are considered:

- Signal:
  - A 24% systematic uncertainty on the signal for the $\bar{\text{S}}$ reconstruction efficiency due to systematic uncertainties on the single track reconstruction efficiencies (section 10.13.1).
  - A 5% systematic uncertainty on the number of pp collisions in the total dataset (section 10.13.2).
  - A systematic uncertainty on the BDT signal efficiency ($\epsilon_{\text{Sig,BDT}}$), which is conservatively taken to be $\frac{1}{2} \times (1 - \epsilon_{\text{Sig,BDT}})$. This systematic uncertainty is mainly related to track and V^0 angular resolutions (the main inputs to the BDT) which are expected to be well controlled in simulation. The bottom left figure in Figure 10.30 and Figure 10.31 also show this similarity between data and MC.
- Background:
  - A 53% systematic uncertainty estimated from simulation to take into account that S are used as $\bar{\text{S}}$ background reference (section 10.13.3).

[10] A 10% sample from all the datasets under consideration.
[11] Polynomial of order 0.

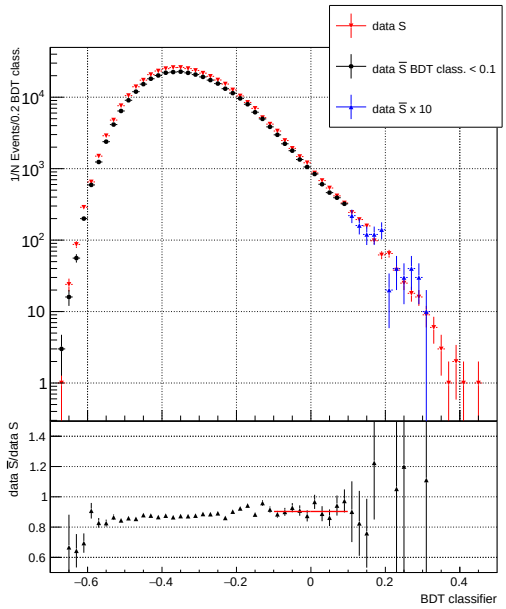


FIGURE 10.38: Distribution of the BDT classifier for: S reconstructed in all datasets under consideration, $\bar{S}$ with a BDT classifier value < 0.1 reconstructed in all datasets under consideration and $\bar{S}$ with a BDT classifier value > 0.1 for 10% of the datasets under consideration. The latter distribution was scaled with a factor 10. Duplicate S and $\bar{S}$, due to overlaps in datasets, were explicitly removed. Only events are shown which have their final state particles within the fiducial region as defined in section 10.13 and which have $|\Delta\phi(K^0_S, \bar{\Lambda}^0/\Lambda^0)| > 0.5$, 2.02 cm $< l_{0,\mathrm{bpc}}(\mathrm{iv}) <$ 2.40 cm and $0 \leq d_{0,\mathrm{bs}}(\bar{S}/S)/l_{0,\mathrm{bs}}(\bar{S}/S) < 0.5$.

- A systematic uncertainty on the extrapolation factor from S background to $\bar{S}$ background. This systematic uncertainty is evaluated by fitting the ratio in Figure 10.38 with a polynomial of order 1 for BDT classifier values from -0.1 to 0.1 and taking the relative difference between the polynomial of order 0 fit (shown in Figure 10.38) and polynomial of order 1 fit at a BDT classifier value of 0.35 as uncertainty on events in the eventually selected signal region (BDT > 0.35). As the background is steeply falling, most events are expected just above that threshold. This introduces an 8% systematic uncertainty.
- A nuisance parameter which takes into account the limited statistics in the S data control sample used to estimate the $\bar{S}$ background. This nuisance parameter is thus dependent on where the cut on the BDT classifier value is placed.

These nuisance parameters, as well as expected signal and background rates serve as input to the *combine* tool which provides an upper limit on the signal strength r:

$$r = \frac{\text{number of signal events}}{\text{number of expected signal events}}. \tag{10.19}$$

The signal strength is evaluated for a range of cuts on the BDT classifier value and the result is shown in Figure 10.39. For this specific result an arbitrary number of 100 events was used as the *signal rate*. This signal rate scales with the signal cross section.

A cut on the BDT classifier value of 0.35 is chosen in order to maximize the sensitivity of the search. At this BDT classifier value, the expected number of background events is 6.3 ± 2.5 (stat) $\pm$ 0.5 (syst), the BDT signal efficiency is 51% and the expected upper limit on the signal strength is 0.21. The expected upper limit on the number of signal events is thus $10.9^{+6.1}_{-3.3}$.

An expected upper limit on $[\sigma(\mathrm{pp} \to \bar{S}) \times \sigma(\bar{S} + \mathrm{n} \to K^0_S + \bar{\Lambda}^0)]$ can now be extracted, expressing the upper limit on the product of the $\bar{S}$ production cross section in LHC pp collisions and the cross section of $\bar{S}$ particles from such pp collisions producing a K^0_S-$\bar{\Lambda}^0$ pair in an interaction with a neutron in material at rest, for $\bar{S}$ with a mass of 1.8 GeV/c^2. The expected upper limit on this product of cross sections is calculated as introduced in section 10.7:

$$[\sigma(\mathrm{pp} \to \bar{S}) \times \sigma(\bar{S} + \mathrm{n} \to K^0_s + \bar{\Lambda}^0)]^{\mathrm{UL}}_{\mathrm{exp}} = \frac{N^{\mathrm{UL}}_{\mathrm{exp}}}{\frac{N_{\mathrm{ev}} \times <\mathrm{PU}> \times (1-C)}{\sigma_{\mathrm{pp}}} \times \epsilon_{\mathrm{reco}} \times \epsilon_{\mathrm{fid}} \times \epsilon_{\mathrm{BkgCuts}} \times N_{\mathrm{n}}}, \tag{10.20}$$

with $N^{\mathrm{UL}}_{\mathrm{exp}}$ the expected upper limit on the number of signal events after all selection cuts, N_{ev} the total number of pp-interactions in the considered dataset, <PU> the average pileup in this dataset, C the overlap rate between datasets, σ_{pp} the proton-proton inelastic cross section, ϵ_{reco} the $\bar{S}$ reconstruction efficiency, ϵ_{fid} the signal efficiency for the fiducial region defined in section 10.13, $\epsilon_{\mathrm{BkgCuts}}$ the efficiency with which the $\bar{S}$ candidates survive the background cuts and N_{n} the neutron column density of the beampipe for $\eta = 0$ [12].

[12] The value at $\eta = 0$ must be used here, as reweighing for the $\bar{S}$ path length through the beampipe was already performed in the reconstruction efficiency calculation.

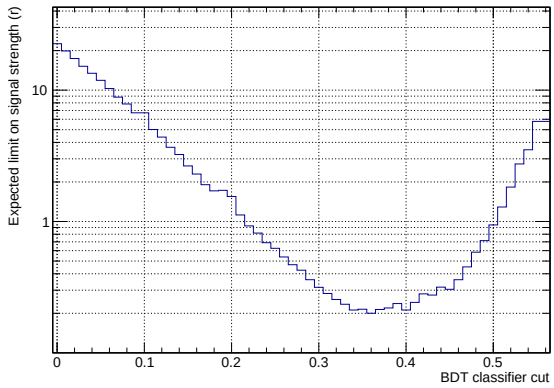


FIGURE 10.39: Expected limits on the signal strength for several cut values on the BDT classifier value.

N_{exp}^{UL}	10.9
Number of pp-collisions (N_{ev})	2.007×10^9
Average pileup (<PU>)	21.4
Dataset overlap rate (C)	12.84%
pp inelastic cross section (σ_{pp})	70 mb
Signal reconstruction efficiency (ϵ_{reco})	0.0014%
Signal efficiency for fiducial region systematic uncertainty (ϵ_{fid})	58.3%
Efficiency of signal for background suppression ($\epsilon_{BkgCuts}$)	48.2%
Neutron column density of CMS beampipe (N_n)	4.97×10^{-5}mb^{-1}

TABLE 10.5: Summary of values used in Equation (10.20) to calculate the observed limit.

This neutron column density is given by:

$$N_n = \frac{N_{neutron,Be} \times \rho_{Be} \times d}{m_{Be}} = \frac{5 \times 1.85 \times 10^{-3} \frac{kg}{cm^3} \times 0.08 \text{ cm}}{14.9 \times 10^{-27} \text{ kg}} = 4.97 \times 10^{-5} \text{ mb}^{-1}, \quad (10.21)$$

with $N_{neutron,Be}$ the number of neutrons in a beryllium nucleus, ρ_{Be} the mass density of beryllium, d the thickness of the beampipe and m_{Be} the mass of a beryllium nucleus.

The values listed in Table 10.5 are used to calculate the expected upper limit. The value for $\epsilon_{BkgCuts}$ shown in this table also takes into account the 94.4% combined signal efficiency for the three background cuts applied before the BDT. This results in an expected upper limit on $[\sigma(pp \to \bar{S}) \times \sigma(\bar{S} + n \to K_S^0 + \bar{\Lambda}^0)]$ of $105^{+57.8}_{-32.4}$ mb^2.

For a partial unblinding, 10% of the studied dataset was used to extract an observed upper limit on $[\sigma(pp \to \bar{S}) \times \sigma(\bar{S} + n \to K_S^0 + \bar{\Lambda}^0)]$. At a BDT classifier value cut of 0.35, with an expected background rate of 0.632 ± 0.795 (stat) $\pm$ 0.051 (syst), a similar calculation as performed for the expected limit with the full dataset results in an expected limit on $[\sigma(pp \to \bar{S}) \times \sigma(\bar{S} + n \to K_S^0 + \bar{\Lambda}^0)]$ of 309^{+273}_{-136} mb^2. The observed background

for this requirement on the BDT classifier value is 0 and the observed limit on the product of cross sections is calculated to be 237 mb[2].

10.15 Summary and outlook

Sexaquarks were proposed in Ref. [88] as a strongly bound state of six quarks. If its mass is in the correct range, the S is a dark matter candidate. CMS, with its excellent tracking detector, and the high luminosity provided by the LHC present an attractive possibility to search for this elusive particle.

The strategy adopted in this first-ever Sexaquark search at the LHC, was to look for $\bar{S}$ produced in pp collisions and annihilating on neutrons in the CMS beampipe, resulting in the production of $\bar{\Lambda}^0$ and K^0_S:

$$\bar{S} + n \rightarrow \bar{\Lambda}^0 + K^0_S \rightarrow \bar{p} + \pi^+ + \pi^- + \pi^+. \tag{10.22}$$

The final state particles in this channel can be reconstructed with the CMS tracker. The vertex of the V^0s gives the annihilation vertex and the invariant mass of the $\bar{S}$ can be approximately reconstructed.

The search strategy which was adopted for this signal is non-standard compared to typical CMS analyses, for the following reasons:

- The $\bar{S}$ needs to be produced and subsequently needs to annihilate on material. As a result, the combination of the production cross section and the annihilation cross section are probed.
- The fact that the $\bar{S}$ first needs to travel from its production vertex before it can interact with material and the fact that the K^0_S and $\bar{\Lambda}^0$ have decay lengths of a few centimetres implies that the final state particles are displaced. This is not beneficial for the track reconstruction, which in CMS is optimised for particles produced in the luminous region. On top of that, the majority of the $\bar{S}$ are produced forward and the signal is furthermore enhanced in the forward region due to the increase of the $\bar{S}$ path length through the beampipe.
- Due to the low momenta of the final state particles, no dedicated trigger is available for this signal. The $\bar{S}$ was therefore searched for in all existing datasets, where in particular the production of $\bar{S}$ in pileup interactions was considered.

The special signal topology required a dedicated signal simulation to develop the search strategy and to estimate the signal reconstruction efficiency. To do so, the $\bar{S}$-neutron interaction was added to the CMS simulation framework. For this simulation and the rest of the work an $\bar{S}$ with a mass of 1.8 GeV/c^2 was investigated. An important, non intuitive, result from this simulation is that the Fermi momentum of the target neutron has a significant effect on the reconstructed invariant mass of the $\bar{S}$. This essentially means that, despite accurate reconstruction of the final state products, no sharp $\bar{S}$ invariant mass peak can be extracted for this channel. On the plus side, it was shown that the location of the neutron-$\bar{S}$ annihilation vertex can be reconstructed with $\mathcal{O}$(mm) accuracy, which is not straightforward as it requires the vertexing of two neutral particles.

Using an adequate signal simulation, it was shown that the reconstruction efficiency for this signal is very low, and estimated to be 0.0014%. Both acceptance and reconstruction effects contribute to this inefficiency, where for the latter the soft pion in the $\bar{\Lambda}^0$ decay proved to be a bottleneck for the reconstruction.

The background study for this signal showed that S candidates can be used as reference sample for the $\bar{S}$ background. Furthermore, the origin of the background was found to be combinatorial. This was proven by reconstructing $\bar{S}$ candidates combining K_S^0 and $\bar{\Lambda}^0$ from different events. Thanks to the special topology of this signal, the background could very efficiently be eliminated using three background rejection cuts and a multivariate classifier.

An expected upper limit (at 95% CL) on $[\sigma(pp \rightarrow \bar{S}) \times \sigma(\bar{S} + n \rightarrow K_S^0 + \bar{\Lambda}^0)]$ was extracted and found to be $105^{+57.8}_{-32.4}$ mb^2. This expresses the upper limit on the product of the $\bar{S}$ production cross section in LHC pp collisions and the cross section of $\bar{S}$ particles from such pp collisions producing a K_S^0-$\bar{\Lambda}^0$ pair in an interaction with a neutron in material at rest, for $\bar{S}$ with a mass of 1.8 GeV/c^2. A 10% unblinding of the dataset under consideration did not lead to the observation of any signal and an observed upper limit (at 95% CL) on the product of cross sections was set at 237 mb^2.

Mainly due to the inefficient signal reconstruction, this analysis does not provide the needed sensitivity to exclude or confirm the existence of the $\bar{S}$. Indeed, in Ref. [88], Farrar estimates the $\bar{S}$ production cross section to be at most 10^{-4} of the pion production cross section. From EPOS-LHC simulations, normalised to the charged hadron yield presented in Ref. [126], the expected average charged pion yield in one pp collision is ≈ 80 over the full η range. Furthermore, Farrar projects the $\bar{S}$-neutron interaction cross section to be at most $10^{-6}\sigma_{pp}$ for proton lab momenta in the 10 GeV range. In this momentum range σ_{pp} is around 40 mb. This leads to a product of production and interaction cross sections which is at most: 80×70 mb $\times 10^{-4} \times 40$ mb $\times 10^{-6} = 2.24 \times 10^{-5}$ mb^2. This shows that the observed limit, extracted from this search, is not sensitive for the considered $\bar{S}$ model.

By considering Equation (10.20), several optimisations can be envisaged to make this search more sensitive: more data can be used, the reconstruction efficiency can be improved, and the search region could be extended from the beampipe volume to include $\bar{S}$ production on detector material as well.

The use of the full 2015, 2016, 2017, 2018 and B-parking [127] dataset would result in an estimated $\mathcal{O}(200)$ increase[13] in the number of pp interactions. It is clear that simply using more data will not bring the large improvement which is required to make this analysis competitive. The effect of the upgraded pixel detector on acceptance and reconstruction efficiency in 2017 and 2018 data should however be evaluated. Track seeding being possible up to larger radii and longitudinal displacement due to the extra detector layers could have a significant effect.

The largest sensitivity gain could be achieved by improving the signal reconstruction efficiency. Whilst a significant inefficiency due to acceptance in the forward region seems inevitable, dedicated tracking algorithms for highly displaced low momentum tracks or V^0s could be developed to augment the reconstruction efficiency for events within acceptance. A complication for these algorithms can be foreseen in performant seeding at high displacement, which relies on stereo-modules in the standard CMS track reconstruction. On the plus side, the high displacement for this signal implies that the track reconstruction for the final state particles in this search is performed in a region of the tracker where the track occupancy is modest compared to the region closer to the luminous region. The implications of modified track reconstruction on track fake rates should of course be evaluated. Efficiently discriminating the background from the signal, as was possible in this search, might become challenging.

[13] Taking into account that only 10% of 2016 Run G and RunH data was used for the partial unblinding performed here.

A third approach to boost the sensitivity would be to include more material to search for $\bar{S}$ annihilation on neutrons. Considering also $\bar{S}$ annihilation on for example the pixel detector and its services would have a significant impact on the N_n factor in Equation (10.20). Furthermore, for this type of more central events the inefficiency due to detector acceptance will implicitly be reduced. Including the pixel detector layers could increase the material budget, in terms of nuclear interaction lengths, with a factor $\mathcal{O}(10)$ [128]. Lower reconstruction efficiency due to the potential loss of these detector layers for seeding and tracking of course needs to be evaluated, but even for the standard CMS track reconstruction, this loss seems to be manageable as illustrated by the top left figures in Figure D.2-D.4.

It is hard to predict if the improvements described above could push the sensitivity into the required range. From the performed analysis it is clear that a dedicated signal simulation is required to evaluate the reconstruction efficiency for a convoluted signal as the one investigated here.

For LHC Run 3, it could be considered to design a dedicated HLT trigger for this signal which could exploit the displacement of this signal and/or use the HLT pixel tracking which will come online for Run 3. If the latter is feasible, an increase of a factor 100 in the data sample could be envisaged if the HLT trigger could be run for each L1 event.

Looking even more ahead, the prospect of the Phase-2 upgrade of the CMS tracking system could add a significant improvement on both the signal acceptance and the tracking efficiency. Indeed, as shown in Figure 4.10, the η range of the current pixel detector and its z coverage will be significantly enlarged. This, together with the vast volume equipped with PS modules (see Figure 5.6), will essentially allow more continuous track seeding up to much larger radii and z. Furthermore, the higher luminosity will provide a larger dataset to look for this signal, but without a dedicated trigger the increase in luminosity will not provide the needed gain.

Other completely different search strategies might also be considered for future Sexaquark searches. Strategies where the production of S or $\bar{S}$ can be probed in a direct way rather than through the succession of production and annihilation are favoured. The latter after all adds an estimated factor 10^{-6} to the product of production and annihilation cross section. A discovery channel which does not have this limitation is the one described in section 10.6.1, where the S and $\bar{S}$ could be searched for in the decay of the Υ. This search furthermore has the advantage that the invariant missing-mass peak is not polluted by the target neutron's Fermi motion as is the case for the channel investigated here. The search performed by BABAR, described in Ref. [108], could be extended to more inclusive channels[14]. The latter is indeed a necessity if a narrow missing-mass peak needs to be extracted. A search at Belle II, taking advantage of the higher luminosity compared to BABAR and Belle, might also be warranted. Of course, the projected factor 50 in integrated luminosity for Belle II compared to Belle will only be attained by 2025.

10.16 Author's contribution

The author was the main developer for this analysis. The author evaluated the reconstruction algorithm for V^0s and developed the reconstruction of S and $\bar{S}$ objects. Starting from the EPOS-LHC package, a signal simulation was set up by the author which included the addition of the neutron-$\bar{S}$ interaction in GEANT and a method to efficiently

[14] $\Upsilon \rightarrow S\, \bar{\Lambda}^0\, \bar{\Lambda}^0\, X$ instead of $\Upsilon \rightarrow S\, \bar{\Lambda}^0\, \bar{\Lambda}^0$.

simulate the $\bar{\Sigma}$ low interaction cross section. The author then used this simulation to estimate the signal reconstruction efficiency. Subsequently the author performed a study of the background to be able to suppress it in the most efficient way and to set a limit on $[\sigma(\mathrm{pp} \to \bar{\Sigma}) \times \sigma(\bar{\Sigma} + \mathrm{n} \to \mathrm{K}^0_S + \bar{\Lambda}^0)]$ with an evaluation of the systematic uncertainties.

Appendix A

MPA and SSA SEU cross sections

This appendix provides the raw data from the MPA and SSA SEU test described in section 8.4.

A.1 SSA high latency

Ion	Angle	Effective LET	Fluence	Duration SEU counter	Duration Full-event	Duration stub	Latency	Full-event Error Upset	Stub Error	Full-event SEU cross section	Error full-event SEU cross section	Stub SEU cross section	Error Stub SEU cross section
	[°]	[MeV/(mg/cm^2)]	[part/s × cm^2]	[s]	[s]	[s]	[# of cycles]	[#]	[#]	[cm^2]	[cm^2]	[cm^2]	[cm^2]
13C4+	0	1.3	14220	30	240	240	500	1	3	2.93015E-07	2.93024E-07	8.79044E-07	5.07567E-07
22Ne7+	0	3.3	14986	30	240	240	500	17	13	4.72663E-06	1.14698E-06	3.61448E-06	1.00288E-06
27Al8+	0	5.7	12000	30	600	600	500	20	12	2.77778E-06	6.21578E-07	1.66667E-06	4.81334E-07
40Ar10+	0	10	15258	30	240	240	500	38	29	1.03771E-05	1.68531E-06	7.91934E-06	1.47187E-06
40Ar10+	30	11.5	13339	30	210	210	500	21	32	7.49681E-06	1.63729E-06	1.14237E-05	2.02199E-06
58Ni18+	0	20.4	14716	30	210	210	500	50	40	1.61793E-05	2.29204E-06	1.29435E-05	2.04936E-06
84Kr25+	0	32.4	15870	30	240	240	500	82	81	2.15291E-05	2.38305E-06	2.12665E-05	2.3684E-06
84Kr25+	30	37.4	13613	30	240	240	500	130	67	3.97904E-05	3.50599E-06	2.05074E-05	2.51135E-06
103Rh31+	0	45.8	15553	30	180	180	500	85	74	3.03621E-05	3.30333E-06	2.64329E-05	3.08096E-06
103Rh31+	30	52.9	14226	30	240	210	500	166	104	4.86199E-05	3.79449E-06	3.48122E-05	3.42639E-06
124Xe35+	0	62.5	15080	30	240	210	500	156	89	4.31034E-05	3.46751E-06	2.81041E-05	2.98785E-06
124Xe35+	30	71.17	13925	30	150	90	500	98	30	4.6918E-05	4.76168E-06	2.39378E-05	4.38188E-06

A.2 SSA low latency

Ion	Angle	Effective LET	Fluence	Duration SEU counter	Duration full-event	Duration stub	Latency	Full-event Error Upset	Stub Error	Full-event SEU cross section	Error full-event SEU cross section	Stub SEU cross section	Error Stub SEU cross section
	[°]	[MeV/(mg/cm^2)]	[part/s × cm^2]	[s]	[s]	[s]	[# of cycles]	[#]	[#]	[cm^2]	[cm^2]	[cm^2]	[cm^2]
13C4+	0	1.3	14220	30	210	210	100	1	6	3.34874E-07	3.34886E-07	2.00924E-06	8.20447E-07
22Ne7+	0	3.3	14986	30	210	210	100	3	13	9.53271E-07	5.50427E-07	4.13084E-06	1.14619E-06
27Al8+	0	5.7	12000	30	570	570	100	6	16	8.77193E-07	3.5819E-07	2.33918E-06	5.85135E-07
40Ar10+	0	10	15258	30	210	210	100	8	25	2.49674E-06	8.82962E-07	7.80231E-06	1.56174E-06
40Ar10+	30	11.5	13339	30	210	210	100	11	24	3.9269E-06	1.18452E-06	8.56779E-06	1.75055E-06
58Ni18+	0	20.4	14716	30	180	210	100	5	54	1.88759E-06	8.44319E-07	1.74737E-05	2.38228E-06
84Kr25+	0	32.4	15870	30	210	210	100	26	93	7.80148E-06	1.53124E-06	2.79053E-05	2.90203E-06
84Kr25+	30	37.4	13613	30	210	210	100	27	47	9.44475E-06	1.81952E-06	1.64409E-05	2.40246E-06
103Rh31+	0	45.8	15553	30	210	210	100	44	85	1.34716E-05	2.03378E-06	2.60247E-05	2.83044E-06
103Rh31+	30	52.9	14226	30	150	90	100	34	38	1.59333E-05	2.73689E-06	2.96796E-05	4.83046E-06
124Xe35+	0	62.5	15080	30	210	210	100	41	68	1.29468E-05	2.02471E-06	2.14728E-05	2.60985E-06
124Xe35+	30	71.17	13925	30	210	210	100	42	88	1.43627E-05	2.21966E-06	3.00932E-05	3.21841E-06

A.3 MPA high latency

Ion	Angle	Effective LET	Fluence	Duration	Latency	SEU counter upset	Full-event Error upset	Stub Error upset	SEU counter SEU cross section	Error SEU counter SEU cross section	Full-event SEU cross section	Error full-event SEU cross section	Stub SEU cross section	Error Stub SEU cross section
	[°]	[MeV/(mg/cm^2)]	[part/(s × cm^2)]	[s]	[# of cycles]	[#]	[#]	[#]	[cm^2]	[cm^2]	[cm^2]	[cm^2]	[cm^2]	[cm^2]
13C4+	0	1.3	15000	282	500	8	3	13	1.89125E-06	1.42811E-08	7.0922E-07	4.09503E-07	3.07329E-06	2.32067E-08
13C4+	30	1.5	13200	282	500	5	4	10	1.34322E-06	1.12355E-08	1.07458E-06	5.37363E-07	2.68644E-06	2.2471E-08
27Al8+	0	5.7	14500	282	500	99	155	93	2.42113E-05	1.87754E-07	3.79066E-05	3.05889E-06	2.27439E-05	1.76375E-07
27Al8+	30	6.6	13000	282	500	85	182	111	2.3186E-05	1.96393E-07	4.96454E-05	3.70391E-06	3.02782E-05	2.56466E-07
40Ar10+	0	10	15100	282	500	187	251	142	4.39153E-05	3.29899E-07	5.89451E-05	3.74684E-06	3.33474E-05	2.50511E-07
40Ar10+	30	11.5	12500	282	500	167	224	136	4.73759E-05	4.14572E-07	6.35461E-05	4.28211E-06	3.85816E-05	3.37616E-07
58Ni18+	0	20.4	15000	282	500	261	415	235	6.17021E-05	4.6592E-07	9.81087E-05	4.87262E-06	5.55556E-05	4.19506E-07
58Ni18+	30	23.6	13200	282	500	297	442	293	7.97872E-05	6.6739E-07	0.000118741	5.73458E-06	7.87127E-05	6.58402E-07
84Kr25+	0	32.4	15000	282	500	414	543	336	9.78723E-05	7.39045E-07	0.000128369	5.59346E-06	7.94326E-05	5.99804E-07
84Kr25+	30	37.4	13100	282	500	431	661	361	0.000116669	9.8201E-07	0.000178929	7.12063E-06	9.77208E-05	8.22519E-07
103Rh31+	0	45.8	15400	282	500	534	850	434	0.000122962	9.09757E-07	0.000195726	6.86776E-06	9.99355E-05	7.39391E-07
103Rh31+	30	52.9	13800	282	500	586	882	487	0.000150581	1.21481E-06	0.000226642	7.84742E-06	0.000125141	1.00958E-06
124Xe35+	0	62.5	14000	282	500	686	975	558	0.000173759	1.38567E-06	0.00024696	8.15058E-06	0.000141337	1.12712E-06
124Xe35+	30	71.17	13300	282	500	575	998	472	0.000153309	1.27447E-06	0.000266091	8.70857E-06	0.000125847	1.04617E-06

A.4 MPA low latency

Ion	Angle	Effective LET	Fluence	Duration	Latency	SEU counter upset	Full-event Error upset	Stub Error upset	SEU counter SEU cross section	Error SEU counter SEU cross section	Full-event SEU cross section	Error full-event SEU cross section	Stub SEU cross section	Error Stub SEU cross section
	[°]	[MeV/(mg/cm^2)]	[part/(s × cm^2)]	[s]	[# of cycles]	[#]	[#]	[#]	[cm^2]	[cm^2]	[cm^2]	[cm^2]	[cm^2]	[cm^2]
13C4+	0	1.3	15000	282	40	5	2	24	1.18203E-06	8.92566E-09	4.72813E-07	3.34349E-07	5.67376E-06	4.28432E-08
13C4+	30	1.5	13700	282	40	6	1	18	1.55304E-06	1.2603E-08	2.58839E-07	2.58848E-07	4.65911E-06	3.78089E-08
27Al8+	0	5.7	14500	282	40	99	110	88	2.42113E-05	1.87754E-07	2.69014E-05	2.57342E-06	2.15212E-05	1.66893E-07
27Al8+	30	6.6	13000	282	40	103	116	100	2.8096E-05	2.37982E-07	3.16421E-05	2.9501E-06	2.72777E-05	2.31051E-07
40Ar10+	0	10	15100	282	40	141	144	129	3.31126E-05	2.48747E-07	3.38171E-05	2.82952E-06	3.02945E-05	2.27577E-07
40Ar10+	30	11.5	12500	282	40	129	180	164	3.65957E-05	3.20239E-07	5.10638E-05	3.83221E-06	4.65248E-05	4.07125E-07
58Ni18+	0	20.4	15000	282	40	263	266	230	6.21749E-05	4.6949E-07	6.28842E-05	3.8848E-06	5.43735E-05	4.1058E-07
58Ni18+	30	23.6	13200	282	40	313	287	256	8.40855E-05	7.03344E-07	7.71008E-05	4.59658E-06	6.87728E-05	5.75259E-07
84Kr25+	0	32.4	15000	282	40	399	342	338	9.43262E-05	7.12268E-07	8.08511E-05	4.41435E-06	7.99054E-05	6.03375E-07
84Kr25+	30	37.4	13100	282	40	414	395	364	0.000112068	9.43277E-07	0.000106924	5.45471E-06	9.85328E-05	8.29354E-07
103Rh31+	0	45.8	14400	282	40	552	526	458	0.000135934	1.05994E-06	0.000129531	5.73743E-06	0.000112786	8.7944E-07
103Rh31+	30	52.9	12800	282	40	457	464	355	0.000126607	1.08624E-06	0.000128546	6.06866E-06	9.83488E-05	8.43796E-07
124Xe35+	0	62.5	14200	282	40	591	608	485	0.000147588	1.16368E-06	0.000151833	6.27294E-06	0.000121117	9.54967E-07
124Xe35+	30	71.17	13300	282	40	658	545	502	0.000175439	1.45843E-06	0.00014531	6.34053E-06	0.000133845	1.11266E-06

Appendix B

Full CIC start-up sequence

Detailed CIC start-up sequence as used in section 9.3:

Start-up step	Action	Note
Preparation	Power FC7 and load FW.	The default state of the firmware should send nothing on the ouput lines, except for the hard reset which should be enabled.
	Power interface card.	
	Enable CIC.	
	Launch 320MHz clock generation for the CIC.	
	Disable the hard reset.	
	Set the scDllConfirmCountSelect to the maximum.	This register sets the time limit the DLLs are allowed to lock. This needs to be set for all PhyPorts. The default value is to give the CIC 4 clock cycles to lock, by setting the maximum this is increased to 31 clock cycles.
Fast command locking	Start the sending of the fast command frame to the CIC.	Sending only the header and trailer of the fast command frame, the CIC uses this header to generate an internal 40MHz clock.
	Check the fast command locking state.	If not locked it means the CIC did not find the fast command header. As there is no phase alignment feature on the fast command line it is possible that this happens. Therefore it is possible to select the 320MHz clock edge on which the CIC samples manually.
	Send a ReSync fast command	
	Check reception of ReSync fast command	To check if the CIC properly interprets fast command payload. If not correct the other 320MHz clock edge can be tried.
Phase alignment	Configure the PhyPorts in 'auto startup' mode.	The other option would be the use of the 'fixed' mode, which can be used ones the phase and word alignment patterns are known from a previous alignment. This would for example be the case when the CIC sits on a hybrid and is connected to the front-end chips directly, the delay between the lines would be constant in that case.
	Send a soft reset to the CIC and release it.	This resets the CIC chip except for the slow control block and registers.
	Check locking of DLLs.	There is 1 status bit for each PhyPort.
	Send phase tuning patterns from the FC7 to the CIC.	The FC7 sends on all the data lines the 1010 1010 repetitive patterns. This allows the CIC to find the correct phase for sampling.
	Check for phase tuning lock on all the input channels.	There is a status bit for each of the 48 data input lines for the CIC.
Word alignment	Configure word alignment patterns in CIC	Programme the word tuning patterns which the CIC will be expecting from the front-end chips into the CIC. One byte can be specified per stub line, therefore there are 40bits in total to configure.
	Disable usage of the external alignment value	By default the CIC is configured to use the alignment values which are stored internally on the chip and which are specific for a certain configuration (e.g. on a hybrid) this will be the default configuration in the final application where these values will be stored in a database. Now we want to run the alignment procedure, so do not use the internal ones. This philosophy is similar to the first step of the phase tuning.
	Enable the automatic word alignment procedure	This is an I^2C command to the chip.
	Send a ReSync fast command	
	Check for word tuning lock on all the stub data lines	There is 1 bit for the status of the word alignment.
	Disable the automatic word alignment procedure	
BX0 alignment	Select the stub line to perform the BX0 alignment	This is a certain line on a certain front-end chip.
	Disable the usage of the external alignment value	The reason for this is similar to the disabling of the usage of the external alignment values of the phase and word alignment.
	Enable the auto BX0 alignment procedure	
	Send a ReSync fast command	The front-end chip should send a recognisable pattern a constant amount of clock cycles after the ReSync.
	Check if the recognisable pattern was seen by the CIC	There is an I^2C status bit to check this. If it is false it means that the pattern was not seen on the line after a given time interval.
	Disable the auto BX0 alignment procedure	
Back-end alignment	Enable the CIC output transmitters	Enable the SLVS pads for data to be sent of the CIC chip.
	Enable the output of the data patterns out of the CIC	These data patterns are used for the phase, word and package alignment. The pattern which is sent is 1110 1010 every 8th BX followed by 7 BXs of 1010 1010 patterns for the stub lines and ...1010... for the full-event line.
	Check if phase and word tuning succeeded on the DAQ	There are status register in the μDTC which specify if the phase and word tuning succeeded. The phase tuning is done on the ...0101010... whereas the word tuning is done on the consecutive three 1's which are sent.
	Send to μDTC stub decoder that the phase and word tuning succeeded.	This tells the μDTC stub decoder that the bytes which arrive from the ISERDESES are correctly phase and word aligned. The μDTC can proceed now to look for the consecutive three 1's to find the start of a stub package, this is the last part of the synchronisation and is referred to as the package alignment.
	Disable the output of the CIC output patterns	

Appendix C

CBC configurations for CBC-CIC start-up

This appendix provides the CBC3 configurations which should be applied during the start-up sequence for the CBC+CIC as explained in section 9.3.

Register name	Page	Address	Phase alignment 1	Phase alignment 2	Phase alignment 3	Word alignment	BX0 alignment
TestPulsePotNodeSel	0x00	0x000D	0xEA	0xEA	0xEA	0xEA	0xC0
MiscTestPulseCtrl&AnalogMux	0x00	0x000F	0x20	0x20	0x20	0x20	0x60
CoincWind&Offset34	0x00	0x0013	0x00	0x00	0x00	0x00	0xCC
CoincWind&Offset12	0x00	0x0014	0x00	0x00	0x00	0x00	0xCC
40MhzClk&Or254	0x00	0x001C	0xEE	0xEE	0xEE	0xEE	0xAE
MaskChannel-008-to-001	0x00	0x0020	0x00	0x00	0xAA	0x00	0x00
MaskChannel-016-to-009	0x00	0x0021	0x00	0x00	0xAA	0x00	0x00
MaskChannel-024-to-017	0x00	0x0022	0x00	0x00	0xAA	0x00	0x03
MaskChannel-032-to-025	0x00	0x0023	0x00	0x00	0xAA	0x00	0x00
MaskChannel-040-to-033	0x00	0x0024	0x00	0x20	0xAA	0x00	0x03
MaskChannel-048-to-041	0x00	0x0025	0x00	0x00	0xAA	0x00	0x00
MaskChannel-056-to-049	0x00	0x0026	0x00	0x01	0xAA	0x00	0x00
MaskChannel-064-to-057	0x00	0x0027	0x00	0x00	0xAA	0x00	0x00
MaskChannel-072-to-065	0x00	0x0028	0x80	0x00	0xAA	0x00	0x03
MaskChannel-080-to-073	0x00	0x0029	0x02	0x00	0xAA	0x00	0x00
MaskChannel-088-to-081	0x00	0x002A	0x14	0x80	0xAA	0x00	0x00
MaskChannel-096-to-089	0x00	0x002B	0x00	0x00	0xAA	0x00	0x00
MaskChannel-104-to-097	0x00	0x002C	0x00	0x04	0xAA	0x00	0x00
MaskChannel-112-to-105	0x00	0x002D	0x00	0x00	0xAA	0x00	0x00
MaskChannel-120-to-113	0x00	0x002E	0x00	0x00	0xAA	0x00	0x00
MaskChannel-128-to-121	0x00	0x002F	0x00	0x00	0xAA	0x09	0x00
MaskChannel-136-to-129	0x00	0x0030	0x00	0x00	0xAA	0x00	0x00
MaskChannel-144-to-137	0x00	0x0031	0x00	0x00	0xAA	0x00	0x00
MaskChannel-152-to-145	0x00	0x0032	0x00	0x00	0xAA	0x00	0x00
MaskChannel-160-to-153	0x00	0x0033	0x20	0x20	0xAA	0x00	0x00
MaskChannel-168-to-161	0x00	0x0034	0x00	0x00	0xAA	0x00	0x00
MaskChannel-176-to-169	0x00	0x0035	0x01	0x01	0xAA	0x00	0x00
MaskChannel-184-to-177	0x00	0x0036	0x00	0x00	0xAA	0x00	0x00
MaskChannel-192-to-185	0x00	0x0037	0x00	0x00	0xAA	0x04	0x00
MaskChannel-200-to-193	0x00	0x0038	0x00	0x00	0xAA	0x02	0x00
MaskChannel-208-to-201	0x00	0x0039	0x00	0x00	0xAA	0x00	0x00
MaskChannel-216-to-209	0x00	0x003A	0x00	0x00	0xAA	0x24	0x00
MaskChannel-224-to-217	0x00	0x003B	0x00	0x00	0xAA	0x00	0x00
MaskChannel-232-to-225	0x00	0x003C	0x00	0x00	0xAA	0x00	0x00
MaskChannel-240-to-233	0x00	0x003D	0x00	0x00	0xAA	0x00	0x00
MaskChannel-248-to-241	0x00	0x003E	0x00	0x00	0xAA	0x00	0x00
MaskChannel-254-to-249	0x00	0x003F	0x00	0x00	0xAA	0x00	0x00
VCth1	0x00	0x004F	0x20	0x20	0x20	0x20	0x3A
VCth2	0x00	0x0050	0x03	0x03	0x03	0x03	0x02

Appendix D

$\bar{S}$ events: reconstruction efficiencies

This appendix gives the reconstruction efficiency for the final state particles, the V^0s and the $\bar{S}$ in reconstructable (section 10.10.5) simulated $\bar{S}$ events. Efficiencies are provided as function of transverse and longitudinal displacement, transverse momentum and pseudorapidity. The actual distribution of these variables are also shown. A discussion on these results is given in section 10.11.

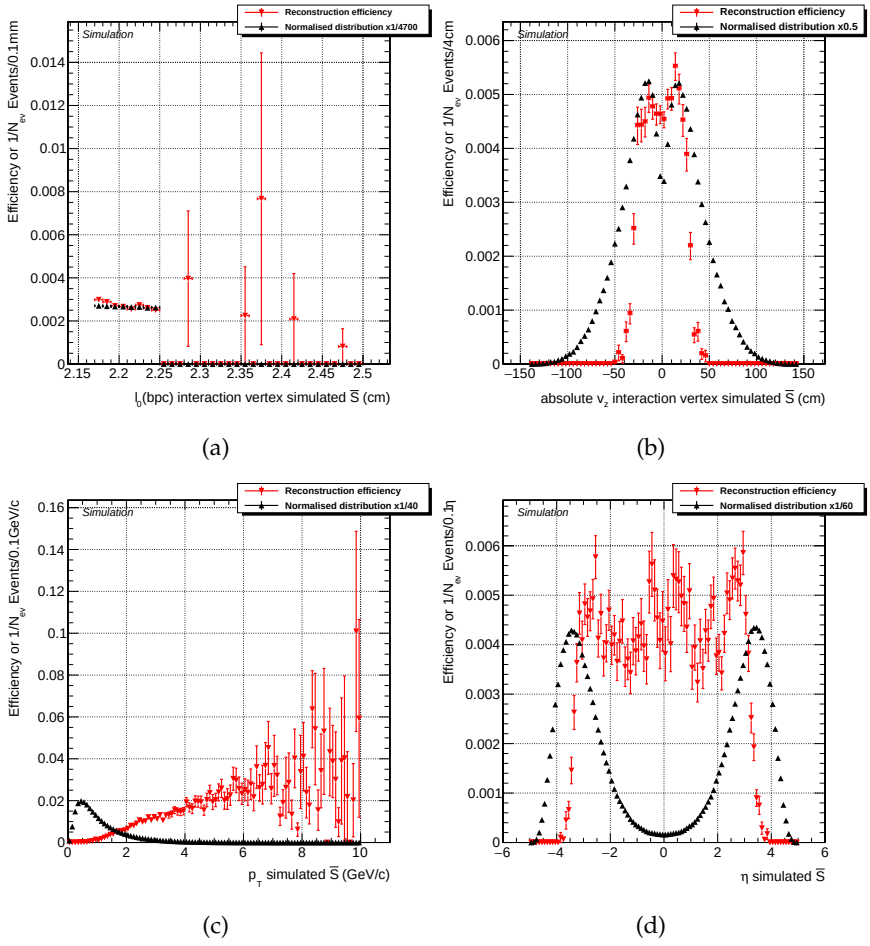
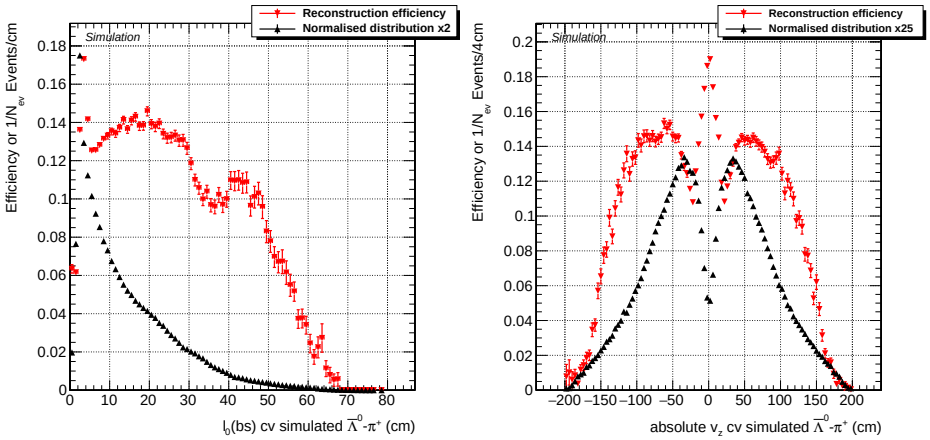
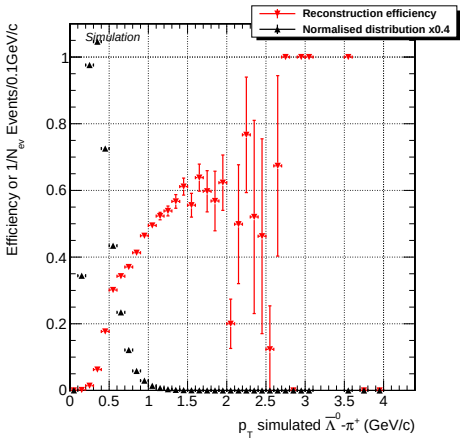


FIGURE D.1: Reconstruction efficiency for the $\bar{S}$ in reconstructable events. The actual distribution of the variables is also shown.

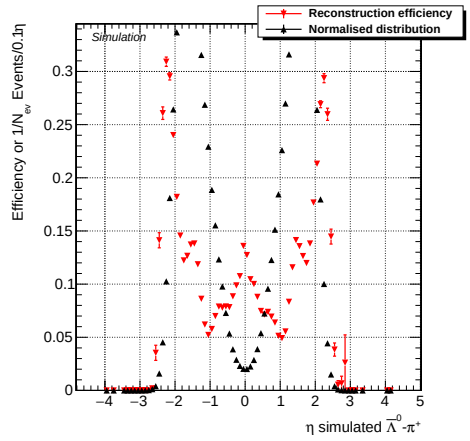


(a)

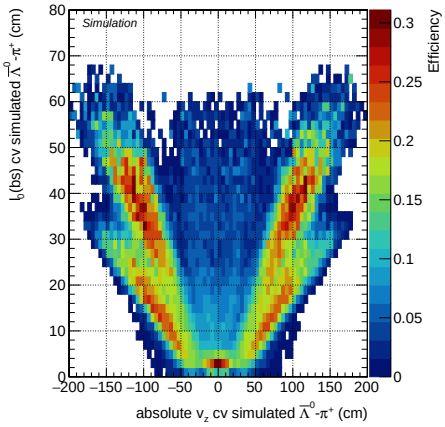
(b)



(c)



(d)



(e)

FIGURE D.2: Tracking efficiency for the π^+ in the $\bar{\Lambda}^0$ decay in reconstructable events. The actual distribution of the variables is also shown in the first four figures.

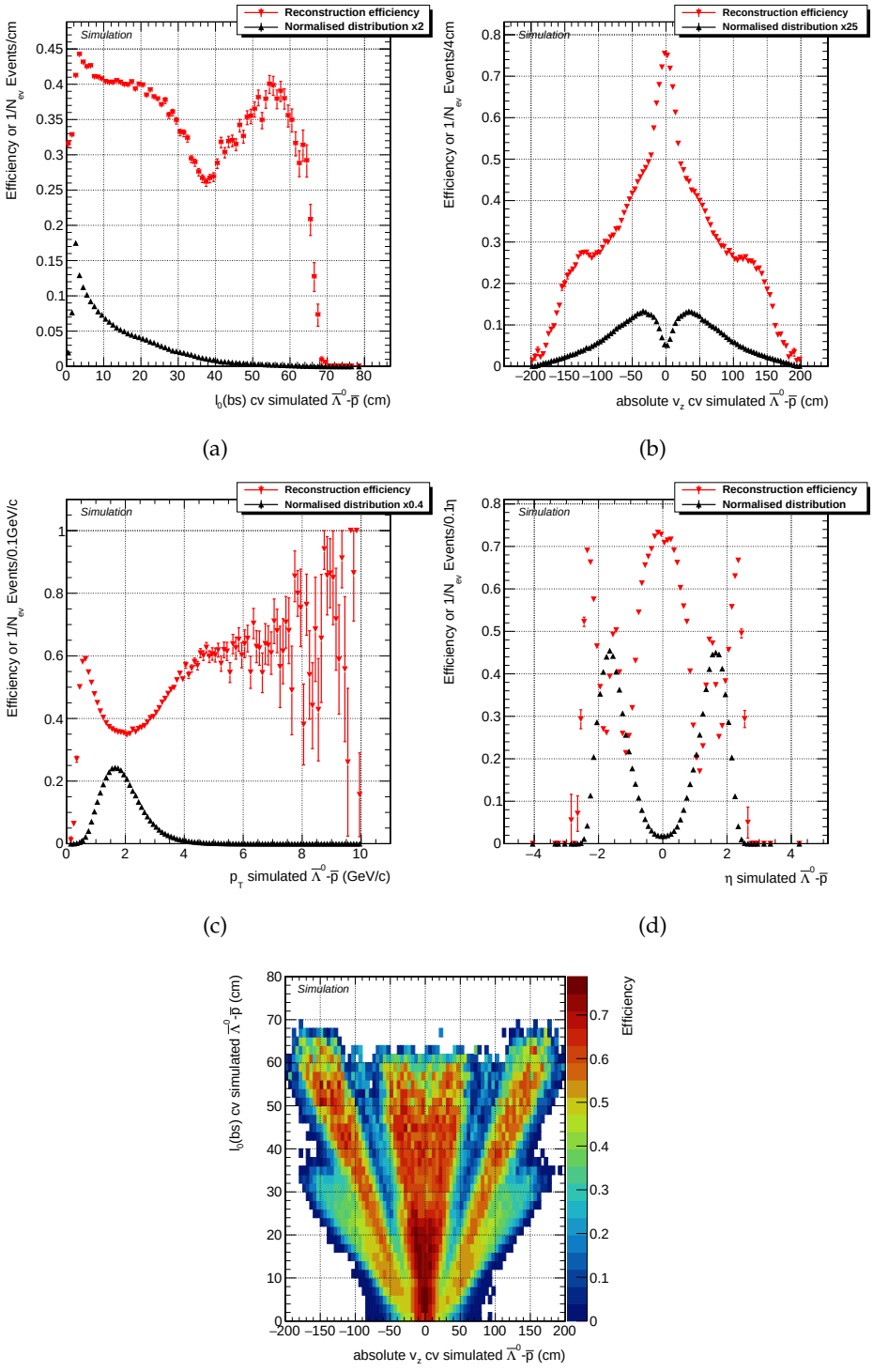
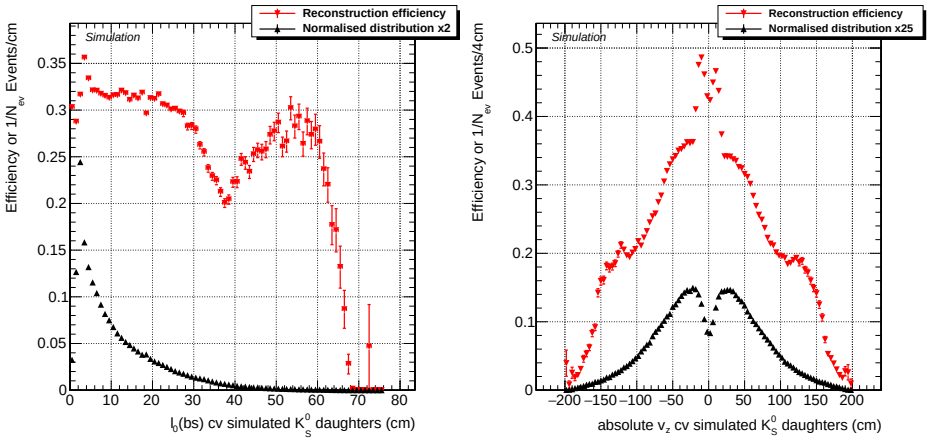
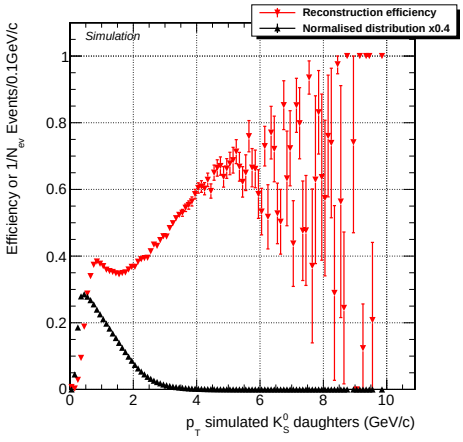


FIGURE D.3: Tracking efficiency for the $\bar{p}$ in the $\bar{\Lambda}^0$ decay in reconstructable events. The actual distribution of the variables is also shown in the first four figures.

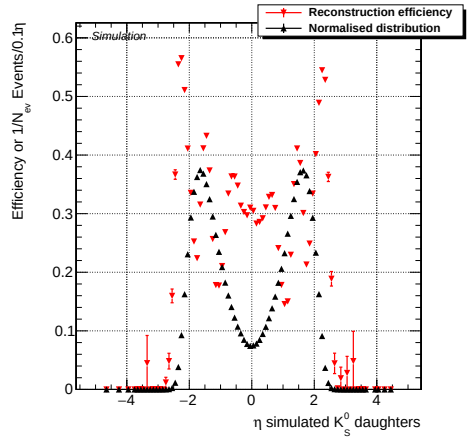


(a)

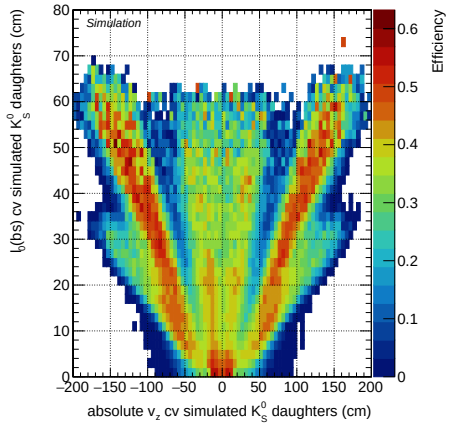
(b)



(c)

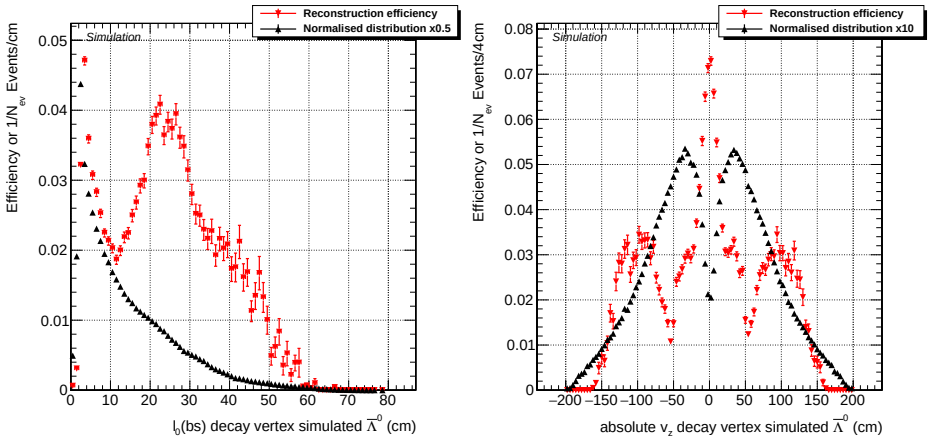


(d)



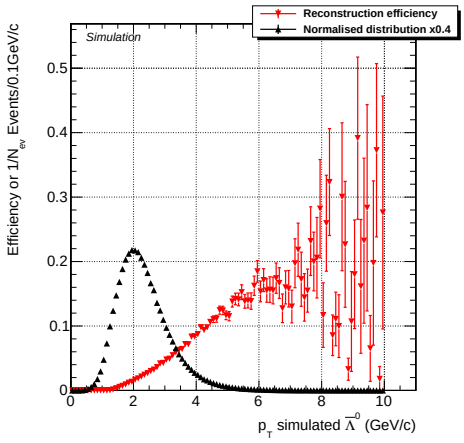
(e)

FIGURE D.4: Tracking efficiency for the pions in the K^0_S decay in reconstructable events. The actual distribution of the variables is also shown in the first four figures.

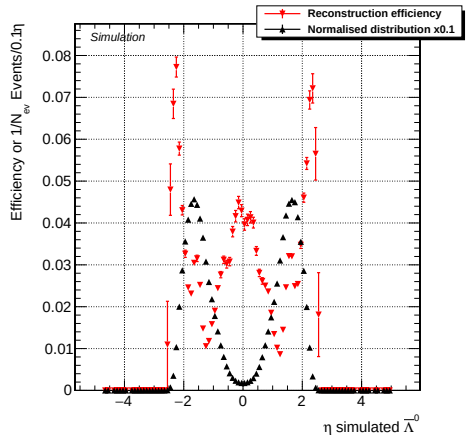


(a)

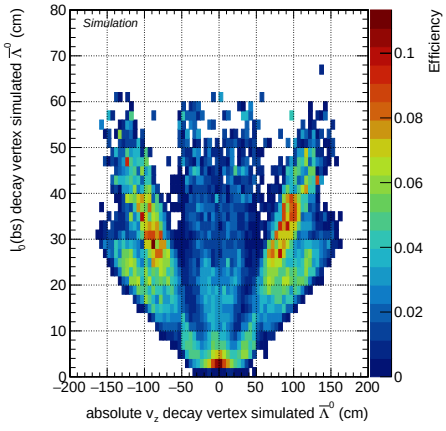
(b)



(c)

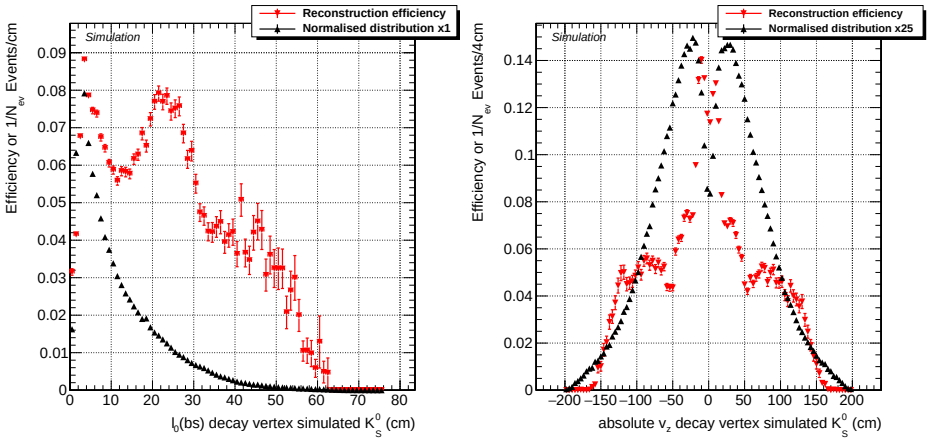


(d)



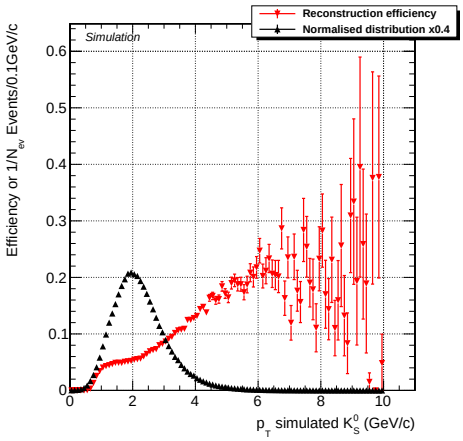
(e)

FIGURE D.5: Reconstruction efficiency for the $\bar{\Lambda}^0$ in reconstructable events. The actual distribution of the variables is also shown in the first four figures.

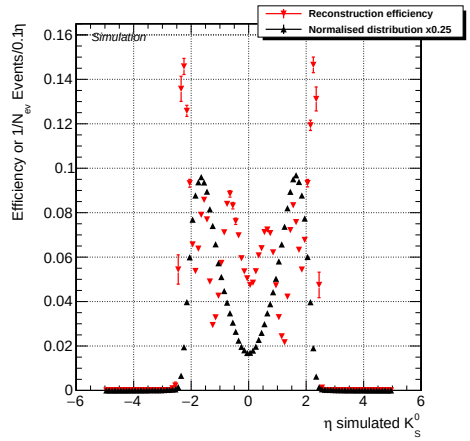


(a)

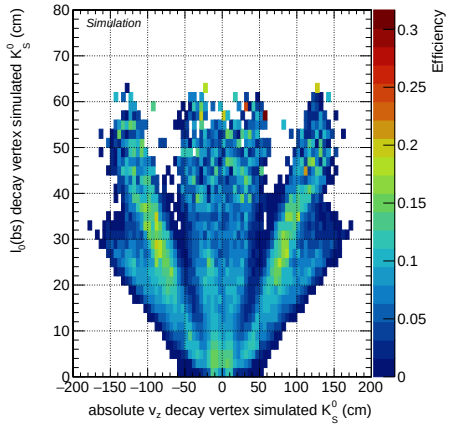
(b)



(c)



(d)



(e)

FIGURE D.6: Reconstruction efficiency for the K^0_S in reconstructable events. The actual distribution of the variables is also shown in the first four figures.

Appendix E

Input variables to the Sexaquark BDT: pre-BDT

This appendix shows the distributions of the input variables to the BDT after applying the cuts described in the figure captions. Six collections are compared to each other:

- Reconstructed background S in DYJets MC.
- Reconstructed background $\bar{S}$ in DYJets MC.
- Reconstructed background S in SingleMuon Run2016H data.
- Reconstructed background $\bar{S}$ in SingleMuon Run2016H data. This is guaranteed to be background by applying an upper limit on the BDT classifier value as mentioned in the legend of the figures.
- $\bar{S}$ reconstructed across events in SingleMuon Run2016H data.
- Reconstructed $\bar{S}$ from signal MC.

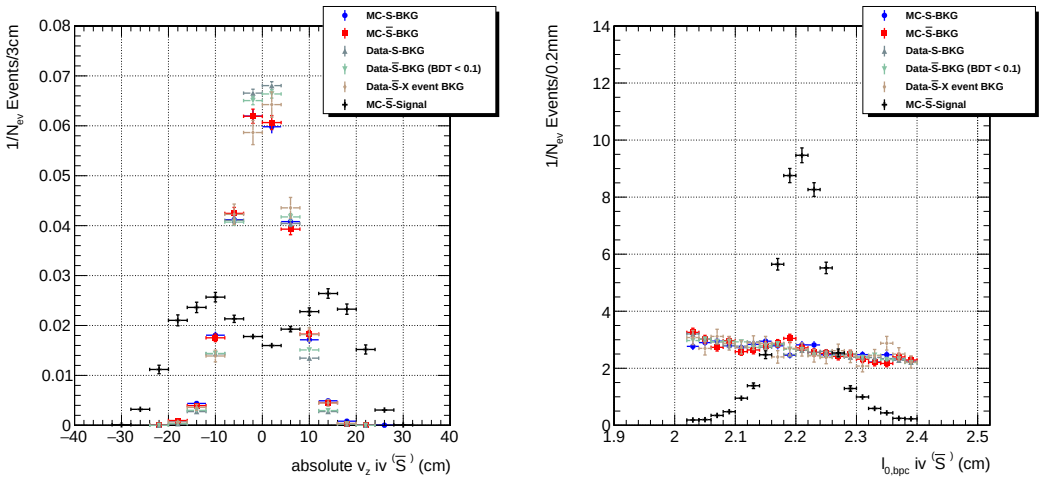


FIGURE E.1: Longitudinal and transverse (w.r.t. the beampipe center) location of the reconstructed interaction vertex of signal MC $\bar{S}$ and S and $\bar{S}$ from the five background samples for S or $\bar{S}$ of which the final state particles lie within the fiducial region as defined in section 10.13 and which have $|\Delta\phi(K^0_S, \bar{\Lambda}^0/\Lambda^0)| > 0.5$, 2.02 cm $< l_{0,\mathrm{bpc}}(\mathrm{iv}) <$ 2.40 cm and $0 \leq d_{0,\mathrm{bs}}(\bar{S}/S)/l_{0,\mathrm{bs}}(\bar{S}/S) < 0.5$.

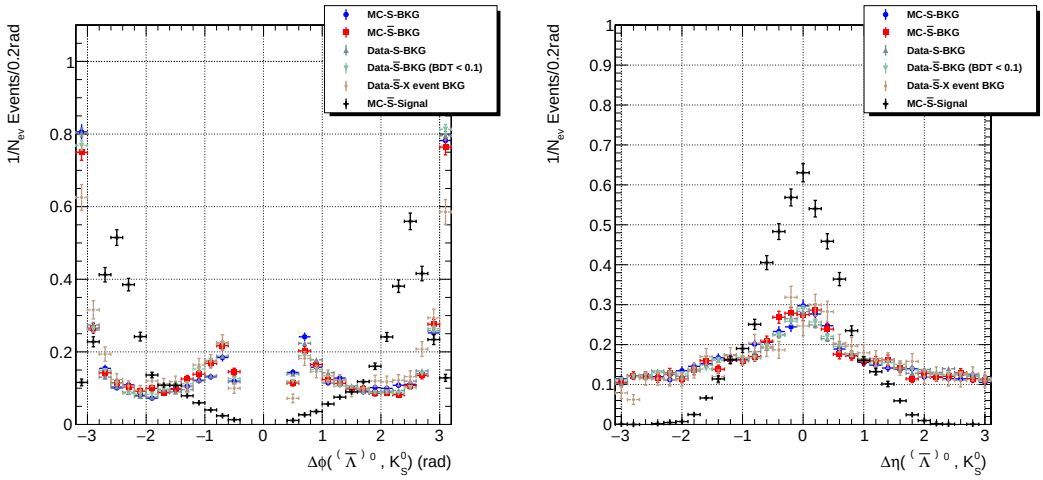


FIGURE E.2: Difference in ϕ and η between the reconstructed momentum vectors of V^0s from signal MC $\bar{S}$ and S and $\bar{S}$ from the five background samples for S or $\bar{S}$ of which the final state particles lie within the fiducial region as defined in section 10.13 and which have $|\Delta\phi(K^0_S, \bar{\Lambda}^0/\Lambda^0)| > 0.5$, 2.02 cm $< l_{0,\mathrm{bpc}}(\mathrm{iv}) <$ 2.40 cm and $0 \leq d_{0,\mathrm{bs}}(\bar{S}/S)/l_{0,\mathrm{bs}}(\bar{S}/S) < 0.5$.

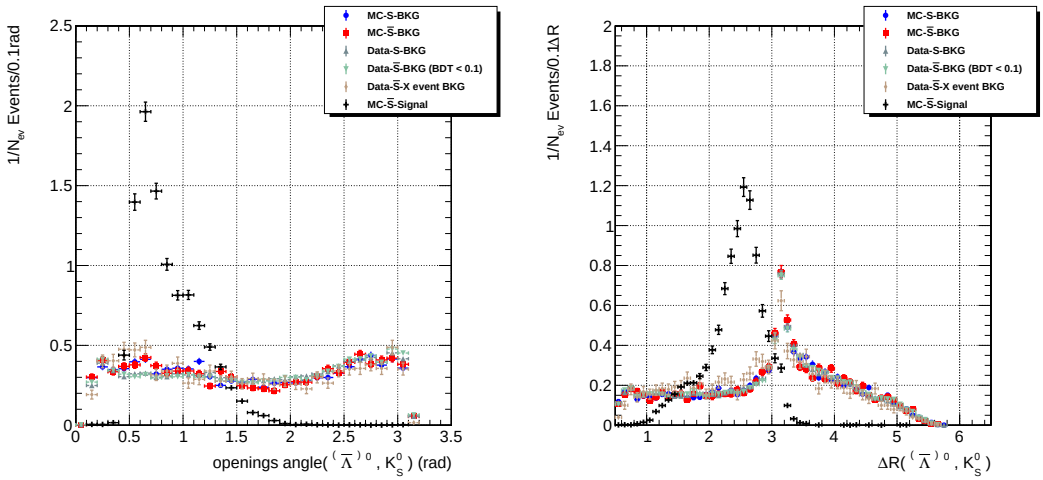


FIGURE E.3: 3D-openingsangle and ΔR between the reconstructed momentum vectors of V^0s from signal MC $\bar{S}$ and S and $\bar{S}$ from the five background samples for S or $\bar{S}$ of which the final state particles lie within the fiducial region as defined in section 10.13 and which have $|\Delta\phi(K^0_S, \bar{\Lambda}^0/\Lambda^0)| > 0.5$, 2.02 cm $< l_{0,\mathrm{bpc}}(\mathrm{iv}) <$ 2.40 cm and $0 \leq d_{0,\mathrm{bs}}(\bar{S}/S)/l_{0,\mathrm{bs}}(\bar{S}/S) < 0.5$.

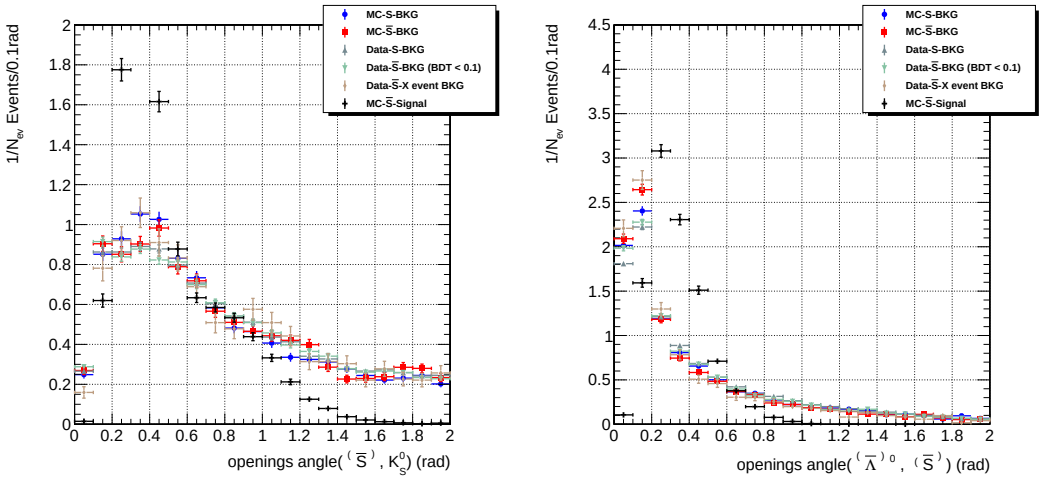


FIGURE E.4: 3D-openingsangles between the reconstructed momentum vectors of signal MC $\bar{S}$ and the V^0s and background S or $\bar{S}$ and the V^0s. Only S or $\bar{S}$ are used of which the final state particles lie within the fiducial region as defined in section 10.13 and which have $|\Delta\phi(K^0_S, \bar{\Lambda}^0/\Lambda^0)| > 0.5$, 2.02 cm $< l_{0,\mathrm{bpc}}(\mathrm{iv}) <$ 2.40 cm and $0 \leq d_{0,\mathrm{bs}}(\bar{S}/S)/l_{0,\mathrm{bs}}(\bar{S}/S) < 0.5$.

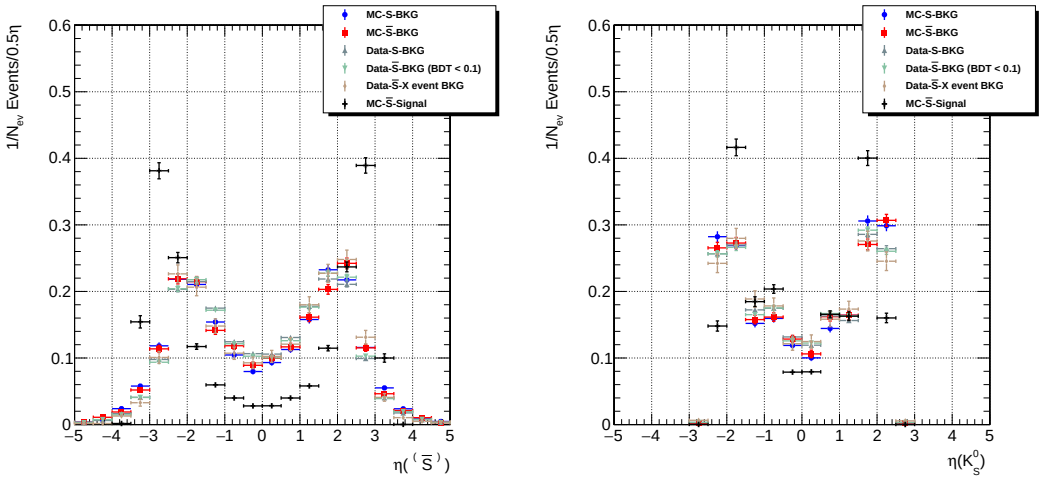


FIGURE E.5: Pseudorapidities of the reconstructed S or $\bar{S}$ and K^0_S from MC signal and the five background samples for S or $\bar{S}$ of which the final state particles lie within the fiducial region as defined in section 10.13 and which have $|\Delta\phi(K^0_S, \bar{\Lambda}^0/\Lambda^0)| > 0.5$, 2.02 cm $< l_{0,\mathrm{bpc}}(\mathrm{iv}) <$ 2.40 cm and $0 \leq d_{0,\mathrm{bs}}(\bar{S}/S)/l_{0,\mathrm{bs}}(\bar{S}/S) < 0.5$.

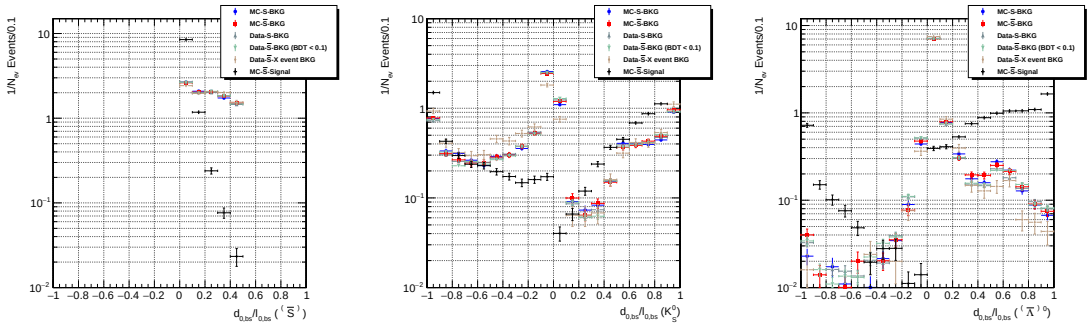


FIGURE E.6: Transverse impact parameter divided by the transverse displacement of the interaction (S/$\bar{\text{S}}$) or decay vertex (V^0s), both calculated with respect to the beamspot of reconstructed S or $\bar{\text{S}}$, K^0_S and Λ^0 or $\bar{\Lambda}^0$ from MC $\bar{\text{S}}$ signal and the five background samples for S or $\bar{\text{S}}$ of which the final state particles lie within the fiducial region as defined in section 10.13 and which have $|\Delta\phi(\text{K}^0_\text{S}, \bar{\Lambda}^0/\Lambda^0)| > 0.5$, 2.02 cm $< l_{0,\text{bpc}}(\text{iv}) <$ 2.40 cm and $0 \leq d_{0,\text{bs}}(\bar{\text{S}}/\text{S})/l_{0,\text{bs}}(\bar{\text{S}}/\text{S}) < 0.5$.

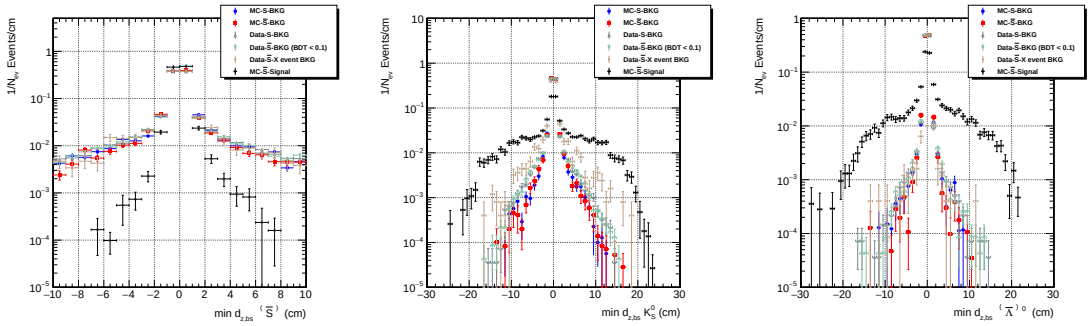


FIGURE E.7: Longitudinal impact parameter calculated with respect to the valid PV which minimises this longitudinal impact parameter for reconstructed S or $\bar{\text{S}}$, K^0_S and Λ^0 or $\bar{\Lambda}^0$ from MC signal and the five background samples for S or $\bar{\text{S}}$ of which the final state particles lie within the fiducial region as defined in section 10.13 and which have $|\Delta\phi(\text{K}^0_\text{S}, \bar{\Lambda}^0/\Lambda^0)| > 0.5$, 2.02 cm $< l_{0,\text{bpc}}(\text{iv}) <$ 2.40 cm and $0 \leq d_{0,\text{bs}}(\bar{\text{S}}/\text{S})/l_{0,\text{bs}}(\bar{\text{S}}/\text{S}) < 0.5$.

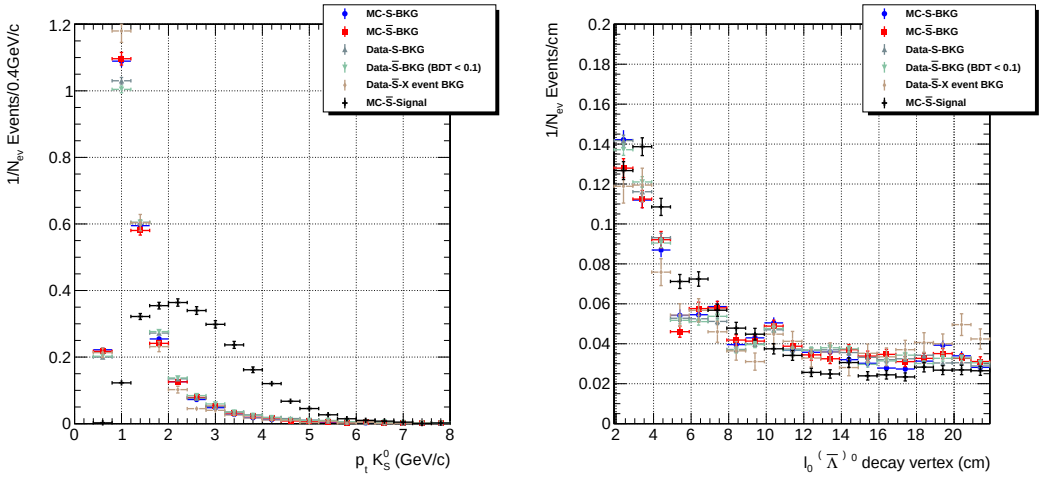


FIGURE E.8: Transverse momenta of the K^0_S and transverse displacement of the Λ^0 or $\bar{\Lambda}^0$ decay vertex from MC $\bar{S}$ signal and S and $\bar{S}$ from the five background samples for S or $\bar{S}$ of which the final state particles lie within the fiducial region as defined in section 10.13 and which have $|\Delta\phi(K^0_S, \bar{\Lambda}^0/\Lambda^0)| > 0.5$, 2.02 cm $< l_{0,\mathrm{bpc}}(\mathrm{iv}) <$ 2.40 cm and $0 \leq d_{0,\mathrm{bs}}(\bar{S}/S)/l_{0,\mathrm{bs}}(\bar{S}/S) < 0.5$.

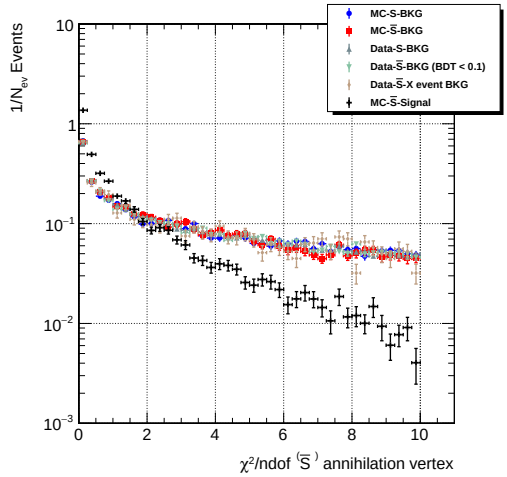


FIGURE E.9: $\chi^2/ndof$ of the S or $\bar{S}$ reconstructed vertex of the MC $\bar{S}$ and S and $\bar{S}$ from the five background samples for S or $\bar{S}$ of which the final state particles lie within the fiducial region as defined in section 10.13 and which have $|\Delta\phi(K^0_S, \bar{\Lambda}^0/\Lambda^0)| > 0.5$, 2.02 cm $< l_{0,bpc} <$ 2.40 cm and $0 \leq d_{0,\mathrm{bs}}(\bar{S}/S)/l_{0,\mathrm{bs}}(\bar{S}/S) < 0.5$.

Appendix F

Input variables to the Sexaquark BDT: post-BDT

This appendix shows the distributions of the input variables to the BDT after applying the cuts described in the figure captions. The difference with Appendix E is that the results shown here are only for events which pass a certain cut on the BDT classifier value (as mentioned in the captions). Six collections are compared to each other:

- Reconstructed background S in DYJets MC.
- Reconstructed background $\bar{S}$ in DYJets MC.
- Reconstructed background S in SingleMuon Run2016H data.
- Reconstructed background $\bar{S}$ in SingleMuon Run2016H data. This is guaranteed to be background by applying an upper limit on the BDT classifier value as mentioned in the legend of the figures.
- $\bar{S}$ reconstructed across events in SingleMuon Run2016H data.
- Reconstructed $\bar{S}$ from signal MC.

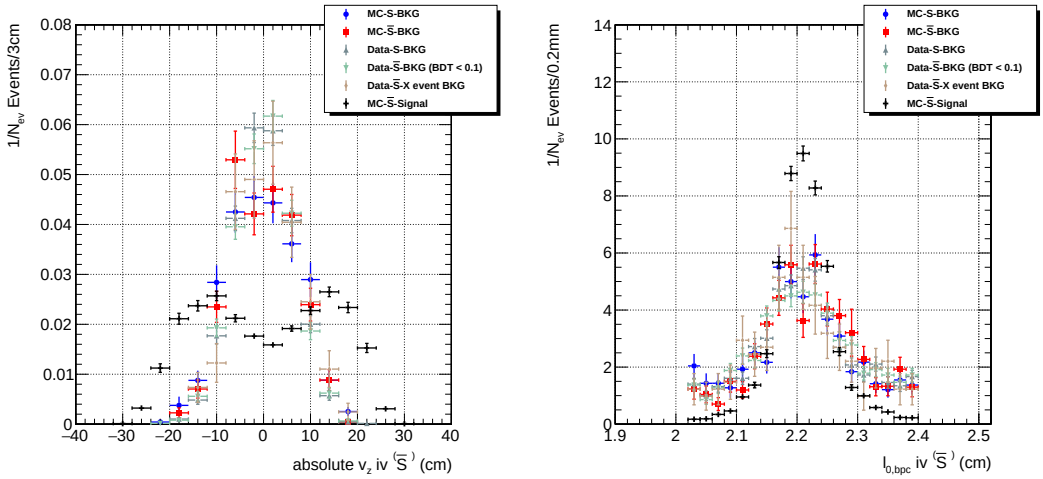


FIGURE F.1: Longitudinal and transverse (w.r.t. the beampipe center) location of the reconstructed interaction vertex of signal MC $\bar{\mathrm{S}}$ and S and $\bar{\mathrm{S}}$ from the five background samples for S or $\bar{\mathrm{S}}$ of which the final state particles lie within the fiducial region as defined in section 10.13 and which have $|\Delta\phi(\mathrm{K}^0_\mathrm{S}, \bar{\Lambda}^0/\Lambda^0)| > 0.5$, 2.02 cm $< l_{0,\mathrm{bpc}}(\mathrm{iv}) <$ 2.40 cm and $0 \leq d_{0,\mathrm{bs}}(\bar{\mathrm{S}}/\mathrm{S})/l_{0,\mathrm{bs}}(\bar{\mathrm{S}}/\mathrm{S}) < 0.5$ and with the BDT classifier value $\geq$ -0.15.

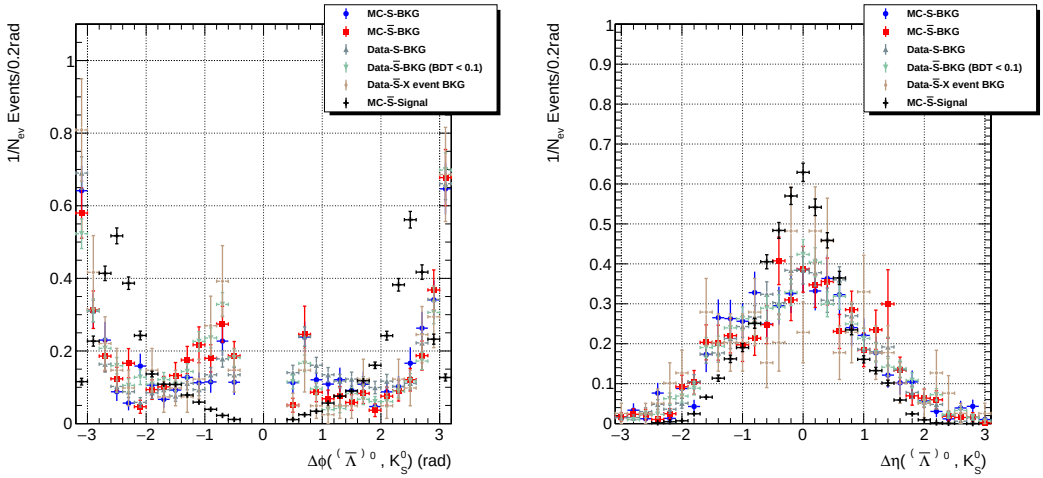


FIGURE F.2: Difference in ϕ and η between the reconstructed momentum vectors of V^0s from signal MC $\bar{\mathrm{S}}$ and S and $\bar{\mathrm{S}}$ from the five background samples for S or $\bar{\mathrm{S}}$ of which the final state particles lie within the fiducial region as defined in section 10.13 and which have $|\Delta\phi(\mathrm{K}^0_\mathrm{S}, \bar{\Lambda}^0/\Lambda^0)| > 0.5$, 2.02 cm $< l_{0,\mathrm{bpc}}(\mathrm{iv}) <$ 2.40 cm and $0 \leq d_{0,\mathrm{bs}}(\bar{\mathrm{S}}/\mathrm{S})/l_{0,\mathrm{bs}}(\bar{\mathrm{S}}/\mathrm{S}) < 0.5$ and with the BDT classifier value $\geq$ -0.15.

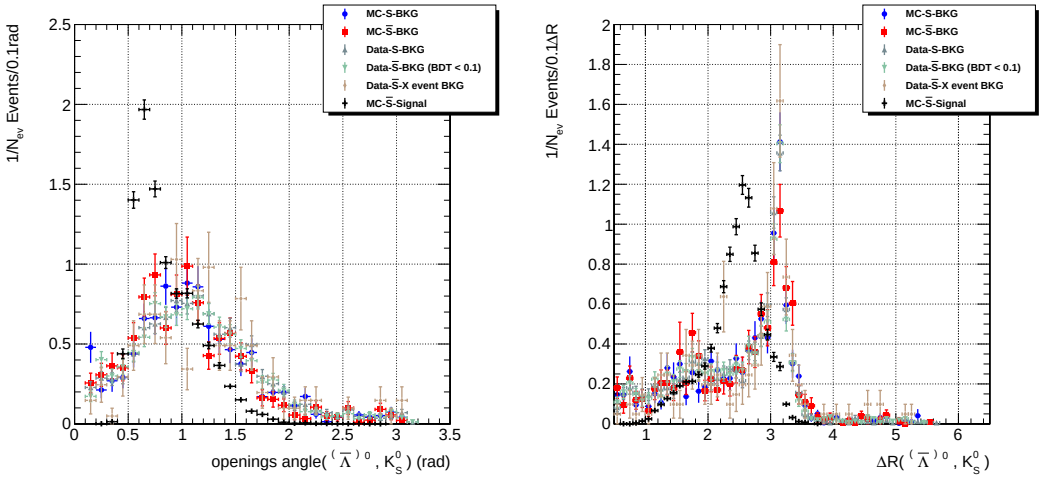


FIGURE F.3: 3D-openingsangle and ΔR between the reconstructed momentum vectors of V^0s from signal MC $\bar{S}$ and S and $\bar{S}$ from the five background samples for S or $\bar{S}$ of which the final state particles lie within the fiducial region as defined in section 10.13 and which have $|\Delta\phi(K^0_S, \bar{\Lambda}^0/\Lambda^0)| > 0.5$, 2.02 cm $< l_{0,\mathrm{bpc}}(\mathrm{iv}) <$ 2.40 cm and $0 \leq d_{0,\mathrm{bs}}(\bar{S}/S)/l_{0,\mathrm{bs}}(\bar{S}/S) < 0.5$ and with the BDT classifier value $\geq$ -0.15.

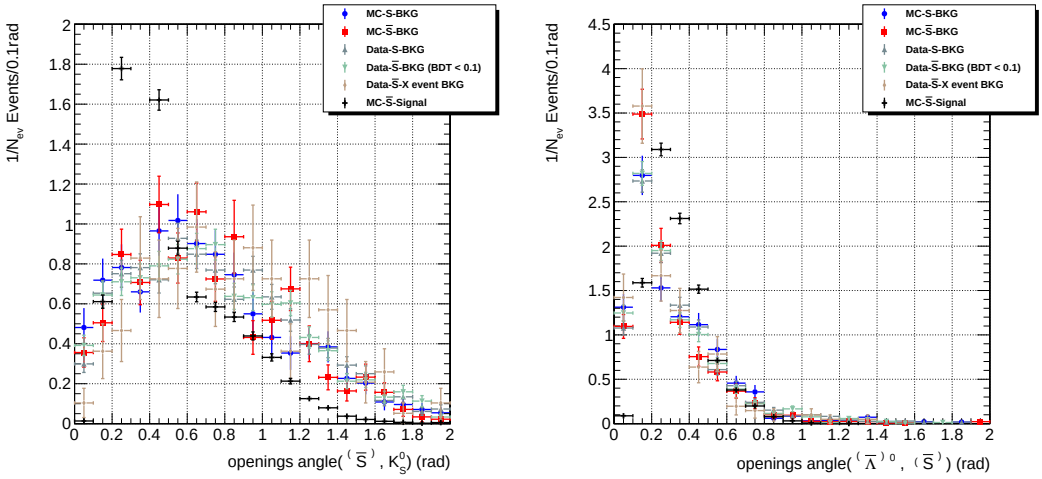


FIGURE F.4: 3D-openingsangles between the reconstructed momentum vectors of signal MC $\bar{S}$ and the V^0s and background S or $\bar{S}$ and the V^0s. Only S or $\bar{S}$ are used of which the final state particles lie within the fiducial region as defined in section 10.13 and which have $|\Delta\phi(K^0_S, \bar{\Lambda}^0/\Lambda^0)| > 0.5$, 2.02 cm $< l_{0,\mathrm{bpc}}(\mathrm{iv}) <$ 2.40 cm and $0 \leq d_{0,\mathrm{bs}}(\bar{S}/S)/l_{0,\mathrm{bs}}(\bar{S}/S) < 0.5$ and with the BDT classifier value $\geq$ -0.15.

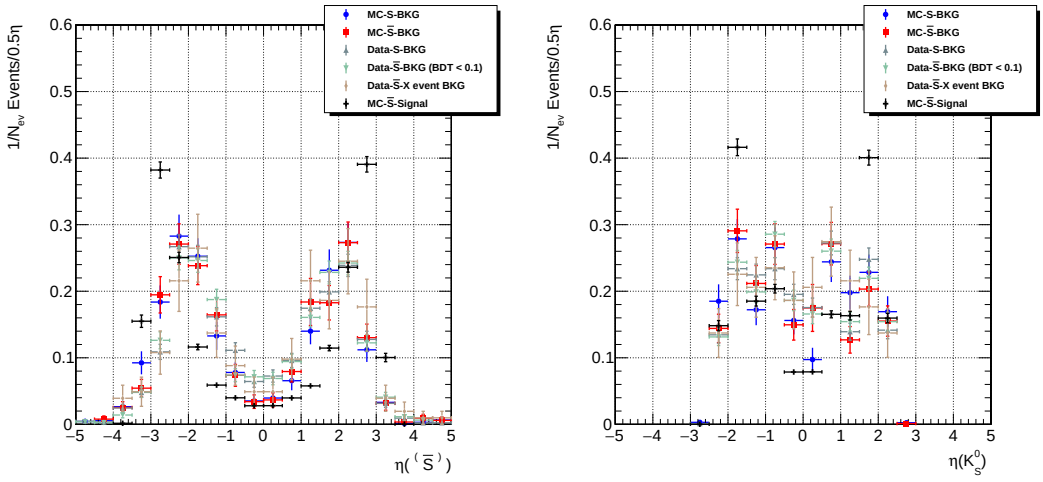


FIGURE F.5: Pseudorapidities of the reconstructed S or $\bar{S}$ and K^0_S from MC signal and the five background samples for S or $\bar{S}$ of which the final state particles lie within the fiducial region as defined in section 10.13 and which have $|\Delta\phi(K^0_S, \bar{\Lambda}^0/\Lambda^0)| > 0.5$, 2.02 cm $< l_{0,\mathrm{bpc}}(\mathrm{iv}) <$ 2.40 cm and $0 \leq d_{0,\mathrm{bs}}(\bar{S}/S)/l_{0,\mathrm{bs}}(\bar{S}/S) < 0.5$ and with the BDT classifier value $\geq$ -0.15.

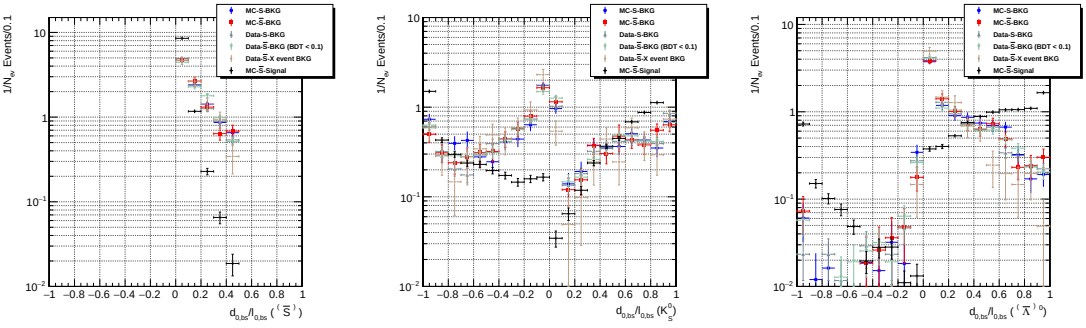


FIGURE F.6: Transverse impact parameter divided by the transverse displacement of the interaction ($S/\bar{S}$) or decay vertex (V^0s), both calculated with respect to the beamspot of reconstructed S or $\bar{S}$, K^0_S and Λ^0 or $\bar{\Lambda}^0$ from MC $\bar{S}$ signal and the five background samples for S or $\bar{S}$ of which the final state particles lie within the fiducial region as defined in section 10.13 and which have $|\Delta\phi(K^0_S, \bar{\Lambda}^0/\Lambda^0)| >$ 0.5, 2.02 cm $< l_{0,\mathrm{bpc}}(\mathrm{iv}) <$ 2.40 cm and $0 \leq d_{0,\mathrm{bs}}(\bar{S}/S)/l_{0,\mathrm{bs}}(\bar{S}/S) < 0.5$ and with the BDT classifier value $\geq$ -0.15.

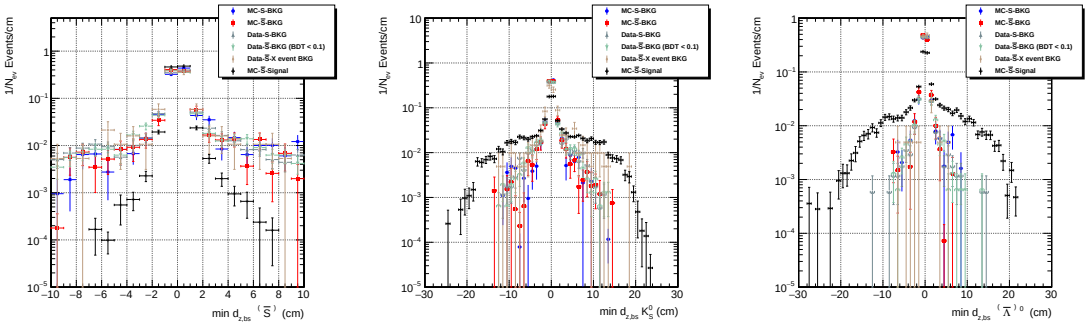


FIGURE F.7: Longitudinal impact parameter calculated with respect to the valid PV which minimises this longitudinal impact parameter for reconstructed S or $\bar{S}$, K^0_S and Λ^0 or $\bar{\Lambda}^0$ from MC signal and the five background samples for S or $\bar{S}$ of which the final state particles lie within the fiducial region as defined in section 10.13 and which have $|\Delta\phi(K^0_S, \bar{\Lambda}^0/\Lambda^0)| > 0.5$, 2.02 cm $< l_{0,\mathrm{bpc}}(\mathrm{iv}) <$ 2.40 cm and $0 \leq d_{0,\mathrm{bs}}(\bar{S}/S)/l_{0,\mathrm{bs}}(\bar{S}/S) < 0.5$ and with the BDT classifier value $\geq$ -0.15.

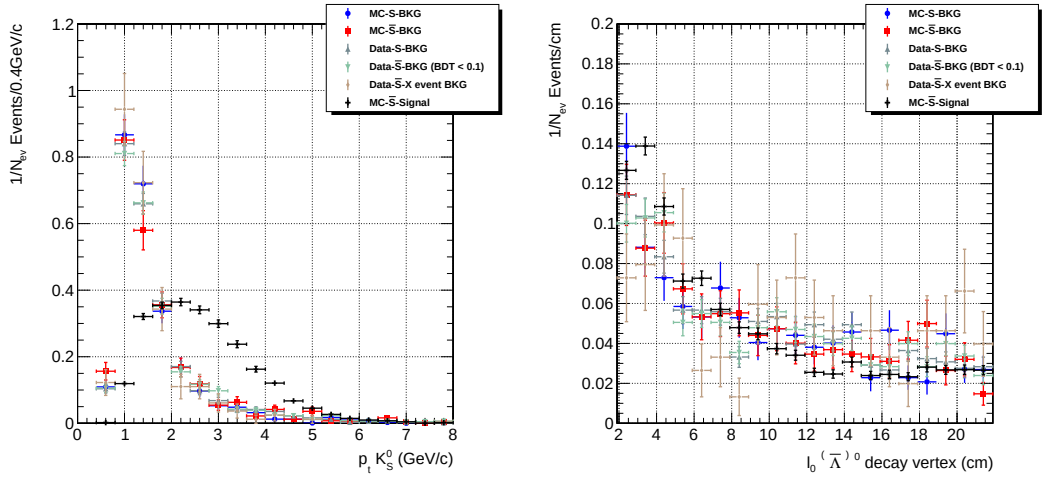


FIGURE F.8: Transverse momenta of the K^0_S and transverse displacement of the Λ^0 or $\bar{\Lambda}^0$ decay vertex from MC $\bar{S}$ signal and S and $\bar{S}$ from the five background samples for S or $\bar{S}$ of which the final state particles lie within the fiducial region as defined in section 10.13 and which have $|\Delta\phi(K^0_S, \bar{\Lambda}^0/\Lambda^0)| > 0.5$, 2.02 cm $< l_{0,\mathrm{bpc}}(\mathrm{iv}) <$ 2.40 cm and $0 \leq d_{0,\mathrm{bs}}(\bar{S}/S)/l_{0,\mathrm{bs}}(\bar{S}/S) < 0.5$ and with the BDT classifier value $\geq$ -0.15.

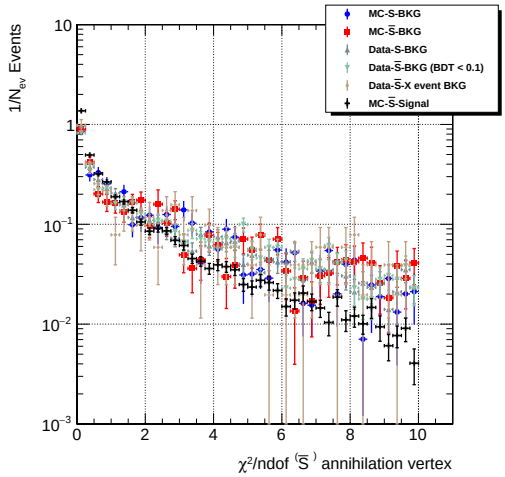


FIGURE F.9: $\chi^2/ndof$ of the S or $\bar{S}$ reconstructed vertex of the MC $\bar{S}$ and S and $\bar{S}$ from the five background samples for S or $\bar{S}$ of which the final state particles lie within the fiducial region as defined in section 10.13 and which have $|\Delta\phi(K^0_S, \bar{\Lambda}^0/\Lambda^0)| > 0.5$, 2.02 cm $< l_{0,bpc} <$ 2.40 cm and $0 \leq d_{0,bs}(\bar{S}/S)/l_{0,bs}(\bar{S}/S) < 0.5$ and with the BDT classifier value $\geq$ -0.15.

List of Abbreviations

2S	Strip-Strip
AC	Alternating Current
ADC	Analogue to Digital Converter
ALICE	A Large Ion Collider Experiment
AMC	Advanced Mezzanine Card
AMUX	Analogue MUltipleXer
APV	Front-end Readout Chip: Analogue Pipeline (Voltage mode)
ASIC	Application-Specific Integrated Circuit
ATCA	Advanced Telecommunications Computing Architecture
ATLAS	A Toroidal LHC ApparatuS
BBN	Big Bang Nucleosynthesis
BDT	Boosted Decision Tree
BE	Binding Energy
BEH	Brout-Englert-Higgs
BRAM	Block Random-Access Memory
BX	Bunch Crossing
CBC	CMS Binary Chip
CE-E	HGCAL Electromagnetic section
CE-H	HGCAL Hadronic section
CE	HGCAL
CIC	Concentrator Integrated Circuit
CL	Confidence Limit
CLB	Configurable Logic Block
CMB	Cosmic Microwave Background
CMOS	Complementary Metal-Oxide-Semiconductor
CMS	Compact Muon Solenoid
CRC	Cyclotron Resource Centre
CSC	Cathode Strip Chamber
DAC	Digital to Analogue Converter
DAQ	Data AcQuisition
DC	Direct Current
DDR3	Double Data Rate 3
DIO5	5-bit port digital IO card in FMC form factor
DLL	Delay-Locked Loop
DM	Dark Matter
DQM	Data Quality Monitoring
DRAM	Dynamic Random-Access Memory
DSP	Digital Signal Processing
DT	Drift Tube
DTC	Data, Trigger and Control
DUT	Device Under Test
DY	Drell-Yan
EB	ECAL Barrel

EC	ECAL Endcap
ECAL	Electromagnetic CALorimter
EP-ESE	Experimental Physics-Electronic Systems for Experiments
FEC	Forward Error Correction
FEH	Front-End Hybrid
FET	Field-Effect Transistor
FIFO	First In, First Out
FLUKA	FLUktuierende KAskade
FMC	FPGA Mezzanine Card
FPGA	Field-Programmable Gate Array
FW	FirmWare
GBT	GigaBit Transceiver
GEM	Gas Electron Multiplier
HB	HCAL Barrel
HCAL	Hadronic CALorimeter
HE	HCAL Endcap
HF	HCAL Forward
HGCAL	High Granularity Calorimeter
HIF	Heavy Ion Facility
HIP	Highly Ionizing Particle
HL-LHC	High Luminosity-Large Hadron Collider
HLT	High-Level Trigger
HO	HCAL Outer
HV	High Voltage
I/0	Input/Output
I^2C	Inter-Integrated Circuit
IC	Integrated Circuit
IDELAY	Input Delay
ILC	International Linear Collider
IP	Internet Protocol
iRPC	improved RPC
ISERDES	Input SERializer/DESerializer
L1	Level1
L1A	L1-Accept
LED	Light Emitting Diode
LET	Linear Energy Transfer
LGAD	Low-Gain-Avalanche-Diode
LHC	Large Hadron Collider
LHCb	Large Hadron Collider beauty
LINAC	LINear ACcelerator
lpGBT	low-power GigaBit Transceiver
LSB	Least Significant Bit
LUT	Look-Up Table
LV	Low Voltage
LVDS	Low-Voltage Differential Signaling
MACHO	Massive Astrophysical Compact Halo Object
MaPSA	Macro-Pixel Sub-Assembly
MBU	Multi-Bit Upset
MC	Monte Carlo

MCH	MicroTCA Carrier Hub
MIP	Minimum Ionizing Particle
MMCM	Mixed-Mode Clock Manager
MOS	Metal-Oxide-Semiconductor
MOSFET	Metal-Oxide-Semiconductor Field-Effect Transistor
MPA	Macro-Pixel ASIC
MPV	Most Probable Value
MSB	Most Significant Bit
MTD	MIP Timing Detector
NIEL	Non-Ionizing Energy Loss
NMOS	N-type channel Metal-Oxide-Semiconductor
OSERDES	Output SERializer/DESerializer
OT	Outer Tracker
PCB	Printed Circuit Board
Ph2 ACF	Phase-2 Acquisition and Control Framework
PMOS	P-type channel Metal-Oxide-Semiconductor
PMT	Photo Multiplier Tube
pp	proton-proton
PS-MCK	Pixel-Strip Mock-Up
PS-p	Pixel-Strip pixel
PS-s	Pixel-Strip strip
PS	Proton Synchrotron or Pixel-Strip
PU	PileUp
PV	Primary Vertex
QCD	Quantum ChromoDynamics
QFT	Quantum Field Theory
QGP	Quark Gluon Plasma
RAM	Random-Access Memory
RECO	RECOnstruction level
RMS	Root Mean Square
RPC	Resistive Plate Chamber
RTL	Register-Transfer Level
SEE	Single-Event Effect
SEH	SErvice Hybrid
SEL	Single-Event Latchup
SET	Single-Event Transient
SEU	Single-Event Upset
SIM	SIMulation level
SiPM	Silicon PhotoMultier
SLVS	Scalable Low-Voltage Signalling
SPS	Super Proton Synchrotron
SRAM	Static Random-Access Memory
SSA	Short Strip ASIC
SW	SoftWare
TB2S	Tracker Barrel with 2S modules
TBPS	Tracker Barrel with PS modules
TDC	Time to Digital Converter
TEC	Tracker EndCap
TEDD	Tracker Endcap Double-Discs

TIB	Tracker Inner Barrel
TID	Total Ionizing Dose or Tracker Inner Disk
TLU	Trigger Logic Unit
TMR	Triple Module Redundancy
TOB	Tracker Outer Barrel
TTC	Trigger, Timing and Control
UCL	Université Catholique de Louvain
UDP	User Datagram Protocol
μDTC	micro Data, Trigger and Control
UE	Underlying Event
μTCA	micro Telecommunications Computing Architecture
VHDCI	Very-High-Density Cable Interconnect
VHDL	Very High Speed Integrated Circuit Hardware Description Language
VLSI	Very Large Scale Integrated
VMEbus	Versa Module Europa or Versa Module Eurocard bus
VTRX+	Versatile TRansceiver plus
WIMP	Weakly Interacting Massive Particle

Acknowledgements

First of all, I would like to thank Steven for giving me the opportunity to perform the research described in this thesis, for his patient guidance and availability to meet up and discuss work, despite a packed agenda. A little over four years ago, you managed to get me excited about the tracker project at the IIHE when we were trying to communicate over a lousy Skype connection whilst being at opposite sides of the globe. This fact is just one of the examples of the enthusiasm you put in your work.

Besides Steven, I would like to thank Prof. P. Vanlaer for the many fruitful discussions and the advice, both on the tracker project and on the Sexaquark analysis. I would also like to thank Prof. G. De Lentdecker for kindly making hardware available any time we asked for something and Prof. J. D'Hondt and Prof. L. Favart for actively supporting the tracker Phase-2 project at the IIHE. Also thanks to the jury members for reading this manuscript in great detail, for the interesting questions, discussions and relevant comments.

During the four years I spent at the IIHE I had two offices, which are referred to below as the *upstairs-greenhouse-sandwich-eating-office* and *downstairs-greenhouse-warm-meal-eating-office*. These offices have two thing in common: steep temperature gradients, but, more importantly, warm people.

The upstairs-greenhouse-sandwich-eating-office:

> Many thanks to Annik, Isabelle, Kevin, Isis, Lieselotte, Karen, Kirill, Doug, Simon, Emil, Petra, Shimaa, Dom, Seth, Dennys, Giannis, Annemie, Audrey, Merel and Quentin whom I shared the floor/office with during the first two years of my thesis. Thanks for all the nice lunch breaks, the absurd humour and the two cherry tomatoes from the indoor garden (it was a successful harvest, indeed). Special thanks also to Marleen for the generous support you provided every day.

The downstairs-greenhouse-warm-meal-eating-office:

> The last two years of my time at the IIHE I spent in the office with Jelena and subsequently Inna, Yannick and Wendi. Thanks for all the candy and chocolate, both got me through quite a few energy glitches. Big thanks as well to Alexandre, Thomas, Barbara, Yifan, Michael, Bugra, David, Hugo, Diego, Qun, Gwen, Louis, Florian, Anastasia, Amandeep, Laurent, Andrey, Olivier, Denis, Stéphane, Romain, Shkelzen, Adriano and all the other members of the computing team.

I am grateful to several colleagues outside of the IIHE as well: Mykyta, Sarah, Davide and Davide, Alessandro, Simone, Mark and Tomasz for sharing many, often long, hours of debugging set-ups or performing tests in some noisy or hot lab at CERN or elsewhere. Furthermore, I am thankful to Prof. G. Farrar and Prof. T. Pierog, respectively for the stimulating discussions on Sexaquarks and for making available a special version of EPOS-LHC to us.

Many thanks to all members of the teaching team which I joined the last four years, especially to Stefan for coordinating the activities and accommodating for my agenda.

Thanks to the Belgian rail company for getting me to work every day! (Calculations show that approximately a quarter of this thesis was written on the train. Whether or not delays had a significant impact on this is still to be investigated.)

Besides work, there was also a bit of personal life in the past four years. I deeply admire Elien for encouraging me when work went stiff, for spending quite some solitary evenings and weekends when I was writing this manuscript and for reminding me, each time I came home, why I took the train back to Ghent every evening. People say that it's only possible to concentrate on work when one is happy ... for sure you had a major impact on that.

Furthermore, thanks to my sister, mother, father and other family members for the encouragement and the food packages. Thanks also to the travelling-friends (for reminding me of the outside world), cycling- and football-friends (for reminding me of those other two limbs called 'legs'), roommates (for reminding me to do the dishes), old study-friends (for reminding me of irrelevant physics paradoxes) and all the other friends from Ghent (for not complaining they did not get a separate category).

Bibliography

[1] A. Purcell. "Go on a particle quest at the first CERN webfest." In: BUL-NA-2012-269. 35/2012 (2012), p. 10. URL: https://cds.cern.ch/record/1473657.

[2] F. Englert and R. Brout. "Broken Symmetry and the Mass of Gauge Vector Mesons". In: *Phys. Rev. Lett.* 13 (9 1964), pp. 321–323. DOI: 10.1103/PhysRevLett.13.321.

[3] P.W. Higgs. "Broken symmetries, massless particles and gauge fields". In: *Physics Letters* 12.2 (1964), pp. 132 –133. ISSN: 0031-9163. DOI: 10.1016/0031-9163(64)91136-9.

[4] M. D. Schwartz. *Quantum Field Theory and the Standard Model*. Cambridge University Press, 2014. ISBN: 1107034736, 9781107034730. URL: http://www.cambridge.org/us/academic/subjects/physics/theoretical-physics-and-mathematical-physics/quantum-field-theory-and-standard-model.

[5] J. Ellis, M. K. Gaillard, and D. V. Nanopoulos. "A Historical Profile of the Higgs Boson. An Updated Historical Profile of the Higgs Boson". In: arXiv:1504.07217 (2015), 22 p. DOI: 10.1142/9789814733519_0014.

[6] Planck Collaboration. "Planck 2018 results. VI. Cosmological parameters". In: *arXiv e-prints* (2018). arXiv: 1807.06209 [astro-ph.CO].

[7] D. Clowe, A. Gonzalez, and M. Markevitch. "Weak-Lensing Mass Reconstruction of the Interacting Cluster 1E 0657-558: Direct Evidence for the Existence of Dark Matter". In: *The Astrophysical Journal* 604.2 (2004), pp. 596–603. DOI: 10.1086/381970.

[8] N. Cibirka et al. "RELICS: Strong Lensing Analysis of the Galaxy Clusters Abell S295, Abell 697, MACS J0025.4-1222, and MACS J0159.8-0849". In: *The Astrophysical Journal* 863.2 (2018), p. 145. DOI: 10.3847/1538-4357/aad2d3.

[9] F. Zwicky. "On the Masses of Nebulae and of Clusters of Nebulae". In: *apj* 86 (1937), p. 217. DOI: 10.1086/143864.

[10] M. Tanabashi et al. "Review of Particle Physics". In: *Phys. Rev. D* 98 (3 2018), p. 030001. DOI: 10.1103/PhysRevD.98.030001.

[11] D. H. Perkins. *Particle Astrophysics*. Oxford University Press, 2009. ISBN: 978–0–19–954546–9.

[12] M. Taoso, G. Bertone, and A. Masiero. "Dark matter candidates: a ten-point test". In: *Journal of Cosmology and Astroparticle Physics* 2008.03 (2008), p. 022. DOI: 10.1088/1475-7516/2008/03/022.

[13] J. I. Read, M. G. Walker, and P. Steger. "Dark matter heats up in dwarf galaxies". In: *Monthly Notices of the Royal Astronomical Society* 484.1 (Jan. 2019), pp. 1401–1420. ISSN: 0035-8711. DOI: 10.1093/mnras/sty3404.

[14] G. F. Knoll. *Radiation Detection and Measurement*. John Wiley and Sons, Inc, 1999. ISBN: 978-0-273-71918-2.

[15] F. Hartmann. *Evolution of Silicon Sensor Technology in Particle Physics*. Springer, 2009. ISBN: 978-3-540-25094-4.

[16] M. Huhtinen and F. Faccio. "Computational method to estimate Single Event Upset rates in an accelerator environment". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 450.1 (2000), pp. 155 –172. ISSN: 0168-9002. DOI: 10.1016/S0168-9002(00)00155-8.

[17] *pn-junction*. https://www.electronics-tutorials.ws/diode/diode_2.html.

[18] D. Schell. "Development of a Macro-Pixel sensor for the Phase-2 Upgrade of the CMS experiment". PhD thesis. Karlsruher Institut für Technologie (KIT), 2019. 196 pp. DOI: 10.5445/IR/1000094129.

[19] A. Vasilescu et al. *Notes on the fluence normalisation based on the NIEL scaling hypothesis*. Tech. rep. 2000. URL: https://rd48.web.cern.ch/RD48/technical-notes/rosetn00-02.ps.

[20] N. Storey. *Electronics, a systems approach*. Fourth edition. PEARSON Education, 2009. ISBN: 978-0-471-07338-3.

[21] A. Caratelli. "Research and development of an intelligent particle tracker detector electronic system". PhD thesis. Lausanne: IEL, 2019, p. 210. DOI: 10.5075/epfl-thesis-9702.

[22] *Layer structure of an ASIC*. https://upload.wikimedia.org/wikipedia/commons/e/ee/Cmos-chip_structure_in_2000s_%28en%29.svg.

[23] J. R. Srour, C. J. Marshall, and P. W. Marshall. "Review of displacement damage effects in silicon devices". In: *IEEE Transactions on Nuclear Science* 50.3 (2003), pp. 653–670. DOI: 10.1109/TNS.2003.813197.

[24] J. R. Schwank et al. "Radiation Effects in MOS Oxides". In: *IEEE Transactions on Nuclear Science* 55.4 (Aug. 2008), pp. 1833 –1853. ISSN: 1558-1578. DOI: 10.1109/TNS.2008.2001040.

[25] M. Shin and M. Kim. "An evaluation of radiation damage to solid state components flown in low earth orbit satellites". In: *Radiation Protection Dosimetry* 108.4 (Apr. 2004), pp. 279–291. ISSN: 0144-8420. DOI: 10.1093/rpd/nch024.

[26] G. Borghello. "Ionizing radiation effects in nanoscale CMOS technologies exposed to ultra-high doses". Presented 08 Mar 2019. 2018. URL: http://cds.cern.ch/record/2680840.

[27] F. Faccio et al. "Influence of LDD Spacers and H^+ Transport on the Total-Ionizing-Dose Response of 65-nm MOSFETs Irradiated to Ultrahigh Doses". In: *IEEE Trans. Nucl. Sci.* 65.1 (2018), pp. 164–174. DOI: 10.1109/TNS.2017.2760629.

[28] V. Bezhenova and A. Michalowska-Forsyth. "Aspect ratio of radiation-hardened MOS transistors". In: *e & i Elektrotechnik und Informationstechnik* 135.1 (2018), pp. 61–68. ISSN: 1613-7620. DOI: 10.1007/s00502-017-0575-2.

[29] C. M. Hsieh, P. C. Murley, and R. R. O'Brien. "A field-funneling effect on the collection of alpha-particle-generated carriers in silicon devices". In: *IEEE Electron Device Letters* 2.4 (1981), pp. 103–105. ISSN: 0741-3106. DOI: 10.1109/EDL.1981.25357.

[30] Altera Corporation. *Introduction to Single-Event Upsets*. Sept. 2013.

[31] J. Hussein and G. Swift. *White Paper: 7 Series FPGAs, Mitigating Single-Event Upsets*. XILINX WP395 v1.1. May 2015.

[32] S. Bonacini et al. "An SEU-Robust Configurable Logic Block for the Implementation of a Radiation-Tolerant FPGA". In: *IEEE Transactions on Nuclear Science* 53.6 (2006), pp. 3408–3416. DOI: 10.1109/TNS.2006.884097.

[33] CMS Collaboration. *The Phase-2 Upgrade of the CMS Tracker*. Tech. rep. CERN-LHCC-2017-009. CMS-TDR-014. Geneva: CERN, 2017. URL: https://cds.cern.ch/record/2272264.

[34] E. Mobs. *The CERN accelerator complex*. URL: https://cds.cern.ch/record/2197559.

[35] O. S. Brüning et al. *LHC Design Report*. CERN Yellow Reports: Monographs. Geneva: CERN, 2004. DOI: 10.5170/CERN-2004-003-V-1.

[36] ATLAS Collaboration. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: *Physics Letters B* 716.1 (2012), pp. 1 –29. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2012.08.020.

[37] CMS Collaboration. "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC". In: *Physics Letters B* 716.1 (2012), pp. 30 –61. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2012.08.021.

[38] Apollinari G. et al. *High-Luminosity Large Hadron Collider (HL-LHC): Technical Design Report V. 0.1*. CERN Yellow Reports: Monographs. Geneva: CERN, 2017. DOI: 10.23731/CYRM-2017-004.

[39] *CMS Luminosity Public Results*. URL: https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults.

[40] T. Sakuma. *Cutaway diagrams of CMS detector*. https://cds.cern.ch/record/2665537. 2019.

[41] *CMS coordinate system*. https://wiki.physik.uzh.ch/cms/_detail/latex:cms_coordinate_system.png?id=latex%3Aexample_spherical_coordinates.

[42] CMS Collaboration. *The Phase-2 Upgrade of the CMS Muon Detectors*. Tech. rep. CERN-LHCC-2017-012. CMS-TDR-016. Geneva: CERN, 2017. URL: https://cds.cern.ch/record/2283189.

[43] L. Viliani. "CMS tracker performance and readiness for LHC Run II". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 824 (2016). Frontier Detectors for Frontier Physics: Proceedings of the 13th Pisa Meeting on Advanced Detectors, pp. 67 –69. ISSN: 0168-9002. DOI: 10.1016/j.nima.2015.09.046.

[44] S. Spannagel. "Status of the CMS Phase I Pixel Detector Upgrade". In: *Nucl. Instrum. Meth.* A831 (2016), pp. 71–75. DOI: 10.1016/j.nima.2016.03.028. arXiv: 1511.06084 [physics.ins-det].

[45] M.J. French et al. "Design and results from the APV25, a deep sub-micron CMOS front-end chip for the CMS tracker". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 466.2 (2001). 4th Int. Symp. on Development and Application of Semiconductor Tracking Detectors, pp. 359 –365. ISSN: 0168-9002. DOI: 10.1016/S0168-9002(01)00589-7.

[46] G. H. Dirkes. "Construction of the CMS silicon strip tracker". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 581.1 (2007). VCI 2007, pp. 299 –302. ISSN: 0168-9002. DOI: 10.1016/j.nima.2007.07.126.

[47] CMS Collaboration. "Description and performance of track and primary-vertex reconstruction with the CMS tracker". In: *JINST* 9.CMS-TRK-11-001. CERN-PH-EP-2014-070. CMS-TRK-11-001 (2014), P10009. DOI: 10.1088/1748-0221/9/10/P10009. arXiv: 1405.6569.

[48] R. E. Kalman and R. S. Bucy. "New Results in Linear Filtering and Prediction Theory". In: *Journal of Fluids Engineering* 83.1 (Mar. 1961), pp. 95–108. ISSN: 0098-2202. DOI: 10.1115/1.3658902. eprint: https://asmedigitalcollection.asme.org/fluidsengineering/article-pdf/83/1/95/4655100/95_1.pdf.

[49] P. Billoir and S. Qian. "Simultaneous pattern recognition and track fitting by the Kalman filtering method". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 294.1 (1990), pp. 219 –228. ISSN: 0168-9002. DOI: 10.1016/0168-9002(90)91835-Y.

[50] R. Frühwirth. "Application of Kalman filtering to track and vertex fitting". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 262.2 (1987), pp. 444 –450. ISSN: 0168-9002. DOI: 10.1016/0168-9002(87)90887-4.

[51] *CMS Tracking Performance Public Results*. URL: https://twiki.cern.ch/twiki/bin/view/CMSPublic/TrackingPOGPerformance2017MC.

[52] CMS Collaboration. "Performance and operation of the CMS electromagnetic calorimeter". In: *JINST* 5.03 (2010), T03010–T03010. DOI: 10.1088/1748-0221/5/03/t03010. arXiv: 0910.3423.

[53] CMS Collaboration. "Performance of the CMS hadron calorimeter with cosmic ray muons and LHC beam data". In: *JINST* 5.03 (2010), T03012–T03012. DOI: 10.1088/1748-0221/5/03/t03012. arXiv: 0911.4991.

[54] S. Chatrchyan et al. "The Performance of the CMS Muon Detector in Proton-Proton Collisions at $\sqrt{s}$ = 7 TeV at the LHC". In: *JINST* 8 (2013), P11002. DOI: 10.1088/1748-0221/8/11/P11002. arXiv: 1306.6905 [physics.ins-det].

[55] CMS Collaboration. "The CMS trigger system". In: *JINST* 12.1 (2017), P01020. DOI: 10.1088/1748-0221/12/01/P01020. arXiv: 1609.02366 [physics.ins-det].

[56] *Report on the Physics at the HL-LHC and Perspectives for the HE-LHC*. Tech. rep. arXiv:1902.10229. Geneva: CERN, 2019. URL: https://cds.cern.ch/record/2651134.

[57] CMS Collaboration. *The Phase-2 Upgrade of the CMS Barrel Calorimeters*. Tech. rep. CERN-LHCC-2017-011. CMS-TDR-015. Geneva: CERN, 2017. URL: https://cds.cern.ch/record/2283187.

[58] CMS Collaboration. *The Phase-2 Upgrade of the CMS Endcap Calorimeter*. Tech. rep. CERN-LHCC-2017-023. CMS-TDR-019. Technical Design Report of the endcap calorimeter for the Phase-2 upgrade of the CMS experiment, in view of the HL-LHC run. Geneva: CERN, 2017. URL: https://cds.cern.ch/record/2293646.

[59] M.I. Pedraza-Morales et al. *RPC upgrade project for CMS Phase II*. Tech. rep. arXiv:1806.11503. 2018. URL: https://cds.cern.ch/record/2628277.

[60] CMS Collaboration. *Technical Proposal for a MIP Timing Detector in the CMS Experiment Phase 2 Upgrade*. Tech. rep. CERN-LHCC-2017-027. LHCC-P-009. Geneva: CERN, 2017. URL: https://cds.cern.ch/record/2296612.

[61] J. G. Hegeman. *The CMS Data Acquisition System for the Phase-2 Upgrade*. Tech. rep. CMS-CR-2018-099. Geneva: CERN, 2018. URL: http://cds.cern.ch/record/2629375.

[62] CMS Collaboration. *The CMS tracker system project: Technical Design Report*. Technical Design Report CMS. Geneva: CERN, 1997. URL: http://cds.cern.ch/record/368412.

[63] D. Contardo et al. *Technical Proposal for the Phase-II Upgrade of the CMS Detector*. Tech. rep. CERN-LHCC-2015-010. LHCC-P-008. CMS-TDR-15-02. Geneva, 2015. URL: https://cds.cern.ch/record/2020886.

[64] CMS Collaboratio. *The Phase-2 Upgrade of the CMS L1 Trigger Interim Technical Design Report*. Tech. rep. CERN-LHCC-2017-013. CMS-TDR-017. Geneva: CERN, 2017. URL: https://cds.cern.ch/record/2283192.

[65] A. Dierlamm. *Characterisation of silicon sensor materials and designs for the CMS Tracker Upgrade*. Tech. rep. AIDA-CONF-2014-006. Geneva: CERN, 2012. URL: https://cds.cern.ch/record/1693190.

[66] D. Ceresa et al. *Macro Pixel ASIC User Manual v0.1*. URL: https://espace.cern.ch/CMS-MPA/SiteAssets/SitePages/Documents/MPA_UserManual_V0.pdf.

[67] A. Caratelli et al. *Short Strip ASIC Specification Document v1*. URL: https://espace.cern.ch/CMS-MPA/SiteAssets/SitePages/Documents/SSA_Manual_V1.0.pdf.

[68] M. Prydderch. *CBC3.1 User Manual*. URL: http://www.hep.ph.ic.ac.uk/ASIC/cbc3.1/CBC3p1_User_Manual_V1p3.pdf.

[69] L. Caponetto et al. *CIC1 technical specification*. URL: https://espace.cern.ch/Tracker-Upgrade/Electronics/CIC/Shared%20Documents/Specifications/CIC_specs_v3.pdf.

[70] *I2C bus specifications for the CMS tracker 2S and PS modules, v2.0, 21/4/2016*. CMS internal document.

[71] P. Moreira. *Inside the LpGBT*. Ecole IN2P3 de Microélectronique. May 2017.

[72] J. Troska et al. *The VTRx+, an Optical Link Module for Data Transmission at HL-LHC*. TWEPP proceeding. 2018. DOI: 10.22323/1.313.0048.

[73] *lpGBT Manual*. URL: https://lpgbt.web.cern.ch/lpgbt/manual/.

[74] Y. Gershtein. "CMS Hardware Track Trigger: New Opportunities for Long-Lived Particle Searches at the HL-LHC". In: *Phys. Rev.* D96.3 (2017), p. 035027. DOI: 10.1103/PhysRevD.96.035027. arXiv: 1705.04321 [hep-ph].

[75] M. Pesaresi et al. "The FC7 AMC for generic DAQ & control applications in CMS". In: *JINST* 10.3 (2015), pp. C03036–C03036. DOI: 10.1088/1748-0221/10/03/C03036.

[76] *XILINX 7 Series FPGAs Configurable Logic Block, UG474*. URL: https://www.xilinx.com/support/documentation/user_guides/ug474_7Series_CLB.pdf.

[77] *MicroTCA Application Guide, A Modular Open Systems Approach for Computer Architecture*. Version 1.1, August 2012. URL: https://www.picmg.org/wp-content/uploads/MicroTCA_Application_Guide_final-_rev1_1.pdf.

[78] C. G. Larrea et al. "IPbus: a flexible Ethernet-based control system for xTCA hardware". In: *JINST* 10 (2015), p. C02019. DOI: 10.1088%2F1748-0221%2F10%2F02%2Fc02019.

[79] *XILINX 7 Series FPGAs Data Sheet: Overview*. URL: https://www.xilinx.com/support/documentation/data_sheets/ds180_7Series_Overview.pdf.

[80] CMS Tracker Collaboration. *Performance of the Prototype CBC3-based Outer Tracker Modules for the Phase II Upgrade of CMS before and after neutron irradiation*. in preparation.

[81] *Update of System Test Activities at KIT*. URL: https://indico.cern.ch/event/818273/contributions/3422488/attachments/1843406/3023409/190513_System_Test_Meeting_-_XTalk.pdf.

[82] A. Ferrari et al. *FLUKA: A multi-particle transport code (Program version 2005)*. 2005. DOI: 10.2172/877507.

[83] S. Mallows et al. "Monte Carlo simulations of the radiation environment for the CMS experiment". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 824 (2016). Frontier Detectors for Frontier Physics: Proceedings of the 13th Pisa Meeting on Advanced Detectors, pp. 30 –32. ISSN: 0168-9002. DOI: 10.1016/j.nima.2015.11.044.

[84] G Bianchi et al. "tkLayout: a design tool for innovative silicon tracking detectors". In: *JINST* 9.03 (2014), pp. C03054–C03054. DOI: 10.1088/1748-0221/9/03/c03054.

[85] D. Cussans. *Description of the JRA1 Trigger Logic Unit (TLU), v0.2c. EUDET-Memo-2009-4*.

[86] H. Perrey. "An EUDET / AIDA pixel beam telescope for detector development ". In: *IEEE Nuclear Science Symposium and Medical Imaging Conference Record (NSS/MIC)* (Oct. 2012). DOI: 10.1109/NSSMIC.2012.6551461.

[87] *Bitonic sorting*. https://commons.wikimedia.org/wiki/File:BitonicSort1.svg.

[88] G. R. Farrar. "Stable Sexaquark". In: *arXiv e-prints* (2017). arXiv: 1708.08951 [hep-ph].

[89] R. L. Jaffe. "Perhaps a Stable Dihyperon". In: *Phys. Rev. Lett.* 38 (1977). [Erratum: Phys. Rev. Lett.38,617(1977)], pp. 195–198. DOI: 10.1103/PhysRevLett.38.617, 10.1103/PhysRevLett.38.195.

[90] Belle Collaboration. "Observation of a Narrow Charmoniumlike State in Exclusive $B^{\pm} \to K^{\pm}\pi^{+}\pi^{-}J/\psi$ Decays". In: *Phys. Rev. Lett.* 91 (26 2003), p. 262001. DOI: 10.1103/PhysRevLett.91.262001.

[91] LHCb Collaboration. "Observation of $J/\psi p$ Resonances Consistent with Pentaquark States in $\Lambda_b^0 \to J/\psi K^- p$ Decays". In: *Phys. Rev. Lett.* 115 (2015), p. 072001. DOI: 10.1103/PhysRevLett.115.072001. arXiv: 1507.03414 [hep-ex].

[92] R. E. Chrien. "H particle searches at Brookhaven". In: *Nuclear Physics A* 629.1 (1998). Quark Lepton Nuclear Physics, pp. 388 –397. ISSN: 0375-9474. DOI: 10.1016/S0375-9474(97)00714-8.

[93] Belle Collaboration. "Search for an H-Dibaryon with a Mass near $2m_\Lambda$ in $\Upsilon(1S)$ and $\Upsilon(2S)$ Decays". In: *Phys. Rev. Lett.* 110 (22 2013), p. 222002. DOI: 10.1103/PhysRevLett.110.222002.

[94] NA 3 Collaboration. "Mass and lifetime limits on new long-lived particles in 300 GeV/c π- interactions". In: *Zeitschrift für Physik C Particles and Fields* 31.1 (1986), pp. 21–32. ISSN: 1431-5858. DOI: 10.1007/BF01559588.

[95] R. H. Bernstein et al. "Search for new, long-lived, neutral particles". In: *Phys. Rev. D* 37 (11 1988), pp. 3103–3106. DOI: 10.1103/PhysRevD.37.3103.

[96] ALICE Collaboration. "p-p, p-Λ and Λ-Λ correlations studied via femtoscopy in pp reactions at $\sqrt{s}$ = 7 TeV". In: *Phys. Rev.* C99.2 (2019), p. 024001. DOI: 10.1103/PhysRevC.99.024001. arXiv: 1805.12455 [nucl-ex].

[97] H. Takahashi et al. "Observation of a ${}^{6}_{\Lambda\Lambda}He$ Double Hypernucleus". In: *Phys. Rev. Lett.* 87 (21 2001), p. 212502. DOI: 10.1103/PhysRevLett.87.212502.

[98] J. K. Ahn et al. "Production of ${}^{4}_{\Lambda\Lambda}H$ Hypernuclei". In: *Phys. Rev. Lett.* 87 (13 2001), p. 132504. DOI: 10.1103/PhysRevLett.87.132504.

[99] G. R. Farrar and G. Zaharijas. "Non-binding of flavor-singlet hadrons to nuclei". In: *Physics Letters, Section B* 559.3-4 (May 2003), pp. 223–228. ISSN: 0370-2693. DOI: 10.1016/S0370-2693(03)00331-9.

[100] E. W. Kolb and M. S. Turner. "Dibaryons cannot be the dark matter". In: *Phys. Rev. D* 99."arXiv:1809.06003" (6 2019), p. 063519. DOI: 10.1103/PhysRevD.99.063519.

[101] S. R. Beane et al. "Light nuclei and hypernuclei from quantum chromodynamics in the limit of SU(3) flavor symmetry". In: *Phys. Rev. D* 87 (3 2013), p. 034506. DOI: 10.1103/PhysRevD.87.034506.

[102] S. A. Yost and C. R. Nappi. "Mass of the H dibaryon in a chiral model". In: *Phys. Rev. D* 32 (3 1985), pp. 816–818. DOI: 10.1103/PhysRevD.32.816.

[103] K.F. Liu and C.W. Wong. "MIT bag model with center-of-mass correction". In: *Physics Letters B* 113.1 (1982), pp. 1 –5. ISSN: 0370-2693. DOI: 10.1016/0370-2693(82)90096-X.

[104] G. R. Farrar and Gabrijela Zaharijas. "Nuclear and nucleon transitions of the H dibaryon". In: *Phys. Rev. D* 70 (1 2004), p. 014008. DOI: 10.1103/PhysRevD.70.014008.

[105] G. R. Farrar. "A precision test of the nature of Dark Matter and a probe of the QCD phase transition". In: *arXiv e-prints* (2018). arXiv: 1805.03723 [hep-ph].

[106] G. R. Farrar. "6-quark Dark Matter". In: *arXiv e-prints* (2017). arXiv: 1711.10971 [hep-ph].

[107] B. D. Wandelt et al. "Self-Interacting Dark Matter". In: *Sources and Detection of Dark Matter and Dark Energy in the Universe*. Ed. by David B. Cline. Berlin, Heidelberg: Springer Berlin Heidelberg, 2001, pp. 263–274. ISBN: 978-3-662-04587-9.

[108] BaBar Collaboration. "Search for a Stable Six-Quark State at BABAR". In: *Phys. Rev. Lett.* 122 (7 2019), p. 072002. DOI: 10.1103/PhysRevLett.122.072002.

[109] *V0Producer Offline Guide*. URL: https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideV0Producer.

[110] T. Pierog et al. "EPOS LHC: Test of collective hadronization with data measured at the CERN Large Hadron Collider". In: *Phys. Rev.* C92.3 (2015), p. 034906. DOI: 10.1103/PhysRevC.92.034906. arXiv: 1306.0121 [hep-ph].

[111] J. Allison et al. "Recent developments in Geant4". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 835 (2016), pp. 186 –225. ISSN: 0168-9002. DOI: 10.1016/j.nima.2016.06.125.

[112] T. Speer et al. *Vertex Fitting in the CMS Tracker*. Tech. rep. CMS-NOTE-2006-032. Geneva: CERN, 2006. URL: https://cds.cern.ch/record/927395.

[113] P. Avery. "Applied Fitting Theory VI Formulas for Kinematic Fitting". CBX 98–37, Apr. 17, 1999 (rev.) URL: http://www.phys.ufl.edu/~avery/fitting/kinematic.pdf.

[114] A. Dominguez et al. *CMS Technical Design Report for the Pixel Detector Upgrade*. Tech. rep. CERN-LHCC-2012-016. CMS-TDR-11. 2012. URL: http://cds.cern.ch/record/1481838.

[115] CMS Collaboration. "Precision measurement of the structure of the CMS inner tracking system using nuclear interactions". In: *JINST* 13.arXiv:1807.03289. CMS-TRK-17-001-003 (2018). Submitted to JINST, P10034. 41 p. DOI: 10.1088/1748-0221/13/10/P10034. URL: https://cds.cern.ch/record/2629890.

[116] C. Colle. "Hunting for nuclear short-range correlations : the connection to electroinduced two-nucleon knockout reactions". eng. PhD thesis. Ghent University, 2017, pp. II, 122.

[117] J. Ryckebusch et al. "The isospin and neutron-to-proton excess dependence of short-range correlations". In: *Phys. Lett.* B792 (2019), pp. 21–28. DOI: 10.1016/j.physletb.2019.03.016. arXiv: 1808.09859 [nucl-th].

[118] A. Hoecker et al. "TMVA: Toolkit for Multivariate Data Analysis". In: *PoS* ACAT (2007), p. 040. arXiv: physics/0703039.

[119] CMS Collaboration. "Measurement of Neutral Strange Particle Production in the Underlying Event in Proton-Proton Collisions at $\sqrt{s}$ = 7 TeV". In: *Phys. Rev.* D88 (2013), p. 052001. DOI: 10.1103/PhysRevD.88.052001. arXiv: 1305.6016 [hep-ex].

[120] CMS Collaboration. "Strange Particle Production in pp Collisions at $\sqrt{s} = 0.9$ and 7 TeV". In: *JHEP* 05 (2011), p. 064. DOI: 10.1007/JHEP05(2011)064. arXiv: 1102.4282 [hep-ex].

[121] CMS Collaboration. "Event generator tunes obtained from underlying event and multiparton scattering measurements". In: *Eur. Phys. J. C* 76.CMS-GEN-14-001. CMS-GEN-14-001-003. CERN-PH-EP-2015-291 (2015), 155. 68 p. URL: http://cds.cern.ch/record/2110213.

[122] G. Cowan et al. "Asymptotic formulae for likelihood-based tests of new physics". In: *The European Physical Journal C* 71.2 (2011). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-011-1554-0.

[123] A. L. Read. "Presentation of search results: The CL(s) technique". In: *J. Phys.* G28 (2002). [,11(2002)], pp. 2693–2704. DOI: 10.1088/0954-3899/28/10/313.

[124] *Procedure for the LHC Higgs boson search combination in Summer 2011*. Tech. rep. CMS-NOTE-2011-005. ATL-PHYS-PUB-2011-11. Geneva: CERN, 2011. URL: https://cds.cern.ch/record/1379837.

[125] *Combine: documentation of the RooStats based statistics tools for Higgs PAG*. http://cms-analysis.github.io/HiggsAnalysis-CombinedLimit/.

[126] CMS Collaboration. "Pseudorapidity distribution of charged hadrons in proton-proton collisions at $\sqrt{s}$ = 13 TeV". In: *Phys. Lett.* B751 (2015), pp. 143–163. DOI: 10.1016/j.physletb.2015.10.004. arXiv: 1507.05915 [hep-ex].

[127] G. Landsberg. private communication. Dec. 4, 2019.

[128] *Studies of Tracker Material*. Tech. rep. CMS-PAS-TRK-10-003. 2010. URL: http://cds.cern.ch/record/1279138.