AGH UNIVERSITY OF SCIENCE AND TECHNOLOGY

FIELD OF SCIENCE: Natural sciences

SCIENTIFIC DISCIPLINE: Physical sciences

DOCTORAL THESIS

Impact of models for multiparton interactions in proton-proton collisions at LHC energies for the prediction of radiation damage of LHCb silicon trackers

Author: Saliha Bashir

Supervisor: dr hab. inż. Agnieszka Obłąkowska-Mucha - prof. AGH

Completed in: AGH University of Krakow,
Department of Particle Interactions and Detection Techniques, Kraków, 2025

Declaration of the author of this dissertation

Aware of legal responsibility for making untrue statements I hereby declare that I have written this dissertation myself and all the contents of the dissertation have been obtained by legal means.

---

signature date

Declaration of the thesis Supervisors

This dissertation is ready to be reviewed.

---

signature of thesis Supervisor date

Contents

Acknowledgments

I extend my heartfelt appreciation to my supervisor, Agnieszka Obłąkowska-Mucha, whose invaluable guidance, patience, motivation, support, encouragement, and insightful knowledge of the subject matter steered me through this research. Her support went far beyond the professional, she didn't just guide a thesis, she helped change a life.

I would also like to express my sincere thanks to Professor Tomasz Szumlak for his invaluable guidance and support, especially during the most demanding and complex parts of my thesis. His willingness to help, no matter how complex the task, gave me both confidence and clarity when I needed it most. Beyond that, his constant motivation and encouragement were truly uplifting. As the head of the department, he not only fulfilled his responsibilities with dedication and excellence but also created an environment where students felt genuinely supported and inspired.

My sincere gratitude also goes to NAWA (Polish National Agency for Academic Exchange) for their generous funding and unwavering support. Their contribution not only made this academic endeavor possible but also allowed me to dedicate myself to my research without financial concerns.

A special thanks to Gloria Corti, my supervisor during my internship at CERN, for her mentorship and unwavering support. She encouraged me to present my work during several LHCb weeks, an experience that significantly boosted my confidence. Our frequent meetings with the simulation group, where she supported and guided me in every possible way, were especially enriching and deeply appreciated. I am truly grateful for the opportunity to learn from her.

To my partner, Fasih, your endless patience and understanding, especially during the tough phases of this research, have been my anchor. You stood by me through every moment of stress, every meltdown, every tear and tantrum, with unwavering love and calm. Your ability to lift me up when I felt like giving up, to bring light into the darkest days, and to simply be there — steady, kind, and present — means more to me than

words can ever express. This journey would have been impossible without you.

I am deeply grateful to my colleagues for their insightful discussions and constant encouragement throughout this process. A special thanks to Paweł, whose generosity in offering his time and expertise at CERN greatly contributed to my thesis. I am also incredibly thankful to Wojciech for his invaluable help in improving my programming skills and to Karol, who assisted me in setting up the necessary environments for my programs to run smoothly. These individuals truly are a gem for our department. Additionally, I'd like to express my gratitude to my colleague Sabin, whose ongoing support and engaging discussions were a constant source of inspiration. I also owe a debt of gratitude to my teachers, who have taught me countless lessons, both directly related to my thesis and beyond. Finally, while my students have sometimes been a challenge, their presence has ultimately shaped and enriched my journey in ways I couldn't have anticipated.

Above all, I am profoundly grateful to my family, whose unconditional love, constant prayers, and unwavering support have been my greatest source of inspiration. Through moments of doubt, stress, and exhaustion, they have always been there to lift me up, offering warmth, comfort, and a reminder of what truly matters. Their belief in me has never wavered, and it is because of them that I have been able to push forward, even when the path seemed uncertain. Thank you for always standing by me and for showing me the true meaning of strength, love, and resilience.

I would like to express my deepest thanks to everyone who has contributed to this thesis, whether directly or indirectly. This research would not have been possible without the support, guidance, and encouragement. The completion of this thesis marks not just the end of an academic chapter but the beginning of new challenges and opportunities. I look forward to carrying the lessons learned and the experiences shared into the future, with heartfelt gratitude for the role each of you has played in this journey.

This work was supported by the National Research Centre, Poland (NCN), grants UMO-2019/35/O/ST2/00546

Abstract

Large Hadron Collider (LHC) is the world's largest and highest-energy particle collider. LHCb is one of the four major experiments at the Large Hadron Collider (CERN). The LHCb experiment has been designed to study CP violation in beauty and charm hadrons, discovering exotic states, with a broad program covering also QCD, searches for long-lived particles, and dark matter.

This thesis concerns a few topics revolving around a common issue related to radiation damage in silicon sensors in one of the LHCb subdetectors, the Vertex Locator (VELO). The analysis is performed with the use of various samples: particles generated in proton-proton collisions at LHC energies with a stand-alone particle generator, events obtained in the simulation within the LHCb detector, a real minimum-bias sample collected in 2018 by LHCb, and finally, special detector data registered in the LHCb VELO in 2024.

The current state-of-the-art method of monitoring and predicting radiation damage is based on simulation. There are different physics models used in generators, and the simulated particle fluence varies between generators, especially when considering low-momentum particles emitted at high rapidity. Therefore, this project's first part deals with comparing general-purpose event generators, Pythia and Herwig, with special attention to parameters describing multiple parton interactions. A comparison of both event generators with respect to LHCb data is also mentioned.

Since the increase in LHC energy, more discrepancies between data and simulation have been observed among experiments, which gave rise to the necessity of an update in the generators' parameters. Therefore, the second task of this thesis focuses on tuning the campaign of the event generators, detailing the methodology and the software tools developed for tuning, such as Rivet and Professor. These are employed for estimating the optimal parameter values of the event generators predominantly used at LHCb. The tuning results indicate the necessity for further refinements to enhance agreement with experimental observations at 13 TeV.

The third section in this project takes into account the minimum-bias LHCb experimental data from Run 2 (Run 3 data were not available during this time). The main purpose of this analysis was to check whether one can use reconstructed and identified charged hadrons to obtain a reliable prediction for the current level of radiation damage in the VELO sensors. This analysis uses data from proton-proton collisions at a centre of mass energy of 13 TeV taken during runs of the LHC in 2018, highlighting discrepancies between data and MC predictions. This study explores the sources of these differences by analysing particle multiplicities and identifying regions where MC models require improvements to better describe the data.

The last stage of the work describes an attempt to determine the particle fluence map in the new pixel detector VELO based on signals from individual sensors. Special samples of detector data collected in the summer of 2024, in the first months when all subdetectors of the modernised LHCb experiment were included for the data taking.

Streszczenie

Wielki Zderzacz Hadronów (LHC) to największy na świecie akcelerator cząstek przyspieszający protony do najwyższych energii. LHCb jest jednym z czterech głównych eksperymentów na LHC, zaprojektowanym do badania parzystości CP w sektorze hadronów pięknych i powabnych, poszukiwań stanów egzotycznych, z szerokim programem obejmującym również QCD, poszukiwanie długożyciowych cząstek i ciemnej materii.

Przedłożona praca dotyczy kilku zagadnień związanych z przewidywaniami i monitoringiem uszkodzeń radiacyjnych w sensorach krzemowych jednego z poddetektorów LHCb, detektora wierzchołka Vertex Locator (VELO). Analiza w każdym temacie wykonywana jest z wykorzystaniem dedykowanych zbiorów danych: cząstek wyprodukowanych przez generatory ogólnego przeznaczenia, przypadków symulowanych i porównaniu ich z danymi rzeczywistymi zebranymi w roku 2018 (Run 2) przez eksperyment LHCb, oraz specjalnych danych detektorowych zarejestrowanych w LHCb VELO w roku 2024 roku (Run 3). Celem przedłożonej analizy jest sprawdzenie, jak można wykorzystać dane w czasie trwania eksperymentu do przewidzenia poziomu zniszczeń radiacyjnych w krzemowym detektorze VELO.

Obecna jedyną metodą służącą do wyznaczenia fluencji cząstek jest symulacja. Wymaga ona stworzenia modelu detektora, co zwykle jest czasochłonne. W dodatku generatory przypadków używane w symulacjach stosują modele fizyczne, a symulowana fluencja różni się w zależności od generatora, zwłaszcza jeśli weźmie się pod uwagę cząstki o niskim pędzie, czy emitowane pod niewielkim kątem w stosunku do osi wiązki. Dlatego też pierwsza część tego projektu dotyczy porównania generatorów zdarzeń ogólnego przeznaczenia, Pythia i Herwig, ze szczególnym uwzględnieniem parametrów opisujących oddziaływania wielopartonowe. Porównanie danych symulowanych z danymi doświadczalnymi, wykonane z użyciem pakietów Rivet oraz Professor, pozwala na wyznaczenie optymalnych wartości parametrów generatorów przypadków (tzw. tuning).

W drugiej części pracy analizowane są dane symulowane i rzeczywiste przypadki

zebrane w roku 2018 przez LHCb w celu wyznaczenia liczby naładowanych hadronów przechodzących przez sensory VELO, z wykorzystaniem standardowych metod rekonstrukcji śladów i identyfikacji cząstek eksperymentu.

Ostatni etap pracy opisuje unikalną próbę wyznaczenia mapy fluencji cząstek w nowym, pikselowym detektorze VELO w oparciu o sygnały z poszczególnych sensorów. Analizie poddano specjalne próbki danych detektorowych zebranych w lecie 2024 roku, w pierwszych miesiącach, kiedy do totu pomiarowego włączone zostały wszystkie podddetektory zmodernizowanego eksperymentu LHCb.

Author's contribution

The author was a researcher of the following project funded by the National Science Centre (NCN) in Poland:

- *Impact of models for multiparton interactions in proton-proton collisions at LHC energies for the prediction of radiation damage of LHCb silicon trackers.*, project number: UMO-2019/35/O/ST2/00546, project type: NCN grant, the amount awarded: 485 040 PLN, project start date: 2020-10-01, project end date: 2025-09-30

The author was an internee funded by the National Science Centre (NCN) at CERN:

- Internship for a period of 6 months at CERN from 01.08.2022–31.01.2023 under the supervision of Gloria Corti (Head of Simulation group at the LHCb experiment.)

The contribution of the author to this thesis:

- As part of the analysis comparing Pythia and Herwig I contributed to the installation and setup of both generators. I was actively involved in configuring and running simulations under a wide range of settings. This required running the generators repeatedly over several days to explore different configurations. I also developed custom software to systematically compare the output from Herwig and Pythia.

- Additionally, during the time period of my internship at CERN, I actively participated in the tuning campaign. My role involved identifying and reporting bugs in the existing tuning software to the relevant experts. I also evaluated whether the current Pythia tuning settings were optimal for the 7 TeV data used by the LHCb experiment. I collaborated closely with experts and presented my findings in multiple meetings and group discussions to gather feedback and improve the

workflow. I was also tasked with verifying the compatibility of different versions of the RIVET software used in the tuning process, ensuring that the versions in use were up-to-date and compatible with our analysis needs, and I communicated any inconsistencies with the RIVET experts. I then ran the software with different settings of PYTHIA and the tuned parameters with the help of PROFESSOR and RIVET, I had various results presented at different LHCb weeks in parallel and plenary sessions.

- For the Run 2 analysis, I was involved in selecting the most suitable datasets for both data and Monte Carlo (MC) simulations. This selection process was carried out in consultation with my supervisor through a series of meetings. For the MC samples, I submitted jobs using Ganga—a job management tool widely used at LHCb for submitting, monitoring, and managing data processing tasks across various computing platforms. Once the jobs were completed, I retrieved the output files for further analysis. Handling the real data was really challenging due to the large file sizes and complexity. To manage this, I initially downloaded a subset of the data and ran it with my analysis settings to extract the events of interest. After finalizing the selection of both data and MC samples, I developed a comparison tool to analyze and contrast the characteristics of the datasets.

- In the final part of my thesis, I visited CERN to collaborate with the VELO experts. With their help, I obtained a suitable option file to extract VELO hits needed for radiation damage studies using real data. I replicated a similar process for Monte Carlo (MC) simulations, where I adapted a sample option file to extract VELO hits from MC. I then created a software tool that enabled direct comparison between data and MC results.

Talks given by the author and posters presented at the international and national conferences:

- S.Bashir (on behalf of LHCb), *Double-parton Scattering at LHCb and Pythia Tunings*, 28th Cracow EPIPHANY Conference on Recent Advances in Astroparticle Physics, Krakow, Poland, 10–14, January 2022, talk;

- S. Bashir (on behalf of LHCb), *Influence of Pythia parameters on event multiplicity*, 14th International Conference on Beauty, Charm and Hyperons, Krakow, Poland, 5-11, June 2022, talk;

- S.Bashir (on behalf of LHCb), *Generators tuning infrastructure*, 106^{th} LHCb Week, CERN, Switzerland, 05-09, December 2022, talk, parallel;
- S.Bashir, *HSF Event Generator Tuning Workshop*, Krakow, Poland, 27-28 June, 2023, online;
- S.Bashir (on behalf of LHCb), *Hadron Production at LHCb Experiment*, EPS-HEP 2023, Hamburg, Germany, 20-25, August 2023, talk;
- S.Bashir, *Tuning Pythia 8 and LHCb tuning infrastructure*, 109th LHCb Week 25 - 29 September 2023, Marseille, France, talk, plenary;
- S.Bashir, *Multiparton Interaction parameters tuning in PYTHIA 8 and LHCb Tuning Infrastructure*, 13th LHC students poster session, 27, November 2023, CERN, Switzerland, poster;

As a member of the LHCb collaboration, the author published the following proceedings on behalf of LHCb:

- S. Bashir, Agnieszka Oblakowska-Mucha *Nucleon Structure and Soft QCD from LHCb*,PoS LHCP2021 (2021) 348;
- S.Bashir, Agnieszka Oblakowska-Mucha *Evidence of Double Parton Scattering at LHCb*, Acta Phys. Pol. B, 15 (2022) 33;
- S.Bashir, Agnieszka Oblakowska-Mucha *Impact on Multiplicity of Particles by Changing Multiparton Interaction Parameters in PYTHIA 8.3 at LHC Energies*, Acta Phys. Pol. B, 16, 3 (2023);
- S.Bashir *QCD Dynamics and Hadron Production in High-Energy Nuclear Collisions at LHCb.*, PoS EPS-HEP2023 (2024) 252.

First author Contributions:

- S.Bashir, Agnieszka Oblakowska-Mucha, Gloria Corti *Comparing particle multiplicity predictions: Insights from* Pythia, Herwig, *and LHCb data* under experiment review

Important seminar talks:

- Probing Soft QCD and Double Parton Scattering at LHCb, March 18, 2022, Seminarium HEP Białasówka / HEP Seminar, AGH University Krakow, Poland.

Participation in important schools:

- 15th MCnet school and 62nd Cracow School of Theoretical Physics, Jun 19–25, 2022.
- HASCO Summer School 2021, July 18-26, 2021, Online.
- PHYSTAT Workshop on "Statistics meets ML" in Particle Physics and Astrophysics, 9–12 Sept 2024, London, UK, Online

Other important activities (including international cooperation):

- XIV International Conference on Beauty, Charm and Hyperon Hadrons, 5-11 June 2022, Kraków, Poland, Organising Committee;
- Krakow Applied Physics and Computer Science Summer School (2020, 2021), June-August, Organising Committee, Co-supervision of a project.
- Participation in 6th Workshop on LHCb upgrade II, 29-31 March, 2023. Barcelona, Spain

Additionally, the author contributed to the LHCb collaboration in the following activities:

- Involvement in Run 3 data taking and taking data manager shifts (2022)
- DQCS (data quality, simulation and computing shifts) (2022-2023).

The author was a teacher and has taught the following tutorials and computer laboratories:

- Introduction to probability, statistics and data handling to Erasmus students and computer laboratories using Statgraphics software (2022-2025).
- Computer Science Statistics to Bachelor's students and computer laboratories in Python (2022-2025).
- Physics 1 for Computer Science students (2025)

- Physics 2 for Computer Science students (2024)

The author was a Volunteer at the following:

- "Accelerating Science" CERN's Flagship Exhibition in Krakow , 11-26 May 2024, Kraków, Poland, Volunteer;

Chapter 1

Introduction to Standard Model of Particle Physics

This chapter describes the Standard Model of particle physics, which is the most successful and comprehensive framework in our quest to understand fundamental constituents of the universe and forces that govern their interactions. It explores the complexities of the Standard Model, offering a conceptual journey into the subatomic realm. From quarks to leptons and force-carrying bosons that mediate their interactions, this model encapsulates the essential building blocks of matter and the underlying principles that shape our physical reality.

1.1 Elementary particles and matter content

The Standard Model (SM) of particle physics describes the constituents of matter. It consists of three generations of quarks and leptons. The interaction between quarks and leptons takes place with the help of force-carrying particles (mediators), which are referred to as "gauge bosons". Both the quarks and leptons are fermions (half-integer spin particles) following Fermi-Dirac statistics and bosons are (integer spin particles) following Bose-Einstein statistics. For all the particles, there are anti-particles which have similar properties to the particles, except that they carry an opposite electric charge. Each lepton family is composed of a charged lepton and a neutrino, whereas for the quarks, there are six "flavors". In total the SM has 24 fermion fields: 18 are quarks, 6 flavors ("up", "down", "charm", "strange", "top", "bottom") times three colors ("red", "green", "blue"), six leptons, 3 charged (electrons, muons and tauons) and the corresponding

neutrinos. The up, charm, and top quarks carry a +2/3e charge, the down, strange, and bottom quarks carry a -2/3e charge, the leptons (electrons) carry a -1e charge (e is an elementary charge), whereas neutrinos carry no charge. Table 1.1 shows the three leptons and quark generations, along with their constituent masses.

Generation	**1st**	**2nd**	**3rd**
Quarks	Up (u) 336 MeV/c^2, Down (d) 340 MeV/c^2	Charm (c) 1550 MeV/c^2, Strange (s) 486 MeV/c^2	Top (t) 177,000 MeV/c^2 , Bottom (b) 4730 MeV/c^2
Leptons	Electron(e^-) 0.5 MeV/c^2, Electron Neutrino (ν_e)	Muon (μ^-) 105 MeV/c^2, Muon Neutrino (ν_μ)	Tau (τ^-) 1776.8 MeV/c^2, Tau Neutrino (ν_τ)

Table 1.1: Three Generations of Quarks and Leptons along with their masses [1].

The quarks and leptons are point-like objects, i.e., they don't have any internal structure, hence, they are called the fundamental particles. However, the quarks combine and form hadrons, which are significantly larger as compared to quarks ($\sim 10^{-13}$ cm). There are two types of hadrons shown in Fig. 1.1:

- **Baryons**: comprising of three colorless quarks. They possess a distinctive attribute known as baryon quantum number, denoted as B=1, and half-integer spin, which makes them fermions. E.g. protons have two u-quarks and a single d-quark ($p = uud$), and neutrons have two d-quarks and a single u-quark ($n = udd$).

- **Mesons**: these are made up of a quark and an anti-quark, their baryon charge is nil (B=0), and their spin is an integer that makes them bosons. E.g. π^0, π^+=$u\bar{d}$ and π^-=$\bar{u}d$.

Beyond the conventional combination of quarks found in baryons and mesons, there exist more complex forms known as exotic quarks, including tetra-quarks and penta-quarks [3]. Tetra-quarks are made up of two quarks and two anti-quarks, while penta-quarks consist of four quarks and one anti-quark. These unique quark combinations illustrate the complex nature of the strong force. The identification and verification of these particles, especially through experiments at the Large Hadron Collider (LHC), have been instrumental in enhancing our understanding of quantum chromodynamics (QCD), the theory that explains how quarks interact under the strong force [4].

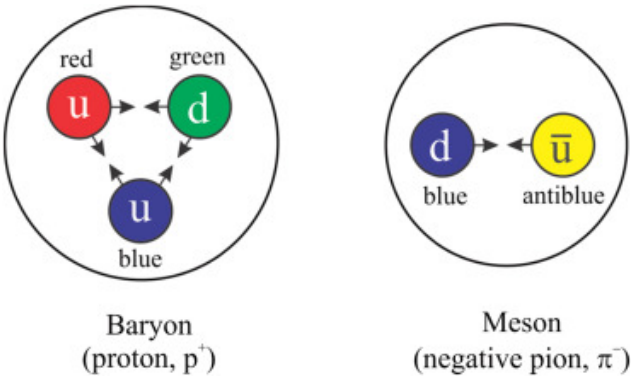


Figure 1.1: Baryons consist of three quarks (red, green, and blue), and mesons contain two quarks, namely a quark and an anti-quark. They are attracted to each other by means of the strong gluon force [2].

1.2 Interactions

The vast variety of physical phenomena is determined by one of the following four types of interactions: electromagnetic, weak, strong, and gravitational. The gauge bosons are mediators of these fundamental interactions. The electromagnetic force is carried by a spin-1 massless vector boson, the photon (γ), and acts between electrically charged particles. On the other hand, the weak interaction is 10^3 times weaker than the electromagnetic force and is ruled by three gauge vector bosons, the $W^{\pm}$ and Z^0 bosons, which have large masses and are spin 1 particles. The strong interactions are responsible for holding the nuclei together and are roughly 100 times stronger than the electromagnetic force. The gauge bosons for the strong force act between the quarks and the eight massless, spin-1 particles called gluons (g). And lastly, the gravitational interaction appears between all types of massive particles and is by far the weakest (the coupling constant is about 10^{42} times smaller than the electromagnetic); therefore, its effect is negligible at the subatomic scale. Graviton (G) is a hypothetical spin-2 massless gauge boson responsible for the gravitational force. Table 1.2 shows all the gauge bosons with charges, masses, spin, and the type of interactions they are responsible for.

There has been an effort to unify all four fundamental interactions by expressing them as different manifestations of a single fundamental interaction field. The unification is partially achieved by SM.

The high energies and the minute distances where the electroweak force and strong force merge are beyond the reach of any current accelerator. For example, to be able to probe a distance of 10^{-30} m, one would require an energy of 10^{14} GeV or beyond, which

Gauge Boson	Involved particles	Q/e	Mass [GeV/c^2]	Spin	Interaction
photon (γ)	charged particles	0	$< 10^{-27}$	1	electromagnetic
W boson (W)	leptons and quarks	$\pm$ 1	80.385 $\pm$ 0.015	1	weak
Z boson (Z)	leptons and quarks	0	91.1876 $\pm$ 0.0023	1	weak
gluon (g)	hadrons	0	$\leq 10^{-3}$	1	strong
graviton (G)	all particles	0	$< 6 \times 10^{-41}$	2	gravitational

Table 1.2: Fundamental properties of six gauge bosons mediating the four fundamental forces [5].

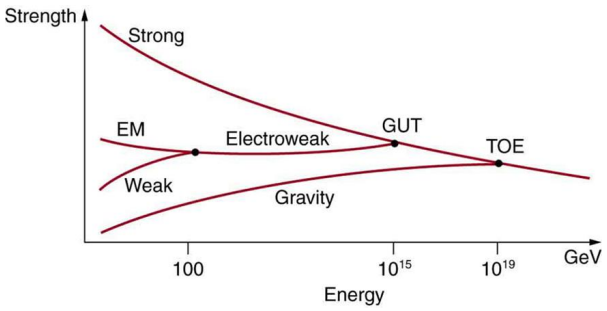


Figure 1.2: The relative strengths of the forces varying with distance plotted against energies [GeV] scale [6]. EM stands for electromagnetic interaction, GUT stands for grand unified theory, and TOE is the theory of everything.

is very far from the capacity of the LHC. Since these conditions are unexplored by the existing experiments, which are essential for testing the Grand Unified Theories (GUTs), they remain ambiguous due to the lack of direct evidence.

1.3 Theoretical foundations of the Standard Model

The SM of particle physics is a relativistic quantum field theory (QFT) that describes the electromagnetic, weak, and strong interactions, governed by the local SU(3)$C \otimes$ SU(2)$L \otimes$ U(1)$_Y$ gauge symmetry. SU(3)$_C$ represents the non-abelian gauge group, with 8 massless gauge bosons (gluons) which are generators of the gauge group and hold quarks together mediating strong force; the letter "C" in SU(3)$_C$ stands for the color. SU(2)$_L \times$ U(1)$_Y$ represents the electroweak symmetry group, which unifies the electromagnetic and weak interactions in the so-called "electroweak theory". The U(1)$_Y$ symmetry group involves both the right and left-handed fermion fields. SU(2)$_L$ and U(1)$_Y$ are the components of electroweak theory that are generated by weak isospin (T) and weak hypercharge (Y), respectively. These two quantum numbers are related to the electric charge (Q) by the

Gell-Mann-Nishijima formula [7], which relates the baryon number (B), the strangeness (S) and the third component of isospin (I_3) of quarks and hadrons to their electric charge:

$$Q = I_3 + \frac{1}{2}(B + S) \tag{1.1}$$

$$Q = T_3 + \frac{Y}{2} \tag{1.2}$$

where T_3 is the third component of weak isospin.

Noether's theorem [8] implies that, if an action is invariant under some group of transformations (symmetry), there exists one or several conserved quantities called "constants of motion", which are associated with these transformations. This theorem establishes that symmetries directly imply conservation laws: as a matter of fact, the invariance under local gauge transformations of the group implies the dynamics itself, as evident in quantum electrodynamics: the interaction between electron and photon naturally follows from the gauge invariance under U(1). Local gauge symmetry transformations govern all quantum field theories, quantum electrodynamics, quantum chromodynamics, and SM of electroweak interactions. These theories describe physical phenomena extremely well, which is confirmed experimentally.

The particles are represented as quantum field excitations. QFT describes particles as mathematical fields, i.e., scalar (spin 0) field, vector (spin1/2) field, or fermions and spin 1 as bosons. The total SM Lagrangian comprises the kinetic term for bosons and fermions, the coupling of the Higgs field to the fermions and bosons, and also the interaction part. The Lagrangian density of the SM can be written as the sum of the electroweak and strong Lagrangians:

$$\mathcal{L}_{\mathcal{SM}} = \mathcal{L}_{\mathcal{EW}} + \mathcal{L}_{\mathcal{QCD}} \tag{1.3}$$

where $\mathcal{L}_{\mathcal{EW}}$ represents the Lagrange density of electroweak interactions, and $\mathcal{L}_{\mathcal{QCD}}$ is the Lagrange density for strong interactions or quantum chromodynamics.

1.3.1 Electroweak interactions

The electroweak interactions can be written as the summation of two parts of the Higgs boson couplings:

$$\mathcal{L}_{\mathcal{EW}} = \mathcal{L}_{\mathcal{S}} + \mathcal{L}_{\mathcal{H}}. \tag{1.4}$$

The $\mathcal{L}_{\mathcal{S}}$ is the Lagrangian for the symmetry which involves only gauge bosons and fermions, and the field of quarks and leptons is represented by ϕ:

$$\mathcal{L}_S = -\frac{1}{4}\sum_{A=1}^{3} F^A_{\mu\nu}F^{A\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} + \bar{\phi_L}i\gamma^\mu D_\mu\phi_L + \bar{\phi_R}i\gamma^\mu D_\mu\phi_R \tag{1.5}$$

where,

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \quad \text{and} \quad F^{A\mu\nu} = \partial_\mu W^A_\nu = \partial_\nu W^A_\mu - g\epsilon_{ABC}W^B_\mu W^C_\nu. \tag{1.6}$$

γ^μ are the Dirac matrices, B_μ is the gauge field which is associated with $U(1)$ symmetry and W^A_μ where A $\in$ 1,2,3 represents three gauge fields of $SU(2)$. The ϵ represents the anti-symmetric tensor of SU(2) group. The couplings of the relevant gauge bosons to the fermion fields ϕ are included in the covariant derivative D_μ. The fermion fields with their left and right-handed chiral components are the projections of the chiral operators:

$$\phi_{L,R} = [1 \pm \gamma^5/2]\phi, \quad \gamma^5 \equiv i\gamma^0\gamma^1\gamma^2\gamma^3. \tag{1.7}$$

The ϕ_L and ϕ_R behave differently under the electroweak interactions. In the SM all the fermion state ϕ_R are singlets (u_R, d_R, ν_R, e_R) and all ϕ_L are doublets.

$$Q_L = \begin{pmatrix} u_L \\ d_L \end{pmatrix}, L_L = \begin{pmatrix} \nu_L \\ e_L \end{pmatrix} \tag{1.8}$$

The requirement of gauge invariance makes both fermions and gauge bosons massless, which is in contradiction with experiments. Therefore, another component was added to the Lagrangian - $\mathcal{L}_{\mathcal{H}}$, which is eventually responsible for the coupling of the Higgs field to the fermions, generating their masses in that way. The Lagrangian, which includes additional potential acting on the scalar field ψ:

$$\mathcal{L}_{\mathcal{H}} = (D_\mu\psi)^\dagger(D_\mu\psi) - V(\psi^\dagger\psi) + \mathcal{L}_{\mathcal{Y}} \tag{1.9}$$

$$V(\psi) = \mu^2\psi^\dagger\psi - \frac{\lambda}{4}(\psi^\dagger\psi)^4, \tag{1.10}$$

where μ, λ are constants. Once V(ψ) takes the minimal value, the $|\psi|$ is non-zero and the symmetry of the Lagrangian is spontaneously broken. The variations around the

minimum of V finally give mass to the fermions and to the field B_μ and W_μ.

The mass term for the lepton is proportional to the Yukawa coupling Y_L and the Higgs vacuum expectation value:

$$\mathcal{L}_\mathcal{Y} = Y_d^{ij} + Y_u^{ij} + Y_c^{ij} + h.c, \tag{1.11}$$

where h.c. is the hermitian conjugate.

The matrices Y_d and Y_u are complex matrices that represent the coupling between different generators ij of quarks and, in general, contain off-diagonal terms, allowing the mixing between families. In order to diagonalize them, mass matrices are unitary transformed, what allows the introduction of Cabibo-Kobayashi-Maskawa matrix V_{CKM}.

$$V_{CKM} = V_{uL}^\dagger V_{dL} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}, \tag{1.12}$$

where each element V_{ij} expresses the coupling strength of the weak interaction between the quark i and j and transforms the quarks from the mass (flavor) eigen-state (d, s, b) into weak eigen-state $(d^{'}, s^{'}, b^{'})$.

Due to the unitary condition, the CKM matrix contains four non-reducible parameters, which in standard parametrisation are three magnitudes of quark transitions and a single phase, which is responsible for the CP violation in the SM. This results in different behavior of particles and anti-particles in weak interactions. The CKM elements can be obtained from the measurements at LHCb [9].

1.3.2 Quantum chromodynamics (QCD)

Quantum Chromodynamics (QCD) describes strong interactions between quarks. The symmetry of QCD is based on the non-abelian $\mathrm{SU}(3)_C$ group acting in a three-dimensional space (represented by the colors) with eight generators. The QCD Lagrangian has the form:

$$\mathcal{L}_{\mathcal{QCD}} = -\frac{1}{4}\sum_{A=1}^{8} F_{\mu\nu}^A \, F^{A\mu\nu} + \sum_j^{n_f} \bar{q}_j \left(i\gamma^\mu \mathcal{D}_\mu - m_q\right)_{ij} q_j, \tag{1.13}$$

where q_j are the quark fields of n_f different flavors with mass m_j and $F_{\mu\nu}$ is the field strength. Covariant derivative in the case of non-abelian gauge theories takes the form:

$$D_\mu = \partial_\mu - ig_s \sum_A T^A g_\mu^A \tag{1.14}$$

In the equation above, g_s is the gauge coupling of QCD. The G_μ^A are the gluon fields with A = 1,...,8 and T^A are the $SU(3)_c$ group generators in the triplet representation of quarks. The generators of $SU(3)_c$ group are often written in a standard basic:

$$T^a = \frac{1}{2}\lambda^a, \tag{1.15}$$

with λ^a matrices (Gell-Mann matrices) of the form:

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$
$$\lambda_4 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \quad \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix},$$
$$\lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}.$$

Because of the requirement of the local gauge invariance for the non-abelian group, there are not only quark-gluon interactions, but also gluon-gluon couplings with three or four gluons, which bring in a lot of new effects, including glueballs and confinement. Therefore, gauge fields in QCD carry color charge and interact strongly.

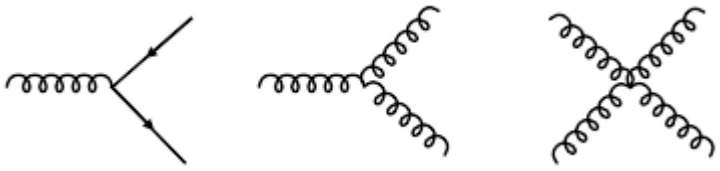


Figure 1.3: The QCD quark-gluon and gluon-gluon vertices [10].

Due to self-interactions of gluons as shown in Fig. 1.3, quarks (colored objects) are always confined to color singlet states and cannot propagate as free particles. Another

aspect of gluons' self-interaction is that the coupling constant α_s, which describes the strength of the interactions, can be large.

QCD could be perturbative (pQCD) if the coupling constant were small, causing higher-order diagrams to asymptotically vanish. This would result in the observation of free quarks, which have not yet been proven.

The perturbative approach to the interaction of quarks is valid on the scale of very small distances or very high energies. At shorter distances, quarks are asymptotically free, and the effective strong interaction constant becomes small. For low low-energy process, pQCD is not a proper theory. The dependency of the α_s on the momentum transfer Q^2 became the main experimental proof that QCD is a renormalized field theory [11].

The strong coupling constant, due to the gluon self-interaction, is not constant but is parameterised as a function of the momentum transfer Q^2 between the particle and the source:

$$\alpha_s(Q^2) = \frac{g^2/4\pi}{1 + \beta\alpha_s Q_0 ln\left(\frac{Q^2}{\Lambda^2_{QCD}}\right)}. \tag{1.16}$$

In equation 1.16, β is a positive number, so α_s decreases as Q^2 increases, and this explains the asymptotic freedom. It incorporates the renormalization of QCD, discovered by David Gross, Frank Wilczek, and David Politzer [11] and introduces the second dimensional parameter Λ_{QCD}. This remarkable parameter indicates the limit of perturbative QCD at $Q^2 = \Lambda_{\mathrm{QCD}}$. Here, α_s becomes infinite, while $\alpha_s \to 0$ as $Q^2 \to \infty$. Small values of Q^2 are associated with larger distances, and $\Lambda^{-1}_{\mathrm{QCD}}$ is called the confinement length. The value of Λ_{QCD} is around 300 MeV [12]. In this way, QCD is divided into a non-perturbative low-energy regime, where calculations using perturbative expansions are not possible, and a high-energy regime where perturbative theory can be effective. At $Q > 100$ GeV (LHC scale), $\alpha_s \approx 0.1$ and perturbative QCD (pQCD) can be used. However, this value is not sufficiently small to neglect higher-order corrections. For this reason, QCD calculations for the LHC involve many diagrams and are always calculated beyond the lowest (leading) order (LO), such as next-to-leading-order (NLO) and next-to-next-to-leading-order (NNLO).

1.4 Limitations of Standard Model

Although the Standard Model (SM) has been highly successful in explaining the physics of strong, weak, and electromagnetic interactions up to the TeV energy scale, with numerous

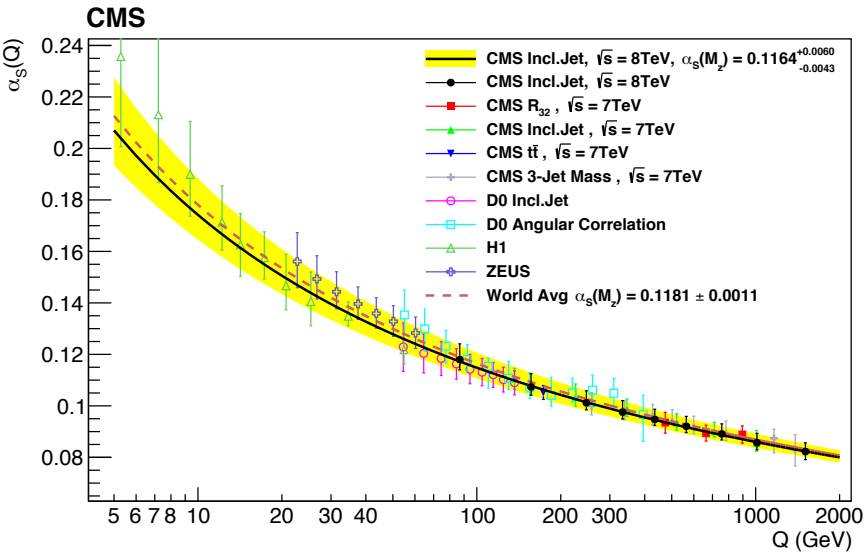


Figure 1.4: Measurements of α_s, including recent LHC results, as a function of energy scale Q [13].

experimental confirmations [13], it still has several notable shortcomings, such as:

- It lacks an explanation of gravity. However, we do have a spin-2 particle, the graviton, giving an effective description of gravitational interactions [14].
- The existence of the neutrino masses requires at least an addition of the ν_R state in the Lagrangian. The non-zero mass of neutrinos has been measured in flavor oscillations [15].
- Assuming that the Universe can be described by an effective quantum field theory up to Planck energies, we can compute the value of the cosmological constant. This determination is larger than what is observed by more than 120 orders of magnitude [16].
- It lacks a sufficient explanation of matter and anti-matter asymmetry. They were expected to be produced in equal amounts during the creation of the universe. The fact that the present-day universe mostly consists of matter is an unexplained phenomenon. The violation of CP in electroweak interactions is a known source of baryon asymmetry, this is the main purpose of the LHCb experiment [17].
- The absence of CP violation in the strong sector is another source of unexplained fine-tuning in the SM that might be clarified in SM extensions.

- Conclusive evidence supports the presence of non-luminous matter as the predominant form in the Universe. The first indications were drawn from measurements of the galactic rotational speeds and orbital velocities within the clusters. Recently, more advanced techniques like gravitational analysis of the cosmic microwave background radiation, primordial nucleosynthesis, and large-scale structures, proven the existence of dark matter. The search for dark matter spans various theories, with potential origins from the SM. The candidates are expected to be heavy, electrically neutral, and weakly interacting; these attributes are labeled as Weakly Interacting Massive Particles (WIMPs) [18].

Chapter 2

Large Hadron Collider at CERN

Humanity's inherent curiosity and desire to understand the world have driven us to explore the smallest components of matter—the fundamental building blocks of everything around us. To aid this pursuit, scientists have constructed massive machines, with one of the most remarkable being the Large Hadron Collider (LHC). This extraordinary device enables researchers to study these tiny particles by accelerating protons to incredibly high energies and colliding them, revealing the secrets of the subatomic world.

2.1 About LHC

The Large Hadron Collider (LHC) [19] is the world's largest and highest-energy particle accelerator located in the European Organisation for Nuclear Research, CERN, near Geneva, Switzerland. The accelerator consists of a 27-kilometre-long tunnel situated on both sides of the France-Swiss border. An overview of the accelerator is shown in Fig. 2.1.

CERN was founded in 1952 to establish a research organisation to explore fundamental physics without any military application. Initially, the focus was to explore atom and nuclear physics, but today the knowledge is far beyond and is now more leaned on the fundamental constituents of matter and the forces that hold them together.

LHC Layout

The 27 km long tunnel consists of eight arcs, and the depth varies from 45 to 170 m. Fig. 2.2 shows a schematic layout of the LHC. There are four intersection points (IP) where the major experiments are located. The IP1, IP2, IP5 and IP8 are ATLAS [22],

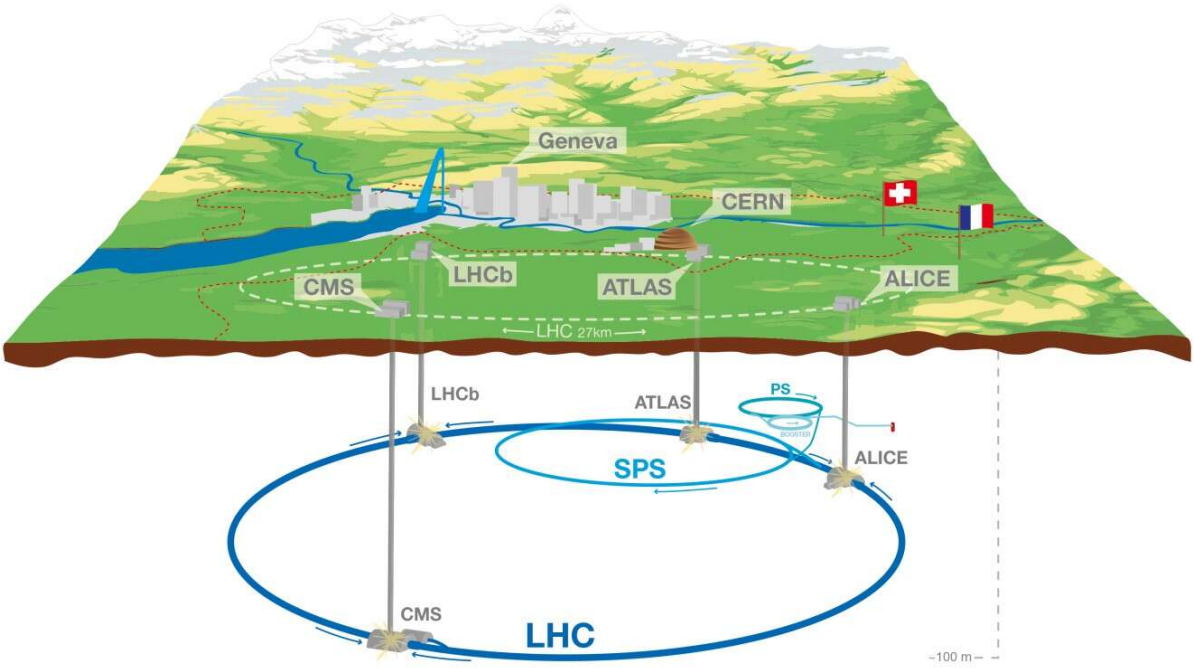


Figure 2.1: An overview of the LHC with its 4 main experiments [20].

ALICE [23], CMS [24] and LHCb [25] respectively. CMS and ATLAS are general-purpose detectors with the physics program ranging from the study of the Standard Model to physics beyond the Standard Model (BSM). ALICE and LHCb have specific physics programmes. ALICE detector is designed for the study of strongly interacting matter at extreme densities, whereas LHCb specialises in studying heavy flavour physics and the differences between matter and anti-matter (CP asymmetry).

Apart from the larger experiments, there are also three small experiments situated in the close vicinity of the main experiments at the LHC. TOTEM [26] is an experiment dedicated to the measurement of proton-proton (pp) interaction cross-sections and in-depth studies of the structure of protons. The location of TOTEM is at IP5 on either side of the CMS. LHCf [27] aims to study the neutral particle production cross-section in the forward region of pp and heavy ion collisions, which is located at IP1 next to the ATLAS experiment. Lastly, MoEDAL [28] is an experiment that searches for magnetic monopoles and is located in front of the LHCb experiment at IP8.

Protons' journey through LHC Accelerators

In the LHC, beams of protons and ions are accelerated. The proton source is the bottle of hydrogen gas, which is injected at the starting point of the linear accelerator (LINAC

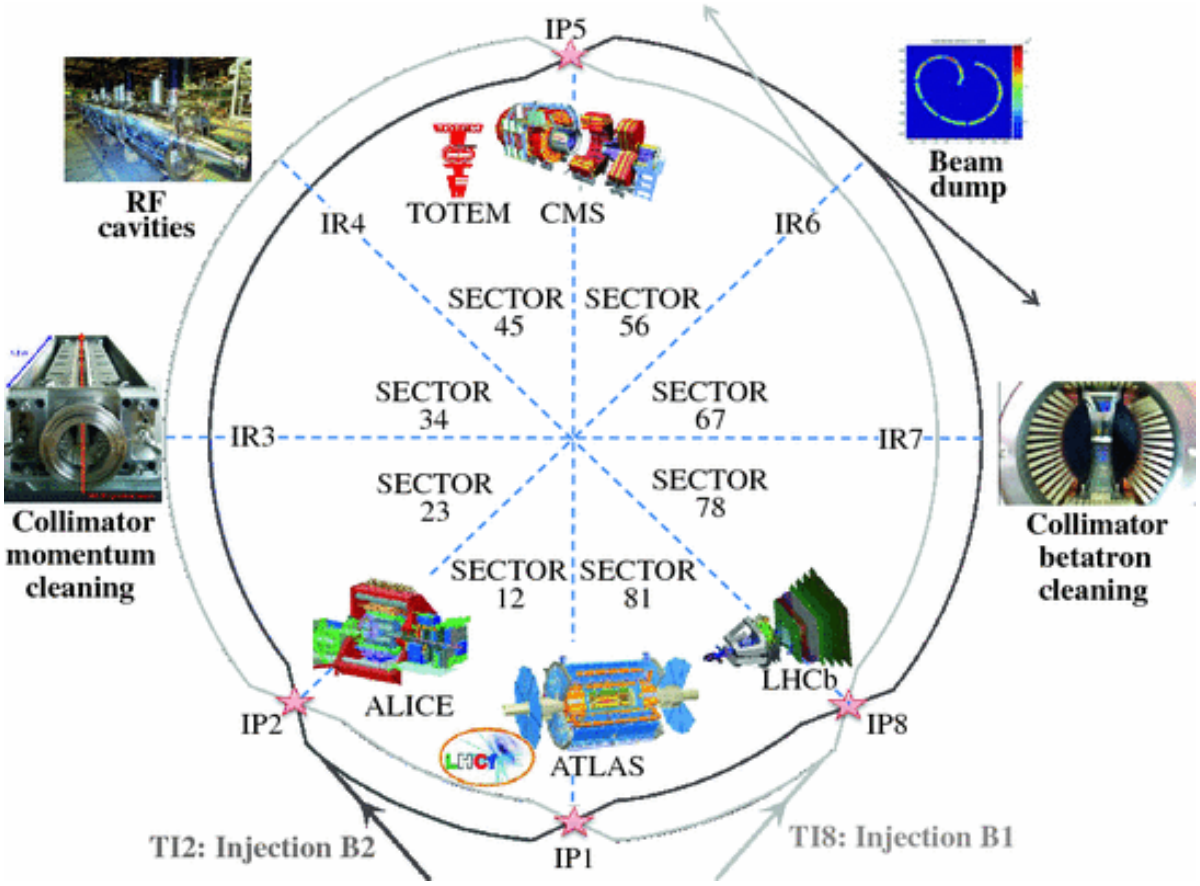


Figure 2.2: Geometry of LHC with eight arcs [21]. The two beams intersect at IP1, IP2, IP5 and IP8.

2), where the electrons are stripped off with the electric field. At the end, the remaining protons reach an energy of 50 MeV. The protons are then accelerated further up to 1.4 GeV in the proton booster. To further increase the energy, the protons are fed to the proton synchrotron (PS), which boosts the energy up to 25 GeV and packs them into 25-ns bunches. Protons are further led to the super proton synchrotron (SPS), where they gain an energy of 450 GeV. Finally, the protons enter the LHC ring, where the energy is increased to almost 7 TeV. The schematic image of the accelerating system is shown in Fig. 2.3.

Apart from the protons, ions are also accelerated and collide at the LHC. The ion source is lead, which is evaporated in the oven. The lead gas then loses electrons by passing through the plasma. The ions are first accelerated in the LINAC 3 and further injected into the low-energy ion ring (LEIR), which accelerates them, and then follow the same chain as protons. Therefore, the particles are initially injected into the PS and subsequently into the SPS, as illustrated in Fig. 2.3.

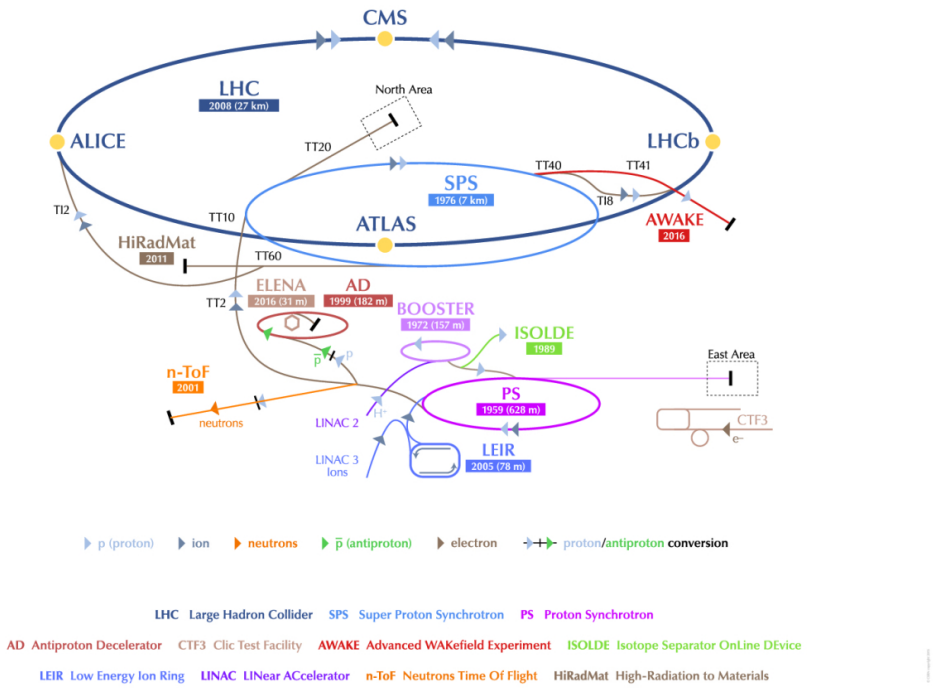


Figure 2.3: CERN accelerator complex with LINAC, PS, SPS, and LHC [29].

2.2 LHCb Experiment

LHCb [25] is one of the four main experiments located at LHC. It is a single arm forward spectrometer, with 20 m length and 10 m height and width. This experiment is dedicated to studying decays of particles containing b and c quarks. Fig. 2.4 shows the schematic overview of the LHCb detector. The LHCb experiment has a unique pseudorapidity coverage of $2< \eta <4.5$, unlike other experiments of LHC, as shown in Fig. 2.5. The detector geometry originates from the fact that b and $\bar{b}$ quarks are produced in pairs in the forward or backwards direction, and the geometry of the detector allows a large fraction of b-hadrons to be reconstructed by covering a small solid angle.

The LHCb experiment has now broadened its physics program. It includes: the study of electroweak gauge bosons in the forward region, since this is the region that is unexplored by the other detectors, measurements of newly discovered exotic particles and their properties, lepton-flavour-violating decays of Higgs-like bosons [32].

The pp collisions in the forward region are dominated by the high particle multiplicities,

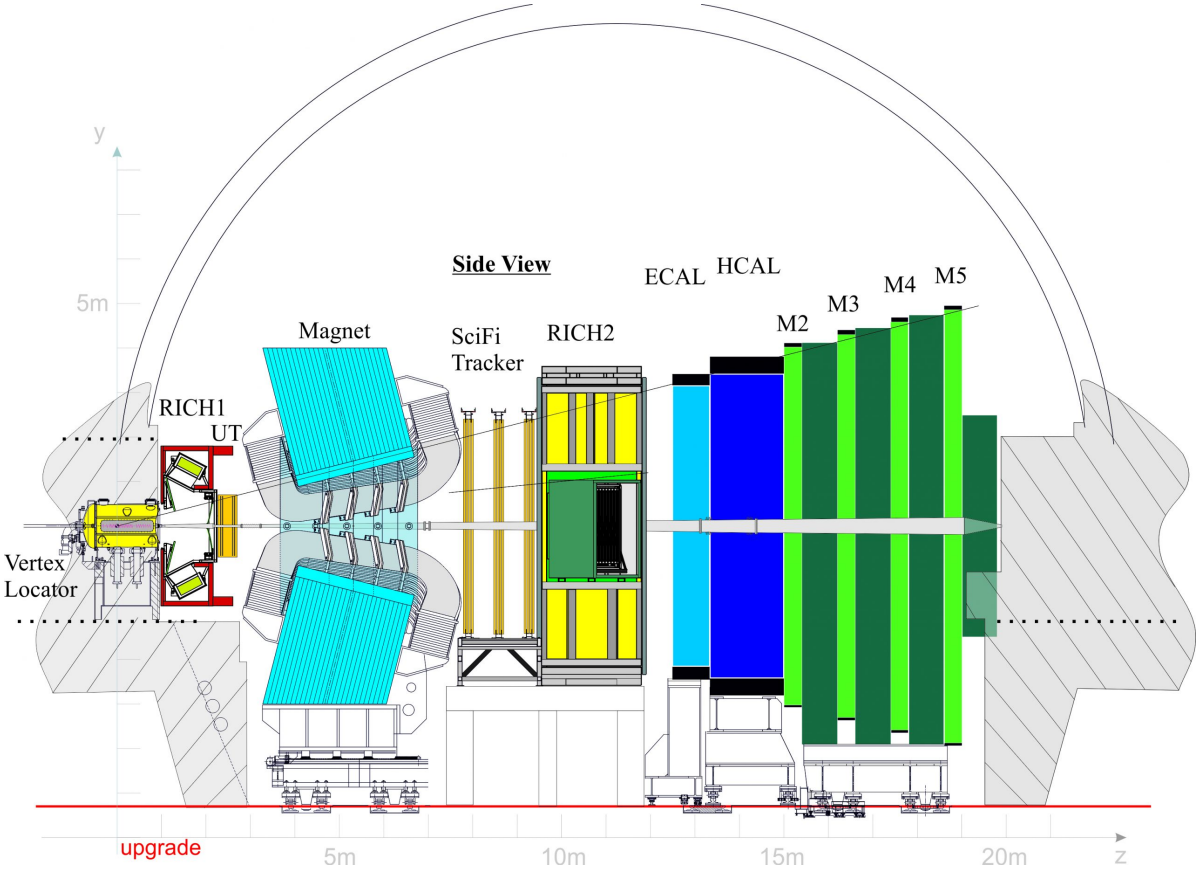


Figure 2.4: Upgraded LHCb detector at LHC [30].

which result in the high occupancy of the detector elements. Hence, it is very challenging to separate the primary and secondary vertices and to reconstruct the tracks that are important for analyses. LHCb has high quality particle identification system that could differentiate the primary from secondary vertices and also precisely track the trajectories of the particles. LHCb is optimised for an average luminosity of 2×10^{32} cm^2 s^{-1}, which is 100 times smaller than CMS and ATLAS.

In 2011, LHCb collected data at luminosity of 3.5×10^{32} cm^2 s^{-1} with $\mu = 1.5$ at $\sqrt{s} = 7$ TeV, and in 2012 at 4×10^{32} cm^2 s^{-1} with $\mu = 1.7$ at $\sqrt{s} = 8$ TeV. An integrated luminosity of 3.19 fb^{-1} was accumulated in Run 1 at $\sqrt{s} = 7$ TeV, and 5.6 additional fb^{-1} were collected at $\sqrt{s} = 13$ TeV at the end of Run 2. Fig. 2.6 shows the integrated luminosity accumulated by the LHCb experiment during the Run 1 (2011-2012) and Run 2 (2015-2018). After the break in data taking, LHCb started taking data in 2022 with a modernised detector (Upgrade I), with the plan to upgrade the "Upgrade II" during LHC High-Luminosity Runs. LHCb detector is supposed to collect data that corresponds to the integrated luminosity of 300 fb^{-1} by the end of 2040,

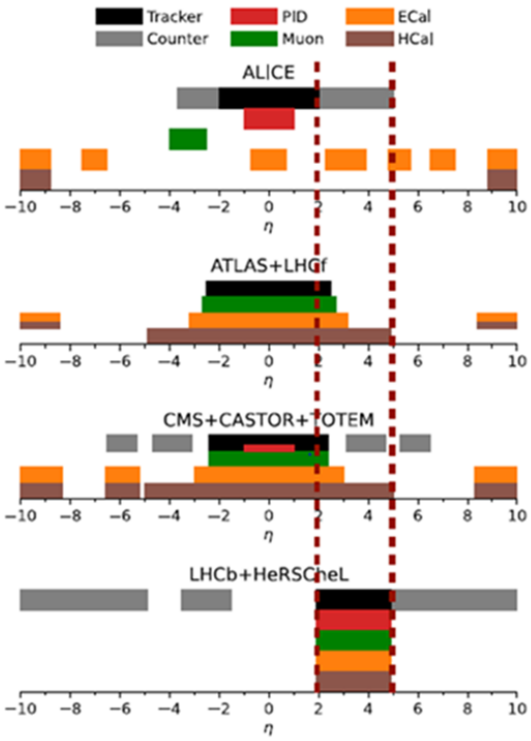


Figure 2.5: The pseudorapidity coverage of the LHCb experiment as compared to other detectors [31].

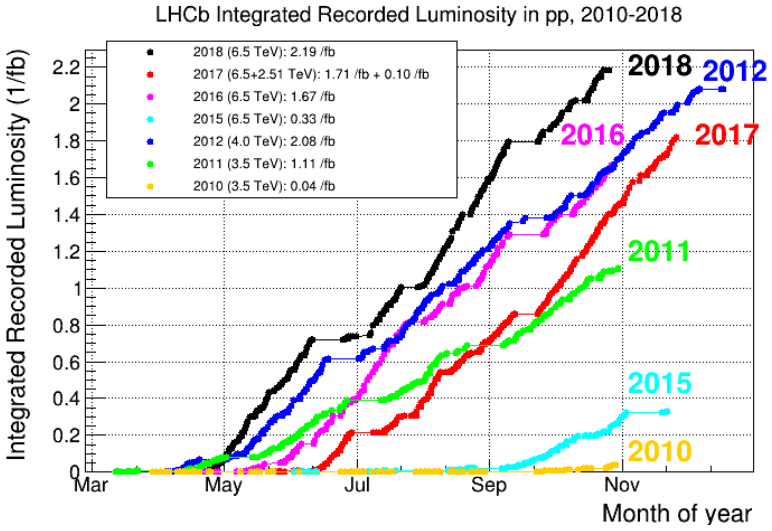


Figure 2.6: Integrated luminosity collected by the LHCb experiment by the end of Run 2 [33].

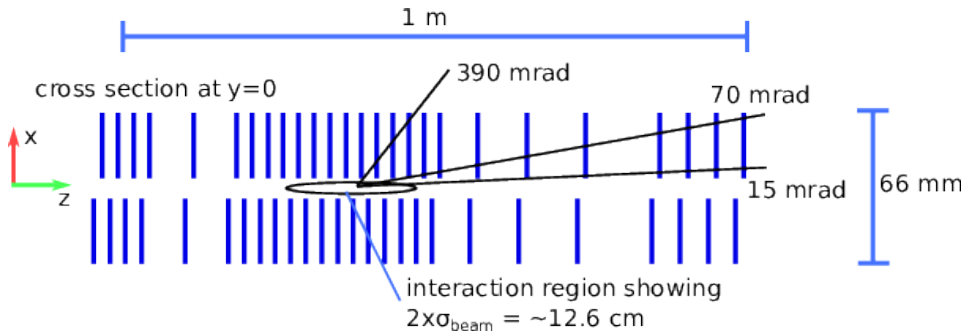


Figure 2.7: Schematic representation of VELO. There are 52 detection modules arranged in two retractable sides. The interaction region is populated with the highest density of modules [37].

2.3 LHCb sub-detectors

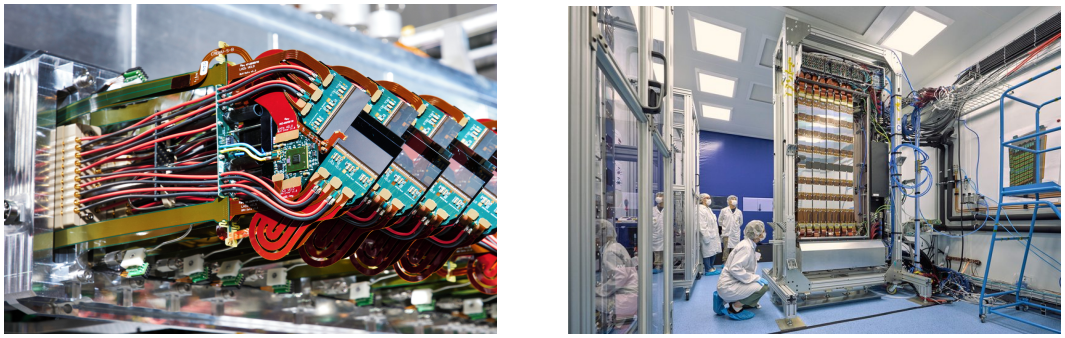
The LHCb spectrometer is composed of several sub-detectors, which have different purposes. These are detectors for tracking, particle identification, and calorimeters. Part of the devices are also used for the triggering. The following sections provide details of these sub-detectors.

2.3.1 Tracking

The tracking system of the LHCb spectrometer consists of three sub-detectors: pixel Vertex Locator (VELO) [34], silicon strip Upstream Tracker (UT) [35], and three stations of scintillating fibres (SciFi) [36] which are placed almost 8 m from the *pp* interaction point (IP). Due to this geometric configuration, not all traversing particles can be reconstructed in every tracking station.

VELO

The VErtex LOcator (VELO) is the closest detector to the collision point and consists of a series of vertically oriented pixel sensors arranged in stations along the beam-line as shown in Fig. 2.7. VELO pixels provide the x-y coordinates, whereas the position of the station gives the z coordinate of the track element. The collisions occur in a 5 cm long region along the z-axis. The VELO is placed upstream of the dipole magnet, in a region practically free from the magnetic field, which makes the tracks in the VELO straight lines. Further details on tracking are mentioned in section 2.6.



(a) (b)

Figure 2.8: Vertex Locators (VELO) arranged in a row [34] (a) and the first half of Upstream Tracker (UT) (b) ready for deployment [38].

UT

The Upstream Tracker (UT) is situated between the first identification detector (RICH1) and the magnet. It is used for the tracking of charged particles. It consists of four planes of silicon strip detectors, which are divided into two stations with a beam pipe hole in the centre. The detectors are arranged in vertical units known as staves within the plane. The initial station consists of two planes: one with vertical strips and the other with stereo strips tilted at 5 degrees. Conversely, the second station follows a similar configuration, but with the stereo layer tilted in the opposite direction. In the first station, both layers are composed of 16 strips, while in the second station, there are 18 strips in each layer. Every component of the detector has been designed with minimal thickness to reduce the exposure of particles to materials. Fig. 2.8 (b) shows the first half of the UT tracker ready for deployment.

SciFI

The former tracking stations, including the gas straw tube tracker and silicon tracker, are now replaced by the scintillating fibre (SciFi) tracker. It is situated downstream of the magnetic dipole and serves the purpose of tracking charged particles and measuring their momentum. This high-resolution detector spans approximately 340 m^2 and employs 0.25 mm diameter scintillating fibers for its construction, with readout facilitated by silicon photo-multipliers (siPMs). Comprising three stations, each equipped with four detector planes, the module consists of eight fibre mats, each 2.4 m long, housing six layers of tightly packed blue-emitting scintillating fibres. Fig. 2.9 shows the tracker aligned and

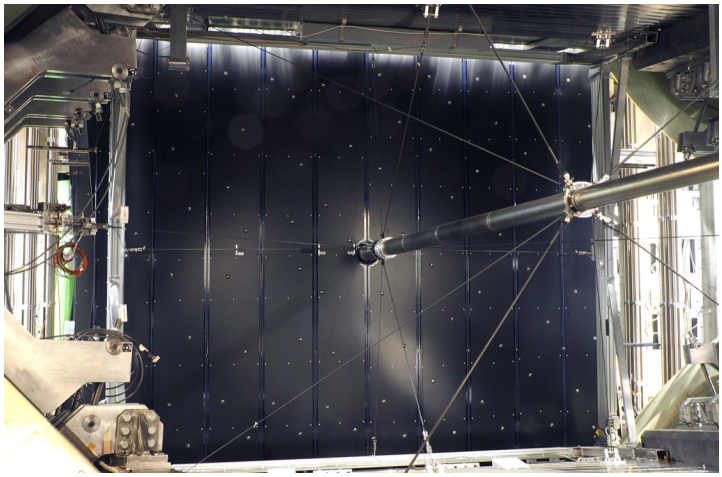


Figure 2.9: The C-frames of the SciFi tracker aligned and fully closed around the beam pipe [39].

fully closed around the beam pipe.

2.3.2 Magnet

The LHCb magnet is a dipole magnet that is placed between the tracking stations and uses non-superconducting technology. It operates at a magnetic field of around 4 Tm. The magnet can switch polarities, which is crucial for understanding the systematic asymmetries of the detector for the examination of factors that might have an impact on the CP-violation measurements. The magnetic field tends to bend the charged particles in the x-direction. Fig. 2.10 shows the complete installation of the LHCb dipole magnet.

2.3.3 Particle identification detectors

Particle identification (PID) is based on the classification of particles according to their masses. For the identification of stable particles, there are three sub-detectors at LHCb: two RICH detectors, calorimeters, and the muon system. The information from these sub-detectors is integrated into a combined likelihood probability function to maximise the identification efficiency. The three sub-detectors used for identification are briefly discussed below.

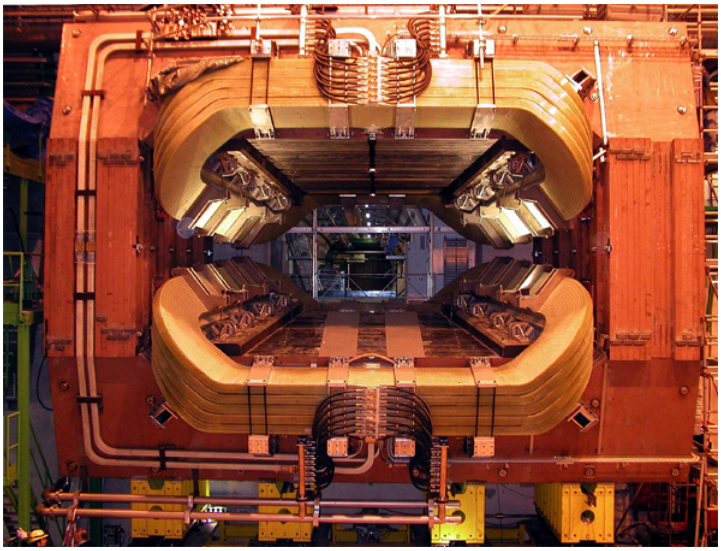


Figure 2.10: The complete installation of LHCb dipole magnet [40].

RICH

LHCb RICH [41] detectors are actually two different detectors, RICH1 and RICH2, located at different positions. RICH1 is located between VELO and UT, whereas RICH2 is between SciFi and ECAL. The primary role of these two sub-detectors is to identify the charged hadrons (particularly protons, pions, and kaons) in a wide range of momenta. The working principle of these detectors is based on the Cherenkov radiation, in which the light rings are produced by the charged particles when passing through the radiator gas with a specific refractive index n. The radius of the rings produced allows the measurement of the Cherenkov angle θ_c: $\cos\,\theta_c = 1/n\beta$, where β is the relative speed of the particle to the speed of light. The two RICH detectors with different radiators provide complete coverage in the momentum range from 2-100 GeV/c. Both of these detectors use different radiator gases. RICH1 uses aerogel and fluorobutane (C_4F_{10}) whereas, RICH2 has CF_4. Fig. 2.11 shows that the charged particles produced in the collisions in LHCb will travel through the mirrors of RICH1 before reaching measurement components further downstream. To reduce the amount of scattering, RICH1 uses special lightweight spherical mirrors constructed from a carbon-fibre reinforced polymer (CFRP), rather than glass [42].

RICH1 covers the low momentum to intermediate ranges, 2-40 GeV/c, whereas RICH2 covers the high momentum region from 15-100 GeV/c. The Cherenkov photons that are emitted by the charged particles when traversing through the RICH detectors are focused into the ring images on the photon detector planes, which are outside the LHCb

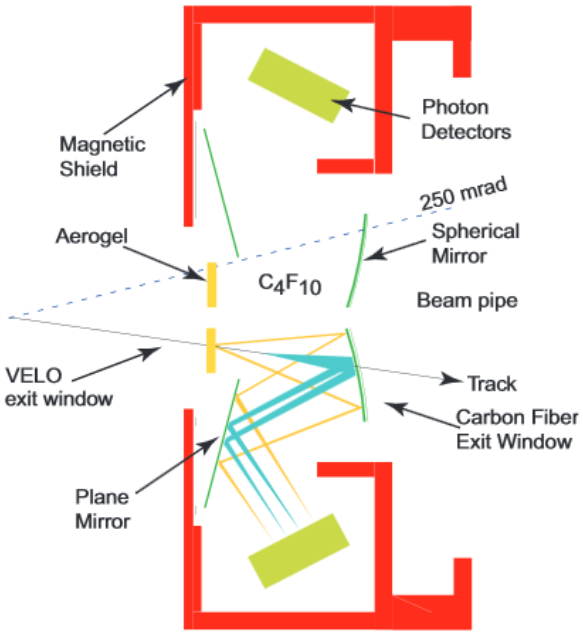


Figure 2.11: Charged particles produced in the collision moving across the RICH1 and scattering with lightweight spherical mirrors constructed with special material [42].

acceptance. A hybrid photo detector (HPD) is designed to capture these photons and is specifically designed for the RICH detectors. HPD is made of a quartz window and a multialkali photo-cathode, where the produced photo electrons are focused onto the silicon pixel array.

Calorimeter system

The calorimeter system [43] at LHCb consists of several sub-detectors, i.e., the Scintillating Pad Detectors (SPD), the Pre-Shower Detector (PrS) used in Run 2 and removed in Run 3, the electromagnetic calorimeter (ECAL), and the Hadronic Calorimeter (HCAL). Each of the above sub-detectors has its specific function. However, the common function among them is the identification of electrons and hadrons by measurement of their energy and location. The scintillation light that is generated when a charged particle passes through these detectors is transmitted to photomultiplier tubes, to enhance the signal and transform it to light and further to an electric signal. Each of the following systems has its own use:

- **SPD and PrS**: These are scintillator pads separated by lead with a thickness of 2.5 radiation lengths. Each SPD cell provides binary information based on the

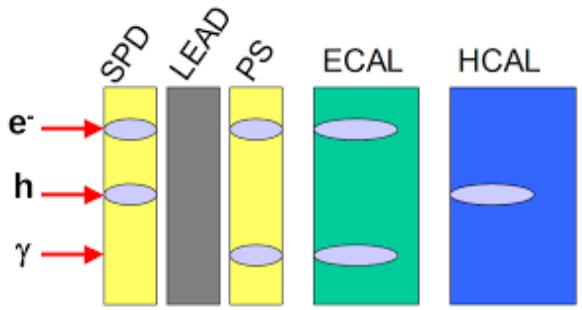


Figure 2.12: Signal deposited on the different parts of the calorimeter by an electron, a hadron, and a photon [44].

comparison between the deposited energy and a predefined threshold, enabling the discrimination between charged particles and neutrals. Subsequently, the energy is quantified within a Photon-Receiving Scintillating (PrS) cell, contributing to the identification of photons and electrons.

- **ECAL**: This detector is situated next to PrS detector. It is placed for the measurement of electron and photon showers. The multiple layers of lead are used as absorbers, the detector possesses a thickness equivalent to 25 radiation lengths. This design helps to achieve complete containment of high-energy electromagnetic showers and to optimize energy resolution. Moreover, the cumulative energy deposition within the Electromagnetic Calorimeter (ECAL) during an event serves the additional purpose of determining the centrality, i.e., the number of participating nucleons, in heavy-ion collision events.

- **HCAL**: This detector is located next to ECAL and provides information about the transverse energy of the hadrons. It has a thickness of 5.6 interaction lengths. Iron is used for absorption, and the scintillating tiles are the active material. Fig. 2.12 shows how various signals travel through the detector, and the signal is deposited on the various sub-detectors.

Muon Stations

The LHCb muon stations [45] are designed for the detection of muons. There are a lot of processes with muons produced in the final state with a crucial meaning in the LHCb physics programme, like for example in $B_s \to \mu^+\mu^-$ [46]. The muon system consists of five rectangular muon chambers, placed after one another, where M1 was placed after

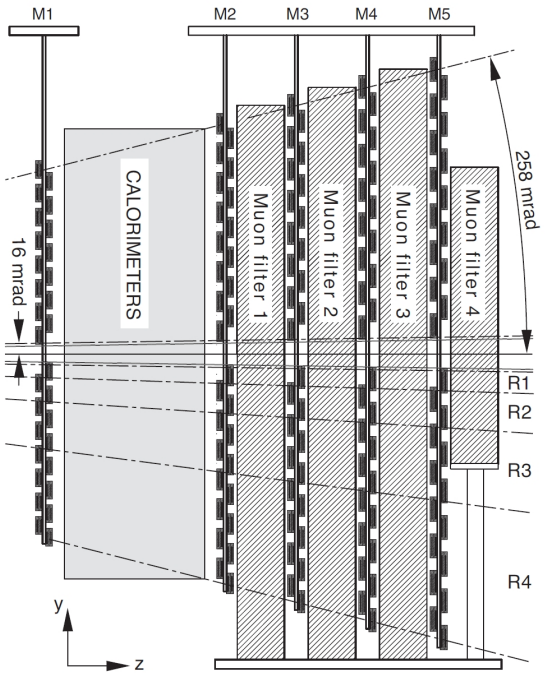


Figure 2.13: A side view of the layout of the muon system [47].

the calorimeters, as shown in Fig. 2.13. These muon stations provide an efficient muon identification system for the use in prominent offline analysis and for fast resonance that can be used in trigger decisions. For a muon to be detected in the muon station, a minimum momentum of 6 GeV/c is required.

2.4 LHCb Upgrade

The measurements obtained from the LHCb spectrometer during Run 1 and 2, although improved results, but there are still some parameters to measure (like CKM angle γ) where experimental precision is below the theoretical value. On the other hand, there are no signs of Physics Beyond the Standard Model (BSM), and one can expect it in the indirect measurements provided by LHCb. To reduce the statistical uncertainties, more data is necessary, which could be obtained by increasing the instantaneous luminosity. The planned nominal instantaneous luminosity during Upgrade I increased by a factor of 5, which was 4×10^{32} to 2×10^{33} cm^{-2} s^{-1}. The main aim of the upgrade is to facilitate recording data in this framework to collect a dataset of at least 50 fb^{-1}. This caused an upgrade of nearly the entire spectrometer. The hardware trigger L0, which utilised data from the calorimeter and muon systems to reduce collision rates to a readout rate

of approximately 1 MHz during Run 1 and 2, had to be eliminated. Consequently, the High-Level Trigger (HLT) was completely reworked to handle the full stream of data [48]. This is where the largest inefficiencies in the entire trigger chain occur; however, in Run 3, the hardware trigger is removed, and the trigger is completely software-based. LHCb will continue data taking with these triggers for Run 3 and Run 4.

There is another upgrade planned called Upgrade II, that would take place after the LHC's Long Shutdown (LS4), and we would begin data taking for Run 5, which is scheduled for 2032. The main aim of this upgrade would be to take advantage from the High Luminosity LHC (HL-LHC) which aims to increase the luminosity to 1.5×10^{34} cm^{-2} s^{-1}, which is 10 times more as compared to what was obtained after Upgrade I [49]. This will have an integrated luminosity of 300 fb^{-1}. It will not only improve the precision of the observable from the flavour physics, but also enlarge the opportunities of the experiment [49]. Fig. 2.14 shows the luminosity trend recorded with the LHCb experiment and how the trend would be improved for the future Runs at LHCb.

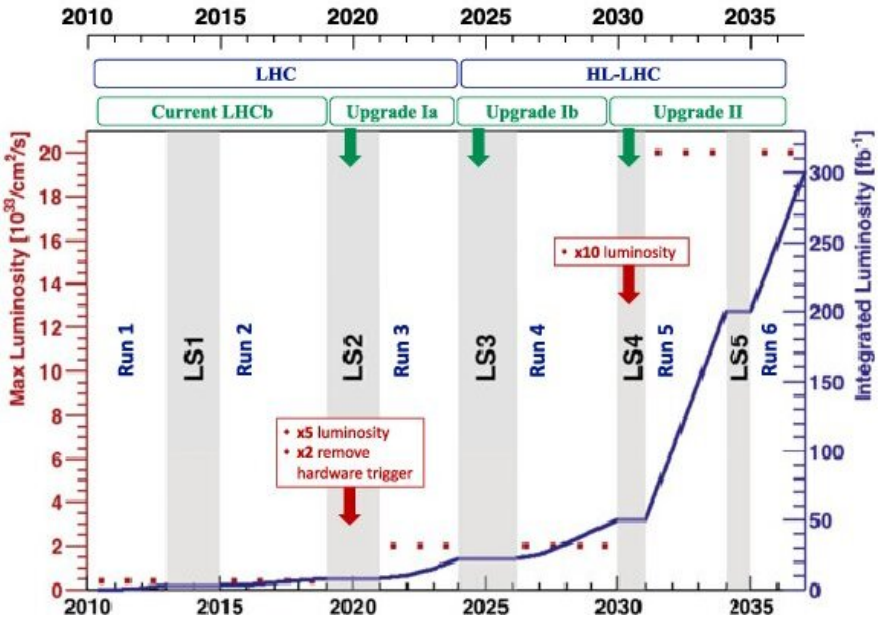


Figure 2.14: Timeline of the current and proposed LHCb upgrades with the expected luminosity [50].

2.5 Trigger

The goal of the LHCb trigger system is to filter the huge amount of incoming data, as we cannot afford to persist multiple terabytes per second of collision information. The designed beam crossing rate inside LHCb is 40 MHz, where to keep only the interesting events, LHCb uses a two-tier trigger system. Due to the high rate of potentially interesting events, the full reconstruction of each event is now used to reduce the data size, rather than relying on partial information as was done in Run 1 and Run 2. The details of the types of tracks and track reconstruction are discussed in Section 2.6.

Until the end of Run 2 in 2018, the first selection step was performed by a hardware trigger, L0, that filtered events with a rate of 1 MHz. The two subsequent High Level Trigger (HLT) software stages thus had to operate at a much lower rate than the nominal LHC collision rate. From 2021 onwards, there was no hardware trigger stage in LHCb's data processing flow to achieve a higher amount of flexibility and efficiency. This poses significant challenges to the HLT processing farm, as it has to operate in real time with a collision rate of 40 million per second. LHCb is the first of the four big experiments at LHC to use a software-only trigger system. Together with the fivefold increase of luminosity, the computational requirements on the HLT computing farm increased by about two orders of magnitude with respect to previous data-taking periods. In contrast to these requirements, the allocated computing budget does not allow for a farm that is orders of magnitude more powerful than the one operated in 2018.

The LHCb trigger system is divided into two stages: High Level Trigger 1 (HLT1) and High Level Trigger 2 (HLT2). The HLT1 reconstruction starts by reconstructing tracks with hit information from the VELO. The VELO is away from the magnetic field, so only straight lines need to be fitted. The track candidates help to reduce the multiplicity in later tracking stages. Based on these candidates and some information about the current beam line position, primary vertices are reconstructed by clustering tracks that have been extrapolated to the beam line. The next HLT1 reconstruction step matches hits in the upstream tracker to a VELO track extrapolation through the first part of the detector. A slightly curved line is fit to the hits in the UT, and the first momentum estimate is extracted from that curvature. Because of the small magnitude of the magnetic field before and in the UT, the momentum estimate has relative uncertainties of about 15 %. Taking into account the magnetic field model and the first momentum estimate, upstream tracks are further extrapolated to the SciFi region, where the track candidates

are matched to SciFi hits. This pushes the momentum estimate to a relative uncertainty of about 0.5 %. A Kalman filter [51] is applied to fit a VELO track candidate, taking into account a momentum estimate from the other tracking stations. This decreases the uncertainty of VELO track parameters. A full Kalman filter application is too expensive for HLT1. These steps conclude the HLT1 upfront reconstruction that serves as input to almost every selection criterion applied in HLT1.

HLT2, on the other hand, takes event input from the disk buffer and aims to perform offline-quality reconstruction and selections to bring the incoming event rate of 1 MHz down to a bandwidth of 10 GB/s. Offline quality reconstruction refers to the most precise calculations, taking into account all calibrations and alignment, even though that might cost more computing resources than simplified calculations in HLT1. HLT2 employs similar reconstruction steps to HLT1 and more. Aside from a high precision trajectory reconstruction, the PID system involving the RICH and both calorimeters helps to form particle hypotheses and neutral particle candidates. Like in HLT1, the output bandwidth in HLT2 is controlled by a set of trigger lines. The average type of HLT2 line, therefore specializes in selecting only a very specific decay structure efficiently, involving requirements on track and vertex topologies as well as PID variables. In HLT2, one mainly differentiates between two types of selection algorithms. The first type is the plain selection that takes a collection of particles and returns a selected collection. These are called "Filters". An example of a Filter is one that collects muons with a high transverse momentum from all the input particles. Algorithms of the second type are called "Combiners". These algorithms combine two or more particles to form a candidate for a decaying particle. Although combiners reconstruct vertices, they will be referred to as selection rather than reconstruction algorithms, because selection criteria with control flow impact are applied at multiple points during the combination process. Fig. 2.15 shows the evolution of trigger schemes at LHCb, starting from Run 1 (left), Run 2 in the center, and the upgrade trigger scheme (right). Run 3 or the upgrade trigger shows the real-time alignment calibration and also the splitting of events into full/turbo streams. Turbo streams are specially designed parts of the data, which allow for fast analysis of processes with high cross-section, like decays of charm hadrons.

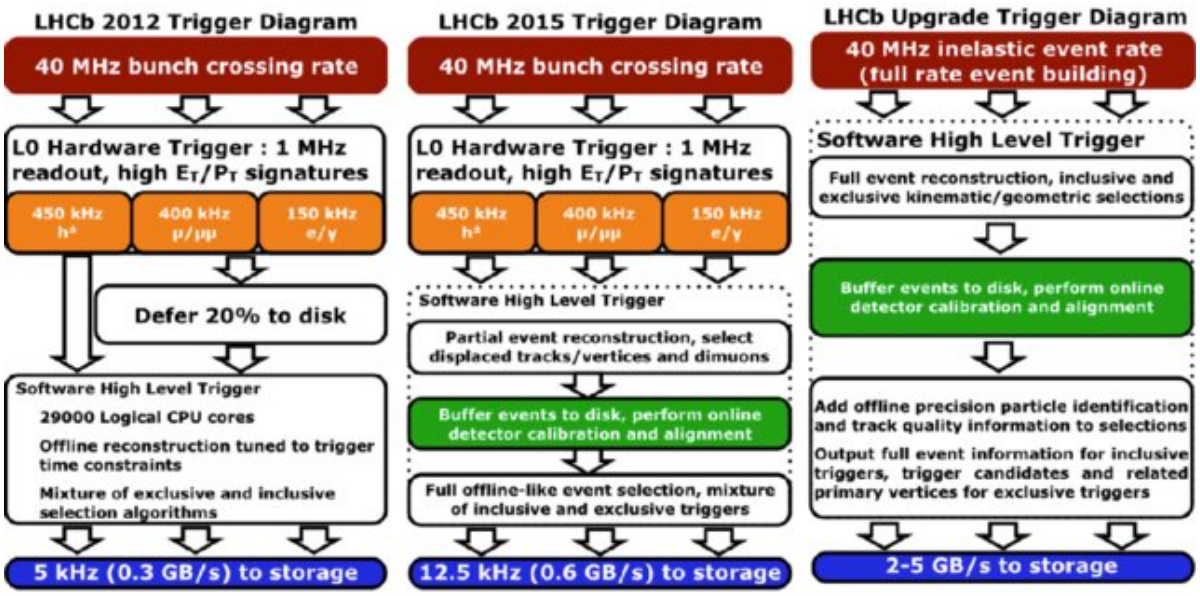


Figure 2.15: LHCb trigger data-flow in Run I (left), Run II (center) and the one foreseen for Run III (right) [48].

2.6 Track reconstruction at LHCb

2.6.1 Track types

The LHCb spectrometer's main task is to record final states originating from the heavy hadrons' decays, and to reconstruct the vertices of their production and decays. The tracking system consists of VELO, UT, and SciFi, but because some particles are quite long-lived, different track types show up in the tracking system:

Long Tracks (L)

These are the tracks that travel across the whole tracking system, starting from VELO up to UT and SciFi stations. Thus, these tracks have the most accurate momentum information of all the track types; therefore, they are used in the majority of the physics analyses at LHCb.

Downstream Tracks (D)

The downstream tracks, on the other hand, leave the hit in UT and SciFi stations, omitting the VELO. They are produced from the charged long-lived particles with displaced origin from the interaction points like K^0 and Λ. The momentum resolution of

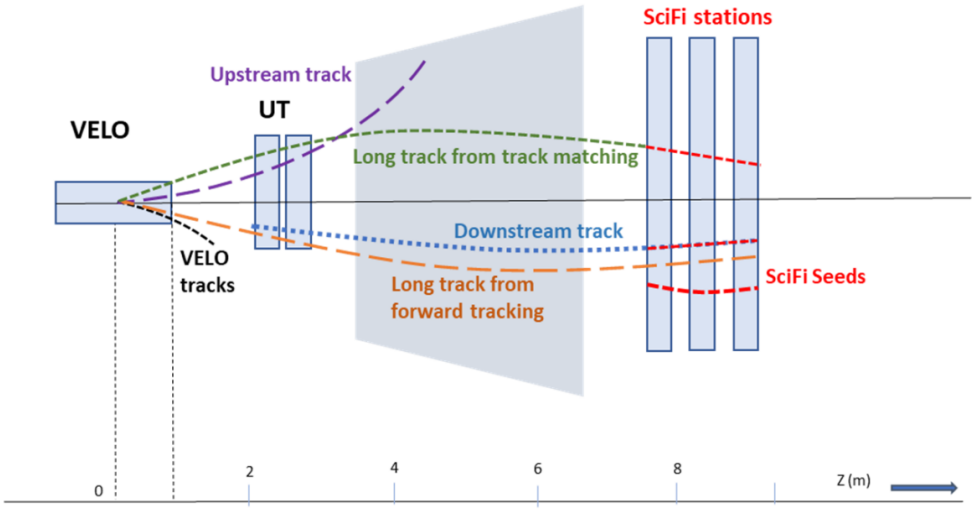


Figure 2.16: A schematic illustration of various track types traversing the tracking system. The magnet is situated in the central part of the drawing. Charged particles are bent in one direction [52].

D tracks is worse than the long tracks, but they play a significant role in many golden channels of B decays, like $B \to K_S^0 J/\psi$ or exotic hadrons with Λ in the final state.

Upstream Tracks

These are the tracks that leave hits in VELO and UT only. These consist of the low-momentum particles that are bent out of the acceptance of the magnetic field. These are used for RICH reconstruction, with the momentum resolution worse than the downstream tracks and long tracks.

VELO Tracks

These are the tracks that leave hits in VELO only. Their momentum cannot be computed, but because of their large polar angles and lack of a magnetic field, they can be used for accurately determining the primary vertices.

T-Tracks

They only appear in the SciFi stations. Their point of origin is either from the material interactions or from the decay of long-lived particles, They are used for RICH2 reconstruction and the alignment of the SciFi stations. Currently, T tracks are considered for

a number of analyses searching for long-lived particles from beyond the SM.

The schematic diagram of the types of tracks at LHCb is shown in Fig. 2.16.

2.6.2 Pattern recognition

There are a number of particles produced after the pp collision, and every charged particle leaves numerous number of hits. Pattern recognition is a method that aims to combine information from the tracking sub-detectors to get information about the track candidates. Different algorithms are used for pattern recognition at LHCb, which are as follows:

VELO seed

The magnetic field inside the VELO is negligible, therefore, the signals inside the VELO are in straight lines. In Run 2, the pattern recognition algorithm first searched for r-z projection using the r-sensors and then added the ϕ-sensor measurement afterwards. However, for Run 3 VELO tracking algorithm is based on an imaging approach using Machine Learning techniques. The new approach leverages Graph Neural Networks (GNNs) and boosted decision trees (BDTs) to improve track reconstruction efficiency, especially in high-occupancy environments.

T-seed

The magnetic field inside the T-stations is small yet present. Therefore, the tracks are more like a parabola in the (x,z) plane. The track candidate is first built by a three-hit combination, for the hits of the x-layers in all three stations. Then the information from the other hits is added if there are hits present in the region around the parabola. At last, the hits from the stereo plane are added. More information on the T-seeding pattern recognition algorithm can be found in [53].

Forward tracking

This is an algorithm that searches particularly for long tracks. The main idea of this algorithm is that one could know the starting and the end point of the track before and after the magnet, then the trajectory in the field could be determined by the equations of motion and the known magnetic field. The VELO seeds are used as an input, since they have a defined position and slope. Track candidates who pass the multiple track quality

requirements, and the best among them is selected, and then finally, the information from the TT hits is also added using selection and clustering procedures.

Track matching

This is an additional method for the long-track reconstruction. This algorithm uses all standalone T tracks and VELO seeds as input, and further propagates through the bending plane of the magnet. The fact that the magnetic field is homogeneous enough that the correct track combinations intersect in the focal plane at a fixed z-position. The difference in the position and slope is used for the compatibility of the VELO-T station track candidate. Finally, the information from the T-hits is added at the end.

Downstream tracking

This is quite similar to the forward tracking algorithm. The starting point is from the T-stations and adds information from UT. There is no information from seeds or hits that is used, since they are already reconstructed as long tracks.

Upstream tracking

This is similar to downstream tracking. Instead of T-station seeds, VELO seeds are used.

Final track selection

The track reconstruction starts by building the VELO and T seeds. The priority is to reconstruct the long tracks for which the forward tracking algorithm and track matching are used. Since there would be a slight chance of saving the same track twice, for this purpose, LHCb uses another algorithm called Clone Killer [54], which uses the technique that if two tracks are clones, then it would keep the track with the most hits assigned. Later, the downstream and the upstream tracking reconstruction is performed, since these are the shortened long tracks, only hits and seeds are used. VELO and T seeds, which are not a part of the long, upstream or downstream tracks, are identified as VELO track or T track.

Performance

The LHCb experiment excelled in particle identification, achieving a 95% efficiency in identifying kaons and a low 5% misidentification rate for pions as kaons. The tracking

system offered a momentum resolution of $\Delta p/p = 0.5\%$ - 1% for momenta between 2 and 200 GeV/c, with excellent time resolution of 50 fs and precise vertexing, characterised by an impact parameter resolution of $15 + 29/p_T$[GeV] μm. The ECAL and HCAL calorimeters provided an energy resolution for ECAL of $\Delta E/E = 1\% + 10\%/\sqrt{E}$[GeV]. This high performance reflects the advanced design and precision of the experiment's tracking and identification systems.

2.7 Radiation damage in silicon sensors

The VELO (Vertex Locator) detector is a silicon pixel detector that is positioned closest to the interaction point within the LHCb experiment. Due to the dense flux of particles traversing through the VELO sensors, it is particularly susceptible to radiation damage. The damage accumulates over time and may finally worsen the performance of the detector, causing the need for replacement.

2.7.1 Silicon detector's principles

Silicon, one of the most abundant elements, forms a crystalline structure through covalent bonding. At absolute zero, it acts as an insulator, as its electrons remain confined to the valence band, separated from the conduction band by an energy gap, as illustrated in Fig. 2.17. As the temperature increases, electrons gain enough energy to transition to the conduction band. The energy gap for silicon at room temperature is approximately 1.12 eV. At room temperature, silicon acts like a semiconductor. Though this value can vary slightly with changes in temperature and environmental conditions.

To improve silicon's conductivity, impurities are introduced through a process known as doping, which produces n-type or p-type semiconductors. When these two types are brought together, a depletion region is formed at the junction, generating an electric field. Applying a reverse bias voltage widens the depletion region, allowing the device to function effectively. When a charged particle passes through the material, it creates electron-hole pairs that are separated by the electric field and subsequently collected by electrodes. For optimal performance, the sensor must be fully depleted to maximize charge collection efficiency. However, care must be taken to avoid exceeding the breakdown voltage, as this can cause permanent damage to the detector [55].

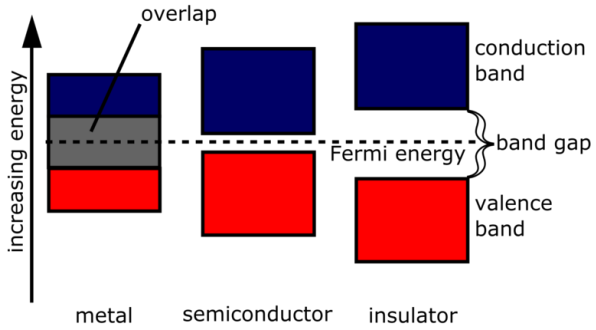


Figure 2.17: A band gap diagram showing the differences between metals, semiconductors, and insulators [56].

2.7.2 Radiation damage and NIEL hypothesis

The impact of radiation on materials, as one of the effects of ageing of the detector, causes so-called *radiation damage*. This occurs when energetic particles collide with atoms, displacing them from their lattice positions, create defects in the material. These collisions lead to energy loss and the creation of defects, degrading the performance of semiconductor devices used in particle detection. The damage caused in the bulk of the sensor is caused by both charged and neutral particles when they lose energy through non-ionising collision with the lattice atoms (silicon or dopant). The recoil atom can migrate through the lattice, causing the dislocation of further atoms. The displacement damage results in permanent changes to the semiconductor material. Created defects influence the energy band structure of the initial, non-irradiated semiconductor.

The extent of displacement damage is directly proportional to the number of particles striking a device per unit area, called fluence ϕ, weighted by the experimental parameter which describes the damage caused by the relevant particle.

The NIEL hypothesis shows that displacement damage scales linearly with the energy transferred during collisions [57]. While this approach does not account for the spatial distribution of defects or annealing processes that may occur post-damage, it provides a foundation for quantifying radiation effects. The displacement damage cross-section, $D(E)$, is a parameter derived from the experiment and is used to calculate the hardness factor, κ, which compares the damage caused by different types of radiation [58].

The hardness factor, κ, is defined as:

$$\kappa = \frac{\int D(E)\phi(E)dE}{D(E_n = 1MeV)\int \phi(E)dE} \tag{2.1}$$

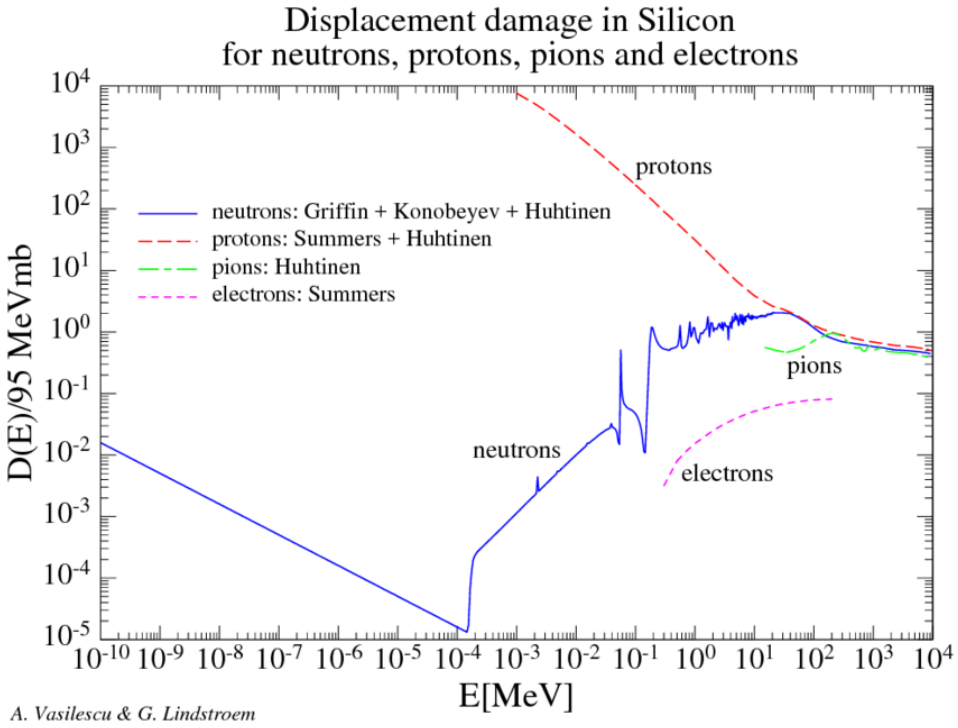


Figure 2.18: Displacement damage functions $D(E)$ normalised to 95 MeVmb for neutrons, protons, pions, and electrons [57].

where $D(E_n = 1 \text{ MeV})$ is normalized to 95 MeVmb. Using κ, the neutron-equivalent fluence can be calculated as:

$$\phi_{eq} = \kappa\phi = \kappa \int \phi(E)dE, \tag{2.2}$$

and is expressed in terms of 1 MeV neutron-equivalent particles/cm^2 [57]. During the data taking, ϕ_{eq} accumulates proportional to the luminosity, finally causing the need for the sensors' replacement.

2.7.3 Radiation damage in the LHCb experiment

In high-energy environments such as the LHC, each pp interaction generates numerous particles across a wide momentum range. Especially in the increased luminosity in upcoming runs, including Run 3 and the High Luminosity LHC (HL-LHC), the particle production rate is expected to be 10 times higher, leading to a significant rise in radiation-induced damage to sub-detectors.

During Run 3, it is expected that the fluence annually will reach levels equivalent to the combined fluence from Runs 1 and 2. Beyond a threshold of 10^{16} n$_{eq}$/cm^2, detector

performance may degrade significantly. To mitigate this, periodic sensor replacement is required, guided by simulations of particle fluence and real-time monitoring of detector conditions [59].

The radiation environment at the LHC is highly heterogeneous, arising from both *pp* collisions and secondary interactions with detector materials. This complex radiation field, comprising charged hadrons, leptons, neutrons, and photons, varies depending on the distance and angle relative to the proton beams and the surrounding materials [60]. Accurate characterisation of this environment relies on Monte Carlo-based simulations, which are essential for evaluating detector performance, planning replacements, and validating radiation damage models. A state-of-the-art cooling system was integrated into the detector modules to counteract radiation-induced effects and manage heat dissipation.

For the radiation damage studies, the LHCb experiment uses FLUKA [61] simulation to evaluate the radiation damage to its detectors and electronic components. The FLUKA package contains a fully integrated code of hadron-hadron, hadron-nucleon, hadron-photon, etc. interactions. Its comprehensive modelling capabilities enable precise simulations of particle interactions and energy deposition within the LHCb environment, facilitating accurate predictions of radiation levels and their potential impact on materials and electronics. These simulations are crucial for designing effective shielding and implementing protective measures, ensuring the reliability of the LHCb detector systems. The LHCb experiment also uses Geant4 [62] for modelling radiation damage by simulating the interaction of high-energy particles with the detector material over time. Geant4 estimates the radiation dose and particle fluence experienced during LHC operations, which helps in assessing long-term detector performance and planning maintenance accordingly.

Chapter 3

Monte Carlo Simulations

To quote: *"The Monte Carlo simulation has become the major means of visualisation of not only detector performance but also of physics phenomena. So far, so good. But it often happens that the physics simulations provided by the Monte Carlo generators carry the authority of data itself. They look like data and feel like data, and if one is not careful, they are accepted as if they were data" - J.D.Bjorken (1992).*

3.1 Physics model for pp interactions

Collisions of high-energy hadrons are evidence that they contain a spectrum of interactions, such as strong, electromagnetic, and weak. Since protons are electrically charged objects with colored internal structure, QCD plays a dominant role in the study of strong interactions between quarks. It might seem surprising that the properties of a proton, a colourless particle, can be deduced by a local theory based on the gauge group SU(3), with the perturbative calculations applied to the colored partons. These calculations are based on a method called the lattice QCD; however, possible but computationally challenging. Physicists at LHC are interested in high-energy proton-proton collisions, the vast variety of different possible elementary interactions between partons, and the formation of the spectrum of final states are the subject of a great deal of physics models embedded in the so-called event generators. The whole process of particles' collision, along with the principles of physics governing the probability of interaction, is computed in simulations with the use of Monte Carlo (MC) methods.

The Monte Carlo techniques are a set of algorithms that employ random sampling to provide approximations or simulations for complex mathematical, statistical, or physical

problems. It operates on the fundamental concept of extracting a substantial number of samples from a distribution that characterises the numerical problem at hand. As an illustrative example, the estimation of π can be achieved by generating N points uniformly distributed between (0,1) in the coordinates of a square, selecting only those falling within a distance of 1 from the origin. By iteratively performing this process a significant number of times (e.g., 10000), averaging the accepted values produces an approximation of $\pi/4$.

High-energy hadron collisions provide a unique opportunity to study a rich spectrum of interactions governed by the strong, electromagnetic, and weak forces. Within particle physics, the collision between two particles is termed an "event," producing a multitude of outgoing particles that can be detected by experimental apparatus. These events adhere to conservation laws, ensuring that the total energy, momentum, and quantum numbers of the final particles are conserved. Despite these principles, the specific properties and number of particles produced in each event can vary significantly. At the Large Hadron Collider (LHC), physicists investigate high-energy proton-proton collisions, which involve a wide array of elementary interactions between partons (quarks and gluons) and result in diverse final states, enriching our understanding of fundamental forces and particle interactions. To study and predict these complex processes, physicists rely on simulation tools known as event generators. These algorithms replicate complex interactions that occur during particle collisions, producing simulated events. Popular Monte Carlo event generators such as Pythia [63], Herwig [64], and Sherpa [65] are continually refined to enhance their accuracy and alignment with experimental observations.

In LHCb, proton-proton event samples are generated using predominantly Pythia [66], a widely-used general-purpose event generator. This chapter provides a concise introduction to both Pythia and Herwig, followed by a comparative analysis of their performance in calculating the multiplicity of particles in minimum bias events. This comparison is crucial for understanding the differences in simulated particle flux through the LHCb VELO sub-detector. To validate the predictions from Pythia and Herwig, experimental data from LHCb is used. This would impact the fluence estimates and predictions of radiation damage, which is discussed in detail in Chapter 6. Fluence studies are essential for estimating the detector's operational lifetime and planning for maintenance or upgrades, which is why accurate modelling through event generators is important. By linking the outputs of event generators with fluence and radiation studies, LHCb ensures that both the theoretical predictions and the physical durability of the

detectors are aligned for optimal performance throughout the experiment's lifetime.

3.2 Proton-proton interactions

The strong force, which governs hadron interactions, is dependent on the distance between particles and becomes weaker at shorter ranges. In high-energy hadron collisions, it is expected that the interactions between free quarks and gluons will dominate, with perturbative QCD (pQCD) offering predictions for these interactions. However, it remains an open question whether pQCD fully accounts for experimental data, and how to extend its applicability to include phenomena at longer distances and lower energy scales. Additionally, experiments like Deep Inelastic Scattering (DIS) have revealed that at higher momentum transfers, the proton exhibits a more complex internal structure, leading to a greater number of potential interactions [67].

In high-energy proton-proton collisions, only a fraction of the proton's energy is typically involved in the hard interactions between quarks, anti-quarks, and gluons, while the remainder contributes to the production of additional hadrons. It is not possible to predict exactly which parton will be involved in a hard interaction, but the probability that a parton carries a fraction x of the proton's momentum can be parameterised using Parton Distribution Functions (PDFs), which capture non-perturbative effects [59]. To model these interactions, the factorisation principle is applied, where perturbative calculations are performed for point-like partons represented by the PDFs.

The differential cross-section that two hadrons h_1 and h_2 containing partons a and b collide and produce c and d states is shown as follows:

$$d\sigma_{h_1h_2\to cd} = \int_0^1 dx_1 dx_2 \sum_{a,b} f_{a/h_1}(x_1,\mu_F^2) f_{b/h_2}(x_2,\mu_F^2) d\hat{\sigma}_{ab\to cd}(\mu_F^2,\mu_R^2). \tag{3.1}$$

Here the partons $a, b \in \{q, \bar{q}, g\}$ originate from hadrons h_1h_2;

- $f_{a/h_1}, f_{b/h_2}$ are the distributions of partons a, b in hadrons h_1h_2 (PDF).
- the four momenta of the partons are $p_a^\mu = x_1 p_{h_1}^\mu$, $p_b^\mu = x_2 p_{h_2}^\mu$, where x_i correspond to the fractions of four momenta of parton i taken from the h_i.
- μ_F is the factorisation scale (hard interaction scale) parameters. The PDFs depend on the factorisation scale parameter μ_F, which is chosen and fixed.

- $d\hat{\sigma}_{ab\to cd}(\mu_F^2, \mu_R^2)$ is a differential cross-section of the elementary hard sub-process between partons $ab \to cd$, which can be calculated in QCD perturbative theory.

Since the PDFs are non-perturbative and cannot be calculated directly, they must be extracted from global fits to experimental data.

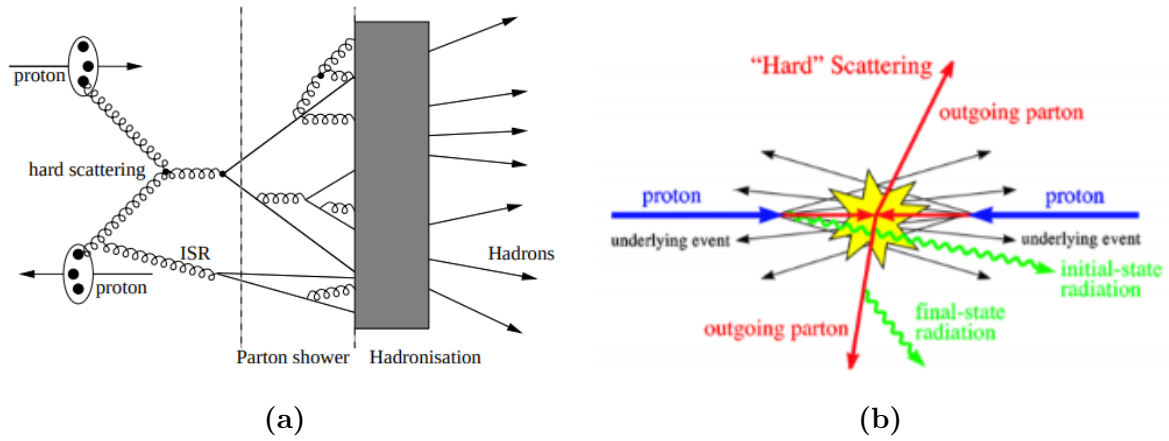
3.3 Total proton-proton cross-section

Protons are electrically charged composite particles that interact via internal colour-charged constituents. Consequently, the total proton-proton cross-section can be divided into two primary components: elastic and inelastic. The inelastic component is further categorised into diffractive interactions (single or double) and the remaining portion, commonly referred to as "inelastic non-diffractive" or simply inelastic. This breakdown of the total cross-section is represented in Eq. 3.2.

$$\sigma_{tot}^{ab}(s) = \sigma_{el}^{ab}(s) + \sigma_{sd(Xb)}^{ab}(s) + \sigma_{sd(aX)}^{ab}(s) + \sigma_{dd}^{ab}(s) + \sigma_{nd}^{ab}(s). \tag{3.2}$$

The first term refers to elastic processes in which protons interact via t-channel exchange, represented as $ab \to ab$, 'sd' refers to single diffractive events, represented as $ab \to aX$ or Xb, in which one proton remians unchanged while the other diffracts, producing other particles. 'dd' refers to the double diffractive processes, depicted as $ab \to$ XX and 'nd' refers to the non-diffractive processes. This classification has an origin because of different types of processes behind the interaction and the physics model that has to be implemented in simulations.

The typical inelastic pp interactions start with at least one hard interaction between partons. This is accompanied by the emission of gluons, production of additional quarks and anti-quarks, generally named as parton shower. These objects, together with the remaining partons from protons, eventually hadronise into final hadrons, in a process called hadronisation. At the LHC energies, it is also possible that two hard interactions might occur, which is known as multiparton interactions (MPI). The schematic diagram showing a pp interaction is shown in Fig. 3.1a.



(a) **(b)**

Figure 3.1: Schematic diagram of a *pp* interaction, modelling initial and final state parton showers, a hard scattering process and hadronisation [68] in an event generator (a) and scattering of an event in *pp* interactions [69] in (b).

3.4 Event generation in PYTHIA

PYTHIA is a general-purpose event generator that contains multiple models for several physics processes like: hard and soft interactions, multiparton interactions (MPIs), PDFs, initial and final state parton showers, fragmentation, decays, and hadronisation model [63]. For the description and implementation of these models, PYTHIA uses many phenomenological parameters that define the applicability of the models. These are included as the default parameter sets in PYTHIA and their values are determined from experimental data.

The majority of processes in PYTHIA are governed by QCD. The perturbative and the non-perturbative parts of the collision could be separated by the hadronisation scale $Q \sim 1$ GeV. The perturbative regime is in the hard scattering region, which is characterised by the large momentum transfers, whereas the non-perturbative regime consists of the physics at low momentum. These could be separated from one another using the factorisation hypothesis, which is shown in Eq. 3.1. A typical *pp* collision can be simulated in the following ordered way:

- Two beams collide at high centre-of-mass energy, producing a hard scattering event involving large momentum transfer between partons. The dynamics of the particles in the beam is characterised by the PDF.

- The initial-state parton showers are created as a sequence of branching by the

shower initiator from each beam.

- One parton from each of the incoming beams interacts via hard scatterings, which produces outgoing partons. The outgoing partons might act as the final-state radiation.
- Due to the confinement, the outgoing partons could not leave the interaction region but instead form colourless hadrons (fragmentation), which is the process where the partons convert to the final hadrons.
- The last step takes care of the decay of unstable particles.

A brief explanation of some physics models implemented in PYTHIA is described below:

3.4.1 Parton showers

The structure of the initial and final state showers is given by $a \to bc$, where a is the mother decaying to two daughters b and c, including emission of quarks, gluons, and photons. This equation further adds the following branches in PYTHIA: $q \to qg$, $q \to q\gamma$, $g \to gg$, $g \to q\bar{q}$, $l \to \gamma$. The branching kinematics is given by Q^2, which is the transverse momentum scale, and other variables that give the sharing of the energy and momentum of the mother (a) between the daughters (b and c). Each process is described by splitting $P_{a\to bc}$, which gives the probability of parton a to branch into b and c. PYTHIA handles the initial state showers itself but uses packages like JETSET [70] for handling the final state showers [63].

3.4.2 Hadronization

QCD works at short distances and fails at larger distances. This is when quarks and gluons experience confinement and combine to form hadrons, in a complex process called "hadronisation." Understanding the hadronisation process, starting from the QCD Lagrangian, remains a challenge. Consequently, various phenomenological models have been developed in the absence of a first-principles understanding. Three primary categories, namely, string fragmentation (SF), independent fragmentation (IF), and cluster fragmentation (CF), are commonly recognised, although numerous variations and hybrids exist. Given their nature as models, none of these approaches can assert absolute correctness, some may be more well-formulated than others [71]. The SF model is widely

employed in PYTHIA. This program utilises the phenomenological Lund string model [72], where confinement is achieved by elongating strings between the outgoing q and $\bar{q}$ partons. As the q and $\bar{q}$ move in opposite directions, stretching the string, the potential energy stored in the string increases and reaches a value hence breaking the string and generating a new pair $q_1\bar{q}_1$ from the vacuum. Consequently, the system separates into two color-singlet systems: $q\bar{q}_1$ and $q_1\bar{q}$. In the Lund string model, this break-up process continues as long as the invariant mass of the string pieces exceeds the on-shell mass of a hadron. Each produced hadron corresponds to a small section of string with a quark and an anti-quark at its ends. Therefore, the fragmentation process is described in terms of branching and can be calculated iteratively [63].

3.4.3 Multiparton Interactions (MPI)

In PYTHIA, MPI plays a critical role in simulating hadron-hadron collisions, such as those occurring in pp interactions at LHC. Because protons are composed of partons, a single collision can involve multiple parton-parton scatterings. The MPI model in PYTHIA captures this complexity by allowing several interactions to occur within one event, contributing significantly to the so-called underlying event — the soft activity not directly related to the hardest scattering. It also incorporates the impact parameter dependence, meaning the likelihood of these interactions varies with how central the collision is. Additionally, features like colour reconnection are included to more accurately simulate how final-state partons hadronise into observable particles [63].

3.5 Event generation in HERWIG

HERWIG is a general-purpose event generator for high-energy hadronic processes [64]. The physics models implemented in HERWIG simulate hard scattering processes involving lepton-lepton, lepton-hadron, and hadron-hadron collisions, as well as soft hadron-hadron collisions. It incorporates colour coherence of partons and handles heavy flavour productions and decays. Additionally, it models QCD jet evolution and accounts for correlations within and between jets arising from interference and gluon polarisation. The package also includes cluster hadronisation of jets, encompassing non-perturbative hadronic events, and utilises a cluster model for soft and underlying hadronic events. However, the theoretical justification of the QCD Monte Carlo simulations lies in the factorisation theorem

for the hard processes mentioned in Eq. 3.1. The processes involved can be divided into a number of stages corresponding to the time and distance scales. HERWIG has a large library of QCD, electroweak, and supersymmetric elementary processes. Generally, this is computed as leading order perturbation theory, which is quite similar to the one in PYTHIA. The energy scale of the hard process Q^2 sets the initial conditions for the production of QCD radiation in the initial and final state parton showers [64]. A brief explanation of some physics models implemented in HERWIG is described below:

3.5.1 Parton showers

The complexity of the initial state parton showers is more than the final state parton showers. There are a few key parameters which are important: hard scale (Q) and x, which refers to the Bjorken variable as discussed in Eq. 3.1. For any value of x, the initial state emission process would factorise and can be described as a coherent branching process suitable for Monte Carlo simulations, incorporating angular constraints between the incoming hadron and emitted partons. On the other hand, for larger scales of x, the coherent branching algorithm accurately sums leading and next-to-leading contributions to help determine the QCD scale in simulations and the parameter Λ_{MS} (QCD scale parameter). Although HERWIG hasn't fully integrated all properties of small x branching, its backwards evolution algorithm ensures smooth integration with valence parton distributions. Matrix-element corrections improve the accuracy of initial-state parton showers in specific processes. To prevent double-counting, radiation with transverse momenta greater than the hard process scale is also taken into account. Initial-state emission can be disabled, limiting radiation to force splitting of non-valence partons [73].

The generation of final-state radiations in HERWIG is through a coherent branching algorithm which is used to simulate the multiple emissions of quarks and gluons in a manner that respects the principles of QCD, such as colour coherence and soft gluon radiation, and has the following key features:

- The energy among the partons is distributed according to Dokshizer Gribov Lipatov Altarelli Parisi (DGLAP) splitting functions [74]. The DGLAP splitting functions describe the probability of a parton splitting into two other partons as a function of the fraction of momentum carried by each resulting parton.

- The available phase space is constrained to an angular-ordered region, accounting for interference effects. This ensures that the angle between emitted partons decreases

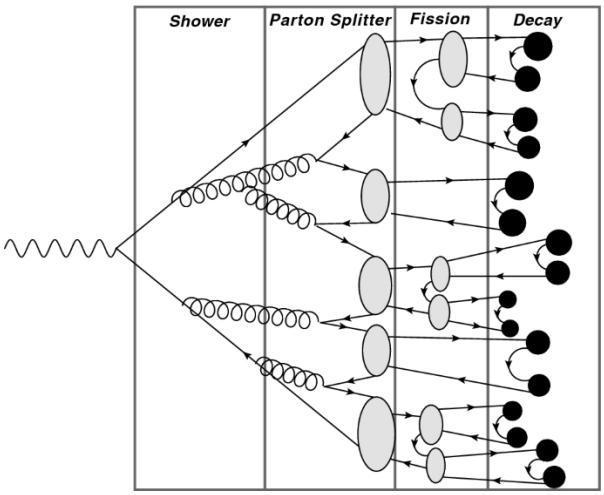


Figure 3.2: A simplified event with the major stages of hadronization in HERWIG [75].

with each successive branching.

- In each branching, the scale of the strong coupling constant α_s is set by the relative transverse momentum of the two daughter partons [73].

3.5.2 Hadronization

The simulation of the general hard process in hadron-hadron collisions consists of a representation of the incoming partons as constituents of the incident hadrons, the conversion of emitted partons into outgoing hadrons, and the underlying soft event associated with spectator partons. The first uses the non-perturbative parton distribution functions; however, for the formation of hadrons and the underlying events, HERWIG uses the pre-confinement property to implement a simple hadronisation model which is independent of hard processes and energy. After the perturbative showering, all the outgoing gluons are split non-perturbatively into light quark-anti-quark or di-quark-anti-diquark pairs, forming jets. These jets are then used to create colour-singlet clusters, which have low mass and size distributions due to pre-confinement. The clusters fragment into hadrons, with lighter clusters representing single hadrons and heavier clusters decaying isotropically into hadron pairs. Some clusters undergo iterative fission until they fall below a specified mass threshold. The parameters controlling these processes ensure accurate modelling of particle production, including high-p_T and heavy particles, while maintaining sensitivity to the details of cluster decay and fission [73].

Fig. 3.2 provides a schematic representation of an event, emphasising the final

state. HERWIG initiates a parton shower, resulting in the production of numerous soft and collinear partons. When the parton shower reaches the GeV scale, the simulation transitions to the hadronisation phase using the cluster model [75].

3.5.3 Multiparton Interactions (MPI)

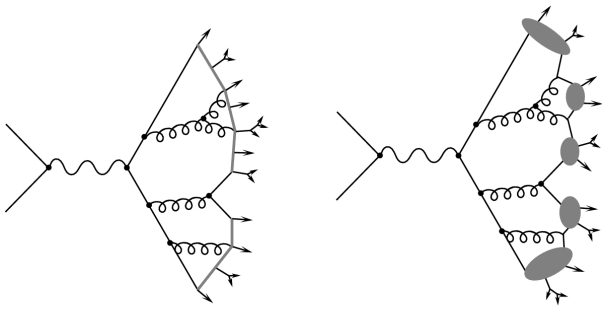
In HERWIG, each event is triggered by a one hard interaction. For minimum-bias events, a dummy process is set up, where two quarks with zero transverse momentum are pulled out of the proton, allowing secondary hard and soft scatters to model a minimum-bias event. The model works well when hard contributions dominate, but when applied to fiducial measurements, where these cuts are loosened, the description of minimum-bias events with HERWIG is bound to fail. The reason is that, in particular, the model for soft interactions is very much ad hoc. It will give the production of soft particles in a way that the "turn-on regions" in the Underlying Event (UE) measurements are well described, but not the correlations among them or with other hard particles. So, the soft model is limited to describing the average soft activity that accompanies a hard event. This failure is more visible when the model for minimum-bias events is applied to observables that have prominent contributions from diffractive events [76].

The comparison between the two event generators can be approached in two distinct ways: the first involves an experiment-independent setup with multiple configurations, while the second involves an experiment-dependent setup, which is LHCb. Both approaches are discussed in detail in the sections 3.6 and 4.9 respectively.

3.6 Comparison of models in PYTHIA and HERWIG

The major difference between PYTHIA and HERWIG lies primarily in the hadronisation: the string model [77,78] in the first and the cluster model [73,79] in the second, respectively. The details of these are mentioned in the section 3.4 and 3.5, respectively. Fig. 3.3 shows the difference in the hadronisation model of the two event generators.

Monte Carlo event generators play an important role in understanding the physics when compared to data. The Monte Carlo models are dependent upon various parameters that correspond to different physics phenomena, which are reported in several studies [81–83] and also discussed for the two general-purpose event generators in the sections 3.4 and 3.5. These parameters could be calibrated or tuned to better describe the data,



(a) String hadronization (b) Cluster hadronization

Figure 3.3: The figure illustrates the differences between the string hadronisation model used in PYTHIA (a) and the cluster hadronisation model used in HERWIG (b) [80].

which are discussed particularly for PYTHIA in Chapter 4. This section describes the comparison of PYTHIA and HERWIG using the pp collisions at $\sqrt{s} = 7$ TeV, and 13 TeV, generating a million events for each MC event generator. The default input parameters used for the event generation for PYTHIA and HERWIG are listed in Table 3.1 and Table 3.2, respectively.

The primary motivation for these comparisons is to enhance our understanding of the particle production mechanism by investigating how a slight variation in the values of parameters can influence the model predictions. Secondly, the comparisons against data are to determine which event generator more accurately describes LHCb data, particularly in the context of multi-parton interactions (MPI). Additionally, the generators can predict the production rates of various particles, such as pions, kaons, protons, and neutrons, which significantly contribute to particle fluence. These predictions are significant since they directly impact estimates of radiation damage within the detector. Furthermore, linking the outputs of event generators to fluence and radiation studies enables LHCb to ensure that both the theoretical predictions and the physical durability of the detectors are aligned for optimal performance throughout the experiment's lifetime.

The PDF settings are crucial in event generators because they determine how the partons are distributed inside protons, which affects the simulated physics. ATLAS and CMS use the CP5 tune with the NNPDF3.1 LO PDF set for PYTHIA. This tune particularly provides a better description for underlying event observables at LHC energies, [84,85]. LHCb, on the other hand, uses CT09MCS for PYTHIA, which is particularly set for forward physics studies. "MCS" stands for modified charm strangeness tailored to match LHCb's kinematic acceptance. For HERWIG CT14NLO provides a good balance between precision and consistency with theoretical calculations, making it suitable for

Parameter	Value
version	PYTHIA8.306
PDF	CT09MCS
softQCD	on
MPI	on
alphaS	0.130
pT0Ref	2.28
ϵ	0.215

Table 3.1: PYTHIA parameters settings.

Parameter	Value
version	HERWIG7.3
PDF	CT14nlo
softQCD	on
MPI	on
DLmode	2
pTmin	3.1
ϵ	0.21

Table 3.2: HERWIG parameters settings.

LHCb's Herwig-based simulations.

The DLmode parameter in HERWIG describes the choice of Donnachie-Landshoff parametrisation for the total cross section [86]. The value set to 2 incorporates a two-pomeron model for a more refined treatment of diffractive scattering.

The other parameters that control the MPI are p_{Ref}^{T0}, ϵ and α_s. These parameters collectively govern the behaviour of the system and shape the distributions of the physical quantities being studied. Eq. 3.1 is divergent at low momentum transfers, and hence it could be regularized by the introduction of threshold parameter p_{T0} as: $1/p_T^4 \rightarrow 1/(p_T^2 + p_{T0}^2)^2$. This approach supports the fact that at low p_T, the partons inside the proton are screened by one another, and at high values, the cross-section is small but nonzero. One also needs to include the centre-of-mass energy dependency, since at higher energies partons are probed at smaller x, fraction of momentum carried away by the parton, where the parton density increases and the distance of the colour screening decreases:

$$p_{T0} = p_{T0}^{ref} \left(\frac{\sqrt{s}}{\sqrt{s_0}} \right)^{\epsilon} \tag{3.3}$$

The parameter $\sqrt{s_0}$ is given at a reference energy, p_{T0}^{ref} is p_{T0} at $\sqrt{s_0}$, p_{T0} is a parameter that acts like a barrier between the hard and soft interactions, ϵ which defines the energy rescaling pace which needs to be tuned to experimental data. On the other hand, the strong coupling constant α_s also becomes divergent at low momentum and has to be regulated by a cut-off parameter.

ATLAS and CMS generally use HERWIG 7.1 event generator with the UE-EE-5 tune for simulating underlying events and MPI, where "UE" stands for underlying events and "EE-5" refers to the energy extrapolation method used in this tune. The parameters used

in HERWIG for MPI are chosen to match the experimental data from ATLAS, including distributions related to the underlying event and the hard scatter. The plots are generated for 1 million events in pp collisions for the default configurations at $\sqrt{s} = 7$ TeV and 13 TeV.

The average number of particles produced in an event for both the default event generators at $\sqrt{s} = 7$ and 13 TeV is shown in Tables 3.3 and 3.4, with the statistical uncertainties, respectively.

	N_π		N_K		N_p		N_{ch}
PYTHIA Default	63.35 ± 0.07	77.1%	8.51 ± 0.01	10.3%	10.21 ± 0.01	12.4%	82.08 ± 0.09
HERWIG Default	54.78 ± 0.04	81.9%	6.80 ± 0.01	10.17%	5.27 ± 0.01	7.86%	66.86 ± 0.06

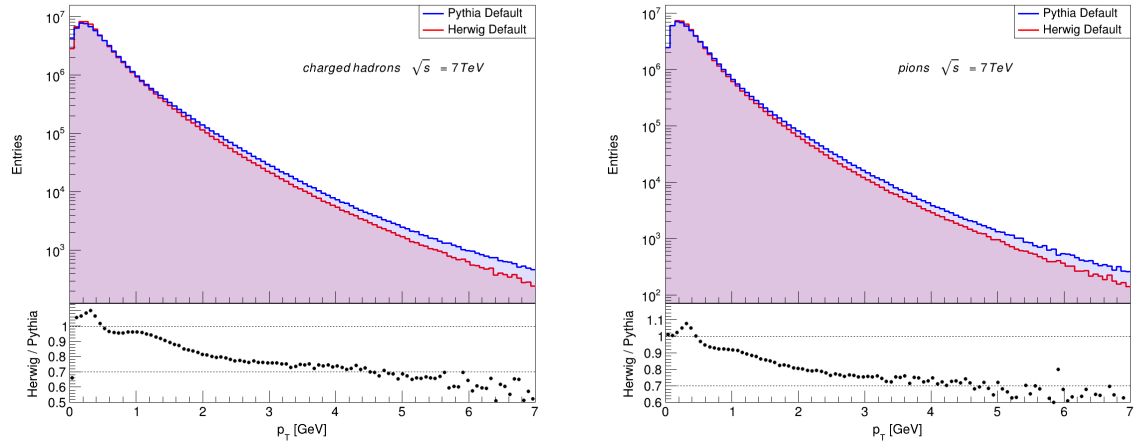
Table 3.3: Average number of charged hadrons in an event for the default settings of PYTHIA and HERWIG at $\sqrt{s} = 7$ TeV with the statistical uncertainties where the percentage indicates its fraction relative to the total average number of charged hadrons N_{ch}.

	N_π		N_K		N_p		N_{ch}
PYTHIA Default	71.76 ± 0.08	77.5%	9.69 ± 0.01	10.47%	11.09 ± 0.01	11.98%	92.55 ± 0.10
HERWIG Default	75.13 ± 0.07	80.9%	9.80 ± 0.01	10.55%	7.89 ± 0.01	8.5%	92.81 ± 0.08

Table 3.4: Average number of charged hadrons in an event for the default settings of PYTHIA and HERWIG at $\sqrt{s} = 13$ TeV with the statistical uncertainties where the percentage indicates its fraction relative to the total average number of charged hadrons N_{ch}.

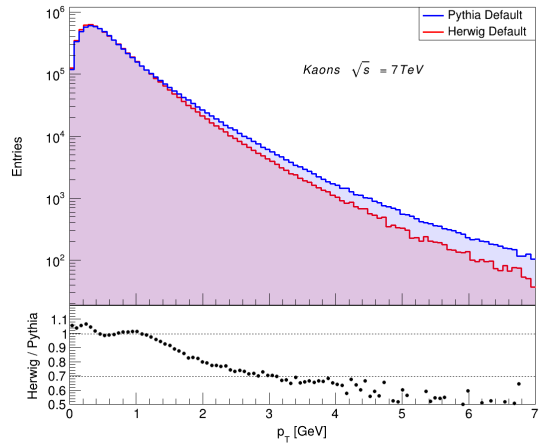
Using the parameters specified in Tables 3.1 and 3.2 for PYTHIA and HERWIG, the following plots are generated for the default configurations.

Fig. 3.4 illustrates the distribution of transverse momentum of charged hadrons produced in pp collisions at $\sqrt{s} = 7$ TeV in (a), while (b), (c), and (d) represent the distribution of pions, kaons, and protons, respectively. The ratio of PYTHIA to HERWIG is shown in the lower panel of each plot, indicating that the ratio subceeds unity at higher momentum scales for the charged hadrons, which is quite evident also in the pions and kaons plots showing that PYTHIA produces more particles at higher momentum values as compared to HERWIG. The case is opposite for protons; the distribution is above unity, indicating that more low-momentum protons are produced by HERWIG than PYTHIA. A similar trend is also observed at 13 TeV, as shown in Fig. 3.5.

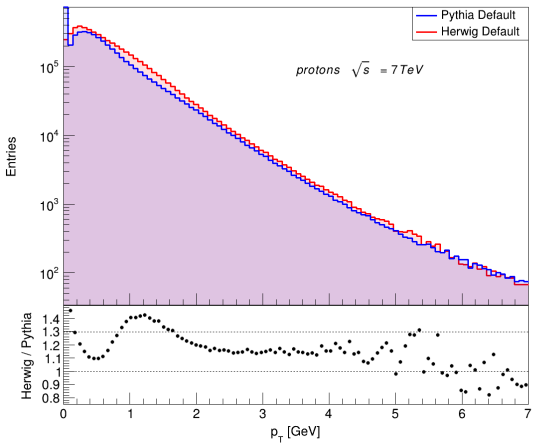


(a)

(b)

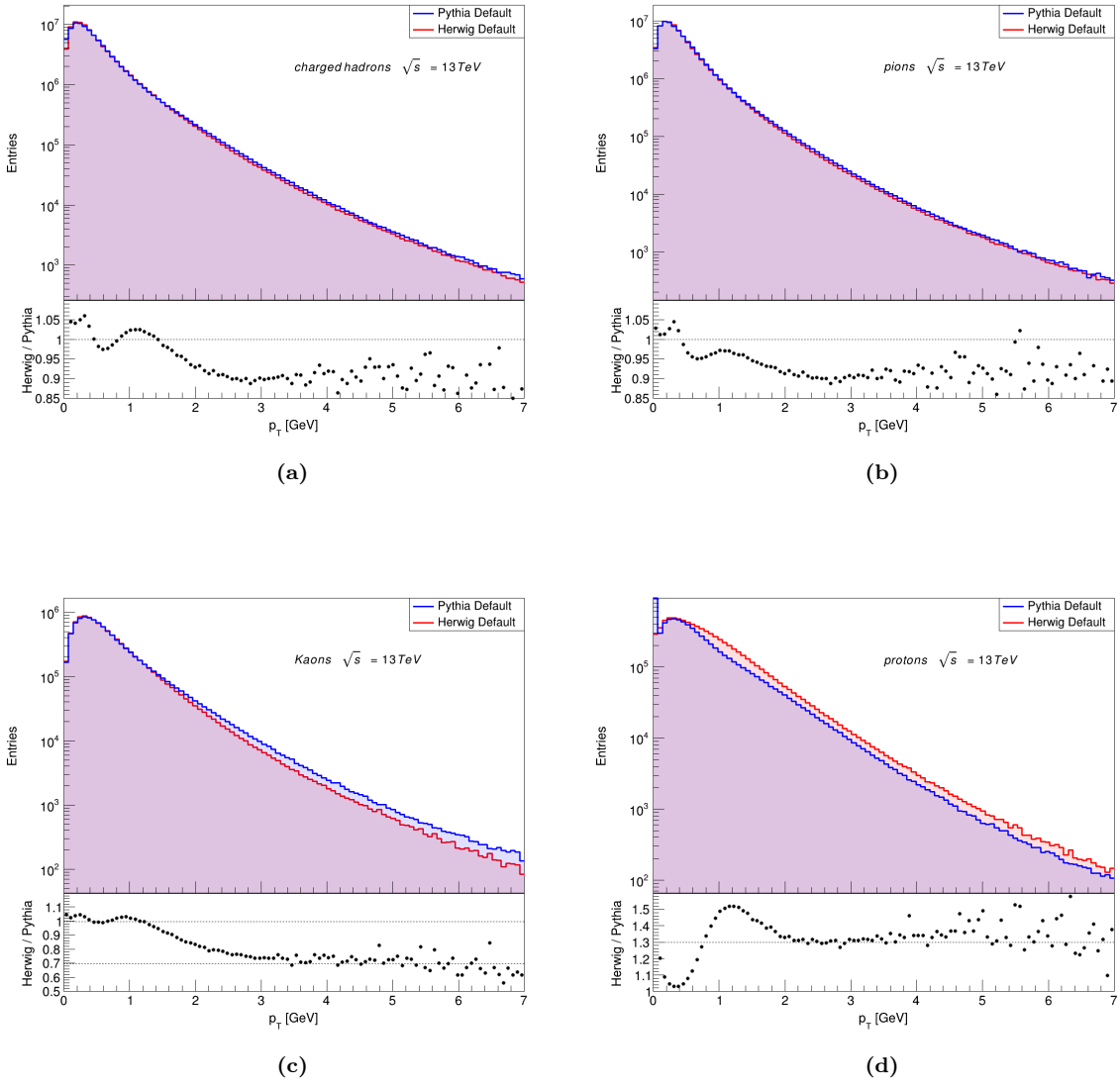


(c)



(d)

Figure 3.4: Comparison of PYTHIA and HERWIG transverse momentum distribution of charged hadrons (a), pions (b), kaons (c), and protons (d) for pp collisions at $\sqrt{s} = 7$ TeV.



(a) **(b)**

(c) **(d)**

Figure 3.5: Comparison of PYTHIA and HERWIG transverse momentum distribution of charged hadrons (a), pions (b), kaons (c), and protons (d) for pp collisions at $\sqrt{s} = 13$ TeV.

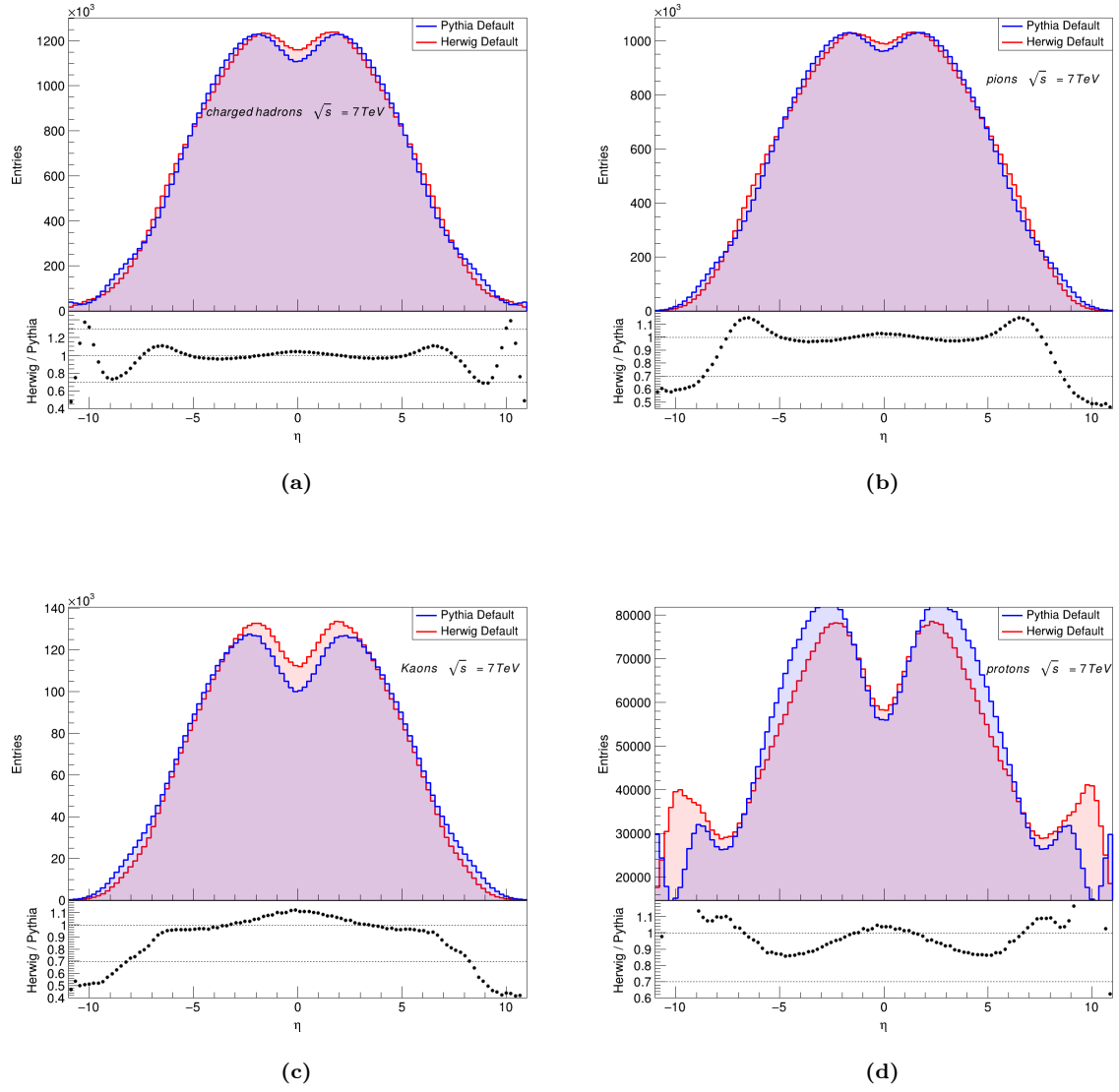


Figure 3.6: Comparison of PYTHIA and HERWIG η distribution of charged hadrons (a), pions (b), kaons (c), and protons (d) for pp collisions at $\sqrt{s} = 7$ TeV.

Fig. 3.6 shows the comparison of PYTHIA and HERWIG at 7 TeV for the pseudorapidity distribution, η of charged hadrons (a), pions (b), kaons (c), and protons (d). PYTHIA and HERWIG predictions show small differences in the forward regions of $|\eta| > 5$ and slight variations near the peak at $\eta = 0$. These differences are likely due to the distinct underlying physics model in hadronisation and parton showering, or to the tuning of model parameters. Since the majority of the particles in the charged distribution are pions, they follow a similar trend. As could be seen from the ratio panel of pions, there is a good general agreement for both generators in the -3< η <3 region. For the case of kaons, the central region is dominated by high-energy particle production from the hardest

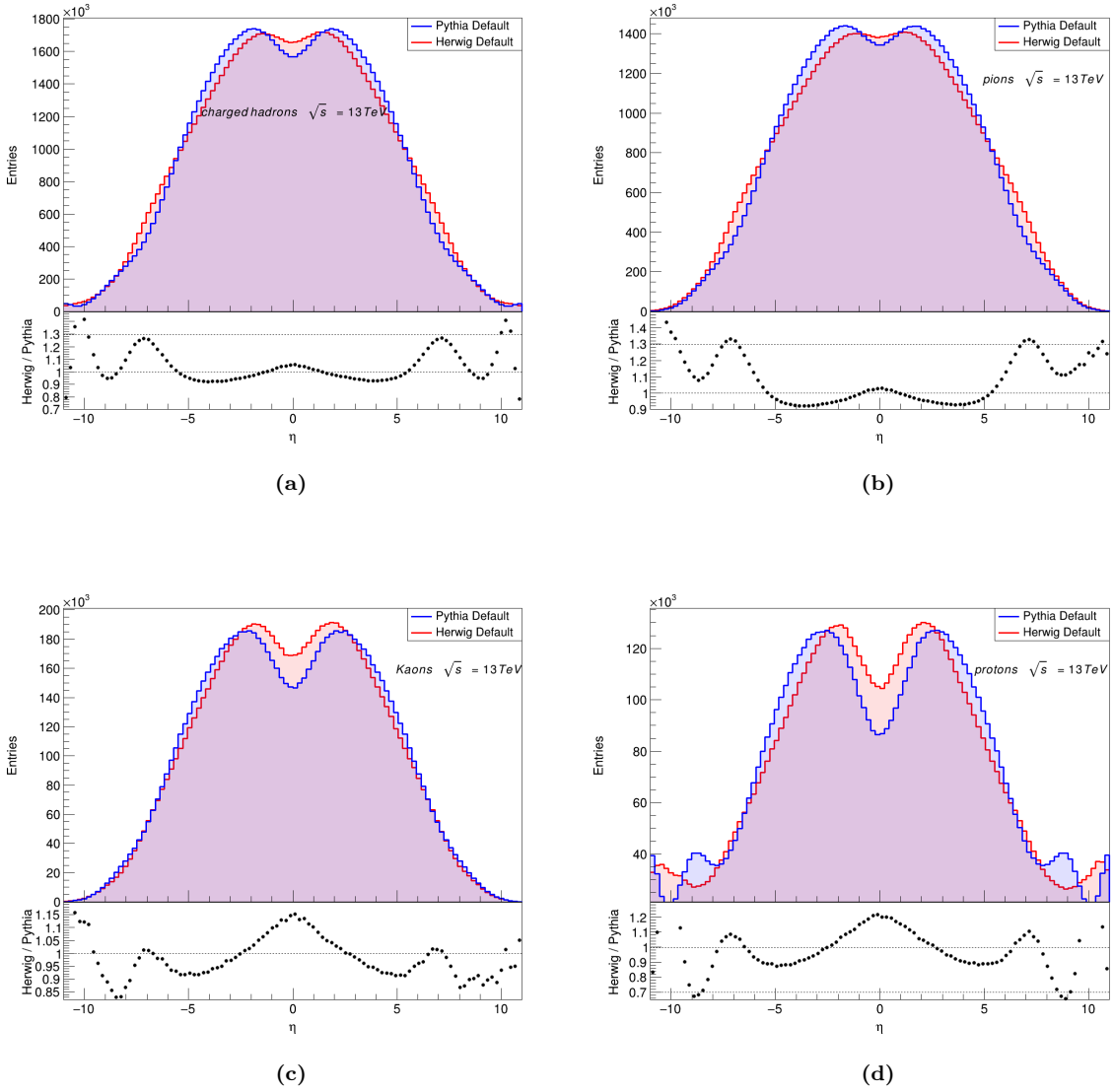


Figure 3.7: Comparison of PYTHIA and HERWIG η distribution of charged hadrons (a), pions (b), kaons (c), and protons (d) for pp collisions at $\sqrt{s} = 13$ TeV.

partonic interactions, where the models tend to disagree. Lastly, for the case of protons, there is a difference between the distribution for HERWIG and PYTHIA, indicating more elastic protons in PYTHIA than HERWIG. There are also some noticeable peaks at $|\eta| \sim 8$ and 10, indicating that a significant fraction of protons are produced in the forward and backward regions. These protons are remnants of the incoming protons from the initial collision, which are typically associated with soft interactions, and because of the difference in the models used in PYTHIA and HERWIG for the treatment of MPI, diffraction, and proton remnants. A similar trend can be seen in the η distributions at 13 TeV as well, which is shown in Fig. 3.7. Both generators are in reasonable agreement in

the central region but diverge in the forward/backward regions due to differing treatments of beam remnants and diffraction models in the two generators.

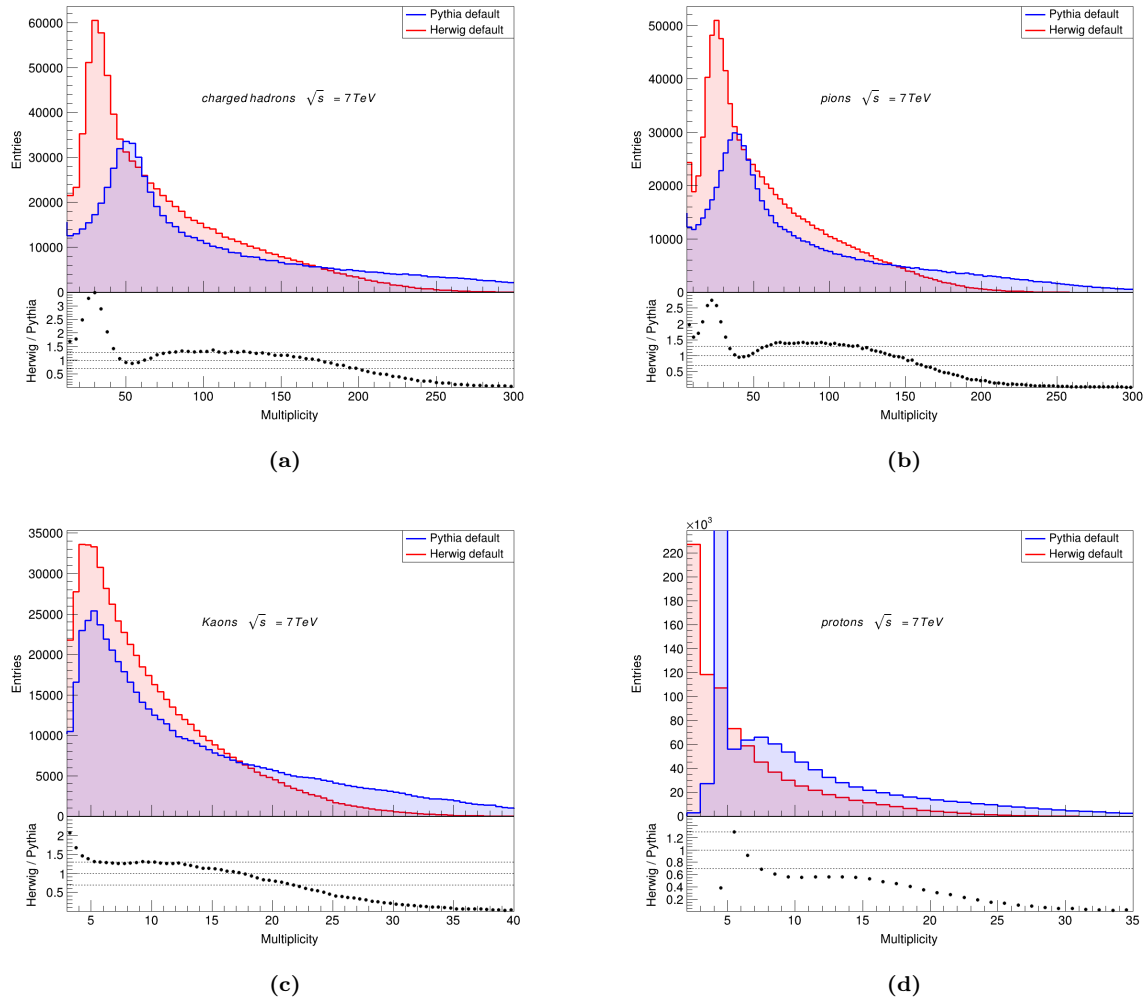


Figure 3.8: Comparison of PYTHIA and HERWIG multiplicity distribution of charged hadrons (a), pions (b), kaons (c), and protons (d) for pp collisions at $\sqrt{s} = 7$ TeV.

Finally, Fig. 3.8 shows the comparison between the two generators in the production of charged hadrons in (a), pions (b), kaons (c) and protons (d) in an event at 7 TeV. There are clearly more charged particles produced in PYTHIA, as compared to HERWIG. At low multiplicity, HERWIG shows a higher particle production rate compared to PYTHIA. This is because HERWIG tends to produce more particles in the initial stages of the parton shower process, which may be a result of differences in the underlying event and parton shower model. HERWIG typically uses a cluster model for hadronisation, where partons are clustered into hadrons in a way that often results in slightly more

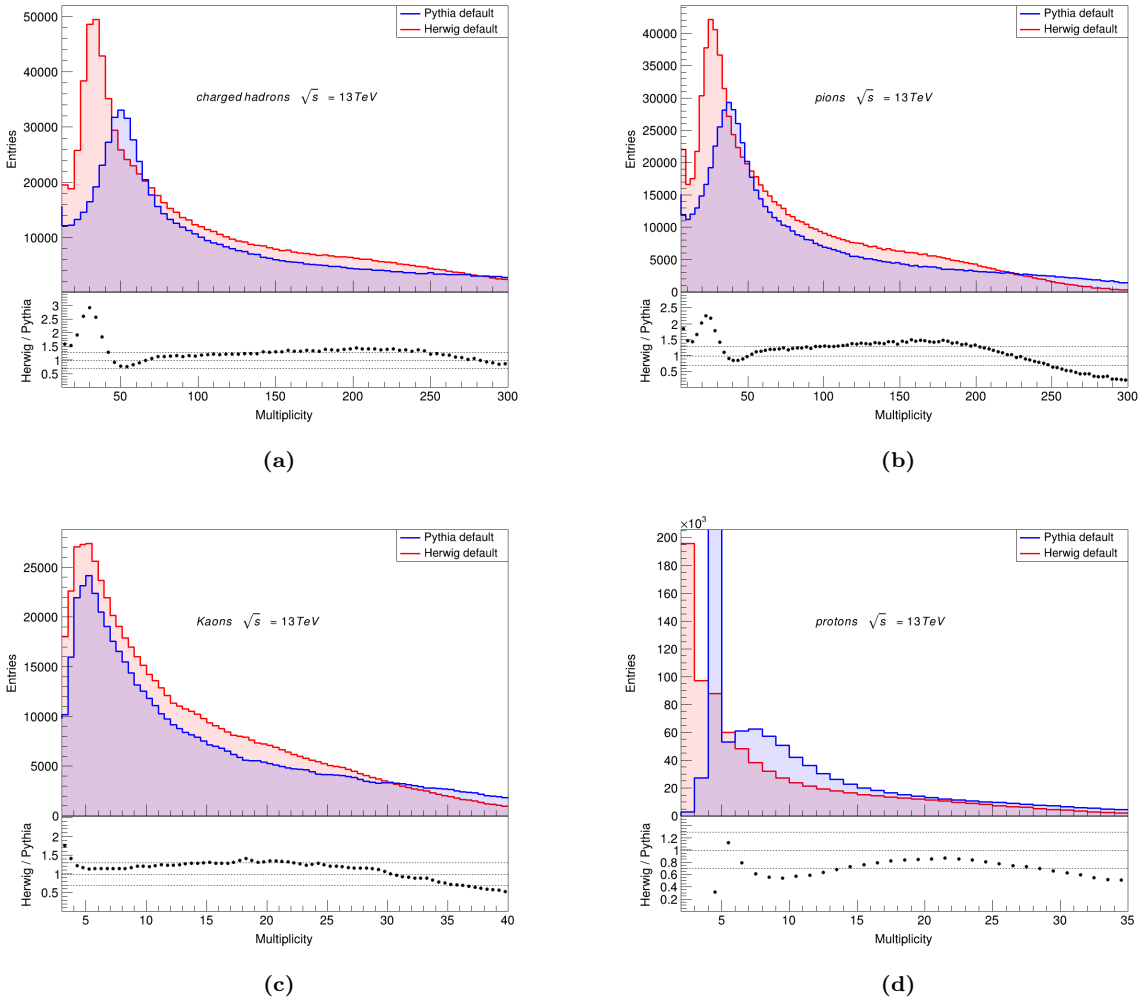


Figure 3.9: Comparison of PYTHIA and HERWIG multiplicity distribution of charged hadrons (a), pions (b), kaons (c) and protons (d) for pp collisions at $\sqrt{s} = 13$ TeV.

particles being produced initially. This can lead to higher multiplicity at low values. PYTHIA, on the other hand, often produces higher multiplicities at high energy due to its string model for hadronisation. In this model, partons (quarks and gluons) create strings that stretch between them and fragment into hadrons, typically resulting in a higher number of final-state particles. MPI in PYTHIA can also significantly contribute to higher multiplicities, as it simulates several hard scatterings in a single event, allowing a broader spectrum of particle production at higher energies. A similar sort of pattern could also be observed at 13 TeV as shown in Fig. 3.9.

Angular production of charged hadrons with respect to their momentum and transverse momentum for PYTHIA and HERWIG at 13 TeV is shown in Fig. 3.10. The resulting

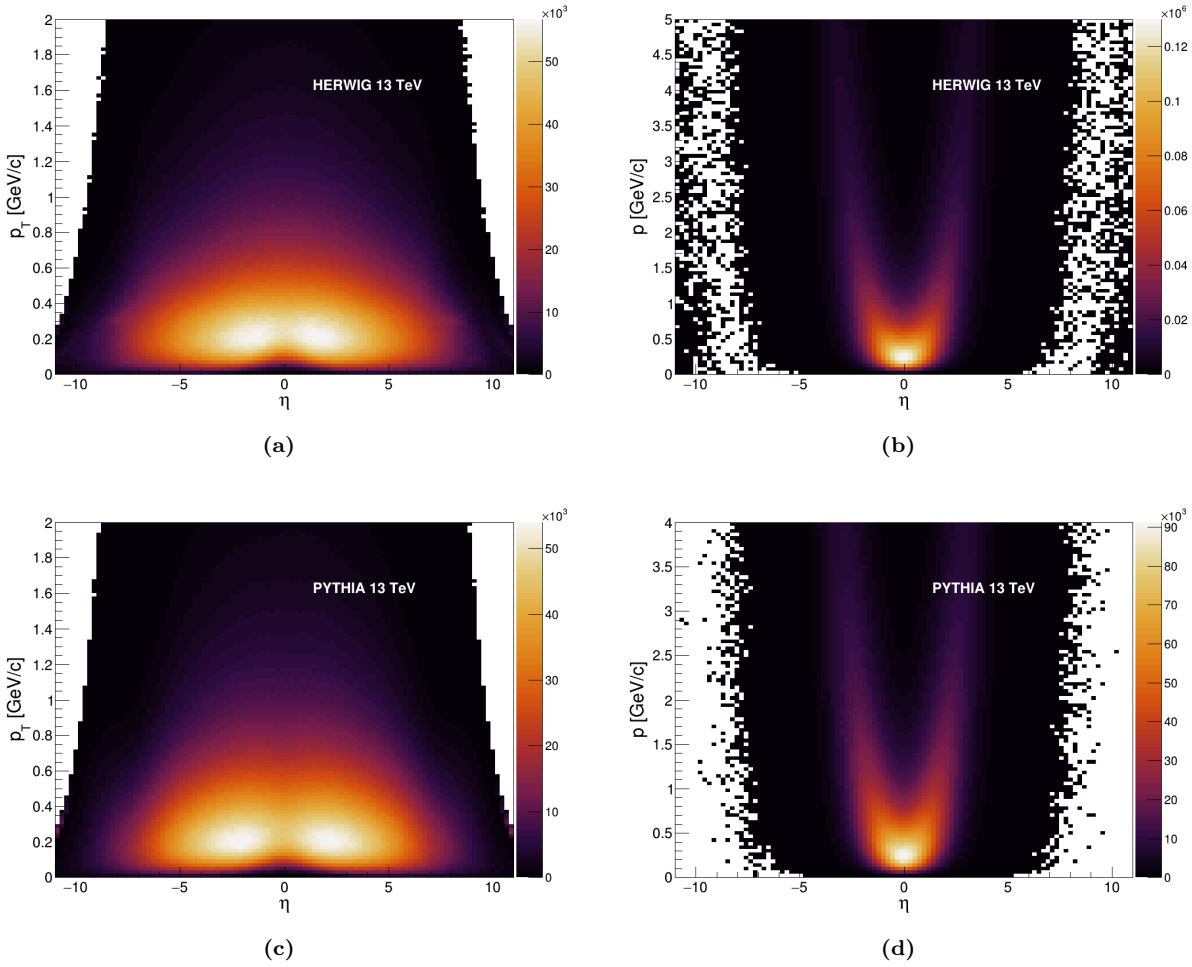


Figure 3.10: Comparison of charged hadron distributions for HERWIG and PYTHIA at $\sqrt{s} = 13$ TeV.(a) and (b) show the relationship between p_T and momentum with respect to η in HERWIG (c) and (d) shows the relationship between p_T and momentum with respect to η in PYTHIA for pp collisions at $\sqrt{s} = 13$ TeV.

plots show a high concentration of low-momentum particles produced predominantly in the central regions. This behaviour is consistent with expectations from soft QCD processes, where the majority of particles are produced with low transverse momenta and at mid-rapidity. Compared to HERWIG, slight differences are observed in the shape and population density for PYTHIA. These differences arise from the distinct hadronisation and parton shower models used in PYTHIA, which can lead to variations in particle multiplicity and kinematic distributions.

3.7 Impact of modification of parameters on particle production

Up to this point, various parameters such as pseudorapidity (η), transverse momentum (p_T), and multiplicity of charged hadrons for the default settings of PYTHIA and HERWIG were analysed. Next, is to investigate the impact on these distributions when any parameter in PYTHIA and HERWIG is slightly modified.

The importance of parameters becomes particularly evident when analysing the plots. Even slight changes in parameter values can have a significant impact on key observables such as transverse momentum, pseudorapidity, and particle production rates. To understand how these parameters influence the results, we refer to Eq. 3.1. Among various parameters, the most critical ones are p_{ref}^{T0}, ϵ and α_s. These parameters, as defined earlier, collectively govern the behaviour of the system and shape the distributions of the physical quantities being studied. These parameters are crucial for the study of multiparton interactions (MPI)

To observe how parameters influence the distributions, Fig. 3.11, 3.12 and 3.13 show a comparison of PYTHIA (a) and HERWIG (b) for transverse momentum, η and multiplicity with respect to the default and two modified settings in the values of parameters. The plots show three distributions, the solid black line refers to the default distribution, whereas the solid blue shows when the value of the cut-off parameter p_{T0}^{ref} is reduced from the default to 2.0 GeV/c and the red line shows when the scaling parameter ϵ is changed to 0.3 for PYTHIA and HERWIG. As mentioned earlier, p_{T0}^{ref} acts like a barrier between hard and soft interactions; therefore, once the value is reduced, there are more soft interactions produced, causing an increase in multiplicity. Since the value of the default parameter for p_{T0}^{ref} is 2.28 GeV/c for PYTHIA and 3.1 GeV/c for HERWIG, reducing this value in PYTHIA shows a significant impact on the number of particles, and the difference is visible in the p_T distribution of HERWIG.

The other important parameter that is modified is ϵ, i.e. the exponent of Eq. 3.3 that controls the energy scaling behaviour of the cut-off parameter and is crucial for ensuring that soft or low-energy emissions do not cause divergences in the perturbative QCD calculations. The comparison of PYTHIA and HERWIG for the charged hadrons with respect to transverse momentum is shown in Fig. 3.11.

Comparison of PYTHIA and HERWIG for the charged hadrons distribution for trans-

verse momentum is shown in Fig. 3.11 However, the distributions show that this parameter does not have a drastic impact and are quite similar to the default distributions. PYTHIA shows not a huge change; however, a slight change is visible at the lower p_T values in HERWIG.

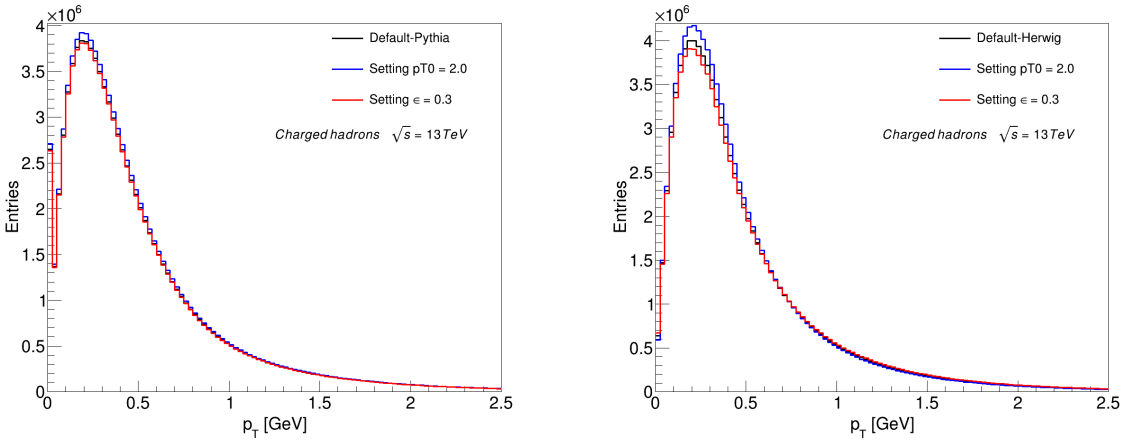


Figure 3.11: Influence of modifying parameters of PYTHIA (a) and HERWIG (b) as a function of transverse momentum of charged hadrons for pp collisions at $\sqrt{s}$= 13TeV.

Fig. 3.12, shows the η distribution for PYTHIA in (a) and HERWIG in (b). The distribution is roughly symmetric around $\eta = 0$, consistent with the symmetry of pp collisions, showing that the entries are concentrated near $\eta = 0$ corresponding to particles emitted perpendicular to the beam axis, with fewer particles at large η where particles travel closer to the beamline. The modifications have more noticeable effects compared to HERWIG, particularly between $\eta = \pm 5$.

The multiplicity plots for both generators are presented in Fig. 3.13. For PYTHIA, the multiplicity demonstrates minimal variation with parameter changes, as the differences in the default values and modified values are relatively small. In contrast, HERWIG exhibits a more noticeable variation. Detailed numerical values are provided in Table 3.5, offering a precise representation of the particle count per event as seen in the plots.

In conclusion, PYTHIA and HERWIG are two significant multi-purpose Monte Carlo event generators, each having its distinct physics models for different processes. This chapter demonstrates how variations in these models change the distributions of the key variables such as transverse momentum, η, and hadron multiplicity, which have a profound meaning in the simulation of the detectors' response. Moreover, it highlights

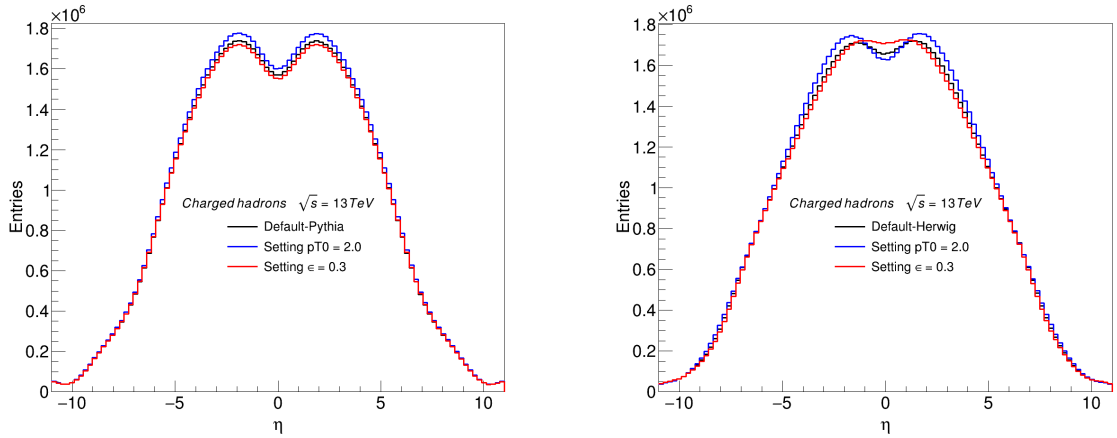


Figure 3.12: Influence of modifying parameters of PYTHIA (a) and HERWIG (b) as a function of η of charged hadrons for pp collisions at $\sqrt{s} = 13$ TeV.

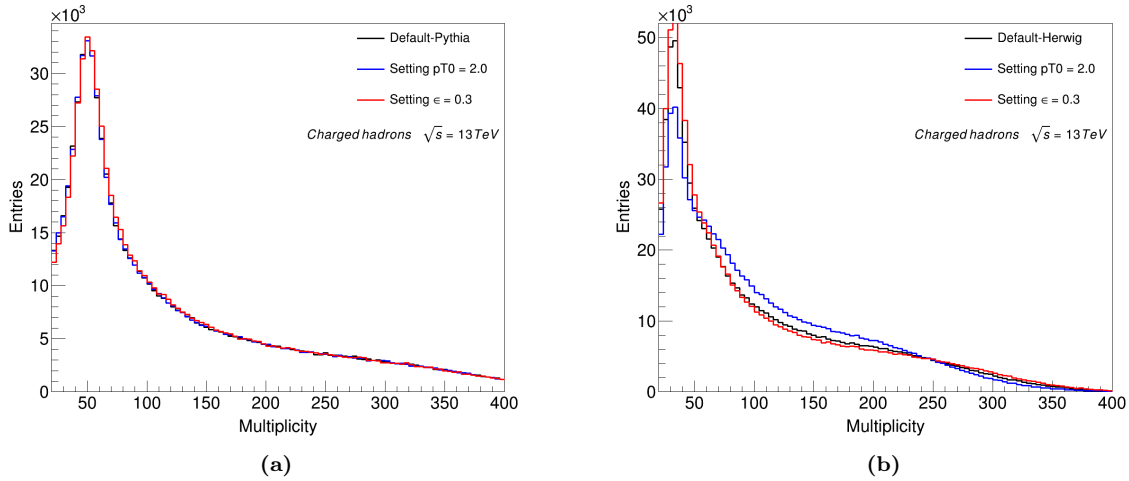


Figure 3.13: Influence of modifying parameters of PYTHIA (a) and HERWIG (b) as a function of multiplicity of charged hadrons for pp collisions at $\sqrt{s} = 13$ TeV.

the sensitivity of these models to the parameters defined within them, emphasising that even small variations in these parameters can substantially affect the results. Therefore, parameter selection must be approached with caution.

Apart from the charge hadron multiplicity, it is essential to know the composition of the produced flux for radiation damage evaluation. The variation in hadron multiplicity due to changes in parameter values directly influences particle fluence, which is a key factor in radiation damage studies. Since the protons (and neutrons) have the most severe damage power, they should be modelled most reliably. This analysis showed that the

	N_π		N_K		N_p		N_{ch}
PYTHIA ($p_{T0}^{ref} = 2.0\ GeV$)	71.66	77.5%	9.68	10.47%	11.077	11.98%	92.42
HERWIG ($p_{T0}^{ref} = 2.0\ GeV$)	77.31	81.55%	9.84	10.38%	7.61	8.02%	94.78
PYTHIA ($\epsilon = 0.3$)	71.62	77.5%	9.68	10.47%	11.075	11.98%	92.38
HERWIG ($\epsilon = 0.3$)	74.36	80.64%	9.78	10.60%	8.06	8.74%	92.21

Table 3.5: Average number of charged hadrons in an event for the modified settings of PYTHIA and HERWIG for pp collisions at $\sqrt{s} = 13$ TeV where the percentage indicates its fraction relative to the total average number of charged hadrons N_{ch}.

largest discrepancies between generations are noted in the case of proton multiplicity.

Fig. 3.14 shows the histogram for HERWIG and PYTHIA, comparing the mean transverse momentum with charged hadrons in Fig. 3.14(a) and (c). For the HERWIG distribution, the mean p_T shows a sharp peak at low p_T values (~ 0.4 GeV/c) for events with moderate multiplicities followed by a rapid drop-off, indicating that most events are soft and low in hadronic activity consistent with HERWIG's modeling of soft interactions, whereas, for PYTHIA, there are slightly higher values of mean p_T observed for events with moderate multiplicities.

On the other hand, PYTHIA's and HERWIG's multiplicity vs η distribution displays a broad, symmetric structure centred at $\eta = 0$, with the highest event densities concentrated at low multiplicities and central rapidities. The comparison between Herwig and Pythia highlights key differences in their underlying physics models, especially in how they generate high-multiplicity or forward events. Such variations are particularly relevant when considering detector acceptance and efficiency in regions like those covered by LHCb.

Eventually, to determine which generator describes the experimental data more accurately, a comparison within the LHCb acceptance regions and corresponding LHCb data is necessary. This comparison is discussed in detail in Section 4.9.

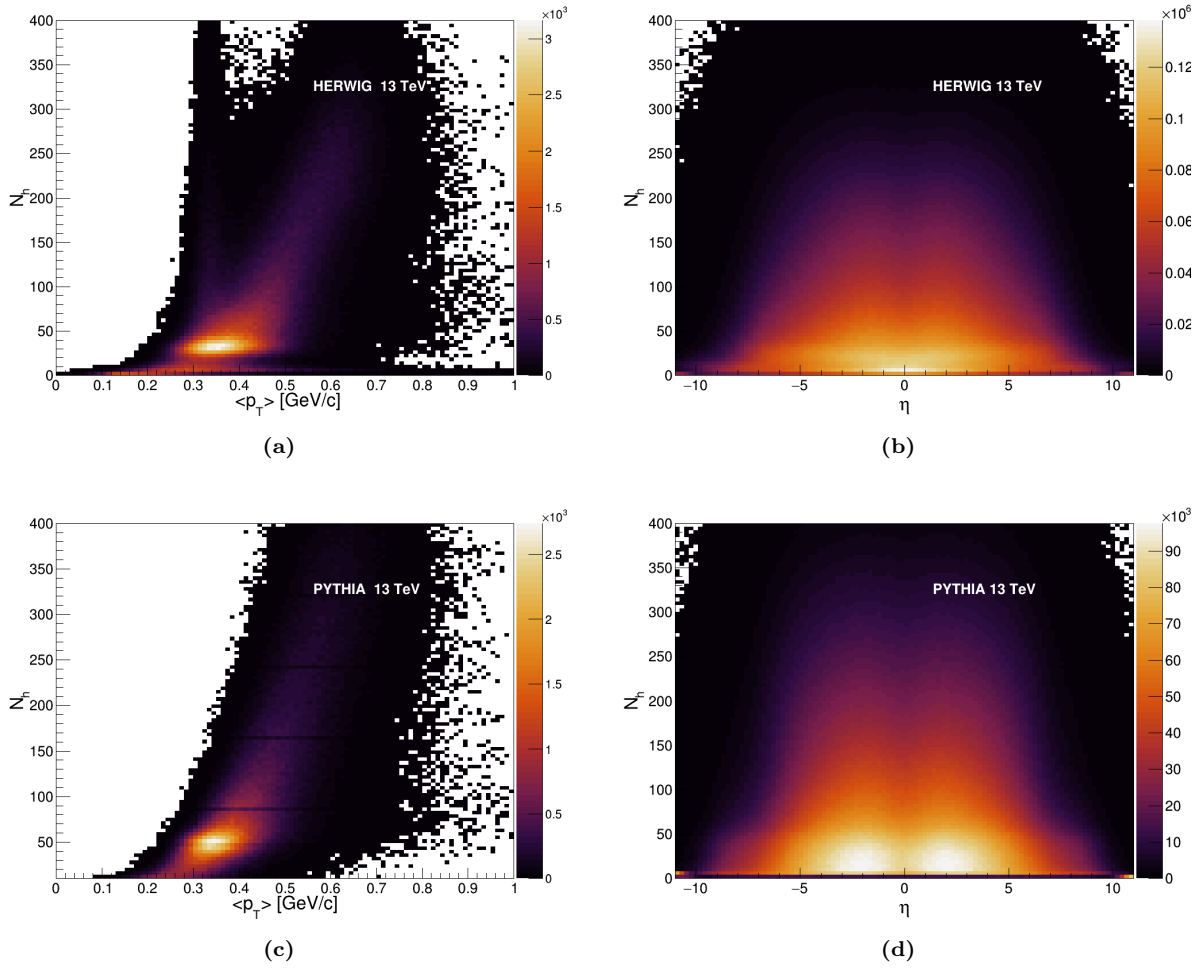


Figure 3.14: Relationship between charged hadrons and mean p_T of charged hadrons is shown for HERWIG in (a) and PYTHIA in (c) for pp collisions at $\sqrt{s} = 13$ TeV and the relationship between angular distribution and charged hadrons multiplicity is shown in (b) for HERWIG and (d) for PYTHIA for pp collisions at $\sqrt{s} = 13$ TeV.

3.8 Multiplicity and kinematic properties of simulated neutrons in the LHCb acceptance

In the LHC environment, particle radiation's main source is prompt particle production in pp collisions. Another radiation source is the production of secondary particles in interactions with the detector and the decay of particles. Tracking detectors are sensitive to charged particles, whereas neutrons are detected in the hadron calorimeter only. Neutrons carry no electric charge, and hence leave no direct signal in the tracking detectors, so they cannot be reconstructed. Since neutrons are especially harmful to the

silicon sensors, the study is further performed, whereas the neutron contribution to the fluence ϕ_{eq} might be evaluated based on the proton contribution. The simulation sample is used for this purpose, and the comparison of the energy spectrum and multiplicities of both hadrons is made, with special attention to the LHCb acceptance region. One should expect the same number of neutrons and protons should be produced in pp interaction, so it might be possible to establish the neutron contribution to the fluence ϕ_{eq}, taking into account the proton flux of particles traversing the VELO sensors.

The multiplicity distributions of protons (p) and neutrons (n) generated within the VELO (which is the same as LHCb) acceptance were compared as shown in Fig. 3.15 (a), revealing a nearly balanced production with a mean p/n ratio of 1.037 $\pm$ 0.001 as shown in Fig. 3.15 (b).

The η distribution of neutrons was studied for both the full generated range (-11 $< \eta <$ 11) and within the LHCb acceptance ($2 < \eta < 4.8$) as shown in Fig. 3.15 (c), highlighting how setting the the detector geometry filters the incoming neutrons.

There is no visible difference in the energy spectrum of neutrons and protons in the LHCb acceptance, see Fig. 3.15 (d).

Finally, a two-dimensional distribution of neutron kinetic energy is spread across the angular distribution within the LHCb acceptance, as shown in Fig. 3.15 (e).

The above study showed that the neutron contribution to the fluence, once only prompt particle production is considered, can be obtained based on the protons reconstructed in the detector, see further Chapter 5.

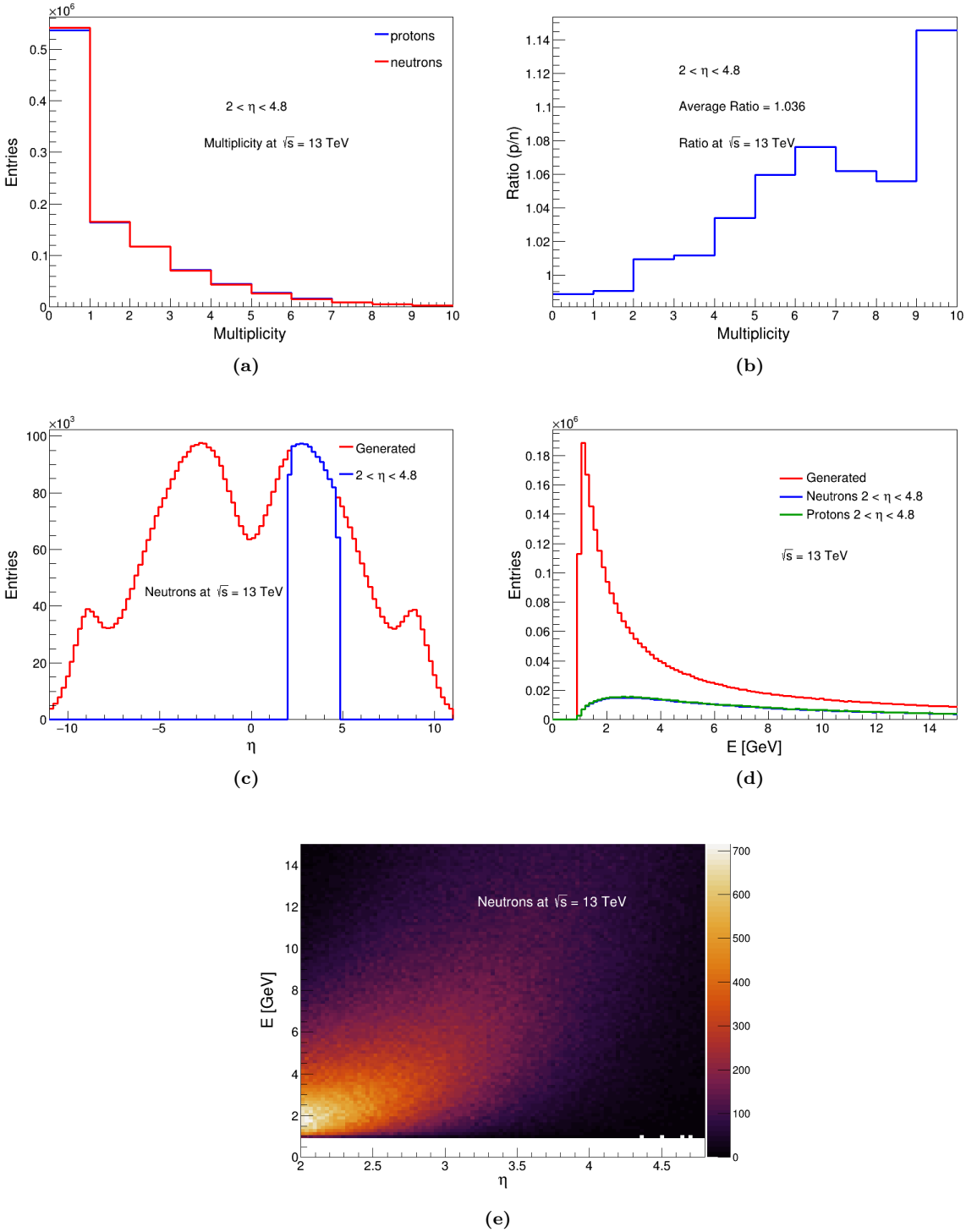


Figure 3.15: Comparision of proton and neutron production in pp collisions at $\sqrt{s}$ =13 TeV. (a) Multiplicity distributions of protons and neutrons within ($2 < \eta < 4.8$), (b) Ratio of proton to neutron multiplicity as a function of total multiplicity, (c) η distribution of neutrons over the full generated range and within the LHCb acceptance, (d) Total energy distribution of neutrons and protons generated and within LHCb acceptance, (e) Two-dimensional distribution of neutron kinetic energy versus η within the LHCb acceptance region.

Chapter 4

Tuning of PYTHIA Parameters

"Abundance is not something we acquire. It is something we tune into." - Wayne Dyer

4.1 Motivation

Monte Carlo event generators (MCEGs), as described in chapter 3, play a crucial role in understanding the physics underlying high-energy particle collisions. They serve as essential components for experimental analyses and are extensively used by both theorists and experimentalists to make predictions and prepare for future experiments. MCEGs incorporate physics models defined by a set of parameters. However, models have to be revised once experimental data are accessible. And, what is very important, the same model should describe all available data, obtained by all experiments, regardless of the conditions of data taking.

This chapter explores the application of the independent libraries RIVET [87] and PROFESSOR [88], within the LHCb simulation framework, specifically for the purpose of tuning various parameters in PYTHIA 8.204, including cross-sections, colour-reconnection schemes, flavour composition, and multiparton interaction parameters. Although the MCEGs come with well-chosen settings, it is often the case that further improvements might be needed to optimise the parameters, which would further improve the description of the data, especially in the forward direction. As mentioned in Chapter 3, the sensitivity of the physics models embedded in the generators can substantially affect the results. So, these parameters can be tuned to match the data in the best possible way.

The adjusting and optimisation of different parameters of MCEGs to the measured data is referred to as "tuning". During the *pp* collisions, there are events dominated by

small transfer of transverse momentum (p_T) which are described by phenomenological models in MCEG. For describing the minimum-bias events at the LHC energies, multiparton interactions (MPI) and colour reconnection (CR) parameters have the most important contribution. In order to better describe the data, it is important to tune the parameters. While the number of parameters in a generator can be large, the majority of the physics is determined by only a few, very important ones. These are the value of α_s, the properties of non-perturbative fragmentation functions, and the parameters related to the modelling of the soft physics and underlying events (UE) [89].

Because of the large and varied data sets available and the high statistics required, having a proper tuning effort can be a bit challenging. This often involves testing the generator against measured data for many observable collision energies and also generator settings. The tuning of the event generators requires dedicated tools such as RIVET [90] and PROFESSOR [88] software packages, which are used for the task in the wider community. These tools contain implementations of the analyses to reproduce measurements on the simulated data and perform some automated evaluation of the parameters of the event generators to match existing measurements in the best possible way.

The PROFESSOR tuning software [91] allows simultaneous tuning of a large number of parameters by approximating the MC response to the parameters analytically, using a parametrisation technique, where the generator is run at randomly sampled points within an interval. The optimal values of parameters can then be obtained with minimisation of the goodness-of-fit function using MINUIT [92].

The RIVET toolkit [90], on the other hand, is an analysis framework for MC generator validation. It allows comparing the MC generator output to data from a variety of physics analyses (plugins).

These tools were used in tuning activity in the past, but currently, these tools have been embedded in the LHCb Simulation framework to facilitate the tuning activity.

4.2 LHCb simulation software

4.2.1 Overview of LHCb simulation software

The LHCb software environment is updated frequently to accommodate the changes taking place in the detector and also to run event reconstruction in real-time. This

topic includes a brief introduction to various software and packages that are used in the LHCb environment. The dominant programming language for the back-end of the code is modern C++, however, the configuration of the components is done in Python. The framework of LHCb is built on Gaudi [93], which is a software package used for data processing applications for high-energy physics experiments.

LHCb real-time analysis strategy might be at risk of losing data in case if there are errors in the software trigger; therefore, a code review system that includes automated unit and integration tests, which checks the system and makes sure working of the new software works. The system is maintained by the LHCb developers. The software is released and deployed often to ensure better physics results.

LHCb code is split into multiple packages, which are versioned and managed on git version control system. The code is available for the public in [94]. The list of projects at LHCb is mentioned below:

- **GAUSS:** This is a project that manages the generation of MC events. It consists of two phases: the generation of events and the simulation of interaction with the detector, and these are handled separately. The first phase, precisely known as the generator phase, is controlled by an application called PYTHIA [63], however, the next phase, called the simulation phase, is taken care of by the application called GEANT 4 [62].

- **BOOLE:** the output of GAUSS is processed by BOOLE, which is responsible for the simulation of detector response, by converting the detector hits into the format of DAQ. In short, it provides the digitisation to mimic the response of real particles.

- **MOORE:** This package is responsible for configuring and running the trigger application. This includes all the information from the configuration of the detector hit decoding, track reconstruction, and PID to event and decay selections.

- **BRUNEL:** This program is responsible for reconstruction, where MC and data are treated identically. This performs the pattern recognition to reconstruct the charged particles and energy clusters deposited by both charged and neutral particles. Moreover, it also performs some initial particle identification.

- **DAVINCI:** This is a software that is used for physics analysis. This package accepts the fully reconstructed tracks and energy clusters as input, and users can manipulate the data and save outputs for further offline analysis.

Fig. 4.1 shows a simplified view of the LHCb data processing applications as used in Run 1 and Run 2.

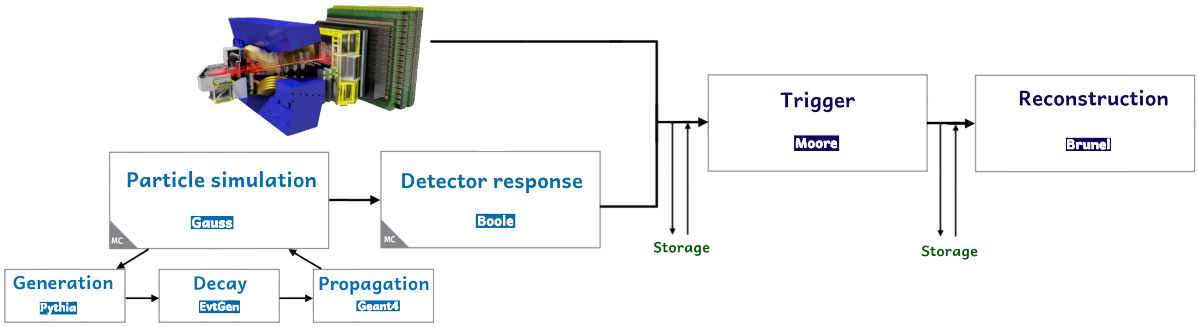
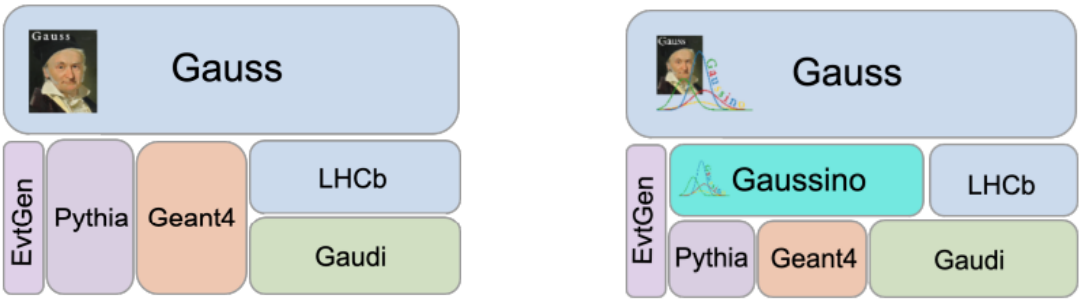


Figure 4.1: LHCb data processing applications as used in Run 1 and Run 2 [95].

4.2.2 LHCb simulation software upgrades for Run 3

LHCb experiment has resumed data taking after a major upgrade for Run 3 and operates with higher luminosity and trigger rates compared to the previous LHC runs. Therefore, there is a need for higher capacity in data storage and computing power to comprehend these changes. These higher rates are very challenging for the processing of the data, and hence, they require some major changes in the software applications.

The whole software at LHCb is being adapted to work in a multi-thread environment to efficiently and effectively use the computing resources. In the last few years, LHCb simulation software GAUSS was rewritten in view of Run 3 to introduce new technologies GAUDI and GEANT4 multi-threading, fast simulation techniques including machine learning based solutions, and new detector descriptions. GAUSSINO is a new experiment-independent simulation framework that would serve as a core simulation framework for GAUSS. GAUSS-ON-GAUSSINO is the new version of GAUSS based on GAUSSINO [96]. GAUSSINO is useful for the prototyping and testing of new technologies, which have an impact on core elements. It exploits the GAUSSINO's infrastructure and provides all the additional functionalities specific to LHCb. However, the major improvement in GAUSSINO is the inter-event parallelism, and this has caused many software elements to be rewritten to guarantee a thread-safe execution and also to have perfect communication with other simulation libraries. Fig. 4.2 shows the simulation software stack before and after the upgrade. For further information on GAUSS-ON-GAUSSINO, please refer to [96].



(a) Gauss Run 1 and Run 2 framework. **(b)** Gauss-on-Gaussino framework.

Figure 4.2: Dependencies in the simulation software stack before (a) and after upgrade (b) [96].

4.3 Methodology and validation tools

A well-prepared framework has been set up for tuning within LHCb. In earlier tuning campaigns, RIVET and PROFESSOR were used independently for tuning, but this framework has embedded them within the Parameter Tuning project. The workflow is shown in Fig. 4.3. The validation tools for tuning are RIVET and PROFESSOR with brief description given below:

RIVET

The RIVET toolkit (Robust Independent Validation of Experiment and Theory) is a versatile and independent software designed for validation, development, and tuning of MCEGs, specifically for SM processes. By reading HepMC event records, RIVET can connect with various event generators and provide a well-defined interface with any specific detector configuration. It has been used extensively by experimentalists for analysis and interpretation purposes, as well as for its applicability in future experimental studies. Moreover, RIVET proves instrumental in the advancement of novel analysis techniques, including machine learning applications, jet substructure analysis, boosted-particle tagging, and pile-up suppression [90].

Standard RIVET plugins used for tuning

RIVET plugins contain a parameterised physics measurement in a dedicated format. The following RIVET plugins from LHCb and other experiments are used for global tuning:

Experiment	Observable
LHCb	V^0 production ratios [97].
LHCb	ϕ production cross-section as a function of p_T and y [98].
LHCb	Prompt hadron production ratios [99].
LHCb	Measurement of energy flow [100].
LHCb	Charged particle multiplicity and density [101].
LHCb	Inelastic cross-section at $\sqrt{s} = 7$ TeV [102].
LHCb	Inelastic cross-section at $\sqrt{s} = 13$ TeV [103].
ATLAS	Track based Underlying Event [104].
ATLAS	K_s^0 and Λ production [105].
CMS	K_s^0, Λ and Cascade-transverse momentum and rapidity spectra [106].
CMS	Study of the UE at forward rapidity [107].
ALICE	p_T of neutral pions and η mesons [108].

Table 4.1: Standard RIVET Plugins used for tuning of parameters.

LbRIVET plugins

In addition to the standard RIVET plugins, a collection of exclusive LHCb RIVET plugins has been developed. This specialised package serves as a dedicated source for LHCb results, allowing to utilisation of patched or unreleased RIVET plugins that are LHCb-specific and therefore not incorporated into the general RIVET framework. The package fulfils a crucial role as a repository for alternative versions of RIVET analysis modules that are preferred for event generator tuning within the LHCb experiment. Furthermore, it facilitates communication between LHCb developers and package maintainers or simulation experts. These custom plugins play an important role in the tuning of MCEGs within the LHCb experiment.

PROFESSOR

PROFESSOR is an acronym for **PRO**cedure **F**or **ES**timating **S**ystematic err**OR**s and it offers a parameterisation-based approach for systematic generator tuning. The conventional formalism for generator tuning involves defining a goodness-of-fit (GoF) function that quantifies the agreement between generated data and reference data, followed by minimising this function. However, the true fit function is often non-analytic, and any iterative approach to minimise will be doomed at the expense of evaluating the fit function at a new parameter space point. To address this challenge, PROFESSOR provides an optimised methodology designed specifically for computationally expensive functions with unknown forms. In this approach, PYTHIA parameters are sampled a specified

number of times (user-defined as n) within a given range. For each parameter sampling, a Monte Carlo (MC) generation is performed, and the resulting predictions are compared to the reference data. Within each data bin, the outcomes of the random samplings are parameterised using polynomials of a certain order (e.g., third order or higher). The coefficients of these polynomials are then numerically compared within the PROFESSOR code. A χ^2 fit is performed using MINUIT to determine the optimal values of each tuning parameter based on the polynomial parameterisations' agreement with the data [91].

Parameter Tuning

LbMCSumbit is another package within the LHCb Simulation that provides a mechanism to define, test, and submit Monte Carlo production requests in LHCb by creating a YAML file, which is passed by the user. This file should include information about the event type, data type, and number of events, along with the simulation version.

In order to optimise the generated data to the reference data, we need to sample the data and run the MC jobs for each sample, which requires the power of the grid. The script "submit-productions.sh" submits the job on the grid with the information about the sample size, parameters, range of parameters, and the rivet analysis is provided by the user. Once the jobs are submitted and completed, they can be downloaded and further interpreted and tuned. This will result in a text file that includes tuned parameters, errors, goodness-of-fit, a correlation matrix, a covariance matrix, weights used, and the mean contribution from each plugin. Additionally, a YODA file is produced, which can be plotted using RIVET. The detailed description of the template of the input files is provided in [109]. The details of the parameters used and the RIVET analysis are individually done for each of the following sections below. The tuning procedure was carried out for cross-section, flavour composition, and multiparton interaction parameters.

4.4 Tuning of cross-section parameters at $\sqrt{s} = 7$ and 13 TeV

Tuning cross-sections is the first and foremost part of tuning, because there are not many parameters involved that contribute to advancing our understanding of fundamental particles and their interactions, supporting the development of new theories and models. The SigmaTotal class, as described in the PYTHIA manual [63], provides the computation

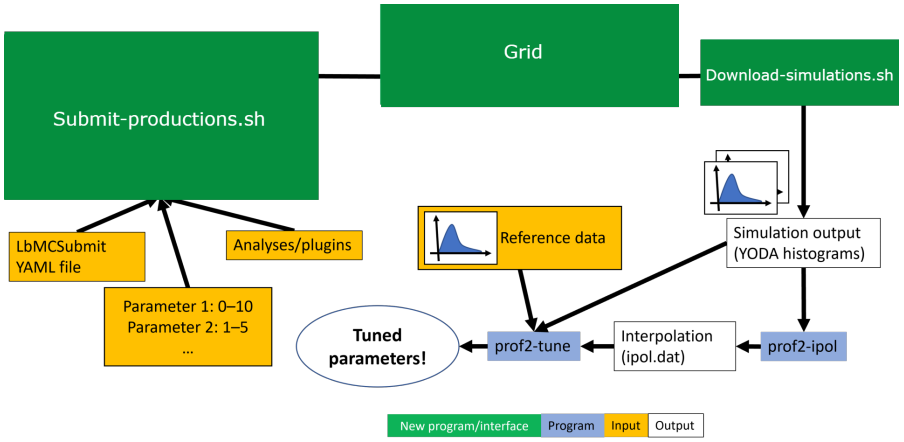


Figure 4.3: The Parameter tuning workflow within LHCb Simulation group [109].

of total, elastic, diffractive, and non-diffractive cross sections in collisions involving hadrons. The optimisation process involves tuning various parameters related to these cross sections, named as: total cross section, elastic cross section, single diffractive, and double diffractive cross section. Table 4.2 presents the sampling range and default values of these parameters, including the values from the PYTHIA 8.204 (P8 default) and LHCb current tunes.

Parameters	Min	Max	P8 default	LHCb current tunes [mb]
σ_{Tot}	85	110	100	94.07715
σ_{El}	20	30	25	29.46597
σ_{SD}	0	20	8	7.391497
σ_{SD}	0	20	8	7.560716
σ_{DD}	1.2	20	4	1.530803

Table 4.2: Tuning parameters, sampling range, PYTHIA default values [63] and LHCb current tunes.

The Sim10 distribution depicted in Fig. 4.4 is generated using a simulation with specific LHCb settings compared to tuning. The option files include the conditions of the beam, i.e., the energy of the beam, as well as other option files to set the data type and the type of event. For our case, we are using minimum bias events. Additionally, the RIVET analysis that we need to use is for the cross-section at 7 and 13 TeV. To achieve the tuning using the PROFESSOR framework, a dataset of one million events was utilised, which was subsequently sampled 200 times within the specified range. The generated samples (jobs) were then submitted to the grid, where each sample was generated one million times. These MC samples were then downloaded, interpolated, and then tuned using the

respective PROFESSOR commands. The optimised parameter values obtained from the PROFESSOR tuning procedure are presented in Table 4.3. Some of the parameters have uncertainty as high as the value of the parameter; this still needs further investigation.

Parameters	Optimized params at $\sqrt{s} = 7$ TeV	Optimized params at $\sqrt{s} = 13$ TeV [mb]
σ_{Tot}	85.228103 ± 19.81280	100.818829 ± 18.11507
σ_{El}	29.954630 ± 8.624173	29.915928 ± 9.687056
σ_{SD}	20.453580 ± 14.52265	19.986077 ± 15.83480
σ_{SD}	20.486980 ± 16.39889	19.892225 ± 19.17122
σ_{DD}	15.972319 ± 1.025035	1.260300 ± 16.16194

Table 4.3: Optimised values of the cross-section parameters with minimization errors at $\sqrt{s} = 7$ and 13 TeV.

A similar tuning procedure for $\sqrt{s} = 7$ TeV and 13 TeV was performed. The Table 4.3 presents parameter values that are closer to the current LHCb tune, particularly at the $\sqrt{s} = 13$ TeV. However, it is noticeable that the double diffraction value at 13 TeV appears anomalous and needs further investigation. Fig. 4.4 illustrates a comparison between tuning achieved with the current simulation (Sim10) at LHCb. The left distribution corresponds to the $\sqrt{s} = 7$ TeV analysis, where the tuning is derived from PROFESSOR, coincides with the current LHCb tune, although with a slightly lower χ^2 value compared to the current simulation (Sim10). Conversely, the right distribution represents the 13 TeV analysis, where current simulation (Sim10) demonstrates closer proximity to the data, while the tuning derived from PROFESSOR yields a χ^2/n of 5.35, indicating the need for further investigation. It is essential to emphasise that a single set of optimised parameters should be utilised for both the $\sqrt{s} = 7$ TeV and 13 TeV tuning processes. Moreover, careful consideration and thorough investigation should be undertaken when selecting RIVET plugins for tuning, as they may potentially impact the obtained parameter values negatively.

4.5 Tuning of flavour composition parameters with colour reconnection model at $\sqrt{s} = 7$ TeV

Tuning of flavour parameters contributes to the understanding of fundamental particle properties, facilitating the exploration of new physics beyond the SM. The StringFlav class is the selection of a new flavour during the fragmentation process and generating a

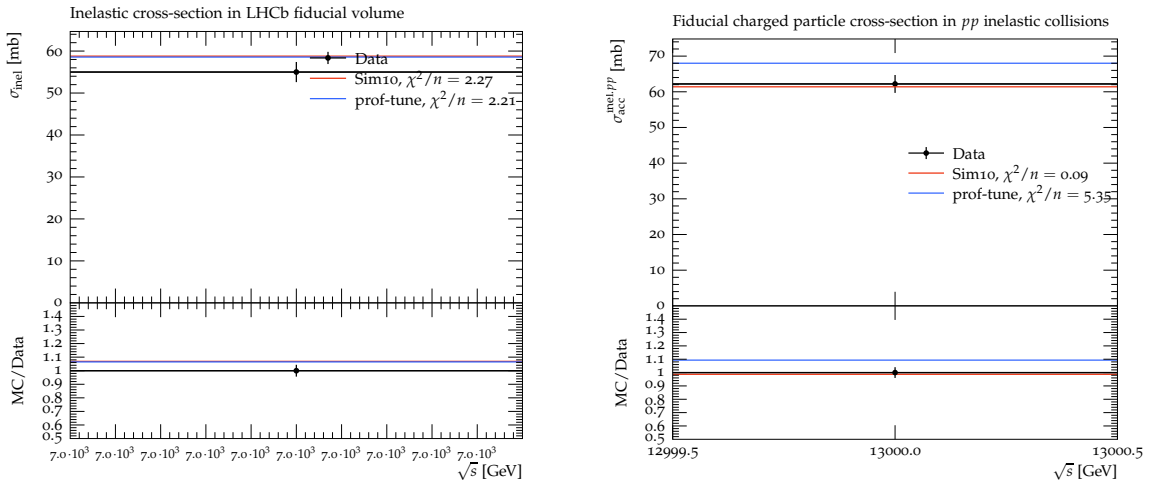


Figure 4.4: Comparison of Sim10 with tuned parameters at $\sqrt{s} = 7$ TeV (left) and 13 TeV (right), along with the χ-squared values indicating the χ^2, where the blue represents my tunes and the red line is the current simulation at LHCb.

new hadron, based on a given set of input flavours [63]. The tuning process is focused on optimising specific parameters associated with flavour handling. The production rate of different particle species is controlled by the following parameters, which are defined in Table 4.4:

Parameters	Definitions
mesonSvector	the relative production ratio for strange mesons
probQQtoQ	the suppression of diquark production relative to quark production.
probStoUD	the suppression of s quark production relative to u or d quark.
probSQtoQQ	the suppression of strange diquark production relative to light diquark production.

Table 4.4: Flavor optimization parameters,and their definitions from the PYTHIA manual [63].

To facilitate the tuning of these parameters, we use RIVET plugins from the following measurements listed in Table 4.5.

Experiment	Measurements
LHCb	ϕ production cross-section as a function of p_{T} and y [98].
LHCb	V^0 production ratios in pp collisions [97].
LHCb	Prompt hadron production ratios in pp collisions [99].
LHCb	Charged particle multiplicities and densities at $\sqrt{s} = 7$ TeV [110].

Table 4.5: Private RIVET plugins from the LHCb measurements used in the tuning procedure for flavour composition parameters at $\sqrt{s} = 7$ TeV.

The RIVET plugins used in the tuning process are from both private and standard plugins. The tuning results are then compared to the current simulation (Sim10) distribution generated using 500,000 events and identical settings for RIVET plugins. The optimised parameter values obtained from PROFESSOR are presented in Table 4.6.

Flavor Parameters	Optimized parameters from PROFESSOR
mesonSvector	0.071212 ± 0.0469
probQQtoQ	0.111079 ± 0.06511
probStoUD	0.707540 ± 0.0911
probSQtoQQ	0.532933 ± 0.0911

Table 4.6: Optimised values of flavour composition parameters with minimisation errors.

Optimising colour reconnection parameters

A new tuning approach was incorporated in PYTHIA to ensure better results, especially for the tuning of the hadron production ratios, which was termed as colour reconnection (CR). In order to track the colour information during the parton shower, partons are connected by colour lines. The quarks and anti-quarks are represented by colour lines with arrows pointing in the direction of colour flow, and gluons are represented by the opposite arrows. The CR model allows the lines to be formed between partons from different interactions and thus allows different colour topologies. This approach also improved the results of multiplicity distribution when compared with data [111]. However, due to its significantly longer computational time compared to PYTHIA, it was not initially included for tuning purposes. Nonetheless, the introduction of colour reconnection parameters offered potential for improving the results. The colour reconnection model has different schemes, which are discussed in detail in 4.7. All of the colour reconnection schemes have a commonality, and that is to use the configuration that minimises the total string length used in Eq. 4.2 between the two partons.

Table 4.7 displays the parameters, along with their definitions, that were used for tuning. This model requires parameters to have a fixed value, which, along with the definitions, are mentioned below:

The process of tuning the colour reconnection scheme requires some parameters to have fixed values. The list of those parameters with the fixed values is:

- **remnantMode = 1**; the model set for the beam remnants, where 1 refers to the new model.

Parameters	Definitions
timeDilationPar	disallow colour reconnection between strings that are not in causal contact
m0	used in the lambda (Eq. 4.2) measure as a protection against small mass systems
junctionCorrection	used in the lambda-measure for junction strings, this parameter allows one to vary how easily junctions form
pT0Ref	pT0Ref is the pT0 at a certain center of mass energy

Table 4.7: clolour Reconnection optimization parameters and their definitions [63].

- **mode = 1**; this would choose the scheme where 1 refers to the QCD-based scheme.
- **lambdaForm = 2**; this allows switching between different options for what lambda-measure to use, set to 2 as per the previous tuning campaigns.

The values obtained for the colour reconnection parameters after the tuning are shown in Table 4.8.

Parameters	Optimized parameters from PROFESSOR
timeDilationPar	2.363346 ± 0.1213 [GeV^{-1}]
m0	0.016538 ± 1.311 [GeV]
junctionCorrection	0.845162 ± 0.1363
pT0Ref	2.251031 ± 0.1319 [GeV/c]

Table 4.8: Optimized results of colour reconnection parameters obtained from PROFESSOR.

The presented optimised parameters are the results of tuning the flavour composition and colour reconnection scheme. The next steps involve tuning the flavour composition parameters while keeping the colour reconnection parameters fixed, followed by tuning the colour reconnection parameters and keeping the flavour composition parameters fixed. This iterative process aims to determine the fixed values of these parameters and further refine the tuning of the multiparton interaction (MPI) model.

Comparison of flavour composition and colour reconnection parameters

Comparison was performed to evaluate the different tuning approaches for the flavour and colour reconnection parameters. The following plots shown in Fig. 4.5 and 4.6 depict various types of tuning in comparison to the current simulation (Sim 10). There were four different tuning scenarios for half a million events, which are as follows:

- TUNE1: Exclusively tuning the flavour composition parameters.

- TUNE2: Exclusively tuning the colour reconnection parameters.
- TUNE3: Tuning the flavor composition parameters while keeping the colour reconnection parameters fixed.
- TUNE4: Tuning colour reconnection parameters while keeping the flavor composition parameters fixed.

The Fig. 4.5 (a) and (b) show the multiplicity distribution of prompt charged particles with respect to η in (a) and p_T in (b), where the data is compared to four different types of settings, the description of the settings is mentioned above. Along with the settings, there is also Sim10, which is the current tuning of PYTHIA. The χ^2 values would indicate the goodness-of-fit. According to the χ^2 values, the current tune settings best describe the data, but since we want to consider the CR and flavour settings together so the important TUNE settings are TUNE3 and TUNE4. Therefore, TUNE3 better describes the data for both p_T and η distributions of prompt charged particles. For the V^0 ratio plots (c) and (d) and hadron ratio distribution plots (e) and (f), TUNE4 is the one with the lowest χ^2 values, in some cases after the current tune settings. However, the exact values of parameters are still under investigation, and one also needs to consider the values and tune them at higher energies.

4.6 Tuning of multiparton interaction parameters at $\sqrt{s}$= 7 TeV

Once the CR and flavor parameters have been tuned, we obtain specific values for these adjustable parameters. With these values fixed, the tuning of MPI parameters is done and explained in section 3.6.

For the purpose of tuning the minimum, maximum values of these parameters, along with the default and current LHCb tune, are shown in Table 4.9:

Parameters	Min	Max	P8 default	Current tune
pT0Ref [GeV/c]	1.5	2.8	2.28	2.742289
ecmPow	0.15	0.3	0.215	0.238
alphaSvalue	0.1	0.2	1.30	0.130

Table 4.9: Multiparton interaction parameters, sampling ranges, PYTHIA default values [63] and values in the LHCb tune.

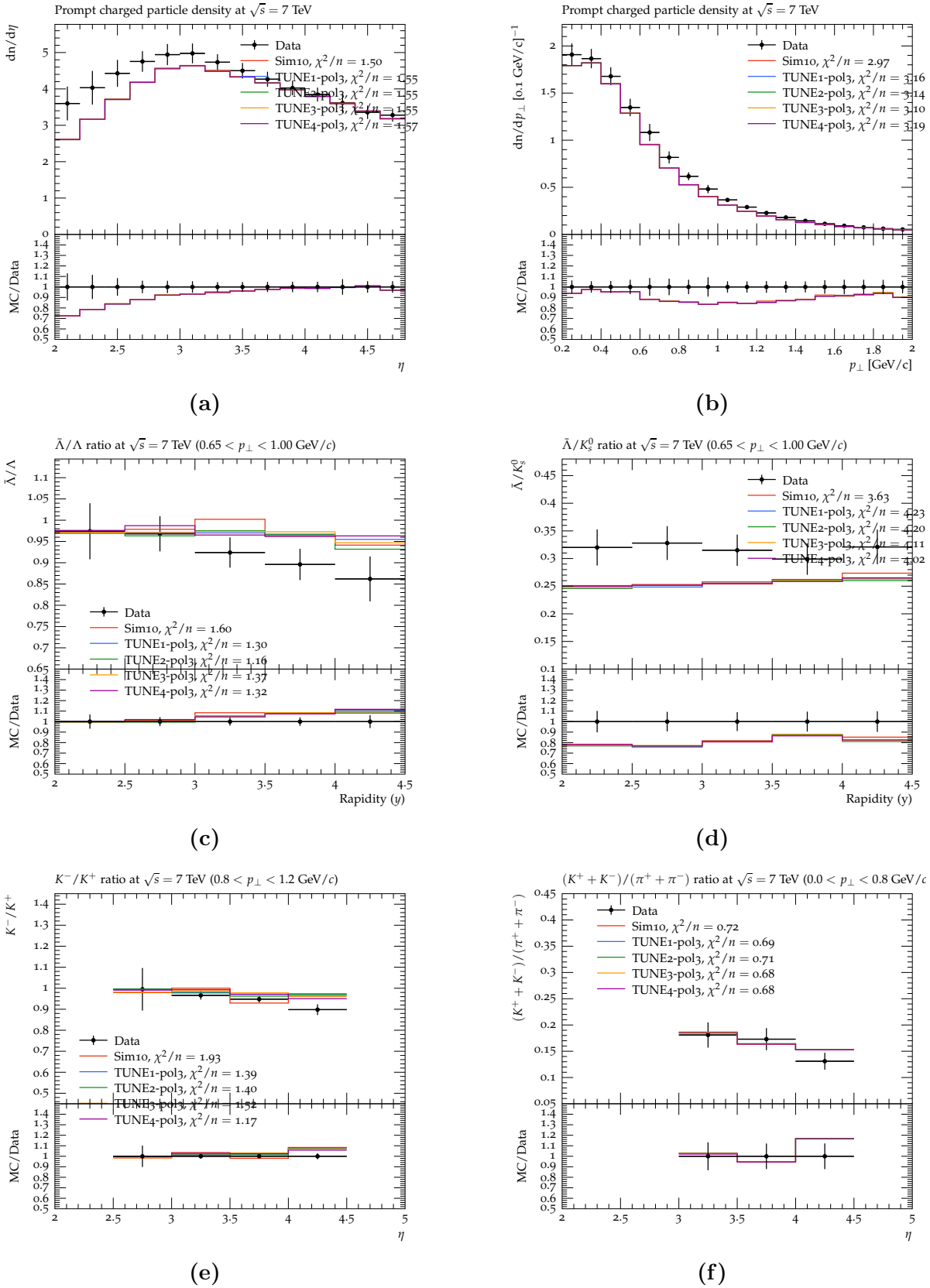
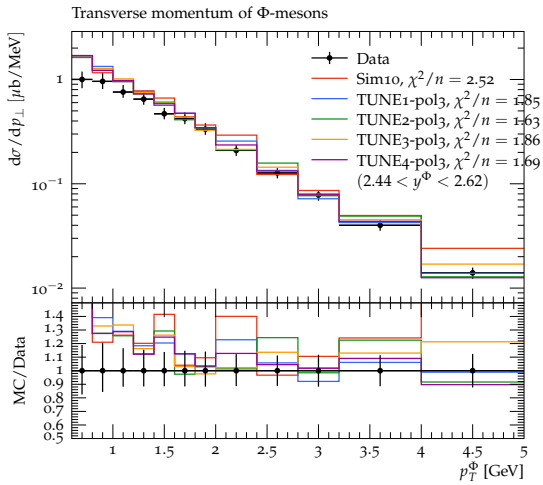
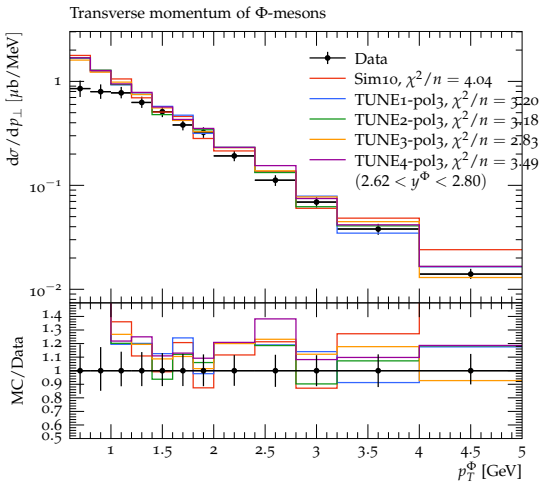


Figure 4.5: Distribution for multiplicity (a) and (b), V^0 and prompt charged particle production ratios (c),(d),(e),(f) illustrating different settings of TUNES and their comparison with Sim10 at $\sqrt{s} = 7$ TeV.



(a) Transverse momentum of ϕ mesons at $\sqrt{s} = 7$ TeV in $2.44 < \mathrm{y} < 2.62$.



(b) Transverse momentum of ϕ mesons at $\sqrt{s} = 7$ TeV in $2.62 < \mathrm{y} < 2.8$.

Figure 4.6: Distribution for ϕ mesons illustrating different settings of TUNES and their comparison with Sim10 at $\sqrt{s} = 7$ TeV.

The RIVET plugins for the charged particle multiplicity and density at $\sqrt{s} = 7$ TeV in pp collisions are used. The plugins are not only from the LHCb analysis but also the track-based minimum bias plugin at $\sqrt{s} = 7$ TeV from the ATLAS experiment, and the underlying events forward rapidity and p_T and η spectra at $\sqrt{s} = 7$ TeV from CMS are also included. There were three separate jobs submitted on the grid with a sample size of 1 million events, with the details as follows:

- TUNE1: This tune refers to the job where all the RIVET plugins are given equal weights.
- TUNE2: this tune refers to the job where different weight is assigned to each plugin (where LHCb is favoured).
- TUNE3: this tune refers to the job with LHCb plugins only, because the results from other plugins might have a negative impact on this measurement, since LHCb is a forward detector, and the central detector tunes may not be suitable for the forward tuning.

The results of these tuning are compared with the current LHCb tune (Sim10). The optimised values of these parameters obtained from the three tunings are shown in Table 4.10.

Parameters	LHCb	TUNE1	TUNE2	TUNE3
pT0Ref [GeV/c]	2.742289	1.790921 ± 0.0284	1.500713 ± 0.0013	2.011927 ± 0.235
alphaSvalue	0.130	0.115543 ± 0.0026	0.199980 ± 0.00018	0.157254 ± 0.029
ecmPow	0.238	0.299414 ± 0.00006	0.299730 ± 0.0029	0.214088 ± 0.052

Table 4.10: The values of the optimised parameters obtained for the three different sets of tunings.

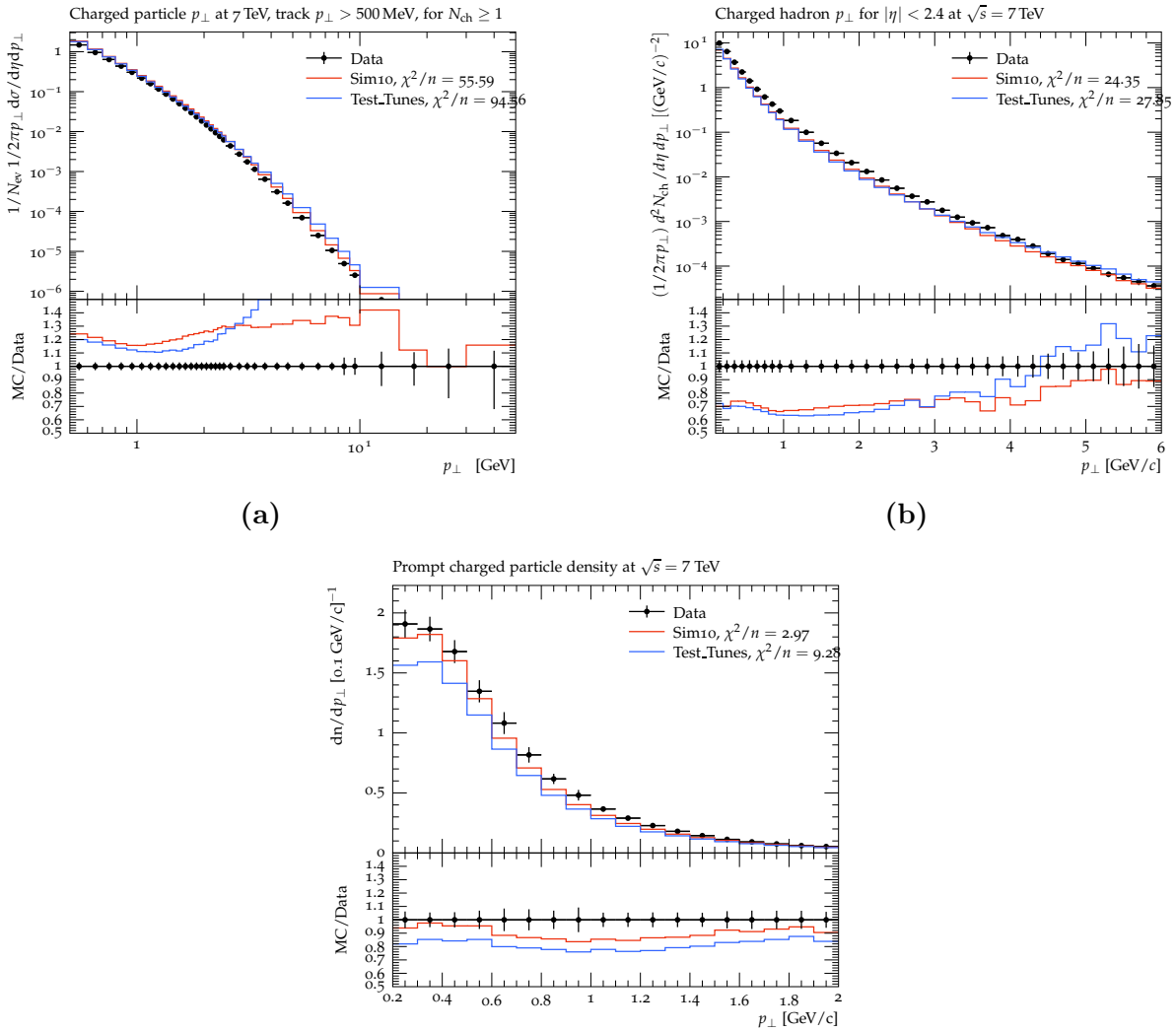


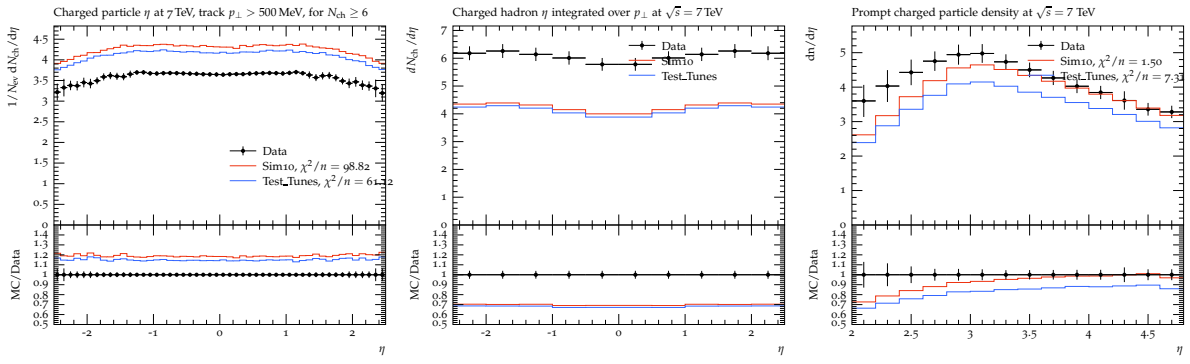
Figure 4.7: Multiplicity distribution plots w.r.t p_T from ATLAS (a), CMS (b), and LHCb (c) for the final tunes compared with the Sim10 (current tune). The χ^2 values are mentioned, which would be useful in analysing which tune better describes the data.

From the values obtained in Table 4.10, pT0Ref has the lowest minimisation uncertainty, although the values of TUNE3 are closer to the current tune and also to the

default value. For the case of coupling constant (α_s), the default and the current tune have the same value, where TUNE2 is closer to the upper limit bound (0.1 - 0.2), but has the least minimisation error. TUNE3, on the other hand, is within the minimisation errors and closer to the LHCb current tune. The last value is for the energy rescaling pace ($ecmPow$); the minimum and maximum values set are 0.15 and 0.3, respectively. TUNE1 and TUNE2 are closer to the upper bound. TUNE3 has values within the minimisation errors.

TUNE3 seems to be better describing the data, since it has its values closer to the current tune as well as the default values of the parameters. So, moving ahead and fixing the MPI parameters to TUNE3 and flavour and CR parameters obtained earlier, we would now compare this tuning locally with the current tune and see if better results are obtained. The RIVET plugins thus picked are for the multiplicity of particles, V^0 production ratios, and prompt charged hadron production ratios at LHCb, ATLAS, and CMS. The results obtained are shown in Fig. 4.7.

Fig. 4.7 shows multiplicity distribution of charged particles from ATLAS (a), CMS (b), and LHCb (c) with respect to p_T. From all three plots, it is quite evident that current tuning fails to explain the data and thus proves the current tune is still a better tune, and the Test_Tunes requires further investigation.

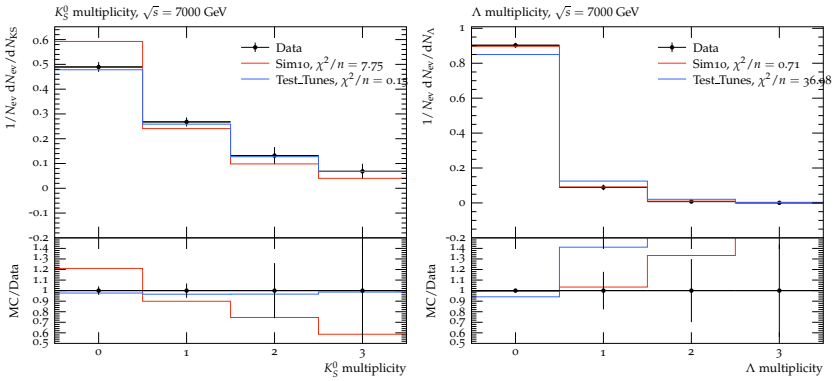


(a) **(b)** **(c)**

Figure 4.8: Multiplicity distribution plots w.r.t η from ATLAS (a), CMS (b), and LHCb (c) for the final tunes compared with the Sim10 (current tune).

Fig. 4.8 shows the multiplicity distribution of charged particles from ATLAS (a), CMS (b), and LHCb (c) with respect to η. For the ATLAS and CMS plot, the Test_Tunes better describes the data, whereas not for the LHCb distribution, where the current tune better explains the data.

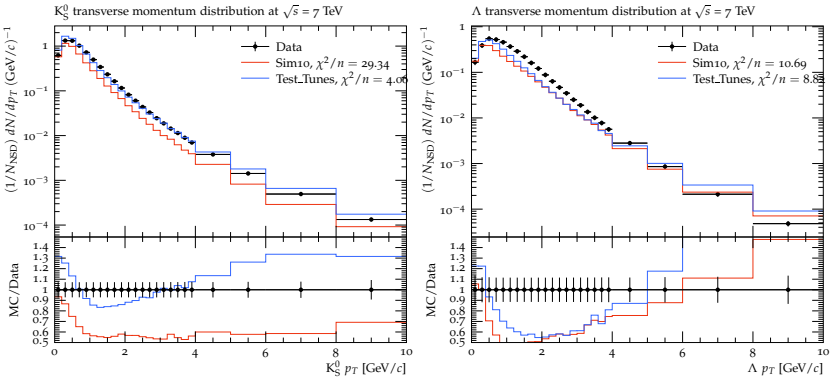
K_s^0 and Λ multiplicity distribution plots are shown in Fig. 4.9 from the ATLAS experiment where it is quite evident from the χ^2 values that the Test_Tunes better describe the ATLAS data for K_s^0 but it fails to describe the Λ multiplicity distribution.



(a) **(b)**

Figure 4.9: Multiplicity distribution for K_s (a) and Λ (b) from ATLAS experiment compared with the Sim10 (current tune) along with χ^2/nvalues.

However, if we look at the p_T distribution of K_s^0 and Λ as shown in Fig. 4.10, the Test_Tunes are far more able to describe the data as compared to the current LHCb tune.

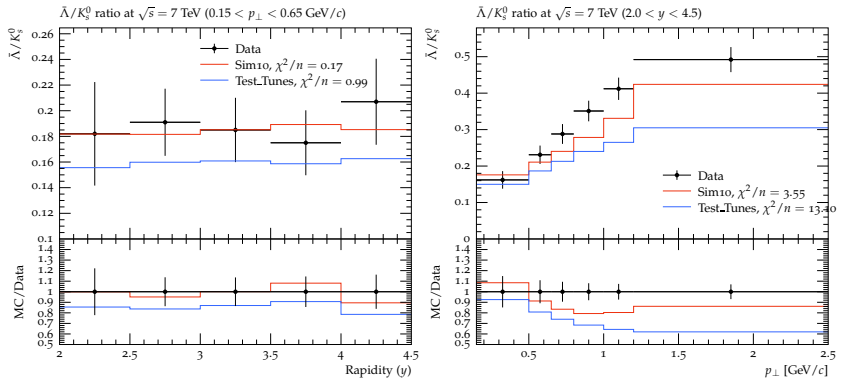


(a) **(b)**

Figure 4.10: p_T distribution for K_s^0 (a) and Λ (b) from CMS experiment compared with the Sim10 (current tune).

Finally, looking at the ratio plots for V^0 ratios from the LHCb experiment shown in Fig. 4.11. The plots show that current tuning is far better than Test_Tunes and hence Test_Tunes needs further investigation. There could be however, be further improvement

made by changing the CR model, since the model used is QCD-inspired, changing the scheme can improve the results. The details of these schemes are shown in the next section 4.7.



(a) **(b)**

Figure 4.11: Ratio distribution for anti particles i.e. $\bar{\Lambda}$ to $\bar{K_s}$ with respect to rapidity (a) and p_T (b) from LHCb experiment compared with the Sim10 (current tune) along with their χ^2-square values.

4.7 Different schemes in the colour reconnection model

This section includes different settings in the colour reconnection model for PYTHIA 8 generator using the minbias event observable at $\sqrt{s} = 7$ TeV. The colour reconnection and MPI (multiparton interaction) effects are strongly correlated; therefore, when tuning both the parameters for MPI and CR should be considered simultaneously. These are not the results from tuning, but a comparison between the different schemes (settings) within the CR model [112].

The MPI-based CR model was implemented in PYTHIA until PYTHIA 8.2 there were two other schemes introduced. Therefore, now the CR model consists of three different schemes, which are MPI-based, QCD-inspired, and gluon-move. A brief introduction to these schemes is given below:

- **MPI-based CR model (CR0)**: In this scheme, the probability for colored partons with transverse momentum p_T from MPI to reconnect is calculated using:

$$P_{rec}(P_T) = \frac{(R_{rec}\, p_{T0})^2}{(R_{rec}\, p_{T0})^2 + p_T^2}, \tag{4.1}$$

where R_{rec} is a tunable parameter, and p_{T0} is the energy dependent parameter as shown in Eq. 3.3, which avoids the divergence of partonic cross-section at low p_T. The MPI system at high p_T would tend to escape from the interaction point, without being colour reconnected to the hard scattering system. The low p_T partons would instead be more likely to exchange colour. Once the systems to be connected are determined, low p_T partons are added to strings to achieve a minimal string length [112].

- **QCD-inspired CR model (CR1)**: This scheme is particularly evolved from the MPI-based model. The main difference concerning CR0 and CR2 is the more complete treatment of QCD multiplet structure, and in particular that reconnections of dipoles can produce structures of three(anti-) colour indices (junctions), and hence enhancing the production of baryons. However, the minimisation of the string length remains. The string length λ is determined in the following way:

$$\lambda = ln\left(1+\sqrt{2}\frac{E_1}{m_0}\right) + ln\left(1+\sqrt{2}\frac{E_2}{m_0}\right), \tag{4.2}$$

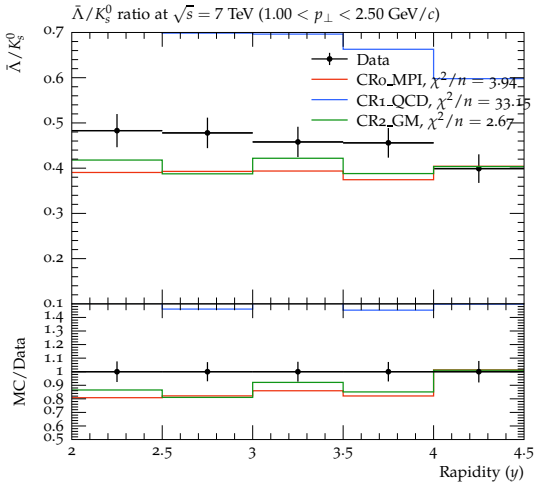
  where E_1 and E_2 represent the energies of the colored partons in the rest frame of the QCD dipole, and m_0 is a constant with the dimension of energy [113].

- **Gluon-move CR model (CR2)**: In this scheme, the reconnections are formed in the same way as in CR0; the main difference is that only gluons are considered for reconnection. The final gluons are recognised along with the pair of colour-reconnected parton pairs. Then an iterative process starts, the difference in string lengths when a final state gluon belongs to the two colour reconnected partons is moved to another connected two-parton system is calculated, and then the gluon is moved to the string which gives the largest reduction in the string length [113].

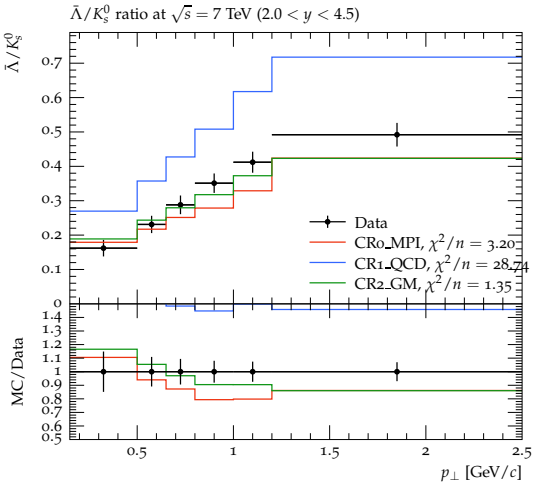
4.7.1 Results from comparison of the three CR schemes

This CR model was particularly introduced to study the ratios (V^0 production ratios and the multiplicity), therefore, the jobs were prepared using Gauss, and the comparison is shown in Fig. 4.12.

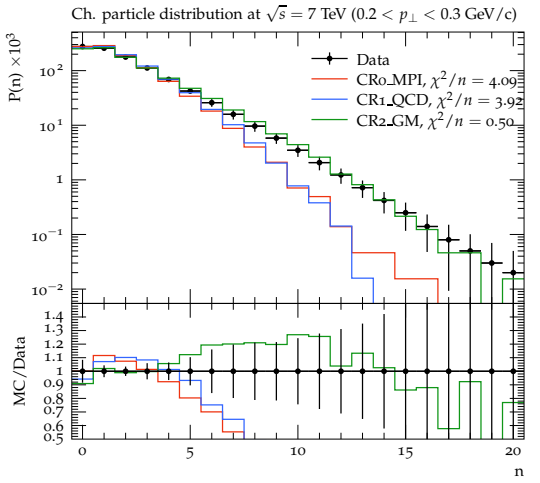
The distributions in Fig. 4.12 show the comparison of different schemes within the CR model. The very first distribution (a) shows the $\bar{\Lambda}/K^0_s$ ratio concerning the



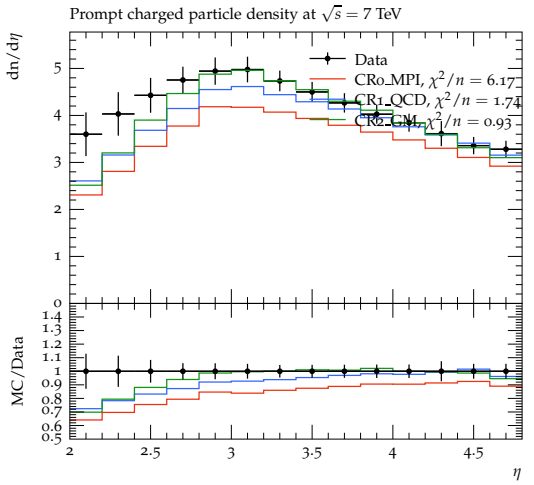
(a)



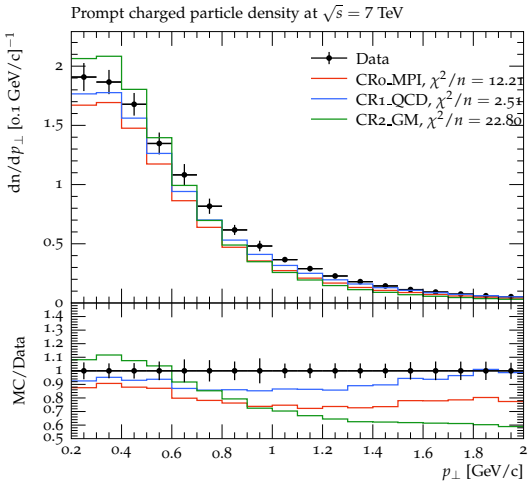
(b)



(c)



(d)



(e)

Figure 4.12: The distributions (a) and (b) represent the V^0 production ratios for rapidity and p_T respectively, where the three schemes are compared with the data. Distribution (c),(d), and (e) represent the multiplicity plots where the comparison is made at $\sqrt{s} = 7$ TeV.

rapidity at $\sqrt{s} = 7$ TeV in a particular range of p_T. The χ^2 values are provided as well, which would help to make a better comparison among the schemes. The χ^2 value of the QCD-inspired scheme describes the data the worst, whereas, gluon move model describes the data best. Similar is the case when the p_T distribution is looked at in a particular range of rapidity in (b). QCD-based scheme fails to explain the data, whereas the gluon move model describes the data the best. Fig. 4.12 (c) shows the multiplicity of charged particles and Fig. 4.12 (d) shows the prompt charged particle density for η; in both cases, the GM scheme better describes the data. But, if we look at the p_T distribution of prompt charged particle density, then things get a little complicated because here the QCD-inspired scheme shows far better results than the other two schemes. However, the right choice of which scheme to implement is still under investigation.

4.8 Summary and conclusion

In this chapter, the workflow of the tuning framework implemented within the LHCb simulation was described along with the latest tuning efforts. The main idea of this work was to optimise the choice of parameters implemented in PYTHIA8.204, specifically related to the cross-section, flavour, colour reconnection, and MPI parameters. To obtain the optimisation of parameters, PROFESSOR with version 2.3.3 and RIVET with version 3.1.4 were used. The primary objective of this analysis was to verify the usage of RIVET and PROFESSOR software together embedded into the LHCb simulation framework. However, the values obtained for the optimised parameters are not final and are still under investigation. Additionally, tracking and monitoring of the timing for colour reconnection parameters are crucial, as previous results obtained with the colour-reconnection model demonstrate minimal χ^2 values. The tuning of multiparton interaction parameters remains challenging and complex due to their interdependencies on other parameters. Finally, the RIVET analyses are limited, and one needs to have measurements at higher energy levels to facilitate the tuning process for Run 3 data.

4.9 Comparison of models in PYTHIA and HERWIG with LHCb settings

Event generation at LHCb is predominantly governed by PYTHIA [66]; therefore, in this section, a comparison is made between these generators, keeping the data in account. The data is taken from the RIVET analysis.

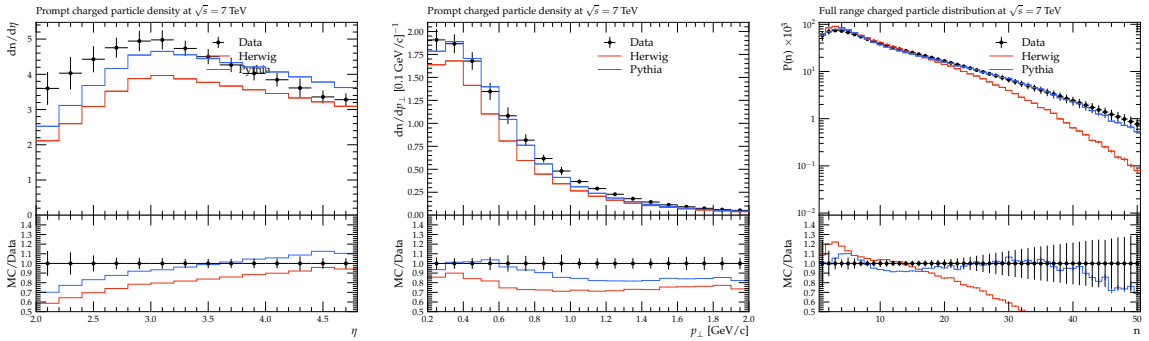
The major difference between the generators lies primarily in the hadronisation: the string model [77,78] and the cluster model [73,79]. The string model is based on the principle of linear confinement, where the potential between quarks at larger length scales is thought to arise from gluon self-interaction and is roughly linear. In this model, quarks split according to this potential to produce hadrons, resulting in well-modelled kinematics but a poorly predicted final state flavour description. PYTHIA uses the string model for hadronisation.

On the other hand, the cluster model, as discussed earlier, utilises pre-confinement, forming proto-hadrons that are independent of the hard process scale and then evolving these proto-hadrons via two-body decays into the final state hadrons. This approach results in a poorly modelled kinematics but a well-predicted flavour description [114]. HERWIG uses the cluster model for hadronisation. The difference in how these models are depicted in the two general-purpose event generators is illustrated in Fig. 3.3.

To evaluate the models, the LHCb measurements of charged particle multiplicity and density [110] in pp collisions at $\sqrt{s} = 7$ TeV with a low interaction rate provide an excellent test. In this case the visible events at the generator level are required to contain at least one charged particle within the pseudorapidity range $2.0 < \eta < 4.8$ and with a transverse momentum $p_T > 0.2$ GeV, a momentum of $p > 2$ GeV and a lifetime of $\tau < 10$ ps among other requirements. A reconstructed event must contain at least one track transversing all LHCb tracking stations as well as passing within 2 mm of the beam line and originating from the luminous region of the collision. The models used are both generated for one million minimum-bias events.

The results from LHCb are compared to simulations from PYTHIA and HERWIG. The plots in the Fig. 4.13 show the differential charged particle density as a function of pseudorapidity(η) (a), and p_T (b), and full range charged particle distribution (c).

For the η distribution (a), HERWIG underestimates the data for the entire η range, whereas PYTHIA underestimates the data at low η range, and overestimates at high η



(a) **(b)** **(c)**

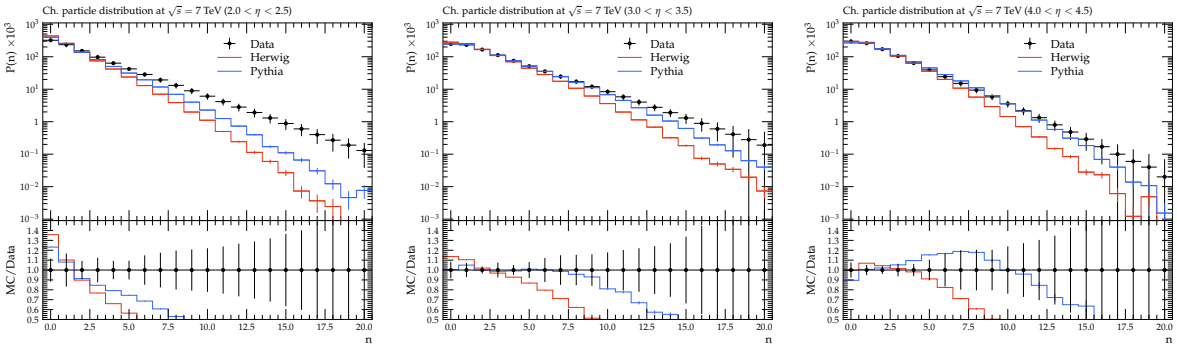
Figure 4.13: Charged particle density as a function of pseudorapidity η in (a), p_T in (b) charged particle multiplicity distribution in (c) at $\sqrt{s} = 7$ TeV for PYTHIA and HERWIG in solid lines compared with the LHCb data in points.

ranges, there is a similar trend observed between the two MC models. Fig. 3.6 seems to be slightly below unit, since there is any cut applied on p_T and momentum of the analysis that could be responsible for the difference. For the p_T distribution, PYTHIA is in good agreement with the data, whereas HERWIG again underestimates the data. This shows a sort of similar pattern to Fig. 3.4 (a), which is the charged distribution of hadrons with respect to the transverse momentum. If we could expand these plots within the same p_T regions, one could observe a similar pattern of PYTHIA dominating the HERWIG distribution. Lastly, for the full range charged particle distribution in Fig. 4.13 (c), it is quite evident from the plot that PYTHIA explains the data better than HERWIG, making PYTHIA a better choice of MC generator to be used at LHCb. The results validate the consistency of the MPI model at LHC energies in the forward region, with no unexpected behaviour observed.

Lastly, for the full range charged particle distribution (c), it is quite evident from the plot that PYTHIA explains the data better than HERWIG, making PYTHIA a better choice of Monte Carlo generator to be used at LHCb.

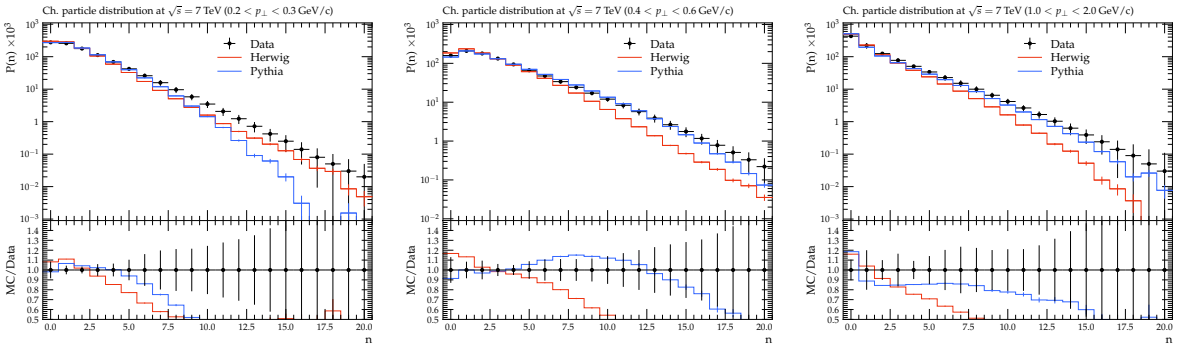
We could also address the charged particle distribution in regions of low, mid, and high η as shown in Fig. 4.14.

As it was quite evident from Fig. 4.13(a) that at low η values both the models underestimate the data, which can be seen in Fog. 4.14(a) as well, but as we move towards the high η values, i.e., Fig. 4.14(b) and (c), PYTHIA seems to be in better agreement with the data as compared to HERWIG. However, the agreement is only good



(a) **(b)** **(c)**

Figure 4.14: Comparison of MC models with LHCb data for the charged particle distribution in low ($2 < \eta < 2.5$) (a), mid ($3 < \eta < 3.5$) (b) and high ($4 < \eta < 4.5$) (c) pseudorapidity η regions at $\sqrt{s} = 7$ TeV.



(a) **(b)** **(c)**

Figure 4.15: Comparison of MC generators with the LHCb data for the charged particle distribution in low $0.2 < p_T < 0.3$ GeV/c (a), mid $0.4 < p_T < 0.6$ GeV/c (b) and high $1.0 < p_T < 2$ GeV/c (c) p_T regions at $\sqrt{s} = 7$ TeV.

with low charged particle multiplicity.

The charged particle multiplicity in p_T region low ($0.2 < p_T < 0.3$) GeV/c, mid ($0.4 < p_T < 0.6$) and high $1.0 < p_T < 2.0$ GeV/c is shown in Fig. 4.15.

In the low p_T regions and high charged particle distribution HERWIG is good at describing the data, but for higher momentum values HERWIG starts to underestimate the data, and PYTHIA is very well in agreement with the data.

From the charged particle multiplicity and density plots, it is quite evident that PYTHIA is in better agreement with the data as compared to HERWIG. The results validate the consistency of the MPI model at LHC energies in the forward region,

with no unexpected behaviour observed. HERWIG in most of the variables is constantly underestimating the data, while PYTHIA is in agreement with the data or is within the uncertainty. However, there is still some space for optimisation of the PYTHIA parameters, which would improve our results.

Chapter 5

Measurement of the charge hadron multiplicity in pp collisions at $\sqrt{s}$=13 TeV

"For every one billion particles of anti-matter, there were one billion and one particles of matter. And when the mutual annihilation was complete, one billionth remained – and that's our present universe." Albert Einstein

5.1 Motivation

LHC was designed to make discoveries and to confirm or extend the boundaries of the Standard Model. Thanks to LHC high energy and high luminosity, a broad physics program can be achieved simultaneously, spanning from soft interactions and studies of nuclear factors to the hard regime when flagship processes, like Higgs production, might be obtained (see Fig. 5.1).

High-energy hadron collisions at the LHC energies are extremely complex. There is are vast amount of accompanying particles produced, which come from parton showers or multiparton interactions, see the discussion in Chapter 3, The substantial part of these particles are soft, and propagate in the forward direction, often traversing through edges of tracking detectors without being reconstructed by the whole tracking system. The soft, non-perturbative regime of the strong interactions is not described by pQCD, and is modelled instead. In the forward regions, for low-transverse-momentum particles, one can see the highest differences in the predictions among the physics models.

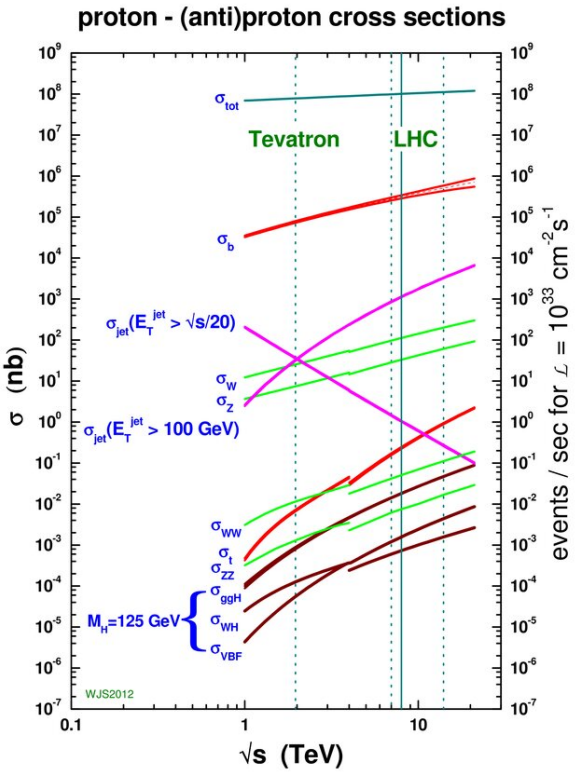


Figure 5.1: Cross-sections of proton-(anti)proton processes are shown as a function of centre-of-mass energy $\sqrt{s}$, with vertical dashed lines indicating energies reached or attainable by the Tevatron and LHC [115].

The theoretical predictions described in Chapter 3 must be cross-checked by the experimental results. In high-energy physics, predictions are made using models embedded in Monte Carlo (MC) simulations, which are then refined using experimental data. The main challenge in this verification is to obtain reliable detector corrections which would mimic the non-ideal performance of the tracking and identification subdetectors.

The main goal of this Chapter is to examine whether reconstructed minimum-bias events from LHCb can be used for online monitoring of the potential risk of radiation damage in VELO silicon sensors. To achieve this, LHCb data from Run 2 (2018) are analysed, and after comparison with MC, the flux of charged hadrons is obtained.

The analysis described in this Chapter considers the LHCb experiment's performance in the momentum resolution and particle identification to obtain multiplicity distributions. These distributions are instrumental in quantifying fluence for assessing radiation damage to detectors.

5.2 Machine configuration and data samples

5.2.1 Configuration of LHC and LHCb

Run 2 was dedicated to operating at the energy 6.5 TeV per beam with 25 ns bunch spacing. The first three months of 2015 focused on magnet powering tests and a magnet training campaign to establish a reliable and reproducible magnet performance at magnetic fields equivalent to 6.5 TeV beam energy. First beams were injected on Easter Sunday April 5^{th}, 2015, and within 12 hours both beams were captured, the first ramp to 6.5 TeV was performed successfully 8 days later and the first stable beams of Run 2 commenced on June 3^{rd}, 2015 [116]. The LHCb detector was fully calibrated, tested, and commissioned for Run 2 during the first six months of 2015. The LHC operation continued to 2018, and LHCb recorded an integrated luminosity of ~ 6.8 fb^{-1} [117] in Run 2, which adds to the luminosity in Run 1, exceeding 9 fb^{-1} of data collected by the LHCb experiment.

5.2.2 Data samples

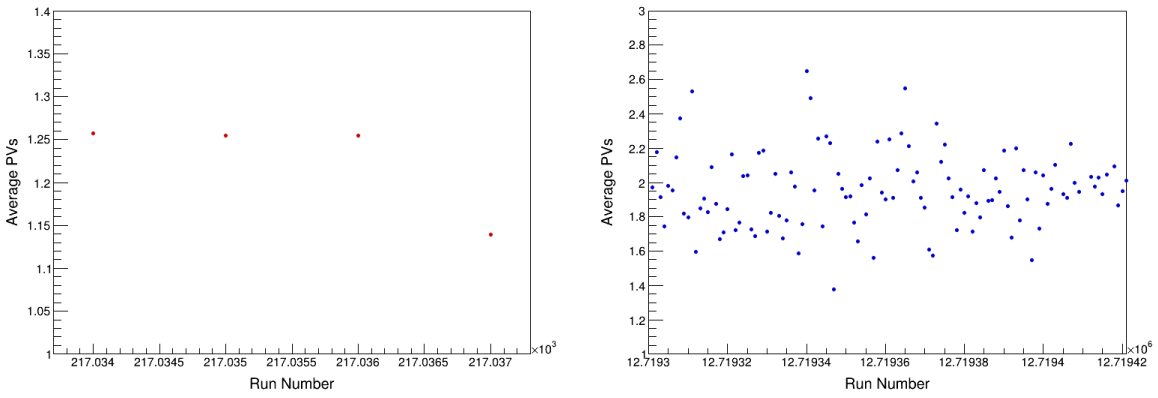
This analysis aims to analyse the data to calculate charged hadron multiplicity in dedicated LHCb runs. For this purpose, the data samples used were collected in October 2018 in pp collisions at $\sqrt{s}$= 13 TeV with VELO closed and magnet polarity up and an instantaneous luminosity of 9.1×10^{31} cm^{-2}s^{-1}. The properties of data and Monte Carlo (MC) simulations used in this analysis are shown in Table 5.1.

ν represents the average number of proton-proton (pp) interactions per bunch crossing, which is calculated for respective integrated Luminosity, total cross-section ~ 110.6 mb^{-1} [118] and the number of colliding bunches (N_{coll}) ~ 2332 with the bunch crossing frequency 11245 Hz.

For data, the average number of visible interactions in the detector acceptance per bunch crossing (μ) is 0.28. The probability of one visible pp collision calculated using the Poisson distribution P(n) is 26% and a pile up of 1.14.

Fig. 5.2 shows the average number of primary vertices for data and MC samples with respect to the run number. The MC sample shows the average PV around 1.992 ± 0.003, whereas for data, it remains much lower at an average of $\sim$1.25, suggesting the pile-up conditions in the MC are significantly higher than in data.

The data consists of 1,097,443 events and corresponds to a minimum-bias event sample with the requirement that at least one track is reconstructed. The current analysis



(a) **(b)**

Figure 5.2: Average number of primary vertices per run for events in pp collisions for Data (a) and MC (b).

is based on charged particles after the reconstruction and later is passed through loose preselection criteria (stripping)[1].

For the case of pions, the stripping line is designed to select and filter "loose pions". All particles that pass the reconstruction chain are considered "pions". Then, additional filters are used; these filters include checks to identify if the particles behave like electrons, kaons, muons, or specifically pions. Additionally, the selection criteria ensure that the selected particles are associated with primary vertices. The input data comes from charged particles that are part of the reconstruction process, and specific algorithms use this to process and select events.

5.2.3 Monte Carlo sample

The Monte-Carlo samples used in this analysis were produced with the LHCb tune Pythia 8 generator [63]. The LHCb default PDF setting, CT09MCS, was used with a value of α_s= 0.130 and p_{T0}^{ref}= 2.742289 GeV/c which are the running coupling constant value and the cut-off transverse momentum mentioned in Chapter 3. A brief outline of the theoretical models used in Pythia 8 MC generator is given in Chapter 3, and how the parameters used in those models can be tuned is mentioned in Chapter 4.

The LHCb simulation process takes generated events and mimics the current experi-

[1]The data produced by LHCb is large in amount and hence difficult to store. To make sure that the data is used efficiently by individuals, it is reprocessed into streams based on specific types of events required by different analysis working groups at the experiment. This is done via a process called stripping, allowing reconstruction of the standard basic particles with no additional selection.

mental conditions (instantaneous luminosity, beam profile, etc.). The detector simulation in this analysis was performed using the LHCb simulation package, further tailored to specifically match the conditions present during the 2018 low μ data taking period. Two minimum-bias samples were created, one for each magnet polarity, containing 502,464 and 500,814 events for magnet up (MU) and magnet down (MD) configurations, respectively. The average number of visible interactions in the simulated samples is $\nu = 1.6$, the probability of observing exactly one visible pp interaction is 67%. The properties of data and MC samples used for this analysis are summarized in Table 5.1.

Property	Magnet Up | Magnet Down
Center of mass energy (CoM)	13 TeV
Bunch Spacing	25 ns
Number of events (data)	1,097,443 | not available
ν (MC)	1.6
μ (data)	0.28
Number of events (MC)	502,464 | 500,814

Table 5.1: Properties of the data and MC sample used in the analysis.

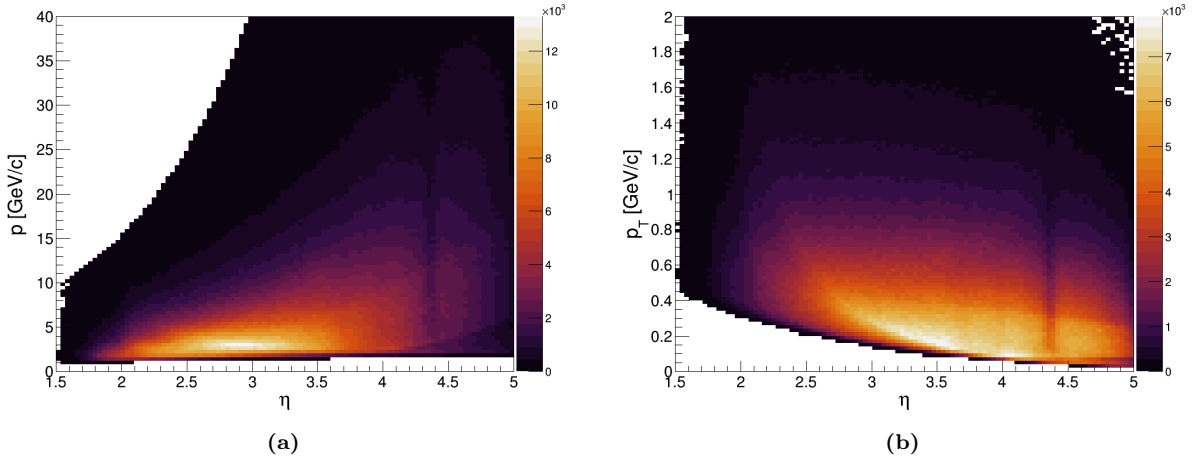
5.3 Event selection

The data and MC sample used in this analysis are described in sections 5.2.2 and 5.2.3. Prior to applying the selection criteria, it is essential to discuss the three types of selection cuts: detector level, generator level, and fiducial selection.

- **Detector level**: refers to the events as recorded by the detector, encompassing the data collected by the LHCb experiment. Similarly, all events generated by the MC that have undergone detector simulation are reconstructed.
- **Generator level**: refers to the MC events produced by the proton-proton process, before detector simulation is applied.
- **Fiducial selection**: involves the sets of cuts restricting event selections to a specific phase space. In this analysis, the fiducial acceptance is defined as $2.0 < \eta < 5.0$ and $2.0 < p < 100$ GeV/c in η and p space. The η cuts align with the detector acceptance, while the lower p bound is set due to the difficulty in reconstructing particles with momentum lower than 2 GeV/c and inefficient PID above 100 GeV.

5.3.1 Data selection

Since the data is to be used for fluence evaluation, we do not aim for the selection of perfect candidates as in searches for B decays. Instead, we need to obtain the flux of charged hadrons, with the best possible PID. The selection of minimum-bias events at the reconstruction level for both data and MC begins by requiring that events contain at least one well-reconstructed track with momentum p $<$ 100 GeV/c and transverse momentum p_T $<$ 4 GeV/c in the fiducial acceptance. To get rid of the background signals, the two basic cuts used are: the particles are not muons and the ghost probability $<$ 0.3, which would reduce most of the background and improve the efficiency. The distribution of particles in momentum and transverse momentum ranges with respect to the η is shown in Fig. 5.3.



(a) (b)

Figure 5.3: Momentum (a) and transverse momentum (b) distribution with respect to η for the charged particles.

The selection of the inclusive minimum-bias events at the reconstruction level for both data and MC begins by requiring how well a track fits with the primary vertex. This is controlled by a parameter PV χ^2/ndof, the lower values of this parameter would indicate a better fit, higher values would indicate that the particle does not originate from the primary vertex. The other variable for the primary vertex is ownPV χ^2, which is typically the primary vertex that best fits the particle track, as there may be multiple primary vertices in an event (e.g., due to pile-up from multiple proton-proton collisions). In general, a good PV fit is essential for accurate particle tracking.

Events are required to contain at least one reconstructed track with $p > 4$ GeV/c and

the track $\chi^2/ndof < 3$. There are several other cuts applied to improve the signal and reduce the background.

The detailed list of all the cuts used in this analysis is shown in Table 5.2.

Pre-selection cuts		
Category	**Variables**	**Pre-selection cuts**
Primary vertex	PV χ^2/ndof	< 1
	ownPV χ^2/ndof	< 1
Tracking	track χ^2/ndof	< 4
	trackmatch χ^2/ndof	< 40
Reduce background	ismuon	$= 0$
	ghost probability	< 0.3
Momentum	p	< 100 GeV/c
	p_T	< 4 GeV/c

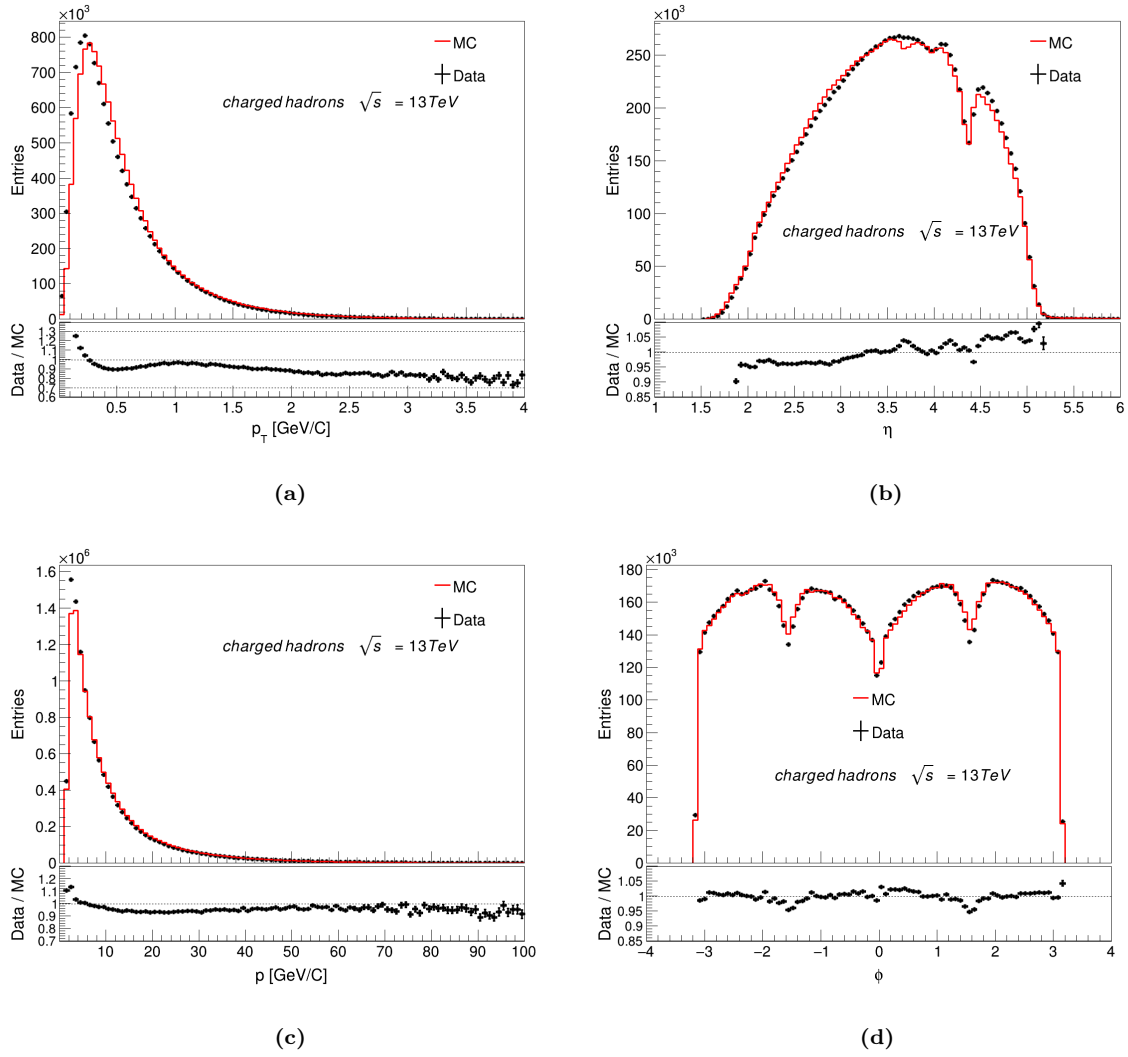
Table 5.2: Selection cuts

For tracking, we need to have a good quality of track fitting. track χ^2/ndof is calculated from the residuals (the differences between the measured and fitted positions) of the track, where lower values indicate a good fit, indicating consistency of the track model with the measurements. Whereas higher values indicate a poor fit. On the other hand, track match measures how well a track matches with a specific detector component or another track's expected trajectory. Different values for vertexing and tracking parameters were tested and are mentioned in Appendix A.0.1 and A.0.2, respectively.

Fig. 5.4 presents the kinematic distributions of charged hadrons after the selection, specifically transverse momentum, η distribution, momentum distribution, and multiplicity, as observed in both data and MC simulations. These distributions characterise the phase space in which the charged particles are reconstructed and provide insight into the typical momentum scales, angular coverage, and energy ranges relevant to the analysis. In addition to outlining the regions where most of the events are concentrated, they also help indicate where performance metrics such as efficiency and purity can be reliably evaluated, and where limited statistics or acceptance may affect their interpretation.

5.3.2 Detector effects

The detector response is also dependent on the occupancy, which is the fraction of detector channels (e.g., pixels, strips, etc.) that record a hit during a given event, but the particles can avoid reconstruction because of the high detector occupancy. Different

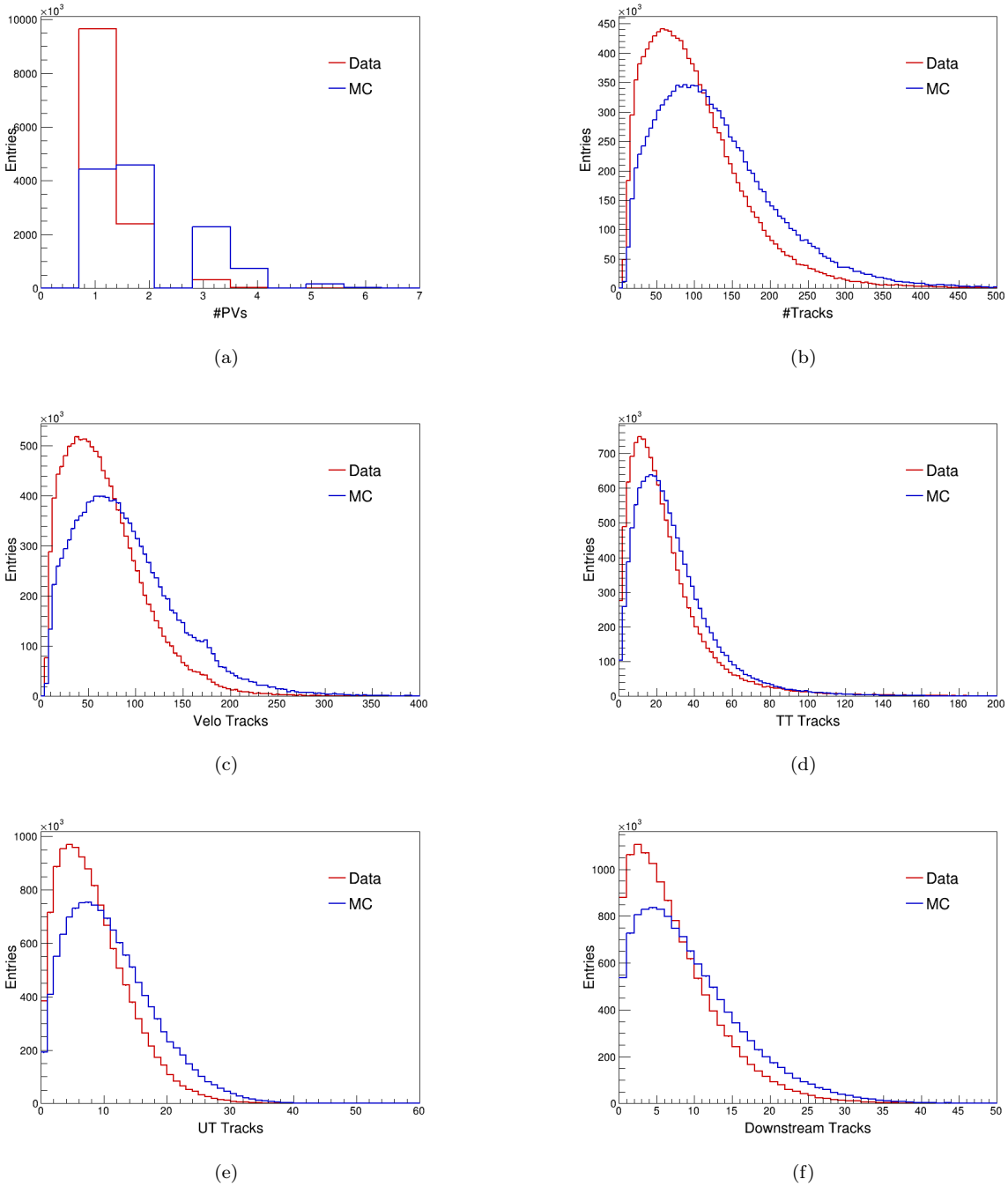


(a) **(b)** **(c)** **(d)**

Figure 5.4: Minimum-bias events for charged particles (a) p_T (b) η (c) p and (d) ϕ. The black points depict the data, and the red solid line depicts the MC distribution. The shape of η and ϕ distributions reflects the LHCb spectrometer geometry.

measurements of occupancy in the subdetectors of LHCb experiment can be considered, like the hits in the detectors, the number of clusters, or the number of tracks. Since this analysis revolves around the charged particles, which are reconstructed as long tracks and according to [119] these tracks go through all the tracking stations of the experiment, the occupancy of the tracking stations would be critical. The tracks and primary vertices for MC and data are shown in Fig. 5.5. The figure shows the number of primary vertices, and the total number of tracks are shown in (a) and (b), whereas the number of tracks in different subdetectors are like Velo, TT, UT and downstream tracks are shown in (c), (d), (e) and (f), respectively in data and MC samples. It is visible that the tracks in MC

overestimate the data tracks; therefore, an agreement with the data can be achieved by weighting the simulation events.



(a) (b) (c) (d) (e) (f)

Figure 5.5: Comparison of the number of primary vertices and tracks in different subdetectors for data and simulation samples in pp collision at 13 TeV.

In order to compute the weights for pp collisions at $\sqrt{s}$ = 13 TeV, different sets of variables were compared, and the one containing the number of reconstructed tracks

and the number of Reconstructed PVs were found to be the ones that gave a better description of the data.

Event-by-event weights were computed based on the ratio of the number of VELO tracks in the data to those in the Monte Carlo (MC) simulation. Specifically, the weights were defined as:

$$w = \frac{N^{\text{data}}_{\text{VELO tracks}}}{N^{\text{MC}}_{\text{VELO tracks}}} \tag{5.1}$$

These weights were then applied to the MC events to reweight the multiplicity distribution and better match the data. The effect of this reweighting can be seen in the multiplicity distribution of charged particles as shown in Fig. 5.6.

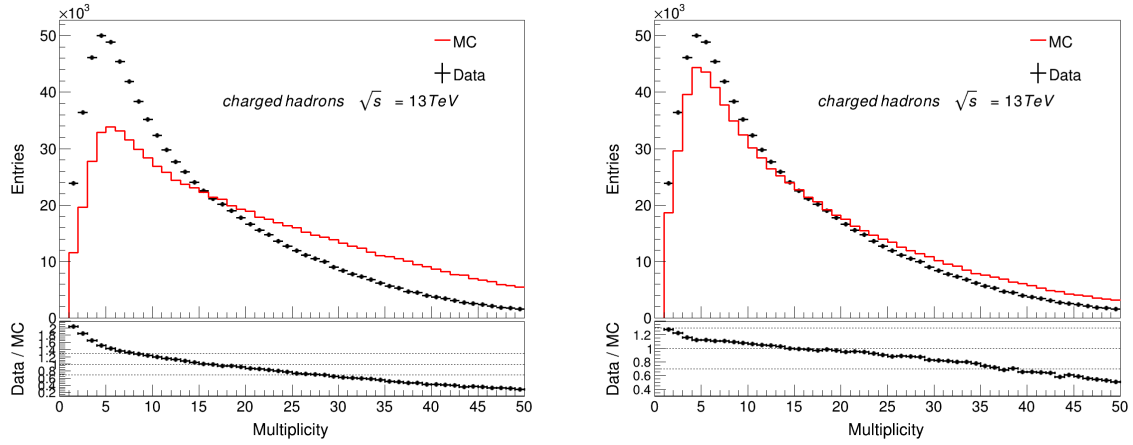


Figure 5.6: Hadron multiplicity distribution at $\sqrt{s}$= 13 TeV for pp collisions before (a) and after(b) applying the weights as mentioned in Eq. 5.1.

5.3.3 Selection efficiency

After applying the data selection cuts, the selection efficiency for MC samples, considering all the cuts, is calculated using Eq. 5.2:

$$\epsilon^{i}_{\text{sel}} = \frac{N^{i}_{\text{MC,pass}}}{N^{i}_{\text{MC,total}}}, \quad \text{where } i \text{ refers to pions, kaons, and protons.} \tag{5.2}$$

In the Eq. 5.2, $N_{MC,pass}$ refers to the number of true matched pions, kaons, and protons that satisfy the selection criteria with PID over $N_{MC,total}$, which refers to the total number of pions, kaons, and protons produced in the MC sample, respectively. The

selection efficiency is calculated using Eq. 5.2 and illustrated in Fig. 5.7 for the MC sample.

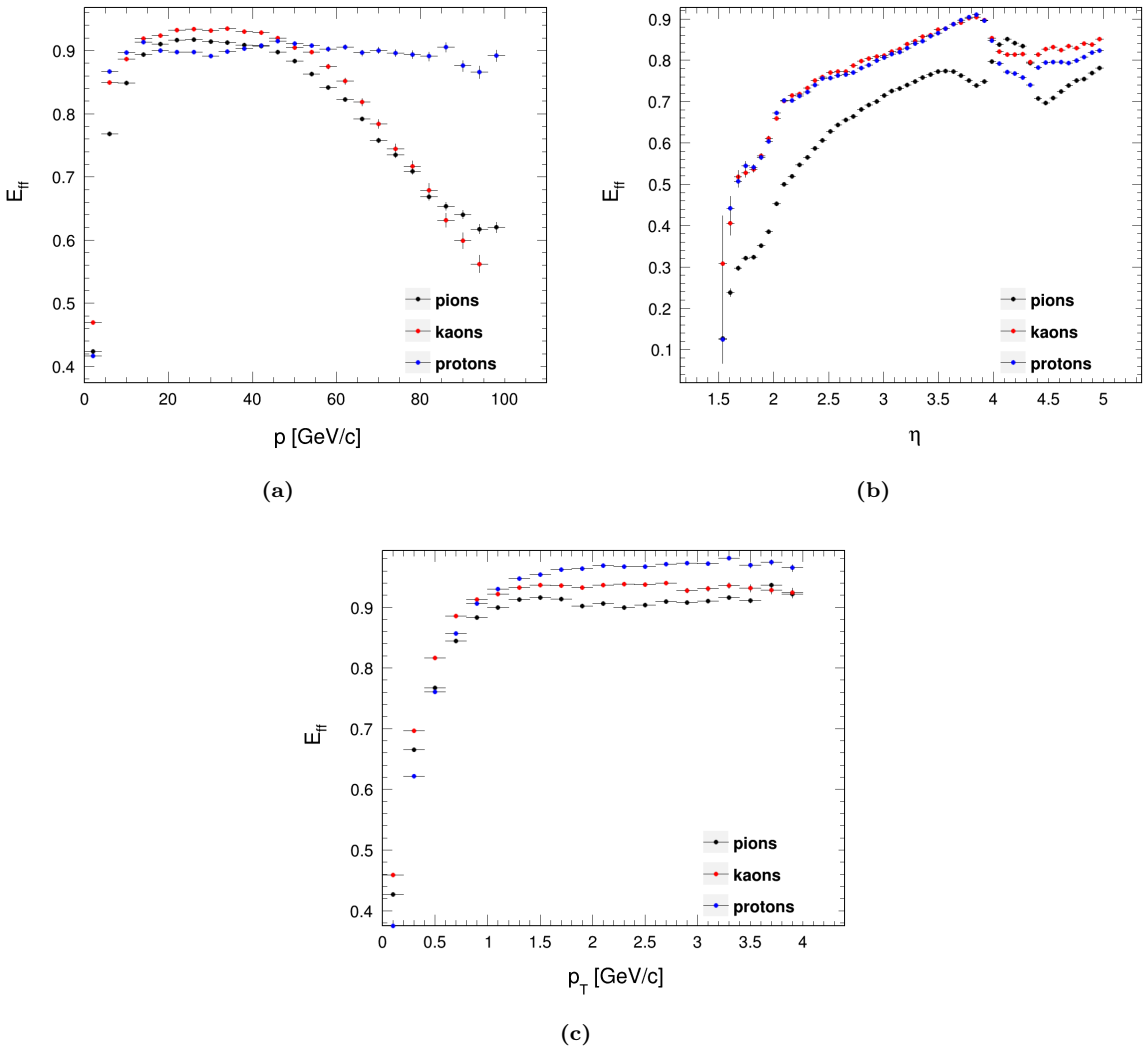


Figure 5.7: Selection efficiency comparison of MC with all the cuts to enhance the signal for different sets of charged particles for p (a) and η in (b) and p_T in (c).

The efficiency is around 90-95% at momentum around 50-55 GeV/c for all the particles as shown in Fig. 5.7 (a) and then decreases for higher momentum values, with some differences between particle types. At around p>50 GeV/c, the efficiency drops to 60% for pions and kaons. Whereas protons maintain an efficiency of 90% even at higher momentum values. Additionally, when examining the energy distribution for charged particles shown in Fig. 5.4 (c), we observe that the energy of these particles decreases beyond 50 GeV, which contributes to the decrease in selection efficiency. However, for the η distribution, in the low η ranges specifically η <2, these particles fall outside the

LHCb acceptance, hence the efficiency is low, and at very high η >4.5, there is reduced coverage in the forward-most regions of the calorimeters and muon system, also limited overlap between sub-detectors, reducing redundancy in track reconstruction. The central region is the most critical, ranging from $2.5 < \eta < 4.5$, where the tracking system is efficient, the efficiency for kaons and protons exceeds to about 90% and then drops to 70-80%, whereas for pions it is reduced to ~ 0.7. LHCb is not optimised for high p_T physics like ATLAS or CMS, but it still maintains good efficiency up to moderate p_T values ($\sim$ 4 GeV/c). The efficiency plateaus for all the charged particles up to 98% for protons, 94% for kaons and 90% for pions.

5.3.4 Particle identification and mis-identification

Particle identification is a crucial step in the LHCb experiment, achieved through the use of various methods and variables. To identify particles, the information from the sub-detectors is combined as described in Sec. 2.3.3. The first method for the purpose of particle identification is the change in log-likelihood that describes various PID hypotheses. There are often different types of hypotheses for multiple types of particles, each characterised by an expected observable. Given a set of measurements, the likelihood is calculated for each particle hypothesis. The likelihood function L expresses the probability of obtaining the observed data under each particle type's hypothesis. The change in log-likelihood is used to compare two competing hypotheses [120]. For instance if the observed data supports hypothesis H_1 over H_2, then the change in log-likelihood is:

$$\Delta lnL = lnL(H_1) - lnL(H_2). \tag{5.3}$$

Positive values would indicate that the data is more likely under hypothesis H_1, and negative values would indicate that the data is more likely under H_2. The absolute value of Δ ln L indicates how strongly the data favors one hypothesis over the other, larger values indicate stronger evidence [121]. When using this technique for the identification of particles, there is a specific algorithm in LHCb [122], designed to assign the most probable Particle ID (PID) to all reconstructed particles by integrating data from various sub-detectors, including the RICH detectors, calorimeters, and muon chambers.

The second method is based on neural network quantity, which is the output of multivariate techniques created by combining tracking and PID information (ProbNN). The ProbNN results are a single probability value for each particle hypothesis.

For the purpose of particle identification, both methods were tested to obtain the optimal selection with the cuts applied are mentioned in Table 5.3.

Both methods are tested at different cut-off parameter values. The tests for different values of PID and the comparison of efficiencies of different techniques for particle identification are shown in Appendix A.0.5 and A.0.7, respectively.

Particle Identification (PID)		
Particle type	**Cuts in Data**	**Cuts in MC**
Pions	PIDK < 0	PIDK < 0
Kaons	PIDK > 5	PIDK > 5
Protons	PIDp > 7	PIDp > 7

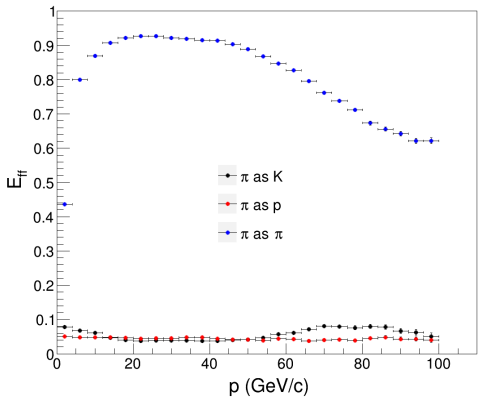
Table 5.3: Particle Identification (PID) cuts for data and MC.

Particle misidentification at the LHCb experiment is a critical factor influencing the accuracy of physics measurements, particularly in studies of rare decays and CP violation. LHCb employs an array of advanced detectors to distinguish among particle species such as pions, kaons, protons, electrons, and muons. Despite the high efficiency of these systems, misidentification can occur when a particle mimics the signature of another due to detector limitations or overlapping signal responses. The mis-identification of particles in the MC sample, for example, the pions treated as kaons in specific momentum bins, is formulated as:

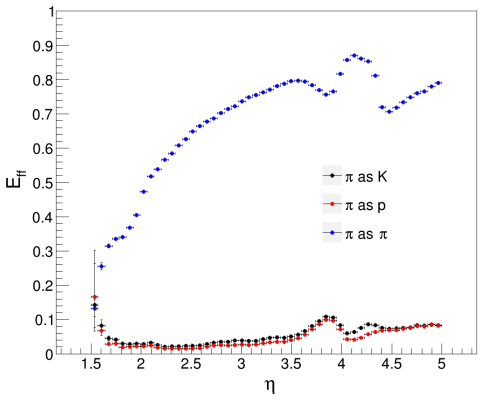
$$\text{MisID Rate}(\pi \to K, p) = \frac{\text{Number of } \pi^{\pm} \text{ misidentified as } K^{\pm} \text{ in momentum bin } p}{\text{Total number of particles identified as } K^{\pm} \text{ in momentum bin } p}$$

The misID efficiency plots with respect to momentum are shown in Fig. 5.8 (a), (c), and (e) and with respect to η are shown in Fig. 5.8 (b), (d), and (f), respectively.

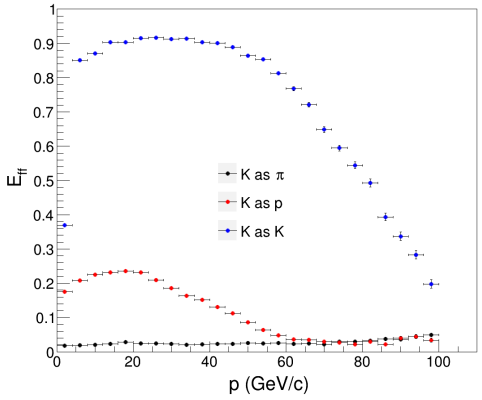
The rate of misidentification in the MC sample of pions as kaons (in black) and protons (in red) is plotted as a function of momentum and η in Fig. 5.8 in (a) and (b) for the MC sample. For the low momentum ranges typically < 20 GeV/c, the misidentification of pions as kaons is decreasing with increasing momentum; however, it remains uniform between 20 - 40 GeV and increases steadily above 40 GeV/c, indicating that at higher momentum ranges the pions can be misidentified as kaons. Whereas, there is a lower chance as seen by the plateaus, where the pions are treated as protons. For the low η ranges typically below 2 and above 4.5, both curves show high misidentification efficiencies, as these are outside the detector tracking capability. However, the fluctuations above $\eta >$



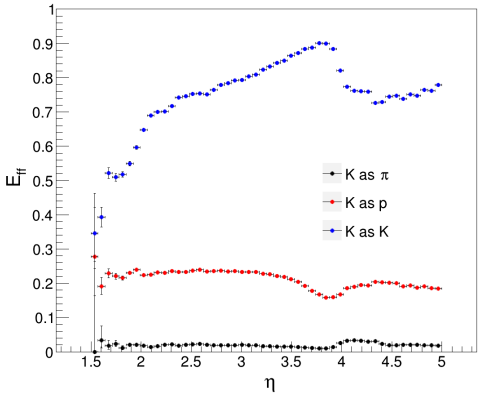
(a)



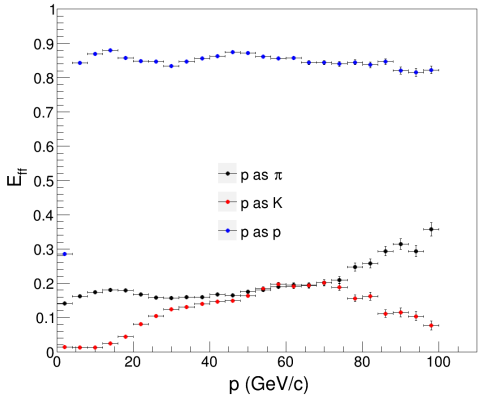
(b)



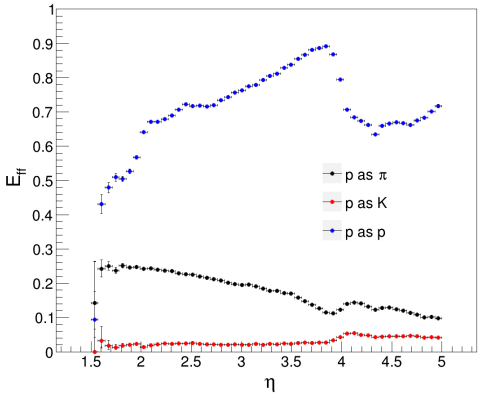
(c)



(d)



(e)



(f)

Figure 5.8: The distributions (a), (c), and (e) show the rate of misidentification of pions, kaons, and protons, each reconstructed as the other two particles based on their momentum. Meanwhile, the distributions (b), (d), and (f) display the rate of misidentification of pions, kaons, and protons, each reconstructed as the other two particles based on their η values in the MC sample.

3.5 occur in specific detector regions or at boundaries where tracking is limited. The slight upward bumps near $\eta = 3.5$ and $\eta = 4.5$ might indicate regions where the separation of particle types becomes more challenging. The overall lower efficiencies indicate good PID performance. The efficiency plots for kaons treated as pions and protons as a function of momentum and η are shown in Fig. 5.8 (c) and (d), respectively. At low momentum values, kaons treated as pions in generally low throughout the momentum ranges, and a slight increase is observed with increasing momenta. However, there is a 20% chance of kaons being treated as protons at low momentum ranges, which decreases with an increase in momentum. From the η distribution, it is visible that kaons are treated as pions with much smaller values as compared to kaons treated as protons. Misidentification of kaons as pions is lower because the separation between kaons and pions is better. The misidentification of protons as pions and kaons is shown as a function of momentum and eta in Fig. 5.8 (e) and (f).

Typically, below p $<$ 20 GeV/c and above 80 GeV/c, the misidentification of protons as pions is higher at lower momenta and increases slightly from 40-80 GeV/c and then increases at higher momenta values. However, for protons treated as kaons, the rate of misidentification is generally low at low momenta and increases from 20-80 GeV/c up to almost 20%, and then there is a decrease at higher momenta values.

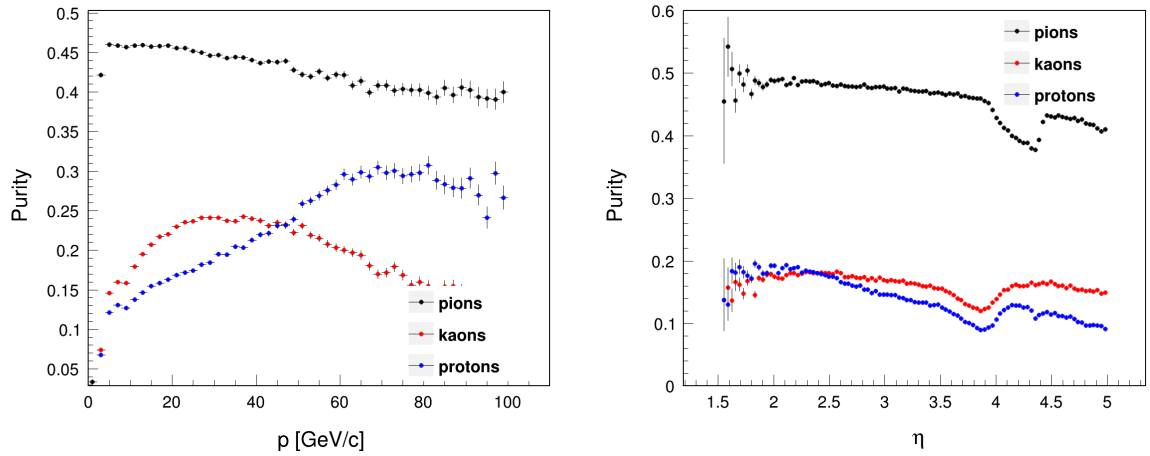
The misid rate as a function of η for protons treated as pions is almost 25% higher at low momentum ranges and shows a decreasing pattern with increasing values of η. Although a different pattern is observed for protons treated as kaons, which is generally low throughout the η ranges, a peak is observed in the $4 < \eta < 5$ range; which could be because of the geometry of the detector.

5.3.5 Purity

Particle Identification (PID) plays an important role in enabling the separation of particle species such as pions, kaons, and protons based on their unique signatures in detector subsystems. The concept of PID purity is defined as the fraction of correctly identified particles of a given type out of the total number identified as that type. The purity of PID is crucial for minimising background contamination in physics analyses and ensuring accurate reconstruction of particle decays. For a sample of particles identified as kaons ($K^{\pm}$), the purity can be defined as:

$$\text{Purity}(K, p) = \frac{\text{Number of correctly identified } K^{\pm} \text{ in momentum bin } p}{\text{Total number of particles identified as } K^{\pm} \text{ in momentum bin } p}$$

The methods employed to estimate and optimise PID purity are, including MC truth-matching. Results from these studies are presented, highlighting the achieved purity for specific particle types as a function of momentum and η shown in Fig. 5.9.



(a) (b)

Figure 5.9: PID purity efficiency comparison of MC for pions, kaons and protons for p (a) and η (b).

It is evident from the plot shown in Fig. 5.9(a) that at low momentum range typically $p < 20$ GeV/c, the purity of pions is around 46% and decreases with increasing momentum. However, for kaons and protons, it is lower at low momentum ranges and increases with increasing momentum. This could be due to saturation in the detector response or reduced separation power of PID variables. Purity with respect to η in (b) shows pions have the highest purity, i.e., around 50%, but reduces with increasing η values, since they are the most abundant particles, making them statistically easier to identify correctly. However, the decrease in purity of pions at $\eta = 4$ is mainly because of the η coverage of the LHCb experiment. Kaons and protons have a purity of around 20% and decrease with increasing values of η. Protons' PID is the lowest, likely because they are less abundant and have a higher likelihood of being misidentified as kaons or pions. Pions generally exhibit the highest purity across all η and p ranges, reflecting their statistical dominance and easier identification. The particle identification variables used for kaons and protons may not be

perfectly separable from those of other particles, especially pions, in the given momentum and pseudorapidity ranges. The purity of the different values for PID criteria, along with the purity of the sample for p_T and η is given in appendix A.0.3 and A.0.4, respectively.

5.4 Results

The objective of this analysis is to compare experimental data from Run 2 with MC simulations to examine whether reconstructed minimum bias events from LHCb can be utilised to evaluate the potential risk of radiation damage. By contrasting the observed data with simulated predictions, the analysis aims to identify areas where experimental results align with or diverge from theoretical expectations. To achieve this goal, the Run 2 data for a minimum bias sample is used. The sample is selected based on the criteria outlined in Section 5.2.2 and 5.2.3 for both data and MC, respectively. Additionally, data-selection requirements, as described in Section 5.3.1, where the cuts are applied on primary vertices, tracking, and also PID, are mentioned.

The results shown in the following section are a comparison of data and MC after applying selection and particle identification criteria. The plots show pions, kaons and protons distribution as a function of various parameters: η in Fig. 5.10, transverse momentum p_T in Fig. 5.11 and their multiplicity distribution in Fig. 5.13. Additional distributions are presented in the Appendix: ϕ in Fig. A.21 and total energy in Fig. A.22.

Fig. 5.10 shows the η distribution for comparison of data and MC. Fig. 5.10 (a) is the η distribution of pions, where MC overestimates the data between $2.5 < \eta < 3.5$. According to the description of the following analysis mentioned in [110] at 7 TeV, the MC distributions plotted with different Pythia models underestimates the data throughout the η ranges, however Pythia 8.180 describes the measurements significantly better than the previous versions. The pattern for the MC to underestimate the data at lower η values is similar to [110]. However, for $\eta > 4$, the data exceeds the MC distribution. For the case of kaons as shown in Fig. 5.10 (b), a slight disagreement between data and MC could be observed at high η values especially in between $3 < \eta < 3.6$, where MC overestimates the data and at $\eta > 4.3$ data exceeds, and for the case of protons, in Fig. 5.10 (c), MC overestimates the data in low η ranges, however, the data overestimates at high η values.

Fig. 5.11 (a) shows pions as a function of transverse momentum; the MC underestimates the data at low p_T scales and overestimates at high p_T scales. For kaons in

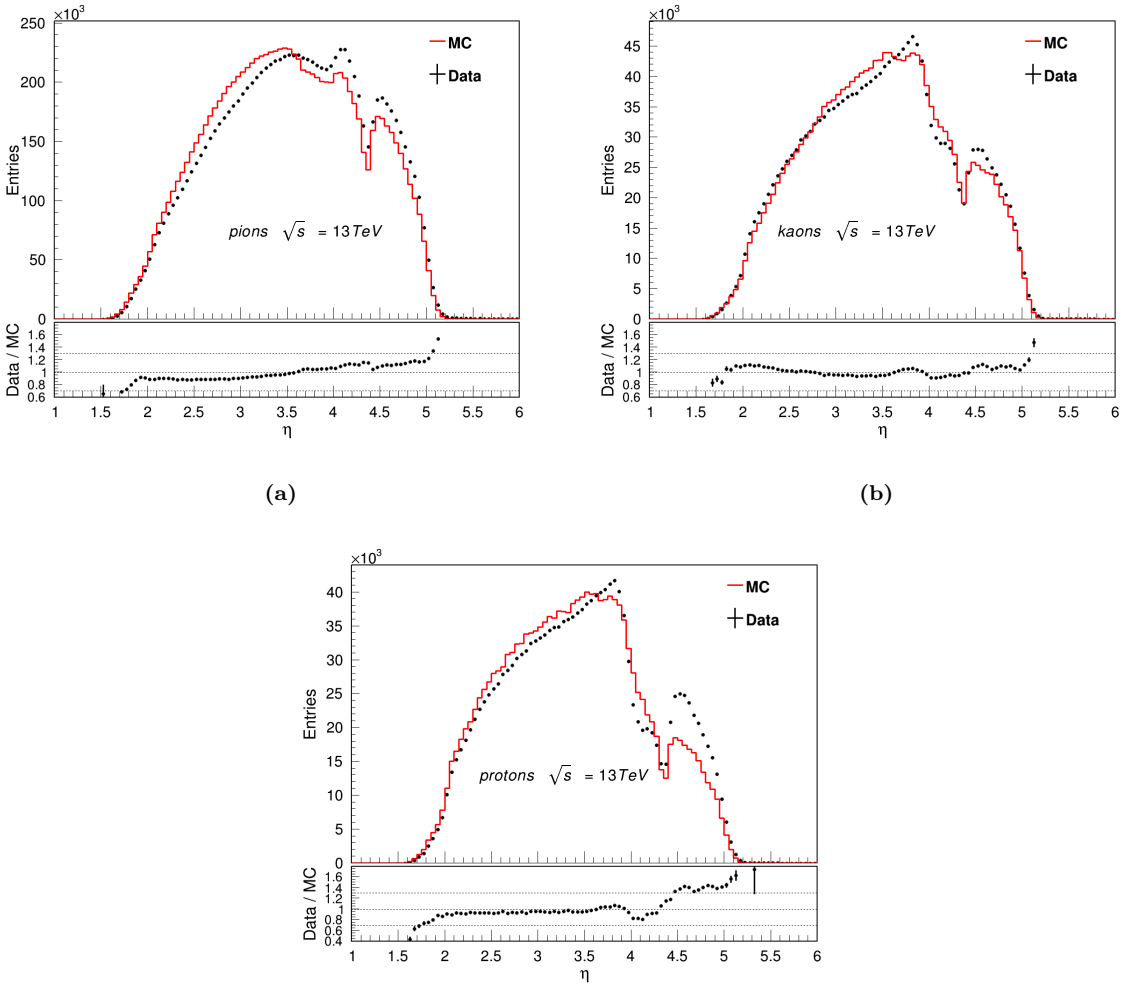


Figure 5.10: η distribution of minimum-bias events for pions (a), kaons (b), and protons (c). The black points depict the data, and the solid red line depicts the MC distribution.

Fig. 5.11 (b) the MC overestimates the data, at low p_T ranges, and for protons the MC underestimates the data at low p_T scale and overestimates at high momentum as shown in Fig. 5.11 (c).

The transverse momentum distribution mentioned in the analysis [110] has various models compared with the data at $\sqrt{s}$= 7 TeV. The pattern observed with the PYTHIA model 8.180 is similar to the one mentioned in Fig. 5.4 (a) for the charged hadrons multiplicity.

Fig. 5.12 (a) shows pions as a function of momentum; the MC underestimates the data at low momentum and overestimates at high momentum scales. Kaons follow a similar trend to that of pions as shown in Fig. 5.12 (b). The low momentum ranges

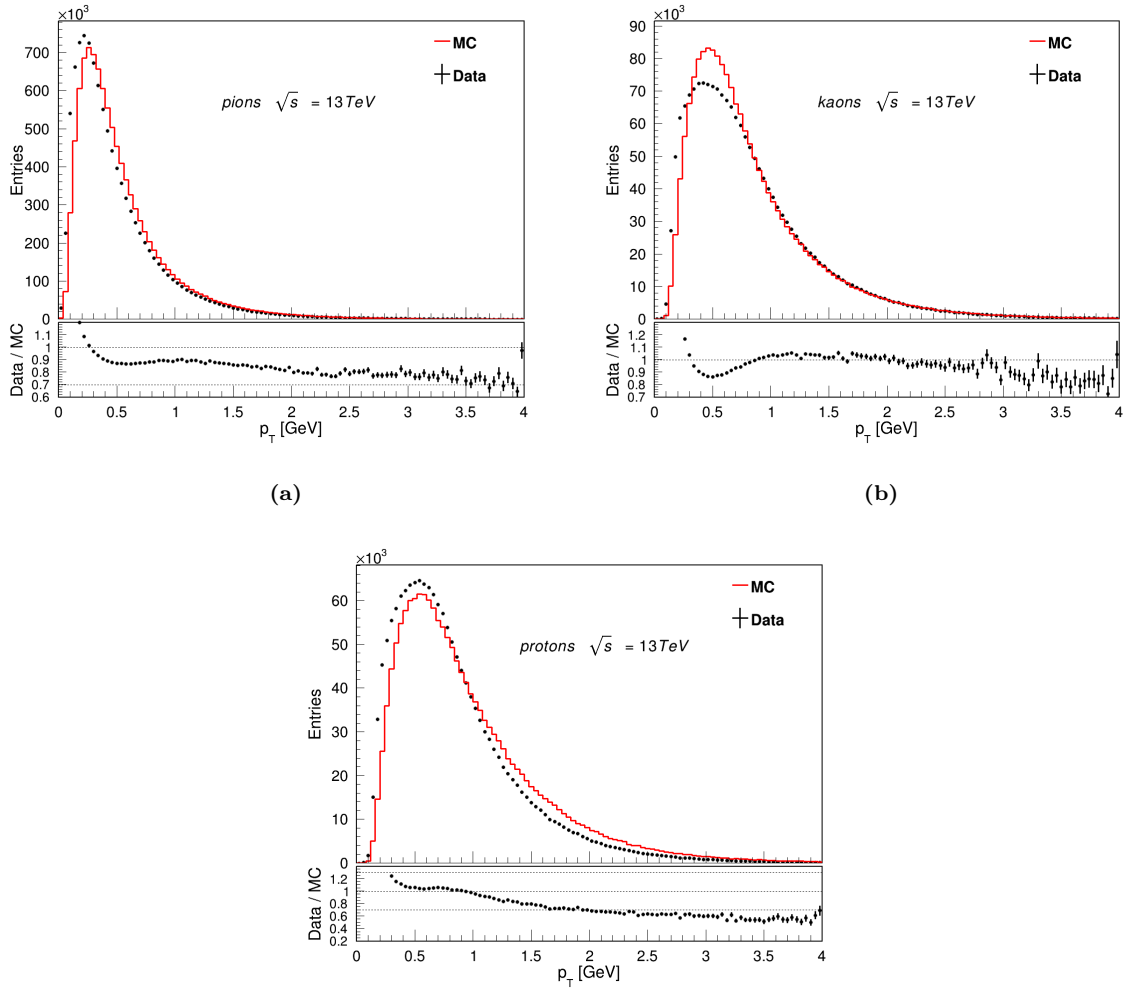


Figure 5.11: p_T distribution of minimum-bias events pions (a), kaons (b), and protons (c). The black points depict the data, and the solid red line depicts the MC distribution.

protons follow a similar trend as that of pions and kaons, however at high momentum the MC exceeds the data as shown in Fig. 5.12 (c).

Fig. 5.13 (a) shows the multiplicity of pions, (b) shows the multiplicity of kaons, and (c) shows the multiplicity of protons. There is a difference between the multiplicity of pions, which is underestimated in the MC sample at lower multiplicity, but overestimated at higher multiplicity values. However, for kaons, and protons, there is not a drastic difference between data and MC for the multiplicity of these particles.

Detector effects such as inefficiencies, miscalibrations, and resolution limitations can distort the measured quantities, leading to shifts in momentum, energy, or particle identification.

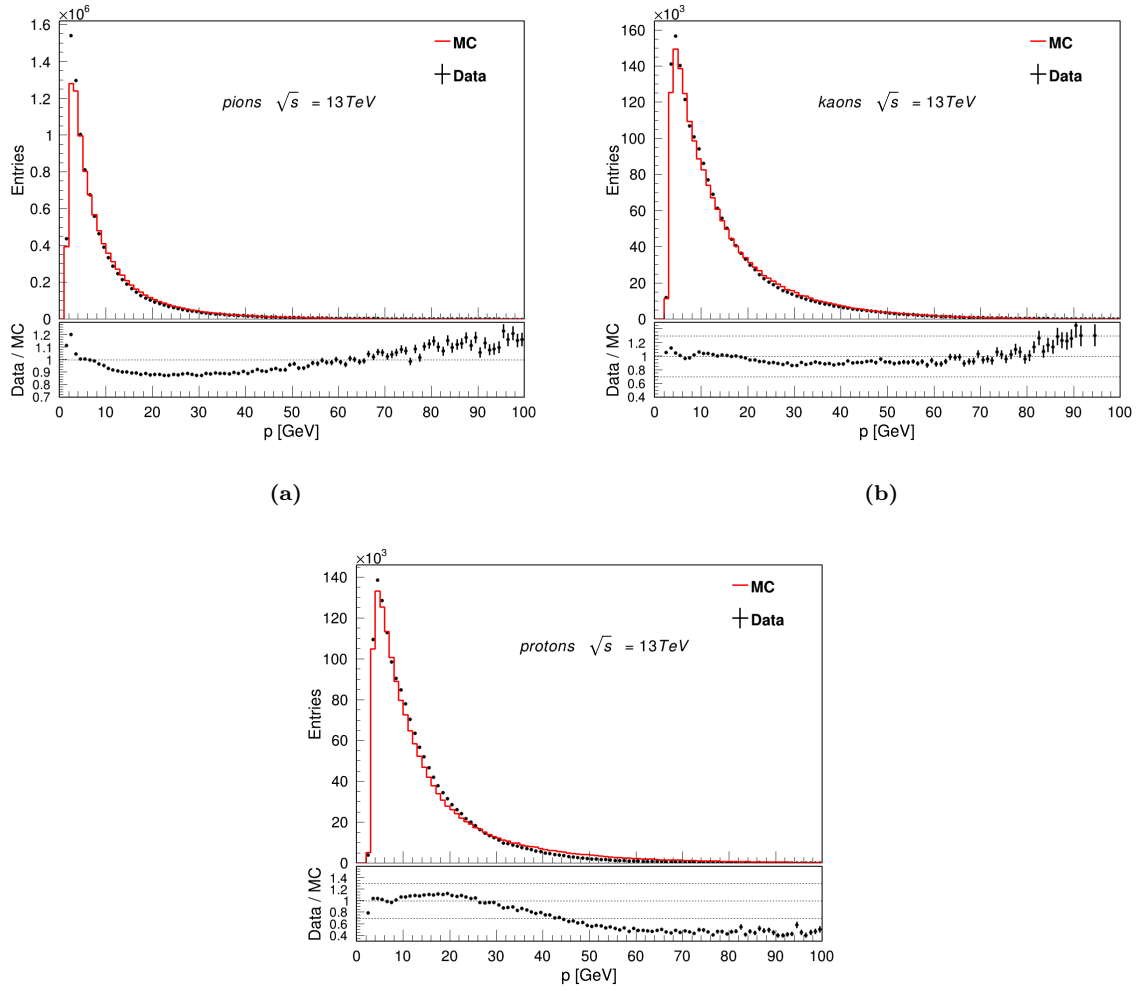


Figure 5.12: p distribution of minimum-bias events pions (a), kaons (b), and protons (c). The black points depict the data, and the solid red line depicts the MC distribution.

5.4.1 Estimation of the fluence in the VELO for Run 2

Once it is known how many charged hadrons traverse through the silicon sensors per one pp interaction at LHC, what is regarded as multiplicity, and the type of hadron is known, one can estimate the neutron equivalence fluence ϕ_{eq} in a data-driven way.

The distribution of kinetic energy of pions, kaons, and protons is shown in Fig. 5.14. In case of neutrons for this study, it was assumed that the same number of neutrons and protons are produced, according to the study shown in Chapter 3.8.

Although there are discrepancies between the data and the simulation in the Fig. 5.13, especially in the case of heavier hadrons, the number of pions is more than two orders

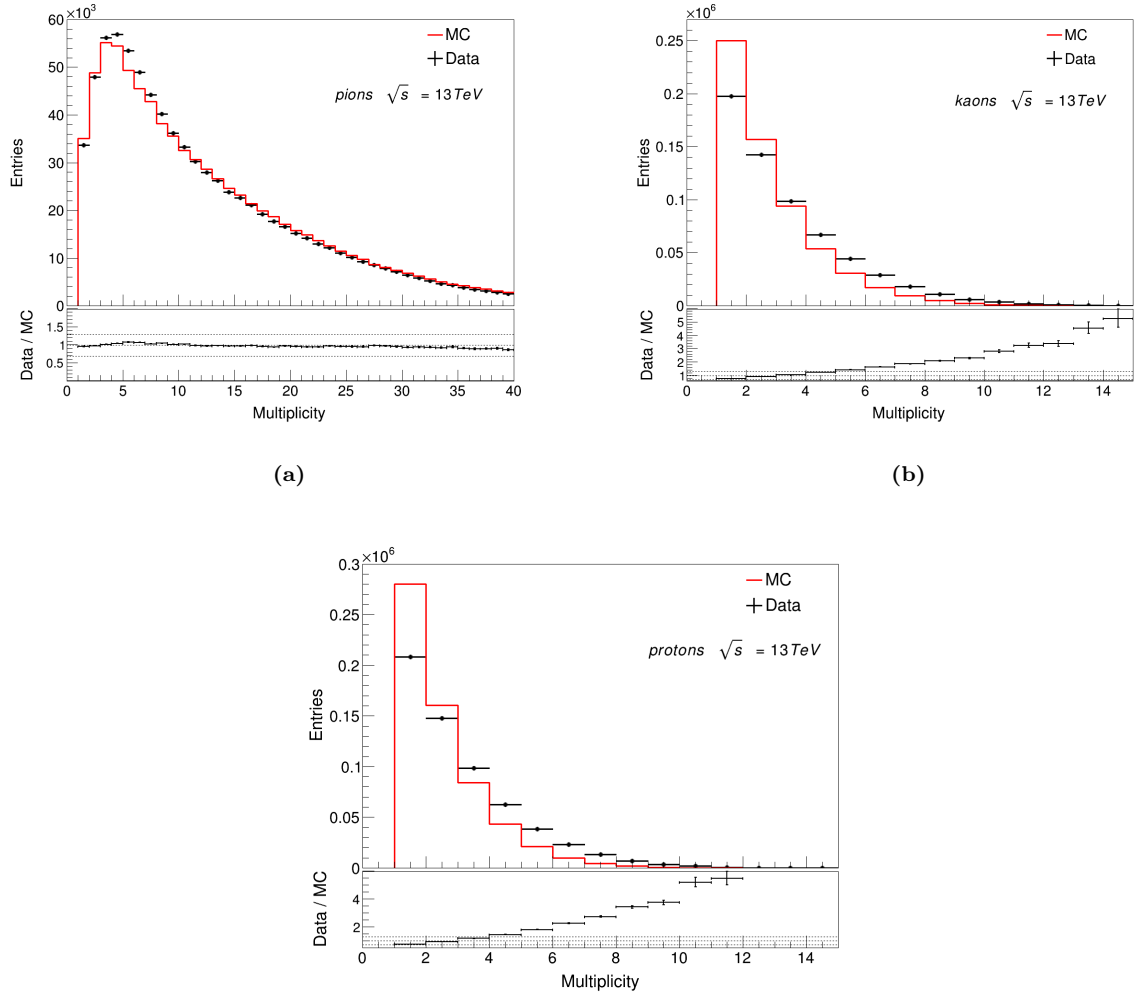


Figure 5.13: Multiplicity distribution of minimum-bias events for pions (a), kaons (b), and protons (c). The black points depict the data, and the solid red line depicts the MC distribution.

of magnitude higher than protons and kaons. Since kaons and protons are produced less frequently, their higher damage power results in a more substantial contribution to radiation damage.

Taking into account the hadron kinetic energy distributions, combining them with the respective damage function, according to the formula 2.2, the fluence ϕ_{eq} might be estimated. Since the VELO radiation field is highly non-uniform (this is known from FLUKA simulation [59]), the fluence is calculated in pseudorapidity bins, see Fig. 5.15.

It is worth remembering here that this analysis is based on *long* tracks, which are reconstructed with a basic condition that a track is reconstructed as it goes through

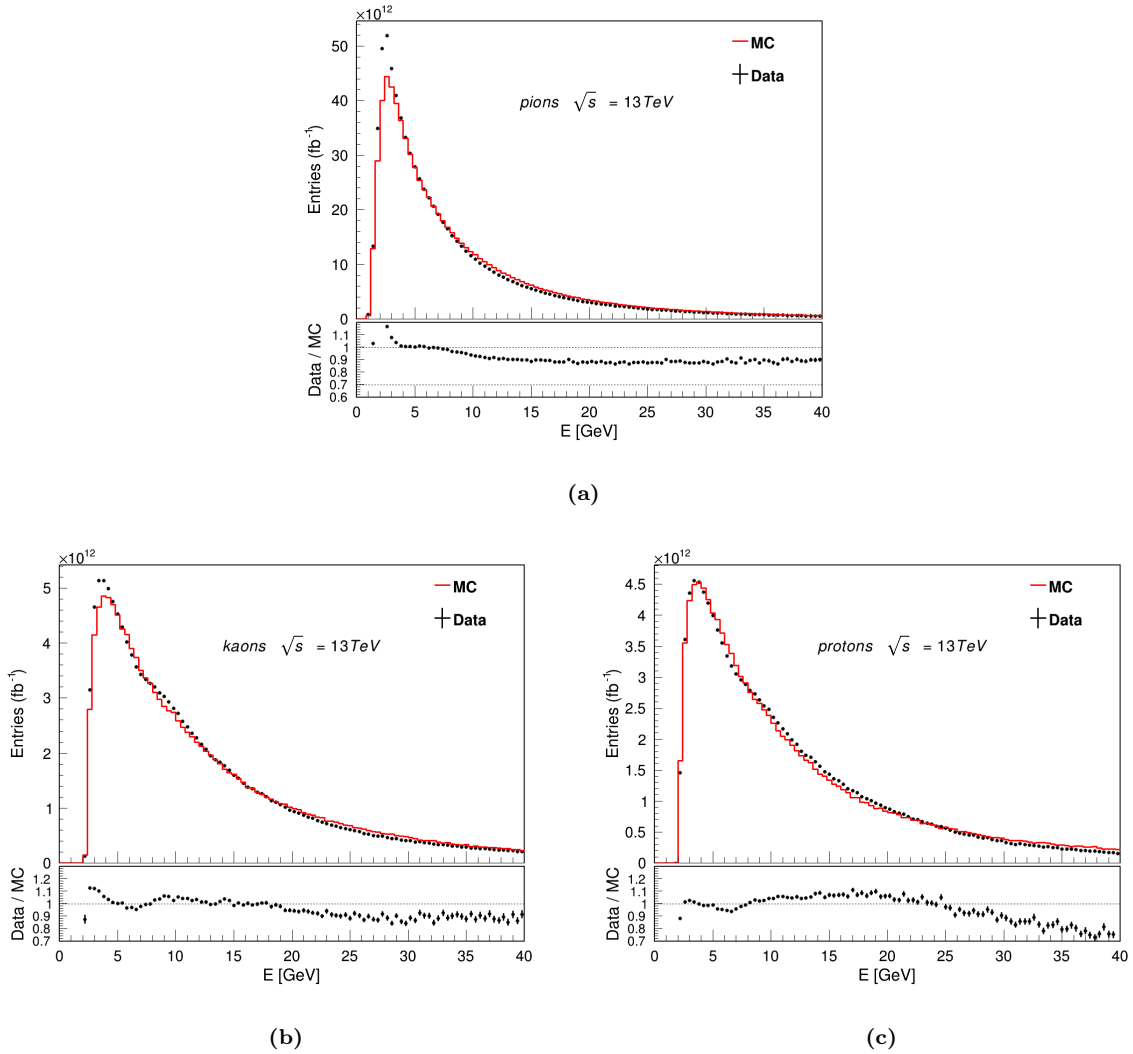


Figure 5.14: Kinetic energy distributions of minimum-bias events for pions (a), kaons (b), and protons (c). The black points depict the data, and the red solid line depicts the MC distribution. Number of events corresponds to 1 fb^{-1} of integrated luminosity.

at least three VELO stations. Therefore, in this study, the VELO acceptance region is sliced into pseudorapidity intervals which correspond to the edges of the acceptance as indicated in Fig. 5.16.

Eventually, the fluence ϕ_{eq} is calculated for a few representative VELO sensors. They are situated at $z = 0.5; 9.5; 43.5; 75.5$ cm (the closest to the IP, the middle two, and the farthest from the IP). The results are shown in the Table 5.4.

The values of ϕ_{eq} obtained with the use of reconstructed long tracks are lower than the values obtained in FLUKA simulation. This is caused, among many other reasons, by the non-ideal reconstruction efficiency and lack of information about secondary particles.

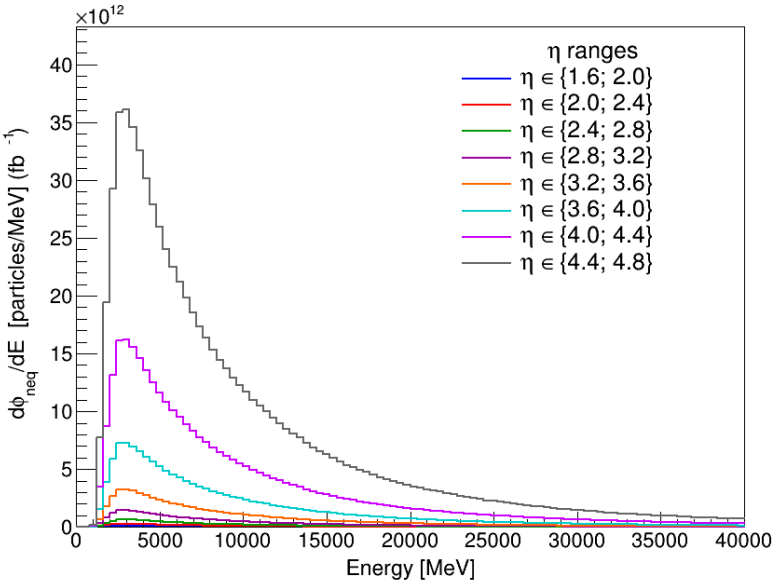


Figure 5.15: The energy spectrum of fluence ϕ_{eq} calculated for 1 fb^{-1} of integrated luminosity for pp collisions at $\sqrt{s}$ = 13 TeV.

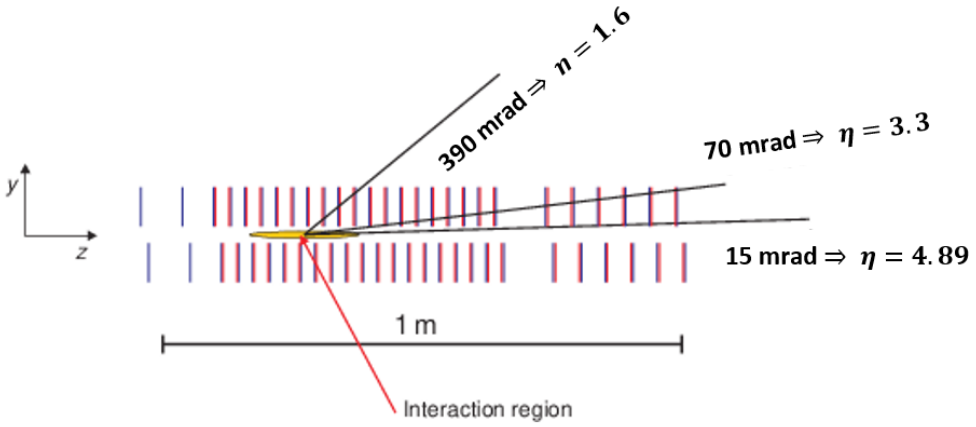


Figure 5.16: View of the yz plane showing the VELO modules and the pp interaction region along with the various η ranges. This figure illustrates in which modules long tracks might have been reconstructed, therefore, it defines the boundaries of the application of this method for the ϕ_{eq} estimation.

Although this is a rough estimation of ϕ_{eq}, it might be useful for comparison of different samples, especially in case of any disturbances or accidents during data taking.

Sensor positions [cm]	ϕ_{eq} [particles/cm^2]
z = 0.5	$(4.96 \pm 1.49) \times 10^{12}$
z = 9.5	$(4.10 \pm 1.23) \times 10^{11}$
z = 43.5	$(2.10 \pm 0.632) \times 10^{10}$
z = 75.5	$(6.59 \pm 1.97) \times 10^{9}$

Table 5.4: Fluence measured at various sensor positions along the z-axis. The numbers correspond to the 1 fb^{-1}.

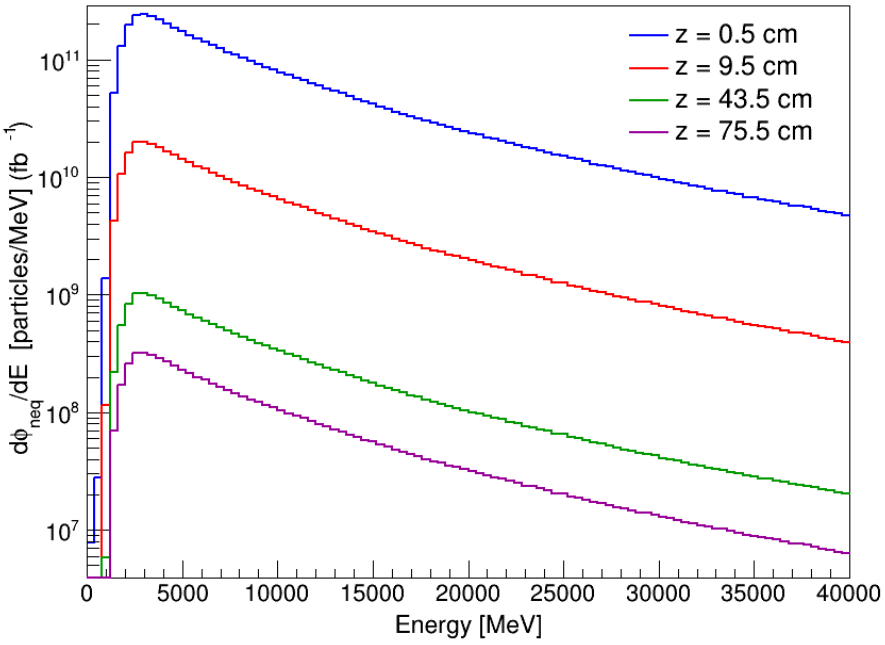


Figure 5.17: Fluence ϕ_{eq} in the representative VELO sensors calculated for 1 fb^{-1} of integrated luminosity for pp collisions at $\sqrt{s}$ = 13 TeV.

Chapter 6

Particle fluence for radiation damage prediction in LHCb silicon trackers

There is no safe amount of radiation. Even small amounts do harm - Linus Carl Pauling (1995).

6.1 Motivation

In particle physics experiments, efficiently tracking charged particles is a critical step for event reconstruction, enabling the identification and characterisation of fundamental interactions. The precise determination of particle trajectories is achieved using advanced detector technologies, usually employing solid-state detectors and gaseous detectors. Among these, semiconductor detectors, especially the silicon-based ones, are highly valued for their superior spatial resolution, fast response, and robustness against high interaction rates. These attributes make them crucial in modern high-energy physics experiments.

At LHCb, silicon-based detectors are strategically deployed in areas demanding the highest spatial precision, particularly near the interaction point where particle densities are highest. This region is instrumented with VErtex LOcator (VELO), introduced in Section 2.3.1, which plays an important role in precise tracking and vertex reconstruction. The VELO and other silicon trackers are integral to the LHCb tracking system, as outlined in Section 2.6, enabling efficient separation of primary and secondary vertices and contributing significantly to the overall physics performance of the experiment.

However, the performance of silicon detectors is not immune to the challenging

operational environment of the Large Hadron Collider (LHC). High radiation levels from intense particle interactions lead to radiation-induced damage in the silicon sensors. This damage manifests in multiple forms, including increased leakage current, charge trapping, and changes in the depletion voltage, all of which degrade the detector's efficiency and resolution over time. Understanding these effects is critical for designing future detectors and optimising existing ones, as discussed in Section 2.7.

In this context, particle fluence studies are essential for predicting radiation damage. These studies quantify the flux of particles traversing the silicon sensors and provide a framework for understanding the cumulative radiation exposure and its impact on detector performance. This chapter thoroughly examines the upgraded VELO detector, including its design and operation. Further, it focuses on the problem of particle fluence in the radiation damage monitoring during data-taking with a use data-driven method.

6.2 Vertex Locator

The most complicated and delicate part of the LHCb experiment is VELO. The new system, after the first upgrade, consists of silicon pixel sensors, increasing the effective number of channels from 180k to almost 41M. The full coverage of LHCb acceptance, in the pseudorapidity range $2 < \eta < 4.5$ is achieved with a series of 26 stations placed along the beam direction. The conceptual layout of the detector within the LHCb coordinate system is shown in Fig. 6.1.

Each VELO station consists of L-shaped VELO modules, which are shown in Fig. 6.2(a). The two retractable L-shaped VELO modules enclose the interaction point, this configuration is shown in Fig. 6.2(b). The detector modules are distributed across multiple stations, as depicted in Fig. 6.2(c), enabling comprehensive coverage of the interaction region. Keeping in mind the increased data rates for Run 3, each VELO module is populated with 41 million hybrid radiation-hard pixel sensors, which are designed to withstand extreme radiation levels and high hit rates. These silicon pixels have a pixel pitch of 55 μm. Each of these sensors is read out by 3 VeloPix ASICs with 256×256 pixels [124]. Fig. 6.3(a) shows the dimensions of the sensor and pixel layout underlying the ASICs. Due to unusual placement of the VELO modules, perpendicular to the LHC beams, sensors are designed to handle highly non-uniform fluences of about ϕ_{eq} = 8× 10^{15} n_{eq} cm^{-2}. Additionally, the sensors are designed to be kept at temperatures below -20°C, which would further safeguard them against the effects of radiation damage.

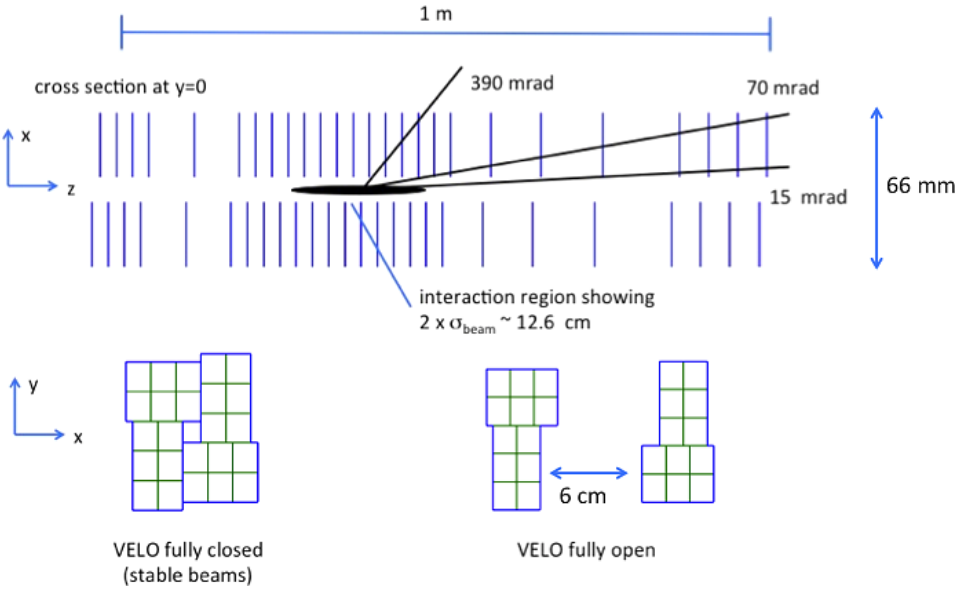
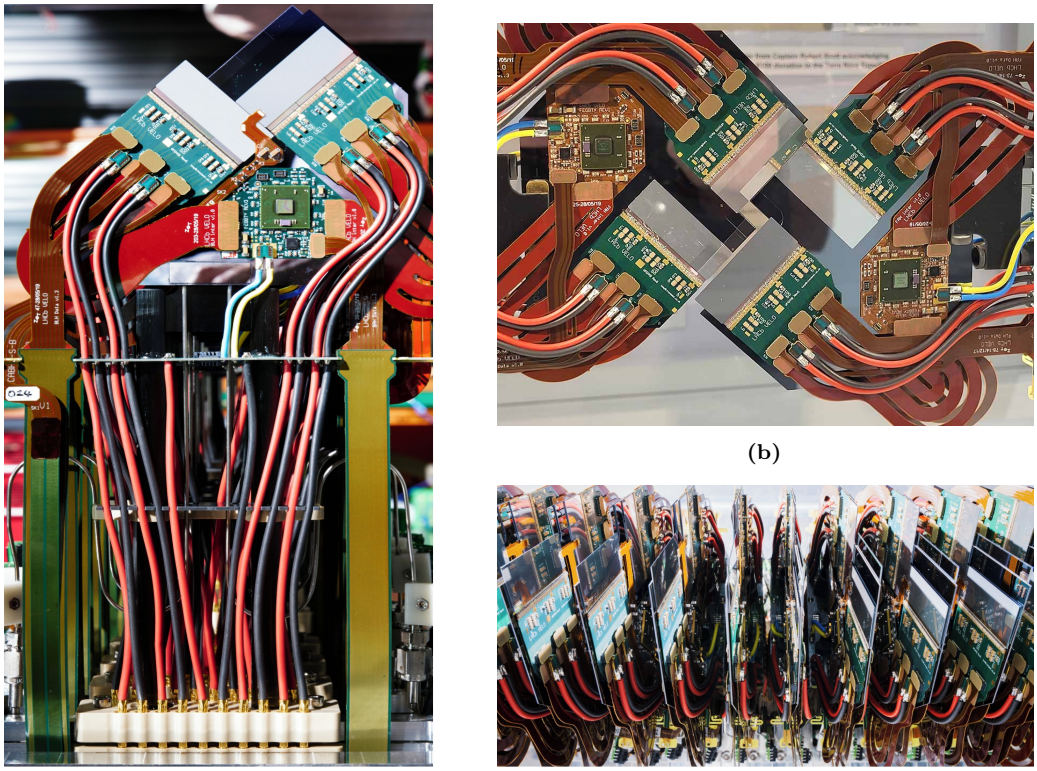


Figure 6.1: Schematic diagram of the LHCb VELO showing the cross-section at y = 0, together with the z extent of the luminous region and the nominal LHCb acceptance. The bottom figure shows the VELO module positions in both the closed (during LHC beam operation) and open positions, are illustrated in the xy plane [123].

The adoption of the pixel-based detector has led to faster reconstruction algorithms, improving hit resolution and vertex reconstruction, and also to a reduced occurrence of ghost tracks. These advancements have significantly enhanced the overall accuracy of VELO. A comparison of the impact parameter resolutions between the upgraded and previous VELO is presented in Fig. 6.3(b) [124].

The new front-end electronics of the VELO are designed to accommodate higher data rates and enhance track reconstruction precision. The retractable transmission lines, as illustrated in Fig. 6.4, are built to be flexible and move in harmony with the module. This design ensures smooth integration and adaptability within the VELO system, facilitating efficient data transfer. These lines utilize aerospace-grade dielectric technology, tested for radiation resistance, which is similar to that used in NASA's spacecraft [126]. VELO operates in a secondary vacuum, which is a requirement set by LHC to prevent contamination of the main LHC vacuum in the event of a leak. This introduces the need for an aluminum foil barrier (RF shielding) to help reduce electromagnetic interference and ensure stable operations.



(a) (b) (c)

Figure 6.2: Illustrations of VELO modules and their arrangements: (a) single module, (b) retractable L-shaped modules, and (c) modules placed in stations.

6.3 Radiation damage in LHCb VELO

VELO operates under extreme conditions close to the point of *pp* interaction, which is just 5.1 mm from the sensors. This proximity makes it more susceptible to intense radiation. But the sensors are equipped with radiation-hardened technology designed to handle such conditions as mentioned in 6.2. The particle flux is highly uneven, with a strong variation depending on both the radius and the z-position of a station. Due to the proximity of VELO to the interaction point, the primary source of particle radiation comes from the prompt production of hadrons (pions, protons, and neutrons) during proton or ion collisions [127].

During irradiation, the electrical properties of silicon undergo significant changes, and new energy levels are induced in the silicon band gap. Also, the charge carriers produced by passing particles can become trapped in defect-induced trapping centres, which could result in a delay at the signal collection electrodes. Overall, the impact of particle radiation leads to three major effects: increased leakage current, altered effective

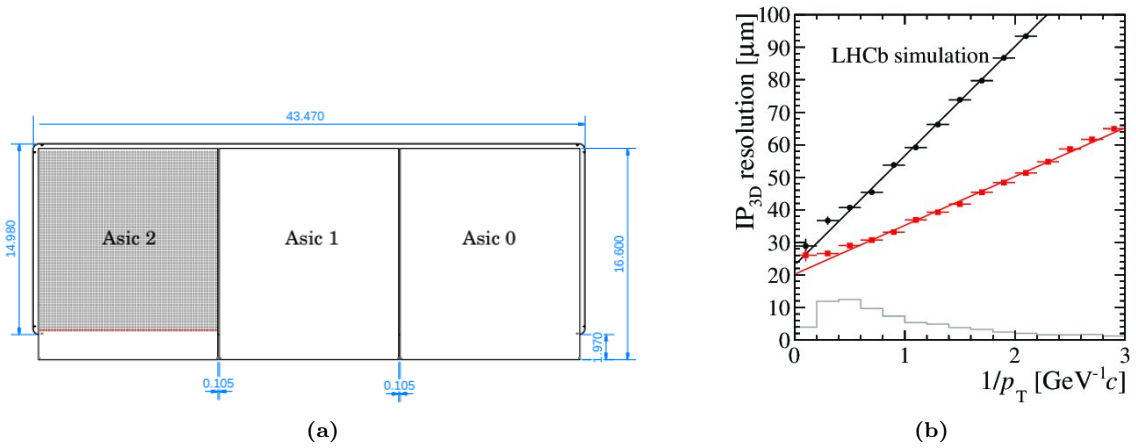


Figure 6.3: (a) shows there are 256 × 256 active bonded pixels [123] and (b) shows the impact parameter (IP) resolution for long tracks traversing the LHCb detector as a function of the inverse of transverse momentum. The gray histogram shows the relative track population in each $1/p_T$ bin. [125].

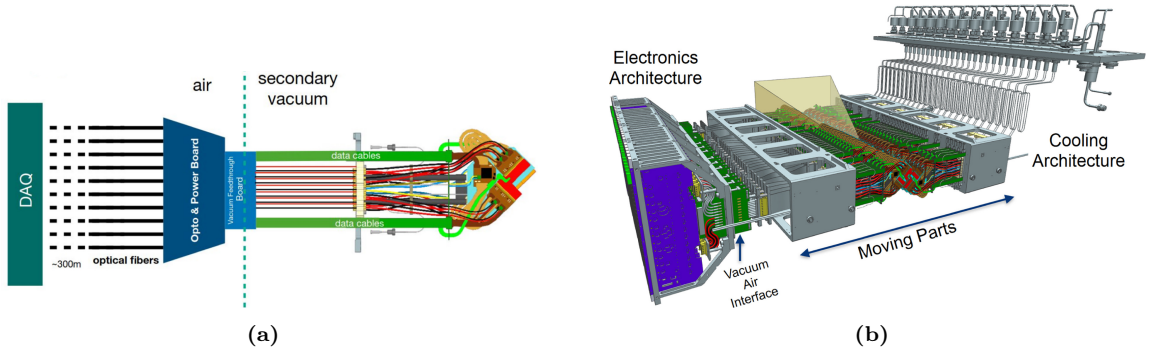


Figure 6.4: The module with DAQ system (a) and 3D model of modules, cooling and electronics architecture of VELO modules (b) [126].

doping concentration (which affects the operating voltage required for full depletion), and reduced charge collection efficiency [60]. These effects build up over time and could lead to deterioration of the detector's performance. Therefore, it is essential to monitor radiation damage in VELO sensors regularly to maintain optimal operating conditions. Since it is not possible to remove the sensors entirely, three monitoring methods are commonly used: current-voltage (IV) scans, current-temperature (IT) scans, and charge collection efficiency (CCE) scans.

Particle irradiation of silicon leads to damage at both the surface and bulk levels. The main cause of bulk damage is the displacement of atoms in silicon from their lattice, which would increase the leakage current. The leakage current varies linearly with fluence,

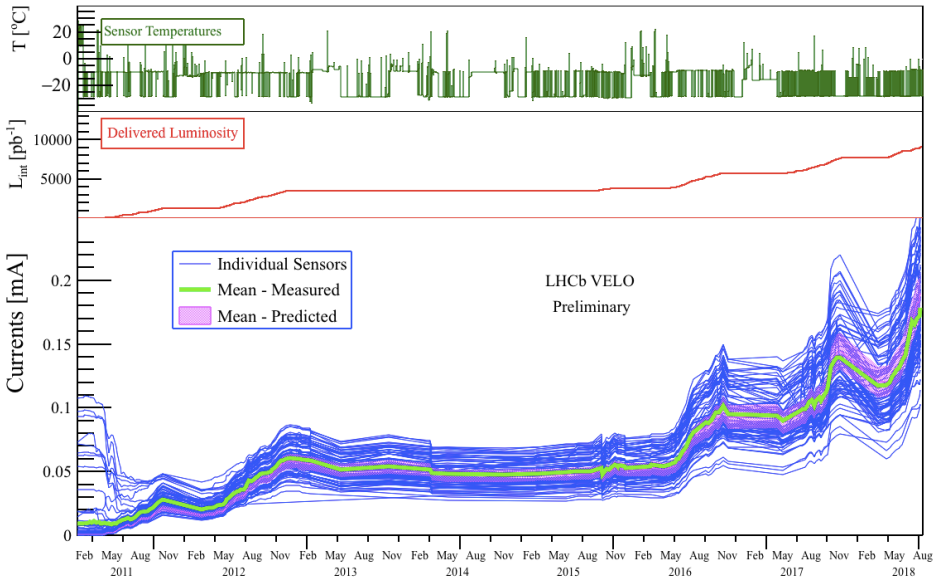


Figure 6.5: Leakage current for VELO in Runs 1 and 2 as a function of time, the plots above show the delivered luminosity and the average temperature variations [128].

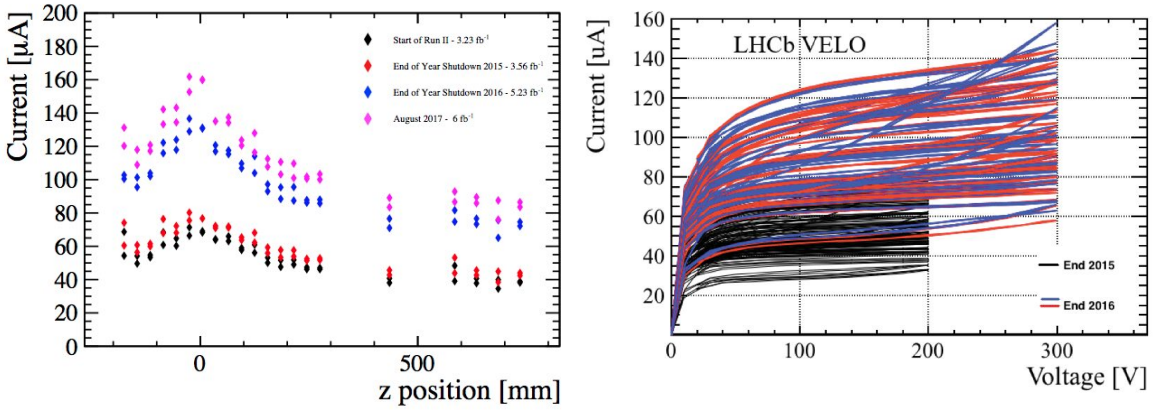
and hence it is a good parameter to measure radiation damage. The example of the increase of leakage currents over time for the previous VELO (Run 1-2) is shown in Fig. 6.5. Leakage currents are constant or decreasing during technical stops due to the annealing processes.

Fig. 6.6(a) illustrates VELO leakage current depending on the z position of the module. The currents are taken at various voltages ranging from (0-300 V), and temperature is normalised to -10°C, The highest leakage current is observed in regions closest to the interaction point at z = 0.

Sensor currents are also analysed as a function of voltage through current-voltage (IV) scans. The primary goal of these scans is to check if the sensor is fully depleted, indicated by reverse-bias current stabilising at higher bias voltages, and also to detect any sudden increase that might signal an impending breakdown. Fig. 6.6(b) provides an example of an IV scan carried out during Run 2 data collection, which clearly shows an increase in the currents caused by radiation damage to the sensor bulk.

6.4 Methods for fluence evaluation

Fluence is estimated using the FLUKA or Geant4 simulation framework, as mentioned in Section 2.7.3. In addition to this simulation-based approach, the present analysis



(a) (b)

Figure 6.6: Leakage current for VELO at Run 1-2 as a function of the z position of the module [128] in (a) and (b) shows the IV scans of selected VELO sensors taken during Run 2. Black lines: currents as a function of applied voltage for both types of sensors at the end of the year 2015, Blue and red lines: correspond to the R-type and Φ type sensors, respectively, measured at the end of the year 2016 [127].

performs a comparative study of the fluence for charged particles, derived from dedicated detector data. These complementary techniques enable a cross-validation of the fluence obtained in simulation.

6.4.1 Machine configuration and data samples for Run 3

Run 3 was commenced in July 2022, following more than three years of upgrades and maintenance. After the upgrade, LHCb became operational with all detectors installed, except for the upstream tracker (UT). This was succeeded by the local commissioning of the sub-detectors and the global commissioning of trigger, alignment, and calibration. During this time, VELO was routinely closed for several months. Finally, in 2024, LHCb was able to collect data for *pp* collisions with a fully instrumented spectrometer. Fig. 6.7 shows the integrated recorded luminosity with respect to the months for different years.

This analysis investigates particles interacting with VELO (VErtex LOcator) sensors, which cause radiation damage to these sensors. The study estimates fluence with a data-driven technique using signals which particles leave in the pixel sensor before the process of reconstruction starts. The specially taken sample consists of detector data, including hits, clusters, and geometrical parameters. Since there is no electrical field in VELO, tracks can be considered as straight lines and momentum is not known. For

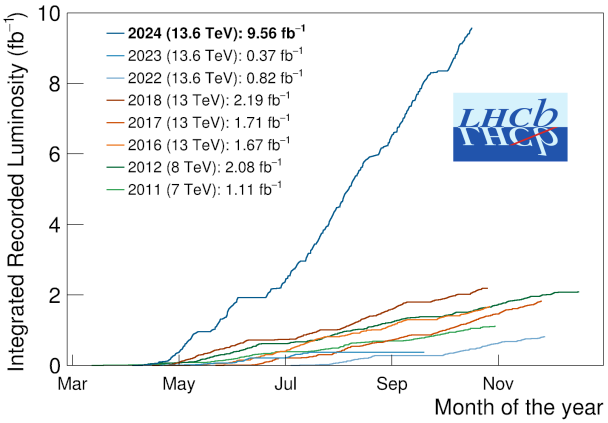


Figure 6.7: Integrated luminosity recorded by LHCb since 2011, with the steepest curve representing data collected in 2024 (Run 3) up to the beginning of September.

the purpose of this study, VELO hits are the primary input. A hit corresponds to the electrical signal produced by an individual sensor when a charged particle passes through, ionising silicon and generating an electrical signal. Hits represent raw, low-level data directly reflecting responses from individual sensor pixels, serving as the foundation for subsequent data reconstruction, such as clustering. Clusters, on the other hand, are higher-level data formed by grouping adjacent hits based on spatial proximity and signal thresholds. This clustering accounts for the charge spread in silicon, where a single particle can activate multiple sensor elements, reducing noise and improving the spatial resolution for reconstructing particle trajectories.

6.4.2 Monte Carlo sample

The MC samples utilised in this analysis were generated using PYTHIA 8 event generator [63]. Minimum bias samples were produced at a centre-of-mass energy of 13.6 TeV, with an average number of visible interactions per event of $\nu = 3.2$. These datasets were subsequently processed using algorithms designed specifically for analysing reconstructed particle tracks. The primary purpose of these algorithms is to extract and store comprehensive details of reconstructed particle tracks, including their geometric properties, kinematic parameters, and hit-level information. These data are preserved for further analyses, such as matching of reconstructed tracks to their corresponding simulated MC truth information. This is the first production for 2022 data, including HLT1 and HLT2. The sample was produced with 100 thousand events.

6.4.3 Data sample

Experimental data complement MC simulations and vice versa, offering a direct measurement of particle distributions and detector responses. This data-driven approach validates simulation results and captures effects sometimes not fully accounted for in simulations. The data were collected from July to September 2024 in pp collisions at $\sqrt{s}$= 13.6 TeV, with VELO closed and magnet polarity up and down with an instantaneous luminosity of 2×10^{33} cm^{-2}s^{-1}. The sample consists of 100 thousand events with a number of colliding bunched (N_{coll}) $\sim$ 2133. The average number of visible interactions in the detector acceptance per bunch crossing (μ) varied across three months of data as mentioned in Table 6.1. The basic fluence-related parameters are extracted from raw detector data, including hit counts, cluster properties, and hit positions etc. These measurements reflect the actual interactions of particles with the silicon sensors and provide a baseline for comparison with simulations. Advanced reconstruction algorithms are used to process the data, ensuring accurate identification of particle trajectories and their spatial distributions.

6.5 Time-dependent performance of the VELO sensor

To ensure an accurate representation of the radiation environment over time, this section describes the specific configuration settings applied to both the data and MC samples. These configurations are designed to account for the time-dependent evolution of the VELO (VErtex LOcator) sensors, which is critical for understanding long-term radiation damage effects.

The VELO detector, as mentioned earlier, is located around the beam pipe and tracks the position of particles that hit the detector pixels. The primary aim is to understand how the hit density changes over time and what this implies for the radiation environment within the detector, particularly in terms of fluence calculations.

Fig. 6.8 shows the VELO hits in a 2D plane for data collected in July and September. The x and y axes represent the transverse positions of hits, while the z-axis corresponds to the bin content. The yellow regions near the centre indicate a high concentration of particles, highlighting areas of the detector exposed to higher radiation levels. Moving farther from the centre, fewer particles hit VELO sensors near the edges of the detector, making these areas less susceptible to radiation damage.

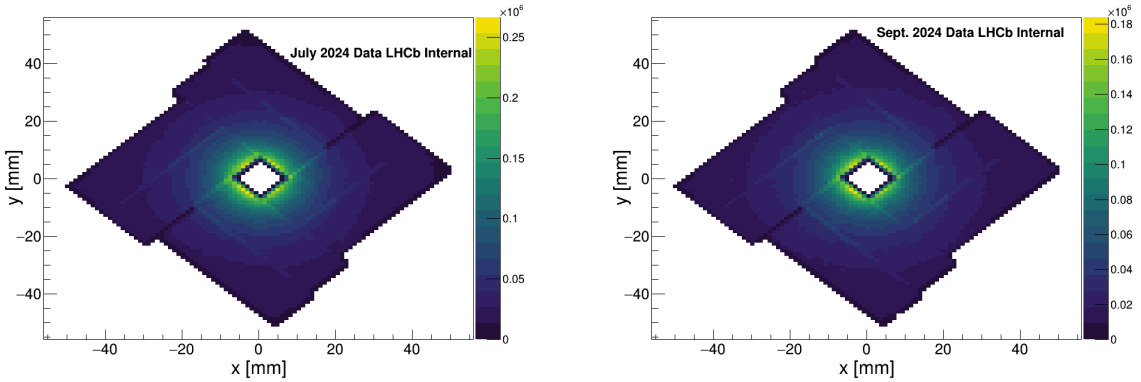


Figure 6.8: The 2D representation of the VELO hits for x and y positions of the sensor for data in July (a) and September (b).

The 2D hit distribution for data collected in July and September shows that the overall trend remains stable, with no significant shift in the concentration of hits across the detector's sensitive areas. Rather than using Cartesian coordinates, the analysis is conducted in spherical coordinates to examine the radial distribution of particle hit density across various VELO modules, highlighting any potential differences.

Figure 6.9(a) presents the radial distribution for the data, and (b) for MC. For this analysis, two modules were selected from the VELO detector's front, middle, and back regions at the following z-positions $-231.25, 118.75, 643.75$ mm, respectively.

The delivered luminosities for the data, along with the average number of visible *pp* interactions in July, August, and September, are shown in Table 6.1.

Months	**Delivered Luminosity [nb^{-1}]**	μ
July	3618.68	4.00
August	5037.96	4.38
September	4959.88	4.38

Table 6.1: The delivered luminosity for July, August, and September measured in nb^{-1} and the average number of visible *pp* interactions.

The hit density in the data reveals that modules 24 and 25 experience the highest hit density, indicating that these modules are more exposed to radiation. This trend is similar to what is observed in MC (b), which aligns with expectations since modules 24 and 25 are in central positions.

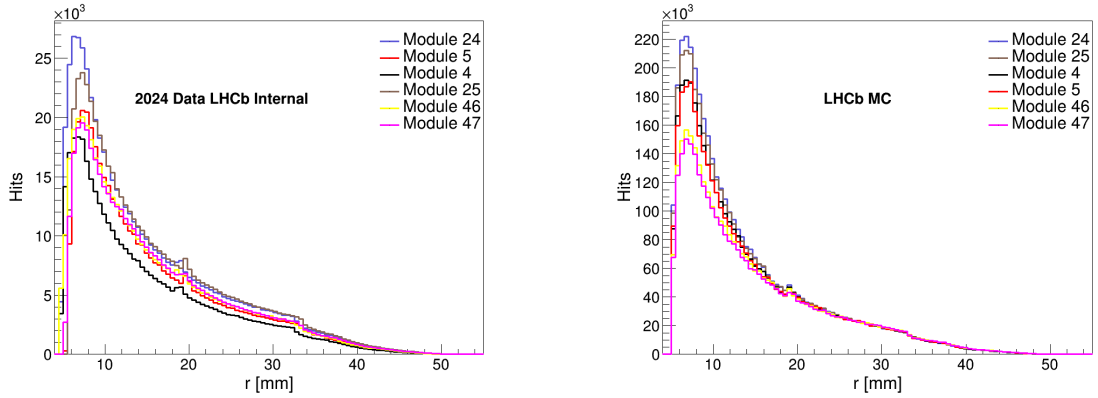


Figure 6.9: The radial representation of the VELO hits in different modules of the VELO for data in (a) and MC in (b).

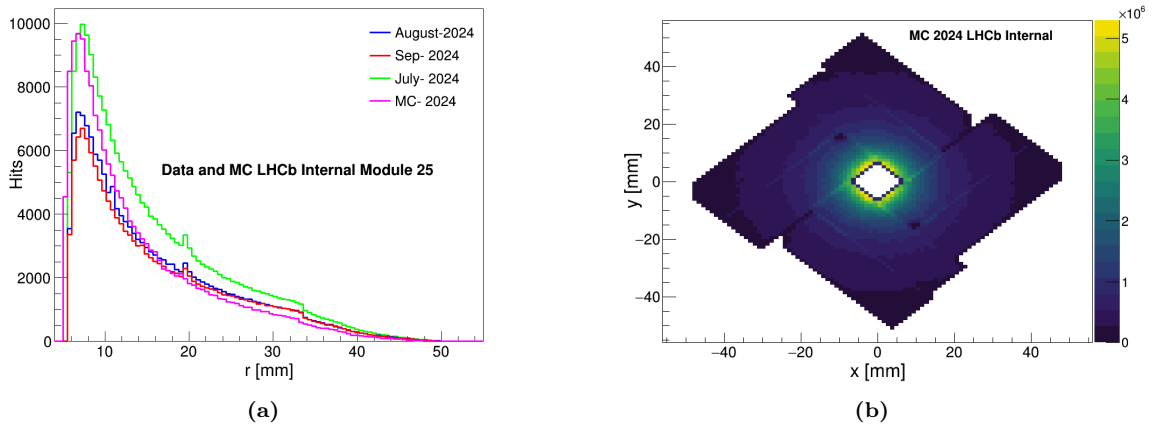


Figure 6.10: The radial representation of the VELO hits for module 25 in data and MC (a) and the 2D hitmap density for the MC sample (b).

Conversely, modules 46 and 47, located at the far end of the VELO station, have the lowest exposure. A similar trend is followed in the simulation. Modules 4 and 5 are located at the beginning and have an average number of hits compared to the ones located in the middle and the back modules.

Notably, there is a discrepancy between the data and MC for modules 5 and 46. In the data, these modules exhibit similar hit densities, whereas in the MC, module 5 registers a higher particle hit density compared to module 46. This difference highlights a potential mismatch in the simulation model or detector response that needs further investigation.

Fig. 6.10(a) shows the comparison of module 25 for data (July, August, and September)

and MC. The MC distribution provides a good reference, as it aligns well with the early data (July), but deviations become noticeable over time, particularly in September. The agreement between MC and July data suggests that the simulation accurately models the initial detector behavior. The divergence over time might highlight the effects of radiation damage, which the static MC model does not fully capture. The radial hit density consistently decreases for all modules, particularly in the inner regions close to the beamline. This trend is a direct consequence of radiation damage, which reduces the efficiency and sensitivity of the silicon sensors. The observed reduction in hit density is consistent with expectations for the VELO detector, where prolonged exposure to high-radiation environments leads to sensor degradation. Fig. 6.10(b) shows the 2D hit density representation for the MC sample.

6.6 Fluence evaluation

Fluence is a key parameter in understanding the radiation environment within the VELO detector of the LHCb experiment. Accurate fluence calculations are crucial for assessing radiation-induced damage to the silicon sensors and predicting the detector's long-term performance. Fluence is described in terms of the total distance that particles travel within a given volume, which provides a more practical approach in certain radiation detection techniques. This alternative definition relates fluence to the total path length of all particles traversing a volume. This relationship allows us to estimate particle fluence from measurable quantities like track lengths. In this chapter, fluence is estimated as track-length density, when a track is counted as a hit in the VELO sensor, multiplied by the sensor width (200 μm) [129].

Hit distribution maps are derived from experimental data and Monte Carlo (MC) simulations. Not only does the number of particles transversing the sensor contribute to the ϕ_{eq}, but also the energy and the angle of crossing are important. The high value of fluence close to the IP is mainly due to the large angles and small particle energy. Fluence in the more distant sensors is lower, even though the particle density is higher at the small angles.

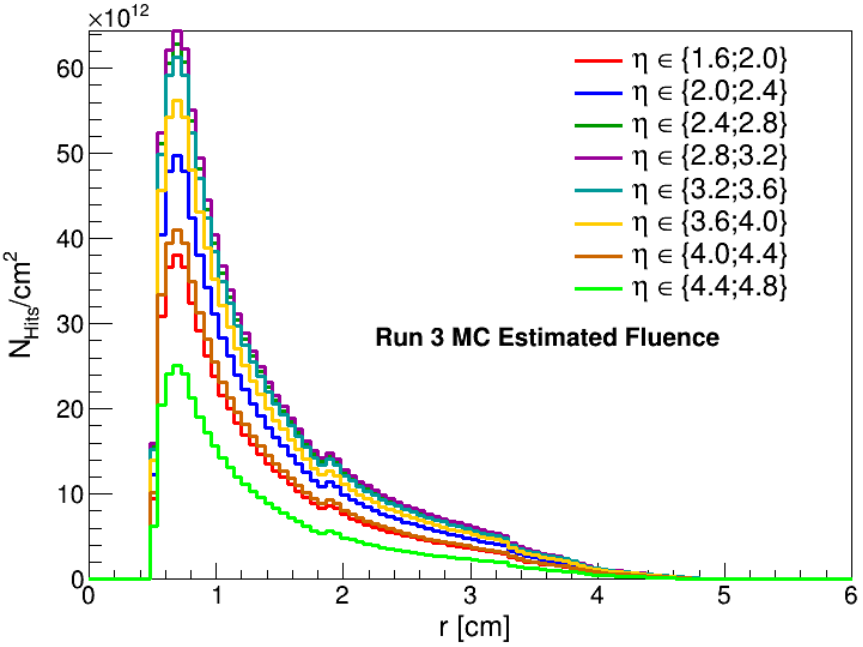
Figure 6.11(a) presents the fluence distribution as a function of the radial coordinate, categorised by different pseudorapidity (η) intervals for the Monte Carlo (MC) sample. These distributions correspond to 1 fb^{-1}. Given that the LHCb detector has an acceptance range of $2 < \eta < 4.5$, the number of hits in regions outside this range is notably low.

Consequently, the majority of particle hits—and therefore higher fluence—are observed within the central η intervals.

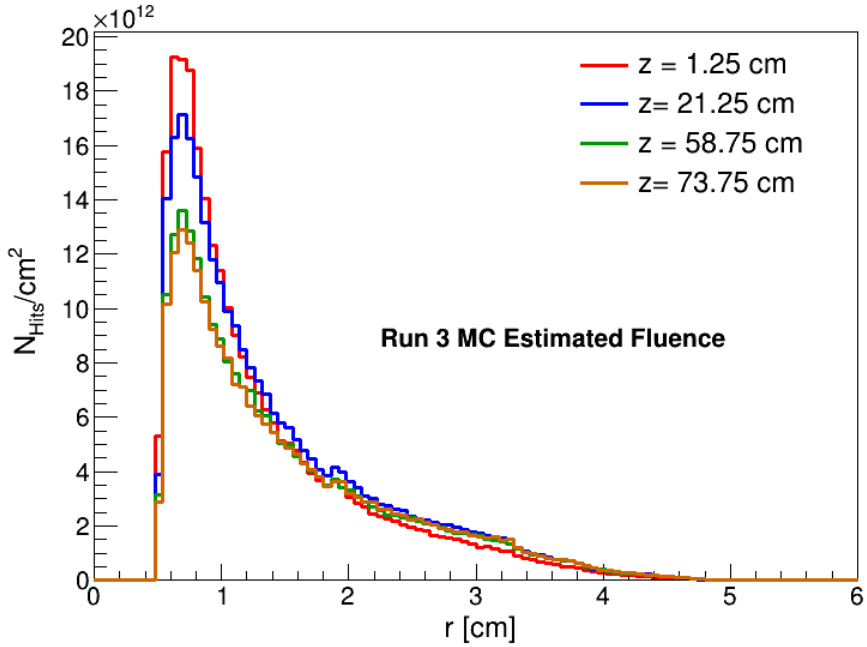
Figure 6.11(b) further illustrates the estimated fluence levels across the detector. It is evident that sensors located closer to the interaction point are subjected to significantly higher fluence, whereas those positioned farther away experience comparatively lower radiation levels. This spatial dependence reflects the expected radiation profile within the detector geometry.

By comparing these visualisations with the delivered luminosity, we can assess the non-uniform distribution of particle fluence across the VELO detector. Figure 6.12(a) displays the radial hit density for various sensors, normalised to 1 fb^{-1}. The sensor positions are indicated, with the red and blue distributions corresponding to sensors located closest to the interaction point. These sensors experience higher particle flux compared to those positioned farther away. Figure 6.12(b) presents the radial hit density, scaled to 1 fb^{-1} for a single sensor located at a z-position of 21.25 cm, across different months. The hit density is highest in July, marking the start of Run 3, when all sensors were fully operational.

This study provided with data-driven method for online, fast fluence estimation based on hit density calculation. By observing the sudden increase of hit density in several pixels, one can spot whether part of the sensor has undergone severe radiation damage. This method might be further improved while adding information from the simulation regarding the association between hits and particles. It would allow for the full evaluation of the ϕ_{eq}. Once the appropriate software is within reach, this task will be fulfilled.

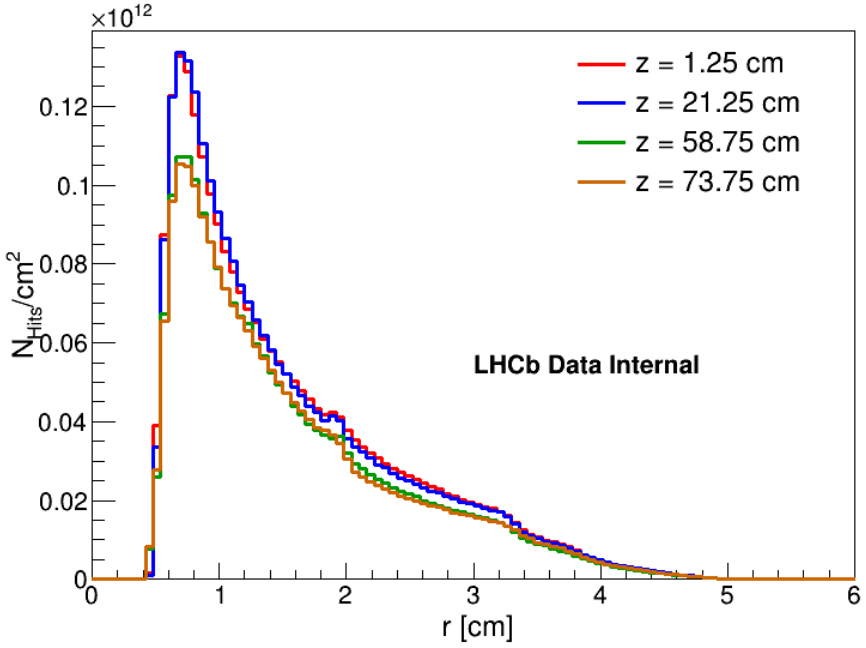


(a)

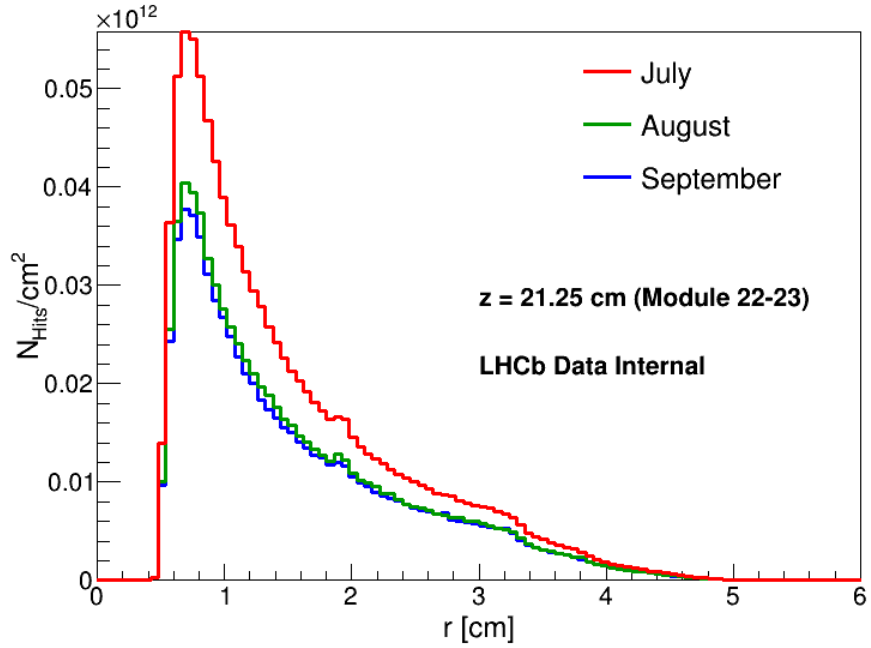


(b)

Figure 6.11: Estimation of fluence for the MC sample plotted as a function of kinetic energy in different η regions in (a). Fluence as a function of energy for the z positions of the MC sample in (b).



(a)



(b)

Figure 6.12: Radial hit density for various VELO sensors, scaled to 1 fb^{-1}; the distributions correspond to different sensor positions (a). Radial hit density for a single VELO sensor at a fixed z-position (21.25 cm) corresponding to module 22-23, scaled to 1 fb^{-1}, for July, August, and September 2024 (b).

Chapter 7

Summary

Art is never finished, only abandoned - Leonardo DaVinci

The Large Hadron Collider (LHC) at CERN stands as the world's most powerful particle accelerator, providing unique opportunities to explore fundamental interactions at unprecedented energy scales. Among its four major experiments—CMS, ATLAS, ALICE, and LHCb—the LHCb detector is specially designed for the study of heavy-flavour hadrons, including beauty and charm quarks, with a focus on forward physics in a particular pseudorapidity region. This unique geometry and physics program makes LHCb a vital contributor to precision tests of the Standard Model and searches for new physics phenomena.

At the heart of precise physics measurements is the ability to accurately simulate particle production and interactions, which is critical for interpreting experimental data. This thesis addresses this challenge by investigating the performance and tuning of event generators, validating their agreement with LHCb data, and evaluating their impact on understanding radiation damage in one of LHCb's most sensitive sub-detectors—the Vertex Locator (VELO).

The first part of the research compares two widely used event generators, Pythia and Herwig, focusing on key observables such as energy, momentum, pseudorapidity (η), and charged particle multiplicity. This comparison highlights intrinsic differences because of their hadronisation models—string fragmentation in Pythia versus the cluster model in Herwig. Confronting both with LHCb experimental data reveals that the LHCb-tuned Pythia 8.212 provides the best description of the recorded events, establishing it as the preferred generator for further analyses within the experiment.

Building upon this foundation, the thesis explores the optimisation of event generator

parameters through advanced tuning techniques. Employing the RIVET and PROFESSOR software frameworks, the study follows a certain methodology to quantify and minimise discrepancies between simulation and data via goodness-of-fit functions. The results demonstrate that while current tunings achieve significant improvement, continued refinements and the development of additional RIVET analysis plugins are necessary to maintain accuracy at higher centre-of-mass energies, such as $\sqrt{s}$= 13 TeV and beyond. This iterative tuning process underscores the dynamic nature of Monte Carlo modelling in the evolving landscape of collider physics.

The third focal point of the thesis examines *pp* collision data from Run 2 (2016–2018) at 13 TeV. The study systematically compares these measurements with Monte Carlo predictions by analysing charged particle multiplicities and their kinematic distributions. Discrepancies in regions sensitive to hadronisation and underlying event modelling indicate that existing theoretical frameworks require further improvements to capture the complexities of high-energy hadronic interactions more accurately. This analysis provides critical guidance for enhancing event generator tuning and modelling strategies, contributing to a more precise understanding of particle production mechanisms.

In parallel, the thesis investigates the Vertex Locator (VELO), a silicon pixel detector positioned a few millimetres from the proton-proton interaction point. VELO's proximity exposes it to intense radiation, necessitating robust radiation-hardened technologies and continuous monitoring of sensor performance. The detector upgrade, which increased the nominal instantaneous luminosity by a factor of five, leads to significantly higher and spatially non-uniform particle fluxes. These fluxes primarily result from prompt hadron production during collisions, creating complex radiation environments that must be well characterised to safeguard detector longevity.

To address this, a novel data-driven approach was developed and applied to estimate particle fluence and flux within VELO, tracking radiation damage over time. Analyses utilising both Run 2 and newly collected Run 3 data from 2024 demonstrate the feasibility of real-time monitoring techniques that complement traditional simulation-based methods. These results provide essential insights into the temporal evolution of sensor degradation, enabling better predictive capabilities for radiation damage during LHC Runs 3 and 4, which will experience even harsher conditions.

Overall, this thesis significantly advances the understanding and modelling of particle interactions relevant to LHCb physics and detector operation. By critically assessing and refining event generator performance, developing robust tuning methodologies, and

implementing innovative data-driven monitoring strategies for the VELO detector, it enhances the precision and reliability of physics analyses under increasingly challenging collider conditions. These contributions are pivotal for sustaining the high-quality data output necessary for probing the Standard Model with greater sensitivity and exploring new physics frontiers.

The work presented herein not only improves the fidelity of simulations and detector performance evaluation at LHCb but also establishes methodological frameworks and tools that will support the experiment's ongoing and future physics programs. In this way, the thesis forms an integral part of the continuous effort to unlock deeper insights into fundamental particles and their interactions, maintaining LHCb's position at the forefront of high-energy physics research.

Appendix A

Efficiencies of the Cuts

A.0.1 Comparison of efficiencies of primary vertex (PV) cuts

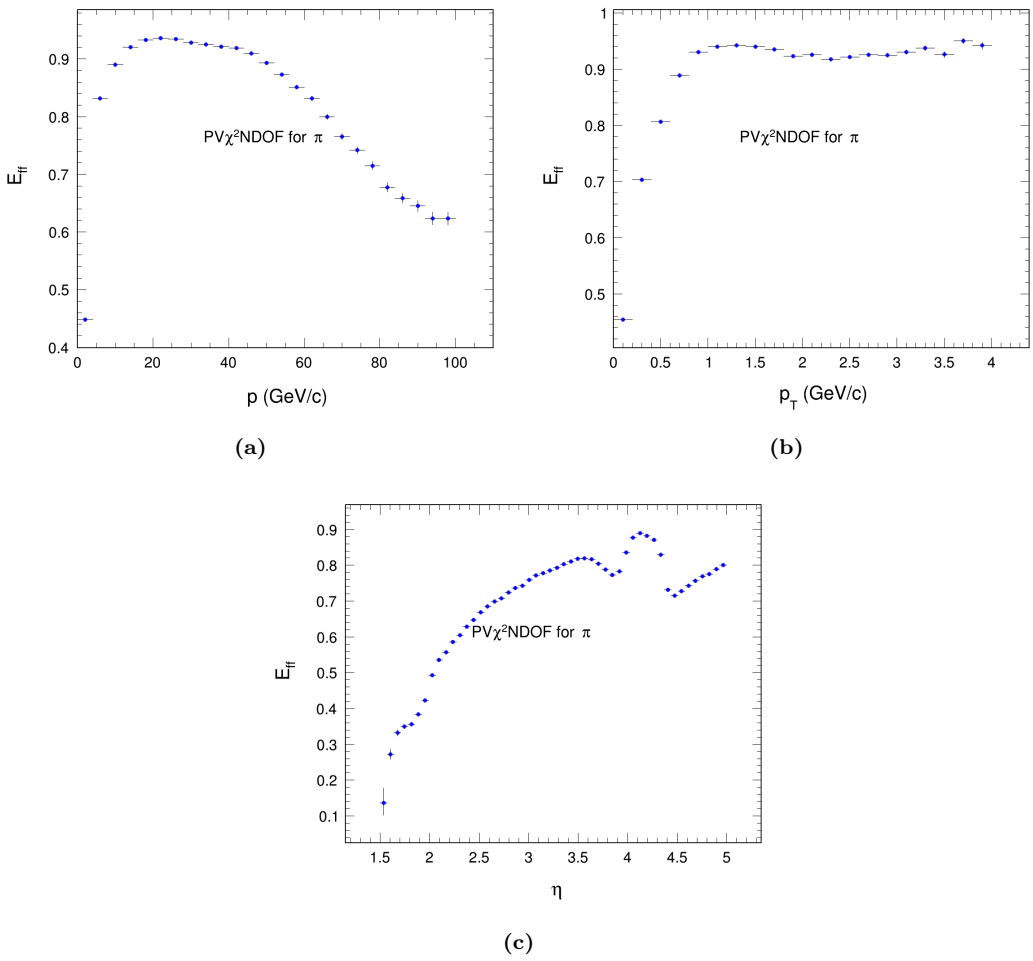
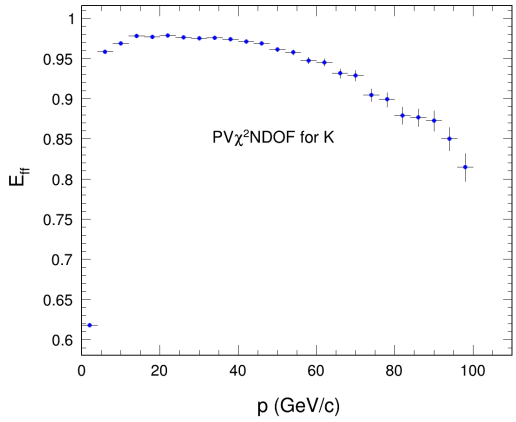
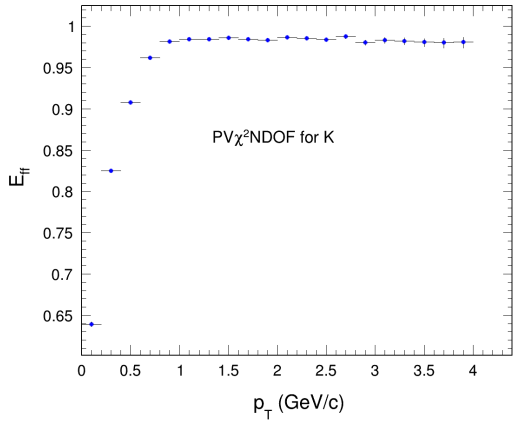


Figure A.1: Comparison of different PVCHI2NDOF for the identification of pions with respect to momentum (a), transverse momentum (b), and η (c).





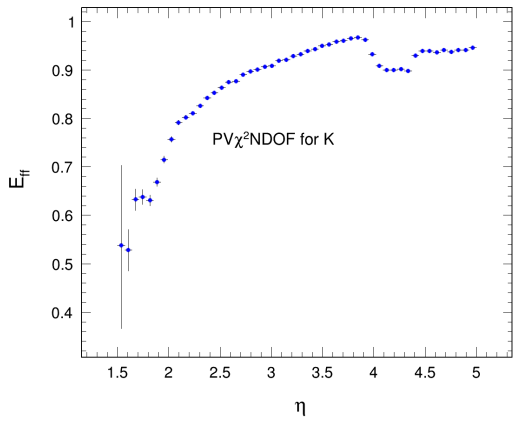
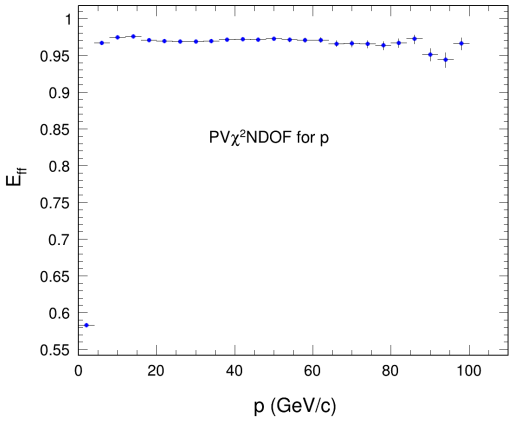
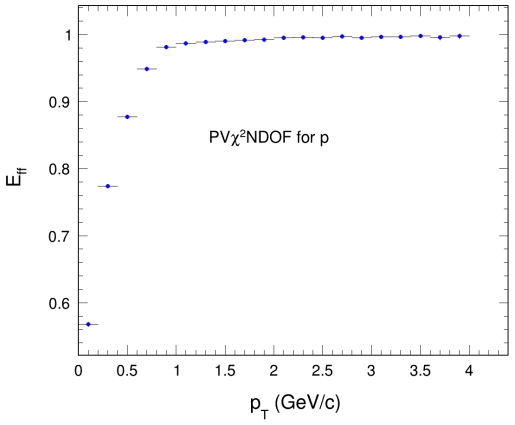


Figure A.2: Comparison of different PVCHI2NDOF for the identification of kaons with respect to momentum (top), transverse momentum (middle), and η (bottom).





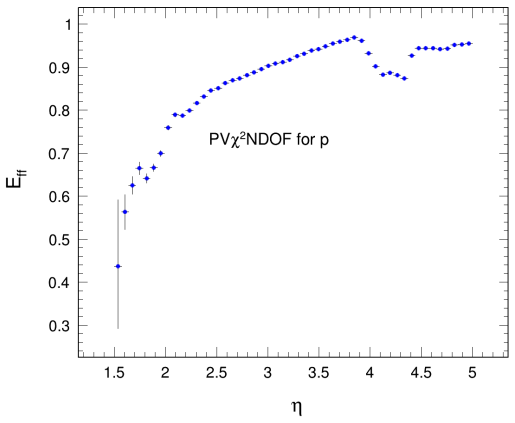


Figure A.3: Comparison of different PVCHI2NDOF for the identification of protons with respect to momentum (top), transverse momentum (middle), and η (bottom).

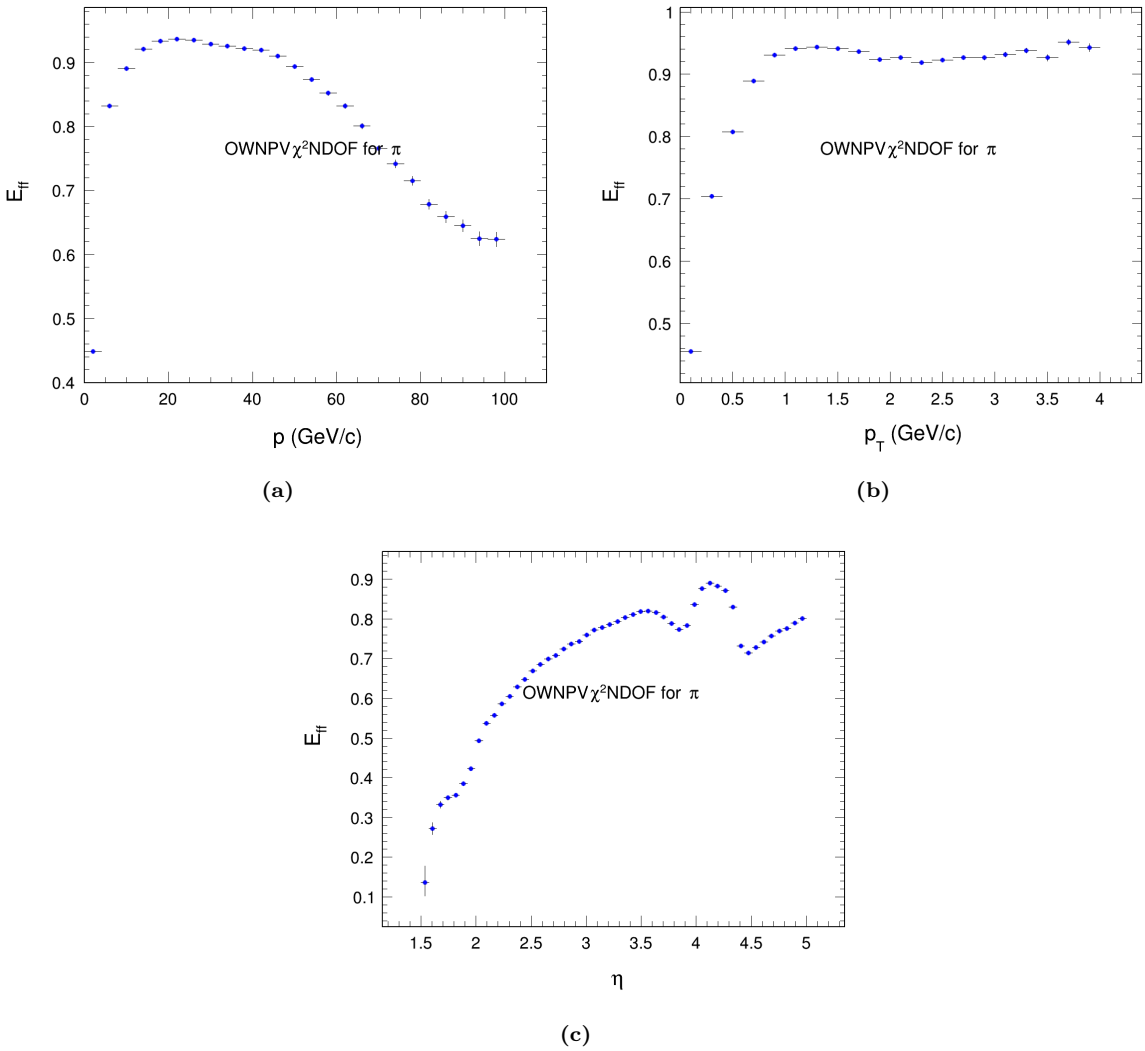
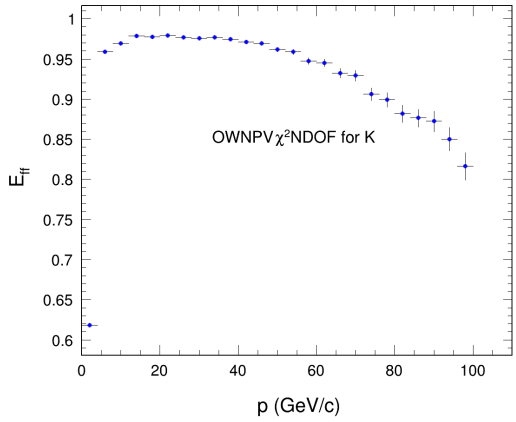
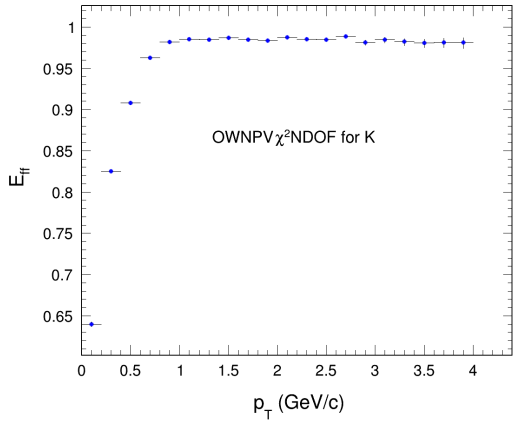


Figure A.4: Comparison of different OWNPVCHI2NDOF for the identification of pions with respect to momentum (a), transverse momentum (b), and η (c).





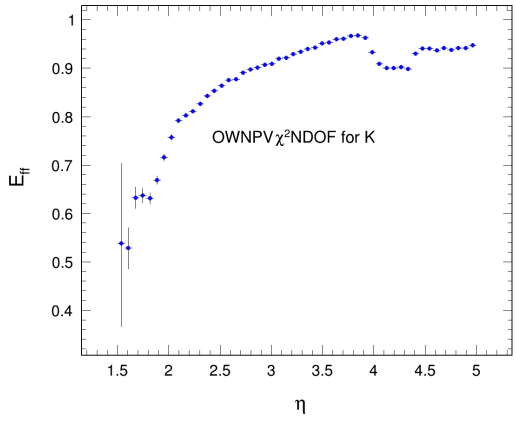
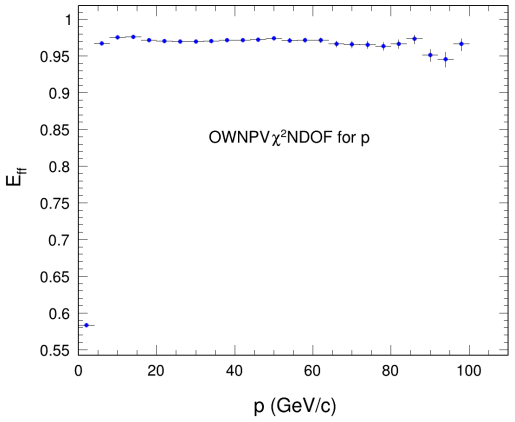
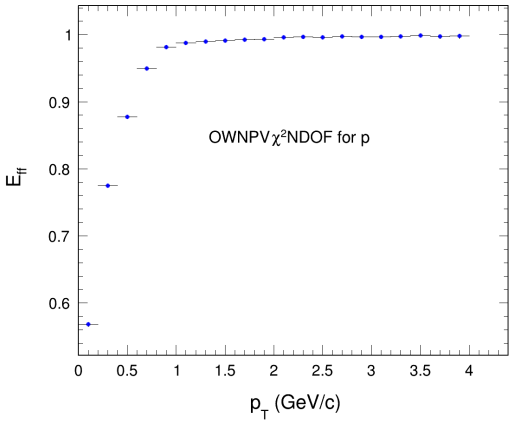


Figure A.5: Comparison of different OWNPVCHI2NDOF for the identification of kaons with respect to momentum (top), transverse momentum (middle), and η (bottom).





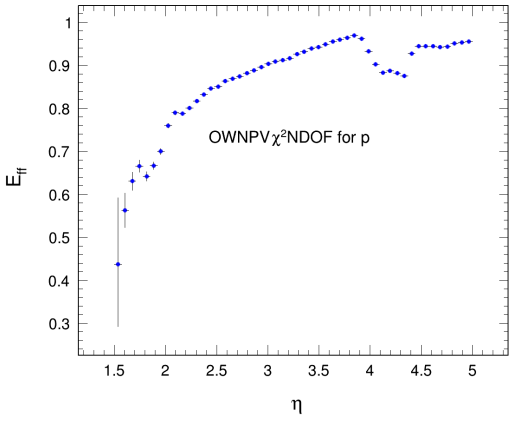
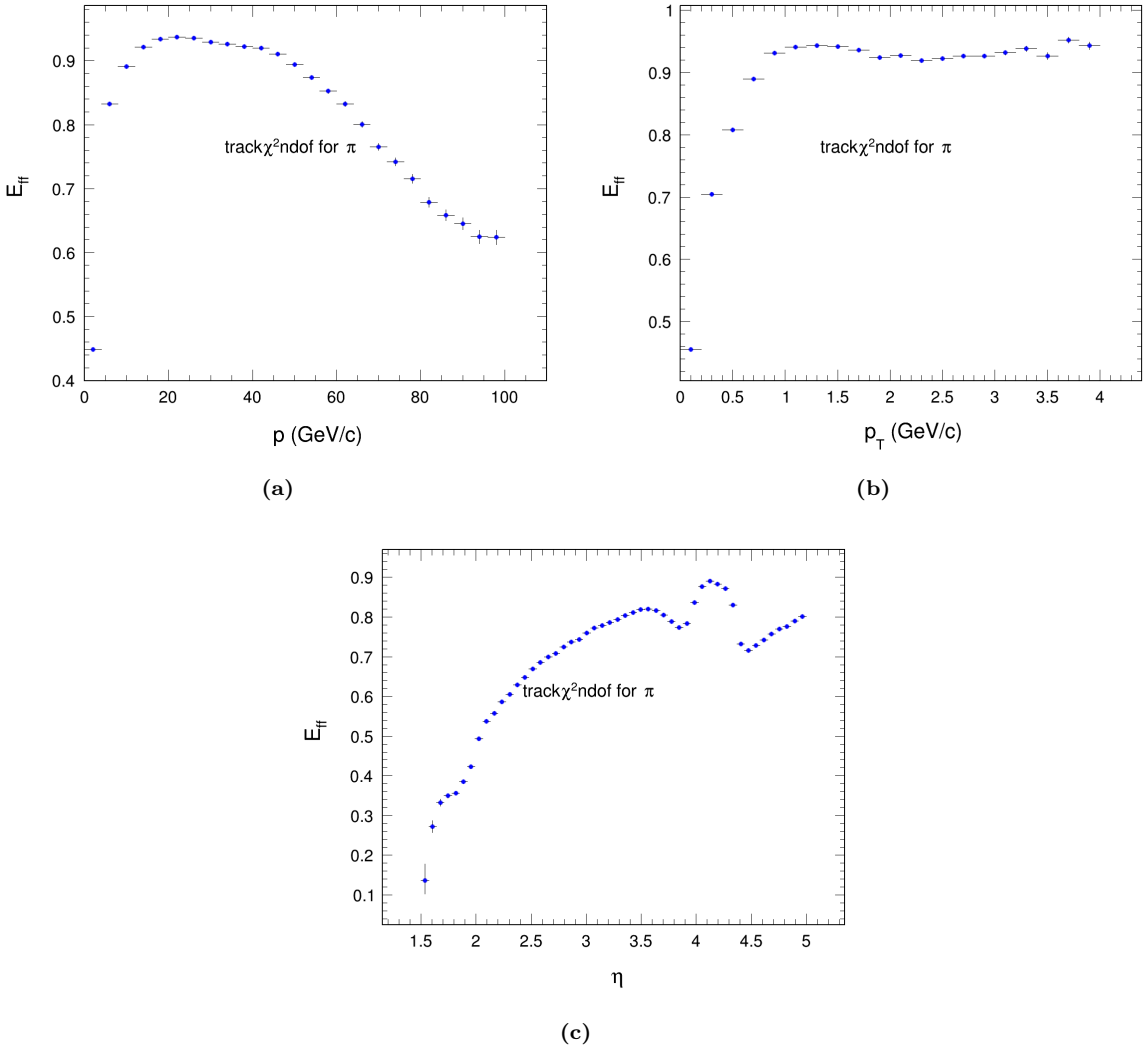


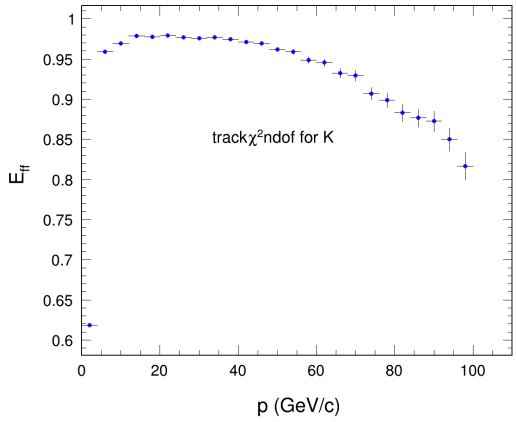
Figure A.6: Comparison of different OWNPVCHI2NDOF for the identification of protons with respect to momentum (top), transverse momentum (middle), and η (bottom).

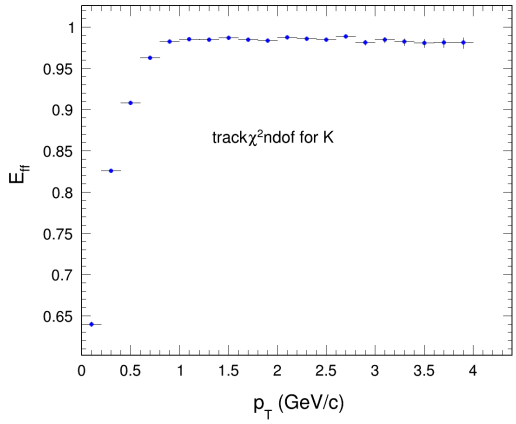
A.0.2 Comparison of the efficiencies of the tracking cuts



(a) **(b)** **(c)**

Figure A.7: Comparison of different trackchi2ndof for the identification of pions with respect to momentum (a), transverse momentum (b) and η (c).





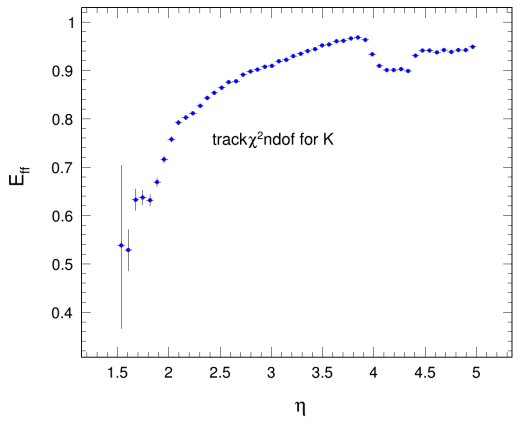
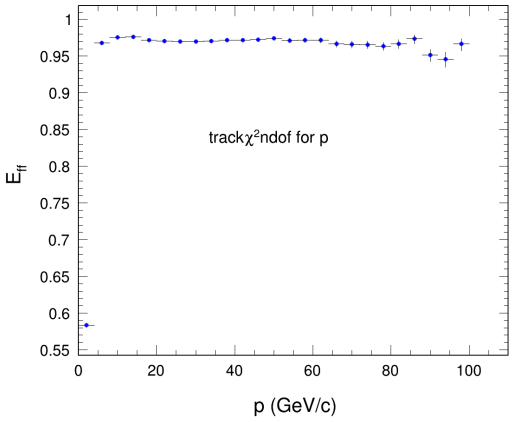
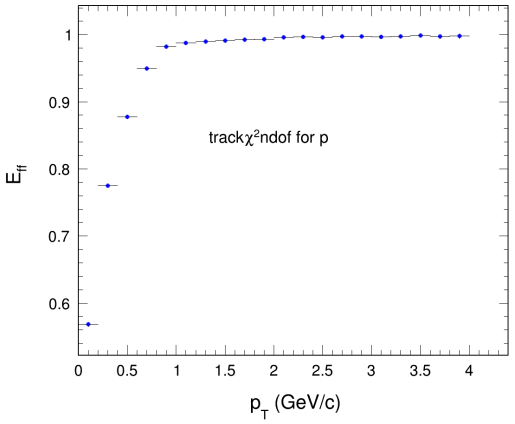


Figure A.8: Comparison of different TRACKCHI2NDOF for the identification of kaons with respect to momentum (top), transverse momentum (middle), and η (bottom).





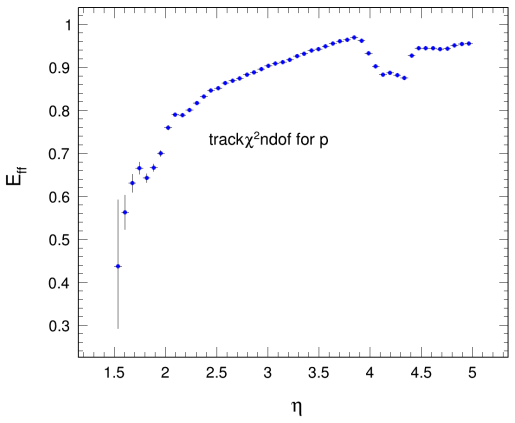


Figure A.9: Comparison of different TRACKCHI2NDOF for the identification of protons with respect to momentum (top), transverse momentum (middle), and η (bottom).

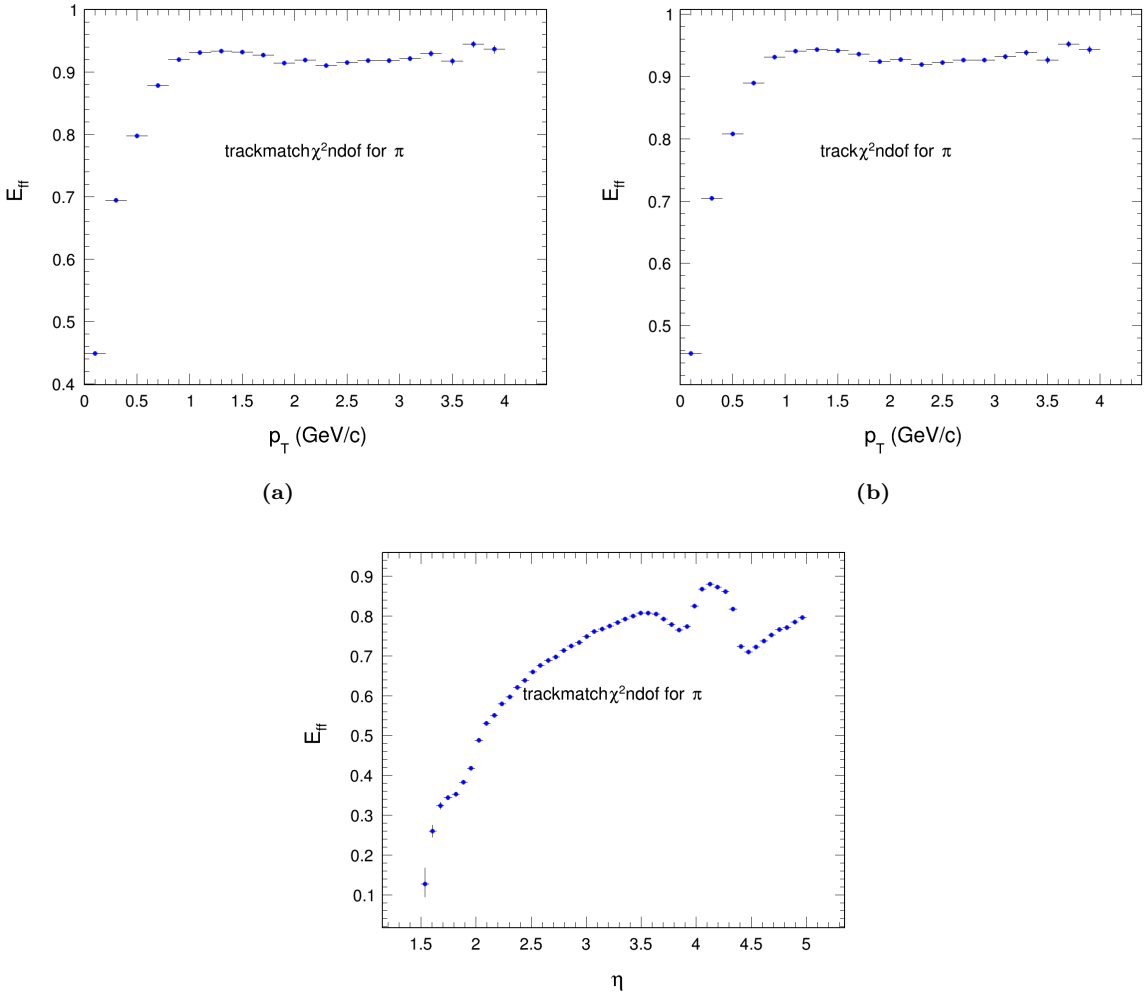
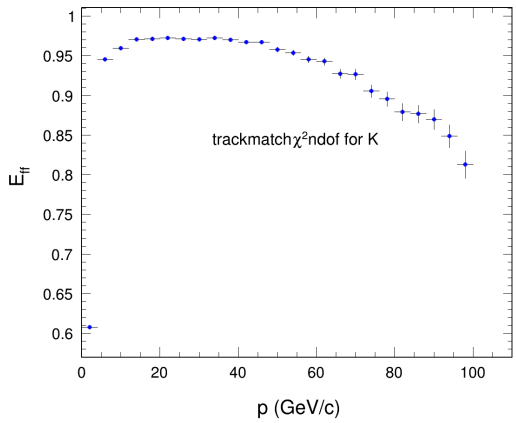
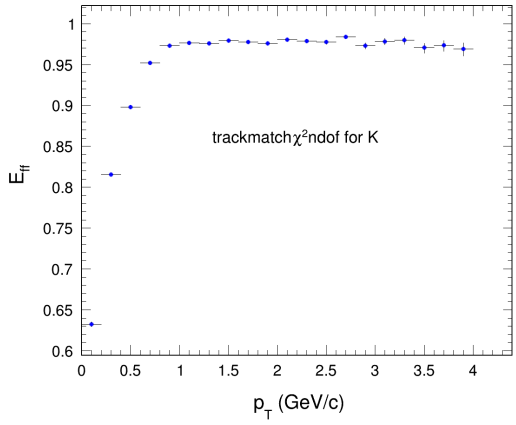


Figure A.10: Comparison of different trackmatchchi2 for the identification of pions with respect to momentum (a), transverse momentum (b) and η (c).





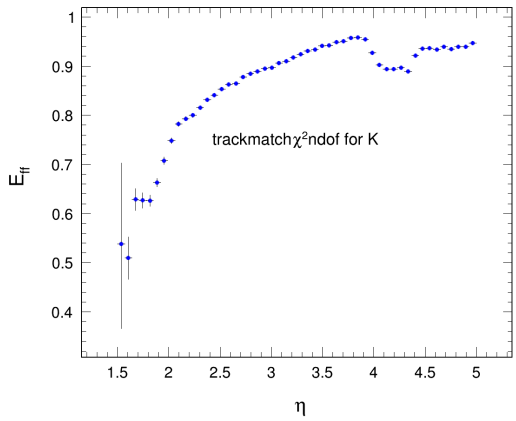
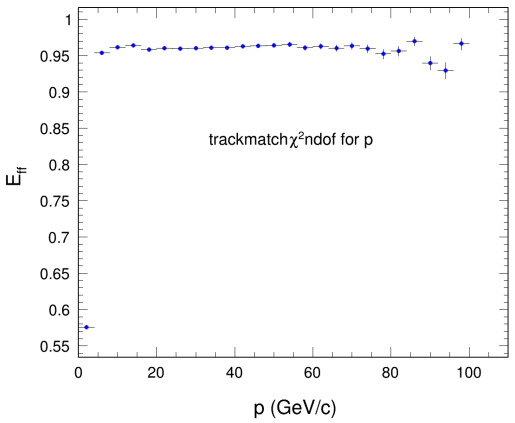
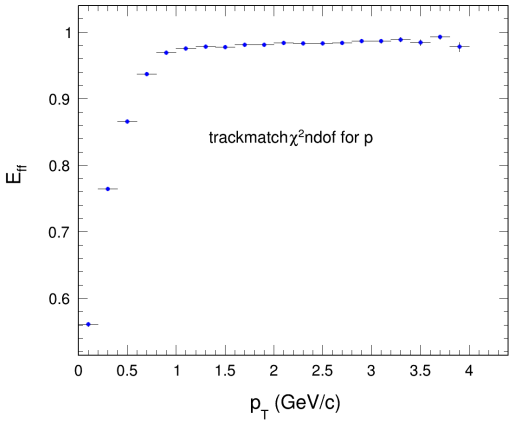


Figure A.11: Comparison of different TRACKMATCHCHI2 for the identification of kaons with respect to momentum (top), transverse momentum (middle), and η (bottom).





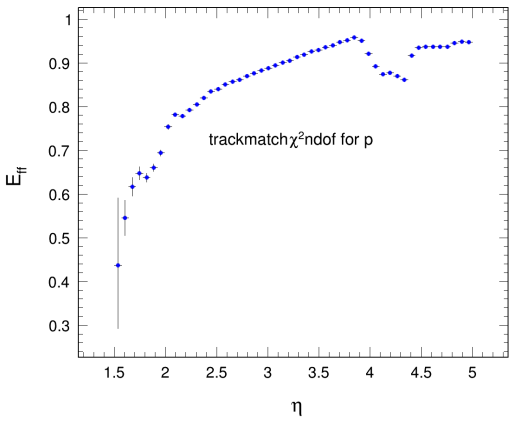
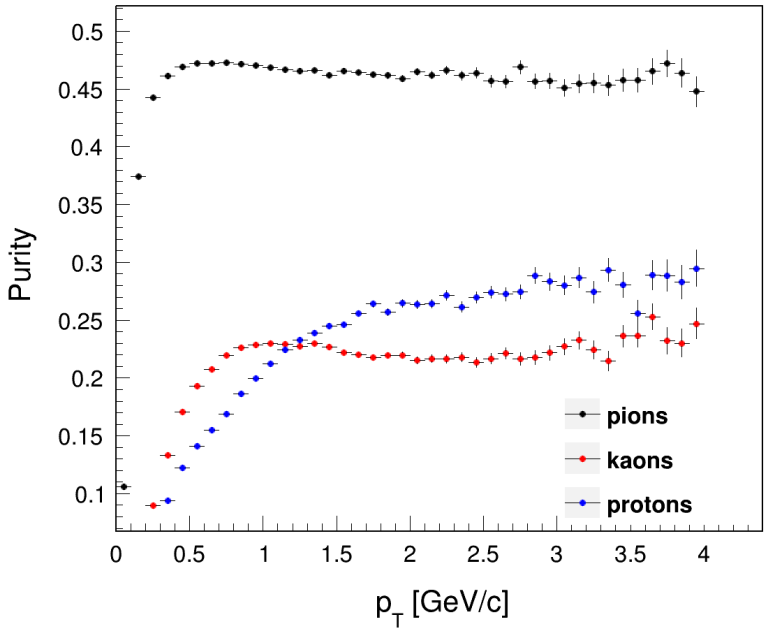


Figure A.12: Comparison of different TRACKMATCHCHI2 for the identification of protons with respect to momentum (top), transverse momentum (middle), and η (bottom).

A.0.3 Purity of the PID with respect to p_T



(a)

Figure A.13: Purity of the PID, with respect to p_T for pions, kaons, and protons.

A.0.4 Purity of the sample with respect to p_T, η and momentum.

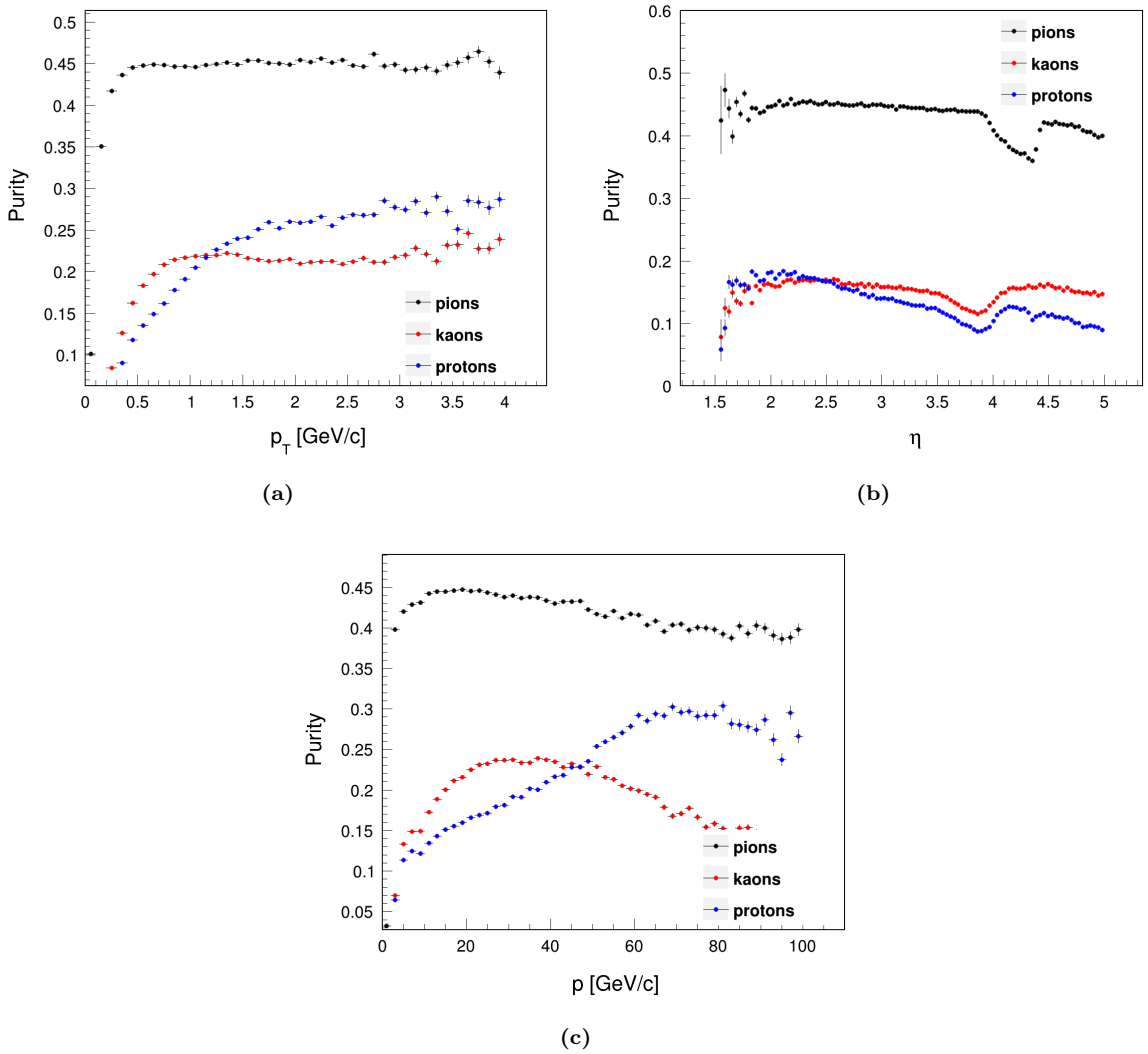


Figure A.14: Purity of the MC sample with respect to p_T, η, and momentum for pions, kaons, and protons.

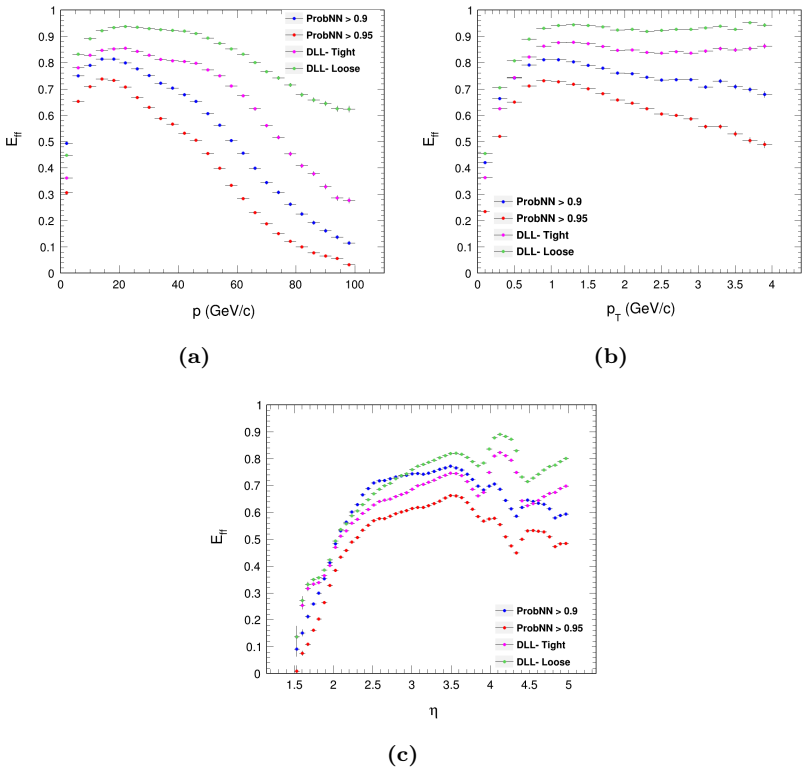
A.0.5 Comparison of the efficiencies of the PID cuts

The particle identification techniques discussed in A.0.5 fall into two categories. The first is the DLL technique, a likelihood-based method that evaluates the probability of a given detector response under different particle hypotheses. The second is ProbNN, a neural network-based approach that assigns probabilities to a particle belonging to a specific type. The applied selection criteria for these methods are detailed in Table A.1.

Cuts Applied	**pions**	**kaons**	**protons**
DLL-Tight	PIDK $<$ -3	PIDK $>$ 5	PIDp $>$ 3
DLL-Loose	PIDK $<$ 0	PIDK $>$ 0	PIDp $>$ 0
probNN-0.9	ProbNNpi $>$ 0.9	ProbNNK $>$ 0.9	ProbNNp $>$ 0.9
probNN-0.95	ProbNNpi $>$ 0.95	ProbNNK $>$ 0.95	ProbNNp $>$ 0.95

Table A.1: Particle Identification (PID) cuts for data and MC.

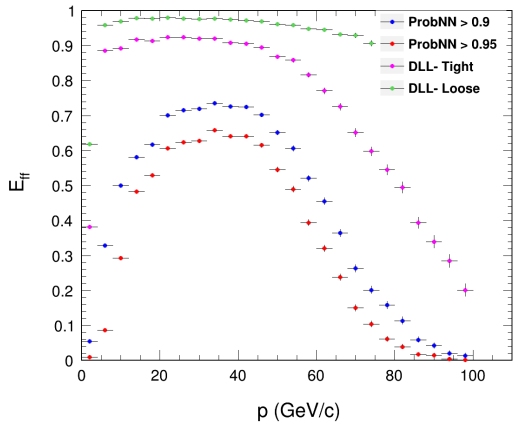
A.0.6 Comparison of the efficiencies of the PID techniques

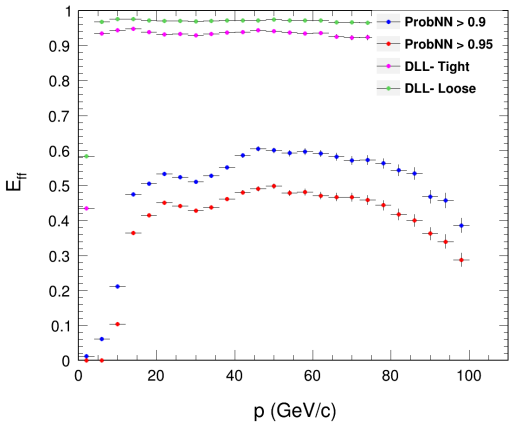


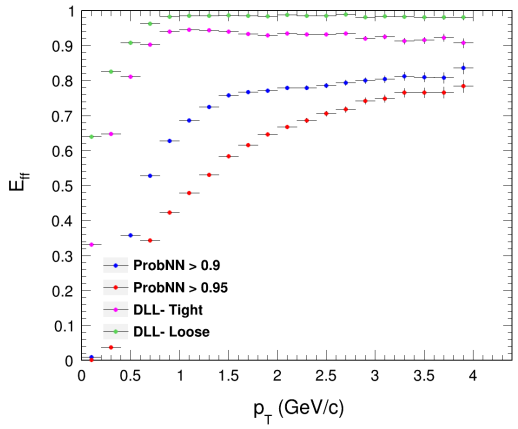
(a) **(b)**

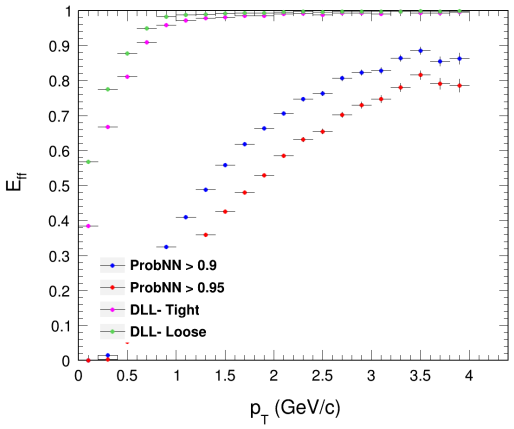
(c)

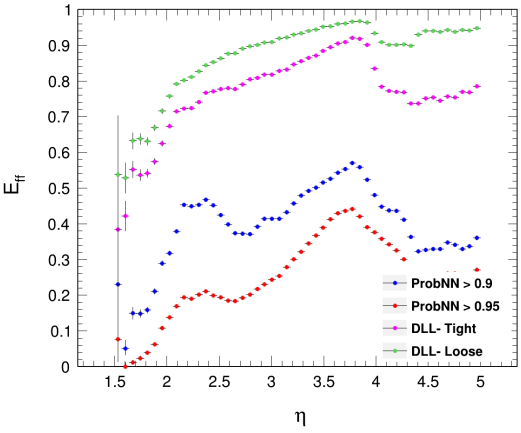
Figure A.15: Comparison of different PID techniques for the identification of pions with respect to momentum (a), transverse momentum (b) and η (c).











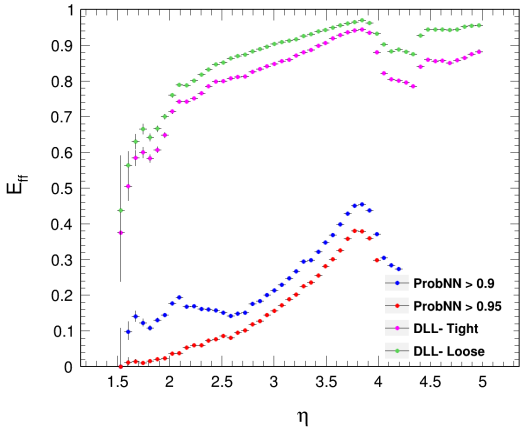


Figure A.16: Comparison of different PID techniques for the identification of kaons with respect to momentum (top), transverse momentum (middle), and η (bottom).

Figure A.17: Comparison of different PID techniques for the identification of protons with respect to momentum (top), transverse momentum (middle), and η (bottom).

A.0.7 Comparison of efficiencies of DLL cuts for pions, kaons and protons at different values

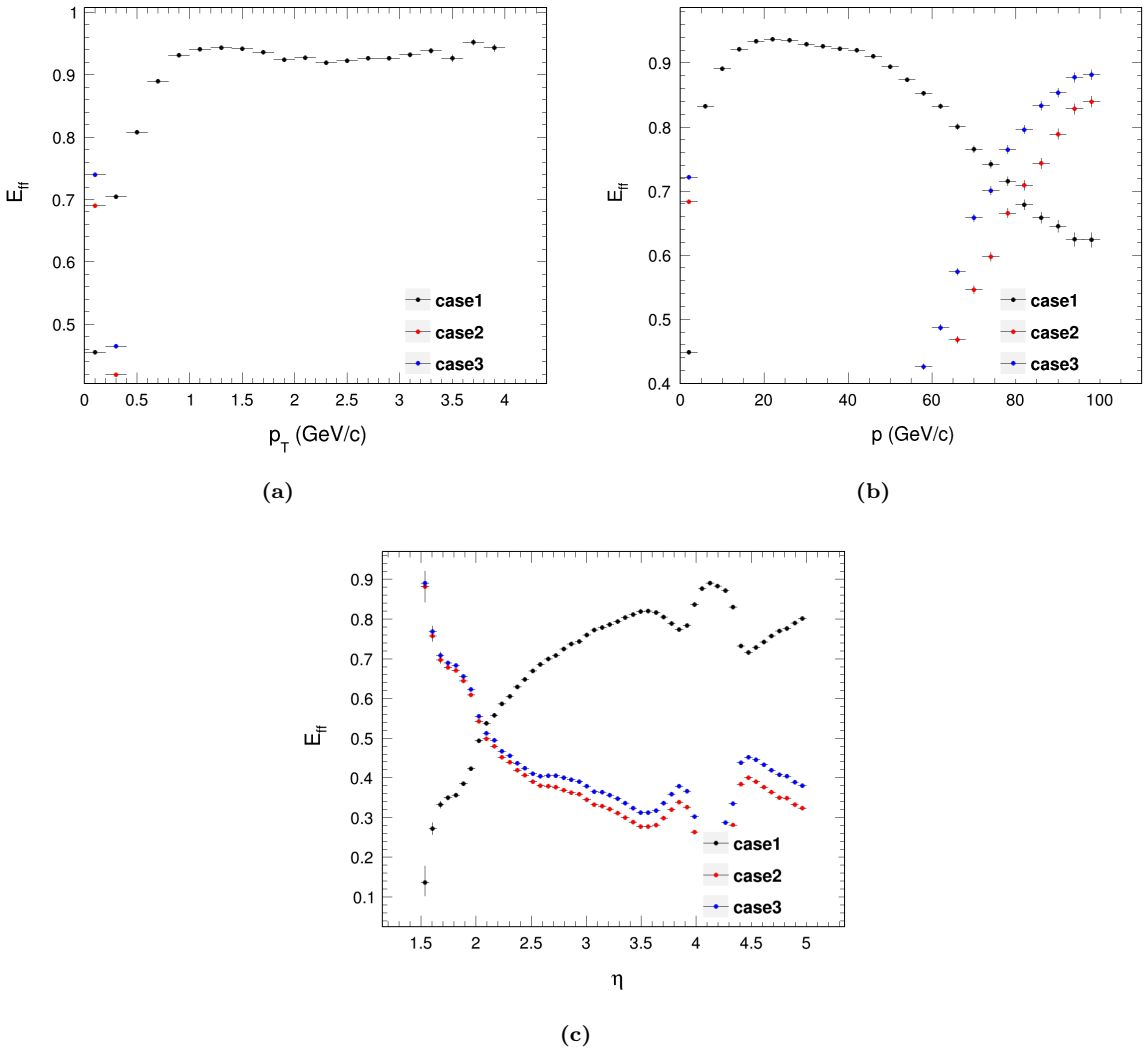


Figure A.18: Comparison of different settings of DLL cuts for the identification of pions with respect to transverse momentum (a), momentum (b) and η (c).

case1	PIDK < 0
case2	PIDK > -5
case3	PIDK > -7

Table A.2: Pions

case1	PIDK > 0
case2	PIDK > 5
case3	PIDK > 7

Table A.3: Kaons

case1	PIDp > 0
case2	PIDp > 5
case3	PIDp > 7

Table A.4: Protons

Table A.5: Different case settings of DLL parameters for the identification of hadrons.

Figure A.18 , A.19 and A.20 shows how the different settings of DLL parameters for the identification of pions, kaons and protons effects the efficiency plots with respect to transverse momentum (a), momentum (b) and eta (c). The cases mentioned here refer to the following settings mentioned in Table A.5

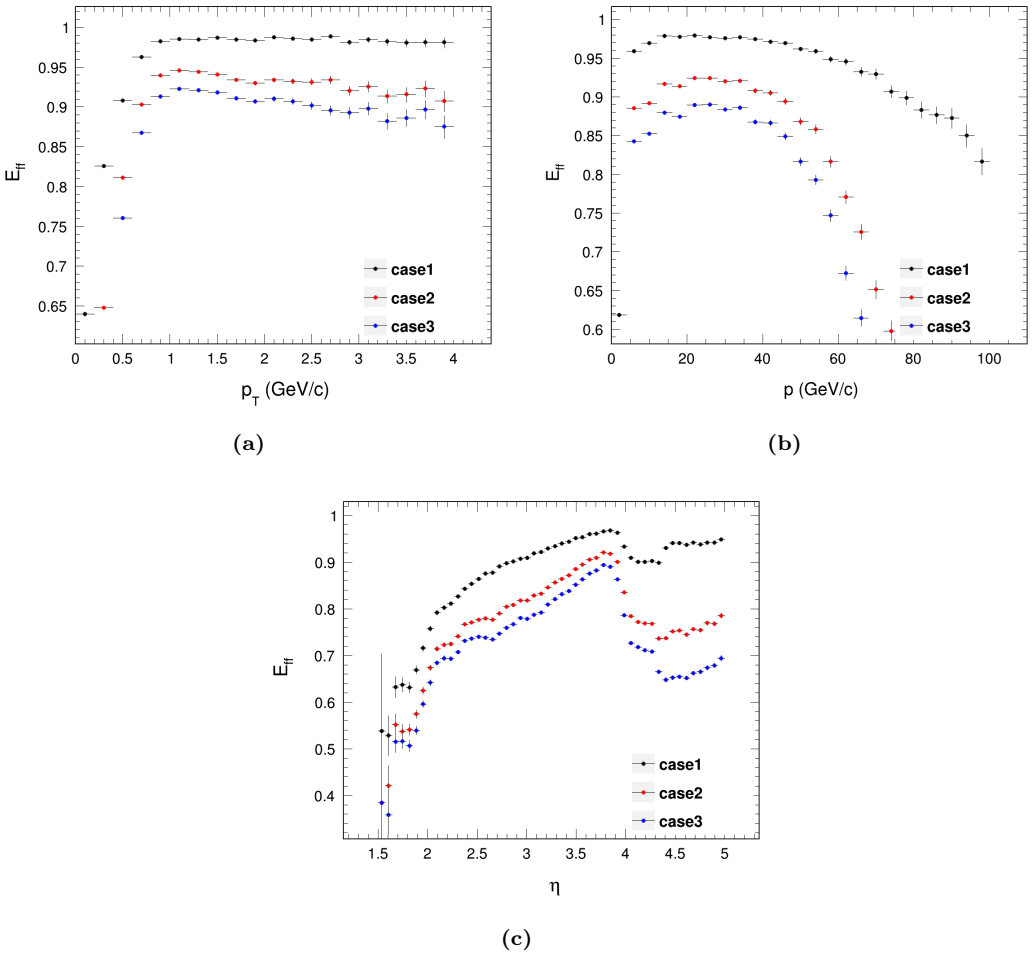


Figure A.19: Comparison of different settings of DLL cuts for the identification of kaons with respect to transverse momentum (a), momentum (b) and η (c).

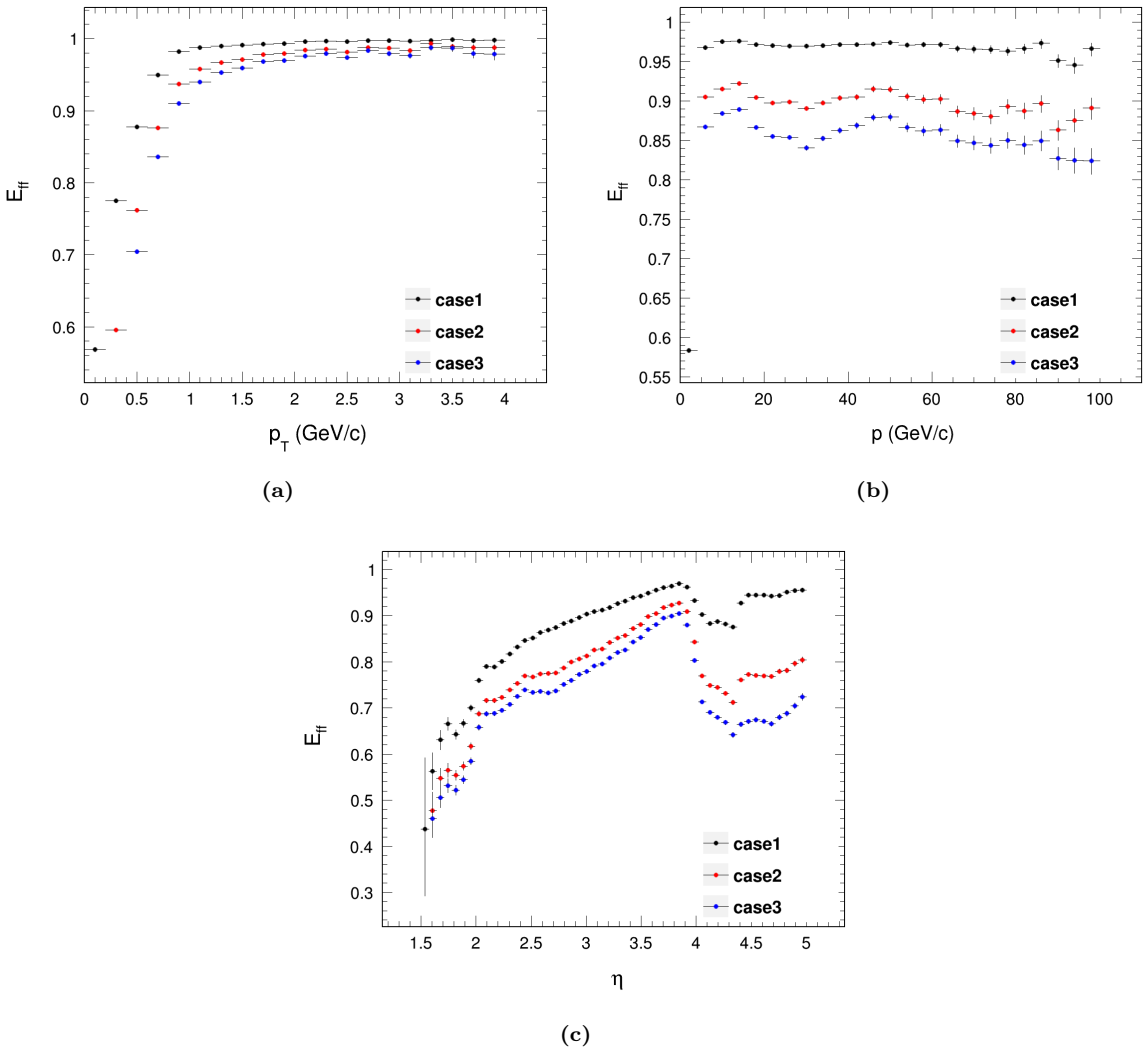
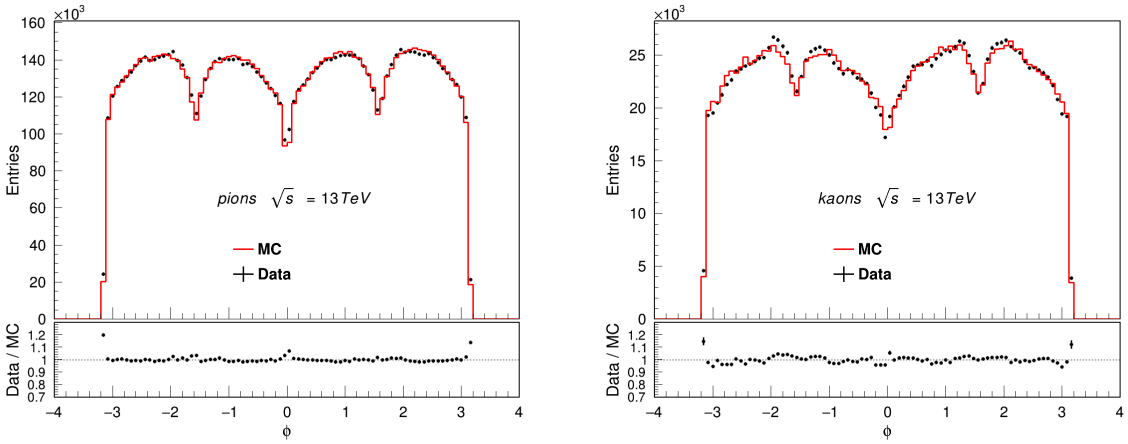


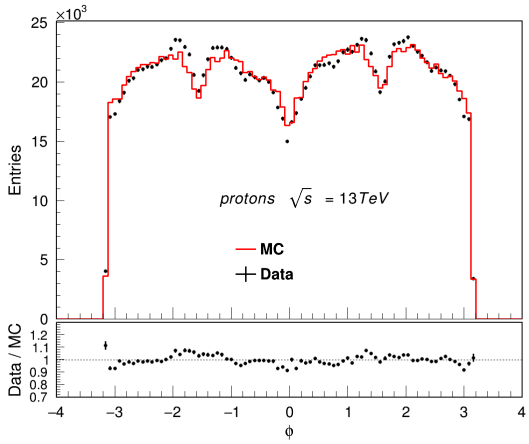
Figure A.20: Comparison of different settings of DLL cuts for the identification of protons with respect to transverse momentum (a), momentum (b) and η (c)..

A.0.8 Primary particles distribution for pions, kaons, and protons



(a)

(b)



(c)

Figure A.21: ϕ distribution of minimum-bias events for pions (a), kaons (b), and protons (c). The black points depict the data, and the red solid line depicts the MC distribution.

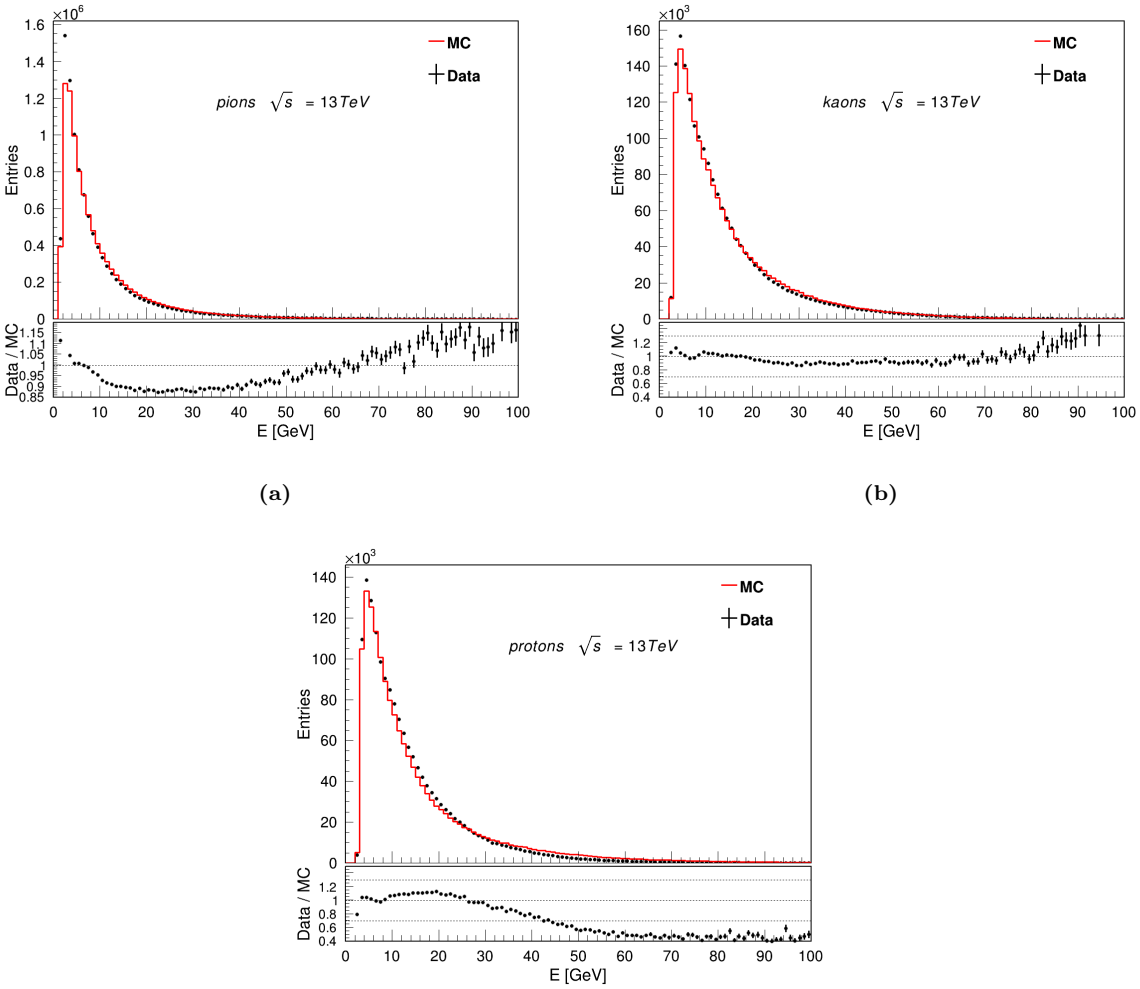


Figure A.22: Total Energy distribution of minimum-bias events for pions (a), kaons (b), and protons (c). The black points depict the data, and the red solid line depicts the MC distribution.

Bibliography

[1] Particle Data Group, T. et al. *Review of particle physics*, Phys. Rev. D **98** (2018) 030001.

[2] M. F. L'Annunziata, *Chapter 21 - hall of fame: Part x*, in *Radioactivity (Second Edition)* (M. F. L'Annunziata, ed.), pp. 729–798. Elsevier, Boston, second edition ed., 2016. doi: https://doi.org/10.1016/B978-0-444-63489-4.00021-6.

[3] M. Gell-Mann, *A schematic model of baryons and mesons*, Physics Letters **8** (1964), no. 3 214.

[4] R. A. et al. *Observation of the resonant character of the $z(4430)^-$ state*, Physical Review Letters **112** (2014), no. 22 222002.

[5] Particle Data Group, M. Tanabashi and H. et al. *Review of particle physics*, Phys. Rev. D **98** (2018) 030001.

[6] P. Langacker, *Grand Unification*, Scholarpedia **7** (2012), no. 10 11419.

[7] S. Hanawa, *Mathematical Formulation of the Gell-Mann-Nishijima Scheme for New Particles*, Progress of Theoretical Physics **17** (1957) 592.

[8] J. Baez, *Noether's theorem in a nutshell*, `https://math.ucr.edu/home/baez/noether.html`, 2, 2020.

[9] LHCb Collaboration, R. Aaij, *First observation of the decay $B_s^0 \to K^-\mu^+\nu_\mu$ and a measurement of $|V_{ub}|/|V_{cb}|$*, Phys. Rev. Lett. **126** (2021) 081804.

[10] B. Povh *et al.*, in *Quarks, Gluons, and the Strong Interaction*, pp. 103–122, Springer Berlin Heidelberg, Berlin, Heidelberg, 2015. doi: 10.1007/978-3-662-46321-5_8.

[11] M. SHIFMAN, in *Historical curiosity: how asymptotic freedom of the Yang-Mills theory could have been discovered three times before Gross, Wilczek, and Politzer, but was not*, pp. 126–130. arXiv:https://www.worldscientific.com/doi/pdf/10.1142/9789812810458_0010.

[12] P. A. Zyla *et al.*, *Review of Particle Physics*, PTEP **2020** (2020) 083C01, arXiv:2007.00642.

[13] A. M. S. et. al, *Running of the top quark mass from proton-proton collisions at* $\sqrt{s}$*= 13 TeV*, Physics Letters B **803** (2020) 135263.

[14] W. Christensen, *Completing the standard model with gravity by general relativizing quantum physics (rqp) (coupling spin-2 gravitons with spin-0 particles to generate higgs mass)*, Journal of Modern Physics **06** (2015) 1969.

[15] A. Dev, *Neutrino Oscillations and Mass Models*, arXiv:2310.17685.

[16] R. Collier, *Quantum Electrodynamics and Planck-Scale*, arXiv:1710.00618.

[17] M. Bauer, *Matter-antimatter Asymmetry, CP Violation and the Time Operator in Relativistic Quantum Mechanics*, arXiv:2208.13864.

[18] R. Durran, A. Neate, and A. Truman, *Newtonion Quantum Gravity, WIMPs and Celestial Mechanics - Summary of Results*, arXiv:2302.01992.

[19] *LHC Machine*, JINST **3** (2008) S08001.

[20] G. Stark, in *The Large Hadron Collider and the ATLAS Detector*, pp. 27–46, Springer International Publishing, Cham, 2020. doi: 10.1007/978-3-030-34548-8_3.

[21] B. Holzer and R. Alemany-Fernandez, in *A Journey to the Heart of the LHC*, T. Schörner-Sadenius, ed., pp. 27–56, Springer International Publishing, Cham, 2015. doi: 10.1007/978-3-319-15001-7_2.

[22] A. Collaboration, *ATLAS Pythia 8 tunes to 7 TeV data*, , https://cds.cern.ch/record/1966419.

[23] T. A. Collaboration, *The ALICE experiment at the CERN LHC*, Journal of Instrumentation **3** (2008) S08002.

[24] C. et al. *The CMS experiment at the CERN LHC*, Journal of Instrumentation **3** (2008) S08004.

[25] T. L. Collaboration, *The LHCb detector at the LHC*, Journal of Instrumentation **3** (2008) S08005.

[26] T. T. Collaboration, *The TOTEM experiment at the CERN LHC*, Journal of Instrumentation **3** (2008) S08007.

[27] LHCf, E. Berti, *The LHCf experiment: present status and physics results*, arXiv:1710.03991.

[28] MoEDAL-MAPP, B. Acharya *et al.*, *MoEDAL-MAPP, an LHC Dedicated Detector Search Facility*, in *Snowmass 2021*, 9, 2022. arXiv:2209.03988.

[29] R. Garoby, *Plans for upgrading the CERN proton accelerator complex*, Journal of Physics: Conference Series **110** (2008) 112003.

[30] L. Collaboration, *LHCb Tracker Upgrade Technical Design Report*, tech. rep., 2014. doi: 10.17181/CERN.OOOE.9O9T.

[31] G. Bruno, *Heavy flavour production at LHC: an overview*, EPJ Web of Conferences **70** (2014) .

[32] R. e. a. Aaij, *Search for lepton-flavour-violating decays of higgs-like bosons*, The European Physical Journal C **78** (2018) .

[33] A. Oyanguren, *B decay anomalies at LHCb*, EPJ Web of Conferences **175** (2018) 01004.

[34] J. Libby, *VELO: the LHCb vertex detector*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **494** (2002), no. 1 113, Proceedings of the 8th International Conference on Instrumentation for Colliding Beam Physics.

[35] LHCb, M. S. Rudolph, *The LHCb Upstream Tracker Upgrade*, PoS **Vertex2019** (2020) 013.

[36] LHCb, T. Kirn, *SciFi – A large scintillating fibre tracker for LHCb*, Nucl. Instrum. Meth. A **845** (2017) 481.

[37] A. Badalov, D. Cámpora, N. Neufeld, and X. Vilasis-Cardona, *LHCb GPU acceleration project*, Journal of Instrumentation **11** (2016) P01001.

[38] E. Halnon, *Nature's Elusive Secrets — artsandsciences.syracuse.edu*, https://artsandsciences.syracuse.edu/news-all/news-2023/natures-elusive-secrets/, 2023.

[39] S. J. (CERN), *Installation of the second half of the LHCb SciFi detector — ep-news.web.cern.ch*, https://ep-news.web.cern.ch/content/installation-second-half-lhcb-scifi-detector, 2022. [Accessed 05-08-2024].

[40] *LHCb Magnet — lhcb-magnet.web.cern.ch*, https://lhcb-magnet.web.cern.ch/. [Accessed 29-01-2025].

[41] S. Gambetta, *The LHCb RICH detectors: Operations and performance*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **952** (2020) 161882, 10th International Workshop on Ring Imaging Cherenkov Detectors (RICH 2018).

[42] *RICH detectors — lhcb-outreach.web.cern.ch*, https://lhcb-outreach.web.cern.ch/detector/rich-detectors/. [Accessed 29-01-2025].

[43] C. Abellán Beteta *et al.*, *Calibration and performance of the LHCb calorimeters in Run 1 and 2 at the LHC*, arXiv:2008.11556.

[44] S. Ek-In, *Reconstruction of semileptonic decays and search for* $B_s^0 \to K^{*+}{}_\mu$ *at the LHCb experiment*, PhD thesis, EPFL - Swiss Federal Technology Institute of Lausanne, 2018, https://www.epfl.ch/labs/lphe/wpcontent/uploads/MasterThesis_Surapat.pdf.

[45] E. Santovetti, *The LHCb muon detector*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **462** (2001), no. 1 297, BEAUTY2000, Proceedings of the 7th Int. Conf. on B-Physics at Hadron Machines.

[46] LHCb, A. et al. *Measurement of the* $B_s^0 \to \mu^+\mu^-$ *decay properties and search for the* $B^0 \to \mu^+\mu^-$ *and* $B_s^0 \to \mu^+\mu^-\gamma$ *decays.* , Phys. Rev. D **105** (2022), no. 1 012010, arXiv:2108.09283.

[47] X. Cid Vidal, *Muon Identification in the LHCb experiment*, arXiv:1005.2585.

[48] R. Quagliani, *Novel real-time alignment and calibration of LHCb detector for Run II and tracking for the upgrade.*, Journal of Physics: Conference Series **762** (2016) 012046.

[49] LHCb, R. Aaij *et al.*, *Physics case for an LHCb Upgrade II - Opportunities in flavour physics, and beyond, in the HL-LHC era*, tech. rep., CERN, Geneva, 2016. ISBN 978-92-9083-494-6, doi: 10.17181/CERN.QZRZ.R4S6.

[50] J. Hammerich, *Towards MightyPix, an HV-MAPS for the LHCb mighty tracker upgrade*, Journal of Instrumentation **17** (2022) C10005.

[51] P. Billoir, M. De Cian, P. A. Günther, and S. Stemmle, *A parametrized kalman filter for fast track fitting at LHCb*, Computer Physics Communications **265** (2021) 108026.

[52] S. Aiola *et al.*, *Hybrid seeding: A standalone track reconstruction algorithm for scintillating fibre tracker at lhcb*, Computer Physics Communications **260** (2021) 107713.

[53] S. Aiola *et al.*, *Hybrid seeding: A standalone track reconstruction algorithm for scintillating fibre tracker at lhcb*, Computer Physics Communications **260** (2021) 107713.

[54] E. Rodrigues, *Dealing with clones in the tracking*, , https://cds.cern.ch/record/1000723.

[55] G. Popescu, *Semiconductor materials*, in *Principles of Biophotonics, Volume 2*, 2053-2563, pp. 12–1 to 12–21. IOP Publishing, 2019. doi: 10.1088/978-0-7503-1644-6ch12.

[56] *Band gap - Energy Education — energyeducation.ca*, https://energyeducation.ca/encyclopedia/Band_gap. [Accessed 19-09-2024].

[57] F. Ravotti, *Development and Characterisation of Radiation Monitoring Sensors for the High Energy Physics Experiments of the CERN LHC Accelerator*, 2006. http://www.documents.cern.ch/cgi-bin/setlink?base=preprint&categ=cern&id=cern-thesis-2007-013.

[58] G. Lindstroem, *Displacement damage in silicon, on-line compilation*, https://rd50.web.cern.ch/niel/. Accessed: 2025-5-8.

[59] A. Oblakowska-Mucha and T. Szumlak, *Comparison of general-purpose events generators for particle fluence simulation in lhc environment*, 2020. doi: https://doi.org/10.1140/epjp/s13360-021-02012-8.

[60] A. Affolder *et al.*, *Radiation damage in the LHCb vertex locator*, Journal of Instrumentation **8** (2013) P08002.

[61] C. Ahdida *et al.*, *New Capabilities of the FLUKA Multi-Purpose Code*, Front. Phys. **9** (2022) 788253.

[62] G. G. Daquino, G. Corti, and G. Folger, *Background radiation studies at LHCb using Geant4*, in *IEEE Symposium Conference Record Nuclear Science 2004.*, vol. 4, pp. 2191–2195 Vol. 4, 2004. doi: 10.1109/NSSMIC.2004.1462697.

[63] T. Sjöstrand and et al. *An introduction to PYTHIA 8.2*, Computer Physics Communications **191** (2015) 159.

[64] G. Bewick, S. F. Ravasio, P. Richardson, and M. H. Seymour, *Initial state radiation in the* Herwig *7 angular-ordered parton shower*, Journal of High Energy Physics **2022** (2022) .

[65] E. Bothmann *et al.*, *Event generation with Sherpa 2.2*, SciPost Phys. **7** (2019) 034.

[66] G. Corti, A. Davis, and M. Kreps, *The Simulation Project*, tech. rep., CERN, Geneva, 2021. https://cds.cern.ch/record/2761006.

[67] O. Brüning, M. Klein, LHeC, and P. Collaboration, *Exploring the energy frontier with deep inelastic scattering at the LHC*, Journal of Physics G: Nuclear and Particle Physics **46** (2019) 123001.

[68] J. Van Tilburg, *Track simulation and reconstruction in LHCb*, 2005. https://cds.cern.ch/record/885750.

[69] J. Leygonie, *Understanding rt using two-particle correlations by simulating proton-proton collisions in a monte-carlo model*, Student Paper, 2021. https://lup.lub.lu.se/lupStat/record/9040372.

[70] J.-P. Guillaud, *Miniguide PYTHIA*, tech. rep., CERN, Geneva, 2000. https://cds.cern.ch/record/687295.

[71] T. Sjostrand, *Status of Fragmentation Models*, Int. J. Mod. Phys. A **3** (1988) 751.

[72] S. Ferreres-Solé and T. Sjöstrand, *The space–time structure of hadronization in the lund model*, The European Physical Journal C **78** (2018) .

[73] G. Corcella *et al.*, HERWIG *6: an event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)*, Journal of High Energy Physics **2001** (2001) 010–010.

[74] M. Markovych and A. Tandogan, *Analytic Evolution of DGLAP Equations*, arXiv:2304.10458.

[75] C. B. Duncan and P. Kirchgaeßer, *Kinematic strangeness production in cluster hadronization*, The European Physical Journal C **79** (2019) .

[76] S. Gieseke, F. Loshaj, and P. Kirchgaeßer, *Soft and diffractive scattering with the cluster model in* Herwig, The European Physical Journal C **77** (2017) .

[77] S. Sugimoto, *Hadrons from string theory*, Few-Body Systems **54** (2012) 71–75.

[78] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjöstrand, *Parton fragmentation and string dynamics*, Physics Reports **97** (1983), no. 2 31.

[79] A. Kupco, *Cluster hadronization in HERWIG 5.9*, in *Workshop on Monte Carlo Generators for HERA Physics (Plenary Starting Meeting)*, pp. 292–300, 4, 1998. arXiv:hep-ph/9906412.

[80] S. Jain, A. K. Gupta, and N. K. Mondal, *Search for quark-lepton compositeness at fermilab tevatron and cern lhc*, Physical Review D **62** (2000) .

[81] C. Bierlich, S. Chakraborty, G. Gustafson, and L. Lönnblad, *Setting the string shoving picture in a new frame*, Journal of High Energy Physics **2021** (2021) .

[82] C. Bierlich, G. Gustafson, L. Lönnblad, and A. Tarasov, *Effects of overlapping strings in pp collisions*, 2015. doi: https://doi.org/10.1007/JHEP03(2015)148.

[83] A. Ortiz Velasquez *et al.*, *Color reconnection and flowlike patterns in pp collisions*, Phys. Rev. Lett. **111** (2013) 042001.

[84] CMS Collaboration *et al.*, *Cms pythia 8 colour reconnection tunes based on underlying-event data*, European Physical Journal C - Particles Fields **83** (2023) .

[85] Sirunyan *et al.*, *Extraction & validation of new set of CMS* Pythia *8 tunes from underlying-event measurements*, The European Physical Journal C **80** (2020) .

[86] M. M. Block and F. Halzen, *Analyticity constraints on the Landshoff-Donnachie fit to total cross sections*, Physical Review D **74** (2006) .

[87] Rivet — *the particle-physics mc analysis toolkit*, https://rivet.hepforge.org/.

[88] Professor, https://professor.hepforge.org/.

[89] S. Amoroso, *Tuning of event generators to measurements of tt production and a general search for new physics with the ATLAS experiment*, PhD thesis, Freiburg U., Freiburg U., 7, 2015, doi: 10.6094/UNIFR/14921.

[90] C. Bierlich and A. B. et al. *Robust independent validation of experiment and theory:* Rivet *version 3*, SciPost Physics **8** (2020) .

[91] H. Schulz, *Systematic Event Generator Tuning with* Professor, PhD thesis, Humboldt-Universität zu Berlin, 2009, "https://professor.hepforge.org/diplomathesis_h_schulz.pdf".

[92] A. Buckley *et al.*, *Systematic event generator tuning for the LHC*, Eur. Phys. J. C **65** (2010) 331, arXiv:0907.2973.

[93] M. Clemencic *et al.*, *Recent developments in the LHCb software framework gaudi*, Journal of Physics: Conference Series **219** (2010) 042006.

[94] *LHCb Gitlab*, https://gitlab.cern.ch/lhcb.

[95] M. Mazurek, G. Corti, and D. Muller, *New simulation software technologies at the LHCb Experiment at CERN*, tech. rep., 2021. doi: doi = 10.1088/1742-6596/2438/1/012108,.

[96] M. Mazurek, M. Clemencic, and G. Corti, *Gauss and Gaussino: the LHCb simulation software and its new experiment agnostic core framework*, PoS **ICHEP2022** (2022) 225.

[97] R. Aaij *et al.*, *Measurement of V^0 production ratios in pp collisions at $\sqrt{s}$ = 0.9 and 7 TeV*, Journal of High Energy Physics **2011** (2011) .

[98] R. Aaij *et al.*, *Measurement of the inclusive ϕ cross-section in pp collisions at $\sqrt{7}$ TeV.*, Physics Letters B **703** (2011) 267–273.

[99] R. Aaij *et al.*, *Measurement of prompt hadron production ratios in pp collisions at $\sqrt{s} = 0.9$ and 7* TeV, The European Physical Journal C **72** (2012) .

[100] R. Aaij *et al.*, *Measurement of the forward energy flow in pp collisions at $\sqrt{s}$ = 7 TeV*, The European Physical Journal C **73** (2013) .

[101] R. Aaij *et al.*, *Measurement of charged particle multiplicities and densities in pp collisions at $\sqrt{s} = 7$ TeV in the forward region*, The European Physical Journal C **74** (2014) .

[102] R. Aaij *et al.*, *Measurement of the inelastic pp cross-section at a centre-of-mass energy of $\sqrt{s}$ = 7 TeV*, Journal of High Energy Physics **2015** (2015) .

[103] R. Aaij *et al.*, *Measurement of the inelastic pp cross-section at a centre-of-mass energy of 13 TeV*, Journal of High Energy Physics **2018** (2018) .

[104] G. Aad *et al.*, *Measurement of underlying event characteristics using charged particles in pp collisions at $\sqrt{s} = 900$ GeV and 7 TeV with the ATLAS detector*, Phys. Rev. D **83** (2011) 112001, arXiv:1012.0791.

[105] G. Aad *et al.*, *Kshort and Λ production in pp interactions at $\sqrt{s} = 0.9$ and 7 TeV measured with the ATLAS detector at the LHC*, Phys. Rev. D **85** (2012) 012001, arXiv:1111.1297.

[106] V. Khachatryan *et al.*, *Strange Particle Production in pp Collisions at $\sqrt{s} = 0.9$ and 7 TeV*, JHEP **05** (2011) 064, arXiv:1102.4282.

[107] S. Chatrchyan *et al.*, *Study of the Underlying Event at Forward Rapidity in pp Collisions at $\sqrt{s}$ = 0.9, 2.76, and 7 TeV*, JHEP **04** (2013) 072, arXiv:1302.2394.

[108] B. Abelev *et al.*, *Neutral pion and η meson production in proton-proton collisions at $\sqrt{s} = 0.9$ TeV and $\sqrt{s} = 7$ TeV*, Phys. Lett. B **717** (2012) 162, arXiv:1205.5724.

[109] S. Thor, *LHCb generator tuning with* Professor *and* Rivet, 2022. https://lhcb-generator-tuning.docs.cern.ch/latest/.

[110] R. Aaij *et al.*, *Measurement of charged particle multiplicities and densities in pp collisions at* $\sqrt{s} = 7$ *TeV in the forward region*, The European Physical Journal C **74** (2014) .

[111] P. Sarma, B. Barman, and B. Bhattacharjee, *Effect of color reconnection on multiplicity dependent charged particle production in* Pythia*-generated pp collisions at the LHC energies*, 2022. doi: 10.1140/epja/s10050-023-00989-7.

[112] C. Collaboration, *CMS* Pythia *8 colour reconnection tunes based on underlying-event data*, The European Physical Journal C **83** (2023), no. 7 Art.

[113] *A study of different colour reconnection settings for* Pythia *8 generator using underlying event observables*, tech. rep., CERN, Geneva, 2017. https://cds.cern.ch/record/2262253.

[114] LHCb, P. Ilten, *Soft QCD Measurements at LHCb*, in *2nd Large Hadron Collider Physics Conference*, 9, 2014. arXiv:1409.0434.

[115] V. Pascuzzi, *Looking for Beyond the Standard Model Physics in Dijet-Plus-Lepton Final-State Events Collected with the ATLAS Detector*, PhD thesis, 11, 2019, doi: 10.13140/RG.2.2.33848.90885.

[116] J. Wenninger, *Operation and Configuration of the LHC in Run 2*, , https://cds.cern.ch/record/2668326.

[117] I. Belyaev *et al.*, *The history of LHCb*, The European Physical Journal H **46** (2021) .

[118] T. T. Collaboration *et al.*, *First measurement of elastic, inelastic and total cross-section at* $\sqrt{s} = 13$ *TeV by TOTEM and overview of cross-section data at LHC energies*, 2018. doi: https://doi.org/10.1140/epjc/s10052-019-6567-0.

[119] *Lhcb detector performance*, International Journal of Modern Physics A **30** (2015), no. 07 1530022, arXiv:https://doi.org/10.1142/S0217751X15300227.

[120] L. Anderlini *et al.*, *Computing strategy for PID calibration samples for LHCb Run 2*, tech. rep., CERN, Geneva, 2016. https://cds.cern.ch/record/2199780.

[121] S. Oggero, *Beauty in the Crowd:: Commissioning of the LHCb Pile-Up detector and First evidence of* $B_s \to \mu^+\mu^-$, phd-thesis - research and graduation internal, Vrije Universiteit Amsterdam, 2013, https://research.vu.nl/ws/portalfiles/portal/42112053/chapter%203_%20The%20LHCb%20experiment.pdf.

[122] R. Aaij *et al.*, *Study of forward Z + jet production in pp collisions at* $\sqrt{s} = 7$ *TeV*, JHEP **01** (2014) 033, arXiv:1310.8197.

[123] V. Coco *et al.*, *VELO Upgrade Module Nomenclature*, tech. rep., CERN, Geneva, 2019. https://cds.cern.ch/record/2683878.

[124] P. Svihra, *The design and construction of lhcb velo upgrades modules*, tech. rep., CERN, 2020. doi: 10.1109/NSS/MIC42101.2019.9059693.

[125] L. Eklund, *The LHCb VELO upgrade*, Nuclear and Particle Physics Proceedings **273-275** (2016) 1079.

[126] E. R. Rodríguez, *The LHCb VELO detector: Design, operation and first results*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **1065** (2024) 169469.

[127] K. Akiba *et al.*, *Radiation Damage Effects and Operation of the LHCb Vertex Locator. Radiation Damage Effects and Operation of the LHCb Vertex Locator*, IEEE Trans. Nucl. Sci. **65** (2018), no. 5 1127.

[128] E. Buchanan, *The LHCb vertex locator and silicon tracker operation performance Run II*, p. 016, 07, 2018. doi: 10.22323/1.309.0016.

[129] A. B. Chilton, *A note on the fluence concept*, Health Phys. **34** (1978) 715–716.