Measurement of Charged-Particle Distributions in Proton–Proton Interactions at $\sqrt{s} = 8$ TeV with the ATLAS Detector at the LHC

Wolfgang Lukas

A dissertation submitted to the University of Innsbruck
for the degree of Doctor of Philosophy (PhD)

Supervisors:

Ao. Univ.-Prof. Dr. Emmerich Kneringer
University of Innsbruck, Institute for Astro- and Particle Physics

Ao. Univ.-Prof. i.R. Dr. Gerald Rudolph
University of Innsbruck, Institute for Astro- and Particle Physics

Dr. Andreas Salzburger
CERN

Graz, Austria — August 16, 2016

The logo on the title page of this thesis is taken from the
official webpage of the University of Innsbruck:
`https://www.uibk.ac.at/`

Leopold-Franzens-Universität Innsbruck

Eidesstattliche Erklärung

Ich erkläre hiermit an Eides statt durch meine eigenhändige Unterschrift, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe. Alle Stellen, die wörtlich oder inhaltlich den angegebenen Quellen entnommen wurden, sind als solche kenntlich gemacht.

Die vorliegende Arbeit wurde bisher in gleicher oder ähnlicher Form noch nicht als Magister-/Master-/Diplomarbeit/Dissertation eingereicht.

16. August 2016
Datum

Unterschrift

Abstract

The majority of proton–proton interactions at the LHC can be described as inelastic soft-QCD processes, characterised by a small momentum transfer between their constituents. These interactions cannot be calculated from first principles with perturbation theory, and must hence be described by phenomenological models that are improved by tuning their parameters against experimental data. Measurements of the hadronic final states, presented as distributions of stable primary charged particles that were produced in the hadron collisions, can be used as an input for the tuning. To this end, distributions of reconstructed events and tracks are extracted from the recorded collision data, which must be corrected and unfolded to account for detector effects such as selection and reconstruction inefficiencies, migration effects and background contaminations. Some of these corrections are derived from Monte Carlo simulations of the full detector response to the particles that were generated by the phenomenological models. The dominant source of systematic uncertainty in this measurement is found to be the uncertainty of the material description used in the simulation.

This thesis describes the successful efforts to improve the detailed material description of the inner tracker of ATLAS, resulting in a more accurate simulation geometry with significantly better constraints on this dominant source of systematic uncertainty. Building upon this knowledge, the measurement of primary-charged-particle distributions is performed on data recorded in 2012 with the ATLAS detector at $\sqrt{s} = 8$ TeV centre-of-mass energy, utilising a minimum-bias trigger and a special low-luminosity experimental setup. This study aims to achieve the highest precision of all ATLAS minimum-bias measurements from the Run-1 data-taking period of the LHC between 2009 and 2013. Results are obtained in five different regions of particle kinematics and multiplicity, with transverse-momentum thresholds of $p_{\mathrm{T}} > 100$ and 500 MeV. Among these, two high-multiplicity phase-space regions are measured by ATLAS for the first time, thus extending the scope of the measurements. Hadronic final-state distributions of primary-charged-particle multiplicity, pseudorapidity and transverse momentum are presented along with carefully evaluated total systematic and statistical uncertainties. The results are compared to predictions made by various phenomenological models and thus provide valuable constraints for their potential further improvement, while also contributing to a better theoretical understanding of the soft-QCD sector of particle physics.

Zusammenfassung

Der überwiegende Anteil von Proton–Proton-Wechselwirkungen am LHC kann durch inelastische Soft-QCD-Prozesse beschrieben werden, welche durch einen geringen Impulstransfer zwischen den Partonen charakterisierbar sind. Diese Gruppe von Wechselwirkungen lässt sich jedoch nicht mittels Störungstheorie berechnen und muss stattdessen durch phänomenologische Modelle beschrieben werden, deren Parameter man durch Vergleiche mit experimentellen Daten feinabstimmt und somit an die experimentellen Ergebnisse angleicht. Hierzu werden Messungen von hadronischen Endzuständen der Teilchenkollisionen herangezogen, welche als Verteilungen von stabilen geladenen Primärteilchen dargestellt werden. Zu diesem Zweck werden aus den gemessenen Daten zunächst Verteilungen von Observablen rekonstruierter Kollisionsereignisse und Teilchenspuren extrahiert. Diese müssen sodann korrigiert und entfaltet werden, um Detektoreffekte wie Ineffizienzen der Datenselektion und -rekonstruktion, Migrationseffekte sowie Kontaminationen durch Messuntergründe zu kompensieren. Einige dieser Korrekturen werden aus komplexen Monte-Carlo-Simulationen gewonnen, indem man die aus den phänomenologischen Modellen erzeugten Teilchen, ihre Wechselwirkung mit dem Detektor und die daraus resultierenden Detektorsignale simuliert und auswertet. Die dominante Quelle systematischer Unsicherheiten dieser Messung besteht in der Unsicherheit der zur Simulation verwendeten Geometrie- und Materialbeschreibung des Detektors.

Die vorliegende Arbeit beschreibt zunächst die erfolgreichen Bemühungen, detaillierte Kenntnis über die Materialverteilung des inneren ATLAS-Spurdetektors zu gewinnen und damit das für Simulationen verwendete Detektormodell weiter zu präzisieren, wodurch eine deutliche Verringerung des Beitrags aus dieser Quelle systematischer Unsicherheiten ermöglicht wird. Auf dieser Grundlage wird eine Präzisionsmessung von Verteilungen geladener Primärteilchen aus experimentellen Daten durchgeführt, die mit dem ATLAS-Detektor im Jahr 2012 bei Teilchenkollisionen mit einer Massenschwerpunktsenergie von $\sqrt{s} = 8$ TeV unter Verwendung eines Minimum-Bias-Triggers und unter speziellen experimentellen Bedingungen bei geringer Luminosität aufgezeichnet wurden. Für die Resultate dieser Studie wird die höchste Präzision aller vergleichbaren Messungen von ATLAS aus der Run-1-Datenaufzeichnungsperiode des LHC (2009-2013) angestrebt. Ergebnisse werden in fünf verschiedenen Regionen von Teilchenkinematik und -multiplizität dargestellt, wobei die untere Grenze der transversalen

Teilchenimpulse mit $p_{\mathrm{T}} > 100$ und 500 MeV definiert wird. Zwei Phasenräume mit hohen Teilchenmultiplizitäten werden dabei erstmalig von ATLAS gemessen und erweitern somit Geltungsbereich und Aussagekraft der Studie. Hadronische Endzustände werden schließlich als Verteilungen von Multiplizität, Pseudorapidität und Transversalimpuls der geladenen Primärteilchen präsentiert, wobei große Sorgfalt auf die Ermittlung und Darstellung systematischer und statistischer Unsicherheiten gelegt wird. Die Ergebnisse werden mit Vorhersagen verschiedener phänomenologischer Modelle verglichen und liefern so eine wertvolle Grundlage für deren künftige weitere Verbesserung und Feinabstimmung der Modellparameter. Die vorliegende Arbeit trägt somit auch zu einem verbesserten theoretischen Verständnis des Soft-QCD-Sektors der Teilchenphysik bei.

Acknowledgements

This work would not have been possible without the knowledge, work and support provided by many individuals in the past and present. I gratefully acknowledge the countless giants upon whose shoulders I stand — in science and all other domains of life. Here I wish to explicitly thank but a few of them: my supervisors, colleagues, friends and family.

Beginning with my institute at the University of Innsbruck, first of all I'd like to express my gratitude to my supervisor Prof. Emmerich Kneringer, for his incredible amount of support, encouragement and expert advice on many levels. I am very grateful to Prof. Gerald Rudolph for his detailed advice and insightful conversations about MC tunes and particle physics in general. Our project leader Prof. Dietmar Kuhn made my ongoing affiliation with our institute possible and gave me enormous support and motivation. Furthermore, he amazingly encouraged me to follow my heart and interests, even though he knew well my passion for detours; it worked out at last! I thank the head of our institute Prof. Olaf Reimer and all other staff (especially Wolfgang, Patrick and Anna) for the welcoming and delightful atmosphere at my *alma mater*. My special gratitude goes to Leonie Werner from our institute secretariat, who has always been tremendously helpful and kind. Gregor often hosted me during my stays in Innsbruck and lifted up my spirits with his uncomplicated, heartfelt nature; I hope we will continue to play chess!

Among my CERN colleagues, I first and foremost wish to thank my supervisor Andreas Salzburger for his never-ending support, encouragement and pragmatic advice, based upon his tremendous expertise in track reconstruction, simulation and physics analyses. It was a pleasure to share the office with you and to repeatedly co-conspire in the dematerialisation of my stash of dark chocolate! I thank Michael Duehrssen-Debling for good advice and his water cooker, which helped me economise during times of excessive coffee and maté consumption. I am grateful for having been able to work under Markus Elsing's professional and supportive group leadership.

Simone Pagan Griso miraculously juggled ten thousand responsibilities and still found time for detailed responses to my questions; I also thank him for his appreciative nature which greatly energised me during exhausting times. Thorsten Kuhl would often spontaneously materialise behind me in the office doorstep, glance over my shoulder, and ingeniously point out helpful solutions to me. I also thank Oldrich Kepka for his advice and expert involvement in the analysis. I gladly recall the coffee conversations with Edward Sarkisyan-Grinbaum after our weekly analysis meetings, which brought me new ideas and enlarged the scope of my understanding of soft-QCD physics. My gratitude extends to Morten Medici for our collaboration on the analysis and our late-night Skype sessions, in which we supported each other during times of struggle. I also thank David Milstead and Wayne Cribbs who came to the Minimum Bias analysis at

the very right time. Dave's support during our paper manuscript write-up and the intricate ATLAS-internal peer-review phase allowed me to finish my thesis while also bringing our analysis to fruition without entirely losing my mind in the process. *Tack så mycket!* Thanks also to my other analysis colleagues, the soft-QCD and Standard Model conveners (especially Tim, Sasha, Alessandro, Jan and Ulla) and our editorial board members for insightful discussions and competent guidance along the road to publication; this process taught me to look much deeper.

Petya Lilova, Nathalie Schwarzbauer and others at the ATLAS secretariat have always been so helpful, patient and friendly that I propose to call them *the elixir of life that runs through the veins of the ATLAS collaboration.* This lengthy title would certainly extend to the administrative, medical and other service personnel, with special mention of the cheerful staff at CERN cafeterias and restaurants. My most inspired thanks go to Claudia Marcelloni and the first TEDxCERN team for an amazing volunteer experience in May 2013 (the T-shirt remains one of my all-time favourites!), which would open many doors later on.

Among my wonderful friends at CERN and in Switzerland and France, I especially thank Elmar, Manuel, Alex, Andrew, Vince, Dan and Danny for truly inspiring times and for standing firmly together in our shared *maison*; Elmar again for friendship, cooking, humour, hospitality and driving instructions; Florian for our friendship; Mario who started it all long ago by infecting me with the idea to apply for the CERN Summer Student programme in 2007; Nina, Michi, Ricardo, Philip, Robert, Roland, and others from our *Los Lunchos* (sic!) group; Jakob, Harris and Olga, Nadia (hugs!), Cat, and many others. I sincerely hope you all will continue to enjoy your lives, no matter how far away; we will stay close in our hearts, and I hope we will reunite often. A special hug goes to my friends Slavi and Achim for being part of my life, and for blessing me with your hospitality during my travels between CERN and my family in Austria. Thank you, Achim, for coaching me to keep on track during the tedious *grande finale.* I also thank Moh and all the cheerful teachers and dancers at the Salsa Geneva school. Taking up salsa classes there was one of the best decisions of my life!

It goes without saying, and yet shall be emphasised no less, that without my family I would not be who I am, nor would I have accomplished this thesis without them. First and foremost I thank my mother and father, my late grandparents, and my brother; my uncle and aunts and cousins; my nephews and nieces; and all their loved ones who have come into our family. Likewise I am grateful for all friends who have co-shaped and inspired my path in life: Ling, Tanya, Tania, Hope, Lea, Tom, Klaus, Newcomb, Scott and Susan *(thank you — may the Yoga flourish!)*, Mark, Stuart, Clare, Eric, Josh, Marie, Will, Beau, Eric and Ronit, and many others. My dear friend Arnold gave me one of the wisest pieces of advice when he encouraged me several months ago to rekindle my Buddhist meditation practice — not *despite of*, but rather *because of* the intensity of the final stages of my work. Thus one of the most significant periods of my life, which I had anticipated to be a horrendous torture and to be endured with gritted teeth, was transformed into a most fruitful experience of curious inquiry, learning and realisation.

With all my heart I thank my wonderful and beloved partner Sylvia. I would not have met you (and your family and friends) if all these people and conditions had not congenially conspired in exactly this way. For someone who enjoys to meander as much as I do, a down-to-earth partner who can easily keep me firmly rooted on the ground is a wonderful blessing. To share our love and lives in such mutual openness is an even greater gift! <3

My time at CERN and the University of Innsbruck has allowed me to experience first-hand the scope and benefits of international collaboration. This experience has reaffirmed my belief that peaceful co-existence and *interbeing* is not only possible, but indeed absolutely essential for the sake of evolution and flourishing of all sentient beings. My earnest hope is that this work and thesis may bring some benefit that outweighs the footprint of spent resources, and may help to increase our collective understanding of the universe in which we live, if only a little.

CERN implements an open-access policy and makes its results publicly available; I profoundly respect this bold approach of sharing our knowledge and am eager to promote it further. Finally, I would like to thank the Austrian Ministry of Science, Research and Economy (BMWFW) for its financial support in the "CERN Austrian Doctoral Student Programme" which made my work at CERN possible.

Contents

"Speculation about [the origin and purpose of] the cosmos
is an imponderable that is not to be speculated about.
Whoever speculates about these things
will go mad and experience vexation."

— Gautama Siddhartha, $\sim$ 500 BCE (Anguttara Nikaya 4.77)

Chapter 1

Introduction

The Standard Model of Particle Physics (SM, Chapter 2) is widely regarded as one of the most successful theoretical frameworks in modern physics. It consistently describes all currently known elementary particles as well as three of their fundamental physics interactions, which are the theories of strong, weak and electromagnetic interactions. The predictions of the SM have been verified by searches for these elementary particles (such as quarks, bosons and mesons) and precise measurements of their properties (such as masses, lifetimes and decay products). These include the discoveries and measurements made at CERN of the Z^0 and $W^\pm$ bosons, the top quark, and recently the Higgs boson.[1]

CERN is an international organisation for nuclear and particle physics research which hosts the world's largest particle physics laboratory complex, including a circular particle accelerator called the Large Hadron Collider (LHC) [1]. The facilities of CERN are located near Geneva, Switzerland and have been designed for the experimental as well as theoretical study of the constituents of matter, mainly through particle collisions at the high-energy frontier under controlled laboratory conditions. Such experiments are designed to probe the regimes of physics phenomena at very small timescales and sub-atomic distances. The masses and energy-dependent production cross sections of many particles under study require to perform the collisions with centre-of-mass energies up to the TeV regime.[2] In order to provide the necessary experimental conditions for such studies, particle colliders (along with measurement devices called particle detectors) have been installed since the mid-20^{th} century in various laboratories around the world. Among these, the LHC machine delivers proton–proton (pp) as well as heavy-ion collisions at the highest centre-of-mass energies that have been achieved in laboratories until today, up to $\sqrt{s} = 13$ TeV.

Despite its large success in describing the fundamental constituents of matter, the SM is confronted with significant limitations. Most notably, it does not account for roughly 95 % of the energy density of our universe, which are hypothesised to exist in forms such as Dark Matter

[1] First evidence for the Higgs boson discovery at CERN has been publicly announced in July 2012. In the original announcement, these results were more cautiously described as statistical evidence for the existence of a Higgs-like particle, i.e. a particle sharing the same characteristics as those attributed to the SM Higgs boson. It has not been clarified until today whether the discovered particle is the only Higgs boson; some theories suggest the existence of other Higgs bosons at higher mass.

[2] For instance, the mass of the Higgs boson has been measured to be approximately 125.1 GeV by the two largest collaborations at the LHC.

and Dark Energy that physicists have not been able to confirm in laboratory settings until today. Furthermore, it cannot directly predict some properties of fundamental particles, such as the Higgs boson mass, which must hence be experimentally determined. The SM also does not account for gravity, which is one of the four known fundamental physics interactions. Several theories have thus been devised to extend the SM, such as Supersymmetry (SUSY) or string theories, and their predictions can be probed by searches for evidence of new physics beyond the SM. If such evidence is found, our most fundamental theoretical framework of particle physics will need to be revised.

In the theory of strong interactions called Quantum Chromodynamics (QCD, Section 2.1.1) high-energy hadron collisions can be described in terms of strong interactions of their constituents, which are called partons. Among these interactions, processes with large momentum transfer, called hard-scattering interactions (which can produce heavy bosons, highly-energetic particle jets, etc.), can be calculated from first principles using a powerful mathematical framework called perturbation theory. However, the vast majority of parton interactions involve low momentum transfers, for which calculations become difficult, as perturbation theory is not applicable at such scales. In order to describe these "soft" interaction processes (Soft QCD, Section 2.3), which also underlie the hard-scattering interactions, phenomenological models have been introduced. These models contain large numbers of free parameters, which must be "tuned" in such a way as to optimally reproduce the existing experimental data.

One way to test and refine these phenomenological models, and thus to obtain a deeper understanding of the soft-QCD sector of SM physics, is the experimental study of charged particles produced in hadron collisions at a wide range of centre-of-mass energies. In order to measure observables of these charged particles, particle detectors are equipped with tracking devices that are built around the designed primary interaction region in which the hadrons collide. Such tracking devices usually employ a magnetic field in order to measure the transverse component of charged-particle momenta, which can be derived from the curvature of their trajectories within the magnetic field. Following a hadron collision, the produced particles (mainly pions, kaons and protons, i.e. hadrons) traverse the detector from the primary interaction region outwards and interact with sensitive detector elements, where they give rise to measurable electronic signals. In ATLAS the central tracking device is called the Inner Detector (ID), which is also the most important component for the measurement of primary charged particles. The experimental setup for the study of charged-particle distributions presented in this thesis uses data from pp collisions at a centre-of-mass energy of $\sqrt{s} = 8$ TeV, recorded with the ATLAS detector at the LHC during a dedicated pp collision run with special settings. In particular, the chosen setup safeguards that (with statistically rare exceptions) only one pp interaction occurs at the same time in the collision region inside the ATLAS detector; this prerequisite allows for a reliable measurement of the multiplicities and kinematic properties of primary charged particles that are produced per single pp interaction.

If particle detectors were 100 % efficient, such a measurement would be rather straightforward.[3] However, in practice the complex detectors and the measurements performed with them are subject to various sources of inefficiencies, which must be thoroughly accounted for. Furthermore, modern particle detectors such as ATLAS merely record abstract collections of measurement signals; they cannot distinguish a priori which of these signals originate from the primary interactions and which signals stem from other sources - such as particles produced by secondary interactions inside the detector, cosmic rays, long-lived particle decays, etc. All of these sources of background must also be understood and taken into account. This is commonly achieved by applying corrections for various detector effects and sources of background, while carefully considering their systematic uncertainties. Not all corrections and systematic uncertainties can be derived from the experimental data alone; in such cases computer simulations are required, which thus become an indispensable part of the measurement (Chapter 4).[4]

In order to successfully perform a computer simulation in high-energy physics, all components of the given problem must be well understood. This involves an accurate modelling of the configuration of physical components (i.e. detector geometry, materials, magnetic field configurations, and initial conditions of primary particles), the dynamics of the system (i.e. models of particle transport within a magnetic field, physics interactions between particles and matter), and the detector responses of sensitive elements. The underlying physics processes are common to all experiments, such that the models describing physics interactions can be validated independently. However, large systematic uncertainties can remain on the configuration side, e.g. due to insufficient modelling of the specific detector geometry of an experiment.

One of the most striking examples for the adverse effects of insufficient modelling of the detector configuration is given by the track reconstruction efficiency. This parameter describes the ability to reconstruct original particle trajectories from an abstract collection of measurement signals by using dedicated pattern recognition algorithms (Chapter 5). In an idealised (and physically unrealistic) scenario, where particles traverse matter along their initial trajectories without any disturbances due to material interactions, and where the physical locations of all measurement points are perfectly reflected by the computer models (referred to as "alignment"), such measurement signals can be easily and unambiguously translated into precisely known trajectories by the pattern recognition algorithms. This would lead to a nearly perfect and uniform track reconstruction efficiency across the entire kinematic range of charged particles. In practice the track reconstruction task becomes more complex, mainly because particles can

[3] Even so, the enormous intellectual, practical and technical expertise and coordination required to construct, maintain and run the entire experimental setup, including the particle collider and detector, represents a highly complex challenge – which must be mastered in order to perform such a measurement in the first place. See Ref. [2] for an introduction into organisational principles within the ATLAS collaboration.

[4] For completeness, it should be emphasised that this is not the only reason for employing computer simulations in particle physics: indeed the whole planning of an experiment such as ATLAS requires enormous simulation efforts e.g. to predict measurement performances, in order to choose a detector layout which optimises the expected performance for the greatest amount of "critical" studies, while still satisfying numerous physical, technological, logistical, financial and other constraints. This process requires the weighing of preferences from various sides and an informed decision-making which converges all the way from the physics-level to the management-level, adding further weight to the number of challenges presented in this introductory chapter.

indeed deviate from their current path when they interact with matter (Section 2.7); in the most drastic example, a hadronic particle (such as a pion, kaon or proton) can be abruptly halted following a hadronic interaction with a nucleus. In such a case only a truncated collection of measurement points between the primary collision vertex and the location of the hadronic interaction has been produced by that particle. If this process happens too early in the detector, the pattern recognition algorithms cannot reconstruct the initial particle trajectory, with the effect that this particle cannot be accounted for and is "lost" to the measurement. The resulting track reconstruction inefficiency is therefore directly related to the amount of material traversed by hadronic particles, as more material implies a higher probability for a particle to undergo such an interaction.

The correction factors to account for such track reconstruction inefficiencies can only be derived with reasonable precision from computer simulations.[5] This means that the accuracy of the entire correction procedure, which is applied to the experimental measurements, is limited by the accuracy of the map (i.e. the description of the location and amount of material in the detector model which was used in that simulation) with respect to the territory (i.e. the actual implementation of the physical detector). Despite the enormous diligence that has been applied during the assembly of the ATLAS detector, including the documentation of the actual physical detector geometry and the materials used, even a small mismatch between description and actual implementation (e.g. a cable, a cooling pipe, or a thin copper plate) can lead to a significant systematic bias in the correction factors described above. Therefore it is absolutely essential for a successful charged-particle measurement to put reliable constraints on the uncertainty of the known material distribution. Since this is the largest source of systematic uncertainty in the entire measurement, it is also crucial to reduce this uncertainty by refining the detector models for simulation wherever possible.

As a consequence, much effort needs to be invested into the validation of the detector model and the evaluation of its material uncertainty. This can be done with dedicated data-driven material studies that are sensitive to certain particle-matter interactions (such as hadronic interactions and photon conversions) and allow to evaluate the actual material distribution in the ID, which can then be compared with results obtained from the simulation (Chapter 6). Other prerequisites for such investigations include: knowledge of the measurement efficiencies of tracking detector components; knowledge of the resolutions of reconstructed track parameters which can limit the resolutions of the material studies; knowledge of reconstruction efficiencies and resolutions of primary and secondary vertices (the points of origin of tracks) along with their systematic uncertainties; and an evaluation of the alignment of the detector components with respect to each other and to the nominal interaction region (to constrain biases of the reconstructed track parameters and differences in the momentum resolution between experimental data and simulation).

[5] Data-driven methods can also be employed to estimate the track reconstruction efficiency, but these suffer from experimental constraints including very small statistics, whereas it is relatively easy to obtain simulation samples with large enough statistics.

The main goal of this thesis is to perform a precise measurement of charged-particle distributions in high-energy proton–proton collisions at a centre-of-mass energy of $\sqrt{s} = 8$ TeV with the ATLAS detector at the LHC (Chapter 7). This study aims to exceed the precision of previously made measurements at lower energies, in order to make a valuable contribution to the understanding of the soft-QCD sector of SM physics. That goal was achieved by a twofold strategy: detailed ID performance studies and material studies were carried out in order to ensure that all of the above requirements were fulfilled; based upon knowledge from these studies, the charged-particle measurement was subsequently performed. The first set of tasks was carried out by several groups within the ATLAS collaboration, with significant contributions by the author; the second measurement was conducted by a small analysis team in which the author took the leading role, from the early stages of the analysis onwards through detailed studies and internal documentation until the recent ATLAS paper publication in Ref. [3].

Structure of this Thesis: This introduction describes the motivation and main challenges of the studies and measurements presented in this thesis within the context of fundamental high-energy particle physics research, with a more general audience in mind.

Theoretical foundations of SM physics and strong interactions in the soft-QCD sector, phenomenological models to describe them, and an overview of particle-matter interactions are laid out in Chapter 2.

The LHC and the ATLAS experiment, along with the special experimental settings that were applied to collect the data that was used for the main studies presented in this thesis, are introduced in Chapter 3.

The ATLAS simulation infrastructure, detector geometry models, and implementation of physics interactions in the simulation are given in Chapter 4; this includes a brief survey of validation studies performed by the author.

In Chapter 5 the principles of charged-particle reconstruction are explained and the performance of the inner tracker is evaluated in terms of measurement efficiencies, track and vertex reconstruction, and detector alignment.

Chapter 6 provides a collection of ID material validation studies, using the SCT extension efficiency, hadronic interaction vertices, and photon conversion vertices; these studies culminated in significant improvements of the simulation geometry along with a drastic reduction of material uncertainties of the ID.

The main work of this thesis is the measurement of primary charged-particle distributions at $\sqrt{s} = 8$ TeV centre-of-mass energy, which is extensively described in Chapter 7 with discussions of the experimental setup and simulations, data selections, corrections for backgrounds and detector inefficiencies, systematic uncertainties, and comparisons of the main results with predictions from phenomenological models.

All of these studies and results are summarised in Chapter 8, and their impact on other ATLAS measurements as well as on the general understanding of soft-QCD physics is discussed.

Additional details on the charged-particle measurement, including comparisons of the results to predictions from further phenomenological models, are given in Appendix A. Finally, Appendix B provides auxiliary material concerning the ID reconstruction geometry and material validation studies, while Appendix C gives further details on the ID reconstruction performance.

Chapter 2

Theoretical Background

2.1 The Standard Model of Particle Physics

The Standard Model of Particle Physics (SM) is a theoretical framework that successfully describes our current understanding of the composition of matter in terms of known elementary particles and their measured properties, as well as three of the four fundamental physics interactions (electromagnetism, strong and weak nuclear force) acting upon and between them. Its predictions have been experimentally verified with high precision up to the TeV energy scale, including the discoveries of the Z^0 and $W^{\pm}$ bosons, gluons, the top quark, and the recent observation of the SM Higgs boson at the LHC.

The limitations of the SM are reflected in its lack to provide phenomenological descriptions of phenomena such as gravity and neutrino oscillations, and in its large number of free parameters such as fermion masses and coupling constants. Furthermore, the SM cannot predict some of the properties of fundamental particles, such as the Higgs boson mass, from first principles; these properties must therefore be obtained from experimental data. For these and other reasons, despite its success to describe many physical phenomena, the SM is currently regarded to be an incomplete theory. Several extensions and alternatives to the SM, such as Supersymmetry (SUSY) and string theories, have been proposed and investigated to improve our current understanding, but such explorations are not relevant to this thesis and shall therefore not be further discussed here.

The SM is a renormalisable gauge theory in which the strong, weak and electromagnetic interactions are derived from the single requirement of local gauge invariance. It can be described with the formalism of Quantum Field Theory (QFT). The gauge group of the SM is the tensor product of three abstract symmetry groups, $\mathrm{SU(3)_C} \times \mathrm{SU(2)_L} \times \mathrm{U(1)_Y}$. The first special unitary group describes the strong interaction between quarks and gluons, while the latter two groups describe the electroweak interaction which unifies the weak force and electromagnetism. Here the subscript C signifies the "colour charge", L indicates the "left-handed" chirality of weak interactions, and Y stands for "hypercharge" which distinguishes the third group from the $\mathrm{U(1)_{EM}}$ group of Quantum Electrodynamics (QED). The electroweak interaction contains $\mathrm{U(1)_{EM}}$ through electroweak mixing.

Fundamental point-like particles can be classified as *fermions* with half-integer spin and *bosons* with integer spin. Fermions are grouped in three generations and comprise *leptons* (electrons (e), muons (μ) and taus (τ) with their neutrino partners ν_e, ν_μ, ν_τ) and colour-charged *quarks* (up and down (u,d), charm and strange (c,s) and top and bottom (t,b)). They carry weak isospin and are therefore affected by the weak force. Charged weak transitions among quark flavours and charge parity (CP) violation in the quark sector are described by the complex unitary Cabibbo–Kobayashi–Maskawa (CKM) mixing matrix.

Fermions couple to gauge bosons which are the mediators of the fundamental forces. The electromagnetic force, described by the theory of Quantum Electrodynamics (QED), is mediated by the exchange of massless photons (γ) and acts upon all charged particles, including charged leptons and quarks. In weak interactions affecting all leptons and quarks, either neutral (Z^0) or charged ($W^\pm$) gauge bosons are exchanged. The masses of the W and Z bosons are generated by the Brout–Englert–Higgs (BEH) mechanism of spontaneous electroweak symmetry breaking. This mechanism also generates the fermion masses that are determined by Yukawa couplings.

The theory of strong interactions, which are mediated by massless and electrically neutral gluons, is called Quantum Chromodynamics (QCD) (see Section 2.1.1) and applies only to quarks and gluons themselves. Quarks naturally appear in bound states as *hadrons*, which are grouped into *baryons* (e.g. protons and neutrons) consisting of three quarks or antiquarks, and *mesons* (e.g. pions and kaons) consisting of quark-antiquark pairs. In the parton model, quarks and gluons are commonly referred to as *partons*. A complete listing of known elementary particles and their properties can be found in Ref. [4].

2.1.1 Quantum Chromodynamics (QCD)

QCD is a non-abelian gauge theory, based on the special unitary symmetry group $\mathrm{SU}(3)_C$, that describes the dynamics of strong interactions between quarks and gluons [5]. In this framework, the strong force is mediated by massless and electrically neutral gluons carrying a colour charge in the form of an octet combination of three basic colour charges, which are referred to as (anti-) *red*, *green* and *blue*. Due to their colour charge, the gluons are also able to interact with each other; this property of self-interaction gives rise to an effective "running" strong coupling constant g_s, usually expressed in the form $\alpha_\mathrm{s} = g_\mathrm{s}^2/4\pi$, which decreases with rising energy. At the reference scale of the Z boson mass, its world average has been recently evaluated as $\alpha_\mathrm{s}(M_Z^2) = 0.1181 \pm 0.0013$ [4, 6, 7]. This value has been obtained by a combination of a number of measurements, including hadronic to leptonic branching fractions of Z and τ decays at e^+e^- colliders, production cross-sections (e.g. $t\bar{t}$ or $b\bar{b}$) at hadron colliders, and distributions of global event shape variables in e^+e^- annihilation. A comparison between the theoretical prediction and latest experimental results of the running of α_s is given in Figure 2.1.

The coupling constant α_s and the quark masses m_q are considered as the fundamental QCD parameters, from which the energy scale of strong interactions, Λ_QCD, can be derived. The commonly used approximation of $\Lambda_\mathrm{QCD} \approx 210$ MeV, consistent with experiments, is close to

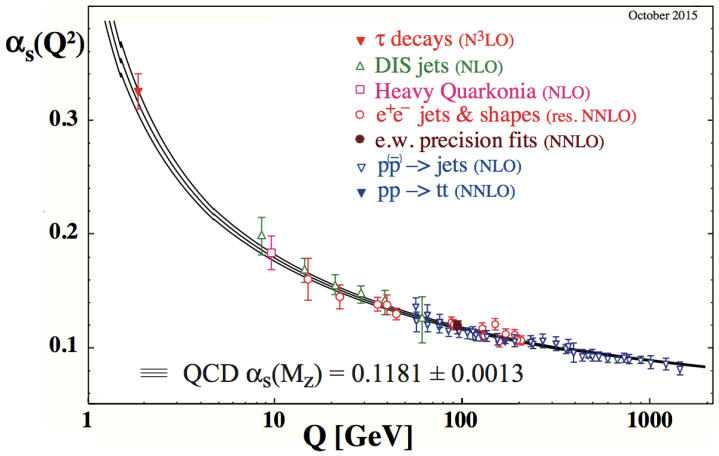


Figure 2.1: Comparison between the theoretical QCD prediction and latest experimental results of the running of α_s as parameterised by the momentum-transfer scale, Q. Plot taken from Ref. [4] (PDG, 2015 update).

the inverse of the hadron size, $1/R_0$, and to the mass scale of hadronic physics: for instance, the masses of pions and protons have been measured as $m_{\pi^0} \approx 135$ MeV and $m_p \approx 938$ MeV.

The framework of QCD can provide a theoretical explanation to several experimental observations, such as confinement, hadronisation and asymptotic freedom, which are related to the energy-scale dependence of α_s:

- *Confinement* describes the observation that particles with colour charge, i.e. quarks and gluons, are confined inside colour-neutral bound states called hadrons, and thus unable to propagate freely over macroscopic distances. This phenomenon can be related to the non-abelian type of the $SU(3)_C$ group.
- *Hadronisation* occurs when the linearly increasing potential of the colour field between two interacting partons at large distances, $V(r) \approx \lambda r$, gives rise to the creation of new partons (quark-antiquark pairs) and the formation of new colour-neutral hadrons.
- *Asymptotic freedom* describes the observation that at large momentum transfers the coupling α_s becomes small enough such that the closely confined partons can be regarded as freely moving inside hadrons. This property allows for the application of perturbation theory to describe high-energy hadron collisions in terms of parton interactions.

Perturbative calculations of QCD processes in orders of α_s are however limited to *hard scattering* events, which involve large momentum transfers between interacting partons. The Lorentz-invariant four-momentum transfer squared between partons can be expressed by the Mandelstam variable $\hat{t} = (p_{\text{in}} - p_{\text{out}})^2 = q^2 = -Q^2$. In the corresponding regime of $Q^2 \gg \Lambda^2_{\text{QCD}}$, the parameter $\alpha_s(Q^2)$ can be approximated at next-to-leading order (NLO) as:

$$\alpha_{\mathrm{s}}(Q^2)|_{Q^2 \gg \Lambda^2_{\mathrm{QCD}}} = \frac{1}{\beta_0 \ln(Q^2/\Lambda^2_{\mathrm{QCD}})} \left(1 - \frac{\beta_1 \ln[\ln(Q^2/\Lambda^2_{\mathrm{QCD}})]}{\beta_0^2 \ln(Q^2/\Lambda^2_{\mathrm{QCD}})}\right) \tag{2.1}$$

For the gauge group SU(3), the β coefficients in Eq. 2.1 are given by $\beta_0 = (33 - 2n_f)/(12\pi)$ and $\beta_1 = (153 - 19n_f)/(24\pi^2)$, where n_f represents the number of active quark flavours (with $m_q < Q$); any value of $n_f \leq 16$ therefore leads to asymptotic freedom.[1]

At long distances corresponding to small momentum transfers Q^2, the coupling strength described by α_{s} increases (leading to confinement), and in the perturbative approximation it would diverge at $Q^2 = \Lambda^2_{\mathrm{QCD}}$. However, the applicability of perturbation theory to calculate the interaction processes breaks down already below the scale of $Q^2 < 1$ GeV^2, where higher-order terms of α_{s} can no longer be neglected. In order to describe the *soft interaction* processes in this low-momentum regime, phenomenological models are introduced. Details are given in Section 2.3.

2.2 Proton-Proton Collisions

A deeper understanding of SM physics and QCD interactions can be obtained by the experimental study of proton-proton (pp) collisions. Protons are colour-neutral objects with a complex substructure, comprising three valence quarks (u,u,d) as well as gluons which can split into quark-antiquark ($q\bar{q}$) pairs termed sea quarks. These point-like substructures in the proton have been discovered via the observation of scaling behaviour of the structure functions for deep inelastic scattering (DIS) processes in ep interactions. In this parton model, each parton is assumed to carry a fraction of the hadron four-momentum, and can thus be described by parton distribution functions (PDFs) that can be determined from measurements in a wide range of DIS processes. These PDFs describe the probability densities $f_{a/A}$ for partons inside a hadron A to carry four-momentum fractions x_a which depend on the energy scale Q^2 of the given process. The probability densities can thus be expressed as $f_{a/A}(x_a, Q^2)$. The PDFs for valence quarks, sea quarks and gluons have been extracted e.g. from ep collision measurements at the HERA collider [8]; examples are shown in Figure 2.2 at two different energy scales. While their shapes have non-perturbative origin, the approximately logarithmic Q^2 dependence of the PDFs can be derived from perturbative QCD calculations. These logarithmic violations of the QCD scaling behaviour can also be understood in terms of an increasing number of additional partons (sea quarks and gluons) that are simultaneously probed at higher energy scales.

The complexity of the proton substructure allows for a large number of possible processes $a+b \to X$ in high energy pp collisions, where a and b represent interacting partons of the colliding protons, while the reaction products X can include quarks, leptons and bosons (e.g. Drell–Yan

[1] At the scale of the Z boson mass, i.e. $Q \approx M_Z$, using $n_f = 5$ the approximation in Eq. 2.1 yields a value of $\Lambda_{\mathrm{QCD}} \approx 230$ MeV. With higher-order corrections, one obtains a result closer to $\Lambda_{\mathrm{QCD}} \approx 210$ MeV.

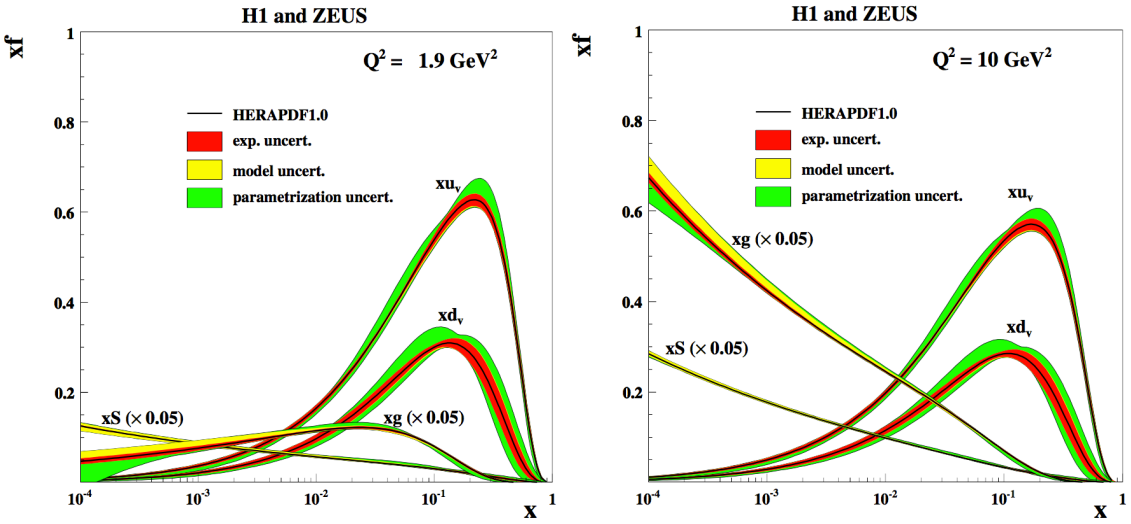


Figure 2.2: Comparison of PDFs of valence quarks (xu_v and xd_v), total sea-quark distribution (xS) and total gluon distribution (xg) at NLO, extracted from ep collision data measured by HERA experiments at (left) $Q^2 = 1.9$ GeV^2 and (right) $Q^2 = 10$ GeV^2. The shaded areas show the experimental uncertainties, as well as model and parameterisation uncertainties of the fit results. Plots taken from Ref. [9].

processes, $q\bar{q} \to l^+l^-$). These interactions are typically divided into *hard* and *soft* processes, depending on the characteristic momentum transfer Q^2 between the two interacting partons.

Hard scattering processes are characterised by a large momentum transfer $Q^2 \gg 1$ GeV^2, well above the scale of strong interactions as given by Λ_{QCD}. Under the assumption that such a process can be factorised into parts of non-perturbative and perturbative origin, its theoretical differential cross-section can be approximated by the *factorisation theorem* [10, 11]:

$$\sigma_{AB} = \int \mathrm{d}x_a \mathrm{d}x_b \, f_{a/A}(x_a, Q^2) f_{b/B}(x_b, Q^2) \, \hat{\sigma}_{ab\to X}. \tag{2.2}$$

Here the non-perturbative part is approximated by the two PDFs $f_{a/A}$ and $f_{b/B}$, whereas the cross-section of the hard interaction, $\sigma_{\mathrm{hard}} = \hat{\sigma}$, can be calculated precisely using perturbative QCD in orders of $\alpha_{\mathrm{s}}(Q^2)$. The factorisation theorem provides the main theoretical basis for experimental high-energy particle physics with large momentum transfer.

Hard QCD interactions comprise $qq \to qq$, $qg \to qg$ and $gg \to gg$ scattering processes, all of which involve gluon exchange at leading order (LO). These interactions lead to the di-jet final states of $2 \to 2$ processes. For a meaningful calculation, a transverse-momentum (p_{T}) cut-off has to be specified to avoid the divergence of the fixed-order cross-section if $p_{\mathrm{T}} \to 0$.

2.3 Soft QCD

At smaller momentum transfer between the interacting partons, below $Q^2 \approx 1\ \mathrm{GeV}^2$, the strong coupling $\alpha_s(Q^2)$ increases and perturbative calculations of QCD can no longer be used to describe strong interactions. As a consequence, the differential cross-section can no longer be approximated by the factorisation theorem in the low-energy regime. These low-p_T scattering processes are referred to as *soft processes*. They can be described quantitatively by phenomenological models with tuneable parameters.

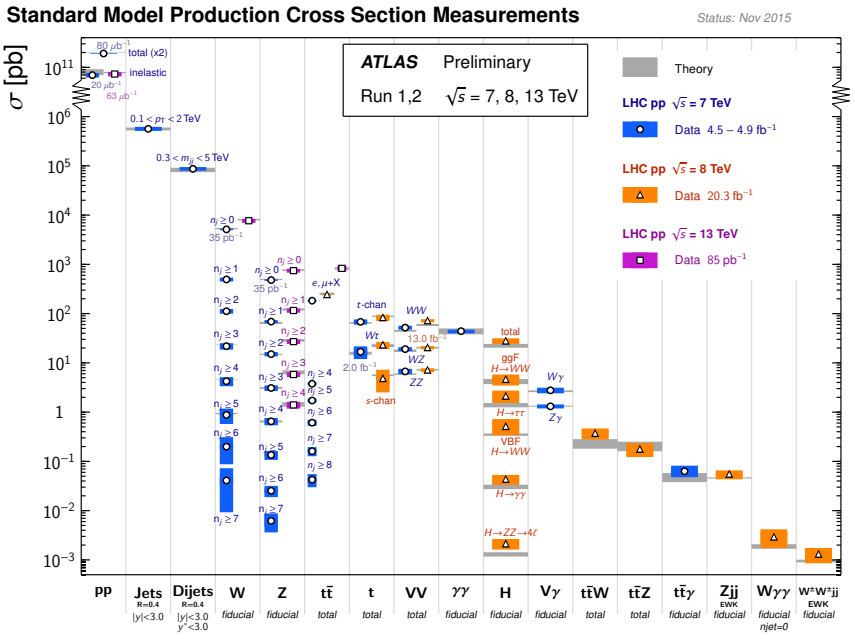


Figure 2.3: Summary of various fiducial SM production cross-section measurements, corrected for leptonic branching fractions, compared to theoretical predictions that were calculated at NLO or higher. The luminosity given in the legend was used for each measurement, except where another specific value is indicated. Plot taken from Ref. [12].

Soft interactions are characterised by large cross-sections with a logarithmic dependence on the centre-of-mass collision energy. The total cross-section σ_{tot} represents the sum of all possible processes in a pp collision and consists of single contributions from elastic and inelastic interactions with mostly small momentum transfers; hard processes also contribute to σ_{tot} with a rate that is several orders of magnitude lower, as shown in Figure 2.3 for various production channels including jet and di-jet events. Inelastic interactions can be subdivided into diffractive and non-diffractive processes, which are described in Section 2.4:

$$\sigma_{\mathrm{tot}} = \sigma_{\mathrm{elastic}} + \sigma_{\mathrm{inelastic}} = \sigma_{\mathrm{elastic}} + \sigma_{\mathrm{diffractive}} + \sigma_{\mathrm{non-diffractive}} \tag{2.3}$$

In order to understand diffractive processes with small momentum transfer, one first needs to consider the theories of the Regge formalism and Pomeron exchange. A description of non-diffractive processes is provided by the theory of Multiple Parton Interactions (MPI).

2.3.1 Regge Theory and Pomerons

Regge theory provides solutions for the non-relativistic scattering equation of a spherically-symmetric potential using non-integer, complex-valued angular momenta [13]. These scattering amplitudes can be parameterised with Regge trajectories, which are associated with families of virtual particles that share the same internal quantum numbers. They contribute to the total scattering cross-section, which was found to rise logarithmically with the centre-of-mass energy at a scale above 10–20 GeV for pp as well as $p\bar{p}$ interactions. This energy dependence was accounted for by a special Regge trajectory with the internal quantum numbers of the vacuum, called the Pomeron trajectory [14, 15]. The Pomeron is understood to be the effective sum of all virtual particles along this trajectory. Under the Pomeron hypothesis, the pp and $p\bar{p}$ scattering cross-sections are expected to converge asymptotically at high energies; this prediction is consistent with available experimental results.

Until today, no known particles have been uniquely associated with the Pomeron, which is hypothesised to correspond to charge- and colour-neutral multi-gluon objects called glueballs. It can be described at leading order by a gluon–gluon exchange process, and corrected at higher orders by a ladder of gluon interactions. Pomerons are thought to contribute to diffractive interactions in which the incident particles exchange one or two colourless objects. This Pomeron exchange can be described with PDFs that are based on measurements of diffractive ep interactions, and which can be used to predict multi-particle production in diffractive hadron–hadron interactions.

2.3.2 Multiple Parton Interactions

The composite nature of hadrons gives rise to a high probability of multiple simultaneous scatterings of partons within a single hadron–hadron interaction. Such processes are known as Multiple Parton Interactions (MPI) [16]. As more additional partons (sea quarks and gluons) can be simultaneously probed at higher energy scales, the MPI cross-section tends to increase with the centre-of-mass collision energy. The most common MPI are $2 \to 2$ processes (that can give rise to di-jet events), for which the cross-section is dominated by $\hat{t}$-channel gluon exchange. To avoid the divergence at low momentum transfers, this cross-section can be regularised in phenomenological models by an energy-scale-dependent cut-off parameter, $p_{\perp 0}^{\mathrm{ref}}$.[2] This tuneable parameter steers the transition between the hard scattering regime, which is calculated using perturbative QCD, and the phenomenologically described soft interactions. It is usually set for

[2] If the calculated interaction cross-section exceeds the total cross-section, this can be interpreted as multiple parton scatterings (MPI) occurring in the same event; conversely, the interaction cross-section divided by the total cross-section yields the average number of parton scatterings per event.

a nominal reference centre-of-mass energy of 1800 GeV and is being used in combination with a tuneable exponent to describe the evolution of the p_{T} cutoff with the centre-of-mass energy.

MPI also contribute to the production cross-sections and final states of multi-jet events. At smaller momentum transfer scales, MPI are essential for the successful phenomenological description of the Underlying Event activity, i.e. the background of parton interactions which occur together with a hard scattering process, and particularly of Minimum Bias events which comprise primarily soft interactions (see Section 2.6). While the MPI lead to an increase in the final-state multiplicity in such events, colour reconnection mechanisms limit the rise of the multiplicity. These phenomena are assumed to result from physical overlaps of colour strings, such that by reconnecting the strings the potential energy of the system is reduced; here low-p_{T} systems correspond to longer strings which are thus more likely to reconnect.

The parton density inside hadrons can be assumed to follow a single or double Gaussian profile. The probabilities of MPI as well as hard scattering processes then rise with a decreasing impact parameter of the hadron–hadron collision, i.e. with an increasing overlap between the interacting hadrons. The simultaneous interactions are therefore no longer assumed to be independent. Furthermore, it is expected that with a small probability the outgoing states from one parton–parton interaction can become the incoming states for another such process, termed rescattering.

The understanding of MPI in hadron–hadron collisions remains an active area of theoretical and experimental research, and represents one of the main challenges in the parameter tuning of phenomenological models, with high relevance for the understanding and simulation of Minimum Bias and Underlying Event physics.

2.4 Diffractive Interactions and Scattering Cross-Sections

The total pp cross-section can be described as the sum of elastic and inelastic cross-sections, see Eq. 2.3. It has been determined by LHC experiments, using the optical theorem and the measurement of the differential elastic cross-section, as between $\sigma_{\mathrm{tot}}(pp \to X) \approx 96\,\mathrm{mb}$ [17] (ATLAS + ALFA, utilising a luminosity-dependent method) and $\sigma_{\mathrm{tot}}(pp \to X) \approx 103\,\mathrm{mb}$ [18,19] (CMS + TOTEM, employing a luminosity-independent method) at the centre-of-mass energy $\sqrt{s} = 8$ TeV, with a contribution from elastic processes of between $\sigma_{\mathrm{el}}(pp \to pp) \approx 24\,\mathrm{mb}$ (ATLAS) and $\sigma_{\mathrm{el}}(pp \to pp) \approx 27\,\mathrm{mb}$ (CMS).

Inelastic events comprise non-diffractive reactions, characterised by gluon (colour-octet) exchange between the interacting partons, and diffractive dissociation, characterised by Pomeron (colour-singlet) exchange. Diffractive events can be further classified as single-, double- or central-diffractive reactions. For these categories, the characteristic event topologies are illustrated in Figure 2.4. The total inelastic cross-section can be calculated by subtracting the elastic contribution from the total pp cross-section, which yields values between $\sigma_{\mathrm{inel}} \approx 72\,\mathrm{mb}$ (ATLAS) and $\sigma_{\mathrm{inel}} \approx 76\,\mathrm{mb}$ (CMS).

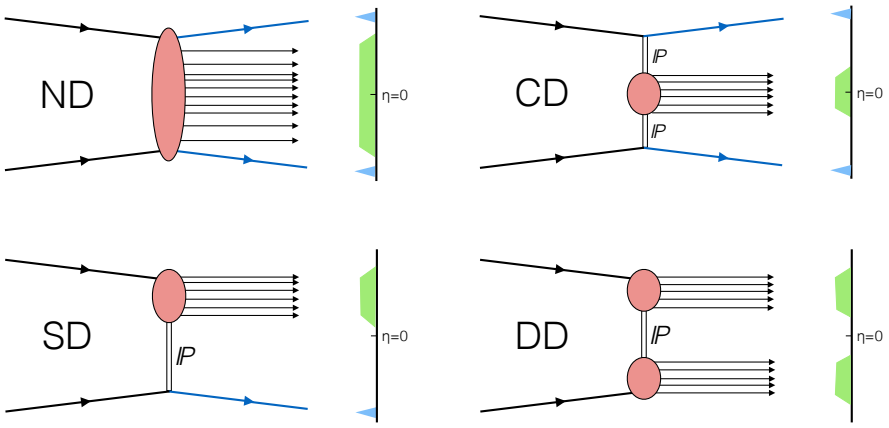


Figure 2.4: Illustration of the characteristic event topologies in non-diffractive (ND), central-diffractive (CD), single-diffractive (SD) and double-diffractive (DD) pp interactions. In non-diffractive processes the final-state hadrons are almost evenly distributed across a wide rapidity interval and have larger average multiplicities. Diffractive scattering involves a single or double Pomeron exchange and exhibits characteristic rapidity gaps in the final-state hadron distributions.

2.4.1 Elastic Scattering

In elastic scattering both interacting particles remain intact, without any exchange of internal quantum numbers, and no additional particles are produced. Elastic scattering is dominated by processes with small momentum transfer Q^2. The outgoing particles can be measured in the forward directions with dedicated detectors located at a larger distance from the interaction region, such as the ALFA roman-pot sub-detectors of the ATLAS experiment at the LHC.

Elastic interactions occur with an expected fraction of 25–27 % of all pp interactions at $\sqrt{s} = 8$ TeV, based on calculations using the optical theorem.

2.4.2 Non-diffractive Interactions

Non-diffractive (ND) interactions are characterised by a gluon (colour-octet) exchange between the interacting partons. The exchange of quantum numbers (colour charge) between the two interacting systems causes the formation of stretching colour fields between the outgoing partons, which lead to hadronisation in the central region. The typically larger number of final-state hadrons, which increases further with rising centre-of-mass energy, can span the full rapidity range, while the transverse momentum spectrum of these particles rapidly decreases from lower to higher momenta.

ND interactions are the dominant process among inelastic interactions at $\sqrt{s} = 8$ TeV, with an expected fraction of 50–55 % of all pp interactions.[3]

[3] For inelastic processes, the given fractions and their uncertainties are based on phenomenological parameterisations of the data using Regge theory, which correspond to the MC models and tunes that are used in this work.

2.4.3 Diffractive Scattering

Diffractive processes [20, 21] are characterised by a small momentum transfer between the two interacting particles and by the absence of exchange of internal quantum numbers. These interactions can be phenomenologically described by an exchange of a Pomeron (colour-singlet) that interacts with gluons and sea quarks of one or both protons, resulting in a system of particles called "diffractive system". In such processes one (single-diffractive, SD) or both excited protons (double-diffractive, DD) dissociate, and the final-state hadronic system retains the same internal quantum numbers as the original protons. As a consequence of this colour-neutral Pomeron exchange, the final-state hadrons tend to propagate in the forward direction, and none or few outgoing particles in the central region are found. This characteristic signature, called *rapidity gap*, can be exploited in specific "forward physics" studies. If two Pomerons are exchanged in one *pp* interaction, the higher-order process is called central-diffractive dissociation (CD) or double Pomeron exchange. In this rare process both colliding protons remain intact, but the Pomeron exchange gives rise to an additional particle shower in the central region, accompanied by two large rapidity gaps on either side.

While Pomeron exchange can successfully describe diffractive processes at high centre-of-mass energies, a second type of interaction called Reggeon exchange, historically formulated earlier, is found to describe the data better at lower collision energies. Reggeons can be understood to interact with the valence quarks of colliding protons. The total *pp* cross-section can thus be described more accurately by a Reggeon–Pomeron model based on a summation of both trajectories [22, 23]. This approach is implemented in some phenomenological QCD models, specifically the Pythia and Herwig++ models that are described in Section 2.5.1.

SD and DD processes occur with expected fractions of 11–15 % and 7–10 % of all *pp* interactions at $\sqrt{s} = 8$ TeV, respectively. Experiments show that with increasing centre-of-mass energy the rise of the DD cross-section exceeds that of SD processes, and the relative contribution of both processes to the total cross-section increases, while it remains below 1 % in the case of CD processes. In this thesis, the CD event category will be disregarded.

2.5 Monte Carlo Event Generators

Monte Carlo (MC) methods can be employed as numerical techniques for the probabilistic simulation of a wide range of phenomena, including complex processes in high-energy physics, such as particle collisions, particle-matter interactions and detector responses, which are not easily amenable to analytical calculation. They are based on computational algorithms that use (pseudo-) random numbers, such as to sample random variables from given probability density functions, or to select among the possible discrete probabilistic outcomes of a given process. MC methods are also extensively applied in particle physics to generate large samples of possible final states of particles that arise from individual collision events, by employing programs called MC event generators. These generators typically calculate hard scattering interactions based on

perturbative QCD, while soft interactions are described by phenomenological models with large numbers of free parameters. They can be configured to utilise different models and settings e.g. for diffraction, parton showering, hadronisation, colour reconnection, MPI energy evolution, as well as PDFs and matrix elements. The degree of compatibility between such models (and their hypothesised theoretical underpinnings) and experimental results can be assessed via comparisons of the MC predictions to experimental data, e.g. by performing minimum-bias and underlying-event measurements (see Section 2.6). These measurements and comparisons are utilised to optimise (tune) the sets of model parameters (also called *tunes*) of the generators.

Several MC event generator models of inclusive hadron–hadron interactions are employed in the Minimum Bias analysis presented in this thesis: these are the Pythia 8, Pythia 6, Epos, Qgsjet-II and Herwig models, all of which are briefly introduced in the following sections.[4]

2.5.1 Pythia 8

The general-purpose MC event generators Pythia 8 [24] (written in C++) and Pythia 6 [25] (written in Fortran) implement several of the phenomenological models that are used to describe inelastic hadron–hadron interactions. These models employ a variety of parameters that have been tuned to minimum-bias and underlying-event data from pp and $p\bar{p}$ collisions at centre-of-mass energies between 200 GeV and 7 TeV. In the Pythia generators, the simulation of complex hadronic high-energy physics processes is separated into ND, SD and DD interactions, which together comprise the total inelastic cross-section. The ND event category is broken down into the following tuneable subprocesses:

- *Hard Scattering*: The hard-scattering cross-section depends on the chosen PDFs, according to the factorisation theorem (Eq. 2.2). The particular physics process that gives rise to a hard scattering event is randomly sampled according to the current knowledge of production cross-sections.
- *Multiple Parton Interactions* (MPI): Between all initial partons, as well as partons that are created by other processes, simultaneously occurring hard or semi-hard $2 \to 2$ interactions within a single hadron–hadron collision can be simulated. In Pythia, these are assumed to be p_{T}-ordered. While MPI contribute significantly to multiplicity fluctuations, the multiplicity rise is limited by colour reconnection phenomena in the final state, as explained below. The additional soft and semi-hard parton interactions that can occur alongside the hard interaction are called the Underlying Event (UE).
- *Initial State Radiation* (ISR): Partons of both colliding hadrons can radiate photons and gluons prior to the hard scattering process.
- *Final State Radiation* (FSR): Final-state partons can radiate additional photons and gluons and thus create parton showers which can even give rise to further MPI. This process is

[4] MC generators that are mostly optimised for other specific studies, such as Phojet, Sybill, Sherpa and others, are not considered in this thesis.

stopped at a tuneable cut-off corresponding to a mass scale of typically 1 GeV. Different generator models use various ordering variables. In the PYTHIA models that were used for the studies presented in this thesis, the initial- and final-state radiations are taken to be p_{T}-ordered.

- *Beam Remnants* (BR): These comprise all partons that are not involved in the above parton interactions. They carry most of the total energy of the initial particles, and compensate the colour of the interacting partons.
- *Colour Reconnection* (CR): The beam remnants and other coloured objects are assumed to be connected by colour strings, which can reconnect with each other when physically overlapping in space-time, such that the total string length decreases and the potential energy of the system is reduced. This phenomenon affects the average transverse momentum of final-state particles.
- *Hadronisation*: Due to confinement the single partons must eventually form colour-neutral bound final states. This hadronisation process can thus far only be modelled phenomenologically. For instance, in the Lund string fragmentation model [26] that is used by both PYTHIA versions, a colour string is created in the field between a pair of colour-anticolour charged partons, with a potential energy that depends linearly on the separation distance (0.9 GeV/fm).[5] Vacuum fluctuations inside the string field can absorb its energy by giving rise to a new quark-antiquark pair, which effectively cuts the original string into two shorter strings between the available colour charges while reducing the total energy of the string system. This mechanism repeats until the potential energy falls below a threshold, resulting in the formation of pairs of bound mesons. The production of baryonic final states can be understood by a similar process involving diquark-antidiquark pairs.
- *Particle decay*: The decay of unstable particles is simulated according to known particle properties, i.e. masses, lifetimes and decay branching ratios. This includes resonance decays which can be simulated along with the hard process.

Compared with PYTHIA 6, PYTHIA 8 interleaves the MPI not only with ISR, but also FSR processes into one common p_{T}-ordered sequence. For diffraction, both PYTHIA generators use the same model to calculate cross-sections, diffractive masses and momentum transfers. Both versions implement diffractive processes based on the Reggeon–Pomeron and the Lund string fragmentation model, employing by default the Schuler–Sjöstrand parameterisation [27]. The probabilities for a Pomeron to couple to the diffracted proton (Pomeron flux) can be separated into a Pomeron–gluon and Pomeron–valence quark contribution, with a tuneable ratio between them. Here PYTHIA 8 employs a mass-dependent implementation, which favours the Pomeron–gluon coupling at higher diffractive masses. Furthermore, the PYTHIA 8 version differs from PYTHIA 6 by its modelling of particle production in high-mass diffractive systems above

[5] By contrast, the HERWIG model uses a hadronisation scheme based on colourless clusters.

$M > 10$ GeV, in that it employs a full perturbative description including diffractive PDFs, MPI, ISR and FSR [20].

Thanks to the improvements with respect to PYTHIA 6 in the diffractive simulation and MPI modelling, PYTHIA 8 predictions were found to be in better agreement with the data, particularly when low-p_{T} final-state particles are included. Therefore PYTHIA 8 has become the standard generator used in ATLAS for the modelling of all simulated pile-up interactions. In the main studies presented in this thesis, the versions PYTHIA 8.185 and 8.186 were used with the A2 [28, 29] and MONASH [30] tunes, respectively.

The MONASH tune, made by authors rather than with a fully automated procedure, is based on a re-analysis of fragmentation-sensitive e^+e^- data, as well as minimum-bias and underlying-event data from hadron collisions, utilising the NNPDF2.3 LO PDF-set for $\alpha_{\mathrm{s}}(M_Z) = 0.130$. The same value is also used for hard-scattering matrix elements, including those for MPI, while an effective value of $\alpha_{\mathrm{s}}(M_Z) = 0.1365$ is applied for both ISR and FSR (in contrast to the default values of 0.137 for ISR and 0.1383 for FSR). For the infrared shower cutoff, a value of $p_{\mathrm{T}}^{\mathrm{min}} = 0.5$ close to Λ_{QCD} is used. Furthermore, an interpolation between an exponential and Gaussian matter overlap is chosen (Option=3).

The A2 tune that was used in this thesis is based on minimum-bias and underlying-event data from hadron collisions up to $\sqrt{s} = 7$ TeV, and utilises the MSTW2008 LO [31] PDF-set. The tune was made with an automated parameter adjustment, using the PROFESSOR tool. It employs the following MPI parameters:

- `MultipleInteractions:pT0Ref` (the p_{T} cutoff used to avoid soft divergences at a reference energy of $\sqrt{s} = 1800$ GeV), equivalent to PARP(82),
- `MultipleInteractions:ecmPow` (the exponent used in its energy evolution to other centre-of-mass energies), equivalent to PARP(90),
- `MultipleInteractions:a1` (which affects the width of a Gaussian that is used for the matter overlap (Option=4); the width depends on x like $1 + \mathrm{a1} \cdot \ln(1/x)$), and
- `BeamRemnants:reconnectRange` (which controls the probabilities of colour reconnection for high-p_{T} and general colour strings), equivalent to PARP(77) and PARP(78).

A comparison of the MPI tuning parameters of the two main PYTHIA 8 minimum bias (A2) tunes is summarised in Table 2.1. For PYTHIA 8 A2, the tune which utilises the MSTW2008 LO PDF produced a better description of minimum-bias observables at $\sqrt{s} = 0.9$ and 7 TeV than the CTEQ6 L1 tune. It was therefore used as the baseline tune for the Minimum Bias analysis at $\sqrt{s} = 8$ TeV presented in this thesis.

2.5.2 PYTHIA 6

The PYTHIA 6 AMBT2B tune with the CTEQ6 L1 PDF has been performed on published ATLAS minimum-bias data from the 2010 data-taking period at $\sqrt{s} = 0.9$ and 7 TeV, including

Pythia 8 tune	PDF	pT0Ref	ecmPow	a1**	b***	reconnectRange
A2	CTEQ6 L1	2.18	0.22	0.06		1.55
A2	MSTW2008 LO	1.90	0.30	0.03		2.28
Monash	NNPDF2.3 LO	2.28*	0.215*		1.85	1.8

Table 2.1: Tuned MPI parameters for the Pythia 8 A2 (minimum bias) and Monash tunes. (*) = for this tune the reference energy for the extrapolation (ecmRef) was chosen to be 7 TeV instead of the default 1.8 TeV. (**) = parameter for Gaussian matter overlap (Option=4). (***) = parameter for interpolation between exponential and Gaussian matter overlap (Option=3).

a complete modelling of hadronisation, final-state showering and fragmentation, initial-state parton cascades, and MPI [32]. It is included in the present work mainly as an alternative MC tune for cross-checks during the Minimum Bias analysis. The Pythia 6 version 6.425 was used to produce the MC samples for this analysis.

The MPI parameters used in this tuning are:

- PARP(82) which is equivalent to `MultipleInteractions:pT0Ref`,
- PARP(90) which is equivalent to `MultipleInteractions:ecmPow`,
- PARP(83) and PARP(84) which are strongly correlated parameters for the double-Gaussian modelling of the hadronic matter distribution,
- PARP(77) and PARP(78) which are equivalent to `BeamRemnants:reconnectRange`.

A Pythia 6 AMBT2B tune with the MSTW2008 LO PDF was also performed, however it was found that the tune utilising the CTEQ6 L1 PDF produced a better description of minimum-bias observables at $\sqrt{s} = 0.9$ and 7 TeV. For both tunes, the PARP(83) parameter was fixed to the value of 0.356 from the AMBT1 tuning.

Pythia 6 tune	PDF	PARP(77)	PARP(78)	PARP(82)	PARP(84)	PARP(90)
AMBT2B	CTEQ6 L1	0.357	0.235	2.34	0.605	0.246
	MSTW2008 LO	0.533	0.263	2.06	0.602	0.284
Innsbruck2013	CTEQ6 L1	0.90*	0.516	2.76**	***	0.219**
	MSTW2008 LO	0.90*	0.550	3.02**	***	0.277**

Table 2.2: Tuned MPI parameters for the Pythia 6 AMBT2B (minimum bias) and Innsbruck2013 tunes. (*) = a fixed value was used for this parameter. (**) = for this tune the reference energy for the extrapolation (PARP(89)) was chosen to be 7 TeV instead of the default 1.8 TeV. (***) = here only a single parameter (PARP(83)) was used for the exponential and Gaussian modelling of the hadronic matter distribution.

An additional Pythia 6 Innsbruck2013 tune [33–35] was used with the CTEQ6 L1 PDF for generator-level comparisons of final distributions of the Minimum Bias analysis. This tune

was performed on published ATLAS and CDF measurements of minimum-bias and underlying-event data up to $\sqrt{s} = 7$ TeV. The parameter values given here are from a re-tuning using the latest Pythia version 6.428, instead of version 6.425.

In the Innsbruck2013 tunes utilising the PDF sets CTEQ6 L1 and MSTW2008 LO, the tuned Λ_{QCD} values for hard scattering (PARP(1)) were found to be 0.165 and 0.226, whereas for the tuned Λ_{QCD} parameter for FSR (PARP(72)) values of 0.389 and 0.580 were obtained, respectively [35]. For the Λ_{QCD} parameter for ISR (PARP(61)) a fixed value of 0.19 was chosen. The beam remnant parameters PARP(79) and PARP(80) were fixed at best fit values of 1.5 and 0.06 in the final tunes. For the exponential and Gaussian modelling of the hadronic matter distribution, only a single parameter (PARP(83)) was used with tuned values of 1.82 and 1.70 for the two PDF sets, respectively.

A comparison of the MPI tuning parameters of the two main Pythia 6 AMBT2B (minimum bias) and Innsbruck2013 tunes, employing the PDF sets CTEQ6 L1 and MSTW2008 LO, is summarised in Table 2.2.

2.5.3 Epos

The hadronic interaction model Epos[6] is based on the hydrodynamic evolution of ultra-relativistic particle collisions, starting from initial flux-tube structures [36]. It has been used primarily for cosmic-ray air shower simulations [37] and heavy-ion collision experiments [38] and can also be used to simulate *pp* interactions. Epos implements an effective QCD-inspired field-theory approach called "parton-based Gribov–Regge theory" [39] that simultaneously describes hard and soft scattering processes. Unlike the Pythia generators, Epos does not utilise the standard PDFs for its calculations.

Epos uses a number of parameters determining soft Pomeron properties, perturbative QCD cutoffs, string dynamics, screening and saturation effects, and projectile and target remnant properties [40]. These parameters have been tuned against available e^+e^-, *pp* and proton/deuteron–nucleus scattering measurements at a wide range of collision energies.

Each individual particle interaction in Epos produces a projectile and target remnant, which carry most of the collision energy, and at least two elementary colour flux tubes (strings) which subsequently break into hadrons. A single scattering amounts to the exchange of a complete parton ladder, including ISR. In high-multiplicity events, many partons ladders participate in the exchange, and a large number of flux tubes are produced, similar to heavy-ion collisions. The Epos model computes the energy density corresponding to these flux tubes, assumes thermalisation, and then performs a hydrodynamic expansion based on these new initial conditions. Recently published results from ALICE support the validity of the hydrodynamic expansion model [41].

The Epos version 3.1 was used with the Epos LHC tune [42] to produce the MC samples for the Minimum Bias analysis presented in this thesis. This tune contains a parameterised

[6] Epos stands for *"Energy conserving quantum mechanical approach, based on Partons, parton ladders, strings, Off-shell remnants, and Splitting of parton ladders"*.

approximation of the hydrodynamic evolution, derived from tuning to previous LHC minimum-bias data. The modelling of radial flow is optimised for the small-volume, high-density topologies that are found in *pp* interactions.

2.5.4 QGSJET-II

The QGSJET-II [43] model is based on a Reggeon field theory framework and simulates soft and semi-hard parton interactions with the "semi-hard Pomeron" approach. Like the EPOS model, QGSJET-II was developed for the simulation of cosmic rays, and does not utilise the standard PDFs in its calculations. This model also differs from the previously described generators by its lack of simulating colour coherence effects. In the work presented here, QGSJET-II is used with the default tune.

2.5.5 HERWIG++

HERWIG++ [44] is a general-purpose MC event generator in which inclusive high-energy hadron-hadron interactions are simulated with particular emphasis on an accurate modelling of ISR and FSR. It employs an eikonal model of independent multi-parton scatterings, a colourless-clusters hadronisation scheme and a colour reconnection mechanism, and allows for a choice between angular-ordered and dipole shower models.

The version 2.7.1 was employed to generate MC samples with the UE-EE-5 [45] tune, performed on underlying-event data at 7 TeV as well as double parton scattering data, while using the CTEQ6 L1 [46] PDF-set.

2.6 Measurements of Charged-Particle Distributions

High-energy *pp* collisions at the LHC and other particle colliders are dominated by non-perturbative soft-QCD processes, which can (as of today) only be described by phenomenological models. These models use a variety of tuneable parameters which must in turn be constrained by measurements of observables over a wide range of energies. To this end, the properties and dynamics of soft-QCD interactions in the low-transverse-momentum regime can be studied experimentally by inclusive measurements of charged-particle distributions (called *minimum bias*) containing only very low rates of hard scattering processes, or via measurements of events that involve a hard interaction with large transverse momentum transfer as well as a soft component (called *underlying event*).[7]

Underlying-event studies distinguish between particles in the *toward*, *away* and *transverse* regions, defined by their azimuthal angle relative to the leading high-p_{T} jet which originated from the hard interaction. The soft particles of interest, which constitute the underlying event,

[7] While neutral particles are relevant for studies involving the measurement of jet energies and of observables regarding the complete event, as well as for physics searches such as $H \to \gamma\gamma$, they are not considered in minimum-bias and underlying-event studies.

are assumed to populate the transverse regions. A special trigger (see Section 3.2.5) is employed to record a sample of only such events in which a hard-scatter interaction was present.

Minimum-bias measurements employ an inclusive minimum-bias trigger (see Section 3.2.6) to record a data sample of the vast majority of collision events that primarily contain particles originating from soft interactions. As the production cross-section for soft-QCD processes is several orders of magnitude larger than for hard processes, the minimum-bias measurement requires a much smaller luminosity than other studies of rare physics processes and particles.

The minimum-bias analysis presented in this thesis (see Chapter 7) is performed with primary charged particles produced in pp collisions which were recorded by the ATLAS experiment [47] at the LHC [1] at a centre-of-mass energy of $\sqrt{s} = 8$ TeV. A special experimental setup (see Section 3.1.1) was used to obtain a data sample with a very high probability of containing only one pp interaction per recorded collision event. In this analysis, as well as in equivalent measurements performed by LHC experiments at centre-of-mass energies between $\sqrt{s} = 0.9$ and 13 TeV, four types of distributions are studied:

1. the charged-particle multiplicity (or number of charged particles) per event,
2. the charged-particle density as a function of pseudorapidity,
3. the charged-particle density as a function of transverse momentum, and
4. the average transverse momentum as a function of the charged-particle multiplicity per event.

In equivalent studies from the pre-LHC era, mainly the first two distributions were measured at centre-of-mass energies up to $\sqrt{s} = 1800$ GeV. The rate of particle production due to parton interactions is dependent on the collision energy, particularly through the MPI mechanism. The measured charged-particle multiplicities are therefore being found to increase with the centre-of-mass energy (at a scale above 10–20 GeV). This energy-dependent multiplicity evolution can be studied by a comparison of the density of centrally produced charged particles over a wide range of collision energies.

The transverse-momentum distribution shows a continuous transition from soft interactions in the low-p_{T} regime to hard processes and jet-like events in the high-p_{T} range of the spectrum. Between these domains the p_{T} spectrum is steeply falling by several orders of magnitude, and its shape in the MC simulation is sensitive to the tuning of MC generator models. While this measurement is usually limited by a low-momentum threshold, the spectrum can be extrapolated to $p_{\mathrm{T}} = 0$ in order to obtain a more inclusive spectrum of charged particles. The distribution of the average transverse momentum versus multiplicity is of particular interest for the tuning of MC generators, as their predictions of this functional dependence are highly sensitive to the modelling of colour reconnection.

Furthermore, the results of these studies have significant implications for the modelling of event backgrounds arising from low-p_{T} processes, which can affect the topologies of hard-scattering events containing an interaction with large momentum transfer, when particles

originating from more than one hadron–hadron interaction are present at the same time. A better understanding and modelling of soft-QCD processes can thus also contribute to the reduction of systematic uncertainties in measurements based on such high-p_{T} phenomena.

The measured distributions are obtained from reconstructed particle trajectories (tracks), corrected for detector effects including measurement inefficiencies and background contaminations, and compared to MC predictions from various phenomenological models, using simulation samples obtained with the MC event generators listed in Section 2.5. In order to apply appropriate corrections for the inefficiencies and backgrounds, a detailed understanding of the particle-matter interactions in high-energy physics, the experimental setup and detector conditions in the given data-taking period, the employed simulation techniques and models, and the reconstruction performance are required. In the remainder of this thesis, these aspects are addressed first, before the measurement of charged-particle distributions at $\sqrt{s} = 8$ TeV is finally presented in Chapter 7.

2.7 Particle-Matter Interactions

Particle detection and particle reconstruction depend on the interactions between particles and detector materials. As charged and neutral particles traverse the ATLAS detector, they interact with sensitive detector elements as well as the insensitive material e.g. from service and support structures. While charged particles immediately interact with the material through Coulomb forces, uncharged particles can often pass through parts of the detector unnoticed. In order to be detected they have to undergo an interaction that often causes the production of charged secondary particles. However, the same types of particle–matter interactions can also lead to a degradation of the resolution of the reconstructed particle properties, and they can reduce the efficiency to reconstruct these particles altogether.

These physics interactions and their probabilities are characterised by the properties of the incident particles, such as their charge z, velocity $\beta = v/c$ and total energy, as well as properties of the traversed materials, such as their atomic number Z, atomic mass A, mean ionisation potential I, density ρ and thickness x, and quantities derived from these properties. This section presents a survey of the physics interactions that describe the passage of high-energy particles through matter, largely following the summary given in the 2015 update of Chapter 32 in the latest available PDG review [4], and discusses their effects on the studies presented in this thesis.

2.7.1 Energy Loss

In high-energy physics applications, the particle energy is typically much larger than the binding energy of electrons in the atoms of which the traversed detector material is composed. This is particularly the case when fast charged particles travel with relativistic speed $v = \beta c$. The atoms of the traversed matter can be approximated as nuclei at rest surrounded by free electrons. Charged particles interact with these nuclei and electrons by electromagnetic forces. These

inelastic collisions lead to mostly small energy losses due to ionisation ($\mathrm{O}(\alpha^2)$),[8] bremsstrahlung ($\mathrm{O}(\alpha^3)$), direct pair production ($\mathrm{O}(\alpha^4)$) and photonuclear interactions, as well as atomic and collective excitations which are are only significant at lower energies. Charged hadrons can also interact with nuclei via single inelastic collisions, which are described in Section 2.7.3.

2.7.1.1 Ionisation

For a moderately relativistic, charged, heavy particle (with mass $M \gg m_e$), the mean rate of energy loss $\langle \mathrm{d}E/\mathrm{d}x \rangle$ per unit length x due to ionisation, also referred to as *mass stopping power*, can be described by Bethe's equation [48] (also known as Bethe–Bloch equation):

$$\langle \frac{\mathrm{d}E}{\mathrm{d}x} \rangle = K z^2 \frac{Z}{A} \frac{1}{\beta^2} \left[\frac{1}{2} \ln \frac{2 m_e c^2 \beta^2 \gamma^2 W_{\max}}{I^2} - \beta^2 - \frac{\delta(\beta\gamma)}{2} \right] \quad , \tag{2.4}$$

where $K = 4\pi N_A r_e^2 m_e c^2$ is a constant, m_e is the electron mass at rest, $r_e \approx 2.818$ fm is the classical electron radius, Z and A denote the atomic number and atomic mass of the traversed material, I is the mean excitation energy of the material in units of eV, $W_{\max} = 2 m_e c^2 \beta^2 \gamma^2 / (1 + 2\gamma m_e / M + (m_e/M)^2)$ is the maximum energy transfer to an electron in a single collision in units of MeV, and $\delta(\beta\gamma)$ describes the density effect correction. The mean rate of energy loss is therefore functionally dependent on β, but not on the total thickness x of the traversed material.

Particles with the same momentum have similar but slowly decreasing rates of energy loss in materials with increasing Z, because the ratio Z/A for stable elements tends to slowly decrease with Z. In mixed and compound materials, the contributions of individual materials can be added proportionally (Bragg additivity) and one obtains

$$\langle \frac{\mathrm{d}E}{\mathrm{d}x} \rangle = \sum_j w_j \langle \frac{\mathrm{d}E}{\mathrm{d}x} \rangle_j \quad , \tag{2.5}$$

where $\langle \mathrm{d}E/\mathrm{d}x \rangle_j$ is the mean rate of energy loss in the j-th component in units of $[\mathrm{MeV\,cm^{-1}}]$. An equivalent treatment can be used to determine the ratio $\langle Z/A \rangle$. This method would underestimate $\langle I \rangle$, though, because the binding energies of electrons in compounds are higher than in the pure elements; a more accurate treatment is given in Ref. [49]. Effective values of $\langle \delta \rangle$ have been tabulated for chemical elements and for a large number of mixtures and compound materials [50, 51].

Ionisation loss is a stochastic process for which the energy transfer probability distribution can be described by the asymmetric Landau distribution [52]. The most probable energy loss for a particle traversing a moderate amount of material, with thickness x in units of $[\mathrm{g\,cm^{-2}}]$, can be calculated as

[8] Here, α denotes the fine-structure constant with $\alpha = e^2/4\pi\epsilon_0\hbar c \approx 1/137$.

$$\Delta_p = \xi \left[\ln \frac{2mc^2\beta^2\gamma^2}{I} + \ln \frac{\xi}{I} + j - \beta^2 - \delta(\beta\gamma) \right] \quad , \tag{2.6}$$

with $\xi = \frac{K}{2}\langle\frac{Z}{A}\rangle\frac{x}{\beta^2}$ and $j = 0.200$. It is therefore functionally dependent on the thickness of the traversed material. The rare events in the long tail of the Landau-shaped energy loss distribution arise physically from a large energy transfer of an ionising incident particle to knock-on electrons (δ-rays). The mean energy loss $\langle \mathrm{d}E/\mathrm{d}x \rangle$ is weighted by all events in the distribution, and is therefore higher than the most probable energy loss. The PDG review emphasises that *"few concepts in high-energy physics are as misused as* $\langle \mathrm{d}E/\mathrm{d}x \rangle$*"* [4]. In many practical applications the most probable energy loss is found to be more representative to describe the process and should be used instead of the mean value; in very thin absorbers, however, only few poisson-distributed collisions take place, and the total energy loss shows a large variance. Ionisation losses affect the measurements presented in this thesis mainly by the degradation of the reconstructed track parameter resolutions, which are discussed in Chapter 5.

At low energies between $0.05 < \beta\gamma < 1$, shell corrections must be included in the Bethe equation, as well as the "Bloch correction" to include higher-order terms, and an additional "Barkas correction" that decreases the mass stopping power for negatively charged particles. At high energies the electric field of a charged particle flattens and extends, and the probability for distant collisions increases, while polarisation effects limit the increase. The density effect correction of the Bethe equation then becomes

$$\frac{\delta(\beta\gamma)}{2} \quad \rightarrow \quad \ln\left(\frac{\hbar\omega_p}{I}\right) + \ln\beta\gamma - \frac{1}{2} \quad , \tag{2.7}$$

with the plasma energy $\hbar\omega_p$ that depends on the electron density of the material. Under such conditions, the mean energy loss grows with $\ln\beta\gamma$ rather than $\ln\beta^2\gamma^2$. This correction is usually calculated using a parameterisation introduced by Sternheimer [50, 53]. With all these corrections the accuracy of Bethe's equation is better than 1% for particles traversing intermediate-Z materials at energies within most of the "Bethe region" that extends roughly between $0.1 < \beta\gamma < 1000$. The total mass stopping power for muons in Cu, taking all these effects and corrections into account, is visualised in Figure 2.5.

Many relativistic particles that are produced in high-energy collision experiments are called *minimum-ionising particles* or MIPs due to their mean energy loss rates being rather close to the minimum of the Bethe region. However, for particles at very high energies above $\beta\gamma > 1000$, radiative losses become relevant, as described in the following section.

2.7.1.2 Bremsstrahlung

Accelerating charged particles emit electromagnetic radiation known as bremsstrahlung photons, approximately in their direction of flight, to which they transfer a part of their energy. This

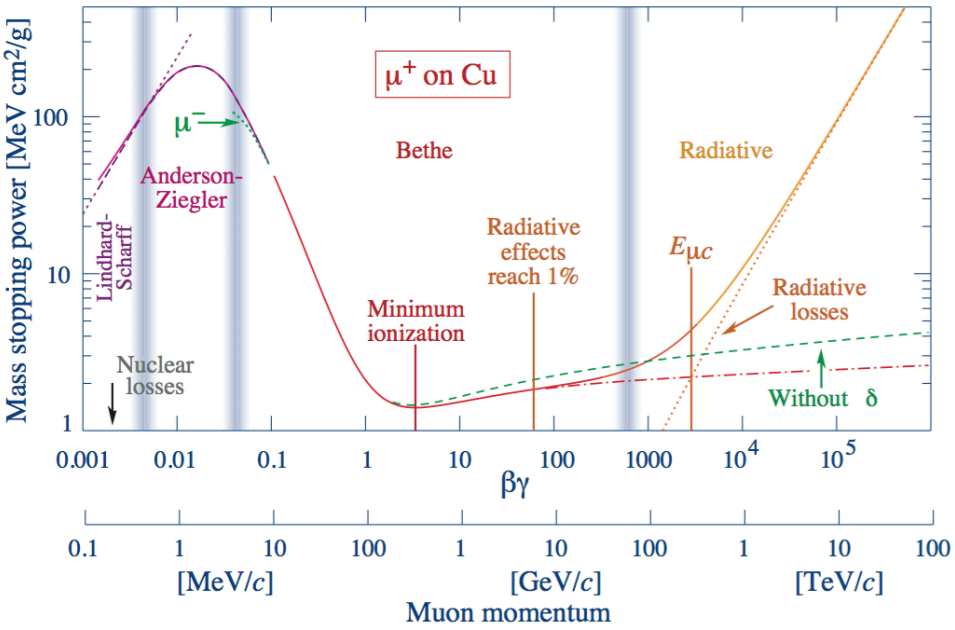


Figure 2.5: Mass stopping power (mean energy loss) of muons in Cu, displayed as a function of $\beta\gamma = p/mc$. The solid curve shows the total stopping power when all effects are added. The vertical shaded bands indicate the transitions between different regions in which specific approximations hold. Plot taken from Ref. [4] (PDG, 2015 update).

effect becomes dominant at high energies, when charged particles interact with magnetic fields and with the Coulomb fields of nuclei. The effect is suppressed by $1/m^2$ and therefore happens predominantly for the light-weight electrons and positrons, where bremsstrahlung exceeds ionisation effects beyond critical energies of $E_c \approx 800\ \mathrm{MeV}/(Z + 1.2)$; for heavier particles such as muons and pions, the effect only becomes significant above a much higher critical energy threshold in the TeV regime (e.g. in cosmic rays). The average rate of muon energy loss can be expressed by

$$-\frac{\mathrm{d}E}{\mathrm{d}x} = a(E) + b(E)E \quad , \tag{2.8}$$

where $a(E)$ is the ionisation energy loss given by Eq. 2.4, and $b(E)E$ comprises the sum of e^+e^- pair production, bremsstrahlung, and photonuclear contributions.

The radiation length X_0 is the characteristic mean distance for a high-energy electron to radiate off all but $1/e$ of its energy by bremsstrahlung while traversing a particular medium. It is given in units of $[\mathrm{g\,cm^{-2}}]$. This is also the characteristic scale length to describe the development of electromagnetic showers (or cascades) at high energies. The Tsai formula [54] can be used to calculate X_0 as

$$\frac{1}{X_0} = 4\alpha r_e^2 \frac{N_A}{A}(Z^2[L_{\text{rad}} - f(Z)] + ZL'_{\text{rad}}) \quad , \tag{2.9}$$

where $f(Z)$ can be expressed by an approximation with αZ dependence [55]. The parameters L_{rad} and L'_{rad} have been tabulated for all elements, which from $Z \geq 5$ onwards follow simple analytic logarithmic expressions with Z dependence, such that the part in parentheses at the end of Eq. 2.9 simplifies to $Z(Z+1)\log(183Z^{1/3})$.[9] For mixtures and compounds, the combined radiation length can be approximated by

$$\frac{1}{X_0} = \sum_j \frac{w_j}{X_{0,j}} \quad , \tag{2.10}$$

where w_j is the fractional weight of the j-th element in the compound, and $X_{0,j}$ is its radiation length.[10]

Ionisation losses of fast electrons can be described by the theory of Bethe and Heitler [56]. Their total mean energy loss per unit length x due to bremsstrahlung can be expressed by

$$\left(\frac{\mathrm{d}E}{\mathrm{d}x}\right)_{\text{rad}} = -\frac{E_{\text{i}}}{X_0} \quad , \tag{2.11}$$

and therefore depends only on the initial energy E_{i} of the particle and the radiation length X_0 of the traversed material. With an approximation using Tsai's formula given in Eq. 2.9, the cross section of the energy loss of electrons due to bremsstrahlung can be expressed as

$$\frac{\mathrm{d}\sigma}{\mathrm{d}k} = \frac{A}{X_0 N_A k}\left(\frac{4}{3} - \frac{4}{3}y + y^2\right) \quad , \tag{2.12}$$

where k is the energy that has been transferred from the electron to the radiated photon, and $y = k/E$ is the fraction of k with respect to the initial electron energy E. The number of photons emitted by an electron, with energies between k_{min} and k_{max}, can be approximated (for short path lengths $d \ll X_0$) by

$$N_\gamma = \frac{d}{X_0}\left[\frac{4}{3}\ln\left(\frac{k_{\text{max}}}{k_{\text{min}}}\right) - \frac{4}{3}\frac{k_{\text{max}} - k_{\text{min}}}{E} + \frac{k_{\text{max}}^2 - k_{\text{min}}^2}{2E^2}\right] \quad . \tag{2.13}$$

This process affects the study of primary-charged-particle distributions presented in this thesis twofold: (1) it affects the reconstruction of the initial properties of primarily produced electrons

[9] Another approximation gives $X_0 = (716.4\,A)/(Z(Z+1)\ln(287/\sqrt{Z}))\,[\mathrm{g\,cm^{-2}}]$.
[10] The same approximation also holds for the nuclear interaction length L_0 which is introduced in Section 2.7.3.

and positrons; (2) the produced bremsstrahlung photons can give rise to further secondary charged particles via the process of pair production, as described in the next section.

2.7.1.3 Pair Production

High-energy photons can lose energy while traversing matter due to the photoelectric effect, Compton scattering, Rayleigh scattering, photonuclear absorption, and pair production. At high energies the process of electron-positron pair production dominates. Based on Tsai's formula from Eq. 2.9, the total cross section for e^+e^- pair production in the high-energy limit can be approximated by

$$\sigma = \frac{7}{9}\frac{A}{X_0 N_A} \quad , \tag{2.14}$$

which is accurate within the level of a few percent for energies of 1 GeV and above, especially for materials with high Z. This effect is relevant for the study of primary-charged-particle distributions presented later in this thesis, as it gives rise to the production of secondary charged particles in the material structures of the measurement device, thus introducing a source of background in the measurement. Details are given in Section 7.4.2.

2.7.2 Multiple Scattering

Any charged particle traversing a medium is subject to successive individual deflections by small angles which mostly arise due to Coulomb scattering from nuclei, as well as due to strong interactions in the case of charged hadrons. Under the central limit theorem, the net scattering angle and the net displacement due to large numbers of individual deflections can be assumed to follow Gaussian distributions which are symmetrically centred around zero. The RMS width of this Gaussian approximation of the projected scattering angle, θ_0, was first expressed by Molière [57] and can be calculated more accurately with the Highland formula [58],

$$\theta_0 = \frac{13.6 \text{ MeV}}{\beta c p} Z \sqrt{t/X_0}(1 + 0.038 \ln(t/X_0)) \quad , \tag{2.15}$$

where p is the momentum and Z is the charge number of the incident particle, while t/X_0 represents the thickness of the scattering medium in units of radiation length. For a layered material of low total thickness, the same equation can be applied when using t/X_0 of the combined material. For electrons traversing materials of higher total thickness, one also needs to take into account that the electron momentum changes continuously due to relatively high radiation losses [59].

The actual probability density function (PDF) of the projected scattering angle contains a Gaussian core describing $\approx 98\,\%$ of the distribution, as well as non-Gaussian tails due to rare

large-angle scattering processes. These can be described with a Gaussian mixture model, in which two Gaussians are utilised to describe the core and the tails, respectively [60]:

$$f(\theta_{\text{ms}}) = (1-\epsilon)\cdot g_0(\theta_{\text{ms}};\sigma_{\text{core}}) + \epsilon \cdot g_0(\theta_{\text{ms}};\sigma_{\text{tail}}) \quad , \tag{2.16}$$

where g_0 signifies Gaussian distributions with mean values of zero and widths of σ_{core} and σ_{tail} for the core and tail component, respectively. Both widths and the parameter ϵ are model parameters which depend on the traversed material in terms of t/X_0. For very thin scatterers, a semi-Gaussian mixture model with a non-Gaussian parameterisation of the single-scattering angle can be applied to describe the tails more accurately [61].

The multiple scattering processes affect the measurements presented in this thesis mainly by the degradation of reconstructed track parameter resolutions, and also by increasing the probability to reconstruct tracks with too high momentum, as described in Section 7.5.3.2.

2.7.3 Hadronic Interactions

Charged hadrons can collide with nuclei with an interaction cross section σ that is proportional to the geometrical cross section $A^{2/3}$ of the nuclei. At high particle energies, the mean free path until such an interaction occurs can can be approximated by

$$\lambda = \frac{A}{N_A \rho \sigma} \approx A^{1/3}\cdot 35\,[\text{g cm}^{-2}] =: L_0 \quad . \tag{2.17}$$

This is also referred to as the nuclear interaction length L_0, which is much longer than the radiation length X_0 in the case of high-Z materials. In such inelastic collisions, the hadron–nucleus interactions produce secondary hadronic showers. These inelastic scattering processes can be classically described by a multi-chain model in which the incident particle penetrates into a nucleus and collides with a nucleon inside, producing a number of hadronic secondary particles. As long as these particles (including the scattered incident particle and recoil nucleon) remain within the nucleus, they can interact again, thus developing an intra-nuclear cascade.

At the end of the interaction, even if the incident particle escapes the nucleus and remains intact, it can no longer follow its former trajectory, as its energy and direction will have changed significantly. This process has major consequences on the reconstruction efficiency of charged hadronic particles, and gives rise to the main source of systematic uncertainty in the measurement of charged-particle distributions. The effects of hadronic interactions due to the presence of detector material will be discussed in several chapters throughout this work.

Chapter 3

The LHC and the ATLAS Experiment

The well-known proverb, *"Extraordinary claims require extraordinary evidence,"* that has been popularised by the late Carl Sagan and can be traced back to earlier scientists such as Laplace [1] and Hume,[2] can be seen as a defining principle for the realm of experimental particle physics: any claim of a new physics discovery is required to be supported by evidence of a staggering 5σ significance. Even for studies that do not aim to discover new physics phenomena, very high precision and diligence are indispensable to build up knowledge and make progress in the field. As the collection of such evidence inevitably requires outstanding effort, from intellectual power to cutting-edge technology, the proverb might thus be extended to a corollary: *"Extraordinary evidence requires extraordinary instrumentation."*

This chapter introduces the extraordinary instrumentation that lies at the heart of the studies presented in the remainder of this thesis. It summarises the up-to-date knowledge of the LHC and the ATLAS experiment, and highlights the main components of the ATLAS detector and the trigger system that are most relevant in this context. The specific experimental setup that was used for the main study presented in this thesis is also discussed.

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [1] is part of a predominantly subterranean accelerator complex at the European Centre for Nuclear Research (CERN) in the Swiss/French border area near Geneva, Switzerland. It is a superconducting circular hadron collider of almost 27 km circumference, installed within the underground tunnel of the previous Large Electron-Positron (LEP) collider experiment at a depth of 50–175 m. The LHC has been designed and built in order to accelerate proton and heavy-ion beams in two adjacent accelerator rings, using 1232 powerful dipole magnets to create two opposite magnetic fields of up to 8.4 T which keep the particles in both rings on track, while quadrupole magnets keep the particles tightly bunched together. These particle bunches can be brought to head-on collision at four interaction points (IPs) that are used by the main high-energy particle physics experiments at the LHC, called ATLAS, CMS, ALICE and LHCb, as shown in Figure 3.1.

[1] "The weight of evidence for an extraordinary claim must be proportioned to its strangeness."
[2] "No testimony is sufficient to establish a miracle, unless the testimony be of such a kind, that its falsehood would be more miraculous than the fact which it endeavours to establish."

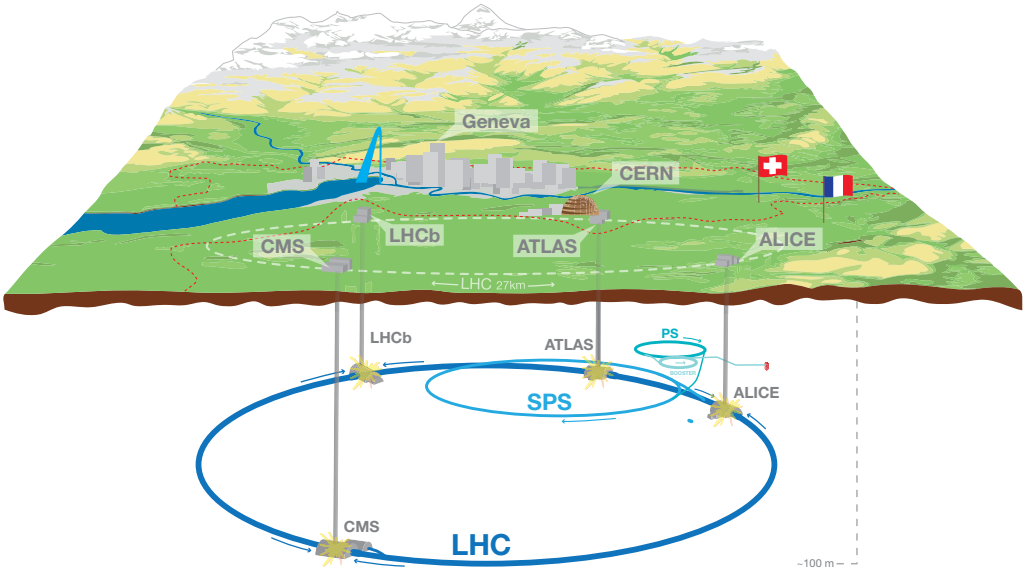


Figure 3.1: Schematic overview of the LHC accelerator complex, including the ALICE, ATLAS, CMS and LHCb experiments. Image taken from Ref. [62].

The LHC machine is able to deliver proton-proton (pp) collisions at a centre-of-mass energies of up to $\sqrt{s} = 14$ TeV. During the Run-1 period of its operation,[3] data was collected by the experiments at pp collision energies of $\sqrt{s} = 0.9$, 2.36, 7 and 8 TeV. In addition, the LHC has also been used to deliver proton-lead (p+Pb) and lead-lead (Pb+Pb) collisions with up to $\sqrt{s} = 5.5$ TeV per colliding nucleon pair, which are of special interest to experiments such as ALICE.

Besides the centre-of-mass energy, the instantaneous luminosity is considered the most important machine parameter. It is defined for pp collisions as:

$$\mathcal{L} = \frac{\mu \, n_b \, f_r}{\sigma_{\text{inel}}} \, , \tag{3.1}$$

where μ is the average number of the Poisson-distributed inelastic pp interactions per proton bunch crossing,[4] n_b is the number of colliding pairs of proton bunches within the LHC storage rings, f_r is the revolution frequency of proton bunches, and σ_{inel} is the cross section for inelastic pp interactions. Each of the proton bunches typically contains $\mathrm{O}(10^{11})$ protons.

[3] The data-taking period referred to as *Run-1* took place between late 2009 and early 2013, delivering pp collisions of up to $\sqrt{s} = 8$ TeV. It was followed by the *Long Shutdown 1* (LS1) period in which the accelerators and experiments were maintained and upgraded to new state-of-the-art technology. The *Run-2* data-taking period commenced in Spring 2015 and is currently scheduled to last until the end of 2018, aiming to deliver pp collisions of up to $\sqrt{s} = 14$ TeV.

[4] The value of μ tends to decrease during a fill due to slowly decreasing instantaneous luminosity corresponding to steadily declining beam intensities.

ATLAS and CMS are general purpose experiments designed to perform new physics searches (including SUSY, dark matter, neutrino physics, and jet substructure studies) as well as precision measurements of the properties and production cross sections of known elementary particles (such as the differential top-quark cross-section, or the W boson mass). Most of these studies require the design high-luminosity operation of up to $\mathcal{L} \approx 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, for which the LHC storage rings can be filled with up to 2808 proton bunches with a bunch spacing of 25 ns, thus providing a collision rate of up to 40 MHz.[5]

3.1.1 Experimental Setup for Charged-Particle Measurements

The high-luminosity operation is needed by most studies at the ATLAS experiment in order to collect as many collision events with signatures of the studied processes as possible, such as rare $H \to ZZ^{(*)} \to \mu\mu\mu\mu$ decays. Under these conditions it is desirable to record several pp interactions simultaneously in one bunch crossing in order to maximise the yield of rare physics processes. The occurrence of simultaneous interactions within one bunch crossing is also referred to as in-time *pile-up*.

For the charged-particle multiplicity measurement presented in this thesis, a different setup is needed in order to minimise the probability of more than one simultaneous pp interaction per bunch crossing,[6] as well as to minimise event backgrounds which can contaminate the recorded data sample. The required low-μ conditions are ensured by a μ-scanning technique in which both focused proton beams are adjusted such as to achieve only a partial overlap between their two approximately Gaussian beam profiles. The proper adjustment for stable low-μ conditions during the run is constantly monitored with dedicated beam luminosity monitors [63].

Parameter	**design operation (2012)**	**low-μ run**
Beam energy (TeV)	4	4
Filled paired bunches per beam n_b	1380	28
Particles per bunch	1.7×10^{11}	$< 9 \times 10^{11}$
Time between collisions (ns)	50	≥ 50
Inst. online luminosity $\mathcal{L}$ ($\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$)	7.7×10^{33}	$< 1.8 \times 10^{28}$
β^* at IP of ATLAS (m)	0.6	0.6
Average pp collisions per bunch crossing $\langle\mu\rangle$	20.7	< 0.004

Table 3.1: Summary of LHC parameters during Run-1 operation at the design luminosity in 2012 (center) and during the part of the low-μ run in April 2012 which was used for the Minimum Bias analysis at $\sqrt{s} = 8$ TeV (right).

[5] The LHCb and ALICE detectors are designed for more specific physics analyses (B-physics and heavy-ion collisions) at a lower luminosity operation of $\mathcal{L} \approx 10^{32}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$.

[6] The actual probability for such pile-up events can be calculated from μ by using the Poisson distribution.

Using a dedicated low-μ setup, the corresponding data sample was recorded in a special run during April 2012 at $\sqrt{s} = 8$ TeV, in which all required detector sub-systems were operational. The delivered peak luminosity was kept below $\mathcal{L} \leq 1.8 \times 10^{28}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ in the analysed part of the data, corresponding to $\mu < 0.004$. The special run furthermore used only 28 paired proton bunches (and 19 unpaired bunches per beam that were used to study beam background effects), grouped into two bunch trains with 50 ns bunch spacing and maximum beam intensities of 8.93×10^{11} and 8.29×10^{11} protons, respectively. Table 3.1 summarises the most important run parameters of the LHC and ATLAS, comparing typical values at the design luminosity operation with the setup for the low-μ run in April 2012.

At the LHC, charged-particle multiplicity distributions and observables have also been measured by the CMS, ALICE and LHCb experiments; a full list of references to available measurements is given in Chapter 7.

3.2 The ATLAS Detector

As of today, the ATLAS experiment [47] is mankind's largest-volume general-purpose particle detector. The immensely complex measurement device comprises several sub-detectors, one solenoidal and one toroidal magnet system, as well as cabling, cooling and supporting structures. It measures approximately 25 m in height and 44 m in length, weighs about 7000 tons, and is located about 100 m below the surface of the French and Swiss countryside close to the main CERN site. A schematic view of the apparatus is shown in Figure 3.2.

The ATLAS detector has been primarily designed to enable the search for *new physics*. It is also used for precision measurements of known elementary particles, as well as studies of various properties of collision events, such as charged-particle multiplicities. The apparatus can be divided into three main sub-detector parts, where the former are encapsulated by the latter:

- Inner Detector (ID)
- Calorimeters
- Muon Spectrometer (MS)

The ID and MS are tracking systems that enable precise momentum measurements of charged particles. The ID is also responsible for precise measurements of primary and secondary interaction vertices as well as parameters of charged-particle trajectories. The calorimeters are utilised to determine the total energies of traversing electromagnetic and hadronic particles. Together these sub-detector components also allow for a precise identification of particles which arise from the high-energy collisions.

As the ID is the only sub-detector of ATLAS which is relevant for the work presented in this thesis, the Calorimeters and Muon Spectrometer will only be described briefly here. Extensive details about these latter two subsystems can be found in the ATLAS Technical Design Report [64].

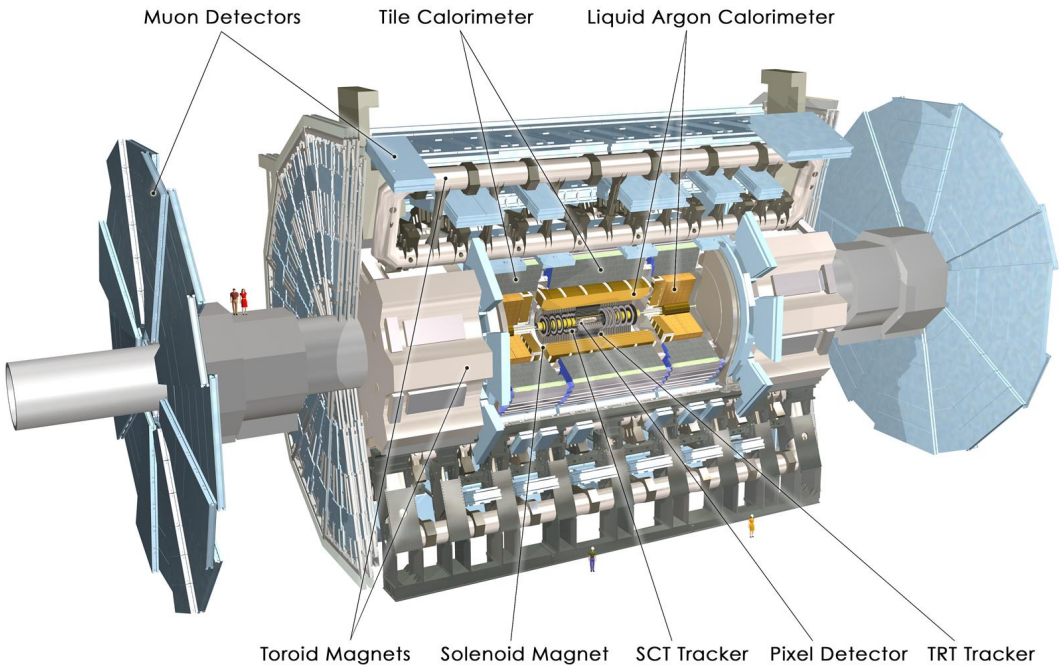


Figure 3.2: Computer-generated image of the ATLAS detector and its main sub-detector systems. Image taken from Ref. [47].

3.2.1 The ATLAS coordinate system

The ATLAS coordinate system is defined as a right-handed Cartesian coordinate system with the origin at the nominal interaction point in the geometrical centre of the detector. The positive direction of the x-axis is defined as pointing from this origin to the centre of the LHC ring, the positive direction of the y-axis points upwards, and the positive z-axis points counter-clockwise along the beam direction when viewed from above. Cylindrical coordinates r and ϕ (the azimuthal angle around the z-axis) are used in the transverse (x, y)-plane. Furthermore, two quantities which are often used in high-energy particle collision experiments are the pseudorapidity η, defined in terms of the polar angle θ as

$$\eta = -\ln\tan(\frac{\theta}{2}), \tag{3.2}$$

and the transverse momentum p_{T} in the (x, y)-plane, which is given by

$$p_{\mathrm{T}} = p\sin\theta. \tag{3.3}$$

The angular distance between trajectories is measured in units of $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$. Individual detector modules in the ID use local right-handed Cartesian reference frames which have their origin in the geometrical centre of each device. Details can be found in Ref. [64].

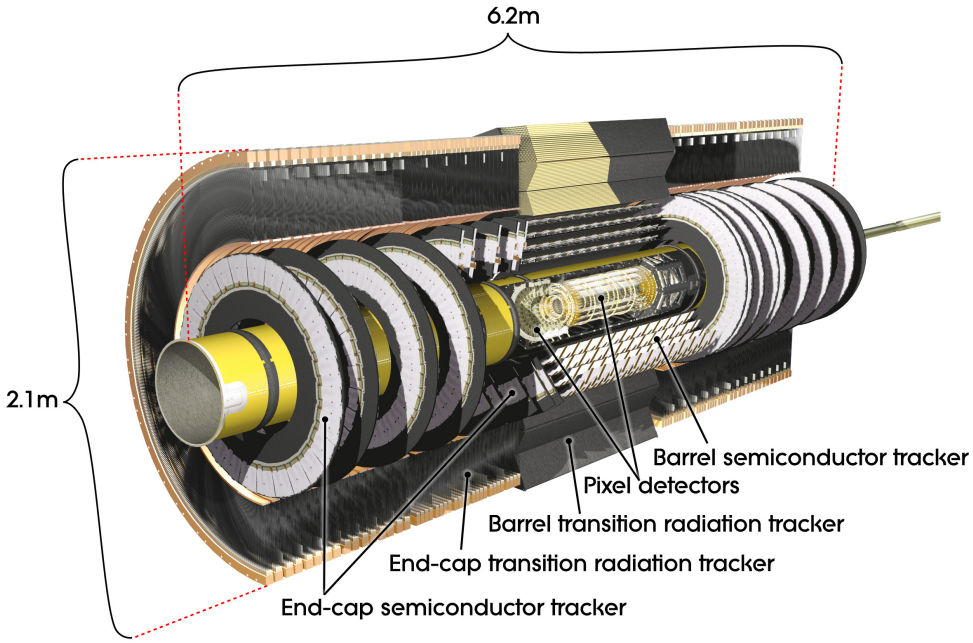


Figure 3.3: Computer-generated image of the ATLAS inner detector (ID) and its subcomponents. Image taken from Ref. [47].

3.2.2 Inner Detector

The Inner Detector is the innermost tracking device of ATLAS. It was designed to provide precise measurements of the trajectories and kinematic properties of charged particles, as well as the locations of primary and secondary vertices. It comprises three sub-detectors employing different technologies: a silicon pixel detector (Pixel), a silicon microstrip detector (Semiconductor Tracker or SCT) and a transition radiation tracker (TRT). A solenoid immerses all three sub-detectors in a 2 T axial magnetic field that causes charged particles to propagate along curved trajectories, from which their momenta in the transverse (x, y)-plane can be derived.

The ID is designed to provide full coverage in ϕ and within $|\eta| < 2.7$, whereas the acceptance for ID track reconstruction is restricted to $|\eta| < 2.5$. Its sub-detectors are organised into a barrel region with cylindrical setups around the interaction point, and two end-cap regions with disc-shaped radial setups in the forward and backward direction, respectively. The ID barrel region comprises three Pixel layers, four SCT double-layers, and 73 layers of TRT straws. Each end-cap region has three Pixel layers, nine SCT double-layers, and 160 TRT layers. The sensitive elements in the barrel cover radial distances from the beam axis of 50.5–150 mm, 299–560 mm and 563–1066 mm, respectively. Typical position resolutions are 10 µm, 17 µm and 130 µm for the r–ϕ coordinate of each sub-detector, and 115 µm and 580 µm for the longitudinal coordinate of the Pixel and SCT sub-detectors, respectively. A schematic view of the ID is presented in Figures 3.3, 3.4 and 3.5.

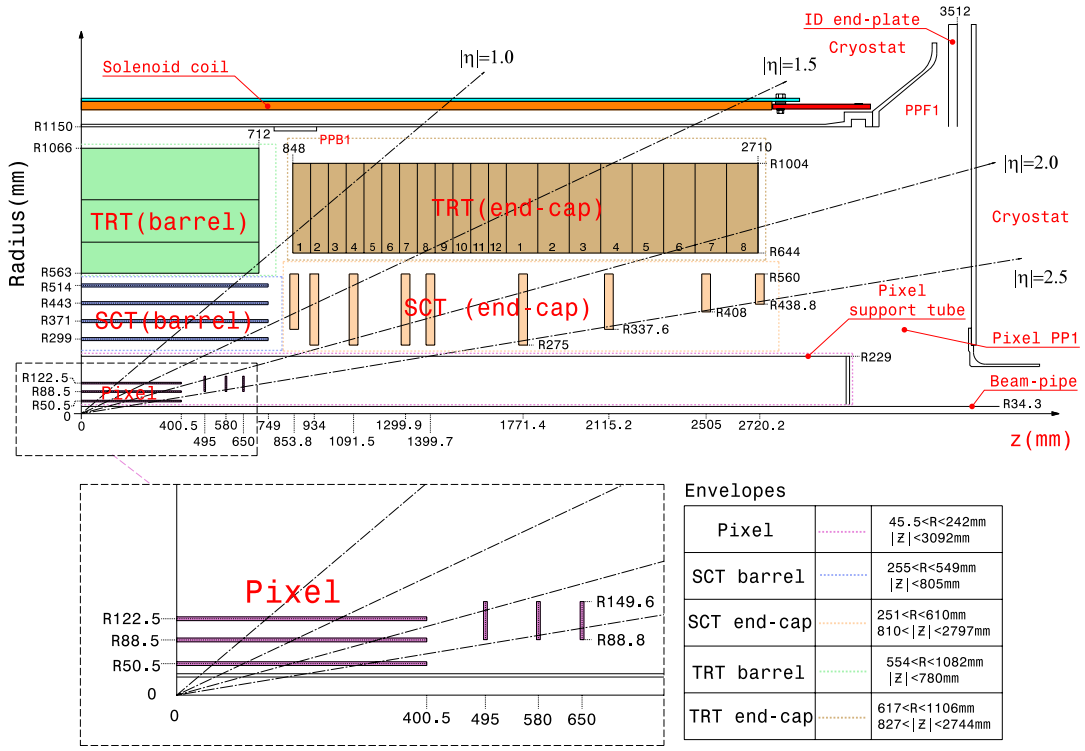


Figure 3.4: Schematic planar view of a quadrant of the ID, highlighting the main subsystems and detector elements along with their active dimensions. The lower part shows a zoom of the Pixel detector region. Image taken from Ref. [47].

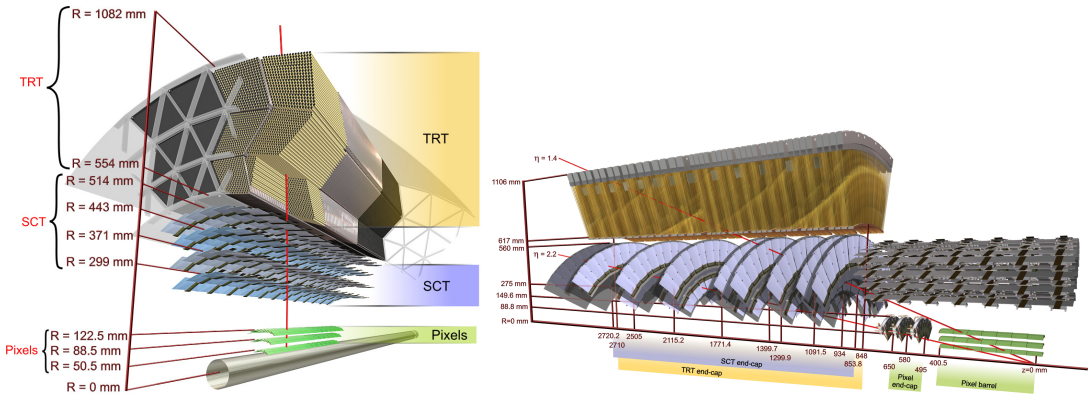


Figure 3.5: Schematic illustration of the sensor layers and structural elements in (a) the barrel and (b) the end-cap region of the ID. Images taken from Ref. [47].

3.2.2.1 Pixel Detector

The Pixel detector is the innermost sub-detector of ATLAS with the highest spatial resolution of all components. It consists of 1744 silicon modules arranged in three cylindrical barrel layers as well as three end-cap discs on each side. The innermost cylindrical layer, closest to the beam axis, comprises modules positioned at a radius of approximately 50.5 mm. It is also called *B-Layer* since it plays an indispensable role in B-physics, which relies upon the reconstruction of

jets produced by bottom quarks. The barrel modules are arranged in a turbine-like layout at an azimuthal angle of 20° [65]. The silicon modules of 250 µm sensor thickness comprise a total of 80.8 million silicon pixels with a typical size of $50 \times 400\,\mu\text{m}^2$ in the transverse and longitudinal direction.[7] Each individual pixel provides a measurement signal if it collects an amount of charge beyond a tuneable threshold, measured by the pulse height using a time-over-threshold (ToT) technique.

A charged particle traversing a pixel sensor deposits energy through ionisation (and the creation of electron–hole pairs) typically in more than one adjacent pixel. Electronic signals in such *pixel clusters* can be formed due to a non-zero incident angle between the particle trajectory and the boundaries of individual pixels, the diffusion of electrons and holes among adjacent pixels within the sensor, or a charge drift caused by the magnetic field. Details about the extraction and utilisation of pixel cluster information by the ATLAS reconstruction software are given in Ref. [66].

The pixel detector is contained within a service structure called the Pixel Support Tube (PST), and requires further service structures including the Patch Panel (PP) regions for its operation and maintenance. It is crucial to minimise the passive or non-sensitive material of such support structures, including cabling, heating and cryogenics, during the construction of the ID, as such material can interfere with the particle reconstruction due to particle–matter interactions, yet does not provide any useful measurable information. These effects and their impact on the presented measurements are investigated in Section 6.1.

3.2.2.2 Semiconductor Tracker (SCT)

The SCT uses a similar layout as the Pixel detector. It consists of silicon strip modules arranged in four cylindrical barrel layers, as well as nine disc layers on each end-cap side. Each measurement layer comprises two silicon strips that are rotated by an angle of 40 mrad with respect to each other, in order to allow a precise position measurement while reducing the total number of required readout channels to approximately 6 million.

The Pixel and SCT detectors provide measurement signals within $|\eta| < 2.5$. They comprise the most important components of the ATLAS detector for the purpose of precise track reconstruction in the studies presented in this thesis. As they employ similar technologies based on silicon modules, the total number of measurement clusters in both systems along a single reconstructed particle trajectory is commonly referred to as *silicon hits*; this number is often used as a track quality criterion by the reconstruction algorithms.

It is possible that a charged particle traverses a Pixel or SCT sensor without generating any measurement signal, such as when the corresponding silicon module is malfunctioning or inactive; this is usually taken into account in the detector conditions during a measurement, to avoid biasing the reconstruction of the particle trajectory and momentum. It is also possible that a measurement signal is generated in a sensor without being traversed by a corresponding

[7] In the regions between the readout chips, so-called *long pixels* with a longitudinal size of 600 µm are used.

charged particle at all; such detector noise can contribute to the reconstruction of so-called fake tracks (see Section 5.1).

3.2.2.3 Transition Radiation Tracker (TRT)

The outermost sub-detector of the inner tracker of ATLAS is known as the TRT. It is a transition radiation detector which also provides a large number of measurements that help improve the track reconstruction, particularly the momentum resolution. The TRT consists of 370 000 drift tubes (called straws) of 4 mm diameter in a dense arrangement, filled with a Xenon-based gas mixture, using a gold-plated Tungsten–Rhenium wire as an anode. The barrel straws are aligned in parallel to the beam pipe and divided by a readout gap at $|\eta| < 0.1$, and are additionally read out at both ends of the tube arrangement. This design limits the measurement resolution in η but provides a good resolution in ϕ. Together with the radially arranged end-cap straws, the TRT acceptance range extends to $|\eta| \lesssim 2.0$.

The TRT straws are interleaved with polypropylene fibres and foils as a transition radiation element, such that when the media boundaries are traversed by ultra-relativistic particles, collinear photons are emitted depending on their Lorentz factor, $\gamma = E/m$, and increase the collected charge that is deposited inside the straw. By using separate signal amplitude thresholds, the TRT can discriminate between (a) signals caused by minimum-ionising particles and (b) transition radiation e.g. from electrons (which have a much higher Lorentz γ-factor due to their lower mass), and can thus be utilised for particle identification.

3.2.3 Calorimeter

The ATLAS calorimeters consist of an electromagnetic (EM) calorimeter, a hadronic calorimeter, and a forward calorimeter. The combination of these components provides an almost hermetical enclosure around the interaction point, up to $|\eta| < 4.9$, which is crucial for precise measurements of the missing transverse energy, $E_{\mathrm{T}}^{\mathrm{miss}}$; this observable is a typical signature of final states including neutrinos and various new physics phenomena. Solenoid magnets and cryostats are located between the ID and the calorimeters; furthermore, a pre-sampler within $|\eta| < 1.8$ accounts for the energy that is lost by particles due to interactions with sensitive as well as passive detector material before they reach the calorimeters.

The EM calorimeter comprises accordion-shaped Pb absorbers and liquid-argon (LAr) samplers, arranged in a three-layer barrel and two end-caps that cover a range of up to $|\eta| < 3.2$. It is designed to absorb the energy of electromagnetic particles and thus to provide a high-resolution measurement of electron, photon and jet energies.

The hadronic calorimeter is a sampling calorimeter which surrounds the EM calorimeter. It consists of steel absorbers and tile scintillators in the barrel region, and copper-plate absorber wheels interleaved with liquid-argon samplers in the end-caps. This apparatus is designed to fully absorb the secondary cascades caused by hadronic particles which penetrate the EM

calorimeter; this is mandatory to keep the rate of contamination due to punch-through particles entering the surrounding Muon Spectrometer as low as possible [67].

The Forward Calorimeter (FCAL) is a sampling calorimeter which covers the forward range of $3.1 < |\eta| < 4.9$. It uses liquid argon as an active sampling material, as well as one copper and two tungsten absorber layers.

Furthermore, a dedicated luminosity monitor (LUCID) is installed in the very forward direction at a distance of $\pm 17\,\mathrm{m}$ from the nominal interaction point of ATLAS. This device aims to detect signatures of inelastic *pp* interactions and to provide an online monitoring of the luminosity and beam conditions, as well as a relative measurement of the integrated luminosity. LUCID is complemented by two other forward physics detectors: the zero degree calorimeter (ZDC) at $\pm 140\,\mathrm{m}$ and the roman-pot ALFA detector at $\pm 240\,\mathrm{m}$ from the nominal interaction point of ATLAS, respectively. ALFA is capable of measuring the absolute luminosity as well as signatures of elastic *pp* scattering in the very forward direction [17, 21].

3.2.4 Muon Spectrometer

The Muon Spectrometer (MS) is the largest component of the ATLAS detector. It is located outside of the calorimeters and extends between approximately $4.5 < r < 11\,\mathrm{m}$ and $7 < |z| < 23\,\mathrm{m}$. Eight large superconducting coils in a barrel as well as two end-cap toroids provide a strong and complex toroidal magnetic field. The forward acceptance of the MS extends up to $|\eta| < 2.7$. Its sensitive chambers are arranged in three cylindric barrel layers and four end-cap discs on each side. In the region up to $|\eta| < 2.4$, Resistive Plate Chambers (RPC) in the barrel and Thin Gap Chambers (TGC) in the end-caps provide fast signals which can also be used for hardware triggering. Between $2.0 < |\eta| < 2.7$, Monitored Drift Tube chambers (MDT) and Cathode Strip Chambers (CSC) measure the η and ϕ coordinates of tracks with high precision. When combined with tracking information from the ID, these measurements enable a precise reconstruction of muon trajectories and their momenta (up to 2–3 % for muons within a range of 10–200 GeV).

3.2.5 Trigger System

At its design operation, the LHC delivers collision events to each of its experiments at a rate of up to 40 MHz. The amount of *raw* data, prior to any further data processing steps, is typically of the order of $\gtrsim 1\,\mathrm{MB}$ per recorded collision event. Therefore, if the readout systems allowed to record all collision events, the total amount of data to be stored would rise up to more than 40 TB per second; this would exceed today's technological capabilities by far.

For most physics studies, however, only a tiny fraction of the initially collected data contains the rare signatures of events which may be actually relevant to that particular study. Filtering the data and recording only such events of interest allows to reduce the rate to manageable levels of few hundred Hz. To this end, the Trigger and Data Acquisition (TDAQ) system [68] is employed. This pre-processing system comprises three stages which are called *Level 1* (L1),

Level 2 (L2) and *Event Filter* (EF). At each stage different selection criteria are applied, which are tailored to the specific requirements of every data-taking run as specified in a trigger menu. Only if all three stages are passed by an event, the complete detector information is read out and the data are transmitted to the central mass storage facility at CERN.

The first-level hardware trigger (L1) selects events with predefined signatures such as large $E_{\mathrm{T}}^{\mathrm{miss}}$, high E_{T} objects, jets, leptons, photons, or signals from the Minimum Bias Trigger Scintillators (MBTS). It uses only data from a subset of the detector with reduced granularity and simplified algorithms, and defines regions of interest (RoIs) that are passed to the higher level triggers. At this stage the event rate is reduced to $\sim 75\,\mathrm{kHz}$. The second-level trigger (L2) analyses the RoIs further, now using data from all sub-detectors at full granularity. The event rate is reduced to $\sim 3.5\,\mathrm{kHz}$. If the event passes both trigger levels, its full information is transferred to a dedicated processing farm (EF). At this stage the event is analysed using detailed track reconstruction algorithms, and the event rate is reduced to approximately 300–400 Hz. From here the datasets are sent to the Worldwide LHC Computing Grid (WLCG) system for full offline reconstruction, reprocessing, distribution, detailed analyses, and data storage [69].

3.2.6 Minimum Bias Trigger Scintillators

For the Minimum Bias measurement which is described in this thesis, the trigger system requires a coincidence of hardware (L1) signals from Beam Pickup Timing devices (BPTX) and Minimum Bias Trigger Scintillators (MBTS). The BPTX devices are located at $\pm 175\,\mathrm{m}$ distance from the interaction region. When a pair of bunches simultaneously passes both locations, coincident signals are triggered by the electrostatic button pick-up detectors that are attached to the beam pipe, thus signalling when precisely a collision can be expected inside ATLAS.

The MBTS modules serve the purpose to select genuine pp collision events with the highest possible efficiency, while safeguarding that the recorded sample be as unbiased as possible. They are installed in both forward regions, between the ID end-caps and the LAr end-cap calorimeter cryostats, at a longitudinal distance of $z = \pm 3.56\,\mathrm{m}$ from the nominal point of origin. Each module consists of 16 plastic scintillation counters, each with 20 mm thickness, arranged in eight azimuthal sectors and two rings covering pseudorapidity range $2.08 < |\eta| < 2.82$ and $2.82 < |\eta| < 3.75$. The `L1_MBTS_1` and `L1_MBTS_2` triggers that were used to collect data for this study require a total of at least one or two MBTS signals above threshold, respectively. In addition, a trigger on unpaired (non-colliding) bunches was used to collect data for dedicated beam-background studies. Details on these studies and on the trigger efficiencies are given in Chapter 7. Further details on the MBTS trigger system within the context of Minimum Bias measurements can be found in Ref. [70].

Chapter 4

Detector and Event Simulation

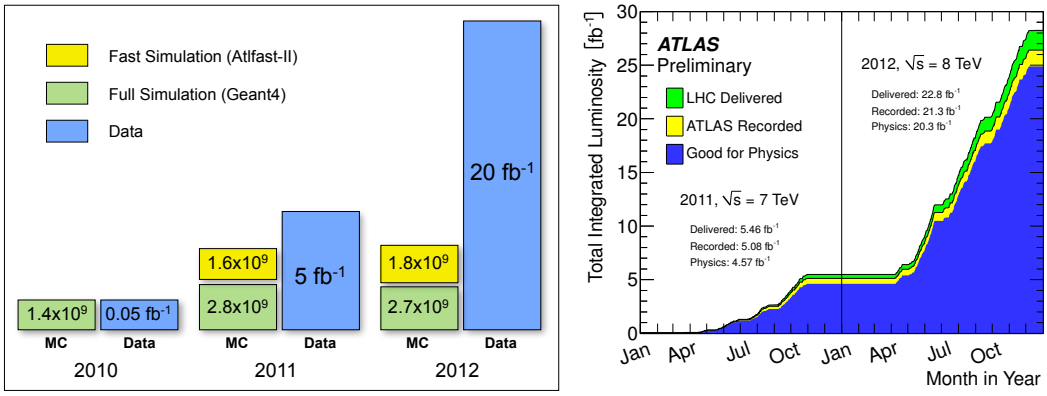
Even with the highest degree of technological sophistication and complexity, physics measurements based on the raw data recorded from the ATLAS detector cannot be perfectly accurate. On one hand, the intrinsic resolution of sensitive detector elements limits the capability to acquire precise knowledge of the particle trajectories. On the other hand, parts of the measurement device can become inefficient or produce noisy signals; particle–matter interactions can degrade the resolutions and reconstruction efficiencies; background processes can contaminate the recorded data samples; parts of the device may not be perfectly calibrated; etc. Some of these effects require corrections to be applied, thus giving rise to systematic uncertainties which must be quantified and accounted for in the final results of physics measurements. In many cases these corrections and systematic uncertainties cannot be derived from data-driven methods, but their precise evaluation is possible with the aid of computer simulations, provided that the models employed in these simulations are accurately reflecting the experimental conditions, detector setup, and underlying physics.

The ATLAS experiment uses an object-oriented, C++-based software framework called ATHENA for the individual or sequential steering of event generation, simulation, digitisation, reconstruction, and dedicated physics analysis tasks.[1] This chapter summarises the ATLAS software framework and simulation infrastructure, introduces the available detector models and simulation strategies, and briefly describes developments during and beyond the Run-1 period of the LHC [74].

4.1 The ATLAS Simulation Framework

The use of Monte Carlo (MC) simulations of particle–matter interactions and the detailed detector response has become a standard technique in most high-energy physics experiments. This involves a wide range of applications, from the commissioning of the highly complex measurement devices to detailed comparisons of measured observables between the simulation and experimental data. Furthermore, many physics analyses depend on large MC datasets in order to estimate systematic uncertainties, model background processes, derive corrections for

[1] ATHENA is an enhanced version of the LHCb GAUDI framework [71, 72]. A comprehensive description of ATHENA is given in Section 3.3 of Ref. [73].



(a) **(b)**

Figure 4.1: (a) Schematic evolution of the number of simulated MC events, using different simulation strategies, and the amount of data recorded by ATLAS in terms of integrated luminosity per year. Plot taken from Ref. [74] with modifications to better reflect the actual values from 2012. (b) Timeline of the cumulative integrated luminosity delivered by the LHC (green), recorded by ATLAS (yellow), and certified as good quality data after reprocessing (blue) using pp collisions with stable beams at $\sqrt{s} = 7$ and 8 TeV in 2011 and 2012. Plot taken from Ref. [75].

detector effects, and study rare processes with with small cross-sections. In many cases a highly detailed simulation is required to reach the necessary precision, which however consumes a large amount of computing resources.

The demand for these large MC samples rises further as the recorded luminosity at the LHC, and hence the amount of ATLAS data to be analysed, increases over time (see Figure 4.1). In order to achieve the goal of large-scale MC production despite the limited computing resources, a variety of simulation techniques as well as detector models with different levels of detail have been developed in recent years [76].

The MC sample generation process is generally divided into three steps that may be combined into a single configurable workflow (termed *job*):

Event generation of primary-particle collections: this can be achieved with MC event generators such as Pythia and Epos (see Section 2.5), which produce primary-particle collections representing final states of pp as well as heavy-ion collision events. These final-state collections in standard HepMC [77] format constitute a major part of the initial conditions which are used as input for the simulation stage.

Simulation of physics processes and detector response: this step can be performed with a variety of simulation strategies that can be characterised by varying degrees of accuracy and simulation speed. It includes the propagation of particles through the magnetic field, the simulation of their physics interactions with the detector material, and the production of hit collections from the simulated detector response to these particle–matter interactions.

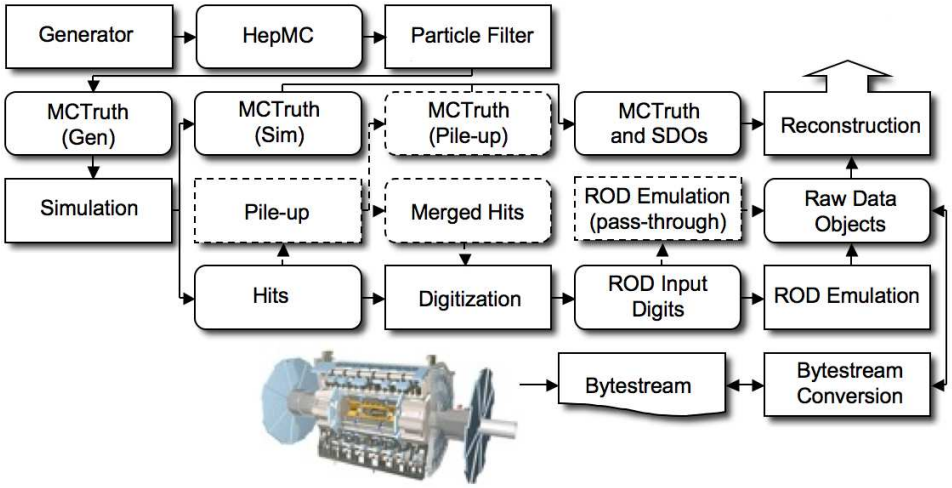


Figure 4.2: Flow diagram of the ATLAS simulation software chain. Algorithms are displayed in rectangular boxes, while persistent data objects are shown in boxes with rounded corners. Dashed boxes represent an optional part of the software chain that can be employed for pile-up simulation. Plot taken from Ref. [76].

Digitisation and pile-up simulation: this step uses the hit collections from the simulation output to produce simulated raw data objects (RDO) for the reconstruction, while optionally including an overlay of simulated pile-up events. These RDOs are roughly equivalent to the collision data recorded directly from the detector.

The simulated objects from the digitisation step become the input for the subsequent reconstruction of particle trajectories (tracks), vertices, and other physics objects. The reconstruction of tracks and vertices is described in Chapter 5; these objects are particularly important in the context of this thesis as they provide the main input to the track-based charged-particle measurement. The output data formats of the reconstruction step are referred to as ESD (Event Summary Data), AOD (Analysis Object Data) and NTUP (n-tuple) [78]; the latter format is used as an input for the two main studies presented in this thesis (see Section 6.1 and Chapter 7). Figure 4.2 presents the ATLAS simulation scheme in a flow diagram.

The remainder of this chapter will be focusing on the simulation step and introduce the corresponding ATLAS detector models (Section 4.2), simulation techniques (Section 4.3), and validation strategies (Section 4.4).

4.2 Detector Models

Detector models with different levels of detail can be used to simulate the response of the ATLAS detector to the particle–matter interactions. The full detector simulation utilises a highly detailed geometry description, while a more simplified geometry is preferred for fast simulation strategies as well as charged-particle reconstruction.

4.2.1 Detailed Detector Models

The entire ATLAS detector geometry can be described by a set of geometrical primitives, which are taken from a library called the `GeoModel` toolkit [79]. This geometry consists of logical volumes corresponding to these objects, with specific dimensions and associated material properties, and of physical volumes which are instantiations of these logical volumes, each with a specified position and orientation.[2] In addition, an abstract readout geometry layer is defined which contains specific information for the simulation of the sensitive detector response (i.e. production of hit collections) and for the digitisation step.

Several versions of the GeoModel detector description can be produced; a persistent detector version tag is assigned to each specific model (such as `ATLAS-GEO-21-01-00`) in order to store these versions in a relational database from which any required configuration can be retrieved. Descriptions of magnetic field configurations are stored in separate maps, whereas other detector conditions (such as alignment (see Section 5.3), calibration, and inactive detector modules) are kept in a dedicated conditions database. These conditions, some of which are evaluated during specific data-taking runs, can be applied as modifications to any compatible detector geometry layout.

During the operation of ATLAS, the detector geometry and material description is repeatedly being validated, and updated and stored in the database if needed. Alternative detector models are produced for specific studies with alternative geometries and/or different material properties. Furthermore, specific detector configurations are used for commissioning, e.g. with cosmic rays and/or dedicated magnetic field setups. Any GeoModel detector description can be converted to a raw simulation geometry that is suitable for full detector simulation with the GEANT4 toolkit, as described in Section 4.3.1.

4.2.2 Simplified Detector Models

The abstraction of a complex detector geometry to a simplified model is a commonly used technique to enable faster track finding and track fitting in the reconstruction of collision events in high-energy physics experiments (see Chapter 5), which is applied primarily due to constraints on available computing resources. Such simplified models can also be utilised for the purpose of fast simulation of particles and detector responses. However, for high-quality track reconstruction and fast simulation, it is essential that descriptions of the material budget and its spatial distribution within the detector are sufficiently accurate. Furthermore, the parameterisation of physics processes that are relevant for both tasks, such as energy loss and multiple scattering, should also enable a fast, efficient and sufficiently accurate integration of material effects along the particle trajectories. Hence a balance between two constraints is required:

[2] The full description of the ATLAS detector comprises more than 300 different materials and more than 300,000 physical volumes, or approximately 5 million total volumes. The ID is modelled with more than 200 materials and more than 50,000 physical volumes. A detailed overview of the ATLAS detector geometry as part of the ATLAS simulation infrastructure is given in Ref. [76].

1. obtaining the required degree of *simplification* of the detector geometry description and the parameterisation of physics processes, to speed up the simulation, while
2. maintaining sufficient *accuracy* of the detector geometry description and modelling of physics interactions.

This challenge comprises three components: the abstraction of the detector geometry, the parameterisation of physics processes, and the integration of material effects. All these requirements have been addressed by developing a dedicated reconstruction geometry, which is described in the following section.

4.2.2.1 Tracking Geometry

For the dual purpose of (1) track reconstruction and (2) the fast simulation of physics events, the complicated ATLAS geometry needs to be simplified while maintaining information on the exact positions, dimensions and individual alignments of the sensitive detector elements, as well as up-to-date detector conditions data. The simplified track reconstruction geometry that has been developed for this dual purpose is commonly referred to as Tracking Geometry [80]. It uses uses a simplified detector description in which the complex distribution of the detector material is mapped onto a discrete number of surfaces (layers).[3] This layout enables the track reconstruction algorithms [81] to quickly extrapolate a track from one layer to the next (within the magnetic field configuration) and to adjust the track parameters and their covariances according to the estimated material traversed along the way. In the fast simulation, the Tracking Geometry enables to simulate the passage of a particle through the entire material within a certain geometrical region, being represented on a single surface, in only one step. Two types of layers can be distinguished:

1. *Measurement layers* (representing sensitive detector elements) contain information on the integrated material that is traversed by any particle which intersects this measurement surface at a given position.
2. *Navigation layers* between these measurement surfaces facilitate the tracking by providing information about the adjacent measurement layers for any given point in the Tracking Geometry.

The same Tracking Geometry can be used for track reconstruction and fast simulation, with notable differences between these two applications. In the track reconstruction it is sufficient to estimate the average material effects, taking only the mean expected values of energy loss (including bremsstrahlung for electrons) and multiple scattering angles into account. In the fast simulation, however, a sufficiently realistic modelling and probabilistic sampling of all physics interactions is essential, and additional physics processes such as hadronic interactions and particle decays need to be considered as well, taking into account the amount and composition of traversed material. These different implementations are elucidated in Section 4.4.4.

[3] The conversion of a full GEANT4 detector geometry to a Tracking Geometry is described in Appendix B.1.1.

4.2.2.2 Inner Detector in the Tracking Geometry

In order to describe the ID in the Tracking Geometry, the sensitive detector elements of the barrel and disc layers of the Pixel detector and the SCT are directly represented by corresponding measurement layers.[4] The almost continuously distributed sensitive material in the TRT is condensed onto a few discrete layers to minimise the memory and CPU footprint. The precise material distributions of the various active (sensitive) and passive components that build the Pixel, SCT and TRT sub-detectors and support structures, as seen from the vantage point of particles traversing these components, are described by two-dimensional material maps with suitable granularity that are projected onto the measurement layers of the Tracking Geometry (see Appendix B.1.1).

The material distribution in the full detector geometry is contrasted with the position of measurement layers in the Tracking Geometry in Figure 4.4b, where hit positions of photon conversions are shown for the full GEANT4 simulation and the fast simulation with the simplified Tracking Geometry model.

4.3 Detector Simulation

This section gives a brief overview of the different available simulation flavours and strategies for ATLAS, including some details on the specific implementations of physics processes to describe particle–matter interactions. A more detailed description of the available simulation flavours and strategies for ATLAS can be found in [74, 82].

4.3.1 Full Simulation: GEANT4

The default simulation strategy of ATLAS is based on the GEANT4 particle simulation toolkit [83] and uses a highly detailed detector description. GEANT4 implements detailed models of particle–matter interactions and other physics processes, and allows to simulate particle transportation through a defined geometry and magnetic field. The detector geometry itself is constructed in the GeoModel format and converted into the GEANT4 format. Each single particle–matter interaction is computed according to the relevant physics models that are typically chosen in the form of physics lists. At the time of the main studies presented in this thesis, the two leading models were: (1) `QGSP_BERT`, which includes the Bertini intra-nuclear cascade model (BERT) [84, 85] for hadron–nucleus interactions of protons, neutrons, pions and kaons below 10 GeV, and the Quark–Gluon String Precompound (QGSP) model at higher energies; and (2) `FTFP_BERT`, which uses the FTF model, based on the FRITIOF description of string excitation and fragmentation [86], instead of the QGSP model. In these two GEANT4 physics models the main particle–matter interactions are implemented as described in Ref. [87]:

[4] Details about the implementation of the other main ATLAS detector components, the Calorimeter and Muon Spectrometer, in the Tracking Geometry can be found in [80].

- *Energy loss* processes (see Section 2.7.1) are implemented by a common description for electrons, muons and charged hadrons. The mean energy loss per unit length after a simulation step is evaluated with the Bethe–Bloch formula (see Eq. 2.4) including additional shell and higher-order corrections. The distribution of the total continuous energy loss takes the form of a straggling function [88], which approaches a Gaussian distribution in thick materials and the Landau distribution in thin materials. The contributions due to excitation and ionisation are sampled using Poisson distributions for the number of individual excitations and ionisations, respectively.
- *Bremsstrahlung* cross-sections (see Section 2.7.1.2) are based on the Berger–Seltzer model, in which the differential cross-section is treated as the sum of cross-sections in the fields of the screened nuclei and the electrons of the traversed material, respectively. Extensive tables of these cross-sections have been published e.g. in Refs. [89] and [90]. For electrons with energies above 1 GeV, additional corrections such as the Landau–Pomeranchuk–Migdal and Ter–Mikaelian effects are taken into account.
- *Pair production* due to high-energy photon conversions into e^+e^- pairs (see Section 2.7.1.3) is implemented via a corrected Bethe–Heitler formula [91] dependent on several orders of Z, which is corrected for various effects, including a Coulomb correction function as well as screening functions. The generation of $\mu^+\mu^-$ pairs is also implemented.
- *Multiple scattering* (see Section 2.7.2) is implemented for electrons, muons and hadrons, using a condensed simulation algorithm that determines the angular and spatial distributions after each simulation step via model functions giving the same moments as the Lewis theory [92]. A path length correction is applied to transform the apparent geometrical path length into the (longer) true path length of a multiply scattered particle. The angular deflection is sampled according to a model function based on a slightly modified Highland formula. For the lateral displacement and correlation only the mean values are used.
- *Hadronic interactions* (see Section 2.7.3) can be implemented with various models and parameterisations, including the Bertini and Fritiof models. Typically the first hadron–nucleon collision is simulated in great detail. The subsequent intra-nuclear cascade results in a highly excited nucleus which can decay by fission or pre-equilibrium emission, and a number of secondary particles escaping the nucleus. Elastic hadron–nucleus interactions can be modelled with the Glauber approach [93].

The GEANT4 toolkit has been extensively validated by ATLAS and other experiments during large-scale MC production of simulated events. It has been shown to provide a highly accurate simulation thanks to its detailed description of physics processes and the precise modelling of the complicated ATLAS detector geometry. However, for the same reasons the full GEANT4 simulation inevitably also requires significant computing resources. Under the steadily growing MC simulation demands (see the previously shown Figure 4.1), it becomes impossible for many physics studies to achieve the required MC simulation statistics with this default simulation

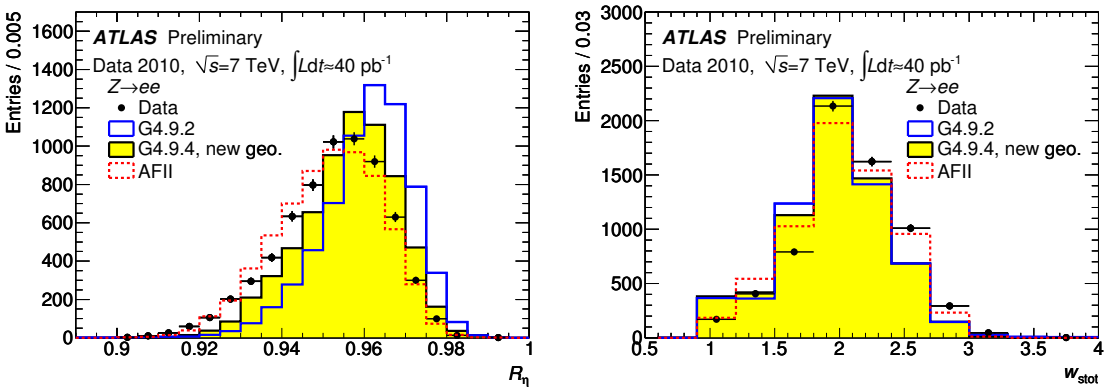


Figure 4.3: Left: deposited energy ratio R_η of particle showers in clusters of $\Delta\eta * \Delta\phi = 3 \times 7$ cells with respect to 7×7 cells in the bulk EM calorimeter layer 2. Right: shower width W_{stot} determined in a region corresponding to the cluster size in the high-granularity strip layer 1. In both validation plots the MC samples have been normalised to match the number of entries in the data. Plots taken from Ref. [96].

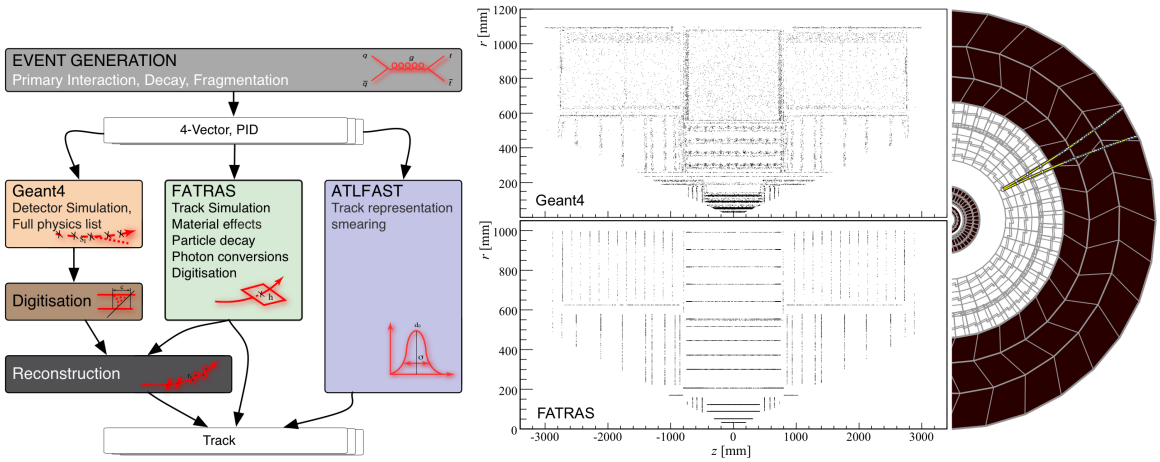
strategy. On the other hand, the highest possible level of detail for the simulation of the entire detector is not always needed. To that end, several fast simulation techniques have been developed to complement the full GEANT4 simulation, which are presented in the following.

4.3.2 Fast Simulation: Fast G4

The simulation of particle cascades (showers) traversing the complex EM calorimetry, which are mainly induced by electrons and photons, represents the slowest part of the full GEANT4 simulation, taking up almost 80% of the simulation time. The Fast G4 simulation replaces low-energy electromagnetic particles in the calorimeter with pre-simulated cascades called *frozen showers*, which were generated with the full GEANT4 simulation and are stored in memory as libraries. Using this approach, the CPU time is reduced by a factor of three in hard scattering events (e.g. $t\bar{t}$ production) with only a small impact on physics results with respect to the full simulation. The Fast G4 simulation is therefore used in the forward EM calorimeter by default for large-scale MC production in ATLAS.

4.3.3 Fast Simulation: ATLFAST-II

The full GEANT4 simulation time can be reduced by more than one order of magnitude by using the ATLFAST-II fast simulation, which employs FASTCALOSIM in the calorimeter [94, 95] and GEANT4 in the other regions of the ATLAS detector. In FASTCALOSIM the energy of single particle showers is deposited directly in the calorimeter cells via parameterisations of the longitudinal and lateral shower profiles, which are based on studies with the full GEANT4 simulation. This approach is intrinsically less accurate, but the parameterisations can be tuned against GEANT4 simulation and data. In some cases, the tuned ATLFAST-II simulation can therefore reproduce the experimental data even more closely than the full GEANT4 simulation.



(a) (b)

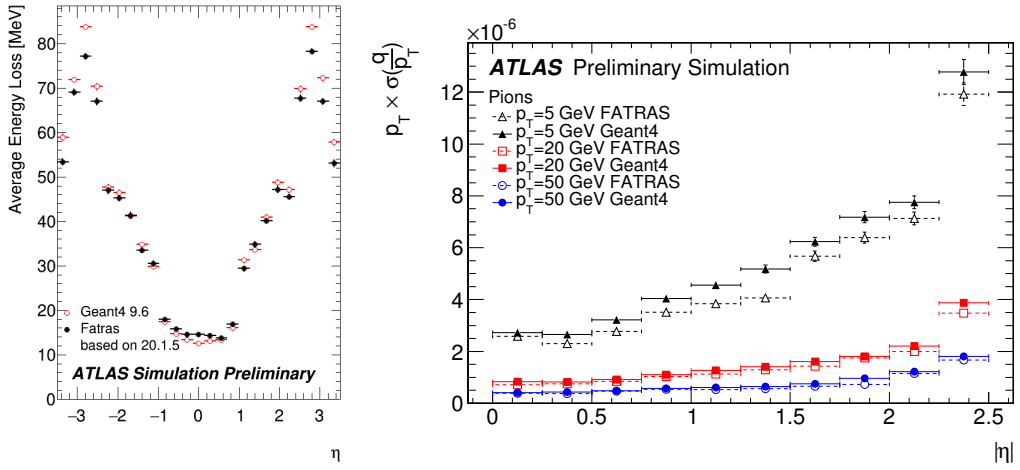
Figure 4.4: Left: overview of several simulation strategies used by ATLAS. Right: visualisation of photon conversion vertices in MC simulation, contrasting the full detector geometry used by GEANT4 (top) with the simplified geometry used by the default ATLAS track reconstruction and FATRAS (bottom). Plots taken from Ref. [97].

ATLFAST-II has been validated against GEANT4 for simulated electrons, jets and missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$). Validation plots of the calorimeter shower shapes of high-E_{T} electrons from $Z \to ee$ events are shown in Figure 4.3. In these plots, data taken in 2010 at $\sqrt{s} = 7$ TeV (black markers) are compared with GEANT4 version 9.4 (yellow histogram) and ATLFAST-II (dashed red histogram). Excellent agreement is found between the data and ATLFAST-II simulation. The result obtained with an older GEANT4 version 9.2 (blue histogram) is also shown to illustrate the impact of a GEANT4 update which is described in Section 4.4.2.

ATLFAST-II has been used since 2011 for the production of large MC samples that are needed for *new physics* searches as well as precision measurements. In the context of this thesis, the ATLFAST-II setup with FASTCALOSIM was utilised for the production of MC samples during the iterative workflow of the SCT extension efficiency study, presented in Section 6.1.4.

4.3.4 Fast Simulation: FATRAS

In the ATLFAST-IIF simulation that employs the Fast ATLAS Track Simulation (FATRAS) for the ID and MS [95, 97, 98], one order of magnitude in event simulation time can be gained in comparison to ATLFAST-II, or two orders of magnitude with respect to the full GEANT4 simulation. FATRAS provides a fast simulation by employing the same or similar algorithms and the same simplified reconstruction geometry as used in the default reconstruction (see Figure 4.4). FATRAS also uses simplified parameterisations of physics processes, which allow for faster computation of particle–matter interactions than in GEANT4. By combining FATRAS with FASTCALOSIM, the ATLFAST-IIF simulation achieves very good agreement with GEANT4-based results (see Figure 4.5).



(a) **(b)**

Figure 4.5: Comparison between FATRAS and GEANT4: (a) energy loss versus η for 1 GeV muons in the pixel and SCT detectors; (b) reconstructed transverse momentum resolution versus η for simulated pions at several values of p_{T}; dashed and solid lines show the p_{T} resolution in GEANT4- and FATRAS-simulated events, respectively. Plots taken from Refs. [95, 99].

An accurate simulation of all physics interactions by FATRAS can only be achieved with a correct description of the traversed detector material that is provided by the simplified reconstruction geometry. The default implementations of main physics interactions in FATRAS [97] can be summarised as follows:

- *Energy loss* (see Section 2.7.1) of heavy particles is implemented by sampling from the Landau distribution around a most probable value (MPV) (see Eq. 2.6) calculated for the average material that has been projected onto a thin surface, which is being traversed in one simulation step. A comparison of the energy loss calculated by GEANT4 and FATRAS is shown in Figure 4.5a.
- *Bremsstrahlung* losses for the much lighter electrons and positrons are implemented according to Eq. 2.11, by sampling from the Bethe–Heitler distribution. The modelling of high-energy photon emission due to bremsstrahlung is simplified with respect to GEANT4.
- *Pair production* due to high-energy photon conversions into e^+e^- pairs (see Section 2.7.1.3) uses parameterisations which were fitted to data, along with scaling parameters for tuning this process to GEANT4 results.
- *Multiple scattering* is implemented via a Gaussian mixture model (see Eq. 2.16 in Section 2.7.2).
- *Hadronic interactions* (see Section 2.7.3) are implemented by simplified parameterisations which are tuned to the full GEANT4 simulation. The nuclear interaction length L_0 is approximated from the radiation length X_0, the average atomic number $\langle Z \rangle$, and a scaling factor. The secondary particle multiplicity is sampled from a fit to the full simulation,

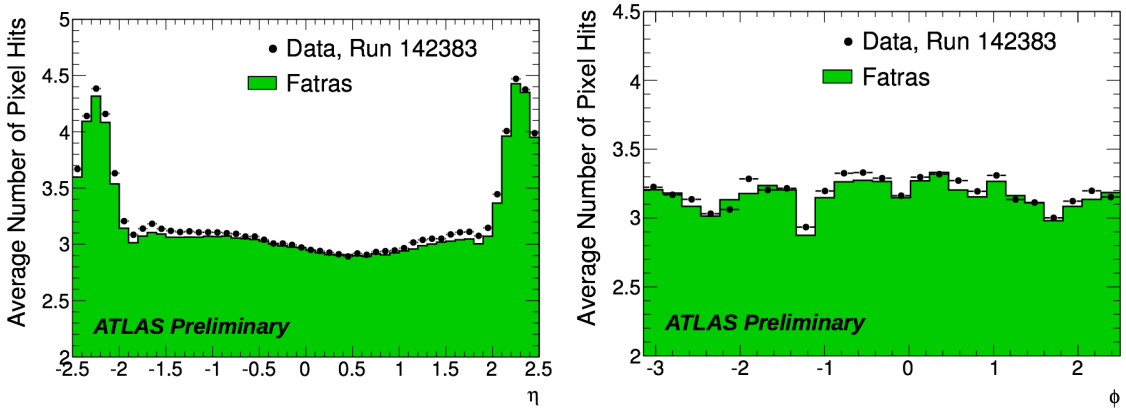


Figure 4.6: Comparison of the η (left) and ϕ (right) distributions of Pixel measurements per reconstructed track, showing good agreement between the 900 GeV data (markers) and FATRAS simulation (histograms). Plots taken from Ref. [100].

and the particle energies are determined iteratively, while their angles are assumed to be proportional to their inverse momentum. Only pions, protons and neutrons are generated according to average fractions predicted by GEANT4. Current developments aim to simulate this process directly with the more accurate GEANT4 routines.

FATRAS facilitates the fast validation, calibration and tuning cycle of the MC simulation to experimental data, including the detector geometry and material distribution, alignment of individual detector elements, and detailed detector conditions. This also makes FATRAS a suitable engine for fast large-scale MC production in various contexts and for specific studies. An example for the validation of FATRAS with $\sqrt{s} = 0.9$ TeV collision data is shown in Figure 4.6 by comparing the number of Pixel measurements per reconstructed track.

4.3.5 Integrated Simulation Framework

Many physics studies require large MC samples for which time-consuming high precision simulation is necessary only for certain particles and regions, while the remaining particles and regions require much less accuracy and could be handled sufficiently precisely by fast simulation methods. The large variety of such individual use-cases can be sensibly accommodated by the recently developed Integrated Simulation Framework (ISF) [74, 82, 101] that allows for a flexible combination of different simulation strategies.

The ISF has been designed to enable all simulation strategies within the same event, and even the same sub-detector region, for different types of particles. This principle is visualised in a use-case example of one simulated event in Figure 4.7. Simulations performed with the ISF are easily configurable for specific physics analyses and user demands. The user can achieve an optimal balance between simulation precision and execution time by selecting faster simulation flavours for all regions and particles that do not require full detail.

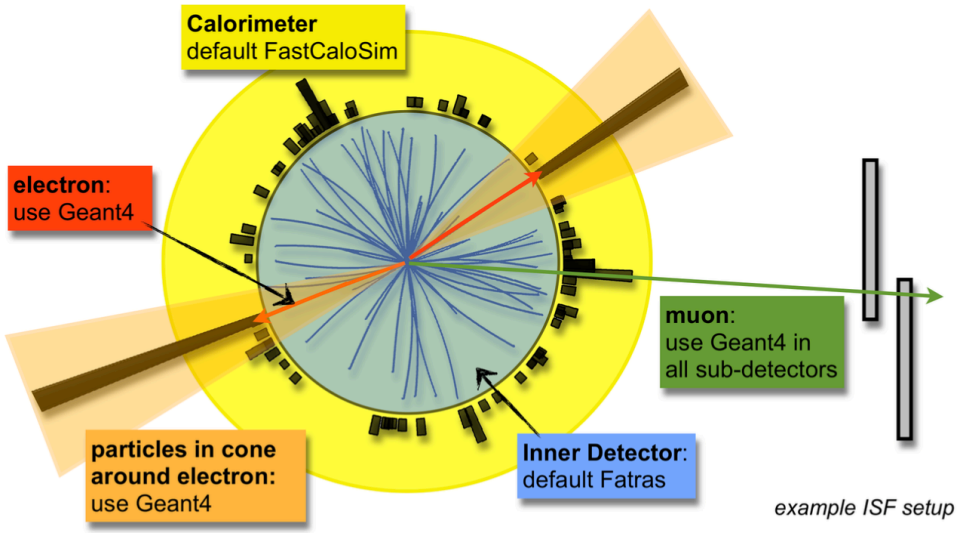


Figure 4.7: The ISF vision in a nutshell: a variety of simulation strategies are employed for specific particle types and/or detector regions within one event. Plot taken from Ref. [82].

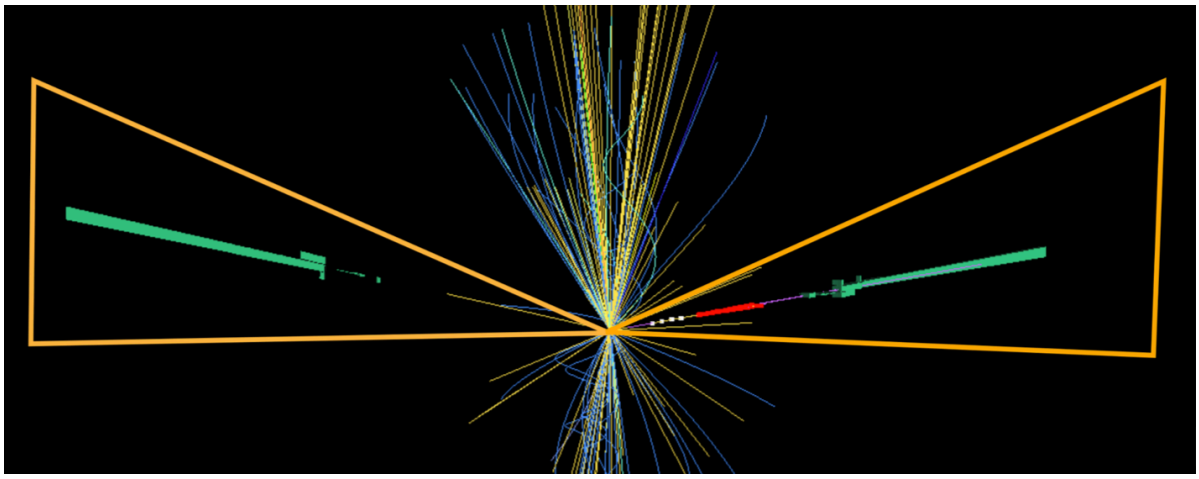


Figure 4.8: Schematic event visualisation in the $x - -y$ plane of an ISF event, which simulates only particles within narrow cones around the two photons from a $H \to \gamma\gamma$ decay, using different simulation engines (GEANT4, FASTCALOSIM and/or FATRAS). Plot taken from Ref. [102].

A typical ISF use case [102] is portrayed in Figure 4.8 for the $H \to \gamma\gamma$ analysis, which requires very high statistics for background shape studies. An optimised ISF setup for this study can be based on partial event simulation, where only the two photons from the Higgs decay and all particles in narrow cones around them are simulated, either by GEANT4 only, or by combinations of full and fast simulation engines, while the remaining particles are discarded as they do not contribute to the measurement. The fastest setup with FASTCALOSIM and FATRAS allows for a CPU time reduction of more than three orders of magnitude compared to the full GEANT4 simulation, while also significantly reducing the output size.

The ISF is fully embedded in the ATLAS software framework, and designed to be extensible in order to include newly developed simulation types as well as the application of parallel computing techniques. It has become the default framework for the large-scale production of MC simulation datasets in ATLAS since 2015.

4.4 Validation of Detector Simulation

The application of any MC simulation strategy in a particle physics experiment requires the careful validation of physics processes as well as detector models used to describe the experimental setup. In this thesis, the assumption is made that the physics processes describing particle propagation, decay, and particle–matter interactions are modelled by the full GEANT4 simulation with sufficiently high precision, such that the uncertainty of material interactions can be combined with the uncertainty of the material itself. The experimental techniques applied to validate the full detector models used by GEANT4 and to constrain their total material uncertainties require the reconstruction of charged-particle trajectories, described in Chapter 5, and are therefore deferred until Chapter 6. Here the prerequisite for charged-particle reconstruction is discussed, namely the validation of the simplified reconstruction geometry and simplified models of physics processes; these are also needed for the validation of fast simulation strategies such as FATRAS which employ these simplified models.

The simplified reconstruction geometry uses maps of the average material properties calculated for predefined detector regions that are traversed by the particles. Thus, the first required step is a verification of whether the properties of compound materials are correctly described in the full simulation. Next, it must be verified that the average material properties used in the reconstruction geometry are consistent with those in the full simulation, from the vantage point of a simulated particle traversing either of these models. This requires a proper strategy to determine these average material properties. Finally, it must be determined whether the physics processes, which are parameterised by these average material properties, are accurately described in the fast simulation, such that both simulations yield comparable physics results.

4.4.1 Toy Detector Setup and Material Averaging

The first validation step was performed with a toy detector model in GEANT4.[5] This toy detector simulates incident particles of a given energy with trajectories perpendicular to a stack of thin material volumes, with or without thin air gaps between them, and records the physics processes in the layers as well as the outgoing primary and secondary particle properties. A simple visualisation of the toy detector setup with incoming and outgoing particle trajectories is given in Figure 4.9.

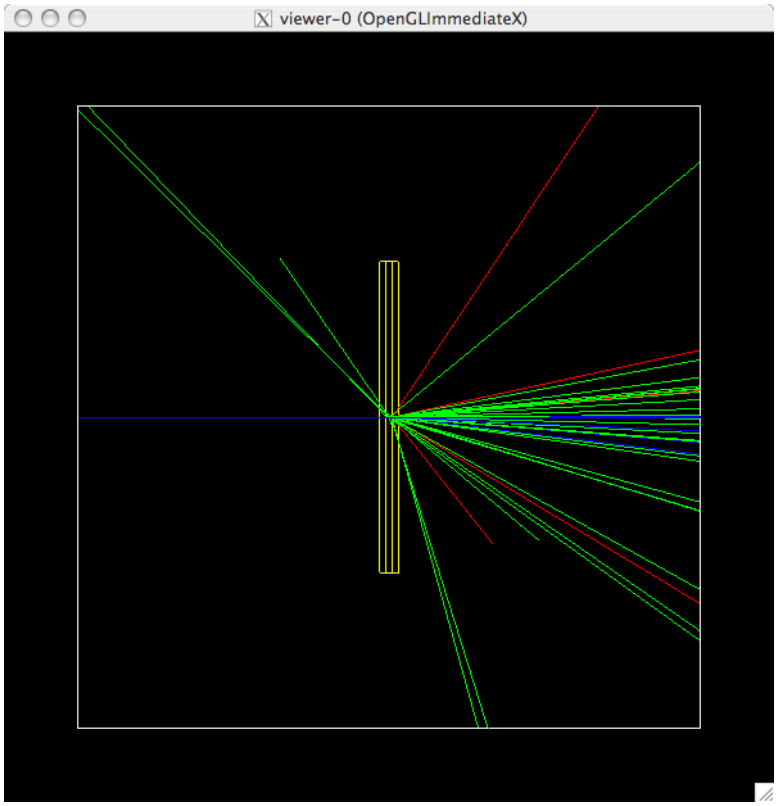


Figure 4.9: Toy Detector setup in GEANT4 and visualisation of an overlay of several events, where an incident π^- particle with $p = 100$ GeV (from left to right) undergoes a hadronic interaction in a thin Pb volume, producing a number of outgoing secondary particles. The secondary particle types are differentiated by colour.

4.4.2 Compound Materials in GEANT4

GEANT4 allows to define compound materials whose properties are calculated directly from the properties of the constituents by sophisticated internal routines. For the combination of several material layers into one compound material, several mixing strategies can be applied individually for each material property. It has been shown that $\langle Z \rangle$ and $\langle A \rangle$ are correctly weighted by the relative number of atoms [4], while the average radiation length $\langle X_0 \rangle$ (important

[5] The GEANT4 version 9.3.p01 was used for the toy detector setup and studies presented in this section.

for multiple scattering, e^+e^- pair production and bremsstrahlung) and nuclear interaction length $\langle L_0 \rangle$ (important for hadronic interactions) can be calculated by Eq. 2.10. This strategy is also followed in the production of material maps for the simplified reconstruction geometry.

In the case of energy loss, non-trivial problems arise with the calculation of $\langle I \rangle$ and $\langle \delta \rangle$ because in compounds the electrons are more tightly bound than in separate elements (see discussion in Ref. [4]). This issue can be illustrated by three possible use-case scenarios in GEANT4 involving the element C:

1. *Homogeneous molecular compound:* different elements are combined by the user to a molecular compound, such as C and Si combined to a SiC crystal structure. GEANT4 uses the mean ionisation potential $I = 78$ eV for C and $I = 173.6$ eV for Si.
2. *Heterogeneous mixture of materials:* different pure materials are combined by the user to a new material comprising a granular mixture or a stack of volumes, without creating new molecular compounds. GEANT4 uses $I = 81$ eV for C and $I = 173$ eV for Si.
3. *Separate material volumes:* the user defines a stack (or other configuration) of separate pure material volumes. GEANT4 uses $I = 81$ eV for C and $I = 173$ eV for Si.

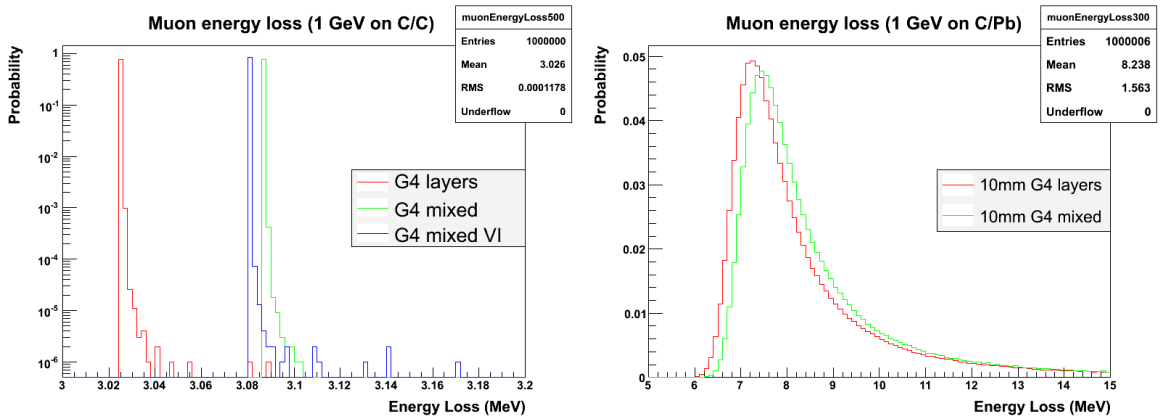


Figure 4.10: Comparison of energy loss of 1 GeV muons traversing stacked layers (red) and mixed material layers (green, blue) before the GEANT4 bug-fix. Left: two carbon layers with a total thickness of 10 mm, energy loss fluctuations turned off. Right: carbon and lead layer with a total thickness of 10 mm.

Due to a programming bug in GEANT4 prior to version 9.4, every user-defined heterogeneous mixture of materials was wrongly treated as a molecular compound. Thus even a heterogeneous mixture of two pure carbon volumes would cause GEANT4 to create a new (hypothetical) compound material with an incorrect (too small) mean ionisation potential. Using the Bethe–Bloch equation (see Eq. 2.4) this results in a too high energy loss for particles traversing the material. The bug was discovered by the author with the Toy Detector setup by comparing the energy loss of muons traversing either a stack of two separate carbon volumes or a single heterogeneous mixed C+C volume with the same thickness. The comparison is visualised in

Figure 4.10. Other combinations, such as Pb and Al, Fe and Si, etc., were also tested with similar results.

In the ATLAS Calorimeter several volumes containing stacks of material layers have been defined (in older geometries up to ATLAS-GEO-16-00-01) by heterogeneous mixed materials in the full detector description. Therefore the bug caused an incorrect mean ionisation potential for these volumes, leading to an incorrect description of the simulated e/gamma (electromagnetic) shower shapes. The discrepancy between MC and Data was recovered after a bug-fix in GEANT4 version 9.4 was implemented by the developers. In addition, from ATLAS-GEO-16-01-00 onwards the Calorimeter LAr description has been refined to reflect the detailed accordion structure. The result of these changes is illustrated in Figure 4.3 via comparison of e/gamma shower shapes. These plots clearly show the improvement between the old (GEANT4 version 9.2, blue histogram, prior to the bug-fix) and new (GEANT4 version 9.4, yellow histogram, after the bug-fix) setup.

4.4.3 Calculation of Average Material Properties

The accurate calculation of average material properties is an indispensable aspect for the creation of a simplified geometry for the purpose of particle simulation as well as track reconstruction. These average material properties affect all physics processes in the simulation and reconstruction, as summarised in Table 4.1. Various approaches were studied with the Toy Detector setup, using GEANT4 simulation with predefined materials from the internal `G4Element` and `G4Material` databases, as well as manually defined materials where the properties are obtained from alternative mixing strategies.

The internal GEANT4 corrections during the calculation of the energy loss account for differences between the results, as shown in Figure 4.11, using various strategies: a setup of separate layers made from predefined `G4Element` and `G4Material` components, a single-layer

physics process	material properties	fast simulation	track reconstruction
energy loss	$Z/A \cdot \rho$	yes	yes
multiple scattering	t/X_0	yes	yes
electromagnetic*	t/X_0	yes	*
hadronic interactions	t/L_0	yes	

Table 4.1: List of contributions of average material properties to the physics processes that are relevant in the fast simulation and track reconstruction, respectively. The variable t refers to the calculated total thickness of the traversed material from the perspective of the particle, i.e. its total path length. (*) The category "electromagnetic" encompasses bremsstrahlung as well as e^+e^- pair production; electron-induced bremsstrahlung is taken into account in the Gaussian sum filter employed by the reconstruction.

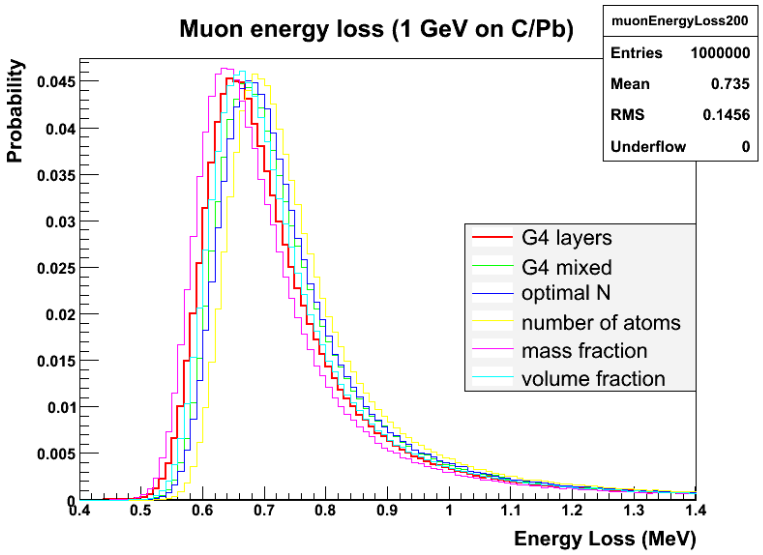


Figure 4.11: Energy loss obtained from GEANT4 simulation of 1 GeV muons traversing a thin material layer made of C and Pb, using the Toy Detector setup for a comparison of different mixing strategies for GEANT4 material components (two separate pure-material C and Pb layers of 50 % thickness each, one mixed-material layer using the GEANT4 routines) or a user-defined custom material (with compound material parameters based on the relative number of nucleons, relative number of atoms, relative mass fractions, or relative volume fraction i.e. thickness).

setup with a `G4Material` that was made from such predefined components using the internal GEANT4 routines, and single-layer setups with manually defined `G4Materials`.

The most relevant of these alternative strategies are (a) weighting by the relative number of atoms, (b) weighting by the relative mass fraction, and (c) weighting by the relative thickness. It was found that for most practical purposes and realistic scenarios the weighting by the relative step length through a material, multiplied by the relative number of atoms per unit volume, is most accurate.

4.4.4 Integration of Material Effects in the Tracking Geometry

The extrapolation of tracks through the Tracking Geometry is a fundamental component for the fast simulation with e.g. FATRAS, and also for the ATLAS track reconstruction that is described in Chapter 5. Here the simplified descriptions of physics processes for the integration of material along extrapolated tracks are described for both applications, and the differences are highlighted.

The directional deflection of stable charged particles is dominated by multiple scattering (see Section 2.7.2). The effect of non-Gaussian tails is negligible in track reconstruction since it is dominated by the loss of accuracy introduced by the simplification of the detector geometry. In the fast simulation however, a Gaussian mixture model is applied to reproduce the tails of the PDF (see Eq. 2.16).

The treatment of energy loss differs between (a) electrons and positrons, which lose a significant fraction of energy due to bremsstrahlung, and (b) other charged particles, for which ionisation losses dominate (see Section 2.7.1). These two particle types are therefore treated separately by the extrapolation. The stochastic energy losses change particle trajectories and their uncertainties in the presence of a magnetic field; this is taken into account by a deterministic mean (based on Eq. 2.4) or most probable value (MPV) (see Eq. 2.6) with a relatively small variance (which added as a Gaussian process noise). In the reconstruction, most track fitting techniques use an approximation to a Gaussian distribution around the deterministic mean instead of the Landau distribution, because the energy loss is usually small compared to the particle momentum. In the fast simulation, where an accurate sampling of the energy loss is crucial, the Landau distribution is chosen for the PDF to determine energy losses of heavy particles. For the much lighter electrons and positrons, above a certain energy threshold bremsstrahlung starts to dominate their energy loss distribution, and therefore the average total mean energy loss is used according to Eq. 2.11, as expressed by Bethe and Heitler.

Two different methods can be used for the integration of material effects during track extrapolation through the reconstruction geometry:

1. a discrete point-like integration method, which requires a purely layer-based description of the ATLAS sub-detector (e.g. Pixel and SCT detectors); here the transport of track parameters and covariances can be decoupled from their corrections due to material effects.
2. a continuous integration method, which accounts for corrections due to material effects as additional terms in the equation of motion, and relies on a volume-based description of the ATLAS sub-detector (e.g. in the Calorimeter that can be modelled approximately as one homogeneous block of material).

In both cases, the integration of material effects due to multiple scattering and energy loss rely on a correct description of the traversed detector material that is provided by the Tracking Geometry. The description can be validated by comparing the integrated material effects of particles traversing the the full geometry and the Tracking Geometry. This validation procedure is described in Appendix B.

Chapter 5

Track and Event Reconstruction in the Inner Detector

One fundamental challenge of experimental particle physics with modern detectors is that the particle trajectories and their kinematic properties can be inferred only indirectly from discrete recorded signals. This calls for reliable methods to accomplish the task of particle and event reconstruction with satisfying precision [103].

Track and vertex reconstruction are essential for the reconstruction and analysis of pp, p+Pb and Pb+Pb collision events recorded by the ATLAS detector. Vertices are defined as the common points of origin of tracks in an event, and they can thus only be found via the reconstructed tracks themselves. Track finding can be understood as a pattern recognition task with the aim to extract combinations of measurement points that are consistent with the hypothesis of particle trajectories. Track fitting comprises the task of estimating the track parameters and their uncertainties from these measurements as accurately as possible. The precise measurement of charged-particle trajectories is fundamental to the reconstruction of basic kinematic properties such as the particle momenta, to the task of particle identification, and in particular to studies of the distributions of charged particles arising from primary collisions.

This chapter describes the reconstruction of charged particles in ATLAS and evaluates the performance of the inner tracking detectors of ATLAS during Run-1, which extended from the LHC startup in late 2009 until early 2013 [104]. Section 5.1 provides basic definitions and principles of charged-particle reconstruction in ATLAS. Reconstructed tracks are used as input for the vertex reconstruction, which is presented in Section 5.2, where differences between low and high pile-up conditions are emphasised. Reconstructed tracks from di-muon events are used for an alignment procedure, which is described in Section 5.3.

5.1 Charged-Particle Reconstruction in ATLAS

Charged-particle trajectories are measured in ATLAS with two tracking detectors, the ID and MS. Only the ID is of practical relevance for the studies presented in Chapter 6 and 7 of this thesis. The ID track reconstruction consists of sequences of different algorithmic strategies [105, 106], which use local and global pattern recognition algorithms to identify the

specific detector signals, referred to as measurements or *hits* (clusters in the Pixel and SCT detector, drift circles in the TRT), that were produced by individual charged particles traversing and interacting with the sensitive detector elements. A track fitting procedure is used to estimate the compatibility of these measurements with a track hypothesis, thus forming track candidates; random signals from detector noise that are sometimes wrongly associated with tracks (called *fake hits*) can be suppressed by evaluating their impact on a fit-quality measure and discarding them if they are found to be incompatible with the fit.

It is possible that along such predicted charged-particle trajectories, a measurement would be expected in a traversed active detector module, but none can be found nor assigned to the track; such expected but missing measurements are referred to as *holes*. Since their occurrence degrades the reliability of track candidates, the reconstruction algorithms constrain the allowed number of holes per reconstructed track. *Shared measurements* occur when the same measurement signal (or cluster) is associated with several track candidates; this can be caused by close-by particles that often appear in collimated systems such as jets. A single particle can also give rise to several track candidates built from almost identical measurements; in this case the best-fitting candidate is selected by an ambiguity processor, and the other candidates are discarded. This algorithm also suppresses *fake tracks* that are built entirely from random measurement combinations and detector noise and cannot be associated with any particles. It is obviously beneficial to keep the rate of fake tracks as low as possible. The remaining tracks are ranked according to a scoring scheme based on the fit quality and the numbers of assigned measurements and holes.

The main track reconstruction sequence of ATLAS is referred to as *inside-out* track finding. Track seeds are formed by combinations of clusters in adjacent silicon detector layers. A combinatorial filter then searches for further measurements that are compatible with a road in the seed direction in order to build track candidates, which are constrained by requirements on the numbers of associated hits and holes and on the track fit quality. The precision of the estimated track parameters, which are obtained by numerical methods due to the inhomogeneous magnetic field, is limited by the intrinsic resolutions of the sub-detectors, described in Section 5.1.1, as well as the imperfect alignment of their sensitive elements, which is addressed in Section 5.3. Furthermore, the amount of material that is traversed by the particles must be accounted for, since particle–matter interactions add to the uncertainty of the track parameters.

The track reconstruction in ATLAS employs a special reconstruction geometry derived from the full detector description, called Tracking Geometry, as described in Section 4.2.2.1. Two commonly employed fitting techniques are a global χ^2 minimisation method and the progressive Kalman Filter [107]. Both techniques aim to find an unbiased representation of the initial track parameters at their vertex of origin; however, detector misalignment can introduce track parameter biases, as explained in Section 5.3. Both are least square estimators and hence assume approximately Gaussian distributions of the track parameters. The first method aims to minimise a global χ^2 function expressed by

$$\chi^2 = \sum_i \frac{[\mathbf{m}_i - h_i(f_i(\mathbf{x}_0, \{\theta_j\}_i))]^2}{\sigma_i^2} + \sum_j \frac{\theta_j^2}{\sigma_{\mathrm{MS},j}^2} \tag{5.1}$$

where i denotes the measurement layers, $\mathbf{m}_i$ describes the i-th measurement parameter vector, h_i is a measurement mapping function, f_i is a function that extrapolates the initial track state $\mathbf{x}_0$ to the i-th layer, σ_i^2 and $\sigma_{\mathrm{MS},j}^2$ are variances (the first referring to the measurements, the latter referring to the distribution of multiple scattering angles), and θ_j is the deflection angle at the j-th scattering surface. In this equation the deterministic energy loss is already included in the first term via the track extrapolation, and the stochastic contribution from multiple scattering is added by the second term, thus introducing the scattering angle θ_j as a fit parameter.

The adaptive Kalman Filter procedure progresses iteratively from measurement to measurement, while updating the knowledge about the state vector at each step. The implementation of this scheme involves three stages: (1) a prediction of the next measurement based on the current state vector, which is based on all previously included measurements; (2) a filtering step that updates the estimate of the state vector by considering the next (current) measurement as well as all previous measurements together; (3) when all measurements are included, a final smoothing step is performed that can remove outliers and also account for perturbations, in this case due to particle–matter interactions. This method can be extended to a Gaussian Sum Filter (GSF) by including non-Gaussian process noise and Gaussian measurement errors [107] that are added to the track parameter covariances.

A detailed treatise on the development of the ATLAS track reconstruction framework, including track finding, fitting and extrapolation techniques, is given in Ref. [108].

5.1.1 Track parameterisation and expected resolution

The particle trajectories are parameterised using five track parameters at the perigee (a representation of the point of closest approach to a reference point or line, such as the primary vertex or the beam line) as follows [78]:

$$\boldsymbol{\tau} = (d_0, z_0, \phi_0, \theta, q/p) \tag{5.2}$$

The impact parameters d_0 and z_0 are defined as the distances of closest approach of a track in the transverse plane and along the beam axis, respectively, measured between the perigee and a reference point or line. ϕ_0 is the azimuthal angle of the track at the perigee, and θ is its polar angle. The ratio q/p is the inverse of the particle momentum p, multiplied by its charge q.

The ID has been designed to provide transverse and longitudinal impact parameter resolutions of approximately $\sigma_{d_0} = 140\,\mu\mathrm{m}/p_\mathrm{T}\,[\mathrm{GeV}] \oplus 10\,\mu\mathrm{m}$ and $\sigma_{z_0 \cdot \sin\theta} = 209\,\mu\mathrm{m}/p_\mathrm{T}\,[\mathrm{GeV}] \oplus 91\,\mu\mathrm{m}$, respectively, for reconstructed tracks [47]. Both resolutions depend on the transverse momentum

(in units of GeV) such that a better resolution is achieved for higher-p_{T} tracks.[1] The closer a measurement surface is located to the beam interaction region, the smaller the lever arm; for this reason the pixel detector, and particularly its innermost layer, provides the measurements that have the highest influence on the impact parameter resolution.[2] The relative precision of ID measurements of transverse momenta was designed to be $\sigma_{p_{\mathrm{T}}}/p_{\mathrm{T}} = 0.05\,\% \, p_{\mathrm{T}}\,[\mathrm{GeV}] \oplus 1\,\%$.

5.1.2 Measurement efficiency of the Inner Detector

The hit efficiencies of the main ID components (barrel and endcap layers/volumes) during 2012 data-taking were measured to be $> 99\%$ for the SCT, $> 97\%$ for the Pixel detector and $> 94\%$ for the TRT [109,110]. The number of measurements-on-track in the Pixel and SCT and the distribution of reconstructed track parameters in 2012 data were well reproduced by MC simulation [3,111], adding confidence that the sensitive detector layers are correctly modelled in the simulation. Detailed comparisons of these distributions between the data and MC simulation are provided in Section 7.3.4.

The accuracy of the reconstruction of charged-particle trajectories is limited by the combined effect of the finite intrinsic resolutions of the sensitive detector elements, as well as imperfect knowledge of the locations of the sensitive detector elements, the exact magnetic field configuration, and the amount of material in the detector. The reconstruction is also affected by the occupancy of detector modules, particularly in dense environments resulting from jets.[3]

The overall track reconstruction performance (also termed *tracking performance*) is most clearly reflected in the measurements-on-track of the sub-detector components and in the track reconstruction efficiency. The track reconstruction efficiency expresses the probability that a track can be successfully reconstructed from available measurements and can also be associated with the corresponding particle that gave rise to these measurements. All these tracking performance indicators are discussed in Chapter 7, where they are shown to yield satisfying results.

5.2 Vertex Reconstruction

Vertices are defined as the locations of physics interactions which are the common points of origin of tracks in an event. They fall into two main categories:

[1] To compare these specifications with experimentally achieved resolutions in 8 TeV data from 2012 during the Run-1 data-taking period of the LHC, see Figure C.1 in Section C.1.

[2] This was one of the main motivations for installing a fourth pixel barrel layer with even smaller radius, called Insertable B-Layer (IBL), into the ID during the detector maintenance phase of 2013 and 2014. The IBL is irrelevant to the studies presented in this thesis, but it does affect the equivalent charged-particle measurements by ATLAS from 2015 onwards. To provide an outlook, a few details about the IBL performance are given in Appendix C.1.

[3] Dense environments are characterised by tracks traversing the pixel layers at distances of the order of the pixel sizes, thus leading to an increased probability of shared measurements.

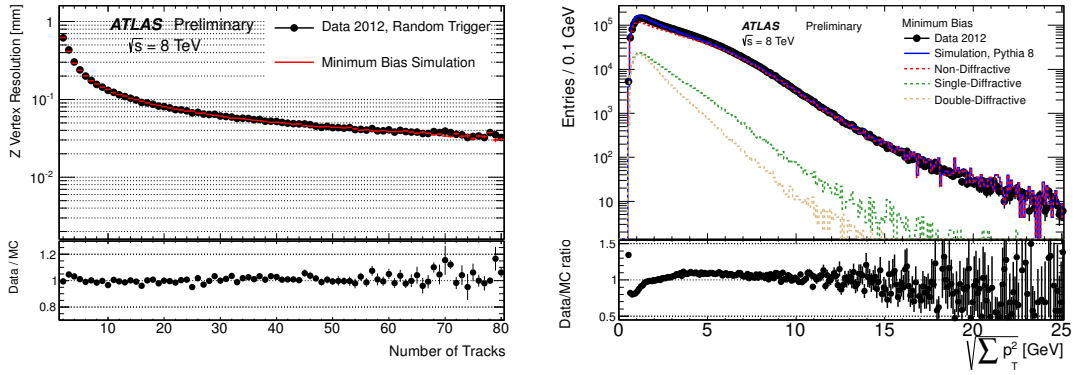
1. *Primary vertices* are found within the luminous beam interaction region, and provide information about the multiplicity and locations of the individual particle–particle collisions within the event. They can be reconstructed from at least two tracks using an iterative vertex-finding algorithm [112–115] that is briefly introduced below.
2. *Secondary vertices* arise outside the luminous region, either from decays of heavy-flavour and long-lived unstable particles, or due to interactions between particles and detector material, such as hadronic interactions and photon conversions. Two types of studies utilising the vertices from the latter categories of secondary interactions are presented in Chapter 6.

For simplicity and owing to the greater relevance in the context of this thesis, the remainder of this section will focus on primary vertices.

The vertex reconstruction algorithms [112, 116] use reconstructed tracks as an input, which are required to meet track quality criteria such as minimum numbers of silicon measurements, a minimum transverse momentum, and limits on the impact parameters d_0^{BS} and z_0^{BS} (with respect to the luminous interaction region) and their uncertainties.[4] The reconstructed positions and covariance matrices of primary vertex candidates are found by an adaptive vertex fitting algorithm [117] that uses only these selected tracks and information about the luminous beam interaction region as an input. At first a vertex seed position is estimated from a selected track and the luminous region. Then further tracks are iteratively added and weighted according to their compatibility with the current vertex position estimate, which is updated after each addition. Finally, all tracks that are incompatible with the best vertex fit (typically by $> 7\sigma$) are removed from this vertex candidate and re-used in a new search for another vertex. Valid primary vertex candidates require at least two associated tracks; good quality vertices with at least five associated tracks can be used to fit the current position and size of the luminous interaction region, which is typically evaluated every ten minutes during a run.

For minimum-bias data recorded in 2012 and the corresponding MC simulation with Pythia 8, the longitudinal vertex position resolution as a function of the number of tracks used for its reconstruction is shown in Figure 5.1a, and the $\sqrt{\sum p_{\mathrm{T}}^2}$ distribution of all tracks used for the reconstructed vertices is shown in Figure 5.1b. The intrinsic vertex position resolution is estimated with the split-vertex method, which separates the tracks associated to a vertex equally into two groups, performs new vertex fits and measures the distance between these two reconstructed vertices. Both comparisons show good agreement between MC simulation and data, thus giving an important validation of the vertex reconstruction performance (also termed *vertexing performance*) and MC modelling in general and particularly for the charged-particle measurements presented in Chapter 7. However, these data were taken in a special experimental setup with minimal pile-up conditions, which do not reflect the typical high-luminosity run

[4] Some of the specific track quality criteria that were used for the vertex reconstruction in the charged-particle measurement presented in this thesis are given in Section 7.3.1.



(a) **(b)**

Figure 5.1: Comparison of reconstructed vertices in 2012 minimum-bias data and MC simulation with Pythia 8. Left: longitudinal vertex position resolution as a function of the number of tracks used for its reconstruction. Right: $\sqrt{\sum p_{\mathrm{T}}^2}$ distribution of all tracks used for each reconstructed vertex. Plots taken from Ref. [118].

settings that are crucial for most of the physics programme of ATLAS and the LHC. Hence, the vertexing performance must generally also be assessed in events with higher pile-up.

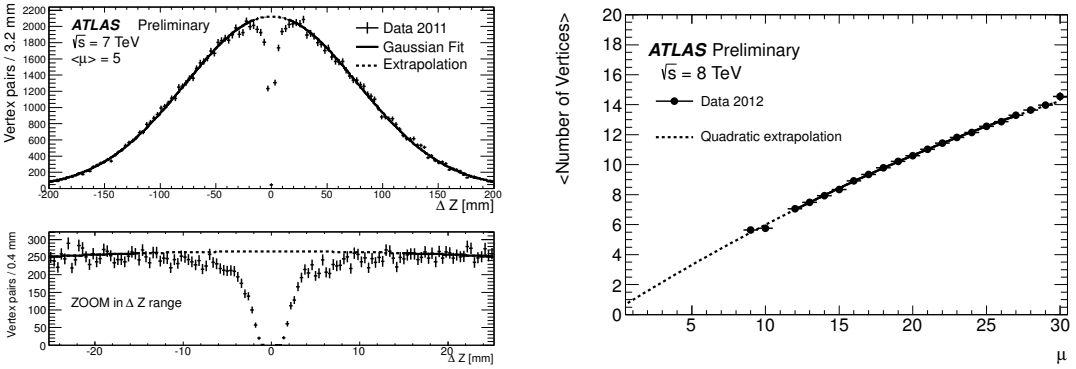
5.2.1 Vertex Reconstruction at higher pile-up conditions

During the Run-1 data taking periods of 2011 and 2012, while the LHC was operating under the design high-luminosity conditions, events with up to $\mu \approx 20$ and 40 average simultaneous pp collisions (termed *in-time pile-up*) were recorded with the ATLAS detector, respectively. Under such pile-up conditions, typically several primary vertices are reconstructed within a single event (see Figure 5.2b). Then the primary vertex with the highest $\sum p_{\mathrm{T}}^2$ is usually identified as the one belonging to a hard-scattering process, while the remaining vertices, typically arising from soft-QCD processes and therefore characterised by lower $\sum p_{\mathrm{T}}^2$, are termed *pile-up vertices*. A precise measurement of the primary vertex and identification of pile-up vertices is essential for the successful reconstruction of hard-scattering events. However, higher pile-up tends to interfere with the identification and reconstruction of vertices, and also deteriorates their position resolution.

Any reconstructed primary vertex can either originate from a true single pp interaction or fall into one of the following categories, especially under high pile-up conditions:

- *Fake primary vertices* are reconstructed mainly from fake tracks, i.e. random combinations of measurements and/or noise signals (see Section 5.1);
- *Split primary vertices* occur when single pp interactions are wrongly reconstructed as composed of two or more primary vertices [116, 119] due to resolution effects, background contaminations from non-primary tracks (e.g. due to particle–matter interactions), jets, etc.;

- *Merged primary vertices* occur when two or more close-by *pp* interactions are wrongly reconstructed as a single primary vertex, or when all tracks belonging to one particular *pp* interaction are wrongly attributed to other close-by vertices.



(a) **(b)**

Figure 5.2: (a) Longitudinal separation of primary-vertex pairs in minimum-bias events recorded in 2011 and containing $n_{\text{vrtx}} = 2$ reconstructed vertices. The difference between the Gaussian extrapolation and the measured distribution corresponds to the fraction of "lost" close-by vertices that have been merged into a single vertex. (b) Multiplicity of reconstructed primary vertices versus μ in minimum-bias data recorded in 2012. Plots taken from Ref. [113].

Vertex merging can already occur in events with only two simultaneous *pp* interactions; these constitute the vast majority of pile-up events in low-luminosity runs such as those that are used for Minimum Bias studies (see Section 7.3.1). The effect size can be studied from distributions of the longitudinal separation Δz_{vrtx} of vertex pairs in minimum-bias events containing $n_{\text{vrtx}} = 2$ reconstructed vertices. As illustrated in Figure 5.2a, a drop in the region of small longitudinal separation Δz_{vrtx} can be observed that corresponds to the fraction of "lost" close-by vertices that have been merged into a single vertex; the hypothetical true distribution would be a Gaussian, as indicated by the fit and extrapolation.

Figure 5.2b shows the relationship between the multiplicity of reconstructed primary vertices and μ in minimum-bias data recorded in 2012. With increasing pile-up, the relationship becomes non-linear primarily due to an increasing rate of merged vertices. The rate of fake vertices can be suppressed to negligible levels by using tight track selection criteria for the vertex reconstruction procedure; however, tighter constraints also reduce the vertex reconstruction efficiency. Detailed studies of the vertexing performance under high and low pile-up conditions of Run-1 are given in [112].[5]

[5] The fourth pixel barrel layer (IBL), which has been installed during LS1 at a smaller radial distance to the interaction point (see Appendix C.1), further improves the vertex resolution and vertexing performance during Run-2.

5.3 Detector Alignment

For any precise physics measurements to be performed with the ID, it must be ensured that the positions and relative alignments of all sensitive detector modules and the magnetic field are correctly described in the reconstruction. Any misalignments can degrade the resolution and potentially give rise to biases on measured track parameters. Such systematic biases can deteriorate measurements of invariant masses, and degrade beam-spot [120] and vertex reconstruction [113, 121] as well as performances of other physics measurements. An alignment procedure was therefore employed [122] to precisely determine the actual positions of all active ID modules, in order to subsequently use that knowledge in the reconstruction. The alignment campaign using 8 TeV data recorded in 2012 is also referred to as *2012 alignment*.

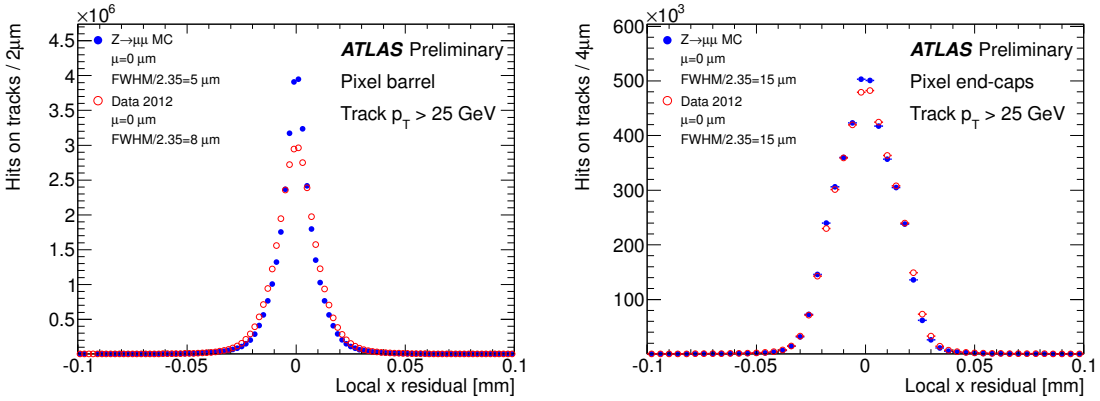


Figure 5.3: Distributions of local x residuals of all measurements in Pixel barrel (left) and end-cap modules (right), measured from $Z \to \mu\mu$ data (open circles) and MC simulation (full circles) when reconstructed with the 2012 alignment constants. In the legend, μ denotes the mean of the distributions. Plots taken from Ref. [122].

The alignment procedure was performed with a technique that uses muon tracks from $Z \to \mu\mu$ and $J/\psi \to \mu^+\mu^-$ decays to minimise track-to-hit residuals; examples are shown in Figure 5.3. The choice of muon candidates with transverse momenta of $p_{\mathrm{T}} > 25$ GeV (for $Z \to \mu\mu$) and $p_{\mathrm{T}} > 7.5$ GeV (for $J/\psi \to \mu^+\mu^-$) helps to minimise the impact of multiple scattering. Additional constraints, derived from $Z \to \mu\mu$ and $Z \to ee$ events, were introduced into the alignment to remove systematic biases from the geometry.

The procedure was applied at three levels of granularity, from main ID subcomponents (Pixel, SCT barrel and end-caps, TRT barrel and end-caps) to individual silicon modules and TRT wires with more than 700k degrees of freedom (DoF). Time-dependent alignment corrections were introduced in 2012 to correct for movements of the ID subcomponents induced by environmental changes, as shown in Figure 5.4.

After successfully applying the alignment procedure, mean values of the distribution of residuals for each individual silicon module were obtained. It was found that the RMS spread of these mean values was reduced to $< 1\,\mu$m. Systematic biases of the track parameters were determined to be less than $\delta d_0 < 1\,\mu$m for the transverse and $\delta z_0 < 10\,\mu$m for the

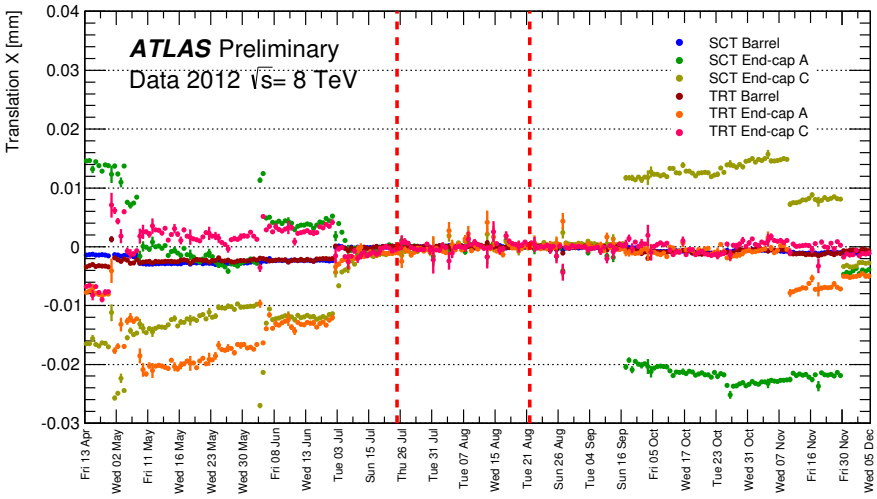


Figure 5.4: Relative shift of the global x-position of the other main ID subcomponents with respect to the pixel detector in 2012. The baseline alignment constants were determined from 30 runs with stable detector conditions during the interval between the two dashed lines. Statistical uncertainties for each run are shown by vertical error bars. Plot taken from Ref. [122].

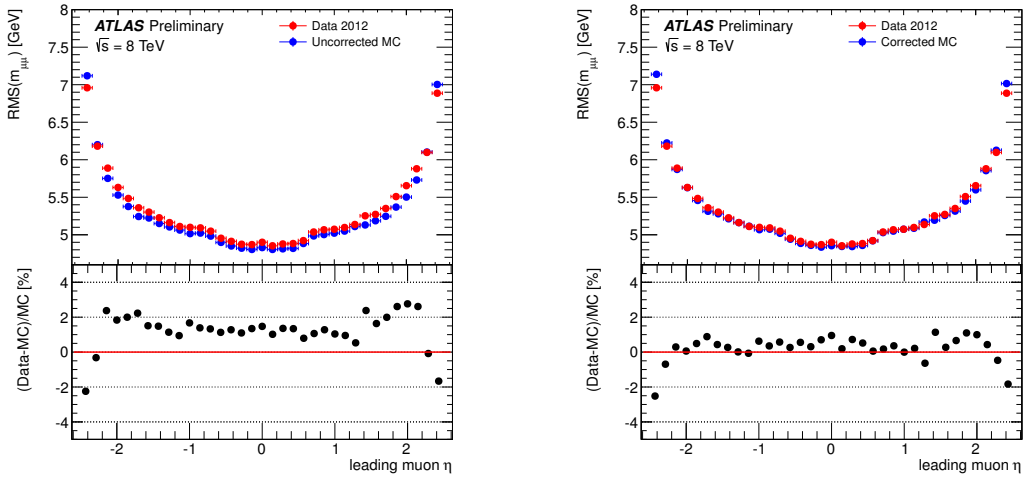


Figure 5.5: RMS of the invariant mass $m_{\mu\mu}$ distribution versus η in data (red) and MC simulation (blue), measured before (left) and after (right) the MC simulation was corrected by a smearing function. Plots taken from Ref. [122].

longitudinal impact parameter, while momentum biases are constrained to better than 0.1 %. The performance of the 2012 alignment is further demonstrated by a measurement of the invariant mass resolution of reconstructed $Z \to \mu\mu$ decays, which yields a discrepancy of less than 1 % between data and MC simulation (see Figure 5.5).

Owing to the successful 2012 alignment campaign, these results and the narrowly constrained track parameter biases demonstrate a satisfying overall ID reconstruction performance. Thereby

a crucial prerequisite to perform the main studies presented in the next two chapters of this thesis is fulfilled.

Chapter 6

Inner Detector Material Studies

In the ATLAS experiment, the ID geometry description plays a crucial role in the simulation of particle–matter interactions and detector responses, but also in the reconstruction of charged particle trajectories, which is essential for the charged-particle density measurement (see Chapter 7) and other physics analyses (e.g. in $H \to \gamma\gamma$ the placement and composition of traversed material determines the density of photon conversions, while in $H \to ZZ \to 4l$ the amount and composition of traversed material corresponds to the magnitude of energy loss of the leptons due to ionisation and bremsstrahlung). In particular, the presence of material in the ID is the main source of track reconstruction inefficiencies of hadrons as they undergo nuclear interactions with the detector components. Validation studies of the ID material were done within the ATLAS collaboration, using several techniques of which the three most important are discussed here:

1. a tag-and-probe method to measure the SCT extension efficiency,[1]
2. a study using hadronic interaction vertices, and
3. a study using photon conversion vertices.

This chapter describes these three techniques and evaluates their impact on the systematic uncertainty of the material description used in the final Run-1 geometry.

Other techniques have proved to be more challenging, such as studies of the invariant mass of K^0_S decays [123], and are not discussed here, even though they were useful in the past to provide constraints on the ID material budget uncertainty.

6.1 SCT Extension Efficiency

The SCT extension efficiency method has been introduced in 2010 [124, 125] as a technique to quantify the material located between the pixel detector and the SCT. The principle of this technique is a standard tag-and-probe measurement of the rate at which Pixel *tracklets* (track

[1] The work presented here using the SCT extension efficiency technique was performed by the author, using a custom re-implementation of the method based on the C++ programming language and the ROOT analysis toolkit. This study has given rise to further ID material studies by the ATLAS collaboration, culminating in an update of the ATLAS detector model that was eventually chosen as the final Run-1 geometry. The studies using the latter two techniques were performed independently by other members of the ATLAS collaboration.

fragments reconstructed from hit clusters in the pixel detector) are extended into the SCT by finding a combined track (tag) that corresponds to the Pixel tracklet (probe). This rate, called SCT extension efficiency, is sensitive to the amount of material in the ID, particularly in the region between the Pixel and SCT sub-detectors. These regions contain significant amounts of passive (non-sensitive) material in the form of support structures, cables and cooling systems. More traversed material in terms of nuclear interaction length, t/L_0, corresponds to a higher probability of hadronic interactions of charged particles with the material, by which the incident particles are destroyed (see Section 2.7.3). If such processes occur before a particle passes through the first SCT layers, no extension of the tracklet into the SCT can be made; in other words, the reconstruction algorithms cannot find the track, since their requirement of a minimum number of SCT measurements is not fulfilled. This type of particle-matter interaction is therefore the main source of track reconstruction inefficiencies for charged hadrons.

Evidently, an inaccurate material description in terms of nuclear interaction lengths thus leads to a systematic bias in the track reconstruction efficiency, which then propagates into measurements of minimum-bias observables such as charged-particle multiplicities. Precise knowledge of the ID material allows to constrain the systematic uncertainty on this efficiency and its contribution to the total systematic uncertainty of the charged-particle measurement. The level of knowledge of ID material between the Pixel and SCT sub-detectors can be assessed from a comparison of the SCT extension efficiencies obtained from MC simulation and data samples, respectively. A high level of agreement between MC simulation and data corresponds to an appropriate modelling of the material budget in this region (in terms of nuclear interaction lengths) in the MC simulation.[2]

The following sections describe how the SCT extension efficiency was assessed for various versions of the ATLAS detector description that were used for physics analyses at $\sqrt{s} = 8$ TeV. This study was performed with a custom software package written by the author, which implements a similar strategy as that used in previous studies. The sensitivity of the SCT extension efficiency to variations of the ID material budget was assessed using dedicated simulation samples with distorted geometries. From comparisons to the nominal geometry, an upper limit for the overall uncertainty on the amount of passive material located between the Pixel and SCT detectors was obtained.

6.1.1 Motivation and History

During Run-1 the region between the pixel detector and the SCT has been difficult to model for a variety of reasons. On one hand, this region contains no sensitive detector modules, and

[2] The SCT extension efficiency is sensitive to the material budget in terms of nuclear interaction length L_0, but not necessarily in terms of radiation length X_0 which is relevant for electromagnetic processes such as energy loss, bremsstrahlung, pair production (mainly photon conversion into e^+e^- pairs), and multiple scattering. As L_0 and X_0 are on average strongly correlated for most materials traversed by particles in the acceptance region of the ID, the results of this study can be understood to correspond also to the quality of description of the material budget in terms of X_0. However, the photon conversion studies, described in Section 6.3, provide more stringent tests of the material budget in terms of X_0.

therefore it cannot be easily probed with techniques which rely on measurement positions in active modules. On the other hand, it contains a significant amount of passive material from Pixel services that are difficult to model, especially in the Patch-Panel 0 (PP0) regions (shown in Figure 6.1) which are located in the forward regions of the inner tracker acceptance at $|\eta| > 2.2$ and within a radius (distance from the beam axis) of $174 < r < 218\,\text{mm}$, i.e. between the outermost sensitive Pixel layer and the innermost sensitive SCT layer. In these regions the detector description has been partially inaccurate since the very beginning of Run-1 due to their complicated geometry (octagonal structures, cable bundles, connectors, etc.) and some oversights in the technical documentation process.

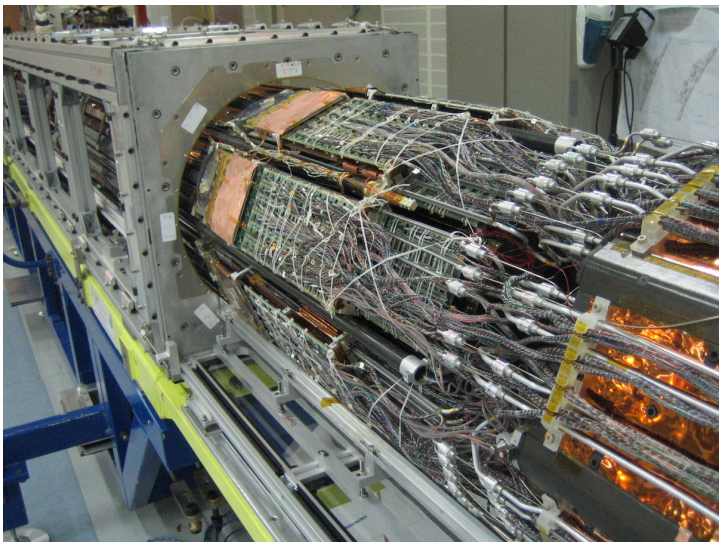


Figure 6.1: Image of part of the pixel detector, including the PP0 region, showing octagonal structures, cabling and connectors. Photograph taken in 2013 during the ID upgrade, provided by Jens Dopke.

The passive material from the PP0 region does not affect the reconstruction of track fragments within the pixel detector; however it contributes strongly to the inefficiency of reconstructing combined tracks which are built by finding extensions from such pixel tracklets into the SCT. This already posed a challenge during the previous charged-particle multiplicity measurements, in particular the Minimum Bias analysis at $\sqrt{s} = 7$ TeV performed in the year 2010. During that study, the SCT extension efficiency method was first employed [124, 125] as a highly effective method to measure discrepancies in terms of ID material budget between MC simulation and data and to constrain the systematic uncertainty on the track reconstruction efficiency. Insights obtained from this procedure led to a significant improvement of the ATLAS geometry description for MC simulation, especially in the PP0 region, which also resulted in a reduction of this dominating source of systematic uncertainty.

Between 2010 and 2012 further changes and refinements of the material description in the PP0 region (and other parts of the ID) were implemented in the detector models, but the effects of these updates were not validated again with the same method. Thus an overestimation of

the PP0 material between $2.3 < |\eta| < 2.5$ was introduced that went unnoticed until the early stages of the Minimum Bias analysis at $\sqrt{s} = 8$ TeV, where significant discrepancies between MC simulation and data were observed in this $|\eta|$ region. These discrepancies were traced back to a mis-estimated track reconstruction efficiency resulting from an inaccurate modelling of the ID material budget, and prompted a new SCT extension efficiency study with the explicit aim to improve the material description in the PP0 region. The main results of this new study are presented in this section.

6.1.2 Method Description

The SCT extension efficiency method is a tag-and-probe measurement of the rate at which a reconstructed pixel tracklet (probe) can be successfully extended into the SCT, thus forming a combined track (tag) that can be associated with the tracklet according to specific matching criteria. A dedicated low-p_{T} reconstruction setup is employed as presented in the following.

6.1.2.1 Track Reconstruction Algorithms

The default offline reconstruction setup for analyses with low-p_{T} tracks, performed within the full acceptance range $|\eta| < 2.7$ of the ID,[3] employs a sequence of two algorithms [105, 106]:

1. a standard *inside-out* algorithm with a low-p_{T} threshold of 100 MeV,
2. an additional *low-p_T* algorithm that reconstructs tracks with a transverse momentum between $100 < p_{\mathrm{T}} < 400$ MeV.

In the inside-out algorithm, candidates for combined tracks are first reconstructed from Pixel and SCT measurements and subsequently extrapolated to include measurements in the TRT, as described in Section 5.1.

The low-p_{T} track reconstruction algorithm follows a similar recipe as the inside-out algorithm, but uses only measurements that have not been assigned to tracks found by the prior inside-out algorithm. It requires fewer silicon clusters (i.e. the sum of pixel and SCT hits), but demands more pixel measurements, and allows fewer pixel holes. Thus the algorithm manages to reconstruct tracks especially from those particles which cannot reach far into the SCT due to the solenoid field and are missed by the initial algorithm. In the 8 TeV studies presented here, the low-p_{T} algorithm reconstructs around 50 % of all tracks within $100 < p_{\mathrm{T}} < 150$ MeV, thus doubling the total yield in this category; above $p_{\mathrm{T}} > 200$ MeV, its relative contribution decreases to $< 2\,\%$.[4] Further details of the track reconstruction algorithms and studies carried out to validate the low-p_{T} tracking algorithm can be found in Refs. [111] and [127].

[3] The ID is nominally designed for reconstruction within $|\eta| < 2.5$, but its acceptance extends to $|\eta| < 2.7$ mainly for combined muons (with a relaxed requirement on SCT measurements). The reconstruction is done per default in the whole ID acceptance region. It is possible (although rare) that hits compatible with $|\eta| > 2.5$ contribute to reconstructed tracks within $|\eta| < 2.5$. Furthermore, the vertex z coordinate is usually not exactly zero and can also vary slightly during a run.

[4] These fractions are similar to those found in the previous studies at $\sqrt{s} = 7$ TeV [126].

Pixel tracklets are reconstructed by a variation of the inside-out algorithm, with tighter requirements on the number of Pixel hits and holes. Detailed settings of all three track reconstruction algorithms are summarised in Table 6.1.

Criteria	**inside-out**	**low-p_{T}**	**Pixel tracklet**
min. p_{T} [MeV]	100	100	100
max. p_{T} [MeV]	–	400	–
min. Si clusters (total)	7	5	3
min. Si clusters (not shared)	6	4	3
max. Si clusters (shared)	1	1	0
min. Pixel hits	0	2	2
max. Si holes	3	2	1
max. Pixel holes	2	1	1
max. SCT holes	2	2	0
max. double holes	1	1	0

Table 6.1: Requirements for the reconstruction of combined tracks using the standard *inside-out* and the special *low-p_T* algorithm and for the reconstruction of pixel tracklets.

6.1.2.2 Data Selection and Matching Criteria

The study presented here used the same minimum-bias data and MC input samples, and the same event and track selections for combined tracks, which were used for the charged-particle measurement at $\sqrt{s} = 8$ TeV in the most inclusive phase space (see Sections 7.2 and 7.3). For pixel tracklets, the following selection criteria were chosen to ensure a high probe purity:

- a minimum transverse momentum of $p_{\mathrm{T}} > 500$ MeV,
- a transverse distance of closest approach to the primary vertex within $|d_0^{\mathrm{PV}}| < 0.2\,\mathrm{mm}$,
- a longitudinal distance of closest approach to the primary vertex within $|z_0^{\mathrm{PV}} \cdot \sin\theta| < 0.2\,\mathrm{mm}$.

These requirements for pixel tracklets are much tighter than the corresponding selection criteria for combined tracks, and were chosen to safeguard a high accuracy of the SCT extension efficiency measurement by suppressing tracks associated with non-primary particles. A matching algorithm with multiple passes was employed, which attempts to find a unique association between each selected pixel tracklet and the best-matching combined track, using the following matching criteria:

- The pixel tracklet is matched to the combined track to which it has the smallest angular distance $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$, with an upper boundary of $\Delta R_{\mathrm{max}} = 0.03$.

- At least two common silicon hits (i.e. pixel hits) are shared between the pixel tracklet and the associated combined track.
- Each pixel tracklet and combined track can only be matched once, i.e. multiple associations are not allowed.

The upper boundary ΔR_{max} for the angular distance of associated tracklet–track pairs was determined from a study of the ΔR distribution of all possible tracklet–track combinations. The tight requirement of two common silicon hits has been chosen in order to suppress track parameter biases of the associated pairs.

6.1.2.3 Track Parameter Biases

Figure 6.2 visualises the level of agreement between track parameters, as well as numbers of pixel hits, in every matched tracklet–track pair. The mean values and widths of these distributions correspond to the relative track parameter biases and resolutions, respectively.

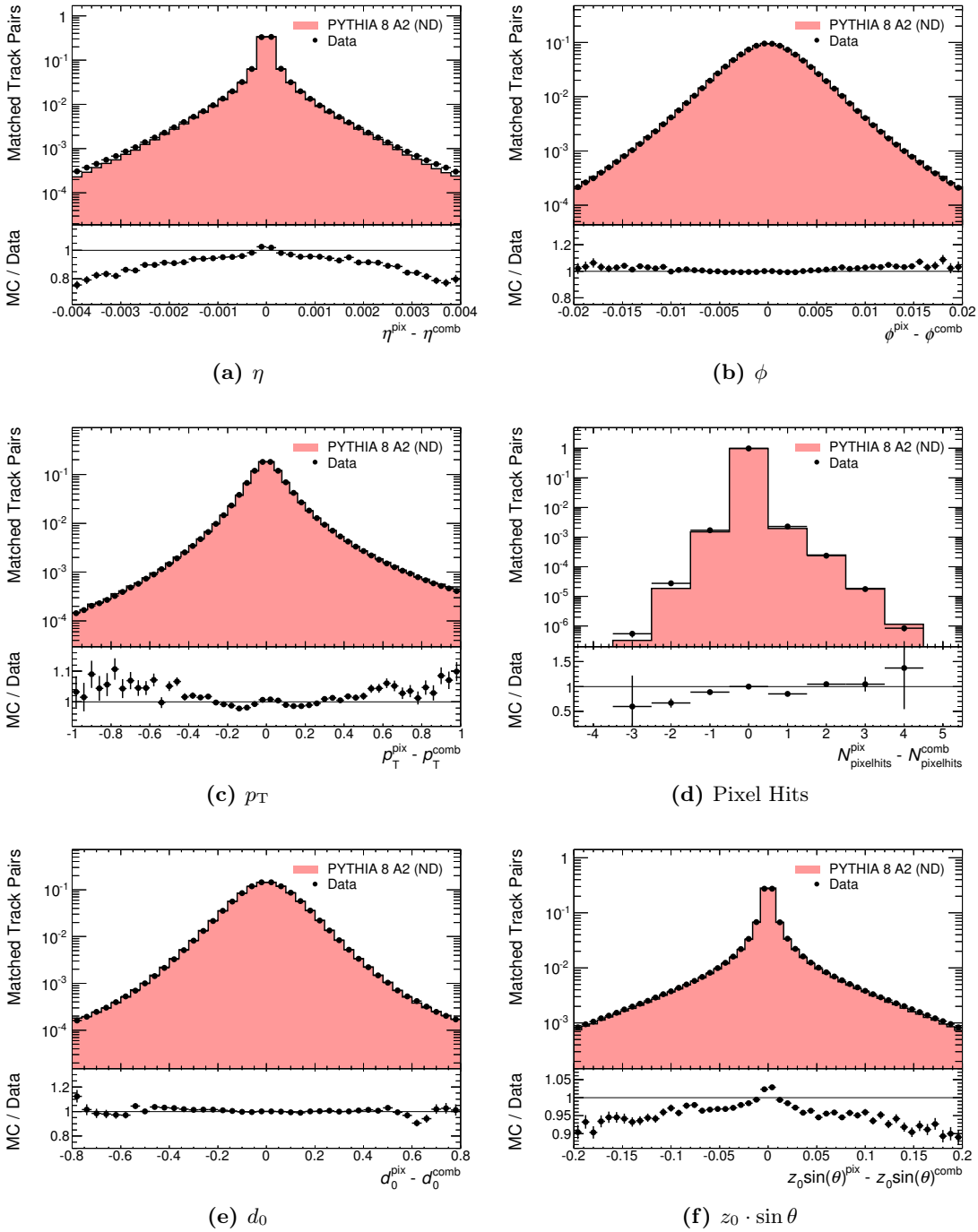
All distributions show good agreement between MC simulation and data, and no significant bias between the track parameters is found, thus validating the matching technique and indicating that the MC simulation describes these parameters and their resolutions well. A slight skewness of the p_{T} distribution (Figure 6.2c) and the pixel hits distribution (Figure 6.2d) is observed similarly in MC simulation and data. The disagreement in the statistically non-significant tails of the pixel hits distribution is expected to result from discrepancies of inactive pixel sensors in MC simulation and data, and can therefore be neglected.

6.1.3 Preliminary Studies

During the early stages of the new Minimum Bias analysis at $\sqrt{s} = 8$ TeV, while using the ATLAS-GEO-20-00-01 geometry model, in the primary-charged-particle pseudorapidity distribution $1/N_{\text{ev}} \cdot \mathrm{d}N_{\text{ch}}/\mathrm{d}\eta$ significant MC/data discrepancies were observed at large $|\eta|$. Figure 6.3a shows a clearly unphysical rise of the data within $2.3 < |\eta| < 2.5$ in this distribution. These shape anomalies were attributed to an inaccurate description of the ID material budget in the utilised detector model, which can distort the track reconstruction efficiency that is derived from a MC simulation using that model, thus leading to the observed overcorrection at large $|\eta|$.[5] In order to test the validity of this causal hypothesis, the amount and distribution of the ID material located between the Pixel and SCT sub-detectors was probed by measuring the SCT extension efficiency.

Figure 6.3b shows the SCT extension efficiency in MC simulation (using the ATLAS-GEO-20-00-01 geometry model) and data as a function of the pixel tracklet parameter η. Due to particle-matter interactions and other detector effects, it is impossible to reach a perfect efficiency of 100 %; the inefficiency corresponds to the amount of material that is traversed by particles

[5] By contrast, the final results of these distributions after improving the detector model (thus using a more accurate track reconstruction efficiency correction) are illustrated in e.g. Figure 7.43a of Section 7.9.1. For details on how these distributions are made, see Section 7.6.



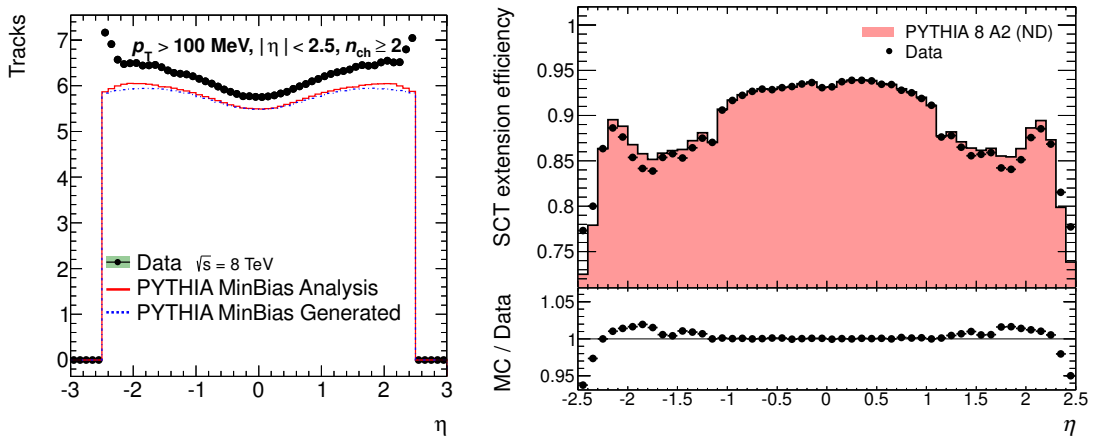
(a) η **(b)** ϕ

(c) p_T **(d)** Pixel Hits

(e) d_0 **(f)** $z_0 \cdot \sin\theta$

Figure 6.2: Comparison of the track properties between pixel tracklets and combined tracks after matching. All distributions are normalised to unity. Statistical errors are shown for the ratio between MC simulation and data.

in terms of nuclear interaction length, t/L_0. In the pseudorapidity range $2.3 < |\eta| < 2.5$, the measured inefficiency is significantly higher in MC simulation than in data, causing discrepancies of more than 6 %; this confirms the hypothesis that the detector model used for the MC



(a) (b)

Figure 6.3: Visualisation of the impact caused by an overestimation of ID material budget in the GEO-20 model of the ATLAS detector description: (a) primary-charged-particle density at $\sqrt{s} = 8$ TeV as a function of η, showing large discrepancies at $|\eta| > 2.3$; (b) SCT extension efficiency as a function of η, with large discrepancies at $|\eta| > 2.3$ due to a mis-estimation of the material budget.

simulation overestimates the material budget in these regions. The discrepancy found with this technique corresponds directly to the non-physical shape of the $1/N_{\text{ev}} \cdot \mathrm{d}N_{\text{ch}}/\mathrm{d}\eta$ distribution of the data as shown in Figure 6.3a.

The SCT extension efficiency method thus provided detailed information on localised discrepancies of the ID material description, and the corresponding detector regions were subsequently investigated further in order to identify the problematic material structures and to improve the material description in the detector models.

6.1.4 Detailed Investigation of ID Material Discrepancies

In order to determine the causes of the observed material discrepancies in the PP0 region and to improve the material description, a detailed re-assessment of the Pixel services was performed via comparisons between (a) the current ATLAS geometry model (GEO-20), (b) original construction plans, and (c) measurements of the actual physical ID in early 2013, while the ID was removed from the ATLAS cavern during the LS1 (Long Shutdown 1) phase for detector maintenance and upgrades. Several passive material structures were identified in the physical detector which had either not been described at all in the construction plans and computer models, or had been described and modelled by different shapes and/or at different locations. These structures include:

- beam pipe support system (BPSS)
- patch panels and support structures
- cooling pipe support and heater blankets

- rails and pixel roller assembly
- endcap endplates
- various cables and connectors
- other octagonal structures (which were assumed to be uniformly distributed over ϕ in the detector geometry models)

More details on these parts of the ID can be found in the ATLAS detector paper [47]. After the identification of these structures as potential sources of the observed material budget discrepancies, the geometry description was updated by the following iterative procedure:

1. a new candidate model of the ATLAS detector geometry was created by members of the Material Working Group, using the best currently available knowledge about the ID material (including insights from previous iterations);
2. MC samples of minimum-bias events using the candidate geometry model and ATLFAST-II simulation (see Section 4.3.3) were produced by the author;
3. the candidate model was validated by the author using SCT extension efficiency studies based on these new MC samples;
4. the candidate model was validated by the author with tools and techniques described in Appendices B.2 and B.3 to quantify the changes in the material budget following the proposed geometry update;
5. the findings were shared and discussed among the involved working groups and other members of the ATLAS collaboration, providing feedback for the next iteration.

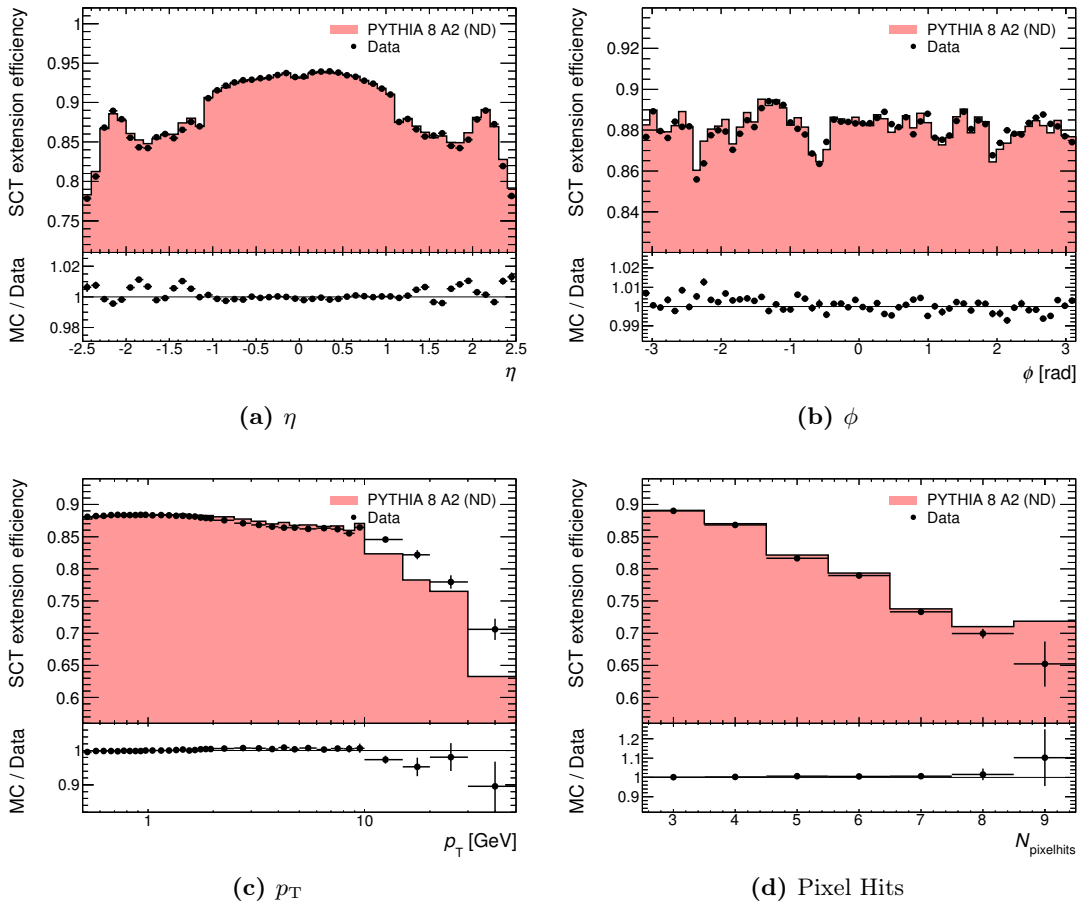
This procedure was repeated until the MC/data ratio of the SCT extension efficiency converged towards unity with residual discrepancies of less than 1.5 %. Eventually all these studies culminated in the creation and validation of a new GEO-21 model. The new geometry version for the Minimum Bias analysis at $\sqrt{s} = 8$ TeV, presented in this thesis, was chosen to be ATLAS-GEO-21-01-00.[6] Most of the changes between GEO-20 and GEO-21 were located in the PP0 region at $|\eta| > 2.2$, while smaller adjustments were also made in the region $1.4 < |\eta| < 2.0$.

6.1.5 SCT Extension Efficiency for Improved Detector Description

Figure 6.4 compares the SCT extension efficiency between MC simulation and data as a function of pixel tracklet parameters when using the new GEO-21 detector description. Here the MC/data discrepancies at large η have almost vanished, and the agreement in $1.4 < |\eta| < 2.0$ has improved as well (see Figure 6.4a). The overall level of agreement is better than 1.5 % across the entire

[6] The ATLAS-GEO-21-01-00 description exactly resembles the later defined baseline ATLAS-GEO-21-02-02 (which then became the final Run-1 geometry ATLAS-R1-2012-03-00-00) in the ID volume, but not in the Calorimeter. The Minimum Bias analysis is not sensitive to the differences between these geometries, as they have an identical impact on the reconstruction of all tracks that are used in the analysis, but other analyses which involve measurement signals from the Calorimeter should use the latter geometry version.

pseudorapidity range. These results provide good evidence that the new GEO-21 model describes the ATLAS detector more accurately than the previous GEO-20 model. This conclusion is strongly supported by the shape of the corrected η distribution of the primary-charged-particle density, shown in Figure 7.43a of Section 7.9.1.



(a) η **(b)** ϕ

(c) p_{T} **(d)** Pixel Hits

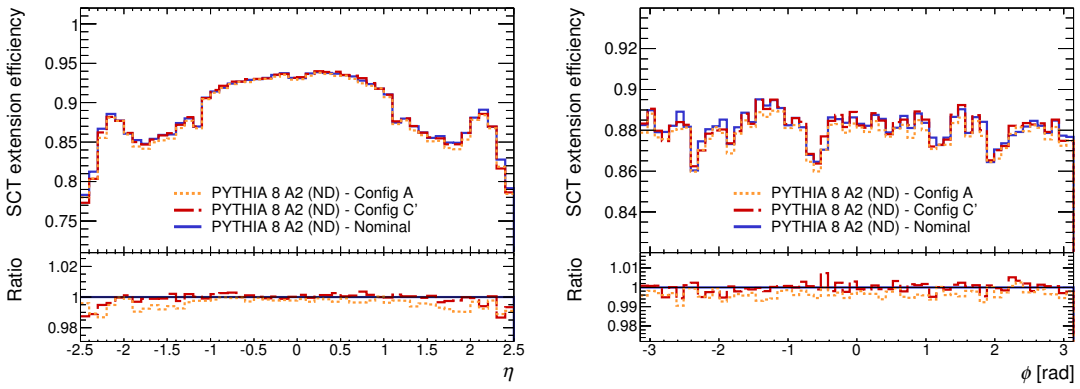
Figure 6.4: SCT extension efficiency as a function of various pixel tracklet parameters using ATLAS-GEO-21-01-00: (a) η, (b) ϕ, (c) p_{T} (for tracklets and tracks above $p_{\mathrm{T}} > 500$ MeV), (d) the number of pixel hits.

Figure 6.4b shows some discrepancies between MC simulation and data in the SCT extension efficiency as a function of ϕ, which is partly related to the track parameter resolution. As these local disagreements in ϕ are below the 1.5 % level and were found to be distributed over η, their residual effect on pseudorapidity distributions of charged-particle densities in the Minimum Bias analysis remains negligibly small. Similarly, good agreement between MC simulation and data at the 1 % level was observed in the efficiency as a function of p_{T} up to 10 GeV (see Figure 6.4c) and as a function of the number of pixel hits (see Figure 6.4d), respectively; the slightly larger discrepancies at high p_{T} arise due to momentum resolution effects as well as mis-measured high transverse momenta arising from particle-matter interactions. Given the steeply falling high-p_{T} spectrum, these effects are not statistically significant for the sake of validating the ID material

budget of the new detector geometry; their treatment in the charged-particle measurement is discussed in Section 7.5.3.2.

6.1.6 SCT Extension Efficiency for Distorted Geometries

Dedicated MC samples with additional ID material were used to determine the sensitivity of the method to variations in the detector description. In one of these samples, designated *Config A* and utilising the geometry version ATLAS-GEO-21-04-02, the density of passive ID material was scaled up by 5 % in terms of radiation length X_0; these structures span the entire pseudorapidity range. In the other sample, designated *Config C'* and employing the geometry version ATLAS-GEO-21-03-02, the density of passive material in the Pixel services was scaled up by 10 % in terms of X_0; these structures are mostly located at larger pseudorapidities beyond $|\eta| > 2$.



(a) η **(b)** ϕ

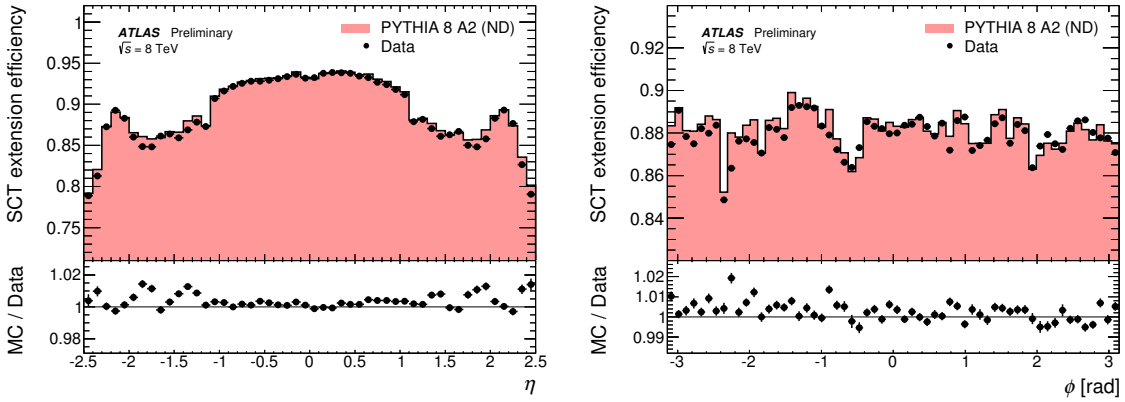
Figure 6.5: Comparison of the SCT extension efficiency in MC simulation as a function of (a) η and (b) ϕ. Samples obtained from the nominal MC simulation are compared with two samples utilising detector descriptions with extra ID material: Config A (+5 % passive ID material) and Config C' (+10 % Pixel Services).

Figure 6.5 shows the SCT extension efficiencies in MC simulation utilising the alternative detector descriptions with additional ID material, compared to the nominal MC simulation. The observed relative efficiency loss corresponds to the localised distribution of extra material as a function of η; it is more pronounced for tracks at higher $|\eta|$, as the corresponding particles traverse more additional material in terms of nuclear interaction length L_0 due to their incident angle than in the central barrel region. The difference between the nominal and distorted geometries remains within 1.5 %. For the +5 % sample, this small variation is found to be on the approximately same level as the residual discrepancy between the nominal MC geometry and data (Figure 6.4a). The result indicates that when the new GEO-21 detector description is used, the +5 % sample provides sufficient constraints for the estimation of systematic uncertainties due to the material uncertainty in the specific ID region to which this method is sensitive.

The sensitivity of the track reconstruction efficiency to the additional material is studied in more detail during the charged-particle measurement, as described in Section 7.5.3.2.

6.1.7 Stability Tests of the Method

Additional studies were performed to assess the stability of the SCT extension efficiency method and to estimate its systematic uncertainties, by studying the impact of variations of the pixel tracklet quality requirements. The MC/data ratio of the SCT extension efficiency for different p_{T} thresholds of the pixel tracklets is shown in Figure 6.6 and was found to be stable when the threshold is varied from 0.5 GeV to 1 GeV, due to the tight impact parameter cuts.



(a) η **(b)** ϕ

Figure 6.6: SCT extension efficiency as a function of pixel tracklet parameters when the transverse-momentum threshold is varied to $p_{\mathrm{T}} > 1$ GeV. The ratio of the MC simulation over data is found to be stable when compared with Figure 6.4. Plots taken from Ref. [128].

The effect of a variation of the cuts on impact parameters $|d_0^{\mathrm{PV}}|$ and $|z_0^{\mathrm{PV}} \cdot \sin\theta|$ of the pixel tracklets was also studied, and its effect on the MC/data ratio of the SCT extension efficiency was found to be negligible. The impact of a variation of the required number of common hits for the association between pixel tracklet and combined track was found to be negligible as well. While such variations change the absolute efficiency values, their relative changes were found to be approximately constant as a function of η as well as ϕ.

It can thus be concluded for the SCT extension efficiency measurement that the level of agreement between MC simulation and data remains stable when the chosen tracklet quality cuts or tracklet-track matching requirements are varied. Therefore no additional systematic uncertainties due to these selections need to be considered.

6.1.8 Summary

The SCT extension efficiency method was successfully employed to assess the impact and constrain the amount of material located between the Pixel and SCT sub-detectors. This study

revealed large discrepancies between MC simulation and data within the regions of $2.3 < |\eta| < 2.5$ that indicated a mis-modelling of the MC simulation geometry. Detailed investigations of detector descriptions, construction plans, database implementations and physical measurements of the ID during LS1 allowed to update the simulation geometry from the GEO-20 to the GEO-21 model, resulting in significantly better agreement between MC simulation and data.

For the improved ID geometry used in the GEO-21 model, it was shown that the remaining discrepancies between MC simulation and data are compatible with a material uncertainty that can be estimated using a distorted geometry with +5% additional passive material in the whole ID (designated Config A).

6.2 Hadronic Interaction Vertices

The precise material distribution in the ID, localised as well as quantified in terms of the nuclear interaction length L_0, can be probed via reconstructed secondary vertices arising from hadronic interactions of primary particles. This was studied by the ATLAS collaboration, comparing MC simulations to data from non-diffractive high-multiplicity pp collisions recorded in low-luminosity runs at $\sqrt{s} = 7$ TeV using a minimum-bias trigger [129].

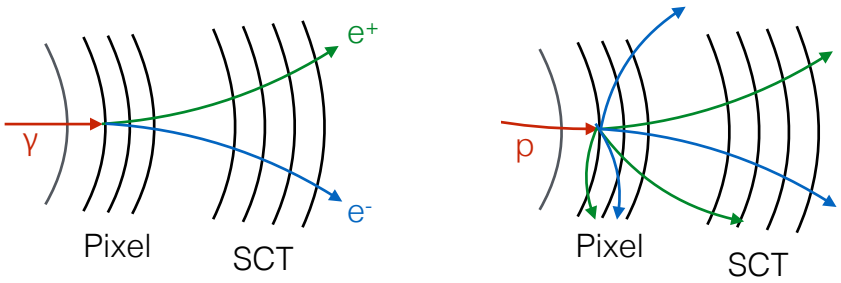
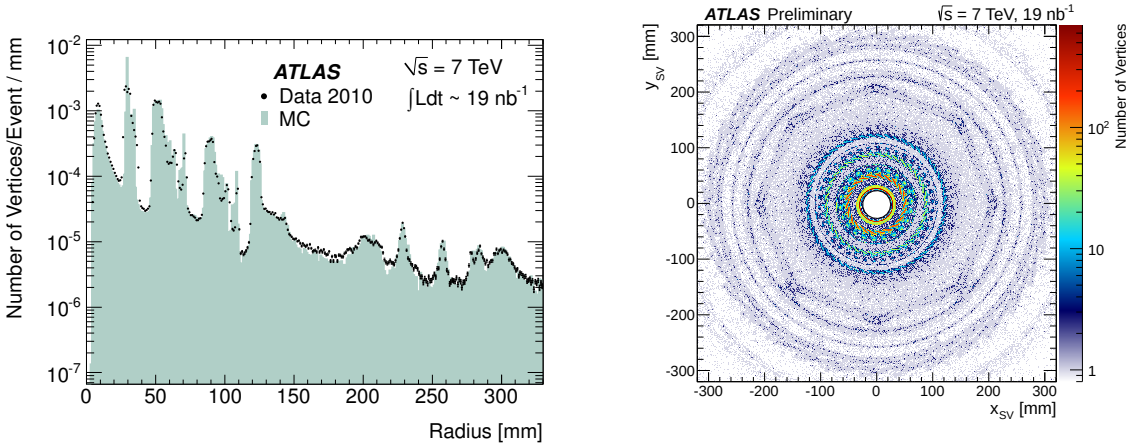


Figure 6.7: Illustration of photon conversions (left) and hadronic interactions (right) in the first Pixel layer. The outgoing secondary particles typically have markedly different opening angles, leading to different radial resolutions of the vertex positions.

This method provides an excellent resolution of vertex positions, within 0.2–1 mm in the radial (r) as well as longitudinal (z) directions, depending on the radial distance of the secondary vertex from the luminous region. The high resolution is possible due to the large opening angles between the secondary particles (illustrated in Figure 6.7) which help constrain the reconstructed vertex positions. The selection of track candidates for this study is performed by requiring a large transverse impact parameter with respect to the primary vertex, $d_0^{\mathrm{PV}} > 5\,\mathrm{mm}$, which removes $> 99\,\%$ of primary tracks and most products from K_{S}^0 decays and γ conversions, and by invariant mass vetoes on γ conversion as well as on K_{S}^0 and Λ candidates which would otherwise contaminate the sample. Only vertices within $|z^{\mathrm{SV}}| < 300\,\mathrm{mm}$ are considered in the study.



(a) (b)

Figure 6.8: (a) Radial distribution of reconstructed secondary vertices (SV) from hadronic interactions in 2010 data (points) and MC simulation (histogram). (b) (x, y)-distribution of reconstructed SV from hadronic interactions in 2010 data. In both plots the longitudinal vertex positions were constrained to $|z^{\mathrm{SV}}| < 300\,\mathrm{mm}$. Plots taken from Refs. [129, 130].

Figure 6.8a compares the radial distributions of reconstructed secondary vertices in MC simulation and data. The beam-pipe at $r \approx 34\,\mathrm{mm}$ as well as the three Pixel layers at $r \approx 50$, 88 and 122 mm are clearly visible, along with their service and support structures in between. Figure 6.8b shows the (x, y) positions of the reconstructed secondary vertices in data and also reveals a beam-pipe shift of $\Delta r \approx 2\,\mathrm{mm}$ with respect to the nominal (0,0) coordinate.

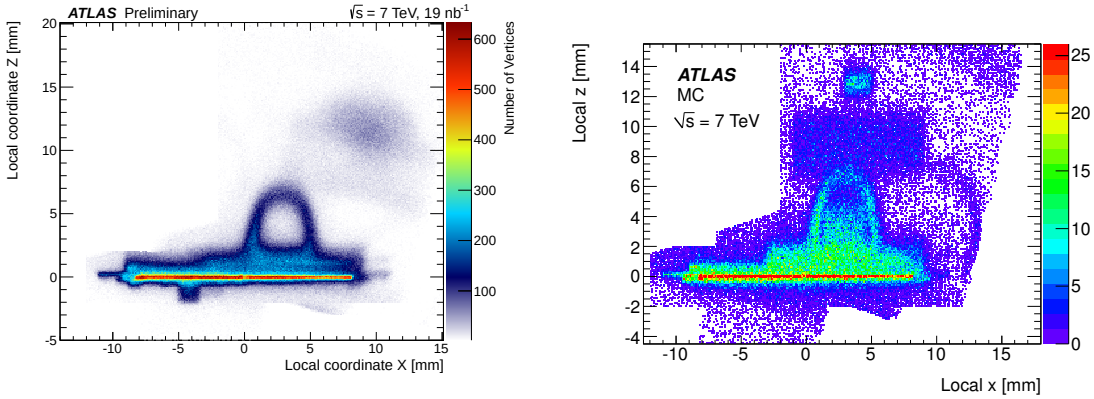


Figure 6.9: Local (x, z) positions of reconstructed secondary vertices (SV) from hadronic interactions in the innermost Pixel layer modules in 2010 data (left) and MC simulation (right), with global $|z^{\mathrm{SV}}| < 300\,\mathrm{mm}$. The hollow half-elliptical structure around $(x, z) = (3, 4)$ corresponds to the cooling pipe. Plots taken from Refs. [129, 130].

Using the same method, the complex structure and the precise distribution of material of individual pixel modules can be made visible. Figure 6.9 shows a comparison of the reconstructed vertices in data and MC simulation, with all modules within the first Pixel layer being overlaid

for better statistics. The observed discrepancies between data and MC, such as the material distribution of the mostly gaseous cooling fluid being described as a liquid in the detector model, directly point to necessary improvements of the ATLAS detector model for accurate MC simulations. These changes have already been incorporated into the detector models that were utilised in the studies presented in Section 6.1 and Chapter 7 of this thesis.

With this technique, the overall ID material description in the simulation geometry was validated to within an experimental uncertainty of about 7 % [129], while not including the systematic uncertainty due to the modelling of hadronic interactions in MC simulation.

6.3 Photon Conversion Vertices

The accuracy of the ID material description in terms of radiation length X_0 affects the reconstruction of high-energy photons and electrons that are sensitive to electromagnetic interactions such as bremsstrahlung, as well as the electromagnetic shower development (mainly due to energy loss in the ID material) [131]. The ID material budget in terms of X_0 can be probed by studying secondary vertices from photon conversions, as it is directly related to the conversion rate [132,133]. About 50 % of the photons convert into lepton pairs inside the ID before reaching the EM calorimeter.

Photon conversion candidates are found by matching oppositely-charged tracks with large fractions of high-threshold TRT hits, which is a typical transition radiation signature of high-energy electrons and positrons. The opening angles of these e^+e^- pairs are very small (illustrated in Figure 6.7), which leads to an intrinsically lower resolution of reconstructed vertex positions ($\sim 4\,\text{mm}$ in the radial direction) with respect to the technique utilising hadronic interaction vertices; energy loss effects due to bremsstrahlung further degrade the precision of the photon conversion method. A main advantage of this method is that it is also applicable at larger radii and transverse distances from the luminous region.

Measurements inside the ID are compared to those in a reference volume with accurately known material and a large rate of conversions; for this purpose the beam-pipe is chosen. The distributions of photon conversion vertices in the radial direction are shown in Figure 6.10, while a comparison in the transverse (x, y) plane is made in Figure 6.11. Both figures reveal the main detector features, with an overall good agreement of data and MC simulation. The results are comparable to those obtained with the method using hadronic interaction vertices, however the vertex resolution of photon conversions is evidently less precise. The measurement accuracy is dominated by systematic uncertainties.

In the course of these studies the description of ID support structures was improved, and good agreement in the distribution of photon conversions was found between data and MC simulation when utilising the improved models. The results of these measurements are found to be not inconsistent with an assumption of a final uncertainty of about 5% of the ID material.

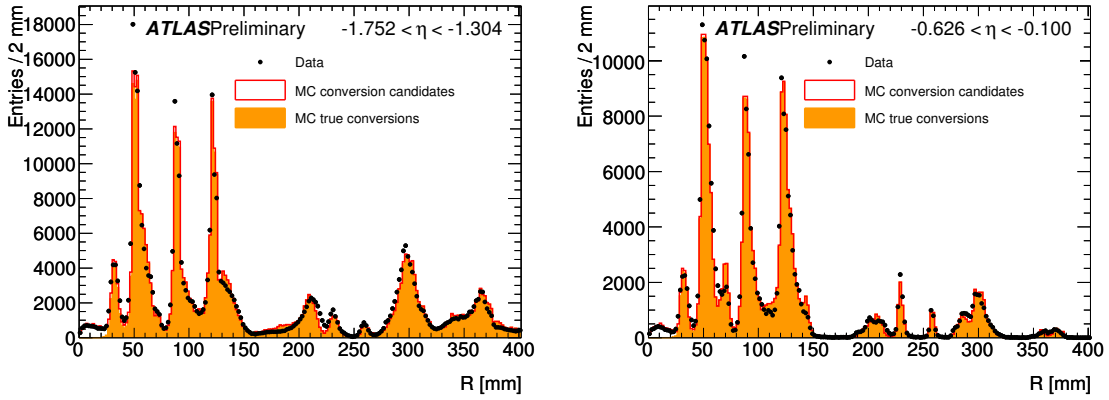


Figure 6.10: Radial distribution of the number of reconstructed photon conversion vertices in the barrel (left) and end-cap (right) sections of the ID. The beam pipe, as well as the Pixel and two SCT layers, can be identified as corresponding to sharp peaks. The distribution of the MC simulation is normalised to the data. Plots taken from Ref. [134].

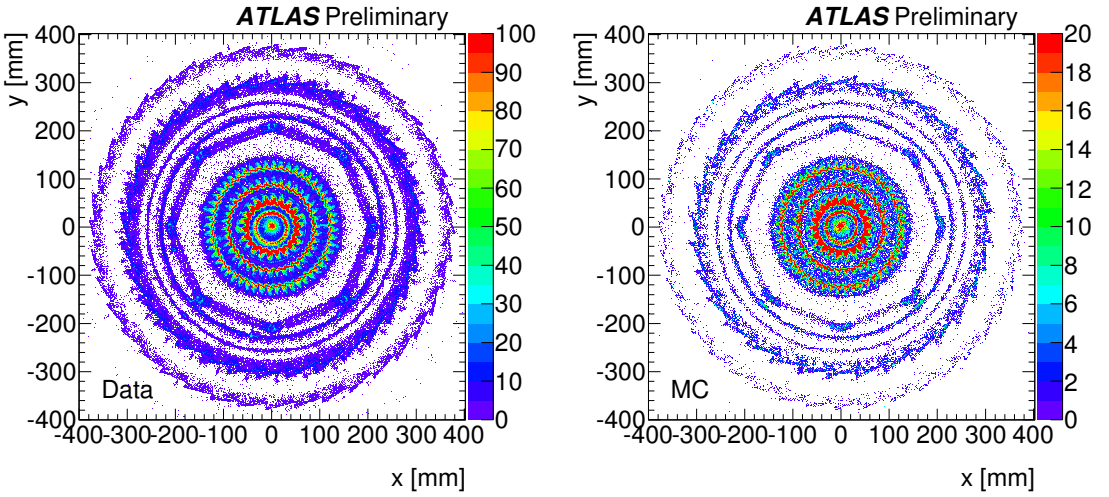


Figure 6.11: Transverse (x, y)-distribution of the number of reconstructed photon conversion vertices within $|\eta| < 1$ in data (left) and MC simulation (right). Plots taken from Ref. [134].

6.4 Summary of Material Studies

The SCT extension efficiency method was successfully employed to constrain the amount of material located between the Pixel and SCT sub-detectors. Following the detailed investigations, a new detector model for MC simulation was produced, leading to significantly better agreement between MC simulation and data.

Furthermore, by combining the results from this study with two other ID material studies that used secondary vertices from hadronic interactions and from photon conversions, it has been determined by the ATLAS collaboration that an alternative simulation geometry with +5 % additional passive material in the whole ID can be used to describe the systematics due

to the material uncertainty of the ID when using the new GEO-21 model and the final Run-1 geometry. This is a significant improvement with respect to the previous standard of using an alternative simulation geometry with +10 % additional passive material, which had been employed to constrain the ID material uncertainty for models from GEO-10 until GEO-20.

Thanks to these combined efforts, the final Run-1 detector geometry has a much higher accuracy and a smaller material uncertainty than previous ATLAS geometries, which benefits several ATLAS measurements, including the Minimum Bias analysis that is described in the next chapter.

Chapter 7

Charged-Particle Distributions in pp Collisions at $\sqrt{s} = 8$ TeV

The Minimum Bias analysis measures inclusive distributions of charged-particle densities in collider experiments such as ATLAS. This chapter describes the analysis procedure and detailed results of the charged-particle measurements at $\sqrt{s} = 8$ TeV, thus building and extending upon the related recent publications [3, 135] as well as previous equivalent studies [126, 136].[1] Comparisons are therefore mostly made with respect to the earlier Run-1 ATLAS measurements at lower centre-of-mass energies ($\sqrt{s} = 0.9$, 2.36 and 7 TeV), and references to the new Run-2 measurements at $\sqrt{s} = 13$ TeV [137,138] are only given in the most relevant contexts. Compared with the earlier ATLAS studies, the presented study covers a wider scope of measurements, uses a more robust fiducial definition of primary charged particles, and has significantly reduced systematic uncertainties, thus achieving the highest precision of all ATLAS Minimum Bias measurements of pp collisions from the Run-1 data-taking period at the LHC. This precision was achieved thanks to a significant reduction of the leading systematic uncertainty which arises due to the imperfect knowledge of the ATLAS detector material, based on the work described in the previous chapter.

7.1 Motivation, Definitions and Scope

Inclusive charged-particle measurements can probe the strong interactions particularly in the non-perturbative domain of QCD governed by soft processes with low momentum transfer, which are dominating hadron collisions at the LHC. The experimental results of such studies help to constrain and optimise the parameters in phenomenological models which are utilised to

[1] The presented analysis employs a custom analysis framework that builds upon the framework employed in previous ATLAS Minimum Bias studies at $\sqrt{s} = 0.9$, 2.36 and 7 TeV [126, 136], based on the C++ programming language and the ROOT analysis toolkit. The author took the leading role in performing and documenting this analysis, which led to publications of the obtained results in Refs. [3] and [135]. However, the presented work could not possibly have been accomplished without the dedicated contribution, ongoing support and expertise of all analysis team members (in alphabetical order: Wayne Cribbs, Deepak Kar, Oldrich Kepka, Thorsten Kuhl, Roman Lysak, Morten Medici, David Milstead, Simone Pagan-Griso, Andreas Salzburger, Edward Sarkisyan-Grinbaum, and Krzysztof Wozniak), as well as expert feedback from several other members of the ATLAS collaboration, particularly the Standard Model group conveners and Editorial Board members.

describe these soft-QCD processes and to predict the hadronic final states. The understanding of strong interactions achieved by charged-particle measurements also plays a crucial role in the prediction of soft-QCD backgrounds and topologies of events with hard-scatter and pile-up interactions, and the reduction of systematic uncertainties and biases on measurements of high-p_{T} phenomena due to these effects.

Charged-particle distributions have been previously measured in pp and $p\bar{p}$ (as well as p+Pb and Pb+Pb) collisions at a wide range of centre-of-mass energies from $\sqrt{s} = 30$ GeV up to $\sqrt{s} = 7$ TeV [126, 136, 139–156] and most recently in pp collisions at $\sqrt{s} = 8$ and 13 TeV [3, 137, 138, 157–160]. The measurements presented in this thesis [3] implement similar methodologies to those in previous and recent ATLAS Minimum Bias publications [126, 136–138], thus complementing the results at $\sqrt{s} = 0.9$, 2.36, 7 and 13 TeV. They are based on $\sim 160\,\mu\mathrm{b}^{-1}$ of data recorded by the ATLAS experiment in 2012 during the low-μ run 200805 with low LHC beam currents, using minimum-bias trigger scintillators (MBTS). This special dataset is of similar size as the datasets used in the other ATLAS minimum-bias publications.

The general properties of soft-QCD processes were studied by measuring the following primary charged-particle multiplicity and density distributions, based on particles within specified kinematic ranges:

$$\frac{1}{N_{\mathrm{ev}}} \cdot \frac{\mathrm{d}N_{\mathrm{ch}}}{\mathrm{d}\eta}, \quad \frac{1}{N_{\mathrm{ev}}} \cdot \frac{1}{2\pi p_{\mathrm{T}}} \cdot \frac{\mathrm{d}^2 N_{\mathrm{ch}}}{\mathrm{d}\eta\,\mathrm{d}p_{\mathrm{T}}}, \quad \frac{1}{N_{\mathrm{ev}}} \cdot \frac{\mathrm{d}N_{\mathrm{ev}}}{\mathrm{d}n_{\mathrm{ch}}} \quad \textit{and} \quad \langle p_{\mathrm{T}} \rangle \text{ versus } n_{\mathrm{ch}},$$

where n_{ch} is the primary-charged-particle multiplicity in a selected event, N_{ch} is the total number of primary charged particles in all selected events within the data sample, N_{ev} is the total number of selected events in the data sample, and $\langle p_{\mathrm{T}} \rangle$ is the average transverse momentum of all particles within a given multiplicity category. Selected events are events corresponding to a required minimum multiplicity of primary charged particles satisfying kinematic selection criteria, as defined in Section 7.1.2.

7.1.1 Object Definitions

A useful distinction can be made between the *particle-level* and *detector-level* objects and their respective distributions. Their corresponding constituents are called *generated particles* and *reconstructed tracks*, which can be further divided into *primary* and *non-primary* charged particles and their associated tracks. Neutral particles are not measured by the ATLAS tracking detectors and are therefore not considered in the analysis.

- *Detector-level* distributions are obtained from *tracks* that are reconstructed from the detector response signals produced by interactions between charged particles and sensitive detector elements. These distributions are subject to detector effects such as particle interactions with the active (sensitive) and passive detector material, the individual conditions of sensitive detector elements, inefficiencies of the reconstruction algorithms,

background contaminations, resolution and migration effects, as well as selection biases due to the event and track selection criteria that are applied in the analysis.

- *Particle-level* distributions of charged particles can be either (a) obtained as *generated particle-level* distributions directly from the output of MC event generators — these particles are also called *generated particles* — or (b) approximated from the detector-level distributions by applying a full correction and unfolding procedure to compensate for all selection and reconstruction inefficiencies, background contaminations and other effects listed above.
- *Primary* charged particles were previously (up to 7 TeV measurements) defined as having a mean lifetime $\tau > 30\,\text{ps}$, produced either directly as final-state particles resulting from hadron collisions, or as decay products of such final-state particles with shorter lifetimes (i.e. with $\tau < 30\,\text{ps}$). In the new analyses performed at $\sqrt{s} = 8$ and 13 TeV, primary charged particles are defined as having a mean lifetime $\tau > 300\,\text{ps}$. Charged particles with a mean lifetime between $30 < \tau < 300\,\text{ps}$ and their decay products are now excluded from this fiducial primary-charged-particle definition as they were found to introduce a significant model-dependence when included (see Section 7.4.3), and are instead treated as a source of background.
- *Non-primary* charged particles comprise particles originating from another source than the primary collision. These are mostly *secondary* particles (sometimes denoted *secondaries*) that are produced from particle–matter interactions and particle decays at a later point in time after the primary collision. Non-primary tracks comprise tracks associated with secondary particles, as well as combinatorial fake tracks which are not associated with any particles; in the data, this category can also include tracks originating from non-collision background (see Section 7.4.1). All non-primary particles and tracks are treated as a source of background (see Section 7.4.2).

7.1.2 Phase-Space Selections

Final results are given for five different phase spaces, i.e. fiducial volumes defined by kinematic selections of primary charged particles with $|\eta| < 2.5$ (within the designed acceptance region of the ID) and either $p_{\text{T}} > 100$ MeV or 500 MeV, corresponding to charged-particle multiplicities between $n_{\text{ch}} \geq 1$ and 50:

- $|\eta| < 2.5$, $p_{\text{T}} > 100$ MeV and $n_{\text{ch}} \geq 2$,
- $|\eta| < 2.5$, $p_{\text{T}} > 500$ MeV and $n_{\text{ch}} \geq 1$,
- $|\eta| < 2.5$, $p_{\text{T}} > 500$ MeV and $n_{\text{ch}} \geq 6$,
- $|\eta| < 2.5$, $p_{\text{T}} > 500$ MeV and $n_{\text{ch}} \geq 20$,
- $|\eta| < 2.5$, $p_{\text{T}} > 500$ MeV and $n_{\text{ch}} \geq 50$.

The first of these phase spaces provides the most inclusive measurement of charged-particle distributions by also including particles in the (rather challenging) low-p_{T} regime. By comparison, the second phase space is easier to analyse, and here most measurements are available at different centre-of-mass energies. The third, fourth and fifth selection apply a higher minimum n_{ch} threshold than the second phase space; this is expected to reduce the fraction of single- and double-diffractive events in the sample. The first three phase spaces allow for direct comparisons of the minimum-bias observables with earlier ATLAS measurements taken at $\sqrt{s} = 0.9$ and 7 TeV; thus they also provide important constraints to the MPI model parameters in MC tunes that describe the energy extrapolation (e.g. PARP(89) and PARP(90) in Pythia 6, ecmRef and ecmPow in Pythia 8; see also Section 2.5.1). The fourth and fifth (high-multiplicity) phase spaces are measured by ATLAS and presented in this work for the first time; these novel results provide additional constraints for the tuning of MC models.

7.1.3 Analysis Method

The final charged-particle multiplicity, pseudorapidity and transverse momentum distributions are obtained with a sequential analysis procedure that comprises the following defined steps:

1. MC samples of hadronic final states corresponding to the measured data are produced with several MC event generator models; these particle collections are propagated through a detailed detector model, and the detector response is simulated (see Section 7.2).
2. Data quality requirements including low pile-up conditions are applied to the recorded data sample, to ensure a high purity of the analysis input (see Section 7.3).
3. For the selection of event candidates from the recorded data to be included in the analysis, a Level-1 Minimum-Bias trigger is used (see Section 3.2.6; furthermore, a reconstructed primary vertex and a minimum number of reconstructed tracks with well-defined kinematic properties are required (see Section 7.3.1).
4. In addition to the event selection, various track quality selection criteria are applied to reconstructed tracks, ensuring that only good-quality tracks are included in the analysis (see Section 7.3.2). When these selections are applied to reconstructed data and MC samples, one first obtains uncorrected "raw" detector-level distributions (see Section 7.3.5).
5. The impact of event-level and track-level backgrounds is evaluated (see Section 7.4). MC and data samples are utilised to determine the fractions of non-primary particles in the data (see Section 7.4.2). Corrections for contaminations due to strange baryons, as well as for migration effects at the borders of the kinematic region in p_{T} and η, are derived from MC simulation.
6. MC and data samples are used to determine selection inefficiencies (see Section 7.5), which comprise the trigger efficiency (see Section 7.5.1), the vertex reconstruction efficiency (see Section 7.5.2) and the track reconstruction efficiency (see Section 7.5.3).

7. All corrections are applied as event- or track-level weights (see Section 7.6.1), and an unfolding procedure is employed to account for migration and resolution effects, in order to obtain corrected and unfolded primary-charged-particle distributions with minimal model dependence (see Section 7.6):
   a) as the multiplicity, pseudorapidity and transverse-momentum distributions must be normalised to a per-event level, the total number of events N_{ev} is derived first by integrating over the corrected and unfolded multiplicity distribution;
   b) subsequently the normalised multiplicity, pseudorapidity and transverse momentum distributions are derived.
8. Closure tests are performed as consistency checks to validate the correction and unfolding procedure, by applying the full correction and unfolding procedure to reconstructed (detector-level) MC samples and comparing the results with the corresponding generated (particle-level) MC distributions (see Section 7.7).
9. Total systematic uncertainties and their individual contributions to the final distributions are assessed (see Section 7.8).
10. Finally, the corrected and unfolded primary-charged-particle distributions are compared to particle-level predictions of hadronic final states that were made with various MC event generator models (see Section 7.9).

7.2 Monte Carlo Simulation

MC samples of inclusive pp collision events were produced to model the corresponding minimum-bias data at $\sqrt{s} = 8$ TeV, using various models of MC event generators with optimised parameter settings (tunes), as described in Section 2.5. This section first gives an overview of the models and datasets that were used to compare the generated hadronic final states to the corrected data. Thereafter, the datasets of generated events that were used to simulate the ATLAS detector response are described.

7.2.1 MC Samples for Particle-Level Distributions

In order to compare the corrected measurements to predictions of hadronic final states made by phenomenological models, the MC samples given in Table 7.1 were generated using the listed MC event generators and tunes. For both Pythia 8 tunes ND, SD and DD samples were produced and weighted with the relative production cross-sections predicted by the generator. The Rivet toolkit [161] was employed to produce the particle-level distributions for final plots, which are presented in Section 7.9.

Generator	Tune	PDF-set	Model	Size [events]	x-section [mb]
Pythia 8	A2	MSTW2008 LO	ND	100M	52.12
			SD	10M	12.48
			DD	10M	8.25
Pythia 8	Monash	NNPDF 2.3 LO	ND	100M	52.12
			SD	10M	12.48
			DD	10M	8.25
Epos	LHC		inclusive	100M	71.32
Qgsjet-II			inclusive	100M	
Pythia 6	Innsbruck2013	CTEQ6 L1	inclusive	100M	
Herwig++	UE-EE-5	CTEQ6	inclusive	100M	
Herwig++	UE-EE-4	CTEQ6	inclusive	100M	

Table 7.1: Sample sizes and cross-sections of MC tunes used for comparisons of particle-level distributions with minimum-bias data at $\sqrt{s} = 8$ TeV.

7.2.2 MC Samples for Simulation of Detector Response

Several datasets of generated events were used to simulate the ATLAS detector response, to derive corrections for inefficiencies and other detector effects, and to quantify the systematic uncertainties. The following MC models and tunes were employed:

- the Pythia 8 A2 tune with the `MSTW2008 LO` PDF, described in Section 2.5.1,
- the Pythia 6 AMBT2B tune with the `CTEQ6 L1` PDF, described in Section 2.5.2,
- the Epos LHC tune [42], described in Section 2.5.3.

An overview of the samples produced from these three MC models is given in Table 7.2. The samples generated with the Pythia 8 A2 tune are used as a baseline for the correction procedure of the analysis. Alternative samples with an increased amount of ID material were also produced for studies of systematic uncertainties due to imperfect knowledge of the detector material (described in Section 7.5.3.2). The Pythia 6 AMBT2B and Epos LHC samples are mainly used to derive model-dependent systematic uncertainties; this includes the different modelling and cross-sections of diffractive events in the two Pythia generators. Furthermore, the Epos LHC model is used to determine the baseline fraction of strange baryons in the data, and to calculate extrapolation factors to facilitate the comparison of final results with earlier studies (see Section 7.4.3).

Generator	Tune	PDF-set	Model	Size [events]	x-section [mb]	Settings
Pythia 8	A2	MSTW2008 LO	ND	20M	52.12	
			SD	5M	12.48	
			DD	5M	8.25	
Pythia 6	AMBT2B	CTEQ6 L1	ND	10M	49.586	
			SD	2M	13.806	
			DD	2M	9.375	
Epos	LHC		inclusive	1M	71.32	
Pythia 8	A2	MSTW2008 LO	ND	5M		config A
Pythia 8	A2	MSTW2008 LO	ND	5M		config C'
Pythia 8	A2	MSTW2008 LO	ND	10M	52.12	QGSP
			SD	2M	12.48	
			DD	2M	8.25	

Table 7.2: Sample sizes and cross-sections of MC tunes used for modelling of minimum-bias data at $\sqrt{s} = 8$ TeV and simulation of the detector response. The alternative Pythia 8 A2 samples using extra ID material are labelled as config A and config C'. The Pythia 8 A2 samples made with an alternative physics list are labelled as QGSP.

7.2.3 Detector Simulation of Minimum-Bias Events

All generated events in the MC samples described in Section 7.2.2 were processed within the ATLAS simulation framework [76] based on the Geant4 toolkit [83] to perform a detailed simulation of the detector response. The detailed detector conditions during the data-taking run were also applied in the simulation, thus accounting for inactive and inefficient regions of the measurement apparatus. The simulation samples and the recorded LHC data were then reconstructed by the same ATLAS reconstruction software framework. The obtained datasets in NTUP format were processed with a dedicated analysis framework to derive corrections and systematic uncertainties.

The ATLAS geometry used for simulation is `ATLAS-GEO-21-01-00`, which has an ID description almost identical to the final Run-1 geometry `ATLAS-R1-2012-03-00-00`, with an identical impact on the presented study. For Geant4 different physics lists can be chosen. At the time of this study, the two leading choices for high-energy physics simulations were labelled `QGSP_BERT` and `FTFP_BERT` (see Section 4.3.1). In previous measurements the `QGSP_BERT` physics list was used, but for the current studies the `FTFP_BERT` physics list was chosen due to its better modelling of hadronic interactions of anti-baryons. The differences between these two lists with respect to the most important corrections, based on the vertex reconstruction efficiency and track reconstruction efficiency, are described in Sections 7.5.2.1 and 7.5.3.2. The influence of the

choice of MC generators and physics lists on the correction for backgrounds from non-primary tracks is described in Section 7.4.2.2.

To assess the generator model dependence of the track reconstruction efficiency and track parameter resolution, MC simulation samples obtained from Pythia 6 AMBT2B and Epos LHC are compared with those obtained from Pythia 8 A2. The dependence of the simulations on the material description is assessed with alternative ATLAS detector descriptions using more passive material than in the nominal material description. Table 7.2 also summarises the MC samples used for these studies, which are based on the Pythia 8 A2 tune. Results of these tests are described later in Section 7.5.3.2.

7.2.4 Truth Matching

In addition to the reconstruction of charged-particle trajectories, one also needs to determine whether the reconstructed tracks can be associated to the generated charged particles. In the work presented here, the association between generated particles and reconstructed tracks was performed using a *cone-matching* algorithm, by determining for each particle the one specific track with the smallest angular distance $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$, while requiring the track to be within a cone of $R = 0.15$ with respect to the particle.[2] In addition, the particle trajectory must be compatible with the position of at least one pixel cluster of the track, by requiring at least one common pixel measurement for the generated particle and the associated reconstructed track. If the track-particle association fails, the reconstructed track is considered to be a *fake* track (see Section 5.1).

In MC simulation, the classification of primary and secondary particles is internally expressed in a barcode. Particles that correspond to the fiducial definition of stable primary particles are assigned a barcode number between $0 < \text{barcode} < 200000$, whereas secondary particles that are produced from decays and particle–matter interactions are assigned a barcode number ≥ 200000. In this way the track-particle association allows to classify reconstructed tracks as either primary or non-primary tracks. In the presented analysis, the only exception from this scheme is that generated particles with a lifetime between $30 < \tau < 300\,\text{ps}$, which are now excluded from the fiducial stable-primary-particle definition, are still assigned a primary barcode by the simulation; this is the reason why reconstructed tracks associated with these particles must be separately classified and excluded from the measurement. Their treatment is explained in more detail in Section 7.4.3.

7.3 Data Quality, Event and Track Selection

The 8 TeV data sample was recorded on the $5^{\text{th}}/6^{\text{th}}$ April 2012 during the special low-luminosity run 200805, using an experimental setup as described in Section 3.1.1. The average number of pp

[2] In contrast to this analysis and previous measurements, the analysis at $\sqrt{s} = 13$ TeV employs a *hit-matching* technique for the association of reconstructed tracks with generated charged particles [137]. This robust technique was not available for the current analysis due to technical reasons.

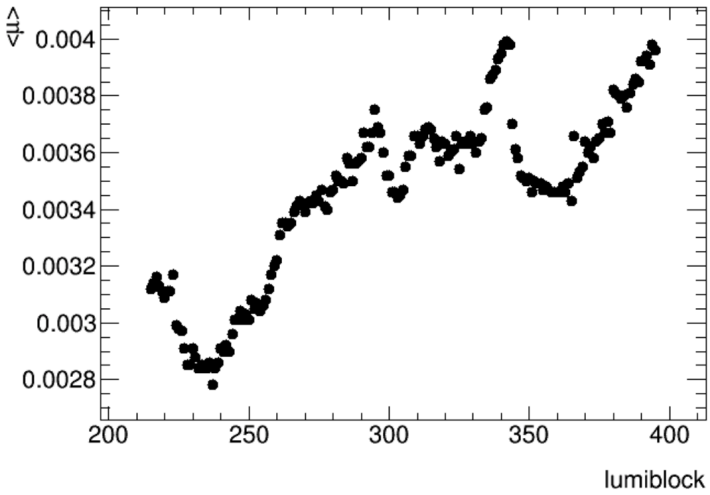


Figure 7.1: Time-dependence of the average multiplicity of pp interactions per bunch crossing, $\langle\mu\rangle$, during the run 200805. The time coordinate is expressed by the luminosity block (LB) number.

interactions $\langle\mu\rangle$ per bunch crossing was kept below $\mu < 0.004$ to minimise pile-up contaminations of the recorded sample, as shown in Figure 7.1. The recorded data sample comprises events in which proton bunches were expected to collide and the MBTS trigger recorded a signal above threshold in at least one or two of the scintillation counters. It corresponds to an integrated luminosity of approximately $160\,\mu\text{b}^{-1}$, similar to analyses at $\sqrt{s} = 7$ and 13 TeV. The possibility of a systematic effect due to the overall increase of μ during the run is discussed in Appendices A.1 and A.6.

The probability for pile-up interactions in the data sample is estimated from μ using Poisson statistics to be $\approx 0.1\,\%$, and further measures are taken to reduce this fraction in the final event selection, as described in Section 7.3.1.

7.3.1 Event Selection

A number of event selection criteria was applied to the data in order to minimise the rates of pile-up and non-collision background events, contaminations from non-primary tracks, and overall systematic uncertainties:

1. Only events that have triggered at least one (two) MBTS signals (`L1_MBTS_1` or `L1_MBTS_2`) were considered.[3]
2. A reconstructed primary vertex [116, 119] was required, made from ≥ 2 associated tracks that were compatible with the measured beam spot (see Section 5.2) and the following criteria:

[3] Pre-scale factors were applied and updated during the run in order to maintain approximately constant readout rates from the MBTS trigger streams. Thus it was also possible that events were contained in the `Mbts_2` stream but not in the `Mbts_1` stream. A combination of both triggers was therefore employed to maximise the overall data yield. The trigger efficiency correction is discussed in Section 7.5.1.

a) $p_{\mathrm{T}} > 100$ MeV,

b) ≥ 1 Pixel hit,

c) ≥ 4 SCT hits,

d) a transverse impact parameter with respect to the measured beam spot of $|d_0^{\mathrm{BS}}| < 4\,\mathrm{mm}$,[4]

e) track-fit uncertainties of $\sigma(d_0^{\mathrm{BS}}) < 5\,\mathrm{mm}$ and $\sigma(z_0^{\mathrm{BS}}) < 10\,\mathrm{mm}$.

3. Pile-up contamination was suppressed by rejecting events that had any additional vertex with ≥ 4 associated tracks satisfying the above requirements (see also Section 7.4.1.3).

4. Additional phase-space selections were applied on the kinematic variables of tracks and on their multiplicity (see Section 7.1.2), .[5]

Selection	**Events**	$\varepsilon_{\mathbf{sel,evt}}$	**Tracks**	$\varepsilon_{\mathbf{sel,trk}}$
All	11 450 084		197 713 297	
Trigger	11 112 189	96.95 %	196 170 889	99.22 %
Vertex	9 289 880	84.60 %	195 782 686	99.80 %
Pile-up	9 238 736	99.45 %	194 525 133	99.36 %
Phase space	9 166 134	99.21 %	194 460 286	99.97 %

Table 7.3: Cut-flow of the numbers of selected events and tracks in the data sample, given in the $n_{\mathrm{ch}} \geq 2$, $p_{\mathrm{T}} > 100$ MeV phase space after passing each stage of the event selection: 1) all events, 2) after trigger selection, 3) after reconstructed-vertex requirement, 4) after pile-up removal, 5) final phase-space selection. The relative efficiencies of each selection step are also quoted.

A cut-flow of the numbers of selected events and tracks in the $n_{\mathrm{ch}} \geq 2$, $p_{\mathrm{T}} > 100$ MeV phase space after passing each stage of the event selection is presented in Table 7.3.

7.3.2 Track Selection

For the measurements at $\sqrt{s} = 8$ TeV a sequence of two track reconstruction algorithms is used; details are given in Section 6.1.2.1. The analysis is performed on reconstructed data and MC simulation samples in the NTUP data format (see Section 4.1 and Ref. [78]) that contains detailed event- and track-level information. The following track selection cuts, mostly the same as in the previous analyses at $\sqrt{s} = 0.9$ and 7 TeV, are applied to tighten the quality of the tracks to be accepted in the analysis:

- $|\eta| < 2.5$,

[4] Additional information on the beam-spot conditions during the run 200805 is given in Appendix A.6.1.

[5] Note that in the $n_{\mathrm{sel}} \geq 1$, $p_{\mathrm{T}} > 500$ MeV phase space the selected events require at least one selected track with $p_{\mathrm{T}} > 500$ MeV, but they must also contain at least one other track with $p_{\mathrm{T}} > 100$ MeV to satisfy the vertex reconstruction requirement.

- $p_{\mathrm{T}} > 100$ MeV or 500 MeV depending on the phase space,
- a measurement in the innermost Pixel layer if the extrapolated track passes through a sensitive region of an operational module therein,
- ≥ 1 measurement in the Pixel detector,
- a sum of $\geq 2, 4$ or 6 hits and inactive modules per track in the SCT for tracks within $100 < p_{\mathrm{T}} \leq 200$ MeV, $200 < p_{\mathrm{T}} \leq 300$ MeV or $p_{\mathrm{T}} > 300$ MeV, respectively,[6]
- a track-fit χ^2 probability > 0.01 for tracks with $p_{\mathrm{T}} > 10$ GeV,
- a transverse and longitudinal impact parameter of $|d_0^{\mathrm{PV}}| < 1.5\,\mathrm{mm}$ and $|z_0^{\mathrm{PV}} \sin\theta| < 1.5\,\mathrm{mm}$, respectively, measured with respect to the primary vertex.

All tracks passing these selection cuts are labelled *selected tracks*. The track-fit χ^2 probability cut is applied to remove tracks with very high reconstructed $p_{\mathrm{T}}^{\mathrm{rec}}$ that has been mis-measured due to detector mis-alignment (see Section 5.3), multiple scattering or hadronic interactions. The transverse d_0^{PV} and longitudinal z_0^{PV} impact parameter requirements suppress the rate of non-primary tracks, as described in Section 7.4.2.

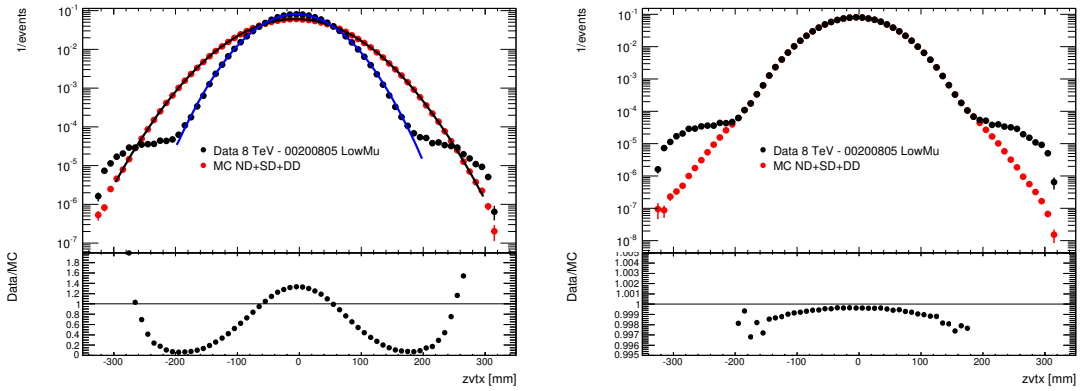
One of the main requirements for events to be considered in the analysis is that the number of selected tracks, labelled as n_{sel}, must be above the threshold of phase-space acceptance. A similar set of track selection criteria is employed to parameterise the trigger and vertex reconstruction efficiency (see Sections 7.5.1 and 7.5.2); here the two impact parameter requirements are replaced by one beam-spot related requirement, $d_0^{\mathrm{BS}} < 1.8\,\mathrm{mm}$, and the multiplicity following these criteria is labelled as $n_{\mathrm{sel}}^{\mathrm{BS}}$.

The total numbers of selected events and tracks in the recorded 8 TeV data sample are listed for each phase space in Table 7.4, and compared to available numbers of analogous ATLAS measurements at 7 and 13 TeV.

Phase Space		$\sqrt{s} = 13$ **TeV**		$\sqrt{s} = 8$ **TeV**		$\sqrt{s} = 7$ **TeV**	
$n_{\mathrm{ch}} \geq$	$p_{\mathrm{T}} >$	Events	Tracks	Events	Tracks	Events	Tracks
2	100 MeV	(n/a)	(n/a)	9 166 134	194 460 286	10 066 072	209 809 430
1	500 MeV	8 870 790	106 353 390	8 769 942	91 595 911	9 619 049	97 224 268
6	500 MeV	–	–	4 989 626	81 172 275	5 395 381	85 587 104
20	500 MeV	–	–	1 411 032	41 871 254	–	–
50	500 MeV	–	–	63 992	3 698 551	–	–

Table 7.4: Numbers of events and tracks in 2012 low-μ data at $\sqrt{s} = 8$ TeV which are selected for the analysis in the five considered phase-space regions. The numbers are compared to those from ATLAS measurements at other centre-of-mass energies, taken from Refs. [126] and [137].

[6] In order to mitigate discrepancies in the track selection due to a slightly different average number of SCT hits on track in data and MC, which is driven by a mis-modelling of inactive SCT modules, the requirement has been adjusted from the formerly used "minimum number of SCT hits" to "minimum number of a combination of SCT hits and inactive modules".



(a) **(b)**

Figure 7.2: Distribution of z-positions of primary vertices in 8 TeV data and MC simulation (a) before and (b) after reweighting the MC distribution to fit the data.

7.3.3 Vertex z reweighting of MC events

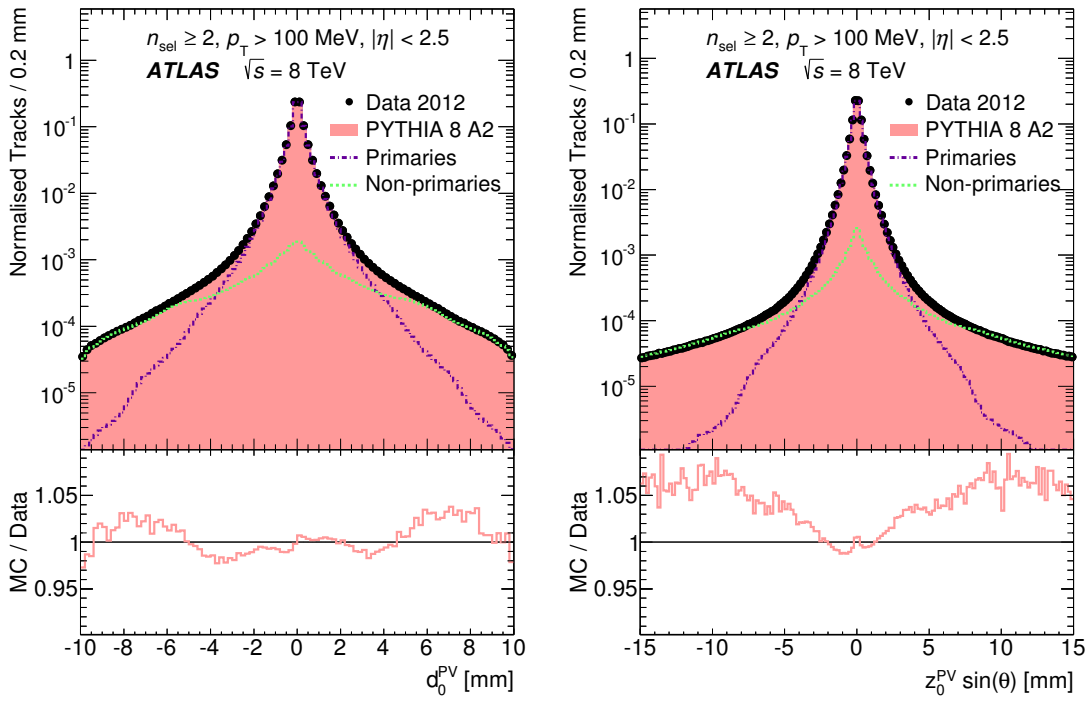
Different beamspot z profiles in data and MC samples result in different distributions of the primary vertex z position. In order to compensate for this effect, the MC sample is reweighted to data using a vertex z dependent factor, w_{zvrtx} (see Figure 7.2). This factor is applied to every MC-based track- and event-level distribution shown in this chapter, unless explicitly stated otherwise.

7.3.4 Comparison of Reconstructed Tracks in Data and MC

The reliability of any distributions and corrections that are derived from MC samples hinges upon the level of agreement between data and MC simulation in terms of detector response and reconstructed track parameters. Extensive studies have been conducted and documented for the previous Minimum Bias measurement [126] which provide a basis for the current analysis.

Figure 7.3 shows the distributions of transverse and longitudinal impact parameters of selected tracks (where only the requirement on the shown parameter is lifted), each measured with respect to the primary vertex. Model-dependent effects were suppressed by reweighting the reconstructed p_{T} spectrum of the MC sample to match the spectrum in data. Furthermore, the MC distributions have been reweighted to match the fractions of reconstructed non-primary tracks in data, using scale factors derived from template fits as described in Section 7.4.2. The normalised impact parameter distributions show good agreement between data and MC simulation. The observed discrepancies in the z_0^{PV} tails are due to the fact that the scale factors were derived from template fits of d_0^{PV} distributions.

Figure 7.4 shows the average number of measurements per reconstructed track in the Pixel and SCT sub-detectors, where the sum of hits and inactive modules per reconstructed track is used for the SCT distribution. Hits and holes are defined as in Section 5.1. Modules that were found to be inactive during data-taking were disabled during the digitisation/reconstruction

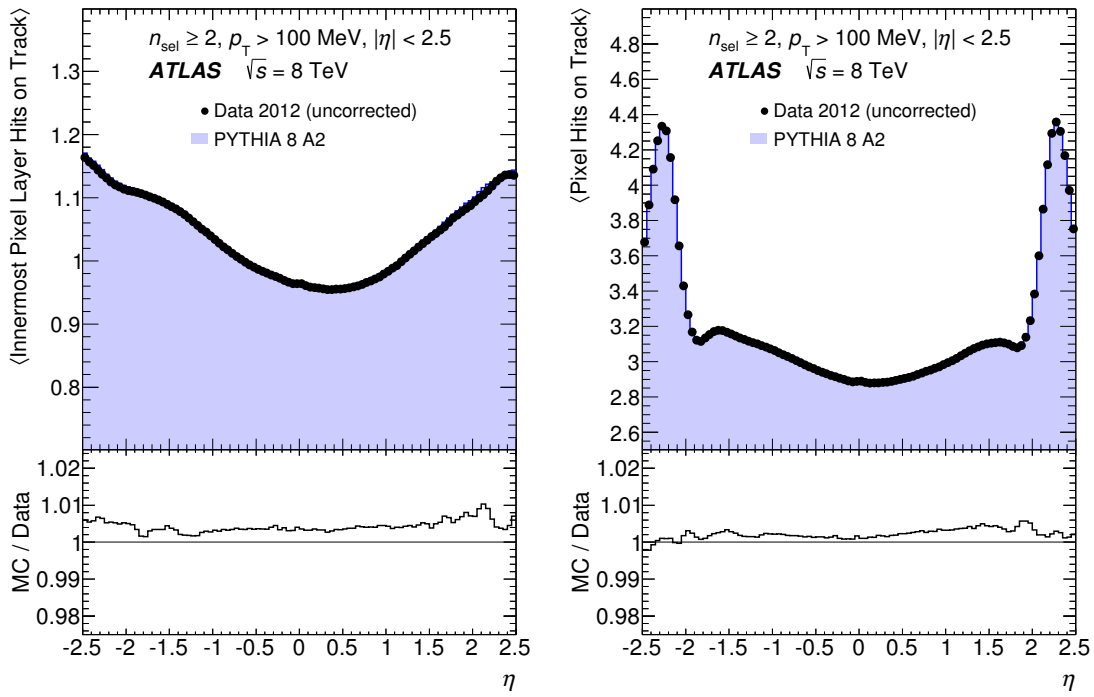


(a) (b)

Figure 7.3: Data and MC distributions of the normalised (a) transverse, d_0^{PV}, and (b) longitudinal, $z_0^{\mathrm{PV}} \cdot \sin\theta$, impact parameters with respect to the reconstructed primary vertex for selected tracks (where only the requirement on the shown parameter is lifted). The MC distributions are reweighted to match the reconstructed data in terms of transverse momentum, as well as the to give the same rates of reconstructed non-primary tracks. Plots taken from Ref. [3].

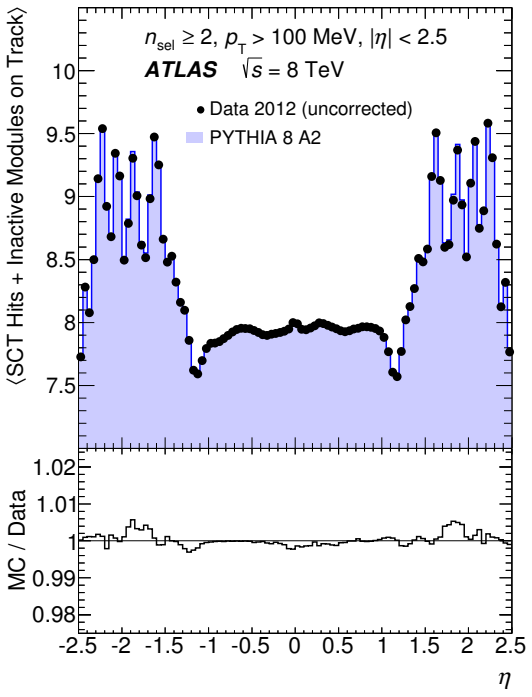
chain and thus contribute neither hits nor holes to the reconstructed tracks traversing them. The distributions show good agreement between data and MC simulation, with slight discrepancies due to imperfect modelling of the detector conditions in MC simulation. Their impact on the track reconstruction efficiency was found to be negligible, and therefore no special systematics were required for them. A detailed comparison of the average number of measurements per track in various p_{T} intervals can be found in Appendix A.4.

Figure 7.5 shows the average number of holes per track in the Pixel and SCT detectors, respectively. The relative difference between data and MC simulation reaches up to a factor > 2, however it is small in absolute terms (mostly < 0.01 holes per track). This causes only $\approx 1\,\%$ of tracks in the most affected η regions to have one additional hole in MC simulation, sometimes at the expense of one hit. If the number of silicon hits and/or holes of such tracks is directly at the threshold of acceptance in the reconstruction algorithms, these tracks could be lost in MC simulation and thus introduce a bias on the track reconstruction efficiency. The average fraction of tracks at this threshold is found to be less than $1\,\%$, and so the overall effect on the track reconstruction efficiency is below the $0.01\,\%$ level in the most affected η regions, which is



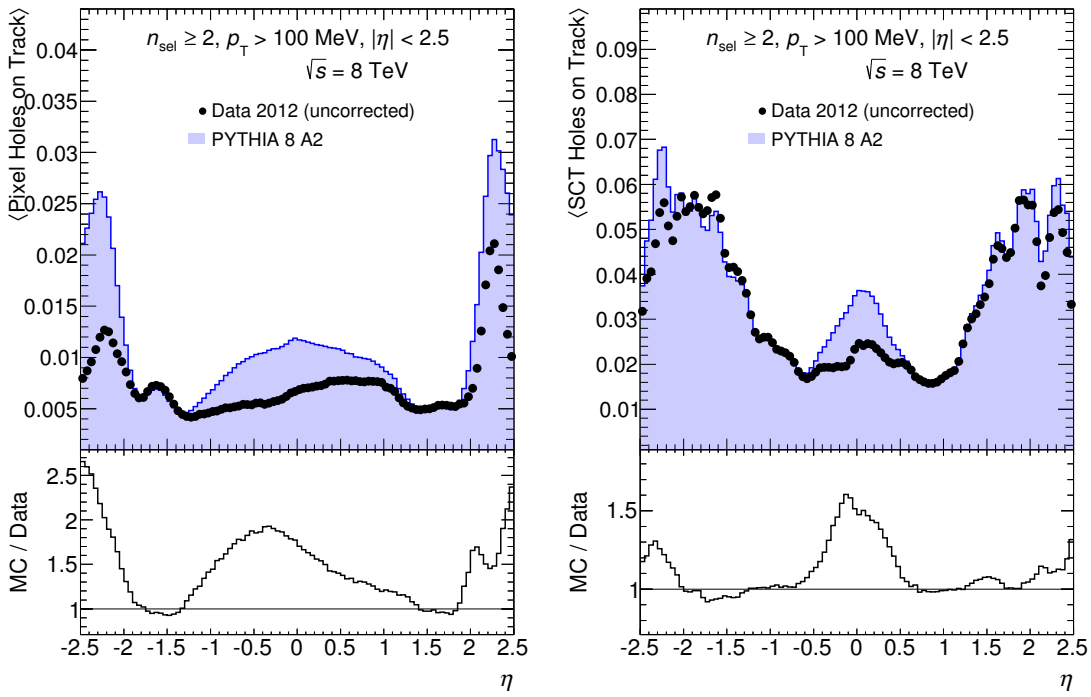
(a) Pixel layer-0

(b) Pixel



(c) SCT

Figure 7.4: Average measurements per reconstructed track in data and MC simulation in (a) the innermost Pixel layer, (b) the entire Pixel detector, and (c) the entire SCT detector (taking the sum of hits and inactive modules). The MC simulation is reweighted to match the detector-level p_T distribution in data. Plots taken from Ref. [3].



(a) Pixel **(b)** SCT

Figure 7.5: Average number of (a) Pixel and (b) SCT holes per reconstructed track in data and MC simulation.

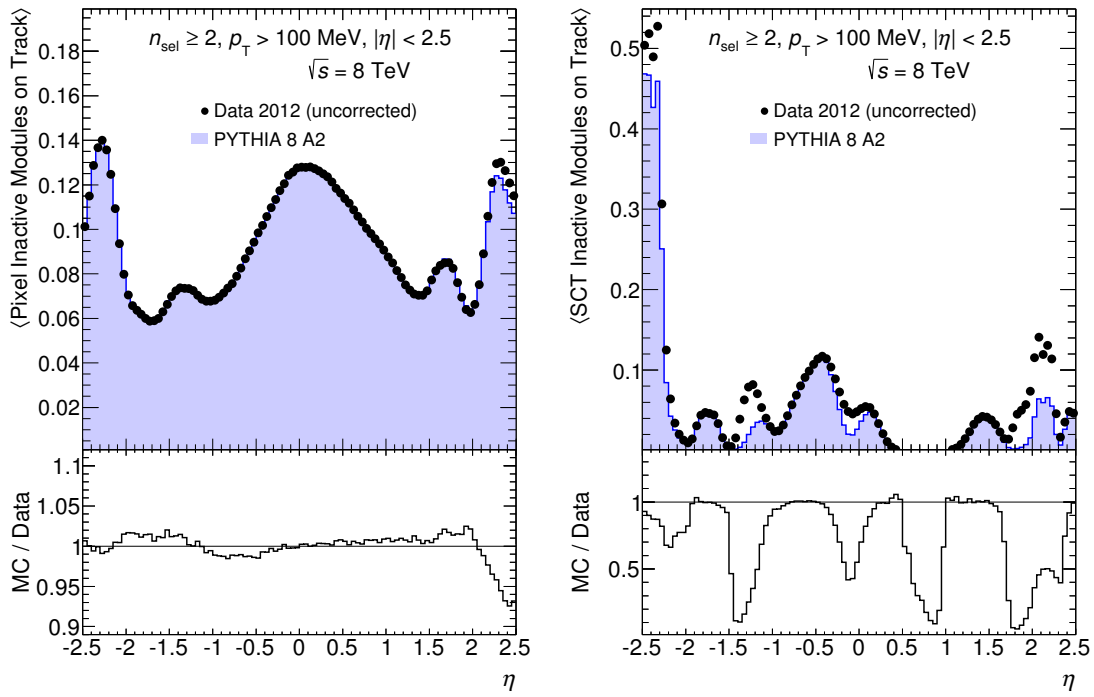
more than two orders of magnitude below the total systematic uncertainties as determined in Section 7.5.3.2. In summary, the effect is negligible.

Figure 7.6 shows the average number of inactive Pixel and SCT modules per track, respectively. These distributions give an indication of whether the run conditions during data-taking were correctly modelled in the MC simulation, but the remaining discrepancy has no direct impact on the track reconstruction which is only sensitive to hits and holes on track.

7.3.5 Detector-Level Distributions

The normalised and uncorrected detector-level distributions for η, p_{T} and n_{sel} are shown in Figure 7.7, using only the vertex z weights w_{zvrtx} for MC events. These distributions are a first validation result of the track reconstruction, which can highlight potential issues such as discrepancies in the material description of the detector,[7] and give an early indication of how well the data are modelled by the MC tunes. Varying performances can be seen with the different generators; at this level the Epos LHC model describes the data better than both

[7] Such material discrepancies became visible in the raw η distribution during the early stages of this analysis, leading to extensive follow-up studies which contributed to significant improvements of the final Run-1 detector geometry description. See also Chapter 6 for details. Comparisons of detector-level distributions obtained from different MC tunes are shown in Appendix A.6.3.1.



(a) Pixel **(b)** SCT

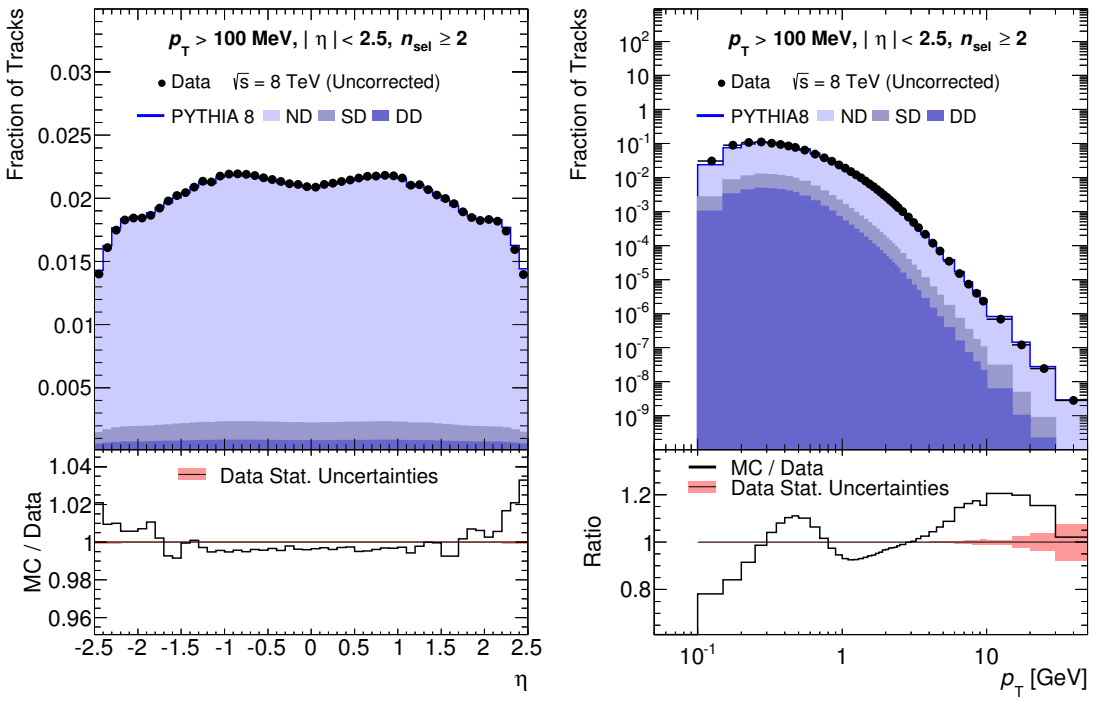
Figure 7.6: Average number of inactive (a) Pixel and (b) SCT modules per reconstructed track in data and MC simulation.

Pythia models. However, in order to obtain the final distributions, further event-level and track-level corrections need to be applied and an unfolding to the particle-level distributions can be performed.

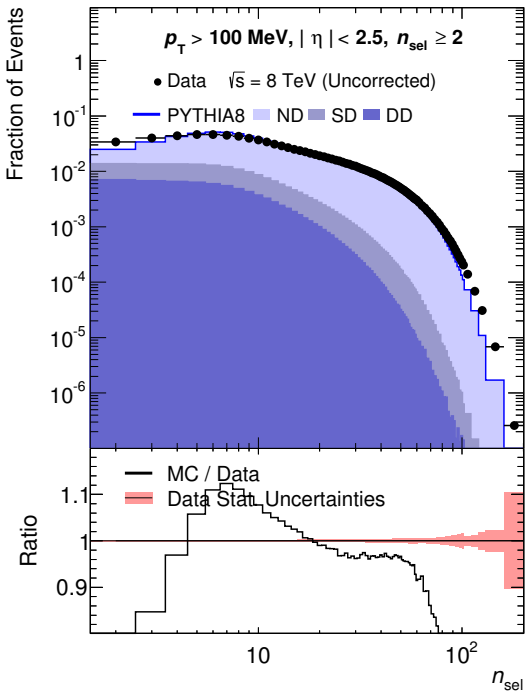
7.4 Background Contribution

The total background contribution to the charged-particle density can be broken down into event-level and track-level backgrounds, which are described in detail in the following sections:

- *Event backgrounds* refer to all classes of non-collision events which do not originate from a single collision of two protons from colliding bunches.
- *Track backgrounds from non-primary tracks* consist of secondary tracks and fake tracks; the former originate mostly from decays or secondary interactions in the detector material, while the latter are combinatorial artefacts from the track reconstruction algorithms with no corresponding generated particles.
- *Track backgrounds from strange baryons* comprise tracks that are reconstructed from measurements produced by the short-lived particles with $30 < \tau < 300\,\text{ps}$, and/or by their decay products which are propagating in a similar direction in both η and ϕ.



(a) (b)



(c)

Figure 7.7: Normalised detector-level distributions of reconstructed tracks as a function of (a) η, (b) p_{T} and (c) reconstructed track multiplicity per event n_{sel}, for Data and Pythia 8 A2 MC simulation. The MC distributions have been reweighted to data to compensate for a different distribution of the primary vertex z position.

The rate of fake tracks was estimated from MC simulation to be very small (below 0.1 %, which is also consistent with earlier measurements [126]). As it is difficult to distinguish between the two categories of secondary and fake tracks, they are simply treated together in the category of *non-primary tracks*.

7.4.1 Background from Non-Collision and Pile-Up Events

The selected events in the data sample can be contaminated due to non-collision events caused by beam-induced background or cosmic-ray background, and due to pile-up effects.

7.4.1.1 Beam-induced background

Beam-induced background (also termed *beam background*) is mainly caused by beam-gas interactions inside the beam pipe or by collisions of protons with upstream collimators [162]. Their ratio is estimated from events produced by non-colliding bunches (isolated unpaired bunches with trigger name `L1_MBTS_2_UNPAIRED_ISO`) [136]. Such background events often lack a reconstructed primary (pseudo-)vertex, and are thus effectively reduced by the vertex requirement in combination with the track selection criteria. In the most inclusive phase space ($n_{\mathrm{ch}} \geq 2$, $p_{\mathrm{T}} > 100$ MeV) the rate of beam-induced background was estimated to be < 0.01 % of all selected events. These remaining background events are more likely to populate specific regions of the measured distributions, such as event categories with very low n_{sel}. They have a significant impact on the data vertex reconstruction efficiency, in particular for low-n_{sel} events with low $p_{\mathrm{T}}^{\mathrm{min}}$, as shown in Section 7.5.2.

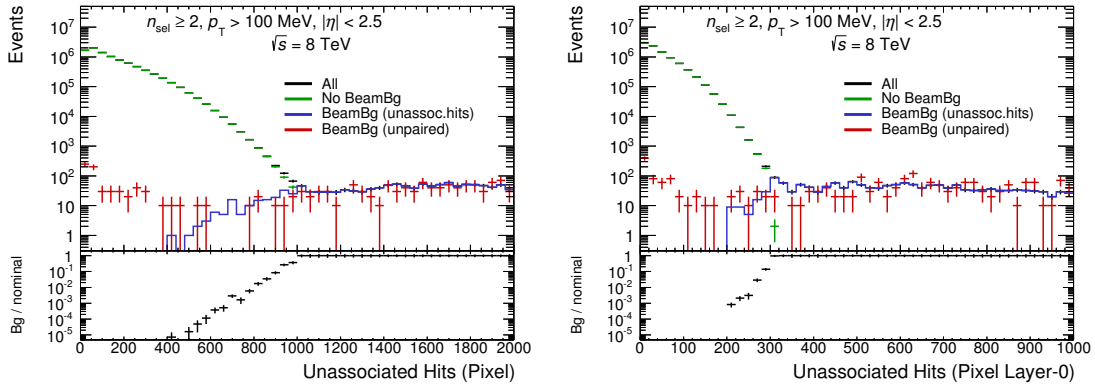
A typical signature for beam-background events is a large number of *unassociated hits* in the Pixel detector which are not associated with any reconstructed tracks. The beam background can therefore be effectively removed with the following cuts on the number of unassociated hits in the whole Pixel detector, $N_{\mathrm{Unassociated}}^{\mathrm{Pixel}}$, and the innermost Pixel layer-0, $N_{\mathrm{Unassociated}}^{\mathrm{Pixel\ layer-0}}$:

$$N_{\mathrm{Unassociated}}^{\mathrm{Pixel}} < 1000, \tag{7.1}$$

$$N_{\mathrm{Unassociated}}^{\mathrm{Pixel\ layer-0}} < 300. \tag{7.2}$$

By rejecting events with a larger number of unassociated hits in either category, ~ 90 % of background and less than 0.01 % of signal events are removed. The distributions of the number of unassociated hits per event are shown in Figure 7.8.

The efficiency of the unassociated-hits cut is validated by comparisons with a data sample using an isolated non-colliding bunches trigger, which shows very good agreement in all distributions, including the MBTS timing difference between the trigger scintillators on the A and C side of the ATLAS detector (see Figure 7.9f). This MBTS timing difference is a characteristic feature of upstream collisions (e.g. with collimators or gas) which are taking place at a distance



(a) Pixel **(b)** Pixel layer-0

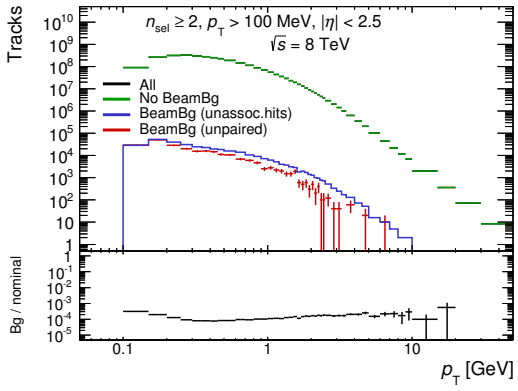
Figure 7.8: Number of unassociated hits per event in nominal (black), nominal minus beam-background (green) and beam-background events (red, blue) in the $\sqrt{s} = 8$ TeV data sample in (a) all layers of the the Pixel detector and (b) the innermost Pixel layer-0. Beam-background events in the nominal data sample (blue) are identified using a selection based on the "unassociated hits" cut in Eq. 7.1, and are compared to events from a special sample (red) that was recorded using a trigger on isolated, non-colliding ("unpaired") bunches. The ratio shows the relative fraction of beam background determined with the former of these two methods.

of several metres from the nominal beam-spot location, producing a number of boosted tracks that can traverse both trigger scintillators with a characteristic time delay of 25–30 ns.

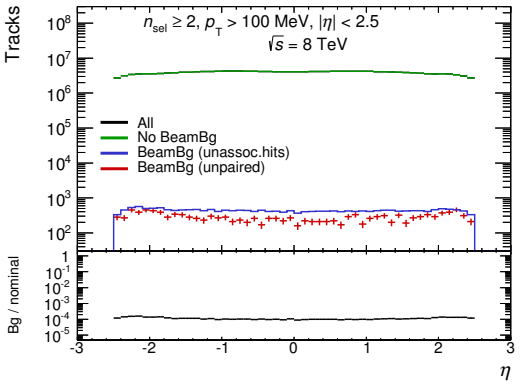
Beam-background events show a functional dependence on the p_T and η of tracks, with higher relative fractions in the low-p_T and high-$|\eta|$ regimes. The number of tracks from beam background is at least three orders of magnitude lower than signal for tracks with $100 < p_T < 200$ MeV, and almost four orders of magnitude lower than signal for tracks above $p_T > 200$ MeV (see Figure 7.9a). The slightly higher fraction of beam background in the high-$|\eta|$ regions (see Figure 7.9b) follows the expected η distribution due to boosted tracks that originated in upstream interactions. Beam-background events are also being found to contribute mostly to very low-n_{sel} events (up to $< 1\,\%$ for $n_{sel} = 2$, see Figure 7.9e). The apparent increase of the fraction of beam background for very high-n_{sel} events above $n_{sel} \geq 100$ is assumed to be partially an artefact of the unassociated-hits cut.[8] Beam-background tracks are found to populate the d_0^{PV} and z_0^{PV} tails more likely than signal events, resulting in slightly flatter impact parameter distributions (see Figures 7.9c and 7.9d).

The beam-background cut from Eq. 7.1 is not applied in the main analysis, as it would introduce additional systematic uncertainties due to the removal of events, including false positives; it is only used for the calculation of the vertex efficiency (see Section 7.5.2) and for

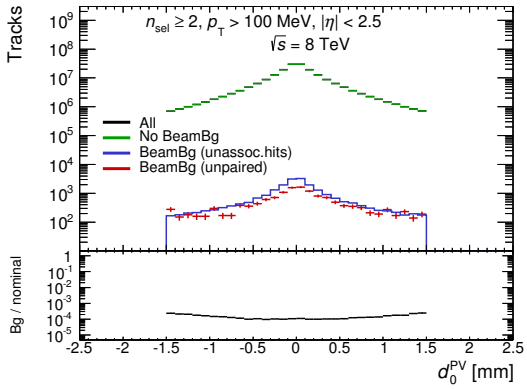
[8] Good collision events with high track multiplicity have a higher number of measurement clusters, and thus also a higher statistical chance to exceed the limit imposed by the cut on the number of unassociated hits. This is assumed to account for a significant fraction of the high-n_{sel} events that are classified as beam background.



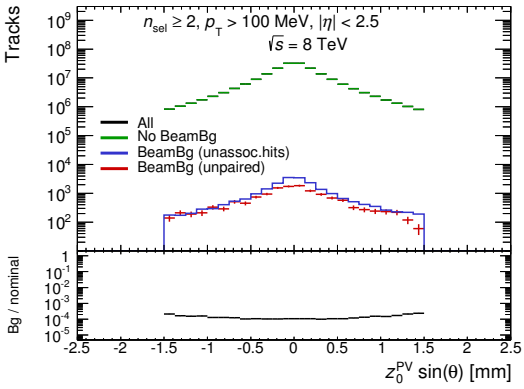
(a) p_{T}



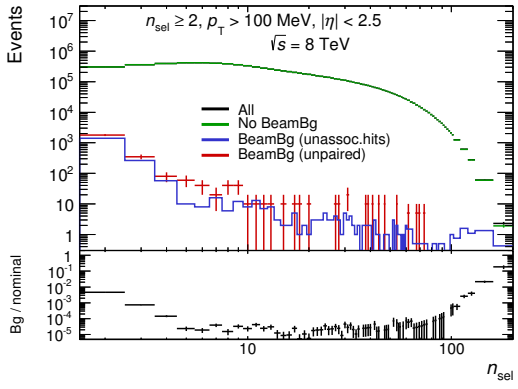
(b) η



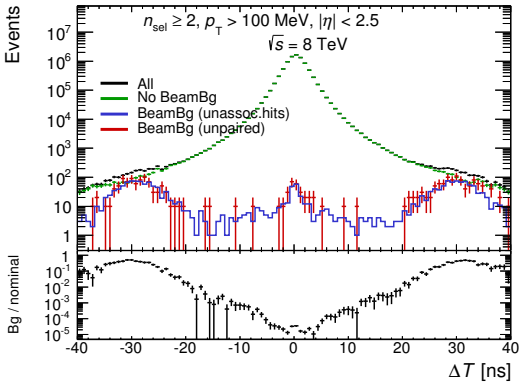
(c) d_0^{PV}



(d) $z_0^{\mathrm{PV}} \cdot \sin\theta$



(e) n_{sel}



(f) MBTS timing

Figure 7.9: Comparison of track (a-d) and event (e-f) distributions in nominal (black), nominal minus beam-background (green) and beam-background events (red, blue) in the $\sqrt{s} = 8$ TeV data sample as a function of (a) p_{T}, (b) η, (c) d_0^{PV}, (d) $z_0^{\mathrm{PV}} \cdot \sin\theta$, (e) n_{sel}, and (f) MBTS timing difference between the two trigger scintillators on each side. Beam-background events are determined either via cuts complementary to the cuts on unassociated hits in Eq. 7.1 (blue) or from a sample using an isolated non-colliding bunches trigger (red). The ratio shows the relative fraction of beam background determined with the former of these two methods.

studies of systematic uncertainties on the non-primary fraction related to the presence of beam background (see Section 7.4.2).

7.4.1.2 Cosmic-ray background

The rate of background events due to cosmic rays was estimated to be less than $< 10^{-6}$ [136], and was therefore neglected.[9]

7.4.1.3 Pile-up

Even though the average number of collisions per bunch crossing was below $\mu < 0.004$ in the data sample used for the analysis (see Section 7.3.1), Poisson statistics suggests that a non-negligible rate of $\approx 0.1\,\%$ pile-up events would remain in the data sample. Pile-up contamination was therefore suppressed by rejecting events with at least one additional reconstructed vertex (beyond the primary vertex) with ≥ 4 associated tracks. This selection reduced the estimated rate of remaining pile-up events in data to a negligible level of about $0.002\,\%$.[10]

The pile-up suppression criterion is found to remove $0.55\,\%$ of data events having passed the MBTS trigger selection, and $0.01\,\%$ of Pythia 8 A2 MC events with a reconstructed primary vertex, respectively; the latter fraction can be interpreted as the rate of false positives of the method, but the resulting possible bias is found to be negligibly small and is therefore ignored in the remainder of the analysis. The rate of events containing additional reconstructed vertices with < 4 tracks is found to be $< 0.01\,\%$. Most of these events, as well as the removed false positives, are either characterised by split vertices (see Section 5.2.1), or they contain tracks from secondary interactions that are wrongly reconstructed as belonging to a second primary vertex.

7.4.2 Background from Non-Primary Tracks

Reconstructed tracks corresponding to non-primary particles can originate mainly from these sources:

1. hadronic interactions of particles with the beam pipe or detector material,
2. decays of long-lived primary particles with strange quark content (mostly K^0 and Λ^0) that were created at the pp collision,
3. photons from electron bremsstrahlung, π_0 and η decays, which are converting to electron-positron pairs via interactions with the beam-pipe or detector material.

[9] The rate of cosmic rays passing through the beam line is expected to be $< 0.001\,\text{Hz}$, while the event readout rate during the low-μ run is on the order of 1–2 kHz.

[10] This evaluation was made by combining the expected fraction of pile-up events from μ, the fraction of events with $n_{\text{sel}} < 4$ selected tracks, and the estimated fraction of events with close-by vertices ($\Delta z < 5\,\text{mm}$). A conservative estimate yields < 200 remaining pile-up events (out of 9.2×10^6 selected events) which may contaminate the data sample by increasing the track multiplicity of the event by up to three tracks. This is clearly negligible.

Due to the prevalence of electron-positron pair production at low p_{T}, and the characteristic shapes of impact parameter distributions of electrons and positrons (which are dominated by their momenta and the detector geometry, and differ from the corresponding shapes due to other particles), the non-primaries are subdivided into the categories of electrons (including positrons) and non-electrons. While it is straightforward to calculate the fractions of reconstructed primary, electron and non-electron tracks in the MC sample by classifying them according to their associated generated particles, the MC models do not necessarily predict these fractions correctly. It is found that the non-primary fractions in the data sample can be different, and hence the scaling factors necessary to compensate for this difference must be evaluated by other means, as described in Section 7.4.2.1.

The category of non-electrons also includes combinatorial *fake tracks* that are described in Section 5.1. The fraction of fake tracks was estimated to be O(0.1%) using MC simulation, which is more than one order of magnitude less than the fraction of secondary tracks. Therefore fake tracks do not need to be treated separately and can be combined with secondary tracks into the category of non-primary (non-electron) tracks.

7.4.2.1 Template fitting method

The fraction of non-primary particles $f_{\mathrm{nonp}}(\eta, p_{\mathrm{T}})$ within the selected sample is determined via a template fitting method in a similar way as in previous analyses [126], exploiting the property of non-primary particles to show more pronounced tails in their impact parameter distributions than primary particles. A refined version of the method described in Ref. [21] is presented here. In principle, templates of the transverse (d_0^{PV}) and longitudinal (z_0^{PV}) impact parameter distributions of primary and non-primary tracks in the MC samples are scaled to fit the distribution in data. The scaling factors obtained from these MC template fits can then be used to determine the actual fraction of non-primary tracks within the data sample. In the current analysis, only the d_0^{PV} distributions were used in the template fitting method, because - as previous studies have shown - they are more sensitive to secondaries than z_0^{PV} distributions due to the way in which a displaced secondary vertex shapes these two distributions.

The template fitting method is applied for various side-band fitting ranges of the d_0^{PV} distributions, which are chosen to be outside of the signal region (given by $|d_0^{\mathrm{PV}}| < 1.5\,\mathrm{mm}$ according to the track selection criteria) in order to acquire an unbiased estimate of the non-primary fractions in the data. Since at lower transverse momenta the impact parameters have broader distributions and the rate of secondary tracks is higher, the method is applied separately for various p_{T} ranges. The d_0^{PV} distributions of primary tracks, as well as electron (including positron) and non-electron non-primaries in the MC samples, are extracted in p_{T} intervals of 50 MeV width between $100 < p_{\mathrm{T}} < 550$ MeV, and one p_{T} interval for tracks with $p_{\mathrm{T}} > 550$ MeV. They are used as templates $t_{\mathrm{i}}(d_0^{\mathrm{PV}})$ with corresponding scale factors f_{i}. For each p_{T}-interval the linear combination of these templates,

$$t_{\text{fit}}(d_0^{\text{PV}}) = f_{\text{prim}} \cdot t_{\text{prim}}(d_0^{\text{PV}}) + f_{\text{el}} \cdot t_{\text{el}}(d_0^{\text{PV}}) + f_{\text{nonel}} \cdot t_{\text{nonel}}(d_0^{\text{PV}}), \tag{7.3}$$

is fitted to the corresponding d_0^{PV} distribution in data in all chosen side-band regions, by varying two of the three scale factors while satisfying the following normalisation constraint,

$$\int \mathrm{d}d_0^{\text{PV}}\ t_{\text{fit}}(d_0^{\text{PV}}) \ = \ \int \mathrm{d}d_0^{\text{PV}}\ (t_{\text{prim}}(d_0^{\text{PV}}) + t_{\text{el}}(d_0^{\text{PV}}) + t_{\text{nonel}}(d_0^{\text{PV}})). \tag{7.4}$$

At higher p_{T} the fraction of electrons decreases and the fitting method becomes insensitive to the electron template. Above $p_{\text{T}} > 500$ MeV a combined template for all non-primaries (electrons and non-electrons) is used instead, given by:

$$t_{\text{fit,combined}}(d_0^{\text{PV}}) = f_{\text{prim}} \cdot t_{\text{prim}}(d_0^{\text{PV}}) + f_{\text{nonp}} \cdot t_{\text{nonp}}(d_0^{\text{PV}}) \tag{7.5}$$

For each combination of scale factors $f_{\text{i}} \in [0.2; 2]$, the χ^2/n_{dof} between the fitted template t_{fit} and the data is calculated within the side-band regions given by $d_{0,\text{min}}^{\text{PV}} < |d_0^{\text{PV}}| < d_{0,\text{max}}^{\text{PV}}$. The lower limit on both sides is varied within $d_{0,\text{min}}^{\text{PV}} \in [1.5; 6.5]$ mm in steps of 0.5 mm. The fitting ranges have an upper limit at $d_{0,\text{max}}^{\text{PV}} = 9.5$ mm due to the track reconstruction algorithm requiring the track to be within $|d_0^{\text{PV}}| < 10$ mm.

These 11 different d_0^{PV} intervals allow to select a range of side-band regions in which the fit results are relatively stable, also requiring the electron template to have a sufficient discrimination power in the overall fit while neither dominating the non-electron fit nor vice versa, i.e. obtaining neither very large nor very small scale factors. The central $d_{0,\text{min}}^{\text{PV}}$ value within this range, where neither of the templates dominate the fit, can be used to derive the nominal scale factors. Their difference to scale factors obtained within the fit intervals at the lower and upper $d_{0,\text{min}}^{\text{PV}}$ edges of the stable range, with either primaries or non-primaries being more dominant in the fit, allow to estimate systematic errors on the scale factors. Eventually these fitted templates allow to extract the corresponding fractions of primaries and non-primaries in data.

Examples of the fitted templates in different p_{T} intervals are shown in Figure 7.11. It can be seen that the sensitivity of the method to the electron template decreases at higher p_{T}, where the d_0^{PV} distribution gradually flattens in the side-band region while showing a decreasing overall rate compared to the non-primary template, and thus loses its discrimination power in the fitting procedure.[11] The scaling factors obtained for each p_{T} bin and non-primary category are listed in Table 7.5.

[11] The peaks of the electron distribution as a function of $|d_0^{\text{PV}}|$, clearly visible in Figure 7.11, correspond to the radial ID material distribution in the beam-pipe and innermost pixel layer, where the presence of more material increases the local probability for photon conversions into e^+e^- pairs; see also Section 6.3.

Category	Template scale factors		
p_{T} [MeV]	f_{nonel}	f_{el}	f_{combined}
100–150	$1.25 \pm 0.00^*$	$0.59 \pm 0.01^*$	1.01 ± 0.01
150–200	$1.25 \pm 0.01^*$	$0.67 \pm 0.04^*$	1.10 ± 0.02
200–250	$1.20 \pm 0.00^*$	$0.71 \pm 0.02^*$	1.09 ± 0.03
250–300	$1.23 \pm 0.00^*$	$0.75 \pm 0.01^*$	1.15 ± 0.02
300–350	$1.25 \pm 0.01^*$	$0.84 \pm 0.09^*$	1.21 ± 0.02
350–400	$1.29 \pm 0.01^*$	$0.95 \pm 0.12^*$	1.26 ± 0.01
400–450	$1.33 \pm 0.03^*$	$1.22 \pm 0.39^*$	1.32 ± 0.01
450–500	$1.36 \pm 0.03^*$	$1.34 \pm 0.43^*$	1.36 ± 0.02
500–550	1.40 ± 0.05	1.38 ± 0.62	$1.39 \pm 0.01^*$
> 550	1.33 ± 0.03	2.00 ± 0.03	$1.34 \pm 0.04^*$

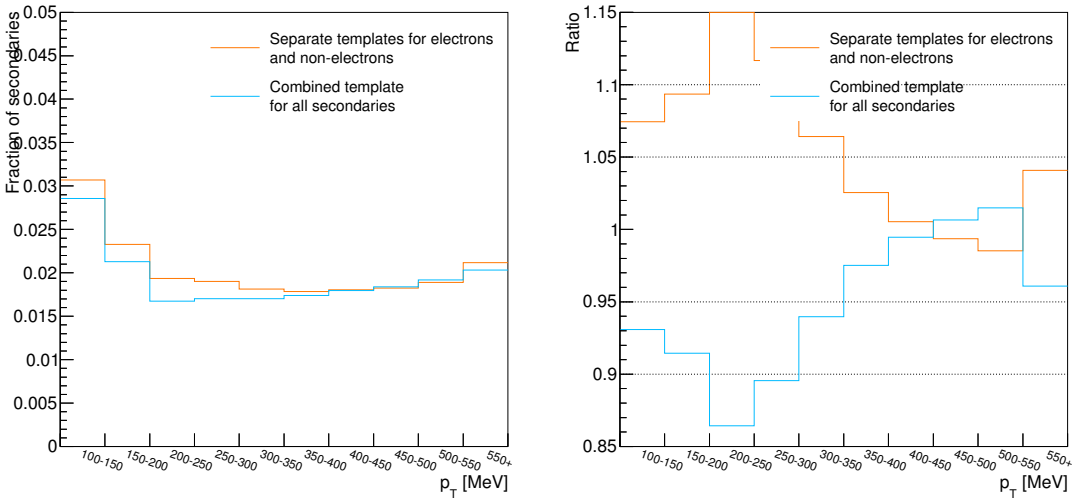
Table 7.5: Scale factors from template fitting method using separate (Eq. 7.3) and combined (Eq. 7.5) non-primary templates, with uncertainties obtained from variations of the fit intervals. (*) = scale factors used in final analysis

7.4.2.2 Systematic Uncertainties

Various aspects of the method to calculate the non-primary fractions in data introduce individual systematic uncertainties which are added in quadrature: these are the choices of separate or combined templates, d_0^{PV} fit intervals, MC tunes, and physics lists, as well as the uncertainty due to different transfer functions for MC tunes and due to non-collision background events. Conservative estimates are made for these individual sources, because the contribution of the total systematic uncertainty on non-primary fractions to the final distributions is completely dominated by the material uncertainty (see Section 7.5.3), thus relieving the necessity to optimise each individual contribution here.

Choice of separate or combined templates At higher p_{T} the fitting ranges with larger $d_{0,\mathrm{min}}^{\mathrm{PV}}$ become insensitive to the electron template, and a combined template for all non-primaries (electrons and non-electrons) can be used instead (see Eq. 7.5). The ratio of the calculated non-primary fraction between using separate and combined templates is shown in Figure 7.10. This allows to estimate that the choice of using either separate or combined templates introduces a systematic uncertainty of up to 15 % in the range between $200 < p_{\mathrm{T}} < 250$ MeV) and as little as 1 % between $400 < p_{\mathrm{T}} < 500$ MeV). However this should be considered an overly conservative estimate for the lower p_{T} bins (below $p_{\mathrm{T}} < 300$ MeV), where a combined non-primary template would be clearly insufficient for the template fitting method.

Choice of $|d_0^{\mathrm{PV}}|$ fit intervals As the fitting method loses sensitivity to the electron template with increasing p_{T}, and the fit/data agreement in the signal region deteriorates with increasing



(a) **(b)**

Figure 7.10: Average non-primary fraction using separate or combined templates for electrons and non-electrons, shown as (a) comparison and (b) ratio.

$d_{0,\mathrm{min}}^{\mathrm{PV}}$, it does not make sense to consider all d_0^{PV} fit intervals, particularly for the higher p_{T} bins. Stable ranges of fit intervals are found by analysing the fit/data ratios for each p_{T} bin and fit interval (see Figure 7.11 for examples) and requiring the stability of scale factors when fit intervals are modified. Thus upper limits on $d_{0,\mathrm{min}}^{\mathrm{PV}}$ are set and a representative nominal value at the centre of that range is found. Results from this procedure are presented in Table 7.6.

The systematic uncertainty introduced by choosing one particular nominal fit interval is estimated by calculating the ratios of non-primary fractions at the upper and lower edges of the selected stable range of d_0^{PV} intervals with respect to the nominal fractions, as shown in Figure 7.12. From these ratios a systematic uncertainty of up to 3 % can be assigned to the specific choice of d_0^{PV} fit intervals in all p_{T} bins between $100 < p_{\mathrm{T}} < 500$ MeV (using separate templates), 3 % from $500 < p_{\mathrm{T}} < 550$ MeV and 5 % above $p_{\mathrm{T}} > 550$ MeV (both using combined templates).

Choice of MC generator and physics list The Pythia 8 A2 tune which is used for the main analysis is also used for the template fitting procedure. Possible model-dependent systematic effects of this specific choice are assessed by repeating the template fitting procedure with alternative MC samples produced by the Pythia 6 AMBT2B and the Epos LHC models. The ratios of the fitted non-primary fractions between these generators and the baseline Pythia 8 A2 tune are shown in Figure 7.13. From these ratios, a systematic uncertainty of up to 5 % is taken to account for the event-generator choice.

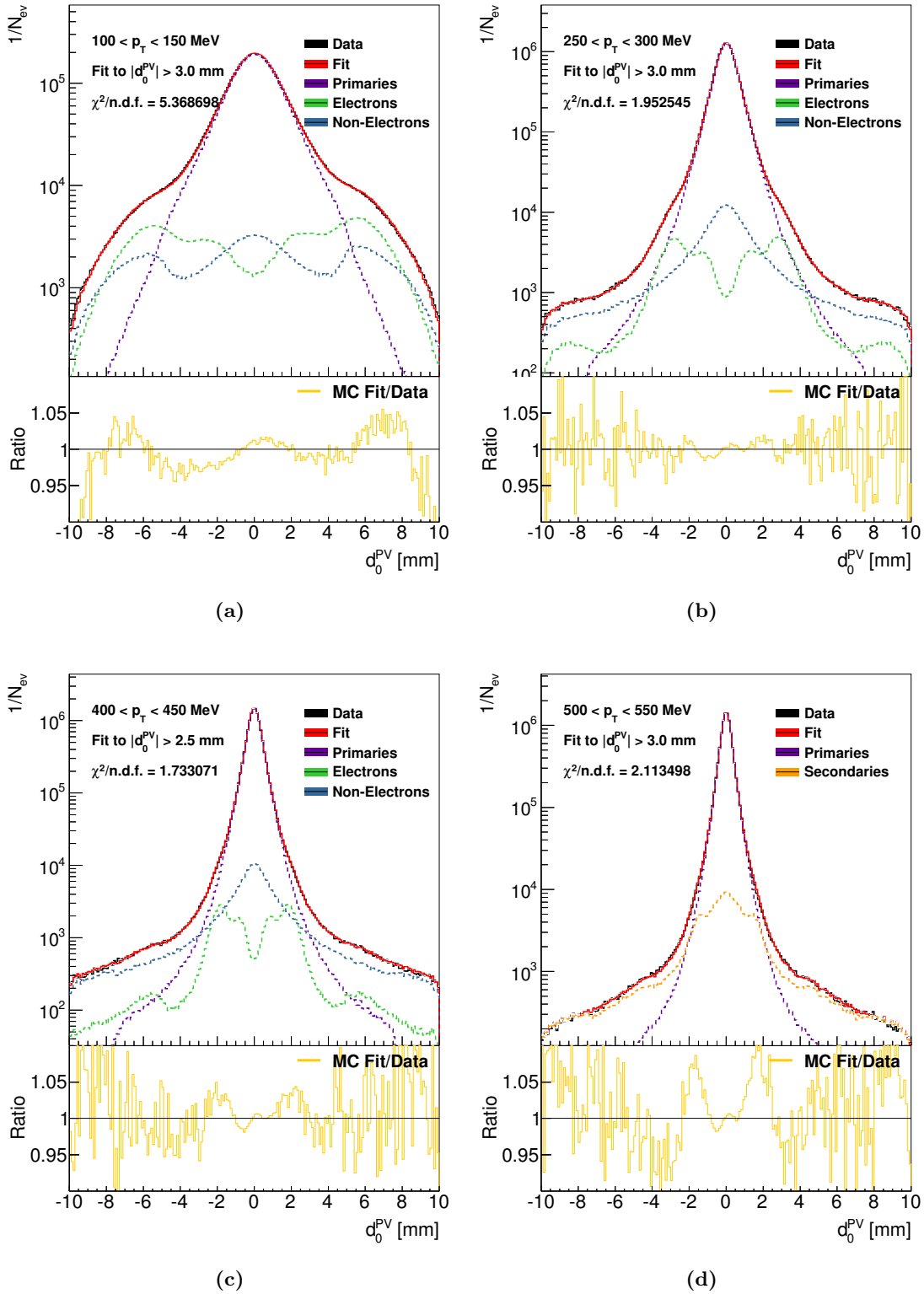
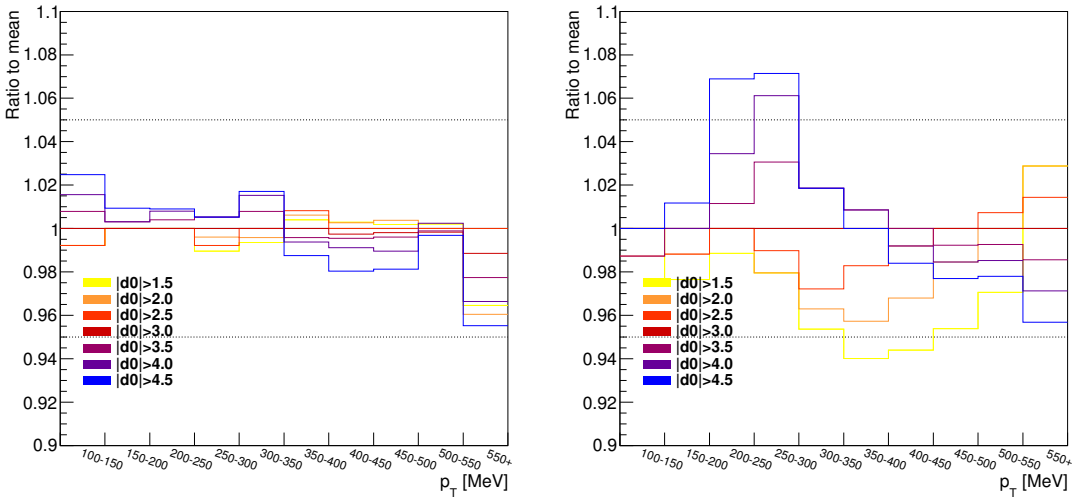
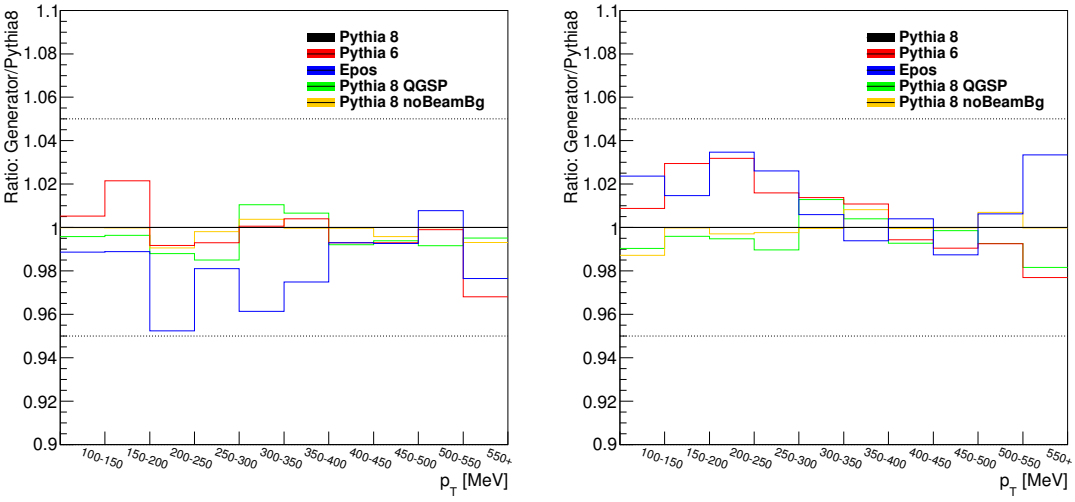


Figure 7.11: Transverse impact parameter d_0^{PV} distribution at $\sqrt{s} = 8$ TeV for primary and non-primary particles (MC templates) after scaling them to the best fit value to data, using tracks within $|d_0^{\mathrm{PV}}| > 3\,\mathrm{mm}$ or $2.5\,\mathrm{mm}$ and (a) $100 < p_{\mathrm{T}} < 150$ MeV, (b) $250 < p_{\mathrm{T}} < 300$ MeV, (c) $400 < p_{\mathrm{T}} < 450$ MeV (all using separate electron and non-electron templates) and (d) $500 < p_{\mathrm{T}} < 550$ MeV (using combined templates).



(a) **(b)**

Figure 7.12: Ratio between the non-primary fraction at the high/low $d_{0,\mathrm{min}}^{\mathrm{PV}}$ fit interval within the preselected stable range and the fraction at the nominal $d_{0,\mathrm{min}}^{\mathrm{PV}}$ fit interval for (a) separate and (b) combined templates.



(a) **(b)**

Figure 7.13: Ratio between the nominal non-primary fraction calculated from the Pythia 8 A2 MC sample and the fraction obtained with different MC generators or physics lists, as well as a Pythia 8 A2 MC sample with removed beam background, for (a) separate and (b) combined templates.

Category	Template fit window		Systematic
p_{T} [MeV]	max $d_{0,\mathrm{min}}^{\mathrm{PV}}$	nominal $d_{0,\mathrm{min}}^{\mathrm{PV}}$	$\sigma_{\mathrm{interval}}$
100–150	4.5 mm	3.0 mm	3 %
150–200	4.5 mm	3.0 mm	1 %
200–250	4.5 mm	3.0 mm	1 %
250–300	4.5 mm	3.0 mm	2 %
300–350	3.5 mm	2.5 mm	2 %
350–400	4.5 mm	3.0 mm	2 %
400–450	4.0 mm	2.5 mm	2 %
450–500	4.0 mm	2.5 mm	2 %
500–550	4.5 mm	3.0 mm	3 %
> 550	4.5 mm	3.0 mm	5 %

Table 7.6: Upper limits on $d_{0,\mathrm{min}}^{\mathrm{PV}}$ for the intervals used to fit the non-primary templates, and nominal fit intervals chosen to derive non-primary fractions and scale factors, together with the systematic uncertainty from the choice of fit intervals. Combined templates are being used for intervals above $p_{\mathrm{T}} > 500$ MeV.

Category	MC Generator			Systematic
p_{T} [MeV]	Pythia 8 A2	Pythia 6 AMBT2B	Epos LHC	$\sigma_{\mathrm{MC,transfer}}$
100–150	0.0559	0.0558	0.0550	1.6 %
150–200	0.0398	0.0402	0.0390	2.1 %
200–250	0.0306	0.0302	0.0289	5.9 %
250–300	0.0283	0.0279	0.0276	2.5 %
300–350	0.0282	0.0281	0.0267	5.6 %
350–400	0.0240	0.0240	0.0232	3.4 %
400–450	0.0252	0.0250	0.0248	1.6 %
450–500	0.0245	0.0242	0.0240	2.1 %
500–550	0.0253	0.0249	0.0248	2.0 %
> 550	0.0268	0.0261	0.0273	2.7 %

Table 7.7: Transfer functions of different MC generators, and systematic uncertainty obtained from the largest differences with respect to the Pythia 8 A2 tune.

Additionally, the impact of different transfer functions was assessed for the different MC generators. The transfer function is defined as the ratio of non-primaries in the sideband fit region to non-primaries in the signal region. The largest difference between the Pythia 8 A2 tune and either Epos LHC or Pythia 6 AMBT2B is taken as an additional systematic

contribution due to the choice of MC generators. The largest of these variations is found to be below 6 %, as shown in Table 7.7.

The impact of the choice of physics list was also assessed, allowing for an additional estimate of the sensitivity of the template fits to slightly changing input parameters. The difference between the new nominal FTFP and the previously used QGSP physics list was found to be less than 2 % as shown in Figure 7.13.

Beam-background contributions Tracks from beam-background events are found to contribute mostly to the low-p_{T} regime, and to be more uniformly distributed in d_0^{PV} than nominal tracks from signal events (see Figure 7.9c), resulting in a relatively enhanced effect in the $|d_0^{\mathrm{PV}}|$ tails. Since this contribution can affect the template fitting procedure, the impact of beam background on the non-primary fractions was studied. After beam-background removal with the cuts described in Eq. 7.1, the template fitting procedure was repeated for the data and MC sample. Figure 7.13 shows a comparison between the non-primary fraction after beam-background removal and the nominal fraction in each p_{T} bin. The systematic uncertainty due to this effect was found to be 1 %, which is negligible in comparison to other systematic contributions.

Other sources of systematic uncertainty The dependence of non-primary fractions and scale factors on n_{sel} was investigated and found to be negligible; a minor dependence on η can be covered by a systematic uncertainty of less than 1 %, which is perfectly negligible in comparison to other contributions, and thus neglected.

Category	**Systematic uncertainty [%]**						
p_{T} [MeV]	$\sigma_{\mathrm{templates}}$	$\sigma_{\mathrm{interval}}$	σ_{MC}	$\sigma_{\mathrm{transfer}}$	$\sigma_{\mathrm{physics}}$	$\sigma_{\mathrm{background}}$	total σ_{syst}
100–150	8	3	1	2	1	1	10
150–200	10	1	2	2	1	1	11
200–250	15	1	5	6	2	1	17
250–300	12	2	2	3	2	1	13
300–350	7	2	4	6	1	1	11
350–400	3	2	3	4	1	1	7
400–450	1	2	1	2	1	1	4
450–500	1	2	1	2	1	1	4
500–550	2	3	1	2	1	1	5
> 550	5	5	4	3	2	1	9

Table 7.8: Individual systematic uncertainty contributions and total systematic uncertainty on the fraction of non-primaries.

Total systematic uncertainty from non-primaries The individual sources of systematic uncertainty are added in quadrature to determine the total systematic uncertainty on the fraction of non-primary tracks, as listed in Table 7.8. The largest total uncertainty of 17 % is found only in the $200 < p_{\mathrm{T}} < 250$ MeV bin, whereas in other bins the uncertainty stays below 13 %. As the largest uncertainties are dominated by an overly conservative estimate of $\sigma_{\mathrm{templates}}$, an average estimate of a symmetrised 15 % uncertainty for all p_{T} bins is chosen for the analysis. The impact on the final p_{T} distributions remains below 1 % in the range up to $p_{\mathrm{T}} < 30$ GeV and rises to as much as 2 % at higher transverse momenta.

7.4.2.3 Fitted fraction of non-primary tracks in data

The fraction of non-primaries in the signal region is calculated from the fitted MC templates for all intervals listed in Table 7.6, and systematic uncertainties are calculated according to Table 7.8. Separate templates for electron and non-electron non-primaries are employed below 500 MeV, whereas above a combined template for all non-primaries is utilised. Table 7.9 lists the calculated fractions of non-primaries in the data and the relative abundance of electrons among them.

Category	**Fraction of tracks in data**		
p_{T} [MeV]	non-primary tracks	electron tracks	electrons [%]
100–150	0.0307 ± 0.0031	0.0114 ± 0.0011	37 ± 7
150–200	0.0233 ± 0.0026	0.0063 ± 0.0007	27 ± 6
200–250	0.0194 ± 0.0033	0.0042 ± 0.0007	22 ± 7
250–300	0.0190 ± 0.0025	0.0039 ± 0.0005	21 ± 5
300–350	0.0181 ± 0.0020	0.0034 ± 0.0004	19 ± 4
350–400	0.0178 ± 0.0012	0.0036 ± 0.0003	20 ± 3
400–450	0.0181 ± 0.0007	0.0034 ± 0.0002	19 ± 2
450–500	0.0183 ± 0.0007	0.0027 ± 0.0001	15 ± 1
500–550	0.0192 ± 0.0010	(combined)	-
> 550	0.0203 ± 0.0018	(combined)	-

Table 7.9: Scaled fractions of all reconstructed non-primary and electron tracks in data, and relative fraction of electrons among all non-primary tracks in data, with total statistical and systematic uncertainties added in quadrature.

7.4.3 Background from Tracks produced by Strange Baryons

A novel aspect of the Minimum Bias measurements at $\sqrt{s} = 8$ and 13 TeV is that stable primary charged particles are now defined as having a mean lifetime $\tau > 300$ ps, instead of the previously used definition $\tau > 30$ ps. All particles with $30 < \tau < 300$ ps are now considered unstable,

as they decay in the ATLAS detector after a flight length of typically $20 < c\tau < 50\,\text{mm}$ (see Table 7.10), and thus rarely reach beyond the Pixel detector into the SCT. These particles were found to belong to the class of strange baryons: Σ^-, Σ^+, Ξ^- and Ω^-.

Particle	**PDG code**	**Lifetime** [ps]	**Decay length** $c\tau$ [mm]	**rel. abundance** [%]
Σ^-	3112	147.9 ± 0.011	44.3	~ 40
Σ^+	3222	80.2 ± 0.003	24.1	~ 35
Ξ^-	3312	163.9 ± 0.015	49.1	~ 20
Ω^-	3334	82.1 ± 0.011	24.6	~ 5

Table 7.10: Summary of the different short-lived particles with a mean lifetime between $30 < \tau < 300\,\text{ps}$, which are now excluded from our fiducial stable-charged-particle definition, while their decay products are considered to be background non-primary particles.

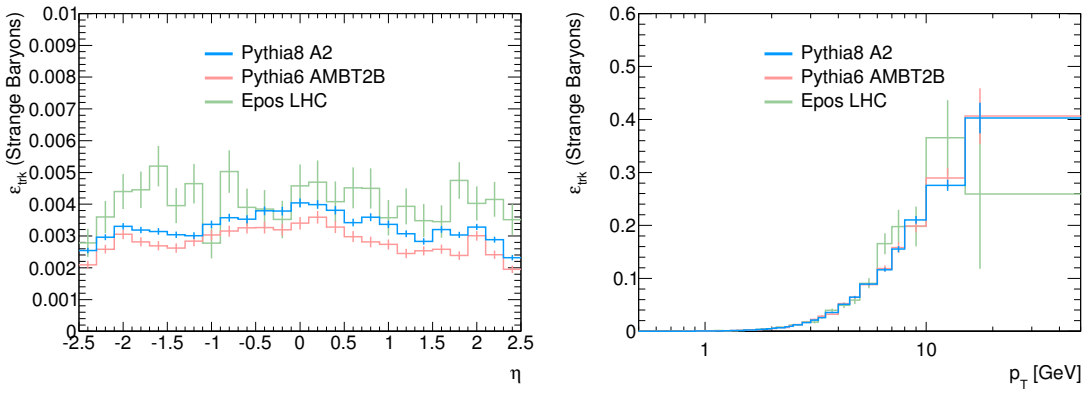
The reconstruction efficiency (see Section 7.5.3 for the definition) of these particles (and/or their decay products which may propagate in a similar direction in both η and ϕ) is close to 0 % for most regions of the phase spaces under consideration, as shown in Figure 7.14, but shows a strongly p_{T}-dependent rise at higher momenta. The production rates of these particles vary significantly between different generators (see Figure 7.15); if included in the analysis, they would present a significant model-dependent systematic error of up to 5 % in the most affected p_{T} range to the overall track reconstruction efficiency. Therefore, in contrast to previous analyses, the fraction of reconstructed tracks associated with particles with a mean lifetime of $30 < \tau < 300\,\text{ps}$ is now evaluated from MC simulation and subtracted from the final distributions, using a track p_{T} and η dependent parametrisation of their rate. This fraction is shown in in Figure 7.16 for different MC generators.

7.4.3.1 Systematic Uncertainties

On the fraction of removed tracks, a systematic variation of 50 % is assigned due to the observed differences between MC generators. In all final η distributions shown in this analysis, the systematic uncertainty due to this removal procedure is below 0.01 % and thus negligible. However, the correction makes a significant impact on final p_{T} distributions in the high-p_{T} range from $\gtrsim 5$ GeV onwards. The combined impact of this correction and the non-primary correction, added in quadrature, on the final p_{T} distributions rises to as much as 3.5 % at higher transverse momenta of $30 < p_{\text{T}} < 50$ GeV (see Figure 7.38 in Section 7.8).

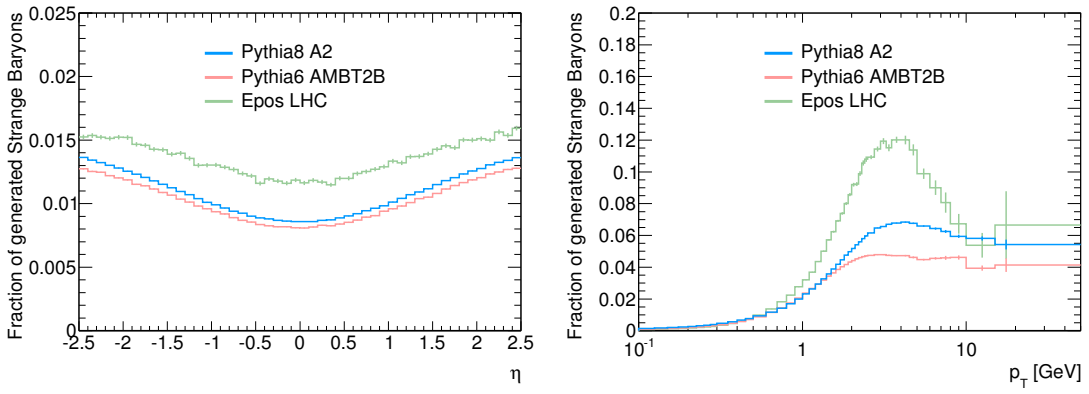
7.4.3.2 Comparison with Previous Measurements

For a comparison with previous measurements which did not remove the strange baryons, their generated fractions (as shown in Figure 7.15) must be added again to the measured distributions.



(a) **(b)**

Figure 7.14: Reconstruction efficiency of tracks associated with strange baryons as a function of (a) η and (b) p_{T} in the phase space $p_{\mathrm{T}} > 100$ MeV, $n_{\mathrm{ch}} \geq 2$, for different MC generators.



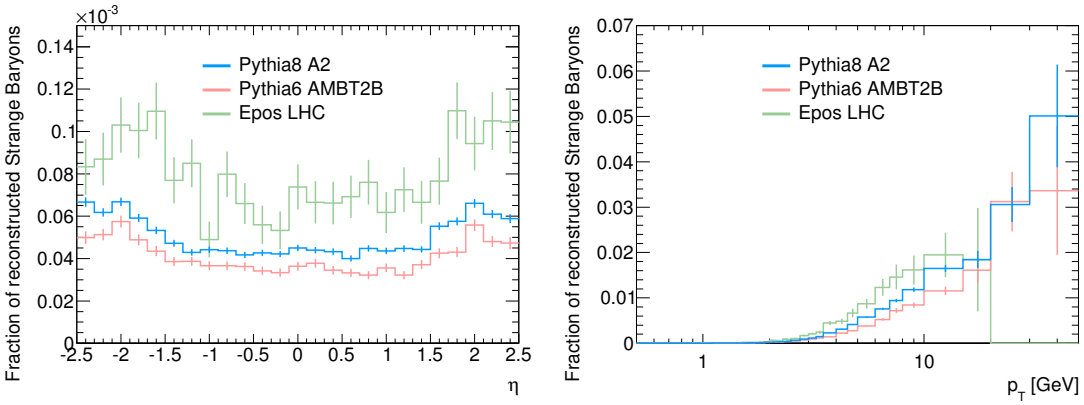
(a) **(b)**

Figure 7.15: Fractions of generated strange baryons as a function of (a) η and (b) p_{T} in the phase space $p_{\mathrm{T}} > 100$ MeV, $n_{\mathrm{ch}} \geq 2$, for different MC generators.

In this analysis and the $\sqrt{s} = 13$ TeV measurement, correction factors were derived from the predicted rates of generated strange baryons using the EPOS LHC tune (as it was shown that these EPOS predictions agree well with ALICE data [163, 164]), and the envelope of differences to other tunes including PYTHIA 8 A2 is taken as a systematic uncertainty. For the corrections of the central charged-particle densities (discussed in Section 7.9.5), the integral over $|\eta| < 0.2$ of the generated fractions of strange baryons is taken, yielding extrapolation factors of 1.012 ± 0.004 in the $p_{\mathrm{T}} > 100$ MeV phase space, and 1.025 ± 0.008 in the $p_{\mathrm{T}} > 500$ MeV phase spaces.

7.5 Selection Efficiencies

Any events and tracks from pp collisions which are not selected for the analysis (e.g. due to hardware detector effects, reconstruction inefficiencies of the software algorithms, etc.) must be



(a) **(b)**

Figure 7.16: Fractions of reconstructed tracks associated with strange baryons as a function of (a) η and (b) p_{T} in the phase space $p_{\mathrm{T}} > 100$ MeV, $n_{\mathrm{ch}} \geq 2$, for different MC generators.

accounted for, by applying event-level and track-level corrections to the measured distributions to compensate for the losses. These corrections and their associated uncertainties are derived from the measured efficiencies of the event selection, comprising the MBTS trigger selection efficiency (Section 7.5.1) and primary-vertex reconstruction efficiency (Section 7.5.2), and of the track selection, i.e. the track reconstruction efficiency (Section 7.5.3). The track-level corrections comprise additional factors to account for background contaminations and migration effects. Higher efficiencies are generally desirable, as larger correction factors to compensate for the inefficiencies tend to introduce larger uncertainties to the measurements.

7.5.1 Trigger Efficiency

The trigger efficiency, $\varepsilon_{\mathrm{trig}}$, of the `L1_MBTS_1` (`L1_MBTS_2`) trigger is calculated from the rate of events triggered by a space-point control trigger, `EF_mbSpTrk`, which also passed the `L1_MBTS_1` (`L1_MBTS_2`) trigger requirement. `EF_mbSpTrk` is a random trigger on colliding bunches at L1, which requires at least two Pixel and three SCT measurements at L2. The trigger efficiency of `L1_MBTS_1` is thus given by

$$\varepsilon_{\mathrm{trig}}(n_{\mathrm{sel}}^{\mathrm{BS}}) = \frac{\text{events triggered by } \texttt{EF\_mbSpTrk} \text{ and passing } \texttt{L1\_MBTS\_1}}{\text{events triggered by } \texttt{EF\_mbSpTrk}}. \tag{7.6}$$

This efficiency is parameterised by $n_{\mathrm{sel}}^{\mathrm{BS}}$, i.e. the number of tracks that were selected using modified impact parameter requirements (see Section 7.3.2). Although the requirements of the `EF_mbSpTrk` trigger are very loose, its efficiency was tested on a sample of fully unbiased events obtained with a purely random `L1_RD0` trigger. The results of this test are shown in Figure 7.17, where indeed no indication of any inefficiency of the `EF_mbSpTrk` trigger for events used in the minbias analysis can be seen. The `EF_mbSpTrk` trigger did not fire in only one event for which

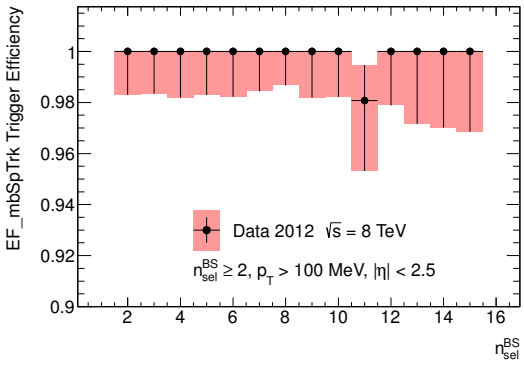


Figure 7.17: Efficiency of the `EF_mbSpTrk` trigger as a function of $n_{\text{sel}}^{\text{BS}}$, measured using 8 TeV data events obtained with a random trigger. `EF_mbSpTrk` is a random trigger on colliding bunches at L1, which requires at least two Pixel and three SCT measurements at L2. The total uncertainty, obtained with a Bayesian method, is represented by shaded areas.

some tracks were later reconstructed, which was interpreted as a beam-halo event.[12] The errors are relatively large due to the small recorded event statistics for the sample selected by the `L1_RD0` trigger. This particular sample of randomly triggered events is therefore not suitable for the other trigger efficiency studies.

The trigger efficiency for the `L1_MBTS_1` requirement, in which at least one signal in the MBTS modules is above threshold, is presented in Figures 7.18a and 7.18b; in the most inclusive phase space, it quickly converges to maximum efficiency from $n_{\text{sel}}^{\text{BS}} \geq 7$ onwards. The trigger efficiency for `L1_MBTS_2` is shown in Figures 7.18c and 7.18d; it requires at least two signals above threshold in the MBTS counters, leading to a slower rise to maximum efficiency. From these results, a trigger efficiency correction factor can be derived for each event in the data sample, which is parameterised by $n_{\text{sel}}^{\text{BS}}$, i.e. applying the same selection criteria as were used for the trigger efficiency determination.[13] Details of this correction are given in Section 7.6.1.

Biases on p_{T} and η distributions due to the trigger selection were studied and found to be negligible in comparison with other systematic effects. The possibility of a systematic effect due to the overall increase of μ during the run was found to be negligible, as described in Appendix A.1.

7.5.1.1 Systematic Uncertainties

In order to evaluate systematic uncertainties on the trigger efficiencies, the $|d_0^{\text{BS}}|$ selection criterium of the tracks utilised to calculate $n_{\text{sel}}^{\text{BS}}$ was either removed or made more restrictive. The difference between the efficiencies obtained with these modified selections was symmetrised and applied as a systematic uncertainty. Following this procedure, a total systematic uncertainty

[12] Beam-halo events occur when an outlier particle in a bunch passing through ATLAS hits a part of the detector, thus producing a spray of particles.

[13] For all events in the `EF_mbMbts_1_NoAlg` trigger stream, an efficiency correction based on `L1_MBTS_1` was applied; for events that were not contained in the first category, but were recorded by the `EF_mbMbts_2_NoAlg` trigger stream, an efficiency correction based on `L1_MBTS_2` was applied.

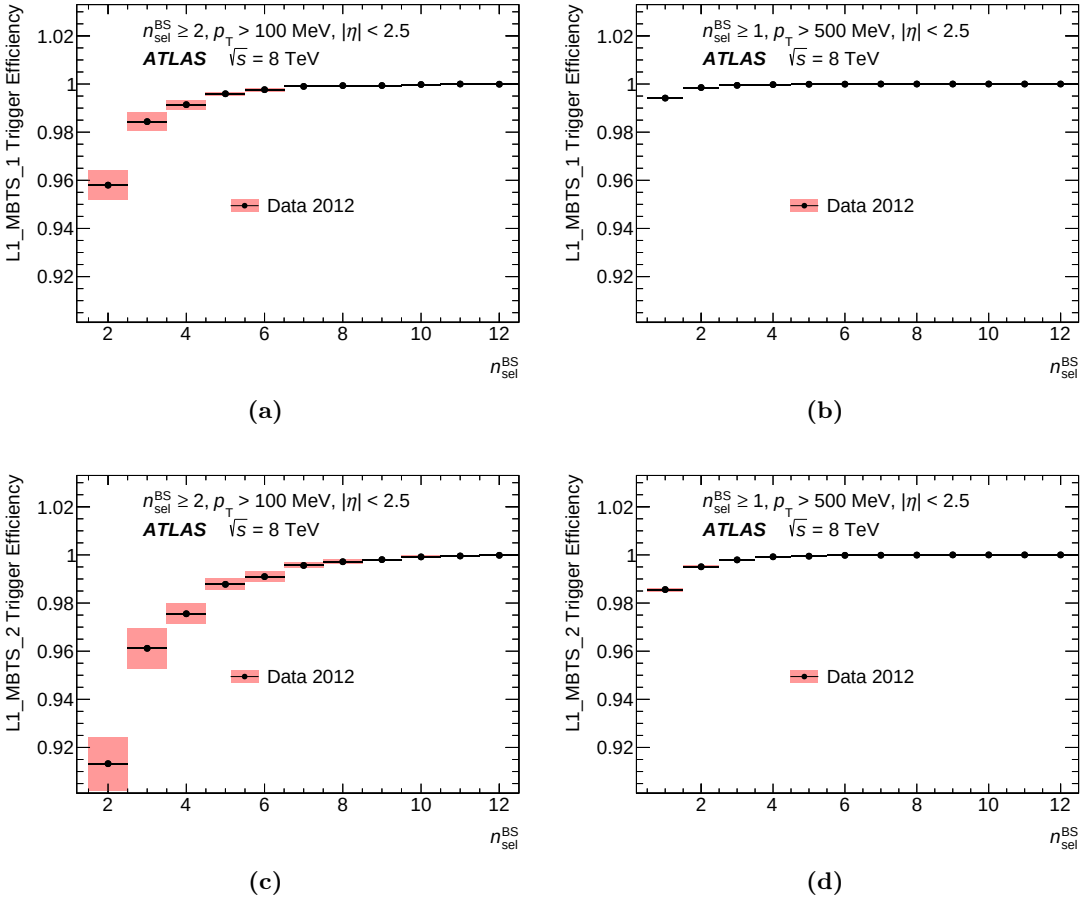


Figure 7.18: Comparison of the (a,b) `L1_MBTS_1` and (c,d) `L1_MBTS_2` trigger efficiencies as a function of (a,c) $n_{\mathrm{sel}}^{\mathrm{BS}}$ (selected tracks with $p_{\mathrm{T}} > 100$ MeV) and (b,d) $n_{\mathrm{sel},500}^{\mathrm{BS}}$ (selected tracks with $p_{\mathrm{T}} > 500$ MeV) for the two main phase spaces used in the analysis. The `L1_MBTS_1` (`L1_MBTS_2`) trigger requires that one (two) or more signals above threshold were registered by the MBTS modules. The total uncertainty is represented by shaded areas. Plots taken from Ref. [3].

in the $n_{\mathrm{ch}} \geq 2, p_{\mathrm{T}} > 100$ MeV phase space of approximately 0.7 % for events with $n_{\mathrm{sel}}^{\mathrm{BS}} = 2$, and decreasing rapidly at higher multiplicities, is applied to the `L1_MBTS_1` trigger efficiency. In the $n_{\mathrm{ch}} \geq 1, p_{\mathrm{T}} > 500$ MeV phase space this systematic uncertainty is even less than 0.1 % in the $n_{\mathrm{sel}}^{\mathrm{BS}} = 1$ event category. For the `L1_MBTS_2` triggers, the systematic uncertainties are found to be approximately twice as large as for the `L1_MBTS_1` trigger.

7.5.2 Vertex Reconstruction Efficiency

Primary vertices are found using reconstructed tracks as described in Section 5.2. However, for some events with very low multiplicity no vertex can be reconstructed, leading to a loss of the event in the analysis. Therefore a vertex reconstruction efficiency, $\varepsilon_{\mathrm{vtx}}$, is calculated separately for data and MC samples, and parameterised by $n_{\mathrm{sel}}^{\mathrm{BS}}$ in exactly the same way as the

trigger efficiency. This modified selection allows to account for events which do not contain a reconstructed primary vertex. Beam-induced background events are suppressed during the efficiency calculation, using the criteria from Eq. 7.1. The efficiency is defined as the ratio of events selected by the MBTS trigger which also have a reconstructed vertex to all events that passed the trigger selection:

$$\varepsilon_{\text{vtx}}(n_{\text{sel}}^{\text{BS}}) = \frac{\text{events passing the MBTS trigger with a reconstructed vertex}}{\text{events passing the MBTS trigger}}. \tag{7.7}$$

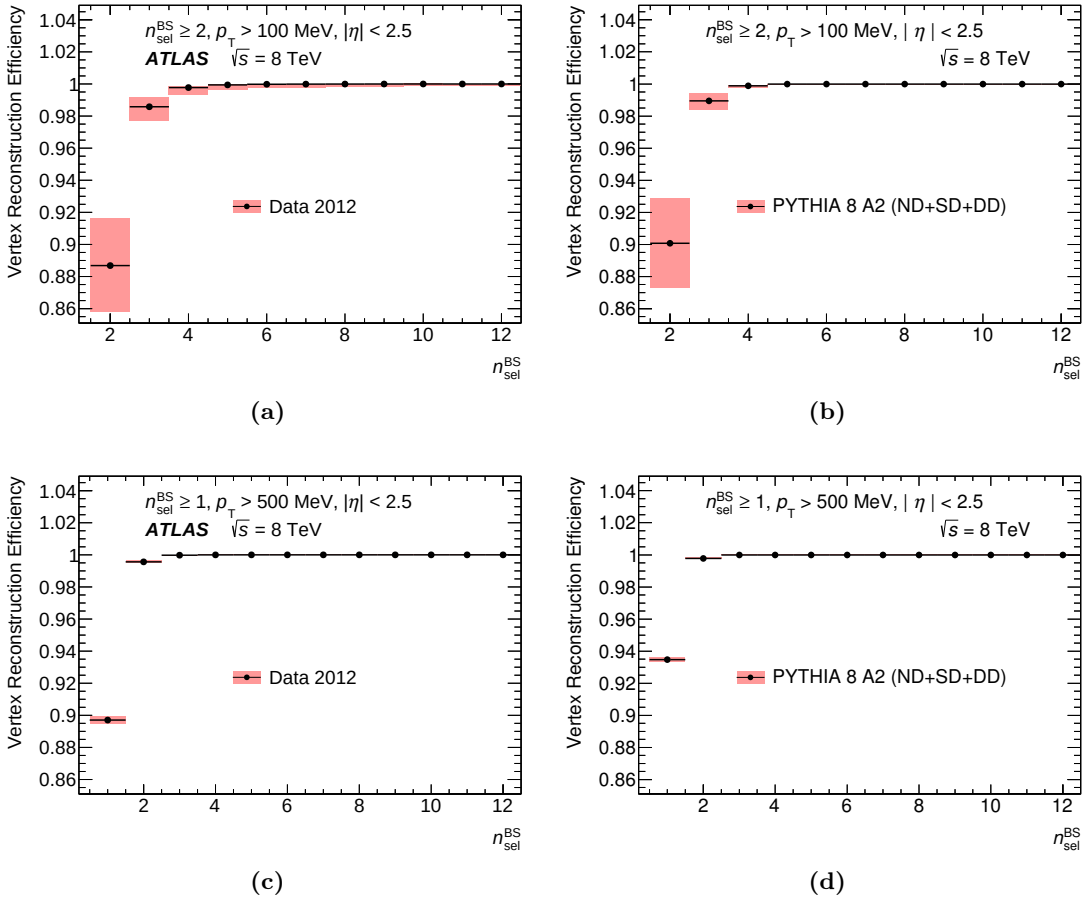


Figure 7.19: Vertex reconstruction efficiency as a function of n_{sel} with final systematic uncertainties, in the $p_{\text{T}} > 100$ MeV, $n_{\text{ch}} \geq 2$ phase space (a) for data and (b) for MC simulation, and in the $p_{\text{T}} > 500$ MeV, $n_{\text{ch}} \geq 1$ phase space (c) for data and (d) for MC simulation. The total uncertainty is represented by shaded areas. Plots taken from Ref. [3].

The final result for the vertex reconstruction efficiency in 8 TeV data and MC simulation is shown in Figure 7.19 with systematic uncertainties that are further described in Section 7.5.2.2. The average efficiency in the $p_{\text{T}} > 100$ MeV phase space is approximately 89 % for data and 90 % for Pythia 8 A2 MC simulation in the $n_{\text{sel}}^{\text{BS}} = 2$ event category, and quickly converges

to maximum efficiency as the multiplicity increases. In the $p_{\mathrm{T}} > 500$ MeV phase space, the average efficiency is close to 90 % for data and 93.5 % for Pythia 8 A2 MC simulation in the $n_{\mathrm{sel}}^{\mathrm{BS}} = 1$ event category, converging even faster to 100 % with rising $n_{\mathrm{sel}}^{\mathrm{BS}}$.

Appendix A.2 shows a comparison of vertex reconstruction efficiencies between different MC tunes and data at 8 TeV, as well as a comparison of efficiencies in data measured at different centre-of-mass energies between 0.9 and 8 TeV.

The vertex reconstruction depends not only on the multiplicity, but also the event topology and specific kinematic properties of reconstructed tracks (for instance, tracks with higher momenta can be better extrapolated to the vertex), such as $\Delta z_0^{\mathrm{min}}$ (the minimal difference between the longitudinal impact parameters of track pairs within an event), $p_{\mathrm{T}}^{\mathrm{min}}$ (the minimal p_{T} of selected tracks in the event) and the track pseudorapidity η (for events with $n_{\mathrm{sel}}^{\mathrm{BS}} = 1$ in the $p_{\mathrm{T}} > 500$ MeV phase spaces).

In previous studies it was found that in the $p_{\mathrm{T}} > 100$ MeV phase space, the $\Delta z_0^{\mathrm{min}}$ dependence in low-multiplicity events increases at lower $p_{\mathrm{T}}^{\mathrm{min}}$. This observation has been confirmed at $\sqrt{s} = 8$ TeV, and thus for events with $n_{\mathrm{sel}}^{\mathrm{BS}} = 2$, a special $\Delta z_0^{\mathrm{min}}$ parameterisation is used for the vertex reconstruction efficiency (subdivided into corrections for events with $100 < p_{\mathrm{T}}^{\mathrm{min}} < 200$ MeV in 25 MeV intervals, and for events with $p_{\mathrm{T}}^{\mathrm{min}} > 200$ MeV).[14] Detailed results for these parameterisations are shown in Appendix A.2.3.

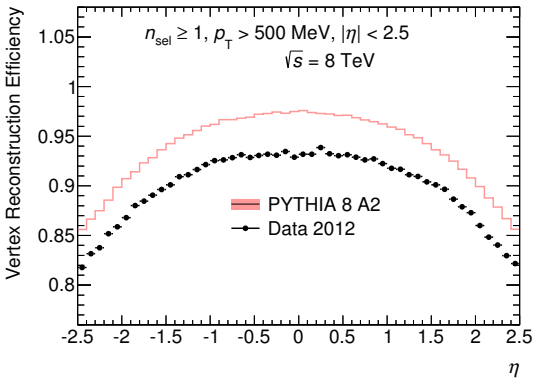


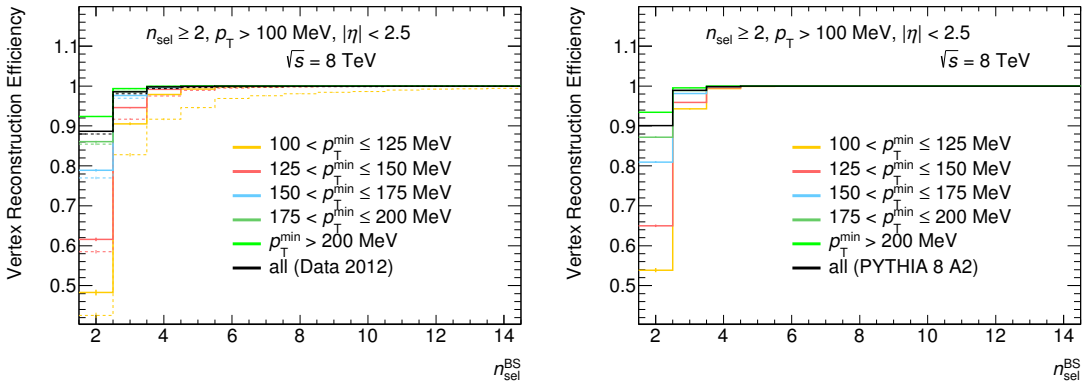
Figure 7.20: Vertex reconstruction efficiency for $n_{\mathrm{sel}}^{\mathrm{BS}} = 1$ as a function of η for data and MC simulation in the $p_{\mathrm{T}} > 500$ MeV, $n_{\mathrm{ch}} \geq 1$ phase space. The errors represent statistical fluctuations in the data sample.

For events with $n_{\mathrm{sel}}^{\mathrm{BS}} = 1$ in the $p_{\mathrm{T}} > 500$ MeV phase space, the vertex reconstruction efficiency is found to vary with η of the selected track as shown in Figure 7.20, with values between 81–94 % for data and 85–98 % for Pythia 8 A2 MC simulation; this dependence is accounted for by a special η parameterisation of the efficiency. Events with $n_{\mathrm{sel}}^{\mathrm{BS}} = 0$ have a number of selected tracks $n_{\mathrm{sel}} = 0$ as well, and are thus ignored as they do not contribute to the final event selection in either of the phase spaces considered in the analysis.

[14] For events with $n_{\mathrm{sel}}^{\mathrm{BS}} \geq 3$, the $\Delta z_0^{\mathrm{min}}$ dependence is already found to be negligibly small, as events with large $\Delta z_0^{\mathrm{min}}$ and no reconstructed vertex become increasingly rare with rising track multiplicities. Therefore no special Δz parameterisation was used for $n_{\mathrm{sel}}^{\mathrm{BS}} \geq 3$ events.

7.5.2.1 Differences to previous analyses

In the previous analyses up to $\sqrt{s} = 7$ TeV only the vertex reconstruction efficiency for data was considered; this is now improved by applying individual efficiencies for data and MC simulation, leading to better results in closure tests (see Section 7.7). Furthermore, the strong dependence on $p_{\mathrm{T}}^{\mathrm{min}}$ between $100 - 200$ MeV for low-multiplicity events in the data sample, which was shown to result from residual beam-induced background, was studied in more detail. As a consequence, the correction for events within $100 < p_{\mathrm{T}}^{\mathrm{min}} < 200$ MeV is now measured in 25 MeV intervals. This more refined subdivision is now also applied for the detailed $\Delta z_0^{\mathrm{min}}$ parameterisation at $n_{\mathrm{sel}}^{\mathrm{BS}} = 2$. The effect of the $p_{\mathrm{T}}^{\mathrm{min}}$ dependence is visualised in Figure 7.21 where the beam background in data is already suppressed.



(a) **(b)**

Figure 7.21: Vertex reconstruction efficiency as a function of n_{sel} in the $p_{\mathrm{T}} > 100$ MeV, $n_{\mathrm{ch}} \geq 2$ phase space in different $p_{\mathrm{T}}^{\mathrm{min}}$ intervals (a) for data and (b) for MC simulation. The dashed lines show the efficiency without beam-background removal in the data sample.

Another change from the old to the new analysis is the new choice of the FTFP physics list for MC simulation. The impact of this choice on the vertex reconstruction efficiency was found to be $< 0.001\,\%$ and was therefore neglected in the analysis.

7.5.2.2 Systematic Uncertainties

The systematic uncertainty on the vertex reconstruction efficiency comprises effects from two sources: (a) the influence of beam-induced background, and (b) variations of the applied impact parameter requirements. The first category can be assessed by comparing the nominal efficiency, in which the beam-induced background is suppressed by the criteria from Eq. 7.1, with an alternative efficiency in which these criteria were not applied. This strategy is similar to the method used in previous measurements [126, 136]. The difference in efficiencies with and without beam-induced background removal is shown in Figure 7.21a by the solid and dashed lines, respectively. This $p_{\mathrm{T}}^{\mathrm{min}}$-dependent difference is assigned as a one-sided systematic uncertainty, with the largest contribution of $0.7\,\%$ in the $n_{\mathrm{sel}}^{\mathrm{BS}} = 2$ event category of the $p_{\mathrm{T}} > 100$ MeV phase

space. In the $p_{\mathrm{T}} > 500$ MeV phase space, both the rate of beam-induced background and its impact on the vertex reconstruction efficiency are found to be negligible.

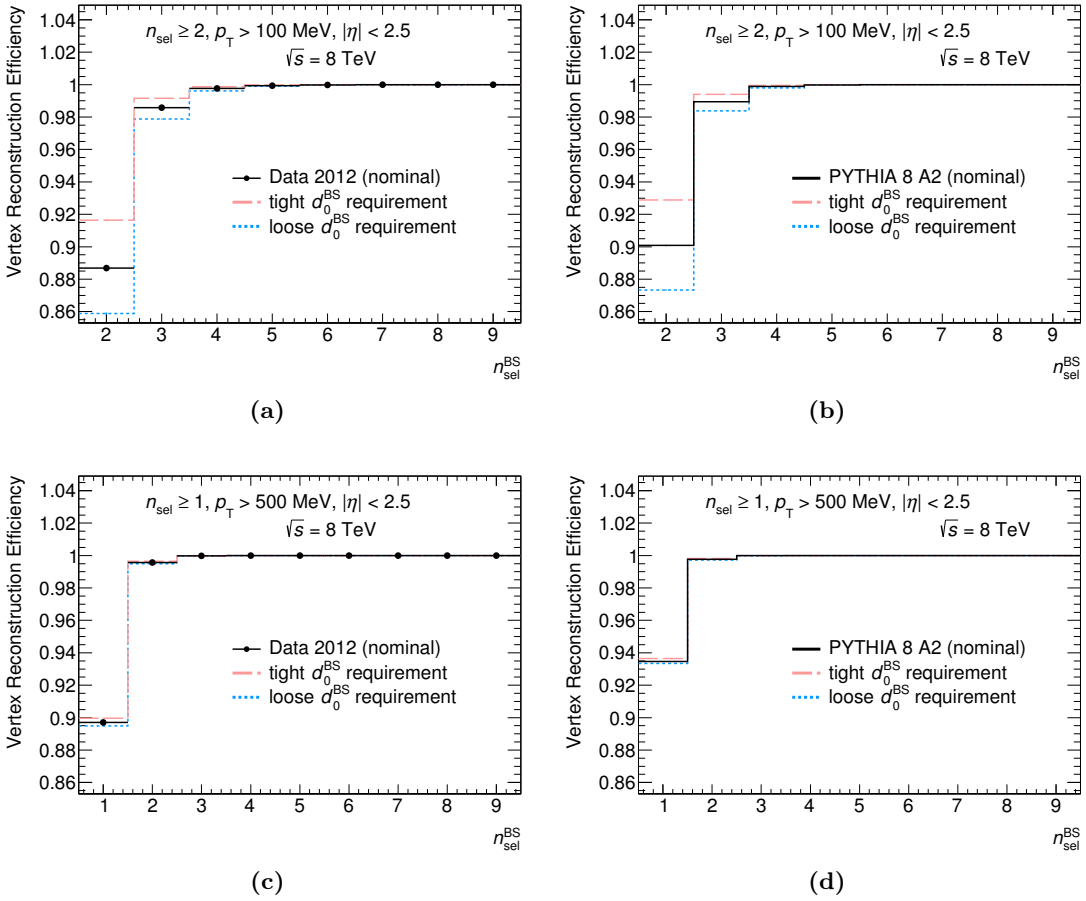


Figure 7.22: Vertex reconstruction efficiency as a function of n_{sel} with variations of the impact parameter requirements, in the $p_{\mathrm{T}} > 100$ MeV, $n_{\mathrm{ch}} \geq 2$ phase space (a) for data and (b) for MC simulation, and in the $p_{\mathrm{T}} > 500$ MeV, $n_{\mathrm{ch}} \geq 1$ phase space (c) for data and (d) for MC simulation. The statistical errors are comparable with the size of the line.

The difference in the efficiency due to a variation of the impact parameter requirement, from $|d_0^{\mathrm{BS}}| < 1.8\,\mathrm{mm}$ to $|d_0^{\mathrm{BS}}| < 1.3\,\mathrm{mm}$ (tight requirement) and $|d_0^{\mathrm{BS}}| < 2.5\,\mathrm{mm}$ (loose requirement), provides a conservative estimate for an additional source of systematic uncertainty, related to non-primary effects, with results shown in Figure 7.22. In the $p_{\mathrm{T}} > 100$ MeV (500 MeV) phase space, the largest difference at $n_{\mathrm{sel}}^{\mathrm{BS}} = 2$ ($n_{\mathrm{sel}}^{\mathrm{BS}} = 1$) is found to be 2.9 % (0.3 %), and it decreases rapidly with rising multiplicity.

All sources of systematic uncertainty are added in quadrature, and shown as asymmetrical errors in Figure 7.19. The treatment of the impact of this total uncertainty on final distributions is described in Section 7.5.3.2.

7.5.3 Track Reconstruction Efficiency

The track reconstruction efficiency, ε_{trk}, is applied as a correction in order to compensate for any effects (such as detector material, selection efficiencies, or track parameter resolutions) which decrease the probability of a charged particle to be successfully reconstructed and to pass the selection criteria. This factor was calculated from the PYTHIA 8 A2 MC sample and parameterised by p_{T} and η. The efficiency is defined by the ratio of the number of all reconstructed tracks uniquely *matched to* (i.e. associated with) generated charged primary particles, $N_{\text{rec}}^{\text{matched}}$, to the number of all generated charged primary particles, N_{gen}:

$$\varepsilon_{\text{trk}}(p_{\text{T}}, \eta) \;=\; \frac{N_{\text{rec}}^{\text{matched}}(p_{\text{T}}^{\text{gen}}, \eta^{\text{gen}})}{N_{\text{gen}}(p_{\text{T}}^{\text{gen}}, \eta^{\text{gen}})}, \tag{7.8}$$

where $p_{\text{T}}^{\text{gen}}$ and η^{gen} are parameters of the generated particles. The track–particle association is performed with a cone-matching algorithm as described in Section 7.2.4.

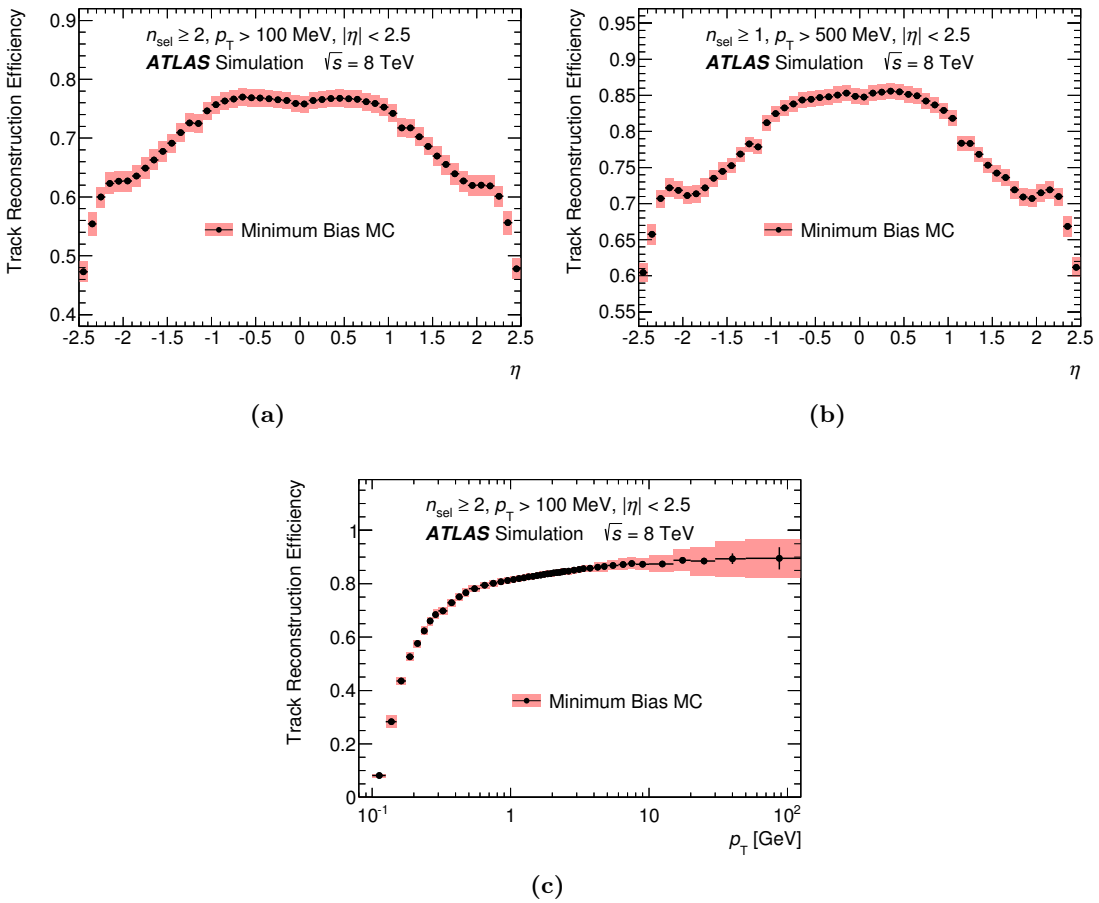
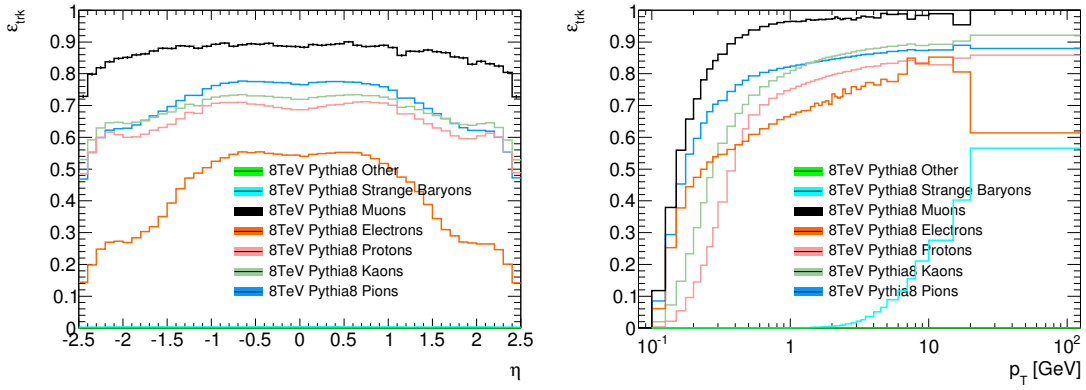


Figure 7.23: Track reconstruction efficiency as a function of (a-b) η and (c) p_{T} of the track, shown in the $p_{\text{T}} > 100$ MeV and 500 MeV phase space. The shaded areas represent the sum of systematic and statistical errors. Plots taken from Ref. [3].

Figure 7.23 shows the η and p_{T} distributions of the track reconstruction efficiency, including all systematic uncertainties that are discussed in Section 7.5.3.2. The η-dependent shape of the efficiency distribution corresponds to the amount of detector material traversed by the primary charged particles, in particular that between the Pixel and SCT sub-detectors, as the presence of material increases the probability of hadronic interactions which undermine the ability to successfully reconstruct tracks (see Section 6.1).



(a) **(b)**

Figure 7.24: Track reconstruction efficiency of different particle types as a function of (a) η or (b) p_{T} of the track in the $p_{\mathrm{T}} > 100$ MeV phase space.

A breakdown of track reconstruction efficiencies of different particle types, of which pions and kaons are dominating the overall charged-particle composition, is illustrated in Figure 7.24. As the generated-particle composition differs between MC tunes, different overall track reconstruction efficiencies can be observed for the Pythia 8, Pythia 6 and Epos tunes. An efficiency comparison of these tunes is shown in Appendix A.3. A systematic uncertainty due to the particle composition is assigned to cover this effect (see Section 7.5.3.2). A comparison of track reconstruction efficiencies of positively and negatively charged particles was made as shown in Appendix A.3 and has negligible impact on the systematic uncertainty.

7.5.3.1 Differences to previous analyses

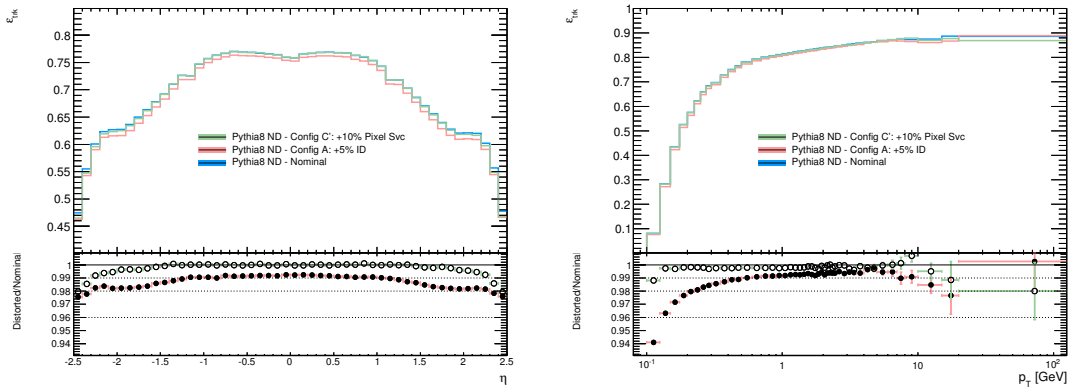
The assessment of systematic uncertainties has been extended with respect to previous analyses. Now the impact of the choice of physics list for MC simulation (which was changed from QGSP to FTFP) is taken into account, while the uncertainties due to vertex and trigger efficiency, particle composition and track quality selection have been parameterised more carefully, rather than treating them as flat uncertainties.

In the measurements at $\sqrt{s} = 8$ and 13 TeV all particles with a mean lifetime between $30 < \tau < 300$ ps are now considered to be non-stable, as described in Section 7.4.3, and were therefore also excluded from the evaluation of the track reconstruction efficiency, thus reducing the model dependence of the correction.

7.5.3.2 Systematic Uncertainties

The various sources of systematic uncertainty on the track reconstruction efficiency were evaluated in separate categories as follows.

Material The uncertainty due to incomplete knowledge of the ID material is the dominant factor of the total systematic uncertainty. It is calculated from a comparison between the nominal track reconstruction efficiency and the efficiency obtained from a MC sample with extra material in the ID simulation. The relative difference between these two efficiencies is symmetrised and taken as a systematic uncertainty. MC samples obtained with two alternative ATLAS geometries were used for this step: (a) ATLAS-GEO-21-04-02 which has a 5% increase of material in the whole ID (excluding sensitive detectors) and is denoted by *Config A*, and (b) ATLAS-GEO-21-03-02 which has a 10% increase of passive material in the Pixel services (SQP, InnerPP0 and OuterPP0) and is denoted by *Config C'*. These alternative geometries reflect the level of knowledge about the detector material budget at the end of Run-1, which was obtained by performing studies with photon conversions, hadronic vertices and the SCT extension efficiency [104, 165] (see Chapter 6).



(a) **(b)**

Figure 7.25: Track reconstruction efficiency as a function of (a) η or (b) p_{T} of the track, comparing distorted geometries (Config A and Config C') with the nominal geometry. The largest relative difference is taken as a systematic uncertainty contribution due to the material uncertainty.

Figure 7.25 shows the comparison of the track reconstruction efficiency between the distorted geometries *Config A* and *Config C'* and the nominal geometry. The largest difference is found for *Config A*; this difference is evaluated in p_{T} and η regions as given in Table 7.11, and taken as a systematic uncertainty due to imperfect knowledge of the ID material. The largest deviations between the nominal and distorted geometries can be found at large pseudorapidity ($|\eta| > 2.3$), where the amount of passive material in the ID is significantly larger than in the remaining acceptance region, and at either small ($100 < p_{\mathrm{T}} < 200$ MeV) or large ($p_{\mathrm{T}} > 10$ GeV) transverse

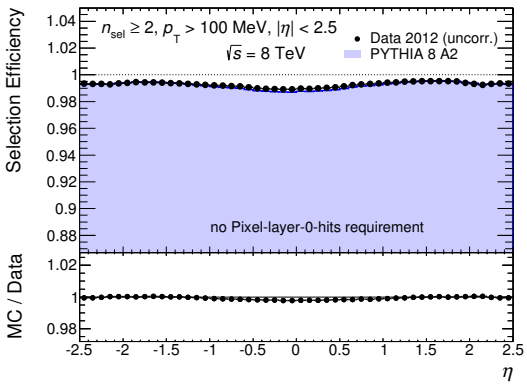
p_{T} [GeV]	$\lvert\eta\rvert < 1.1$	$1.1 \leq \lvert\eta\rvert < 1.5$	$1.5 \leq \lvert\eta\rvert < 1.7$	$1.7 \leq \lvert\eta\rvert < 2.3$	$2.3 \leq \lvert\eta\rvert \leq 2.5$
$0.1 - 0.125$	7	7	7	6	8
$0.125 - 0.15$	3.5	4	4.5	4.5	5
$0.15 - 0.2$	2.5	3	3.5	3.5	4.5
$0.2 - 0.25$	2	2	2.5	3	4.5
$0.25 - 0.3$	1.5	1.5	2	2.5	3.5
$0.3 - 0.4$	1	1.5	2	2.5	3.5
$0.4 - 0.5$	1	1.5	1.5	2	2.5
$0.5 - 1$	1	1	1.5	1.5	2
$1 - 2$	1	1	1	1.5	2
$2 - 6$	1	1	1	1.5	1.5
$6 - 10$	1	1	1.5	1.5	5
$10 - 20$	2	3	4	4	5
> 20	3.5	4	4.5	5	5

Table 7.11: Systematic uncertainties (in %) on the track reconstruction efficiency, originating from incomplete knowledge of the ID material.

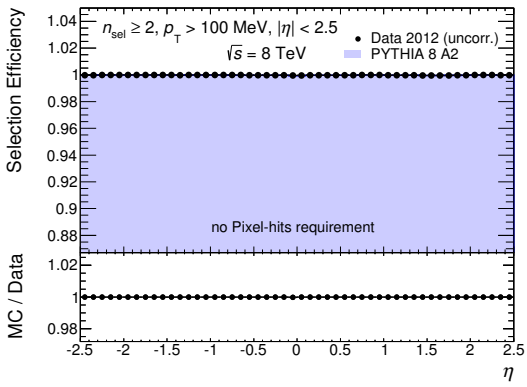
momentum, of which the latter effect may be an artefact due to the large statistical uncertainty in the high-p_{T} region, caused by the limited event statistics in the alternative MC samples.

Track Selection The systematic uncertainties from different effects of track quality selection cuts on data and MC simulation are evaluated by $N - 1$ tests. Here only one of the track selection cuts is either relaxed (e.g. IP cuts are relaxed from 1.5 to 2.5 mm) or entirely disabled, and the selection efficiency is calculated separately for data and MC simulation as the ratio between the number of tracks using all N cuts divided by the number of tracks using only $N - 1$ cuts. The deviations of the data/MC double ratio of these selection efficiencies from unity are taken as systematic uncertainties in suitably defined p_{T} and η bins. Finally all the systematic contributions from each of the N track quality cuts are taken in quadrature for each bin. Inclusive η and p_{T} distributions of selection efficiencies for the track quality cuts used in this analysis are shown in Figure 7.26 and Figure 7.27. Detailed examples of η distributions for the SCT hits requirement in various p_{T} bins are shown in Figure 7.28.

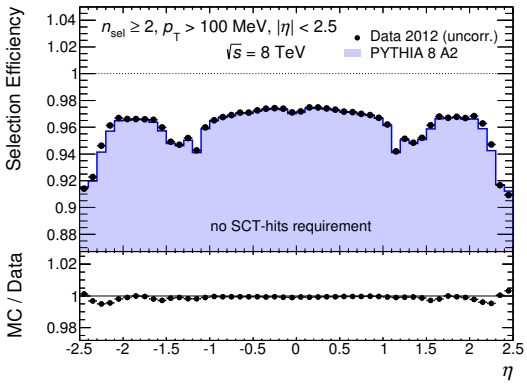
The final systematic uncertainties of track selection cuts were evaluated in much greater detail for $\sqrt{s} = 8$ TeV than in the previous analysis at 7 TeV, where a combined uncertainty of 1 % was assigned to the entire kinematic region [126]. Table 7.12 shows the uncertainties in η and p_{T} categories that were chosen after performing a careful analysis of the detailed results for each cut. Due to the limited high-p_{T} statistics, the quoted uncertainties for $p_{\mathrm{T}} > 10$ GeV were chosen as rather conservative estimates.



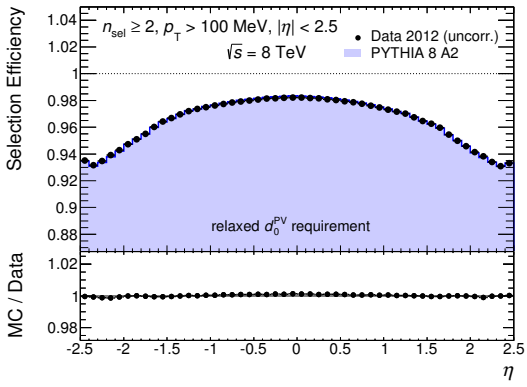
(a) Pixel layer-0 hits



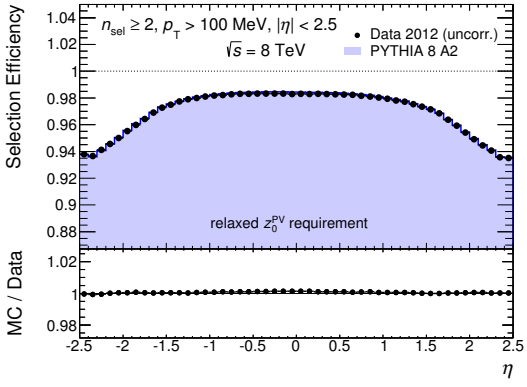
(b) Pixel hits



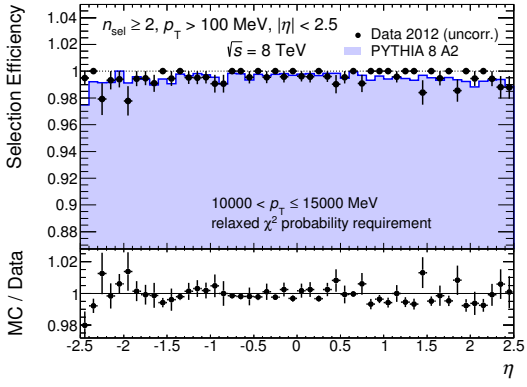
(c) SCT hits



(d) d_0^{PV}

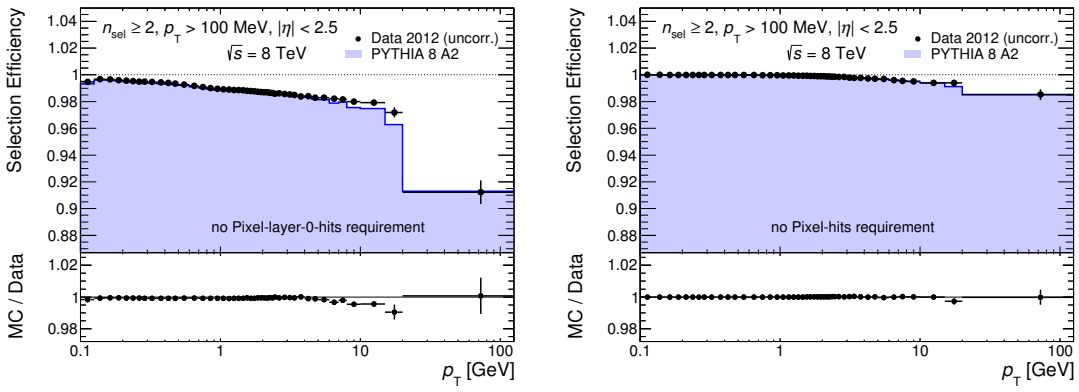


(e) z_0^{PV}



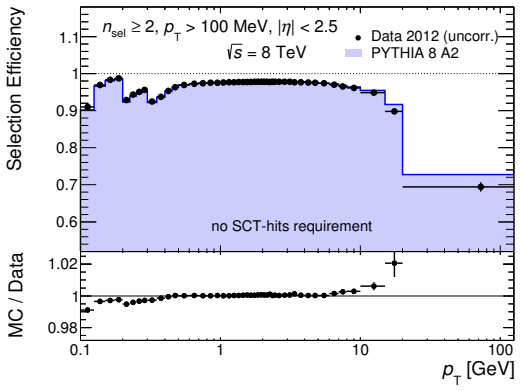
(f) χ^2 probability

Figure 7.26: Inclusive results in η for the systematic uncertainties on track quality cuts using $N-1$ tests: (a) no Pixel Layer-0 hit requirement, (b) no Pixel hits requirement, (c) no SCT hits requirement, (d) relaxed d_0^{PV} cut, (e) relaxed z_0^{PV} cut, (f) relaxed χ^2 probability cut (applied from $p_{\mathrm{T}} > 10$ GeV onwards).

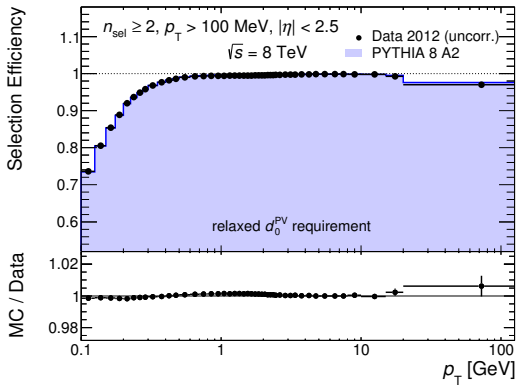


(a) Pixel layer-0 hits

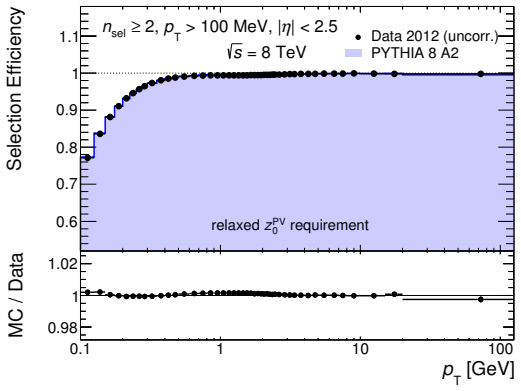
(b) Pixel hits



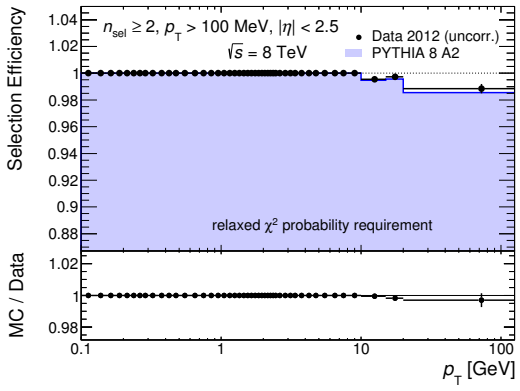
(c) SCThits



(d) d_0^{PV}

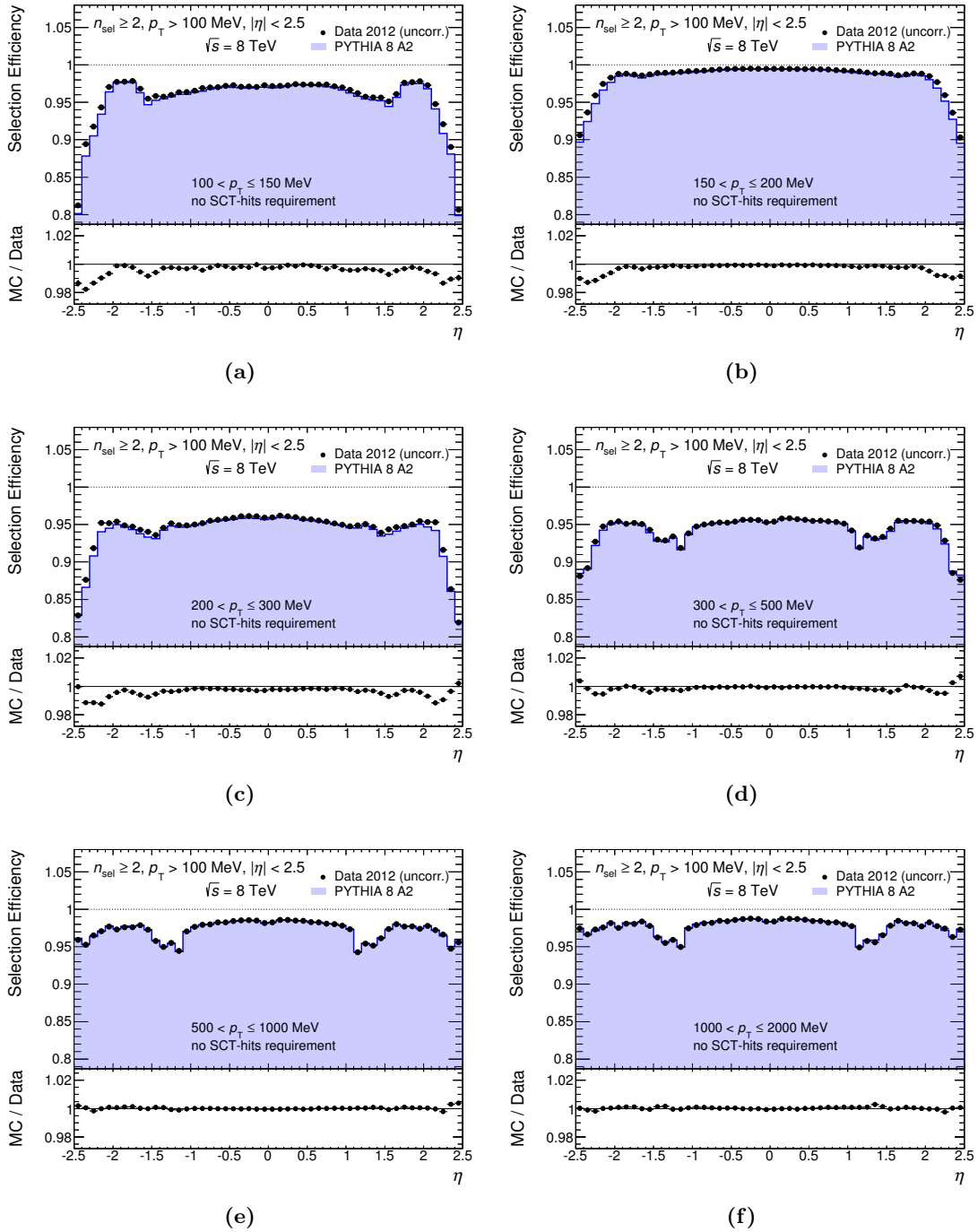


(e) z_0^{PV}



(f) χ^2 probability

Figure 7.27: Inclusive results in p_{T} for the systematic uncertainties on track quality cuts using $N-1$ tests: (a) no Pixel Layer-0 hit requirement, (b) no Pixel hits requirement, (c) no SCT hits requirement, (d) relaxed d_0^{PV} cut, (e) relaxed z_0^{PV} cut, (f) relaxed χ^2 probability cut (applied from $p_{\mathrm{T}} > 10$ GeV onwards).



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

Figure 7.28: Detailed results in η for the systematic uncertainties on track quality cuts using $N-1$ tests: no SCT hits requirement in p_{T} bins (a) $100 < p_{\mathrm{T}} < 150$ MeV, (b) $150 < p_{\mathrm{T}} < 200$ MeV, (c) $200 < p_{\mathrm{T}} < 300$ MeV, (d) $300 < p_{\mathrm{T}} < 500$ MeV, (e) $500 < p_{\mathrm{T}} < 1000$ MeV, (f) $1000 < p_{\mathrm{T}} < 2000$ MeV.

p_{T} [GeV]	$\lvert\eta\rvert < 1.1$	$1.1 \leq \lvert\eta\rvert < 1.5$	$1.5 \leq \lvert\eta\rvert < 1.8$	$1.8 \leq \lvert\eta\rvert < 2.3$	$2.3 \leq \lvert\eta\rvert \leq 2.5$
$0.1 - 0.15$	0.5	0.6	0.8	0.9	1.6
$0.15 - 0.2$	0.5	0.6	0.8	0.9	1.2
$0.2 - 0.3$	0.5	0.6	0.8	0.5	0.9
$0.3 - 0.5$	0.5	0.5	0.5	0.5	0.6
$0.5 - 1$	0.5	0.5	0.5	0.5	0.5
$1 - 2$	0.5	0.5	0.5	0.5	0.5
$2 - 6$	0.5	0.5	0.5	0.5	0.5
$6 - 10$	0.7	0.5	0.5	0.5	1.5
$10 - 15$	0.8	1.0	1.8	2.5	4.0
$15 - 20$	2.0	3.0	4.0	4.5	6.5
> 20	2.5	4.5	6.5	7.5	8.0

Table 7.12: Systematic uncertainties (in %) on the track reconstruction efficiency, originating from track selection cuts (d_0^{PV}, z_0^{PV}, Pixel Layer-0 hit, Pixel hits, SCT hits, χ^2 probability) used in the $\sqrt{s} = 8$ TeV analysis.

χ^2 probability cut The systematic uncertainties from the χ^2 probability cut are derived via the same $N - 1$ test method as described for the other track selection cuts. Here the χ^2 probability threshold for tracks with $p_{\mathrm{T}} > 10$ GeV is relaxed from 0.01 to 0.005, and again the deviations of the data/MC double ratio from unity are taken as systematic uncertainties. Due to small differences in χ^2 distributions of data and MC simulation, attributed to residual misalignment, a *fudge factor* of 1.01 is applied on the χ^2 values of MC tracks before calculating their χ^2 probabilities. This fudge factor ensures that the probability distributions of data and MC simulation are compatible with each other, particularly in the region affected by the cut, and thus safeguards that similar fractions of tracks are removed from both samples. Detailed results in various η regions are shown in Figure 7.29. The systematic uncertainties derived from this method are found to be 1 % for $10 < p_{\mathrm{T}} < 20$ GeV and 3 % for $p_{\mathrm{T}} > 20$ GeV. This is a smaller uncertainty than in the previous study at 7 TeV, where it was found to be 10 % for all tracks above $p_{\mathrm{T}} > 10$ GeV [126]. The reduction of this systematic uncertainty can be attributed to the improved detector alignment in the 2012 analysis.

p_{T} Resolution Migration effects due to momentum resolution differences between data and MC simulation were accounted for by an uncertainty of 5 % for low-transverse-momentum tracks between $100 < p_{\mathrm{T}} < 150$ MeV, while at high p_{T} a one-sided uncertainty of -7 % for tracks within $10 < p_{\mathrm{T}} < 30$ GeV and -9 % for tracks with $p_{\mathrm{T}} > 30$ GeV was applied to cover migration effects arising from the steeply falling p_{T} spectrum and the lower p_{T} resolution in data, in a

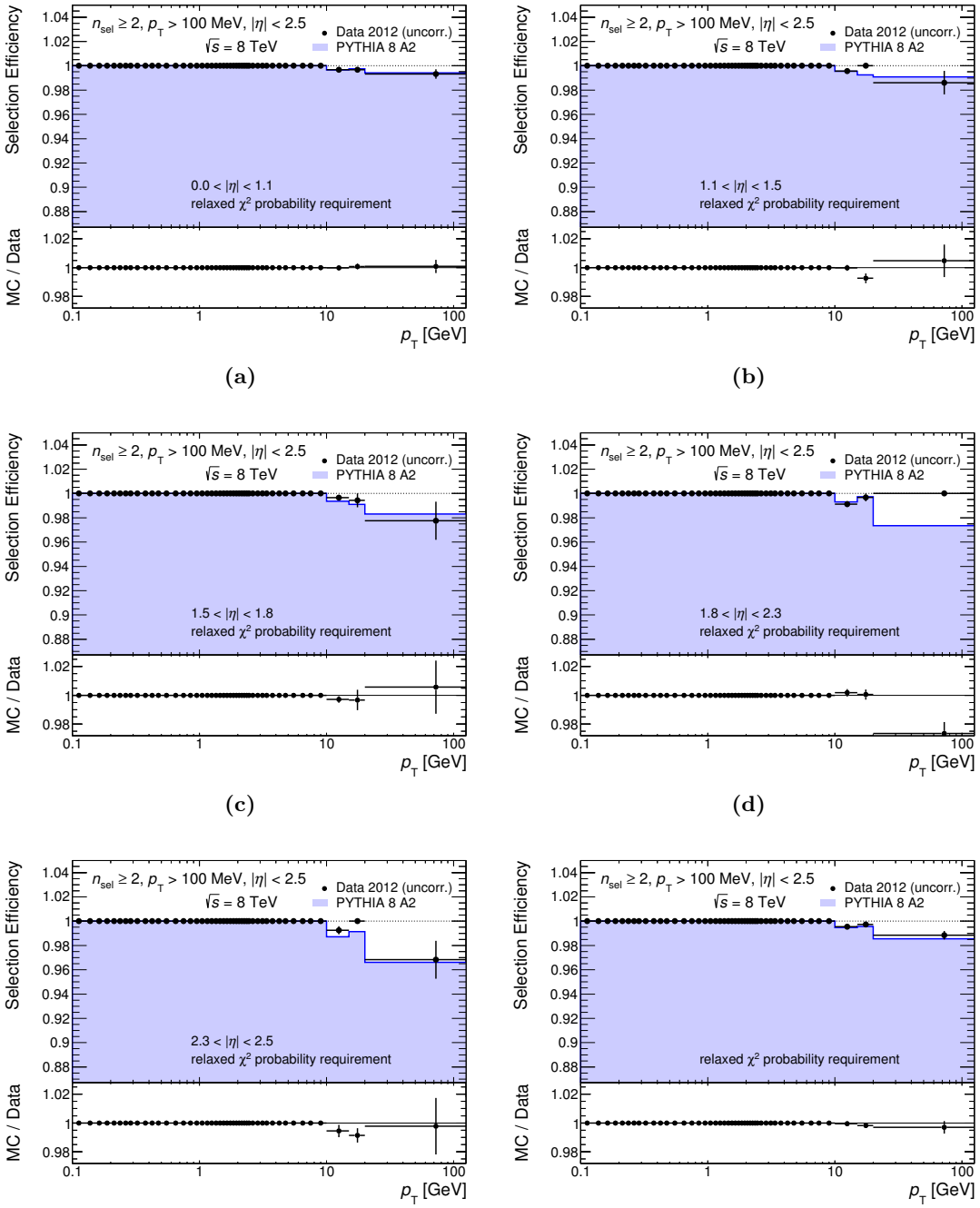


Figure 7.29: Detailed results in p_T for the systematic uncertainties on track quality cuts using $N-1$ tests: relaxed χ^2 probability cut in (a) $|\eta| < 1.1$, (b) $1.1 < |\eta| < 1.5$, (c) $1.5 < |\eta| < 1.8$, (d) $1.8 < |\eta| < 2.3$, (e) $2.3 < |\eta| < 2.5$, (f) inclusive result over the whole fiducial η range.

similar manner as for the 7 TeV analysis [126].[15] This one-sided uncertainty is combined linearly with the systematic errors from the following category.

Mis-measured high-p_{T} Tracks The fraction of tracks with grossly over- or underestimated p_{T} increases with p_{T} and η and becomes significant from $p_{\mathrm{T}} \gtrsim 5$ GeV onwards, even though the SCT hit requirement and the χ^2 probability cut mitigate this effect. The tails in the track momentum resolution correspond to a p_{T} migration in both directions, but due to the steeply falling p_{T} spectrum (and a longer non-Gaussian tail towards high-p_{T}) the net effect is a stronger migration of low-p_{T} particles to high-p_{T} reconstructed tracks. The actual fraction of mis-measured tracks can only be derived from MC simulation, where the $p_{\mathrm{T}}^{\mathrm{rec}}$ of the reconstructed track can be compared to the $p_{\mathrm{T}}^{\mathrm{gen}}$ of its associated (truth-matched) generated particle. A track is considered to be mis-measured if

$$|p_{\mathrm{T}}^{\mathrm{rec}} - p_{\mathrm{T}}^{\mathrm{gen}}|/p_{\mathrm{T}}^{\mathrm{gen}} > 0.5, \tag{7.9}$$

i.e. if the reconstructed transverse momentum differs by more than 50 % from the transverse momentum of its associated generated particle. This fraction rises with $|\eta|$ due to a degradation of the momentum resolution, caused by (a) the increasing amount of traversed material at larger pseudorapidity, which in turn increases the contributions from multiple scattering and hadronic interactions, and (b) the increasing flight distances of particles between the measurement layers.

p_{T} [GeV]	$-2.5 \leq \eta < -2.3$	$-2.3 \leq \eta < -1.7$	$\lvert\eta\rvert \leq 1.7$	$1.7 < \eta \leq 2.3$	$2.3 < \eta \leq 2.5$
$5-6$	1.8	0.5	0.1	0.5	2.5
$6-7$	3.5	0.7	0.2	0.7	3.5
$7-8$	4.5	1.0	0.2	1.0	4.5
$8-10$	7.5	1.2	0.3	1.4	6.0
$10-15$	9.0	1.6	0.4	1.6	9.0
$15-20$	26.5	4.5	0.8	5.0	15.0
$20-30$	39.0	9.0	2.2	9.5	30.0
$30-50$	88.0	28.0	6.5	22.0	60.0
50	96.0	75.0	28.5	48.0	92.0

Table 7.13: Estimated fraction (in %) of mis-measured tracks in MC simulation in bins of p_{T} and η. These values are added linearly with the p_{T} resolution uncertainty, and applied together as a relative systematic uncertainty on the track reconstruction efficiency.

[15] This systematic uncertainty for high-p_{T} tracks can be considered as very conservative, as the momentum resolution differences between data and MC simulation were found to be much better constrained with the 2012 alignment than in the previous study from 2010.

The extrapolation of mis-measured tracks back to their point of origin results in a badly reconstructed d_0^{PV} parameter, and hence this class of tracks exhibits more pronounced tails in the d_0^{PV} distributions. This can be exploited for an estimate of the rate of mis-measured tracks in data, by comparing the d_0^{PV} tails in data and MC simulation in various high-p_{T} and η bins. Differences in the tails are compensated by scaling the fractions of mis-measured tracks in MC simulation, which thus become the estimates of these fractions in data. At $\sqrt{s} = 7$ TeV scaling factors were required and the overall fraction of mis-measured tracks in data was found to be higher than in MC simulation, mainly due to misalignment. At $\sqrt{s} = 8$ TeV similar d_0^{PV} distributions in data and MC simulation were found, suggesting similar fractions of mis-measured tracks in both samples, and no scaling factors were needed.

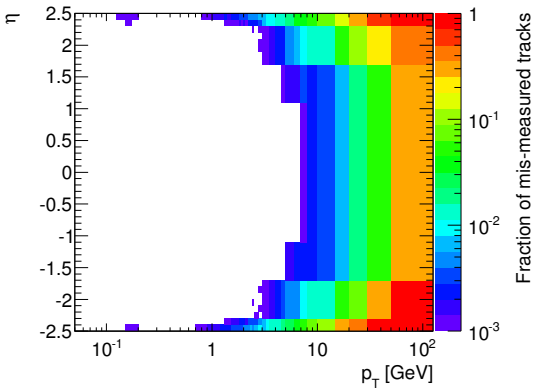


Figure 7.30: Log-scale distribution of fraction of mis-measured tracks, which are defined by $|p_{\mathrm{T}}^{\mathrm{rec}} - p_{\mathrm{T}}^{\mathrm{gen}}|/p_{\mathrm{T}}^{\mathrm{gen}} > 0.5$.

In the presented 8 TeV study the range of systematics from mis-measured tracks was extended from the previously used region, $p_{\mathrm{T}} > 10$ GeV, down to $p_{\mathrm{T}} > 5$ GeV, and a finer p_{T} and η binning was chosen. The estimated fractions in data, as determined from MC simulation, are shown in Table 7.13 and in Figure 7.30. Due to the small track statistics in the high-p_{T} regime, these results were used only to derive a systematic uncertainty, rather than for a correction of the charged-particle spectra.

p_{T} Spectrum As a different p_{T} spectrum of reconstructed particles is found in data and MC simulation (see uncorrected detector-level p_{T} distributions in Section 7.3.5), and the track reconstruction efficiency depends strongly on p_{T}, events belonging to a certain n_{ch} category can have different average track reconstruction efficiencies in data and MC simulation.

Figure 7.31 shows a comparison of the average track reconstruction efficiency (derived from MC simulation and parametrised by η and p_{T}) of all tracks in the data and various MC samples, respectively, as a function of n_{sel} of the event from which these tracks are taken. A systematic uncertainty due to the different p_{T} spectra is derived by symmetrising the absolute difference between data and MC simulation. This difference is applied as a systematic uncertainty only

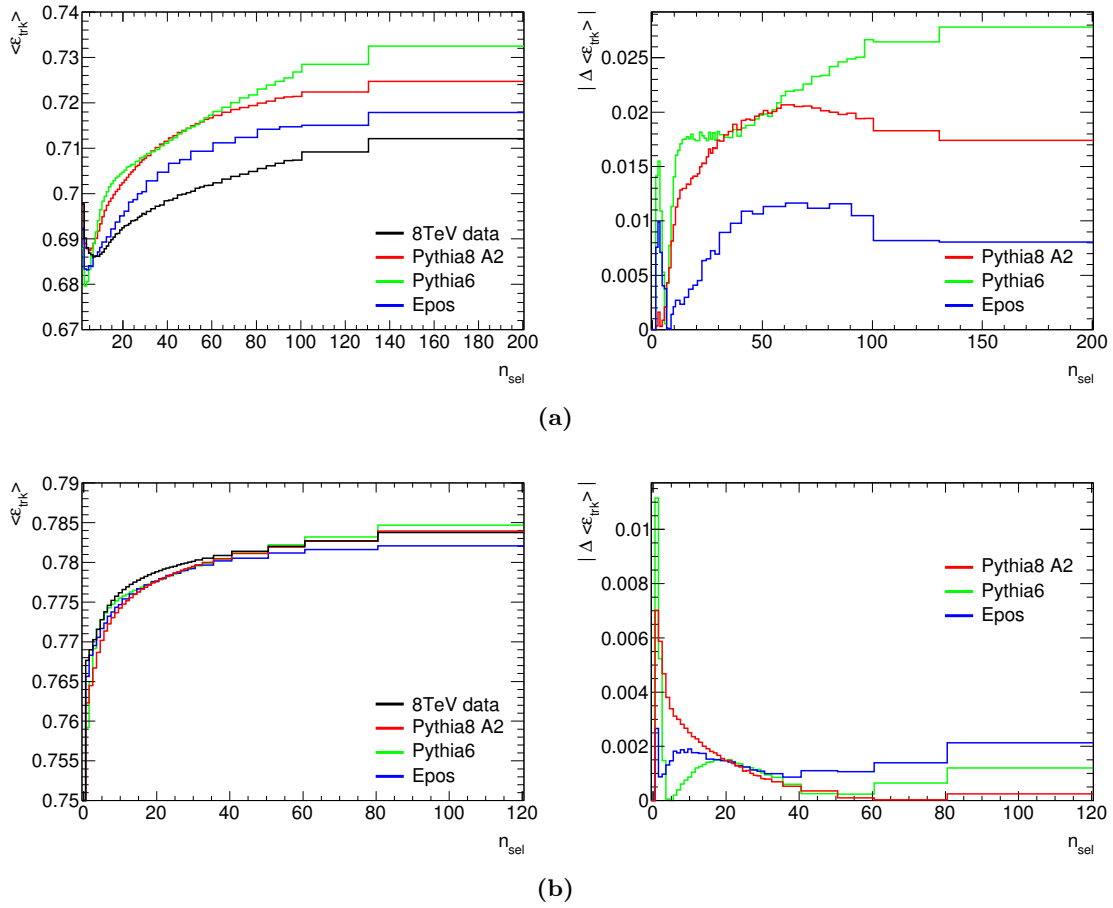


Figure 7.31: Comparison of the average track reconstruction efficiency (as derived from MC simulation and parametrised by η and p_{T}) of all tracks in the data and MC samples, respectively, as a function of n_{sel} of the event from which these tracks are taken (left), and systematics contribution due to the track reconstruction efficiency as a function of n_{sel} due to different p_{T} spectrum in data and MC simulation (right). Results are shown for the (a) $p_{\mathrm{T}} > 100$ MeV, $n_{\mathrm{ch}} \geq 2$ and (b) $p_{\mathrm{T}} > 500$ MeV, $n_{\mathrm{ch}} \geq 1$ phase spaces. In the analysis the absolute difference between data and MC simulation with Pythia 8 A2 is taken as a systematic uncertainty.

to the n_{ch}-dependent final distributions, i.e. the n_{ch} spectrum itself and the $\langle p_{\mathrm{T}} \rangle$ versus n_{ch} distribution.

The p_{T}-spectrum systematics can be reduced by choosing a MC tune with a better modelling of the p_{T} spectrum that is observed in the data. Figure 7.31 indicates that the Epos LHC tune, which models the low-p_{T} spectrum in the data quite well, would yield a smaller systematic uncertainty than the Pythia 8 A2 tune in the phase space with the $p_{\mathrm{T}} > 100$ MeV threshold. The Epos LHC tune thus appears to be a promising candidate model for the baseline correction procedure in future analyses.

Vertex and Trigger Efficiency The systematics due to vertex reconstruction and trigger efficiency uncertainties are parameterised by n_{sel} and applied as an uncertainty on the track

weights in an event, reaching values as high as 3 % at low multiplicity $n_{\text{sel}} = 2$ in the most inclusive $p_{\text{T}} > 100$ MeV phase space. They are combined into the "flat" category along with the following sources.

Truth Matching The systematic uncertainty due to the cone-matching algorithm that associates reconstructed tracks with generated particles (see Section 7.2.4) is conservatively estimated with 0.4 %, thus using the same flat uncertainty as in the previous measurement at 7 TeV. It is combined into the "flat" category along with the following sources.

Particle Composition and Choice of Physics List Different particle types have different reconstruction efficiencies (see Figure 7.24). The majority of stable charged particles produced in pp collisions are pions, kaons and protons. At transverse momenta of $p_{\text{T}} \sim 1$ GeV, their reconstruction efficiencies are approximately 82 %, ~ 80 % and ~ 75 %, respectively. However, different MC models generate different relative fractions of these particles. At the same transverse momenta, the fractions of the three particle types generated by Pythia 8 A2 are 77 %, 14 % and 9 %, whereas for the Epos LHC tune these fractions are 72 %, 18 % and 10 %. These differences in particle composition therefore result in a model-dependent uncertainty on the overall track reconstruction efficiency. This uncertainty is parameterised by the transverse momentum p_{T} of tracks and varies between 0.2 % and 1 %. Details on the differences of overall track reconstruction efficiencies between MC generators are shown in Appendix A.3 in Figures A.5 and A.6.

The impact of the new choice of the physics list for MC simulation, which was changed from FTFP to QGSP, on the track reconstruction efficiency is shown in Figure A.8 in Appendix A.3. The difference between the two physics lists is assigned as a flat systematic uncertainty of 0.3 %. The main source for this difference is the improved simulation of antiprotons in the FTFP physics list, resulting in a higher track reconstruction efficiency of these particles.

The systematic effects due to trigger and vertex reconstruction efficiency, truth matching, particle composition, and choice of physics list are combined into the "flat" systematic uncertainty category with n_{sel} and p_{T} dependence.[16]

Total systematic uncertainty All individual sources of systematic uncertainty on the track reconstruction, their parameterisations and affected phase-space regions are summarised in Table 7.14. Symmetric errors are used for all listed categories, except for p_{T} resolution effects and mis-measured high-p_{T} tracks. The shaded areas in Figure 7.23 show the total systematic uncertainty on the track reconstruction efficiency due to these categories, calculated by adding all contributions from the individual sources in quadrature, except for the asymmetric high-p_{T} effects which were not included in these plots. This total systematic uncertainty reaches highest

[16] In previous analyses, the same sources of uncertainty (except the new source of physics list) used to be combined into a non-parameterised "flat" uncertainty which was equally applied for all tracks, regardless of the particular event and track properties. In the current analysis, this has been changed to a more detailed treatment. The same category name is kept for legacy reasons.

Systematic Uncertainty	Param.	Phase Space	Size	Region
Material	p_{T}, η	all	± 1 to $\pm 8\,\%$	increases with $\lvert\eta\rvert$ and low/high p_{T}
Track Selection(*)	p_{T}, η	all	± 0.5 to $\pm 8\,\%$	increases with $\lvert\eta\rvert$ and low/high p_{T}
p_{T} Resolution	p_{T}	all	$\pm 5\,\%$ negligible $-7\,\%$ $-9\,\%$	$100 < p_{\mathrm{T}} < 150$ MeV $0.15 < p_{\mathrm{T}} < 10$ GeV $10 < p_{\mathrm{T}} < 30$ GeV $p_{\mathrm{T}} > 30$ GeV
Mis-measured high p_{T}	p_{T}, η	all	up to $-16\,\%$	only for $p_{\mathrm{T}} > 5$ GeV, increases with $\lvert\eta\rvert$ and p_{T}
p_{T} Spectrum	n_{sel}	$p_{\mathrm{T}} > 100$ MeV $p_{\mathrm{T}} > 500$ MeV	up to $\pm 2.1\,\%$ up to $\pm 0.7\,\%$	increases with n_{sel} decreases with n_{sel}
Vertex, Trigger	n_{sel}	$p_{\mathrm{T}} > 100$ MeV $p_{\mathrm{T}} > 500$ MeV	up to $\pm 3\,\%$ up to $\pm 0.2\,\%$	decreases with n_{sel}
Particle composition	p_{T}	all	± 0.2 to $\pm 1.0\,\%$	increases with low/high p_{T}
Physics List	–	all	$\pm 0.3\,\%$	flat
Truth Matching	–	all	$\pm 0.4\,\%$	flat

Table 7.14: List of sources of systematic uncertainty on the track reconstruction efficiency for $\sqrt{s} = 8$ TeV, including information on the parameterisation used for each contribution. (*) The track selection uncertainty includes the χ^2 probability cut, which by itself contributes 1–3 % to the overall uncertainty.

values in the low- and high-p_{T} regions as well as at large $|\eta|$ and is clearly dominated by the material uncertainty. The impact of all individual sources on the final distributions is discussed in Section 7.8.

7.6 Correction and Unfolding Procedure

In order to measure general observables which can be compared with other charged-particle density measurements, the reconstructed data distributions must be corrected for selection and reconstruction inefficiencies, model dependencies, migration effects, and contributions from background and secondary particles. This is done by applying event- and track-level weights to the distributions, followed by an unfolding procedure as described in the following.

7.6.1 Event- and Track-Level Weights

Every selected event is corrected with an event-level weight to compensate for "lost" events due to inefficiencies of the MBTS trigger selection (see Section 7.5.1) and the vertex reconstruction algorithm (see Section 7.5.2). Thus the total event weight, parameterised by $n_{\mathrm{sel}}^{\mathrm{BS}}$ and other event-related parameters, becomes:

$$w_{\mathrm{ev}}(n_{\mathrm{sel}}^{\mathrm{BS}}, x) = \frac{1}{\varepsilon_{\mathrm{trig}}(n_{\mathrm{sel}}^{\mathrm{BS}})} \cdot \frac{1}{\varepsilon_{\mathrm{vtx}}(n_{\mathrm{sel}}^{\mathrm{BS}}, x)} \cdot w_{\mathrm{zvrtx}}. \tag{7.10}$$

Here, x denotes a combination of $p_{\mathrm{T}}^{\mathrm{min}}$ of all selected tracks, $\Delta z_0^{\mathrm{min}}$ of track pairs in the event (for events with $n_{\mathrm{sel}} = 2$), and η (for events with $n_{\mathrm{sel}} = 1$). The factor w_{zvrtx} reweights the vertex z distribution of MC events to data, as described in Section 7.3.3. Eq. 7.10 simplifies for MC simulation due to $\varepsilon_{\mathrm{trig}} \equiv 1$ and for data due to $w_{\mathrm{zvrtx}} \equiv 1$, respectively.

Each selected track is corrected for the inefficiency of the track reconstruction algorithm (see Section 7.5.3) and three additional factors:

- f_{nonp} accounts for the background due to non-primary tracks (see Section 7.4.2);
- f_{sb} is the fraction of reconstructed tracks which have been associated with strange baryons (see Section 7.4.3);
- f_{okr} is the fraction of additional tracks, associated with particles from outside the selected kinematic region, that have migrated into the phase space due to the imperfect p_{T} and η resolution; this factor is determined from MC simulation.

These factors are combined into an individual track-level weight for each track as follows:

$$w_{\mathrm{trk}}(p_{\mathrm{T}}, \eta) = \frac{1}{\varepsilon_{\mathrm{trk}}(p_{\mathrm{T}}, \eta)} \cdot (1 - f_{\mathrm{nonp}}(p_{\mathrm{T}}, \eta) - f_{\mathrm{sb}}(p_{\mathrm{T}}, \eta) - f_{\mathrm{okr}}(p_{\mathrm{T}}, \eta)). \tag{7.11}$$

The distributions of all factors in Eq. 7.11 are presented in Figure 7.32. Differences in the momentum resolution between data and MC simulation contribute to a systematic uncertainty on f_{okr} at the low-p_{T} threshold, but this effect was found to be negligible in comparison to the systematic uncertainty due to the p_{T} resolution. Differences in the η resolution, as well as track–track correlations inside events, were found to be negligible in this analysis; the latter effect would only become significant for events with very high multiplicities of several hundred tracks.

7.6.2 Corrections to $\mathrm{d}N_{\mathrm{ev}}/\mathrm{d}n_{\mathrm{ch}}$

The correction procedure to obtain the charged-particle multiplicity distribution, $\mathrm{d}N_{\mathrm{ev}}/\mathrm{d}n_{\mathrm{ch}}$, as well as the other final distributions, follows the strategy from the previous analysis [126].

The multiplicity distribution of selected tracks, n_{sel}, is corrected with the event-level weights. Due to the track reconstruction inefficiency, n_{sel} is typically lower than the true number of primary particles, n_{ch}, that gave rise to the reconstructed tracks; in rare cases, n_{sel} can be higher than n_{ch} due to background contaminations. These relations can be described in a matrix, $M_{\mathrm{ch,sel}}$, as the probabilities that events with an observed track multiplicity n_{sel} (corrected with their event-level weights) correspond to events with various possible primary-particle multiplicities n_{ch}. As the true number of primary particles in the data is not known a priori, this matrix must be initially populated from MC simulation; in the presented study, the Pythia 8 A2 tune was used for this step. It is normalised such that the probabilities for each individual n_{sel} event category add up to unity, in oder to safeguard that the following procedure preserves

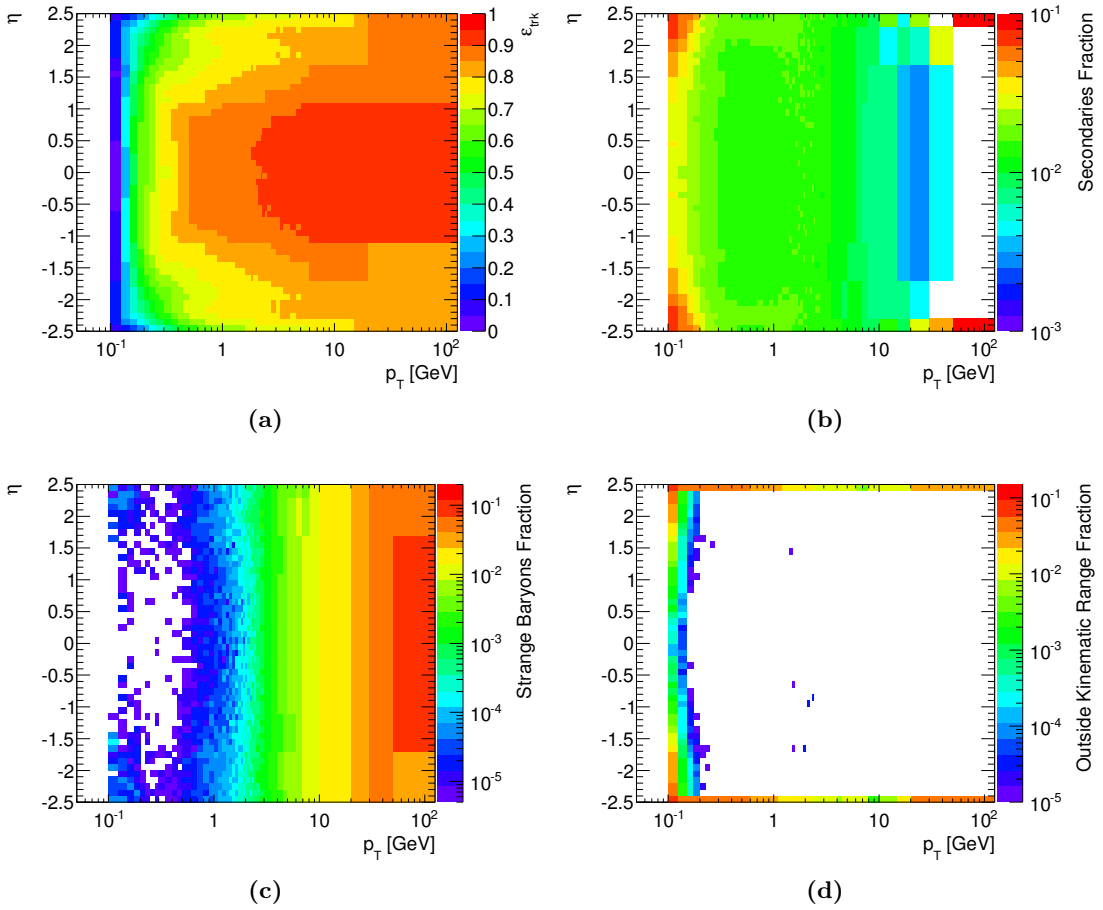


Figure 7.32: (a) Distribution of the track reconstruction efficiency, $\varepsilon_{\mathrm{trk}}(p_{\mathrm{T}}, \eta)$. (b) Log-scale distribution of fraction of non-primary tracks (shown here without scale factors), $f_{\mathrm{nonp}}(p_{\mathrm{T}}, \eta)$. (c) Log-scale distribution of fraction of tracks associated with strange baryons, $f_{\mathrm{sb}}(p_{\mathrm{T}}, \eta)$. (d) Log-scale distribution of fraction of selected tracks, $f_{\mathrm{okr}}(p_{\mathrm{T}}, \eta)$, which are migrating from outside the kinematic region into the phase space limited by $|\eta| < 2.5$ and $p_{\mathrm{T}} > 100$ MeV. The same distributions are also used for the $p_{\mathrm{T}} > 500$ MeV phase spaces. All distributions shown here are derived from MC simulation using the Pythia 8 A2 tune.

the total event yield, N_{ev}.[17] The matrix can then be used as the initial input to an iterative Bayesian unfolding [166] procedure. This technique allows to converge from an observed to a true distribution (in the given case, from n_{sel} to n_{ch}), typically after several iterations. In each subsequent iteration step the matrix $M_{\mathrm{ch,sel}}$ is applied to reweight the n_{sel} distribution in the data. Then the χ^2 difference between the distributions prior to and following this unfolding step is calculated to test for convergence of the method. If it has not yet converged, i.e. $\chi^2/n_{\mathrm{dof}} > 1$, the matrix is modified by weights derived from the resulting unfolded distribution, and the

[17] For some events with $n_{\mathrm{sel}} \geq n_{\mathrm{ch}}^{\mathrm{min}}$ above the acceptance limit $n_{\mathrm{ch}}^{\mathrm{min}}$, the true multiplicity $n_{\mathrm{ch}} < n_{\mathrm{ch}}^{\mathrm{min}}$ is found to be below the threshold. Such rare events will migrate out of the phase-space region during the unfolding procedure, but the effect is negligible with respect to other systematic uncertainties.

next iteration is performed. In the presented study, convergence was reached after five (four) iterations in the $p_{\mathrm{T}} > 100$ MeV (500 MeV) phase space.

However, this procedure cannot account for any events in which n_{sel} is below the threshold of acceptance in the phase space under study, $n_{\mathrm{sel}} < n_{\mathrm{ch}}^{\mathrm{min}}$, that were already filtered out at the event-selection stage prior to the unfolding. Among these, all events that would actually correspond to $n_{\mathrm{ch}} \geq n_{\mathrm{ch}}^{\mathrm{min}}$ charged particles have thus migrated out of the phase-space acceptance, mainly due to the track reconstruction inefficiency, and are effectively lost to the measurement. To account for these lost events in the $n_{\mathrm{ch}} \geq 2$ phase space, an additional correction term was applied to the unfolded n_{ch} distribution:

$$w_{\mathrm{lost}} = 1/(1 - (1 - \bar{\varepsilon}_{\mathrm{trk}})^{n_{\mathrm{ch}}} - n_{\mathrm{ch}} \cdot \bar{\varepsilon}_{\mathrm{trk}} \cdot (1 - \bar{\varepsilon}_{\mathrm{trk}})^{(n_{\mathrm{ch}}-1)}) \tag{7.12}$$

where $\bar{\varepsilon}_{\mathrm{trk}}$ is the mean effective track reconstruction efficiency for each given n_{ch} event category. A simpler correction term, $w_{\mathrm{lost}} = 1/(1 - (1 - \bar{\varepsilon}_{\mathrm{trk}})^{n_{\mathrm{ch}}})$, was used for the phase space with $n_{\mathrm{ch}} \geq 1$. In the analysis, only the efficiency $\bar{\varepsilon}_{\mathrm{trk}}$ corresponding to the event category with lowest multiplicity was applied, because variations in the different event categories were sufficiently small. Various other effects can bias the low-n_{ch} region of the multiplicity distribution; for MC studies such as closure tests (see Section 7.7), these effects can be accounted for by tuning $\bar{\varepsilon}_{\mathrm{trk}}$ (thus assuming that all given effects can be expressed by this single factor) such that a perfect closure is achieved in the lowest n_{ch} event category.

Phase Space		**Data**	**MC generator tune**		
$n_{\mathrm{ch}} \geq$	$p_{\mathrm{T}} >$		Pythia 8 A2	Pythia 6 AMBT2B	Epos LHC
2	100 MeV	0.675	0.662	0.658	0.668
1	500 MeV	0.808	0.779	0.782	0.778

Table 7.15: Mean effective track reconstruction efficiencies for data and various MC tunes, determined in events with the minimum amount of tracks that are required by the given phase space. These average efficiencies are used to correct the unfolded n_{ch} spectrum for events that have migrated out of the phase space.

Between different MC models, the tuned efficiency varies slightly due to different predictions of their p_{T} and n_{ch} spectra, combined with the p_{T}-dependence of the track reconstruction efficiency. For data the mean track reconstruction efficiency of all tracks falling into the lowest n_{sel} category of the phase space is used. An overview of the tuned efficiency factors for data and various MC tunes in the two main phase spaces considered in this analysis is given in Table 7.15.

7.6.2.1 Systematic Uncertainties

The systematic uncertainty due to the unfolding procedure is determined with the following sequence: (1) either the input distributions or the unfolding matrix are modified, (2) the same unfolding procedure is applied, and (3) the obtained alternative results are compared to the nominal results. This is done separately for various individual sources of this systematic uncertainty, which can be classified by the following categories:

1. systematic variations of the input distribution due to uncertainties of track weights, e.g. due to:
   a) material uncertainty on the track reconstruction efficiency
   b) track selection cuts
   c) fraction non-primary tracks and strange baryons
   d) trigger and vertex efficiency, particle composition, physics list, truth matching
   e) different p_{T} spectrum of reconstructed tracks in data and MC simulation
2. alternative input distributions, e.g.:
   a) reweighted MC distributions as pseudo-data (including a data-driven closure test)
   b) distributions from alternative MC tunes as pseudo-data
3. statistical variations of the unfolding matrix

In the first main category, the applied strategy is to create a modified n_{sel} distribution by randomly removing (or adding) tracks from the events according to the uncertainty on track weights. This is separately done for each sub-category: for instance, in the "material uncertainty" sub-category, each individual track in an event is randomly removed with a probability reflecting the material uncertainty on the track reconstruction efficiency, parameterised by the p_{T} and η of the track. The same principle is followed for the other sub-categories. The obtained modified n_{sel} distribution is then used as an alternative input to the same unfolding procedure that was described earlier. The systematic effect due to this source of uncertainty is finally taken as the difference between the alternative and nominal unfolded multiplicity distributions, and is symmetrised. In the most inclusive phase space given by $n_{\mathrm{ch}} \geq 2$ and $p_{\mathrm{T}} > 100$ MeV, the systematic effect due to the material uncertainty is found to be $< 1\,\%$ at low multiplicity and reaches as much as $11\,\%$ at high multiplicity; similar results are obtained in the $n_{\mathrm{ch}} \geq 1$, $p_{\mathrm{T}} > 500$ MeV phase space (see Section 7.8). This is a significant reduction with respect to the previous measurement at $\sqrt{s} = 7$ TeV where this source of uncertainty reached as much as $25\,\%$. The improvement in the current analysis is due to the more accurate modelling of the ATLAS detector and its corresponding strongly reduced material uncertainty in the final Run-1 geometry.

A significant contribution to the total systematic uncertainty on the unfolded n_{ch} distribution comes from the uncertainty due to different p_{T} spectra in data and MC simulation. As the track reconstruction efficiency depends strongly on p_{T} (particularly in the low-p_{T} range of the most

inclusive phase space), the statistical relations between n_{sel} and n_{ch} that are expressed in the unfolding matrix depend on the particular model-dependent p_{T} spectrum of the MC sample that was employed to generate the matrix; the further the spectrum predicted by the model deviates from that observed in the data, the larger the systematic effect. The impact of this model-dependent effect is investigated by randomly removing tracks in each event according to the difference between the average track reconstruction efficiencies in the given n_{sel} event category, derived using the p_{T} spectrum in data and MC simulation (see Figure 7.31). The obtained modified n_{sel} distribution is again used as an alternative input to the unfolding procedure; this time the difference between the alternative and nominal distributions is taken as an asymmetric uncertainty. In the $n_{\mathrm{ch}} \geq 2$, $p_{\mathrm{T}} > 100$ MeV phase space, this uncertainty ranges from $-2\,\%$ to $13\,\%$, with largest uncertainties in the high-n_{ch} region, whereas it almost vanishes completely in the $n_{\mathrm{ch}} \geq 1$, $p_{\mathrm{T}} > 500$ MeV phase space (see Section 7.8). This is a significant reduction with respect to the previous measurement at $\sqrt{s} = 7$ TeV where this source of uncertainty reached values of as much as $40\,\%$ in the most inclusive phase space [126]. The improvement in the presented study is mainly due to the much better modelling of the p_{T} spectrum by the new baseline Pythia 8 A2 MC tune, whereas in the previous study a Pythia 6 tune was used.[18]

The systematic uncertainties of the unfolding procedure due to the sources listed in the second (alternative input distributions, described in Appendix A.5) and third category (random statistical fluctuations of the unfolding matrix) are combined in quadrature and later referred to as the "Unfolding" systematic (see Section 7.8); this combined uncertainty ranges between 1–14 % in the most inclusive phase space, and between 1–9 % in the $n_{\mathrm{ch}} \geq 1$, $p_{\mathrm{T}} > 500$ MeV phase space. It thus becomes the leading or next-to-leading contribution to the total systematic uncertainty.

7.6.3 Calculation of N_{ev}

The total event yield N_{ev} in the data sample is calculated by integrating the corrected and unfolded n_{ch} distribution. The corrected η and p_{T} distributions, as well as the n_{ch} spectrum itself, are then normalised by this number.

7.6.3.1 Systematic Uncertainties

For the total systematic uncertainty on N_{ev}, all sources of systematic uncertainty that can lead to a migration of events into or out of the phase-space acceptance region must be considered. The dominant contribution arises from the material uncertainty on the track reconstruction efficiency. Alternative event yields are calculated from the alternative n_{ch} distributions that were produced before, and differences to the nominal value of N_{ev} are taken in quadrature. The total

[18] Here it should be noted again that the p_{T} spectrum in 8 TeV data was modelled even better by Epos LHC (see Section 7.5.3.2), and thus by employing it in the unfolding procedure, a further reduction of this source of systematic uncertainty could possibly be achieved; this was however not done due to the low available event statistics for this tune.

uncertainty on N_{ev} at $\sqrt{s} = 8$ TeV was found to be 0.19 % (0.13 %) in the $p_{\mathrm{T}} > 100$ MeV, $n_{\mathrm{ch}} \geq 2$ ($p_{\mathrm{T}} > 500$ MeV, $n_{\mathrm{ch}} \geq 1$) phase space.

7.6.4 Corrections to $1/N_{\mathrm{ev}} \cdot \mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta$

Corrected η distributions are obtained by weighting the individual tracks with the event- and track-level factors described in Eqs. 7.10 and 7.11. No unfolding procedure needs to be applied, because the resolution width of the reconstructed η track parameter is much smaller than the bin width used for the η distributions, and the distributions are sufficiently flat within the selected kinematic boundaries. Additional corrections are needed in the higher-multiplicity phase spaces; these are described in Section 7.6.7.

7.6.4.1 Systematic Uncertainties

The systematic uncertainties on the η distribution are obtained by modifying the correction weights by their respective systematic uncertainties. The dominant categories are:

1. material uncertainty of the track reconstruction efficiency
2. uncertainty due to track selection cuts
3. fraction of non-primary tracks and strange baryons
4. trigger and vertex efficiency, particle composition, physics list, truth matching

The correction weights are varied by systematic uncertainties from each of the individual categories, and the symmetrised differences between the resulting alternative distributions and the nominal corrected distribution are evaluated. In addition, a conservatively shaped systematic uncertainty due to non-closure (see Section 7.7) is taken into account for each η distribution. The uncertainties from all contributions are added in quadrature and the total systematic uncertainty, which is dominated by the material uncertainty, is applied to the final distribution.

The final uncertainty on the central charged-particle density at $\eta = 0$ is found to be on the order of 1.7 % in the $p_{\mathrm{T}} > 100$ MeV, $n_{\mathrm{ch}} \geq 2$ phase space, and 1.2 % in the $p_{\mathrm{T}} > 500$ MeV, $n_{\mathrm{ch}} \geq 1$ phase space, respectively. These uncertainties are 30–40 % smaller than in the previous analysis at $\sqrt{s} = 7$ TeV, thanks to the improved constraints on the material uncertainty.

7.6.5 Corrections to $1/(2\pi p_{\mathrm{T}} N_{\mathrm{ev}}) \cdot \mathrm{d}^2 N_{\mathrm{ch}}/(\mathrm{d}\eta\, \mathrm{d}p_{\mathrm{T}})$

As for the η distributions, corrected p_{T} distributions are obtained by weighting the individual tracks with the event- and track-level factors described in Eqs. 7.10 and 7.11. They are then used as an input to an iterative Bayesian unfolding [166] procedure, similar to the procedure used for the n_{ch} distribution, in order to account for resolution and migration effects. Here, the unfolding matrix relates the reconstructed track p_{T} values to the true p_{T} spectrum of primary particles, and the method ensures that the reconstructed initial p_{T} distribution converges towards the

latter. A fine binning of the distributions is used during the iterative procedure in order to suppress fluctuations. After four (five) iterations in the $p_{\mathrm{T}} > 100\,(500)$ MeV phase space, the method converges as indicated by $\chi^2/n_{\mathrm{dof}} < 1$. As for the η distributions, additional corrections are needed in the higher-multiplicity phase spaces (see Section 7.6.7).

7.6.5.1 Systematic Uncertainties

The systematic uncertainty on the p_{T} unfolding procedure due to alternative input distributions and unfolding matrices was estimated in the same way as for the second and third category of the n_{ch} systematics (see Section 7.6.2.1). In addition, alternative input distributions were obtained by modifying the track correction weights by their respective systematic uncertainties, as was done for the η distribution. The dominant source of systematic uncertainty is again found to be the material uncertainty. The obtained modified p_{T} distribution is then used as an alternative input to the same unfolding procedure that was described earlier, and the difference between the alternative and nominal unfolded p_{T} distributions is evaluated. All individual contributions to the systematic uncertainty were combined in quadrature; furthermore, an asymmetric contribution due to mis-measured high-p_{T} tracks and the momentum resolution was added. The final systematic uncertainty on the p_{T} distribution reaches from a minimum of less than 2 % up to 30 %, with the highest values in the $30 < p_{\mathrm{T}} < 50$ GeV range, dominated by systematic effects due to the material and the unfolding procedure (see Section 7.8).

7.6.6 Corrections to $\langle p_{\mathrm{T}} \rangle$ versus n_{ch}

The correction and unfolding strategy for the $\langle p_{\mathrm{T}} \rangle$ versus n_{ch} distribution follows the procedure from the previous analysis [126]. The distribution is defined by $(\sum_i p_{\mathrm{T}}(i) / \sum_i 1)$ versus n_{ch}, where i denotes all tracks from events within a given n_{ch} category. Before this ratio can be calculated, the two distributions in the numerator (the sum of all transverse momenta of selected tracks) and denominator (the total number of selected tracks) must be individually evaluated as functions of n_{sel} with all track-level corrections applied; then they must be separately unfolded from detector level (function of n_{sel}) to particle level (function of n_{ch}), utilising the same unfolding matrix that has been obtained in the final iteration of the n_{ch} unfolding (see Section 7.6.2). The final $\langle p_{\mathrm{T}} \rangle$ versus n_{ch} distribution is then obtained by taking the ratio of these two components. All reconstructed tracks with extremely high momentum, $p_{\mathrm{T}} > \sqrt{s}/2$ are considered to be unphysical and therefore removed, in order to prevent biases on the average transverse momenta in affected event categories; only one track in data is thus excluded from the measurement, with negligible impact on the systematic uncertainty.

7.6.6.1 Systematic Uncertainties

The dominant source of systematic uncertainty comprises effects caused by two simplifications that are made during the unfolding procedure, which are regarded as responsible for a non-closure effect (see Figures 7.33d and 7.33d). Firstly, the track reconstruction efficiency is only

parameterised by p_{T} and η, but not by the multiplicity of the event; a more accurate procedure would employ a full three-dimensional parametrisation, but this was highly unfeasible due to statistical limitations. Secondly, the p_{T} spectrum of the tracks in the various n_{sel} and n_{ch} event categories is not necessarily the same, but this is not reflected in the unfolding procedure for events that migrate from an n_{sel} category to another n_{ch} category; this might be resolved by a two-dimensional p_{T} and n_{ch} unfolding, however this would again be unfeasible due to statistical limitations and high complexity; while such efforts might be worthwhile to reduce the total uncertainty in future analyses, they were clearly beyond the scope of the presented work, given than the net effect from the non-closure due to the chosen strategy is quite small. A possible model-dependence of the non-closure, due to different p_{T} spectra of the MC samples used for the unfolding, was investigated by repeating the procedure with Pythia 8 A2, Pythia 6 AMBT2B and Epos LHC samples, thus varying the input distribution and unfolding matrix. A similar level of non-closure was observed for each of these MC generator models, indicating no significant model-dependence. This result justifies that the non-closure of the baseline Pythia 8 A2 tune can be taken as a systematic uncertainty. The non-closure effects due to the unfolding procedure (see Figures 7.33d and 7.33d) are covered by applying a systematic uncertainty of 1 % over the to the entire n_{ch} range, except for the low-multiplicity regions in the $p_{\mathrm{T}} > 100$ MeV and 500 MeV phase spaces where a shaped systematic uncertainty of up to 4 % is applied.

Effects due to systematic uncertainties on the event- and track-level weights that are used in the correction procedure are accounted for by varying the weights and thus modifying the input distributions for the numerator and denominator before applying the unfolding procedure. The same uncertainty categories as for the unfolding of the n_{ch} distribution are used; similarly, the largest effects are obtained for the material uncertainty and for the different p_{T} spectra in data and MC simulation. However, all these systematic uncertainties have a very small effect on the final distribution, as they mostly cancel out between numerator and denominator. Taken in quadrature, they remain below 1 % (0.4 %) across the entire multiplicity spectrum in the $p_{\mathrm{T}} > 100$ MeV, $n_{\mathrm{ch}} \geq 2$ ($p_{\mathrm{T}} > 500$ MeV, $n_{\mathrm{ch}} \geq 1$) phase space.

Effects due to uncertainties on the unfolding matrix, which are obtained by varying the $n_{\mathrm{sel}} \to n_{\mathrm{ch}}$ migration probabilities in the matrix before the unfolding procedure is applied, were found to be dominated by contributions from the other two categories. Statistical uncertainties were calculated for the numerator and denominator distributions and added in quadrature for their ratio, but this was also found to have only a negligible effect on the total uncertainty.

7.6.7 Corrections for Higher-Multiplicity Phase Spaces

For η and p_{T} distributions in higher-multiplicity phase spaces, it must be taken into account that the phase-space requirement $n_{\mathrm{ch}} \geq n_{\mathrm{ch}}^{\mathrm{min}}$ on the number of particles was initially applied on the number of reconstructed tracks ($n_{\mathrm{sel}} \geq n_{\mathrm{ch}}^{\mathrm{min}}$). A separate correction is required to account for events with $n_{\mathrm{ch}} \geq n_{\mathrm{ch}}^{\mathrm{min}}$ and $n_{\mathrm{sel}} < n_{\mathrm{ch}}^{\mathrm{min}}$, which have migrated out of the higher-multiplicity acceptance region due to track losses caused by the track reconstruction inefficiency.

The standard procedure from the nominal phase space is used for all tracks from events with $n_{\text{sel}} \geq n_{\text{ch}}^{\text{min}}$, for which it can be assumed that $n_{\text{ch}} \geq n_{\text{ch}}^{\text{min}}$. Tracks belonging to events with $n_{\text{sel}} < n_{\text{ch}}^{\text{min}}$ are also added to the η and p_{T} distributions, weighted according to the probability $p(n_{\text{ch}} \geq n_{\text{ch}}^{\text{min}} \,|\, n_{\text{sel}})$ that they correspond to an event with $n_{\text{ch}} \geq n_{\text{ch}}^{\text{min}}$ particles. The weighting factors $w_{n_{\text{sel}}}$ are extracted from the final iteration of the n_{ch} unfolding matrix, that describes the probabilities that events with an observed track multiplicity n_{sel} correspond to events with n_{ch} primary particles. For any category $n_{\text{sel}} < n_{\text{ch}}^{\text{min}}$, the migration probability ($p(n_{\text{ch}} \geq n_{\text{ch}}^{\text{min}} \,|\, n_{\text{sel}})$) can thus be calculated as the sum over all probabilities in the matrix corresponding to $n_{\text{ch}} \geq n_{\text{ch}}^{\text{min}}$ and n_{sel},

$$w_{n_{\text{sel}}} = p(n_{\text{ch}} \geq n_{\text{ch}}^{\text{min}} \,|\, n_{\text{sel}}) = \sum_{n_{\text{ch}} \geq n_{\text{ch}}^{\text{min}}} M_{\text{ch,sel}}. \tag{7.13}$$

In the $n_{\text{ch}} \geq 6$ phase space, the obtained weights range from $w_{n_{\text{sel}}=5} \approx 69\,\%$ to less than $2\,\%$ for the $n_{\text{sel}} = 2$ category. In the $n_{\text{ch}} \geq 20$ (50) phase space, the weight for events with $n_{\text{sel}} = 19$ (49) is found to be about $96\,\%$ ($> 99.8\,\%$). The event normalisation is based on the same procedure as for distributions in the nominal phase space; the normalisation factor $N_{\text{ev}}(n_{\text{ch}} \geq n_{\text{ch}}^{\text{min}})$ is calculated from the n_{ch} distribution by restricting the integral to the $n_{\text{ch}} \geq n_{\text{ch}}^{\text{min}}$ interval.

7.6.7.1 Systematic Uncertainties

The higher-multiplicity phase spaces are subject to the same sources of systematic uncertainty that were also considered for the distributions using the nominal $n_{\text{ch}} \geq 1$ selection. The effect of variations of the track reconstruction efficiency on the migration probabilities in the correction procedure was additionally assessed by investigating the impact of the procedure on detector-level η distributions of the Pythia 8 A2 MC sample with $5\,\%$ extra material in the ID (Config A), which reflect the systematic variation due to the material uncertainty. These distributions were normalised to the number of selected events in each high-multiplicity phase space under consideration, and compared to the results obtained with the nominal Pythia 8 A2 sample, both for the case with and without the phase-space migration corrections. As a result, it was found that in these high-multiplicity phase spaces, a variation in the track reconstruction efficiency leads to highly correlated changes in both the numbers of selected tracks and selected events which effectively cancel out, such that residual differences in the normalised track distributions are found to be on the level of the material uncertainties in the same way as for the nominal phase space. Therefore it was concluded that no additional systematic uncertainty is needed for the high-multiplicity phase spaces.

7.7 Closure Tests

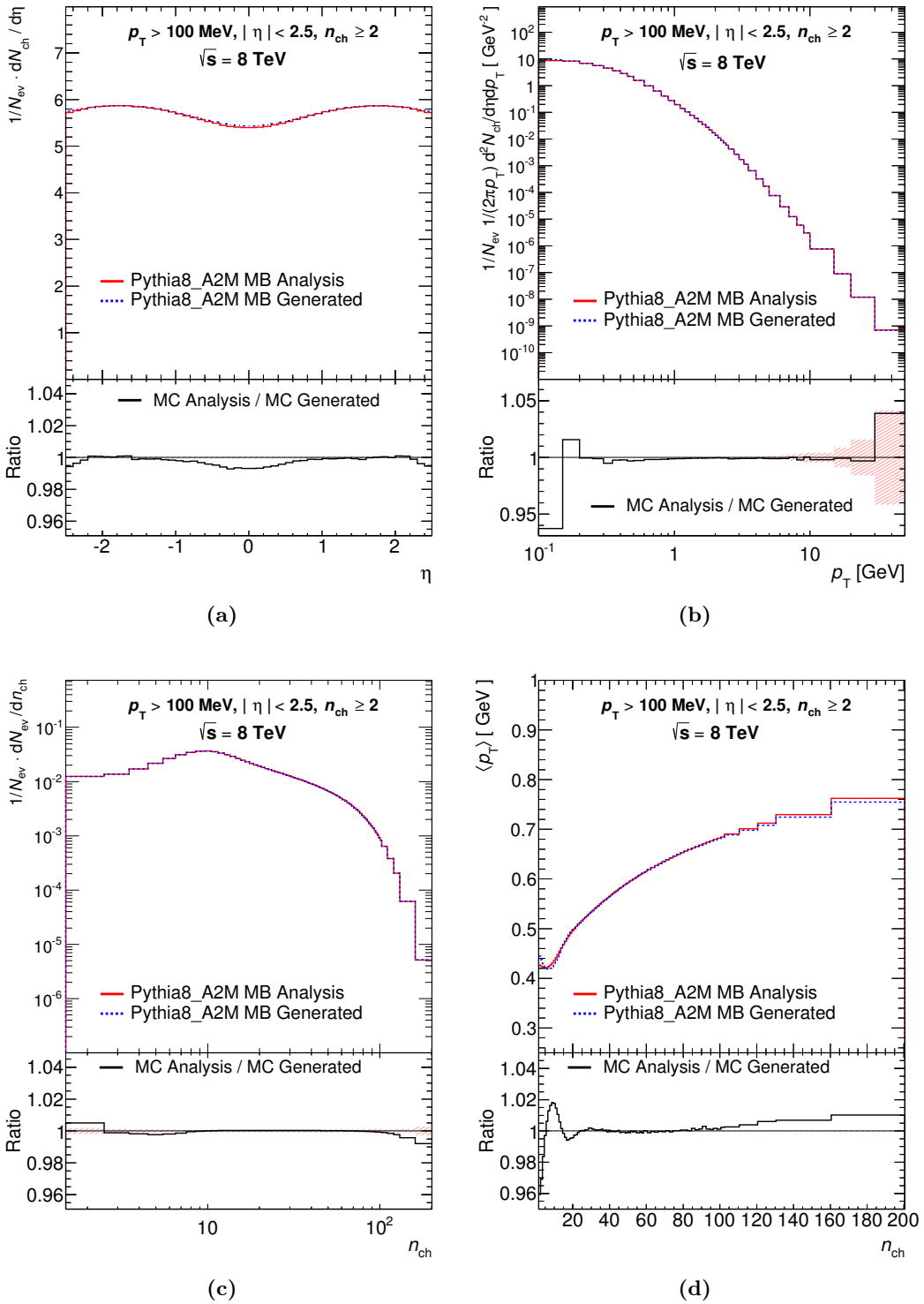
Closure tests were performed, as in previous measurements [126], to validate the robustness of the correction and unfolding procedure for all distributions and in all phase spaces used in this analysis. For these tests, the full nominal correction procedure is applied to the MC-based detector-level distributions, such as to reproduce the particle-level distribution that was generated by the same MC tune. The residual difference between the generated and corrected distributions is termed non-closure. A level of non-closure below 1 % is typically considered to be sufficient for the validation of the procedure; larger non-closures should be well understood, and must then be taken as a source of systematic uncertainty on the given distribution.[19] In the following, the Pythia 8 A2 tune is employed for these tests; the equivalent results obtained with different MC generators are given in Appendix A.6.3.2.

Figure 7.33 shows the closure tests for the most inclusive phase space at $p_{\mathrm{T}} > 100$ MeV and $n_{\mathrm{ch}} \geq 2$. In the p_{T} distribution shown in Figure 7.33c, the non-closure of $\sim 6\,\%$ at $100 < p_{\mathrm{T}} < 150$ MeV is taken as an additional one-sided systematic uncertainty; this non-closure is caused by momentum resolution and migration effects that are not fully accounted for by the correction procedure, and not fully covered by the systematic uncertainty that was already included to cover such effects. The non-closure of $\sim 2\,\%$ at $150 < p_{\mathrm{T}} < 200$ MeV is symmetrised and also added to the systematic uncertainty, while the non-closure at high p_{T} is covered by the statistical uncertainty. In addition, the η non-closure, even though well below 1 %, is taken as an additional conservatively shaped systematic uncertainty; it is most pronounced in the central and forward regions. The non-closure in the low-multiplicity region of the $\langle p_{\mathrm{T}} \rangle$ versus n_{ch} distribution, which arises due to simplifications made in the unfolding procedure, is fully taken as a conservatively shaped systematic uncertainty ranging from 1 % to 4 %.

Figure 7.34 shows the closure tests for the phase space limited by $p_{\mathrm{T}} > 500$ MeV and $n_{\mathrm{ch}} \geq 1$. In the p_{T} distribution shown in Figure 7.34c, the non-closure of $\sim 1\,\%$ at $500 < p_{\mathrm{T}} < 600$ MeV is taken as an additional systematic uncertainty, while the non-closure at high p_{T} is covered by the statistical uncertainty. In addition, the η non-closure, even though well below 1 %, is taken as an additional conservatively shaped systematic uncertainty. The non-closure of $\sim 1.6\,\%$ in the multiplicity distribution at $n_{\mathrm{ch}} > 100$, understood to be a consequence of the imperfect performance of the unfolding procedure in the tails of the distribution, can be taken as a systematic uncertainty, but it is completely negligible with respect to other systematic contributions in this region. The non-closure in the $\langle p_{\mathrm{T}} \rangle$ versus n_{ch} distribution, which is due to simplifications made in the unfolding procedure, is fully taken as a shaped systematic uncertainty ranging from 0.5 % to 2 %.

Figures 7.35 to 7.37 show the closure tests for the high-multiplicity phase spaces at $p_{\mathrm{T}} > 500$ MeV and $n_{\mathrm{ch}} \geq 6$, 20 and 50. In the phase space limited by $n_{\mathrm{ch}} \geq 6$, the non-closure of both η and p_{T} remains well below 1 %, except at high p_{T} where the non-closure is covered by

[19] With a hypothetically flawless correction and unfolding procedure, perfect agreement (i.e. zero non-closure) would be achieved.



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

Figure 7.33: Comparison of generated and fully corrected primary-charged-particle distributions for simulated events (using PYTHIA 8 A2) with $n_{\mathrm{ch}} \geq 2$ and $p_{\mathrm{T}} > 100$ MeV as a function of (a) pseudorapidity, η, (b) transverse momentum, p_{T}, (c) the charged-particle multiplicity per event, n_{ch}, and (d) average transverse momentum, $\langle p_{\mathrm{T}} \rangle$, versus multiplicity. The non-closure at $100 < p_{\mathrm{T}} < 150$ MeV is taken as an additional systematic uncertainty, while the non-closure at high p_{T} is covered by the statistical uncertainty. The non-closure in the $\langle p_{\mathrm{T}} \rangle$ distribution is fully taken as a systematic uncertainty.

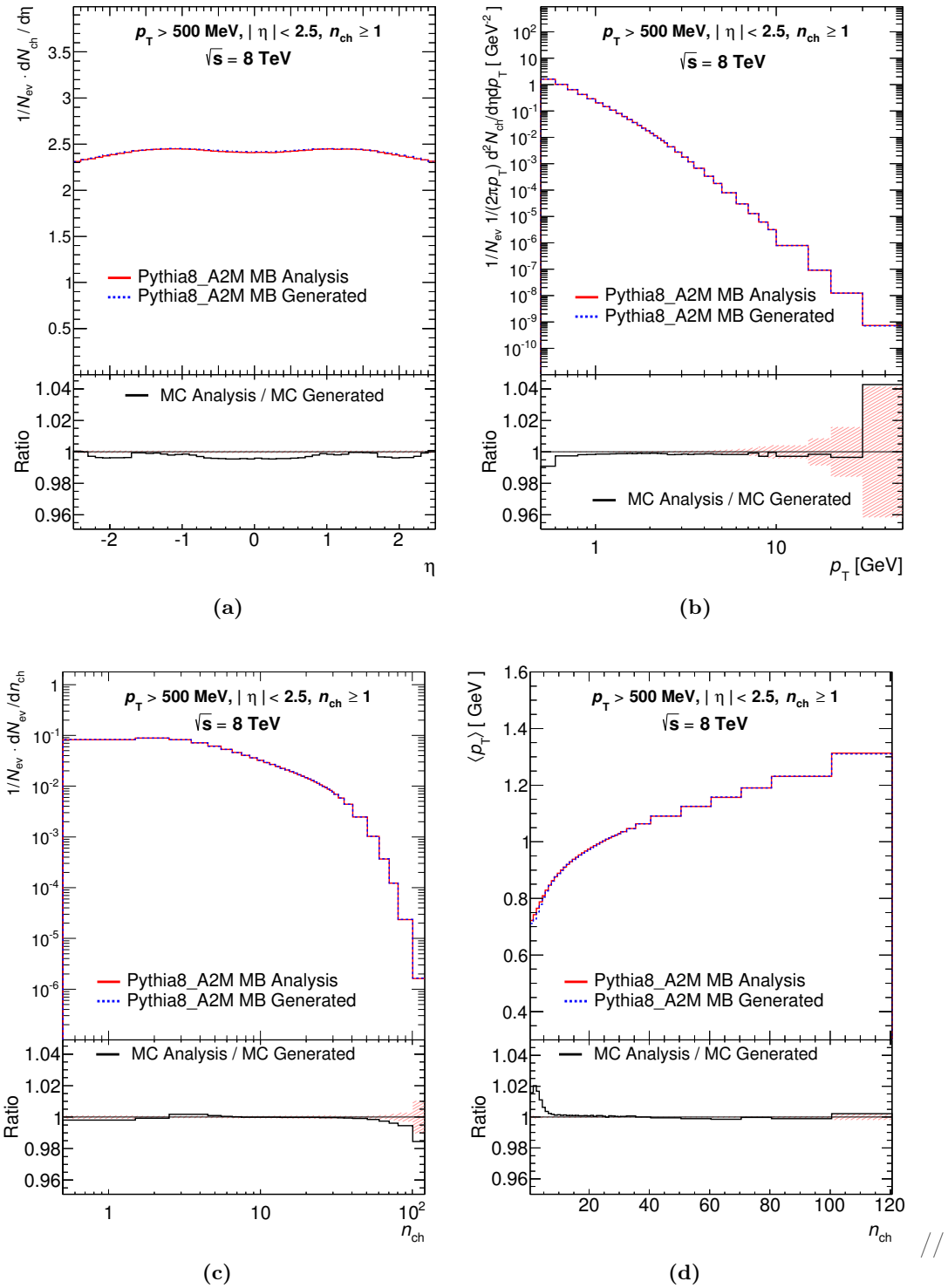
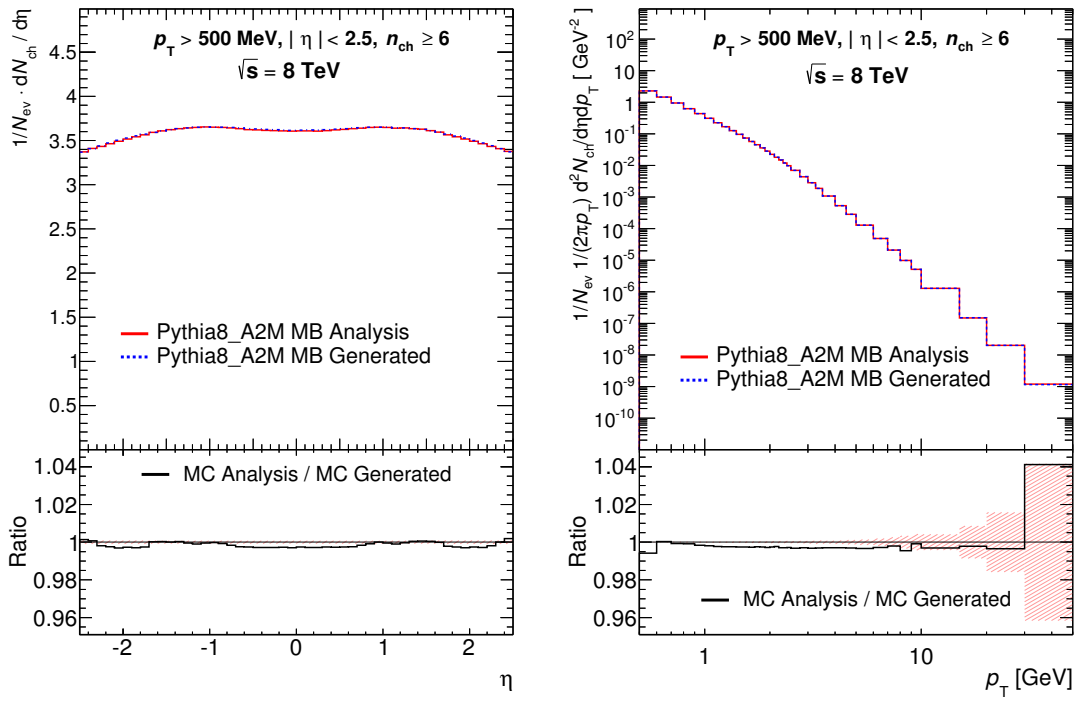


Figure 7.34: Comparison of generated and fully corrected primary-charged-particle distributions for simulated events (using PYTHIA 8 A2) with $n_{\mathrm{ch}} \geq 1$ and $p_{\mathrm{T}} > 500$ MeV as a function of (a) pseudorapidity, η, (b) transverse momentum, p_{T}, (c) the charged-particle multiplicity per event, n_{ch}, and (d) average transverse momentum, $\langle p_{\mathrm{T}} \rangle$, versus multiplicity. The non-closure at high p_{T} is covered by the statistical uncertainty. The non-closure in the $\langle p_{\mathrm{T}} \rangle$ distribution is fully taken as a systematic uncertainty.



(a) (b)

Figure 7.35: Comparison of generated and fully corrected primary-charged-particle distributions for simulated events (using PYTHIA 8 A2) with $n_{\mathrm{ch}} \geq 6$ and $p_{\mathrm{T}} > 500$ MeV as a function of (a) pseudorapidity, η, and (b) transverse momentum, p_{T}. The non-closure at high p_{T} is covered by the statistical uncertainty.

the statistical uncertainty. In the phase space limited by $n_{\mathrm{ch}} \geq 20$, the non-closure of both η and p_{T} increases slightly, and in the phase space limited by $n_{\mathrm{ch}} \geq 50$, non-closures of up to $\sim 1\,\%$ in the forward regions of the η distribution, and up to $\sim 2\,\%$ in the p_{T} distribution at several GeV, are found. All these non-closures are propagated to their final distributions as conservatively shaped systematic uncertainties.

7.8 Total Systematic Uncertainties

The impact of individual sources of systematic uncertainty on the final results was assessed by producing alternative results, mostly based on variations of the event weights, track weights or unfolding matrices that were used in the correction procedure. The differences between these alternative results and the nominal results were then taken as systematic uncertainties. Ultimately, all these individual contributions, as well as the statistical uncertainties of the data (which are always clearly dominated by the systematic uncertainties), were added in quadrature. The resulting total uncertainties are shown in the final distributions in Section 7.9, as well as in Appendix A.

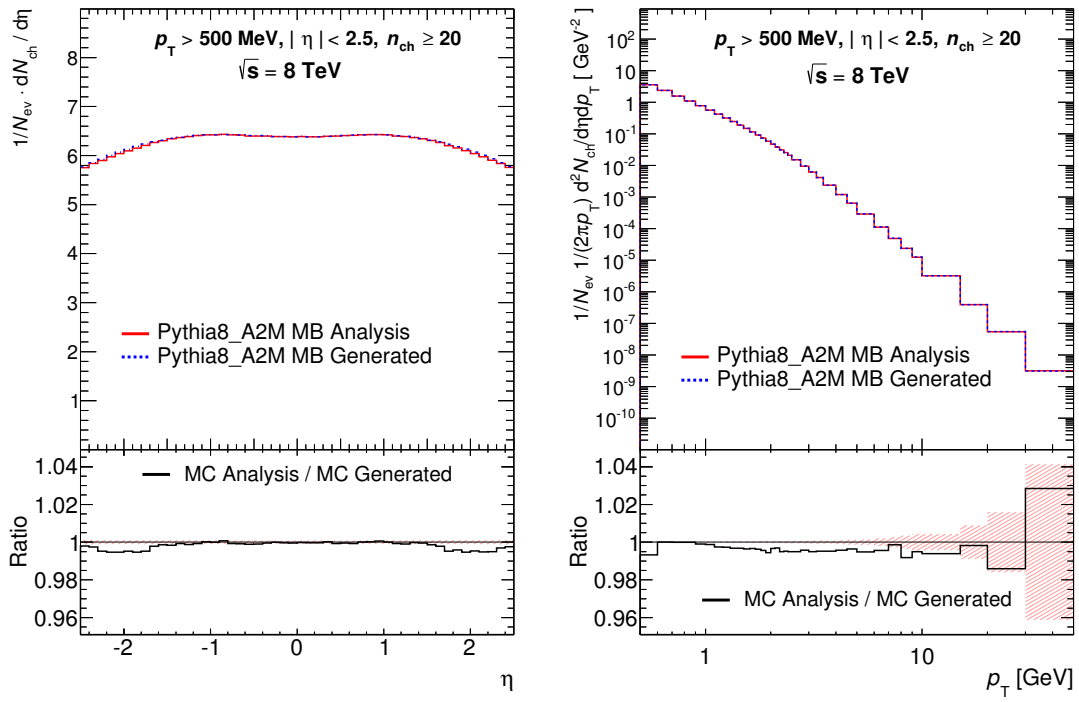


Figure 7.36: Comparison of generated and fully corrected primary-charged-particle distributions for simulated events (using PYTHIA 8 A2) with $n_{\mathrm{ch}} \geq 20$ and $p_{\mathrm{T}} > 500$ MeV as a function of (a) pseudorapidity, η, and (b) transverse momentum, p_{T}. The non-closure at high p_{T} is covered by the statistical uncertainty.

A breakdown of the sources of systematic uncertainty into the dominant components for different distributions and phase spaces is shown in Figures 7.38 to 7.42. The labels in these figures are referring to the following categories of systematic effects:

Material refers to the systematic uncertainty on the track reconstruction efficiency due to incomplete knowledge of the material distribution in the ID (see Section 7.5.3.2). This was found to be the leading uncertainty on the η distribution, and among the two leading uncertainties in all other distributions. It reaches highest values at large pseudorapidities, low and high transverse momenta, and high multiplicities.

Secondaries refers to the combined systematic uncertainty related to the estimated fractions of non-primary tracks (see Section 7.4.2.2) and reconstructed strange baryons (see Section 7.4.3), both of which contaminate the sample and were therefore subtracted. The total uncertainty of both sources in the corrected p_{T} distributions rises to as much as 3.5 % at the highest transverse momenta, but it is never a dominant uncertainty.

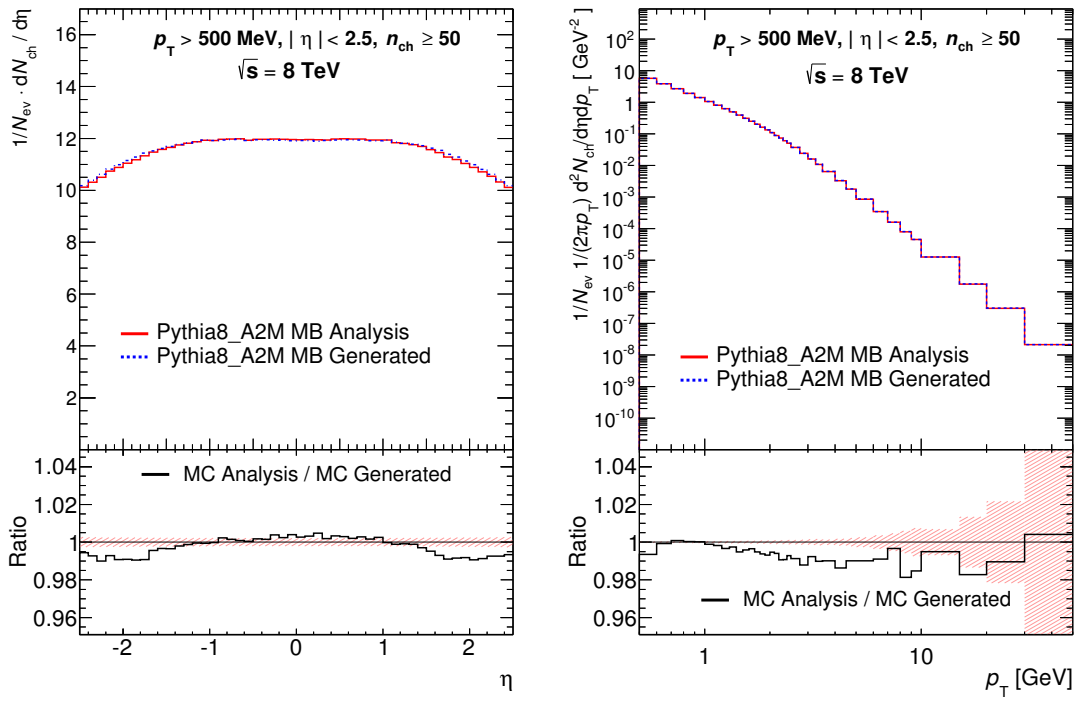


Figure 7.37: Comparison of generated and fully corrected primary-charged-particle distributions for simulated events (using PYTHIA 8 A2) with $n_{\mathrm{ch}} \geq 50$ and $p_{\mathrm{T}} > 500$ MeV as a function of (a) pseudorapidity, η, and (b) transverse momentum, p_{T}. The non-closure at high p_{T} is covered by the statistical uncertainty.

Cuts refers to the systematic uncertainty associated with different track selection efficiencies in data and MC simulation (see Section 7.5.3.2). The total uncertainty from this source reaches as much as 6 % at highest transverse momenta, and 5 % at high multiplicities, which is however still dominated by other categories.

Flat combines the systematic uncertainties associated with the trigger and vertex reconstruction efficiencies (see Sections 7.5.1.1 and 7.5.2.2), the truth-matching algorithm (see Section 7.5.3.2), the particle composition and the choice of physics list Section 7.5.3.2. However, this category is dominated by other sources in all regions of the measured distributions.

Statistics refers to the statistical uncertainty on the track reconstruction efficiency, which only affects the high-p_{T} region with up to 2 %. This is clearly not a dominant uncertainty.

pT-Spectrum refers to the systematic uncertainty associated to the different p_{T} spectra in the MC models and data (see Section 7.5.3.2). This effect dominates only in the high-multiplicity range of the final n_{ch} distribution in the $n_{\mathrm{ch}} \geq 2$, $p_{\mathrm{T}} > 100$ MeV phase space, reaching as much as 13 %.

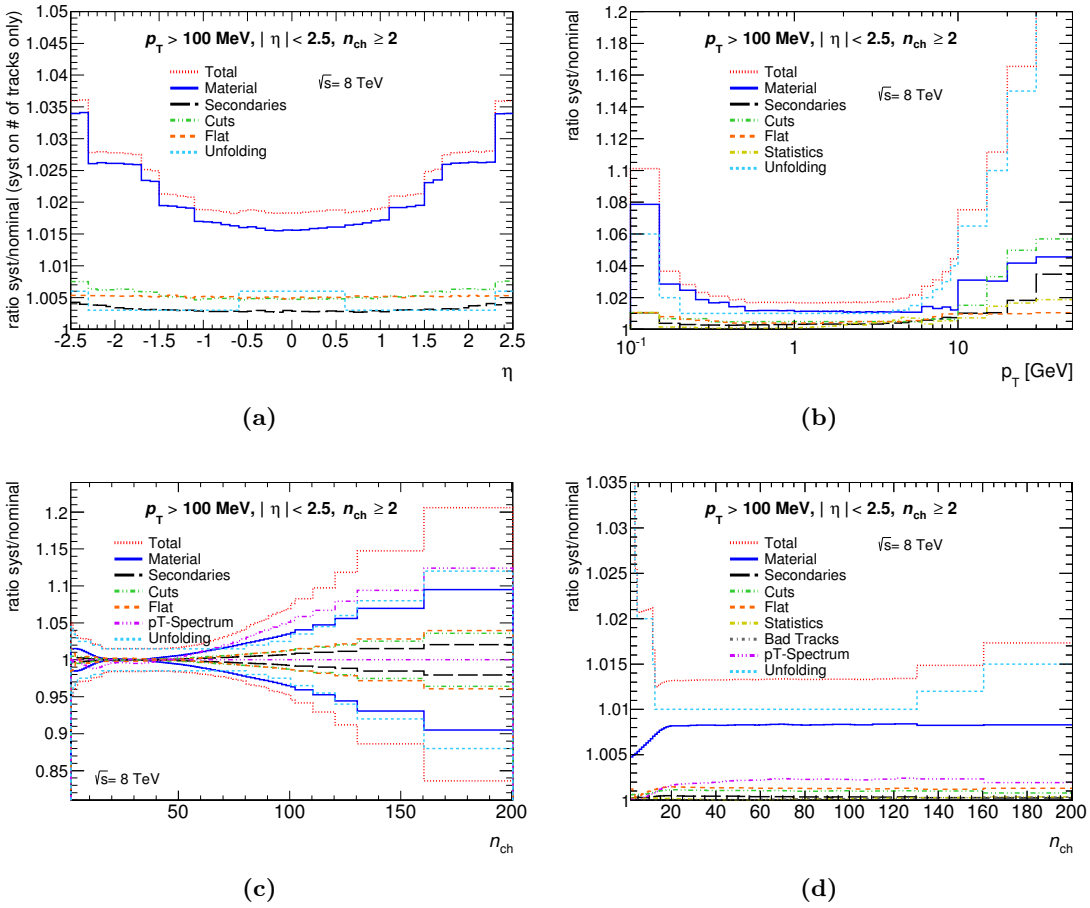


Figure 7.38: Overview of different sources of systematic uncertainty on the primary-charged-particle distributions as a function of (a) the pseudorapidity, η, (b) transverse momentum, p_{T}, (c) multiplicity, n_{ch}, and (d) $\langle p_{\mathrm{T}} \rangle$ versus n_{ch}, for events with $n_{\mathrm{ch}} \geq 2$ and $p_{\mathrm{T}} > 100$ MeV.

Bad Tracks refers to the combination of the one-sided systematic uncertainties associated to the fraction of mis-measured high-p_{T} tracks, which increases with transverse momentum and pseudorapidity, and to high-p_{T} momentum resolution effects. This category dominates in the high-p_{T} regions of the transverse-momentum distributions, and is negligible in all other distributions.

Unfolding refers to the systematic uncertainty associated with the unfolding technique, obtained by adding in quadrature the individual contributions from variations of the input spectra and the unfolding matrix (including statistical variations of the matrix), as well as non-closure effects in the nominal unfolding procedure (see Sections 7.6.2.1, 7.6.5.1 and 7.6.6.1 and Appendix A.5). This uncertainty dominates in the high-p_{T} regions of the transverse-momentum distributions where it reaches as much as 20 %, in the low- and high-n_{ch} regions of the multiplicity distributions with up to 12 %, and in the entire range of the $\langle p_{\mathrm{T}} \rangle$ versus n_{ch} distribution.

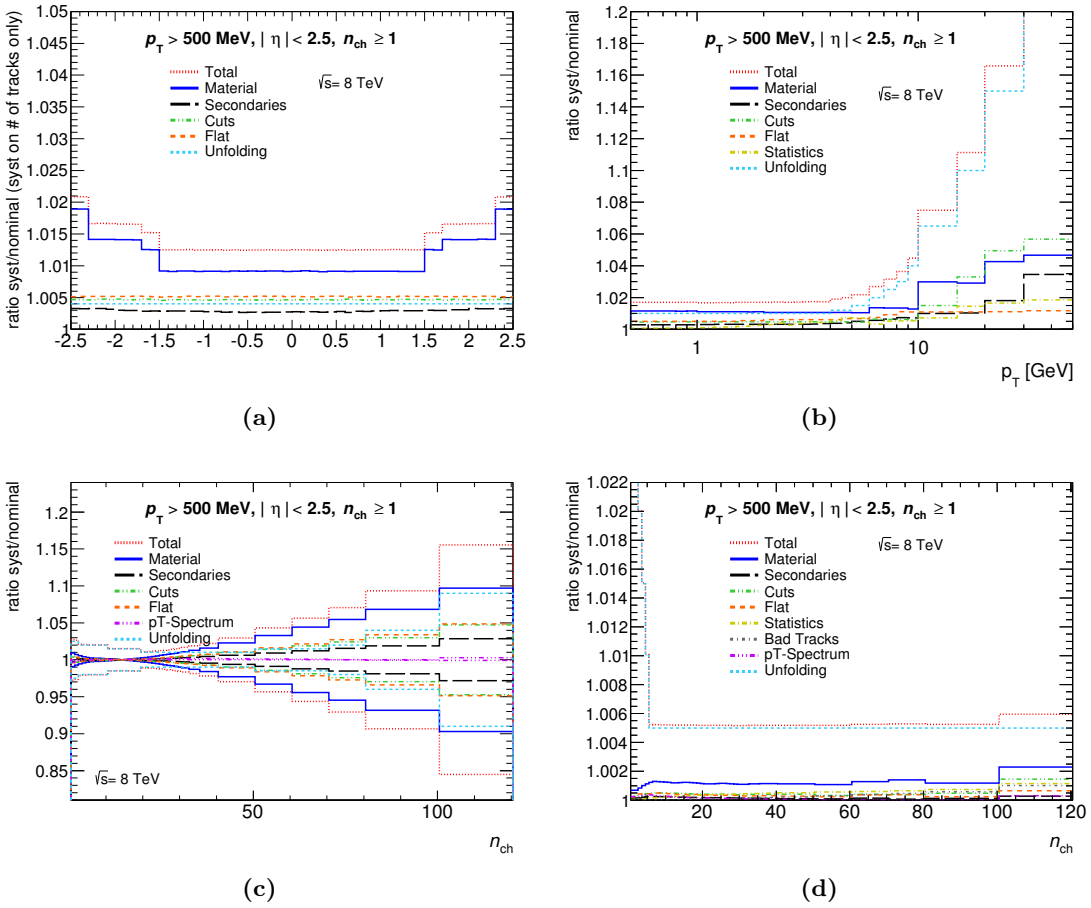


Figure 7.39: Overview of different sources of systematic uncertainty on the primary-charged-particle distributions as a function of (a) the pseudorapidity, η, (b) transverse momentum, p_{T}, (c) multiplicity, n_{ch}, and (d) $\langle p_{\mathrm{T}} \rangle$ versus n_{ch}, for events with $n_{\mathrm{ch}} \geq 1$ and $p_{\mathrm{T}} > 500$ MeV.

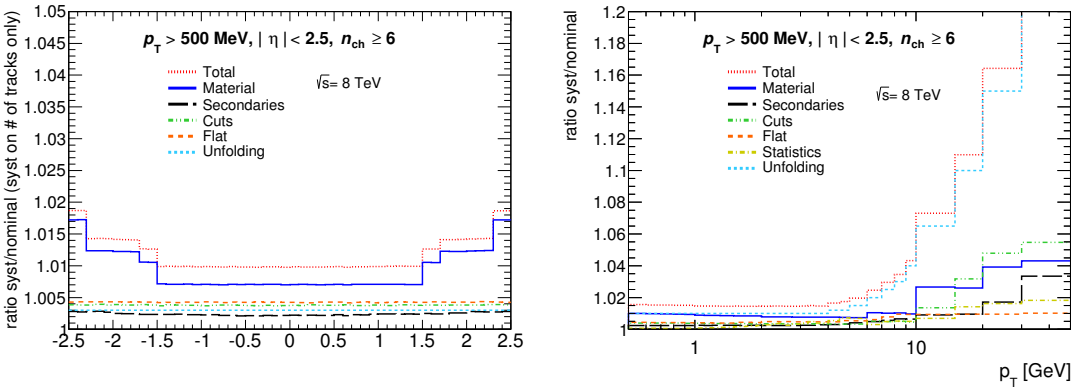
A summary of the effect sizes of all categories on the final distributions is presented in Table 7.16.

7.9 Results

Final-state distributions of stable charged particles are shown in Figures 7.43 to 7.47 for all phase spaces listed in Section 7.3.1, and are compared to predictions of various QCD-based phenomenological models [3]. The performance of three additional models (Pythia 6 Innsbruck2013 and two Herwig++ tunes), which were not included in the 8 TeV publication, is discussed in Appendix A.8.

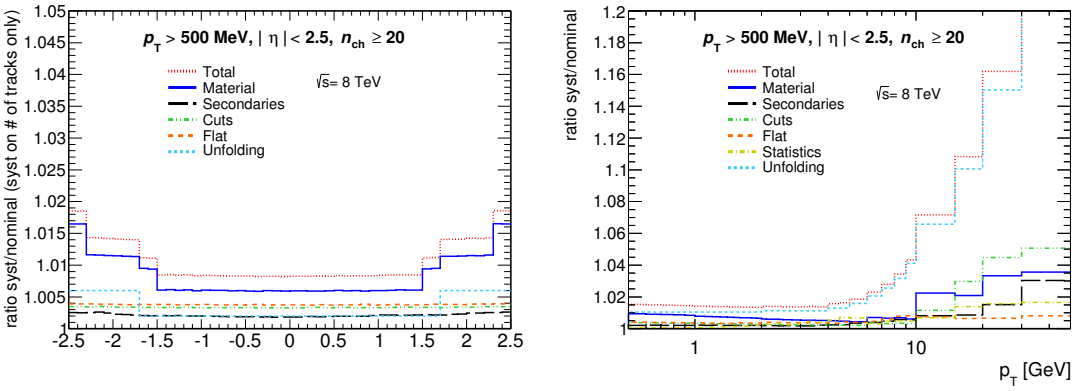
7.9.1 Pseudorapidity Distribution

Figures 7.43a, 7.44a, 7.45a, 7.46a and 7.47a show distributions of the primary-charged-particle yield as a function of pseudorapidity, $1/N_{\mathrm{ev}} \cdot \mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta$. The distribution measured in the $n_{\mathrm{ch}} \geq 2$



(a) (b)

Figure 7.40: Overview of different sources of systematic uncertainty on the primary-charged-particle distributions as a function of (a) the pseudorapidity, η, (b) transverse momentum, p_{T}, for events with $n_{\mathrm{ch}} \geq 6$ and $p_{\mathrm{T}} > 500$ MeV.

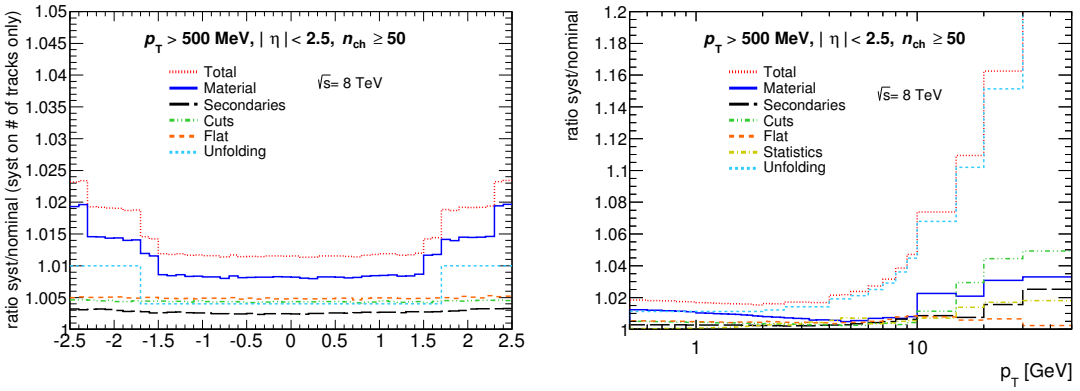


(a) (b)

Figure 7.41: Overview of different sources of systematic uncertainty on the primary-charged-particle distributions as a function of (a) the pseudorapidity, η, (b) transverse momentum, p_{T}, for events with $n_{\mathrm{ch}} \geq 20$ and $p_{\mathrm{T}} > 500$ MeV.

and $p_{\mathrm{T}} > 100$ MeV phase space has a more variable shape than those in the phase-space regions with $p_{\mathrm{T}} > 500$ MeV, and reaches maxima around $|\eta| \approx 1.7$. In the latter phase spaces, the distribution shows a plateau within $|\eta| \lesssim 1.5$ that shrinks to $|\eta| \lesssim 1.0$ at the $n_{\mathrm{ch}} \geq 50$ threshold. The small fluctuations at higher absolute pseudorapidities (where a smooth distribution would be expected) suggest imperfections in the material description; these possible effects are however compatible with the total systematic uncertainty.

In the most inclusive selection ($n_{\mathrm{ch}} \geq 2$, $p_{\mathrm{T}} > 100$ MeV), the EPOS LHC and PYTHIA 8 MONASH models describe the data well, whereas PYTHIA 8 A2 and QGSJET-II underestimate the charged-particle density. Such discrepancies are not surprising, though, as the given MC models were mostly tuned to measurements at higher transverse-momentum thresholds of



(a) **(b)**

Figure 7.42: Overview of different sources of systematic uncertainty on the primary-charged-particle distributions as a function of (a) the pseudorapidity, η, (b) transverse momentum, p_{T}, for events with $n_{\mathrm{ch}} \geq 50$ and $p_{\mathrm{T}} > 500$ MeV.

Category	$n_{\mathrm{ch}} \geq 2$, $p_{\mathrm{T}} > 100$ MeV				$n_{\mathrm{ch}} \geq 1$, $p_{\mathrm{T}} > 500$ MeV			
	n_{ch}	η	p_{T}	$\langle p_{\mathrm{T}} \rangle$	n_{ch}	η	p_{T}	$\langle p_{\mathrm{T}} \rangle$
Material	0–10	1.6–3.4	1–8	0.5–0.85	0–10	0.9–1.8	1.5–4.7	0–0.25
Secondaries*	0–2	0.3–0.4	0.3–3.5	0–0.1	0–3	0.3–0.35	0.3–3.5	0–0.05
Cuts	0–4	0.5–0.8	0.5–6	0–0.15	0–5	0.4–0.5	0.5–6	0–0.15
Flat**	0–4	0.5–0.6	0.4–1.1	0–0.15	0–5	0.5 –0.55	0.5–1.2	0–0.1
Statistics			0–2				0–2	0–0.1
pT-Spectrum	0–13			0–0.25	0–0.5			0–0.05
Bad Tracks			0–25				0–25	0–0.1
Unfolding***	1.5–12	0.3–0.6	1–20	1–4	1–9	0.4	1–20	0.5–2.2

Table 7.16: Summary of total systematic uncertainties on final corrected distributions in the $n_{\mathrm{ch}} \geq 2$, $p_{\mathrm{T}} > 100$ MeV and $n_{\mathrm{ch}} \geq 1$, $p_{\mathrm{T}} > 500$ MeV phase spaces. For each systematic category and type of distribution, the range of the relative systematic errors is given in percent. (*) This category combines effects due to non-primary and strange baryon corrections. (**) This category combines effects due to trigger and vertex reconstruction efficiency, truth matching, particle composition, and choice of physics list. (***) This category includes non-closure effects.

$p_{\mathrm{T}} > 500$ MeV. EPOS LHC also describes the plateau region of the distributions well in the phase spaces limited by $p_{\mathrm{T}} > 500$ MeV, except for the $n_{\mathrm{ch}} \geq 20$ selection; it predicts higher densities at $|\eta| > 1.7$. The QGSJET-II model consistently overestimates the data at $p_{\mathrm{T}} > 500$ MeV, as does the PYTHIA 8 MONASH tune except at $n_{\mathrm{ch}} \geq 50$. PYTHIA 8 A2 tends to underestimate the data at $n_{\mathrm{ch}} \geq 1, 50$ but gives a very good description of the plateau regions at $n_{\mathrm{ch}} \geq 6, 20$.

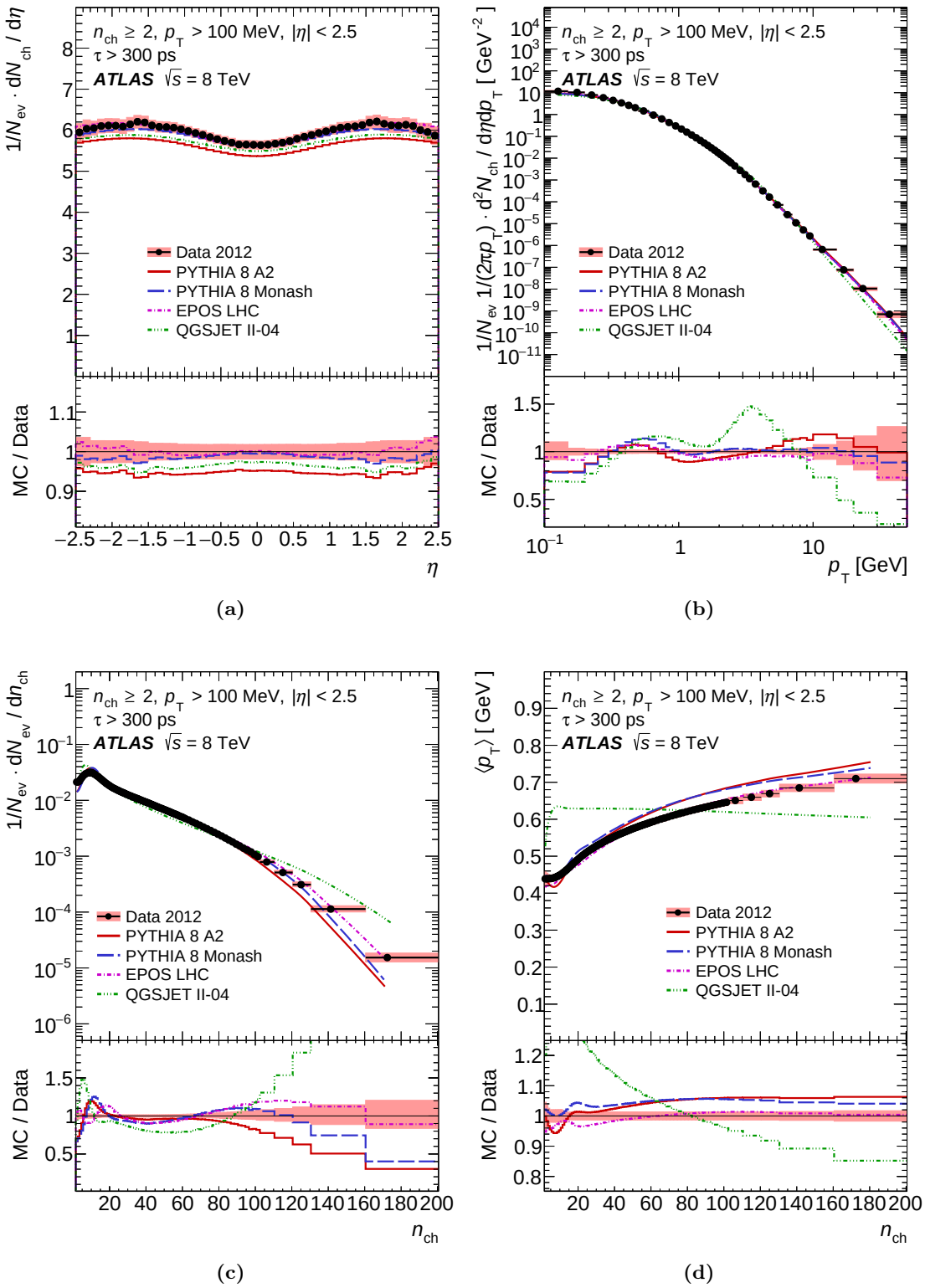


Figure 7.43: Distributions of primary charged particles in events for which $n_{ch} \geq 2$, $p_T > 100$ MeV and $|\eta| < 2.5$ as a function of (a) pseudorapidity, η, (b) transverse momentum, p_T, (c) multiplicity, n_{ch}, and (d) average transverse momentum, $\langle p_T \rangle$, versus multiplicity. The data (markers) are compared to several particle-level predictions made by MC generator models (curves). The total uncertainty on the data is represented by shaded areas. Plots taken from Ref. [3].

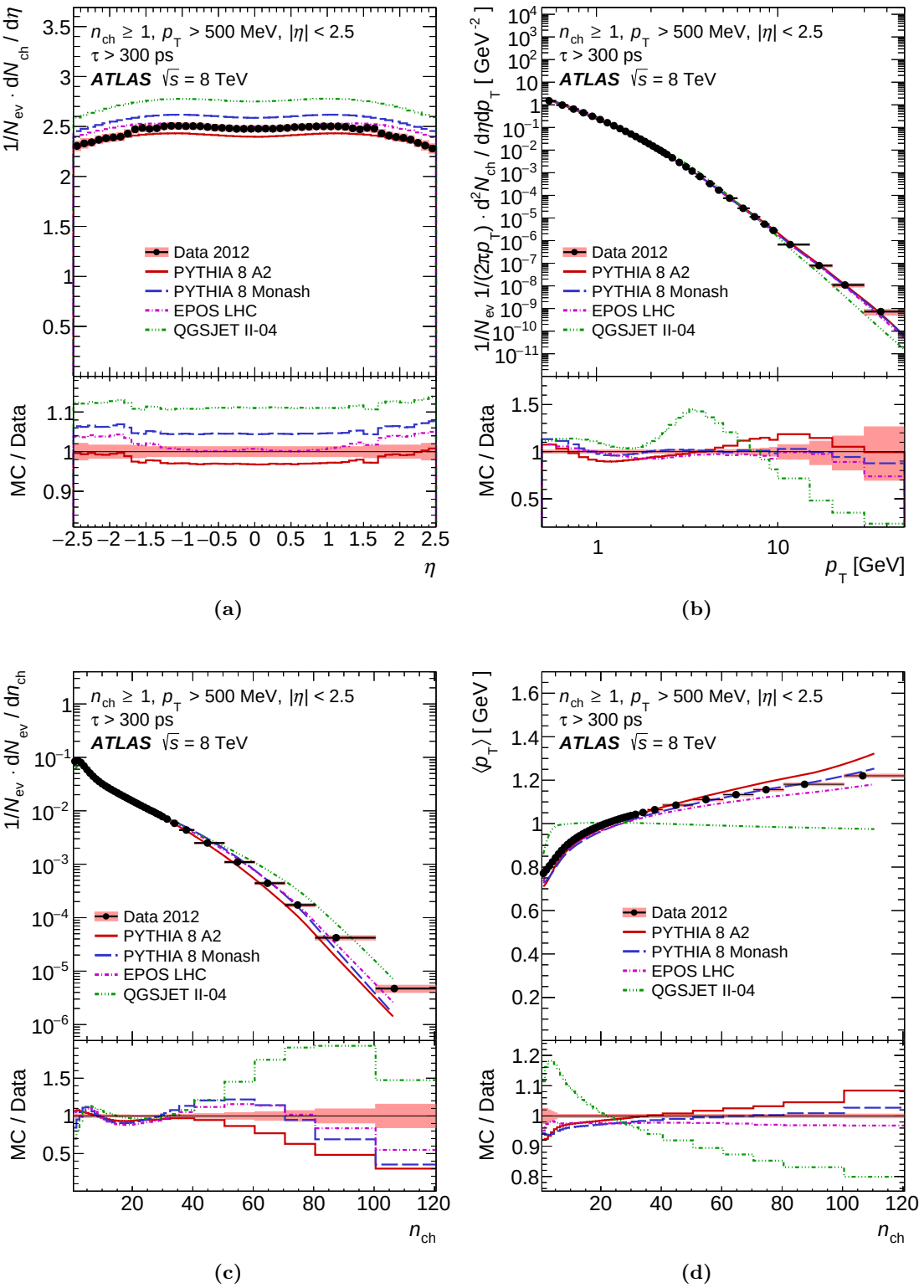
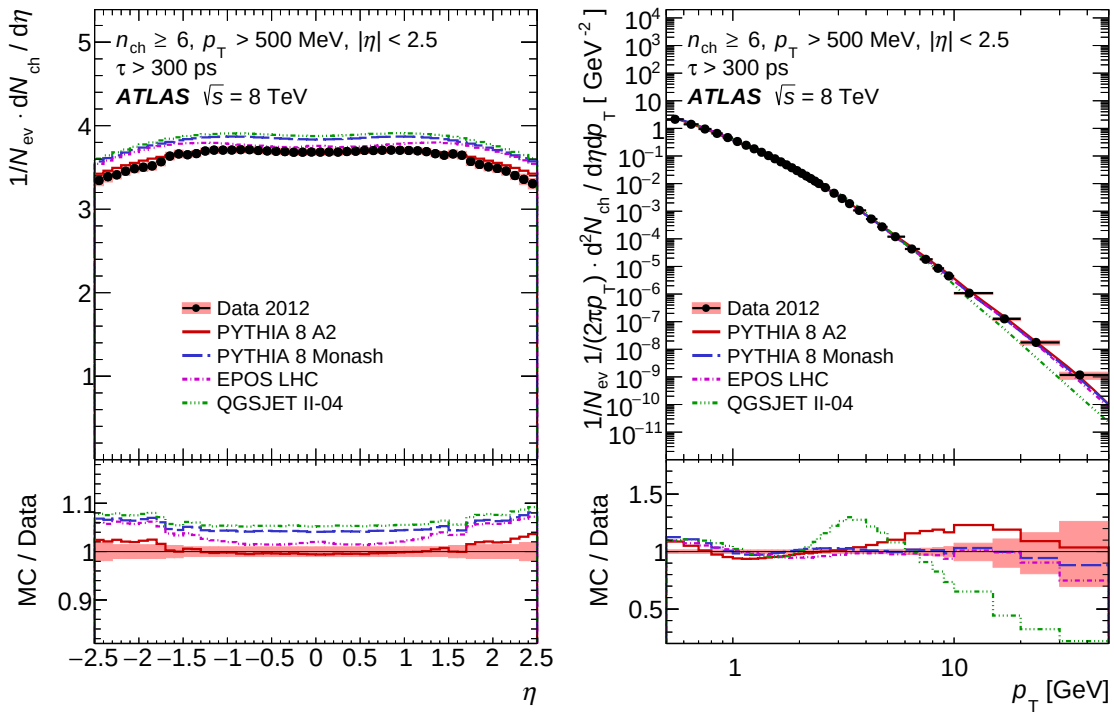


Figure 7.44: Distributions of primary charged particles in events for which $n_{\mathrm{ch}} \geq 1$, $p_{\mathrm{T}} > 500$ MeV and $|\eta| < 2.5$ as a function of (a) pseudorapidity, η, (b) transverse momentum, p_{T}, (c) multiplicity, n_{ch}, and (d) average transverse momentum, $\langle p_{\mathrm{T}} \rangle$, versus multiplicity. The data (markers) are compared to several particle-level predictions made by MC generator models (curves). The total uncertainty on the data is represented by shaded areas. Plots taken from Ref. [3].



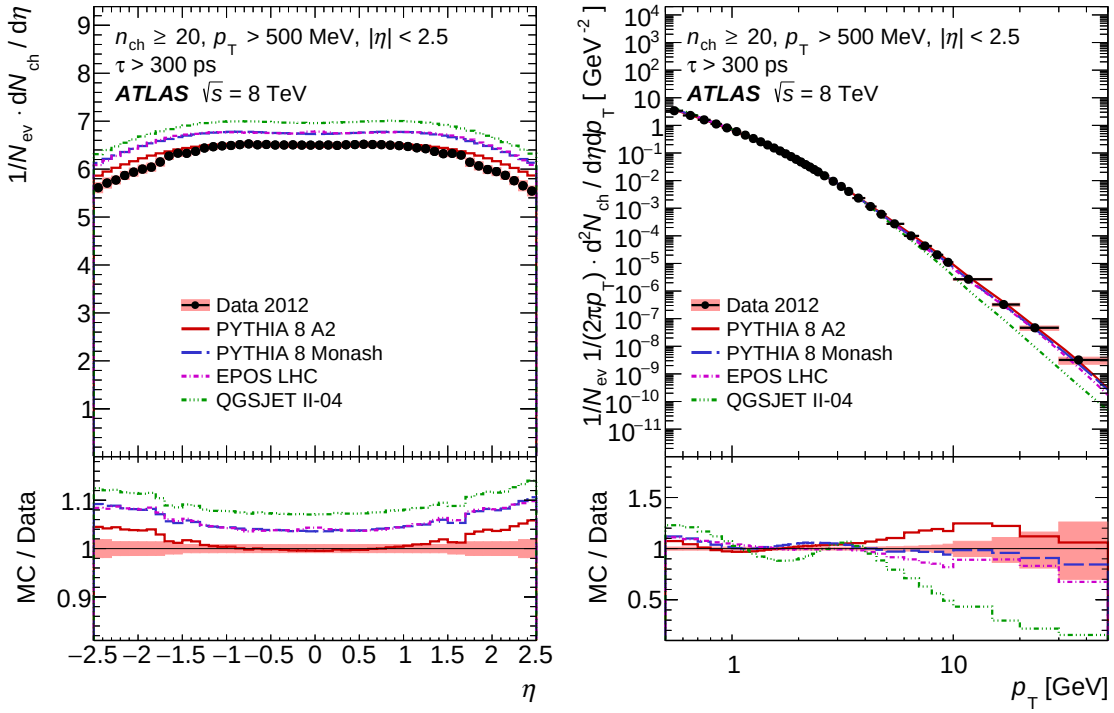
(a) (b)

Figure 7.45: Distributions of primary charged particles in events for which $n_{ch} \geq 6$, $p_T > 500$ MeV and $|\eta| < 2.5$ as a function of (a) pseudorapidity, η, and (b) transverse momentum, p_T. The data (markers) are compared to several particle-level predictions made by MC generator models (curves). The total uncertainty on the data is represented by shaded areas. Plots taken from Ref. [3].

7.9.2 Transverse-Momentum Distribution

Figures 7.43b, 7.44b, 7.45b, 7.46b and 7.47b show normalised distributions of the primary-charged-particle yield as a function of transverse momentum, $1/(2\pi p_T N_{ev}) \cdot d^2N_{ch}/(d\eta\, dp_T)$. While none of the considered models is found to be fully consistent with the experimental data, the Pythia 8 Monash tune provides a very good description above $p_T > 1$ GeV, including the $n_{ch} \geq 50$ selection for which the other models provide significantly worse descriptions with rising p_T. The low-p_T region of the data (below $p_T < 1$ GeV) is best described by Epos LHC, and underestimated by the other models; this is particularly the case for the $n_{ch} \geq 2$, $p_T > 100$ MeV phase space. However, Epos LHC begins to underestimate the data from $p_T \gtrsim 5$ GeV onwards in the phase spaces limited by $n_{ch} \geq 20$ and 50. Fair predictions are made by Pythia 8 A2 below $p_T \lesssim 6$ GeV, however at higher transverse momenta large deviations of as much as 20–30 % are found. The Qgsjet-II approach does not describe the distributions well, and gives particularly large deviations at higher transverse momenta from $p_T > 2$ GeV onwards.

Additional linear and logarithmic variations of these transverse-momentum distributions are shown in Appendix A.7.



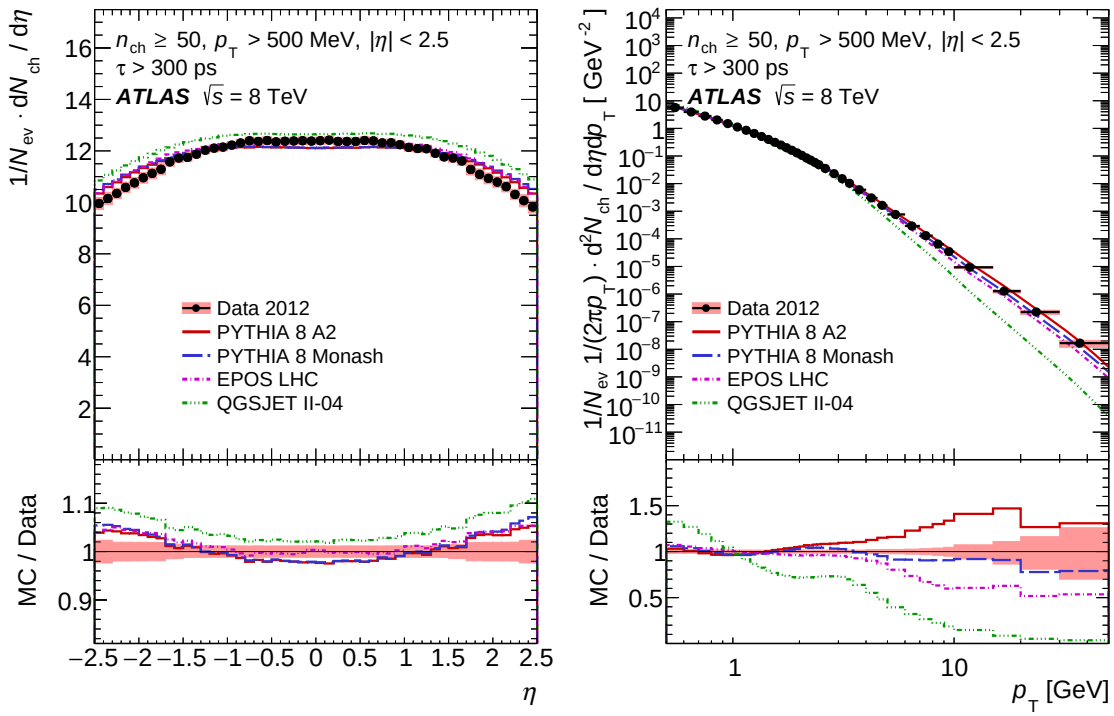
(a) (b)

Figure 7.46: Distributions of primary charged particles in events for which $n_{ch} \geq 20$, $p_T > 500$ MeV and $|\eta| < 2.5$ as a function of (a) pseudorapidity, η, and (b) transverse momentum, p_T. The data (markers) are compared to several particle-level predictions made by MC generator models (curves). The total uncertainty on the data is represented by shaded areas. Plots taken from Ref. [3].

7.9.3 Charged-Particle Multiplicity Distribution

The multiplicity distributions of primary charged particles, $1/N_{ev} \cdot dN_{ev}/dn_{ch}$, are presented in Figures 7.43c and 7.44c for the two main phase spaces considered in this analysis: (1) $n_{ch} \geq 2$, $p_T > 100$ MeV and (2) $n_{ch} \geq 1$, $p_T > 500$ MeV. After an initial rise, both distributions reach their maxima around $n_{ch} \approx 9$ and $n_{ch} \approx 2$, respectively, followed by a steeply falling spectrum. Again, none of the considered models is found to be fully consistent with the experimental results. However, the predictions made by EPOS LHC are in good agreement with the data in the $n_{ch} \geq 2$, $p_T > 100$ MeV phase space, with deviations not exceeding 20 %; this is also the case in the $n_{ch} \geq 1$, $p_T > 500$ MeV phase space for multiplicities up to $n_{ch} \approx 100$. The PYTHIA 8 tunes do not predict the low-n_{ch} regions very well, provide reasonable descriptions at intermediate multiplicities, and significantly underestimate the high-multiplicity regions. The QGSJET-II model shows the largest deviations and clearly overestimates the event yield at high multiplicities.

Additional linear and logarithmic variations of these charged-particle multiplicity distributions are shown in Appendix A.7.



(a) (b)

Figure 7.47: Distributions of primary charged particles in events for which $n_{ch} \geq 50$, $p_T > 500$ MeV and $|\eta| < 2.5$ as a function of (a) pseudorapidity, η, and (b) transverse momentum, p_T. The data (markers) are compared to several particle-level predictions made by MC generator models (curves). The total uncertainty on the data is represented by shaded areas. Plots taken from Ref. [3].

7.9.4 Average Transverse Momentum as a Function of n_{ch}

Final results for the $\langle p_T \rangle$ versus n_{ch} distribution are shown in Figures 7.43d and 7.44d for the two main phase spaces. In both cases the $\langle p_T \rangle$ is found to increase with multiplicity, with a steeper rise within $10 \lesssim n_{ch} \lesssim 40$ and $n_{ch} \lesssim 20$, respectively. This dependence can be attributed to colour coherence effects that are more prominent at higher multiplicities, where typically several MPI mechanisms contribute to the event. Without colour coherence effects, the produced final-state particles would share approximately the same p_T on average, regardless of the total multiplicity. Though if such effects are considered, the multiplicity rise due to MPI is limited by the the process of colour reconnection (see Sections 2.3.2 and 2.5.1). The same energy is then shared among fewer additional particles, which effectively increases the average p_T. Colour reconnection is simulated by the Pythia 8 and Epos models, which predict a $\langle p_T \rangle$ versus n_{ch} distribution similar to that observed in the data. Here Epos LHC performs particularly well in the $p_T > 100$ MeV phase space. The effect is not accounted for by Qgsjet-II, which therefore wrongly predicts an almost constant $\langle p_T \rangle$.

7.9.5 Comparison of Central Charged-Particle Densities at $|\eta| < 0.2$

The central primary-charged-particle density per unit pseudorapidity is considered to be a highly significant observable in minimum-bias studies, as its evolution with centre-of-mass energy can be easily studied and compared with predictions made by MC event-generator models. Here it is calculated as the average density within the $|\eta| < 0.2$ interval of the $\frac{1}{N_{\text{ev}}}\frac{\mathrm{d}N_{\text{ch}}}{\mathrm{d}\eta}$ distribution. The obtained values for all phase-space regions considered in this study, for particles with mean lifetime $\tau > 300\,\text{ps}$ (thus excluding strange baryons), are shown in Table 7.17. The values are compared to those obtained with the previously used fiducial definition, $\tau > 30\,\text{ps}$, using an extrapolation factor of 1.012 ± 0.004 (for $p_{\text{T}} > 100$ MeV) or 1.025 ± 0.008 (for $p_{\text{T}} > 500$ MeV), which accounts for the fraction of charged strange baryons predicted by EPOS LHC simulation (see Section 7.4.3.2). The results of the measurements at $\sqrt{s} = 13$ TeV [137, 138] are also shown for comparison.

Energy	**Phase Space**		$\mathbf{1/N_{ev} \cdot dN_{ch}/d\eta}$ **at** $\mathbf{|\eta| < 0.2}$	
$\sqrt{s}$	$n_{\text{ch}} \geq$	$p_{\text{T}} >$	$\tau > 300\,\text{ps}$ (fiducial)	$\tau > 30\,\text{ps}$ (extrapolated)
8 TeV	2	100 MeV	5.644 ± 0.003 (stat) ± 0.103 (syst)	5.710 ± 0.003 (stat) ± 0.105 (syst)
8 TeV	1	500 MeV	2.477 ± 0.001 (stat) ± 0.031 (syst)	2.538 ± 0.001 (stat) ± 0.036 (syst)
8 TeV	6	500 MeV	3.684 ± 0.002 (stat) ± 0.036 (syst)	3.775 ± 0.002 (stat) ± 0.046 (syst)
8 TeV	20	500 MeV	6.501 ± 0.005 (stat) ± 0.054 (syst)	6.662 ± 0.005 (stat) ± 0.074 (syst)
8 TeV	50	500 MeV	12.398 ± 0.032 (stat) ± 0.142 (syst)	12.705 ± 0.032 (stat) ± 0.174 (syst)
13 TeV	2	100 MeV	6.422 ± 0.002 (stat) ± 0.096 (syst)	6.500 ± 0.002 (stat) ± 0.099 (syst)
13 TeV	1	500 MeV	2.874 ± 0.001 (stat) ± 0.033 (syst)	2.943 ± 0.001 (stat) ± 0.042 (syst)

Table 7.17: Central primary-charged-particle density $1/N_{\text{ev}} \cdot \mathrm{d}N_{\text{ch}}/\mathrm{d}\eta$ at $|\eta| < 0.2$ for the five different phase spaces considered in this study at $\sqrt{s} = 8$ TeV [3]. The results are given for the fiducial definition $\tau > 300\,\text{ps}$, as well as for the previously used fiducial definition $\tau > 30\,\text{ps}$ using extrapolation factors to account for the fraction of charged strange baryons predicted by EPOS LHC simulation. Available numbers from new ATLAS studies at $\sqrt{s} = 13$ TeV are added for comparison [137, 138].

Using only the previously used definition $\tau > 30\,\text{ps}$ (thus including strange baryons), a comparison of results for all phase-space regions and all energies available from ATLAS Minimum Bias analyses [3, 126, 137, 138] is shown in Table 7.18. In the two main phase spaces the total systematic uncertainties have been reduced by as much as 30–40 % in the presented 8 TeV study with respect to the previous measurement at $\sqrt{s} = 7$ TeV.

The centre-of-mass energy evolution of the charged-particle multiplicity per unit pseudorapidity at $|\eta| < 0.2$ is shown in Figure 7.48 for the three main phase spaces ($p_{\text{T}} > 100$ MeV, $n_{\text{ch}} \geq 2$), ($p_{\text{T}} > 500$ MeV, $n_{\text{ch}} \geq 1$), and ($p_{\text{T}} > 500$ MeV, $n_{\text{ch}} \geq 6$). Results from other available ATLAS measurements are shown in addition to those measured in this study. In order to make consistent comparisons with the previous measurements, the results at $\sqrt{s} = 8$ and 13 TeV are extrapolated for compatibility with the previously used fiducial definition $\tau > 30\,\text{ps}$. The results

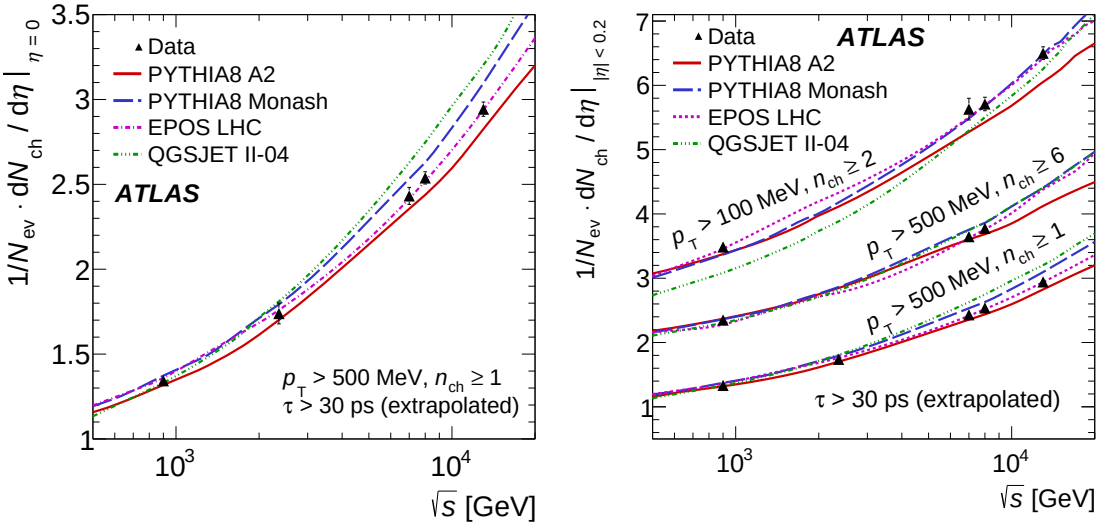
Phase Space		**Energy**	$\mathbf{1/N_{ev} \cdot dN_{ch}/d\eta}$ **at** $\mathbf{|\eta| < 0.2}$
$n_{ch} \geq$	$p_T >$	$\sqrt{s}$ [TeV]	$\tau > 30$ ps (extrapolated)
2	100 MeV	0.9	3.483 ± 0.009 (stat) ± 0.106 (syst)
		7	5.630 ± 0.003 (stat) ± 0.169 (syst)
		8	5.710 ± 0.003 (stat) ± 0.105 (syst)
		13	6.500 ± 0.002 (stat) ± 0.099 (syst)
1	500 MeV	0.9	1.343 ± 0.004 (stat) ± 0.027 (syst)
		2.36	1.740 ± 0.019 (stat) ± 0.058 (syst)
		7	2.423 ± 0.001 (stat) ± 0.050 (syst)
		8	2.538 ± 0.001 (stat) ± 0.036 (syst)
		13	2.943 ± 0.001 (stat) ± 0.042 (syst)
6	500 MeV	0.9	2.380 ± 0.009 (stat) ± 0.027 (syst)
		7	3.647 ± 0.002 (stat) ± 0.052 (syst)
		8	3.775 ± 0.002 (stat) ± 0.046 (syst)
20	500 MeV	8	6.662 ± 0.005 (stat) ± 0.074 (syst)
50	500 MeV	8	12.705 ± 0.032 (stat) ± 0.174 (syst)

Table 7.18: Central charged-particle density $1/N_{ev} \cdot dN_{ch}/d\eta$ at $|\eta| < 0.2$ (for particles with mean lifetime $\tau > 30$ ps, thus including strange baryons) for the five different phase-space regions considered in this study [3], compared with values from studies at different centre-of-mass energies where results are available [126, 137, 138]. The results at $\sqrt{s} = 8$ and 13 TeV are extrapolated by factors to account for the fractions of strange baryons predicted by Epos LHC simulation.

are compared with predictions made by various MC event-generator models, and among these the Epos LHC tune is found to describe the data best, followed by predictions made by the Pythia 8 tunes. As for previous spectra, the predictions made by the Qgsjet-II model tend to disagree with the data, although this tune also generally reproduces the energy dependence of the central charged-particle density.

7.10 Conclusion

This study presents measurements of distributions of primary-charged-particle multiplicities as well as pseudorapidity and transverse momentum spectra, based on 160 μb^{-1} of data produced in pp collisions at $\sqrt{s} = 8$ TeV that were recorded by ATLAS using a minimum-bias trigger. Novel results in high-multiplicity phase spaces are also presented, which have hitherto not been measured by ATLAS. Compared with previous studies at lower collision energies, a new fiducial definition of stable primary charged particles was employed to reduce model dependences. The presented measurements are found to have the highest precision of all ATLAS Minimum Bias



(a) (b)

Figure 7.48: Evolution of the average primary-charged-particle multiplicity per unit of pseudorapidity at $|\eta| < 0.2$ with rising centre-of-mass energy. Results are shown for the phase spaces (a) ($p_{\mathrm{T}} > 500$ MeV, $n_{\mathrm{ch}} \geq 1$) and (b) ($p_{\mathrm{T}} > 100$ MeV, $n_{\mathrm{ch}} \geq 2$), ($p_{\mathrm{T}} > 500$ MeV, $n_{\mathrm{ch}} \geq 1$) and ($p_{\mathrm{T}} > 500$ MeV, $n_{\mathrm{ch}} \geq 6$). The results at $\sqrt{s} = 8$ and 13 TeV are extrapolated to include strange baryons, using factors derived from Epos LHC simulation. The data (markers) are compared to several particle-level predictions made by MC generator models (curves). The total uncertainty on the data is shown by vertical error bars. Plots taken from Refs. [3, 138].

measurements of pp collisions from the Run-1 data-taking period at the LHC, thanks to a significant reduction of the leading systematic uncertainty, thus providing a powerful test of QCD-based models and contributing to their further improvement. The data were compared with other available measurements and with predictions made by four different MC models: Pythia 8 A2, Pythia 8 Monash, Epos LHC and Qgsjet-II.

In the most inclusive phase space at $p_{\mathrm{T}} > 100$ MeV and $n_{\mathrm{ch}} \geq 2$, the Epos LHC tune models all measured distributions of the 8 TeV data far more accurately than Qgsjet-II and both Pythia tunes, among which the Pythia 8 Monash tune makes significantly better predictions than Pythia 8 A2. The advantage of Epos LHC over its competitors is most pronounced in the multiplicity distribution (where both Pythia tunes fail to describe the low-n_{ch} as well as high-n_{ch} regions of the spectrum) and in the lower end of the p_{T} distribution. The Epos LHC tune also does an excellent job at modelling the η distribution and the $\langle p_{\mathrm{T}} \rangle$ versus n_{ch} distribution.

In the phase spaces with higher momentum threshold at $p_{\mathrm{T}} > 500$ MeV, the Pythia 8 A2 and Epos LHC tunes describe the pseudorapidity distributions best, while among all considered models Qgsjet-II consistently shows the worst performance; all models tend to predict a flatter η distribution than found in the data. In the high-multiplicity phase spaces at $n_{\mathrm{ch}} \geq 20$ and especially $n_{\mathrm{ch}} \geq 50$, the high-p_{T} tails are increasingly misrepresented by all generators except

Pythia 8 Monash, which exhibits a stable performance in modelling the p_{T} distribution with good accuracy. The multiplicity distribution in the $p_{\mathrm{T}} > 500$ MeV phase spaces is well described by Pythia 8 A2 up to $n_{\mathrm{ch}} \approx 60$, while Pythia 8 Monash exhibits difficulties to model the low-n_{ch} region; again both Pythia 8 tunes fail to describe the high-multiplicity tails. On average Epos LHC does a fair job in modelling this distribution, however it predicts a too high event rate between $40 < n_{\mathrm{ch}} < 70$ and a too low event rate in the high-multiplicity tails beyond $n_{\mathrm{ch}} > 100$. The $\langle p_{\mathrm{T}} \rangle$ versus n_{ch} distribution is modelled reasonably well by Pythia 8 Monash followed by Epos LHC (which predicts a slightly too low average transverse momentum) and Pythia 8 A2. The Qgsjet-II tune shows the worst performance of all considered models.

In summary, none of the compared tunes clearly stands out for the $p_{\mathrm{T}} > 500$ MeV phase spaces; the lower multiplicity thresholds favour Pythia 8 A2 over Pythia 8 Monash, and vice versa for the higher multiplicity thresholds, while Epos LHC shows a fair performance in modelling all final distributions. However in the most inclusive $p_{\mathrm{T}} > 100$ MeV phase space, the best performance is clearly achieved by the Epos LHC tune, followed by the Pythia 8 Monash and Pythia 8 A2 tunes.

The performances of three additional MC tunes are discussed in Appendix A.8.

Chapter 8

Summary and Outlook

Summary of Results This thesis presents measurements of primary-charged-particle distributions from inelastic pp collisions at $\sqrt{s} = 8$ TeV centre-of-mass energy recorded with the ATLAS detector at the LHC in 2012. A minimum-bias trigger was utilised to select events that correspond to a dataset with $160\,\mu\text{b}^{-1}$ integrated luminosity. Data-driven as well as simulation-based corrections for detector effects were applied to the reconstructed data.

The presented measurements introduce a new fiducial definition for stable primary charged particles, given by a mean lifetime of $\tau > 300\,\text{ps}$ instead of the previously used definition $\tau > 30\,\text{ps}$. This change, which has also been implemented by the analogous new 13 TeV measurements, was required to reduce model-dependent uncertainties due to different predicted fractions of strange baryons by the MC generators, as these particles have very low reconstruction efficiencies. This source of systematic uncertainty, arising from particle composition, has previously not been taken into account. For comparison with previously published results, a simulation-based extrapolation was performed to obtain central charged-particle densities following the previously used fiducial definition.

A high precision of the results was made possible by detailed studies of the ID material, including measurements of the SCT extension efficiency to probe the material between the Pixel and SCT detectors. These studies led to an improved final Run-1 geometry description of ATLAS with a decreased material uncertainty of $\pm 5\,\%$, which benefits a wide range of Run-1 as well as Run-2 physics analyses. Owing to these improvements in understanding the ID material, a reduction of the leading systematic uncertainty of the presented 8 TeV measurements by as much as $50\,\%$ with respect to earlier measurements at lower energies up to 7 TeV was achieved.

Distributions of primary-charged-particle multiplicity (n_{ch}), pseudorapidity (η) and transverse-momentum (p_{T}) distributions, and average transverse-momentum spectra as a function of multiplicity, are presented in five different phase spaces. These include event selections at high multiplicities which have hitherto not been measured by ATLAS. In the most inclusive phase space, $p_{\text{T}} > 100$ MeV and $n_{\text{ch}} \geq 2$, the central charged-particle density $1/N_{\text{ev}} \cdot \mathrm{d}N_{\text{ch}}/\mathrm{d}\eta|_{|\eta|<0.2}$ was measured to be 5.64 ± 0.10. In the phase spaces with a higher transverse-momentum threshold, given by $p_{\text{T}} > 500$ MeV and $n_{\text{ch}} \geq 1$, 6, 20 and 50, the central charged-particle densities were measured to be 2.477 ± 0.031, 3.68 ± 0.04, 6.50 ± 0.05 and 12.40 ± 0.15, respectively.

The extrapolated result corresponding to the previously used fiducial definition, $\tau > 30\,\text{ps}$, was 5.71 ± 0.11 for the most inclusive phase space. In the phase spaces with the higher transverse-momentum threshold, the central charged-particle densities were determined to be 2.54 ± 0.04, 3.78 ± 0.05, 6.66 ± 0.07 and 12.71 ± 0.18, respectively. These results provide the most precise ATLAS measurements of primary-charged-particle distributions using data from Run-1 of the LHC. The total systematic uncertainties of the main results quoted above are as much as 30–40 % smaller than the previous highest-precision ATLAS measurements at lower energies.

The data were compared with predictions made by various MC generator models, and the overall best description of the data was found to be given by the EPOS LHC model, followed by the PYTHIA 8 A2 and MONASH models. EPOS LHC predictions describe the data better than other tunes in the low-momentum regime of the most inclusive measured phase space, $p_{\text{T}} > 100$ MeV and $n_{\text{ch}} \geq 2$, and therefore also give the most accurate prediction of the central charged-particle density in this phase space. However, the PYTHIA 8 MONASH tune provides the most accurate description of the data above $p_{\text{T}} > 1$ GeV in all measured transverse-momentum distributions, where other models show large discrepancies especially towards high p_{T}. The EPOS LHC model makes the best predictions of the energy evolution of the central charged-particle density in the three phase spaces which have been measured by ATLAS at centre-of-mass energies between 0.9 and 13 TeV, followed by both PYTHIA 8 tunes. Among all considered models, the QGSJET-II tune shows the largest discrepancies with respect to the data.

Impact and Outlook The results of the presented 8 TeV measurements of primary-charged-particle distributions were recently published in Ref. [3], and provide valuable constraints for future tuning efforts and a better understanding of soft-QCD physics models. Owing to their better precision with respect to earlier measurements, these measurements can thus be regarded as partially superseding previous results at 7 TeV centre-of-mass energy. They are therefore expected to provide a new standard for the tuning of MC generator models together with the analogous new Run-2 measurements at 13 TeV [135, 137, 138]. The Run-2 measurements face new challenges compared to Run-1, in particular a modified ID geometry with the new IBL. While imperfect knowledge of the updated geometry and material description introduced a new source of systematic uncertainty, the new Run-2 measurements benefit from the improved reconstruction performance due to the IBL, and from the reduced overall material uncertainty of the ID thanks to the efforts made during Run-1, including the material studies presented in this thesis. It can therefore be expected that future measurements of primary-charged-particle distributions will yield results with even higher precision, which can provide better constraints for the tuning of phenomenological soft-QCD models. All of these studies are expected to make significant contributions to our understanding of non-perturbative QCD physics.

Appendix A

Additional Plots for the Charged-Particle Density Measurement

A.1 Additional studies of MBTS trigger efficiencies

During the low-mu run 200805 all MBTS trigger modules were fully operational, and the small differences between their individual efficiencies were thus neglected.

The stability of the MBTS trigger efficiencies during the low-μ run 200805 was assessed by studying the possible dependence of average trigger efficiencies on the luminosity blocks (LB). The efficiencies of both triggers (`L1_MBTS_1` and `L1_MBTS_2`) were calculated separately for each LB from the range used in the analysis ($215 \leq \text{LB} \leq 395$), using a standard sample of events triggered by `EF_mbSpTrk`. Within uncertainties, no obvious systematic deviation from a common constant value was found in any LB subrange. The fit of a possible linear dependence shows only a tiny increase of the efficiency: during the whole run, the average efficiency increases from 0.99758 to 0.99790 (i.e. by 0.00032) for `L1_MBTS_1`, and from 0.99382 to 0.99417 (i.e. by 0.00035) for `L1_MBTS_2`. These systematic increases in the sub-per-mille region can be considered negligible, as they are smaller than statistical fluctuations and other uncertainties of the calculated efficiencies, and would only have a completely negligible impact on the final results of this analysis.

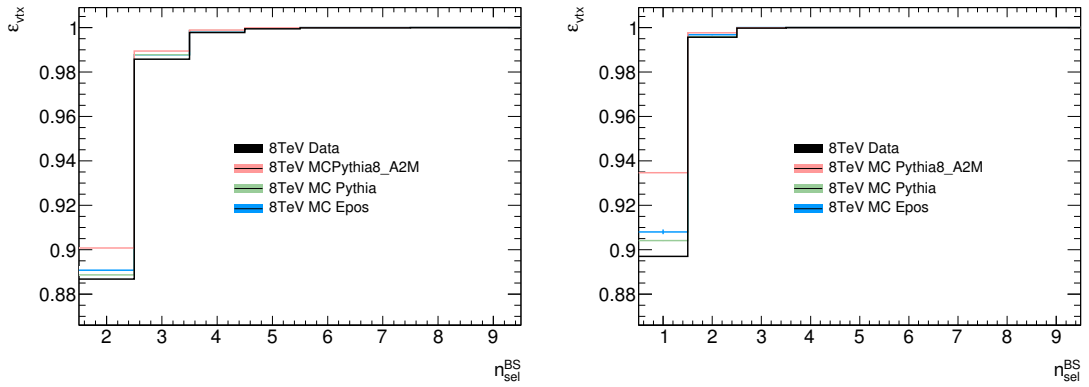
A.2 Additional studies of vertex reconstruction efficiency

A.2.1 Comparison of MC tunes

A comparison of vertex reconstruction efficiencies between the different MC tunes and data is shown in Figure A.1. The differences between MC tunes stem from different double-, single- and non-diffractive compositions in each tune. Non-diffractive events have the highest efficiency, which is significantly higher than the efficiencies in single- and double-diffractive events.

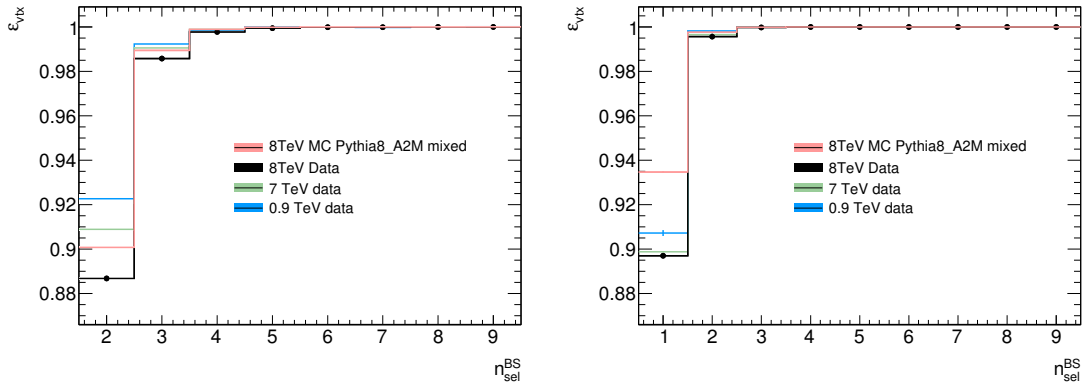
A.2.2 Comparison with previous analyses

A comparison of vertex reconstruction efficiencies between the current analysis at $\sqrt{s} = 8$ TeV and previous analyses at $\sqrt{s} = 0.9$ and 7 TeV is presented in Figure A.2.



(a) **(b)**

Figure A.1: Vertex reconstruction efficiency for different MC tunes and data in (a) the $p_{\mathrm{T}} > 100, n_{\mathrm{ch}} \geq 2$ and (b) the $p_{\mathrm{T}} > 500, n_{\mathrm{ch}} \geq 1$ phase space. The errors represent statistical fluctuations in the samples.



(a) **(b)**

Figure A.2: Vertex reconstruction efficiency as a function of n_{sel} compared with previously studied centre-of-mass energies for (a) the $p_{\mathrm{T}} > 100, n_{\mathrm{ch}} \geq 2$ and (b) the $p_{\mathrm{T}} > 500, n_{\mathrm{ch}} \geq 1$ phase space. The statistical errors are comparable with the size of the line.

A.2.3 Detailed study of efficiencies in low-multiplicity events

The $\Delta z_0^{\mathrm{min}}$ parameterisation of the vertex reconstruction efficiency in events with $n_{\mathrm{sel}} = 2$ in the most inclusive phase space ($p_{\mathrm{T}} > 100$ MeV) is shown in Figure A.3 for data and Figure A.4 for MC simulation. The $\Delta z_0^{\mathrm{min}}$ dependence increases at lower $p_{\mathrm{T}}^{\mathrm{min}}$. The parameterisation is therefore subdivided into several $p_{\mathrm{T}}^{\mathrm{min}}$ categories, and inclusive distributions for the entire range of events are also given.

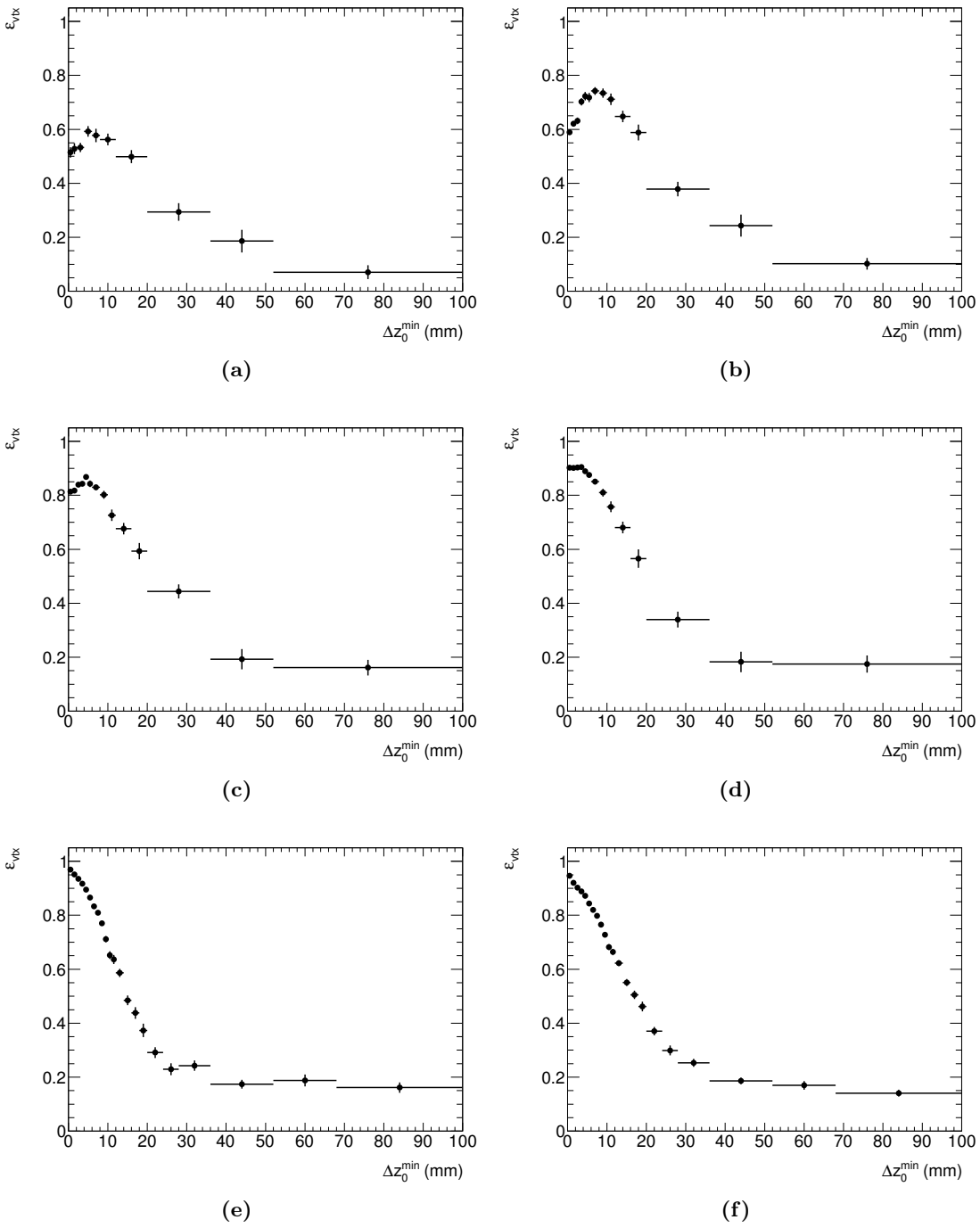


Figure A.3: Vertex reconstruction efficiency in data events with $n_{\mathrm{sel}} = 2$ in the most inclusive phase space ($p_{\mathrm{T}} > 100$ MeV) as a function of $\Delta z_0^{\mathrm{min}}$, separated by $p_{\mathrm{T}}^{\mathrm{min}}$ categories: (a) $100 < p_{\mathrm{T}}^{\mathrm{min}} < 125$ MeV, (b) $125 < p_{\mathrm{T}}^{\mathrm{min}} < 150$ MeV, (c) $150 < p_{\mathrm{T}}^{\mathrm{min}} < 175$ MeV, (d) $175 < p_{\mathrm{T}}^{\mathrm{min}} < 200$ MeV, (e) $p_{\mathrm{T}}^{\mathrm{min}} > 200$ MeV, and (f) inclusive distribution of all events.

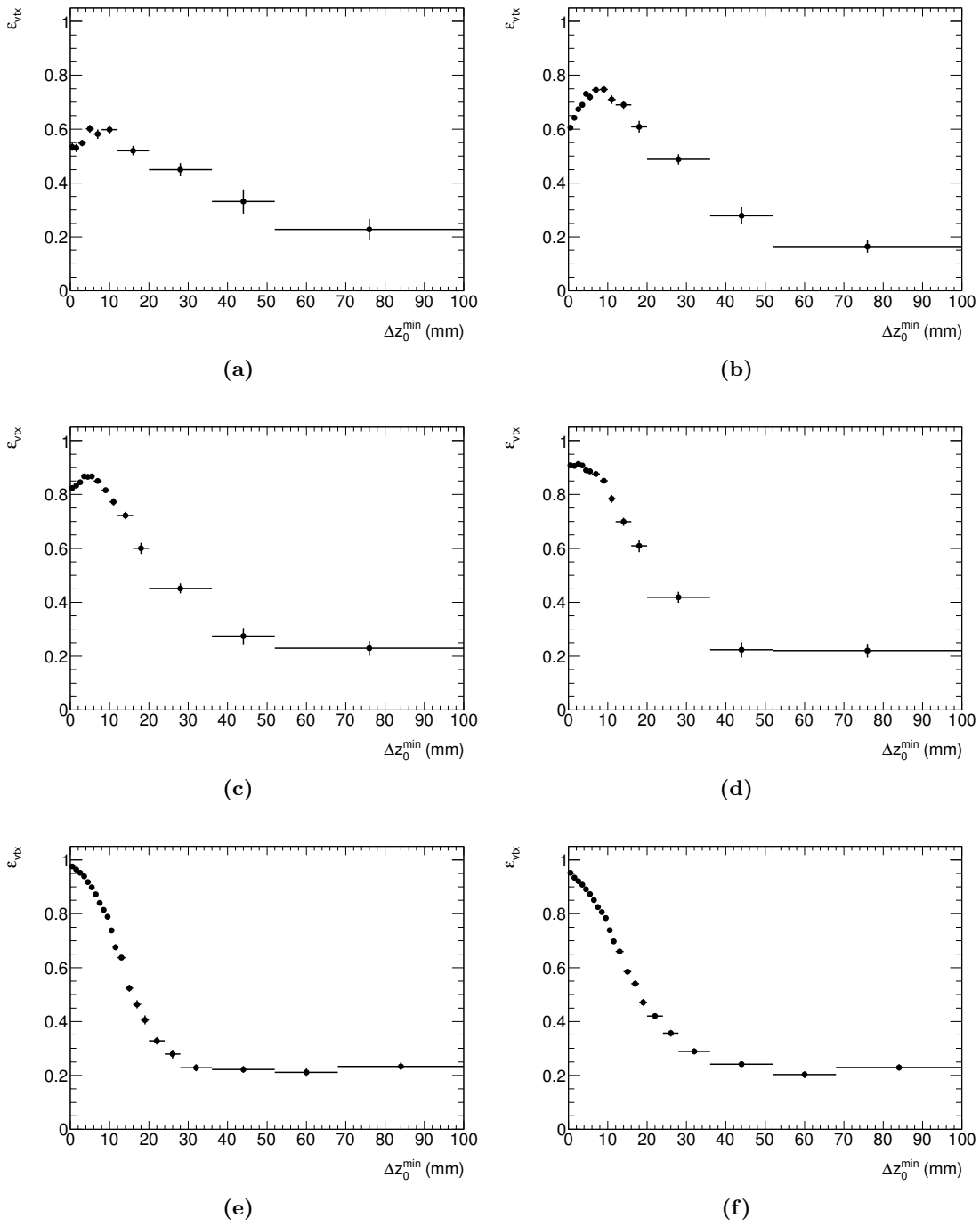
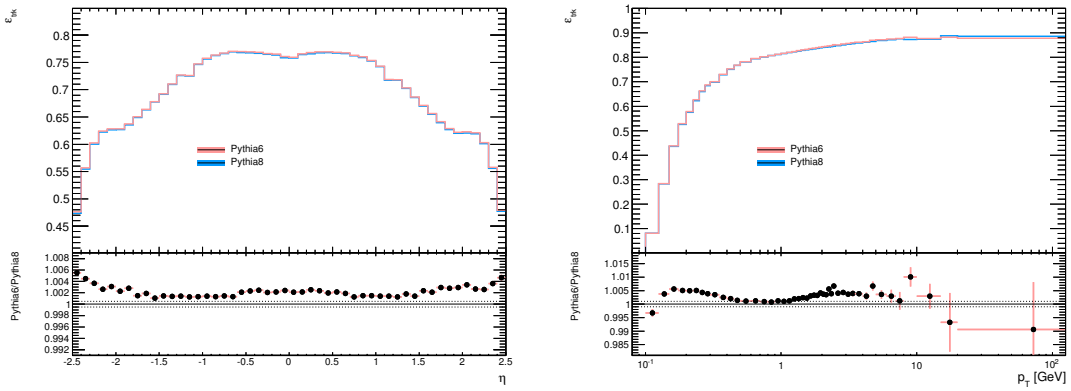
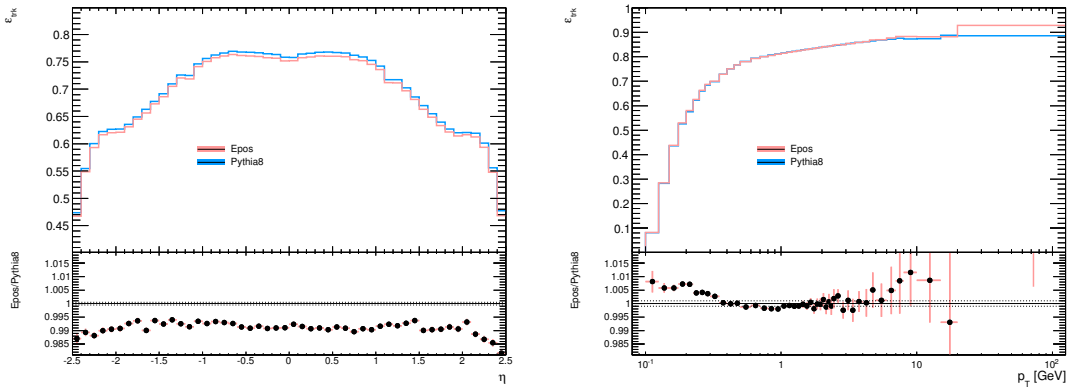


Figure A.4: Vertex reconstruction efficiency in MC simulation events with $n_{\mathrm{sel}} = 2$ in the most inclusive phase space ($p_{\mathrm{T}} > 100$ MeV) as a function of $\Delta z_0^{\mathrm{min}}$, separated by $p_{\mathrm{T}}^{\mathrm{min}}$ categories: (a) $100 < p_{\mathrm{T}}^{\mathrm{min}} < 125$ MeV, (b) $125 < p_{\mathrm{T}}^{\mathrm{min}} < 150$ MeV, (c) $150 < p_{\mathrm{T}}^{\mathrm{min}} < 175$ MeV, (d) $175 < p_{\mathrm{T}}^{\mathrm{min}} < 200$ MeV, (e) $p_{\mathrm{T}}^{\mathrm{min}} > 200$ MeV, and (f) inclusive distribution of all events.



(a) (b)

Figure A.5: Track reconstruction efficiency as a function of (a) η or (b) p_{T} of the track, comparing the Pythia 6 AMBT2B and Pythia 8 A2 tunes in the $p_{\mathrm{T}} > 100$ MeV phase space.



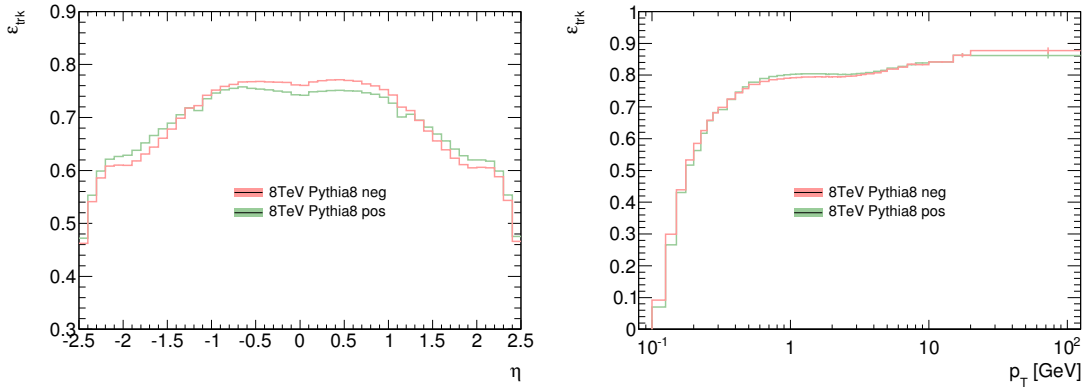
(a) (b)

Figure A.6: Track reconstruction efficiency as a function of (a) η or (b) p_{T} of the track, comparing the Epos LHC and Pythia 8 A2 tunes in the $p_{\mathrm{T}} > 100$ MeV phase space.

A.3 Additional studies of track reconstruction efficiency

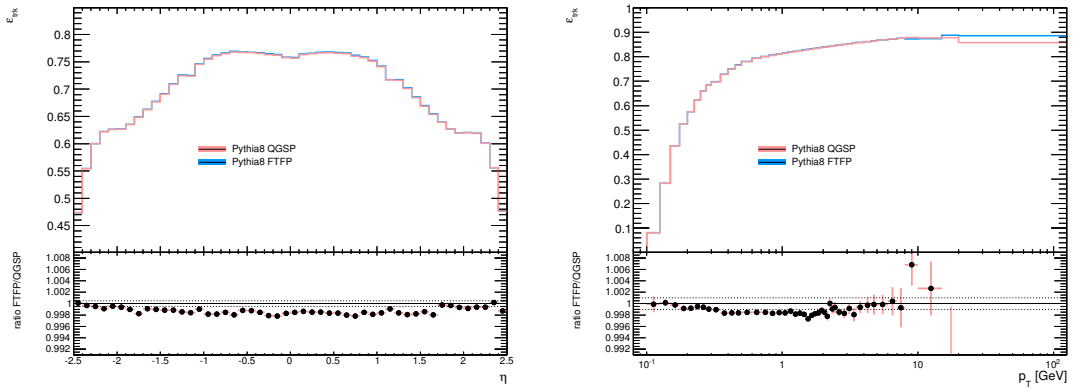
A comparison of track reconstruction efficiencies between the different MC tunes is shown in Figure A.5 and Figure A.6. From the differences in p_{T} to the baseline Pythia 8 A2 tune, a systematic uncertainty between $0.2 - 1.0\%$ due to particle composition can be derived. The differences in η result mainly from the different p_{T} spectra of the MC generators, which are folded into the η distribution.

A comparison of track reconstruction efficiencies of positively and negatively charged particles is shown in Figure A.7. The systematic uncertainties due to different ratios of positively and negatively charged reconstructed tracks in data and MC have been evaluated and were found to be negligible.



(a) (b)

Figure A.7: Track reconstruction efficiency of positively and negatively charged particles as a function of (a) η or (b) p_{T} of the track in the $p_{\mathrm{T}} > 100$ MeV phase space.



(a) (b)

Figure A.8: Difference in track reconstruction efficiency between the old and new physics lists, QGSP and FTFP, as a function of (a) η or (b) p_{T} of the track.

The impact of the new choice of the physics list for MC simulation, which was changed from FTFP to QGSP, on the track reconstruction efficiency is shown in Figure A.8. The difference between the two physics lists is assigned as a flat systematic uncertainty of 0.3 %, which is a conservative estimate. The main source for this difference is found in the improved simulation of antiprotons in the FTFP physics list, resulting in a higher track reconstruction efficiency of these particles.

A.4 Additional detector performance plots

Figure A.9 shows the average number of hits per track in the innermost Pixel layer-0 and the whole Pixel detector, as well as the average number of hits and inactive modules per track in the SCT detector, respectively, for detailed transverse momentum p_{T} regions of reconstructed

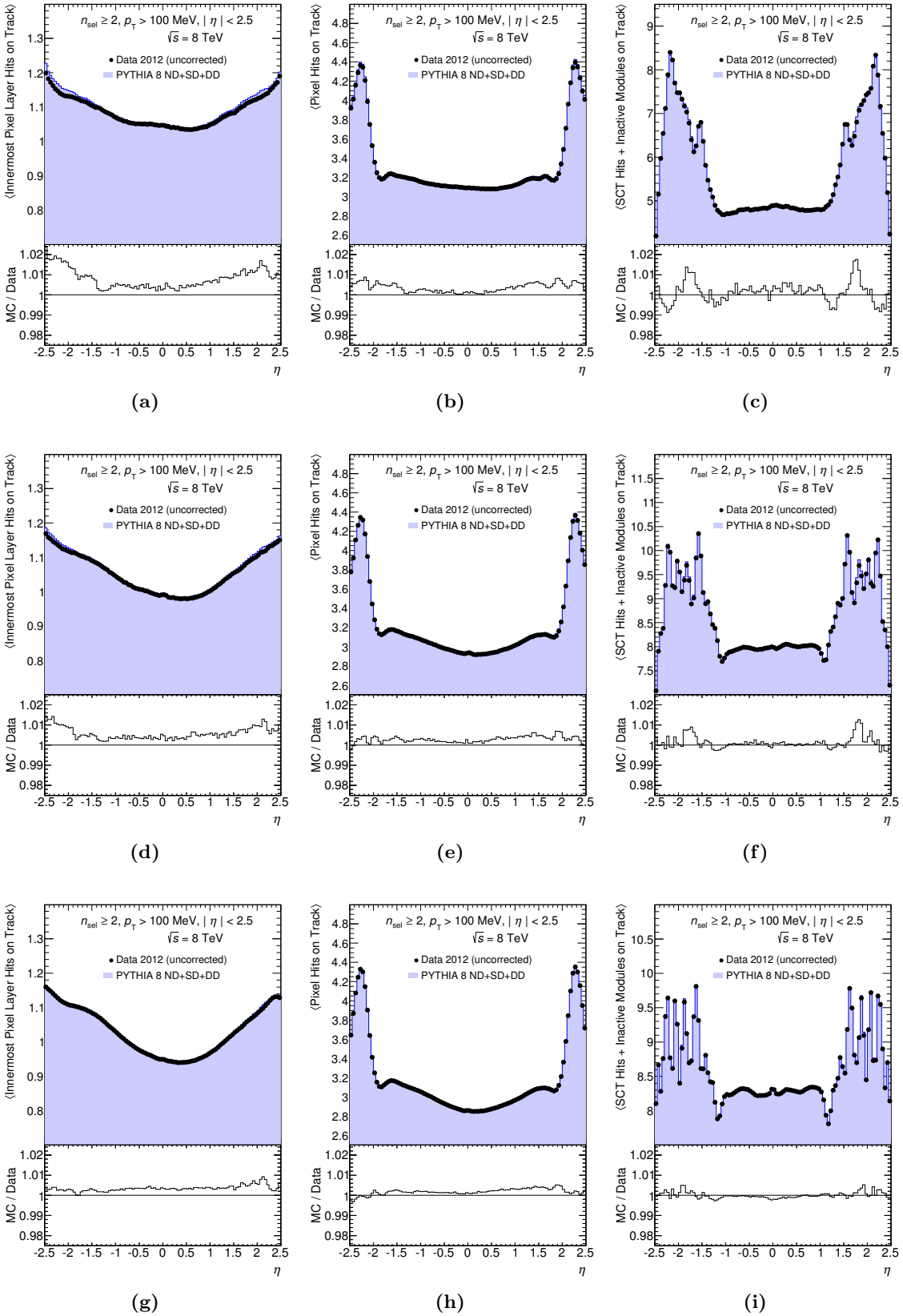


Figure A.9: Data/MC comparison of the average number of hits per track in the (a,d,g) innermost layer of the Pixel detector, (b,e,h) whole Pixel detector, and (c,f,i) average number of hits and inactive modules per track in the SCT detector, for detailed transverse momentum p_{T} regions of reconstructed tracks: (a-c) $100 < p_{\mathrm{T}} < 200$ MeV, (d-f) $200 < p_{\mathrm{T}} < 300$ MeV, (g-i) $p_{\mathrm{T}} > 300$ MeV.

tracks. The distributions show good agreement between data and MC simulation, with slight discrepancies due to imperfect modelling of detector conditions in MC simulation. The effect of these discrepancies on the track reconstruction efficiency and its systematics is found to be negligible.

A.5 Unfolding tests

The full correction procedure for the charged-particle multiplicity spectrum includes a Bayesian unfolding technique as described in Section 7.6.2. While this procedure is designed to be independent of the choice of employed physics models, migration effects into and out of the region of analysis can still introduce non-negligible model-dependence effects. For example, a model-dependent variation in the expected transverse-momentum distributions with multiplicity can change the probabilities to observe various multiplicities of selected tracks, n_{sel}, for a given true multiplicity value, n_{ch}, thus affecting the migration matrix.

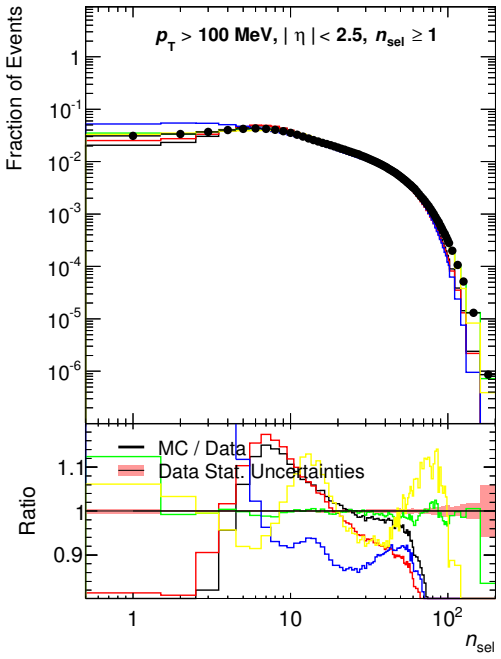


Figure A.10: Distributions of charged-particle multiplicity before unfolding. Data are compared to Pythia 8 A2 (black), Pythia 8 A2 for which the true multiplicity distribution is reweighted (green, see Section A.5.1), Pythia 8 A2 for which the diffractive contributions are rescaled (red, see Section A.5.2), Pythia 6 AMBT2B (blue), Epos LHC (yellow).

Figure A.10 presents multiplicity distributions of reconstructed tracks in various MC models and data prior to the unfolding step. None of the models can perfectly reproduce the data. The robustness of the correction procedure to different model assumptions was therefore evaluated via closure tests, using MC spectra as pseudo-data, with different techniques:

1. a data-driven method,

2. a variation of diffractive components,
3. unfolding QCD-based models with other models.

The effective efficiency factor used in the final step as a correction for migrations over the boundary of the analysis region is adjusted for each considered model; the values used range from 0.658 to 0.668 in the most inclusive phase space, as given in Table 7.15.

A.5.1 Data-driven reweighting of multiplicity spectra

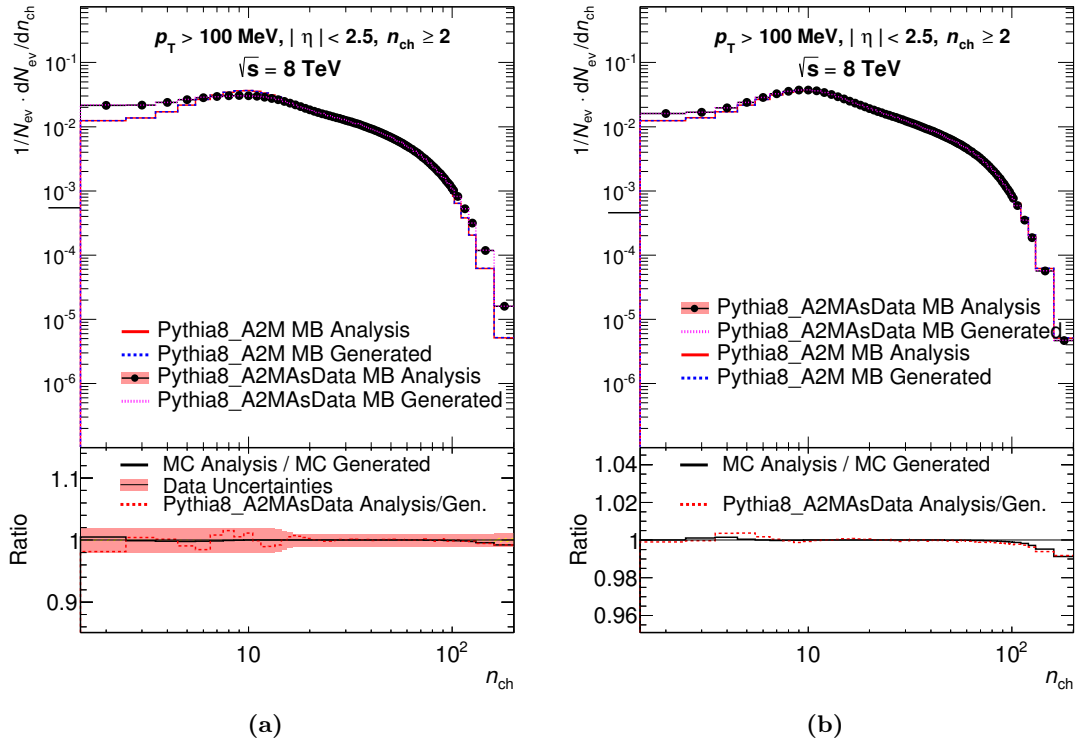


Figure A.11: The charged-particle-multiplicity spectra for reweighted Pythia 8 A2-generated events: (a) Unfolded with the nominal migration matrix obtained with unweighted events; the dashed line represents the unfolded pseudo-data following full convergence of the method, and the solid line represents the closure test of the nominal distribution. (b) Unfolded with adjusted relative diffractive contributions; the data are unfolded with the nominal migration matrix obtained with unweighted events.

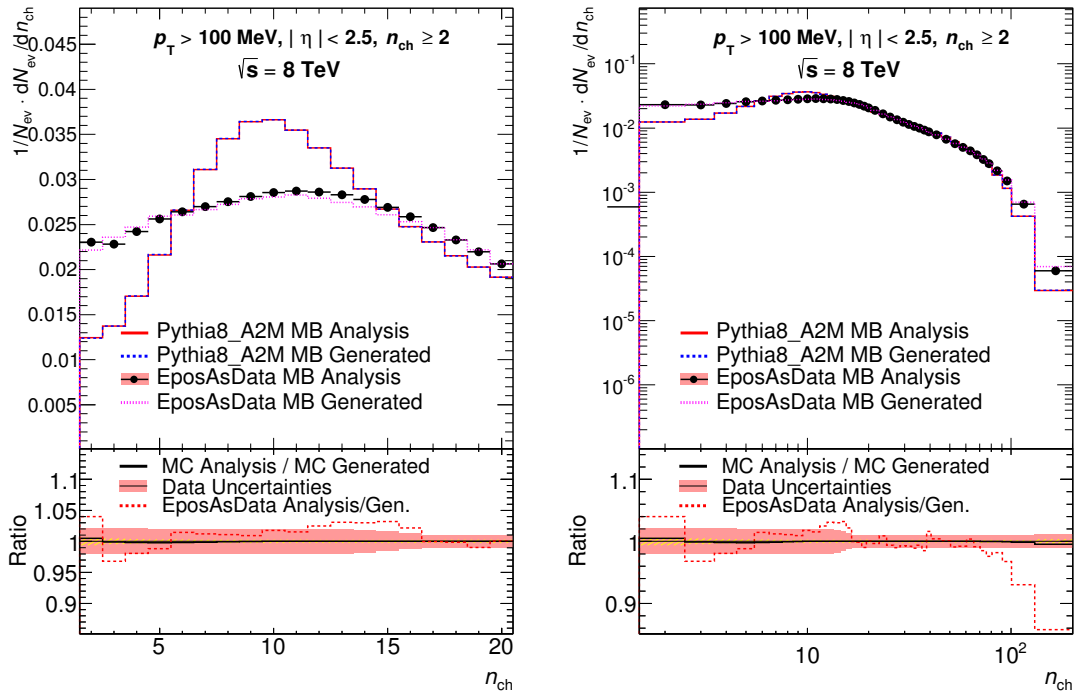
A truth-level reweighting of the Pythia 8 A2 multiplicity distribution, using a smooth reweighting function, was made such that a good description of the data was achieved at the detector level. Figure A.10 shows the multiplicity spectrum, obtained after applying event and track weights, compared with multiplicity-weighted and unweighted events generated with Pythia 8 A2. The nominal spectrum was used to extract the migration matrix, and the reweighted spectrum was used as pseudo-data. Figure A.11a shows the results of the reweighting procedure. After five iterations of the unfolding procedure, the unfolded distribution agrees well

with the reweighted truth-level distribution, implying a sensitivity to variations in the migration matrix of up to $\sim 2\,\%$ for each bin.

A.5.2 Variation of diffractive contributions

A reweighting of the different diffractive components can also change the shape of the predicted multiplicity distribution. For this work, a rescaling of the relative contributions (by a factor 1.5) of the single- and double-diffractive cross sections were made to the Pythia 8 A2 model. As shown in Figure A.10, the reweighted model achieves a better description of the data. This diffraction-adjusted multiplicity distribution was used as pseudo-data, and the unfolding was performed in three iterations, using the nominal migration matrix. Figure A.11b shows that the unfolding reproduces the true spectra to less than $\sim 1\,\%$. The apparent non-closure at high multiplicity ($n_{\mathrm{ch}} > 100$) can be attributed to statistical uncertainties.

A.5.3 Unfolding QCD-based models with other models



(a) (b)

Figure A.12: Unfolded distributions of charged-particle multiplicities at low multiplicities (a) and over the full multiplicity range (b). Data from Epos LHC are used as pseudo-data and the unfolding is made with the nominal migration matrix from Pythia 8 A2. The unfolded distribution of Pythia 8 A2 using the nominal migration mixture is also shown as a cross check of the unfolding procedure.

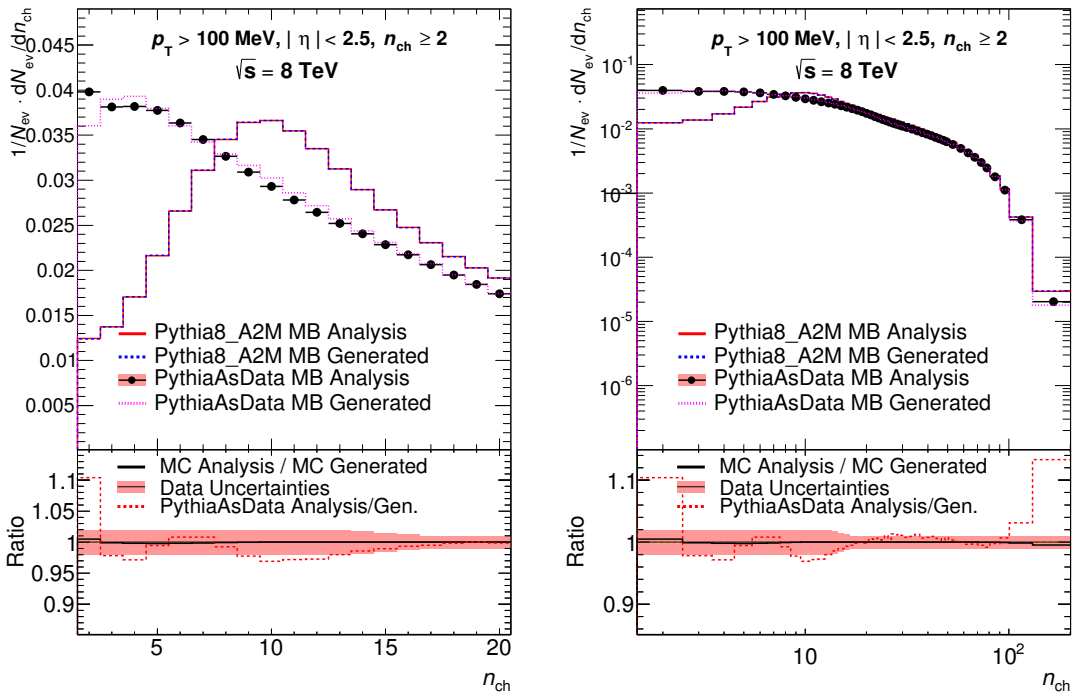


Figure A.13: Unfolded distributions of charged-particle multiplicities at low multiplicities (a) and over the full multiplicity range (b). Data from Pythia 6 AMBT2B are used as pseudo-data and the unfolding is made with the nominal migration matrix from Pythia 8 A2. The unfolded distribution of Pythia 8 A2 using the nominal migration mixture is also shown as a cross check of the unfolding procedure.

As a further test of the unfolding robustness, one MC model can be used as pseudo-data while another model provides the migration matrix for the unfolding procedure. Using the nominal migration matrix derived from the unweighted Pythia 8 A2 data, pseudo-data from Epos LHC and Pythia 6 AMBT2B were unfolded with three and five iterations, respectively. The results of the unfolding for Epos LHC (Pythia 6 AMBT2B) are shown in Figure A.12 (A.13). A non-closure within $< 3\,\%$ is obtained, except at very low and high multiplicities. As a cross-check of the unfolding procedure, the results of the unfolding of Pythia 8 A2 using the nominal matrix are also shown, giving almost complete closure.

A.5.4 Conclusion of Unfolding Tests

The robustness of the unfolding procedure has been tested and quantified using three techniques, of which the first (data-driven) method was found to be the most representative estimate for the systematics of the unfolding procedure. Taking also the other techniques into account, the systematic uncertainty due to unfolding (including systematics due to statistical fluctuations in the unfolding matrix) is estimated by 1.5 % up to 12 % in the most inclusive phase space, and

1 % up to 9 % in the phase space with the higher transverse-momentum threshold $p_{\mathrm{T}} > 500$ MeV, with highest systematic uncertainties applied to the highest-multiplicity event categories (see Figure 7.38c and Figure 7.39c in Section 7.8). It is taken into account as one of the contributions to the overall systematic uncertainty of the unfolded charged-particle multiplicity spectra.

The same techniques are applied to determine the unfolding systematics of the final p_{T} spectra; their total unfolding systematics are also shown in Figures 7.38b to 7.42b in Section 7.8 and are taken into account in the final results.

A.6 Stability Tests of Charged-Particle Density Measurements

A.6.1 Beamspot Properties during data-taking in Run 200805

In order to assess the stability of the results in the charged particle density measurement at $\sqrt{s} = 8$ TeV, as a preliminary step the beam-spot properties from the low-μ run 200805 were analysed. It was found that the beam-spot z coordinate moves by a few mm during the luminosity block range with good data quality ($215 \leq \mathrm{lbn} \leq 395$) while the beam-spot size remains approximately constant. Figure A.14 shows various parameters of the beam-spot during run 200805 as a function of the luminosity block.

A.6.2 Analysis Results with Subsets of $\sqrt{s} = 8$ TeV Data

In order to assess the stability of the results in the charged particle density measurement at $\sqrt{s} = 8$ TeV, the analysis was conducted with subsets of the data from the low-μ run 200805, based on the average number of interactions per bunch crossing μ in each luminosity block, as shown in Figure 7.1. It was found that the inclusive as well as the central charged particle multiplicity in the $p_{\mathrm{T}} > 100$ MeV phase space remains stable during the run, as well as in different μ-intervals.

A.6.2.1 Detector-Level Distributions for Subsets of the Data

Figure A.15 shows the raw detector-level distributions of charged particles using different subsets of the data. The distributions show good agreement, except for fluctuations at high p_{T}.

A.6.2.2 Comparison of Final Distributions for Subsets of the Data

Figure A.16 shows a comparison of final distributions in the most inclusive phase space, comparing different subsets of the data (based on the average number of interactions per bunch crossing $\langle\mu\rangle$ in each luminosity block) against the full data sample. The small fluctuations in the η distribution are found to be negligible,[1] while the visible discrepancies in the high-momentum

[1] The small fluctuations in the comparison of η distributions stem from not having applied a dedicated vertex z reweighting on the MC samples for each subset of the data (the beamspot z-coordinate moves slightly during the run, yielding slightly different vertex z distributions in the data for the different μ-subsets). Ideally one would have to determine dedicated corrections (especially the η-dependent track reconstruction efficiency) for

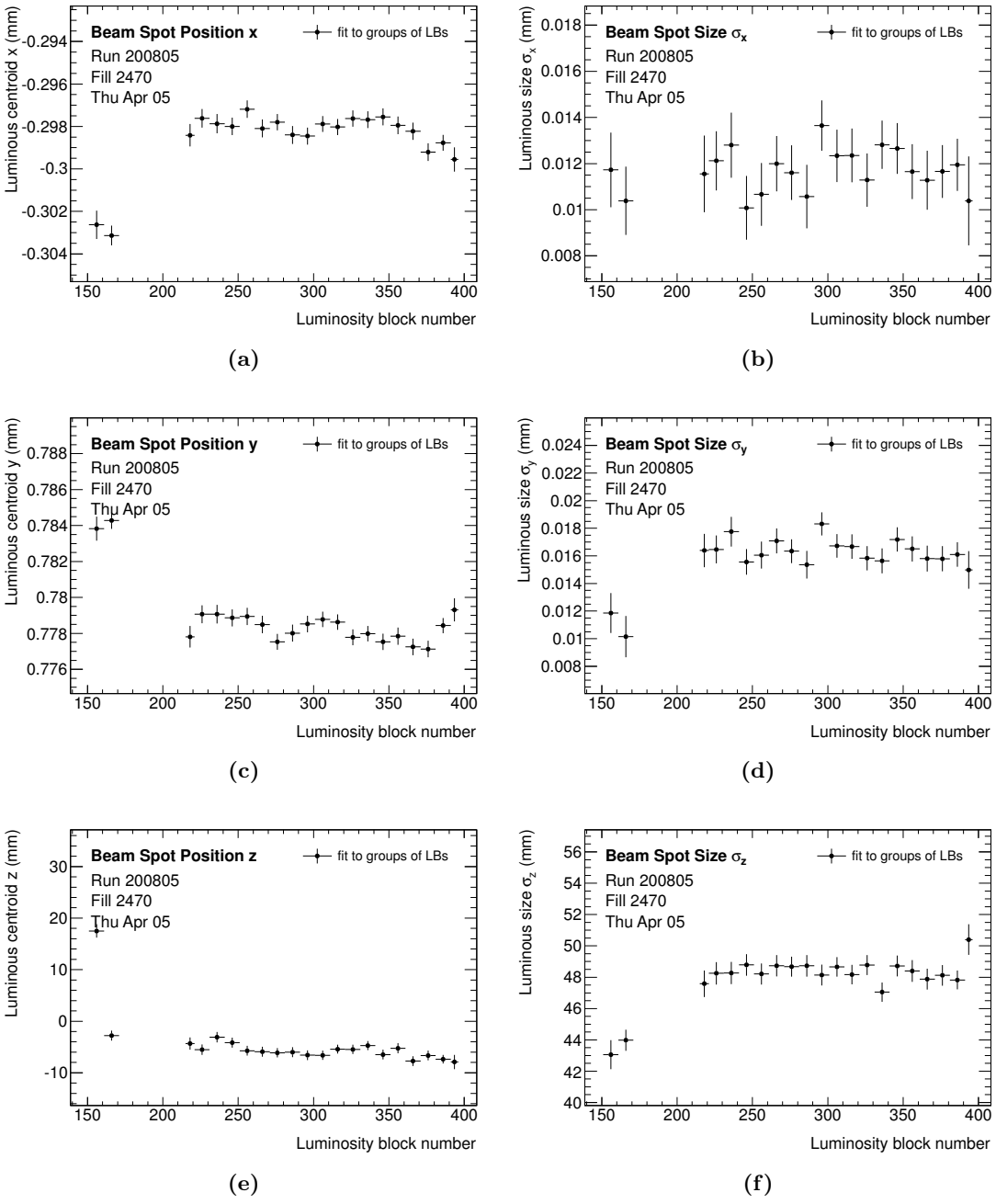


Figure A.14: Beam-spot parameters of low-μ run 200805 as a function of the luminosity block.

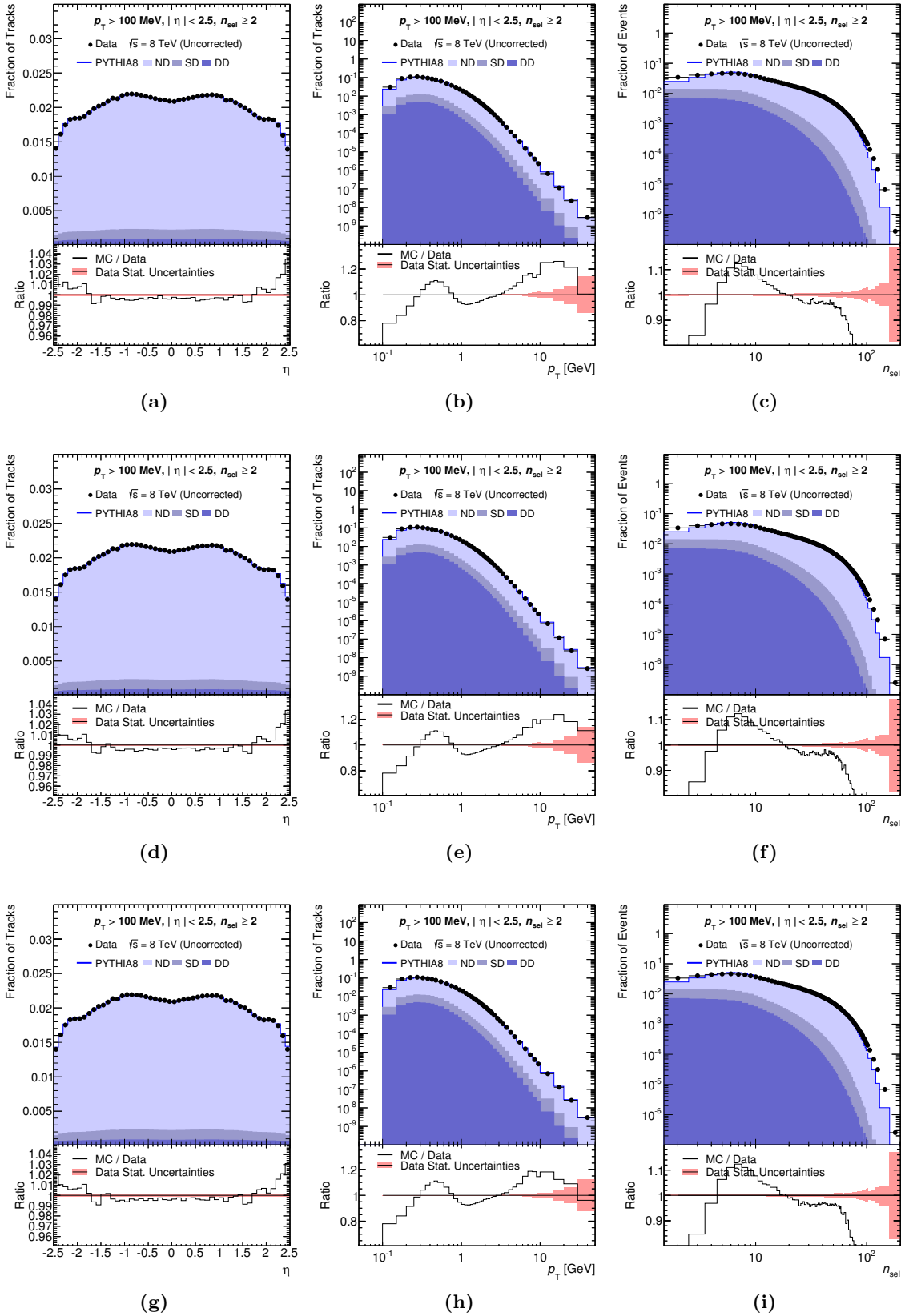
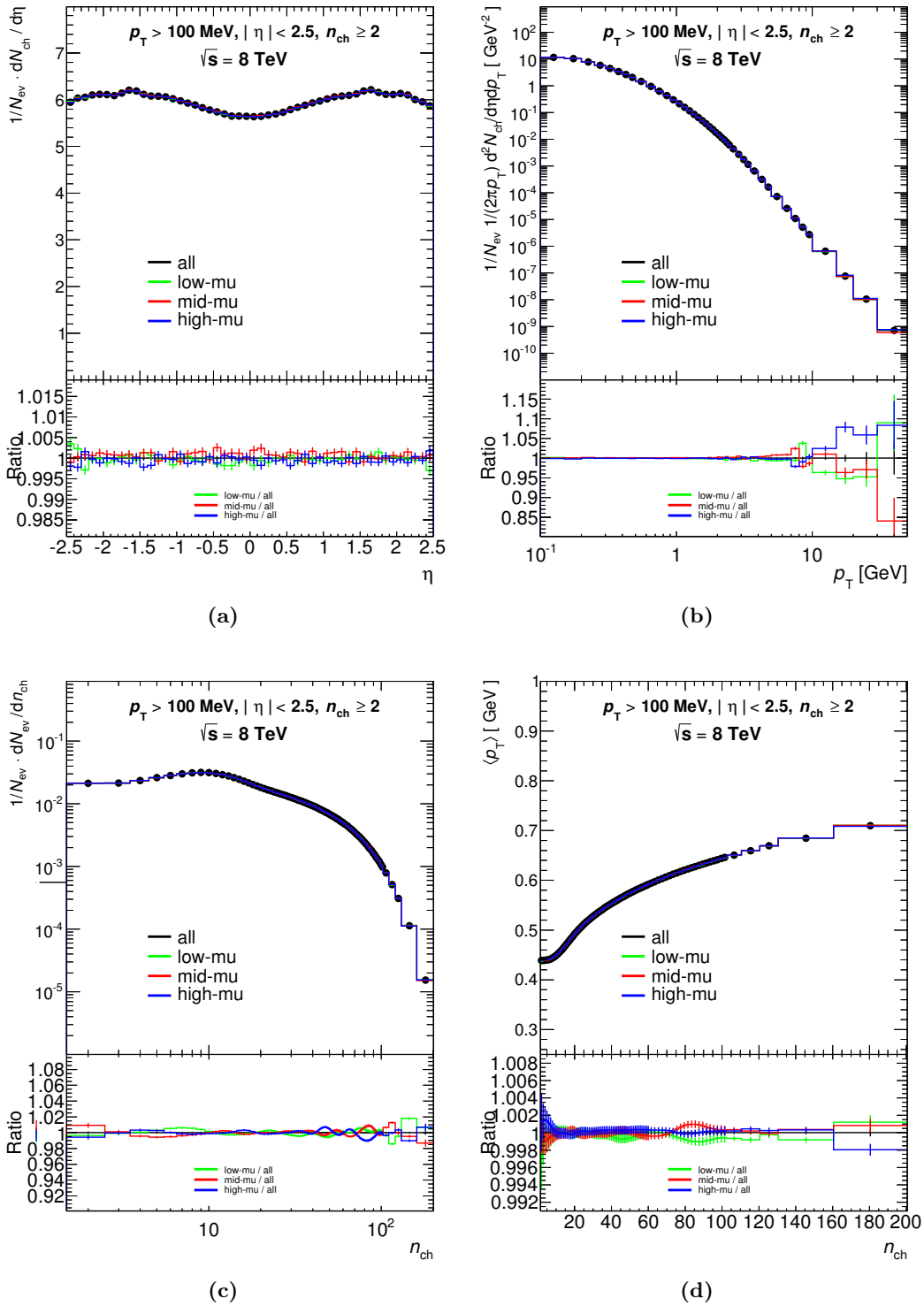


Figure A.15: Raw detector-level distributions of reconstructed tracks as a function of (a,d,g) η, (b,e,h) p_{T} and (c,f,i) reconstructed track multiplicity per event n_{sel}, for Data and MC, in three μ-ranges of the data: (a-c) $\mu \leq 0.0034$, (d-f) $0.0034 < \mu \leq 0.0036$, (g-i) $0.0036 < \mu$.



(a) (b)

(c) (d)

Figure A.16: A comparison of fully corrected charged-particle distributions for events with $n_{\mathrm{ch}} \geq 2$ within the kinematic range $p_{\mathrm{T}} > 100$ MeV and $|\eta| < 2.5$ at $\sqrt{s} = 8$ TeV as a function of the pseudorapidity (a), the transverse momentum in log-log scale (b), the charged-particle multiplicity per event in log-log scale (c), and the average transverse momentum as a function of the number of charged particles in the event (d), comparing different subsets of the data (depending on the average $\langle\mu\rangle$ per lumiblock) and using the Pythia 8 A2 tune for the full correction and unfolding procedure.

tails of the p_{T} distribution are compatible with statistical uncertainties and are smaller than the total systematic uncertainties of the final distribution. The fluctuations in both n_{ch}-dependent distributions are due to statistical effects of the Bayesian unfolding method and remain negligible in comparison with other systematic uncertainties due to the unfolding procedure.

μ Range	Events	$1/N_{\mathrm{ev}} \cdot \mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta$ at $\lvert\eta\rvert < 0.2$
all	9 166 134	5.644 ± 0.003 (stat) ± 0.103 (syst)
$0.0028 < \mu \leq 0.0034$	2 692 413	5.639 ± 0.005 (stat) ± 0.103 (syst)
$0.0034 < \mu \leq 0.0036$	3 082 858	5.652 ± 0.005 (stat) ± 0.103 (syst)
$0.0036 < \mu \leq 0.0040$	3 390 863	5.641 ± 0.005 (stat) ± 0.103 (syst)

Table A.1: Comparison of $1/N_{\mathrm{ev}} \cdot \mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta$ at $|\eta| < 0.2$ in the most inclusive phase space for the full dataset considered in the charged particle density measurement at $\sqrt{s} = 8$ TeV (luminosity blocks $215 \leq \mathrm{lbn} \leq 395$) and subsets of the full dataset based on the average number of interactions per bunch crossing μ.

Table A.1 shows the central charged-particle density per event in the most inclusive phase space, using different subsets of the data. The results are found to be stable at the level of statistical uncertainties and remain well below the total systematic uncertainties.

A.6.3 Comparison of Results obtained with different MC Tunes

A.6.3.1 Detector-level Distributions for Pythia 6 AMBT2B and Epos LHC Tunes

Using only the z-vertex correction weights, the raw η, p_{T} and n_{sel} distributions are obtained as shown in Figure A.17 for three different MC tunes in the $p_{\mathrm{T}} > 100$ MeV, $n_{\mathrm{sel}} \geq 2$ phase space. Except for the η distribution modelled by Pythia 8 A2, none of the models describe the data very well at the detector level. However it can be seen that in this phase space Epos LHC gives a much better description of the data than both Pythia tunes in terms of p_{T} and n_{sel} distributions.

A.6.3.2 Closure Tests for Pythia 6 AMBT2B and Epos LHC Tunes

Closure tests with alternative MC generator models are shown in the $p_{\mathrm{T}} > 100$ MeV, $n_{\mathrm{ch}} \geq 2$ phase space for the Pythia 6 AMBT2B tune in Figure A.18 and for the Epos LHC tune in Figure A.19.

each of these subsets from MC after this dedicated reweighting, leading to smaller fluctuations between the final η distributions, while leaving the other final distributions almost unchanged. However, not applying these dedicated corrections was found to be still sufficient for the purpose of the investigation presented here.

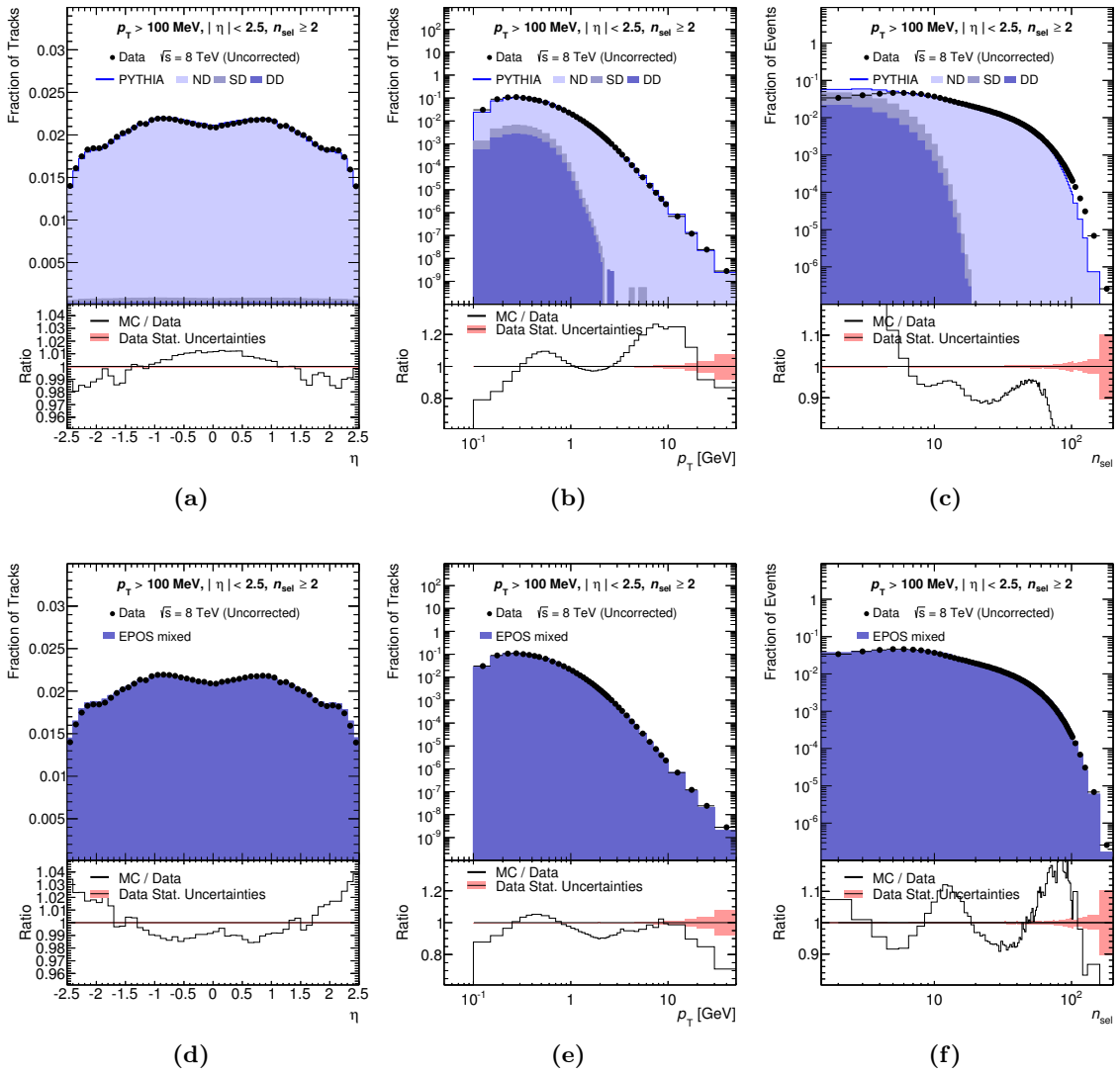
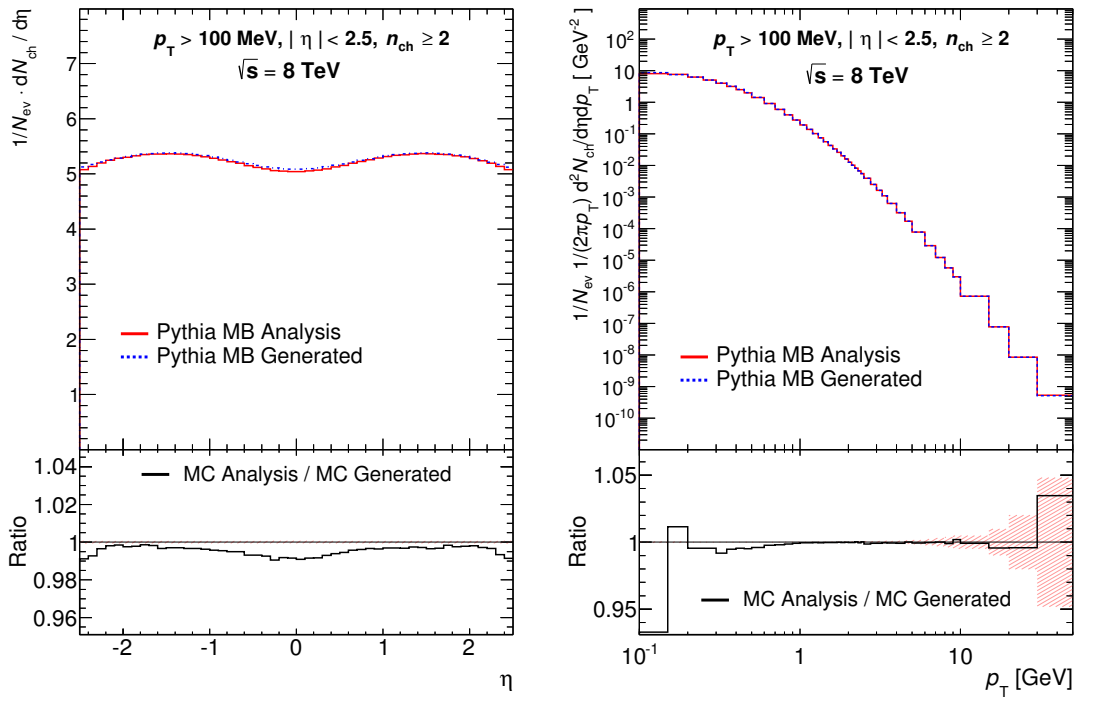
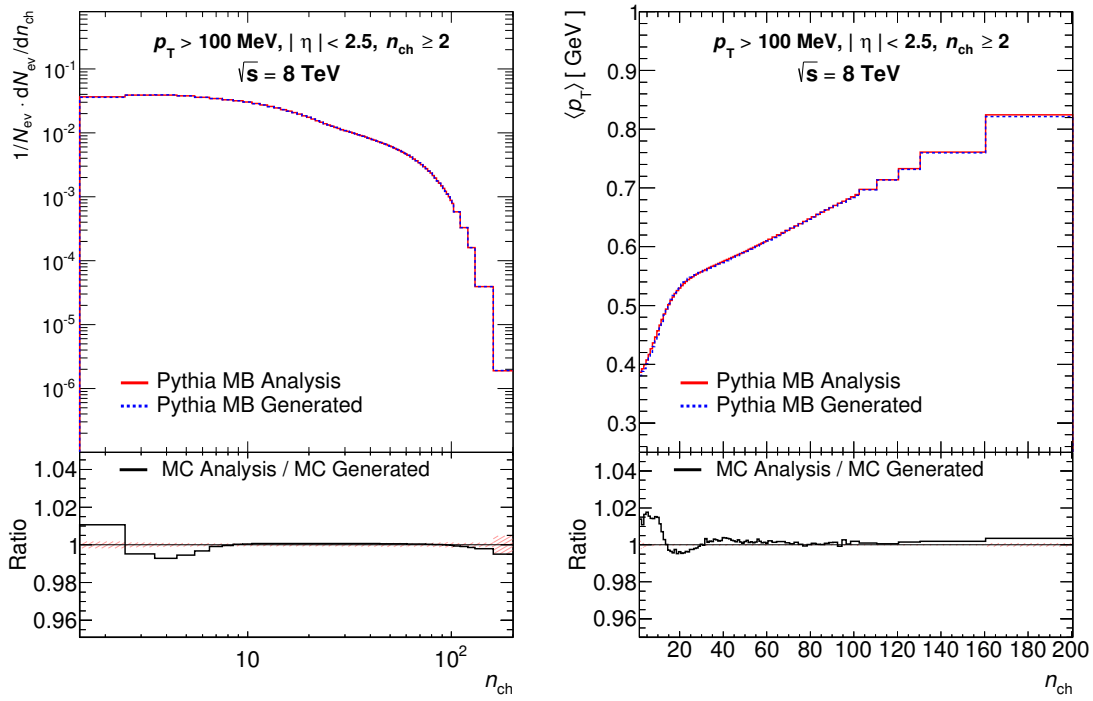


Figure A.17: Raw detector-level distributions of reconstructed tracks (events) in the $p_T > 100$ MeV, $n_{sel} \geq 2$ phase space as a function of (a) η, (b) p_T and (c) n_{sel} for data and PYTHIA 6 AMBT2B MC; (d) η, (e) p_T and (f) n_{sel} for data and EPOS LHC MC. All distributions are normalised to unity.



(a) (b)



(c) (d)

Figure A.18: Comparison of generated and fully corrected primary-charged-particle distributions for simulated events (using the PYTHIA 6 AMBT2B tune) with $n_{\mathrm{ch}} \geq 2$ within the kinematic range $p_{\mathrm{T}} > 100$ MeV and $|\eta| < 2.5$ as a function of (a) pseudorapidity, η, (b) transverse momentum, p_{T}, (c) the charged-particle multiplicity per event, n_{ch}, and (d) average transverse momentum, $\langle p_{\mathrm{T}} \rangle$, versus multiplicity.

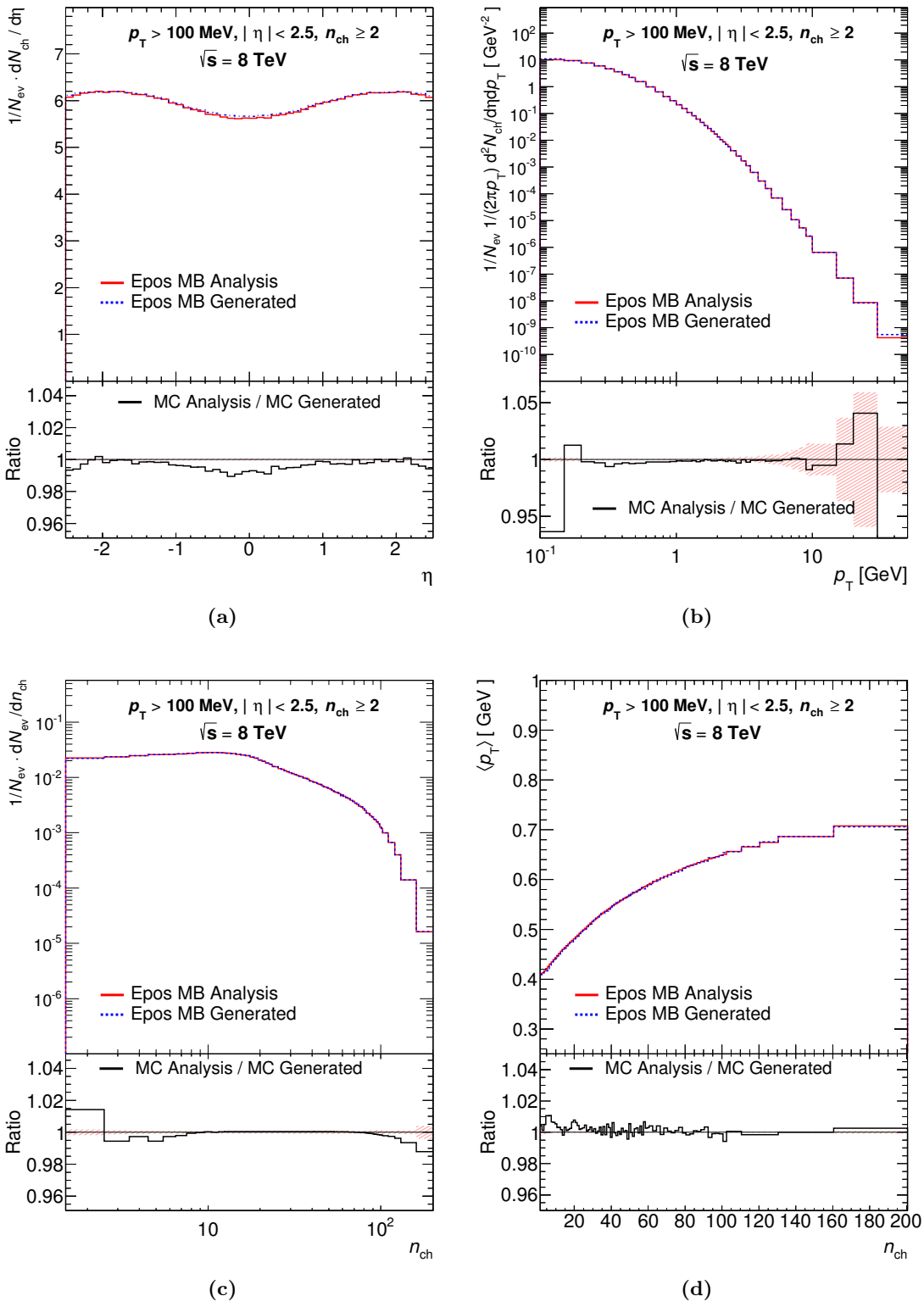


Figure A.19: Comparison of generated and fully corrected primary-charged-particle distributions for simulated events (using the Epos LHC tune) with $n_{\mathrm{ch}} \geq 2$ within the kinematic range $p_{\mathrm{T}} > 100$ MeV and $|\eta| < 2.5$ as a function of (a) pseudorapidity, η, (b) transverse momentum, p_{T}, (c) the charged-particle multiplicity per event, n_{ch}, and (d) average transverse momentum, $\langle p_{\mathrm{T}} \rangle$, versus multiplicity.

A.6.3.3 Comparison of Results for Pythia 8 A2, Pythia 6 AMBT2B and Epos LHC Tunes

Phase Space	MC tune	$1/N_{\mathrm{ev}} \cdot \mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta$ at $\lvert\eta\rvert < 0.2$
$p_{\mathrm{T}} > 100$ MeV, $n_{\mathrm{ch}} \geq 2$	Pythia 8 A2	5.644 ± 0.003 (stat) ± 0.103 (syst)
	Pythia 6 AMBT2B	5.651 ± 0.003 (stat) ± 0.104 (syst)
	Epos LHC	5.647 ± 0.003 (stat) ± 0.103 (syst)
$p_{\mathrm{T}} > 500$ MeV, $n_{\mathrm{ch}} \geq 1$	Pythia 8 A2	2.477 ± 0.001 (stat) ± 0.031 (syst)
	Pythia 6 AMBT2B	2.476 ± 0.001 (stat) ± 0.031 (syst)
	Epos LHC	2.476 ± 0.001 (stat) ± 0.031 (syst)

Table A.2: Comparison of $\mathrm{d}n_{\mathrm{ch}}/\mathrm{d}\eta$ at $|\eta| < 0.2$ for the full dataset considered in the charged particle density measurement at $\sqrt{s} = 8$ TeV, corrected and unfolded with different MC tunes.

A comparison of central charged particle densities $1/N_{\mathrm{ev}} \cdot \mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta$ obtained using corrections derived from the Pythia 8 A2, Pythia 6 AMBT2B and Epos LHC MC tunes, respectively, is presented in Table A.2 for the main phase spaces considered in this analysis.

The differences within each phase space are found to be at the level of statistical uncertainties and remain well below the total systematic uncertainties.

A.7 Variations of final plots

Figures A.20a and A.20b show linear and logarithmic variations of Figure 7.43b.

Figures A.20c and A.20d show linear and logarithmic variations of Figure 7.43c.

Figures A.21a and A.21b show linear and logarithmic variations of Figure 7.44b.

Figures A.21c and A.21d show linear and logarithmic variations of Figure 7.44c.

Figures A.22a and A.22b show linear and logarithmic variations of Figure 7.45b.

Figures A.23a and A.23b show linear and logarithmic variations of Figure 7.46b.

Figures A.24a and A.24b show linear and logarithmic variations of Figure 7.47b.

A.8 Final plots with alternative MC tunes

The presented measurements of primary charged particles were compared to additional phenomenological models describing minimum-bias events. For this purpose, additional particle-level MC samples were generated with the following models that are based on the CTEQ6 L1 PDF-set, with a sample size of 100M events each:

- Pythia 6 Innsbruck2013
- Herwig++ UE-EE-5

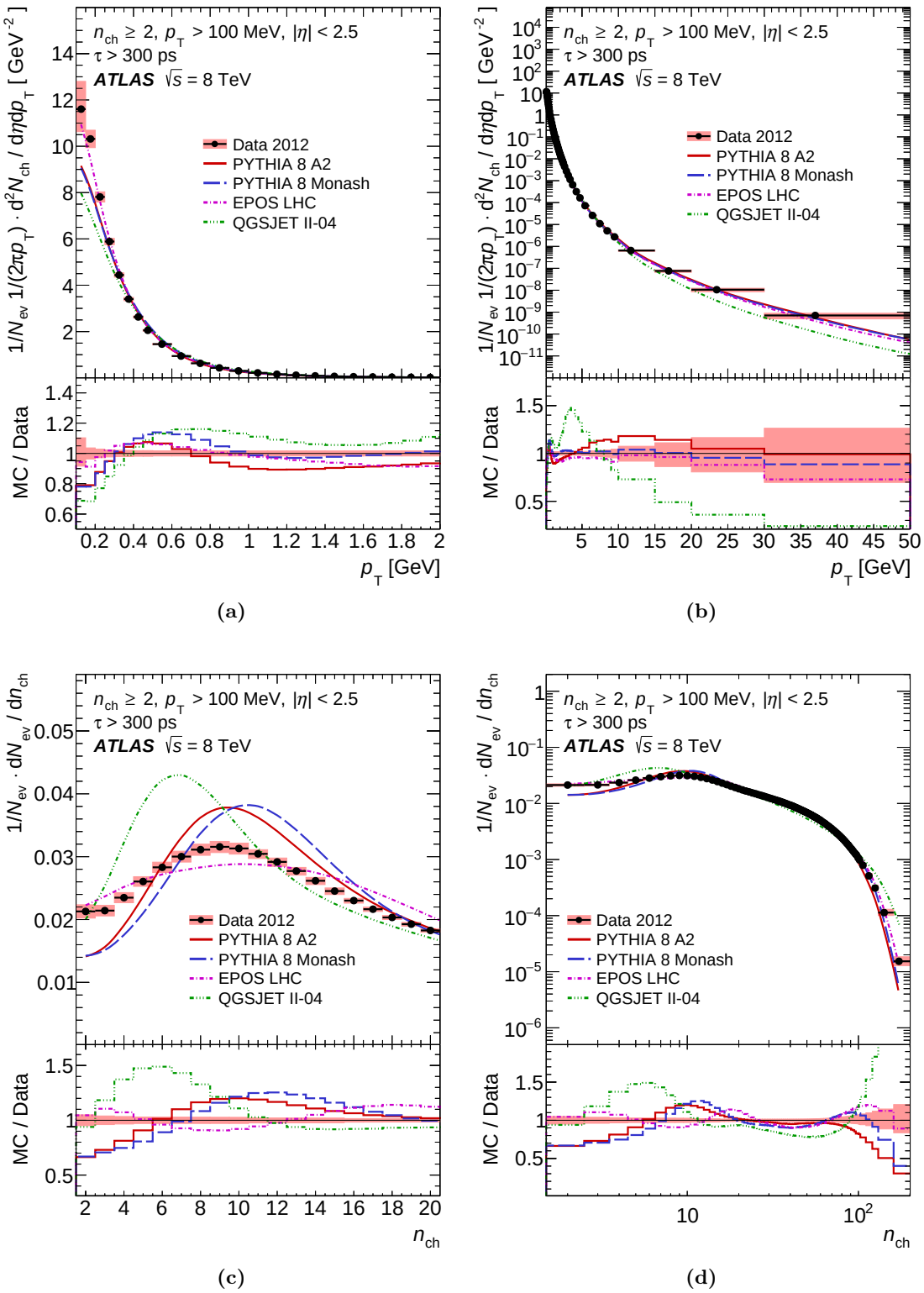


Figure A.20: Alternative distributions of primary charged particles in events for which $n_{ch} \geq 2$, $p_T > 100$ MeV and $|\eta| < 2.5$ as a function of (a-b) transverse momentum, p_T, and (c-d) multiplicity, n_{ch}. The data (markers) are compared to several particle-level predictions made by MC generator models (curves). The total uncertainty on the data is represented by shaded areas. Plots taken from Ref. [3].

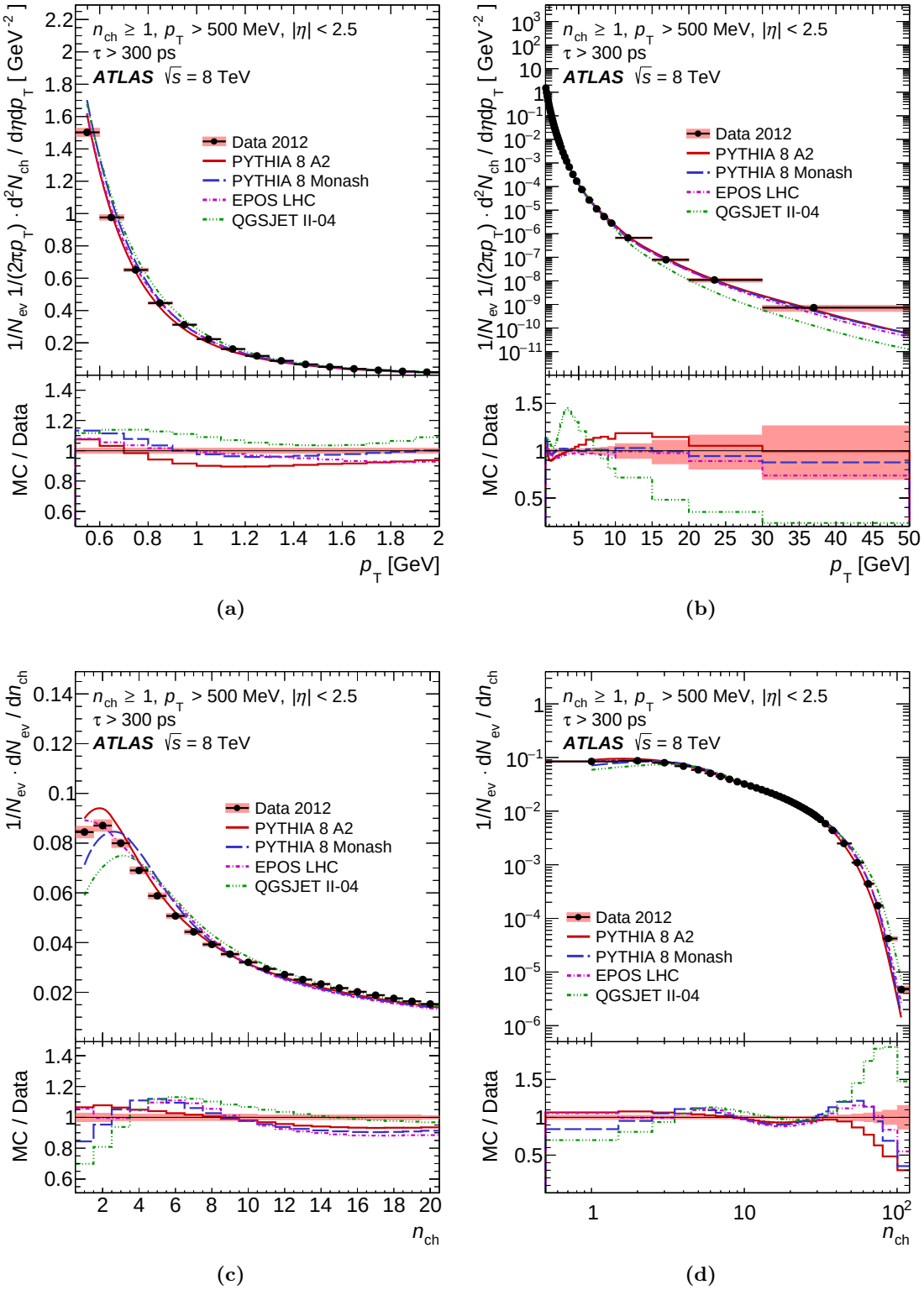
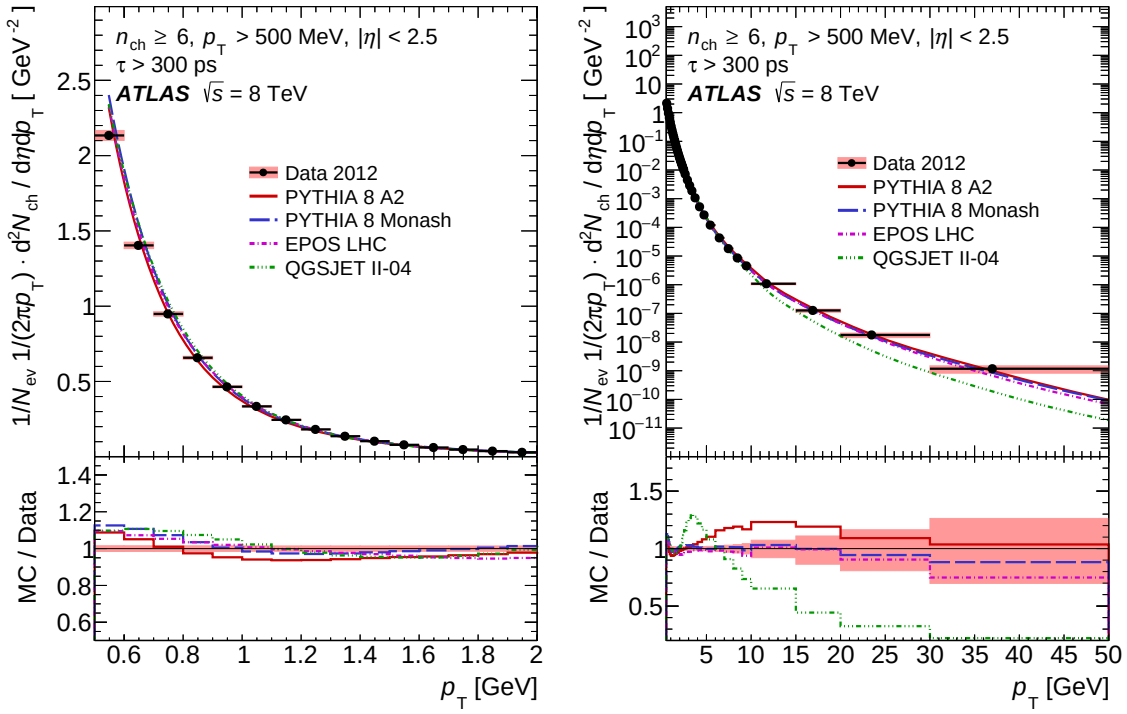


Figure A.21: Alternative distributions of primary charged particles in events for which $n_{\mathrm{ch}} \geq 1$, $p_{\mathrm{T}} > 500$ MeV and $|\eta| < 2.5$ as a function of (a-b) transverse momentum, p_{T}, and (c-d) multiplicity, n_{ch}. The data (markers) are compared to several particle-level predictions made by MC generator models (curves). The total uncertainty on the data is represented by shaded areas. Plots taken from Ref. [3].



(a) (b)

Figure A.22: Alternative distributions of primary charged particles in events for which $n_{\mathrm{ch}} \geq 6$, $p_{\mathrm{T}} > 500$ MeV and $|\eta| < 2.5$ as a function of the transverse momentum, p_{T}. The data (markers) are compared to several particle-level predictions made by MC generator models (curves). The total uncertainty on the data is represented by shaded areas. Plots taken from Ref. [3].

- Herwig++ UE-EE-4

Distributions of stable charged primary particles in 8 TeV data compared with predictions from these additional models are shown in Figures A.25 to A.29.

A.8.1 Pseudorapidity Distribution

Figures A.25a, A.26a, A.27a, A.28a and A.29a show distributions of the primary-charged-particle yield as a function of pseudorapidity, $1/N_{\mathrm{ev}} \cdot \mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta$. The distribution measured in the $n_{\mathrm{ch}} \geq 2$ and $p_{\mathrm{T}} > 100$ MeV phase space is poorly described by all additional tunes. In the $p_{\mathrm{T}} > 500$ MeV phase spaces, the Pythia 6 Innsbruck2013 tune describes the data well for all multiplicity thresholds, $n_{\mathrm{ch}} \geq 1, 6, 20, 50$, while the shape of differences between the prediction and the data is inverted with respect to the Pythia 8 A2 and Epos LHC tunes.

A.8.2 Transverse-Momentum Distribution

Figures A.25b, A.26b, A.27b, A.28b and A.29b show normalised distributions of the primary-charged-particle yield as a function of transverse momentum, $1/(2\pi p_{\mathrm{T}} N_{\mathrm{ev}}) \cdot \mathrm{d}^2 N_{\mathrm{ch}}/(\mathrm{d}\eta\, \mathrm{d}p_{\mathrm{T}})$. All

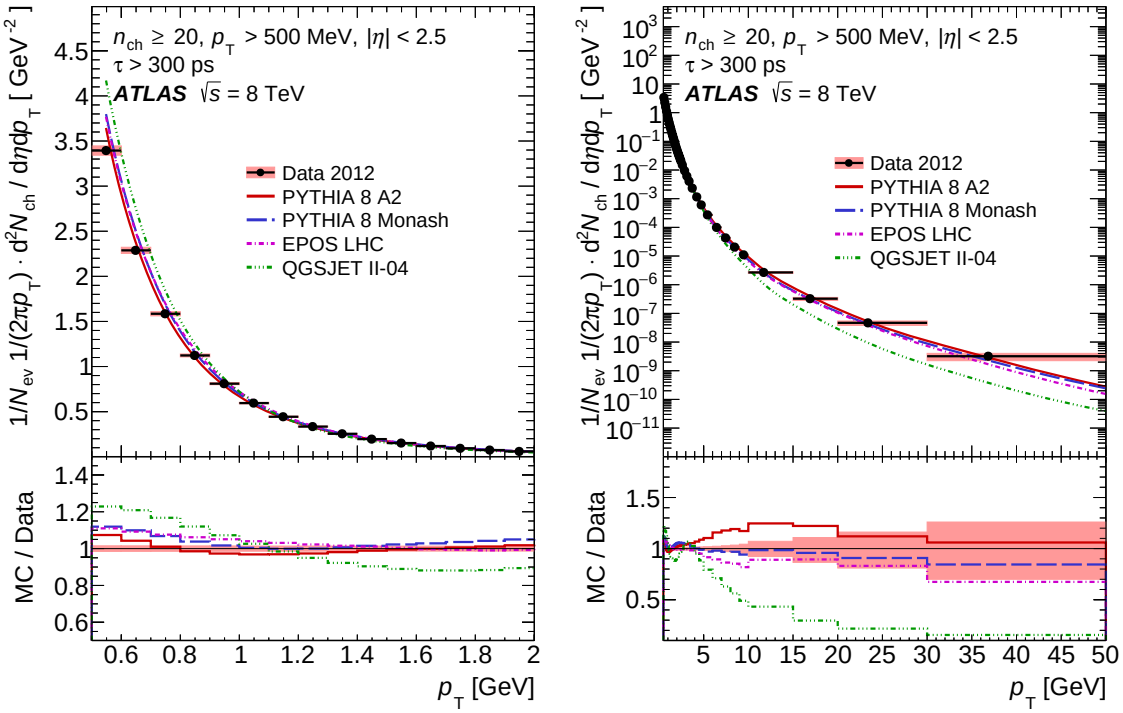
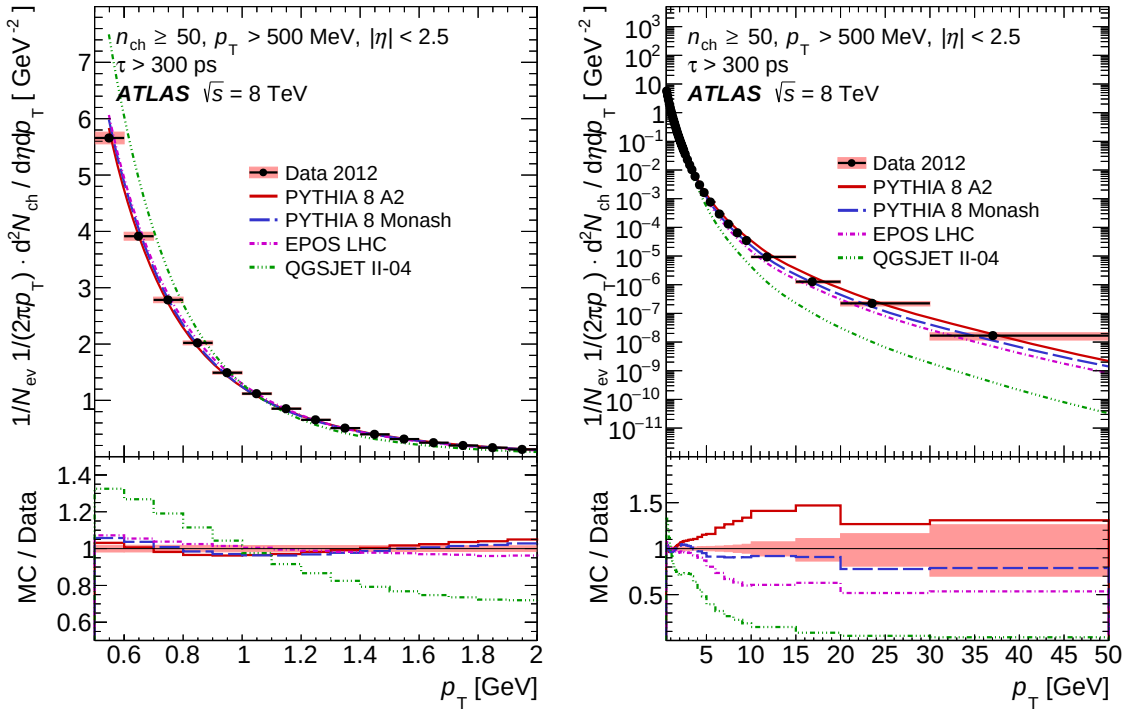


Figure A.23: Alternative distributions of primary charged particles in events for which $n_{\text{ch}} \geq 20$, $p_{\text{T}} > 500$ MeV and $|\eta| < 2.5$ as a function of the transverse momentum, p_{T}. The data (markers) are compared to several particle-level predictions made by MC generator models (curves). The total uncertainty on the data is represented by shaded areas. Plots taken from Ref. [3].

additional tunes fail to describe the distribution corresponding to the most inclusive selection, $n_{\text{ch}} \geq 2$, $p_{\text{T}} > 100$ MeV. The Pythia 6 Innsbruck2013 tune gives an excellent description of the data for the $p_{\text{T}} > 500$ MeV selections with multiplicity thresholds $n_{\text{ch}} \geq 1, 6$, and makes a very good prediction of the distributions with $n_{\text{ch}} \geq 20, 50$.

A.8.3 Charged-Particle Multiplicity Distribution

The multiplicity distributions of primary charged particles, $1/N_{\text{ev}} \cdot \mathrm{d}N_{\text{ev}}/\mathrm{d}n_{\text{ch}}$, are presented in Figures A.25c and A.26c for the two main phase spaces considered in this analysis: (1) $n_{\text{ch}} \geq 2$, $p_{\text{T}} > 100$ MeV and (2) $n_{\text{ch}} \geq 1$, $p_{\text{T}} > 500$ MeV. None of the additional models are consistent with the data in the most inclusive selection at $p_{\text{T}} > 100$ MeV. The Pythia 6 Innsbruck2013 tune provides a fair description of the data in the $p_{\text{T}} > 500$ MeV phase space between $5 < n_{\text{ch}} < 50$.



(a) (b)

Figure A.24: Alternative distributions of primary charged particles in events for which $n_{\text{ch}} \geq 50$, $p_{\text{T}} > 500$ MeV and $|\eta| < 2.5$ as a function of the transverse momentum, p_{T}. The data (markers) are compared to several particle-level predictions made by MC generator models (curves). The total uncertainty on the data is represented by shaded areas. Plots taken from Ref. [3].

A.8.4 Average Transverse Momentum as a Function of n_{ch}

Final results for the $\langle p_{\text{T}} \rangle$ versus n_{ch} distribution are shown in Figures A.25d and A.26d for the two main phase spaces. None of the additional models are consistent with the data in the most inclusive selection at $p_{\text{T}} > 100$ MeV. The Pythia 6 Innsbruck2013 tune and the Herwig++ UE-EE-4 tune provide a fair description of the data in the $p_{\text{T}} > 500$ MeV phase space for multiplicities below $n_{\text{ch}} < 80$.

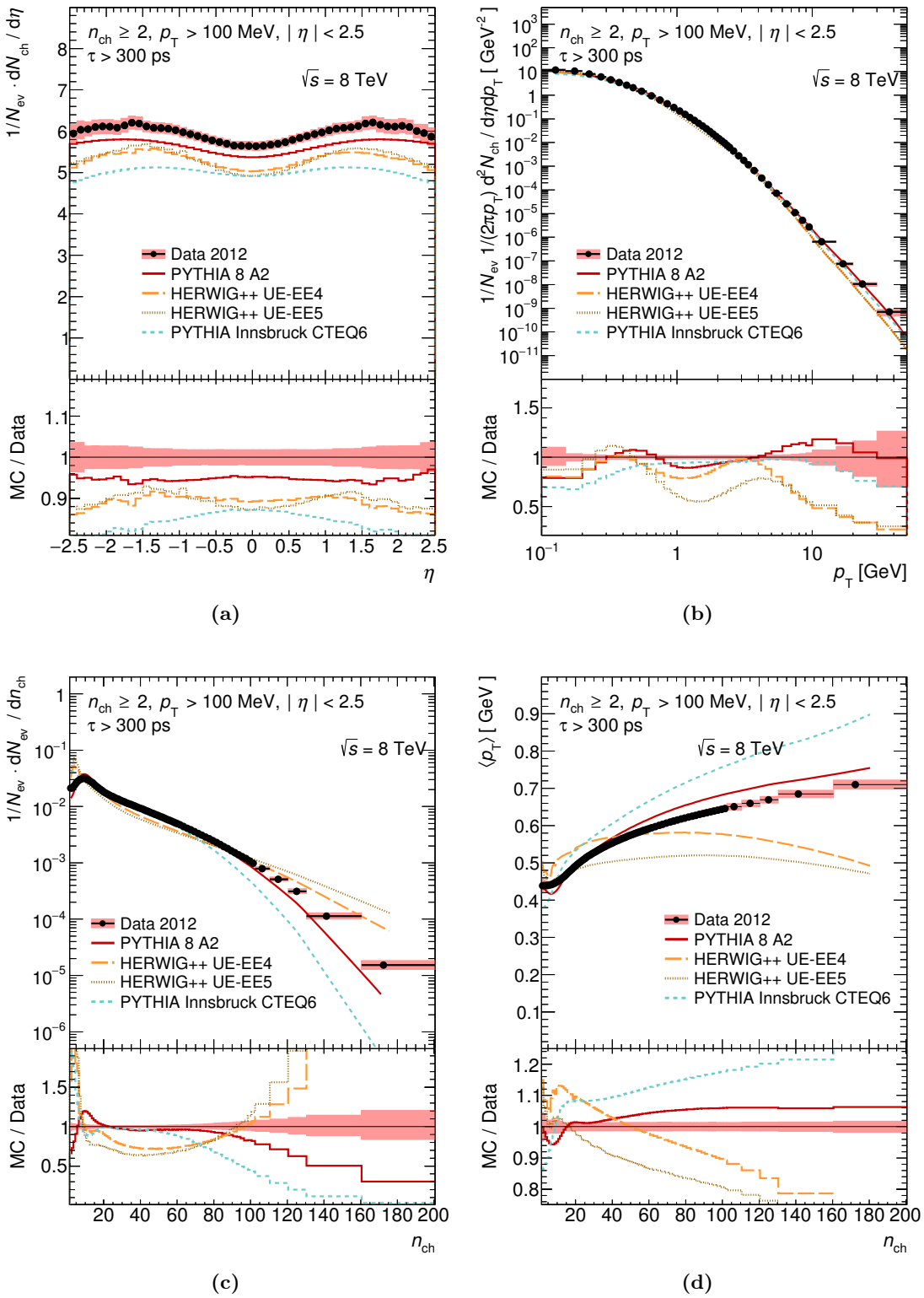


Figure A.25: Distributions of primary charged particles in events for which $n_{\mathrm{ch}} \geq 2$, $p_{\mathrm{T}} > 100$ MeV and $|\eta| < 2.5$ as a function of (a) pseudorapidity, η, (b) transverse momentum, p_{T}, (c) multiplicity, n_{ch}, and (d) average transverse momentum, $\langle p_{\mathrm{T}} \rangle$, versus multiplicity. The data (markers) are compared to several particle-level predictions made by MC generator models (curves). The total uncertainty on the data is represented by shaded areas.

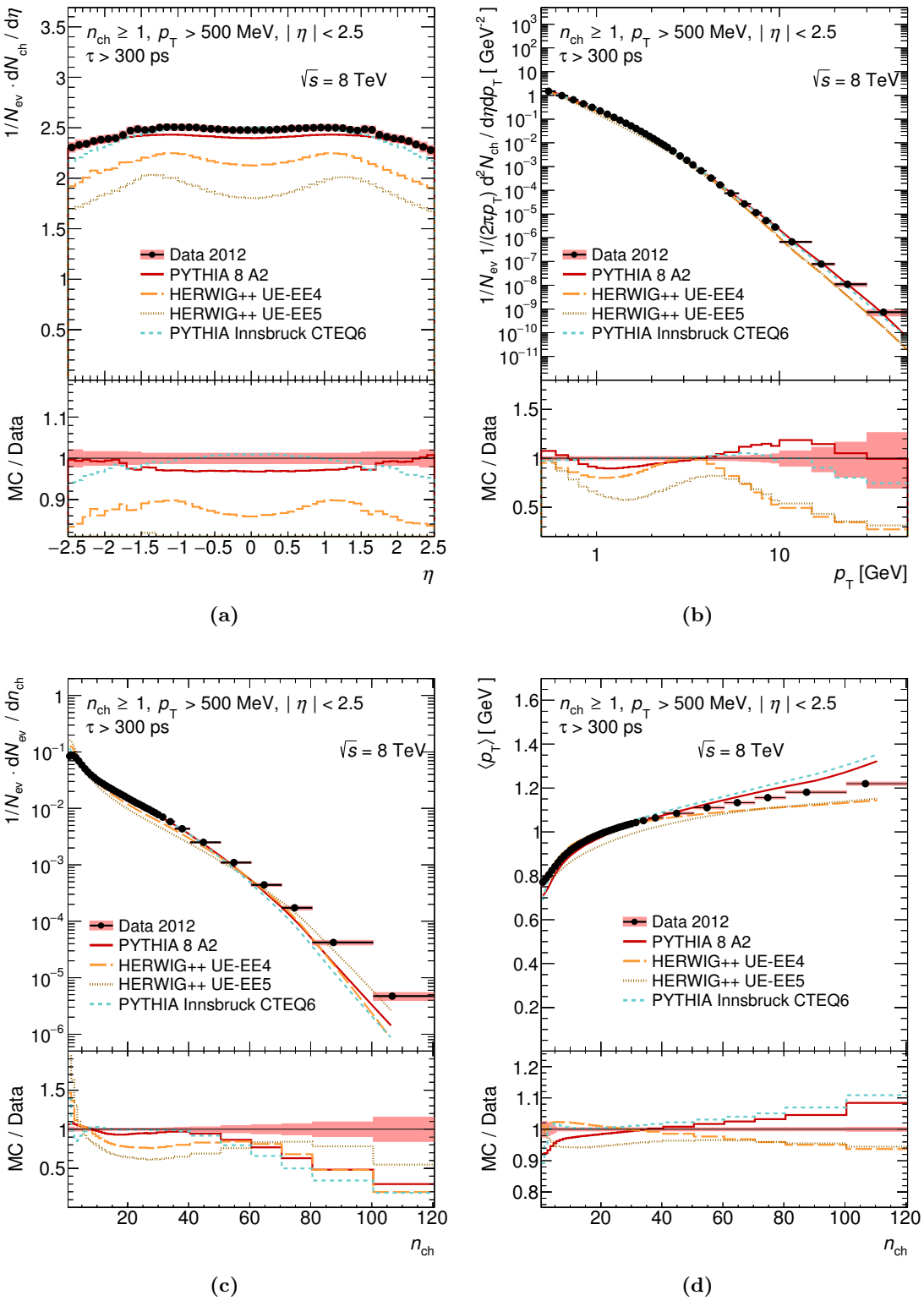
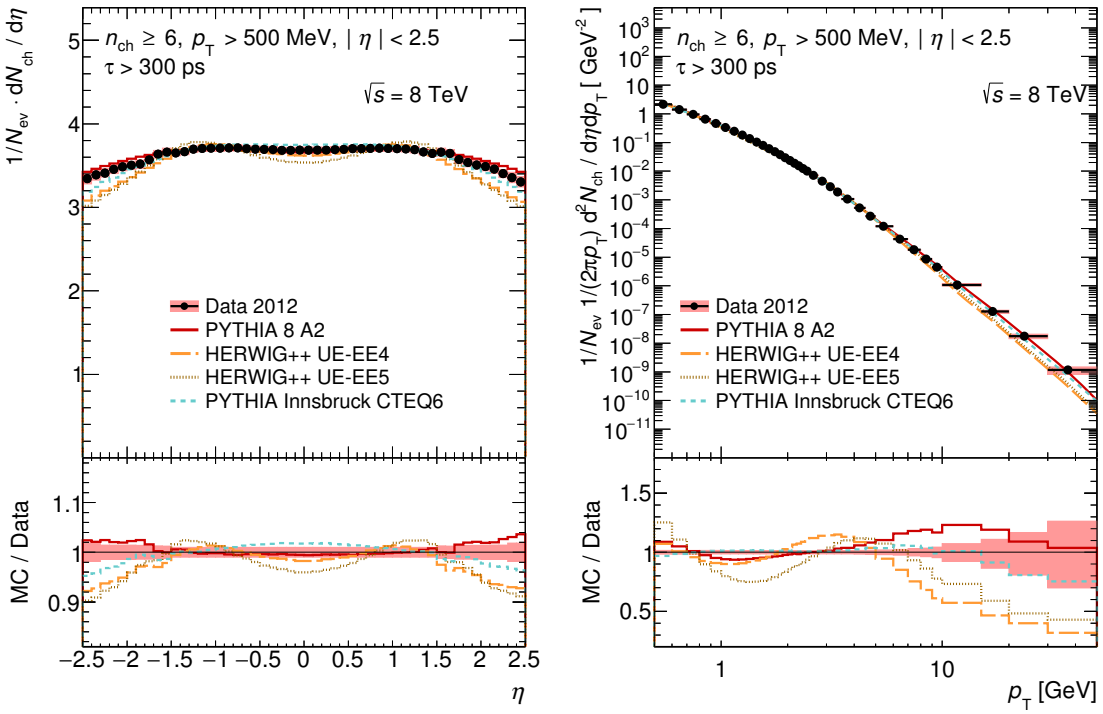
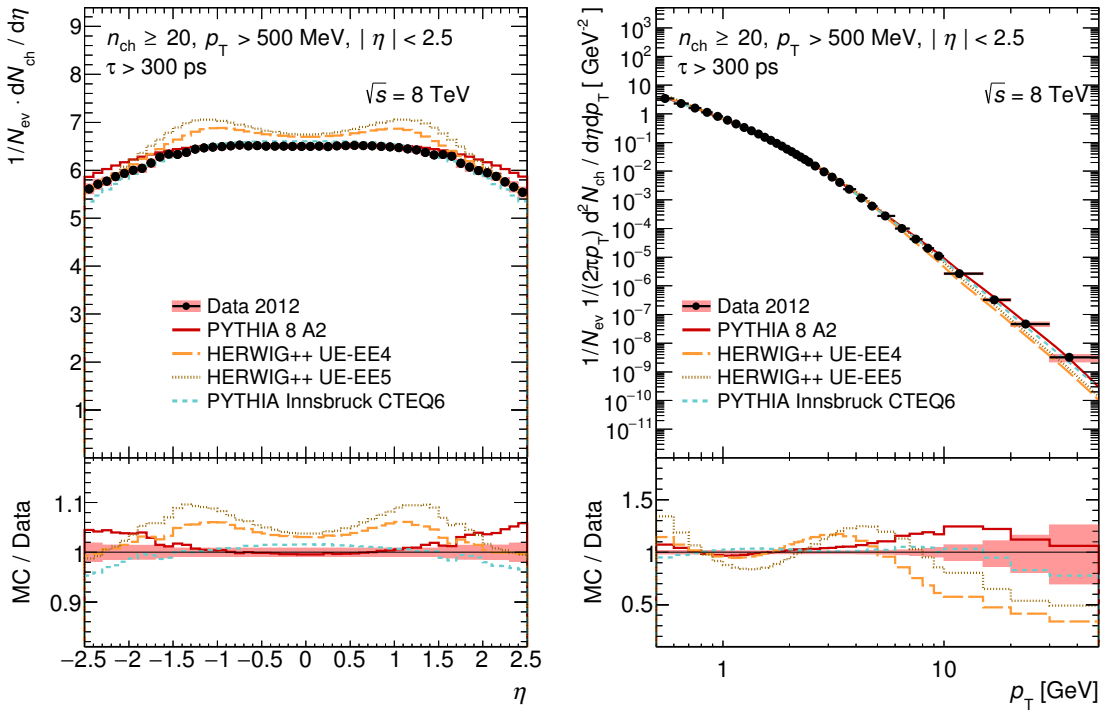


Figure A.26: Distributions of primary charged particles in events for which $n_{\mathrm{ch}} \geq 1$, $p_{\mathrm{T}} > 500$ MeV and $|\eta| < 2.5$ as a function of (a) pseudorapidity, η, (b) transverse momentum, p_{T}, (c) multiplicity, n_{ch}, and (d) average transverse momentum, $\langle p_{\mathrm{T}} \rangle$, versus multiplicity. The data (markers) are compared to several particle-level predictions made by MC generator models (curves). The total uncertainty on the data is represented by shaded areas.



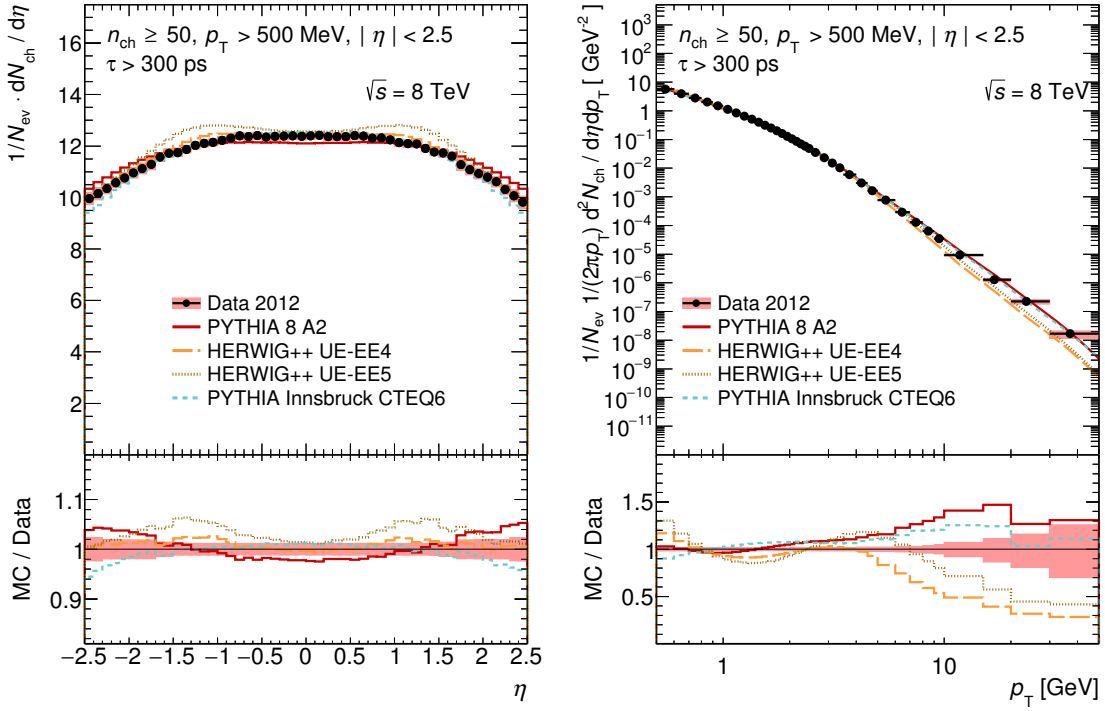
(a) **(b)**

Figure A.27: Distributions of primary charged particles in events for which $n_{ch} \geq 6$, $p_T > 500$ MeV and $|\eta| < 2.5$ as a function of (a) pseudorapidity, η, and (b) transverse momentum, p_T. The data (markers) are compared to several particle-level predictions made by MC generator models (curves). The total uncertainty on the data is represented by shaded areas.



(a) **(b)**

Figure A.28: Distributions of primary charged particles in events for which $n_{ch} \geq 20$, $p_T > 500$ MeV and $|\eta| < 2.5$ as a function of (a) pseudorapidity, η, and (b) transverse momentum, p_T. The data (markers) are compared to several particle-level predictions made by MC generator models (curves). The total uncertainty on the data is represented by shaded areas.



(a) **(b)**

Figure A.29: Distributions of primary charged particles in events for which $n_{ch} \geq 50$, $p_T > 500$ MeV and $|\eta| < 2.5$ as a function of (a) pseudorapidity, η, and (b) transverse momentum, p_T. The data (markers) are compared to several particle-level predictions made by MC generator models (curves). The total uncertainty on the data is represented by shaded areas.

Appendix B

Technical details of the Detector Geometry Validation

B.1 Implementation and Validation of the Tracking Geometry

B.1.1 Creation of Material Maps

In order to describe the material distributions from various ID subcomponents in the Tracking Geometry, the material traversed along the most typical particle trajectories must be associated with the nearby material layers and stored in binned material maps. Here the steps performed to produce these material maps are outlined.

B.1.1.1 Step 1: Geantino Mapping

The association of ID material with Tracking Geometry layers is done by mapping the material described in the full simulation geometry onto the simplified Tracking Geometry. This is done by propagating a non-interacting virtual particle called *Geantino* along a straight line from the nominal interaction point outwards through the simulation geometry. A detailed map of the simulation geometry is used as an input.[1]

For each one of a sufficiently large number of simulated Geantinos, the spatial information (position, step length) and the material properties (radiation length X_0, nuclear interaction length L_0, density ρ, atomic number Z and mass number A) of the GEANT4 simulation steps are stored in an output file.[2]

B.1.1.2 Step 2: Material Mapping

The output file obtained in this first step is subsequently used as an input for the Material Mapping step,[3] which:

1. takes the information of every single recorded Geantino step,

[1] This detailed map of the simulation geometry is produced by a dedicated GEANT4 UserAction called `GeantinoMapping`.
[2] The information is translated into a persistent class and written to a ROOT file by the `MaterialStepRecorder`.
[3] This procedure is implemented via the `MaterialMapping` algorithm from the `TrkDetDescrAlgs` package.

2. filters it (applying validation cuts),
3. associates it to the closest cylindric or disc-shaped material layer of the Tracking Geometry,
4. finds the intersection point of the Geantino on that associated layer, and
5. records the associated material hit on the intersected layer within its enclosing tracking volume.[4]

In this way the material maps are produced by associating and mapping one recorded GEANT4 step after another. If the association process fails for any of the following reasons,

- no associated layer can be found at all, or
- the associated layer cannot be found in the collection (map) of registered layers, or
- no valid intersection with the associated layer can be made,

the hit is recorded in an "unmapped hits" collection for each associated tracking volume (requiring that an enclosing tracking volume was found for the given position of the GEANT4 step).

During the finalisation of each event in the Material Mapping step, the recorded hit collections on each respective layer are filled into two-dimensional histograms. These histograms are using z/ϕ bins for cylinder layers and r/ϕ bins for disc layers. For a full coverage of the tracking acceptance region of the ID ($|\eta| < 2.7$) with sufficient statistics, O(10^6) simulated Geantinos are required.

B.1.1.3 Step 3: Material Validation

The material maps are then validated by a dedicated Material Validation job,[5] which extrapolates non-interacting particles (called *Fatrasinos*, the FATRAS equivalent to Geantinos) on straight tracks through the Tracking Geometry and again records the material seen by them, both on the individual layers and within each enclosing tracking volume. A comparison plot of the results from the mapping and validation job for the pixel detector is shown in Figure B.1, and for the SCT in Figure B.2. In these examples, very good agreement between the two geometry models is seen in the region $|\eta| < 2.5$.

If all the previous steps were successful, the new Tracking Geometry material file for the ATLAS geometry layout can be uploaded and registered to the ATLAS geometry database.

The `MaterialStepRecorder` class which is used for the production of material maps has been extended by the author to include the nuclear interaction length L_0, and (optionally) the recording of the energy loss and the multiple scattering angles when simulating interacting particles like muons for separate studies. A new `EnergyLossRecorder` class produces output files for direct comparisons of the integrated X_0, L_0 and energy loss between full simulation (GEANT4), fast simulation (FATRAS) and standard reconstruction, recorded when particles traverse specified "virtual" surfaces (see section B.2).[6]

[4] This step is implemented in the `MaterialMapper` tool from the `TrkDetDescrTools` package.
[5] Implemented via the `MaterialValidation` algorithm using the `MaterialMapper` tool.
[6] The `MaterialStepRecorder` and `EnergyLossRecorder` class reside in the `TrkG4UserActions` package

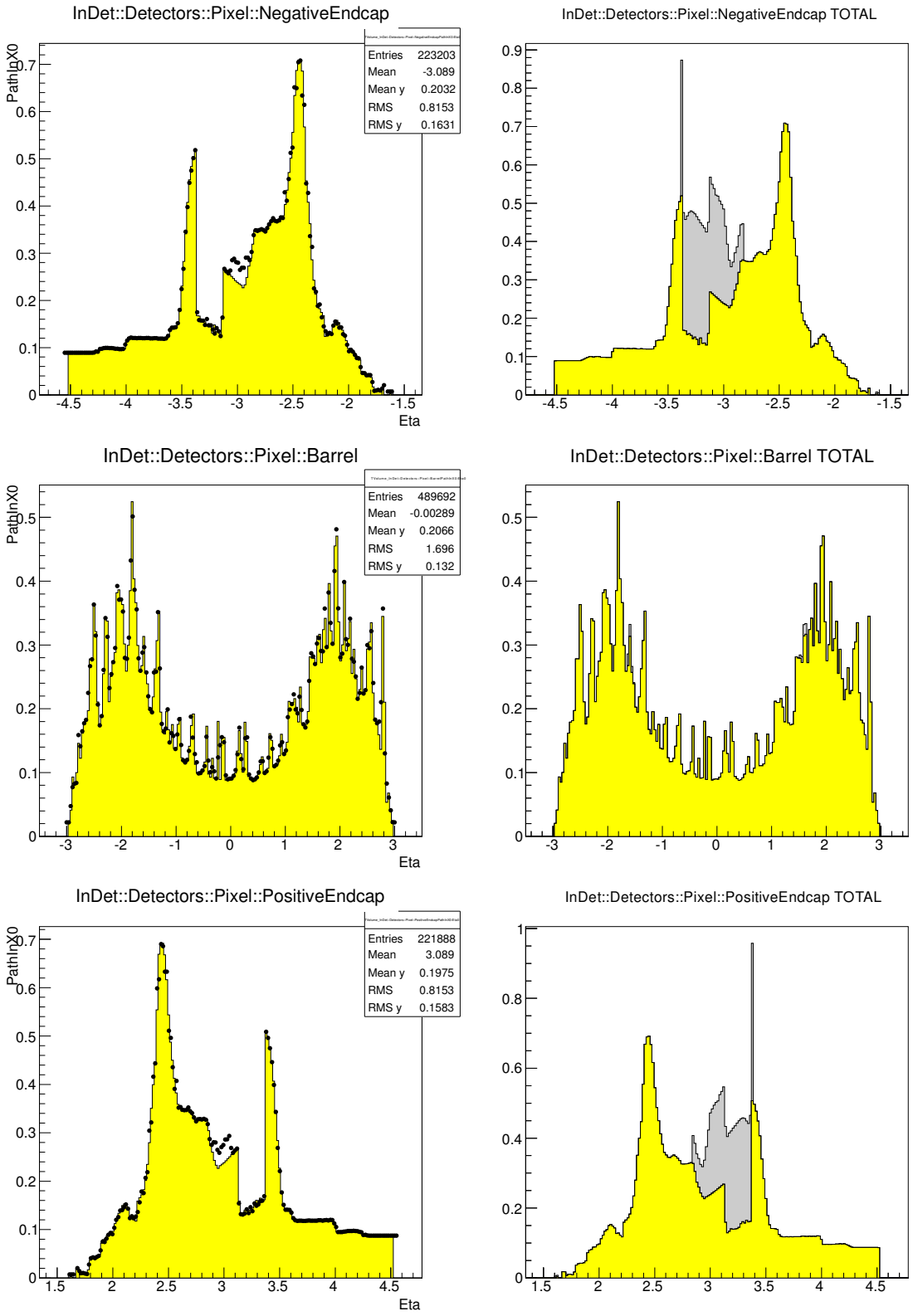


Figure B.1: Mapping validation results for the mapped (left) and total (right) hits in the Pixel barrel and end-caps for ATLAS-GEO-18-01-00. The simulation geometry is shown as a histogram (with yellow areas representing the mapped material and grey areas representing the unmapped material), the reconstruction geometry as black markers.

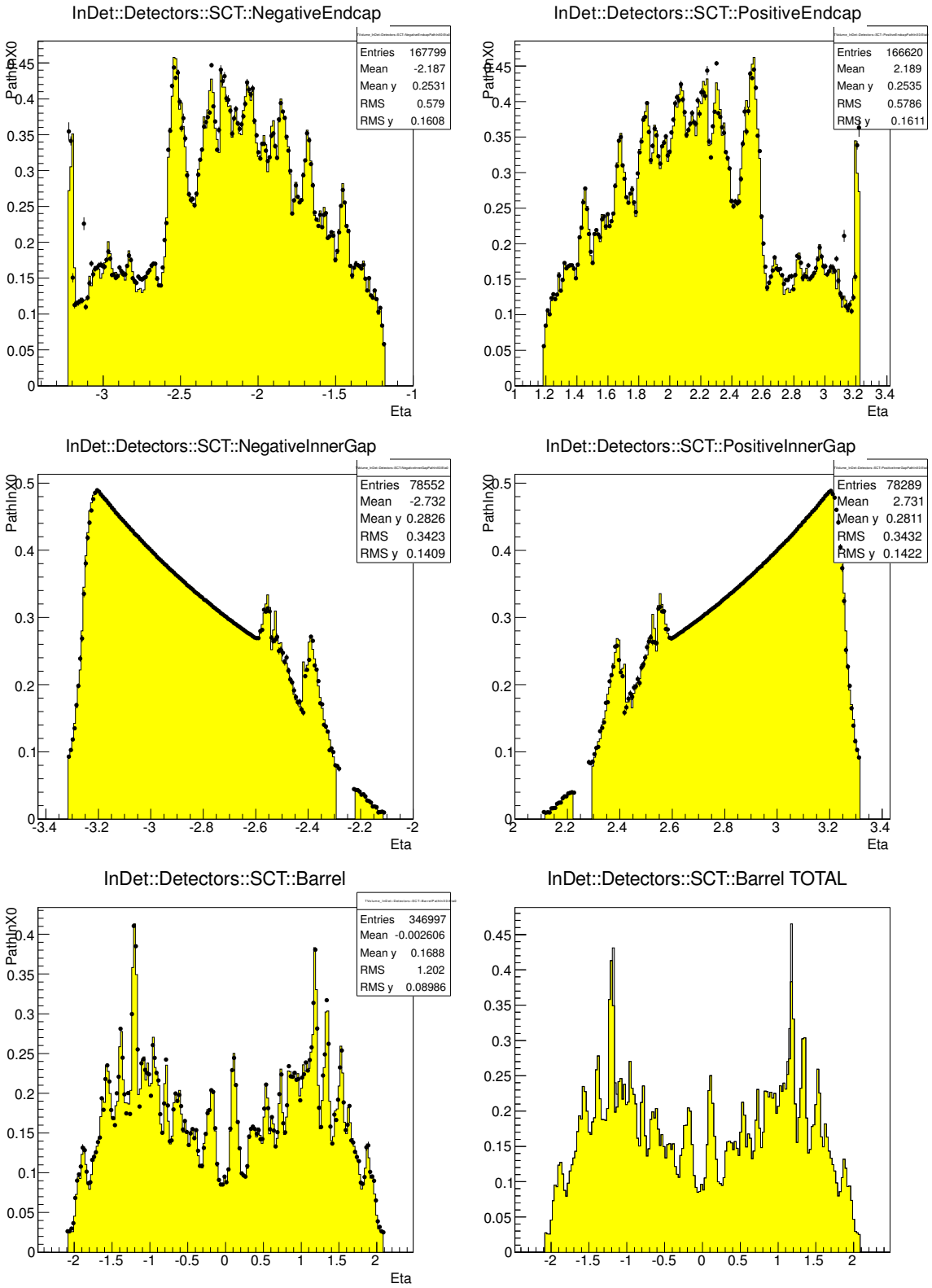


Figure B.2: Mapping validation results for the mapped and total hits in the SCT barrel and end-caps for ATLAS-GEO-18-01-00. The simulation geometry is shown as a histogram (with yellow areas representing the mapped material and grey areas representing the unmapped material), the reconstruction geometry as black markers.

B.2 Layer-based Validation Framework

A dedicated and flexible validation framework is essential in order to verify quickly that the Tracking Geometry material seen by the track extrapolation engine matches the material traversed by a particle in the full simulation. This comparison can and should be done not only at the border regions between the main ATLAS detector components - ID, Calorimeter and MS - but also for various regions within these components. A good agreement between the geometries in terms of material properties such as t/X_0, t/L_0 and energy loss indicates a successful material mapping process and a suitable Tracking Geometry description as well as proper material integration of the physics processes.

Beyond the validation of the material maps described in the previous section, a new validation framework has been implemented by the author. This technique is based on a comparison of X_0, L_0 and energy loss between full simulation (GEANT4), fast simulation (FATRAS) and standard reconstruction when particles traverse specified virtual surfaces. The easiest and most straightforward way to implement these virtual surfaces is via definitions of cylinder and disc surfaces, which can be placed arbitrarily between/around main ATLAS detector components (e.g. Beam Pipe, ID, Calorimeter, MS), detector subsystems (e.g. Pixel, SCT, TRT) or individual sensitive and passive regions within subsystems (e.g. pixel layers, PST). These virtual surfaces act as checkpoints for particles which are propagating outwards from the interaction point. As soon as each primary particle reaches or crosses the surface, several properties are written into an output file: the position (x, y, z and r), momentum (p), energy, η, ϕ, accumulated t/X_0 and t/L_0.

B.2.1 Layer-based Validation with GEANT4

In the case of GEANT4 simulation through the full detector geometry, the recording of this information is performed by the `EnergyLossRecorder`. In every step the path length in units of X_0 and L_0, t/X_0 and t/L_0, is accumulated (unless disabled in the jobOptions) and the post-step position is compared to the virtual surface dimensions specified in the jobOptions. As soon as a particle has crossed the boundaries of a virtual surface that has not yet been recorded, the pre-step parameters mentioned above are recorded (see Figure B.3), and at the end of the particle lifecycle this collection is written to the output file. Any secondary particles arising from e.g. hadronic interactions or bremsstrahlung are immediately removed from the simulation.

B.2.2 Layer-based Validation with Tracking Geometry

In the case of track extrapolation through the Tracking Geometry with either FATRAS (using more detailed physics processes, e.g. energy loss with Landau sampling) or the standard reconstruction tools, the recording is done by the `EnergyLossExtrapolationValidation` class.[7] Here the tracks are directly extrapolated in one step to the next virtual surface specified in the

[7] The `EnergyLossExtrapolationValidation` class resides in the `TrkExAlgs` package.

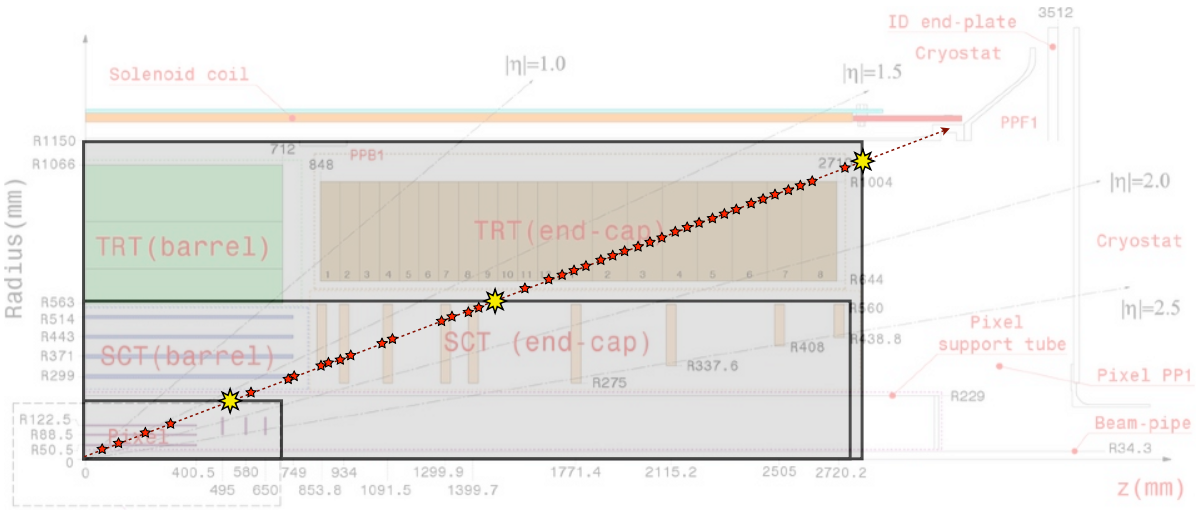


Figure B.3: Sketch illustrating the principle of the layer-based validation technique with GEANT4 and the full detector description. Whenever a specified surface is reached, the material parameters of the track, collected from all previous steps, are stored.

jobOptions (see Figure B.4), either with the current track parameters as start parameters (in the so-called "onion mode" which assumes a strictly hierarchical order of virtual surfaces) or again with the same initial start parameters at the nominal interaction point. If the extrapolation has been successful, the new track parameters are stored. After the final extrapolation the entire collection of track parameters is written to an output file. For interacting particles the physics processes can be individually enabled or disabled.

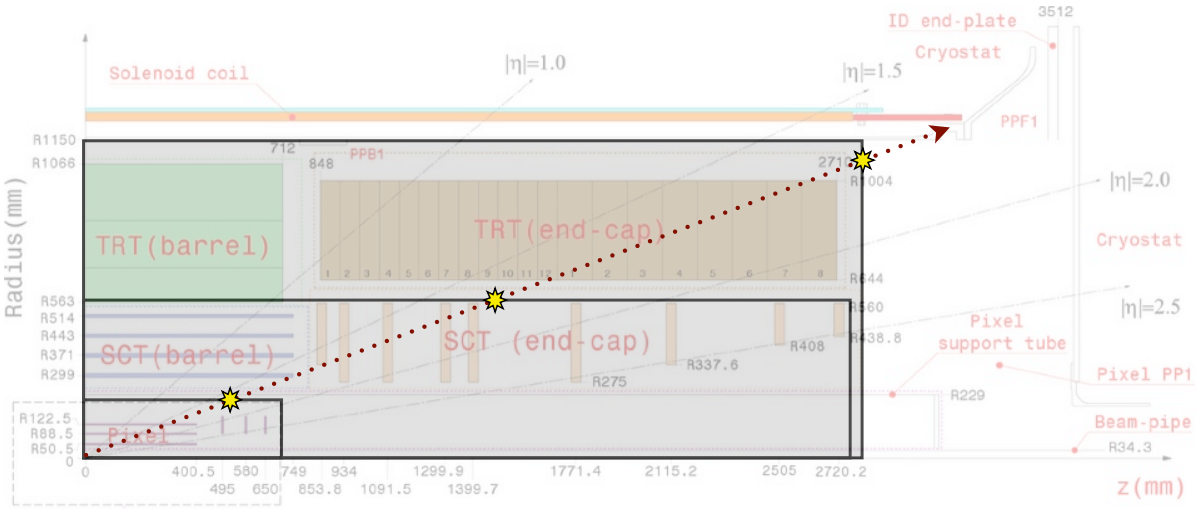


Figure B.4: Sketch illustrating the principle of the layer-based validation technique with the Tracking Geometry. The track is extrapolated from one specified surface to the next, and the updated track parameters are stored after every successful extrapolation step.

B.2.2.1 Comparison and Options

The starting parameters are in both cases determined by sampling η and ϕ from a uniform random distribution between the minimum and maximum values specified in the jobOptions. Either p or p_{T} can be set constant. The actual comparison is then performed by a dedicated ROOT script. This validation technique is not restricted to comparisons between different simulation flavours as described above. It can also be used for direct comparisons of detector descriptions of the same type, i.e. two different GEO models or Tracking Geometries with different material maps.

B.3 Validation of Material Budget in Different Detector Models

The differences in the material budget - the amount of material seen by a particle as it passes through the detector geometry - can be visualised for a twofold purpose: (a) a validation of the Tracking Geometry and the produced material maps with respect to the full ATLAS geometry, (b) a direct comparison between different versions of GEO models or TrackingGeometries.

B.3.1 Validation of Material Maps in Tracking Geometry

A precise material budget comparison on user-defined surfaces can be achieved with the cylinder-based validation framework. The user can define the cylinder dimensions arbitrarily, however it is advisable to choose such dimensions that the surfaces lie somewhere in the "empty" region between detector elements: while the GEANT4 simulation is very precise, the Tracking Geometry is only an approximation based on a projection of all the material within certain regions onto specific tracking surfaces. In regions close to tracking surfaces, the material seen by the extrapolation up to the given physical location can be over- or underestimated with respect to the full simulation.

B.3.2 Comparison of ATLAS Geometry Models

The cylinder-based validation technique can be employed to visualise material changes between different GEO models. Here the user can define the cylinder dimensions arbitrarily. As an example, the redistribution of cables in the PP0 region from ATLAS-GEO-16-01-00 to ATLAS-GEO-18-00-00 is shown in Figure B.5. The layer-by-layer visualisation reveals in which regions (i.e. between which virtual surfaces) the material changes take place.

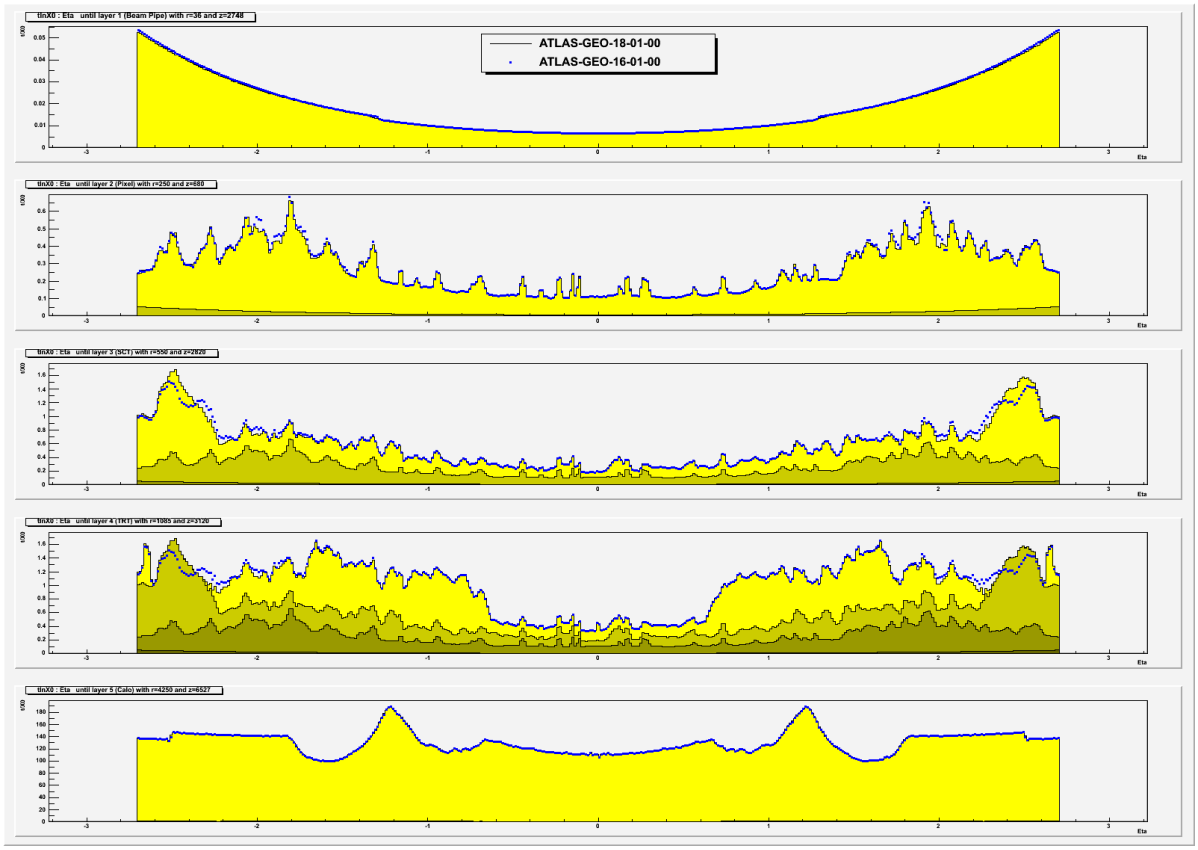


Figure B.5: GEO model comparison plots of material in t/X_0 (with cylinder-based validation framework) between ATLAS-GEO-18-01-00 and ATLAS-GEO-16-01-00. The visible differences around $1.9 < |\eta| < 2.5$ validate the restructuring of cables in the PP0 region outside of the pixel detector.

Appendix C

Additional details of Inner Detector performance

C.1 Pixel Detector Upgrade (IBL)

The performance of the innermost Pixel detector layer (B-layer) is critical for the entire physics program of the ATLAS experiment. During the LS1 phase in 2013 and 2014, a fourth Pixel layer (Insertable B-Layer or IBL) was added to the Pixel detector between a new, smaller beam pipe and the previous innermost Pixel layer [64], and has been fully commissioned. This layer brings a number of benefits, such as increased tracking robustness against failure of pixel modules, higher b-tagging efficiency, and better vertexing performance and tracking precision due to the closer location to the interaction point ($r \approx 33\,\text{mm}$.

Figure C.1 shows the improvement in the transverse and longitudinal impact parameter resolutions, σ_{d_0} and σ_{z_0}, due to the installed IBL. Similar results are obtained for other track parameters. The figure also shows the excellent experimental performance of the track parameter resolution in the Run-1 period with 8 TeV data using the 2012 alignment.

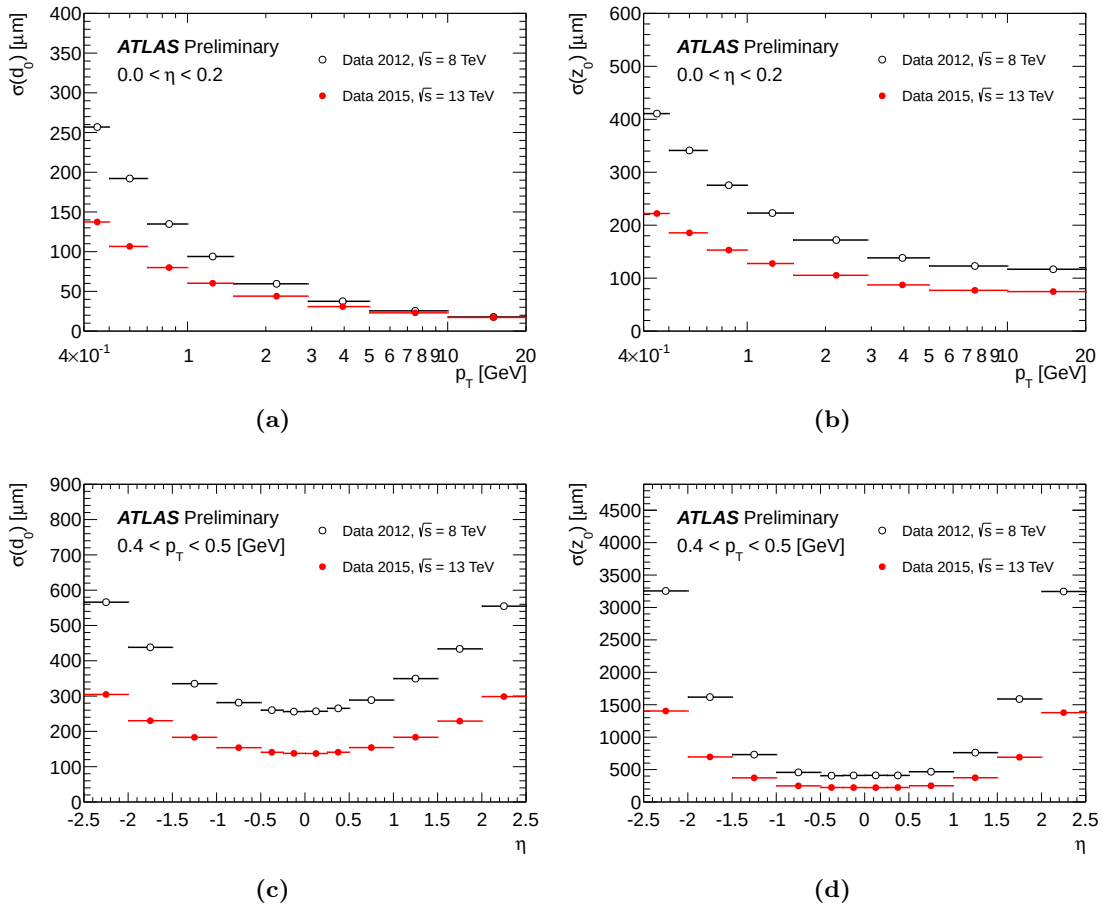


Figure C.1: Unfolded resolution of the transverse and longitudinal track parameters, σ_{d_0} and σ_{z_0}, for single muons as a function of (a,b) p_{T} and (c,d) η. The results measured from 2015 data (red) at $\sqrt{s} = 13$ TeV with the installed IBL, using a minimum-bias trigger, are compared against the results from 2012 data (black) at $\sqrt{s} = 8$ TeV without the IBL, using jet, tau and missing ET triggers. Plots taken from Ref. [167].

Bibliography

[1] L. Evans, (ed.) and P. Bryant, (ed.), *LHC Machine*, JINST **3** (2008) S08001.

[2] M. Boisot, M. Nordberg, S. Yami, and B. Nicquevert, *Collisions and Collaboration: The Organization of Learning in the ATLAS Experiment at the LHC*. Oxford University Press, Oxford, 2011.

[3] ATLAS Collaboration, G. Aad et al., *Charged-particle distributions in pp interactions at $\sqrt{s} = 8$ TeV measured with the ATLAS detector*, Eur. Phys. J. **C76** (2016) 403, arXiv:1603.02439 [hep-ex].

[4] Particle Data Group Collaboration, K. Olive et al., *Review of Particle Physics*, Chin. Phys. C **38** (2014) 090001.

[5] G. Dissertori, I. G. Knowles, and M. Schmelling, *Quantum chromodynamics: high energy experiments and theory*, vol. 115 of *International series of monographs on physics*. Oxford Univ. Press, Oxford, 2003.

[6] G. Dissertori, *The Determination of the Strong Coupling Constant*, arXiv:1506.05407 [hep-ex].

[7] S. Bethke, *Experimental tests of asymptotic freedom*, Prog. Part. Nucl. Phys. **58** (2007) 351–386, arXiv:hep-ex/0606035 [hep-ex].

[8] ZEUS, H1 Collaboration, H. Abramowicz et al., *Combination of measurements of inclusive deep inelastic $e^{\pm}p$ scattering cross sections and QCD analysis of HERA data*, Eur. Phys. J. C **75** (2015) 580, arXiv:1506.06042 [hep-ex].

[9] ZEUS, H1 Collaboration, F. D. Aaron et al., *Combined Measurement and QCD Analysis of the Inclusive e+- p Scattering Cross Sections at HERA*, JHEP **01** (2010) 109, arXiv:0911.0884 [hep-ex].

[10] J. C. Collins, D. E. Soper, and G. F. Sterman, *Factorization of Hard Processes in QCD*, Adv.Ser.Direct.High Energy Phys. **5** (1988) 1–91, arXiv:hep-ph/0409313 [hep-ph].

[11] J. M. Campbell, J. W. Huston, and W. J. Stirling, *Hard Interactions of Quarks and Gluons: A Primer for LHC Physics*, Rept. Prog. Phys. **70** (2007) 89, arXiv:hep-ph/0611148 [hep-ph].

[12] ATLAS Collaboration, *Summary plots from the ATLAS Standard Model physics group*, 2015 (accessed April 3, 2016), https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CombinedSummaryPlots/SM/.

[13] T. Regge, *Introduction to complex orbital momenta*, Nuovo Cimento **14** (1959) 951.

[14] V. N. Gribov, *A Reggeon diagram technique*, Sov. Phys. JETP **26** (1968) 414–422.

[15] S. Donnachie, H. G. Dosch, O. Nachtmann, and P. Landshoff, *Pomeron physics and QCD*, Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol. **19** (2002) 1–347.

[16] P. Bartalini et al., *Multi-Parton Interactions at the LHC*, arXiv:1111.0469 [hep-ph], http://inspirehep.net/record/944170/files/arXiv:1111.0469.pdf.

[17] ATLAS Collaboration, M. Aaboud et al., *Measurement of the total cross section from elastic scattering in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, arXiv:1607.06605 [hep-ex].

[18] TOTEM Collaboration, G. Antchev et al., *Measurement of Elastic pp Scattering at $\sqrt{s} = 8$ TeV in the Coulomb-Nuclear Interference Region – Determination of the ρ Parameter and the Total Cross-Section*, Tech. Rep. CERN-PH-EP-2015-325, CERN, Geneva, Dec, 2015. https://cdsweb.cern.ch/record/2114603.

[19] TOTEM Collaboration, G. Antchev et al., *Luminosity-Independent Measurement of the Proton-Proton Total Cross Section at $\sqrt{s} = 8$ TeV*, Phys. Rev. Lett. **111** (2013) 012001.

[20] S. Navin, *Diffraction in Pythia*, arXiv:1005.3894 [hep-ph].

[21] M. A. Medici, *Diffraction with ALFA and ATLAS at $\sqrt{s} = 8$* TeV, Master's thesis, Niels Bohr Institute, University of Copenhagen, 2013. http://discoverycenter.nbi.ku.dk/teaching/thesis_page/Medici-thesis.pdf. Presented 27 Sep 2013.

[22] L.-J. Zhou, Z.-H. Hu, and W.-X. Ma, *High energy proton-proton elastic scattering in reggeon-pomeron exchange model*, Commun. Theor. Phys. **45** (2006) 1069–1072.

[23] Z.-H. Hu, L.-J. Zhou, and W.-X. Ma, *Reggeon, pomeron and glueball, odderon-hadron-hadron interaction at high energies: From Regge theory to quantum chromodynamics*, Commun. Theor. Phys. **49** (2008) 729–738.

[24] T. Sjostrand, S. Mrenna, and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, Comp. Phys. Comm. **178** (2008) 852–867, arXiv:0710.3820 [hep-ph].

[25] T. Sjostrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, JHEP **05** (2006) 026, arXiv:hep-ph/0603175 [hep-ph].

[26] B. Andersson, G. Gustafson, and B. Soderberg, *A General Model for Jet Fragmentation*, Z. Phys. C **20** (1983) 317.

[27] G. A. Schuler and T. Sjöstrand, *Hadronic diffractive cross sections and the rise of the total cross section*, Phys. Rev. D **49** (1994) 2257–2267, http://link.aps.org/doi/10.1103/PhysRevD.49.2257.

[28] ATLAS Collaboration, *Summary of ATLAS Pythia 8 tunes*, Tech. Rep. ATL-PHYS-PUB-2012-003, CERN, Geneva, Aug, 2012. https://cdsweb.cern.ch/record/1474107.

[29] ATLAS Collaboration, *Further ATLAS tunes of PYTHIA6 and Pythia 8*, ATL-PHYS-PUB-2011-014, 2011, https://cdsweb.cern.ch/record/1400677.

[30] P. Skands, S. Carrazza, and J. Rojo, *Tuning PYTHIA 8.1: the Monash 2013 Tune*, Eur. Phys. J. C **74** (2014) 3024, arXiv:1404.5630 [hep-ph].

[31] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, *Parton distributions for the LHC*, Eur. Phys. J. C **63** (2009) 189–285, arXiv:0901.0002 [hep-ph].

[32] ATLAS Collaboration, G. Aad et al., *ATLAS tunes of PYTHIA 6 and Pythia 8 for MC11*, tech. rep., CERN, Geneva, Jul, 2011. https://cdsweb.cern.ch/record/1363300.

[33] N. Firdous and G. Rudolph, *Tuning of PYTHIA6 to Minimum Bias Data*, EPJ Web Conf. **60** (2013) 20056.

[34] N. Firdous, *Tuning of the PYTHIA 6.4 Multiple Parton Interaction model to Minimum Bias and Underlying Event data*. PhD thesis, University of Innsbruck, 2013. http://physik.uibk.ac.at/hephy/theses/diss_nf.pdf. Presented 6 Feb 2014.

[35] G. Rudolph, *Private communications*, 2015-2016.

[36] S. Porteboeuf, T. Pierog, and K. Werner, *Producing Hard Processes Regarding the Complete Event: The EPOS Event Generator*, arXiv:1006.2967 [hep-ph].

[37] T. Pierog and K. Werner, *Muon Production in Extended Air Shower Simulations*, Phys. Rev. Lett. **101** (2008) 171101, arXiv:astro-ph/0611311 [astro-ph].

[38] K. Werner et al., *Event-by-Event Simulation of the Three-Dimensional Hydrodynamic Evolution from Flux Tube Initial Conditions in Ultrarelativistic Heavy Ion Collisions*, Phys.Rev. **C82** (2010) 044904, arXiv:1004.0805 [nucl-th], https://cdsweb.cern.ch/record/1258083.

[39] H. J. Drescher, M. Hladik, S. Ostapchenko, T. Pierog, and K. Werner, *Parton based Gribov-Regge theory*, Phys. Rep. **350** (2001) 93–289, arXiv:hep-ph/0007198 [hep-ph].

[40] K. Werner et al., *Evidence for Hydrodynamic Evolution in Proton-Proton Scattering at LHC Energies*, Phys.Rev. **C83** (2011) 044915, arXiv:1010.0400 [nucl-th].

[41] Y. Corrales Morales, *Identified hadron distributions in p-Pb collisions at* $\sqrt{s_{NN}}$ *= 5.02 TeV with the Inner Tracking System of ALICE at the LHC*. PhD thesis, Turin U., 2015. http://inspirehep.net/record/1381354/files/CERN-THESIS-2015-028.pdf.

[42] T. Pierog, I. Karpenko, J. M. Katzy, E. Yatsenko, and K. Werner, *EPOS LHC: Test of collective hadronization with data measured at the CERN Large Hadron Collider*, Phys. Rev. C **92** (2015) 034906, arXiv:1306.0121 [hep-ph].

[43] S. Ostapchenko, *Monte Carlo treatment of hadronic interactions in enhanced Pomeron scheme: I. QGSJET-II model*, Phys. Rev. D **83** (2011) 014018, arXiv:1010.1869 [hep-ph].

[44] M. Bahr et al., *Herwig++ Physics and Manual*, Eur. Phys. J. **C58** (2008) 639–707, arXiv:0803.0883 [hep-ph].

[45] S. Gieseke, C. Rohr, and A. Siodmok, *Colour reconnections in Herwig++*, Eur. Phys. J. C **72** (2012) 2225, arXiv:1206.0041 [hep-ph].

[46] J. Pumplin, D. R. Stump, J. Huston, H. L. Lai, P. M. Nadolsky, and W. K. Tung, *New generation of parton distributions with uncertainties from global QCD analysis*, JHEP **07** (2002) 012, arXiv:hep-ph/0201195 [hep-ph].

[47] ATLAS Collaboration, G. Aad et al., *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.

[48] H. Bethe, *Zur Theorie des Durchgangs schneller Korpuskularstrahlen durch Materie*, Annalen der Physik **397** (1930) 325–400.

[49] S. M. Seltzer and M. J. Berger, *Evaluation of the collision stopping power of elements and compounds for electrons and positrons*, The International Journal of Applied Radiation and Isotopes **33** (1982) 1189–1218, http://www.sciencedirect.com/science/article/pii/0020708X82902447.

[50] R. Sternheimer, M. Berger, and S. Seltzer, *Density effect for the ionization loss of charged particles in various substances*, Atomic Data and Nuclear Data Tables **30** (1984) 261–271, http://www.sciencedirect.com/science/article/pii/0092640X84900020.

[51] S. M. Seltzer and M. J. Berger, *Improved procedure for calculating the collision stopping power of elements and compounds for electrons and positrons*, The International Journal of Applied Radiation and Isotopes **35** (1984) 665–676, http://www.sciencedirect.com/science/article/pii/0020708X84901133.

[52] L. Landau, *On the energy loss of fast particles by ionization*, J. Phys. (USSR) **8** (1944) 201–205.

[53] R. M. Sternheimer, *The Density Effect for the Ionization Loss in Various Materials*, Phys. Rev. **88** (1952) 851–859, http://link.aps.org/doi/10.1103/PhysRev.88.851.

[54] Y.-S. Tsai, *Pair production and bremsstrahlung of charged leptons*, Rev. Mod. Phys. **46** (1974) 815–851, http://link.aps.org/doi/10.1103/RevModPhys.46.815.

[55] H. Davies, H. A. Bethe, and L. C. Maximon, *Theory of Bremsstrahlung and Pair Production. II. Integral Cross Section for Pair Production*, Phys. Rev. **93** (1954) 788–795, http://link.aps.org/doi/10.1103/PhysRev.93.788.

[56] H. Bethe and W. Heitler, *On the Stopping of fast particles and on the creation of positive electrons*, Proc. Roy. Soc. Lond. A **146** (1934) 83–112.

[57] H. A. Bethe, *Molière's Theory of Multiple Scattering*, Phys. Rev. **89** (1953) 1256–1266, http://link.aps.org/doi/10.1103/PhysRev.89.1256.

[58] G. R. Lynch and O. I. Dahl, *Approximations to multiple Coulomb scattering*, Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms **58** (1991) 6–10, http://www.sciencedirect.com/science/article/pii/0168583X9195671Y.

[59] B. Rossi and K. Greisen, *Cosmic-Ray Theory*, Rev. Mod. Phys. **13** (1941) 240–309, http://link.aps.org/doi/10.1103/RevModPhys.13.240.

[60] R. Frühwirth and M. Regler, *On the quantitative modelling of core and tails of multiple scattering by Gaussian mixtures*, Nucl. Instrum. Meth. A **456** (2001) 369–389, http://www.sciencedirect.com/science/article/pii/S0168900200005891.

[61] R. Frühwirth and M. Liendl, *Mixture models of multiple scattering: computation and simulation*, Computer Physics Communications **141** (2001) 230–246, http://www.sciencedirect.com/science/article/pii/S0010465501004039.

[62] Service graphique, CERN, *Overall view of the LHC. Vue d'ensemble du LHC*, https://cdsweb.cern.ch/record/1708849, General Photo.

[63] ATLAS Collaboration, M. Aaboud et al., *Luminosity determination in pp collisions at* $\sqrt{s} = 8$ TeV *using the ATLAS detector at the LHC*, to be submitted to Eur. Phys. J. C (2016), https://cdsweb.cern.ch/record/2150760/.

[64] M. Capeans et al., *ATLAS Insertable B-Layer Technical Design Report*, Tech. Rep. CERN-LHCC-2010-013. ATLAS-TDR-19, CERN, Geneva, Sep, 2010. https://cdsweb.cern.ch/record/1291633.

[65] G. Aad et al., *ATLAS pixel detector electronics and sensors*, JINST **3** (2008) P07007, http://iopscience.iop.org/1748-0221/3/07/P07007/.

[66] ATLAS Collaboration, G. Aad et al., *A neural network clustering algorithm for the ATLAS silicon pixel detector*, JINST **9** (2014) P09009, arXiv:1406.7690 [hep-ex].

[67] E. Ritsch, *Fast Calorimeter Punch-Through Simulation for the ATLAS Experiment*, Master's thesis, University of Innsbruck, Innsbruck, 2011. https://inspirehep.net/record/1088082/files/CERN-THESIS-2011-112.pdf. Presented 28 Sep 2011.

[68] ATLAS Collaboration, P. Jenni, M. Nessi, M. Nordberg, and K. Smith, *ATLAS high-level trigger, data-acquisition and controls: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 2003. https://cdsweb.cern.ch/record/616089.

[69] I. Bird et al., *Update of the Computing Models of the WLCG and the LHC Experiments*, Tech. Rep. CERN-LHCC-2014-014. LCG-TDR-002, CERN, Geneva, Apr, 2014. https://cdsweb.cern.ch/record/1695401.

[70] R. E. Kwee, N. Ellis, and H. Kolanoski, *Development and Deployment of an Inner Detector Minimum Bias Trigger and Analysis of Minimum Bias Data of the ATLAS Experiment at the Large Hadron Collider*. PhD thesis, Humboldt U., Berlin, Sep, 2011. https://cdsweb.cern.ch/record/1456529. Presented 13 Jan 2012.

[71] M. Clemencic et al., *Recent developments in the LHCb software framework Gaudi*, J. Phys.: Conf. Ser. **219** (2010) 042006.

[72] G. Barrand et al., *GAUDI - A software architecture and framework for building HEP data processing applications*, Comp. Phys. Comm. **140** (2001) 45–55.

[73] ATLAS Collaboration, *ATLAS Computing: technical design report*. Technical Design Report ATLAS. CERN, Geneva, 2005. https://cdsweb.cern.ch/record/837738.

[74] W. Lukas, *Fast Simulation for ATLAS: Atlfast-II and ISF*, J. Phys.: Conf. Ser. **396** (2012) 022031, http://stacks.iop.org/1742-6596/396/i=2/a=022031.

[75] ATLAS Collaboration, *Luminosity Results for Run-1*, (accessed January 2, 2016), https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults.

[76] ATLAS Collaboration, G. Aad et al., *The ATLAS Simulation Infrastructure*, Eur. Phys. J. C **70** (2010) 823–874, arXiv:1005.4568.

[77] M. Dobbs and J. B. Hansen, *The HepMC C++ Monte Carlo event record for High Energy Physics*, Comp. Phys. Comm. **134** (2001) 41–46.

[78] F. Akesson et al., *ATLAS Tracking Event Data Model*, Tech. Rep. ATL-SOFT-PUB-2006-004, CERN, Geneva, Jul, 2006. https://cdsweb.cern.ch/record/973401.

[79] J. Boudreau and V. Tsulaia, *The GeoModel toolkit for detector description*, http://doc.cern.ch/yellowrep/2005/2005-002/p353.pdf.

[80] A. Salzburger, S. Todorova, and M. Wolter, *The ATLAS Tracking Geometry Description*, Tech. Rep. ATL-SOFT-PUB-2007-004, CERN, Geneva, Jun, 2007. http://cdsweb.cern.ch/record/1038098.

[81] A. Salzburger, *The ATLAS Track Extrapolation Package*, Tech. Rep. ATL-SOFT-PUB-2007-005, CERN, Geneva, Jun, 2007. http://cdsweb.cern.ch/record/1038100.

[82] ATLAS Collaboration, E. Ritsch, *Concepts and Plans towards fast large scale Monte Carlo production for the ATLAS Experiment*, J. Phys.: Conf. Ser. **523** (2014) 012035, http://inspirehep.net/record/1299910.

[83] GEANT4 Collaboration, S. Agostinelli, et al., *GEANT4: A Simulation toolkit*, Nucl. Instrum. Meth. A **506** (2003) 250–303.

[84] H. W. Bertini, *Intranuclear-cascade calculation of the secondary nucleon spectra from nucleon-nucleus interactions in the energy range 340 to 2900 mev and comparisons with experiment*, Phys. Rev. **188** (1969) 1711–1730.

[85] H. W. Bertini and M. P. Guthrie, *Results from medium-energy intranuclear-cascade calculation*, Nucl. Phys. A **169** (1971) 670–672.

[86] B. Andersson, G. Gustafson, and B. Nilsson-Almqvist, *A model for low-pT hadronic reactions with generalizations to hadron-nucleus and nucleus-nucleus collisions*, Nucl. Phys. B **281** (1987) 289–309, http://www.sciencedirect.com/science/article/pii/0550321387902574.

[87] GEANT4 Collaboration, *GEANT4 Physics Reference Manual*, 2012 (accessed January 6, 2016), http://geant4.web.cern.ch/geant4/UserDocumentation/UsersGuides/PhysicsReferenceManual/BackupVersions/V9.6/fo/PhysicsReferenceManual.pdf.

[88] H. Bichsel, *Straggling in thin silicon detectors*, Rev. Mod. Phys. **60** (1988) 663–699, http://link.aps.org/doi/10.1103/RevModPhys.60.663.

[89] S. M. Seltzer and M. J. Berger, *Bremsstrahlung spectra from electron interactions with screened atomic nuclei and orbital electrons*, Nucl. Instrum. Meth. B **12** (1985) 95–134.

[90] S. M. Seltzer and M. J. Berger, *Bremsstrahlung Energy Spectra from Electrons with Kinetic Energy 1 keV-10 GeV Incident on Screened Nuclei and Orbital Electrons of Neutral Atoms with Z = 1-100*, Atomic Data and Nuclear Data Tables **35** (1986) 345.

[91] W. Heitler, *The Quantum Theory of Radiation*. Oxford University Press, London, 1957.

[92] H. W. Lewis, *Multiple Scattering in an Infinite Medium*, Phys. Rev. **78** (1950) 526–529, http://link.aps.org/doi/10.1103/PhysRev.78.526.

[93] R. J. Glauber, *Theory of High Energy Hadron–Nucleus Collisions*. Springer US, Boston, MA, 1970. http://dx.doi.org/10.1007/978-1-4684-1827-9_43.

[94] ATLAS Collaboration, T. Yamanaka, *The ATLAS calorimeter simulation FastCaloSim*, J. Phys.: Conf. Ser. **331** (2011) 032053.

[95] ATLAS Collaboration, *Performance of the Fast ATLAS Tracking Simulation (FATRAS) and the ATLAS Fast Calorimeter Simulation (FastCaloSim) with single particles*, Tech. Rep. ATL-SOFT-PUB-2014-001, CERN, Geneva, Mar, 2014. https://cds.cern.ch/record/1669341.

[96] ATLAS Collaboration, *Data/MC Comparison for Calorimeter Shower Shapes of High E_T Electrons*, 2011 (accessed January 3, 2016), http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/EGAMMA/PublicPlots/20111005/ATL-COM-PHYS-2011-1299/index.html.

[97] K. Edmonds et al., *The Fast ATLAS Track Simulation (FATRAS)*, Tech. Rep. ATL-SOFT-PUB-2008-001, CERN, Geneva, Mar, 2008. https://cdsweb.cern.ch/record/1091969.

[98] ATLAS Collaboration, S. Hamilton et al., *The ATLAS Fast Track Simulation project (FATRAS)*, IEEE Nucl. Sci. Symp. Conf. **2010** (2010) 311–316.

[99] N. Calace and A. Salzburger, *ATLAS Tracking Detector Upgrade studies using the Fast Simulation Engine*, J. Phys.: Conf. Ser. **664** (2015) 072005, https://cdsweb.cern.ch/record/2016123.

[100] J. Mechnich, *FATRAS - the ATLAS Fast Track Simulation project*, J. Phys.: Conf. Ser. **331** (2011) 032046, http://stacks.iop.org/1742-6596/331/i=3/a=032046.

[101] E. Ritsch, *ATLAS Detector Simulation in the Integrated Simulation Framework applied to the W Boson Mass Measurement*. PhD thesis, University of Innsbruck, Innsbruck, 2014. http://physik.uibk.ac.at/hephy/theses/diss_er.pdf.

[102] J. Chapman et al., *Concepts and Plans towards fast large scale Monte Carlo production for the ATLAS Experiment*, https://cdsweb.cern.ch/record/1547511.

[103] R. Mankel, *Pattern recognition and event reconstruction in particle physics experiments*, Rept. Prog. Phys. **67** (2004) 553, arXiv:physics/0402039 [physics].

[104] W. Lukas, *ATLAS inner tracking detectors: Run 1 performance and developments for Run 2*, Nucl. Part. Phys. Proc. **273–275** (2016) 1134–1140, http://www.sciencedirect.com/science/article/pii/S2405601415006677, 37th International Conference on High Energy Physics (ICHEP).

[105] T. Cornelissen et al., *Concepts, Design and Implementation of the ATLAS New Tracking (NEWT)*, Tech. Rep. ATL-SOFT-PUB-2007-007, CERN, Geneva, Mar, 2007. https://cdsweb.cern.ch/record/1020106.

[106] T. Cornelissen et al., *The new ATLAS track reconstruction (NEWT)*, J. Phys.: Conf. Ser. **119** (2008) 032014.

[107] A. Strandlie and R. Frühwirth, *Track and vertex reconstruction: From classical to adaptive methods*, Rev. Mod. Phys. **82** (2010) 1419–1458.

[108] A. Salzburger, *Track Simulation and Reconstruction in the ATLAS Experiment*. PhD thesis, University of Innsbruck, Innsbruck, 2008. http://physik.uibk.ac.at/hephy/theses/diss_as.pdf.

[109] ATLAS Collaboration, G. Aad et al., *Operation and performance of the ATLAS semiconductor tracker*, JINST **9** (2014) P08009, arXiv:1404.7473 [hep-ex].

[110] ATLAS Collaboration, G. Aad et al., *Basic ATLAS TRT performance studies of Run 1*, https://cdsweb.cern.ch/record/1669603.

[111] A. Andreazza et al., *Validation of MBTS trigger performance and low-pT tracking performance in low-luminosity 8 TeV data*, Tech. Rep. ATL-COM-INDET-2012-052, CERN, Geneva, Jul, 2012. https://cdsweb.cern.ch/record/1463926.

[112] ATLAS Collaboration, K. Grimm et al., *Reconstruction of primary vertices at the ATLAS experiment in Run 1 proton-proton collisions at the LHC*, Tech. Rep. CERN-EP-2016-150, CERN, Geneva, Jun, 2016. https://cds.cern.ch/record/2158820.

[113] K. Grimm et al., *Methods to quantify the performance of the primary vertex reconstruction in the ATLAS experiment under high luminosity conditions*, https://cdsweb.cern.ch/record/1457597.

[114] ATLAS Collaboration, *Performance of the ATLAS Inner Detector Track and Vertex Reconstruction in the High Pile-Up LHC Environment*, https://cdsweb.cern.ch/record/1435196.

[115] K. Prokofiev, *Reconstruction of primary vertices in pp collisions at energies of 900 GeV and 7 TeV with the ATLAS detector*, arXiv:1010.1483 [physics.ins-det].

[116] ATLAS Collaboration, *Performance of primary vertex reconstruction in proton–proton collisions at $\sqrt{s}$ = 7 TeV in the ATLAS experiment*, ATLAS-CONF-2010-069, 2010, https://cdsweb.cern.ch/record/1281344.

[117] R. Frühwirth, W. Waltenberger, and P. Vanlaer, *Adaptive vertex fitting*, J.Phys. **G34** (2007) N343.

[118] S. Pagan Griso, K. Prokofiev, A. Andreazza, K. Grimm, E. Guido, F. Meloni, M. Rudolph, A. Salzburger, and A. Wildauer, *Vertex reconstruction plots*. Geneva, Apr, 2012, https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/IDTRACKING/PublicPlots/ATL-COM-PHYS-2012-474/.

[119] G. Piacquadio, K. Prokofiev, and A. Wildauer, *Primary vertex reconstruction in the ATLAS experiment at LHC*, J. Phys.: Conf. Ser. **119** (2008) 032033.

[120] ATLAS Collaboration, *Characterization of Interaction-Point Beam Parameters Using the pp Event-Vertex Distribution Reconstructed in the ATLAS Detector at the LHC*, Tech. Rep. ATLAS-CONF-2010-027, CERN, Geneva, May, 2010. https://cdsweb.cern.ch/record/1277659.

[121] ATLAS Collaboration, *Performance of the ATLAS Secondary Vertex b-tagging Algorithm in 7 TeV Collision Data*, Tech. Rep. ATLAS-CONF-2010-042, CERN, Geneva, Jul, 2010. https://cdsweb.cern.ch/record/1277682.

[122] ATLAS Collaboration, *Alignment of the ATLAS Inner Detector and its Performance in 2012*, Tech. Rep. ATLAS-CONF-2014-047, CERN, Geneva, Jul, 2014. https://cdsweb.cern.ch/record/1741021.

[123] J. Biesiada, T. Eifert, A. Salzburger, and M. Schott, *Study of the Material Budget in the ATLAS Inner Detector with K0s in 900GeV Collision Data*, Tech. Rep. ATL-INDET-INT-2010-019, CERN, Geneva, Oct, 2010. https://cdsweb.cern.ch/record/1297667.

[124] H. Gray, G. Gorfine, M. Limper, and A. Salzburger, *Using the SCT Extension Efficiency to Probe Material between the Pixel Detector and the SCT*, Tech. Rep. ATL-PHYS-INT-2010-095, CERN, Geneva, Sep, 2010. https://cdsweb.cern.ch/record/1290364.

[125] H. M. Gray, *The Charged Particle Multiplicity at Center of Mass Energies from 900 GeV to 7 TeV measured with the ATLAS Experiment at the Large Hadron Collider*. PhD thesis, Caltech, 2010. http://inspirehep.net/record/886915/files/CERN-THESIS-2010-163.pdf. Presented 9 Nov 2010.

[126] ATLAS Collaboration, *Charged-particle multiplicities in pp interactions measured with the ATLAS detector at the LHC*, New J. Phys. **13** (2011) 053033, arXiv:1012.5104 [hep-ex].

[127] G. Aad et al., *Track Reconstruction Efficiency in $\sqrt{s} = 7$ TeV Data for Tracks with $p_T > 100$ MeV*, Tech. Rep. ATL-PHYS-INT-2011-001, CERN, Geneva, Jan, 2011. https://cdsweb.cern.ch/record/1319318.

[128] A. Salzburger, S. Pagan Griso, and A. Morley, *TrackingCP plots for approval: Plots for approval for LHCC*, Tech. Rep. ATL-COM-PHYS-2014-156, CERN, Geneva, Feb, 2014. http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/IDTRACKING/PublicPlots/ATL-COM-PHYS-2014-156/.

[129] ATLAS Collaboration, G. Aad et al., *A study of the material in the ATLAS inner detector using secondary hadronic interactions*, JINST **7** (2012) P01013, arXiv:1110.6191 [hep-ex].

[130] ATLAS Collaboration, *Two Dimensional Views of the Run-1 ATLAS Inner Detector Produced by the Use of Hadronic Interaction Vertices*, 2016 (accessed July 7, 2016), https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/IDTR-2016-001/.

[131] D. Banfi, *Electron and photon reconstruction and identification with the ATLAS detector and performance with $\pi^0 \to \gamma\gamma$ and $\eta \to \gamma\gamma$ with $\sqrt{s} = 0.9$ TeV data*, Tech. Rep. ATL-LARG-PROC-2013-017, CERN, Geneva, Dec, 2013. https://cdsweb.cern.ch/record/1639611.

[132] ATLAS Collaboration, A. Morley, *Early Material Studies at the ATLAS Experiment*, Tech. Rep. ATL-PHYS-SLIDE-2010-273, CERN, Geneva, Sep, 2010. https://cdsweb.cern.ch/record/1290107.

[133] ATLAS Collaboration, K. Tackmann, *ATLAS Inner Detector Studies with Photon Conversions*, Tech. Rep. ATL-PHYS-SLIDE-2010-154, CERN, Geneva, Jun, 2010. https://cdsweb.cern.ch/record/1272442.

[134] T. Koffas and K. Tackmann, *ATLAS Tracker Material using Photon Conversions*, https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/EGAMMA/PublicPlots/20110216/ATL-COM-PHYS-2011-120/.

[135] W. Lukas, *Measurements of charged-particle distributions at $\sqrt{s} = 8$ and 13 TeV with the ATLAS detector*, Tech. Rep. ATL-PHYS-PROC-2016-086, CERN, Geneva, Jun, 2016. https://cdsweb.cern.ch/record/2195278.

[136] ATLAS Collaboration, *Charged-particle multiplicities in pp interactions at $\sqrt{s} = 900$ GeV measured with the ATLAS detector at the LHC*, Phys. Lett. B **688** (2010) 21–42, arXiv:1003.3124 [hep-ex].

[137] ATLAS Collaboration, G. Aad et al., *Charged-particle distributions in $\sqrt{s} = 13$ TeV pp interactions measured with the ATLAS detector at the LHC*, Phys. Lett. B **758** (2016) 67–88, arXiv:1602.01633 [hep-ex].

[138] ATLAS Collaboration, M. Aaboud et al., *Charged-particle distributions at low transverse momentum in $\sqrt{s}$=13 TeV pp interactions measured with the ATLAS detector at the LHC*, arXiv:1606.01133 [hep-ex], https://cds.cern.ch/record/2157893.

[139] UA1 Collaboration, G. Arnison, et al., *Transverse Momentum Spectra for Charged Particles at the CERN Proton anti-Proton Collider*, Phys. Lett. B **118** (1982) 167.

[140] ABCDHW Collaboration, A. Breakstone, et al., *Charged Multiplicity Distribution in pp Interactions at ISR Energies*, Phys. Rev. D **30** (1984) 528–535.

[141] UA5 Collaboration, R. E. Ansorge, et al., *Diffraction Dissociation at the CERN Pulsed Collider at CM Energies of 900 GeV and 200 GeV*, Z. Phys. C **33** (1986) 175.

[142] UA5 Collaboration, G. J. Alner, et al., *UA5: A general study of proton–antiproton physics at $\sqrt{s}$ = 546 GeV*, Phys. Rep.**154** (1987) 247–383.

[143] UA5 Collaboration, R. E. Ansorge, et al., *Charged Particle Correlations in $\bar{p}p$ Collisions at c.m. Energies of 200 GeV, 546 GeV and 900 GeV*, Z. Phys. C **37** (1988) 191–213.

[144] CDF Collaboration, F. Abe, et al., *Transverse Momentum Distributions of Charged Particles Produced in $\bar{p}p$ Interactions at $\sqrt{s}$ = 630 GeV and 1800 GeV*, Phys. Rev. Lett. **61** (1988) 1819–1822.

[145] UA5 Collaboration, R. E. Ansorge, et al., *Charged Particle Multiplicity Distributions at 200 GeV and 900 GeV Center-Of-Mass Energy*, Z. Phys. C **43** (1989) 357–374.

[146] UA1 Collaboration, C. Albajar, et al., *A Study of the General Characteristics of $p\bar{p}$ Collisions at $\sqrt{s}$ = 0.2 TeV to 0.9 TeV*, Nucl. Phys. B **335** (1990) 261–287.

[147] CDF Collaboration, F. Abe, et al., *Pseudorapidity distributions of charged particles produced in $\bar{p}p$ interactions at $\sqrt{s}$ = 630 GeV and 1800 GeV*, Phys. Rev. D **41** (1990) 2330–2333.

[148] E735 Collaboration, T. Alexopoulos, et al., *Multiplicity dependence of transverse momentum spectra of centrally produced hadrons in $\bar{p}p$ collisions at 0.3 TeV, 0.54 TeV, 0.9 TeV, and 1.8 TeV center-of-mass energy*, Phys. Lett. B **336** (1994) 599–604.

[149] CDF Collaboration, D. E. Acosta et al., *Soft and hard interactions in $p\bar{p}$ collisions at $\sqrt{s}$ = 1800-GeV and 630-GeV*, Phys. Rev. D **65** (2002) 072005.

[150] CDF Collaboration, T. Aaltonen, et al., *Measurement of Particle Production and Inclusive Differential Cross Sections in $p\bar{p}$ Collisions at $\sqrt{s}$ = 1.96 TeV*, Phys. Rev. D **79** (2009) 112005, `arXiv:0904.1098 [hep-ex]`, [Erratum: Phys. Rev. D **82** (2010) 119903].

[151] ALICE Collaboration, K. Aamodt, et al., *First proton–proton collisions at the LHC as observed with the ALICE detector: Measurement of the charged particle pseudorapidity density at $\sqrt{s}$ = 900 GeV*, Eur. Phys. J. C **65** (2010) 111–125, `arXiv:0911.5430 [hep-ex]`.

[152] CMS Collaboration, *Transverse momentum and pseudorapidity distributions of charged hadrons in pp collisions at $\sqrt{s}$ = 0.9 and 2.36 TeV*, JHEP **02** (2010) 041, `arXiv:1002.0621 [hep-ex]`.

[153] ALICE Collaboration, K. Aamodt, et al., *Charged-particle multiplicity measurement in proton–proton collisions at $\sqrt{s} = 7$ TeV with ALICE at LHC*, Eur. Phys. J. C **68** (2010) 345–354, `arXiv:1004.3514 [hep-ex]`.

[154] CMS Collaboration, *Transverse-momentum and pseudorapidity distributions of charged hadrons in pp collisions at $\sqrt{s}$ = 7 TeV*, Phys. Rev. Lett. **105** (2010) 022002, arXiv:1005.3299 [hep-ex].

[155] CMS Collaboration, *Charged particle multiplicities in pp interactions at $\sqrt{s}$ = 0.9, 2.36, and 7 TeV*, JHEP **01** (2011) 079, arXiv:1011.5531 [hep-ex].

[156] LHCb Collaboration, R. Aaij, et al., *Measurement of charged particle multiplicities and densities in pp collisions at $\sqrt{s}$ = 7 TeV in the forward region*, Eur. Phys. J. C **74** (2014) 2888, arXiv:1402.4430 [hep-ex].

[157] CMS and TOTEM Collaborations, S. Chatrchyan, et al., *Measurement of pseudorapidity distributions of charged particles in proton–proton collisions at $\sqrt{s}$ = 8 TeV by the CMS and TOTEM experiments*, Eur. Phys. J. C **74** (2014) 3053, arXiv:1405.0722 [hep-ex].

[158] ALICE Collaboration, J. Adam, et al., *Charged-particle multiplicities in proton–proton collisions at $\sqrt{s}$ = 0.9 to 8 TeV*, arXiv:1509.07541 [nucl-ex], submitted to Eur. Phys. J. C.

[159] CMS Collaboration, *Pseudorapidity distribution of charged hadrons in proton–proton collisions at $\sqrt{s}$ = 13 TeV*, Phys. Lett. B **751** (2015) 143–163, arXiv:1507.05915 [hep-ex].

[160] ALICE Collaboration, J. Adam, et al., *Pseudorapidity and transverse-momentum distributions of charged particles in proton–proton collisions at $\sqrt{s}$ = 13 TeV*, Phys. Lett. B **753** (2016) 319–329, arXiv:1509.08734 [nucl-ex].

[161] A. Buckley, J. Butterworth, L. Lonnblad, D. Grellscheid, and H. Hoeth, *Rivet user manual*, Comp. Phys. Comm. **184** (2013) 2803–2819, arXiv:1003.0694 [hep-ph].

[162] ATLAS Collaboration, G. Aad et al., *Beam-induced and cosmic-ray backgrounds observed in the ATLAS detector during the LHC 2012 proton-proton running period*, JINST **11** (2016) P05013, arXiv:1603.09202 [hep-ex].

[163] ALICE Collaboration, B. Abelev et al., *Multi-strange baryon production in pp collisions at $\sqrt{s} = 7$ TeV with ALICE*, Phys. Lett. **B712** (2012) 309–318, arXiv:1204.0282 [nucl-ex].

[164] ALICE Collaboration, B. B. Abelev et al., *Production of $\Sigma(1385)^{\pm}$ and $\Xi(1530)^{0}$ in proton-proton collisions at $\sqrt{s} = 7$ TeV*, Eur. Phys. J. **C75** (2015) 1, arXiv:1406.3206 [nucl-ex].

[165] A. K. Morley and S. Pagan Griso, *Summary of Run 1 ID material description discussion*, Tech. Rep. ATL-COM-INDET-2015-019, CERN, Geneva, Nov, 2013. https://cdsweb.cern.ch/record/2008451.

[166] G. D'Agostini, *A Multidimensional unfolding method based on Bayes' theorem*, Nucl. Instrum. Meth. A **362** (1995) 487–498.

[167] ATLAS Collaboration, *Impact Parameter Resolution*, 2016 (accessed July 7, 2016), https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/IDTR-2015-007/.

[168] A. Buckley, *The hepthesis LaTeX class*, https://www.ctan.org/tex-archive/macros/latex/contrib/hepthesis.

List of figures

List of tables

Glossary

Abbreviation	Meaning
ALFA	**A**bsolute **L**uminosity **f**or **A**TLAS
ALICE	**A** **L**arge **I**on **C**ollider **E**xperiment
ATLAS	**A** **T**oroidal **L**HC **A**pparatu**S**
CERN	**C**onseil **E**uropéen pour la **R**echerche **N**ucléaire
CMS	**C**ompact **M**uon **S**olenoid
DD	**D**ouble-**D**iffractive interaction
EF	**E**vent **F**ilter
FATRAS	**F**ast **A**TLAS **Tra**ck **S**imulation
FSR	**F**inal-**S**tate **R**adiation
GEANT	**Ge**ometry **an**d **T**racking
ID	**I**nner **D**etector
ISF	**I**ntegrated **S**imulation **F**ramework
ISR	**I**nitial-**S**tate **R**adiation
L1	**L**evel-**1** trigger
L2	**L**evel-**2** trigger
LB	**L**uminosity **B**lock
LHC	**L**arge **H**adron **C**ollider
LHCb	**L**arge **H**adron **C**ollider **b**eauty experiment
LS1	**L**ong **S**hutdown **1** (2013-2014)
MB	**M**inimum **B**ias
MBTS	**M**inimum **B**ias **T**rigger **S**cintillator
MC	**M**onte **C**arlo
MPI	**M**ultiple **P**arton **I**nteractions
MS	**M**uon **S**pectrometer

Abbreviation	**Meaning** (cont'd)
ND	**N**on-**D**iffractive interaction
PDF	**P**arton **D**istribution **F**unction
pp	**P**roton-**P**roton interaction
PP0	**P**atch-**P**anel **0**
QCD	**Q**uantum **C**hromo**d**ynamics
Run-1	first data-taking period of the LHC (2009-2013)
Run-2	second data-taking period of the LHC (2015-2018)
SCT	**S**emi**c**onductor **T**racker
SD	**S**ingle-**D**iffractive interaction
SM	**S**tandard **M**odel of particle physics
TRT	**T**ransition **R**adiation **T**racker
UE	**U**nderlying **E**vent

Colophon

This thesis was made in LaTeX 2_ε using the "hepthesis" class [168].