UNIVERSITÀ DELLA CALABRIA

SCUOLA DI DOTTORATO IN SCIENZE E INGEGNERIA DELL'AMBIENTE, DELLE COSTRUZIONI E DELL'ENERGIA, CICLO XXIX

DOCTORAL THESIS

Charged-Particle Distributions and Material Measurements in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS Inner Detector

---

Author

Valentina Maria Martina CAIRO

Supervisors

Dr. Andrea DELL'ACQUA – CERN

Prof. Marco SCHIOPPA – UNICAL

Doctoral School Coordinator

Prof. Pietro Salvatore PANTANO

A thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy

in the

INFN Gruppo Collegato di Cosenza, Laboratori Nazionali di Frascati
Department of Physics, Università Della Calabria, Rende, Italy

Settore Scientifico-Disciplinare: FIS/01

Academic Year 2016-2017

A Nino,
Franca,
Federica
perché sono e saranno sempre loro
la mia vera Vittoria.

UNIVERSITÀ DELLA CALABRIA

Abstract

Department of Physics

Doctor of Philosophy

Charged-Particle Distributions and Material Measurements in $\sqrt{s}$ = 13 TeV pp collisions with the ATLAS Inner Detector

by Valentina Maria Martina CAIRO

ATLAS INNER DETECTOR, TRACKING, MATERIAL STUDIES, SOFT QCD, STANDARD MODEL

The Run 2 of the Large Hadron Collider, which began in Spring 2015, offers new challenges to the Experiments with its unprecedented energy scale and high luminosity regime. To cope with the new experimental conditions, the ATLAS Experiment was upgraded during the first long shutdown of the collider, in the period 2013-2014.

The most relevant change which occurred in the ATLAS Inner Detector was the installation of a fourth pixel layer, the Insertable B-Layer, at a radius of 33 mm together with a new thinner beam pipe. The Pixel Services, located between the Pixel and SCT detectors, were also modified. Owing to the radically modified ID layout, many aspects of the track reconstruction programs had to be re-optimized.

In this thesis, the improvements to the tracking algorithms and the studies of the material distribution in the Inner Detector are described in detail, together with the improvements introduced in the geometry model description in simulation as well as the re-evaluation and the reduction of the systematic uncertainty on the estimate of the track reconstruction efficiency.

The results of these studies were applied to the measurement of Charged-Particle Multiplicity in proton–proton collisions at a centre-of-mass energy of 13 TeV. The charged-particle multiplicity, its dependence on transverse momentum and pseudorapidity and the dependence of the mean transverse momentum on the charged-particle multiplicity are presented for various fiducial phase spaces. The measurements are corrected for detector effects, presented as particle-level distributions and are compared to the predictions of different Monte Carlo event generators.

New sets of recommended performance figures along with the related systematic uncertainties were also derived for several aspects of the ATLAS tracking, such as track reconstruction efficiency, fake rate and impact parameter resolution. These recommendations provide information on appropriate working points, i.e. track selection criteria with well-understood performance. They apply to physics analyses using Inner Detector tracks in Run 2 data and are important inputs for other objects based on tracks, such as jets. A simulation-based method which uses the tracking recommendations to calibrate light-jets mis-tagged as b-jets it is also presented in the context of the measurement of the cross-section of the W-boson produced in association with b-jets at 13 TeV, together with an overview of the inclusive W-boson cross-section analysis.

UNIVERSITÀ DELLA CALABRIA

Abstract

Department of Physics

Doctor of Philosophy

Charged-Particle Distributions and Material Measurements in $\sqrt{s}$ = 13 TeV pp collisions with the ATLAS Inner Detector

by Valentina Maria Martina CAIRO

ATLAS INNER DETECTOR, TRACKING, MATERIAL STUDIES, SOFT QCD, STANDARD MODEL

Il Run 2 del Large Hadron Collider, iniziato nella Primavera del 2015, permette agli Esperimenti di esplorare regioni di energia e luminositá mai raggiunte prima. Al contempo, però, gli apparati sperimentali devono essere aggiornati per poter operare in condizioni sempre più estreme.

Per queste ragioni, un quarto strato di rivelatore a pixel, chiamato Insertable B-Layer, è stato aggiunto nel tracciatore interno dell'Esperimento ATLAS. Al contempo, anche il tubo a vuoto, in cui circolano i fasci di protoni, è stato sostituito con uno più sottile del precedente. La regione dei servizi dei rivelatori a Pixel é stata ugualmente modificata ed ottimizzata nello stesso periodo. Questi cambiamenti hanno richiesto uno sforzo considerevole della Collaboration ATLAS nella ri-ottimizzazione degli algoritmi di tracciamento che tenesse conto delle modiche apportate.

In questa tesi viene riportato il lavoro originale dell'Autore per quanto concerne i miglioramenti degli algoritmi di tracciamento e gli studi sulle distribuzione dei materiali nel tracciatore interno dell'Esperimento ATLAS. L'applicazione di tale lavoro nell'ambito di analisi di Fisica è a sua volta descritta.

In particolare, gli studi dei materiali nel tracciatore interno hanno determinato un miglioramento nella descrizione dei modelli geometrici nella simulazione dell'esperimento, oltre che una notevole riduzione delle incertezze sistematiche nella stima dell'efficienza di ricostruzione delle tracce, che rappresenta l'ingrediente principale per le misure di Molteplicità di Particelle Cariche.

Tali risultati sono di validità generale e sono stati usati anche per calcolare valori di riferimento ed incertezze sistematiche per vari aspetti del tracciamento di particelle cariche in ATLAS. Queste raccomandazioni si applicano ad analisi di fisica che vedono direttamente coinvolte tracce oppure oggetti più complessi che si basano a loro volta su tracce, come i jets. Un esempio puo' essere l'applicazione di tali raccomandazioni per calibrare jets provenienti da quark leggeri che vengono erroneamente associati con il quark beauty. Questa calibrazione è estremamente importante nelle misure di sezione d'urto di produzione del bosone vettore W in associazione con jets da quark beauty, per le quali rappresenta una delle principali sorgenti di errore sistematico.

Acknowledgements

To the University of Calabria and in particular to Marco Schioppa, for being the first Professor trusting in me and for guiding me towards the CERN Main Entrance.
All'Università della Calabria ed in particolare a A Marco Schioppa, per essere stato il primo Professore a credere in me e per avermi guidata verso l'ingresso principale del CERN.

To Andrea Dell'Acqua, for being there waiting for me, supporting me strongly in every choice I made, being mad at me when I could not stop asking questions. And for that day, when I was sad and you said: "In this field as well as in the real life, you will always receive a lot of slaps. The only reasonable reaction is to get angry like a wild beast and then get back to doing what you were doing with much more motivation than before." Thank you for letting me grow up.
Ad Andrea Dell'Acqua, per essere stato lì ad aspettarmi, per avermi supportata in qualsiasi scelta abbia fatto, per essersi arrabbiato con me quando non riuscivo a smettere di fare domande. E per quel giorno in cui ero triste ed hai detto: "In questo ambiente, così come nella vita vera, di sberle ne riceverai sempre parecchie. L'unica reazione ragionevole è incazzarsi come una belva e poi tornare a fare ciò che stavi facendo con una motivazione in più rispetto a prima". Grazie per avermi fatta crescere.

To Heather Gray, who is a brilliant physicist and the key person who made me loving tracking. Thank you for everything you taught me and for all the suggestions you gave me when I was in panic mode.
Ad Heather Gray, che è un fisico brillante ed anche la persona chiave che mi ha fatta appassionare al tracking. Grazie per tutto ciò che mi hai insegnato e per tutti i suggerimenti che mi hai dato quando entravo in panico.

To Anthony Morley, the best colleague I have ever had, an infinite source of help, a scientist with a great and broad preparation and, above all, a rare kindness. For all the times we made a mess and for all the times we fixed it, thank you, it has been a lot of fun!
Ad Anthony Morley, il migliore collega che abbia mai avuto, una sorgente infinita di aiuto, uno scienziato dalla grande preparazione e soprattutto dalla rara gentilezza. Per tutte le volte in cui abbiamo fatto dei pasticci e per tutte le volte in cui li abbiamo risolti, grazie, è stato molto divertente!

To Daniel Froidevaux, for being an undiscussed mentor and make me feel, since the first day, part of big CERN ATLAS Team family.
A Daniel Froidevaux, per essere stato un mentore indiscusso e per avermi fatta sentire parte della grande famiglia del gruppo CERN di ATLAS sin dal primo giorno.

To Christoph Rembser, for being always positive and for being able to communicate it even when supporting me. It is always a pleasure to talk to you.
A Christoph Rember, per la sua positività e per la capicità di comunicarlo anche nel supportare e sostenere me. È sempre un piacere parlare con te.

To Andy Chisholm and Hideyuki Oide, for our beloved radiographies of the ATLAS Inner Detector. I am honored to have the possibility to share this experience with you both.
Ad Andy Chisholm e Hideyuki Oide, per le nostre care radiografie dell'Inner Detector di ATLAS. Sono onorata di avere la possibilità di condividere questa esperienza con voi.

To Oldrich Kepka, Christoph Eckardt and Filip Nechansky for all our iterations on the charged-particle multiplicity analysis.
Ad Oldrich Kepka, Christoph Eckardt e Filip Nechansky, per tutte le nostre iterazioni sull'analisi di molteplicità di particelle cariche.

To Simone Pagan Griso and Andi Salzburger, who seem to be able to explain everything one wants to know on the performance of an inner tracker
A Simone Pagan Griso ed Andi Salzburger, che sembrano saper spiegare tutto ciò si voglia sapere sulle performance di un tracciatore interno.

To Federico Sforza, with whom I share an office and also many coffees and physics chats. I always appreciate your diagrams on the whiteboard
A Federico Sforza, con cui condivido un ufficio ed anche molti caffè e chiacchierate di fisica. Apprezzo sempre i tuoi diagrammi alla lavagna!

To Vivek Jain, for letting me work on this thesis without being worried about my near future.
A Vivek Jain, per avermi dato l'opportunità di lavorare su questa tesi senza preoccuparmi del mio prossimo futuro.

To my friends, Chilufya Mwewa and Verena Martinez, for your energy and enthusiasm which always traps me.
Alle mie amiche, Chilufya Mwewa and Verena Martinez, per la vostra energia ed il vostro entusiamo che mi cattura sempre.

To Vincenzo Catanzaro, for all the times he makes me laughing.
A Vincenzo Catanzaro, per tutte le volte in cui mi fa ridere.

To the CERN ATLAS Muon section, to the Tracking Combined Performance and to the Standard Model groups, because they did not make me feel alone.
Alla sezione dei Muoni di ATLAS ed ai gruppi di Tracking Combined Performance e Standard Model, per non avermi fatta sentire sola.

To the members of the ATLAS Early Career Scientist Board: Gabriel Facini, Claire Lee, Ruth Pöttgen, Kate Pachal, Miha Muskinja, Rustem Ospanov with whom I am sharing a new adventure and who well know how stressful can be preparing a thesis!
Ai membri del Comitato dei Giovani Scienziati di ATLAS: Gabriel Facini, Claire Lee, Ruth Pöttgen, Kate Pachal, Miha Muskinja, Rustem Ospanov con cui sto condividendo una nuova avventura ed i quali sanno bene quanto possa essere stressante preparare una tesi!

To the wonderful ladies who make the ATLAS secretariat a special place where one can get any help.
Alle meravigliose ragazze del segretariato di ATLAS, che lo rendono un posto speciale dove rivolgersi per qualsiasi necessità.

And, especially,

To my mom Franca, for being the strongest Woman in the World, my absolute example to never give up. For every time you know what I think with no need to talk.
A mia mamma, la Donna più forte del Mondo, il mio esempio assoluto per non mollare mai. Per tutte quelle volte in cui sai cosa penso, senza bisogno di parlare.

To my dad Nino, for the tireless support you always give me. For all our discussions, for trusting each other and for being, always, so similar.
A mio padre Nino, per l'instancabile supporto che mi dai sempre. Per tutte le nostre discussioni, per la fiducia reciproca e per essere sempre così simili.

To my sister Federica, Princess, who will hate me for having said that. For your comics and your photos which make this thesis unique. For being absolutely my favourite artist, for never agreeing on a compromise, for being silent and brave.
A mia sorella Federica, Princi, che mi odierà per averlo detto. Per le tue vignette e le tue foto che rendono questa tesi unica. Per essere la mia artista preferita in assoluto, per non scendere mai a compromessi, per essere silenziosa e coraggiosa.

To Purpie, my inseparable travelling companion, for all the destinations which are just waiting for us.
A Purpie, il mio inseparabile compagno di viaggio, per tutte le destinazioni che ci stanno aspettando.

To Domenico e Pietro, because they always hold my hand.
A Domenico e Pietro, perché mi tengono sempre per mano.

And, as usual,

To Albert Einstein, born on March 14th 1879. Ad Albert Einstein, nato il 14 Marzo 1879.

To Valentina Cairo, born on March 14th 1989. A Valentina Cairo, nata il 14 Marzo 1989.

Conn(ll)ecting the dots, by Federica Cairo.

Contents

List of Figures

List of Tables

Introduction

The Large Hadron Collider (LHC), the largest accelerator built so far, allowed to reach major advances in physics already during its first period of operation, started in 2009. The list of highlights is led by the discovery of the Higgs boson, measured by the ATLAS and CMS Experiments and announced on 4 July 2012, approximately 50 years after it was hypothesised.

In the period 2013-2014, during the first long shutdown of the collider, the accelerator complex was consolidated to allow the energy to be increased and the experiments were improved to meet the increasingly challenging experimental conditions of the LHC. In the ATLAS Experiment, a fourth pixel layer, the Insertable B-Layer, was installed in the Inner Detector at a radius of 33 mm, together with a new thinner beam pipe. The Pixel Services, located between the Pixel and SCT detectors, were also modified. The Run 2 of the Large Hadron Collider began in Spring 2015 at the unprecedented centre-of-mass energy of 13 TeV and with record instantaneous luminosities, which constantly exceed the design value.

The structure of this thesis reflects the following.

The theoretical aspects are presented in Chapter 1, with a brief description of the key elements of the Standard Model of Particle Physics and an overview of the soft Quantum Chromodynamics. The crucial aspects of the Monte Carlo Event Generator and the tuning strategies are also described.

The Large Hadron Collider, its main Run 1 performance and Run 2 objectives are shown in Chapter 2 with a description of the ATLAS Experiment and its components.

The improvements to the tracking algorithms, which were made during the long shutdown, together with the performance of the upgraded ATLAS Inner Detector and the recommended benchmark values to be used in the ATLAS Run 2 analyses are presented in Chapter 3. These recommendations, which provide information on appropriate track selection criteria with well-understood performance, apply to physics analyses using Inner Detector tracks in Run 2 data and are important inputs for other objects based on tracks, such as jets.

The material distribution in the revised ID has thoroughly been studied, which led to an improvement of the geometry model description in simulation as well as to a significant reduction of the systematic uncertainty on the estimate of the track reconstruction efficiency, which constitutes the main ingredient of the Charged Particle Multiplicity measurements. The material distribution in the forward pseudorapidity region of the Run 2 ATLAS Inner Detector was studied in detail with the Track-Extension Efficiency method. For the first time in ATLAS, this measurement and the related uncertainties were explicitely expressed in terms of nuclear interaction lenghts. These results are presented in detail in Chapter 4, together with a brief overview of the other two methods used to study the central region, namely Hadronic Interactions and Photon Conversions.

In Chapter 5, the charged-particle distributions measured with the ATLAS detector at 13 TeV are presented, with a particular emphasis on the tracking related aspects. The results

are shown for three different fiducial phase spaces, which include requiring events with at least one charged particle with $|\eta| < 2.5$ and $p_{\mathrm{T}} > 500$ MeV together with a measurement performed in the low transverse momentum regime by requiring events with at least two charged particle with $|\eta| < 2.5$ and $p_{\mathrm{T}} > 100$ MeV and an additional measurement performed in a phase space common to the ATLAS, CMS and ALICE experiments, in which events with at least 1 charged particle with $|\eta| < 0.8$ and $p_{\mathrm{T}} > 500$ MeV are selected.

Applications of the tracking performance, described in Chapter 3, are summarised in Appendix A. In particular, jets from light-flavour quarks are sometimes mis-tagged as jets from b-quark, and this effect is usually evaluated by means of data-driven techniques. By using the tracking recommendations, instead, a simulation-based approach can be developed, which consists in evaluating the systematics on tracks and seeing how much this affects the mis-tag rate in simulation by using the modified track collection in the b-tagging algorithms. The measurement of the cross-section of the W-boson produced in association with b-jets, to which this method will be applied, is also described in Appendix A together with an overview of the inclusive W-boson cross-section analysis.

Auxiliary distributions related to the Track-Extension Efficiency, presented in detail in Chapter 4, are collected in Appendix B.

The ATLAS Experiment is undergoing a continuous upgrade program aimed at coping with the high luminosity performance, which will be delivered by the upgraded LHC during Run 3 and beyond (2020). One of these upgrades consists in the replacement of the first station of the ATLAS muon end-cap system, Small Wheels, during the second long shutdown (2018). The New Small Wheel will have to operate in a high background radiation region while reconstructing muon tracks with high precision as well as furnishing information for the Level-1 trigger. Work on the simulation of the New Small Wheel chambers layout, together with the validation procedures put in place, is briefly reported in Appendix C.

Chapter 1

Theoretical and Phenomenological Foundations

People like you and I,
though mortal of course like everyone else,
do not grow old no matter how long we live...
We never cease to stand like curious children
before great mystery into which we were born.

-Albert Einstein in a letter to Otto Juliusburger-

In this Chapter, a brief overview of the Standard Model of Particle Physics will be given with particular emphasis on the Soft QCD aspects, which constitute the theoretical bases of the results presented in this thesis. The Monte Carlo event generators will also be introduced with a note on validation and tuning strategies.

1.1 The Standard Model of Particle Physics

The Standard Model (SM) of Particle Physics [1–3] is built around mathematical formalisms to describe three of the four known fundamental forces and all the elementary particles associated with them. It incorporates the electromagnetic, the weak and the strong forces but it does not describe the gravitational theory. It is a renormalisable quantum field theory coherent with quantum mechanics and special relativity. The Standard Model is furthermore a gauge theory in which a particle is defined by a local quantum field. If no interaction is present, the free field is described by only two quantum numbers, the spin and the mass; if interactions are present, gauge symmetries can be used to describe them: new quantum numbers classify the type and the strength of force while new particles, force-mediator vector bosons, are used to propagate the interaction.

The gauge group of the Standard Model is SU(3)$\times$ SU(2) $\times$ U(1), whose components and fields will be described in the following.

Elementary particles with spin = 1/2 are known as fermions. According to the spin-statistics theorem, fermions respect the Pauli exclusion principle [4]. Fermions are further classified as leptons (electrons, muons and taus and their anti-particles) and quarks (up, down, charm, strange, top and bottom) and they are organized in three families with identical quantum numbers and different masses. The charged heavier components of the second and third family are unstable and decay into the lighter charged particles of the first family, which constitute most of the ordinary matter. The four fermions in each family are distinguished by their charges under strong and electromagnetic interactions: two of them are quarks, charged under strong interaction (*colour* being the quantum number of the strong interaction), and two are leptons, not charged under strong interaction. From the electromagnetic point of view, the two quarks have charge 2/3 (quark *up* considering the first family) or -1/3 (quark *down* considering the first family), while the two leptons

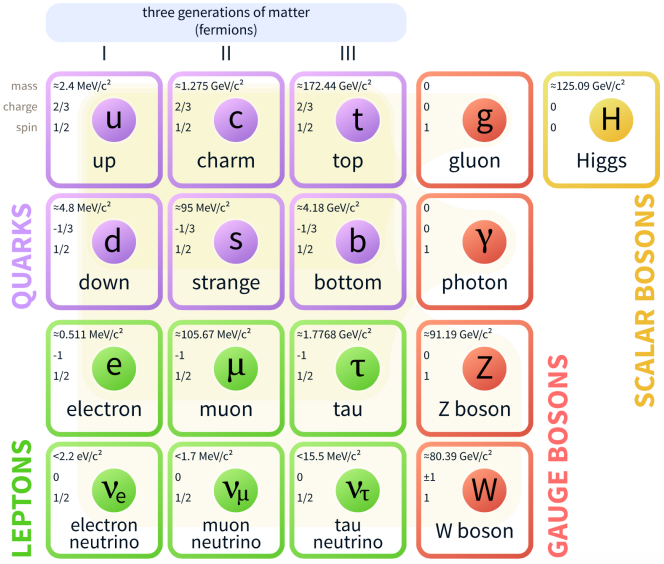


FIGURE 1.1: The building blocks of the Standard Model of Particle Physics.

have charge -1 (electron e considering the first family) or 0 (neutrino ν_e considering the first family, meaning that neutrinos are neutral under both strong and electromagnetic forces).

Interactions between particles occur by means of the exchange of particles with spin = 1, known as vector bosons: the photon (massless) mediates electromagnetic interactions, eight gluons (massless) mediate the strong interactions and the Z and W bosons (massive) carry the weak force. The Higgs boson [5–8], recently observed by the ATLAS and CMS experiments [9–12], is a massive scalar particle and it is a key building block of the Standard Model. It has no intrinsic spin, and for that reason is classified as a boson (like the gauge bosons, which have integer spin). The Higgs boson plays a unique role in the Standard Model, by explaining why the other elementary particles, except the photon and gluon, are massive. A problem in the Standard Model theory is that the symmetry of the electroweak Lagrangian cannot be exact because this would require the weak gauge bosons to be massless. The solution was the introduction of the Higgs mechanism which spontaneously breaks this symmetry by introducing an electroweak doublet of complex scalar fields. A summary of the SM fundamental particles and their properties is shown in Figure 1.1.

In the following sections, the theory of the Electroweak interactions as well as the perturbative Quantum Chromodynamics will be shortly described for completeness, while more details will be given about the soft Quantum Chromodynamics since the topic is more strongly related to this thesis.

1.1.1 Electroweak Model

In the Standard Model of Particle Physics, the Weak and the Electromagnetic interactions have been combined into a unified Electroweak Theory. Although these two forces appear very different at everyday low energies, the theory models them as two different aspects of the same force. Above the unification energy, on the order of 246 GeV [1], they would merge into a single electroweak force. Sheldon Glashow, Abdus Salam, and Steven Weinberg were awarded the 1979 Nobel Prize in Physics for their contributions to the unification of the weak and electromagnetic interaction between elementary particles.

[1] The particular number 246 GeV is taken to be the vacuum expectation value $v = (G_F\sqrt{2})^{-1/2}$ of the Higgs field (where G_F is the Fermi coupling constant).

Mathematically, the unification is accomplished under an SU(2) $\times$ U(1) gauge group, where SU(2) is the non-Abelian group used in the spin algebra, and U(1) is the Abelian group equivalent to the one used in Quantum Electrodynamics. The quantum number arising from SU(2) is the *weak isospin* (similarly to the *isospin* for strong interactions), $\vec{I}$, and the one arising from U(1) is the *hypercharge*, Y. They are related to the observed charge of real particles, Q, by the the Gell-Mann-Nishijima equation $Q = I_3 + Y/2$. The corresponding gauge bosons are the three W bosons of weak isospin from SU(2) (W_1, W_2, and W_3), and the B boson of weak hypercharge from U(1), respectively, all of which are massless. In the Standard Model, the $W^{\pm}$ and Z^0 bosons, and the photon, are the results of spontaneous symmetry breaking of the electroweak symmetry from SU(2) $\times$ $U(1)_Y$ to $U(1)_{em}$, caused by the Higgs mechanism. $U(1)_Y$ and $U(1)_{em}$ are different copies of U(1); the generator of $U(1)_{em}$ is given by Q.

The spontaneous symmetry breaking makes the W_3 and B bosons coalesce into two different bosons: the Z_0 boson and the photon (γ),

$$\begin{pmatrix} \gamma \\ Z^0 \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} B \\ W_3 \end{pmatrix}$$

where θ_W is the weak mixing angle. The axes representing the particles have essentially just been rotated, in the (W_3, B) plane, by the angle θ_W. This also introduces a mismatch between the mass of the Z_0 (M_Z) and the mass of the $W\pm$ (M_W) particles.

$$M_Z = \frac{M_W}{\cos\theta_W}.$$

The W_1 and W_2 bosons, in turn, combine to give massive charged bosons

$$W^{\pm} = \frac{1}{\sqrt{2}}(W_1 \mp iW_2).$$

The distinction between electromagnetism and the weak force arises because there is a (nontrivial) linear combination of Y and I_3 that vanishes for the Higgs boson (it is an eigenstate of both Y and I_3, so the coefficients may be taken as $-I_3$ and Y): $U(1)_{em}$ is defined to be the group generated by this linear combination, and is unbroken because it does not interact with the Higgs field.

The existence of the electroweak interactions was experimentally established in two stages, the first being the discovery of neutral currents in neutrino scattering by the Gargamelle collaboration [13] in 1973, and the second in 1983 by the UA1 [14] and the UA2 [15] collaborations that involved the discovery of the W [16, 17] and Z [18, 19] gauge bosons in proton-antiproton ($p\bar{p}$) collisions at the converted Super Proton Synchrotron, for which Carlo Rubbia and Simon van der Meer were promptly awarded the 1984 Nobel Prize in Physics. In 1999, Gerardus't Hooft and Martinus Veltman were awarded the Nobel Prize for showing that the electroweak theory is renormalisable.

1.1.2 Quantum Chromodynamics

In the SM, strong interactions of coloured quarks and gluons are described by the gauge field theory of the Quantum Chromodynamics (QCD) [20, 21]. QCD is the SU(3) component of the SU(3)$\times$ SU(2) $\times$ U(1) Standard Model of Particle Physics: strong interactions between quarks and gluons are related to the invariance of the Lagrangian under SU(3) rotations in the colour space. This invariance requires the introduction of 8 gauge fields corresponding to the 8 gluons.

Neither quarks nor gluons can be observed as free particles: colour-neutral combinations of quarks, anti-quarks, and gluons are called *hadrons* and they can be further classified into *mesons* (a quark plus and anti-quark) and *baryons* (three quarks or three anti-quarks). The hadrons' constituents are generally referred to as *partons*.

In a comparison to the electromagnetic force, it must be recognized that the strong force is very different in nature and this is reflected when trying to define a coupling constant [2] for the strong interaction: the electromagnetic force is infinite in range and obeys to the inverse square law, while the strong force involves the exchange of massive particles and it has therefore a very short range.

At subatomic distances, the strong force not only determines the binding of quarks and gluons inside hadrons, it also determines the cohesion of protons and neutrons inside atomic nuclei. Hadrons like protons and neutrons are responsible for more than 99% of the mass of all visible matter in the universe, and that mass is mainly generated by the strong binding of quarks inside hadrons, rather than by the (generally small) masses of the quarks themselves.

The restriction of the strong force to subatomic distances is related to two features called *asymptotic freedom* and *confinement*.

As it can be seen in Figure 1.2 and it can be observed by analysing the expression of the strong coupling constant [3] at leading order:

$$\alpha_s(Q) = \frac{22\pi}{(33 - 2n_f)ln((Q^2)/\lambda^2_{QCD})} \tag{1.1}$$

at high energy (or high momentum transfers or small distances, i.e. *hard processes*) $\alpha_s(Q)$ becomes smaller and it goes to zero at asymptotically high energies (leading to the term *asymptotic freedom*) and, as such, it can therefore be treated by means of perturbative QCD (pQCD) approaches. At low energy (or low momentum transfers or large distance, i.e. *soft processes*) $\alpha_s(Q)$ increases and it prevents the existence of free quarks which can then only exist in bound states.

1.2 Hadronic Processes in proton-proton collisions

Proton-proton (pp) collisions (and hadron-hadron collisions in general) are extremely complex phenomena owing to the composite nature of the incoming protons and the very complicated dynamics of the interactions between partons. The partonic content of the hadrons is energy dependent, continuously changing and, a priori, unknown: the momentum distribution of those partons is normally parameterised in the so-called parton distribution functions (PDFs), which are then extracted from experimental distributions and phenomenological extrapolations. When the two hadrons collide several partons may interact and be scattered in different directions, according to their momenta. Either the incoming or outgoing partons may radiate and all outgoing partons fragment to produce observable particles: in this chain, both perturbative and non-perturbative QCD processes are involved.

[2] A coupling constant is a quantity that determines the strength of the force exerted in an interaction. The coupling constant in itself is not a physical observable, but it is rather defined in the context of perturbation theory, which enters predictions for experimentally measurable observables.

[3] The strong coupling constant, α_s, is the only free parameter of the Lagrangian of quantum chromodynamics, if the quark masses are considered as fixed. As such, this coupling constant, or equivalently $g_s = \sqrt{4\pi\alpha_s}$, is one of the three fundamental coupling constants of the Standard Model of particle physics, together with g and g' which indicate the coupling strengths relevant for weak isospin and weak hypercharge, and can be rewritten in terms of the Weinberg mixing angle $tan\theta_W = g'/g$ and the fine-structure constant $\alpha = e^2/(4\pi)$, where the electric charge is given by $e = g \sin\theta_W$.

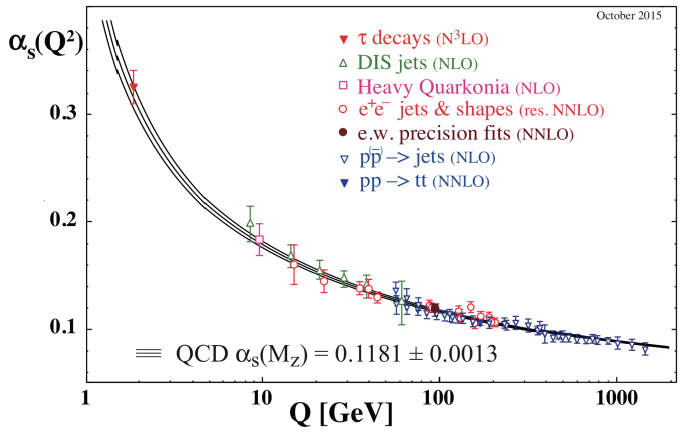


FIGURE 1.2: Summary of measurements of $\alpha_s(Q)$ as a function of the energy scale Q. The respective degree of QCD perturbation theory used in the extraction of $\alpha_s(M_{Z^0})$ is indicated in brackets (NLO: next-to-leading order; NNLO: next-to-next-to leading order; res. NNLO: NNLO matched with resummed next-to-leading logs; N3LO: next-to-NNLO). Ref. [22].

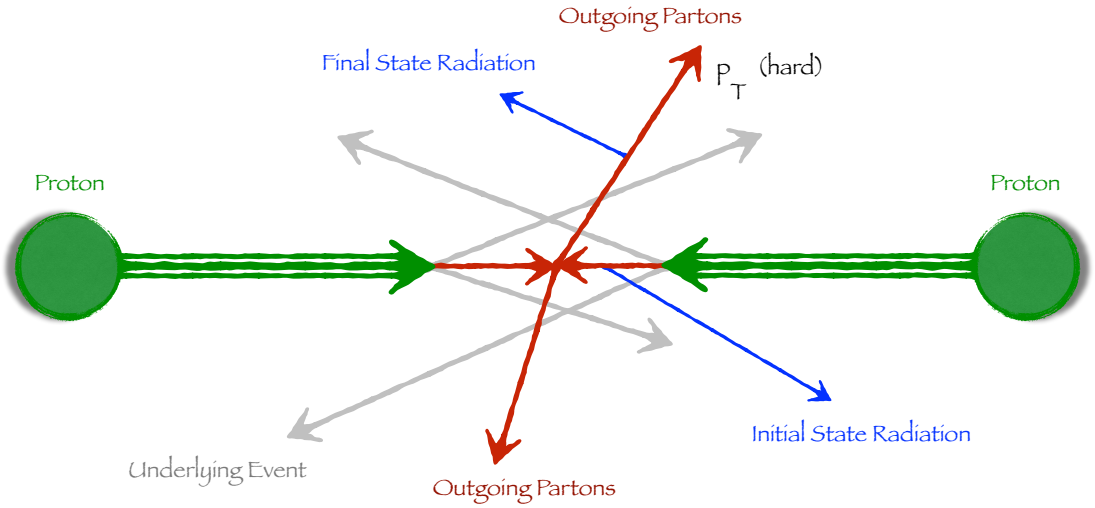


FIGURE 1.3: Pictorial llustration of the components of a hard scattering process in a proton-proton collision.

Most of the intervening processes are soft and give rise to low-momentum final state particles, but occasionally a hard parton-parton scattering occurs through a large momentum transfer. All interesting physics processes, such as Higgs or Vector Boson production, proceed via hard scatterings, while the soft components of the event, normally referred to as *underlying event* (UE), introduce an additional level of complexity.

Thanks to the QCD factorization theorem (Section 1.2.1 and Ref. [23]), physicists are able to separate the cross-section of hadron-hadron collisions into two independent terms: the process-dependent, perturbative QCD calculable short-distance parton cross-section, and the universal long-distance functions which can be measured with global fits to experiments. This functions include the PDFs, fragmentation functions, multi-parton correlation functions, generalized parton distributions, generalized distribution amplitudes and many kinds of form factors.

1.2.1 Hard QCD

Asymptotic freedom, which characterises the QCD at high energies, allows to compute many features of the hard subprocesses via perturbation theory by using, for instance, Feynman diagrams.

The factorization formula [24] to compute the cross-section for any process $ab \to n$ at a hadron collider is the following:

$$\sigma = \sum_{a,b} \int_0^1 dx_a dx_b \int f_a^{h_1}(x_a, \mu_F) f_b^{h_2}(x_b, \mu_F) d\hat{\sigma}_{ab\to n}(\mu_F, \mu_R) = \tag{1.2}$$

$$= \sum_{a,b} \int_0^1 dx_a dx_b \int d\Phi_n f_a^{h_1}(x_a, \mu_F) f_b^{h_2}(x_b, \mu_F) \times \frac{1}{2\hat{s}} |M_{ab\to n}|^2(\Phi_n; \mu_F, \mu_R).$$

In this equation, $f_{a,b}^{h_{1,2}}(x, \mu_F)$ are the PDFs, which depend on the factorization scale μ_F and on the momentum fraction $x_{a(b)}$ of a parton $a(b)$ with respect to its parent hadron h; $d\hat{\sigma}_{ab\to n}(\mu_F, \mu_R)$ is the parton-level cross-section for the production of the final state n through the initial partons a and b which depends on the momenta given by the final-state phase space Φ_n ($d\Phi_n$ is the differential phase-space element over the n final state particles), on the renormalization μ_R and on the factorization μ_F scales. To obtain the fully differential parton-level cross-section it is necessary to take the product of the corresponding matrix element squared $|M_{ab\to n}|^2$ and the parton flux $\frac{1}{2\hat{s}}$ with $\hat{s}$ being the hadronic centre-of-mass energy squared. The matrix element can be written as the sum of all Feynman diagrams:

$$M_{ab\to n} = \sum_i F_{ab\to n}^{(i)}. \tag{1.3}$$

Any summation over quantum numbers can be moved outside the square, allowing one to sum over helicity and colour orderings in such a way that in the computation of the cross-section-section, one can sample not only over phase-space but also on helicities and colour configurations and the choice determines the starting conditions for the subsequent parton showering. The integration over the phase-space may contain cuts, which reflect the acceptance of the detector and also physical necessities, such as a cut on the transverse momentum of particles produced in the t-channel [25].

For a given PDF set and a choice of unphysical factorization and renormalization scales (which are not based on first principles but the result of a careful phenomenological procedure), the cross-section calculated by Eq. 1.2 is fully specified.

1.2.2 Soft QCD

The total cross-section is a measure of the probability that a pair of hadrons undergo any interaction. As such, it is the sum of the cross-sections of all possible proton-proton interaction. Hard QCD events, as described above, constitute only a tiny fraction of the total cross-section, which is then dominated by events in which a relatively small momentum transfer occurs (peripheral processes). This is shown in Figure 1.4, in which σ_{tot} is orders of magnitude larger than very abundant hard QCD processes such as the production of b-quarks [26].

Soft interactions can be either elastic or inelastic. In elastic scattering both protons emerge intact and no additional particles are produced. Inelastic scattering occurs when the interaction causes at least one of the incoming protons to dissociate, i.e., the outgoing particles differ from the incoming particles. The first theory of the soft interactions was formalised by Isaak Pomeranchuck [28], who predicted that if the cross-section scales logarithmically with the centre-of-mass energy, the cross-sections of proton-proton and

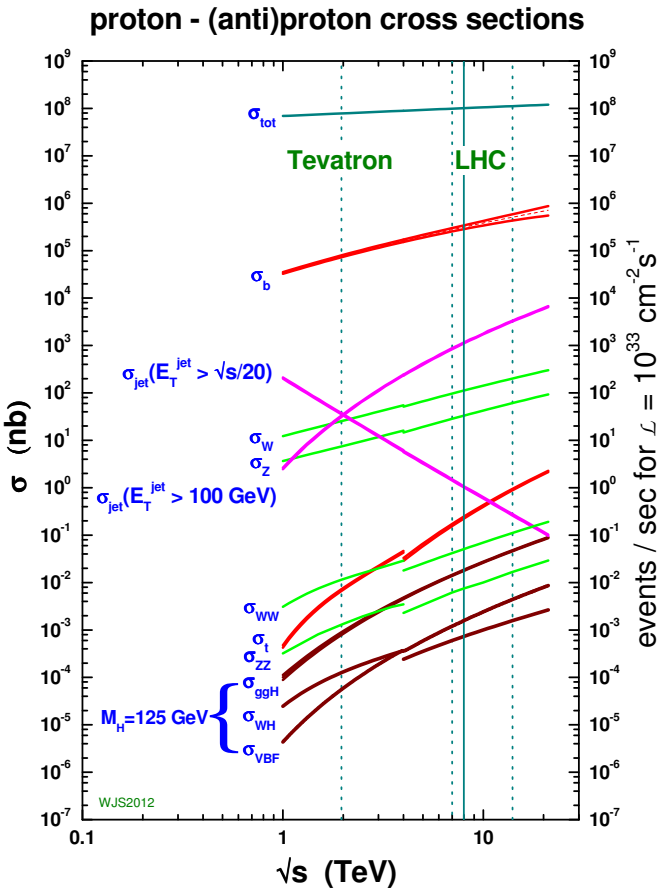


FIGURE 1.4: Standard Model cross-sections as a function of the collider energy. Cross-sections are calculated either at NLO or NNLO pQCD, using MSTW2008 [27] (NLO or NNLO) parton distributions, with the exception of the total hadronic cross-section which is based on a parametrisation of the Particle Data Group. The discontinuity in some of the cross-sections at 4 TeV is due to the switch from proton-antiproton to proton-proton collisions at that energy. Ref. [26].

proton-antiproton should become equal at large energies, as shown in Figure 1.5. Gribov then introduced the concept of the exchange of a Regge trajectory ensuring such behaviour [29–32]. The particles on the Regge trajectory are virtual and have the quantum numbers of the vacuum. The concept of pomeron was, thus, introduced to be the sum of all particles on a Regge trajectory: in the context of QCD, it can be regarded as a colourless and flavourless combination of gluons.

Inelastic interactions can be classified in three different categories. If colour charge is exchanged, the interactions are referred to as non-diffractive (ND). In single diffraction (SD) a single pomeron is exchanged between the two protons and one of the incoming protons forms a diffractive system. In double diffraction (DD) a single pomeron is also exchanged, but both of the protons form diffractive systems. The exchange of two pomerons is referred to as central diffraction (CD). Non-diffractive interactions constitute the inelastic process with the largest cross-section. The various components have different multiplicities and topologies. The exchange of colour charge in non-diffractive events results in

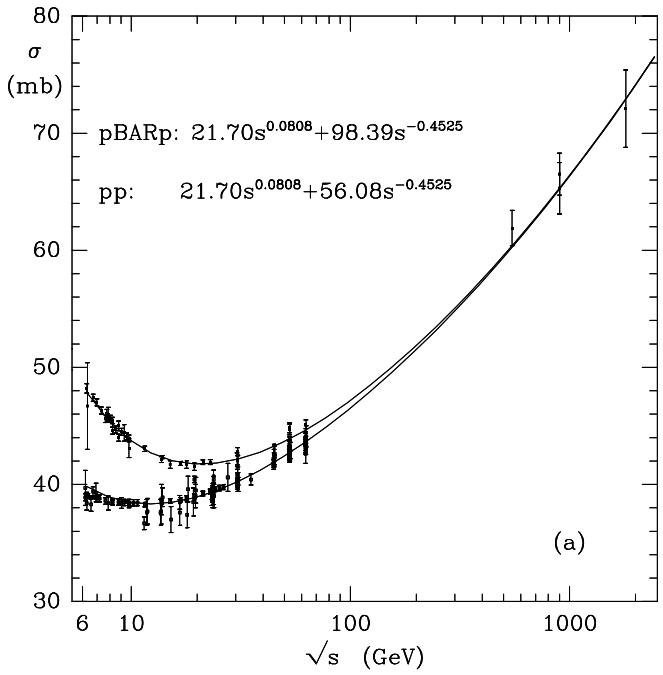


FIGURE 1.5: Measurements of the total cross-section in pp and $p\bar{p}$ scattering as a function of the centre-of-mass energy. The curves show the results of a two component fit, in which the first term describes pomeron exchange and the second describes Reggion exchange. Ref. [33].

a large number of particles being produced at central rapidity [4]. However, the multiplicity falls sharply at forward rapidity. In a single-diffractive event, the proton which breaks up produces particles at high rapidity. The other incoming particle is essentially undisturbed and has the rapidity of the beam. In a double-diffractive event particles are produced symmetrically at positive and negative forward rapidity, with fewer particles produced in the central region. Both non-diffractive and double-diffractive interactions are symmetric with respect to the beam line.

1.3 Monte Carlo Event Generators

The development of specialised software libraries based on Monte Carlo Methods, Monte Carlo (MC) event generators, to describe phenomenologically particle interactions began shortly after the discovery of the partonic structure of hadrons and the formalisation of QCD as the theory of strong interactions. For hard processes, the main event can be explained in terms of parton interactions. The primary partons can be directly fragmented into hadrons to describe final states, but this does not account for the transverse broadening of jets and lepton pair distributions when going to hard interactions. It was understood that the primary partons, being coloured, can emit gluons as in the case of electron Bremsstrahlung, and that these gluons, unlike the Bremsstrahlung photons, can themselves radiate, leading to a parton cascade or shower that accounts for the broadening. It is clear that hadron formation would occur naturally as the end point of parton showering, when transferred momentum is low and the corresponding value of the QCD running coupling constant is large. The hadronisation process then becomes non-perturbative, hence

[4] The rapidity of a particle is defined as $y = \frac{1}{2}\ln(\frac{E+p_L}{E-p_L})$ with E and p_L being, respectively, the energy and the component of momentum of the particle along the beam axis.

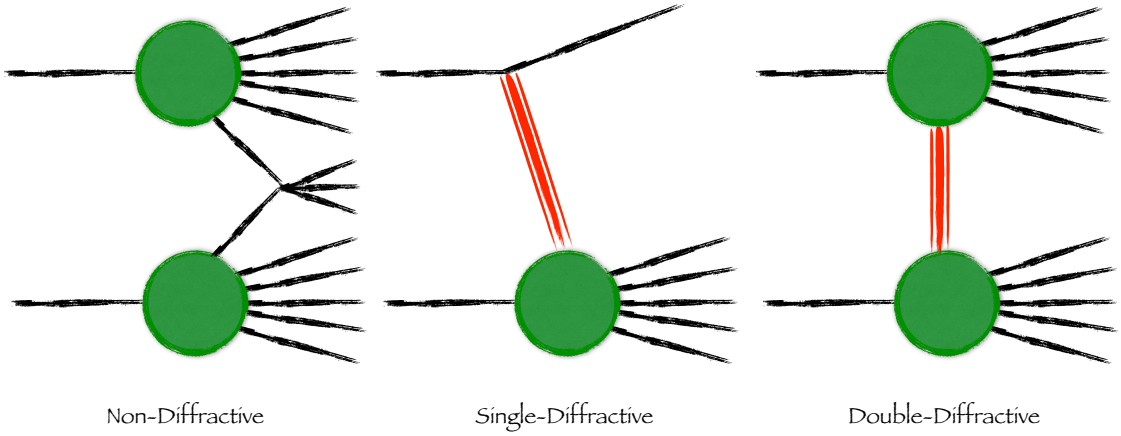


FIGURE 1.6: Inelastic proton-proton collisions classified in three different categories. Non-diffractive interactions occur when colour charge is exchanged between the two protons. In single (double)-diffraction a pomeron is exchanged between the two protons and one (two) of the protons forms a diffractive system, which dissociates.

hadronisation models, inspired by QCD but not derivable from it, had to be developed with a set of tunable parameters to describe the hadron-level properties of final states.

The signal processes of interest are often part of the hard interaction category and thus can be treated with perturbation models, but, as shown in figure 1.4, the majority of collisions are soft, leading to diffractive scattering or multi-particle production with low transverse momenta. These soft processes need to be simulated but, as in the case of hadronisation, their non-perturbative nature means that models with tunable parameters to describe the data are needed. A related phenomenon is the underlying event, the component of the final state in hard interactions that is not associated with the primary hard-process. The hard components of these interactions are described by perturbative QCD, but again the soft component has to be modelled. The same multiple-parton interaction model can serve for the simulation of soft collisions, provided that there is no conflict between the parameter values needed to describe the two phenomena.

As described in details in Ref. [22, 24] and in the references therein, many stages are needed when simulating a MC event. These steps are summarised below, while a pictorial representation can be seen in Figure 1.7 (Ref. [34]) and a scheme is shown in fig 1.8:

- Random generation of kinematics and partonic channels of a hard scattering process;
- Parton shower, usually based on the successive random generation of gluon emissions (lower energy scale with respect to previous emissions following a perturbative QCD distribution, which depends on the momenta of all previous emissions);
- Parton showering stops at a scale of ~ 1 GeV, hadronisation models are then used to convert the resulting partons into hadrons:

  stretching a color string across quarks and gluons and breaking it up into hadrons;

  breaking each gluon into $q\bar{q}$ pair and subsequently grouping quarks and antiquarks into colorless clusters and eventually hadrons;
- In the case of pp collisions, it is also needed to describe the collision between the two hadron remnants by generating the underlying event, usually implemented via additional $2 \to 2$ scattering (multi-parton interaction at a few GeV scale);

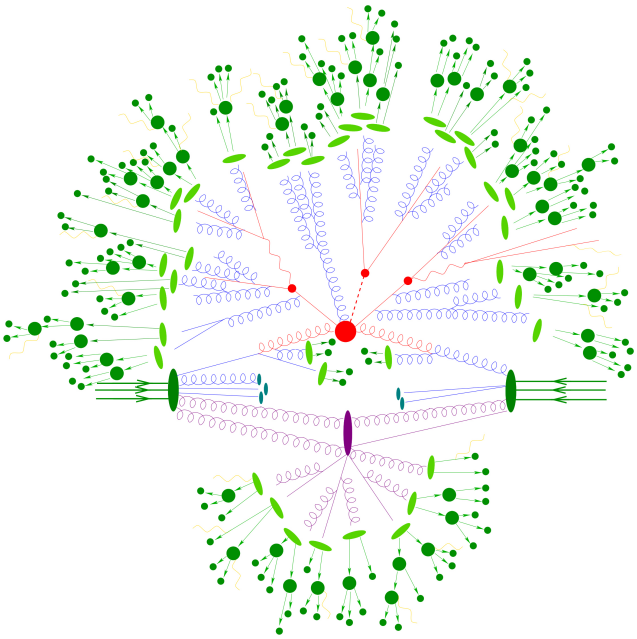


FIGURE 1.7: Pictorial representation of a hadron-hadron collision as simulated by a Monte-Carlo event generator for a $t\bar{t}h$ event [34]. The red dot in the center represents the hard collision, surrounded by a tree-like structure representing gluon Bremsstrahlung as simulated by parton showers. The purple dot indicates a secondary hard scattering event. Parton-to-hadron transitions are represented by light green dots, dark green dots indicate hadron decays, while yellow lines signal soft photon radiation.

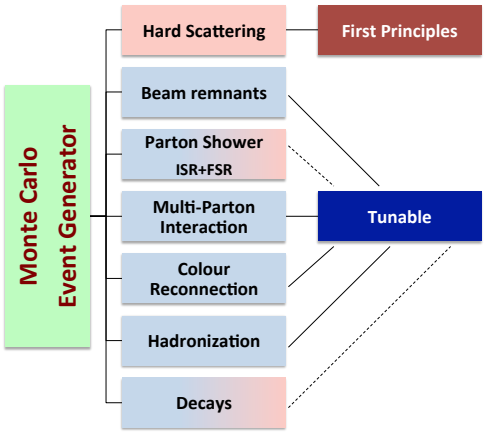


FIGURE 1.8: The different steps in a Monte Carlo Event Generator: the hard scattering, completely based on first principles; the parton shower and the decays, which are partly based on first principles; the beam remnants, multi-parton interactions, colour reconnection and hadronisation, which are mainly based on tunable parameters.

- Finally, many of the hadrons that are produced during the hadronisation step are unstable resonances. Sophisticated models are used to simulate their decay to the lighter hadrons that are long-lived enough to be considered stable on the time-scales of particle detectors;

- Since many of the particles involved at all stages of the simulation are charged, QED radiation effects can also be activated in the event chain at various stages.

Not all these steps are relevant in all processes and, in fact, the majority of events that constitute the total hadron-hadron cross-section are of soft QCD type and rely more on phenomenological models, while signal events and their SM backgrounds rely basically on all the components.

The Leading-Order (LO) multi-parton matrix elements are sometimes used to generate hard high-multiplicity partonic configurations as additional starting points for the showering (supplemented with some prescription for consistently merging samples with different initial multiplicities) because the soft and collinear approximations that underlie parton showers may fail to reproduce the full pattern at wide angle emissions which are important in new physics searches.

MC event generators provide LO cross-sections for the hard process, but a variety of processes are also available at Next to Leading Order (NLO). When NLO predictions are available, the ratio of NLO and LO prediction (referred to as *k-factor*) is taken into account in complex processes involving, for instance, jet observables associated with a vector boson or high-p_{T} jets containing B-hadrons.

1.3.1 Monte Carlo Validation and Tuning

The results from a Monte Carlo event generator have to be compared with all relevant data that it is supposed to describe. This process is known as *validation*. It is essential for the validation to be global: in order to have a predictive power, the model should describe the underlying physics and not just parameterise the data. In this sense, validation is important to develop models as well as for debugging both code and physics models. The procedure in which the free parameters of the model are adjusted within their allowed ranges to improve the description of the relevant data is called *tuning* and the set of parameters is defined *tune*. A set of observables is required for predictive power and to come up with a generally usable single set of parameters. Depending on the suitability of observables to the model being tuned, it may or may not be possible to describe all data simultaneously. The optimization of MC parameters to the chosen observables may be performed manually, according to the expected physical behaviour of the models, or by a more automated method driven by the quality of the fit to data. In any case, some sampling of the parameter space is typical to observe the generator behaviour in response to parameter changes. The parameter values typically span a range in which the underlying physical picture is valid, but scans of more restricted ranges are used to produce a final tune. Since all simulations lack some known physics effects, the choice of reference data is important in order to have a generator tune which optimises phenomenological simulation aspects and do not make up aspects which are described by calculable QCD. The parameter space of a generic event-generator is usually too big for a complete investigation, even with automated methods. In fact, a solution is the factorization of the parameters into minimal sets, which are suitable for each group of observables. For this reason, the tuning of a Monte Carlo event generator is done in steps, as explained in detail in Ref. [24].

To summarise:

- Hadronisation and final-state fragmentation: flavour and kinematic structures of the hadronisation and final-state shower mechanisms are assumed to be universal between hadron and e^+e^- collisions. The e^+e^- observables, such as event shapes, identified particle rates and transverse momentum spectra, may be described by a generator without first requiring the tuning of initial-state hadron collider effects. Thus, observables are used to tune final-state shower and hadronisation parameters. Typical parameters are α_s/λ_{QCD}, the InfraRed-Cutoff for the final state showering, the string tension and fragmentation function parameters for string hadronisation models, and the gluon constituent mass and cluster momentum smearing in cluster hadronisation models. The kinematics of light and heavy quark fragmentation are often treated separately. This permits further factorization to charm- and bottom-specific observables without compromising the statistically dominant light fragmentation.

- Initial-state parton shower: after a reasonable tune of the final-state parameters is obtained, the initial-state parton shower parameters are tuned. This step is performed before tuning the soft QCD effects because the goal is to tune the shower to observables with little Multi-Parton Interaction (MPI) and beam-remnant contamination. Then, full flexibility of the heavily-parameterized MPI mechanism can be used to perform the final best fit to data. Some typical observables for initial-state shower tuning are di-jet azimuthal decorrelations and hadron collider jet shapes. Typical parameters are the shower α_s/λ_{QCD}, the shower InfraRed-Cutoff and the scaling factor for the α_s evaluation scale and the starting scale for the parton cascade.

- MPI and beam remnant effects: MPI modelling is the aspect of Monte Carlo modelling which is least constrained by QCD calculation and it is left as the final stage of the tuning process. Many parameters are related to MPI, with the most important ones being the minimum transverse momentum $p_\perp$ cutoff/regulator for perturbative $2 \to 2$ scattering, the parameterization of the scaling of this cutoff with collision energy, the hadronic matter distribution/overlap and any parameters relating to colour reconnection of either strings or clusters. The primordial $k_\perp$ width is often considered as part of this tuning step, as it may affect soft QCD observables. As MPI models generate multiple scattering from low-x gluons extracted from the beam-remnants, they are affected by the choice of PDF and for this reason each PDF requires a specific MPI tune.

A relevant aspect to be noticed is that the observables for MPI tuning are drawn from events characterised by very low transferred momentum (known as *minimum bias* events from the fact that a minimal set of trigger conditions are required at data taking) and from the underlying event data from as many hadron colliders as possible. Since a key feature of soft QCD modelling is the scaling of MPI activity with the collider centre-of-mass energy, having access to a broad range of collision energies is desirable. Current focus is on the Large Hadron Collider and the experimental tunings place emphasis on it with the goal of best describing the soft QCD backgrounds to hard-process simulations. Thus, the work reported in this thesis, which concentrates on measurements of charged-particle distributions, is of big importance for Monte Carlo developers because it allows to improve the MPI tuning of Monte Carlo event generators.

The tools commonly used in the generator validation and tuning processes are the Rivet [35] package for MC validation and the Professor [36] system for tuning. Their strength is in systematically verifying event simulations and optimizing their parameters, where required and physically sensible.

Rivet is considered a validation tool because it does not produce the tuning, but it provides a standard set of analyses by which to verify the accuracy of a given generator

with a given tuning. Rivet analyses are based on a comprehensive set of computational tools called *projections*, which perform standard computations such as jet algorithms, event shape observables, and a variety of other common tasks.

Then, the Professor tool builds up the output of these analyses by performing the real tune, i.e. by optimising the generator parameters to best match the generator predictions to the reference data. The process of generator tuning can be seen as the optimization of a very complicated function with many parameters. The volume of the parameter space increases exponentially with the number of parameters and the process is so CPU-intense even for a single evaluation of the function that any attempt to scan the parameter space will fail for more than a few parameters. Thus, Professor parameterises the function based on a non-exhaustive scan of the space: this is an approximation, but its accuracy is systematically verifiable. The MC parameterization is generated by independently fitting a function to each of the observable bin values, approximating how they vary in response to changes in the parameter vector.

Different applications place different emphases on different observables, e.g. LHC vs. Tevatron data or underlying event vs. soft interactions data. Once a set of observable has been chosen, it is a matter of logistics to create equivalent tunes for different PDFs: this permits a more accurate measure of the systematic effect of the PDF choice than was previously possible. In the particular case of MPI model tuning to soft QCD observables, this approach has shown that much of the effect of PDF changes can be absorbed into typical MPI model parameter choices.

Professor is not only used for MC tuning, but also for studies about extra-solar planets and supersymmetric model phenomenology.

With the LHC operations entering maturity, the demand will be more focused on systematic uncertainty evaluation, than on tuning, and the Professor tool will still be suitable for construction of tune-error estimates.

1.3.2 Monte Carlo Event Generator used in this Thesis

Inclusive charged-particle measurements in proton-proton (pp) collisions (presented in details in the main body of this thesis) provide insight into the strong interaction in the low-energy non-perturbative QCD region. Particle interactions in soft processes are typically described by QCD-inspired models implemented in Monte Carlo event generators with free parameters, which can be constrained by such measurements. An accurate description of low-energy strong interaction processes is essential for simulating single pp interactions as well as the effects of multiple pp interactions at high instantaneous luminosities in hadron colliders.

The PYTHIA 8 [37] event generator is used as a default to perform the track reconstruction studies described in Chapter 3. For the Inner Detector Material Measurement described in Chapter 4 both PYTHIA 8 and EPOS [38] generators are used. In the analysis of Charged-Particle distributions presented in in Chapter 5 also the HERWIG++ [39] (only shown in preliminary distributions), and QGSJET-II [40] event generators are used. Few details about these Monte Carlo generators will be given in the following to help the reader to gain a general overview, while specific comments, when relevant, will be found in each analysis section.

- In PYTHIA 8 [5], inclusive hadron-hadron interactions are described by a model which splits the total inelastic cross-section into non-diffractive and diffractive processes. The non-diffractive part is dominated by t-channel gluon exchange. Its simulation includes multiple parton-parton interactions. The diffractive part involves a color-singlet exchange. It is further divided into single-diffractive dissociation, in which

[5] PYTHIA version 8.18x

one of the initial hadrons remains intact and the other is diffractively excited and dissociates, and double-diffractive dissociation in which both hadrons dissociate. The simulated sample contains ∼22% SD and ∼12% DD processes, with the remaining ∼ 66% being ND events. To reproduce experimental data, the ATLAS minimum-bias tune A2 [41] is used, which is based on the MSTW2008LO PDF [27]. It provides a good description of minimum bias events and of the transverse energy flow data, a calorimeter-based minimum bias analysis performed with $\sqrt{s} = 7$ TeV data [42]. An alternative tune, Monash [43], is used for comparison. It uses updated fragmentation parameters compared to A2 and minimum bias, Drell-Yan, and underlying event data from the LHC to constrain ISR and MPI parameters. In addition, it uses SPS and Tevatron data to constrain the centre-of-mass energy scaling. It uses the NNPDF2.3LO PDF [44]. This tune gives an excellent description of the 7 TeV minimum bias p_{T} spectrum.

- In HERWIG++ , inclusive hadron-hadron interactions are simulated by applying a MPI model for the non-diffractive processes to events without hard scattering. It is therefore possible to generate an event with no 2→2 partonic scattering, in which only beam remnants are produced with no additional activity. These types of events look similar to double-diffractive dissociation, even though HERWIG++ does not have any explicit model for diffractive processes. Version 2.7.1 is used with a 7 TeV underlying event tune, UE-EE-5-CTEQ6 [45] with CTEQ6L1 PDF [46, 47], which utilises colour reconnection and energy dependent MPI minimum p_{T} cutoff. It provides simultaneously a good description of both the underlying event and the double parton scattering data by tuning the deep parton interaction (DPI) effective cross-section.

- EPOS stands for *Energy conserving quantum mechanical approach, based on Partons, parton ladders, strings, Off-shell remnants, and Splitting of parton ladders*. The latest version 3.4 is used, which is equivalent to 1.99 version with the so called LHC tune [48]. It provides an implementation of a parton-based Gribov-Regge theory [49], which is an effective QCD-inspired field theory describing the hard and soft scattering simultaneously. Hence, the calculations do not rely on the standard parton distribution functions (PDFs) as used in generators like PYTHIA 8 and HERWIG++ .

- QGSJET-II offers a phenomenological treatment of hadronic and nuclear collisions at high energies, being developed in the Reggeon Field Theory framework [50]. The soft and semi hard parton processes are included in the model within the semi-hard pomeron approach. Nonlinear interaction effects are treated by means of pomeron-pomeron interaction diagrams. The latest model version comprises three important updates: treatment of all significant enhanced diagram contributions to the underlying dynamics, including pomeron loops, re-calibration of the model with new LHC data, and improved treatment of charge exchange processes in pion-proton and pion-nucleus collisions.

A summary of these Monte Carlo event generators and tunes is given in Table 1.1.

Generator	Version	Tune	PDF	Focus	Data	From
PYTHIA 8	8.185	A2	MSTW2008LO	MB	LHC	ATLAS [41]
PYTHIA 8	8.186	Monash	NNPDF23LO	MB/UE	LHC	Authors [43]
HERWIG++	2.7.1	UEEE5	CTEQ6L1	UE	LHC	Authors [45]
EPOS	3.4	LHC	-	MB	LHC	Authors [48]
QGSJET-II	II-04	LHC	-	MB	LHC	Authors [40]

TABLE 1.1: Details of the Monte Carlo models used. It should be noted that the tunes use data from different experiments for constraining different processes, but only the data which had the most emphasis on each specific tune are shown. Here "LHC" indicates data taken at $\sqrt{s} = 7$ TeV, although $\sqrt{s} = 900$ GeV data were also included in ATLAS tunes, with much smaller emphasis. Some tunes are focused on describing the minimum-bias distributions better, while the rest are tuned to describe the underlying event distributions, as indicated. For the specific measurements and observables used in each tune, please refer to the references reported in the last column.

Chapter 2

The Large Hadron Collider and the ATLAS Experiment

I'm burnin' through the sky yeah
Two hundred degrees
That's why they call me Mister Fahrenheit
I'm trav'ling at the speed of light
I wanna make a supersonic man out of you
Don't stop me now
I'm having such a good time
I'm having a ball
Don't stop me now...

-Queen-

In this Chapter, the Large Hadron Collider will be introduced, with a brief report of the performance reached during Run 1 and during the first Run 2 period. The ATLAS experiment will also be presented together with an overview of its subdetectors.

2.1 Introduction

The Large Hadron Collider (LHC) [51] is the world's largest collider. It is located at the Conseil Européen pour la Recherche Nucléaire (CERN), across the France-Switzerland border, as shown in Figure 2.1. Two multi-purpose detectors, ATLAS (A Toroidal LHC ApparatuS) [52] and CMS (Compact Muon Solenoid) [53] were designed and built for a complete probe of proton-proton and ion-ion collisions, while two specific experiments LHCb (Large Hadron Collider beauty) [54] and ALICE (A Large Ion Collider Experiment) [55] are focused respectively on studies of the b-quark and of the heavy ion collisions. The LHC offers an unprecedented physics potential, from precise measurements of Standard Model parameters as well as the Higgs boson properties to the search for new physics phenomena. Furthermore, nucleus-nucleus collisions at the LHC provide a unique opportunity to study the properties of strongly interacting matter at extreme energy density, including the possible phase transition to a colour-deconfined state, referred to as quark-gluon plasma.

2.2 CERN Accelerator Complex

The Large Hadron Collider is only the last element of the big CERN accelerator complex, shown in Figure 2.2.

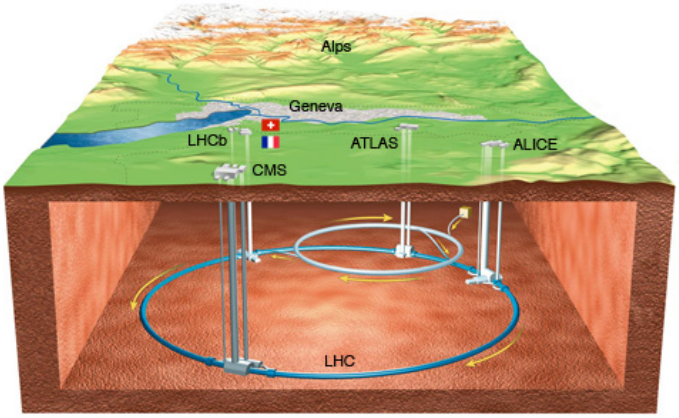


FIGURE 2.1: CERN Map showing the LHC and the four main experiments: ATLAS, CMS, LHCb and ALICE.

The first element is a bottle of hydrogen gas which constitutes the proton source. The electron are stripped away from the hydrogen atoms to obtain the protons [1]. The protons are pulsed from the hydrogen bottle for up to 100 microseconds per pulse. Pulses are repeated again and again until enough protons are produced.

The hydrogen bottle is placed at one end of the Linear Accelerator 2 (Linac 2 [2]) which uses radio-frequency cavities to charge cylindrical conductors. The protons pass through the conductors, which are alternately positively or negatively charged. The conductors behind them push the particles and the conductors ahead of them pull, causing the particles to accelerate. Small quadrupole magnets ensure that the protons are arranged in a tight beam. At the end of the Linac 2, the protons have an energy of 50 MeV and have reached 31.4 % of the speed of light and gained 5% of mass.

With these features they enter the Proton Synchrotron Booster which is made up of four superimposed synchrotron rings which accelerate protons up to 1.4 GeV ($\sim 91.6\%$ of the speed of light) for injection into the Proton Synchrotron (PS). The Booster became operational in 1972. Before that, the protons were injected directly from the Linac to the PS, but the number of protons to be injected in the PS was limited by the maximum energy of 50 MeV achieved by the Linac. With the higher energy at the end of the booster, the PS can be fed with up to 100 times more protons and this makes the beam much more suitable for the experiments.

The PS is one of the key elements of the accelerator complex, as it accelerates not only protons from the booster but also ions for the Low Energy Ion Ring (LEIR). The PS began operations during 1959. It is 628 m long and it is made of 277 conventional (room-temperature) electromagnets, including 100 dipoles to bend the beams round the ring. The accelerator operates up to 25 GeV (99.93% of the speed of light), when the protons are injected into the Super Proton Syncrotron (SPS) which acts as a link between the PS and the LHC.

The SPS was switched on in 1976 and a major highlight came in 1983 with the Nobel-prize-winning discovery of W and Z particles, with the SPS running as a proton-antiproton

[1] With the total mass of the protons calculated at rest, even though the proton beams are very intense, only about 2 ng of hydrogen are accelerated every day. The LHC would need around 1 million years to accelerate 1 g of hydrogen!

[2] Linac 2 started up in 1978, when it replaced Linac 1. It was originally built to allow higher intensity beams for the accelerators in CERN's accelerator complex. Linac 2 will be replaced by Linac 4 in 2020. After Linac 2 was built, Linac 1 was used to deliver ions for experiments at the Super Proton Synchrotron. There was soon pressure to provide heavier ions to study quark-gluon plasma, so the dedicated Linac 3 was built for that scope and expected to be operational until 2022.

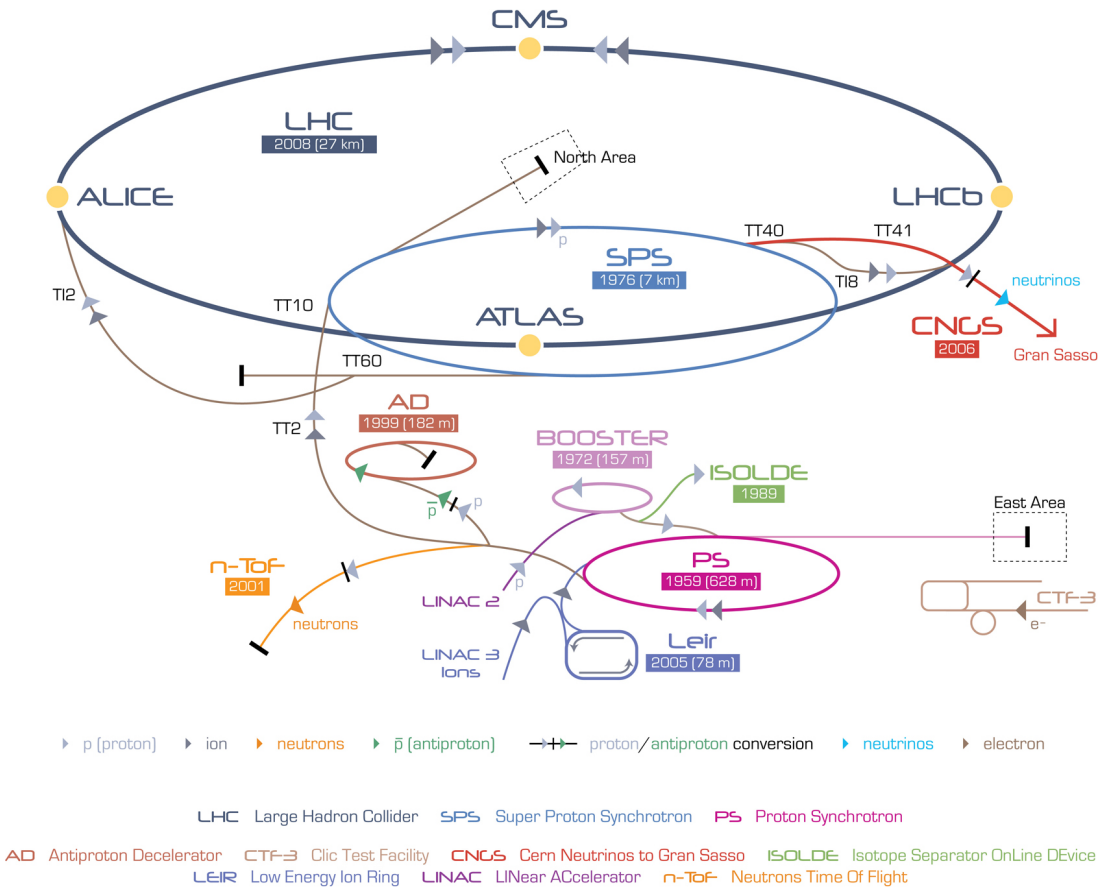


FIGURE 2.2: The CERN accelerator complex.

collider. The SPS has a circumference of 7 km and it accelerates protons up to 450 GeV (99.98% of the speed of light). In the case of the SPS, 1317 conventional electromagnets, including 744 dipoles to bend the beams around the ring, are used.

The LHC, the most powerful and largest accelerator in the world, closes the chain receiving the proton bunches from the SPS. It uses superconducting magnets with a number of accelerating structures to boost the energy of the particles along the 27 km ring. Two high-energy particle beams are accelerated to an energy of 7 TeV (nominal value) which means 99.9999991% the speed of light before they are made to collide. The beams travel in opposite directions in separate beam pipes (two tubes kept at ultra-high vacuum) and they are guided around the accelerator ring by a strong magnetic field maintained by superconducting electromagnets. The electromagnets are built from coils of special electric cable that operates in a superconducting state (efficiently conducting electricity without resistance or loss of energy). This requires the temperature of the magnets, -271.3 C, to be lower than in the outer space. To keep this temperature, the accelerator is connected to a distribution system of liquid helium, which cools the magnets and other supply services.

2.3 The Large Hadron Collider

The Large Hadron Collider [51] was built between 1998 and 2008 and installed in the 27 km long tunnel previously housing the Large Electron Positron Collider (LEP). The main objective of such an ambitious project was the discovery of the Higgs boson, which was eventually announced in 2012 [9–12], and the study of rare events at centre-of-mass energies of up to 14 TeV. The number of events per second, N_{event}, generated in the LHC

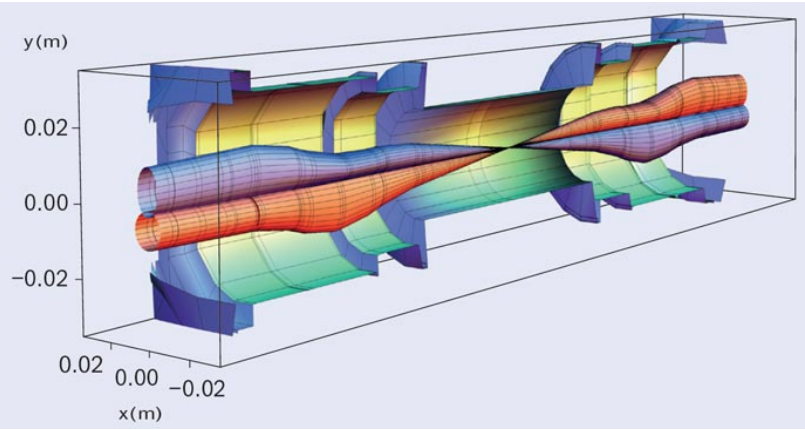


FIGURE 2.3: Beam envelopes in the interaction region around point 1 (ATLAS) showing how the beam sizes are reduced in the last 60 m on each side of the interaction point following the squeeze. Note the different transverse scale: the radius of the cut-away beam pipe is just 18 mm at the collision point. The clockwise beam is in blue and the anti-clockwise beam is in red. Picture taken from [56].

collisions is given by $N_{event} = L \cdot \sigma_{event}$, where σ_{event} is the cross-section of the process under study and L the machine luminosity, which depends on the beam parameters and which, for a Gaussian beam distribution, can be written as:

$$L = \frac{N_b^2 n_b f_{rev} \gamma_r}{4\pi \epsilon_n \beta^*} F \tag{2.1}$$

where N_b is the number of particles per bunch, n_b is the number of bunches per beam, f_{rev} is the revolution frequency, γ_r is the relativistic factor.

F is the geometric luminosity reduction factor due to the crossing angle at the interaction point which, assuming equal beam parameters for both circulating beam, can be written as follows:

$$F = 1/\sqrt{1 + \left(\frac{\theta_c \sigma_z}{2\sigma^*}\right)^2} \tag{2.2}$$

where θ_c is the full crossing angle at the IP, σ_z the RMS bunch length and σ^* the transverse RMS beam size at the IP.

The normalized transverse beam emittance, ϵ_n, is a property of a particle beam within an accelerator defined as $\epsilon_n = \beta\gamma_r\epsilon$ where β and γ_r are the relativistic functions and ϵ is the emittance. The emittance measures the average spread of the particle coordinates in position-and-momentum phase space. As a particle beam propagates along the accelerator, the position spread may change, but in a way that does not change the emittance, if radiation has a negligible effect, as in the case of heavy particles like protons. In a colliding beam accelerator, keeping the emittance small means that the likelihood of particle interactions will be greater resulting in higher luminosity values. The emittance is inversely proportional to the beam momentum: increasing the momentum of the beam reduces the emittance and hence the physical size of the beam. This reduction is called adiabatic damping. The normalized emittance does not change as a function of energy and so it can be used to gauge beam degradation if the particles are accelerated. If β is close to one then the emittance is approximately inversely proportional to the energy and so the physical width of the beam varies inversely to the square root of the energy.

β^* is the value of the β function at an interaction point. The β function is related to the transverse size of the particle beam at the location s along the nominal beam trajectory. Assuming a Gaussian shape in the beam transverse direction, it can be expressed as $\sigma(s) = \sqrt{\epsilon \cdot \beta(s)}$ with $\sigma(s)$ being the width of the Gaussian. The β function is typically adjusted to have a local minimum at the interaction points, as can be seen in Figure 2.3, in order to minimize the beam size and thus maximise the interaction rate. The smaller the beam size at the interaction point, the faster the rise of the β function (and thus the beam size) farther away from the interaction point. The nominal LHC triplet magnet design is just compatible with $\beta^* = 0.5$ m.

A commonly used quantity to indicate the amount of data delivered by the accelerator and collected by the experiments, thus to measure their performance, is the integrated luminosity, L_{int}, defined as the integral of L over a certain time, $L_{int} = \int L dt$.

The search for rare events at the LHC requires both high collision energies and high beam intensities.

Contrarily to previous hadron accelerators, which collided protons and anti-protons in a common vacuum and magnet system for both circulating beams, the use of anti-protons for a high beam intensity machine like the LHC, is made unpractical by the extremely inefficient anti-proton production mechanism and the need of accumulating these anti-particles. Two separate proton beams have to be used and to steer two beams of same-sign charged-particles requires opposite magnet dipole fields for both beams. The LHC is therefore designed as a proton-proton collider with separate magnet fields and vacuum chambers in the main arcs and with common sections only at the intersection regions where the experiments are located. The two beams share an approximately 130 m long common beam pipe along the interaction regions (IR).

Together with the large number of bunches (2808 for each proton beam), and a nominal bunch spacing of 25 ns, the long straight section and the common beam pipe introduce 34 (a total of 136 for the four IR) parasitic collision points for each experimental intersection region. Dedicated crossing angle orbit "bumps" separate the two LHC beams left and right with respect to the central interaction point in order to avoid collisions at these parasitic collision points. In addition, the use of the pre-existing LEP tunnel sets strong constraints on the LHC parameters: the peak beam energy in a storage ring depends on the integrated dipole field along the storage ring circumference. Considering ultra-relativistic particles traveling in a ring, the maximum particle energy (in TeV) can be expressed as $E = 0.00030 \cdot B \cdot R$ where B is the magnetic field strength (in T) and R is the radius of the collider (in m). The use of superconducting magnet technology and a peak dipole field of about 8 T, allows to reach a peak beam energy of about 7 TeV. Due to technical limitations imposed after the 2008 incident [57], current beam energies are limited to about 6.5 TeV and the design centre-of-mass energy of 14 TeV will be reached only in the next data taking period after 2020. A detailed description of the performance limitations of the LHC can be found in [51] and in the reference therein.

A summary of the LHC parameters is given in Table 2.1.

2.3.1 Pile-up

The high total pp cross-section at the centre-of-mass energies of the LHC and the high luminosity values makes it such that at every beam crossing one has to deal with many simultaneous interactions. These constitute an irreducible background to any physics process of interest as well as an experimental challenge. Due to the finite dimensions of a proton bunch, the vertices of the events are spread over a Gaussian distribution along the beam line (with a Full Width at Half Maximum of the order of 10-12 cm). Given the average detector dimensions of an LHC experiment and a bunch spacing of 25 ns, at any moment

Nominal Beam Parameters at the Large Hadron Collider	
Nominal Energy (protons)	7 TeV
Peak magnetic dipole field	8.33 T
Design Luminosity	$10^{34}\mathrm{cm}^{-2}\mathrm{s}^{-1}$
Circumference	26.66 km
Revolution frequency	11.246 kHz
Number of bunches per proton beam	2808
Number of protons per bunch (at start)	$1.15 \cdot 10^{11}$
Number of collisions per second	600 millions
Minimum distance between bunches	~ 7 m
Bending radius	2803.95 m
Circulating Current	0.53 A
Bunch spacing	25 ns
Normalized transverse beam emittance	3.75 nrad
β^*	0.55 m
Total crossing angle (collision point)	285 μrad
Beam lifetime	10 h
Luminosity lifetime	15 h
Radiated Power per beam (syncrotron radiation)	~ 6 kW

TABLE 2.1: LHC nominal parameters. Ref. [51].

the experiment sees the events belonging up about 3 bunch crossings. This poses stringent conditions on the detector:

- the inner tracker has to be able to discriminate vertices with high resolution, $\mathcal{O}(1)$ mm;
- a complex and robust bunch crossing identification system has to be established.

Every subdetector has a characteristic time window, gate, in which it is active. During this gate signals from different events from different bunch crossings accumulate (*pile-up*, as normally named). In addition, other effects can contribute to the event, such as machine background (beam halo), cavern background (from neutrons and photons), cosmic rays. Thus, any single event is actually the combination of many bunch crossings, many events per bunch crossing and many other contributions due to experimental conditions. All this makes the pile-up a key feature of the LHC physics. Figure 2.4 shows a typical event registered by the ATLAS Experiment in which 25 pile-up vertices were reconstructed from minimum bias interactions.

2.3.2 The LHC performance and future programme

Since the first pp collisions at the end of 2009, the LHC has operated until 2013 (Run 1, at different values of the centre-of-mass energy $\sqrt{s} = 0.9, 2.76, 7, 8\mathrm{TeV}$), with a continuously increasing instantaneous luminosity. By the time that the Run 1 proton physics run ended, the total integrated proton-proton luminosity delivered to ATLAS and CMS had reached nearly 30 fb^{-1}, which permitted the discovery of a Higgs boson. A wealth of lead-lead and proton-lead collisions was also delivered. After the Long Shutdown 1 (LS1, from March 2013 to March 2015), Run 2 began in June 2015 and the LHC delivered for the first time proton-proton collisions at a centre-of-mass energy of 13 TeV, the highest to date.

The timeline of the LHC, including High Luminosity LHC (HL-LHC), and ATLAS upgrade phases can be seen in Figure 2.5.

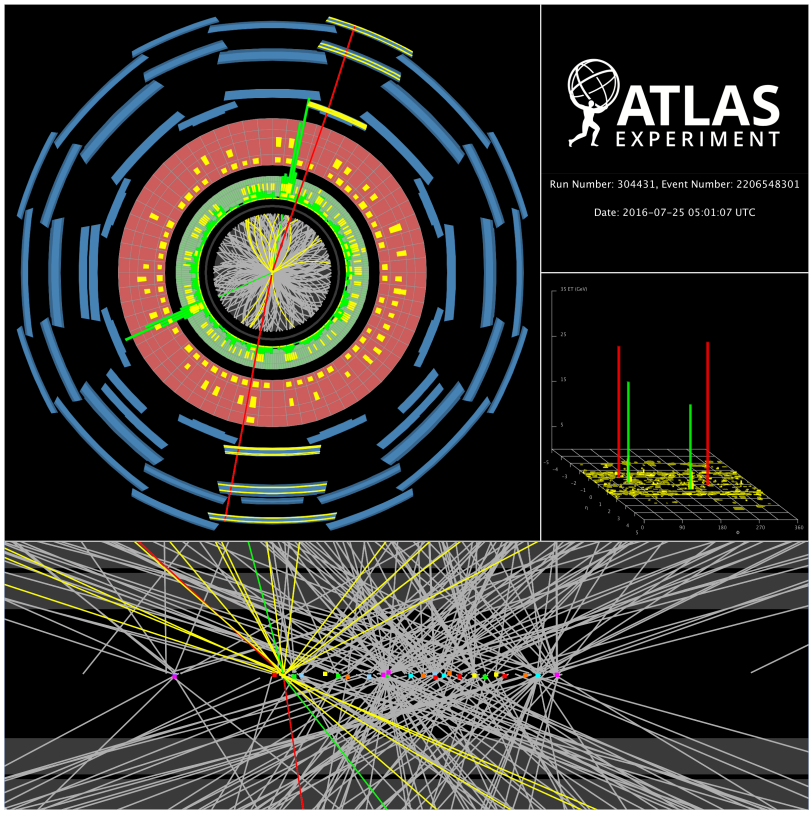


FIGURE 2.4: A display of a candidate Higgs boson event from proton-proton collisions recorded by ATLAS with LHC stable beams at a collision energy of 13 TeV. The Higgs boson candidate is reconstructed in a beam crossing with 25 additionally reconstructed primary vertices from the minimum bias interactions. The candidate event is reconstructed in the $2\mu 2e$ final state. In the left display, the red lines show the path of the two muons including the hits in the muon spectrometer, the green lines show the paths of the two electrons together with the energy deposit in the electromagnetic calorimeter, and the yellow tracks are the remaining charged particles from the Higgs boson candidate vertex. No jet has been reconstructed with transverse momentum above 30 GeV. The grey tracks correspond to the charged particles from the 25 pile-up vertices with transverse momentum above 0.5 GeV and the colored squares in the lower display correspond to the position of the reconstructed vertices. As shown in the lego plot on the right, the muons have a transverse momentum of 29 and 31 GeV, and the electrons 23 and 19 GeV . The invariant mass of the four lepton system is 119 GeV, the dimuon invariant mass is 69 GeV, the dielectron invariant mass is 42 GeV. This event falls within the standard mass selection window of 118 to 129 GeV and has a BDT discriminant response which gives a high probability that this event comes from a Higgs decay as opposed to a background ZZ^* decay. Ref. [58].

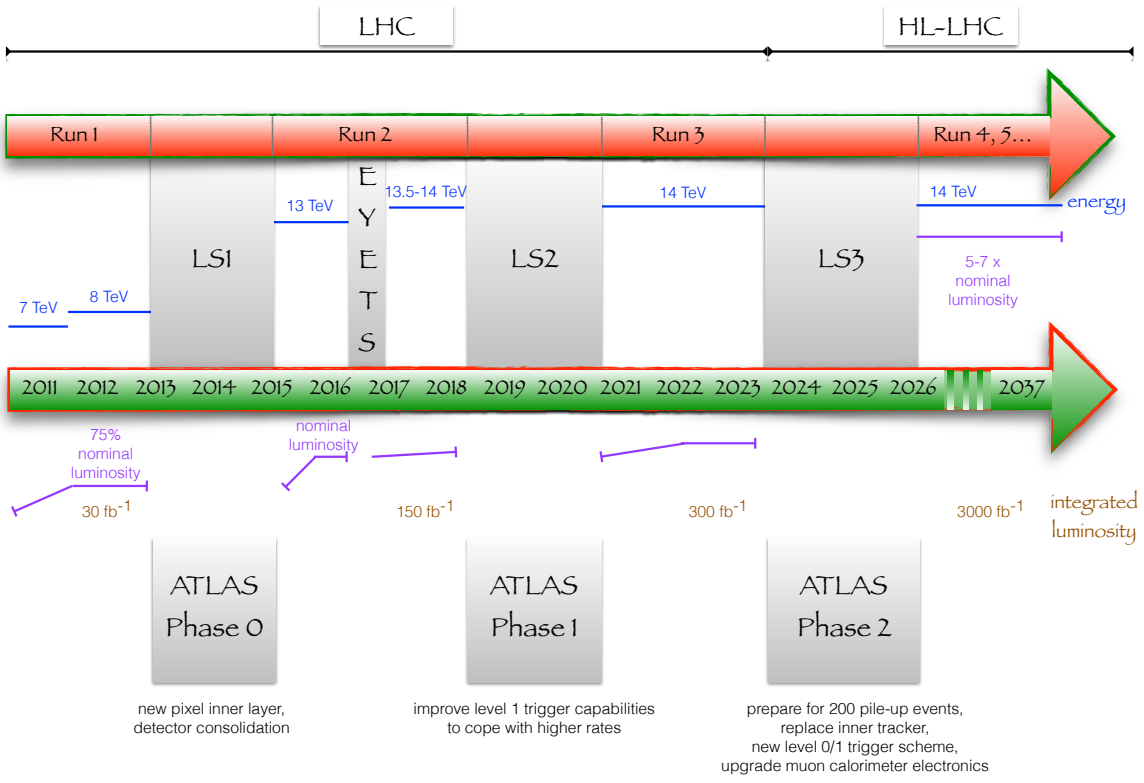


FIGURE 2.5: The LHC and ATLAS upgrade phases with indicated the typical luminosity and centre-of-mass energy values.

Run 1 The first period (2009-2010) was devoted mainly to commissioning and establishing confidence in operational procedures and in the machine protection system, laying the foundations for the physics program which was to follow. Commissioning of the ramp to 3.5 TeV went smoothly and the first (unsqueezed) collisions were established on 30 March. Squeeze commissioning then successfully reduced β^* to 2 m in all four IPs. The decision to go for bunches of nominal intensity came in June 2010. This involved an extended commissioning period and subsequent operation with beams of up to $\sim$ 50 separate bunches and in the next steps the number of bunches was further increased. This required to move to bunch trains with a gap of 150 ns between bunches and the introduction of well defined beam-crossing angles in the interaction regions to avoid parasitic collisions. There was also a back-off in the squeeze to a β^* of 3.5 m. These changes required to set up the tertiary collimators again and to recommission the process of injection, ramp and squeeze. A phased increase in total intensity followed, and at each step, few days of running were necessary to check system performance. At the end of 2010, beams of 368 bunches of about $1.15 \cdot 10^{11}$ protons per bunch and a peak luminosity of $2.1 \cdot 10^{32}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ were achieved. The total integrated luminosity for ATLAS in 2010 was around 0.04 fb^{-1} The beam energy remained at 3.5 TeV in 2011 and during the year the LHC's performance limits were exploited. After a series of tests, encouraging performances led to the decision to run with 50 ns as bunch spacing. The maximum number of bunches at 50 ns (1380) was reached by the end of June 2011. The LHC's performance was increased further by reducing the emittances and by increasing the bunch intensity. The result was a peak luminosity of $2.4 \cdot 10^{33}\ \mathrm{cm}^{-2}\mathrm{s}^{-1}$. In ATLAS and CMS β^* was reduced to 1 m[3] and with an increased bunch inttensity, the peak luminosity went up to $3.7 \cdot 10^{33}\ \mathrm{cm}^{-2}\mathrm{s}^{-1}$. By the end of 2011, ATLAS registered a total integrated luminosity of $\sim 5\ \mathrm{fb}^{-1}$, as it can be seen in Figure 2.6. At the beginning of 2012, the beam energy was increased to 4 TeV and the bunch spacing remained at 50 ns. Together with the use of tight collimator settings, the aperture at the IPs allowed for a more aggressive squeeze to reach β^* 0.6 m, but as a consequence, the sensitivity to orbit movements increased together with the impedance, which had an appreciable effect on beam stability. Peak luminosity soon came close to its highest for the year, although there were determined and long-lasting attempts to further improve performance. Integrated luminosity rates were of the order of 1 fb^{-1} per week, with a total of $\sim$ 20 fb^{-1} registered by ATLAS (proton-proton collisions were extended until Dicember 2012), as it shown in Figure 2.6.

In addition, during Run 1, the LHC also delivered many special runs, such as lead-lead collisions, special fills for LHCb and also runs with β^* = 1 km for TOTEM[4] and the ATLAS forward detectors. In addition, the period culminated in a successful proton-lead run at the beginning of 2013, with ALICE, ATLAS, CMS and LHCb all taking data. Considering the size, complexity and operating principles of the LHC, very good performance was achieved during Run 1. In 2012 there were $\sim$ 200 days of proton-proton collisions with 36.5% of this time spent in stable-beams conditions.

[3] As previously explained, smaller beam size at an IP implies bigger beam sizes in the neighbouring inner triplet magnets, but careful measurements revealed a better-than-expected aperture in the interaction regions, opening the way for this further reduction in β^*.

[4] TOTEM stands for TOTal, Elastic and diffractive cross-section Measurement. With its detectors spread across almost 0.5 km around the CMS interaction point, TOTEM is the CERN's longest experiment. It has $\sim$ 3000 kg of equipment, including four particle telescopes as well as 26 Roman pot detectors. The telescopes use cathode-strip chambers and Gas Electron Multipliers (GEM) to track the particles emerging from collisions at the CMS interaction point. Meanwhile, Roman Pots with silicon sensors perform measurements of scattered protons.

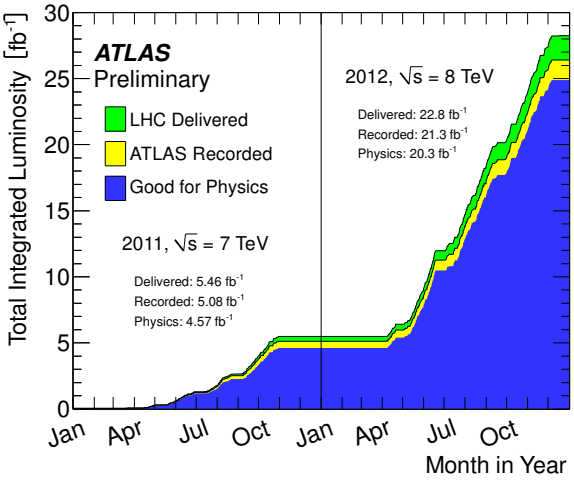


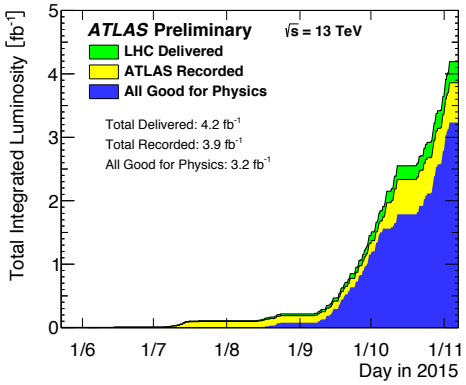
FIGURE 2.6: Total Integrated Luminosity collected by ATLAS during the Run 1 of the LHC. Ref. [59].

Run 2 After two years of LS1 and several months of re-commissioning activities with and without beam, in June 2015 the LHC delivered for the first time proton-proton collisions with stable beams at a centre-of-mass energy of 13 TeV, the highest to date.

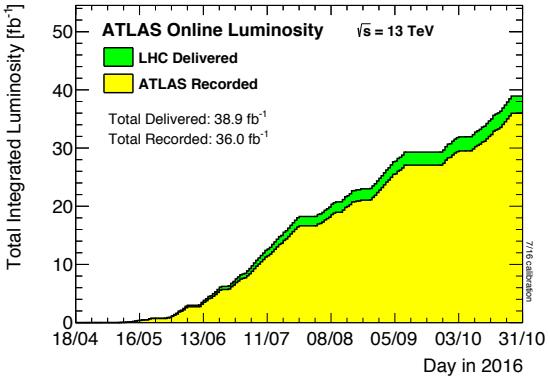
An early goal of the experiments consisted in the collection of inelastic pp collisions with a special LHC configuration, which allows for a very low expected mean number of interactions per bunch crossing, $< \mu > = 0.005$. By exploiting this data-set, inelastic pp collisions were, in fact, characterised in terms of charged-particle production (see Chapter 5) and underlying event [60].

During the first phases of Run 2 in 2015, the performance of the accelerator was exploited at the new centre-of-mass energy by gradually increasing the intensity of the beams. The number of proton bunches increased up to 2244 per beam, spaced at intervals of 25 nanoseconds and the delivered integrated luminosity at the end of 2015 was about 4 fb^{-1}, as shown in Figures 2.7(a) and 2.7(d). After a short technical stop, from December 2015 to March 2016, the LHC resumed operations with the aim of collecting about 25 fb^{-1} at the end of 2016, but the excellent performance achieved during the year, allowed the accelerator limits to pushed further. The design value of the instantaneous luminosity was surpassed in June 2016 reaching up 10% more than the nominal value in October 2016, as shown in Figure 2.7(c). Some runs were characterised by very long beam lifetime with each proton beam circulating in the accelerator for more than 30 hours and at the end to the 2016 data-taking period the recorded integrated luminosity was about 40 fb^{-1}, as it is shown in Figure 2.7(b), which corresponds to almost a factor of 2 more than the expected value. The mean number of interactions per crossing (the mean of the Poisson distribution of the number of interactions per crossing calculated for each bunch) reached an average value of 13.7 in 2015 and 24.2 in 2016, as shown in Figure 2.7(e).

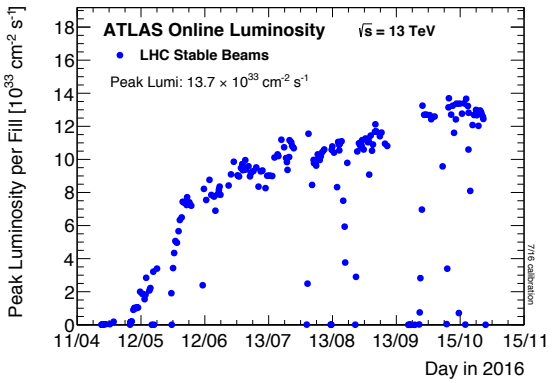
The 2016 pp data-taking ended on the 26th October and, as it happened in previous years, the LHC prepared to deliver special heavy-ions runs. After a successful lead-lead (Pb-Pb) collision run in 2015 at 5 TeV, the accelerator was reconfigured at the end of 2016 for a month-long heavy-ion run, devoted to colliding beams of protons and lead nuclei. Following a feasibility test in 2012 and an initial month-long run in 2013, pPb collisions remain a novel mode of operation at the LHC. A first data sample was collected in November at the centre-of-mass energy of 5.02 TeV. One of the first fills also turned out to be the longest LHC fill ever, lasting almost 38 hours The accelerator was then prepared to deliver such a collisions at a higher centre-of-mass energy corresponding to 8.16 TeV, the highest



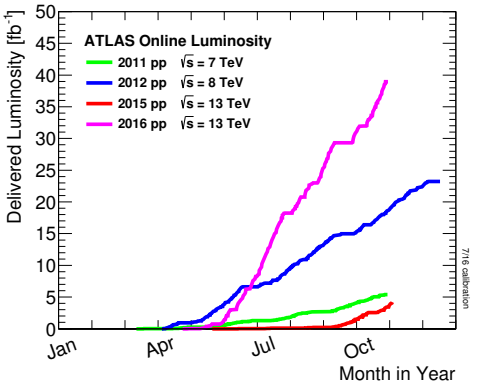
(a) Total Integrated Luminosity and Data Quality in 2015



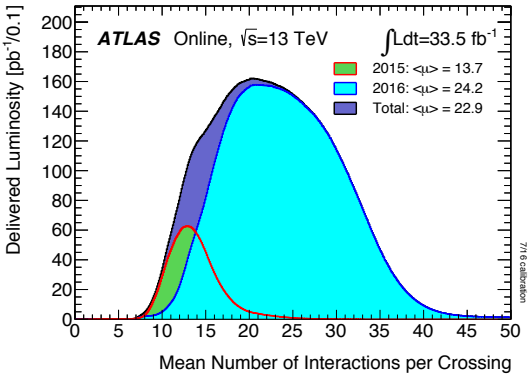
(b) Total Integrated Luminosity in 2016



(c) Peak Luminosity per Fill in 2016



(d) Delivered Luminosity versus time for 2011-2016 (*pp* data only)



(e) Number of Interactions per Crossing

FIGURE 2.7: ATLAS Run 2 Luminosity Plots. Ref. [59].

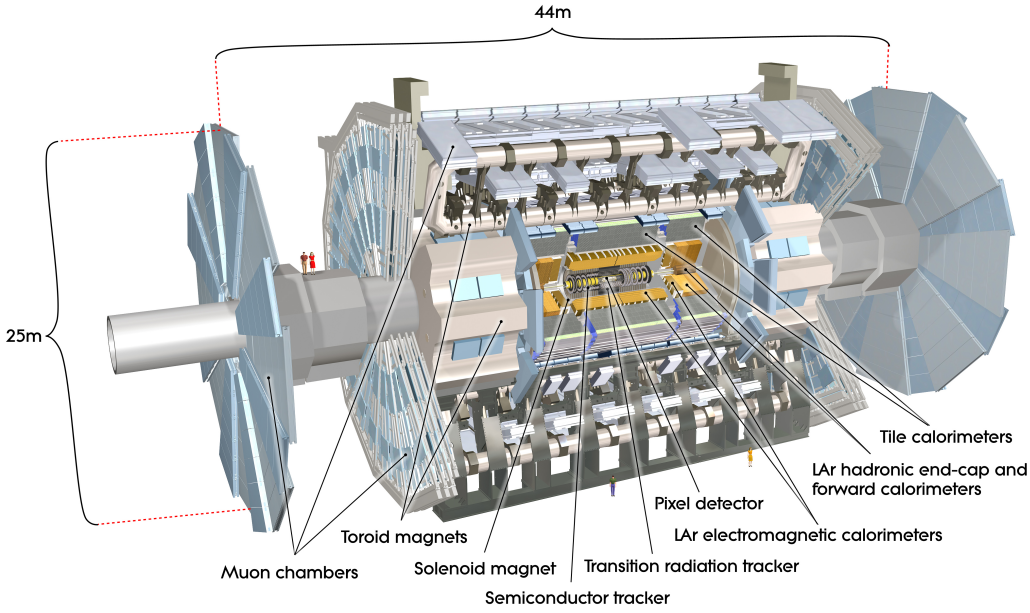


FIGURE 2.8: ATLAS Experiment. The dimensions and the different subdetectors are shown.

to date for these asymmetric collisions. The energy was lowered again to 5 TeV before the technical stop period, to allow the ALICE Experiment to collect more events.

The LHC Run 2 also included successful forward-physics runs for the TOTEM/CT-PPS, ALFA and AFP experiments.

During the 2015-2016 data-taking periods, the LHC operated in stable-beams mode for 60% of its operational time, an unprecedented record as no other energy-frontier accelerator ever reached more than 50% (the Large Electron Positron was in stable beam for about 30% of its operational lifetime).

2.4 The ATLAS Experiment

The ATLAS Experiment [52] is an international collaboration involving 170 institutes spread all over the world. Given the experimental conditions provided by the LHC (even harsher in Run 2), ATLAS needs fast and radiation-hard electronics and sensor elements. In addition, high detector granularity is necessary to handle the particle fluxes and to reduce the influence of overlapping events. The solid angle coverage is required to be as large as possible. Good charged-particle momentum resolution and reconstruction efficiency in the inner tracker are essential and for offline tagging of b-jets, vertex detectors close to the interaction region are required to observe secondary vertices. Electromagnetic and hadronic calorimeters have to provide respectively good photon/electron identification and accurate jet and missing transverse energy measurements. In addition, good muon identification and momentum resolution over a large range of momenta are fundamental requirements. Highly efficient triggering on low transverse-momentum objects with sufficient background rejection, is a prerequisite to achieve an acceptable trigger rate for most physics processes of interest. The overall structure of the ATLAS detector can be seen in Figure 2.8. The ATLAS experiment covers nearly the entire solid angle around the beam interaction point, It is built with an onion structure and the positions of the different subdetectors is determined by the particles to be detected and their interaction with

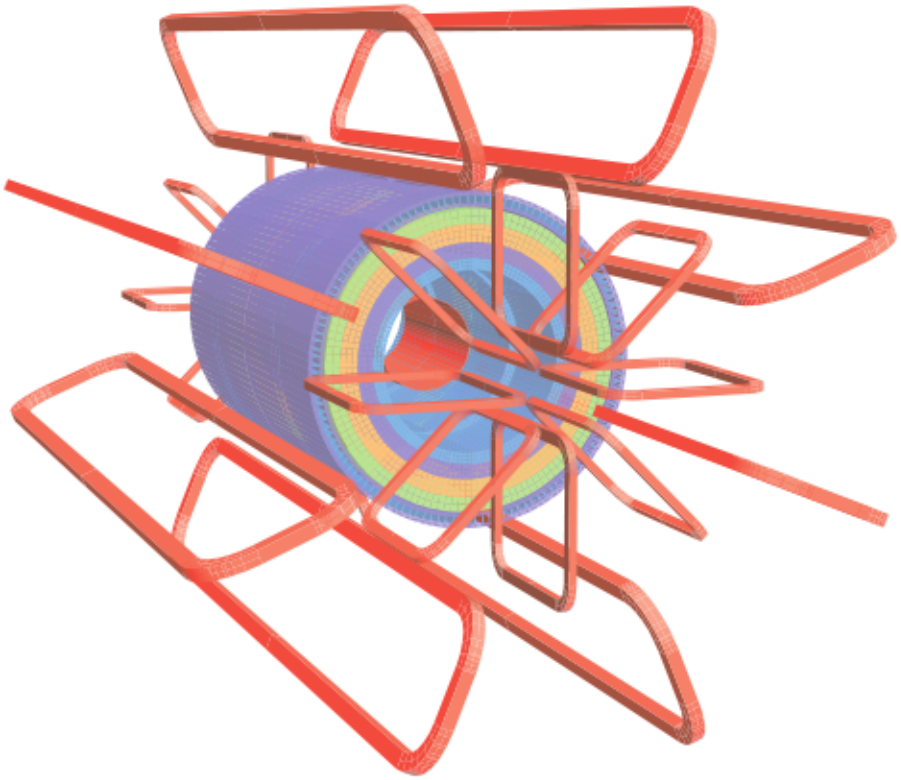


FIGURE 2.9: Schema of the magnetic system geometry. The eight barrel coils are visible together with the endcap coils and the central solenoid.

matter. The Inner Detector used to track charged-particles is the closest detector to the interaction point and it is placed in a 2 T magnetic field generated by a superconductive solenoid. Going further out in the detector, an electromagnetic and a hadronic calorimeter are used to identify, via showering, electromagnetic particles and hadrons and to measure their energy. The system of hadronic calorimeters lies into a toroidal magnet made by 8 coils which determines the entire ATLAS outer structure. Into this magnetic system and in the furthermost region from the interaction point, the detectors to trigger and track muons are found.

2.4.1 The Magnets

In high energy physics experiments, magnetic fields are used to bend charged-particles into helical paths. Measuring the curvature radius allows to evaluate the particles' charge and momentum. The magnetic system [61] of the ATLAS Detector consists of four different superconductive magnets: a central solenoid surrounding the Inner Detector region, a barrel toroid and two end cap toroids which form the outer muon spectrometer. A schema can be seen in Figure 2.9.

The solenoid

The ATLAS solenoid [62], shown in Figure 2.10, produces a magnetic field approximatively uniform in the ATLAS central region and oriented along its Z axis. The coil has a diameter of 2.4 m and a length of 5.3 m. When a 8kA current circulates in the superconductive Nb-Ti windings, it generates a 2 T field. To be working, the coil needs to be

FIGURE 2.10: The ATLAS Detector Solenoid.

cooled down to 4.5 K by using liquid helium. To minimize the particle path through inactive material, the cryostat containing liquid helium is coupled together with the one which contains the Liquid Argon (LAr) electromagnetic calorimeter.

The barrel toroid

In a simplified view, a toroid can be imagined as a solenoid with the two ends coupled together and, in the ideal case, the magnetic field is contained in the toroid itself. In the ATLAS experiment, the barrel toroid [63] consists of 8 big superconductive coils and as many cryostats. With its enormous structure, the ATLAS toroid is the biggest ever built magnetic system. As in the case of the solenoid, each coil is made of Nb-Ti wire windings in an aluminum matrix and it is nominally operating at 20 kA giving a magnetic field of $\sim$ 3.9 T close to the coils and $\sim$ 0.5 T at the centre of the coils.

The end-cap toroids

The end-cap toroids [64], one of them being shown in Figure 2.12, generate a magnetic field close to the beam pipe in order to bend particles being produced at very small angles with respect to the beam direction. As for the barrel toroids, the end-cap toroids are made by 8 Nb-Ti coils, but in this case they are coupled together in a single cryostat. Each end-cap toroid is $\sim$ 5 m long with an inner bore of 1.65 m and an outer diameter of 10.7 m. The magnetic field is not uniform and it can reach up to 4.1 T in the coils. The end-cap toroid coil system is rotated by 22.5 with respect to the barrel toroid coil system in order to provide radial overlap and to optimise the bending power at the interface between the two coil systems.

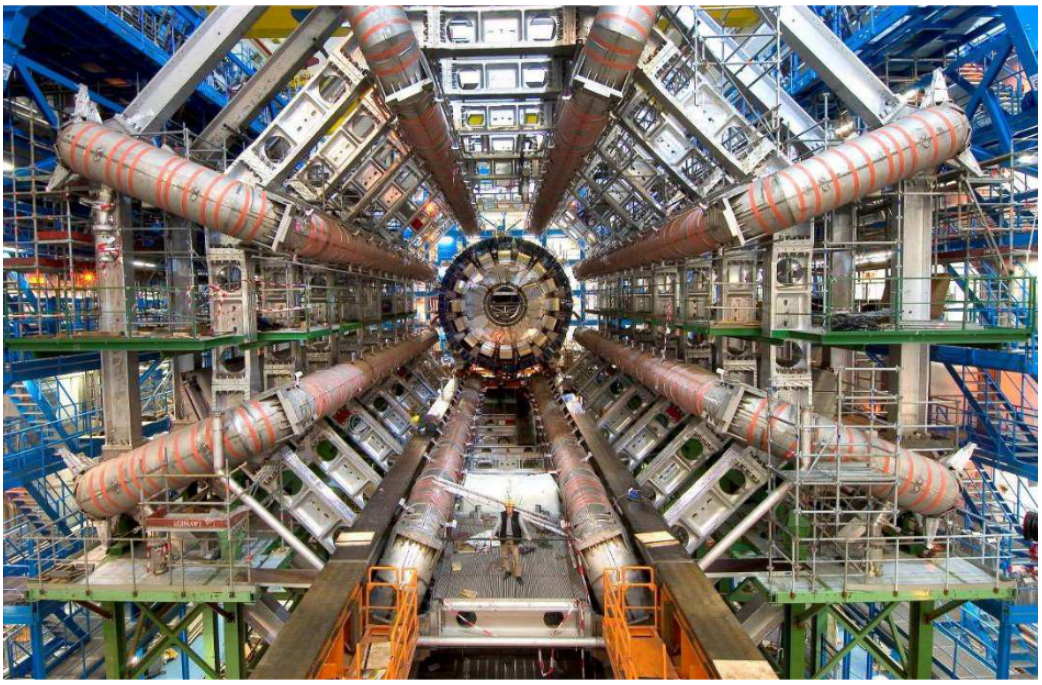


FIGURE 2.11: A picture of the barrel toroid in a comparison with the human dimensions.

FIGURE 2.12: A picture of one of the two ATLAS end-cap toroids.

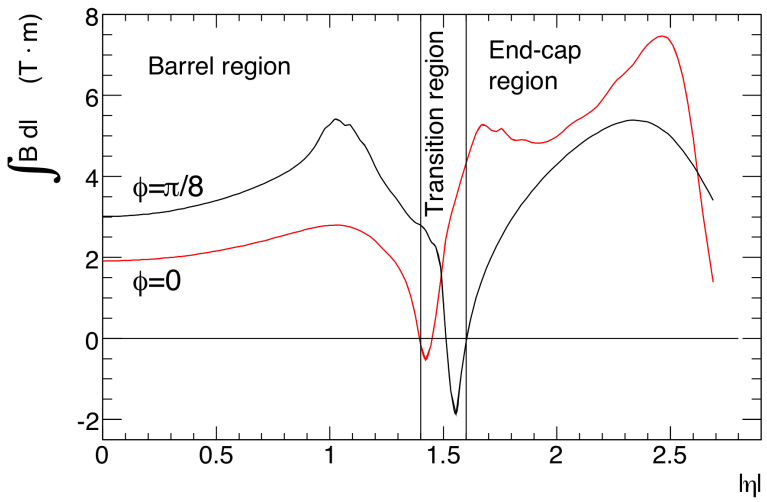


FIGURE 2.13: $\int \mathrm{Bdl}$ as a function of $|\eta|$, for infinite-momentum muons. Two curves are shown corresponding to the azimuthal angles of $\phi = 0$ (red line) and $\phi = \pi/8$ (black line).

The toroidal field

The magnetic field, which is mainly orthogonal to the muon trajectories, is provided by the barrel toroid in the region of pseudorapidity $|\eta| < 1.4$ [5], by a combination of the barrel and end-cap toroids in the so-called transition region defined by $1.4 < |\eta| < 1.6$ and by the end-cap toroids for $1.6 < |\eta| < 2.7$. The performance in terms of bending power is characterised by the field integral $\int \mathrm{Bdl}$, with B representing the field component normal to the muon direction and the integral being computed along an infinite-momentum muon trajectory, between the innermost and outermost muon-chamber planes. As can be seen in Figure 2.13, the barrel toroid provides between 1.5 and 5.5 Tm of bending power in the pseudorapidity range $|\eta| < 1.4$ and the end-cap toroids approximately 1 to 7.5 Tm in the region $1.6 < |\eta| < 2.7$. Within the transition region, where the two magnets overlap, the bending power is lower and not uniform.

2.4.2 The Inner Detector

The ATLAS Inner Detector (ID) [65, 66], shown in Figure 2.14, is designed to reconstruct the trajectory of charged-particles, to measure their momentum and to determine the positions of primary vertices, i.e. positions of the inelastic proton-proton interactions, and secondary vertices, i.e. vertices generated from particle decays and from interactions with matter. It covers the region of pseudorapidity $|\eta| < 2.7$. A particle traversing the Inner Detector deposits a certain energy (it needs to be tiny not to affect the energy measurements in the calorimeters) which is collected by the detector's sensors: these *hits* allow to identify the particle's path and they can be used to reconstruct the particles' trajectories. The ATLAS ID is contained in a 6 m long cylindrical region with a radius of about 1 m. It encloses the beam pipe and it is surrounded by the central solenoid.

[5] ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the centre of the detector and the z-axis along the beam pipe. The x-axis points from the IP to the centre of the LHC ring, and the y-axis points upwards. Cylindrical coordinates (r, ϕ) are used in the transverse plane, ϕ being the azimuthal angle around the z-axis. The pseudorapidity is defined in terms of the polar angle θ as $\eta = -\ln\tan(\theta/2)$. Angular distance is measured in units of $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.

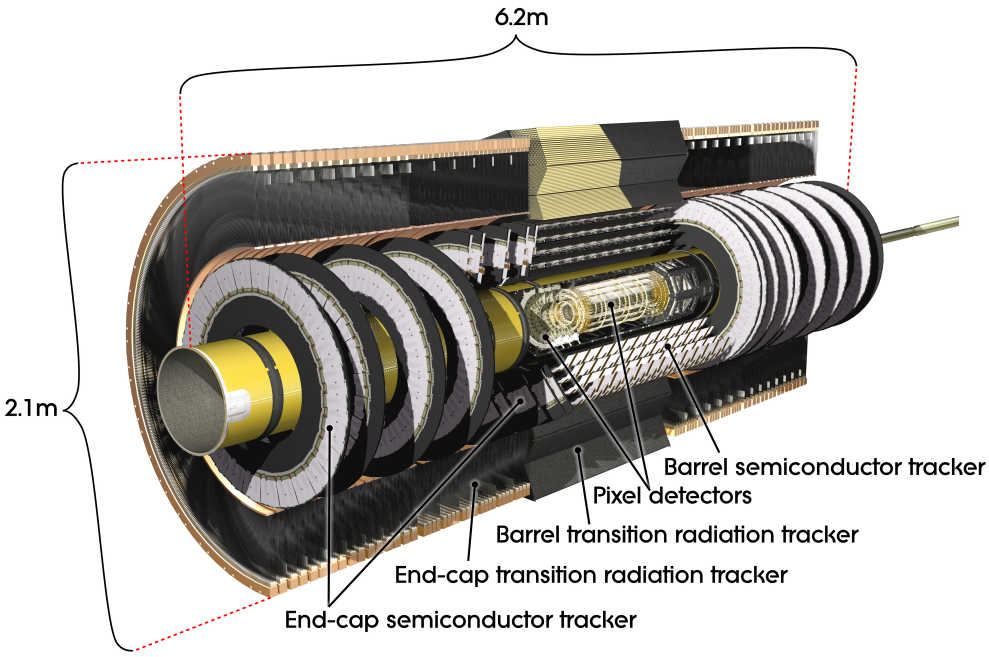


FIGURE 2.14: The ATLAS Inner Detector with the subdetectors shown.

The ATLAS ID comprises three independent but complementary detector systems: the Pixel detector (Pixel), the Semiconductor Tracker (SCT) and the Transition-Radiation drift-tube Tracker (TRT). In the innermost region, discrete space-points from silicon pixel layers and stereo pairs of SCT layers allow for high-resolution pattern recognition capabilities, while at larger radii, the TRT, composed of many layers of gaseous straw tube elements interleaved with polyethylene foils to generate transition radiation by electrons, records up to 40 space points per track. The combination of precision trackers at small radii with the TRT at larger radius allows for a very robust pattern recognition and high precision in both $R-\phi$ and z coordinates. For what concerns the measurements in the TRT, a lower per-point precision with respect to that achieved with the silicon detectors is compensated by the large number of measurements and longer measured track length. The electron identification capabilities are enhanced by the detection of transition radiation photons in the xenon-based gas mixture of the straw tubes, while the semiconductor detectors also allow for impact parameter measurements and vertexing for heavy-flavour and $\tau-$lepton tagging.

During Long Shutdown 1, the ATLAS Inner Detector was upgraded, by inserting a new, innermost barrel layer of Pixel called the Insertable B-Layer (IBL) [67], as shown in Figure 2.15. The upgraded Pixel detector spans over the radial region of 33–150 mm (measured from the interaction point), while the SCT and TRT detectors cover the radial regions of 299–560 mm and 563–1066 mm, respectively. The $r-z$ cross-section view of the ID can be seen in Figure 2.16.

Pixel Detector

The Run 1 Pixel detector [68] consists of three barrel layers (referred to as PIX1, PIX2, PIX3 from inner to outer) and three end-cap disk layers. It hosts 1744 pixel-sensor modules, and each module contains 46 080 pixels with a typical size of 50 μm (r–ϕ) $\times$ 400 μm (z). The detector contains around 80 million pixels in total. The radii of the three barrel layers are 50.5 mm, 88.5 mm and 122.5 mm respectively. The barrel and end-cap layers of the pixel

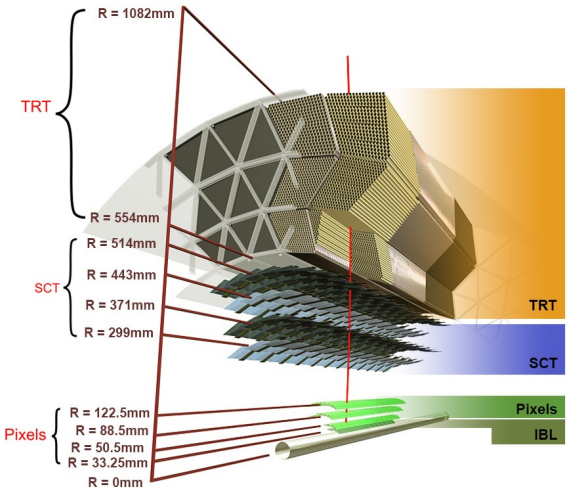


FIGURE 2.15: Sketch of the ATLAS inner detector showing all its components, including the new insertable B-Layer (IBL). The distances to the interaction point are also shown.

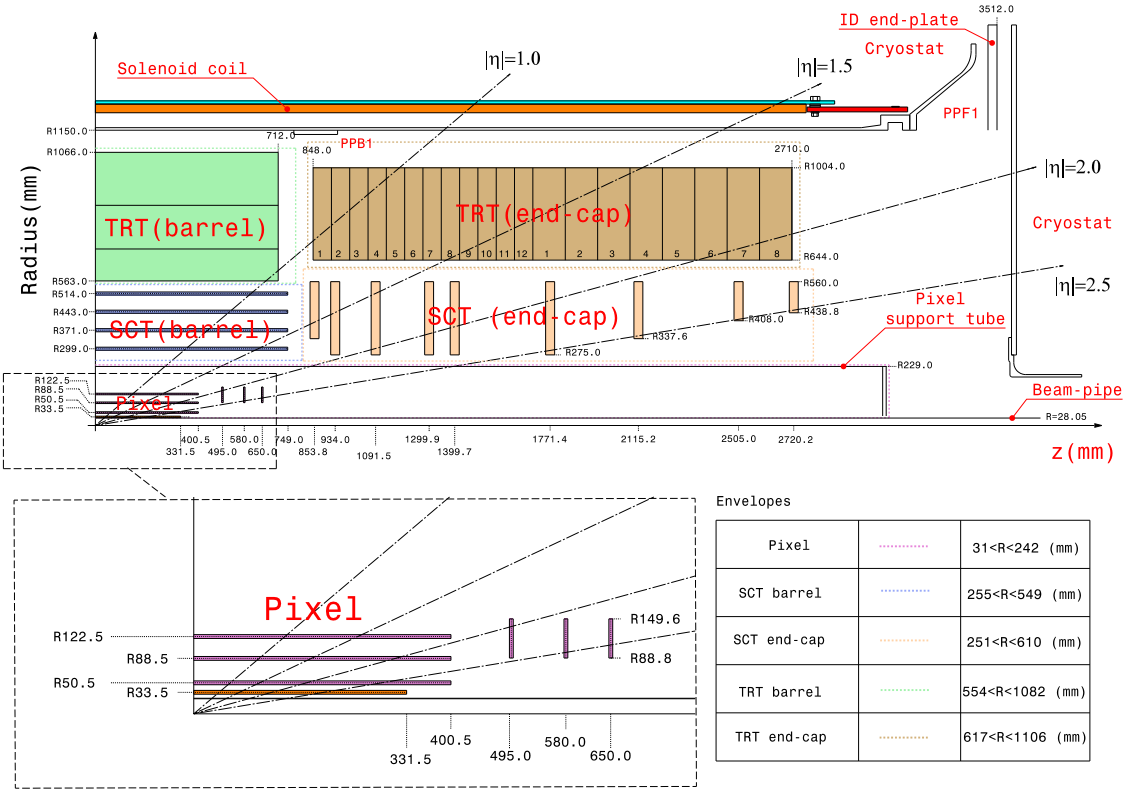


FIGURE 2.16: The r–z cross-section view of the layout of the ATLAS Inner Detector for Run 2. Compared to Run 1, the IBL (shown in red) and its services, together with the new beam pipe, were added. The top panel shows the whole inner detector, whereas the bottom-left panel shows a magnified view of the pixel detector region.

detector are supported by an octagonal prism structure knows as the pixel support frame (PSF) with a radius of $r \simeq 200$ mm. It is inserted inside the pixel support tube (PST) which has a radius of 229 mm.

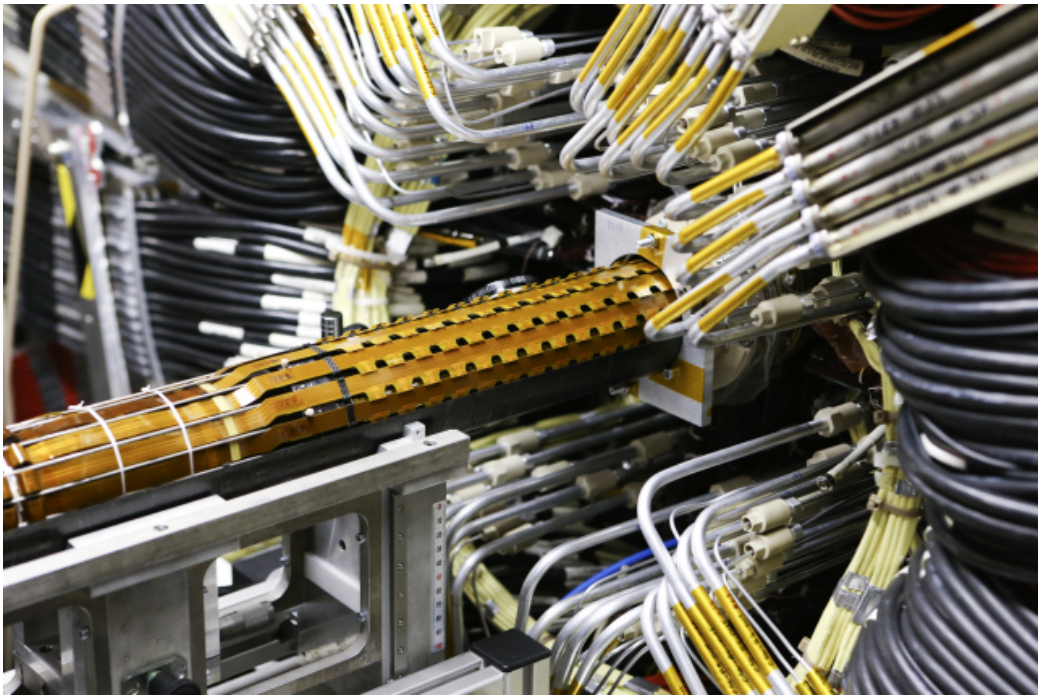


FIGURE 2.17: IBL installation into the inner detector of the ATLAS Experiment. Ref. [69].

The high-radiation environment imposes stringent conditions on the Pixel sensors, on detector electronics, mechanical structure and services. To maintain an adequate noise performance after radiation damage, the silicon sensors must be kept at low temperature (in the range [-5, -10] C°). That notwithstanding, the performance of the first pixel layer degrades with time because of the high radiation dose and after years of continuous running at design luminosity the PIX1 layer will most likely fail.

For Run 2, the Pixel detector was refurbished and improved by inserting the IBL, shown in Figure 3.9, which allows for a fourth measurement point in the Pixel detector. The IBL is installed at a radius of $r = 33.5$ mm and it covers the pseudorapidity region of $|\eta| < 3.03$. It consists of 14 staves with $\sim$ 12 million 50×250 μm^2 pixel sensors. The IBL staves are placed between the inner positioning tube (IPT) at $r = 29.0$ mm and the inner support tube (IST) at $r = 42.5$ mm. The IPT and IST are made of carbon fibre and resin. The thickness of the IPT goes from 0.325 mm for $|z| < 311$ mm, to 0.455 mm at the farthest edge.

To minimise the distance of the IBL from the interaction point, the beam pipe used during Run 1 was replaced with a new, thinner one. The new beam pipe consists of a 0.8 mm-thick beryllium pipe wrapped with polyamide tapes and aerogel thermal insulators. It has an inner radius of 23.5 mm, and an outer radius which goes from 24.3 mm for $|z| < 30$ mm to 28.2 mm for $|z| > 311$ mm. The thermal insulator has been removed in the central part of the new beam pipe at $|z| < 311$ mm, in order to reduce material thickness. The material composition of the new beam pipe was measured using X-rays [70] as well as by mass measurements to a precision of 1% before installation: these measurements were used to update the geometry model description and simulation, as explained in Chapter 4. In total, the extra material introduced by the IBL in the ID corresponds to roughly 1.5% of a radiation length averaged over ϕ for tracks which are perpendicular to the IBL staves and originating from $r = 0$.

Semiconductor Tracker

The Semiconductor Tracker consists of 4088 silicon strip modules, arranged in four barrel layers (referred to as SCT1, SCT2, SCT3, SCT4 from inner to outer) and two end-caps comprising of nine wheels each. Each module is composed of two layers of silicon strip detector sensors glued back-to-back with a relative stereo angle of $40\,\mathrm{mrad}$. Each silicon detector is 6.36 x 6.40 cm^2 with 780 readout strips of 80 $\mu\mathrm{m}$ pitch. The barrel modules are mounted on carbon-fibre cylinders at radii of 30.0, 37.3, 44.7, and 52.0 cm. The end-cap modules are very similar in construction but use tapered strips with one set aligned radially. The SCT barrel layers are supported by the inner and outer thermal enclosures, referred to as the SCT-ITE and SCT-OTE respectively, which are located at $r \simeq 255\ \mathrm{mm}$ and $r \simeq 550\ \mathrm{mm}$. The SCT system was not modified since Run 1.

Transition Radiation Tracker

The TRT is the outermost of the ID subdetectors. It consists of more than 350 thousand polyamide drift (straw) tubes filled with a gas mixture of 70% Xe, 27% CO_2 and 3% O_2. The straw tube wall, designed to have good electrical and mechanical properties with minimal wall thickness, is made of two 35 μm thick multi-layer films bonded back-to-back. Each straw is 4 mm in diameter and equipped with a 30 $\mu\mathrm{m}$ diameter gold-plated W-Re wire. The straw tubes length is 144 cm for the barrel and 37 cm for the end-caps. Each channel provides a drift time measurement, giving a spatial resolution of 170 $\mu\mathrm{m}$ per straw, and two independent thresholds. These allow the detector to discriminate between tracking hits, which pass the lower threshold, and transition radiation hits, which pass the higher one. In contrast to the silicon sensors, the TRT is designed to operate at room temperature. The operating specifications also imply requirements on the alignment precision and on the stability between alignment periods. Quantitatively, the track precision should not deteriorate by more than 20% between alignment periods. For this reason, benchmark values for the alignment are regularly provided by the ATLAS Tracking group, as mentioned in Chapter 3. As for the SCT, the TRT system was not modified after Run 1.

Pixel Services

When the Pixel detector was extracted from ATLAS during LS1, refurbished and reinstalled, together with replacing the disabled modules and inserting the IBL, also the services between the Pixel and the SCT detectors were changed. The service panels located at $|z| \geq 780$ mm inside the PST were replaced with the new Service Quarter Panels (nSQP) [71]. The optical–electrical signal conversion boards, which were previously placed on the old service panels at $(r,\ z) \simeq (174,\ 1070)\ \mathrm{mm}$, were moved to a region referred to as the ID end-plate, which is located between the ID and the end-cap calorimeters. These devices are relatively far from the interaction point, to reduce ionising radiation dose and to facilitate the access for maintenance. This also helps to reduce the material amount of the nSQPs compared to the previous service panels. The service areas where these changes occurred are called patch panel-0 (PP0) of the Pixel services. Photographs of the nSQPs are shown in Figure 2.18.

2.4.3 The Calorimeter System

The ATLAS Calorimeter system [73], shown in Figure 2.19, is composed of a collection of electromagnetic and hadronic calorimeters which covers a pseudorapidity range of $|\eta| < 4.9$. It is based on different techniques in order to fulfill a variety of requirements dictated by the physics processes to be studied and to cope with the radiation environment

(a) (b)

FIGURE 2.18: (a) A quadrant of the bare new Service Quarter Panel (nSQP). The front side is connected to the detector. (b) Pixel services after re-assembly in LS1. In this photograph, the right-hand-side is the detector, and the left-hand-side represents the services. Ref. [72].

over the whole η range. In the liquid argon (LAr) electromagnetic sampling calorimeters, which surround the ID, the fine granularity was chosen to precisely measure electrons and photons, while the hadronic calorimeters, with a coarser granularity, meet the requirements set for hadronic resolution, jet reconstruction and missing transverse energy ($E_T{}^{\text{miss}}$) measurements. Calorimeters must be hermetic, provide good containment for electromagnetic and hadronic showers, and also limit hadronic punch-through into the muon system. Therefore, calorimeter thickness is an important parameter. The total thickness of the EM calorimeter is larger than 22 radiation lengths (X_0) over the whole barrel and 24 radiation lengths in the end-caps. The ~ 9.7 interaction lengths (λ) of active calorimeter in the barrel (~ 10 in the end-caps) are sufficient to provide good resolution for high-energy jetswhile ensuring almost complete shower containment. Including also a 1.3 λ for the outer support, the total thickness of 11 λ at $\eta = 0$ is adequate to reduce the punch-through well below the irreducible level of prompt or decay muons. The numbers of interaction lengths in front of and in the electromagnetic and hadronic calorimeters is shown in Figure 2.20.

Electromagnetic Calorimeter

The Electromagnetic (EM) Calorimeter consists of a barrel (EMB, $|\eta| < 1.475$) and two end-cap (EMEC, $1.475 < |\eta| < 3.2$) calorimeters, each of them positioned in a separate cryostat. In order to achieve the targeted performance of the EM calorimeter in the barrel region, the choice of having the central solenoid in front of the EM calorimeter requires a careful optimization of the material budget. For this reason, the central solenoid and the LAr calorimeter [74] share a common vacuum vessel, eliminating the need of two vacuum walls. The barrel calorimeter is composed of two identical half-barrels, separated by a small gap (4 mm) at z=0. Each end-cap calorimeter is mechanically divided into two coaxial wheels: an outer wheel covering the region $1.375 < |\eta| < 2.5$ and an inner wheel to cover the region $2.5 < |\eta| < 3.2$. The EM calorimeter is a lead-LAr sampling calorimeter, featuring particular accordion-shaped Kapton electrodes and lead absorber plates over its full coverage: this peculiar geometry provides complete coverage around ϕ without azimuthal cracks. In the region of $|\eta| < 2.5$ where most of the physics analyses are restricted, the EM calorimeter is segmented in three sections in depth, while in the end-cap inner wheel, the calorimeter is segmented in only two sections in depth and has a

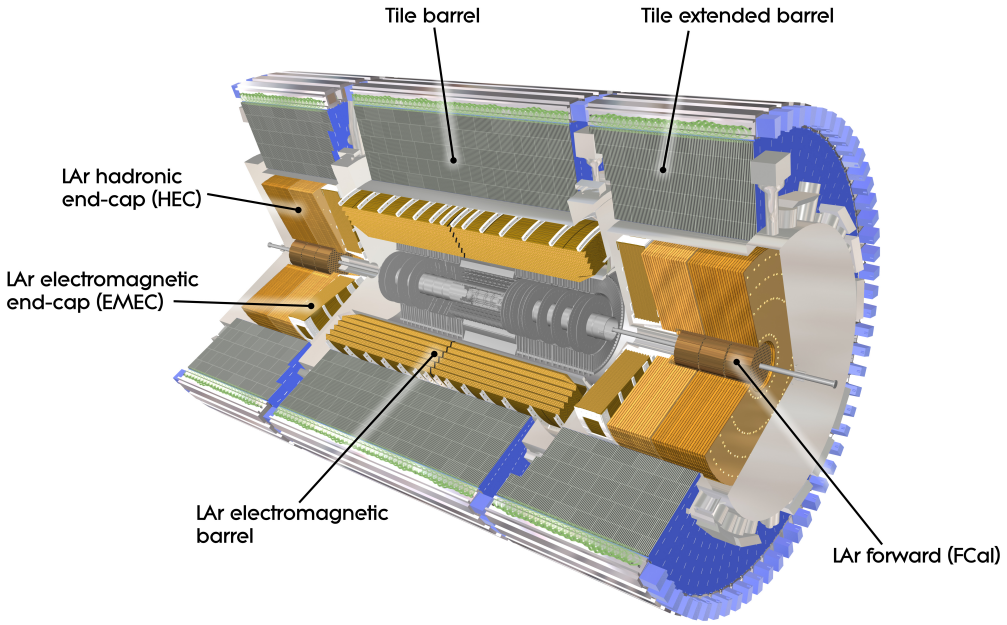


FIGURE 2.19: The ATLAS calorimeters with the subdetectors shown.

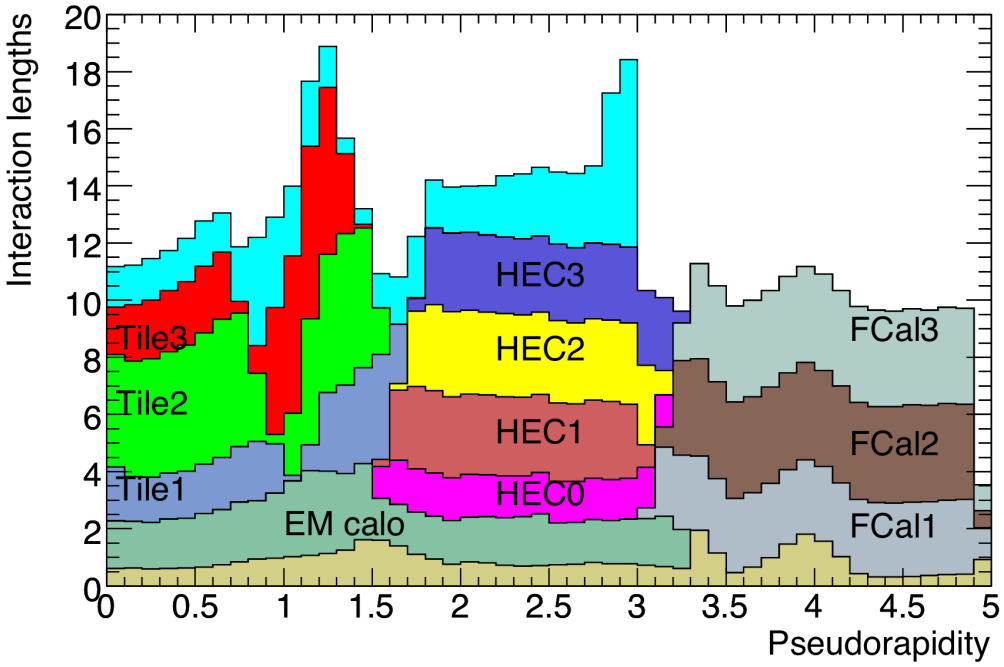


FIGURE 2.20: The number of interaction lengths in the subdetectors of the ATLAS calorimeters.

coarser lateral granularity. To improve the e-π separation and to increase the prompt $\gamma-\pi_0$ discrimination as well as to take into account and correct for the energy lost by electrons and photons upstream of the calorimeter, in the region of $|\eta| < 1.8$ a pre-sampler made of an active LAr layer (1.1 cm in the barrel and 0.5 cm in the end-caps) is used.

Hadronic Calorimeter

A scintillator-tile calorimeter [75] placed outside the EM calorimeter cryostat in the barrel region, provides hadronic calorimetry in the range of $|\eta| < 1.7$. The tile calorimeter is divided into a central barrel and two smaller extended barrel sectionss, one on either side of the central barrel, segmented in 64 azimuthal modules. Radially, the tile calorimeter extends from an inner radius of 2.28 m to an outer radius of 4.25 m. The total detector thickness at the outer edge of the tile-instrumented region is 9.7 λ at $\eta = 0$. Two sides of the scintillating tiles are read out by wavelength shifting fibres into two separate photomultiplier tubes. Readout cells, defined by grouping fibers into the photomultipliers, are pseudo-projective towards the interaction region.

Owing to the radiation levels which are foreseen to be reached in the end-caps, LAr technology is used in stead of scintillator tiles for hadronic calorimetry in the end-caps (hadronic end-cap calorimeter, HEC): the HEC matches the pseudorapidity coverage of the end-cap EM calorimeters, with which it ishares a common cryostat. In order to reduce any gap and maximise hermeticity, the HEC covers a region which extends from $|\eta| = 1.5$ (overlapping with the tile calorimeter) up to $|\eta| = 3.2$, overlapping with the forward calorimeter which starts at $|\eta| = 3.1$. Each wheel is composed of 32 identical wedge-shaped modules, assembled with fixtures at the periphery and at the central bore. Each wheel is divided into two segments in depth, for a total of four layers per end-cap. The wheels closest to the interaction point are built out of 25 mm parallel copper plates, while those further away use 50 mm copper plates (for all wheels the first plate is half-thickness). The outer radius of the copper plates is 2.03 m, while the inner radius is 0.475 m (except in the overlap region with the forward calorimeter where this radius becomes 0.372 m). The copper plates are interleaved with 8.5 mm LAr gaps, providing the active medium for this sampling calorimeter.

Forward Calorimeter

The LAr forward calorimeters (FCal) provide both electromagnetic and hadronic energy measurements into an extended range of pseudorapidity up to $|\eta| = 4.9$. Since the depth of the calorimeter is limited by the fact that, in order to reduce the amount of neutron albedo in the inner detector cavity, the FCal front face is recessed by about 1.2 m with respect to the EMEC front face, a high-density design was necessary for the FCal. With a total of ~ 10 λ, each end-cap is made of three modules: the first one, made of copper, measures the electromagnetic component of the showers, while the other two, made of tungsten are optimised for hadronic energy measurements. Each module is constituted of a metal matrix, with regularly spaced longitudinal channels filled with the electrode structure, consisting of concentric rods and tubes parallel to the beam axis. The liquid Argon which fills the gap between the rod and the tube acts as sensitive medium. This geometry allows for excellent control of the gaps, which are as small as 0.25 mm in the first section in order to avoid problems due to ion buildup.

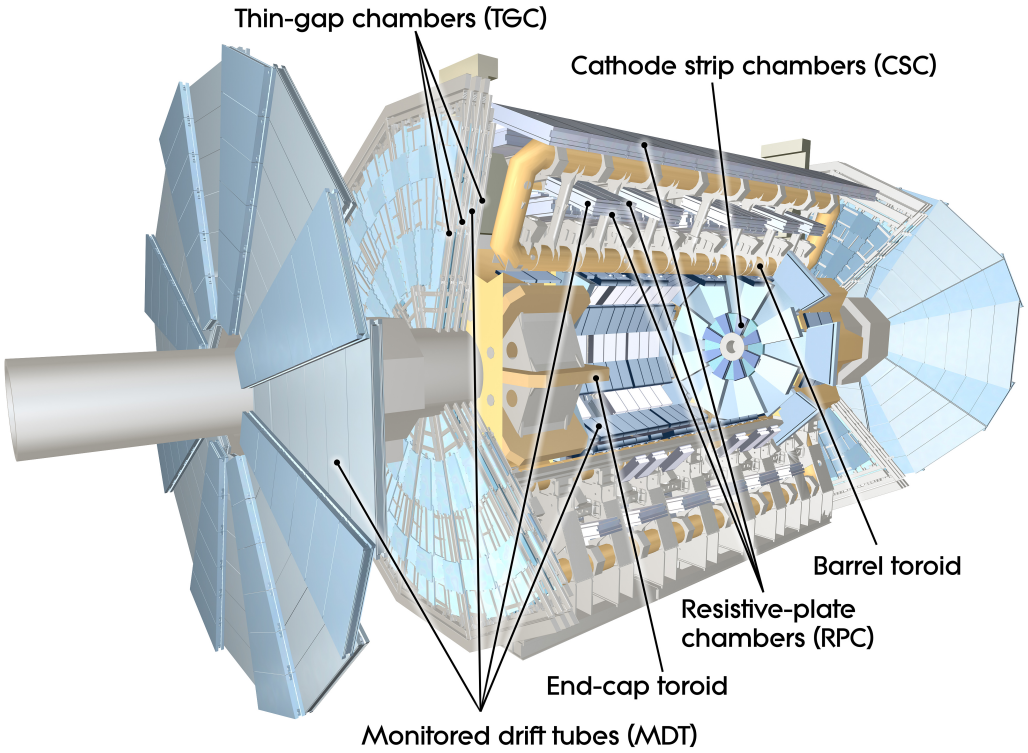


FIGURE 2.21: The ATLAS Muon Spectrometer with the various subdetectors and the magnetic system also shown.

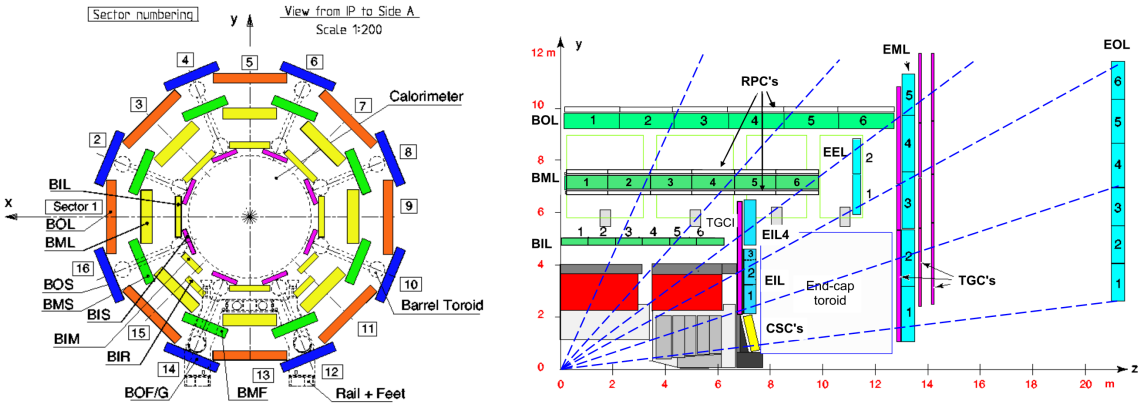


FIGURE 2.22: Cross-section of the Muon system perpendicular to the beam axis, non-bending plane, in the barrel (left-hand picture), and in a plane containing the beam axis, bending plane (right-hand picture).

2.4.4 The Muon Spectrometer

The Muon Spectrometer [76], shown in Figure 2.21, works out of measuring the magnetic deflection of muon tracks in the large superconducting air-core toroid magnets, instrumented with separate trigger and high-precision tracking chambers. The performance goal is a stand-alone transverse momentum resolution of approximately 10% for 1 TeV muons, which translates into a sagitta along the beam axis of about 500 μm to be measured with a resolution of < 50 μm.

In the barrel region, tracks are measured in three cylindrical layers of chambers arranged around the beam axis; in the transition and end-cap regions, the chambers are installed in three layers of planes perpendicular to the beam line. The cross-section of the Muon Spectrometer in both the non-bending and the bending planes can be seen is Figure 2.22.

The ATLAS Muon Spectrometer is a composite system of different subdetectors. Precision measurements of the track coordinates in the main bending direction of the magnetic field are provided by Monitored Drift Tubes (MDT) over most of the η range, while at large pseudorapidities ($2.0 < |\eta| < 2.7$) specific multi-wire proportional chambers with cathodes segmented into strip, Cathode Strip Chambers (CSC), are used in order to cope with the high rate and background conditions. The independent trigger system is based on two different technologies: Resistive Plate Chambers (RPC) in the barrel and Thin Gap Chambers (TGC) in the end-caps. It covers the pseudorapidity range $|\eta| < 2.4$ and it is used for different purposes: provide bunch-crossing identification, filter out tracks below a pre-defined p_{T} threshold, and measure the muon ϕ coordinate in the direction orthogonal to that determined by the precision-tracking chambers.

Monitored Drift Tubes These detectors provide precision momentum measurements by combining high measurement accuracy, predictability of mechanical deformations and simplicity of construction. A chamber consists of three to eight layers of pressurised drift tubes, with a diameter of 29.970 mm, operating with Ar/CO_2 gas (93/7) at an absolute pressure of 3 bar, with an average resolution of 80 μm per tube or 35 μm per chamber. The electrons resulting from ionisation are collected at the central tungsten-rhenium wire with a diameter of 50 μm at a potential of 3080 V. The wire is held in position at the tube ends by a cylindrical end-plug which guarantees the concentricity of the wire with respect to the tube with an accuracy of < 10 μm. The central conductor holding the wire also serves for the gas transfer in and out of the tube Signal transmission to the electronics and connection to the HV supply system are at opposite ends. Concerning the structure, the layer dimensions and the chamber sizes increase proportionally to their distance from the interaction point. The chambers are rectangular in the barrel and trapezoidal in the end-cap. Their shapes and dimensions were chosen to optimise solid angle coverage, while respecting the envelopes of the magnet coils, support structures and access ducts The naming of the chambers is based on their location in the barrel or end-cap (B,E), their assignment to inner, middle, or outer chamber layer (I, M, O) and their belonging to a large or a small sector (L,S). Furthermore, the sector number (1-16) and the sequence number of the chamber in a row of chambers in a sector are added to completely specify a MDT chamber. Special chambers also exist: in sector 12 and 14, to keep the acceptance losses due to the ATLAS support structure (feet) to a minimum, modified BOS chambers,called BOF or BOG, and the modified BMS chambers called BMF, are used. The BEE (Barrel End-cap Extra) are other special chambers located in the castellations of the end-cap toroid cryostats. In addition, the outer part (in the radial direction) of the EI wheel does not project into the EO wheel. In order to allow for momentum measurement in this region, an intermediate ring of chambers, the EES and EEL (the "E" stands from Extra) chambers, have been introduced with an offset of about 3-3.6 m with respect to the corresponding EI wheel chambers. All

regular MDT chambers consist of two groups of tube layers, called multi-layers, separated by a mechanical spacer. In the innermost layer of the muon detector, each multi-layer consists of four tube layers to enhance the pattern-recognition performance; in the middle and outer layer of the muon detector, each multi-layer consists of three tube layers only. The height of the support beam between the multi-layers (spacer) depends on the chamber type, varying from 6.5 mm to 317 mm. The MDT chambers cover the region of $|\eta| < 2.7$, except for the end-cap innermost layer where they are replaced by CSC within the range $2.2 < |\eta| < 2.7$.

Cathode Strip Chambers The MDT's can safely operate at maximum counting rates of about 150 $\mathrm{Hz/cm^2}$. This rate is exceeded for $|\eta| > 2.2$ in the first end-cap layers. Therefore,at a distance of ~ 7 m from the interaction point, the MDT's are replaced by Cathode-Strip Chambers, which are characterised by high spatial, time and double track resolution with high-rate (up to ~ 1000 $\mathrm{Hz/cm^2}$) capability and low neutron sensitivity. The CSC's are multi-wire proportional chambers with the wires oriented in the radial direction. Both cathodes are segmented, one with the strips perpendicular to the wire-sand provides the precision coordinate, while the other is parallel to the wires providing the transverse coordinate. The position of the track is obtained by interpolation between the charges induced on neighboring cathode strips. The CSC wire signals are not read out. In the bending direction, the readout pitch is 5.31 mm for the large and 5.56 mm for the small chambers and allows the CSC to reach a resolution of 60 μm per plane to be compared with the resolution of the MDT tube layer which corresponds to 80 μm. In the non-bending direction the cathode segmentation is coarser leading to a resolution of 5 mm. As for the MDT's, the CSC's are modulated in large and small chambers in ϕ. The entire CSC system consists of two wheels with eight chambers (eight small and eight large) each, and every chamber contains four CSC planes resulting in four independent measurements in $\eta - \phi$ along the track. This leads to a similar configuration like in the multi-layer of the MDT system, but with finer granularity. The gas mixture is Ar/CO_2 gas (80/20) and the operating voltage is 1900 V.

Resistive Plate Chambers The trigger system in the barrel consists of three concentric cylindrical layers around the beam axis, referred to as the three trigger stations. The large lever arm between inner and outer RPC's allows for a high (9-35 GeV) p_{T} trigger, while the two inner chambers provide the low p_{T} (6-9 GeV) trigger. Each station consists of two independent detector layers which measure η and ϕ of the track. The system is redundant because each track going through the three stations leaves six measurements in η and ϕ. This redundancy allows to use a coincidence system which rejects tracks and improves the trigger efficiency. The naming convention for the RPC's follows the one of the MDT's with a RPC/MDT pair in the outer layer called a station and a RPC/MDT/RPC packages in the middle layer called a superstation. A RPC trigger chamber is made of two rectangular detectors, contiguous to each other, called units. Each unit consists of two independent detector layers, called gas volumes, which are each read out by two orthogonal sets of pick-up strips. The principle of operation of the RPC is different with respect to the one of the detectors previously described. The RPC is not a wire-detector, but it is a gaseous parallel electrode-plate detector with two resistive plates, made of phenolic-melaminic plastic laminate, kept parallel to each other at a distance of 2 mm by insulating spacers. Between the plates an electric field of ~ 4.9 kV/mm allows avalanches to form along the ionising tracks towards the anode. The read out system is based on capacitive coupling to metallic strips, which are mounted on the outer faces of the resistive plates. A mixture of $C_2H_2F_4/Iso - C_4H_{10}/SF_6$ (94.7/5/0.3) is used as a gas. The RPC's can operate in both

avalanche and streamer mode, but given the LHC high rate environment the avalanche mode was selected because of its higher rate capability.

Thin Gap Chambers These detectors are part of the trigger system but they also provide a second, azimuthal coordinate to complement the MDT's measurements in the bending direction. The TGC's are multi-wire proportional chambers with a wire-to-cathode distance of 1.4 mm, smaller than the wire-to-wire distance of 1.8 mm. This geometry, if used together with a highly quenching gas mixture of CO_2 and $n - C_5H_{12}$ (n-pentane), allows for a gain of $\sim 3 \cdot 10^5$, which means quasi-saturated mode. The high electric field around the TGC wires and the small wire-to-wire distance lead to very good time resolution for a large majority of the tracks. Concerning the mechanical structure, the seven detector layers in the middle stations (big wheels) are arranged in one triplet and two doublets. The triplet is to cope with false coincidences from background hits, which mainly occurs in the end-cap region. The trigger detectors are assembled in two concentric rings: outer (or end-cap) ring covering the pseudorapidity region $1.05 \leq |\eta| \leq 1.92$ and an inner (or forward) ring covering $1.92 \leq |\eta| \leq 2.4$. Moreover, the TGC's are redundant and thus the fakes are rejected and the efficiency is improved.

2.4.5 Forward Detectors

In the forward regions of the ATLAS experiment, three detectors were installed at different stages:

- LUCID (LUminosity measurement using Cherenkov Integrating Detector) is placed at $\pm$ 17 m from the interaction point and it is the main online relative-luminosity monitor for ATLAS;
- ALFA (Absolute Luminosity For ATLAS) is located at $\pm$ 240 m, it consists of scintillating fibre trackers positioned inside Roman pots which are designed to approach as close as 1 mm to the beam;
- Zero-Degree Calorimeter (ZDC), at $\pm$ 140 m from the interaction point where the common beam pipe divides back in two independent beam pipes, plays a key role in determining the centrality of heavy-ion collisions. Each ZDC module consists of layers of quartz rods and tungsten plates which measure neutral particles at $|\eta| > 8.2$.

2.4.6 Trigger and Data Acquisition System

The proton-proton interaction rate at the design luminosity of $10^{34}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ is approximately 1 GHz, while the event data recording, based on technology and resource limitations, is limited to about 1 kHz [77].This requires an overall rejection factor of about 10^6 against minimum-bias processes while maintaining maximum efficiency for the interesting physics.

The ATLAS trigger system is an essential component of the experiment as it is responsible for deciding whether or not to keep an event from a given bunch-crossing interaction for later study by so achieving the desired rejection factor.

A comprehensive description of the Run 1 Trigger and Data Acquisition (TDAQ) system of the ATLAS Experiment can be found in Refs. [78, 79], while the performance of the Run 2 TDAQ system are detailed in Ref. [77].

The TDAQ system, which is schematically shown in Figure 2.23, consists of a hardware-based first-level trigger (L1) and a software-based high-level trigger (HLT). The L1 trigger decision is formed by the Central Trigger Processor (CTP), which receives inputs from

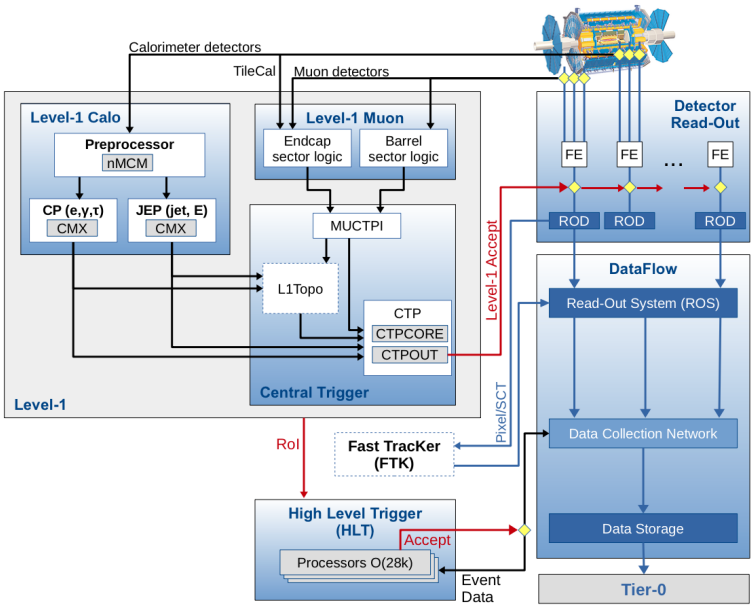


FIGURE 2.23: The ATLAS Trigger and Data Acquisition system in Run 2 with emphasis on the components relevant for triggering. Ref. [77].

the L1 calorimeter (L1Calo) and L1 muon (L1Muon) triggers as well as several other subsystems such as the Minimum Bias Trigger Scintillators (MBTS), the LUCID Cherenkov counter and the Zero-Degree Calorimeter (ZDC). After the L1 trigger acceptance, the events are buffered in the Read-Out System (ROS) and processed by the HLT. The HLT receives Region-of-Interest (RoI) information from L1, which can be used for regional reconstruction in the trigger algorithms. After the events are accepted by the HLT, they are transferred to local storage at the experimental site and exported to the Tier-0 facility at CERN's computing centre for offline reconstruction.

In Run 1 the HLT consisted of separate Level-2 (L2) and Event Filter (EF) farms. For Run 2 these were merged into a single homogeneous farm allowing better resource sharing and overall simplification on both the hardware and software sides.

Trigger

The L1 trigger system uses a subset of the total detector information to make a decision in less than 2.5 μs on whether or not to continue processing an event, reducing the data rate to approximately 100 kHz. The HLT provides the reduction to a final data-taking rate of approximately 1 kHz with an event size of approximately 2 Mbyte. High p_{T} muons are identified at L1 trigger by using the information of the trigger chambers in the Muon Spectrometer, while calorimeter selections (electrons, photons, jets, missing energy) are based on reduced-granularity information from all the calorimeters. L1 results are then processed by the central trigger processor which uses a trigger menu made up of combinations of trigger selections. To have an optimal use of the bandwidth when luminosity and background conditions change, a pre-scaling of trigger menu items is also available. Events passing the L1 trigger selection are transferred to the detector-specific electronics and to the data acquisition system. The L1 trigger defines, for each event, Regions-of-Interest (RoI's) which give the $\eta - \phi$ coordinates of the detector regions where the trigger identified interesting features. This information is then used by the high-level trigger. The

L2 selection is seeded by the RoI information provided by the L1 trigger: at full granularity and precision, all the available detector data within the RoI's is used. The L2 menus reduce the trigger rate to $\sim$ 3.5 kHz, with an average event processing time of about 40 ms. The final stage of the event selection is carried out by the event filter, which reduces the event rate to roughly 1 kHz with offline analysis procedures within an average event processing time of the order of four seconds.

Minimum-bias triggers Studies of the total cross-section and involving hadronisation, diffraction, hadrons containing strange quarks and other non-perturbative properties of pp interactions require the use of a high-efficiency trigger for selecting all inelastic interactions that result in particle production within the detector. The Minimum Bias Trigger Scintillators (MBTS), positioned on the front face of the end cap calorimeter cryostats, are used as the primary L1 hardware triggers for recording inelastic events by requesting a minimal set of conditions (minimum bias trigger) [80–82]. The plastic scintillation counters composing the system (which were replaced during LS1) consist of two planes of twelve counters, each plane formed of an inner ring of eight counters and an outer ring of four counters. These rings are sensitive to charged-particles in the interval $2.07 < |\eta| < 3.86$. Each counter is connected to a photomultiplier tube and provides a fast trigger via a constant fraction discriminator and is read out through the Tile calorimeter data acquisition system. The MBTS triggers require a certain multiplicity of counters to be above threshold in a bunch-crossing with colliding beams. The `L1_MBTS_1` and `L1_MBTS_2` triggers require any one or two of the 24 counters to be above threshold, respectively. The coincidence of two hits in the latter suppresses beam-induced backgrounds from low-energy neutrons and photons. The `L1_MBTS_1_1` trigger requires at least one counter to be above threshold in both the $+z$ and $-z$ hemispheres of the detector and is used to seed the high-multiplicity HLT triggers. The same trigger selections are also applied to empty (no beam present) and unpaired (one beam present) beam-crossings to investigate beam-induced backgrounds. No additional HLT selection is applied to `L1_MBTS_1` and `L1_MBTS_2` triggered events. The `mb_sptrk` trigger is used to determine the efficiency of the MBTS. It is seeded using a random trigger on filled bunches and requires at least two reconstructed space-points in the Pixel system and three in the SCT, along with at least one reconstructed track with $p_{\mathrm{T}} > 200$ MeV.

The MBTS minimum bias trigger is highly efficient, even for events containing only two charged-particles with $p_{\mathrm{T}} > 100$ MeV and $|\eta| < 2.5$. As previously mentioned, the primary minimum-bias and high-multiplicity data set at the centre-of-mass energy of 13 TeV was recorded in June 2015. The average number of interaction per bunch crossing varied between 0.003 and 0.03, and the interaction rate had a maximum of about 15 kHz. More than 200 million interactions were recorded during a one-week data-taking period. Most of the readout bandwidth was dedicated to the loosest `L1_MBTS_1` trigger recording events at 1.0 to 1.5 kHz on average.

Data Acquisition

After an event is accepted by the L1 trigger, the data from the pipe-lines are transferred from the detector to the Readout Drivers (ROD's), which are detector-specific functional elements of the front-end systems. ROD's achieve a higher level of data concentration and multiplexing by gathering information from several front-end data streams. Digitised signals are formatted as raw data prior to being transferred to the Data Acquisition (DAQ) system. The ROD's follow some general ATLAS rules, for instance the definition of the event data format. The readout system, which is the first stage of the DAQ, receives and stores the data in local buffers. Events selected by the L2 trigger are transferred to the event-building system and then to the event filter for the final selection. At this stage, the

data are moved to a permanent storage at the CERN computing centre. The data acquisition system also takes care of the configuration, control and monitoring of the hardware and software components which provide the data-taking functionality. The DCS allows for a coherent and safe operation of the ATLAS detector hardware and provides a homogeneous interface to the technical infrastructure and the subdetectors of the experiment. In addition, it controls the operational parameters, such as high and low voltages, gas, temperature, etc, and allows for automatic or manual recovery actions. The synchronization of the state of detector with the data-taking as well as the communication between subdetectors and other systems such as the LHC, is also done via DCS.

2.4.7 Computing and Software

The ATLAS Computing Model

The ATLAS Computing Model [83] is based on the Grid [84] paradigm and on a high degree of decentralization and sharing of computing resources: off-site facilities are vital to the operation of ATLAS in a way that was not the case for previous CERN-based experiments.

The CERN Tier-0 facility is the place where the primary event processing occurs. The RAW data are archived at CERN and copied to the Tier-1 facilities around the world. The Tier-1 provides also the reprocessing capacity, the access to the data in various processed versions, and allows scheduled analysis of the processed data by physics analysis groups.

Derived datasets produced by the physics groups are copied to the Tier-2 facilities for further analysis. The Tier-2 facilities also provide the simulation capacity for the experiment, with the simulated data housed at Tier-1s. In addition, Tier-2 centers will provide analysis facilities, and some will provide the capacity to produce calibration sets based on raw data processing. Additional computing resources will be available for data processing and analysis at Tier-3 centers and other computing facilities to which ATLAS may have access. The structure of the ATLAS Tiers is shown in Figure 2.24. A CERN Analysis Facility provides an additional analysis capacity, with an important role in the calibration and algorithmic development work.

Event Data Model Large collaborations such as ATLAS have the necessity to put in place common interfaces and data objects to insure easy maintenance and coherence of the experiments software platform over long periods of time. The ATLAS Event Data Model [85] improves commonality across the detector subsystems and subgroups such as trigger, test beam reconstruction, combined event reconstruction and physics analysis. Moreover, it fosters the use of common software between online data processing and offline reconstruction. The ATLAS detector produces up to one PByte of data per year, a vast amount of information which prohibits the distribution of raw data to collaborators around the World. To enable physicists to analyse the data at remote sites two additional stages of datasets were introduced in Run 1:

- The Event Summary Data (ESD) which contains the detailed output of the detector reconstruction and will be produced from the raw data. It will contain sufficient information to allow particle identification, track re-fitting, jet calibration etc. thus allowing for the rapid tuning of reconstruction algorithms and calibrations. The size per event is about 500 kB.
- The Analysis Object Data (AOD) which is a summary of the reconstructed events, and contains sufficient information for common analyses. The size per event is about 100 kB.

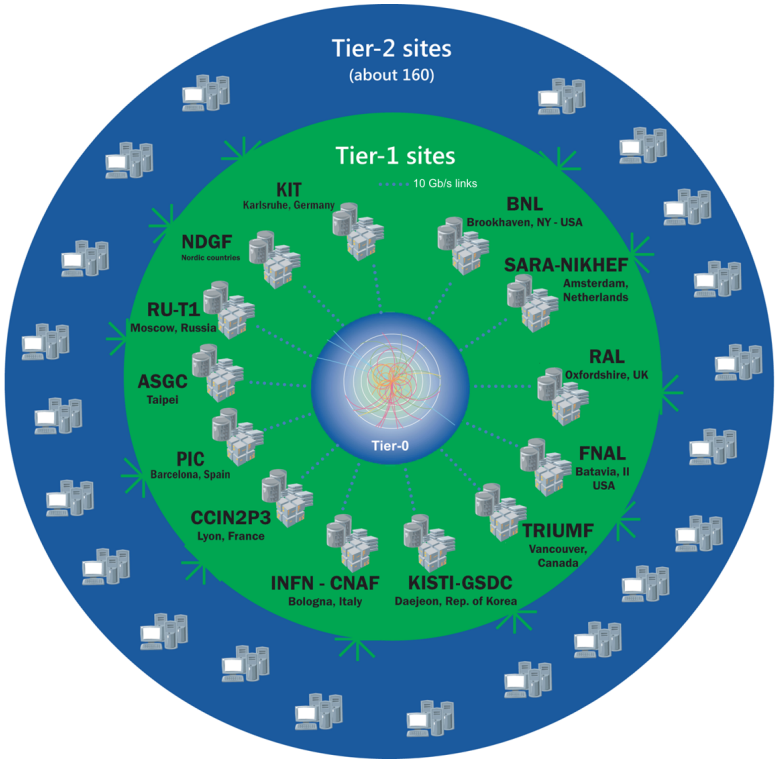


FIGURE 2.24: Structure of the ATLAS Tiers as it was in June 2014.

The design of the Run 1 EDM had some limitations. In particular, the data structures used were complicated and relied on expensive C++ [86] features such as virtual inheritance. During LS1, ATLAS redesigned the event data model for Run 2 to simplify it and make it more easily readable directly from ROOT [87], which was not possible with the pre-existing EDM. The implementation is based on the concept of an *auxiliary store*, which stores object data in a set of vectors separate from the objects themselves. Additional features include allowing user analysis code to add additional information to objects, on-demand reading of parts of objects, and shallow copies. This data model is named xAOD and it is being extensively used during the Run 2 data-taking. To streamline the data and to introduce specific information which might not be available in the general xAOD, and at the mean time to keep the size of the sample small enough, xAOD *derivations*, DxAOD, can be produced.

The ATLAS Offline Software

The main aims of the ATLAS Offline Software are:

- to process events delivered by the ATLAS trigger and data acquisition system;
- to deliver the processed results to physicists within the ATLAS Collaboration;
- to provide tools for them to analyse the processed information and produce physics results.

The complexity and scale of ATLAS mean that the software must be highly modular and robust, and must furthermore be flexible enough to meet the needs of the experiment throughout its operational lifetime.

The ATLAS software is focused on an object-oriented approach, based primarily on the C++ programming language, but with some components implemented using FORTRAN

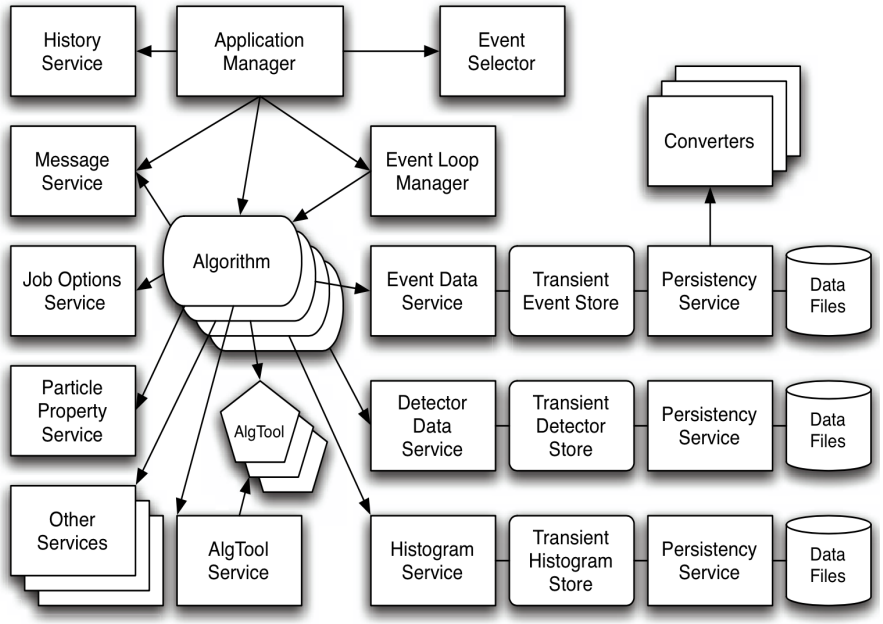


FIGURE 2.25: The major components that have been identified within the Athena architecture. Ref. [83].

[88] and Java [89]. A component-based model has been adopted, whereby applications are built up from collections of plug-compatible components based on a variety of configuration files. This capability is supported by a common framework that provides common data-processing support. This approach results in great flexibility in meeting the basic processing needs of the experiment, and also for responding to changing requirements throughout its lifetime. The heavy use of abstract interfaces allows for different implementations to be provided, supporting different persistency technologies, or optimized for the offline or high-level trigger environments.

The LHCb experiment originally developed the Gaudi framework [90], which became a common LHCb-ATLAS project. The current Athena framework used in ATLAS is an enhanced version of Gaudi of which it maintains the basic design choices (like separation between data and algorithmic code) and whose abstract interfaces to user defined objects such as Algorithm, Tool and Service allowed maximum flexibility in developing both a range of shared components and components that are specific to the experiments and better meet their particular requirements. In ATLAS, all the data processing, from high-level trigger to event simulation, reconstruction and analysis, takes place within the Athena framework. In this way it is easier for code developers and users to test and run algorithmic code, with the assurance that all geometry and conditions data will be the same for all types of applications (simulation, reconstruction, analysis, visualization). The Athena component model is shown in Figure 2.25.

For an easier understanding of specific terms used in the following chapters, few crucial terms are summarised below, while a more comprehensive description of the ATLAS software releases can be found in Ref. [91] and references therein.

- An independently developed piece of software providing a well-defined functionality in form of libraries, applications and data files is referred to as a *package*.
- Typically the source code of a package resides in a specific area in a code repository (e.g. SVN). To uniquely define a specific version of a package in the code repository, a reference called *tag* is used.

- A combination of hardware and software tags which identifies the type of binaries to be built is a *platform* (e.g. Linux slc6, gcc48, etc.).
- A *release* is a set of tagged packages built together for a number of platforms, providing some well defined and documented functionality for end users and made available for distribution. The release policy should provide a good balance between software development and stability for end users. Major and minor releases are built. Major releases may contain important changes in the architecture of software, new functionality changes in packages, database schema changes and other similar changes which require action from end users. Minor releases may contain some internal changes which do not require code changes at the user's side. Regular nightly builds are performed on a shared file system during the night hours, so every day developers can access the most recent versions of all packages built together as nightly releases. Nightly builds are the main arena for the integration of new developments and the target for regular automatic validation tests.

2.4.8 Upgrade Programme

From January to May 2017, the LHC will undergo a short technical stop to prepare for the final phase of Run 2, which is supposed to last until Fall 2017, when the LHC will enter a three-year Long Shutdown 2. The ATLAS Experiment is facing a continuous upgrade program aimed at coping with the high luminosity scenarios, which will be characteristic of an upgraded LHC during Run 3 and beyond (2020).

One of the key Phase I upgrades consists in the replacement of the first end-cap station in the ATLAS muon system during the second Long Shutdown. The muon detectors in the end-cap regions are arranged into symmetric configurations around the z-axis, the so-called Big and Small (MDT's, TGC's and CSC's) Wheels. Due to the expected overwhelming background rates at high η values foreseen for the ATLAS Phase 2, the detector technologies being used during the Run 1 and Run 2 of the LHC in the Small Wheels will have to be replaced and new high-precision and radiation-hard detectors must be developed. The new detectors, based on the Micro MEsh GAseous Structure (MicroMeGaS's) and small Thin Gap Chambers (sTGC's) technologies, will be again assembled to build a wheel: for this reason, the upgrade programme is called New Small Wheel. These detectors will be operated at very high rates (~ 15 kHz/cm^2) with a good resolution at the level of 100 μm. The project is described in detail in the Appendix A.

The High-Luminosity LHC, which will accelerate proton beams at the centre-of-mass energy of 14 TeV, is expected to begin operation in the second half of 2026 and to achieve instantaneous luminosities of the order of $L = 7.5 \cdot 10^{34}\mathrm{cm}^{-2}\mathrm{s}^{-1}$, corresponding to an average number of simultaneous interation of $< \mu > \approx 200$ [92, 93]. To cope with the high pile-up and radiation doses as well as with the data rate, the current ATLAS Inner Detector will be completely replaced during the third long shutdown by an all-silicon Inner Tracker (ITk). ITk is being designed to yield at least 13 points per charged-particle in the central region, as well as to provide uniform coverage in the forward region where the design requirements are still evolving [94]: in any scenario, ITk will cover the pseudorapidity region of $|\eta| < 4.0$, providing tracking and vertexing capabilities in the forward region, and thereby enabling to reject forward pile-up jets, whose tracks mostly originate from pile-up vertices rather than from the hard-scatter vertex. The extended η coverage of the ITk will have a strong impact on physics measurements, such as vector boson fusion (VBF) and vector boson scattering (VBS) because these processes are characterised by forward jets in the pseudorapidity region $|\eta| > 2.5$. To be ready for the HL-LHC, during ATLAS Phase II other upgrades are also foreseen. A Level 0 trigger scheme will be implemented, the

Level 1 trigger scheme will be optimised and the electronics for the muon and calorimeter systems will be replaced.

Chapter 3

Inner Detector Tracking Performance

Why do I write?
Out of fear.
Out of fear that the memory of the people I write about might go lost.
Out of fear that the memory of myself might get lost.
Or even just to be shielded by a story,
to slip inside a story and
stop being recognizable,
controllable,
subject to blackmail.

-Fabrizio de André-

The reconstruction of charged-particle trajectories is a fundamental aspect of modern high energy physics experiments and it is, in general, a very complex task that puts stringent requirements onto the software realisation. In this Chapter, a brief review of the improvements to the track reconstruction algorithms, in particular for low momentum tracks, which were implemented during the Long Shutdown 1, is given. In addition, a first study on the performance of the ATLAS Inner Detector with Run 2 data at 13 TeV is reported together with a set of benchmark values to be applied to Run 2 analyses using objects based on tracks.

3.1 Introduction

The Run 2 of the LHC, as described in detail in Chapter 2, began in June 2015, after the two years of Long Shutdown 1 during which the Pixel detector was refurbished and improved by inserting the IBL [67]. Addind the IBL to the ID significantly improved the track reconstruction performance, in particular for what concerns the reconstruction of low momentum tracks and both the transverse and longitudinal impact parameter [1] resolution [95].

3.2 Track Reconstruction

The reconstructed trajectories of charged-particles, known as tracks, are used as input to the identification of a wide range of final-state physics objects, from charged leptons to converted photons decaying into pairs of electrons, and to determine precisely the location of both primary vertices reconstructed along the beam line and secondary vertices significantly displaced from the beam line.

[1] The transverse impact parameter, d_0, is defined as the shortest distance between a track and the beam line in the transverse plane – this defines the *point of closest approach*. The longitudinal impact parameter, z_0, is defined as the distance in z between the primary vertex and the point on the track which corresponds to the closest approach in the transverse plane.

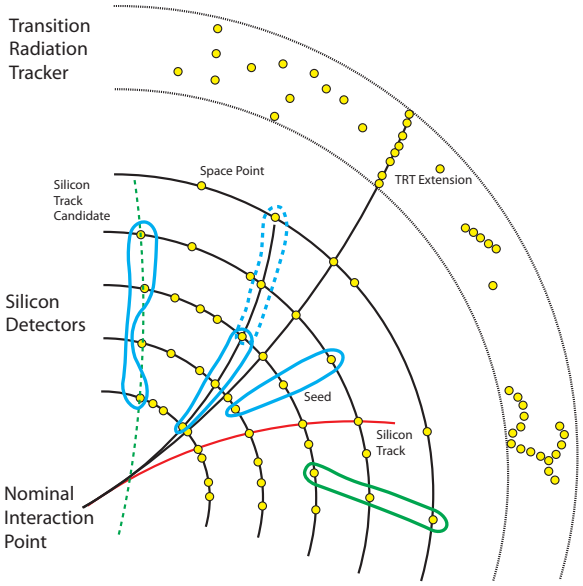


FIGURE 3.1: Sketch of the different elements used to reconstruct the track of a charged particle within the ID.

In the ATLAS Experiment, the ID is responsible for the track measurements [96, 97]: this is a multi-step process, utilising early candidate rejection in order to single out efficiently a high-purity sample of charged particle tracks with precisely determined parameters. Charged particles which travel in a (constant) solenoidal magnetic field follow a helical path, which can be represented by five parameters with respect to a given surface. In the ATLAS track reconstruction, these parameters, schematically represented in Figure 3.2, are chosen to be: the transverse and longitudinal impact parameters (z_0 and d_0 respectively), the transverse and azimuthal angles (θ and ϕ respectively), and the curvature parameter (q/p, where q represents the electric charge of the particle and p its momentum). The algorithms used for the various steps of the track reconstruction with the ATLAS Inner Detector have been optimised prior to the beginning of Run 2 of the LHC [98] in order to allow them to cope with the anticipated increase of pile-up collisions up to $< \mu > \sim 40$.

In the track reconstruction process, it is needed to identify patterns which are consistent with the path of a charged-particle traversing the ID. This is referred to as *Pattern recognition* and it is based on several steps starting from combining space points to create track seeds which are then used to start building track candidates. These elements can be seen in Figure 3.1.

The "inside-out" sequence [100] begins in the silicon layers. Space points, shown as yellow markers in 3.1, are created by combining energy depositions, *hits*, in the Pixel (where a space point is simply a pixel cluster) and in the SCT modules (where strip clusters on either side of a module are combined using the stereo angle between the microstrips as ameans of calculating where along the strip direction the charged particle crossed into the module). Finally, an extension to the TRT is performed and the hit collection is fitted to calculate the final track parameters.

A collection of three space points is then identified as a *track seed*. It provides a first estimate of candidate particle trajectories. Considering that each charged particle going through the ID creates more that 4 space points and given the average charged particle multiplicity (details in Chapter 4) of an event at the LHC Run 2, the number of possible seeds is much higher than the number of real tracks because of combinatorial effects. Thus, in order to reduce the seed sample, it is highly important to reject as soon as possible those

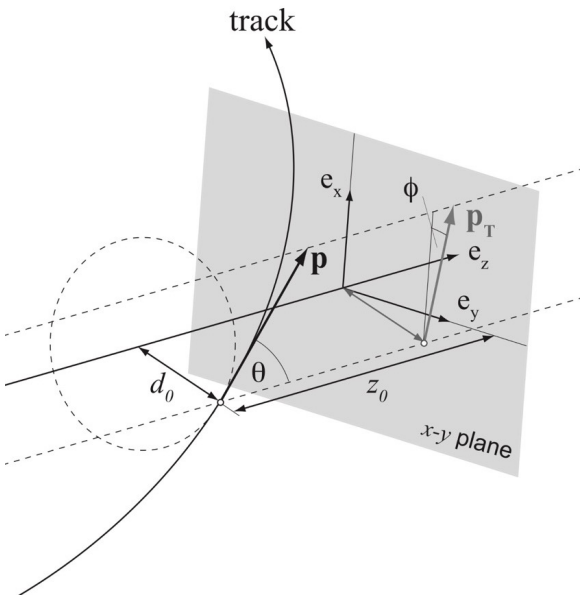


FIGURE 3.2: The perigee representation expressed in the ATLAS track parameterisation. The local expression of the point of closest approach is given by the signed transverse impact parameter d_0 and the longitudinal impact parameter z_0. The momentum direction is expressed in global coordinates using the azimuthal angle ϕ that is defined in the projected $x-y$ plane and the polar angle θ, which is measured with respect to the global z axis. Ref. [99].

seeds which are not physically compatible with a certain charged particle trajectory. The procedure starts from the case of three space points all located in the SCT detector (SSS case). The SSS case presents the best purity, i.e. the highest efficiency to find a good track. The resulting track seed is then extrapolated to the beamline and the intersection is used to perform a fast one dimensional vertex-finding. A search region is then defined by establishing the maximum and minimum vertex z-positions. Cases with seeds in the Pixel detector (PPP) or in either Pixel and SCT detectors (PSS, PPS), which cannot be extrapolated back to the identified search region are rejected at this stage. On the other hand, the accepted seeds are processed further, to be turned into so-called track-candidates. If a fourth space point can be found in the layer before or after the space points used in the seed and if this is compatible with the initial trajectory estimate, then priority is given to this track-candidate to be processed first. The initial seed parameters determine a search region and hits are located within this region. If a sufficient number of hits not already associated with an existing track candidate is found, then a Kalman filter [101] is applied in order to identify compatible hits, which are added to the track candidate: the track parameter estimation is updated accordingly.

At this point, the track candidate collection still contains tracks arising from combinatorial effects (usually referred to as "fake tracks" or simply "fakes") along with track duplicates, which can be identified by measurements that they share with other track candidates. Thus, an ambiguity solving step is performed. This can remove track candidates at an early stage if they do not match pre-defined quality criteria, such as a minimum number of hits, minimum transverse momentum, etc.

After that, the ambiguity solving algorithm assigns a score to each track candidate: positive scores are applied for unique measurements and good fit quality, while penalties are introduced for missing measurements (where they would be expected) or shared

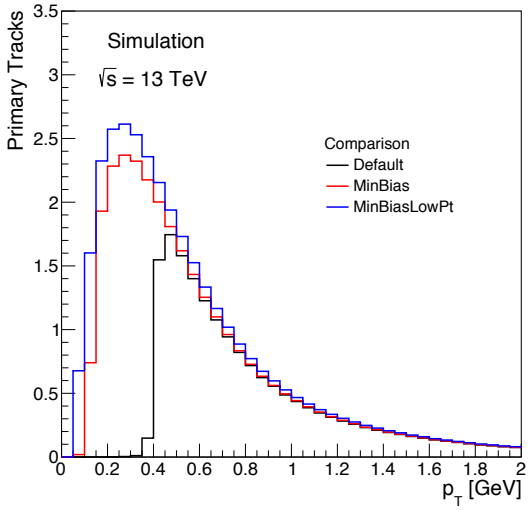


FIGURE 3.3: Primary Tracks as a function of the transverse momentum p_{T} for three different setups: Default, MinBias, MinBiasLowPt.

measurements with other track candidates (determined by the number of hits in each sub-detector).

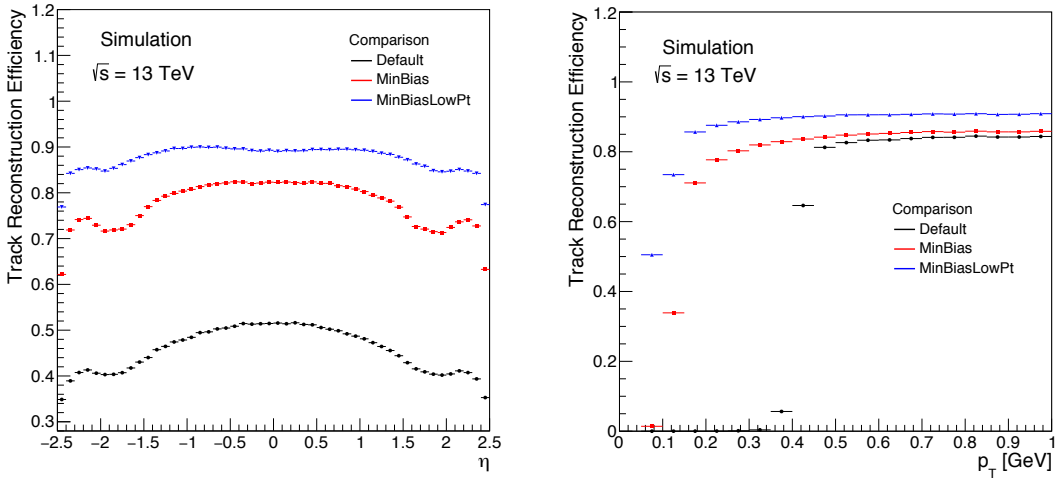
Tracks that successfully pass the ambiguity solving stage and are within the coverage of the TRT detector are then extended into the TRT and completed with TRT measurements. A successful TRT extension increases the momentum resolution significantly by exploiting the longer lever arm for field integration.

This silicon-seeded inside-out track reconstruction algorithm produces the majority of tracks in most ATLAS events. A second pass TRT-seeded reconstruction is then performed using drift circles and silicon hits which were not assigned to any track, in order to reconstruct tracks which were missed by the first pass.

Since these tracks are most likely the result of photon conversions, back-tracking is only performed in Regions of Interest around energy deposits in the Electromagnetic Calorimeter. Similarly, additional tracking passes can be performed on e.g. unassigned pixel hits to produce stand-alone pixel track segments, usually referred to as *tracklets* which are useful for specific physics analyses, as will be described in Chapter 4.

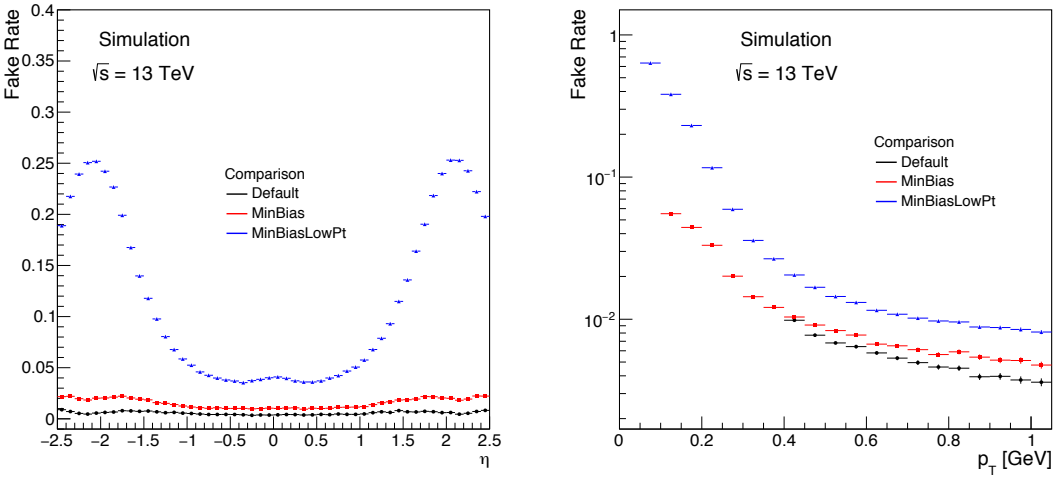
3.3 Low-p_{T} reconstruction

Studies performed in the context of the charged particle multiplicity measurement at $\sqrt{s} =$ 13 TeV [81] motivated various modifications to the track reconstruction. Most significantly, the minimum transverse momentum for tracks to be reconstructed by a single pass of the inside-out sequence was lowered from 500 to 100 MeV and the number of silicon hits was lowered accordingly from 7 to 5. To determine the optimal setup for the track reconstruction, different configurations of the track reconstruction algorithms were tried by studying effects on the efficiency and the fake rate. The track reconstruction efficiency is evaluated from simulation using a hit-based track-to-truth particle association to associate reconstructed tracks to simulated primary particles [80]. Each of the clusters is connected to the truth particle which has the largest energy deposition in the MC simulation. The clusters are then weighted according to their importance in the track reconstruction: if the clusters are from the pixel detector (including IBL) the weight is set to 10 for each cluster. If they are from the SCT, the weight is set to 5 and if they are from the TRT, the weight is set to 1. A



(a) (b)

FIGURE 3.4: Track Reconstruction Efficiency in three different setups: Default, MinBias, MinBiasLowPt as a function of (a) pseudorapidity η and (b) transverse momentum p_{T}.



(a) (b)

FIGURE 3.5: Fake Rate in three different setups: Default, MinBias, MinBiasLowPt as a function of (a) pseudorapidity η and (b) transverse momentum p_{T}.

weighted matching probability P_{match} can be defined using the ratio of the number of hits which are common to a given track and the corresponding truth particle ($N^{\mathrm{common}}_{\mathrm{Pixel,\ SCT,\ TRT}}$) and the number of hits which form the track ($N^{\mathrm{track}}_{\mathrm{Pixel,\ SCT,\ TRT}}$):

$$P_{\mathrm{match}} = \frac{10 \cdot N^{\mathrm{common}}_{\mathrm{Pixel}} + 5 \cdot N^{\mathrm{common}}_{\mathrm{SCT}} + 1 \cdot N^{\mathrm{common}}_{\mathrm{TRT}}}{10 \cdot N^{\mathrm{track}}_{\mathrm{Pixel}} + 5 \cdot N^{\mathrm{track}}_{\mathrm{SCT}} + 1 \cdot N^{\mathrm{track}}_{\mathrm{TRT}}}. \tag{3.1}$$

The hit-matching technique is very robust since it exploits the relation between tracks

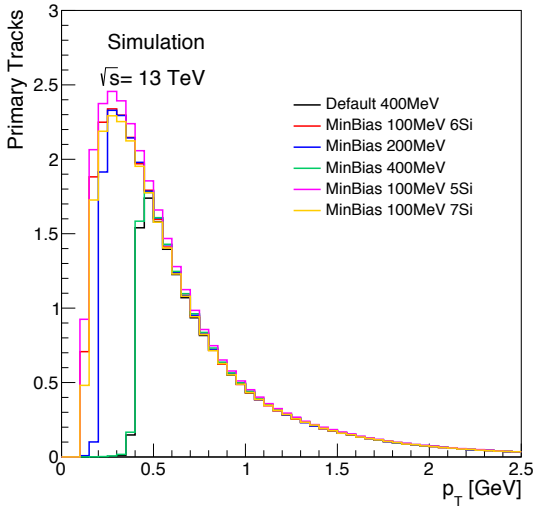
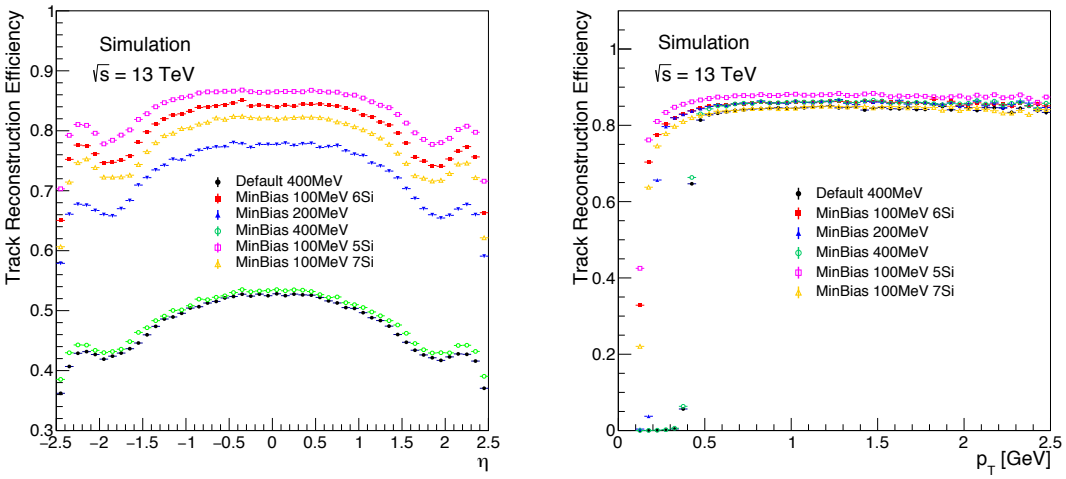


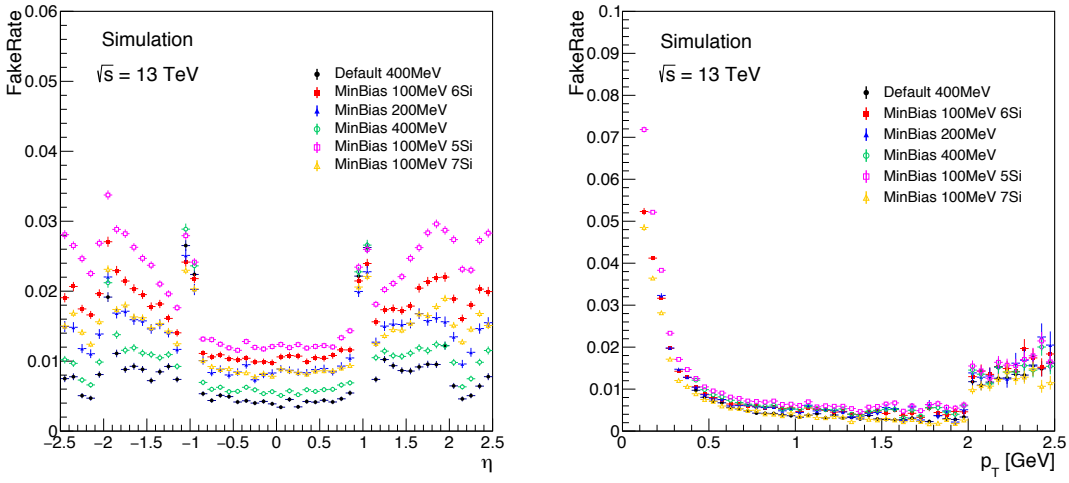
FIGURE 3.6: Primary Tracks as a function of the transverse momentum p_{T} for six different setups: Default (400 MeV and 7 Silicon Hits), MinBias (400 MeV and 7 Silicon Hits, 200 MeV and 7 Silicon Hits, 100 MeV and 5, 6 or 7 Silicon Hits).



(a) (b)

FIGURE 3.7: Track Reconstruction Efficiency in two different setups, namely Default and MinBias with different requirements on the minumum transverse momentum and on the number of silicon clusters. The distributions are shown as a function of (a) pseudorapidity η and (b) transverse momentum p_{T}.

and truth particles at the GEANT4 energy deposition level and is mostly independent of the detector resolution. A reconstructed track with $P_{\mathrm{match}} > 0.7$ (unless differently stated) is tagged as a primary track if it emerges from the Monte Carlo primary vertex and as a secondary track if it is the result of simulated physics interactions within the detector.



(a) (b)

FIGURE 3.8: Fake Rate in two different setups, namely Default and MinBias with different requirements on the minumum transverse momentum and on the number of silicon clusters. The distributions are shown as a function of (a) pseudorapidity η and (b) transverse momentum p_T.

The track reconstruction efficiency $\epsilon_{trk}(p_T, \eta)$ is then measured as a function of p_T and η with the following definition:

$$\epsilon_{trk}(p_T, \eta) = \frac{N_{rec}^{matched}(p_T, \eta)}{N_{gen}(p_T, \eta)} \tag{3.2}$$

where p_T and η are truth particle properties, $N_{rec}^{matched}(p_T, \eta)$ is the number of reconstructed tracks matched to truth charged particles and $N_{gen}(p_T, \eta)$ is the number of truth charged particles in that (p_T, η) range.

The fake rate, $r_{fake}(p_T, \eta)$ is measured as a function of p_T and η with the following definition:

$$r_{fake}(p_T, \eta) = \frac{N_{rec}^{unmatched}(p_T, \eta)}{N_{gen}(p_T, \eta)} \tag{3.3}$$

Figure 3.3 compares the momentum distributions of primary tracks, in Monte Carlo simulation as reconstructed with different setups. As it can be noticed in Figure 3.4, the *Default* setup requires the p_T of the track to be larger than 400 MeV. The *MinBias* setup, applies a p_T threshold of 100 MeV. The *MinBiasLowPt* setup is based on a p_T threshold of 50 MeV. In all these setups, the required minimum number of silicon hits is 6.

While the Default and the MinBias setups are based on a single-pass of the previously described inside-out track reconstruction algorithm, to reconstruct tracks down to 50 MeV in the MinBiasLowPt setup, a second-pass is applied for tracks with $p_T < 100$ MeV. Figures 3.4(a) and 3.4(b) show, respectively, the track reconstruction efficiency as a function of the pseudorapidity η and the transverse momentum p_T. A minimum p_T of 50 MeV is always required for the generated truth particles: it is then clear that the efficiency, being p_T-dependent is much lower for those setups which imply higher reconstructed p_T thresholds, such as Default and MinBias. In general, the higher the p_T threshold the higher is the track reconstruction efficiency, ϵ_{trk}, as effects introduced by e.g. multiple scattering or dead material become less severe. As a function of pseudorapidity, the efficiency is lower

in the region $|\eta| > 1$ due to particles passing through thicker material. The slight increase in efficiency at $|\eta| \sim 2.2$ is due to particles passing through a larger number of layers in the end-cap Pixel regions of the ID. A high reconstruction efficiency is not the only parameter to be taken into account when deciding which setups are good to be used for certain scopes. Another important parameter is the rate of fake tracks $rate_{\text{fakes}}$, shown in Figures 3.5(a) and 3.5(b) as a function of the pseudorapidity η and the transverse momentum p_{T}. For the Default setup, which has the highest p_{T} threshold, the rate of fake tracks remains always below the level of 1%, while for the MinBias setup it is of the order of 2% over the whole η range. The worst scenario is represented by the MinBiasLowPt setup. In this case, the fake rate reaches peaks of $\sim 25\%$ for $|\eta| > 1.5$, maily due to the very low p_{T} tracks as it can be seen in Figure 3.5(b) where the fake rate rises up to 60% in the first p_{T} bin for tracks with transverse momentum in the range [50, 100] MeV: the effect of multiple scattering and passive material is so large that the majority of tracks originates from random combinations of hits and is characterized by very low P_{match}, thus it is correctly classified as fakes. In addition, Figure 3.5(b) shows that also when considering tracks with p_{T} of about 1 GeV, the fake rate is almost a factor of 2 higher than for the other two setups.

Therefore, after the analysis of the track reconstruction efficiency and the rate of fake tracks in the three setups, the MinBiasLowPt setup was discarded owing to the large fake rate, which resulted particularly high in the very low p_{T} region and made the reconstruction setup not suitable for physics studies.

On the other hand, the Default and MinBias setups were further investigated. Different transverse momentum thresholds (400, 200, 100 MeV) and number of silicon hits (5, 6, 7) required at reconstruction level were tested with the MinBias setup, as shown in Figure 3.6. Six new setups, only based on a single-pass of the main track reconstruction algorithm, were established and the corresponding tracking performance figures were studied.

As it can be seen in Figures 3.7 and 3.8, the setup which shows the highest track reconstruction efficiency by keeping, at the same time, the fake rate below 3% in the entire η range, is the MinBias setup with a p_{T} threshold of at least 100 MeV and at least 5 Silicon clusters.

This setup was then chosen for the Run 2 ATLAS track reconstruction down to 100 MeV and it was used for all the tracking-related studies based on samples of minimum bias events.

3.4 Track Reconstruction Performance in Run 2

In June 2015, for the first time the LHC delivered proton-proton collisions at a center-of-mass energy of 13 TeV, the highest of any collider to date. The performance of the ATLAS track reconstruction algorithm was checked in data and compared to Monte Carlo simulations, as extensively described in [102]. Among the comparisons studied by the Author of this thesis are the pure hit content as well as more detailed performance measurements like the SCT and TRT extension efficiency in which the probability of extending a track segment respectively in the SCT and TRT detector is measured.

A measurement of the track impact parameter resolution and its comparison to simulation will be reported for completeness.

The properties of hits on track were also studied in the core of high momentum jets and details can be found in [102].

3.4.1 Data and Simulated samples

Data recorded during stable LHC running periods on June 9 and June 10 2015, when the tracking detectors and the solenoid were in fully operational mode are used to evaluate the track reconstruction performance. The LHC delivered collisions during special fills with low beam currents and reduced focusing to give a mean number of interactions per bunch crossing below 0.005. Inelastic proton-proton collisions were collected with a minimum-bias trigger, requiring at least one hit in the MBTS. In total about 10 M events were recorded. Data for the first period is summarized in Table 3.1. For what concerns the simulated sample, Pythia 8 - A2 was used to generate events containing a single inelastic proton-proton interaction. The detector response is simulated in a detailed model implemented in GEANT4 [103].

3.4.2 Track Quality Selection

Tracks were reconstructed by using the MinBias setup and they are selected by applying specific track quality criteria. Two sets of quality criteria are studied in this note, based on transverse momentum (p_{T}), pseudorapidity (η) and the number of hits in the subsystems. As previously explained, the tracking algorithm can efficiently reconstruct tracks down to a transverse momentum of 100 MeV, but a higher p_{T} threshold of 400 MeV was chosen for these studies in order to reduce bin-by-bin migration effects due to finite p_{T} resolution and to keep the rate of fake tracks as low as possible.

Loose

- $p_{\mathrm{T}} > 400$ MeV
- $|\eta| < 2.5$
- Number of Pixel and SCT clusters on track ('silicon hits') ≥ 7
- Number of shared modules ≤ 1
- Number of silicon holes ≤ 2
- Number of pixel holes ≤ 1

Tight Primary (in addition to the Loose selection requirements)

- Number of silicon hits ≥ 9 (if $|\eta| \leq 1.65$)
- Number of silicon hits ≥ 11 (if $|\eta| \geq 1.65$)
- At least one hit on one of the two innermost pixel layers
- No pixel holes

A shared hit is a hit used by more than one track. A Pixel module is considered to be shared if it has one or more shared hits. In the case of the SCT, a shared module has at least two shared hits[2]. Holes are defined as intersections of the reconstructed track trajectory with a sensitive detector element which do not result in a hit - essentially a "missing hit" on a track. Holes are only counted between the first and last hit assigned to the track, in order not to penalise late-decaying/secondary particles or tracks that do not reach the outer silicon layers of the detector. They are estimated by following closely the track trajectory and comparing the hits-on-track with the intersected modules. Inactive modules or regions such as edge areas on the silicon sensors are excluded from the hole

[2]No requirement is made that the two hits should be on opposite sides of the double-sided modules.

definition. The tracks selected by the above working points can have extensions into the TRT, but no explicit criteria are applied on the number of TRT hits.

The Loose track selection corresponds to the default track requirements applied during inside-out track reconstruction and aims to achieve a high charged particle reconstruction efficiency at the cost of a non-negligible fraction of fake tracks. The name Tight Primary reflects the fact that this set of criteria is designed and optimized for selecting primary tracks, which are defined in Section 3.3, and rejecting fake tracks at the cost of a reduced track reconstruction efficiency.

TABLE 3.1: Summary of used data from low luminosity runs at $\sqrt{s} = 13$ TeV. The column N_{Evt} represents the total number of events passing data quality and trigger selection. $N_{\text{Evt,sel}}$ represent the number of these events passing event selection, especially the requirement of one reconstructed vertex. N_{Trk} is the total number of tracks in the selected events, while $N_{\text{Trk,loose}}$ and $N_{\text{Trk,tight}}$ are the tracks passing the loose and tight-primary selection respectively. A reconstruction setup specific for this low luminosity requirement was used for this data, reconstructing tracks down to a p_{T} of 100 MeV. Therefore, most tracks fail selection because of the requirement on p_{T} to be at least 400 MeV in the used track selections. Ref. [102].

Luminosity	N_{Evt}	$N_{\text{Evt,sel}}$	N_{Trk}	$N_{\text{Trk,loose}}$	$N_{\text{Trk,tight}}$
168 μb^{-1}	10 838 000	9 302 410	346 653 000	165 141 000	144 900 000

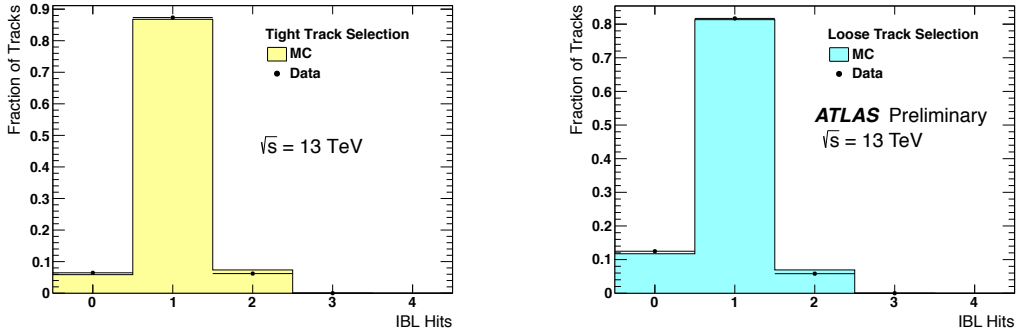
3.4.3 Basic Performance

The reconstructed track performance for the low luminosity run is presented for the loose and tight-primary track selection in events satisfying the following criteria:

- a MBTS trigger on at least one side of the ATLAS detector.
- the presence of at least one reconstructed vertex [104] with at least two associated tracks.
- the rejection of events with a second vertex with four or more tracks, in order to remove multiple interactions within a beam crossing.
- at least one track passing the loose selection.

The selected tracks in simulation are reweighted to ensure that their kinematics match those observed in data. The primary $z-$vertex distribution in simulation is also matched to the data.

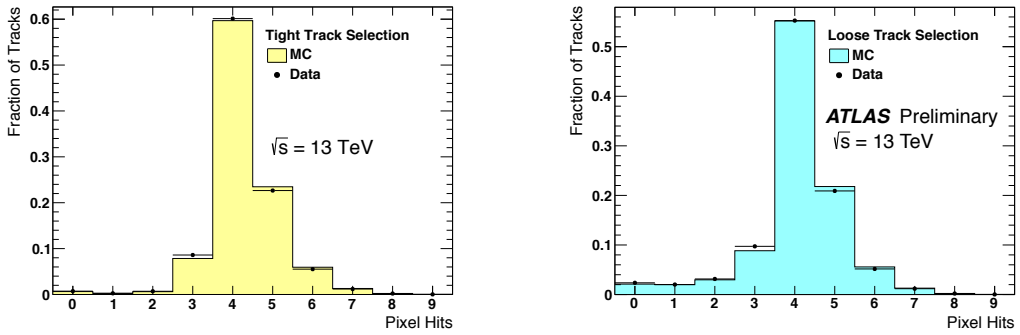
In Figures 3.9, 3.10, 3.11 and 3.12 the number of hits on track are compared for the individual detector technologies for tracks passing the loose and tight track selection. Data and MC simulation agree fairly well for all quantities. A similar comparison of the average number of hits versus η is shown in Figures 3.13, 3.14, 3.16 and 3.17. The evident discrepancies can be attributed to multiple sources. The dominant ones are wrong descriptions of dead and inefficient pixel modules in the MC simulation and residual mis-alignment of the detector elements of the ID [105]. In particular, the asymmetry observed in Figure 3.13 can be attributed to dead modules in one of the runs used. This is demonstrated in Figure 3.15 where, as in Figure 3.14(b), the average number of pixel hits versus η is shown, but the description of dead and inefficient pixel modules in simulation has been adjusted to match the one in the corresponding data set.



(a) Tight Track Selection

(b) Loose Track Selection, Ref. [102]

FIGURE 3.9: Comparison of the number of IBL hits distributions in data and simulation for the (a) tight and (b) loose track selection. The distributions are normalized to one.

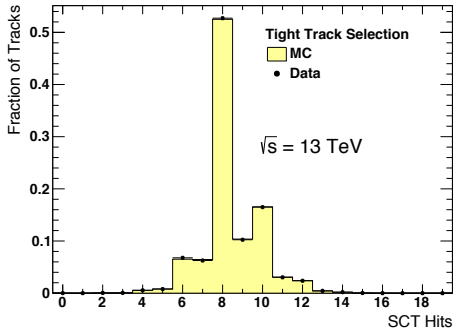


(a) Tight Track Selection

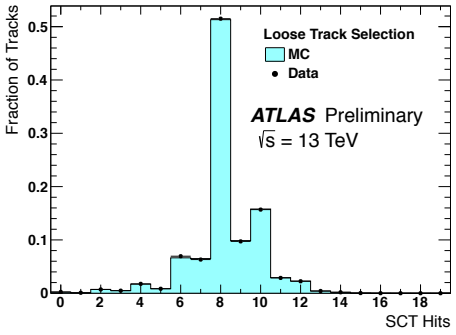
(b) Loose Track Selection, Ref. [102].

FIGURE 3.10: Comparison of the number of pixel hits distributions in data and simulation for the (a) tight and (b) loose track selection. The distributions are normalized to one.

The efficiency to extend a track to the SCT (which will be described in detail in Chapter 4), i.e. to add hits from the SCT to a track with only pixel hits, is presented in Figure 3.18(a). For this plot different track selection criteria are applied, specifically the track to be extended into the SCT is required to have a minimum of four pixel hits. For the extension to be successful it is required to have at least two SCT hits on the extended track. Similarly, the efficiency to extend a track with silicon hits into the TRT is shown in Figure 3.18(b) for both the loose and tight-primary selections. A successful extension into the TRT requires at least nine TRT hits to be added to the track. While both efficiencies are affected by the amount of material between the respective detector volumes, this effect dominates for the SCT extension efficiency, and makes the method suitable for studies of the Inner Detector passive material. The SCT extension efficiency is, in fact, a valid figure to check for differences in the detector description between data and simulation. The observed discrepancies are sizable, growing up to 2.5% at $|\eta| \sim 2.4$ and they have been extensively investigated and described in Chapter 4.

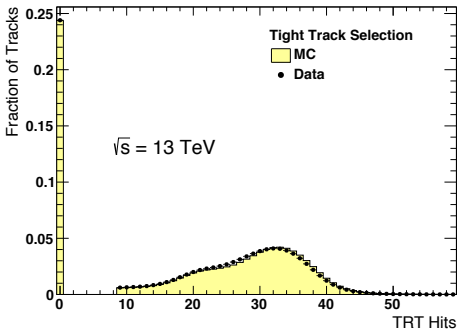


(a) Tight Track Selection

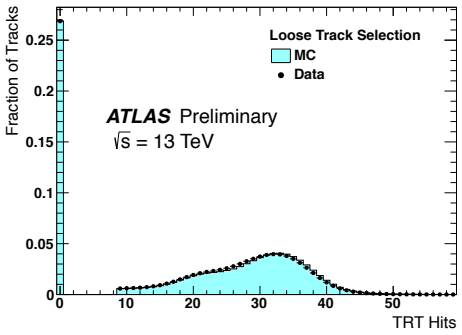


(b) Loose Track Selection, Ref. [102].

FIGURE 3.11: Comparison of the number of SCT hits distributions in data and simulation for the (a) tight and (b) loose track selection. The distributions are normalized to one.



(a) Tight Track Selection



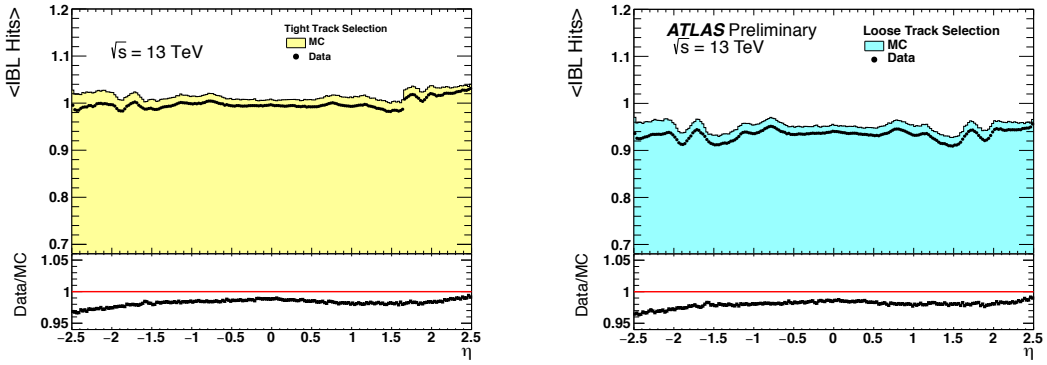
(b) Loose Track Selection, Ref. [102].

FIGURE 3.12: Comparison of the number of TRT hits distributions in data and simulation for the (a) tight and (b) loose track selection. The distributions are normalized to one.

3.4.4 Impact Parameter Resolution

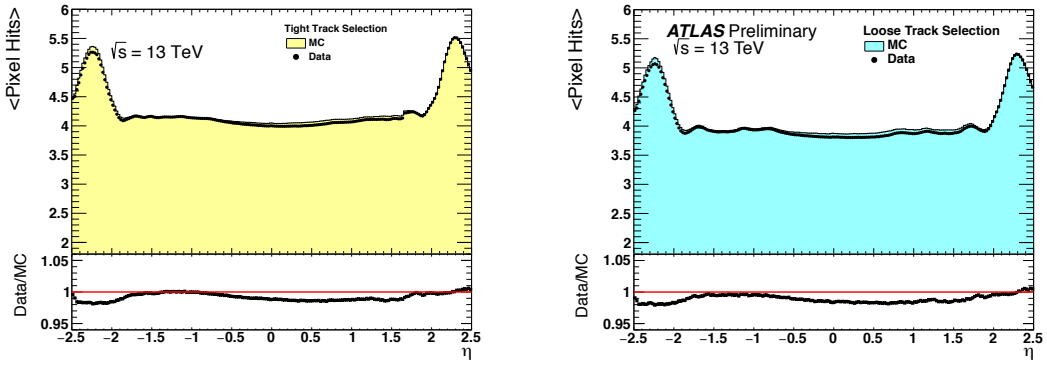
Figure 3.19 shows the transverse and longitudinal impact parameters in data and simulation with respect to the primary vertex A subset of the previously discussed data from low μ runs at $\sqrt{s} = 13$ TeV is used. The simulation is reweighted to take into account the (p_T, η) kinematics of tracks, as well as the z-vertex distribution in data. The observed differences in Figure 3.19(a) arise from discrepancies in the material description used in simulation for the IBL and this effect results to be dominant at low p_{T}. At higher p_{T}, the residual misalignment of the detector components contributes to the data-simulation discrepancy. The impact parameter resolution is driven by the resolution of the individual measurements in the pixel detector. In simulation, the resolution of the longitudinal measurements is known to be better to that in data due to a simplified energy deposit model. This effect is visible in Figure 3.19(b).

Since the impact parameter is measured with respect to the primary vertex, the resolution of the primary vertex is unavoidably folded into the impact parameter's intrinsic



(a) Tight Track Selection (b) Loose Track Selection, Ref. [102].

FIGURE 3.13: Comparison of the number of IBL hits as a function of the pseudorapidity η in data and simulation for the (a) tight and (b) loose track selection.



(a) Tight Track Selection (b) Loose Track Selection, Ref. [102].

FIGURE 3.14: Comparison of the number of Pixel hits as a function of the pseudorapidity η in data and simulation for the (a) tight and (b) loose track selection.

resolution. In addition, also the unbiased[3] primary vertex resolution depends on the p_{T} and η of the track, due to the correlations of this track with the remaining tracks present in the same event – though this is a minor effect. As the impact parameter resolution depends on the tracks kinematics, its distribution is further convoluted as a function of p_{T} and η. For a proper comparison between data and simulation the impact parameter resolution needs to be unfolded from the resolution of its reference point. This iterative deconvolution procedure is described elsewhere [106]. It was validated by comparing the unfolded resolution in simulation to the one known for the generated particles and was shown to be in general in agreement with the intrinsic resolution. In the unfolded z_0 resolution a discrepancy can be observed for tracks above $|\eta| \sim 0.5$. This residual discrepancy is added as a systematic error to the MC simulation in the comparison below.

The unfolded results can be seen in Figure 3.20. The results from data are drawn alongside the values expected for particles from simulation. In both Figures 3.20(a) and

[3]The unbiased resolution of a vertex measurement can be obtained by removing a specific track from the inputs used to determine the vertex's position and then recalculating it.

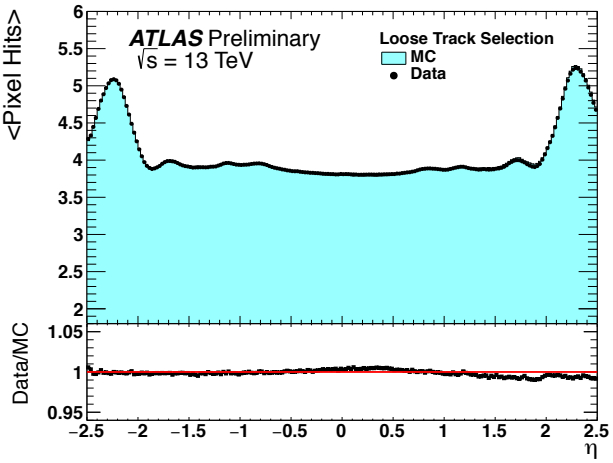
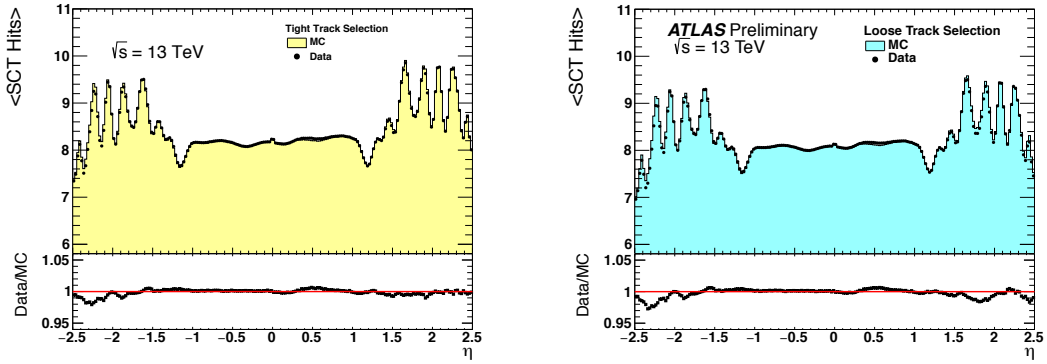


FIGURE 3.15: Comparison of the average number of pixel hits per selected track as a function of pseudorapdity of the track in data and simulation for the loose track selection. The Monte Carlo simulation uses an updated description of dead and inefficient modules derived specifically for the data set which is plotted. Ref. [102].



(a) Tight Track Selection

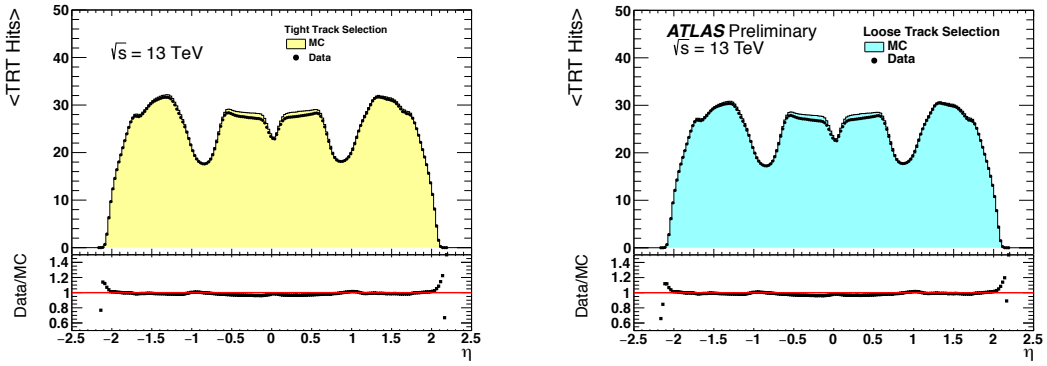
(b) Loose Track Selection, Ref. [102].

FIGURE 3.16: Comparison of the number of SCT hits as a function of the pseudorapidity η in data and simulation for the (a) tight and (b) loose track selection.

3.20(c) a fairly constant difference from low to higher p_{T} is observed in a central η region of $0.2 < \eta < 0.4$, which accounts for roughly a 16% and 5% worsened resolution of d_0 and z_0 respectively. At high η, the z_0 resolution differs vastly from the expected value, as visible in Figure 3.20(d). For the d_0, this behavior can not be observed, but it rather shows a 15−20% difference in the selected p_{T} range. The source of this differences was already discussed for Figure 3.19.

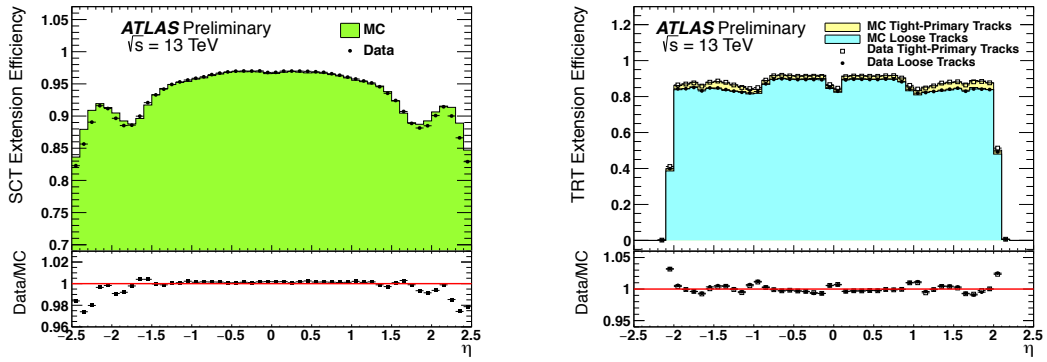
3.5 Early Inner Detector Tracking Performance in the 2015 data

The studies undertaken to provide a recommended set of Inner Detector tracking performance benchmarks (referred to as *tracking recommendation*) are summarised. These recommendations provide information on appropriate track selection criteria (referred to as



(a) Tight Track Selection (b) Loose Track Selection, Ref. [102].

FIGURE 3.17: Comparison of the number of TRT hits as a function of the pseudorapidity η in data and simulation for the (a) tight and (b) loose track selection.



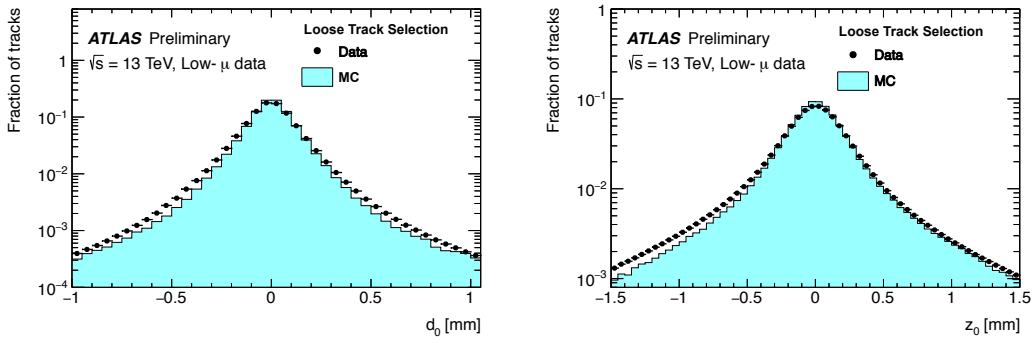
(a) SCT Extension Efficiency (b) TRT Extension Efficiency

FIGURE 3.18: Efficiency to extend a track reconstructed in the pixel detector to the (a) SCT and the efficiency to extend a track reconstructed in the silicon detectors to the (b) TRT in data and simulation. Ref. [102].

working points) with well-understood performance to be applied in ATLAS Run 2 physics analyses using ID tracks, along with the estimated systematic uncertainties from a variety of sources which should be applied. The track reconstruction efficiencies measured by the Author of this thesis are presented for the Loose and Tight track quality selections described in 3.4.2. The evaluation of the fake rate as well as the impact parameter resolution is also shown for completeness, while the analysis of the alignment weak modes is not discussed here and details can be found in [107].

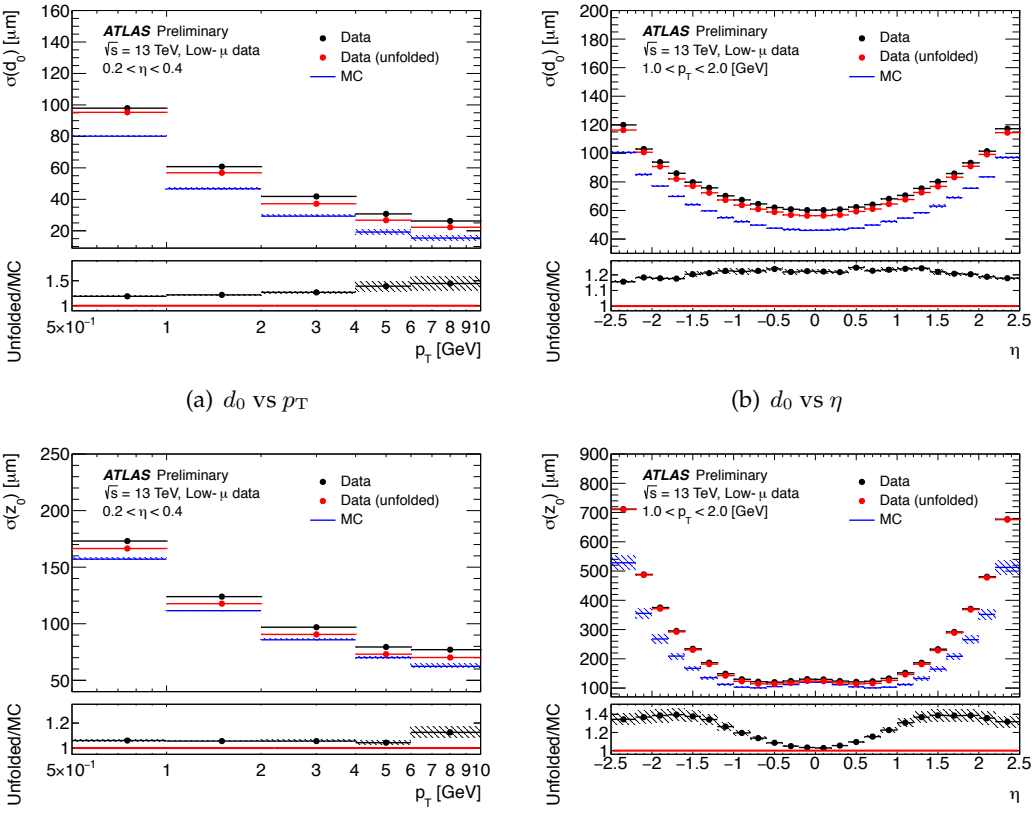
3.5.1 Data and Simulated events

The simulation-based track reconstruction efficiency shown in Section 3.5.2 is evaluated by using the same minimum bias simulated events described in Section 3.4.1 generated with Pythia 8 - A2. The ATLAS detector geometry on which such an event sample is based is the standard geometry (referred to as *original*) used as baseline for the production of Monte Carlo simulated samples in 2015. The main source of uncertainty on the tracking efficiency is the imprecise knowledge of the ID material distribution. Therefore, based on



(a) Transverse (b) Longitudinal

FIGURE 3.19: (a) Transverse, d_0, and (b) longitudinal, z_0, impact parameter distributions expressed with respect to the primary vertex. Ref. [102].



(a) d_0 vs p_{T} (b) d_0 vs η

(c) z_0 vs p_{T} (d) z_0 vs η

FIGURE 3.20: Unfolded (a,b) transverse and (c,d) longitudinal impact parameter resolution measured in data as a function of p_{T} (for values of $0.2 < \eta < 0.4$) and η (for values of $1 < p_{\mathrm{T}} < 2$ GeV), compared to the expectation from Monte Carlo simulation. The error on Monte Carlo simulation includes the statistical uncertainty and the non-closure of unfolded Monte Carlo data with respect to the true resolution known from simulation. For reference, all figures also show the resolution in data before unfolding. Ref. [102].

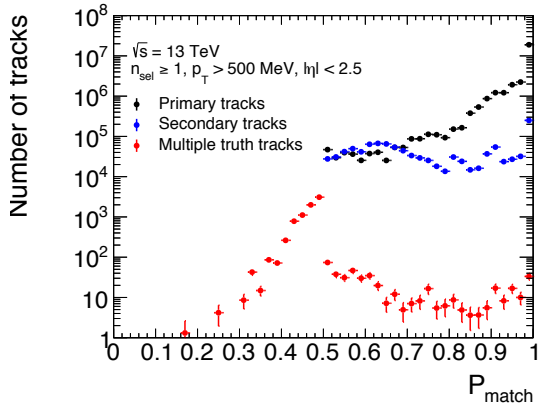


FIGURE 3.21: Distribution of hit-based track-truth matching P_{match} for primary, secondary and fake tracks (here called "multiple truth" tracks).

the ID material studies described in detail in Chapter 4, a number of distorted geometries were produced on top of the original geometry in order to study this source of uncertainty:

- $+5\%$*Extra* = 5% of extra material uniformily distributed in the ID.
- $+25\%$*PixServExtra* = 25% of extra material uniformily distributed in the Pixel service regions.
- $+50\%$*PixServExtra* = 50% of extra material uniformily distributed in the Pixel service regions.
- $+10\%$*IBLExtra* = 10% of extra material uniformily distributed in the IBL region.

For the fake rate studies shown in Section 3.5.3, data taken between September and November 2015 are used, with an integrated luminosity of around 1.74 fb^{-1}. During this period, the bunch spacing was 25 ns and the average number of interactions per bunch crossing, μ, reached approximately 23. A zero-bias trigger[4] was used to select the events. The data are compared to Monte Carlo simulation in which a varying number of minimum bias events, generated with Pythia 8 - A2 as described earlier, were overlaid on top of an 'empty' event, produced by simulating events containing single neutrinos, in order to provide simulated 'pile-up' events with a range of values of μ.

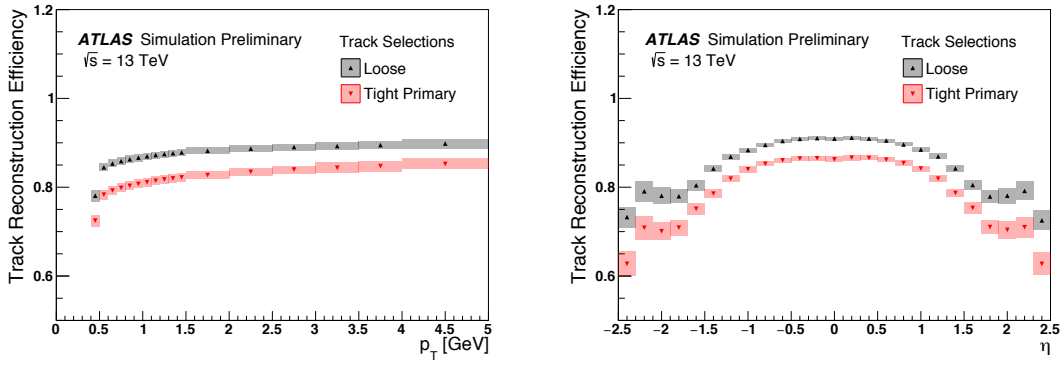
For the impact parameter resolution studies presented in Section 3.5.4, the minimum bias data set collected in June 2015 and already described in Section 3.4.1, is used. Data are compared to simulation and, in addition to the minimum bias simulated sample, a sample of Z bosons decaying into muon pairs is used. This allows to improve the studies in the high track momentum regime where minimum bias events are statistically limited. The Z $\rightarrow \mu\mu$ sample was generated by using POWHEG [108–110] with a dedicated modeling of single boson production [111], interfaced to PYTHIA 8 for showering.

In all cases, the detector response is simulated with GEANT4.

3.5.2 Track Reconstruction Efficiency and Systematics

For the studies described here, a $P_{\text{match}} > 0.5$ (instead of 0.7, as used in the studies described in Section 3.3) is required for a primary track. As it can be seen in Figure 3.21, it was checked that using this looser cut allows to gain in statistics without increasing the rate of fakes.

[4]This trigger selects events one LHC orbit after a reference trigger, to give a completely unbiased sample of collision events.



(a) (b)

FIGURE 3.22: Track reconstruction efficiency, evaluated by using minimum bias simulated events, as a function of truth (a) p_{T} and (b) η for Loose and Tight Primary track selections. The bands indicate the total systematic uncertainty. Ref. [107].

Primary charged particles are then defined as charged particles with a mean lifetime $\tau > 300$ ps, either directly produced in proton-proton interactions or from decays of directly produced particles with mean lifetime $\tau < 30$ ps. Particles produced from decays of particles with $\tau > 30$ ps are flagged as secondary particles and are excluded, as will be explained in details in Chapter 5.

The track reconstruction efficiency for primary tracks, as defined in Equation 3.2, is shown in Figure 3.22 as a function of p_{T} and η for both Loose and Tight Primary selections, as defined in Section 3.4.2. The evaluated Tight Primary track reconstruction efficiency is overall lower (up to $\sim 10\%$ for $1.5 \leq |\eta| \leq 2.5$, $\sim 5\%$ for $|\eta| < 1.5$ and $\sim 5\%$ in the full p_{T} range) than the Loose track reconstruction efficiency because of the more stringent requirements, but the Tight Primary selection improves fake rejection, as detailed in Section 3.5.3. In general, the lower track reconstruction efficiency in the region $|\eta| > 1$ is due to the increasing amounts of material that the particles must traverse. The slight increase in efficiency for $|\eta| > 2$ is due to the particles passing through a larger number of sensitive layers. For $p_{\mathrm{T}} \geq 5$ GeV, the track reconstruction efficiency reaches a plateau equal to ∼90% and ∼85% respectively for the Loose and Tight Primary selections. The accuracy with which the amount of material in the ID is known is the largest source of uncertainty on the simulation-based estimate of the track reconstruction efficiency. Three main components to the systematic uncertainty (referred to as $\mathrm{Sys}_{+5\%Extra}(\mathrm{p_T}, \eta)$, $\mathrm{Sys}_{PixServExtra}(\mathrm{p_T}, \eta)$ and $\mathrm{Sys}_{+30\%IBLExtra}(\mathrm{p_T}, \eta)$ in the following) are designed to cover disagreements between data and MC simulation in the context of the ID material studies described both in Ref. [112, 113] and in Chapter 4. These components have been evaluated for both Loose and Tight Primary selections as a function of p_{T} and η by subtracting unity from the ratio of the track reconstruction efficiencies measured in the original and distorted geometry samples listed in Section 3.5.1:

$$\mathrm{Sys}_{ExtraMaterial}(\mathrm{p_T}, \eta) = \frac{\epsilon_{\mathrm{trk}}^{original}(\mathrm{p_T}, \eta)}{\epsilon_{\mathrm{trk}}^{ExtraMaterial}(\mathrm{p_T}, \eta)} - 1 \tag{3.4}$$

where *ExtraMaterial* means +5%*Extra*, +25%*PixServExtra*, +50%*PixServExtra* or +10%*IBLExtra*.

The contribution $\mathrm{Sys}_{+5\%Extra}(\mathrm{p_T}, \eta)$ is directly measured from the +5%*Extra* sample, to cover the systematic uncertainty due to an overall 5% material mis-modeling which

Track Reconstruction Efficiencies and Systematic Uncertainties				
Track Quality Selection	Loose		Tight Primary	
η Range	$\lvert\eta\rvert \leq 0.1$	$2.3 \leq \lvert\eta\rvert \leq 2.5$	$\lvert\eta\rvert \leq 0.1$	$2.3 \leq \lvert\eta\rvert \leq 2.5$
Track Reconstruction Efficiency	91%	73%	86%	63%
$\mathrm{Sys}_{+5\%Extra}$	0.4%	0.9%	0.5%	1.1%
$\mathrm{Sys}_{PixServExtra}$	—	2.0%	—	2.3%
$\mathrm{Sys}_{+30\%IBLExtra}$	0.2%	0.5%	0.2%	0.5%
Total Systematic Uncertainty	0.4%	2.2%	0.5%	2.6%

TABLE 3.2: Track reconstruction efficiencies and absolute systematic uncertainties for both Loose and Tight Primary track quality selections. Two η intervals were selected to demonstrate the range of the systematic uncertainties: $|\eta| \leq 0.1$ (which is characterised by the lowest systematic uncertainty values) and $2.3 \leq |\eta| \leq 2.5$ (which is characterised by the largest systematic uncertainty values). The total systematic uncertainty is estimated by quadratically combining the three different contrbutions: $\mathrm{Sys}_{+5\%Extra}(\mathrm{p_T}, \eta)$, $\mathrm{Sys}_{PixServExtra}(\mathrm{p_T}, \eta)$ and $\mathrm{Sys}_{+30\%IBLExtra}(\mathrm{p_T}, \eta)$. The results shown are obtained by integrating over p_{T} in the range [0.4,20] GeV. Ref. [107].

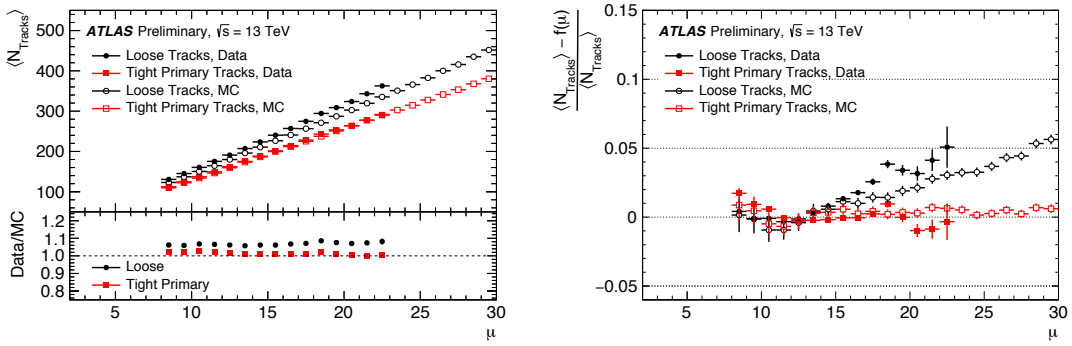
is the assumed upper limit on the precision with which the Run 1 material budget is known. Based on the hadronic interactions and photon conversions studies (as described in Ref. [112] and in Chapter 4), the contribution $\mathrm{Sys}_{+10\%IBLExtra}(\mathrm{p_T}, \eta)$ corresponding to the results from the +10%*IBLExtra* sample, was increased by a factor of 3 to fully cover the disagreement between data and simulation, and it is denominated as $\mathrm{Sys}_{+30\%IBLExtra}(\mathrm{p_T}, \eta)$. An uncertainty denoted $\mathrm{Sys}_{PixServExtra}(\mathrm{p_T}, \eta)$ is applied in the region with $|\eta| > 1.5$ to cover the data/simulation discrepancy in the forward region of the detector. For $1.5 < |\eta| < 2.3$, the full difference between the original and $+50\%$*PixServExtra* samples is conservatively used to evaluate this contribution to the systematic uncertainty, while, in the region $|\eta| > 2.3$, since the missing pixel services material is estimated to be in the order of 30% [112], it is necessary to carry out a linear interpolation between the $+25\%$*PixServExtra* and $+50\%$*PixServExtra* samples to evaluate the uncertainty. The error bars in Figure 3.22 represent the total systematic uncertainty obtained by combining in quadrature the three contributions discussed above. An overview of the track reconstruction efficiency and the absolute values of the systematic uncertainties is shown in Table 3.2.

In the simulation samples used to evaluate the track reconstrunction efficiency, pions represent the dominant particle type with a fraction of more than 98%. Therefore, since the track reconstruction efficiency is sensitive to particle type, the above-described track systematic uncertainties are intended to be applied to charged pions.

3.5.3 Fake Rate

Not all tracks reconstructed in the ID will actually correspond to a charged particle traversing the detector; in particular, combinatorial effects which increase with pile-up can give rise to fake tracks. As a cross-check on the modelling of the tracking fake rate, an estimate of this quantity is made by assuming that the number of non-fake tracks is to first order proportional to the number of pile-up interactions. Any deviation from linearity is therefore assumed to be due to fakes. Figure 3.23(a) shows the average number of tracks passing the Loose and Tight Primary selections, as a function of μ.

The data in Figure 3.23(a) are fitted with a linear function $f(\mu) = m\mu$ in the region $10 \leq \mu \leq 15$ in which both data and Monte Carlo demonstrate approximately linear



(a) (b)

FIGURE 3.23: Average number of reconstructed tracks as a function of μ for data and minimum bias simulation, as obtained by applying the Loose and Tight Primary selections (a), and an estimation of the tracking fake rate, derived from the deviation from linearity of a fit to N_{tracks} as a function of μ (b). In Figure (a), the statistical uncertainty on the mean number of tracks is considered, but the error bars are too small to be seen. The displayed μ range for data, in contrast to simulation, is limited by the available μ values in the data sample. Ref. [107].

behaviour. The relative deviation from the linear fit:

$$\frac{\langle N_{\text{tracks}}\rangle(\mu) - f(\mu)}{\langle N_{\text{tracks}}\rangle(\mu)}$$

provides an estimate of the fake rate, which is shown in Figure 3.23(b). It is a relative measure, and by construction close to zero in the region of the fit, $10 \leq \mu \leq 15$. By approximating the fake rate in this manner, the contribution from fake tracks is assumed to be negligible in the fit region; this is a more reliable assumption for the Tight Primary track selection than for the Loose track selection. This method also assumes that the tracking efficiency is independent of μ, and that the fake contribution from sources other than pile-up combinatorics are negligible; these latter two assumptions have been verified to hold in Monte Carlo. The Tight Primary track selection is estimated to have a smaller fake rate, which is also more stable versus μ, than the Loose selection. The recommended systematic uncertainty to be applied on the fake rate is 50%, which is typical of the size of the differences between Data and MC seen in these studies.

3.5.4 Impact Parameter resolution

The intrinsic resolutions of the track impact parameters with respect to the primary vertex are measured from data and simulation. As already mentioned in Section 3.5.1, to extrapolate the systematic uncertainty on the impact parameters to higher track p_{T}, the results from minimum bias simulation are compared with those using muon tracks from Z boson decays.

For minimum bias data and simulation, events are required to have a reconstructed primary vertex with at least 10 tracks, while events with one or more additional reconstructed primary vertices with more than 4 tracks are removed to reduce the effect of pile-up. For $Z \to \mu\mu$ simulation, the track selection, as described in Ref. [114], requires at least four silicon hits, at least 25 hits in the TRT and the p_{T} of the track to be higher than 15 GeV. To extract the resolution figures, the impact parameter distribution is fitted within 2σ of

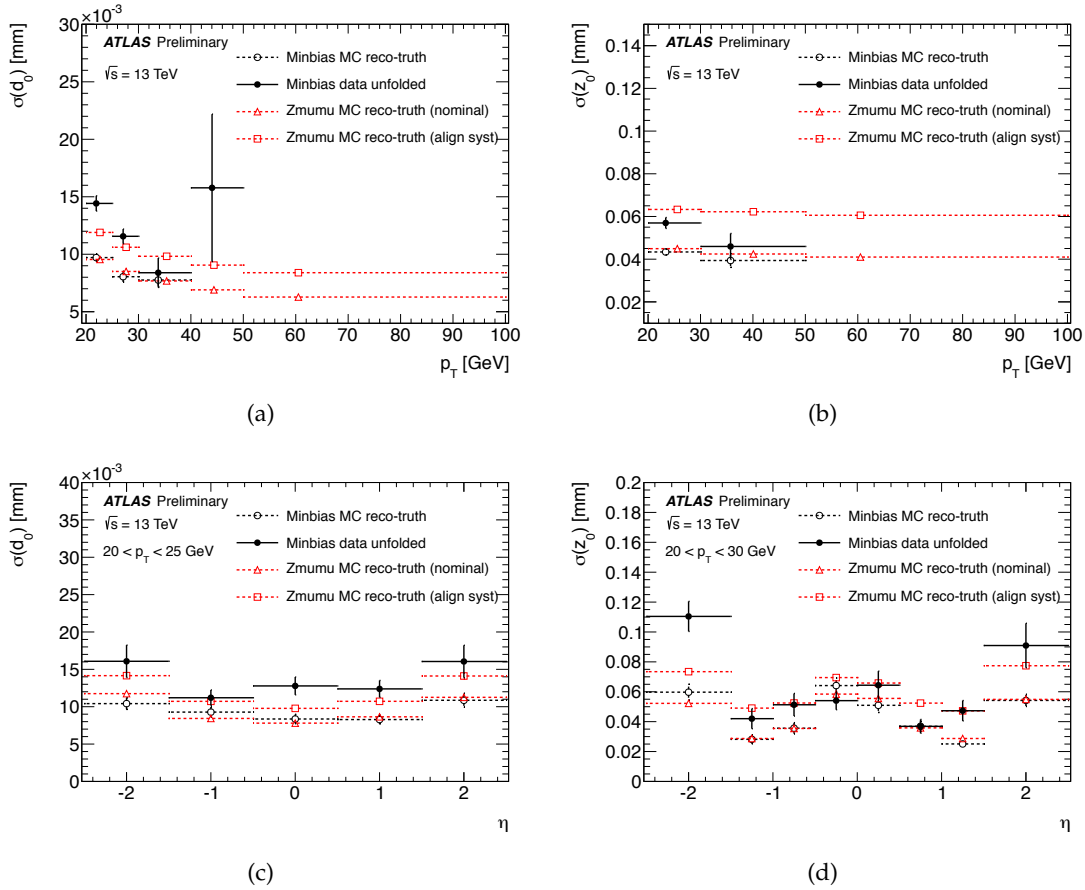
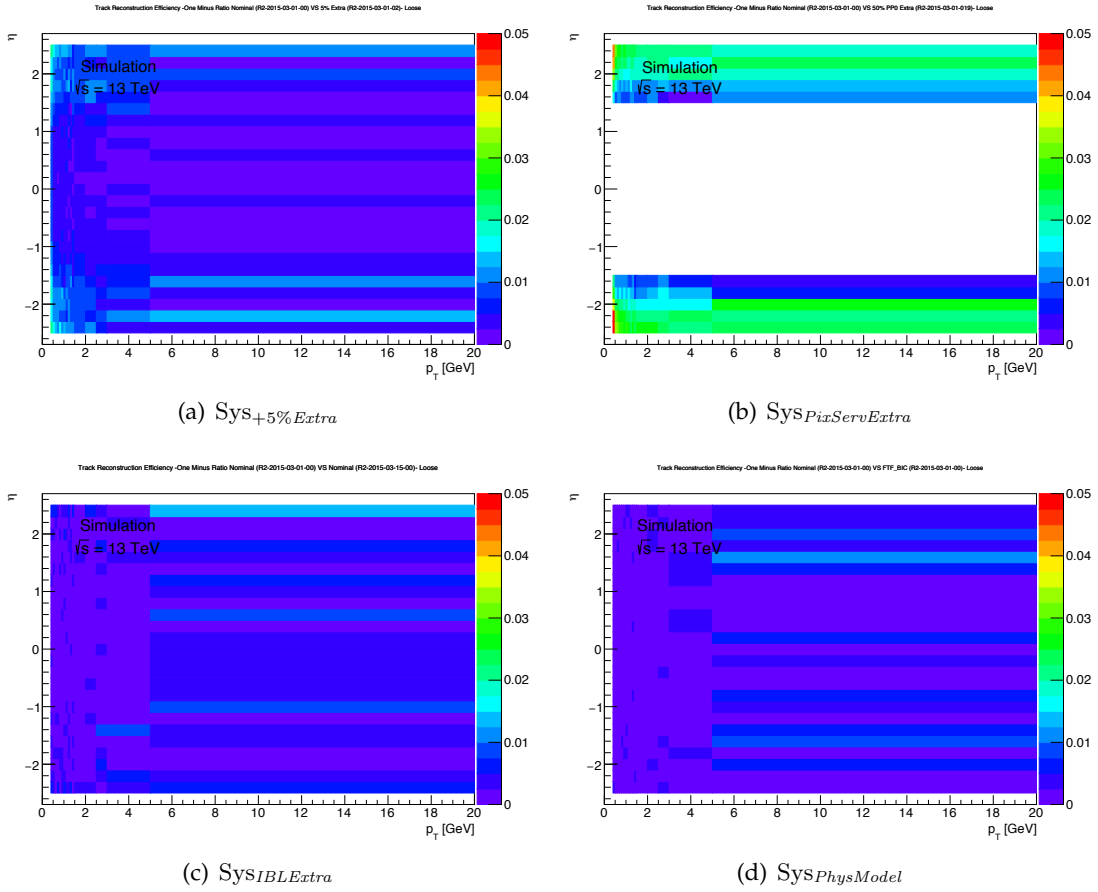


FIGURE 3.24: Unfolded transverse (a)(c) and longitudinal (b)(d) impact parameter resolution measured in data as a function of p_{T} and η, compared to the expectation from minimum bias and $Z \to \mu\mu$ simulation. For simulation, the resolution is taken from the differences between the reconstructed and truth quantities (reco-truth) while for data the resolution is derived by using an iterative deconvolution procedure. In the $Z \to \mu\mu$ simulation, *nominal* shows the resolutions with perfect alignment, while *align syst* shows the effect of June-2015 alignment configuration, with uncertainties obtained by applying Gaussian-distributed misalignments to the various detector components. The d_0 and z_0 resolutions in both simulations agree with each other to within 20%. Ref. [107].

its mean with a Gaussian function for each bin in η and p_{T} of tracks. In data, the width of the Gaussian function includes the contribution from the uncertainty on the position of the reconstructed primary vertices. The intrinsic resolutions of tracks are obtained by deconvoluting the primary vertex uncertainty using an iterative deconvolution procedure described in Ref. [106]. In simulation, the values of the resolutions are calculated from the difference between the reconstructed impact parameters and those from the Monte Carlo truth particles. The values of σ representing the intrinsic resolution are shown in Figure 3.24. The resolutions from simulation with the nominal and June-2015 alignment [105] in $Z \to \mu\mu$ configurations are also shown. The nominal alignment configuration represents a perfectly aligned detector, while the June-2015 alignment configuration reproduces the uncertainties on the alignment by applying error scalings (random, Gaussian-distributed misalignments with σ corresponding to the relevant alignment uncertainties) to various detector components.



(a) $\mathrm{Sys}_{+5\%Extra}$

(b) $\mathrm{Sys}_{PixServExtra}$

(c) $\mathrm{Sys}_{IBLExtra}$

(d) $\mathrm{Sys}_{PhysModel}$

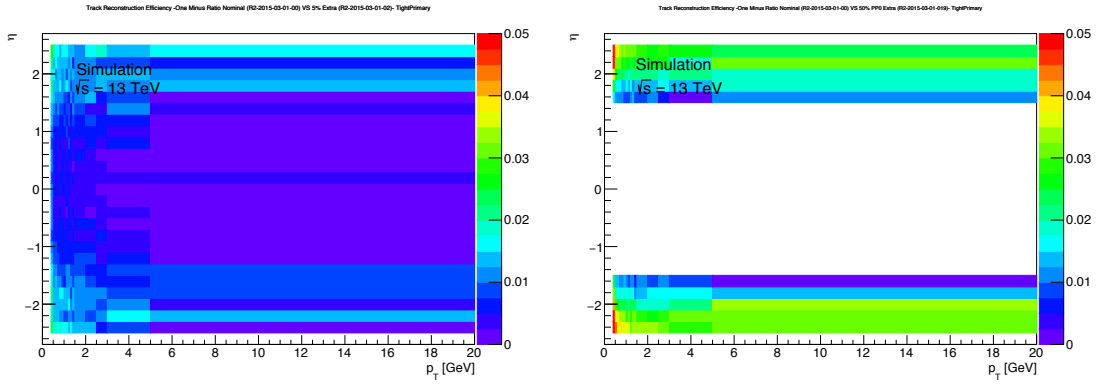
FIGURE 3.25: Individual components of the systematic uncertainty on track reconstruction efficiency, evaluated by using minimum bias simulated events reprocessed with release 20.7, as a function of truth p_{T} and η for the Loose track selection.

The d_0 and z_0 resolutions in both simulation samples agree with each other to within 20%. The difference of the resolutions between minimum bias data and simulation is taken to be the systematic uncertainty on the impact parameter resolutions. The systematic uncertainties are obtained by taking the square root of the quadrature difference between data and simulation in regions with sufficient statistics. A parameterization of the resolutions as a function of η and p_T is used to extrapolate to higher p_{T} regions which are statistically limited. The results derived from the $Z \to \mu\mu$ sample offer a cross-check of the validity of the extrapolations to these regions.

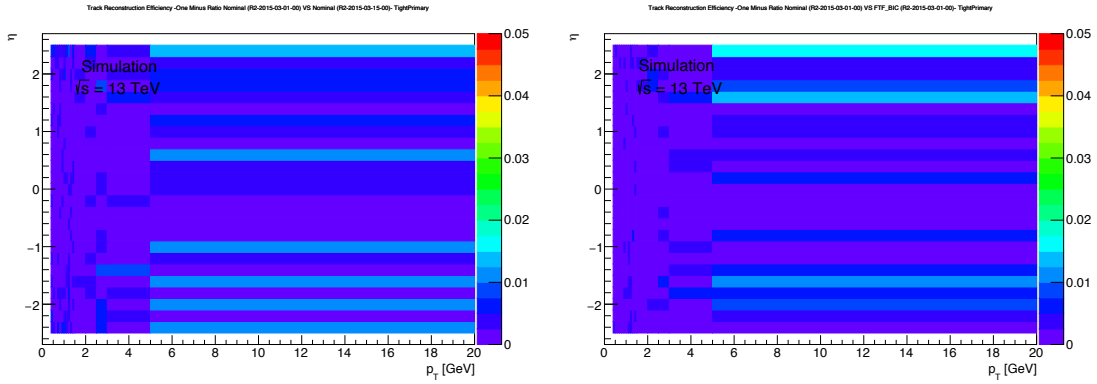
3.5.5 Updated Tracking Efficiency Recommendations in release 20.7

At the beginning of 2016, an updated ID geometry description, referred to as *updated*, (which will be detailed in Chapter 4) was implemented in a dedicated release of the ATLAS software (release 20.7) used to reprocess data and simulation samples. Thus, the tracking recommendations described above (based on a previous release of the ATLAS software, release 20.1) needed to be updated accordingly.

To provide new appropriate benchmark values, the systematic uncertainties have been evaluated by using Eq. 3.4. The ID material in the simulated samples used to measure $\mathrm{Sys}_{+5\%Extra}$ and $\mathrm{Sys}_{PixServExtra}$ is scaled as described in Section 3.5.1, while to estimate the systematic uncertainty due to mis-modeling of the IBL $\mathrm{Sys}_{IBLExtra}$, a comparison between

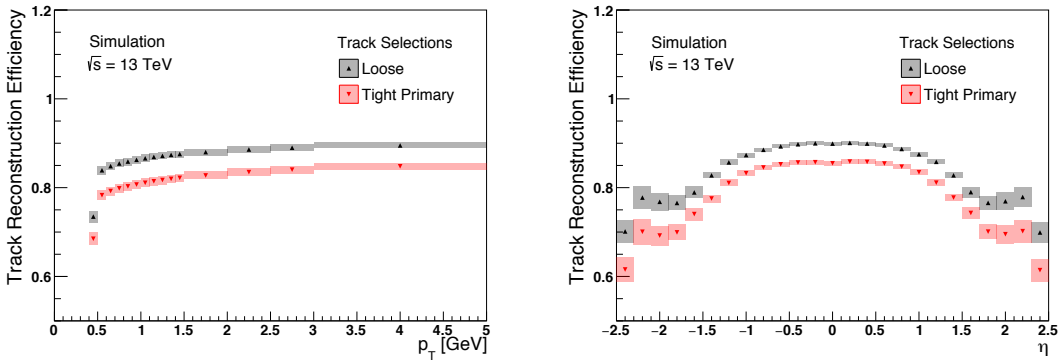


(a) $\mathrm{Sys}_{+5\%Extra}$ (b) $\mathrm{Sys}_{PixServExtra}$



(c) $\mathrm{Sys}_{IBLExtra}$ (d) $\mathrm{Sys}_{PhysModel}$

FIGURE 3.26: Individual components of the systematic uncertainty on track reconstruction efficiency, evaluated by using minimum bias simulated events reprocessed with release 20.7, as a function of truth p_{T} and η for the Tight Primary track selection.



(a) (b)

FIGURE 3.27: Updated track reconstruction efficiency, evaluated by using minimum bias simulated events reprocessed with release 20.7, as a function of truth (a) p_{T} and (b) η for Loose and Tight Primary track selections. The bands indicate the total systematic uncertainty.

the original and the updated geometry was used. The difference in the track reconstruction efficiency between the results from the original and the updated geometry, was increased by a factor of 1.3 to fully cover the disagreement between data and simulation still measured with the hadronic interaction and photon conversion rates (see Chapter 4) after improving the geometry description. This corresponds to roughly scaling up the passive material of the IBL by 40%.

In addition, a new source of systematic uncertainty, $\mathrm{Sys}_{PhysModel}$, has been studied to encompass the differences in the tracking efficiency due to the usage of different GEANT 4 physics models, known as *physics lists*, to describe the interactions of particles with material. The accuracy of modelling is particularly important to describe hadronic interactions. Currently, 19 physics lists are available in the ATLAS detector simulation suite and six of them (`FTFP_BERT`, `FTFP_BERT_HP`, `QGSP_BERT`, `QGSP_BERT_HP`, `QGSP_BIC`, `QGSP_FTFP_BERT`) are used as reference physics lists, which are routinely validated and updated.

The default physics list used in this analysis is `FTFP_BERT`. This is a recommended model by the GEANT4 collaboration for collider experiments. For hadronic interactions this model uses the Fritiof model ("`FTF`") [115, 116] for particle kinetic energies larger than 4 GeV, and a Bertini-like cascade ("`BERT`") for hadrons with kinetic energy below 5 GeV [117]. In the overlap energy region, the model is randomly selected. In addition, the pre-compound model is used for nuclear de-excitation ("`P`"). The majority of the hadronic interactions takes place at low energy and therefore the Bertini model is relatively important. Bertini employs classical scattering without matrix elements. Cross sections and angular distributions are free and derived from experiments. The nuclear density distributions and potentials are step-like. The pre-compound model helps to extend the range of Bertini model down to zero initial energy.

In this study, `FTFP_BERT` is compared to an alternative physics list `FTF_BIC`. Between `FTFP_BERT` and `FTF_BIC`, the inelastic scattering cross section is identical. For the final-state model of pion-, proton- and neutron-nucleus inelastic interactions below 5 GeV `FTF_BIC` uses the binary cascade model ("`BIC`") [118]. The pre-compound model ("`P`") is not included. `BIC` is a hybrid between a classical cascade and a full quantum-molecular dynamics model, and the implementation is more theoretically motivated compared to the `BERT` model.

The observed differences in the track reconstruction efficiency are very small ($\mathcal{O}(0.1\%)$).

In Figures 3.25 and 3.26 the individual components of the systematic uncertainty are shown, respectively for Loose and Tight Primary selections, in bins of p_{T} and η. By combining these contributions in quadrature, the final values shown in Figure 3.27 can be derived. The central value of the tracking efficiency integrated over p_{T} is slightly lower with respect to the previous recommendations and this is due to the different reconstruction setups: the Default reconstruction, with a minimum p_{T} of 400 MeV was used in place of the MinBias setup with a minimum p_{T} of 100 MeV which was used so far. This made the reconstruction efficiency to be lower in the very first p_{T} bin. In addition, the systematic uncertainty due to the imprecise description of the material in the IBL was reduced by about 50% thanks to the improved IBL description. All the updated values for the track reconstruction efficiency and for the systematics are summarised in Table 3.3.

Updated (release 20.7) Track Reconstruction Efficiencies and Systematic Uncertainties				
Track Quality Selection	Loose		Tight Primary	
η Range	$\lvert\eta\rvert \leq 0.1$	$2.3 \leq \lvert\eta\rvert \leq 2.5$	$\lvert\eta\rvert \leq 0.1$	$2.3 \leq \lvert\eta\rvert \leq 2.5$
Track Reconstruction Efficiency	90%	70%	86%	62%
$\mathrm{Sys}_{+5\%Extra}$	0.4%	1.0%	0.5%	1.1%
$\mathrm{Sys}_{PixServExtra}$	—	2.2%	—	2.5%
$\mathrm{Sys}_{IBLExtra}$	0.1%	0.3%	0.1%	0.3%
$\mathrm{Sys}_{PhysModel}$	0.1%	0.1%	0.1%	0.1%
Total Systematic Uncertainty	0.4%	2.4%	0.5%	2.7%

TABLE 3.3: Updated (release 20.7) track reconstruction efficiencies and absolute systematic uncertainties for both Loose and Tight Primary track quality selections. Two η intervals were selected to demonstrate the range of the systematic uncertainties: $|\eta| \leq 0.1$ (which is characterised by the lowest systematic uncertainty values) and $2.3 \leq |\eta| \leq 2.5$ (which is characterised by the largest systematic uncertainty values). The total systematic uncertainty is estimated by quadratically combining the four different contributions: $\mathrm{Sys}_{+5\%Extra}(\mathrm{p_T}, \eta)$, $\mathrm{Sys}_{PixServExtra}(\mathrm{p_T}, \eta)$, $\mathrm{Sys}_{IBLExtra}(\mathrm{p_T}, \eta)$ and $\mathrm{Sys}_{PhysModel}$. The results shown are obtained by integrating over p_{T} in the range [0.4,20] GeV.

Chapter 4

Material studies with the ATLAS Run 2 Inner Detector

Seeing in the air things that the others did not see...

-Alessandro Baricco, Silk-

In this chapter the focus will be on the methodologies developed to study the passive material distribution in the ATLAS Run 2 Inner detector. The emphasis will be on a method called track-extension efficiency, which was developed by the Author of this thesis to probe the material in the pixel service region. Two other methods, not directly implemented by the author of this thesis, were used to study the material in the barrel pixel region. They are based on the reconstruction of secondary vertices identifying the location of hadronic interactions or photon conversions. A brief description will also be given for these two studies, because the three methods are complementary and strongly related to each other, all aiming to a better understanding of the ID material and to finally improve the simulation of the detector components.

4.1 Introduction

Tracking detectors reconstruct the trajectories of charged particles and determine their momenta by measuring their curvature in magnetic field. Momentum resolution and tracking efficiency are strongly affected by the amount of material particles traverse both because they can be deviated from their ideal trajectory by e.g. multiple scattering and because they can disappear in a destructive interaction e.g. inelastic scattering. An accurate description of the material is essential to understand the performance of the detector. The material in the ATLAS ID was studied with several methods, using $\sqrt{s} = 13$ TeV pp collision data. Nuclear interactions by primary particles produced in pp collisions with material are the most important source of secondary particles; hence the uncertainty on the track reconstruction efficiency is directly coupled to the accuracy by which the material distribution is known. For electromagnetic calorimeters, the knowledge of the material situated between the collision point and the calorimeter front-face is essential to calibrate the energy of reconstructed electrons, un-converted and converted photons [119]. Furthermore, searches for new physics which imply reconstructing the decay vertex of long-lived particles require a precise description of the material to define decay volumes with minimal backgrounds [120]. The accuracy of the description of the material distribution is thus an essential requirement for physics analyses with the ATLAS detector, at the same level as other key ingredients required for particle reconstruction, e.g. the magnetic field description, the understanding of the processes occurring inside semiconductor sensors or gases, and the alignment of the components.

The characteristics of a material in terms of interaction with high-energy particles are expressed by properties known as the radiation length, X_0 and nuclear interaction length,

λ_I, which are quantified in units of mm in this thesis[1]. The X_0 is the mean path length over which a high-energy ($E \gg 2m_e$) electron loses all but $1/\mathrm{e}$ of its energy by bremsstrahlung. Similarly, λ_I is the mean path length to reduce the flux of relativistic primary hadrons to $1/\mathrm{e}$. In terms of electromagnetic interactions, the material thickness along a particular trajectory C is expressed by a dimensionless number N_{X_0}, frequently referred to as the *number of radiation lengths* in the literature. This is calculated as a path integral:

$$N_{X_0}^{[C]} = \int_C \mathrm{d}s \, \frac{1}{X_0(s)} \,, \tag{4.1}$$

where $X_0(s)$ is the local radiation length of the material at the position s along the trajectory C respectively[2]. Similarly, for what concerns hadronic interactions, the material thickness is expressed by a dimensionless number (*number of interaction lengths*) denoted by N_{λ_I}:

$$N_{\lambda_I}^{[C]} = \int_C \mathrm{d}s \, \frac{1}{\lambda_I(s)} \,. \tag{4.2}$$

4.2 Overview of the Methodologies

Several *in-situ* methods using collision data have been developed to estimate the material amount in the tracking detectors of collider experiments [122–124].

Probing material in the innermost barrel regions of the ATLAS Inner Detector by reconstructing photon conversion vertices is a common method to measure the material thickness of tracking detectors by taking advantage of the precise theoretical understanding of electromagnetic interaction processes [125].

The reconstruction of hadronic interaction vertices instead of photon conversions is a complementary approach. Studies of the material by using hadronic interactions were already performed during the LHC Run 1 [126, 127]. Even if the interaction dynamics is complex and is only phenomenologically modelled in the simulation, this method allows for a radial position resolution of the vertex which is much better than that achieved with photon conversions.

Another complementary approach is to measure the stopping rate of charged hadrons through hadronic interactions, referred to as the *track-extension efficiency* method.

The precision of each measurement varies depending on the detector region. All of these approaches are used together to cover the full inner detector volume and cross-check individual measurements. As shown in Figure 4.1, the hadronic interaction and photon conversion methodologies can probe the material in the barrel Pixel region and up to the first layer of the SCT detector[3], while by using the track-extension efficiency technique, the passive material in the Pixel services region between the Pixel and the SCT detectors can be investigated.

The measurements presented here are performed by comparing observables which are sensitive to the material distribution between data and Monte Carlo simulation.

[1] The common convention is of using $\mathrm{g/cm^2}$ in literature [121].

[2] The bracket "$[C]$" represents that the value is defined with respect to the trajectory C, but this can be omitted if the specified trajectory is clear.

[3] if the re-tracking is enabled, the material can be probed by hadronic interactions further out in the SCT detector, as explained in 4.4.1

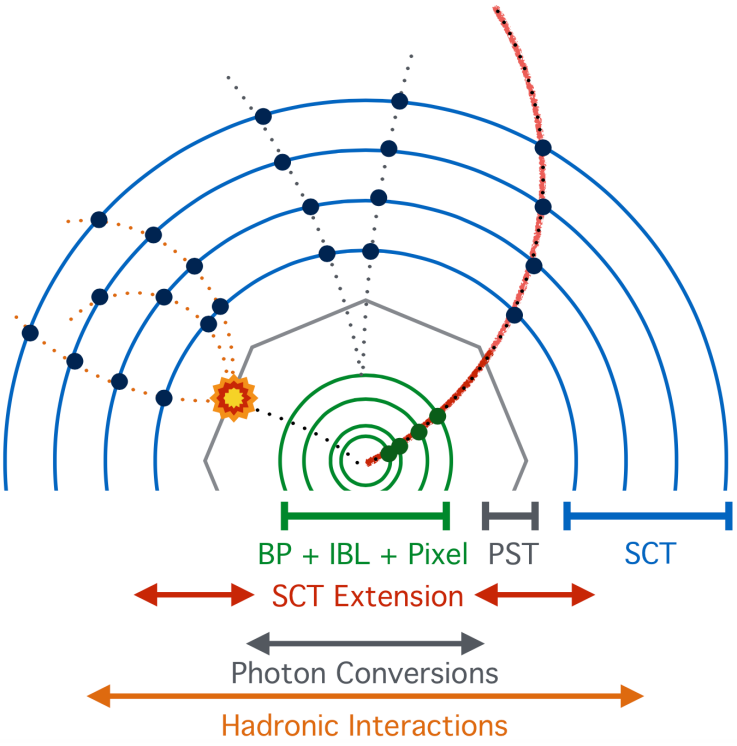


FIGURE 4.1: Sketch of the three methods used to study the material in the ATLAS Run 2 Inner Detector together with their acceptance limits.

4.3 Data and Simulation Samples

4.3.1 Data samples

The pp collision data sample used to perform the measurements described in this Chapter was collected in June 2015 at a centre-of-mass energy of $\sqrt{s} = 13$ TeV as it was already described in Chapter 3.

4.3.2 Simulation samples

The Pythia8-A2 and the EPOS-LHC Monte Carlo event generators were used to simulate inelastic minimum bias pp collisions, as described in Section 1.3.2. As it will be shown in Chapter 5 and as presented in Ref. [80], both PYTHIA 8 and EPOS are found to provide reasonable descriptions of the charged particle multiplicity distributions measured in pp collisions at $\sqrt{s} = 13$ TeV. The events are processed through the standard ATLAS detector simulation [128] programs based on GEANT4 to reproduce detector effects and are reconstructed by the same software as used to process the real data. The `FTFP_BERT` physics list is used to parameterise hadronic interactions (see Section 3.5.5 for details about physics lists). The ATLAS detector model is fed into GEANT4 by means of a collection of geometry models, each describing a specific sub-detector.

Geometry models

The material description including the atomic composition and geometrical layout, hereafter referred to as the *geometry model*, is created on the bases on engineering design drawings of the detector, together with supporting measurements of masses, dimensions and

compositions of detector components. During the construction of the Inner Detector, detailed measurements of the mass of detector components were undertaken, and the corresponding volumes in the geometry model were adapted to agree with the measurement as accurately as possible [52, 129]. However, the implementation of an accurate geometry model is challenging, because of the complexity of the detector design and the necessity of a thorough validation of the description (which implies a certain level of simplification, again).

As described in Chapter 2, the ATLAS ID was upgraded during LS1 with the installation of the IBL coupled together with a new thinner beam pipe and with the replacing the Pixel Service panels (cables, cooling pipes and support structures) located in the forward region of the Pixel detector. All these changes motivated the re-evaluation and the creation of a new ID geometry model.

The geometry model for the Inner detector describes both the active elements of the detectors (e.g. the silicon pixel sensors) and the passive material (e.g. support structures, connectors, cooling pipes and cables). The measurements presented in this thesis utilise several alternative ID geometry models, summarised below:

Original - This ID geometry model represents the nominal geometry model used to generate ATLAS 2015 simulation samples as well as in the samples used for the studies reported in Section 3.4 and 3.5. The studies presented in this thesis identified a number of missing components in the simulated description of the IBL.

Updated - A modified version of the *original* geometry model in which additional components were added to the simulated description of the IBL. These additional components include flex buses and a number of surface-mounted devices on the front-end of the modules. Small modifications to the positioning of each IBL stave and the material densities of the IBL support structures have also been introduced.

Figure 4.2 shows the radial and z material distributions in radiation length for both the *original* and *updated* geometry models. For the *updated* geometry, Figure 4.3 shows the material distribution in radiation length in the r–z view, and Figure 4.4 shows N_{λ_I} as a function of η.

Based on the *original* and *updated* geometry models, collections of distorted geometry models were also created, in which the density of a variety of components is artificially scaled by a known amount. Furthermore, three specific distorted models were personally created by the Author of this thesis to study the effect that the introduction of additional material may have on the track-extension efficiency. In these models, an additional ring of Carbon was added to the *original* geometry model and positioned at different r–z coordinates and orientations ("ring layout 1"[4] , "ring layout 2"[5] and "ring layout 3"[6]), but covering the same region of $2.2 < |\eta| < 2.3$.

A pictorial representation for ring layout 1 and 2 is shown in Figure 4.5.

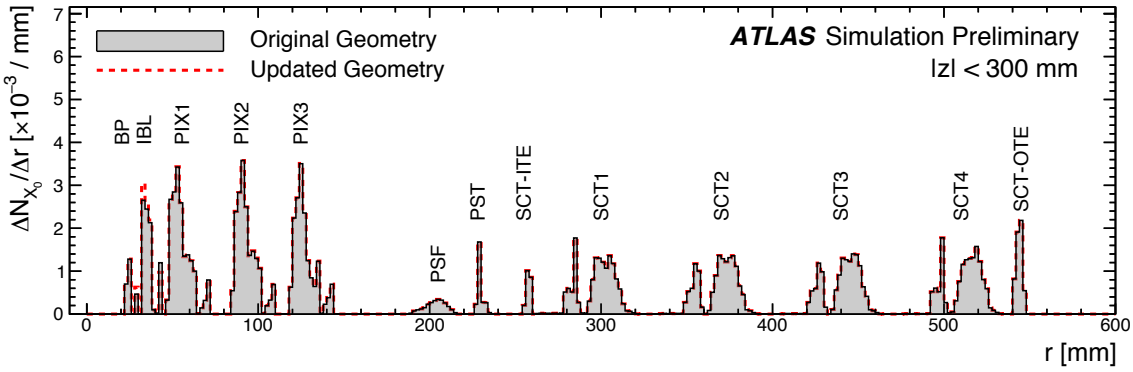
The distorted models are used to calibrate the material measurement methods and assess the systematic uncertainties associated with the measurements.

Table 4.1 summarises the collection of MC samples used in this thesis. PYTHIA 8 is used as the nominal event generator for all of the studies except for the hadronic interaction study, which uses EPOS as the nominal event generator since it is found to provide a better description of background processes than PYTHIA 8.

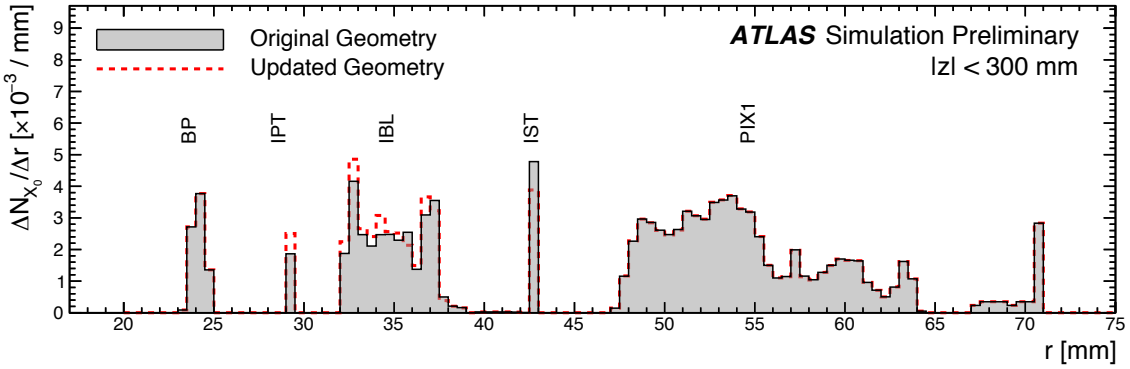
[4] In the ring layout 1, the ranges of the r–z coordinates of the two rings of passive material are: r_{ring1} = [235 mm; 238.4 mm]; z_{ring1} = [1047 mm; 1160 mm]; r_{ring2} = [141 mm; 157 mm]; z_{ring2} = [-700 mm; -683.6 mm].

[5] In the ring layout 2, the ranges of the r–z coordinates of the two rings of passive material are: r_{ring1} = [146 mm; 149.4 mm]; z_{ring1} = [650.7 mm; 716 mm]; r_{ring2} = [141 mm; 157 mm]; z_{ring2} = [-700 mm; -683.6 mm].

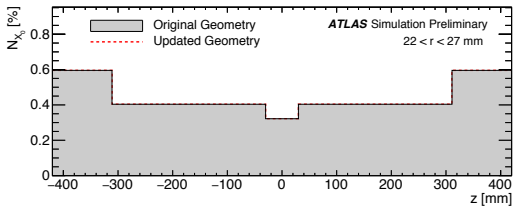
[6] In the ring layout 2, the ranges of the r–z coordinates of the two rings of passive material are: r_{ring1} = [235 mm; 241 mm]; z_{ring1} = [900 mm; 1300 mm]; r_{ring2} = [141 mm; 157 mm]; z_{ring2} = [-700 mm; -683.6 mm].



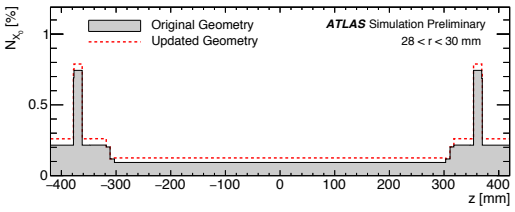
(a) r < 600 mm



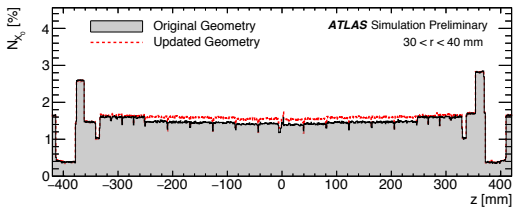
(b) 20 mm < r < 75 mm



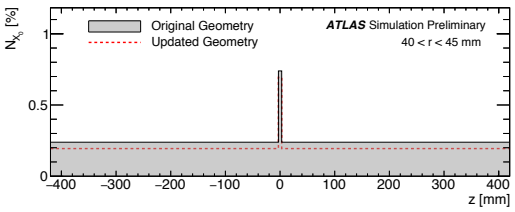
(c) Beam Pipe



(d) IPT



(e) IBL



(f) IST

FIGURE 4.2: The radial distribution of the differential radiation length, $\Delta N_{X_0}/\Delta r$, for $|z| < 300$ mm and (a) for $r < 600$ mm and (b) for $20\,\mathrm{mm} < r < 75$ mm for the *original* geometry and the *updated* geometry. The simulated material is sampled for each z-position along a straight radial path (perpendicular to the beam line). Number of radiation lengths as a function of the z-coordinate for different radial sections are shown in (c)–(f). Ref. [113].

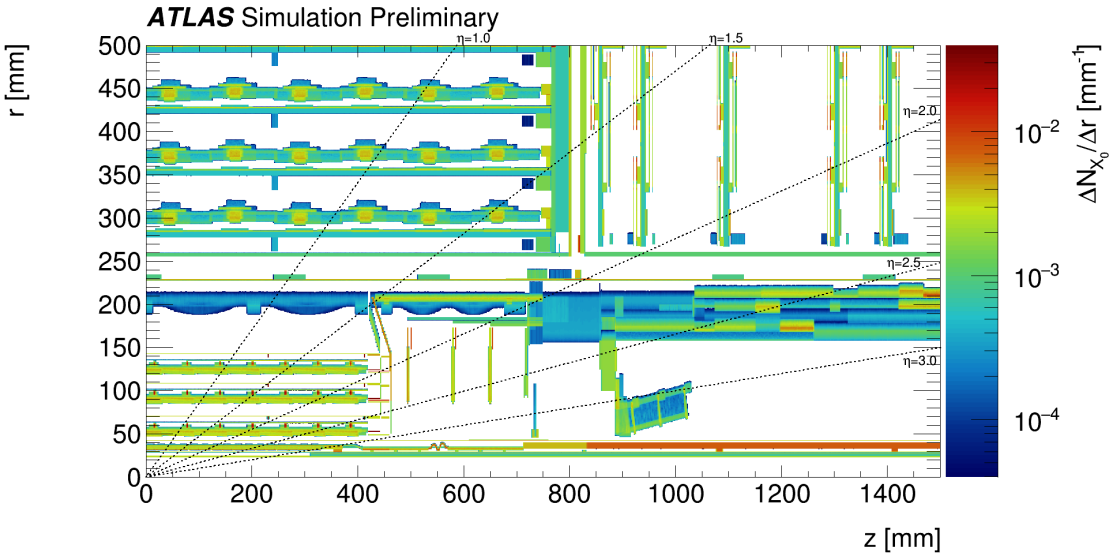


FIGURE 4.3: The r–z distribution of the differential number of radiation lengths, $\Delta N_{X_0}/\Delta r$, for the *updated* geometry model of a quadrant of the inner detector barrel region of the pixel detector and the SCT. The simulated material is sampled for each z-position along a straight radial path (perpendicular to the beam line). From [113].

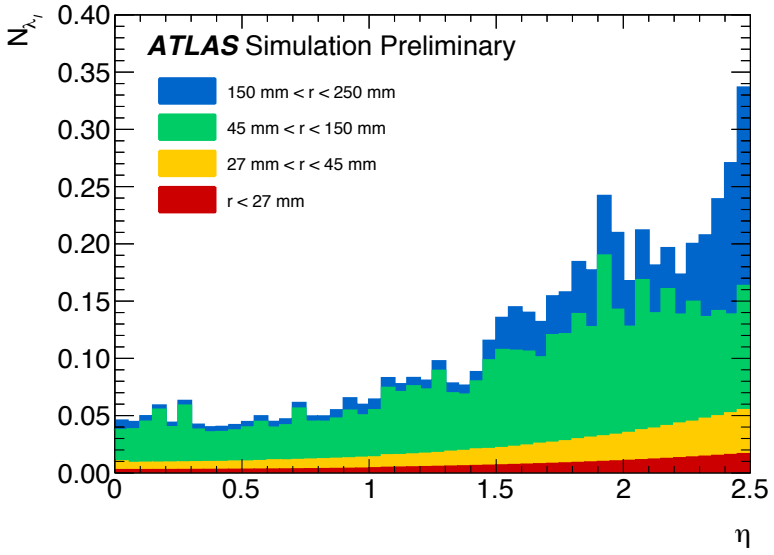
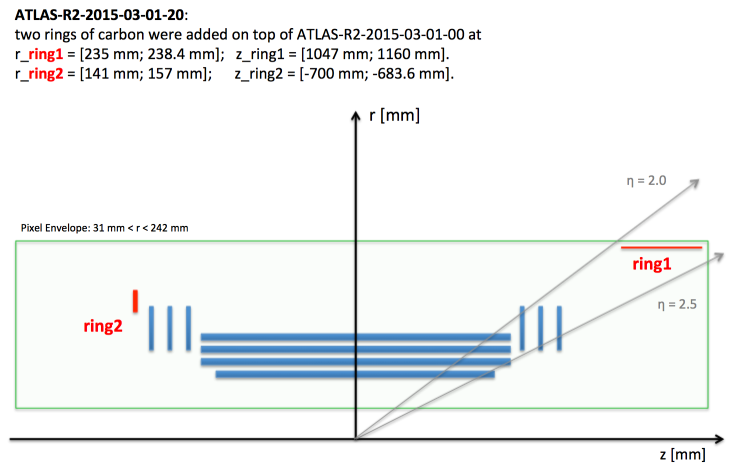
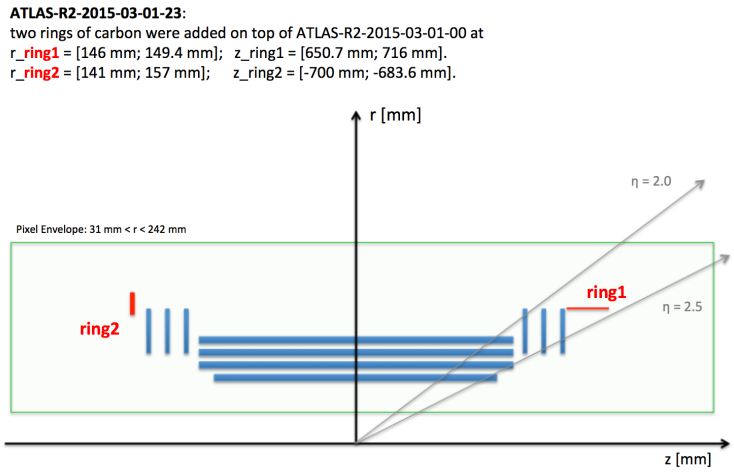


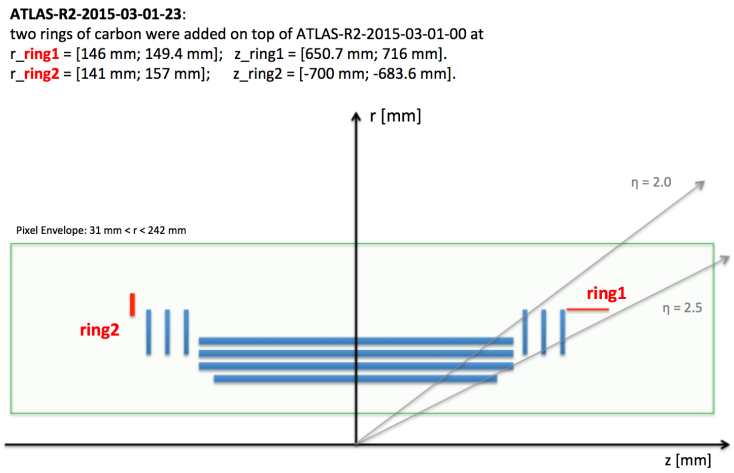
FIGURE 4.4: The amount of material associated with nuclear interactions, $N_{\lambda_I} = \int \mathrm{d}s\, \lambda_I^{-1}$, averaged over ϕ, as a function of η in the positive η range integrated up to $r = 250$ mm for the *updated* geometry model. The simulated material is sampled from $z = 0$ along a straight path with fixed ϕ. The material within the inner detector is shown separately for the regions $r <$ 27 mm, 27 mm $< r <$ 45 mm, 45 mm $< r <$ 150 mm and 150 mm $< r <$ 250 mm, corresponding approximately to the beam pipe, IBL, pixel barrel and pixel service region, respectively. The statistical uncertainty in each bin is negligible. Ref. [113].



(a) Ring Layout 1 (ATLAS-R2-2015-03-01-20)



(b) Ring Layout 2 (ATLAS-R2-2015-03-01-23)



(c) Ring Layout 3 (ATLAS-R2-2015-03-01-27)

FIGURE 4.5: Sketch of the (a) ring layout 1 (geometry model `ATLAS-R2-2015-03-01-20`), (b) ring layout 2 (geometry model `ATLAS-R2-2015-03-01-23`) and (c) ring layout 3 (geometry model `ATLAS-R2-2015-03-01-27`) in which the location of the two rings of carbons is highlighted.

TABLE 4.1: The list of MC samples used in the analyses, with the base geometry model, type distortion and the event generator used.

Base Geometry	Distortion	Event Generator
updated	nominal	PYTHIA 8 (A2) ND
	nominal	EPOS (LHC)
	IBL $+10\%$	PYTHIA 8 (A2) ND
	IBL -10%	PYTHIA 8 (A2) ND
	IST $+10\%$	PYTHIA 8 (A2) ND
	IST -10%	PYTHIA 8 (A2) ND
original	nominal	PYTHIA 8 (A2) ND
	nominal	EPOS (LHC)
	overall $+2.5\%$	PYTHIA 8 (A2) ND
	overall $+5\%$	PYTHIA 8 (A2) ND
	overall $+10\%$	PYTHIA 8 (A2) ND
	pixel service $+10\%$	PYTHIA 8 (A2) ND
	pixel service $+25\%$	PYTHIA 8 (A2) ND
	pixel service $+50\%$	PYTHIA 8 (A2) ND
	pixel service $+50\%$	EPOS (LHC)
	ring layout 1, Figure 4.5(a)	PYTHIA 8 (A2) ND
	ring layout 2, Figure 4.5(b)	PYTHIA 8 (A2) ND
	ring layout 3, Figure 4.5(c)	PYTHIA 8 (A2) ND

4.4 Secondary vertices

The rate and the features of secondary vertices (see Section 3.2 for the definition of primary and secondary vertices) from hadronic interactions and photon conversions in real data samples can be used to probe the accuracy of the ID material distribution when compared, in both qualitative and quantitative manner, to analogous samples of simulated pp collisions.

Photon conversions are clean electromagnetic processes which exhibit a high reconstruction purity. The radial vertex position resolution is around 2 mm, limited by the zero opening angle of the electron–positron pair.

In contrast, hadronic interactions are a complex phenomenon which is difficult to model in simulation and their reconstruction suffers from backgrounds associated with hadron decays and combinatoric fake vertices. However, resolutions of $\mathcal{O}(0.1)$ mm can be achieved due to large opening angles between the outgoing particles. This facilitates a detailed radiography of the material, including minute components, e.g. the capacitors mounted on the surfaces of the pixel modules, allowing their location to be determined precisely.

Qualitative comparisons of the distributions of reconstructed vertices from photon conversions and hadronic interactions in data and simulation samples can identify poorly described or inaccurately-positioned components within the ID geometry model. This method is effectively able to probe the central barrel region of $|z| < 400$ mm in a radial range from the beam pipe up to the first layer of the SCT at $r \simeq 300$ mm, and is suitable to probe the barrel structures including the IBL and the new beam pipe. The measurable ID volumes are divided into several groups by radii, which are listed in Table 4.2 and are hereafter referred to as radial regions. The boundaries are chosen to naturally classify different barrel layers of the ID. For what concerns the photon conversion analysis, the regions of IPT, IBL and IST are combined into one region denoted by IBL since the method does not have a good enough resolution to differentiate these components.

These two analyses do not attempt to make any quantitative statements on the material distribution of the ATLAS ID in terms of the radiation length X_0 nor the nuclear interaction

TABLE 4.2: Definition of the radial regions used for data/MC comparison. Note that in the case of the photon conversion analysis, the IPT, IBL and IST regions are always considered together, due to the limited resolution of the approach. The corresponding z region used for the data/MC comparison is $|z| < 400\,\mathrm{mm}$ for all of the radial regions listed.

Radial regions	Radial range [mm]	Description
BP	22.5–26.5	beam pipe
IPT	28.5–30.0	inner positioning tube
IBL	30.0–40.0	IBL staves (for photon conversion: IPT+IBL+IST)
IST	41.5–45.0	inner support tube
PIX1	45.0–75.0	first pixel barrel layer (pixel B-layer)
PIX2	83–110	second pixel barrel layer (pixel layer-1)
PIX3	118–145	third pixel barrel layer (pixel layer-2)
PSF	180–225	pixel support frame
PST	225–240	pixel support tube
SCT-ITE	245–265	SCT inner thermal enclosure
SCT1	276–320	first SCT barrel layer
SCT2	347–390	second SCT barrel layer
Gap1	73–83	material gap between PIX1 and PIX2
Gap2	155–185	material gap between PIX3 and PST

length λ_I. Such measurements would require an unfolding of the reconstructed position distributions which, in the case of the photon conversions, for example, relies upon a detailed understanding of the behaviour of the photon flux through the detector as a function of a variable downstream material integral. To obtain such an understanding additional studies would be required.

4.4.1 Hadronic Interaction Analysis

Hadronic interactions are the result of inelastic scattering between hadrons and atoms in detector materials. Daughter particles may have large opening angles with respect to the incoming hadron. When multiple daughter particles of a hadronic interaction are reconstructed as tracks, it is possible to reconstruct accurately the position of the interaction vertex. Compared to photon conversion, the radial position resolution of the vertex reconstruction is by far better due to large opening angles between daughter particles. This enables a precise radiography of the detector material and to determine the location of minute material components of sub-millimeter size, e.g. the surface-mount capacitors on the Pixel front-end chips, as can be seen in Figure 4.6 which shows the distribution of hadronic-interaction candidate vertices in the x–y plane for the data and the PYTHIA 8 MC simulation for the *updated* geometry simulation. The qualitative features of the two distributions indicate that the geometry model description is generally accurate.

The reconstructed tracks from the hadronic-interaction daughter particles tend to have relatively large impact parameters. This enables to differentiate between candidate tracks and primary particles by requiring large impact parameters. On the other hand, the reconstruction efficiency at large impact parameters is smaller. In particular, the default silicon track reconstruction has a hard cutoff on the transverse impact parameter at $|d_0^{\mathrm{BL}}| < 10$ mm, and only the TRT back-tracking is available for $|d_0^{\mathrm{BL}}|$ larger than 10 mm. In the Run 1 material studies based on hadronic-interactions [127], the so-called re-tracking technique was applied by re-running the track reconstruction on ESD files (see Section 2.4.7 for details about the event data models), and the impact parameter cutoff was effectively removed from silicon tracks. With this approach, the volume in which hadronic-interaction

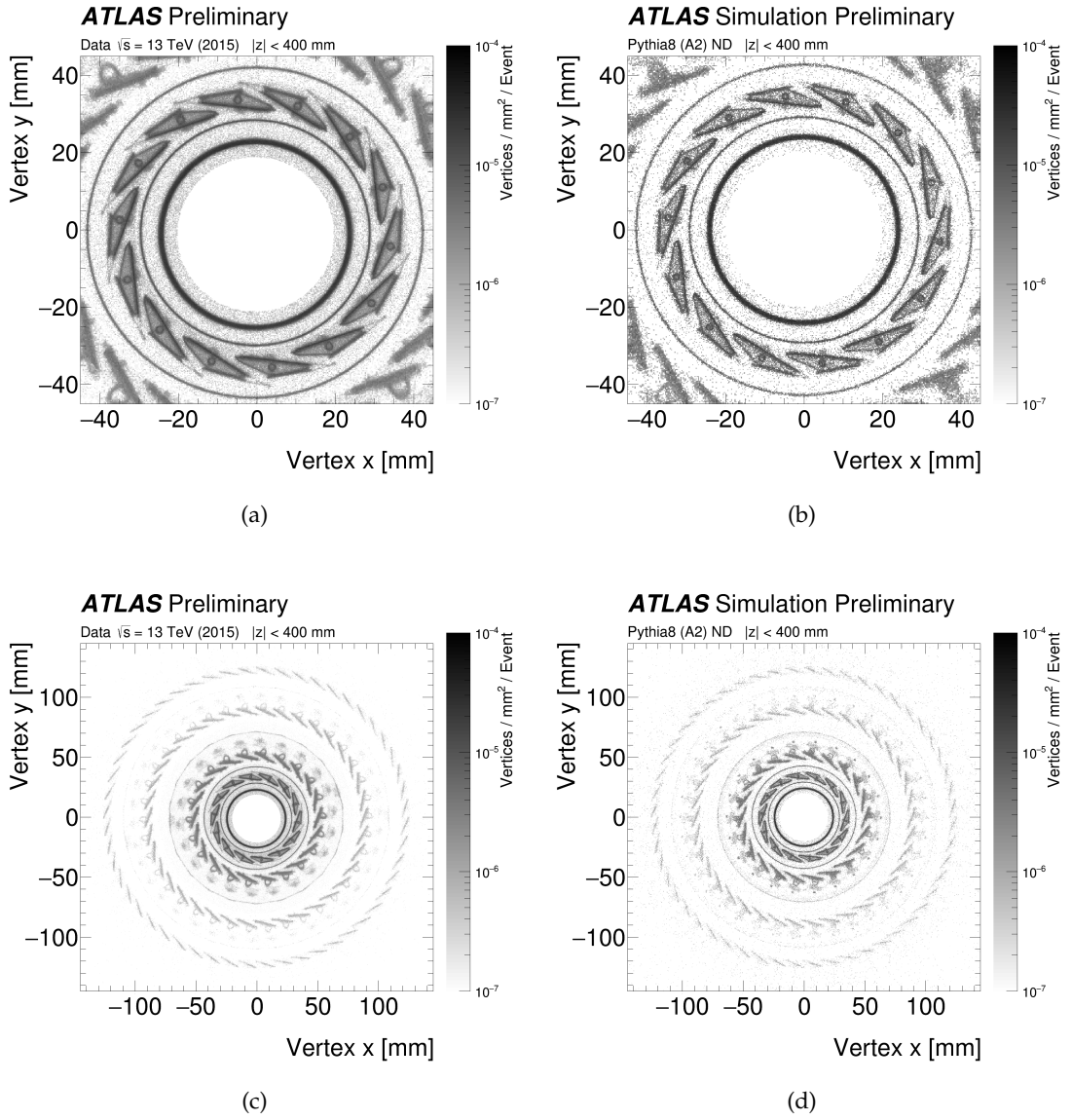


FIGURE 4.6: Distribution of hadronic-interaction vertex candidates in data in $|\eta| < 2.4$ and $|z| < 400$ mm for data and the PYTHIA 8 MC simulation with the *updated* geometry model. (a), (b) The x–y view zooming in to the beam pipe, IPT, IBL staves and IST, and (c), (d) of the pixel detector. Ref. [113].

vertexing can probe covers up to $r \lesssim 400$ mm and $|z| \lesssim 1000$ mm with much improved efficiency compared to the standard track reconstruction.

However, as mentioned in Chapter 2, during LS1 a significant change was applied to the event data model which resulted in the xAOD format to be used for Run 2 analyses. A technical barrier was accidentally introduced and this prevented to perform re-tracking. Although this problem was resolved in Athena release 20.7, this analysis does not employ the re-tracking technique, and only uses the standard track reconstruction. Therefore the probed region is approximately the same as the initial hadronic-interaction study in Run 1 [126], which does not reach further than the first SCT barrel layer.

In order to understand the tracking efficiency of large impact parameter tracks, studies of K^0_S samples were performed, as explained in [113]. The event selection is summarised

in Table 4.5.

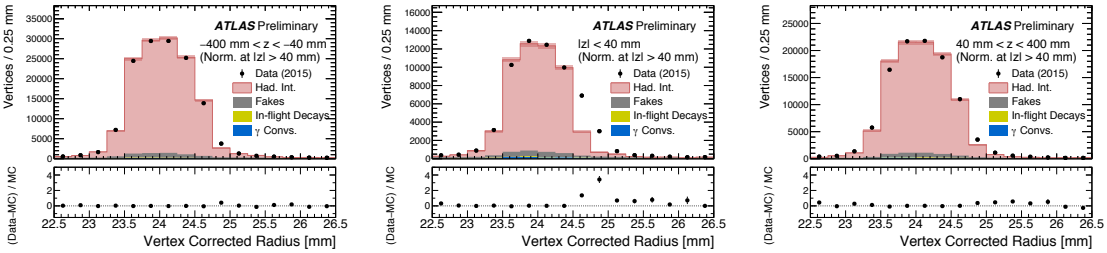
The validity of the inner detector description is assessed by comparing the yields of the hadronic-interaction candidates between data and the MC simulations, as it can be seen in Figure 4.8.

The acceptance of the hadronic interaction reconstruction is such that interactions within the beam pipe are only reconstructed within the range $|z| < 250\,\mathrm{mm}$. The description of the geometry model is generally good, but an excess of candidates is observed in data at the centremost part of the beam pipe within $|z| < 40\,\mathrm{mm}$. The radial distributions of the beam pipe in different z-ranges are shown in Figure 4.7 normalised to the rate in the beam pipe at $|z| > 40\,\mathrm{mm}$. While the radial distribution is well described for $|z| > 40\,\mathrm{mm}$, there is a significant excess within $|z| < 40\,\mathrm{mm}$, which appears to be localised to the outer surface of the beam pipe. The excess is 12% of the rate at $|z| < 40\,\mathrm{mm}$, corresponding to approximately $N_{X_0} = 0.03$–0.04%. This excess is also observed in the photon conversion case, as shown in Figure 4.9(a) in the region $|z| < 50\,\mathrm{mm}$. Investigations of engineering records suggest that several 60-μm-thick polyimide tape layers are missing in the simulated description of the beam pipe in the *updated* geometry model. Since the geometry model description of the beam pipe at $|z| > 40\,\mathrm{mm}$ is assumed to be accurate within 1% precision, the rate observed in this region is used to derive a normalisation factor for the MC simulation common to the all radial regions. For the IBL staves, the rate in the simulation with the *original* geometry model is found to be significantly smaller around $r \simeq 32\,\mathrm{mm}$, as shown in Figure 4.8(a). The corresponding deficit is also observed for photon conversions. Investigations clarified that some surface mounted components, e.g. capacitors, located on the front-end chips of the IBL modules are missing in the *original* geometry model. The *updated* geometry model was created to solve this issue and it improved the agreement with the data significantly. The radial position of the IPT matches very well with that of the MC geometry, while the thickness of the tube is larger than the nominal value of 0.325 mm. For the IST, the rate of the data is approximately 16% smaller than the *original* geometry model. In the *updated* geometry, the density of both the IPT and IST is scaled and agreement of rate ratio of the data to the *updated* geometry simulation is improved. The description of the three outer barrel layers of the Pixel detector was investigated and improved during in Run 1, thus the distribution is reasonably described by the MC simulation in all three layers. Some small deficit in the MC simulation is observed around $r \simeq 50\,\mathrm{mm}$ and $r \simeq 85\,\mathrm{mm}$ which may indicate some missing components on the pixel modules. Moreover, a clear discrepancy in the region of the staves and cabling structures is observed at $58\,\mathrm{mm} < r < 72\,\mathrm{mm}$ and $96\,\mathrm{mm} < r < 112\,\mathrm{mm}$. An excess in the MC simulation is also observed in the photon conversion measurements in this region. The material composition of the PSF, PST and SCT barrel layers are also unchanged since Run 1 and data and simulation are in good agreement.

Systematic Uncertainties

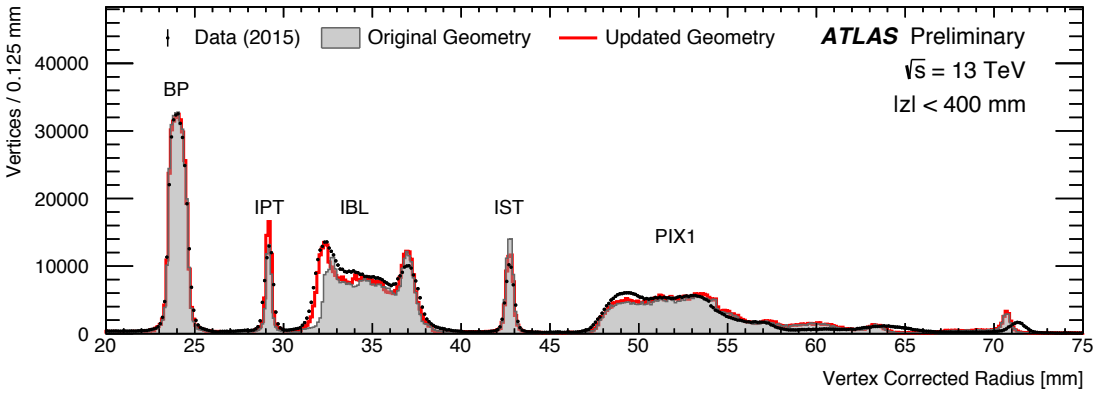
In contrast to the electromagnetic photon conversion, hadronic interactions are a complex phenomenon and they are only phenomenologically described in the simulation. Inevitably, the accuracy of the hadronic-interaction modeling is a source of uncertainty in this measurement.

Reducible background vertices from decays in-flight of K^0_S, Λ^0 or photon conversions are rejected by vetoing the invariant mass regions close to the masses of these particles. Some decay vertices however cannot be excluded, e.g. $K^\pm$ decaying to three charged pions in which one pion is not reconstructed. Fake vertices, which are reconstructed by either accidental vertex reconstruction from random combination of tracks originating from different vertices or contamination of fake tracks, are the dominant background contribution.

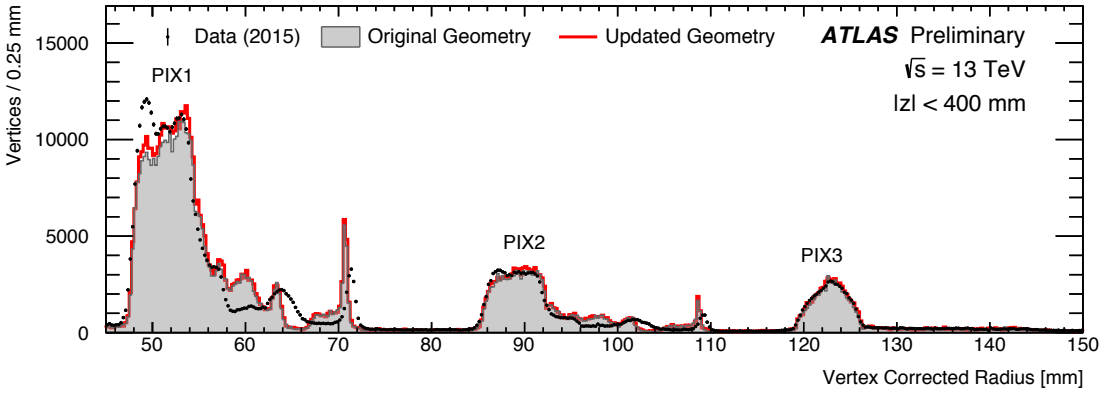


(a) $-400\ \text{mm} < z < -40\ \text{mm}$ (b) $|z| < 40\ \text{mm}$ (c) $40\ \text{mm} < z < 400\ \text{mm}$

FIGURE 4.7: Comparison between data and simulation of the r-distribution of the hadronic interaction candidates at the beam pipe ($22.5\ \text{mm} < r < 26.5\ \text{mm}$) in different z sections. The MC simulation is normalised to the data using the rate at $|z| > 40\ \text{mm}$. An excess is observed at the outer surface of the beam pipe for $|z| < 40\ \text{mm}$. Ref. [113].



(a) 20 mm < r < 45 mm



(b) 45 mm < r < 150 mm

FIGURE 4.8: Comparison of the radial distribution of hadronic interaction candidates between data and simulation (*original* and *updated* simulations) for(a) 20 mm < r < 45 mm and (b) 45 mm < r < 150 mm. Ref. [113].

Hadronic-interaction vertices mostly originate from hadrons coming directly from the pp collision. The predicted flux of the primary particles varies with Monte Carlo generators. In this study, the flux was normalised *in-situ* by comparing the yield of hadronic

interactions in the beam pipe where the material composition and bugdet is considered to be well understood. This allows to eliminate the uncertainty of the primary-particle flux.

The systematic uncertainties and the correction procedures applied to this analysis are described in details in [113]. To summarise, the main components which contribute to the systematic uncertainty are:

- the physics modelling of the hadronic interactions in GEANT4 and the `FTFP_BERT` physics list, assessed via a data-driven technique;
- the background components, such as fake vertices for smaller radii up to around the pixel support tube and for the SCT region, contamination of in-flight decays and photon conversions;
- primary particle flux, estimated by comparing Pythia8 and EPOS results;
- reconstruction efficiency;
- measurement closure.

Hadronic Interaction Results

The ratio of the hadronic interaction vertices measured in data and in the *updated* simulation, after background subtraction, is shown in Figure 4.10(a) together with the related uncertainties.

The total uncertainty is dominated by systematic uncertainties. The physics modelling is the major source of systematic uncertainty for radii smaller than the IST. At larger radii, the background uncertainty becomes significant as the purity of the hadronic interactions decreases.

The total measurement uncertainty for the background subtracted comparison is estimated to be 7%–13% for the inner radial regions from the beam pipe up to the PIX3, and 22%–42% for outer radial regions from the PSF to the SCT1. The large variance in the size of the uncertainty between the radial regions arises mainly from variations in the purity of the hadronic interaction candidate vertices. The uncertainty on the background rate is enhanced at the PSF and outer layers. The uncertainty of the physics modelling is smaller in the IPT–PIX2 regions relatively to the beam pipe, reflecting the large correlation in the values of the rate ratio resulting from changes associated with the cuts used in the uncertainty evaluation. The beam pipe data/simulation ratio is not equal to unity due to the presence of the excess within $|z| < 40$ mm, which is excluded from the normalisation.

The results were also obtained from the background-inclusive rate ratio and for the layers which are unchanged since Run 1 (PIX1–PIX3 and SCT1). They are found to be consistent with the previous analyses presented in Refs. [126, 127].

The results for all the radial regions before and after background subtraction are summarised in Table 4.3.

4.4.2 Photon Conversion Analysis

The probability, P_{conv}, for a photon (with energy $E_\gamma > 2m_e$, where m_e is the electron mass) to convert into an electron pair while traversing an amount of material characterised by a fraction of x of a radiation length X_0 is given by

$$P_{\text{conv}} = 1 - \exp\left(-\frac{7}{9}\frac{x}{X_0}\left[1-\xi\right]\right), \tag{4.3}$$

where ξ is a parameter which depends on the energy of the incident photon, E_γ [130]. For photons with $E_\gamma \approx 1$GeV, $\xi \approx 0.05$, tending to zero as $E_\gamma \to \infty$. A corresponding

photon survival probability, $P_{\text{surv}} = 1 - P_{\text{conv}}$ can also be defined. One can also expand Equation 4.3 for small x/X_0 to give,

$$P_{\text{conv}} = \frac{7}{9}\frac{x}{X_0}\left[1-\xi\right] - \frac{49}{162}\left(\frac{x}{X_0}\left[1-\xi\right]\right)^2 + \mathcal{O}\left(\left(\frac{x}{X_0}\right)^3\right), \tag{4.4}$$

thus at leading term in x/X_0 (the next-to-leading term is small, around 1% for $x/X_0 = 0.03$, the approximate material content of a single pixel layer), the photon conversion probability is directly proportional to the amount of material it traverses. This behaviour makes a measurement of the conversion rate of photons produced in pp collisions an ideal probe of the material content of the ATLAS detector. The di-photon decays of copiously produced light neutral mesons (particularly the π^0) for minimum bias pp provide a high statistics source of photon conversions with which to perform such a measurement.

The aim of this analysis is to use photon conversions to probe the accuracy of the ID material content within the ATLAS detector simulation. The analysis is performed through the reconstruction of a high purity sample of photon conversion candidates within a pp collision data sample and compare it, in both a qualitative and quantitative manner, to an analogous sample of photon conversion candidates reconstructed in simulated pp collisions. The qualitative comparisons of reconstructed photon conversions in data and simulation samples aims to identify absent or inaccurately positioned components of the ATLAS ID that may affect the performance of the ATLAS simulation as a tool for physics analysis.

The analysis aims to perform a quantitative assessment of the accuracy of the ID material content within the ATLAS detector simulation through a comparison of the yields of reconstructed photon conversions within different regions of the detector, based on the position of the reconstructed photon conversion vertex. The event selection is summarised in Table 4.5. All quantitative comparisons are presented in terms of the direct observable, the relative reconstructed photon conversion yield in data and simulation, as it can be seen in Figure 4.9.

Similar studies with photon conversions were performed with the Run 1 ATLAS Inner Detector, and the study presented here aims to build upon the experience gained from these studies [131].

Systematic Uncertainties

As for the hadronic interactions studies, the systematic uncertainties in this analysis are described in details in [113]. To summarise, the main components are:

- the simulated photon flux normalisation;
- modelling of primary photon flux, estimated by comparing Pythia8 with EPOS in place of the data;
- photon conversion reconstruction efficiency;
- measurement closure.

Photon Conversion Results

The ratio between data and *updated* simulation for photon conversions is shown in Figure 4.10(b) together with the related uncertainties. The ratio remains within 1.00 ± 0.17 for all of the radial sections, spanning the cylindrical region $r < 325\,\text{mm}$ and $|z| < 400\,\text{mm}$. Good agreement between data and simulation is observed for the IBL region and the first SCT layer. An excess in the observed conversion rate in data with respect to simulation of

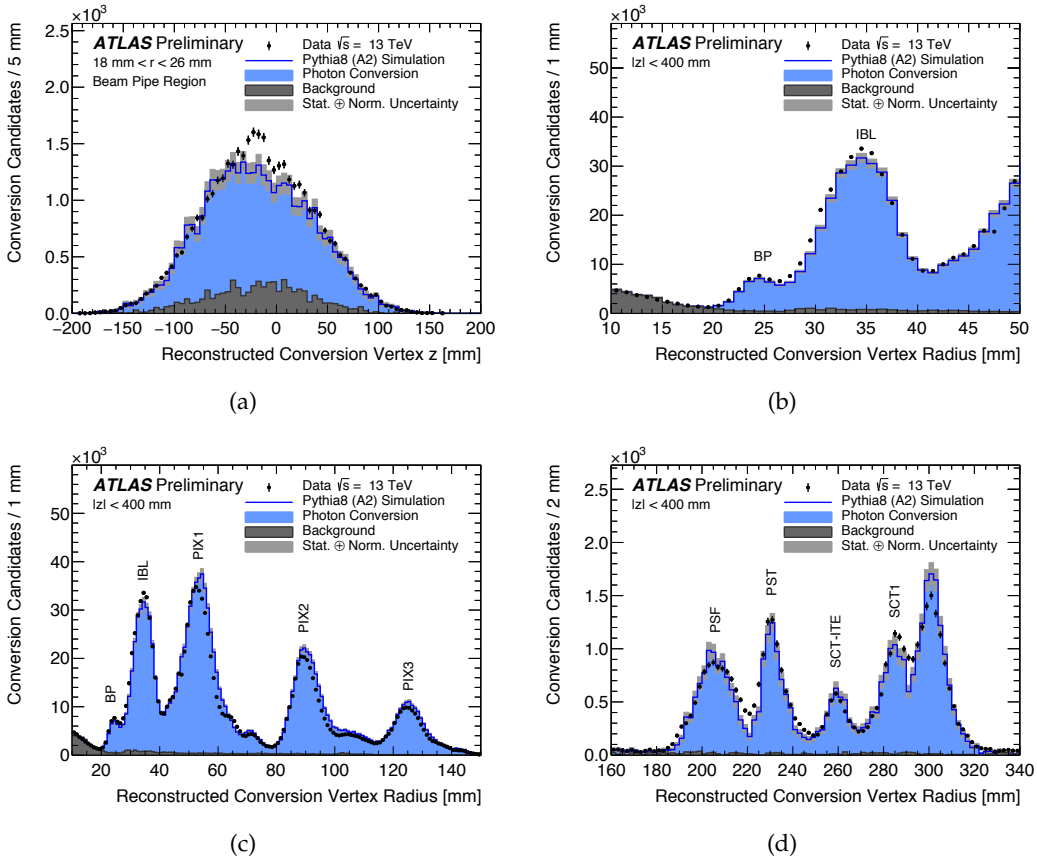


FIGURE 4.9: Conversion vertex position distributions for PYTHIA 8 simulation compared to data, including: (a) The conversion vertex z position distribution in the beam pipe radial region and the conversion vertex radial distributions in (b) the beam pipe, (c) pixel and (d) inner SCT regions. Ref. [113].

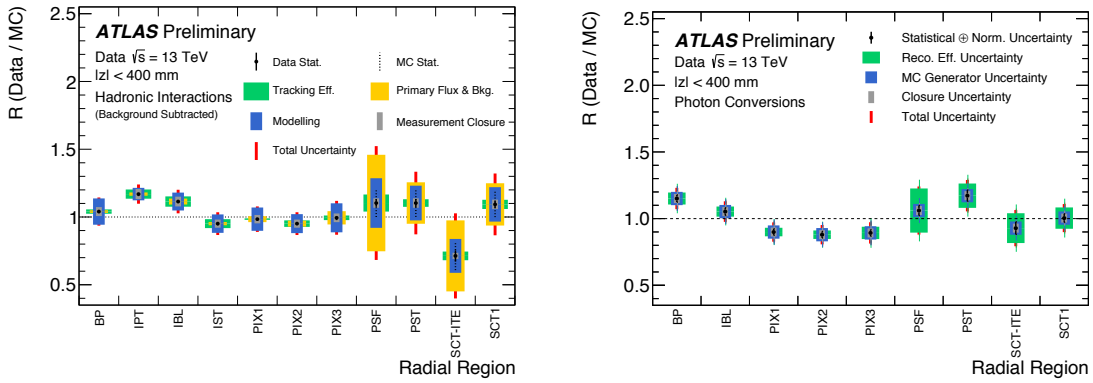
around 10%–15% is observed in the beam-pipe region. This excess is observed also with hadronic interactions and the cause is thought to be due to a localised region of material missing from the *updated* geometry. The largest deviations from unity in the measured ratio values are observed in all three original pixel layers, which exhibit a systematic deficit in the conversion rate of between 10%–12% lower than that predicted by the simulation. The results for all the radial regions are summarised in Table 4.4.

4.4.3 Combined Results

Figure 4.11 shows both the hadronic interaction and photon conversion measurements. While sensitive to slightly different properties of the ID material (the nuclear interaction and radiation lengths), the two measurements are compatible in all radial regions within the uncertainties.

4.5 Track-Extension Efficiency Analysis

Nota Bene: The extensive description of the Track-Extension efficiency analysis reported here follows the chronological evolution of this study.



(a) (b)

FIGURE 4.10: Rate ratio betwen data and simulation for (a) hadronic interactions and (b) photon conversions using the *updated* geometry model. The hadronic interaction results are shown after background subtraction. The systematic uncertainties for each radial section are also shown. Ref. [113].

TABLE 4.3: Hadronic interaction rate ratio of data with respect the EPOS MC simulation using the *updated* geometry model for different radial sections. *Syst.(model)* is the uncertainty of the physics modelling of hadronic interactions, *Syst.(flux & bkg.)* is the primary particle flux uncertainty and the uncertainty of the fakes and decays backgrounds, *Syst.(eff.)* is the systematic uncertainty of track reconstruction efficiency, and *Syst.(closure)* is the uncertainty of the closure of the measurement. The total uncertainty is calculated from the sum in quadrature of the statistical and systematic uncertainties.

Hadronic interaction inclusive rate ratio: R_i^{incl}								
Radial region	Value	Stat.(data)	Stat.(MC)	Syst.(model)	Sys(flux & bkg.)	Syst.(eff.)	Syst.(closure)	Total uncertainty
BP	1.04	±0.00	±0.01	±0.09	±0.01	±0.02	±0.03	±0.10
IPT	1.16	±0.01	±0.01	±0.05	±0.01	±0.03	±0.04	±0.07
IBL	1.10	±0.00	±0.01	±0.07	±0.02	±0.03	±0.04	±0.09
IST	0.96	±0.01	±0.01	±0.07	±0.01	±0.03	±0.03	±0.08
PIX1	0.99	±0.00	±0.01	±0.08	±0.02	±0.01	±0.03	±0.09
PIX2	0.96	±0.00	±0.01	±0.07	±0.02	±0.02	±0.03	±0.08
PIX3	1.00	±0.00	±0.01	±0.10	±0.03	±0.01	±0.03	±0.11
PSF	1.03	±0.01	±0.03	±0.12	±0.11	±0.02	±0.03	±0.17
PST	1.06	±0.02	±0.05	±0.14	±0.09	±0.02	±0.03	±0.17
SCT-ITE	0.89	±0.02	±0.05	±0.07	±0.09	±0.01	±0.03	±0.13
SCT1	1.04	±0.01	±0.04	±0.11	±0.07	±0.02	±0.03	±0.14
Hadronic interaction background-subtracted rate ratio: R_i^{subtr}								
Radial region	Value	Stat.(data)	Stat.(MC)	Syst.(model)	Syst.(flux & bkg.)	Syst.(eff.)	Syst.(closure)	Total uncertainty
BP	1.04	±0.00	±0.01	±0.10	±0.01	±0.02	±0.03	±0.10
IPT	1.17	±0.01	±0.02	±0.05	±0.01	±0.03	±0.04	±0.07
IBL	1.11	±0.00	±0.01	±0.07	±0.02	±0.04	±0.04	±0.09
IST	0.95	±0.01	±0.01	±0.07	±0.01	±0.03	±0.03	±0.08
PIX1	0.98	±0.00	±0.01	±0.09	±0.02	±0.01	±0.03	±0.09
PIX2	0.95	±0.01	±0.01	±0.07	±0.02	±0.02	±0.03	±0.08
PIX3	1.00	±0.01	±0.02	±0.11	±0.05	±0.02	±0.03	±0.13
PSF	1.10	±0.03	±0.10	±0.18	±0.36	±0.06	±0.04	±0.42
PST	1.10	±0.03	±0.10	±0.13	±0.15	±0.03	±0.04	±0.23
SCT-ITE	0.71	±0.04	±0.11	±0.13	±0.27	±0.03	±0.02	±0.31
SCT1	1.09	±0.03	±0.09	±0.13	±0.16	±0.03	±0.04	±0.22

Some track- or tracklet-level cuts, as reported in Table 4.5, were not originally introduced and they were found to improve the method only after a more thorough optimization. The differences can be noticed when comparing the results reported in the preliminary publication [112] and in the final one [113]. For this reason, the data-driven correction applied to the tracking efficiency in the Charged Particle Multiplicity measurement [80], which was based on the preliminary material studies [112], has been updated to take into account the new results [113]. This will be explained in Chapter 5.

In addition, the name of the method itself changed when the measurement converged

TABLE 4.4: Photon conversion rate ratio and associated uncertainties in data measured with respect to simulation with the *updated* geometry. Measurements are presented in nine radial regions of the detector in the cylindrical region $r < 325\,\mathrm{mm}$ and $|z| < 400\,\mathrm{mm}$. *Stat.(data & MC)* is the statistical and normalisation uncertainty, *Syst.(eff.)* is the systematic uncertainty associated with the reconstruction efficiency, *Syst.(MC gen.)* is the systematic uncertainty associated with the choice of MC Generator and *Syst.(closure)* is the systematic uncertainty associated with the closure of the measurement. The total uncertainty is calculated from the sum in quadrature of the statistical and systematic uncertainties.

Photon conversion rate ratio: R_i							
Radial region	Value	Stat.(data & MC)	Syst.(MC gen.)	Syst.(eff.)	Syst.(closure)	Syst.(total)	Total uncertainty
BP	1.15	±0.03	±0.05	±0.04	±0.03	±0.08	±0.08
IBL	1.05	±0.03	±0.05	±0.04	±0.03	±0.07	±0.08
PIX1	0.90	±0.02	±0.05	±0.03	±0.03	±0.07	±0.07
PIX2	0.88	±0.02	±0.05	±0.03	±0.03	±0.07	±0.07
PIX3	0.89	±0.02	±0.05	±0.05	±0.03	±0.08	±0.08
PSF	1.06	±0.04	±0.05	±0.17	±0.03	±0.18	±0.18
PST	1.17	±0.04	±0.05	±0.09	±0.03	±0.11	±0.12
SCT-ITE	0.93	±0.05	±0.05	±0.11	±0.03	±0.13	±0.14
SCT1	1.00	±0.03	±0.05	±0.08	±0.03	±0.10	±0.11

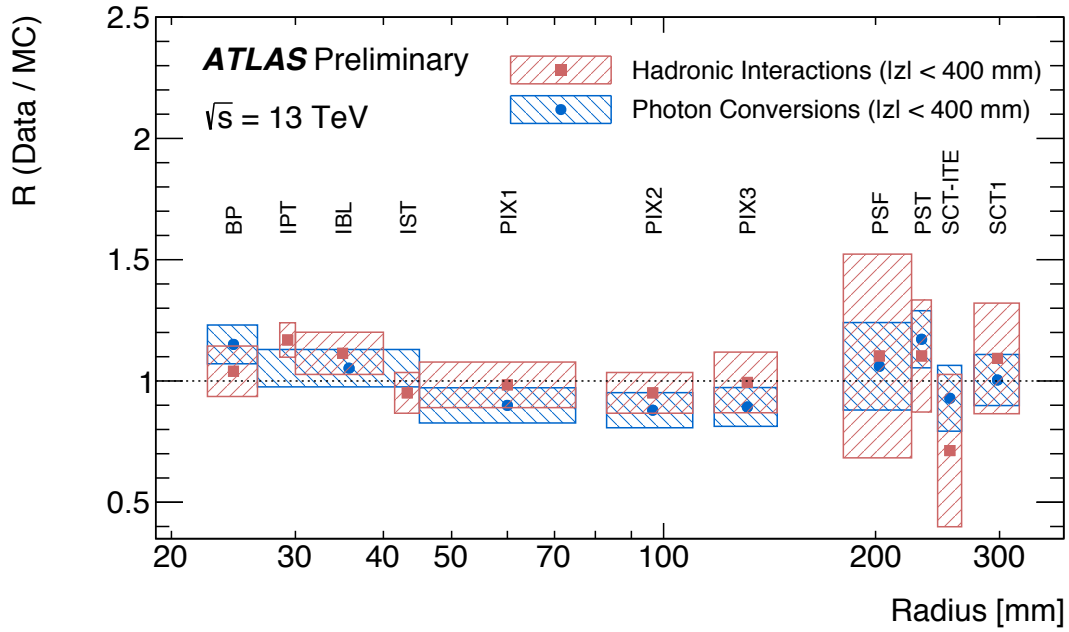


FIGURE 4.11: Comparison of the rate ratio measurements for data versus simulation for hadronic interactions and photon conversions as a function of radius. The horizontal range of each marker represents radial range vertices used in each measurement, while the vertical range represents the total uncertainty. Ref. [113].

towards a publication: the original SCT-Extension Efficiency, which can be still found in certain plots or description, was replaced by Track-Extension Efficiency. The term Track-Extension Efficiency was considered more appropriate because it indicates more explicitely the goal of this study, i.e. measuring the efficiency of extending a track from the Pixel to the SCT detector.

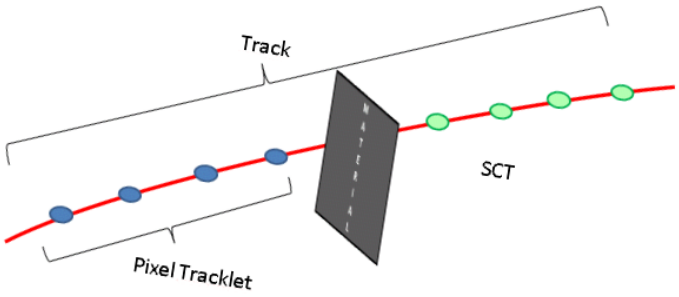


FIGURE 4.12: Schema of the Track-Extension Efficiency method: Pixel tracklets and SCT extended tracks are shown.

4.5.1 Introduction

The pixel services, as described in Section 2.4.2, are positioned between the pixel and SCT detectors in the forward region (approximately $1.0 < |\eta| < 2.5$). This region of the inner detector is more challenging to model within the simulation than the central barrel region, due to the complexity of the structure and the amount as well as the diversity of material, as shown in Figures 2.18, 4.3 and 4.4.

The pixel service region is also expected to exhibit a relatively high rate of hadronic interactions due to the high density of material and the longer path that hadrons produced at high pseudo-rapidity travel. If a charged hadron traversing the region between the pixel and SCT detectors undergoes a hadronic interaction and disappears, it will only leave hits in the pixel detector. A track associated with the particle's trajectory can be reconstructed from pixel detector hits alone (referred to as a *tracklet* hereafter). Hits in the SCT and TRT detectors deposited by the products of the hadronic interaction are unlikely to be associated to this tracklet. The rate of hadronic interactions can then be related to the so-called *track-extension efficiency*, denoted by $\mathcal{E}_{\text{ext}}$. A very simplified sketch of track and tracklets used in the calculation of the track-extension efficiency is shown in Figure 4.12.

The $\mathcal{E}_{\text{ext}}$ is defined as:

$$\mathcal{E}_{\text{ext}} \equiv \frac{n^{\text{matched}}_{\text{tracklet}}}{n_{\text{tracklet}}}\,, \tag{4.5}$$

where n_{tracklet} is the number of tracklets satisfying certain selection criteria and $n^{\text{matched}}_{\text{tracklet}}$ is the number of tracklets that are matched to a track which is reconstructed using all ID hits (referred to as a *combined* track). $\mathcal{E}_{\text{ext}}$ is related to the amount of material crossed by a particle and is therefore dependent on the kinematics and origin of the particle. For sufficiently high-p_{T} particles, when averaging over ϕ and restricting the production vertex z_{vtx} to a sufficiently narrow range, the particle trajectory C can be approximately described as a function of η alone.

The tracklet sample will be contaminated by particles associated with the by-products of long-lived weakly-decaying hadrons and combinatorial fakes in addition to general detector inefficiencies that affect the reconstruction of combined tracks. After accounting for such effects, the difference of $\mathcal{E}_{\text{ext}}(\eta)$ between data and simulation, $\Delta\mathcal{E}^{\text{data}-\text{MC}}_{\text{ext}}(\eta)$, is proportional to the difference in the amount of material between the real detector and the geometry model in terms of the number of nuclear interaction lengths, $\Delta N^{\text{data}-\text{MC}}_{\lambda_I}$:

$$\Delta N^{\text{Data}-\text{MC}}_{\lambda_I}(\eta) \simeq -K(\eta)\cdot\Delta\mathcal{E}^{\text{Data}-\text{MC}}_{\text{ext}}(\eta)\,, \tag{4.6}$$

where $K(\eta)$ is a scale factor and $\Delta N^{\text{Data}-\text{MC}}_{\lambda_I}$ is the difference in the number of nuclear interaction lengths in the real detector and the geometry model. The sensitivity of this

method is proportional to the amount of the material along the track path length but it is impossible to identify accurately the radial position of the material. In-flight decays of short-lived charged particles, e.g. weakly-decaying strange baryons, and fake tracks bias $\mathcal{E}_{\text{ext}}$ to lower values. The factor $K(\eta)$ is introduced, which accounts for the algorithmic reconstruction efficiency of the combined track as well as the fact that the tracklets arise from not only a sample of stable hadrons, but also contain contributions from weakly-decaying hadrons and fake tracks.

4.5.2 Track-Extension Algorithm

Since this analysis is based on Pixel tracklets and on the matching between them and complete tracks reconstructed in the silicon detectors, following the procedure described in Chapter 3, the SCT-extension algorithm `SCTExtension` was written by the author of this thesis. It is used to run over the `ESD` samples and produce a derived `DxAOD` with the additional information necessary to perform the track-extension efficiency analysis (see Section 2.4.7 for details about the event data models). This algorithm allows the number of hits in common between the tracks and the tracklets to be stored, and, if there is at least one common hit between a track and a tracklet, an element link between the two is also stored. This last step allows to retrieve all the information related to the track which is matched to the tracklet.

4.5.3 Data Selection

Tracklets are generally reconstructed by using the inside-out approach down to p_{T} 50 MeV. The tracklets used in the analysis have to pass certain selection criteria: they are required to have at least four pixel hits, and to satisfy $p_{\text{T}} > 500$ MeV and $|\eta| < 2.5$. The requirement on the number of pixel hits is imposed to suppress the contribution from fake tracklets, and the p_{T} requirement to suppress the contamination from non-primary charged particles and weakly-decaying hadrons. To further reject non-primary charged particles, the transverse and longitudinal impact parameters of tracklets are required to satisfy $|d_0| < 2$ mm and $|z_0 \sin\theta| < 2$ mm. Once these requirements have been applied, the fraction of non-primary charged particles in the tracklet sample is approximately 3%. A detailed description of the tracklet selection is summarised in Table 4.5. Combined tracks are required to have at least two SCT hits. A tracklet is classified as *matched*, if the tracklet and a selected combined track share at least one common pixel hit.

In order to reduce the variation in track trajectories associated with a single value of η arising from the variation of the z-position of the primary vertex (z_{vtx}), a requirement of $|z_{\text{vtx}}| < 10$ mm is imposed.

The requirement on the z_{vtx} as well as a tighter cut on the number of SCT hits (increased from 2 to 4) required on the combined track was found to be useful only after many of the analysis checks were already performed, thus it is applied only on the measurements relevant to derive the final results, as it will be better explained in Section 4.5.10.

4.5.4 Sensitivity of the method

To establish a measurement of the material in the ID, based on the track-extension efficiency, it is necessary to ensure that the method has enough sensitivity to catch possible discrepancies between data and Monte Carlo simulation and to discriminate actual discrepancies from statistical fluctuations. Thus, the sensitivity of the method has been checked at the very beginning of these studies by using samples of one million non-diffractive minimum bias events. These events, simulated with the Pythia 8 A2 event

TABLE 4.5: Summary of selection criteria for different methods of hadronic interaction vertex reconstruction, photon conversion vertex reconstruction, track-extension efficiency and transverse impact parameter studies.

Notations:
N_{Si}: number of hits on the track within the pixel and SCT layers;
N_{pix}: number of hits on the track within the pixel layers;
N_{SCT}: number of hits on the track within the SCT layers;
$N_{\mathrm{Si}}^{\mathrm{sh}}$: number of hits on the track within the pixel and SCT layers that are shared with other tracks;
$N_{\mathrm{Si}}^{\mathrm{hole}}$: number of sensors crossed by the track within the pixel and SCT detectors where expected hits are missing.
$N_{\mathrm{pix}}^{\mathrm{hole}}$: number of sensors crossed by the track within the pixel detector where expected hits are missing.
Hadronic Interactions
Requirements applied to tracks associated with primary vertices: the *loose-primary* requirement
$p_{\mathrm{T}} > 400\mathrm{MeV}$ and $\lvert\eta\rvert < 2.5$;
$N_{\mathrm{Si}} \geq 7$; $N_{\mathrm{Si}}^{\mathrm{sh}} \leq 1$; $N_{\mathrm{Si}}^{\mathrm{hole}} \leq 2$; $N_{\mathrm{pix}}^{\mathrm{hole}} \leq 1$; either ($N_{\mathrm{Si}} \geq 7$ and $N_{\mathrm{Si}}^{\mathrm{sh}} = 0$) or $N_{\mathrm{Si}} \geq 10$.
Requirement on primary vertices
at least five tracks satisfying the loose-primary selection criteria are associated with the primary vertex;
pile-up veto.
Acceptance
$\lvert d_0\rvert > 5$ mm and at least one SCT hit, $\chi^2/N_{\mathrm{dof}} < 5.0$ for tracks associated with secondary vertices;
hit pattern recognition for combinatorial fake rejection: see Section **??** for details;
primary vertex position -160 mm $< z_{\mathrm{pv}} < 120$ mm;
secondary vertex $\lvert\eta\rvert < 2.4$, $\lvert z\rvert < 400$ mm and $r > 20$ mm;
number of tracks associated with the secondary vertex is two.
In-flight decay veto
K_{s}^0veto: $\lvert m_{sv}(\pi\pi) - m_{K_{\mathrm{s}}^0}\rvert > 50\mathrm{MeV}$;
photon conversion veto: $m_{sv}(ee) > 100\mathrm{MeV}$;
Λ^0 veto: $\lvert m_{sv}(p\pi) - m_{\Lambda^0}\rvert > 15\mathrm{MeV}$.
Fake rejection
tracks associated with secondary vertex: $p_{\mathrm{T}} > 300\mathrm{MeV}$;
secondary vertex $\chi^2/N_{\mathrm{dof}} < 4.5$.
Photon Conversions
Requirement on primary vertices
at least 15 tracks are associated with the primary vertex;
pile-up veto.
Acceptance
primary vertex position -160 mm $< z_{\mathrm{pv}} < 120$ mm;
tracks associated with secondary vertex: $p_{\mathrm{T}} > 250\mathrm{MeV}$ and $N_{\mathrm{SCT}} \geq 4$;
conversion $p_{\mathrm{T}}^{\gamma} > 1\mathrm{GeV}$ and $\lvert\eta^{\gamma}\rvert < 1.5$.
Quality selection criteria
conversion vertex $\chi^2/N_{\mathrm{dof}} < 1.0$;
the photon trajectory must point back to the primary vertex to within 15 mm in the longitudinal plane and within 4.5 mm. in the transverse plane.

Track-Extension Efficiency	
Tracklet reconstruction	**Requirement on tracklets**
$N_{\mathrm{pix}}^{\mathrm{hole}} \leq 1$;	$p_{\mathrm{T}} > 500\mathrm{MeV}$ and $\lvert\eta\rvert < 2.5$;
at least three non-shared hits;	$\lvert z_{\mathrm{vtx}}\rvert < 10$ mm (introduced in the last iteration);
$p_{\mathrm{T}} > 50\mathrm{MeV}$.	at least four pixel hits: $N_{\mathrm{pix}} \geq 4$;
Requirement on primary vertices	$\lvert d_0\rvert < 2$ mm and $\lvert z_0 \sin\theta\rvert < 2$ mm.
pile-up veto.	**Requirement on combined tracks**
	$p_{\mathrm{T}} > 100\mathrm{MeV}$ and $N_{\mathrm{SCT}} \geq 2$ (4 in the last iteration);
	at least one shared hit with the matched tracklet.

generator, are based on the original geometry, but with the global ID material scaled up by 2.5% (overall +2.5%), 5% (overall +5%) and 10% (overall +10%).

As it can be seen in Figure 4.13, the $\mathcal{E}_{\mathrm{ext}}$ uniformly decreases as the material scale increases for all expansions in p_{T}, η, and ϕ. This decrease can be seen also in Figure 4.15 and Figure 4.16 which show respectively the ratio and the difference in $\mathcal{E}_{\mathrm{ext}}$ for the different amount of material, as function of p_{T}, η, ϕ. However, the change in the efficiency is generally small. This is because the material in the support structures of the pixel detector and the SCT was scaled uniformly, but no additional material was actually added between the sub-detectors.

An average efficiency for the barrel and in the endcaps can also be evaluated, as shown

in Figure 4.14, which makes the statistical separation among the three samples clearly visible.

In total, the observed drop in efficiency is $< 2\%$ at high η for the `Py8(Orig_All+10%)` sample.

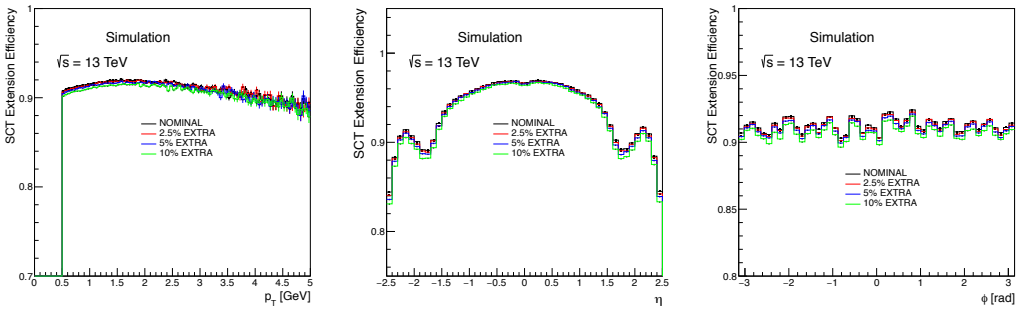


FIGURE 4.13: $\mathcal{E}_{\text{ext}}$ with respect to the pixel-tracklet parameters p_{T}, η, ϕ for the four different analysed samples: nominal (black), overall +2.5% (red), overall +5% (blue), overall +10% (green).

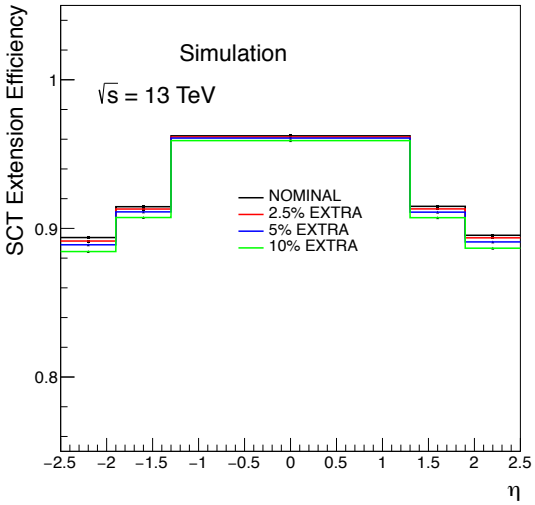


FIGURE 4.14: $\mathcal{E}_{\text{ext}}$ with respect to the pixel-tracklet parameters η in five η bins ($-2.5 < \eta < -1.9$, $-1.9 < \eta < -1.3$, $-1.3 < \eta < 1.3$, $1.3 < \eta < 1.9$, $1.9 < \eta < 2.5$) for the four different geometry model samples: nominal (black), overall +2.5% (red), overall +5% (blue), overall +10% (green).

The sensitivity of $\mathcal{E}_{\text{ext}}$ to additional material has been assessed as well by using non-diffractive Pythia 8 A2 minimum-bias MC samples with only the material in the pixel services regions scaled up by 10%, 25%, 50%, shortly referred to as pixel service +10%, pixel service +25% and pixel service +50% in Section 4.3.

A comparison of $\mathcal{E}_{\text{ext}}$ between the nominal sample and the samples with extra material in the pixel services region is shown in Figure 4.17.

The $\mathcal{E}_{\text{ext}}$ decreases as the material increases for all expansions in p_{T}, η, and ϕ. This decrease can be seen in Figure 4.18 and Figure 4.19 which, as before, show respectively the ratio and the difference in $\mathcal{E}_{\text{ext}}$ for the different amount of material, as function of p_{T}, η, ϕ. The change in the efficiency is most clearly visible at high η values because the material was scaled in the pixel services which are predominant at high η . Not surprisingly, the

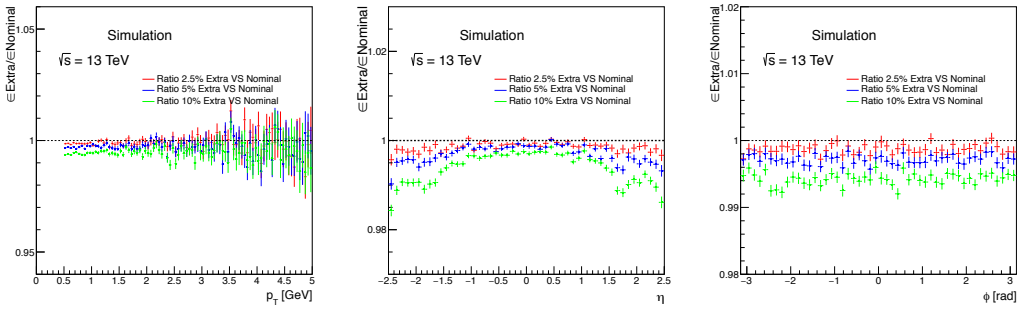


FIGURE 4.15: Ratio $\mathcal{E}_{\text{ext}}$ between nominal and overall +2.5%, overall +5%, and overall +10% geometry samples, with respect to the pixel-tracklet parameters p_{T}, η, ϕ.

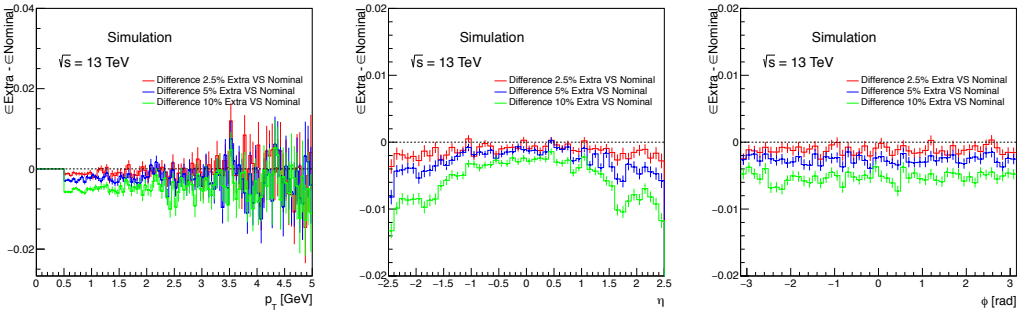


FIGURE 4.16: Difference of $\mathcal{E}_{\text{ext}}$ between nominal and overall +2.5%, overall +5%, and overall +10% geometry samples, with respect to the pixel-tracklet parameters p_{T}, η, ϕ.

$\mathcal{E}_{\text{ext}}$ displays a linear dependence on the amount of material added at high η, as will be discussed in the next paragraph and as shown in Figure 4.24.

In total, the observed drop in efficiency is again $< 2\%$ at high η for the sample with the pixel services material scaled up by 10%.

Given the clear sensitivity of the method to differences in the ID material distribution, which was reached by using one million events, the results obtained in simulation can be safely compared with those obtained in data.

4.5.5 Comparison of data and simulation

In Figure 4.20 the comparison of data, PYTHIA 8 and EPOS simulation based on the *original* nominal geometry is shown for the track-extension efficiency as function of p_{T} and η of the tracklets [112]. The distribution is approximately flat around a value of 97% within $|\eta| < 0.5$, gradually falling towards a local minimum of around 87% at $|\eta| \simeq 1.9$. The efficiency recovers to around 91% at $|\eta| \simeq 2.2$, and continues to fall as $|\eta|$ increases further. This structure of $\mathcal{E}_{\text{ext}}(\eta)$ reflects the distribution of material as a function of η, as shown in Figure 4.4. The MC simulation describes the overall structure of the η dependence, and the agreement is good in the central region of $|\eta| < 1$. Nevertheless, discrepancies at the level of a few percent are observed in the forward region. Figure 4.37(a) shows the average track-extension efficiency as a function of p_{T}, integrated over η and ϕ. The p_{T} dependence also shows discrepances between data and MC simulation, and the data points lie between those corresponding to the MC generators, PYTHIA8 and EPOS. Differences between the two generators will be described in Section 4.5.7.

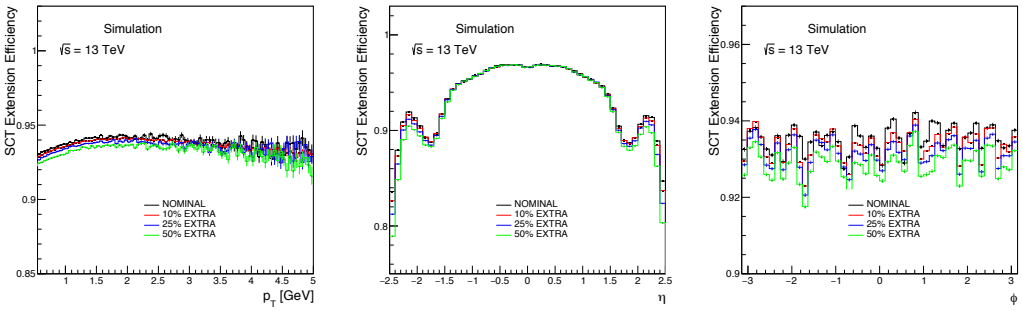


FIGURE 4.17: $\mathcal{E}_{\text{ext}}$ with respect to the pixel-tracklet parameters p_{T}, η, ϕ for the four different geometry models: nominal (black), pixel service +10% (red), pixel service +25% (blue), pixel service +50% (green). The errors are computed with a Bayesian method.

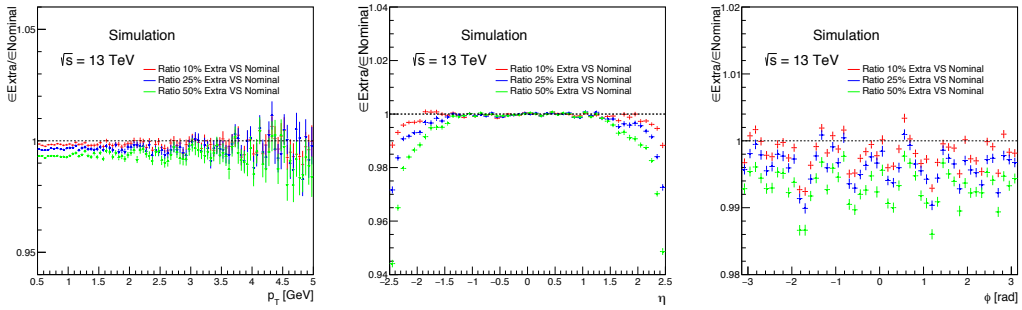


FIGURE 4.18: Ratio of $\mathcal{E}_{\text{ext}}$ between nominal and pixel service +10%, pixel service +25%, and pixel service +50% geometry samples with respect to the pixel-tracklet parameters p_{T}, η, ϕ.

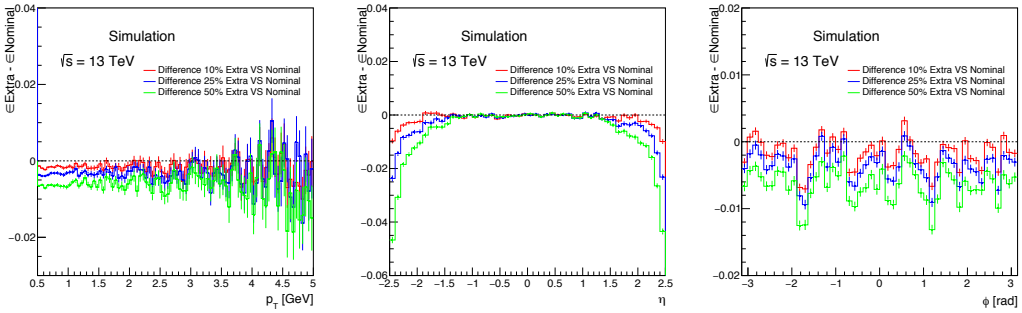
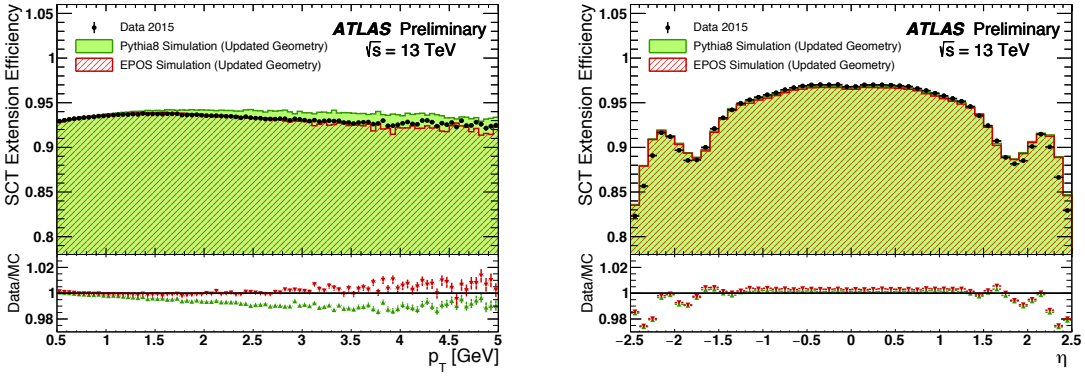


FIGURE 4.19: Difference of $\mathcal{E}_{\text{ext}}$ between nominal and pixel service +10%, pixel service +25%, and pixel service +50% geometry samples with respect to the pixel-tracklet parameters p_{T}, η, ϕ.



(a) (b)

FIGURE 4.20: SCT extension efficiency as a function of (a) p_{T} and (b) η of the tracklets in a comparison among data, PYTHIA 8 and EPOS. Ref. [112].

The three Pythia 8 A2 samples (pixel service +10%, pixel service +25% and pixel service +50%) were also compared to data, as shown in Figures 4.21, 4.22 and 4.23 in order to check if any of them gave a better description of the data.

Unfortunately, none of these simulated samples describes the data perfectly, thus to estimate the missing amount of material in simulation, a linear interpolation between $\mathcal{E}_{\mathrm{ext}}$ as calculated in the different samples has to be performed, as discussed in Section 4.5.9.

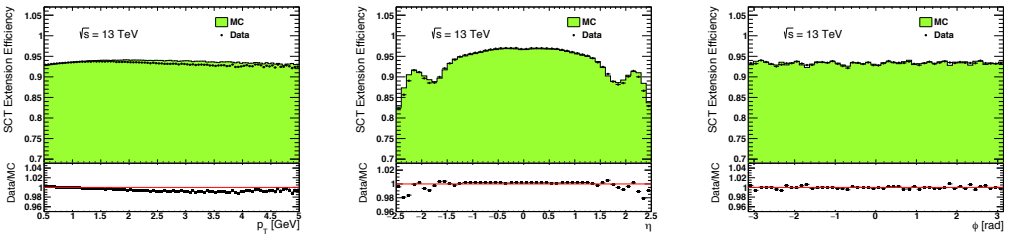


FIGURE 4.21: $\mathcal{E}_{\mathrm{ext}}$ with respect to the pixel-tracklet parameters p_{T}, η, ϕ for data and the pixel service +10% MC.

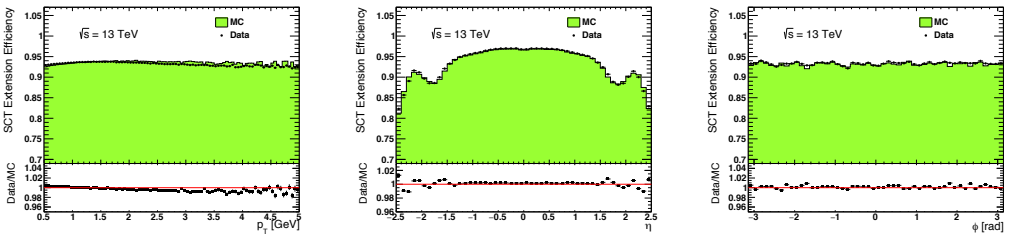


FIGURE 4.22: $\mathcal{E}_{\mathrm{ext}}$ with respect to the pixel-tracklet parameters p_{T}, η, ϕ for data and the pixel service +25% MC.

4.5.6 Linearity of the Track-Extension efficiency

The linearity of the change in the track-extension efficiency as a function of the material added in the Pixel services was checked by performing a linear interpolation of the values obtained for each η bin when using a set of MC simulation samples in which the material

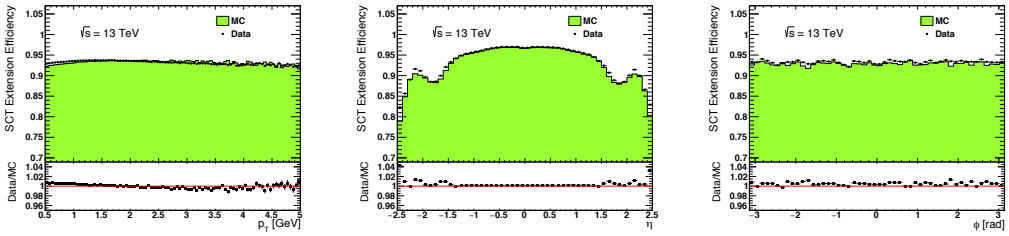


FIGURE 4.23: $\mathcal{E}_{\text{ext}}$ with respect to the pixel-tracklet parameters p_{T}, η, ϕ for data and the pixel service +50% MC.

density in the pixel service was scaled by a factor (w), where $w = 0,\ 10\%,\ 25\%,\ 50\%$ (see Table 4.1). $\mathcal{E}_{\text{ext}}^{\text{MC}}(w)$ is then calculated for each distorted sample, and the slope $\mathrm{d}\mathcal{E}_{\text{ext}}^{\text{MC}}/\mathrm{d}w \equiv \alpha^{\text{MC}}$ is obtained from a linear fit for each η bin, as shown in Figure 4.24. Compared to Equation 4.6, K is estimated as:

$$K = -\frac{1}{\alpha^{\text{MC}}} \cdot N_{\lambda_I}^{\text{MC}}(w=0)\ , \tag{4.7}$$

where $N_{\lambda_I}^{\text{MC}}(w=0)$ is the amount of material associated with nuclear interactions within the pixel service region corresponding to the η bin without weighting. Therefore:

$$\Delta N_{\lambda_I}^{\text{data-MC}} = \frac{N_{\lambda_I}^{\text{MC}}(w=0)}{\alpha^{\text{MC}}} \cdot \Delta\mathcal{E}_{\text{ext}}^{\text{data-MC}}\ . \tag{4.8}$$

From the linear interpolation, the percentage of material supposed to be missing in simulation was extracted. A closure test of the procedure was also performed to ensure that the calculated amount of material was accurate. This was performed by using the sample where the material in the pixel services was scaled by 25%. This sample was treated like data. First it was removed from all fits used to determine the relation between the amount of material and changes in $\mathcal{E}_{\text{ext}}$. Secondly the measured change in $\mathcal{E}_{\text{ext}}$ was calculated and converted to an effective amount of missing material. Figure 4.25 shows the accuracy of the estimate of the amount of missing material. No additional systematic is added for non-closure of the method as all measurements are within their statistical uncertainty.

4.5.7 Systematic Uncertainties

The following effects which affect directly the $\mathcal{E}_{\text{ext}}$ are considered:

- Modelling of particle interactions - GEANT 4 Physics Lists
- p_{T}-dependence and particle composition of $\mathcal{E}_{\text{ext}}$ measurement
- Particle kinematics
- Pixel-tracklet fake rate
- Gradient of the material scaling in the two generators
- Number of Hits used in the matching between track and tracklet
- SCT detector efficiency and dead modules
- Pixel detector efficiency and dead modules

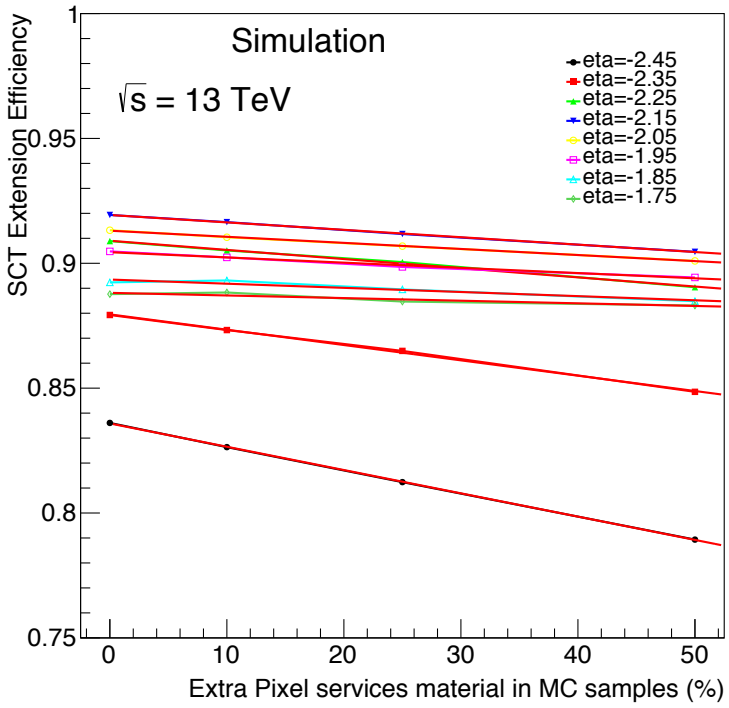


FIGURE 4.24: The relationship between the track-extension efficiency and the amount of material in the pixel service region for a subset of η regions. For each η region a linear fit is applied.

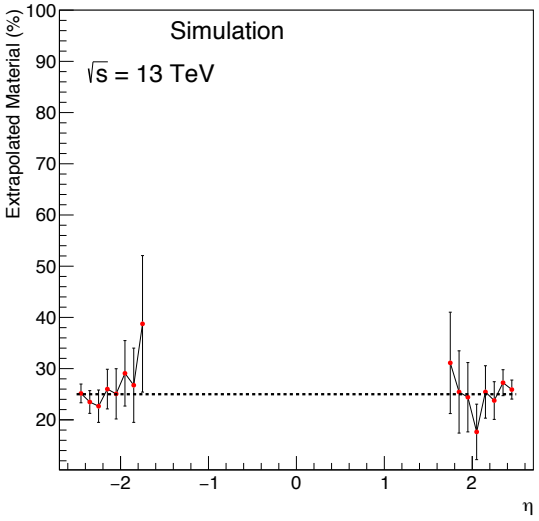


FIGURE 4.25: Results of the closure test. All measurement are within their statistical uncertainty so no additional systematic is added for non-closure of the method.

In addition, to convert the measurement from the $\mathcal{E}_{\text{ext}}$ into an evaluation of the missing material in nuclear interaction lengths, the method needs to be calibrated and this also introduces an additional source of systematic uncertainty on the final measurements which will be discussed in the following.

As summarised below, some of these effects give sizable results while others are negligible.

Particle composition As the tracklets used in the measurement of $\mathcal{E}_{\text{ext}}$ originate from a variety of particles, the final $\mathcal{E}_{\text{ext}}$ is actually the weighted sum of the $\mathcal{E}_{\text{ext}}$ for all particle types:

$$\mathcal{E}_{\text{ext}} = \sum_{i=\text{species}} f_i \mathcal{E}^i_{\text{ext}} \,, \tag{4.9}$$

where f_i is the fraction of reconstructed tracklets associated to a particular particle species and $\mathcal{E}^i_{\text{ext}}$ is its associated track-extension efficiency.

The probability that a particle interacts with matter, and hence its $\mathcal{E}_{\text{ext}}$, depends on the species of the particle. In addition, in-flight decays of short-lived charged hadrons, e.g. weakly-decaying strange baryons, represent a source of irreducible background to the $\mathcal{E}_{\text{ext}}$ measurement as they exhibit an experimental signature identical to stable particles interacting with matter. This induces a dependence of the $\mathcal{E}_{\text{ext}}$ measurement on the particle composition in the simulation.

Considerable differences in the predicted rate of particles in various event generators are observed. For example, the predicted cross section of weakly-decaying strange baryons in EPOS is twice bigger than that of PYTHIA 8 [80]. As such, the impact of the particle composition on the measurement of $\mathcal{E}_{\text{ext}}$ needs to be evaluated in data or estimated from simulation.

Figures 4.26 and 4.27 show $\mathcal{E}_{\text{ext}}$ for various particles and generators. There is a large particle flavour dependent variation in the track-extension efficiency especially for what is labelled as "other". Other particles (which represent less than 0.5% of the particles analysed, as will be shown later) include short-lived particles such as strange baryons which have a significantly lower efficiency than all other categories. There is a large variation in $\mathcal{E}_{\text{ext}}$ for particles classified as "other" between EPOS and PYTHIA 8 which indicates that the composition of this category is significantly different between the two. This further indicates that the understanding of the composition of particles which create the tracklets is poorly modelled and is likely to be the cause of the variation of the material estimate as a function of the tracklets p_{T}. Some other detailed pictures can be found in Appendix B.

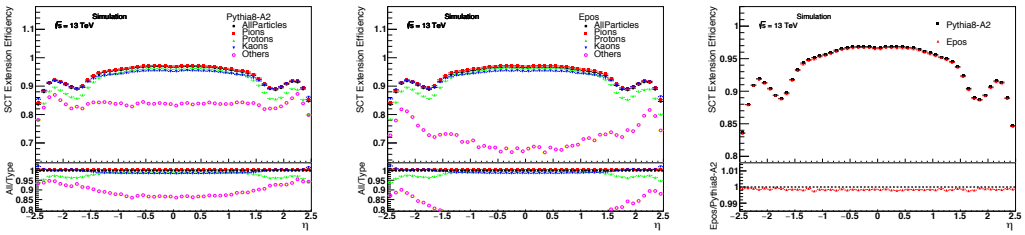


FIGURE 4.26: $\mathcal{E}_{\text{ext}}$ with respect to η of the tracklets in a comparison among different particle selections generated by PYTHIA 8-A2 (left) and EPOS (middle) and in a comparison between the two generators without any particle cut applied (right).

Both the particle composition and the material interaction probability vary as a function of the hadron p_{T} and η. If there were a perfect description of the particle composition

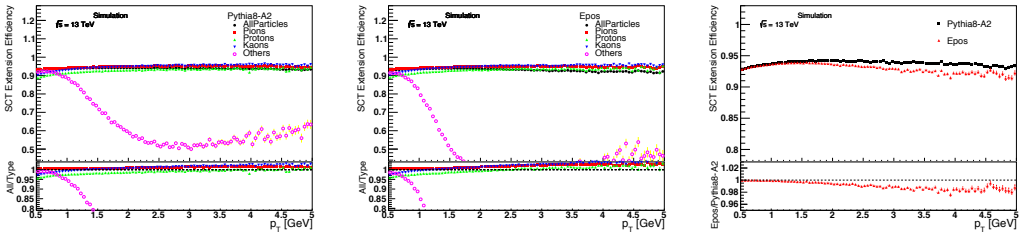


FIGURE 4.27: $\mathcal{E}_{\text{ext}}$ with respect to p_{T} of the tracklets in a comparison among different particle selections generated by PYTHIA 8-A2 (left) and EPOS (middle) and in a comparison between the two generators without any particle cut applied (right).

and material interactions in the simulation, no difference in the relative change of the $\mathcal{E}_{\text{ext}}$ between data and simulation would be observed when varying the p_{T} range used to perform the measurement. However, as just shown, notable deficiencies in the modelling are present in all MC generators, such as the fraction of short-lived charged hadrons as a function of p_{T} and thus the distribution of their decay length. To assess the effect of these discrepancies on the measurement, the relative change in $\mathcal{E}_{\text{ext}}$ evaluated in data with respect to simulation (PYTHIA 8) is measured in four regions of p_{T} within the range $0.5\ \text{GeV} < p_{\text{T}} < 5\ \text{GeV}$. The maximum deviation from the inclusive value is used to estimate a systematic uncertainty of 0.33%. Furthermore, the difference in $\mathcal{E}_{\text{ext}}(\eta)$ integrated over $|\eta| < 2.5$ between the EPOS and the PYTHIA 8 samples (0.15%), shown in Figure 4.20, is treated as a systematic uncertainty. To encompass both of these effects, broken down in Table 4.6 a systematic uncertainty of 0.36%, the sum in quadrature of 0.33% and 0.15%, is applied in each η bin.

Lower p_{T} bins (< 500 MeV) were also initially investigated, but in these regions, as the momentum is reduced, the rate of secondary tracks (that is the ratio between the number of tracks associated with a secondary particle and the total number of tracks entering the selection) increases (see Figure 4.29). This motivated the restriction of the track-extension efficiency study in a well defined phase space between 0.5 and 5 GeV since in this region the fraction of secondaries is relatively stable and the effect on the final measurement can be considered negligible.

Fake tracklets Fake tracklets are another source of bias to the measurement of $\mathcal{E}_{\text{ext}}$. In all simulated samples, the fraction of fake tracklets is less than 0.3%. To estimate the uncertainty on $\mathcal{E}_{\text{ext}}$ associated with fake tracklets, a variation of $\pm 50\%$ of the fake tracklet rate is considered, as described in Section 3.5.3 in the context of the tracking recommendations, and the corresponding variation of $\mathcal{E}_{\text{ext}}(\eta)$ is assigned as an uncertainty for each η bin. Figure 4.30 shows the rate of fake pixel-tracklets as a function of η in the Pythia 8 A2 sample based on the *original* geometry. No sizable differences are seen with respect to other samples.

Other sources Many other aspects which may potentially contribute to the uncertainty were investigated and found to be negligible ($\ll 0.01\%$). These include the differences in the GEANT4 physics lists, hit efficiency in the pixel and SCT detectors, residual misalignment between the pixel and SCT detectors, gradient of the material scaling in the two generators. Furthermore, adjusting the requirement on the minimum number of shared hits between the tracklet and selected combined track was found to have a negligible effect on the final measurement.

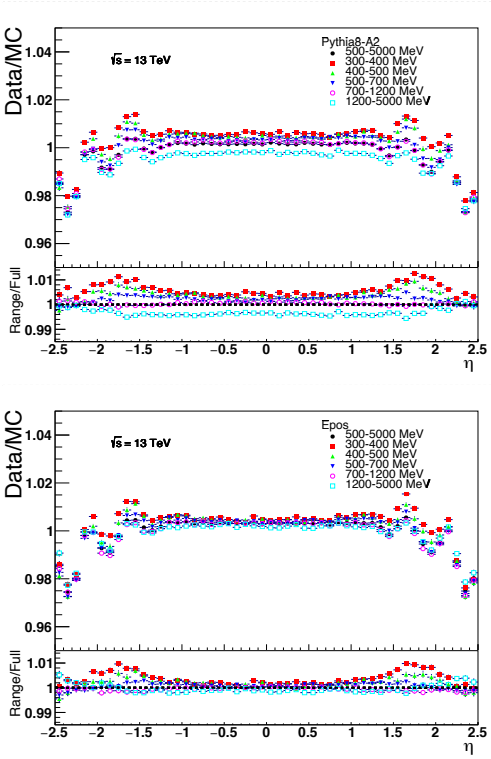


FIGURE 4.28: p_{T}-dependence of the ratio of $\mathcal{E}_{\mathrm{ext}}$ in data and simulation for PYTHIA 8 (left) and EPOS (right).

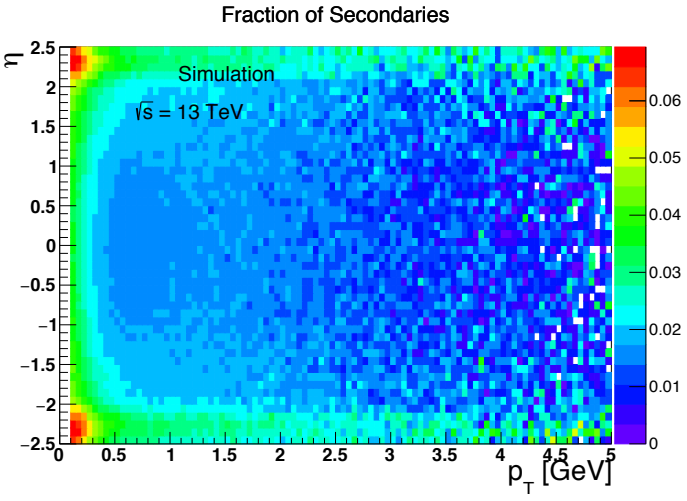


FIGURE 4.29: The fraction of secondary tracks entering the full track selection in Monte Carlo Simulation as a function of η and p_{T}.

η	$\mathcal{E}_{\text{ext}}$		p_{T}-dependence Uncertainty	
± 0.05	(Data - MC) PYTHIA 8 [%]	(Data - MC) EPOS [%]	PYTHIA 8 [%]	EPOS [%]
-2.45	-1.20	-1.20	-0.02	-0.33
-2.35	-2.37	-2.23	-0.01	-0.21
-2.25	-1.82	-1.79	0.027	-0.14
-2.15	-0.27	-0.17	0.10	-0.11
-2.05	-0.14	-0.05	0.16	-0.08
-1.95	-0.75	-0.65	0.20	-0.05
-1.85	-0.81	-0.81	0.24	-0.04
-1.75	-0.26	-0.22	0.33	0.03
-1.65	0.34	0.39	0.39	0.13
-1.55	0.34	0.41	0.39	0.20
-1.45	-0.03	0.13	0.35	0.20
-1.35	-0.17	0.03	0.33	0.16
-1.25	-0.00	0.17	0.33	0.15
-1.15	0.12	0.37	0.33	0.15
-1.05	0.19	0.32	0.33	0.15
-0.95	0.16	0.37	0.33	0.13
-0.85	0.12	0.30	0.32	0.11
-0.75	0.18	0.33	0.30	0.09
-0.65	0.15	0.27	0.30	0.09
-0.55	0.14	0.35	0.29	0.09
-0.45	0.14	0.31	0.26	0.09
-0.35	0.15	0.30	0.23	0.09
-0.25	0.14	0.30	0.24	0.09
-0.15	0.16	0.32	0.25	0.10
-0.05	0.13	0.26	0.25	0.10
0.05	0.21	0.30	0.25	0.10
0.15	0.17	0.28	0.25	0.11
0.25	0.14	0.30	0.26	0.13
0.35	0.14	0.33	0.26	0.13
0.45	0.20	0.30	0.28	0.09
0.55	0.16	0.29	0.30	0.08
0.65	0.13	0.31	0.32	0.08
0.75	0.11	0.25	0.32	0.07
0.85	0.17	0.35	0.33	0.09
0.95	0.20	0.32	0.34	0.12
1.05	0.17	0.32	0.33	0.14
1.15	0.16	0.31	0.33	0.15
1.25	0.12	0.31	0.33	0.15
1.35	-0.04	0.15	0.32	0.14
1.45	-0.30	-0.13	0.31	0.13
1.55	0.02	0.18	0.30	0.12
1.65	0.31	0.50	0.28	0.10
1.75	-0.11	0.02	0.27	0.05
1.85	-0.63	-0.49	0.23	0.00
1.95	-0.90	-0.79	0.15	-0.04
2.05	-0.54	-0.43	0.12	-0.09
2.15	-0.11	0.03	0.07	-0.14
2.25	-1.35	-1.20	-0.02	-0.19
2.35	-2.37	-2.23	-0.06	-0.21
2.45	-1.84	-1.67	-0.06	-0.20

TABLE 4.6: The measured $\mathcal{E}_{\text{ext}}$ in different regions of the detector compared to different event generators. The p_{T}-dependent uncertainties are evaluated from half of the maximum difference in the $\mathcal{E}_{\text{ext}}$ measurement in each of the p_{T} bins.

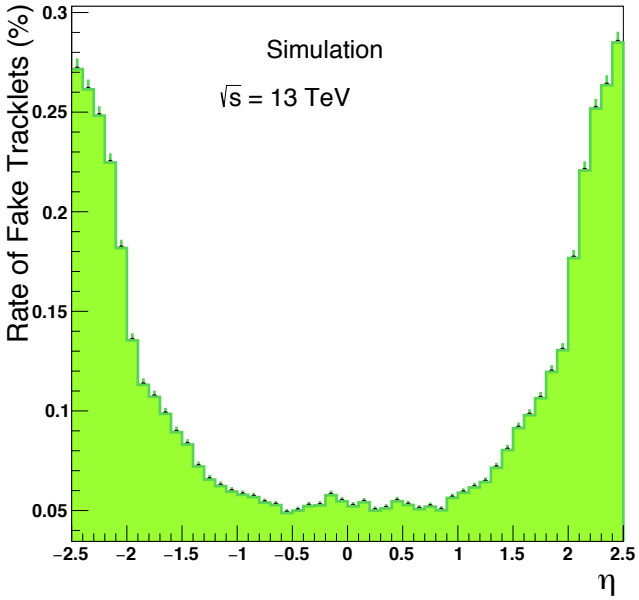
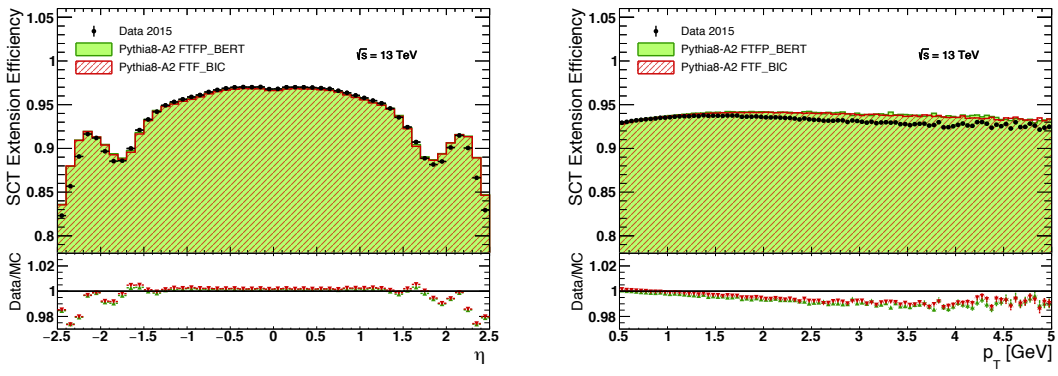


FIGURE 4.30: Rate of fake pixel-tracklets as a function of the pseudorapidity η in `Py8(Orig_Alt)`.

Particle kinematics After correcting the p_{T} and η spectra of tracks in simulation to match the ones in data, there is a negligible residual impact on the measured SCT-extension efficiency ($< 0.05\%$) so no further systematics are included.

Modelling of particle interactions The impact of the differences in the physics process description in the different physics lists on the track-extension efficiency has been checked and was found to be negligible, as shown in Figure 4.31. Thus, no systematics uncertainties are taken into account for the usage of a specific physics list.



(a) (b)

FIGURE 4.31: SCT-extension efficiency as function of (a) pseudorapidity, η, and (b) transverse momentum, p_T, in a comparison between PYTHIA 8 based on `FTFP_BERT`, PYTHIA 8 based on `FIT_BIC` and data.

Gradient of the material scaling in the two generators It has been supposed that $\mathcal{E}_{\mathrm{ext}}$ scales linearly with material and the missing material in simulation has been extrapolated by using the PYTHIA 8 samples. It is thus necessary to check that the scaling of $\mathcal{E}_{\mathrm{ext}}$ with increasing material is the same in PYTHIA 8 and EPOS, otherwise a systematic uncertainty

on the gradient needs to be considered. By comparing the difference in $\mathcal{E}_{\text{ext}}$ between the Pythia 8 nominal and pixel service +50% samples with the difference between the EPOS nominal and pixel service +50% samples, it can be seen in Figure 4.32 that the effect of the gradient is totally negligible, thus no systematics uncertainty has been added.

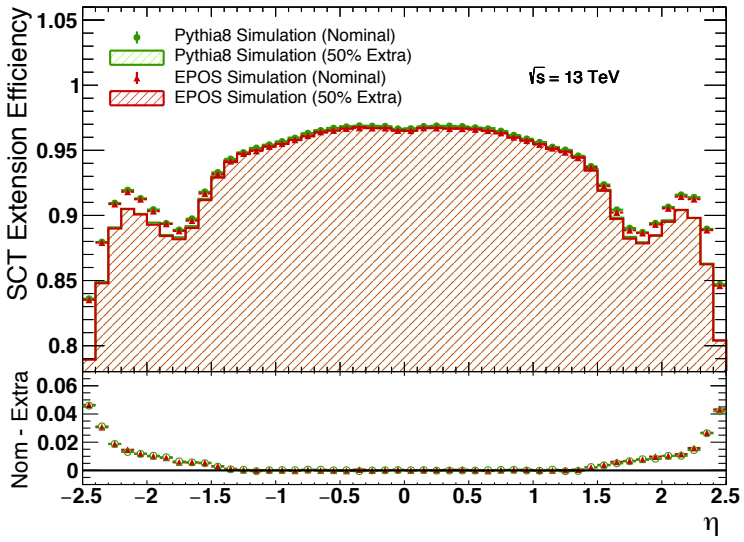


FIGURE 4.32: $\mathcal{E}_{\text{ext}}$ with respect to η in `Py8(Orig_Alt)`, `Py8(Orig_PP0+50%)`, `EPOS(Orig)` and `EPOS(Orig_PP0+50%)`.

Number of Hits used in the matching between track and tracklet The matching between the track and the pixel-tracklet is considered valid when they share at least one common hit. The impact of changing the number of hits has been investigated and it is found to be fully negligible. As it can be seen in Figure 4.33(a) and 4.33(b), there is a decrease in the SCT-Extension efficiency of the order of $\sim 0.5\%$ when moving from 1 to 2 hits and of $\sim 1\%$ when moving from 1 to 3 hits, but this applies to both data and simulation, so the impact on the difference is fully negligible within the statistical uncertainties. In addition, a comparison between the Pythia 8 nominal and pixel service +50% samples is shown in Figure 4.33(c) and 4.33(d) to ensure also that the scaling is identical when using different number of hits for the matching.

SCT hit efficiency The SCT hit efficiency was measured in data to be 99.9% (a similar value to that used in the simulation) and so it is not considered to be a systematic.

Masked SCT modules If an SCT module located just after the pixel detector is disabled for whatever reason, the $\mathcal{E}_{\text{ext}}$ measurement would probe more material and thereby would have a lower value. SCT conditions were not identical in data and simulation, however only 11 of the 4088 SCT modules differed between data and MC (2 modules are off in MC but on in data; 9 modules are off in data but on in MC (see Table 4.7 for details)).

If a module was disabled, all tracks passing through that module would traverse approximately 0.05 X_0 of additional passive material or at most a 1% change in the track-extension efficiency. Only one of the modules listed in Table 4.7 is likely to be the first SCT module struck by a track originating from the IP. As there are over 40 modules in the innermost layer of the SCT in each η region, the impact on $\mathcal{E}_{\text{ext}}$ would be very small ($\simeq 0.025\%$) and hence is ignored.

Masked pixel modules In the simulation, conditions in the pixel detector are nearly identical to those in the data. The only known difference occurs in the IBL, where one module has one inactive chip. The region that this chip covers is $1 < \eta < 2$. $\mathcal{E}_{\text{ext}}$ is nearly identical for both the positive and negative η regions and hence this module does not impact the measurement and is not considered (see Table 4.6). It should be noted

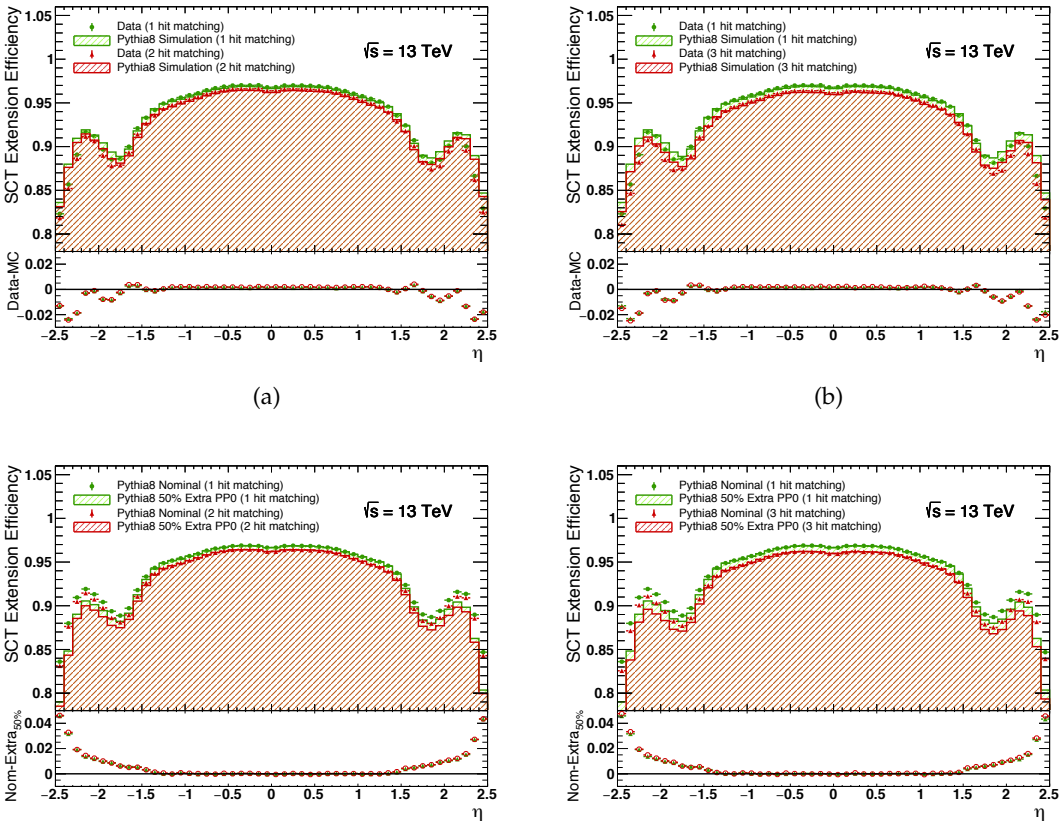


FIGURE 4.33: SCT-extension efficiency as function of (a) pseudorapidity, in a comparison between 4.33(a) PYTHIA 8 and data and 4.33(c) when requiring 1, 2 or 3 hits in common between the track and the pixel-tracklet.

Status	Approximate η	Approximate ϕ	BEC	Layer	Ring
Dead in MC	-1.7	-2.9	ECC	3	0
Dead in MC	0.5	-0.5	Barrel	1	2
Dead in data	-1.8	-0.2	ECC	2	1
Dead in data	-2.2	-0.9	ECC	4	2
Dead in data	-2.2	0.9	ECC	5	1
Dead in data	1.4	0.5	ECA	1	0
Dead in data	1.8	-0.9	ECA	4	0
Inefficient in data	2.5	-2.8	ECA	7	1
Inefficient in data	1.8	-1.2	ECA	2	1
Inefficient in data	-2.3	-2.0	ECC	7	0
Inefficient in data	1.5	-2.6	ECA	0	1

TABLE 4.7: List of dead or inefficient SCT modules that differ between data and simulation conditions.

that changes in pixel conditions would only have a second-order effect on the efficiency measurement as any change will alter both numerator and denominator in the efficiency measurement in the same way.

4.5.8 Track-Extension Efficiency Preliminary Results

η ± 0.05	$\mathcal{E}_{\text{ext}}$ (Data - MC) [%]	Measurement Uncertainties Statistical [%]	p_{T}-dependence [%]	Fake Rate [%]	Extrapolated Material [%]
-2.45	-1.204	0.035	0.36	0.136	13.61
-2.35	-2.368	0.031	0.36	0.131	37.78
-2.25	-1.823	0.027	0.36	0.124	51.05
-2.15	-0.270	0.024	0.36	0.112	10.06
-2.05	-0.138	0.025	0.36	0.091	4.95
-1.95	-0.750	0.027	0.36	0.068	37.64
-1.85	-0.809	0.029	0.36	0.057	49.69
-1.75	-0.263	0.028	0.36	0.054	25.71
-1.65	0.335	0.026	0.36	0.049	-30.40
-1.55	0.343	0.023	0.36	0.045	-32.65
1.55	0.020	0.023	0.36	0.046	-4.43
1.65	0.310	0.025	0.36	0.049	-28.28
1.75	-0.109	0.028	0.36	0.053	10.33
1.85	-0.632	0.029	0.36	0.060	37.33
1.95	-0.898	0.029	0.36	0.065	47.04
2.05	-0.540	0.027	0.36	0.088	28.17
2.15	-0.106	0.025	0.36	0.110	4.30
2.25	-1.349	0.026	0.36	0.126	42.26
2.35	-2.371	0.030	0.36	0.132	42.79
2.45	-1.843	0.034	0.36	0.142	20.11

TABLE 4.8: The difference between the measured $\mathcal{E}_{\text{ext}}$ in different regions of the detector in data and simulation and the different sources of uncertainty.

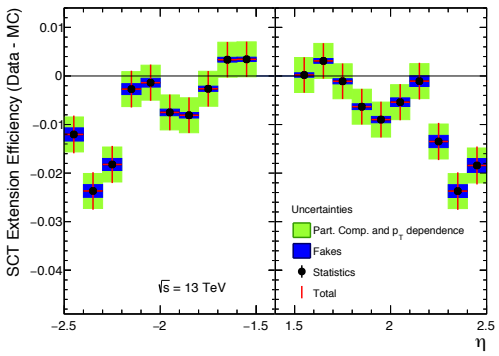


FIGURE 4.34: The difference between the $\mathcal{E}_{\text{ext}}$ measured in data and in simulation is shown together with the uncertainties. The total uncertainty includes the uncertainty from fake pixel tracklets, the uncertainty from the particle composition and p_{T}-dependence and the statistical uncertainty combined in quadrature.

The results together with the related uncertainties are summarised in Table 4.8 and shown in Figure 4.34. The discrepancy between data and simulation reaches its maximum in the range $2.3 < |\eta| < 2.4$. It is clear that the main source of uncertainty is the systematic one (dominated by the p_{T} dependence), while the statistical uncertainty remains very small in the full pseudorapidity range.

4.5.9 Material Estimation

As already stated in Equation 4.6, the difference in $\mathcal{E}_{\text{ext}}(\eta)$ between data and simulation, $\Delta\mathcal{E}_{\text{ext}}^{\text{Data}-\text{MC}}(\eta)$, can be obtained from:

$$\Delta N_{\lambda_I}^{\text{Data}-\text{MC}}(\eta) \simeq -K(\eta) \cdot \Delta\mathcal{E}_{\text{ext}}^{\text{Data}-\text{MC}}(\eta)\,,$$

To convert the track-extension efficiency to estimated nuclear interaction lengths, it is necessary to establish the appropriate value of $K(\eta)$ to calculate the difference in material
.

The fraction of particles that remain on average after traveling through material is given by:

$$f(N_{\lambda_I}) = e^{-N_{\lambda_I}}. \tag{4.10}$$

Assuming the only loss is from interactions with the material and considering all the material located between the pixel and SCT detectors, then:

$$f(N_{\lambda_I}) = \mathcal{E}_{\text{ext}}. \tag{4.11}$$

For $N_{\lambda_I} \ll 1$, Eq. 4.10 corresponds to:

$$f(N_{\lambda_I}) = 1 - N_{\lambda_I} + \mathcal{O}(N_{\lambda_I}^2). \tag{4.12}$$

In presence of any additional passive material (ΔN_{λ_I}) and as a consequence of Eqs. 4.10 and 4.12, the difference in the fraction of particles after passing through that material can be expressed as:

$$f(N_{\lambda_{I,0}} + \Delta N_{\lambda_I}) - f(N_{\lambda_{I,0}}) = f(N_{\lambda_{I,0}})f(\Delta N_{\lambda_I}) - f(N_{\lambda_{I,0}}) = \tag{4.13}$$

$$= f(N_{\lambda_{I,0}})[f(\Delta N_{\lambda_I}) - 1] \approx f(N_{\lambda_{I,0}})(-\Delta N_{\lambda_I}).$$

where $N_{\lambda_{I,0}}$ is the nominal material. Summarising:

$$\frac{f(N_{\lambda_{I,0}} + \Delta N_{\lambda_I}) - f(N_{\lambda_{I,0}})}{f(N_{\lambda_{I,0}})} \approx -\Delta N_{\lambda_I},$$

which, as assumed in Eq. 4.11, for material located between the pixel and SCT detectors corresponds to:

$$\frac{\mathcal{E}_{\text{ext}}^{\text{mod}} - \mathcal{E}_{\text{ext}}^{\text{nom}}}{\mathcal{E}_{\text{ext}}^{\text{nom}}} \approx -\Delta N_{\lambda_I}^{\text{mod}-\text{nom}}. \tag{4.14}$$

This means that the normalised difference in track-extension efficiency in the nominal (*nom*) geometry and in the modified (*mod*) geometry (left hand side of Eq. 4.14), $\Delta\mathcal{E}_{\text{norm ext}}^{\text{mod}-\text{nom}}$, is approximately equal to the change in material $-\Delta N_{\lambda_I}^{\text{mod}-\text{nom}}$ and $K(\eta) \approx 1/\mathcal{E}_{\text{ext}}^{\text{nom}}$:

$$\Delta\mathcal{E}_{\text{norm ext}}^{\text{mod}-\text{nom}} \approx -\Delta N_{\lambda_I}^{\text{mod}-\text{nom}}. \tag{4.15}$$

To verify the value of $K(\eta)$, a set of MC simulation samples in which the material in the pixel service region is modified was created (see Table 4.1). In particular, the samples based on the distorted geometries pixel service +50%, ring layout 1 and ring layout 2 where used. This allowed to check differences in the track-extension efficiency when the material is uniformly added in a wide region, as in the case of the pixel services, or it is well localised in r-z as in the geometries where only a ring of material was added. Any deviation of $K(\eta)$ from the expected dependence needs to be taken as a systematic uncertainty.

4.5.10 z_{vtx} cut and tighter SCT Hits cut

The η value at which material is sampled is a function of the z-position of the primary vertex (z_{vtx}). A more refined information can thus be obtained by applying a tight cut:

- $|z_{\text{vtx}}| < 10$ mm.

This reduces vertex smearing which, in turn, introduces fuzziness in the location of extra material.

This cut was not originally used in order to propagate consistently the track-extension efficiency results to the data-driven correction applied to the Charged Particle Multiplicity measurement [80], because in that analysis no cut was imposed to the z_{vtx} distribution. In the context of the Run 2 ID Material Studies, this method was improved and the methodology to convert the track-extension efficiency to nuclear interaction lengths slightly changed.

As explained in Section 4.5.9, the relationship between the track-extension efficiency and the amount of material is linear. The same procedure shown in Figure 4.24 was repeated for tracklets passing the new z_{vtx} requirement to ensure that the linearity was still guaranteed and the approximation made in Eq. 4.12 was still valid, as shown in Figure 4.35.

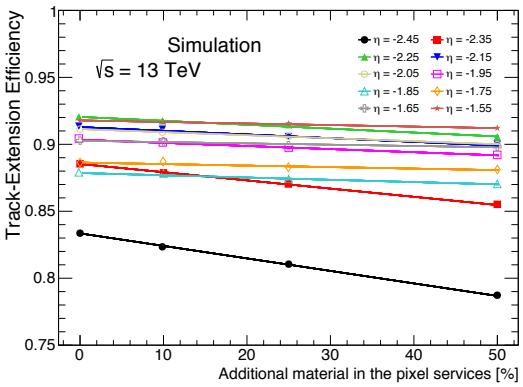


FIGURE 4.35: The relationship between $\mathcal{E}_{\text{ext}}$ (η) and the amount of material in the pixel services. $\mathcal{E}_{\text{ext}}$ (η) was measured using tracklets in $0.5 < p_{\text{T}} < 5$ GeVwith $|z_{\text{vtx}}| < 10$ mm. For each η region, a linear fit is applied.

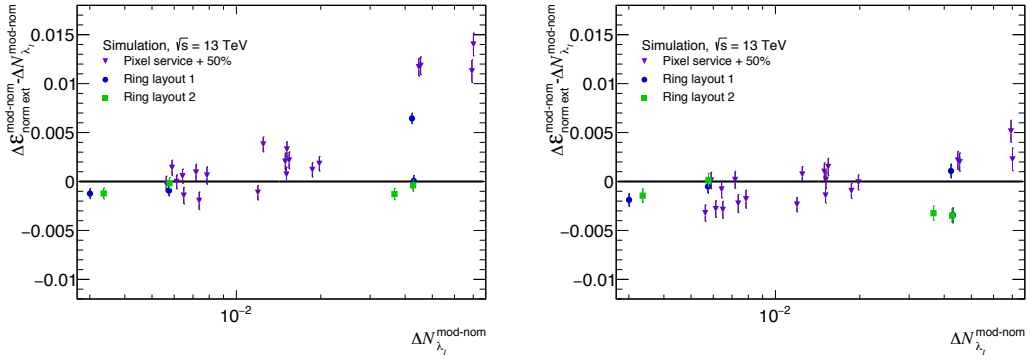
In addition, to reduce the possibility of randomly associatings 2 SCT hits to the combined track (no requirement is made that the 2 SCT hits have to be on different layers or disks) and thus, match it with a tracklet with the result of an artificially increased track-extension efficiency, the number of SCT hits was increased from 2 to 4. The number of SCT hits cannot be increased further because of the risk to probe also material in the region of the SCT detector, thus exposing the method to SCT module inefficiencies. As shown if Figures 4.36, the closure test is affected by much smaller deviations from the expected value when requiring at least 4 SCT hits instead of 2. Thus, the tighter requirement improves the calibration of the method and reduces its systematic uncertainties.

A comparison between data and simulation after including both the cut on the z_{vtx} and the tigher cut on the number of SCT hits is shown in Figure 4.37.

All the relevant systematics have been re-evaluated accordingly.

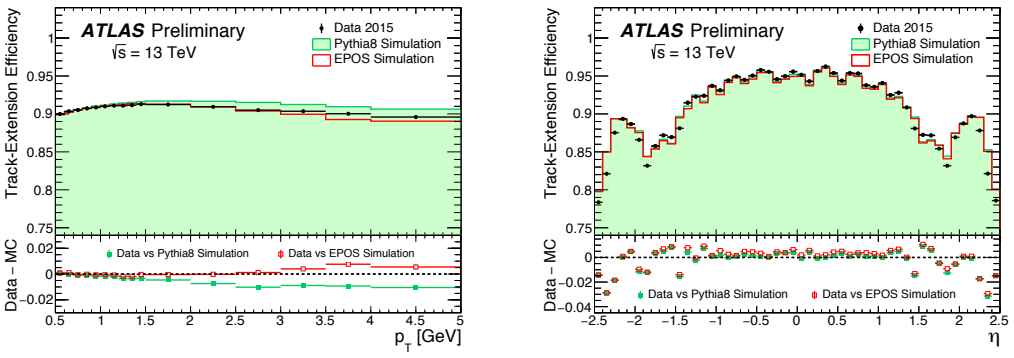
The p_{T}-dependence and particle composition of $\mathcal{E}_{\text{ext}}$ measurement described in 4.5.7 is shown in Figure 4.38.

The values of the systematic uncertainty accounting for p_{T}-dependence and particle composition evaluated after the z_{vtx} cut and by requiring 4 SCT hits on the matched track



(a) (b)

FIGURE 4.36: The difference between the normalised track-extension efficiency, $\Delta\mathcal{E}^{\text{mod}-\text{nom}}_{\text{norm ext}}$, and the nuclear interaction lengths in the nominal geometry and in the modified geometries, as a function of the difference of nuclear interaction length between the same samples, $\Delta N^{\text{mod}-\text{nom}}_{\lambda_I}$. A modified geometry results in different values of $\Delta N^{\text{mod}-\text{nom}}_{\lambda_I}$ over several η bins. Figure (a) for the 2 SCT hits and (b) for the 4 SCT hits requirement. The method to estimate $\Delta N^{\text{mod}-\text{nom}}_{\lambda_I}$ will be detailed in Section 4.5.11.



(a) (b)

FIGURE 4.37: Track-extension efficiency as a function of (a) p_{T} and (b) η of the tracklets in a comparison between data, PYTHIA 8 and EPOS. Ref. [113].

are bigger than the ones shown in Table 4.6, so to be conservative a systematic uncertainty of 0.50% (instead of 0.33% previously used) is applied to all the pseudorapidity bins. In addition, to fully cover the differences between PYTHIA 8 and EPOS in the barrel region, as shown in Figure 4.37, a systematic uncertainty of 0.21% (instead of the 0.20% previously used) is applied to all the pseudorapidity bins. This leads to a combined uncertainty of 0.54% instead of 0.36% previously used.

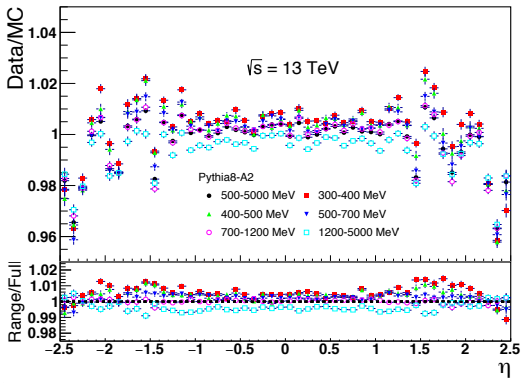


FIGURE 4.38: p_{T}-dependence of the ratio of $\mathcal{E}_{\mathrm{ext}}$ in data and simulation for PYTHIA 8.

The difference of the $\mathcal{E}_{\mathrm{ext}}$ between the data and the PYTHIA 8 simulation sample as a function of η, $\Delta\mathcal{E}_{\mathrm{ext}}^{\mathrm{Data-MC}}(\eta)$, together with the associated statistical and systematic uncertainties is shown in Figure 4.39. The optimised selection criteria allowed also to expand the region probed by this method from $1.5 < |\eta| < 2.5$ to $|\eta| < 2.5$.

4.5.11 Calculation of the material in the geometry model

The difference of the $\mathcal{E}_{\mathrm{ext}}$ between the data and the PYTHIA 8 simulation is translated to an amount of material using Equation 4.8.

The amount of material in the pixel service region associated with nuclear interactions for each η bin, $N_{\lambda_I}(\eta)$, is established from the difference between the *original* geometry model and the modified geometry models, namely the sample with the density of the material in the pixel service region scaled by 50% (pixel service + 50%) and the two samples with rings of passive material added between the pixel and the SCT detector (ring layout 1 and 2). The ring layout 3 geometry model was created with a procedure analogues to the one used for ring layout 1 and 2 in order to check the validity of the systematics envelopes, which resulted to cover conservatively all the variations and thus correctly evaluated.

For each geometry model, N_{λ_I} is calculated by integrating the material along the path of a virtual neutral non-interacting particle, known to as *geantino* in the GEANT4 simulation. The difference between the two geometry models is taken to be contribution of the weighted material in the pixel service region.

For the geantino-based calibration to be accurate, the path travelled by the geantinos must match that of the particles used in the measurement. For this reason, the distribution of geantino production verices in z should match the distribution of z_{vtx} observed in data. Thus, the geantino vertex z-distribution is re-weighted to match the z_{vtx}-distribution observed in the data.

In addition, the track-extension efficiency is affected by the radial position and orientation of material in the detector. The dependence on the location and orientation of the material can be simply explained: if the missing material is located nearer to the first layer of the SCT there is a higher probability that one of the secondary particles arising from

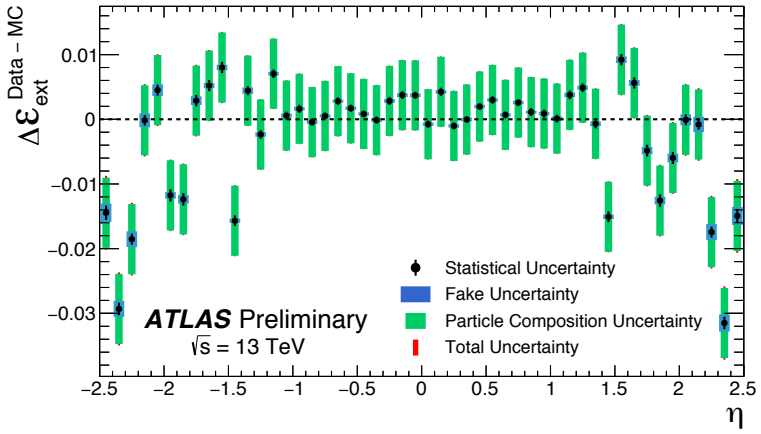


FIGURE 4.39: The difference between the track-extension efficiency measured in data and in simulation, $\Delta\mathcal{E}_{\text{ext}}^{\text{Data}-\text{MC}}(\eta)$, is shown together with the uncertainties. The total uncertainty includes the uncertainty from fake tracklets, the uncertainty from the particle composition and p_{T}-dependence and the statistical uncertainty are all summed in quadrature. Ref. [113].

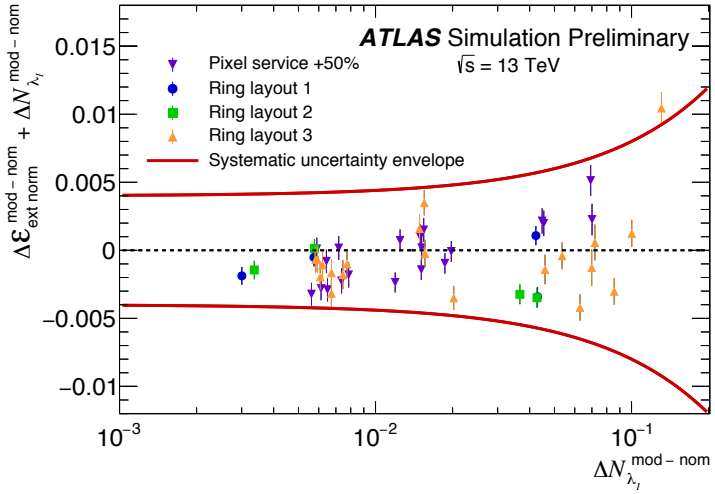


FIGURE 4.40: The sum of $\Delta\mathcal{E}_{\text{norm ext}}^{\text{mod}-\text{nom}}$ (normalised track-extension efficiency in the nominal geometry and in the modified geometries) and $\Delta N_{\lambda_I}^{\text{mod}-\text{nom}}$ (number of nuclear interaction lengths in the nominal geometry and in the modified geometries). A modified geometry results in different values of $\Delta N_{\lambda_I}^{\text{mod}-\text{nom}}$ over several η bins. The solid red lines are envelopes described by the curves $\pm(0.004 + 0.04 \cdot \Delta N_{\lambda_I}^{\text{mod}-\text{nom}})$. Ref. [113].

hadronic interactions produces a hit in the SCT compatible with the tracklet and thus considered as an extension. This artificial increase of the track-extension efficiency is highly suppressed by requiring four SCT hits (which correspond to hits on at least two layers of the SCT detector) on the combined track. As described in Section 4.3.2, a series of detector geometries were created in which an additional ring of passive material was added to the detector at different radii covering $2.2 < |\eta| < 2.3$. Simulated samples were created based on these geometries and the track-extension efficiency per unit of material was calculated. The variation in this quantity is shown in Figure 4.40. This variation is constrained by an envelope described by the equation $\pm(0.004 + 0.04 \cdot \Delta N_{\lambda_I}^{\text{mod}-\text{nom}})$. The value calculated

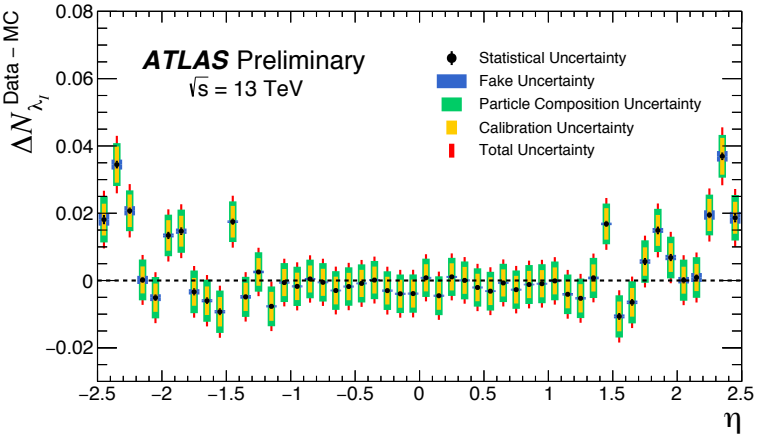


FIGURE 4.41: The excess amount of material between the pixel and SCT detector associated with nuclear interactions in data, $\Delta N_{\lambda_I}^{\mathrm{Data-MC}}$, based on the track-extension efficiency measurement. The uncertainties shown include the uncertainty from fake pixel-tracklets, the uncertainty from the particle composition and p_{T}-dependence, the uncertainty from the material location and the statistical uncertainties. Ref. [113].

as $0.004 + 0.04 \cdot \Delta N_{\lambda_I}^{\mathrm{Data-MC}}$ is taken as an additional systematic uncertainty in the final results.

4.5.12 Track-Extension Efficiency Results

Figure 4.41 shows the number of nuclear interaction lengths missing in simulation as a function of η. Detailed values can be read in Table 4.9. It can be observed that the main sources of systematic uncertainty derive from particle composition, which has an impact on the track-extension efficiency, and from the calibration procedure which affects directly the final measurements of $\Delta N_{\lambda_I}^{\mathrm{Data-MC}}$.

The results of the track extension efficiency method exhibit, within the uncertainties, good agreement between data and simulation in the pseudorapidity region $|\eta| < 1$. The geometry model of the pixel services in this region was highly optimised for Run 1 and no major changes occurred between Run 1 and Run 2. For the forward region, $\Delta N_{\lambda_I}^{\mathrm{Data-MC}}(\eta)$ is greater than zero for $1.4 < |\eta| < 1.5$, $1.8 < |\eta| < 2.0$ and $2.3 < |\eta| < 2.5$. This indicates some missing material in the corresponding regions of the geometry model. The maximum of $\Delta N_{\lambda_I}^{\mathrm{Data-MC}}$ of $(3.7 \pm 0.9)\%$ is observed at $2.3 < |\eta| < 2.4$, as can be seen in Table 4.9. This corresponds to approximately 10% more material in the pixel service region at the corresponding location in η.

Auxiliary plots showing features of the track-extension efficiency are collected in Appendix B.

4.6 Summary and Conclusions

A good description of the distribution of material in the inner detector is crucial for understanding the performance of track reconstruction within ATLAS. Three complementary techniques, hadronic interaction and photon conversion vertex reconstruction together with an estimation using the track-extension efficiency, are applied to measure the inner detector's material using low-luminosity $\sqrt{s} = 13\mathrm{TeV}$ pp collisions at the LHC. While the

TABLE 4.9: Excess amount of material associated with nuclear interactions in data compared to simulation, $\Delta N_{\lambda_I}^{\mathrm{Data-MC}}$, derived from the track-extension efficiency as a function of η together with the uncertainties. *Syst.(particle comp.)* is the systematic uncertainty related to particle composition, *Syst.(fake)* is the uncertainty of the fake rate and *Syst.(calibration)* is the uncertainty associated with the calibration procedure. The total uncertainty is the sum in quadrature of the statistical and systematic components. The measurements are labelled by their bin centre in η.

η	Excess amount of material in data: $\Delta N_{\lambda_I}^{\mathrm{Data-MC}}$ [%]					
	Value	Stat.	Syst.(particle comp.)	Syst.(fake)	Syst.(calibration)	Total uncertainty
-2.45	1.81	±0.14	±0.68	±0.18	±0.47	±0.86
-2.35	3.45	±0.11	±0.64	±0.15	±0.54	±0.85
-2.25	2.07	±0.09	±0.61	±0.14	±0.48	±0.79
-2.15	0.02	±0.09	±0.61	±0.12	±0.40	±0.74
-2.05	-0.51	±0.10	±0.61	±0.10	±0.42	±0.76
-1.95	1.34	±0.11	±0.62	±0.06	±0.45	±0.78
-1.85	1.47	±0.11	±0.64	±0.08	±0.46	±0.80
-1.75	-0.34	±0.10	±0.63	±0.08	±0.41	±0.77
-1.65	-0.60	±0.10	±0.63	±0.05	±0.42	±0.76
-1.55	-0.93	±0.10	±0.63	±0.05	±0.44	±0.77
-1.45	1.75	±0.09	±0.60	±0.04	±0.47	±0.77
-1.35	-0.49	±0.08	±0.60	±0.05	±0.42	±0.73
-1.25	0.25	±0.07	±0.59	±0.04	±0.41	±0.72
-1.15	-0.77	±0.07	±0.59	±0.03	±0.43	±0.74
-1.05	-0.06	±0.06	±0.58	±0.03	±0.40	±0.71
-0.95	-0.17	±0.07	±0.58	±0.03	±0.41	±0.71
-0.85	0.05	±0.06	±0.57	±0.03	±0.40	±0.70
-0.75	-0.05	±0.06	±0.57	±0.03	±0.40	±0.70
-0.65	-0.30	±0.06	±0.58	±0.02	±0.41	±0.71
-0.55	-0.18	±0.06	±0.57	±0.03	±0.41	±0.70
-0.45	-0.09	±0.05	±0.57	±0.03	±0.40	±0.70
-0.35	0.01	±0.06	±0.57	±0.03	±0.40	±0.70
-0.25	-0.30	±0.07	±0.57	±0.03	±0.41	±0.71
-0.15	-0.39	±0.06	±0.57	±0.03	±0.42	±0.71
-0.05	-0.39	±0.06	±0.57	±0.03	±0.42	±0.71
0.05	0.08	±0.06	±0.57	±0.03	±0.40	±0.70
0.15	-0.45	±0.07	±0.58	±0.03	±0.42	±0.72
0.25	0.11	±0.06	±0.57	±0.03	±0.40	±0.70
0.35	-0.00	±0.05	±0.56	±0.03	±0.40	±0.69
0.45	-0.21	±0.06	±0.57	±0.03	±0.41	±0.70
0.55	-0.32	±0.06	±0.58	±0.03	±0.41	±0.71
0.65	-0.07	±0.06	±0.57	±0.03	±0.40	±0.70
0.75	-0.27	±0.06	±0.57	±0.03	±0.41	±0.71
0.85	-0.12	±0.07	±0.58	±0.02	±0.40	±0.71
0.95	-0.09	±0.07	±0.58	±0.03	±0.40	±0.71
1.05	-0.01	±0.06	±0.58	±0.04	±0.40	±0.71
1.15	-0.41	±0.07	±0.59	±0.04	±0.42	±0.73
1.25	-0.53	±0.07	±0.59	±0.03	±0.42	±0.73
1.35	0.08	±0.08	±0.60	±0.04	±0.40	±0.73
1.45	1.68	±0.09	±0.60	±0.04	±0.47	±0.77
1.55	-1.07	±0.10	±0.63	±0.06	±0.44	±0.78
1.65	-0.65	±0.10	±0.63	±0.05	±0.43	±0.77
1.75	0.56	±0.10	±0.63	±0.07	±0.42	±0.77
1.85	1.49	±0.11	±0.64	±0.07	±0.46	±0.80
1.95	0.68	±0.11	±0.62	±0.08	±0.43	±0.76
2.05	0.01	±0.09	±0.61	±0.10	±0.40	±0.74
2.15	0.09	±0.09	±0.60	±0.13	±0.40	±0.74
2.25	1.95	±0.09	±0.61	±0.14	±0.48	±0.79
2.35	3.69	±0.11	±0.64	±0.16	±0.55	±0.86
2.45	1.87	±0.14	±0.68	±0.18	±0.47	±0.86

first two methods probe the barrel region of the inner detector, in particular the new detector components installed in Run 2 (the beam pipe, the IBL and the supporting tubes of IPT and IST), the track-extension efficiency method has sensitivity also in the forward η region of $1.0 < |\eta| < 2.5$, in which most of the refurbished pixel services reside.

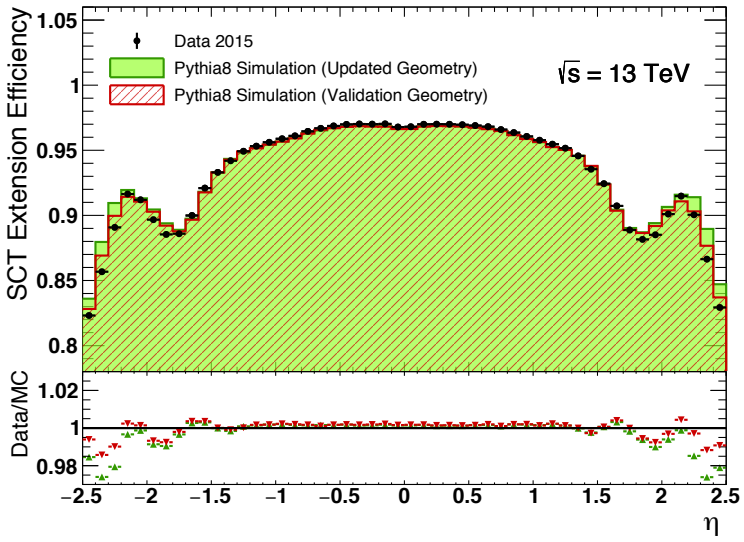


FIGURE 4.42: Track-extension efficiency as a function of η of the tracklets in a comparison between data and Pythia 8 simulation for the *updated* and *validation* geometry model. The only difference between these two geometries is an improved description of the pixel services implemented in the *validation* geometry.

The description of the geometry model was examined in detail both in radial and longitudinal distributions of the rate of reconstructed hadronic interaction and photon conversion vertices. In the central barrel region, a significant amount of missing material in the IBL front-end electronics for the flex bus, surface-mounted devices on the front-end chips and the IPT and IST was identified in the *original* geometry model (the geometry model used for ATLAS MC simulation in 2015). The *updated* geometry model, which was created to resolve the above discrepancies, provides a much better description. The beam pipe is found to be very accurately described except the centremost region of $|z| < 40$ mm. The simulated material in the IBL within the *updated* geometry model is found to be consistent with that in observed data, within the less than 10% uncertainties of the hadronic interaction and conversion measurements. The existing pixel barrel layers are found to be described well, and the results from the analyses using the hadronic interactions and photon conversions agree within the systematic uncertainties. This result confirms the results of the previous hadronic interaction analysis using the Run 1 data set.

While sensitive to slightly different material properties, both the hadronic interactions and photon conversions provide a consistent understanding of the barrel detector material. The *updated* geometry model provides reasonable agreement with the data in the ratio of the rate measurements of hadronic interactions and photon conversions within the uncertainties of the measurements. The measured rates of photon conversions and hadronic interactions reconstructed in data are found to agree to within 7–18% with those predicted by simulation, based on the *updated* geometry model, up to the outer envelope of the pixel detector.

The results of these studies were of crucial importance to implement an improved description of the material in the ATLAS inner detector simulation, already used in the 13 TeV charged-particle analysis presented in Chapter 5.

In the forward region, the material in the pixel service region is found to be underestimated in the geometry model by up to $\Delta N_{\lambda_I} = (3.7 \pm 0.9)\%$ at some values of η. This corresponds roughly to 10% of the material in the pixel services in the corresponding

regions. The results shown here have already been used to improve the systematic uncertainties on the 13 TeV charged-particle analysis presented in Chapter 5. At the mean time, a thorough investigation of the geometry model was required to improve the complex description of the material in the pixel service region and a better description has been produced which allows for about 50% reduction in the data/simulation discrepancies, as shown in Figure 4.42. For clarity, it must be stated that the selection criteria used to measure the track-extension efficiency presented in 4.42 do not include the cuts introduced in the last iterations of the analysis as described in Section 4.5.10.

This new description of the pixel services region was added on top of an improved *updated* geometry, which implies the fix of the description of the beam pipe in the central region of $z < 40$ mm described in Section 4.4.1. The resulting ID geometry will be used as a baseline in future Run 2 ATLAS analyses.

Chapter 5

Charged-Particle Multiplicities

Mistakes are
the portals
of discovery

-James Joyce-

This chapter presents the most recent charged-particle spectra measured by using data collected with the ATLAS detector at the centre-of-mass energy of 13 TeV, with a particular emphasis on the analysis performed in a fiducial phase space in which tracks with transverse momentum p_{T} higher than 500 MeV and in a pseudorapidity region $|\eta| < 2.5$ are selected. Differences in procedures and results with respect to the analysis performed with very low momentum tracks ($p_{\mathrm{T}} > 100$ MeV) will be explained when relevant. In general, the accent will be on the tracking-related aspects in which the Author of this thesis led the work, but other important aspects of the analysis, such as trigger and vertexing efficiency, background evaluation and unfolding of the final results, will be reported for completeness.

5.1 Introduction

The measurements of inclusive charged-particle spectra provide insight into the low energy non-perturbative region of quantum chromodynamics. A description of low-energy processes within a perturbative framework is not possible in this regime, thus charged-particle interactions are typically described by QCD-like models implemented in Monte Carlo event generators and constrained with data-driven measurements. Furthermore, soft processes, arising from pile-up at high luminosity, which leads to more than one interaction per beam crossing, may also affect the event topologies triggered by a specific hard-scattering interaction. An understanding of soft QCD processes is therefore important both in its own right and as a means of reducing systematic uncertainties in measurements of high transverse momentum phenomena.

In the analysis of charged-particles, the most commonly measured distributions are:

- $\frac{1}{N_{\mathrm{ev}}} \cdot \frac{\mathrm{d}N_{\mathrm{ch}}}{\mathrm{d}\eta}$
- $\frac{1}{N_{\mathrm{ev}}} \cdot \frac{1}{2\pi p_{\mathrm{T}}} \cdot \frac{\mathrm{d}^2 N_{\mathrm{ch}}}{\mathrm{d}\eta \mathrm{d}p_{\mathrm{T}}}$
- $\frac{1}{N_{\mathrm{ev}}} \cdot \frac{\mathrm{d}N_{\mathrm{ev}}}{\mathrm{d}n_{\mathrm{ch}}}$
- $\langle p_{\mathrm{T}} \rangle$ as a function of n_{ch}

where p_{T} is the track momentum component that is transverse to the beam direction, η is the track pseudo rapidity, n_{ch} is the number of charged-particles in an event, N_{ev} is the

number of selected minimum bias events, N_{ch} is the total number of charged-particles in the data sample and $\langle p_{\mathrm{T}} \rangle$ is the average p_{T} for a given number of charged-particles.

The average charged-particle density as a function of the pseudorapidity $\frac{1}{N_{\mathrm{ev}}} \cdot \frac{\mathrm{d}N_{\mathrm{ch}}}{\mathrm{d}\eta}$ is sensitive to the fraction of energy of the collision that is converted into soft particles, which are produced at central pseudorapidity. As the collision energy increases the rate of multiple parton interactions per collision rises, which increases the density at central pseudorapidity. Particle production becomes more central at higher collision energies.

The distribution of the average momentum as a function of the number of charged-particles was only introduced in recent years because it was found to be useful in constraining the colour reconnection parameters in Pythia (see Section 1.3.1).

A way to study the dependence of the multiplicity on the centre-of-mass energy of the collision is to compare the average number of charged particles at central pseudorapidity. The mean number of primary charged-particles in the very central region $|\eta| < 0.2$ can be compared between various measurements at different centre-of-mass energies and between various experiments which have limited acceptance.

The charged-particle distributions are one of the first measurements at a collider experiment, as little integrated luminosity is required due to the large production cross-section (see Section 1.2.2). Since one of the largest uncertainties in the phenomenological models is the extrapolation of the multiplicity as a function of energy, the more the energy values at which the data is collected, the more accurate the tuning can be.

Various collider experiments have measured charged-particle distributions, including the Split Field Detector (SFM [132]) [133] and the Streamer Chamber Detector [134, 135] at the Intersecting Storage Ring (ISR [136]); the Underground Area 1 (UA1) [137, 138], Underground Area 5 (UA5 [139]) [140–148] and the P238 experiment [149] at the Super Proton Synchrotron; and E735 [150, 151] and the Collider Detector at Fermilab (CDF [152]) [153–157] at the Tevatron. In the LHC era, charged-particle distributions have been measured by the CMS Collaboration, as detailed in [158–161] and in the references therein, as well as by the ALICE Collaboration [162–164] and by the LHCb collaboration [165]. The ATLAS Collaboration has also performed measurements of charged-particle observables at various centre-of-mass energies [122, 166, 167].

It is interesting to see how different measurements can be combined and which is the level of agreement with the Monte Carlo simulation.

Figure 5.1 shows the 13 TeV CMS measurement [161] of $\frac{\mathrm{d}N_{\mathrm{ch}}}{\mathrm{d}\eta}$ in $|\eta| < 0.5$ in a comparison with PYTHIA 8 and EPOS LHC predictions and with previous results from ISR [135, 168], UA5 [142, 144], PHOBOS [169], and ALICE [162] data at various centre-of-mass energies. The measured values are empirically fitted using a second-order polynomial in $\ln(s)$ as $1.55 - 0.113\ln(s) + 0.0168\ln(s)^2$, where s has the units GeV^2, which provides a good description of the available data over the full energy range. The PYTHIA 8 and EPOS LHC event generators globally reproduce the collision-energy dependence of hadron production in inelastic pp collisions.

Figure 5.2 is a compilation of data-only results on pseudorapidity density of charged-particles measured in $|\eta| < 0.5$. The measurements done by the ALICE collaboration reported here have been obtained for inelastic events (INEL) and for events having at least one charged particle produced with $p_{\mathrm{T}} > 0$ in the pseudorapidity interval $|\eta| < 1$ (INEL>0). These measurements are compared with results at different proton-proton collider energies [164]. The energy dependence of $\frac{\mathrm{d}N_{\mathrm{ch}}}{\mathrm{d}\eta}$ is parametrised by the power law as^b fitted to data, where a and b are free parameters. By combining the data at lower energies with ALICE and CMS results at $\sqrt{s}$=13 TeV, the values obtained are $b = 0.103 \pm 0.002$ and $b = 0.111 \pm 0.004$ for INEL and INEL > 0 event classes, respectively.

The most recent charged-particle measurements produced by the ATLAS collaboration with the analysis of pp collisions at the centre-of-mass energy of 13 TeV [80, 81] will be

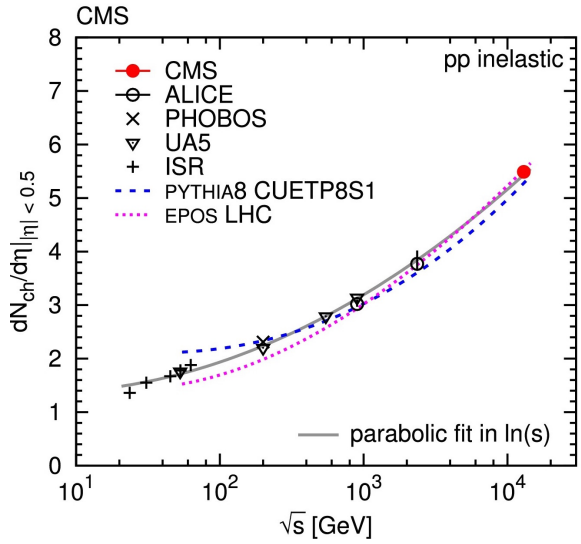


FIGURE 5.1: Centre-of-mass energy dependence of $\frac{\mathrm{d}N_{\mathrm{ch}}}{\mathrm{d}\eta}$ for $|\eta| < 0.5$, including ISR [135, 168], UA5 [142, 144], PHOBOS [169], and ALICE [162] data. The solid curve shows a second-order polynomial in $\ln(s)$ fit to the data points, including the new result at $\sqrt{s} = 13$ TeV. The dashed and dotted curves show the PYTHIA 8 and EPOS LHC predictions, respectively. From [161].

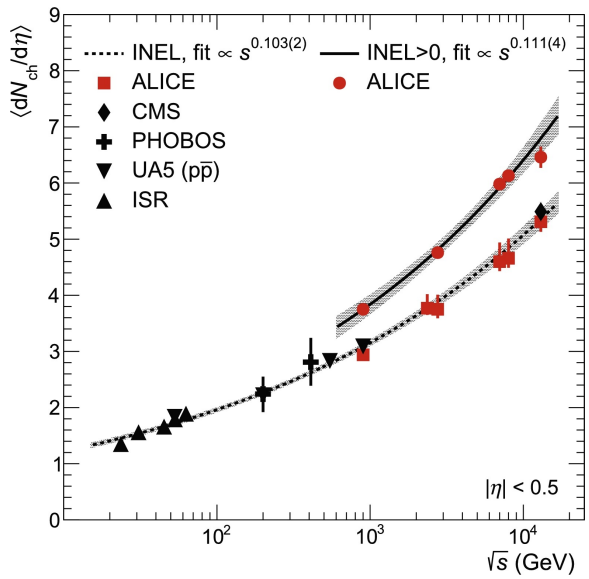


FIGURE 5.2: Charged-particle pseudorapidity density measured in the central pseudorapidity region $|\eta| < 0.5$ for INEL and INEL > 0 events. The uncertainties are the quadratic sum of statistical and systematic contributions. The lines are power-law fits of the energy dependence of the data and the grey bands represent the standard deviation of the fits. From [164].

reported in this Chapter with comparison to previous ATLAS results.

5.2 Charged-Particles in $\sqrt{s} = 13$ TeV pp collisions with ATLAS

The distributions of primary charged-particles were measured by using ~151 μb^{-1} of data recorded by the ATLAS experiment in June 2015 at a center-of-mass energy of 13 TeV. This dataset is of similar size as the one used in the previous ATLAS minimum-bias publications [122, 166]. The analysis strategy also follows closely what was done in the previous analyses, but with some differences that will be highlighted in the following. Primary charged-particles are defined as charged-particles with a mean lifetime $\tau > 300$ ps, which are either directly produced in pp interactions or from subsequent decays of directly produced particles with a shorter lifetime $\tau < 30$ ps. In this analysis, differently than earlier measurements, particles with a mean lifetime between $30 < \tau < 300$ ps and their decay products are not included in the primary particle definition due of experimental limitations. In fact, these particles are charged strange baryons which tend to decay within the detector volume and, especially if they have low momentum, they often do not leave enough hits to reconstruct a track, leading to a very low track reconstruction efficiency of approximately 0.3%. Details will be given in Section 5.9 and a summary of these requirements is reported in Table 5.1.

Particle	Charged	Requirement 1	Requirement 2	Fiducial charged primary	Decay prod. as fiducial charged primary
B^0,D^0	No	No	No	No	Yes
$B^\pm$,$D^\pm$	Yes	No	No	No	Yes
strong resonance	?	No	No	No	Yes
strong resonance from decays	?	No	No	No	Yes
charged strange hadrons(except $K^\pm$)	yes	Yes	No	No	No
neutral strange hadrons	no	Yes	No	No	No
charged stables ($\pi^\pm$,$K^\pm$,p,$e^\pm$,$\mu^\pm$)	Yes	Yes	Yes	Yes	no decay
charged stables (n)	No	Yes	Yes	No	no decay

TABLE 5.1: Look up table for the different cases for particle classification. Requirement 1 is on primary particles with a life time below $\tau < 30$ ps. These particles have been excluded from the stable particle definition and their decay products are considered to be primary particles. Requirement 2 is a new life time requirement added to this analysis (a change with respect to previous publications [122, 166]). All particles with a mean lifetime between 30 and 300 ps as well as their decay products are considered to be background secondary particles. Fiducial charged primaries are the particles used as signal in this analysis.

The *nominal* fiducial phase space includes events with

- at least 1 charged-particle with $|\eta| < 2.5$ and $p_{\mathrm{T}} > 500$ MeV,

but two other phase-space selections were considered in the studies of charged-particle distributions at 13 TeV:

- at least 2 charged-particle with $|\eta| < 2.5$ and $p_{\mathrm{T}} > 100$ MeV (*extended*);
- at least 1 charged-particle with $|\eta| < 0.8$ and $p_{\mathrm{T}} > 500$ MeV (*reduced*).

Diffractive physics is expected to mostly contribute at low number of charged-particles and at low track momentum. Therefore varying the selection cuts on n_{ch} (not done at 13 TeV, but extensively studied at 8 TeV [167]) and p_{T} in effect affects the relative contribution from diffractive events.

The quantities introduced in Section 5.1, $\frac{1}{N_{\mathrm{ev}}} \cdot \frac{\mathrm{d}N_{\mathrm{ch}}}{\mathrm{d}\eta}$, $\frac{1}{N_{\mathrm{ev}}} \cdot \frac{1}{2\pi p_{\mathrm{T}}} \cdot \frac{\mathrm{d}^2 N_{\mathrm{ch}}}{\mathrm{d}\eta \mathrm{d}p_{\mathrm{T}}}$, $\frac{1}{N_{\mathrm{ev}}} \cdot \frac{\mathrm{d}N_{\mathrm{ev}}}{\mathrm{d}n_{\mathrm{ch}}}$ and $\langle p_{\mathrm{T}} \rangle$ as a function of n_{ch}, are presented here for data and compared to Monte Carlo simulation predictions. The mean number of primary charged-particles in the very central region $|\eta| < 0.2$ was also measured and compared with results at lower centre-of-mass energies.

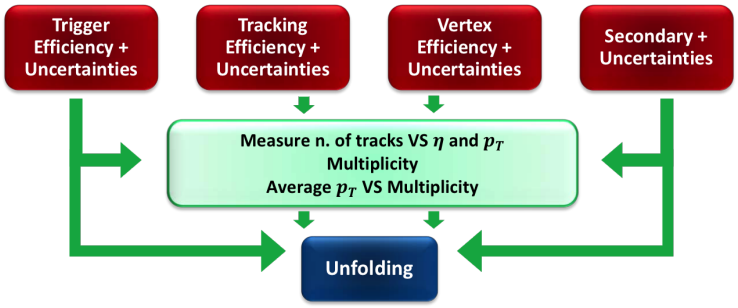


FIGURE 5.3: Schema representing the main steps of the Minimum Bias analysis, such as trigger, tracking, vertex efficiency and secondary evaluation together with the respective uncertainties, and the main observables to be measured. Once these quantities are measured, the data are unfolded and the final results can be obtained.

Monte Carlo simulation is used extensively throughout this analysis and the differences in the modeling of the detector response to various effects is a leading uncertainty. A summary of the Monte Carlo samples used in the analysis will be given in Section 5.3.

The data needs to be corrected for a number of detector related effects to obtain the true primary charged-particle distributions (unfolding) and compare the measurements to particle-level MC predictions. After selecting appropriate events and ensuring that the data within those events are of good quality (described in Section 5.4), the event selection efficiency needs to be measured and corrected. In particular the trigger efficiency and its uncertainties (which are presented in Section 5.5), and the vertex reconstruction efficiency and uncertainties (Section 5.6) are of crucial importance.

Also the contribution of non-collision background events and non-primary tracks to the final event and track selections, respectively, need to be estimated and corrected. The estimation of these backgrounds on the measurement will be discussed in Section 5.8.

Track reconstruction will be discussed in Section 5.7, where an estimate of the track reconstruction efficiency and its uncertainty and performance will be given.

Section 5.10 will review the procedure applied to the selected data to correct for detector effects.

The final results and comparisons to Monte Carlo generators will be presented in Section 5.11.

All these steps are also schematically represented in Figure 5.3.

5.3 Simulated samples

As already mentioned in Chapter 1 and summarised in Table 1.1, the PYTHIA 8 (both A2 and Monash tunes), HERWIG++ (only shown in preliminary plots), EPOS and QGSJET-II event generators are used in the analysis of charged-particle distributions.

The PYTHIA 8 generator with the A2 tune is used as baseline setting to produce MC samples of *non-diffractive*, *single-diffractive* and *double-diffractive* processes. The contributions of these three processes are mixed together in the generated samples according to the generator cross sections. All the events are processed through the ATLAS detector simulation program [170, 171], which is based on GEANT4 [103]. They are then reconstructed and analyzed using the same program chain as the one used for data.

To assess the dependence on the material description of the inner detector simulations, geometry models with distorted passive material with respect to the nominal material description are used, in agreement with the material studies presented in Chapter 4. The

charged-particles analysis is the only analysis published by ATLAS in the period 2015-2016 which uses the *updated* geometry model as a baseline. Table 5.2 summarizes the samples used for these studies with short details about the version of the Monte Carlo event generator and the tunes used.

In all Monte Carlo samples, the primary vertex distribution in the Monte Carlo is re-weighted to match the primary vertex distribution in data. In addition, based on the weighted matching probability P_{match} introduced in Chapter 3, a reconstructed track with $P_{\text{match}} > 0.5$ is tagged as a primary track if it emerges from the Monte Carlo primary vertex, as a secondary track if it is the result of simulated physics interactions within the detector.

5.4 Data Selection

The $13\,\text{TeV}$ minimum bias data sample used in this analysis is the same as the one described in Section 3.4.1, which after applying the data quality requirements corresponds to an integrated luminosity of $151\mu\text{b}^{-1}$. Only events for which the solenoid magnet was turned on are used for this analysis. During the data-taking period, more than 99% of the IBL, 97% of the pixel detector, 99% of the SCT and 98% of the TRT were operational. Events are collected from colliding proton bunches using a trigger which required has recorded one or more MBTS counters above threshold on either side of the detector.

To reduce the contribution from background events and non-primary tracks, events are required to satisfy the following criteria:

- to fire at L1 the MBTS trigger on at least one side of the ATLAS detector (called *single-arm trigger*, technically L1_MBTS_1). This trigger is passed through the event filter and the event is always recorded;
- contain a primary vertex [104];
- not contain a second vertex with four or more tracks (in order to remove events with more than one interaction per beam crossing)
- at least one good track in the event ($n_{\text{sel}} \geq 1$), where the track is defined as follows:
  - A transverse momentum above $500\,\text{MeV}$ and $|\eta| < 2.5$ (momentum and pseudorapidity limits changed accondingly to the analised phase space, as described above);
  - At least 1 pixel hit;
  - If one hit is expected in the IBL, then one hit required. The active area coverage of the IBL layer is more than 99%. If a track passes through an inactive IBL module, then one hit is required in the next layer if one hit is expected. This requirement suppresses tracks from secondary particles;
  - At least 6 SCT hits for tracks with a transverse momentum above $500\,\text{MeV}$ and $|\eta| < 2.5$ (the requirement is set to 6, 4 and 2 SCT hits respectively for $p_{\text{T}} > 400$, $p_{\text{T}} < 400$ or 300 MeV in the case of the low-p_{T} analysis). If a track passes through an inactive layer, it is counted as a hit. This makes the analysis less sensitive to differences in the number of dead module in data and simulation;
  - The transverse impact parameter d_0 calculated with respect to the LHC beam line is required to be less than 1.5 mm ($|d_0| < 1.5\,\text{mm}$).
  - The longitudinal impact parameter z_0 is calculated with respect to the primary vertex. It is required that the distance between the primary vertex and the track at the point where we measure the d_0 multiplied by $\sin(\theta)$ is $z_0 \cdot \sin\theta < 1.5\,\text{mm}$;

– If the track p_{T} exceeds $10\,\mathrm{GeV}$, then the track χ^2 probability must be bigger than 0.01. This is done to suppress mis-measured tracks.

Generator	Version/Tune	Model	Generated	Simulated	Geometry
			Updated Geometry		
PYTHIA 8	8.185 A2	Non diffractive	125M	20M	ATLAS-R2-2015-03-06-00
PYTHIA 8	8.185 A2	Single diffractive	15M	5M	ATLAS-R2-2015-03-06-00
PYTHIA 8	8.185 A2	Double diffractive	15M	5M	ATLAS-R2-2015-03-06-00
EPOS	1.3 LHC	inclusive	100M	10M	ATLAS-R2-2015-03-06-00
PYTHIA 8	8.185 A2	Non diffractive	0	5M	+10% IBL , ATLAS-R2-2015-03-06-01
PYTHIA 8	8.185 A2	ND, $n_{charged} > 120$	1M	1M	ATLAS-R2-2015-03-06-00
PYTHIA 8	8.185 A2	ND, $n_{charged} > 160$	1M	1M	ATLAS-R2-2015-03-06-00
PYTHIA 8	8.185 A2	ND, $n_{charged} > 200$	1M	ATLAS-R2-2015-03-06-00	
particle gun		Pi-plus	0	5M	ATLAS-R2-2015-03-06-00
particle gun		Pi-minus	0	5M	ATLAS-R2-2015-03-06-00
particle gun		K-plus	0	1M	ATLAS-R2-2015-03-06-00
particle gun		K-minus	0	1M	ATLAS-R2-2015-03-06-00
particle gun		Proton	0	500k	ATLAS-R2-2015-03-06-00
particle gun		Anti-Proton	0	500k	ATLAS-R2-2015-03-06-00
			Original Geometry		
PYTHIA 8	8.185 A2	Non diffractive	125M	20M	ATLAS-R2-2015-03-01-00
PYTHIA 8	8.185 A2	Single diffractive	15M	5M	ATLAS-R2-2015-03-01-00
PYTHIA 8	8.185 A2	Double diffractive	15M	5M	ATLAS-R2-2015-03-01-00
PYTHIA 8	8.186 Monash	Non diffractive	105M	20M	ATLAS-R2-2015-03-01-00
PYTHIA 8	8.186 Monash	Single diffractive	15M	5M	ATLAS-R2-2015-03-01-00
PYTHIA 8	8.186 Monash	Double diffractive	15M	5M	ATLAS-R2-2015-03-01-00
PYTHIA 8	8.185 A2	Non diffractive	0	5M	+2.5%, ATLAS-R2-2015-03-01-01
PYTHIA 8	8.185 A2	Non diffractive	0	5M	+5% , ATLAS-R2-2015-03-01-02
PYTHIA 8	8.185 A2	Non diffractive	0	5M	+10%,ATLAS-R2-2015-03-01-03
PYTHIA 8	8.185 A2	Non diffractive	0	5M	+10% PP0,ATLAS-R2-2015-03-01-04
PYTHIA 8	8.185 A2	Non diffractive	0	5M	+25% PP0,ATLAS-R2-2015-03-01-18
PYTHIA 8	8.185 A2	Non diffractive	0	5M	+50% PP0,ATLAS-R2-2015-03-01-19
PYTHIA 8	8.185 A2	Non diffractive	0	5M	+10% IBL, ATLAS-R2-2015-03-01-16
PYTHIA 8	8.185 A2	Non diffractive	0	5M	+20% IBL, ATLAS-R2-2015-03-01-17
EPOS	1.3 LHC	inclusive	100M	10M	ATLAS-R2-2015-03-01-00
HERWIG++	2.7.1 UEEE5	inclusive	100M	10M	ATLAS-R2-2015-03-01-00
HERWIG++	2.7.1 UEEE4	inclusive	100M	0	No
QGSJET-II	II-04 LHC	inclusive	100M	0	No
PYTHIA 8	8.185 A2	ND, $n_{charged} > 120$	1M	1M	ATLAS-R2-2015-03-01-00
PYTHIA 8	8.185 A2	ND, $n_{charged} > 160$	1M	1M	ATLAS-R2-2015-03-01-00
PYTHIA 8	8.185 A2	ND, $n_{charged} > 200$	1M	1M	ATLAS-R2-2015-03-01-00
particle gun		Pi-plus	0	5M	ATLAS-R2-2015-03-01-00
particle gun		Pi-minus	0	5M	ATLAS-R2-2015-03-01-00
particle gun		K-plus	0	1M	ATLAS-R2-2015-03-01-00
particle gun		K-minus	0	1M	ATLAS-R2-2015-03-01-00
particle gun		Proton	0	500k	ATLAS-R2-2015-03-01-00
particle gun		Anti-Proton	0	500k	ATLAS-R2-2015-03-01-00

TABLE 5.2: Details of the MC events used to assess the track reconstruction efficiency and its dependence on the material description used in the inner detector simulation. Two principal geometries were used for these samples (referred to as *updated* and *original*) with the only significant difference between them being description of the IBL. More details were provided in Chapter 4

.

	Simulation **Events**	% passing	Data **Events**	% passing
Before cuts	10000000		10855221	
After requiring a primary vertex is present	8323843	83.24%	9343833	86.08%
After pileup suppression	8323504	99.99%	9316964	99.71%
After requiring 1 track above p_{T} > 500 MeV	7984705	95.93%	8870790	95.21%

TABLE 5.3: Number of events before and after applying all the event selection criteria. For data the first row displays the data after only the trigger requirement is applied and the Monte Carlo with no cuts applied.

	Simulation **Tracks**	% passing	Data **Tracks**	% passing
All Tracks	293317640		347194760	
$p_{\mathrm{T}} > 500$ MeV	137242338	46.79%	155759984	44.86%
$\lvert\eta\rvert < 2.5$	134359043	97.90%	152557877	97.94%
Innermost hit requirement	129037285	96.04%	140837448	92.32%
Pixel Hit > 1	112547309	87.22%	121075669	85.97%
SCT Hits > 6	103795189	92.22%	111705230	92.26%
$\lvert d_0\rvert < 1.5\,\mathrm{mm}$	98074303	94.49%	107545661	96.28%
$\lvert\Delta(z_0 \cdot \sin\theta)\rvert < 1.5\,\mathrm{mm}$	96726666	98.63%	106354071	98.89%
χ^2 probability	96726099	100.00%	106353390	100.00%

TABLE 5.4: Number of tracks after event selection before and after applying all the track selection criteria.

The numbers of events and tracks selected before and after applying all the requirements are given in Tables 5.3 and 5.4. The differences in requiring a primary vertex between data and MC in Table 5.3 are a result of requiring a trigger in data and not in the simulation. This results in a larger fraction of low multiplicity events in the simulation than in the data leading to the lower cut efficiency in the simulation. The simulation does not include pile-up, hence differences are expected at the level of ≈0.3%, which is about the difference seen on the pileup suppression monitoring.

Small differences in the track selection efficiency can be seen for a number of the criteria, however, for the majority of them it is difficult to disentangle detector efficiency issues from physics differences at this stage and as such differences in the fraction of tracks passing each criteria should not be considered as an overall systematic uncertainty. For example, the requirement of a hit in the innermost layer and on the number of pixel hits may differ because there is higher rate of secondary particles in data than simulation. The impact parameter cuts could differ because of different kinematic properties of primaries and secondaries in data and simulation. More details will be given in the following.

5.4.1 Inner Detector performance distributions

In the following, several representative distributions of Inner Detector performance are shown which, compared to the ones presented in Section 3.4.3, were obtained by applying the track selection described above (Section 5.4). This selection, differently than the Loose and Tight Primary selection (Section 3.4.2), includes specific requirements, for instance, on the transverse and longitudinal impact parameters as well as on the track χ^2 probability.

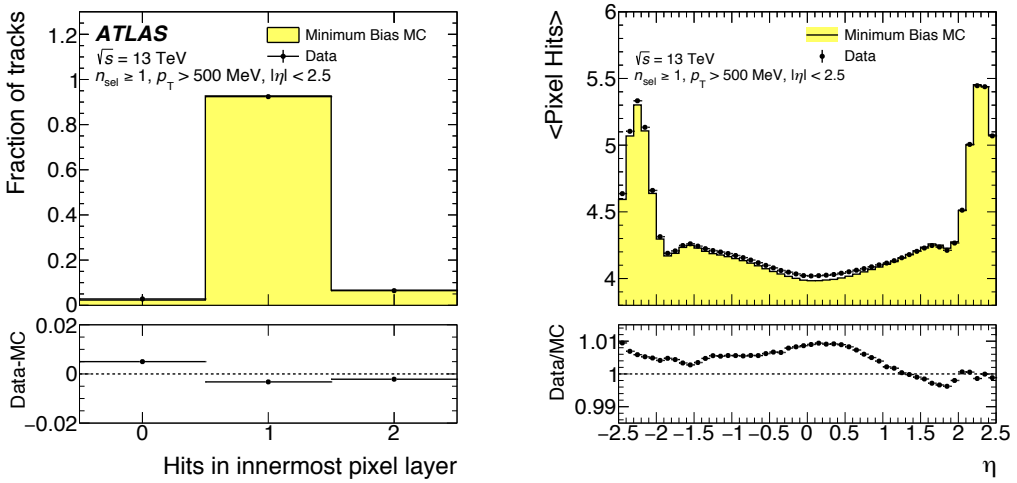
Figure 5.4 shows the average number of IBL ((a)) and pixel (b) hits per track as a function of η. The average number of IBL hits has two distinct features. The first is the reduced average number of hits seen in data, this is due to a difference in the number of tracks passing through two modules in the overlap region of the IBL staves. In addition, the fraction of background tracks rejected with the IBL hit cut is different in data and Monte Carlo. The second is that in the region ($1.0 < \eta < 2.0$) an IBL module had a dead readout chip for both runs leading to a reduced efficiency (50%). These differences are seen in the average number of pixel hits as well.

Figure 5.5 shows the average number of hits on reconstructed tracks as a function of η in the entire SCT ((a)) and silicon tracker ((b)) detectors. The number of SCT hits on track for all but the end cap region of the detector matches the simulation very precisely. In the end-cap region differences in the number of active modules and strips are the expected cause of the disagreements (see Table 4.7).

How these differences in the detector conditions and performance are taken into account in the analysis is discussed in Section 5.7.

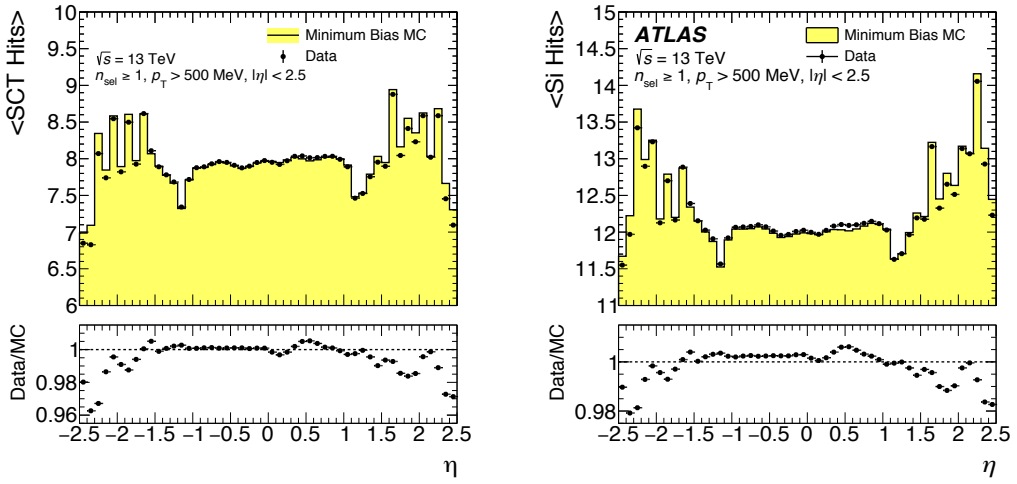
Figure 5.6 shows the transverse impact parameter ((a)) and longitudinal impact parameter ((b)) for tracks with transverse momentum larger than 500 MeV. The width of the central distribution is larger in data than in simulation, partly due to a degradation of the impact parameter resolution caused by imperfect detector alignment and partly caused by having a wider beam spot in data for 50% of the data than the simulation. A detailed study of the impact parameter resolution to calculate the rate of secondaries in data is discussed in Section 5.8.

A special configuration of the track reconstruction algorithms, the MinBias setup described in Chapter 3, was used to reconstruct all the events analysed in the context of the charged-particle multiplicity measurements (independently on the specific phase-space). The most critical change between this setup and the default one, besides lowering the p_{T} threshold to 100 MeV, is a reduced minimum number of silicon hits from seven to five. All tracks, irrespective of their transverse momentum, are reconstructed in a single pass of the track reconstruction algorithm. This allowed to reconstruct low-momentum tracks with



(a) (b)

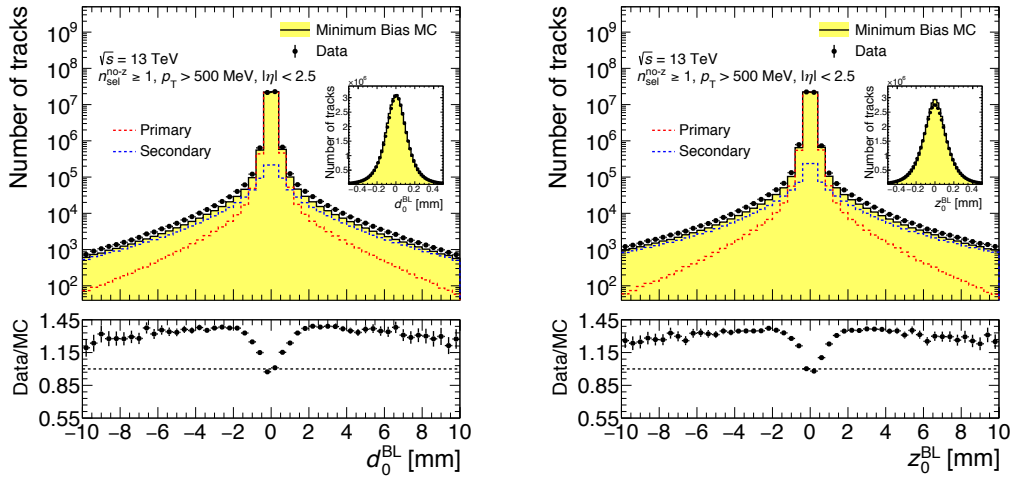
FIGURE 5.4: Comparison between data and simulation at $\sqrt{s} = 13$ TeV for tracks with transverse momentum $p_{\mathrm{T}} > 500$ MeV: number of hits on reconstructed track in the IBL (left) and the average number of hits on reconstructed track as a function of η in the IBL (right). The p_{T} and η distribution of the tracks in MC is re-weighted to match the data and the number of tracks is normalised to the data.



(a) (b)

FIGURE 5.5: Comparison between data and simulation at $\sqrt{s} = 13$ TeV for tracks with transverse momentum $p_{\mathrm{T}} > 500$ MeV: the average number of silicon hits on reconstructed track as a function of η in the SCT (left) and all silicon (right) detectors. The requirement on the distribution shown is removed from the requirements on the reconstructed tracks. The p_{T} distribution and η of the tracks in MC is re-weighted to match the data and the number of tracks is normalised to the data. Ref. [80].

good efficiency and low fake rate. Figure 5.7 shows the comparison between data and simulation for what concerns the distribution of the number of pixel hits associated with a track in the low-momentum region. Data and simulation agree reasonably well given the



(a) (b)

FIGURE 5.6: Comparison between data and simulation at $\sqrt{s} = 13$ TeV for tracks with transverse momentum $p_{\mathrm{T}} > 500$ MeV: the transverse impact parameter d_0 (left) and the longitudinal impact parameter z_0(right). This is before a secondary scale factor is applied. The p_{T} distribution and η of the tracks in MC is re-weighted to match the data and the number of tracks is normalised to the data.

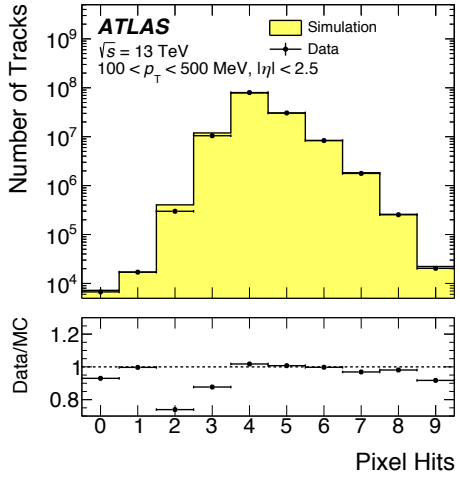


FIGURE 5.7: Comparison between data and Pythia 8 A2 simulation for what concerns the distribution of the number of pixel hits associated with a track. The distribution is shown before the requirement on the number of pixel hits is applied, for tracks with $100 < p_{\mathrm{T}} < 500$ MeV and $\eta < 2.5$. The p_{T} and η distribution of the tracks in MC is re-weighted to match the data and the number of tracks is normalised to the data. Ref. [81].

known imperfections in the simulation of inactive pixel modules. These differences are taken into account in the systematic uncertainty on the tracking efficiency by comparing the efficiency of the pixel hit requirements in data and simulation after applying all other track selection requirements.

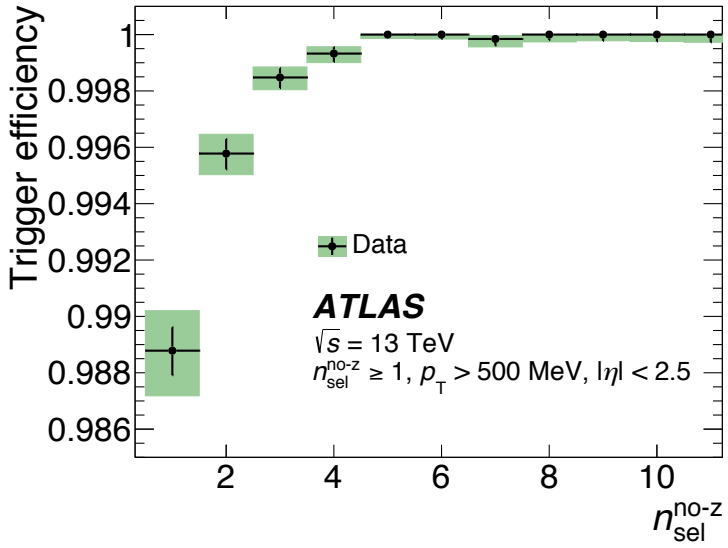


FIGURE 5.8: Trigger efficiency with respect to the event selection as a function of the number of reconstructed tracks ($n_{\text{sel}}^{\text{no-z}}$) with $p_{\text{T}} > 500$ MeV and $|\eta| < 2.5$, without the $z_0^{BL} \cdot \sin\theta$ constraint ($n_{\text{sel}}^{\text{no-z}}$). A reconstructed vertex is not required. The statistical errors are shown as black lines, the total errors as green shaded areas. Ref. [80].

5.5 Trigger

The trigger efficiency, ϵ_{trig}, is measured from a data sample selected using a randomly-seeded control trigger. It is defined as the ratio of events from the control trigger in which the (full trigger chain of) L1 MBTS (L1 term: L1_MBTS_1, full chain: EF_noalg_L1_MBTS_1) also accepts the event, over the total number of events in the control sample. The control trigger used for this analysis selects events from a random filled bunch crossing on L1 and filtered by the HLT. The first step of filtering asks for at least two pixel hits with a time over threshold of at least 20 bunch crossings [1] and at least three SCT hits. Then, the trigger applies a track reconstruction in the event filter. At least one track with $p_{\text{T}} > 200$ MeV and a longitudinal impact parameter with respect to the nominal center of ATLAS of less than $40\,\text{mm}$ needs to be reconstructed. The vertex requirement for selected tracks is removed for this study, to account for correlations between the trigger and vertex reconstruction efficiency. Since there is no vertex requirement, it is not possible to apply the same impact-parameter cuts as for the selected final tracks. Therefore the trigger efficiency is measured using impact parameter constraints with respect to the beam line.

The trigger efficiency is parameterized as a function of the number of selected tracks with $p_{\text{T}} > 500$ MeV, which is labeled $n_{\text{sel}}^{\text{no-z}}$. The definition of $n_{\text{sel}}^{\text{no-z}}$ differs from the analysis selection because the $|z_0^{\text{BS}}|$ selection is loosened to 1000 mm to avoid biasing the vertex position along the beam line. The effiency is 98.8% in the first$n_{\text{sel}}^{\text{no-z}}$ bin and rapidly increases, reaching 100%, for $n_{\text{sel}}^{\text{no-z}} = 6$. The trigger requirement is found to introduce no observable bias in the p_{T} and η distributions of selected tracks within the statistical uncertainties of the data recorded with the control trigger. The resulting trigger efficiency is shown in Fig. 5.8 [80].

[1] This corresponds to a measure proportional to the deposited charge in the pixel cell.

5.5.1 Systematic uncertainty

Systematic uncertainties on the trigger efficiency are calculated by varying the track selection criteria and taking into account the non-uniform response of the A and C sides of the MBTS detector. The variations on the track selection are as follows:

1. The constraint on the transverse impact parameter is removed, increasing the number of accepted tracks. This causes migration of events to higher $n_{\text{sel}}^{\text{no-z}}$ bins, leading to a decrease of trigger efficiency at low multiplicities.

2. For tracks in events with a primary vertex, a cut on the z coordinate with respect to the primary vertex is applied. This reduces the number of accepted tracks causing a migration of events to lower $n_{\text{sel}}^{\text{no-z}}$ bins.

The difference in efficiency obtained for the modified definition of $n_{\text{sel}}^{\text{no-z}}$ and the default value is less than 0.1%. The variation provides a conservative estimate of the effect of beam-induced background and non-primary tracks on the trigger efficiency at low values of $n_{\text{sel}}^{\text{no-z}}$. Another source of systematic uncertainties is the non-uniform response of the MBTS detector. The difference between the efficiency is calculated separately for events triggered by the MBTS A and C side and it reaches 1% for $n_{\text{sel}}^{\text{no-z}} = 1$. To translate this into a systematic error on the L1_MBTS_1 trigger efficiency it has to be multiplied by 1-$\epsilon_{\text{trig,oneside}}$, where $\epsilon_{\text{trig,oneside}}$ denotes the efficiency of the MBTS A (or MBTS C) trigger. For the $n_{\text{sel}}^{\text{no-z}}$ =1 bin the systematic error from this source is thus only about 0.12%.

The two sources of systematic uncertainties (track selection and non-uniform response) are combined in quadrature. The total systematic uncertainty on the trigger efficiency determination is less than 0.15% in the first $n_{\text{sel}}^{\text{no-z}}$ bin, decreasing rapidly as $n_{\text{sel}}^{\text{no-z}}$ increases.

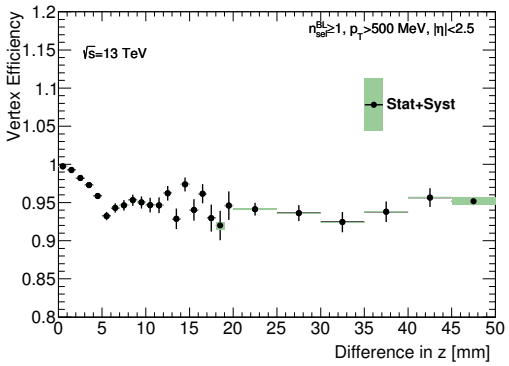
Additionally, the effect of events failing both the physics and control triggers for the trigger-efficiency determination is studied using the Pythia 8 MC samples and the effect on the trigger efficiency is found to be negligible: less than 1×10^{-9} for the $n_{\text{sel}}^{\text{no-z}} = 1$ bin and less than 1×10^{-10} integrated across all $n_{\text{sel}}^{\text{no-z}}$ bins.

5.6 Vertex Reconstruction

The vertex reconstruction efficiency is calculated differently in data and simulation. For data, the systematic uncertainties are dominated by the subtraction of non-collision beam background, and the measurement of this background is discussed in Section 5.6.1. The primary vertex algorithm [104] is unchanged compared to the vertex finder used in previous measurements. A track used in vertex reconstruction is defined as follows:

- The track transverse momentum is above 100 MeV;
- At least 1 Pixel hit;
- At least 4 SCT hits;
- At least 6 Silicon hits;
- Transverse and longitudinal impact parameters as well as their uncertainties calculated with respect to the beam spot (BS): $|d_0^{\text{BS}}| < 4\,\text{mm}$, $|z_0^{\text{BS}}| < 1000\,\text{mm}$, $\sigma_{d_0}^{\text{BS}} < 5\,\text{mm}$, and $\sigma_{z_0}^{\text{BS}} < 10\,\text{mm}$.

The vertex reconstruction efficiency, ϵ_{vtx}, is determined from data by taking the ratio of triggered events with a reconstructed vertex to the total number of triggered events as



(a)

FIGURE 5.9: Vertex efficiency in data with respect to the Δz between reconstructed tracks in events with $n_{\text{sel}}^{\text{no-z}} = 2$. The statistical errors are shown as black lines, the total errors as green shaded areas. The plot shows the efficiency in data at 13 TeV. The systematic errors are from the subtraction of the non-collision beam background.

defined in Equation 5.1. The non-collision beam background, which mostly affects the denominator, is subtracted.

$$\epsilon_{\text{vtx}} = \frac{\text{events passing MBTS\_1 with } n_{\text{sel}}^{\text{no-z}} \geq 1 \text{ and having a reconstructed vertex}}{\text{events passing MBTS\_1 with } n_{\text{sel}}^{\text{no-z}} \geq 1} \tag{5.1}$$

The subtraction of non-collision beam background, a source of systematic uncertainty, is discussed in Section 5.6.1. The dependence of the vertex reconstruction efficiency on η and $n_{\text{sel}}^{\text{no-z}}$ of the selected tracks is studied as well as the dependence on the projection along the beam-axis of the separation between the perigees[2] of the tracks (Δz), for events with more than one track. The η dependence is corrected for events with one reconstructed track. For events with higher multiplicities the Δz dependence is found to be small as shown in Figure 5.9 for $n_{\text{sel}}^{\text{no-z}} = 2$, and it is neglected. Events with $n_{\text{sel}}^{\text{no-z}} > 1$ have an efficiency very near 100%, so additional corrections are not needed.

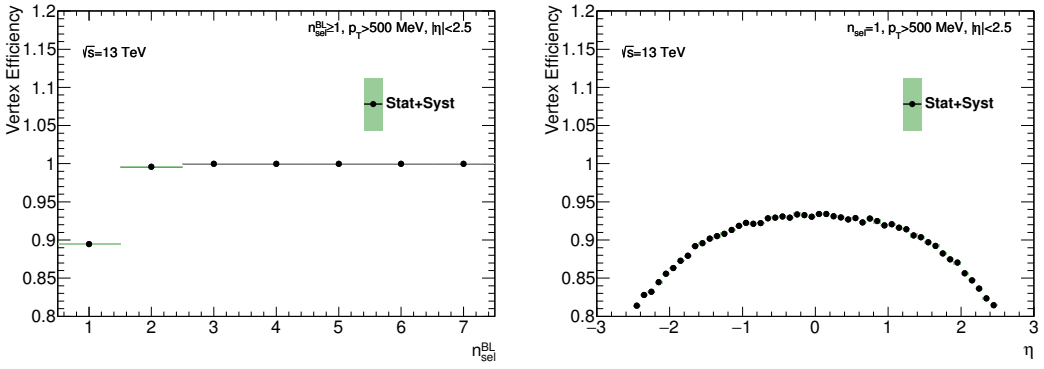
The efficiency, as shown in Figure 5.10, is measured to be $\sim$ 89% for events with $n_{\text{sel}}^{\text{no-z}} = 1$.

It rapidly rises to 100% at higher track multiplicities. The η dependence in $n_{\text{sel}}^{\text{no-z}} = 1$ varies from 81-93% for large $|\eta|$ to small values. Because the vertex efficiency is data-driven the result is independent of the Monte Carlo description of the physics events and the detector. A study of how ϵ_{vtx} would change with different material distributions in the detector and a study of the sensitivity to the different physics processes in Minimum Bias events was also performed. It was found that, while the vertexing efficiency is not suspectible to the material budget for non-diffractive samples, the difference between non-diffractive and single-diffractive or double-diffractive is about 10-15% for $n_{\text{sel}}^{\text{no-z}}$ =1. A difference of up to 10% for $n_{\text{sel}}^{\text{no-z}}$ =1 is also seen when comparing the two PYTHIA 8 tunes with HERWIG++ .

5.6.1 Systematic Uncertainties

The non-collision beam (NCB) background comes from beam halo and other beam phenomena, which can be reconstructed as real collisions within the ATLAS detector. The

[2]The perigee of a track is defined here as the point of closest approach of the track and the coordinate origin (0,0,0).



(a) (b)

FIGURE 5.10: Vertex efficiency in data with respect to the event selection as a function of the number of reconstructed tracks (a) and as function of η for exactly one selected track (b). The statistical errors are shown as black lines and they are too small to be visible. The total errors are shown as green shaded areas and they are also very small. The plot shows the efficiency in data at 13 TeV for $n_{\text{sel}}^{\text{no-z}} \geq 1$, $p_{\text{T}} > 500$ MeV, $|\eta| < 2.5$. The systematic errors are from the subtraction of the non-collision beam background.

contribution of beam-related backgrounds to the sample selected both with and without a vertex requirement is estimated using non-colliding bunches as in previous minimum bias ATLAS publications. The difference in the vertex reconstruction efficiency with and without beam background removal is assigned as a systematic uncertainty. The beam background contribution is larger in the vertex efficiency measurement than in the analysis phase space because a reconstructed primary vertex is not required in the denominator.

The normalization of the beam background is computed with several methods. Since the beam background does not come from collisions but rather only from 1 proton beam, the process is expected to cause hits in one side of the detector first and then propagate to the other side. This results in a large time difference between A and C side hits in the MBTS. The MBTS trigger provides sub-25 ns timing resolution: the difference in the average A and C side hit times is shown in Figure 5.11 in data runs 267358-9. The large central peak around 0 mostly comes from collisions, but the smaller peaks away from 0 are caused by non-collision beam background. A trigger, which selects on events in which only one filled bunch is in the ATLAS detector ("UNPAIRED"), is used to extract a shape for the beam background. The normalization is then fitted using this shape. The normalization comes from events with hits on both A and C-sides and is applied to the UNPAIRED trigger, which has the same hit requirements as the analysis. The MBTS was fitted with the sum of 3 Gaussians + 2 more Gaussians for the NCB background peaks. The central peak distribution (sum of 3 Gaussians shown with a dotted red line) was subtracted, and the remaining contributions were divided by the number of observed events from the UNPAIRED trigger.

The UNPAIRED events are used to model the kinematic distributions ($n_{\text{sel}}^{\text{no-z}}$ and track-η) after being normalized to the above fit. Other estimates including using number of pixel hits that are not associated to tracks as well as other methods which do not require hits on both A and C-sides were also used to cross-check the normalization of the non-collision beam background.

The effect on the vertex efficiency from subtracting the non-collision beam background is shown in Figure 5.12, and the full subtraction is assigned as the systematic uncertainty on the vertex efficiency. The $n_{\text{sel}}^{\text{no-z}} = 1$ events with the 500 MeV track p_{T} threshold is affected by at most 0.1%.

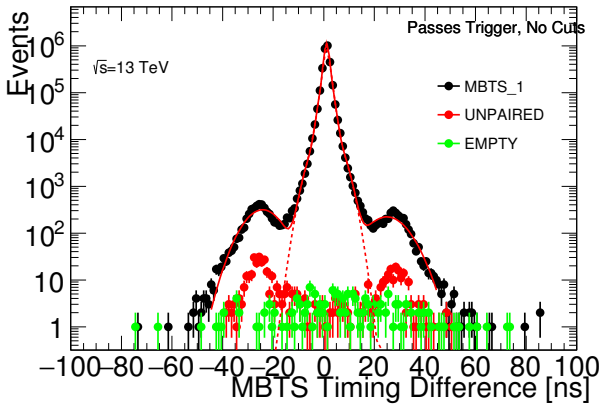
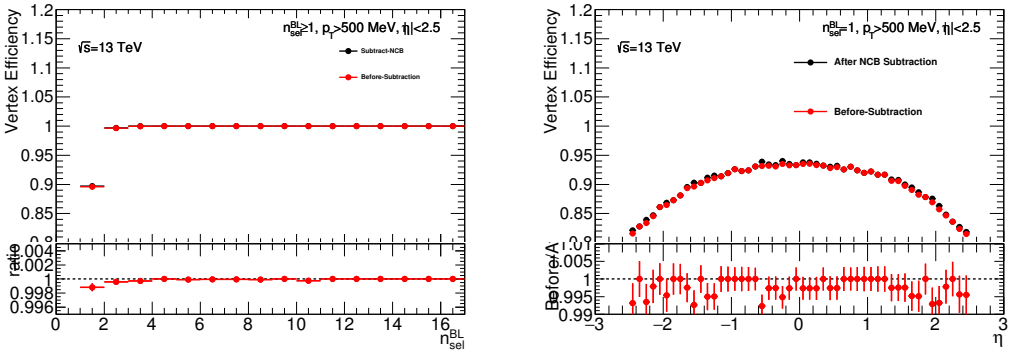


FIGURE 5.11: Time difference between A and C sides of the MBTS trigger with the HLT_MBTS_1 trigger in runs 267358-9.



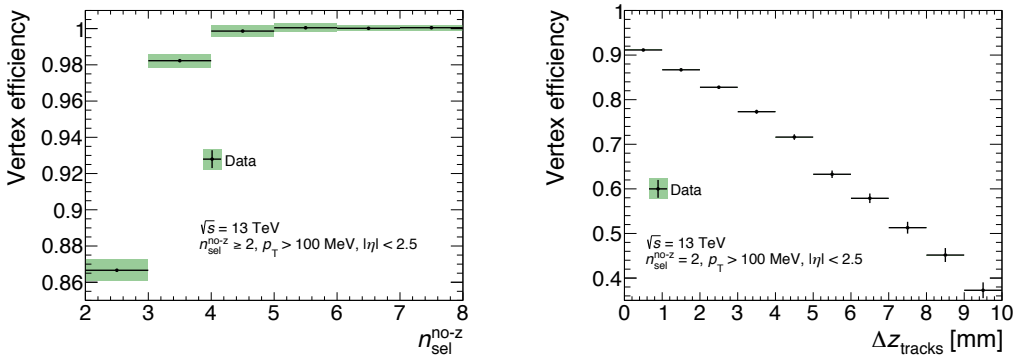
(a) (b)

FIGURE 5.12: The vertex efficiency for data in runs 267358-9 is shown for $n_{\text{sel}}^{\text{no-z}} \geq 1$ (left) and $n_{\text{sel}}^{\text{no-z}} = 1$ (right) before (red) and after (black) the NCB background subtraction.

This η dependence of the vertex reconstruction efficiency is taken into account in the correction procedure to derive the unfolded final results, which will be described in Section 5.10.

5.6.2 Vertex Efficiency in the Extended Phase Space

In contrast to the baseline analysis, in the case of the extended phase space, at least two tracks are required, as the p_{T} requirement of the analysis track selection is the same as in the selection used in the vertex reconstruction. In Figure 5.13 the vertex efficiency is shown as a function of the number of selected tracks (left) and as a function of Δz_{tracks} for exactly two selected tracks (right). The Δz_{tracks} is the distance between the two selected tracks in the event at the beamline. The efficiency in the first two bins is about 12% lower than for the 500 MeV phase-space. The vertex is reconstructed for tracks with a minimum transverse momentum of $p_{\text{T}} > 100$ MeV as in the baseline analysis. The vertex efficiency is then calculated as a function of the number of selected tracks $\text{n}_{\text{sel}}^{\text{no-z}}$, which means all tracks with all the selection cuts applied, except the z_0 cut. This means two tracks with a transverse momentum of $p_{\text{T}} > 100$ MeV in this analysis and one track with $p_{\text{T}} > 500$ MeV for the baseline analysis. This leads to a lower efficiency in the first bin, reaching a plateau around 100% for higher number of reconstructed tracks. Events with only two tracks are



(a) (b)

FIGURE 5.13: Vertex efficiency in data with respect to the event selection as a function of the number of reconstructed tracks (left) and as function of Δz_{tracks} for exactly two selected tracks (right). The statistical errors are shown as black lines, the total errors as green shaded areas, which are very small and only hard to see for the right plot. The plot shows the efficiency in data at 13 TeV for $n_{\text{sel}}^{\text{no-z}} \geq 2$, $p_{\text{T}} > 100$ MeV, $|\eta| < 2.5$. The systematic errors are from the subtraction of the non-collision beam background.

corrected using a Δz_{tracks} dependent correction because a dependence on the separation was observed for events with low track multiplicity. This means that in case of an event with only two tracks, the value for the vertex efficiency is computed in dependence of the Δz_{tracks} as shown in Figure 5.13(a). For events with more than two tracks, the vertex efficiency is computed in dependence of the number of selected track as shown in Figure 5.13(b).

In this case, it is the Δz_{tracks} dependence of the vertex reconstruction efficiency which needs to be taken into account in the correction procedure to derive the unfolded final results, as will be specified in Section 5.10.

5.7 Track Reconstruction Efficiency

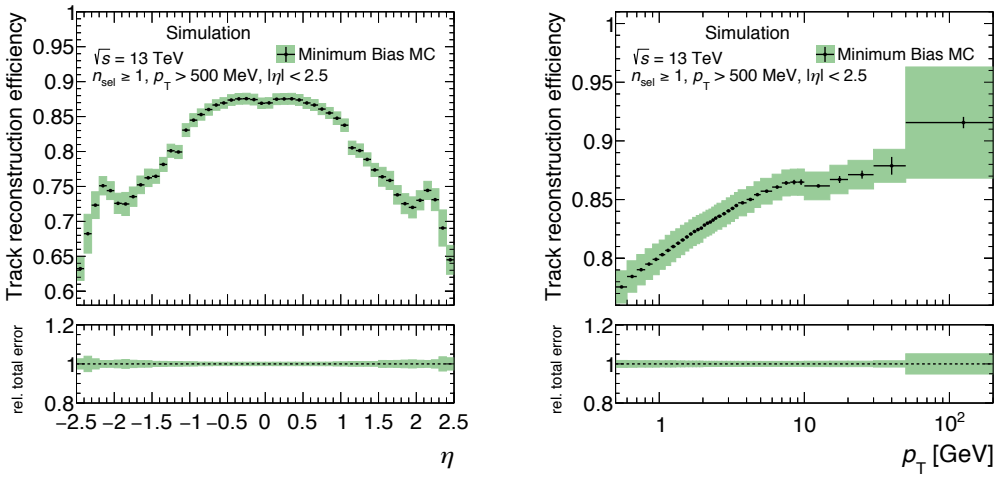
The track reconstruction efficiency, extensively described in Chapter 3, is one of the main ingredient in the charged-particle multiplicities measurements, because its systematic uncertainties represent the main source of uncertainty for this analysis.

The tracking efficiency for primary particles used in this analysis is based on simulation with a correction to take into account material description deficiencies. The systematic uncertainties on the track reconstruction efficiency are based primarily on the knowledge of the material in the detector as well as on the differences in data and simulation in the track selection and the alignment.

The tracking efficiency for primary particles is based on simulation using the primary particle definition and track-to-truth particle association described in Chapter 3 and in the above Section 5.3. The efficiency is measured in simulation as a function of p_{T} and η and it is defined, as already reported in Equation 3.2, as follows:

$$\epsilon_{\text{trk}}(p_{\text{T}}, \eta) = \frac{N_{\text{rec}}^{\text{matched}}(p_{\text{T}}, \eta)}{N_{\text{gen}}(p_{\text{T}}, \eta)}$$

where p_{T} and η are generated particle properties, $N_{\text{rec}}^{\text{matched}}(p_{\text{T}}, \eta)$ is the number of reconstructed tracks matched ($P_{\text{match}} > 0.5$) to a given generated charged-particle and $N_{\text{gen}}(p_{\text{T}}, \eta)$ is the number of generated charged-particles in that (p_{T}, η) bin. Only events



(a) (b)

FIGURE 5.14: The primary track efficiency as a function of (a) η and (b) p_{T} The green error bands show the total uncertainty.

which pass the event selection defined in Section 5.4 are used. It was tested that the measured efficiency is independent of the number of generated particles in a given event and is therefore independent of the trigger and vertex requirements. This allows to populate phase space regions with a low number of tracks in physics events using a single particle sample and to apply the selection cuts with respect to the generated vertex in these events.

The tracking efficiency for primary tracks after a data-driven correction, which will be discussed shortly, is shown in Figure 5.14 as a function of p_{T} and η. The lower track reconstruction efficiency in the region $|\eta| > 1$ is due to the increasing amounts of material that the particles must traverse. The slight increase in efficiency for $|\eta| > 2$ is due to particles passing through a larger number of sensitive layers.

5.7.1 Corrections to and Systematic uncertainties on the track reconstruction efficiency

As the track reconstruction efficiency is based on simulation with a correction to take into account material description deficiencies in the simulation, the systematic uncertainties result from the understanding of the compatibility of the detector geometry model in data and simulation.

The cuts applied in the track selection itself as well as the modelling of the passive material in the ID can systematically affect the track reconstruction efficiency. The various sources are summarized in Table 5.5 which contains the systematic uncertainties applied to the track reconstruction. An explanation of each element in the table is given in the following. All effects are assumed to be Gaussian and are added in quadrature.

In addition, Table 5.6 summarises the impact of other track fit quality related issues on the final result. These effects do not impact the track reconstruction efficiency but do change the reliability of the simulation to emulate the data. An explanation of each element in the table is also given in the following.

Systematic Uncertainty	Size	Region
Track Selection (Sec. 5.7.1)	0.5 %	flat in p_{T} and η
Material (Sec. 5.7.1)	0.6 % - 3 %	decreases with p_{T}, increases with $\lvert\eta\rvert$
χ^2 prob. cut efficiency (Sec. 5.7.1)	0.5% - 5%	only for $p_{\mathrm{T}} > 10$ GeV increases with p_{T}

TABLE 5.5: The systematic uncertainties on the track reconstruction efficiency. All uncertainties are quoted relative to the track reconstruction efficiency.

Systematic Uncertainty	Size	Region
χ^2 prob. cut remaining bad tracks (Sec. 5.7.1)	0.2 % - 7%	only for $p_{\mathrm{T}} > 10$ GeV increases with $\lvert\eta\rvert$ and p_{T}
Alignment and other high p_{T} (Sec. 5.7.1)	0.1% - 10%	only for $p_{\mathrm{T}} > 10$ GeV averaged over η, increases with p_{T}

TABLE 5.6: The systematic uncertainties on the final distributions due to track performance in data. All uncertainties are quoted relative to the total number of reconstructed or unfolded tracks.

Track Selection

Although the detector conditions used in MC simulation are described as precisely as reasonably possible, there are still small differences between data and simulation in the number of detector hits counted. To remove dependence on the kinematics of tracks in η and p_{T}, a two-dimensional re-weighting of the MC is used. The systematic uncertainty is estimated by comparing the efficiency of each cut in data and MC, defined as:

$$\epsilon_{\mathrm{cut}}(p_{\mathrm{T}}, \eta) = \frac{N^{\mathrm{tracks}}_{\mathrm{all\,cuts}}(p_{\mathrm{T}}, \eta)}{N^{\mathrm{tracks}}_{\mathrm{N-1\,cuts}}(p_{\mathrm{T}}, \eta)} \tag{5.2}$$

where $N^{\mathrm{tracks}}_{\mathrm{all\,cuts}}(p_{\mathrm{T}}, \eta)$ is the number of reconstructed tracks with all cuts applied and $N^{\mathrm{tracks}}_{\mathrm{N-1\,cuts}}(p_{\mathrm{T}}, \eta)$ is the number of tracks after a certain hit requirement was removed. The cut efficiency for primary tracks in data and MC is shown in Figure 5.15 as a function of η and p_{T} for two cases: all Pixel hit requirements (Figures 5.15(a) and 5.15(b)) and all SCT hit requirements removed (Figures 5.15(c) and 5.15(d)). Large differences are observed at high p_{T} for the efficiencies of both cuts: this is results of a high fraction ($> 10\%$) of poorly measured tracks entering the denominator when loosening the cuts. The uncertainty associated to mis-measured tracks are detailed in the χ^2 probability cut Section.

The track selection uncertainty is found to be at most 0.5% for the range $\lvert\eta\rvert < 2.5$ and the entire p_{T} range as shown in figure 5.16.

Material

As extensively detailed in Chapter 4, the track reconstruction efficiency is heavily dependent on the material distribution in the detector since hadrons have a sizable probability of interacting with the material. Studies during Run 1 indicate that the material in the entire Inner Detector is known within an accuracy of 5% or better. This limit is predominantly based on detailed surveys during the construction of the Inner Detector [52] in addition to a number of studies which were performed to verify the material description in simulation [172]. Thus, 5% uncertainty on the material in the entire Inner Detector is assessed by using a simulated sample with the overall ID material uniformly scaled up by 5%. The corresponding systematic uncertainty on the track reconstruction efficiency is shown in Figure 5.17.

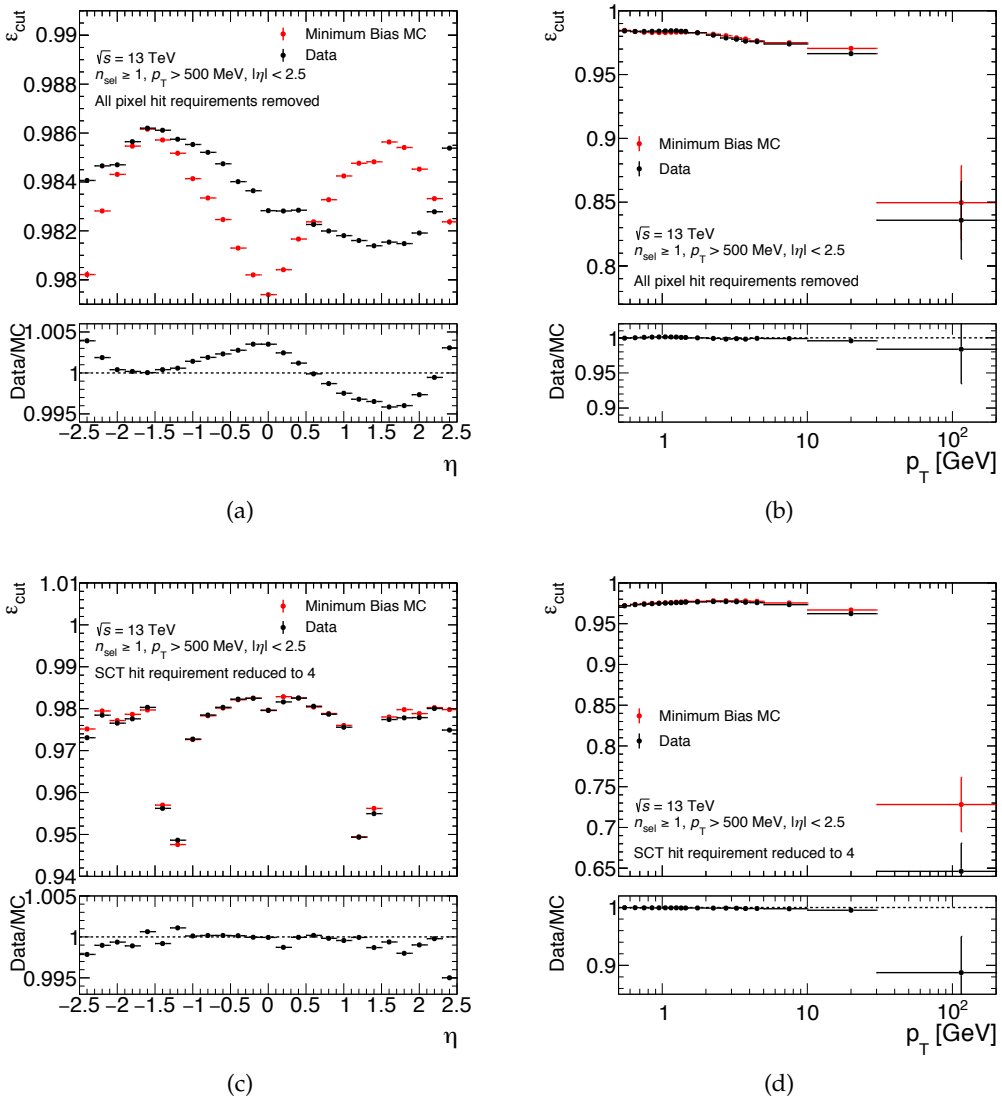


FIGURE 5.15: Cut selection efficiency for data and MC for the Pixel detector (top) and the SCT (bottom) as a function of η (left) and p_{T} (right). ϵ_{cut} with all Pixel hit requirements removed is shown in the top row and the effect of removing all SCT hit requirements is shown in the bottom row. The ratio between data and MC is used to determine the uncertainty due to the track selection.

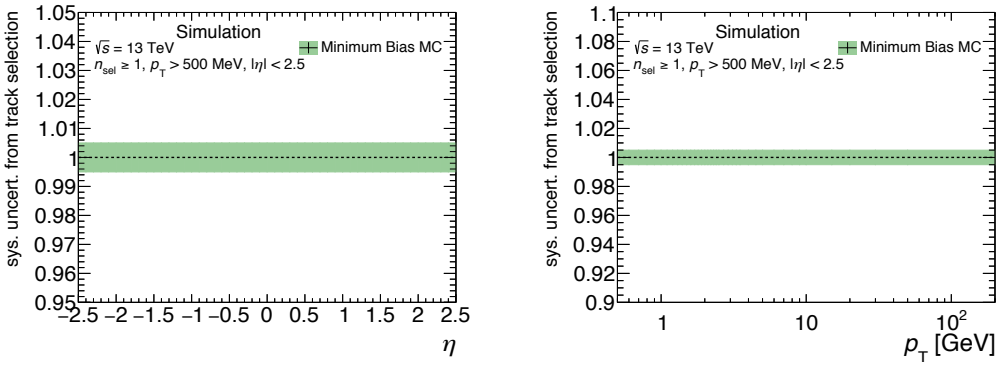
For what concerns the passive material localised in specific ID regions, the related systematic uncertainty on the track reconstruction efficiency were estimated by using the three data-driven methods described in Chapter 4:

1. Hadronic interactions and photon conversions in the barrel pixel region;

2. Track-Extension Efficiency in the end-cap pixel service region.

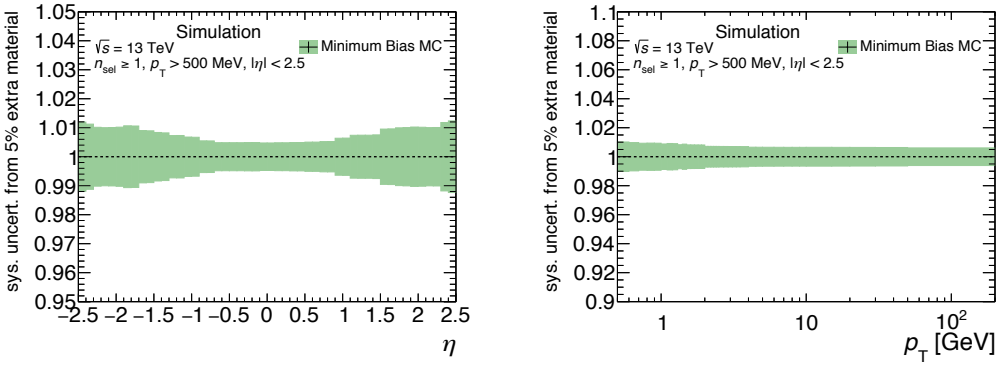
The details of these methods have been already given in Chapter 4, thus here only the information relevant to the charged-particle multiplicity will be reported.

Studies of the rate of hadronic interactions (Section 4.4.1) and photon conversions (Section 4.4.2) showed missing material in the IBL as modelled in the *original* geometry. In fact, when normalised to the rate of secondary vertices in the beam pipe, the observed rate of



(a) (b)

FIGURE 5.16: Relative systematic uncertainty from track selection on the primary tracking efficiency as a function of (a) η and (b) p_{T}.



(a) (b)

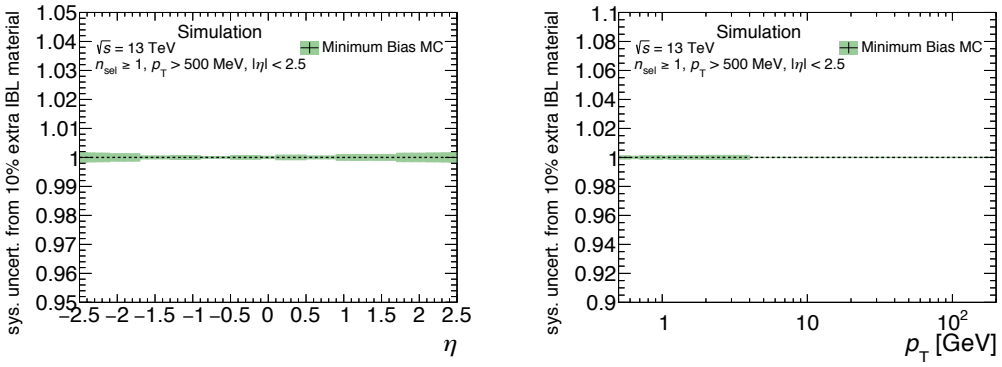
FIGURE 5.17: Systematic uncertainty due to 5% extra material in the inner detector on the primary tracking efficiency as a function of (a) η and (b) p_{T}.

hadronic interactions in the IBL region was approximately 30% higher (50% higher in the low mass IPT region) in data than in simulation and the rate of photon conversions in the IPT and IBL region was approximately 40% higher than what was seen in data. After correcting the description of the IBL, IST and IPT in the *updated* geometry, simulation agreed with both measurements to the level of about 10% (See Figure 4.11).

The charged-particle multiplicity was performed by using as a baseline the *updated* geometry and this allowed for a reduction of the systematic uncertainty. To be conservative and to cover the remaining discrepancies between data and simulation, the uncertainty on the tracking efficiency introduced by the mis-modelling of the IBL is estimated by using the IBL +10% sample (passive material in the IBL scaled up by 10%, corresponding to an increase of 7.5% and 8% of the total radiation and interaction length of the IBL material respectively). This introduces an additional 0.2% uncertainty on the tracking efficiency, Figure 5.18 shows the p_{T} and η dependence of this. This uncertainty is added linearly to the other efficiency uncertainties.

For what concerns the pixel services region, the track-extension efficiency method [3] shows differences between data and simulation of up to -3%, which can roughly identify a deficit in the pixel service material description of about 10%, with a strong dependence on η.

[3] In this case, the z_{vtx} cut described in Chapter 4 was removed in order to be consistent with the event selection used in the charged-particle multiplicity analysis.



(a) (b)

FIGURE 5.18: Systematic uncertainty due to 10% extra material in the IBL, on the primary tracking efficiency as a function of (a) η and (b) p_{T}.

As explained in Sections 4.5.11 and 4.5.12, such a difference in track-extension efficiency was first used to estimate an effective amount of missing material, in number of interaction lengths, in each pseudorapidity bin, as in Equation 4.15:

$$\Delta\mathcal{E}^{\mathrm{Data-MC}}_{\mathrm{norm\ ext}} \approx -\Delta N^{\mathrm{Data-MC}}_{\lambda_I}.$$

Then, by using this effective amount of material, a data-driven correction to the track reconstruction efficiency ($\Delta\epsilon^{PixServCorrection}_{\mathrm{trk}}$) can be estimated by normalising the change in the track reconstruction efficiency when comparing the nominal sample with one of the artificially modified samples ($\Delta\epsilon^{\mathrm{mod-nom}}_{\mathrm{trk}}$):

$$\Delta\epsilon^{PixServCorrection}_{\mathrm{trk}} = \frac{\Delta\epsilon^{\mathrm{mod-nom}}_{\mathrm{trk}}}{\Delta N^{\mathrm{mod-nom}}_{\lambda_I}} \cdot \Delta N^{\mathrm{Data-MC}}_{\lambda_I}. \tag{5.3}$$

In this case, the pixel service +50% sample was choosen. With an analogous procedure, the uncertainties related to the track-extension efficiency can be translated into uncertainties on the data-driven correction. The results are shown in Figure 5.19. The dominant uncertainty on the track-reconstruction efficiency correction is the uncertainty on the particle composition used to make the measurement.

Summarising, the total uncertainty introduced by the imprecise knowledge of the material in the detector is taken to be the sum of:

- 5% uncertainty on the material in the entire inner detector
- 10% uncertainty on the IBL material
- η dependent uncertainty on the correction to the track reconstruction efficiency in the PP0 region ($|\eta| > 1.5$).

To be conservative, the uncertainties of the entire detector and IBL are added linearly since the effect on the tracking efficiency is additive. The uncertainty of the correction to the track reconstruction efficiency is added in quadrature as this is an already assigned uncertainty on the SCT extension efficiency. This leads to a total systematic uncertainty due to material of 0.6-3%, which decreases with p_{T} and increases with $|\eta|$.

These values are larger than those reported in Ref. [80], where the maximum systematic uncertainty due to material was below 2%. In fact, after a thorough examination of the track-extension efficiency, it was found that the value of the correction was overestimated due to an issue in the calibration procedure, while the magnitude of the uncertainty was underestimated:

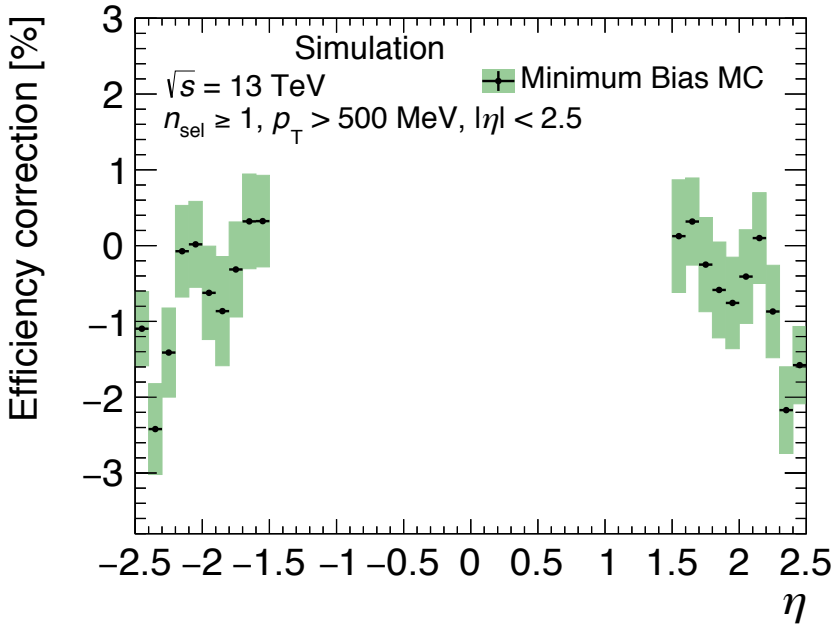


FIGURE 5.19: The tracking efficiency correction based on the SCT efficiency measurement. The uncertainties shown includes all systematic uncertainties considered and the statistical uncertainty.

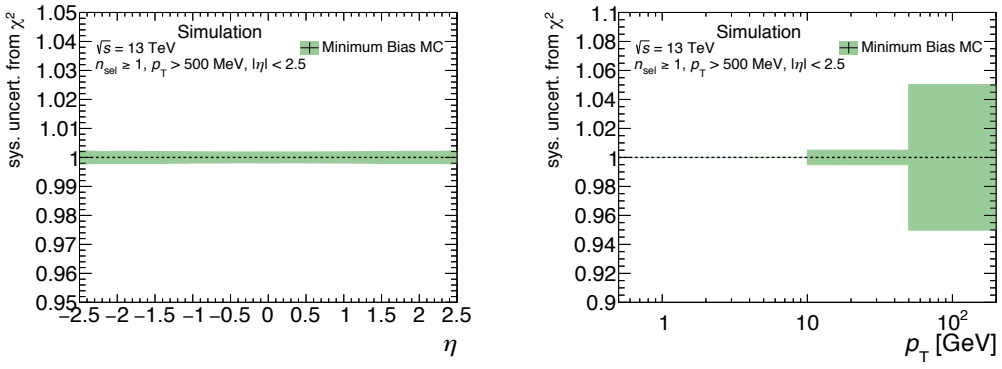
- It was assumed that in the samples in which the material was scaled up by a certain amount in the pixel service region, there was no additional material elsewhere. After a careful exploration of the geometry database, it was instead found that also the material in the IBL was scaled up.
- The dependency of the track-extension efficiency on the radial position of the material was missing and only assessed in the last iterations. The tighter 4 SCT hit requirement allowed to reduce this systematic uncertainty, which would be larger if the loose 2 SCT hit requirement was used, as previously shown in Figure 4.36.

In the final unfolding procedure, the data-driven correction has a sizable impact only on the charged-particle distribution as a function of η. Thus, this distribution has been updated to account for the new correction and it will be shown in Section 5.11. .

χ^2 probability cut

Badly measured low momentum charged-particles are sometimes reconstructed as high momentum tracks. While the fraction of these particles is very low at low p_{T}, these tracks constitute a sizable fraction at high reconstructed p_{T} because of the steeply falling p_{T} distribution. Such badly measured tracks are caused by interactions and multiple scattering with the material and usually have a bad χ^2 fit probability. To suppress them, a cut on the χ^2 probability of $P\left(\chi^2, n_{dof}\right) > 0.01$ is applied for tracks with $p_T > 10\,\mathrm{GeV}$.

Two methods were used to estimate the uncertainty of the fraction of mis-measured tracks in data and simulation. Excellent agreement between data and simulation was found for both methods. Details will not be given here, but they can be found in Ref. [80] and the supporting documentation. The uncertainty on the remaining amount of mis-measured tracks has been determined to be less than 0.2% at 10 GeV increasing to 7% for p_{T} above 50 GeV. The uncertainty in the efficiency of the cut is assessed to be to 0.5% below 50 GeV and 5% above 50 GeV (as shown in figure 5.20). These uncertainties are added in quadrature, as double sided uncertainties, to other sources.



(a) (b)

FIGURE 5.20: Relative systematic uncertainty from χ^2 probability cut on the primary tracking efficiency integrated over p_{T} (left) and η (right).

Alignment

The charged-particle multiplicity analysis is a challenging track-based analysis because it is performed in the very first phases of data taking, when effects from e.g. a bad description of the material in simulation or detector alignment are not yet optimised can have a very big impact. There is, indeed, the possibility that the alignment of the ID is suboptimal due to limited data and time needed to refine the alignment. The detector alignment is performed by minimising a χ^2 of the track-to-hit residuals with respect to the alignment parameters. Thus, detector misalignments can be divided into two general categories:

- χ^2 invariant deformations;
- deformations that impact the fit quality.

The former, the so-called "weak modes", as already mentioned in Chapter 3, can not be assessed with limited amounts of data. As such, a conservative estimate based on the weak modes initially calculated in Run 1 is used in this analysis. Detector misalignment that affects the track fit quality are largely removed with early Run 2 data, and thus remaining misalignments can be quantified by comparing the track-hit residuals in data and MC simulation.

The misalignment of the detector has been shown to have no measurable effect on the track reconstruction efficiency or the track acceptance in the selection. The largest impact of ID misalignments is on the measured momentum of the reconstructed tracks. Random detector misalignments cause a smearing of the momentum resolution, while systematic misalignments will bias the reconstructed momentum. Both categories impact the measured track momentum distribution. Four effects are considered to estimate the systematic uncertainty due to detector alignment:

- Random misalignment of detector modules;
- temperature dependent IBL bowing;
- p_{T} dependent sagitta bias;
- p_{T} dependent sagitta smearing.

More details on this can be found in Refs. [80, 107] and supporting documentations. These modified p_{T} spectra are used as input in the unfolding analysis to determine their impact on the final unfolded p_{T} distribution, as will be described in Section 5.10.2).

Figure 5.21 shows the unfolded p_{T} distributions for all considered alignment systematics compared to the nominal distribution. The effect of misalignment is negligible for tracks with a p_{T} below 10 GeV. Above 10 GeV, the effect due to random module misalignment is dominant and rises up to 10% for tracks with a p_{T} of 100 GeV. In addition, these results confirm that the uncertainty due to the bowing of the IBL caused by an absolute change of 1 degree Celsius in the IBL operating temperature[4] has no impact on this analysis. A random misalignment, which is an approximation for the worse case alignment scenario, should only have an impact on the analysis in the highest p_{T} region.

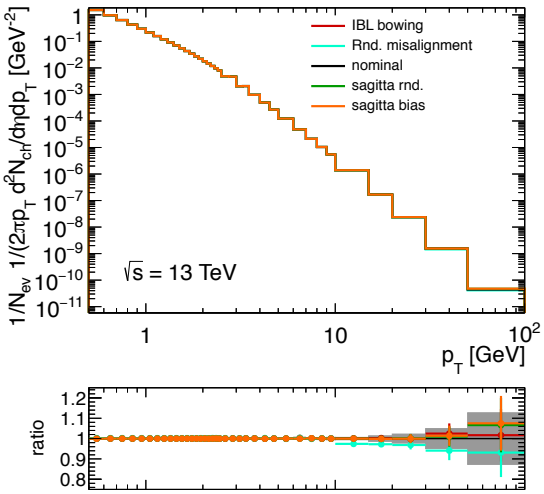


FIGURE 5.21: Unfolded p_{T} distributions for the four different alignment systematics (colored). The differences in the unfolded distributions are mostly within the statistical uncertainty of the method (grey band).

5.7.2 Track Reconstruction Efficiency in the Extended Phase Space

The tracking efficiency for the selected primary tracks at low p_{T} is shown in Fig 5.22 as function of p_{T} and η. For $100 < p_{\mathrm{T}} < 125$ MeV and integrated over η, the primary-track reconstruction efficiency is 27.5%.

The systematic uncertainties on the track reconstruction were calculated according to the baseline analysis, except for the material-related uncertainty.

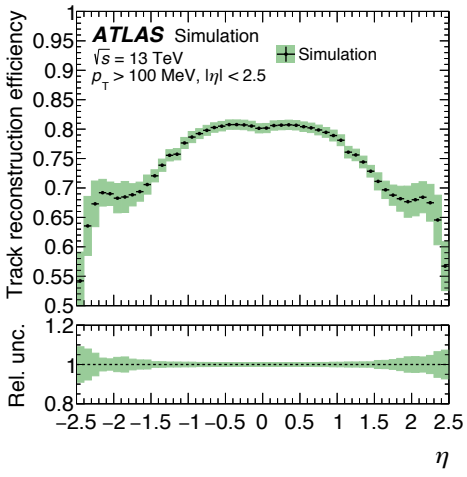
TABLE 5.7: The systematic uncertainties on the track reconstruction efficiency. All uncertainties are quoted relative to the track reconstruction efficiency.

Systematic Uncertainty	Size	Region
Track Selection	0.5%	flat in p_{T} and η
χ^2 probability	0.2-7%	increasing with p_{T}
Material	1-9%	decreases with p_{T}; increases with $\lvert\eta\rvert$

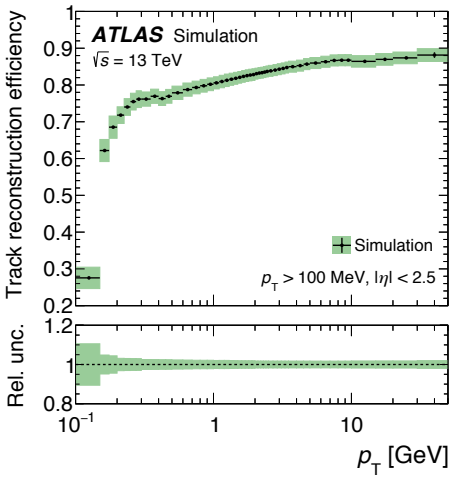
Material correction and systematic uncertainty

The uncertainty due to imprecise knowledge of the material in the inner detector is taken as the sum of a 5% uncertainty on the material in the entire inner detector, an additional

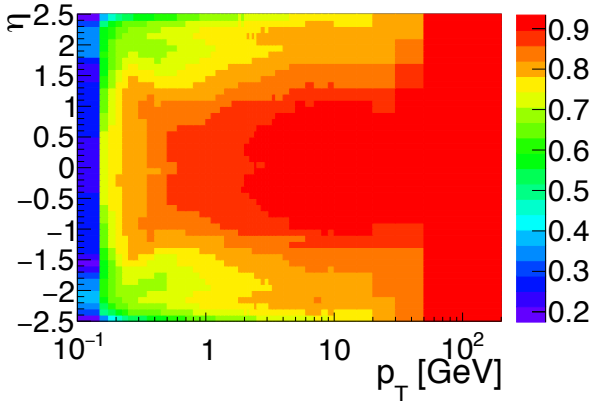
[4]The temperature was stable to better than 0.1 degree Celsius during the data collection period



(a) Ref. [81].



(b) Ref. [81].



(c)

FIGURE 5.22: The primary track efficiency as a function of (a) η and (b) p_{T} and as a function of (c) p_{T} and η. The green error bands includes the total systematic and statistical uncertainty.

10% uncertainty on the IBL material and a 50% uncertainty coming from the pixel service region ($|\eta| > 1.5$).

The various components of the tracking efficiency systematics can be seen in Fig. 5.25. The total systematic uncertainty ranges from 1 to 9%, decreasing with p_{T} and increasing with $|\eta|$.

As it was done in the baseline analysis, the components due to material in the entire inner detector and IBL are added linearly, but in in this case, the data-driven correction derived from the SCT extension efficiency was not applied. The calculation of this correction was restricted to tracks with $p_{\mathrm{T}} > 500$ MeV and only extrapolated to low p_{T} tracks. In the chosen regime, every source of uncertainty affecting the method is under control, while it was observed that the SCT extension efficiency differs strongly in some eta bins, such as $|\eta| = 1.75$, when lowering the p_{T} threshold. Many effects can contribute to this phenomenon, e.g. an increased rate of the secondaries.

In addition, other effects such as track bending become relevant below 500 MeV, and track interactions with the material are more significant and more difficult to control by using an ad hoc correction.

Nonetheless, the results obtained after the application of the data-driven correction

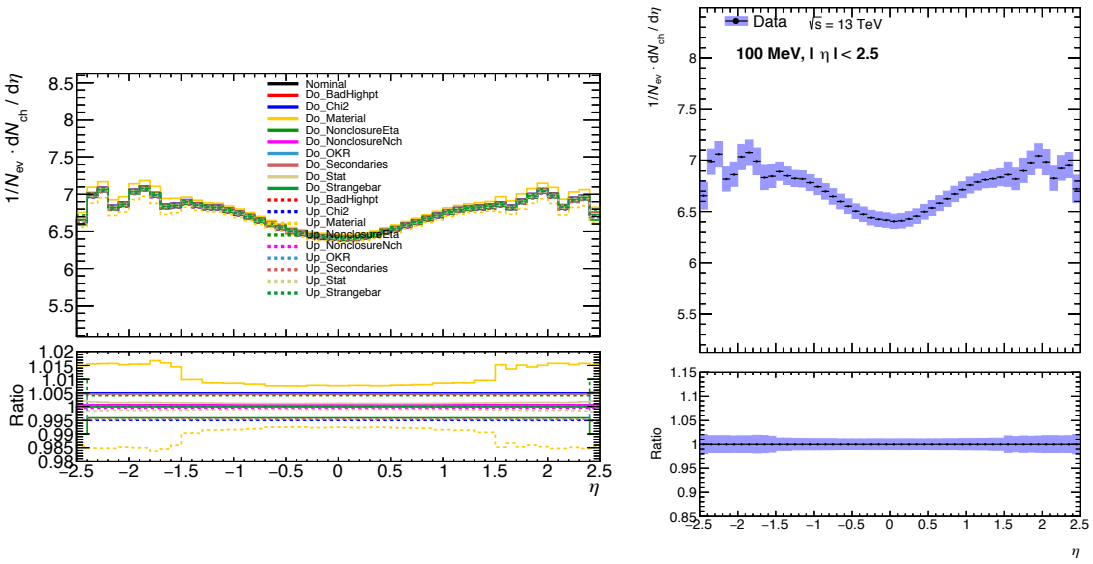


FIGURE 5.23: Breakdown of all different systematics for the η distribution (left). The corrected data with the total systematic uncertainty (right).

have also been derived, for comparison and to illustrate the issues observed. Only the η plots are shown and discussed in the following as there was no visible effect for the other distributions.

Figure 5.23 shows the breakdown of the systematic uncertainties and the corrected data. The systematic uncertainty on the data-driven correction is derived in the same way as in the baseline analysis. For low p_{T} tracks this procedure would need to be reassessed, likely resulting in a larger uncertainty. The corrected data distribution shows significant fluctuations for $|\eta| > 1.5$. This is exactly the region where the track-extension efficiency is relevant.

Several studies were done in order to understand these fluctuations. A possible bin-to-bin migration between truth and reconstructed η was taken into account. The changes when taking this migration into account were below 0.05% and only slightly affected the η distribution. The location of the material is not well known, therefore an additional correction was defined which considers the change in η depending on the material location of the detector. The changes in the results were again below 0.05% and did not change the η distribution.

The η distribution after applying the data-driven correction is shown in Figure 5.24. The fluctuations already described are visible in this plot.

Measured with tracks above 500 MeV, the correction is only extrapolated to the low p_{T} phase-space. The amount of secondary particles is not well predicted in the Monte Carlo event generators (as it will be discussed in Section 5.8, the secondary rate differs by about 40% between data and simulation) and may affect low p_{T} tracks more deeply. Due to these issues, it was finally decided to not use the data-driven correction for the low momentum phase-space. Therefore the 50% material variation for the pixel services was used in order to derive a simple systematic uncertainty which is in agreement with the tracking recommendations presented in Chapter 3.

Significant work would be required to obtain a reliable track-extension efficiency and the related correction in the low momentum range.

It must be noted that the application tests of the data-driven correction down to $p_{\mathrm{T}} =$ 100 MeV presented here were performed before updating the correction as shown earlier

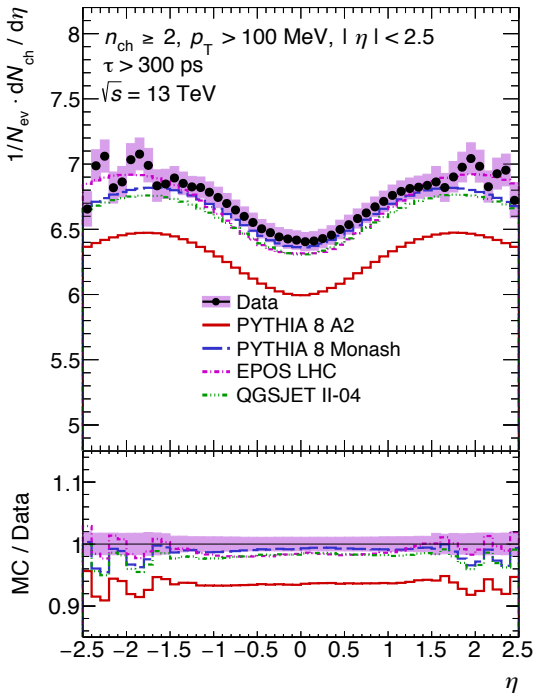


FIGURE 5.24: Pseudorapidity dependence of the charged-particles after applying a data-driven correction to the track reconstruction efficiency.

in this chapter. The improved correction, which is characterised by lower magnitude and larger uncertainties, might allow for a better result with smaller fluctuations. This work is left for future studies.

Total systematic uncertainty

Figure 5.26 shows the total systematic uncertainty on the primary tracking efficiency as function of η and p_{T}. The size of the systematic uncertainty is also shown in Table 5.8 for some specific p_{T} and η bins. The uncertainties are larger for the low p_{T} analysis due to the different approach taken to evaluate the sensitivity to the material modelling.

TABLE 5.8: The systematic uncertainty on the track reconstruction efficiency for some specific p_{T} and η bins.

p_{T} bin range [GeV]	η bin range	size of the systematic uncertainty [%]
0.1 – 0.15	0 – 0.1	2.77
0.1 – 0.15	2.4 – 2.5	3.30
0.3 – 0.35	0 – 0.1	1.13
0.3 – 0.35	2.4 – 2.5	4.34
0.5 – 0.6	0 – 0.1	0.77
0.5 – 0.6	2.4 – 2.5	4.81

5.8 Backgrounds

Two different sources of background have to be considered, one coming from non-collision events such as proton interactions with residual gas molecules in the beam pile (beam

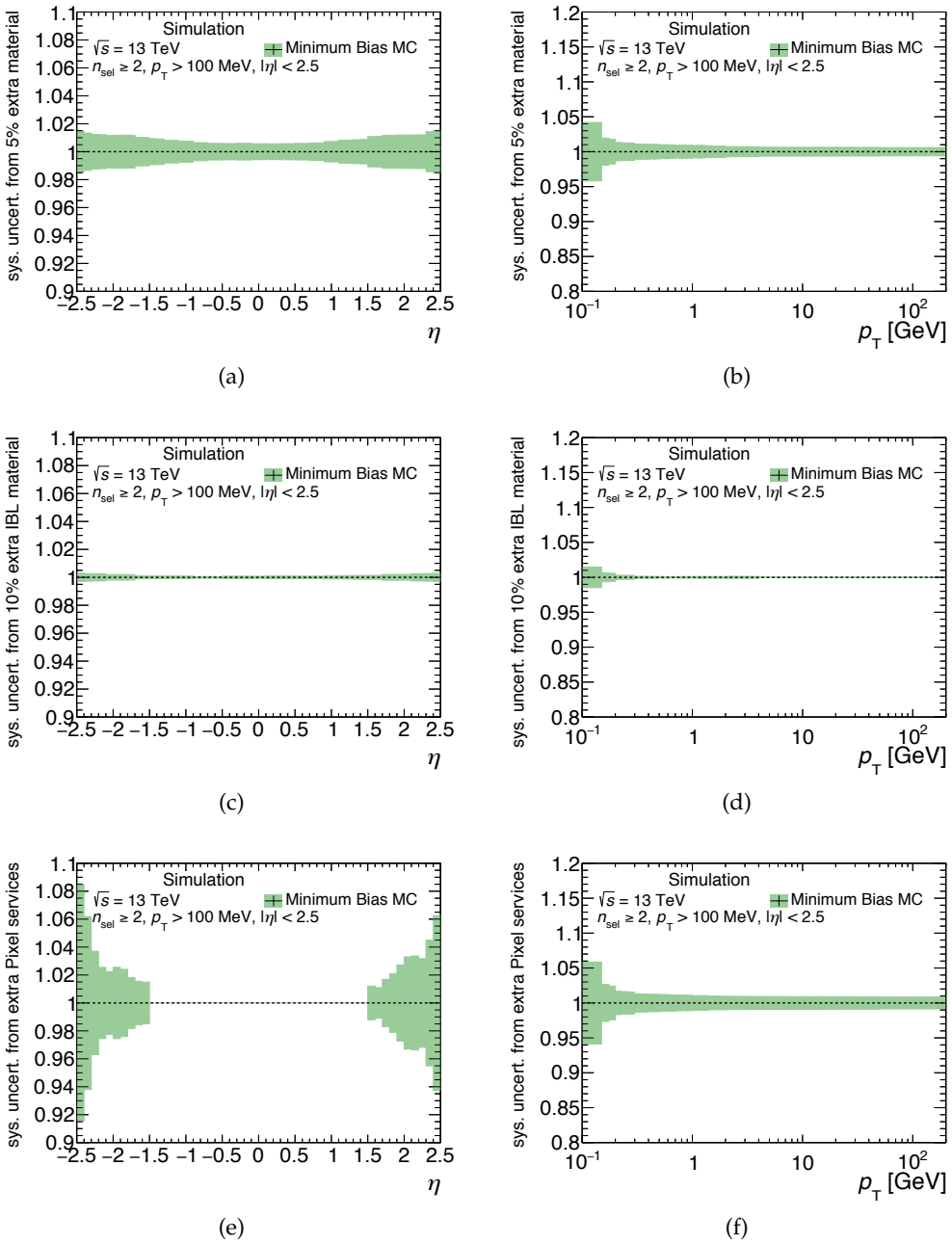
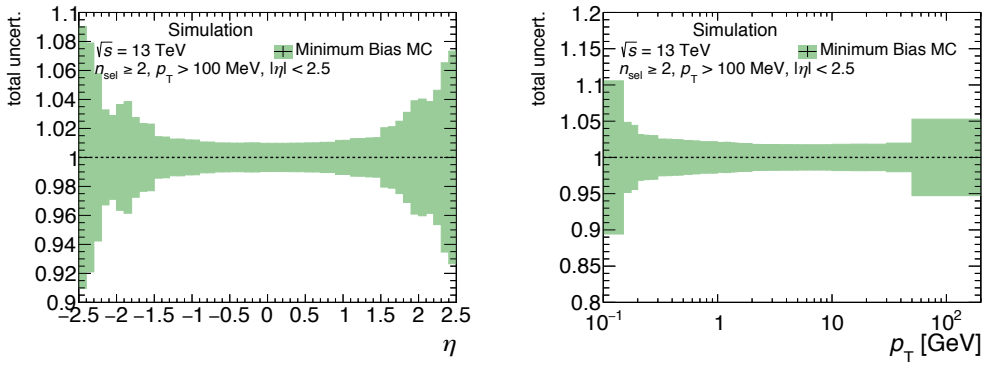


FIGURE 5.25: Relative uncertainty due to material in the inner detector on the primary tracking efficiency integrated over p_{T} (on the left) and η (on the right). (a) and (b) show the systematic uncertainty from 5% extra material; (c) and (d) show the systematic uncertainty from 10% extra material in the IBL; (e) and (f) show the systematic uncertainty from extra material in the patch panel region $|\eta|$ >1.5

pipe) and cosmic ray interactions and one deriving from non-primary tracks, such as decay and hadronic interaction products.

The beam halo contribution is evaluated by considering events which pass the full event selection but occur when only one of the two beams is present. After normalising to the contribution expected in the selected data sample (using the difference in the time of the MBTS hits on each side of the detector) a contribution of less than 0.01% of events is found from this source, which is considered negligible.



(a) (b)

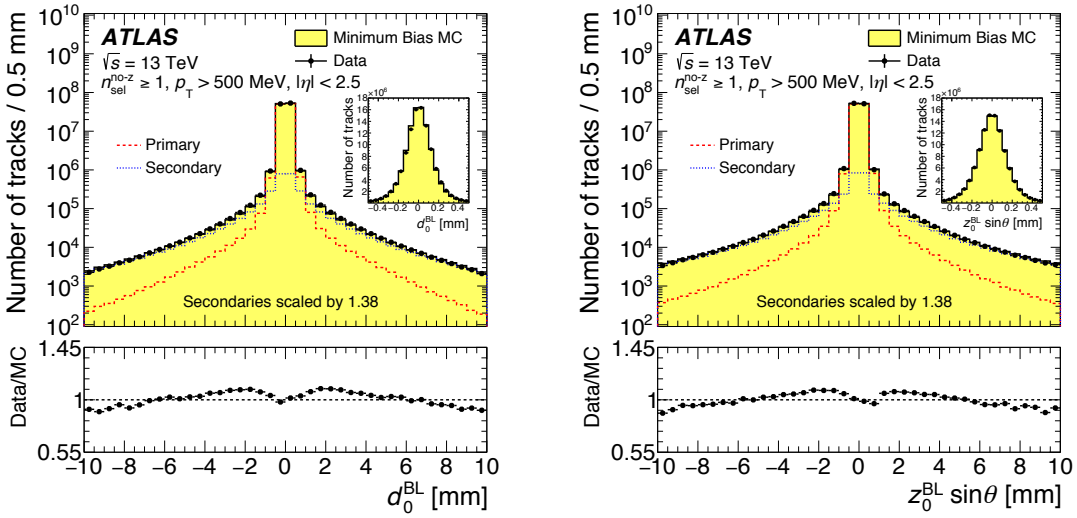
FIGURE 5.26: Relative uncertainty on the primary tracking efficiency integrated over p_{T} and η. The statistical uncertainty is included here, but it is negligible.

The expected rate of cosmic ray events (less than 0.001 Hz) is irrelevant with respect to the event readout rate (2 kHz): this source of background is therefore neglected.

Events with more than one interaction in the same bunch-crossing (pile-up events) must be rejected. This is accomplished by vetoing on any event with more than one primary vertex. Some events may survive because the interactions are very close in z and are merged together. The probability to merge vertices is estimated by inspecting the distribution of the difference in the z position of pairs of vertices (Δz). This distribution displays a deficit around $\Delta z = 0$ due to vertex merging. The magnitude of this effect is used to estimate the probability of merging vertices, to be 3.2%. When this is combined with the number of expected additional interactions for $\langle \mu \rangle = 0.005$ (which is typical of the data-set used in this analysis), the remaining contribution from tracks from additional interactions is found to be less than 0.01%, which is negligible. The additional tracks in events in which the second vertex has fewer than four associated tracks are mostly rejected by the $|\Delta(z_0 \cdot \sin\theta)|$ requirement, and the remaining contribution is also neglected.

The contribution from tracks originating from secondary particles are due to hadronic interactions, photon conversions and decays of long-lived particles. This contribution is subtracted from the number of reconstructed tracks before correcting for other detector effects. The contribution of tracks from secondary particles is estimated using simulation predictions for the shapes of the d_0^{BL} distributions, shown in Figure 5.27, for tracks from primary and secondary particles satisfying all track selection criteria except the one on d_0^{BL}. These predictions form templates which are fitted to the data in order to extract the relative contribution of tracks from secondary particles. The Gaussian core of the distribution is dominated by the tracks from primary particles, with a width determined by their d_0^{BL} resolution; tracks from secondary particles dominate the tails. The fit is performed in the region $4 < |d_0^{\mathrm{BL}}| < 9.5$ mm, in order to reduce the dependence on the description of the d_0^{BL} resolution, which affects the core of the distribution. From the fit, it was determined that the fraction of tracks from secondary particles in simulation needs to be scaled by a factor 1.38 ± 0.14. This indicates that $(2.3 \pm 0.6)\,\%$ of tracks satisfying the final track selection criteria ($|d_0^{\mathrm{BL}}| < 1.5$ mm) originate from secondary particles, where systematic uncertainties are dominant and are discussed below. Of these tracks 6% come from photon conversions and the rest from hadronic interactions or long-lived decays. The description of the η and p_{T} dependence of this contribution is modelled sufficiently accurately by the simulation that no additional correction is required.

Figure 5.27(a) displays a small disagreement in the core of the d_0^{BL} distribution. This has no impact in the tail of the distribution used for the fit. The dominant systematic



(a) (b)

FIGURE 5.27: Comparison between data and PYTHIA A2 simulation for (a) the transverse impact parameter distribution of the tracks, prior to any requirement on the transverse impact parameter, calculated with respect to the average beam position, d_0^{BL}; and (b) the difference between the longitudinal position of the track along the beam line at the point where d_0^{BL} is measured and the longitudinal position of the primary vertex projected to the plane transverse to the track direction, $|\Delta(z_0 \cdot \sin\theta)|$, prior to any requirement on $|\Delta(z_0 \cdot \sin\theta)|$. Only statistical uncertainties of the data are shown. The separate contributions from tracks coming from primary and secondary particles are also shown and the fraction of secondary particles in the simulation is scaled by 1.38 to match that seen in the data, with the final simulation distributions normalised to the number of tracks in the data. The inserts in the panels show the distributions on a linear scale. Ref. [80].

uncertainty stems from the interpolation of the number of tracks from secondary particles from the fit region to the region $|d_0^{\mathrm{BL}}| < 1.5$ mm. Different generators are used to estimate the interpolation and differences between data and simulation in the shape of the d_0^{BL} distribution in the fit region are considered. Additional, much smaller, systematic uncertainties arise from a variation of the fit range, considering the η dependence of the fitted fractions and from using special simulation samples with varying amounts of detector material. There is a second source of non-primary particles: charged-particles with a mean lifetime $30 < \tau < 300$ ps which, unlike in previous analyses [122, 166], are excluded from the primary-particle definition. These are charged strange baryons which decay after a short flight length and have a very low track reconstruction efficiency. Reconstructed tracks from these particles are treated as background and are subtracted. The fraction of reconstructed tracks coming from strange baryons is estimated from simulation with EPOS to be $(0.01 \pm 0.01)\%$ on average, with the fraction increasing with track p_{T} to be $(3 \pm 1)\%$ above 20 GeV. The fraction is much smaller at low p_{T} due to the extremely low efficiency of reconstructing a track from a particle that decays early in the detector. The systematic uncertainty is taken as the maximum difference between the nominal EPOS prediction and that of PYTHIA A2 or PYTHIA MONASH, which is then symmetrised.

Studies on strange baryons will be given in Section 5.9.

There is also a contribution of less than 0.1% from fake tracks (those formed by a random combination of hits or from a combination of hits from several particles); these are neglected.

5.8.1 Backgrounds in the Extended Phase Space

Background contributions to the tracks from primary particles include, as in the case of the nominal phase-space, fake tracks, strange baryons and secondary particles. Also in this case, the background contributions are subtracted on a statistical basis from the number of reconstructed tracks before correcting for other detector effects. The contribution of fake tracks, estimated from simulation, is higher in the low-pt regime and not negligible since it goes up to a 1% for all p_{T} and η intervals with a relative uncertainty of $\pm$50% determined from dedicated comparisons of data with simulation, as discussed in Chapter 3.

The charged strange baryons are treated as background as in the nominal phase space, but the fraction is much smaller at low p_{T} due to the extremely low track reconstruction efficiency.

The contribution from secondary particles is estimated by performing a template fit to the distribution of the track transverse impact parameter d_0^{BL}as for the nominal phase-space, using templates for primary and secondary particles created from PYTHIA A2 simulation. The shape of the transverse impact parameter distribution differs for electron and non-electron secondary particles, as the d_0^{BL} reflects the radial location at which the secondaries were produced. The processes for conversions and hadronic interactions are rather different, which leads to differences in the radial distributions. The electrons are more often produced from conversions in the beam pipe. Furthermore, the fraction of electrons increases as p_{T} decreases. Therefore, separate templates are used for electrons and non-electron secondary particles in the region $p_{\mathrm{T}} < 500\,\mathrm{MeV}$. The rate of secondary tracks is the sum of these two contributions and is measured with the fit. The background normalisation for fake tracks and strange baryons is determined from the prediction of the simulation. The fit is performed in nine p_{T} intervals, each of width 50 MeV, in the region $4 < |d_0^{\mathrm{BL}}| < 9.5$ mm. The fitted distribution for $100 < p_{\mathrm{T}} < 150$ MeV is shown in Figure 5.28. For this p_{T} interval, the fraction of secondary tracks within the region $|d_0^{\mathrm{BL}}| < 1.5$ mm is measured to be $(3.6 \pm 0.7)\%$, equally distributed between electrons and non-electrons. For tracks with $p_{\mathrm{T}} > 500$ MeV, the fraction of secondary particles is

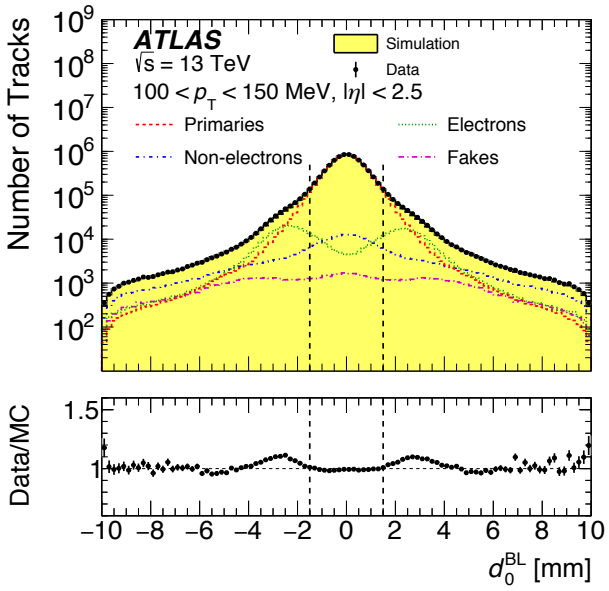


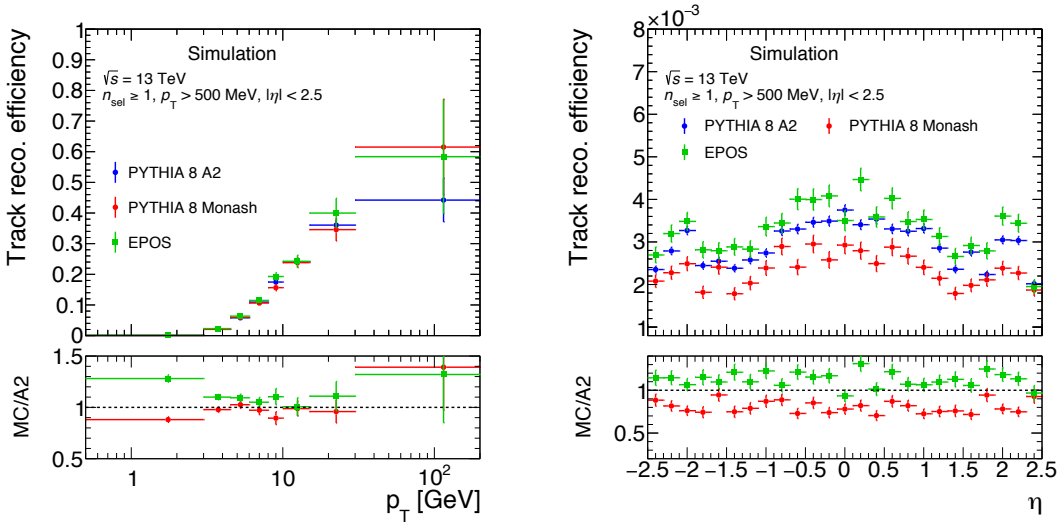
FIGURE 5.28: Comparison between data and PYTHIA A2 simulation for the transverse impact parameter d_0^{BL} distribution. The d_0^{BL} distribution is shown for $100 < p_{\mathrm{T}} < 150\mathrm{MeV}$ without applying the cut on the transverse impact parameter. The position where the cut is applied is shown as dashed black lines at ± 1.5 mm. The simulated d_0^{BL} distribution is normalised to the number of tracks in data and the separate contributions from primary, fake, electron and non-electron tracks are shown as lines using various combinations of dots and dashes. The secondary particles are scaled by the fitted fractions as described in the text. Only statistical uncertainties of the data are shown. The lower panel shows the ratio of data to MC prediction. From [81].

measured to be $(2.3 \pm 0.6)\%$; these are mostly non-electron secondary particles. The uncertainties are evaluated by using different generators to estimate the interpolation from the fit region to $|d_0^{\mathrm{BL}}| < 1.5$ mm, changing the fit range and checking the η dependence of the fraction of tracks originating from secondaries. This last study is performed by fits integrated over different η ranges, because the η dependence could be different in data and simulation, as most of the secondary particles are produced in the material of the detector. The systematic uncertainties arising from imperfect knowledge of the passive material in the detector are also included; these are estimated using the same material variations as used in the estimation of the uncertainty on the tracking efficiency, described in Section 5.7.2.

5.9 Strange Baryons

In contrast to previous analyses, in the measurements of charged-particle multiplicities at 13 TeV (and at 8 TeV [167]) all particles with mean lifetime in the range $30 < \tau < 300$ ps and their decay products were removed from the fiducial phase space definition. These are mostly charged strange baryons which decay after a short distance.

The efficiency as a function of the reconstructed track p_T (right) and η (left) is shown in Figure 5.29 for different simulations. The efficiency is about 5% at $p_T = 5$ GeV and increases to 50% at $p_T = 30$ GeV in PYTHIA 8 A2.



(a) (b)

FIGURE 5.29: The efficiency of strange baryons at $\sqrt{s} = 13$ TeV for tracks with transverse momentum $p_{\mathrm{T}} > 500$ MeV as a function of the (a) transverse momentum p_{T} and of the (b) pseudorapidity η.

The fraction of reconstructed tracks associated with strange baryons is shown in Figure 5.30 as a function of the reconstructed track p_T (right) and η (left) . The fraction increases from 0.5% at $p_T = 5$ GeV to about 3% at $p_T = 30$ GeV in PYTHIA 8 A2.

The fraction of strange baryons is also shown as a function of the kinematic variables at generator level, which can be seen in Figure 5.31.

A very small rate is observed with a systematic variation of about 80% in different generators.

The track reconstruction efficiency with and without the strange baryon contribution was measured for different generators and it is shown in Figure 5.32 for PYTHIA 8 A2 (left) and EPOS (right). After removing the strange baryon contribution, the primary track efficiency is smoothly increasing from 72% at $p_T = 500$ MeV to about 88% for $p_T = 30$ GeV.

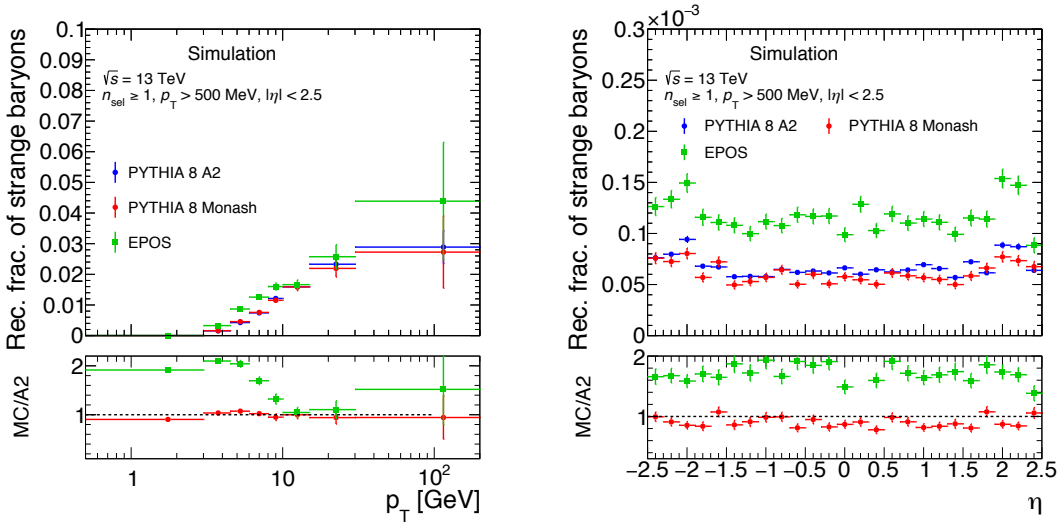
Including the strange baryons in the tracking efficiency would lead to a large systematic error of the order of $\mathcal{O}(1\%)$. Therefore these particles are removed from the primary particle sample, the fiducial definition and the unfolded distributions.

The remaining low rate of accepted tracks from strange baryons is removed using a track p_{T}-dependent parametrisation from the EPOS generator, since this agrees better with the ALICE measurement [173], as it will be shown in Section 5.9.1.

As systematic uncertainty, the full difference of the rate of strange baryons between the EPOS and PYTHIA 8 generator is used. No η dependence bigger than the statistical error was found for these tracks and therefore this was not taken into account in the systematic error. Figure 5.33 shows both distributions. The systematic error shown in the accepted p_{T} spectrum is the difference in the rate between PYTHIA 8 and EPOS .

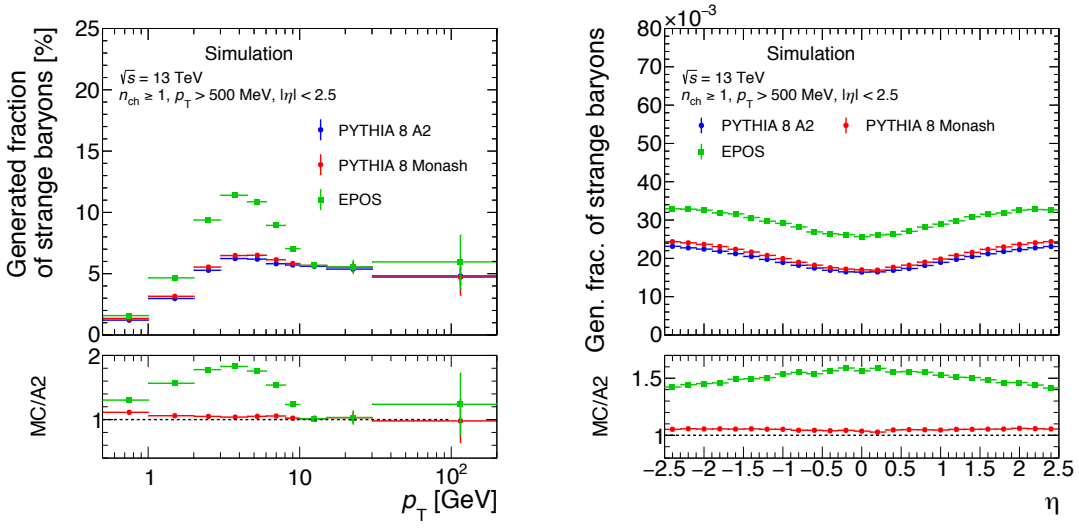
5.9.1 Comparison to ALICE Results

Even if the the strange baryons are excluded from the definition of primary charged-particles in this analysis, the final distributions are produced also with the strange baryons included in order to compare the results with other experiments. These distributions are obtained by deriving the extrapolation factors from MC and applying them to the nominal distributions which do not include strange baryons. The extrapolation factors are derived by comparing the MC distributions with/without strange baryons included.



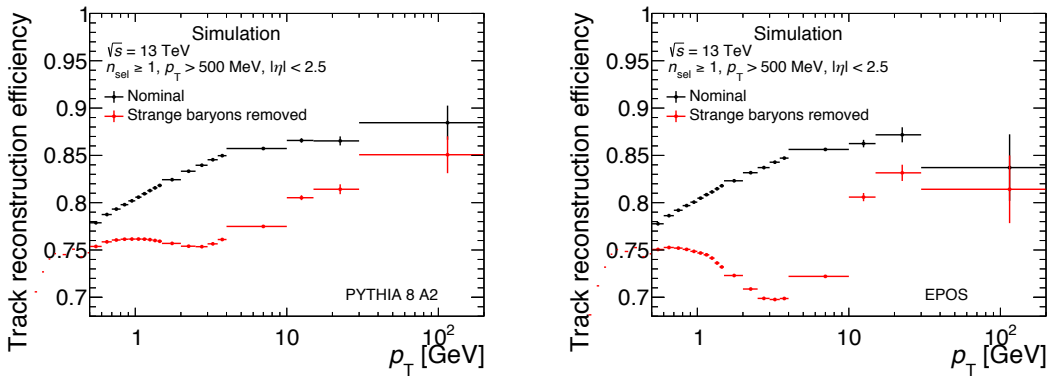
(a) (b)

FIGURE 5.30: The reconstructed fraction of strange baryons at $\sqrt{s} = 13$ TeV for tracks with transverse momentum $p_{\mathrm{T}} > 500$ MeV as a function of the (a) transverse momentum p_{T} and of the (b) pseudorapidity η.



(a) (b)

FIGURE 5.31: The generated fraction of strange baryons at $\sqrt{s} = 13$ TeV for tracks with transverse momentum $p_{\mathrm{T}} > 500$ MeV as a function of the (a) transverse momentum p_{T} and of the (b) pseudorapidity η.



(a) (b)

FIGURE 5.32: The efficiency of (non)strange baryons at $\sqrt{s} = 13$ TeV for tracks with transverse momentum $p_{\mathrm{T}} > 500$ MeV: (a) PYTHIA 8 A2 and (b) EPOS.

The ALICE experiment measures the rate of strange baryons Ω^-, $\bar{\Omega}^+$, Ξ^-, $\bar{\Xi}^+$ in pp collisons at 7 TeV [174] and their resonances $\Sigma(1385)^+, \Xi(1530)^0, \bar{\Xi}(1530)^0$ [175]. These data are used here to compare with the MC models and accordingly decide which MC sample to use for the extrapolation.

The Rivet analyses of the ALICE measurement are publicly available [173, 176], thus high-statistics Monte Carlo samples are generated at 7 TeV for PYTHIA 8 Monash and A2 and for EPOS with the exact same settings as the samples produced at 13 TeV. Afterwards, the Rivet analysis was run on these samples.

The comparison of strange baryon distributions is shown Fig. 5.34. It shows that both PYTHIA 8 tunes significantly underestimate the strange baryon production for Ω (top) and Ξ (middle) baryons and their ratio (bottom). The EPOS prediction is within 20% of the the ALICE data.

In Fig 5.35 left, the production rate of $\Sigma^0(1385)$ baryons as a function of their p_{T} per unit rapidity is shown while the right plot shows the production of $\Xi(1530)$ and $\bar{\Xi}^0(1530)$ baryons. Both PYTHIA 8 tunes are similar and they heavily underpredict the data (up to factor of 3-5 in certain bins). EPOS describes the data best although it is also not perfect and underestimates the data up to 20% at low p_{T} while overestimates the data up to 80% at high p_{T}.

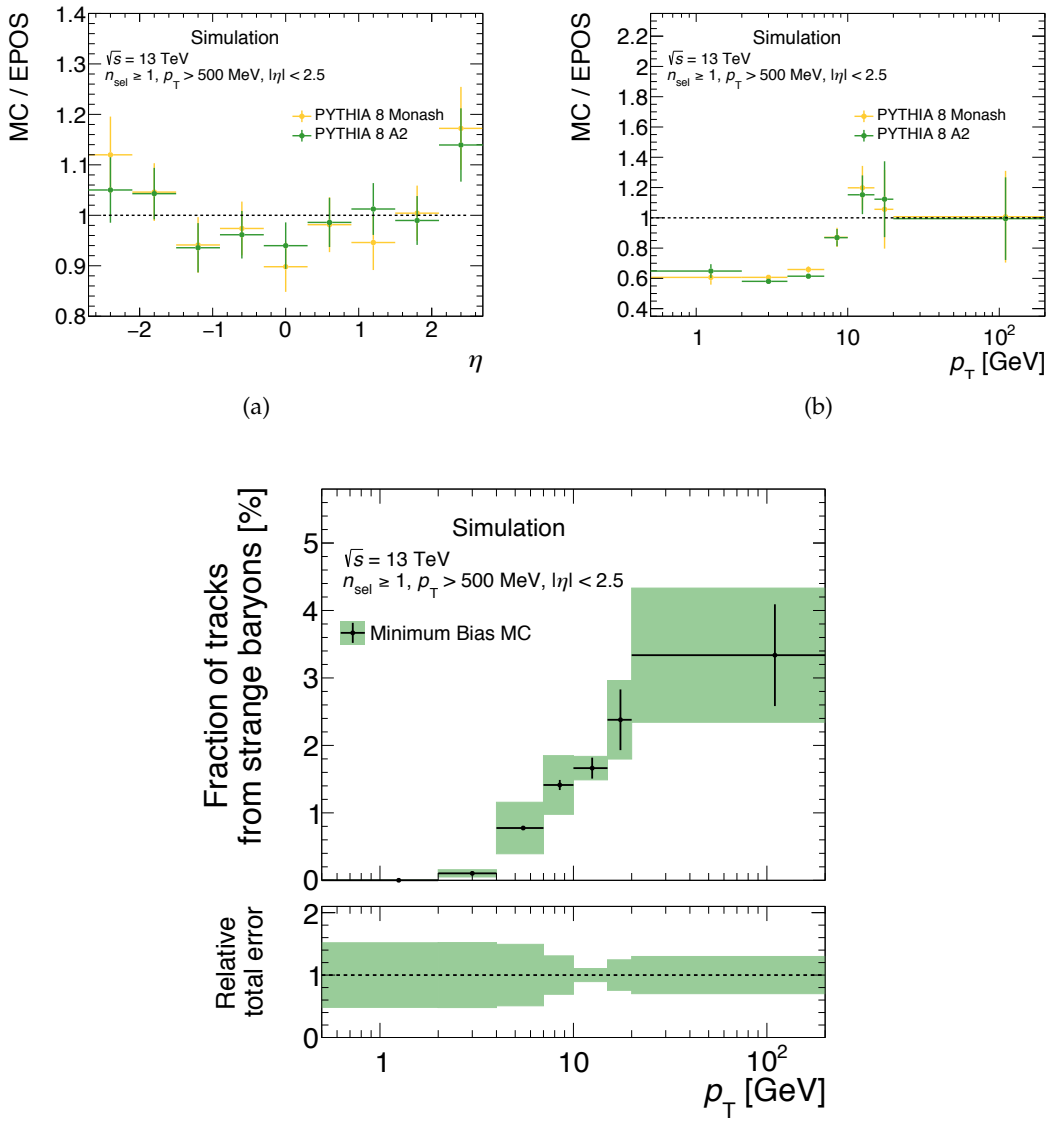
Since EPOS describes the data best, this generator is chosen to derive the correction and to obtain the results with the strange baryons included. The largest difference between the EPOS and other MC generators are used as a systematic uncertainty.

5.10 Correction Procedure

The data must be corrected to obtain inclusive spectra of charged primary particles satisfying the particle-level phase-space requirement. These corrections include inefficiencies due to trigger selection, vertex and track reconstruction. The effect of events lost due to the trigger and vertex requirements is corrected by an event-by-event weight:

$$w_{\mathrm{ev}}(n_{\mathrm{sel}}, \eta) = \frac{1}{\epsilon_{\mathrm{trig}}(n_{\mathrm{sel}}^{\mathrm{no-z}})} \cdot \frac{1}{\epsilon_{\mathrm{vtx}}(n_{\mathrm{sel}}^{\mathrm{no-z}}, \eta)}, \tag{5.4}$$

where $n_{\mathrm{sel}}^{\mathrm{no-z}}$ is the multiplicity of selected tracks with the relaxed selection (no cut on vertex and hence no cut on longitudinal impact parameter, see Section 5.6). The η dependence



(a) (b)

(c)

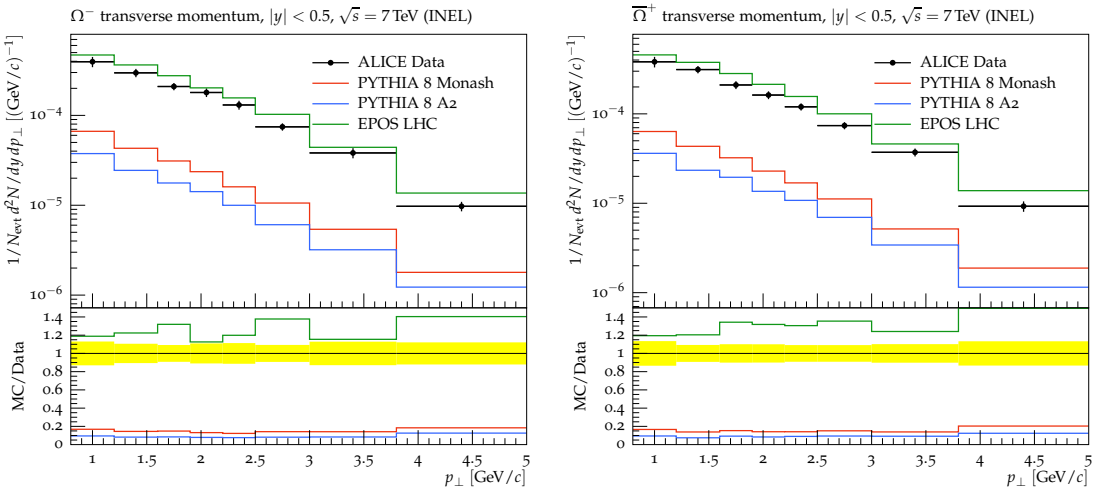
FIGURE 5.33: Ratio of the selected strange baryon fraction in PYTHIA 8 with Monash tune and EPOS over PYTHIA 8 with A2 tune as function of (a) η and (b) p_{T}. (c) Selected strange baryon fraction in A2. The systematic error in green includes the difference in the generated fraction for the different generators and the statistical error of the estimate.

is relevant only for the first analysis bin $n_{\mathrm{sel}}^{\mathrm{no-z}} = 1$, and corresponds to the η of the selected track. The calculation of the trigger efficiency ϵ_{trig} is explained in Section 5.5. The calculation of ϵ_{vtx} is explained in Section 5.6.

To correct for inefficiencies in the track reconstruction, the p_{T}, η, and mean p_{T} distributions of the selected tracks are corrected for with a track-by-track weight:

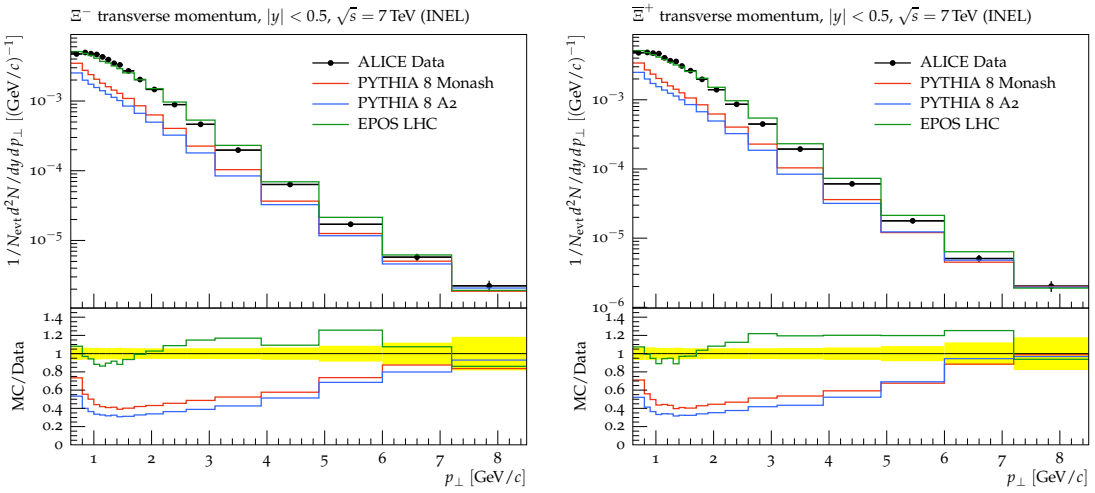
$$w_{trk}(p_{\mathrm{T}}, \eta) = \frac{1}{\epsilon_{\mathrm{trk}}(p_{\mathrm{T}}, \eta)} \cdot (1 - f_{\mathrm{nonp}}(p_{\mathrm{T}}, \eta) - f_{\mathrm{okr}}(p_{\mathrm{T}}, \eta) - f_{\mathrm{sb}}(p_{\mathrm{T}})), \tag{5.5}$$

where f_{nonp} and f_{sb} are the fractions of non-primary tracks and weakly decaying strange hadrons determined as described in Sections 5.8 and 5.9. The fraction of selected tracks for



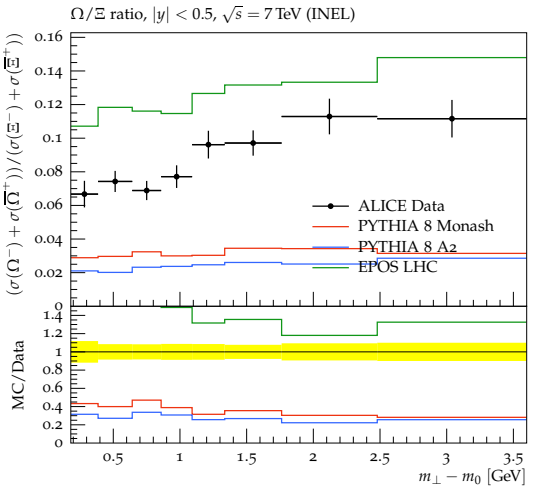
(a) Ω^-

(b) $\bar{\Omega}^+$



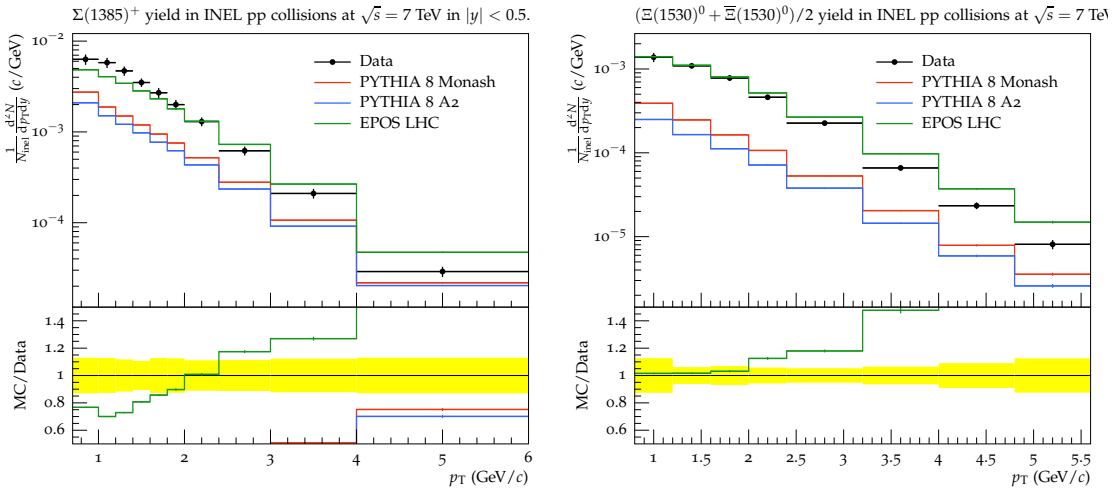
(c) Ξ^-

(d) $\bar{\Xi}^+$



(e) Ω/Ξ ratio

FIGURE 5.34: The rate of strange baryons (a) Ω^-, (b) $\bar{\Omega}^+$, (c) Ξ^-, (d) $\bar{\Xi}^+$ and (e) the average rates of double-strange baryons Ω^-+ $\bar{\Omega}^+$ over Ξ^-+ $\bar{\Xi}^+$ as a function of their transverse momentum. The data points correspond to ALICE measurement [174].



(a) Σ (b) Ξ

FIGURE 5.35: The rate of strange baryons (a) $\Sigma(1385)^+$ and (b) the average rates of double-strange baryons $\Xi(1530)^0$ and $\overline{\Xi}(1530)^0$ (right) as a function of their transverse momentum. The data points correspond to ALICE measurement [175].

which the corresponding primary particle is outside the kinematic range, $f_{\text{okr}}(p_{\text{T}}, \eta)$, originates from resolution effects and is estimated using MC. Note that due to the definition of the primary efficiency down to $p_{\text{T}} > 100$ MeV, this correction is non-zero only at $|\eta| \sim 2.5$, not at the $p_{\text{T}} = 500$MeV boundary At $p_{\text{T}} > 500$ MeV only the residual difference between the migration in the phase space and out of the phase space needs to be studied. The uncertainty on f_{okr} is mostly due to the uncertainty on the resolution difference between data and MC which is studied using extra material samples, and a small residual effect (<1%) of migration between bins was observed. These migrations largely cancel due to the fact that the primary tracking efficiency is derived in terms of particle p_{T} and η, and it is applied to tracks with reconstructed p_{T} and η. The last term, $f_{\text{sb}}(p_{\text{T}})$ indicates the removal of a fraction of tracks due to strange baryons seen in Figure 5.33.

The correlations between particles in term of primary track efficiency are expected to be small (unless looking at events with more than several hundreds of particles) and, thus, neglected when correcting the distributions.

The correction procedure for each of the measured distributions is discussed separately in the following Sections. All the event and track weights applied are derived from the PYTHIA 8 A2 Monte Carlo generator.

Figure 5.36 shows the raw distributions, in a wider range than presented on final distributions.

5.10.1 Brief summary of systematics

The status of the experimental systematic uncertainties and systematic uncertainties due to non-closure of the unfolding method is shown in Table 5.9. By default, the following uncertainties are propagated to all distributions:

- Material systematics - includes the uncertainty due to material modeling as well as the uncertainty due to the N-1 tests
- Statistical uncertainty of the tracking efficiency

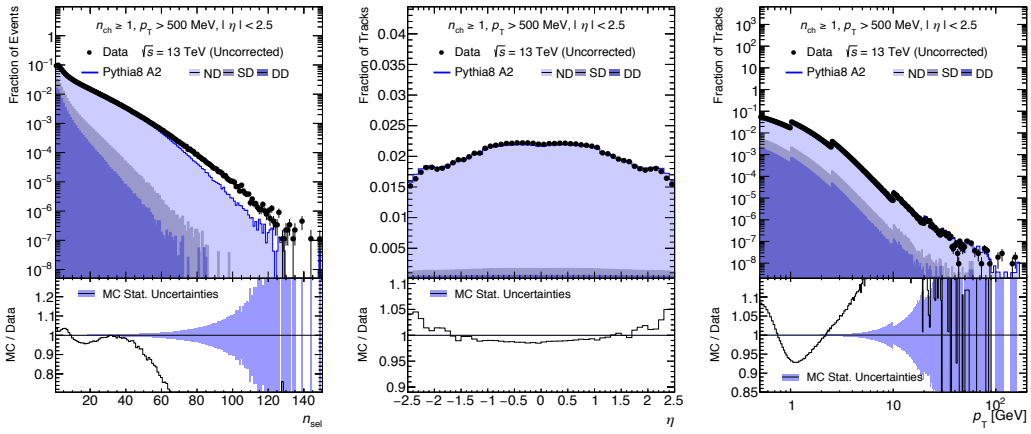


FIGURE 5.36: Raw distributions normalized to 1. For p_{T} distribution, the structure is induced by intermediate binning. The same is used in the construction of p_{T} unfolding matrix.

- Fraction of non-primary tracks
- Fraction of badly measured tracks which are not removed by the χ^2 cut
- Systematic uncertainty related to the removal of strange baryons.

Systematic uncertainties due to vertex and trigger efficiency are neglected as the largest contribution is for $n_{\mathrm{sel}}^{\mathrm{no-z}}$=1, where each of them is smaller than 0.2%, while the systematic due to tracking efficiency and non-closure is 4% in the $n_{\mathrm{sel}}^{\mathrm{no-z}}$=1 bin. In addition, a systematic uncertainty due to non-closure observed in particular distributions is included.

Table 5.9

Distribution	Systematic uncertainty
all	material
all	secondaries
all	strange baryons
all	effect badly measured tracks
p_{T}	aligment
n_{ch}	uncertainty on $\varepsilon^{\mathrm{miss}}$ - taken into account by varying input distribution
n_{ch},p_{T},$\langle p_{\mathrm{T}} \rangle$	propagation of stat uncert. due to matrix
	non-closures
n_{ch}	4,2,1% in n_{ch} =1,2,3, cover vtx, trigger syst
η	0.5% - flat
p_{T}	1% for p_{T} >3GeV, 2% for p_{T} >20 GeV
$\langle p_{\mathrm{T}} \rangle$	2% in n_{ch} =1,2,3 and 0.5% everywhere

TABLE 5.9: Experimental systematic uncertainties and non-closure uncertainties.

5.10.2 Correction to $\frac{\mathrm{d}N_{\mathrm{ev}}}{\mathrm{d}n_{\mathrm{ch}}}$

The observed n_{sel} distribution is first corrected for the vertex and trigger reconstruction efficiencies. To remove the detector effects in the final results, it is needed to express

the multiplicity distribution in terms of the number of charged primary particles n_{ch} instead of the number of selected charged tracks n_{sel}. For this, a Bayesian unfolding procedure [177] is used. Integrating the probability relation given by the Bayes theorem $P(n_{\mathrm{ch}}) \cdot P(n_{\mathrm{sel}}|n_{\mathrm{ch}}) = P(n_{\mathrm{ch}}|n_{\mathrm{sel}}) \cdot P(n_{\mathrm{sel}})$ over n_{sel}, one gets the distribution of primary particles

$$N_{\mathrm{ev}}(n_{\mathrm{ch}}) = \sum_{n_{\mathrm{sel}} \geq 0} P(n_{\mathrm{ch}}|n_{\mathrm{sel}}) \cdot N_{\mathrm{ev}}(n_{\mathrm{sel}}) = \frac{1}{\varepsilon^{\mathrm{miss}}(n_{\mathrm{ch}})} \sum_{n_{\mathrm{sel}} \geq 1} P(n_{\mathrm{ch}}|n_{\mathrm{sel}}) \cdot N_{\mathrm{ev}}(n_{\mathrm{sel}}). \quad (5.6)$$

The second relation factorizes the contribution of events that are lost due to track reconstruction inefficiency but would pass the particle level phase-space cuts, i.e. those with $n_{\mathrm{ch}} \geq 1$ but n_{sel} =0. They are corrected for by a special factor $\varepsilon^{\mathrm{miss}}(n_{\mathrm{ch}})$:

$$\varepsilon^{\mathrm{miss}}(n_{\mathrm{ch}}) = 1 - (1 - \bar{\epsilon}_{\mathrm{trk}})^{n_{\mathrm{ch}}} \quad (5.7)$$

where $\bar{\epsilon}_{\mathrm{trk}}$ is the mean tracking efficiency in the $n_{\mathrm{sel}} = 1$ bin. The correction is significant in the first n_{ch} bin, with significantly smaller contributions in higher bins. The procedure assumes that the mean tracking efficiency is constant as a function of n_{ch}. This is only an approximation as a small dependence can be seen from Figure 5.42, however the residual effect on a non-closure due to this effect should be smaller than 1%. The average tracking efficiency in $n_{\mathrm{sel}} = 1$ is found to be $\bar{\epsilon}_{\mathrm{trk}}$=0.79 (Data), 0.78 (PYTHIA 8 A2), 0.78 (PYTHIA 8 MONASH), 0.76 (HERWIG++), 0.79 (EPOS).

The conditional probability $P(n_{\mathrm{ch}}|n_{\mathrm{sel}})$ of observing an event with true n_{ch}, given the observed n_{sel}, is unknown. The central idea of the Bayesian unfolding is to get an estimate for $P(n_{\mathrm{ch}}|n_{\mathrm{sel}})$ from MC, and improve the estimate iteratively. For r iteration:

$$P(n_{\mathrm{ch}}|n_{\mathrm{sel}})^{r+1} = P(n_{\mathrm{sel}}|n_{\mathrm{ch}}) N^r_{\mathrm{ev}}(n_{\mathrm{ch}}) / N^r_{\mathrm{ev}}(n_{\mathrm{sel}}). \quad (5.8)$$

where $P(n_{\mathrm{sel}}|n_{\mathrm{ch}})$ is the *resolution function* and $N^r_{\mathrm{ev}}(n_{\mathrm{ch}})$ and $N^r_{\mathrm{ev}}(n_{\mathrm{sel}})$ are distributions of primary particles and of selected tracks, respectively. For $r = 0$, these are the prior estimates from MC. The matrix $P(n_{\mathrm{ch}}|n_{\mathrm{sel}})^r$ is normalized such that the number of events does not change except for the rare case ($\sim 2\%$ for low multiplicities) where $n_{\mathrm{sel}} > n_{\mathrm{ch}}$ and n_{ch} is below our acceptance cut, $n_{\mathrm{ch}} \geq 1$. In this way, a small correction is applied to remove events due to secondary tracks that are characterized by $n_{\mathrm{ch}} = 0$, but are reconstructed in the fiducial definition.

This matrix is obtained using Monte Carlo and is applied to data to obtain the observed n_{ch} distribution. The unfolding matrix from Eq. 5.8 is shown in Figure 5.37. In order to keep statistical uncertainties coming from the matrix small, samples filtered on multiplicity are used (see Table 5.2). They were produced using a multiplicity filter counting primary charged-particles above $p_{\mathrm{T}} > 100$ MeV with the following thresholds $n_{\mathrm{ch}} > 120, 160, 200$.

While the matrix is mostly populated along the diagonal, it shows events with large multiplicity at particle level, reconstructed with significantly smaller multiplicity at detector level. These events contain tracks with p_{T} similar to the p_{T} of true particles, but show low number of silicon hits. This observation points to the problem of the track reconstruction. The rate of these events is, however, small (in the sub-permill level), and is neglected in the analysis.

Closure tests

Several closure tests are performed using different Monte-Carlo generators (PYTHIA 8 A2, PYTHIA 8 MONASH, HERWIG++ , EPOS) for the input pseudo-data and the matrix. Here,

two closure tests are described.

In the first closure test, PYTHIA 8 A2 is used to construct the unfolding matrix and also to derive the input distributions. Results are shown in Figure 5.38. A comparison is given for events with $n_{\mathrm{sel}} \geq 1$ that do not migrate outside of the detector phase space, which is labeled "visible" in Figure 5.38 (left). This corresponds to Eq. 5.6 without the $\varepsilon^{\mathrm{miss}}(n_{\mathrm{ch}})$ factor and probes specifically the deconvolution step of the correction.

The second closure test consists of using PYTHIA 8 A2 to construct the unfolding matrix and apply it to EPOS as pseudo-data. Results are shown in Figure 5.39. The number of iterations was determined in closure tests using different models for both the pseudodata and the matrix, requiring that a closure within the statistical uncertainty of the method is obtained. Four iterations were used. Figure 5.39 (right) shows how the results improve with increasing the number of iterations.

The non-closure is observed mainly at low-n_{ch}. In the tail, a non-closure of several percents is negligible with respect to the systematic error, and is neglected. The assigned non-closure systematics is 4, 2, 1% in n_{ch} =1,2,3, respectively.

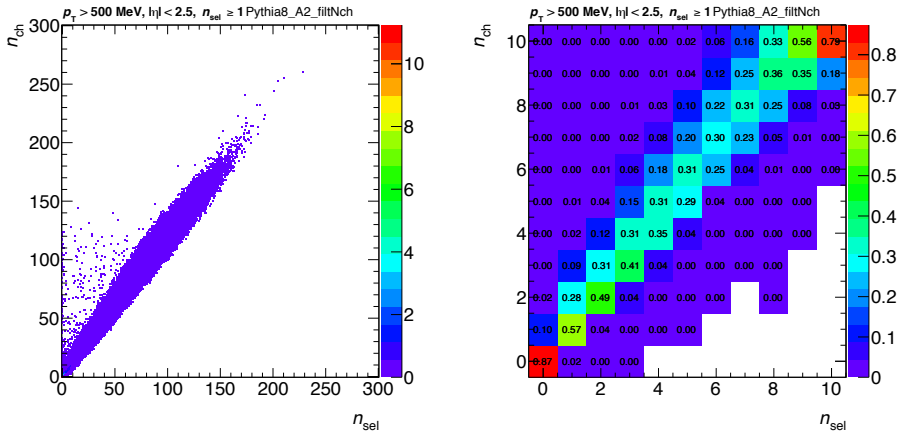


FIGURE 5.37: Migration matrix derived from PYTHIA 8 A2 combined with filtered samples (left). On the right, columns are normalized to 1, as in Eq. 5.8.

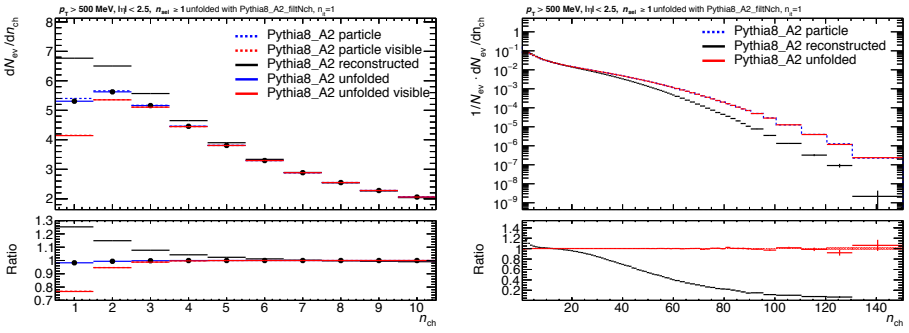


FIGURE 5.38: Closure test using PYTHIA 8 A2 as the input distribution and PYTHIA 8 A2 generator to construct unfolding matrix. Left side show a zoom in for $n_{\mathrm{ch}} \leq 10$ The label "visible" is used for the subset of events with $n_{\mathrm{sel}} \geq 1$ that do not migrate outside of the detector phase space.

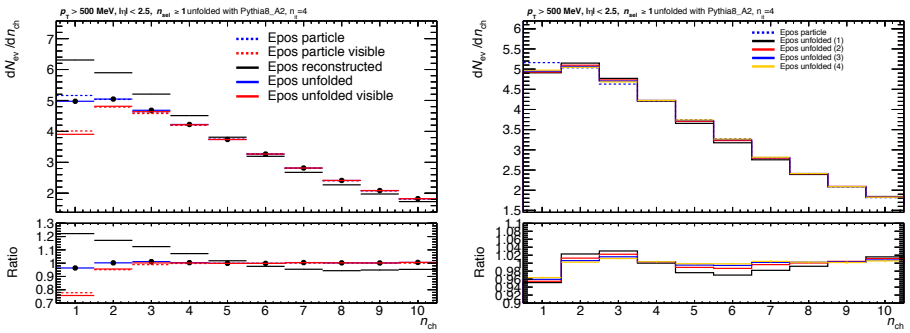


FIGURE 5.39: Closure test using EPOS for the input and PYTHIA 8 A2 generator to construct unfolding matrix. The number of iterations used on the left is $n_{it} = 4$. The ratio is calculated with respect to n_{ch} spectrum at particle level. The right plot shows how the results improve with increasing the number of iterations. The denominator in the ratio corresponds to the fourth and last iteration.

Data

The same procedure as described above for MC is applied for data, using the nominal PYTHIA 8 A2 with filtered samples for the unfolding. The number of iterations to convergence is small thanks to the good modeling of the multiplicity distribution, as seen already at detector level Figure 5.36. The next-to-last and last iterations have small differences within the stat uncertainty of the method.

The MC deviates from data for n_{sel} >60, and that is where the iterations improve as seen from figure 5.40 (right). In the same figure, the unfolded result using different models for unfolding of the data is shown, demonstrating a good stability of the unfolding procedure.

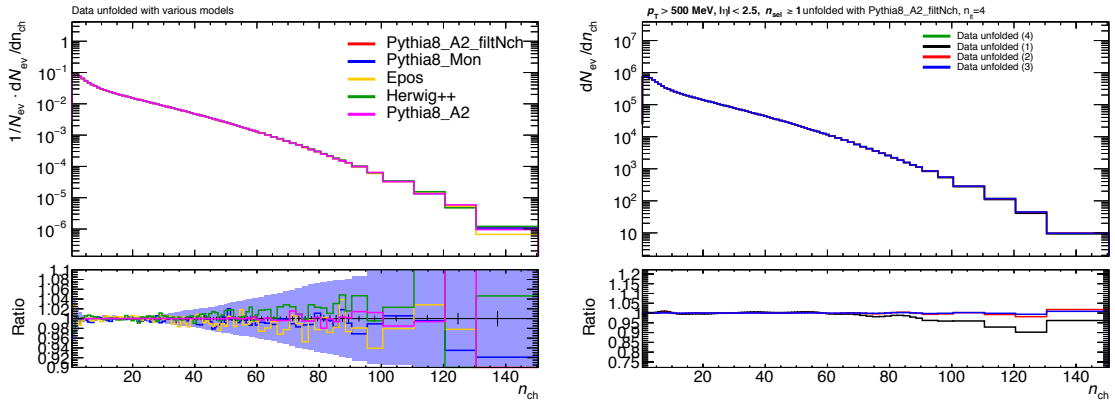


FIGURE 5.40: Comparison of unfolded data using the migration matrix from different models to the default model PYTHIA 8 A2 including filtered samples, see legend (left). The filled area denotes the total statistical plus systematic uncertainty. On the right, convergence of the unfolding for different iterations compared to the fourth and last iteration.

Systematic uncertainties

Systematic uncertainties on the unfolded distributions are obtained by using modified input distributions, to which the unfolding procedure is applied and results compared to

that obtained with the nominal input distribution. The matrix and the correction factors are not modified. A study was carried out in which the matrix was varied instead of the input with the conclusion that the two types of propagation of uncertainties yield a similar result. The unfolding procedure relies on the MC modeling of the $n_{\rm ch} \rightarrow n_{\rm sel}$ migration (resolution), which requires a good modeling of the tracking efficiency and of the particle $p_{\rm T}$-spectrum in different multiplicity bins. The systematic uncertainties are propagated starting from the observed $n_{\rm sel}$ spectrum in data and randomly removing and adding tracks in the $n_{\rm sel}$ distribution. The statistic and systematic uncertainties, which will be discussed in the following, are shown in Figure 5.41.

- **Track weight uncertainties**
  The number of tracks, which are removed/added depends on the $p_{\rm T}$ and η values of the existing tracks and their associated track reconstruction efficiency uncertainties $\Delta\epsilon_{\rm trk}(p_{\rm T}, \eta)$. For each track, a random number is thrown, and if smaller/larger than $\Delta\epsilon_{\rm trk}(p_{\rm T}, \eta)$, $n_{\rm sel}$ is increased/decreased by one. The systematic uncertainty due to the down variation of the tracking efficiency is slightly underestimated for low multiplicities, because it is not possible to create tracks from $n_{\rm sel}$ = 0. Therefore, the size of the uncertainty obtained from up variation is symmetrised and taken also as for the down variation in $n_{\rm ch}$ < 5 (very small effect, more significant for $p_{\rm T} > 100$ MeV phase-space). In this way, the following uncertainties are propagated: due to modeling of material, statistical error of the tracking efficiency, strange baryon fraction, fraction of secondaries. The uncertainties for different track efficiency components are added in quadrature, by summing positive and negative deviations from the nominal values separately.

- **$p_{\rm T}$-spectrum difference between Data and MC**
  In the 7 TeV analysis **??**, the systematic uncertainty due to differences in the $p_{\rm T}$-spectrum between data and MC was estimated by varying the track weight in the $n_{\rm sel}$ distribution. The variation was obtained using the difference in the average tracking efficiencies in data (derived from the $p_{\rm T}$-spectrum) and MC.

  As the average tracking efficiency is the ratio of the sum over all reconstructed tracks in data over all true tracks, i.e. tracks from charged-particles, the denominator part can be expressed by the sum of reconstructed tracks weighted by the track reconstruction efficiency $\epsilon_{\rm trk}$, $\bar{\epsilon}_{\rm trk}(n_{\rm sel}) = n(n_{\rm sel}) / \sum_{i \in n_{\rm sel}} \frac{1}{\epsilon_{\rm eff,i}}$. The efficiency $\bar{\epsilon}_{\rm trk}(n_{\rm sel})$ is computed by either integrating over the $p_{\rm T}$-spectrum in data or MC, considering thus the shape difference of the $p_{\rm T}$-spectrum in data and MC. The mean tracking efficiency versus multiplicity is seen in Figure 5.42. Because the largest deviation is only 0.2% for low $n_{\rm sel}$, and at large multiplicity, where the effect of event migration is the strongest, the agreement is very good, no systematic uncertainty is considered.

- **Non-closure** A systematic uncertainty due to non-closure discussed in previous section is added: 4, 2, 1% in $n_{\rm ch}$ =1,2,3, respectively.

5.10.3 Correction to $N_{\rm ev}$

The total number of events, $N_{\rm ev}$, is defined as the integral of the $n_{\rm ch}$ distributions after all corrections are applied.

Systematic Uncertainties

The systematic uncertainties on $N_{\rm ev}$ are obtained in the same way as for the $n_{\rm ch}$ distributions. Only the systematic uncertainties affecting the events to enter or leave the phase

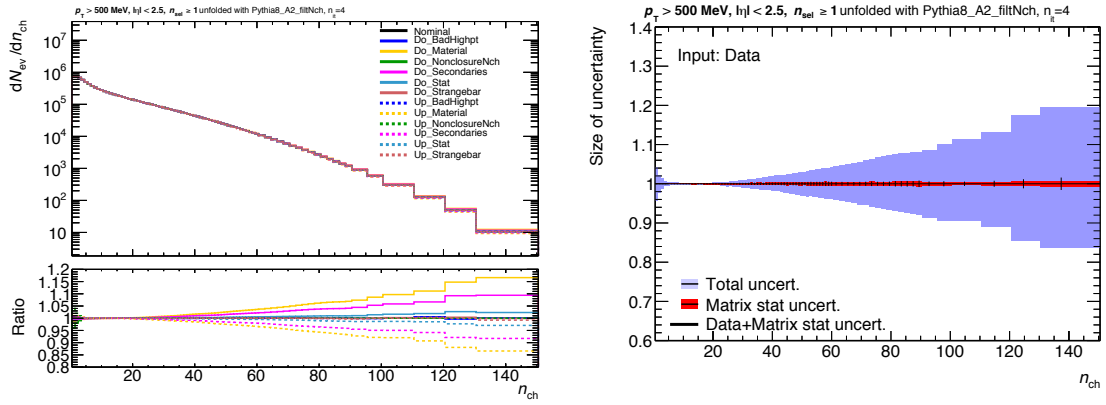


FIGURE 5.41: Left: Breakdown of systematic uncertainties to components of the tracking efficiency. Right: Breakdown of the statistical uncertainties showing the contribution of statistical uncertainty from the unfolding matrix.

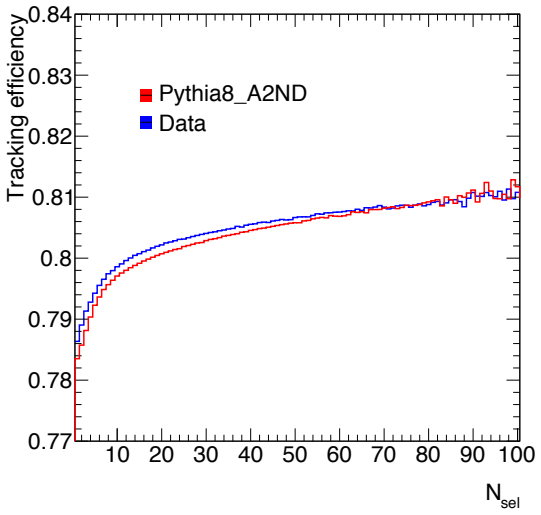


FIGURE 5.42: Measured mean tracking efficiency (see text) versus n_{sel} for data and PYTHIA 8 A2.

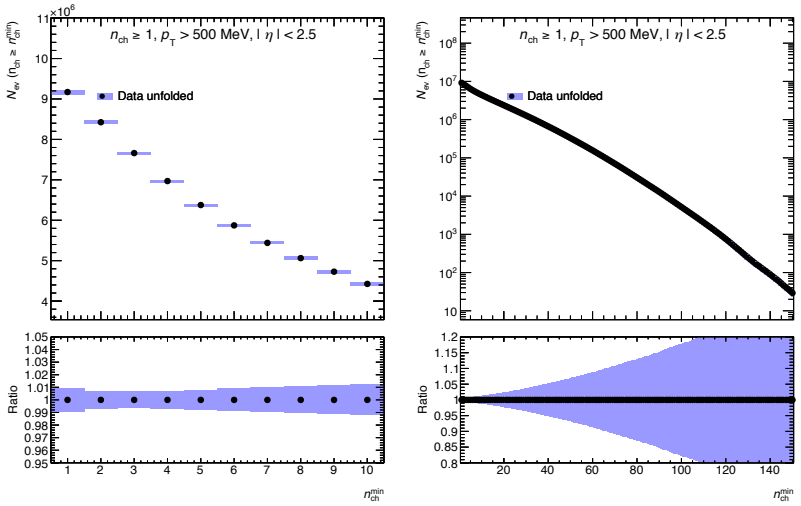


FIGURE 5.43: Total systematic uncertainty for $n_{\mathrm{ch}} > n_{\mathrm{ch}}^{\mathrm{min}}$ zoomed to low multiplicity (left), full range (right).

space have an impact on N_{ev}. Figure 5.43 shows the total systematic error on the total number of events $N_{\mathrm{ev}}(n_{\mathrm{ch}} > n_{\mathrm{nch}}^{\mathrm{min}})$ as a function of $n_{\mathrm{nch}}^{\mathrm{min}}$ which denotes a selection on particle multiplicity. Thus, the plot shows what is the expected number of events and its uncertainty for that phase-space (previously ATLAS measured $n_{\mathrm{ch}} \geq 6$,20).

The total number of events after unfolding is $N_{\mathrm{ev}} = 9.171 \cdot 10^6 \pm 0.060 \cdot 10^6$ (syst) for the nominal phase-space $n_{\mathrm{ch}} \geq 1$ with a relative uncertainty of 0.65%, due mostly to the track reconstruction efficiency and non-closure.

5.10.4 Corrections to $\frac{1}{N_{\mathrm{ev}}} \cdot \frac{\mathrm{d}N_{\mathrm{ch}}}{\mathrm{d}\eta}$

The tracks are first corrected for the event level inefficiencies of Eq. 5.4 of the trigger and the vertex. Then the tracks are corrected for the track reconstruction inefficiencies, non-primary track contamination, remaining strange baryons and out of kinematic range factors, see Eq. 5.5. The η distribution is normalized by N_{ev} obtained as discussed above in Section 5.10.3.

The mean particle density at $\eta = 0$ is calculated as an average over 4 bins of $\Delta\eta = 0.1$ contained in the range $|\eta| < 0.2$. The obtained charged-particle density per unit of pseudo-rapidity reads

$$\frac{1}{N_{\mathrm{ev}}} \frac{dn_{\mathrm{ch}}}{d\eta}\bigg|_{\eta=0} = 2.874 \pm 0.001(\mathrm{stat.}) \pm 0.033(\mathrm{syst.}).$$

The discussion of the correction and systematic uncertainties follows.

Closure tests

Figure 5.44 shows the closure test with the PYTHIA 8 A2 generator used both as the input and also as the MC to derive the tracking efficiency. All tests close within 1%. Removing the overall shift, which is taken into account by the systematic uncertainty on total number of events, the shape variation is within 0.5% for all generators. The source of the residual non-closure might be related to an observed poor momentum resolution for particles with momentum $p = 1.4\,\mathrm{GeV}$ (or, more precisely, ten times the pion mass). This comes from the reconstruction programs and is due to an issue in the treatment of energy loss in the track extrapolation(a discontinuous behavior of the Bethe-Bloch formula in the tracking

algorithms). Due to incorrectly reconstructed kinematics, these particles can then obtain an inappropriate tracking correction, leading to a non-closure.

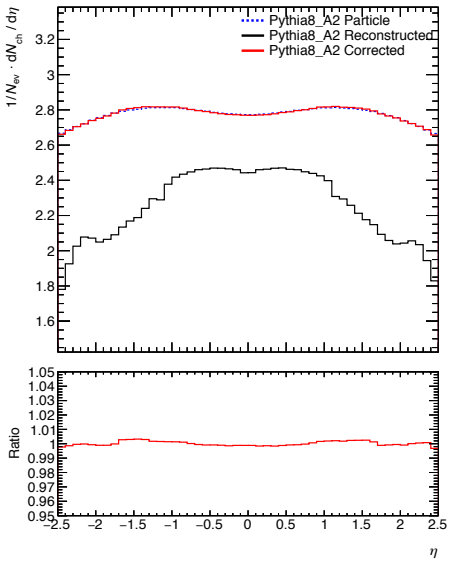


FIGURE 5.44: Closure test of the η distribution using PYTHIA 8 A2 MC.

Systematic uncertainties

In order to estimate the effect on the final η distributions of the uncertainties related to the correction procedure, the input detector level is modified by varying the track weights according to their systematic uncertainties. For each systematic uncertainty, the η distribution of the number of tracks is normalized by the N_{ev} derived for the same systematic variation from the multiplicity distribution, as mentioned in Section 5.10.3. The systematic uncertainties on the distribution and the normalization are therefore treated in a correlated way. The systematic uncertainties of the η densities are taken as the difference in each bin to the nominal distribution, summed in quadrature. Figure 5.45 shows the components contributing to the total systematic uncertainty (left), the largest being due to material (right). In addition to experimental uncertainties, a non-closure observed in multiplicity distribution is propagated. This is however negligible.

5.10.5 Corrections to $\frac{1}{N_{\mathrm{ev}}} \cdot \frac{\mathrm{d}N_{\mathrm{ch}}}{\mathrm{d}p_{\mathrm{T}}}$

The tracks are first corrected for the event level inefficiencies from the trigger and the vertex reconstruction. Then the tracks are corrected for the track reconstruction inefficiencies, non-primary track contamination and out of kinematic range factors. Finally, a similar unfolding method to that used on the n_{ch} distribution is used to correct the measured track p_{T} of the primary particle momentum. In order to populate the high p_{T} tails of the unfolding matrix, all available reprocessed non-diffractive models (PYTHIA 8 A2, EPOS) are added. In addition, MC samples generated with flat $\log(p_{\mathrm{T}})$ spectra are added to the samples ($\pi^{\pm}, p^{\pm}, K^{\pm}$). More bins are used for the unfolding than are shown in the final distributions; this is necessary in order to avoid amplification of small data-MC differences that would get amplified with successive iterations, causing large fluctuations. The p_{T} migration matrix is shown in Figure 5.46 for PYTHIA 8 A2 (left) and filled with additional single particle samples (right).

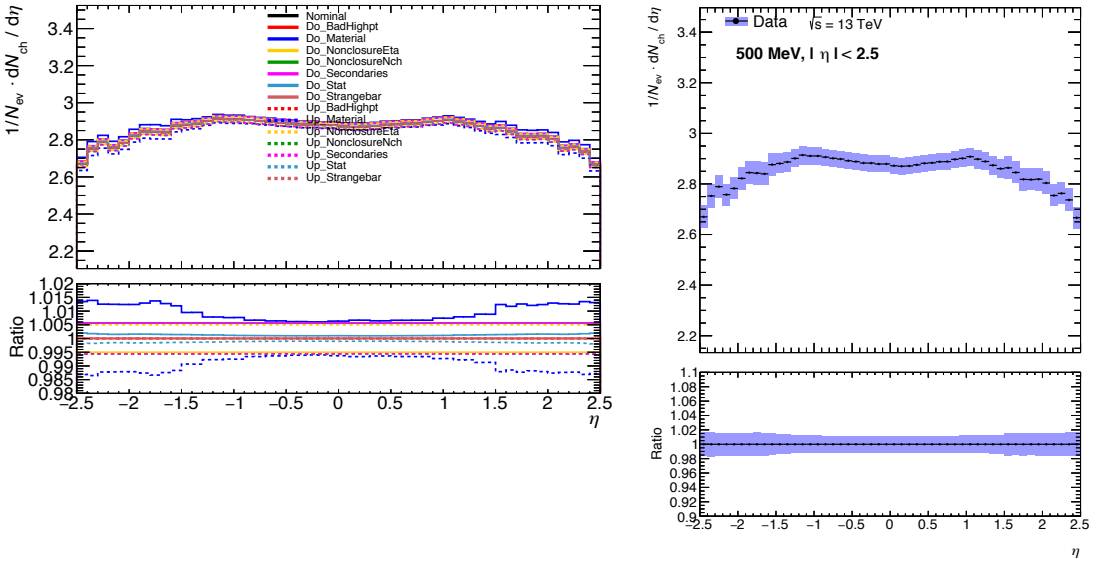


FIGURE 5.45: Components of the systematic uncertainties of the η distribution (left). Corrected data distribution with the total systematic uncertainty (right).

For this distribution five iterations are required before convergence is reached; convergence is defined as for the n_{ch} distribution: it is the first iteration for which stability between iterations is achieved.

Closure test

The correction procedure is tested by using different MC for the input distribution and for the unfolding matrix. The track weights are the default ones derived from the nominal PYTHIA 8 A2.

As already outlined in Section (5.9), removing strange baryons from the primary particle definition reduces the MC dependence on the correction method on the MC. An example of such a MC dependence inducing a bad closure test is seen in Appendix in Figure **??**

Figure 5.47 shows the closure test for PYTHIA 8 A2 used both as the input to be corrected and as the model used to build the matrix (left) and EPOS unfolded with PYTHIA 8 A2 for the unfolding matrix. The track correction accounts for most of the detector effects, except at high p_{T}, which is improved by the unfolding. A non-closure within 1% is seen after track weight correction for EPOS corrected with nominal PYTHIA 8 A2 MC which may be due to a residual particle composition difference between the generators in addition to the strange baryons.

Correction of the data

Figure 5.48 shows unfolding iterations when unfolding data (right). For comparison, the unfolding iterations are also displayed for EPOS generator (left).

Systematic Uncertainties

In order to estimate the effect on the final p_{T} distributions, the track weights are varied by their systematic uncertainties. For each variation a new p_{T} distribution is obtained and put

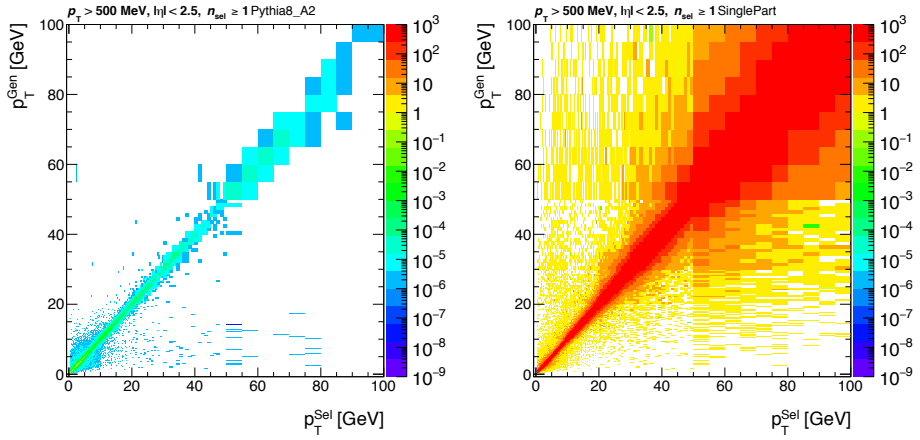


FIGURE 5.46: Unfolding matrix used for the p_T distribution filled using PYTHIA 8 A2 (left) and using ND contributions from several MC as well as single particle samples (right).

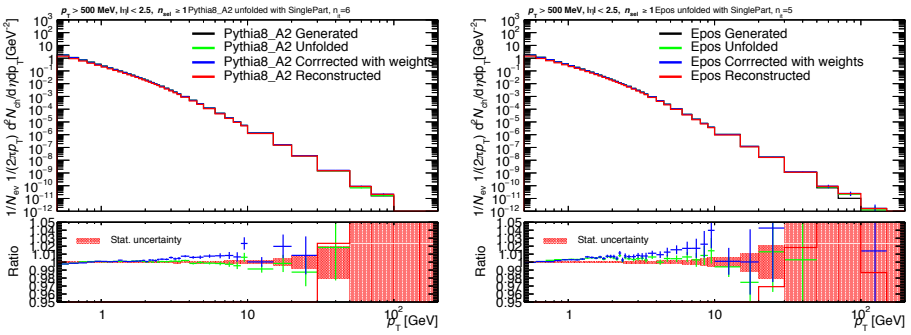


FIGURE 5.47: Left: Closure test derived using PYTHIA 8 A2 as the MC for both the input and matrix. Right: Closure test with EPOS input and PYTHIA 8 A2 as the matrix. Track wise corrections are derived from PYTHIA 8 A2 in both cases. The blue distribution is obtained by applying track weights only. The green distribution is the final one and includes also the unfolding. The shaded region corresponds to the statistical uncertainty obtained with pseudo experiments and the error bars in the ratio plot include only the statistical uncertainty of the reference particle level spectrum.

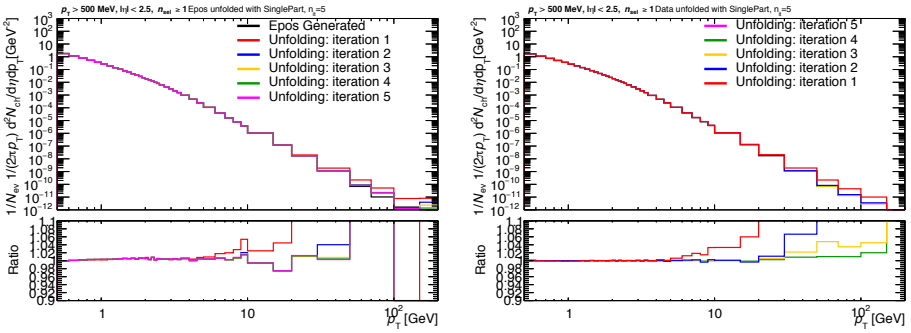


FIGURE 5.48: Iterations in p_T unfolding for EPOS (left) and data (right). The ratio shows difference of various iterations with respect to the last one. The EPOS sample contains less statistics and the closure test has significant statistical uncertainty for high p_T as seen in Figure 5.47.

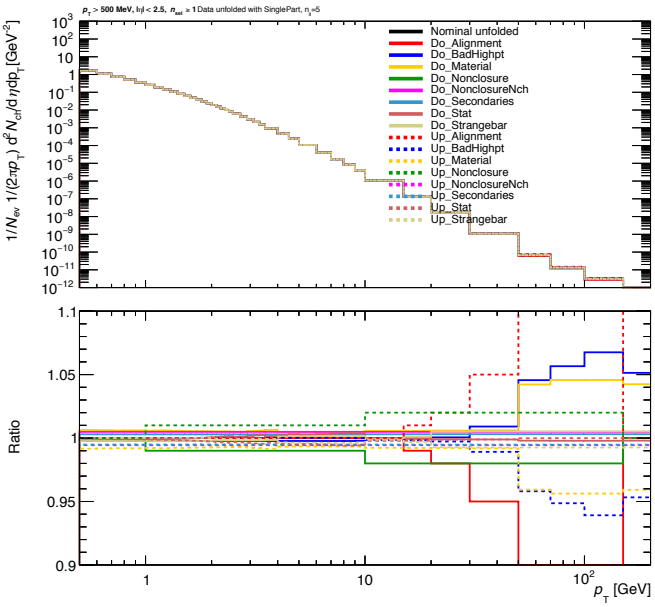


FIGURE 5.49: Contribution of different systematic variations of the tracking efficiency to the uncertainty of the unfolded spectrum.

through the unfolding procedure. The difference with respect to the nominal data particle p_{T} spectrum is taken as a systematic uncertainty.

The contributions of the systematic uncertainties are shown in Figure 5.49 together with the total uncertainty of the p_{T} distribution. There are two statistical uncertainty components: due to data and due to matrix. The statistical uncertainty due to matrix is obtained by smearing the migration matrix using 50 pseudo-experiments. The improvement of the statistic uncertainty using additional single particle samples is seen in Figure 5.50 unfolding. For example, at p_{T} =35GeV, the statistical error of the unfolding reduces from 8% to 1%. The discussion of systematic uncertainties follows.

The final p_{T} distribution is shown in Figure 5.51. It also displays cross-check unfolding data with different generators in addition to PYTHIA 8 A2 + single particle samples. The different unfolded results are consistent other than for EPOS, the difference being due to small statistic to construct the matrix in the MC (only 5 M events).

5.10.6 Corrections to <p_{T} > as a function of n_{ch}

The idea of the correction procedure for the $\langle p_{\mathrm{T}} \rangle$ vs. n_{ch} distribution is to correct separately two components: $\sum_i p_{\mathrm{T}}(i)$ vs. n_{ch} and $\sum_i 1$ vs. n_{ch}. The ratio of these two quantities, which is the $\langle p_{\mathrm{T}} \rangle$ vs. n_{ch} is calculated after all corrections on these terms are applied. The sum is over all tracks and all events; the second sum represents the total number of tracks in each bin of n_{ch}. Each of these distributions, $\sum_i p_{\mathrm{T}}(i)$ and $\sum_i 1$, is corrected in the two steps

- Start with the two distributions as a function of n_{sel} and first correct them on a track-by-track basis applying the appropriate track weights; this track-by-track correction is applied to the distributions of the data and thus does not rely on the p_{T} spectrum of the MC
- Second, apply the same unfolding procedure as for the n_{ch} unfolding to each of the distributions.

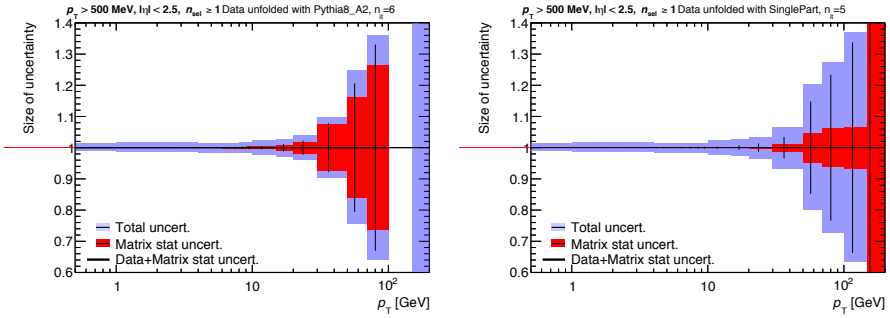


FIGURE 5.50: Comparison of statistical and systematic uncertainties. The error bars show the uncertainty due to data statistics. The red band reflects the uncertainty of the migration matrix, while the blue band is obtained as a sum of the statistic and systematic uncertainties. The two plots illustrate how single particle MC helps to reduce the statistical uncertainty due to matrix. Left: uses PYTHIA 8 A2 MC. Right: PYTHIA 8 A2 + single particle MC included.

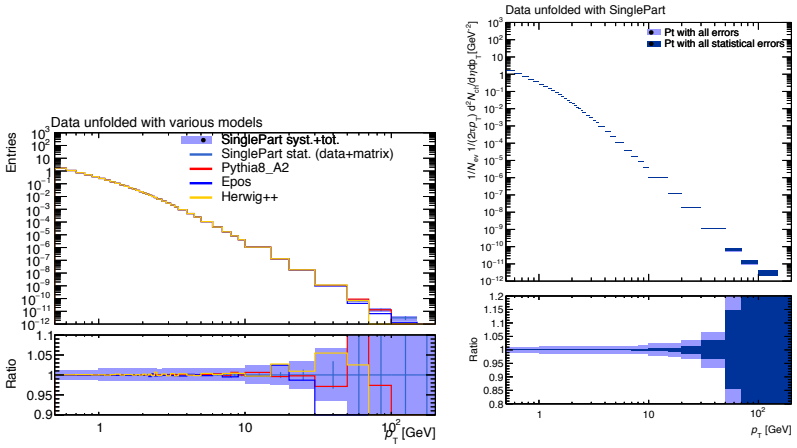


FIGURE 5.51: Left: Cross-check of p_{T} unfolding of the data using a migration matrix constructed with different models. Right: The final unfolded distribution of the data in a wider p_{T} range than reported in the final results. SinglePart in the figure corresponds to PYTHIA 8 A2 + single particle MC included. The light blue area depicts the total uncertainty, while the dark blue in the right figure shows the statistical uncertainty only.

Finally, take the ratio of the two distributions to obtain the corrected $\langle p_{\mathrm{T}} \rangle$ vs n_{ch}. The matrix used for the correction is the one obtained after the final iteration of the n_{ch} unfolding, as described in Sec. 5.10.2, and represents a matrix which would be obtained using MC in which the true n_{ch} spectrum would be the same as in data.

The statistical uncertainty of the result is properly calculated by propagating the uncertainty on the mean of the track-weight corrected $\langle p_{\mathrm{T}} \rangle$ distribution vs. n_{sel} through the unfolding using the migration matrix.

Mathematical details of the procedure First of all, the relations for event and track-wise corrected distributions $\sum_i p_{\mathrm{T}}(i)$ and $\sum_i 1$ (the charged multiplicity in that bin) are written as a function of n_{sel}. They are denoted as the numerator N^{cor} and the denominator D^{cor} for a given track multiplicity n_{sel}. The integration over η in n_{sel} =1 due to vertex efficiency

is not explicitly mentioned for simplicity. The relations read:

$$N^{cor}(n_{\rm sel}) = N_{\rm ev} \int dp_{\rm T} \cdot P(n_{\rm sel}) \cdot \rho_{\rm sel}(p_{\rm T}|n_{\rm sel}) \cdot w_{\rm ev}(n_{recotr}) \cdot w_{\rm trk}(p_{\rm T},\eta) \cdot p_{\rm T} \quad (5.9)$$

$$D^{cor}(n_{\rm sel}) = N_{\rm ev} \int dp_{\rm T} \cdot P(n_{\rm sel}) \cdot \rho_{\rm sel}(p_{\rm T}|n_{\rm sel}) \cdot w_{\rm ev}(n_{recotr}) \cdot w_{\rm trk}(p_{\rm T},\eta). \quad (5.10)$$

$\rho_{sel}(p_{\rm T}|n_{\rm sel})$ is the number of tracks per event per unit of $p_{\rm T}$ in a given $n_{\rm sel}$ bin, $P(n_{\rm sel})$ is the probability that an event is observed with $n_{\rm sel}$ tracks and $w_{\rm trk}(p_{\rm T},\eta)$ and $w_{\rm ev}(n_{recotr})$ are the track correction and event correction weights, respectively. In the same manner as for η and $p_{\rm T}$ distribution corrections, the track weight is parameterized as a function of n_{recotr} through vertex and trigger efficiency corrections and as function of $p_{\rm T}$ and η through the track reconstruction efficiency corrections. Concerning the N^{cor} and D^{cor}, the track correction factor is supposed to correct back the number of selected tracks in a given $p_{\rm T}$ and $n_{\rm sel}$ bin to the number of charged-particles in the same bin. For this to happen, we have to assume that the track-wise correction works independently of $n_{\rm sel}$, or in other words that the following condition is satisfied for all $n_{\rm sel}$ bins:

$$\rho_{ch}(p_{\rm T}|n_{\rm sel}) = \rho_{\rm sel}(p_{\rm T}|n_{\rm sel}) \cdot w_{\rm ev}(n_{recotr}) \cdot w_{\rm trk}(p_{\rm T},\eta), \quad (5.11)$$

where $\rho_{ch}(p_{\rm T}|n_{\rm sel})$ is the number of charged-particles per unit of $p_{\rm T}$ in a given $n_{\rm ch}$ bin. This assumption is, however, not true in reality, because the actual track reconstruction efficiency (not to be confused with the average track reconstruction efficiency) depends slightly on $n_{\rm sel}$ through the resolution of the primary vertex and the impact parameter cut. This is not an issue for the η and $p_{\rm T}$ distributions which are measured inclusively, integrated over $n_{\rm sel}$ but is important for an observable which is measured as a function of $n_{\rm sel}$. In the ideal case, all weights should be parameterized as a function of $n_{\rm sel}$, $p_{\rm T}$ and η.

After applying the unfolding matrix, the numerator and denominator are expressed as function of $n_{\rm ch}$:

$$N^{unf}(n_{\rm ch}) = N_{\rm ev} \int dp_{\rm T} \sum_{n_{\rm sel}} P(n_{\rm ch}|n_{\rm sel}) \cdot P(n_{\rm sel}) \cdot \rho_{\rm ch}(p_{\rm T}|n_{\rm sel}) \cdot p_{\rm T} \quad (5.12)$$

$$D^{unf}(n_{\rm ch}) = N_{\rm ev} \int dp_{\rm T} \sum_{n_{\rm sel}} P(n_{\rm ch}|n_{\rm sel}) \cdot P(n_{\rm sel}) \cdot \rho_{\rm ch}(p_{\rm T}|n_{\rm sel}). \quad (5.13)$$

where $P(n_{\rm ch}|n_{\rm sel})$ is a probability that an event with $n_{\rm ch}$ charged-particles is observed as an event with $n_{\rm sel}$ tracks. The second assumption is that we have to make is that the tracks do not change their $p_{\rm T}$ spectrum when migrated from $n_{\rm sel}$ to $n_{\rm ch}$ bin. This implies that the number of tracks per event per unit of $p_{\rm T}$ in events that are migrated from a given $n_{\rm sel}$ to $n_{\rm ch}$ bin $\rho_{\rm ch}(p_{\rm T}|n_{\rm ch}|n_{\rm sel})$, is given by the number of tracks per event per unit of $p_{\rm T}$ in events in the corresponding $n_{\rm sel}$ bin $\rho_{\rm ch}(p_{\rm T}|n_{\rm sel})$ multiplied by the migration probability $P(n_{\rm ch}|n_{\rm sel})$:

$$\rho_{\rm ch}(p_{\rm T}|n_{\rm ch}|n_{\rm sel}) = P(n_{\rm ch}|n_{\rm sel}) \cdot \rho_{\rm ch}(p_{\rm T}|n_{\rm sel}). \quad (5.14)$$

The final form of the numerator and denominator can then be written as:

$$N^{unf}(n_{\rm ch}) = \int dp_{\rm T} \cdot \rho_{\rm ch}(p_{\rm T}|n_{\rm ch}) \cdot p_{\rm T} \quad (5.15)$$

$$D^{unf}(n_{\rm ch}) = \int dp_{\rm T} \cdot \rho_{\rm ch}(p_{\rm T}|n_{\rm ch}) \quad (5.16)$$

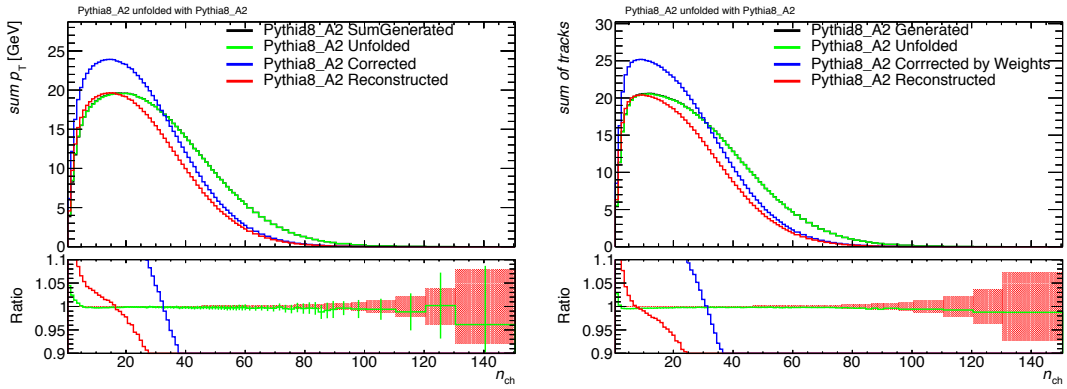


FIGURE 5.52: Closure test for $N^{unf}(n_{\mathrm{ch}})$ and $D^{unf}(n_{\mathrm{ch}})$. PYTHIA 8 A2 is used as the input spectrum as well as the MC to construct the n_{ch} migration matrix. The distributions are normalized to cross-section in mb. This means that they show sum of the p_{T} and sum of the number of tracks in the dataset of $1\mathrm{mb}^{-1}$.

where we have performed the following summation over n_{sel}:

$$\rho_{ch}(p_{\mathrm{T}}|n_{\mathrm{ch}}) = \sum_{n_{\mathrm{sel}}} P(n_{\mathrm{sel}}) \cdot \rho_{\mathrm{ch}}(p_{\mathrm{T}}|n_{\mathrm{ch}}|n_{\mathrm{sel}}) \tag{5.17}$$

Taking the ratio of Eq. 5.15 and Eq. 5.16 we get the $\langle p_{\mathrm{T}} \rangle$ vs. n_{ch}:

$$\langle p_{\mathrm{T}} \rangle (n_{\mathrm{ch}}) = \frac{\int dp_{\mathrm{T}} \cdot \rho_{\mathrm{ch}}(p_{\mathrm{T}}|n_{\mathrm{ch}}) \cdot p_{\mathrm{T}}}{\int dp_{\mathrm{T}} \cdot \rho_{\mathrm{ch}}(p_{\mathrm{T}}|n_{\mathrm{ch}})}. \tag{5.18}$$

Closure test

As described above, the unfolding procedure assumes that

- The tracking efficiency depends only on p_{T} and η and is independent of the track particle multiplicity.
- The p_{T} spectrum of the tracks in events that migrate back from a given n_{sel} bin to a given n_{ch} bin is the same as the p_{T} spectrum of tracks in events in the corresponding n_{sel} bin.

The fact that these assumptions are not valid leads to a non-closure of the corrected distribution that can be tested on the MC. The performance of the correction procedure is summarized in Figure 5.52, which shows the closure test for $N^{unf}(n_{\mathrm{ch}})$ and $D^{unf}(n_{\mathrm{ch}})$ using PYTHIA 8 A2. In Figure 5.53 the closure test of $\langle p_{\mathrm{T}} \rangle (n_{\mathrm{ch}})$ is shown for PYTHIA 8 A2 and EPOS models unfolded with PYTHIA 8 A2 migration matrix. While there are trends observed in the closure of the individual components, they largely cancel in the ratio $\langle p_{\mathrm{T}} \rangle (n_{\mathrm{ch}})$, where a residual non-closure of 2% is observed in n_{sel} =1,2,3 and 1% is observed in n_{sel} =3,4.

Systematic uncertainties

The systematic uncertainty on the $\langle p_{\mathrm{T}} \rangle$ distribution is calculated by modifying the input distributions according to each systematic variation, keeping the unfolding matrix the

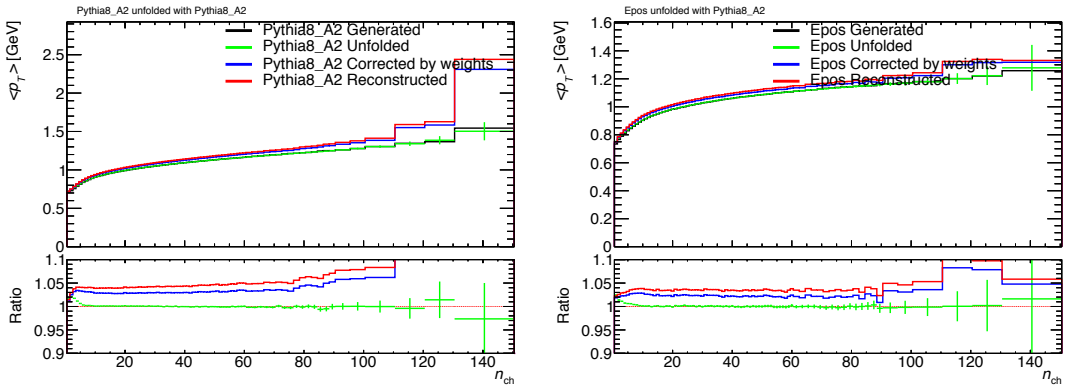


FIGURE 5.53: Closure test for $\langle p_{\mathrm{T}} \rangle (n_{\mathrm{ch}})$. Left: PYTHIA 8 A2 is used as the input spectrum and for migration matrix. Right: EPOS is unfolded with PYTHIA 8 A2.

same. The uncertainty is propagated by varying the track weight w_{trk} according to the uncertainties of tracking efficiency, and, at the same time, changing the number of selected tracks n_{sel} by randomly creating/removing tracks. This random procedure is exactly the same as used to derive the systematic uncertainty of multiplicity distribution, and is described in Section 5.10.2. The measurement of $\langle p_{\mathrm{T}} \rangle$ is very precise, as many effects cancel in the ratio. The size of all propagated systematic uncertainties shown in 5.54 is smaller than 0.5%. An additional systematic uncertainty comes from the non-closure: 0.25% applied across the whole range, except n_{ch} =1,2,3 where it is taken to be 2% and 1% in n_{sel} =3, 4.

5.10.7 Strange Baryons Extrapolation

The fiducial measurement is defined, as mentioned earlier, using a specific charged primary particle selection which excludes strange baryons. However, in order to be able to compare the results with other experiments or previous measurements, the data are also presented for a selection, which includes the strange baryons. These distributions are obtained by extrapolating the nominal unfolded distribution to a more inclusive definition. The extrapolation is performed bin-by-bin, and the size of the extrapolation factors is derived from MC by comparing the MC distribution with/without strange baryons.

As described in section 5.9.1, a study was carried out to find out which generator describes the strange baryon spectra the best. ATLAS MC tunes were compared to the ALICE measurement of strange baryons at 7 TeV using Rivet. The study showed that EPOS describes the strange baryon rate within tens of percent, contrary to PYTHIA tunes which underestimate the strange baryon rate by factor of 2-3. The extrapolation factors are therefore derived using EPOS, while the systematic uncertainty is calculated as the largest difference in the extrapolation factor between EPOS and PYTHIA 8 A2 or PYTHIA 8 MONASH tunes and symmetrised. Because it does not produce the strange baryons at all, QGSJET-II generator is excluded. The statistical uncertainty of the generator used to assess the systematic uncertainty is added in quadrature to the total systematic on the extrapolation factor.

The statistical and systematic uncertainties are propagated to the extrapolated measured data. Note that when the multiplicity distribution is corrected by the bin-by-bin extrapolation factors, it is no longer normalized to one. The difference in the normalization is 0.4% and is neglected.

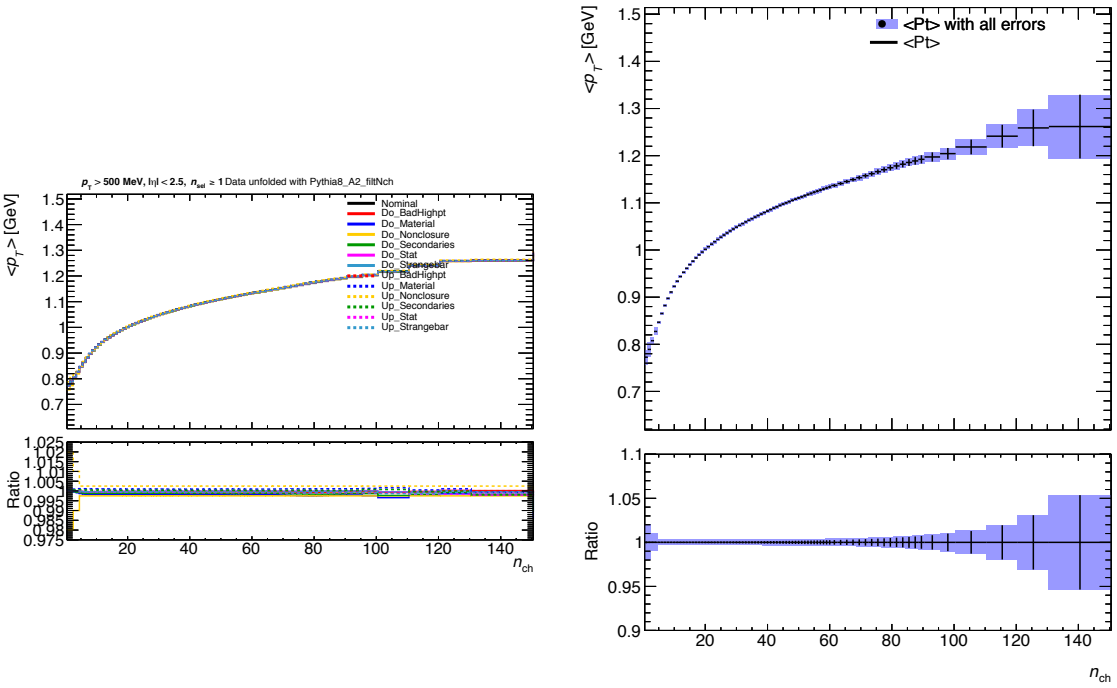


FIGURE 5.54: Left: Contribution of systematic uncertainties to the $\langle p_{\mathrm{T}} \rangle$ distribution. Right: Final unfolded distribution with statistic and total uncertainty.

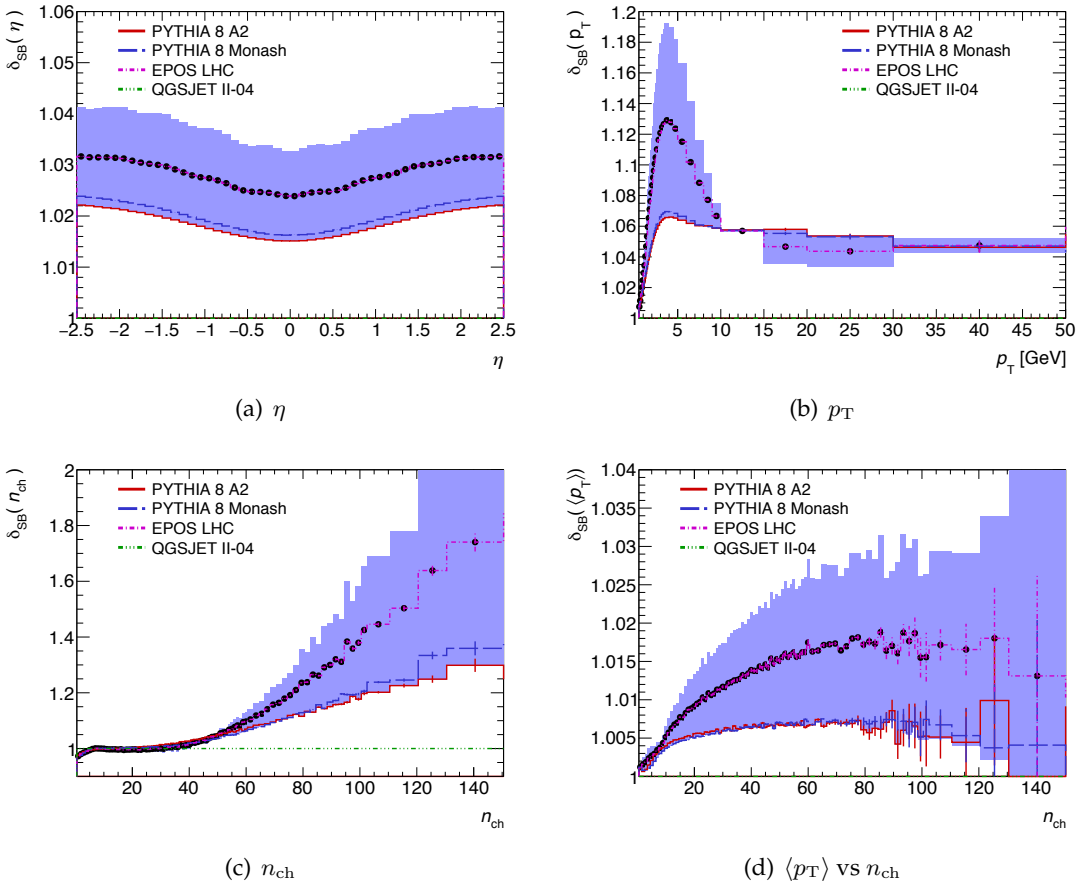
In Figure 5.55, it can be seen that the statistical uncertainty at high multiplicity for the mean-p_{T} vs. multiplicity distribution is considerable. The correction factors were therefore smeared by means of the Root TH1F::Smear() function (Figure). Moreover, the systematic uncertainty is found to be zero for 10 GeV. This is unphysical; there is no good reason to assume that the description of the strange production is better constrained at this particular p_{T}, and a good agreement between generators for high p_{T} is merely a coincidence. In fact, ALICE measured the strange baryons component only up to 5 GeV. A pragmatic approach has been chosen, the systematic uncertainty for $p_{\mathrm{T}} > 10$ GeV is therefore taken as the largest uncertainty in any of the bin above 10 GeV.

5.10.8 Correction Procedure in the Extended Phase Space

Similarly to the nominal phase-space, the data are corrected to obtain inclusive spectra for primary charged-particles satisfying the particle-level phase space requirement. The inefficiencies due to the trigger selection and vertex reconstruction are applied to all distributions as event weights:

$$w_{\mathrm{ev}}(n_{\mathrm{sel}}^{\text{no-z}}, \Delta z_{\mathrm{tracks}}) = \frac{1}{\varepsilon_{\mathrm{trig}}(n_{\mathrm{sel}}^{\text{no-z}})} \cdot \frac{1}{\varepsilon_{\mathrm{vtx}}(n_{\mathrm{sel}}^{\text{no-z}}, \Delta z_{\mathrm{tracks}})}. \tag{5.19}$$

Distributions of the selected tracks are corrected for inefficiencies in the track reconstruction with a track weight using the tracking efficiency ($\varepsilon_{\mathrm{trk}}$) and after subtracting the fractions of fake tracks (f_{fake}), of strange baryons (f_{sb}), of secondary particles (f_{sec}) and of



(a) η (b) p_{T}

(c) n_{ch} (d) $\langle p_{\mathrm{T}} \rangle$ vs n_{ch}

FIGURE 5.55: Strange-baryon extrapolation factors as predicted by EPOS, PYTHIA 8 A2,PYTHIA 8 MONASH, QGSJET-II, as a function of (a) pseudo-rapidity, η, (b) transverse momentum, p_{T}, (c) multiplicity, n_{ch}, and (d) the mean transverse momentum, $\langle p_{\mathrm{T}} \rangle$, versus n_{ch}. The black points display the nominal value of the correction, while the shaded area displays the associated systematic uncertainty. Note:QGSJET-II predicts no strange baryons, and the factors are one.

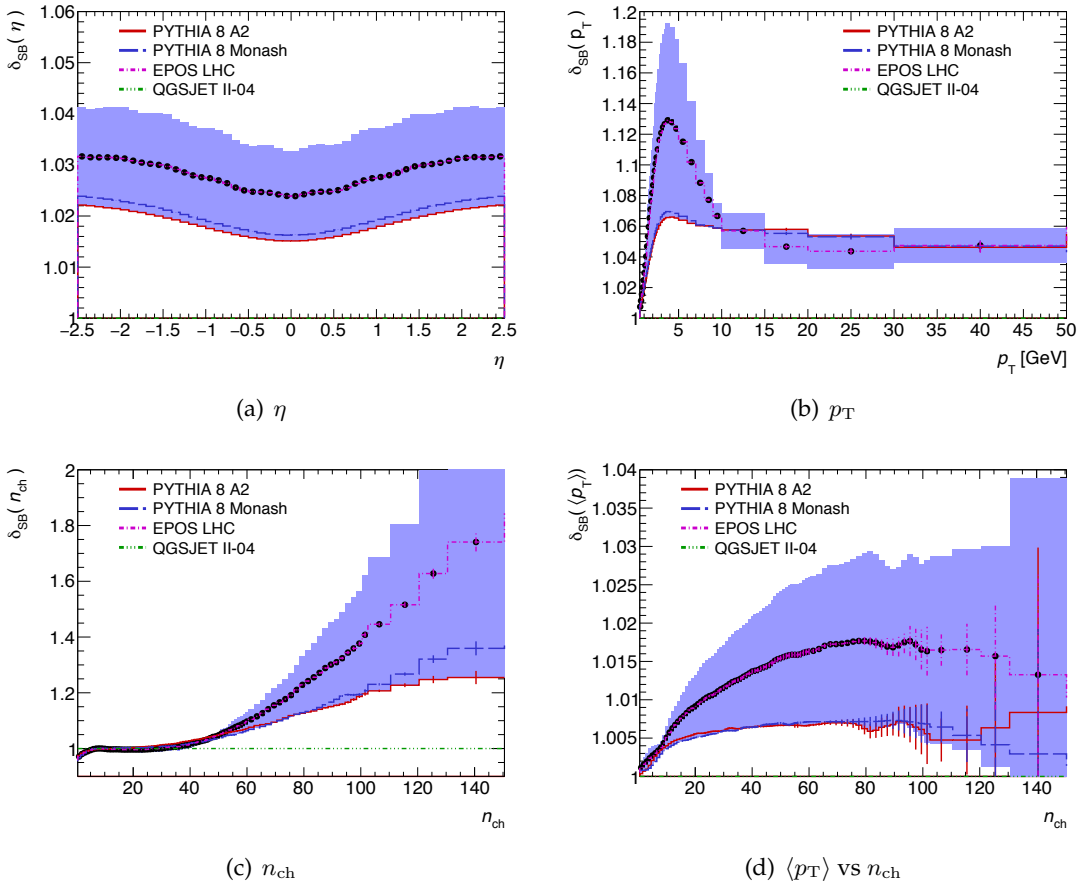
particles outside the kinematic range (f_{okr}):

$$w_{\mathrm{trk}}(p_{\mathrm{T}}, \eta) = \frac{1}{\varepsilon_{\mathrm{trk}}(p_{\mathrm{T}}, \eta)} \cdot [1 - f_{\mathrm{fake}}(p_{\mathrm{T}}, \eta) - f_{\mathrm{sb}}(p_{\mathrm{T}}, \eta) - f_{\mathrm{sec}}(p_{\mathrm{T}}, \eta) - f_{\mathrm{okr}}(p_{\mathrm{T}}, \eta)] \,. \quad (5.20)$$

The event and the track weights are slightly different with respect to those applied in the nominal phase space because of the differences in the vertexing and tracking efficiency as well as in the background components, as described in the previous sections. The general unfolding procedure to obtain the final results is identical to the one described above and used in the nominal phase space. More details can be found in Ref. [81].

5.11 Final Results

The corrected distributions for primary charged-particles in events with $n_{\mathrm{ch}} \geq 1$ in the kinematic range $p_{\mathrm{T}} > 500$ MeV and $|\eta| < 2.5$ are shown in Figure 5.57.



(a) η (b) p_{T}

(c) n_{ch} (d) $\langle p_{\mathrm{T}} \rangle$ vs n_{ch}

FIGURE 5.56: Same as Figure 5.55, with smearing applied and special high-p_{T} systematic treatment as described in the text.

A summary of the main systematic uncertainties affecting the η, p_{T} and n_{ch} distributions is given in Table 5.10. In most regions of all distributions the dominant uncertainty comes from the track reconstruction efficiency. Uncertainties in the $\langle p_{\mathrm{T}} \rangle$ vs. n_{ch} measurement are found in the same way as those in the n_{ch} distribution. The dominant uncertainty is from non-closure which varies from $\pm 2\%$ at low n_{ch} to $\pm 0.5\%$ at high n_{ch}. All other uncertainties largely cancel in the ratio and are negligible. At high n_{ch} the total uncertainty is dominated by the statistical uncertainty.

The results are compared to predictions of models tuned to a wide range of measurements. The measured distributions are presented as inclusive distributions with corrections that rely minimally on the MC model used, in order to facilitate an accurate comparison with predictions.

Figure 5.57(a) shows the multiplicity of charged-particles as a function of pseudorapidity. It must be noted that this distribution was updated with respect to the one presented in Ref. [80] by using the improved data-driven correction described in Section 5.7.1. The mean particle density is roughly constant at 2.9 for $|\eta| < 1.0$ and decreases at higher values of $|\eta|$. EPOS describes the data for $|\eta| < 1.0$, and predicts a slightly larger multiplicity at larger $|\eta|$ values. QGSJET-II and PYTHIA MONASH predict multiplicities that are too large by approximately 15% and 5% respectively. PYTHIA A2 predicts a multiplicity that is 3% too low in the central region, but describes the data well in the forward region.

Figure 5.57(b) [80] shows the charged-particle transverse momentum distribution. EPOS describes the data well over the entire p_{T} spectrum. The PYTHIA tunes describe the data

TABLE 5.10: Summary of systematic uncertainties on the η, p_{T} and n_{ch} distributions. Values for p_{T} and n_{ch} as in **??**.

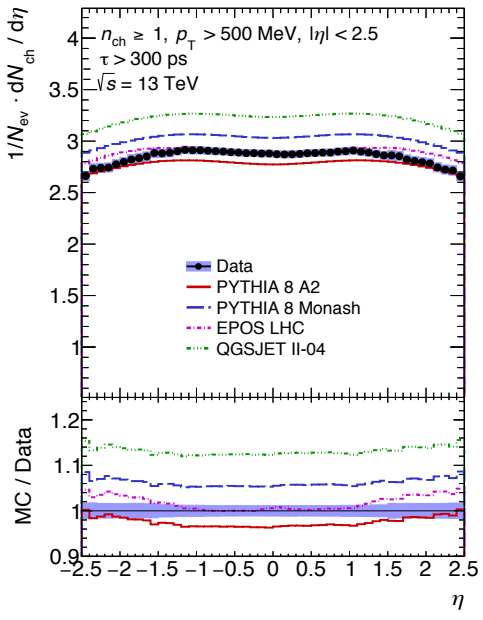
Source	Distribution	Range of values
Track reconstruction efficiency	η	0.5% – 1.8%
	p_{T}	0.7%
	n_{ch}	0% – $^{+17\%}_{-14\%}$
Non-primaries	η	0.5%
	p_{T}	0.5% – 0.9%
	n_{ch}	0% – $^{+10\%}_{-8\%}$
Non-closure	η	0.7%
	p_{T}	0% – 2%
	n_{ch}	0% – 4%
p_{T}-bias	p_{T}	0% – 5%
High-p_{T}	p_{T}	0% – 1%

reasonably well, but are slightly above the data in the high-p_{T} region. QGSJET-II gives a poor prediction over the entire spectrum, overshooting the data in the low-p_{T} region and undershooting it in the high-p_{T} region.

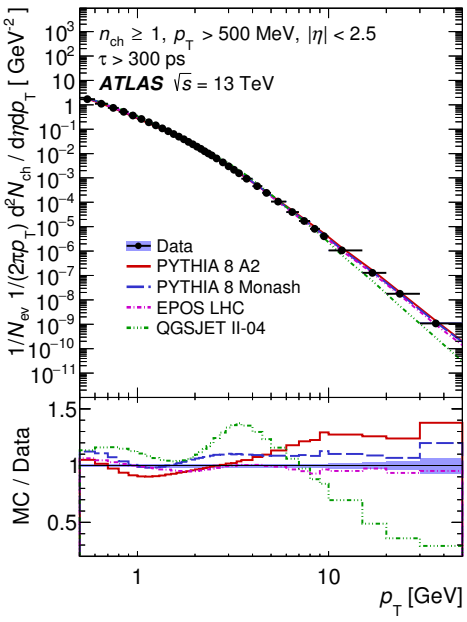
Figure 5.57(c) [80] shows the charged-particle multiplicity distribution. The high-n_{ch} region has significant contributions from events with numerous MPI. PYTHIA A2 describes the data in the region $n_{\mathrm{ch}} < 50$, but predicts too few events at larger n_{ch} values. PYTHIA MONASH, EPOS and QGSJET-II describe the data reasonably well in the region $n_{\mathrm{ch}} < 30$ but predict too many events in the mid-n_{ch} region, with PYTHIA MONASH and EPOS predicting too few events in the region $n_{\mathrm{ch}} > 100$ while QGSJET-II continues to be above the data.

Figure 5.57(d) [80] shows the mean transverse momentum versus the charged-particle multiplicity. The $\langle p_{\mathrm{T}} \rangle$ rises with n_{ch}, from 0.8 to 1.2 GeV. This increase is expected due to colour coherence effects being important in dense parton environments and is modelled by a colour reconnection mechanism in PYTHIA or by the hydrodynamical evolution model used in EPOS. If the high-n_{ch} region is assumed to be dominated by events with numerous MPI, without colour coherence effects the $\langle p_{\mathrm{T}} \rangle$ is approximately independent of n_{ch}. Including colour coherence effects leads to fewer additional charged-particles produced with every additional MPI, with an equally large p_{T} to be shared among the produced hadrons [178]. EPOS predicts a slightly lower $\langle p_{\mathrm{T}} \rangle$, but describes the dependence on n_{ch} very well. The PYTHIA tunes predict a steeper rise of $\langle p_{\mathrm{T}} \rangle$ with n_{ch} than the data, predicting lower values in the low-n_{ch} region and higher values in the high-n_{ch} region. QGSJET-II predicts a $\langle p_{\mathrm{T}} \rangle$ of ~ 1 GeV, with very little dependence on n_{ch}; this is expected as it contains no model for colour coherence effects. In summary, EPOS and the PYTHIA tunes describe the data most accurately, with EPOS reproducing the η and p_{T} distributions and the $\langle p_{\mathrm{T}} \rangle$ vs. n_{ch} the best and PYTHIA A2 describing the multiplicity the best in the low- and mid-n_{ch} regions. QGSJET-II provides an inferior description of the data.

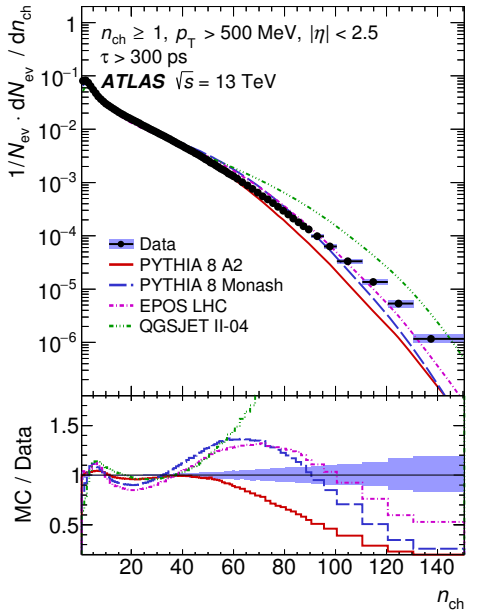
The mean number of primary charged-particles in the central region is computed by averaging over $|\eta| < 0.2$ to be 2.874 $\pm$ 0.001 (stat) $\pm$ 0.033 (sys). This measurement is then corrected for the contribution from strange baryons as shown in Figure 5.58 [80] and compared to previous measurements [166] at different $\sqrt{s}$ values in Figure 5.59 together with the MC predictions. The correction factor for strange baryons depends on the MC model used and is found to be 1.0241 ± 0.0003 (EPOS), 1.0150 ± 0.0004 (PYTHIA MONASH) and 1.0151 ± 0.0002 (PYTHIA A2), where the uncertainties are statistical. QGSJET-II does not include charged strange baryons. The prediction from EPOS is used to perform the



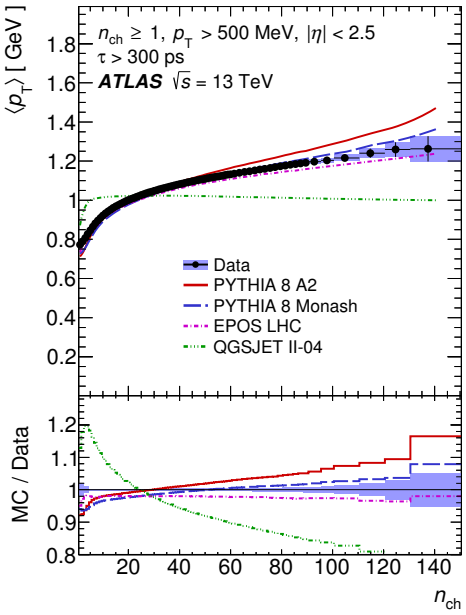
(a)



(b) Ref. [80].



(c) Ref. [80].



(d) Ref. [80].

FIGURE 5.57: Primary-charged-particle multiplicities as a function of (a) pseudorapidity, η, and (b) transverse momentum, p_{T}; (c) the multiplicity, n_{ch}, distribution and (d) the mean transverse momentum, $\langle p_{\mathrm{T}} \rangle$, versus n_{ch} in events with $n_{\mathrm{ch}} \geq 1$, $p_{\mathrm{T}} > 500$ MeV and $|\eta| < 2.5$. The dots represent the data and the curves the predictions from different MC models. The x-value in each bin corresponds to the bin centroid. The vertical bars represent the statistical uncertainties, while the shaded areas show statistical and systematic uncertainties added in quadrature. The bottom panel in each figure shows the ratio of the MC simulation to data. Since the bin centroid is different for data and simulation, the values of the ratio correspond to the averages of the bin content.

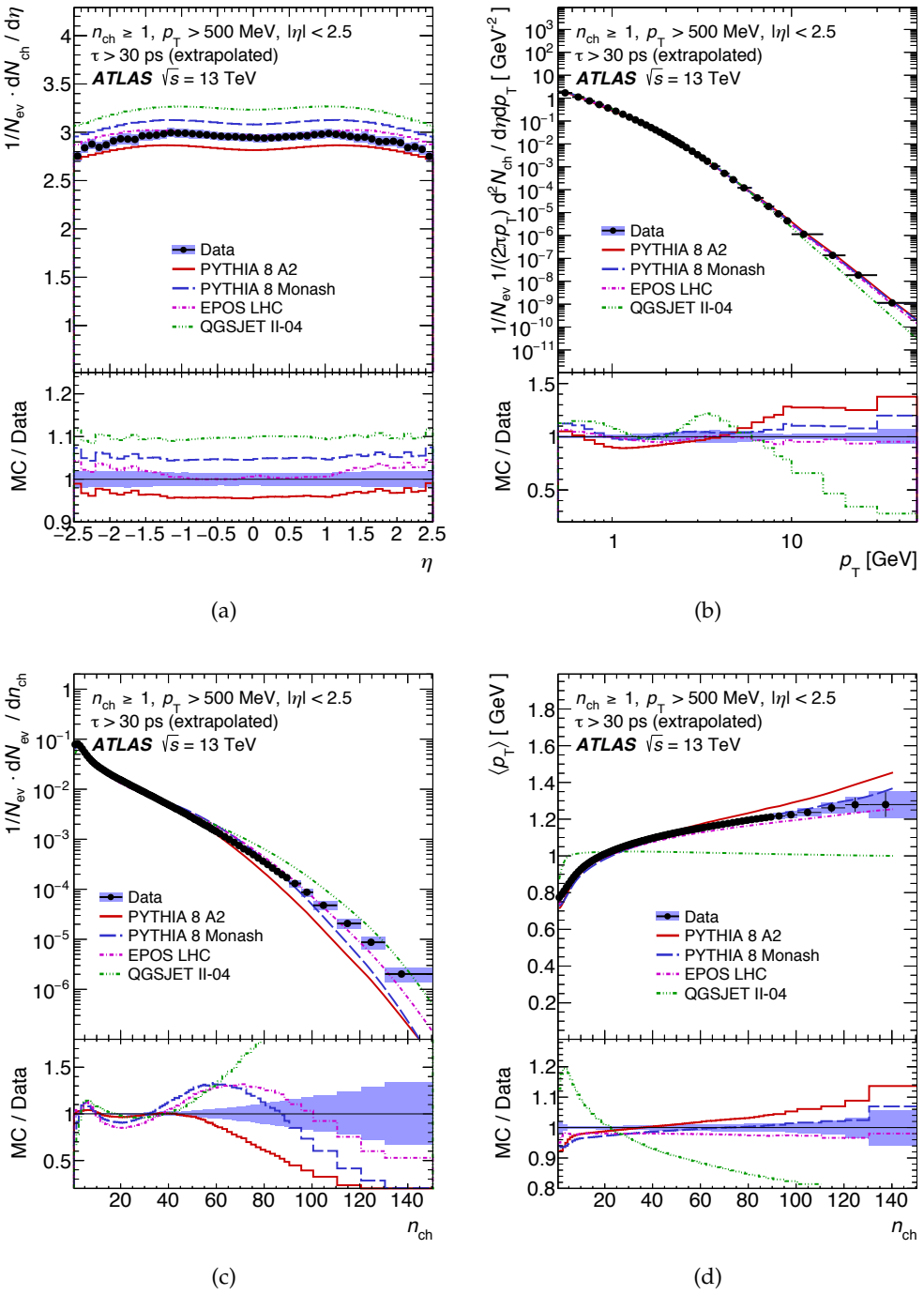


FIGURE 5.58: Charged-particle multiplicities as a function of (a) pseudorapidity, η, and (b) transverse momentum, p_{T}, (c) the multiplicity, n_{ch}, distribution and (d) the mean transverse momentum, $\langle p_{\mathrm{T}} \rangle$, versus n_{ch} for events with $n_{\mathrm{ch}} \geq 1$, $p_{\mathrm{T}} > 500$ MeV and $|\eta| < 2.5$. The results have been extrapolated to include charged-particles with a mean lifetime $30 < \tau < 300$ ps (charged strange baryons). The extrapolation is taken from EPOS, with an uncertainty coming from the statistics on the sample combined in quadrature with the symmetrised maximum difference in the extrapolation when performed with PYTHIA MONASH or PYTHIA A2. The dots represent the data and the curves the predictions from different MC models. The x-value in each bin corresponds to the bin centroid. The vertical bars represent the statistical uncertainties, while the shaded areas show statistical and systematic uncertainties added in quadrature. The bottom inserts show the ratio of the MC over the data. Since the bin centroid is different for data and MC, the values of the ratio correspond to the averages of the bin content. Ref. [80].

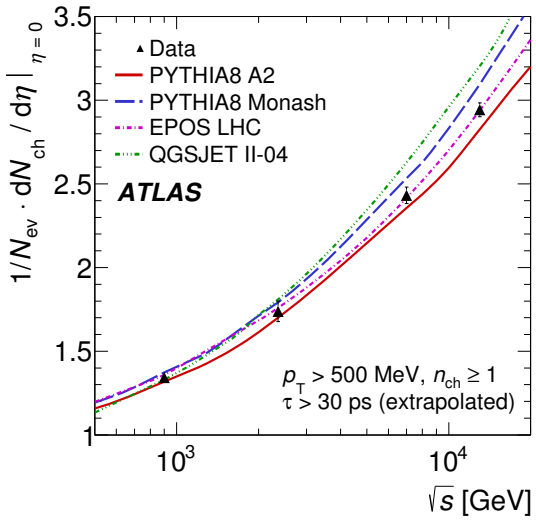


FIGURE 5.59: The average primary-charged-particle multiplicity in pp interactions per unit of pseudorapidity, η, for $|\eta| < 0.2$ as a function of the centre-of-mass energy. The values at centre-of-mass energies other than 13 TeVare taken from Ref. [166]. Charged strange baryons are included in the definition of primary particles. The data are compared to various particle-level MC predictions. The vertical error bars on the data represent the total uncertainty. Ref. [80].

extrapolation and the deviation from the PYTHIA MONASH prediction is taken as a systematic uncertainty and symmetrised to give 1.024 ± 0.009. The mean number of primary charged-particles increases by a factor of 2.2 when $\sqrt{s}$ increases by a factor of about 14 from 0.9 TeV to 13 TeV. EPOS and PYTHIA A2 describe the dependence on $\sqrt{s}$ very well, while PYTHIA MONASH and QGSJET-II predict a steeper rise in multiplicity with $\sqrt{s}$.

5.11.1 Final Results in the Extended Phase Space

The measured charged-particle multiplicities in events containing at least two charged particles with $p_{\mathrm{T}} > 100$ MeV and $|\eta| < 2.5$ are shown in Figure 5.60 [80].

A summary of the systematic uncertainties, are larger than the statistical uncertainties, is given in Table 5.11 [81] for all observables. The dominant uncertainty is due to material effects on the track reconstruction efficiency. Uncertainties due to imperfect detector alignment are taken into account and are less than 5% at the highest track p_{T} values. In addition, resolution effects on the transverse momentum can result in low-p_{T} particles being reconstructed as high-p_{T} tracks. All these effects are considered as systematic uncertainty on the track reconstruction. The track background uncertainty is dominated by systematic effects in the estimation of the contribution from secondary particles. The track reconstruction efficiency determined in simulation can differ from the one in data if the p_{T} spectrum is different for data and simulation, as the efficiency depends strongly on the track p_{T}. This effect can alter the number of primary charged-particles and is taken into account as a systematic uncertainty on the multiplicity distribution and $\langle p_{\mathrm{T}} \rangle$ vs n_{ch}. The non-closure systematic uncertainty is estimated from differences in the unfolding results using PYTHIA A2 and EPOS simulations. For this, all combinations of these MC generators are used to simulate the distribution and the input to the unfolding.

The corrected data are compared to predictions from various generators. Figure 5.60(a) shows the charged-particle multiplicity as a function of the pseudorapidity η. PYTHIA MONASH, EPOS and QGSJET-II give a good description for $|\eta| < 1.5$. The prediction from

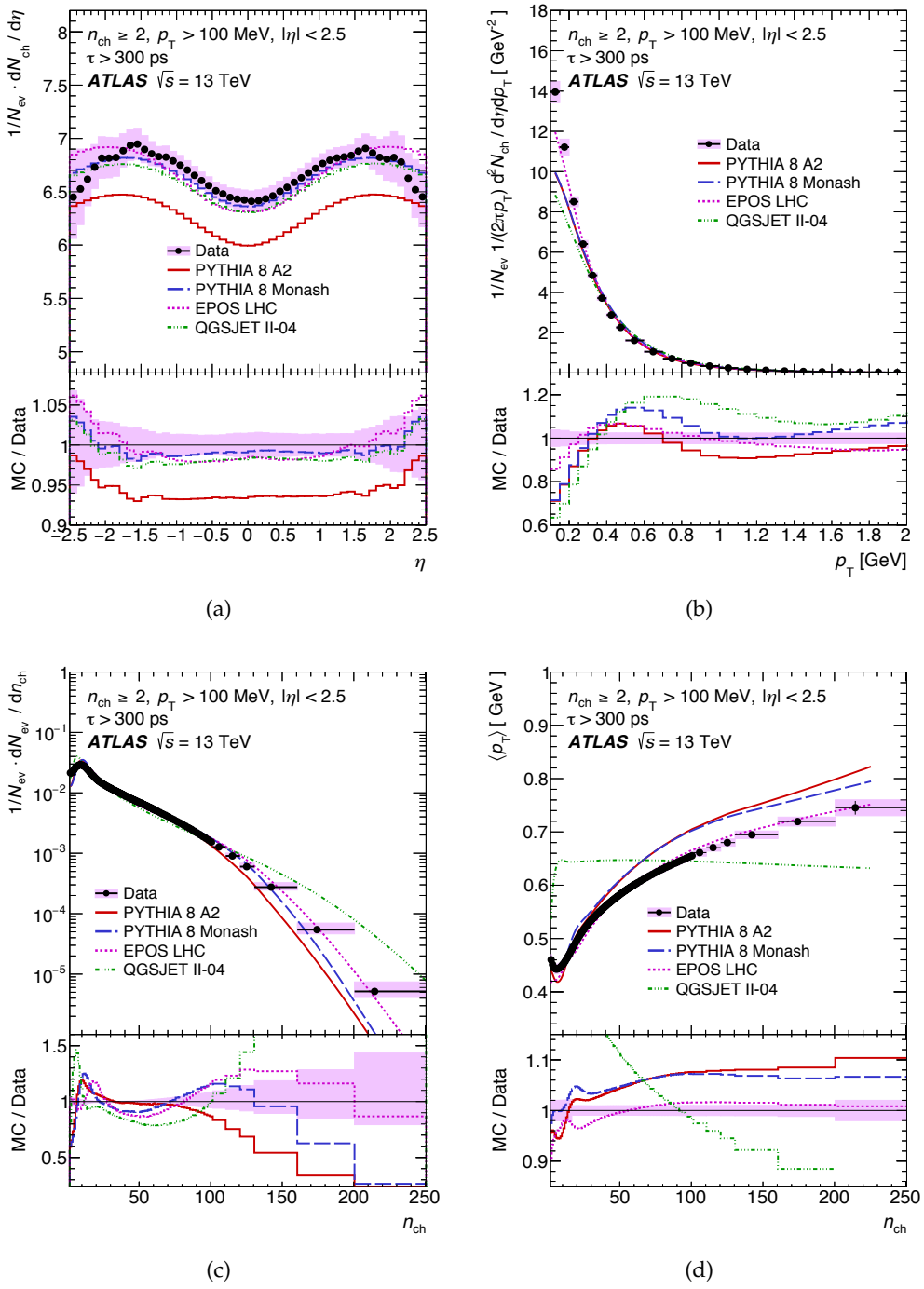


FIGURE 5.60: Primary charged-particle multiplicities as a function of (a) pseudorapidity η and (b) transverse momentum p_{T}, (c) the primary charged-particle multiplicity n_{ch} and (d) the mean transverse momentum $\langle p_{\mathrm{T}} \rangle$ versus n_{ch} for events with at least two primary charged particles with $p_{\mathrm{T}} > 100\mathrm{MeV}$ and $|\eta| < 2.5$, each with a lifetime $\tau > 300\,\mathrm{ps}$. The black dots represent the data and the coloured curves the different MC model predictions.The vertical bars represent the statistical uncertainties, while the shaded areas show statistical and systematic uncertainties added in quadrature. The lower panel in each figure shows the ratio of the MC simulation to data. As the bin centroid is different for data and simulation, the values of the ratio correspond to the averages of the bin content. Ref. [81].

TABLE 5.11: Summary of the systematic uncertainties in the η, p_{T}, n_{ch} and $\langle p_{\mathrm{T}} \rangle$ vs. n_{ch} observables. The uncertainties are given at the minimum and the maximum of the phase space. Ref. [81]

Distribution	$\frac{1}{N_{\mathrm{ev}}} \cdot \frac{\mathrm{d}N_{\mathrm{ch}}}{\mathrm{d}\lvert\eta\rvert}$	$\frac{1}{N_{\mathrm{ev}}} \cdot \frac{1}{2\pi p_{\mathrm{T}}} \cdot \frac{\mathrm{d}^2 N_{\mathrm{ch}}}{\mathrm{d}\eta \mathrm{d}p_{\mathrm{T}}}$	$\frac{1}{N_{\mathrm{ev}}} \cdot \frac{\mathrm{d}N_{\mathrm{ev}}}{\mathrm{d}n_{\mathrm{ch}}}$	$\langle p_{\mathrm{T}} \rangle$ vs. n_{ch}
Range	0–2.5	0.1–50 GeV	2–250	0–160 GeV
Track reconstruction	1%–7%	1%–6%	0%–$^{+38\%}_{-20\%}$	0%–0.7%
Track background	0.5%	0.5%–1%	0%–$^{+7\%}_{-1\%}$	0%–0.1%
p_{T} spectrum	–	–	0%–$^{+3\%}_{-9\%}$	0%–$^{+0.3\%}_{-0.1\%}$
Non-closure	0.4%–1%	1%–3%	0%–4%	0.5%–2%

PYTHIA A2 has the same shape as predictions from the other generators, but lies below the data. The charged-particle transverse momentum is shown in Figure 5.60(b). EPOS describes the data well for $p_{\mathrm{T}} > 300$ MeV. For $p_{\mathrm{T}} < 300$ MeV, the data are underestimated by up to 15%. The other generators show similar mismodelling at low momentum but with larger discrepancies up to 35% for QGSJET-II. In addition, they mostly overestimate the charged-particle multiplicity for $p_{\mathrm{T}} > 400$ MeV; PYTHIA A2 overestimates only in the intermediate p_{T} region and underestimates the data slightly for $p_{\mathrm{T}} > 800$ MeV.

Figure 5.60(c) shows the charged-particle multiplicity. Overall, the form of the measured distribution is reproduced reasonably by all models. PYTHIA A2 describes the data well for $30 < n_{\mathrm{ch}} < 80$, but underestimates it for higher n_{ch}. For $30 < n_{\mathrm{ch}} < 80$, PYTHIA MONASH, EPOS and QGSJET-II underestimate the data by up to 20%.
PYTHIA MONASH and EPOS overestimate the data for $n_{\mathrm{ch}} > 80$ and drop below the measurement in the high-n_{ch} region, starting from $n_{\mathrm{ch}} > 130$ and $n_{\mathrm{ch}} > 200$ respectively. QGSJET-II overestimates the data significantly for $n_{\mathrm{ch}} > 100$. The mean transverse momentum versus the primary charged-particle multiplicity is shown in Figure 5.60(d). It increases towards higher n_{ch}, as modelled by a colour reconnection mechanism in PYTHIA and by the hydrodynamical evolution model in EPOS. The QGSJET-II generator, which has no model for colour coherence effects, describes the data poorly. For low n_{ch}, PYTHIA A2 and EPOS underestimate the data, where PYTHIA MONASH agrees within the uncertainties. For higher n_{ch} all generators overestimate the data, but for $n_{\mathrm{ch}} > 40$, there is a constant offset for both PYTHIA tunes, which describe the data to within 10%. EPOS describes the data reasonably well and to within 2%. The mean number of primary charged-particles per unit pseudorapidity in the central η region is measured to be 6.422 ± 0.096, by averaging over $|\eta| < 0.2$; the quoted error is the systematic uncertainty, the statistical uncertainty is negligible. In order to compare with other measurements, all measurements are corrected for the contribution from strange baryons (and therefore extrapolated to primary charged-particles with $\tau > 30$ ps) by a correction factor of 1.0121 ± 0.0035. Results can be seen in Figure 5.61. The central value is taken from EPOS; the systematic uncertainty is taken from the difference between EPOS and PYTHIA A2 (the largest difference was observed between EPOS and PYTHIA A2) and the statistical uncertainty is negligible.

The mean number of primary charged-particles after the correction is 6.500 ± 0.099. This result is compared to previous measurements [80, 166, 167] at different $\sqrt{s}$ values in Figure 5.62. The predictions from EPOS and PYTHIA MONASH match the data well. For PYTHIA A2, the match is not as good as was observed when measuring particles with $p_{\mathrm{T}} >$ 500 MeV [80].

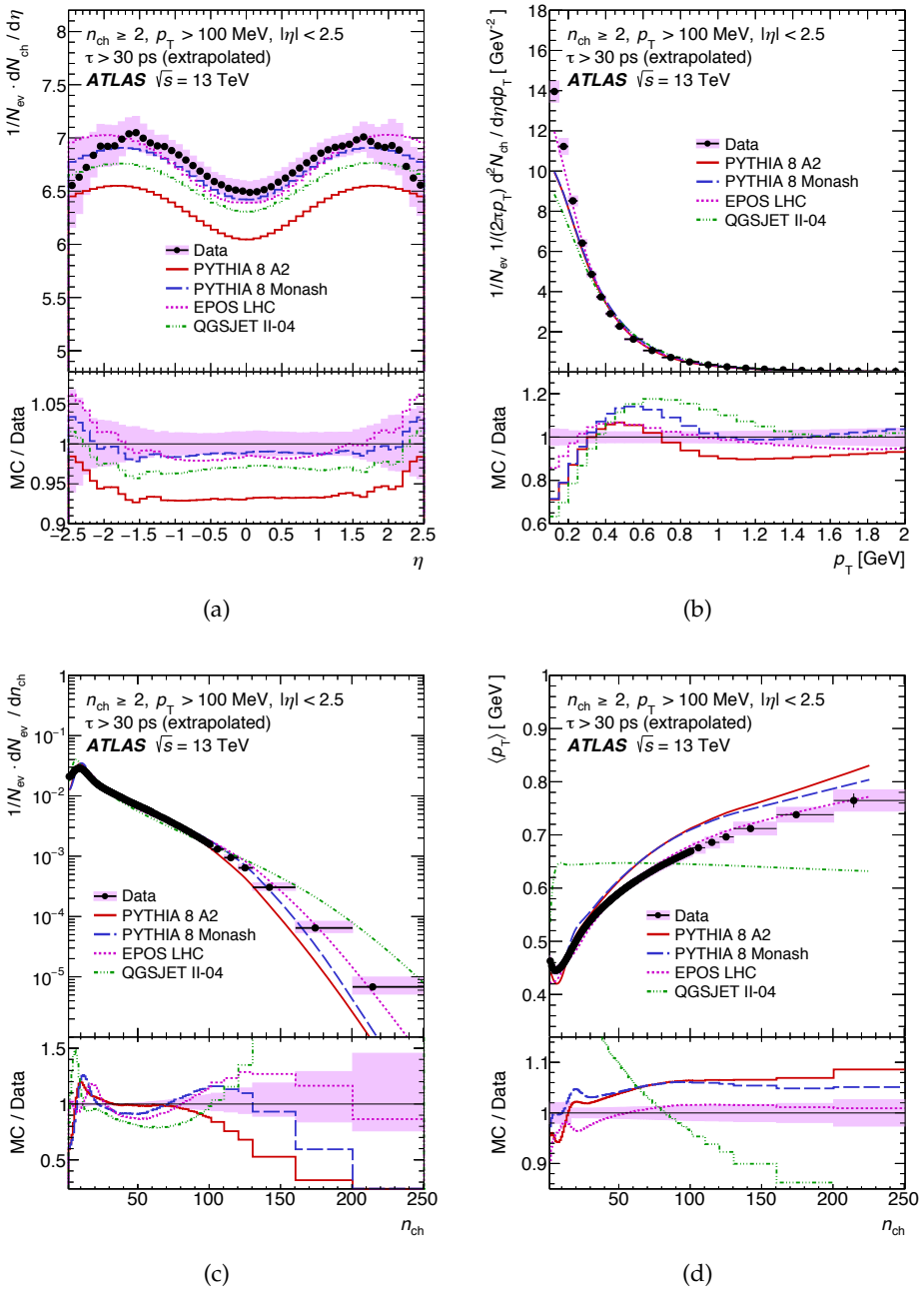


FIGURE 5.61: Primary charged-particle multiplicities as a function of (a) pseudorapidity η and (b) transverse momentum p_{T}, (c) the primary charged-particle multiplicity n_{ch} and (d) the mean transverse momentum $\langle p_{\mathrm{T}} \rangle$ versus n_{ch} for events with at least two primary charged-particles with $p_{\mathrm{T}} > 100\mathrm{MeV}$ and $|\eta| < 2.5$. The results have been extrapolated to include charged strange baryons (charged particles with a mean lifetime of $30 < \tau < 300$ ps). The extrapolation factor is taken from EPOS. The statistical uncertainty on the extrapolation factor from EPOS is added quadratically with the systematic uncertainty which is determined by taking the difference to the extrapolation factor from PYTHIA A2. The black dots represent the data and the coloured curves the different MC model predictions. The vertical bars represent the statistical uncertainties, while the shaded areas show statistical and systematic uncertainties added in quadrature. The lower panel in each figure shows the ratio of the MC simulation to data. As the bin centroid is different for data and simulation, the values of the ratio correspond to the averages of the bin content. Ref. [81].

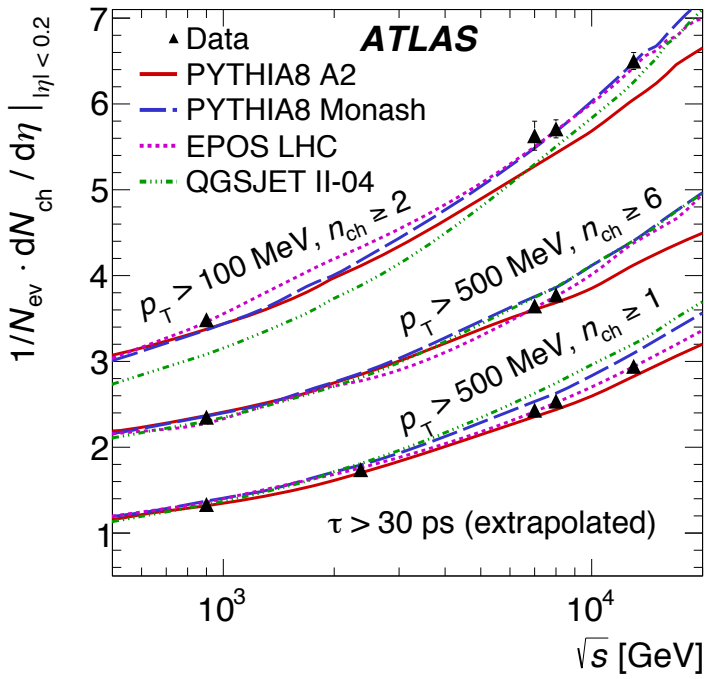


FIGURE 5.62: The average primary charged-particle multiplicity in pp interactions per unit of pseudorapidity η for $|\eta| < 0.2$ as a function of the centre-of-mass energy $\sqrt{s}$. The values for the other pp centre-of-mass energies are taken from previous ATLAS analyses [80, 166, 167]. The value for particles with $p_{\mathrm{T}} > 500$, MeV for a $\sqrt{s} = 13$, TeV is taken from Ref. [80]. The results have been extrapolated to include charged strange baryons (charged-particles with a mean lifetime of $30 < \tau < 300$ ps). The data are shown as black triangles with vertical errors bars representing the total uncertainty. They are compared to various MC predictions which are shown as coloured lines. Ref. [81].

5.12 Results in a common phase space

The corrected distributions for primary charged-particles in events with $n_{\mathrm{ch}} \geq 1$ in the kinematic range $p_{\mathrm{T}} > 500$ MeV and $|\eta| < 0.8$ are shown in Figure 5.63 [80]. This is the phase space that is common to the ATLAS, CMS and ALICE experiments. The method used to correct the distributions and obtain the systematic uncertainties is exactly the same as that used for the results with $|\eta| < 2.5$, but obtained using the $|\eta| < 0.8$ selection, thus no correction is applied to account for the imprecise description of the material in forward η region. Figure 5.63(a) shows the primary-charged-particle multiplicity as a function of pseudorapidity, where the mean particle density is roughly 3.5, larger than in the main phase space due to the tighter restriction of at least one primary charged-particle with $|\eta| < 0.8$. The p_{T} and n_{ch} distributions are shown in Figures 5.63(b) and 5.63(c) respectively and the $\langle p_{\mathrm{T}} \rangle$ as a function of n_{ch} is shown in Figure 5.63(d). The level of agreement between the data and MC generator predictions follows the same pattern as seen in the main phase space.

5.13 Comparison with other MC generators

In this section, comparisons with some additional generator and tune predictions are shown in order to give a more comprehensive overview of the generator behaviour, even though the Author of this thesis did not work directly on those studies. In the original publications [80, 81] as well as in this thesis, data are compared with the predictions given by the generators listed in 5.3. In this section, instead, the data are compared to PYTHIA 6 [179] AMBT2B [41] (with CTEQ6L1 [47] PDF) and Perugia2011 [180] (with CTEQ6L1 PDF) tunes, PYTHIA 8 A14 [181] tune and HERWIG++ [39] UE-EE-4 tune (with MRSTLO [182] PDF).

PYTHIA 6 as well as PYTHIA 8 and HERWIG++ are leading-logarithm parton shower models matched to leading-order matrix element, including multiple parton–parton interactions (MPI). The PYTHIA 6 generator, similarly to PYTHIA 8 , use a p_{T}-ordered or virtuality ordered shower, and a hadronization model based on the fragmentation of color strings. The HERWIG++ generator implements a cluster hadronisation scheme with parton showering ordered by emission angle. The PYTHIA 8 generator adds to the established PYTHIA 6 MPI model by interleaving not only the ISR emission sequence with the MPI scatters, but also the FSR emission. All three processes therefore compete against each other for emission phase space in the resulting evolution. Different settings of model parameters, optimised, or tuned, to reproduce the existing experimental data, have been used for the MC generators.

The same selection as the one used in the nominal phase space at 13 TeV is applied. The results can be seen in Figure 5.64.

5.14 Conclusion

Primary-charged-particle multiplicity measurements with the ATLAS detector using pp collisions delivered by the LHC at $\sqrt{s} = 13$ TeV were presented. From a data sample corresponding to an integrated luminosity of 151 $\mu\mathrm{b}^{-1}$, nearly nine million inelastic interactions with at least one reconstructed track with $|\eta| < 2.5$ and $p_{\mathrm{T}} > 500$ MeV are analysed. The results highlight clear differences between MC models and the measured distributions. Among the models considered EPOS reproduces the data the best, PYTHIA A2 and MONASH give reasonable descriptions of the data and QGSJET-II provides the worst description of the data. The charged-particle multiplicities were also studied in an

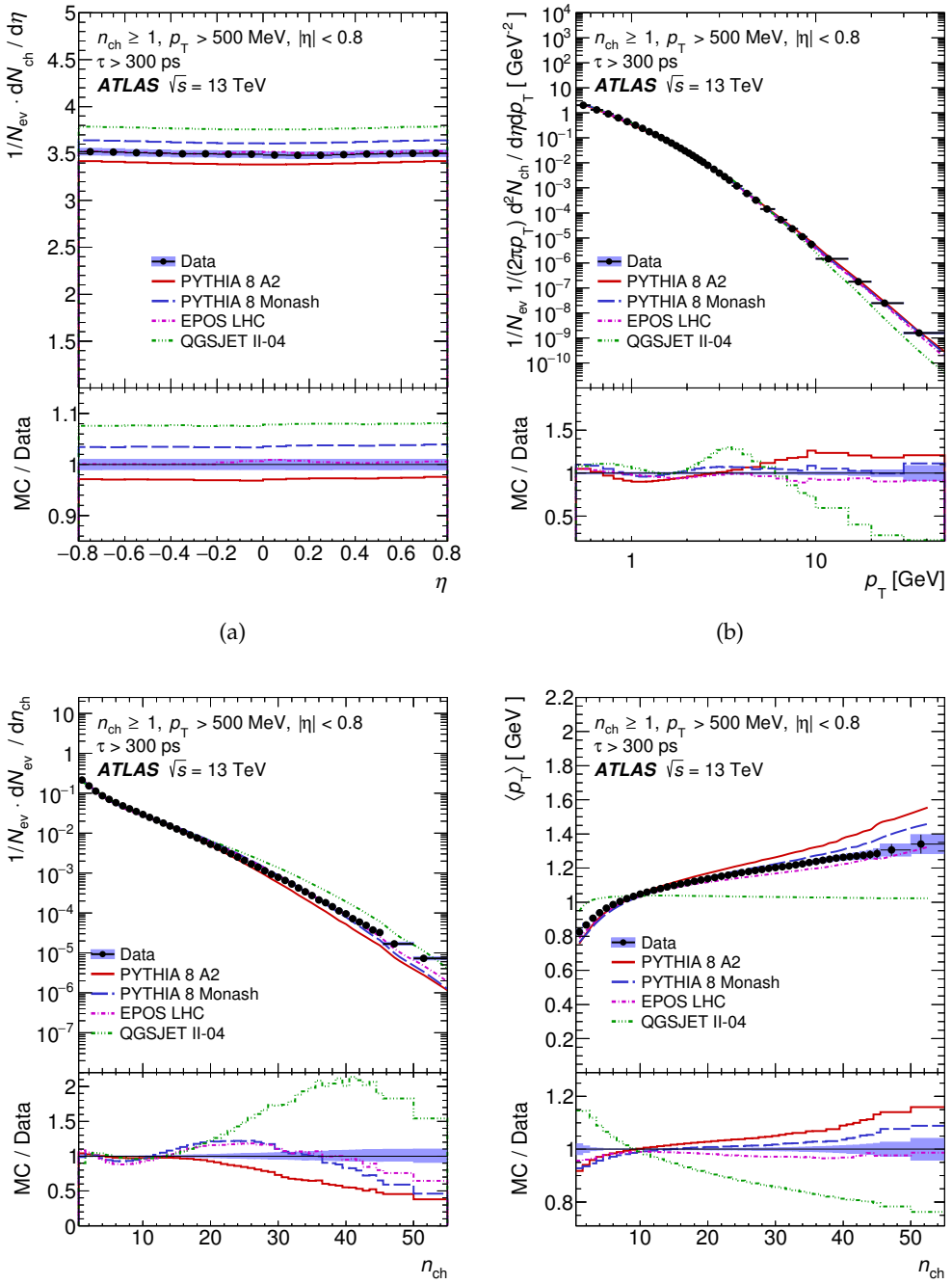


FIGURE 5.63: Primary-charged-particle multiplicities as a function of (a) pseudorapidity, η, and (b) transverse momentum, p_{T}; (c) the multiplicity, n_{ch}, distribution and (d) the mean transverse momentum, $\langle p_{\mathrm{T}} \rangle$, versus n_{ch} in events with $n_{\mathrm{ch}} \geq 1$, $p_{\mathrm{T}} > 500$ MeV and $|\eta| < 0.8$. The dots represent the data and the curves the predictions from different MC models. The x-value in each bin corresponds to the bin centroid. The vertical bars represent the statistical uncertainties, while the shaded areas show statistical and systematic uncertainties added in quadrature. The bottom panel in each figure shows the ratio of the MC simulation over the data. Since the bin centroid is different for data and simulation, the values of the ratio correspond to the averages of the bin content. Ref. [80].

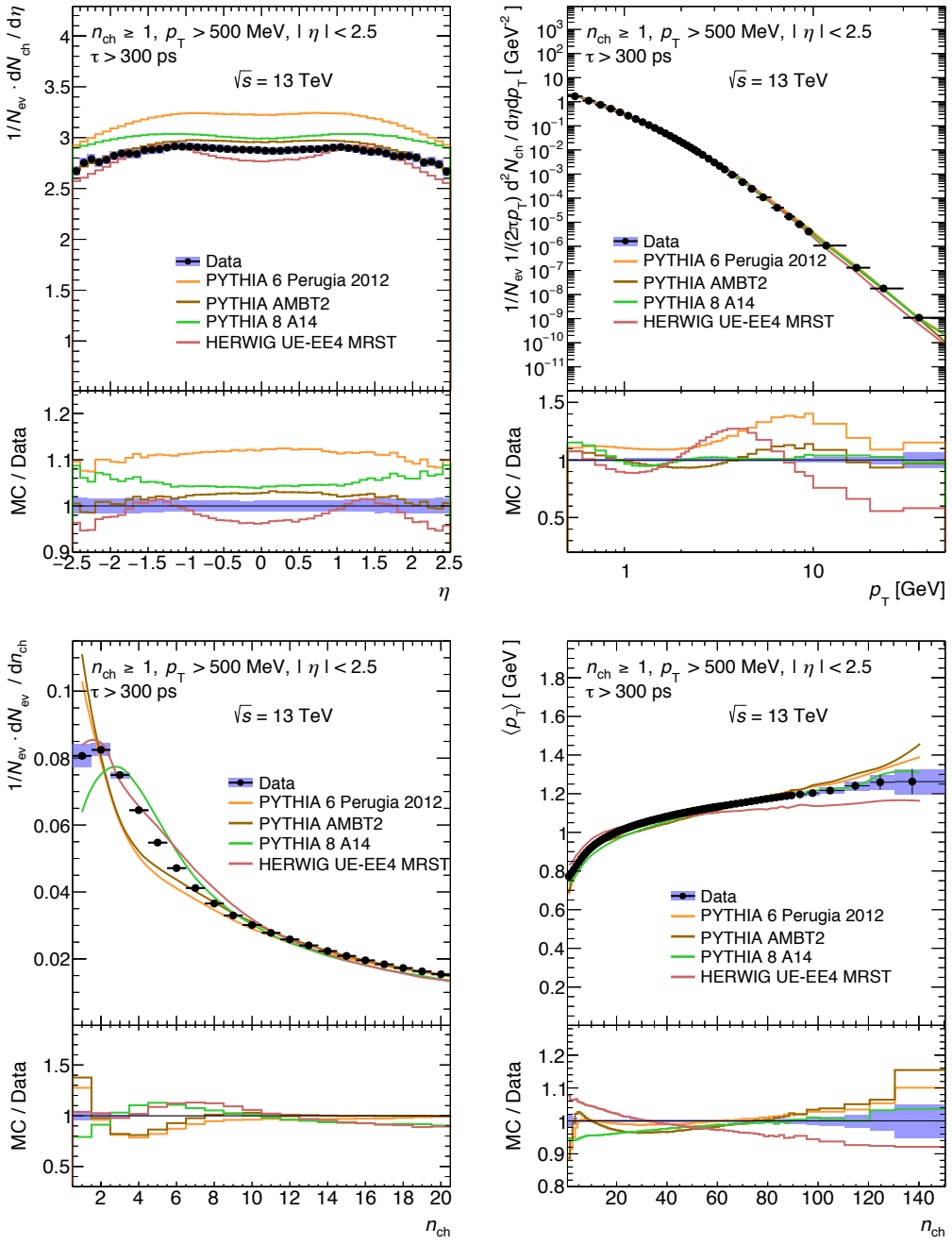


FIGURE 5.64: Charged-particle multiplicities as a function of (a) pseudorapidity, η, and (b) transverse momentum, p_{T}, (c) the multiplicity, n_{ch}, distribution and (d) the mean transverse momentum, $\langle p_{\mathrm{T}} \rangle$, versus n_{ch} for events with $n_{\mathrm{ch}} \geq 1$, $p_{\mathrm{T}} > 500$ MeV and $|\eta| < 2.5$ in a comparison with an extra set on Monte Carlo Event Generators.

extended and a reduced phase space. The conclusions are similar to those of the nominal phase space, with a better behaviour of the EPOS generator even more clear.

These results have been already used by the ATLAS Collaboration in order to produce a new PYTHIA 8 tune, referred to as PYTHIA A3 [183]. The A2 tune of PYTHIA gives reasonable results in the prediction of the charged-particle multiplicity at 13 TeV, but an overestimation of the fiducial inelastic cross-section [184] was found compared to 13 TeV data. Since both measurements are important in order to define an optimal setup for the pile-up simulation, the PYTHIA A3 tune was developed with the aim of improving the prediction of the visible inelastic cross-section and, in the mean time, keeping a good prediction of the charged-particle multiplicity. All the details can be found in [183], where it can be seen that the results are promising, because as wanted, the prediction of the visible inelastic cross-section has improved and the charged-particle multiplicity is still reasonable described, even though discrepancies in the very low and very high n_{ch} region remain.

Conclusions

After a successful data-taking period during 2010-2012, which led the discovery of the Higgs Boson, the LHC accelerator complex and the experiments were consolidated and improved in the period 2013-2014, during the first long shutdown of the collider.

To meet the more challenging experimental conditions of the LHC, a fourth pixel layer, the Insertable B-Layer, was installed in the ATLAS Inner Detector together with a new thinner beam pipe, at a radius of 33 mm. The pixel services, located in the end-cap region between the Pixel and the SCT detector were also modified.

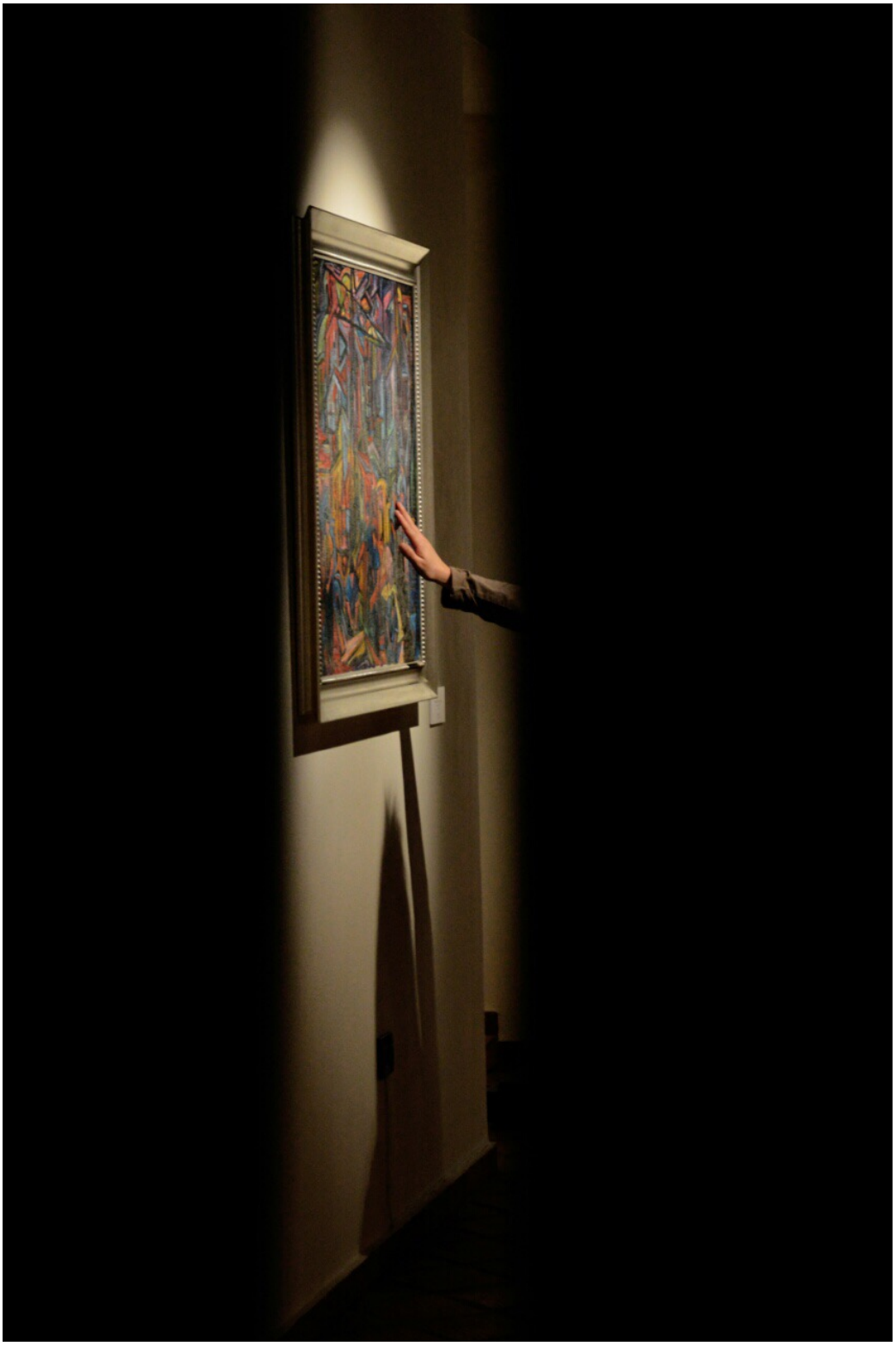
The tracking algorithms were optimised to account for these changes and a robust reconstruction of low momentum tracks was established. In addition, the performance figures of the upgraded Inner Detector were extensively studied and benchmark values were recommended. These recommendations apply to physics analyses using Inner Detector tracks in Run II data and constitutes an important input for other objects based on tracks, such as jets. A simulation-based approach was developed, which consists in evaluating the systematics on tracks and seeing how much this affects the mis-tag rate of light-jets in simulation by using the modified track collection in the b-tagging algorithms. To obtain a precise calibration of the light-jets is crucial in the measurement of the cross-section of the W-boson produced in association with b-jets.

The material in the Inner Detector was evaluated with three complementary techniques. Secondary vertices from hadronic interactions and photon conversions were used to probe the material in the barrel pixel region, while the stopping rate of charged hadrons in the end-cap pixel service region was measured with the Track-Extension efficiency method. The three methodologies led to a big improvement in the geometry models used in the ATLAS detector simulation programs which will be used as a baseline for the ATLAS Run II physics analyses. The results of hadronic interactions and photon conversions derived from simulated samples based on the *updated* geometry model agrees with data to within 7-18%. In particular, for the recently installed IBL, data and simulation agree to within 10%. The track-extension efficiency allowed to estimate that the geometry model used in 2015 Monte Carlo production lacked up to $(3.69 \pm 0.86)\%$ of nuclear interaction lengths in the pseudorapidity region $1 < |\eta| < 2.5$, which corresponds roughly to 10% of the material in the pixel services in the corresponding regions.

The results obtained with the Run II ID Material Studies played a big role in the measurement of charged-particle multiplicities at $\sqrt{s} = 13$ TeV, because the largest source of uncertainty on the tracking efficiency, main ingredient of this measurements, is the lack of knowledge about the material distribution. The analysis was performed in three different phase-spaces: *nominal* ($p_{\mathrm{T}} > 500$ MeV, $|\eta| < 2.5$), *reduced* ($p_{\mathrm{T}} > 500$ MeV, $|\eta| < 0.8$) and *extended* ($p_{\mathrm{T}} > 100$ MeV, $|\eta| < 2.5$). The analysis in the low-momentum regime was possible thanks to the improved Run II reconstruction algorithms. For what concerns the analysis in the nominal phase-space, the new aspect with respect to previous measurements is the use of the track-extension efficiency to perform a data driven correction to the simulation-based track-reconstruction efficiency. This correction allowed to reduce the systematic

uncertainty and to obtain better results to be used in the tuning of Monte Carlo event generators. In addition, differently than in earlier analyses, strange baryons were removed from the primary-charged-particle definition because of the very low reconstruction efficiency and of the differences in the predicted rates in different generator, which would lead to a model dependence of the measurement if those particles were included. The mean number of primary charged particles in the central region is computed by averaging over $|\eta| < 0.2$ and it is found to increase by a factor of 2.2 when $\sqrt{s}$ increases by a factor of about 14 from 0.9 TeV to 13 TeV. The predictions of various Monte Carlo event generators were compared to data. Overall, the best results were given by EPOS, followed by PYTHIA MONASH and PYTHIA A2, while QGSJET-II gives worse predictions. The charged-particle multiplicities results at $\sqrt{s} = 13$ TeV have been already used to release a new Pythia 8 tune, namely Pythia 8 A3.

The ATLAS Experiment is undergoing a continuous upgrade program aimed at coping with the high luminosity performance, which will be delivered by the upgraded LHC during Run 3 and beyond (2020). One of these upgrades consists in the replacement of the first station of the ATLAS muon end-cap system, Small Wheels, during the second long shutdown (2018). The New Small Wheel will have to operate in a high background radiation region while reconstructing muon tracks with high precision as well as furnishing information for the Level-1 trigger. The simulation of the New Small Wheel chamber layout was implemented into the general ATLAS Simulation framework and a set of automatic validation procedures was put in place.



Relative view point, by Federica Cairo.

–INVICTUS–
William Ernest Henley

Out of the night that covers me,
Black as the pit from pole to pole,
I thank whatever gods may be
For my unconquerable soul.

In the fell clutch of circumstance
I have not winced nor cried aloud.
Under the bludgeonings of chance
My head is bloody, but unbowed.

Beyond this place of wrath and tears
Looms but the Horror of the shade,
And yet the menace of the years
Finds and shall find me unafraid.

It matters not how strait the gate,
How charged with punishments the scroll,
I am the master of my fate,
I am the captain of my soul.

Appendix A

Application of tracking studies to W-boson Physics

The aim of this Appendix is to give examples of the importance of tracking-related studies in the context of physics measurements.

Discovered in 1983, the W boson is a fundamental particle and, together with the Z boson, it is responsible for the weak force. Their properties have been extensively studied over the last 30 years at different colliders ($Sp\bar{p}S$, Tevatron, LEP). W and Z boson production still represents one of the most sought after signatures at the LHC and one of the first analyses to be performed. Measurements of vector boson production cross-section either by themselves or in association with other objects constitute important tests of the Standard Model and its fundamental parameters.

One of the recent highlights in vector boson physics was the measurement by the ATLAS collaboration of the W-boson mass [185] with a precision to the level of better than 20 MeV.

Two examples related to W-boson physics to which the Author of this thesis personally contributed will be illustrated.

The measurement of W-boson production provides benchmark for the understanding of quantum chromodynamics and electroweak processes. One of the largest sources of background on the inclusive measurement of the W-boson production cross section is due to the pile-up modeling. The studies done to carefully reweight the pile-up distribution in Monte Carlo simulation to match those in data will be discussed.

Another interesting signature which can be measured with high precision at the LHC is the associated production of W-bosons with b-quark jets. This measurement can provide insights in the next-to-leading order QCD predictions and also on the b-quark content in PDFs. In the following, the calibration of light-quark jets in the context of the W-boson production in association with b-jets will be presented. This study is of crucial importance to determine the systematic uncertainty arising from the rate of light-jets mis-tagged as b-jets.

A.1 Measurement of W boson production cross section in pp collisions at $\sqrt{s} = 13\text{TeV}$ with the ATLAS detector

A.1.1 Introduction

Measurements of electroweak vector-boson production at hadron colliders provide a benchmark for the understanding of quantum chromodynamics and electroweak processes. Predictions are available up to next-to-next-to-leading-order (NNLO) accuracy in QCD and include EW corrections at next-to-leading-order (NLO) accuracy [186]. The cross-section predictions depend on the parton distribution functions (PDFs) and are thus sensitive to the underlying dynamics of strongly interacting particles. Therefore, measurements of

$W^{\pm}$ and Z-boson production offer a unique opportunity to test models of parton dynamics at the Large Hadron Collider's new higher centre-of-mass energy of $\sqrt{s} = 13$ TeV.

The studies summarised here and described in detail in Ref. [187] present measurements of the inclusive production cross sections times leptonic branching ratios for the $W^{\pm} \to e^{\pm}\nu$, $W^{\pm} \to \mu^{\pm}\nu$, $Z \to e^{+}e^{-}$, and $Z \to \mu^{+}\mu^{-}$ processes. Measurements of the cross-section ratios of W^{+} to W^{-} production and of $W^{\pm}$ to Z production are also presented. All measurements are performed with 13 TeV pp collision data corresponding to an integrated luminosity of $81\ \mathrm{pb}^{-1}$, The data were collected during the period of June 13 to July 16, 2015, at which point the LHC circulated $6.5\,\mathrm{TeV}$ beams with $50\,\mathrm{ns}$ bunch spacing. The peak delivered instantaneous luminosity was $L = 1.7 \times 10^{33}\,\mathrm{cm}^{-1}\,\mathrm{s}^{-1}$ and the mean number of pp interactions per bunch crossing (hard scattering and pile-up events) was $\langle\mu\rangle = 19$.

A detailed description of data and simulated samples used in the analysis, as well as a comprehensive presentation of the methodologies can be found in Ref. [187] and they will not be reported in this thesis. What will be mainly discussed is, instead, the impact of the pile-up modeling on the final measurement.

A.1.2 Production cross-section

The total production cross section for the $W^{\pm}$ boson times the branching ratio for decays into a single-lepton flavour $\ell^{\pm} = e^{\pm}, \mu^{\pm}$ ($\sigma^{\mathrm{tot}}_{W^{\pm}}$,$\sigma^{\mathrm{tot}}_{W^{+}}$, and $\sigma^{\mathrm{tot}}_{W^{-}}$) can be expressed as a ratio of the numbers of background-subtracted data events N to the product of the integrated luminosity of the data $\mathcal{L}$, an acceptance factor A, and a correction factor C:

$$\sigma^{\mathrm{tot}} = \frac{N}{\mathcal{L} \cdot A \cdot C}. \tag{A.1}$$

The acceptance factor A is expressed as the fraction of decays satisfying the fiducial acceptance (geometric and kinematic requirements) at the Monte Carlo generator level. The correction factor C is the ratio of the total number of generated events which pass the final selection requirements after reconstruction to the total number of generated events within the fiducial acceptance. This factor, defined before the decay leptons emit photons via final-state radiation (Born-level FSR), includes the efficiencies for triggering on, reconstructing, and identifying the $W^{\pm}$ decay products within the acceptance, and also accounts for the slight difference between the fiducial and reconstructed phase spaces. The production cross sections defined without the acceptance factors ($\sigma^{\mathrm{tot}} \cdot A$) are referred to as the fiducial cross sections ($\sigma^{\mathrm{fid}}_{W^{\pm}}$, $\sigma^{\mathrm{fid}}_{W^{+}}$, $\sigma^{\mathrm{fid}}_{W^{-}}$, and $\sigma^{\mathrm{fid}}_{Z}$). For the $W^{\pm}$-boson measurement, the fiducial phase space is defined by the lepton transverse momentum $p^{\ell}_{\mathrm{T}} > 25$ GeV, the lepton pseudorapidity $|\eta_{\ell}| < 2.5$, the neutrino transverse momentum $p^{\nu}_{\mathrm{T}} > 25$ GeV, and the $W^{\pm}$-boson transverse mass [1] $m_{\mathrm{T}} > 50$ GeV.

A.1.3 Evaluation of systematic uncertainties

The experimental systematic uncertainties in the measurements of the cross sections enter via the evaluation of the correction factor and the luminosity in the denominator of Eq. (A.1), and through the estimation of the background subtracted from the candidate events in its numerator.

The sources of systematic uncertainties in the correction factors C, summarised in Table A.1, are as follows. *Trigger:* The lepton trigger efficiency is estimated in simulation, with a dedicated data-driven analysis performed to obtain the simulation-to-data trigger

[1] $m_{\mathrm{T}} = \sqrt{2p^{\ell}_{\mathrm{T}}\, p^{\nu}_{\mathrm{T}}\, [1 - \cos(\phi_{\ell} - \phi_{\nu})]}$ with azimuthal angle of the charged lepton ϕ_{ℓ} and azimuthal angle of the neutrino ϕ_{ν}.

$\delta C/C$ [%]	$Z \to e^+e^-$	$W^+ \to e^+\nu$	$W^- \to e^-\overline{\nu}$	$Z \to \mu^+\mu^-$	$W^+ \to \mu^+\nu$	$W^- \to \mu^-\overline{\nu}$
Lepton trigger	0.1	0.3	0.3	0.2	0.6	0.6
Lepton reconstruction, identification	0.9	0.5	0.6	0.9	0.4	0.4
Lepton isolation	0.3	0.1	0.1	0.5	0.3	0.3
Lepton scale and resolution	0.2	0.4	0.4	0.1	0.1	0.1
Charge identification	0.1	0.1	0.1	–	–	–
JES and JER	–	1.7	1.7	–	1.6	1.7
E_T^{miss}	–	0.1	0.1	–	0.1	0.1
Pile-up modelling	< 0.1	0.4	0.3	< 0.1	0.2	0.2
PDF	0.1	0.1	0.1	< 0.1	0.1	0.1
Total	1.0	1.9	1.9	1.1	1.8	1.8

TABLE A.1: Relative systematic uncertainties (%) in the correction factors C in the different channels. Ref. [187].

correction factors and the corresponding uncertainties. *Reconstruction, Identification, and Isolation:* The lepton selection efficiencies as determined from simulation are corrected with simulation-to-data correction factors and their associated uncertainties [188, 189]. *Energy, Momentum Scale/Resolution:* Uncertainties in the lepton calibrations are applied as they can cause a change of acceptance because of migration of events across the p_{T} threshold and $m_{\ell\ell}$ boundaries. *Charge Identification:* Electron charge misidentication may occur when electrons radiate early in the detector and the resulting photons subsequently convert and are reconstructed as high p_{T} tracks. A particle with reconstructed charge opposite to the parent electron may then accidentally be associated with the calorimeter cluster. The effect of electrons having their charge reconstructed wrongly is studied using a control sample of $Z \to ee$ events in which both electrons are reconstructed with the same charge and is found to be well described by the Monte Carlo simulation, within the statistical uncertainty of the control sample. An uncertainty is assessed to cover any small residual differences between data and simulation. The probability of charge misidentification is negligible in the muon channel. *Jet-Energy Scale/Resolution (JES and JER):* The corresponding uncertainties, described in Ref. [190], are propagated to the calculation of the missing transverse momentum. E_T^{miss}*:* Uncertainties in the soft component of the E_T^{miss} scale and resolution evaluated as described in Ref. [191] are included. *Pile-up:* Incorrect modelling of pile-up effects can lead to acceptance changes and is accounted for with dedicated studies. *PDF:* The impact of PDF eigenvector variations is propagated to the correction factor.

In the Z-boson channel, the systematic uncertainties from the background evaluation contribute negligibly to the experimental cross-section uncertainty. This is not the case for the $W^{\pm}$-boson channel; the multijet background evaluation results in uncertainties of up to 3.4% on the cross-section measurements in the electron channel and up to 1.4% in the muon channel.

The cross sections have a 2.1% uncertainty in the measurement of the integrated luminosity, which is derived, following a methodology similar to that detailed in Ref. [192], from a preliminary calibration of the luminosity scale using a pair of x–y beam separation scans performed in August 2015. Apart from the determination of the luminosity, the dominant experimental systematic uncertainties in the cross-section evaluations are the jet-energy scale/resolution and the multijet background for the $W^{\pm}$-boson measurements while they are lepton reconstruction and identification efficiencies for the Z-boson measurements.

It is clearly visible that the systematic uncertainty aring from the pile-up modeling is one of the largest systematics in the W-channel. The Author of this thesis studied the effect of different pile-up reweighting methods.

A.1.4 Pile-Up Reweighting Studies

As the Monte Carlo samples are generated with a generic spectrum of average interactions per crossing $\langle\mu\rangle$, the simulation should be corrected to take into account the actual values in data. This improves the data/simulation agreement in several kinematic variables sensitive to the amount of secondary interactions deriving from the collisions.

Different pile-up reweighting methods have been investigated, in the context of $W \to \mu\nu$ analysis using period A4 data, in order to choose the best procedure to be applied to the full analysis.

The distributions for μ, vertices, transverse mass and missing energy without any reweighting applied are shown in Figure A.1.

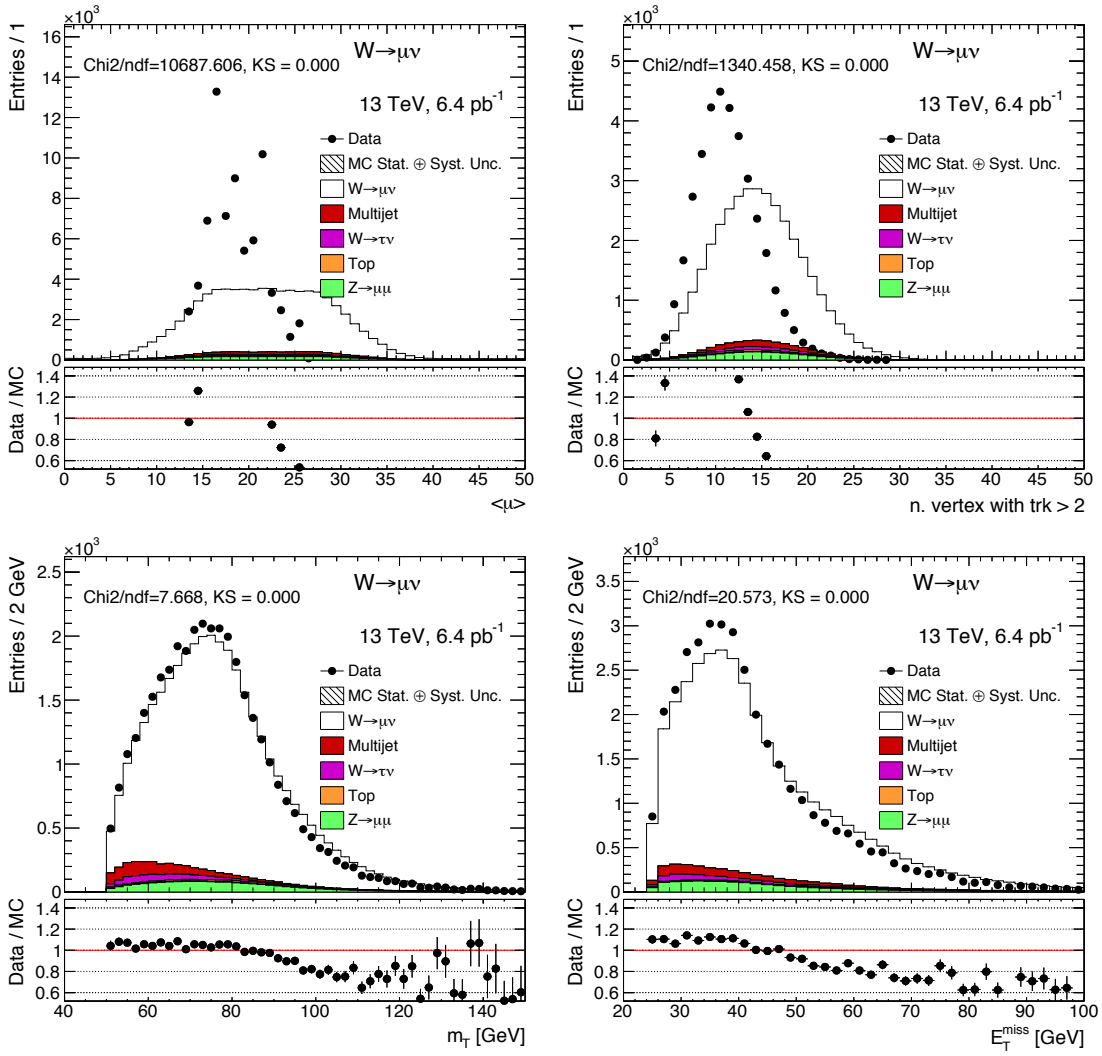


FIGURE A.1: Mu (top-left), vertices (top-right), transverse mass (bottom-left) and missing energy (bottom-right) distributions, calculated from the $W \to \mu\nu$ selection without applying any reweighting.

Four methods have been then investigated:

- $\langle\mu\rangle$-reweighting method: the average interactions per bunch crossing distribution in Monte Carlo simulation has been reweighted to match the corresponding distribution in data
- vertex-reweighting method: the vertices distribution in Monte Carlo simulation has been reweighted to match the corresponding distribution in data

- 2D $\langle\mu\rangle$-vertex reweighting method: the 2D distribution of $\langle\mu\rangle$ vs number of vertices has been reweighted in Monte Carlo simulation to match the corresponding distribution in data

- $\langle\mu\rangle$-reweighting method with an additional $1/1.16$ correction factor: the average interactions per bunch crossing distribution in Monte Carlo simulation has been reweighted to match the corresponding distribution in data after that the $\langle\mu\rangle$ values have been scaled by a factor of $1/1.16$. This correction factor has been evaluated in the context of the ATLAS Tracking CP group [193] in order to take into account the fraction of inelastic activity differently described in the MC simulation with respect to the data and to match the number of vertices vs $\langle\mu\rangle$ in data and MC simulation.

The effect of the different methods on the main affected kinematic variables (transverse mass and missing transverse energy) can be seen in Figures A.2 and A.3. The agreement between data and simulation, quantitatively described by the reduced χ^2 and the Kolmogorov-Smirnov (KS) tests reported on the distributions, clearly indicate that the best procedure to be be applied is the $\langle\mu\rangle$-reweighting method with an additional $1/1.16$ correction factor.

The meaning and the effect of the additional $1/1.16$ correction factor can be easily understood by looking at Figures A.4: by scaling the top-left mu distribution by a factor of $1/1.16$, the bottom-right vertices distribution reaches an improved agreement.

A.1.5 Estimation of the systematic uncertainty due to pileup reweighting variation

As it is discussed in Section A.1.4, the analysis uses the standard pileup reweighting tool and performs additional scaling of the $\langle\mu\rangle$ distribution in data, by the factor $s_p = 1/1.16$, to achieve better agreement between data and Monte Carlo simulation. For the conference note analysis, the uncertainty of the C_Z factor due to this additional scaling was evaluated by varying the scaling factor between 1 and the nominal value of $1/1.16$, and when symmetrized. This appendix describes an improved methodology of the uncertainty estimation.

The improvements come from the two sources:

- Updated studies from the tracking group show that the upper variation should be limited to 1.23 [193]. Additional discussions in the context of the measurement of the inelastic cross-section [184] suggest that the lower variation should be reduced to 1.09 instead of 1.00 This change in the prescription leads to significantly reduced pileup reweighting uncertainty which has been re-evaluated for the W and Z events. For example, for the $W \to \mu\nu$ events the uncertainty on the C_W factor is reduced from 1.2% to 0.4%.

- The variation of the pileup scaling factor yields not only in change of the E_T^{miss} distribution, which is a genuine systematic uncertainty, but also in changes of the isolation and identification efficiencies. The latter are determined in a data-driven way and have dedicated uncertainties. The change in $C_{W,Z}$ due to change in the efficiency scale factors should be thus subtracted, to avoid double counting of the uncertainties.

  A dedicated tag-and-probe study was performed using $Z \to e^+e^-$ and $Z \to \mu^+\mu^-$ samples to evaluate the impact of the pileup reweighting variation of the isolation scale factors. The efficiencies measured in data were compared to those estimated in MC, using official correction factors. MC samples are considered with the standard $s_p = 1/1.16$ and modified pileup scaling factors ($s_p = 1.0, 1/1.09, 1/1.23$). It

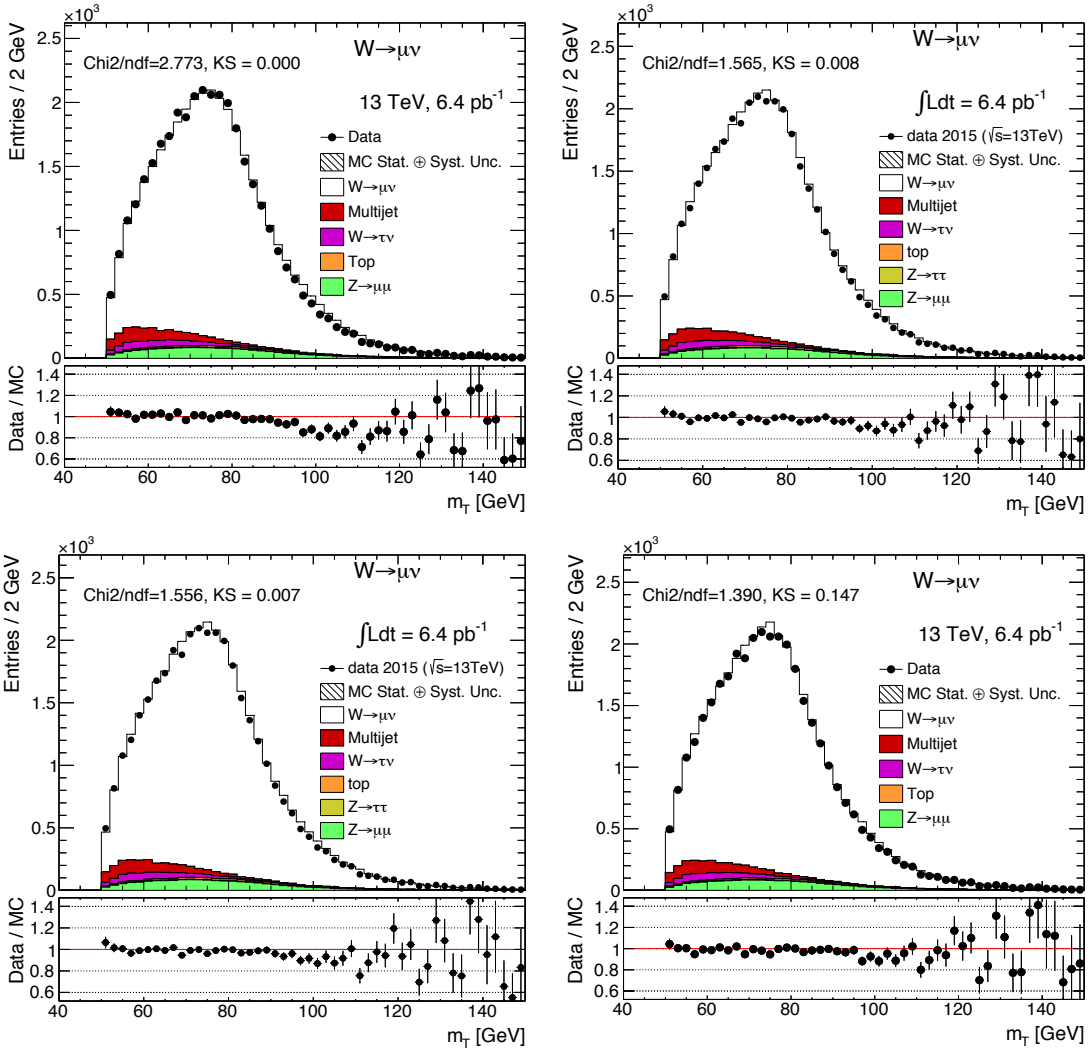


FIGURE A.2: Transverse mass distribution, calculated from the $W \to \mu\nu$ selection, using (top-left) the $\langle\mu\rangle$-reweighting method, (top-right) the vertex-reweighting method, (bottom-left) the 2D $\langle\mu\rangle$-vertex rewighting method, (bottom-right) the $\langle\mu\rangle$-reweighting method with a further $1/1.16$ correction factor.

is observed that changes in the s_p factor leads to significant change in the isolation efficiency while other efficiencies are not affected significantly.

Figure A.6 shows isolation efficiency ϵ_{ISO} as a function of lepton pseudorapidity measured in data and MC with the pileup scaling factors of $s_p = 1/1.16$ and $s_p = 1$ for $Z \to \ell\ell$ events. One can see consistent change of the isolation efficiency of about 0.4%, similar for electrons and muons. The overall shift is determined using a constant function to fit the ratio of the two efficiencies.

The corrected systematic uncertainties in C_Z and C_W factors due to pile up scaling factor variation are determined as

$$\begin{aligned} \Delta C_Z^{s_p,corr} &= \Delta C_Z^{s_p} - 2 \times (\Delta\epsilon_{ID} + \Delta\epsilon_{ISO}) \,, \\ \Delta C_W^{s_p,corr} &= \Delta C_W^{s_p} - \Delta\epsilon_{ID} - \Delta\epsilon_{ISO} \,, \end{aligned} \tag{A.2}$$

where $\Delta C_{Z,W}^{s_p}$ stands for the raw change of $C_{W,Z}$ due to variation in s_p and $\Delta\epsilon_{ID}, \Delta\epsilon_{ISO}$ stand for the change in the ID and isolation efficiencies. Table A.2 summarises the changes in the isolation efficiency, as well as raw and corrected changes of $C_{Z,W}$ for the electron and

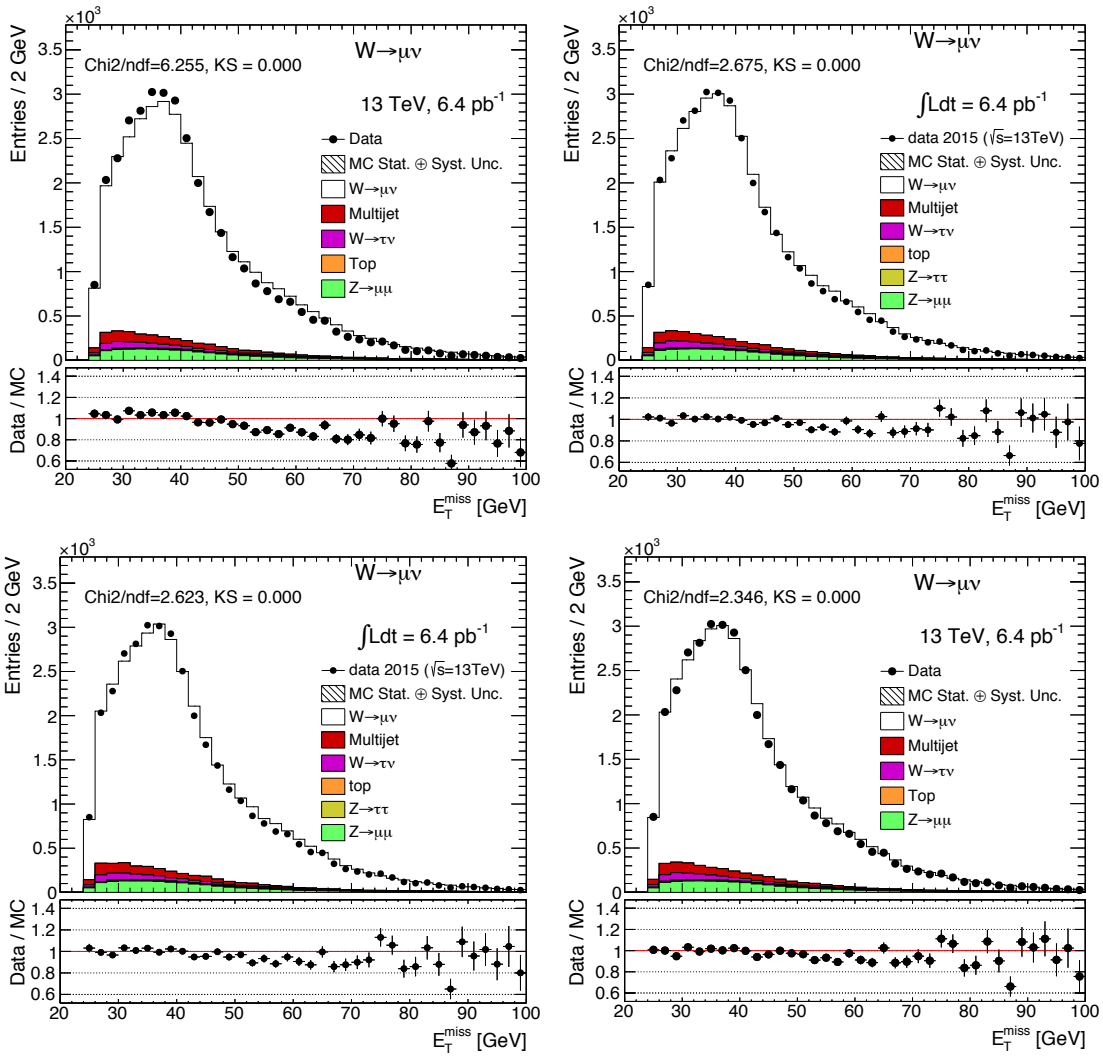


FIGURE A.3: Missing transverse energy distribution from the $W \to \mu\nu$ selection,using (top-left) the $\langle\mu\rangle$-reweighting method, (top-right) the vertex-reweighting method, (bottom-left) the 2D $\langle\mu\rangle$-vertex rewighting method, (bottom-right) the $\langle\mu\rangle$-reweighting method with a further 1/1.16 correction factor.

muon channel for different values of s_p. Based on these variations, the resulting systematic uncertainty is summarised in table A.3.

A.1.6 Results

The m_{T} distributions after the final selection are shown in Figure A.7 for the $W \to e\nu$ and $W \to \mu\nu$ channels, respectively, for the data compared to the predictions, normalised to the measured cross section.

The ratios of the fiducial electron and muon channel measurements in the $W^{\pm}$ ($R_{W^{\pm}} = \sigma^{\mathrm{fid}}_{W^{\pm}\to e\nu}/\sigma^{\mathrm{fid}}_{W^{\pm}\to\mu\nu}$) and the Z-boson ($R_Z = \sigma^{\mathrm{fid}}_{Z\to e^+e^-}/\sigma^{\mathrm{fid}}_{Z\to\mu^+\mu^-}$) channels, evaluated taking into account correlations in the systematic uncertainties, are shown in Figure A.8. Since these results agree well with Standard Model expectations of lepton universality, a simultaneous combination of the W^+, W^- and Z-boson fiducial cross sections using the `HERAverager` program [196, 197] is performed.

s_p variation	1.00 %	1.09 %	1.23 %
$\Delta\epsilon_{ID}$ (electrons)	-0.025	-0.013	0.001
$\Delta\epsilon_{ISO}$ (electrons)	-0.423	-0.182	0.139
$\Delta\epsilon_{ISO}$ (muons)	-0.436	-0.172	0.170
$\Delta C_Z^{s_p}$ ($Z \to ee$)	-0.908	-0.381	0.274
$\Delta C_Z^{s_p}$ ($Z \to \mu\mu$)	-0.919	-0.354	0.307
$\Delta C_W^{s_p}$ ($W^+ \to e^+\nu$)	-1.41	-0.50	0.47
$\Delta C_W^{s_p}$ ($W^- \to e^-\nu$)	-1.40	-0.56	0.48
$\Delta C_W^{s_p}$ ($W^+ \to \mu^+\nu$)	-1.15	-0.40	0.40
$\Delta C_W^{s_p}$ ($W^- \to \mu^-\nu$)	-1.20	-0.41	0.37
$\Delta C_Z^{s_p,corr}$ ($Z \to ee$)	-0.012	0.009	-0.006
$\Delta C_Z^{s_p,corr}$ ($Z \to \mu\mu$)	-0.047	-0.010	-0.034
$\Delta C_W^{s_p,corr}$ ($W^+ \to e^+\nu$)	-0.96	-0.31	0.33
$\Delta C_W^{s_p,corr}$ ($W^- \to e^-\nu$)	-0.95	-0.37	0.34
$\Delta C_W^{s_p,corr}$ ($W^+ \to \mu^+\nu$)	-0.71	-0.23	0.23
$\Delta C_W^{s_p,corr}$ ($W^- \to \mu^-\nu$)	-0.76	-0.24	0.20

TABLE A.2: Changes of the identification and isolation efficiencies, C_W and C_Z factors, before and after correction using equation A.2.

Channel	Uncertainty up %	Uncertainty down %
$Z \to e^+e^-$	0.009	-0.006
$Z \to \mu^+\mu^-$	-0.010	-0.034
$W^+ \to e^+\nu$	0.33	-0.31
$W^- \to e^-\nu$	0.34	-0.37
$W^+ \to \mu^+\nu$	0.23	-0.23
$W^- \to \mu^-\nu$	0.20	-0.24

TABLE A.3: Summary of the systematic uncertainties in $C_{W,Z}$ due to pileup reweighting.

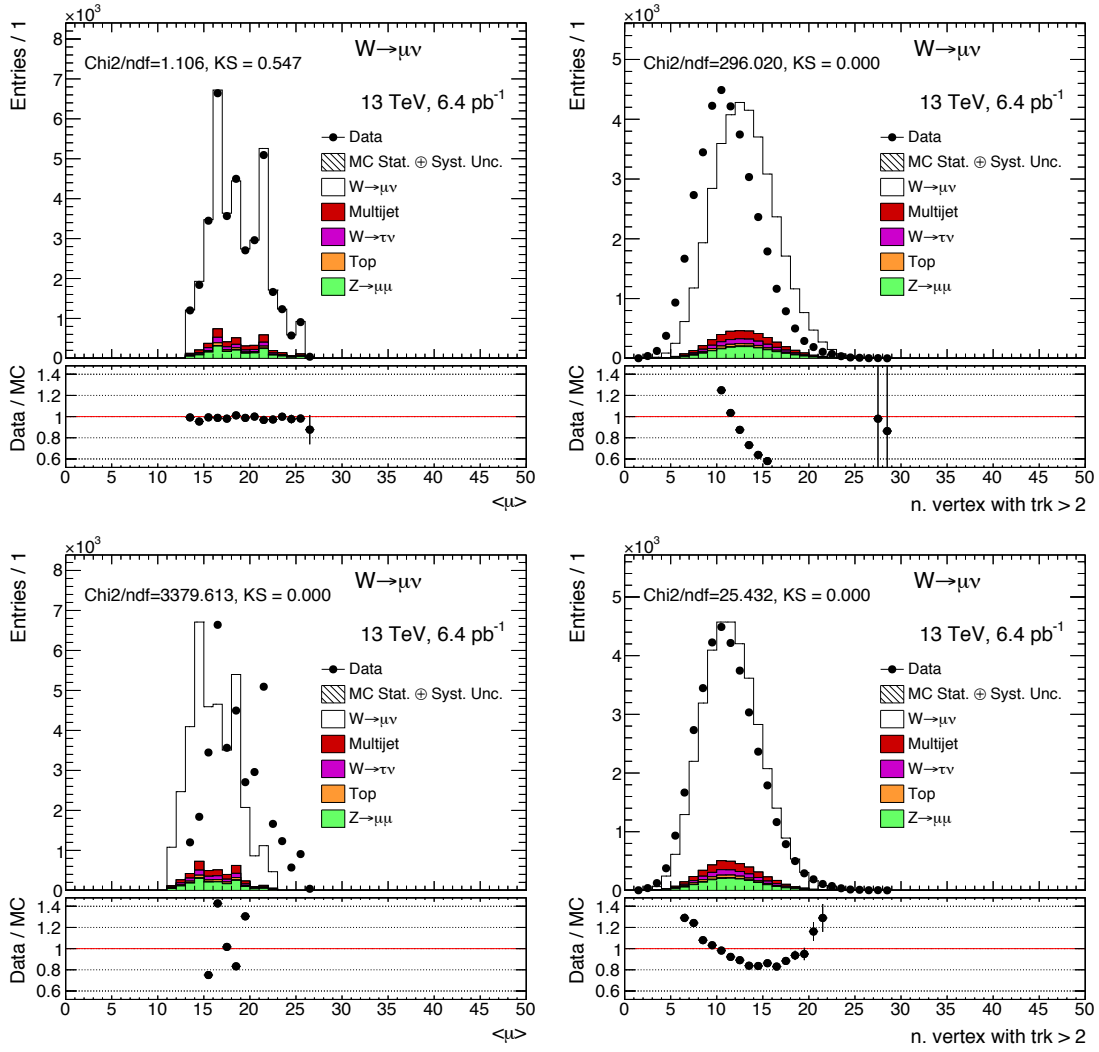


FIGURE A.4: Mu (left) and Vertices (right) distributions from the $W \to \mu\nu$ selection, using (top) the $\langle\mu\rangle$-reweighting method, (bottom) the $\langle\mu\rangle$-reweighting method with a further $1/1.16$ correction factor.

The combination uses the individual sources of the systematic uncertainties in addition to uncertainties in the background evaluations. Sources corresponding to lepton reconstruction and identification are uncorrelated between the electron and muon channels. Some sources, such as JES, JER, $E_T^{\text{miss}}$and multijet background, only affect $W^{\pm}$-boson measurements. The correlation model used for combining the multijet W^+ and W^- uncertainties in each lepton channel is defined by:

$$\delta(W^{\pm})^2 = \delta(W^+)^2 + \delta(W^-)^2 + 2\rho\delta(W^+)\delta(W^-), \tag{A.3}$$

performed separately for each source of systematic uncertainty considered for this background. All such uncertainties are considered to be uncorrelated between the electron and muon channels except that of the jet-energy-scale variation. The correlation coefficient, ρ, is obtained from the uncertainties evaluated separately for the W^+ and W^- channels ($\delta(W^+),\ \delta(W^-)$), and repeating the multijet background extraction without selecting the final-state charge ($\delta(W^{\pm})$). The correlations of the systematic uncertainties vary from 0.2 to 1 (fully correlated), depending on the lepton channel and type of uncertainty. The common normalisation uncertainty due to the luminosity calibration is excluded from the combination procedure and applied separately to the result.

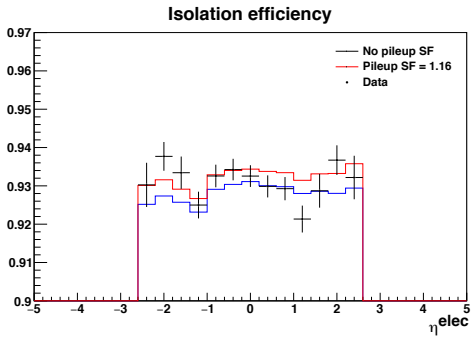


FIGURE A.5: Isolation requirement efficiency for electrons as a function of the rapidity as determined using tag-and-probe method. The data are shown as dots with error bars. The solid red (blue) histogram shows MC efficiency using $s_p = 1/1.16$ ($s_p = 1$).

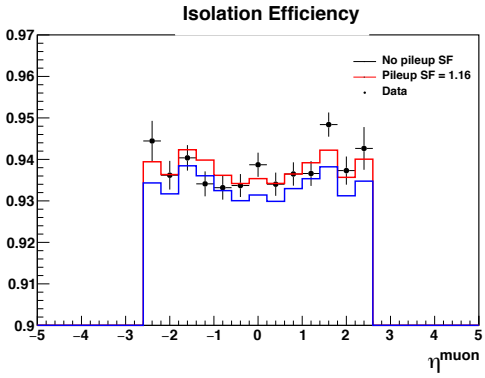


FIGURE A.6: Isolation requirement efficiency for muons as a function of the rapidity as determined using tag-and-probe method. The data are shown as dots with error bars. The solid red (blue) histogram shows MC efficiency using $s_p = 1/1.16$ ($s_p = 1$).

The combination yields a $\chi^2/N_{\text{d.o.f.}} = 3.0/3$, indicating excellent compatibility of the measurements. There is a reduction of uncertainty compared to individual electron and muon channel measurements since many of the systematic error sources are uncorrelated.

The combined fiducial cross sections are compared in Figure A.9 to the predictions, which are calculated using different PDF sets. The measurements agree well with the predictions and the experimental precision is comparable to the PDF uncertainties; however, the total precision is diluted by the uncertainty of the preliminary luminosity calibration.

Ratios of the measured cross sections benefit from the cancellation of some experimental uncertainties. The ratios of W^+ to W^- and $W^\pm$ to Z-boson production, measured by the ATLAS, CMS, and LHCb collaborations in the past [198–202], proved to be powerful tools to constrain PDF fits. The ratio of W^+ to W^--boson cross sections is mostly sensitive to the difference of u_v and d_v valence-quark distributions at low Bjorken-x, while the ratio of $W^\pm$ to Z constrains the strange-quark distribution [203].

The systematic uncertainties of the ratio measurements are largely uncorrelated between the electron and muon channels, apart from the common luminosity uncertainty. However, there is a strong correlation between W^+ and W^--boson measurements and between the $W^\pm$ and Z-boson results for the same-flavour measurement. The results for the W^+/W^- and $W^\pm/Z$ ratios of fiducial production cross sections in the combined electron and muon channels are compared to predictions in Figure A.10. The dominant components of the systematic uncertainty in the $W^\pm/Z$ ratio are from both the multijet

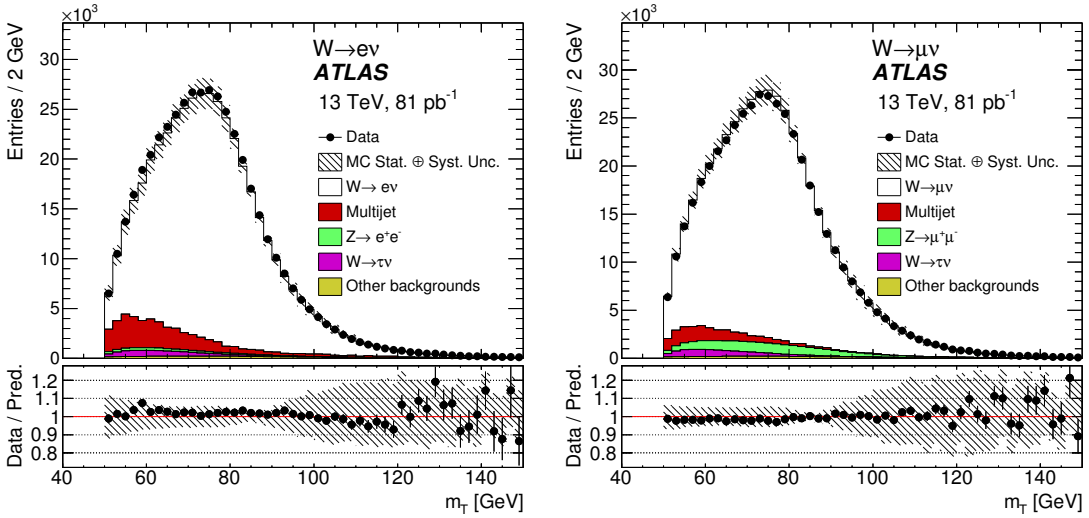


FIGURE A.7: Transverse mass distributions from the $W \to e\nu$ and $W \to \mu\nu$ selections (top) The predicted signal distributions are normalised to the measured cross sections as presented in this paper. The shaded bands in the histograms encompass the uncertainties described in Table A.1. In addition to these uncertainties in the correction factors, the uncertainties in the evaluation of the multi-jet background in the $W^{\pm}$-boson analysis are included in the shaded bands. Ref. [187].

background and the jet-energy scale/resolution while that of the W^+/W^- ratio is from the uncorrelated part of the multijet background uncertainty. For the ratios $R_{W^+/W^-} = \sigma^{\text{fid}}_{W^+}/\sigma^{\text{fid}}_{W^-}$ and $R_{W/Z} = \sigma^{\text{fid}}_{W^\pm}/\sigma^{\text{fid}}_{Z}$, several predictions agree within quoted uncertainties, although all predictions are above the central value for the data in both cases.

A.2 Measurement of the cross-section of W-boson produced in association with b-jets in $\sqrt{s} = 13\text{TeV}$ pp collisions with the ATLAS detector

In this Appendix, preliminary studies for the measurement of the cross-section of W-bosons production in association with b-jets are presented. A short introduction about the importance of this measurement will be given as well as a summary of the data set and event generators to be used, but no details will be given on the analysis due to its very preliminary stage.

Particular emphasis will be given to some background studies based on Monte Carlo simulations, such as the light-flavour jet calibration, in which the tracking recommendations, presented in Chapter 3 have been used to perform a simulation-based calibration. Some studies about the top-background rejection are also briefly presented.

A.2.1 Introduction

The measurement of W-bosons produced in association with b-jets provides an important test of perturbative quantum chromodynamics at next-to-leading order, as it is sensitive to heavy-flavour quarks in the initial state. Several processes contribute to W+b-jets production at NLO. In the four-flavour number scheme (4FNS), where only u, d, c, s quarks are considered as initial-state quarks, the processes which contributes to the final cross section are $q\bar{q} \to Wb\bar{b}(g)$ and $gq \to Wb\bar{b}q$.

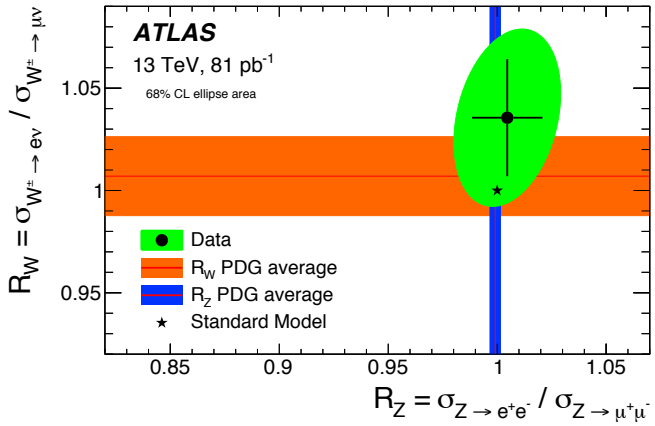


FIGURE A.8: Ratio of the electron- and muon-channel $W^{\pm}$ and Z-boson production fiducial cross sections, compared to the expected values of the Standard Model and previous experimental verifications of lepton universality for on-shell $W^{\pm}$ and Z bosons, shown as PDG average bands [194, 195]. The PDG average values and the result are shown with total uncertainties. Ref. [187].

When considering the presence of b-quarks in the initial state (5FNS), the single b-quark processes $bq \rightarrow Wbq(g)$ and $bg \rightarrow Wbq\bar{q}$, also play a significant role at the LHC energies. The description of the b-quark in the 5FNS is also approximated because the b-quark is considered as massless. The 4 and 5 Flavour Number scheme diagrams can be seen in Figure A.11.

In addition, double-parton interactions (DPI), in which W-boson and b-jets are produced from different parton-parton interactions within the same pp collision, are also expected to contribute to the total observed W+b-jets cross section.

Furthermore, events with a W-boson produced in association with b-jets signal constitute a dominant background to many other processes with smaller cross sections, from top production, to searches for the Standard Model Higgs Boson, such as the Higgs boson associated-production process WH with $H \rightarrow b\bar{b}$ decays and many Beyond the Standard Model processes including SUSY and other exotica.

A.2.2 Data and Simulated samples

The data used in this measurement were recorded in 2015 and 2016, when LHC circulated $6.5\,\mathrm{TeV}$ proton beams with $25\,\mathrm{ns}$ bunch-spacing. In 2015 data taking, the peak delivered instantaneous luminosity was L = 5x10$^{33}\mathrm{cm}^{-1}\,\mathrm{s}^{-1}$ and the mean number of pp interactions per bunch crossing corresponds to $\langle\mu\rangle = 13$; the peak delivered luminosity for 2016 pp collisions was L = 1.37x10$^{34}\mathrm{cm}^{-1}\,\mathrm{s}^{-1}$ with $\langle\mu\rangle = 25$.

Only events recorded during stable beam conditions and with all ATLAS sub-systems fully operational are considered. The analysed data passing these basic data-quality requirements correspond to a total luminosity of $36.5\,/\mathrm{fb}^{-1}$ for the full 2015+2016 dataset.

Normalised to the highest-order theoretical calculations available, Monte Carlo simulations are used to emulate signal events, to estimate the contribution of the background events, to extract the differential Wb and Zb yields, to unfold the data to the particle level and to compare with the unfolded data contributions.

The samples for the $V+b(\bar{b})$ signal and the V+light and c-jets are generated with Sherpa 2.2.1 MC [204], which is doing both the matrix-element generation and the parton showering Matrix elements are calculated up to two additional partons at NLO and up to four

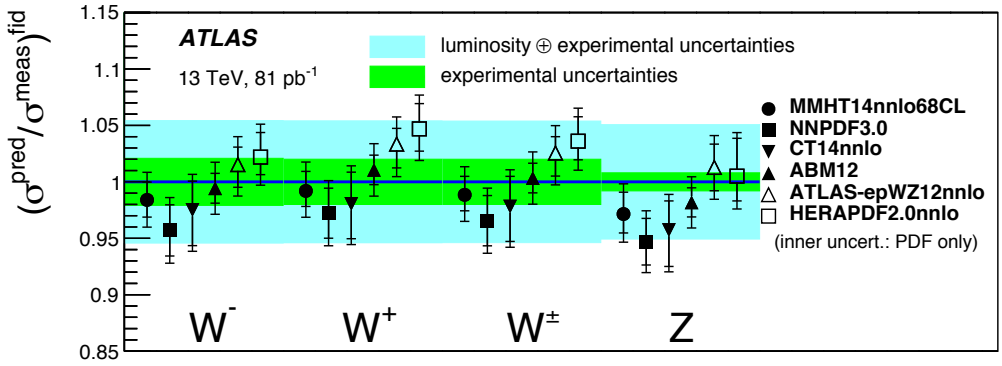


FIGURE A.9: Ratio of the predicted to measured fiducial cross section for the combined electron and muon channels using various PDFs. The inner (outer) band corresponds to the experimental uncertainty without (with) the luminosity uncertainty. The inner error bar of the predictions represents the PDF uncertainty while the outer error bar includes the sum in quadrature of all other systematic uncertainties. Ref. [187].

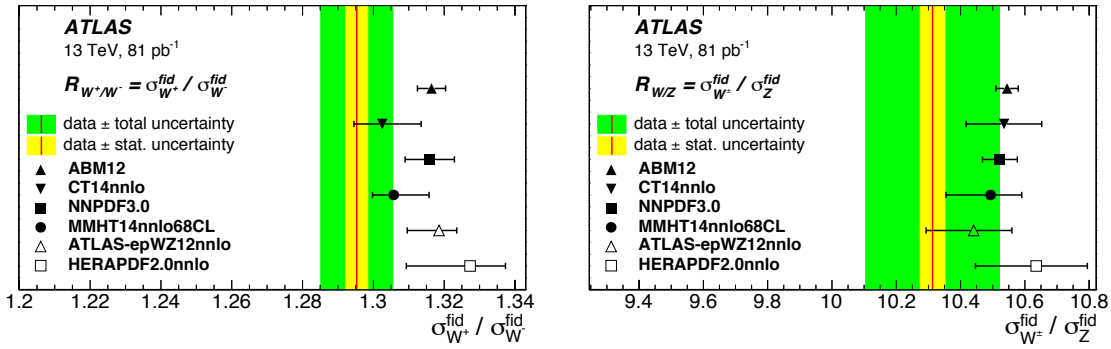


FIGURE A.10: Ratios (red line) of W^+ to W^- boson (left) and $W^\pm$ to Z boson (right) combined production cross sections in the fiducial region compared to predictions based on different PDF sets. The inner (yellow) shaded band corresponds to the statistical uncertainty while the outer (green) band shows statistical and systematic uncertainties added in quadrature. The theory predictions are given with only the corresponding PDF uncertainties shown as error bars. Ref. [187].

partons at LO; all additional jets are produced by the parton shower. Sherpa 2.2.1 uses the 5-flavour scheme with massless b- and c- quarks in the matrix element and massive quarks in the parton shower. The PDF set used is NNPDF 3.0 NLO with α_s = 0.018.

Alternative samples for the V+jets simulation are produced using Alpgen [205], a leading order (LO) matrix-element generator, interfaced with Pythia 6 to model parton shower (Perugia2012 tune [180]). Matrix elements are calculated up to five additional partons; the $V + b(\bar{b})$ and c-jets ME are explicitly included and the heavy flavour overlap procedure (HFOR) is applied to face for the double counting of heavy quarks from gluon splitting in the parton shower.

The measured cross sections will also be compared to predictions from MadGraph MG5_aMC@NLO v2.2.2 generator [206], interfaced to Pythia v8.186 for the modelling of parton shower and underlying events with A14 tune [181]. The LO matrix-element generates up to 4 partons and all additional jets are produced by the parton shower. The NNPDF v2.3 LO PDF set is used with α_s = 0.13. MadGraph5_aMC@NLO provides the 5-flavour

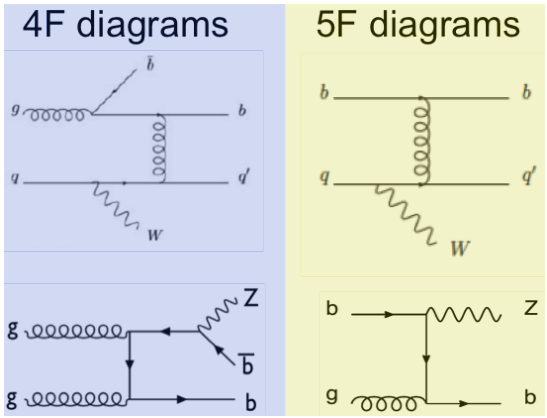


FIGURE A.11: The 4 and 5 Flavour Number scheme diagrams.

calculation with massless b- and c-quarks in the matrix element, and massive quarks in the Pythia 8 shower.

The distributions of top quark production (both $t\bar{t}$ and single top) were generated with the POWHEG -Box v2 generator [207] and PYTHIA v.6.4 [208] (Perugia 2012 tune), and the pair production is normalised to the cross section calculated at NNLO+NNLL with the Top++2.0 program [209]. For the $t\bar{t}$ sample, the POWHEG model parameter hdamp, which controls matrix element to parton shower matching in POWHEG and effectively regulates the high-p_{T} radiation, was set to the top quark mass, 172.5 GeV, a setting which was found to give the best modelling of the $t\bar{t}$ system at 7 TeV [41].

Diboson process with one of the bosons decaying hadronically and the other leptonically are simulated using using the Sherpa v2.1.1 generator with the CT10 PDF set. The matrix element generates ZZ, WW and WZ processes with up to zero or one parton at NLO and up to three additional partons at LO. The matrix element is merged with Sherpa parton shower following the ME+PS@NLO prescription.

Multiple overlaid proton-proton collisions are simulated with the soft QCD processes of PYTHIA using tune A2 [41] and the MSTW2008LO PDF. The pile-up distributions of the Monte Carlo samples have been reweighted so that the μ distribution matches the observed distribution in the data. All of the samples were processed with the Geant 4-based simulation [103] of the ATLAS detector.

A.2.3 Event Selection and Data/Simulation Comparison

The event selection, which depends on the different objects used in the analysis, such as electrons, muons, jets and missing energy, is not yet finalised in the analysis at the time of writing. For this reason it will not be reported. Some plots showing preliminary data-simulation comparison will be shown only with the aim of introducing the main source of backgrounds in this measurements and, thus, explain the contribution of the Author.

Data-Simulation comparison

In this section the detector level kinematic plots for the main variables used in the analysis are shown, for the 2015 dataset only, corresponding to $3.2\,/\mathrm{fb}^{-1}$, processed with the Athena release 20.1. Figure A.12 shows the transverse momentum distribution of electron and muon decays of the W vector boson. Figure A.13 shows the missing transverse energy distribution of electron and muon decays of the W vector boson, while in Figure fig:InclMTWReco the distribution of the transverse mass is reported. Finally, in Figure

A.15, the inclusive number of jets distribution is reported. Only the inclusive distributions are shown, where no selection on the b-jets is performed

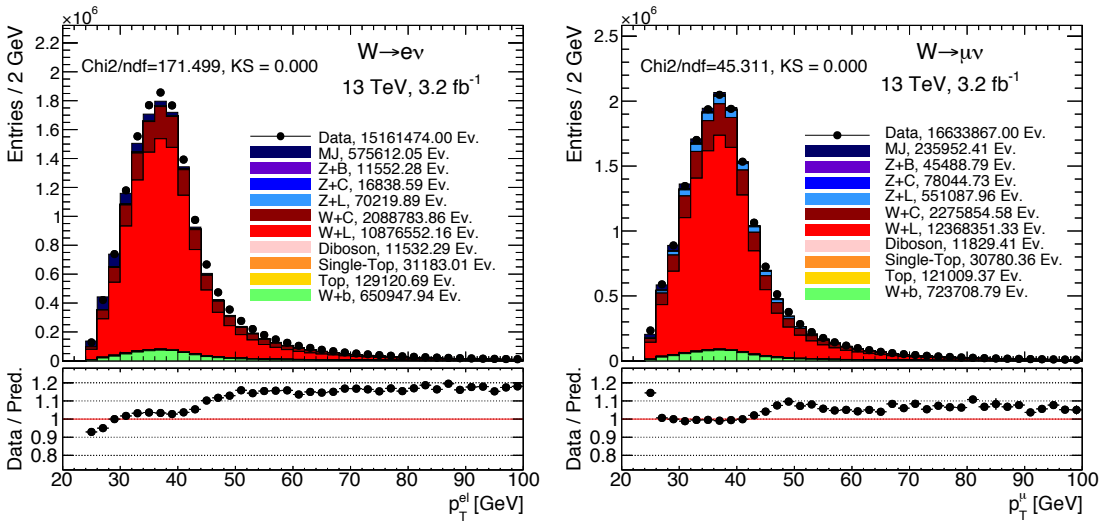


FIGURE A.12: The lepton transverse momentum before applying any selection on the number of jets or b-jets.

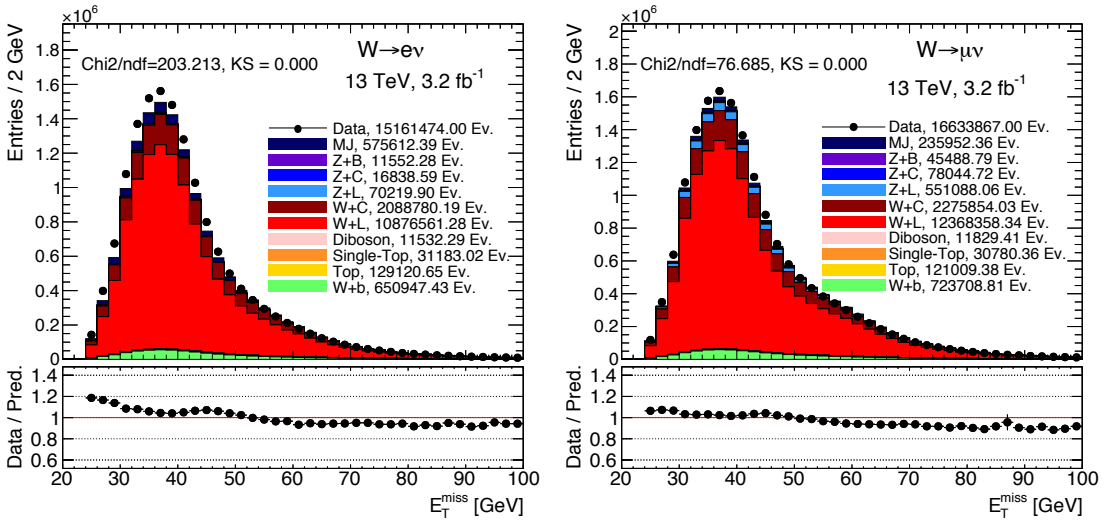


FIGURE A.13: The missing transverse energy before applying any selection on the number of jets or b-jets.

A.2.4 Backgrounds

As it can be noticed in the Figures shown in Section A.2.3, the main backgrounds to this analysis come from top production (both single top and $t\bar{t}$), QCD multijet production faking the signal, and other electroweak-produced backgrounds, such as: diboson production with semi-leptonic decays, $Z+b(\bar{b})$ for the W-channel and $W+b(\bar{b})$ for the Z-channel, V+jet production with c- and light-flavour jets, and $V + b(\bar{b})$ where the V decays to τ-leptons.

In the following section, emphasis will be given on the light-jet calibration to which the Author of this thesis contributed.

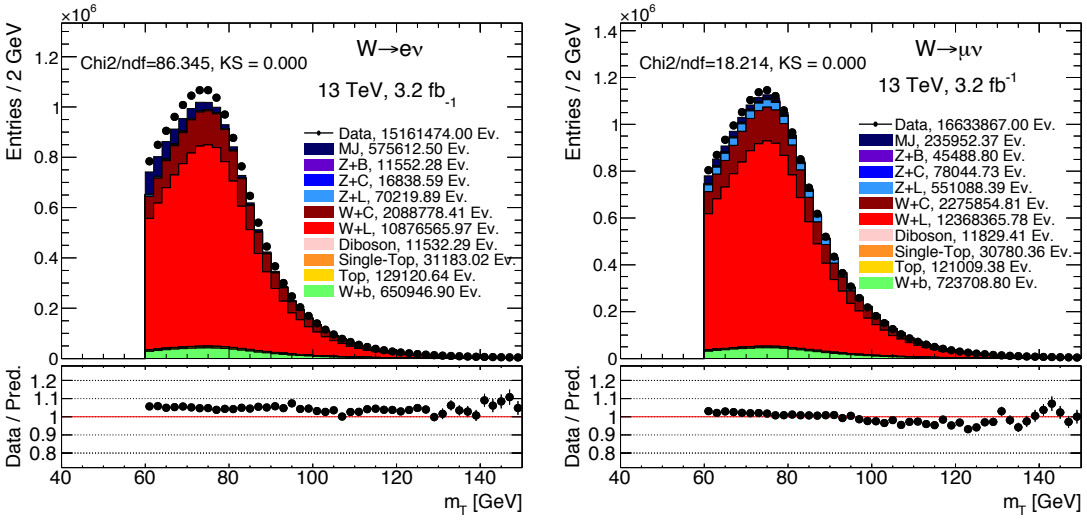


FIGURE A.14: The transverse mass before applying any selection on the number of jets or b-jets.

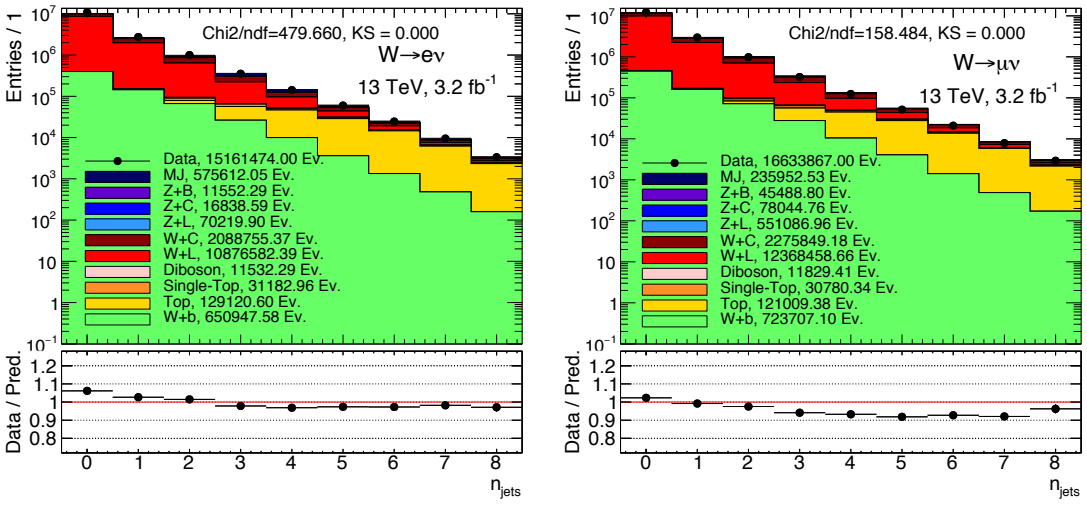


FIGURE A.15: The jet multiplicity for jets in the calorimeter fiducial region.

QCD Background

A combination of the seletion requirements used in the analysis, for both channels (high E_T^{miss}, large m_{T} or Z mass selection, isolated leptons) is indeed efficient in rejecting the QCD background. Nevertheless, given the very large cross section of QCD jet production, residual events with very different characteristics can contaminate the selected events. With respect to the Z-channel, this effect is more pronouced in the W-channelwhere the requirement of only one isolated lepton, and large missing energy can lead to a QCD contamination. The type of QCD events that can contaminate the two signal selections can arise from heavy quark leptonic decays, conversions in the detector material, and hadrons in general, and can vary in composition based on the p_{T} range considered. The simulation of such events is very difficult, and in any case leads to too low statistics. One of the techniques that can be used to reject backgrounds coming from fake leptons is based on the selection of a fake-enriched data sample. This sample can be obtained by relaxing or inverting the isolation or identification requirements on the leptons (*fake template selection*), and subsequently normalising this template from data in a control region (CR), which is selected to be similar to the signal region of the measurement (SR), but background-enriched. To extract the normalisation, a maximum-likelihood template fit to the data can be used, with the templates form the signal and other backgrounds than the fake one being estimated are taken from MC simulation. The scale factor obtained from this fit is then applied to the number of fake template events satisfying the SR selection.

This procedure has of course some weaknesses, that arise from:

- the arbitrariness of the template selection;
- the arbitrariness of the choice of the discriminant variable to fit;
- possible biases in the composition and kinematics of the chosen template with respect to the events containing non-prompt leptons or fakes passing the signal selection;
- the subtraction from the fake template of the contamination coming from prompt leptons produced by signal events or other electroweak processes.

To take into account these weaknesses, systematic uncertainties needs to be evaluated.

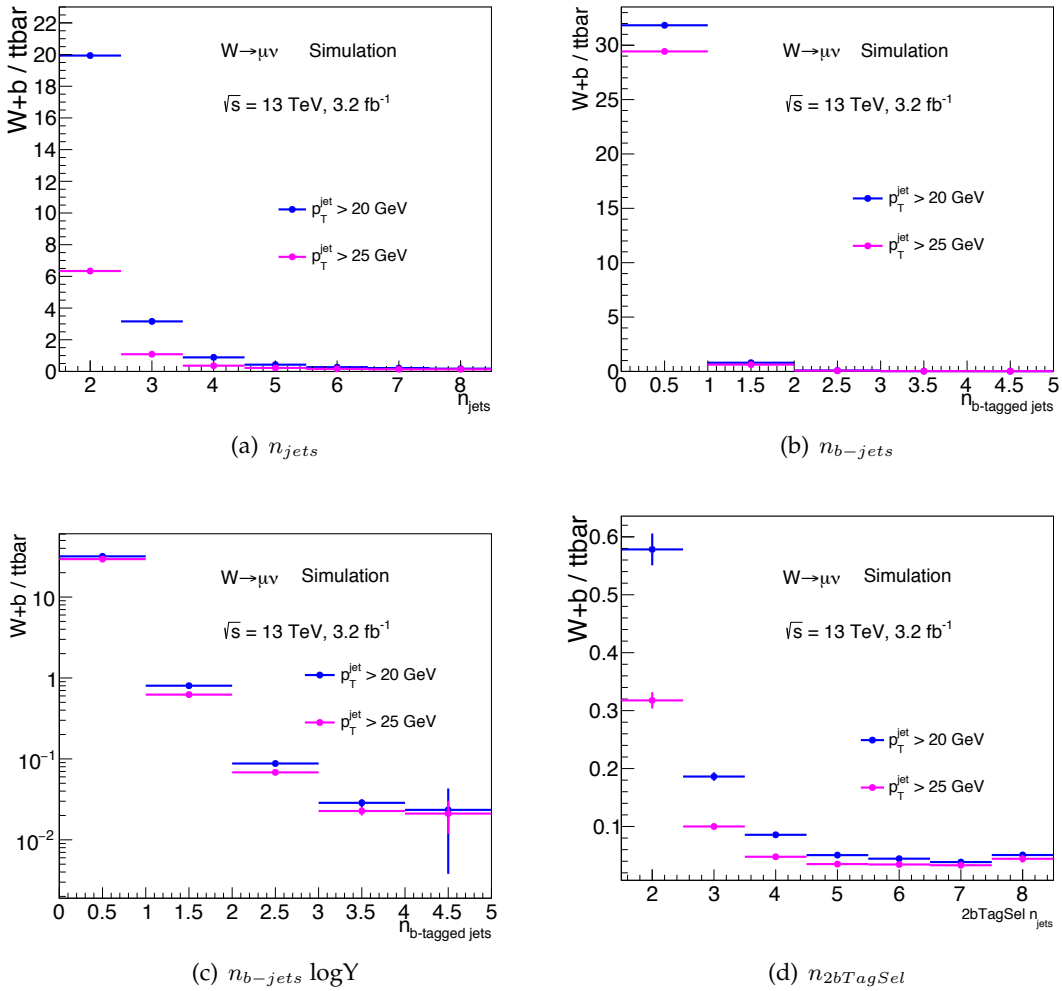
Top Background

One of the first test, which was performed by the Author of this thesis with the aim of reducing the top-background contamination in the signal region, was a study on the p_{T} threshold in the jet selection. It was originally set to be greater than 25 GeV, but it was found that by loosening this cut from 25 GeV to 20 GeV, a big gain can be obtained in the signal/background ratio, as shown in Figure A.16. Thus, in the final object selection, a $p_{\text{T}} > 20$ GeV will be required to the jets.

Light-jet calibration

The identification of jets originating from the b-quark hadronization (b-tagging) is an important element of a number of prominent analyses performed with the ATLAS detector at the Large Hadron Collider (LHC): precise Standard Model measurement, study of Higgs-boson and top-quark properties, exploration of New Physics scenarios.

The b-tagging of a jet relies on the property of the b-hadrons to have long lifetime τ ($\tau \sim 1.5$ ps, corresponding to a free path of about $c\tau \sim 450$ μm) and large mass, resulting in the production of secondary decay vertices inside the jet cone and a large multiplicity



(a) n_{jets} (b) n_{b-jets}

(c) n_{b-jets} logY (d) $n_{2bTagSel}$

FIGURE A.16: Ratio between the W+b-jets and the $t\bar{t}$ sample as a function of (a) n_{jets}, (b) n_{b-jets}, (b) n_{b-jets} in log scale and (d) $n_{2bTagSel}$.

of decay products. These observables are reconstructed thanks to the charged-particle tracking capability of the ATLAS Inner Detector. The information is then combined using a multi-variate algorithm able to discriminate in an optimal way a jet originating from a b-quark with respect to one originating from a c-quark or a light-flavour parton (LF), i.e. u, d, s-quarks or gluons. Specific selections on the output weight distribution of a given b-tagging algorithm are called working points (WPs) and defined as a function of the average efficiency of tagging a b-jet as measured in simulation.

The objects used in the measurements are defined as follows.

- **Primary vertex:** the events are required to have at least one reconstructed primary vertex, with at least two tracks of $p_{\mathrm{T}} > 0.4$ GeV, consistent with the average beam-spot position. If more than one primary vertex candidate is present in the same event, the one with the highest $\sum p_{\mathrm{T}}^2$ of the associated tracks is chosen.

- **Calorimeter jets:** jets are reconstructed from topological clusters [210] formed from energy deposits in the calorimeter with the anti-k_t algorithm and a radius parameter of 0.4 [211]. The clusters are calibrated at the electromagnetic (EM) scale. The measured jet energies are corrected using the jet area method [212, 213] to reduce

the effects due to additional proton-proton interactions in the same or neighbouring bunch crossings (pileup) and the underlying event. The jet energy at the hadronic scale is then obtained by applying calibration factors depending on transverse momentum and pseudorapidity [214]. Inner detector tracks are associated with the calorimeter jets based on their angular distance $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$. The exact track selection criteria depend on the b-tagging algorithm and are detailed in Ref. [215, 216]. Calorimeter jets entering this analysis are required to have a transverse momentum $p_{\text{T}}^{\text{jet}} > 20$ GeV and to be within the acceptance of the inner detector, limited to pseudorapidities of $|\eta^{\text{jet}}| < 2.5$, where b-jet identification (which relies on the high spatial resolution of the inner detector) is possible. In order to reduce the contamination of jets arising from additional pileup interactions, the jet-vertex tagger tool (JVT) [217], building a 2-dimensional likelihood from tracking and vertexing information, is used. The JVT output is required to be greater than 0.59 for jets with $p_{\text{T}}^{\text{jet}} < 60$ GeV and $|\eta^{\text{jet}}| < 2.4$, which corresponds to the *medium* JVT working point. Additional cleaning requirements are applied to ensure a minimum non-collision background while keeping a sufficiently large sample [218].

- **Track jets:** track jets are reconstructed from inner detector tracks using the anti-k_t algorithm with a radius parameter of 0.2. The tracks are required to have $p_{\text{T}} > 0.5$ GeV with at least one hit in the pixel detector and six hits in the silicon strip detector, and to be tightly matched to the primary vertex using impact parameter thresholds on the tracks. Such thresholds greatly reduce the number of tracks from pileup vertices whilst being highly efficient for tracks truly originating from the hard-scatter vertex. Once the track jet axis is determined, a second step of track association is performed and tracks with looser impact parameter requirements are associated with the jet. This allows to collect the tracks needed for effectively running the b-tagging algorithms. Only track jets associated with at least two tracks, $p_{\text{T}}^{\text{jet}} > 10$ GeV and $|\eta^{\text{jet}}| < 2.5$ are used for the analysis. No calibration nor further corrections are applied.

- **Truth jets:** truth calorimeter jets are reconstructed from all truth particles excluding muons and neutrinos with the anti-k_t algorithm and a radius parameter of 0.4. Corrections based on the truth particle record and excluding muons, neutrinos and particles from pileup interactions are computed from the jet area method to reduce the effects due to the underlying event. Only truth jets with $p_{\text{T}}^{\text{jet}} > 10$ GeV are kept in the truth record. MC events are rejected if they do not include at least two truth jets.

- **Jet flavour:** in the simulation, a jet flavour label is assigned using particle level information by matching jets to weakly decaying b and c-hadrons with $p_{\text{T}} > 5$ GeV, in a cone of radius $\Delta R = 0.3$. The flavour labelling is exclusive with the hadron matched only to the closest jet. If a b-hadron is found within the cone, the jet is labelled as a b-jet. If no b-hadron is found, the search is repeated for c-hadrons, then for τ-leptons. If no match is found for b-hadrons, c-hadrons or τ-leptons, the jet is labelled as a light-flavour jet.

- **B-tagging:** experimentally, jets are identified as originating from a b-quark using the MV2c10 discriminant, a multivariate analysis which spans values from -1 to 1, with b-flavour (light-flavour) jets yields values closer to 1 (-1), respectively. Different selections on the MV2c10 value correspond to different nominal values of the efficiency for b-jets produced in $t\bar{t}$ events and are called working points. The WPs that are calibrated in this study have efficiencies of 60%, 70%, 77% and 85% when using the nominal $t\bar{t}$ simulated sample.

The performance of b-tagging is characterised by the probability of tagging a true b-jet (ε_b) and the probabilities of mistakenly tagging as a b-jet a jet originating from a c-quark (ε_c) or a LF (ε_l), referred to as "mistag rates'" in the following. Ideally, Monte Carlo (MC) simulations including the various quark flavours could be used to evaluate the performance of b-tagging and the mistag rates. However, additional calibration of the MC is often needed to account for differences between data and simulation.

The calibration results are provided as data/MC efficiency scale factors:

$$SF = \frac{\varepsilon^{\text{data}}}{\varepsilon^{\text{MC}}}, \tag{A.4}$$

which depend on the b-tagging algorithm, the associated working point, the quark flavour and the jet transverse momentum, $p_{\text{T}}^{\text{jet}}$. The SFs for the LF mistag rate have also been studied as a function of the absolute value of the jet pseudorapidity, $|\eta^{\text{jet}}|$.

To measure the mistag rate of light-flavour jets, the ATLAS collaboration usually relies on a data-driven technique, the so-called "negative tag method". In this thesis, an alternative method called "adjusted-MC method" is presented since the Author of this thesis has contributed to its implementation, together with others. A similar approach was successfully used already in Run 1 for high-p_{T} b-jets tagging efficiency and it is being re-applied in Run 2.

The method is based on a bottom-up approach where the underlying tracking variables are adjusted to match the data and the effect is then propagated to the b-tagging observables. This can be summarized in the following steps:

1. Accurate analysis of ID tracking performance, in both data and simulation, in terms of track impact-parameter resolution, track reconstruction efficiency, and track fake rate, as detailed in Chapter 3.

2. The MC is adjusted, producing what can be dubbed a "MC-adj'" sample, to reproduce the given tracking performance in data.

3. Re-evaluation of the b-tagging algorithm on the adjusted MC and extraction of the new mistag rate $\varepsilon_l^{\text{MC-adj}}$. The LF-jet scale scale factors, SF_l, with respect to the standard (i.e. not modified) MC are then given by:

$$SF_l \equiv \frac{\varepsilon_l^{\text{MC-adj}}}{\varepsilon_l^{\text{MC}}} \tag{A.5}$$

Basically the adjusted-MC calibration method relies on data-driven calibration, but it is performed one step before the b-tagging algorithm evaluation. This is a complementary approach to the negative-tag method, described in the previous sections, where the data-driven calibration is performed using a control sample enriched in LF-jets.

The adjusted-MC method is particularly suitable for the calibration of the mistag rate because experimental effects, like the greater-than-zero IP-resolution for prompt tracks, are supposed to be the major contributors to the LF-jet tag-rate. Systematic uncertainties are applied to cover for additional effects connected to the physics modeling of the LF-jets, like expected number of tracks inside a LF-jet or their momentum spectrum.

The procedure consist into applying the track systematic uncertainties on the track collection, and giving the modified version as an input to the b-tagging algorithms. The implementation of this procedure is based two software tools:

- `InDetTrackSystematicsAlgs`, an Athena algorithm able take as input a collection of tracks from the xAOD data-format and to produce a new collection of tracks where the effects of the tracking recommendations are propagated;

Tracking Performance Variation Name	Label	LF-Calibration notes
TRK_EFF_LOOSE/TIGHT_GLOBAL	eff-GLOBAL-L/T	Not relevant
TRK_EFF_LOOSE/TIGHT_IBL	eff-IBL-L/T	Not relevant
TRK_EFF_LOOSE/TIGHT_PP0	eff-PP0-L/T	Not relevant
TRK_BIAS_QOVERP_SAGITTA_WM	sagitta-bias WM	Not relevant
TRK_RES_D0/Z0_DEAD	dead modules	Not relevant
TRK_RES_D0_MEAS	d_0-smearing	Relevant
TRK_RES_Z0_MEAS	z_0-smearing	Relevant
TRK_FAKE_RATE_LOOSE	fake-rate	Relevant, Syst. variation
TRK_RES_D0_MEAS_UP	d_0-smearing up	Syst. variation
TRK_RES_Z0_MEAS_UP	z_0-smearing up	Syst. variation
TRK_RES_D0_MEAS_DOWN	d_0-smearing dw	Syst. variation
TRK_RES_Z0_MEAS_DOWN	z_0-smearing dw	Syst. variation
TRK_RES_D0Z0_MEAS	$d_0 z_0$-smearing	Syst. variation

TABLE A.4: List of effects considered for the adjusted-MC calibration of the LF-jet tag rate. For details about the tracking performance variation see [107] and Chapter 3.

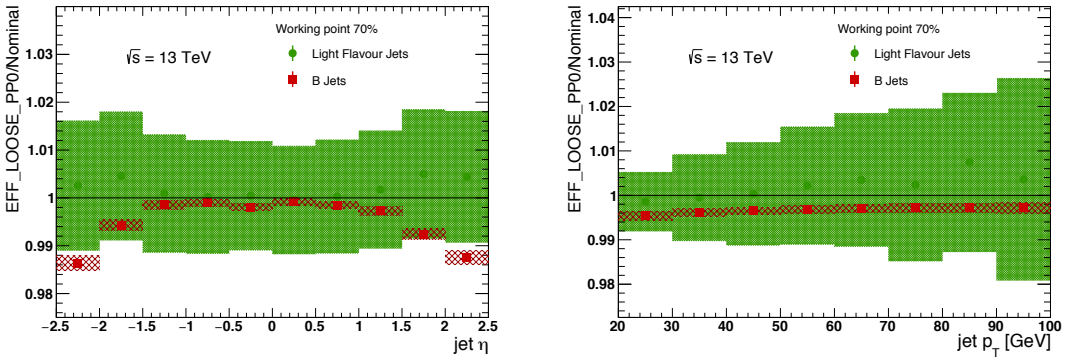
- `Run2BtagOptimisationFramework` or "re-tagging" framework, an Athena algorithm which takes as an input xAOD track and jet collections and re-evaluate the b-tagging algorithms (i.e. the so called re-tagging) on them.

Table A.4 list all the studied tracking performance variations, implemented in the `InDetTrackSystematicsAlgs` package. For details about the tracking performance variation see [107] and Chapter 3 . After some preliminary check on two different Monte Carlo event generators, namely Herwig++ [39] and Powheg [109], the majority of the effects were found to be negligible for the light jets within the uncertainties, as it can be seen for example in Figure **??** in which the effect of a 50% extra material in the forward region of the Inner Detector was studied for the loose track selection (see Chapters3 and 4). On the other hand, some of the tracking uncertainties were found to have a big impact on this procedure. In particular the impact of the fake rate was up to 5%, as shown in Figure A.18 and in Figure A.18, and the effect of the d_0 impact parameter resolution, shown in Figure A.20, was of up to 70% and, as it can be seen in Figure A.21, for z_0 it was of up to 40%.

Updated Light-jet calibration The studies shown above have been re-done after the data and simulation samples were reprocessed with the Athena release 20.7. As already mentioned in Chapter 3, the tracking recommendations have been updated. The most relevant ones for the adjusted-MC method are related to the impact parameters and are shown in Figures A.22 and A.23.

Each of the tracking performance variations is supposed produce a "small" change on the observables predicted by the ATLAS simulation, in this way the effects can be treated as a systematic uncertainties with respect to the baseline MC and each effect can be studied as a single element because, at first order, factorizing from the others.

However the preceding assumptions do not hold in the case of the adjusted-MC LF-jet calibration because it was observed that the impact of z_0 and d_0 smearing on the mistag-rate is large; as expected because IP resolution has a major role in the tagging of LF-jets. Furthermore several variations must be combined to give the total mistag-rate calibration, SF_l, therefore correlations may need to be studied and taken into account.



(a) Powheg (b) Powheg

FIGURE A.17: Effect of the tracking systematic uncertainty due to 50% of extra material in the PP0 region of the ID for the Powheg event generator as a function of (a) η and (b) p_{T} respectively. The distributions are shown for light-flavour jets in green and for b-jets in red.

In the case of perfect decorrelation between the tracking performance variations, the combination of them resulting in the total mistag-rate calibration scale-factor is:

$$SF_l = \prod_{var} SF_{l,var}, \tag{A.6}$$

where var indicates each of the tracking performance variations relevant for the final computation of SF_l.

Between the tracking performance variations reported in Tab. A.4, as already discussed three of them have a sizable effect on the mistag-rate: d_0-smearing, z_0-smearing, fake-rate. Therefore, following Eq. A.2.4, we have:

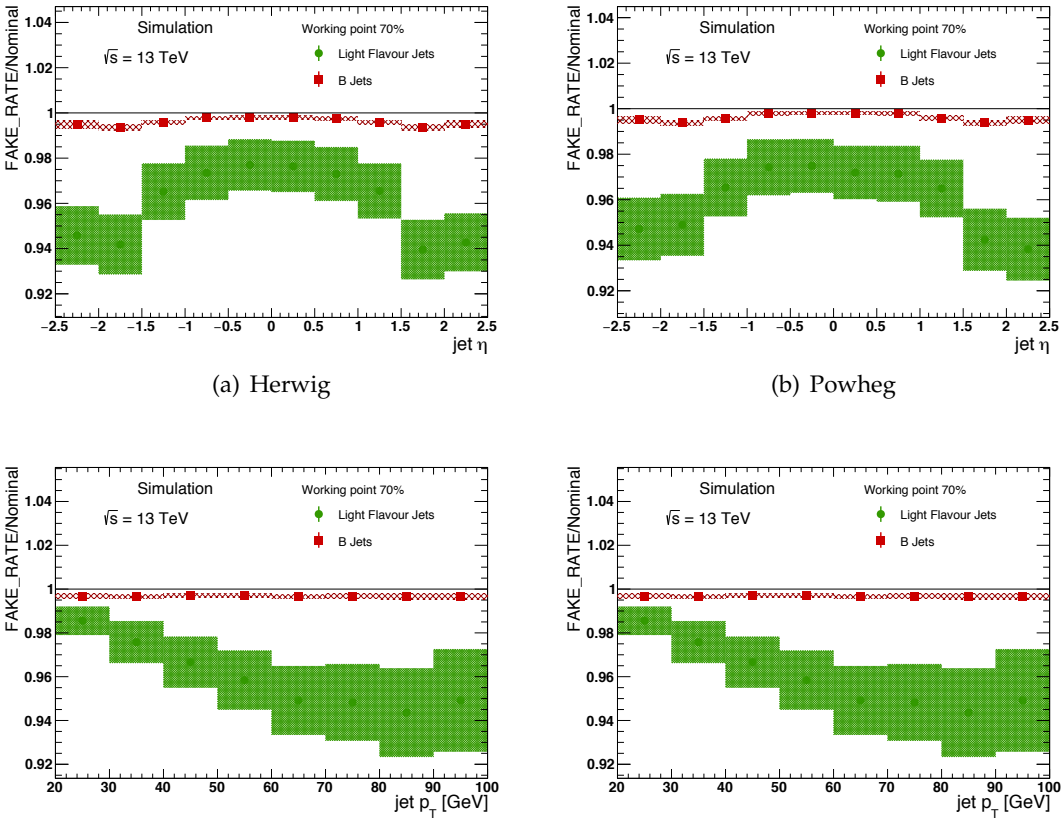
$$SF_l = SF_{l,d_0\text{-smearing}} \cdot SF_{l,z_0\text{-smearing}} \cdot SF_{l,\text{fake-rate}}. \tag{A.7}$$

After the combination, correlations need to be studied and, if needed, taken into account. The focus is on the study on the IP smearing variations because the effect on the mistag-rate is about one order of magnitude larger than the fake-rate variation, so the effect of correlations between IP smearing and fake-rate should be negligible.

Correlations due to intrinsic nature of the multivariate optimization of the MV2c10 discriminant, have been studied by producing a new sample where both the d_0 and the z_0 smearing are applied simultaneously, but independently, on each track to the events. The results from the $d_0 z_0$-smearing and from the combination of the d_0 and the z_0 smearing using Eq. A.2.4 are in relatively good agreement although the correlations introduce a difference of $\approx 5\%$. This difference is taken as a systematic uncertainty on the final result.

Other sources of systematic uncertainties are considered on the final mistag-rate calibration scale-factor measured with the adjusted-MC method:

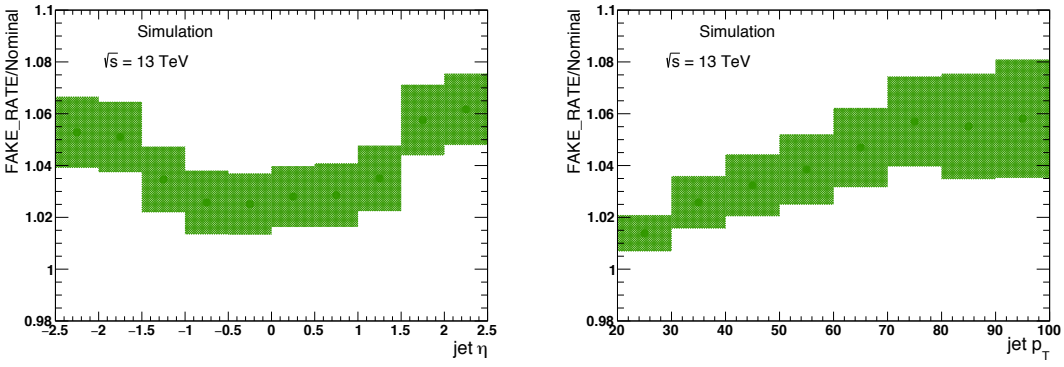
- **MC statistics:** a total of about $20 \cdot 10^6$ events of $t\bar{t}$ MC (one half of the total available). This uncertainty is relatively large for the high-purity WPs where a large statistics of simulated LF-jets is needed to obtain a precise evaluation of SF_l as a function of p_T and η.



(a) Herwig (b) Powheg

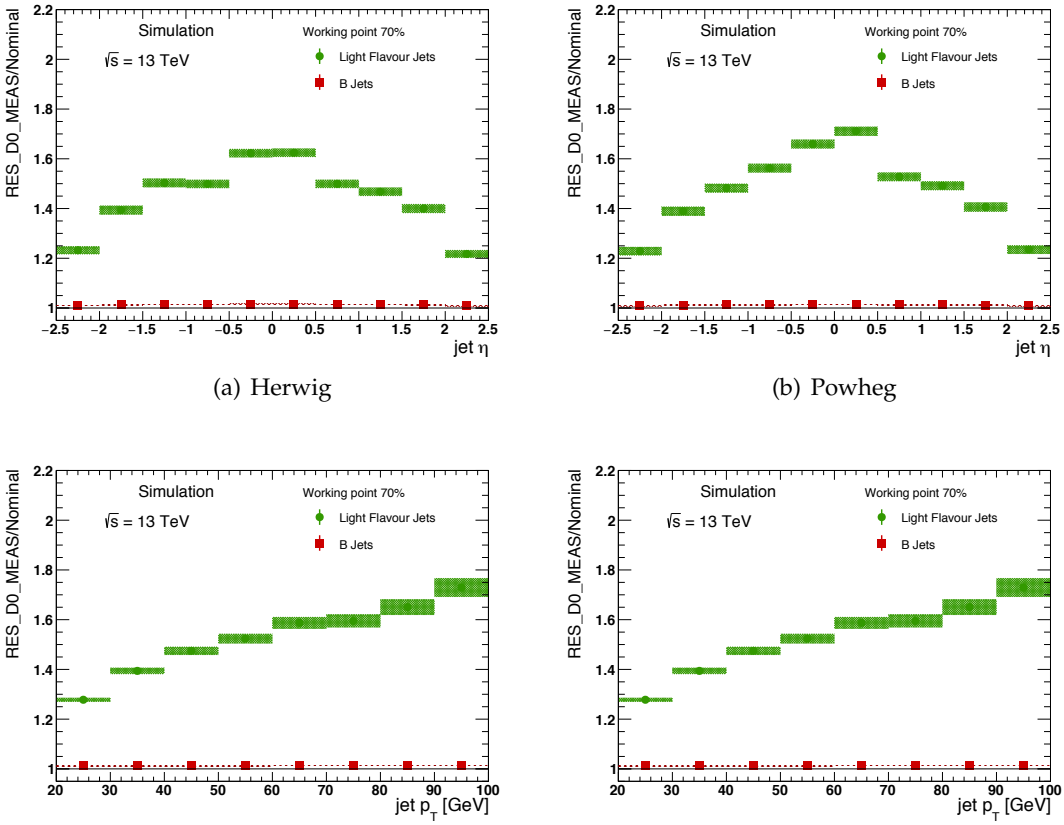
(c) Herwig (d) Powheg

FIGURE A.18: Effect of the tracking systematic uncertainty due to the fake rate in (a) and (c) for the Herwig event generator as a function of η and p_{T} respectively, and (b) and (d) for the Powheg event generator as a function of η and p_{T} respectively. The distributions are shown for light-flavour jets in green and for b-jets in red. The effect needs to be symmeterised as it can be seen in Figure A.19.



(a) Powheg (b) Powheg

FIGURE A.19: Effect of the tracking systematic uncertainty due to the fake rate for the Powheg event generator as a function of (a) η and (b) p_{T} respectively. The distributions are shown only for light-flavour jets in green.



(a) Herwig

(b) Powheg

(c) Herwig

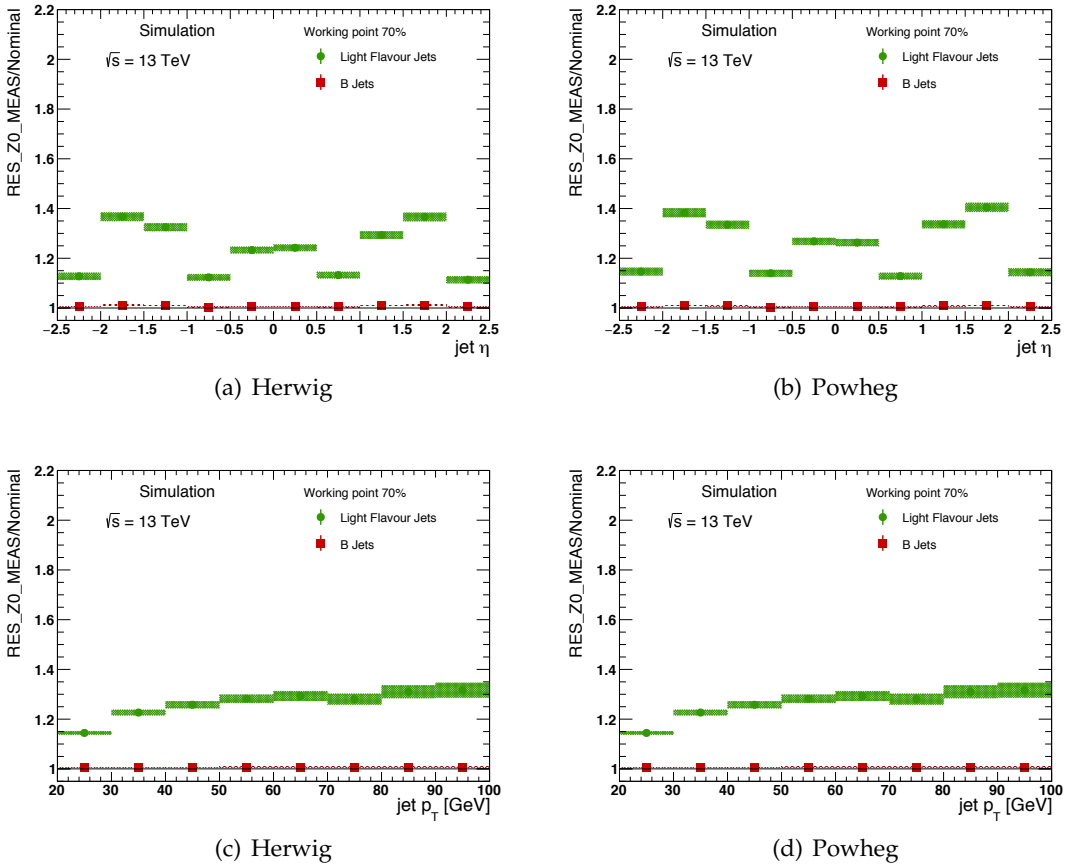
(d) Powheg

FIGURE A.20: Effect of the tracking systematic uncertainty due to the d_0 impact parameter resolution in (a) and (c) for the Herwig event generator as a function of η and p_{T} respectively, and (b) and (d) for the Powheg event generator as a function of η and p_{T} respectively. The distributions are shown for light-flavour jets in green and for b-jets in red.

- **IP-smearing uncertainty:** Propagation of d_0-smearing and z_0-smearing systematic uncertainties. The variations are provided as `UP/DOWN` of the smearing parameter maps. It is interesting to notice that, although the IP-smearing uncertainties are roughly symmetric, their impact on SF_l is not. This is due to the non-linear effect of the IP-smearing on the mistag-rate. An example of this is reported in Fig. A.24 where the d_0-smearing parameters have been artificially multiplied or divided by a factor of two, and the ratio between the resulting SF_l is reported.

- **IP-smearing correlations:** Comparison of SF_l obtained from simultaneous d_0z_0-smearing versus the combination of d_0 and z_0 smearings using Eq. A.2.4. The derived uncertainty is symmetrized around the nomial value of SF_l.

- **Fake-tracks:** the total effect of the fake-rate is removed from SF_l to evaluate a corresponding uncertainty. The uncertainty is symmeterised around the nominal value of SF_l. The uncertainty is along the lines of the track fake-rate calibration prescriptions by the tracking-CP group. The conservative approach was also motivated by the lack of precise study of the correlations between track fake-rate and IP smearing was performed, so this uncertainty.



(a) Herwig (b) Powheg

(c) Herwig (d) Powheg

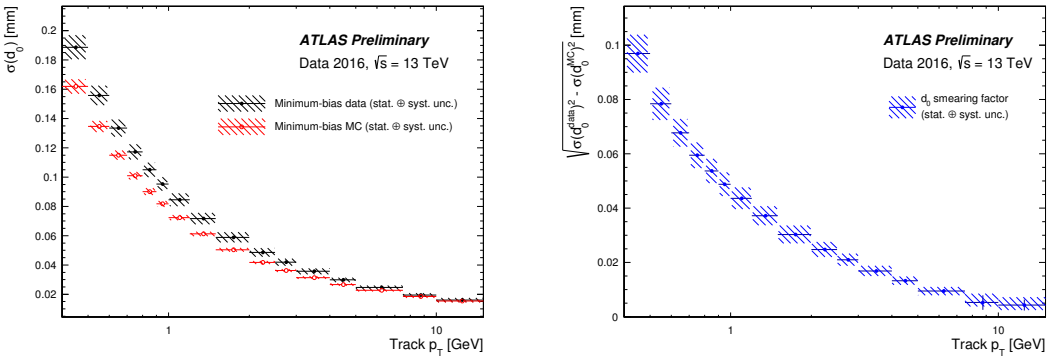
FIGURE A.21: Effect of the tracking systematic uncertainty due to the z_0 impact parameter resolution in (a) and (c) for the Herwig event generator as a function of η and p_{T} respectively, and (b) and (d) for the Powheg event generator as a function of η and p_{T} respectively. The distributions are shown for light-flavour jets in green and for b-jets in red.

- **LF-jet hadronization:** the SF_l results from two $t\bar{t}$ simulated samples, one using Pythia6 and the other Herwig as parton shower, are compared to asses an uncertainty on the LF-jet hadronisation model (charged-particle multiplicity and spectum, LF-hadron characteristics, etc.). The uncertainty is symmeterised around the nominal SF_l value. Although the two MCs share the same ME calculation, it is impossible to keep track of the same events and jets between the two samples, therefore this systematic uncertainty retains also a not negligible statistical component.

For each of the listed systematic uncertainties, a modified version of Eq. A.2.4 is evaluated and the difference with respect to the nominal value of SF_l is taken as uncertainty. The breakdown of the different systematic uncertainties can be seen in the bottom part of each of the Figures from A.25 to A.26.

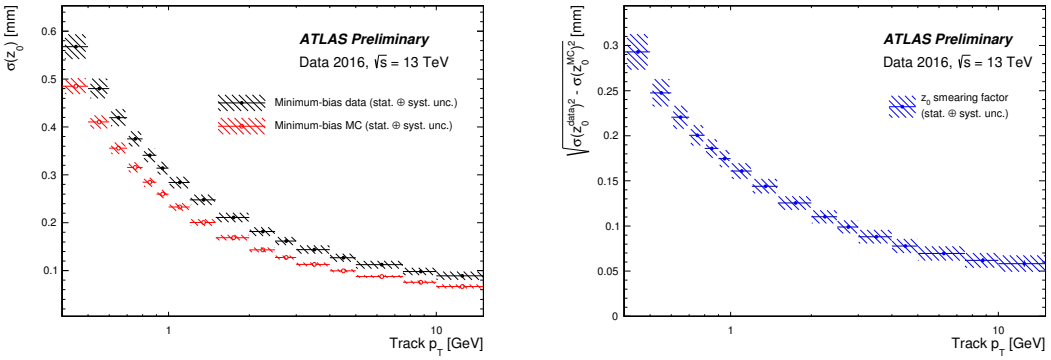
The preliminary results of the adjusted-MC mistag rate calibration are extracted using same p_T-η binning of negative tag method in order to compare the results. Equation A.2.4 and the corresponding systematic variations are applied for each of the p_T-η bins. Results are reported here in Figures A.25 and A.26 for fixed-cut-efficiency WPs 85% and 60%.

Figures A.27 and A.28 show the comparison of the negative-tag and of the adjusted-MC calibration methods. The results have been found in good agreement within the uncertainties.



(a) (b)

FIGURE A.22: Intrinsic transverse impact parameter resolution of tracks, $\sigma(d_0)$, measured in minimum-bias simulation and in 2016 data (a) and quadrature difference between the two (b). The transverse impact parameter resolution is extracted in fine bins of p_{T} and η of the tracks using iterative Gaussian fits, then averaged over η. Systematic uncertainties take into account: primary vertex resolution uncertainty, unfolding procedure, non-Gaussian tails of the resolution fit, data-taking period dependence assessed by comparing events collected in 2015 and in 2016 runs. Ref. [219].



(a) (b)

FIGURE A.23: Intrinsic longitudinal impact parameter resolution of tracks, $\sigma(z_0)$, measured in minimum-bias simulation and in 2016 data (a) and quadrature difference between the two (b). The transverse impact parameter resolution is extracted in fine bins of p_{T} and η of the tracks using iterative Gaussian fits, then averaged over η. Systematic uncertainties take into account: primary vertex resolution uncertainty, unfolding procedure, non-Gaussian tails of the resolution fit, data-taking period dependence assessed by comparing events collected in 2015 and in 2016 runs. Ref. [219].

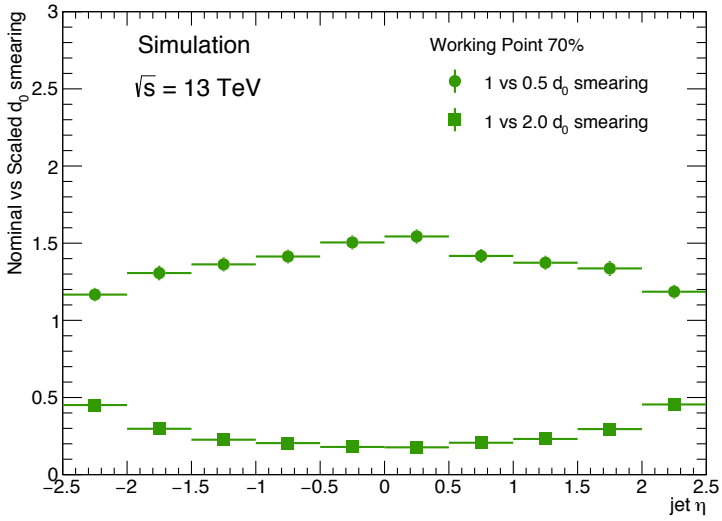


FIGURE A.24: Ratio of the SF_l value obtained from the d_0-smearing and the ones obtained after multiplying the d_0-smearing parameters by a factor of 2 (green squares) or 0.5 (green circle), for the 70% efficiency WP and as a function of jet η.

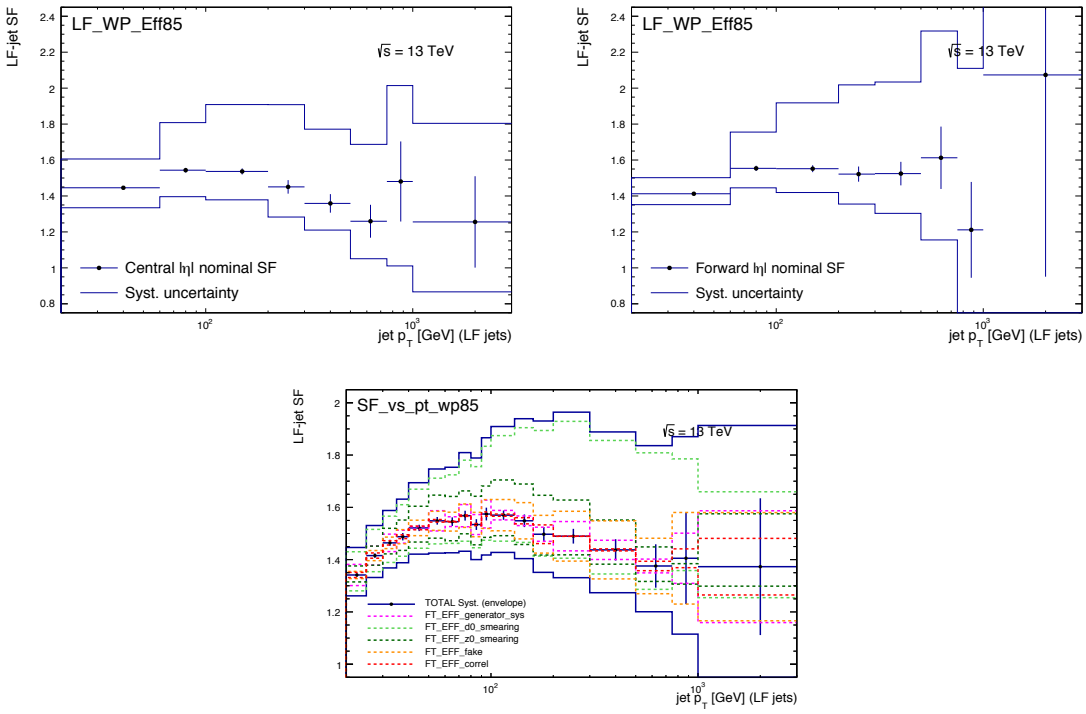


FIGURE A.25: Adjusted-MC LF-SF vs p_T for 85% efficiency WP. Top left: central eta ($|\eta| < 1.2$). Top right: forward eta ($1.2 < |\eta| < 2.5$). Bottom left: central eta ($|\eta| < 1.2$) syst. breakdown.

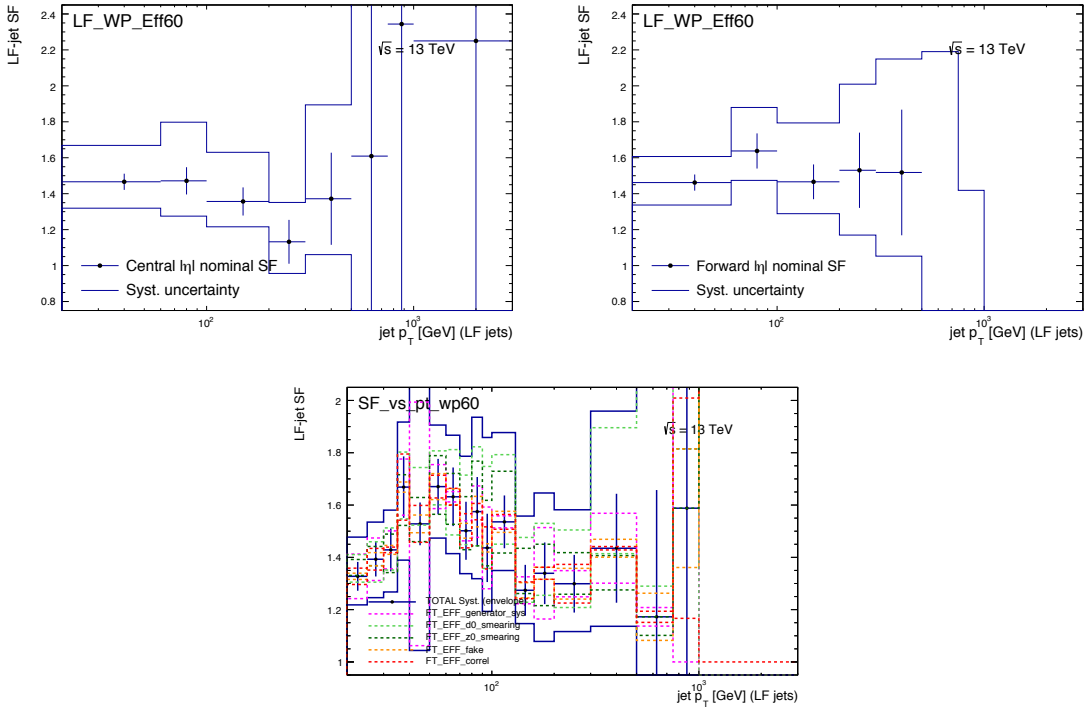


FIGURE A.26: Adjusted-MC LF-SF vs p_T for 60% efficiency WP. Top left: central eta ($|\eta| < 1.2$). Top right: forward eta ($1.2 < |\eta| < 2.5$). Bottom left: central eta ($|\eta| < 1.2$) syst. breakdown.

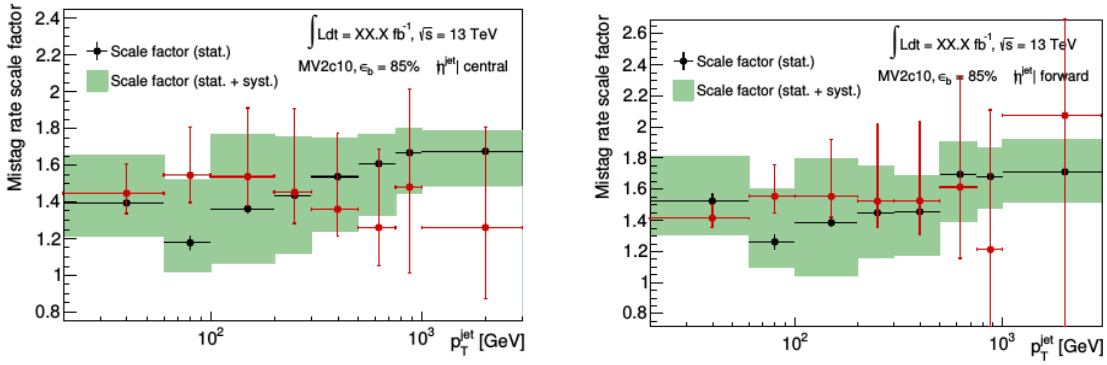


FIGURE A.27: Comparison of LF-SF vs p_T for 85% efficiency WP for the Negative-tag method and adjusted-MC method. Results from the Negative-tag method are shown with black lines (nominal values and statistical uncertainties) and green bands (total systematic uncertainties), results from the adjusted-MC method are shown with red lines (crosses showing total systematic uncertainties). Left: central eta ($|\eta| < 1.2$). Right: forward eta ($1.2 < |\eta| < 2.5$).

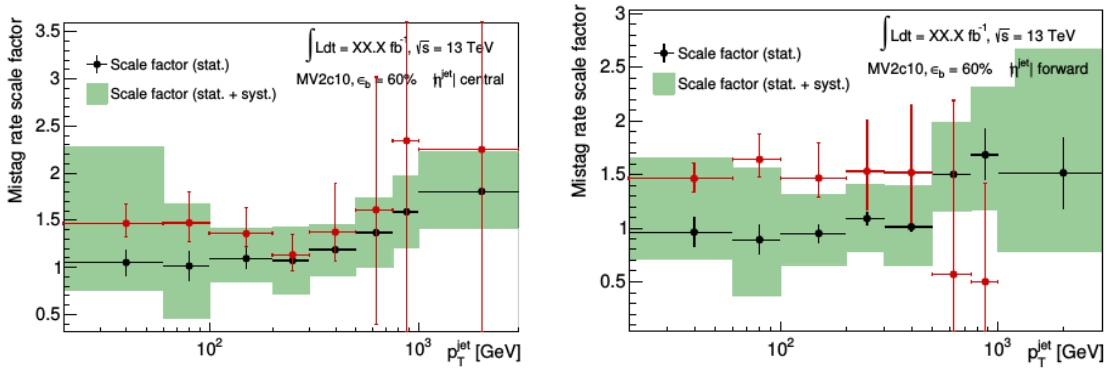


FIGURE A.28: Comparison of LF-SF vs p_T for 60% efficiency WP for the Negative-tag method and adjusted-MC method. Results from the Negative-tag method are shown with black lines (nominal values and statistical uncertainties) and green bands (total systematic uncertainties), results from the adjusted-MC method are shown with red lines (crosses showing total systematic uncertainties). Left: central eta ($|\eta| < 1.2$). Right: forward eta ($1.2 < |\eta| < 2.5$).

Appendix B

Track-Extension Efficiency Auxiliary Plots

In this Appendix auxiliary plots are presented to support the discussion in Chapter 4.

Section B.1 shows a collection of plots based on the preliminary selection criteria. These plots document differences and features related to the choice of Monte Carlo event generators and physics lists in the ATLAS detector simulation programs.

In Section B.2 the effects of the final selection criteria will be shown on distributions such as the track-extension efficiency as a function of ϕ and both the track-extension efficiency and the fraction of reconstructed tracklets for different particle species.

B.1 Without z_{vtx} cut and with 2 SCT Hits requirement

B.1.1 Primary Particle Flux: PYTHIA 8 and EPOS

Figures B.1 and B.2 show the SCT-extension efficiency as a function of pseudorapidity, η and transverse momentum, p_T, in a comparison between Pythia-8 (A2) and EPOS-LHC for different particles: pions, protons, kaons and others, where "others" are dominated by strange baryons. The main difference between Pythia 8 (A2) and EPOS-LHC can be seen in the descriptions of strange baryons.

B.1.2 Tracklet Composition: PYTHIA 8 and EPOS

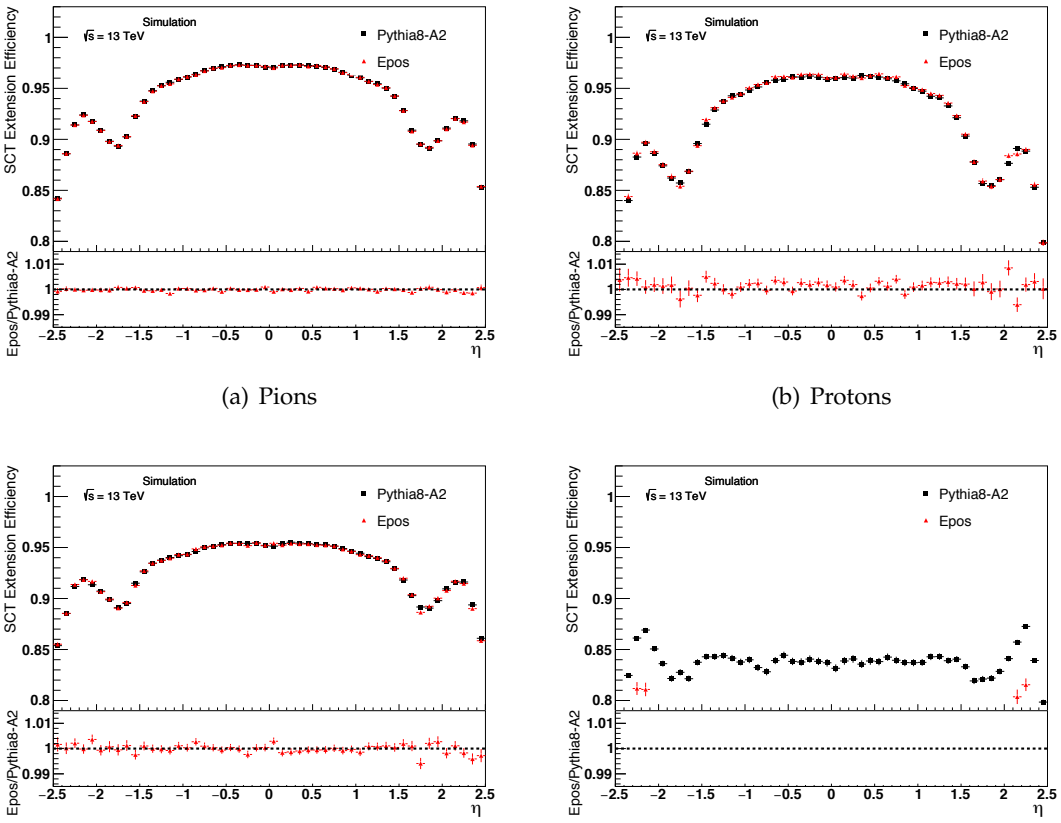
The reconstructed fraction of pixel track segments as function of p_{T}, η divided by category (all, primary, secondary) and by type (pions, protons, kaons, electrons, others) is shown for both PYTHIA 8 B.3 and EPOS B.4.

B.1.3 Primary and Secondary Tracklet Fraction per flavour: PYTHIA 8 and EPOS

The reconstructed fraction of pixel track segments as function of p_{T}, η divided by category (primary, secondary) and by type (all, pions, protons, kaons, electrons, others) is shown for both PYTHIA 8 B.5, B.6 and EPOS B.7, B.8.

B.1.4 SCT-Extension Efficiency per category and per flavour: PYTHIA 8 and EPOS

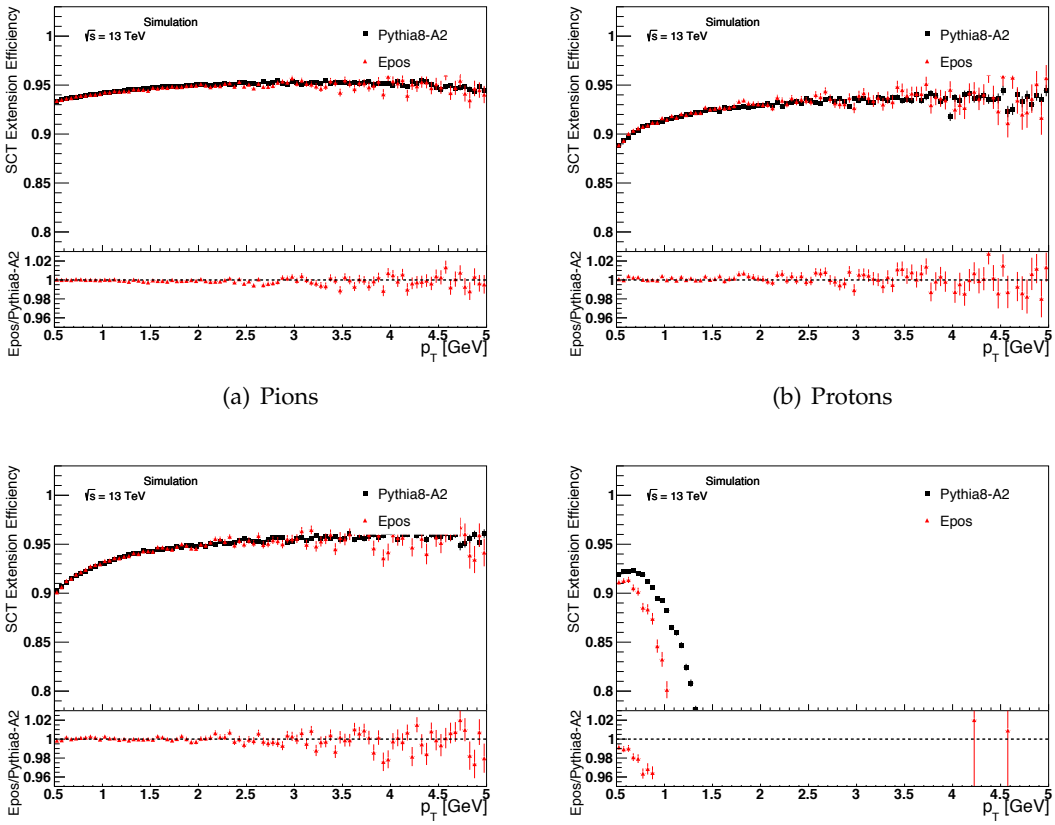
The $\mathcal{E}_{\mathrm{ext}}$ as a function of p_{T}, η divided by type (all, pions, protons, kaons, electrons, others) and by category (all, primary, secondary) is shown for both PYTHIA 8 B.9, B.10 and EPOS B.11, B.12.



(a) Pions (b) Protons

(c) Kaons (d) Others

FIGURE B.1: SCT-extension efficiency as a function of pseudorapidity η in a comparison between Pythia-8 (A2) and EPOS-LHC for (a) pions, (b) protons, (c) kaons and (d) others.



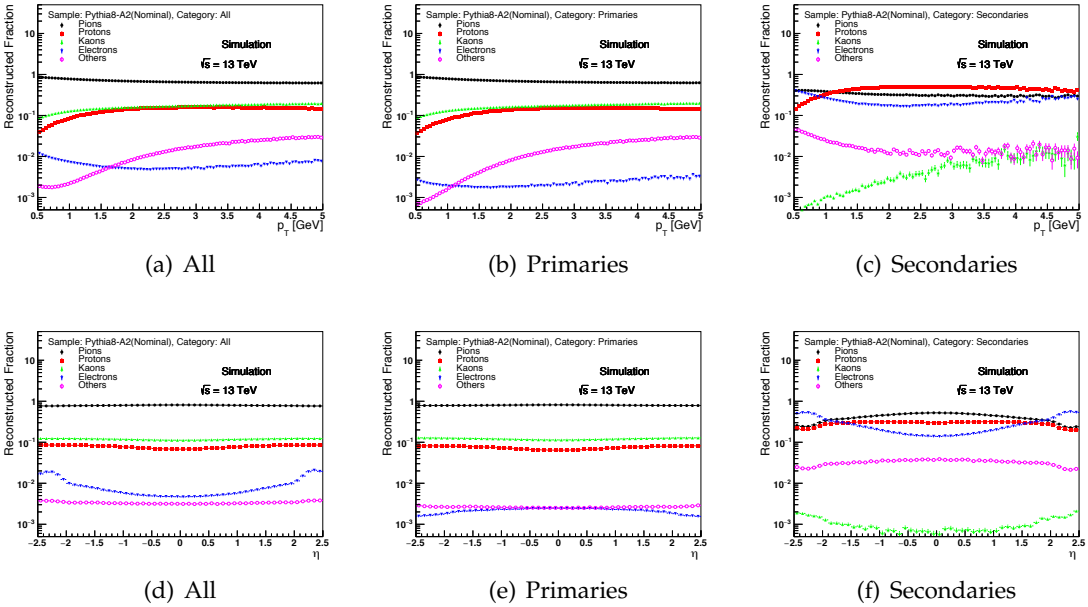
(a) Pions

(b) Protons

(c) Kaons

(d) Others

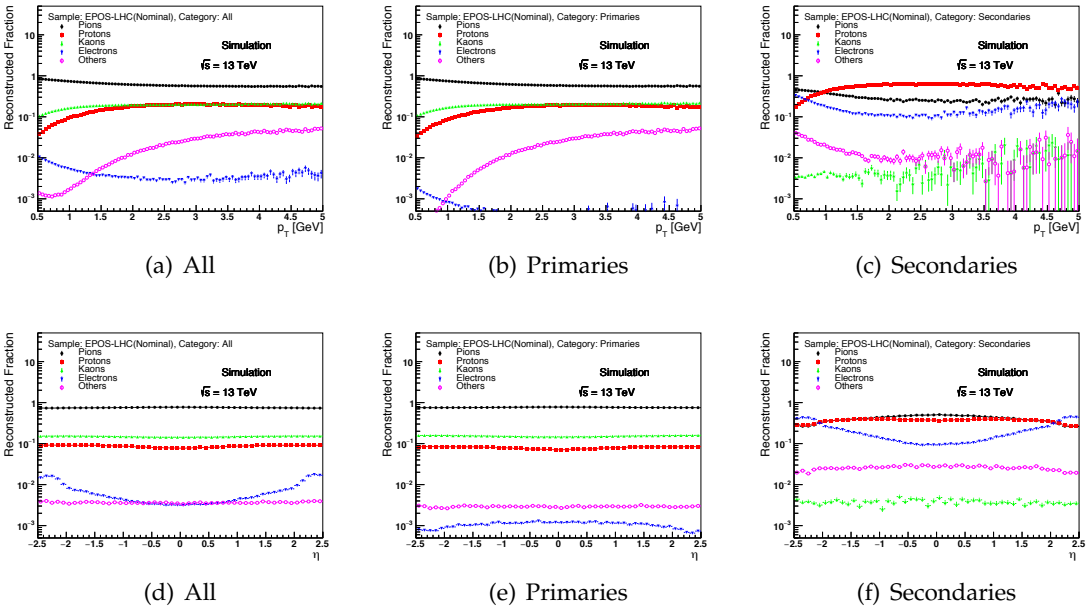
FIGURE B.2: SCT-extension efficiency as a function of transverse momentum p_T in a comparison between Pythia-8 (A2) and EPOS-LHC for (a) pions, (b) protons, (c) kaons and (d) others.



(a) All (b) Primaries (c) Secondaries

(d) All (e) Primaries (f) Secondaries

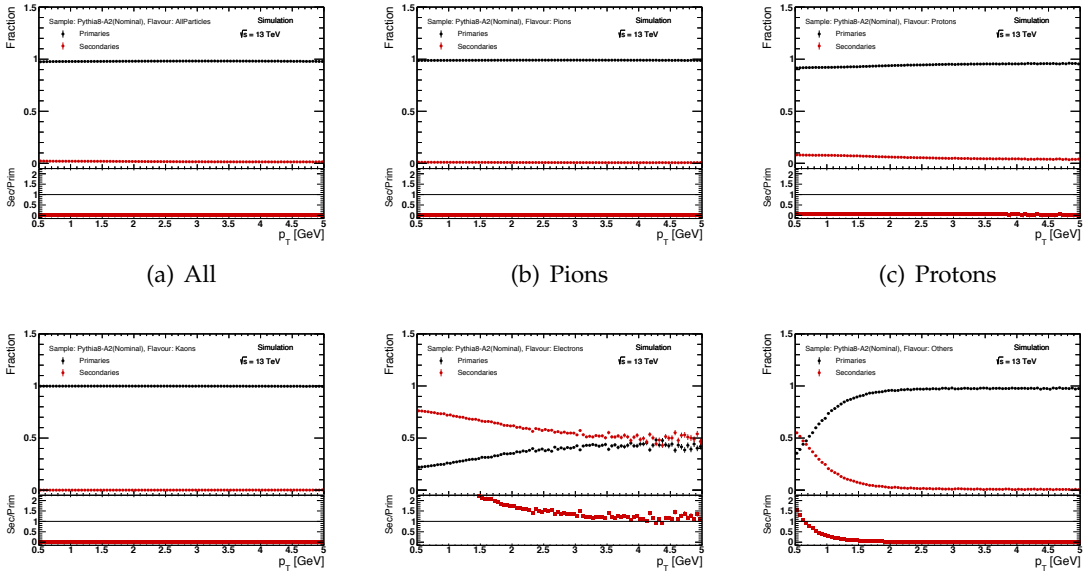
FIGURE B.3: Reconstructed fraction of pixel track segments as function of p_{T}, η divided by category (all, primary, secondary) and by type (pions, protons, kaons, electrons, others) for PYTHIA 8.



(a) All (b) Primaries (c) Secondaries

(d) All (e) Primaries (f) Secondaries

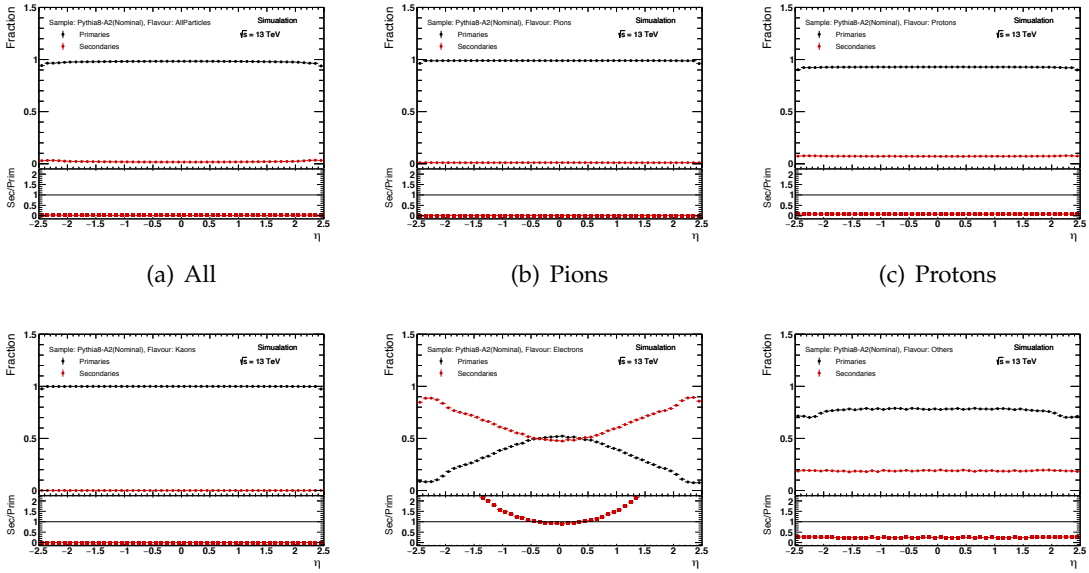
FIGURE B.4: Reconstructed fraction of pixel track segments as function of p_{T}, η divided by category (all, primary, secondary) and by type (pions, protons, kaons, electrons, others) for EPOS.



(a) All (b) Pions (c) Protons

(d) Kaons (e) Electrons (f) Others

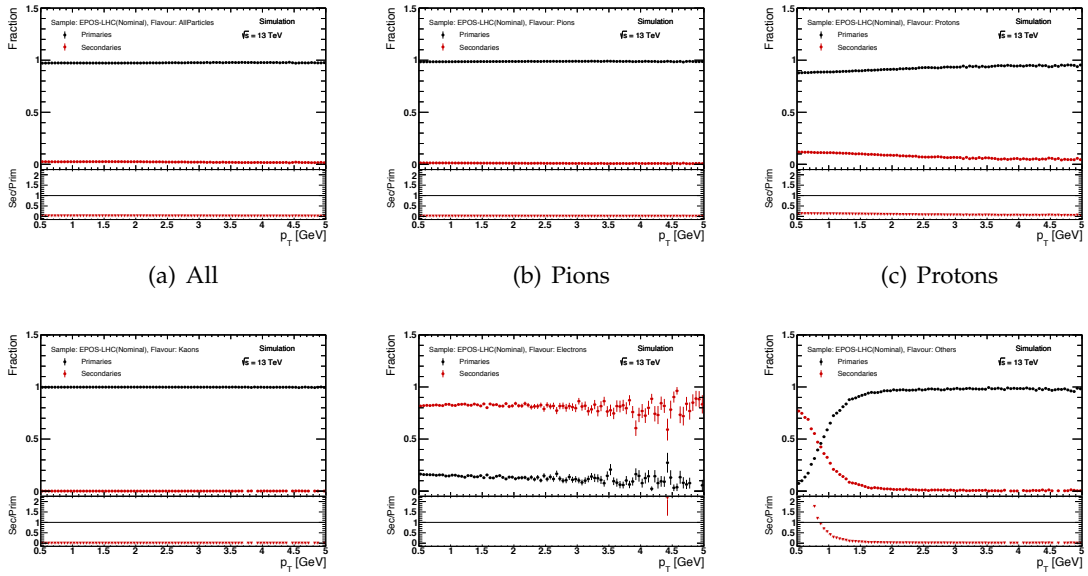
FIGURE B.5: Reconstructed fraction of pixel track segments as function of p_{T} divided by type (all, pions, protons, kaons, electrons, others) and by category (primary, secondary) for PYTHIA 8.



(a) All (b) Pions (c) Protons

(d) Kaons (e) Electrons (f) Others

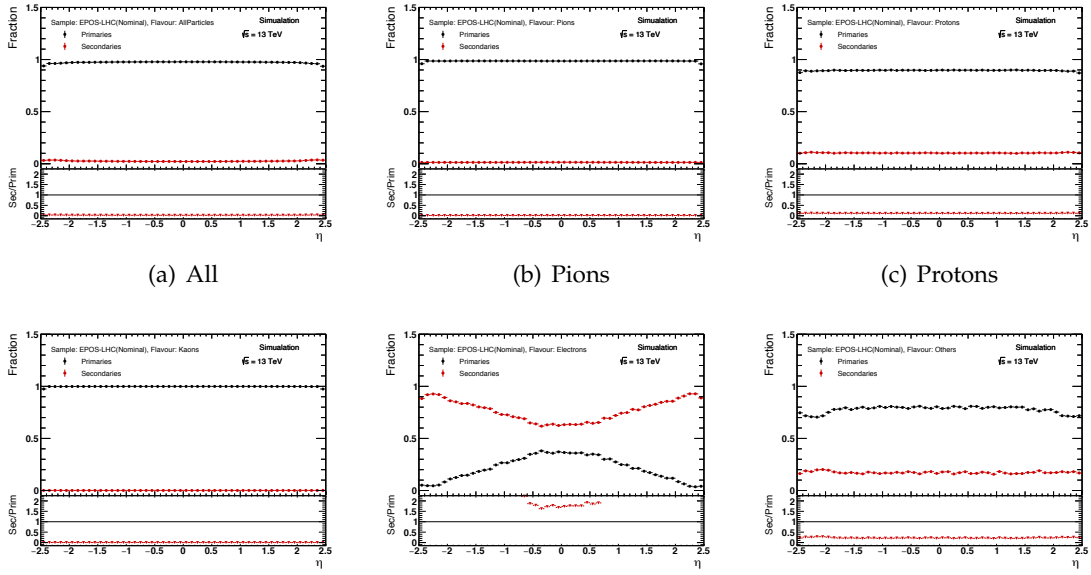
FIGURE B.6: Reconstructed fraction of pixel track segments as function of η divided by type (all, pions, protons, kaons, electrons, others) and by category (primary, secondary) for PYTHIA 8.



(a) All (b) Pions (c) Protons

(d) Kaons (e) Electrons (f) Others

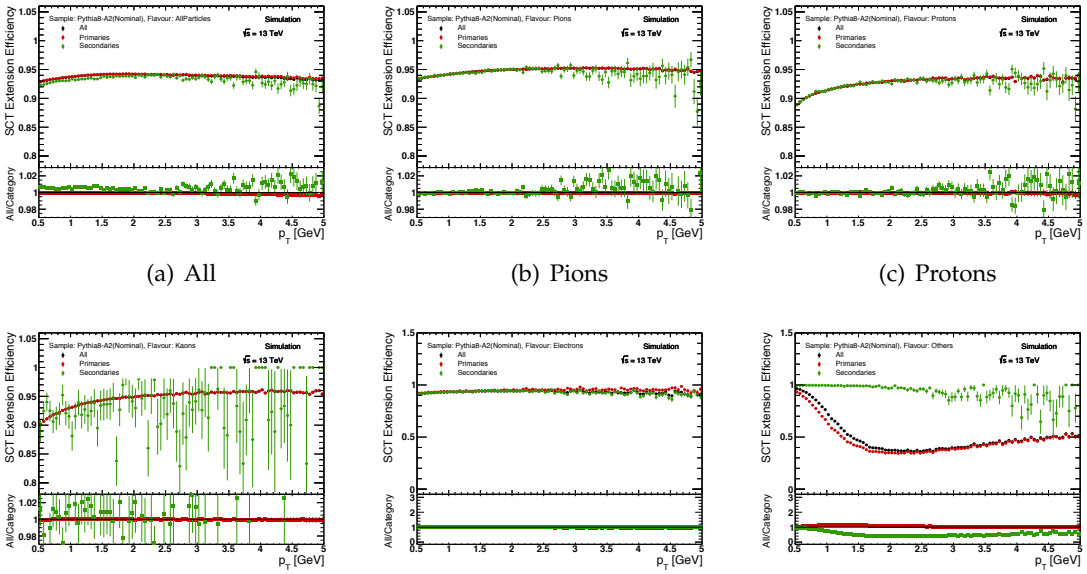
FIGURE B.7: Reconstructed fraction of pixel track segments as function of p_{T} divided by type (all, pions, protons, kaons, electrons, others) and by category (primary, secondary) for EPOS.



(a) All (b) Pions (c) Protons

(d) Kaons (e) Electrons (f) Others

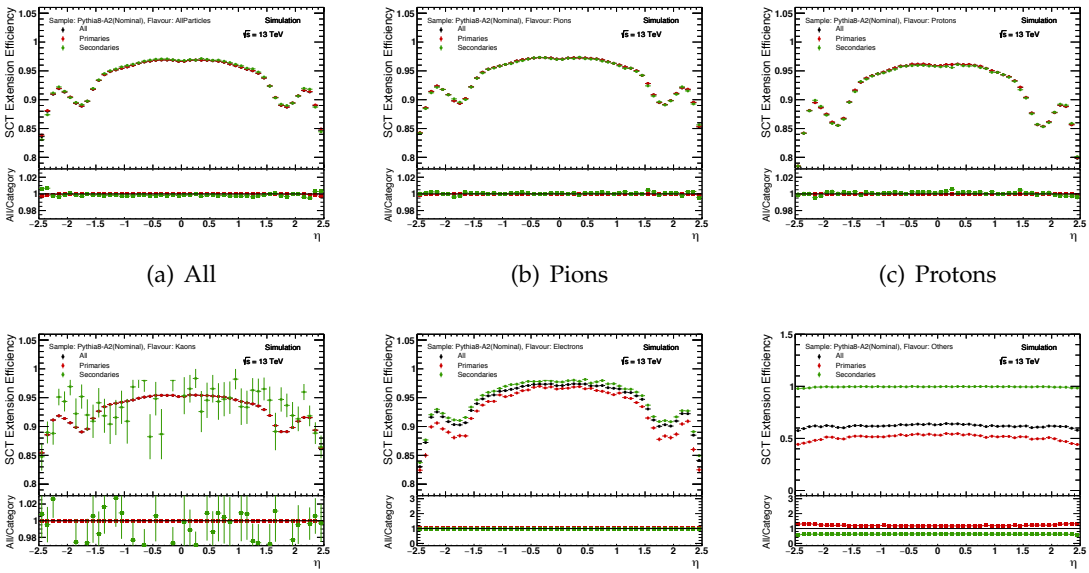
FIGURE B.8: Reconstructed fraction of pixel track segments as function of η divided by type (all, pions, protons, kaons, electrons, others) and by category (primary, secondary) for EPOS.



(a) All (b) Pions (c) Protons

(d) Kaons (e) Electrons (f) Others

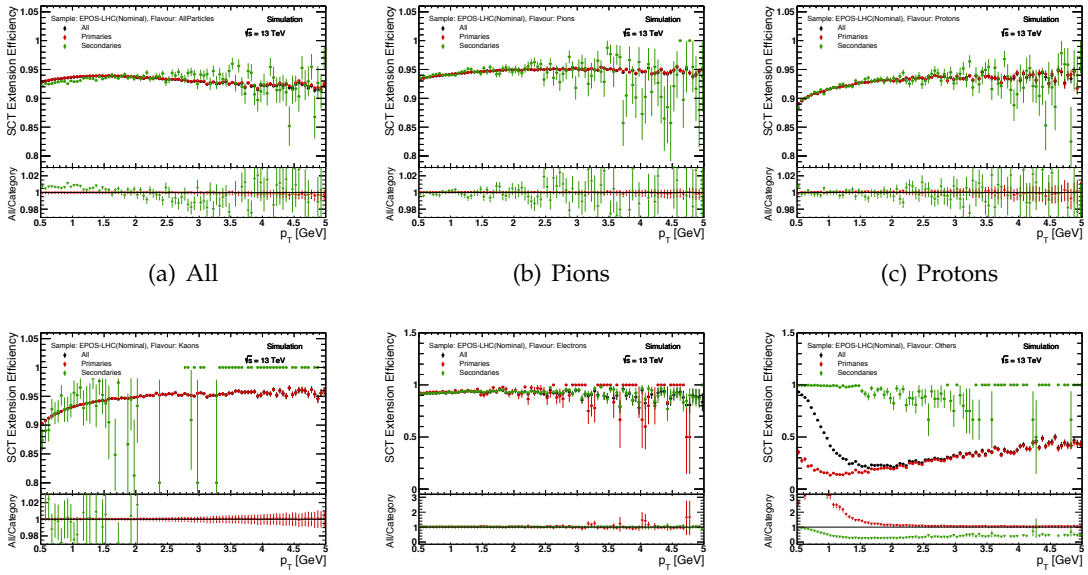
FIGURE B.9: $\mathcal{E}_{\text{ext}}$ as a function of p_{T} divided by type (all, pions, protons, kaons, electrons, others) and by category (all, primary, secondary) for PYTHIA 8.



(a) All (b) Pions (c) Protons

(d) Kaons (e) Electrons (f) Others

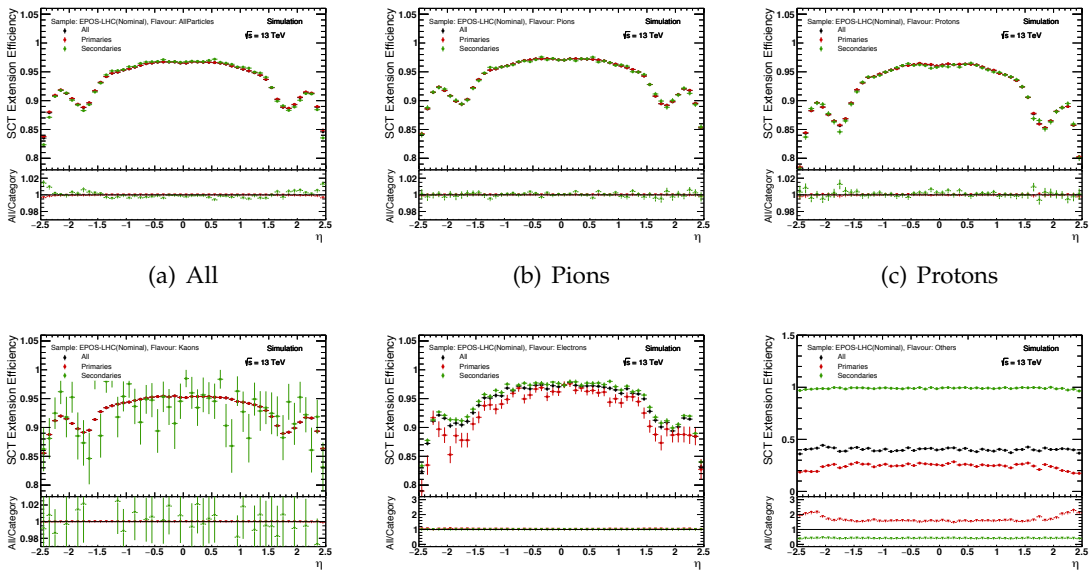
FIGURE B.10: $\mathcal{E}_{\text{ext}}$ as a function of η divided by type (all, pions, protons, kaons, electrons, others) and by category (all, primary, secondary) for PYTHIA 8.



(a) All (b) Pions (c) Protons

(d) Kaons (e) Electrons (f) Others

FIGURE B.11: $\mathcal{E}_{\text{ext}}$ as a function of p_{T} divided by type (all, pions, protons, kaons, electrons, others) and by category (all, primary, secondary) for EPOS.



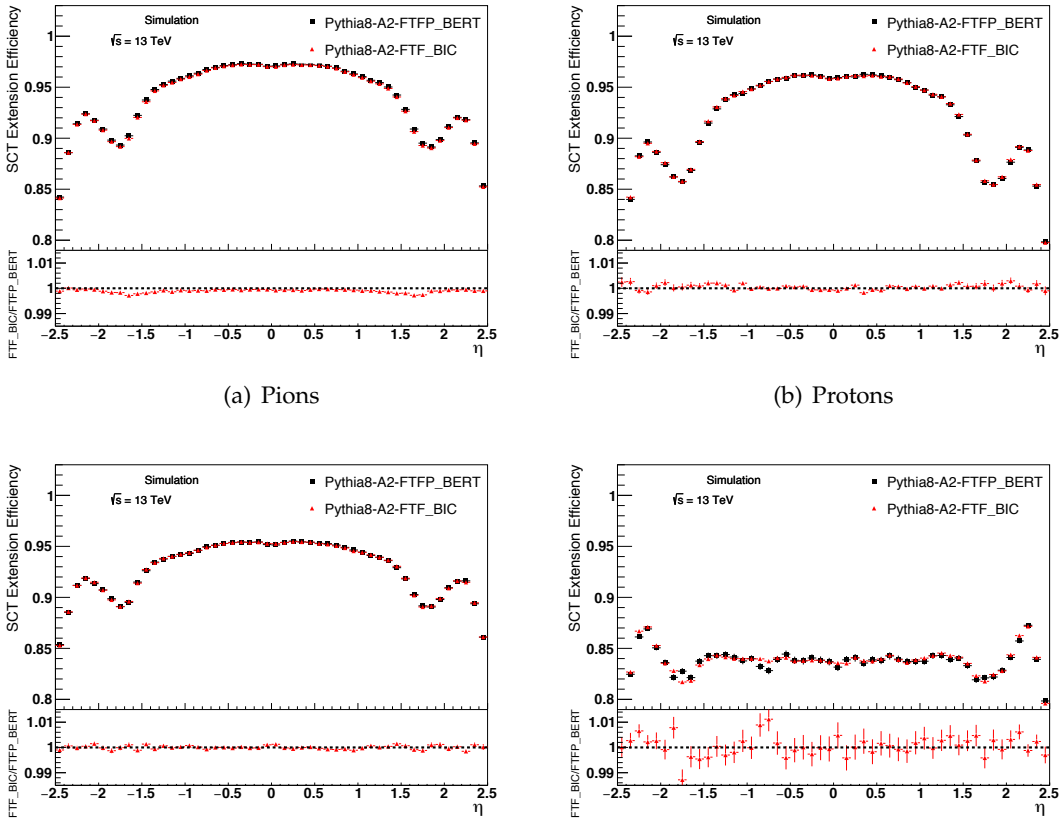
(a) All (b) Pions (c) Protons

(d) Kaons (e) Electrons (f) Others

FIGURE B.12: $\mathcal{E}_{\text{ext}}$ as a function of η divided by type (all, pions, protons, kaons, electrons, others) and by category (all, primary, secondary) for EPOS.

B.1.5 Physics List: FTFP_BERT and FTF_BIC

Figures B.13 and B.14 show the SCT-extension efficiency as a function of pseudorapidity, η, and transverse momentum, p_T, in a comparison between Pythia-8 (A2) based on FTFP_BERT and Pythia-8 (A2) based on FIT_BIC for different particles: pions, protons, kaons and others, where "others" are dominated by strange baryons. A comparison between the two Monte Carlo samples with different physics lists and data is also shown on Figure 4.31. No big differences ($> 0.1\%$) are visible between the two physics lists.



(a) Pions (b) Protons

(c) Kaons (d) Others

FIGURE B.13: SCT-extension efficiency as a function of pseudorapidity η in a comparison between Pythia-8 (A2) based on FTFP_BERT and Pythia-8 (A2) based on FIT_BIC for (a) pions, (b) protons, (c) kaons and (d) others.

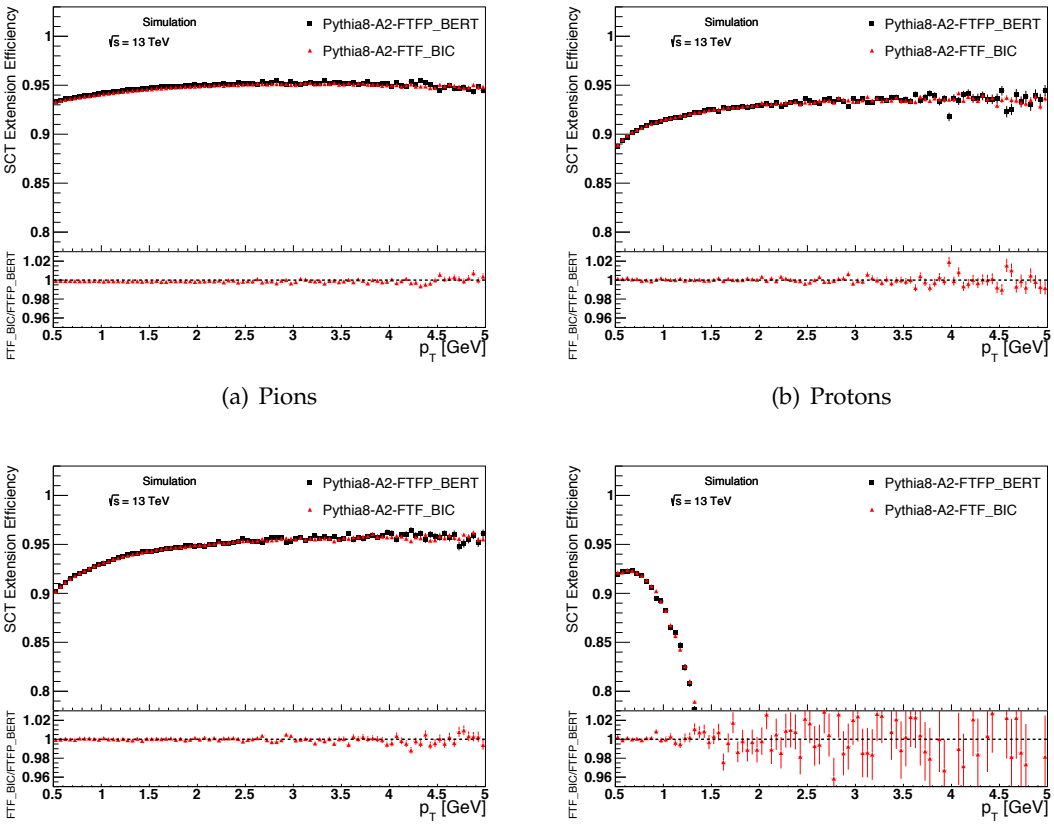
B.2 With the final event selection

B.2.1 Track-extension Efficiency as a function of ϕ

For completeness, figures B.15 and B.16 show the track-extension efficiency respectively as a function of ϕ and in 2D maps as a function of η and ϕ.

B.2.2 Track-extension Efficiency for particle species

Figure B.17 shows the track-Extension efficiency exclusively for different particle species.



(a) Pions (b) Protons

(c) Kaons (d) Others

FIGURE B.14: SCT-extension efficiency as function of transverse momentum p_T in a comparison between Pythia-8 (A2) based on `FTFP_BERT` and Pythia-8 (A2) based on `FIT_BIC` for (a) pions, (b) protons, (c) kaons and (d) others.

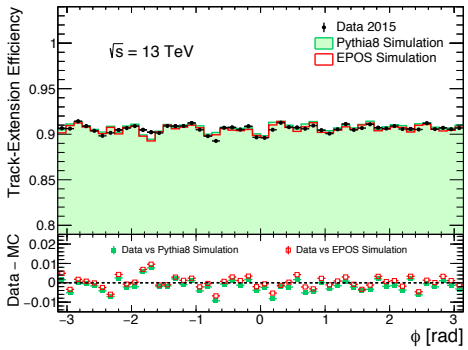
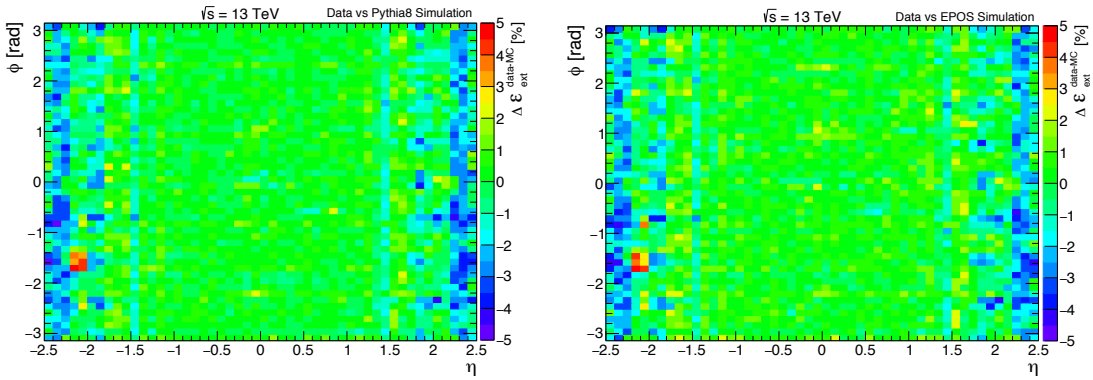
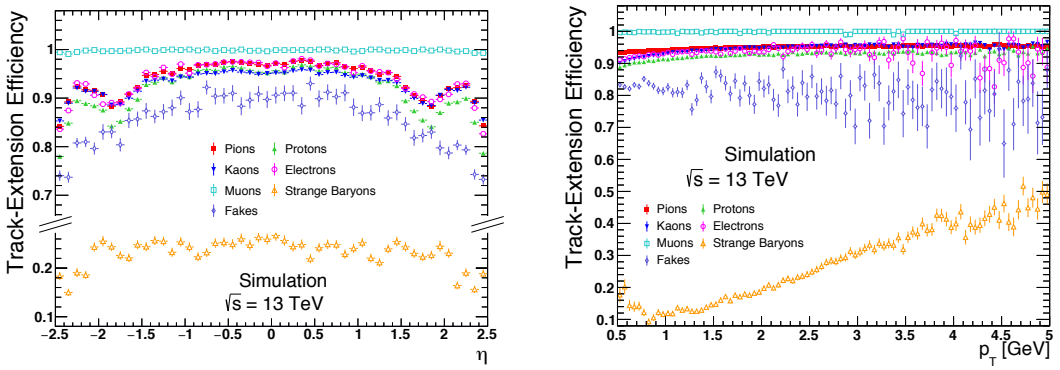


FIGURE B.15: Track-extension efficiency as a function of ϕ of the tracklets in a comparison between data, PYTHIA 8 and EPOS.



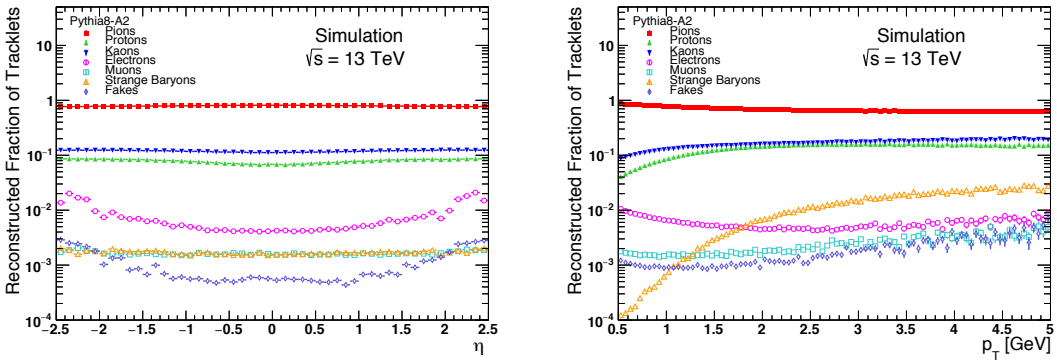
(a) Pythia 8 - A2 (b) Epos LHC

FIGURE B.16: Track-extension efficiency as a function of η and ϕ of the tracklets in a comparison between (a) data and PYTHIA 8 and between (b) data and EPOS.



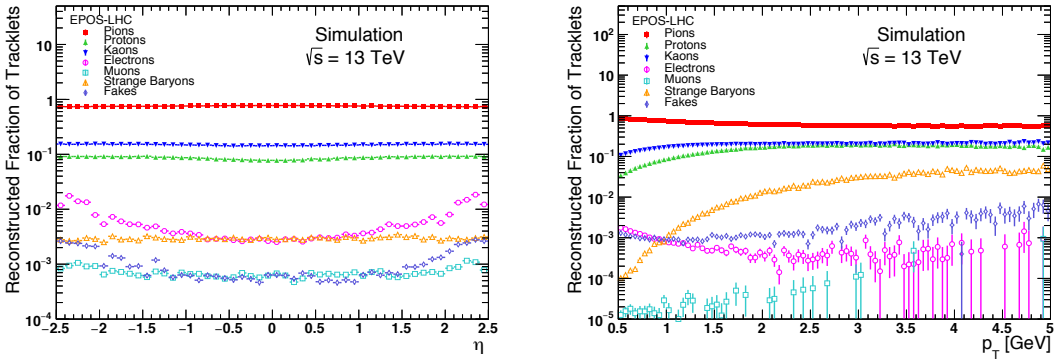
(a) η (b) p_{T}

FIGURE B.17: Track-Extension efficiency exclusive for pions, protons, kaons, electrons, muons, weakly-decaying strange baryons ($\Sigma^{\pm}$, Ξ^{-}, Ω^{-} and their antiparticles) or fake tracklets as a function of (a) η and (b) p_{T} in simulation.



(a) η (b) p_{T}

FIGURE B.18: Fraction of reconstructed tracklets associated with pions, protons, kaons, electrons, muons, weakly-decaying strange baryons ($\Sigma^{\pm}$, Ξ^{-} and Ω^{-}) or fake tracklets with respect to the total number of reconstructed tracklets as a function of (a) η and (b) p_{T} in the PYTHIA 8 simulated sample.



(a) η (b) p_{T}

FIGURE B.19: Fraction of reconstructed tracklets associated with pions, protons, kaons, electrons, muons, weakly-decaying strange baryons ($\Sigma^{\pm}$, Ξ^{-} and Ω^{-}) or fake tracklets with respect to the total number of reconstructed tracklets as a function of (a) η and (b) p_{T} in the EPOS simulated sample.

B.2.3 Fraction of tracklets for particle species

Figures B.18 and B.19 show the fraction of tracklets exclusively for different particle species.

Appendix C

New Small Wheel Simulation

C.1 Introduction

The ATLAS Experiment is embarking on a continuous upgrade program aimed at improving the detector performance to cope with the high luminosity conditions, which will characterise by the upgraded LHC [220] during Run 3 and beyond (2020), as briefly summarised in Chapter 2.

The main focus of the ATLAS Phase-I upgrade [221] to be undertaken in 2018 is on the Level-1 trigger. The objective is to sharpen the trigger threshold turn-on as well as discriminate against background while maintaining the low transverse momentum threshold for electrons and muons and keeping the Level-1 rate at a reasonable level level to be handled by the Experiments. Upgrades are planned for both the muon and the calorimeter trigger systems, without which the single lepton Level-1 triggers would either have to be prescaled or use a higher p_{T} threshold, which would result in a significant loss of acceptance for many physics processes to be explored.

The Phase-II upgrade [222] of ATLAS includes a full replacement of the central tracking system as well as major upgrades of the trigger and readout systems.

The recent discovery of the Higgs boson has started detailed studies of the nature of this new particle. Continuing these studies with higher statistics and higher energy will be a major topic for the forthcoming upgraded LHC physics programme together with searches for new physics beyond the Standard Model.

While high luminosity will provide more data, it is essential that the ATLAS detector is still able to operate in the higher background environment and maintaining as good a performance as at lower luminosities.

C.2 Phase-I Upgrade of the ATLAS Muon System

One of the major Phase-I upgrades of the ATLAS Detector consists in the replacement of the first station of the ATLAS muon end-cap system, the so-called Small Wheels, during the second Long Shutdown starting in 2018.

Figure C.2 shows a cross section of the ATLAS detector in the z-y plane. The barrel system covers the pseudorapidity region of $|\eta| < 1.0$ whereas the end-cap system covers the range $1.0 < |\eta| < 2.7$ for the muon tracking and $1.0 < |\eta| < 2.4$ for Level-1 trigger. The barrel and end-cap systems consist of three stations each, measuring the muon momentum based on the curvature in the ATLAS toroidal field.

There are two main issues, which represent a serious limitation on the ATLAS performance beyond design luminosity: cherge build up in the MDTs due to high rate, which will seriously deteriorate muon tracking in the end-cap muon spectrometer and an unacceptable rate of fake high p_{T} Level-1 muon triggers in the forward direction.

In order to fix both of these problems, the Small Wheels will be replaced by the New Small Wheels (NSW) [223], which will have to operate in a high rate region while reconstructing muon tracks with high precision as well as furnishing information for the Level-1

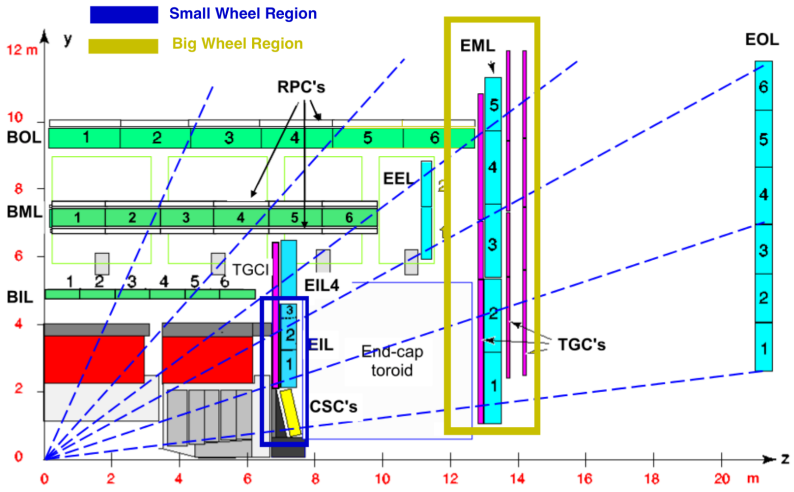


FIGURE C.1: A z-y view of 1/4 of the ATLAS detector. The blue boxes indicate the end-cap Monitored Drift Tube chambers (MDT) and the yellow box in the Small Wheel area the Cathode Strip Chambers (CSC). The green boxes are barrel MDT chambers. The trigger chambers, Resistive Plate chambers (RPC) and Thin Gap Chambers (TGC), are indicated by the outlined white and the magenta boxes. This is a cut-out on the muon spectrometer at the large sectors, hence the names End-cap Inner Large (EIL), End-cap Middle Large (EML) and End-cap Outer Large (EOL). The detector regions of the Small Wheel and Big Wheel are also outlined.

trigger. The NSW is, in fact, a set of precision tracking and trigger detectors able to work at high rates with excellent real-time spatial and time resolution. These detectors can provide the muon Level-1 trigger system with online track segments with a good angular resolution to confirm that muon tracks originate from the IP. In this way, the end-cap fake trigger rate will be considerably reduced. With the proposed NSW, the ATLAS muon system will maintain full muon tracking capabilities at the highest LHC luminosities. At the same time the Level-1 low p_{T} (typically $p_{\mathrm{T}} > 20$ GeV) single muon trigger rate will be kept at an acceptable level.

The η coverage of the proposed NSW (and the existing Small Wheel) is $1.3 < |\eta| < 2.7$.

More details about the NSW features and on the physics impact of this upgrade can be found in Ref. [223], while in the next sections a brief description of the NSW simulation in the general context of the ATLAS software will be given.

C.3 New Small Wheel Layout

The NSW consists of 16 detector planes in two multilayers. Each multilayer comprises four small Thin Gap Chambers (sTGC) and four Micro Mesh Gas Structures (MicroMeGaS, MM) detector planes. The sTGC are primarily deployed for triggering given their single bunch crossing identification capability. The detectors are arranged in such a way (sTGC - MM - MM - sTGC) as to maximize the distance between the sTGCs of the two multilayers. The naming convention adopted for the NSW is shown in Figure C.2. This configuration has been optimised for the online track resolution: increased distance between detector multilayers leads to an improved online resolution of the track segment angle. The MM detectors have exceptional precision tracking capabilities due to their small gap (5 mm)

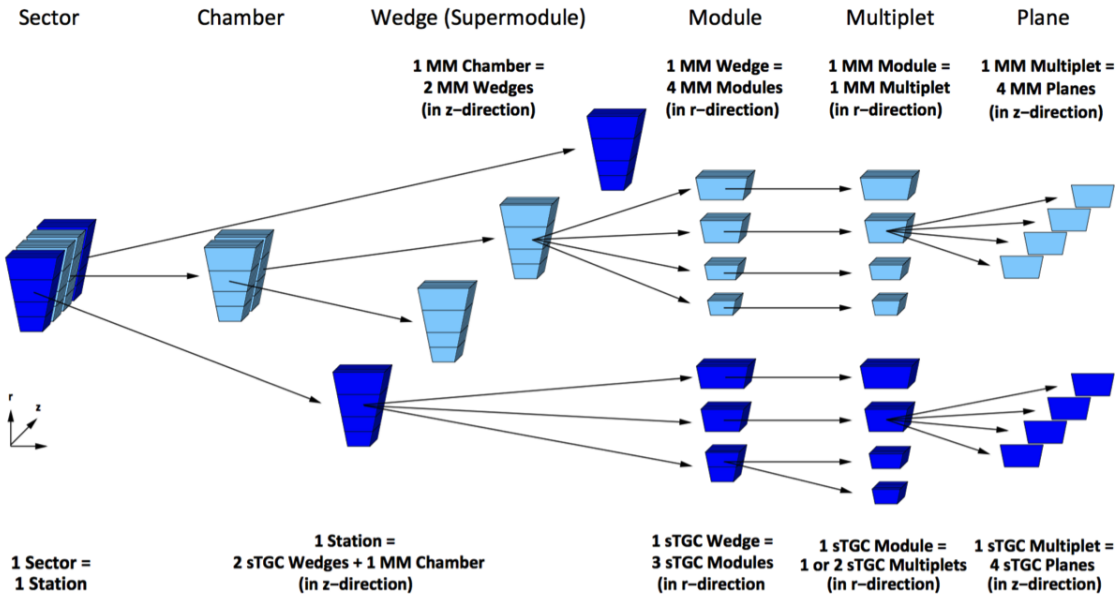


FIGURE C.2: New Small Wheel naming convention.

and strip pitch (0.5 mm). The choice of eight planes per detector was dictated by the need to provide a redundant, fully functional detector system over its whole lifetime. The main issues that have been addressed are operation in a high background environment and detector deterioration with time and its influence on the track segment reconstruction efficiency and resolution. In addition, the large number of planes will ensure an appropriate detector performance even if some planes fail to work properly. Access and repair opportunities will be rare, if at all possible. Furthermore, it may be required for long-term stability reasons to operate the detector planes with lower high voltage settings than it is optimal for full efficiency. Eight planes per detector will once more ensure an overall efficiency close to 100% both for online and offline track reconstruction. The two NSW detector technologies also complement each other in their corresponding primary functions: the sTGC may contribute to offline precision tracking, as they are able to measure track hits with a resolution better than 150 μm (depending on the incident track angle). For triggering, experience has shown that redundancy is highly important in the forward direction at high luminosities. The MM detectors will be utilised as an additional trigger system to provide improved redundancy, robustness and coverage of the forward trigger.

C.4 Simulation of the New Small Wheel

Accurate simulation of the New Small Wheel system needs to be implemented in the ATLAS Simulation programmes described in Chapter 2 to evaluate the improved performance and to optimize the layout. The new simulation codes must be usable within the existing muon system description and not perturbe ongoing simulation activities which support data taking. The description of the muon chambers is currently hard-coded in C++ with the chamber dimensions and other geometry-related infomation taken from the ATLAS Muon database (AMDB). The new chamber types (sTGC, MM) are not supported in AMDB, thus the choice was to move the NSW description to an XML format (parsed into geometry primitives at run time), which is totally independent from the code used for the chamber description. The XML-based approach allowed to begin with a very simplified

geometry description, which can be detailed in iterative stages any time new information becomes available, whitout breaking the simulation chain. At the same time, automatic XML generation is possible, starting e.g. from an Excel-based parameter book. Human intervention is thus reduced to a minimum making the full procedure less error prone. Description of chamber details (like frames, spacer) in C++ hides implementation details and allows for simpler detector description. This approach is much more convenient than having the parameters into a database and the NSW experience will be also used to design a complete replacement of the muon system simulation suite.

In addition to the geometry layout, also the detector response needs to be simulated and this is done in a dedicated *digitization* step. Strip pitch, wire etc. are introduced at digitization level and are not simulated by GEANT 4. Other experimental effects (cluster size, Lorentz angle, drift velocity etc.) are parameterized as a function of test-beam results or dedicated simulations (Garfield [224], etc). A *pseudo-digitization* package was developed which allows to by-pass accurate digitization and speed up the reconstruction time while still providing strip and wire response. For what concerns the background, the main sources in the muon chambers are:

- particles from showers from primary events (in-time pile-up);
- particles from out-of-time interactions in the detector or beam-line elements (cavern background)

The background events are simulated separately and overlayed on top of physics events at digitization level, according to the chosen instantaneous luminosity levels. Standalone background generators were also developed for the NSW TDR studies.

The existing reconstruction chain was also modified in order to include MM and sTGC hits in pattern recognition and track reconstruction. In parallel, studies on the implementation of new trigger algorithms,which utilize New Small Wheel trigger information, are ongoing with the aim of including it in the global muon trigger logic.

C.4.1 Geometry Implementation

The implementation of latest layout, shown in Figure C.3 according to technical drawings and the sTGC parameters book required partial re-implementation of both sTGC and Micromegas geometry primitives. sTGC geometry was driven by automatic XML generation while at the time of this implementation (August 2014) the Micromegas dimensions and positions were still provisional and/or driven by sTGC. The implementation of new positioner elements made description more rational and compact. The new description was validated, to check for overlaps and possible inconsistencies. At the first implementation, the new naming conventions and XML organization broke completely the whole NSW software chain.

sTGC layout

The main geometrical parameters which charatcterise the layout of the sTGCs are included into an Excel file, referred to as *Parameter book*. This parameter book can be automatically converted into an XML file to be read in by the simulation code. Digitization parameters such as strip pitch, pads scheme etc., are also in XML and a link between simulation and digitization was put in place, to ensure that the same parameters are used throughout the whole chain. Links to Readout Geometry were re-established after changes in shapes and naming conventions broke the chain implemented for the TDR. The most external sTGC chamber has an exagonal shape, as shown in Figure C.4, which was originally implemented in the NSW TDR geometry by means of complicated boolean subtractions. Unfortunately, boolean operations are not suitable for frame implementation and they constitute

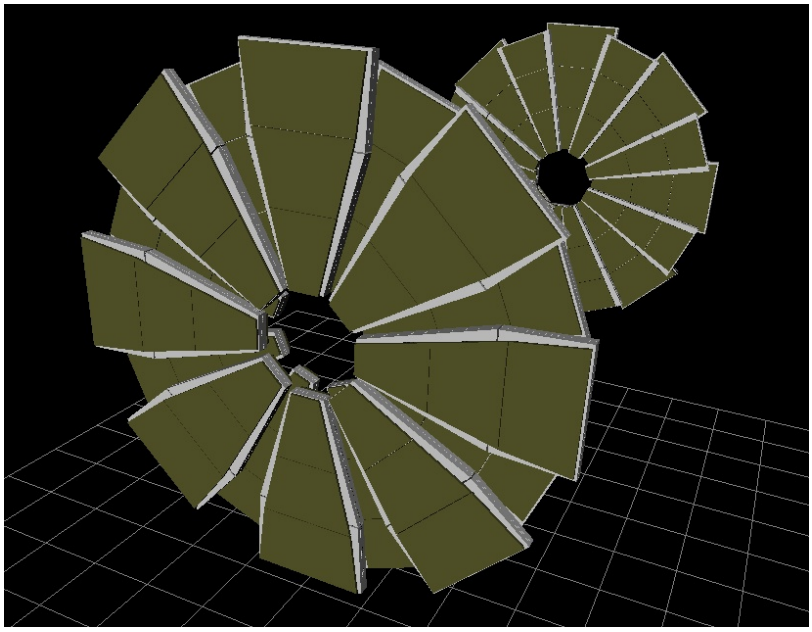


FIGURE C.3: The New Small Wheels: in green the sTGCs and in grey the Micromegas chambers. Image generated with VP1 [225].

a source of endless problems at tracking time. Thus, these particular sTGC chambers were reimplemented by means of Boundary REPresentation solids [103], meaning that the sTGC are not trapezes any more and this caused the Readout Elements and geometry to break, which required a prompt fix in order to not stop the simulation development. In addition, a frame which surrounds each chamber and which was missing at the time of the TDR, was introduced.

MM layout

The MM chambers of the NSW can be seen in Figure C.5. As in the case of the sTGCs, also for MM chambers, the main geometrical parameters are collected in the parameter book. The chamber frame was introduced, as it can be seen in Figure C.6.

C.4.2 Validation procedures

A series of tests and validation suites is currently being developed to check code integrity and to monitor the correct program functioning when conditions (e.g. layout) change. For instance, when the level of layout details increases it is important to check that everything is under control, e.g. frames are visible, gaps are correctly reproduced, etc. Effects from frame material and chamber overlaps can be clearly seen in chamber *hit radiographies*, as shown in Figure C.7. A list of histograms to be filled automatically for every nigtly Ahena release was implemented in the general framework of the ATLAS Run Time Tester (RTT) [226].

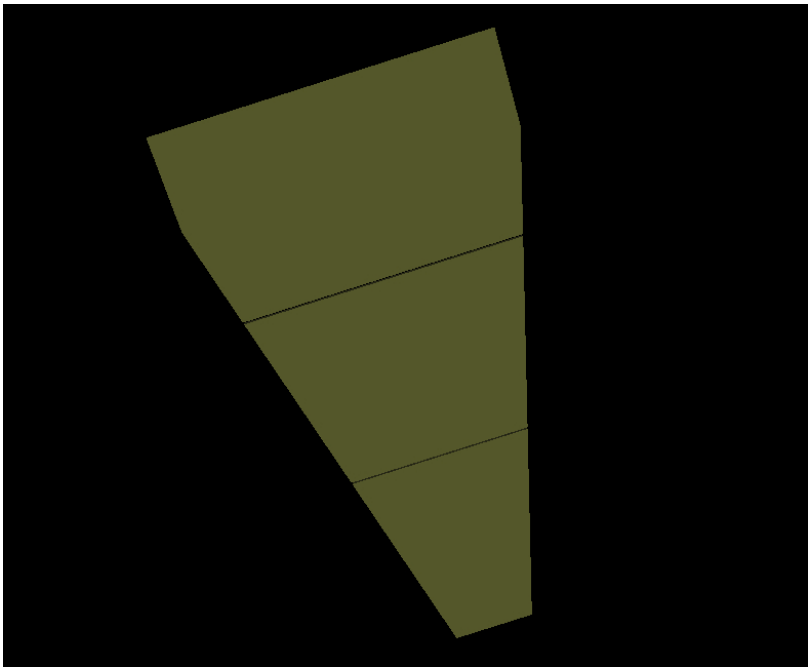


FIGURE C.4: An sTCC sector in the New Small Wheels. Image generated with VP1 [225].

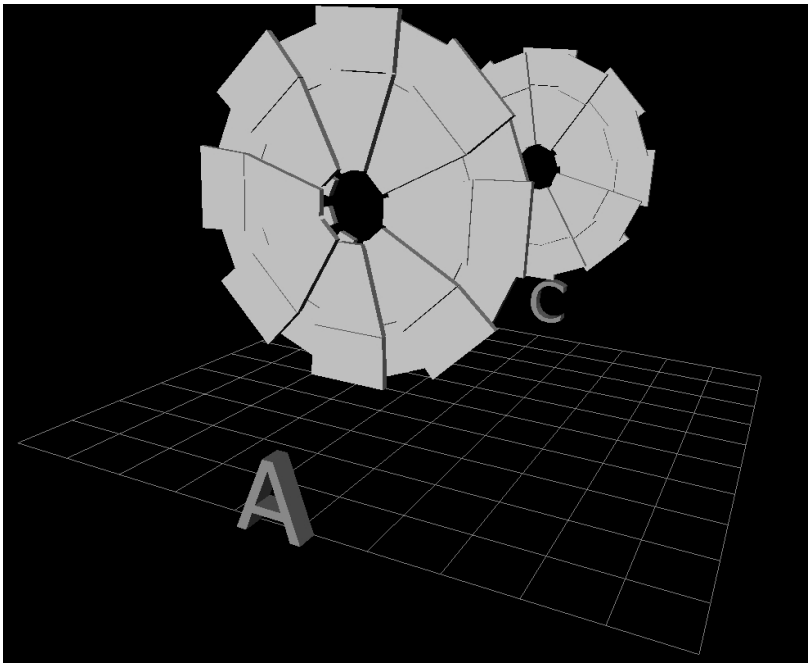


FIGURE C.5: The Micromegas detectors in the New Small Wheels. Image generated with VP1 [225].

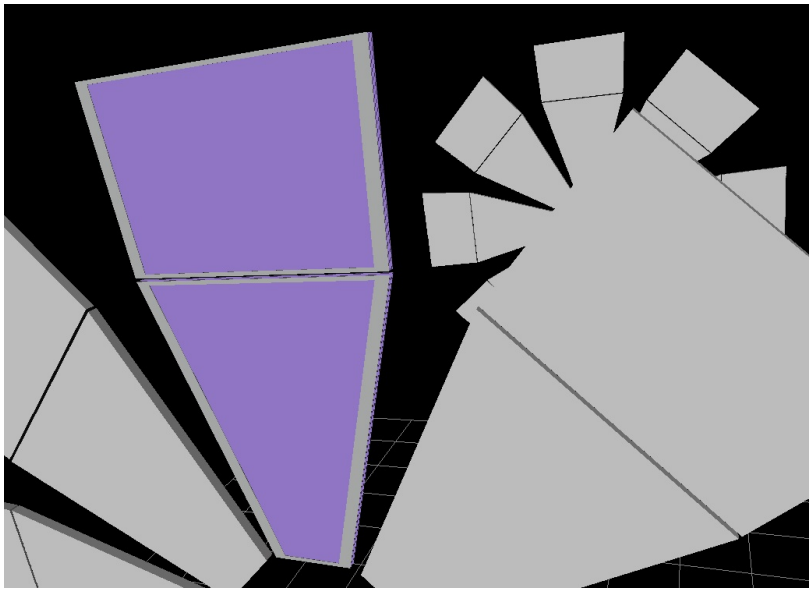
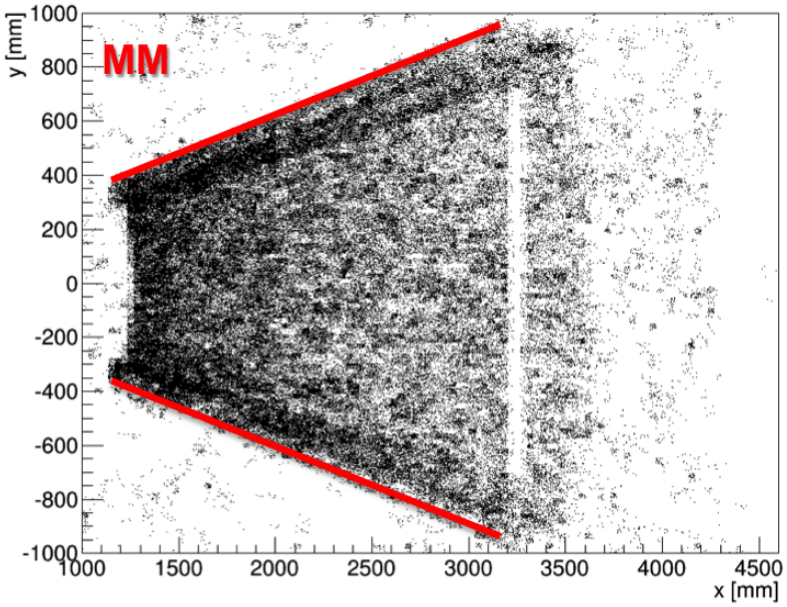
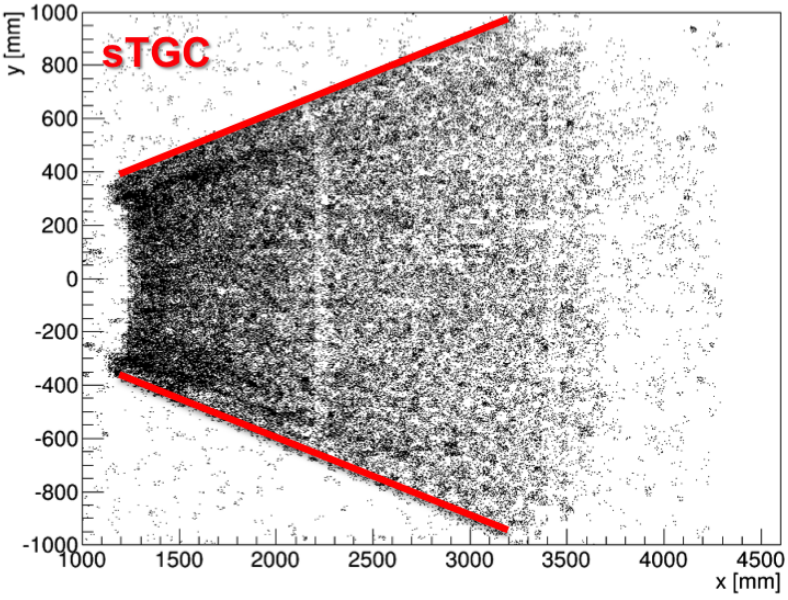


FIGURE C.6: The frames which surround the Micromegas detectors in the New Small Wheels. Image generated with VP1 [225].



(a) MM



(b) sTGC

FIGURE C.7: Hit radiographies of a sector of (a) Micromegas and (b) sTGC chambers. The red lines represent the chamber edges.

Bibliography

[1] A. Salam. "Elementary Particle Theory". In: (1969), p. 367.

[2] S. Weinberg. "A Model of Leptons". In: *Phys. Rev. Lett.* 19 (1967), pp. 1264–1266.

[3] S. L. Glashow. "A Model of Leptons". In: *Nucl. Phys.* 22 (1961), pp. 579–588.

[4] W. Pauli. "Über den Zusammenhang des Abschlusses der Elektronengruppen im Atom mit der Komplexstruktur der Spektren". In: *Zeitschrift für Physik* 31, Issue 1 (1925), pp. 765–783.

[5] F. Englert and R. Brout. "Broken Symmetry and the Mass of Gauge Vector Mesons". In: *Physical Review Letters* 13 (1964), p. 321.

[6] P. Higgs. "Broken Symmetries, Massless Particles and Gauge Fields". In: *Physicas Letters* 12 (1964), p. 132.

[7] P. Higgs. "Broken Symmetries and the Masses of Gauge Bosons". In: *Physical Review Letters* 13 (1964), p. 508.

[8] P. Higgs. "Spontaneous Symmetry Breakdown without Massless Bosons". In: *Physical Review* 145 (1966), p. 1156.

[9] ATLAS Collaboration. "Observation of a New Particle in the Search for the Standard Model Higgs Boson with the ATLAS detector at the LHC". In: *Physics Letters B* 716 (2012), pp. 1–29.

[10] CMS Collaboration. "Observation of a new boson at a mass of 125 GeV with the CMS experiment detector at the LHC". In: *Physics Letters B* 716 (2012), pp. 30–61.

[11] ATLAS and CMS Collaborations. "Combined Measurement of a the Higgs boson mass at $\sqrt{s}$=7 and 8 TeV with the ATLAS and CMS experiment Experiments". In: *Phys. Rev. Lett.* 716 (2012), pp. 30–61.

[12] ATLAS and CMS Collaborations. "Measurements of the Higgs boson production and decay rates and constraints on its coupling from a combined ATLAS and CMS analysis at the LHC pp collision data at $\sqrt{s}$=7 and 8 TeV". In: *arXiv:1606.02266 [hep-ex]* (2016), p. 70.

[13] F. J. Hasert et al. "Observation of neutrino-like interactions without muon or electron in the gargamelle neutrino experiment". In: *Physics Letters B* 43 (1973), pp. 138–145. DOI: https://doi.org/10.1016/0370-2693(73)90499-1.

[14] *UA1*. URL: https://home.cern/about/experiments/ua1.

[15] *UA2*. URL: https://home.cern/about/experiments/ua2.

[16] G. et al.) UA1 Collaboration (Arnison. "Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at $\sqrt{s} = 540\ GeV$". In: *Phys.Lett.* 122B (1983), pp. 103–116.

[17] M. et al.) UA2 Collaboration(Banner. "Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti-p p Collider". In: *Phys.Lett.* 122B (1983), pp. 476–485.

[18] G. et al.) UA1 Collaboration (Arnison. "Experimental Observation of Lepton Pairs of Invariant Mass Around $95\ GeV/c^2$ at the CERN SPS Collider". In: *Phys.Lett.* 126B (1983), pp. 398–410.

[19] P. et al.) UA2 Collaboration(Bagnaia. "Evidence for Z0 $\rightarrow e^+e^-$ at the CERN anti-p p Collider". In: *Phys.Lett.* 129B (1983), pp. 130–140.

[20] H. Leutwyler H. Fritzsch M. Gell-Mann. "Advantages of the Color Octet Gluon Picture". In: *Phys. Lett. B* 47 (1973), p. 365.

[21] M. E. Peskin and D. V. Schroeder. "An Introduction to quantum field theory". In: *Westview Press* (1995).

[22] K.A. Olive et al. "Particle Data Group". In: *Chin. Phys. C* 38, 090001 (2014 and 2015 update).

[23] Davison E. Soper John C. Collins and George Sterman. "Factorization of Hard Processes in QCD". In: *arXiv:hep-ph/0409313* (2004).

[24] Andy Buckley et al. "General-purpose event generators for LHC physics". In: arXiv:1101.2599v1 [hep-ph] (2011).

[25] S. Mandelstam. "Determination of the Pion-Nucleon Scattering Amplitude from Dispersion Relations and Unitarity". In: *Physical Review* (1958). DOI: 10.1103/PhysRev.112.1344.

[26] W.J. Stirling. *Private communication*. URL: http://www.hep.ph.ic.ac.uk/~wstirlin/plots/plots.html.

[27] A. Sherstnev and R.S. Thorne. "Parton Distributions for LO Generators". In: *Eur.Phys.J.* C55 (2008), pp. 553–575.

[28] I. Pomeranchuck. "Stability of a Fermi Liquid". In: *Sov. Phys. JETP* 35 (1958).

[29] V. N. Gribov. "A Reggeon Diagram Technique". In: *Sov. Phys. JETP* 26:414-422 (1968).

[30] V. A. Abramovsky, V. N. Gribov, and O. V. Kancheli. "Character of Inclusive Spectra and Fluctuations produced in Inelastic Processes by Multi-Pomeron Exchange". In: *Yad. Fiz.* 18:595-616 (1973).

[31] G. Veneziano. "Regge Intercepts and Unitarity in Planar Dual Models". In: *Nucl. Phys.* 1B74:365 (1974).

[32] J. Bartels, T. Gehrmann, and M. G. Ryskin. "Spin dependence of Deep Inelastic Diffraction: An Outline". In: *Prepared for 2nd Topical Workshop on Deep Inelastic Scattering off Polarized, Targets: Theory Meets Experiment (SPIN 97)* Zeuthen, Germany (1-5 Sep 1997).

[33] A. Donnachie and P. V. Landshoff. "Total Cross Sections". In: *Phys. Lett.* B296:227-232 (1992).

[34] Sherpa Collaboration) (F. Krauss. *Sketch of a $t\bar{t}h$ event*. URL: http://www.opensciencegrid.org/wp-content/uploads/2014/05/event.jpg.

[35] Andy Buckley et al. "Rivet user manual". In: arXiv:1003.0694 (2013), p. 50.

[36] Andy Buckley et al. "Systematic event generator tuning for the LHC". In: *Eur. Phys. J. C65* arXiv:0907.2973 (2010), pp. 331–357. DOI: 10.1140/epjc/s10052-009-1196-7.

[37] S. Mrenna T.Sjöstrand and P. Skands. "Brief Introduction to PYTHIA 8.1". In: *Comput. Phys. Comm* 178 (2008). eprint: arXiv:0710.3820.

[38] S. Porteboeuf, T. Pierog, and K. Werner. "Producing Hard Processes Regarding the Complete Event: The EPOS Event Generator". In: (2010). arXiv: 1006.2967 [hep-ph].

[39] M. Bahr et al. "Herwig++ Physics and Manual". In: *Eur.Phys.J.* C58 (2008), pp. 639–707.

[40] S. Ostapchenko. "Monte Carlo treatment of hadronic interactions in enhanced Pomeron scheme: QGSJET-II model". In: *Phys. Rev. D* 83 (1 2011), p. 014018. DOI: 10.1103/PhysRevD.83.014018. URL: http://link.aps.org/doi/10.1103/PhysRevD.83.014018.

[41] The ATLAS Collaboration. "Further ATLAS tunes of PYTHIA6 and Pythia 8".

[42] ATLAS Collaboration. "Measurements of the pseudorapidity dependence of the total transverse energy in proton-proton collisions at $\sqrt{s} = 7$ TeV with ATLAS". In: *JHEP* 1211 (2012), p. 033. DOI: 10.1007/JHEP11(2012)033. arXiv: 1208.6256 [hep-ex].

[43] Peter Skands, Stefano Carrazza, and Juan Rojo. "Tuning PYTHIA 8.1: the Monash 2013 Tune". In: *Eur.Phys.J.* C74.8 (2014), p. 3024.

[44] Richard D. Ball et al. "Parton distributions with LHC data". In: *Nucl.Phys.* B867 (2013), pp. 244–289.

[45] Stefan Gieseke, Christian Rohr, and Andrzej Siodmok. "Colour reconnections in Herwig++". In: *Eur.Phys.J.* C72 (2012), p. 2225. DOI: 10.1140/epjc/s10052-012-2225-5.

[46] J. Pumplin et al. "New generation of parton distributions with uncertainties from global QCD analysis". In: *JHEP* 0207 (2002), p. 012.

[47] J. Pumplin et al. "New generation of parton distributions with uncertainties from global QCD analysis". In: *JHEP* 07 (2002), p. 012. DOI: 10.1088/1126-6708/2002/07/012. arXiv: hep-ph/0201195 [hep-ph].

[48] T. Pierog, Iu. Karpenko, J. M. Katzy, E. Yatsenko, and K. Werner. "EPOS LHC: Test of collective hadronization with data measured at the CERN Large Hadron Collider". In: *Phys. Rev. C* 92 (2015). DOI: 10.1103/PhysRevC.92.034906.

[49] H. J. Drescher et al. "Parton-based Gribov-Regge theory". In: *Phys. Rept.* (). URL: http://dx.doi.org/10.1016/S0370-1573(00)00122-8.

[50] Henry D.I. Abarbanel. "Reggeon field theory: Formulation and use". In: *Phys. Rep.* 21 (1975), pp. 19–182. DOI: 10.1016/0370-1573(75)90034-4.

[51] Bruning, Oliver S. et al. "LHC Design Report Vol.1: The LHC Main Ring". In: CERN-2004-003-V1, CERN-2004-003, CERN-2004-003-V-1 (2004).

[52] ATLAS Collaboration. "The ATLAS Experiment at the CERN Large Hadron Collider". In: *JINST* 3 (2008), S08003. DOI: 10.1088/1748-0221/3/08/S08003.

[53] CMS Collaboration. "The CMS experiment at the CERN LHC". In: *JINST* 3 (2008), S08004. DOI: 10.1088/1748-0221/3/08/S08004.

[54] LHCb Collaboration. "The LHCb Detector at the LHC". In: *JINST* 3 (2008), S08005. DOI: 10.1088/1748-0221/3/08/S08005.

[55] ALICE Collaboration. "The ALICE Experiment at the CERN LHC". In: *JINST* 3 (2008), S08002. DOI: 10.1088/1748-0221/3/08/S08002.

[56] Mike Lamont. "The LHC's first long run". In: *CERN Courier* (August 19, 2013). URL: http://cerncourier.com/cws/article/cern/54381.

[57] "Summary of the Analysis of the 19 September 2008 Incident at the LHC, Technical report". In: *CERN* (Oct 2008).

[58] ATLAS Collaboration. *Event Displays from Run 2 physics analyses*. URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/EventDisplayRun2Physics.

[59] *ATLAS Luminosity Public Results*. URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2.

[60] ATLAS Collaboration. "Measurement of charged-particle distributions sensitive to the underlying event in $\sqrt{s} = 13$ TeV proton-proton collisions with the ATLAS detector at the LHC". In: *JHEP* 03 (2017), p. 157. DOI: 10.1007/JHEP03(2017)157. arXiv: 1701.05390 [hep-ex].

[61] ATLAS Collaboration. "ATLAS magnet system: Technical design report". In: (1997).

[62] ATLAS Collaboration. "ATLAS central solenoid: Technical design report". In: (1997).

[63] ATLAS Collaboration. *ATLAS barrel toroid: Technical Design Report*. 1997.

[64] ATLAS Collaboration. "ATLAS endcap toroids: Technical design report". In: (1997).

[65] ATLAS Collaboration. "ATLAS inner detector: Technical design report. Vol. 1". In: (1997).

[66] ATLAS Collaboration. "ATLAS inner detector: Technical design report. Vol. 2". In: (1997).

[67] M. Capeans et al. "ATLAS Insertable B-Layer Technical Design Report". In: CERN-LHCC-2010-013. ATLAS-TDR-19 (2010). CERN-LHCC-2010-013, ATLAS-TDR-19. URL: https://cds.cern.ch/record/1291633.

[68] ATLAS Collaboration. *ATLAS pixel detector: Technical Design Report*. Technical Design Report ATLAS. Geneva: CERN, 1998. URL: https://cds.cern.ch/record/381263.

[69] *ATLAS IBL Photos*. URL: https://cds.cern.ch/record/1702006.

[70] H. Oide, H. Pernegger, M. Garcia Sciveres, F. Cerutti, J. Dopke, D. Dobos, B. Di Girolamoa, D. Ferrere, M.A. Gallileee. "Study of material budget of the ATLAS Run-2 beam pipe and IBL components". In: ATL-COM-INDET-2014-064 (). URL: https://cds.cern.ch/record/1757462.

[71] S. Welch and J. Dopke. "The ATLAS Pixel nSQP Readout Chain". In: *ATL-INDET-PROC-2012-018* (2012).

[72] Jens Dopke. "Private communication".

[73] ATLAS Collaboration. "ATLAS calorimeter performance Technical Design Report". In: (1996).

[74] ATLAS Collaboration. "ATLAS liquid argon calorimeter: Technical design report". In: (1996).

[75] "ATLAS tile calorimeter: Technical design report". In: (1996).

[76] ATLAS Collaboration. "ATLAS muon spectrometer: Technical design report". In: (1997).

[77] ATLAS Collaboration. "Performance of the ATLAS Trigger System in 2015". In: *Eur. Phys. J.* C77.5 (2017), p. 317. DOI: 10.1140/epjc/s10052-017-4852-3. arXiv: 1611.09661 [hep-ex].

[78] ATLAS Collaboration. "Performance of the ATLAS Trigger System in 2010". In: *Eur. Phys. J.* C72 (2012), p. 1849. DOI: 10.1140/epjc/s10052-011-1849-1. arXiv: 1110.1530 [hep-ex].

[79] The ATLAS TDAQ Collaboration. "The ATLAS Data Acquisition and High Level Trigger system". In: *Journal of Instrumentation* 11.06 (2016), P06008. URL: http://stacks.iop.org/1748-0221/11/i=06/a=P06008.

[80] ATLAS Collaboration. "Charged-particle distributions in $\sqrt{s} = 13$ TeV pp interactions measured with the ATLAS detector at the LHC". In: *Phys. Lett. B* 758 (2016), p. 67. DOI: 10.1016/j.physletb.2016.04.050. arXiv: 1602.01633 [hep-ex].

[81] ATLAS Collaboration. "Charged-particle distributions at low transverse momentum in $\sqrt{s} = 13$ TeV *pp* interactions measured with the ATLAS detector at the LHC". In: *Eur. Phys. J. C* 502 (2016), p. 76. DOI: 10.1140/epjc/s10052-016-4335-y. arXiv: 1606.01133 [hep-ex].

[82] ATLAS Collaboration. "Measurement of the Inelastic Proton-Proton Cross Section at $\sqrt{s} = 13$ TeV with the ATLAS Detector at the LHC". In: *Phys. Rev. Lett.* 117 (18 2016), p. 182002. DOI: 10.1103/PhysRevLett.117.182002. URL: https://link.aps.org/doi/10.1103/PhysRevLett.117.182002.

[83] ATLAS Collaboration. "ATLAS Computing Technical Design Report". In: CERN-LHCC-2005-022. ATLAS-TDR-017 (2005). CERN-LHCC-2013-022, ATLAS-TDR-017. URL: https://cds.cern.ch/record/837738.

[84] Ian Bird. "Computing for the Large Hadron Collider". In: *Annual Review of Nuclear and Particle Science* 61 (2011), pp. 99–118. URL: http://www.annualreviews.org/doi/abs/10.1146/annurev-nucl-102010-130059.

[85] P.F. Åkesson, E. Moyse. "Software releases management in the Trigger and Data Acquisition of the ATLAS Experiment". In: (). DOI: 10.5170/CERN-2005-002.255. URL: https://cds.cern.ch/record/865580/.

[86] Bjarne Stroustrup. *The C++ Programming Language*.

[87] *Root data analysis framework*. URL: https://root.cern.ch.

[88] J. Reid. "The Future of Fortran". In: *Computing in Science and Engineering* 5 (2003), pp. 59–67. DOI: 10.1109/MCISE.2003.1208645.

[89] James Gosling, Bill Joy, Guy Steele, Gilad Bracha, Alex Buckley. "The Java Language Specification". In: *Oracle America* (2015). URL: https://docs.oracle.com/javase/specs/jls/se8/jls8.pdf.

[90] *The Gaudi Project*. URL: http://gaudi.web.cern.ch/.

[91] A. Kazarov, M. Caprini, R. Hauser, I. Soloviev. "Software releases management in the Trigger and Data Acquisition of the ATLAS Experiment". In: ATL-DAQ-PROC-2011-014 (). URL: https://cds.cern.ch/record/1337261/.

[92] ATLAS Collaboration. "ATLAS Phase-II Upgrade Scoping Document". In: CERN-LHCC-2015-020, LHCC-G-166 (2015) (). URL: https://cds.cern.ch/record/2055248.

[93] ATLAS Collaboration. *Technical Design Report for the ATLAS Inner Tracker Strip Detector*. Tech. rep. CERN-LHCC-2017-005. ATLAS-TDR-025. Geneva: CERN, 2017. URL: https://cds.cern.ch/record/2257755.

[94] ATLAS Collaboration. "Expected Performance of the ATLAS Inner Tracker at the High-Luminosity LHC". In: ATL-PHYS-PUB-2016-025 (). URL: https://cds.cern.ch/record/2222304/.

[95] ATLAS Collaboration. "Impact Parameter Resolution". In: *IDTR-2015-007* (2015).

[96] T. Cornelissen et al. *Concepts, Design and implementation of the ATLAS New Tracking (NewT)*. ATL-INDET-PUB-2008-002. 2007. URL: https://cds.cern.ch/record/1020106.

[97] T. Cornelissen et al. *Single Track Performance of the Inner Detector New Track Reconstruction (NEWT)*. ATL-INDET-PUB-2008-002. 2008. URL: https://cds.cern.ch/record/1092934.

[98] Andreas Salzburger. "Optimisation of the ATLAS Track Reconstruction Software for Run-2". In: *Journal of Physics: Conference Series* 664.7 (2015), p. 072042. URL: http://stacks.iop.org/1742-6596/664/i=7/a=072042.

[99] Andreas Salzburger. "Doctoral Thesis: Track Simulation and Reconstruction in the ATLAS Experiment". In: (). URL: http://physik.uibk.ac.at/hephy/theses/diss_as.pdf.

[100] T. Cornelissen et al. "The new ATLAS track reconstruction (NEWT)". In: *J. Phys. Conf. Ser.* 119 (2008), p. 032014. DOI: 10.1088/1742-6596/119/3/032014. URL: http://stacks.iop.org/1742-6596/119/i=3/a=032014.

[101] R. E. Kalman. "A New Approach to Linear Filtering and Prediction Problems". In: *Journal of Basic Engineering 82:35* (1960).

[102] ATLAS Collaboration. *Track Reconstruction Performance of the ATLAS Inner Detector at $\sqrt{s} = 13$ TeV*. ATL-PHYS-PUB-2015-018. 2015. URL: http://cdsweb.cern.ch/record/2037683.

[103] S. Agostinelli et al. "GEANT4: A simulation toolkit". In: *Nucl. Instr. Meth.* A506 (2003), pp. 250–303.

[104] ATLAS Collaboration. "Reconstruction of primary vertices at the ATLAS experiment in Run 1 proton-proton collisions at the LHC". In: *Eur. Phys. J.* C77.5 (2017), p. 332. DOI: 10.1140/epjc/s10052-017-4887-5. arXiv: 1611.10235 [physics.ins-det].

[105] ATLAS Collaboration. *Alignment of the ATLAS Inner Detector with the initial LHC data at $\sqrt{s} = 13$ TeV*. ATL-PHYS-PUB-2015-031. 2015. URL: http://cdsweb.cern.ch/record/2038139.

[106] ATLAS Collaboration. "Tracking Studies for b-tagging with 7 TeV Collision Data with the ATLAS Detector". In: ATLAS-CONF-2010-070 (2010). URL: http://cdsweb.cern.ch/record/1281352.

[107] ATLAS Collaboration. *Early Inner Detector Tracking Performance in the 2015 Data at $\sqrt{s} = 13$ TeV*. ATL-PHYS-PUB-2015-051. 2015. URL: http://cdsweb.cern.ch/record/2110140.

[108] Paolo Nason. "A New method for combining NLO QCD with shower Monte Carlo algorithms". In: *JHEP* 11 (2004), p. 040. DOI: 10.1088/1126-6708/2004/11/040. arXiv: hep-ph/0409146 [hep-ph].

[109] Stefano Frixione, Paolo Nason, and Carlo Oleari. "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method". In: *JHEP* 11 (2007), p. 070. DOI: 10.1088/1126-6708/2007/11/070. arXiv: 0709.2092 [hep-ph].

[110] Simone Alioli et al. "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX". In: *JHEP* 06 (2010), p. 043. DOI: 10.1007/JHEP06(2010)043. arXiv: 1002.2581 [hep-ph].

[111] Simone Alioli et al. "NLO vector-boson production matched with shower in POWHEG". In: *JHEP* 0807 (2008), p. 060. DOI: 10.1088/1126-6708/2008/07/060. arXiv: 0805.4802 [hep-ph].

[112] ATLAS Collaboration. *Studies of the ATLAS Inner Detector material using $\sqrt{s} = 13$ TeV pp collision data*. ATL-PHYS-PUB-2015-050. 2015. URL: http://cdsweb.cern.ch/record/2109010.

[113] ATLAS Collaboration. "Study of the Material of the ATLAS Inner Detector for Run 2 of the LHC". In: (2017).

[114] ATLAS Collaboration. "Study of the mechanical stability of the ATLAS Insertable B-Layer". In: *CERN* (). URL: https://cds.cern.ch/record/2022587/.

[115] B. Andersson, G. Gustafson and B. Nilsson-Almqvist. "A model for low-pT hadronic reactions with generalizations to hadron-nucleus and nucleus-nucleus collisions". In: *Nucl. Phys., B* 281 (1987).

[116] Nilsson-Almqvist, Bo and Stenlund, Evert. "Interactions Between Hadrons and Nuclei: The Lund Monte Carlo, Fritiof Version 1.6". In: *Comput. Phys. Commun.* 43 (1987), p. 387. DOI: 10.1016/0010-4655(87)90056-7.

[117] Wright, D. H. and Kelsey, M. H. "The Geant4 Bertini Cascade". In: *Nucl. Instrum. Meth.* A804 (2015), pp. 175–188. DOI: 10.1016/j.nima.2015.09.058.

[118] G. Folger and V.N. Ivanchenko and J.-P. Wellisch. "The Binary Cascade - Nucleon nuclear reactions". In: *Eur. Phys. Jour. A* 21 (2004).

[119] ATLAS Collaboration. "Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data". In: *Eur. Phys. J. C* 74 (2014), p. 3071. DOI: 10.1140/epjc/s10052-014-3071-4. arXiv: 1407.5063 [hep-ex].

[120] ATLAS Collaboration. "Search for massive, long-lived particles using multitrack displaced vertices or displaced lepton pairs in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector". In: *Phys. Rev. D* 92 (2015), p. 072004. DOI: 10.1103/PhysRevD.92.072004. arXiv: 1504.05162 [hep-ex].

[121] K. A. Olive et al. "Review of Particle Physics". In: *Chin. Phys.* C38 (2014), p. 090001. DOI: 10.1088/1674-1137/38/9/090001.

[122] ATLAS Collaboration. "Charged-particle multiplicities in pp interactions at $\sqrt{s} = 900$ GeV measured with the ATLAS detector at the LHC". In: *Phys. Lett. B* 688 (2010), p. 21. DOI: 10.1016/j.physletb.2010.03.064. arXiv: 1003.3124 [hep-ex].

[123] ATLAS Collaboration. *Study of the Material Budget in the ATLAS Inner Detector with K^0_S decays in collision data at $\sqrt{s} = 900$ GeV*. ATLAS-CONF-2010-019. 2010. URL: https://cds.cern.ch/record/1277651.

[124] ATLAS Collaboration. *Probing the response of the ATLAS electromagnetic calorimeter and material upstream with energy flow from $\sqrt{s} = 7$ TeV minimum bias events*. ATLAS-CONF-2010-037. 2010. URL: https://cds.cern.ch/record/1277677.

[125] ATLAS Collaboration. *Photon Conversions at $\sqrt{s} = 900$ GeV measured with the ATLAS Detector*. ATLAS-CONF-2010-007. 2010. URL: https://cds.cern.ch/record/1274001.

[126] ATLAS Collaboration. "A study of the material in the ATLAS inner detector using secondary hadronic interactions". In: *JINST* 7 (2012), P01013. DOI: 10.1088/1748-0221/7/01/P01013. arXiv: 1110.6191 [hep-ex].

[127] ATLAS Collaboration. "A study of the material in the ATLAS inner tracker using secondary hadronic interactions in 7 TeV pp collisions". In: *JINST* 11 (2016), P11020. DOI: 10.1088/1748-0221/11/11/P11020. arXiv: 1609.04305 [hep-ex].

[128] ATLAS Collaboration. "The ATLAS Simulation Infrastructure". In: *Eur. Phys. J.* C70 (2010), pp. 823–874. DOI: 10.1140/epjc/s10052-010-1429-9. arXiv: 1005.4568 [physics.ins-det].

[129] A. Parker et al. "The material budget of the ATLAS Inner Detector". In: (1998).

[130] Y.-S. Tsai. "Pair Production and Bremsstrahlung of Charged Leptons". In: *Rev. Mod. Phys.* 46 (1974), p. 815.

[131] Maria Josefina Alconada Verzini et al. *Lessons learned from inner detector material studies with photon conversions using 7 TeV data*. Tech. rep. ATL-COM-PHYS-2016-596. Geneva: CERN, 2016. URL: https://cds.cern.ch/record/2155363.

[132] J. Billan, R. Perin, and V. Sergo. "The Split Field Magnet of the CERN Intersecting Storage Rings". In: *eConf* C720919 (1972), p. 433.

[133] A. Breakstone et al. "Charged Multiplicity Distribution in pp Interactions at CERN ISR Energies". In: *Phys. Rev.* D30:528 (1984). DOI: 10.1103/PhysRevD.30.528.

[134] K. Eggert et al. "A Streamer Chamber Detector at the CERN Intersecting Storage Rings". In: *Nucl. Instrum. Meth.* 126 (1975), pp. 477–491. DOI: 10.1016/0029-554X(75)90798-3.

[135] W. Thome et al. "Charged Particle Multiplicity Distributions in pp Collisions at ISR Energies". In: *Nucl. Phys.* D129:365 (1977).

[136] E. Fisher. "Two Kilometers at 10^{-10} Torr. The CERN Intersecting Storage Rings for Protons". In: *Journal of Vacuum Science and Technology* 9 (1972). DOI: 10.1116/1.1317013.

[137] 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". In: *Nucl. Phys.* B335:261 (1990).

[138] G. Arnison et al. "Charged Particle Multiplicity Distributions in Proton Anti-Proton Collisions at 540 GeV Center Of Mass Energy". In: *Phys. Lett.* B123:108 (1983).

[139] *UA5*. URL: http://cerncourier.com/cws/article/cern/44856.

[140] G. J. Alner et al. "Scaling Violation Favoring High Multiplicity Events at 540 GeV CMS Energy". In: *Phys. Lett.* B138:304 (1982).

[141] G. J. Alner et al. "An Investigation of Multiplicity Distributions in Different Pseudorapidity Intervals in anti-p p Reactions at a CMS Energy of 540 GeV". In: *Phys. Lett.* B160:193 (1985).

[142] G. J. Alner et al. "Scaling of Pseudorapidity Distributions at C.M. Energies Up to 0.9 TeV". In: *Z. Phys.* C:331-6 (1986).

[143] G. J. Alner et al. "Scaling Violations in Multiplicity Distributions at 200 GeV and 900 GeV". In: *Phys. Lett.* B167:476-480 (1986).

[144] G. J. Alner et al. "UA5: A General Study of Proton-Antiproton Physics at $\sqrt{s}$ = 546 GeV". In: *Phys. Rept.* 154:247-383 (1987).

[145] K. Alpgard et al. "Comparison of $p\bar{p}$ and pp Interactions at $\sqrt{s}$ = 53 GeV". In: *Phys. Lett.*. B112:183 (1982).

[146] R. E. Ansorgeet et al. "Diffraction Dissociation at the CERN pulsed $p\bar{p}$ collider at C.M. Energies of 900 and 200 GeV". In: *Z. Phys.* C33:175 (1986).

[147] R. E. Ansorgeet et al. "Charged Particle Correlations in at the Tevatron Collider with $\sqrt{s}$ = 1.8 TeV Collisions at C.M. Energies of 200, 546, and 900 GeV". In: *Z. Phys.* C37:191-213 (1988).

[148] R. E. Ansorgeet et al. "Charged Particle Multiplicity Distributions at 200 and 900 GeV C.M. Energy". In: *Z. Phys.* C43:357 (1989).

[149] R. Harr et al. "Pseudorapidity Distribution of Charged Particles in pp collisions at $\sqrt{s}$ = 630 GeV". In: *Phys. Lett.* B401:176-180 (1997).

[150] T. Alexopoulos et al. "The Role of Double Parton Collisions in Soft Hadron Interactions". In: *Phys. Lett.* B435:453-457 (1998).

[151] Clark S. Lindsey et al. "Results from E735 at the Tevatron $p\bar{p}$ Collider with $p\bar{p}$ = 1.8 TeV". In: *Nucl. Phys.* A544:343-356 (1992).

[152] CDF Collaboration. "The CDF Detector: An Overview". In: *Nucl.Instrum.Meth. A271* FERMILAB-PUB-88-025-E (1988), p. 40. DOI: 10.1016/0168-9002(88)90298-7.

[153] F. Abe et al. "Transverse Momentum Distributions of Charged Particles Produced in $p\bar{p}$ Interactions at $\sqrt{s}$ = 630 GeV and 1800 GeV". In: *Phys. Rev. Lett.* 61:1819 (1988).

[154] F. Abe et al. "Pseudorapidity Distributions of Charged Particles Produced in $p\bar{p}$ Interactions at $\sqrt{s}$ = 630 GeV and 1800 GeV". In: *Phys. Rev.* D41:2330 (1990).

[155] D. Acosta et al. "Soft and Hard Interactions in $p\bar{p}$ Colliions at $\sqrt{s}$ = 1800 GeV and 630 GeV". In: *Phys. Rev.* D65:072005 (2002).

[156] T. Aaltonen et al. "Measurement of Particle Production and Inclusive Differential Cross Sections in $p\bar{p}$ Collisions at $\sqrt{s}$ = 1.96 TeV". In: *Phys. Rev.* D79:112005 (2009).

[157] CDF Collaboration. "Measurement of Particle Production and Inclusive Differential Cross Sections in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV ". In: *Phys. Rev. D* 79 (2009), p. 112005. DOI: 10.1103/PhysRevD.79.112005. arXiv: 0904.1098 [hep-ex].

[158] CMS Collaboration. "Transverse momentum and pseudorapidity distributions of charged hadrons in pp collisions at $\sqrt{s} = 0.9$ and 2.36 TeV". In: *JHEP* 02 (2010), p. 041. DOI: 10.1007/JHEP02(2010)041. arXiv: 1002.0621 [hep-ex].

[159] CMS Collaboration. "Charged particle multiplicities in pp interactions at $\sqrt{s} = 0.9$, 2.36, and 7 TeV". In: *JHEP* 01 (2011), p. 079. DOI: 10.1007/JHEP01(2011)079. arXiv: 1011.5531 [hep-ex].

[160] CMS Collaboration. "Transverse-Momentum and Pseudorapidity Distributions of Charged Hadrons in pp Collisions at $\sqrt{s} = 7$ TeV". In: *Phys. Rev. Lett.* 105 (2010), p. 022002. DOI: 10.1103/PhysRevLett.105.022002. arXiv: 1005.3299 [hep-ex].

[161] CMS Collaboration. "Pseudorapidity distribution of charged hadrons in proton–proton collisions at $\sqrt{s} = 13$ TeV". In: *Phys. Lett. B* 751 (2015), p. 143. DOI: 10.1016/j.physletb.2015.10.004. arXiv: 1507.05915 [hep-ex].

[162] ALICE Collaboration. "Charged-particle multiplicity measurement in proton-proton collisions at $\sqrt{s} = 0.9 and 2.36$ TeV with ALICE at LHC". In: *Eur. Phys. J. C* 68 (2010), p. 89. DOI: 10.1140/epjc/s10052-010-1339-xarXiv:1004.3034. arXiv: 1004.3034 [hep-ex].

[163] ALICE Collaboration. "Charged-particle multiplicity measurement in proton–proton collisions at $\sqrt{s} = 7$ TeV with ALICE at LHC". In: *Eur. Phys. J. C* 68 (2010), pp. 345–354. DOI: 10.1140/epjc/s10052-010-1350-2. arXiv: 1004.3514 [hep-ex].

[164] ALICE Collaboration. "Pseudorapidity and transverse-momentum distributions of charged particles in proton-proton collisions at $\sqrt{s}$ = 13 TeV". In: *Phys. Lett. B* 753 (2016), pp. 319–329. DOI: 10.1016/j.physletb.2015.12.030. arXiv: 1509.08734 [hep-ex].

[165] LHCb Collaboration. "Measurement of charged particle multiplicities in pp collisions at $\sqrt{s}$ = 7 TeV in the forward region". In: *Eur.Phys.J. C* 72 (2012), pp. 1947–48. DOI: 10.1140/epjc/s10052-012-1947-8. arXiv: 1112.4592 [hep-ex].

[166] ATLAS Collaboration. "Charged-particle multiplicities in pp interactions measured with the ATLAS detector at the LHC". In: *New J. Phys.* 13 (2011), p. 053033. DOI: 10.1088/1367-2630/13/5/053033. arXiv: 1012.5104 [hep-ex].

[167] ATLAS Collaboration. "Charged-particle distributions in pp interactions at $\sqrt{s} = 8$ TeV measured with the ATLAS detector at the LHC". In: *Eur. Phys. J. C* 76 (2016), p. 403. DOI: 10.1140/epjc/s10052-016-4203-9. arXiv: 1603.02439 [hep-ex].

[168] A.M. Rossi, G. Vannini, A. Bussière, E. Albini, D. D'Alessandro, G. Giacomelli. "Experimental study of the energy dependence in proton-proton inclusive reactions". In: *Nucl. Phys. B* 84 (1975). DOI: 10.1016/0550-3213(75)90307-7.

[169] PHOBOS Collaboration. "Pseudorapidity distribution of charged particles in d+Au collisions at $\sqrt{s_{NN}} = 200GeV$". In: *Phys. Rev. Lett.* 93 (2004). DOI: 10.1103/PhysRevLett.93.082301.

[170] Z. Marshall. "The ATLAS simulation software". In: (2008). ATL-SOFT-PROC-2008-001.

[171] The ATLAS collaboration. "The ATLAS Simulation Infrastructure". English. In: *The European Physical Journal C* 70.3 (2010), pp. 823–874. ISSN: 1434-6044. DOI: 10.1140/epjc/s10052-010-1429-9. URL: http://dx.doi.org/10.1140/epjc/s10052-010-1429-9.

[172] A. K. Morley and S. Pagan Griso. "Summary of Run 1 ID material description discussion". In: ().

[173] A. Buckley at al. "Multi-strange baryon production in *pp* collisions at $\sqrt{s} = 7$ TeV with ALICE". In: (). URL: http://rivet.hepforge.org/analyses#ALICE_2012_I1097057.

[174] Betty Abelev et al. "Multi-strange baryon production in *pp* collisions at $\sqrt{s} = 7$ TeV with ALICE". In: *Phys. Lett.* B712 (2012), pp. 309–318. DOI: 10.1016/j.physletb.2012.05.011. arXiv: 1204.0282 [nucl-ex].

[175] Betty Bezverkhny Abelev et al. "Production of $\Sigma(1385)^{\pm}$ and $\Xi(1530)^{0}$ in proton-proton collisions at $\sqrt{s} = 7$ TeV". In: *Eur. Phys. J.* C75.1 (2015), p. 1. DOI: 10.1140/epjc/s10052-014-3191-x. arXiv: 1406.3206 [nucl-ex].

[176] Andy Buckley et al. "Production of $\Sigma(1385)^{\pm}$ and $\Xi(1530)^{0}$ in proton-proton collisions at $\sqrt{s} = 7$ TeV." In: (). eprint: https://www.hepforge.org/archive/rivet/contrib/pre2.3.0/#ALICE_2014_I1300380.

[177] G. D'Agostini. "A Multidimensional unfolding method based on Bayes' theorem". In: *Nucl. Instr. Meth.* A362 (1995), pp. 487–498. DOI: 10.1016/0168-9002(95)00274-X.

[178] Torbjörn Sjöstrand. "Colour reconnection and its effects on precise measurements at the LHC, " in: (2013). arXiv: 1310.8073 [hep-ph].

[179] T.Sjöstrand, S. Mrenna, and P. Skands. "PYTHIA 6.4 physics and manual". In: *JHEP 05* arXiv:hep-ph/0603175 (2006).

[180] Peter Zeiler Skands. "Tuning Monte Carlo Generators: The Perugia Tunes". In: *Phys. Rev.* D82 (2010), p. 074018. DOI: 10.1103/PhysRevD.82.074018. arXiv: 1005.3457 [hep-ph].

[181] ATLAS Collaboration. "ATLAS Pythia 8 tunes to 7 TeV data". URL: https://cds.cern.ch/record/1966419.

[182] A. Sherstnev and R. S. Thorne. "Different PDF approximations useful for LO Monte Carlo generators". In: *Proceedings, 16th International Workshop on Deep Inelastic Scattering and Related Subjects (DIS 2008): London, UK, April 7-11, 2008*. 2008, p. 149. DOI: 10.3360/dis.2008.149. arXiv: 0807.2132 [hep-ph]. URL: https://inspirehep.net/record/790666/files/arXiv:0807.2132.pdf.

[183] ATLAS Collaboration. *A study of the Pythia 8 description of ATLAS minimum bias measurements with the Donnachie-Landshoff diffractive model*. ATL-PHYS-PUB-2016-017. 2016. URL: https://cds.cern.ch/record/2154449.

[184] ATLAS Collaboration. "Measurement of the Inelastic Proton-Proton Cross Section at $\sqrt{s} = 13$ TeV with the ATLAS detector at the LHC". In: *Phys. Rev. Lett.* 117 (2016), p. 30. DOI: 10.1103/PhysRevLett.117.182002. arXiv: 1606.02625v4 [hep-ex].

[185] ATLAS Collaboration. "Measurement of the W-boson mass in pp collisions at $\sqrt{s} =$ TeV with the ATLAS detector". In: CERN-EP-2016-305 (2016). arXiv: 1701.07240 [hep-ph].

[186] Charalampos Anastasiou et al. "High precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at NNLO". In: *Phys. Rev. D* 69 (2004), p. 094008. DOI: 10.1103/PhysRevD.69.094008. arXiv: hep-ph/0312266 [hep-ph].

[187] ATLAS Collaboration. "Measurement of $W^{\pm}$ and Z-boson production cross sections in *pp* collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Phys. Lett. B* 759 (2016), p. 601. DOI: 10.1016/j.physletb.2016.06.023. arXiv: 1603.09222 [hep-ex].

[188] ATLAS Collaboration. *Electron identification measurements in ATLAS using* $\sqrt{s} =$ *13 TeV data with* 50 *ns bunch spacing*. ATL-PHYS-PUB-2015-041. 2015. URL: http://cdsweb.cern.ch/record/2048202.

[189] ATLAS Collaboration. *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at* $\sqrt{s} = 13$ *TeV*. arXiv: 1603.05598 [hep-ex].

[190] ATLAS Collaboration. *Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at* $\sqrt{s} = 13$ *TeV*. ATL-PHYS-PUB-2015-015. 2015. URL: https://cds.cern.ch/record/2037613.

[191] ATLAS Collaboration. *Expected performance of missing transverse momentum reconstruction for the ATLAS detector at* $sqrts = 13$ *TeV*. ATL-PHYS-PUB-2015-04. 2015. URL: https://cds.cern.ch/record/2013489.

[192] ATLAS Collaboration. "Improved luminosity determination in pp collisions at sqrt(s) = 7 TeV using the ATLAS detector at the LHC". In: *Eur. Phys. J.* C73.8 (2013), p. 2518. DOI: 10.1140/epjc/s10052-013-2518-3. arXiv: 1302.4393 [hep-ex].

[193] Proposal for pileup reweighting, https://indico.cern.ch/event/437993/contribution/1/attachments/1138739/1630981/spagan_MuRescaling.pdf.

[194] K. A. Olive et al. "Review of Particle Physics". In: *Chin. Phys. C* 38 (2014), p. 090001. DOI: 10.1088/1674-1137/38/9/090001.

[195] The ALEPH, DELPHI, L3, OPAL Collaborations, the LEP Electroweak Working Group. "Electroweak Measurements in Electron-Positron Collisions at W-Boson-Pair Energies at LEP". In: *Phys. Rept.* 532 (2013), p. 119. DOI: 10.1016/j.physrep.2013.07.004. arXiv: 1302.3415 [hep-ex].

[196] A. Glazov. "Averaging of DIS cross section data". In: *AIP Conf. Proc.* 792 (2005), p. 237. DOI: 10.1063/1.2122026.

[197] F. D. Aaron et al. "Measurement of the Inclusive ep Scattering Cross Section at Low Q^2 and x at HERA". In: *Eur. Phys. J. C* 63 (2009), p. 625. DOI: 10.1140/epjc/s10052-009-1128-6. arXiv: 0904.0929 [hep-ex].

[198] ATLAS Collaboration. "Measurement of the inclusive $W^{\pm}$ and Z/γ cross sections in the electron and muon decay channels in *pp* collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector". In: *Phys. Rev. D* 85 (2012), p. 072004. DOI: 10.1103/PhysRevD.85.072004. arXiv: 1109.5141 [hep-ex].

[199] CMS Collaboration. "Measurement of the Inclusive W and Z Production Cross Sections in *pp* Collisions at $\sqrt{s} = 7$ TeV". In: *JHEP* 1110 (2011), p. 132. DOI: 10.1007/JHEP10(2011)132. arXiv: 1107.4789 [hep-ex].

[200] CMS Collaboration. "Measurement of inclusive W and Z boson production cross sections in pp collisions at $\sqrt{s}$ = 8 TeV". In: *Phys. Rev. Lett.* 112 (2014), p. 191802. DOI: 10.1103/PhysRevLett.112.191802. arXiv: 1402.0923 [hep-ex].

[201] LHCb Collaboration, R. Aaij et al. "Measurement of the forward Z boson production cross-section in pp collisions at $\sqrt{s} = 7$ TeV". In: *JHEP* 1508 (2015), p. 039. DOI: 10.1007/JHEP08(2015)039. arXiv: 1505.07024 [hep-ex].

[202] LHCb Collaboration, R. Aaij et al. "Measurement of forward W and Z boson production in pp collisions at $\sqrt{s} = 8\,\text{TeV}$". In: *JHEP* 1601 (2016), p. 155. DOI: 10.1007/JHEP01(2016)155. arXiv: 1511.08039 [hep-ex].

[203] ATLAS Collaboration. "Determination of the strange quark density of the proton from ATLAS measurements of the $W \to \ell\nu$ and $Z \to \ell\ell$ cross sections". In: *Phys. Rev. Lett.* 109 (2012), p. 012001. DOI: 10.1103/PhysRevLett.109.012001. arXiv: 1203.4051 [hep-ex].

[204] T. Gleisberg, S. Höche, F. Krauss, M. Schönherr, S. Schumann, F. Siegert, J. Winter. "Event generation with Sherpa 1.1". In: *JHEP 02 (2009) 007* (). DOI: 10.1088/1126-6708/2009/02/007. arXiv: 0811.4622.

[205] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, A. D. Polosa. "ALPGEN, a generator for hard multiparton processes in hadronic collisions". In: *JHEP 07 (2003) 001* (). arXiv: 0206293 [hep-ph].

[206] T. Stelzer, W. F. Long. "Automatic generation of tree level helicity amplitudes". In: *Comput. Phys. Commun. 81 (1994) 357–371* (). arXiv: 9401258.

[207] Stefano Frixione, Paolo Nason, and Giovanni Ridolfi. "A Positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction". In: *JHEP* 0709 (2007), p. 126. DOI: 10.1088/1126-6708/2007/09/126. arXiv: 0707.3088 [hep-ph].

[208] Torbjörn Sjöstrand, Stephen Mrenna, and Peter Z. Skands. "PYTHIA 6.4 Physics and Manual". In: *JHEP* 05 (2006), p. 026. DOI: 10.1088/1126-6708/2006/05/026. arXiv: hep-ph/0603175 [hep-ph].

[209] Michal Czakon and Alexander Mitov. "Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders". In: *Comput. Phys. Commun.* 185 (2014), p. 2930. DOI: 10.1016/j.cpc.2014.06.021. arXiv: 1112.5675 [hep-ph].

[210] ATLAS Collaboration. "Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1". In: (2016). arXiv: 1603.02934 [hep-ex].

[211] Matteo Cacciari and Gavin P. Salam. "Dispelling the N^3 myth for the k_t jet-finder". In: *Phys. Lett.* B641 (2006), pp. 57–61. DOI: 10.1016/j.physletb.2006.08.037. arXiv: hep-ph/0512210 [hep-ph].

[212] Matteo Cacciari and Gavin P. Salam. "Pileup subtraction using jet areas". In: *Phys.Lett. B* 659 (2008), pp. 119–126. DOI: 10.1016/j.physletb.2007.09.077. arXiv: 0707.1378 [hep-ph].

[213] Matteo Cacciari, Gavin P. Salam, and Sebastian Sapeta. "On the characterisation of the underlying event". In: *JHEP* 1004 (2010), p. 065. DOI: 10.1007/JHEP04(2010)065. arXiv: 0912.4926 [hep-ph].

[214] Morad Aaboud et al. "Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: (2017). arXiv: 1703.09665 [hep-ex].

[215] ATLAS Collaboration. *Expected performance of the ATLAS b-tagging algorithms in Run-2*. ATL-PHYS-PUB-2015-022. 2015. URL: https://cds.cern.ch/record/2037697.

[216] ATLAS Collaboration. *Commissioning of the ATLAS high performance b-tagging algorithms in the 7 TeV collision data*. ATLAS-CONF-2011-102. 2011. URL: https://cds.cern.ch/record/1369219.

[217] ATLAS Collaboration. *Tagging and suppression of pileup jets with the ATLAS detector*. ATLAS-CONF-2014-018. 2014. URL: https://cds.cern.ch/record/1700870.

[218] ATLAS Collaboration. *Selection of jets produced in 13TeV proton-proton collisions with the ATLAS detector*. Tech. rep. ATLAS-COM-CONF-2015-024. Geneva: CERN, 2015. URL: https://cds.cern.ch/record/2016323.

[219] ATLAS Collaboration. *Impact Parameter Resolution Using 2016 MB Data*. URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/IDTR-2016-018/.

[220] L. Rossi and O. Brüning. "High Luminosity Large Hadron Collider A description for the European Strategy Preparatory Group". In: Tech. Rep. CERN-ATS-2012-236 (2012). URL: https://cdsweb.cern.ch/record/1471000.

[221] ATLAS Collaboration. "Letter of Intent for the Phase-I Upgrade of the ATLAS Experiment". In: CERN-LHCC-2011-012. LHCC-I-020 (2011). URL: https://cds.cern.ch/record/1402470.

[222] ATLAS Collaboration. "Letter of Intent for the Phase-II Upgrade of the ATLAS Experiment". In: CERN-LHCC-2012-022. LHCC-I-023 (2012). URL: https://cds.cern.ch/record/1502664.

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

[224] *Garfield - simulation of gaseous detector*. URL: http://garfield.web.cern.ch/.

[225] *ATLAS Virtual Point One - VP1*. URL: http://atlas-vp1.web.cern.ch/atlas-vp1/home/.

[226] *ATLAS Run Time Tester*. URL: https://atlas-rtt.cern.ch/.