

Charged particle spectra in different final states at $\sqrt{s} = 13$ TeV with the CMS Experiment

Dissertation

zur Erlangung des Doktorgrades
des Departments Physik
der Universität Hamburg

vorgelegt von

Juan Manuel Grados Luyando
aus Puebla, Mexiko

Hamburg

2017



Gutachter der Dissertation:

Dr. Hannes Jung

Prof. Dr. Johannes Haller

Gutachter der Disputation:

Prof. Dr. Daniela Pfannkuche

Prof. Dr. Erika Garutti

Dr. Paolo Gunnellini

Prof. Dr. Johannes Haller

Dr. Hannes Jung

Datum der Disputation:

2. November 2017

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Dissertationsschrift selbst verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

I hereby declare, on oath, that I have written the present dissertation by my own and have not used other than the acknowledged resources and aids.

Hamburg, den 17 August 2017

ABSTRACT

In this thesis, measurements of different charged-particle distributions performed with the CMS detector in proton-proton collisions at a centre-of-mass energy $\sqrt{s} = 13$ TeV are presented. In particular, the charged-particle pseudorapidity distributions, also referred to as charged-particle pseudorapidity densities, the multiplicity distributions of charged particles per event, and different transverse-momentum distributions corresponding to all charged-particles, the leading charged-particle, and the integrated spectrum of the latter are presented. The charged-particles are selected with transverse momenta $p_T > 0.5$ GeV in the range $|\eta| < 2.4$. The data are corrected for the different detector effects such as resolution and efficiency. The measured distributions are presented for four different samples corresponding to an Inclusive event sample, an Inelastic-enhanced event sample, a sample dominated by non-single diffractive dissociation events (NSD-enhanced sample), and a sample enriched by single-diffractive dissociation events (SD-enhanced sample).

The measurements presented in this thesis provide extensive and unique insights into low-energy exchange processes that dominate the proton-proton interactions. The rich variety of distributions presented for different event samples, especially those enhanced in diffractive processes, are valuable sources of information to understand the transition from the perturbative to the non-perturbative regions and to adjust the model parameters present in modern MC event generators.

ZUSAMMENFASSUNG

In dieser Arbeit werden verschiedene Verteilungen von geladenen Teilchen aus Daten bestimmt, die mit dem CMS-Detektor bei Proton-Proton-Kollisionen bei einer Schwerpunktsenergie von $\sqrt{s} = 13$ TeV aufgezeichnet wurden. Es wurden Verteilungen gemessen als Funktion der Pseudorapidität, der Teilchenmultiplizität und des Transversalimpulses für alle geladenen Teilchen und für die geladenen Teilchen, welche den größten Transversalimpuls haben. Zusätzlich werden auch Spektren, integriert über den Transversalimpuls vorgestellt. In den Analysen werden geladene Teilchen mit Transversalimpuls $p_T > 0.5$ GeV und eine Pseudorapidität $|\eta| < 2.4$ untersucht. Die gemessenen Verteilungen werden für die verschiedenen Detektoreffekte wie Auflösung und Effizienz korrigiert. Es werden verschiedene Ereigniskategorien untersucht: Inclusive Ereignisse, Ereignisse angereichert mit Inelastischen Ereignissen, Ereignisse angereichert mit sogenannten "Non-Single Diffractive Events", und Ereignisse angereichert mit diffraktiven Ereignissen.

Die in dieser Arbeit vorgestellten Messungen liefern umfangreiche und neue Einblicke in Prozesse, welche die Proton-Proton-Wechselwirkungen dominieren. Die verschiedenen Verteilungen, die für die verschiedenen Ereigniskategorien präsentiert wurden, speziell diejenigen für diffraktive Prozesse, sind essentiell um den Übergang vom perturbativen (bei relativem hohem Transversalimpuls) zum nicht-perturbativen (bei sehr kleinem Transversalimpuls) Bereich zu untersuchen und um die Modellparameter in modernen MC-Event-Generatoren besser bestimmen zu können.

CONTENTS

1	Introduction	13
2	Theory	17
2.1	The Standard Model of Particle Physics	17
2.1.1	The electroweak interaction	20
2.1.2	The strong interaction	23
2.2	Hadron-hadron collisions	25
2.2.1	The Quark Parton Model	26
2.2.2	Parton Distribution Functions	27
2.2.3	Soft and hard interactions	30
2.2.4	Multi parton interactions	32
2.2.5	Diffraction events	33
2.3	Monte Carlo event generators	36
2.3.1	MPI models	37
2.3.2	Hadronization models	39
2.4	Minimum Bias events	41
3	Experiment	43
3.1	The Large Hadron Collider	43
3.2	The CMS experiment	45
3.2.1	Trigger system	47
3.2.2	Silicon tracker	47
3.2.3	Hadronic Forward calorimeter	49
4	Event reconstruction	51
4.1	Track reconstruction	52
4.2	Primary vertex	53
4.3	HF towers	54

4.4	Detector simulation	55
4.4.1	Pile-up simulation	55
5	Event selection	57
5.1	Data samples	57
5.2	MC samples	58
5.3	Detector level selection	59
5.3.1	Trigger selection	59
5.3.2	Vertex selection	59
5.3.3	Track selection	59
5.3.4	Event selection	60
5.4	Stable particle level definition	60
6	Detector performance	65
6.1	Monte Carlo reweighting	65
6.1.1	Vertex multiplicity reweighting	65
6.1.2	Vertex z position reweighting	68
6.2	Tracks	70
6.3	HF towers	80
7	Detector effects	85
7.1	Matching	86
7.2	Resolution effects	87
7.3	Response matrices	90
7.4	Acceptance, background, purity and stability	91
7.4.1	Pileup effects	93
7.4.2	Full event selection	97
8	Efficiency corrections and unfolding	105
8.1	D'Agostini unfolding method	106
8.2	Monte Carlo unfolding	107
8.3	Toy Monte Carlo unfolding	112
8.4	Validation of the procedure	113
8.5	Unfolding conclusions	121
9	Uncertainties	133
9.1	Model dependence	133
9.2	Event selection	134
9.3	Tracking efficiency	137
9.4	Pile-up contribution	137
9.5	Total systematic uncertainties	140
10	Results	145
10.1	Comparison to Monte Carlo event generators	146
10.2	MPI effects	154
10.3	Saturation effects	159
10.4	Hard contribution to diffractive events	164

11 Summary and conclusions	171
Appendices	175
A Tracker alignment	177
A.1 Track-based alignment	177
A.1.1 MillePede	179
A.2 Personal contributions	180
B Additional control plots	183
B.1 Vertex	183
B.2 Tracks	187
12 Bibliography	207

CHAPTER

1

INTRODUCTION

The Standard Model of particle physics is the theoretical framework developed to explain the most fundamental building blocks of matter and the interactions between them. Within this theory a huge variety of phenomena are successfully described up to unprecedented accuracy. In order to test the theoretical predictions of the Standard Model, physicists have been finding ingenious ways to probe the matter at the smallest possible scales, just being limited by the technology of their time. Nowadays, the Large Hadron Collider (LHC) at the European Organization for Nuclear Research (CERN from its French initials) is the largest and most powerful ever constructed particle accelerator in the world. It consists of two rings with superconducting magnets in which two proton beams circulate in opposite directions. There are four interaction points located along the circumference, each of them corresponding to one of the four largest experiments at the LHC. These experiments allow to measure the properties of the different particles emerging from the collisions, making it possible to test the theoretical predictions of the Standard Model and to search for new particles predicted by theories going beyond our present understanding.

One of the fundamental forces in nature is the *strong force*, which only affects the particles carrying color charge. The strong force is described within the Standard Model by the theory of Quantum Chromodynamics (QCD), which describes the interaction between quarks and gluons, also known as *partons*. Since the protons are composite particles made out of quarks and gluons, the proton-proton collisions are primarily described by QCD. Even though the theory of QCD provides the exact dynamical equations, the current mathematical methods are not able to solve these equations exactly. Nevertheless, the use of perturbation theory allows us to obtain approximate solutions as a power series expansion in function of the strong force running coupling constant α_s . The perturbation theory can be applied at large energy transfers corresponding to values of the running coupling constant smaller than unity. At lower energy transfers, α_s grows to values larger than unity, and the perturbation theory is not valid anymore.

The bulk of particles produced in proton-proton collisions are predominantly characterised by a small transverse momentum p_T and arise from soft to semi-hard (multi)parton interactions characterised by low energy transfers. This makes a somehow inverted situation where it is possible to calculate (at least approximately) the most rare interactions at large exchanged energies, but where the dominant soft to semi-hard contributions cannot be predicted from first principles. Hence, QCD-inspired phenomenological models have been developed to obtain predictions in this region of the phase space. The measurement of particle production without any bias of a hard scattering is one of the very first and basic measurements in hadronic colliders at high energies. Especially the measurement of low- p_T particles is particularly important to constrain and improve the modelling of particle production mechanisms and to probe the transition from the non-perturbative to the perturbative QCD region.

This thesis describes measurements that can help to better understand the transition region from the perturbative to the non-perturbative regime, and to constrain the modelling of the different processes involved in particle production. Different charged-particle distributions measured with the Compact Muon Solenoid (CMS) detector in proton-proton collisions at a centre-of-mass energy $\sqrt{s} = 13$ TeV are presented, namely the charged-particle pseudorapidity distributions, also referred to as charged-particle pseudorapidity densities, the multiplicity distributions of charged particles per event, and different transverse-momentum distributions corresponding to all charged-particles, the leading charged-particle, and the integrated spectrum of the latter. The charged particles are selected with transverse momenta $p_T > 0.5$ GeV in the range $|\eta| < 2.4$.

Inclusive measurements of charged-particle pseudorapidity distributions, $dN_{\text{ch}}/d\eta$, and transverse momentum distributions, dN_{ch}/dp_T , have previously been performed in proton-proton (pp) and proton-antiproton ($p\bar{p}$) collisions for different centre-of-mass energies and phase space regions [1–15]. The first measurement of the charged-hadron pseudorapidity distribution, $dN_{\text{had}}/d\eta$, at $\sqrt{s} = 13$ TeV by CMS is reported in [1].

In this thesis, the measured distributions are presented for four different samples corresponding to an Inclusive event sample, an Inelastic-enhanced event sample, a sample dominated by non-single diffractive dissociation events (NSD-enhanced sample), and a sample enriched by single-diffractive dissociation events (SD-enhanced sample). Single-diffractive dissociation events are characterised by the absence of hadronic activity in a large region of the detector and interpreted in a framework of QCD as resulting from a colorless interaction between the two colliding proton beams. The study of the previously mentioned distributions for different event selections provides richer information to further constrain the phenomenological models and enlightens the path to a better understanding of the complex processes involved in the soft QCD regime and in the transition to the perturbative QCD region.

The thesis is structured as follows. In Chapter 2, an introduction to the basic concepts of the Standard Model with special focus on the strong interaction is presented. In addition, the main concepts of hadronic collisions and diffractive processes are introduced, as well as their simulation by Monte Carlo event generators. The description of the experimental setup, namely the LHC and the different experiments located at the interaction points are introduced

in Chapter 3, a special attention being paid to the CMS experiment and the subdetectors relevant in this analysis. In Chapter 4 the reconstruction of the different physics objects, as well as the detector and pileup simulations are described. The Monte Carlo and data samples used in the analysis, the detector level and stable particle level event selections are described in Chapter 5. The performance of the track and vertex reconstruction, and the performance of the Hadronic Forward calorimeters are studied in Chapter 6. In Chapter 7, resolution studies are presented, as well as the efficiency, fake rate, purity and stability estimates associated to the track reconstruction for each of the event selections. The description of the procedure to correct the data for detector effects is explained in Chapter 8 while the uncertainties associated to the measurements and corrections are outlined in Chapter 9. The final results of this thesis are presented in Chapter 10. First, a comparison of the data to different Monte Carlo event generators is discussed. Then, some of the most important parameters involved in the description of soft particle production and the transition region in modern Monte Carlo event generators are studied along with various comparisons to different models. The last chapter is devoted to the summary and outlook of the presented analysis.

CHAPTER

2

THEORY

Contents

2.1	The Standard Model of Particle Physics	17
2.1.1	The electroweak interaction	20
2.1.2	The strong interaction	23
2.2	Hadron-hadron collisions	25
2.2.1	The Quark Parton Model	26
2.2.2	Parton Distribution Functions	27
2.2.3	Soft and hard interactions	30
2.2.4	Multi parton interactions	32
2.2.5	Diffraction events	33
2.3	Monte Carlo event generators	36
2.3.1	MPI models	37
2.3.2	Hadronization models	39
2.4	Minimum Bias events	41

2.1 The Standard Model of Particle Physics

The Standard Model of particle physics (SM) is the accepted theory of the fundamental building blocks of matter and the interactions between them.

The definition of what is considered to be the most basic building blocks of matter has evolved thanks to the improvement of our understanding of nature. An example of the evolution of our comprehension of nature is the word Atom. It originally comes from the ancient Greek words *a*–“not” and *tomos*–“a cutting”, i.e., an non-divisible elementary object. The

Chapter 2. Theory

word atom was used again in the XIX century by John Dalton to explain the elements and their reactions. However, we nowadays know that atoms are not elementary particles, but are build from electrons, protons and neutrons. It is though interesting to notice how antique the idea of all matter being constructed out of indivisible basic entities is.

But it is not until the XIX and XX centuries that the real revolution of particle physics took place. In this period the structure of atoms is unveiled with the discovery of the electron (J.J. Thomson - 1897 [16]) and the discovery of the proton (Rutherford - 1920). In the following years, a series of revolutionary theories, such as the quantum theory, and the discovery of new particles, such as the neutrons, muons, pions, etc., paved the road to what is nowadays considered the most successful and precise theory of physics.

The different elementary particles in the SM are grouped in three distinctive groups: two fermion groups (of spin- $\frac{1}{2}$) called *leptons* and *quarks*, and one boson group (of spin-1) called *gauge-bosons*. The first two groups form the elementary building blocks of matter. The latter are responsible for carrying three of the four fundamental forces known in nature: the electromagnetic, weak and strong forces. The electromagnetic and weak interactions are nowadays unified in a single description called the electroweak interaction. Leptons are only sensitive to the electroweak interaction, while the quarks are sensitive to the electroweak and strong interactions. The fourth fundamental force in nature is gravity, but since the theoretical ground for a quantum theory of gravity is still not successful, the inclusion to the SM is not possible. However, the effects of the gravitational interaction between the elementary particles at the current experimental energies can be neglected in comparison with the other three forces. Figure 2.1 shows the relative strength of the four fundamental forces relative to the electromagnetic force, its range and which particles are affected by each force.

Property \ Interaction	Gravitational	Weak	Electromagnetic	Strong	
		(Electroweak)		Fundamental	Residual
Acts on:	Mass – Energy	Flavor	Electric Charge	Color Charge	Footnote 1
Particles experiencing:	All	Quarks, Leptons	Electrically charged	Quarks, Gluons	Hadrons
Particles mediating:	Graviton (not yet observed)	$W^+ W^- Z^0$	γ	Gluons	Mesons
Strength relative to electromag for two u quarks at: $\left\{ \begin{array}{l} 10^{-18} \text{ m} \\ 3 \times 10^{-17} \text{ m} \end{array} \right.$ for two protons in nucleus	10^{-41}	0.8	1	25	Not applicable to quarks
	10^{-41}	10^{-4}	1	60	
	10^{-36}	10^{-7}	1	Not applicable to hadrons	20

Figure 2.1: Summary of the properties of the four fundamental forces in nature. The relative strength relative to the electromagnetic force, its range and the particles affected by each force are listed [17].

In addition to the mentioned groups of particles, one more spin-0 boson, the so-called *Higgs boson*, is needed to explain the spontaneous symmetry breaking of the electroweak interaction, and the origin of mass within the theory. Figure 2.2 shows a summary of all the elementary constituents of the SM and their properties.

¹ The strong binding of color-neutral protons and neutrons to form nuclei is due residual strong interactions between their color-charged constituents. It is similar to the residual electrical interaction that binds electrically neutral atoms to form molecules. It can be viewed as the exchange of mesons between the hadrons.

Fermions are grouped in three *generations*, each of them containing four particles: one up-type quark, one down-type quark, one charged lepton and one neutral lepton. The differences between the generations are the masses and flavours of the quarks and the masses of the charged leptons, with the masses increasing from generation to generation. For the case of the neutral leptons (neutrinos) the mass hierarchy is uncertain, since their masses have not been yet experimentally determined. Currently, upper limits for the neutrino masses and relations between the masses exist, which let different possibilities open.

FERMIONS				BOSONS	
generation	I	II	III		
QUARKS	up type	u_p $S=\frac{1}{2}$ $Q=+\frac{2}{3}e$ $m=2.3 \text{ MeV}$	c_{harm} $S=\frac{1}{2}$ $Q=+\frac{2}{3}e$ $m=1.28 \text{ GeV}$	t_{op} $S=\frac{1}{2}$ $Q=+\frac{2}{3}e$ $m=173.2 \text{ GeV}$	g_{luon} $S=1$ $Q=0$ $m=0$
	down type	d_{own} $S=\frac{1}{2}$ $Q=-\frac{1}{3}e$ $m=4.8 \text{ MeV}$	s_{trange} $S=\frac{1}{2}$ $Q=-\frac{1}{3}e$ $m=95 \text{ MeV}$	b_{ottom} $S=\frac{1}{2}$ $Q=-\frac{1}{3}e$ $m=4.18 \text{ GeV}$	γ_{photon} $S=1$ $Q=0$ $m<1\cdot10^{-18} \text{ eV}$
LEPTONS	charged	e_{lectron} $S=\frac{1}{2}$ $Q=-1e$ $m=0.51 \text{ MeV}$	μ_{uon} $S=\frac{1}{2}$ $Q=-1e$ $m=105.7 \text{ MeV}$	τ_{au} $S=\frac{1}{2}$ $Q=-1e$ $m=1.777 \text{ GeV}$	W[±]_{boson} $S=1$ $Q=\pm 1e$ $m=80.4 \text{ GeV}$
	neutral	ν_e^{neutrino} $S=\frac{1}{2}$ $Q=0$ $m < 2 \text{ eV}$	ν_μ^{neutrino} $S=\frac{1}{2}$ $Q=0$ $m < 2 \text{ eV}$	ν_τ^{neutrino} $S=\frac{1}{2}$ $Q=0$ $m < 2 \text{ eV}$	Z⁰_{boson} $S=1$ $Q=0$ $m=91.2 \text{ GeV}$
				H_{iggs} $S=0$ $Q=0$ $m=125.7 \text{ GeV}$	

Figure 2.2: Elementary constituents of the Standard Model of particle physics. Fermions are shown on the left side, subdivided in three generations and in the quark and lepton families. The last column on the right shows the boson group divided in the force-carriers and the Higgs boson [18].

Every fermion has a corresponding anti-particle, i.e. a particle which has all the additive quantum numbers reversed while the mass and spin remain the same. The usual notation for an anti-particle is a bar on top of the corresponding symbol of the particle, e.g., for a given quark q , the corresponding anti-quark is denoted as $\bar{q}$.

The quarks can be combined following certain conservation rules, to create composite particles. The combination of three quarks (anti-quarks) forms a *baryon* (qqq) (*anti-baryon* ($\bar{q}\bar{q}\bar{q}$)), and the combination of a quark and an anti-quark gives a *meson* ($q\bar{q}$). All baryons and mesons are grouped together in the category of *hadrons*. The most common type of

baryons are the proton and the neutron, (uud) and (udd) respectively. In the case of the mesons, one of the most common is the pion, e.g., π^+ ($u\bar{d}$). Since atoms are composite states of protons, neutrons and electrons, all the visible matter in the universe is made out of leptons and quarks from the first generation.

An important tool to represent the interactions between particles are the Feynman diagrams. In these diagrams the different particles are represented with different types of lines. These graphical representations can be directly translated to equations by substituting each line and vertex by its corresponding mathematical object. Hence, it is a widely used representation in particle physics. Figure 2.3 shows an example of a Feynman diagram representing a electron-positron pair annihilation into a photon, which then splits into a $q\bar{q}$ pair with a gluon being radiated by the $\bar{q}$.

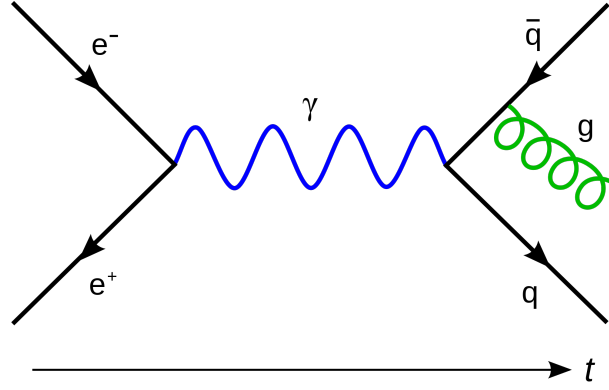


Figure 2.3: Example of a Feynman diagram representing an electron-positron pair annihilation into a photon, with subsequent splitting of the photon into a quark-antiquark pair, with the antiquark radiating a gluon. Time is represented in the x-axis.

The SM is a gauge theory, based on the symmetry group $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ which describe the strong, weak and electromagnetic interactions. In the following sections, the different interactions between particles within the SM are described.

2.1.1 The electroweak interaction

The electroweak interaction (EW) is the unification of the electromagnetic and weak interactions in a single theoretical description. The unification occurs for collision energies larger than ~ 100 GeV, usually referred to as the EW scale. Below the EW scale, the theory predicts different behaviours for the electromagnetic and weak interactions. Figure 2.4 demonstrates the unification in e^-p and e^+p collisions at the HERA collider. At low momentum transfer Q^2 , neutral current processes with photon exchange dominate over the charged current processes mediated via W bosons [19]. Over 100 GeV both contributions become of similar size, and the electroweak unification occurs.

The electroweak theory is described by the non-abelian gauge symmetry group $SU(2)_L \otimes U(1)_Y$. The hyper-charge operator Y defined as $Y = 2(Q - T^3)$ is the generator of the $U(1)_Y$ symmetry, where Q is the conserved electric charge. The weak isospin operators $T^{1,2,3}$ are

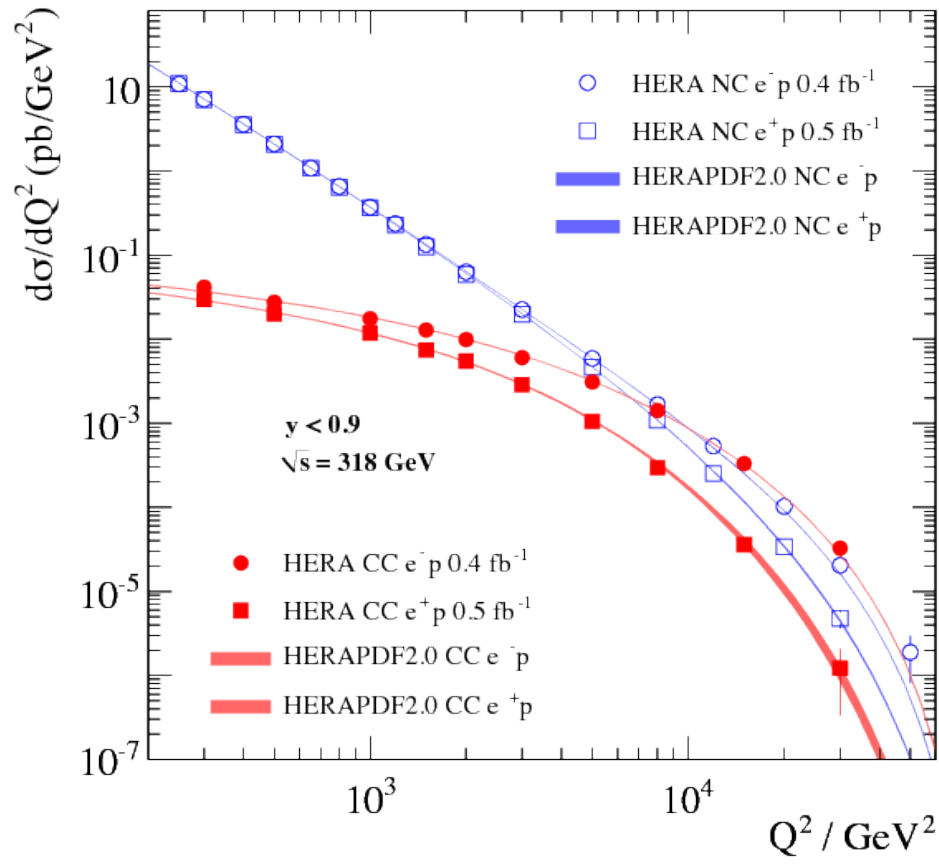


Figure 2.4: Differential cross-section versus momentum transfer-squared measured at the HERA collider for neutral (blue) and charged (red) current deep inelastic scattering (DIS) cross section [19].

the generators of the $SU(2)_L$ symmetry, where the subscript L denotes the fact that weak interactions only affect the left handed components of the Dirac spinors. The electroweak Lagrangian is defined as:

$$\mathcal{L}_{SU(2)_L \otimes U(1)_Y} = \bar{\psi} i \gamma_\mu D^\mu \psi - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} - \frac{1}{4} W_{\mu\nu}^i W_i^{\mu\nu}, \quad (2.1)$$

where in the first term, ψ represents the Dirac field of a spin 1/2 particle, γ_μ the gamma matrices and D_μ the gauge covariant derivative which ensures the conservation of local symmetries, defined as:

$$D_\mu = \partial_\mu + i \frac{g'}{2} Y B_\mu + i \frac{g}{2} \tau_a W_\mu^a. \quad (2.2)$$

τ_a are the Pauli matrices related to the isospin operators as $T^a = \tau_a/2$. W_μ^i and B_μ are the $SU(2)$ and $U(1)$ gauge fields respectively, with field strength tensors

$$\begin{aligned} B_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu, \\ W_{\mu\nu}^i &= \partial_\mu W_\nu^i - \partial_\nu W_\mu^i - g \epsilon_{ijk} W_\mu^j W_\nu^k, \end{aligned} \quad (2.3)$$

where $g(g')$ is the $SU(2)$ ($U(1)$) gauge coupling and ϵ_{ijk} is the totally antisymmetric symbol.

A key feature of the electroweak theory is the concept of Spontaneous Symmetry Breaking (SSB). When constructing the electroweak Lagrangian, the gauge bosons are required to be massless in order to preserve the gauge invariance. This requirement is in contradiction with the experimental results that the weak gauge bosons are massive. The addition of the masses to the Lagrangian while keeping the theory finite is achieved by the so called *Higgs Mechanism*, a direct application of the SSB. As a consequence of the Higgs Mechanism, a massive scalar Higgs boson is produced, the W and Z bosons acquire a mass, while the photon remains massless. The Higgs boson was the last piece of the SM to be simultaneously discovered in 2012 by two of the main experiments of the Large Hadron Collider at CERN.

In the next two sections the electromagnetic and weak interactions are discussed separately.

Electromagnetic interaction

The electromagnetic force is carried by the exchange of one or more photon gauge-bosons (γ) and it is described by Quantum Electro Dynamics (QED). The photon only couples to electrically-charged particles, and as a consequence it does not interact with itself. Due to the massless nature of the photon, the electromagnetic force decreases proportionally to the inverse of the distance between the interacting particles with an infinite range.

An interesting feature of the electromagnetic force is that the strength of the interaction (coupling strength) increases with the energy of the interaction, or equivalently, when the interaction occurs at smaller distances. This is an effect of the quantum fluctuations of photons being emitted by the charged particle, which fluctuate into e^-e^+ pairs forming a cloud surrounding the source.

This cloud is polarised by the electric field of the original charge creating a screening effect. Low energy (large wave length) probes measure a smaller electric field, while large energy (small wave length) probes penetrate the cloud and measure a larger charge as the cloud is penetrated. The effect of the coupling strength varying as a function of the energy of the interaction is known as *running coupling*. The electromagnetic coupling strength has a logarithmic dependence on the energy of the interaction getting the value $\alpha_e \sim 1/137$ at low interaction energies. The running coupling is not a unique feature of the electromagnetic interaction, it is also present in the other forces described by the SM, yet with some important differences. The differences are discussed in the corresponding sections.

Weak interaction

The weak interaction affects the quarks and leptons through the weak isospin by the exchange of the gauge bosons $W^\pm$ and Z^0 . In contrast to the photon of the electromagnetic force, the gauge bosons of the weak interaction are massive, with masses equal to 80.4 GeV and 91.2 GeV for the $W^\pm$ and Z^0 [20], respectively.

The large masses of the $W^\pm$ and Z^0 gauge bosons make the weak interaction both, short ranged, and weak with respect to the other two forces of the SM (at energies below the EW scale). The typical range of the weak force is of the order of $\sim 10^{-16} - 10^{-17}$ m. At the beginning, this short range was described by a contact interaction within the Fermi theory. The weak force is the responsible of the *Beta* and *Gamma* radiations, and is what drives the nuclear reactions inside the Sun together with the strong force.

2.1.2 The strong interaction

As explained at the beginning of this chapter, the quarks can be combined to form composite particles containing either one quark and one anti-quark or three quarks (anti-quarks). An immediate problem of the description of hadrons by the combination of different quarks is the apparent possible construction of quantum-systems containing different fermions in the same quantum state, leading to a violation of the Pauli exclusion principle. This problem was solved by introducing a new additive quantum number, the *color charge*. The color charge can take three possible values, in contrast with the electric charge which can take only one. In that way, each of the quarks inside a hadron carries a different color charge, and the Pauli exclusion principle is satisfied. The three color charges are commonly named as red, green and blue. The combination of three different color charges or a color-anticolor combination creates a color-neutral particle, a property of all free particles in nature.

The strong force is responsible for the interaction between color-charged particles, giving its name to the theory explaining it: *Quantum Chromodynamics* (QCD). In QCD the force carriers are the *gluons*. These are massless gauge bosons just as the photon, but with the difference that they carry the charge of the interaction they mediate and therefore, interacting with themselves. Since the gluon is massless, one could expect to have a long range interaction, just as for the electromagnetic force. The empirical experience is in contradiction with this fact. It is the color-charge property of the gluon that makes the interaction short ranged, giving rise to two of the main features of QCD: *the asymptotic freedom* and *the color*

confinement.

In contrast with the EM screening effect, where the photon is neutral, the gluon can interact with itself due to its color charge. This creates an *anti*-screening effect in which the strength of the interaction, represented by the strong coupling α_s , decreases logarithmically for increasing interaction energies. This means that the strong force decreases at short distances, making the quarks that are inside composite particles behave asymptotically free as they come closer [21], a property known as the asymptotic freedom of QCD.

On the contrary, the interaction strength becomes large at large distances or small interaction energies. As α_s increases, the potential energy of the color field grows, until pairs of quark and antiquarks are created from the vacuum and new neutral mesons are created instead of free quarks. In that way the color-charge is always confined into color-neutral particles.

The fact that α_s decreases as the energy of the interaction increases makes QCD a theory calculable in a perturbative expansion in powers of α_s for energies where $\alpha_s \ll 1$. The expression for α_s as a function of the momentum transfer Q^2 , calculated from the perturbative expansion at LO can be written as:

$$\alpha_s(Q^2) = \frac{12\pi}{(33 - 2N_f) \ln \frac{Q^2}{\Lambda_{\text{QCD}}^2}}, \quad (2.4)$$

where N_f is the number of active quark flavours. The QCD scale or confinement scale Λ_{QCD} is an experimentally determined parameter and has been determined to be $\Lambda_{\text{QCD}} \sim 200 \text{ MeV}$. Figure 2.5 shows a summary of several α_s measurements as a function of the momentum transfer Q [22].

In the low energy regime, also known as the non-perturbative regime, QCD cannot be used to calculate observables from first principles. The non-perturbative QCD regime and the transition between the two regimes can only be described by phenomenological models.

The QCD theory is a non-abelian gauge theory described by the symmetry group $SU(3)_C$. The QCD Lagrangian can be written as:

$$\mathcal{L}_{SU(3)_C} = -\frac{1}{4} F_{\mu\nu}^i F_i^{\mu\nu} + \sum_f \bar{q}_f (i\gamma^\mu D_\mu - m_q) q_f, \quad (2.5)$$

where $F_{\mu\nu}^i$ is the field strength tensor for the gluon fields G_μ^i ($i = 1, \dots, 8$),

$$F_{\mu\nu}^i = \partial_\mu G_\nu^i - \partial_\nu G_\mu^i - g_s f_{ijk} G_\mu^j G_\nu^k, \quad (2.6)$$

with g_s the QCD gauge coupling constant and f_{ijk} the structure functions of the Lie algebra of the $SU(3)_C$ group defined by

$$[\lambda^i, \lambda^j] = 2if_{ijk}\lambda^k, \quad (2.7)$$

with λ^i the $SU(3)$ generator matrices. The term $F_{\mu\nu}^i F_i^{\mu\nu}$ in $\mathcal{L}_{SU(3)_C}$ contains the three and four gluon self-interactions which lead to the anti-screening effect discussed above and to the

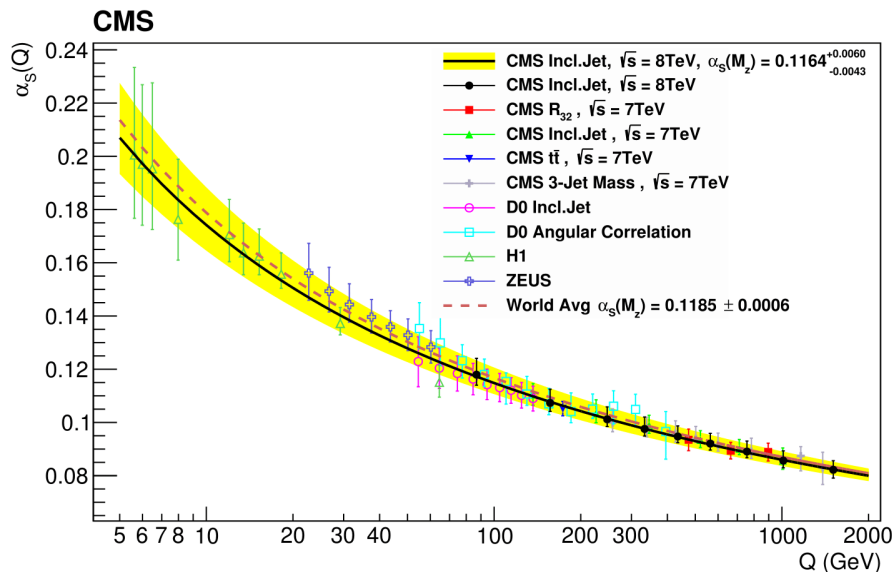


Figure 2.5: Summary of measurements of α_s as a function of Q . The black dots in the figure show the number obtained from the $\sqrt{s} = 8$ TeV inclusive jet measurement. Measurements from H1, ZEUS, and D0 experiments at the HERA and Tevatron colliders are displayed as well [22].

asymptotic freedom of QCD.

The second term in equation 2.5 is the gauge covariant derivative for the quark fields where the summation runs over the quark flavours, α and γ being the color indices

$$D_\mu = \partial_\mu - ig_s \frac{\lambda_\alpha}{2} G_\mu^\alpha. \quad (2.8)$$

2.2 Hadron-hadron collisions

Hadron-hadron collisions are one of the means developed to study the internal structure of hadrons, the properties of QCD and the SM in general. The confinement of color-charged particles into hadrons makes the experimental study of QCD an indirect task. In any kind of experiment involving collisions of particles, the detectors can only directly detect final state particles and their decays, which can be of hadronic or leptonic nature. Only hadrons which are colorless bound states of quarks and gluons can be detected, while the access to the quark and gluon degrees of freedom can only be inferred from the theory.

Before describing the properties and characteristics of hadron-hadron collisions, it is important to introduce the Quark Parton Model and the Parton Distribution Functions.

2.2.1 The Quark Parton Model

The Quark Parton Model (QPM) was gradually developed by gathering different ideas as new experimental data became available. The first piece of the QPM was the Quark Model (QM) developed during the 60s by Gell-Mann and Ne’eman to classify all the particles known by that time in a theoretical group. This idea resembles the construction of the periodic table of elements by Mendeleev, in which the large amount of particles discovered by then (referred by that time as “the particle zoo”) could be accommodated in the different positions of the arrangement.

A great success of the QM was the prediction of new particles needed to fill vacant states predicted by the group theory, e.g. the Ω^- at Brookhaven in 1964.

A peculiarity of this new description was that all particles were sitting in higher representations of the theoretical group, suggesting the presence of a more basic underlying structure. Gell-Mann and Zweig made a proposal of three new particles which could represent the underlying basic structure of the group. Gell-Mann named them *quarks*.

Later in mid 60’s, in electron-proton (ep) collisions at the SLAC experiment [23] and at DESY [24], in analogy to the Rutherford experiment to unveil the structure of the atom, it was found that the proton has an internal structure of point-like particles which were called *partons*. During this period, Feynman developed the Parton Model (PM) which gave the best description of new experimental data. In the subsequent years, new experiments contributed to indicate that the partons of the PM had the properties predicted by the QM for the quarks. It was clear that the partons could be identified as the quarks.

The combination of these two ideas confirmed by the experiments is commonly known as the *Naive Quark Parton Model* (nQPM). In the nQPM the proton is composed only by three loosely bound quarks (*valence quarks*), and ep collisions can be seen as the scattering of an electron off a quark by a photon exchange. Each of the constituents of the proton shares some fraction x_i (*x-Bjorken* or x_b) of the total four-momentum of the proton, with a probability $f_i(x_i)$ given by the so-called *Parton Distribution Functions* (PDF) (more in Sec. 2.2.2). The nQPM predicts that the interaction probability between the electron and a quark inside the proton only depends on x_b and not on the virtuality Q^2 carried by the probe. This property, known as the Bjorken scaling, was later shown to be violated at the EMC [25] and HERA experiments [26,27]. The scaling violation is due to the existence of gluons. As the virtuality of the probe increases, its spatial and temporal resolution also increase, and the probe becomes able to resolve the vacuum fluctuations inside the proton. This means that a probe with large virtuality Q^2 will see a quark sharing its momentum with radiated gluons. Gluons can also fluctuate into $q\bar{q}$ pairs and further interact. As a consequence, the total four-momentum of the proton is distributed over more constituents and the PDFs become dependent on the scale Q^2 at which the interaction occurs.

The inclusion of gluons and $q\bar{q}$ pairs into the description of the internal structure of hadrons, in addition to the valence quarks already introduced by the nQPM, is known as the *Improved Quark Parton Model* (iQPM). The gluon and quark fluctuations inside the hadrons are called “gluons and quarks of the sea”.

2.2.2 Parton Distribution Functions

As mentioned in Sec. 2.2.1, the hadrons are internally composed by many partons. They include the valence quarks, and the gluons and quarks of the sea. Because of momentum conservation, all these internal partons have to share some fraction of the total momentum of the parent hadron. The analytical description of the internal structure of hadrons is done with the PDFs. At leading order, the PDF is the probability density to find a specific kind of parton with a given momentum fraction inside the hadron. Figure 2.6 shows the PDFs of different partons inside the proton as a function of the momentum fraction x at a value of the virtuality $Q^2 = 10 \text{ GeV}^2$. The measurements were obtained by the ZEUS and H1 experiments of the HERA collider at DESY.

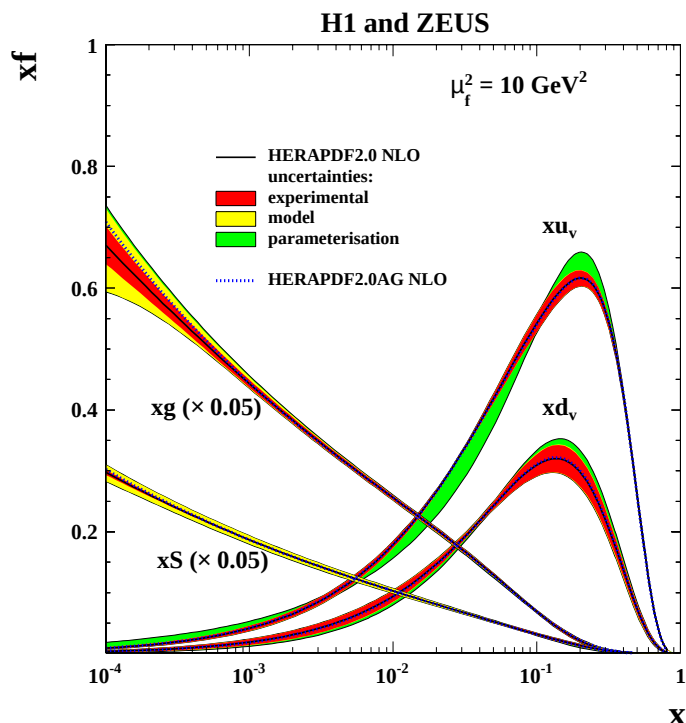


Figure 2.6: The parton distribution functions xu_v , xd_v , $xS = 2x(\bar{u} + \bar{d})$ and xg of HERA-PDF2.0NLO at $\mu_f^2 = 10 \text{ GeV}^2$ with $Q_{min}^2 = 3.5 \text{ GeV}^2$ [28].

A crucial feature of the internal structure of hadrons is that the amount of partons depends on the energy at which they are probed. Two main effects are seen. The first one corresponds to the increase of the low- x parton densities with increasing energy. As the energy at which the hadron is probed increases, the spatial resolution also increases. By doing this, the probe is able to interact with short-distance gluon emissions coming from the valence quarks. Furthermore, the emitted gluons can split into $q\bar{q}$ pairs. Each time there are emissions and splittings, the resulting partons have to share the momentum of the original parton. This process increases the amount of low- x partons at very short-distances or high energies.

The second effect is that the amount of large- x partons, dominated by valence quarks, decreases as the interaction energy increases. This is another manifestation of the same process explained above. As the energy of the interaction increases, the valence quarks are probed at shorter distances. This increases the probability of the quark to have radiated a gluon just before the interaction with the probe, effectively measuring a lower density of partons at large- x . Figure 2.7 shows the reduced cross section for neutral current interaction in e^+p and e^-p collisions at HERA in combination with measurements from fixed-target experiments [29,30]. The reduced cross section is proportional to the structure function F_2 , which is directly related to the quark PDFs of the proton. The Bjorken scaling is observed at medium and large- x for medium Q^2 values, and is violated when going to smaller values of x .

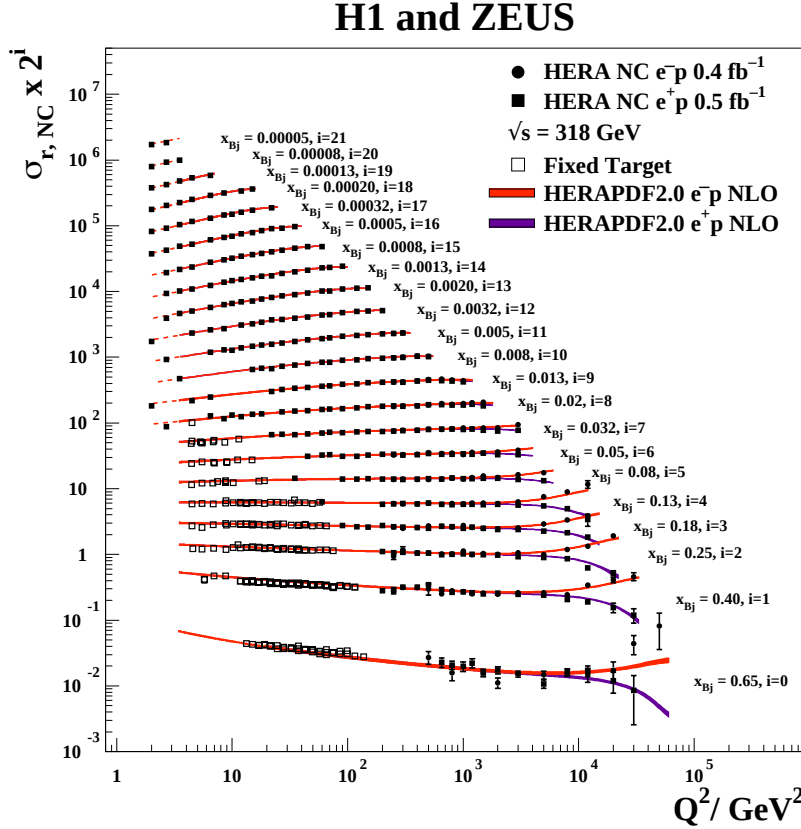


Figure 2.7: Proton structure function as a function of the energy scale Q^2 for multiple x values. Data points correspond to the combined HERA data for the inclusive NC e^+p and e^-p [28].

The PDFs are not accessible within pQCD, and need therefore to be extracted from measurements like the ones shown in figure 2.7. The densities of partons are experimentally extracted in processes which are: i) sensitive to one or more type of partons, ii) at a specific energy scale Q^2 , and iii) at a specific value of x . This means that the PDFs are only extracted

in a limited phase-space in a discrete manner. An extrapolation is therefore needed to obtain these functions and be able to make predictions at different values of Q^2 and x . There are currently three complementary formalisms to carry out the extrapolation of PDFs, usually referred to as *evolution equations*:

- **DGLAP** Named after its authors Dokshitzer, Gribov, Lipatov, Altarelli and Parisi (DGLAP) [31–33]. It makes possible to evolve parton densities from an initial relatively small energy scale Q_0^2 and x not too small, to a higher scale $Q^2 > Q_0^2$ by resumming terms of the form $(\alpha_s \ln(Q^2/Q_0^2))^n$. It is also referred as a Leading Log Approximation on Q^2 (LL_QA). Within this formalism, successive parton emissions come with higher transverse momentum as the evolution is carried out from the proton to the hard scattering. This type of evolution is said to be ordered in transverse momentum. The DGLAP evolution equation is able to successfully describe the available data for not too small values of x , but is expected to break at very low- x values for which gluon splittings start to dominate and eventually recombination effects occur.
- **BFKL** Introduced by Balitsky, Fadin, Kuraev, and Lipatov (BFKL) [34,35]. It evolves unintegrated PDFs (uPDF) at fixed virtuality Q^2 from a high value of the momentum fraction x_0 to smaller values $x < x_0$, by resumming terms of the form $(\alpha_s \ln(1/x))^n$. The uPDFs $\mathcal{G}(x, k_T^2)$ depend explicitly on the transverse momentum of the partons and are related to the usual PDF via

$$xg(x, Q^2) \cong \int^{Q^2} \frac{dk_T^2}{k_T^2} \mathcal{G}(x, k_T^2). \quad (2.9)$$

The scale of the evolution Q^2 is of the order of the hard scattering scale and is kept unchanged, while the longitudinal momentum fraction x is evolved, leading to an ordering in x , $x_0^2 \gg x_1^2 \gg \dots \gg x_n^2$.

In this formalism, the radiated partons are not restricted in transverse momentum and can take any kinematically allowed value. The unrestricted values in transverse momentum make them non-negligible with respect to the longitudinal momenta of the partons, and have therefore to be taken into account.

The BFKL evolution equations are also referred to as a Leading Log Approximation on x (LL_xA).

- **CCFM** The Ciafaloni-Catani-Fiorani-Marchesini evolution equation (CCFM) [36,37] treats the parton evolution in both the large x and small x regions. It combines the DGLAP and BFKL formalisms by applying an angular ordering which ensures at large x a treatment like the former and at small x a treatment equivalent to the latter.

The imposed angular ordering implemented in the CCFM evolution equations ensures energy-momentum conservation, making this formalism suited for simulations with Monte Carlo event generators (see Sec. 2.3).

Figure 2.8 shows the $\ln(Q^2) - \ln(1/x)$ plane indicating the phase space regions where each of the different evolution equations are valid.

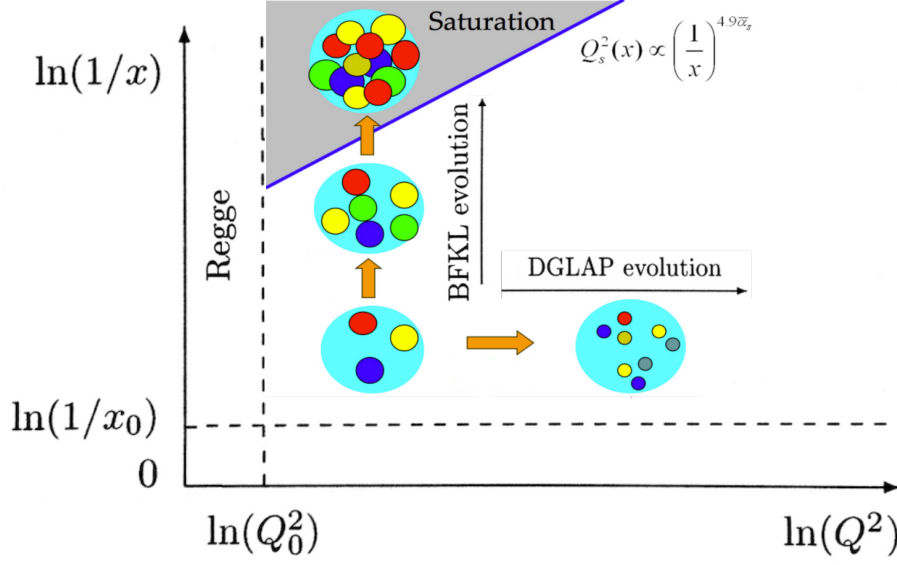


Figure 2.8: The $\ln(Q^2) - \ln(1/x)$ plane indicating the phase space regions where each of the different evolution equations are valid. Also shown are the regions in which Regge phenomenology applies (see Sec.2.2.5) and where the saturation/recombination effects have to be taken into account. Based on [38] and [39].

2.2.3 Soft and hard interactions

The hadron-hadron collisions are often divided into two broad categories: soft interactions and hard interactions. The soft interactions are defined as those in which the momentum transfer between the interacting hadrons is small compared to the centre-of-mass energy ($\sqrt{s}$) of the collision. This kind of interactions are sensitive to long-distance effects and are sensitive to the hadron as a coherent whole. The total, elastic and single/double diffractive cross sections are among the soft processes. In Sec. 2.2.5 the single/double diffraction are addressed in more detail.

Hard interactions are the interactions in which a large momentum transfer is present. In hard interactions the energy of the exchanged particle is large enough to probe the internal structure of the hadron. As a consequence, these kinds of events contain either high transverse momentum objects, or high-mass objects, or a combination of both. Hard interactions are calculable using pQCD due to the presence of a high energy scale that makes $\alpha_s < 1$.

Most of the scattering processes in a collider experiment fall in the soft regime with typical cross section values of $\mathcal{O}(10 \text{ mb})$, while the cross section for events where a hard scale is present are of $\mathcal{O}(10 \text{ } \mu\text{b})$.

In the following sections these two kind of interactions are discussed.

Hard interactions

The time scale at which an interaction with a large momentum transfer occurs is much shorter than the time needed by the partons inside the hadron to interact with each other. This difference of time scales makes it possible to separate the description of collisions with hadrons into two kinds of processes: the short-range processes and the long-range processes. In the former, a process-dependent short-ranged hard process involving partons is calculated. In the latter, a process-independent long-ranged description of the hadron through the use of PDFs is performed.

This separation into short- and long-range processes is commonly known as the *Factorisation theorem*. Figure 2.9 shows a schematic diagram representing the factorised parts on a hadron-hadron scattering calculation

$$d\sigma(h_1 h_2 \rightarrow cd) = \int_0^1 dx_1 dx_2 \sum_{a,b} f_{a/h_1}(x_1, \mu_F^2) f_{b/h_2}(x_2, \mu_F^2) d\hat{\sigma}^{ab \rightarrow cd}(Q^2, \mu_F^2). \quad (2.10)$$

The function f_{a/h_1} (f_{b/h_2}) represents the PDF of a parton of type $a(b) \in \{q, \bar{q}, g\}$ inside the interacting hadron h_1 (h_2). The momenta carried by the interacting partons are given by $p_a^\mu = x_1 p_{h_1}^\mu$ and $p_b^\mu = x_2 p_{h_2}^\mu$. In general, $x_1 \neq x_2$, which gives a boost along the beam axis in the direction of the lower x by $\beta = (x_1 - x_2)/(x_1 + x_2)$. The sum is performed over all the subprocesses which contribute to the cd final state. The scale μ_F is known as the factorisation scale, the value at which the short- and long-range processes are separated. The value of μ_F is usually set to values close to the scale of the hard subprocess Q .

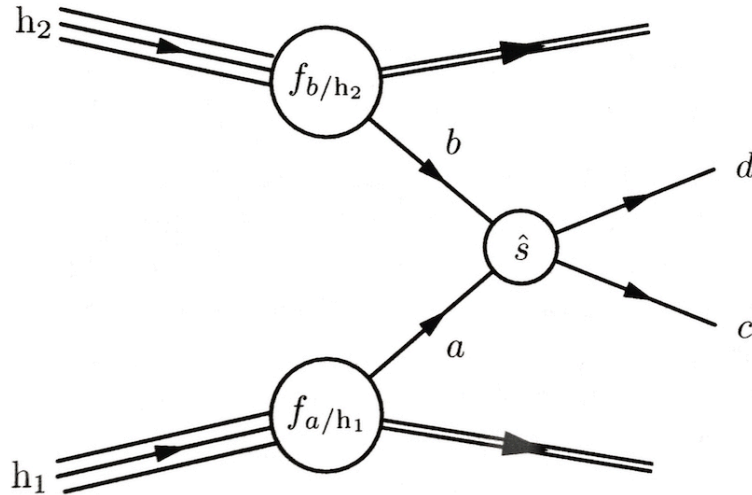


Figure 2.9: Schematic diagram representing the factorisation of a hadron-hadron scattering into process-independent PDFs, $f_{x/h}$, and a process-dependent hard-scattering subprocess between partons a and b . As a result, two partons d and c are produced [38].

It is important to mention that the factorisation theorem is only mathematically proven for few processes, e.g. electron-proton Deep Inelastic Scattering (DIS) and hadron-hadron

Drell-Yan processes [40]. Even if the factorisation theorem has not been proved for all kind of processes, it still gives an excellent description of the experimental data.

The power of the factorisation theorem resides in the independence of the PDFs from the hard subprocess, which means that once the PDFs have been measured and evolved by the evolution equations, they can be used in any other kind of scattering process involving the same initial hadrons.

Soft interactions

As mentioned above, these processes are characterised by a low momentum transfer with respect to the centre-of-mass energy. For very low values of $|t| < 1 \text{ GeV}^2$, the soft interactions are dominated by the elastic scattering. In the elastic scattering, the interacting hadrons remain intact without excitation of any internal degrees of freedom.

In addition to the elastic scattering, a new kind of reactions take place. These reactions are known as *Diffraction dissociation* processes and are a key concept for this thesis. The discussion of diffraction is developed on its own in Sec 2.2.5.

It is important to stress that soft interactions are characterised by small momentum transfers with respect to the centre-of-mass energy ($\sqrt{s}$). In general, the low momentum transfer involved in these kinds of interactions leaves pQCD out of reach and more phenomenological models have to be implemented. However, for large enough $\sqrt{s}$ values, a large scale process can occur with $\mu \gg \Lambda_{QCD}$, making it possible to implement pQCD. This case is denominated *hard diffraction* and it only represents a tiny fraction of the diffractive cross section.

The phenomenological models used to describe soft interactions have multiple free parameters which cannot be inferred from the theory. These parameters are estimated from the measurement of observables sensitive to the soft region. The parameter estimation is performed by fitting multiple observables simultaneously, a process called *tuning* (see Sec 2.3). Hence, this thesis focuses on the study of the soft region and is complemented with a discussion on the transition region between the soft and the hard components of the particle production.

2.2.4 Multi parton interactions

The description of hadrons by the iQPM as a collection of free quarks and gluons leads to the possibility of having more than a single parton-parton interaction within the same hadron-hadron collision. This feature of hadronic collisions is known as *Multi Parton Interactions* (MPI). As higher collision energies become accessible in different high-energy collision facilities, partons with successively lower momentum fraction x , which are more and more abundant, can be probed. As a consequence, events with softer particles in the final state coming from MPI are enhanced.

For instance, the QCD $2 \rightarrow 2$ partonic cross section is of special interest to study the transition region between the soft and the hard regimes. This cross section is dominated by

the t -channel gluon exchange with the following behaviour

$$\sigma_{2 \rightarrow 2}(p_{T_{min}}) \propto \int_{p_{T_{min}}^2}^{s/4} \frac{d\hat{\sigma}}{dp_T^2} dp_T^2 \propto \frac{1}{p_{T_{min}}^2}, \quad (2.11)$$

which is divergent for $p_{T_{min}} \rightarrow 0$. At a $\sqrt{s} = 8$ TeV, this partonic cross section exceeds the inelastic cross section at transverse momentum values around 4-5 GeV [41], which is still in the perturbative region and no breakdown of perturbation theory can be claimed. Even though there are different approaches to solve this paradox, all of them share a close relation with MPI. In Sec 2.3.1 the most representative approaches to solve this problem are discussed.

2.2.5 Diffractive events

Soft diffractive collisions have been successfully described by phenomenological models based on Regge theory. In this theory, the t -channel exchange is interpreted as the exchange of so-called *Regge trajectories*. These trajectories are parametrised as $\alpha(t) = \alpha(0) + \alpha' t$, where $\alpha(0)$ is the intercept with the y -axis and α' is an almost universal slope ($\alpha' \approx 1 \text{ GeV}^{-2}$).

The parameters of some trajectories describing soft hadron-hadron interactions can be found by fitting the spins and masses of real mesons and baryons. Soft diffraction is characterised by the exchange of a specific trajectory, the *Pomeron trajectory*, which contains a single object called the *Pomeron*. The Pomeron is a color neutral object with the quantum numbers of the vacuum, i.e., the characteristics needed to describe the observed distinctive signatures of diffractive events in hadron-hadron collisions. Two main features are observed in diffractive events:

- none, one or both of the interacting hadrons might survive or be scattered into a low-mass system. In any case, the outgoing hadrons or low-mass systems have approximately equal energy to that of the incoming hadrons and the same quantum numbers. Figure 2.10 shows a schematic representation of the four main components of a hadronic inelastic scattering: non-diffractive, single diffractive, double diffractive and central diffractive.
- at least one large rapidity region Δy devoid of hadronic activity is present, the so-called large rapidity gap (LRG). The rapidity y is defined as $y = (1/2) \cdot \ln[(E + p_z)/(E - p_z)]$, where E and p_z are the energy and longitudinal momentum of the final state particle, respectively.

The contribution of a trajectory to the total inelastic cross section is given by $\sigma_{tot}(s) \propto s^{\alpha-1}$, where s is the centre-of-mass energy squared of the collision and α the intercept of the trajectory.

The parameters of the Pomeron trajectory have been estimated from fits to a wide range of data to be $\alpha(0) \approx 1.08$ [42] and $\alpha' \approx 0.251 \text{ GeV}^{-2}$ [43].

Fig 2.8 shows the kinematic region where the Regge theory is used to describe hadronic collisions. The energy scale in this region is too small to apply pQCD, hence the need to use phenomenological models. When a hard scale is present, the description of (hard)

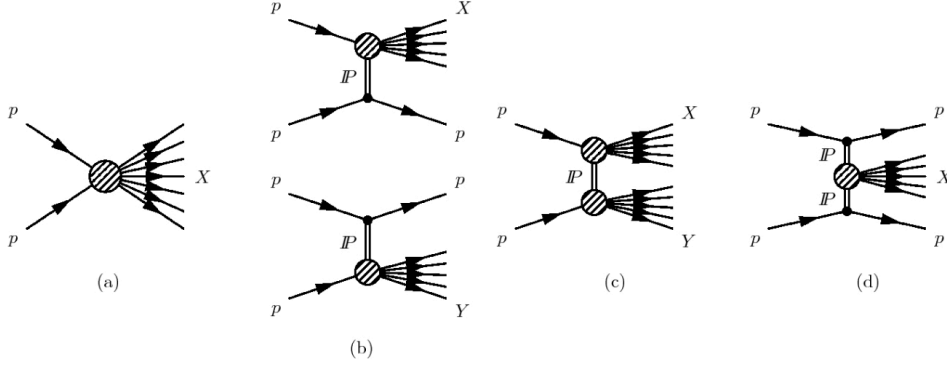


Figure 2.10: Schematic representation of the four main components of hadronic inelastic collisions. From left to right: non-diffractive, single diffractive, double diffractive and central diffractive components. Single and double diffractive components are related with the exchange of a single Pomeron, while central diffractive is characterised by a double Pomeron exchange.

diffraction in terms of quarks and gluons by pQCD is better understood. Diffraction, both in soft and hard regimes, is a powerful tool to investigate the low- x partons inside the proton.

The idea of diffractive PDFs (dPDF) in the DIS experiments was developed for hard diffraction. In diffractive DIS (DDIS), two levels of factorisation are implemented, following an analogous process to the factorisation scheme explained in section 2.2.3, where a cross section is calculated by the convolution of a process independent non-perturbative PDF and a process dependent perturbative partonic cross section. The first level in the DDIS factorisation procedure is the requirement of the diffractive structure function F_2^D to satisfy collinear factorisation:

$$F_2^D = \sum_{a=q,g} C_2^a \otimes f_{a/p}^D, \quad (2.12)$$

where the coefficient C_2 is the perturbative calculable partonic cross section of interest, and $f_{a/p}^D$ are the dPDFs that satisfy the same evolution equations as the inclusive PDFs. The second level of factorisation is the so-called *Regge factorisation* and it is assumed to be [20]:

$$f_{a/p}^D(x_{\mathbb{P}}, t, z, \mu^2) = f_{\mathbb{P}/p}(x_{\mathbb{P}}, t) f_{a/\mathbb{P}}(z, \mu^2), \quad (2.13)$$

where $f_{a/\mathbb{P}}$ are the parton densities of the Pomeron, and $z \in [x/x_{\mathbb{P}}, 1]$ is the momentum fraction of the Pomeron carried by the parton entering the hard subprocess. The Pomeron flux factor $f_{\mathbb{P}/p}(x_{\mathbb{P}}, t)$ is a phenomenological factor calculated using the Regge theory. The Regge factorisation is a convenient ansatz that is not strictly necessary but simplifies the interpretation and calculations.

The DDIS factorisation has been proven to be successful in describing DDIS processes [44]. However, the factorisation cannot be applied to hard diffraction processes happening in hadron-hadron collisions [44, 45]. In figure 2.11 an attempt to use the DPDFs extracted

in DDIS to describe diffraction in $p\bar{p}$ as measured by the CDF experiment is shown. It can be seen, that there is a clear overestimation of the predictions by a factor of 3 up to a factor of 10.

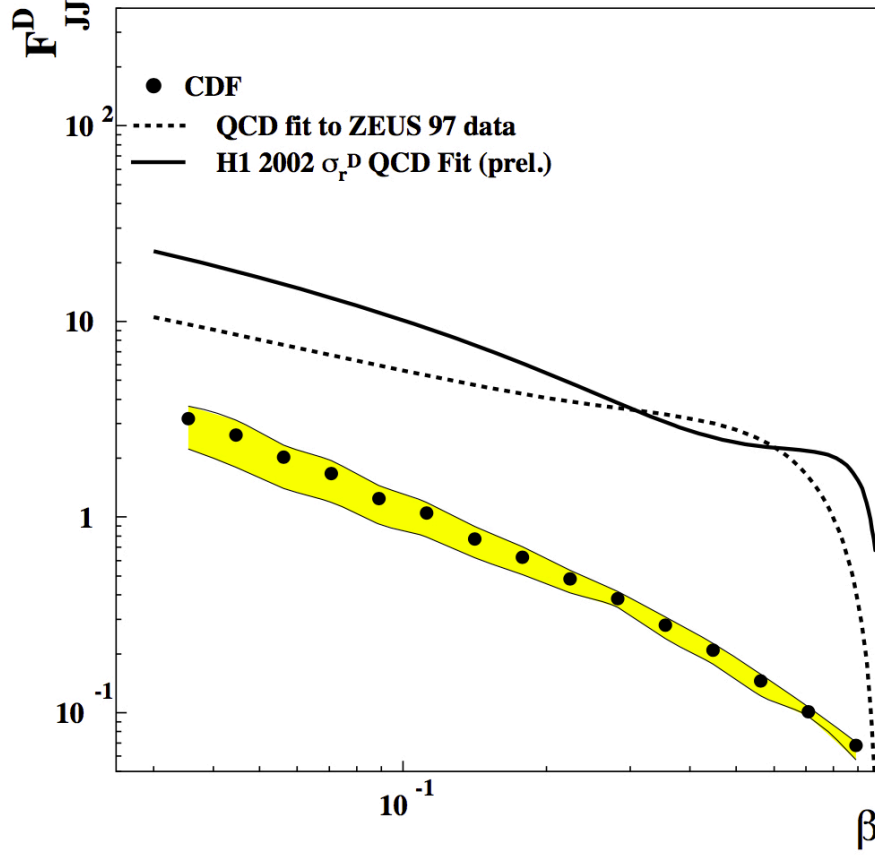


Figure 2.11: CDF results for the cross sections of diffractive dijet production with a leading antiproton in $p\bar{p}$ collisions [46], compared with the prediction obtained from the diffractive PDFs extracted at HERA [47]. Source [48]

The reason of the factorisation breaking is the presence of MPI. The additional interactions stemming from the MPI contribute to lower the expected number of diffractive events due to the creation of particles which fill the would-be rapidity gap. This phenomenon gave rise to the idea of *rapidity gap survival probability* as a method to calculate hard diffractive cross sections.

Figure 2.12 shows three diagrams for different MPI contributions in hadron-hadron collisions. Figure 2.12a shows two simultaneous hard interaction chains between the two incident protons. Figure 2.12b shows two hard interaction chains, the left chain corresponding to a diffractive interaction where the hard scattering occurs between the two exchanged Pomerons. The right chain corresponds to a hard interaction between the two incident protons, while figure 2.12c shows a single diffractive interaction where two gg hard interactions occur between the Pomeron coming from one proton and the other proton.

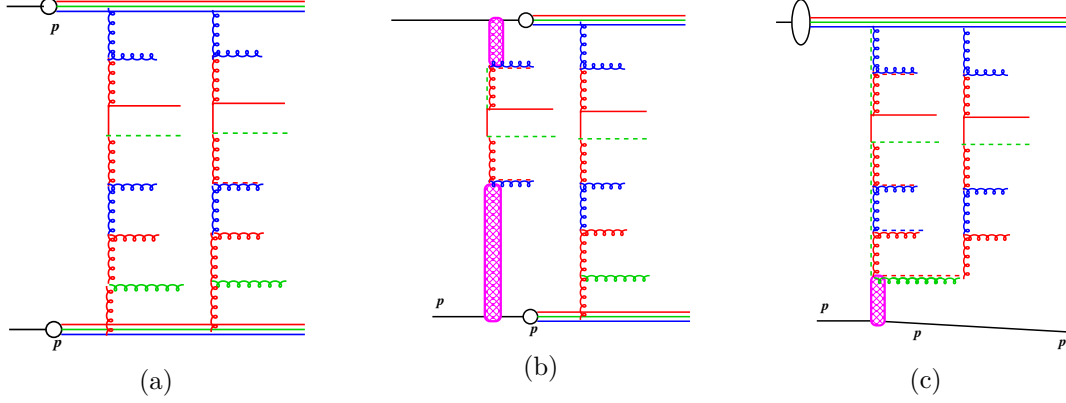


Figure 2.12: Three different MPI contributions in pp scattering. (a) Two gg hard scatterings between the same incident protons. (b) Two gg hard scatterings, the left chain corresponds to a central diffractive interaction (double Pomeron exchange) where the two Pomerons exchange a gluon each one. The right chain correspond to a gg hard interaction between the protons. (c) Two gg hard interaction between the Pomeron coming from one proton and the other incident proton.

2.3 Monte Carlo event generators

Monte Carlo (MC) event generators are an essential tool to perform experimental measurements as well as to make theoretical predictions. They are built to describe the physics of particle interactions from the very high scales, up to the typical low scale of hadron formation and decays. At high scales they are based on perturbation theory, while at low scales where the soft phenomena takes place and no computation from first principles is available, phenomenological models are used instead.

In high energy physics (HEP) there are numerous MC event generators available specialised on different kind of processes. The *general-purpose* MC (GPMC) event generators like PYTHIA [49], HERWIG [50] and SHERPA [51], are of great importance since they are the main tool used to simulate all aspects of a particle collision:

- **Initial and final state radiation** The incoming interacting partons are subject to radiate gluons, due to their color charge, and photons in the case of quarks, which are electrically charged. This kind of radiation is generally known as *parton showers* (PS) and can be calculated with the evolution equations introduced in Sec 2.2.2. The radiation emitted from the partons prior to the hard interaction is called *initial state radiation* or ISR. On the other hand, the radiation emitted from the partons produced in the hard interaction is known as *final state radiation* or FSR. The PS simulation is performed perturbatively from the hard scale of the main interaction of the event, down to a low scale where hadronization happens. (see Sec 2.2.3).
- **Hard process** The main hard interaction calculated with pQCD. The GPMC usually only include leading order (LO) perturbative matrix element calculations for the hard scattering process. In case a higher order pQCD calculation is needed, different options are available, e.g. MADGRAPH [52] for LO and Next-to-LO (NLO) and POWHEG [53,54]

Matrix Element (ME) for NLO calculations. Higher order matrix element calculators need to be interfaced to one of the GPMC to fully simulate the hadron-hadron collision. The interface is not a trivial task and different problems like double-counting have to be sorted out. The details on current approaches can be found elsewhere [55].

- **Multi parton interactions** The MPI is an essential part of hadron-hadron collision simulations. Without the implementation of MPI, the predicted multiplicities for inclusive events cannot reproduce the measurements at all. The implementation of MPI in modern GPMC event generators is addressed in Sec 2.3.1.
- **Hadronization** After generating all the final state partons, either coming directly from the hard scattering or from the PS, the rearrangement of the partons into color-neutral hadrons needs to be performed. This process is known as *hadronization* and it is a non-perturbative procedure for which QCD-inspired phenomenological models are needed. The two main models used nowadays are the *String Model* used by PYTHIA and the *Cluster Model* implemented in HERWIG and SHERPA. In Sec 2.3.2 these two models are discussed in more detail.
- **Decay** Some of the created hadrons in the hadronization step might be resonances or heavy unstable hadrons which have to be further decayed into stable long-lived particles.

Figure 2.13 shows a schematic representation of a hadron-hadron collision where the different components described above are highlighted in different colors.

The non-perturbative components involved in the simulation of high energy particle collisions contain important free parameters which have to be obtained from fits to a large number of measurements. Depending on the nature of the parameter under study, different kinds of observables are used. The MPI and hadronization are among the least understood components of a high energy particle collision and multiple parameters need to be constrained. Some of these parameters are presented in Sec. 2.3.1 and Sec 2.3.2. The process of parameter determination is often called tuning. A special kind of events defined experimentally as Minimum Bias (MB) and Underlying Event (UE) are specially sensitive to the production mechanisms of the majority of the particles in a event. MB events are aimed to capture everything coming from the interaction of two particles as inclusive as experimentally possible, where the restrictions come mainly from detector capabilities like acceptance and efficiencies. In UE the focus is on all the processes that take place in a collision but the hard scattering.

2.3.1 MPI models

As mentioned earlier, the increase of the $2 \rightarrow 2$ QCD partonic cross section for $p_{T,\min} \rightarrow 0$ is closely related to the idea of MPI. One has to notice that this partonic cross section is an inclusive cross section, thus, an event with two parton-parton interactions will count twice in $\sigma_{2 \rightarrow 2}$ but only once in the total hadron-hadron cross section σ_{tot} . From these statements, it is possible to calculate the average number of interaction per hadron-hadron collision above a given $p_{T,\min}$ as

$$\langle n \rangle(p_{T,\min}) = \frac{\sigma_{2 \rightarrow 2}(p_{T,\min})}{\sigma_{tot}}, \quad (2.14)$$

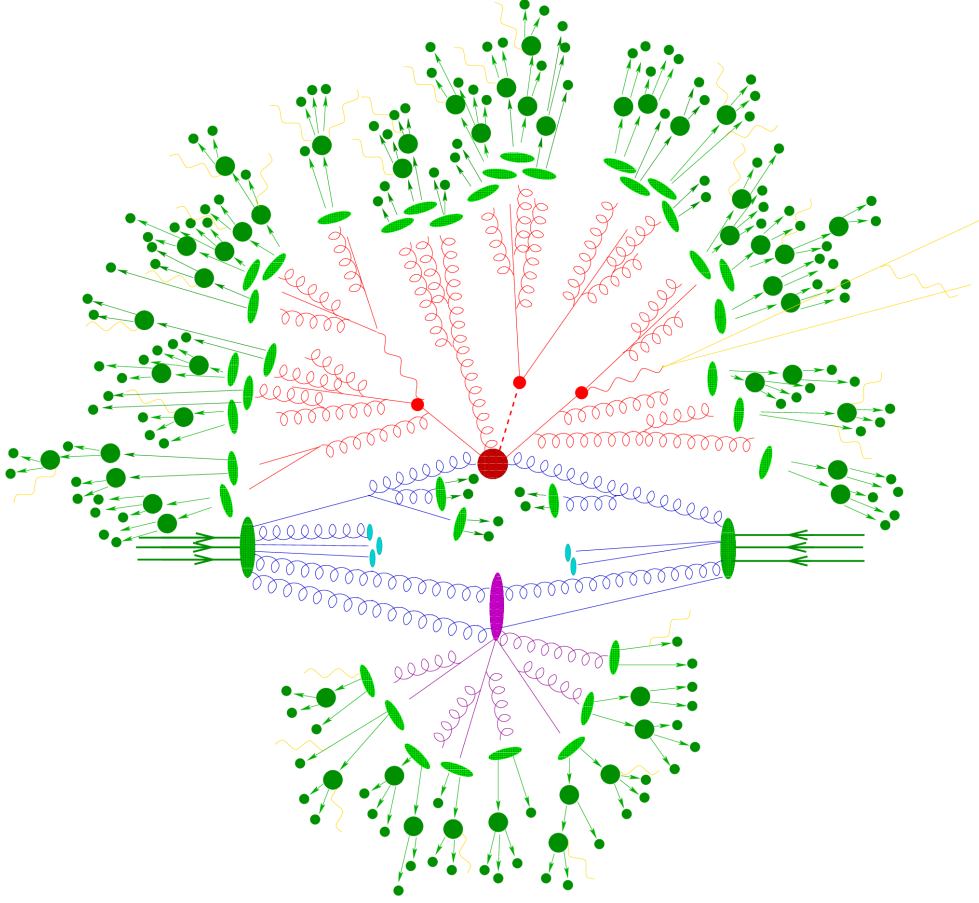


Figure 2.13: Matrix element generators simulate the hard scattering (shown in red). Parton showers produce Bremsstrahlung (blue). Multipartonic interactions (purple). Hadronization of QCD partons (green). Hadron decay packages simulate unstable hadron decay (dark green). QED Bremsstrahlung (yellow). [56]

assuming that all the individual parton-parton interactions are independent from each other and equivalent among themselves. This interpretation makes the total hadronic cross section finite by moving the large jet cross section to the number of interaction per hadron-hadron collision.

Two additional features are needed to solve this issue. First, it has to be explicitly ensured that all multiple interactions in the event fulfil energy-momentum conservation, i.e. the interactions cannot use more momentum than available. By requesting this, the predicted events with large numbers of MPI will be suppressed. In practice, the energy-momentum conservation for MPI simulations is handled in slightly different ways by the GPMC. For instance, in PYTHIA and SHERPA the MPIs are iteratively calculated in order of the hardness of the interaction. In each iteration, the PDFs are rescaled by removing the momentum fraction x taken by the interacting partons of the previous iteration, ensuring the energy-momentum conservation. In HERWIG, the generation of further interactions is allowed until the energy-momentum conservation is violated.

The second feature to fully suppress the rapid grow of the cross section is the reduction of the number of interactions at low- p_T and low- x . This is achieved by including a screening effect, because as the p_T of the exchanged parton becomes smaller (i.e. its wavelength $\sim 1/p_T$ becomes larger), its ability to resolve typical color-anticolor separations also reduces. Hence, its coupling to color charges vanishes in the limit $p_T \rightarrow 0$, inducing an interaction suppression. As also seen in the previous section, the actual implementation in the GPMC comes with some differences. HERWIG implements a simple version by choosing a step function such that the perturbative $2 \rightarrow 2$ partonic cross section completely vanishes below a $p_{T,\min}$ scale. Below the $p_{T,\min}$ scale a set of soft scatterings are added to cover this phase space region. A more elaborated model is implemented in PYTHIA and SHERPA, where instead of a step function, a smoothing factor of the form

$$\left(\frac{\alpha_s(p_{T0}^2 + p_T^2)}{\alpha_s(p_T^2)} \frac{p_T^2}{p_{T0}^2 + p_T^2} \right)^2, \quad (2.15)$$

is included in the $2 \rightarrow 2$ partonic cross section in order to regularize the divergences. $p_{T,\min}$ and p_{T0} are free parameters of the models to be extracted from fits to experimental data.

Furthermore, as the $\sqrt{s}$ of the interaction increases, the density of low- x gluons increases making the screening distance smaller. This means that the parameter p_{T0} has a dependence on the centre-of-mass energy (c.m.e.). In this model, a power-like behaviour inspired in the proton-proton cross section scaling with $\sqrt{s}$ is used

$$p_{T0}(s) = p_{T0}^{Ref} \left(\frac{\sqrt{s}}{\sqrt{s}^{Ref}} \right)^{\epsilon^{Ref}}, \quad (2.16)$$

where all *Ref*(erence) parameters are free and have to be obtained from comparisons with the experiments.

An additional feature of MPI is their dependency on the centrality of the collision [57]. The hadrons are composite particles with many partons inside which are spatially extended. The PDF describes the share of momentum by the partons inside the hadron but says nothing about their spatial distribution. When two hadrons collide, its transversal overlap can vary from very low or null to completely overlapped. This overlap is known as the *impact parameter* b . The amount of MPI is dependent on the choice of the matter distribution function and on the impact parameter. In PYTHIA several different options for the description of the matter distribution inside the hadron are available, such as Gaussians, double Gaussians and exponentials. HERWIG, on the other hand, is based on the electromagnetic form factor. In all the current models a factorisation between the the PDF and the matter distribution is assumed, $f(x, b) = f(x)g(b)$, where x is the momentum fraction and b is the impact parameter.

2.3.2 Hadronization models

The hadronization step on a particle collision simulation relies on phenomenological models. In the following, the two main models implemented in the current GPMC event generators

are presented.

String model

In this model, the color field between a $q\bar{q}$ pair is pictured as a color flux tube connecting both particles. Mathematically this tube can be represented by a string at its axis. As two color-charged quarks move apart, the color flux tube is stretched resulting on the increase of its potential. Linear confinement implies a potential for the string of the form $V(r) = \kappa r$, where $\kappa \sim 1 \text{ GeV/fm} \approx 0.2 \text{ GeV}^2$ is the tension of the string as determined from hadron mass spectroscopy experiments. If the moving q and $\bar{q}$ have enough energy, the string potential between them can grow enough to create a new $q'\bar{q}'$ pair in between the original quarks. Each of these new quarks attaches to each of the originally moving particles by a color string, creating two new systems similar to the starting one. Figure 2.14 shows a schematic representation of this process. If the energy of the new systems is again large enough, the process is repeated until the kinetic energy is too small to break a string.

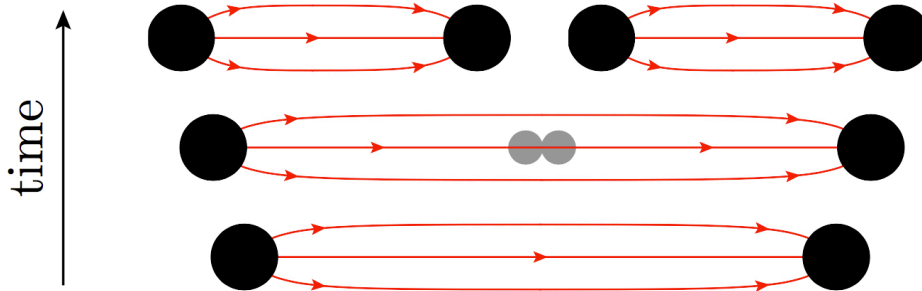


Figure 2.14: Schematic representation of the string breaking by quark pair creation in the string field. [20]

Additionally, baryon production can also be taken into account by producing pairs of diquarks at the string break.

Cluster model

The cluster model is based on the so-called *preconfinement*. Preconfinement is a characteristic of the PS based on the principle that if the color structure of the produced particles is followed, it is possible to find the matching partner in color space for every external parton. Furthermore, the matched pairs tend to end up close in phase space. All these color-singlet pairs form clusters with mesonic quantum numbers which later decay to lighter well-known resonances and stable hadrons.

The cluster model emphasises the perturbative phase of the parton evolution by taking the output of the PS as its starting point. In contrast, the string model emphasises the non-perturbative dynamics of the confinement of partons.

2.4 Minimum Bias events

In Sec. 2.3 the concept of MB events has already been mentioned. MB events are experimentally defined and therefore experiment dependent. The way of selecting an event with the minimum experimental bias as possible depends on the capabilities and acceptance of the experiment in use. The measured event can be further corrected for the deviations introduced by the experiment. However, the corrections for acceptance limitations inevitably induce a bias towards the model used for the extrapolation, known as "model dependent correction". That is why even though the measurements can be corrected for the experimental limitations, these corrections should not extrapolate too much with respect to the reconstructed phase space.

In this thesis, the different observables defined in Sec. 5.3 are measured for MB events and further corrected to the stable particle level as defined in Sec. 5.4. In addition to the MB selection, the events are further categorised into different event samples. As mentioned in Sec. 2.3, all the MC event generators rely on phenomenological models to describe the phenomena unreachable to pQCD. All these models have multiple free parameters which have to be determined by comparing the predictions to the different MB and UE measurements. The diffractive-enhanced event samples open the possibility to study the impact of these parameters on the description of different diffractive processes. This allows for a better constraint of the models, giving insights on the true nature of the processes behind the phenomenological models.

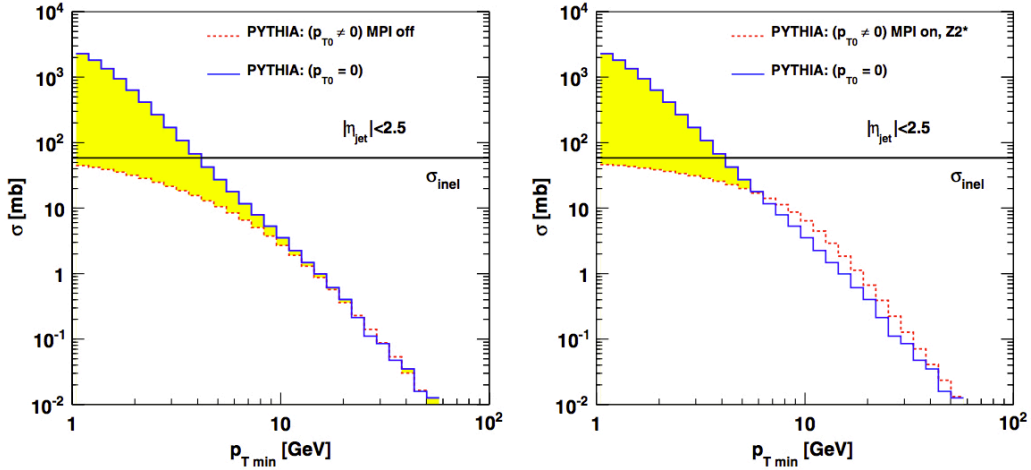


Figure 2.15: The $2 \rightarrow 2$ partonic QCD cross-section as a function of $p_{T,min}$ as predicted by PYTHIA in the range $|\eta| < 2.5$. The solid line shows the prediction applying $p_{T0} = 0$ including parton shower and hadronization, while the dashed line shows the prediction with $p_{T0} \neq 0$, shown without multi-parton interactions (left) and including multi-parton interactions (right). [58]

One of the most important parameters is the p_{T0} (see Sec. 2.3.1), which is intimately related to the amount of produced MPI in a hadron-hadron collision. Also important is the

transition between the phenomenological and perturbative description of diffractive processes (see Sec. 2.2.5), which is regularised in the MC event generator PYTHIA8 by the mass of the diffractively produced system as well as by the p_{T0} parameter. An extensive discussion on these studies is performed in chapter 10.

The motivation to study the effects of p_{T0} in the taming of the $2 \rightarrow 2$ partonic QCD cross-section at low- p_T values has been presented in [58]. In figure 2.15 the effects of the $2 \rightarrow 2$ partonic cross-section at low- p_T values compared to the results when the cross-section is tamed, are shown. Both scenarios are presented for the case when MPI is turned off (left plot) and for when MPI is included (right plot). The inclusion of MPI only shifts the distribution towards larger p_T values but it is not fixing the large growth in the very low- p_T region. It is the taming of the cross-section, which is dependent on the p_{T0} parameter, the responsible to cure this unphysical behaviour. This phenomenological solution has been extensively tested, e.g. in the measurements of the leading charged-particle and leading jet transverse-momentum distributions integrated as a function of $p_{T,\min}$ by the CMS collaboration at $\sqrt{s} = 8 \text{ TeV}$ [41]. Even though this phenomenological solution provides a great improvement of the description of the measurements, it is far from being perfect [41]. However, the ability of a phenomenological model to describe the measurements is only one step towards the understanding of the true underlying nature of the low- p_T effects, i.e. it does not provide a full understanding of the physics undergoing the process.

CHAPTER

3

EXPERIMENT

Contents

3.1	The Large Hadron Collider	43
3.2	The CMS experiment	45
3.2.1	Trigger system	47
3.2.2	Silicon tracker	47
3.2.3	Hadronic Forward calorimeter	49

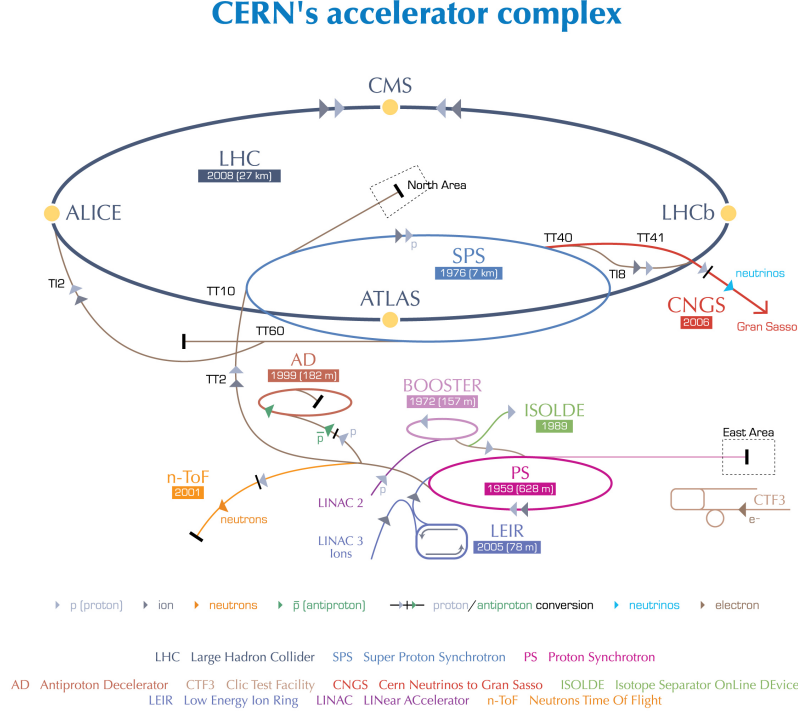
Since the beginning of the 20th century there has been a revolution in the way the matter is studied in order to find its most elemental building blocks. From fix target experiments to bubble chambers, and finally to the nowadays particle colliders, a great number of techniques have been developed to measure the properties of the particles resulting from the studied interactions. In this chapter a description of the Large Hadron Collider (LHC) and the Compact Muon Solenoid (CMS) is presented, putting special attention on the subdetectors relevant for the measurement presented in this thesis.

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is located near Geneva, crossing borders between Switzerland and France. It is operated by the European Organization of Nuclear Research (CERN). It is installed in a 27 km long circular tunnel, that hosted the former Large Electron-Positron Collider (LEP) from 1989 to 2000.

The LHC is a superconducting hadron-accelerator and collider composed of two rings, where two independent proton-proton, proton-lead or lead-lead beams circulate in opposite directions. The particle beams are supplied by the already existing accelerator complex at

CERN. For the proton beams the injection process consists of four steps: the protons are subsequently injected into the Linac2, Proton Synchrotron Booster (PSB), Proton Synchrotron (PS) and Super Proton Synchrotron (SPS), acquiring an energy increase from the initial 50 MeV to 1.4 GeV, 25 GeV and 450 GeV in each step. Finally the LHC gives the last boost to the proton beams from the injection energy of 450 GeV to the desired collision energy up to 7 TeV. Shown in figure 3.1 the accelerator complex at CERN used to inject beams into the LHC.



European Organization for Nuclear Research | Organisation européenne pour la recherche nucléaire

© CERN 2008

Figure 3.1: Accelerator complex at CERN, some of which are used to inject particle beams to the LHC accelerator [59].

Each of the proton beams is designed to provide 2808 bunches containing 1.15×10^{11} protons each.

Two of the key features that determine a particle accelerator are its c.m.e.¹ and its instantaneous luminosity. In the case of the LHC, the nominal value of the c.m.e is $\sqrt{s} = 14$ TeV with an instantaneous luminosity of $10^{34} \text{ cm}^{-2} \text{ sec}^{-1}$. The latter is defined as:

$$\mathcal{L} = f \cdot \frac{k_B N_p^2}{a} \quad (3.1)$$

where k_B denotes the number of bunches per beam, N_p the number of protons per bunch and f the bunch revolution frequency. The factor a describes the geometric size of the bunch

¹As defined in Chapter 2.

in the plane transverse to the proton beam direction [60]. The design luminosity leads to around 10^9 proton-proton interactions per second.

During the first period of the LHC (referred to as Run1) which lasted from 2010 until 2012, the machine operated at a centre-of-mass energy of $\sqrt{s} = 7$ TeV and 8 TeV for pp collisions, and after the first programmed long upgrade period, the so called *Long Shutdown 1* that lasted from 2013 to 2015, the current second period (Run2) started with a centre-of-mass energy of $\sqrt{s} = 13$ TeV.

The LHC accelerator provides collisions at four points along its circumference where six different experiments are installed.

- ALICE (A Large Ion Collider Experiment) is a heavy-ion detector that focuses on QCD, the strong-interaction sector of the SM. It is designed to address the physics of strongly interacting matter and the quark-gluon plasma at extreme values of energy density and temperature in nucleus-nucleus collisions [61].
- ATLAS (A Toroidal Lhc ApparatuS) is a general-purpose detector designed for the discovery of new physics and precise measurement in the SM. Along with the CMS detector, one of its main subjects of study is the Higgs boson sector. It is the largest detector for high-energy physics ever constructed with 46 m long and 25 m of diameter [62].
- CMS (Compact Muon Solenoid) is the second of the general-purpose detectors at the LHC. More information in section 3.2.
- LHCb (LHC beauty) is an experiment dedicated to heavy flavour physics. Its primary goal is to look for indirect evidence of new physics in CP violation and rare decays of beauty and charm hadrons [63].
- LHCf (LHC forward) is an experiment dedicated to the measurement of neutral particles emitted in the very forward region $\eta > 8.4$. The physics goal is to provide data for calibrating the hadron interaction models that are used in the study of Extremely High-Energy Cosmic-Rays [64]. It is located at 140 m either side of the ATLAS collision point.
- TOTEM (TOTal, Elastic and diffractive cross-section Measurement) is an experiment dedicated to measure the total pp cross-section with an luminosity-independent method and study elastic and diffractive scattering [65]. It is located at 147 m either side of the CMS collision point.

3.2 The CMS experiment

The CMS experiment [66] is located at the interaction point 5 near the french village of Cessy. With a weight of 14 000 t and dimensions of 21.6 m long by 14.6 m on diameter, it is one of the biggest experiments ever constructed. The CMS experiment uses a right-handed coordinate system, with the origin at the nominal interaction point (IP), the x -axis pointing to the centre of the LHC ring, the y -axis pointing upward, and the z -axis along the anticlockwise beam-direction. The azimuthal angle φ is measured in the (x, y) plane, where $\varphi = 0$

is the positive x and $\varphi = \pi/2$ is the positive y direction. The polar angle θ is defined with respect to the z -axis, where $\theta = \pi/2$ represents the perpendicular direction to the axis and $\theta = 0$ points along the negative z direction.

A quantity often used in high-energy physics is the pseudorapidity (η), it is defined as:

$$\eta = -\ln \tan(\theta/2). \quad (3.2)$$

In the relativistic limit, where a particle travels close to the speed of light and its mass can be neglected, the pseudorapidity becomes equivalent to the rapidity. This is convenient since the rapidity is a lorentz-invariant quantity but the calculation of the pseudorapidity remains related to the polar angle, which makes its visualisation more intuitive.

The central feature of the CMS detector is a superconducting solenoid of 6 m internal diameter, that provides a uniform magnetic field of 3.8 T along the beam axis. Inside the magnet volume the silicon-pixel tracker, the silicon-strip tracker, the lead-tungstate electromagnetic calorimeter and the brass/scintillator hadron calorimeter are located. Outside the magnet volume, the second part of hadronic calorimeters and the muon system consisting of gaseous chambers are located. Figure 3.2 shows a schematic view of the CMS detector.

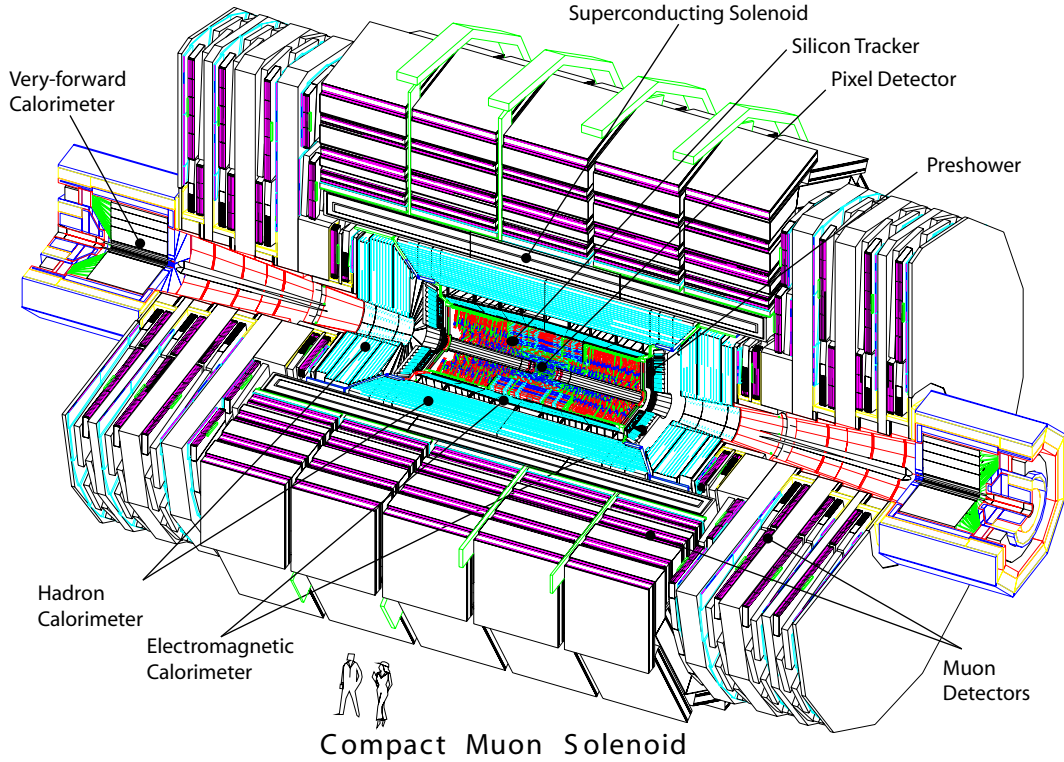


Figure 3.2: Schematic view of the CMS detector showing the different components.

For this analysis, the silicon tracker detector and the hadronic forward calorimeter (HF) are used. Therefore a detail description is given in the following.

3.2.1 Trigger system

The trigger system of the CMS experiment is responsible for selecting events of potential physics interest among the vast amount of collisions delivered by the LHC. There are two important reasons for that. The rate at which the collisions are delivered by the LHC machine is in GHz, that is $\mathcal{O}(10^6)$ larger than the storage capabilities. And, not all the collisions are of interest from the physics point of view.

The trigger system is divided in two stages: the Level-1 Trigger (L1T) and the High-Level Trigger (HLT). The L1T is based on pure hardware signals that have to fulfil certain conditions in order to select the event. Those conditions are known as trigger menus. After the L1T, the event rate is decreased to the order of kHz. The HLT takes the output of the L1T and process it further in a software basis. It makes use a computer farm just next to the experiment with more than 10000 CPU cores [67]. The HLT performs a fast reconstruction of physics objects, like muons, electrons, photons, etc. and identification criteria are applied to select events of potential physics interest. In Sec. 3.2.2, are addressed some of the requirements of the CMS tracking system to be part of the HLT. After this stage the rate is decreased from 100 kHz to 400 Hz, consistent with an archival storage capability of $\mathcal{O}(100)$ MB/s.

Beam Pickup Timing for the eXperiment

The Beam Pickup Timing for the eXperiment (BPTX) [68], is composed of two detectors located at 175 m either side of the interaction point. It provides timing from opposite beams entering the interaction region, from which precise trigger information and z-vertex position are obtained. This device is used to provide what is called a ZeroBias trigger as is explained in Sec. 5.3.1.

3.2.2 Silicon tracker

The silicon tracker of the CMS experiment is one of its key features, since it is the responsible to reconstruct the vertices and trajectories of all charged particles in the collision.

At design energy and luminosity conditions the LHC is expected to create collision environments with approximately 40 proton-proton interactions per bunch crossing with an average of 50 charged particles each one. This large amount of collisions creates an extreme environment with over 1000 charged particles emerging from the IP. As mentioned in Sec. 3.2.1, the silicon tracker is one of the subdetectors of the CMS experiment that contributes to the HLT decision. For that, the response of the tracker has to be fast and efficient while keeping the fake rate low. In order to avoid big energy losses and ionisation of the particles coming from the collision when crossing the detector, the material budget of the detector needs to be as small as possible to still successfully measure the energy and direction of the particles.

The tracking system is the closest subdetector to the IP. It has a cylindrical shape of 5.8 m long and a diameter of 2.5 m. It is further divided into two main sections, the silicon pixel and the silicon strip parts. The silicon pixel detector consists of 3 concentric barrel layers

with radii 4.4 cm, 7.3 cm and 10.2 cm called Barrel Pixel (BPIX), and 2 endcap disks called Forward Pixel (FPX) resulting in an acceptance of $|\eta| < 2.5$. The radial region between 20 cm to 116 cm is occupied by the silicon strip detector, and is subdivided in the Inner Barrel, Inner Disks, Outer Barrel and at both sides the End Caps closing the full tracker barrel. Figure 3.3 illustrates the tracking detector.

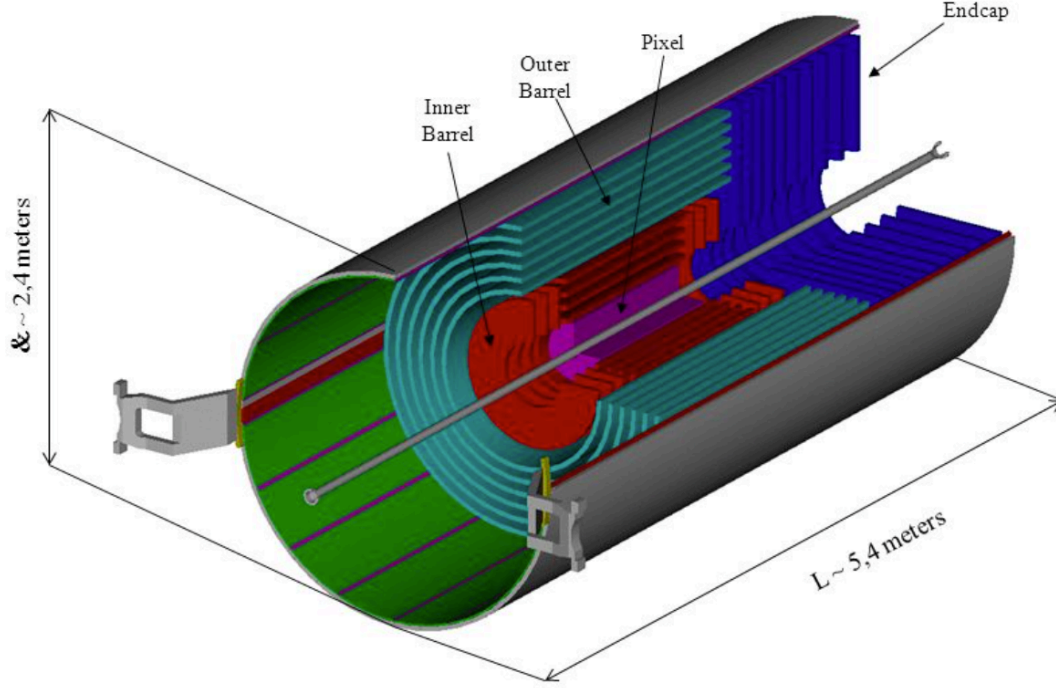


Figure 3.3: Schematic view of the CMS tracker detector. The different components are visible, the Pixel Barrel and the Strip sections, Inner Barrel, Outer Barrel and Endcaps.

These two sections are built from carbo-fibre structures, in which 1440 pixel modules and 15148 strip modules are mounted, giving an active silicon area of 200 m^2 . The reason to have two kinds of detector units is that the pixels provide a 2-dimensional spacial measurement in their plane and are designed to identify tracks very close to the IP, where the track density is high. As the particles move towards the exterior layers of the tracker, the particle density decreases and 1-dimensional silicon modules are sufficient to provide a reliable measurement.

From vertices and particle trajectories reconstruction, the transverse momentum of the particles can be reconstructed. For measured particles with p_T values, ranging between 1 to 10 GeV and pseudorapidity values of $|\eta| < 1.4$, the track resolutions are typically 1.5 % in p_T and 25–90 (45–150) μm in the transverse (longitudinal) impact parameter [69].

3.2.3 Hadronic Forward calorimeter

The Hadronic Forward Calorimeter (HF) is one of the components of the Calorimeter system of the CMS experiment. The calorimeter system of CMS is composed of an Electro-magnetic (ECAL) and a Hadronic (HCAL) sections. The HCAL is further subdivided in four parts, the Barrel (HB) and the Outer Barrel (HO), where the former is embedded inside the solenoid and the later resides just after the solenoid volume.

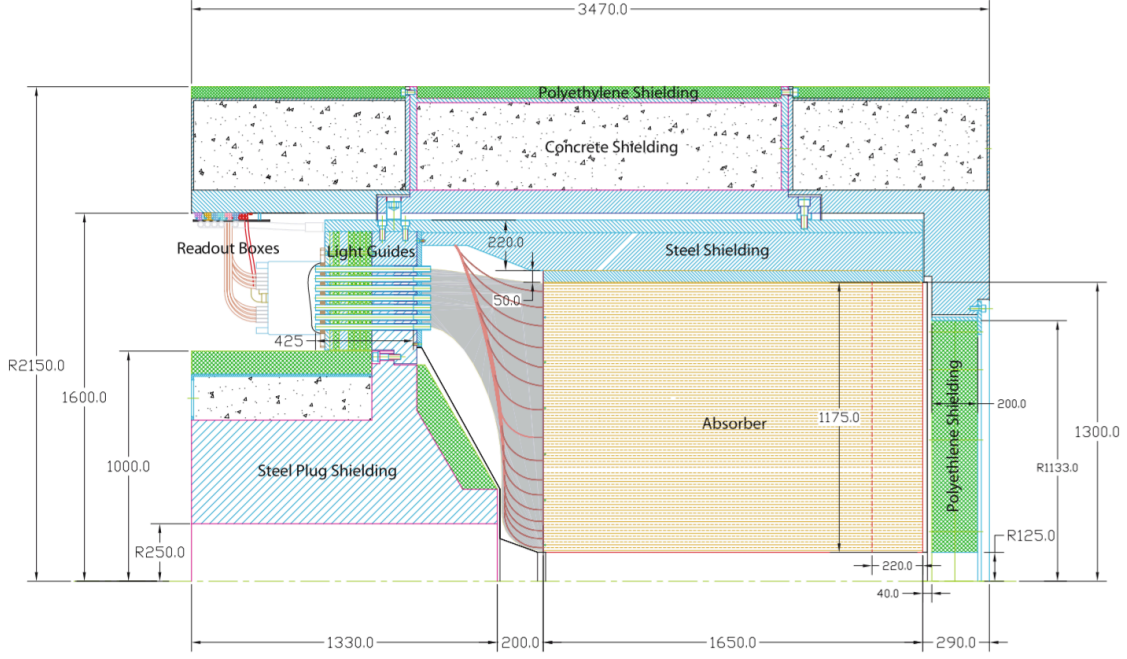


Figure 3.4: Schematic view of the CMS Hadronic Forward (HF) calorimeter. It is visible the absorber, the optic fibres and the readout boxes that contain the PMT's. The CMS interaction point is located to the right of the schematic.

At both sides of the HB barrel, the End-cap Calorimeter (HE) is located, closing the barrel volume. The most distant calorimeter with respect to the IP is the HF calorimeter, which is located at 11.2m either side along the beam direction. The HF calorimeter is one of the CMS subdetectors that plays a crucial role in this analysis.

The HF calorimeter covers a pseudorapidity acceptance of $3 < |\eta| < 5$, being one of the most forward detectors in CMS, just after the Zero Degree Calorimeter ($|\eta| < 8.4$) and CASTOR ($-6.6 < \eta < -5.2$).

It is constructed in a sampling fashion, i.e. there are two kinds of material involved in the construction: an absorber material, and the active material. This way of construction permits to have a compact detector while reducing the amount of active material and consequently the construction price. In the case of the HF, the absorber material are steel plates, while the active material are radiation-hard quartz fibres summing up to 1000 km of fibres with a weight of about 500 t.

Each of the HF detectors has a length of 1.65 m and a cylindrical shape with outer and inner radii of 130 cm and 12.5 cm respectively, leaving just enough space to let the beam pipe of 10 cm radius fit inside. The cylinders are divided in 18 azimuthal wedges of 20-degrees, where a matrix of longitudinal grooves spaced by 5 mm horizontally and vertically hosts the 800 μm radius quartz fibres [70]. Half of the fibres run over the full depth of the absorber, while the other half starts at a depth of 22 cm from the front of the detector. These two sets of fibres are read out separately. This arrangement makes it possible to distinguish showers generated by electrons and photons, which deposit a large fraction of their energy in the first 22 cm, from those generated by hadrons, which produce signals in both calorimeter segments. All these fibres are grouped with a typical segmentation of 0.175×0.175 ($\Delta\eta \times \Delta\varphi$), except for the first and last η segmentations which have a size of $\Delta\eta \simeq 0.1$ and $\Delta\eta \simeq 0.3$, respectively, and the two highest- $|\eta|$ segments, which have $\Delta\varphi = 20$ degrees each one. The energy collected in form of Cherenkov light by all the fibres in each of these segmentation corresponds to what is called an HF tower, or more generally, calorimeter tower.

CHAPTER

4

EVENT RECONSTRUCTION

Contents

4.1	Track reconstruction	52
4.2	Primary vertex	53
4.3	HF towers	54
4.4	Detector simulation	55
4.4.1	Pile-up simulation	55

The CMS detector is capable to measure charged and neutral particles, and further differentiate their nature by combining the information provided by the different subdetectors.

Figure 4.1 shows the characteristic signals left by each kind of particle while traversing the different layers of subdetectors. Neutral particles are only detected by the calorimeters, where photons deposit all their energy in the ECAL and hadrons in the HCAL. Charged particles leave a track in the inner tracking-system, where the track is bended by the magnetic field inside of the detector, according to the charge and momentum of the particle. Electrons deposit all their energy in the ECAL and charged hadrons in the HCAL, while muons cross the full detector up to the iron yoke, where dedicated muon chambers are embedded.

In this chapter the reconstruction of the different physics objects used for this thesis are described, in particular: tracks, vertices and HF towers. Also the detector and pileup simulations are covered. Detailed information on other physics objects reconstructed by the CMS experiment and not used for this thesis can be found elsewhere [60, 71, 72].

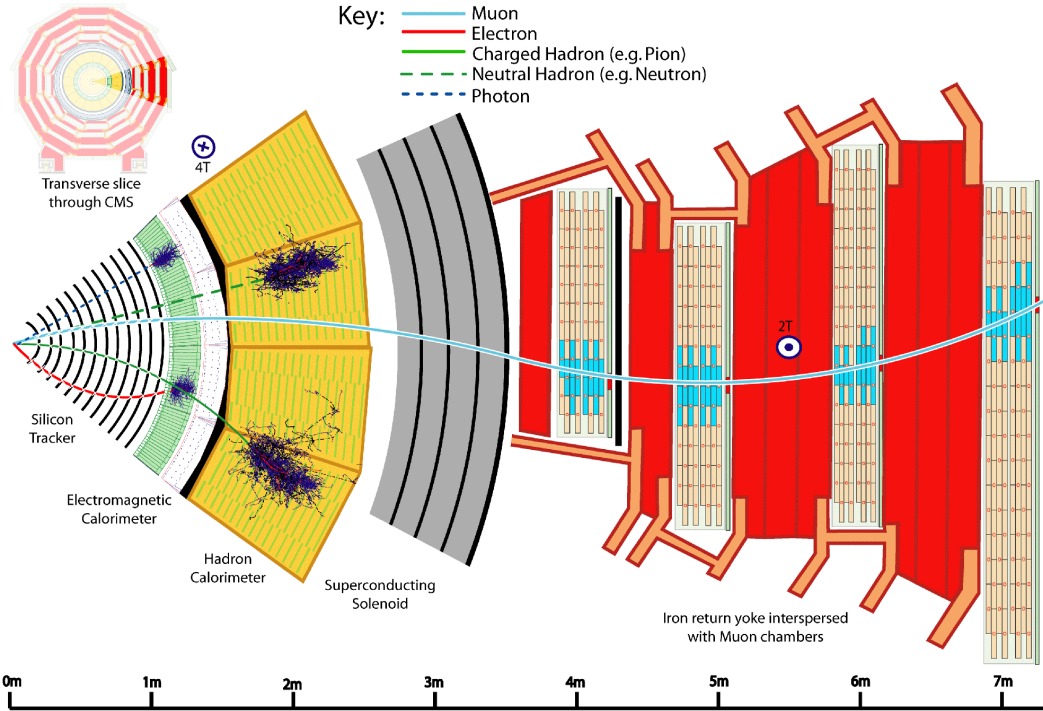


Figure 4.1: Transverse slice of the CMS detector showing the trace left in the subdetectors by different characteristic particles.

4.1 Track reconstruction

The tracker is designed to reconstruct tracks from events with over 1000 charged particles emerging from the IP at each bunch crossing, where each bunch crossing has, in average, 20 simultaneous proton-proton interactions (the so-called pileup). In addition, signals in the detector coming from earlier and latter bunch crossings are expected due to finite time resolution of the detector (known as out-of-time pileup).

As mentioned in Sec. 3.2.1, the tracker system has to achieve a high efficiency with a low fake rate at frequencies of up to 100 kHz to be used by the HLT, where the fake tracks are often coming from unrelated hits wrongly combined together or from related hits with additional spurious hits.

The software used for the track reconstruction takes as input the hits reconstructed in the planes of each of the silicon modules using local coordinates. Then, the local coordinates have to be translated to the global tracking coordinate system, taking into account the discrepancies between the assumed tracker geometry and the actual (current) geometry, where the actual geometry of the tracking detector has been estimated with an impressive accuracy as described in appendix A.

The tracking software at CMS is commonly referred to as the Combinatorial Track Finder (CTF), which is an adaptation of the combinatorial Kalman filter [69, 73].

The CTF reconstructs the tracks using an iterative procedure. At each iteration the easiest tracks to find (e.g. tracks with large- p_T values, tracks produced near the interaction point) are reconstructed and the associated hits removed from the collection. In this way, after each iteration the combinatorial complexity of hits is reduced, letting the CTF find more difficult tracks (e.g. tracks with low- p_T values, displaced tracks).

The process consists of five iterations:

- Iteration 0: In this iteration, tracks originated near the interaction point with $p_T > 0.8$ GeV that have three pixel hits (prompt tracks) are found. Most of the tracks are found at this step.
- Iteration 1: Used to recover prompt tracks that have only two pixel hits.
- Iteration 2: Configured to find low- p_T prompt tracks.
- Iterations 3-5: Intended to find tracks that originate outside the beam spot and to recover tracks not found during the previous iterations.

At each iteration four steps are performed:

- Step 1: A seed is generated from the information provided by 2 or 3 hits to estimate the trajectory and the corresponding uncertainties.
- Step 2: An extrapolation of the trajectory along the flight path of the charged particle is performed, and by making use of the Kalman filter, additional hits compatible with the trajectory are assigned.
- Step 3: The best fitted and smoothest trajectory based in the Kalman filter is found.
- Step 4: Quality flags are applied to the reconstructed tracks, such that only those fulfilling certain criteria are selected.

The difference between iterations are the generation of seeds and the criteria of the quality flags.

4.2 Primary vertex

A central feature of each proton-proton collision is the primary vertex (PV). Per bunch crossing multiple interactions can occur, which have to be properly separated by the reconstruction procedure. The procedure centrally used by the CMS experiment consists of three steps [69]:

- Track selection

Aims to select tracks produced promptly in the interaction-region by imposing requirements on the maximum allowed transverse impact-parameter-significance with respect to the beamspot, on the number of strip and pixel hits associated to the tracks, and on their normalized χ^2 . To ensure high reconstruction efficiency in Minimum Bias events, there is no requirement on the minimum allowed track p_T .

- Track clustering

Carried by a Deterministic Annealing procedure (DA) [74], in which each track is associated to one vertex according to its z coordinate and p_T . The DA process not only finds positions and assignments of tracks to vertices, but also finds the number of vertices.

- Fitting the position of each vertex

Vertices with at least two assigned tracks are fitted using the Adaptive Vertex Fitter (AVF) [75]. The AVF computes the best estimate of vertex parameters, including its x , y and z position and covariance matrix, as well as the indicators for the success of the fit, such as the number of degrees of freedom for the vertex, and the weights of the tracks used in the vertex.

After the reconstruction of all the vertices in the event, the vertex with the largest sum $\sum_i p_{T,i}$, where $p_{T,i}$ is the transverse momentum of the i th-track associated to the vertex, is selected as the primary vertex.

4.3 HF towers

The event selection that will be described in section 5.3.4 relies on the energy depositions left by particles in the HF calorimeters.

As mentioned in Sec. 3.2.3, the HF calorimeter has two different types of fibres (according to their length) running through it. Long fibres traversing the full depth, and short fibres starting at 22 cm from the front face of the detector. These long and short fibres are interleaved and connected to two different readout channels. This way, the signals from the electromagnetic and hadronic parts can be separated as shown in figure 4.2.

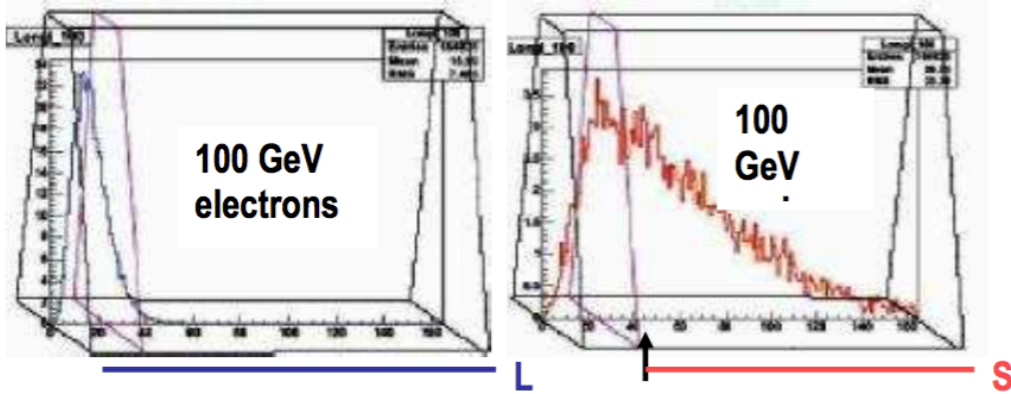


Figure 4.2: The development of Electromagnetic (left) and Hadronic (right) showers traveling from left to right in one HF wedge [70].

Each HF side is divided into 18 wedges of 20 degrees in azimuthal angle. Each wedge consists of 13 sections in η and 2 sections in φ , with $\Delta\varphi = 10$ degrees, except for the two last η segments which are of $\Delta\varphi = 20$ degrees. This segmentation creates 432 sections, each of them with two independent channels, one for the long fibres and the other for the short fibres.

Energy deposition measurements in the HF have two basic building blocks, the calorimeter cells (HF rechits), and the calorimeter towers (HF towers). The rechit is a data structure that contains energy, time and flags indicating the quality of the reconstructed quantities. Once the rechit energies are collected, different quality cuts are applied and the rechits are further combined into HF towers, which contain an electromagnetic and a hadronic component.

4.4 Detector simulation

In order to compare the simulated MC samples with the measured data at detector level, and to correct the data to the stable particle level, a proper simulation of the detector response is needed. The detector simulation is done with the GEometry ANd Tracking (GEANT4) toolkit [76].

A detailed computational model of the detector is taken by GEANT4 in order to simulate the interactions of the particles crossing the different materials. This includes the active detector materials, such as the silicon modules of the tracker, or the quartz fibres of the HF calorimeter, but also all the structural material like supports, cooling pipes, etc. In that way, GEANT4 can simulate the electronic signals of each of the CMS subdetectors in exactly the same format as the data, in such a way that the same reconstruction software can be used in both, MC and data.

In general, the MC and detector simulations are centrally provided [77] by the CMS collaboration through the CMS software, CMSSW [78]. It is separated in different stages. First the MC simulation is generated at stable particle level, also known as generator level. Then, the detector response of the generated particles interacting with the detector materials, both active and inactive, is simulated with GEANT4. These two steps together are the so-called GEN-SIM step. Right after, the simulated detector signals are used to reconstruct physics objects like tracks and jets. This step is known as RECO, and it is the same for both data and MC simulations. Finally, a skimming of the RECO information can be performed in order to reduce the size of the files.

4.4.1 Pile-up simulation

At the LHC, the proton bunches that collide in the centre of the detector can contain up to 1.15×10^{11} protons. With this large amount of protons and depending on the beam conditions, i.e. crossing angle and beam size, each bunch crossing can produce from zero proton-proton interactions up to 50 or even 100 proton-proton pileup interactions. In this section, the CMS procedure to simulate the pileup is explained.

Two Monte Carlo event samples are generated with exactly one collision per bunch crossing. One of the samples is regarded as the signal, while the events of the other sample are considered as additional collisions to be added to each signal event according to a Poisson distribution. As a result, each Monte Carlo event has at least one collision per bunch crossing. In this analysis, the additional collisions are distributed according to a Poisson distribution of mean $\alpha = 0.3$. Therefore, the average number of collisions per bunch crossing in the Monte

Carlo event samples with at least one collision per bunch crossing, $\langle n \rangle|_{n \geq 1}$, is given by:

$$\langle n \rangle|_{n \geq 1}(\text{Monte Carlo}) = 1 + \alpha = 1.3. \quad (4.1)$$

This implementation of the pileup in the simulation does not coincide with the pileup definition commonly used for the data samples. The number n of collisions per bunch crossing in the data samples, with $n \geq 0$, is distributed according to a Poisson distribution whose mean μ defines the average number of collisions per bunch crossing. For the run used in this analysis, the value of μ is equal to 1.3 [79], and the average number of collisions per bunch crossing in the data sample is given by:

$$\langle n \rangle|_{n \geq 0}(\text{data}) = \mu = 1.3. \quad (4.2)$$

The average number of collisions per bunch crossing for the events in the data sample with at least one collision per bunch crossing, $\langle n \rangle|_{n \geq 1}$, is therefore given by:

$$\langle n \rangle|_{n \geq 1}(\text{data}) = \frac{\mu}{1 - e^{-\mu}} = 1.79. \quad (4.3)$$

The Monte Carlo event samples are therefore reweighted in order to describe the distribution of the vertex multiplicity in the data. At the end of the reweighting procedure, the value of α should be equal to 0.79. The reweighting procedure is described in Sec 6.1.

CHAPTER

5

EVENT SELECTION

Contents

5.1	Data samples	57
5.2	MC samples	58
5.3	Detector level selection	59
5.3.1	Trigger selection	59
5.3.2	Vertex selection	59
5.3.3	Track selection	59
5.3.4	Event selection	60
5.4	Stable particle level definition	60

In this chapter, the data sets used for the analysis, the selection criteria at detector level and the stable particle level are presented. Sec. 5.3 is separated in four subsections explaining the implemented trigger, the vertex selection, the track selection and the different enhanced sample definitions. In Sec. 5.4, the stable particle level to which the measurements are corrected is defined.

5.1 Data samples

The analysis presented in this thesis is based on the data recorded by the CMS detector in July 2015 at a centre-of-mass energy of $\sqrt{s} = 13$ TeV. The Zero Bias datasets used correspond to 3.9 million events triggered by the presence of both beams passing at the IP¹. The average number of collisions per bunch crossing, so called pileup, is equal to 1.3 [79].

¹As defined in Sec. 3.2, the IP is the interaction point.

5.2 MC samples

Different MC event generators are used to simulate proton-proton collisions. The Monte Carlo event generator PYTHIA8 (version 8.153) is used with the recent tune CUETP8M1, which was determined from underlying event data at $\sqrt{s} = 7$ TeV. PYTHIA8 uses a model in which partonic interactions are interleaved with parton showering. In PYTHIA, the parton showers are modelled according to the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) prescription and hadronization is based on the Lund string fragmentation model. Diffractive cross sections are described by the Schuler-Sjöstrand model. In PYTHIA8, particle production from a low mass diffractive system X with a mass $M_X < 10$ GeV is described by the string model, while for states with higher masses $M_X > 10$ GeV a perturbative description of the Pomeron-proton scattering is introduced. This latter is based on the diffractive parton distribution functions (dPDFs), which represent the probability distributions for partons in the proton under the constraint that the proton emerges intact from the collision. The non-perturbative string model introduces a mass dependence in the relative probability for a Pomeron to couple to a quark or a gluon.

The MC event generator HERWIG++ (version 2.7.0) is used with the recent tune CUETHS1, which was determined from underlying event data at $\sqrt{s} = 7$ TeV. It is based on matrix-element calculations similar to those used in PYTHIA and the parton showering is also driven by the DGLAP evolution equations. However, HERWIG++ features angular-ordered parton showers and uses cluster fragmentation for the hadronization. Due to the bad description of the MB data as shown in Chapter 6, this MC is not further used in the correction of the measurement, neither in the discussion of the results.

The data are also compared to the predictions of the MC event generator EPOS(version 1.99) used in cosmic ray physics with the LHC tune. The model includes contributions from soft- and hard-parton dynamics. The soft component is described in terms of the exchange of virtual quasi-particle states, as defined in Gribov's Reggeon field theory, with multi-pomeron exchanges accounting for underlying-event effects. At higher energies, the interaction is described in terms of the same degrees of freedom, but generalised to include hard processes via hard-pomeron scattering diagrams, which are equivalent to a leading-order pQCD approach with DGLAP evolution. These models are retuned to LHC data, including cross section measurements by TOTEM, and charged particle multiplicity measurements in the central region at $\sqrt{s} = 7$ TeV by ALICE and CMS.

Event samples obtained from Monte Carlo event generators are passed through the CMS detector simulation based on GEANT4 and are processed and reconstructed in the same manner as collision data. Simulated event samples are used to determine the event selection efficiency and the tracking efficiency.

5.3 Detector level selection

In this section, the trigger implemented in this analysis, as well as the offline selection of events and tracks are presented. The offline event selection consists on different requirements being applied to the reconstructed vertices, tracks and the maximum energy deposition on the HF calorimeters.

5.3.1 Trigger selection

As explained in Sec. 3.2.1, a trigger system is needed to select events of interest. For this analysis, the subject of study are the so-called Minimum Bias events (MB). The characteristic of the MB events is to have a selection as loose and inclusive as possible. In contrast with other analyses, where specific objects are under study, like Drell-Yan [80] or jet cross sections [81], in this analysis the interest is focused on studying for instance processes involved in the inclusive particle production or the saturation effects in the low transverse momenta. For that reason, the most unbiased trigger available is used, the so-called ZeroBias trigger.

The ZeroBias trigger requires the coincidence of signals from both BPTX devices, indicating the presence of both proton bunches crossing the interaction point.

In addition to the nominal analysis trigger of coincidence of signals in both BPTX devices, three other selections using BPTX information have been used for background studies, as described in Sec. 6.3. A selection corresponding to the absence of circulating beams in the LHC is performed, known as *no BPTX* trigger. Additionally, two single side selections, where only one beam is circulating in the LHC, *BPTX minus only* and *BPTX plus only*, are requested.

5.3.2 Vertex selection

After the online event selection, the events are selected offline. The first requirement for the events is to have one and only one good reconstructed vertex, i.e. a vertex located within 15 cm of the beam spot along the beam line, $|z| < 15$ cm. The distance to the beam spot in the transverse plane must be smaller than 0.2 cm, and it has to contain at least two tracks fulfilling the nominal track selection of the analysis, except for the transverse momentum requirement of $p_T > 0.5$ GeV.

5.3.3 Track selection

The collection of tracks is defined as follows: high purity tracks are selected with a transverse momentum $p_T > 0.5$ GeV and a relative transverse momentum uncertainty smaller than 10 %. Tracks are measured within the pseudorapidity range $|\eta| < 2.4$ corresponding to the fiducial acceptance of the tracker, in order to avoid effects from tracks very close to the geometric edge of the tracking detector. A track-vertex association is applied by requiring a small impact parameter with respect to the primary-vertex position, both in the transverse plane, and along the z -axis. In the case of the impact parameter with respect to the vertex position in the transverse plane, d_{xy} , the significance is required to satisfy $|d_{xy}|/\sigma_{xy} < 3$, while for the point of closest approach to the primary vertex along the beam line, d_z , the

requirement $|d_z|/\sigma_z < 3$ is applied. The number of pixel hits associated to a track has to be greater or equal to 3 in the central region of the tracker acceptance $|\eta| < 1$ and greater or equal to 2 in the forward region $|\eta| > 1$.

5.3.4 Event selection

The event selection based on the HF calorimeters uses the HF towers with an energy threshold of 5 GeV in the fiducial acceptance $3 < |\eta| < 5$. The activity in the calorimeters is defined by the presence of at least one tower with an energy above the threshold value. The absence of calorimeter activity (veto condition) is defined by the absence of towers with an energy above the threshold value. In Sec. 6.3 the motivation for the energy threshold of 5 GeV is described.

An Inelastic-enhanced event sample is defined by requiring activity in at least one of the HF calorimeters, corresponding to the presence of at least one HF tower with an energy $E > 5$ GeV in any of the HF calorimeters. A sample dominated by non-single-diffractive dissociative events (NSD-enhanced sample) is defined by requiring activity in both of the HF calorimeters, corresponding to the presence of at least one HF tower with an energy $E > 5$ GeV on both sides of the calorimeter. A sample enhanced in single diffractive dissociative events (SD-enhanced sample) is defined by requiring activity in only one of the HF calorimeters, the veto condition being applied to the other side. The Inclusive sample obtained by only requiring the presence of a good reconstructed primary vertex contains 2.25 million events. The Inelastic-enhanced, NSD-enhanced and SD-enhanced event samples contain 2.23, 2 and 0.23 million events, respectively.

5.4 Stable particle level definition

The stable particle level is defined by the presence of at least one stable charged particle with a transverse momentum $p_T > 0.5$ GeV within the pseudorapidity region $|\eta| < 2.4$. The Inelastic-enhanced event sample is defined by the presence of at least one stable particle with an energy $E > 5$ GeV in the range $3 < |\eta| < 5$. The NSD-enhanced event sample is defined by the presence of at least one stable particle with an energy $E > 5$ GeV in both regions $3 < \eta < 5$ and $-5 < \eta < -3$. The SD-enhanced event sample is defined by the presence of at least one stable particle with an energy $E > 5$ GeV in one of the regions $3 < \eta < 5$ or $-5 < \eta < -3$, while the other region has to be devoided of particles with an energy $E > 5$ GeV.

The definition of the Inelastic-enhanced, NSD-enhanced and SD-enhanced event samples being based on the activity in the HF calorimeters at detector level has been investigated using predictions of the MC event generator PYTHIA8. The presence or absence of stable particles in the corresponding η range at stable particle level, the correlation between the leading HF calorimeter tower energy and the leading stable-particle energy in the corresponding η range have also been studied using the MC predictions from PYTHIA8.

The response of the calorimeter to the leading tower is presented in figure 5.1 for the HF-plus (i.e. $3 < \eta < 5$) and HF-minus (i.e. $-5 < \eta < -3$) calorimeters separately. The response

is defined by the ratio between the leading tower energy and the leading stable-particle energy, and is shown as a function of the leading stable particle energy. The response observed is in reasonable good agreement with the non-linear behaviour of the calorimeter. Figure 5.2 shows the stable particle occupancy for the HF-plus and HF-minus calorimeters separately, for stable particles with an energy $E > 5$ GeV and a segmentation in pseudorapidity and azimuthal angle corresponding to the actual tower granularity defined in table 5.1. The low values of the occupancy validate the assumption to associate leading towers with leading particles in the HF calorimeters. Figure 5.3 shows the correlation in energy, pseudorapidity and azimuthal angle between the leading HF calorimeter towers and the leading stable particles with an energy $E > 5$ GeV for the HF-plus and HF-minus calorimeters separately. The reasonable correlation observed validates again the association between leading towers and leading stable particles in the HF pseudorapidity range.

HF ring number	$ \eta $ lower edge	$ \eta $ upper edge	η width	φ width
1	2.866	2.976	0.111	10°
2	2.976	3.152	0.175	10°
3	3.152	3.327	0.175	10°
4	3.327	3.503	0.175	10°
5	3.503	3.677	0.175	10°
6	3.677	3.853	0.175	10°
7	3.853	4.027	0.175	10°
8	4.027	4.204	0.175	10°
9	4.204	4.377	0.175	10°
10	4.377	4.552	0.175	10°
11	4.552	4.730	0.175	10°
12	4.730	4.903	0.175	20°
13	4.903	5.205	0.300	20°

Table 5.1: Pseudorapidity and azimuthal segmentation of the HF calorimeters. The first column shows the ring number of each HF segmentation, the second and third columns show the lower and upper η edges respectively. The last two columns show the width in η and the width in φ , respectively.

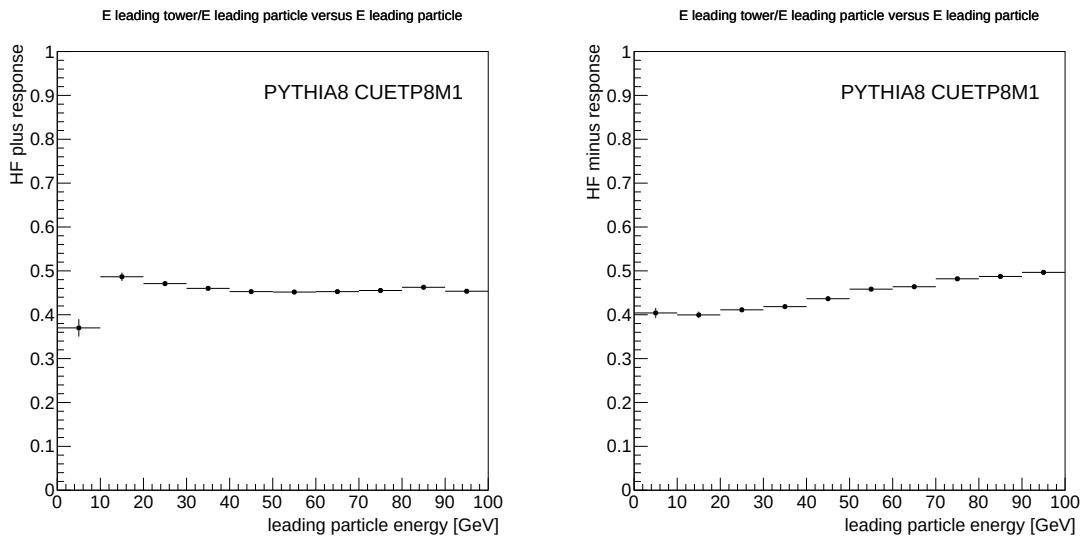


Figure 5.1: Distributions of the response for the leading tower for the HF-plus (left) and HF-minus (right) calorimeters. The distributions correspond to the predictions of the Monte Carlo event generator PYTHIA8.

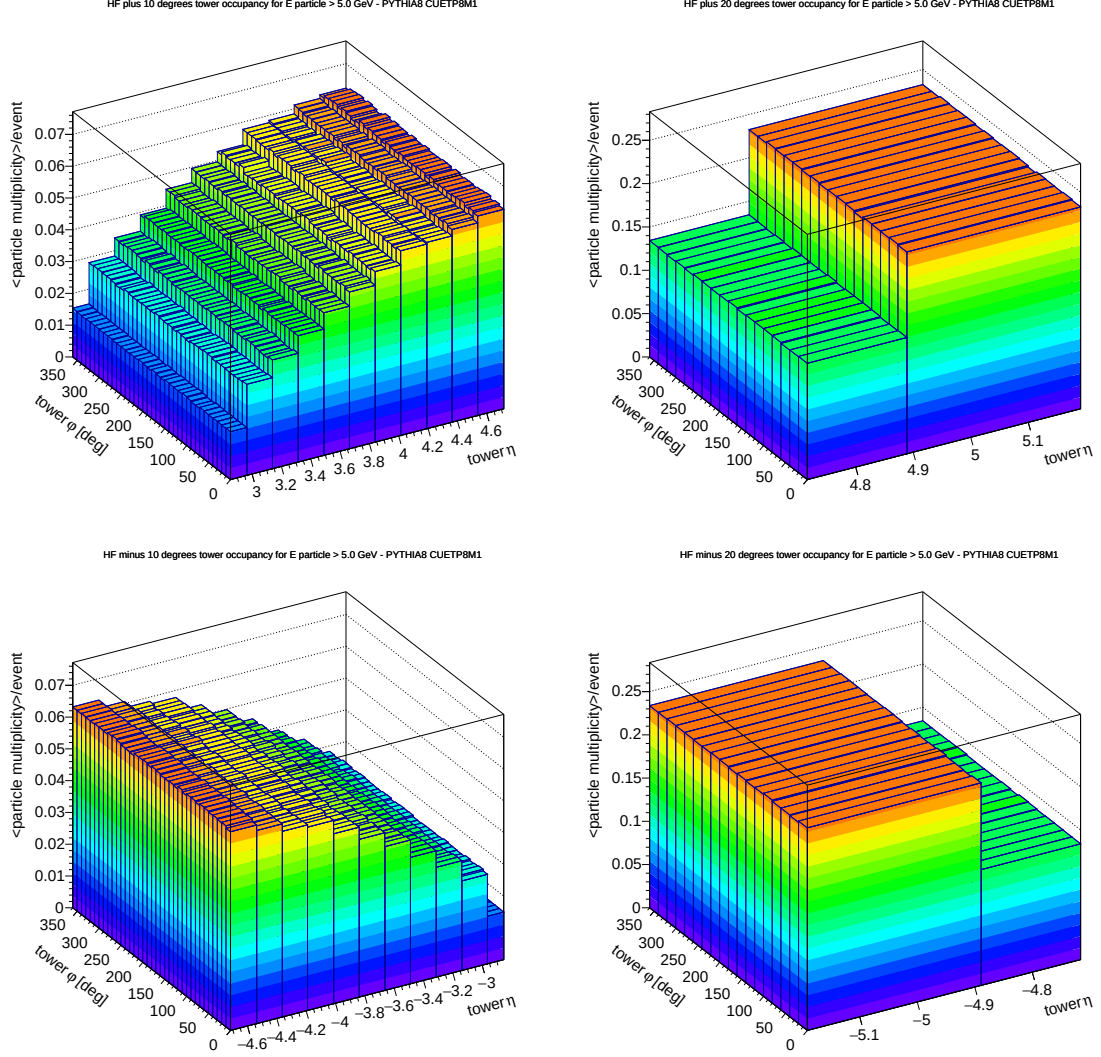


Figure 5.2: Distributions of the stable particle occupancy for the HF-plus (top) and HF-minus (bottom) calorimeters separately, for stable particles with an energy $E > 5$ GeV and a segmentation in pseudorapidity and azimuthal angle corresponding to the actual tower granularity. The left plots correspond to $2.86 < |\eta| < 4.73$ and the right plots to $4.73 < |\eta| < 5.20$. The distributions correspond to the predictions of the Monte Carlo event generator PYTHIA8.

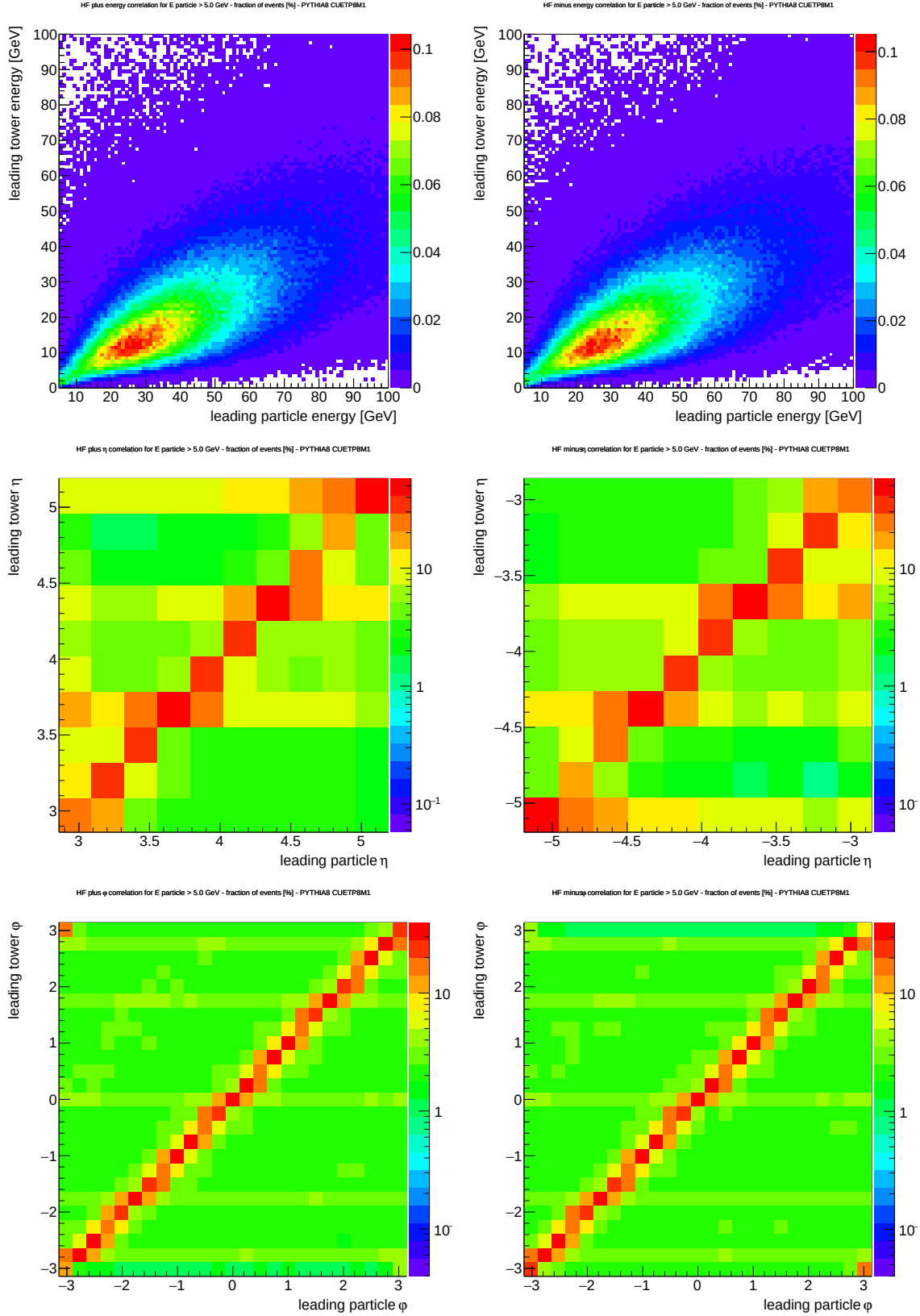


Figure 5.3: From top to bottom: distributions of the correlation in energy, pseudorapidity and azimuthal angle between the leading HF calorimeter towers and the leading stable particles with an energy $E > 5$ GeV for the HF-plus (left) and HF-minus (right) calorimeters separately. The distributions correspond to the predictions of the Monte Carlo event generator PYTHIA8.

CHAPTER

6

DETECTOR PERFORMANCE

Contents

6.1	Monte Carlo reweighting	65
6.1.1	Vertex multiplicity reweighting	65
6.1.2	Vertex z position reweighting	68
6.2	Tracks	70
6.3	HF towers	80

In this chapter, the description of the data at detector level by the Monte Carlo simulations that have been reconstructed with GEANT4 is presented. Also the study for the determination of the energy threshold for the HF towers is described.

6.1 Monte Carlo reweighting

It is usual that the conditions implemented in the simulation do not match the ones on the experimental data. For instance, the number of pile-up interactions, luminosity, luminous region, etc. can be some of the quantities that have small to medium discrepancies. Hence, a reweighting procedure to match the Monte Carlo simulation to the data is implemented.

The Monte Carlo samples used for this analysis showed differences to the data on the vertex multiplicity distribution and on the position of the interaction point along the z-axis. The two different reweighting procedures implemented in this analysis are described in the following.

6.1.1 Vertex multiplicity reweighting

Figure 6.1 shows the comparison of the vertex multiplicity distribution between data and Monte Carlo for events with at least one vertex per bunch crossing.

The differences in the vertex multiplicity distribution are due to the fact that the average number of collisions per bunch crossing for events with at least one collision per bunch crossing, $\langle n \rangle_{|n \geq 1}$, is 1.79 in the data sample but 1.3 in the Monte Carlo samples.

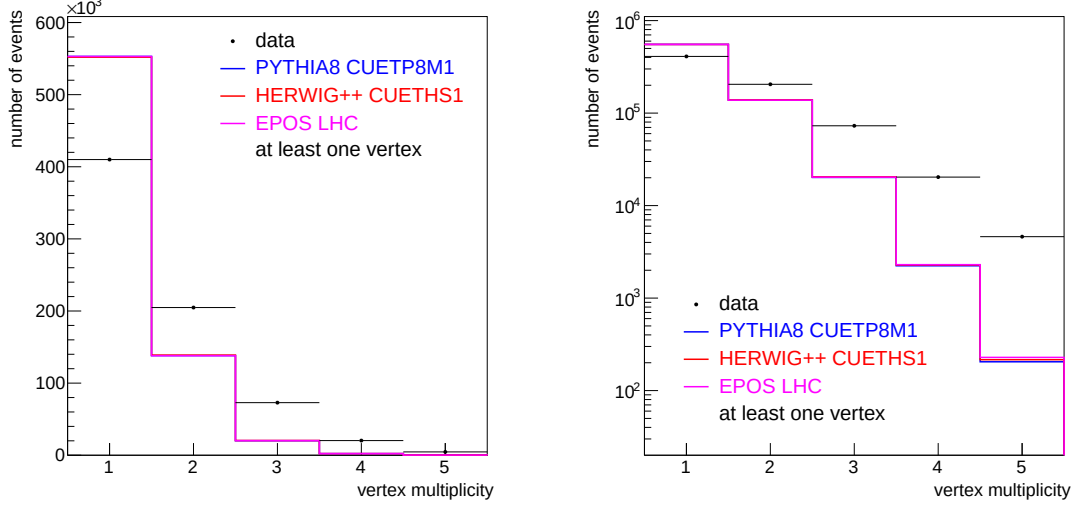


Figure 6.1: Distributions of the good vertex multiplicity for events with at least one good reconstructed vertex before the reweighting procedure. The data are compared to the predictions of the Monte Carlo event generators PYTHIA8, HERWIG++ and EPOS.

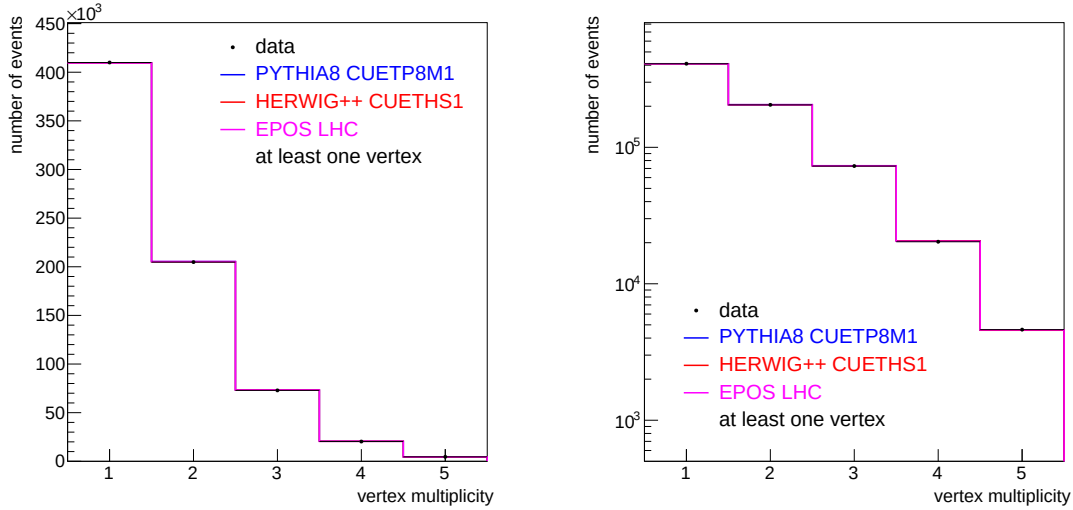


Figure 6.2: Distributions of the good vertex multiplicity for events with at least one good reconstructed vertex after the reweighting procedure. The data are compared to the predictions of the Monte Carlo event generators PYTHIA8, HERWIG++ and EPOS.

To bring the Monte Carlo simulation to the data, a weight defined as the ratio of the

Vertex reconstruction efficiency [%]							
number of collisions per bunch crossing	number of reconstructed vertices						
	0	1	2	3	4	5	6
1	15.4 %	83.3 %	1.2 %	0 %	0 %	0 %	0 %
2	2.4 %	28.0 %	67.2 %	2.3 %	0.1 %	0 %	0 %
3	0.4 %	7.1 %	34.6 %	54.8 %	3.0 %	0.1 %	0 %
4	0 %	1.8 %	12.5 %	37.1 %	45.1 %	3.3 %	0.2 %
5	0 %	0.4 %	3.6 %	17.5 %	39.8 %	35.4 %	3.3 %

Table 6.1: Vertex reconstruction efficiency as a function of the number of collisions per bunch crossing determined with the Monte Carlo event generator PYTHIA8.

vertex multiplicity distribution in data over Monte Carlo,

$$w = \frac{nPV_{data}}{nPV_{MC}} \quad (6.1)$$

is applied as a function of the additional number of generated collisions per bunch crossing. This is done iteratively until a good agreement between data and Monte Carlo is found. In this case five iterations were necessary. Figure 6.2 shows the vertex multiplicity distribution after the reweighting procedure.

The correlation between the number of vertices and the number of generated interactions per bunch crossing is shown in figure 6.3. This can be translated as a vertex reconstruction efficiency as a function of the number of collisions per bunch crossing, as shown in table 6.1.

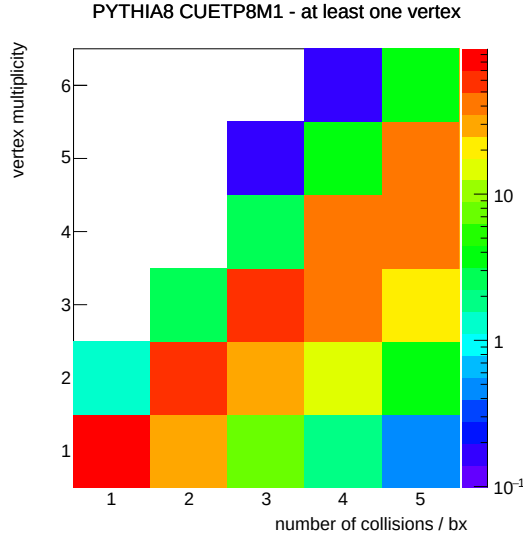


Figure 6.3: Correlation between the good vertex multiplicity and the number of collisions per bunch crossing for the predictions of the Monte Carlo event generator PYTHIA8.

In figure 6.4, the simulated number of collisions per bunch crossing before and after the reweighting is shown. Each time the MC distributions are compared to a Poisson distribution

in order to estimate the average number of additional interactions per bunch crossing. Before the reweighting procedure a Poisson of mean $\alpha = 0.3$ describes the MC distributions, while a Poisson of mean $\alpha = 0.79$ describes the MC distributions after the reweighting procedure, only showing small discrepancies in the last bin.

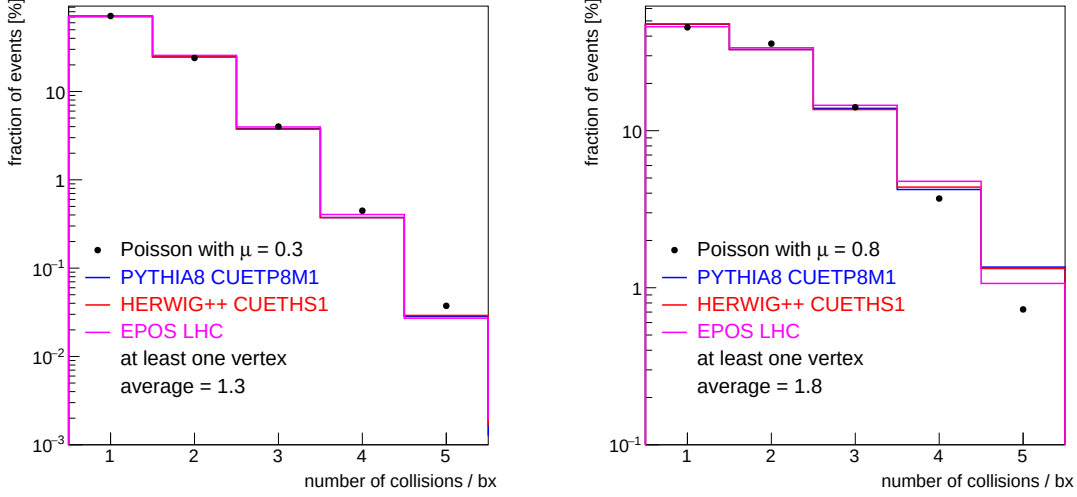


Figure 6.4: Distributions of the number of collisions per bunch crossing before (left) and after (right) the reweighting procedure for the predictions of the Monte Carlo event generators PYTHIA8, HERWIG++ and EPOS.

The effects of the reweighting procedure on the separation between the first and second vertex along the beam line, Δz , are shown in figure 6.5. The description is improved, specially in the small separation region. The description of this region is important to properly select events with exactly one good vertex.

Figure 6.6 shows the negligible effect of the reweighting procedure on the pseudorapidity of the tracks for events with exactly one vertex.

6.1.2 Vertex z position reweighting

When the beams traveling in opposite directions collide inside of the detector, the actual longitudinal position along the beam pipe where this occurs can vary depending on specific beam configurations of the run. Ideally, the position would always be exactly at the centre of the detector. In practice, the position varies according to a gaussian distribution that can also have a mean different than zero. The spread of the distributions is known as luminous region.

To fix the discrepancy between the recorded data and the Monte Carlo simulation, both distributions are fitted with a gaussian distribution. These two distributions are normalised to unity. Then a weight calculated as the ratio between the two gaussians, $w = f(z_{data})/f(z_{MC})$, is applied as a function of the z coordinate in the simulation. Figure 6.7 shows the vertex z coordinate distribution before and after the reweighting.

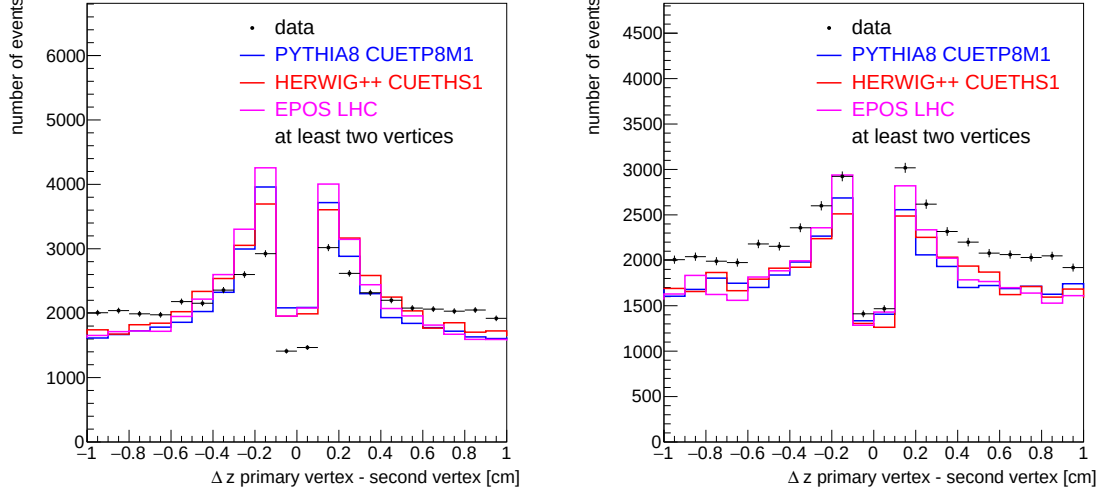


Figure 6.5: Distributions of the Δz separation along the beam line between the first and second vertices for events with at least two good reconstructed vertices, before (left) and after (right) the reweighting procedure. The data are compared to the predictions of the Monte Carlo event generators PYTHIA8, HERWIG++ and EPOS.

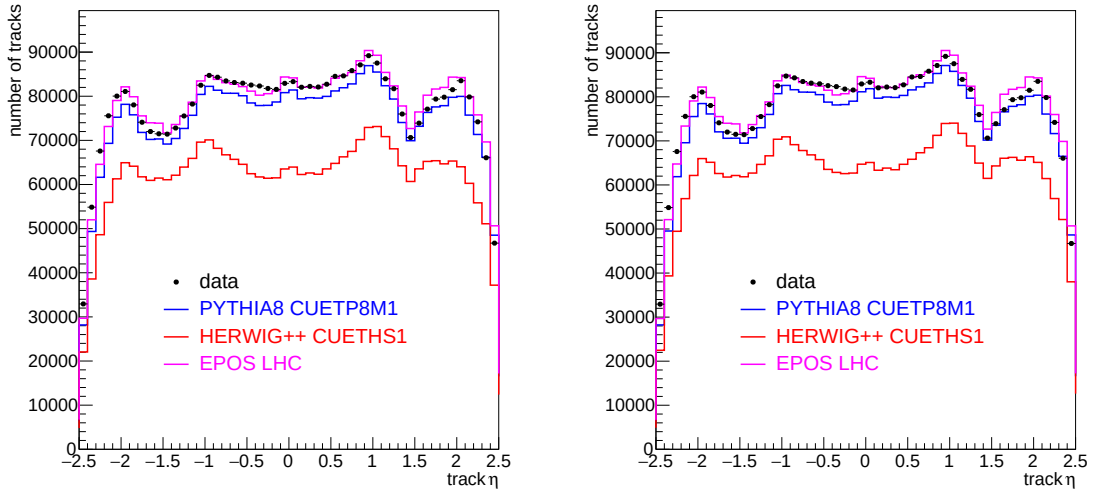


Figure 6.6: Pseudorapidity distributions of the tracks for events with exactly one good reconstructed vertex before (left) and after (right) the reweighting procedure. The data are compared to the predictions of the Monte Carlo event generators PYTHIA8, HERWIG++ and EPOS.

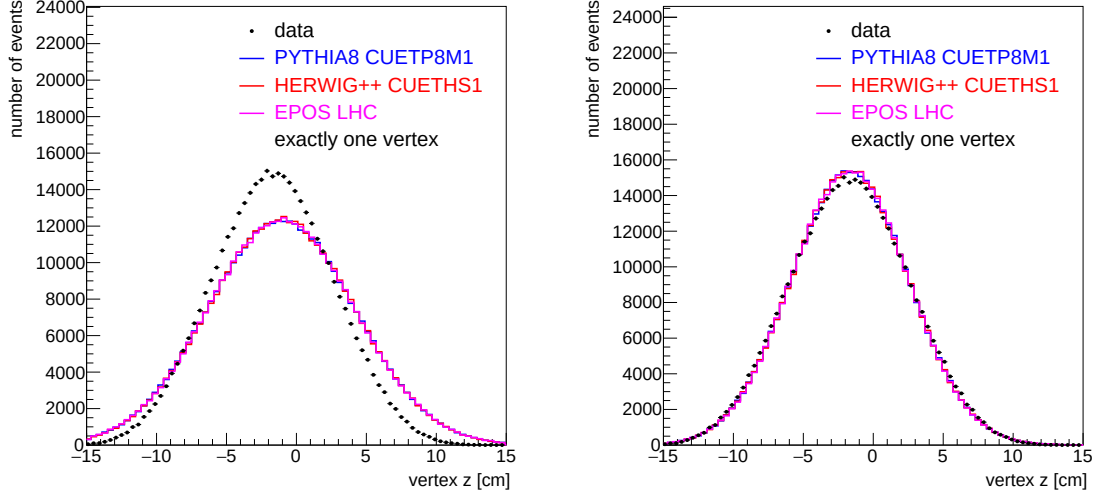


Figure 6.7: Distributions of the z coordinate of the reconstructed primary vertex before (left) and after (right) the reweighting procedure. The data are compared to the predictions of the Monte Carlo event generators PYTHIA8, HERWIG++ and EPOS.

After the reweighting, a Gaussian fit to the distribution of the z coordinate gives an average value of $\langle v_z \rangle = -1.68$ cm and a standard deviation of $\sigma z = 4.20$ cm for the size of the luminous region.

6.2 Tracks

In order to demonstrate the performance of the track and vertex reconstructions, several control distributions are shown for the primary vertices and the tracks associated to them. In particular, the distributions of the variables used in the track selection are presented. Additional control distributions are included in appendix B. The tracks have to fulfil the conditions specified in chapter 5, except for the requirements on the impact parameters, on the relative transverse momentum uncertainty and on the number of pixel hits associated to a track. A good agreement is observed between the data and the predictions from PYTHIA8 and EPOS, while HERWIG++ fails to describe most of the distributions.

The distributions of the pseudorapidity and azimuthal angle, and multiplicity and transverse momentum of the tracks associated to the primary vertex are shown in figures 6.8 and 6.9 respectively. The distributions of the pseudorapidity and azimuthal angle of the tracks are shown in figure 6.10 before and after the requirement to have at least 3 pixel hits.

Figures 6.11 and 6.12 show the transverse momentum distributions in different η regions, $|\eta| < 0.6$, $0.6 < |\eta| < 1.2$, $1.2 < |\eta| < 1.8$ and $1.8 < |\eta| < 2.4$, in linear and logarithmic scales respectively. The pseudorapidity, azimuthal angle and transverse momentum distributions of the leading tracks are presented in figure 6.13. Figure 6.14 shows the relative transverse momentum resolution σ_{p_T}/p_T and the resolution on the pseudorapidity and azimuthal angle

of the tracks.

The significance of the impact parameter with respect to the vertex position in the transverse plane, d_{xy}/σ_{xy} , and the significance of the point of closest approach to the primary vertex along the beam line, d_z/σ_z are presented in figure 6.15. The distributions of the track hit multiplicity are shown in figure 6.16 for all the tracker layers and for the pixel and strip layers separately.

The agreement between the data and the simulation is satisfactory, except for the region $-2 < \eta < -0.2$, $2.4 < \varphi < 2.7$ of the pixel barrel after the requirement to have at least 3 pixel hits. The distributions of the tracks from figure 6.10 show that the data and the simulation are nevertheless in reasonable agreement in the affected $\eta - \varphi$ range.

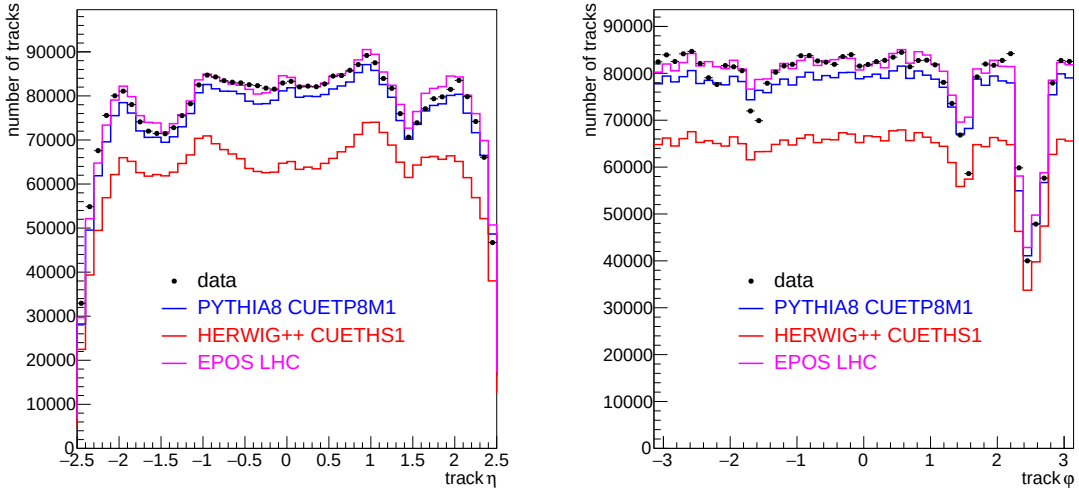


Figure 6.8: Distributions of the pseudorapidity and azimuthal angle of the tracks associated to the primary vertex. The data are compared to the predictions of the Monte Carlo event generators PYTHIA8, HERWIG++ and EPOS.

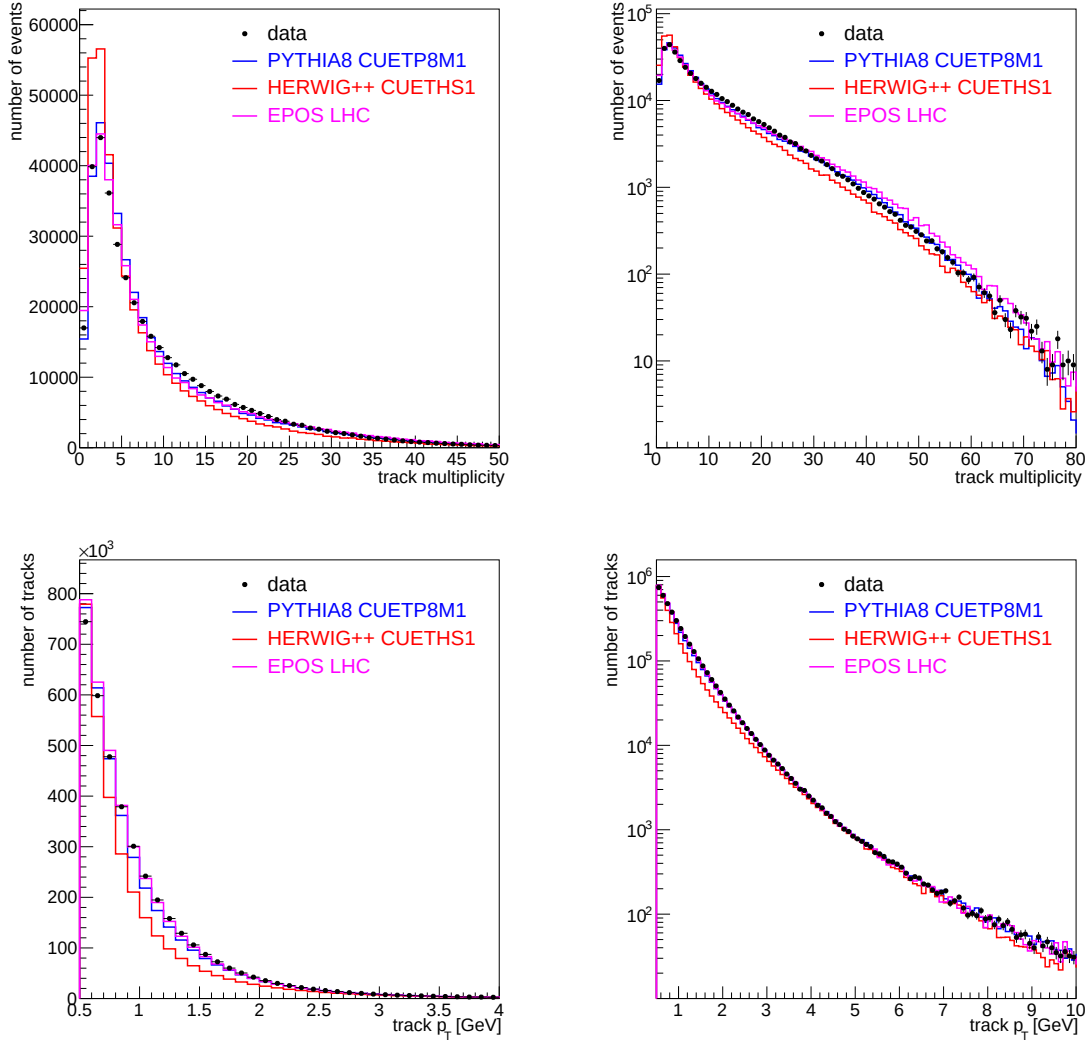


Figure 6.9: Distributions of the multiplicity and transverse momentum of the tracks associated to the primary vertex. Linear scale on the left, logarithmic scale on the right. The data are compared to the predictions of the Monte Carlo event generators PYTHIA8, HERWIG++ and EPOS.

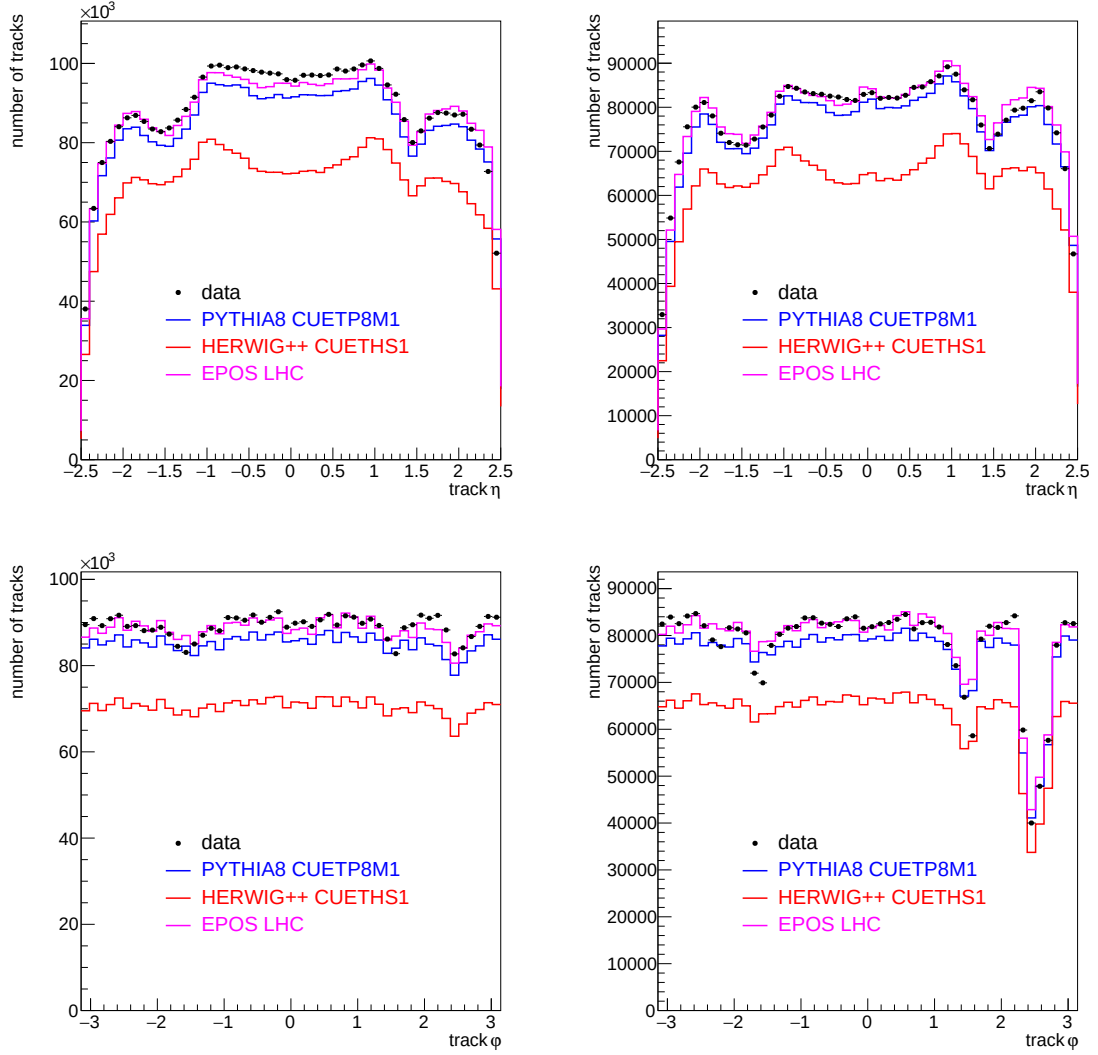


Figure 6.10: Distributions of the pseudorapidity and azimuthal angle of the tracks associated to the primary vertex before (left) and after (right) the requirement to have at least 3 pixel hits. The data are compared to the predictions of the Monte Carlo event generators PYTHIA8, HERWIG++ and EPOS.

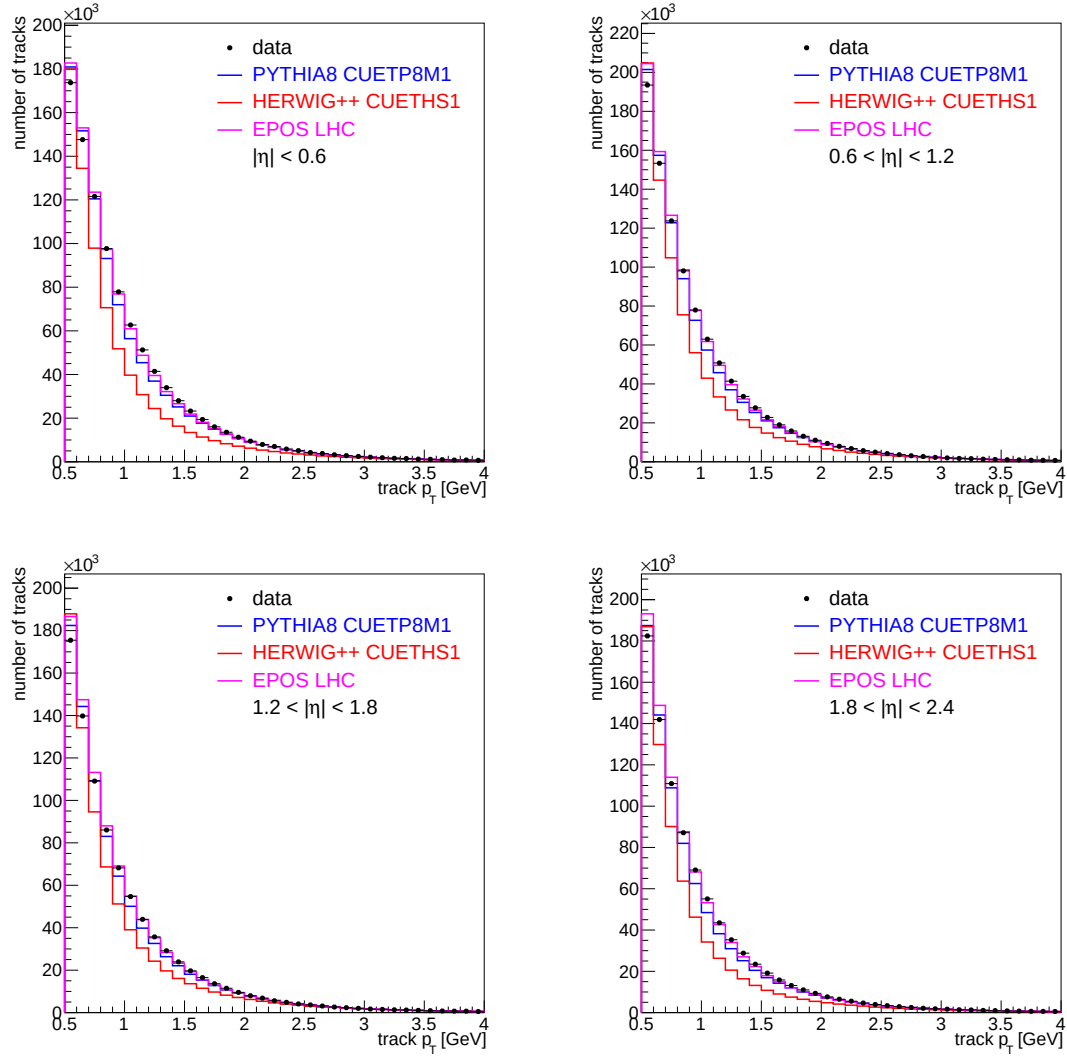


Figure 6.11: Transverse momentum distributions of the tracks in different η regions, $|\eta| < 0.6$, $0.6 < |\eta| < 1.2$, $1.2 < |\eta| < 1.8$ and $1.8 < |\eta| < 2.4$, in a linear scale. The data are compared to the predictions of the Monte Carlo event generators PYTHIA8, HERWIG++ and EPOS.

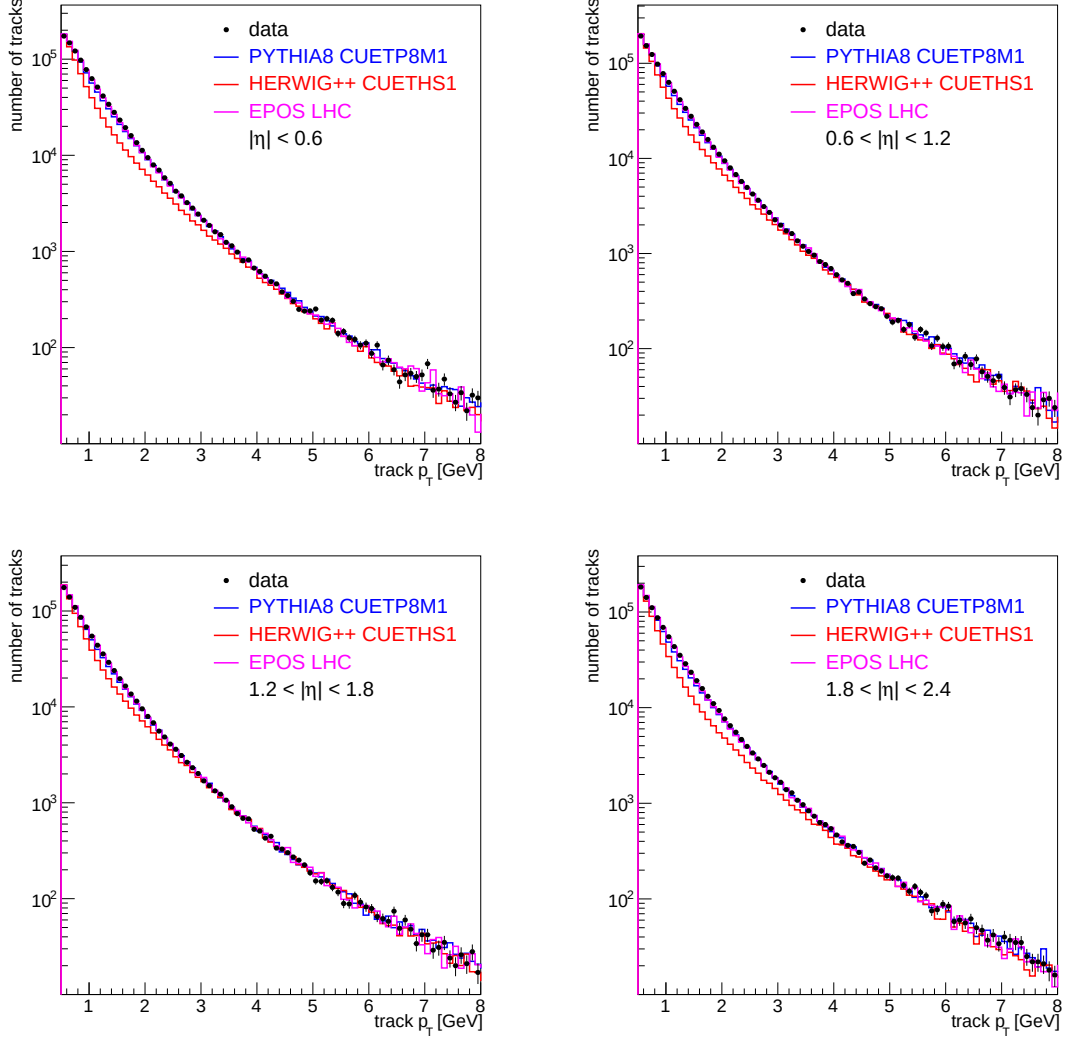


Figure 6.12: Transverse momentum distributions of the tracks associated to the primary vertex in different η regions, $|\eta| < 0.6$, $0.6 < |\eta| < 1.2$, $1.2 < |\eta| < 1.8$ and $1.8 < |\eta| < 2.4$, in a logarithmic scale. The data are compared to the predictions of the Monte Carlo event generators PYTHIA8, HERWIG++ and EPOS.

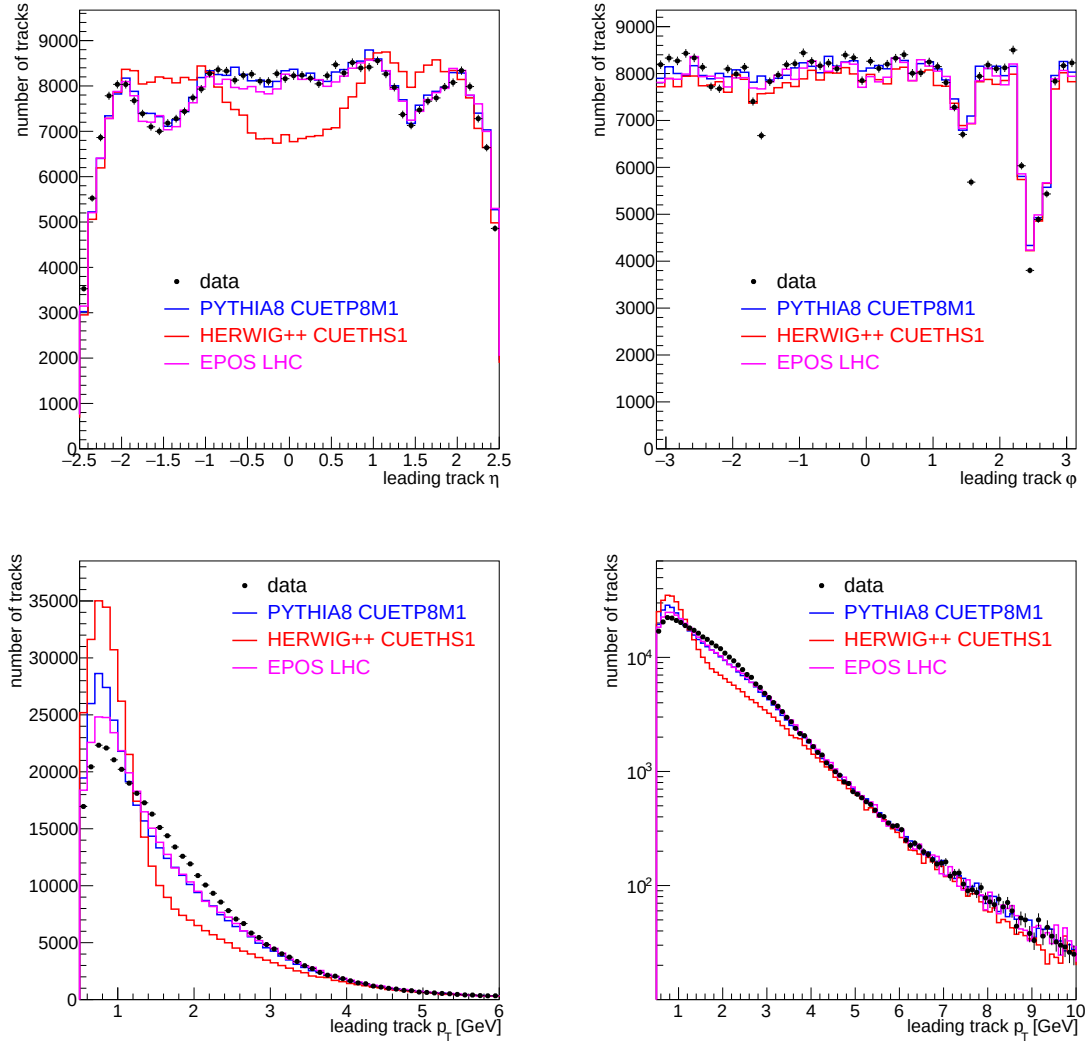


Figure 6.13: Pseudorapidity, azimuthal angle and transverse momentum distributions of the leading tracks. The data are compared to the predictions of the Monte Carlo event generators PYTHIA8, HERWIG++ and EPOS.

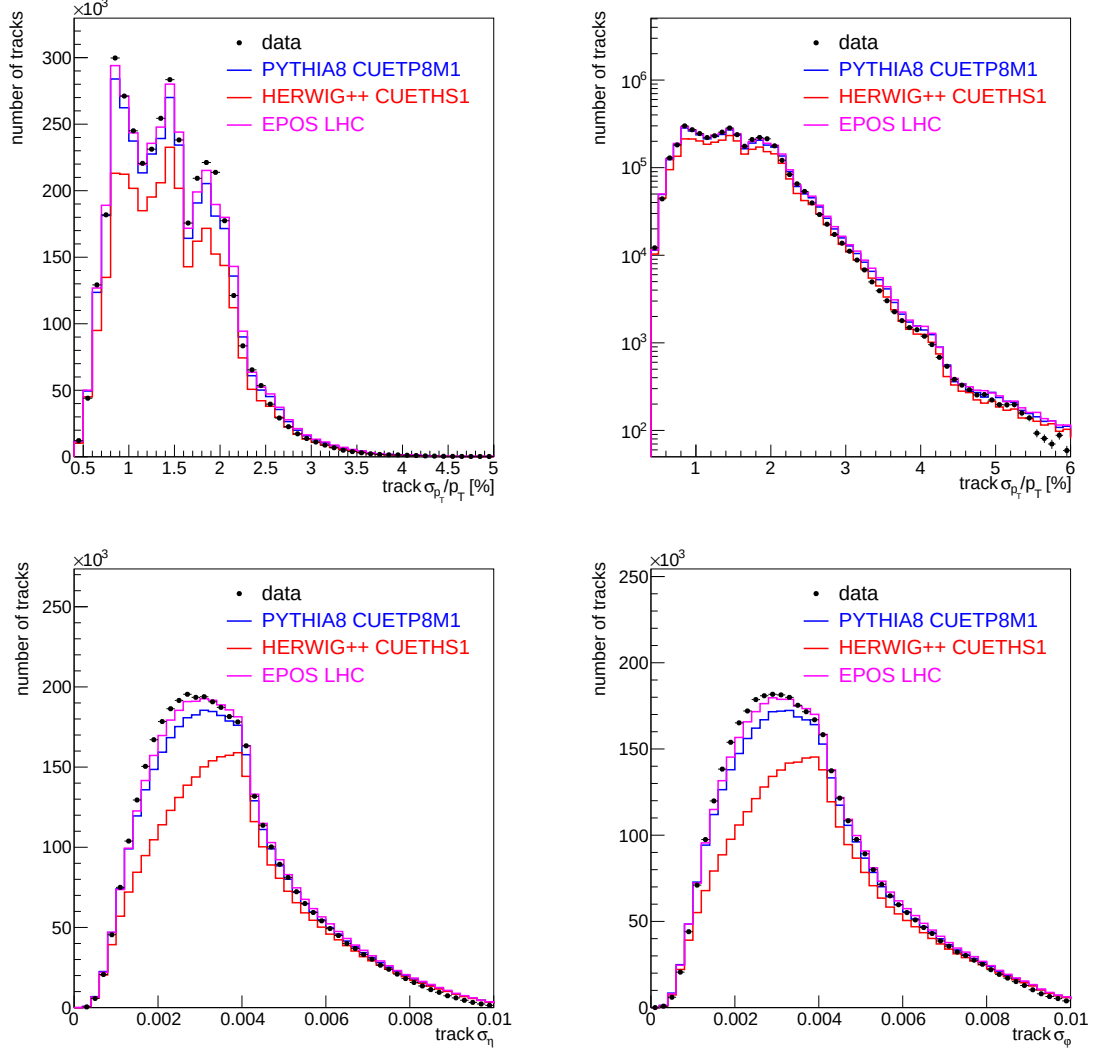


Figure 6.14: Relative transverse momentum resolution σ_{p_T}/p_T and resolution on the pseudo-rapidity and azimuthal angle of the tracks. The data are compared to the predictions of the Monte Carlo event generators PYTHIA8, HERWIG++ and EPOS.

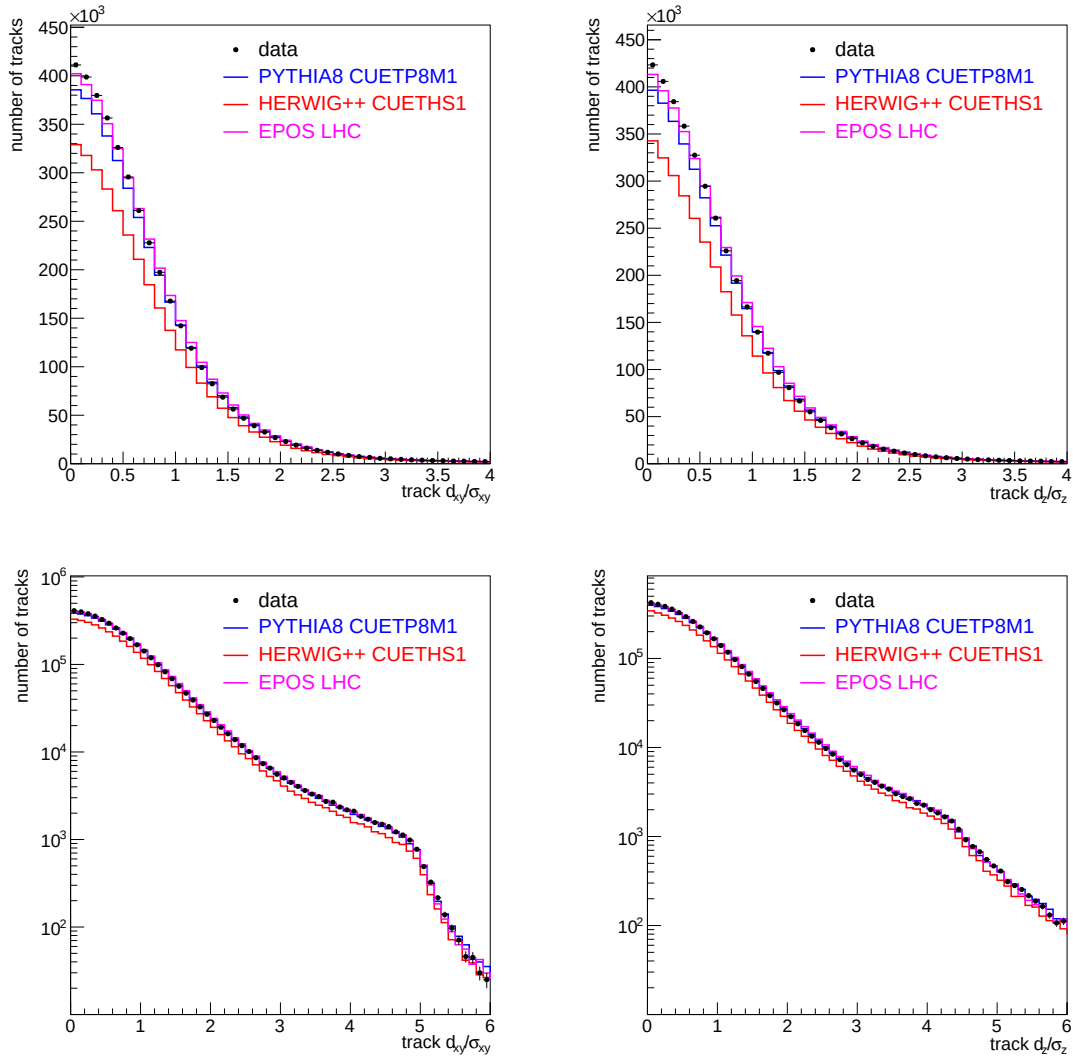


Figure 6.15: Significance of the impact parameter with respect to the vertex position in the transverse plane, d_{xy}/σ_{xy} , and significance of the point of closest approach to the primary vertex along the beam line, d_z/σ_z . The data are compared to the predictions of the Monte Carlo event generators PYTHIA8, HERWIG++ and EPOS.

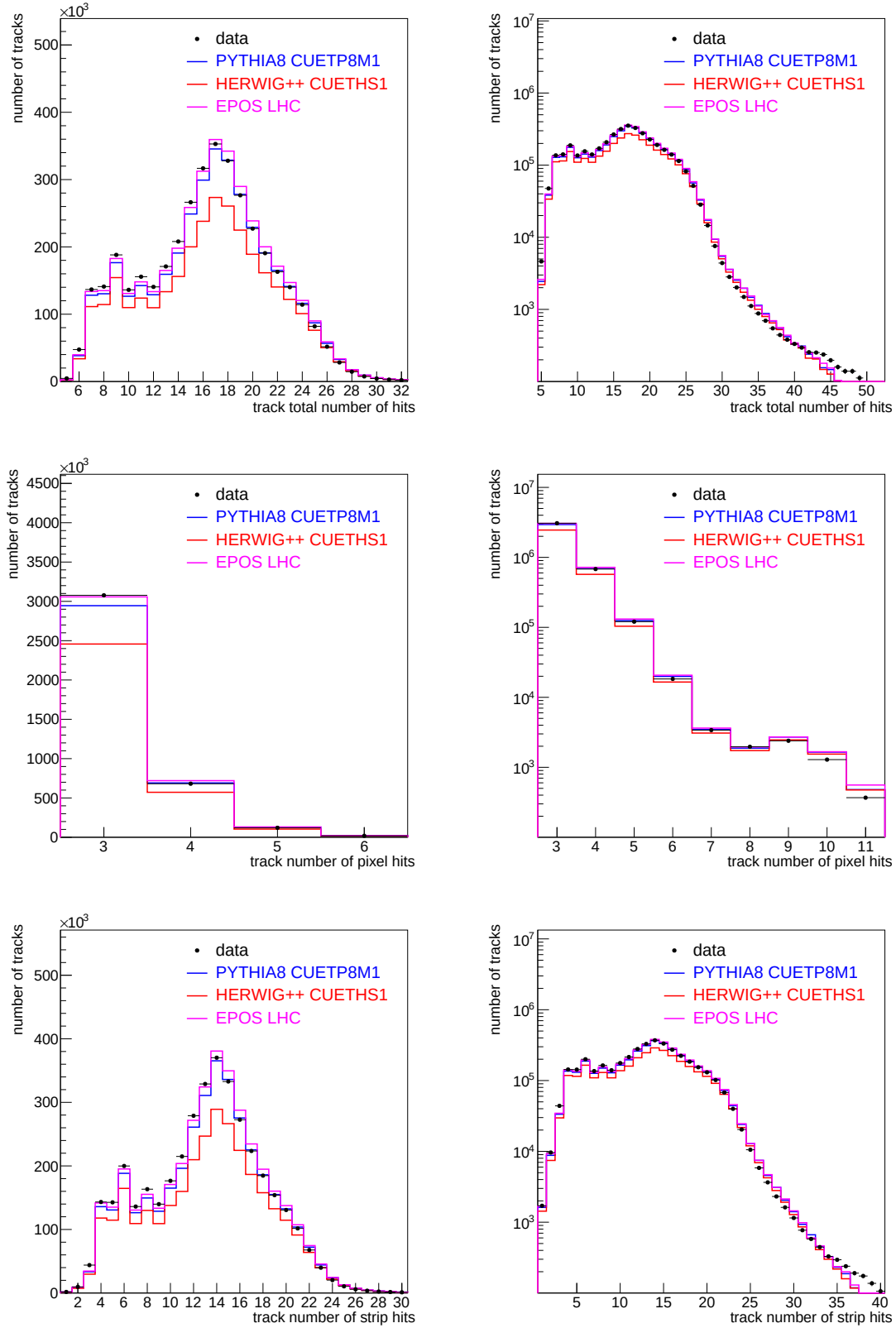


Figure 6.16: Distributions of the track hit multiplicity for all the tracker layers and for the pixel and strip layers separately. The data are compared to the predictions of the Monte Carlo event generators PYTHIA8, HERWIG++ and EPOS.

6.3 HF towers

The performance of the HF rechits and HF towers at low energies is affected by electronic noise and beam-induced background. It is important to determine the lowest reliable measured value for the energy deposition coming from the actual pp interaction and not from the aforementioned sources. For that purpose, three different event samples selected with the different BPTX triggers described before, are used. A sample of events selected online by the *no BPTX* trigger and corresponding to the absence of circulating beams in the LHC is used to determine the electronic noise in the HF cells and towers. Samples of events selected online by the *BPTX plus* only or *BPTX minus* only trigger and corresponding to the presence of only one beam circulating in the LHC in one or the other direction are used to determine the beam-induced background.

In the run 251721 used in the analysis, these triggers show some correlation with the Zero Bias trigger, and the activity in HF can not be determined in a reliable way. Another run for which the HF conditions are identical has therefore been used. The run 247324 satisfies this request and is used to determine the electronic noise and beam-induced background.

Figure 6.17 shows the comparison between the three event samples mentioned above, for the HF rechits and HF towers, separately for the HF with $\eta < 0$ (HFminus) and for HF with $\eta > 0$ (HFplus). Figure 6.18 shows the comparison of the signal in the different samples for noise and background studies, the nominal signal sample and the Monte Carlo simulations. In the same figure the corresponding energy cuts are represented with vertical dashed lines.

It is important to notice that the Monte Carlo simulations do not take into account the electronic noise nor the beam-induced background of the HF calorimeters. That gives large discrepancies between the simulation and the recorded data as shown in figures 6.19 and 6.20, where the multiplicity distribution of rechits and towers for increasing energy threshold is presented.

Table 6.2(6.3) shows the fraction of cells(towers) above a energy threshold of 3 GeV(5 GeV) in each of the event samples.

Fractions of events with at least one active calorimeter cell [%]			
sample	activity in HF plus	activity in HF minus	activity in HF plus OR HF minus
no BPTX	0.38 ± 0.03	0.15 ± 0.02	0.53 ± 0.04
BPTX plus	0.38 ± 0.03	0.17 ± 0.02	0.54 ± 0.04
BPTX minus	0.50 ± 0.03	0.12 ± 0.02	0.62 ± 0.04

Table 6.2: Fractions of events with at least one calorimeter cell with an energy value above a threshold of 3 GeV in the background dominated samples.

The present analysis makes use of the HF towers for the event selection as described in section 5.3.4, and from the studies it is concluded that 5 GeV is the most suitable tower energy threshold value.

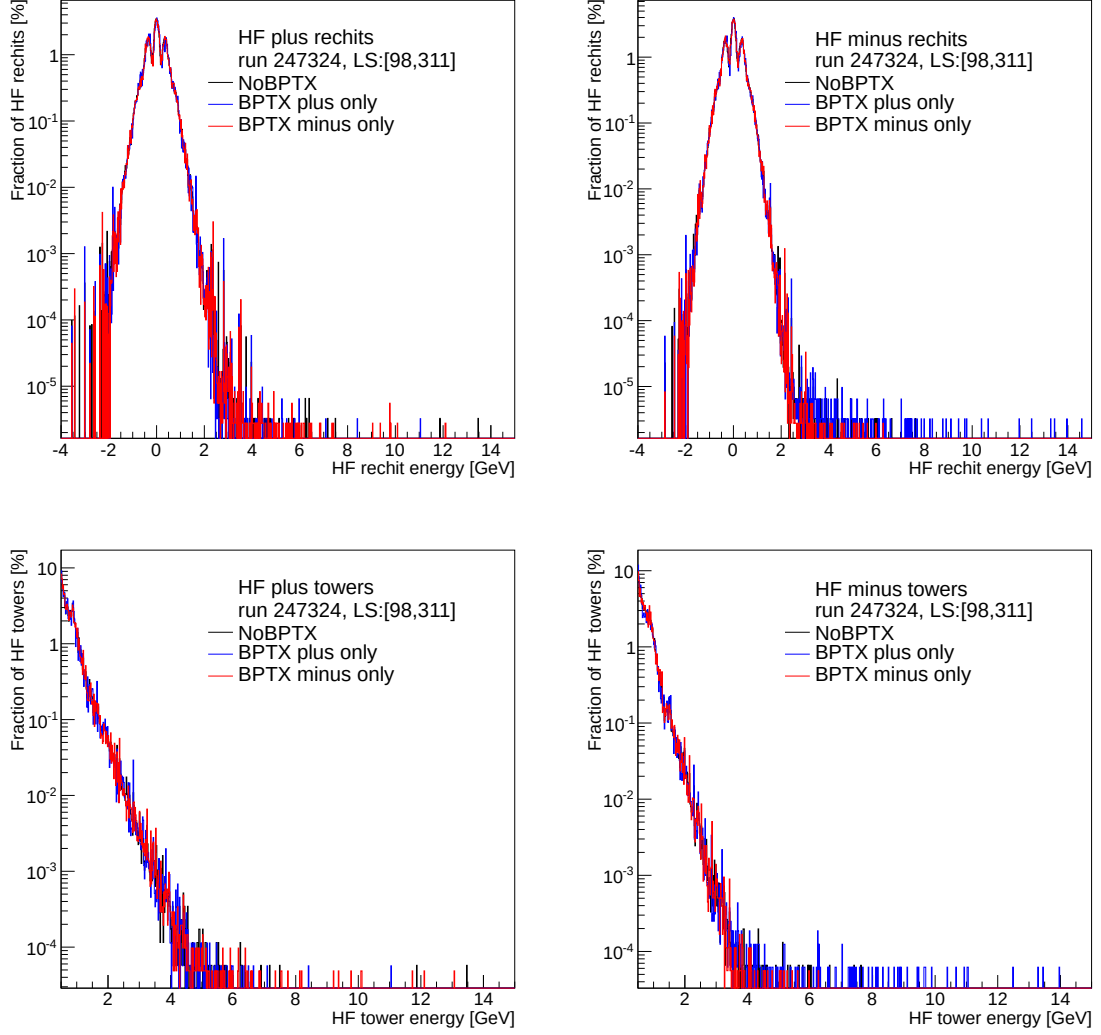


Figure 6.17: Distributions of the calorimeter cell energy (top) and calorimeter tower energy (bottom) for both HF calorimeters (HF minus at the left and HF plus at the right), for the event samples dominated by the electronic noise (*no BPTX*) and the beam-induced background (*BPTX plus only* or *BPTX minus only*).

Fractions of events with at least one active calorimeter tower [%]			
sample	activity in HF plus	activity in HF minus	activity in HF plus OR HF minus
no BPTX	0.08 ± 0.02	0.04 ± 0.01	0.12 ± 0.02
BPTX plus	0.08 ± 0.01	0.05 ± 0.01	0.13 ± 0.02
BPTX minus	0.09 ± 0.01	0.02 ± 0.01	0.11 ± 0.02

Table 6.3: Fractions of events with at least one calorimeter tower with energy above a threshold of 5 GeV in the background dominated samples.

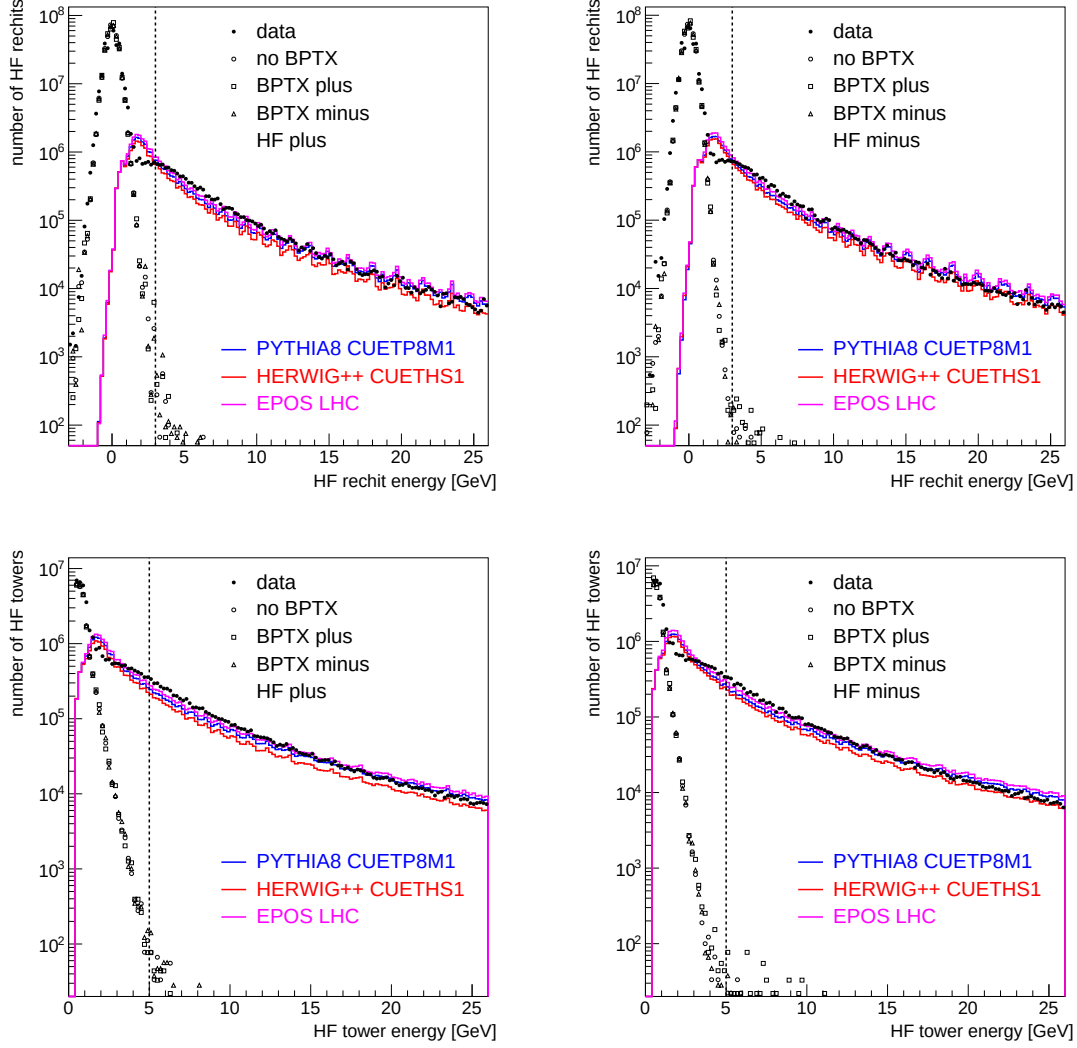


Figure 6.18: Distributions of the cell energy (top) and tower energy (bottom) for both HF calorimeters (HF minus at the left and HF plus at the right), for the signal event sample and the event samples dominated by the electronic noise (*no BPTX*) and the beam-induced background (*BPTX plus only* or *BPTX minus only*). The data are compared to the predictions of the Monte Carlo event generators PYTHIA8, HERWIG++ and EPOS.

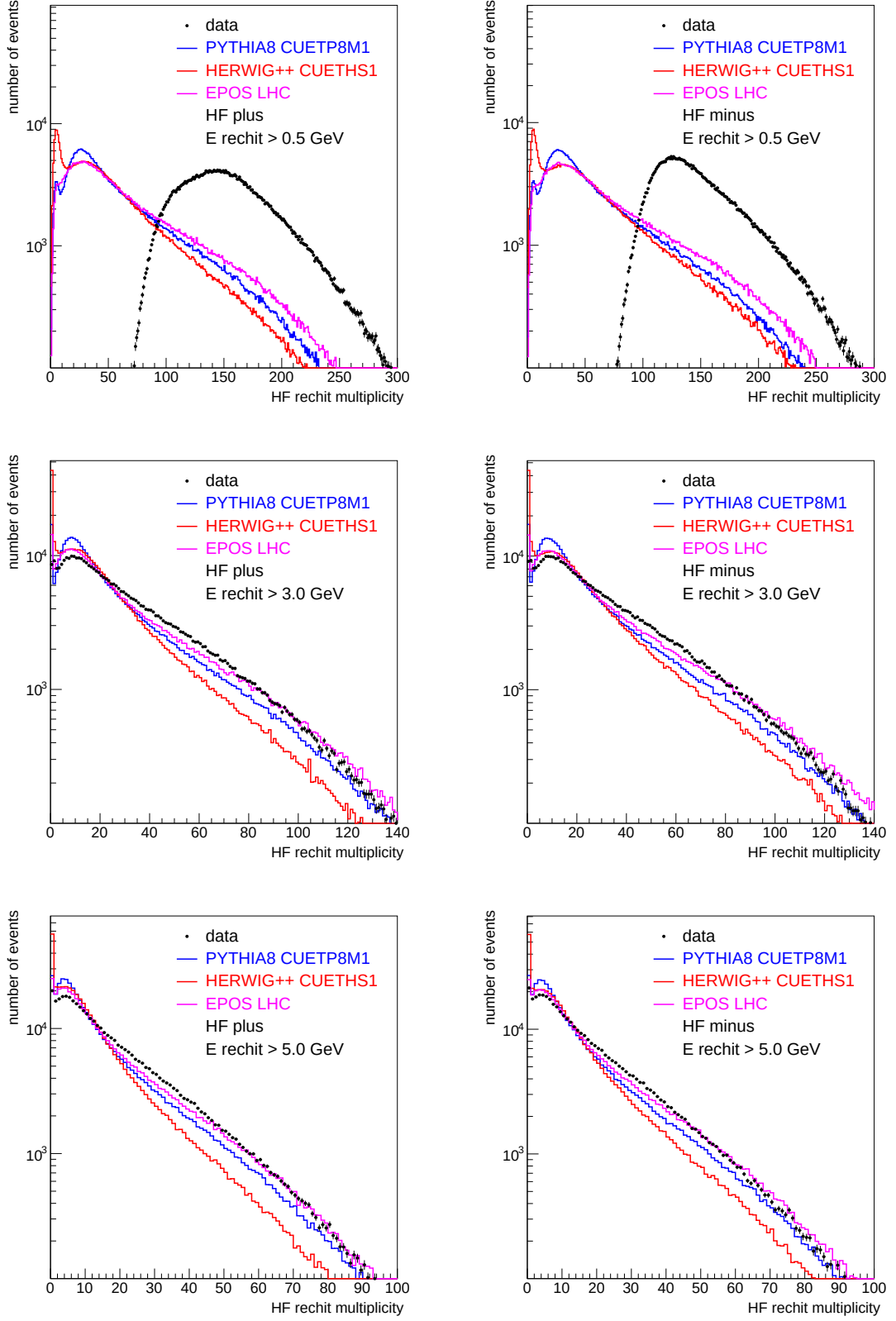


Figure 6.19: Distributions of the cell multiplicity for both HF calorimeters (HF minus at the left and HF plus at the right) for increasing values of the threshold applied to the cell energy. From top to bottom the thresholds are set to 0.5 GeV, 3 GeV and 5 GeV. The data are compared to the predictions of the Monte Carlo event generators PYTHIA8, HERWIG++ and EPOS.

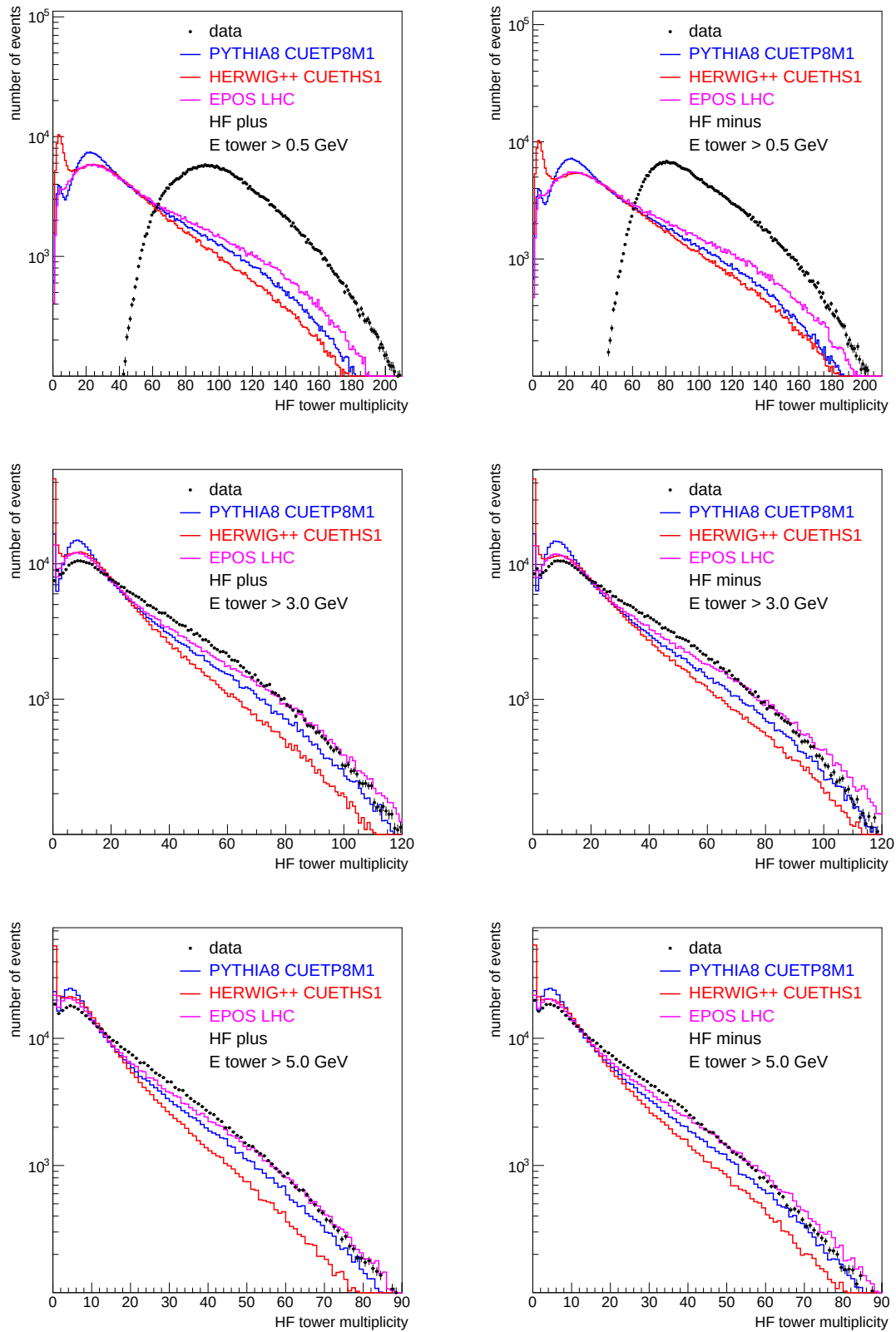


Figure 6.20: Distributions of the tower multiplicity for both HF calorimeters (HF minus at the left and HF plus at the right) for increasing values of the threshold applied to the tower energy. From top to bottom the thresholds are set to 0.5 GeV, 3 GeV and 5 GeV. The data are compared to the predictions of the Monte Carlo event generators PYTHIA8, HERWIG++ and EPOS.

CHAPTER

7

DETECTOR EFFECTS

Contents

7.1	Matching	86
7.2	Resolution effects	87
7.3	Response matrices	90
7.4	Acceptance, background, purity and stability	91
7.4.1	Pileup effects	93
7.4.2	Full event selection	97

In this chapter, the effects induced by the detector on the measured quantities are studied.

Due to the imperfect nature of the detectors, every time a measurement is performed, the measured value will be in general different from the true value. This deviation can be attributed to different sources. For instance, the miscalibration of the detector, the wrong or not functioning detector sections, or the limited detector resolution are some of the most common sources. The miscalibration of the detector is among the simplest detector effects to tackle, because it only induces a systematic shift of the measurement, which introduces non to very small limitations on the measurements. On the other hand, the detector resolution does impose limitations in the measurement. One way to determine these effects is with the help of a MC simulation that contains the full detector simulation. Then, a matching between the generator and detector levels is performed as explained in Sec. 7.1. After the matching is performed, the detector resolution is studied (Sec. 7.2). The concepts of response matrices (Sec 7.3), as well as acceptance, background, purity and stability (Sec. 7.4) are as well introduced in this section.

7.1 Matching

In order to study the different detector effects, a correspondence between generator level (i.e. stable particle level) and reconstructed level (i.e. detector level) needs to be established. This correspondence is commonly known as "matching". The matching is done in a 2-dimensional space, where the variables are different than the observable of interest. For instance, the study of the detector effects on the pseudorapidity of the particles is done by matching particles to tracks in a $p_T - \varphi$ space. The analogous study as a function of the transverse momentum of the particles is done by matching particles to tracks in a $\eta - \varphi$ space.

Figure 7.1 shows the distribution of the separation $\Delta R = \sqrt{\Delta p_T^2 + \Delta \varphi^2}$ and $\Delta R = \sqrt{\Delta \eta^2 + \Delta \varphi^2}$ between all the tracks associated to the primary vertex and all the stable charged particles obtained from the predictions of the Monte Carlo event generator PYTHIA8. Δp_T , $\Delta \varphi$ and $\Delta \eta$ respectively represent the separation in transverse momentum, azimuthal angle and pseudorapidity between the tracks and the stable charged-particles with a transverse momentum $p_T > 0.5$ GeV/c in the pseudorapidity region $|\eta| < 2.4$. The peak at low values of ΔR corresponds to those tracks matched to the actual stable charged particles from which they originate, while the tail extending to high ΔR values corresponds to all the possible wrong matching combinations, known as combinatorial background. The matching criteria to associate a track to a stable charged particle is chosen to be $\Delta R(\eta, \varphi) < 0.05$ and $\Delta R(p_T, \varphi) < 0.07$.

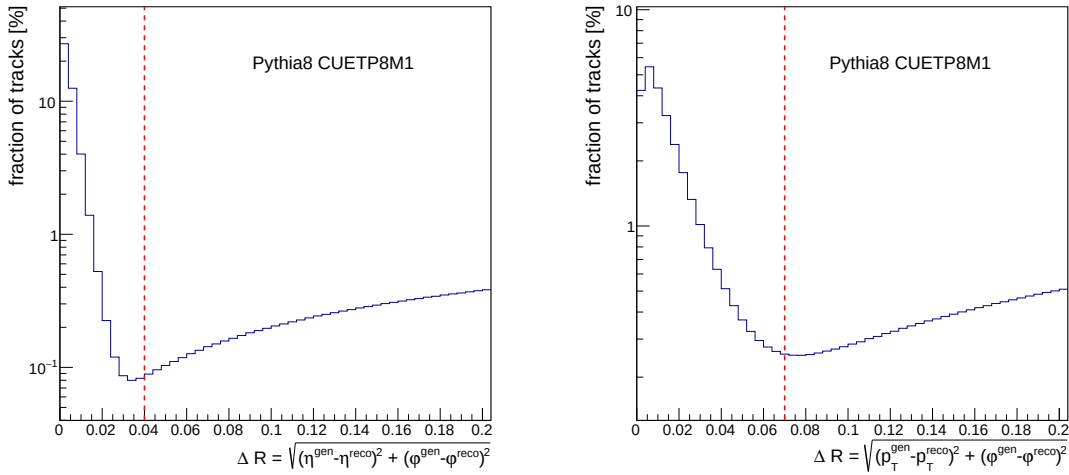


Figure 7.1: Distribution of the $\Delta R = \sqrt{\Delta \eta^2 + \Delta \varphi^2}$ (left) and $\Delta R = \sqrt{\Delta p_T^2 + \Delta \varphi^2}$ separations (right) between all the tracks associated to the primary vertex and all the stable charged particles obtained from the predictions of the Monte Carlo event generator PYTHIA8. The dashed vertical lines show the cut value chosen to select a matched pair.

In the case of studying detector effects as a function of observables that are event based, the concept of matching varies, as it is the case for the multiplicity distributions. For these kind of observables, the matching is done by requiring the events to fulfil the selection criteria

at stable particle and detector level, simultaneously. From the fact that the events are either selected or rejected, and no wrong matchings can be done, there is no combinatorial background and no cut value is needed.

7.2 Resolution effects

Every time a well calibrated detector measures a quantity, the obtained value will be distributed according to a gaussian distribution centred at the true value. In the case of a miscalibration, the same distribution would be obtained with a shift proportional to the value of the miscalibration. The width of the gaussian distribution is regarded as the resolution of the detector. As the resolution of the detector increases, the width of the gaussian becomes smaller. In general, the resolution behaves differently for different observables. It can be a constant value over the full measured range, or it can vary. For that reason, different approaches are implemented for the different observables of this analysis.

The limited resolution of the detector imposes an important constraint in the measurements, dictating the minimum size of the bin-width that can be measured. If the bin width chosen for the distribution is too small with respect to the corresponding resolution, then a large migration between bins is observed. One way to avoid this problem is to chose a bin-width at least two to three times larger than the resolution.

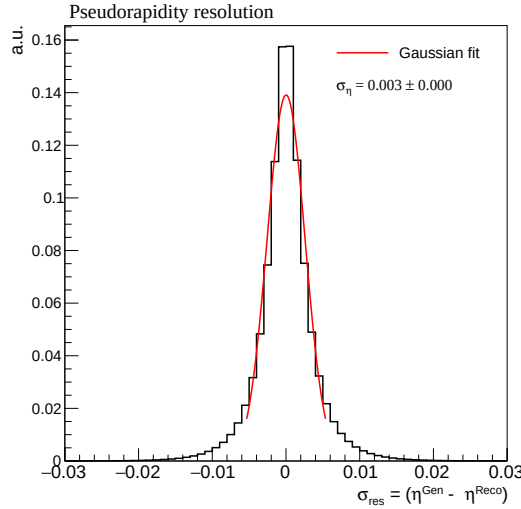


Figure 7.2: Detector resolution as a function of the η of charged particles calculated from Monte Carlo simulations with PYTHIA8 CUETP8M1.

Pseudorapidity resolution

In the case of pseudorapidity, the resolution is calculated as the difference between the generated and the reconstructed values, $\eta_{gen} - \eta_{reco}$, for the matched particle-track pairs with $\Delta R = \sqrt{\Delta p_T^2 + \Delta \varphi^2} < 0.05$.

Multiplicity ranges for resolution			
Range	[min - max]	Range	[min - max]
1	1 – 5	11	51 – 55
2	6 – 10	12	56 – 60
3	11 – 15	13	61 – 65
4	16 – 20	14	66 – 70
5	21 – 25	15	71 – 75
6	26 – 30	16	76 – 80
7	31 – 35	17	81 – 85
8	36 – 40	18	86 – 90
9	41 – 45	19	91 – 95
10	46 – 50	20	96 – ∞

Table 7.1: Multiplicity ranges for the study of detector resolution.

The detector has an approximately flat response as a function of η . Hence, the resolution as a function of the η of the generated charged particle is also flat, and only a single resolution value for the full η range is needed.

Figure 7.2 shows the η resolution of charged particles. A gaussian of $\sigma_\eta = 0.002$ describes the curve.

Multiplicity resolution

The detector resolution is not homogeneous in terms of the multiplicity, i.e. it deteriorates as the number of particles increases. This is partially related to the increasing complexity of assigning all the hits in the tracker to the actual tracks. As the multiplicity increases, the probability to assign unrelated hits to a track becomes higher, which affects the correct calculation of the number of true particles, in general reconstructing less tracks than real particles.

Therefore, for the multiplicity distribution, the resolution is obtained the same way as for the pseudorapidity resolution, but the distribution is split in 20 ranges of size $\Delta M = 5$, where M is the multiplicity. Only the last bin has a different size covering the range from 96 to infinity. The split of the resolution in multiple ranges helps to obtain simpler curves to fit instead of a very complex, wide range extended, single curve. In table 7.1 the multiplicity ranges are shown.

For instance, the mean of the resolution for the last multiplicity range, $96 - \infty$, is around 32. That means that in average, the events with true multiplicities between $96 - \infty$, are reconstructed with in average, 32 tracks less.

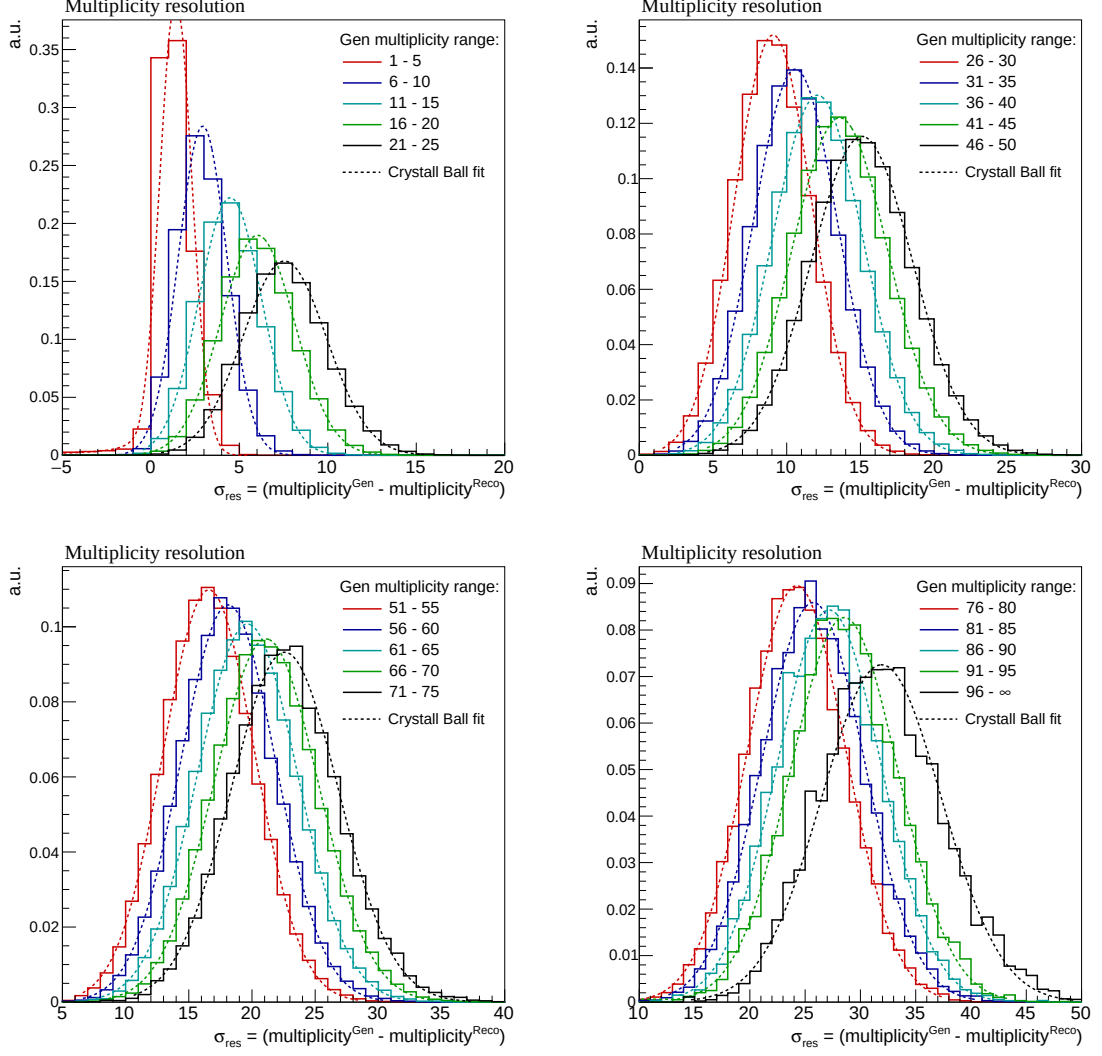


Figure 7.3: Detector resolution for the multiplicity of charged particles calculated from Monte Carlo simulations with PYTHIA8 CUETP8M1 for the different multiplicity ranges.

p_T ranges for resolution		
Range	[min - max) [GeV]	$\langle p_T \rangle$ [GeV]
1	0.5 – 1	0.70
2	1 – 2	1.35
3	2 – 3	2.38
4	3 – 4	3.40
5	4 – 5	4.41
6	5 – 6	5.43
7	6 – 7	6.44
8	7 – 8	7.45
9	8 – 9	8.45
10	9 – 10	9.46
11	10 – ∞	13.32

Table 7.2: Transverse momentum ranges for the study of detector resolution. The mean transverse-momentum in each range is displayed in the third column.

Transverse momentum resolution

For the p_T resolution a slightly different method is used. The first part is analogous to the procedure followed for the multiplicity, i.e. the resolution is studied in small ranges. In this case, 11 ranges are set as a function of p_T^{gen} , and the mean p_T per range is calculated. Table 7.2 shows these values.

Then, the resolution, relative to the true p_T^{gen} , is calculated as: $(p_T^{gen} - p_T^{reco})/p_T^{gen}$, for every range. For each of the relative resolution curves a gaussian distribution is fitted, and the corresponding $\sigma_{p_T}^i$ is extracted ($i=[1,11]$). All the extracted $\sigma_{p_T}^i$ are plotted as a function of the corresponding p_T mean, $\langle p_T \rangle^i$.

Finally, the plot is fitted and the obtained function is used to describe the relative transverse-momentum resolution. Figure 7.4 shows the relative transverse-momentum resolution as a function of p_T^{gen} , and the fitted function as a continuous line.

7.3 Response matrices

Response Matrices (RM) are a very useful way of representing the detector response as a function of a given observable. They do not only provide a graphical representation of what the detector measures with respect to the truth, but also play an important role in the correction of the measurement to the stable-particle level, as explained later in Chapter 8.

Once the collection of matched particle-track pairs has been created, the construction of the RM is straight forward. It only requires to fill a 2-dimensional histogram, in which the y-axis corresponds to the stable-particle level and the x-axis to the detector level, with respect to the observable of interest.

The RMs are constructed with a binning according to the criterion of the resolution stud-

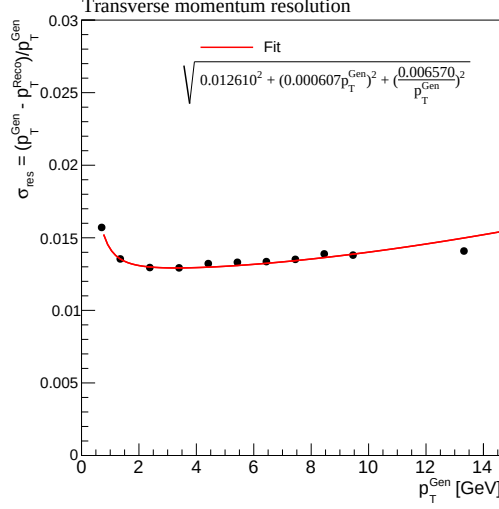


Figure 7.4: Relative detector resolution as a function of the transverse momentum of charged particles calculated from Monte Carlo simulations with PYTHIA8 CUETP8M1. The points represent the relative detector resolution in each range, and the continuous line shows the fitted function.

ies of Sec. 7.2. Still, it is important to notice that if large migrations coming from other sources than the resolution are present, as in the case of the multiplicity, an appropriate binning will only partially mitigate the problem, and those remaining migrations are corrected using the method explained in Chapter 8.

In figure 7.5, different RMs are shown for the “Inclusive” selection as a function of η , multiplicity and p_T of charged-particles, and for the leading p_T of charged-particles. The shown RMs are normalised to the number of entries in each row. In that way the migration fraction per stable particle bin is visible.

The RMs corresponding to the pseudorapidity and transverse momentum show a highly diagonal behaviour. This is a consequence of low migrations obtained through the choice of the binning and the exceptional tracking capabilities of the CMS experiment. This implies that more than 90% of the tracks measured in a given bin are particles that have actually been produced in that bin.

In the case of the multiplicity RM, a non-diagonal behaviour is observed. The explanation for this has already been discussed in Sec. 7.2.

7.4 Acceptance, background, purity and stability

Once the RMs are constructed, the variables to study the quality of the tracks as a function of the measured variables are calculated. The tracking efficiency, or acceptance, is defined as the fraction of stable charged-particles matched to a track compared to all stable particles.

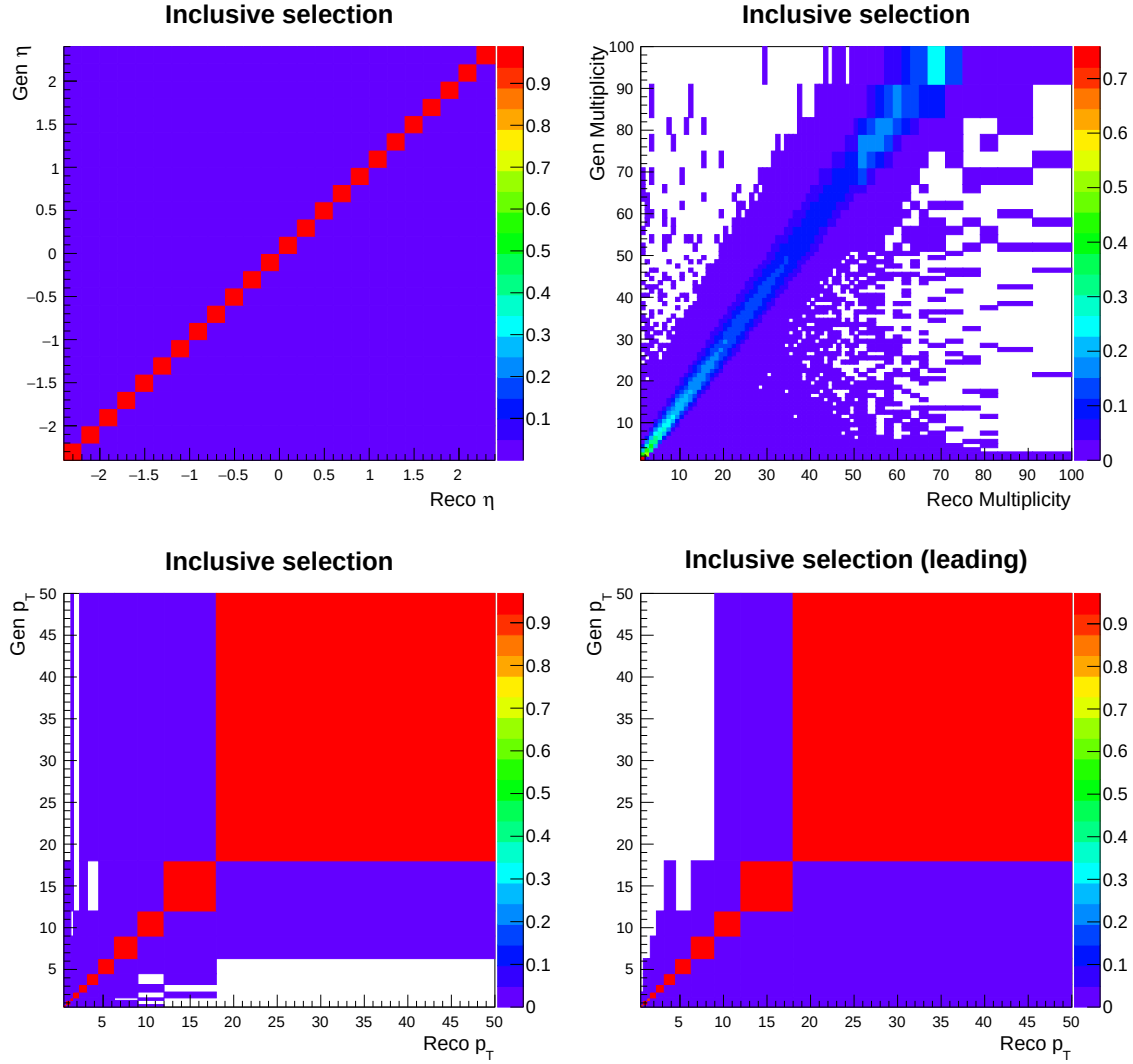


Figure 7.5: Response matrices for pseudorapidity (top left), multiplicity (top right) and transverse momentum of charged particles (bottom left), as well for the leading transverse-momentum charged-particles (bottom right) calculated from Monte Carlo simulations with PYTHIA8 CUETP8M1.

In general, the acceptance can be expressed as a function of the measured observables of the stable charged-particles, as follows:

$$\text{tracking efficiency}(X_{gen}) = \frac{N \text{ matched particles}(X_{gen})}{N \text{ all particles}(X_{gen})} \Big|_{\text{vertex selection}}, \quad (7.1)$$

where “N all particles” is the total number of stable charged-particles, “N matched particles” the number of stable charged-particles matched to a track and X_{gen} the observable of interest of the stable charged-particles. The subscript “vertex selection” indicates that the selection at detector level is only based on the vertex requirement. The tracking fake rate, or background, is defined as the fraction of tracks that cannot be matched to a stable charged-particle and it is expressed as a function of the measured observable of the tracks:

$$\text{tracking fake rate}(X_{reco}) = \frac{N \text{ unmatched tracks}(X_{reco})}{N \text{ all tracks}(X_{reco})} \Big|_{\text{vertex selection}}, \quad (7.2)$$

where “N all tracks” is the total number of tracks, “N unmatched tracks” the number of tracks that cannot be matched to a stable charged-particle and X_{reco} the measured observable of the tracks. The tracking purity is defined as the fraction of tracks that are reconstructed in the same bin as the stable charged-particle to which they are matched:

$$\text{tracking purity}(X) = \frac{N \text{ matched tracks}(X_{reco}, X_{gen} \in X \text{ bin})}{N \text{ matched tracks}(X_{reco})} \Big|_{\text{vertex selection}}, \quad (7.3)$$

where “N matched tracks” is the number of tracks matched to a stable charged-particle. The tracking stability is defined as the fraction of stable charged-particles that are generated in the same bin as the track to which they are matched:

$$\text{tracking stability}(X) = \frac{N \text{ matched particles}(X_{gen}, X_{reco} \in X \text{ bin})}{N \text{ matched particles}(X_{gen})} \Big|_{\text{vertex selection}}. \quad (7.4)$$

These quality variables are used in the following to study two important effects: the effects of the pileup on the tracking, and the effects of the event selection based on the HF activity, both as a function of the observables of interest.

7.4.1 Pileup effects

In order to illustrate the effects of the pileup on the track reconstruction and the effects of the requirement to have exactly one reconstructed primary vertex on the track and event reconstruction, three different scenarios are shown.

In the scenario corresponding to the nominal conditions of the analysis, the Monte Carlo simulations have an average number of collisions per bunch crossing equal to 1.8 and the requirement to have exactly one reconstructed primary vertex is applied. In the second scenario, the simulations have the same pileup conditions but the vertex requirement is replaced by the condition to have at least one reconstructed primary vertex in the event. Finally in the third scenario, the simulations have no contributions from pileup and the requirement to have exactly one reconstructed primary vertex is applied.

The tracking efficiency, fake rate, purity, and stability are shown in figures 7.6, 7.7 and 7.8 for the 3 scenarios, respectively, as a function of η , multiplicity and transverse momentum obtained from the Monte Carlo event generator PYTHIA8. The corresponding plots for the leading p_T charged-particles are excluded since the behaviour is the same as for the p_T charged-particles but with less statistics.

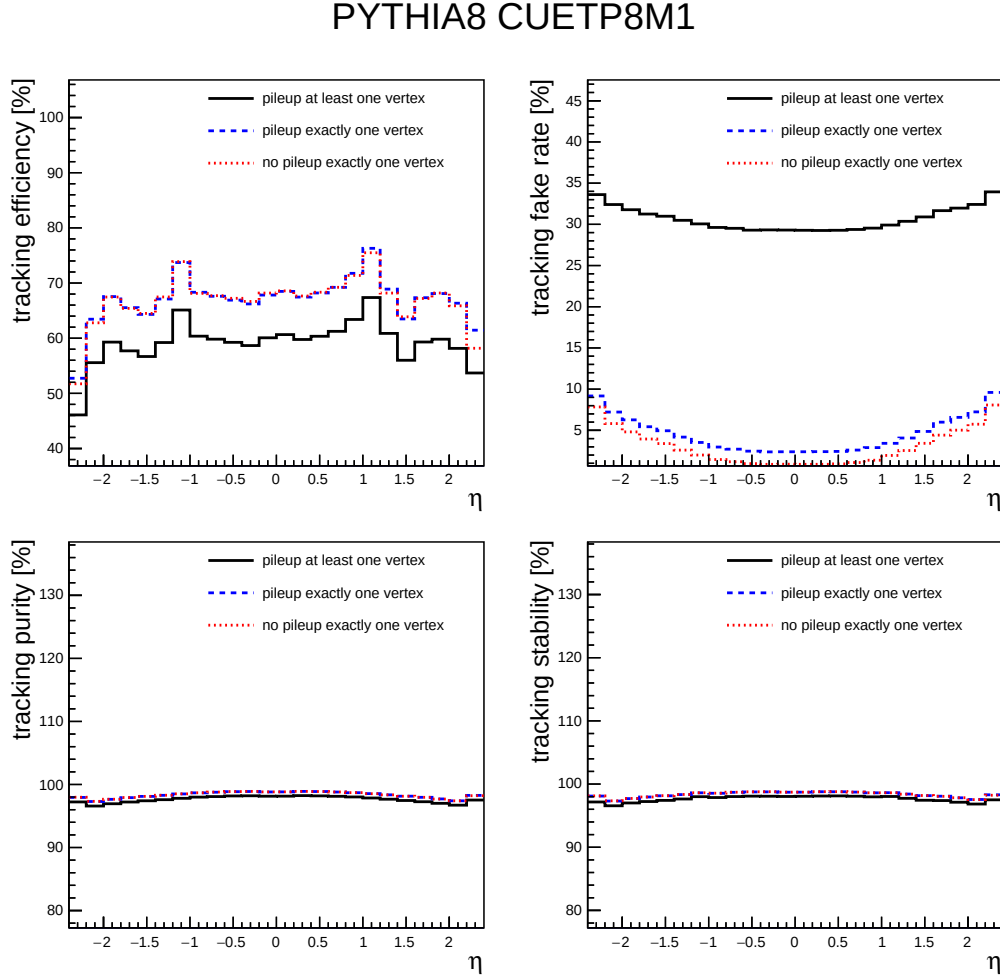


Figure 7.6: Efficiency, fake rate, purity and stability of the track reconstruction for the 3 scenarios described in the text as a function of pseudorapidity and for the predictions of the Monte Carlo event generator PYTHIA8.

In the case of η , the purity and stability are by definition almost insensitive to the pileup conditions, showing values larger than 98 % over the full η range. The efficiency shows a dependence on the η of the track, with an average value of the order of 70% in the central region $|\eta| < 1$, decreasing to 50 – 60% at the edge of the tracker acceptance for the scenario corresponding to the nominal conditions of the analysis. While the efficiency and fake rate are affected by the pileup contributions, the requirement to have exactly one reconstructed

PYTHIA8 CUETP8M1

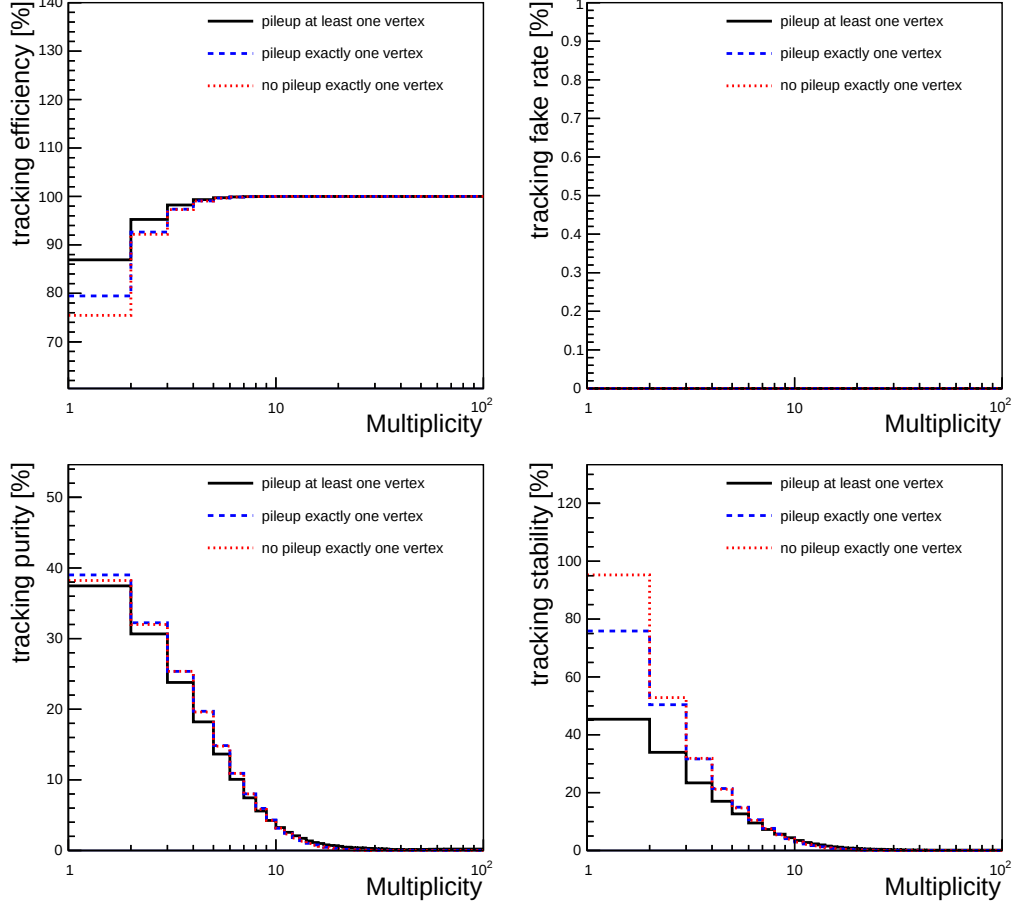


Figure 7.7: Efficiency, fake rate, purity and stability of the track reconstruction for the 3 scenarios described in the text as a function of multiplicity and for the predictions of the Monte Carlo event generator PYTHIA8.

vertex in the event enables to efficiently suppress their effects. The efficiency is restored to the value observed in the scenario without pileup contributions, and the fake rate is suppressed by a factor of the order of 4.

For the multiplicity, the efficiency reaches 100% starting from multiplicity values around 6 and drops down to 80% for lower values. The reason for this is, that there are more pileup events with lower multiplicities from which miss-matching can occur in each bunch crossing. In contrast, there are fewer pileup events with large multiplicities per bunch crossing, and a miss-match is less probable. The nominal selection gives slightly lower efficiency values than in the case without pileup. Since the low region of the spectrum has a greater pileup contamination, the matching between the detector- and the stable-particle level degenerates. Given that the multiplicity is an event-by-event, observable there are no "unmatched" objects in the event. Hence, the fake rate is always zero. The purity and stability do not provide

PYTHIA8 CUETP8M1

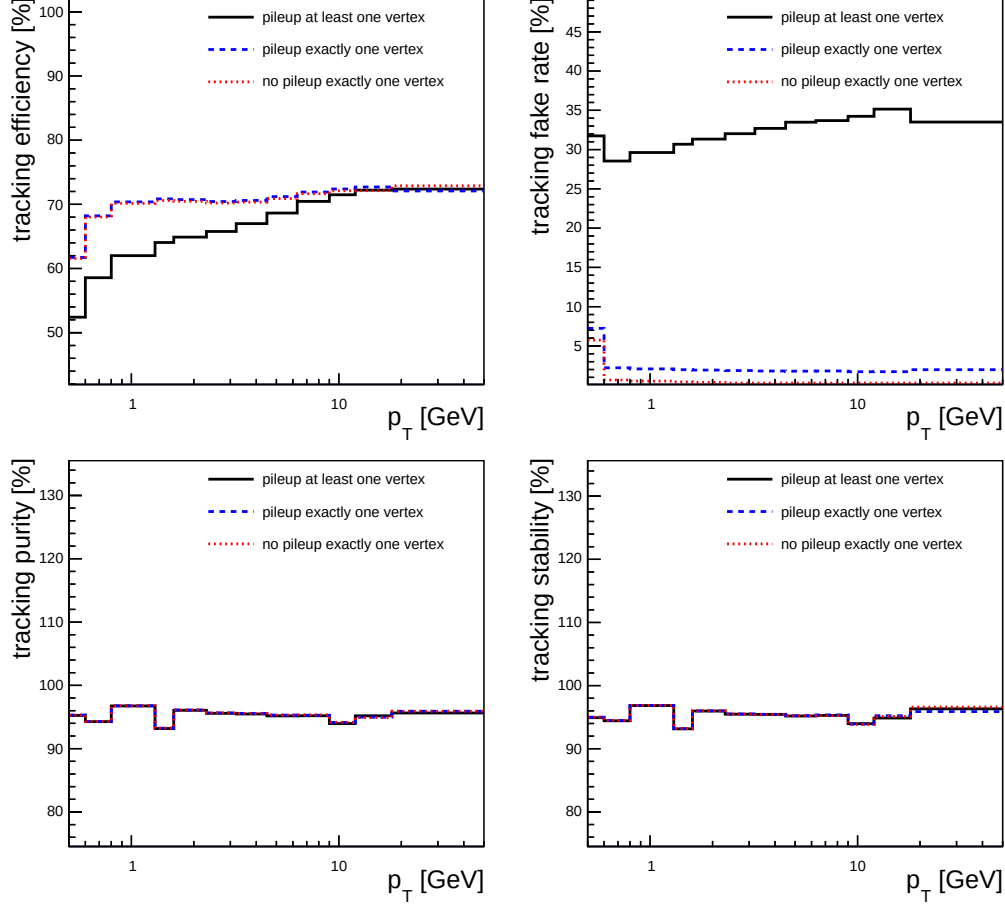


Figure 7.8: Efficiency, fake rate, purity and stability of the track reconstruction for the 3 scenarios described in the text as a function of transverse momentum and for the predictions of the Monte Carlo event generator PYTHIA8.

much information here, since these quantities are calculated with respect to the entries in the diagonal of the response matrix, and as shown in Sec. 7.3, the multiplicity suffers from large migration effects that increase with the multiplicity.

In the case of the p_T , the tracking efficiency reaches a plateau at 70% starting from 0.8 GeV. The efficiency is more affected by the pileup as the p_T decreases, going down to 50% in the scenario with at least one reconstructed vertex. This is the expected behaviour for events with multiple vertices, due to the increased number of produced particles for which the tracking algorithm has to reconstruct their trajectories. The reconstruction of the trajectories also becomes more complicated as the p_T of the particles decreases, due to the multiple scatterings with the detector material and the more pronounce trajectory curvatures within the magnetic field. The nominal selection of exactly one reconstructed vertex is able to suppress this effect. The fake rate for the scenario with at least one reconstructed vertex

goes up to 35% and again the nominal selection makes it close to the no pileup scenario, with an average value of 2%. In the case of the purity and stability the conclusions are the same as for the pseudorapidity case.

7.4.2 Full event selection

The quality of the event selection based on the activity in the HF calorimeters is studied in this section in terms of efficiency, fake rate, purity, and stability. The full event selection is now applied in order to define the Inelastic-enhanced, NSD-enhanced and SD-enhanced event samples. The global efficiency, fake rate, purity, and stability combine thus the behaviour of the track reconstruction and the behaviour of the selection based on the HF activity. The quantities are also presented for an inclusive event sample, for which the selection of events at detector level is only based on the vertex requirement. In this case the efficiency, fake rate, purity, and stability are equivalent to those of the tracking reconstruction presented in the previous section. The different quantities are shown in figures 7.9 to 7.14 for the 3 different selections previously mentioned. The plots show the predictions of the MC event generator PYTHIA8 as a function of the charged-particle η , multiplicity and transverse momentum, such that the effects of the pileup on the HF selection and the effect of the requirement to have exactly one reconstructed primary vertex in the event are illustrated.

The global efficiency shows a dependence on the track η . It reflects the tracking efficiency, which has an average value of the order of 70% in the central region $|\eta| < 1$, and decreases to 50 – 60% at the edge of the tracker acceptance, for the scenario corresponding to the nominal conditions of the analysis. The global efficiency is nearly insensitive to the HF selection. The global fake rate is of the order of 2 – 9% for PYTHIA8, reaching the lowest value at the central part of the tracker detector. The global purity and stability are insensitive to the pileup conditions and HF selection, and have values larger than 96% over the full η range. While the efficiency and fake rate are again affected by the pileup contributions, the requirement to have exactly one reconstructed vertex in the event enables to efficiently suppress their effects. The efficiency is restored to the value observed in the scenario without pileup contributions, and the fake rate is suppressed by a factor of the order of 3 to 4.

In the case of the multiplicity distributions, the effects of the different pileup conditions for the efficiency and purity are small, and they are only noticeable in the low end of the spectra. The stability presents the largest variations between the pileup conditions, mainly in the low multiplicity region with variations of up to 40%. The global efficiency, purity and stability do not present a large dependence on the HF selection.

For the p_T distributions, the global efficiency, purity, and stability are insensitive to the HF selection. The average value as a function of the charged-particle p_T for the global efficiency is around 70% dropping down to 60% at the lowest bin. The fake rate has an average value of 2% for the Inclusive, Inelastic-enhanced and NSD-enhanced samples and around 1.5% for the SD-enhanced sample. All the HF selection have a peak in the fake rate for the first bin reaching 7% fake rate. This behaviour at low- p_T shows that the tracking algorithm becomes inefficient to reconstruct the tracks correctly at such low- p_T values. Finally, the purity and stability show a constant behaviour as a function of the p_T with an average value of 95%.

The SD sample starts to suffer from low statistics in the high- p_T region where fluctuations start to be present. The nominal selection is able to bring the efficiency up to the values corresponding to the sample without pileup.

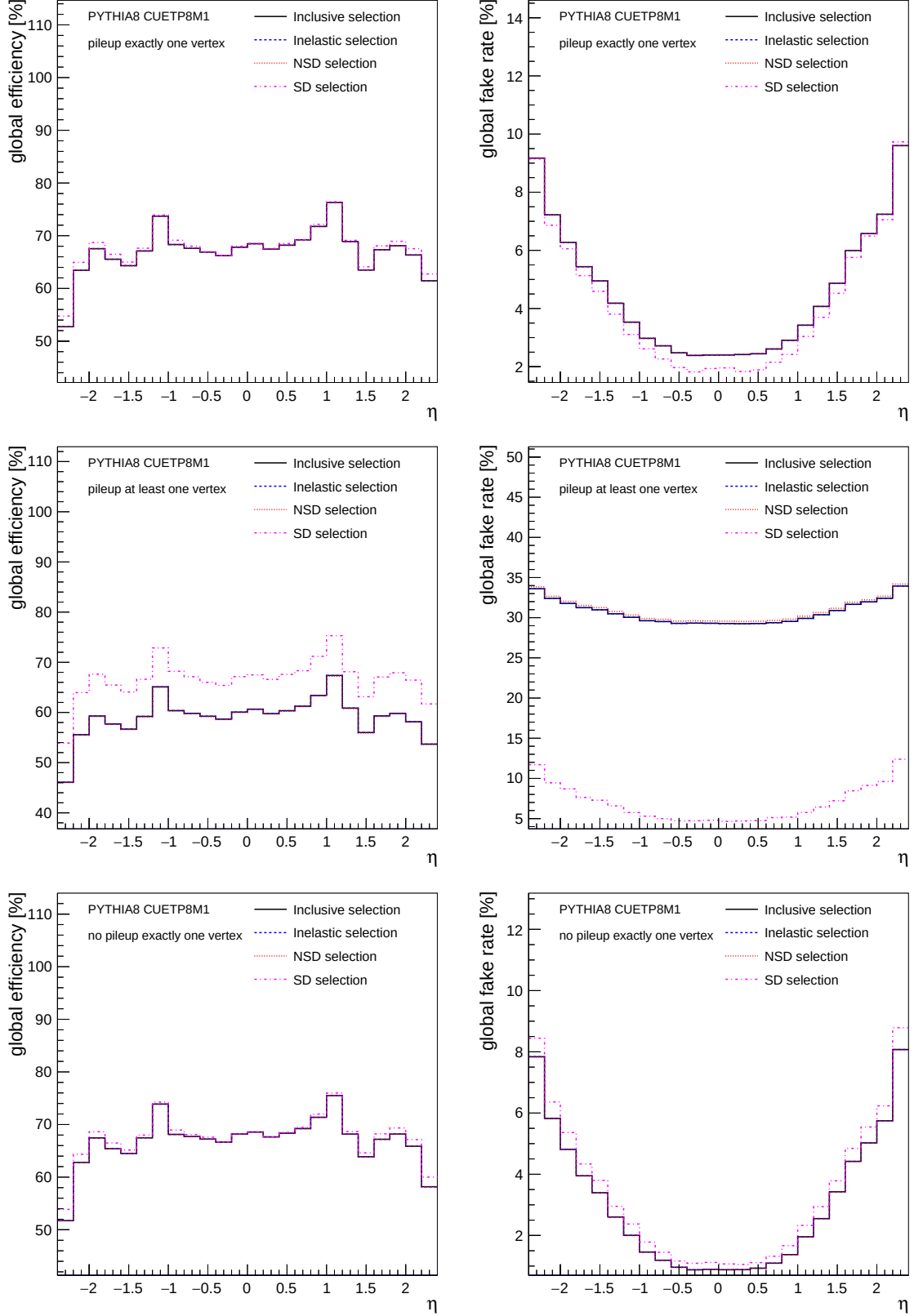


Figure 7.9: Efficiency (left) and fake rate (right) of the η distributions for the full event selection. The predictions of the Monte Carlo event generators PYTHIA8 are shown for the scenario corresponding to the nominal analysis, the scenario with an average number of collisions per bunch crossing equal to 1.8 but the condition to have at least one vertex in the event, and the scenario without pileup contributions and the condition to have exactly one vertex, respectively from top to bottom.

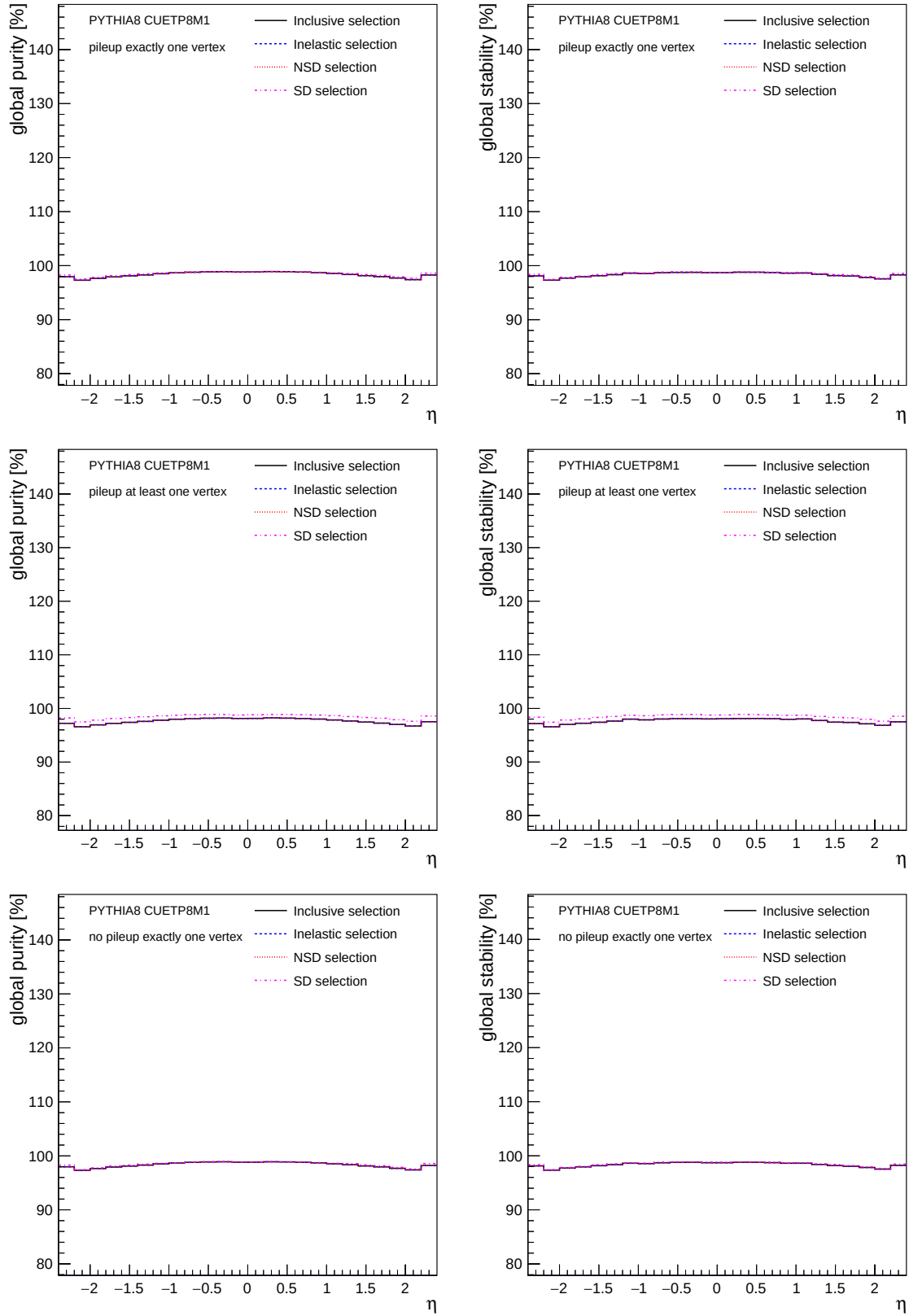


Figure 7.10: Purity (left) and stability (right) of the pseudorapidity distributions for the full event selection. The predictions of the Monte Carlo event generators PYTHIA8 are shown for the scenario corresponding to the nominal analysis, the scenario with an average number of collisions per bunch crossing equal to 1.8 but the condition to have at least one vertex in the event, and the scenario without pileup contributions and the condition to have exactly one vertex, respectively from top to bottom.

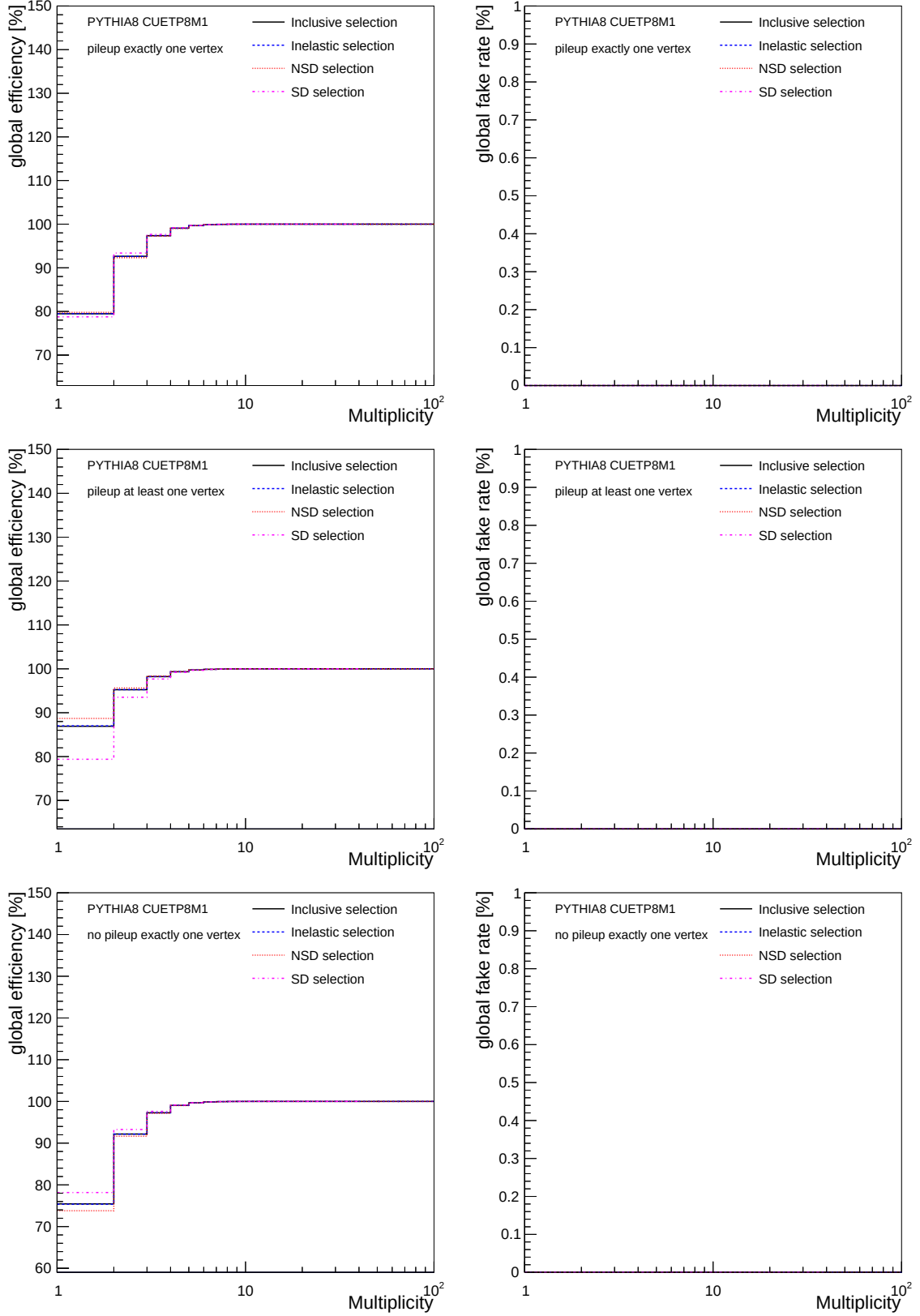


Figure 7.11: Efficiency (left) and fake rate (right) of the multiplicity distributions for the full event selection. The predictions of the Monte Carlo event generators PYTHIA8 are shown for the scenario corresponding to the nominal analysis, the scenario with an average number of collisions per bunch crossing equal to 1.8 but the condition to have at least one vertex in the event, and the scenario without pileup contributions and the condition to have exactly one vertex, respectively from top to bottom.

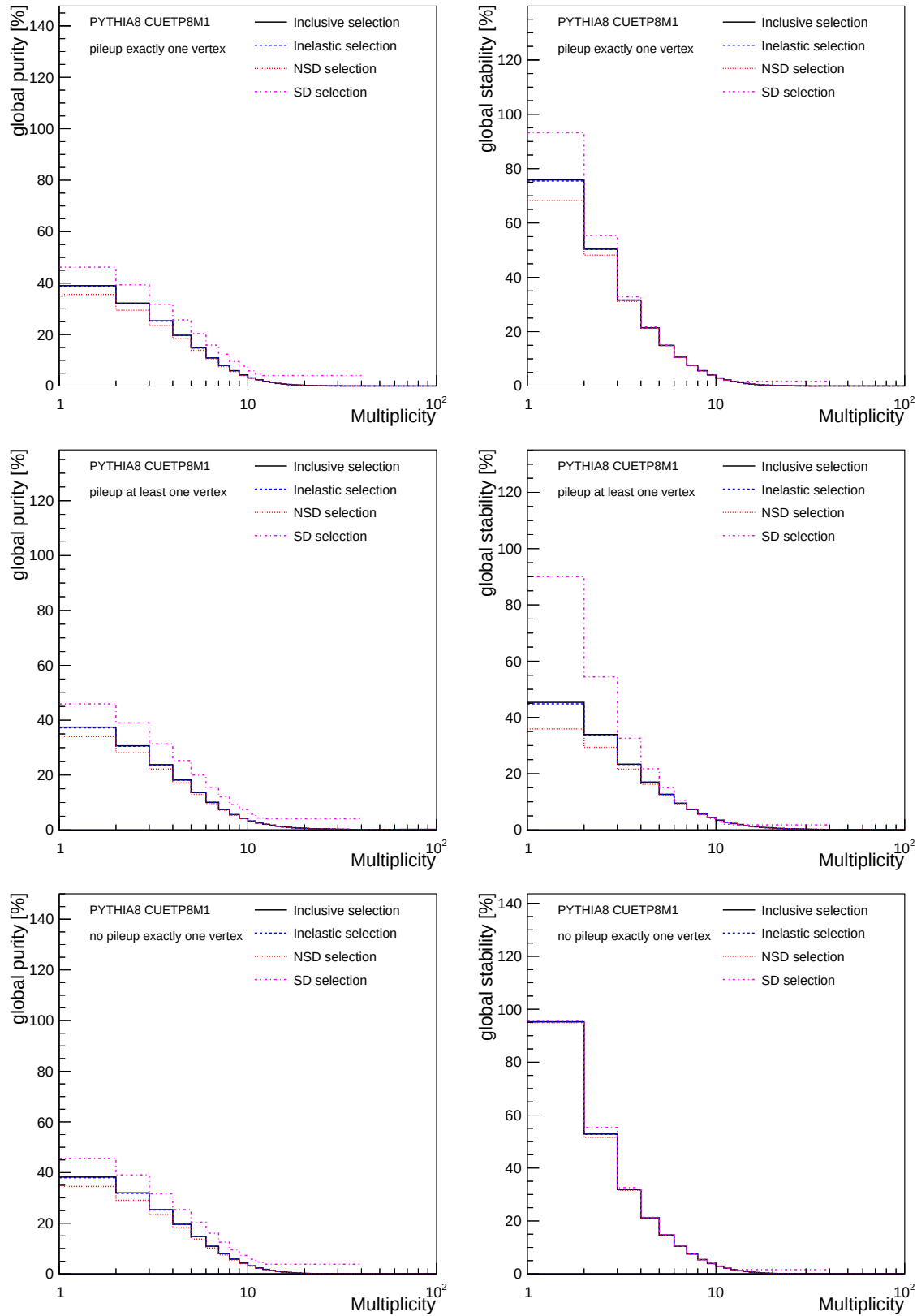


Figure 7.12: Purity (left) and stability (right) of the multiplicity distributions for the full event selection. The predictions of the Monte Carlo event generators PYTHIA8 are shown for the scenario corresponding to the nominal analysis, the scenario with an average number of collisions per bunch crossing equal to 1.8 but the condition to have at least one vertex in the event, and the scenario without pileup contributions and the condition to have exactly one vertex, respectively from top to bottom.

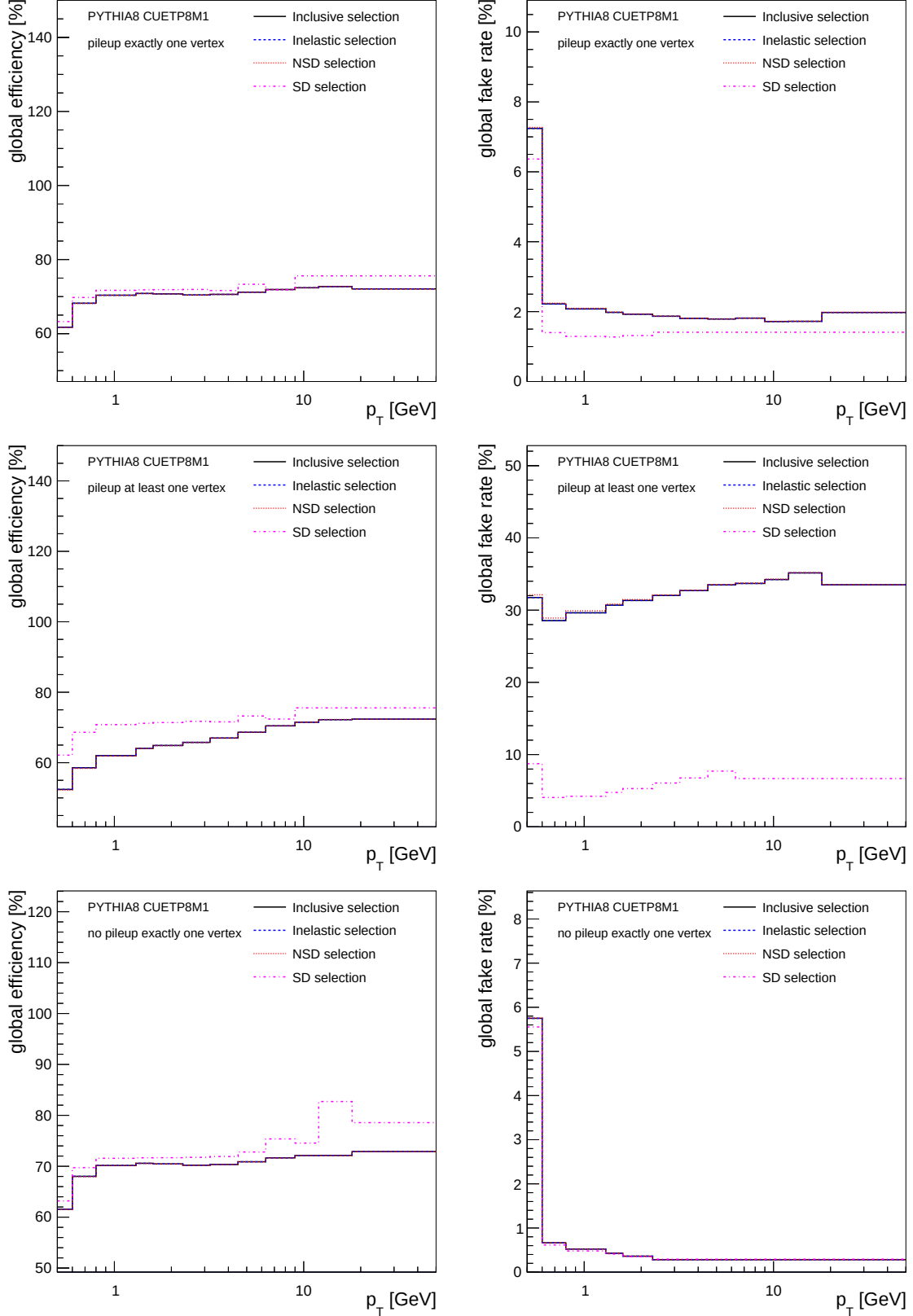


Figure 7.13: Efficiency (left) and fake rate (right) of the transverse-momentum distribution for the full event selection. The predictions of the Monte Carlo event generators PYTHIA8 are shown for the scenario corresponding to the nominal analysis, the scenario with an average number of collisions per bunch crossing equal to 1.8 but the condition to have at least one vertex in the event, and the scenario without pileup contributions and the condition to have exactly one vertex, respectively from top to bottom.

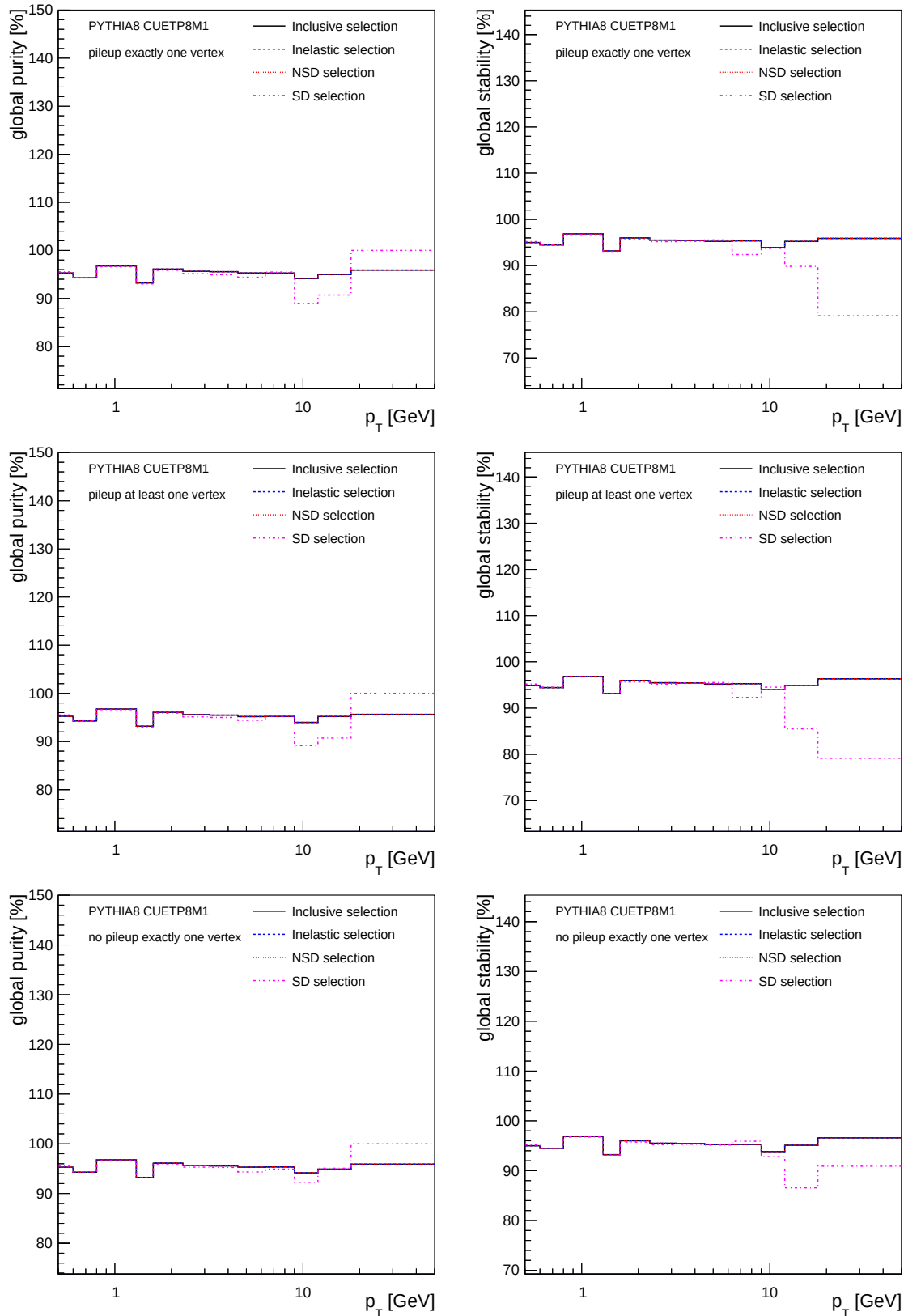


Figure 7.14: Purity (left) and stability (right) of the transverse-momentum distribution for the full event selection. The predictions of the Monte Carlo event generators PYTHIA8 are shown for the scenario corresponding to the nominal analysis, the scenario with an average number of collisions per bunch crossing equal to 1.8 but the condition to have at least one vertex in the event, and the scenario without pileup contributions and the condition to have exactly one vertex, respectively from top to bottom.

CHAPTER

8

EFFICIENCY CORRECTIONS AND UNFOLDING

Contents

8.1	D’Agostini unfolding method	106
8.2	Monte Carlo unfolding	107
8.3	Toy Monte Carlo unfolding	112
8.4	Validation of the procedure	113
8.5	Unfolding conclusions	121

The data are corrected to the stable particle level as defined in Section 5.4, for tracking and event selection efficiencies, and for detector smearing effects through the use of the Monte Carlo event generators PYTHIA8 and EPOS.

A response matrix, R , is constructed to describe the detector response by using the information provided by the Monte Carlo event generators and the simulation of the detector reconstruction with the software GEANT4.

The elements of the response matrix (R_{ji}) represent the conditional probability that for a given observable, a simulated value i is measured with a value j .

Given the true spectrum T , the measured spectrum M can be calculated as:

$$M = RT. \quad (8.1)$$

From the experimental perspective, the goal is to infer T from M :

$$T = R^{-1}M. \quad (8.2)$$

In practice, the matrix R cannot always be inverted. In addition, the limited statistics of the Monte Carlo samples used to obtain the response matrix lead to large fluctuations in the resulted distribution T .

In the following, the D’Agostini unfolding method used to overcome the mentioned problem is described.

8.1 D’Agostini unfolding method

The D’Agostini unfolding method is based on the Bayes’ theorem, where the definite relation between the probability of an event A conditional on another event B (that is generally different from the probability of B conditional on A) is described as:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}, \quad (8.3)$$

where $P(A)$ and $P(B)$ are the independent probabilities for the event A and B , respectively. $P(B|A)$ is the probability of the event B under the condition that A is true. Given these three quantities, $P(A|B)$ can be inferred. If A are the (C) auses and B the (E) ffects, or in other words, A are the true and B the measured quantities, the Bayes’ theorem can be rewritten in a discretised form as:

$$P(C_i|E_j) = \frac{P(E_j|C_i)P_0(C_i)}{\sum_{l=1}^{n_C} P(E_j|C_l)P_0(C_l)}, \quad (8.4)$$

where the index i runs over all (C) bins and j over all (E) bins. $P(E_j|C_i)$ is the response matrix R_{ji} , $P_0(C_i)$ is the prior or assumed true probability spectrum and $\sum_{l=1}^{n_C} P(E_j|C_l)P_0(C_l)$ is the folded prior or smeared prior spectrum, where the sum runs over all the cause bins. In the left side of Equation 8.4, the conditional probability that a measured value E_j comes from a given true value C_i is obtained. Figure 8.1 shows those relations between the causes and the effects, where each cause can be measured in different effect bins, or sometimes not even measured at all, due to inefficiencies (represented as T - “trash”).

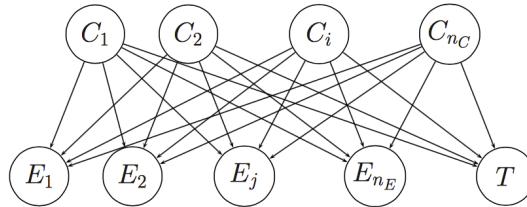


Figure 8.1: Probabilistic links from causes to effects. The node indicated by T (“trash”) stands for the inefficiency bin.

Once $P(C_i|E_j)$ is calculated, the corrected spectrum can be obtained as:

$$x(C_i) = \frac{1}{\epsilon_i} P(C_i|E_j) \cdot x(E_j), \quad (8.5)$$

where $x(C_i)$ ($x(E_j)$) is the number of entries in the i -th (j -th) bin of the unfolded (measured) distribution, and $\epsilon_i = \sum_{l=1}^{n_E} P(E_l|C_i)$ is the efficiency of a true value to be detected in any measured bin, i.e. inside the fiducial phase space.

A problem with this method as described above is that it assumes that the true spectrum is known a priori, which seems inconsistent. This problem can be solved by iterating over the previous steps for a finite number of times using as a prior at each iteration the unfolded distribution of the previous iteration until a convergence is reached. A convergence criterion to stop the iterative procedure is later explained in Section 8.4.

As discussed in [82], this procedure converges to the true spectrum because the unfolded distribution at each iteration always falls between the prior (of that iteration) and the true spectrum.

8.2 Monte Carlo unfolding

In this section, the unfolding procedure based in Monte Carlo samples with the full simulation of the detector effects is presented. The correction procedure is separated in two parts. The first part consists of the correction of smearing effects, taking into account at the same time the miss and fake, i.e., the objects that are simulated but not reconstructed, and the objects that are reconstructed but cannot be identified with any object at generator level. The second step consists of a correction factor for the detector efficiency to select events. This includes the trigger efficiency, the vertex selection efficiency and the enhanced event selection efficiency based on HF. The event-selection correction factor is an event-based scaling and is independent of the observables under study. It corrects for the overall efficiency to select an event at reconstructed level given that it was a selected event at generated level.

Figures 8.2, 8.3 and 8.4, show the obtained response matrices normalised by rows (particle level), for η , p_T and multiplicity, respectively. The response matrices are presented for the Inelastic-enhanced, NSD-enhanced and the SD-enhanced event samples in the top, middle and bottom rows, respectively. The different response matrices obtained from the event generator PYTHIA8 CUETP8M1 are shown on the left and EPOS LHC on the right.

Table 8.1 shows the correction factor for the event reconstruction efficiency. The correction factor takes into account the efficiency of the trigger, the selection of a good vertex, and the enhanced event selection based on HF (as defined in Chapter 5).

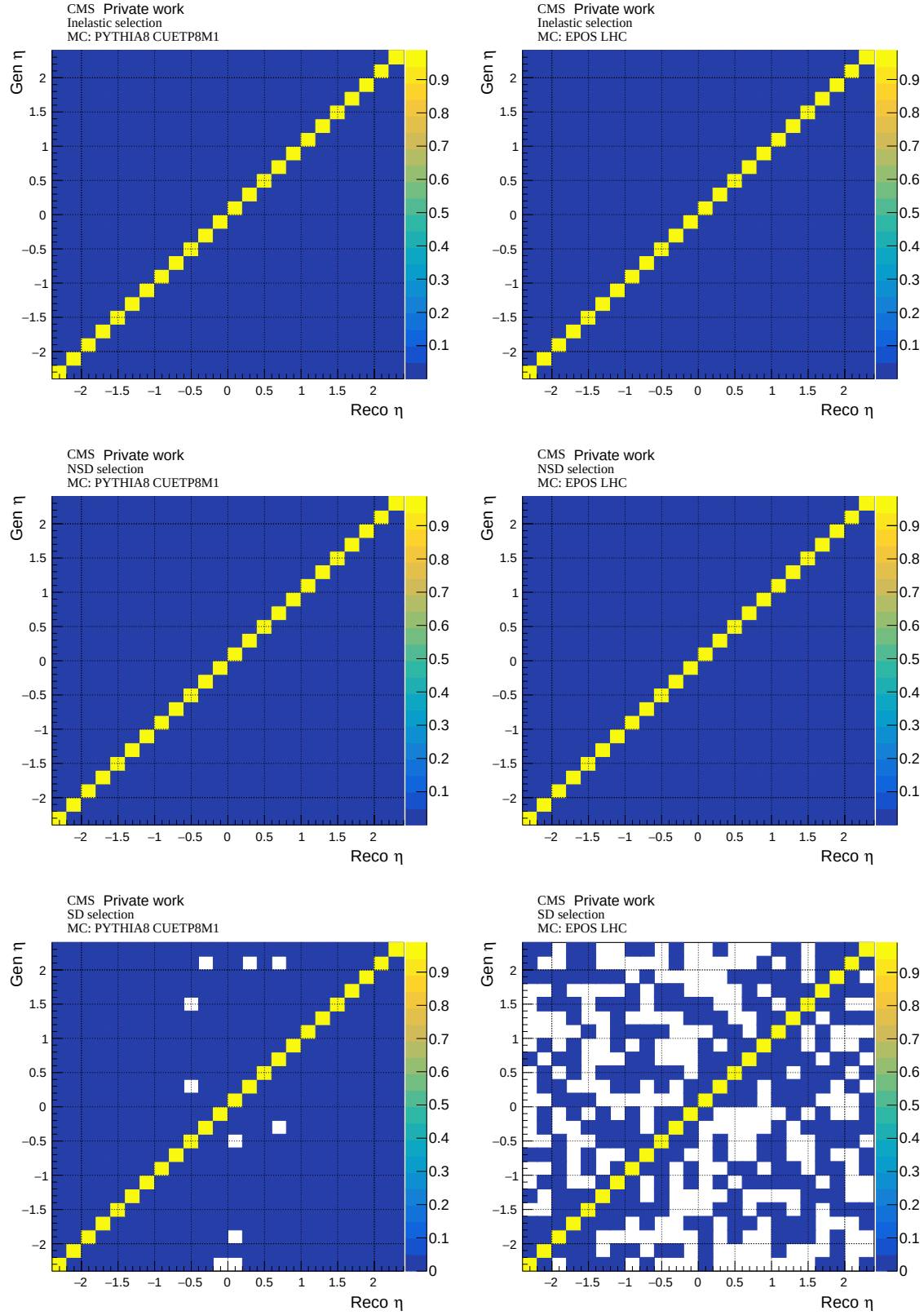


Figure 8.2: Response matrices for the charged-particle pseudorapidity, corresponding to the Inelastic-enhanced (top), the NSD-enhanced (middle), and the SD-enhanced event sample (bottom). The predictions of PYTHIA8 CUETP8M1 and EPOS LHC are shown on the left and right respectively.

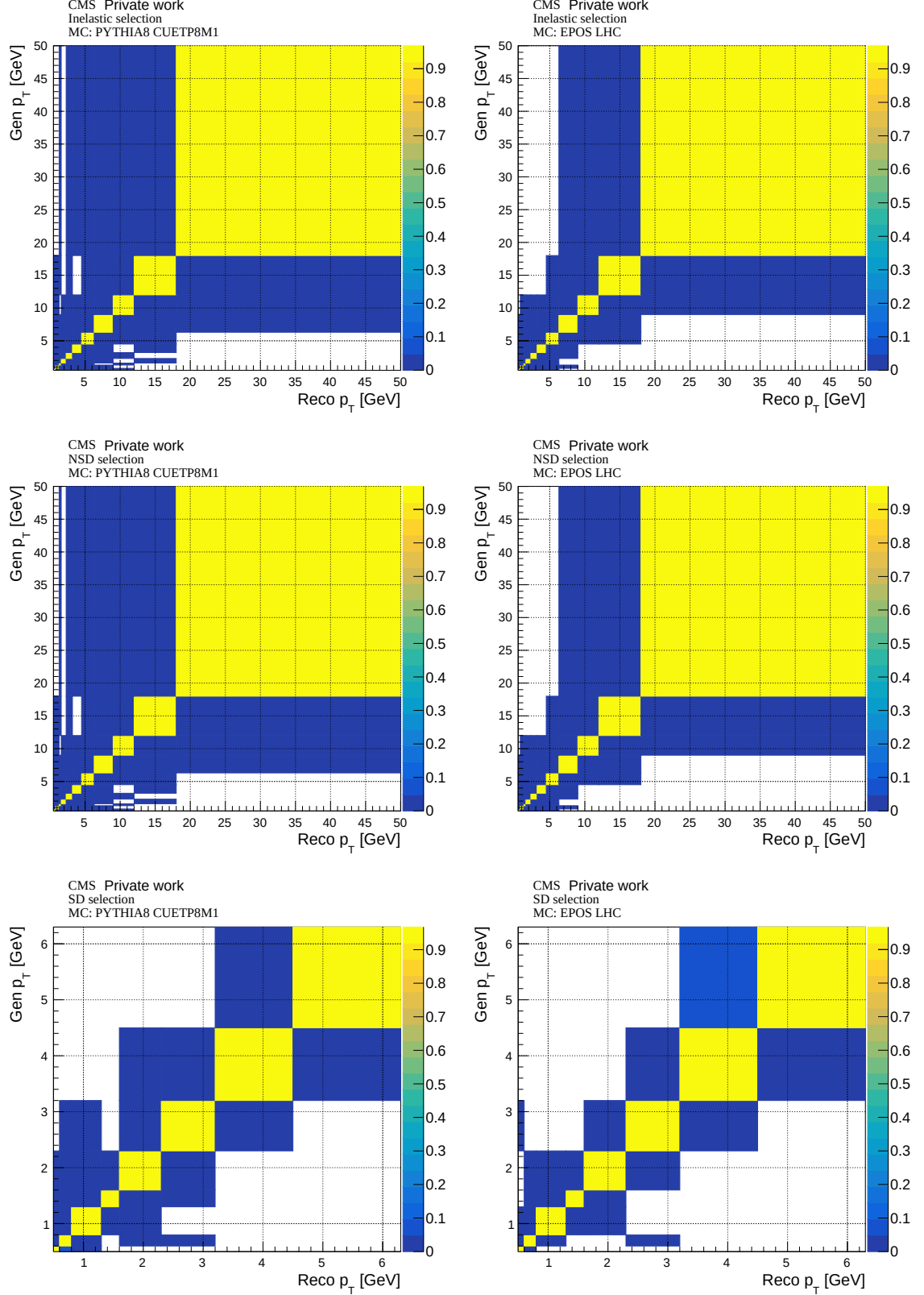


Figure 8.3: Response matrices for the charged-particle transverse momentum, corresponding to the Inelastic-enhanced (top), the NSD-enhanced (middle), and the SD-enhanced event sample (bottom). The predictions of PYTHIA8 CUETP8M1 and EPOS LHC are shown on the left and right respectively.

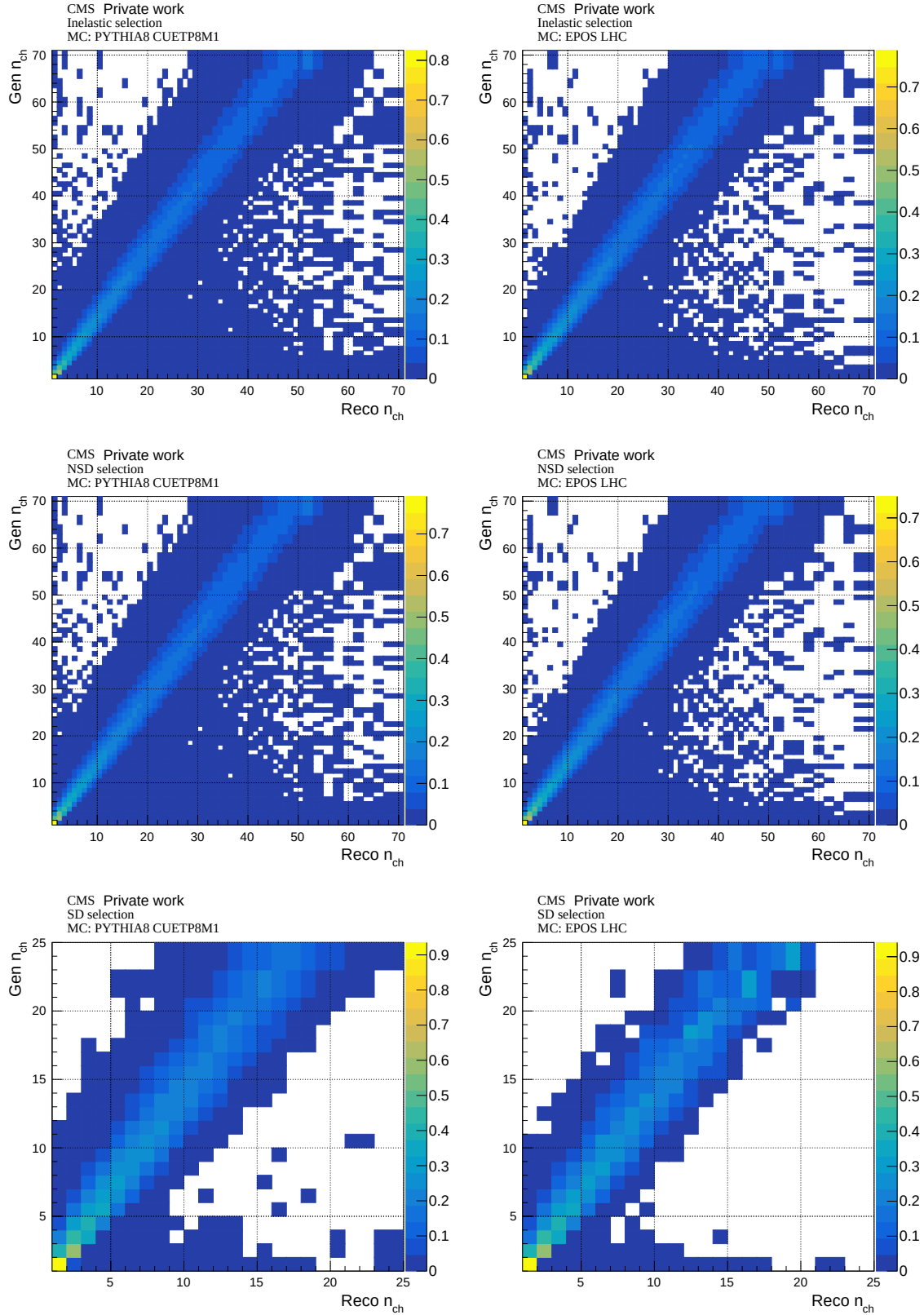


Figure 8.4: Response matrices for the charged-particle multiplicity per event, corresponding to the Inelastic-enhanced (top), the NSD-enhanced (middle), and the SD-enhanced event sample (bottom). The predictions of PYTHIA8 CUETP8M1 and EPOS LHC are shown on the left and right respectively.

	Correction factor for event selection efficiency		
Selection	EPOS LHC	PYTHIA8 CUETM1	PYTHIA8 MBR
Inelastic	0.62	0.61	0.62
NSD	0.60	0.60	0.58
SD	0.92	0.77	1.1
SD-One-Side	0.92	0.77	1.1

Table 8.1: Correction factor for event selection efficiency. The event selection includes the trigger, vertex and the HF enhanced event selection. The values are presented for three Monte Carlo event generators: EPOS LHC, PYTHIA8 CUETP8M1, and PYTHIA8 MBR.

8.3 Toy Monte Carlo unfolding

As the simulation of reconstructed events takes a considerable amount of computational time, it is often the case that only a limited amount of statistics is available. This leads to statistical fluctuations and not populated enough response matrices.

In the present analysis a method called “Toy Monte Carlo” is used to solve this problem. In this method, a reasonable amount of simulated Monte Carlo events at stable particle (*MC Gen*) and at reconstructed levels (*MC Reco*) is still needed, basically for two purposes:

- To calculate the resolution of the detector as a function of the observable of interest. (*Resolution Curves* Sec 7.2).
- To generate random numbers according to the original Monte Carlo distributions.

The stable particle level spectrum with larger statistics is (re)generated using the *Gen MC*, and for each of the (re)generated “particles” a smearing using the previously calculated *Resolution Curves* is applied. The (re)generated distribution will be referred to as *Toy Gen* and the smeared distribution to as *Toy Gen Smeared*. The response matrix is constructed using these two distributions.

In the case where the measurement has a reconstruction and event selection efficiencies of $\sim 100\%$, the unfolding with this response matrix is the only step needed to correct the measurement. In this analysis, the efficiency is not of the order of $\sim 100\%$ and an intermediate step is needed. In this step the *Reco MC* is used to generate the *Toy Reco* distribution. And with the *Toy Gen Smeared* and *Toy Reco* distributions a correction factor is calculated:

$$\epsilon_{Eff}(x_i^{reco}) = \frac{Toy\ Gen\ Smeared(x_i^{reco})}{Toy\ Reco(x_i^{reco})}, \quad (8.6)$$

where x_i is the i -th bin of the variable of interest at reconstructed level. This step corrects the measured distribution for reconstruction and event selection inefficiencies but leaves the smearing effects to the unfolding step. Figure 8.5 shows a diagram with the full procedure described here.

Figures 8.6, 8.7 and 8.8, show the response matrices normalised by rows (particle level) obtained with the “Toy Monte Carlo” method. The results of two Monte Carlo event generators are presented, on the left for PYTHIA8 CUETP8M1 and on the right for EPOS LHC.

The detector efficiency correction factor, ϵ_{Eff} , corresponding the Inclusive-enhanced, NSD-enhanced, and SD-enhanced sample are presented in figures 8.9, 8.10 and 8.11 for the pseudorapidity, transverse momentum and multiplicity of charged particles, respectively.

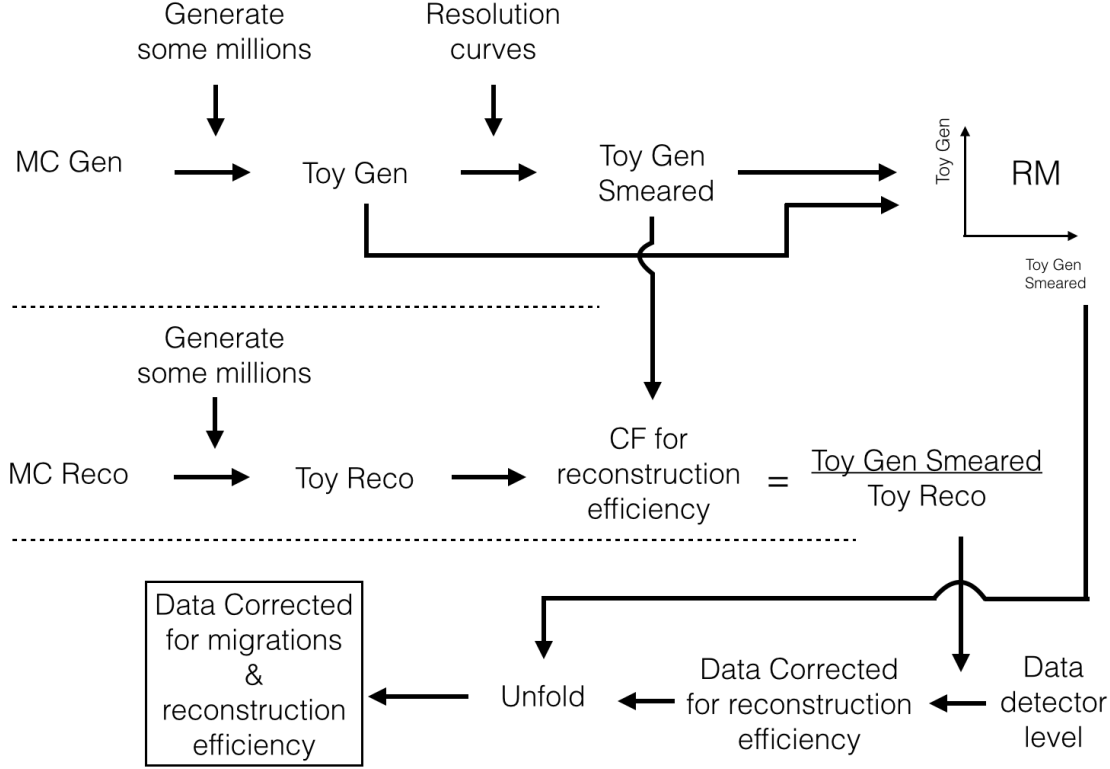


Figure 8.5: Diagram of the procedure to correct the measured distributions taking into account the efficiency effects as well the migrations between bins

8.4 Validation of the procedure

In this section the different tests used to validate the method are described. An important difference between the “*Toy MC*” method and the “*Full MC*” method, is the stage of the procedure at which the unfolding is applied, and also what is the unfolding correcting for. In the case of the “*Toy MC*” method, the unfolding is applied after correcting for the detector reconstruction efficiency. This is due to the fact that the “Toy response matrix” does not include information on the reconstruction efficiency but only information on smearing effects. While in the “*Full MC*” method, the unfolding is first applied and after the data is corrected for the event selection efficiency. The reason for this is that in contrast to the “Toy response matrix”, the “Full response matrix” does include all the information on track reconstruction efficiency and smearing effects. These features on the order of the corrections by the different methods employed, become evident when comparing figures 8.15 and 8.16.

- **Closure test** The measured distribution is replaced by a Monte Carlo simulation that passes through the full detector reconstruction as the real data. In this way, the reconstructed Monte Carlo that has been corrected, can be compared with the “truth”, i.e. the stable particle level of the simulation. There are two variants of this test, one where the same sample to make the correction is the one used as simulated data, and the other, in which two independent Monte Carlo samples are used. In the

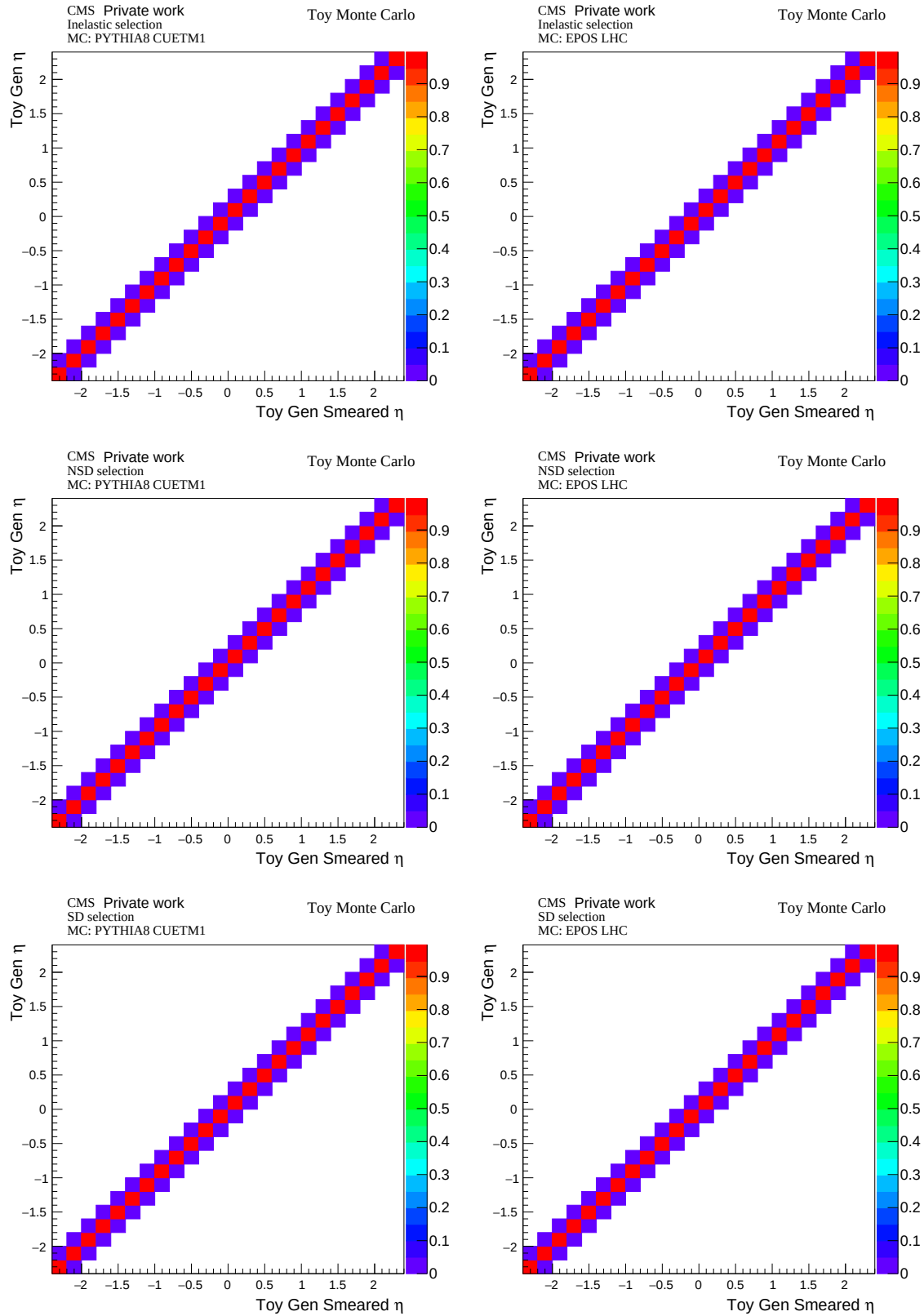


Figure 8.6: Reproduction of response matrices using the “Toy Monte Carlo” method for the pseudorapidity of charged particles. At the top row the Inelastic-enhanced sample, in the middle for the NSD-enhanced sample, and at the bottom for the SD-enhanced sample. At the left the PYTHIA8 CUETP8M1 Monte Carlo event generator and at the right EPOS LHC.

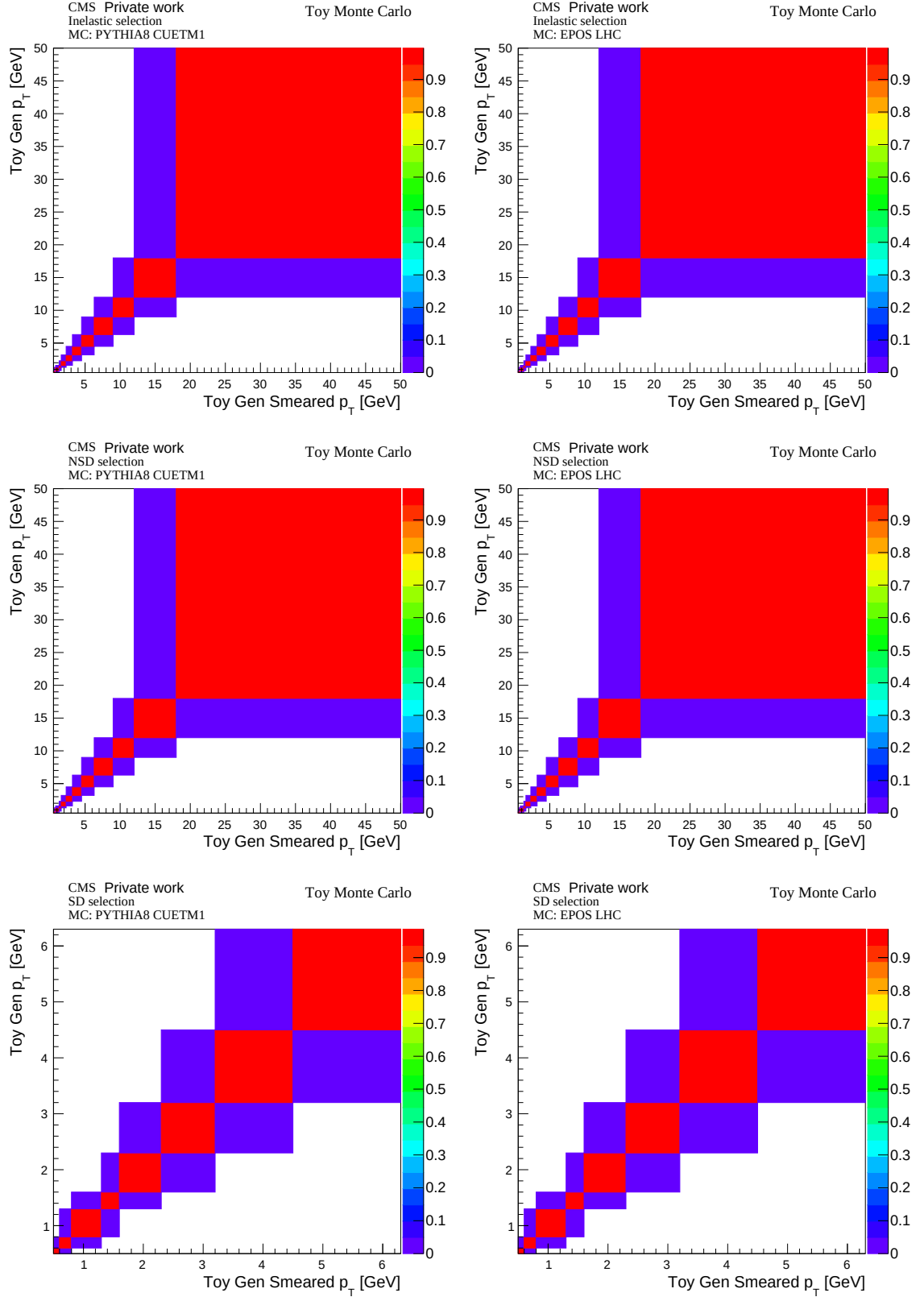


Figure 8.7: Reproduction of response matrices using the “Toy Monte Carlo” method for the transverse momentum of charged particles. At the top row the Inelastic-enhanced sample, in the middle for the NSD-enhanced sample, and at the bottom for the SD-enhanced sample. At the left the PYTHIA8 CUETP8M1 Monte Carlo event generator and at the right EPOS LHC.

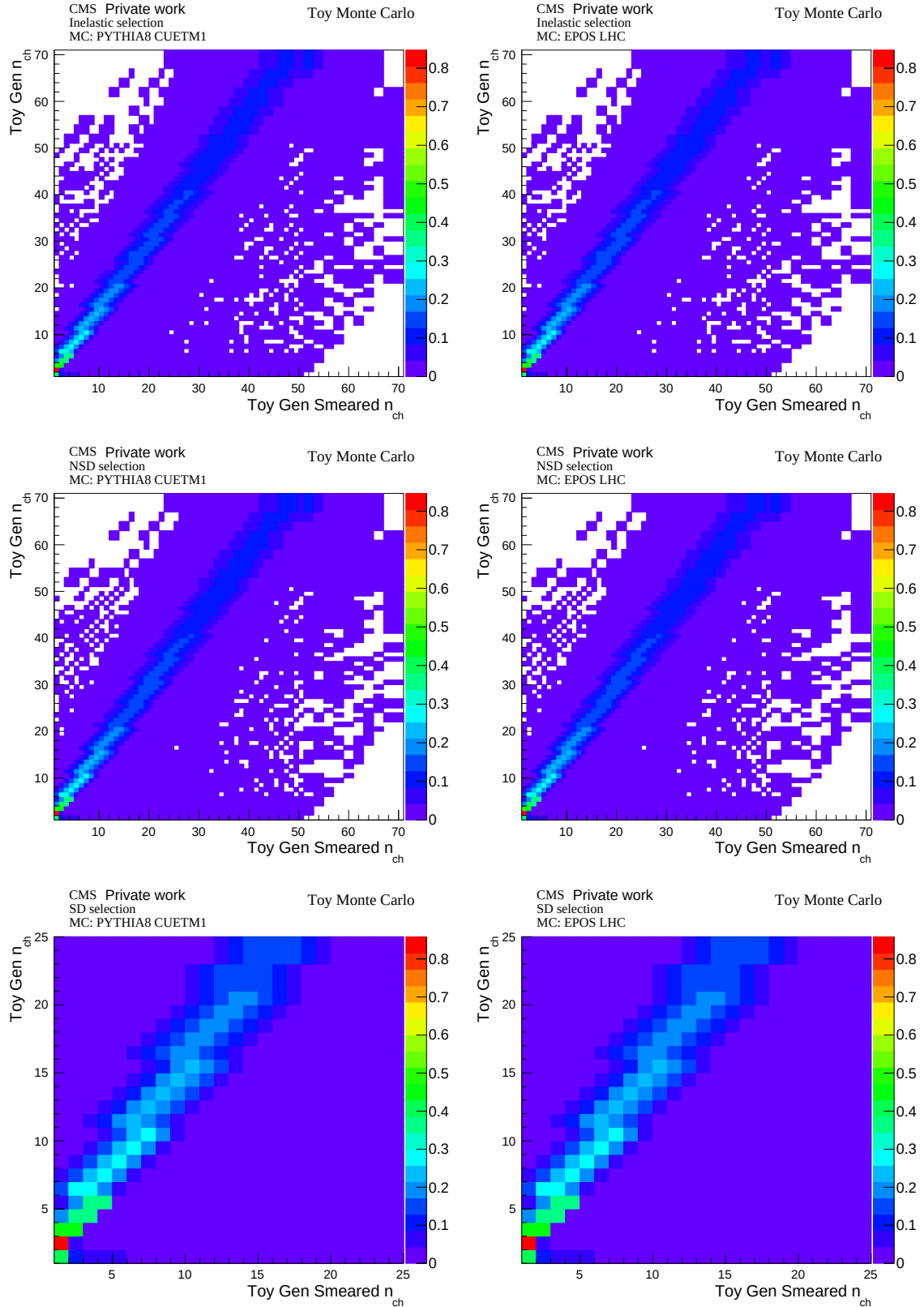


Figure 8.8: Reproduction of response matrices using the "Toy Monte Carlo" method for the charged particle multiplicity per event. At the top row the Inelastic-enhanced sample, in the middle for the NSD-enhanced sample, and at the bottom for the SD-enhanced sample. At the left the PYTHIA8 CUETP8M1 Monte Carlo event generator and at the right EPOS LHC.

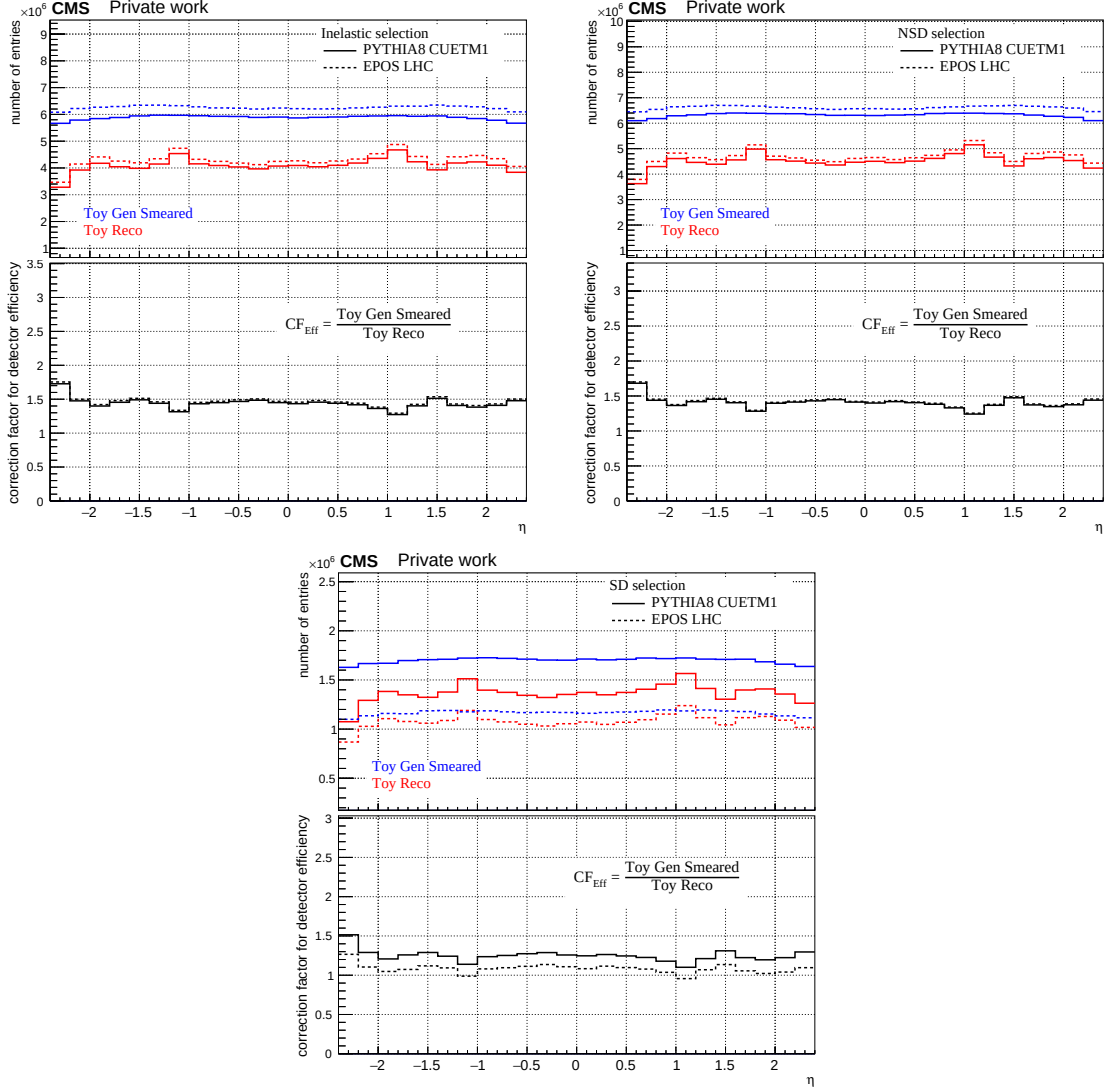


Figure 8.9: At the top of each plot the “Toy Gen Smeared” and “Toy Reco” predictions for the pseudorapidity of charged particles from the Monte Carlo event generators PYTHIA8 CUETP8M1 and EPOS LHC are shown. At the bottom of each plot the detector efficiency correction factors from the “Toy Monte Carlo” method for the pseudorapidity of charged particles. The top left plot corresponds to the Inelastic-enhanced sample, the top-right to the NSD-enhanced sample and at the bottom for the SD-enhanced sample.

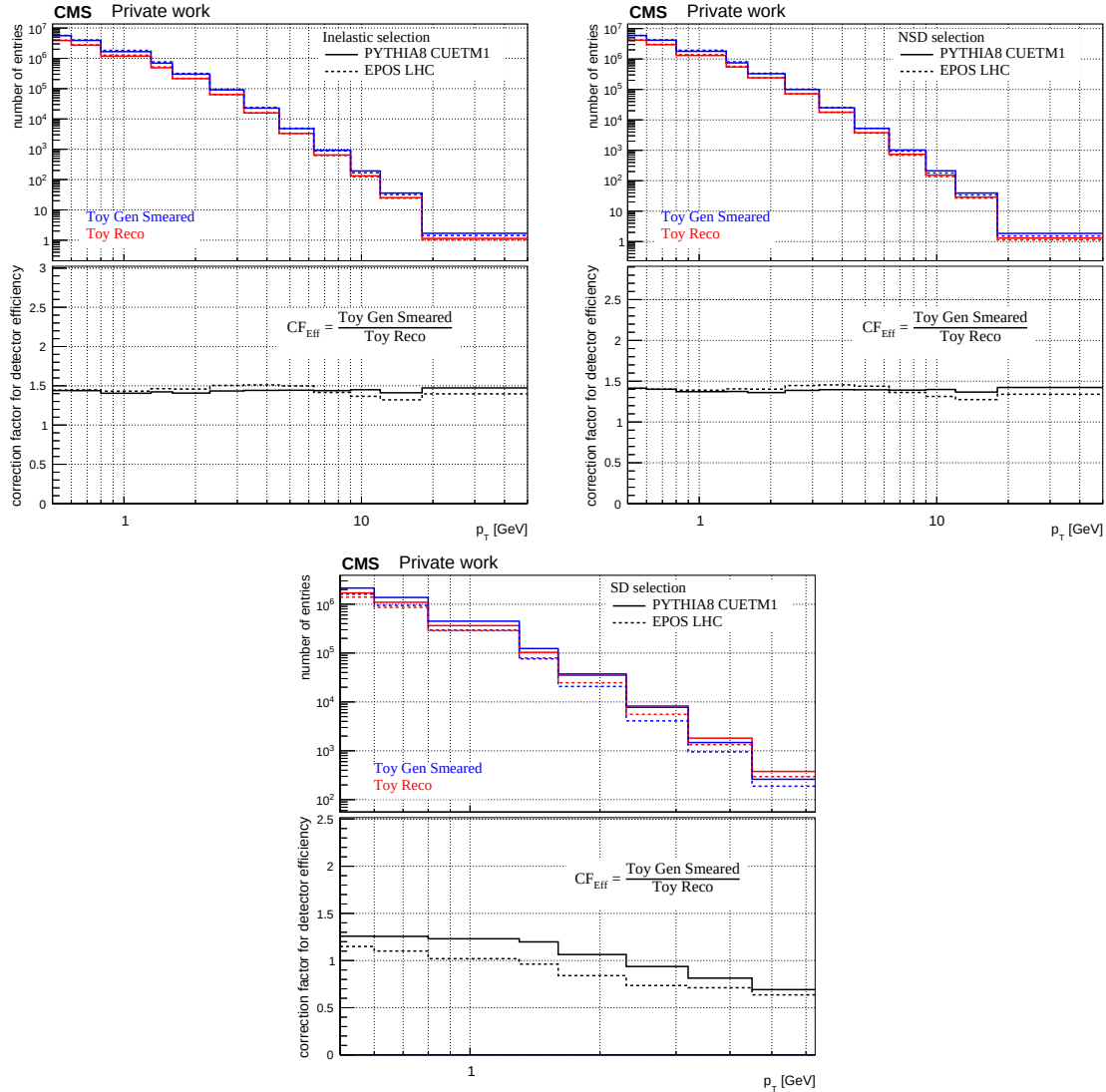


Figure 8.10: On top the “Top Gen Smeared” and “Toy Reco” predictions for the transverse momentum of charged particles from the Monte Carlo event generators PYTHIA8 CUETP8M1 and EPOS LHC are shown. At the bottom the detector efficiency correction factors from the “Toy Monte Carlo” method for the transverse momentum of charged particles. The top left plot corresponds to the Inelastic-enhanced sample, the top-right to the NSD-enhanced sample and at the bottom for the SD-enhanced sample.

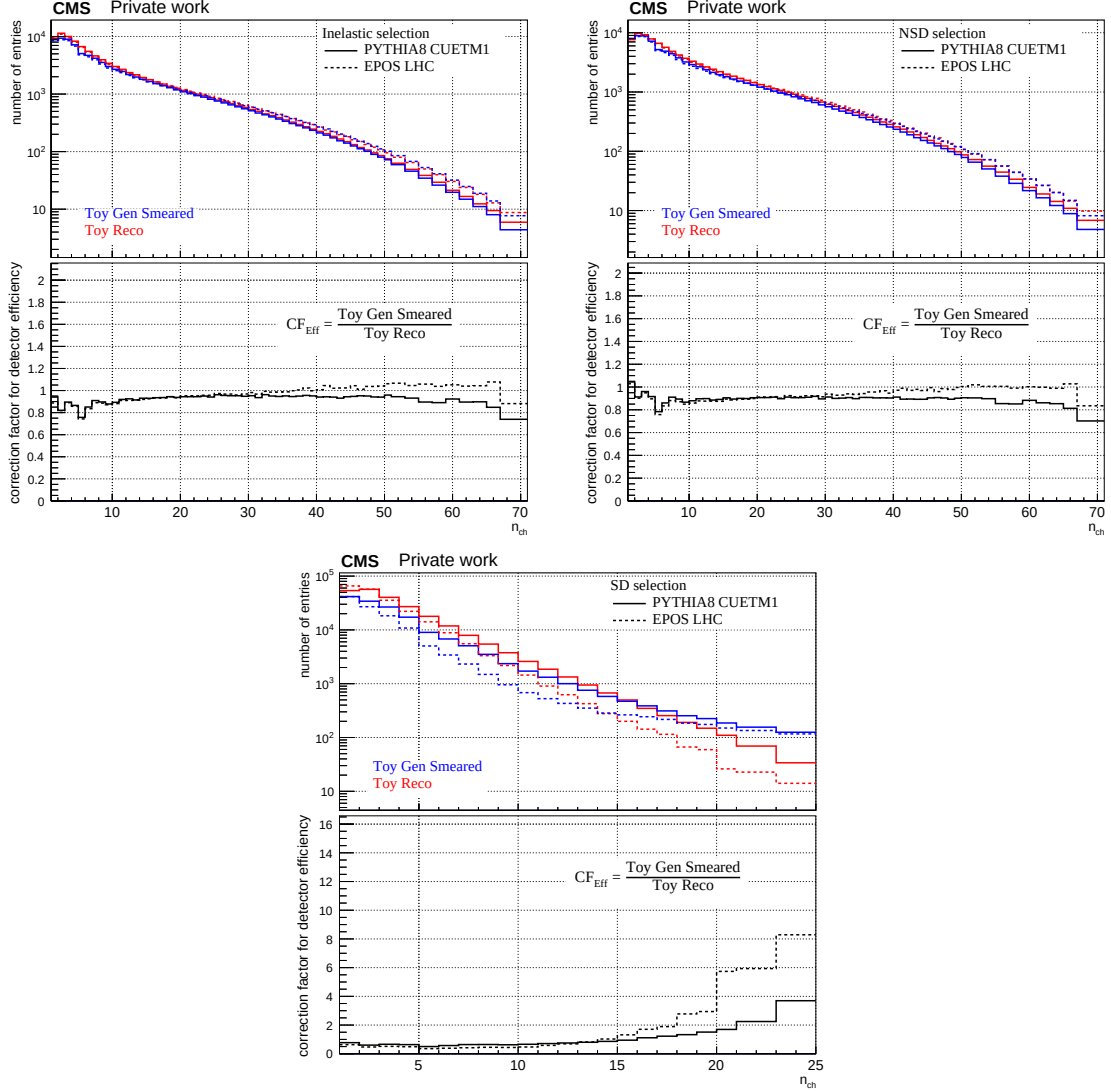


Figure 8.11: On top the “Toy Gen Smeared” and “Toy Reco” predictions for the charged particle multiplicity per event from the Monte Carlo event generators PYTHIA8 CUETP8M1 and EPOS LHC are shown. At the bottom the detector efficiency correction factors from the “Toy Monte Carlo” method for the charged particle multiplicity per event. The top left plot corresponds to the Inelastic-enhanced sample, the top-right to the NSD-enhanced sample and at the bottom for the SD-enhanced sample.

former one, it is expected that the corrected distribution and the stable particle level of the Monte Carlo are the same, after a number of iterations. In the latter case, small differences can be expected, but not very significant. Figures 8.12 and 8.13 show the closure test for the “*Toy MC*” method for the Non-Single Diffractive enhanced sample, using the same sample and independent samples as pseudo data and for the corrections, respectively. Figure 8.14 shows the closure test for the “*Full MC*” method for the Non-Single Diffractive enhanced sample.

- **Backfolding** This test extends the closure test by checking the stability of the method. Once a measurement is corrected, it is possible to use the response matrix and the efficiency correction factor ϵ_{Eff} to “*foldback*” the corrected distribution to the reconstructed level. Even though the unfolded distribution is expected to fall between the prior and the true spectrum, the fluctuations in the measured spectrum due to limited statistics can get amplified as the number of iterations increases.

Performing a χ^2 test between the uncorrected and foldedback distributions as a function of the number of iterations, provides a criterion to determine where to stop the iterative procedure after reaching convergence and before starting to get large fluctuations. In general, the unfolding converges after 4-5 iterations.

Figure 8.15 shows the pseudorapidity, transverse momentum and multiplicity distributions of charged particles for the case of Non-single Diffractive enhanced sample in each of the different steps of the unfolding-backfolding procedure of the “*Toy MC*” method, namely the uncorrected data, the data with the efficiency correction factor being applied, the unfolded data with the efficiency correction factor being applied, the backfolded data, the Monte Carlo reconstructed level, the Monte Carlo generated level, and the folding of the Monte Carlo generated level which is in agreement with the Monte Carlo at reconstructed level.

Similarly, figure 8.16 shows the pseudorapidity, transverse momentum and multiplicity distributions of charged particles for the case of Non-single Diffractive enhanced sample in each of the different steps of the unfolding-backfolding procedure of the “*Full MC*”, namely the uncorrected data, the unfolded data without the efficiency correction factor being applied, the unfolded data with the efficiency correction factor being applied, the backfolded data, the Monte Carlo reconstructed level, the Monte Carlo generated level, and the folding of the Monte Carlo generated level which is in agreement with the Monte Carlo at reconstructed level.

Figures 8.17, 8.18 and 8.19 show the χ^2 over the number of degrees of freedom (NDF) as a function of the number of iterations (left), and the ratio (*Folded/Uncorrected Data*) as a function of the observable of interest (right), all of them for the Non-Single Diffractive enhanced sample.

- **Bottom-line test** The Bottom-line test is a sanity check of the procedure in which the level of agreement between the data and some model, before (*reco*) and after the unfolding procedure (*truth*) is compared. In the case in which the unfolding procedure is able to correct the measurement without making use of any regularisation, the level of agreement before and after the unfolding should be exactly the same. In practice, the unfolding procedures make use of different tricks to regularise the minimisation

in order to avoid high fluctuations in the results. Due to the regularisation in the unfolding procedure, the level of agreement between the model and the data at (*reco*) and (*truth*) levels should remain close but not necessarily equal. The level of agreement is quantified by the χ^2 test:

$$\chi^2 = (y - Kx)^T V_y^{-1} (y - Kx), \quad (8.7)$$

where in the case of unfolded level, y is the corrected data, K the response matrix, x the prediction, and V_y the covariance matrix of the corrected data. Values of $\chi_{truth}^2 < \chi_{reco}^2$ can be expected, but values $\chi_{truth}^2 \gg \chi_{reco}^2$ give indications of too strong regularisation.

As mentioned before, the two methods to correct the data implemented in this analysis apply the unfolding at different stages of the correction procedure. Hence, what is considered the (*reco*) and (*truth*) levels are not the same when comparing the two methods. What has to be evaluated is the level of agreement within each method.

In tables 8.2, 8.3, and 8.4, the bottom-line test for the “*Full MC*” and “*Toy MC*” methods is summarised for the pseudorapidity, transverse momentum and multiplicity of charged particles, respectively. Each time the results are shown for the Inelastic-enhanced, NSD-enhanced and SD-enhanced samples by columns, and for three different Monte Carlo event generators EPOS LHC, PYTHIA8 CUETP8M1 and PYTHIA8 MBR on the top, middle and bottom of the table, respectively.

8.5 Unfolding conclusions

Following the results of the Bottom-line test performed for both, the “*Full MC*” method and the “*Toy MC*” method, it is concluded that:

- The “*Toy MC*” method is a reliable tool when the detector smearing effects are moderate.
- In the case of strong migrations, the “*Toy MC*” method becomes unstable. This is in part due to the bias introduced by the resolution binning and the fitting to the resolution curves.
- The “*Full MC*” method has been chosen for the unfolding of the multiplicity distributions, leaving the SD-enhanced selection out of reach due to lack of statistics in the MC samples.

The criteria to choose the final number of iterations is defined from the results obtained in the backfolding and bottom-line tests. The backfolding test requires that the convergence has been reached, while the bottom-line test validates that particular choice by requiring that the χ^2 at reco and truth levels are close to each other, where “close” is defined as to be of the same order of magnitude. Following these criteria, the number of iterations is chosen to be 2 for all the pseudorapidity distributions, as well as for all the transverse-momentum distributions, while for all the multiplicity distributions the optimal choice is to stop after 5 iterations. The χ^2 values of the bottom-line test for these particular number of iterations are within $\sim 5\%$ agreement.

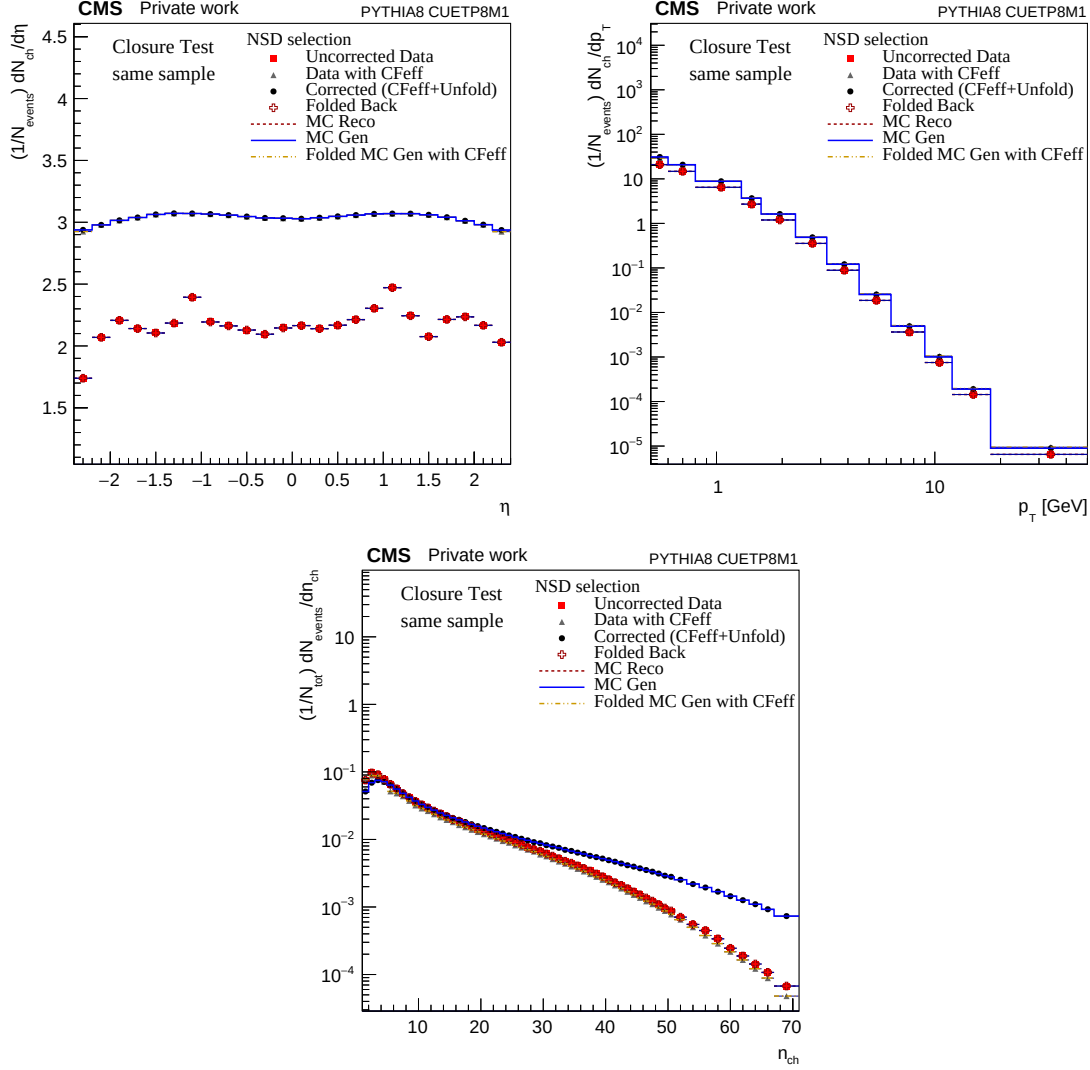


Figure 8.12: Closure test with the same sample for the “*Toy MC*” method for the pseudorapidity, transverse momentum and multiplicity of charged particles in the case of the Non-Single Diffractive enhanced sample. The corrections correspond to the Monte Carlo event generator PYTHIA8 CUETP8M1.

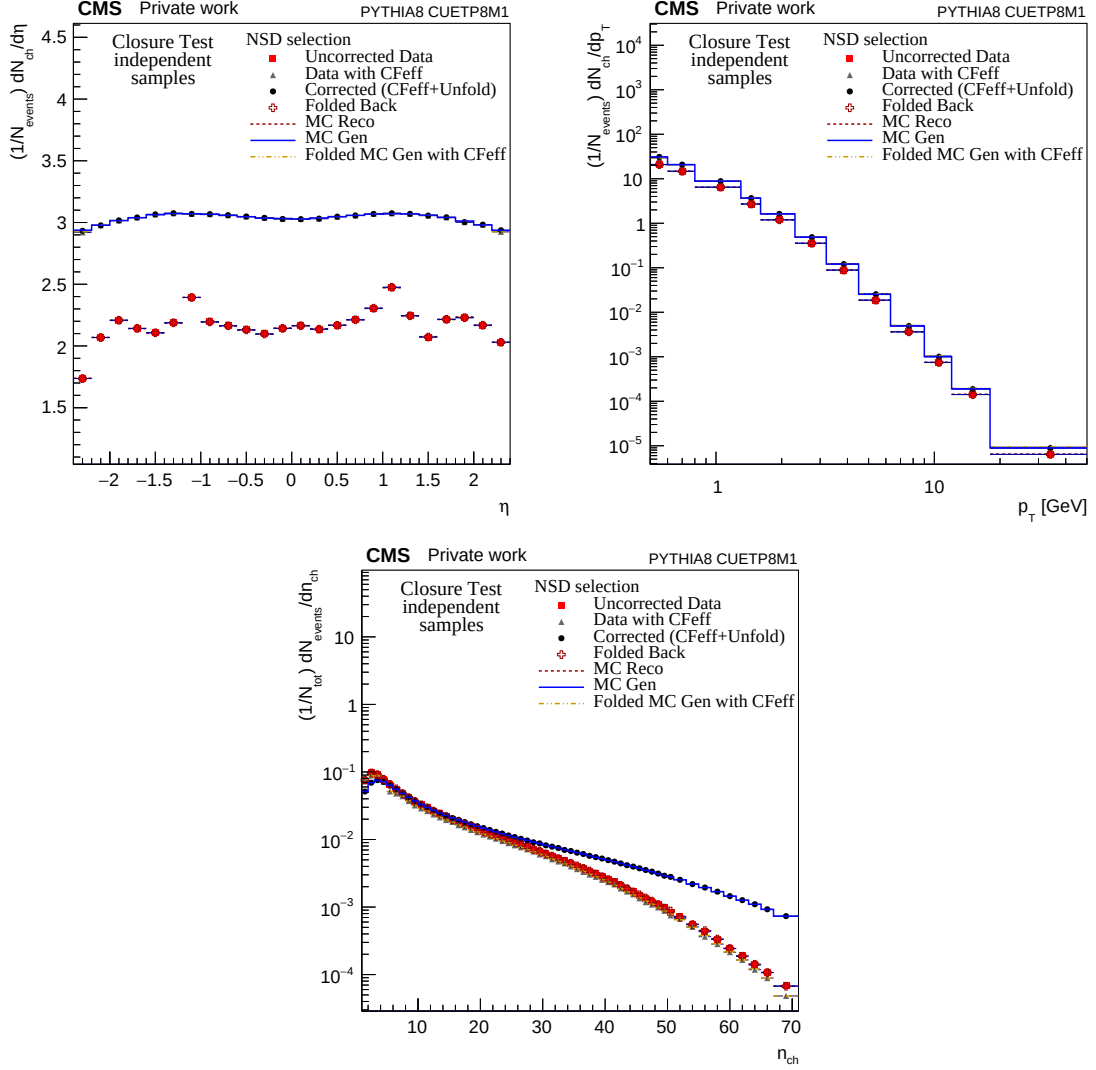


Figure 8.13: Closure test with independent samples for the “*Toy MC*” method for the pseudorapidity, transverse momentum and multiplicity of charged particles in the case of the Non-Single Diffractive enhanced sample. The corrections correspond to the Monte Carlo event generator PYTHIA8 CUETP8M1.

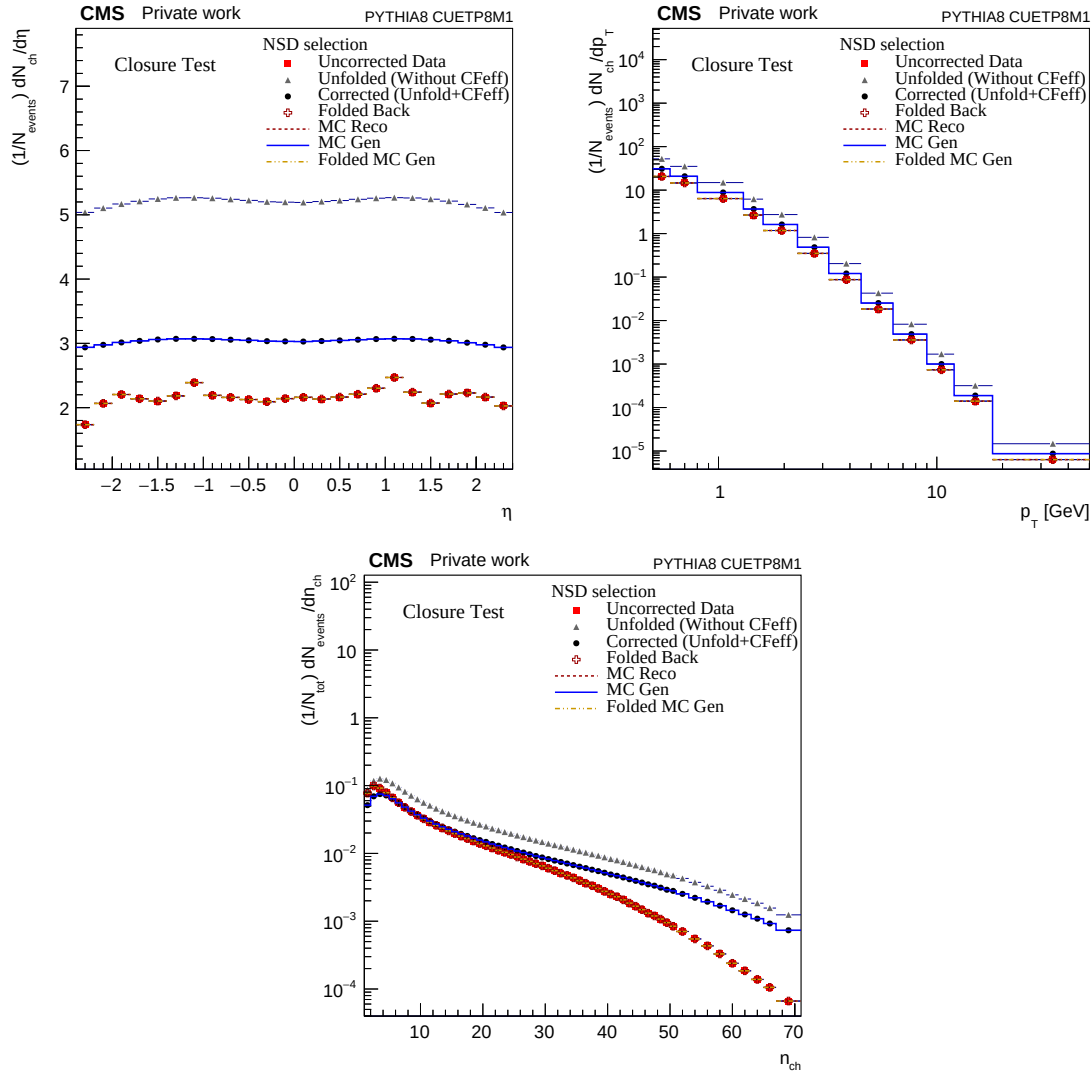


Figure 8.14: Closure test with same sample for the “Full MC” method for the pseudorapidity, transverse momentum and multiplicity of charged particles in the case of the Non-Single Diffractive enhanced sample. The corrections correspond to the Monte Carlo event generator PYTHIA8 CUETP8M1.

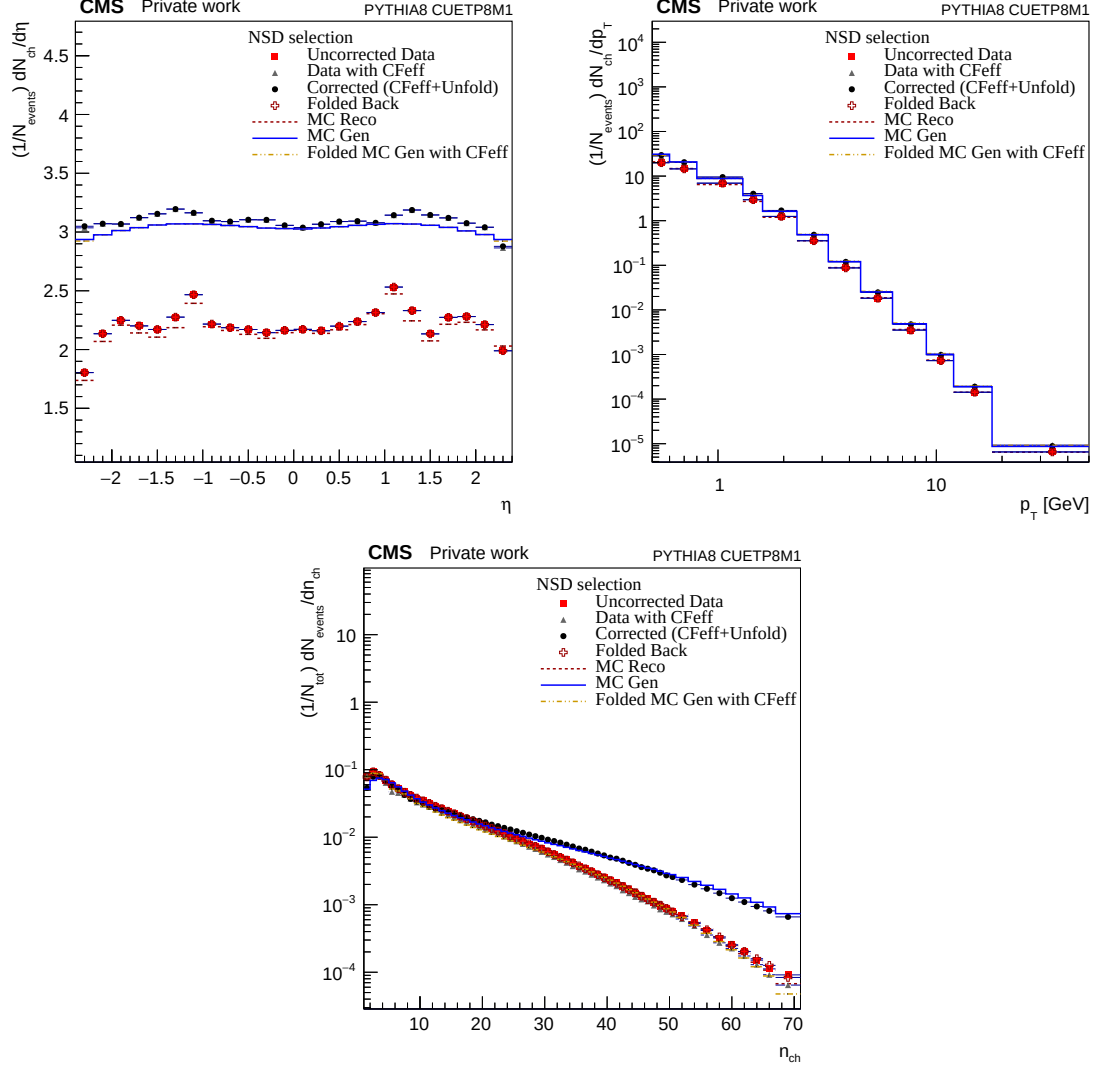


Figure 8.15: Summary plot with all the steps in the correction of the data and the folding tests for the “*Toy MC*” method. From top to bottom of the legend, the uncorrected data, the data with the efficiency correction factor being applied, the unfolded data with the efficiency correction factor being applied, the backfolded data, the Monte Carlo reconstructed level, the Monte Carlo generated level, and the folding of the Monte Carlo generated level, which is comparable to the Data with the efficiency corrections being applied.

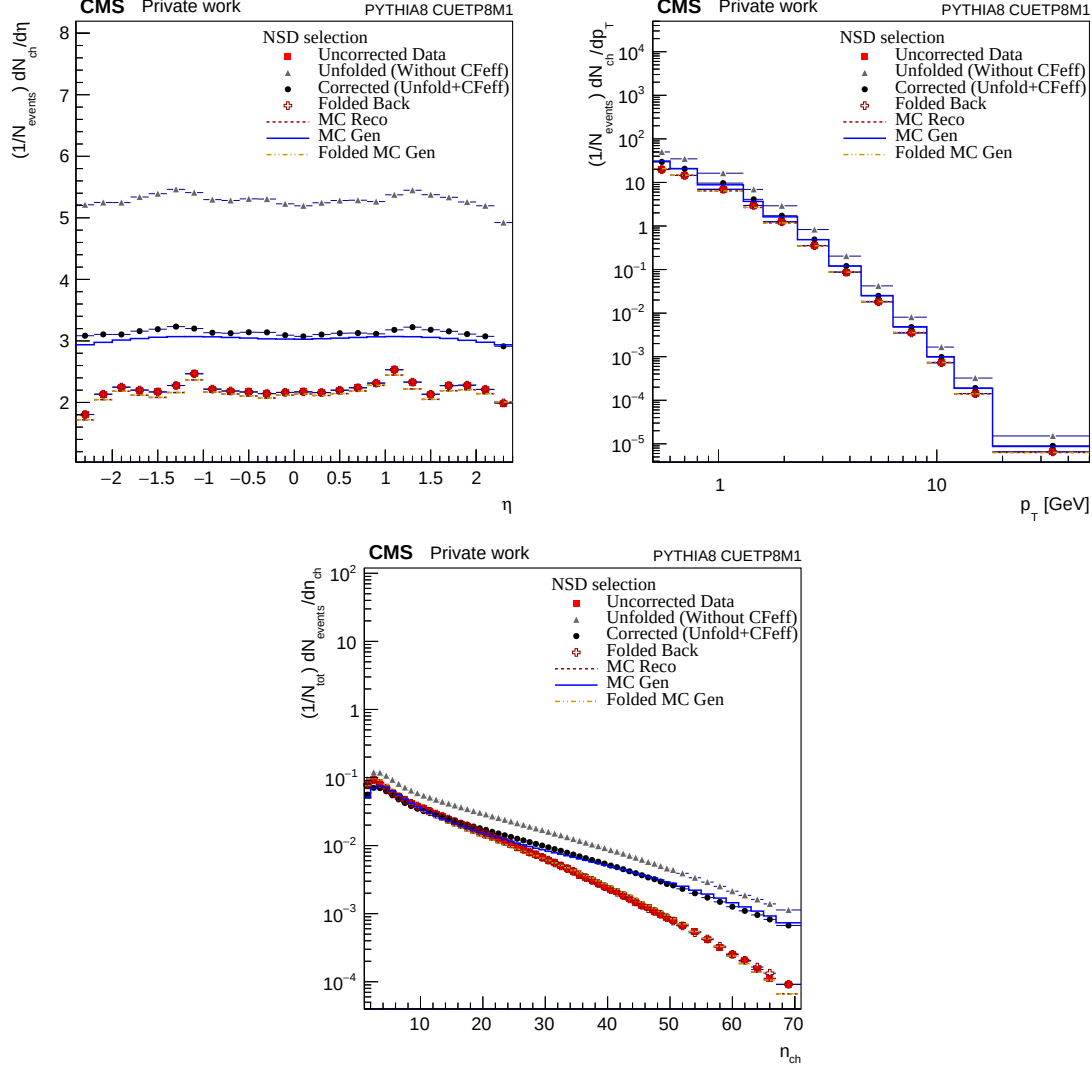


Figure 8.16: Summary plot with all the steps in the correction of the data and the folding tests for the “*Full MC*” method . From top to bottom of the legend, the uncorrected data, the unfolded data without the efficiency correction factor being applied, the unfolded data with the efficiency correction factor being applied, the backfolded data, the Monte Carlo reconstructed level, the Monte Carlo generated level, and the folding of the Monte Carlo generated level, which is in agreement with the Monte Carlo at reconstructed level.

χ^2 values for Bottom-line test										
		Inelastic selection			NSD selection			SD selection		
Observable	iter	ndf	reco	truth	ndf	reco	truth	ndf	reco	truth
η	EPOS LHC									
		“Full MC” method								
	1	23	4529.4	4529.4	23	4332.5	4332.5	23	18597.4	18597.4
	2			4538.2			4340.8			18550.7
	5			4532.4			4334.9			18422.4
	10			4532.0			4334.5			18357.4
	20			4532.0			4334.5			18347.0
	50			4532.0			4334.5			18346.8
		“Toy MC” method								
	1	23	3756.2	3756.2	23	6820.1	6820.9	23	11492.7	11492.7
	2			3729.6			6747.9			11308.5
	5			3728.0			6743.5			11297.4
	10			3728.0			6743.5			11297.4
	20			3728.0			6743.5			11297.4
	50			3728.0			6743.5			11297.4
	PYTHIA8 CUETP8M1									
		“Full MC” method								
	1	23	35415.8	35415.8	23	23290.6	23290.6	23	4661.3	4661.3
	2			35411.0			23288.4			4688.5
	5			35406.5			23284.5			4725.3
	10			35406.3			23284.4			4730.9
	20			35406.3			23284.4			4731.1
	50			35406.3			23284.4			4731.1
		“Toy MC” method								
	1	23	23000.5	23000.4	23	8094.4	8095.2	23	7139.4	139.3
	2			22645.8			7985.4			7022.1
	5			22624.0			7978.6			7014.8
	10			22624.0			7978.6			7014.8
	20			22624.0			7978.6			7014.8
	50			22624.0			7978.6			7014.8
	PYTHIA8 MBR									
		“Full MC” method								
	1	23	128666.4	128666.4	23	221290.7	221290.7	23	1559.9	1559.9
	2			128714.6			221339.2			1533.2
	5			128723.6			221347.5			1461.7
	10			128723.7			221347.6			1424.3
	20			128723.7			221347.6			1414.7
	50			128723.7			221347.6			1413.8
		“Toy MC” method								
	1	23	98431.5	98430.9	23	226885.3	226886.9	23	1547.3	1547.2
	2			96850.1			223216.9			1525.1
	5			96752.8			222992.7			1523.7
	10			96752.8			222992.6			1523.7
	20			96752.8			222992.6			1523.7
	50			96752.8			222992.6			1523.7

Table 8.2: Bottom-line test for the unfolding of the pseudorapidity distribution of charged particles shown for three different Monte Carlo event generators EPOS LHC, PYTHIA8 CUETP8M1 and PYTHIA8 MBR. The three different enhanced samples are shown, the Inelastic-enhanced at the left column, the NSD-enhanced in the middle column and the SD-enhanced at the right column. Each time two unfolding methods are compared, the “*Full MC*” method on the top and the “*Toy MC*” method on the bottom. The Bottom-line test requires that the χ^2 value before the unfolding (*reco*) and after the unfolding (*truth*) are similar.

χ^2 values for Bottom-line test										
		Inelastic selection			NSD selection			SD selection		
Observable	iter	ndf	reco	truth	ndf	reco	truth	ndf	reco	truth
p_T	EPOS LHC									
		“Full MC” method								
	1	11	29548.0	29548.0	11	29332.6	29332.6	7	22720.3	22720.2
	2			29562.1			29344.0			21672.6
	5			29572.2			29351.0			19982.7
	10			29571.7			29350.5			19548.8
	20			29571.7			29350.5			19182.8
	50			29571.7			29350.5			17467.7
		“Toy MC” method								
	1	11	2610.0	2608.8	11	2069.7	2069.2	7	11933.7	11935.4
	2			2482.4			1977.8			11228.6
	5			2436.5			1942.5			10966.8
	10			2436.4			1942.4			10965.4
	20			2436.4			1942.4			10965.4
	50			2436.4			1942.4			10965.4
	PYTHIA8 CUETP8M1									
		“Full MC” method								
	1	11	87842.2	87842.2	11	71196.5	71196.5	7	7373.7	7373.7
	2			87764.6			71151.0			7406.5
	5			87750.5			71144.2			7606.7
	10			87750.1			71144.0			7659.4
	20			87750.1			71144.0			7630.0
	50			87750.1			71144.0			7401.6
		“Toy MC” method								
	1	11	8950.3	8950.4	11	5205.5	5204.7	7	4420.9	4425.5
	2			8463.1			4980.4			4192.4
	5			8304.7			4897.0			4113.8
	10			8304.1			4896.7			4113.5
	20			8304.1			4896.7			4113.5
	50			8304.1			4896.7			4113.5
	PYTHIA8 MBR									
		“Full MC” method								
	1	11	129585.3	129585.3	11	221462.8	221462.8	7	2778.8	2785.5
	2			129668.1			221540.2			3379.0
	5			129679.0			221551.0			6892.2
	10			129678.1			221550.5			58941.4
	20			129678.1			221550.5			-
	50			129678.1			221550.5			-
		“Toy MC” method								
	1	11	16041.9	16044.2	11	36454.3	36451.3	7	3537.3	3536.4
	2			14985.5			34011.2			3356.5
	5			14663.9			33291.7			3275.1
	10			14662.5			33288.6			3274.6
	20			14662.5			33288.6			3274.6
	50			14662.5			33288.6			3274.6

Table 8.3: Bottom-line test for the unfolding of the transverse momentum distribution of charged particles shown for three different Monte Carlo event generators EPOS LHC, PYTHIA8 CUETP8M1 and PYTHIA8 MBR. The three different enhanced samples are shown, the Inelastic-enhanced at the left column, the NSD-enhanced in the middle column and the SD-enhanced at the right column. Each time two unfolding methods are compared, the “Full MC” method on the top and the “Toy MC” method on the bottom. The Bottom-line test requires that the χ^2 value before the unfolding (*reco*) and after the unfolding (*truth*) are similar.

χ^2 values for Bottom-line test										
		Inelastic selection			NSD selection			SD selection		
Observable	iter	ndf	reco	truth	ndf	reco	truth	ndf	reco	truth
n_{ch}	EPOS LHC									
	1 2 5 10 20 50	“Full MC” method								
		58	19478.3	19230.1	58	20450.3	20222.0	21	4422.8	1339.0
				19050.5			20129.0			-
				18503.7			19647.4			-
				18179.7			19267.6			-
				17926.9			19001.1			-
				17627.9			18722.5			-
	“Toy MC” method									
	1 2 5 10 20 50	58	17704.4	-	58	17595.1	-	21	2781.4	5537.6
				16734.2			16692.4			3533.6
				14094.8			14819.3			4060.2
				9515.5			8661.6			6480.3
				7347.6			7492.6			4512.5
				12368.4			16857.1			-
		PYTHIA8 CUETP8M1								
	1 2 5 10 20 50	“Full MC” method								
		58	18648.7	18822.6	58	14636.5	14803.3	21	3158.1	3151.4
				19093.7			14988.4			4318.2
				19016.3			14919.5			13961.3
				18905.2			14825.7			18589.0
				18871.4			14779.1			18078.4
				18857.9			14749.1			19252.8
	“Toy MC” method									
	1 2 5 10 20 50	58	15581.1	-	58	10542.8	-	21	4213.0	-
				14653.5			9871.9			4781.7
				14116.8			9524.6			4568.8
				11475.9			7046.3			3990.9
				10102.6			6718.5			3897.4
				48871.4			13157.2			4092.4
		PYTHIA8 MBR								
	1 2 5 10 20 50	“Full MC” method								
		58	33311.2	33160.8	58	38085.2	37900.1	21	1411.3	1288.8
				33258.1			38023.8			-
				31923.4			36416.1			-
				31287.5			35689.6			-
				30926.3			35269.1			-
				30781.6			35019.3			-
	“Toy MC” method									
	1 2 5 10 20 50	58	31780.2	-	58	43637.9	-	21	2125.2	-
				28717.2			40501.8			2477.4
				21345.1			28773.7			2204.7
				13638.8			15996.1			2456.8
				7283.2			9336.1			1334.5
				4200.1			-			1450.2

Table 8.4: Bottom-line test for the unfolding of the multiplicity distribution of charged particles shown for three different Monte Carlo event generators EPOS LHC, PYTHIA8 CUETP8M1 and PYTHIA8 MBR. The three different enhanced samples are shown, the Inelastic-enhanced at the left column, the NSD-enhanced in the middle column and the SD-enhanced at the right column. Each time two unfolding methods are compared, the “Full MC” method on the top and the “Toy MC” method on the bottom. The Bottom-line test requires that the χ^2 value before the unfolding (*reco*) and after the unfolding (*truth*) are similar.

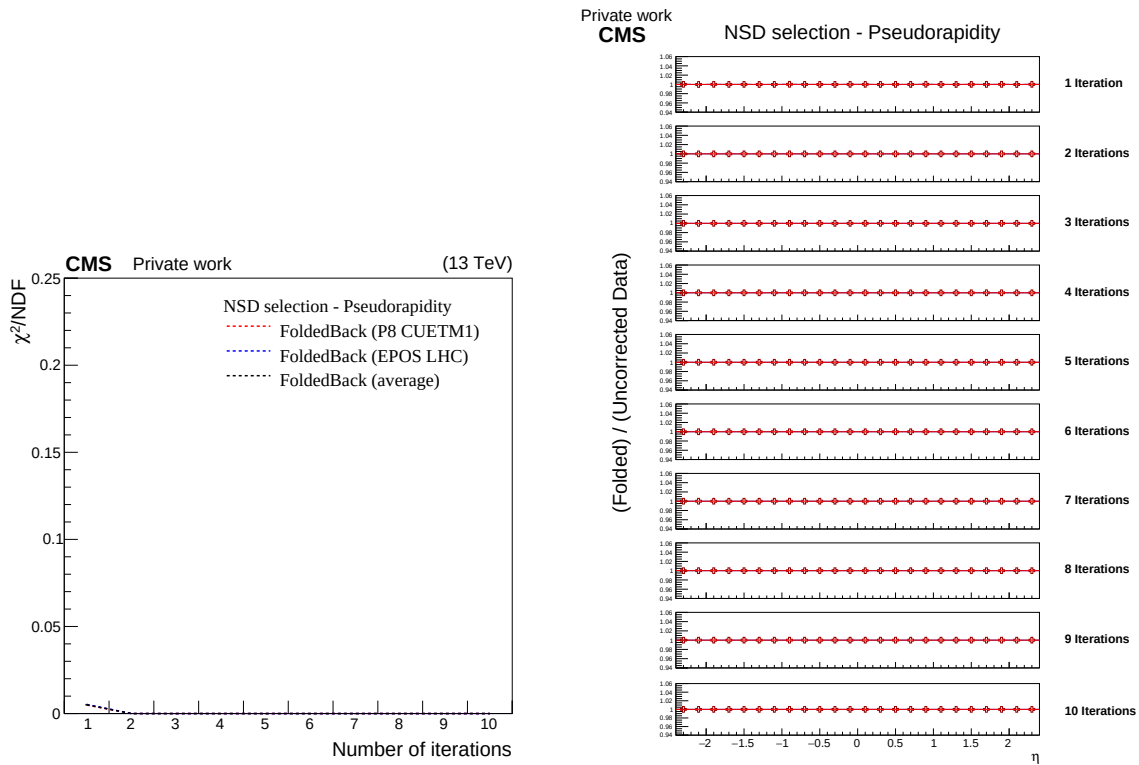


Figure 8.17: On the left, the χ^2 over the number of degrees of freedom (NDF). On the right, the ratio (*Folded*/*Uncorrected Data*) as a function of the pseudorapidity of charged particles for the Non-Single Diffractive enhanced sample, shown for different number of iterations of the D'Agostini Unfolding method.

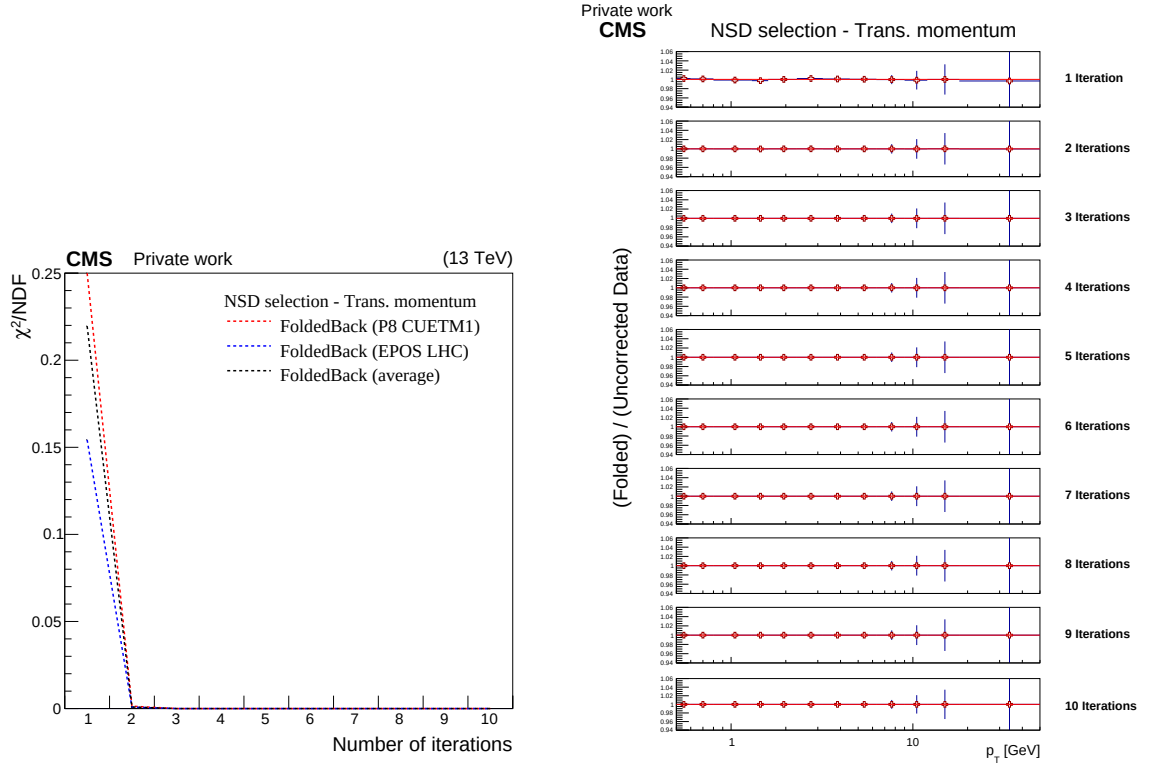


Figure 8.18: On the left, the χ^2 over the number of degrees of freedom (NDF). On the right, the ratio $(\text{Folded}/\text{Uncorrected Data})$ as a function of the transverse momentum of charged particles for the Non-Single Diffractive enhanced sample, shown for different number of iterations of the D'Agostini Unfolding method.

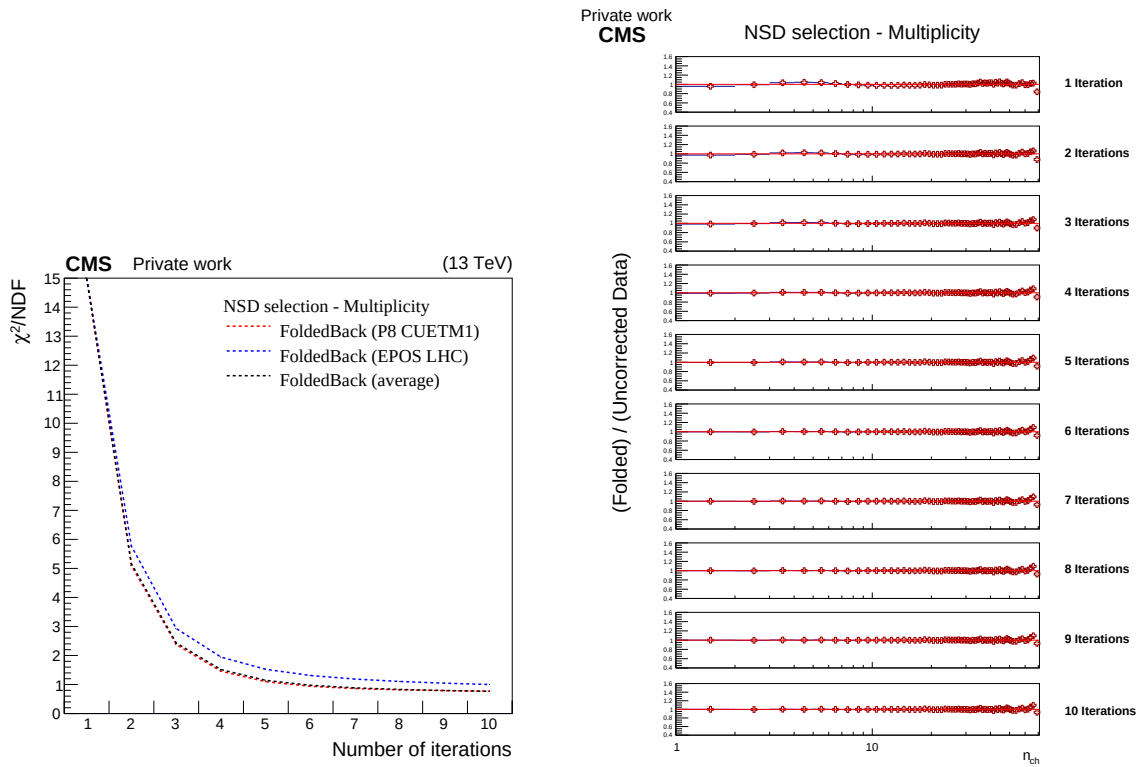


Figure 8.19: On the left, the χ^2 over the number of degrees of freedom (NDF). On the right, the ratio (*Folded*/*Uncorrected Data*) as a function of the multiplicity of charged particles for the Non-Single Diffractive enhanced sample, shown for different number of iterations of the D'Agostini Unfolding method.

CHAPTER

9

UNCERTAINTIES

Contents

9.1	Model dependence	133
9.2	Event selection	134
9.3	Tracking efficiency	137
9.4	Pile-up contribution	137
9.5	Total systematic uncertainties	140

In the following the different sources of systematic uncertainties are presented, as well as how they have been estimated. It is worth to mention that all the histograms shown here for the uncertainties are in absolute value. For the final total uncertainty calculated from the different sources, the sign of each of the uncertainties is taken into account and an asymmetric error band is obtained.

9.1 Model dependence

The systematic uncertainties associated to the model dependence of the correction procedure are estimated by taking half of the difference of the corrected distributions determined with PYTHIA8 and EPOS. It is important to notice that this uncertainty is introduced when the measurement is corrected for the efficiency with ϵ_{Eff} as defined in 8.6. The reason for this is that the calculation of this correction factor depends on the MC predictions.

Figure 9.1 shows the systematic uncertainties due to the model dependence of the correction procedure as a function of the pseudorapidity η , multiplicity n_{ch} and transverse momentum p_T . In the case of the transverse momentum, the leading and integrated leading distributions are also shown.

The uncertainties due to the model dependence are approximately constant over the full η range for all the event selections, with values of the order of 0.5 – 1% for the Inclusive, Inelastic-enhanced and NSD-enhanced event samples, and of the order of 7 – 8% for the SD-enhanced selection.

In the case of the transverse momentum, the behaviour is slightly increasing towards the high p_T , from 0.5% to a maximum of 4% in the case of the Inclusive, Inelastic-enhanced and NSD-enhanced event samples, and the opposite behaviour is found in the case of the SD-enhanced sample, where it peaks at 10% for p_T of around 2 GeV and decreases as the p_T grows.

The model dependence uncertainty for the multiplicity distributions ranges from 0% to $\sim 6\%$ in the case of the Inclusive, Inelastic-enhanced and NSD-enhanced event samples, and shows a large impact in the SD-enhanced sample case, with values of up to 42%. This large model dependence in the SD-enhanced sample is coming from the different description of the detector efficiency of the two Monte Carlo used.

9.2 Event selection

The systematic uncertainties associated to the event selection are determined by varying the threshold applied to the energy of the HF calorimeter towers. The default value of the energy threshold is varied from 5 GeV to 4 GeV and 6 GeV, corresponding to an uncertainty of 20% on the HF energy scale. Figure 9.2 shows the systematic uncertainties associated to the event selection as a function of the pseudorapidity η , multiplicity and transverse momentum p_T . The uncertainty is also shown as a function of the leading charged particle p_T and the min leading p_T for the case of the integrated distribution. The inclusive sample is excluded from these plots since it does not depend on the HF event selection.

The uncertainties due to the event selection are approximately constant over the full η range for all the event selections, with values of the order of 0.1%, 0.8% for the Inelastic-enhanced and NSD-enhanced event samples, and of the order of 5 – 7% for the SD-enhanced selection.

The transverse momentum distributions also show a rather flat behaviour for the case of the Inelastic-enhanced and NSD-enhanced samples, and a rising behaviour in the case of the SD-enhanced sample, with values as low as 5% to values up to 50%.

In the case of the multiplicity distributions, the Inelastic-enhanced and NSD-enhanced samples show again a flat behaviour at around 1 – 2%, with the NSD-enhanced sample showing a small rise at low multiplicities up to 5%. The SD-enhanced sample is the one suffering the largest impact of this uncertainty source, showing an increasing trend from 0% to up to 50%.

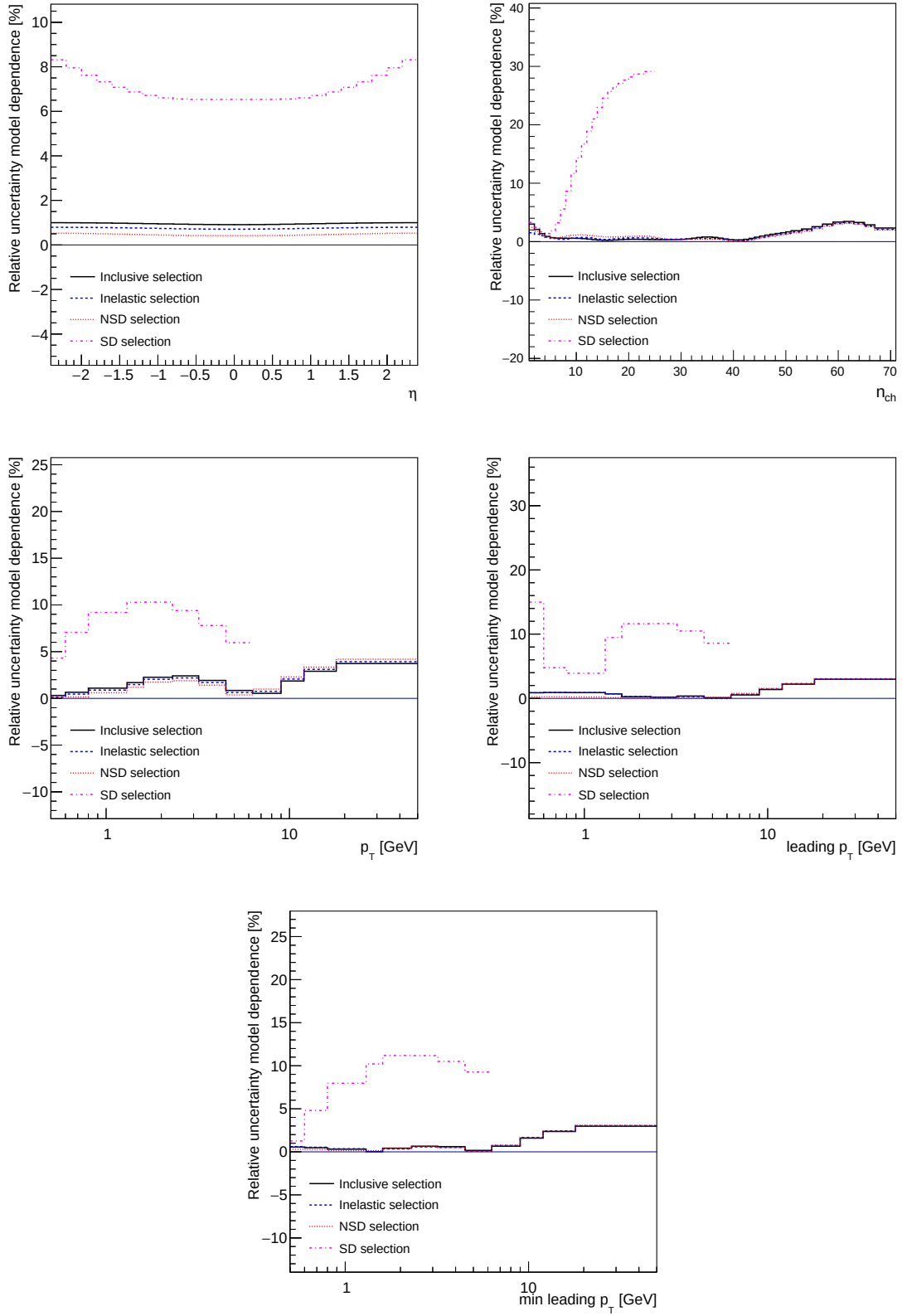


Figure 9.1: Systematic uncertainties due to the model dependence of the correction procedure as a function of η , n_{ch} and p_T . For the p_T , the leading and integrated leading distributions are also shown.

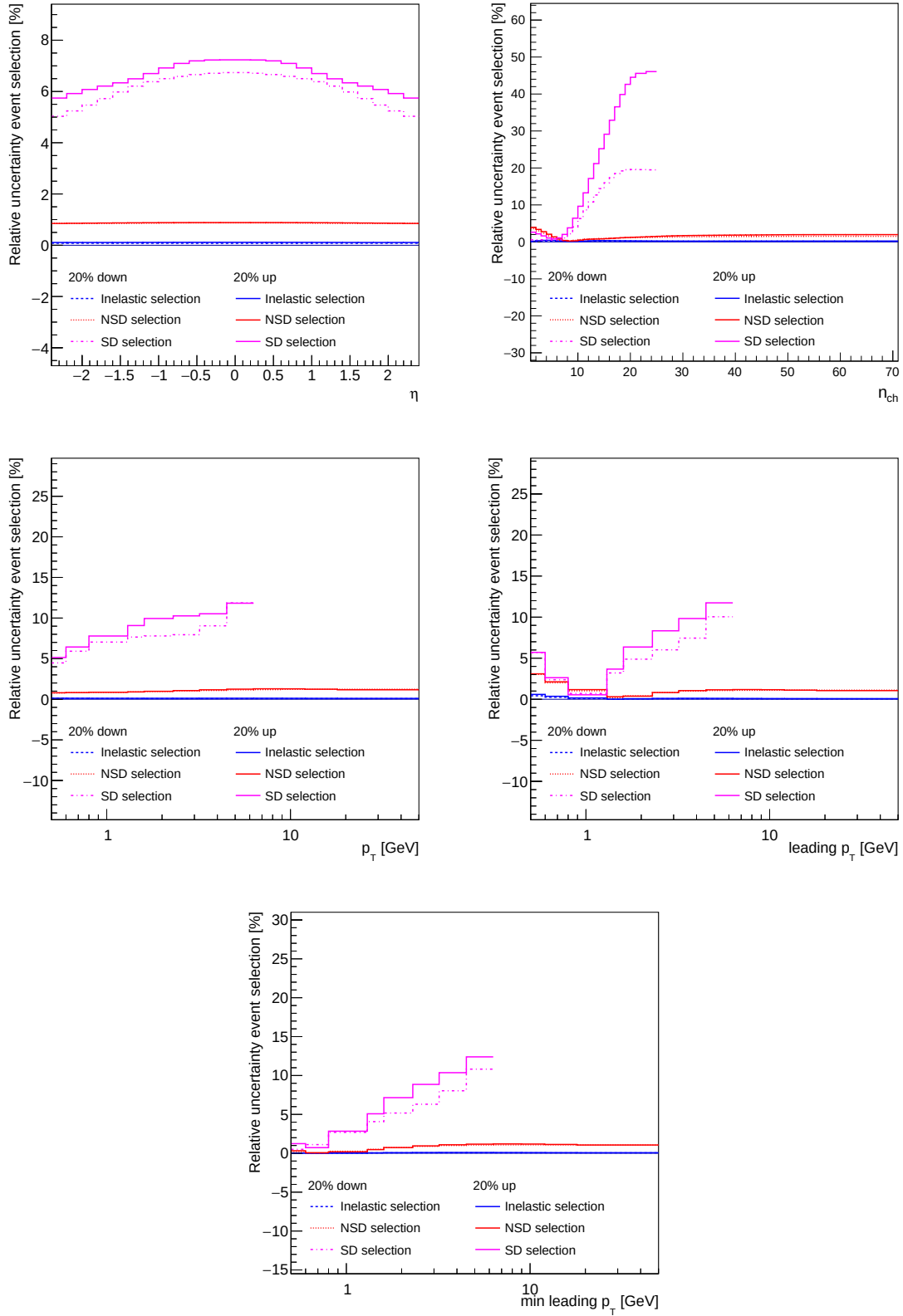


Figure 9.2: Systematic uncertainties due to the event selection based on HF as a function of η , n_{ch} and p_T . The uncertainty is also shown as a function of the leading charged-particle p_T and the min leading p_T for the case of the integrated distribution.

9.3 Tracking efficiency

A value of 4% is assigned to the systematic uncertainty associated to the difference between the track reconstruction efficiency in data and simulation. This value was used in the previous analysis at a c.m.e.¹ of $\sqrt{s} = 7$ TeV. In order to validate this choice, the performance of the track reconstruction in the Run II data taking period is compared to the performance observed in Run I. The distributions of the pseudorapidity, azimuthal angle and transverse momentum of the tracks are shown in figure 9.3 for the data sample used in the analysis, using the nominal Run II conditions and the previous Run I conditions for the track reconstruction respectively. The distributions of the relative difference between the results obtained with the Run I and the Run II conditions are shown in figure 9.4. The relative difference is of the order of a few percent and nearly constant as a function of the azimuthal angle and transverse momentum of the tracks. Some structure, nearly symmetrical around the central value, is observed for the pseudorapidity distribution. The agreement between the performance of the track reconstruction obtained using the Run I and the Run II conditions validates the value of 4% to estimate the systematic uncertainty.

9.4 Pile-up contribution

The average number of collisions per bunch crossing is equal to 1.3 [79] in the data sample used for the analysis, and the measurements can be affected in several ways by the pileup contribution. The track multiplicity at the primary vertex can be biased to higher values by the merging of two or several interactions into a single reconstructed vertex at detector level. The activity in the HF calorimeters is also affected. Particles produced in different interactions can overlay in a same calorimeter tower, producing migrations between the different event topologies which are defined at detector level by the energy deposition in the HF η range. An estimation of the systematic uncertainties due to the modelling of the pileup contribution is described in the following. The nominal event selection requires the presence of exactly one reconstructed vertex in order to reject beam background events and events with more than one collision per bunch crossing. The full analysis is repeated with an alternative event selection requiring the presence of at least one reconstructed vertex in the event. Then, the different measured distributions are determined from the track collection associated to the first primary vertex defined as the one with the highest value of the sum of the magnitude of the transverse momenta of its associated tracks.

All distributions are corrected to the stable particle level defined in section 5.4, and the difference between the default and alternative event selection at detector level should be taken into account by the respective correction procedures. The difference between the corrected distributions therefore originates from the mismodelling of the pileup contribution in the simulation, and can be taken as an estimate of the systematic uncertainties associated to the pileup description in the Monte Carlo.

Figure 9.5 shows the systematic uncertainties associated to the treatment of the pileup as a function of the pseudorapidity η , multiplicity n_{ch} and transverse momentum p_{T} . The

¹As defined in Chapter 2.

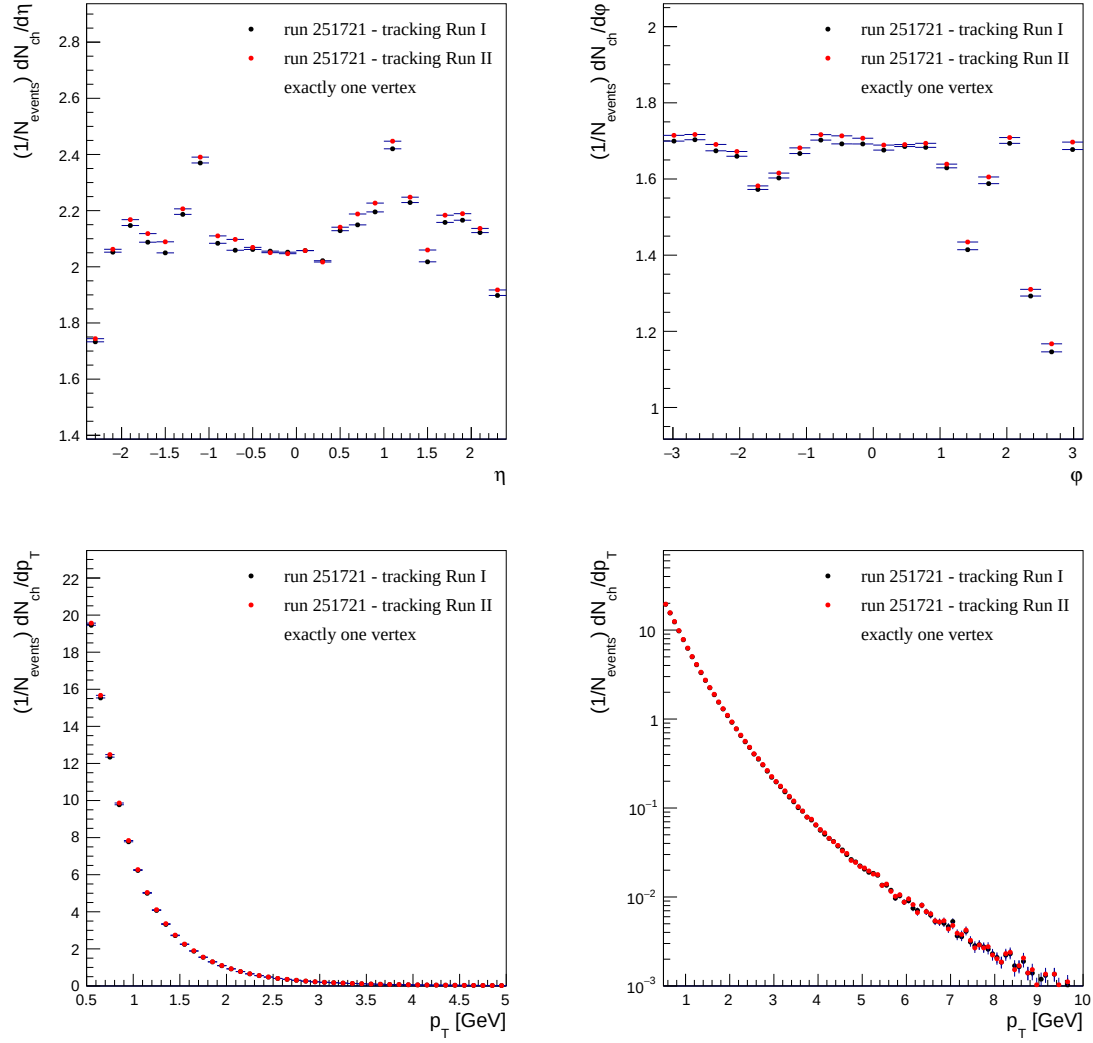


Figure 9.3: Distributions of the η (top left), azimuthal angle (top right) and p_T (bottom - linear and logarithmic scales) of the tracks for the data sample used in the analysis, using the nominal Run II conditions and the previous Run I conditions for the track reconstruction respectively.

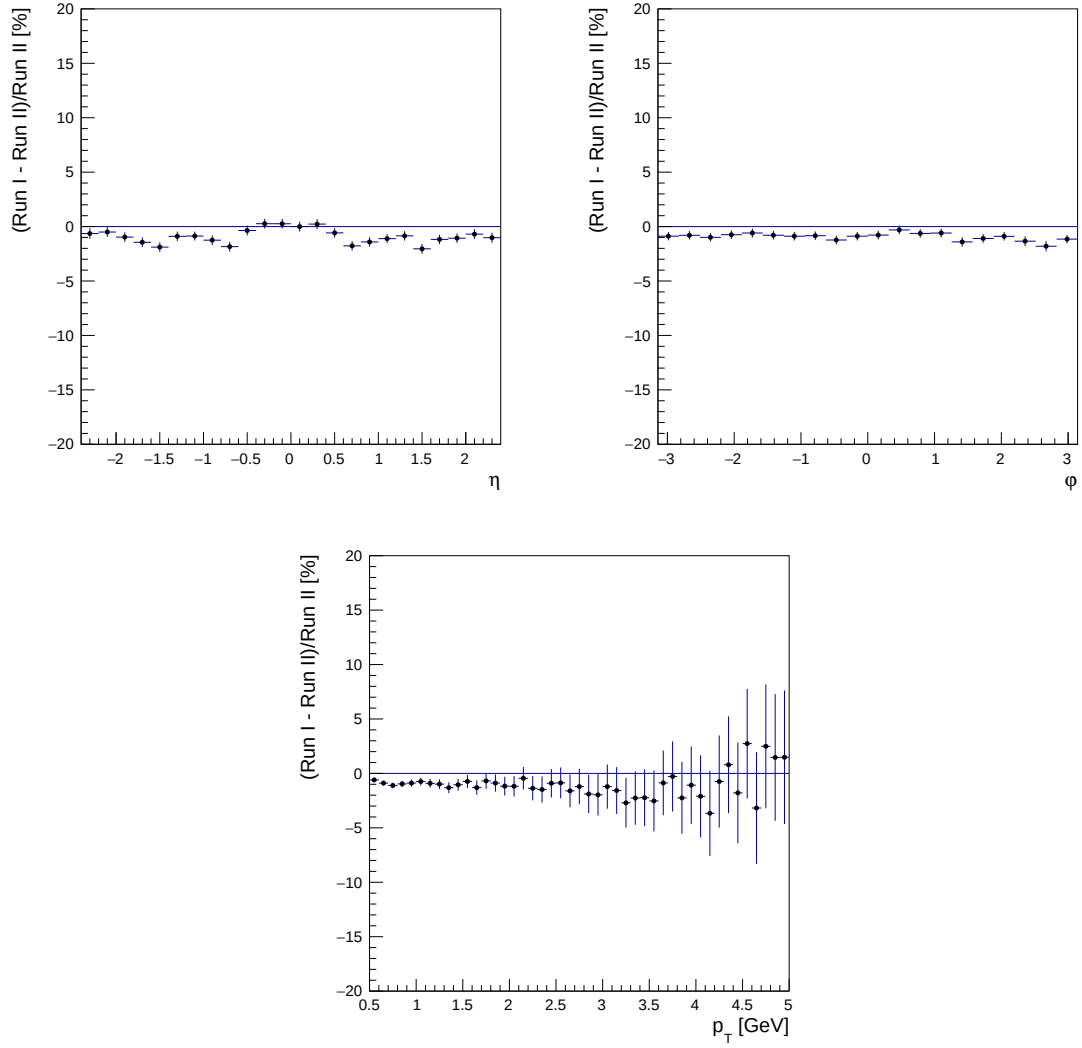


Figure 9.4: Distributions of the relative difference between the results obtained with the Run I and the Run II conditions for η , azimuthal angle and p_T of the tracks.

uncertainty is also shown as a function of the leading charged particle p_T and the min leading p_T for the case of the integrated distribution.

The uncertainties associated to the pileup description as a function of pseudorapidity, show an approximate flat behaviour over the full range, with average value of 0.2, 1.0 and 1.6% for the SD-, Inelastic- (and Inclusive) and NSD-enhanced samples respectively.

For the case of the transverse momentum, an average value of 1% over the full range for the Inclusive, Inelastic-enhanced and NSD-enhanced samples is obtained, while for the SD-enhanced sample there is an increase up to 7% for high p_T values.

In the multiplicity distributions, the modelling of the pileup leads to uncertainty values ranging from 0% to 8% in the case of the Inclusive, Inelastic-enhanced and NSD-enhanced samples. The apparent change of direction around multiplicities of 24 comes from taking the absolute value of the uncertainty curves. The SD-enhanced sample presents a roughly constant behaviour around 1%.

9.5 Total systematic uncertainties

Figure 9.6 shows the total systematic uncertainties as a function of the pseudorapidity η , multiplicity and transverse momentum p_T obtained by taking into account the different sources of uncertainties due to the model dependence, the event selection, the difference of tracking efficiency in data and simulation and the treatment of the pileup contribution. The total uncertainty is also shown as a function of the leading charged particle p_T and the min leading p_T for the case of the integrated distribution.

The total systematic uncertainties show a constant behaviour as a function of η , with values of the order of 4% for the Inclusive, Inelastic-enhanced and NSD-enhanced selections, and 9% for the SD-enhanced sample. After correcting to stable particle level, the charged-particle pseudorapidity densities are averaged and symmetrized over the η range of the measurement, and half of the difference between the corresponding values on the positive and negative η sides is taken as an additional systematic uncertainty. The relative difference between the corresponding densities on the positive and negative η sides is smaller than 0.5% in the central pseudorapidity region, and reaches values of 1% at most forward pseudorapidities.

The total uncertainty for the multiplicity distributions show an approximate flat behaviour over the full range for the Inclusive, Inelastic-enhanced and NSD-enhanced selections with values of the order of 4 – 8%, while the SD-enhanced selection shows a steep growth from 10% to 45%.

Table 9.1 summarises all the contributions per observable and per event selection. The total uncertainty is also reported for each case.

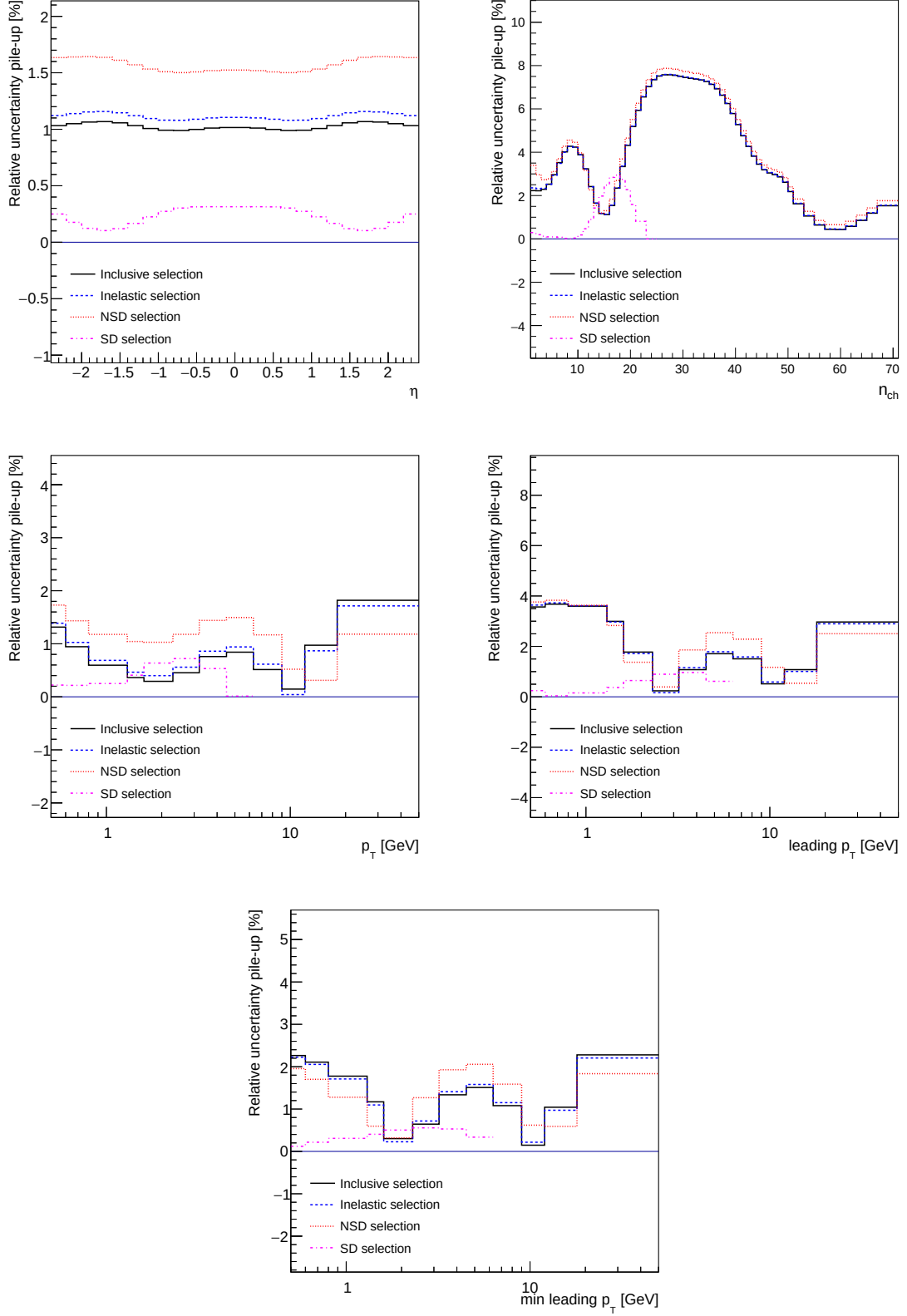


Figure 9.5: Systematic uncertainties associated to the description of the pileup contribution as a function of η , n_{ch} and p_T . The uncertainty is also shown as a function of the leading charged-particle p_T and the min leading p_T for the case of the integrated distribution.

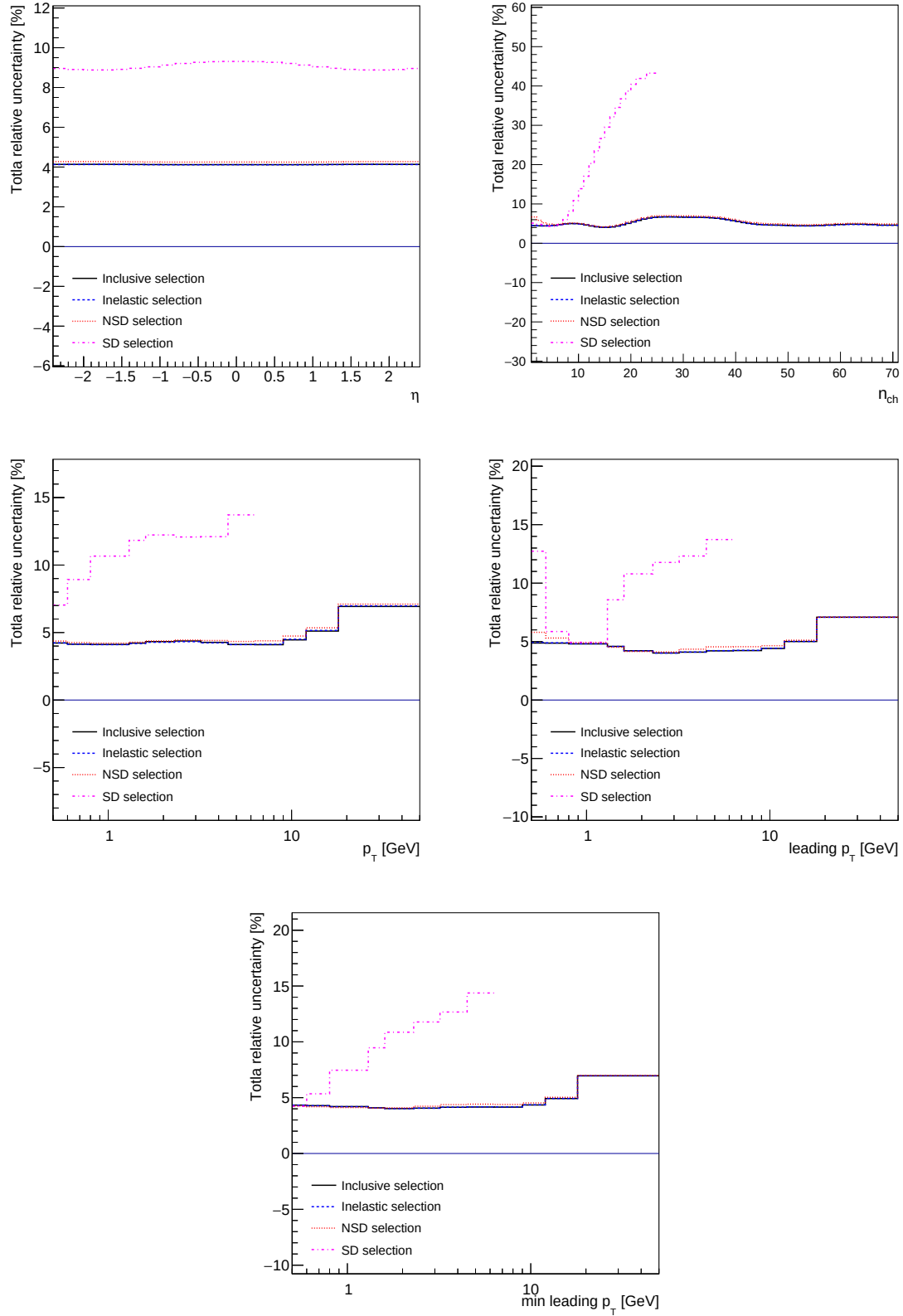


Figure 9.6: Total systematic uncertainties as a function of η , n_{ch} and p_T . Also shown as a function of the leading charged-particle p_T and the min leading p_T for the case of the integrated distribution.

Systematic uncertainties [%]					
Observable	Systematic uncertainty	Inclusive selection	Inelastic selection	NSD selection	SD selection
η	Model dependence	1	0.8	0.5	7
	Event selection	N.A.	0.2>	1	7
	Pileup modelling	1	1	1.6	0.3
	Tracking efficiency	4	4	4	4
	Total	4	4	4	9
n_{ch}	Model dependence	0 – 4	0 – 4	0 – 4	4 – 30
	Event selection	N.A.	0 – 2	0 – 2	0 – 50
	Pileup modelling	0 – 8	0 – 8	0 – 8	0 – 3
	Tracking efficiency	4	4	4	4
	Total	4 – 8	4 – 8	4 – 8	5 – 45
p_T	Model dependence	0 – 4	0 – 4	0 – 4	4 – 10
	Event selection	N.A.	0.2>	1	1 – 12
	Pileup modelling	0 – 2	0 – 2	0.5 – 2	0 – 1
	Tracking efficiency	4	4	4	4
	Total	4 – 7	4 – 7	4 – 7	7 – 14
$p_T^{leading}$	Model dependence	0 – 3	0 – 3	0 – 3	4 – 14
	Event selection	N.A.	0.2>	1	1 – 12
	Pileup modelling	0 – 4	0 – 4	0.5 – 4	0 – 1.5
	Tracking efficiency	4	4	4	4
	Total	4 – 7	4 – 7	4 – 7	5 – 14
$\int p_T^{leading}$	Model dependence	0 – 4	0 – 4	0 – 4	1 – 11
	Event selection	N.A.	0.2>	1	1 – 12
	Pileup modelling	0 – 2	0 – 2	0.5 – 2	0 – 0.5
	Tracking efficiency	4	4	4	4
	Total	4 – 7	4 – 7	4 – 7	4 – 15

Table 9.1: Summary of systematic uncertainties for each observables for the Inclusive, Inelastic-enhanced, NSD-enhanced and SD-enhanced event samples.

CHAPTER

10

RESULTS

Contents

10.1 Comparison to Monte Carlo event generators	146
10.2 MPI effects	154
10.3 Saturation effects	159
10.4 Hard contribution to diffractive events	164

In this chapter, the results obtained from the study performed in this thesis are presented corrected to stable particle level. The total systematic uncertainties are obtained by taking into account the different sources of uncertainties discussed in Chapter 9. They are shown as a shaded band encompassing the corrected data points. At the bottom of each plot, the ratios of the predictions to the data are shown. It is important to notice that the scales of the ratio plots are adjusted accordingly each time.

Figure 10.1 shows the event classes corresponding to the Inclusive, Inelastic-enhanced, NSD-enhanced and SD-enhanced event samples. They are shown as a function of each of the measured observables, namely, η , p_T , leading p_T , the integrated leading p_T as a function of $p_{T,\min}$ and n_{ch} , as defined in the previous sections. The superposition of the different samples allows to compare how the particle density changes from the Inclusive to the diffractive-enhanced event samples.

The η distribution corresponding to the SD-enhanced event sample is also presented as a symmetrized distribution between the SD-minus and SD-plus enhanced event samples and is referred to as SD-One-Side enhanced event sample. The η distributions are averaged and symmetrized over the positive and negative η ranges to suppress statistical fluctuations. The symmetrization is done by taking the mirror-like reflection with respect to the y-axis at $\eta = 0$. SD-minus and SD-plus correspond to the mutually exclusive event samples depending on the

side of the forward-detector that contains the hadronic activity.

In the case of the η densities, a reduction by a factor of $\sim 1/3$ is observed in the SD-enhanced sample with respect to the most inclusive samples. In the p_T distributions (i.e. p_T , leading p_T and integrated leading p_T), the SD-enhanced event sample presents the most steeply falling spectrum at large p_T values. For the n_{ch} distributions the NSD-enhanced event sample shows a depletion of low-multiplicity events and an increase of high-multiplicity events with respect to the Inclusive and Inelastic-enhanced event samples. In the case of the n_{ch} distributions, the SD-enhanced sample is not included as concluded from the unfolding validation studies performed in Chapter 8.

In the following section, the predictions from different MC event generators are compared to the corrected data. In addition, some of the most relevant parameters related to the particle production and the transition between soft and hard diffraction are studied for the MC event generator PYTHIA8. The Inclusive event sample is omitted in the following since it does not provide additional information with respect to the Inelastic-enhanced event sample.

10.1 Comparison to Monte Carlo event generators

The corrected data are compared to the predictions of the Monte Carlo event generators PYTHIA8 tune CUETP8M1, PYTHIA8 MBR tune 4C and EPOS LHC.

Figure 10.2 shows the η densities of charged-particles. For the Inelastic-enhanced event sample, the predictions of EPOS LHC give the best description of the data. The predictions of PYTHIA8 CUETP8M1 slightly underestimate the measurements, while those of PYTHIA8 MBR tune 4C overestimate them. For the NSD-enhanced event sample, the predictions of EPOS LHC and PYTHIA8 CUETP8M1 both give a reasonable description of the data, while the predictions of PYTHIA8 MBR tune 4C still overestimate the measurements. For the SD-enhanced samples, the opposite behaviour is observed, with the predictions of EPOS LHC underestimating the data, and the predictions of PYTHIA8 CUETP8M1 overestimating them. The predictions of PYTHIA8 MBR tune 4C nicely describe the data, showing small deviations at the edges of the phase space. For the SD-One-Side enhanced event sample, the predictions of EPOS LHC and PYTHIA8 CUETP8M1 are able to describe the charged-particle η density on the η side associated with the direction of the diffractively scattered proton, but fail to describe the activity on the η side of the diffractively produced final state. The description of the SD-One-Side by the predictions of PYTHIA8 MBR tune 4C is excellent over almost the full range, only showing some deviations at the side of the diffractive system.

Figure 10.3 shows the charged-particle multiplicity distributions for the Inelastic-enhanced and NSD-enhanced event samples. The multiplicity distribution corresponding to the SD-enhanced sample has been omitted as concluded in Sec. 8.5.

A similar description of the Inelastic-enhanced and the NSD-enhanced event samples by the MC event generators is observed, only with slight differences at low multiplicities. This shows that the SD dissociation events mainly contribute at low multiplicities. The predic-

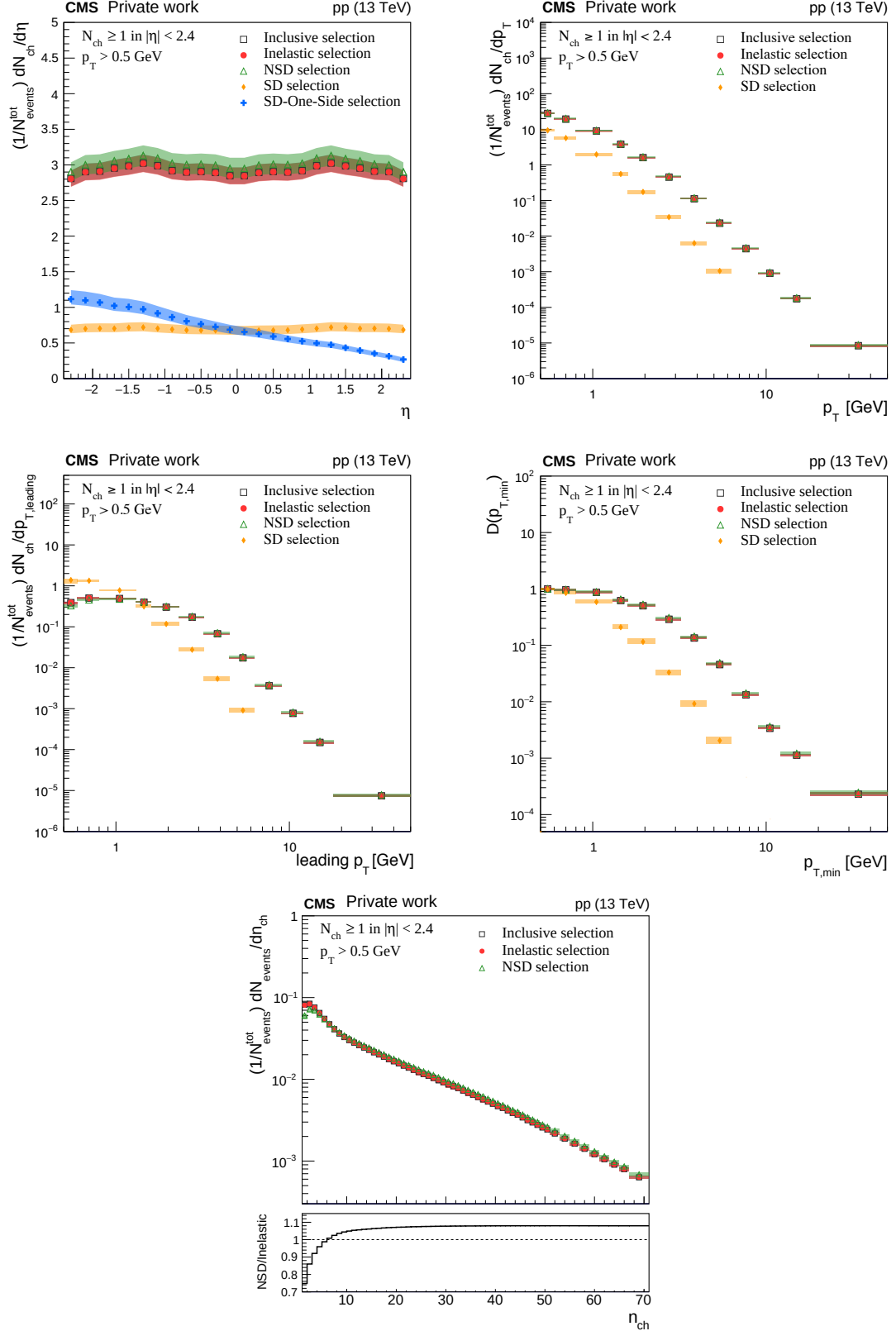


Figure 10.1: From top to bottom, left to right: η , p_T , leading p_T , integrated leading p_T and n_{ch} distributions for the charged particles present in the Inclusive (squares), Inelastic-enhanced (circles), NSD-enhanced (triangles), SD-enhanced (diamonds) and SD-One-Side enhanced (crosses) event samples. The band around the data points corresponds to the statistical and systematics uncertainties.

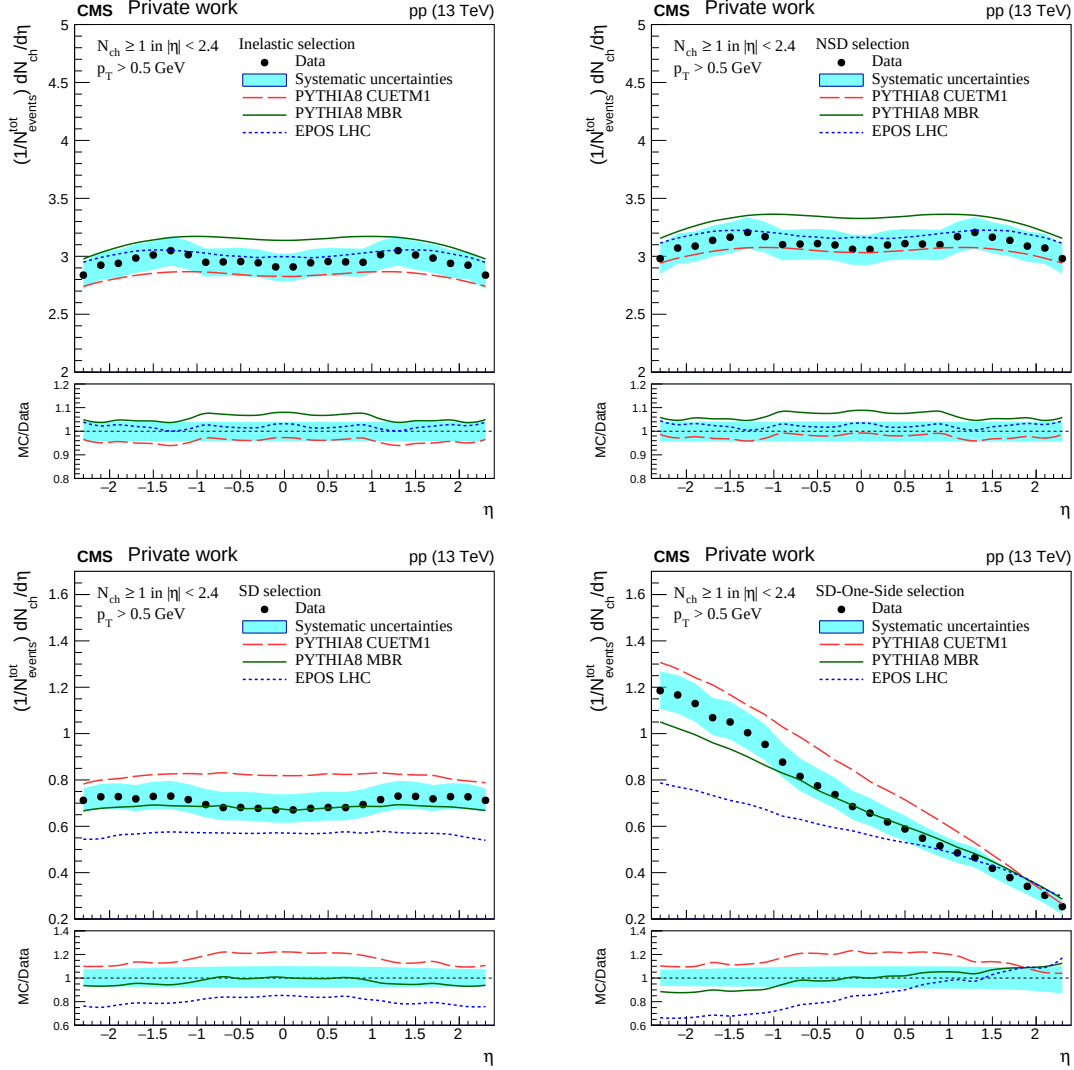


Figure 10.2: Charged-particle pseudorapidity densities averaged and symmetrized over the positive and negative η ranges. Corresponding to, from top to bottom, left to right: Inelastic-enhanced, NSD-enhanced, SD-enhanced and SD-One-Side enhanced event samples. The corrected data are compared to the predictions of the Monte Carlo event generators PYTHIA8 CUETP8M1 (long dashes), PYTHIA8 MBR tune 4C (continuous line) and EPOS LHC (short dashes). The lower plot shows the ratios of the data-to-MC predictions for the different event generators.

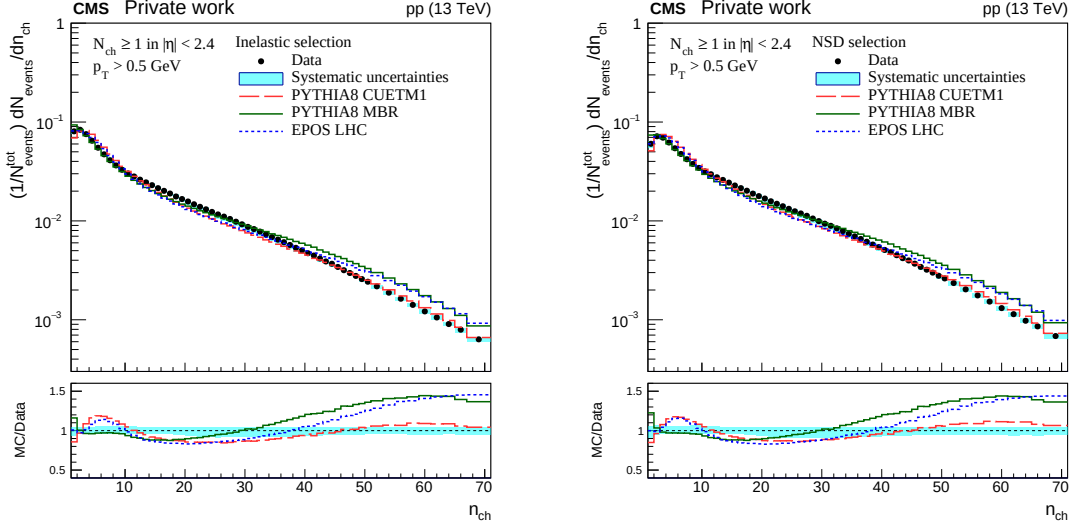


Figure 10.3: Charged-particle multiplicity distributions corresponding to, from left to right: Inelastic-enhanced and NSD-enhanced event samples. The corrected data are compared to the predictions of the Monte Carlo event generators PYTHIA8 CUETP8M1 (long dashes), PYTHIA8 MBR tune 4C (continuous line) and EPOS LHC (short dashes). The lower plot shows the ratios of the data-to-MC predictions for the different event generators.

tions of PYTHIA8 MBR tune 4C give a good description of the data for the low multiplicity region, while PYTHIA8 CUETP8M1 and EPOS LHC overestimate the data by around 20%. This is in agreement with the results of the SD pseudorapidity distribution shown before, where PYTHIA8 MBR tune 4C is the model that best describes the SD-enhanced sample. At multiplicities above 35-40, the predictions of PYTHIA8 CUETP8M1 give the best description of the data, while those of PYTHIA8 MBR tune 4C and EPOS LHC are off by up to 50%. The high multiplicity region is specially sensitive to MPI (see Sec.10.2) and improving this description would lead to a better understanding of the MPI.

Figures 10.4 to 10.6 show the charged-particle p_T distributions for all the particles, the leading particle and the integrated spectrum of the leading particle, respectively. These three distributions are shown for the Inelastic-enhanced, NSD-enhanced and SD-enhanced event samples. The p_T range for the SD-enhanced event sample is reduced in comparison to the other event samples, reaching a maximum of 6.3 GeV instead of 50 GeV. This is a consequence of the more steeply falling p_T spectrum of the SD-enhanced event sample (see figure 10.1), which requires larger event samples to reach the same p_T values as in the other two event categories.

The p_T distribution of the charged-particles in the Inelastic-enhanced event sample is best described by the predictions of PYTHIA8 CUETP8M1 almost over the complete p_T range. Small deviations of up to 10% at the low p_T region are observed. This region is dominated by particles coming from the MPI. The predictions of PYTHIA8 MBR tune 4C are able to describe the low p_T region but rapidly start to overestimate the particle production by up

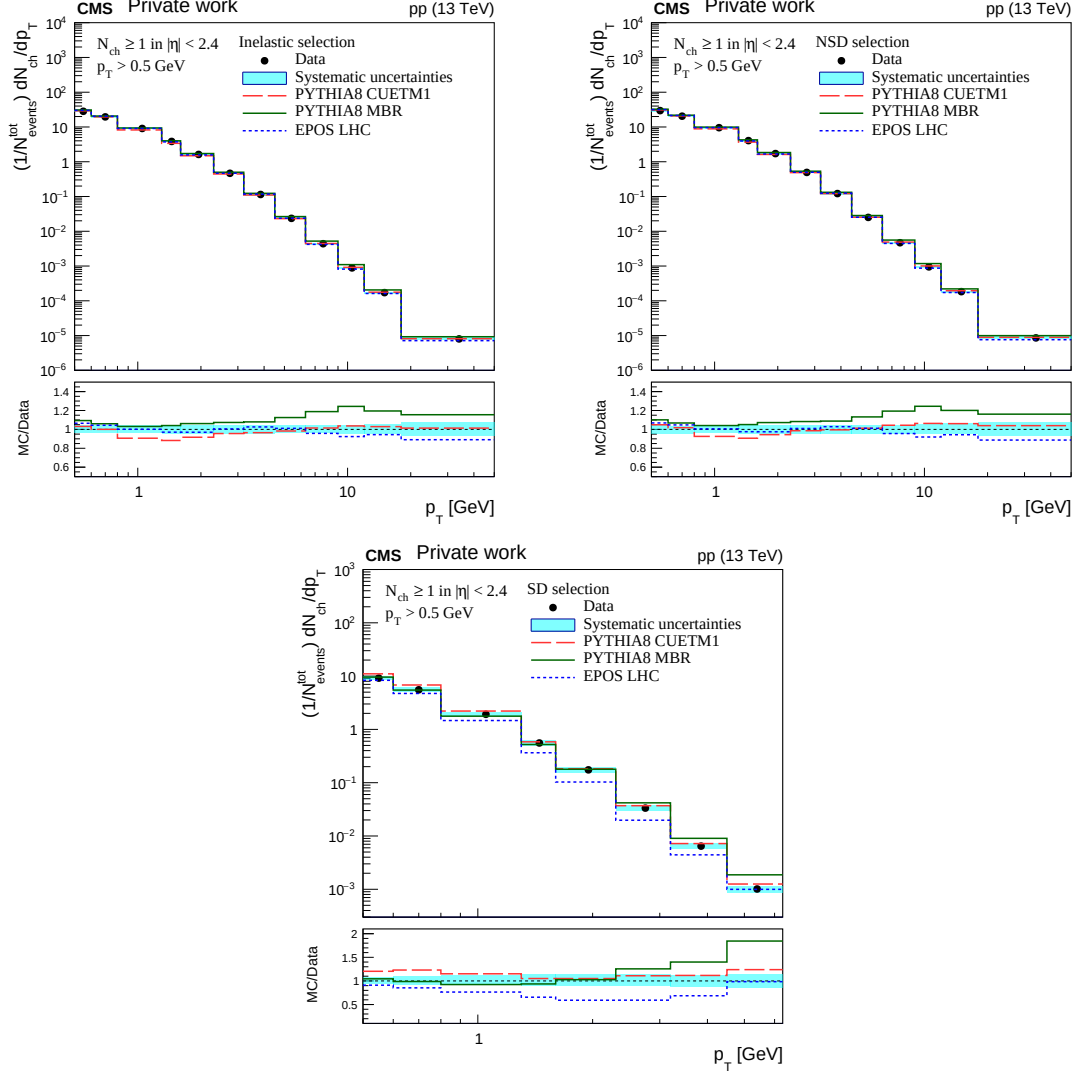


Figure 10.4: Charged-particle p_T densities corresponding to, from top to bottom, left to right: Inelastic-enhanced, NSD-enhanced and SD-enhanced event samples. The corrected data are compared to the predictions of the Monte Carlo event generators PYTHIA8 CUETP8M1 (long dashes), PYTHIA8 MBR tune 4C (continuous line) and EPOS LHC (short dashes). The lower plot shows the ratios of the data-to-MC predictions for the different event generators.

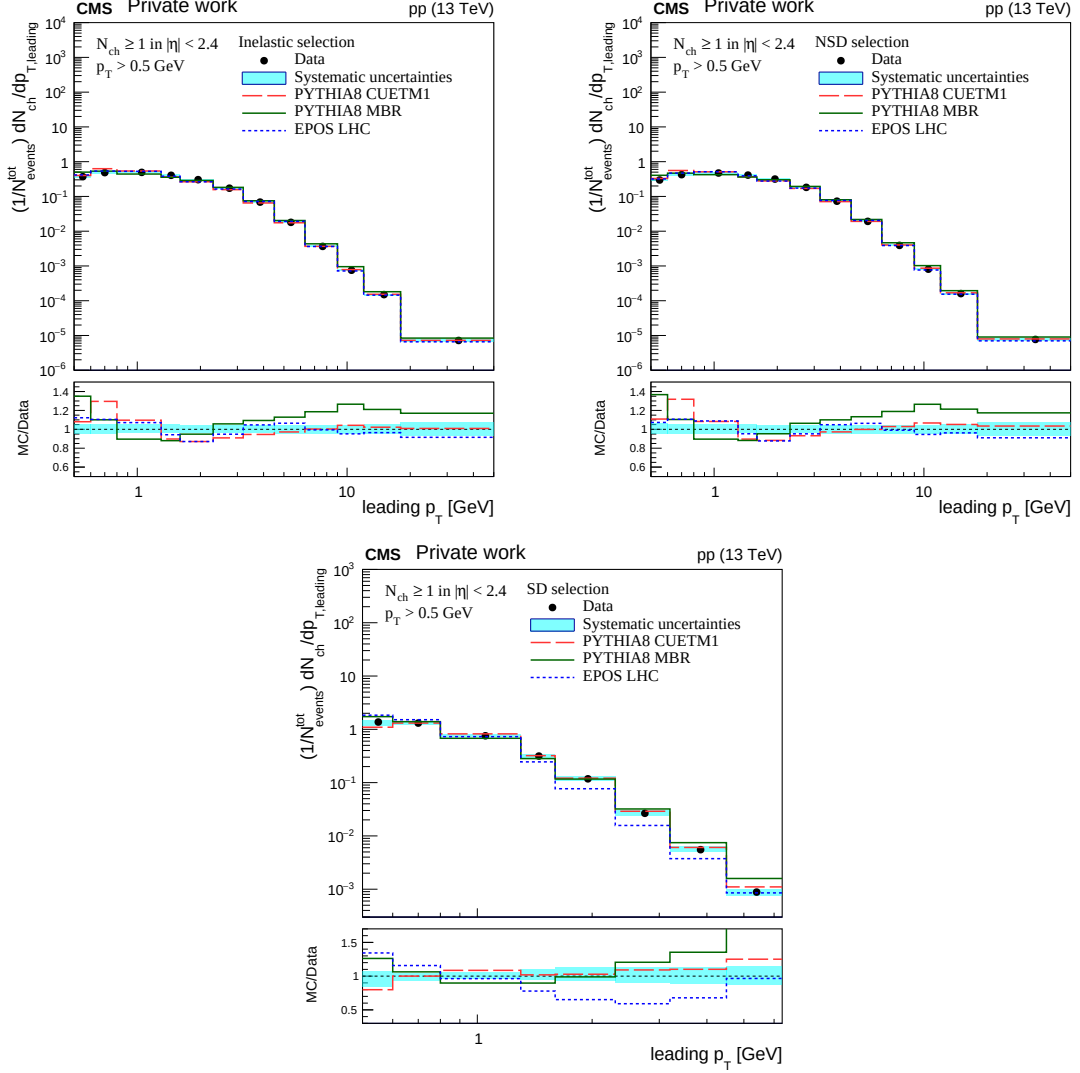


Figure 10.5: Leading charged-particle p_T densities corresponding to, from top to bottom, left to right: Inelastic-enhanced, NSD-enhanced and SD-enhanced event samples. The corrected data are compared to the predictions of the Monte Carlo event generators PYTHIA8 CUETP8M1 (long dashes), PYTHIA8 MBR tune 4C (continuous line) and EPOS LHC (short dashes). The lower plot shows the ratios of the data-to-MC predictions for the different event generators.

to 40% for $p_T > 5$ GeV. The predictions of EPOS LHC give a reasonable description of the data for transverse-momenta up to $p_T \sim 10$ GeV, but above this value they underestimate it by $\sim 10\%$. It is interesting to observe how complicated it is to correctly describe the bulk of soft particles mainly coming from MPI and the hard p_T particles primarily coming from the main hard scattering at the same time within a model.

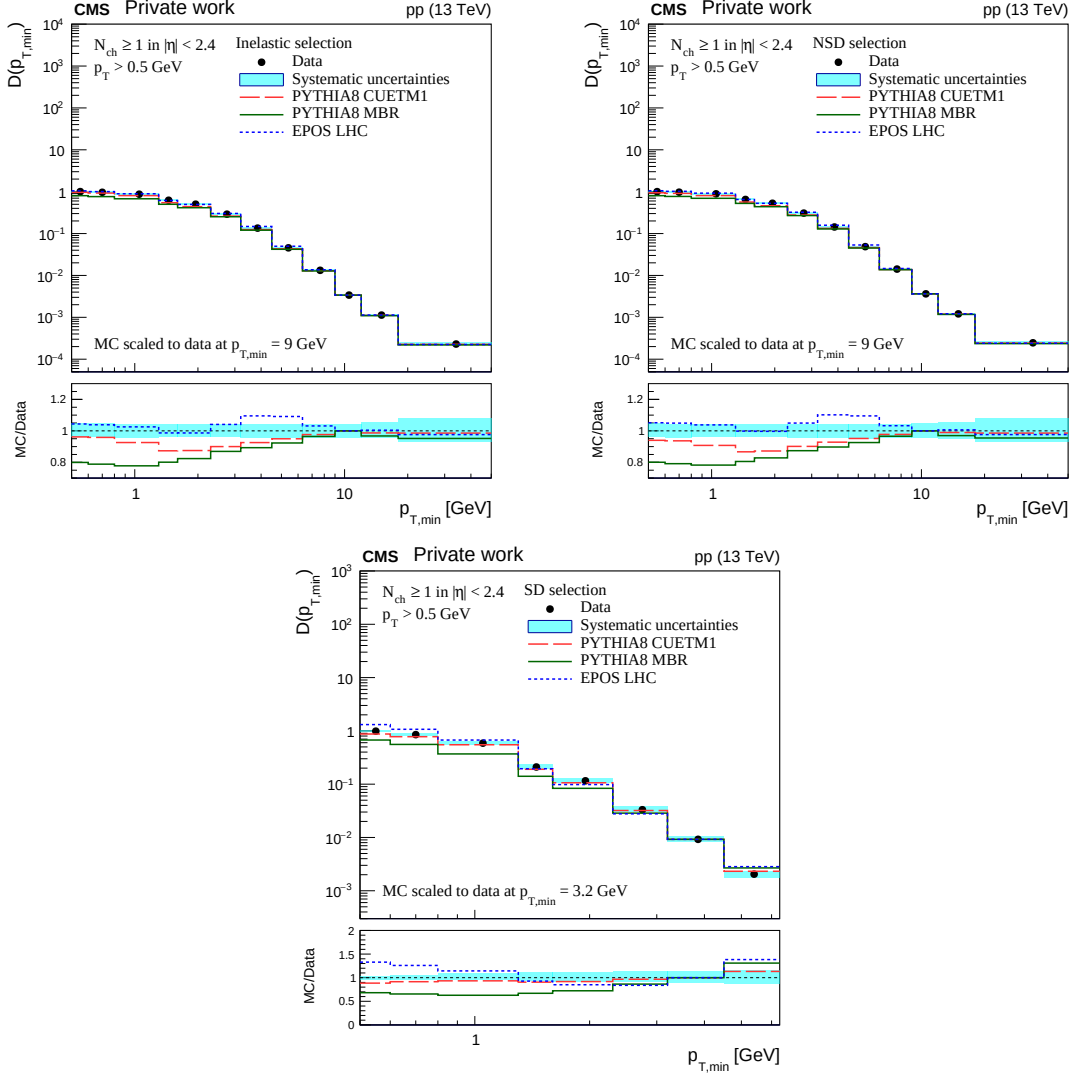


Figure 10.6: Integrated leading charged-particle p_T densities as a function of $p_{T,min}$ corresponding to, from top to bottom, left to right: Inelastic-enhanced, NSD-enhanced and SD-enhanced event samples. The corrected data are compared to the predictions of the Monte Carlo event generators PYTHIA8 CUETP8M1 (long dashes), PYTHIA8 MBR tune 4C (continuous line) and EPOS LHC (short dashes). The lower plot shows the ratios of the data-to-MC predictions for the different event generators.

The leading p_T distributions of charged particles and their integration as a function of $p_{T,min}$ are presented in figures 10.5 and 10.6, respectively. These two distributions provide

valuable information on the modelling of the transition between the non-perturbative and perturbative descriptions of the $2 \rightarrow 2$ partonic cross-section and the modelling of MPI [58]. The predictions of EPOS LHC give the best description of the data in the case of the Inelastic-enhanced and NSD-enhanced event samples, only deviating at low $p_{T,\min}$ values by up to $\sim 10\%$. For $p_{T,\min} > 4 \text{ GeV}$, the predictions of EPOS LHC and PYTHIA8 CUETP8M1 are able to reproduce the data. The predictions of PYTHIA8 MBR tune 4C are not able to describe the data neither at low nor at high $p_{T,\min}$ for any of the presented event samples. In the case of the SD-enhanced event sample, the predictions of PYTHIA8 CUETP8M1 provide the best description of the data, while those of EPOS LHC are off by up to $\sim 40\%$.

For the distribution of the integrated leading p_T charged-particle as a function of $p_{T,\min}$, the MC predictions are rescaled to the data in order to describe the high- $p_{T,\min}$ region. The motivation for the rescaling is that the theoretical predictions based on pQCD are more robust and agree better with data in this region. The exact choice of the normalisation point is arbitrary. In the case of the Inelastic-enhanced and NSD-enhanced event samples, the chosen value is $p_{T,\min} = 9 \text{ GeV}$ whereas for the SD-enhanced event sample, with a reduced range, it is $p_{T,\min} = 3.2 \text{ GeV}$. As mentioned in Sec. 2.2.4, the reinterpretation of the divergent behaviour of the $2 \rightarrow 2$ QCD partonic cross-section is intimately related with MPI. This is why a good description of the taming of $2 \rightarrow 2$ QCD partonic cross-section at low- p_T is of great interest. The Inelastic-enhanced and NSD-enhanced event samples are best described by the predictions of EPOS LHC and PYTHIA8 CUETP8M1, where the former overestimates the particle production by $\sim 10\%$ at around 4-5 GeV, and the latter underestimates it by the same amount at a lower $p_{T,\min}$ value of $\sim 1 \text{ GeV}$. The predictions of PYTHIA8 MBR tune 4C are only able to describe the data in the high- p_T region above 9 GeV but increasingly underestimate the data as moving towards lower p_T values, where discrepancies of up to $\sim 20\%$ are visible. In the case of the SD-enhanced event sample, the PYTHIA8 CUETP8M1 predictions are the best to describe the data, while those of EPOS LHC and PYTHIA8 MBR tune 4C overestimate and underestimate the data by up to $\sim 40\%$, respectively. The failure of the MC models to describe the transition from the high- to the low- p_T region suggests that the saturation of the $2 \rightarrow 2$ partonic cross-section and the MPI effects at low p_T need to be better understood.

Up to now different MC event generators have been compared to the measurements presented in this analysis. In the following, further studies to investigate the behaviour of key free parameters in PYTHIA8 are provided, namely, the simulation of MPI, the p_{T0} and diffraction.

10.2 MPI effects

Multipartonic interactions play a mayor role in the proper description of the particle densities in hadron-hadron collisions. It is also a key piece to understand the internal structure of the proton. In this section the different observables shown in the previous section are compared to the predictions of PYTHIA8 CUETP8M1 with and without the inclusion of MPI.

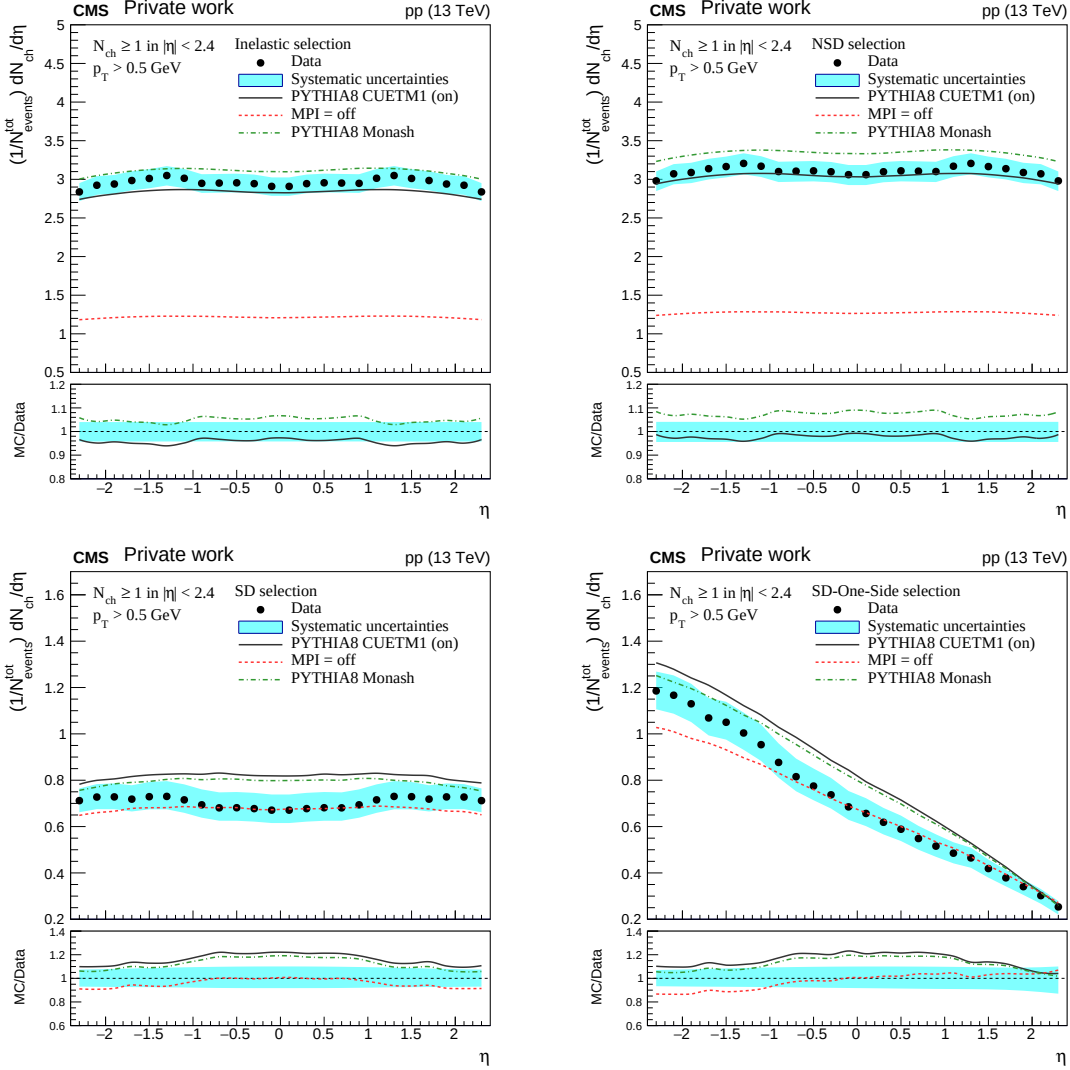


Figure 10.7: Charged-particle η densities averaged and symmetrized over the positive and negative η ranges. Corresponding to, from top to bottom, left to right: Inelastic-enhanced, NSD-enhanced, SD-enhanced and SD-One-Side enhanced event samples. The corrected data are compared to the predictions of the Monte Carlo event generator PYTHIA8 CUETP8M1 with MPI on (continuous line) and MPI off (dashed line). The predictions of PYTHIA8 with the Monash tune (dash-dot line) are also included to evaluate the contribution of the MPI independently of the specific parameters of the tune. The lower plot shows the ratios of the data-to-MC predictions for the different event generators.

The predictions of PYTHIA8 with the Monash tune are also included to evaluate the contribution of the MPI independently of the specific parameters of the tune. It is interesting to observe how different effects appear by removing the MPI from the description of the various measured observables. In some cases a pedestal effect is observed, while in others it affects the shape of the distribution.

In figure 10.7 it can be seen that for the Inelastic-enhanced and NSD-enhanced event samples about 2/3 of the particles are coming from the MPI, while in the SD-enhanced event sample it is only around 1/5. In all these cases, the effect is almost constant over the full η range. In the same figure the pseudorapidity distribution for the SD-One-Side enhanced event sample is also shown. There the MPI is mainly affecting the side of the diffractively produced system, while the side of the scattered proton remains almost unchanged. The more pronounced effect on the side of the diffractive system, suggest the possibility to study MPI in the Pomeron-proton system.

Figure 10.8 shows the n_{ch} for the Inelastic-enhanced and NSD-enhanced event samples. They show the crucial role MPI plays in order to properly describe the high multiplicity tail and to tame the low multiplicity region.

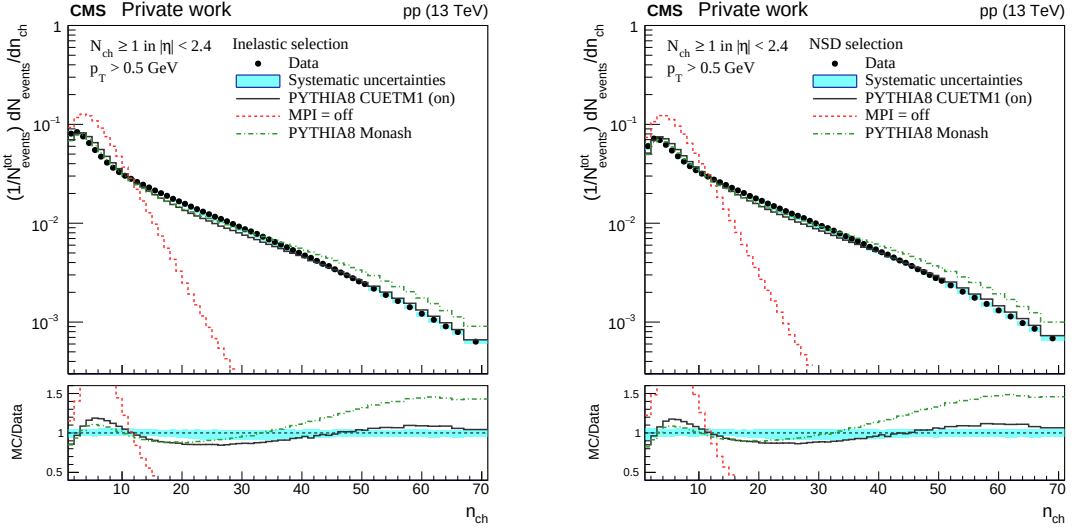


Figure 10.8: Charged-particle multiplicity distributions corresponding to, from left to right: Inelastic-enhanced and NSD-enhanced event samples. The corrected data are compared to the predictions of the Monte Carlo event generator PYTHIA8 CUETP8M1 with MPI on (continuous line) and MPI off (dashed line). The predictions of PYTHIA8 with the Monash tune (dash-dot line) are also included to evaluate the contribution of the MPI independently of the specific parameters of the tune. The lower plot shows the ratios of the data-to-MC predictions for the different event generators.

The p_{T} distributions for all the charged particles, the leading charged-particle and the integrated distribution of the latter are presented in figures 10.9, 10.10 and 10.11, respectively. In general the effects are similar for the Inelastic-enhanced and NSD-enhanced event samples, but slightly different in the case of the SD-enhanced event sample.

In figure 10.9 the Inelastic-enhanced and NSD-enhanced event samples show that switching off the MPI provokes a significant depletion of particles ($\sim 70\%$ less) at low- p_T , but no effect is observed at $p_T > 10$ GeV. On the other hand, the SD-enhanced event sample shows a smaller decrease ($\sim 20\%$ less) at low- p_T and an interesting steady increase at large- p_T .

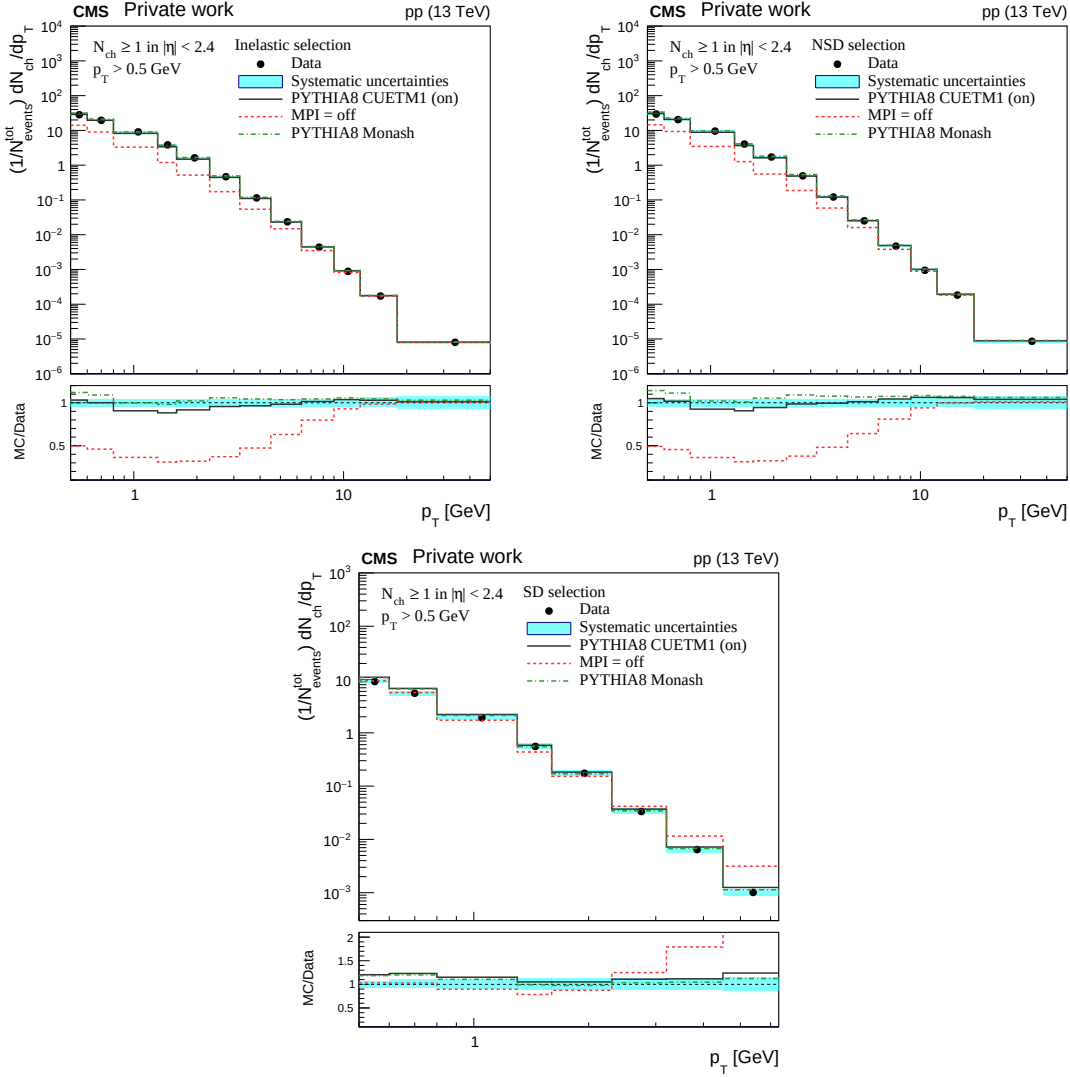


Figure 10.9: Charged-particle p_T densities corresponding to, from top to bottom, left to right: Inelastic-enhanced, NSD-enhanced and SD-enhanced event samples. The corrected data are compared to the predictions of the Monte Carlo event generator PYTHIA8 CUETP8M1 with MPI on (continuous line) and MPI off (dashed line). The predictions of PYTHIA8 with the Monash tune (dash-dot line) are also included to evaluate the contribution of the MPI independently of the specific parameters of the tune. The lower plot shows the ratios of the data-to-MC predictions for the different event generators.

Figures 10.10 and 10.11 show the leading p_T charged-particle distribution and the integral of it as a function of $p_{T,\text{min}}$, respectively. In the most inclusive selections a depletion of par-

ticles of up to $\sim 40\%$ is observed at medium transverse-momenta $2 < p_T < 10$ GeV, followed by an increase of particles at very low- p_T ($p_T < 2$ GeV) with respect to the sample with MPI included. The leading p_T particle distribution for the SD-enhanced event sample presents a similar behaviour as in the case of all-charged-particles distribution of figure 10.9, where a very small impact is observed at low- p_T and an enhancement of particles at large- p_T is present.

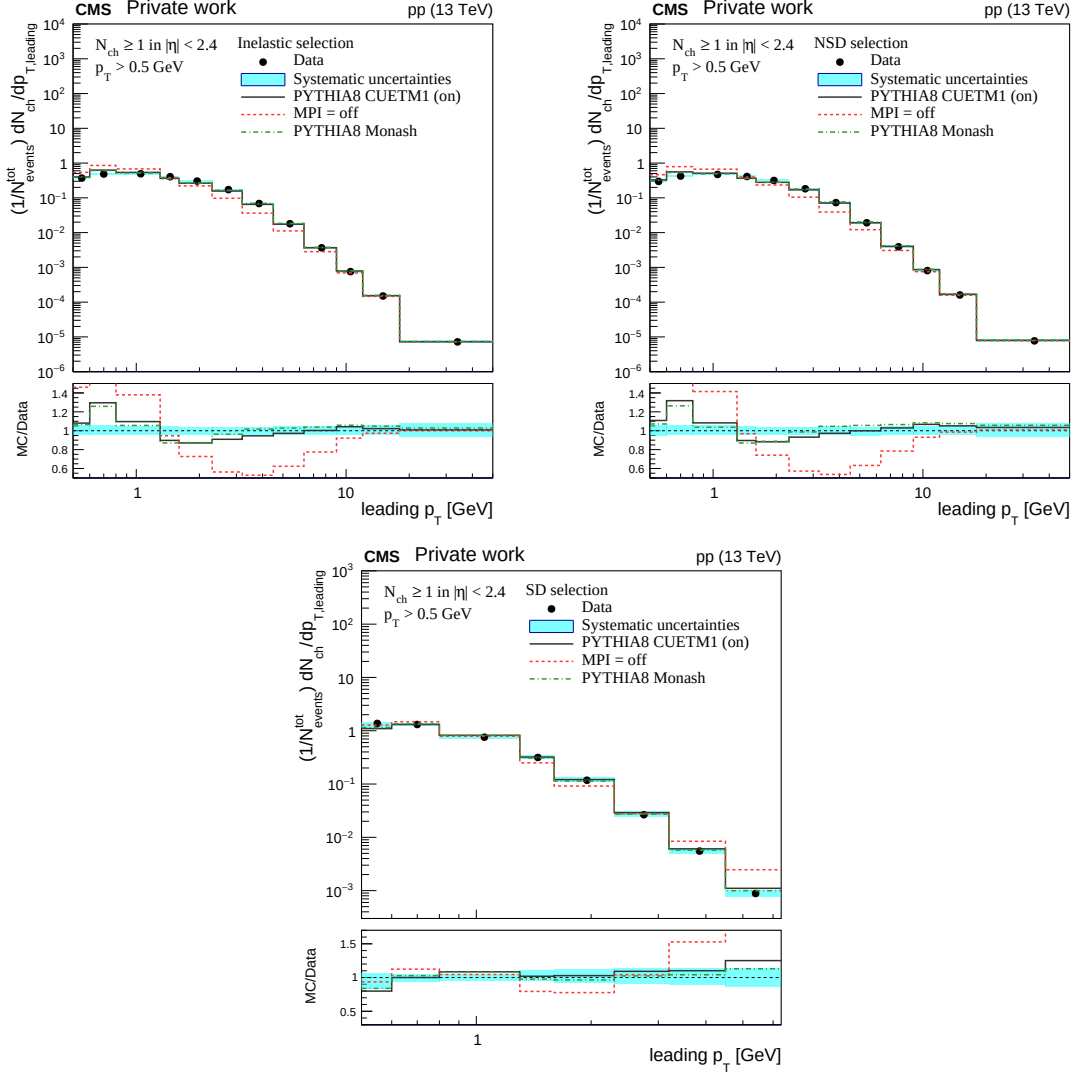


Figure 10.10: Leading charged-particle p_T densities corresponding to, from top to bottom, left to right: Inelastic-enhanced, NSD-enhanced and SD-enhanced event samples. The corrected data are compared to the predictions of the Monte Carlo event generator PYTHIA8 CUETP8M1 with MPI on (continuous line) and MPI off (dashed line). The predictions of PYTHIA8 with the Monash tune (dash-dot line) are also included to evaluate the contribution of the MPI independently of the specific parameters of the tune. The lower plot shows the ratios of the data-to-MC predictions for the different event generators.

The integrated leading p_T charged-particle distribution as a function of $p_{T,\min}$ for the different event classes shows how the inclusion of MPI in the simulation of hadronic collisions significantly improves the description of the transition region regularised in PYTHIA8 by the p_{T0} parameter.

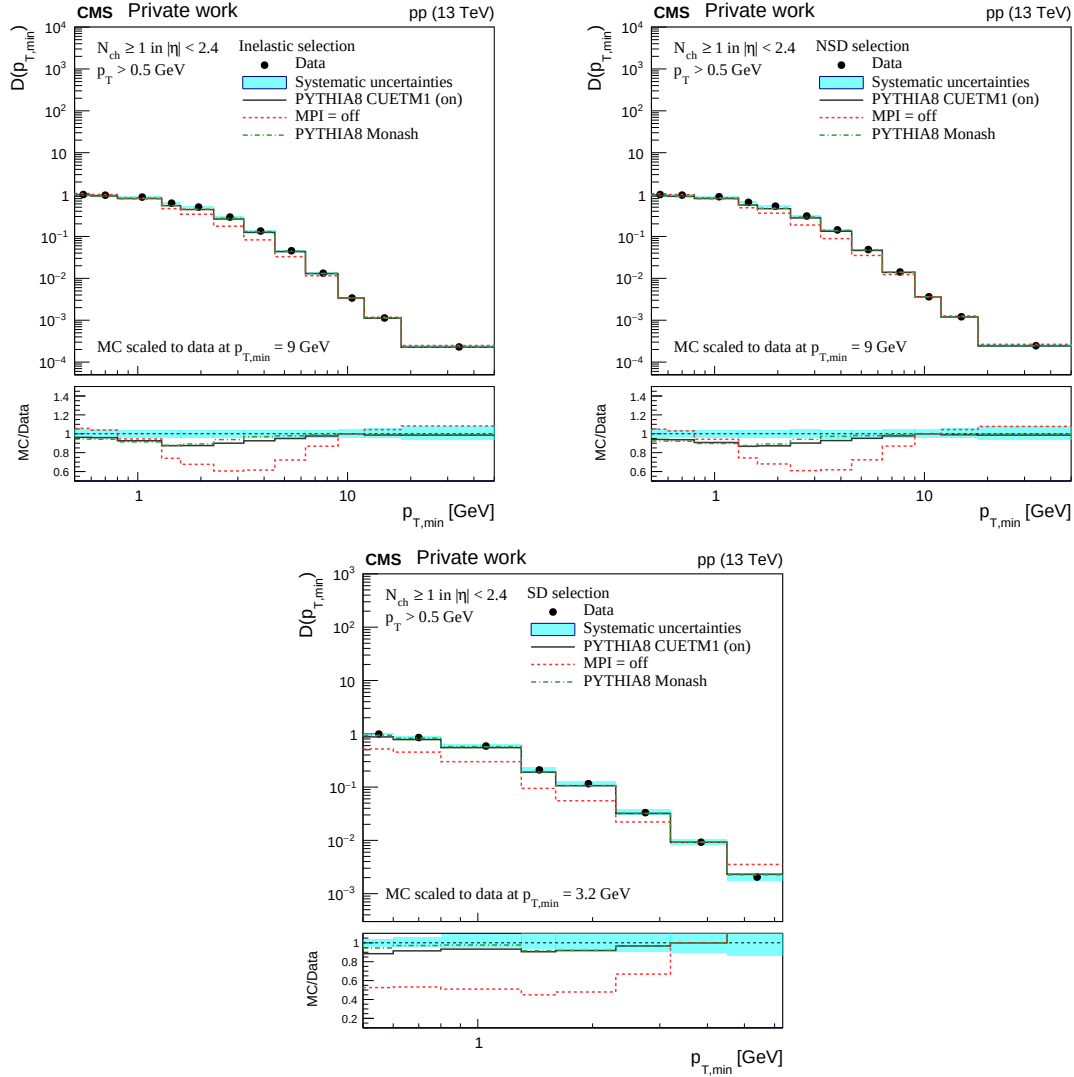


Figure 10.11: Integrated leading charged-particle p_T densities as a function of $p_{T,\min}$ corresponding to, from top to bottom, left to right: Inelastic-enhanced, NSD-enhanced and SD-enhanced event samples. The corrected data are compared to the predictions of the Monte Carlo event generator PYTHIA8 CUETP8M1 with MPI on (continuous line) and MPI off (dashed line). The predictions of PYTHIA8 with the Monash tune (dash-dot line) are also included to evaluate the contribution of the MPI independently of the specific parameters of the tune. The lower plot shows the ratios of the data-to-MC predictions for the different event generators.

10.3 Saturation effects

In Sec. 2.3.1 some of the most popular MPI models available in the modern MC event generators have been described, in particular the one implemented in PYTHIA8. In this model, the value of p_{T0} at which the taming of the $2 \rightarrow 2$ partonic cross-section takes place is one of the key free parameters. This parameter regulates both, the amount of MPI present in the event, as well as the production of soft particles by pQCD. The p_{T0} parameter is commonly included in most of the Minimum Bias and Underlying Event PYTHIA8 tunes.

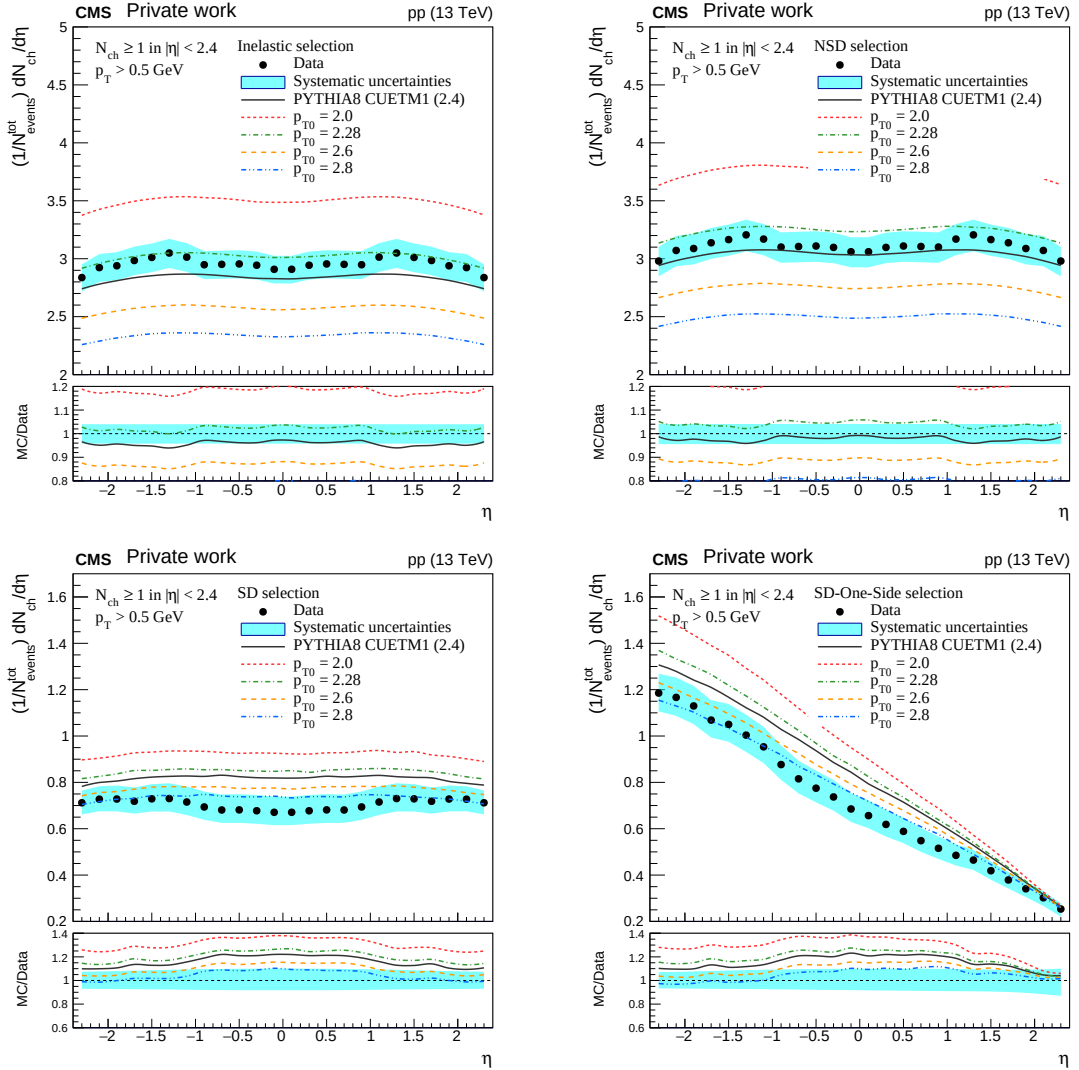


Figure 10.12: Charged-particle η densities averaged and symmetrized over the positive and negative η ranges. Corresponding to, from top to bottom, left to right: Inelastic-enhanced, NSD-enhanced, SD-enhanced and SD-One-Side enhanced event samples. The corrected data are compared to the predictions of the Monte Carlo event generator PYTHIA8 CUETP8M1 with different values of the p_{T0} parameter, 2.0, 2.28 (default of PYTHIA8), 2.4 (default of CUETP8M1 tune), 2.6 and 2.8. The lower plot shows the ratios of the data-to-MC predictions for the different p_{T0} values.

In this section the p_{T0} value is varied from 2.4 GeV (the value in the PYTHIA8 CUETP8M1 tune), down to 2.0 and up to 2.8 GeV, including the default PYTHIA8 value of 2.28 GeV.

In figure 10.12 the η distributions corresponding to all the event classes of this analysis are shown. In there it is possible to appreciate the significant impact on the amount of particles produced by the MPI. For the case of the most inclusive event selections, varying the p_{T0} parameter from 2.4 to 2.0 GeV enhances the amount of particles by $\sim 25\%$, while by increasing it to 2.8 GeV the amount of particles is depleted by $\sim 15\%$. This is the expected behaviour since the p_{T0} value regulates the amount of MPI by unitarizing the rapidly growing $2 \rightarrow 2$ partonic QCD cross-section at low- p_T . Hence, towards low- p_T values the $2 \rightarrow 2$ partonic QCD cross-sections increases rapidly allowing a larger MPI contribution. On the other hand, the cross-section decreases towards high- p_T making the MPI contribution smaller. This is precisely what is observed in figure 10.12. The effects on the SD-enhanced sample are less prominent with respect to the other two samples. The reason for this is the dependency of p_{T0} on the c.m.e. of the system, and since the c.m.e. of the Pomeron-proton system is lower than the c.m.e. of the proton-proton system, the range of the tested p_{T0} variations is above the transition region of the SD processes. In the case of the SD-One-Side enhanced event sample, the effects are more prominent on the side of the diffractively produced system, while the side of the scattered proton is less affected.

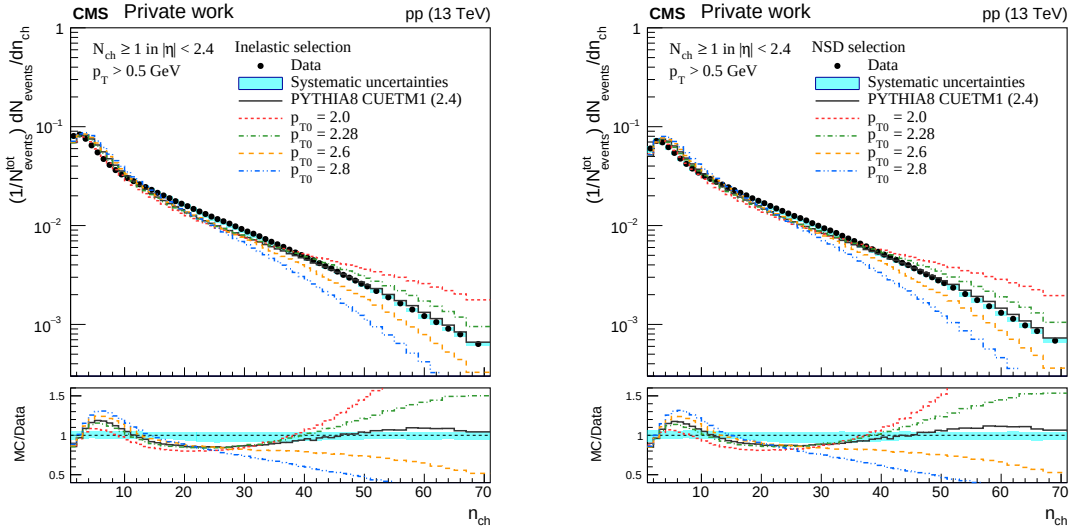


Figure 10.13: Charged-particle multiplicity distributions corresponding to, from left to right: Inelastic-enhanced and NSD-enhanced event samples. The corrected data are compared to the predictions of the Monte Carlo event generator PYTHIA8 CUETP8M1 with different values of the p_{T0} parameter, 2.0, 2.28 (default of PYTHIA8), 2.4 (default of CUETP8M1 tune), 2.6 and 2.8. The lower plot shows the ratios of the data-to-MC predictions for the different p_{T0} values.

Figure 10.13 presents the n_{ch} for the Inelastic-enhanced and NSD-enhanced event samples. It is precisely the high multiplicity tail, which is populated by the MPI (see figure 10.8), the most affected by the variations of the p_{T0} parameter.

Figure 10.14 shows the p_T distribution for charged-particles corresponding to the Inelastic-enhanced, NSD-enhanced and SD-enhanced event samples. For the most inclusive event samples, the effects of varying the p_{T0} parameter are mostly present at low- p_T values, while the large- p_T region remains almost without changes. In the case of the SD-enhanced event sample the effects are similar to those on the other two event samples for the corresponding p_T range.

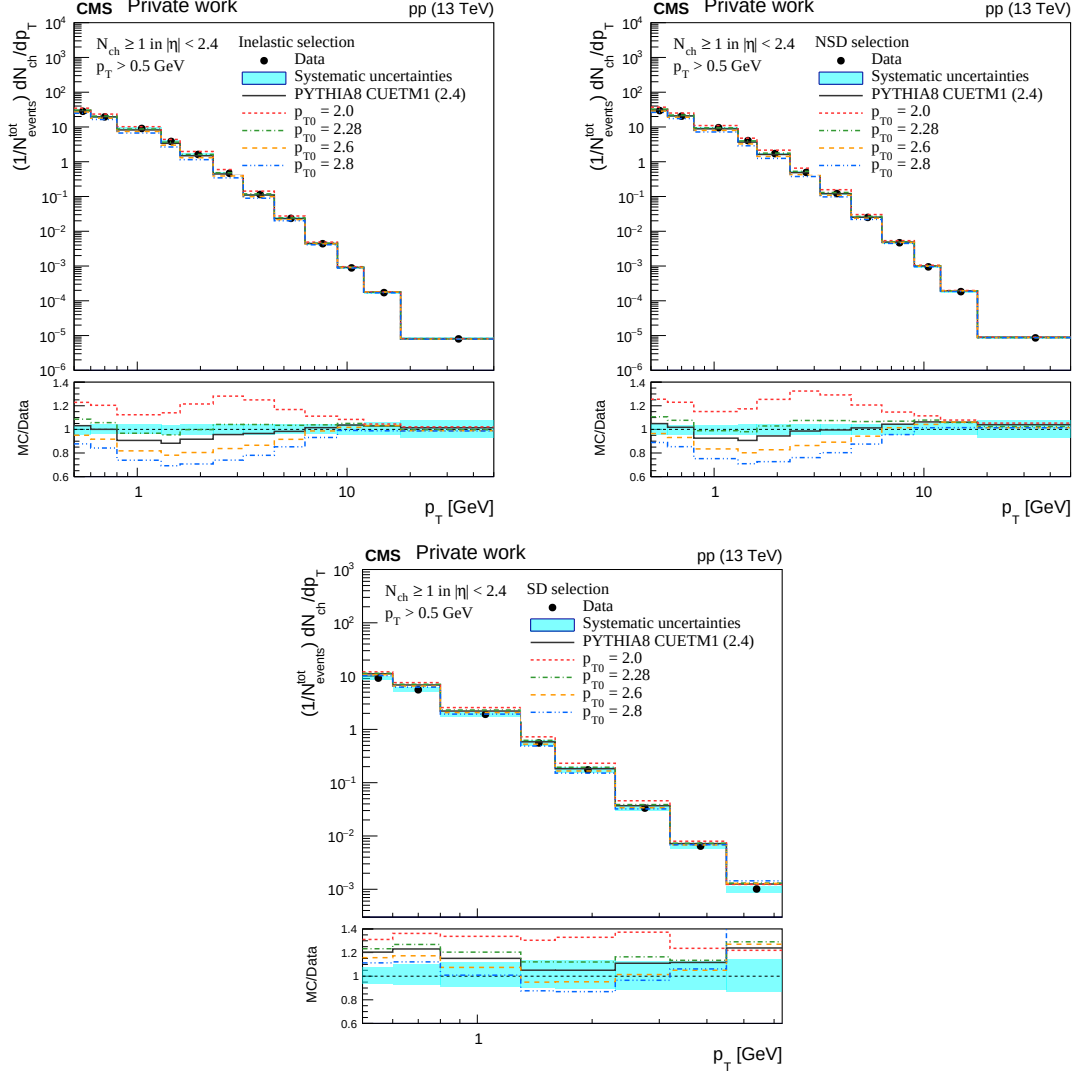


Figure 10.14: Charged-particle p_T densities corresponding to, from top to bottom, left to right: Inelastic-enhanced, NSD-enhanced and SD-enhanced event samples. The corrected data are compared to the predictions of the Monte Carlo event generator PYTHIA8 CUETP8M1 with different values of the p_{T0} parameter, 2.0, 2.28 (default of PYTHIA8), 2.4 (default of CUETP8M1 tune), 2.6 and 2.8. The lower plot shows the ratios of the data-to-MC predictions for the different p_{T0} values.

Figures 10.15 and 10.16 show the leading p_T charged-particle distribution and the integration of it as a function of $p_{T,\min}$. The distributions are presented for the same three event samples as before. It is precisely the middle p_T region, around the typical p_{T0} values, the region that presents the most prominent effects. This observable, now presented in different enhanced event samples, provides valuable information to further constrain this important parameter of the model. Still, it is important to notice that adjusting p_{T0} alone is not enough to properly describe these distributions.

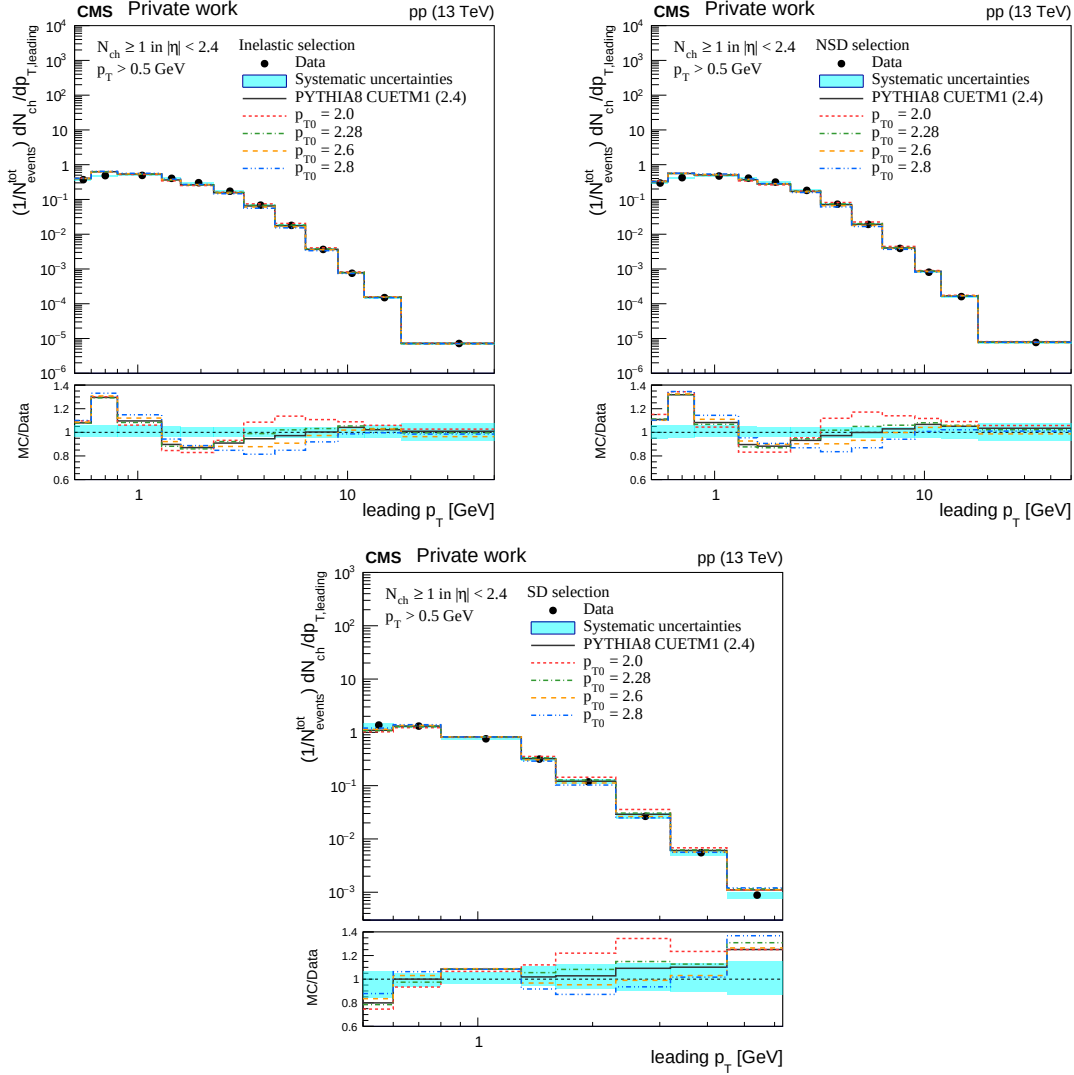


Figure 10.15: Leading charged-particle p_T densities corresponding to, from top to bottom, left to right: Inelastic-enhanced, NSD-enhanced and SD-enhanced event samples. The corrected data are compared to the predictions of the Monte Carlo event generator PYTHIA8 CUETP8M1 with different values of the p_{T0} parameter, 2.0, 2.28 (default of PYTHIA8), 2.4 (default of CUETP8M1 tune), 2.6 and 2.8. The lower plot shows the ratios of the data-to-MC predictions for the different p_{T0} values.

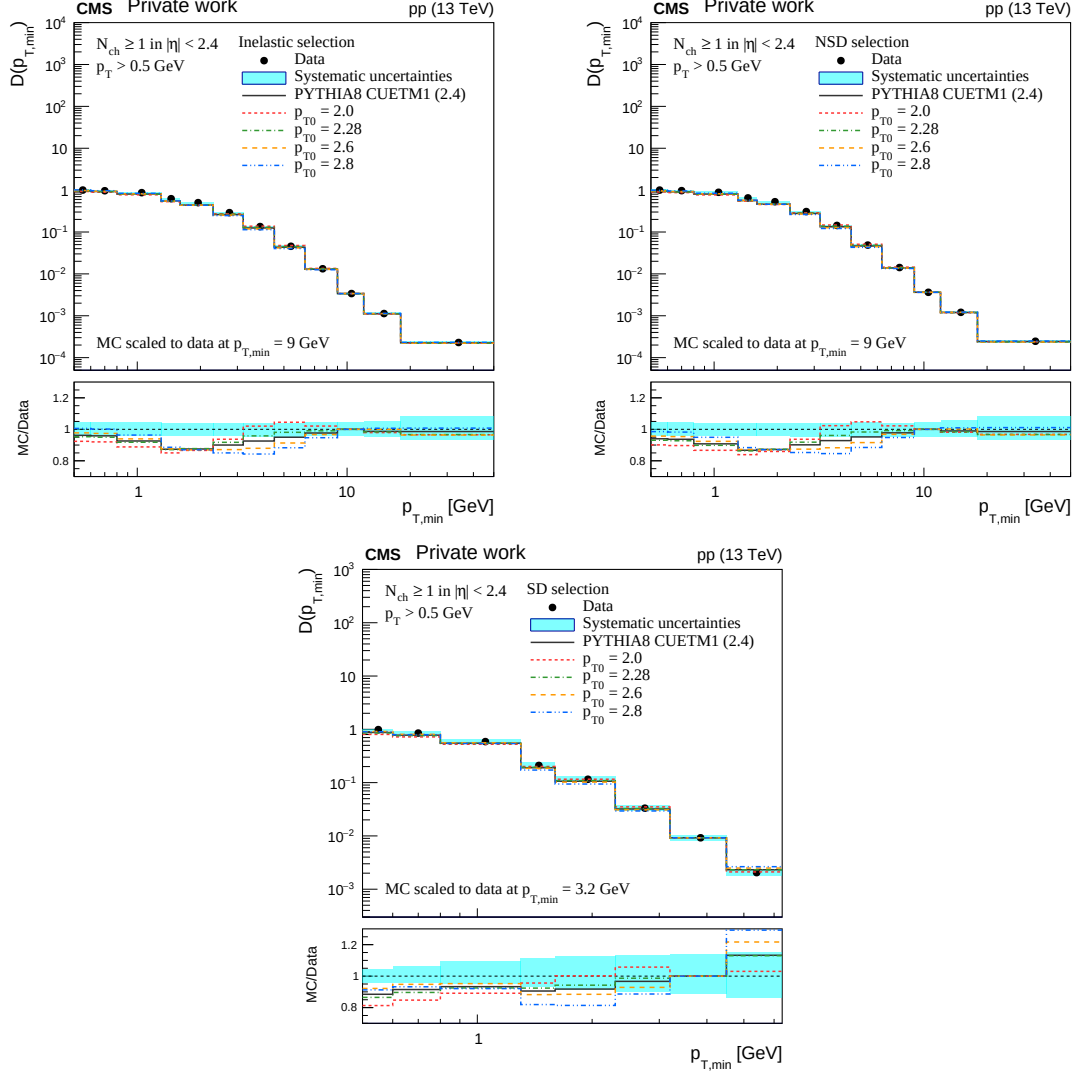


Figure 10.16: Integrated leading charged-particle p_T densities as a function of $p_{T,min}$ corresponding to, from top to bottom, left to right: Inelastic-enhanced, NSD-enhanced and SD-enhanced event samples. The corrected data are compared to the predictions of the Monte Carlo event generator PYTHIA8 CUETP8M1 with different values of the p_{T0} parameter, 2.0, 2.28 (default of PYTHIA8), 2.4 (default of CUETP8M1 tune), 2.6 and 2.8. The lower plot shows the ratios of the data-to-MC predictions for the different p_{T0} values.

10.4 Hard contribution to diffractive events

In PYTHIA8 the diffractively produced systems are classified either as low-mass or high-mass system. PYTHIA8 produces these diffractive masses through the parametrisation of a Pomeron flux that can generate masses down to 1.2 GeV. However, PYTHIA8 only allows the perturbative description of collisions with c.m.e. larger than 10 GeV. Hence, a simple model for low-mass systems is included to simulate the interactions between the Pomeron and the proton, as well as the usual perturbative treatment for high-mass systems.

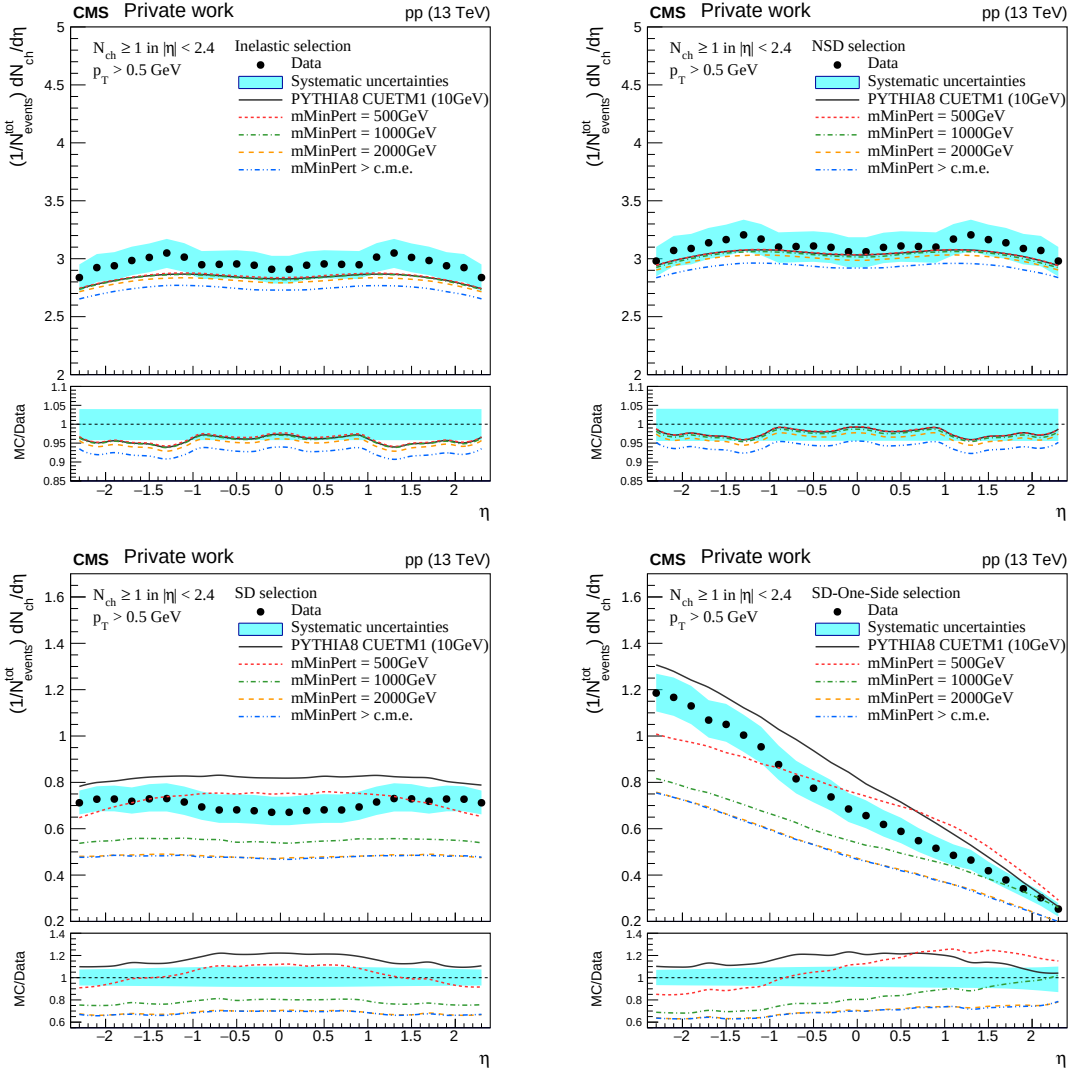


Figure 10.17: Charged-particle η densities averaged and symmetrized over the positive and negative η ranges. Corresponding to, from top to bottom, left to right: Inelastic-enhanced, NSD-enhanced, SD-enhanced and SD-One-Side enhanced event samples. The corrected data are compared to the predictions of the Monte Carlo event generator PYTHIA8 CUETP8M1 with different values of the mMinPert parameter, 10 GeV (CUETP8M1 and PYTHIA8 default), 500 GeV, 1000 GeV, 2000 GeV and mMinPert > c.m.e. The lower plot shows the ratios of the data-to-MC predictions for the different mMinPert values.

The transition between these two descriptions is regularised by a parameter which is set to the minimum mass of the diffractive system to be produced perturbatively (mMinPert). If, for instance, the value of mMinPert is set to a larger value than the c.m.e., then all the diffractive interactions are calculated by the non-perturbative model. In this section the mMinPert parameter is varied from the default value of 10 GeV to 500 GeV, 1000 GeV, 2000 GeV and a value larger than the c.m.e.

Figure 10.17 shows the η distribution of charged-particles for the Inelastic-enhanced, NSD-enhanced, SD-enhanced and SD-One-Side enhanced event samples compared to the predictions of PYTHIA8 CUETP8M1 with the aforementioned mMinPert variations. The Inelastic-enhanced and NSD-enhanced event samples are almost insensitive to the variations of mMinPert until it is set above the c.m.e. Only when diffraction is entirely simulated by the non-perturbative model included in PYTHIA8, a depletion of particles of $\sim 5\%$ is obtained. It is interesting to note, that the description of the SD-enhanced event sample has a more gradual change. For mMinPert values lower than the c.m.e. a depletion of particles on the side of the diffractive system is observed. The depletion is gradually extended towards the side of the scattered proton as the mMinPert value is increased. At the same time, the side of the scattered proton shows a small but constant enhancement of particles which suddenly drops as mMinPert increases. When mMinPert is larger than the 2000 GeV, a substantial decrease on the particle production across the full η range is observed, where the side of the diffractive system suffers a depletion of up to 40% and the side of the scattered proton of 20%.

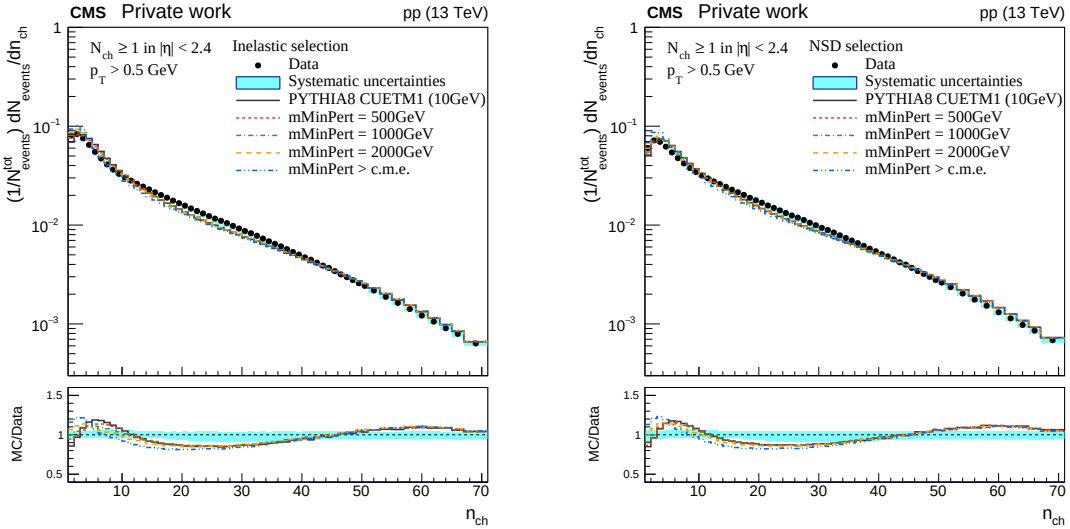


Figure 10.18: Charged-particle multiplicity distributions corresponding to, from left to right: Inelastic-enhanced and NSD-enhanced event samples. The corrected data are compared to the predictions of the Monte Carlo event generator PYTHIA8 CUETP8M1 with different values of the mMinPert parameter, 10 GeV (CUETP8M1 and PYTHIA8 default), 500 GeV, 1000 GeV, 2000 GeV and mMinPert > c.m.e. The lower plot shows the ratios of the data-to-MC predictions for the different mMinPert values.

Figure 10.18 shows n_{ch} for the Inelastic-enhanced and NSD-enhanced event samples. A migration of events from medium to low multiplicities is observed as $m\text{MinPert}$ is increased, causing a depletion for multiplicities between 5 and 30, and an enhancement of the peak at multiplicities below 5.

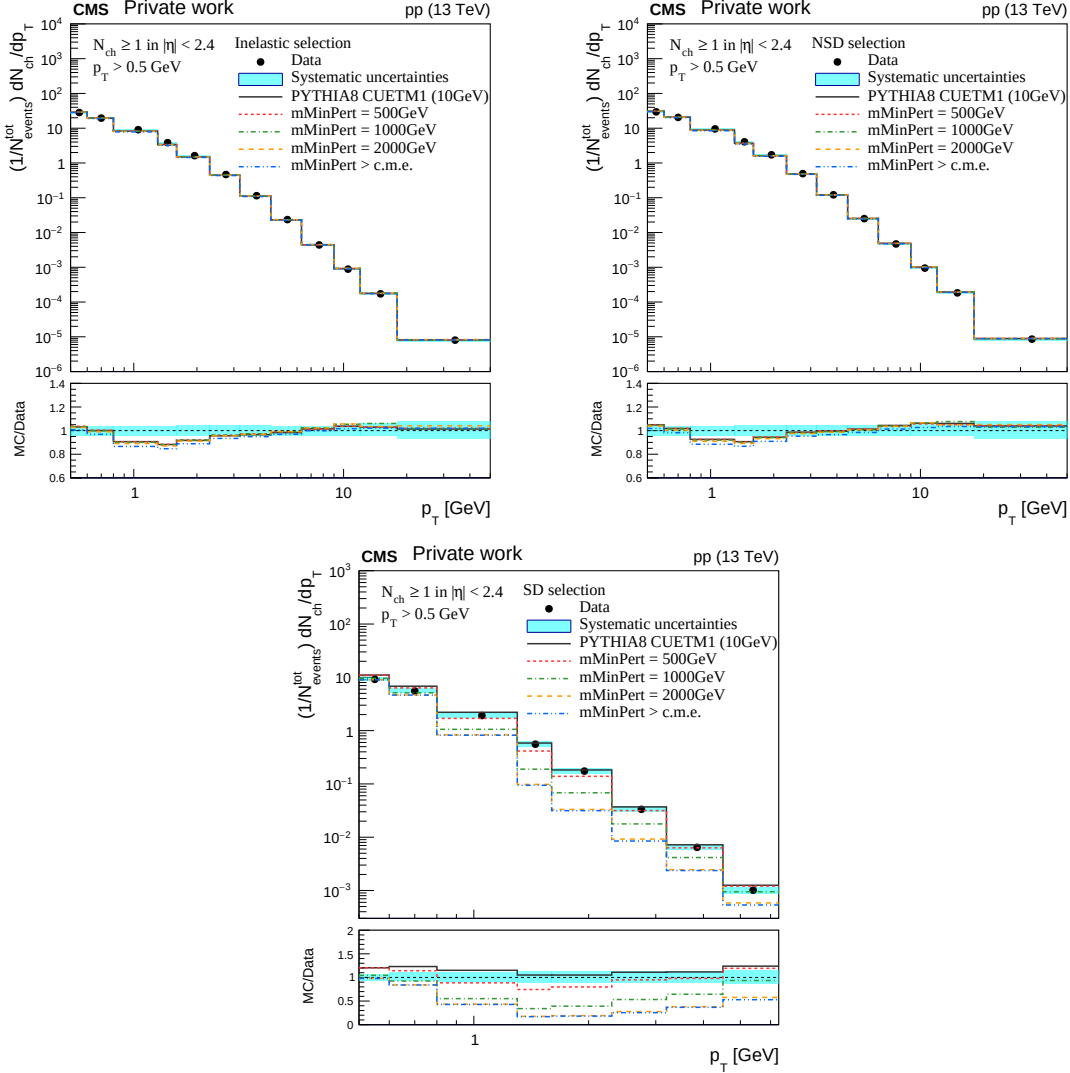


Figure 10.19: Charged-particle p_T densities corresponding to, from top to bottom, left to right: Inelastic-enhanced, NSD-enhanced and SD-enhanced event samples. The corrected data are compared to the predictions of the Monte Carlo event generator PYTHIA8 CUETP8M1 with different values of the $m\text{MinPert}$ parameter, 10 GeV (CUETP8M1 and PYTHIA8 default), 500 GeV, 1000 GeV, 2000 GeV and $m\text{MinPert} > \text{c.m.e.}$. The lower plot shows the ratios of the data-to-MC predictions for the different $m\text{MinPert}$ values.

Figure 10.19 shows the p_T of all charged-particles for the Inelastic-enhanced, NSD-enhanced and SD-enhanced event samples. The first two mentioned event samples are almost insensitive to the $m\text{MinPert}$ variations. It is only when $m\text{MinPert} > \text{c.m.e.}$ that a depletion on the particle density of $\sim 5\%$ for $p_T < 3$ GeV is visible. For larger p_T values no effect is visible.

On the other hand, the description of the SD-enhanced event sample presents a significant impact on the mMinPert value, including values below the c.m.e. All the variations show a larger deviation for middle p_T values (~ 2 GeV) than at the edges of the phase space. For the variations of 500 and 1000 GeV, the decrease on the particle density at mid- p_T is around 15 and 50%, respectively. When mMinPert is larger than 1000 GeV, a substantial depletion of particles is obtained, from 20 to 80% at low- p_T and mid/high- p_T , respectively.

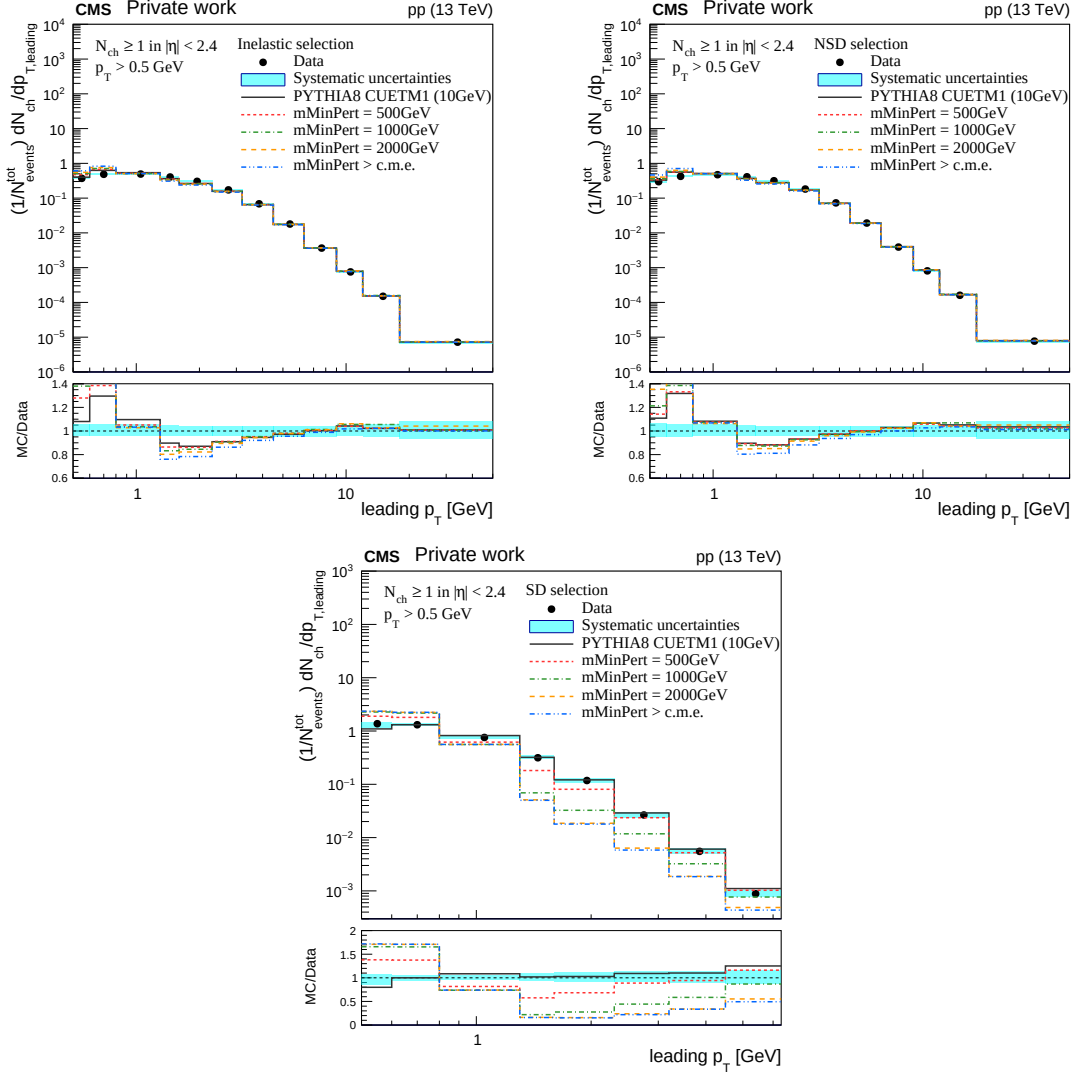


Figure 10.20: Leading charged-particle p_T densities corresponding to, from top to bottom, left to right: Inelastic-enhanced, NSD-enhanced and SD-enhanced event samples. The corrected data are compared to the predictions of the Monte Carlo event generator PYTHIA8 CUETP8M1 with different values of the mMinPert parameter, 10 GeV (CUETP8M1 and PYTHIA8 default), 500 GeV, 1000 GeV, 2000 GeV and mMinPert > c.m.e. The lower plot shows the ratios of the data-to-MC predictions for the different mMinPert values.

In figure 10.20 the leading p_T charged-particle distributions for the same three event samples are shown. All the event samples show the same trend as in figure 10.19 for $p_T > 1$ GeV, while for $p_T < 1$ GeV a systematic increase in the density of particles as the mMinPert value is observed.

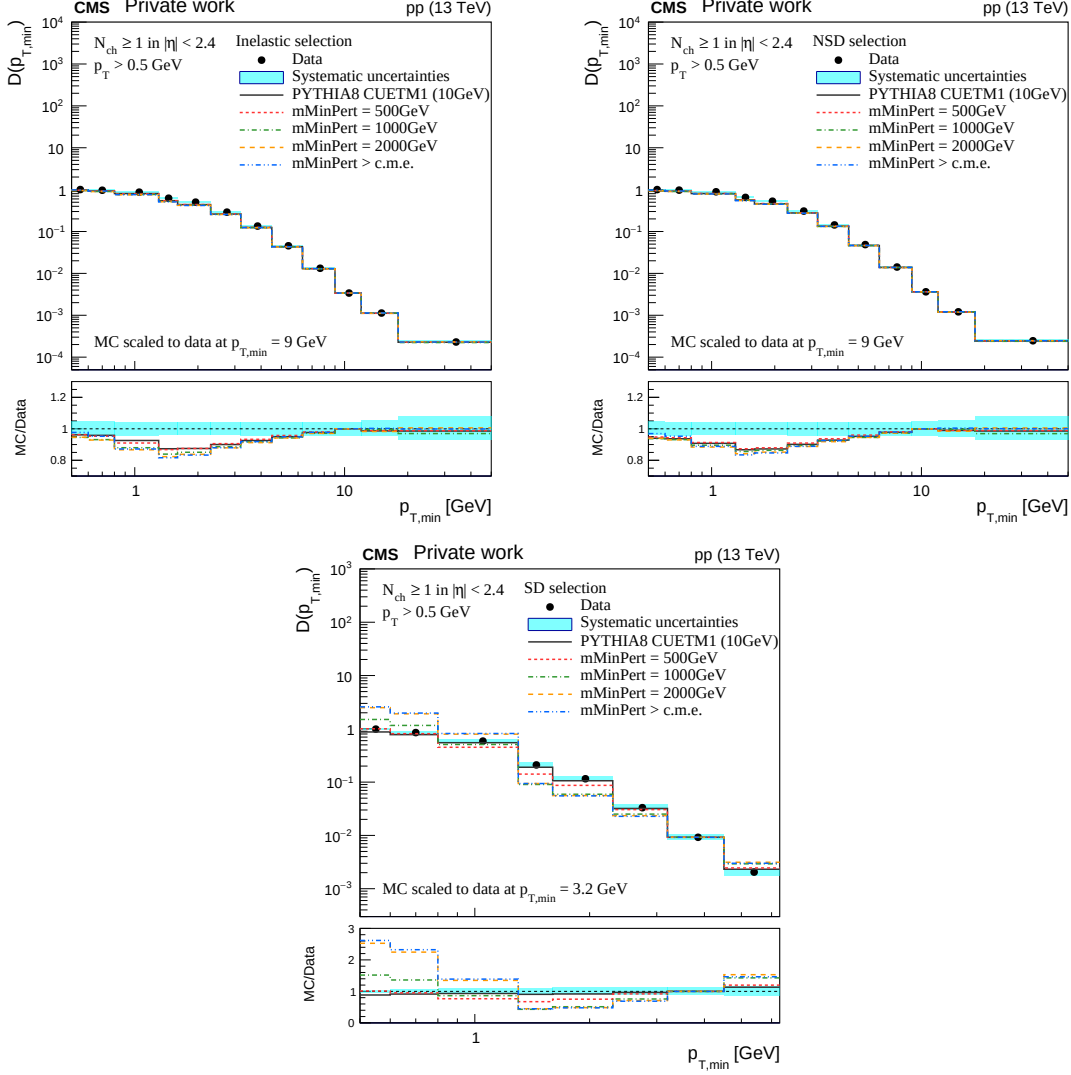


Figure 10.21: Integrated leading charged-particle p_T densities as a function of $p_{T,min}$ corresponding to, from top to bottom, left to right: Inelastic-enhanced, NSD-enhanced and SD-enhanced event samples. The corrected data are compared to the predictions of the Monte Carlo event generator PYTHIA8 CUETP8M1 with different values of the mMinPert parameter, 10 GeV (CUETP8M1 and PYTHIA8 default), 500 GeV, 1000 GeV, 2000 GeV and mMinPert > c.m.e. The lower plot shows the ratios of the data-to-MC predictions for the different mMinPert values.

All the mentioned variations on the description of the leading p_T charged-particle distributions are reflected in the integrated distributions shown in figure 10.21. In the cases of the Inelastic-enhanced and NSD-enhanced event samples, the predictions corresponding to the $m_{\text{MinPert}} > \text{c.m.e.}$ show the largest variation of $\sim 5\%$ difference for $0.8 < p_T < 2 \text{ GeV}$, the other m_{MinPert} variation being in between. The predictions for the SD-enhanced event sample corresponding to the m_{MinPert} variation are off by up to 260% at low- p_T values. For $p_{T,\text{min}}$ values between 0.8 and 2 GeV, the variation corresponding to 500 GeV underestimates the data by up to 40%, while all the other variations agree to each other underestimating the data by up to 60%.

CHAPTER

11

SUMMARY AND CONCLUSIONS

Proton-proton collisions involve multiple processes taking place at different energy scales, which makes their description a complex task. For instance, the main hard scattering occurs at a large-energy transfer between the interacting protons, whereas the bulk of charged-particles produced in the collision are characterised by small transverse momentum and arise from soft to semi-hard (multi)parton interactions characterised by low-energy transfers. The soft to semi-hard QCD regime is a rich source of interesting and important phenomena that unfortunately cannot be calculated from first principles, yet. Instead, a variety of QCD-inspired phenomenological models have been developed to simulate this energy regime and provide predictions. A drawback of these phenomenological models is that they include by construction multiple free parameters that have to be extracted from the comparison to experimental measurements. In addition, some of these parameters have intricate dependencies to the centre-of-mass energy, hence their extraction from measurements at different collision energies is of primary importance.

The understanding and proper simulation of the processes at low-energy transfers are essential to perform precise measurements in hadron colliders. For instance, the simulation of MPI, Initial and Final State Radiation (ISR/FSR), and hadronization, rely at some degree on phenomenological models. Furthermore, the beams provided by the LHC have such an intense luminosity that multiple proton-proton interactions within the same bunch crossing are expected, the so-called pileup (PU). Most of the PU interactions lay into the soft to semi-hard energy regime, and also include their own ISR, FSR and hadronization processes. Hence, even for studies where a specific hard interaction is experimentally requested, most of the activity in the detectors is coming from soft to semi-hard QCD processes originating from the PU interactions. The understanding of the processes at low-energy transfers is therefore of primary importance to be able to correct for the presence of PU.

Even though the contribution of the soft regime appears in most of the analyses as a relevant background, the understanding and study of these processes is also crucial by itself. They are intimately related to the internal structure of the proton, and to how the parton densities saturate at low-momentum fractions. At low x , the strong rise of the gluon density in the proton is expected to be tamed by recombination processes, leading to a saturation of the gluon density. The observation of this effect, which is not included in the DGLAP evolution equations, would shed light on the asymptotic high-energy behaviour of QCD. The study of these processes also provides information on how to connect the perturbative QCD regime to the phenomenological non-perturbative QCD regime.

In this thesis, different charged-particle distributions measured with the CMS detector in proton-proton collisions at a centre-of-mass energy $\sqrt{s} = 13 \text{ TeV}$ have been presented, namely the charged-particle pseudorapidity distributions, the multiplicity distributions of charged particles per event, and different transverse momenta distributions corresponding to all charged-particles, the leading charged-particle, and the integrated spectrum of the latter. The charged-particles are selected with transverse momenta $p_T > 0.5 \text{ GeV}$ in the range $|\eta| < 2.4$. The data are corrected for the different detector effects such as resolution and efficiency. The measured distributions are presented for four different samples selected according to the maximum particle energy in the range $3 < |\eta| < 5$. The samples correspond to an Inclusive event sample, an Inelastic-enhanced event sample, a sample dominated by non-single diffractive dissociation events (NSD-enhanced sample), and a sample enriched by single-diffractive dissociation events (SD-enhanced sample).

The measurements presented in this thesis provide extensive and unique insights into low-energy exchange processes that dominate the proton-proton interactions. The rich variety of distributions presented for different event samples, specially those enhanced in diffractive processes, are valuable sources of information to further constrain the models implemented in modern MC event generators. This helps to either improve the description of the models or to indicate where new formulations of certain components of the models are needed.

Diffractive in pp collisions is an active area of research with many complications to solve, e.g. the intricate coexistence of MPI and diffraction. The contribution of MPI tends to suppress the presence of diffractive processes, and the phenomenological description of this suppression is far to be clearly understood. The increase of MPI is directly related to the increase of soft gluon density inside the proton, which eventually saturates due to the presence of recombination effects. These saturation effects at some degree are modelled by the dampening of the $2 \rightarrow 2$ partonic cross-section at low- p_T values. In PYTHIA8 the region at which the dampening takes place is dictated by the p_{T0} parameter. This suppression is also intimately related with the allowed amount of MPI to be produced in a pp collision. In sections 10.2 and 10.3, these two effects are studied for each of the presented event samples. From these comparisons it is possible to conclude that the presence of MPI is extremely important to describe the saturation rate, but is not sufficient to accurately model it. Also, the p_{T0} parameter alone is not able to completely cure the observed discrepancy. These two observations suggest that the current MPI and regularisation models are not able to give a complete description of the transition region. Furthermore, many of the different components in a proton-proton collision present intricate dependencies on one another. Therefore, the

observed effects cannot be uniquely associated to a single parameter. Additional investigation of further components of the models, such as the parametrisation of the hadronic matter distribution, color reconnection effects and soft diffraction models, could help to improve the understanding of this region.

It is also interesting to notice that having two hadrons in the initial state opens the possibility to have MPI contributions within the diffractive Pomeron-proton system. Effects of these processes are observed in the comparison of the results corresponding to the SD-enhanced event sample to the MC predictions with and without MPI. In this event selection, the MPI contribution coming from the proton-proton system is highly suppressed by the requirement of low particle activity on one forward side of the detector. Hence, the observed depletion of particles by switching off the MPI is mainly coming from the decrease of the MPI activity in the Pomeron-proton system. Is particularly interesting to observe these effects on the SD-One-Side enhanced event selection, where the activity coming from the diffractive system is directly measured. The difference of MPI activities in the proton-proton and Pomeron-proton systems can be quantified from the comparison of the Inclusive-enhanced and NSD-enhanced event samples to the SD-enhanced event sample. In the first two event samples, switching off the MPI reduces the particle density by 2/3 while in the SD-enhanced event sample the depletion is of 1/5. The study of these effects also gives access to the parton densities of the Pomeron, also known as diffractive-PDFs.

The presented studies on the saturation rate by varying the p_{T0} parameter (Sec. 10.3), in combination with the studies on the contribution of the soft- and hard-diffractive components (Sec. 10.4), provide useful information related to the typical scales of the different diffractive processes, as well as on the interplay between their soft and hard components. Furthermore, the dependency of the p_{T0} parameter on the centre-of-mass energy of the interaction is visible from the comparison of leading- p_T distributions corresponding to the most inclusive selections to the selection enhanced in SD dissociation. This effect comes from the reduced c.m.e. of the Pomeron-proton system present in SD dissociation events with respect to the c.m.e. of the proton-proton system. From these studies it is possible to extract that the main diffractive processes in the Inelastic-enhanced and NSD-enhanced event samples are primarily described by soft-diffraction. On the other hand, the SD-enhanced event sample presents an equal contribution from the soft- and hard-diffractive descriptions, for which the former dominates at scales below 500–1000 GeV, and the latter above those same scales.

Within the CMS collaboration, the Monte Carlo event generator working group is the responsible to provide more accurate set of parameters for the MC models by simultaneously fitting the predictions to a large variety of experimental data. The different distributions presented in this thesis are planned to be used for these purposes in the next release of MC tunes.

The categorisation of events in samples enhanced in different diffractive processes could be fruitfully applied to the study of the Underlying Event. In these measurements, different contributions of the particle production are separated according to the azimuthal separation $\Delta\varphi$ with respect to the leading p_T object in the event [83–85]. The measured particle activity in the azimuthal regions $60 < |\Delta\varphi| < 120$ enables to separate the contributions from the MPI from those of the ISR and FSR. It would be interesting to study these individual

contributions in the different event samples considered in this thesis.

Appendices

APPENDIX

A

TRACKER ALIGNMENT

As part of the CMS collaboration every member is required to provide certain time to develop and operate the experiment. In my particular case I have worked in the software based alignment of the silicon tracking detector of the CMS experiment. This appendix is dedicated to explain the basic idea behind the alignment procedure, as well as to present the main areas where I have contributed.

A.1 Track-based alignment

In order to exploit the capabilities of the silicon tracking detector installed in the CMS experiment, the determination of the set of parameters describing the geometrical properties of the silicon modules is needed. These parameters are commonly referred to as *alignables* and all together define what is known as the *tracker geometry*. In the case of the CMS tracker detector the number of parameters that have to be determined goes up to 200,000. These parameters correspond to the spacial position of each of the individual modules and their deformations (curvatures), and the relative position of the structures holding the modules, while taking into account things like radiation damage and not functioning modules, among others.

Figure A.1 shows the relative coordinate system of each individual module, as well as the incident angles with which the tracks hit the module. Figure A.2 shows a schematic representation of one quarter of the CMS tracker in the z - r plane. In this figure the different subdetectors of the tracking detector are represented. From inside out, first the subdetectors composed by pixel modules are found, namely the Barrel Pixel (BPIX) and the Forward Pixel (FPIX). Immediately after and surrounding the BPIX and FPIX the subdetectors formed by strip modules are located. These are the Tracker Outer Barrel (TOB), the Tracker End Cap (TEC), the Tracker Inner Barrel (TIB) and the Tracker Inner Disks (TID).

Each of the mentioned subdetectors is further divided in different mechanical structures. For instance, the barrel subdetectors are subdivided into two half-shells which are further (sub)divided into ladders, in which the individual modules are mounted.

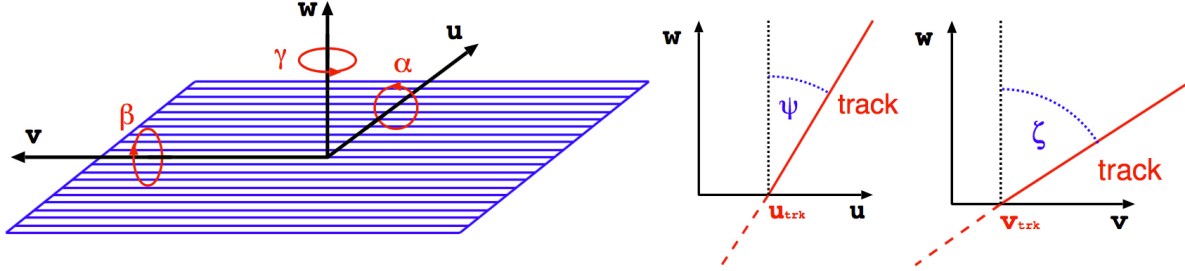


Figure A.1: Sketch of a silicon strip module showing the axes of its local coordinate system, u , v , and w , and the respective local rotations α , β , γ (left), together with illustrations of the local track angles ψ and ζ (right). Figure taken from [86].

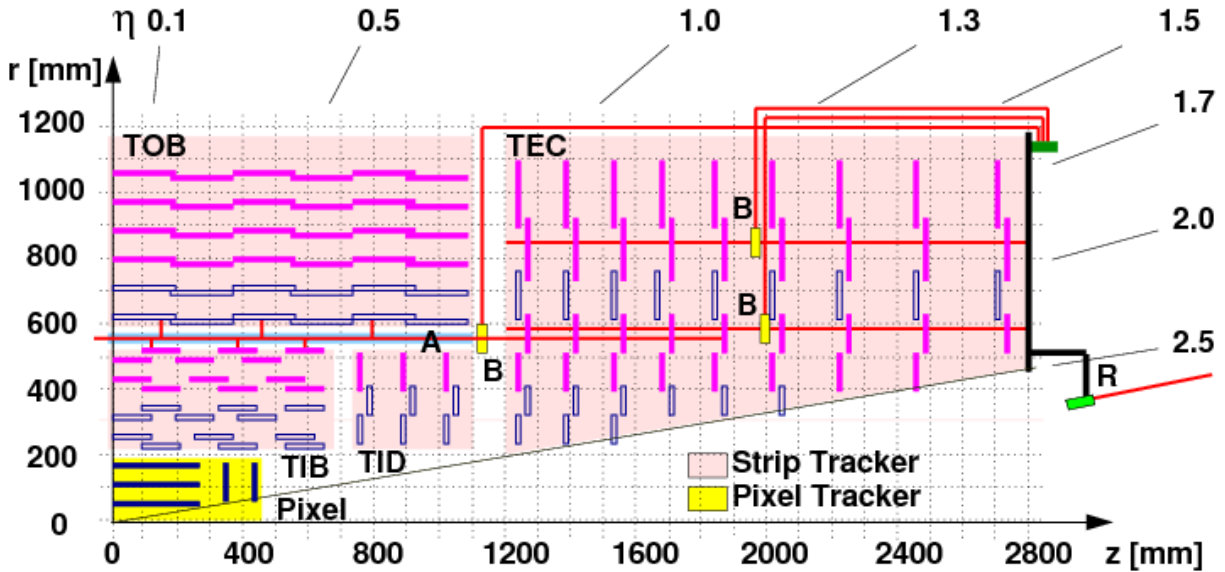


Figure A.2: Schematic view of one quarter of the silicon tracker in the r - z plane. The positions of the pixel modules are indicated within the hatched area. At larger radii within the lightly shaded areas, solid rectangles represent single strip modules, while hollow rectangles indicate pairs of strip modules mounted back-to-back with a relative stereo angle. The figure also illustrates the paths of the laser rays (R), the alignment tubes (A) and the beam splitters (B) of the laser alignment system. Figure taken from [86].

Due to the limited accessibility of the tracker inside the CMS detector and the required level of precision, the alignment is based on the tracks reconstructed by the tracker in situ. The alignment procedure is divided into different accuracy levels; the least precise accuracy level corresponds to the determination of the positions and rotations of the subdetectors,

also known as high-level structures (BPIX, FPIX, TOB, etc.). The next level of accuracy corresponds to the determination of more precise alignables corresponding to the position and rotations of the mechanical substructures of the subdetectors (half-shells, ladders, etc.). Finally, the most precise alignment accuracy is reached when the alignables corresponding to the individual modules are determined. At each level it is possible to keep fixed any arbitrary number of alignables in order to focus in those of interest. This feature is particularly important since the ability to determine the alignment parameters depends on the available type of track samples, as well as on their statistical size. The type of track samples used for the alignment procedure include tracks from: isolated muons, minimum bias events, muons from Z-boson decays and cosmic ray events.

Each of these samples provides different constraints that help to cure misalignment effects which could not be fixed otherwise. Among these misalignment scenarios the so-called *weak modes* are specially difficult to treat. The weak modes are deformations in the tracker structures that provide equally good fitted tracks independently of the degree of the deformation, but still affect the measured physical outcome. For instance, the so called *twist* weak mode, corresponding to the coherent movement of the modules in the azimuthal φ direction by an amount proportional to their longitudinal position ($\Delta\varphi = \tau \cdot z$), changes the measured transverse momentum of the tracks. This effect can be cured by using a sample in which the tracks originate from the decay of a known resonance, constraining the expected reconstructed mass and hence the reconstructed transverse momentum of the tracks.

A.1.1 MillePede

The track-based alignment of the silicon tracker detector in the CMS experiment is performed with the MillePede-II software [87] (MPII). MPII applies the least squares method to simultaneously minimise the residuals of all the positions of the measured track hits with respect to the predicted ones. The χ^2 implemented in MPII is shown in Equation A.1.

$$\chi^2(\mathbf{p}, \mathbf{q}) = \sum_j^{\text{tracks}} \sum_i^{\text{measurements}} \left(\frac{m_{ij} - f_{ij}(\mathbf{p}, \mathbf{q}_j)}{\sigma_{ij}} \right)^2 \quad (\text{A.1})$$

where $\mathbf{p}$ are the alignment parameters (alignables), $\mathbf{q}$ are the track parameters, m are the measured hit positions and f are the predicted hit positions. σ are the corresponding uncertainties. In order to provide the initial trajectory predictions, it is necessary to assume an initial tracker geometry. Therefore, the obtained alignment parameters are corrections to that initial assumption.

MPII uses the informations provided by millions of tracks to obtain the actual geometry parameters. By taking into account the information of such a large number of tracks, the system of equations becomes too large for the computational capabilities available nowadays. However, MPII solves this huge system of equations by using a property present on certain least square problems where the involved parameters can be divided into *local* and *global*. Local parameters are those only related with a subset of the data, while global parameters are those affecting the global behaviour. When this property is present, the matrices of the

problem can be subdivided into smaller matrices in order to focus only on the solution of the global parameters. In the case of the alignment of the tracker, such conditions are fulfilled, the global parameters are those related with the actual geometry of the tracking detector, and the local parameters correspond to the individual track parameters. The interest is to determine the global parameters related with the properties of the tracking detector, i.e. the position, rotation and curvature of all the modules and the position and rotations of the mechanical structures.

A.2 Personal contributions

During my service work for the CMS collaboration, I have actively participated in the different tracker alignment projects listed in the following:

- First alignment of the CMS Run 2 tracking detector.

During the first long shutdown (LS1) of the LHC accelerator from 2013 to the end of 2014, the CMS experiment received extensive maintenance work and upgrades. The tracker detector was specially manipulated, as for instance, multiple modules of the pixel barrel were replaced. This led to new unknown misaligned geometry of the pixel modules and the structures they are mounted on. Therefore, in order to provide an acceptable tracking performance from day one of the LHC Run 2 program, a tracker alignment campaign using MC simulations and cosmic ray data was performed.

In order to achieve this, first a realistic misaligned geometry had to be created. The information on the mechanical assembly accuracy of the tracker modules, in combination with the known precision for mounting the tracker inside the CMS experiment are used. As a second step, an alignment strategy is created. An alignment strategy consists in choosing which components of the tracker have to be aligned and the estimate of the minimum amount of data needed. By using different MC samples of cosmic rays and MB events, it is possible to test different combinations in order to find a realistic and optimal solution. This step gives estimates on the amount of cosmic ray data to collect prior to the start of the collisions, as well as on the amount of MB events needed to achieve the nominal tracking performance.

Using the first cosmic ray data collected with the magnetic field of the CMS experiment turned off, the so-called Cosmic Run at Zero Tesla (CRUZET), the alignment of the manipulated structures during the LS1 has been obtained. The determination of the alignment parameters was further improved using the cosmic ray data collected at the Cosmic Run At Four Tesla (CRAFT) in addition to the CRUZET data.

- Phase I Pixel detector

During the extended end-of-year shutdown 2016-2017, the pixel detector received an additional barrel layer and one additional end-cap disk. Prior to their installation extensive work was done within the tracker alignment group of CMS. During this time, the consolidation of the alignment workflow to operate with the new pixel geometry took place. Among the different topics to investigate, I have contributed in the creation of realistic start-up scenarios of the new barrel pixel, which have been used to test

different alignment strategies as well as to quantify the impact on the tracking efficiency of multiples scenarios.

APPENDIX

B

ADDITIONAL CONTROL PLOTS

B.1 Vertex

The distributions of the uncertainty σ_z on the z coordinate, of the χ^2 , number of degrees of freedom and reduced χ^2 associated to the primary vertex reconstruction are shown in figure B.1 for events with exactly one reconstructed primary vertex.

The distributions of the coordinates x and y of the reconstructed primary vertex are shown in figure B.2 for events with exactly one reconstructed primary vertex, together with the position of the vertex in the transverse plane, ρ . The distributions are shown with respect to the origin of the coordinate systems used in the reconstruction of the data and the simulation, and with respect to the position of the beam spot. The dependence of the uncertainty σ_z on the position z and on the track multiplicity is shown in figure B.3.

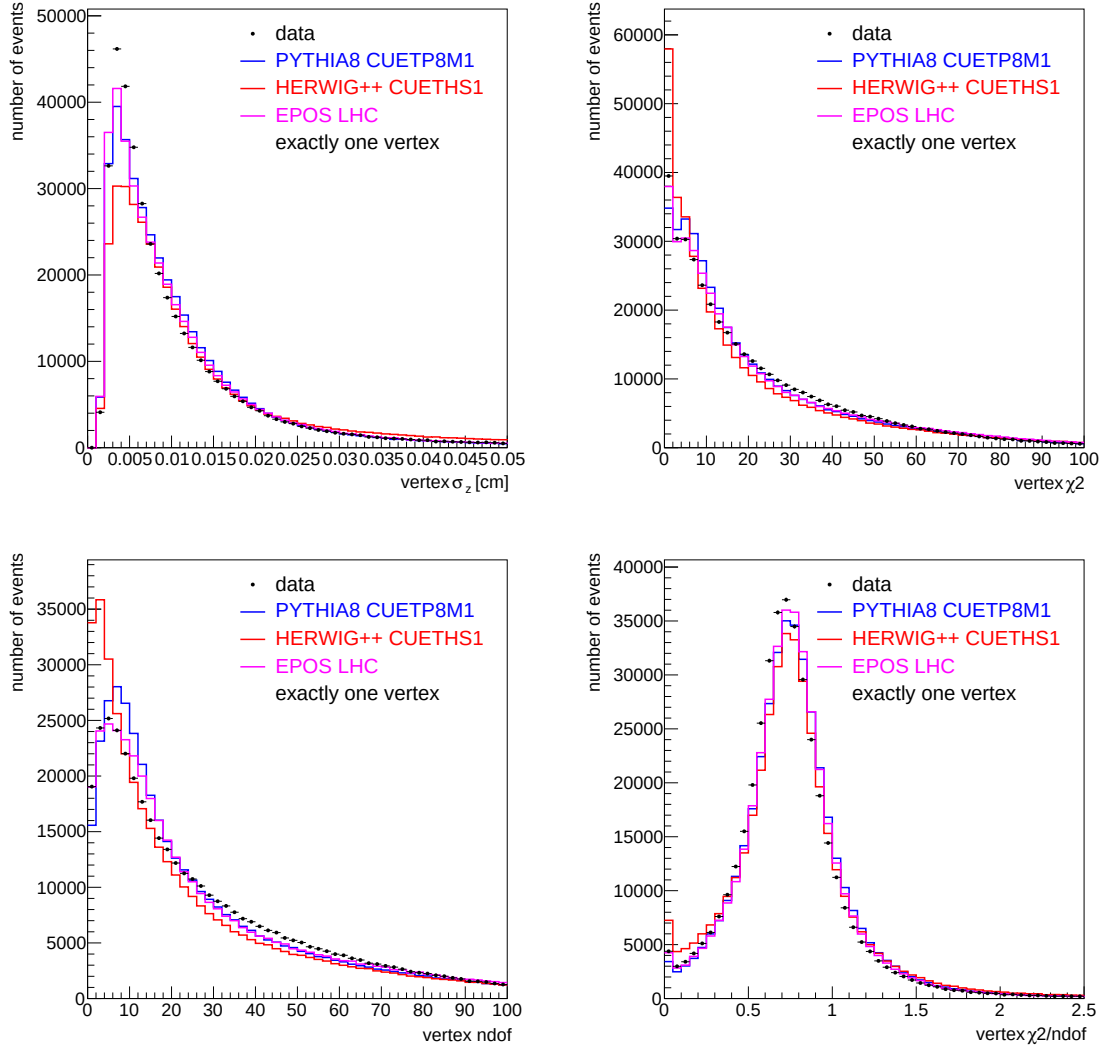


Figure B.1: Distributions of the uncertainty σ_z on the z coordinate, of the χ^2 , number of degrees of freedom and reduced χ^2 associated to the primary vertex reconstruction for events with exactly one reconstructed primary vertex. The data are compared to the predictions of the Monte Carlo event generators PYTHIA8, HERWIG++ and EPOS.

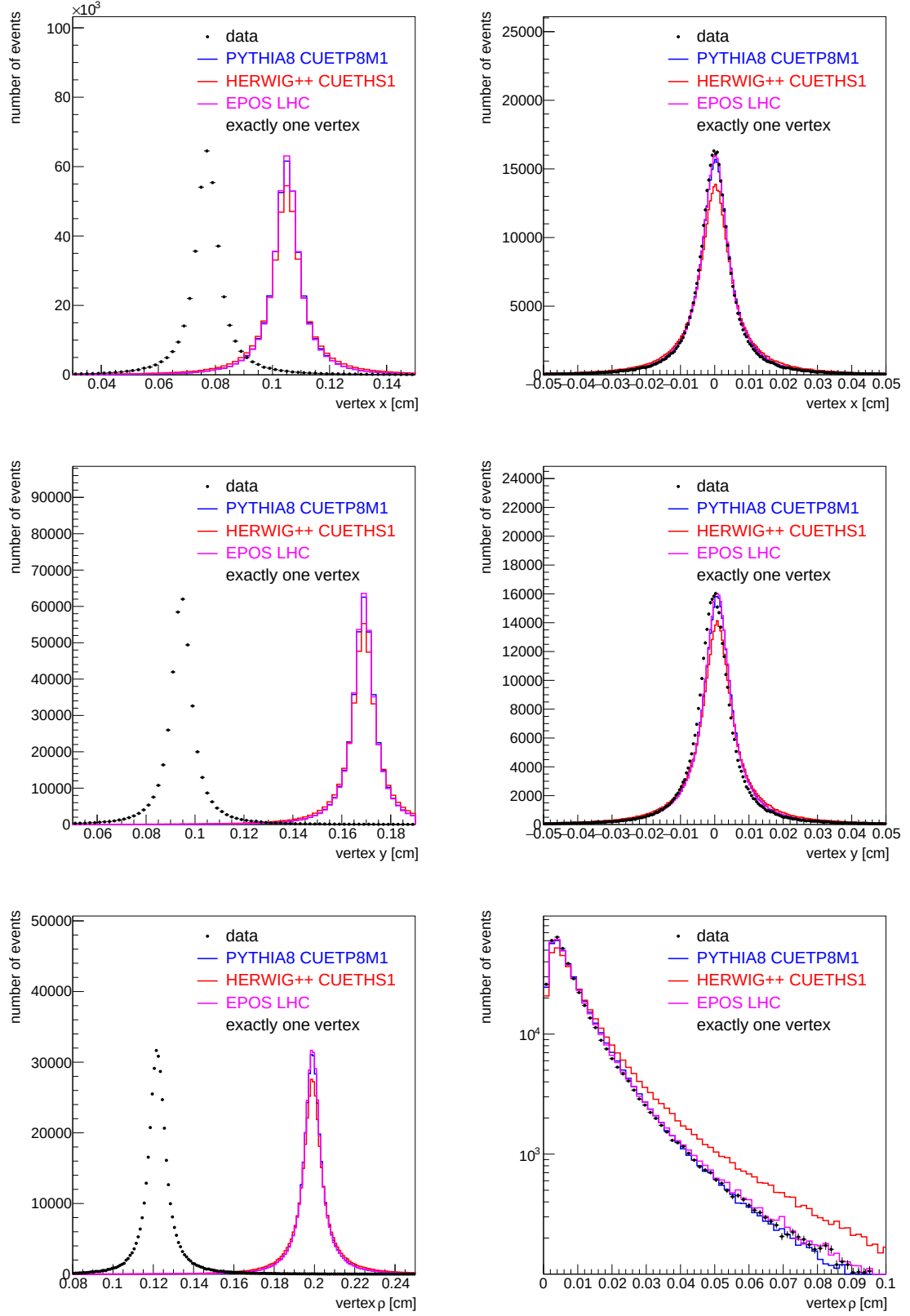


Figure B.2: Distributions of the coordinates x and y of the reconstructed primary vertex and position ρ of the vertex in the transverse plane for events with exactly one reconstructed primary vertex. The distributions are shown with respect to the origin of the coordinate systems used in the reconstruction of the data and the simulation (left), and with respect to the position of the beam spot (right). The data are compared to the predictions of the Monte Carlo event generators PYTHIA8, HERWIG++ and EPOS.

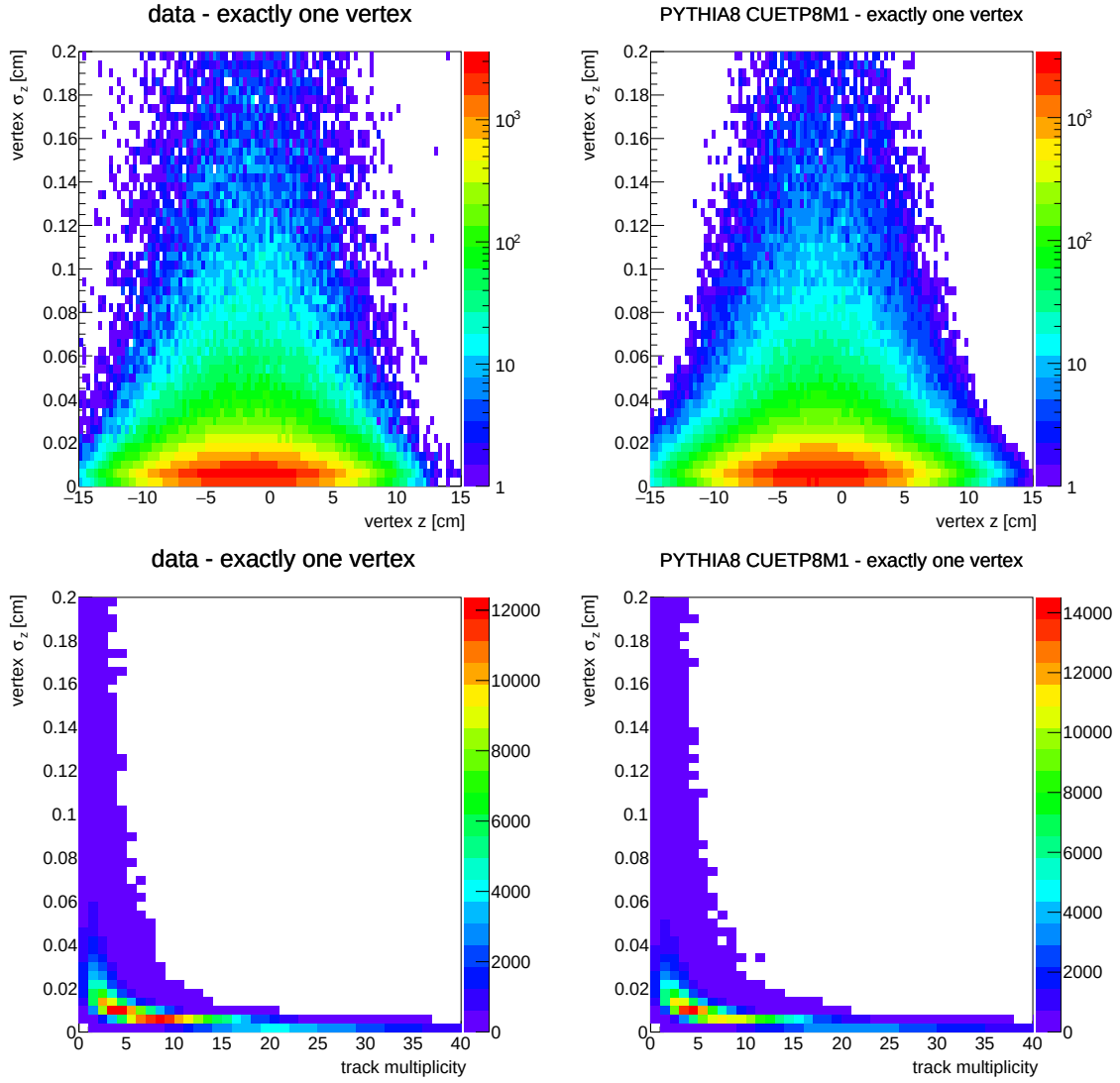


Figure B.3: Uncertainty on the z coordinate of the reconstructed primary vertex as a function of the position z and the track multiplicity for events with exactly one reconstructed primary vertex. The data are compared to the predictions of the Monte Carlo event generator PYTHIA8.

B.2 Tracks

Figure B.4 shows the distributions of the azimuthal angle of the tracks in the central ($|\eta| < 1$) and forward ($|\eta| > 1$) regions of the tracker acceptance, and figure B.5 the two dimensional distributions of the azimuthal angle and pseudorapidity of the tracks before and after the requirement. The ratios between the two dimensional $\eta - \varphi$ distribution after the requirement and the distribution before the requirement are shown in figure B.6. The distributions of the transverse momentum of the tracks before and after the requirement are shown in figure B.7. The level of agreement between the data and the simulation remains unchanged after the requirement to have at least 3 pixel hits. The ratios between the transverse momentum distribution after the requirement and the distribution before the requirement are shown in figure B.8 for the tracks (left) and the matched charged particles (right). The suppression of tracks by the condition on the number of pixel hits follows the same behaviour in the data and the simulation, with a slightly lower rejection at low transverse momentum.

The dependence of the relative transverse momentum resolution on the track transverse momentum, pseudorapidity and number of associated hits is shown in figure B.9. The significance of the impact parameter with respect to the vertex position in the transverse plane, d_{xy}/σ_{xy} , and the significance of the point of closest approach to the primary vertex along the beam line, d_z/σ_z were presented in figure 6.15. Their dependence on the track pseudorapidity and number of associated hits is shown in figure B.10.

Figures B.11 and B.12 show the dependence of the significances d_{xy}/σ_{xy} and d_z/σ_z on the track pseudorapidity and multiplicity in a two dimensional view. The distributions of the χ^2 , number of degrees of freedom, reduced χ^2 and weights associated to the track reconstruction are presented in figure B.13.

Figure B.14 shows eventually the dependence of the hit multiplicity on the pseudorapidity of the associated track, for all the tracker layers and for the pixel and strip layers separately.

The distributions of the hits from the tracks associated to the primary vertex have also been studied. The collection of hits being not available in the RECO data format used for the analysis, the track reconstruction algorithm has to be rerun in order to access to the properties of the hits. Figures B.15 and B.16 show the distributions of the x, y coordinates of the hits from the tracks associated to the primary vertex in the pixel barrel (PXB), pixel forward (PXF), tracker inner disk (TID), tracker inner barrel (TIB), tracker outer barrel (TOB) and tracker endcap (TEC). The data are compared to the predictions of the Monte Carlo event generator PYTHIA8. The distributions of the pseudorapidity and azimuthal angle of the hits from the tracks associated to the primary vertex are presented in figures B.17 to B.22 for the different components of the tracking detector, before and after the requirement to have at least 3 pixel hits. The data are compared to the predictions of the Monte Carlo event generator PYTHIA8.

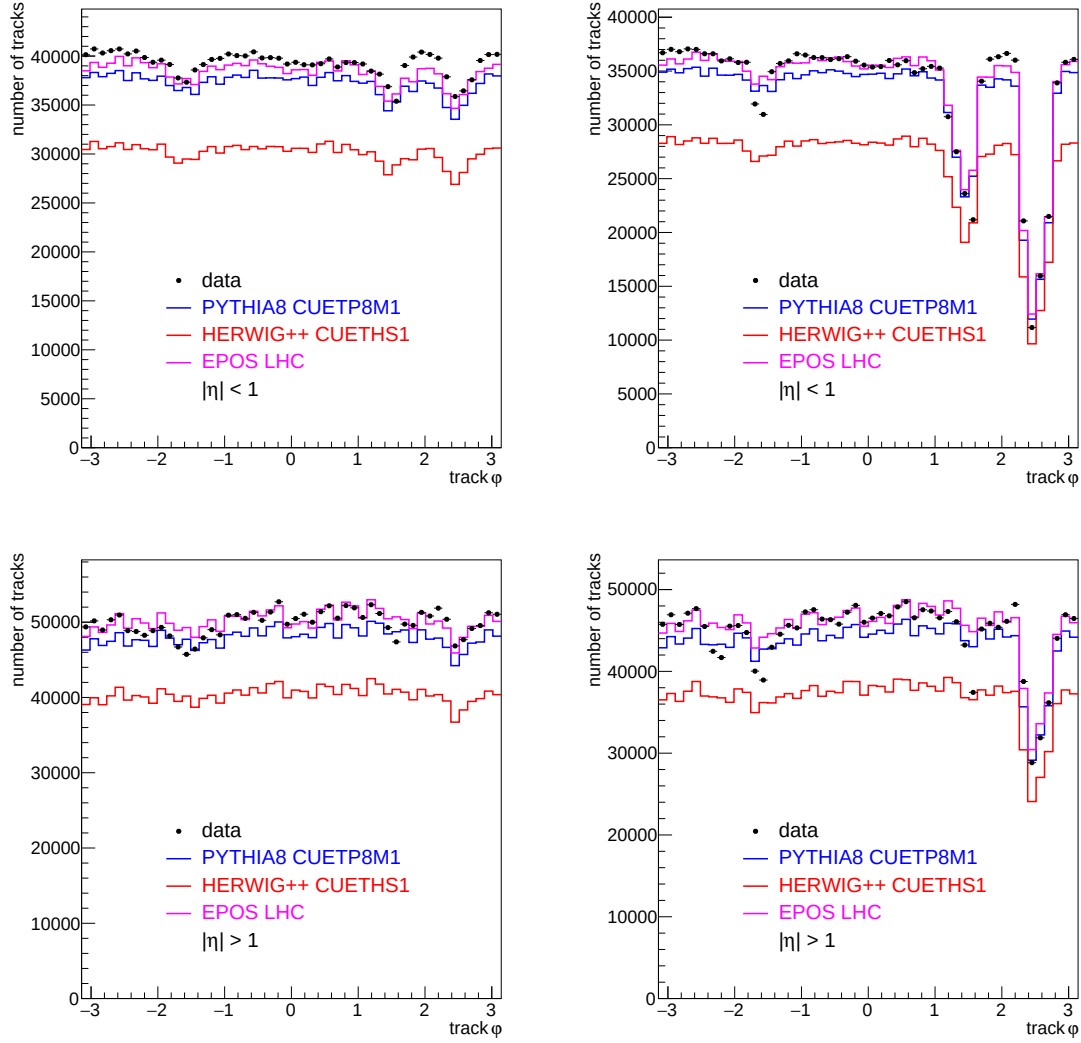


Figure B.4: Distributions of the azimuthal angle of the tracks associated to the primary vertex in the central ($|\eta| < 1$) and forward ($|\eta| > 1$) regions of the tracker acceptance, before (left) and after (right) the requirement to have at least 3 pixel hits. The data are compared to the predictions of the Monte Carlo event generators PYTHIA8, HERWIG++ and EPOS.

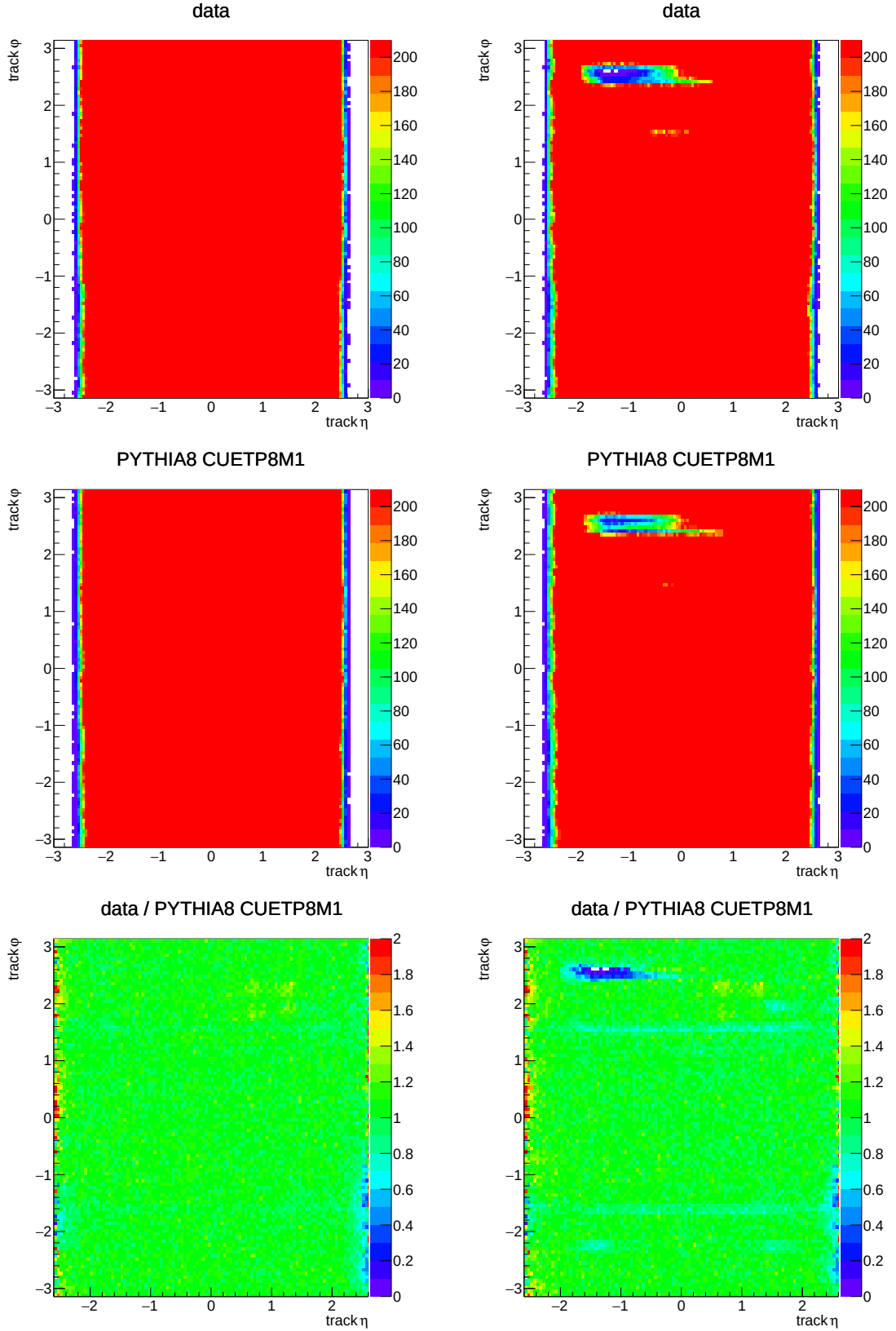


Figure B.5: Two dimensional distributions of the azimuthal angle and pseudorapidity of the tracks associated to the primary vertex before (left) and after (right) the requirement to have at least 3 pixel hits. The data are compared to the predictions of the Monte Carlo event generator PYTHIA8.

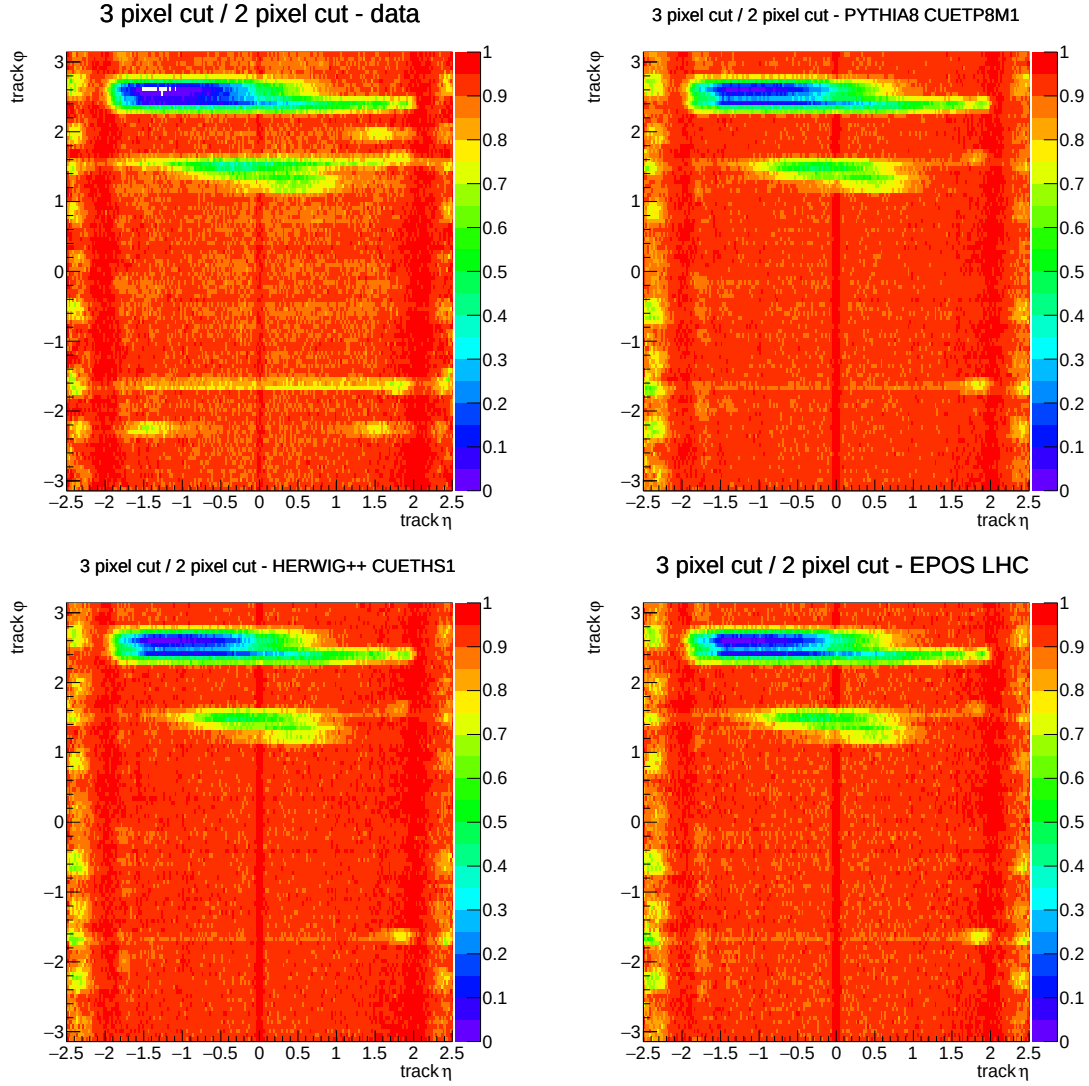


Figure B.6: Ratios between the two dimensional $\eta - \varphi$ distribution of the tracks after the requirement to have at least 3 pixel hits and the distribution before the requirement. The data are compared to the predictions of the Monte Carlo event generators PYTHIA8, HERWIG++ and EPOS.

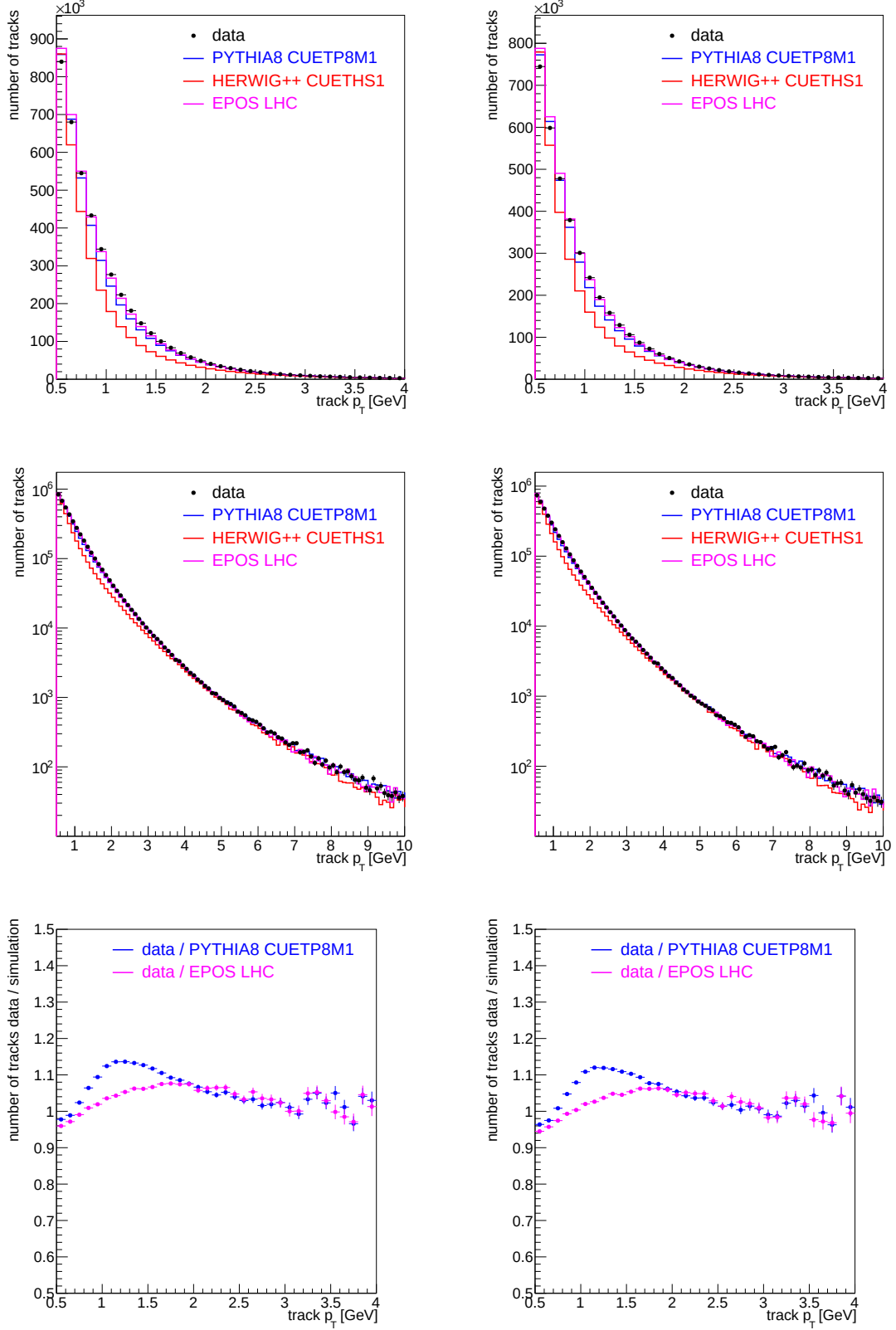


Figure B.7: Distributions of the transverse momentum of the tracks associated to the primary vertex before (left) and after (right) the requirement to have at least 3 pixel hits. The data are compared to the predictions of the Monte Carlo event generators PYTHIA8, HERWIG++ and EPOS.

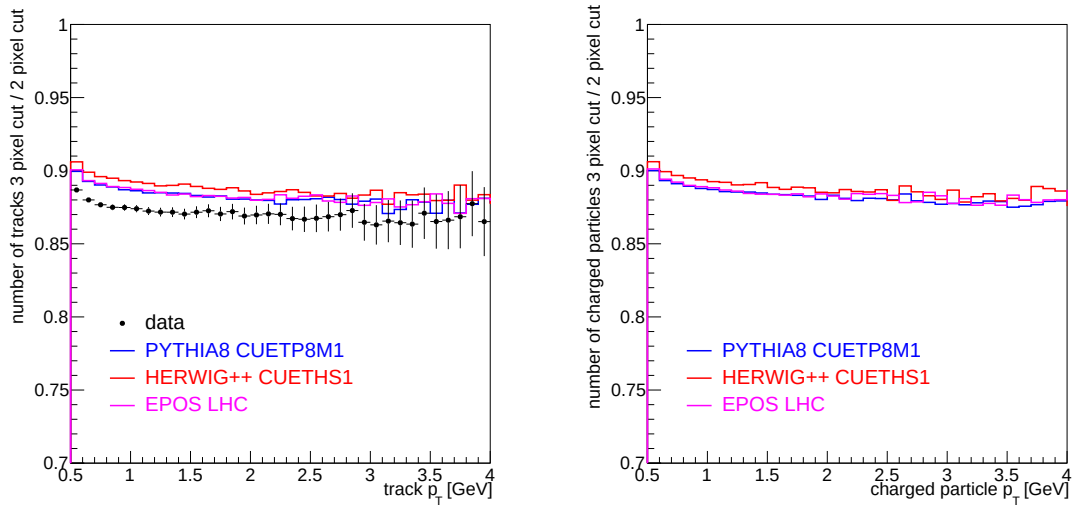


Figure B.8: Ratios between the transverse momentum distribution of the tracks (left) and the matched charged particles (right) after the requirement to have at least 3 pixel hits and the distribution before the requirement. The data are compared to the predictions of the Monte Carlo event generators PYTHIA8, HERWIG++ and EPOS.

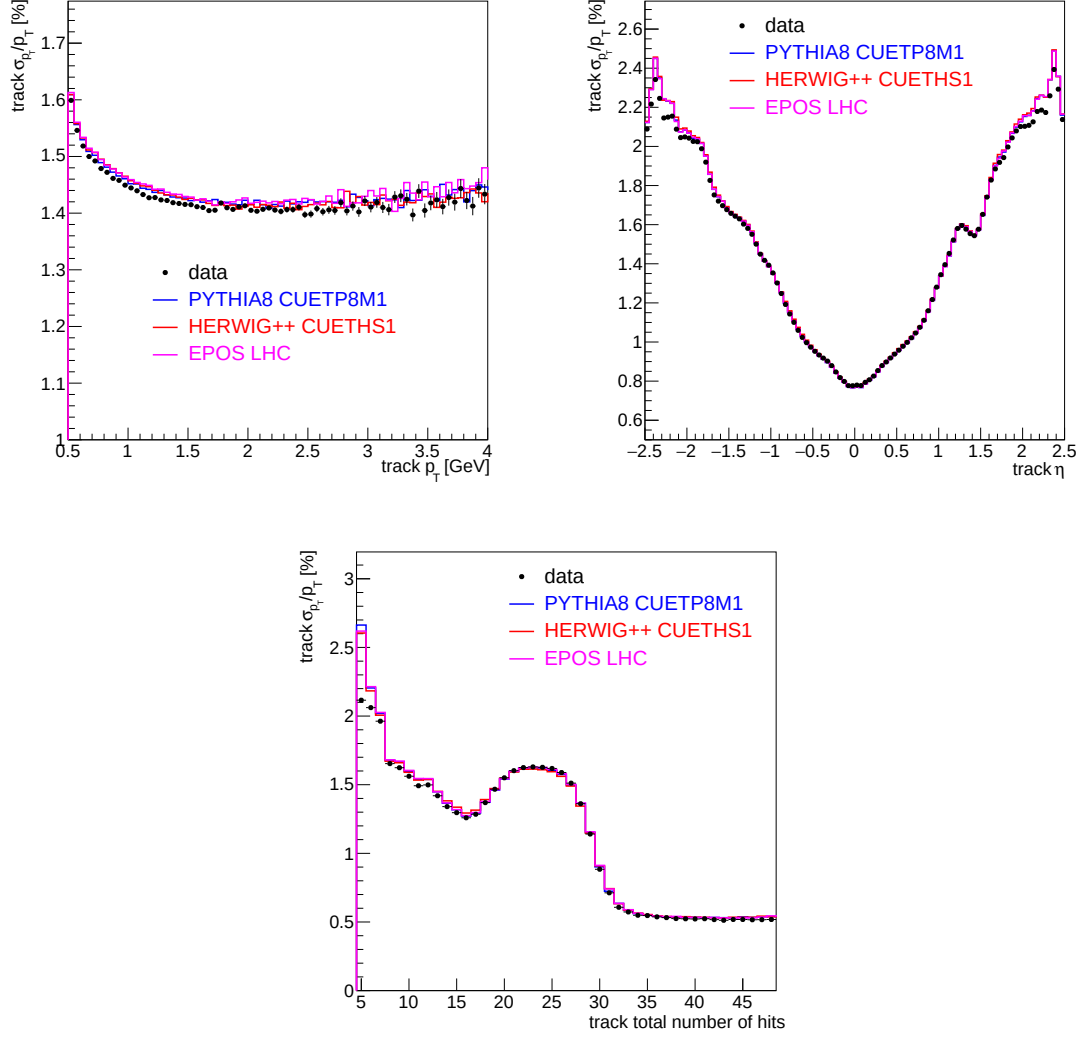


Figure B.9: Dependence of the relative transverse momentum resolution on the track transverse momentum, pseudorapidity and number of associated hits. The data are compared to the predictions of the Monte Carlo event generators PYTHIA8, HERWIG++ and EPOS.

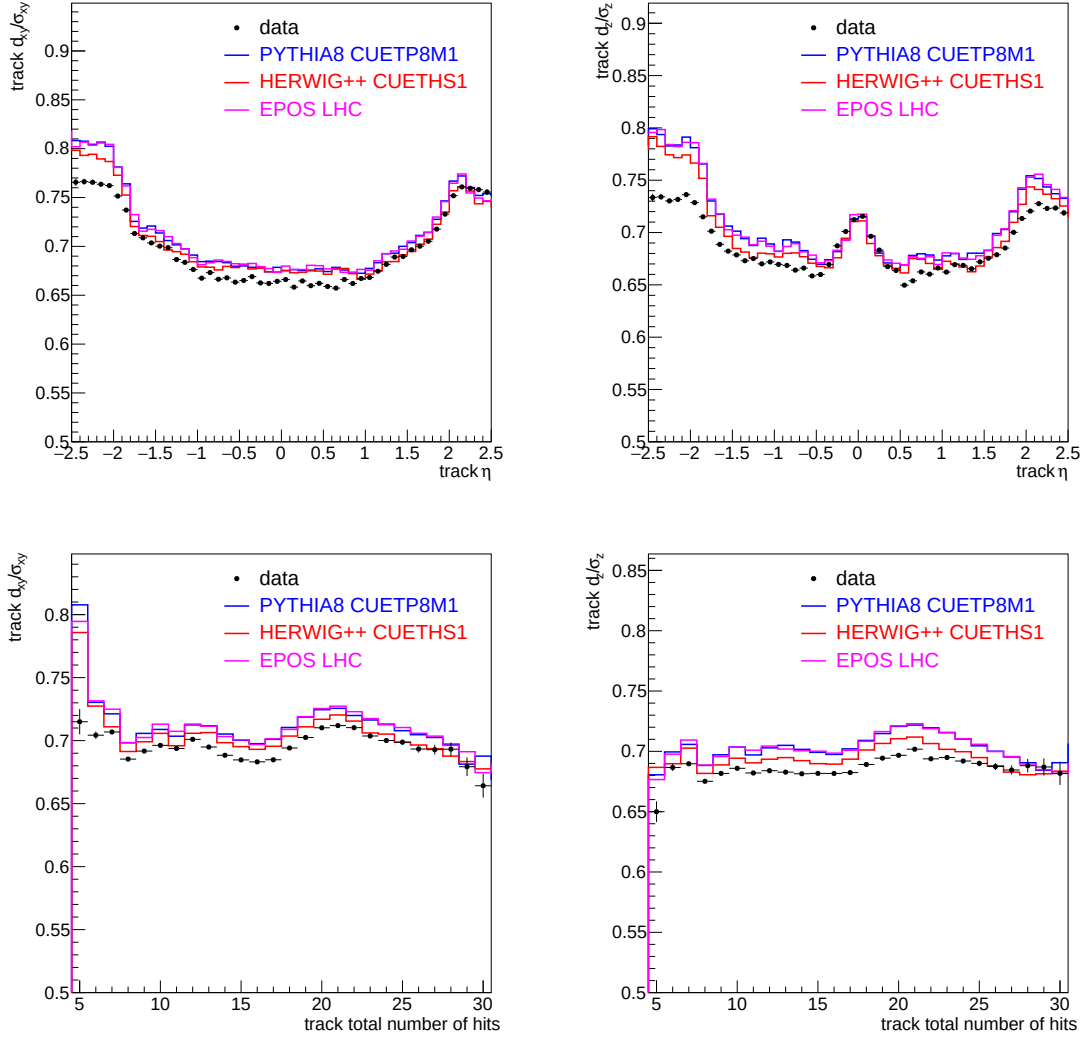


Figure B.10: Dependence of the significances d_{xy}/σ_{xy} and d_z/σ_z on the track pseudorapidity and total number of associated hits. The data are compared to the predictions of the Monte Carlo event generators PYTHIA8, HERWIG++ and EPOS.

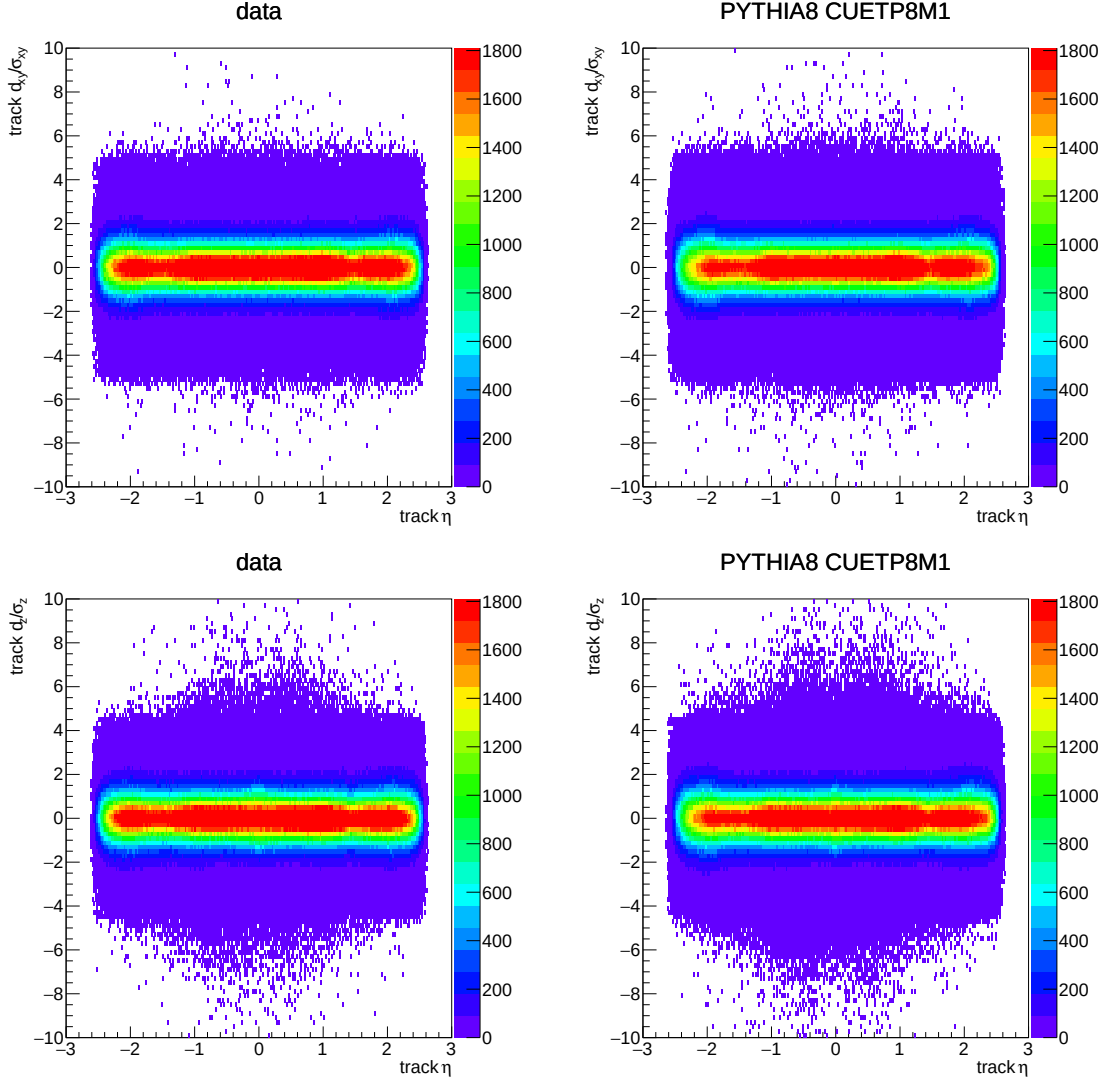


Figure B.11: Dependence of the significances d_{xy}/σ_{xy} and d_z/σ_z on the track pseudorapidity in a two dimensional view. The data are compared to the predictions of the Monte Carlo event generator PYTHIA8.

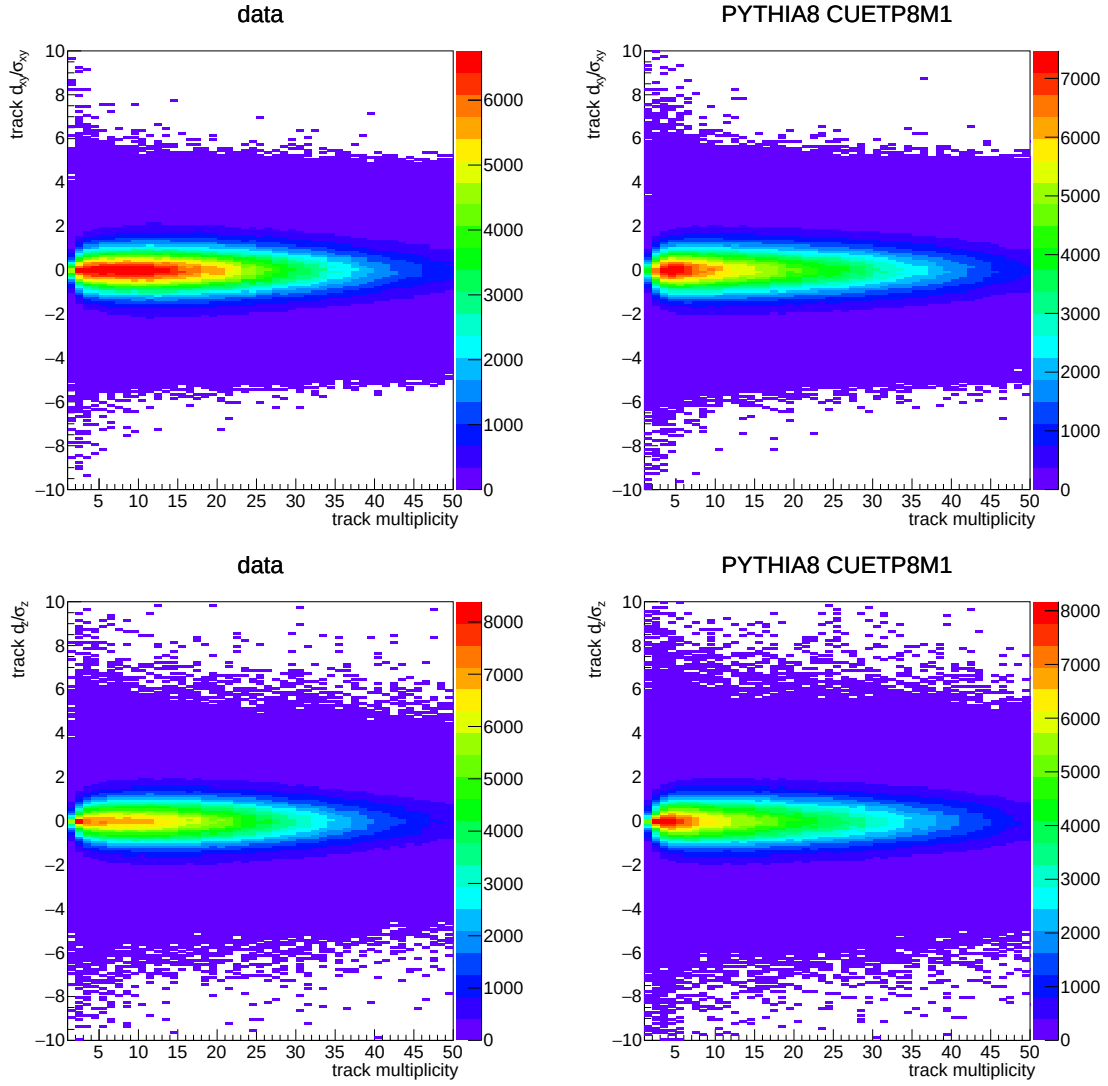


Figure B.12: Dependence of the significances d_{xy}/σ_{xy} and d_z/σ_z on the track multiplicity in a two dimensional view. The data are compared to the predictions of the Monte Carlo event generator PYTHIA8.

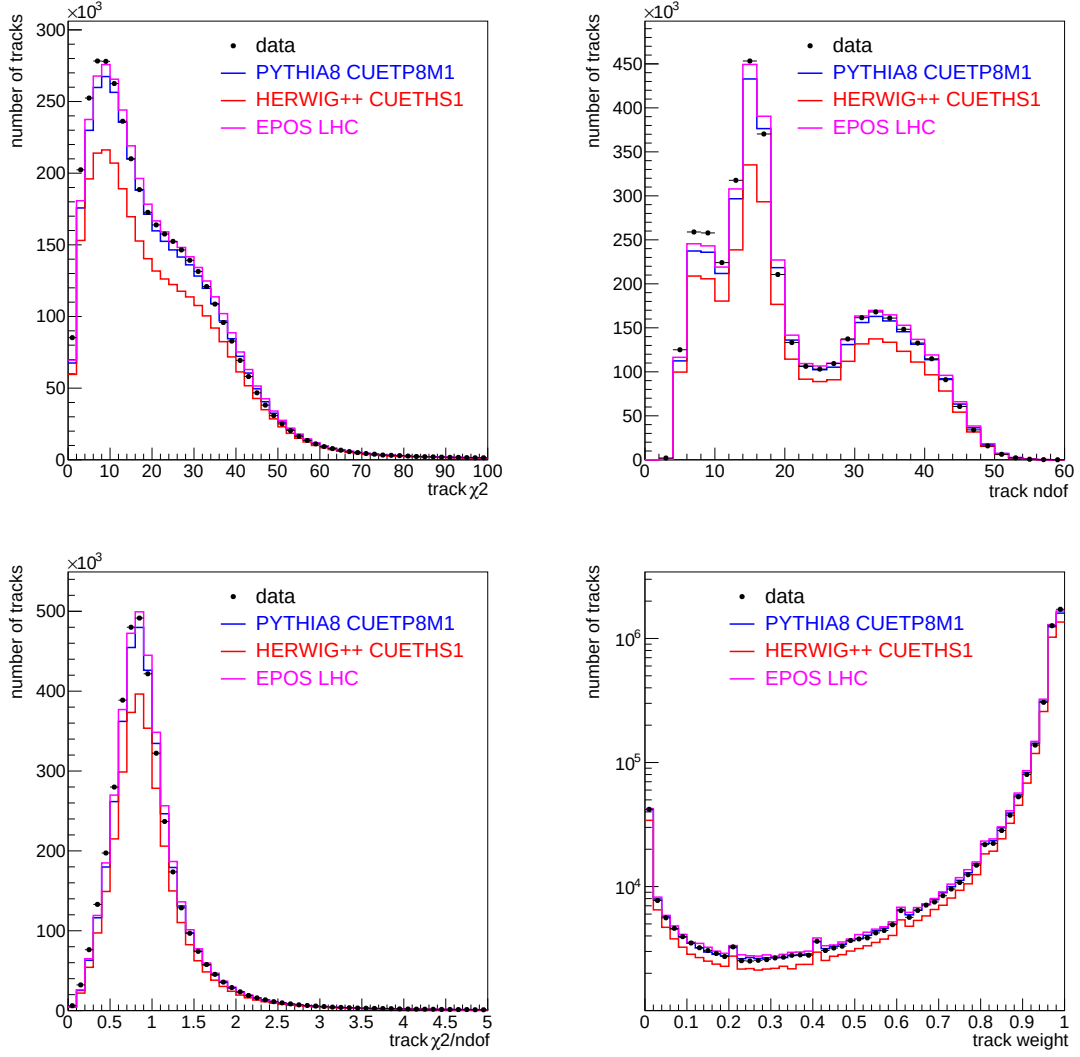


Figure B.13: Distributions of the χ^2 , number of degrees of freedom, reduced χ^2 and weights associated to the track reconstruction. The data are compared to the predictions of the Monte Carlo event generators PYTHIA8, HERWIG++ and EPOS.

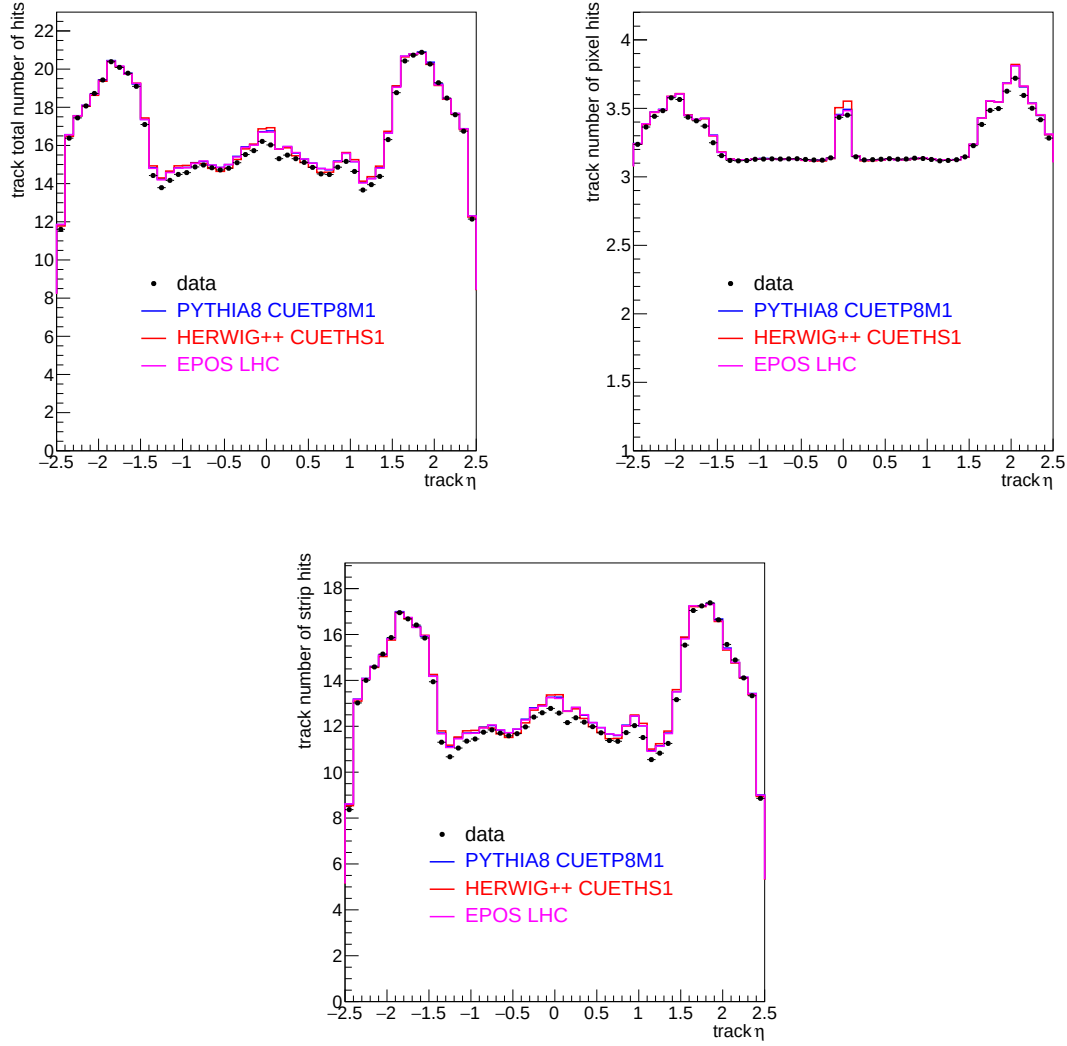


Figure B.14: Dependence of the hit multiplicity on the pseudorapidity of the associated track, for all the tracker layers and for the pixel and strip layers separately. The data are compared to the predictions of the Monte Carlo event generators PYTHIA8, HERWIG++ and EPOS.

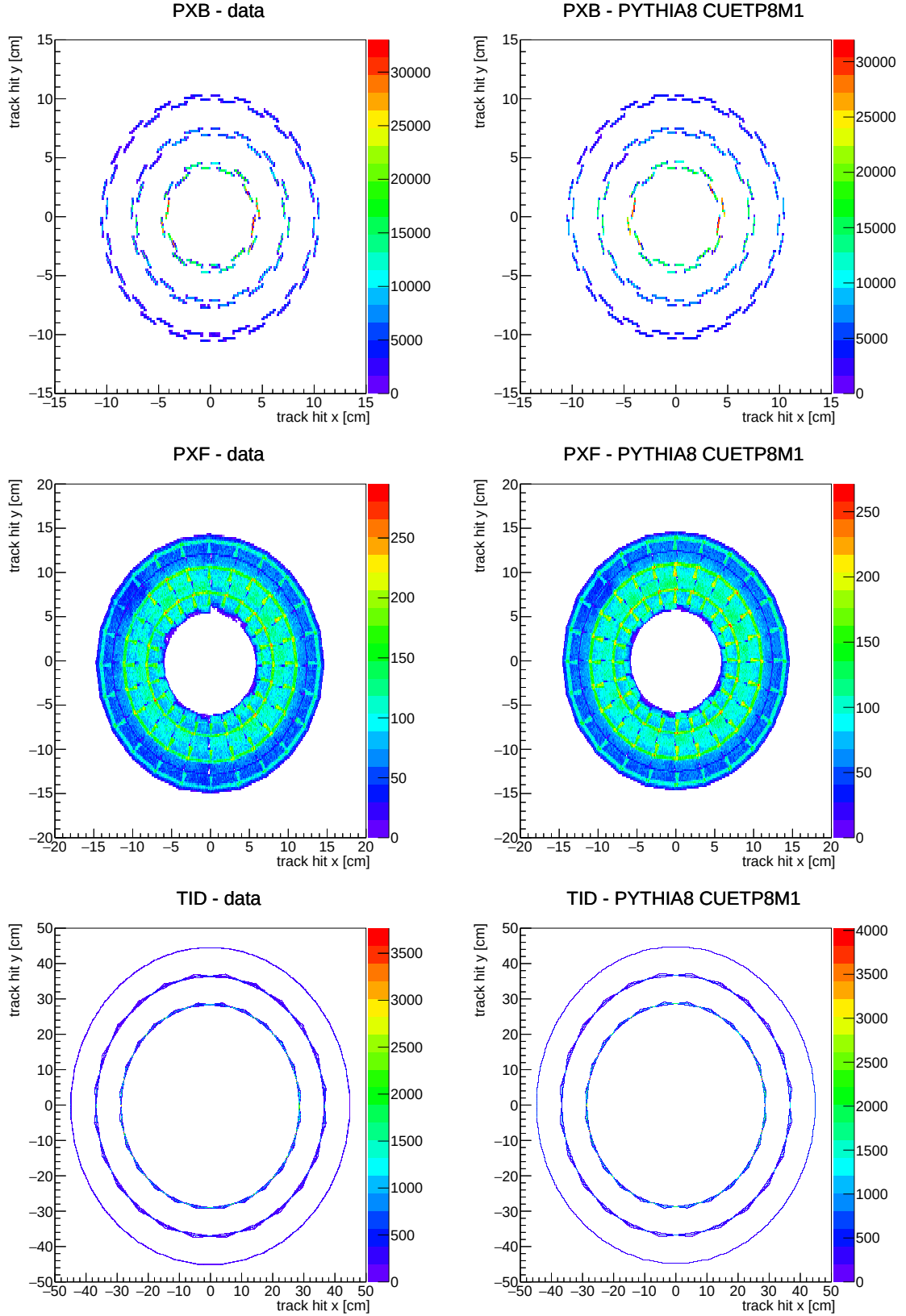


Figure B.15: Distributions of the x, y coordinates of the hits from the tracks associated to the primary vertex in the pixel barrel (PXB), pixel forward (PXF) and tracker inner disk (TID). The data are compared to the predictions of the Monte Carlo event generator PYTHIA8.

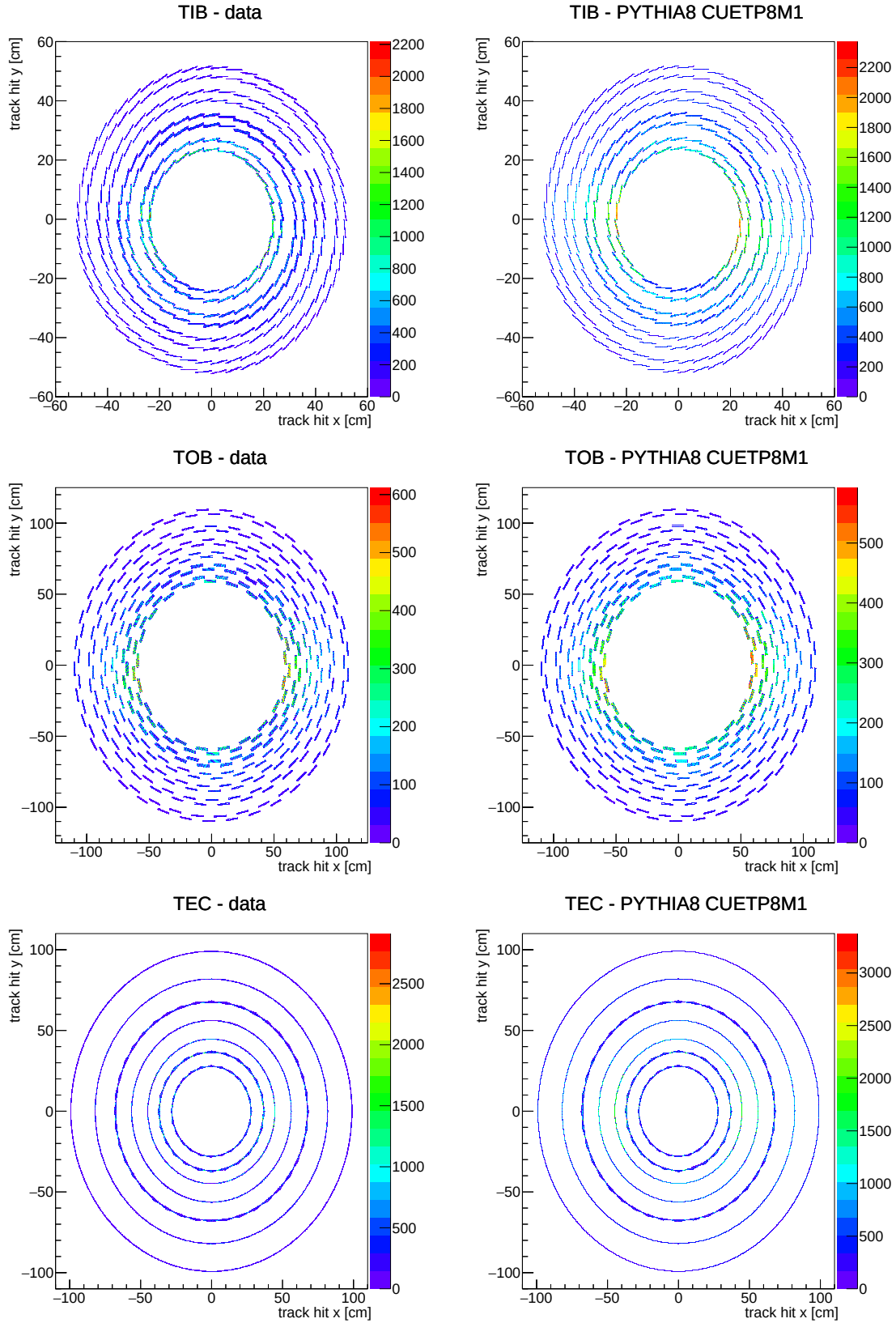


Figure B.16: Distributions of the x, y coordinates of the hits from the tracks associated to the primary vertex in the tracker inner barrel (TIB), tracker outer barrel (TOB) and tracker endcap (TEC). The data are compared to the predictions of the Monte Carlo event generator PYTHIA8.

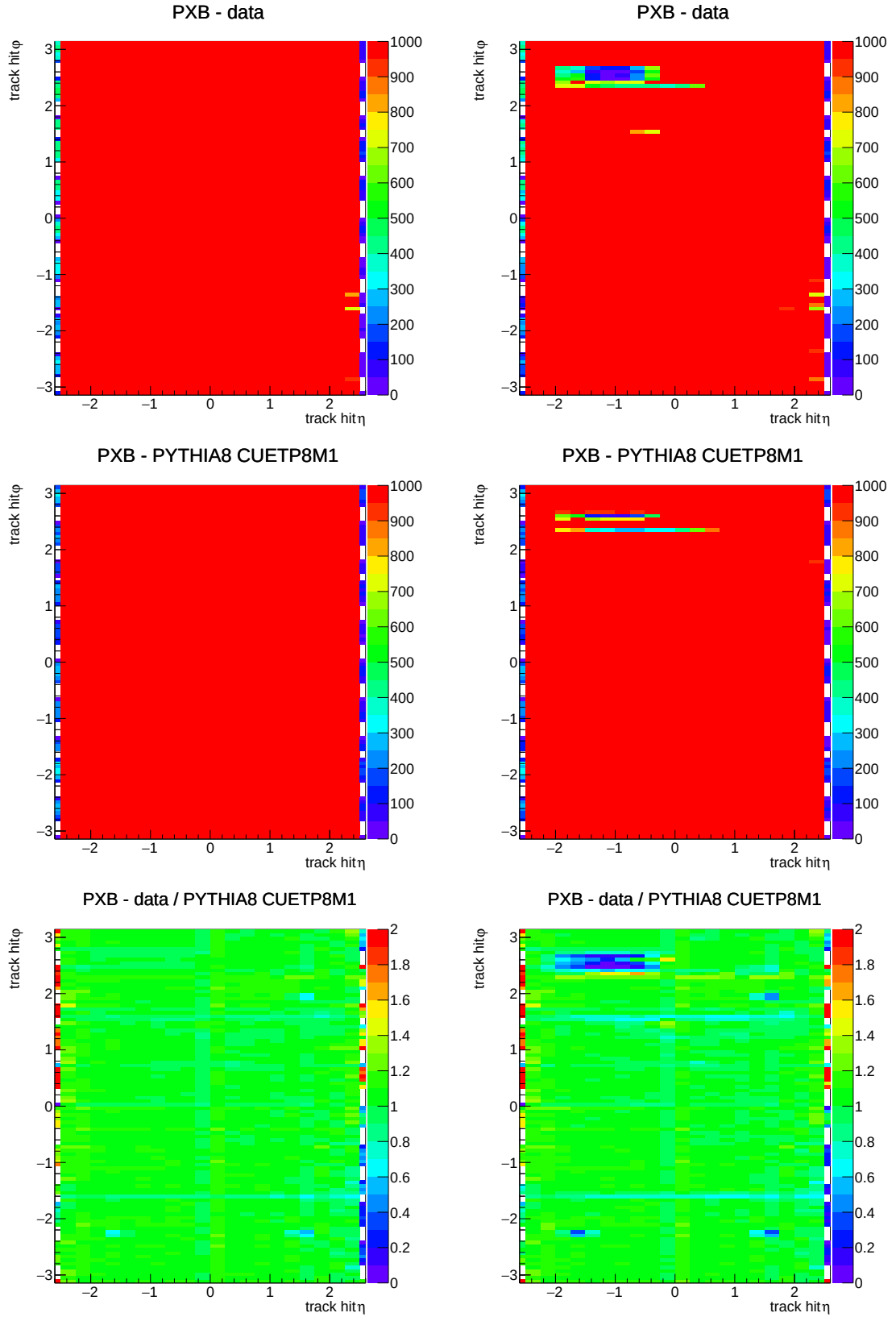


Figure B.17: Distributions of the pseudorapidity and azimuthal angle of the hits from the tracks associated to the primary vertex in the pixel barrel before (left) and after (right) the requirement to have at least 3 pixel hits. The data are compared to the predictions of the Monte Carlo event generator PYTHIA8.

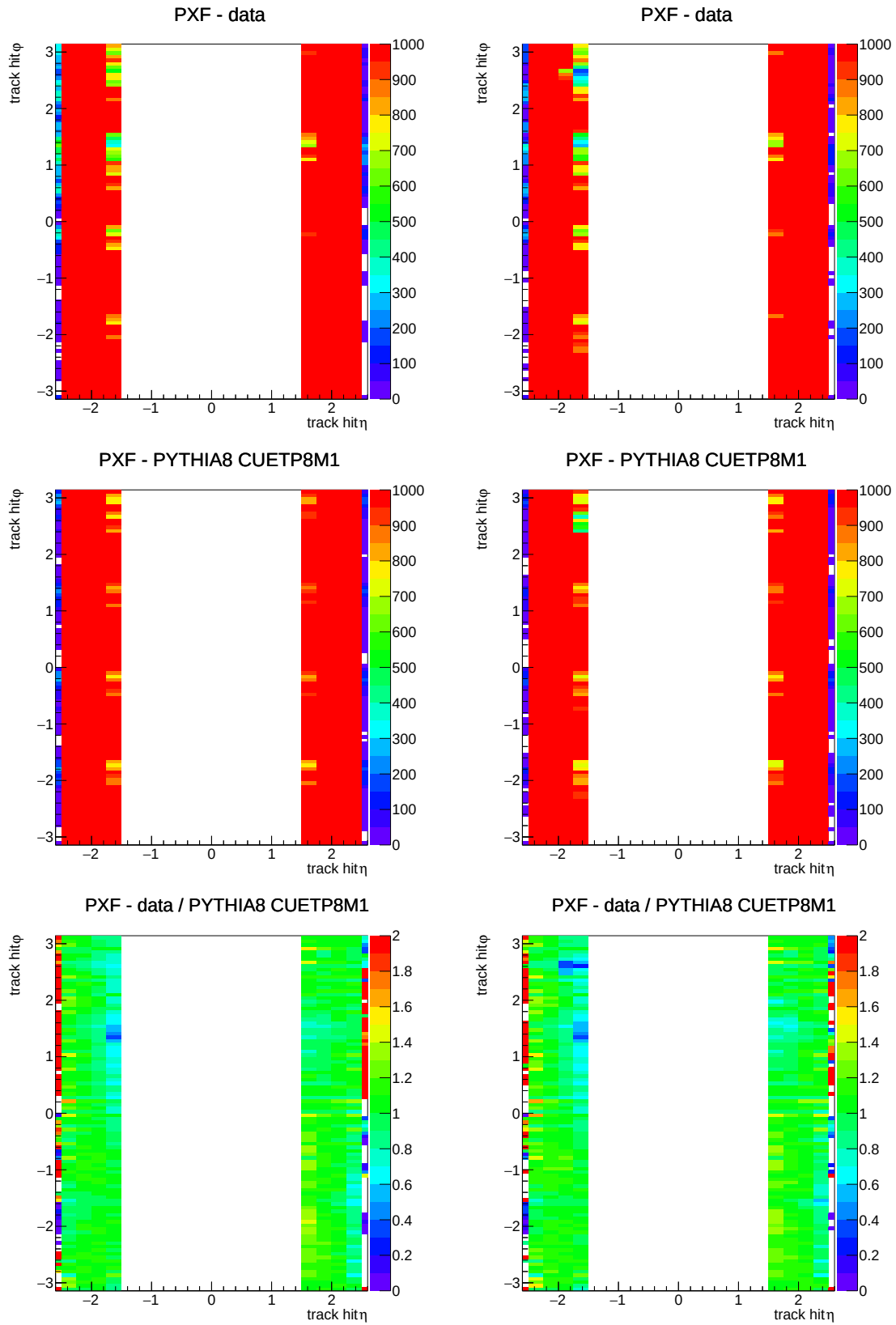


Figure B.18: Distributions of the pseudorapidity and azimuthal angle of the hits from the tracks associated to the primary vertex in the pixel forward before (left) and after (right) the requirement to have at least 3 pixel hits. The data are compared to the predictions of the Monte Carlo event generator PYTHIA8.

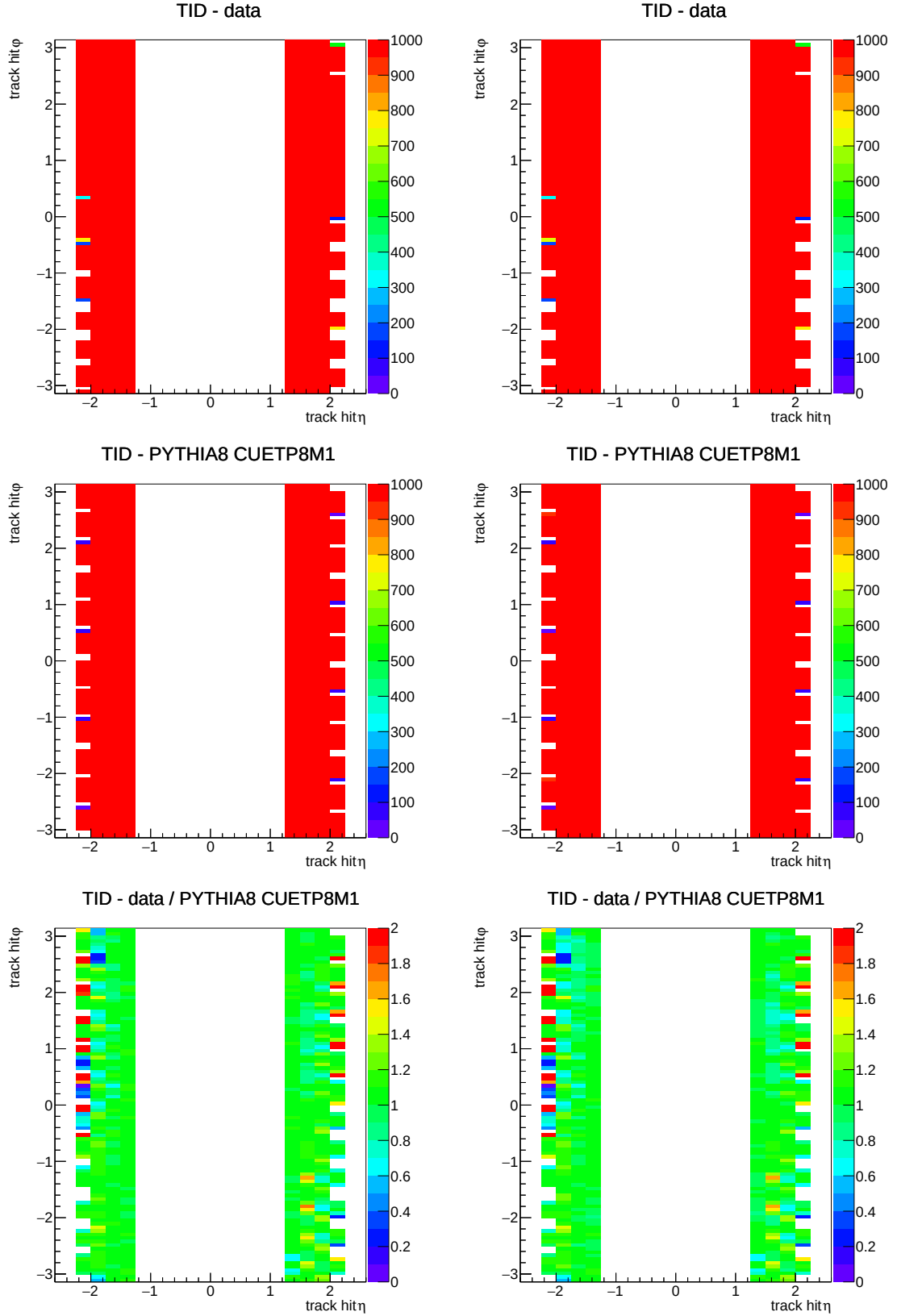


Figure B.19: Distributions of the pseudorapidity and azimuthal angle of the hits from the tracks associated to the primary vertex in the tracker inner disk before (left) and after (right) the requirement to have at least 3 pixel hits. The data are compared to the predictions of the Monte Carlo event generator PYTHIA8.

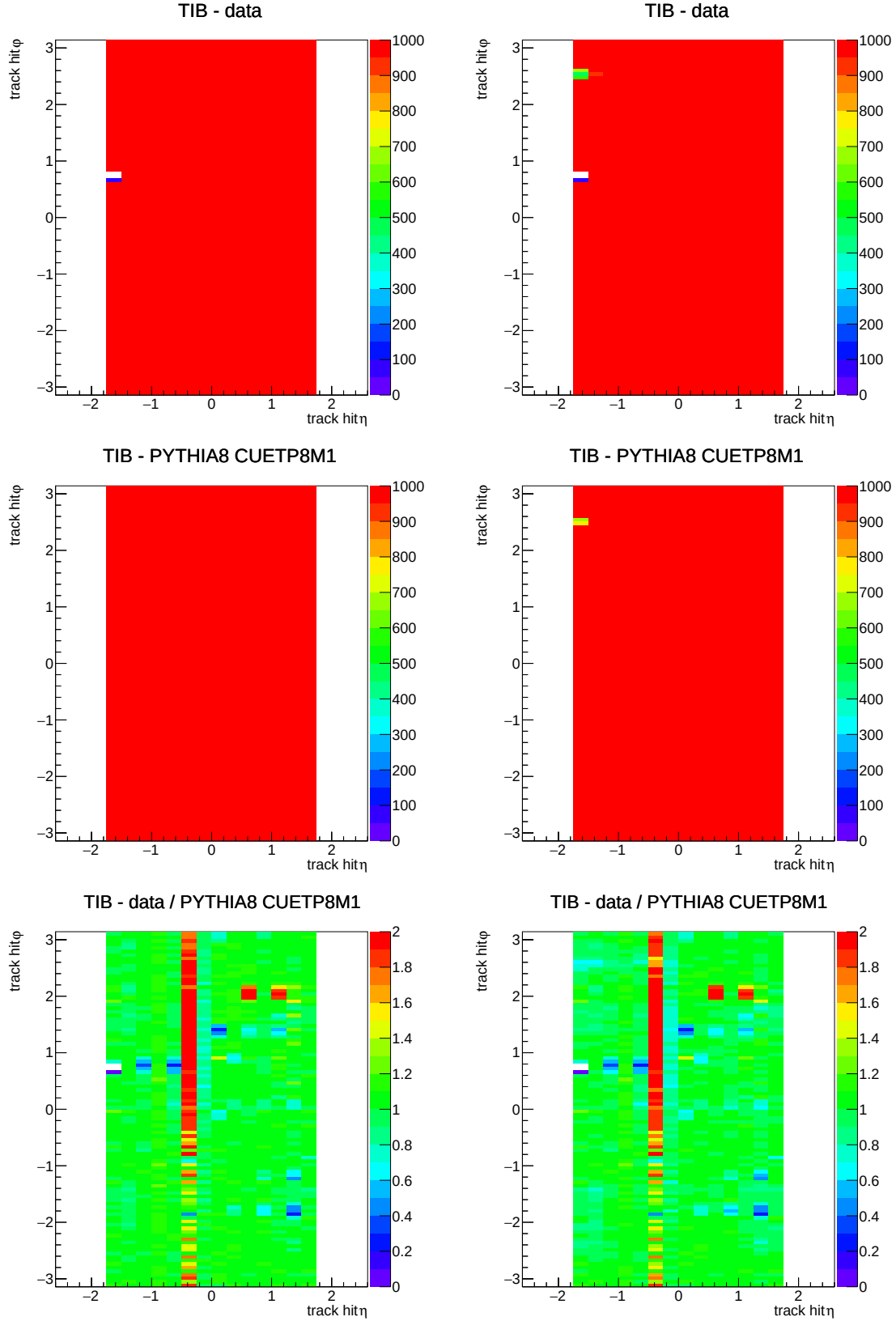


Figure B.20: Distributions of the pseudorapidity and azimuthal angle of the hits from the tracks associated to the primary vertex in the tracker inner barrel before (left) and after (right) the requirement to have at least 3 pixel hits. The data are compared to the predictions of the Monte Carlo event generator PYTHIA8.

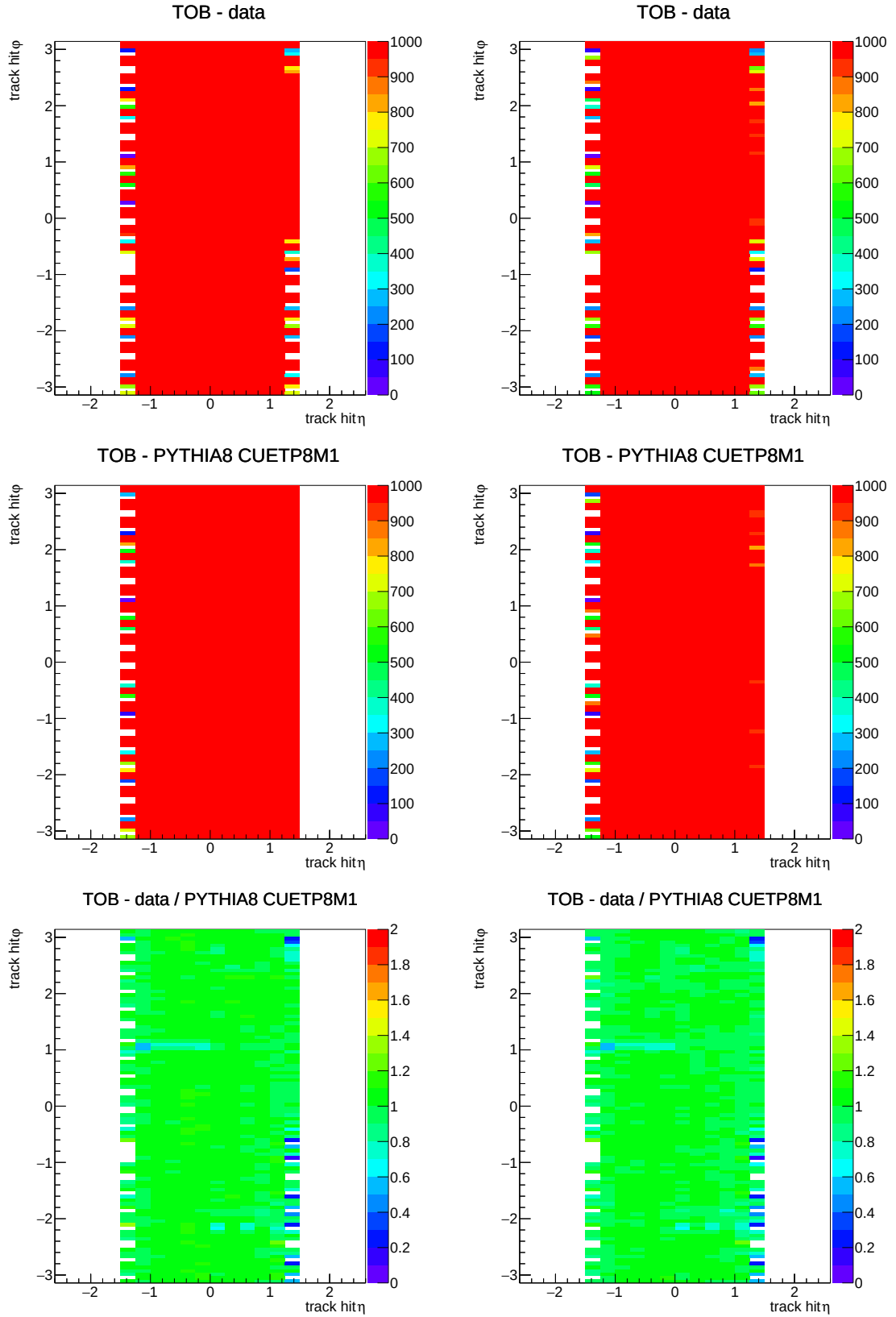


Figure B.21: Distributions of the pseudorapidity and azimuthal angle of the hits from the tracks associated to the primary vertex in the tracker outer barrel before (left) and after (right) the requirement to have at least 3 pixel hits. The data are compared to the predictions of the Monte Carlo event generator PYTHIA8.

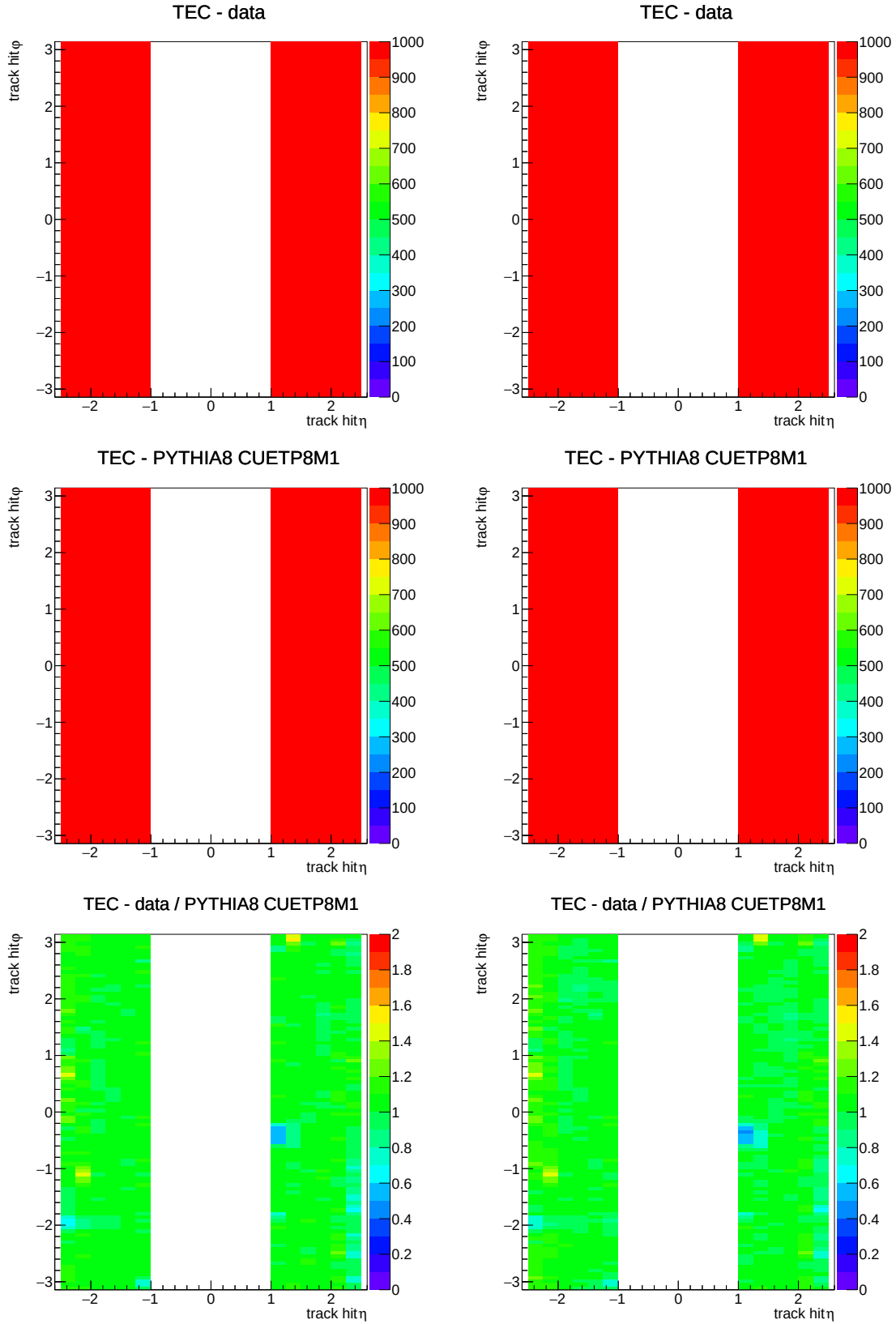


Figure B.22: Distributions of the pseudorapidity and azimuthal angle of the hits from the tracks associated to the primary vertex in the tracker endcap before (left) and after (right) the requirement to have at least 3 pixel hits. The data are compared to the predictions of the Monte Carlo event generator PYTHIA8.

CHAPTER

12

BIBLIOGRAPHY

- [1] CMS Collaboration, “Pseudorapidity distribution of charged hadrons in proton-proton collisions at $\sqrt{s} = 13$ TeV”, *Phys. Lett.* **B751** (2015) 143–163, doi:10.1016/j.physletb.2015.10.004, arXiv:1507.05915. 14
- [2] ALICE Collaboration, “Pseudorapidity and transverse-momentum distributions of charged particles in proton-proton collisions at $\sqrt{s} = 13$ TeV”, *Phys. Lett.* **B753** (2016) 319–329, doi:10.1016/j.physletb.2015.12.030, arXiv:1509.08734. 14
- [3] CMS Collaboration, “Measurement of pseudorapidity distributions of charged particles in proton-proton collisions at $\sqrt{s} = 8$ TeV by the CMS and TOTEM experiments”, *Eur. Phys. J. C* **74** (2014) 3053, doi:10.1140/epjc/s10052-014-3053-6. 14
- [4] CMS Collaboration, “Transverse momentum and pseudorapidity distributions of charged hadrons in pp collisions at $\sqrt{s} = 0.9$ and 2.36 TeV”, *JHEP* **02** (2010) 041, doi:10.1007/JHEP02(2010)041. 14
- [5] CMS Collaboration, “Transverse-momentum and pseudorapidity distributions of charged hadrons in pp collisions at $\sqrt{s} = 7$ TeV”, *Phys. Rev. Lett.* **105** (2010) 022002, doi:10.1103/PhysRevLett.105.022002. 14
- [6] ATLAS Collaboration, “Charged-particle multiplicities in pp interactions measured with the ATLAS detector at the LHC”, *New J. Phys.* **13** (2011) 053033, doi:10.1088/1367-2630/13/5/053033. 14
- [7] ALICE Collaboration, “Charged-particle multiplicity measurement in proton-proton collisions at $\sqrt{s} = 0.9$ and 2.36 TeV with ALICE at LHC”, *Eur. Phys. J. C* **68** (2010) 89, doi:10.1140/epjc/s10052-010-1339-x. 14

- [8] ALICE Collaboration, “Charged-particle multiplicity measurement in proton–proton collisions at $\sqrt{s} = 7$ TeV with ALICE at LHC”, *Eur. Phys. J. C* **68** (2010) 345, doi:10.1140/epjc/s10052-010-1350-2. 14
- [9] LHCb Collaboration, “Measurement of charged particle multiplicities in pp collisions at $\sqrt{s} = 7$ TeV in the forward region”, *Eur. Phys. J. C* **72** (2012) 1947, doi:10.1140/epjc/s10052-012-1947-8. 14
- [10] TOTEM Collaboration, “Measurement of the forward charged-particle pseudorapidity density in pp collisions at $\sqrt{s} = 7$ TeV with the TOTEM experiment”, *Europhys. Lett.* **98** (2012) 31002, doi:10.1209/0295-5075/98/31002. 14
- [11] CDF Collaboration, “Measurement of particle production and Inclusive Differential Cross Sections in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV”, *Phys. Rev. D* **79** (2009) 112005, doi:10.1103/PhysRevD.79.112005. [Erratum doi:10.1103/PhysRevD.82.119903]. 14
- [12] CDF Collaboration, “Pseudorapidity distributions of charged particles produced in $p\bar{p}$ interactions at $\sqrt{s} = 630$ GeV and 1800 GeV”, *Phys. Rev. D* **41** (1990) 2330, doi:10.1103/PhysRevD.41.2330. 14
- [13] UA1 Collaboration, “A study of the general characteristics of proton-antiproton collisions at $\sqrt{s} = 0.2$ to 0.9 TeV”, *Nucl. Phys. B* **335** (1990) 261, doi:10.1016/0550-3213(90)90493-W. 14
- [14] UA4 Collaboration, “Pseudorapidity distribution of charged particles in diffraction dissociation events at the CERN SPS collider”, *Phys. Lett. B* **166** (1986) 459, doi:10.1016/0370-2693(86)91598-4. 14
- [15] UA5 Collaboration, “Scaling of pseudorapidity distributions at c.m. energies up to 0.9 TeV”, *Z. Phys. C* **33** (1986) 1, doi:10.1007/BF01410446. 14
- [16] J. J. Thomson, “Cathode rays”, *Phil. Mag. Ser.5* **44** (1897) 293–316, doi:10.1080/14786449708621070. 18
- [17] J. H. University, “Standard Model Poster”, August, 2016.
<https://web.archive.org/web/20160304133522/https://www.pha.jhu.edu/%7Edfehling/particle.gif>. 18
- [18] N. Bartosik, “SM-particles table”, Jun, 2016.
http://bartosik.pp.ua/hep_sketches/sm_particles. 19
- [19] A. Hoecker, “Physics at the LHC Run-2 and Beyond”, in *2016 European School of High-Energy Physics (ESHEP 2016) Skeikampen, Norway, June 15-28, 2016*. 2016. arXiv:1611.07864. 20, 21
- [20] Particle Data Group Collaboration, “Review of Particle Physics”, *Chin. Phys.* **C40** (2016), no. 10, 100001, doi:10.1088/1674-1137/40/10/100001. 23, 34, 40
- [21] G. ALTARELLI, “The Standard model of particle physics”, arXiv:hep-ph/0510281. 24

-
- [22] CMS Collaboration, “Measurement and QCD analysis of double-differential inclusive jet cross sections in pp collisions at $\sqrt{s} = 8$ TeV and cross section ratios to 2.76 and 7 TeV”, *Journal of High Energy Physics* **2017** (Mar, 2017) 156, doi:10.1007/JHEP03(2017)156. 24, 25
 - [23] M. Breidenbach et al., “Observed Behavior of Highly Inelastic Electron-Proton Scattering”, *Phys. Rev. Lett.* **23** (Oct, 1969) 935–939, doi:10.1103/PhysRevLett.23.935. 26
 - [24] W. Albrecht et al., “Elastic Electron-Proton Scattering at Momentum Transfers up to 245 F^{-2} ”, *Phys. Rev. Lett.* **17** (Dec, 1966) 1192–1195, doi:10.1103/PhysRevLett.17.1192. 26
 - [25] J. Aubert et al., “The ratio of the nucleon structure functions F_2^N for iron and deuterium”, *Physics Letters B* **123** (1983), no. 3, 275 – 278, doi:http://dx.doi.org/10.1016/0370-2693(83)90437-9. 26
 - [26] C. Adloff et al., “A measurement of the proton structure function $F_2(x, Q^2)$ at low x and low $\{Q^2\}$ at $\{HERA\}$ ”, *Nuclear Physics B* **497** (1997), no. 1–2, 3 – 28, doi:https://doi.org/10.1016/S0550-3213(97)00301-5. 26
 - [27] J. Breitweg et al., “Measurement of the proton structure function F_2 and $\sigma_{tot}^{\gamma^*p}$ at low Q^2 and very low x at HERA”, *Physics Letters B* **407** (1997), no. 3–4, 432 – 448, doi:https://doi.org/10.1016/S0370-2693(97)00905-2. 26
 - [28] H. Abramowicz et al., “Combination of measurements of inclusive deep inelastic $e^\pm p$ scattering cross sections and QCD analysis of HERA data”, *The European Physical Journal C* **75** (2015), no. 12, 580, doi:10.1140/epjc/s10052-015-3710-4. 27, 28
 - [29] A. Benvenuti et al., “A high statistics measurement of the proton structure functions $F_2(x, Q^2)$ and R from deep inelastic muon scattering at high Q^2 ”, *Physics Letters B* **223** (1989), no. 3, 485 – 489, doi:http://dx.doi.org/10.1016/0370-2693(89)91637-7. 28
 - [30] The New Muon Collaboration Collaboration, “Measurement of the proton and deuteron structure functions, F_2^p and F_2^d , and of the ratio σ_L/σ_T ”, *Nuclear Physics B* **483** (1997), no. 1–2, 3 – 43, doi:https://doi.org/10.1016/S0550-3213(96)00538-X. 28
 - [31] V. N. Gribov and L. N. Lipatov, “Deep inelastic $e p$ scattering in perturbation theory”, *Sov. J. Nucl. Phys.* **15** (1972) 438. 29
 - [32] G. Altarelli and G. Parisi, “Asymptotic freedom in parton language”, *Nuclear Physics B* **126** (1977), no. 2, 298 – 318, doi:http://dx.doi.org/10.1016/0550-3213(77)90384-4. 29
 - [33] Y. L. Dokshitzer, “Calculation of the Structure Functions for Deep Inelastic Scattering and $e^+ e^-$ Annihilation by Perturbation Theory in Quantum Chromodynamics.”, *Sov. Phys. JETP* **46** (1977) 641. 29

- [34] E. A. Kuraev, L. N. Lipatov, and V. S. Fadin, “Multi - Reggeon Processes in the Yang-Mills Theory”, *Sov. Phys. JETP* **44** (1976) 443–450. [Zh. Eksp. Teor. Fiz.71,840(1976)]. 29
- [35] E. A. Kuraev, L. N. Lipatov, and V. S. Fadin, “The Pomeranchuk Singularity in Nonabelian Gauge Theories”, *Sov. Phys. JETP* **45** (1977) 199–204. [Zh. Eksp. Teor. Fiz.72,377(1977)]. 29
- [36] M. Ciafaloni, “Coherence effects in initial jets at small Q^2/s ”, *Nuclear Physics B* **296** (1988), no. 1, 49 – 74, doi:10.1016/0550-3213(88)90380-X. 29
- [37] S. Catani, F. Fiorani, and G. Marchesini, “QCD coherence in initial state radiation”, *Physics Letters B* **234** (1990), no. 3, 339 – 345, doi:10.1016/0370-2693(90)91938-8. 29
- [38] G. Dissertori, I. G. Knowles, and M. Schmelling, “Quantum Chromodynamics: High Energy Experiments and Theory (International Series of Monographs on Physics)”. Oxford University Press, 1 edition, 5, 2009. 30, 31
- [39] D. d’Enterria, “Forward Physics at the LHC: within and beyond the Standard Model”, *Brazilian Journal of Physics* **38** (09, 2008) 381 – 390. 30
- [40] M. Shifman, ed., “At the Frontier of Particle Physics : Handbook of QCD (3 Volumes)”. World Scientific Publishing Company, 4, 2001. 32
- [41] CMS Collaboration, “Production of leading charged particles and leading charged-particle jets at small transverse momenta in pp collisions at $\sqrt{s} = 8$ TeV”, *Phys. Rev.* **D92** (2015), no. 11, 112001, doi:10.1103/PhysRevD.92.112001. 33, 42
- [42] A. Donnachie and P. Landshoff, “Total cross sections”, *Physics Letters B* **296** (1992), no. 1, 227 – 232, doi:http://dx.doi.org/10.1016/0370-2693(92)90832-0. 33
- [43] A. Donnachie and P. Landshoff, “pp and pp elastic scattering”, *Nuclear Physics B* **231** (1984), no. 2, 189 – 204, doi:http://dx.doi.org/10.1016/0550-3213(84)90283-9. 33
- [44] J. C. Collins, “Proof of factorization for diffractive hard scattering”, *Phys. Rev.* **D57** (1998) 3051–3056, doi:10.1103/PhysRevD.61.019902, 10.1103/PhysRevD.57.3051, arXiv:hep-ph/9709499. [Erratum: Phys. Rev.D61,019902(2000)]. 34
- [45] A. Berera and D. E. Soper, “Behavior of diffractive parton distribution functions”, *Phys. Rev.* **D53** (1996) 6162–6179, doi:10.1103/PhysRevD.53.6162, arXiv:hep-ph/9509239. 34
- [46] M. Arneodo and M. Diehl, “Diffraction for non-believers”, in *HERA and the LHC: A Workshop on the implications of HERA for LHC physics. Proceedings, Part B*, pp. 425–446. 2005. arXiv:hep-ph/0511047. 35
- [47] ZEUS Collaboration, “Dissociation of virtual photons in events with a leading proton at HERA”, *Eur. Phys. J.* **C38** (2004) 43–67, doi:10.1140/epjc/s2004-02047-4, arXiv:hep-ex/0408009. 35

-
- [48] CDF Collaboration, “Diffractive dijets with a leading antiproton in $\bar{p}p$ collisions at $\sqrt{s} = 1800$ GeV”, *Phys. Rev. Lett.* **84** (2000) 5043–5048, doi:10.1103/PhysRevLett.84.5043. 35
 - [49] T. Sjöstrand et al., “An Introduction to PYTHIA 8.2”, *Comput. Phys. Commun.* **191** (2015) 159–177, doi:10.1016/j.cpc.2015.01.024, arXiv:1410.3012. 36
 - [50] M. Bähr et al., “Herwig++ physics and manual”, *Eur. Phys. J. C* **58** (2008) 639, doi:10.1140/epjc/s10052-008-0798-9, arXiv:0803.0883. 36
 - [51] T. Gleisberg et al., “Event generation with SHERPA 1.1”, *JHEP* **02** (2009) 007, doi:10.1088/1126-6708/2009/02/007, arXiv:0811.4622. 36
 - [52] J. Alwall et al., “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations”, *JHEP* **07** (2014) 079, doi:10.1007/JHEP07(2014)079, arXiv:1405.0301. 36
 - [53] C. Oleari, “The POWHEG-BOX”, *Nucl. Phys. Proc. Suppl.* **205-206** (2010) 36–41, doi:10.1016/j.nuclphysbps.2010.08.016, arXiv:1007.3893. 36
 - [54] S. Frixione, P. Nason, and C. Oleari, “Matching NLO QCD computations with Parton Shower simulations: the POWHEG method”, *JHEP* **11** (2007) 070, doi:10.1088/1126-6708/2007/11/070, arXiv:0709.2092. 36
 - [55] S. Hoeche et al., “Matching parton showers and matrix elements”, in *HERA and the LHC: A Workshop on the implications of HERA for LHC physics: Proceedings Part A*, pp. 288–289. 2005. arXiv:hep-ph/0602031. doi:10.5170/CERN-2005-014.288. 37
 - [56] S. Höche, “Introduction to parton-shower event generators”, in *Theoretical Advanced Study Institute in Elementary Particle Physics: Journeys Through the Precision Frontier: Amplitudes for Colliders (TASI 2014) Boulder, Colorado, June 2-27, 2014*. 2014. arXiv:1411.4085. 38
 - [57] A. Buckley et al., “General-purpose event generators for LHC physics”, *Phys. Rept.* **504** (2011) 145–233, doi:10.1016/j.physrep.2011.03.005, arXiv:1101.2599. 39
 - [58] A. Grebenyuk et al., “Jet production and the inelastic pp cross section at the LHC”, *Phys.Rev.* **D86** (2012) 117501, doi:10.1103/PhysRevD.86.117501. 41, 42, 153
 - [59] C. Lefèvre, “The CERN accelerator complex. Complexe des accélérateurs du CERN”, (Dec, 2008). <https://cds.cern.ch/record/1260465>. 44
 - [60] CMS Collaboration, “CMS Physics: Technical Design Report Volume 1: Detector Performance and Software”,. There is an error on cover due to a technical problem for some items. 45, 51
 - [61] ALICE Collaboration, “The ALICE experiment at the CERN LHC”, *Journal of Instrumentation* **3** (2008), no. 08, S08002, doi:10.1088/1748-0221/3/08/S08002. 45
 - [62] ATLAS Collaboration, “The ATLAS Experiment at the CERN Large Hadron Collider”, *Journal of Instrumentation* **3** (2008), no. 08, S08003, doi:10.1088/1748-0221/3/08/S08003. 45

- [63] LHCb Collaboration, “The LHCb Detector at the LHC”, *Journal of Instrumentation* **3** (2008), no. 08, S08005, doi:10.1088/1748-0221/3/08/S08005. 45
- [64] LHCf Collaboration, “The LHCf detector at the CERN Large Hadron Collider”, *Journal of Instrumentation* **3** (2008), no. 08, S08006, doi:10.1088/1748-0221/3/08/S08006. 45
- [65] TOTEM Collaboration, “The TOTEM Experiment at the CERN Large Hadron Collider”, *Journal of Instrumentation* **3** (2008), no. 08, S08007, doi:10.1088/1748-0221/3/08/S08007. 45
- [66] CMS Collaboration, “The CMS experiment at the CERN LHC”, *JINST* **3** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004. 45
- [67] G. Bauer et al., “The CMS High Level Trigger System: Experience and future development”, *J. Phys. Conf. Ser.* **396** (2012) 012008, doi:10.1088/1742-6596/396/1/012008. 47
- [68] CMS Collaboration, “The CMS trigger system”, *JINST* **12** (2017), no. 01, P01020, doi:10.1088/1748-0221/12/01/P01020, arXiv:1609.02366. 47
- [69] CMS Collaboration, “Description and performance of track and primary-vertex reconstruction with the CMS tracker”, *J. Instrum.* **9** (2014) P10009, doi:10.1088/1748-0221/9/10/P10009. 48, 52, 53
- [70] A. Penzo, Y. Onel, and the CMS Collaboration, “The CMS-HF quartz fiber calorimeters”, *Journal of Physics: Conference Series* **160** (2009), no. 1, 012014. 50, 54
- [71] CMS Collaboration, “CMS technical design report, volume II: Physics performance”, *J. Phys. G* **34** (2007) 995, doi:10.1088/0954-3899/34/6/S01. 51
- [72] CMS Collaboration, F. Beaudette, “The CMS Particle Flow Algorithm”, in *Proceedings, International Conference on Calorimetry for the High Energy Frontier (CHEF 2013): Paris, France, April 22-25, 2013*, pp. 295–304. 2013. arXiv:1401.8155. 51
- [73] W. Adam, B. Mangano, T. Speer and T. Todorov, “Track Reconstruction in the CMS tracker”, *Technical Report CMS-NOTE-2006-041* (2006). 52
- [74] CMS Collaboration, “Track and vertex reconstruction in CMS”, *Nucl. Inst. Meth. A* **582** (2007) 781, doi:10.1016/j.nima.2007.07.091. 54
- [75] R. Fruhwirth, W. Waltenberger, and P. Vanlaer, “Adaptive vertex fitting”, *J. Phys. G* **34** (2007) N343, doi:10.1088/0954-3899/34/12/N01. 54
- [76] GEANT4 Collaboration, “GEANT4—a simulation toolkit”, *Nucl. Instrum. Meth. A* **506** (2003) 250, doi:10.1016/S0168-9002(03)01368-8. 55
- [77] G. Boudoul et al., “Monte Carlo Production Management at CMS”, *Journal of Physics: Conference Series* **664** (2015), no. 7, 072018. 55

-
- [78] CMS Collaboration, “CMSSW Workbook Offline Twiki Page”, (2008).
<https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBook>. 55
 - [79] CMS Collaboration, “Measurement of long-range near-side two-particle angular correlations in pp collisions at $\sqrt{s} = 13$ TeV”, [arXiv:1510.03068](#). 56, 57, 137
 - [80] CMS Collaboration, “Measurement of the underlying event in the Drell-Yan process in proton-proton collisions at $\sqrt{s} = 7$ TeV”, *Eur. Phys. J. C* **72** (2012) 2080, doi:10.1140/epjc/s10052-012-2080-4, [arXiv:1204.1411](#). 59
 - [81] V. Khachatryan et al., “Measurement of the double-differential inclusive jet cross section in proton-proton collisions at $\sqrt{s} = 13$ TeV”, *The European Physical Journal C* **76** (2016), no. 8, 451, doi:10.1140/epjc/s10052-016-4286-3. 59
 - [82] G. D’Agostini, “A Multidimensional unfolding method based on Bayes’ theorem”, *Nucl. Instrum. Meth. A* **362** (1995) 487–498, doi:10.1016/0168-9002(95)00274-X. 107
 - [83] CDF Collaboration, “The Underlying event in hard scattering processes”, *eConf C010630* (2001) P501, [arXiv:hep-ph/0201192](#). 173
 - [84] S. Chatrchyan et al., “Measurement of the underlying event in the Drell-Yan process in proton-proton collisions at $\sqrt{s} = 7$ TeV”, *The European Physical Journal C* **72** (Sep, 2012) 2080, doi:10.1140/epjc/s10052-012-2080-4. 173
 - [85] CMS Collaboration, “Underlying Event Measurements with Leading Particles and Jets in pp collisions at $\sqrt{s} = 13$ TeV”,. 173
 - [86] T. C. collaboration, “Alignment of the CMS tracker with LHC and cosmic ray data”, *Journal of Instrumentation* **9** (2014), no. 06, P06009. 178
 - [87] V. Blobel and C. Kleinwort, “A New method for the high precision alignment of track detectors”, in *Advanced Statistical Techniques in Particle Physics. Proceedings, Conference, Durham, UK, March 18-22, 2002*, pp. URL-STR(9). 2002. [arXiv:hep-ex/0208021](#). 179

ACKNOWLEDGMENTS

I would like to thank to all the people who made this journey not only possible but also a joyful and unforgettable experience. First of all I would like to thank my supervisor Dr. Hannes Jung for giving me the opportunity to join his group and for all the invaluable scientific and non-scientific advices. He is a great person who always tries to create a friendly and fruitful atmosphere.

I also want to express my gratitude to Dr. Benoît Roland for all his support and rich physics discussions which made this work possible. Special thanks go to Prof. Dr. Johannes Haller and Dr. Hannes Jung for accepting to be the referees of my thesis and disputation, and also thanks to Prof. Dr. Daniela Pfannkuche, Prof. Dr. Erika Garutti and Dr. Paolo Gunnellini for being part of my disputation committee.

Many thanks to all the current and former members of the QCD group at DESY - Hannes, Paolo, Samantha, Nastja, Ben, Hans, Tomasz, Vardan, Patrick, Ola, Daniela, Armando, Svenia, Engin, Radek and Jindrich - for their willingness to help at any moment and for all the laughs during the lunch time.

I want to thank to the DESY-Alignment group - Rainer, Matthias, Gregor and many more - for letting me collaborate in this important and interesting task.

I would also like to thank all the DESY colleagues that have become friends during these last years, Artur, Ksenia, Francesco, Özgür, Carmen, María, Eleni, Ivan, Andrej, James, Kelly, Gerrit, Till, Matteo, Marek and many more. Also thanks to my friends at CERN, Luis and Abraham, for those nice dinners each time I was there.

Very special thanks go to Jasone for all her support during the most difficult moments of the PhD and for all the countless hours she spent correcting thesis chapters to make them sound a little bit less robotic (proceso que recibió el nombre de “descavernicolización”).

Outside DESY I have met a great amount of wonderful people who have shown me the best of Hamburg and with whom I have enjoyed my free time to the maximum. I would like to thank them all: Alessandro, Maite, Federica, Nicola, Tobia, Antonio Farre, Mattia Rizzitek, Diana, Laura, Pedro, Mi Do, Francesco, Mattia Gallucci, Antonio Soto, Manuel, Stefan, Dimitri, Jonathan, Magda, Ann-Marie, James, Felix, Nikhil, Nicloe, Jhonalan, David, Ian, Phill, Thomas, Hendrik, Edina, Dirk, George, Bernd, Sven and many many more. Last but not least, to my flatmate and friend Ariane Forkel.

Finally I would like to express my deepest gratitude to my family for all the unconditional support they have always given to me.
