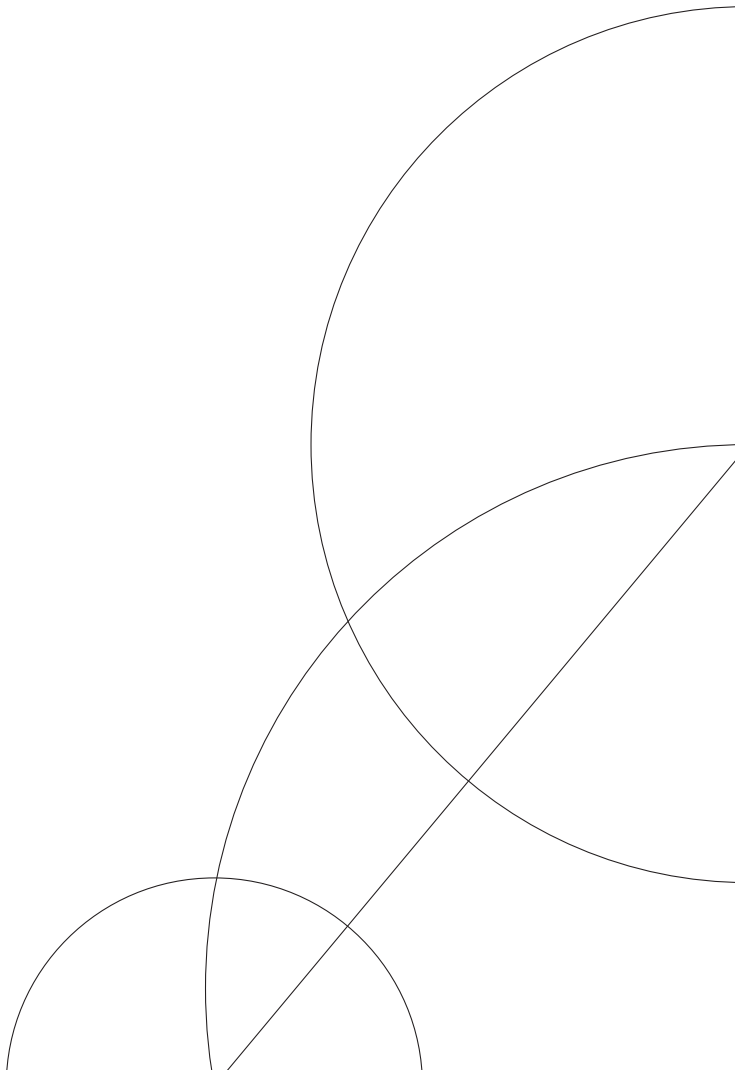
FACULTY OF SCIENCE
UNIVERSITY OF COPENHAGEN

Thesis submitted for the PhD degree (Doctor of Philosophy) in Physics to the PhD School of The Faculty of Science, University of Copenhagen

Charged-Particle Multiplicity Distributions over Wide Pseudorapidity Range in Proton-Proton and Proton-Lead Collisions with ALICE

Valentina Zaccolo
Niels Bohr Institute
University of Copenhagen



Academic Advisor:
Jens Jørgen Gaardhøje
Niels Bohr Institute
University of Copenhagen

April 2015

Contact Information:

Address:	Valentina Zaccolo Niels Bohr Institute Blegdamsvej 17 DK-2100 Copenhagen Ø
Email:	zaccolo@nbi.ku.dk
Supervisor:	Prof. Jens Jørgen Gaardhøje
Email:	gardhoje@nbi.ku.dk
Chair of Opponents:	Prof. Peter Henrik Hansen, Niels Bohr Institute
Opponent:	Prof. Larissa Bravina, University of Oslo
Opponent:	Dr. Flemming Videbæk, Brookhaven National Laboratory

Abstract

The charged–particle distribution ($P(N_{ch})$ as a function of N_{ch}), produced in high energy collisions between protons (pp) and between protons and heavy nucleus (pPb), depends on the fundamental processes, which lead to the formation of the observed particles. In particular, the so–called multiplicity distribution is sensitive to the number of collisions between quarks and gluons contained in the colliding systems.

In this thesis, data using the Forward Multiplicity Detector and the Silicon Pixel Detector of ALICE at CERN's Large Hadron Collider (LHC) are presented, for pp and pPb collisions. For the first time the multiplicity distributions are performed over such a wide kinematic range at the LHC (pseudorapidity coverage of $-3.4 < \eta < +5.1$) and at the highest energies ever, i.e. all available energies at the LHC's first run: at $\sqrt{s}$ =0.9, 2.76, 7 and 8 TeV for pp collisions and at $\sqrt{s}$ =5.02 TeV for pPb and Pbp collisions.

The results are compared, where possible, with the results of other LHC experiments and with theoretical Monte Carlo simulations (including PYTHIA, PHOJET and DMPJET) and with the IP–Glasma model. Moreover, data are compared using the Koba–Nielsen–Olesen (KNO) model, and it appears that this scaling is broken at energies from 0.9 GeV for pp collisions. Results for pPb collisions suggest that current models which include Color Glass Condensate effects, i.e. models based on the assumption of the saturation density of gluons, cannot reproduce the data.

Resumé (in Danish)

Fordelingen af antallet af elektrisk ladede partikler ($P(N_{ch})$ som funktion af N_{ch}), der produceres i energirige kollisioner mellem protoner (pp) og mellem protoner og tunge atomkerner (pPb), afhænger af de fundamentale processer der fører til dannelsen af observerbare partikler. I særdeleshed er den såkaldte multiplicitetsfordeling følsom overfor antallet af kollisioner mellem quarker og gluoner indeholdt i de kolliderende systemer.

I denne afhandling præsenteres data indsamlet med brug af Forward Multiplicity Detector og Silicon Pixel Detector i ALICE ved CERN's Large Hadron Collider (LHC) for pp og pPb kollisoner. For første gang er sådanne målinger udført over et bredt kinematisk område ved LHC (pseudorapiditet dækning $-3.4 < \eta < +5.1$) og ved de højeste energier nogen sinde, dvs. alle tilgængelige energier ved LHC's første forsøgsrunde: ved $\sqrt{s}$ =0.9, 2.76, 7 og 8 TeV for pp kollisioner og ved $\sqrt{s_{NN}}$ =5.02 TeV for pPb og Pbp kollisioner.

Resultaterne sammenlignes, hvor det er muligt, med resultaterne fra andre LHC eksperimenter, samt gængse teoretiske Monte Carlo simuleringer (herunder PYTHIA, PHOJET and DMPJET) og IP–Glasma model. Desuden sammenlignes data med Koba–Nielsen–Olesen (KNO) modellen, og det vises, at denne skalering er brudt ved energier fra 0.9 GeV for pp kollisioner. Resultater for pPb kollisioner antyder at gængse modeller der inkluderer Color Glass Condensate effekter, dvs. modeller der bygger på antagelsen af en mætningstæthed af gluoner, ikke kan reproducere data.

I am among those who think that science has great beauty. A scientist in his laboratory is not only a technician: he is also a child placed before natural phenomena which impress him like a fairy tale. We should not allow it to be believed that all scientific progress can be reduced to mechanisms, machines, gearings, even though such machinery also has its beauty.

Neither do I believe that the spirit of adventure runs any risk of disappearing in our world. If I see anything vital around me, it is precisely that spirit of adventure, which seems indestructible and is akin to curiosity.

– Marie Curie

as quoted in *Madame Curie: A Biography (1937)* by Eve Curie Labouisse

Contents

Preface

The first time I had the chance to go to CERN was for a quick visit as a young bachelor student. At that time, I was very attracted to particle physics, although I was more charmed by the theoretical mechanisms which drive and bind the fundamental particles. Due to this, after my bachelor I decided to explore the phenomenology of the fundamental particles. Caught by the beauty of the mechanisms, I developed a master thesis with the ATLAS experiment, which was somehow in between theory and experiment. At that point, I discovered that really putting the hands on the experimental data was the most exciting side of the work. I then decided to apply for a PhD in experimental physics and I was both enthusiastic about particle physics and by heavy–ion physics done at the LHC. I am really grateful that I was hired by Professor Jens Jørgen Gaarhøje as a PhD student in the heavy–ion group at Copenhagen. Heavy–ion physics is still a wild field, in some sense, in which experimental physicists sort of have *carte blanche*, and in which the theory is a full new world to discover.
The past three years have been extremely exciting, I had a chance to develop my analysis on charged–particle multiplicity distributions. Also, it was stimulating to be directly involved in the FMD work, doing regularly shifts at CERN for the first year or so, when we were going through Run I at LHC, and to do the QA of the detector for this period and the reconstruction of the data right after.

Now that my PhD is nearly over, I am presenting in this thesis all my work on charged–particle multiplicity distributions. The focus is on the pp collisions results, but I also devoted some time, in the end, to the study of the pPb distributions. I thought it was charming and challenging at the same time to look into distributions which were not yet explored at LHC. All the results for the available four pp energies in LHC Run I are in this thesis, together with the pPb results which were taken in the last period of data taking.
The pp results are almost complete, and a paper is in preparation, just a few more things need to be polished. After that, it will be possible for the ALICE collaboration to publish a paper on charged–multiplicity distributions in a unique pseudorapidity range, the widest available at LHC. I hope to have a paper draft and paper proposal to the collaboration in June 2015.

The first part, containing four chapters, presents the framework on which the charged–particle multiplicity distributions are built. In particular, in the first chapter the fun-

damental blocks of the high–energy particle physics are described, paying particular attention to soft diffraction, which is very important to describe the multiplicity distributions. In the second chapter, the heavy–ion collisions are explored. Although they are not strictly part of this thesis, in my opinion it was important to analyze the theories which are behind the ALICE collaboration physics program. Also, some of the new states of matter described in the second chapter are produced on pPb collisions and pp collisions, at high–energy. One of them is the Color Class Condensate. In the third chapter I go quickly through some of the very new discoveries of the heavy–ion program. The fourth chapter is, by all means, the most important of the four. In that one, the models which describe the charged–particle multiplicity distributions are presented. In particular the KNO scaling and the double NBD fits are described, together with the clan model, which provides a conceptual framework for understanding the multiplicity distributions.
The second part describes the tools which are needed to measure the charged–particle multiplicity distributions. Firstly, the hardware part is explored, namely the LHC and ALICE, paying particular attention to the two detectors used in this thesis (FMD and SPD). Secondly, the software is described, which means Monte Carlo simulations and the digitalization of the detectors response.
The last part is devoted to the description of the analysis procedure and the results. First the analysis procedure for pp data is explained, going through the selection of the events and to the unfolding procedure, in detail. Then the systematic uncertainties are described, and in chapter 9 all the results for pp are presented. The pPb analysis is very similar to the one carried on for pp. In chapter 10 analysis and results for pPb are shown.

I would like to thank everybody who helped me during my PhD. Special thanks go to my advisor and group leader of the HEHI group, Professor Jens Jørgen Gaardhøje, for many useful discussions in the past years and for a very careful proofreading of my thesis. I would like to thank very much all the members of the HEHI group, from Hans Bøggild, Børge Svane Nielsen and Ian Bearden, to Kristjan Gulbrandsen and Christian Holm Christensen. Many thanks also to Marek Chojnacki and Ante Bilandzic. Finally, I thank a lot also our former PhD student Alexander Colliander Hansen.
Special thanks go also to the ultra–diffractive physics working group members of ALICE. In particular I thank the current convenors, Jesus Guillermo Contreras Nuno and Orlando Villalobos Baillie, and the former convenor, Eugenio Scapparone, for the help in the last year or so, and the trust and priority they have given to my work. Also, I would specially like to thank the multiplicity physics analysis group convenor of ALICE, Anton Alkin, for countless discussions about this work, and for the great help he has given me, from suggestions about coding issues, to interpretation of the systematics. I would also like to thank Martin Poghosyan, Beomkyu Kim, Evgeny Kryshen and Cvetan Valeriev Cheskov for discussions along the way. I thank also Björn Schenke and Raju Venugopalan, for giving me a quick insight into their theory model.
Doing a PhD has been a lot of fun, but it has also been very demanding, above all in the last year. I would like to thank my family and especially Chiara, who has supported and helped me in the difficult times.

Happy Reading!

Part I

Multiplicity Distributions Framework

Chapter 1

High–Energy Physics

This first chapter is meant to be a brief introduction to the concepts necessary to understand the charged–particle multiplicity for proton–proton collisions. The chapter starts with the Standard Model section, which will shortly describe the building blocks of this particle physics model. After that, the proton–proton collision dynamics will be pictured, paying particular attention to the description of the soft processes, called also the underlying event. In fact, an important fraction of the measured events come from soft diffraction, described in detail in the last section.

1.1 The Standard Model

The Standard Model (SM) of Elementary Particles is a theory that describes the fundamental constituents of the universe (elementary particles) and the interaction between these constituents.

In the SM, all fundamental constituents of matter are spin $1/2$ particles. These particles are divided into two categories: leptons and quarks. Leptons are divided into neutrinos that are electrically neutral, and charged leptons with charge -1. Also quarks are divided into two categories, type *up* and type *down* quarks, with electric charge of $+2/3$ and $-1/3$ respectively. Every one of these particles has a corresponding anti–particle, with the same properties but opposite electric charge and quantum numbers. Each of these categories involves exactly three particles, and so it is possible to group them into three families, or generations, as shown in Table 1.1.

The quarks carry a color charge that is related to the strong interaction and they cannot be found isolated, since they have to form bound colorless states. A quark can be bound with an anti–quark forming a meson, or three (anti–)quarks can form a (anti–)baryon; these are grouped under the name of hadrons.

To introduce interactions between particles, it is necessary to impose that the Lagrangian of the system respects gauge local symmetries; this request leads to the introduction of spin 1 particles, called gauge bosons, which mediate the interactions. The

Fermions	Family			Charge
	I	II	III	
u–type quarks	u	c	t	$+2/3$
d–type quarks	d	s	b	$-1/3$
Neutrinos	ν_e	ν_μ	ν_τ	0
Leptons	e	μ	τ	-1

Table 1.1: Three fundamental fermions families.

symmetry group SU(2)$\otimes$ U(1) gives the electroweak interactions, while SU(3) gives the strong interaction. The procedure to introduce the electroweak interaction is allowed only if all the particles involved are massless, otherwise mass terms will not respect gauge symmetry. The masses are then reintroduced as the result of the interaction between these particles and the Higgs scalar field and with the spontaneous breaking of gauge symmetry.

In detail, on the basis of the free Dirac Lagrangian for fermions:

$$\mathcal{L} = \sum_f \overline{\Psi}_f (i\partial\!\!\!/ - m_f)\Psi_f \tag{1.1}$$

it is possible to introduce the interactions imposing the invariance of the Lagrangian respect to a local gauge symmetry group. The particles corresponding to the interaction fields are called gauge bosons.

The SU(2) invariance requires three vector bosons. A linear combination of them gives the $W^\pm$ charged bosons which are responsible of the charged currents, while the third combines with the U(1) invariance boson to produce both electromagnetic interactions (mediated by the photon γ) and the weak neutral interactions (mediated by the Z_0 boson).
The strong interactions are described by the Quantum Chromodynamics (QCD) theory and are generated by the SU(3) symmetry group with eight gauge bosons called gluons (Table 1.2.).

Interactions	Bosons	Electric Charge	Mass (GeV)
Electromagnetic	γ	0	0
Weak	$W^\pm$	± 1	80.399 ± 0.023
	Z	0	91.1876 ± 0.0021
Strong	g	0	0

Table 1.2: The interactions between particles together with the gauge bosons and their properties.

The Discovery of the Higgs Boson

The Higgs mechanism provides a way to break the electroweak symmetry and to give masses to the particles, without the introduction of an explicit mass term into the Lagrangian. As a consequence of this mechanism a scalar particle, the Higgs boson, is predicted, whose mass is not specified by the theory.

The Standard Model predictions have been well verified by the LEP, SLC and Tevatron experiments. On the contrary, some aspects of the theory remained obscure until the announcement on July 4th, 2012 of the observation at the Large Hadron Collider (LHC) of a narrow resonance with a mass of about 125 GeV. The broadcast and subsequent publication by ATLAS and CMS experiments provided an important new direction in the decades–long search for the SM Higgs boson [1, 2].

1.2 Protons and High–Energy Collisions

A high–energy collision is characterized by colliding particles, which have a total relativistic energy much higher than their rest mass. A typical event of two colliding hadrons, e.g. protons, can originate tens or hundreds of particles, from a variety of processes.

The proton is composed of partons, namely by three valence quarks, plus gluons and sea quarks. In particular, the parton content of the proton is characterized by the Parton Distribution Functions (PDFs) $f_i(x, Q^2)$. They describe the probability of finding a parton i that carries a fraction of the proton momentum x, called Bjorken x. The PDFs depend on the values of the Bjorken x and the momentum scale that characterizes the collision, Q^2, and are obtained by global fits to data measurements from different experiments. The PDF fits from ZEUS and H1 experiments at HERA can be seen in Figure 1.1. It is clear that at high energies the gluon and sea quark PDFs are dominant (left plot).

Figure 1.2 shows a schematic view of a proton–proton (pp) collision, or scattering. When two of the constituent partons collide head–on (red and green arrows), the scattering is called hard and is described using perturbative QCD. However, the majority of the processes in pp collisions are soft, and are often referred to as the underlying event, and namely are elastic and inelastic scatterings divided into single, double and non–diffractive events. These soft processes are much more difficult to describe and include also initial and final state radiation, as well as interactions of other remnant partons.

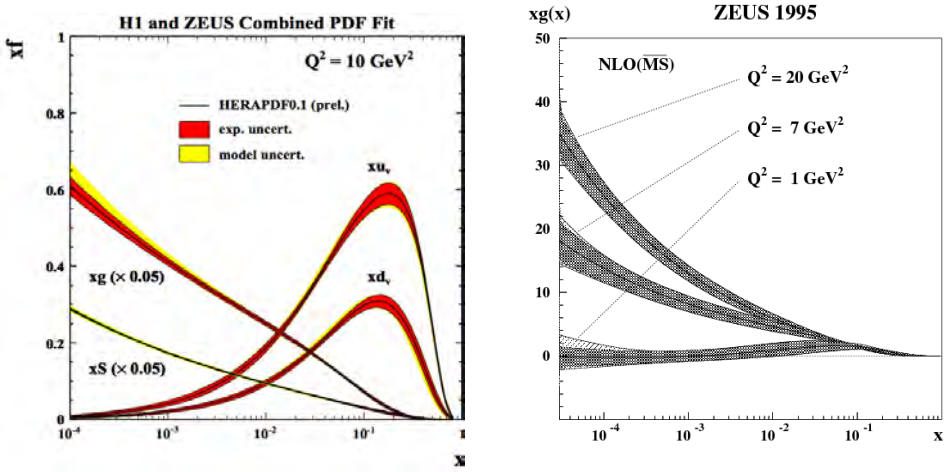


Figure 1.1: (left) PDFs from the analysis of the combined data set of ZEUS and H1, evolution of gluons, sea and valence quarks is shown (gluon and sea quarks distributions are divided by 20 to fit the figure) [3]; (right) The gluon momentum distribution as a function of x at fixed values of Q^2 from the ZEUS QCD fit [4].

If, e.g., the hard parton–level process considered involves a quark and an antiquark scattering into a final state Y, the leading order QCD prediction for the sub–process cross section is

$$\sigma(p(P_1) + p(P_2) \to Y + X) = \int_0^1 dx_1 \int_0^1 dx_2 \sum_f f_f(x_1) f_{\overline{f}}(x_2) \cdot \cdot \sigma(q_f(x_1 P) + \overline{q}_f(x_2 P) \to Y) \tag{1.2}$$

where the sum is over all the quark and antiquark species, and X is any hadronic final state.

In general, a pp collision can be described like:

Incoming Beams Two beams, made of partons, characterized by PDFs, move towards each other.

Initial–State Radiation The partons from each beam can shower before the interaction, i.e. they can branch like $q \to qg$, $q \to \gamma g$ or $g \to gg$.

Hard Interaction Occurs between two partons from the beam and produces outgoing partons. It is also possible to have more than one hard scattering in the same collision, that case in called multiple–parton interaction. The interacting partons carry only a fraction of the total beam energy, and some of the partons are called beam remnants, since they do not participate to the collision.

Semi–Hard and Soft Interactions (Underlying Event) Several semi–hard interactions occur between other partons. Soft processes like single, double and non–diffractive processes will be discussed in detail in the next section.

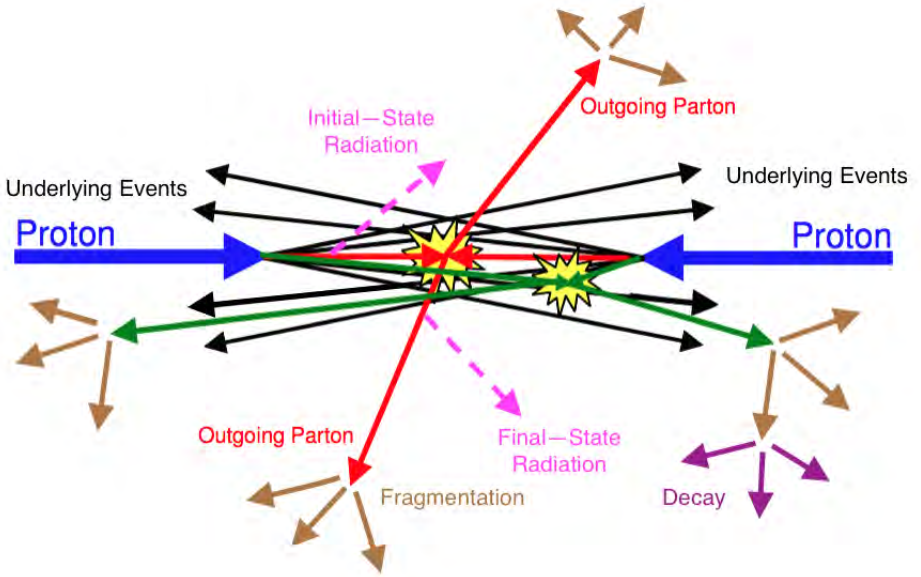


Figure 1.2: Schematic view of a high–energy proton–proton collision (Adapted from [5]).

Final–State Radiation Branchings of the outgoing partons. These processes become more relevant going to higher energies

Fragmentation Also called hadronization, it is the fragmentation process from outgoing quarks and gluons into colorless hadrons due to QCD confinement (see Section 2.1 and Figure 2.2 for details on confinement). This process is described phenomenologically, since it is not completely understood, and, if the energy is high enough, mesons and baryons are formed.

Decay If the hadrons produced during fragmentation are unstable, they can decay.

1.3 Soft Diffraction

As mentioned in the previous section, the hadronic processes can either be soft of hard. The latter are well described by perturbative QCD (pQCD). The soft sector instead, which dominates, cannot be described by pQCD, since the higher terms of the strong coupling constant (α_s) are large enough than the theory is not perturbative anymore, when the transverse momentum is small.

If the total cross section behaves asymptotically like a power of $\ln s$, then the particle and antiparticle cross sections become asymptotically equal, this concept was introduced by Pomeranchuk[1]. While Gribov predicted that, in order to ensure this behavior, the exchange of a Regge trajectory has to happen. The particles which lie on this trajectory are virtual and have the same internal numbers of the vacuum. The effective sum of all these particles is known as the Pomeron $\mathbb{P}$ (firstly introduced in 1961 [7]). In

[1]For several systems, the total cross section falls at low energies and rises slowly at high energies, but it is constant over a large range of s [6].

QCD, the $\mathbb{P}$ is a multiple gluon or a gluonball exchange, colorless and flavorless. For further details, a deep analysis and description of the Pomeron and the Regge trajectories can be found in [8].

Looking to the final states of the hadron scattering, it is possible to further divide the interactions into elastic and inelastic.

Elastic Interaction $p_1 + p_2 \rightarrow p_1' + p_2'$, both protons emerge intact after the collision and there are no other particles produced. Elastic scattering can happen with an exchange of a Pomeron $\mathbb{P}$.

Inelastic Interaction $p_1 + p_2 \rightarrow x_1 + ...$ Interaction in which there are other particles produced during the collision. Namely, the hadrons are color singlets, during the collision they can exchange a color octet gluon, becoming colored. Thereafter, the particles need to be colorless again, if the connected system gets separated. The final state can also be different from the initial one.

Inelastic collisions can be diffractive or non–diffractive. Non–diffractive scatterings (ND) are the dominant processes in pp collisions, more hadrons are produced as a consequence of an exchange of color charge. On the other side, diffractive events can be single, double or central diffractive. Diffraction is described in terms of exchange of a Pomeron using Regge's theory; it occurs when the Pomeron interacts with the proton and produces a system of particles, called the diffractive system. These multi–particle final states have the same internal quantum numbers of the colliding partons.

The case in which only one of the proton dissociates is called Single Diffractive (SD) ($p_1 + p_2 \rightarrow p_1' + X_2$ or $p_1 + p_2 \rightarrow X_1 + p_2'$) and is shown in Figure 1.3.

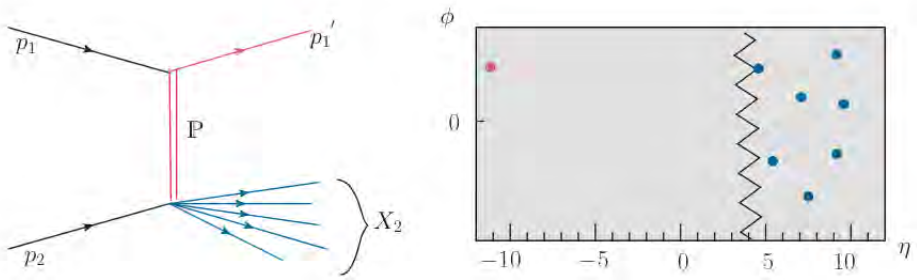


Figure 1.3: Single Diffractive diagram. In the $\eta - \phi$ diagram, the blue dots represent the proton dissociation, while the pink dot represents the non–dissociated proton [9].

In a Single Diffractive collision, a Pomeron is emitted from a proton p_i (Figure 1.4) with a transferred momentum $t = (p_i - p_i')^2$. Thereafter, the emitted Pomeron interacts with the other proton p_j and the produced system X is the diffractive system. Since the Pomeron has only a fractional momentum of the proton, X is scattered forward and, as pictured in Figure 1.3, there is a large rapidity gap (LRG) between the outgoing proton p_i' and X. This means that there is a large separation in rapidity in forward regions between the quasi–elastic scattered proton and the diffractive system, in which particles are not emitted.

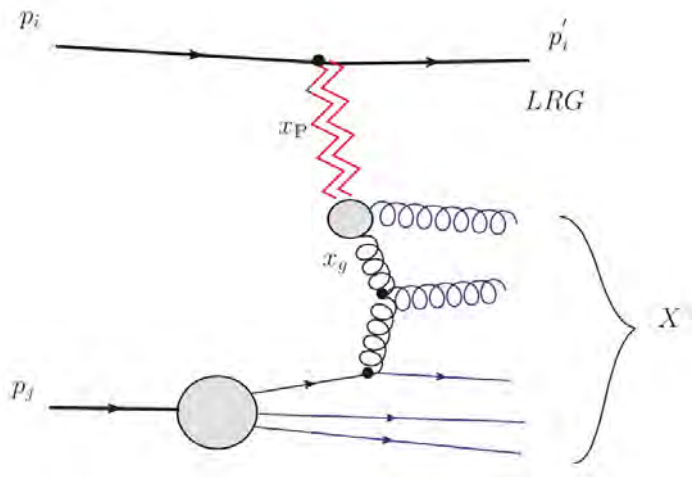


Figure 1.4: Exchange diagram for Single Diffraction. p_i and p_j are the incoming protons, while p_i' is the quasi–elastic outcoming proton (pink dot in Figure 1.3) and X is the diffractive system. x_{IP} is the fraction of the proton's momentum carried by the Pomeron, while x_g is the fraction of the Pomeron's momentum carried by its constituent gluon [9].

The Pomeron flux is defined as the probability that a Pomeron, with fractional momentum of the proton carried by the Pomeron x_{IP} and transferred momentum t, couples with the proton. A diffractive system X is characterized by its diffractive mass (the invariant X mass): $M_X^2 = x_{IP}s$, where $\sqrt{s}$ is the center of mass energy.

The case in which both the colliding protons dissociate, is called Double Diffractive process ($p_1 + p_2 \to X_1 + X_2$) Figure 1.5. There are also other possible scenarios, if two Pomerons are exchanged, Central Diffraction (CD) occurs. However, this process has a very low cross section at the LHC (around 1 mb [10]), and therefore, is not considered further in this thesis, together with other multiple Pomeron exchange processes.

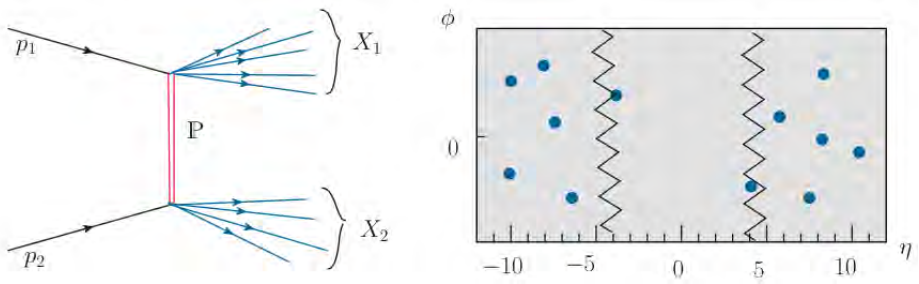


Figure 1.5: Double Diffractive diagram. In the $\eta - \phi$ diagram, the blue dots represent the proton dissociation [9].

The tot pp cross section is then the sum over several processes:

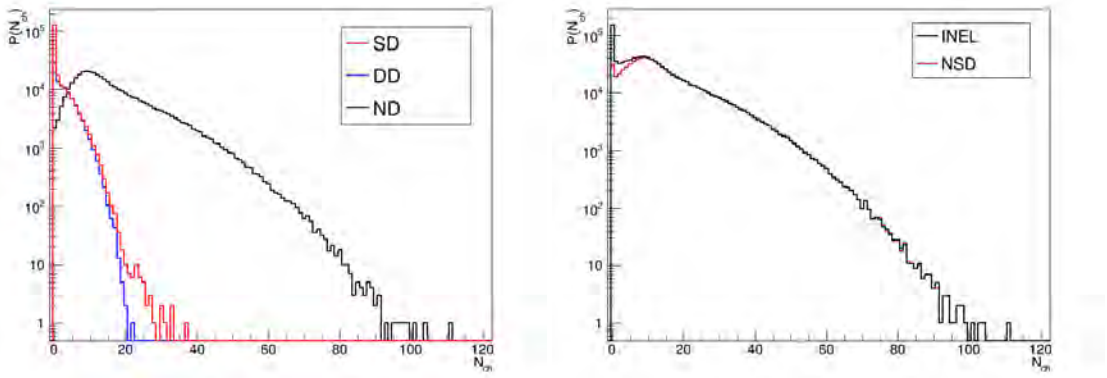
$$\begin{aligned} \sigma_{tot} &= \sigma_{el} + \sigma_{inel} = \sigma_{el} + (\sigma_D + \sigma_{ND}) = \\ &= \sigma_{el} + (\sigma_{SD} + \sigma_{DD} + \sigma_{CD} + \sigma_{mult}) + \sigma_{ND} \end{aligned} \tag{1.3}$$

Another exchange mediator called the Reggeon $\mathbb{R}$ is needed to reproduce diffractive

data and total cross sections correctly. Reggeons couple most likely to the valence quarks of a proton, while Pomerons couple mostly with sea quarks and gluons. Being those more abundant in high–energy collisions, the Pomeron is the predominant exchange mediator at LHC energies. In any case, both mediators are needed to properly describe σ_{tot}.

It is important to notice that ALICE has limited capabilities to distinguish among those different processes, therefore, it is necessary to correct for trigger inefficiencies. These trigger inefficiencies are very different for diffractive and non–diffractive systems and will be presented and discussed extensively in the last part of this thesis.

The probability distribution of the produced charged particles is very different between the diffraction processes. In Figure 1.6a, single, double and non–diffractive events are shown (derived with PYTHIA Monte Carlo generator [11], which will be described in Section 6.1). As expected the SD processes (in red) involve mainly the first bins of the distribution, peaking in zero and one charged particle. In fact, only one of the particles involved in the collision breaks up and produces the X diffractive system. The DD events are shown in blue, both particles produce a diffractive system, therefore, more charged particles are produced. The ND events are the majority (the cross section in more than doubled respect to the diffractive events one in the PYTHIA generator), and they produce more particles. It should be noted that the few events in the SD tail which are above the DD curve are statistical fluctuations. In fact, the DD events are more, since they generate form diffraction of two particles.



(a) (b)

Figure 1.6: Number of events, in logarithmic scale, obtained with the PYTHIA generator, for a pseudorapidity range $|\eta| < 2.4$. (a) SD, DD and ND processes shown; (b) INEL and NSD events. See text for detailed explanation.

In Figure 1.6b, distributions for two of the trigger classes used to present the results in this thesis are shown. Namely the inelastic class (INEL) in black, which includes all events: single, double plus non–diffractive, and the non–single diffractive class (NSD), involves double and non–diffractive events. The INEL class requires smallest corrections for inefficiencies (since it is the class which contains all the events triggered) re-

spect to the NSD class, but the latter is used to compare with previous experiments[2]. It is important to notice that the first bins in the NSD sample are dominated by the double–diffractive spectrum, consequences of this will be discussed later on, in the presented results.

[2]Triggering detectors in previous experiments were, in fact, designed to target these events.

Chapter 2

Heavy–Ion Physics

In the previous chapter I briefly brought you through the trip in which the experimental particle physicists have gone through in the last decades. If the discovery of the Higgs boson gave an answer to the last open question about the validity of the SM, we cannot say the same about the heavy–ion Standard Model which is still a hot topic.

Heavy–Ion collisions have been firstly considered in the late 60's and this subject was introduced at the workshop on "BEV/Nucleon Collisions of Heavy Ions – How and Why" at Bear Mountain, New York in the fall of 1974. In the beginning of the 80's, it was planned to build the Relativistic Heavy Ion Collider (RHIC) at Brookhaven National Laboratory, in order to investigate a new state of matter made of quarks and gluons, the Quark Gluon Plasma (QGP) which is thought to have prevailed in early Universe in the first 30 micro seconds.

The stages of the heavy–ion collision will be investigated in detail, with emphasis on the partonic forms of Quantum Chromodynamics (QCD) matter in the stages of the collision and of the theories that are used for the matter description.
Particular stress will be given to the forms of QCD matter that can be produced during a proton–proton (pp) or proton–nucleus (pA) collision, which are matter of study in this thesis.

2.1 Collisions and New States of Matter

The Quantum Cromodymamics (QCD) theory started to be formulated around 40 years ago, however, although the simplicity of its Lagrangian in equation (1.1), the QCD dynamics is a lot more complicated. It is especially difficult to describe matter at very high densities and with strong collective phenomena, like the one created in heavy–ion collisions. During these collisions, various forms of QCD matter are created, as can be seen in Figure 2.1.

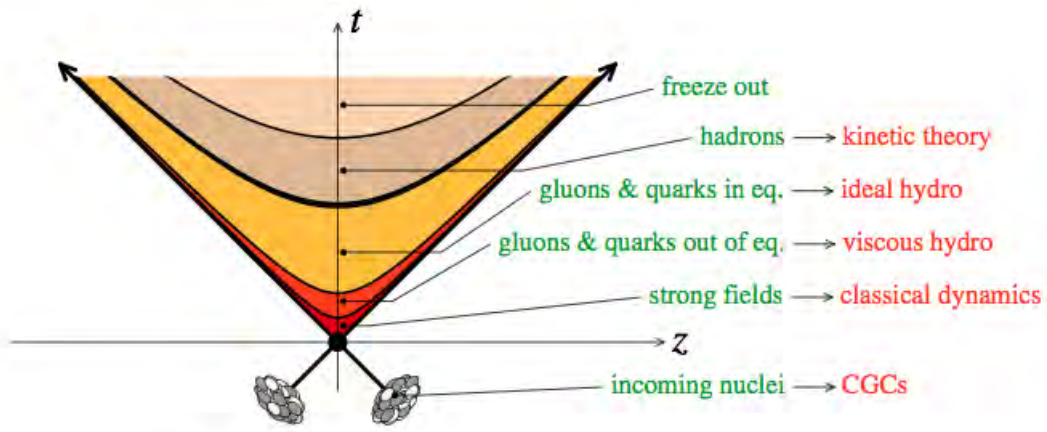


Figure 2.1: Various stages of the heavy–ion collisions as a function of time t and longitudinal collision axis z [12]

Incoming Nuclei Prior to the collision, the two incoming nuclei look like Lorentz–contracted pancakes mostly composed by gluons (as can be see in the left plot of Figure 1.1 in Section 1.2), which carry a small fraction of the longitudinal momentum of the nucleons and have an increasing density, those gluons carry a large transverse momentum. The gauge coupling that governs the interactions among the gluons is relatively weak. This kind of matter is called Color Glass Condensate (CGC) and will be described extensively in Section 2.2.

Hard Particles The collision takes place at proper time $\tau \equiv 0$, where $\tau \equiv \sqrt{t^2 - z^2}$ with t being the time coordinate (in the lab frame) and z the longitudinal coordinate. The first processes that occur are the ones that produce hard particles (hadronic jets, dilepton pairs, direct photons, up to heavy quarks) which are the most numerous compounds of the final measured spectrum.

Glasma At $\tau \backsim 0.2$ fm/c the gluons composing the CGC are liberated by the collision. Most of the multiplicity that is measured in the final state by the detectors, is produced by hadronization and fragmentation of these gluons. After being set free, these partons form a relatively dense medium. This non–equilibrium state of partonic matter is called Glasma, and will be analyzed in detail in Section 2.3.

Quark–Gluon Plasma In pp collisions, at that point, the produced partons evolve, without interacting, towards the final state. Conversely, in heavy–ion collisions, the partons liberated have been found to interact strongly with each other, at around $\tau \backsim 1$ fm/c, the matter approaches local thermal equilibrium. This strongly interacting fluid, almost perfect, which absorbs much of the energy of any fast parton which travels through is called Quark–Gluon Plasma (QGP) and will be described in Section 2.4. The detailed study of this phase, is the reason why heavy–ion collisions have been planned and are examined in depth.

Hot Hadron Gas At $\tau \backsim 10$ fm/c (in heavy–ion collisions) and critical temperature $T_c \simeq 150 \div 180$ MeV hadronization occurs, which means that quarks and gluons get confined inside colorless hadrons. This system is still dense, it preserves local thermal equilibrium and is called Hot Hadron Gas.

Frozen–Out Matter The hadrons stop interacting with each other at around $\tau \backsim 20$ fm/c, since the density becomes low. This transition phase from hot interacting liquid to a system of free particles is called freeze-out. After that moment, the hadrons propagate and direct towards the detectors. The freeze–out temperature in heavy–ion and pp collisions, is around $T_f \simeq 170$ MeV.

The different stages of the heavy–ion collision described above involve different hadron matter forms and, as a consequence, different effective theories are needed[1]. In QCD, the coupling constant runs with the momentum exchanged in the interaction, moreover, the coupling constant $\alpha_s(R) \equiv g^2/4\pi$ increases with the distance R or if the momentum transferred Q decreases, this property is called asymptotic freedom (see Figure 2.2).

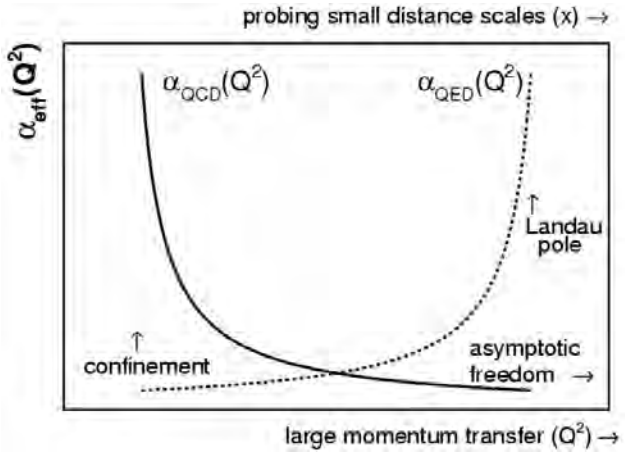


Figure 2.2: Running of the QED and QCD coupling contants. In QCD, asympotic freedom is reached for lange momentum transferred Q, while confinement occurs at small Q.

Based on the properties of QCD, it is possible to analyze the different effective theories for the stages above:

Incoming Nuclei and Glasma As mentioned above, the gauge coupling that governs the interactions among the gluons is relatively weak $\alpha_s \backsim 0.3$, the parton density is very high and perturbation theory (theory based on the expansion in powers of α_s) is valid. An effective theory aimed to describe hadronic matter under these conditions, called CGC theory, has been developed in the last 10 years.

Hard Particles The processes that produce hard particles can be described by kinetic theory, derived from QCD at weak coupling. This theory is still under development, because it cannot describe the rapid thermalization suggested by the RHIC and LHC data. Part of the hard particles with large transverse energies, called hard probes, do interact with the matter liberated after the hadronization process. By measuring the effects of these interactions, it is possible to obtain information on the matter that they traversed. In fact, semi–hard partons can

[1] An effective theory differs form a model and is guided by an underlying theory, like QCD, while a model simply provides a description of the data.

loose part of the transverse energy via interaction in the medium (jet quenching) suggesting that these interactions are quite strong.

Quark–Gluon Plasma This fluid is moderately strongly coupled, $\alpha_s \backsim 0.3 \div 0.4$. The thermal equilibrium is only local, in fact the system keeps expanding, and the theory which describes this flow is hydrodynamics. The strength of the coupling can be determined by a parameter called viscosity–over–entropy–density ratio η/s and the RHIC and LHC data suggest a small ratio, which identifies a strong coupling. Therefore, there is strong evidence that the QGP is a strongly coupled sQGP.

We shall return to some of the relevant experimental results in Chapter 3, while in the following sections the states of matter produce in a heavy–ion collision are analyzed in detail.

2.2 The Color Glass Condensate

This section is devoted to the study of the early stages of the hadron and heavy–ion collisions, characterized by high density QCD matter with high gluon occupation numbers.

The Color Glass Condensate (CGC) is a dense but weakly coupled system and its idea was first motivated by the ZEUS data collected at DESY [4], in Hamburg, already shown in Figure 1.1 to the right, in Section 1.2. The density of partons becomes very large and leads to a saturation of the partons at very high energies, or small values of Bjorken x variable, which identifies the fractional momentum carried by the gluon. When the saturation scale Q_s, corresponding to the density per unit transverse area, becomes large, the coupling constant becomes weak. Therefore, this was a hint that QCD at high energies could be described as a theory of partons which are weakly coupled and this system was called a Color Glass Condensate (CGC) since [13]:

Color: The gluons carry color charge of SU(3).

Glass: The glass is a disordered system which behaves like a solid on short time scales, or like a liquid on longer time scales. The partons in the CGC system are disordered and act in a similar way.

Condensate: The saturated gluons and their sources have high occupation numbers, and their color charges add coherently to each other (this concept will be further explained in Section 2.2, referring to the BFKL $Y-$evolution). A coherent quantum state with high occupancy can be described as a condensate.

The QCD partons

To analyze and explain the behavior of the particles produced in hadronic collisions, it is possible to start from the Deep Inelastic Scattering (DIS) of a lepton onto a hadronic target. In DIS, the empirical Bjorken variable x is defined as

$$x \equiv \frac{Q^2}{2(P \cdot q)} = \frac{Q^2}{s + Q^2 - M^2} \tag{2.1}$$

where Q^2 is the momentum transferred, and $s \equiv (P+q)^2$ is the invariant energy squared of the system. The x scales like Q^2/s. The two kinematical quantities, Q^2 and x, determine the transverse size of the parton $1/Q$, a quark in this DIS case, and the longitudinal momentum fraction x.
From equation (2.1) it is clear that small longitudinal momentum fraction x corresponds to the high–energy regime $s \gg Q^2$. The most interesting partons are the ones with small $x \ll 1$, since the bulk of the particles are produced at small $x \backsim 10^{-6}$ and hard momentum transferred $Q^2 \geq 10\,\mathrm{GeV}^2$ at RHIC and LHC.
It might be useful, at that point, to introduce a couple of kinematical variables, used also in the following. The rapidity is an alternative to the longitudinal momentum and for a particle with $4-$momentum $p^\mu = (E, \overline{p_\perp}, p_z)$ is defined as

$$y \equiv \frac{1}{2} \ln \frac{E + p_z}{E - p_z} \tag{2.2}$$

where $E^2 = m_\perp^2 + p_z^2$ and $m_\perp \equiv \sqrt{m^2 + p_\perp^2}$ is the transverse mass. A similar quantity, but more used in the experiments, is the pseudorapidity

$$\eta \equiv \frac{1}{2} \ln \frac{p + p_z}{p - p_z} = -\ln \tan \frac{\theta}{2} \tag{2.3}$$

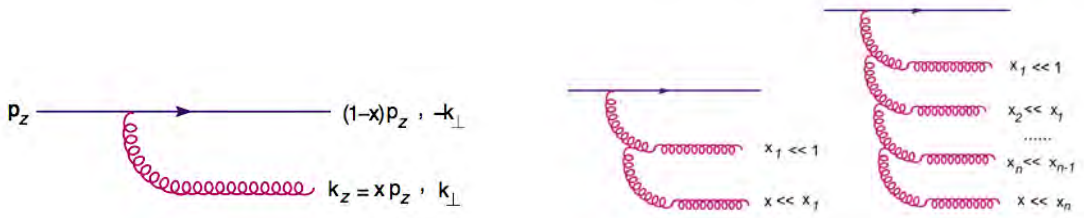
where $p = \mid \overline{p} \mid$ is the magnitude of the $3-$momentum vector. The pseudo–rapidity is directly related to the polar angle θ of the particle relative to the longitudinal axis, $\cos\theta = p_z/p$.

Small x physics

In perturbative QCD, the parton evolution proceeds via gluon bremsstrahlung, favoring the emission of short and collinear gluons, which carry only a small longitudinal momentum fraction $x \ll 1$ and a small transverse momentum $k_\perp$. The gluon emission carries $x = k_z/p_z$ of the longitudinal momentum of its parent quark or gluon, and is represented in Figure 2.3a. If $x \ll 1$ and to the first order in α_s, the differential probability for the gluon emission is [12]

$$dP_{Brem} \backsim \frac{\alpha_s(k_\perp^2)}{\pi^2} \frac{d^2k_\perp}{k_\perp^2} \frac{dx}{x} \tag{2.4}$$

As shown in the above equation, the parton branching is suppressed by a power of $\alpha_s(k_\perp^2)$, which is small when $k_\perp^2 \gg \Lambda_{QCD}^2$. This suppression is compensated by the phase–space available for the quark and gluon emission, if they have transverse momentum $k_\perp^2 \ll Q^2$ or longitudinal momentum fraction from 1 to x. Therefore, for large values of momentum transferred $Q^2 \gg \Lambda_{QCD}^2$ and/or small longitudinal momentum fraction $x \ll 1$, radiative processes of gluons and quarks at all orders must be re–summed (not suppressed anymore). The relevant equations are then computed as power series of α_s, starting with the leading order in equation (2.4).



(a) (b)

Figure 2.3: (a) Gluon bremsstrahlung to lowest order; (b) Gluon cascades [12].

There are two cases of evolution:

Q^2-**evolution** When Q^2 increases, the occupation number of the partons emitted in the transverse plane decreases. Since the decrease in the partonic area is stronger than the increase in their number, the partonic system becomes more and more dilute. This evolution is described by the DGLAP equation [14–16]. The partons occupy a smaller transverse area $1/Q^2$ and can be viewed as independent. The parton occupation number gives the proper measurement of the parton density in the hadron and the number of partons at a given value of x times the area occupied by one parton is the gluon distribution:

$$xg(x,Q^2) \equiv x\frac{dN_g}{dx} \tag{2.5}$$

$Y-$**evolution** When x decreases at fixed Q^2, mostly gluons are emitted with small longitudinal momentum fractions, but they occupy approximately the same transverse area as their parent gluons. This evolution is described by the BFKL equation [17–19]. The gluon occupation number, $n(x,Q^2) \simeq xg(x,Q^2)/Q^2R^2$ where R is the hadron radius in its rest frame, increases, and the gluonic system evolves towards increasing density. This evolution is fast and leads to a breakdown of the picture of independent partons.
The slow gluons that are in the end of the cascade, shown in Figure 2.3b, have a shorter life time than the preceding fast gluons, since the life time of a parton is proportional to its x (for a complete derivation see e.g. [12]). Therefore, a fast gluon with $x' \gg x$ behaves like a frozen color source emitting gluons at the scale x. Their color charges add coherently, giving rise to a large color charge density. Using a variable called shortly capital rapidity, identified by rapidity difference between the final gluon and the original valence quark: $Y \equiv \ln(1/x)$, the fluctuations in the color charge density increase rapidly with Y.

This growth is seen in the ZEUS data in Figure 1.1. The BKLF equation assumes that the radiated gluons do not interact with each other, which leads to a larger and larger gluon density. As soon as the gluon occupation number $n(x,Q^2) \ll 1$ the system is dilute; when $n(x,Q^2) \backsim \mathcal{O}(1)$ the gluons start overlapping. Finally, when $n(x,Q^2) \backsim \mathcal{O}(1/\alpha_s)$ non–linear effects (gluon saturation) start to be important and stop the growth of the gluon distribution. The gluon saturation occurs

and the saturation transverse–momentum $Q_s(x)$ is defined at:

$$n(x, Q^2 = Q_s^2(x)) \backsim \frac{1}{\alpha_s} \tag{2.6}$$

as

$$Q_s^2(x) = \alpha_s N_c \frac{xg(x, Q_s^2(x))}{\pi R^2} \simeq \alpha_s \frac{xg(x, Q_s^2(x))}{R^2} \tag{2.7}$$

where $\alpha_s N_c$ is the color charge squared of a single gluon and πR^2 is the hadron area in the transverse plane. The saturation momentum is simply the density of gluons per unit transverse area multiplied by the color charge squared, and it helps to divide the gluons into two regimes:

- Gluons with $k_\perp \leq Q_s(x)$ which are at saturation and therefore do not grow anymore when x decreases further;
- Gluons with $k_\perp \gg Q_s(x)$ are still in a dilute regime, but with rapidly increasing occupation numbers.

For a schematic view of the saturation and dilute regime, look at Figure 2.4.

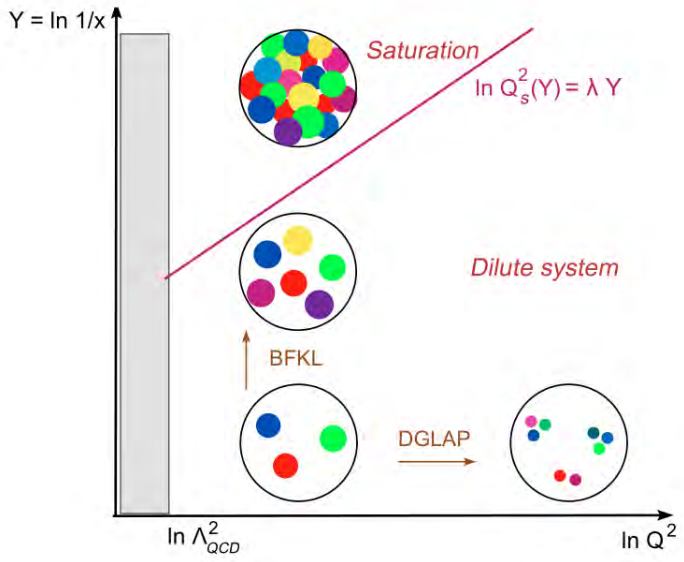


Figure 2.4: Phase diagram of parton evolution. Each colored blob represents a parton, and the straight line separates between dense and dilute regimes, derived for equation (2.8) [12].

The gluon occupancy is further amplified if a nucleus with atomic number $A \gg 1$ is considered instead of a proton. The saturation effects become important at larger values of x than for a proton, and the heavy–ion collisions are a rich soil for saturation studies, in fact:

$$Q_s^2(x, A) \simeq Q_0^2 A^{1/3} \left(\frac{x_0}{x}\right)^\lambda \tag{2.8}$$

with $\lambda = 0.2 \div 0.3$ called saturation exponent. The proton saturation scale Q_0^2 at the original value x_0 is non–perturbative and cannot be calculated with CGC effective theory, that parameter is fitted from the data.

The CGC Theory and Particle Formation

As mentioned in the beginning of this section, the partonic form of matter made with the saturated gluons is know as the Color Glass Condensate (CGC) and is weakly coupled. For sufficiently small x values with $Q_s^2 \gg \Lambda_{QCD}^2$ and $\alpha_s(Q_s^2) \ll 1$, the CGC can be described with perturbative QCD. However, when the saturation momentum is semi–hard like in LHC and RHIC phenomenology, Q_s supplants Λ_{QCD} as an infrared cutoff and the theory cannot be simple perturbative QCD. The CGC theory, therefore, must include non–linear physics of gluon saturation. This theory describes the gluons with a small x while the ones with a larger values $x' \gg x$ are integrated into perturbative QCD.

The CGC weight function is the probability to find a specific configuration with a certain color charge density. Every new step in this evolution involves the emission of a new gluon which, unlike the BFKL case, is allowed to interact with the strong color field of the source (gluons and valence quarks with large values of $x' \gg x$). This procedure generates an evolution equation described by the JIMWLK Hamiltonian (from Jalilian-Marian, Iancu, McLerran, Weigert, Leonidov and Kovner), in which, for $g\times$density $\ll 1$ (dilute regime), can be expanded to quadratic order of the density and describes the BFKL evolution, while, for $g\times$density ~ 1, contains non linear effects that prevent the emission of new gluons (gluon saturation). For details see e.g. [12].

In order to describe scattering cross–sections, the CGC theory needs to be combined with a factorization scheme. When probing dilute or dense parts of the hadron wavefunction, the factorization formula is different:

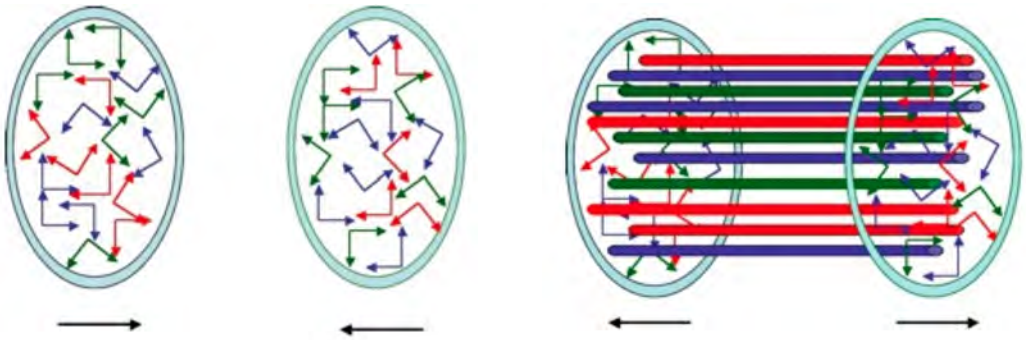
Dilute Regime In this case, one parton in one projectile interacts with one parton in the other projectile, to produce the final state. The hadronic cross-section is the convolution of the two parton distribution functions times the cross–section of the partonic subprocess.

Dense Regime When the collision probes the saturation effects of at least one of the incoming hadrons. The partons from one projectile scatter off a dense gluonic system, a CGC, in the other projectile undergoing multiple scattering. Each additional scattering implies a correction of order $\alpha_s n$ to the cross–section. It is possible to apply the eikonal approximation, i.e., in brief, assume that an energetic parton is not significantly deflected by the scattering and, therefore, can preserve a straight line trajectory through the collision. In a pA dilute dense scattering system, this means that the light quark produced from the proton, which scatters off the gluons from the nucleus, keeps a fixed transverse coordinate while crossing the nucleus.

The eikonal approximation can be applied both to pA dilute–dense scattering, and nucleus–nucleus (AA) collisions dense–dense scattering. In the pA collisions the particles produced by the collision evolve via fragmentation and hadronization towards the final hadrons observed in the detectors, without interacting. While for AA collisions the partons keep interacting influencing the final state spectrum and this makes it difficult to verify the CGC theory which involves the early–stage particle composition. In Chapter 3, some of the typical signatures of CGC are described.

2.3 The Glasma

In the previous Section, we have seen that prior to the collision the incoming nuclei are in the form of two Color Glass Condensates. Due to the high–energy kinematics, the chromo–electric and magnetic fields inside the CGCs have only transverse components, orthogonal respect to the collision axis, as can be seen in Figure 2.5a. The fields at different points are correlated, due to saturation and the high–energy evolution, and these correlations are restricted to a saturation disk with transverse area of radius $r \sim 1/Q_s$. In fact, saturation prohibits emission of gluons with $k_\perp \ll Q_s$ (gluons with $k_\perp \leq Q_s(x)$ are at saturation and therefore do not grow anymore when x decreases further). Therefore, there are domains separated by transverse distances $\Delta x_\perp \gg 1/Q_s$. Gluons arrange themselves in a way to minimize their mutual repulsion inside a saturation disk, this means that the disk has a zero overall color charge. Moreover, the gluons belong to the same parton cascade, therefore the correlations extend over a rapidity interval $\Delta Y \sim 1/\alpha_s$.



(a) (b)

Figure 2.5: (a) Field Configurations in the two CGCs' nuclei before the collision; (b) Color Flux Tubes, the Glasma in early stages of the collision [20].

A heavy–ion collision can be viewed as the scattering between two sheets of colored glass. It takes around $\Delta t \sim 1/Q_s$ to liberate and release a small$-x$ gluon with transverse momentum $k_\perp \sim Q_s$ at midrapidity (when $\eta \simeq 0$).

After the collision, the transverse fields on the two sheets are separated from each other and, in addition to those fields, there are longitudinal electric and magnetic fields which extend along the collision axis. This longitudinal components give rise to color flux tubes with the endpoints on the two sheets and transverse radius $r \sim 1/Q_s$, as it is pictured in Figure 2.5b. Right after the collision, the fields are strong, since they carry most of the energy of the original CGC. With increasing time, the system expands and so does the strength of the fields. When the magnitude of both transverse and longitudinal fields becomes of order one, these fields can be approximated to an incoherent superposition of particles. These particles can interact between each other, since their density is still high enough, and the interactions will then lead to a phase of local thermal equilibrium, called the QGP.

The intermediate form of matter, which is in non–equilibrium state between CGC and QGP is know as the Glasma and has as main feature the presence of color flux tubes with transverse area $\sim 1/Q_s^2$. The color fields in the flux tubes depend upon the proper time τ, but not on the rapidity η, therefore, are boost invariant.
The correlation between the fields in also present in pp collisions, for sufficiently high energy in a limited range in $k_\perp$ [21], an in–depth discussion about this specific case can be found in Section 3.1

2.4 The Quark Gluon Plasma

Heavy–Ion collisions have been firstly evaluated in the late 60's and this subject was introduced at the workshop on "BEV/nucleon collisions of heavy ions" at Bear Mountain, New York in the fall of 1974. In particular, as Lee and Wick speculated, new forms of matter could be explored by high–energy collision between very heavy nuclei [22]. At the same time, Collins and Perry, among others, stated that an ultra–dense form of matter with deconfined quarks and gluons could exists, as a consequence of the asymptotic freedom of QCD [23]. Right after Cabibbo ad Parisi defined the critical temperature T_c as the one of the phase transition from hadronic to quark matter [24]. This temperature was previously identified by Hagedorn as a limiting temperature above which hadronic matter could not exist [25]. Secondly, they also schematized the first phase diagram of hadronic matter, distinguishing between the confined and unconfined phases of the quarks. The first time in which this form of matter was called Quark Gluon Plasma (QGP) was in a paper by Shuryak in the late 70's, namely [26]. With all these premises, in the beginning of the 80's, it was planned to build the Relativistic Heavy Ion Collider (RHIC), in order to explore the properties of the QGP over the deconfinement point. In the following we will have an in–depth look to the QGP and its properties, also from the numerical point of view.

The Quark Gluon Plasma (QGP) is an ultra dense form of matter, with deconfined quarks and gluons. When the energy density of matter is of the order of that inside the proton, the quark and gluon degrees of freedom, which normally are confined into hadrons, are liberated, and the QGP is formed. Numerical lattice QCD computations (see e.g. [27]) show that, when the temperature reaches the critical temperature $T \approx T_c \simeq 150 \div 180$ MeV, the energy density rises by about an order of magnitude, as can be see in Figure 2.6.
The realistic case is the one with 2+1 flavors, namely 2 u and d light quarks with $m_{u,d}/T = 0.4$ and a heavier s quark with $m_s/T = 1$. Below T_c, specifically at low temperatures and small baryon densities, the medium is constituted by a low density gas of confined hadrons. As the temperature increases, the lightest hadrons, the pions, are thermally excited, therefore, the degrees of freedom of the system can be assumed to be three, which namely correspond to a dilute gas of 3 charged states of pions [28]. Above T_c, in the QGP regime, there are around 40 degrees of freedom. First they come from the eight colors of gluons and the two helicity degrees, plus 2 or 3 depending on the light flavors of quarks considered. Furthermore, the anti–quarks must be taken into account, with two spin states and three colors. Finally, the energy density, the pressure, and the temperature are proportional to the number of degrees of freedom. Therefore, it is easy to understand that the rapid change in energy density over a narrow range

of temperature is due to more than an order of magnitude increase in the degrees of magnitude [29].

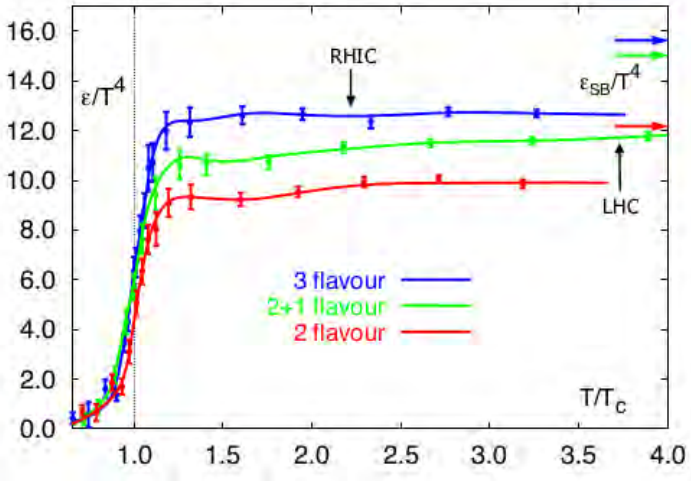


Figure 2.6: The energy density as a function of T/T_c scaled by T^4 from lattice QCD [27]. The realistic case is for 2+1 flavors. The estimate for LHC and RHIC temperature are included. At high temperature is expect that ϵ/T^4 will asymptotically approach the ideal gas values, as shown by the arrows.

Cabibbo and Parisi (in [24]) hypothesized that the phase space transition between confined and deconfined quarks was a statistical mechanical phase transition, i.e. there is a mathematical discontinuity in the energy density. In fact this is the case if the system can have both high temperature and high baryon density. The QCD phase space is shown in Figure 2.7, the various phases are shown as a function of temperature and net baryon density, μ_B .

The position of the critical point and the nature of the phase transition are still under study and discussion.

It is important to search for this form of matter, because it is a form of strongly interacting QCD matter at high temperature and baryon number density, and was present during the first few microseconds of the Big Bang. The QGP may occur naturally, nowadays, in supernovae and neutron stars, which is matter at high temperature and low baryon number density [30]. Therefore it can provide important information about the origin of mass for matter and about how quarks are confined into hadrons.

The matter produced in high–energy–heavy ion colliders, like RHIC and LHC, is a strongly coupled Quark Gluon Plasma. In fact after the collision of two ultra relativistic heavy ions, the matter is at very high energy density, and as time passes, the matter expands. Therefore the density of gluons decreases while the interaction strength increases, at this point there is enough time for the matter to thermalize and form a QGP.

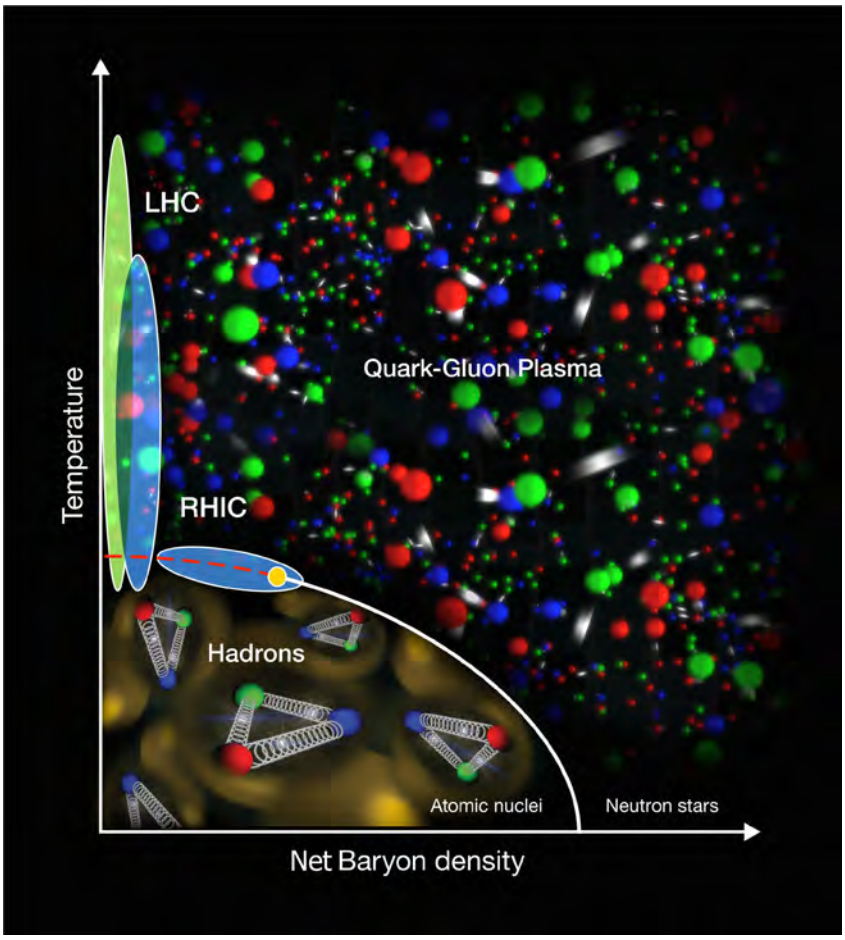


Figure 2.7: The phase space diagram of QCD. LHC and RHIC sit in the energy spot for exploring the transition between confined and deconfined matter. Image Credit: Brookhaven National Laboratory.

Chapter 3

State of the Art

In this chapter, the status of the pp and pA relativistic collisions is presented, the AA case is not included here, since it is not matter of study of this thesis. However, the focus will be on recent results relevant for the heavy-ion community. During the month of May 2014, there was the 24th Quark Matter Conference in Darmstadt, and the latest and more interesting results were presented. Therefore, part of the results shown are taken from that conference proceedings.

3.1 Proton–Proton Collisions

Before the high–energy runs at LHC, pp collisions were mostly treated by the heavy–ion community as reference measurements to use to compare to the heavy–ion results. In fact, important signatures of the quark–gluon plasma need a comparison of heavy-ion and pp collisions, therefore, it is essential to have a accurate understanding of pp collisions.

Since hadron collisions were only studied as a reference by the heavy–ion community, there was a big surprise when in September 2010 the CMS collaboration published an observation of near–side correlations in pp collisions [21]. The community started to question if this could be a signal of collectivity in small systems, but, soon thereafter it was suggested that the near–side ridge could be explained within the Color Glass Condensate framework [31].

The ridge–like structure observed by CMS in high multiplicity events with momenta range from $1-3$ GeV at 7 TeV can be seen in Figure 3.1. $\Delta\eta$ is the difference in pseudorapidity between the two particles, while $\Delta\phi$ is the difference in their azimuthal angle. The long–range structure, with pseudorapidity $2.0 < \Delta\eta < 4.8$ of two–particle correlation functions, was studied by CMS. The correlation is defined as

$$R(\Delta\eta, \Delta\phi) = \left\langle (\langle N \rangle - 1)\left(\frac{S(\Delta\eta, \Delta\phi)}{B(\Delta\eta, \Delta\phi)} - 1\right)\right\rangle \tag{3.1}$$

where S and B are the signal and random background distributions respectively, $\langle N \rangle$ is the number of tracks per event averaged over the multiplicity bin, and the final $R(\Delta\eta, \Delta\phi)$ is found by averaging over multiplicity bins. For details about the procedure refer to [21].

A similar ridge–like structure was observed before in AA collisions at RHIC. Long range correlations of particles occur at very early proper times (for details see [31]), nearly instantaneously after the collision or must preexist in the incoming projectiles.

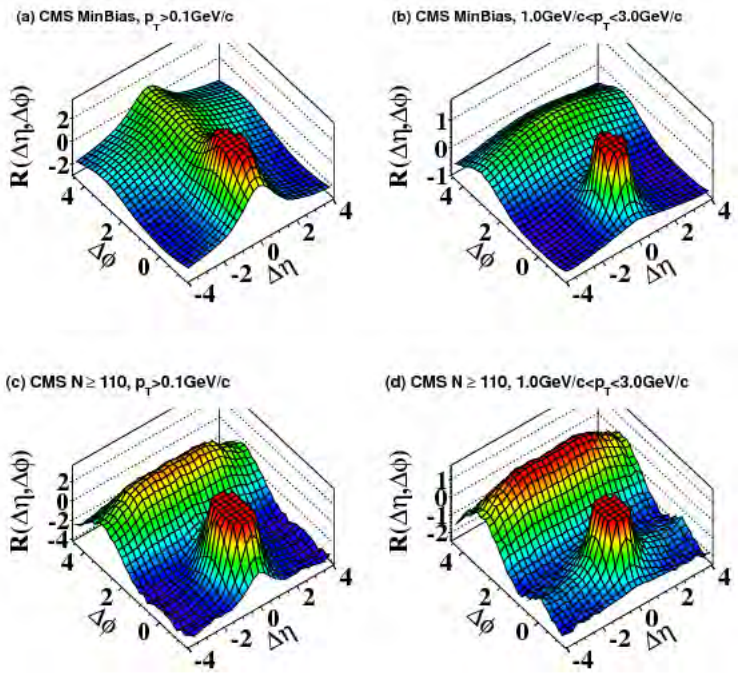


Figure 3.1: CMS 3–D display of two–particle correlation for minimum bias high multiplicity events [21].

The long range correlations arise as a consequence of the saturation of gluons with momenta $k_\perp \leq Q_s$ before the collision. This regime is semi–hard, $Q_s > \Lambda_{QCD}$ is large due to high energies and/or large nuclei. As discussed in Section 2.2, the dynamics of the saturated gluons can be described using weak gauge couplings as it is done in the CGC effective theory. In particular, as described in Section 2.3, the Glasma, strong longitudinal chromo–electric and magnetic boost–invariant fields, is predicted in high–energy collisions. The correlations are computed in the Glasma flux–tube picture and the results, shown in [31], exhibit the same features of the CMS results in Figure 3.1.
It can therefore be concluded that the physics of gluon saturation, like the ridge in pp collisions, revealed many surprises and will help understand the fundamental properties of strong interactions.

3.2 Proton–Lead Collisions

Proton–Lead (pPb) collisions were mostly considered, like the pp ones, to be a control and reference measurement for heavy–ion collisions. In fact some of the typical observables analyzed in heavy–ions do not show significant deviation from an incoherent superposition of AA collisions. However, some measurements hint at effects at low p_T, which might be of collective origin and are presented in the following.

Multiplicity Dependence of the Transverse Momentum

The Figure 3.2 shows the average transverse momentum as a function of charged–particle multiplicity in pp, pPb and $PbPb$ collisions from data collected respectively at energies in the center of mass of 7 TeV, 5.02 TeV and 2.76 TeV. In particular, in pp collisions (upper panel) a strong increase of $\langle p_T \rangle$ with N_{ch} is seen and can be understood as an effect of color reconnections between strings produced in multiple parton interactions [32].

The behavior observed in pPb (middle panel) is still not well understood. The only model which can describe the data within systematics is EPOS [33] and assumes collective flow in pPb [1]. Whether initial state effects are compatible with the data remains to be further studied.

Multiplicity Dependence of Identified–Particle Spectra

The interpretation of fluid Quark–Gluon Plasma is supported by the measurements of the elliptic flow, which represents a correlation between space and momentum variables. Different flow patterns are observed in AA collisions and are quantified in terms of a Fourier decomposition. The isotropic (zero–order) harmonic is called radial flow, while the elliptic flow, mentioned above, is the second order harmonic. In particular, radial effects have been observed also in pPb collisions at LHC.

The transverse momentum distributions of pions, kaons, protons and lambdas in pPb collision at 5.02 TeV (Figure 3.3) show an evolution with multiplicity similar to the pattern observed in AA collisions. The hardening of the spectra as a function of event multiplicity, which is stronger for heavier particles, is attributed to collective radial expansion of the system, in the heavy–ion collisions. The models, like EPOS, which incorporate final state effects, give a more accurate description on the data, like in the previous measurement shown in Figure 3.2.

Long Range Angular Correlations

Long range correlation structures were observed. In particular, Figure 3.4 shows the associated yield per trigger particle as a function of $\Delta\phi$ averaged over $|\Delta\eta| < 1.8$. The yield is for pairs of charged particles in pPb collisions at $\sqrt{s_{NN}} = 5.02$ TeV, for different event classes, compared to pp collisions at $\sqrt{s} = 2.76$ and 7 TeV. When subtracting

[1] In EPOS, collective flow is build both for pp systems (flow that comes from fast expansion of a small volume) and AA (slow expansion of a large volume). For pA, EPOS uses pp flow parameters for small number of nucleon pairs and AA flow for large systems instead [33].

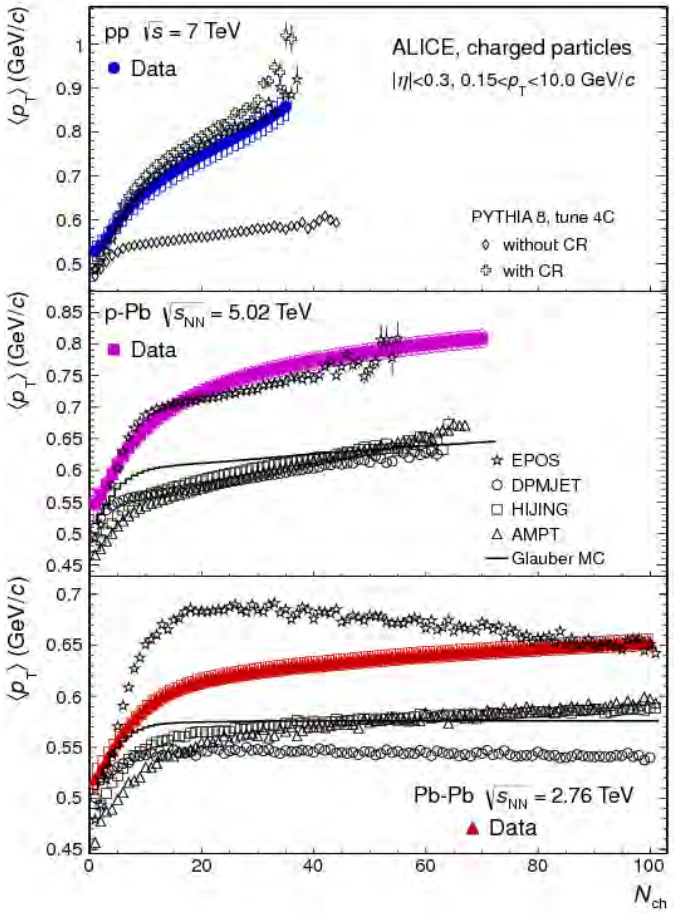


Figure 3.2: Average transverse momentum $\langle p_T \rangle$ as a function of charged–particle multiplicity N_{ch} measured in pp (upper panel), pPb (middle panel), and $PbPb$ (lower panel) collisions in comparison to model calculations [34].

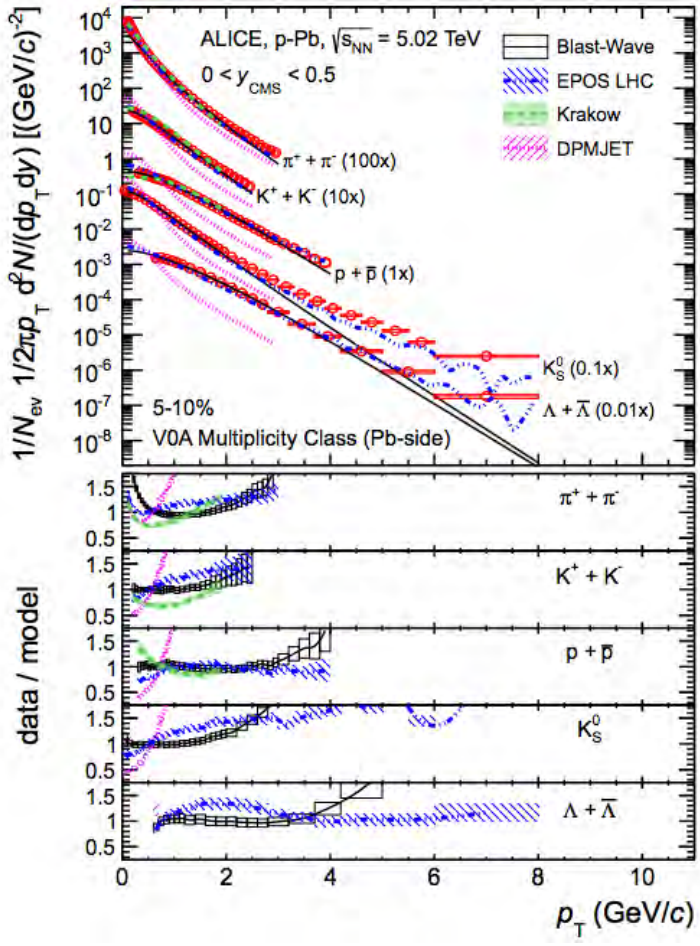


Figure 3.3: Pion, kaon, and proton transverse momentum distributions in the $5 - 10\%$ V0A multiplicity class, compared to several models [35].

the near–side ridge from the near side and away side symmetrically, the correlation shape becomes similar to those of pp collisions at 7 TeV (for the detailed procedure refer to [36]). Similar results have been observed by CMS [37] (only for the near–side) and by ATLAS [38].

The observation of a similar near–side and away–side ridge–like structure is consistent with CGC model calculations. The overall good agreement of this theory with data is strongly suggestive of gluon saturation, a detailed analysis is discussed in [39]. Certainly, further data from the LHC will clarify if these structures are only due to gluon saturation and initial state effects, or if there are other QCD effects which play a role. E.g., some models like [40] explained the long range angular correlations in the light of hydrodynamic expansion of the fireball in pPb collisions. Further pPb data will provide clarity and challenge either or both scenarios.

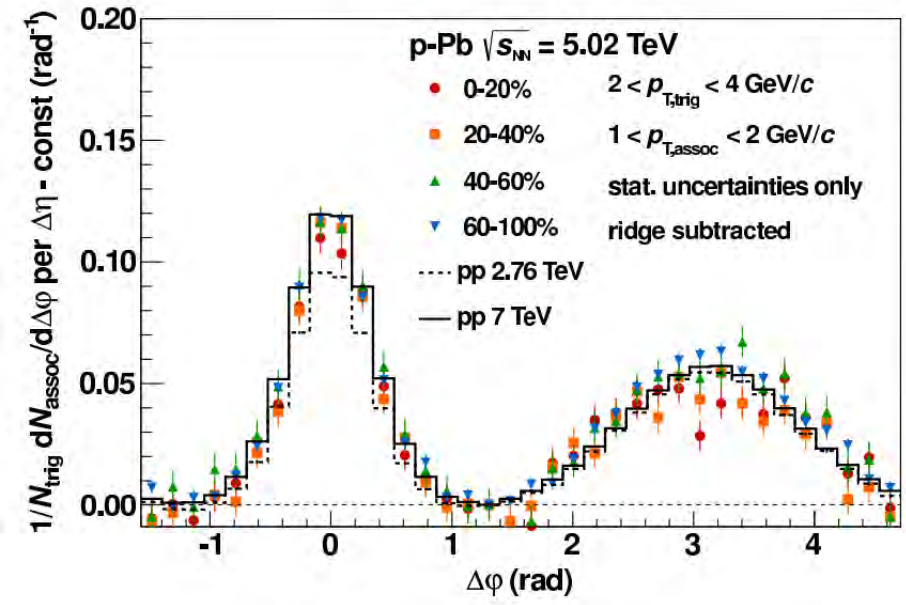


Figure 3.4: Associated yield per trigger particle as a function of $\Delta\phi$ averaged over $|\Delta\eta| < 1.8$ for pairs of charged particles in pPb collisions at $\sqrt{s_{NN}} = 5.02$ TeV for different event classes, compared to pp collisions at $\sqrt{s} = 2.76$ and 7 TeV [36].

J/ψ and $\psi(2S)$ suppression

The J/ψ suppression is one of the classical signals analyzed in heavy–ions, because the J/ψ, and its excited states, are expected to melt, when the temperature is raised well beyond the critical temperature. This is one of the predicted signals of the formation of the quark–gluon plasma [41]. The nuclear modification factors for J/ψ and its first exited state ψ' or $\psi(2S)$ identify the rates of the mesons measured over the expected production rate measurements in pp.

The inclusive J/ψ and $\psi(2S)$ production has been studied by the ALICE Collaboration [42, 43], both on the p–going (direction in which the proton comes from the A-side

of ALICE) and in the Pb–going direction (direction in which the proton comes from the C-side of ALICE, muon side). Looking at Figure 3.5 it is clear that the J/ψ is suppressed in the p–going direction (left panel), increasing with transverse momentum (event activity). Conversely, it is not suppressed in the Pb–going direction (right panel). What it is observed for $\psi(2S)$ is the same for the p–going direction (red squares), but $\psi(2S)$ is also suppressed in the Pb–going direction (blue squares).

Both quantities indicate that the $\psi(2S)$ is significantly more suppressed than the J/ψ, but initial state effects alone, like gluon saturation and shadowing, cannot account for the observed difference between the two. Final–state effects, like charmonium interaction with the final state hadronic medium created in pPb collisions, can affect the unexpected yields.

This section can be concluded stating that in the latest 2012-2013 LHC results for pPb collisions there were some major surprises, and some of the measurements shown here still need to find a theoretical explanation. According to the latest LHC schedule, it will be possible to take pPb data at $\sqrt{s_{NN}} = 5.1$ TeV again in 2016. Before that, the community looks forward to new developments in the theoretical reconciliation with some of the above mentioned findings.

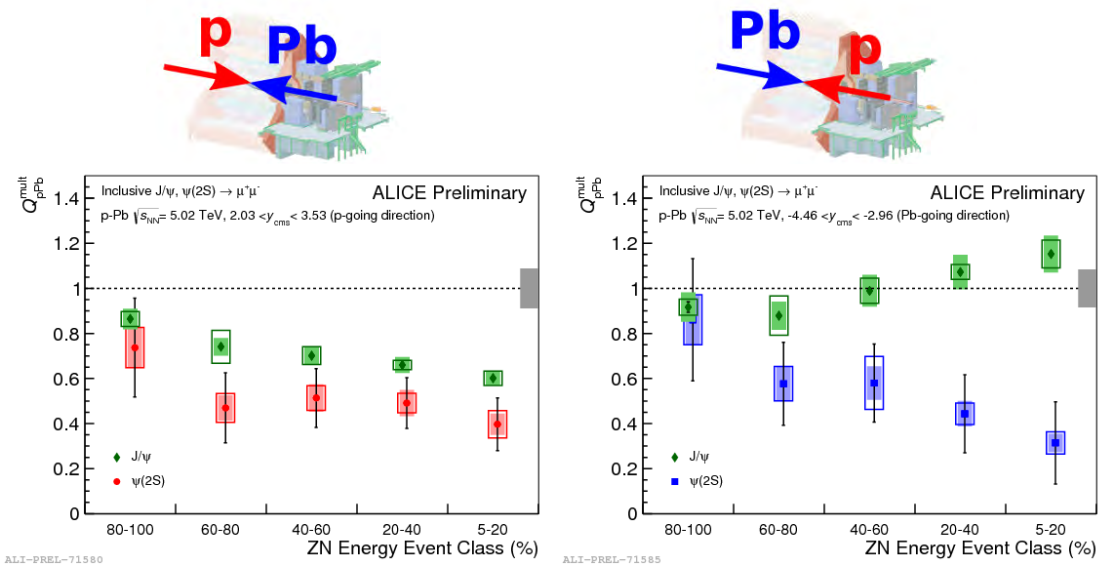


Figure 3.5: Nuclear modification factor of J/ψ and $\psi(2S)$ for pPb collisions, as a function of the event activity in p–going direction (left panel) and Pb–going direction (right panel) [44].

Chapter 4

Multiplicity of Charged-Particles

The charged–particle multiplicity is one of the basic measurements in high–energy collisions. In particular, this thesis aims to measure the charged–particle multiplicity distribution $P(N_{ch})$ as a function of N_{ch}, number of charged particles, in a specific pseudorapidity η range, as will be described later on. An other multiplicity quantity measured is the pseudorapidity density ($dN_{ch}/d\eta$ vs. N_{ch}, the integrated number of charged particles produced in a given pseudorapidity range), but will not be discussed in the following. Usually these global observables are measured at the beginning of the data taking, but the measurements can be performed again in a later stage, when the detector is known better and developing better techniques.

This chapter presents several models that try to describe the multiplicity distributions. In the last part of this thesis, the models will be compared to the actual measured data. After that, Section 4.2 will contain previous measurements of the charged–particles multiplicity distributions, performed by ALICE, CMS and, the older, UA5 Collaboration, as well as some from Tevatron. These measurements will be compared with the present work in the results chapter.

4.1 Models

There are several models that attempt to describe the multiplicity distributions and they will be presented in this section, in chronological order. The first step towards a successful understanding of the multiplicity distributions was done by Feynman in 1969 [45]. Moving to higher energies, deviations from those first models were observed, and the data were described using Negative Binomial Distributions.

Feynman Scaling

In this paper [45], Feynman states that the invariant cross–section of an inclusive process can be rewritten using a scaling function (Feynman scaling) which is independent of the collision energy $\sqrt{s}$. Although he does not give any mathematical proof, his rea-

soning is not difficult to follow. Using the total energy in the center of mass $s = 2W^2$, Feynman defined a new variable $x = p_z/W$, where p_z is the longitudinal momentum. To compare experiments at various values of W, the variables x and the transverse momentum Q can be used.

Feynman proceeds by noting that, in an inclusive experiment (in which there is search for special particles with a certain (x, Q), but also other particles can be produced), the fields that are radiated in the z direction, $\psi(z)$, get narrower as W rises. Therefore, the energy in this field is distributed like a delta function in z, which means that the field energy is uniform in momentum dp_z. Moreover, every particle of mass m carries an energy $E = \sqrt{m^2 + p_z^2 + Q^2}$.

If the field energy is distributed in fixed ratios, independently of energy W, among the particles, he concludes that the mean number of particles $\langle N \rangle$ is distributed as dp_z/E, for not too large x and at fixed Q. At this point, he defines the Feynman scaling functions $f_i(x, Q)$, with $\lim_{x\to 0} f_i(x, Q) = F_i(Q)$. The probability P_i of finding a particle of kind i and energy E, with transverse momentum Q and mass m_i is:

$$P_i(E) = f_i(x, Q)\frac{dp_z d^2Q}{E} \tag{4.1}$$

If $W \to \infty$:

$$\frac{dp_z}{E} = \frac{dp_z}{W} \cdot \frac{W}{E} = dx \cdot \frac{W}{\sqrt{m^2 + p_z^2 + Q^2}} = dx \cdot \frac{W}{W\sqrt{\frac{m^2+p_z^2+Q^2}{W^2}}} = \frac{dx}{x} \tag{4.2}$$

since $W \to \infty$, $W \gg m$ and $W \gg Q$.

The conclusion of Feynman's paper is then that $\langle N \rangle \propto dx/x \propto \ln W$, but we shall demonstrate this mathematically, since it is not intuitively understandable [1].

The Feynman scaling function is defined using the invariant cross section, dividing by the total cross section, since it has the meaning of a probability:

$$f_i(x, Q) = \frac{1}{\sigma_{tot}} E \frac{d^3\sigma}{dp_z d^2Q} \tag{4.3}$$

Assuming that, if $W \to \infty$, Q is limited, it is possible to factorize the scaling function

$$\widetilde{f}_i(x) = \int d^2Q f_i(x, Q) \tag{4.4}$$

At this point, the definition of probability of finding a particle, with a given energy, which is the average number of particles with a determined energy divided by the

[1] It is interesting and fascinating to notice how, until 50 years ago, the borders between philosophy and physics were not so well defined. In this paper, Feynman states "I have difficulty in writing this note because it is not in the nature of a deductive paper, but is the result of an induction. I am more sure of the conclusions than of any single argument which suggested them to me for they have an internal consistency which surprises me and exceeds the consistency of my deductive arguments which hinted at their existence." [45].

total number of particles, can be used

$$P_i(E) \equiv \frac{\langle N \rangle}{n_{tot}} \quad \Longrightarrow \quad \langle N \rangle = P_i(E) \cdot n_{tot} \tag{4.5}$$

Integrating equation (4.3) and using the equation in which the scaling functions were defined (4.1)

$$\int \frac{1}{\sigma_{tot}} E \frac{d^3\sigma}{dp_z d^2Q} \frac{d^3p}{E} = \int f_i(x,Q) \frac{d^3p}{E} = \int P_i(E) \cdot dn = \langle N \rangle \tag{4.6}$$

and with the factorization (4.4)

$$\langle N \rangle = \int f_i(x,Q) \frac{d^3p}{E} = \int \widetilde{f}_i(x) \frac{dp_z}{E} \tag{4.7}$$

Solving this integral will lead to the proof of the hypothesis ($\langle N \rangle \propto \ln W$).

The first step would be to rewrite it in x using the definition $x = p_z/W$:

$$\langle N \rangle = \int \widetilde{f}_i(x) \frac{dp_z}{E} = \int \widetilde{f}_i(x) \frac{dp_z}{W\sqrt{\frac{Q^2+m^2+p_z^2}{W^2}}} = \int_{-1}^{+1} \widetilde{f}_i(x) dx \frac{1}{\sqrt{\frac{Q^2+m^2}{W^2} + x^2}} \tag{4.8}$$

The function $\widetilde{f}_i(x)$ is even[2], therefore, it is possible to consider only positive x. Integrating by parts:

$$\begin{aligned} \langle N \rangle =& 2 \int_0^1 \widetilde{f}_i(x) dx \frac{1}{\sqrt{\frac{Q^2+m^2}{W^2} + x^2}} = \\ =& 2\widetilde{f}_i(x) \ln \left(2x + 2\sqrt{x^2 + \frac{Q^2+m^2}{W^2}} \right) \Bigg|_0^1 \\ & - 2 \int_0^1 \frac{\partial \widetilde{f}_i(x)}{\partial x} \ln \left(2x + 2\sqrt{x^2 + \frac{Q^2+m^2}{W^2}} \right) \end{aligned} \tag{4.9}$$

where the first term is:

$$= 2\widetilde{f}_i(1) \ln \left(2 + 2\sqrt{1 + \frac{Q^2+m^2}{W^2}} \right) - 2\widetilde{f}_i(0) \ln \left(2\sqrt{\frac{Q^2+m^2}{W^2}} \right) \tag{4.10}$$

Looking into the asymptotic behavior for the upper limit of the integral, for $W \to \infty$:

$$\begin{aligned} \lim_{W\to\infty} 2\widetilde{f}_i(1) \ln \left(2 + 2\sqrt{1 + \frac{Q^2+m^2}{W^2}} \right) \sim \lim_{W\to\infty} \ln \left(2 + 2\sqrt{1 + \frac{Q^2+m^2}{W^2}} \right) \sim \\ \sim \ln 4 \end{aligned} \tag{4.11}$$

[2]It is symmetric for collisions of identical particles, in other cases the integration can be performed in two different stages for positive and negative x, giving the same scaling result.

while for the lower limit

$$\lim_{W\to\infty} -2\widetilde{f}_i(0)\ln\left(2\sqrt{\frac{Q^2+m^2}{W^2}}\right) \sim \lim_{W\to\infty} 2\widetilde{f}_i(0)\left(2\ln W - 2\ln\sqrt{Q^2+m^2}\right) \sim \quad (4.12)$$
$$\sim \ln W$$

Therefore, for $W \to \infty$, the first part of the integral in equation (4.10) scales as $\ln W$.

The second is slightly more complicated, it is possible to use the Lebesgue's theorem, which allows to bring the limit inside the integral, if a function is monotonically increasing. This is the case, because $\widetilde{f}_i(x)$ is limited, due to the energy conservation, and so it is its derivative for only positive x. Therefore, with $\partial\widetilde{f}_i(x)/\partial x < C$:

$$\int_0^1 \frac{\partial\widetilde{f}_i(x)}{\partial x}\ln\left(2x+2\sqrt{x^2+\frac{Q^2+m^2}{W^2}}\right)dx < \quad (4.13)$$
$$< \int_0^1 C\ln\left(2x+2\sqrt{x^2+\frac{Q^2+m^2}{W^2}}\right)dx$$

and, for the limit in which $W \to \infty$, it scales like $\ln 4$ (as already calculated above).

It is possible then to conclude that

$$\langle N\rangle \propto \ln W \propto \ln\sqrt{s} \quad (4.14)$$

Koba–Nielsen–Olesen Scaling

Koba, Nielsen and Olesen demonstrated in 1972 [46] that Feynman's scaling law implies that the normalized multiplicity distribution keeps its form, independently of the energy of the beam and just scales as $\ln s$. This is found to be valid for energies up to 30 GeV.

The multiplicity distribution is defined as

$$P_n(s) = \frac{\sigma_n(s)}{\sigma_{tot}(s)} \quad (4.15)$$

where $\sigma_n(s)$ is the cross section for multiplicity n at center of mass energy $\sqrt{s}$. They demonstrated that it has the form

$$P_n(s) = \frac{1}{\langle n\rangle}\psi\left(\frac{n}{\langle n\rangle}\right) \quad (4.16)$$

where $\psi(n/\langle n\rangle)$ is an energy–independent function that can be different for different particles and reaction types. Basically, the starting point is the same integral just demonstrated for the Feynman scaling (4.10), but using q number of particles. Therefore, the scaling function is of the form:

$$f^{(q)}(x_1, Q_1; \ldots; x_q, Q_q) = \frac{1}{\sigma_{tot}} E_1 \cdot \ldots \cdot E_q \frac{d^3\sigma}{d^3p_1 \ldots d^3p_q} \quad (4.17)$$

and the integral looks like

$$\langle n(n-1)\dots(n-q-1)\rangle \equiv \sum_{n}^{\infty} P_n(s)n(n-1)\dots(n-q-1) = \tag{4.18}$$
$$= \int f^{(q)}(x_1,Q_1;\dots;x_q,Q_q)\frac{dx_1d^2Q_1}{\sqrt{x_1^2+\frac{Q_1^2+m^2}{W^2}}}\dots\frac{dx_qd^2Q_q}{\sqrt{x_q^2+\frac{Q_q^2+m^2}{W^2}}}$$

Making use of the factorization of the scaling function, like in equation (4.4), extending it to q number of particles, and performing the integral by parts in a recursive way (integrate once by parts and then apply it again to the resulting integral), it is obtained that:

$$\langle n(n-1)\dots(n-q-1)\rangle = \widetilde{f}^{(q)}(0;\dots;0)(\ln s)^q + \mathcal{O}((\ln s)^{q-1}) \tag{4.19}$$

Using now the fact that the average of any $\mathcal{O}(q-1)$ polynomial goes like $\mathcal{O}((\ln s)^{q-1})$ [3]:

$$\langle n^q\rangle = \widetilde{f}^{(q)}(0;\dots;0)(\ln s)^q + \mathcal{O}((\ln s)^{q-1}) \tag{4.20}$$

therefore

$$\sum_n P_n(s)n^q \approx \int_0^{\infty} n^qP_n(s)dn = \widetilde{f}^{(q)}(0;\dots;0)(\ln s)^q + \mathcal{O}((\ln s)^{q-1}) \tag{4.21}$$

Dividing both terms of the equation by $(\ln s)^q[\widetilde{f}^{(1)}(0)]^q$, the scaling result is obtained:

$$P_n(s) = \frac{1}{\widetilde{f}^{(1)}(0)\ln s}\psi\left(\frac{n}{\widetilde{f}^{(1)}(0)\ln s}\right) + \mathcal{O}\left(\frac{1}{(\ln s)^2}\right) \tag{4.22}$$

and since $\langle n\rangle \approx \widetilde{f}^{(1)}(0)\ln s$ (Feynman's scaling):

$$P_n(s) = \frac{1}{\langle n\rangle}\psi\left(\frac{n}{\langle n\rangle}\right) + \mathcal{O}\left(\frac{1}{\langle n\rangle^2}\right) \tag{4.23}$$

In the Koba, Nielsen and Olesen paper [46], there are only two assumptions made to derive the scaling function in the equation above: the Feynman scaling and the fact that the moments of the distribution c_q:

$$c_q = \frac{\widetilde{f}^{(q)}(0;\dots;0)}{[\widetilde{f}^{(1)}(0)]^q} = \frac{\langle n^q\rangle}{\langle n\rangle^q} \tag{4.24}$$

determine the distribution uniquely. Therefore, by assumption, all the moments together do define uniquely the distribution.

[3] Considering $q=1$, the Feynman's scaling is demonstrated: $\langle n\rangle \propto (\ln s)^1$; if $q=2$, $\langle n(n-1)\rangle \propto (\ln s)^2 + (\ln s)^1$ and so on so forth.

Negative Binomial Distributions

The KNO scaling is not valid as a description of multiplicity distributions above 30 GeV, in fact, it was found by the UA5 experiment in 1985, that the Negative Binomial Distributions (NBD), instead, can describe the data at 540 GeV [47, 48]. If one of the parameters of the NBD is kept free, the KNO scaling is obtained.

The NBD is the probability distribution of the successes before a specified number of failures k when Bernoulli trials are performed [4], and it is defined as

$$P(n;p;k) = \left(\begin{array}{c} n+k-1 \\ n \end{array} \right) (1-p)^k p^n \tag{4.25}$$

where the binomial coefficient can be rewritten like

$$\left(\begin{array}{c} n+k-1 \\ n \end{array} \right) = \frac{(n+k-1)!}{n!(k-1)!} = \frac{(n+k-1)(n+k-2)\cdots(k)}{n!} \tag{4.26}$$

In particular, the probability for every specific sequence of n successes and k failures is $(1-p)^k p^n$, because the outcomes of the $n+k$ trials are independent. The k^{th} failure comes in the end, therefore, the n trials with successes are free to choose out of the remaining $n+k-1$ trials. The above binomial coefficient gives the number of all these sequences of length $n+k-1$. If $k=1$ the NBD is a geometrical distribution, while if $k \to \infty$, it is the Poisson distribution.

The UA5 observed a violation of the KNO scaling [47, 48] in their multiplicity distributions. Namely the scaling implies that, if P_n is the probability of finding n particles in the final state of the interaction, and $\langle n \rangle$ is the mean multiplicity, $\langle n \rangle P_n = \psi(n/\langle n \rangle)$ is energy independent at very high energy. UA5 found, instead, that the distributions were following a NBD of the form

$$\begin{aligned} P(n;\langle n \rangle;k) &= \left(\begin{array}{c} n+k-1 \\ k-1 \end{array} \right) \left(\frac{1}{1+\langle n \rangle/k} \right)^k \left(\frac{\langle n \rangle/k}{1+\langle n \rangle/k} \right)^n = \\ &= \frac{k(k+1)\cdots(k+n-1)}{n!} \frac{\langle n \rangle k^k}{(\langle n \rangle + k)^{n+k}} \end{aligned} \tag{4.27}$$

where the k parameter affects the shape. If k is constant and does not depend on the energy, the KNO scaling is valid.

In the past years, several trials to understand why the NBD approximates the multiplicity distribution have been done, e.g. by Giovannini and Van Hove in [49], right after UA5's publication. They tried to interpret the NBD behavior in terms of a simple form of cascade process, which leads to the concept of clusters, that will be explained in the following paragraphs. Other models, like e.g. [50], aimed to explain the NBD assuming stimulated emission of identical bosons by identical cells, but those model would produce an integer value of k, which is not in agreement with the UA5 results.

[4] A Bernoulli trial is a random experiment with only success or failure possibility as an outcome; moreover, the probability of success is the same every time the experiment is conducted.

The cascade processes, instead, allow to redefine the parameter k, which will be continuously varying.

A first step in characterizing the multiplicity distribution could be to relate by a recursive relation $P(n)$ and $P(n+1)$, this can be simply understood thinking that a multiplicity of $n+1$ collision is related to $n+1$ collisions of multiplicity n (collisions obtained removing one of the $n+1$ particles). Therefore, the relation between the two distributions is between $P(n)$ and $(n+1)P(n+1)$, and this ratio is defined:

$$g(n) = \frac{(n+1)P(n+1)}{P(n)} \tag{4.28}$$

at this point with $a = \langle n \rangle k/(\langle n \rangle + k)$ and $b = \langle n \rangle/(\langle n \rangle + k)$, it is obtained that the NBD has a linear dependence of $g(n)$ on n:

$$g(n) = a + bn \tag{4.29}$$

A linear ratio $g(n)$ is expected for cascade effects, cascading happens when an already produced particle emits additional particles, which change the momentum and quantum numbers respect to the original particle and refers to e.g. fragmentation and decay. All the particles which originate from the same one form a cluster, and the original particle is called ancestor. In the model considered by [49], it is assumed that clusters are produced independently.

If the case of a multiplicity distribution in a limited phase space (limited pseudorapidity interval) is considered, each cluster can have all, some or none of its particles inside the interval. If n particles are included, the $(n+1)^{th}$ particle can be originally produced (i.e. be an ancestor), or be emitted by a particle outside the interval, therefore, it forms an extra one–particle cluster and then be included in the a term in eq. (4.29). On the other side, the $(n+1)^{th}$ particle can be emitted by a preexisting cluster, so that the average effect of this cascade emission can be assumed to be proportional to the number n of particles, and already present in the b coefficient in (4.29).

Considering a fixed center of mass energy, and increasing the pseudorapidity interval, for a given n, it will be more probable that the $(n+1)^{th}$ particle is an extra one–particle cluster. The consequence is that there will be an increase in $a/bn = k/n$, which implies and increase in k, since n is the given fixed parameter.

The multiplicity distribution of a single cluster $P_c(n_c)$ is characterized by $P_c(0) = 0$, because by definition a cluster contains at least one particle, moreover:

$$g(n) \equiv (n_c+1)P_c(n_c+1)/P_c(n_c) = pn_c \quad \text{with} \quad n_c \geq 1 \tag{4.30}$$

where p is a constant which underlines the fact that the emission is $\propto n_c$ particles already inside the cluster. At this point, iterating:

$$P_c(n_c) = P_c(1)\frac{p^{n_c-1}}{n_c} \tag{4.31}$$

Using the fact that for independent emission of particles, the distribution is Poissonian [5] $\widetilde{P}(N) \propto \langle N\rangle^N/N!$. Combining the Poissonian with the above equation, the overall multiplicity distribution is obtained:

$$P(n) \propto \sum_{N=1}^{n}\left[\left(\frac{\langle N\rangle^N}{N!}\right)\sum^{*} P_c(n_1)\dots P_c(n_N)\right] = \tag{4.32}$$
$$= p^n \sum_{N=1}^{n}\left[\left(\frac{\langle N\rangle^N P_c(1)^N}{p^N N!}\right)\sum^{*}(n_1\dots n_N)^{-1}\right]$$

where $\sum^*$ is the sum over all $n = n_1 \dots n_N$ with $n_i \geq 0$ integers.

With the Taylor expansion of an exponential:

$$P(n) \propto \frac{a(a+b)\cdots\big(a+b(n-1)\big)}{n!} \tag{4.33}$$

where the parameters $a = \langle N\rangle P_c(1)$ and $b = p$ in agreement with their physical meaning expressed in the text above.

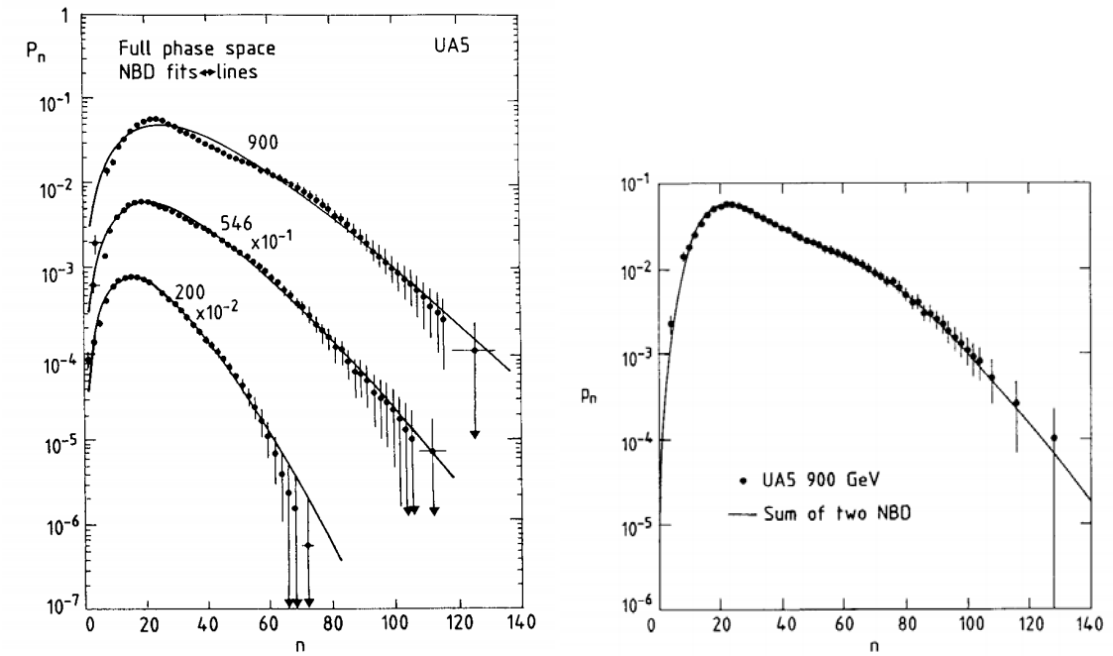
Sum of Soft and Semi–Hard NBD Components

As underlined in the previous section, the UA5 Collaboration found that the multiplicity distributions for the NSD class at 540 GeV was described by the NBD [47, 48]. However, as it shall be seen, deviations from the data were observed for data at 900 GeV, and the deviations brought to a two component model described by the combination of two NBDs. Some fits of various functions to the UA5 multiplicity distributions were done in [51]. In Figure 4.1a, the multiplicities for 200, 540 and 900 GeV are fitted to a single NBD, using formula (4.28). It can be seen with the naked eye that the curve does not fit properly the data for 900 GeV, in fact the reduced $\chi^2 > 2$, in that case. In Figure 4.1b instead, the fit is done using the sum of two NBDs.

In the case of two independent particle production mechanisms, the two–fold NBD contribution would have been a convolution on two NBDs. In that case, interference terms are therefore present. However, in the case in which the events belong either to the semi–hard or to the soft spectrum, i.e. have different physical origins, the resulting distribution is a weighted sum between them:

$$P_n^{total}(w, \langle n_{soft}\rangle, k_{soft}, \langle n_{semi-hard}\rangle, k_{semi-hard}) = \tag{4.34}$$
$$= (1-w)P_{NBD}^{soft}(n; \langle n_{soft}\rangle, k_{soft}) + wP_{NBD}^{semi-hard}(n; \langle n_{semi-hard}\rangle, k_{semi-hard})$$

[5] In a Poissonian distribution, $P(n) \propto a^n/n!$, where $a = \langle n\rangle$. Of course in this case $g(n) = a$ and the production of an additional particle is independent form the already present particles (independent emission).



(a) (b)

Figure 4.1: (a) NBD fit to the UA5 data; (b) Fit using the sum of two NBDs, explanation in the text [51].

There are five free parameters in this fit model, and the final reduced χ^2 to the fit shown in Figure 4.1b is 1.16 [51]. The value of the parameter w is similar to the events defined as events with mini–jets by the UA1 Collaboration (groups of particles with total transverse energy larger then 5 GeV [52]). Moreover, $\langle n_{semi-hard} \rangle \approx 2\langle n_{soft} \rangle$, like UA1 results. From the fit computed by [51], it turns out that the soft component is compatible with obeying KNO scaling. On the contrary, the semi–hard component grows in relative strength starting at around 100 GeV and becomes wider with energy.

In this work from Giovannini and Ugoccioni [53], the possibility of a new third component was explored in addition to the soft and semi–hard components analyzed above. Recalling the clusters model presented in the previous section, it turns out that the average number of clusters for the semi–hard component becomes smaller as the center of mass energy increases, while the corresponding average number of particles per cluster becomes much larger, favoring aggregation within clusters. This behavior may indicate a class of events with high particle density. The eventual contribution from this third (harder) component might happen already at 14 TeV, and needs experimental proof. The total charged–particle multiplicity distribution would then have the following expression:

$$\begin{aligned} P_n^{total} =& \alpha_{soft} P_{NBD}^{soft} + \alpha_{semi-hard} P_{NBD}^{semi-hard} + \\ &+ (1 - \alpha_{soft} - \alpha_{semi-hard}) P_{NBD}^{third} \end{aligned} \tag{4.35}$$

The third NB component is found, from MC simulations with PYTHIA, to be harder than the semi–hard component. Moreover, from [54], the energy density has been com-

puted in order to compare predictions from pp collisions with heavy–ion collisions. The result shows, surprisingly, that for the third component the energy density of collisions at 14 TeV is comparable to the one for $AuAu$ collisions at RHIC.

4.2 Previous Measurements

In this section the multiplicity distributions measured by previous experiments are presented. Some of these distributions will then be compared to the results of this work in the very end of the thesis.

The first results presented are from the Super Proton Synchrotron (SPS) which operated at CERN since 1976. After becoming a collider (not only an accelerator) it was called $Sp\bar{p}S$ and collided proton–antiproton at a maximum center of mass energy of 900 GeV. The results were already shown in the previous section in Figure 4.1. As previously discussed, a discrepancy from the NBD description is underlined in the NSD class at 900 GeV in the pseudorapidity interval in $|\eta| < 1.5$.

It is also interesting to have a look to the results from the Tevatron experiment at the Fermi National Accelerator Laboratory, which started to take $p-\bar{p}$ collisions in 1983 and reached a center of mass energy of 1.96 TeV. The Figure 4.2 shows the multiplicity distributions measured by the Collider Detector at Fermilab (CDF) for NSD events at 1800 GeV. The fits are to a single NBD, and, in the publication, it is stated that they fit properly only the data for the narrower pseudorapidity interval $\eta < |0.5|$.

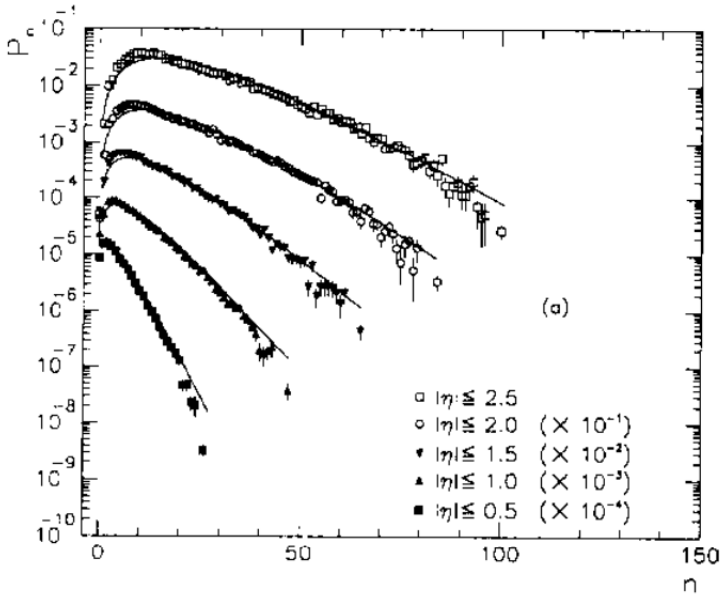


Figure 4.2: Multiplicity distributions at 1800 GeV in 5 pseudorapidity intervals for NSD events measured by CDF, solid lines are the results of the single NBD fit [55].

Moving to the LHC results, the first ones shown in Figure 4.3 are from CMS, from a NSD sample. Instead of showing the full distributions (which will be compared to this work in the end), the Figure illustrates the KNO scaling violation which is pictured in the left panel for a wide pseudorapidity range $|\eta| < 2.4$. While for narrower eta ranges $|\eta| < 0.5$, KNO scaling between 0.9 and 7 TeV holds.

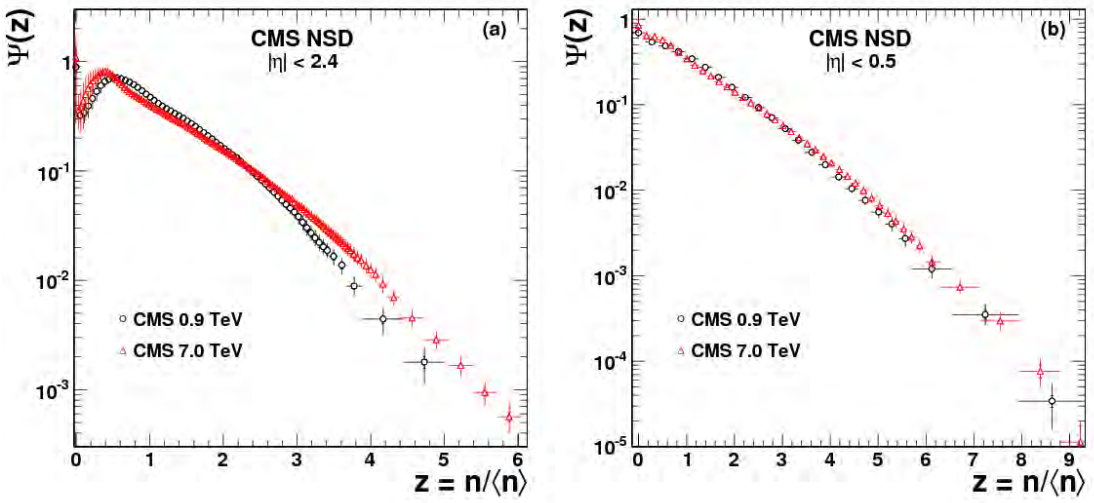


Figure 4.3: The charged–particle multiplicity distributions in KNO variables $z = n/\langle n \rangle$ at 0.9 and 7 TeV in measured by CMS in $|\eta| < 2.4$ (left panel) and $|\eta| < 0.5$ (right panel) [56].

The most interesting results for this work are of course the ones published by ALICE. Although they are done for the central pseudorapidity intervals, $|\eta| < 1.0$, they will be compared to this work in the end. In fact, in this thesis the SPD is used for the central measurements, to be able to explore the widest pseudorapidity range available at LHC. As in the CMS case, the full distributions are spared for the final chapters, and here only the KNO scaling validity and the fits to a single NBD are shown.

In the left panel of Figure 4.4 the UA5 results for proton–antiproton at 0.2 and 0.9 TeV [57] are compared to the ALICE 0.9 and 2.36 TeV pp in KNO variables $z = n/\langle n \rangle$. The events are shown in the NSD class and for the very central pseudorapidity range $|\eta| < 0.5$. The ratio between ALICE at 0.9 and 2.36 TeV shows a slight discrepancy after $z = 4$. The right panel in Figure 4.4, shows the unfolded distributions at 0.9, 2.36 and 7 TeV for the INEL class with at least one hit in the interval $|\eta| < 1.0$ compared to a single NBD fit. While for the first two energies the fit holds, for 7 TeV it is clear the it underestimates the data for $N_{ch} < 5$ and overestimates them for $N_{ch} > 55$.

In the results chapter, this work will be compared to some of the distributions mentioned above, and KNO scaling tests plus NBD fits will be performed and discussed again.

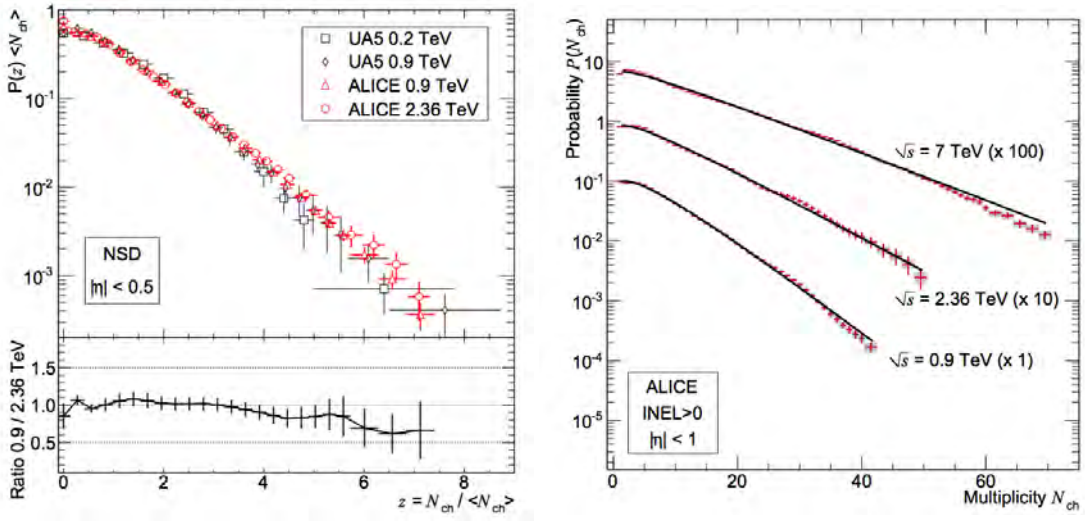


Figure 4.4: (left panel) Comparison of multiplicity distributions in KNO variables measured by UA5 in collisions at 0.2 and 0.9 TeV, and by ALICE at 0.9 and 2.36 TeV, for NSD events in $|\eta| < 0.5$ [58]; (right panel) Multiplicity distributions in $|\eta| < 1.0$ for the INEL> 0 event class. Data shown with the NBD fits [59].

Part II

Hardware and Software Tools

Chapter 5

The LHC and the ALICE Detector

In this chapter, the LHC and the ALICE detector will be briefly described, paying particular attention to the two sub–detectors used for this analysis. Using the FMD and the first layer of the SPD allows full pseudorapidity coverage when measuring the charged–particle multiplicity. The hardware of the detectors will be described, together with the data processing.

5.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is located at CERN and was installed inside the existing 26.7 km tunnel of the Large Electron-Positron Collider (LEP). It is a pp collider, which collides also $PbPb$ and pPb ions. It sits from 50 to 175 meters underground between the borders of France and Switzerland and is currently the world's largest and most powerful particle collider.

Figure 5.1: View of the LHC tunnel in Sector 81.

The LHC started to be operative the 10th of September 2008, but the 19th of September a serious damage of the superconducting magnets required a long technical stop. Superconducting magnets are an important component on the LHC ring, in fact they are 1232 in total, each 14.3 m long. The magnets are used to bend the beams. For p beams at 7 TeV they have to produce a magnetic field of 8.4 Tesla. Beams started to circulate again only in November 2009.

The 30th of March 2010, the center of mass energy of 7 TeV was reached for pp collisions, while the 5th of April 2012 the record energy of 8 TeV was registered. The LHC is capable of colliding Pb ions, and the first $PbPb$ collision was registered the 7th of November 2010, at 2.76 TeV per nucleon in the center of mass. The Pb ion beams start with a piece of pure Pb 2 cm long, which is vaporized into a small number of atoms at around 500 degrees Celsius. The ions, created removing the electrons using an electrical current, are then accelerated towards the LHC. The 13th of September 2012, the LHC has collided for the first time also pPb at 5.02 TeV.

On the 16th of February 2013, the successful LHC Run I ended and a long shutdown period started to prepare the machine to be able to run at 13 TeV. It has been decided to run LHC initially at 13 TeV instead of the 14 TeV which was the designed energy, because the superconducting magnets require a significant amount of time to reach the nominal field. To avoid taking too much time from physics data taking, it has been decided to operate at 13 TeV energy, at least for the majority of the Run II operations.

Operations at LHC have started again and are proceeding with powering tests since the beginning of 2015 (ALICE is taking cosmics now). It is expected to start the Run II with beam in the end of March and collect data useful for physics from May/June 2015 on. The heavy–ion run is planned for November. There will be no pPb collisions in 2015, but perhaps in 2016.

The LHC Experiments

There are four interaction points at the LHC. To each interaction point corresponds the location of one of the four LHC experiments, listed below. ATLAS is positioned at Point 1, ALICE at Point 2, while CMS is located at Point 5 and LHCb at Point 8, as can be seen in Figure 5.2.

ATLAS and CMS are general-purpose detectors at the LHC, which have the same scientific goals, but use different technical solutions and design to achieve them. They investigate a wide range of physics, from studying the Standard Model and the Higgs boson to searching for extra dimensions and particles that could make up dark matter.

ALICE is designed to study the physics of strongly interacting matter at extreme energy densities, where a state of matter known as quark-gluon plasma forms. ALICE will be extensively described in the following section.

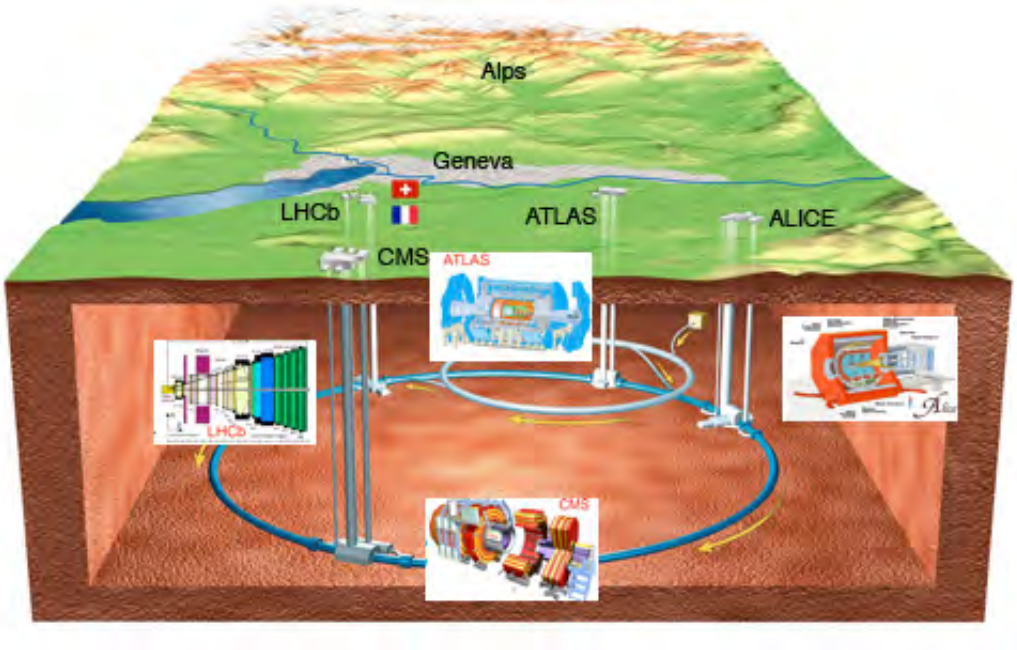


Figure 5.2: The four main LHC experiments and the LHC accelerator.

LHCb investigates the differences between matter and anti–matter by studying the $b-$quark.

There are also three smaller experiments, TOTEM, which investigates directly the proton and is located around the CMS interaction point, while LHCf studies cosmic rays and sits close to ATLAS. The newest seventh LHC experiment is MoEDAL, which searches directly for the magnetic pole and is located around LHCb.

5.2 A Large Ion Collider Experiment

The ALICE Collaboration, which counts more than 1500 scientists from over 150 institutes in 37 countries of the world, uses the A Large Ion Collider Experiment [60] to study heavy–ion and hadron collisions. The collaboration was established in 1993, year in which a Letter of Intent was submitted. Its focus is on the strongly interacting matter at extreme energy densities, produced in heavy–ion collisions, where the quark–gluon plasma (already discussed in Section 2.4) is expected to be created. The ALICE detector weights 10000 tons, and it is 26 m long, 16 m high and 16 m wide. It is situated 56 m underground in LHC interaction Point 2. In Figure 5.3, an overall view of ALICE with the red magnets' door open is shown.

The ALICE detector aims in characterizing and subsequently reconstructing every particle which is produced during the collision. This means that every one of the 18 sub–detectors which constitute ALICE, pictured in Figure 5.4, have a specific role and are designed to give information about the mass, the electric sign and the velocity of the particles.

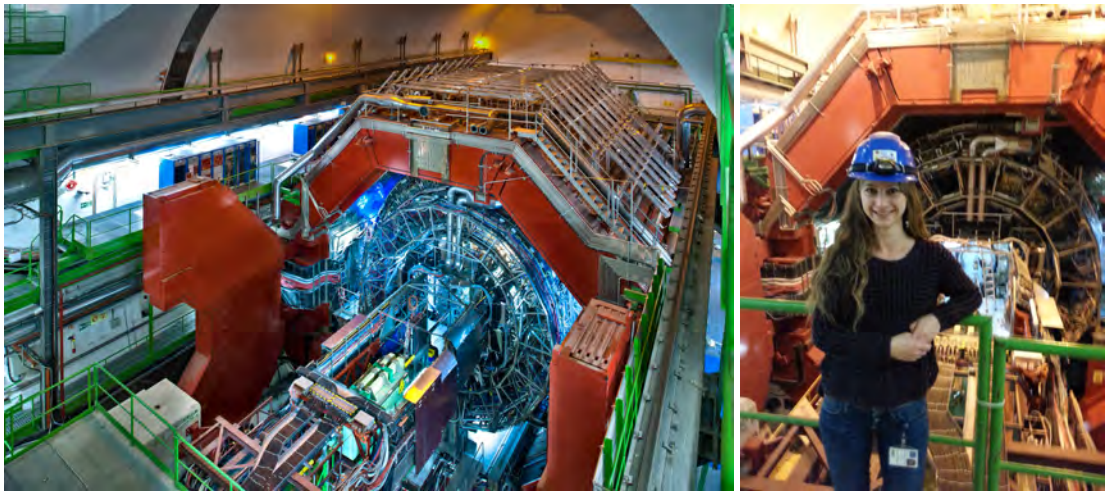


Figure 5.3: Overview of ALICE and a picture of myself in the cavern in 2013, during the long shutdown.

After the collision, the produced particles traverse the detector layers outwards. Firstly, they go through the tracking system, then through the electromagnetic and hadronic calorimeters and then they traverse the muon spectrometer. Clearly, the detectors are embedded into a magnetic field, which bends the trajectories of the particles, in order to determine the charge and the momentum of the particles. The various detectors are described in the following, paying particular attention to the detectors used to develop this analysis, the Forward Multiplicity Detector (FMD) and the Silicon Pixel Detector (SPD).

Before describing ALICE in detail, one needs to discuss the coordinates system which will be used in the following chapters:

- The nominal interaction point is defined as the origin of the coordinate system. The beam direction defines the z–axis, and the x–y plane is transverse to the beam direction. The positive x–axis is defined as pointing from the interaction point to the center of the LHC ring and the positive y–axis is defined as pointing upwards.
- The azimuthal angle φ is measured around the beam axis.
- The polar angle θ is the angle from the beam axis.
- The pseudorapidity is defined as $\eta = -\ln\tan(\theta/2)$.

Collision Overview

A reconstruction of the particles production after a pPb collision in ALICE is shown in Figure 5.5. The first thing to do after a collision is to evaluate the multiplicity of the event, i.e. how many particles were produced, and their spatial distribution. This can be done using the following detectors:

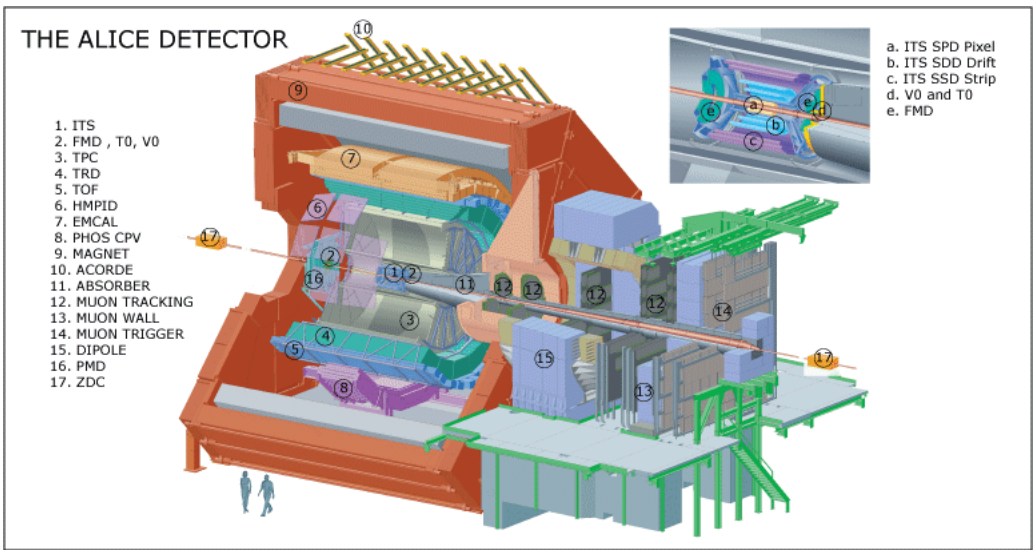


Figure 5.4: The ALICE detector set–up for Run I data taking, with zoomed view of the central barrel.

FMD The Forward Multiplicity Detector is a silicon strip detector which gives to ALICE the widest pseudorapidity coverage among the four main LHC experiments. It is constituted of five silicon rings and provides a coverage from $-3.4 < \eta < -1.7$ on the C–side of the experiment, to $+1.70 < \eta < +5.03$ on the A–side. Since it is the major detector used in this thesis work, it will be described better in Section 5.3.

T0 Is the trigger detector for timing, it can measure the timing of the collision with a very good precision (< 25 ps). Also it is used as a start and luminosity monitoring detector by the accelerator team. It consists of two arrays or Cherenkov counters, one on the A–Side and one on the C–side, each made of 12 cylindrical counters mounted as close as possible around the beam pipe, to maximize the trigger efficiency. The major objective of T0 is to provide the timing signal to the Time Of Flight (TOF), used for particle identification.

V0 Is made of two arrays of segmented oscillator counters installed in both sides of the interaction point. Its primary use is as a minimum bias trigger and to reject the beam–gas background, but it also identifies the centrality of the collision summing up the total multiplicity in the two counters. Together with the T0 it gives the luminosity delivered to the experiment. Both T0 and V0 have nearly the same pseudorapidity coverage of the FMD.

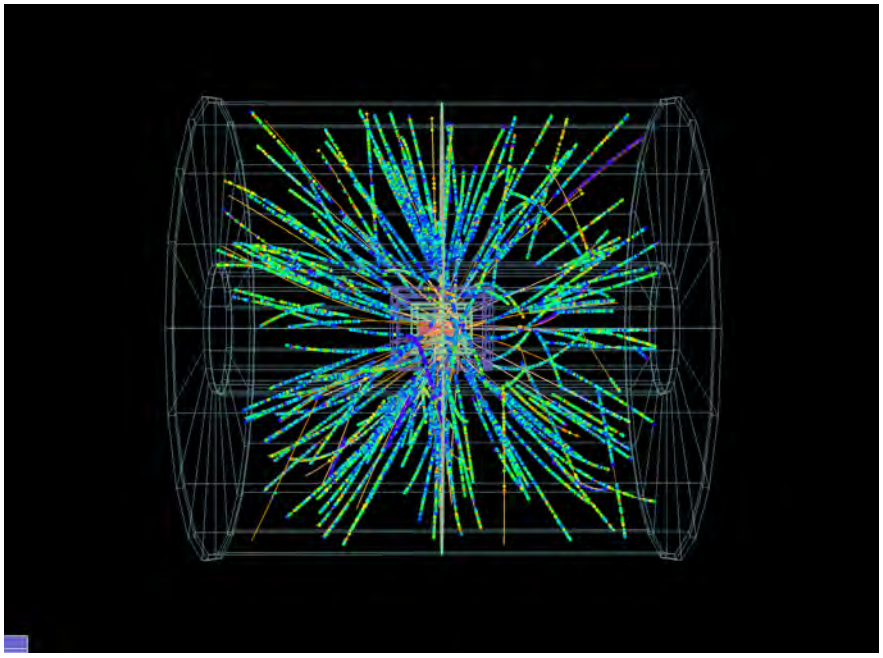


Figure 5.5: pPb collision form the pA run in January 2013. Image Credit: CERN.

ZDC The Zero Degree Calorimeter is located 116 m from the interaction point, in both sides of ALICE along the beam line. It measures the remnants and the spectators of the collision to give the measure the overlap region of two colliding nuclei and the number of participants to the collision. It is made by two hadronic calorimeters to detect protons and two to detect neutrons, which are located at around zero degrees, as the name says.

Tracking Particles

An ensemble of cylindrical detectors measure the trajectory of the charged particles. Neutral particles are not detected as it is schematically represented in Figure 5.6. The tracking detectors are embedded in a magnetic field of 0.5 Tesla, to be able to reconstruct the momentum of the particles from the curvature of the bend of the trajectories. The magnetic field is produced by the huge L3 solenoid magnet, which is 12 m long and a radius of around 6 m. In detail, the tracking detectors are:

ITS The Inner Tracking System, pictured in Figure 5.7, is made of six layers of silicon detectors and covers a pseudorapidity range of around $|\eta| < 0.9$. The first two layers constitute the Silicon Pixel Detector (SPD), which has an extended coverage up to $|\eta| < 1.98$ to provide a continuous coverage together with the FMD for multiplicity measurements. Since it is used for this work, the SPD will be discussed further in Section 5.4. The third and forth layers make up the Silicon Drift Detector, the fifth and sixth are the Silicon Strip Detector. The ITS holds the beam pipe, so that there are no relative movements, and it is coaxial with it. Its purpose is to track and identify particles which have momentum < 200 MeV, to reconstruct the primary vertex with extremely high precision (100 μm). Moreover, it improves momentum and angle resolution for particles reconstructed by

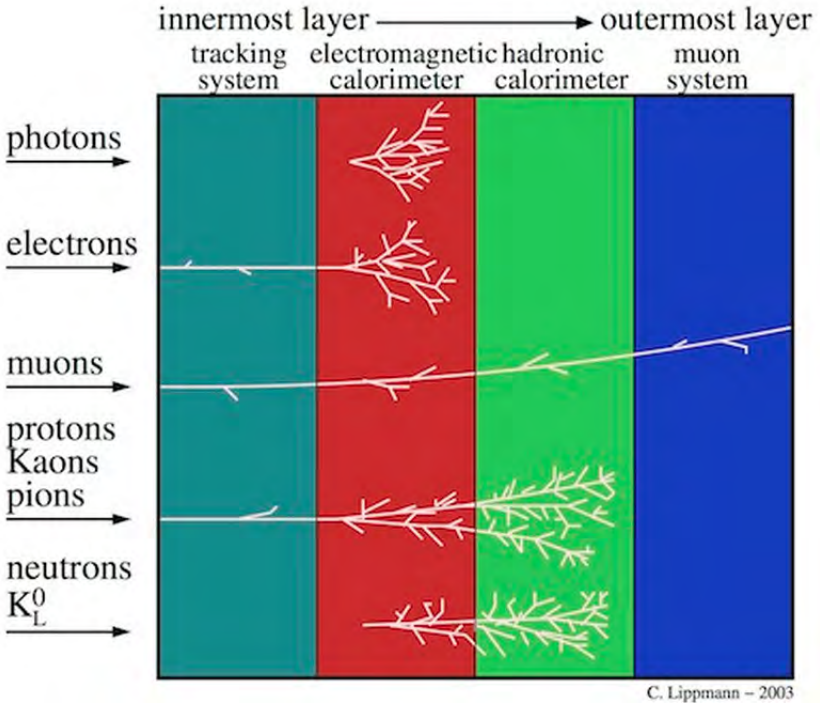


Figure 5.6: Signatures of different particle types in the different detectors. Image Credit: ALICE Matters.

the TPC, and it is able to reconstruct and identify particles which come from decays (of Hyperons and D or B mesons) and not form the primary collision.

TPC The Time–Projection Chamber is the main tracking detector of the central barrel, and it is a large volume filled with a mixture of Ne, CO_2 and N_2 gas. Its pseudorapidity coverage matches the ITS, TOF and TRD one. Charged particles from the collisions cross the gas and ionize it, liberating electrons that then drift towards the end plates of the detector. It can also be used for particle identification, checking the characteristics of the ionization process. The TPC provides also charged–particle momentum measurements and vertex determination.

TRD The Transition Radiation Detector main purpose is to identify positrons and electrons with momentum > 1 GeV via X–ray emission when they traverse the gas mixture of Xe/CO_2. It consists of 18 super modules, which contain 30 modules each. The super modules are directly visible in the ALICE overview in Figure 5.3. Together with the information from the ITS, it allows to reconstruct products of semi–leptonic decays. Moreover, in conjunction with the TPC, has the ability of studying the production of light and heavy vector-meson resonances.

Particle Identification

Stable charged particles are uniquely identified by their mass, obtained from momentum and velocity measurements, and their charge. The momentum and the sign of the charge are derived from the curvature of the track in the magnetic field. On the other

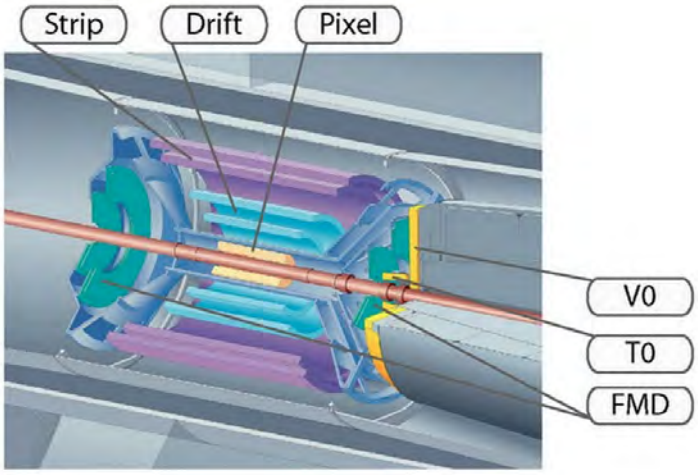


Figure 5.7: Six layers of silicon detectors which constitute the Inner Tracking System of ALICE. Image Credit: ALICE Matters.

side, the velocity of the particle can be obtained using several methods (measurements of time–of–flight, ionization, detection of transition radiation and Cherenkov radiation) according to different momentum ranges. Therefore, in addition to the information given by ITS, TPC and TRD, other information are needed to identify the particles in ALICE.

TOF The Time of Flight is devoted to measure the speed of a particle, determining the time that a particle takes to travel from the vertex to reach it. The TOF is based on multigap resistive plate chambers (MRPCs), which cover a large area of 141 m^2. The MRPCs are made of parallel thin layers of glass, which create a narrow gas gaps with high electric fields. It identifies kaons, pions and protons in a momentum range from around 0.6 to 3 GeV.

HMPID The High Momentum Particle Identification Detector is made to enhance the particle identification of charged hadrons in the momentum range between $3 - 5$ GeV, which is not covered by energy–loss measurements by ITS and TPC nor by time–of–flight in TOF. It consists of seven modules of Ring Imaging Cherenkov (RICH) chambers mounted on a support cradle. The ring-shaped image of the focused Cherenkov radiation enables a measurement of the Cherenkov angle (angle of propagation of the Cherenkov shock wave resulting from charged particles which move into a material faster than the velocity of the light in that specific material) and therefore the particle velocity.

Calorimeters in ALICE

Special detectors, called electromagnetic calorimeters, are needed in order to measure photons, neutral mesons and enhance the capabilities to measure jet properties. Photons, electrons and positrons deposit all their energy in calorimeters via hadronic or electromagnetic showers. In particular it is possible to distinguish photons, since they do not leave tracks before depositing energy. The three calorimeters in ALICE are:

PHOS The PHOton Spectrometer is a single arm spectrometer consisting of an electromagnetic calorimeter and a Charged–Particle Veto (CPV) detector. It has central pseudorapidity rapidity coverage $|\eta| < 0.12$ and it is mounted 4.6 m from the interaction point. PHOS major aim is to measure with very high accuracy photons. Each PHOS module is divided into detection cells made of lead–tungsten which is a very dense material capable to stop and detect photons.

EMCAL The ElectroMagnetic CALorimeter is a lead calorimeter with cylindrical geometry and its main objective is to study the interactions of energetic partons with dense matter produced in heavy–ion collisions. It is located close to the L3 magnet and has pseudorapidity coverage of $|\eta| < 0.7$.

PMD The Photon Multiplicity Detector is devoted to measure the multiplicity and the spacial distribution of photons on forward rapidities $2.3 < \eta < 3.7$. It is made of two plates of gas counters with a thick converter in between, moreover, it utilizes a first layer CPV.

Muon Spectrometer

Muons penetrate matter easier with respect to other particles [1], therefore, a set of absorbers is placed at forward rapidities $-4.0 < \eta < -2.5$. The Muon spectrometer, in Figure 5.8, allows to study the muon decays which produce heavy–quark vector–mesons resonances and other heavy mesons. In order to absorb hadrons and photons from the interaction vertex, it has a passive front absorber plus an additional 1.2 m thick iron wall. At this point the muons, which are able to pass those filters, are detected by a series of tracking detectors.

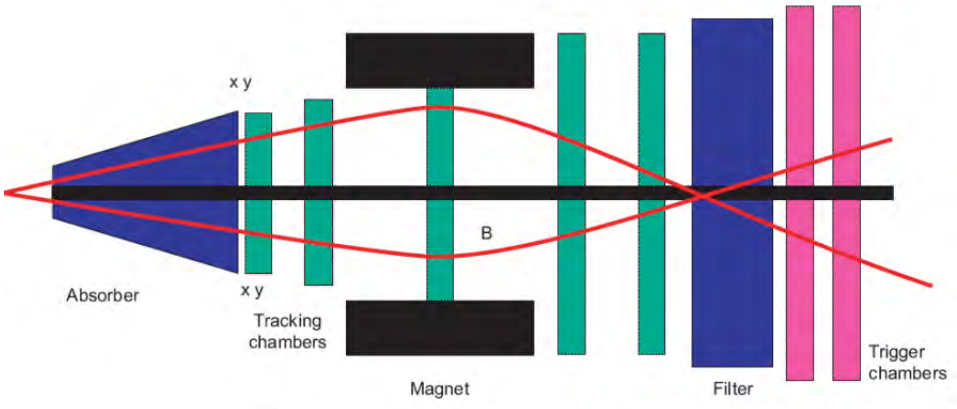


Figure 5.8: Muon Sperctrometer in ALICE composed of an absorber, a set of tracking chambers before, inside and after the magnet and a set of trigger chambers. Image Credit: CERN.

[1] Muons have a greater mass (105 MeV/c^2) respect to other particles, like electrons, and they do not emit much bremsstrahlung. This means that they do not decelerate much and can penetrate deeper in matter.

ACORDE

ALICE has also an ALICE Cosmic Rays DEtector which detects high–energy cosmic rays, in conjunction with TPC, TRD and TOF, by triggering the arrival of atmospheric muons. It is an array of plastic scintillator counters mounted on the top surface of the L3 magnet. It also provides a fast trigger signal to some of the ALICE tracking detectors.

The description of the ALICE sub–detectors has been only brief in this thesis work, for a deep understanding of the system one should refer to [60].

Upgrades to ALICE during LS1

During the Long Shutdown LS1, just concluded in the beginning of 2015, some upgrades have been implemented. Also, major upgrades are planned during LS2 in 2018, which include the replacement of some of the above mentioned detectors. In particular, during LS1 a new ALICE Diffractive AD detector has been installed and will allow better detection of diffractive events. Two scintillator paddles have been installed in the A–side and in the C-side at $4.8 < \eta < 6.3$ and $-7.0 < \eta < -4.9$ respectively. Moreover, a Dijet CALorimeter (DCAL) has been installed, which will extend the existing EMCAL system.

Trigger and Data Acquisition

There are several fast trigger detectors in ALICE: T0, V0, ZDC, SPD, TOF TRD, PHOS, EMCal, Muon Spectrometer and ACORDE. The first step is the activation of the pre-trigger in the TRD after < 900 ns, then there is the L0 trigger at 1.2 μs and the L1 at 6.5 μs. The final trigger L2 waits for the end of the past-future (P/F) protection interval of 88 μs and is issued by the slower TPC at 100 μs.

The aim of P/F protection circuit is to remove some of the pile–up in the events, to do so, it does not simply reject events. If the TPC is included in the triggering detectors it is possible to require, for $PbPb$ events, no more than 3 peripheral events, and no more than 2 semi–central events in the same 88 μs time window. On the other hand, some pile–up is accepted in pp collisions, since the multiplicities are much lower than in heavy–ion collisions. Anyway, the pile–up in the ITS is a more serious issue that biases the reconstruction of the event, therefore, the time window for P/F protection in the ITS is or around 10 μs.

The trigger and Data AcQuisition (DAQ) systems of ALICE were designed in order to share the resources among frequent triggers and detector lifetime for rare triggers. Therefore, the systems have to be able to balance among recording central collisions, which generate very large number of events, and acquire a big amount of rare trigger events.

The task of the High–Level Trigger (HLT) is to collect and process the information for the major detectors of ALICE in computer clusters. Basically, its task is divided into three steps: trigger (accept or reject an event according to online analysis), select (perform partial readout and select a region of interest within the event) and compress (apply compression algorithms to reduce the event size, without excluding physics information).

With this brief description of the triggers' systems, the general description of ALICE is concluded. In the following sections, the detectors used to collect the data for this thesis will be described more in detail.

5.3 Forward Multiplicity Detector

In this Section, the specifics of the FMD will be described together with the signal processing.

The main purpose of the FMD is to measure charged–particles multiplicity in a forward pseudorapidity range, namely $-3.4 < \eta < -1.7$ in the A–side of ALICE and $+1.70 < \eta < +5.03$ in the C–side. The FMD has full azimuthal coverage. Looking at Figure 5.9, which shows the pseudorapidity density derived with the PYTHIA generator, it is clear that there is overlap between the FMD silicon rings and the SPD.

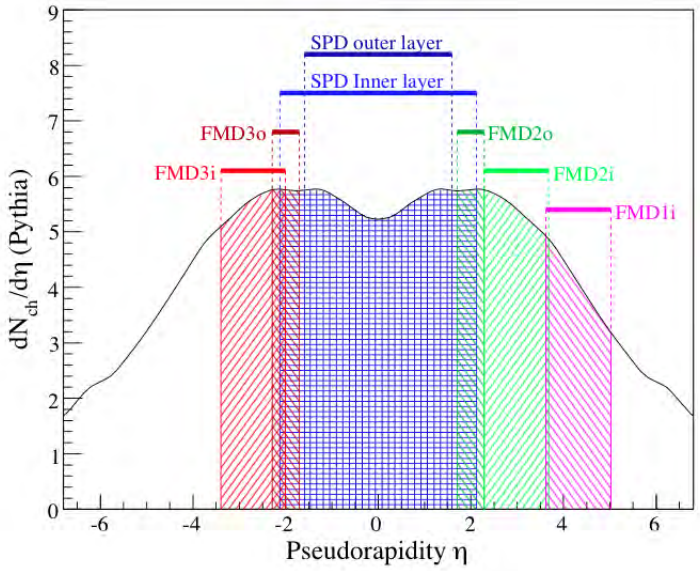


Figure 5.9: Pseudorapidity coverage of the FMD rings along with the two SPD layers seen from the nominal vertex position on simulated multiplicity distributions produced by the PYTHIA event generator with pp collisions at 14 TeV. The regions measured by each detector are illustrated [60].

The segmentation of the detector was chosen in a way that, for central $PbPb$ collisions, each primary charged particle, would occupy one strip. It must be noted that

the majority of the particles which reach the FMD are secondaries coming from interactions in the beampipe, the ITS, the T0 and V0, as well as cables and supports. The hit density in anyway never higher than three charged particles per strip. The FMD provides only offline information, since the system readout time is around 2 μs, making the detector unusable for triggering.

Figure 5.10 shows the FMD layout and the location of each ring. Both FMD2 and FMD3 have an inner and outer ring silicon sensor, located on both sides of the ITS, while the FMD1 has only one inner ring and is located further from the interaction point, opposite to the muon spectrometer. The inner rings consist of 10 silicon sensors, made of two azimuthal sectors, each with 512 silicon strips with radii from 4.2 to 17.2 cm. On the other hand, the outer rings contain 20 silicon sensors and each sector contains 256 silicon strips with radii from 15.4 to 28.4 cm; in this way, each inner or outer ring contains 10240 silicon strips. Figure 5.11 shows one of the inner rings.

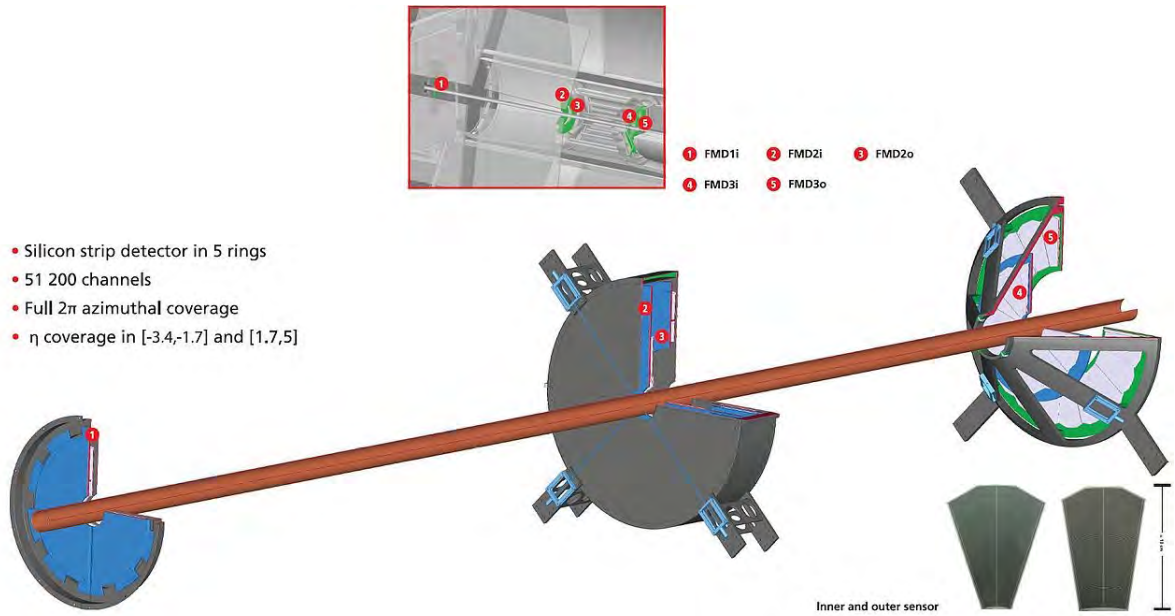


Figure 5.10: FMD layout and positioning. Image Credit: Panos Charitos.

Data Processing

The data and signal processing in ALICE is done in the same way for all the detectors. The raw data are collected and saved, but they are still not usable for physics analysis. Then, the first calibration step, called $cpass$, provides input from the calibration of some detectors, like TPC, TRD, TOF and others. After the data taking period is finished, the calibration is validated, in $vpass$, and, afterwards, the physics reconstruction pass is performed, followed by the Quality Assurance (QA) of the data [61].

Once the QA has been checked, the Event Summary Data (ESD) files are produced and are in principle usable for physics analysis, although they are large. To reduce the

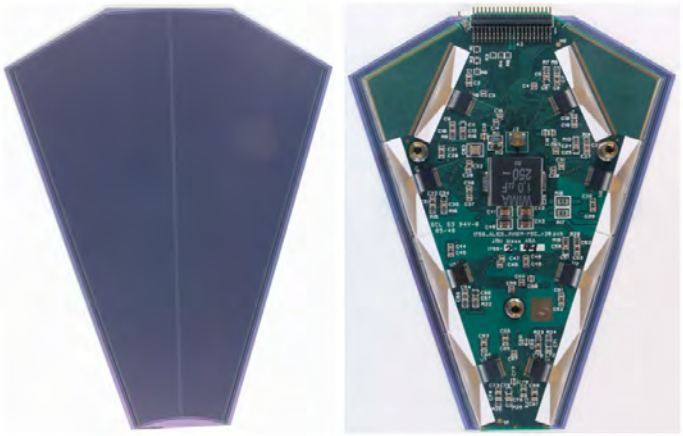


Figure 5.11: Inner silicon FMD module [60].

files size, and optimize them for the analysis, the Analysis Data Objects (AOD) files are produced. For the FMD, a big portion of the signal processing is done after the ESD files are created.

The raw data treatment begins with removal of the detector noise. Looking at the calibrated response for a single FMD strip to accumulated charged particles, in Figure 5.12, the pedestal peak at zero is evident, and it represents the intrinsic noise of the detector. The noise can be easily removed by measuring the offset and the width of the noise, when there are no beams in the accelerator.

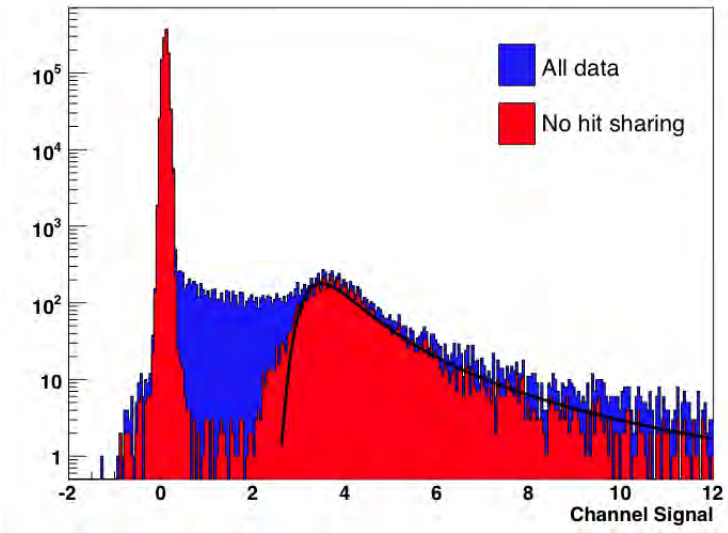


Figure 5.12: The distribution of signals from a single silicon strip. In blue all the signals seen in that strip are shown, while in red the shared hits are removed [60].

The Processing of the ESD for the FMD consists of three steps:

Sharing Filter A single particle, at a certain angle, can deposit energy in two adjacent strips of the FMD. To group strips that might have shared energy, a sharing algo-

rithm is applied. The algorithm loops over all the FMD strips, if $E_{strip} < E_{low}$ (cut set to remove remaining noise), E_{strip} is set to zero. If $E_{low} < E_{strip} < E_{high}$ (cut set to minimize the probability of merging the signals of two particles in adjacent strips into a single cluster, representing a single particle), the adjacent strips are checked. At this point, if $E_{adj} > E_{low}$, the two signals are merged into the strip with higher energy. Lastly, if $E_{strip} > E_{high}$, nothing is done. This algorithm is explained extensively in [62].

Calculation of Number of Charged Particles After the sharing filter, another cut, E_{hit}, is necessary to determine above which energy a signal should be counted as a particle. Therefore, if $E_{strip} < E_{hit}$, the strip is considered empty, otherwise it has been hit by one or more particles. N_{hit} is the number of strips with $E_{strip} < E_{hit}$. When more than one particle hits the detector, it can be difficult to fit the distribution, although it is still a convolution of a Landau and a Gaussian. The peak of the Landau will be shifted by $n(E_{MPV} + \xi \ln n)$ with n number of particles traversing the strip, E_{MPV} value of the single–particle peak and ξ the width of the Landau distribution. To estimate the most probable number of particles in a strip in a single event is complicated, and a simplified approach can be taken. It is possible to use the fit parameters to estimate the values of the three cuts, implying that the cut will only depend on E_{MPV}, first peak of the distribution. Assuming that the particles which traverse the same region of the detector are randomly distributed, the distribution of multiple particles hitting the detector is estimated using a Poissonian:

$$P(n) = \frac{\mu^n e^{-\mu}}{n!} \tag{5.1}$$

with $\mu = N_{ch}/N_{channel}$ number of charged particles per each channel (true occupancy). Therefore, the average number of particles, when a channel is hit, is:

$$\begin{aligned} N_{aver} &= \frac{\sum_{n>0} nP(n)}{\sum_{n>0} P(N)} = \frac{1}{1-P(0)} \sum_{n>0} \frac{n\mu^n e^{-\mu}}{n!} = \\ &= \frac{e^{-\mu}}{1-e^{-\mu}} \mu \sum_{n>0} \frac{\mu^{n-1}}{(n-1)!} = \frac{e^{-\mu}}{1-e^{-\mu}} \mu \sum_{m} \frac{\mu^m}{m!} = \\ &= \frac{e^{-\mu}}{1-e^{-\mu}} \mu e^{\mu} = \frac{\mu}{1-e^{-\mu}} \end{aligned} \tag{5.2}$$

Using 256 or more strips gives a reasonable estimate of N_{aver} (Accurate tests have been done in [63] and here [64]). Once this value is determined by the equation above, all non–zero channels are set to contain N_{aver} in the selected region. Finally, the value of μ is used to estimate the measured occupancy (probability of having any number of hits in the region $P(n > 0)$):

$$\begin{aligned} N_{hit}/N_{channels} \approx & P(n > 0) = 1 - P(0) = 1 - e^{-\mu} \\ \Rightarrow \quad & \mu \approx \ln(1 - N_{hit}/N_{channels})^{-1} \end{aligned} \tag{5.3}$$

Merging Finally, using the z–coordinate of the primary vertex, it is possible to calculate the (η,φ)–coordinate of each channel. All the FMD rings are then combined into a single 2–dimensional (η,φ)–coordinate histogram, which is then stored in the AOD files.

5.4 Silicon Pixel Detector

The SPD constitutes the two innermost layers of the ITS, as it can be seen from Figure 5.13. It aims to determine the position of the primary vertex and to measure the impact parameter of the secondary tracks from weak decays. The SPD operates in a region where the track density can be of 50 tracks per cm^2.

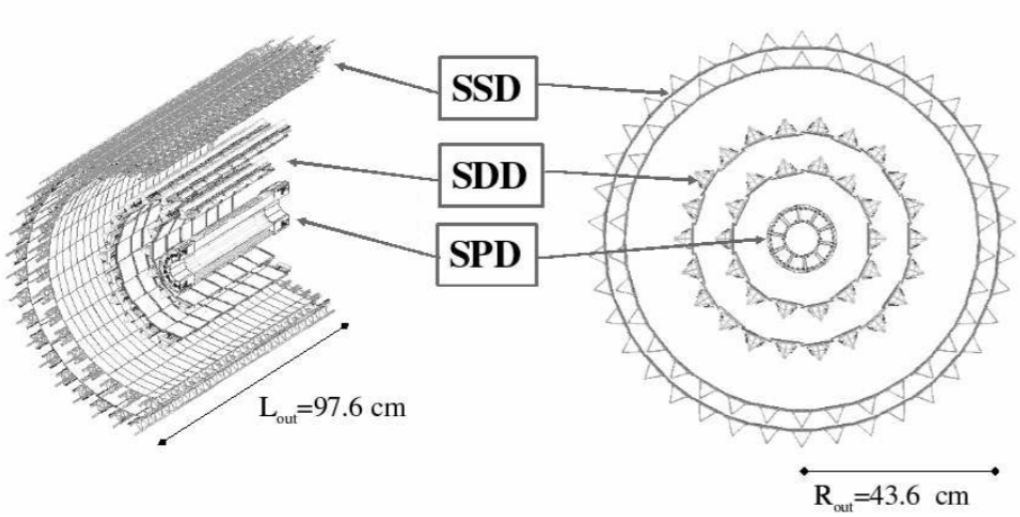


Figure 5.13: Layout of the ITS detectors, constutited by the SPD, SDD and the SSD [60].

As previously mentioned, the first layer of the SPD has more extended pseudorapidity coverage, respect to the SDD and SSD which cover $|\eta| < 0.9$. In fact the first layer of the SPD has coverage $|\eta| < 1.98$ to provide continuous measurements of charged–particle multiplicity with the FMD. The basic module of the SPD is the sensor ladder, which consists of silicon diodes, two sensor ladders plus readout electronics form a half–stave. There are 20 full staves in the inner layer and 40 in the outer layer of the SPD. The SPD records hit pixels with signals above a certain threshold, and it has a fast readout, therefore, it can be used as a triggering system to provide L0 trigger.

The current analysis only uses the clusters from the inner layer of the SPD, because it can be used as an additional FMD ring and provides full coverage between the two detectors. Therefore, the analysis procedure for the inner layer of the SPD is very similar to the one used for FMD and the innermost layer is used as a silicon layer with no additional tracking information.

Chapter 6

Monte Carlo Event Generators

Monte Carlo event generators (MC) use random number generators to sample outcomes from known physical contributions and are a very useful tool to provide simulated events to compare to the real interactions that occur in the experiment. They are generally used to understand the data, but also to estimate the corrections which are needed to be applied to the raw data, in order to derive the underlying true signal.

The description of hard and soft processes in pp collisions is different. Hard interactions (parton interactions with large–momentum transferred, see Section 1.2) can be described by perturbative QCD, while for the soft processes (the majority of the particles in LHC collisions are of this kind) the perturbative approach is not valid and a phenomenological way in is used.

In this Chapter the two MC generators used for pp analysis are described: PYTHIA and PHOJET. These event generators are used to estimate corrections as well as systematic uncertainties. Also DPMJET, which is used for pPb analysis is pictured. Finally, the transport program GEANT is described together with the reconstruction method of the FMD and SPD data.

6.1 PYTHIA

The PYTHIA [11] MC event generator is used to generate high–energy–physics events, providing a representation as accurate as possible of event properties. It is based on a combination of QCD models and analytical results. The physics input goes from hard subprocesses to initial and final–state radiations, going through underlying events descriptions and beam remnants (those processes are already described in Section 1.2). Finally, also fragmentation and decays are included.

The phenomenological models are connected to many tunable parameters, one is $p_{T,min}$, where p_T is the momentum transferred in the hard interaction. In PYTHIA, hard interactions are defined as interactions with a momentum transferred larger than

$p_{T,min}$. Moreover, the multiplicity distributions are sensitive to the tuning of this parameter. To estimate the efficiency corrections in this specific work, a particular tune with $p_{T,min} = 2$ GeV has been used.

The first step in the generation of an event is to randomize the process that has to be generated, and the selection of the process in governed by the different cross–sections already discussed in Section 1.3. The total cross–section is the sum over several components written in eq. (1.3). The optical theorem is used to derive the elastic contribution (these processes cannot be measured by ALICE), while Regge theory is adopted to determine the diffractive cross–sections. The non–diffractive fraction represents the majority of the events at LHC energy, and is handled using QCD.

In the case of a diffractive process, a diffractive mass is selected and the system is handled as a string with the quantum numbers of the original hadron. For more information on how the PYTHIA generator works, one should refer to the PYTHIA 6.4 manual [11].

Perugia 0 (320) Tune

With respect to the older S0 tunes of PYTHIA, the Perugia Tunes [65] include updates from LEP data in the fragmentation and flavor parameters. In particular, the Perugia 0 tune, which was developed in February 2009, uses peculiar parameters that describe minimum–bias results, underlying–event and Drell–Yan data sets taken at Tevatron, and a cut–off at $p_{T,min} = 2$ GeV [66]. The specific tune used for the estimation of the trigger and vertex inefficiency (which will be described extensively in the last part) was adjusted using the Kaidalov–Poghosyan model [67], basically the measured diffraction cross–sections and the shapes of the diffractive masses M_X (defined in Section 1.3) were tuned.

ATLAS–CSC (306) Tune

The Flat Tune of ATLAS–CSC [66] is used to derive the response matrices to unfold the measured distribution. It is not tuned to describe the underlying–event and has a cut–off at $p_{T,min} = 1.9$ GeV. It is possible to create the response matrices also from a regular MC, although it would not be possible to populate the higher multiplicities due to excessive computing time. Therefore, a flat MC, which does not have physical meaning, is used; the multiplicity is set to have the same value of particles up till a certain value, therefore it is flat.

6.2 PHOJET

The PHOJET event generator [68] describes high–energy collisions handling the hard and the soft component separately. The hard component is described by perturbative QCD, like in PYTHIA, while the soft part uses the Dual Parton Model (DPM). This model is based on Regge theory. Basically the cross–sections are a sum of a Reggeon exchange part, a soft Pomeron exchange and a hard part.

The momentum distributions of non–diffractive events are described by an exponential below the $p_{T,cut-off}$ and by perturbative QCD above. If the event is diffractive, the momentum transferred and the M_X are described using the triple–Pomeron approximation (see [68] for details). One of the major differences with PYTHIA in the diffraction description is that also hard interactions can occur between the Pomeron and the proton.

Also for PHOJET, the diffractive masses and cross–sections were tuned with the Kaidalov–Poghosyan model [67]. This particular tune is then used to estimate inefficiencies in our data and for the systematic uncertainties calculation.

Comparison between Generators and Data

In Figure 6.1, the comparison between the CMS published data for NSD sample at $|\eta| < 2.4$ [56] and the generated events for PYTHIA Perugia 0 and PHOJET, both tuned with the [67] model, is shown. PHOJET does agree well with the data, although not in the first bins, where diffractive events play a major role. On the other hand, PYTHIA does a better job in the first region, but does not reproduce the tail of the distribution.

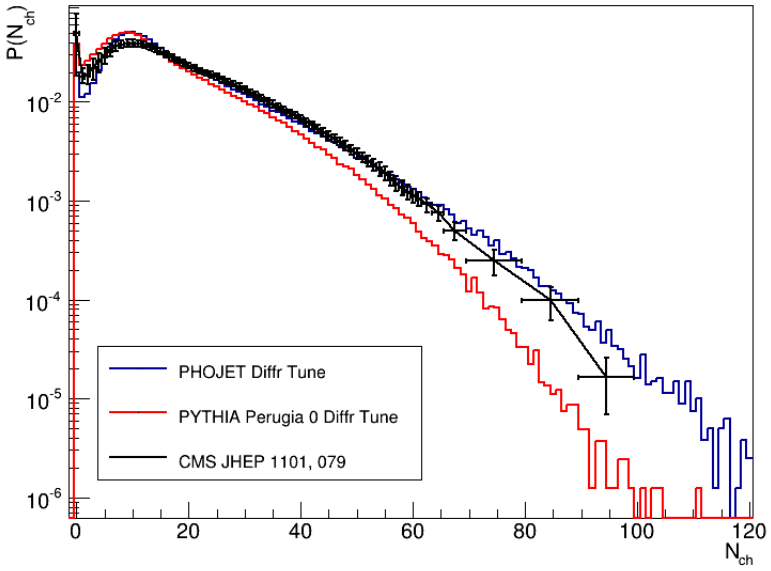


Figure 6.1: Comparison between PYTHIA and PHOJET generators, tuned for diffraction, for NSD events at 900 GeV and $|\eta| < 2.4$ with CMS data taken from [56].

6.3 DPMJET

The DPMJET–III [69] MC generator is based on the Glauber–Gribov formalism to calculate the nuclear cross sections [70]. Also it treats soft and hard scattering interactions, using the Duon Parton Model (like PHOJET) as a two–component model. It uses Reggeon theory for soft interactions and perturbative QCD for hard interactions.

6.4 GEANT

The GEometry ANd Tracking [71] program is designed to simulate the passage of particles through matter and originally was developed for High–Energy Physics, but now it is used also for other kinds of experiments and applications (e.g. in medicine). Its principal usages are to simulate the detector response transporting the particles through an experimental setup and to represent graphically the setup and the particle trajectories.

GEANT allows to describe a structure of geometrical volumes, moreover, each volume has a medium number which relates it to the specific material of the volume. It allows also to accept events from MC generators and transport particles through the volumes of the setup. This is done taking into account boundaries between volumes and physics effects from particles, their interaction with the magnetic field and with matter (secondary particle production, scattering and energy loss). Also it records particle trajectories and visualizes the setup and the trajectories. ALICE uses GEANT version 3 (written in Fortran), GEANT4 (tuned to newer experimental data and written in C++) is also available, and GEANT5 (which will speed up the simulation process using modern CPUs parallelism) is under development. To learn more about GEANT, one should have a look to the manual [71] and also here [72] for a description of tracking references (which contain information on MC and trajectories).

6.5 Detector Response and Digitalization

In order to calculate the detector response, it is necessary to couple the GEANT program with a proper description of the ALICE system, which includes all the wires, supports, detector elements, cooling pipes etc... in ALICE. After that, the simulated detector responses are processed through the same code and algorithms of the real data. This procedure allows to cross–check the data and compute systematics.

In the past, the description of some parts of ALICE were not accurate (e.g. the ITS support and cooling pipes), this affected the analyzes of the forward detectors like the FMD. In fact, also in this work, the subtraction of the secondary particles to unfold the primary multiplicity distributions, relies of MC. This means that if MCs do not contain the correct amount of material, the secondary subtraction will not be accurate enough. The geometry description has been notably improved in the last 3 years, especially during 2013, and the final distributions changed by 5-10% using better material description. The material budget will be further discussed in the last part of this work.

FMD

The FMD response is simulated using the transport code, which evaluates the energy loss E_{loss} of the particles which hit a specific channel. Before converting the signal into ADC counts, noise is added. This is done using results from actual calibration measurements. The ADC count c_i for a strip i, is the sum over:

$$c_i = p_i + x + g_i E_{loss} C \tag{6.1}$$

where p_i is the pedestal from an actual pedestal run and x is the noise taken randomly according to a Gaussian distributions which has a width of the known detector noise. Furthermore, the E_{loss} is multiplied by the real gain g_i and C, a conversion factor from gained signals to ADC counts. A detailed description of this process for FMD can be found in [73]. Now, the MC simulated data are propagated and reconstructed in the same way as for real data, as described in Section 5.3.

SPD

Some additional words must be said about the SPD response. Charged–particles can deposit energy in neighboring pixels when they cross the SPD layers, and therefore, the particles reconstructions can be done counting group of hit pixels, called clusters. Tracklets are obtained combining clusters of the two SPD layers and the primary vertex, allowing one to select primary particles with very high efficiency. As can be seen in Figure 6.2, the vertex is the point of intersection between the most of the tracklets. In this work, only the clusters of the inner layer of the SPD are used, as already explained in Section 5.4.

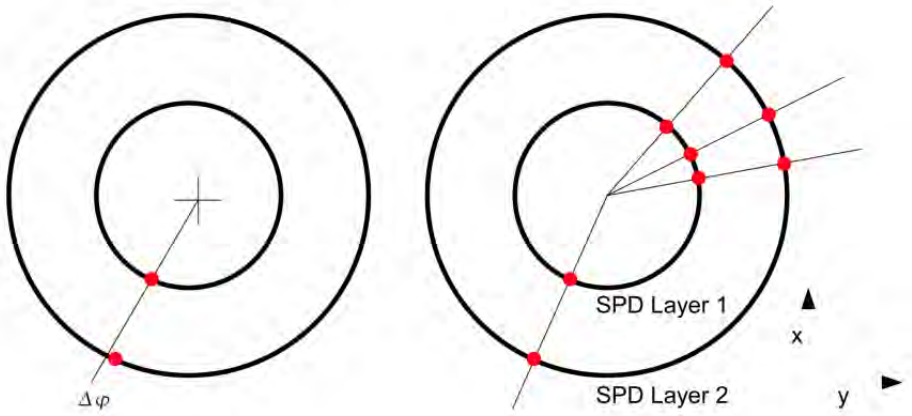


Figure 6.2: (left) Tracklet defined by a cluster in the first later coupled with one in the second layer, in a $\Delta\varphi$ window. (right) Interaction vertex determined by tracklets pointing back [63].

Part III

Data Analysis and Results

Chapter 7

Analysis for pp Collisions

In the previous chapters, hadron and heavy–ion collisions have been described, and also the charged–particle multiplicity distributions have been portrayed from the theoretical point of view. Moreover, the hardware of the ALICE detector was also described together with the software and MC generators. At this point, the procedure to analyze the pp multiplicity distribution can be described, and it will be done in several sections corresponding to the steps which are performed.

As already pointed out, the detectors used for the analysis are the FMD (Section 5.3) for the forward pseudorapidity range and the SPD (Section 5.4) for the central region. The pp distributions are analyzed for four different available center of mass energies: 0.9 TeV, 2.76 TeV, 7 TeV and finally 8 TeV. The steps to perform the measurements are:

Event Selection The trigger and vertex selection of the events are described.

Particles Selection The cuts on the detectors and the counting of the particles procedure are analyzed.

Unfolding Procedure The main ingredients of this analysis are the raw data and the response of the detector, which allows to unfold, i.e. to correct the raw distribution from the effects of secondary particles and acceptance gaps, in order to obtain the true primary multiplicity distribution. The detailed procedure used and the two methods are described.

Efficiency Correction The distributions have to be corrected for inefficiencies in the lower bins, due to biases in the vertex and trigger evaluation and detection in case of production of few particles.

Three different MC generators are used to derive the corrections to the raw data. Firstly, the flat tune of ATLAS–CSC (Section 6.1) is used to build the response of the detector. Afterwards, two different MCs, PYTHIA Perugia 0 (Section 6.1) and PHOJET (Section 6.2), both diffraction tuned using the Kaidalov–Poghosyan model [67], are used to compute the efficiency corrections.

All the analyzes performed by ALICE, therefore also this thesis, are based on the AliROOT software [74], which is build in ROOT [75], a C++ based software developed at CERN.

7.1 Event Selection

For pp analysis, four different energies were considered in this thesis, which is therefore complete and includes all the center of mass energies provided by the LHC in Run I. Namely, the energies go from 0.9 TeV and 2.76 TeV, to 7 and 8 TeV. There was a fifth energy available (2.36 TeV) for a very short while, but it is not considered in this thesis, since there were no stable beams in the LHC and only the SPD was turned on.

The 900 GeV data were collected in May 2010 with positive polarity of the solenoid magnet; the 2.76 TeV data, instead, were collected in 2011. The latter data were taken at the same energy of the $PbPb$ collision. As it will discussed in the following, the 2.76 TeV data are more problematic than the others, as will be described later. The 7 TeV data used in this thesis were collected in two different periods in 2010, one with positive and one with negative polarity of the L3 magnet. This gives rise to higher period–to–period fluctuation, as will be underlined in the systematics chapter. The 8 TeV data had to be chosen very carefully, due to the very high pile–up contamination (this phenomenon will be discussed in the following). Also in this case, analyzed data are taken from two different periods in 2012, but with the same polarity of the magnet.

In Table 7.1, the properties of the data sample are presented. The average number of interactions per bunch crossing μ gives a first idea of the amount of pile–up present in the data, and it is measured by the experiment. Keeping μ low minimizes the probability of having more than one particle interaction in the same bunch crossing. For the values considered in this work, the average probability of having more than one interaction, in a single bunch crossing, is of $\backsim 0.1\%$.

Energy (TeV)	Data Period	Selected MB Events	$\langle\mu\rangle$
0.9	LHC10c	7.4×10^6	0.04 ± 0.01
2.76	LHC11a	25×10^6	0.03 ± 0.01
7	LHC10c and LHC10d	61×10^6	0.04 ± 0.01
8	LHC12b and LHC12d	26×10^6	0.02 ± 0.01

Table 7.1: Data samples used for the analysis. For each center of mass energy, the period used is written together with the total number of selected MB events and the average number of interactions per bunch crossing

In the table, the period identifies an ensemble of runs which are taken with the same conditions (e.g. same luminosity, or same colliding system). On the other hand, inside the same period, the data taking stops for several reasons, usually because one of the ALICE detectors goes into error. Therefore, the period is divided into different runs.

Trigger Selection

The online trigger selection (in the reconstructed data) for all the pp analysis included in this work is MB$_{OR}$. This online trigger is an OR requirement between the V0 in the A–side, the V0 in the C–side of ALICE and the SPD (V0A || V0C || SPD), basically at least one charged particle hits one of these detectors.

However, three different offline triggers have been used in the analysis. Firstly, the INEL trigger which is equivalent to the online MB$_{OR}$, then the INEL> 0 trigger, which is the same, but requiring one charged particle in the SPD in the interval $|\eta| < 1.0$ ((V0A || V0C) & SPD).
The last offline trigger is used to remove the majority of the Single–Diffractive events, and is therefore called Non–Single–Diffractive NSD trigger. To do so, it requires the coincidence of V0A & V0C (often it is called V0AND or MB$_{AND}$), in this way, there is always at least one particle hitting the V0s, and the Single–Diffractive events are most likely removed. It is possible to understand this looking at the Figure 1.6b, where it is clear that the SD diffractive events populate the very first bins and the one particle requirement partially removes them. Unfortunately, the efficiency of detection of so few particles is not very high in ALICE, and the samples used in this thesis have to be corrected as it will be described in the following sections. The installation of the new AD during LS1, Section 5.2, will allow ALICE to detect with better precision the NSD sample, using a new trigger definition V0A & ADA & ADC & V0C.

Background Rejection

There are different sources of background events, which can be counted in the multiplicity distributions if they are not removed. Contamination comes from beam gas and beam halo, the first is due to the fact that the vacuum in the beam pipe is not perfect. Some of the beam particles may collide with a gas molecule and be misidentified as regular beam–beam collisions. The beam halo, instead, comes from actual particles from the beam, which escape from it and collide afterwards with a particle of a new bunch. The most efficient way of correcting for these effects is to require that the V0 timing from the two V0A and V0C coincides with the interaction timing.

The most relevant background comes from pile–up, i.e. when two or more particles collide in the same bunch crossing. A pile–up event is recognized and afterwards removed according to different selection criteria; the distance between the main vertex and an additional one has to be more than 0.8 cm, and, any additional vertex has to be identified by at least 3 tracks in the SPD. Those pile–up events are removed.

To avoid possible non–removed contamination coming from non–identified events, the considered data samples are anyway taken with a very low average number of interactions per bunch crossing $\langle\mu\rangle < 0.04$. The maximum μ, in the pp data considered, is 0.06 in the period LHC10c for 7 TeV data, where the interaction rate was of about 800 Hz. Also in this extreme case, relying on Poissonian statistics, the probability of having zero particles interacting is $P(0) = e^{-0.06}$, while $P(1) = e^{-0.06} \cdot 0.06$. Therefore, the probability of having more than one particle interacting in the bunch crossing is $P(n \geq 2) = 1 - P(0) - P(1) = 1 - 9.4 \times 10^{-1} - 5.6 \times 10^{-2} \backsim 4 \times 10^{-3}$, around 0.4%.

The maximum probability of non–identifying a pile–up event with the algorithm explained in the above paragraph is of 10%, i.e. maximum 10% of pile–up events might escape [76]. This means that, with the algorithm, only 0.04% of pile–up risk to remain in the distribution. For this reason, the pile–up mis–identification is shadowed by the systematic uncertainties and no correction for it is applied in this thesis.

Vertex Selection

The vertex determination used in this thesis is obtained from the SPD, as already shown in Section 5.4. Removal of beam halo can be performed selecting a vertex $-10 < v_z < +10$ cm, while the beam gas does not depend on the position of the vertex, and can be removed as already described in the previous section. Anyway, a more stringent requirement is necessary in order to maximize the coverage between the SPD and the FMD. By all means, it is important to have nearly full coverage when measuring the multiplicity distributions over a certain pseudorapidity range. The coverage of the detectors depends on the vertex selection performed, as can be seen in Figure 7.1, selecting $|v_z| < 4$ cm the non–covered areas filled in green are mostly excluded.

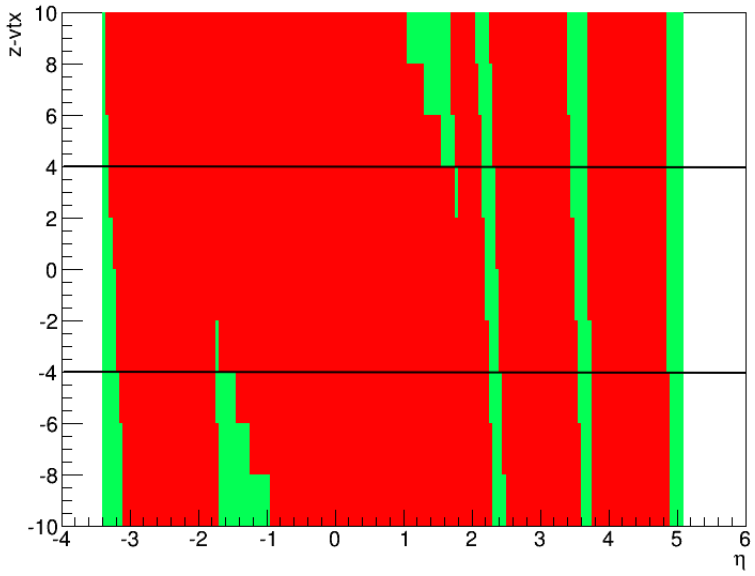


Figure 7.1: Green: maximum acceptance of the SPD and FMD. Red: actual coverage as a function of vertex position. The black lines on the y axis identify the vertex selection used in this thesis $|v_z| < 4$ cm.

To have a determination of vertex as accurate as possible in this analysis, it is required that the maximum error on the measurement of the vertex is of 0.2 cm. The vertex distribution along the z axis is shown in Figure 7.2, it is clear the the distribution is peaked $\backsim 0$, although slightly shifted and centered at $\backsim -0.6$ cm. In the figure, also the selection $|v_z| < 4$ cm is shown.

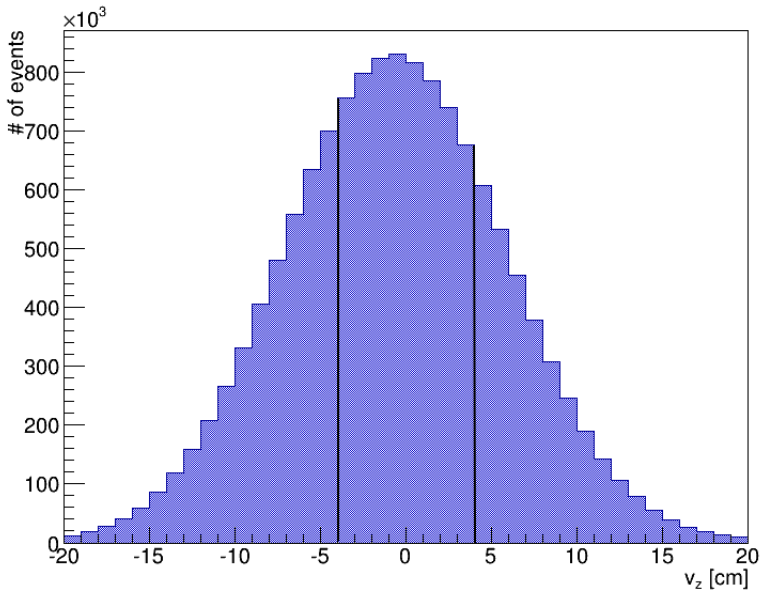


Figure 7.2: Number of events per vertex bin. The black lines identify the vertex selection used in this thesis: $|v_z| < 4$ cm. The v_z-distribution shown is for pp events at $\sqrt{s} = 2.76$ TeV in period LHC11a.

The SPD is also capable of measuring the x-y vertex distribution, which is shown in Figure 7.3 for 900 GeV and 7 TeV. In both cases, it is visible that the vertex is shifted from the (0,0) position by some millimeters, the shift is very small and does not affect the multiplicity measurements. The important point is that the MC generators, used to unfold the distributions, are anchored to the specific runs and have the same shift in the vertex. For this reason, also the x-y vertex distributions for MCs are shown, and they have the exact same shift.

7.2 Particle Selection

In this section, the cuts on the energy distributions for the FMD are presented. Particular attention will be devoted to the $\sqrt{s} = 2.76$ TeV data, which are problematic, both for SPD and FMD. For the SPD, the clusters of the inner layer have been used, as described in Section 5.4. The clusters are stored in a 2–dimensional $(\eta - \varphi)$–histogram, exactly like it is done for the FMD, to ensure that the two analyzes are carried on in the same way.

FMD Cuts

The cuts which are implemented in the FMD energy distribution come from the algorithms which were described in Section 5.3. In particular, the sharing filter needs a low, E_{low}, and a high cut, E_{high}.
In Figure 7.4, the energy distributions for all the five FMD rings are shown for a run at 7 TeV, taken in period LHC10d. The distributions show a pedestal peak at low values

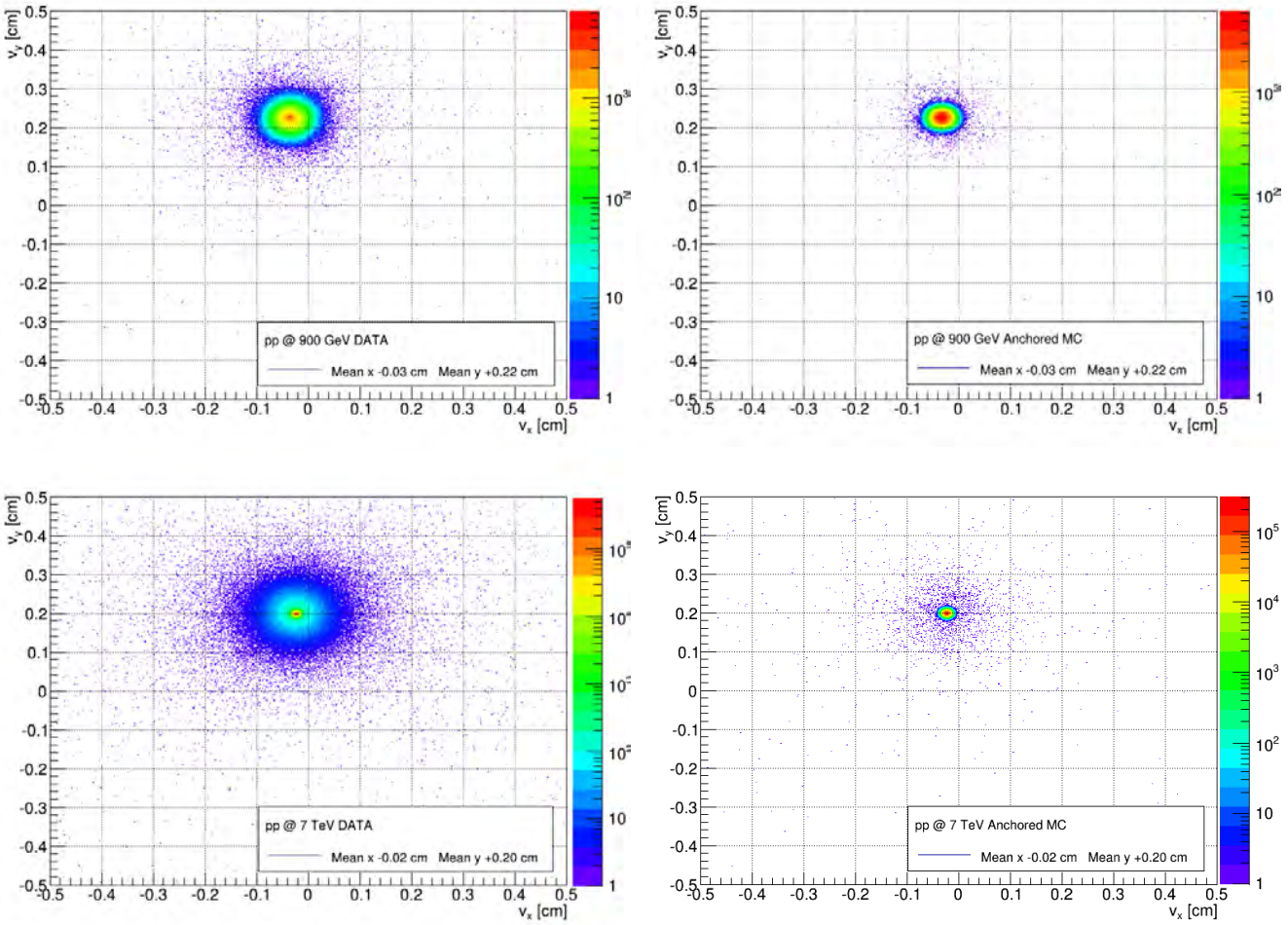


Figure 7.3: x-y vertex distribution of pp at $\sqrt{s} = 0.9$ and 7 TeV data collected in period LHC10c. On the right hand side, the corresponding anchored MCs are shown. The MC have the same exact shift in x and y, although, being MC, they are narrower and more accurate.

(in the red shaded region), a sharing region (shadowed in green, with a plateau of energy deposition) and a convoluted Landau–Gaussian.

The E_{low} cut is fixed at 0.15·$\Delta E/\Delta E_{MIP}$ for the inner and at 0.18 for the outer rings. This cut is applied in order to remove the remaining pedestal peak. The E_{high} and the cut for the Poisson algorithm E_{hit} (see Section 5.3) are identical and are defined as the most probable value minus the width of the Landau distribution: $E_{high} \equiv E_{hit} = E_{MPV} - \xi$. These values are obtained from fits to the energy distributions, and are different from data to MC.

The cuts are computed in pseudorapidity bins, with a width of 0.05 units in η, in which the value of the cut is constant. Since each plot is for an entire ring, there is a slight shift and the cut looks like a narrow region, instead of a clear cut. If the region in which the signal is merged ends at $E_{loss} \sim 0.5\Delta E_{MIP}$, the maximum variation of

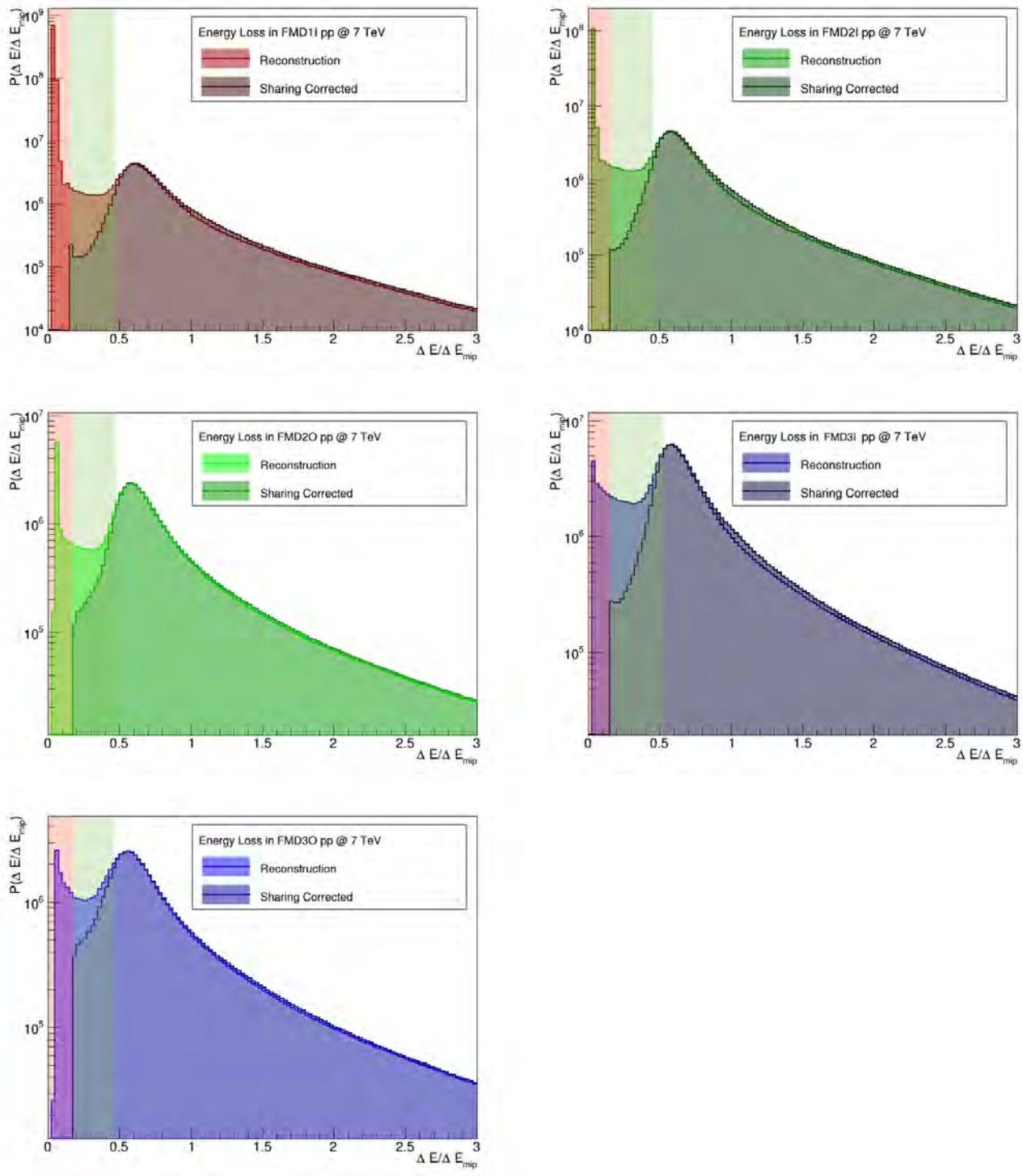


Figure 7.4: Energy loss distributions for $\sqrt{s} = 7$ TeV data collected in period LHC10d for run 126284, for the five FMD rings. The lighter distribution is the ESD reconstructed energy loss, while the darker on top is the sharing corrected one used in the analysis. In red the removed pedestal region is delimited, while in green the sharing region is underlined.

the cut within the entire ring is of about 8%[1]. For the particular run considered in the figure (126284 in LHC10d), E_{MPV} goes from 0.48, for FMD2I, to 0.58, for FMD3I, while ξ goes from 0.025 to 0.095 for the same rings. This means that the effective cut $E_{MPV} - \xi$ goes from 0.46 to 0.53. Particles in the red region, below E_{low} are discarded, particles from E_{low} to E_{high} are considered for merging, while particles above E_{high} are not merged.

These cuts are chosen to ensure better agreement between MC and data. The occupancies, for all four energies considered for pp analysis in this work, are shown in Figure 7.5. As expected the mean occupancy (probability of having one of more hits in the detector) increases, increasing the energy $\sqrt{s}$.

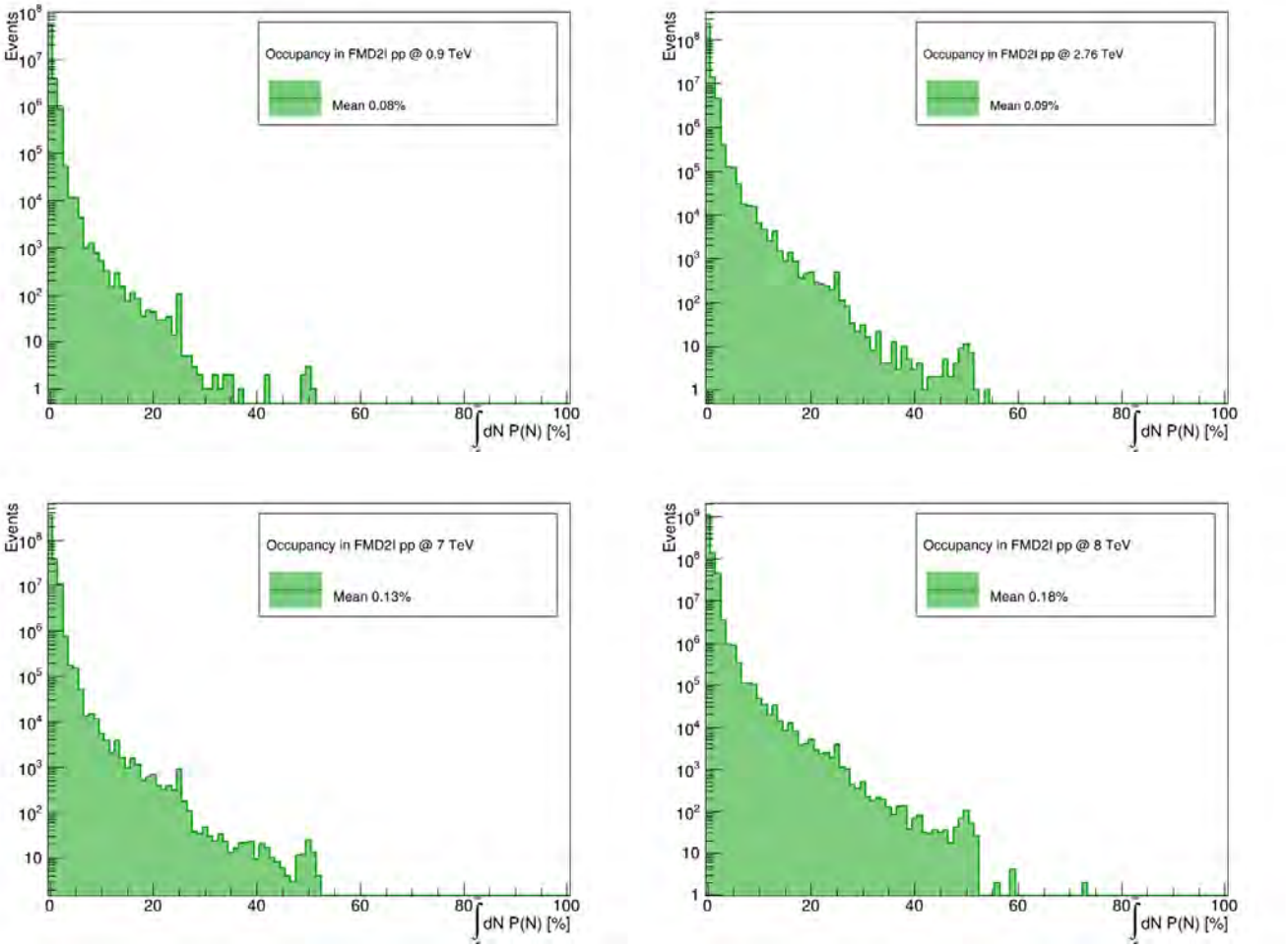


Figure 7.5: Occupancy for the four analyzed energies for FMD2I ring. The small peak at bin 50 and 25 is due to rounding issues when binning a histogram with double–precision floating–point numbers.

[1] MIP stands for Minimum Ionization Particle, defined as the traversing particle depositing the minimum ionization energy. See [63] for details.

Discussion about pp collisions at $\sqrt{s} = 2.76$ TeV

The data quality both for SPD and FMD is very good for the collisions at 0.9, 7 and 8 TeV, unfortunately the same cannot be said for 2.76 TeV data, as shall be seen in the following. Issues are found both for SPD and for FMD, a way to solve the bad quality for SPD was found, but the same cannot be said for the FMD, for which a solution is still not found. The SPD problem and solution are presented first.

The main ingredients to obtain the final true distributions are, of course, the raw data. When looking at the SPD clusters distributions for LHC11a data n Figure 7.6, it is clear that there is a weird bump behavior which is not observed in the tracklets. Normally, there is good agreement between clusters and tracklets, as demonstrated by the very good consistency between this analysis and a parallel one carried on with SPD tracklets in ALICE [76].

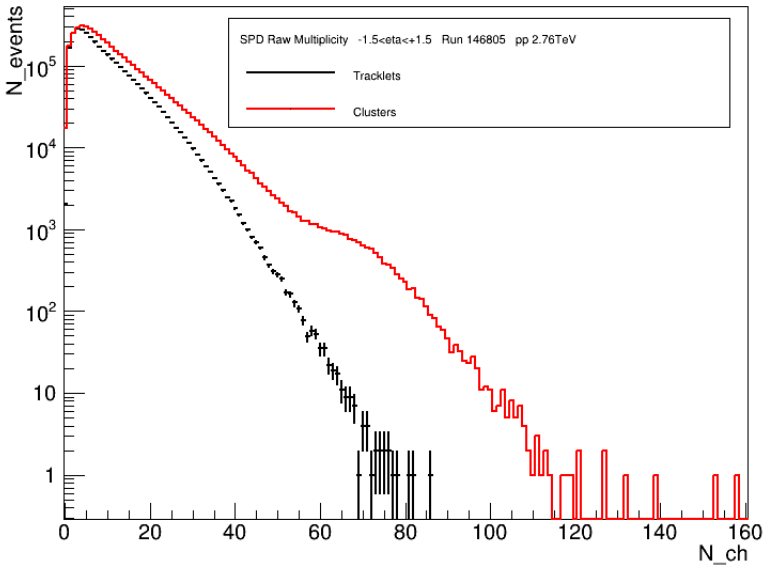
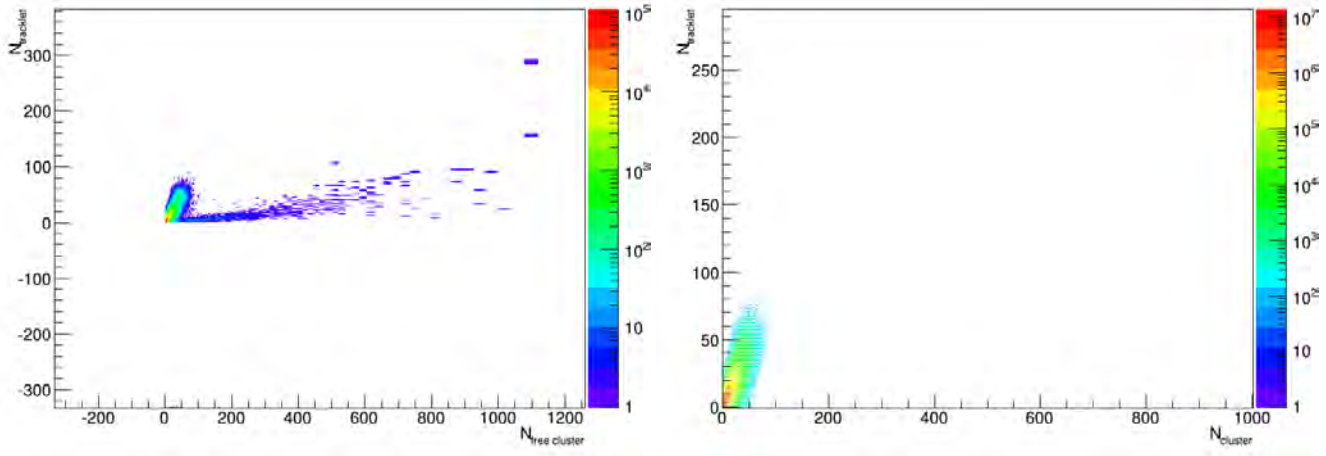


Figure 7.6: SPD tracklets in black (done by Anton Alkin) and clusters in red, used in this analysis, for run 146824 in LHC11a. A bump is present in the clusters, also the slope of the curve is different.

A way to try to debug this problem is to plot tracklets versus clusters to see if there is some halo in the distribution. In fact, as can be seen in Figure 7.7a, there are a lot of outliers, i.e. clusters which are not counted as tracklets. At this point, the idea is to cut out these outliers using a diagonal cut and see if they are causing the bump. The cut is defined as:

$$C = 65 + N_{tracklets} \cdot 4 \tag{7.1}$$

and if the events lay below the cut C they are discarded. From this, Figure 7.7b is obtained, and it is clear that the outliers are cut away. Unfortunately, also applying this procedure, the bump was still present.



(a) (b)

Figure 7.7: (a) Tracklets versus free clusters showing a few hundreds of events firing clusters, but not tracklets; (b) Tracklets versus free clusters after applying the cut defined in eq.(7.1), the halo is removed.

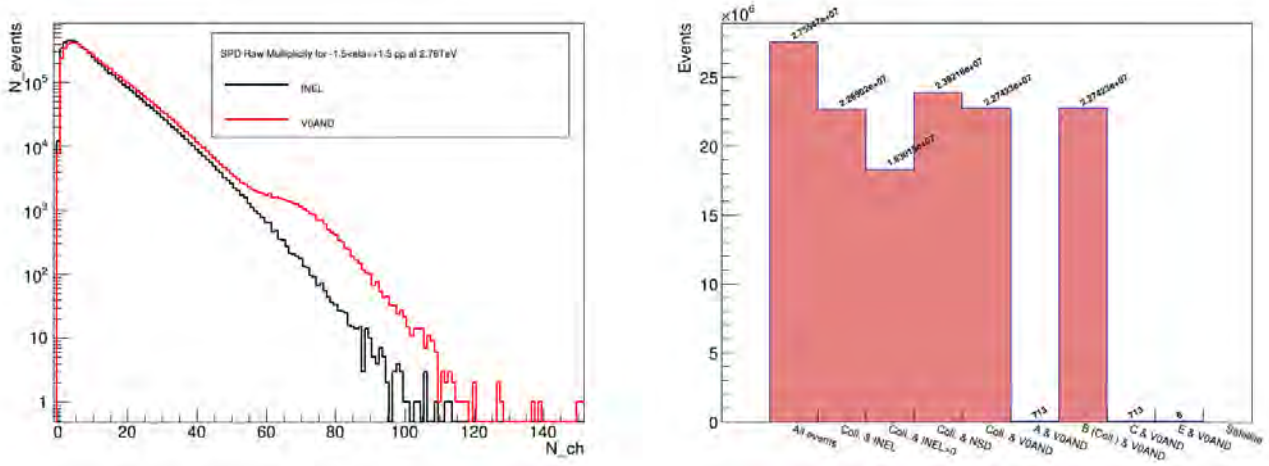
Plotting two different raw samples, Figure 7.8a, one obtained with the INEL trigger definition (MB_{OR}) and one using NSD (V0A & V0C), it was clear that the bump is present only in the NSD trigger sample. Moreover, as explained in Section 7.1, the NSD, defined as V0AND, must be a sub–sample of the INEL trigger, but looking at the triggers histogram (which specifies how many events belong to a specific trigger), the V0AND sample contained more events than the INEL (Figure 7.8b).

After making sure that the timing of the two V0s was fine [77], the solution came from boolean analysis:

$$\begin{aligned}(INEL)\,\&\,(V0AND) \equiv (V0A||V0C||SPD)\,\&\,(V0A\&V0C) \equiv \\ \equiv (V0A)\,\&\,(V0C) \equiv V0AND\end{aligned} \tag{7.2}$$

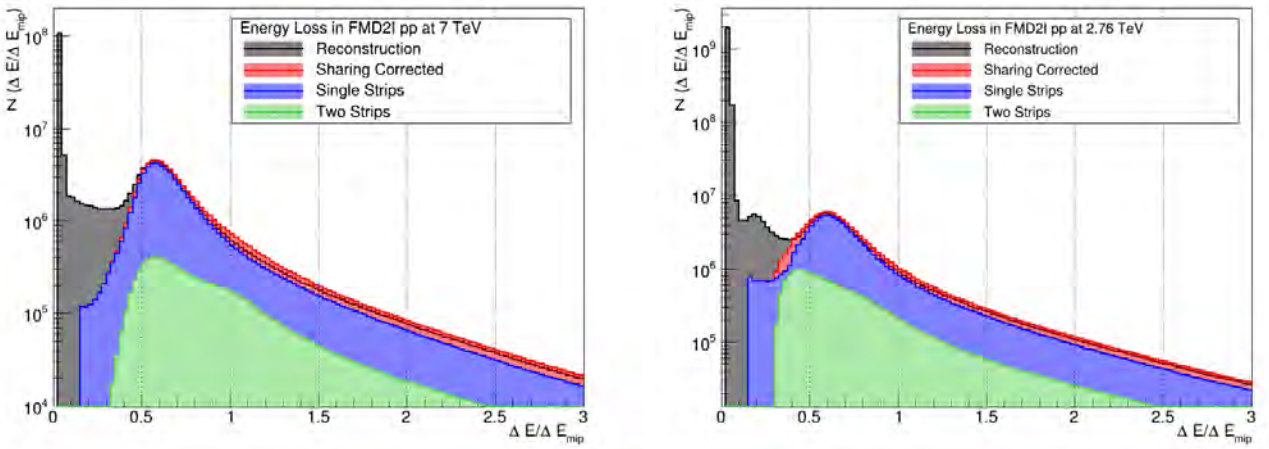
With this simple equivalence, the data are fixed, as it will be shown in the results chapter. Anyway, still remains to be understood were this problem in the trigger definition came from.

For the FMD the problem seems to be slightly more serious, coming from the detector itself and not from a missing and/or faulty trigger definition. Looking at the energy loss in FMD2I at Figure 7.9b, it is clear that something is wrong in the 2.76 TeV data. In the figure, the ESD reconstructed distribution is plotted in black, while the red one is the one used in the analysis, which is sharing corrected. Two additional distributions are shown in the plot, the blue one is the single strip signal, while the green one shows the two strips merged signal, which constitute roughly the 10% of the total events. As shown in Figure 7.9a, the first peak of the Landau, in good data, coincides with the peak of the single strip, meaning that the signals had to be merged. The second Landau peak is instead at around 1.



(a) (b)

Figure 7.8: (a) INEL sample in black and V0AND in red, the bump is present only in the latter; (b) Triggered events, V0AND must be a sub–sample of INEL, because it is an offline trigger, while INEL≡MB$_{OR}$, which is defined online (Section 7.1).



(a) (b)

Figure 7.9: (a) Energy loss distributions in FMD2I for run 126284 with $\sqrt{s} = 7$ TeV in LHC10d, the black distribution is the ESD reconstructed, the red is the analyzed distribution, while blue and green are the single strip and double strip respectively; (b) Energy Loss in FMD2I for run 146824 in LHC11a with $\sqrt{s} = 2.76$ TeV. The bump is present for all data taken at 2.76 TeV in LHC11a, and in all FMD rings.

Starting from the ESD reconstructed distribution in black, there is a small energy deposit right after the pedestal in collisions at $\sqrt{s} = 2.76$ TeV. This is visible also in the single strips deposit, where there is a plateau in the first bins. It must be noted that this small peak is not present in the MC generator energy deposits.

Moreover, looking at the green distribution for energy deposit in two strips, the first peak does not coincide with the peak in single strips. The first peak of the double strip distribution should coincide with the one for the single strip if it has to be merged with the sharing filter algorithm (Section 5.3). Otherwise, if it does not coincide, it means that it comes from a different source. This behavior affects the final raw distribution, making it more flat and moving the mean of it towards the low bins, as it will be shown in the results section.

The source of the bump in the low region still has no explanation. The beam halo cannot be, because the FMD1 is too far away from the interaction point ($3.7 < \eta < 5.1$) to be able to see it, while the bump is present also for it. By now, the most accredited cause is noise internal to the detector, although, still, the shift in the second peak on the two strips distribution is weird. In fact, this might indicate that there is a different signal–like source. In general, the overall 2.76 TeV ALICE runs had numerous problems, also in other detectors.

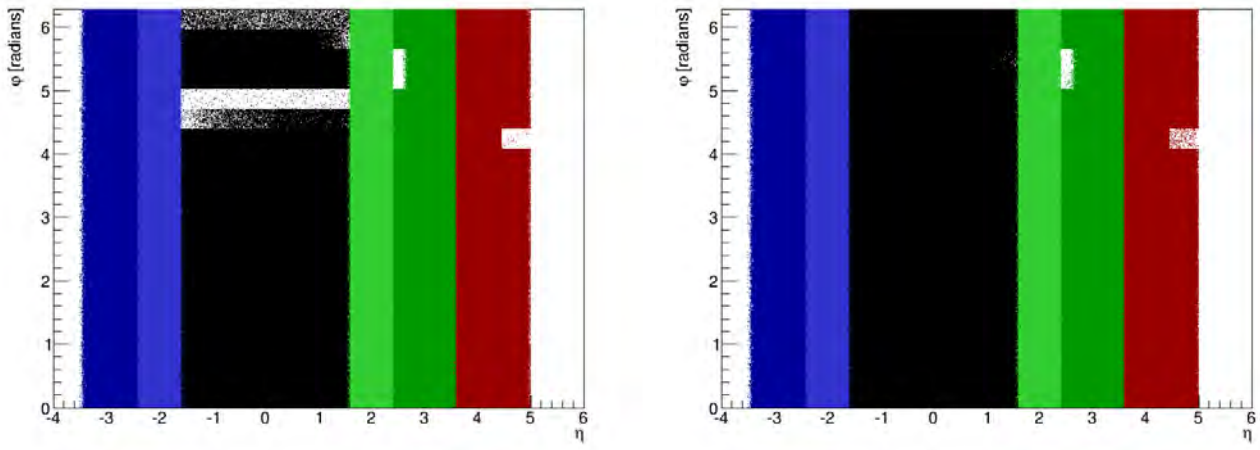
The results with these data and these cuts will be shown in the results chapter, some possible solutions have to be found in the close future, before publication. One idea can be to raise the E_{low} cut for the sharing filter, up to a value of $\backsim 0.3$ (instead of 0.15). Secondly, it is also possible to consider only the single strips without taking into account the merging with the double strips. These possibilities will be tried very soon, but are not included in the thesis.

7.3 Acceptance

Both the FMD and SPD were originally designed to have full azimuthal coverage, but, as can be seen in Figures 7.10a and 7.10b, this is not the case. Concerning the FMD, there is a chip in FMD1I which is not working, and one in FMD2I which is masked out, because the produced signal is faulty and gets closer and closer to the pedestal peak. Additionally, some strips here and there (especially in FMD2I) are not working. For the SPD the issue is more serious for data taken in 2010 and 2011. At that time there were some cooling issues (clogging of the cooling pipes), and part of the SPD had to be turned off to avoid overheating. The problem was solved in 2012, therefore, the azimuthal coverage is fixed in 8 TeV data, as can be seen in Figure 7.10b.

7.4 Secondary Particles

The understanding of the amount of secondaries which hit the FMD and the SPD is crucial, because both FMD and the inner layer of SPD do not have tracking capability. This means that it is impossible to track back a particle which hits the detector and identify if it is a primary or a secondary particle. While for the SPD the secondary con-



(a) (b)

Figure 7.10: (a) Acceptance shown in $(\eta - \varphi)$–histogram for 0.9 TeV, in black the SPD showing holes in 2010 data, in red FMD1I, in green FMD2 (darker inner, lighter outer), in blue FMD3 (darker inner, lighter outer). The acceptance is shown for $-10 < v_z < +10$; (b) Acceptance for 8 TeV data collected in 2012, only the two FMD chips which were tuned off due to malfunctioning are missing.

tribution is minor, many of the supports and cooling systems of the central detectors contribute heavily on the secondary particles which hit the FMD. Figure 7.11 shows a PYTHIA simulation used in conjunction with GEANT3 as transport code to understand the origin of the particles which hit the detector.

In Figure 7.12 the pseudorapidity densities are shown, to get an idea about the percentage of particles which hit the detector. This figure illustrates which are the relative contributions to the measured charged particles. As already pointed out, the major contributors are the beam pipe and the ITS. There is a peak at around $\eta \backsim -2$ were the FMD3I and O are superimposed. Early decays means particles which decay inside the beam pipe, and others include all the other source, which anyway are negligible respect to the above mentioned sources.

It is important to correct for the secondary contributions when deriving the primary spectrum; in this thesis, this is done during the unfolding process, deeply explained in the following sections. In practice, the unfolding corrects for secondaries in the FMD region and for acceptance gaps in the SPD. To be able to unfold properly the primary spectrum, the material composition included in the transport code must be as accurate as possible, and many efforts have been done from 2010 to 2013 to implement it. At this point the knowledge of the material is believed to be correct up to 5-10%. The unknown material contribution is then taken into account in the systematic errors.

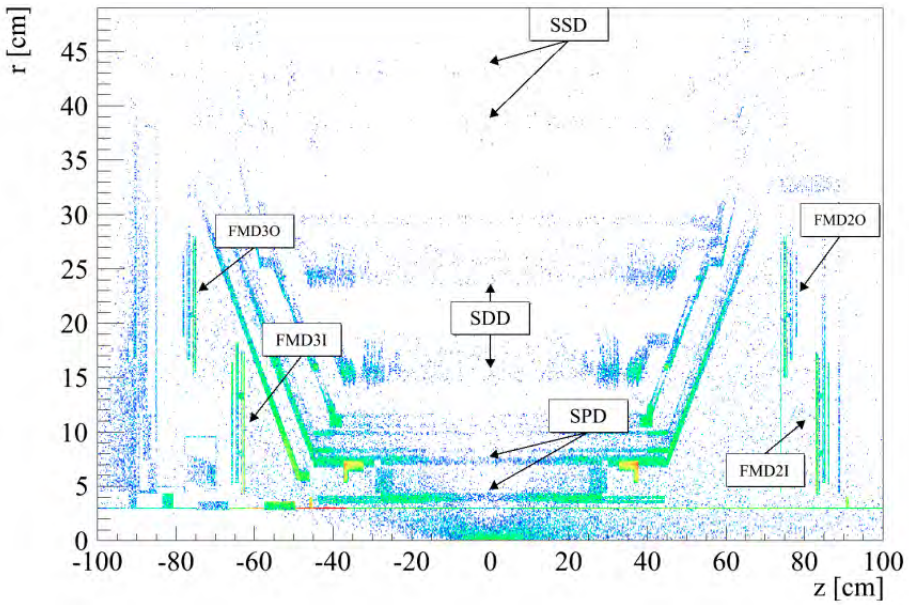


Figure 7.11: Origin of the secondary particles which hit the FMD. The SPD is clearly visible (at $r = 3.9$ and 7.6 cm), as well as the ITS support system and the beam pipe (at 3 cm), together with the FMD rings themselves [63].

7.5 Strange Particles

It is well known that the relative amount of strange particles described in the MC generators is not proper [78]. The major strange particles underestimated are K^0_S by a factor of 2. Λ and $\Xi^{\pm}$ are underestimated by a factor of 3. At this point it is important to look into the decays of these particles: $K^0_S \rightarrow \pi^+\pi^-$, $\Lambda \rightarrow p\pi^-$, $\Xi^+ \rightarrow p\pi^+\pi^-$ and, finally, $\Xi^- \rightarrow p\pi^-\pi^-$. Silicon detectors record only charged particles, therefore, in the primary spectrum only $\Xi^{\pm}$ are missing and not K^0_S and Λ which are neutral. The most relevant contribution is in the secondaries, since in that case all decays are underestimated (p and π).

According to [78], the yields at $y = 0$ are $\langle dN/dy \rangle = 0.184 \pm 0.002 \pm 0.006$ for K^0_S, $0.095 \pm 0.002 \pm 0.006$ for Λ and $\overline{\Lambda}$. For charged $\Xi^{\pm}$ it is $0.0101 \pm 0.0020 \pm 0.0009$. Therefore, only 3% of the total strange–particle yields are charged and in the primary spectrum, it is clear that the effect is negligible. The same cannot be said for the secondary content, where all the decays are included.

The matrix which corrects the raw distribution is build with MC truth primaries on the x–axis and MC reconstructed primaries + secondaries on the y–axis. The flat MC production used in the analysis does not have physical meaning (as explained in Section 6.1), therefore the information of secondaries from weak decayed particles is not present. On the other hand, from studies done in [62], the overall effect in the mean of the distribution should be less than 2% total enhancement of particles (these studies were done for the pseudorapidity density). Unfortunately it is not proper to throw 2% more random particles, because the distribution of only strange particles is not uniform like the regular one, as can be seen in Figure 7.13.

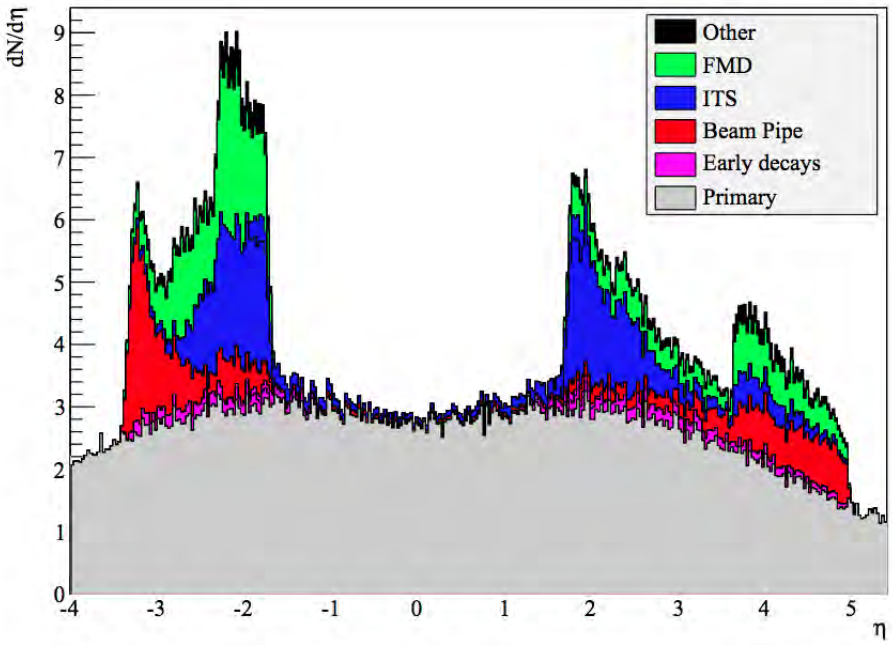


Figure 7.12: PYTHIA and GEANT3 simulated pseudorapidity densities, which identify the percentages of particles hitting SPD and FMD. For SPD the contribution from secondaries is very small, while for FMD, the secondaries can be more than the double of the primaries [62].

From the figure, it is clear that the two normalized distributions are different and it is not possible to add randomly more particles in the global one, because strange particles (in red) are distributed differently. At this point, there is no correction which can be applied to take into account this effect. It must be anyway noted that the effect is not big, but, certainly, for Run II data it would be better to have improved description of strange particles in MC generators. There is presently an analysis group in ALICE which makes such studies.

7.6 Unfolding

The main ingredients for this analysis are the raw data distributions and a matrix which allows to unfold the true primary spectrum from the raw one. In this section, the procedure to unfold and the matrix itself will be deeply described. Ideally, if a detector would be perfect, it would be possible directly to measure the true primary spectrum, but, of course, this is not the case. As seen in Section 7.3, the SPD has several acceptance gaps which must be filled in the final spectrum; in addition, FMD is hit by a lot of secondaries which have to be removed from the measured spectrum (Section 7.4).

Detector Response

The response of the detector is determined by a matrix R_{mt} which identifies the probability that an event with true multiplicity t is measured as an event with the multiplicity m. This matrix, which will be called response matrix in the following, is simply derived

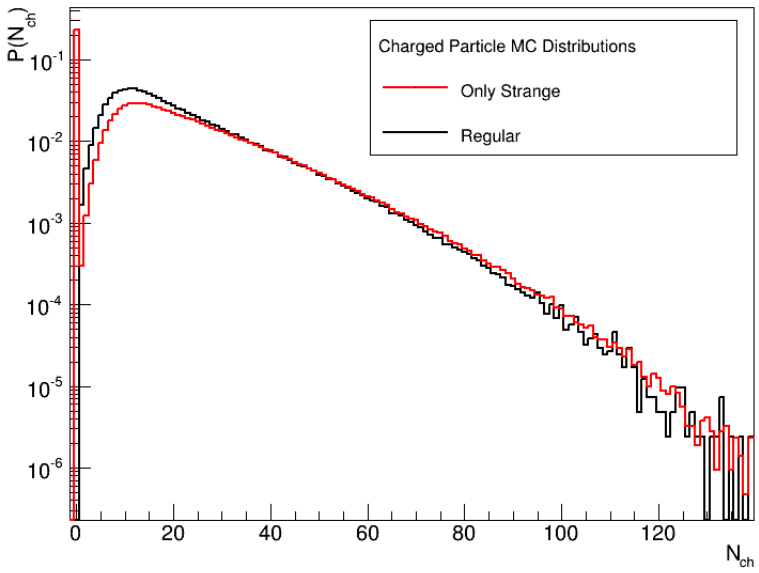


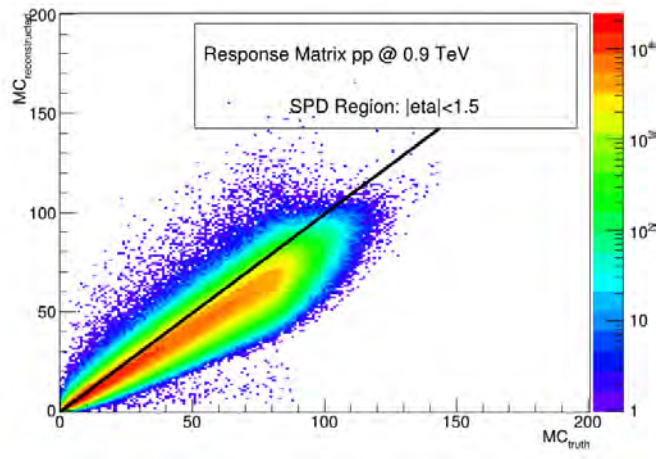
Figure 7.13: In red only strange particles from PYTHIA are included, while in black all particles are in. The mean is lowered by 12% and there is a peak at zero. In the red distribution, the peak identifies the probability of finding 0–strange particles.

by a MC generator distribution (in this case a flat one 6.1) in the same pseudorapidity range and vertex selection (true vertex selection from the MC). Then, it is propagated through the code and through the same reconstruction steps as data themselves.

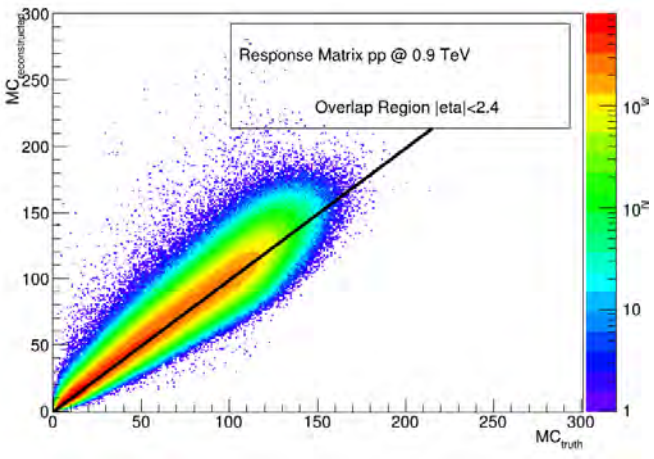
In Figures 7.14, three different response matrices are plotted for different pseudorapidity η ranges. The first, 7.14a, is for the region which will be called pure SPD in the following, because it is derived in the widest range where there is no FMD counts in, $|\eta| < 1.5$. Here the response matrix corrects for the acceptance gaps of the SPD, in fact there are less particles in the reconstructed MC (y–axis) with respect to the true MC (x–axis). The slope is 0.82, which means that 18% of SPD acceptance holes will be filled during the unfolding procedure.

Figure 7.14b shows the response matrix obtained for what will be called the overlap region, namely the region where $|\eta| < 2.4$ and around half of the counts come from SPD and half from the FMD.
In this region the unfolding does not change the raw distribution much. In fact, the slope is 1.02. The matrix which certainly produces the major difference from raw data to unfolded ones is the one for the full range, $-3.4 < \eta < +5.1$, shown in Figure 7.14c. In this case the slope is 1.20.

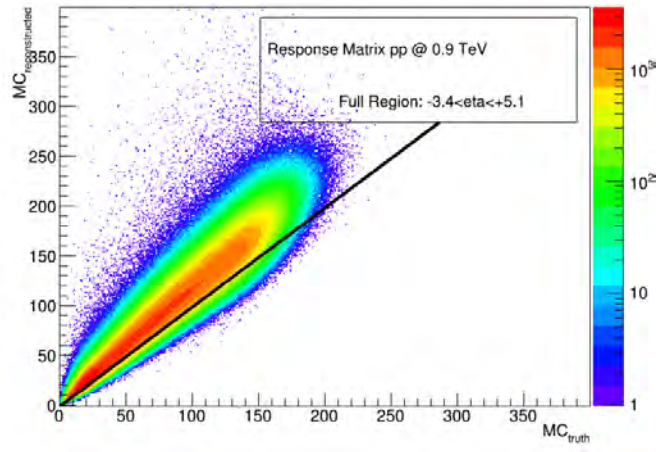
Ideally, given a square response matrix R and a true distribution T, the measured distribution can be written as $M = RT$.
Therefore, formally, in the case in which M and R are available, to obtain the true distribution a simple inversion of the matrix has to be applied $T = R^{-1}M$. But this simple procedure creates problems, firstly, the response matrix might be singular and there-



(a)



(b)



(c)

Figure 7.14: Response matrices obtained propagating MC generated events, in this case wiht PYTHIA ATLAS CSC tune flat, through the reconstruction code. (a) Matrix for the pure SPD region; (b) Overlap between SPD and FMD region; (c) Matrix for region where the majority of the counts are from the FMD. The diagonal is plotted as a black line.

fore one cannot invert it. Also if it is not singular, the resolution of the spectrum will create non–physical fluctuations as we shall see in the following.

Assuming a known true distribution T and a non–singular R, the derivation of M is straight forward ($M = RT$), then, according to the M distribution a sample of N events is produced, $\widetilde{M}$. Now, deriving $\widetilde{T} = R^{-1}\widetilde{M}$, the sample, as illustrated in Figure 7.15, will have a lot or oscillations due to the resolution of $\widetilde{M}$ [79]. It is clear that $\widetilde{T}$ is not the true distribution.

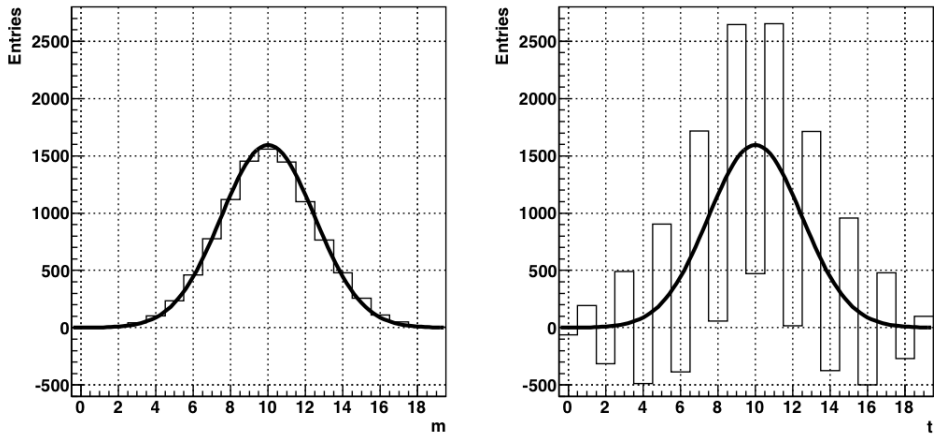


Figure 7.15: In the left plot, the measured distribution and the binned sample produced from it. On the right, the true initially known distribution and the binned histogram with $\widetilde{T}$ derived from the inversion if the response matrix [79].

In the following, two methods to smooth $\widetilde{T}$ are presented, and they both rely on an a priori knowledge of the distribution. The first method, based on Bayes Theorem is used to derive the final results, while the χ^2 minimization is used to estimate the systematic uncertainty generated by the unfolding method itself.

Bayesian Unfolding

The first unfolding method which is used is the one based on the Bayes' Theorem, which states that the conditional probability $P(A|B)$of A if B is true can be written as:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{7.3}$$

where $P(A)$ and $P(B)$ are the independent probabilities of A and B being true, and $P(B|A)$ is the probability of B if A is true. A can be identified as a collision with true multiplicity, while B is a measured event with measured multiplicity; clearly, the conditional probability $P(B|A)$ is the response matrix of the detector. Knowing these quantities, $P(A|B)$ can be computed.

Rewriting the Bayes' Theorem using the nomenclature employed in the previous section, $P(B|A) \equiv R_{mt}$, $P(A) \equiv T$ and $P(B) \equiv M$. Therefore, the theorem can be

restated as:

$$\widetilde{R}_{tm} = \frac{R_{mt}P_t}{\sum_{t'} R_{mt'}P_{t'}} \tag{7.4}$$

where P_t is an a priori guess of the true distribution, which is obviously not known in advance. Now the unfolding procedure simply computes

$$U_t = \sum_m \widetilde{R}_{tm}M_m \tag{7.5}$$

If $U_t \equiv P_t$ then U_t is the true distribution, otherwise, the process is reiterated using U_t instead of P_t. The process can be iterated until convergence or for a fixed number of iterations.

There is another parameter which can be defined in the Bayesian unfolding, the smoothing parameter α. It is, indeed, possible to introduce a smoothing from one iteration to another to smooth possible statistical fluctuations:

$$\widetilde{U}_t = (1-\alpha)U_t + \frac{\alpha}{3}(U_{t-1} + U_t + U_{t+1}) \tag{7.6}$$

Whether to apply the smoothing or not, or to fix the number of iterations will be discussed in the following sections. In Figure 7.16, a test of the Bayesian unfolding is performed. To do so, a MC reconstructed (brought through the code) distribution has been unfolded using a response matrix where the true distribution is well known. After this, the resulting unfolded distribution must be equivalent to the known true, and this is the case. Actually the agreement is perfectly 1, showing that this methodology is very stable and accurate.

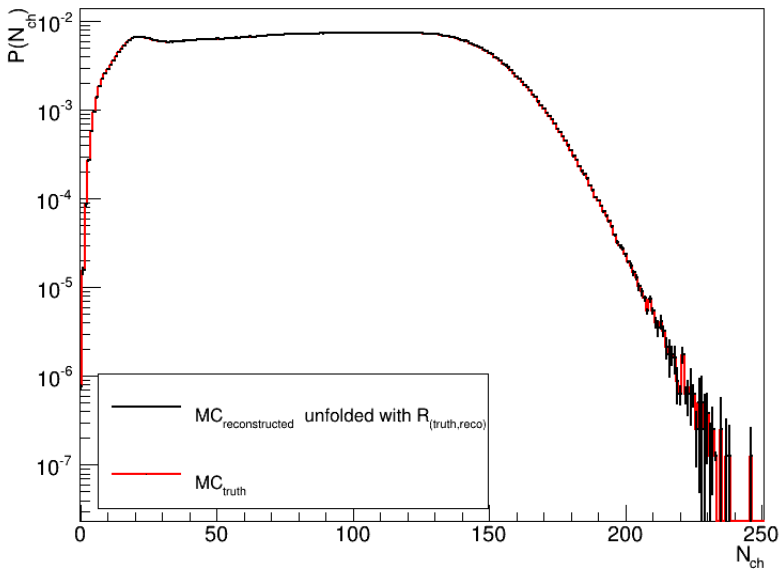


Figure 7.16: Bayesian unfolding test, performed using a flat distribution. The methodology works very well, since the agreement between the two distributions is one.

χ^2 Minimization Unfolding

The second unfolding methodology which was studied in this thesis is the minimization of a χ^2 function of the form:

$$\widehat{\chi^2}(U) = \sum_m \left(\frac{M_m - \sum_t R_{mt} U_t}{e_m} \right)^2 \tag{7.7}$$

where e_m is the error on the measured spectrum and the other quantities are the ones already defined in the previous section. This simple function numerically inverts the response matrix. Therefore, the fluctuations described in the previous sections are present. To solve this problem, a regularization function $P(U)$ can be added to stabilize the distribution, and the final function will look like:

$$\chi^2(U) = \widehat{\chi^2}(U) + \beta P(U) \tag{7.8}$$

where β is a weight applied to regularization term $P(U)$.

Three functions have been tested in this thesis, a constant, a linear and a logarithmic one:

$$P_1(U) = \sum_t \left(\frac{U'_t}{U_t} \right)^2 = \sum_t \left(\frac{U_t - U_{t-1}}{U_t} \right)^2 \tag{7.9}$$

$$P_2(U) = \sum_t \left(\frac{U''_t}{U_t} \right)^2 = \sum_t \left(\frac{U_{t+1} + 2U_t - U_{t-1}}{U_t} \right)^2 \tag{7.10}$$

$$P_3(U) = \sum_t \left(\frac{(\ln U_t)''}{\ln U_t} \right)^2 \tag{7.11}$$

the choices of the weight β will be discussed in the quality tests, in the following section. Also this method works well to unfold, although, especially in the case in which the pseudorapidity range is very wide, which means that the distribution spans many bins, it is very difficult to stabilize it. The risk is to include too many fluctuations, or in the contrary, to overweight the unfolded distribution, making it too flat in the low bins. In Figure 7.17, can be see that there is very good agreement between MC truth and the unfolded distribution with the minimization of the χ^2.

Unfolding Quality Checks

Several tests have been done to check the quality of the unfolding methods. The distribution behaves differently according to the region considered, therefore, to check the quality, it is very useful to divide the distribution into three regions:

Peak Region It is the region in which the distribution rises, reaches the peak and decreases rapidly. Here these parameters are chosen: $N_{ch} = [0; 2 \cdot N_{ch,peak}]$

Exponential Region Is where the exponential slope of the distribution is more or less constant: $N_{ch} = (2 \cdot N_{ch,peak}; N_{ch}(P(N_{ch}) < 10^{-4})]$

Low Statistics Region As the name says, it is the region with low statistics: $N_{ch} = (N_{ch}(P(N_{ch}) < 10^{-4}); N_{ch,tot}]$

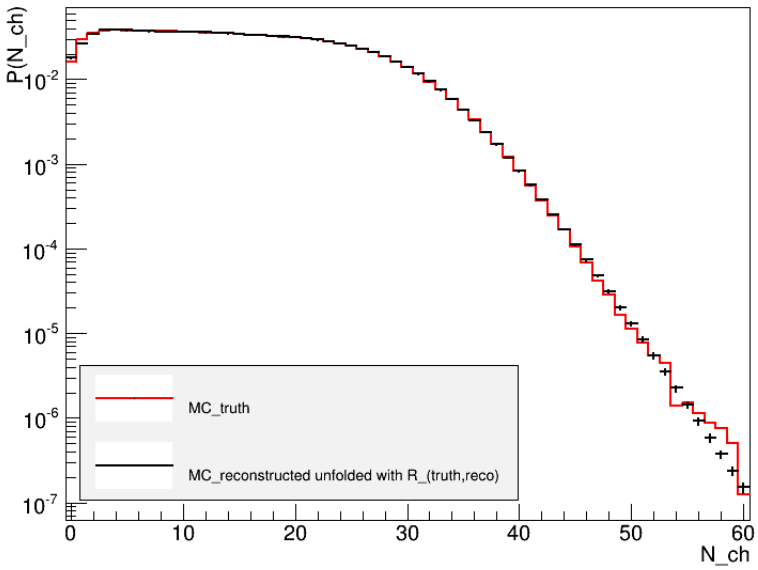


Figure 7.17: χ^2 minimization unfolding test, performed using a flat distribution. The methodology works very well, since the agreement between the two distributions is nearly one.

The main parameter to evaluate the unfolding methods is:

$$Q_1 = \frac{1}{DOF} \sum_t \left| \frac{T_t - U_t}{e_t} \right| \tag{7.12}$$

where e_t are the errors on U_t, Q_1 is normalized to the number of degrees of freedom, which vary according to the number of bins in the selected region. Secondly, a traditional reduced χ^2 test has been done. Both Q_1 and χ^2/DOF should be around 1, if the methods are working well.

In Figure 7.18, both parameters are plotted for Bayesian unfolding using 10 iterations and smoothing parameter $= 1$ (better settings). For the χ^2 minimization unfolding, only the logarithmic function has been chosen, because the linear and constant did not perform as well as it. For Q_1 on the peak and central region, as well as for the reduced χ^2, the minimization of χ^2 function unfolding, using a logarithmic parameterization function, behaves better than the Bayesian unfolding, for weights $< 10^4$. Instead, for the low statistics region the χ^2 minimization unfolding performs better for all β weights, but it is comparable with the Bayesian performances.

It must be said that in order to cancel the fluctuations coming for the χ^2 unfolding, a weight $> 10^3$ must be chosen. Therefore, it can be stated that, for $\beta = 10^3$ to 10^5, the Bayesian performs better or comparable to the χ^2 minimization unfolding. In Figure 7.19, the data for $\sqrt{s}$ =7 TeV are unfolded with the better settings for both methods. It can be seen that residual fluctuations are present in the χ^2 minimization unfolding, since it is difficult to stabilize it for distributions with a lot of bins. The previous ALICE publications [58, 59] used a lower number of bins. Therefore χ^2 unfolding was used.

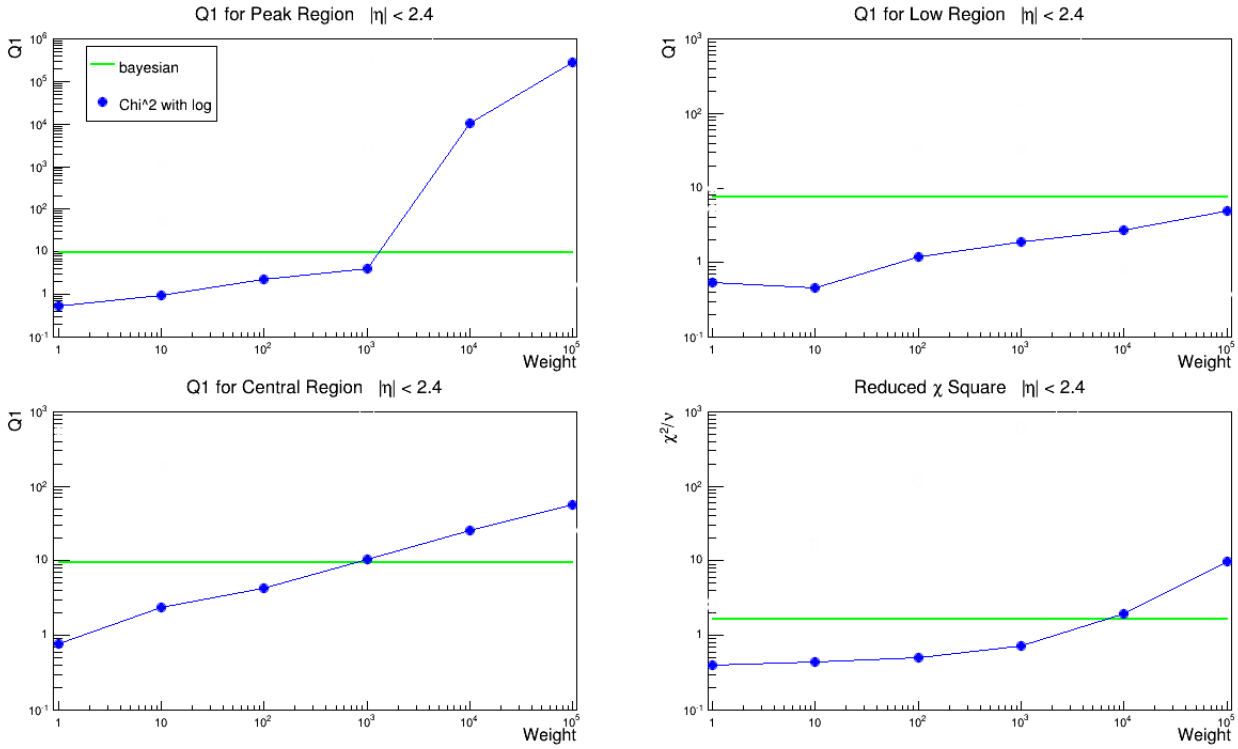


Figure 7.18: Quality parameters Q_1 and χ^2/DOF for χ^2 minimization and Bayesian unfolding. On the x-axis, the weight of the χ^2 is plotted and several order of magnitudes have been probed.

CMS, instead, used Bayesian [56]. In this thesis, the Bayesian unfolding is chosen to derive the results, while the χ^2 is computed for systematic checks.

7.7 Event Selection Efficiency

The efficiency of the event selection is shown in Figure 7.20 for the four energies measured in this thesis. Both INEL and NSD efficiencies are plotted, and the values are the average between the PYTHIA and PHOJET generators. The error bars are the difference between the two. Regarding the INEL sample, the error bars for 0.9 and 8 TeV are very large, because PYTHIA goes to efficiency one very quickly, the opposite of what PHOJET does.

On the other side, for NSD sample there are a lot of statistical fluctuations for 8 TeV, in fact, for the other energies the tuning for diffraction done using the using the Kaidalov–Poghosyan model [67] (the measured diffraction cross–sections and the shapes of the diffractive masses M_X defined in Section 1.3 were tuned) was done centrally. For 8 TeV, the tuning had to be done afterwards, in this work. Since the energy difference between 7 and 8 TeV is fairly small, the MC generators are tuned with the 7 TeV diffractive mass weights. Basically, one has to tune the diffractive mass distribution, which, otherwise are not properly implemented in regular MC. Doing this procedure offline reduces the events sample, and therefore, there is a increase of the

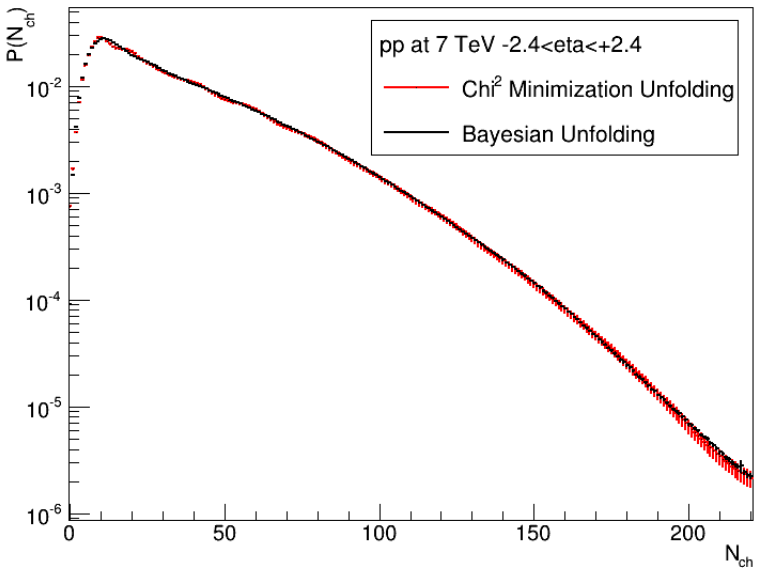


Figure 7.19: 7 TeV data unfolded with both methods, the χ^2 unfolding shows fluctuations in the peak region. The ratio to these two distributions are employed in the systematic uncertainties computation.

statistical fluctuations for 8 TeV.

The efficiency is computed in this way:

$$\epsilon_{NSD} = \frac{N_{ch,reco}(NSD\,\&\,|v_{z,reco}| < 4)}{N_{ch,gen}(NSD\,\&\,|v_{z,gen}| < 4)} \tag{7.13}$$

$$\epsilon_{INEL} = \frac{N_{ch,reco}(INEL\,\&\,|v_{z,reco}| < 4)}{N_{ch,gen}(INEL\,\&\,|v_{z,gen}| < 4)} \tag{7.14}$$

where the nominator contains the reconstructed events with the selected trigger (NSD or INEL) and the selected cut in the reconstructed vertex. The denominator, instead, contains the same quantities for MC truth, therefore from the generator.

At this point, the unfolded distribution has to be corrected for vertex and trigger inefficiency:

$$U_t = \frac{U_t^*}{\epsilon} \tag{7.15}$$

As pointed out several times by now, ALICE is a lot more inefficient for NSD than for INEL, because the triggering on Non–Single-Diffractive events is not very accurate. With the new ALICE Diffractive detector the better trigger efficiency for Single Diffractive events will improve the capabilities of ALICE to detect low mass diffractive systems, getting closer to the values which CMS is also capable to get.

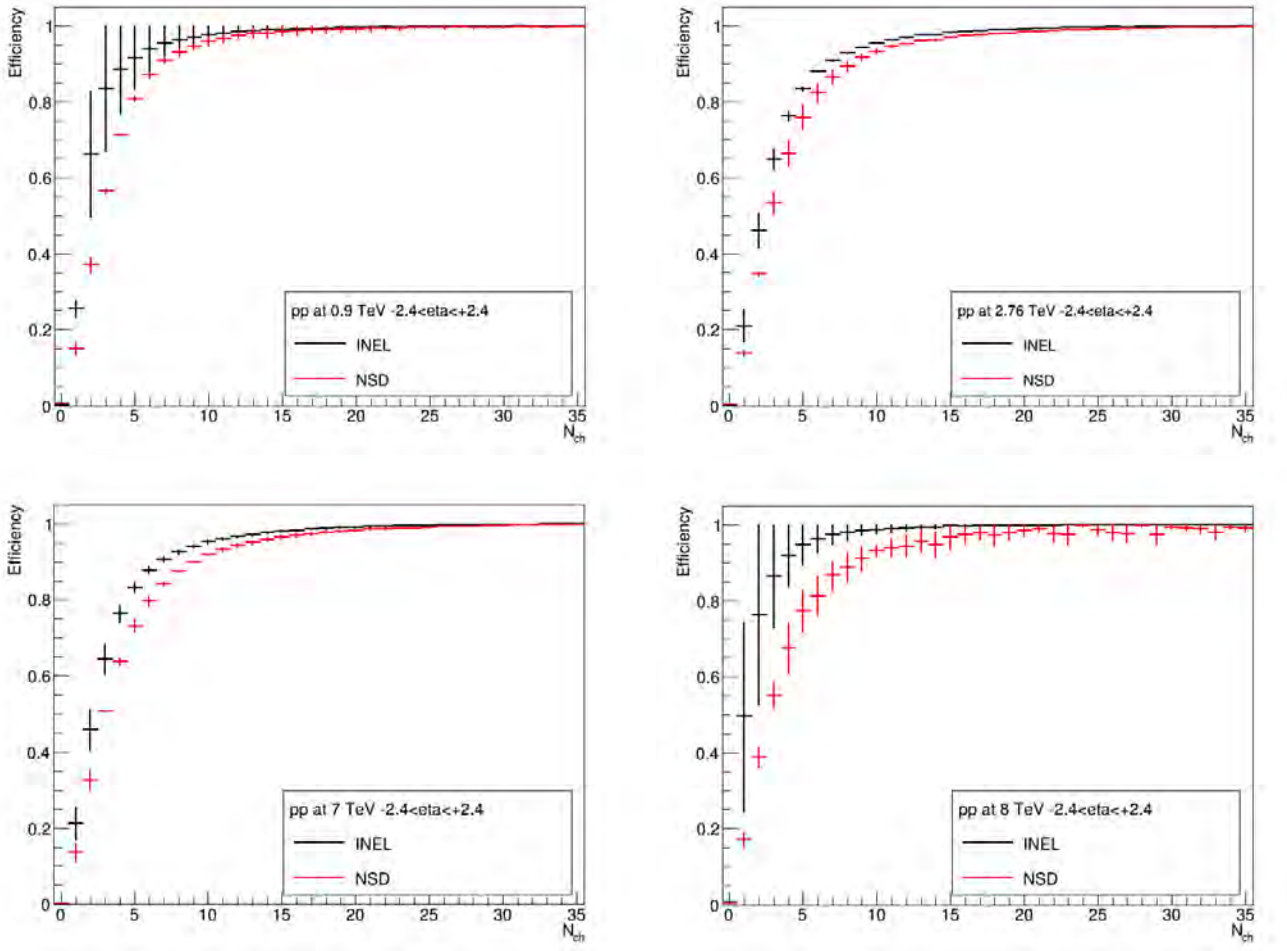


Figure 7.20: Efficiencies for all four energies, for INEL and NSD sample as a funcion of the primary charged particles for $|\eta| < 2.4$. The $\sqrt{s}$ =8 TeV sample shows more statistical fluctuations, because the NSD efficiency needed to be tuned and weighted by hand, therefore, was not directly centrally available.

7.8 Summary of Analysis Parameters

Due to the fact that this chapter is fairly long, the parameters used for pp are summarized here:

Trigger Selection Three trigger classes are considered, INEL, which corresponds to the online Minimum Bias requirement, INEL> 0, which asks for at least one hit in the SPD, and NSD, which requires a coincidence of the two V0s, removing the majority of Single–Diffractive events.

Background Selection Pile up events are removed requiring that the additional vertex is at least 0.8 cm away from the primary vertex, and that it is reconstructed with at least 3 tracks in the SPD.

Acceptance and Vertex Selection The acceptance between SPD and FMD is maximized requiring $-4 < v_z < +4$.

FMD Cuts The cuts in the energy distribution are $E_{low} = 0.15$ for Inner rings and 0.18 for Outer rings. While $E_{high} \equiv E_{hit} = E_{MPV} - \xi$

Secondary Particles The secondary particles are removed with the unfolding procedure.

Strange Particles There is not correction for the lack of strange particles in the MC generators, but the effect is very small ($\backsim 2\%$) and, therefore, contained in the systematic uncertainties.

Detector Response Is computed using the PYTHIA generator flat ATLAS–CSC tune, which provides very good statistics.

Unfolding Performed with the Bayesian method which behaves better in MC truth comparisons and does not show fluctuations, also for large number of bins.

Efficiency Correction Computed from the average efficiencies for PYTHIA Perugia 0 and PHOJET for INEL and INEL> 0 sample. Additionally, for NSD, the MC generators were tuned for diffraction mass distributions.

Chapter 8

Systematic Uncertainties for pp Collisions

In this chapter the systematic uncertainties for pp collisions are discussed. There are several sources of error in the analysis, which are treated in slightly different ways according depending on whether they come from the raw data or from the unfolding method. The error on the efficiency correction is also treated separately, as it was described in Section 7.7. At the end, the systematic uncertainties are summed in quadrature, and the results presented in the following chapter will contain the total systematic uncertainty combined with the statistical uncertainty.

8.1 Run–to–Run Fluctuations

The fluctuations from run to run have been checked for all four energies in the following way: the raw data are split into two separate samples which then are unfolded with two different response matrices composed of only the runs which are anchored to the raw–data–sampled runs. The two different distributions are then averaged bin by bin and divided by the sum of the two distributions. For $\sqrt{s}$ =0.9 and 2.76 TeV, the two samples of raw data are chosen from the same period, while for 8 TeV are chosen from two different periods, where data are taken with the same polarity of the L3 magnet. In these three cases, the run–to–run fluctuations are found to be negligible up to a certain value where the statistical fluctuations play a role.

For $\sqrt{s}$ =7 TeV, instead, data are sampled within two different periods where there was a change in the polarity of the magnet. For this reason, the run–to–run fluctuations are relevant ($\backsim 10 - 15\%$) in the low bins. In Figure 8.1, the two samples used for 0.9 TeV are plotted, both the raw distributions and the unfolded spectrum. In the second sample in black, there is clearly more pile–up in the last bins. In fact, this sample contains runs which have higher average number of interaction per bunch μ of 0.03, while for the red sample it is 0.02.

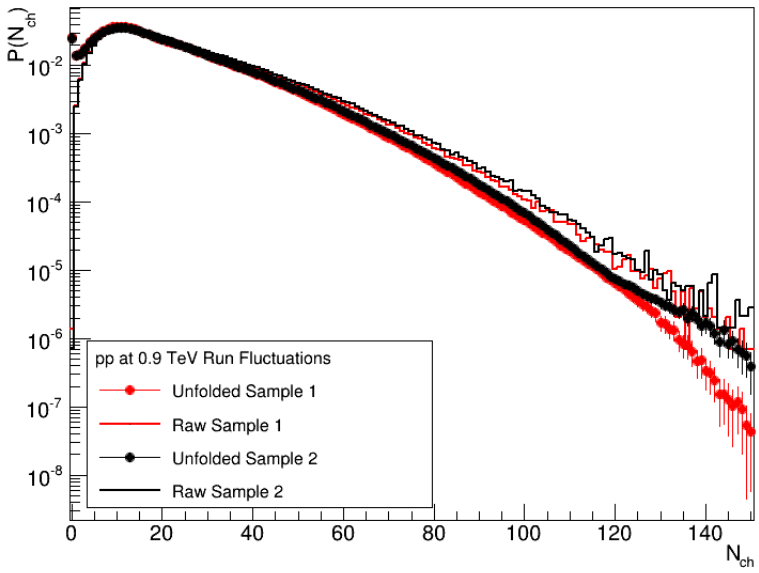


Figure 8.1: The 10 runs used to obtain the spectrum for 0.9 TeV in $|\eta| < 2.4$ are splitted into two samples, in order to observe the run–to–run fluctuations.

8.2 Material Budget

As already shown in Section 7.4 and Figure 7.12, there is a fraction of unknown material between 5 and 10%. To derive the uncertainty due to this effect, the raw distribution is unfolded with a response matrix obtained adding 10% material, uniformly in pseudo-rapidity. Then, the systematic effect is obtained by dividing the new distribution with more material by the regular one.

There are a certain number of caveats in the derivation of this uncertainty, firstly, the +10% material distribution was done only for 0.9 TeV, therefore, in this case, the uncertainty has to be propagated to all the other three energies. This does not create problems, because the missing information on the material and the percentage of secondaries produced respect to primaries are basically the same for all energies[1]. The second issue comes from the fact that the production used to derive it is fairly old (from 2011), when the latest implementations on the material geometry (done by December 2013) were not yet implemented. This affects the production by 5–10% in the tails.

The third problem comes from the fact that the production is not flat, but is produced using the regular PYTHIA Perugia 0 tune generator. This implies that the response matrix population and statistics is not good enough to reach the high multiplicity bins, as can be seen in Figure 8.2. A temporary solution has been to interpolate between the last bins of the systematics plot. Of course, in light of the imminent publication of this work, a new flat production needs to be required, since the fit approxi-

[1]This can be said looking at the slope of the response matrices for different energies, it is the same among all.

mation is not accurate enough and, in particular, the latest geometry implementations need to be taken into account.

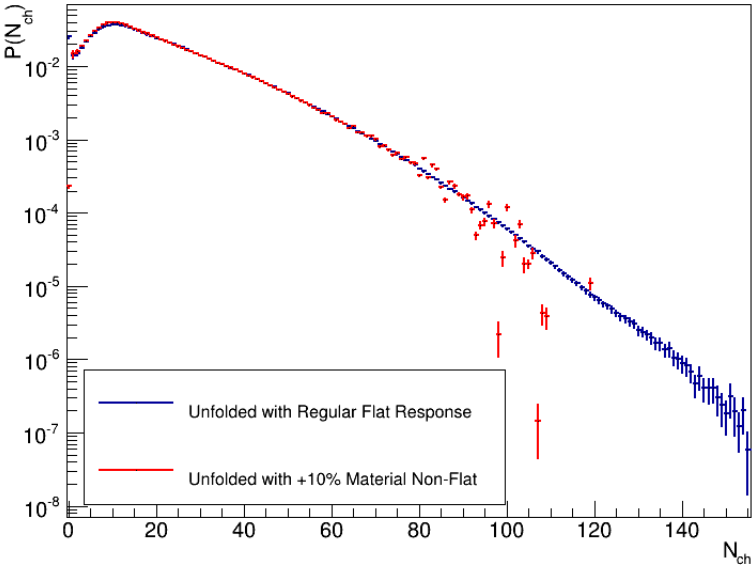


Figure 8.2: +10% material spectrum versus regular spectrum, computed with a flat production. In the red distribution is not possible to reach the high multiplicities. These distributions are derived for 0.9 TeV in $|\eta| < 2.4$. Moreover, the old material production does not reproduce the 0–bin. The ratio between the two distributions is plotted in the systematic uncertainties Figure 8.6.

8.3 Unfolding Method

The uncertainty in the unfolding method has been evaluated by comparing the distributions obtained using the Bayesian (Section 7.6) and the χ^2–minimization (Section 7.6) unfolding methods. Due to the fact that the second methodology produces a lot of oscillations in the final spectrum, the systematic uncertainty was derived finding the maximum oscillation between the two distributions and propagating it to all the bins. Therefore, the error estimated from this source is not dependent from the multiplicity N_{ch}. The comparison between the distributions unfolded with the two methods with better settings are shown in Figure 8.3. The oscillations in the χ^2–minimization unfolding are visible in the peak region.

8.4 Sharing/Poisson Cuts

The last systematic check was done by varying the cuts used for the sharing filter and the Poisson counting method. As explained in Section 5.3, there is a sharing filter cut which is equivalent to the cut for the Poisson algorithm $E_{high} \equiv E_{hit} \equiv E_{MPV} - \xi$. Looking at the behavior of E_{MPV} and ξ within one period (Figure 8.4), it is estimated that the maximum variation is of about 5%. Therefore two distributions have been

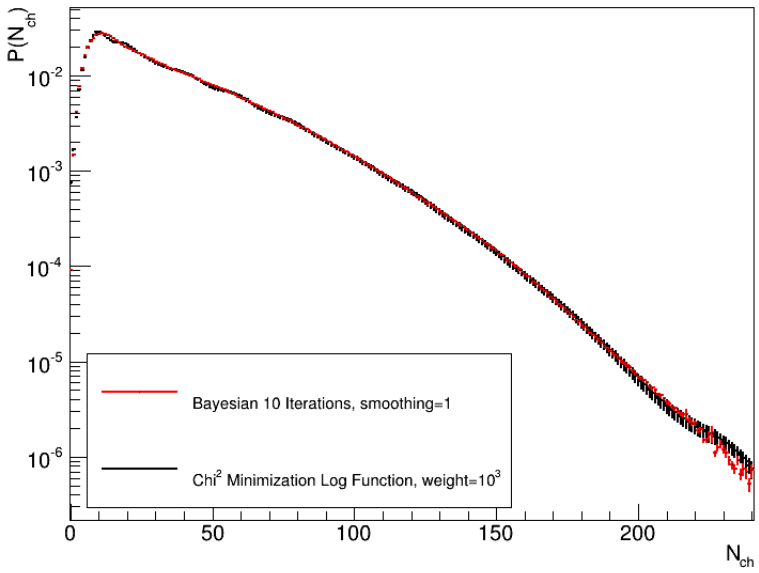


Figure 8.3: Bayesian and χ^2–minimization unfolding with better settings for 7 TeV data and $|\eta| < 2.4$. The maximum difference between the two methods is plotted in the systematic uncertainties Figure A.2

studied, one by raising the Sharing/Poisson cut to $1.05 \cdot (E_{MPV} - \xi)$ and one by lowering it to $0.95 \cdot (E_{MPV} - \xi)$. The systematic uncertainty is derived comparing these two distributions (Figure 8.5).

8.5 Systematic Uncertainties Overview

The four systematic errors are then summed in quadrature, together with the uncertainty from the efficiency correction. The latter was already described in Section 7.7. It is derived by averaging the efficiency calculated from PYTHIA Perugia 0 (6.1) and PHOJET (6.2) diffraction tuned generators and will influence only the first bins, since it is in these bins that the detection is inefficient. Since it behaves differently from the other uncertainty sources and the plots were already shown in the above mentioned section, it is not included in the overview below. The Figures 8.6 show the single contributions from run–to–run fluctuations, material, unfolding and cut variation uncertainties and the total systematic effect for 0.9 TeV data, in this form:

$$\delta_{TOT} = \sqrt{\delta^2_{fluct} + \delta^2_{mat} + \delta^2_{unf} + \delta^2_{cut}} \tag{8.1}$$

with

$$\delta_x = \left| 1 - \frac{P(N_{ch,x})}{P(N_{ch,ref})} \right| \tag{8.2}$$

where $P(N_{ch,ref})$ is the nominal reference distribution, while $P(N_{ch,x})$ is the one obtained varying the uncertainty parameters.
Three particular regions are considered: pure SPD ($|\eta| < 1.5$), overlap region ($|\eta| < 2.4$)

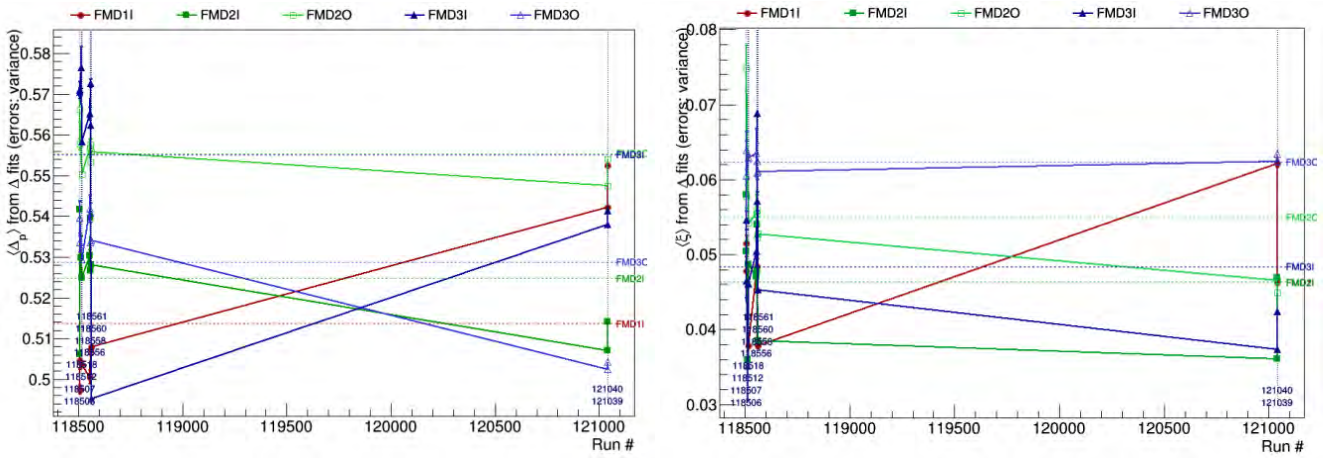


Figure 8.4: Maximum variations of E_{MPV} and ξ in the same period (LHC10c for 0.9 TeV data). The variation has been evaluated among the single FMD rings.

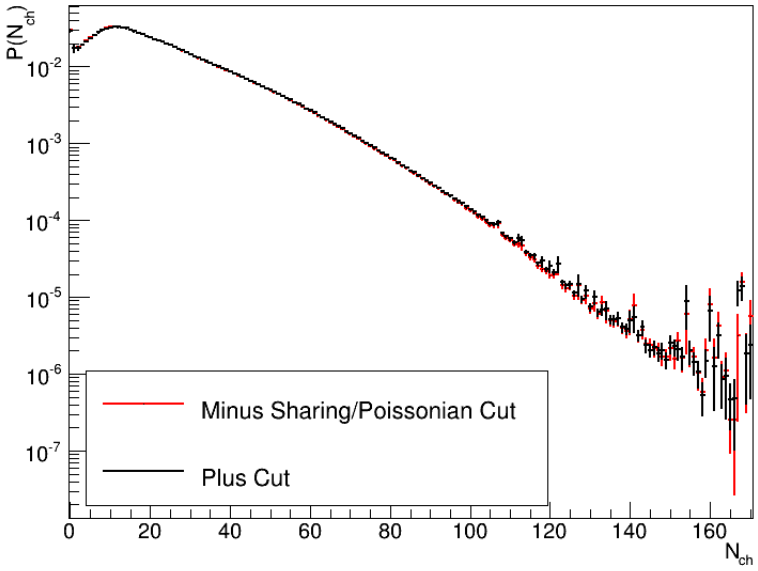


Figure 8.5: Distributions obtained lowering or raising the sharing/Poisson cut by 5% for 0.9 TeV data in $|\eta| < 2.4$. The ratio between the two distributions is plotted in the systematic uncertainties Figure 8.6.

and full region ($-3.4 < \eta < +5.1$). The plots for other pp energies can be found in Appendix A.

One can notice that, overall, the most significant error source comes from the material budget, pointing out the need of a flat production as accurate as possible. In fact, the fit which was done for this thesis is certainly not desirable for the imminent publication. Moreover, the systematic effects depend strongly on the multiplicity bin considered. In particular, the systematics of the 0–bin do rely on inefficiencies and the quantities were presented in Section 7.7. In the event efficiency calculations, the dependence on the generator used to compute it (PYTHIA or PHOJET) is taken into account.

The $P(N_{ch})$ region with lower uncertainty is the one around the mean. In the systematics plots, the region with low statistics (where $P(N_{ch}) < 10^{-4}$) is underlined. After this value, the uncertainties are very large. This happens when studying probability distributions, which are normalized. In fact, a small change ($\backsim 5\%$) in the mean which has a probability of at least three orders of magnitude higher than the low statistics region, will easily produce 30% change in the high multiplicity tail. Nevertheless, the systematics computed in this thesis are comparable to the ones of the already published ALICE papers and with the on–going analysis in the central SPD region [76].

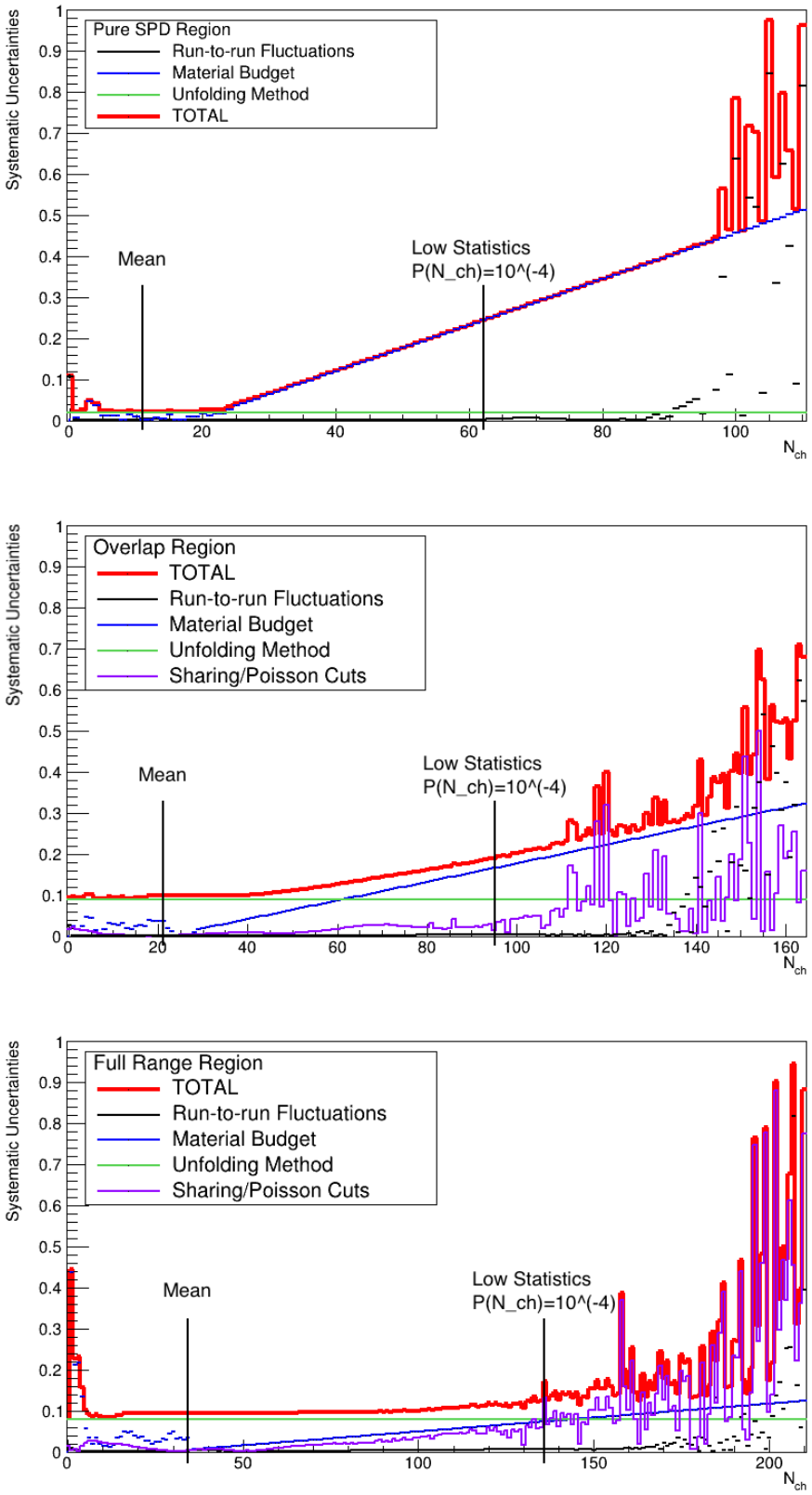


Figure 8.6: Systematic uncertainties overview plots for 0.9 TeV data:
$\delta_x = |1 - P(N_{ch,x})/P(N_{ch,ref})|$.
Top plot: $|\eta| < 1.5$, middle: $|\eta| < 2.4$ and bottom plot: $-3.4 < \eta < +5.1$. See text for details.

Chapter 9

Results for pp Collisions

This chapter contains the main results of this thesis for pp collisions, and it is divided into three sections. Firstly, the multiplicity distributions will be shown for the three trigger classes analyzed for each of the four pp collision energies. The results will be then compared with the previous LHC measurements computed by ALICE in the central region (e.g. obtained only with the SPD) and CMS. They will be compared also with two MC generators, PYTHIA and PHOJET and with the IP–Glasma model. The second section will show the results of KNO scaling, while the third one will show the double NBD fits.

9.1 Multiplicity Distributions for pp

In Figure 9.1, the obtained multiplicity distributions for 0.9 TeV with NSD trigger are shown for eight pseudorapidity ranges. The first three, $|\eta| < 0.5$, $|\eta| < 1.0$ and $|\eta| < 1.5$, are obtained using the SPD clusters. The other five, $|\eta| < 2.0$, $|\eta| < 2.4$, $|\eta| < 3.0$, $|\eta| < 3.4$ and $-3.4 < \eta < +5.1$, are computed with SPD clusters and FMD. The distributions are multiplied by multiples of 10, in order to fit in the same plot without overlapping.

Two things can be noticed from the figure. Firstly, using the FMD detector to enlarge the pseudorapidity window allows to probe twice as large multiplicity. In fact, using only the SPD, it is possible to measure up to 100 particles, with the FMD, 200 particles at 0.9 TeV are reached. Secondly, one can notice that the 0–bin of the full range distribution in black is lower relative to $N_{ch} = 1$.
In the INEL sample, the 0–bin and 1–bin are highly dominated by Single–Diffractive events, while the 1–bin and 2–bin are dominated also by Double–Diffractive events, as can be seen in Figure 1.6a. In the Non–Single–Diffractive trigger, most of the SD events are removed, therefore, the remaining inefficiencies come from DD events. The Figures for INEL and INEL> 0 trigger classes for 0.9 TeV can be found in Appendix B.

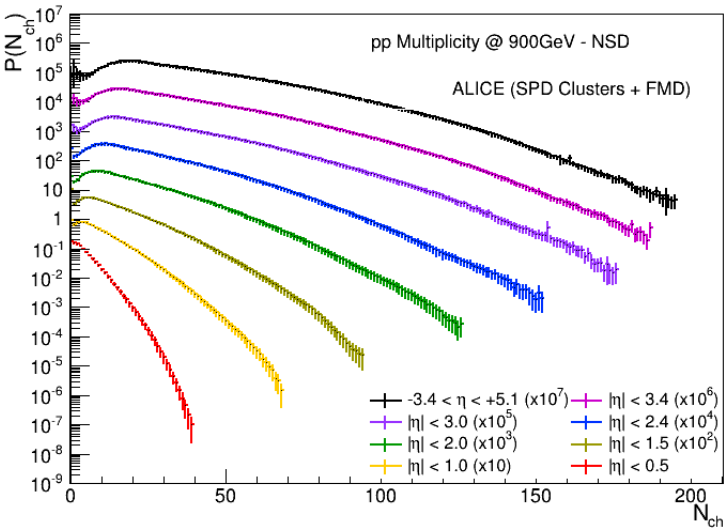


Figure 9.1: Charged–particle multiplicity distributions for NSD pp collisions at $\sqrt{s} = 0.9$ TeV.

Going to higher energies, Figure 9.2 shows the INEL trigger class results for 2.76 TeV data. As already discussed in Section 7.2, these data have bad quality both for SPD and FMD. The SPD problems are fixed, therefore the first three distributions are accurate, while no fix is still found and applied to the FMD data. For this reason, the last five distributions look very flat and their mean is pushed a lot towards the lower bins of the distributions. Certainly, for publication, a fix must be found or this energy will be unfortunately discarded. Additional plots for 2.76 TeV for NSD class and INEL> 0 can be found in Appendix B.

Figure 9.3 shows the results for NSD trigger class collisions at 7 TeV. This is the sample which contains more data in total, 61 million MB events, an order of magnitude more events than available and used for 0.9 TeV. Looking at the Quality Assurance parameters (stability of Landau and Gaussian fit parameters from all the runs used), it can be said that the 7 TeV data from periods LHC10c and 10d have very good quality. For this reason the final distributions look nice and smooth. As before, additional plots are stored in the appendix.

The 8 TeV INEL trigger class events multiplicity distributions are shown in Figure 9.4, while the other trigger classes are shown in Appendix B. The shape of this distribution is very similar to the one for 7 TeV, although the reduced statistics creates a problem at the highest ends of the distribution. In fact, with smaller data sample, the unfolding procedure works less well in the low statistics tails. This effect will be seen also in the following. The data can be trusted up to $N_{ch} \backsim 275$, while for 7 TeV it is possible to reach $N_{ch} \backsim 350$ in the full range distribution.

Figures 9.5a-d, show the evolution of the multiplicity distributions with respect to energy and pseudorapidity range for the INEL trigger class (NSD class is shown in Appendix B). In Figure 9.5a, only the SPD distributions are shown for the four avail-

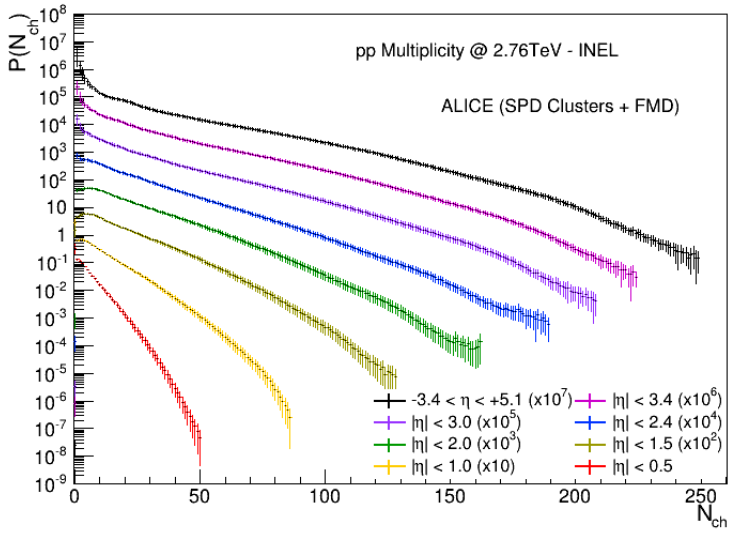


Figure 9.2: Charged–particle multiplicity distributions for INEL pp collisions at $\sqrt{s} = 2.76$ TeV. The combined FMD and SPD distributions, from the green to the black ones, are problematic, since they include FMD bad quality data (see Section 7.2).

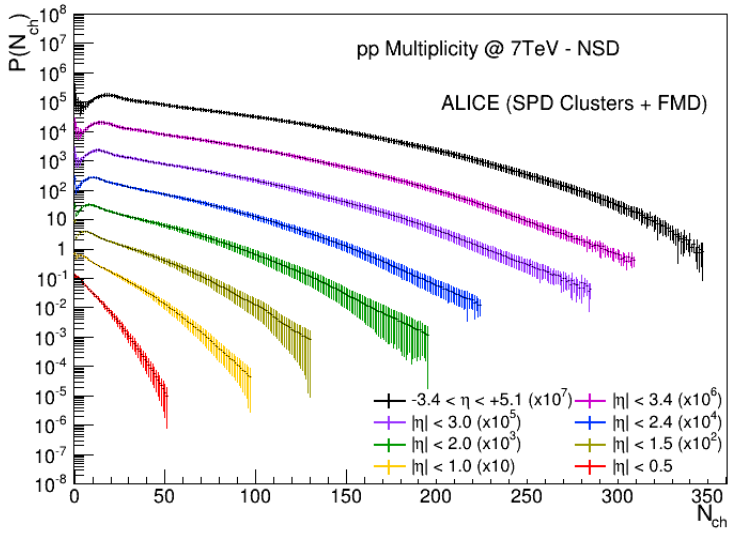


Figure 9.3: Charged–particle multiplicity distributions for NSD pp collisions at $\sqrt{s} = 7$ TeV.

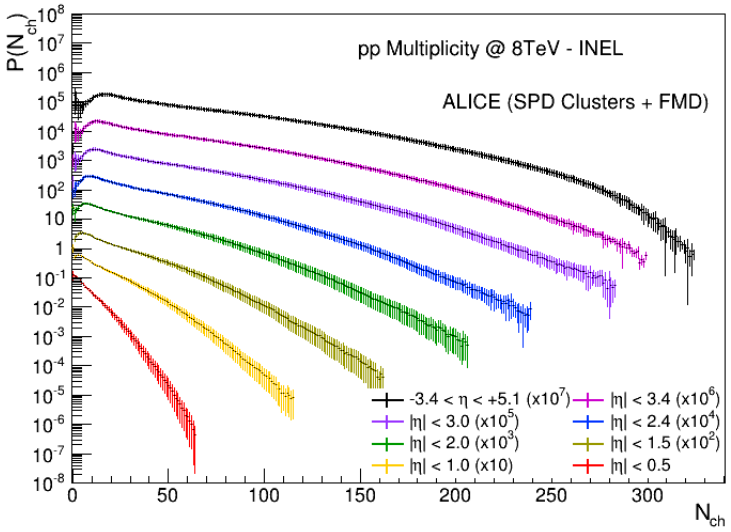


Figure 9.4: Charged–particle multiplicity distributions for INEL pp collisions at $\sqrt{s} = 8$ TeV.

able energies, the two curves for 7 and 8 TeV are very similar, as expected. In Figures 9.5b and 9.5c, the distribution for 2.76 TeV is not included, due to the fact that the FMD is present and the quality of these data is compromised. In Figure 9.5c, it can be noted that there is an overlap between 7 and 8 TeV distributions. From approximately $N_{ch} = 280$ onward, the distribution for 8 TeV does not have physical meaning due to the lack of statistics and, therefore, should be cutoff.

Figure 9.5d shows the evolution in pseudorapidity for 7 TeV collisions, this is a very important plot, because it shows the importance and the relevance of this work. The multiplicity can be extended by $2/3$ by ALICE using the FMD, increasing the multiplicity range to about $N_{ch} = 350$. It is also evident that the distributions shape is very different, the probability of detecting few particles decreases extending the pseudorapidity range. On the contrary, the probability of detecting more particles, e.g. 130, increases by 2 orders of magnitude by widening the η range.

Comparisons with ALICE and CMS Results

A very important step, in every analysis, is the comparison to similar studies already done. Certainly, an important one is the comparison with other results obtained inside the ALICE collaboration. Currently, there is an on–going analysis using the SPD tracklets in ALICE, which allows to compare to the most up–to–date methodology and simulations. This analysis is still not published, therefore, due to ALICE publication policy, it is not possible to show in this thesis the results which are obtained. Of course, the comparisons have been checked internally in ALICE and there is very good agreement within systematics.

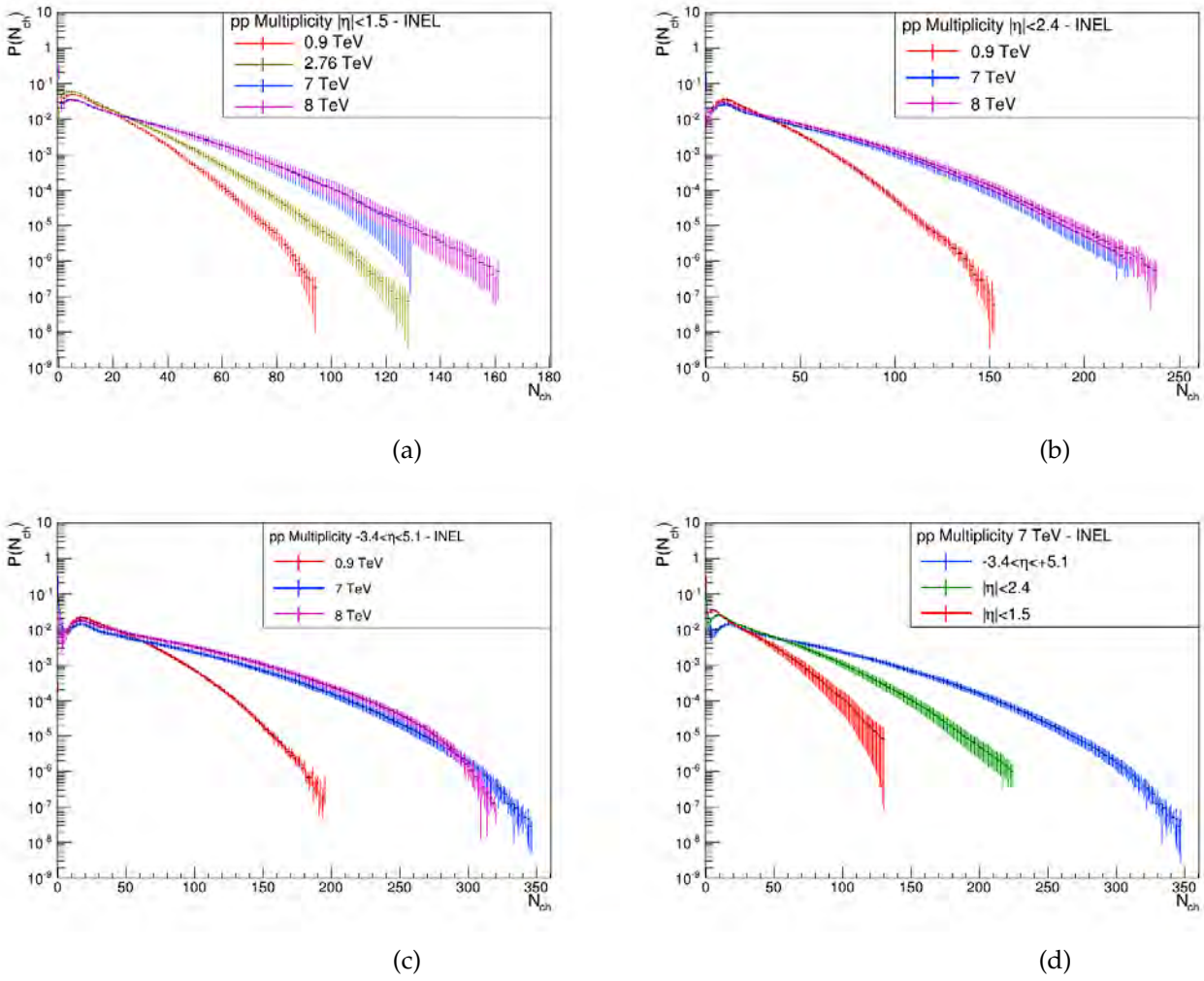


Figure 9.5: (a-b-c) Evolution of charged–particle multiplicity distributions as a function of $\sqrt{s}$, for INEL event classes for three different η ranges; (d) evolution of the distributions as a function of the η ranges for INEL events in $\sqrt{s} = 7$ TeV.

The on–going–analysis uses only SPD, relying on tracklets (which allow to lower the pile–up contamination) and reproduces, therefore, only the first three distributions: $|\eta| < 0.5$, $|\eta| < 1.0$ and $|\eta| < 1.5$. For this reason, these three distributions are derived in the thesis and will be carried on only for internal comparisons and checks, but will not be part of the final imminent publication.

In Figure 9.6, the comparison with ALICE and CMS ([58] and [56] respectively) published data is shown, for the NSD event class at $\sqrt{s} = 0.9$ TeV. Concerning the comparison between this work with previous ALICE results, there is very good agreement between the measurements within systematics. There is disagreement in the tails of for the $|\eta| < 1.0$ distribution, where the published result starts to oscillate (it is the published result which oscillates, as can be seen in the derivative of the distributions, Figure B.10). This is a consequence of the χ^2 minimization method used to unfold the raw data, which, as already explained in Section 8.3, tends to produce oscillations in the distribution.

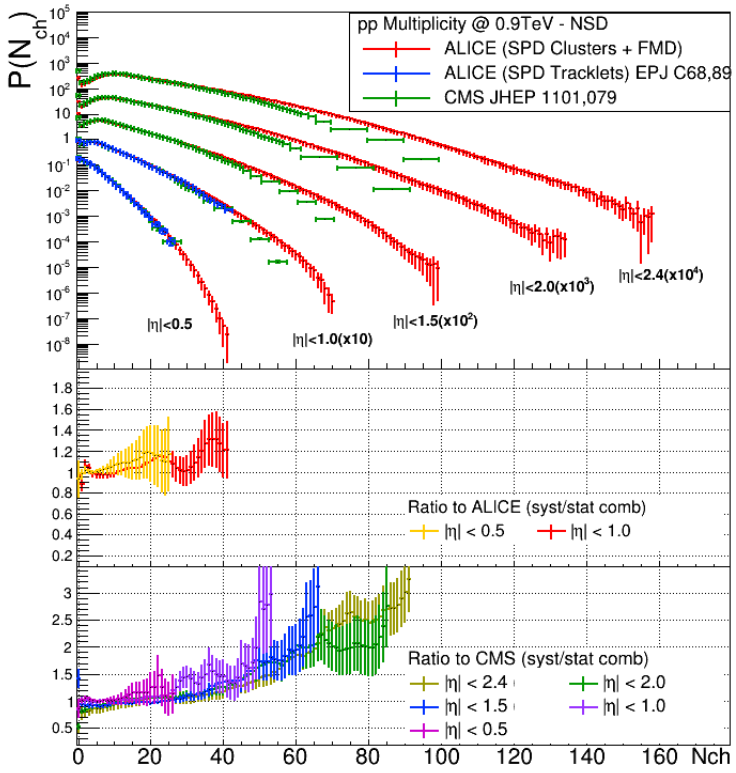


Figure 9.6: Comparison of the multiplicity distributions for NSD pp collisions at $\sqrt{s} = 0.9$ TeV of this work with ALICE [58] and CMS [56]. See text for details.

Discrepancies are observed in the comparison with CMS, both in the first bins and especially in the tails. The first bins are different because the models used to describe the diffraction masses are different, also, CMS can detect better Single and Double–Diffractive events (the detection capability of ALICE for diffraction events will be in-

creased in Run II, see Section 5.2). Concerning the tails, CMS data are systematically lower, and this behavior is observed also when comparing the other on–going analysis carried on with the SPD tracklets with CMS. Therefore, it didn't come as a surprise. It must be noted that differences are less prominent when comparing ALICE and CMS at 7 TeV. This might be due to the fact that more particles are detected in 7 TeV with respect to 0.9 TeV, therefore there is less inefficiency in ALICE.

In Figure 9.7, this work is compared with ALICE published results for INEL>0 trigger class [59]. There is good agreement within systematics, but in the first 3 bins. This is due to the fact that at the time of this publication (2010) the PYTHIA MC generator (Section 6.1) used for the derivation of the efficiency correction (Section 7.7), was not tuned for diffraction using the Kaidalov–Poghosyan model [67]. Moreover, in that case, only PYTHIA was used to derive the efficiency correction, while also PHOJET is used in this thesis.

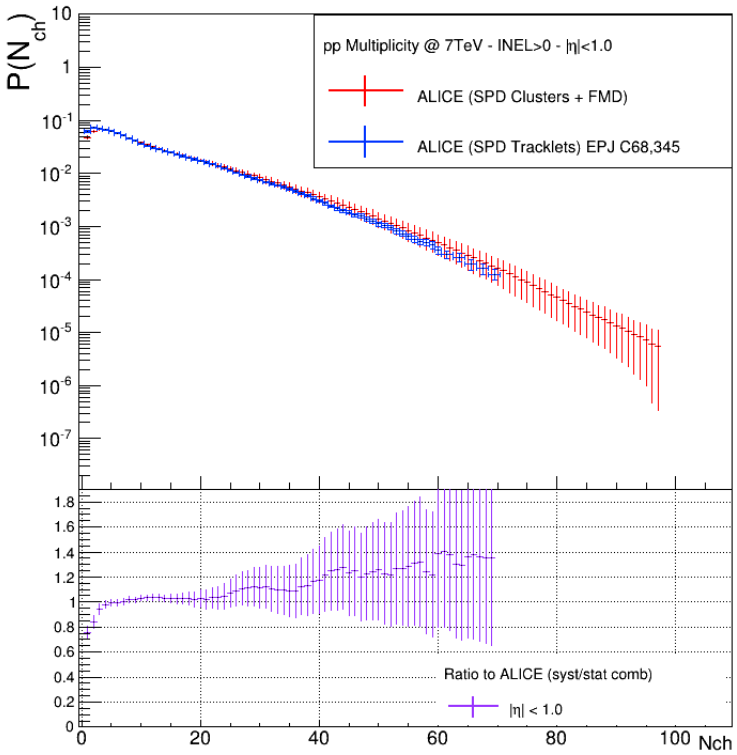


Figure 9.7: Comparison of the multiplicity distributions for INEL>0 pp collisions at $\sqrt{s} = 7$ TeV obtained in this work with ALICE [59]. See text for details.

The last comparison shown is the one in Figure 9.8 with CMS at 7 TeV [56]. In this case there is very good agreement with CMS, but in the very first bins due to the definition of diffraction masses already discussed above. There are no published data for 2.76 and 8 TeV to compare to. Anyway, this work is compared with the on–going analysis with SPD tracklets, showing very good agreement and, as expected, much

better agreement with respect to the comparisons shown here with older results form ALICE.

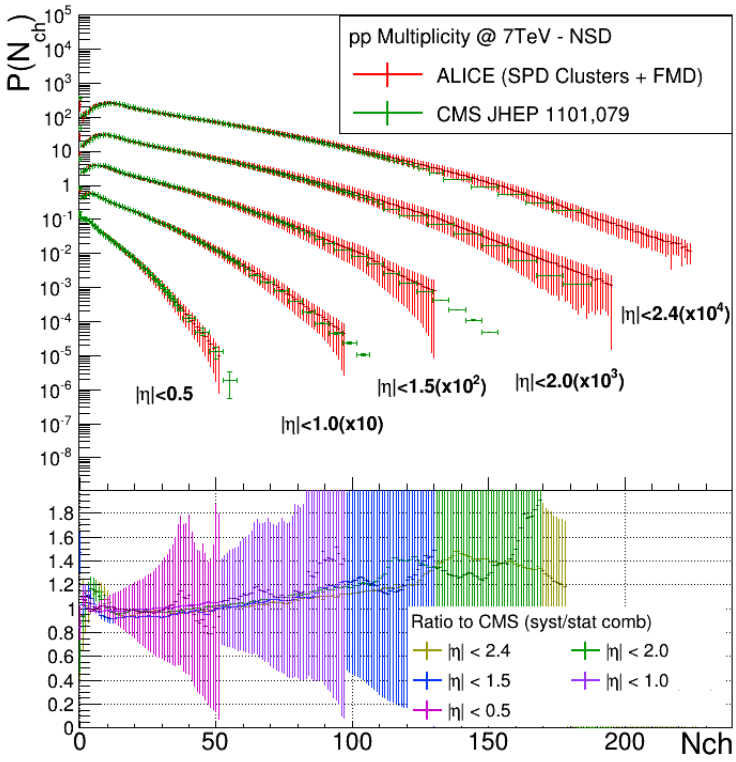


Figure 9.8: Comparison of the multiplicity distributions for NSD pp collisions at $\sqrt{s} = 7$ TeV ontained in this work with CMS [56]. See text for details.

Comparisons with MC Models

In Figure 9.9, comparisons with PYTHIA (6.1) and PHOJET (6.2) MC generator distributions, both diffraction tuned, are shown for INEL events at 0.9 TeV. It can be seen that PHOJET does a good job in the SPD distributions, but cannot reproduce the tails and the first bins for higher pseudorapidity ranges. On the other side, PYTHIA 6 Perugia 0 tune does not reproduce the data at all.

In Figure 9.10, the distributions at 7 TeV for INEL events sample are compared to the same MC generators as above. In this case, neither PHOJET nor PYTHIA Perugia 0 reproduce the data (they both strongly underestimate them). Other MC models will be tested before publication.

Comparisons with IP–Glasma Model

It is possible to compare the multiplicity distributions with the IP–Glasma model, which is a model based on Color Glass Condensate (see Section 2.2). One can notice that NBDs (Section 4.1) are generated within the CGC framework [80]. Moreover,

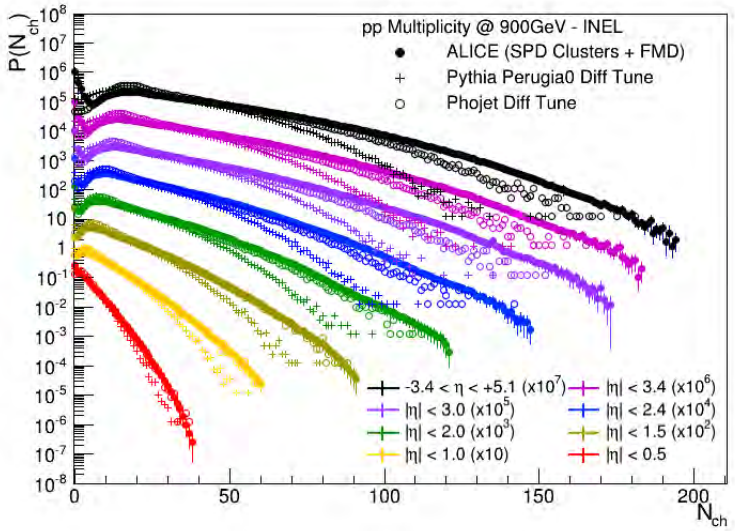


Figure 9.9: Comparison with models for INEL events PYTHIA 6 Perugia 0 and PHOJET at 0.9 TeV.

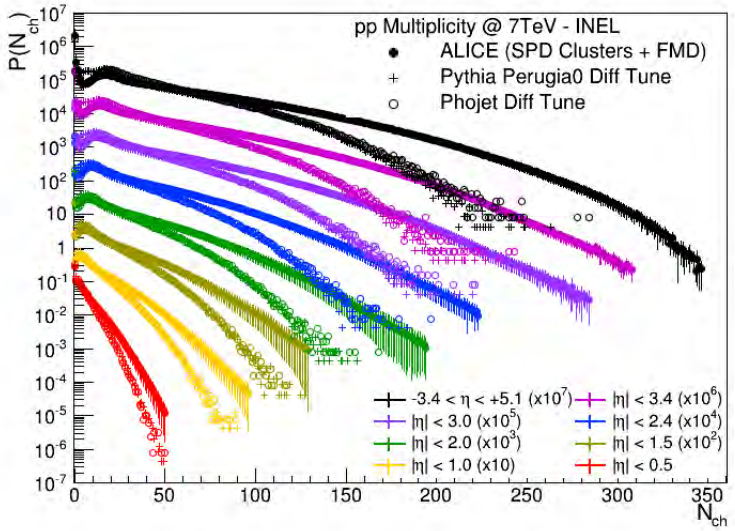


Figure 9.10: Comparison with models for INEL events PYTHIA 6 Perugia 0 and PHOJET at 7 TeV.

the multiplicity distribution produced by the decay of the Glasma flux tubes (see Section 2.3) is a NBD with $k \propto Q_s^2 S_\perp$, where Q_s is the gluon saturation scale and $S_\perp$ is the transverse overlap area of the collision [81]. This means that the NBD naturally–present fluctuations do not have to be added by hand into the CGC bases IP–Glasma model.

Details about this model can be found in [81]. In short, it is the only framework which includes event–by–event sub–nucleon scale color charge fluctuations derived from HERA ee deep inelastic scattering data, together with solutions of Yang–Mills equations for the Glasma fields. Moreover, it includes viscous hydrodynamics.

In Figure 9.11, SPD clusters distributions for $|\eta| < 0.5$ are shown together with the IP–Glasma model distributions. The difference between the blue and the green IP–Glasma distributions is that in the green one the ration between Q_s and density of color charge is fixed. In this way, no fluctuations are introduced. In the blue distribution instead, the color charge density fluctuates around the mean following a Gaussian distribution with width $\sigma = 0.09$. The result is closer to the experimental data, but there is still a discrepancy, which indicates that the color charge density should fluctuate with a non–Gaussian distribution. Additionally, as pointed out in [81], there are still some sources of QCD fluctuations missing in the IP–Glasma framework.

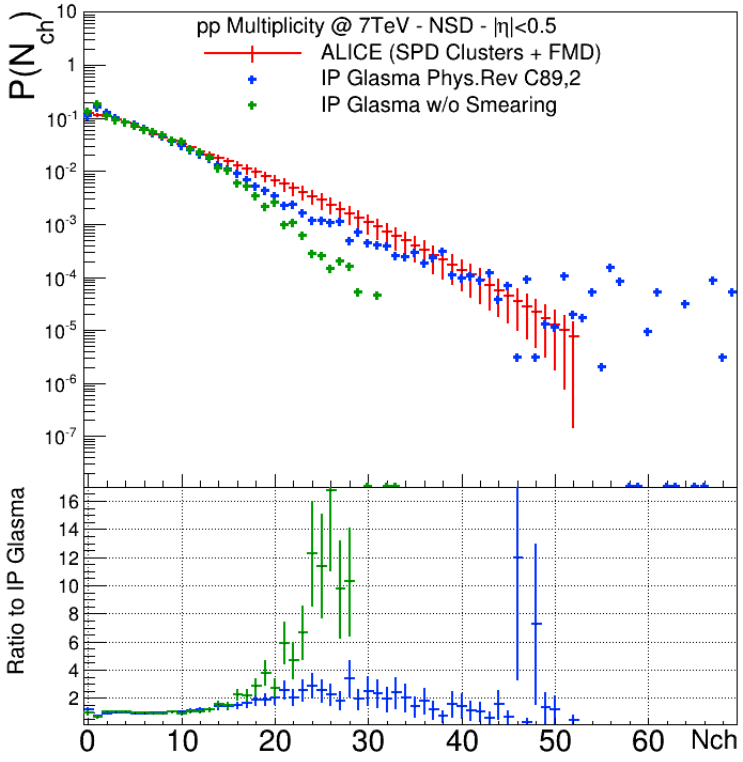


Figure 9.11: Charged–particle multiplicity distributions for pp at $\sqrt{s} = 7$ TeV compared with distributions from the IP–Glasma model, fixing (in green) or letting the ratio between Q_s and the color charge density fluctuate (blue) [81]. The IP–Glasma points are taken from [82].

9.2 KNO Scaling

The evolution of the multiplicity distributions with the center of mass energy $\sqrt{s}$ can be studied also using the KNO variable $N_{ch}/\langle N_{ch}\rangle$ (see Section 4.1). This study is shown in Figures 9.12 for the NSD class in three pseudorapidity regions. In all cases, the tails increase with higher energy, therefore there is KNO scaling violation. Moreover, the violation increases increasing the pseudorapidity range.

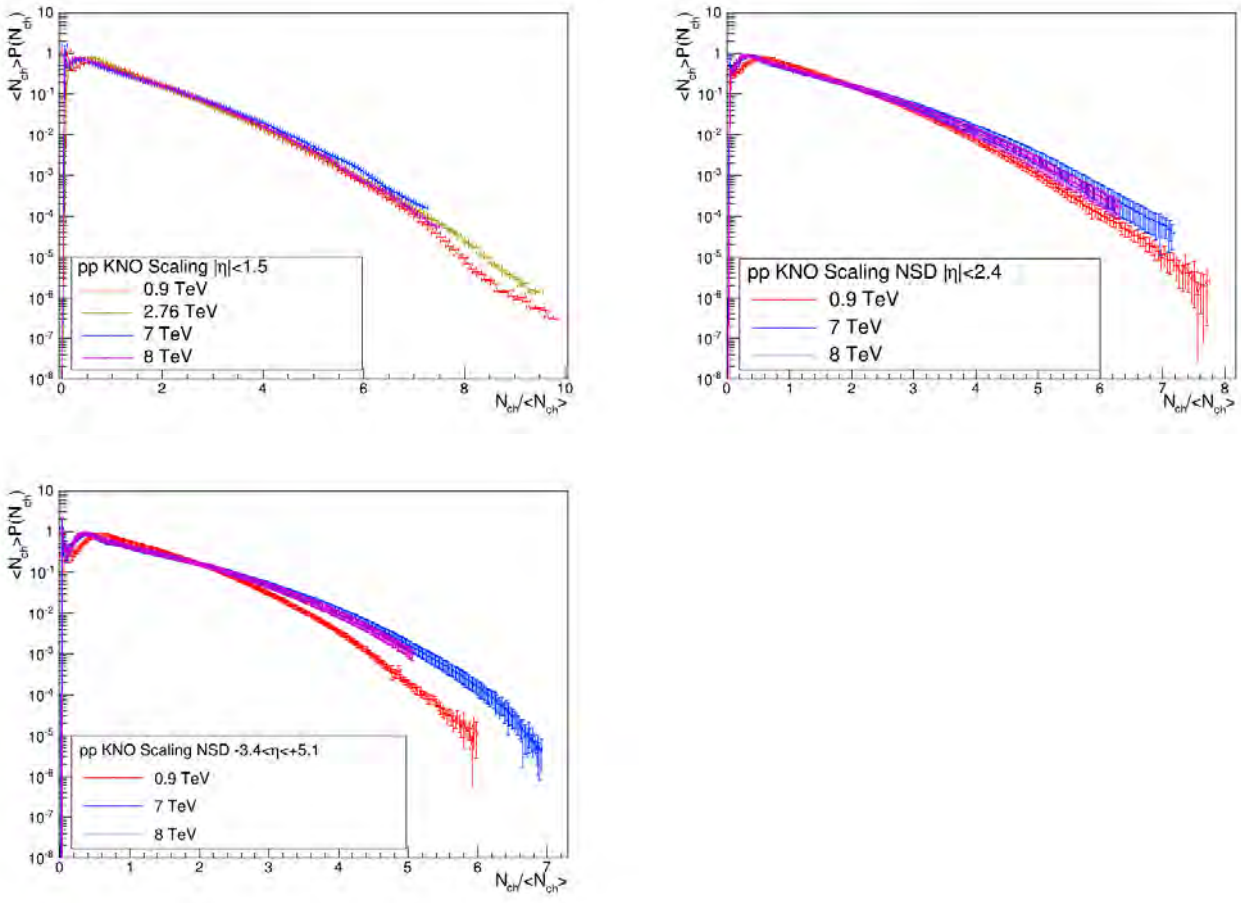


Figure 9.12: KNO scaling, top left: pure SPD region, top right: overlap zone between SPD and FMD, bottom: region for where the majority of the counts are from the FMD. See text for details.

The distribution for $\sqrt{s} = 2.76$ TeV is included only in the SPD case, since the data for FMD at this particular energy are corrupted (Section 7.2). It must be noted that the curves for 7 and 8 TeV are overlapping within systematics, since the two energies are very close. Anyway, it seems that the 8 TeV lies slightly below the 7, and the cause of it has to be investigated further. In any case, this effect is negligible and the two distributions overlap within systematics, as mentioned above.

9.3 Double NBD Fits

As extensively described in Section 4.1, the multiplicity distributions can be fitted with a Negative Binomial Distribution. Already at 0.9 TeV for $|\eta| < 1.3$ in the NSD sample, ALICE observed a departure from the single NBD fit [58]. For this reason, fits to the sum of two NBDs (see Section 4.1) have been performed here for the NSD events sample. Three distributions are plotted for 0.9, 7 and 8 TeV, to see what is the behavior of the fit parameters widening the pseudorapidity range, namely the ones for $|\eta| < 1.5$, $|\eta| < 2.4$ and $-3.4 < \eta < +5.1$. The distributions at 2.76 TeV are not included for reasons previously mentioned (Section 7.2). Anyway, since the SPD data have good quality, the fit parameters for $|\eta| < 1.5$ are included in Table 9.1.

The distributions have been fitted using this function:

$$P(n) = \lambda[\alpha P_{NBD}(n, \langle n\rangle_1, k_1) + (1-\alpha)P_{NBD}(n, \langle n\rangle_2, k_2)] \tag{9.1}$$

where

$$\begin{aligned} P_{NBD}(n, \langle n\rangle, k) &= \frac{\Gamma(n+k)}{\Gamma(k)\Gamma(n+1)}\left(\frac{\langle n\rangle}{k+\langle n\rangle}\right)^n\left(\frac{k}{k+\langle n\rangle}\right)^k = \\ &= \frac{1}{n\mathrm{B}(n,k)}\left(\frac{\langle n\rangle}{k+\langle n\rangle}\right)^n\left(\frac{k}{k+\langle n\rangle}\right)^k \end{aligned} \tag{9.2}$$

To take into account the fact that this function does not describe the 0–bin (which is therefore taken out from the fit), a normalization factor λ is introduced. The fit parameters are written in Table 9.1.

Having a closer look at the parameters, their behavior is mostly consistent with the one observed in [83], where the CMS distributions [56] are analyzed. Looking at the soft component (the first), it can be noticed that $\langle n\rangle_1$ increases with increasing energy and pseudorapidity range. Secondly, the parameter k_1 increases with increasing the pseudorapidity range, but it is constant increasing the center of mass energy within systematics, this fact will be described in the following.

For the semi–hard component (the second), the $\langle n\rangle_2$ parameter behaves in the same way with respect to pseudorapidity and energy increase. Moreover, it can be noted that $\langle n\rangle_2 \simeq 3\langle n\rangle_1$, as already observed for the CMS distributions [83]. All the χ^2/DOF values are less than unity, this comes from the fact that the distributions have been fitted with combined systematic and statistical errors, lowering the χ^2/DOF values. α is a very important parameter to underline the soft events fraction. For 0.9 TeV the average soft event fraction is 86%, for 2.76 TeV is 76% and for 7 and 8 TeV is the same, within systematics, 60% on average. Looking at the fitted distributions, it is clear that it is better to analyze the fits for the widest rapidity ranges, because more particles are included producing a shoulder–like distribution, which is more visible in that range. This helps in including both soft and semi–hard particles.

In particular, as already explained in 4.1, if the parameter k is constant with energy, the KNO scaling is valid. In this case, k_1 is constant with energy within systematics for $|\eta| < 1.5$ and 2.4, indicating that KNO scaling is valid for the soft component of

$\sqrt{s}$ (TeV)	η Range	λ	α	$\langle n \rangle_1$	k_1	$\langle n \rangle_2$	k_2	χ^2/DOF
0.9	$\lvert\eta\rvert < 1.5$	0.91±0.01	0.87±0.03	10.17±0.35	2.20±0.08	26.24±1.11	7.92±0.73	28.41/104
	$\lvert\eta\rvert < 2.4$	0.97±0.01	0.80±0.09	16.57±1.55	2.66±0.23	39.61±3.63	6.57±1.05	11.77/127
	$-3.4 < \eta < +5.1$	0.96±0.01	0.81±0.04	27.80±1.29	3.59±0.27	64.62±3.03	9.57±0.93	31.60/164
2.76	$\lvert\eta\rvert < 1.5$	1.01±0.01	0.76±0.05	9.19±0.54	2.45±0.20	26.72±1.88	4.92±0.59	12.41/123
7	$\lvert\eta\rvert < 1.5$	0.93±0.01	0.51±0.06	8.82±0.77	2.28±0.18	30.41±2.13	3.26±0.46	5.39/124
	$\lvert\eta\rvert < 2.4$	0.98±0.01	0.57±0.05	16.18±1.03	2.36±0.18	52.96±2.42	4.03±0.34	24.11/219
	$-3.4 < \eta < +5.1$	0.95±0.01	0.68±0.04	32.44±2.06	2.35±0.18	96.32±4.21	5.76±0.50	47.75/289
8	$\lvert\eta\rvert < 1.5$	0.97±0.01	0.66±0.07	12.02±1.21	1.91±0.17	37.68±2.84	4.48±0.60	6.33/159
	$\lvert\eta\rvert < 2.4$	1.00±0.01	0.59±0.05	18.15±1.26	2.40±0.15	56.84±2.74	4.19±0.37	15.49/232
	$-3.4 < \eta < +5.1$	0.96±0.01	0.58±0.05	30.56±1.92	2.91±0.22	91.42±3.98	4.70±0.43	40.56/267

Table 9.1: Double NBD fit parameters for multiplicity distributions, NSD events.

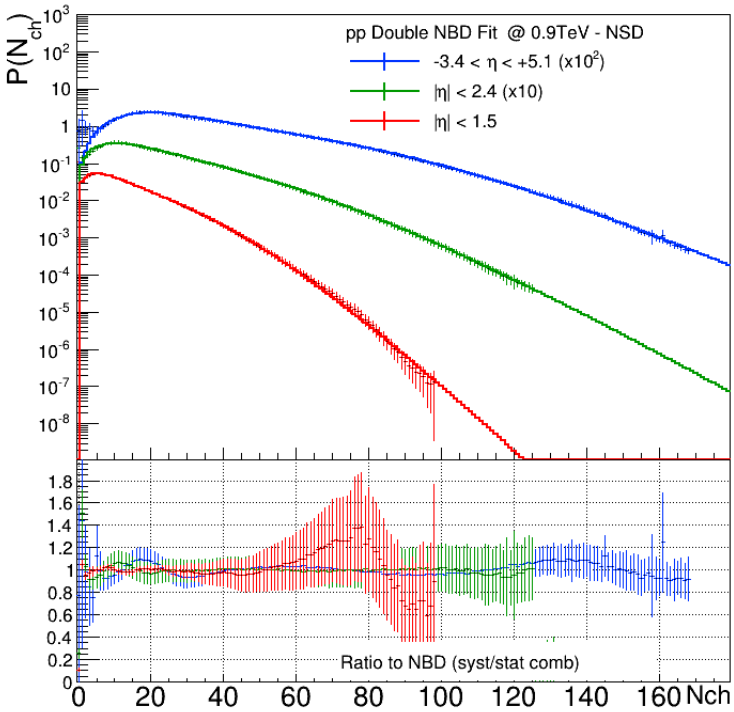


Figure 9.13: Multiplicity distributions for NSD pp collisions at $\sqrt{s} = 0.9$ TeV, the line shows the double NBD fit. Ratios of data to the fits are shown. See text for details.

the distribution, but it is violated only for the semi–hard one. There is anyway KNO scaling violation in the soft component for the full range distributions.

Certainly, the KNO scaling is strongly violated for the semi–hard component. There are three scenarios proposed in the model by Ugoccioni and Giovannini (Section 4.1) summarized here [83]: either KNO scaling is conserved for the semi–hard component, and this is not the case, or the k_2 decreases linearly with increasing energy showing violation, or, finally, KNO violation is not very strong and k_2 starts decreasing with energy but asymptotically tending to a constant value. The last case seems to fit the obtained parameters, since the decrease from 0.9 to 2.76 is very strong, while it seems to saturate from 2.76 to 7 and 8 TeV. Of course, with 13 TeV coming data there will be more space and statistics to speculate about the validity of the clan model (Section 4.1).

The fitted distributions are shown in Figures 9.13, 9.14 and 9.15 for 0.9, 7 and 8 TeV (2.76 TeV fits are not computed for FMD, see 7.2, but SPD region parameters are included in Table 9.1). All the ratios between the distributions and the fits to double NBDs are consistent with one, although two things need to be remarked. The first bins of the full pseudorapidity range cannot be fitted with the double NBDs, because of the raise in the first bins. In fact, the first bins have a different slope. Therefore, the fit for the distributions at $-3.4 < \eta < +5.1$ starts at around bin 4.

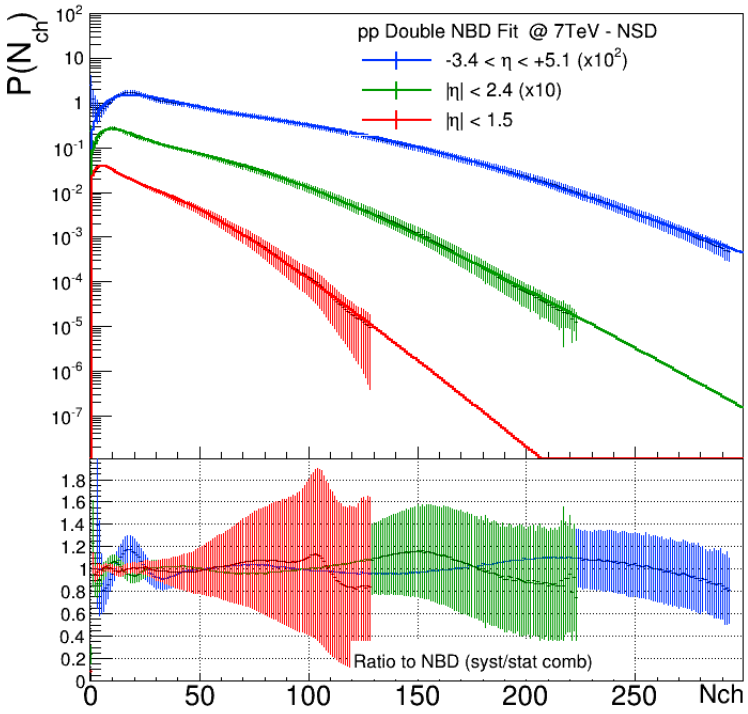


Figure 9.14: Multiplicity distributions for NSD pp collisions at $\sqrt{s} = 7$ TeV, the line shows the double NBD fit. Ratios of data to the fits are shown. See text for details.

Moreover, there is a lowering of the tails in the full range distributions ($-3.4 < \eta < +5.1$) for 7 and 8 TeV, which can be seen in Figures 9.3 and 9.4 that cannot be fitted with the double NBDs. It must be understood if this is a tendency which comes from the unfolding procedure not holding for low statistics and high multiplicity, or if it has physical meaning. Again, more information on this direction will come from 13 TeV data. In fact, they will tell if a third, even harder, component must be added to be able to fit a two–shoulders–like distribution (already discussed in Section 4.1). It is, by now, not possible to fit with a triple NBD the $-3.4 < \eta < +5.1$ due to the very low statistics at high multiplicity.

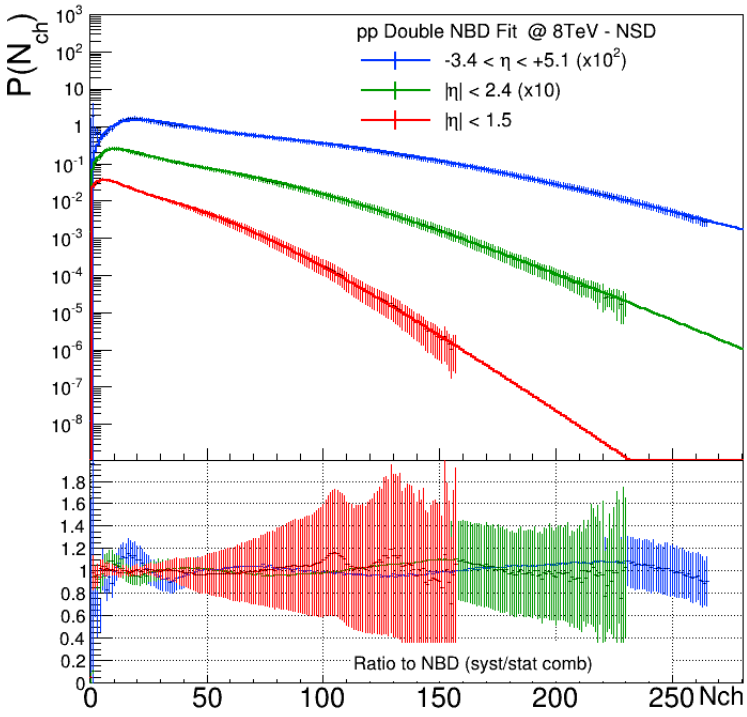


Figure 9.15: Multiplicity distributions for NSD pp collisions at $\sqrt{s} = 8$ TeV, the line shows the double NBD fit. Ratios of data to the fits are shown. See text for details.

Chapter 10

Analysis and Results for pPb Collisions

In the previous chapters , the analysis and results for pp multiplicity distributions have been presented. The procedure for pPb collisions is the same, although using different settings which will be summarized in this chapter. For the sake of clarity, the pPb multiplicity distributions results are presented here, separately from the pp ones.

10.1 Analysis

Also in the pPb distribution case, the detectors used are the FMD for the forward region and SPD for the central pseudorapidity ranges. The pPb distributions are analyzed for the only energy available provided at LHC which is $\sqrt{s_{NN}} = 5.02$ TeV, colliding a p beam at 4 TeV with a Pb beam at 4 TeV. They are referred as collisions in the p–going direction (p towards the muon arm and FMD3, negative η). Secondly, the Pbp collisions in the Pb–going direction (Pb towards the muon arm and FMD3) at $\sqrt{s_{NN}} = 5.02$ TeV are analyzed. The DPMJET [69] MC generator is used to build the response matrices.

Event Selection

The pPb and Pbp data were both collected at the beginning of 2013. The p–going ones are collected with negative polarity of the L3 magnet, while the Pb–going with positive polarity. The properties of the data samples are presented in Table 10.1.

System	Energy (TeV)	Data Period	Selected MB Events
pPb	5.02	LHC13b	15×10^6
Pbp	5.02	LHC13f	1.4×10^6

Table 10.1: Data samples used for the analysis. For each system, the center of mass energy and the data period used are written together with the total number of selected MB events.

The online trigger selection is kINT7 which means an AND requirement between the V0 in the A–side and in the C–side of ALICE. The offline trigger required is minimum bias MB$_{AND}$, which is equivalent to the online kINT7 and therefore equivalent with the NSD trigger used for pp collisions.

The background rejection in pPb is implemented in the same way as for the pp collisions (see Section 7.1). Moreover, like in the pp case, the vertex v_z is selected from -4 to +4 cm, like shown in Figure 10.1.

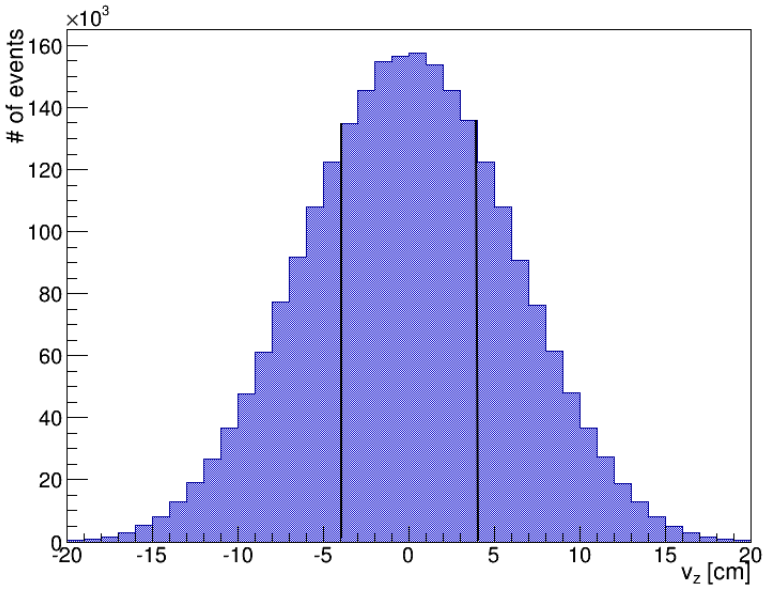


Figure 10.1: Number of events per vertex bin for Pbp events. The black lines identify the vertex selection: $|v_z|$ <4 cm.

Particle Selection

The FMD cuts applied are the same as in the pp case, E_{low} is fixed at 0.15·$\Delta E/\Delta E_{MIP}$ for the inner and 0.18 for the outer rings. While $E_{high} \equiv E_{hit} = E_{MPV} - \xi$. The occupancy for pPb is shown in Figure 10.2, and, as expected, the mean occupancy is higher with respect to the pp occupancy (Figure 7.5).

The azimuthal coverage is basically the same as in Figure 7.10b, the only not working chips in the FMD are one in FMD2I and one in FMD2I. In SPD there is a chip which is switched off, because FMD2I overlaps with it.

The correction for secondaries hitting the FMD is computed processing the unfolding, like it it done for pp collisions (see Section 7.4).

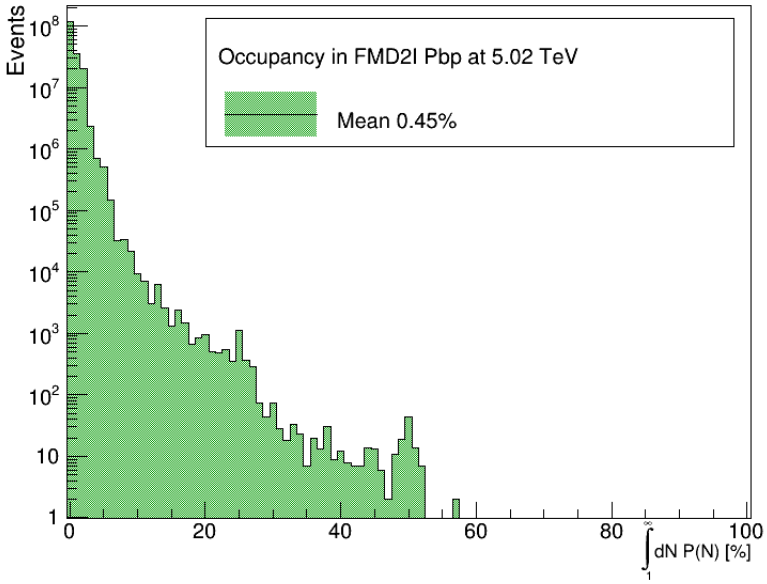


Figure 10.2: Occupancy in the FMD2I ring for Pbp.

Unfolding

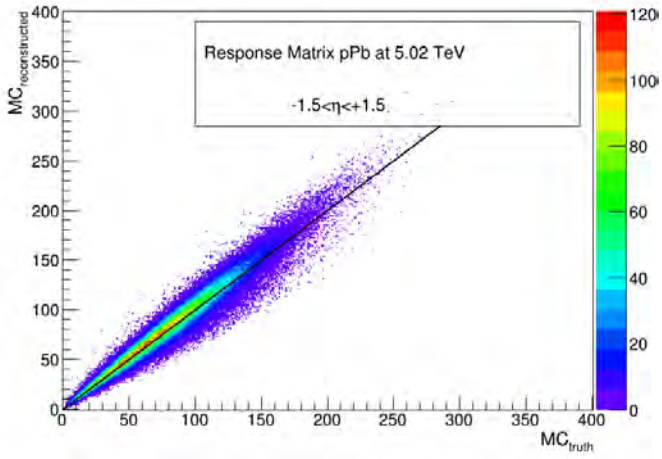
As extensively explained in Section 7.6, it is necessary to unfold to correct for SPD acceptance gaps and FMD secondary hits. The unfolding needs, of course, to be applied also in the pPb case. Figures 10.3a-b-c show the three response matrices for different pseudorapidity ranges. The matrices are derived using the DPMJET MC generator, propagated through the code and the reconstruction steps. Like in the pp case, the major correction is in the full η range (10.3c), where the matrix corrects for the highest amount of secondaries.

In the pPb case, only the Bayesian unfolding technique is used(Section 7.6), since it takes around 7 hours to unfold the eight distributions locally, with full CPU usage. Unfolding by minimizing the χ^2 function takes a very long time, probably around 3 days with full CPU usage, therefore, it is not been performed in this thesis.

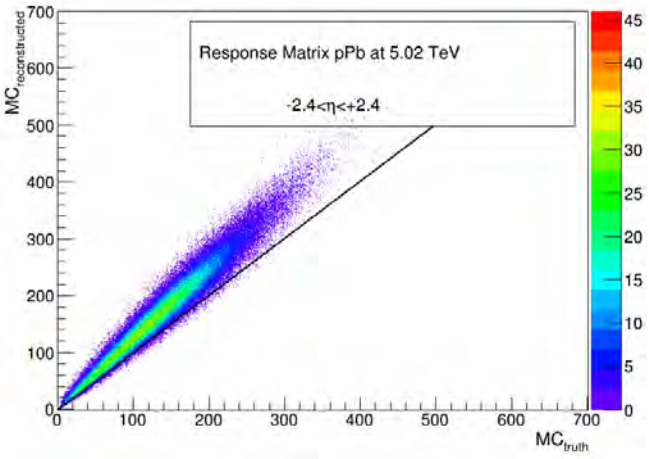
For the pp collisions, it was very important to correct for detection inefficiencies (see Section 7.7), while for pPb the probability of producing 0 or 1 charged particles is $P(0,1) = 0$ for $-3.4 < \eta < +5.1$ in DPMJET MC truth. For this reason, the correction for detection inefficiency is not been computed.

10.2 Systematic Uncertainties

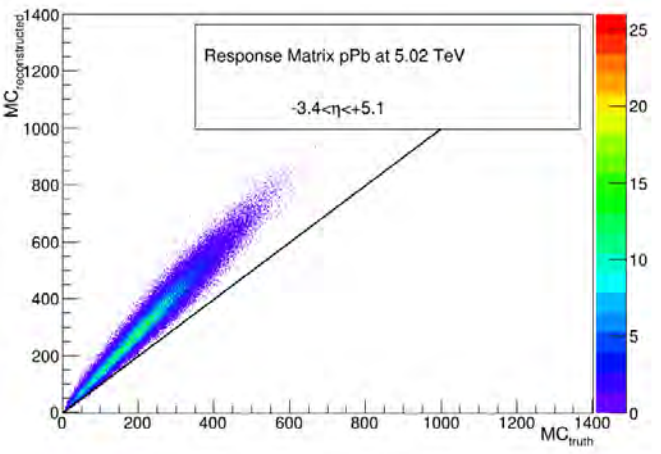
In this section, the systematic uncertainties are discussed. The only two uncertainties which are performed are the run–to–run fluctuations and the variation of the sharing filter cuts. This is because, firstly, the material budget cannot be checked, since there is no available production with increased amount of material; secondly, as underlined in



(a)



(b)



(c)

Figure 10.3: Response matrices obtained propagating MC generated events, in this case using DPMJET, through the reconstruction code. (a) Matrix for the pure SPD region; (b) Overlap between SPD and FMD region; (c) Matrix for region where the majority of the counts are from the FMD. The diagonal is plotted as a black line.

the previous section, the unfolding is performed with only one method, the Bayesian one.

Run–To–Run Fluctuations

Two runs have been analyzed in this thesis, run 195483 for pPb , and run 196433 for Pbp. Therefore, in principle, to derive the run–to–run fluctuations, the raw two distributions have to be unfolded with two different response matrices anchored to the respective runs.

The run–to–run fluctuations test is performed for seven symmetrical ranges from the smallest $|\eta| < 0.5$ to the widest $|\eta| < 3.4$, but it has not be computed for the asymmetrical range $-3.4 < \eta + 5.1$. That is because in this last range the two distributions measure different events. For pPb, the p goes towards the muon spectrometer, while the Pb goes through FMD1, which is located at $+3.68 < \eta + 5.03$. In the other case, Pbp, is the p which goes through the FMD1, therefore less particles are detected. For this reason, the seven symmetrical distributions can be merged together, while, for the widest asymmetrical range, the two distributions are not merged together and will be treated separately.

In Figure 10.4, the pPb and Pbp both raw and unfolded distributions are shown. It can be seen that the Pbp sample contains more pile–up, visible in the high–multiplicity region. The bin–by–bin difference between the two unfolded distributions will give the run–to–run uncertainty. It must be noted that the distributions are re–binned in bins of $5 \cdot N_{ch}$ particles due to the low statistics. The reason of the statistical fluctuations comes form the low statistics of the response matrix shown in Figure 10.3. The matrices are not obtained from flat MC productions, like in the pp case and this affects considerably the resulting unfolded distribution.

Sharing/Poisson Cuts

This systematic check was done varying the cuts used for the sharing filter and the Poisson counting method. Therefore, as explained in Section 8.4, two distributions were obtained, one raising the Sharing/Poisson cut by 5% and one lowering it by 5%. Afterwards, the systematic uncertainty is derived comparing these two distributions in Figure 10.5 bin–by–bin.

Systematic Uncertainties Overview

The two systematic errors are then summed in quadrature:

$$\delta_{TOT} = \sqrt{\delta^2_{fluct} + \delta^2_{cut}} \tag{10.1}$$

and Figures 10.6 show the single contributions from run–to–run fluctuations, cut variation uncertainties and the total systematic effect. Three regions are considered $|\eta| < 1.5$ (where only SPD is present, therefore, the FMD cut variations are not present), $|\eta| < 2.4$ and $|\eta| < 3.4$ (the full range is considered afterwards, since the physics between pPb

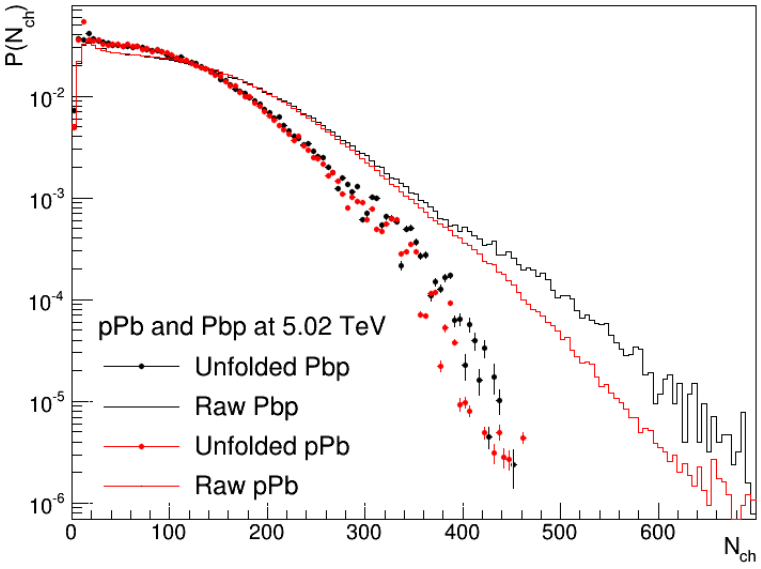


Figure 10.4: The Pbp and pPb for 5.02 TeV in $|\eta| < 2.4$ are displayed, in order to observe the run–to–run fluctuations.

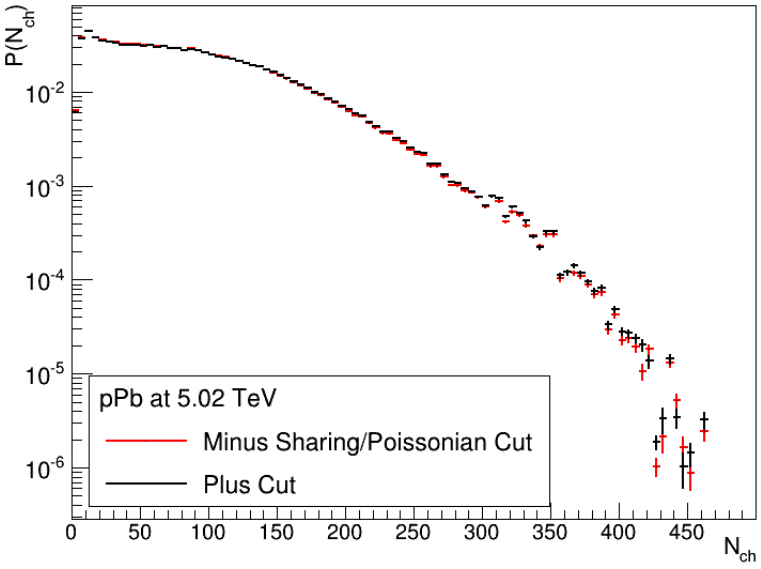


Figure 10.5: Distributions obtained by lowering or raising the sharing/Poissonian cut by 5% for pPb at 5.02 TeV data in $|\eta| < 2.4$.

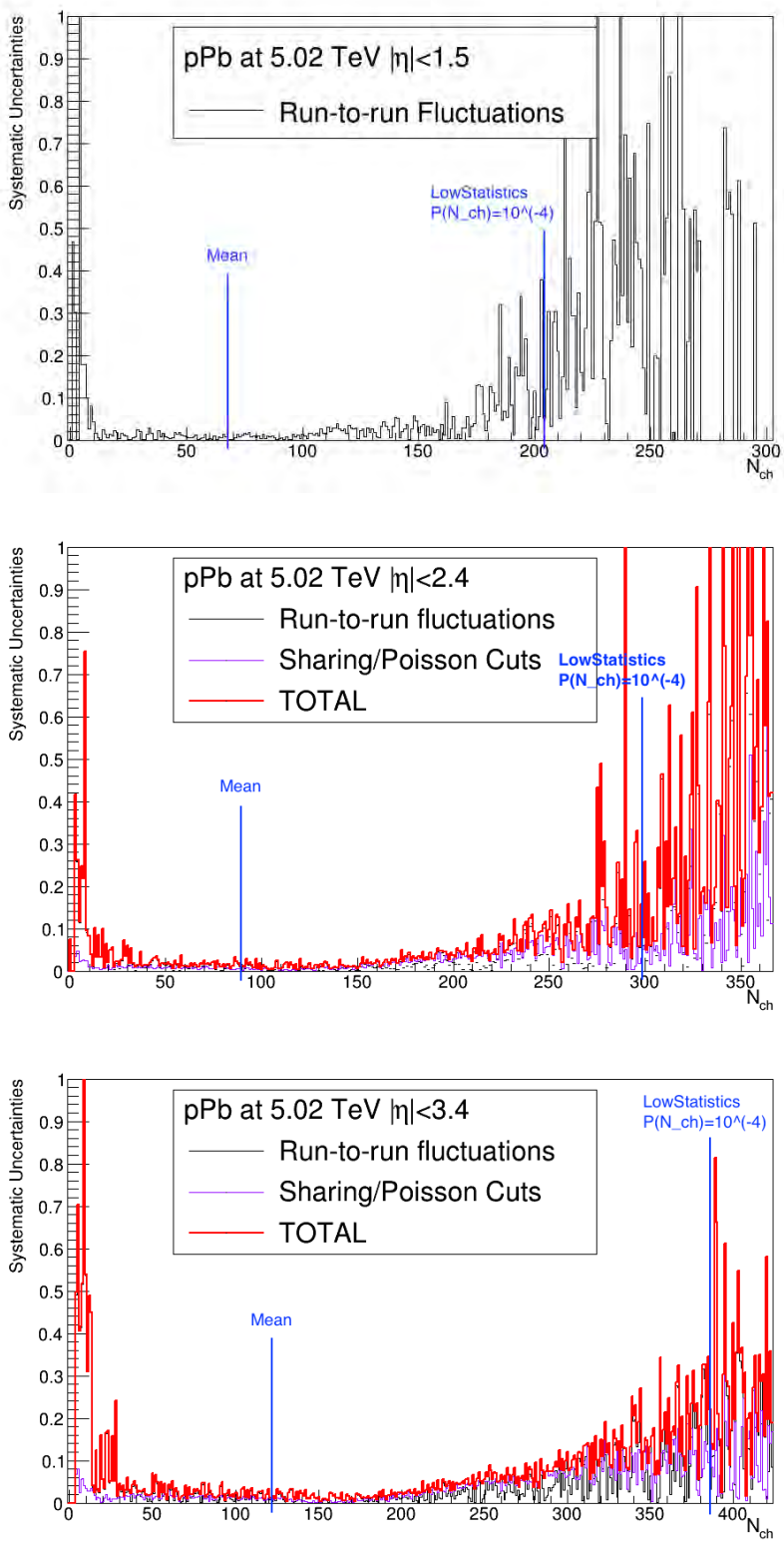


Figure 10.6: Systematic uncertainties overview plots for pPb data:
$\delta_x = |1 - P(N_{ch,x})/P(N_{ch,ref})|$. Top plot: pure SPD, middle: overlap region, bottom plot: full range. See text for details.

and Pbp is different, as mentioned in the previous sections).

The overall systematic uncertainties are underestimated, since the material budget uncertainty is missing, due to the non–availability of the MC production with additional material. Therefore, the systematic error around the mean is very low. Moreover, there are a lot of fluctuations which come from the fact that the unfolded spectrum is derived using non flat MC production.

10.3 Results

In this section, the results for pPb multiplicity distributions will be presented. They will be divided into two parts. In the first part, the symmetric pseudorapidity range result will be shown, from $|\eta| < 0.5$ to $|\eta| < 3.4$. In this case, the results for the pPb and the Pbp runs are merged. In the second part, instead, the separate pPb and Pbp results will be presented, for the asymmetric pseudorapidity range $-3.4 < \eta < +5.1$. All the results are re–binned in $5 \cdot N_{ch}$ in order to reduce the fluctuations originating from the unfolding procedure. In the third section, the results will be compared with the CGC based IP–Glasma model.

Multiplicity Distributions for Symmetric η Ranges

In Figure 10.7, the multiplicity distributions for merged pPb and Pbp are shown for seven symmetric eta ranges. Like in the pp case, the first three $|\eta| < 0.5$, $|\eta| < 1.0$ and $|\eta| < 1.5$ are obtained using only the SPD clusters. On the other side, the last four $|\eta| < 2.0$, $|\eta| < 2.4$, $|\eta| < 3.0$ and $|\eta| < 3.4$ are computed using SPD and FMD. The distributions are multiplied by multiples of 10, in order to not overlap.

It can be noted that using the FMD allows to measure around 600 charged particles, doubling the amount of particles detected in the central pseudorapidity region with the SPD. The amount of detected particles can be further increased with better statistics using a flat MC production to unfold.

Multiplicity Distributions for Asymmetric η Ranges

As mentioned in the previous sections, the asymmetric distributions for pPb and Pbp cannot be merged. The reason can be understood looking at Figure 10.8. In pPb (p–going direction) the Pb goes through FMD1 which is located opposite with respect to the muon spectrometer. In that case more charged particles are detected in the pseudorapidity range $-3.4 < \eta < +5.1$ which includes FMD1 ($+3.68 < \eta < +5.03$). On the contrary, when it is the p which hits the FMD1, fewer charged particles are measured in $-3.4 < \eta < +5.1$.

In Figure 10.9, the two distributions for Pbp, in black, and pPb, in red, are shown. The $P(0)$ is clearly lower in the pPb case, while the mean is higher. In fact the mean for pPb is 146.2 N_{ch} particles, and for Pbp it is 141.2 N_{ch} particles. Moreover, the red pPb distribution reaches higher multiplicity with respect to the black Pbp one.

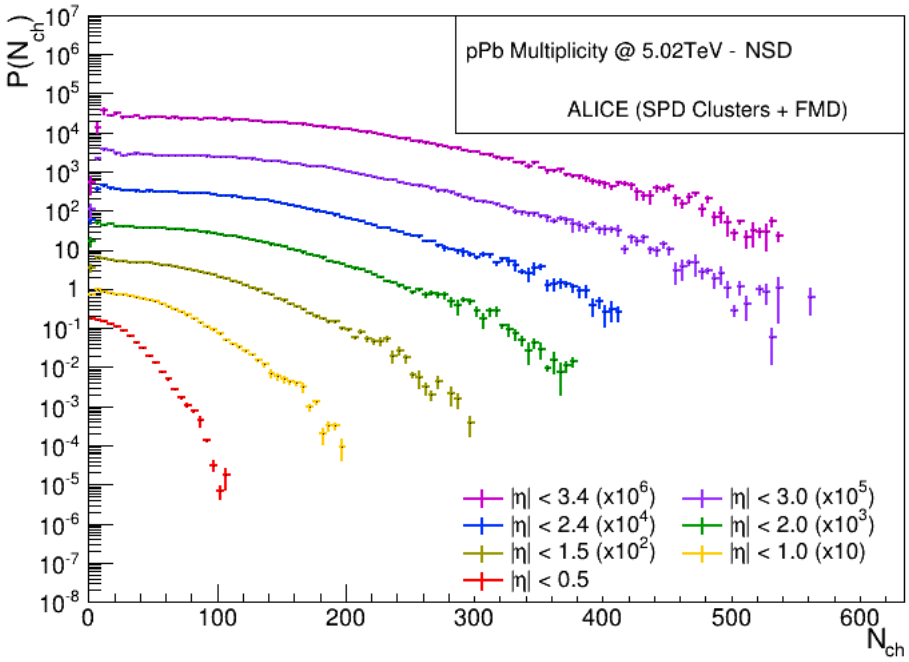


Figure 10.7: Charged–particle multiplicity distributions for pPb collisions at $\sqrt{s_{NN}} = 5.02$ TeV.

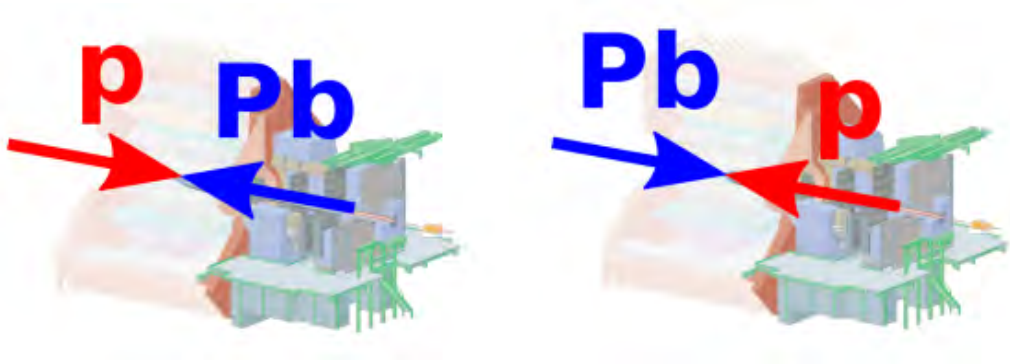


Figure 10.8: p–going (left) and Pb–going (right) collision scheme in ALICE.

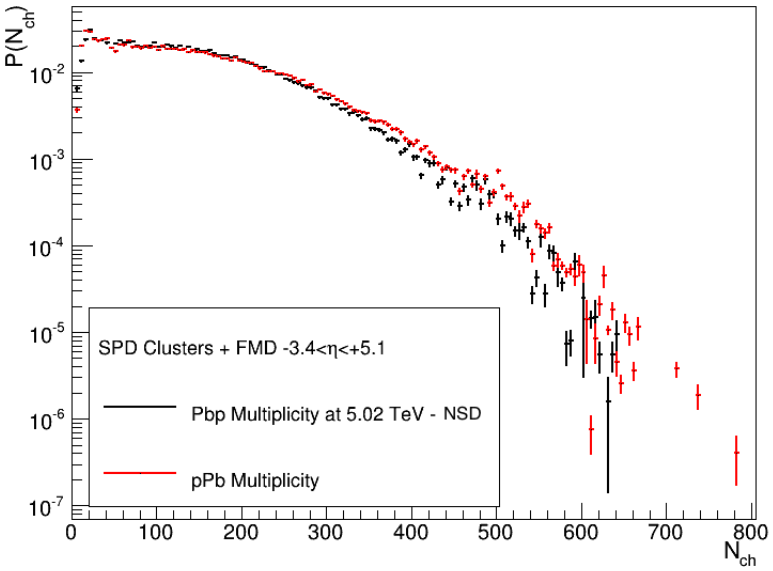


Figure 10.9: Charged–particle multiplicity distributions for pPb and Pbp collisions for $-3.4 < \eta < +5.1$.

Comparisons with IP–Glasma Model

As previously done for pp, it is possible to compare the results for pPb distributions at $|\eta| < 0.5$ with the ones obtained in the IP–Glasma framework. This CGC based model is already briefly explained in Section 9.1. Looking at Figure 10.10, it is clear that the distributions from the model agree with the data only for low multiplicities, up to $N_{ch} = 2 \cdot N_{ch,mean}$. Moreover, letting the color charge density fluctuate around the gluon saturation scale Q_s mean with a Gaussian with a wider variance $\sigma = 0.11$ (instead of the default value 0.09) gets the model closer to the experimental data.

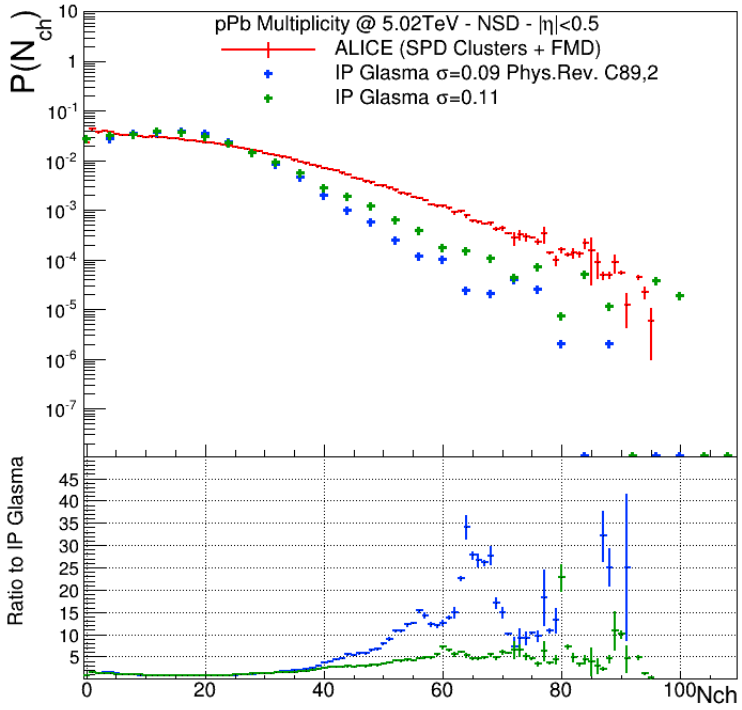


Figure 10.10: Charged–particle multiplicity distributions for pPb at $\sqrt{s_{NN}} = 5.02$ TeV compared with distributions from the IP–Glasma model, with $\sigma = 0.11$ (green) and $\sigma = 0.09$ (blue) [81]. The IP–Glasma points are taken from [82].

Chapter 11

Conclusions and Outlook

In this analysis, data collected combining the Forward Multiplicity Detector (FMD) and the Silicon Pixel Detector (SPD) were used (Chapter 5), resulting in a unique wide pseudorapidity coverage at LHC. In fact, using FMD and SPD allows to get a coverage of more the eight η units, from $-3.4 < \eta < +5.1$. The resulting charged–particle multiplicity distributions $P(N_{ch})$ are presented for four pp collision energies, from $\sqrt{s}$ =0.9 to 8 TeV and for three different trigger classes, INEL, INEL> 0 and NSD. Moreover, $P(N_{ch})$ are presented also for pPb and Pbp collisions at $\sqrt{s_{NN}} = 5.02$ TeV for NSD.

The results obtained for pp highly extend the pseudorapidity coverage of earliest results published by ALICE and CMS at LHC at midrapidity (Chapter 9), allowing to double the higher multiplicity. The η extension has higher systematic uncertainties due to the unknown fraction of material budget in front of the FMD, estimated to be around $5 - 10\%$ (Chapter 8).
The comparisons with previous pp ALICE and CMS published multiplicity distributions around $\eta = 0$ show very good agreement with the present work. The comparisons with ALICE distributions, computed only using the SPD, show very good agreement for both NSD and INEL> 0 at $\sqrt{s} = 0.9$ and 7 TeV. The only discrepancies are in the first bins, which are highly dependent on the triggering method and efficiency, and are due to different techniques used to derive the efficiency corrections. On the other side, the comparison with CMS for the NSD sample at 7 TeV is very sound. For 0.9 TeV the results of this work are systematically higher in the high multiplicity tails with respect to CMS. This came as no surprise, since CMS uses a different definition of the diffractive masses used to tune the Monte Carlo for the efficiency correction.
Both the Monte Carlo models (PYTHIA and PHOJET) and the Color Glass Condensate based IP–Glasma model used to compare with the pp experimental data underestimate the tails of the distributions. This comes from the fact that the multiplicity distributions are described mostly by non–perturbative physics, which turns out to be very difficult to model. The only exception is PHOJET at $\sqrt{s} = 0.9$ TeV which does a good job for the SPD region, around $\eta = 0$.
KNO scaling and double Negative Binomial distribution fits (Chapter 4) have been performed. It is found that KNO scaling is violated for all the three pseudorapidity

regions considered, $|\eta| < 1.5$, $|\eta| < 2.4$ and $-3.4 < \eta < +5.1$. The violation is stronger the wider the η range is. For all energies, double NBD fits are carried out and presented. The double fits are composed of a soft and a hard component. The fraction of soft events in pp 0.9 TeV collisions is about 86%, while it decreases to 60 % for 7 and 8 TeV. It is shown that, for the soft component, the KNO scaling is conserved both for $|\eta| < 1.5$ and $|\eta| < 2.4$. However, KNO is violated for the hard component.

The pPb collisions multiplicity distributions have not been published by other experiments at LHC, nor by ALICE, therefore the results shown here are brand new (Chapter 10). The pPb and Pbp distributions are shown for symmetrical and asymmetrical pseudorapidity ranges and are also compared with the CGC based IP–Glasma model. The model cannot reproduce the data, but it is shown that enhancing the color charge density fluctuations increases the tails of the calculations, improving the agreement with the data.
For the pPb distributions, additional work will still be needed to produce a final result, similar to what was done here for pp. A MC production with more statistics is needed to properly unfold the measured distributions, without them suffering form the present fluctuations.

A paper which contains the pp multiplicity distributions shown in this thesis is in preparation, a few more checks are needed in order for the paper to be ready. The most important one would be to require a flat MC production with additional material content, which will be needed in order to compute properly the systematic uncertainties from the material budget. Also, it needs to be understood the origin of the issues for the $\sqrt{s} = 2.76$ TeV data (Chapter 7).
The pp distributions will also be a hot topic in a few months, when LHC will restart with pp collisions at $\sqrt{s} = 13$ TeV. Now that the methodologies presented here are solid, it will be possible to quickly analyze the new data. The new data will increase of approximately 30% the present N_{ch} amount, allowing to study new exciting physics and explore harder regimes.

Appendices

Appendix A

Systematic Uncertainty Plots for pp collisions

In this appendix, the systematics for 2.76, 7 and 8 TeV are presented. The ones for 2.76 TeV must be taken with a grain of salt, because the quality of the data is quite compromised for the FMD, as pointed out in Section 7.2. For an extensive discussion of the systematic uncertainties, one should refer to Chapter 8

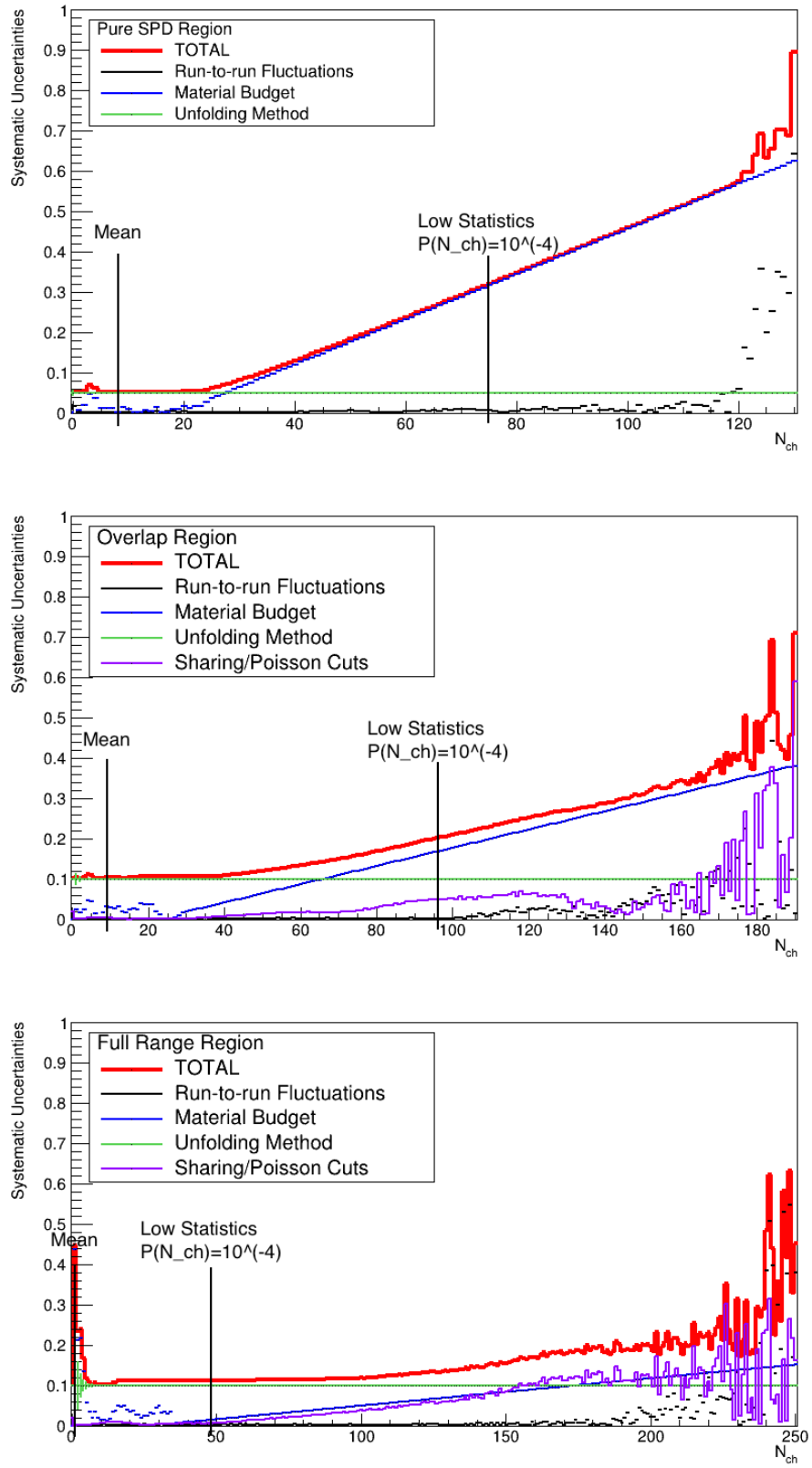
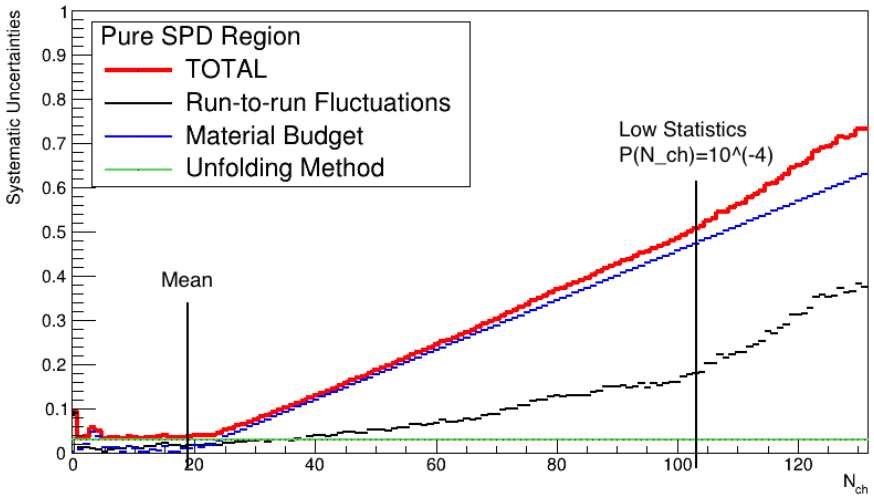
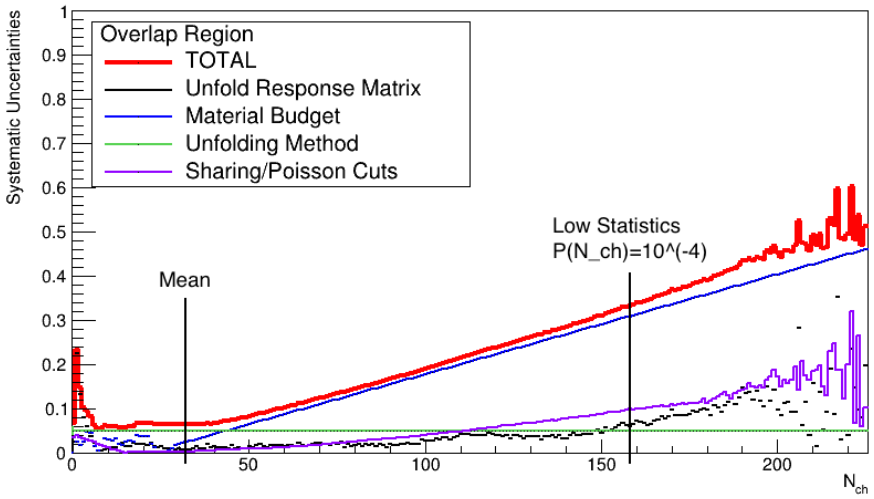


Figure A.1: Systematic uncertainties overview plots for 2.76 TeV data:
$\delta_x = |1 - P(N_{ch,x})/P(N_{ch,ref})|$.
Top plot: $|\eta| < 1.5$, middle: $|\eta| < 2.4$ and bottom plot: $-3.4 < \eta < +5.1$. See Section 7.2 for details on why the bottom distribution mean is so low.





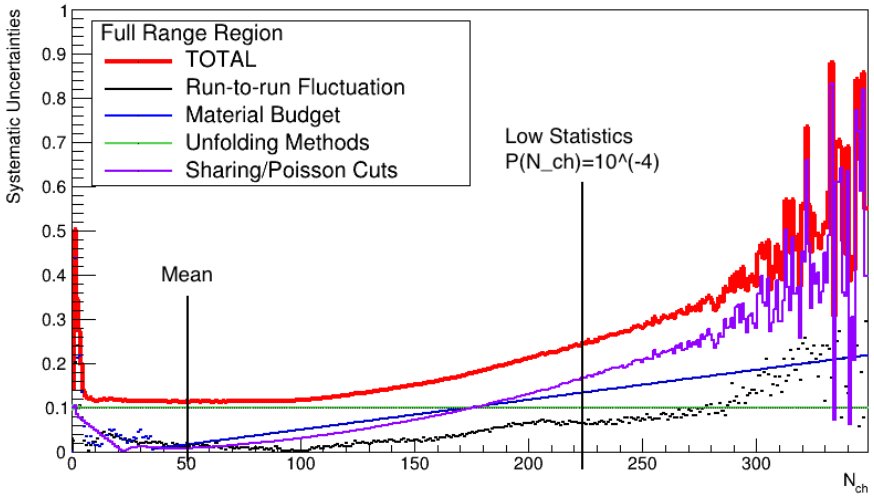


Figure A.2: Systematic uncertainties overview plots for 7 TeV data:
$\delta_x = |1 - P(N_{ch,x})/P(N_{ch,ref})|$.
Top plot: $|\eta| < 1.5$, middle: $|\eta| < 2.4$ and bottom plot: $-3.4 < \eta < +5.1$.

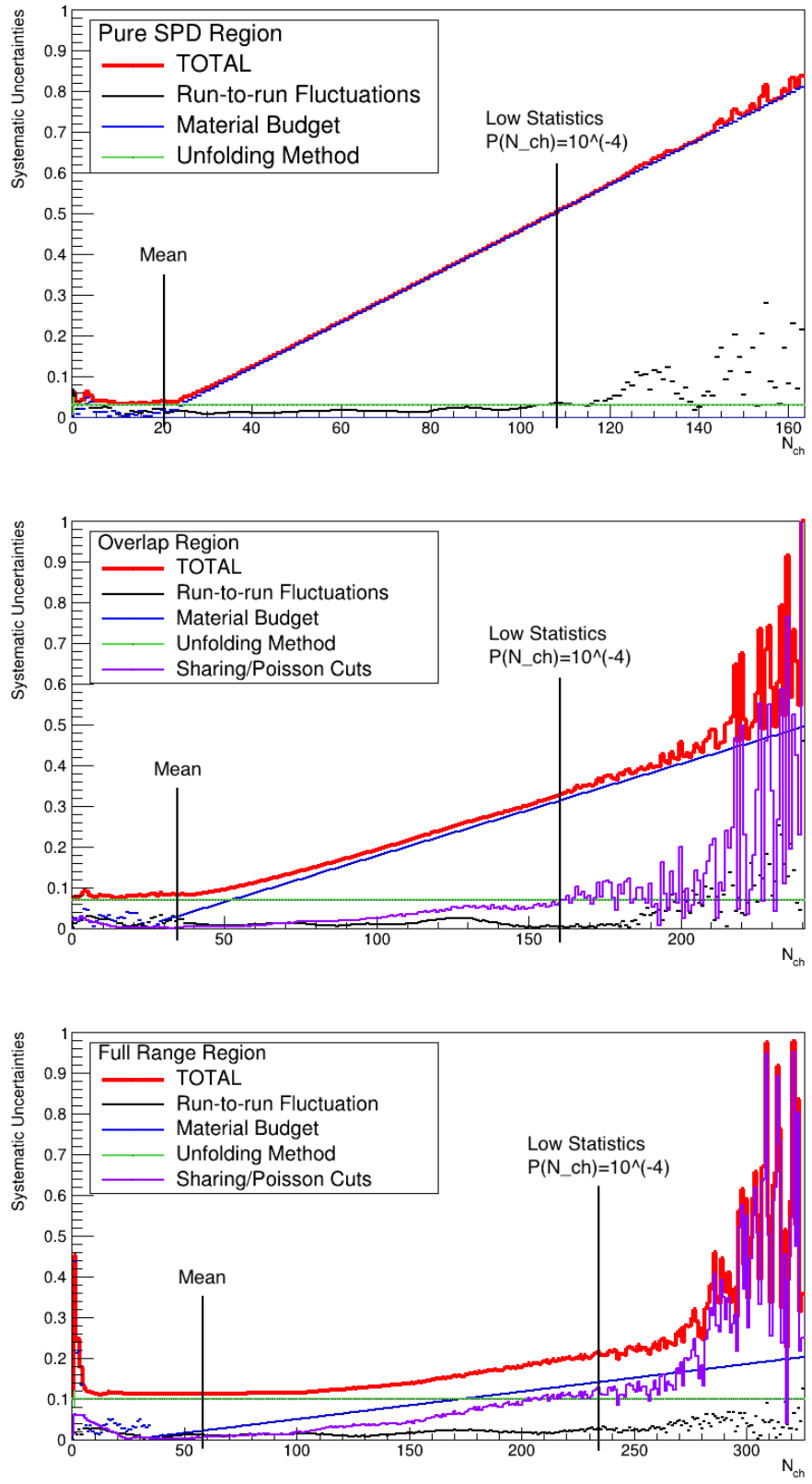


Figure A.3: Systematic uncertainties overview plots for 8 TeV data:
$\delta_x = |1 - P(N_{ch,x})/P(N_{ch,ref})|$.
Top plot: $|\eta| < 1.5$, middle: $|\eta| < 2.4$ and bottom plot: $-3.4 < \eta < +5.1$.

Appendix B

Additional Results for pp collisions

In this appendix, additional results for multiplicity distributions are shown.

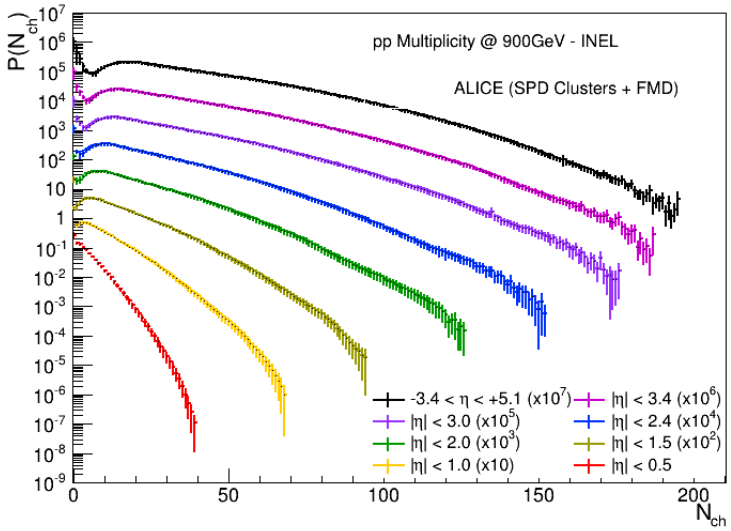


Figure B.1: Charged–particle multiplicity distributions for INEL pp collisions at $\sqrt{s} = 0.9$ TeV.

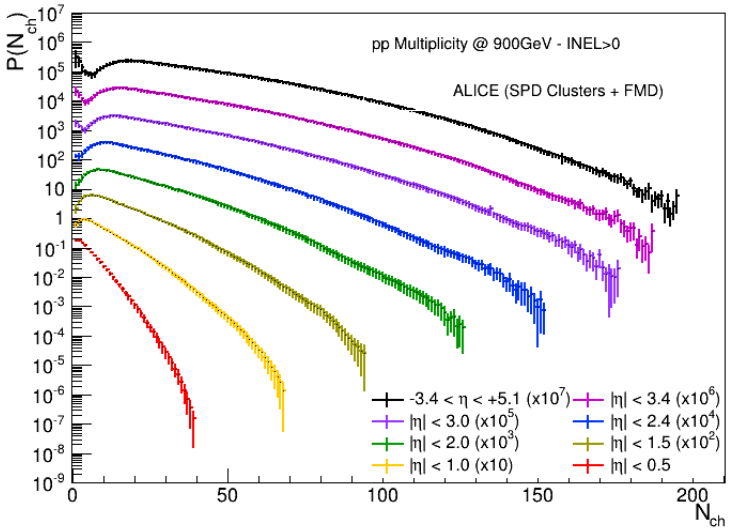


Figure B.2: Charged–particle multiplicity distributions for INEL> 0 pp collisions at $\sqrt{s} = 0.9$ TeV.

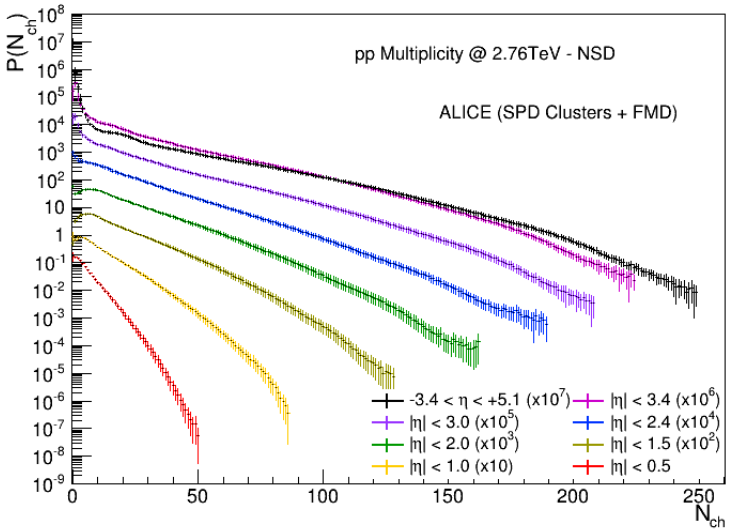


Figure B.3: Charged particle multiplicity distributions for NSD pp collisions at $\sqrt{s} = 2.76$ TeV. The combined FMD and SPD distributions, from the green to the black ones, are problematic, since they include FMD bad quality data (see Section 7.2).

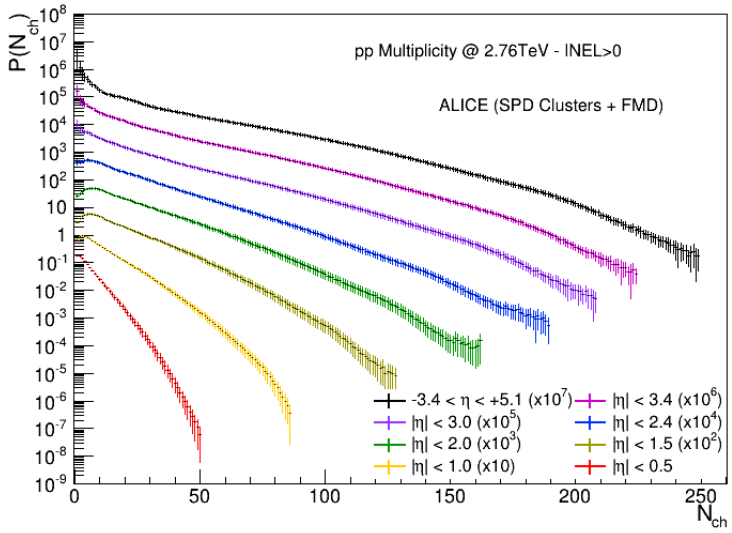


Figure B.4: Charged particle multiplicity distributions for INEL>0 pp collisions at $\sqrt{s} = 2.76$ TeV. The combined FMD and SPD distributions, from the green to the black ones, are problematic, since they include FMD bad quality data (see Section 7.2).

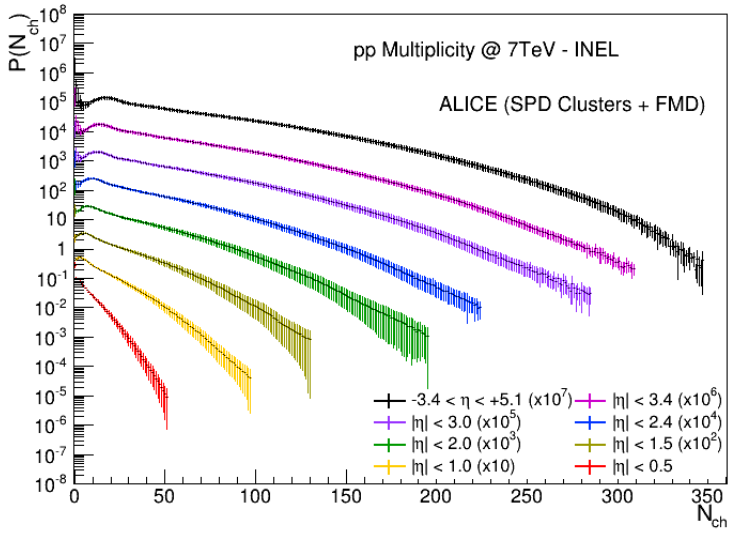


Figure B.5: Charged particle multiplicity distributions for INEL pp collisions at $\sqrt{s} = 7$ TeV.

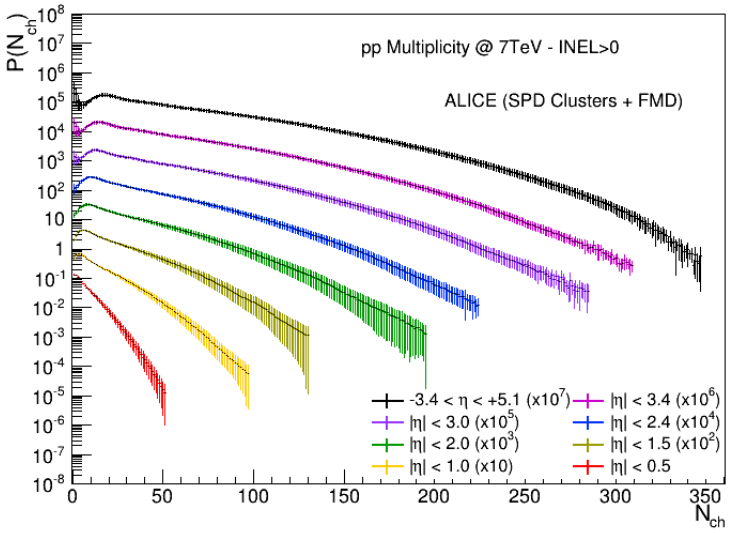


Figure B.6: Charged particle multiplicity distributions for INEL> 0 pp collisions at $\sqrt{s} = 7$ TeV.

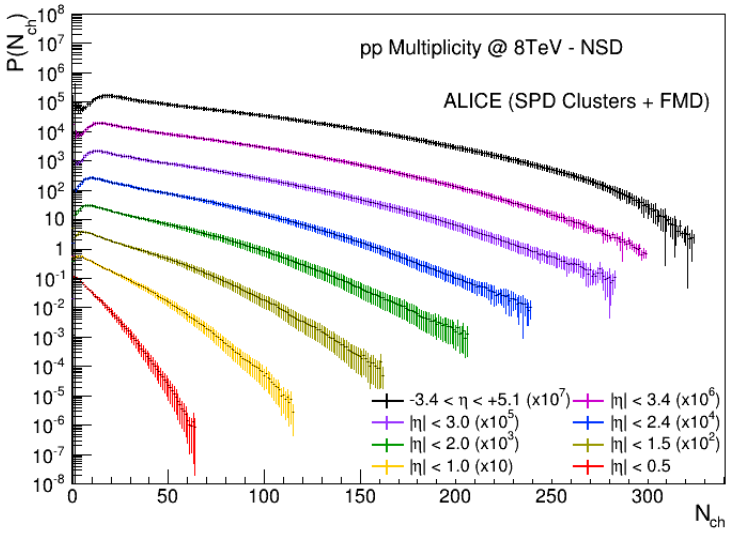


Figure B.7: Charged particle multiplicity distributions for NSD pp collisions at $\sqrt{s} = 8$ TeV.

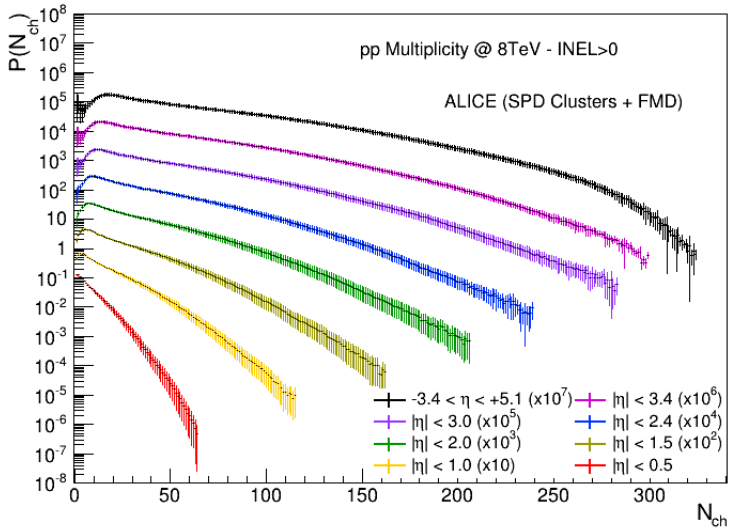
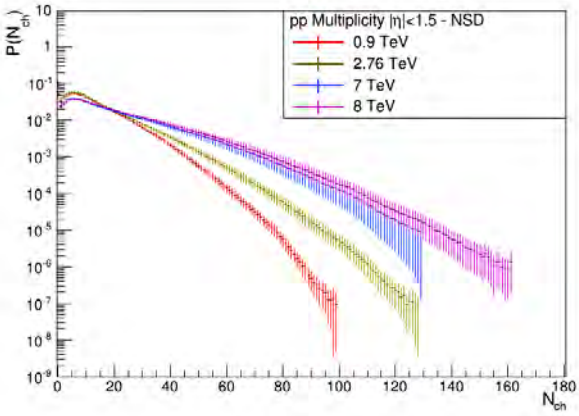
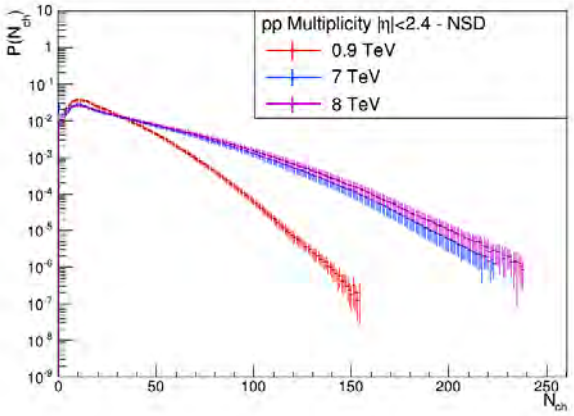


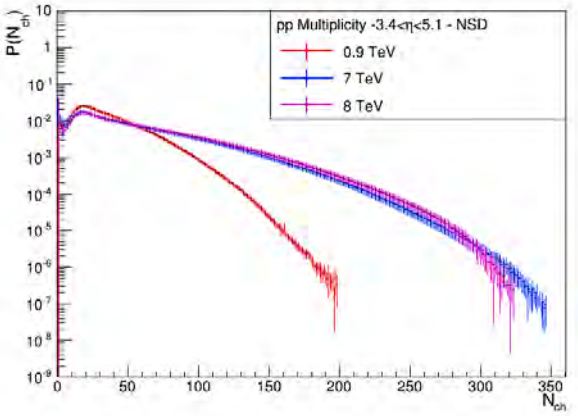
Figure B.8: Charged particle multiplicity distributions for INEL> 0 pp collisions at $\sqrt{s} = 8$ TeV.



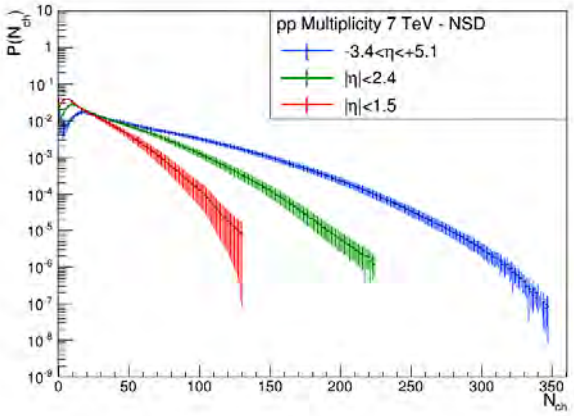
(a)



(b)



(c)



(d)

Figure B.9: (a-b-c) Evolution of charged multiplicity distributions as a function of $\sqrt{s}$, for NSD event classes for three different η ranges; (d) evolution of the distributions as a function of the η ranges for NSD events in $\sqrt{s} = 7$ TeV.

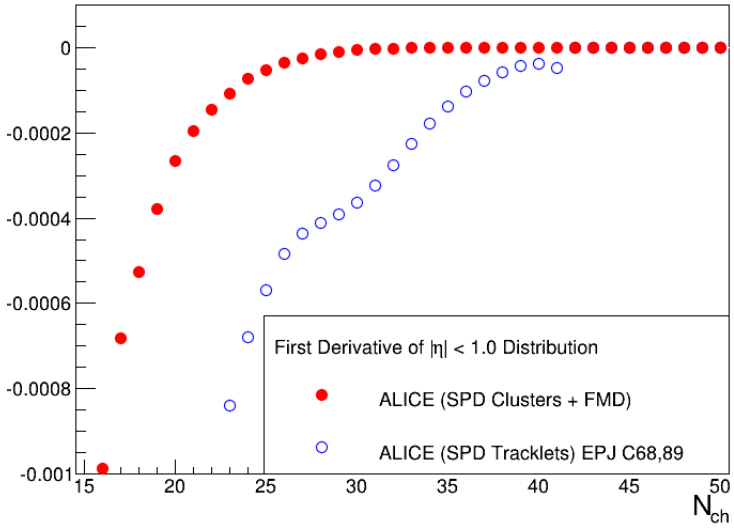


Figure B.10: First derivative of the $|\eta| < 1.0$ distribution for NSD events in 0.9 TeV. Comparison between this work and ALICE [58]. The published results oscillate, this comes most likely from the unfolding methodology used, χ^2 minimization, which produces oscillations in the tails of the distributions.

Bibliography

[1] G. Aad *et al.*, ATLAS Collaboration, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC," *Phys.Lett.*, vol. B716, pp. 1–29, 2012. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex].

[2] S. Chatrchyan *et al.*, CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC," *Phys.Lett.*, vol. B716, pp. 30–61, 2012. DOI: 10.1016/j.physletb.2012.08.021. arXiv: 1207.7235 [hep-ex].

[3] A. Cooper-Sarkar, ZEUS and H1 Collaboration, "Extraction of the proton parton density functions using a NLO-QCD fit of the combined H1 and ZEUS inclusive DIS cross sections," p. 25, 2008. DOI: 10.3360/dis.2008.25. arXiv: 0808.1854 [hep-ph].

[4] J. Breitweg *et al.*, ZEUS Collaboration, "ZEUS results on the measurement and phenomenology of F(2) at low x and low Q^2," *Eur.Phys.J.*, vol. C7, pp. 609–630, 1999. DOI: 10.1007/s100529901084. arXiv: hep-ex/9809005 [hep-ex].

[5] B. Hippolyte, "Bulk matter physics and its future at the Large Hadron Collider," *Eur.Phys.J.*, vol. C62, pp. 237–242, 2009. DOI: 10.1140/epjc/s10052-009-0910-9. arXiv: 0901.3176 [hep-ex].

[6] P. Collins, "An Introduction to Regge Theory and High-Energy Physics," 1977.

[7] G. Chew and S. C. Frautschi, "Principle of Equivalence for All Strongly Interacting Particles Within the S Matrix Framework," *Phys.Rev.Lett.*, vol. 7, pp. 394–397, 1961. DOI: 10.1103/PhysRevLett.7.394.

[8] J. R. Forshaw and D. Ross, "Quantum chromodynamics and the pomeron," *Cambridge Lect.Notes Phys.*, vol. 9, pp. 1–248, 1997.

[9] S. Navin, "Diffraction in Pythia," 2010. arXiv: 1005.3894 [hep-ph].

[10] M. Deile *et al.*, TOTEM Collaboration, "Diffraction and total cross-section at the Tevatron and the LHC," *Springer Proc.Phys.*, vol. 108, pp. 40–45, 2006. DOI: 10.1007/978-3-540-32841-4_7. arXiv: hep-ex/0602021 [hep-ex].

[11] T. Sjostrand, S. Mrenna, and P. Z. Skands, "PYTHIA 6.4 Physics and Manual," *JHEP*, vol. 0605, p. 026, 2006. DOI: 10.1088/1126-6708/2006/05/026. arXiv: hep-ph/0603175 [hep-ph].

[12] E. Iancu, "QCD in heavy ion collisions," 2012. arXiv: 1205.0579 [hep-ph].

[13] E. Iancu and R. Venugopalan, "The Color glass condensate and high-energy scattering in QCD," 2003. arXiv: hep-ph/0303204 [hep-ph].

[14] V. Gribov and L. Lipatov, "Deep inelastic e p scattering in perturbation theory," *Sov.J.Nucl.Phys.*, vol. 15, pp. 438–450, 1972.

[15] G. Altarelli and G. Parisi, "Asymptotic Freedom in Parton Language," *Nucl.Phys.*, vol. B126, p. 298, 1977. DOI: 10.1016/0550-3213(77)90384-4.

[16] 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*, vol. 46, pp. 641–653, 1977.

[17] L. Lipatov, "Reggeization of the Vector Meson and the Vacuum Singularity in Nonabelian Gauge Theories," *Sov.J.Nucl.Phys.*, vol. 23, pp. 338–345, 1976.

[18] E. Kuraev, L. Lipatov, and V. S. Fadin, "The Pomeranchuk Singularity in Nonabelian Gauge Theories," *Sov.Phys.JETP*, vol. 45, pp. 199–204, 1977.

[19] I. Balitsky and L. Lipatov, "The Pomeranchuk Singularity in Quantum Chromodynamics," *Sov.J.Nucl.Phys.*, vol. 28, pp. 822–829, 1978.

[20] L. McLerran, "A Brief Introduction to the Color Glass Condensate and the Glasma," pp. 3–18, 2008. DOI: 10.3204/DESY-PROC-2009-01/26. arXiv: 0812.4989 [hep-ph].

[21] V. Khachatryan *et al.*, CMS Collaboration, "Observation of Long-Range Near-Side Angular Correlations in Proton-Proton Collisions at the LHC," *JHEP*, vol. 1009, p. 091, 2010. DOI: 10.1007/JHEP09(2010)091. arXiv: 1009.4122 [hep-ex].

[22] T. Lee and G. Wick, "Vacuum Stability and Vacuum Excitation in a Spin 0 Field Theory," *Phys.Rev.*, vol. D9, pp. 2291–2316, 1974. DOI: 10.1103/PhysRevD.9.2291.

[23] J. C. Collins and M. Perry, "Superdense Matter: Neutrons Or Asymptotically Free Quarks?" *Phys.Rev.Lett.*, vol. 34, p. 1353, 1975. DOI: 10.1103/PhysRevLett.34.1353.

[24] N. Cabibbo and G. Parisi, "Exponential Hadronic Spectrum and Quark Liberation," *Phys.Lett.*, vol. B59, pp. 67–69, 1975. DOI: 10.1016/0370-2693(75)90158-6.

[25] R. Hagedorn, "Statistical thermodynamics of strong interactions at high-energies," *Nuovo Cim.Suppl.*, vol. 3, pp. 147–186, 1965.

[26] E. V. Shuryak, "Quantum Chromodynamics and the Theory of Superdense Matter," *Phys.Rept.*, vol. 61, pp. 71–158, 1980. DOI: 10.1016/0370-1573(80)90105-2.

[27] F. Karsch, "Lattice results on QCD thermodynamics," *Nucl.Phys.*, vol. A698, pp. 199–208, 2002. DOI: 10.1016/S0375-9474(01)01365-3. arXiv: hep-ph/0103314 [hep-ph].

[28] R. Rapp and J. Wambach, "Chiral symmetry restoration and dileptons in relativistic heavy ion collisions," *Adv.Nucl.Phys.*, vol. 25, p. 1, 2000. DOI: 10.1007/0-306-47101-9_1. arXiv: hep-ph/9909229 [hep-ph].

[29] M. Gyulassy and L. McLerran, "New forms of QCD matter discovered at RHIC," *Nucl.Phys.*, vol. A750, pp. 30–63, 2005. DOI: 10.1016/j.nuclphysa.2004.10.034. arXiv: nucl-th/0405013 [nucl-th].

[30] M. G. Alford, "QCD at high density / temperature," *Nucl.Phys.Proc.Suppl.*, vol. 117, pp. 65–82, 2003. DOI: 10.1016/S0920-5632(03)01411-7. arXiv: hep-ph/0209287 [hep-ph].

[31] A. Dumitru, K. Dusling, F. Gelis, J. Jalilian-Marian, T. Lappi, *et al.*, "The Ridge in proton-proton collisions at the LHC," *Phys.Lett.*, vol. B697, pp. 21–25, 2011. DOI: 10.1016/j.physletb.2011.01.024. arXiv: 1009.5295 [hep-ph].

[32] P. Z. Skands, "Tuning Monte Carlo Generators: The Perugia Tunes," *Phys.Rev.*, vol. D82, p. 074 018, 2010. DOI: 10.1103/PhysRevD.82.074018. arXiv: 1005.3457 [hep-ph].

[33] T. Pierog, I. Karpenko, J. Katzy, E. Yatsenko, and K. Werner, "EPOS LHC : test of collective hadronization with LHC data," 2013. arXiv: 1306.0121 [hep-ph].

[34] B. B. Abelev *et al.*, ALICE Collaboration, "Multiplicity dependence of the average transverse momentum in pp, p-Pb, and Pb-Pb collisions at the LHC," *Phys.Lett.*, vol. B727, pp. 371–380, 2013. DOI: 10.1016/j.physletb.2013.10.054. arXiv: 1307.1094 [nucl-ex].

[35] B. B. Abelev *et al.*, ALICE Collaboration, "Multiplicity Dependence of Pion, Kaon, Proton and Lambda Production in p-Pb Collisions at $\sqrt{s_{NN}}$ = 5.02 TeV," *Phys.Lett.*, vol. B728, pp. 25–38, 2014. DOI: 10.1016/j.physletb.2013.11.020. arXiv: 1307.6796 [nucl-ex].

[36] B. B. Abelev *et al.*, ALICE Collaboration, "Long-range angular correlations on the near and away side in p-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV," *Phys.Lett.*, vol. B719, pp. 29–41, 2013. DOI: 10.1016/j.physletb.2013.01.012. arXiv: 1212.2001 [nucl-ex].

[37] S. Chatrchyan *et al.*, CMS Collaboration, "Observation of long-range near-side angular correlations in proton-lead collisions at the LHC," *Phys.Lett.*, vol. B718, pp. 795–814, 2013. DOI: 10.1016/j.physletb.2012.11.025. arXiv: 1210.5482 [nucl-ex].

[38] G. Aad *et al.*, ATLAS Collaboration, "Observation of Associated Near-Side and Away-Side Long-Range Correlations in $\sqrt{s_{NN}}$=5.02TeV Proton-Lead Collisions with the ATLAS Detector," *Phys.Rev.Lett.*, vol. 110, no. 18, p. 182 302, 2013. DOI: 10.1103/PhysRevLett.110.182302. arXiv: 1212.5198 [hep-ex].

[39] K. Dusling and R. Venugopalan, "Comparison of the color glass condensate to dihadron correlations in proton-proton and proton-nucleus collisions," *Phys.Rev.*, vol. D87, no. 9, p. 094 034, 2013. DOI: 10.1103/PhysRevD.87.094034. arXiv: 1302.7018 [hep-ph].

[40] P. Bozek and W. Broniowski, "Correlations from hydrodynamic flow in p-Pb collisions," *Phys.Lett.*, vol. B718, pp. 1557–1561, 2013. DOI: 10.1016/j.physletb.2012.12.051. arXiv: 1211.0845 [nucl-th].

[41] T. Matsui and H. Satz, "J/ψ Suppression by Quark-Gluon Plasma Formation," *Phys.Lett.*, vol. B178, p. 416, 1986. DOI: 10.1016/0370-2693(86)91404-8.

[42] B. B. Abelev *et al.*, ALICE Collaboration, "J/ψ production and nuclear effects in p-Pb collisions at $\sqrt{S_{NN}}$ = 5.02 TeV," *JHEP*, vol. 1402, p. 073, 2014. DOI: 10.1007/JHEP02(2014)073. arXiv: 1308.6726 [nucl-ex].

[43] B. B. Abelev *et al.*, ALICE Collaboration, "Suppression of ψ(2S) production in p-Pb collisions at $\sqrt{s_{NN}}$ = 5.02 TeV," 2014. arXiv: 1405.3796 [nucl-ex].

[44] R. Arnaldi, ALICE Collaboration, "Inclusive ψ(2S) production in p-Pb collisions with ALICE," *Nucl.Phys.*, vol. A, 2014. DOI: 10.1016/j.nuclphysa.2014.08.019. arXiv: 1407.7451 [nucl-ex].

[45] R. P. Feynman, "Very high-energy collisions of hadrons," *Phys.Rev.Lett.*, vol. 23, pp. 1415–1417, 1969. DOI: 10.1103/PhysRevLett.23.1415.

[46] Z. Koba, H. B. Nielsen, and P. Olesen, "Scaling of multiplicity distributions in high-energy hadron collisions," *Nucl.Phys.*, vol. B40, pp. 317–334, 1972. DOI: 10.1016/0550-3213(72)90551-2.

[47] G. Alner *et al.*, UA5 Collaboration, "An Investigation of Multiplicity Distributions in Different Pseudorapidity Intervals in anti-p p Reactions at a CMS Energy of 540-GeV," *Phys.Lett.*, vol. B160, p. 193, 1985. DOI: 10.1016/0370-2693(85)91491-1.

[48] G. Alner *et al.*, UA5 Collaboration, "Scaling Violations in Multiplicity Distributions at 200-GeV and 900-GeV," *Phys.Lett.*, vol. B167, pp. 476–480, 1986. DOI: 10.1016/0370-2693(86)91304-3.

[49] A. Giovannini and L. Van Hove, "Negative Binomial Multiplicity Distributions in High-Energy Hadron Collisions," *Z.Phys.*, vol. C30, p. 391, 1986. DOI: 10.1007/BF01557602.

[50] A. Giovannini, "*thermal chaos* and *coherence* in multiplicity distributions at high energies," *Nuovo Cim.*, vol. A15, pp. 543–551, 1973. DOI: 10.1007/BF02734689.

[51] C. Fuglesang, "UA5 Multiplcity Distributions and Fits of Various Functions," 1989.

[52] A. Giovannini and R. Ugoccioni, "Possible scenarios for soft and semihard components structure in central hadron hadron collisions in the TeV region," *Phys.Rev.*, vol. D59, p. 094 020, 1999. DOI: 10.1103/PhysRevD.59.094020, 10.1103/PhysRevD.59.09402010.1103/PhysRevD.69.059903, 10.1103/PhysRevD.69.059903. arXiv: hep-ph/9810446 [hep-ph].

[53] A. Giovannini and R. Ugoccioni, "On signals of new physics in global event properties in pp collisions in the TeV energy domain," *Phys.Rev.*, vol. D68, p. 034 009, 2003. DOI: 10.1103/PhysRevD.68.034009. arXiv: hep-ph/0304128 [hep-ph].

[54] R. Ugoccioni and A. Giovannini, "Main results of a search on multiplicity distributions in pp collisions: Is anybody afraid of a new class of hard events?" *Acta Phys.Polon.*, vol. B36, pp. 755–760, 2005. arXiv: hep-ph/0410202 [hep-ph].

[55] F. Rimondi, CDF Collaboration, "Multiplicity distributions in $\bar{p}p$ interactions at $\sqrt{s}$ = 1800-GeV," 1993.

[56] V. Khachatryan *et al.*, CMS Collaboration, "Charged particle multiplicities in pp interactions at $\sqrt{s} = 0.9$, 2.36, and 7 TeV," *JHEP*, vol. 1101, p. 079, 2011. DOI: 10.1007/JHEP01(2011)079. arXiv: 1011.5531 [hep-ex].

[57] R. Ansorge *et al.*, UA5 Collaboration, "Charged Particle Multiplicity Distributions at 200-GeV and 900-GeV Center-Of-Mass Energy," *Z.Phys.*, vol. C43, p. 357, 1989. DOI: 10.1007/BF01506531.

[58] K. Aamodt *et al.*, 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.*, vol. C68, pp. 89–108, 2010. DOI: 10.1140/epjc/s10052-010-1339-x. arXiv: 1004.3034 [hep-ex].

[59] K. Aamodt *et al.*, ALICE Collaboration, "Charged-particle multiplicity measurement in proton-proton collisions at $\sqrt{s} = 7$ TeV with ALICE at LHC," *Eur.Phys.J.*, vol. C68, pp. 345–354, 2010. DOI: 10.1140/epjc/s10052-010-1350-2. arXiv: 1004.3514 [hep-ex].

[60] K. Aamodt *et al.*, ALICE Collaboration, "The ALICE experiment at the CERN LHC," *JINST*, vol. 3, S08002, 2008. DOI: 10.1088/1748-0221/3/08/S08002.

[61] B. B. Abelev *et al.*, ALICE Collaboration, "Performance of the ALICE Experiment at the CERN LHC," *Int.J.Mod.Phys.*, vol. A29, p. 1 430 044, 2014. DOI: 10.1142/S0217751X14300440. arXiv: 1402.4476 [nucl-ex].

[62] H. H. Dalsgaard, "Pseudorapidity Densities in p+p and Pb+Pb collisions at LHC measured with the ALICE experiment," PhD thesis, Niels Bohr Institute, University of Copenhagen, 2012. [Online]. Available: http://www.nbi.dk/~canute/thesis.pdf.

[63] C. Nygaard, "Charged particle multiplicity distributions into forward pseudorapidities in pp and PbPb collisions at the LHC," PhD thesis, Niels Bohr Institute, University of Copenhagen, 2011. [Online]. Available: http://www.nbi.dk/~cnygaard/cnygaard_phd.pdf.

[64] C. H. Christensen and H. H. Dalsgaard, "Analysing the FMD data for $1/N\ dN_{ch}/d\eta$," *ALICE Internal Note*, 2012. [Online]. Available: https://aliceinfo.cern.ch/Notes/node/107.

[65] P. Z. Skands, "The Perugia Tunes," pp. 284–297, 2009. arXiv: 0905.3418 [hep-ph].

[66] D. d'Enterria *et al.*, "Constraints from the first LHC data on hadronic event generators for ultra-high energy cosmic-ray physics," *Astropart.Phys.*, vol. 35, pp. 98–113, 2011. DOI: 10.1016/j.astropartphys.2011.05.002. arXiv: 1101.5596 [astro-ph].

[67] A. B. Kaidalov and M. G. Poghosyan, "Description of soft diffraction in the framework of reggeon calculus: Predictions for LHC," pp. 91–98, 2009. arXiv: 0909.5156 [hep-ph].

[68] F. W. Bopp, R. Engel, and J. Ranft, "Rapidity gaps and the PHOJET Monte Carlo," pp. 729–741, 1998. arXiv: hep-ph/9803437 [hep-ph].

[69] S. Roesler, R. Engel, and J. Ranft, "The Monte Carlo event generator DPMJET-III," pp. 1033–1038, 2000. arXiv: hep-ph/0012252 [hep-ph].

[70] S. Y. Shmakov, V. Uzhinsky, and A. Zadorozhnyi, "DIAGEN: Generator of Inelastic Nucleus–Nucleus Interaction Diagrams," *Comput. Phys. Commun.*, 1988.

[71] R. Brun, F. Carminati, and S. Giani, "GEANT Detector Description and Simulation Tool," 1994.

[72] A. C. Hansen, "Pseudorapidity Dependence of Anisotropic Azimuthal Flow with the ALICE Detector," PhD thesis, Niels Bohr Institute, University of Copenhagen, 2014. [Online]. Available: http://www.nbi.dk/~alex/pdf/thesis.pdf.

[73] C. H. Christensen, "ALICE Forward Multiplicity Detector," PhD thesis, Niels Bohr Institute, University of Copenhagen, 2007. [Online]. Available: http://www.nbi.dk/~cholm/phd.pdf.

[74] Carminati and Bruckner. (2007). The ALICE Offline Bible, [Online]. Available: http://aliweb.cern.ch/secure/Offline/sites/aliweb.cern.ch.Offline/files/uploads/OfflineBible.pdf.

[75] The ROOT Team. (1995-2015). ROOT Data Analysis Framework, [Online]. Available: https://root.cern.ch/.

[76] A. Alkin, Private Communication, 2015.

[77] C. V. Cheshkov and E. Kryshen, Private Communication, 2015.

[78] K. Aamodt *et al.*, ALICE, "Strange particle production in proton-proton collisions at sqrt(s) = 0.9 TeV with ALICE at the LHC," *Eur.Phys.J.*, vol. C71, p. 1594, 2011. DOI: 10.1140/epjc/s10052-011-1594-5. arXiv: 1012.3257 [hep-ex].

[79] J. F. Grosse-Oetringhaus, "Measurement of the Charged-Particle Multiplicity in ProtonProton Collisions with the ALICE Detector," PhD thesis, University of Muensten, 2009. [Online]. Available: http://cds.cern.ch/record/1175646/files/CERN-THESIS-2009-033.pdf.

[80] F. Gelis, T. Lappi, and L. McLerran, "Glittering Glasmas," *Nucl.Phys.*, vol. A828, pp. 149–160, 2009. DOI: 10.1016/j.nuclphysa.2009.07.004. arXiv: 0905.3234 [hep-ph].

[81] B. Schenke, P. Tribedy, and R. Venugopalan, "Multiplicity distributions in p+p, p+A and A+A collisions from Yang-Mills dynamics," *Phys.Rev.*, vol. C89, no. 2, p. 024 901, 2014. DOI: 10.1103/PhysRevC.89.024901. arXiv: 1311.3636 [hep-ph].

[82] B. Schenke, Private Communication, 2015.

[83] P. Ghosh, "Negative binomial multiplicity distribution in proton-proton collisions in limited pseudorapidity intervals at LHC up to $\sqrt{s} = 7$ TeV and the clan model," *Phys.Rev.*, vol. D85, p. 054 017, 2012. DOI: 10.1103/PhysRevD.85.054017. arXiv: 1202.4221 [hep-ph].

List of Figures

List of Tables