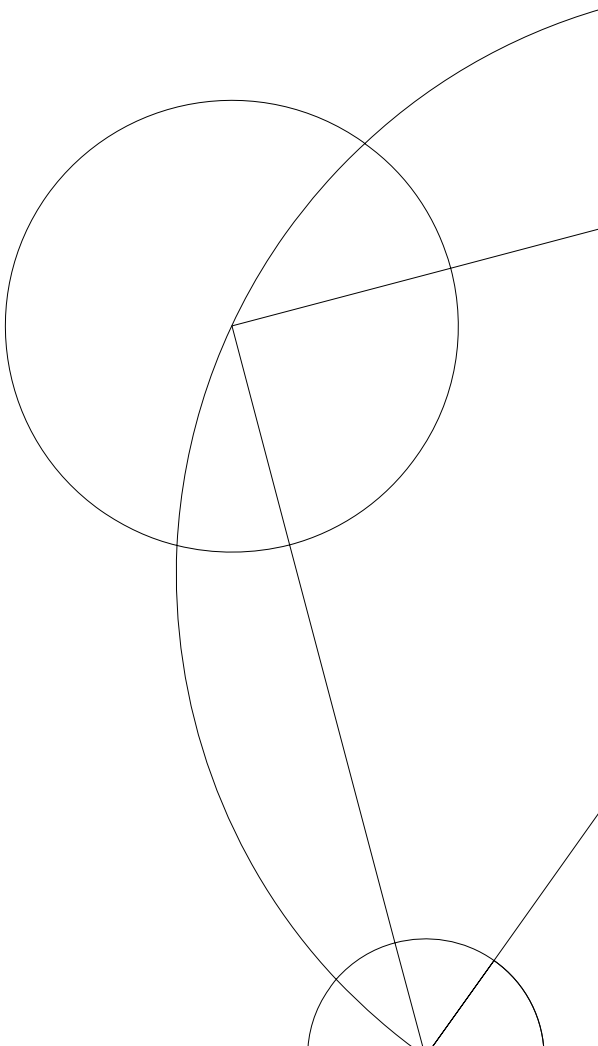


FACULTY OF SCIENCE
UNIVERSITY OF COPENHAGEN

PhD Thesis

Laís Ozelin de Lima Pimentel

Charged-particle multiplicity distributions in p-Pb collisions at $\sqrt{s_{NN}}$ = 5.02 TeV with ALICE



Academic Advisor
Jens Jørgen Gaardhøje
Co-Advisor
Valentina Zaccolo
November 1, 2018

Charged-particle multiplicity in p-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV with ALICE

by *Laís Ozelin de Lima Pimentel*

Academic advisor:

Prof. Jens Jørgen Gaardhøje

Co-advisor:

Dr. Valentina Zaccolo

High Energy Heavy-Ion Group

Niels Bohr Institute
University of Copenhagen



This PhD thesis has been submitted to the PhD School of The Faculty of Science, University of Copenhagen on the 1st of November 2018

Abstract

Approximately 1 μs after the Big Bang, the Universe was in the state of a Quark-Gluon Plasma (QGP), an ultra hot and dense state of matter, consisting of quarks and gluons. With the Large Hadron Collider (LHC) at CERN, in Switzerland, the properties of the QGP are studied by recreating this state of matter in collisions between lead ions. Studies of collisions between protons and between protons and lead ions have shown that smaller systems can also possess features of collectivity reminiscent of a QGP. This is a surprising and interesting discovery since it was thought until recently that only large overlap volumes, reached in collisions between nuclei, could present such a feature.

The charged-particle multiplicity is one of the most fundamental measurements. It provides insights into the mechanisms of particle production, and it is sensitive to the number of collisions between quarks and gluons. In this work, the charged-particle multiplicity distribution in p-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV over a wide kinematic range ($-3.4 < \eta < 5.1$) is measured and presented, using LHC data from run 2, taken in 2016. The data were obtained using the Forward Multiplicity Detector (FMD) and the Silicon Pixel Detector (SPD) of ALICE (A Large Ion Collider Experiment) at the LHC. Differential results are also obtained by dividing the multiplicity distributions into centrality classes (classes that are proportional to the number of interactions the proton undergoes as it collides with the lead-ion), using the V0A and ZNA centrality estimators.

The results obtained are compared with Monte Carlo models available, EPOS LHC and DPMJET, and IP-glasma data available in the literature. It is suggested that EPOS LHC and DPMJET reproduce the measured distributions well, while the IP-Glasma model fails.

Dansk resumé

Omkring 1 μs efter Big Bang var der i universet et kvark-gluon-plasma ("QGP" efter "Quark Gluon Plasma"), en kompakt og utrolig varm tilstand, som bestod af frie kvarker og gluoner. Med Large Hadron Collider (LHC) på CERN, i Schweiz, de QGPs egenskaber bliver undersøgt ved at genskabe denne stofstilstande gennem kollisioner mellem blykerner. Undersøgelser af kollisioner mellem protoner og blykerner har vist, at små systemer også har kollektive egenskaber der minder om et QGP. Det er en overraskende og interessant opdagelse da man indtil for nyligt troede at blot store overlapnings volumener, nået ved at kollidere blykerner, kunne vise en sådan egenskab.

Elektrisk ladede partikel multiplicitetsfordeling er en af de meste grundlæggende målinger. Det giver indsigt i partikelproduktions mekanisme og er følsom over for antallet af kollisioner mellem kvarker og gluoner. I denne afhandling måles og præsenteres elektrisk ladede partikel multiplicitetsfordelingen ved $\sqrt{s_{NN}} = 5.02$ TeV udført over et bredt kinematisk område ($-3.4 < \eta < 5.1$). Dataene var indsamlet ved brug af Forward Multiplicity Detector(FMD) og Silicon Pixel Detector (SPD) af ALICE (A Large Ion Collider Experiment) i LHC. Differentierede resultater opnås ved at dele multiplicitetsfordelingen op i centralitetsklasser (klasser der er proportionale til antallet af vekselvirkninger som protonen gennemgår da den støder sammen med blykernen) ved brug af V0A og ZNA centralitetsestimatorer.

Resultaterne sammenlignes med foreliggende Monte Carlo modeller, EPOS LHC og DPMJET, samt IP-glasma data, som er tilgængelig i litteraturen. Det antyder, at EPOS LHC og DPMJET beskriver de målte fordelinger godt mens IP-Glasma modellen mislykkes.

Acknowledgements

Even though people very often work alone too much during their Ph.D., it is definitely impossible to go through it completely alone. Many people were part of my Ph.D. journey and, without them, this work would not be possible.

First of all, I would like to thank my supervisor, Jens Jørgen Gaardhøje, for accepting me in the HEHI group and giving me an unique opportunity, for always trying to put me on the right track, for reading my thesis carefully and giving me advices. I would never have progressed with my analysis without the help from Valentina Zaccolo, who guided me since the beginning. A super thanks to Christian Holm for solving the mysterious software issues, to Kristjan Guldbrandsen for the frequent discussions during the first half of my PhD, and to Ian Bearden for being supportive and always a good person to talk to in moments of need. I also thank Børge, for teaching about FMD and helping with the crazy errors during the shifts, You for the friendship and Marek for always sharing his very optimistic views and advices. I can not forget Jimmy and Axel, who were always ready to help with all sorts of technical problems, printing posters, good conversation in the kitchen and who would often offer me cake. Cake is always appreciated.

During my Ph.D. I also made incredible friends. I would like to thank my Ph.D., Master and Post-Doc friends, Freja, Katka, Christian, Vojta, Helene and Vytautas for all the beers, trips and shared experiences. An extra thanks to Meera for being a good friend, for the extra beers and long conversations about everything. To Jorge for the friendship, dinners and for making fun of me every Friday for going to Danish classes every Saturday morning. You all made this journey much more enjoyable! I also thank all my friends from Brazil, especially Marina, for their support, even at long distances.

A very important part of it was also my family. I thank my parents, Lucélia and José Rubens, who always supported me and my decisions, and my two caring brothers, Humberto and Haroldo. A thanks to the Schiøtt family for being so receptive and helpful and, finally, a very special big thanks to Emilie, who was extremely patient, loving and who helped me through the hardest times.

Preface

In 2009, I started my Bachelor's studies in Biological Physics at the State University of Sao Paulo and finished in 2012. I followed up with my Master's studies in Biophysics, in the same research field - Protein folding. After 6 years of study, I finally decided that I wanted to do research in a different area and I was fascinated by high energy physics. This whole fascination lead me to search for available Ph.D. positions in this area across Brazil and around the world and, somehow, in 2015, I started my Ph.D. at the Niels Bohr Institute in Denmark.

The initial idea was to study PbPb collisions, however, a previous student in our group started to study the multiplicity distributions in p-Pb collisions and could not finish. Since it is a very interesting system to study and there was a lot of work to do, the direction of my project changed towards the study of charged-particle multiplicity distributions in p-Pb collisions. This new topic was very different from my previous academic experiences and I had to learn a complete new world. Although some words were very similar - protein unfolding, data unfolding - they surely had nothing to do with each other. Three years is hardly enough to learn all we have to learn about the fascinating world of high energy heavy-ion physics, however, throughout my Ph.D. I learned so much and the results of all this effort are summarized here.

This thesis starts with basic concepts in heavy-Ion physics, shows previous results, the motivation to study p-Pb collisions and why it became so interesting recently. After the theoretical background, it is defined what are charged-particle multiplicity distributions, when its studies started and it is shown the main theoretical contributions and past results. It follows up with the experimental setup, the main detectors used in this analysis, and the relevant analysis software. Finally, the results are presented, discussed and summed up.

I wish you a happy reading.

Contents

Introduction **23**

1 Particle World: The Standard Model **25**

1.1 The theory of strong interactions: Quantum Chromodynamics **28**

2 Variables and Concepts **31**

2.1 Coordinates **31**

2.2 Kinematics **31**

3 High-Energy Relativistic Collisions **35**

3.1 Evolution of heavy-ion collisions **36**

3.1.1 Initial Stage: Color Glass Condensate 37

3.1.2 After-Collision state: Glasma 38

3.1.3 Quark-Gluon Plasma 40

4 Evidence of QGP: Previous results from Heavy-ions Collisions **43**

4.1 Highlights from Heavy-ion Collisions **43**

4.1.1 Jet Quenching 43

4.1.2 Flow 45

4.2 Small collision systems: Signs of collectivity **47**

5 Charged-particle Multiplicity Distributions 51

5.1 KNO Scaling 52

5.2 Negative Binomial Distribution 53

5.3 Previous Measurements 55

5.4 Models 57

5.4.1 EPOS LHC . . . 58

5.4.2 DPMJET . . . 59

5.4.3 Double NBDs . . . 59

6 Experiment 61

6.1 The Large Hadron Collider 61

6.2 ALICE Experiment 62

6.3 Central Barrel Detectors 63

6.3.1 Inner Tracking System - ITS . . . 63

6.3.2 Time Projection Chamber - TPC . . . 64

6.3.3 Transition Radiation Detector - TRD . . . 65

6.3.4 Time of Flight - TOF . . . 65

6.3.5 High Momentum Particle Identification Detector- HMPID . . . 66

6.3.6 Electromagnetic Calorimeter - EMCal . . . 66

6.3.7 Photon Spectrometer - PHOS . . . 66

6.4 Forward Detectors 67

6.4.1 Photon Multiplicity Detector - PMD . . . 67

6.4.2 V0 detector . . . 67

6.4.3 T0 detector . . . 68

6.4.4 Zero Degree Calorimeter - ZDC . . . 68

6.5 Muon Spectrometer 68

7 The FMD and SPD 71

7.1 Silicon Detectors 71

7.1.1 Energy Loss in Silicon . . . 71

7.1.2 Energy Straggling . . . 74

7.2 Forward Multiplicity Detector 76

7.2.1 Electronics . . . 76

7.3 SPD 78

8 Data Processing . . . 81

8.1 ROOT, AliRoot and AliPhysics 81

8.2 Monte Carlo Simulations 81

8.2.1 Generators . . . 82

8.2.2 Digitisation . . . 82

8.3 Reconstruction 83

8.4 GRID and AliEn 85

9 Analysis . . . 87

9.1 Event Selection 88

9.1.1 Trigger Selection . . . 88

9.1.2 Vertex Selection . . . 89

9.1.3 Centrality Selection . . . 90

9.2 Particle Counting 92

9.2.1 Energy Distributions . . . 92

9.2.2 Hit Sharing . . . 94

9.2.3 Particle Counting Methods . . . 95

9.3 Unfolding 98

9.3.1 Bayesian Unfolding . . . 98

9.3.2 $\chi^2 - \mathbf{Minimization}$ . . . 101

9.3.3 Unfolding performance . . . 102

9.3.4 Problems in the Unfolding procedure . . . 104

9.4 Trigger Correction 104

10 Systematic Uncertainties . . . 107

10.1 Minimum Bias Data 107

10.1.1 Sharing cuts in the raw distributions . . . 107

10.1.2 Vertex . . . 108

10.1.3 Material Budget . . . 109

10.2 Total Systematic Error 111

11 Results . . . 115

11.1 Minimum Bias Multiplicity Distributions in p-Pb Collisions 115

11.1.1 Comparison to Models . . . 116

11.2 Centrality Selection in p-Pb Collisions 121

11.2.1 Centrality Selection in Data and MC comparison . . . 121

12 Summary **127**

Appendices **137**

A Monte Carlo Studies in Centrality Selection **139**

List of Figures

1.1 Standard Model of elementary particles. Figure taken from[82]. 26
1.2 Quark-quark potential. Figure taken from[16]. 29

2.1 ALICE experiment set up for run II with a zoomed view of the innermost detectors on the top right corner with the coordinate system. Figured modified from [84]. 32

3.1 Representation of a Heavy-ion collision as a function of time. Figure taken from [58]. 36
3.2 Results from ZEUS experiment at HERA showing the proton PDF of quarks (u and d), gluons (g) and sea quarks (S). Figure taken from [58]. . 37
3.3 Representation of the gluon cascade produced by high-energy evolution. Figure taken from [58]. 38
3.4 Schematic representation of parton evolution in QCD, showing the scenarios described by BFKL and DGLAP equations. Each colored circle represents a parton with transversa area $\Delta x_\perp \sim 1/Q^2$ and $k_z = xP$. Figure taken from [58]. 39
3.5 Representation of the saturation disk. Figure taken from [58]. 39
3.6 Left: Field configurations in the nuclues prior to the collision. Right: Longitudinal color flux tubes which develop in between the two nuclei right after the collision. Figure taken from [58]. 40
3.7 QCD phase diagram. Figure taken from [76]. 41

4.1 Two-particle azimuthal distributions for a) minimum bias d+Au collisions and b) central Au+Au and minimum bias pp collisions. Figure taken from [9]. 44
4.2 Unbalanced dijet in PbPb at $\sqrt{S_{NN}} = 2.76$ TeV observed by CMS. Figure taken from[32]. 44
4.3 A non-central Pb-Pb collisions producing an almond shaped interaction. Figure taken from [81]. 45
4.4 a) Anisotropic flow v_n as a function of event centrality including hydrodynamic calculations. Figure taken from [88]. 46

4.5 Results from CMS on 2D two-particle correlation functions for 7 TeV pp. (a) Selection of minimum bias events with $p_t > 0.1$ GeV/c, (b) Selection of minimum bias events with $1 < p_t < 3$ GeV/c, (c) Selection of high multiplicity events with $p_t > 0.1$ GeV/c and (d) Selection of high multiplicity events with $1 < p_t < 3$ GeV/c. Figure taken from [61]. . 47
4.6 Two-particle correlation for pPb collisions at 5.02 TeV for pairs of charged particles with $1 <$ pT < 3 GeV/c. On the figure (a) the results for low-multiplicity events are shown and on figure (b) for a high-multiplicity selection. Figure taken from [33]. 48
4.7 The nuclear modification factor of charged particles as a function of transverse momentum in p-Pb collisions at 5.02 TeV. Figure taken from [4]. 49

5.1 Left: Negative binomial distributions fitted to the true multiplicity at 200, 546 and 900 GeV. Right: Sum of two NBDs fitted to the true distribution at 900 GeV. Figure taken from [50]. 53
5.2 Examples of negative binomial distributions with different parameter values. 54
5.3 Charged particle multiplicity distributions obtained at ISR at four energies. The left plot show the normalized multiplicity distribution and the right plot shows the KNO scaling. Data from ISR[25] and figure from [70]. 55
5.4 Charged particle multiplicity distributions obtained at UA1 at four energies. The left plot show the normalized multiplicity distribution and the right plot shows the KNO scaling. Data from UA1[37, 38] and figure from [70]. 56
5.5 Charged particle multiplicity distributions obtained at UA1 at four energies. The left plot show the normalized multiplicity distribution and the right plot shows the KNO scaling. Data from UA1[37, 38] and figure from [70]. 56
5.6 Charged particle multiplicity distributions obtained at CDF in four pseudorapidity ranges at a)$\sqrt{s} = 1800$ GeV and b)$\sqrt{s} = 630$ GeV [77]. . . 57
5.7 Charged-particle multiplicity distribution published by the CMS experiment[60]. Left: Distributions in 5 pseudorapidity ranges at 0.9 TeV. Right: Distributions in KNO variables at 0.9 and 7.0 TeV in the range $|\eta| < 2.4$. 58
5.8 Charged-particle multiplicity distributions published by the ALICE experiment[5]. Left: Distributions at 0.9 TeV. Right: Distributions at 7 TeV. 59
5.9 Comparison of the multiplicity distributions between ALICE and CMS results for pp collision at 0.9 TeV [5]. 60
5.10 Elementary scattering in EPOS model. Figure from [74]. 60

6.1 The CERN accelerator complex. Credit: CERN. 63
6.2 An overview of ALICE's central barrel detectors. Figure from [57]. 64

6.3 Schematic figure of the forward detectors. Figure from [57]. 67
6.4 The main components of the Muon Spectrometer. Figure taken from [11]. 68

7.1 Stopping power for positive muons in copper as a function of $\beta\gamma$. Figure taken from [71]. 73
7.2 dE/dx for different particles as a function of energy. Figure taken from [64].73
7.3 (a) Energy loss distribution in thin absorbers. (b) Calculations of the energy loss distribution by Landau, Syman, and Vavilov compared to data. Figure taken from[57]. 75
7.4 Overview of the FMD system. Figure taken from [70]. 77
7.5 The Alice FMD electronics. Figure taken from [57]. 77
7.6 Inner (left) and outer (right) sensor types. Each sensor is divided azimuthally into 2 sectors. The dark gray area along the edges are used for bonding pads. Figure taken from[57]. 78
7.7 Arrangement of inner (left) and outer (right) sensor into rings. Figure taken from[57]. 78
7.8 The position of the SPD in the ITS. Figure taken from[36]. 79

8.1 Reconstruction chain. Figure from [57]. 83
8.2 Left: The primary-vertex finding method. The vertex is found in the point where there are more intersections. Right: To define a tracklet, $\Delta\phi$ between the line defined by the vertex and the first layer cluster and the line defined by the second layer cluster should be under a threshold. Figure taken from [53]. 84
8.3 Schematic drawingf of the analysis train structure. Figure taken from [12]. 85
8.4 This figure shows the AliEn centers through the world. The circles correspond to the current activities of each center. Green means that there are jobs running on that site, yellow means server problems and red means service down or no jobs running. Figure taken from [31]. 86

9.1 The measured z-vertex distribution for p-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV88
9.2 Acceptance of FMD and SPD: η-acceptance as a function of v_z vertex. The red areas denote the acceptance, and the blue areas denote no acceptance. Left: Selection in the range $-10 < v_z < +10$cm. Right: Cut to $-4 < v_z < +4$ cm, minimizing the blue areas. 89
9.3 Measured xy-vertex distribution of $\sqrt{s_{NN}} = 5.02$ TeV. 90
9.4 Representation of the impact parameter in a A-A collision[80]. 90
9.5 Distribution of V0A multiplicity, Glauber fits and centrality classes[6]. . . 91
9.6 Distribution of the neutron energy measured in the Pb-remnant side and the centrality classes obtained by the method described in[6]. 92
9.7 Energy loss distributions. The light colors represent the distributions obtained in the reconstruction, while the dark portion is obtained after the hit sharing algorithm is applied. 93

9.8 Signals from a single strip from FMD. The blue distribution includes all the signals detected in that strip, whereas the red one shows the distribution after removing the shared hits, a large set of signals that have energies between the pedestal peak and the maximum of the Landau distribution [36]. 93
9.9 Hit sharing schematic. Figure taken from [57]. 94
9.10 The diagram shows the hit merging algorithm.The subscript c indicates current strip and the subscript n indicates next strip. E_{low} and E_{high} are defined in the text. 95
9.11 Correlation between methods. The energy fits are represented along the x axis and the poisson method along the y axis. The solid red line is the regression, and the dashed x=y line is plotted to guide the eye. . . . 97
9.12 Response matrix obtained through a general purpose Monte Carlo production for p-Pb collisions at 5.02 TeV in the range $-3.4 < \eta < +5.1$. 99
9.13 Distributions obtained using different number of iterations. 100
9.14 Closure test using 100 iterations, Left: Unfolding of MC reconstructed in a narrow η range. Right: Unfolding of MC reconstructed in a wide η range. 101
9.15 Closure test using 10 iterations, Left: Unfolding of MC reconstructed in a narrow η range. Right: Unfolding of MC reconstructed in a wide η range. 101
9.16 χ^2 - minimization unfolding closure test performed without regularization 103
9.17 Q values for χ^2-minimization method using different weights and regularization functions for each region, in which the top plot corresponds to the peak, the middle plot corresponds to the slope and the bottom plot corresponds to the low statistics. The blue line represents the best value for the Bayesian method. 104
9.18 Trigger efficiency as a function of N_{ch}. Left: Efficiency calculated in the range $|\eta| < 3.4$. Right: Efficiency calculated in the range $|\eta| < 0.5$. . . 105

10.1 Multiplicity distributions obtained by lowering(blue) and increasing(red) the sharing cut. Left: Distributions in the range $|\eta| < 5.0$. Right: Distributions in the range $|\eta| < 3.4$. 108
10.2 Fitted Multiplicity distribution obtained by lowering(blue) and increasing(red) the sharing cut. Left: Distributions in the range $|\eta| <$ 5.0. Right: Distributions in the range $|\eta| < 3.4$. 108
10.3 Fitted multiplicity distributions using different vertex selection (top) and the ratio of the distributions obtained in different vertex ranges to the distribution obtained in the standard vertex range (-4cm < vx < +4 cm). Left: $|\eta| < 5.0$. Right: Ratio for $|\eta| < 3.4$. 109
10.4 Fitted response matrices. Top: Response matrix obtained from a flat Monte Carlo production anchored to a pp collision run at 900GeV. Bottom: Response matrix obtained from a Monte Carlo production anchored to a pPb collision run at 5.02TeV. 110

10.5 The top panels show the distribution obtained unfolding pp raw distributions using a response matrix generated from a regular flat Monte Carlo together with the distributions obtained unfolding the raw distributions using a response matrix generated from a flat Monte Carlo increasing the material budget in 14%. The bottom panels show the ratio of the distributions obtained with increased material budget to the regular distributions. Top: Narrow η range. Bottom: Wide η range. . 112
10.6 Multiplicity distributions calculated for pp collisions at 900 GeV with nominal material budget and material increased by 14%(top panel) and ratio of the multiplicity distributions (bottom panel). Top-left: Fit to the ratio in $\eta < 0.5$. Top-right: Fits to the ratio in $\eta < 3.4$. Bottom-left: Extrapolation to the fit in $\eta < 0.5$. Bottom-right: Extrapolation to the fit in $\eta < 3.4$. 113
10.7 Total systematic uncertainties for $-3.4 < \eta < +5.1$, $|\eta| < 3.4$, $|\eta| < 0.5$. 114

11.1 Unfolded and trigger-corrected charged-particle multiplicity distributions in 8 different η ranges for p-Pb collisions at $\sqrt{S_{NN}} = 5.02$ TeV. . . . 116
11.2 Unfolded and trigger-corrected charged-particle multiplicity distributions in 8 different η ranges including systematic uncertainties from Figure 11.1 for MB p-Pb collisions at $\sqrt{S_{NN}} = 5.02$ TeV. 117
11.3 Unfolded and trigger-corrected charged-particle multiplicity distributions in $|\eta| < 3.4$ for MB p-Pb collisions at $\sqrt{S_{NN}} = 5.02$ TeV. The threshold where the distributions are cut is defined accordingly to the errors. 118
11.4 Charged-particle multiplicity distribution accordingly to EPOS LHC and DPMJET in $|\eta| < 0.5$ (left) and $-3.4 < \eta < +5.1$ (right) for p-Pb collisions at $\sqrt{S_{NN}} = 5.02$ TeV 118
11.5 Comparison of the results with the calculated distributions using two different Monte Carlo models - EPOS LHC and DPMJET - for MB p-Pb collisions at $\sqrt{S_{NN}} = 5.02$ TeV. 119
11.6 Closer view of the comparison data-MC models (EPOS LHC and DPMJET). The plot on the left shows closer the distributions in the three widest eta ranges and the plot on the left shows the distributions in the three narrowest eta ranges. 120
11.7 Charged-particle multiplicity distributions for p-Pb at $\sqrt{S_{NN}}$= 5.02 TeV compared to distributions from the IP-Glasma model. The IP-Glasma points are taken from [87]. 120
11.8 Comparison of charged-particle multiplicity distributions from p-Pb data in different centrality classes (using the estimator CL1) with the distributions produced using Monte Carlo simulations based on EPOS LHC (Top) and DPJMET (Bottom) in the pseudorapidity range $|\eta| < 0.5$. The bottom panels show the ratio of the model to the data. 122

11.9 Comparison of charged-particle multiplicity distributions from p-Pb data in different centrality classes (using the estimator CL1) with the distributions produced using Monte Carlo simulations based on EPOS LHC (Top) and DPJMET (Bottom) in the pseudorapidity range $|\eta| <$ 3.4. The bottom panels show the ratio of the model to the data. 123
11.10 Comparison of charged-particle multiplicity distributions from p-Pb data in two different centrality classes (using the estimator CL1) , 00-05%(left) and 80-100%(right),with the distributions produced using Monte Carlo simulations based on EPOS LHC (Top) and DPJMET (Bottom) in the pseudorapidity range $|\eta| < 0.5$. The bottom panels show the ratio of the model to the data. 124
11.11 Comparison of charged-particle multiplicity distributions from p-Pb data in two different centrality classes (using the estimator CL1) , 00-05%(left) and 80-100%(right),with the distributions produced using Monte Carlo simulations based on EPOS LHC (Top) and DPJMET (Bottom) in the pseudorapidity range $|\eta| < 3.4$. The bottom panels show the ratio of the model to the data. 124
11.12 Charged-particle multiplicity distributions from pPb collisions data at $\sqrt{S_{NN}} = 5.02$ TeV in two different centrality classes (using the estimator ZNA), 00-05%(left) and 80-100%(right). 125
11.13 Charged-particle multiplicity distributions from pPb collisions data at $\sqrt{S_{NN}} = 5.02$ TeV in different centrality classes (using the estimator ZNA) in the ranges $|\eta| < 0.5$(left) and $|\eta| < 3.4$(right). 125

A.1 Comparison between Charged-particle multiplicity distributions according to EPOS LHC and DPMJET divided in centrality classes using the estimators V0A (top plots) and CL1 (bottom plots) in $|\eta| < 3.4$. . . . 140
A.2 Comparison between Charged-particle multiplicity distributions according to EPOS LHC and DPMJET divided in centrality classes using the estimators V0A (top plots) and CL1 (bottom plots) in $|\eta| < 1.0$. . . . 141
A.3 Comparison of Charged-particle multiplicity distributions according to EPOS LHC and DPMJET divided in centrality classes using the estimators V0A (top plots) and CL1 (bottom plots) in central collisions (00-05%). 142
A.4 Comparison of Charged-particle multiplicity distributions according to EPOS LHC and DPMJET divided in centrality classes using the estimators V0A (top plots) and CL1 (bottom plots) in peripheral collisions (80-100%). 143

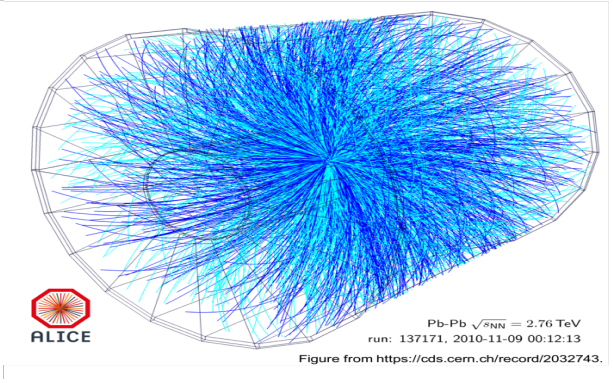


Figure from https://cds.cern.ch/record/2032743.

Introduction

The major goal of heavy-ion experiments is to understand and characterize the hot and dense medium made of deconfined quarks and gluons, called quark-gluon plasma (QPG), produced in heavy-ion collisions. For a long time, it was thought that only large collision systems, such as nucleus-nucleus collisions, could produce the QGP. However, recent investigations suggest that there is collectivity, a feature observed in nucleus-nucleus collisions, also in smaller systems, such as proton-proton and proton-nucleus collisions. Therefore, to properly understand the QPG, further investigation of small systems is needed.

The measurement of charged-particle multiplicity probability distributions, or $P(N_{ch})$ as a function of N_{ch}, provide insights into the mechanism underlying particle production in various collision systems, therefore it can contribute to the characterization of small systems.

In heavy-ion experiments, it is not possible to measure directly the properties of the medium produced in each collision, only the final products. The multiplicity distribution works like a picture taken of the system by the end of the experiment, and, just like a photograph, one can obtain information about the initial state of the system by looking at the final state. In the case of photographs, we are talking about the photons source and in the case of charged-particle multiplicity distributions, we are talking about the projectiles participating in the collision. Measuring $P(N_{ch})$ can be used to evaluate the correct description of the initial state of hadrons and nuclei at high energies.

The measurement of the $P(N_{ch})$ for p-Pb collisions at LHC energies hasn't yet been done, thus being of great interest since it corresponds to an unusual configuration: a single proton "drilling" a hole through the lead nucleus. However, to understand these measurements, we have to start with elementary concepts. The introductory chapters of this thesis provide the reader with the basics of high-energy physics and relativistic heavy-ion collisions. The knowledge will be built up with essential blocks, starting with elementary particles, portraying the standard model, going through the theory behind Relativistic collisions, until the main focus of this work - the charged-particle multiplicity distribution - is introduced. The experimental setup and the analysis methods are explained in details prior to the results, which are presented in the last chapter followed by the conclusion and discussion.


Figure modified from astrorhysy.blogspot.com

Chapter 1

Particle World: The Standard Model

Humanity has always wondered about what the world - or even the Universe - is made of. Around the 5th century BCE, Democritus attempted to explain the composition of the natural world through the atomist theory, in which all matter is composed by indivisible atoms [19]. The efforts continued in the 19th and 20th centuries with scientists such as J.J. Thomson[85] and Bohr [23], among others, and it was discovered that the atom is made of subatomic particles. In the early 1970s, the theory of fundamental particles, the building blocks of matter, and their interactions was put together and it was called the Standard Model, term coined by Pais and Treiman in 1975[28]. Since then, the Standard Model theory has been improved and tested through many experiments, and it is settled as a well tested and strong theory. However, even a strong theory contains flaws or is not complete. Today, we know the model is incomplete since it does not provide any explanations for dark matter and dark energy[86], for why we observe only left-handed neutrinos[67] and it also does not include gravity[66]. Nevertheless, the Standard Model is the currently the best picture of the building blocks of matter we have so far.

Charge and spin are normally used to characterize a particle, in addition to other quantum numbers. The spin is the intrisinc angular momentum of the particle. It can assume the values $0, \frac{1}{2}, 1, \frac{3}{2}$,..., and it is conserved for each particle. When a particle has a half-integer spin, it is called a *fermion*, and the particles with integer spin are called *bosons*. In the standard model, there are 3 categories of particles: two with spin-$\frac{1}{2}$, *leptons* and *quarks*, and one with integer spins, *bosons*.

The Standard Model's components are presented in the following section.

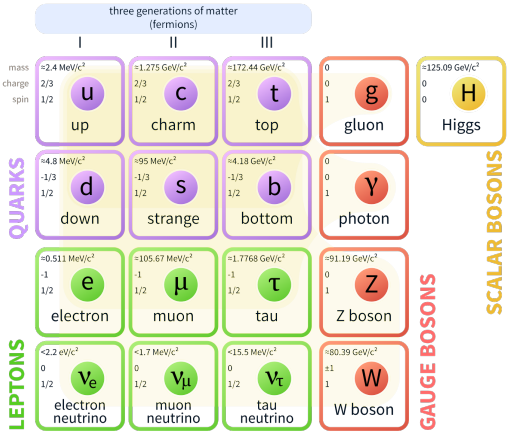


Figure 1.1: Standard Model of elementary particles. Figure taken from[82].

Fermions

Quarks

Quarks are elementary particles of spin $\frac{1}{2}$. They are classified into 6 different flavors and 3 generations, each generation consisting of a quark with charge $+\frac{2}{3}e$ and a quark with charge $-\frac{2}{3}e$. The quarks have an additional and important feature, the color charge. Even though the color charges do not have any relation to our macroscopic perception of colors, they are classified into 3 different color categories, namely red, green, and blue. Every quark (q) has a corresponding antiparticle, anti-quarks ($\bar{q}$), which carry anti-color charge, named anti-red, anti-green, and anti-blue. Just as in the analogy to the RGB scheme, a combination of 3 quarks with different color charges results in a color neutral state.

Quarks have never been observerd as free particles. They interact via the strong force to form bound states called hadrons. There are two families of hadrons: baryons, composed of three quarks, and mesons, composed of one quark and one antiquark. The dominantly occurring quarks in nature today are the up (u) and down (d), which are the constituents of the well-known baryons proton (uud) and neutron (ddu).

Leptons

The lepton that everybody is familiar with is the electron, which is one of the six leptons in the standard model. The electron is an elementary particle bound in the atoms via the electromagnetic interaction. Two other leptons, tau and muon, have the same properties as the electron, apart from their masses. Both muon and tau are allowed to decay into other particles, while the electron, the lightest of the 3, is stable.

The next three leptons are the neutrinos. They are electrically neutral and, as far as we know, only participate in weak interactions. There are three different types of neutrinos, each of them linked to a specific charged lepton. Their masses are extremely small, $<$ 1eV, and they hardly interact with matter, making it almost imperceptible.

In total, there are six leptons and six antileptons.

Bosons

As mentioned previously, bosons have integer spins, therefore they do not follow the Pauli exclusion principle; there is no restriction on the number of bosons which can occupy the same quantum state. In the Standard Model, there are four *gauge bosons*, which are carriers of fundamental forces. A fundamental force is not derived from any other force, and four forces are known to belong into this classification: gravity, electromagnetic, the weak force, and the strong force.

Electroweak bosons: W^+, W^-, Z^0

Discovered in 1983, they are the three very massive vector bosons which mediate the weak force, with masses $m_W = 80.6 GeV/c^2$ and $m_Z = 91.2 GeV/c^2$. Emitting or absorbing a W boson can change the particle's charge and spin, for example a down quark becomes an up quark. The Z^0 boson can only act on the spin, momentum, and energy of a particle, not changing its flavor during the interaction.

Photons

Photons are the force carriers for the electromagnetic interaction. They are massless and move at the speed of light in vacuum. The theory which describes the electromagnetic interaction is the Quantum Electrodynamics (QED). It was observed that the electromagnetic force has a relation to the weak force, therefore the theories were "unified" into the Electroweak force.

Higgs

In 1967, Abdus Salam and Steven Weinberg incorporated the unified theory of electromagnetic and weak forces into the Standard Model. The theory predicted a new type of electrically neutral and unstable particle, now known as the Higgs boson . In 2012, the awaited Higgs boson was observed by the ATLAS and CMS experiments in the Large Hadron Collider (LHC)[35]. It has a lifetime of approximately $3.10^{-25} s$, mass of $m_H = 125.18 GeV/c^2$ and decay mostly to jets of hadrons[14]. Through the Brout-Englert-Higgs mechanism, the bosons W and Z gain mass during the interaction with the "Higgs field", the Higgs boson being the manifestation of the field. The Higgs boson is found to be a scalar boson, since its spin and parity is mainly 0^+, in contrast to the other bosons of the Standard Model, which have spin 1[1].

Gluons

Gluons are the particles which bind quarks together. In other words, they are the mediators of the strong force between quarks of different color charge. In addition to mediate the strong force, the gluons can also participate in the interaction themselves, being also carriers of color charge. That is in contrast to QED, where photons have no electrical charge and do not interact with other photons. Unlike the quarks, gluons are considered to be "bi-colored", carrying a unit of color and a unit of anti-color, resulting in 9 possible effective states. However, their are not the actual number of gluons - the state of a gluon is seen as a combination of different possible states with a certain probability. In this case, there are only 8 different types of gluons because the 9^{th} linear combination is color neutral.

The theory that describes the strong interactions involving quarks and gluons is called Quantum Chromodynamics (QCD), which is defined in the next section.

1.1 The theory of strong interactions: Quantum Chromodynamics

Quantum Chromodynamics is the quantum field theory which describes the strong interaction. The only particles who can feel the strong interaction are the ones possessing color charge - quarks and gluons.

In QCD, the force carriers, gluons, also possess color charge and can interact with themselves. A consequence of this property is the first key concept in QCD, called *confinement*. Quarks and gluons are bundled together in order to form hadrons and there is no experimental evidence that they can be found free in nature [73].

Quarks interact very weakly at short distances, a feature described by the second key concept in QCD, the *asymptotic freedom*. The screening effect in QED is very well known: virtual particles polarize the vacuum, attracting particles with opposite charges and repelling the ones with same charge. The net effect cancels out part of the field in a limited distance. The same effect can happen in QCD, however since gluons also have color charge, the polarization of gluons in the vacuum does not cancel out the field, but increases it. This effect is called color *antiscreening*, which makes the quark-quark interaction increase with distance.

The quark-quark potential can be expressed as following [15]:

$$V = -\frac{4}{3}\frac{\alpha_s(r)\hbar.c}{r} + k.r \tag{1.1}$$

The strong interaction gets weaker at short distances, but when $r \to 0$, $V \to -\infty$. On the other hand, as the distance between quarks increases, the second term of the potential, k.r, is dominant and the interaction is stronger. The strong coupling constant $\alpha_s(r)$ is actually not constant, but depends on the separation r between the quarks, becoming smaller as r increases. The quarks are held together by gluons, and when the distance between the quarks increases, the gluon field stores more and more energy until it reaches a limit where there is enough energy to create new

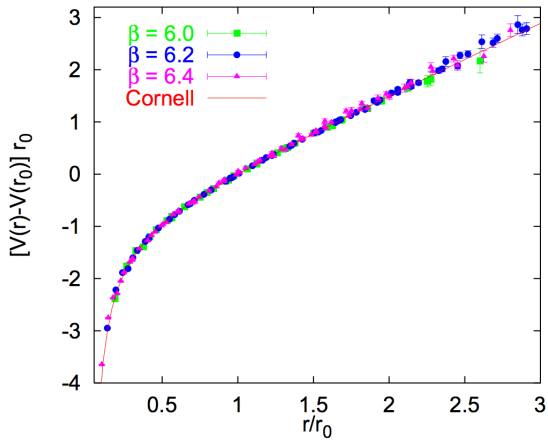
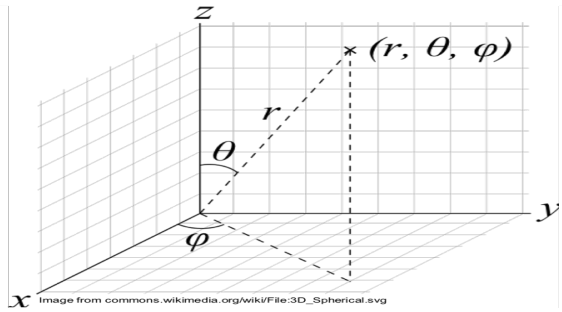


Figure 1.2: Quark-quark potential. Figure taken from[16].

quark- antiquark pairs. The second term in the equation 1.1 represents the field lines between two quarks when they are pulled apart, which are called *flux tubes*, and the constant k is the *string tension*.

There are basically two main different approaches to make calculations in QCD. In the regime where the coupling constant is small, perturbative theory methods work well, thus having perturbative QCD (pQCD). On the other hand, when the coupling constant is fairly high, pertubative methods fail and a method formulated on a discrete Euclidean space time grid[55], called lattice QCD (lQCD), is used.



Chapter 2

Variables and Concepts

To understand heavy-ion experiments and theory, it is important to know the basic variables which are used in this context. In this chapter, the important variables and concepts are presented.

2.1 Coordinates

In a collider, particles are accelerated in opposite directions and made to collide in specific points. In the case of the LHC, the particles collide in the center of the four major detectors: ALICE, ATLAS, CMS and LHCb. The exact point in space where the particles collide is called the *interaction point* (IP). The IP is an important definition since the detectors need the information of the measured IP, which is also known as the primary vertex, in order to reconstruct tracks of charged particles, for instance.

In order the describe the physics before and after the collision, a coordinate system has to be established. The coordinate system adopted by ALICE is a right-handed orthogonal Cartesian system with origin x,y,z = 0 at the IP. The system is depicted in Figure 2.1.

The x axis is defined perpendicular to the beam direction and points towards the accelerator center. The y axis is also perpendicular to the beam direction, but points upwards. By convention, the beam of particles travels along the z axis. The polar angle is given by θ, and the azimuthal angle is given by ϕ.

2.2 Kinematics

Here, the variables necessary to describe relativistic collisions are defined.

Firstly, we can define the four-momentum of a particle of rest mass m_0 traveling

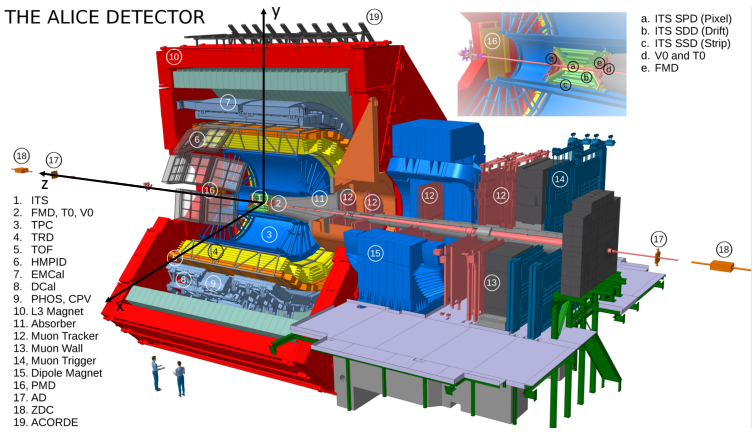


Figure 2.1: ALICE experiment set up for run II with a zoomed view of the innermost detectors on the top right corner with the coordinate system. Figured modified from [84].

with momentum **p** and energy E as

$$P = (E, \overrightarrow{p}) = (E, p_x, p_y, p_z), \tag{2.1}$$

here, the natural constants are defined as c = $\hbar$ = 1, where c is the speed of light and $\hbar$ is Planck's constant. Using the ALICE coordinate system: z is the beam axis and x-y is the transverse plane. The longitudinal momenta of the particles is then given by the z component, p_z, whereas the transverse momentum, which is orthogonal to the beam is given by

$$p_t = \sqrt{p_x^2 + p_y^2} \tag{2.2}$$

and the transverse mass

$$m_t = \sqrt{m^2 + p_t^2} \tag{2.3}$$

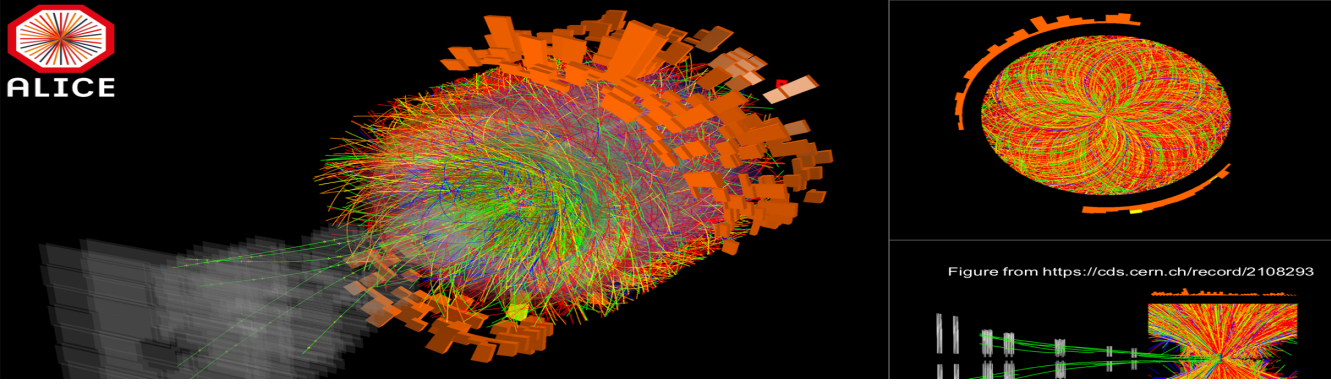
The latter 2 variables are Lorentz invariant, which makes them very convenient to use in relativistic collisions experiments. The rapidity of the particles is also a good variable since it is dimensionless and follows a simple addition in the Lorentz transformation between frames of reference, $y = y_1 + y_2$. It is expressed as

$$y = \frac{1}{2} ln \frac{E + p_z}{E - p_z} = \frac{1}{2} ln \left(\frac{1 + \beta \cos\theta}{1 - \beta \cos\theta} \right), \tag{2.4}$$

where E = p^2+m^2. Even though the rapidity is a very important variable in heavy-ion physics, the calculation of it requires the knowledge of the mass of the particle. For non-identified particles, another variable is useful, which is called pseudorapidity and defined as

$$\eta = -ln(tan\frac{\theta}{2}) \tag{2.5}$$

Here, θ is the angle between the direction of the emitted particle and the z axis.



Chapter 3

High-Energy Relativistic Collisions

The physics at subatomic scales is normally explored by the use of colliders. Almost our whole knowledge and understanding of this subject was obtained analyzing data from high energy collisions between $e^+e^-, p\bar{p}$, pp or heavy nuclei. Throughout the years, different projectiles were used in these collisions, even though the main experimental aparatus is essentially the same since the 1960's. The basic idea is to use a particle accelerator in order to produce two high energy beams of particles, which collide heads-on at a point in space surrounded by detectors designed to detect the outcome of the collisions [72]. This same procedure is used at the LHC, and the details of it are described in section 6.1. In case of fixed target experiments, only one beam of particles is accelerated and shot at a stationary target.

The colliding systems in the LHC program are described in this chapter. First, we have **pp** collisions, which are collisions between protons. This system is the main focus at LHC in general, since the goal of the LHC was to discover the Higgs boson and to find super symmetric (SUSY) particles or other new particles beyond the Standard Model. This is, therefore, an important system to general purpose detectors like ATLAS and CMS, which, typically, can deal with very high momentum particles . ALICE, on the other hand, is a heavy-ion detector, meaning that it was designed to study the strongly interacting matter produced in heavy-ion collisions, consisting of many low momentum particles per event. Proton-proton collisions under ALICE's eyes work to search for new physics, and as a baseline for PbPb collisions.

PbPb collisions are collisions between two lead nuclei, and the main focus of ALICE. In each PbPb collisions, thousands of particles are created, and a QGP is produced in central PbPb collisions where there is a higher density of particles. Therefore, this system is very important to study the main properties of the QGP.

The last collision system is the primary focus of this work. Proton-Lead collisions, **p-Pb**, are collisions between protons and lead nuclei. The heavy-ion community considered this system as an important measurement in order to provide insights into the initial state of collisions and disentangle cold nuclear matter (matter, that was thought not being the QGP) from hot nuclear matter (QGP) effects. This asymmetric

system, however, showed very interesting features, e.g. hints of collective behavior. Even though this system was initially thought as being a baseline, it is now seen as a very interesting system on its own.

3.1 Evolution of heavy-ion collisions

Heavy-ion collisions involve many stages, as shown in Figure 3.1, which can be divided into the four following different stages:

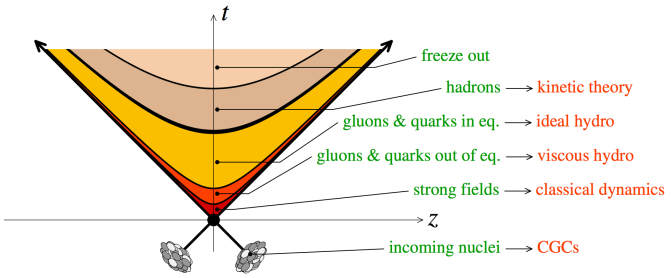


Figure 3.1: Representation of a Heavy-ion collision as a function of time. Figure taken from [58].

- **Initial stage**: The state prior to the collision where the wavefunction of the two Lorentz-contracted nuclei is dominated by a dense and weakly coupled gluonic matter. In one of the theories which explain the state prior to the collision, this matter is called *Color Glass Condensate*(CGC).
- **Non-equilibrium**: Right after the collision, processes involving high momentum transfer take place, which give origin to hard particles such as jets and heavy quarks. Around 0.2 fm/c after the collision, a dense partonic matter in non-equilibrium is formed.
- **Thermal equilibrium and hydrodynamic expansion**: The partonic density and temperature is very high forming the QPG. The partons interact with each other, which leads the medium to local thermal equilibrium since it keeps expanding.This medium behaves like a strongly coupled fluid and its expansion can be described by viscous hydrodynamics.
- **Freeze-out**: The systems keeps expanding and cooling down. When t $\sim$ 10 fm/c, hadronization starts, however the system is still dense and hot preserving the local thermal equilibrium during the expansion. This hadronic system is referred to as hadron gas, which keeps expanding until the density becomes low

enough and the hadrons stop interacting with each other. This stage is called freeze-out.

Each stage of the collision involves a different form of QCD matter. The three initial states deserve extra attention, therefore they are explained in more detail in the next sections of this chapter.

3.1.1 Initial Stage: Color Glass Condensate

Right before the collision, the incoming nuclei appear to an observer at rest in the laboratory frame like two Lorentz-contracted 'pancakes'. The 'pancakes', as we know, are made of quarks and many gluons, each of them carrying a fraction of the total longitudinal momenta of the nuclei, called Bjorken x and defined as

$$x \equiv \frac{Q^2}{2(P.q)} = \frac{Q^2}{s + Q^2 - M^2}, \tag{3.1}$$

where s is the invariant energy squared, $s = (P+q)^2$ and Q^2 the momentum transfer.

In this early stage of the collision, the parton density is very high and dominated by gluons. A proton, for instance, is composed of partons - three valence quarks and gluons. It is possible to obtain the probability of finding a parton i carrying a momentum fraction x, which is the Parton Distribution Function (PDF) $f_i(x, Q^2)$.The HERA experiment showed in DIS data that the gluon density grow with lower x, confirming that the proton wavefunction at $x < 0.01$ is dominated by gluons (g) with some signs of sea quarks (S), as shown in Figure 3.2.

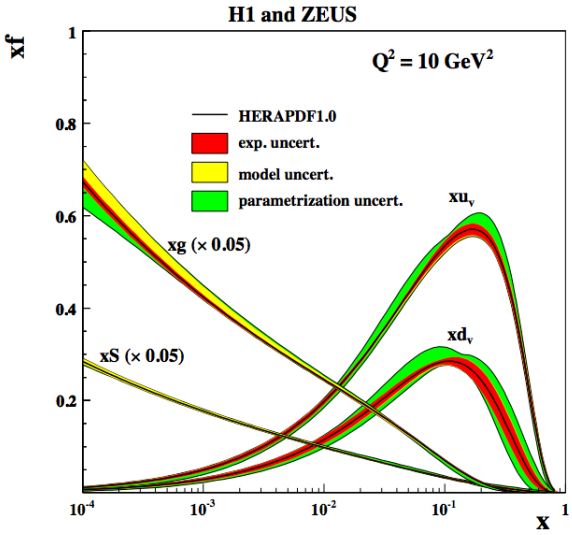


Figure 3.2: Results from ZEUS experiment at HERA showing the proton PDF of quarks (u and d), gluons (g) and sea quarks (S). Figure taken from [58].

When x decreases ,at a fixed Q^2, the emitted gluons have smaller longitudinal momentum fractions, however they occupy the same transverse area as their parent nucleons. The number of gluons therefore increases, which means that the gluonic system evolves towards increasing density. For $x \ll 1$, the valence quark emits a soft gluon, which can radiate a softer gluon that can also radiate again and so on. This process can be visualized in Figure 3.3.

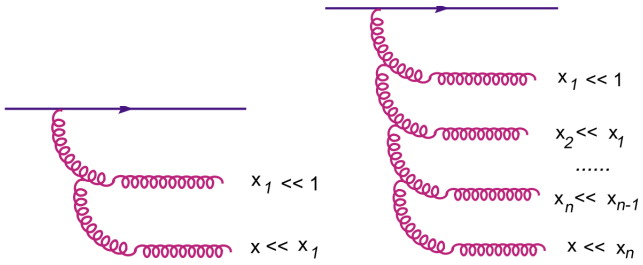


Figure 3.3: Representation of the gluon cascade produced by high-energy evolution. Figure taken from [58].

The gluon distribution at leading order can be obtained by summing the contribution from all ladders. By the end of the cascade, we can find very 'slow' gluons, with small momentum fraction, which have shorter lifetimes in comparison to the 'fast' preceding gluons. The fast gluons with x' $\gg$ x function as frozen color sources emitting gluons at scale x[58]. This results in a large color change density since the sources happen to overlap in the transverse plane, adding their color charges coherently. When the gluon occupation number n, estimated by the number of gluons with a given x times the area occupied by one gluon divided by the transverse area of the hadron[58], is approximately 1, the gluons start overlapping, however the interactions are still weak enough. Non-linear effects become important and the growth of the gluon distribution stops when the interactions are of one, and $n \approx 1/\alpha_s$. This is known as gluon saturation.

The saturation state, or the partonic form of matter made by saturated gluons is the so-called Color Glass Condensate. It is called **color** because gluons carry calor charge, **glass** since there is a separation in time scales between the slow gluons and their fast parent sources at larger x due to Lorentz time dilation, therefore, the sources appear "frozen", like previously mentioned. This kind of system works like a solid in short time scales, but like a fluid in longer time scales, similar to glass. Lastly, it is a **condensate** due to the high occupation numbers of the saturated gluons and their parents and the coherently addition of their color charges[58].

3.1.2 After-Collision state: Glasma

The CGC is the state of the incoming projectiles prior to the collision. After the moment of collision, at $\tau \sim 0.2 fm/c$, the gluons composing the CGC are liberated, forming a high density nuclear matter in a non-equilibrium state called **glasma**[63].

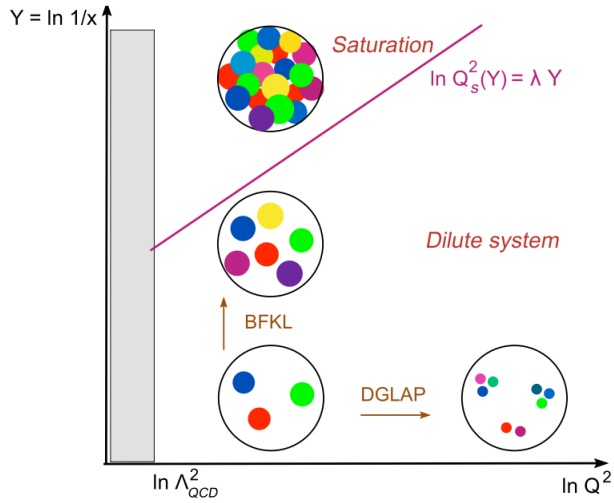


Figure 3.4: Schematic representation of parton evolution in QCD, showing the scenarios described by BFKL and DGLAP equations. Each colored circle represents a parton with transversa area $\Delta x_\perp \sim 1/Q^2$ and $k_z = xP$. Figure taken from [58].

This state is neither a CGC nor a Quark Gluon Plasma, which is going to be described in the next section, but contains certain features of both states.

The incoming hadrons can be seen as two sheets of colored glass. One can solve the classical Yang-Mills equations[1] associated with chromo-electric and magnetic fields, **E** and **B**. Due to high-energy kinematics, these fields can be described in a simple manner - the vectors **E** and **B** only have transverse components, being themselves transverse to the motion direction. On average, the fields are correlated due to high-energy evolution and saturation, and are restricted to a transverse area with radius $\sim 1/Q_s$, which is called saturation disk. A saturation disk, shown in Figure 3.5, has zero overall color charge. The gluons are also correlated due to the fact that they came from the same ancestors, belonging to the same parton cascade, therefore the correlations are extended over a rapidity interval $\Delta Y \sim 1/\alpha_s$.

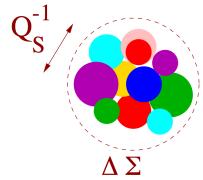


Figure 3.5: Representation of the saturation disk. Figure taken from [58].

[1]The Yang-mills theory is the core gauge theory of the unification of the electromagnetic and weak forces and it is also used to describe the strong force

After the collision, in addition to the transverse fiels, longitudinal fields also arise, extending along the collision axis like color flux tubes. These fields are strong since they carry part of the CGC energy in early moments after the collision, however the intensity of the fields decreases as soon as the system expands and the density decreases. When the magnitudes of the fields become of order one, around $t \sim 1/Q_s$, these fields can now be seen as incoherent superposition of particles, which can interact with each other since the density is still high.

This non-equilibrium state is a form of matter that in initial stages is similar to CGC and later stages similar to the Quark Gluon Plasma, therefore takes the name glasma (**G**lass + P**lasma**). The glasma should lead to the Quark Gluon Plasma, a phase of local equilibrium[58].

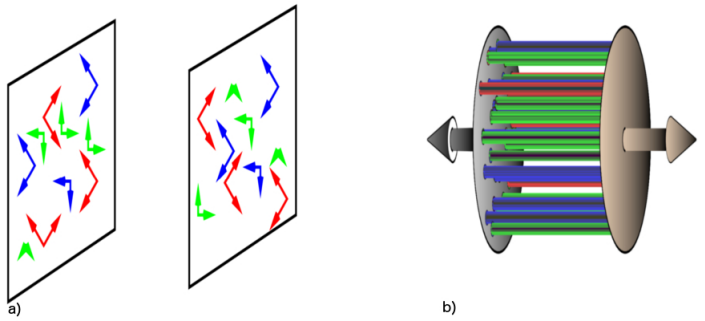


Figure 3.6: Left: Field configurations in the nuclues prior to the collision. Right: Longitudinal color flux tubes which develop in between the two nuclei right after the collision. Figure taken from [58].

3.1.3 Quark-Gluon Plasma

In the first microsecond after the Big Bang, the Universe was completely different than it is today. It is believed that the Universe was very dense and filled with deconfined quarks, anti-quarks and gluons, in a state of matter called Quark-Gluon Plasma (QGP).

The history of the QGP goes back to 1970, when Itoh mentioned the concept of quark matter in neutron stars[59]. In 1974, during the workshop on "Bev/nucleon collisions of heavy ion" at Bear Mountain, New York, heavy ion physics was suggested as a research tool that could create "abnormal dense states of nuclear matter"[17], but it was in 1975 that Collins and Perry stated that, based on asymptotic freedom, at sufficiently high temperatures or baryon density, there could exist a state of matter where quarks and gluons are free, and they called it "quark soup"[39]. This "quark soup" was called "quark-gluon plasma" for the first time in 1979[79].

Initially, it was expected that the QGP would behave like a gas, however, it was found at the Relativistic Heavy Ion Collider (RHIC), located at the Bookhaven National Laboratory, that the QGP at that energy is strongly interacting, meaning

that it behaves more like a perfect fluid[17].

The first Flow studies from Pb-Pb data at 2.76 TeV done by the ALICE experiment from the LHC agreed with that prediction, and the QGP was found to be a strongly coupled liquid, with small viscosity[34]. The matter transitions to the QGP can occur by either increasing the baryon density or the temperature. Many theoretical studies of the QCD phase diagram suggest that a critical point exists[51, 68, 83], where the first order transition line ends, and it is still a subject under investigation. An illustration of the possible QCD phase diagram is shown in the figure 3.7

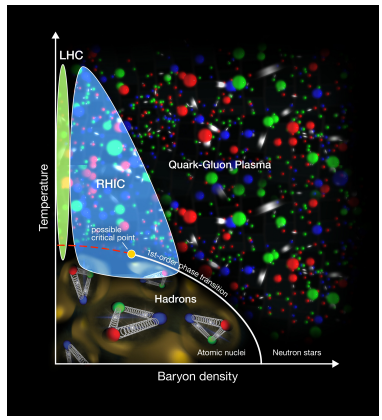
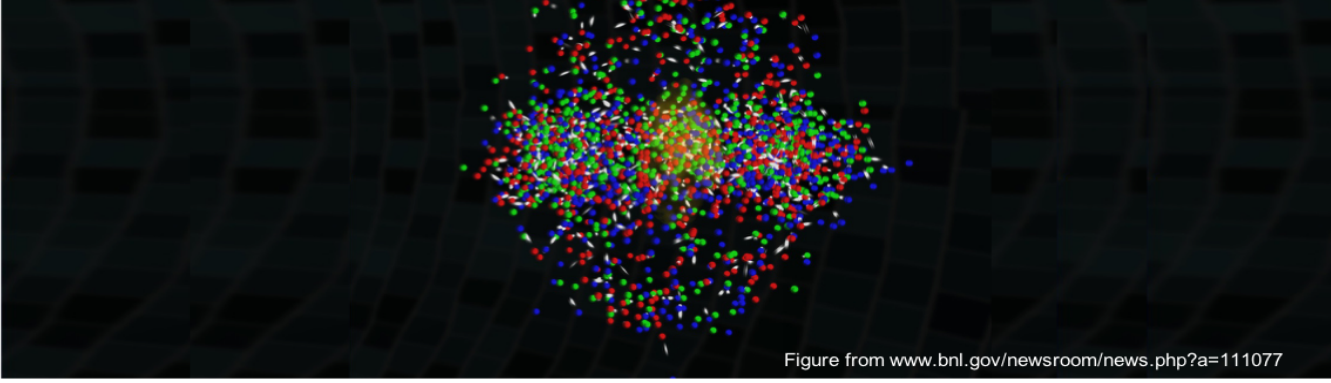


Figure 3.7: QCD phase diagram. Figure taken from [76].

The QGP created in heavy-ion collisions is very similar to the QGP from our early universe, however the size of our laboratory QGP is considerably is smaller and it has a faster dynamical evolution, which leads to faster hadronization.

The study of the QGP and its properties is one of the main goals of high energy heavy-ion collisions. Even though it's believed that the QGP from early universe and the heavy-ion QGP are not completely the same, the study and understanding of the latter might give us a broader comprehension of the early Universe and its expansion process that lead to the Universe we know today.



Chapter 4

Evidence of QGP: Previous results from Heavy-ions Collisions

4.1 Highlights from Heavy-ion Collisions

Now that the collision systems and what QGP is are defined , the answer that still has to be answered is: how do we probe the QGP? There are different observations that indicates the creation of this hot and dense medium and the two most remarkable probes of the QGP are jet quenching and flow, which the main results are presented in this chapter.

4.1.1 Jet Quenching

Hard partonic interactions, with high transverse momentum(p_t), produce two virtual back-to-back partons, which give origin to parton showers that subsequently hadronizes and can be observed through detectors as back-to-back di-jet events. If partons traverse the hot and dense medium produced in heavy-ion collisions, they can lose energy which can be observed as modification of jets, an effect called jet quenching.

Bjorken predicted that jets could be "extinguished" in pp collisions due to elastic final state energy loss of partons[21]. It turned out that the parton energy loss in the hot and dense medium was too small and the data from pp collisions at energies up to Tevatron did not show any jet extinction - or jet quenching. However, other studies suggested that the effect could be observed in heavy-ion collisions, since the induced radiative energy loss would be much larger than the elastic energy loss, thus allowing jet quenching[56].

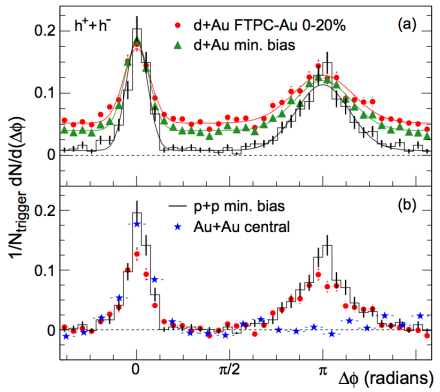


Figure 4.1: Two-particle azimuthal distributions for a) minimum bias d+Au collisions and b) central Au+Au and minimum bias pp collisions. Figure taken from [9].

Results on two-particle azimuthal distributions were obtained by STAR in 2003, as shown in Figure 4.2. It can be seen very clearly that, in nuclear collisions, there is a strong suppression in the back-to-back peak ($\Delta\phi = \pi$), whereas the near-side ($\Delta\phi = 0$) exhibit features similar to the ones observed in pp collisions. There are different possible explanations to this observation. The first suggests that it happens exactly because of the energetic partons lose energy when transversing through the medium, causing the predicted "extinction" of the jets. The peak observed would be originated by jets of partons produced near the surface. An alternative explanation to that could be used, in which initial-effect are in play, such as gluon saturation in the incoming nuclei.

Right after the first Pb-Pb run in the LHC, results on jet quenching were published by CMS [32]. They reported an imbalance in the observed dijet energy, which could be seen even in at the level of event displays during the data taking.

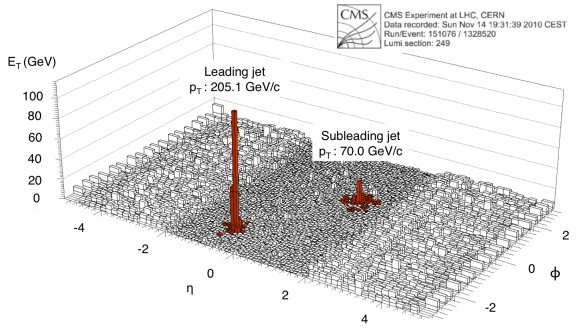


Figure 4.2: Unbalanced dijet in PbPb at $\sqrt{S_{NN}} = 2.76$ TeV observed by CMS. Figure taken from[32].

Studying asymmetric systems, namely d+Au or p-Pb for LHC, can help to separate which effect interplays in the observed jet quenching. In Figure 4.2 a) one can see that the distributions for d+Au collisions are very similar to those observed in pp collisions. This result is inconsistent with the explanation attributing the suppression in Au+Au collisions to initial-state effects. That is one of the great evidences of the formation of the QGP.

4.1.2 Flow

The system created in heavy-ion collisions is very hot and dense, which means that it goes through an expansion and cooling down process, making the system dilute enough and the color interactions strong enough to hadronize. The expansion that the system undergoes is collective, and it is called collective flow, a very important subject in heavy-ions, which can provide information of the transport properties of the QGP, such as its specific viscosity, and its EoS.

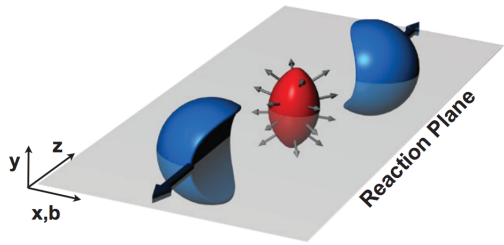


Figure 4.3: A non-central Pb-Pb collisions producing an almond shaped interaction. Figure taken from [81].

Collisions between two nuclei can occur with different geometry. The impact parameter, b, characterizes the distance between the centers of the incoming nuclei in the plane transverse to the beam direction. When the impact parameter is close to zero, the collision is central, where the two nuclei collide heads-on and the overlap region is high, thus having a high number of participating nucleons. When the impact parameter is close to 1, the collision is said to be peripheral and the overlap region is small.

Non-central collisions have an anisotropic geometrical overlap region. The initial spatial asymmetry in this region gives rise to a pressure gradient that enhances the particle production in the reaction plane, defined by the impact parameter b and the beam direction z (Figure 4.4). Even though peripheral collisions are depicted with a perfect almond shape, such as Figure 4.4, in reality the shape is a superposition of different geometries The expression to quantify flow is a Fourier expansion of the azimuthal distribution of the particles emitted in respect to the reaction plane[75]

$$\frac{dN}{d\phi} \propto 1 + \sum_{n=1}^{\infty} 2v_n cos(n(\phi - \Psi_r)) \tag{4.1}$$

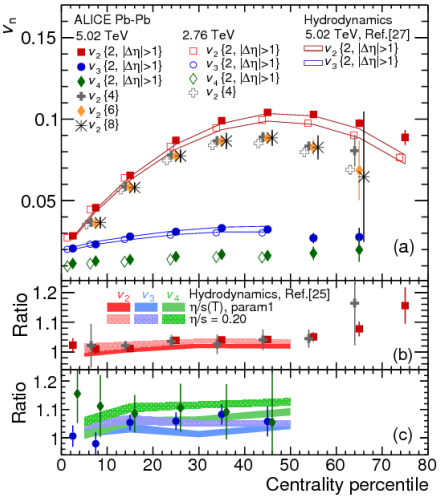


Figure 4.4: a) Anisotropic flow v_n as a function of event centrality including hydrodynamic calculations. Figure taken from [88].

Here, Φ_r is the reaction place angle. It is represented only by cosine terms, since sine terms vanish due to reflection symmetry to the reaction plane for symmetric collisions. In the absence of flow, the above function is constant and the Fourier coefficients should be zero. Therefore, the non-zero harmonic coefficients of the Fourier expansion carry information of the type of flow. The first harmonic coefficient is related to the overall shift of the distribution in the transverse plane, and it is called directed flow, where its magnitude is given by v_1. The second harmonic describes the eccentricity of an elliptic-shaped distribution with magnitude v_2 and it is called elliptic flow. As we go through the v_n coefficients, we get triangular flow v_3, quadrangular flow v_4 and so on, implying that all the coefficients have an actual physical meaning.

Due to the elliptical shape of the interaction region in non-central collisions, the dominant flow coefficient is the elliptic flow v_2. In central collisions, however, the interaction region gets closer to the nucleus shape and the magnitude of v_2 is not so dominant anymore compared to higher order harmonics, which are dominated by fluctuations. In the transition from RHIC to LHC energies, v_2 increased 30%, as predicted by hydrodynamic models which include viscous corrections. The large values of v_2 provide evidence that the medium produced in heavy-ion collisions behaves like an almost perfect fluid[2].

Figure 4.4 shows the centrality dependence of v_2, v_3 and v_4. The left of the plot represents central collisions and the centrality decreases to the right. It is possible to see that v_2 increases in peripheral collisions. Higher harmonics are not as sensitive to centrality as v_2. Figure 4.4b) and c) shows the ratio of the measured coefficients at 5.02 TeV in comparison to 2.76 TeV[88]. The results are in agreement with hydrodynamic predictions.

4.2 Small collision systems: Signs of collectivity

The focus of relativistic heavy-ion collisions is to study the QGP. However, it is not possible to measure the QGP itself directly and, experimentally, it is only feasible to measure the particles produced in the final stages of the collisions. To understand what is seen in heavy-ion collisions, proton-proton collisions are used as a reference experiment - a collision system where the QGP production is not expected - and proton-nucleus collisions is the control experiment, used primarily to disentangle cold nuclear matter (CNM) from hot nuclear matter effects.

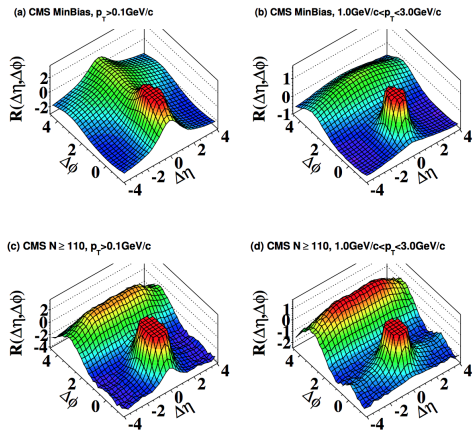


Figure 4.5: Results from CMS on 2D two-particle correlation functions for 7 TeV pp. (a) Selection of minimum bias events with $p_t > 0.1$ GeV/c, (b) Selection of minimum bias events with $1 < p_t < 3$ GeV/c, (c) Selection of high multiplicity events with $p_t > 0.1$ GeV/c and (d) Selection of high multiplicity events with $1 < p_t < 3$ GeV/c. Figure taken from [61].

The initial spatial anisotropy in heavy-ion collisions is transferred to the final state particle distribution through the hydrodynamic expansion. The system undergoes a collective expansion which generates correlated multiparticle production, therefore measurements of azimuthal correlations might reveal important information on whether the system behaves collectivelly or not. Taken this information into consideration, something unexpected happened when CMS published in 2010 results

of long-range ($2 < |\Delta\eta| < 4$), near-side ($\Delta\phi \approx 0$) angular correlations in high multiplicity pp collisions, shown in Figure 4.5, showing a similar behavior as the one observed in Pb-Pb collisions.

That could indicate collective behavior in high multiplicity pp collisions, which can point out to the production of the QPG in the reference experiment. Therefore, the investigation of flow in small systems became a hot topic within the heavy-ion community.

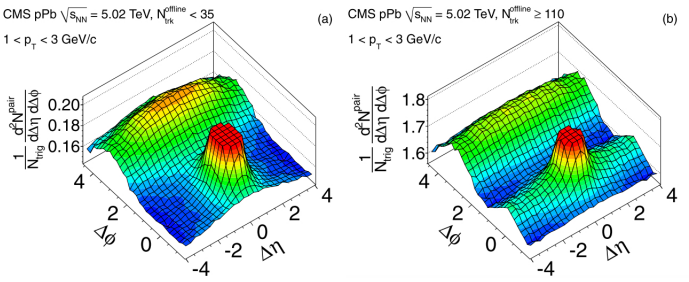


Figure 4.6: Two-particle correlation for pPb collisions at 5.02 TeV for pairs of charged particles with 1 < pT < 3 GeV/c. On the figure (a) the results for low-multiplicity events are shown and on figure (b) for a high-multiplicity selection. Figure taken from [33].

Another unexpected scenario was shown in 2012, when CMS published results on two-particle angular correlations for charged particles emitted in p-Pb collisions at 5.02 TeV. They studied the correlations over a broad η range and full azimuth coverage as a function of the charged particle multiplicity and p_t. The results are shown in Figure 4.6

The two-particle angular correlations are performed using two dimensional correlation functions in $(\Delta\eta, \Delta\phi)$. Here, $\Delta\phi$ is the difference in azimuthal angle between two emitted particles, and $\Delta\eta$ is the difference in pseudorapidity. It is possible to see in Figure 4.6-a, the low multiplicity selection, that the prominent feature is the peak in the near side at $(\Delta\eta, \Delta\phi) = (0,0)$, which are particles coming from the same jet. The elongated part in the away side region at $\Delta\phi = \pi$ arises from particles originated from back-to-back jets.

When we go to the high multiplicity events, it is possible to see a ridge-like structure in the near side, in addition to the peak. This observed ridge is very interesting because it is similar to the one observed in high multiplicity pp collisions and to PbPb collisions.

One way of quantifying the nuclear effects in p-Pb collisions is through the p_t-differential yield relative to pp[4], again the nuclear modification factor R_{pPb}. The R_{pPb} is calculated as

$$R_{pPb}(p_t) = \frac{d^2N_{ch}^{pPb}/d\eta dp_t}{< T_{pPb} > d^2\sigma_{ch}^{pp}/d\eta dp_t} \tag{4.2}$$

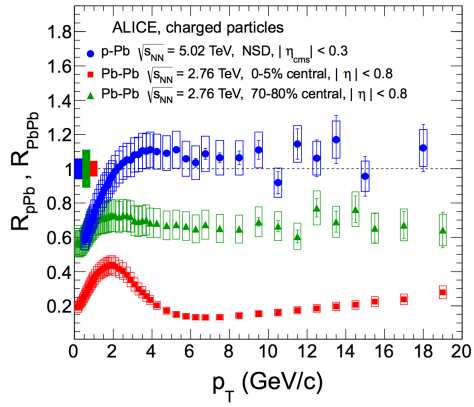
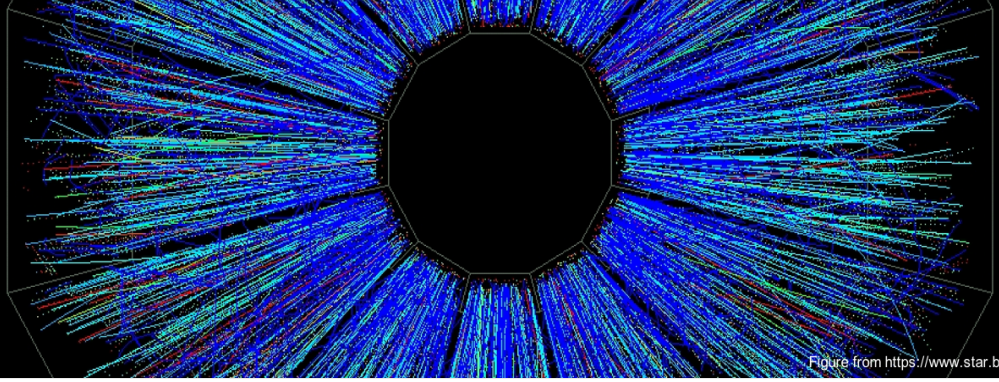


Figure 4.7: The nuclear modification factor of charged particles as a function of transverse momentum in p-Pb collisions at 5.02 TeV. Figure taken from [4].

The nuclear modification factor R_{pPb} was measured and the results were compared to R_{PbPb}, as it is shown in Figure 4.7

The results show that the nuclear modification, R_{pPb}, is consistent with unity. On the other hand, for PbPb collisions, R_{PbPb} exhibit a strong suppression in central collisions. The results for p-Pb collisions indicate that this strong suppression in PbPb data is not caused by initial-state effects, but it could be a mark of the hot matter created in heavy-ion collisions.

Accordingly to these and other results on p-Pb collisions, it is still not clear whether p-Pb collisions can indeed produce the QGP, however it is clear that there is more to this system than previously thought and therefore it is the subject of study of this work. It is expected that the charged-particle multiplicity distributions is sensitive to the color glass condensate, therefore its study can shed some light on this collision system.



Chapter 5

Charged-particle Multiplicity Distributions

After all the basic concepts were introduced, it is possible now to present the main goal of this work. This chapter is dedicated to explain what charged-particle multiplicity distributions are, why they are interesting and how they can be translated into physical concepts through discussing the models used to describe theoretically the distributions, and show previous results from different colliding systems.

The measurement of the charged-particle multiplicity of relativistic high energy collisions is one of the most fundamental observables. The two most important charged-particle multiplicity measurements are the pseudorapidity density $\left(\frac{dN_{ch}}{d\eta}\right)$ vs. η and the charged-particle multiplicity distribution $P(N_{ch})$ as a function of N_{ch}. The goal of this work is to measure the charged-particle multiplicity distribution in proton-lead collisions and, in this text, the shorter term "multiplicity distributions" refers to "charged-particle multiplicity distributions".

The multiplicity distributions are a valuable measurement since the probability P(N_{ch}) of producing a specific number N_{ch} of charged particles reflects directly the particle production mechanisms. For example, if a multiplicity distribution follows a Poisson distribution, it can be inferred from this result that the final state particles were produced independently, since that's a property of a Poisson distribution. If the distribution deviates from a Poisson, then there are signs of correlations in the particle production. The measurement of multiplicity distributions can therefore impose constraints in theoretical models of particle production, such as the available Monte Carlo generators, serving the purpose of improving or even rejecting them.

5.1 KNO Scaling

The interest in multiplicity distributions grew after the famous paper of Koba, Nielsen and Olesen[62], in 1972. In their work, they derived theoretically, based on the Feynman scaling[49], that the multiplicity distributions, when normalized, show the same shape, independently of the energy of the beam. Feynman used phenomenological arguments in his paper[49], in 1969, to define the probability of finding a particle i with a specific transverse momentum p_T and $x_F = \frac{p_Z}{W}$, where x_F is known as Feynman-x and $W = \sqrt{s}/2$, as

$$f_i(p_T, x)\frac{dp_z}{E}d^2p_T, E = \sqrt{m^2 + p_T^2 + p_z^2} \tag{5.1}$$

Feynman deduced that f_i becomes independent of W at high energies and this is known as the Feynman scaling, where f_i is the scaling function. Integrating eq.5.1 gives us the result $\langle N_{ch}\rangle \propto ln\sqrt{s}$, which means that the average number of particles increases logarithmically with W. If we consider that the maximum rapidity also follows the same condition, increasing with $ln\sqrt{s}$, then dN/dy = constant. This means that, in a rapidity distribution, the plateau around mid-rapidity is independent of $\sqrt{s}$, which also extends to the pseudorapidity distribution, $\mathrm{dN/d}\eta$ at mid-rapidity, considering that Feynman scaling holds.

Koba, Nielsen and Olesen generalized Feynman scaling to a q-dimensional scaling function. In this scenario, we have q particles with energy E_q, momenta $p_{T,q}$ and Feynman-x $x_{F,q}$ and the KNO scaling is obtained by calculating

$$\begin{gathered}\langle N_{ch}(N_{ch}-1)....(N_{ch}-q+1)\rangle = \\ \int f^{(q)}(x_1, p_{T,1}; ...; x_q, p_{T,q})\frac{dp_{z,1}}{E_1}dp_{T,1}^2...\frac{dp_{z,q}}{E_q}dp_{T,q}^2\end{gathered} \tag{5.2}$$

Integrating eq.5.2 and with a substitution of $\langle N_{ch}\rangle \propto lns$, we find that the multiplicity distribution scales as

$$P(N_{ch}) = \frac{1}{N_{ch}}\psi\left(\frac{N_{ch}}{\langle N_{ch}\rangle}\right) + \mathcal{O}\left(\frac{1}{\langle N_{ch}\rangle^2}\right) \tag{5.3}$$

Here, we define the KNO variable $z = N_{ch}/\langle N_{ch}\rangle$, and $\psi(z)$ is then a function independent of energy. That implies that all multiplicity distributions would follow a specific curve when plotted as a function of the KNO variable z.

In 1985, the UA5 collaboration observed that KNO scaling was violated in pp collisions at $\sqrt{s} = 540GeV$[13]. Furthermore, they noted that the distributions were "remarkably well described by a negative binomial distribution"[13]. However, when the energies reached up to$\sqrt{s} = 900GeV$, deviations also from a negative binomial distribution (NBD) were observed, both by UA5 and later by Tevatron at $\sqrt{s} = 1800GeV$. The deviations were due to a "shoulder" structure that shows up at around $N_{ch} \approx 2\langle N_{ch}\rangle$, which brought up a model combining two NBDs to describe the data.

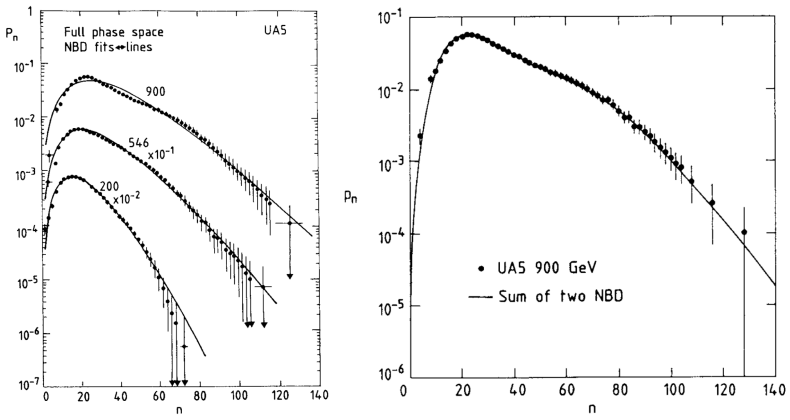


Figure 5.1: Left: Negative binomial distributions fitted to the true multiplicity at 200, 546 and 900 GeV. Right: Sum of two NBDs fitted to the true distribution at 900 GeV. Figure taken from [50].

UA5 collaboration fitted multiplicity distributions at different energies using a single NBD as shown in Figure 5.1 left.

It's clearly possible to see that a single NBD fails to describe the data at 900 GeV. The UA5 experiment tried to fit eight different functions to the data, being successful with the one with the most parameters; a sum of two NBDs, which improves drastically the fit, as we can see in Figure 5.1 right. The Negative binomial distribution and some theoretical explanations for its good fit to data is described in the next section.

5.2 Negative Binomial Distribution

There is, generally, a certain familiarity with the binomial experiment, where there is a fixed number of independent trials and only two possible outcomes. When we say negative binomial distribution, it is almost intuitive to think of a binomial distribution flipped over the x-axis, turning it into "negative". However, the name "Negative binomial distribution" has a different connotation. In a regular binomial experiment, one looks at the number of successes in a specific number of trials, whereas in a negative binomial experiment, we are counting the failures until we reach a specific number of successes.

The negative binomial distribution is defined as

$$P_{p,k}^{NBD} = \binom{n+k-1}{n}(1-p)^n p^k, \tag{5.4}$$

where it is described the probability of obtaining n failures and k-1 successes before the k^{th} success in a Bernoulli experiment with probability of success of p. In the

limit of $1/k \to 0$, the NBD turns into a Poisson distribution and for k=1, we obtain a geometrical distribution. The binomial coefficient can be written as

$$\binom{n+k-1}{n} = \frac{(n+k-1)!}{n!(k-1)!} = \frac{(n+k-1)(n+k-2)...(k)!}{n!} \tag{5.5}$$

The mean $\langle n \rangle$ of the distribution is related to the probability p through the expression $\frac{1}{p} = 1 + \frac{\langle n \rangle}{n}$. The form of the NDB used to describe multiplicity distributions is then

$$P^{NBD}_{\langle n \rangle, k}(n) = \binom{n+k-1}{n} \left(\frac{\langle n \rangle / k}{1 + \langle n \rangle / k} \right)^n \frac{1}{(1 + \langle n \rangle / k)^k} \tag{5.6}$$

Figure 5.2 shows NBDs for different parameters, also showing the extreme cases, eg. for large k where it approaches to a Poisson distribution. k is a parameter which affects the shape of the distribution, therefore, if k is constant and does not depend on energy, the KNO scaling would then be valid.

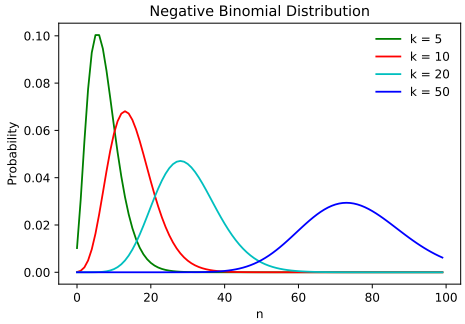


Figure 5.2: Examples of negative binomial distributions with different parameter values.

It's not completely understood why multiplicity distributions follow negative binomial distributions, though some explanations have been developed. The most commonly referred theoretical attempt to obtain a NBD-shaped multiplicity distribution is the clan model. This model describes the underlying production through a cascading process, where all the particles stemming from the same ancestor are called a *clan*, or cluster. The production of the ancestors is independent, following a Poisson distribution $P(N, \langle N \rangle)$, where $\langle N \rangle$ is the average number of clans. To obtain the probability $P_c(n_c)$ of producing n_c particles in a clan, one has to consider that if there are no particles, there is no clan, therefore $P_c(0) = 0$. The probability p of producing extra particles in a clan is proportional to the number of particles which already exist

$$pn_c = \frac{(n+1)P_c(n_c+1)}{P_c(n_c)} \tag{5.7}$$

It can therefore be shown that

$$P_c(n_c) = P_c(1)\frac{p^{n_c-1}}{n_c} \tag{5.8}$$

The overall multiplicity distribution is then given by

$$P(n) = \sum_{N=1}^{n} P(N, \langle N \rangle) \sum^{*} P_c(n_1)P_c(n_2)...P_c(n_N), \tag{5.9}$$

Where $P(n)$ accounts for the distribution of clans and the particles among different clans[54]. Equation 5.9 can be shown to be a NBD, where $\langle n \rangle = \langle N \rangle P_c(1)/(1-p)$ and $k = \langle N \rangle P_c(1)/p$.

5.3 Previous Measurements

In order to provide some background for the results in this work, the highlights of previous multiplicity distributions measurements are shown in this section, starting from initial measurements from before the LHC.

Before the LHC

From 1971 until 1984, the Intersecting Storage Ring (ISR) was operating at CERN, colliding pp, p$\overline{p}$ and pα at a center-of-mass energy of $\sqrt{s} = 63GeV$. The Split Field Detector (SFD) was used to measure multiplicity distributions for NSD and inelastic pp events at $\sqrt{s}$ = 30.4, 44.5, 52.6 and 62.2 GeV. The results are shown in Figure 5.3, and it is clear that KNO scaling is still valid at these energies for NSD events.

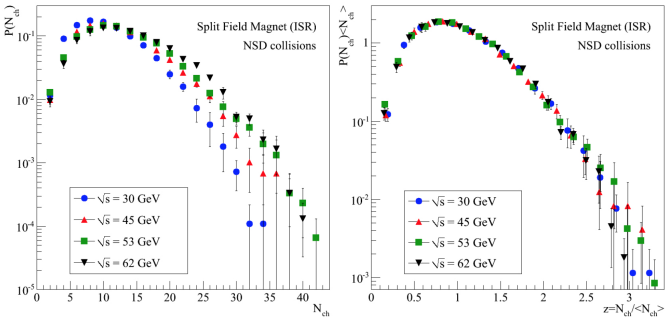


Figure 5.3: Charged particle multiplicity distributions obtained at ISR at four energies. The left plot show the normalized multiplicity distribution and the right plot shows the KNO scaling. Data from ISR[25] and figure from [70].

The Super Proton Synchrotron (SPS) operated at CERN from 1976, accelerating different combinations of e^+, e^-, p, $\overline{p}$ and ions. Later on, it took the name Sp$\overline{p}$S and collided p$\overline{p}$ at $\sqrt{s}$ = 900 GeV. The Underground Area 1 (UA1) and the Underground

Area 5(UA5) published multiplicity distribution measurements at $\sqrt{s} = 200$, 540 and 590 GeV. The results obtained by UA1 are shown in Figure 5.4.

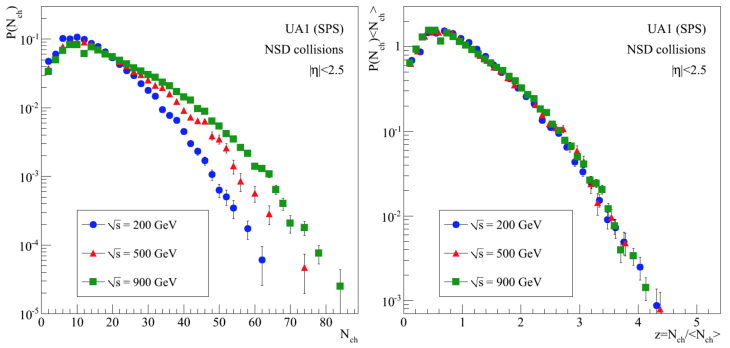


Figure 5.4: Charged particle multiplicity distributions obtained at UA1 at four energies. The left plot show the normalized multiplicity distribution and the right plot shows the KNO scaling. Data from UA1[37, 38] and figure from [70].

Just like ISR, the results from UA1 show that KNO scaling holds. On the other hand, as Figure 5.5 shows, this is not true for the measurements done by UA5. As mentioned previously, UA5 observed that KNO scaling is violated at $\sqrt{s} = 900$ GeV. UA5's results were questioned by UA1 at the time, since their conclusions disagree. The first difference one may notice about the two publications is the η interval, being $|\eta| < 2.5$ and $|\eta| < 1.5$ for UA1 and UA5, respectively. This η difference can lead to different conclusions and a direct comparison between the two[37] showed that, in the interval $|\eta| < 0.5$, both experiments are in agreement and, in the interval $|\eta < 1.5$, they disagree.

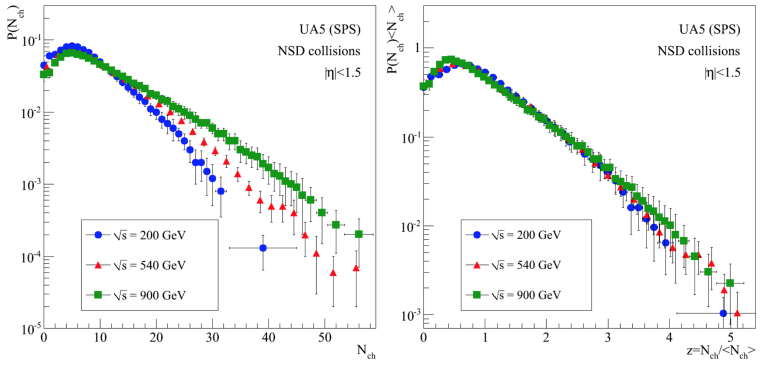


Figure 5.5: Charged particle multiplicity distributions obtained at UA1 at four energies. The left plot show the normalized multiplicity distribution and the right plot shows the KNO scaling. Data from UA1[37, 38] and figure from [70].

Tevatron started operating in 1983 at Fermi National Accelerator Laboratory, publishing $p\bar{p}$ results measured by the Collider Detector Fermilab at $\sqrt{s} = 1800$ GeV for different pseudorapidity ranges in 1993[77]. Figure 5.6 shows the original plot published by the CDF collaboration. The authors observe a weak KNO violation, however claim that more details on the systematics are necessary in order to conclude anything. Their results are also well described by a combination of two NBDs.

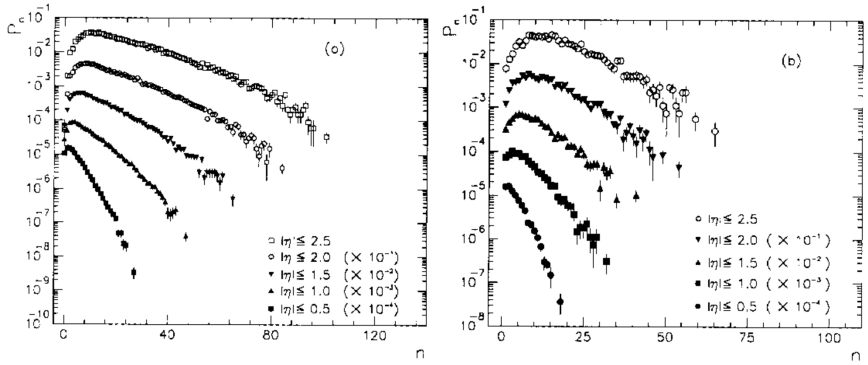


Figure 5.6: Charged particle multiplicity distributions obtained at CDF in four pseudorapidity ranges at a)$\sqrt{s} = 1800$ GeV and b)$\sqrt{s} = 630$ GeV [77].

LHC Era

The Compact Muon Solenoid (CMS) at LHC published in 2010 multiplicity distribution measurements in pp collisions at $\sqrt{s} = 0.9$, 2.36 and 7 TeV[60]. The results were obtained in different pseudorapidity intervals from $|\eta| < 0.5$ to $|\eta| < 2.4$. Figure 5.7 shows on the left the full multiplicity distribution for all the five pseudorapidity intervals at $\sqrt{s} = 0.9$ TeV, where there is a comparison with previous UA5 data and also with ALICE data. On the right, the KNO scaling appears strongly violated when comparing results at $\sqrt{s} = 0.9$ TeV and $\sqrt{s} = 7$ TeV.

More recently, the ALICE experiment published results at $\sqrt{s} = 0.9$, 7 and 8 TeV. A comparison with CMS results was made, showing that there is a disagreement both in the low multiplicity region and the tails of the distribution, which the authors attribute to possible differences in the models used to describe diffraction masses[5].

5.4 Models

To really understand the physics behind the distributions, one has to compare the data to theoretical models. In this work, three different models are used and they are presented in this section.

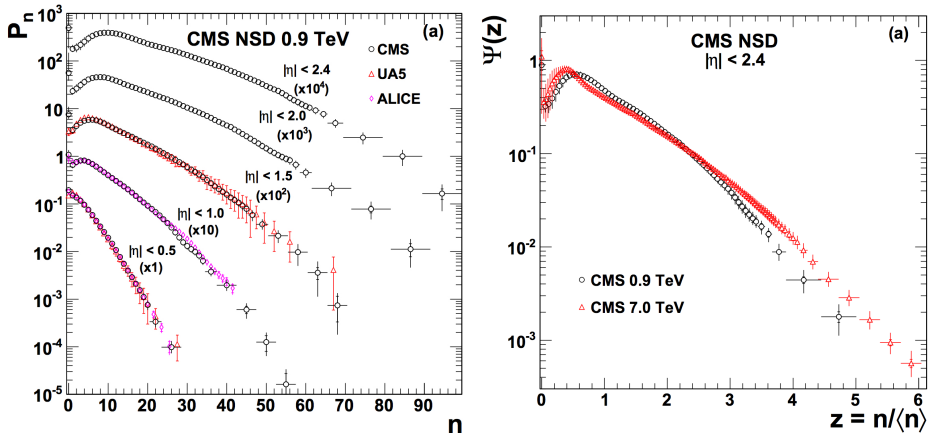


Figure 5.7: Charged-particle multiplicity distribution published by the CMS experiment[60]. Left: Distributions in 5 pseudorapidity ranges at 0.9 TeV. Right: Distributions in KNO variables at 0.9 and 7.0 TeV in the range $|\eta| < 2.4$.

5.4.1 EPOS LHC

EPOS[74] is a Monte Carlo model for minimum bias hadronic interactions, which was already well tested in previous results from RHIC and TEVATRON. EPOS treats the parton evolution in terms of "parton ladders", shown in Figure 5.10. The parton ladder can be seen as a longitudinal color field, therefore treated as relativistic string. The decay of these strings lead to quark-antiquark production, which are the hadrons.Their evolution can be described by simple evolution equations, such as DGLAP, and, at high energies, the gluon densities start to play a big role in non-linear effects have to be taken into account, which is done using BK evolution.

The elementary scatterings form strings, which break and form hadrons. In Heavy-ion collisions, before the hadronization, EPOS distinguishes between strings in high density areas (core) and in low density areas (corona), which is a very unique feature of EPOS. The corona plays an important role in the centrality dependance of some observables in Heavy-ion collisions, and the strings in this region hadronize normally, event-by-event. The rest of the string form the core, which hadronize colectivelly. Having the initial conditions of the system, it evolves according to the equations of hydrodynamics.

EPOS was changed in order to reproduce better the data at LHC energies, and this specific tune is called EPOS LHC. The most important change from EPOS to EPOS LHC is the concept of collective flow, where the flow intensity depends only on the total mass of the core in EPOS LHC.

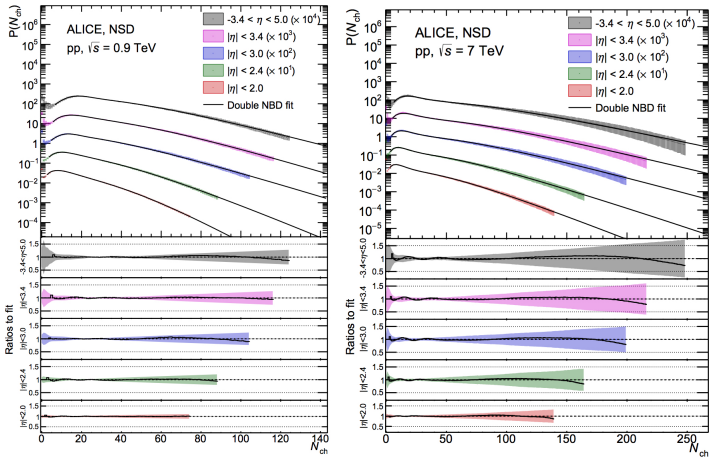


Figure 5.8: Charged-particle multiplicity distributions published by the ALICE experiment[5]. Left: Distributions at 0.9 TeV. Right: Distributions at 7 TeV.

5.4.2 DPMJET

DPMJET is a Monte Carlo model based on Dual Parton Model(DPM)[29] capable of simulating hadron-hadron, hadron-nucleus, photon-hadron, photon-photon and photon-nucleus interactions. DPMJET treats individual nucleon-nucleon collisions and it uses PHOJET[47] to do so.

The model deals with soft interactions implementing the DPM as a two-component model using Reggeon theory and perturbative QCD is used to deal with hard interactions. Since the model simulates individual nucleon-nucleon collisions, the extension to hadron-nucleus collision is done using the Glauber-Gribov multiple scattering formalism. Particles are created through string fragmentation processes.

5.4.3 Double NBDs

As mentioned in section 5.2, negative binomial distributions can fit to charged-particle multiplicity distributions. However it was observed that at some point the data starts to deviate from a single NBD, which prompts to try a double NBD fit. The double NBD fit has 5 free parameters and is given by

$$P(n) = \alpha P_{NBD}(n, \langle n\rangle_1, k_1) + (1-\alpha) P_{NBD}(n, \langle n\rangle_2, k_2), \tag{5.10}$$

The first NBD, $P_{\mathrm{NBD}}(n, \langle n\rangle_1, k_1)$, describes the soft interactions, the most dominant region, whereas the second NBD, $P_{\mathrm{NBD}}(n, \langle n\rangle_2, k_2)$, describes the semi-hard interactions, where mini-jets are observed and α is a parameter which describes the fraction of soft events.

By fitting UA5 data at $\sqrt{s} = 200$, 546 and 900 GeV, it was found that the mean of the two components show a dependence on the energy like $\langle n\rangle_1 = -5.54 + 4.72 ln(\sqrt{s})$,

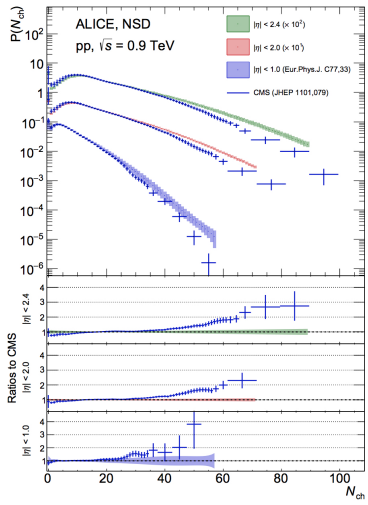


Figure 5.9: Comparison of the multiplicity distributions between ALICE and CMS results for pp collision at 0.9 TeV [5].

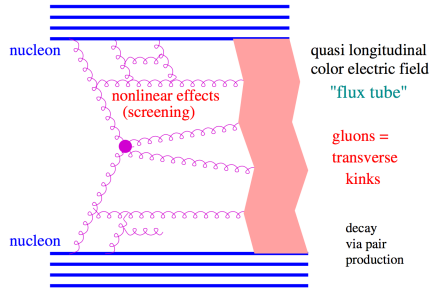
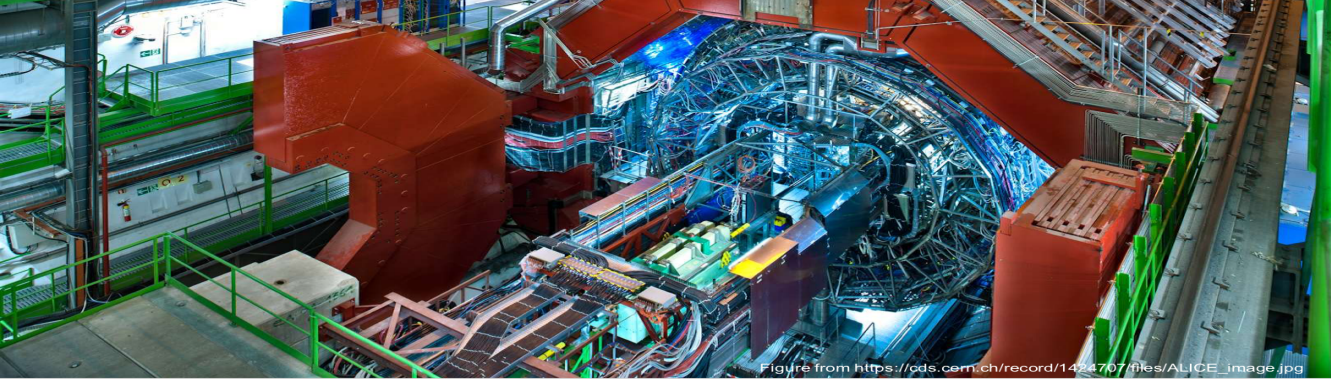


Figure 5.10: Elementary scattering in EPOS model. Figure from [74].

where $\langle n \rangle_2 \sim 2\langle n \rangle_1$[52]. More recent results from CMS and ALICE[5, 60] for pp collisions at different $\sqrt{s}$ found that $\langle n \rangle_1$ and $\langle n \rangle_2$ increase with $\sqrt{s}$ and $\langle n \rangle_2 \simeq 3\langle n \rangle_1$.



Chapter 6

Experiment

The data analyzed in this work were collected by the ALICE experiment[36] at the Large Hadron Collider[30]. In this chapter, the experiments and facilities used for data taking are presented, starting with the Large Hadron Collider and ALICE experiment in general, and going through two specific sub-detectors of ALICE, the Forward Multiplicity Detector (FMD) and the Silicon Pixel Detector (SPD), since they have a major role for my analysis. The chapter finishes presenting the details of the used trigger detectors.

6.1 The Large Hadron Collider

The Large Hadron Collider (LHC), located at CERN, the European Organization for Nuclear Research, on the Franco-Swiss border, is currently the largest and most powerful collider in the world. The LHC consists of a ring, measuring 27 km in circumference, made of superconducting magnets placed approximately 100 m under the ground. Inside the ring, there are two tubes, called beam pipes, kept at ultrahigh vacuum, where two beams of particles or nuclei travel in opposite direction before they are made to collide in 4 different specific points of LHC.

The collision points are the four main experiments of LHC: ALICE, ATLAS, CMS and LHCb. ATLAS and CMS are general purpose detectors, where the goal is to investigate a very wide range of physics, from discoveries of new particles to the study of dark matter and physics beyond the Standard Model. These two detectors were responsible for the discovery of the Higgs boson in 2012. LHCb is an experiment specialized in investigating differences in matter and antimatter studying the b quark. The ALICE experiment, which is the experiment used in this work, is specialized in studying heavy-ion collisions and it is described in details in this chapter.

The LHC collides protons and Lead ions, and to accelerate the particles that will collide inside the LHC tunnels, a complex structure is necessary, and the process is done in different steps, since the LHC itself can't accelerate stationary particles and

bring them to speeds close to the speed of light. A pre-acceleration system is used to accelerate particles until they have enough speed to enter the LHC. An illustration of the CERN experiment complex can be seen in Fig 6.2.

The first part of the acceleration is done in a linear accelerator, LINAC 2, where the protons are accelerated until they reach 1/3 of the speed of light and and energy of 50 MeV. In sequence, the protons go to the booster, a circular accelerator that will bring the protons to 91% of the speed of light and an energy of , and send them to the third step of the acceleration, the PS (Proton Synchrotron), which is a circular accelerator with 628m in circumference. The PS accelerates the protons until they acquire 99.9% of the speed of light and reach an energy of 25 GeV. At this point, a limit is reach and it is not possible to increase the speed of the protons anymore. Even though the speed can't be increased, it is still added more energy to the protons in PS and that is translated into mass, meaning that the protons get heavier than they are when at rest. After acquiring an energy of 25 GeV, the protons are sent to the fourth step of the acceleration, the SPS (Super Proton Synchrotron), a ring with 7km in circumference, which increases the energy of the protons to 450 GeV. The protons are ready to be injected into the LHC, where they are going to gain the necessary extra energy. The process to accelerate heavy lead ions is very similar, but it starts at LINAC 3, followed by LEIR, and then going to PS, SPS and, finally, the LHC.

6.2 ALICE Experiment

ALICE (A Large Ion Collider Experiment)[36] is a detector optimized to study heavy-ion collisions. It is designed to study the properties of the QGP and the physics of strongly interacting matter. The detector has been built by a collaboration of more than 1000 physicists and engineers from institutes in 30 countries in the former L3 LEP experimental cavern. It is a massive detector, being 16m high, 16m wide, 26m long, weight 10000t, and has 18 different detector systems.

ALICE is designed to cope with high multiplicities and low p_t particles. The subsystems provide high-momentum resolution and Particle Identification (PID) over a broad momentum range[36]. The Time Projection Chamber (TPC) is capable of tracking and identifying the thousands of particles that can be created in each heavy-ion collision. The ALICE detector can be divided into 3 main parts: The central barrel detectors, which are located at mid rapidities (ITS, TPC, TRD, TOF, HMPID, EMCAL and PHOS), the forward detectors, used for triggering, centrality estimation and measurement of multiplicities of charged particles or photons(FMD, T0, V0, ZDC and PMD), and the Muon Spectrometer, which tracks and identifies muons. The location of each subsystem is depicted in Fig 2.1.

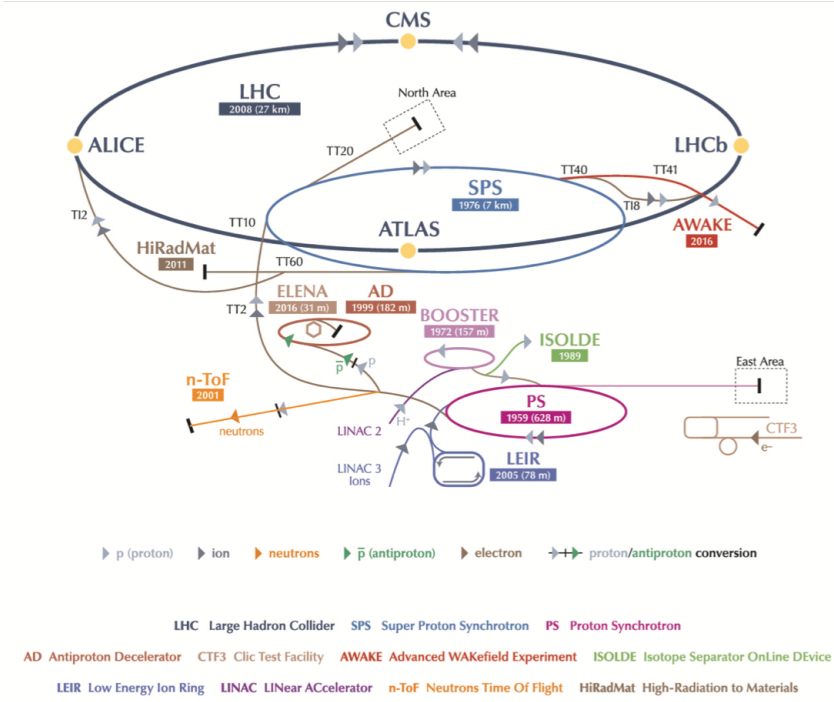


Figure 6.1: The CERN accelerator complex. Credit: CERN.

6.3 Central Barrel Detectors

The detectors covering the central part are the main components of ALICE since they are responsible for tracking and PID, which are essential features for the physics addressed by ALICE.

6.3.1 Inner Tracking System - ITS

The ITS is the innermost detector of ALICE. Its basic function is to determine primary and secondary vertices, and PID and tracking of low momentum particles[44]. The ITS is composed of 6 concentric layers of silicon detectors, and consists of 3 subsystems: the innermost 2 layers of ITS constitute the **Silicon Pixel Detector** (SPD), in the middle there is the **Silicon Drift Detector** (SDD) and the outermost two layers are the **Silicon Strip Detector** (SSD).

- **Silicon Pixel Detector** - It is the detector located closest to the interaction point (IP), being ALICE's main vertexing detector, and one of the triggering detectors, due to its fast read-out capabilities. The SPD barrel consists of two layers of silicon pixel around the beam pipe at a radius of 3.9 cm and 7.6 cm.

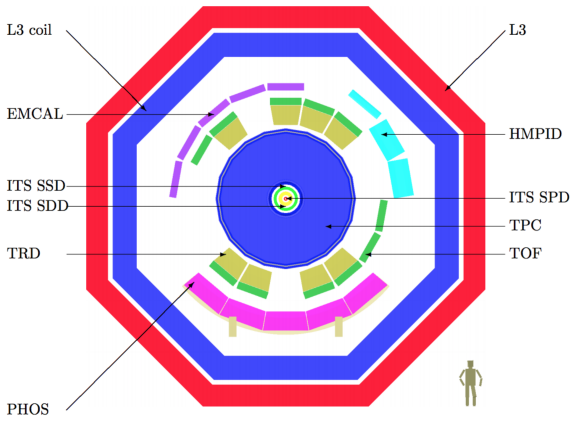


Figure 6.2: An overview of ALICE's central barrel detectors. Figure from [57].

Because of the high density of particles produced in heavy-ion collisions, the SPD is a detector with high granularity. It is used for measuring charged-particle multiplicity in the range $|\eta| < 2.0$, and therefore an important detector for this work in particular. The SPD is discussed in more details in the next chapter.

- **Silicon Drift Detector** - It is the third and the fourth layers of the ITS, made of silicon drift detectors with very high resolution two dimensional sensors, and analogue readout, which, combined with SPD, is used for PID. During the passage of a particle, electrons are released and they drift under the effect of an electric field applied along a parallel direction to the surface of the wafer, being collected by an array of anodes.[69]
- **Silicon Strip Detector** - The two outer layers of ITS, consisting of 1698 double-sided microstrip silicon devices. The SSD has the role of matching the tracks from the TPC, Time Projection Chamber, to the ITS.

6.3.2 Time Projection Chamber - TPC

In central Pb-Pb collisions, thousands of particles are produced [3, 7], and, according to the study purposes of ALICE, those particles and they trajectory have to be identified. The TPC is main responsible for tracking in the central barrel, providing, besides tracking, momentum measurement and PID at $0.1 < p_t < 100$ GeV/c and $|\eta| < 0.9$[65]. The ALICE's TPC is an enormous and a very ambitious detector, currently being the largest in the world. It is a $88m^3$ cylinder, about 5m in diameter and 5m long, filled with gas and divided in two drift regions by a central electrode.

When a charged particle traverses the TPC, it ionizes the atoms of the gas mixture. The knocked electrons drift through the electric field towards the anode plane, where they are readout. At this point, it is possible to get the x,y projection of the tracks, and the z coordinate can be reconstructed from the drift time. The Kalman

filtering method [20] is used to obtain the reconstructed (x,y,z).

6.3.3 Transition Radiation Detector - TRD

The TRD uses the principle of transition radiation, the radiation emitted when a charged particle passes through a boundary between two different media, to identify electrons at p > 1 GeV/c, being the main electron detector in ALICE, working in conjunction with the ITS and TPC. The detector is also capable of triggering processes with high transverse momentum ($p_t > 3$ GeV).

The TRD fills the space between the TPC and TOF[41] and is arranged in 18 sectors in azimuth, following the segmentation of TPC, where each module consists of 5 stacks of 6 layers of tracking chambers, resulting in 540 readout chambers. Each readout chamber is composed by a multi-wire proportional chamber with a drift region, filled with XeCO2 gas mixture, and a radiator. When electrons traverse the radiator, a transition photon is produced and subsequently absorbed by the entrance window to the drift region, producing signal rise at late times that is inexistent for pions, making it possible to distinguish both particles.

6.3.4 Time of Flight - TOF

The TOF detector[46] is responsible for particle identification in a intermediate momentum range, 0.5 - 2.5 GeV, using Multigap Resistive Plate Chamber (MRPC) technology. It is a cylindrical detector located around the TRD, also having 18 sectors in azimuth, which are divided in 5 modules, containing 1638 MRPC strips in total. It does not have full azimuthal coverage though, in order to minimize the material in front of other detectors.

The principal component of TOF, the MRPC, consists of a stack of resistive plates with spaces in between filled with gas [10]. There is an anode and a cathode electrode placed on the outer surface of the two outermost resistive plates, and an uniform electric field exists between them. When a charged particle crosses the MRPC, it ionizes the gas and an amplification of it occurs as an electron avalanche due to the electric field.

The TOF obtains the information of the time of flight of particles going from the interaction point to TOF. In combination with the tracking information obtained by the ITS, TPC and TRD, it is possible to use the time, t, to obtain the mass of the particle through

$$m = p.\sqrt{\frac{t^2}{l^2} - 1}, \tag{6.1}$$

where m is the mass, p is the momentum, t the time of flight, and l the track length. Since the masses of the particles are known, that's how the particle identification is done using TOF.

6.3.5 High Momentum Particle Identification Detector- HMPID

The PID capability of ITS, TPC and TOF is limited by the energy loss measurements, allowing an intermediate range of momentum. The HMPID[18] extend this capability to the high momentum range, 1 - 5 GeV, exploring the cherenkov radiation phenomenon, which is an emission that occurs when a charged particle passes through a medium at speeds greater than the speed of light in that medium. The detector consists of Ring Imaging Cherenkov (RICH) counters in 7 modules placed after the TOF detector at the two o?clock position.

A charged particle with high momentum passes through the Cherenkov radiator, which uses liquid perfluorohexane (C_6F_14). The Cherenkov photons form a constructive wave front at a specific angle with the particle's track, resulting in a light cone around the particle, which is detected by a photon counter and read out as a ring . The identification of the particle is therefore related to the angle of emission, θ_c, which can be related to the particle's velocity by

$$\cos\theta_c = \frac{1}{n\beta}, \tag{6.2}$$

with $\beta = c/v$, where c is the speed of light and v is the particle's velocity, and n is the refractive index of the medium. The tracking information obtained by other detectors is used to calculate the mass of the particle, thus identifying it like previously mentioned.

6.3.6 Electromagnetic Calorimeter - EMCal

Jet quenching measurements play an important role in high energy nuclear collisions, making is possible to probe the medium generated. The EMCal[40] enhances significantly ALICE's capabilities for measuring jets, where a combination with ITS and TPC's tracking measurements make the jet reconstruction possible for pp and PbPb collisions. Furthermore, the detector has a fast readout, acting as a L0/L1trigger for high energy jets.

The detector uses a layered Pb-scintilator with longitudinal wavelength shifting fiber organized in modules consisting of 2x2 towers. The detector has a segmentation of 12288 towers which are grouped into two super modules: the full size, which covers $|\eta| < 0.7$ and $|\phi| < 20°$, and the one-third size, which covers $|\eta| < 0.7$ and $|\phi| < 7°$. Due to design limitations of ALICE, the EMCal is limited to a region of about 110 degrees in azimuth.

6.3.7 Photon Spectrometer - PHOS

PHOS[45] is the outermost detector in the central barrel of ALICE, being spectrometer designed to detect high-energy photons and neutral mesons. The detector consists of 17280 channels of lead-tungstate crystals coupled to PIN-diodes, and it is placed at the bottom of the ALICE. The photon spectrometer is capable of detecting direct photons and di-photons, which provide information of the initial phase of the collision.

6.4 Forward Detectors

A full event covers around 20 unities in rapidity, and the large rapidities regions are usually called as forward regions. The forward detectors were designed to explore the forward region physics and, in ALICE, those detectors are the PMD, FMD, V0 and T0, which are described following in this section. Since FMD is a key detector in this analysis, a very thorough description of the detector is done in the next chapter.

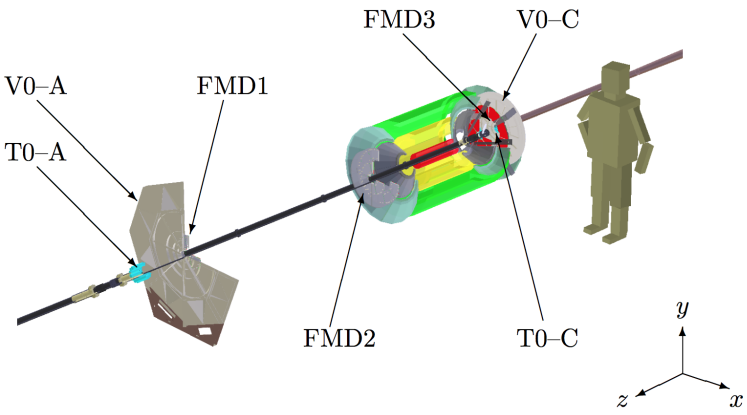


Figure 6.3: Schematic figure of the forward detectors. Figure from [57].

6.4.1 Photon Multiplicity Detector - PMD

The PMD is a preshower detector, where photons pass through a lead converter and initiate an electromagnetic shower. The shower produced gives rise to a larger signal in the plastic scintillators behind the converter. There is a veto detector positioned in front of the converter in order to reject charged particles, which are of no interest here. The preshower detector and the veto detector are based on honeycomb proportional chambers [48]. The PMD measure the photons multiplicities in the $2.3 < \eta < 3.7$ region.

6.4.2 V0 detector

The V0 system consists of two arrays of scintillator counters, V0A and V0C, positioned at both sides of the interaction point and covering the pseudorapidity ranges $2.8 < \eta < 5.1$ and $-3.7 < \eta < -1.7$ respectively. The system has an overlap with the FMD coverage, though it has smaller segmentation. The V0 system works as a minimum bias or centrality trigger, separates beam-beam interactions from background events and measures the beam luminosity. It is also able to measure charged particle multiplicity, which is used for centrality calculations.

6.4.3 T0 detector

The fast trigger roles are played by V0 and T0 detectors. Both detectors are capable of distinguishing between central, semi-central and minimum bias collisions, and of determining the vertex position, but V0 does not have the time precision required by the TOF system or to determine the vertex position accurately. That is when T0 comes in handy - a detector capable of delivering a start signal for TOF with high precision, which has to fulfill the time resolution requirement of 50 ps. The T0 is also subdivided into 2 units installed on both sides of the interaction point and made by quartz Cherenkov radiators coupled to photomultiplier tubes(PMT). As a convention, the left side, located at 370cm from the interaction point, is called T0-A, and the right side, located at 70cm from the interaction point, is called T0-C[24].

6.4.4 Zero Degree Calorimeter - ZDC

In heavy ion collisions, it is very important to distinguish peripheral collisions from central collisions. In peripheral collisions, only a few nucleons participate, which means that the nucleons not involved in the collisions (also called spectators) are going to fly away untouched along the beam line. The ZDC has the role of "collecting" the spectators, measuring their energy in order to estimate the number of nucleons participating in the collisions, hence determining their centrality.

The ZDC is composed of 2 proton calorimeters and 2 neutron calorimeters placed at 115 m away from the interaction point, having two calorimeters on both sides. It is placed along the beam line, at zero degree, therefore the detector system is called zero degree calorimeter.

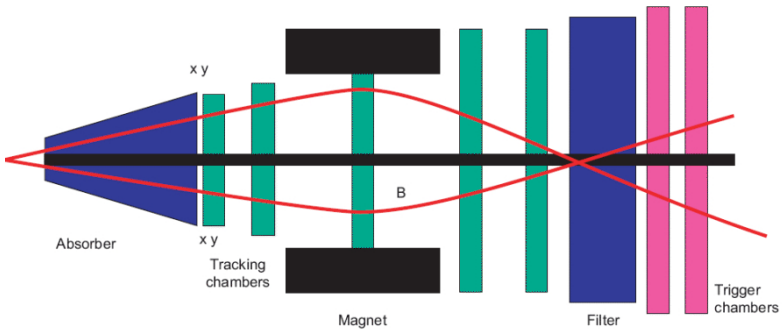


Figure 6.4: The main components of the Muon Spectrometer. Figure taken from [11].

6.5 Muon Spectrometer

The hard probes, such as quarkonium states, are used as a tool to study the hot and dense medium (QGP) created in heavy ion collisions. In the presence of the

QGP, the deconfinement would induce color screening, leading to a supression of quarkonium states. The goal of the muon spectrometer is to study the charmonium states (J/ψ, ψ') and bottomonium states (Υ, Υ', Υ'') via the $\mu^+\mu^-$ decay channel. The spectrometer occupies a big part of the C side of ALICE and covers a the pseudorapidity region -4.0 $< \eta <$ -2.5. It consists of a muon magnet, a trigger and tracking system, and 3 absorbers as it can be seen in the Figure 6.4. Since muons are very penetrating particles, the absorbers filter the background, reducing the flux of hadrons produced in the collisions.

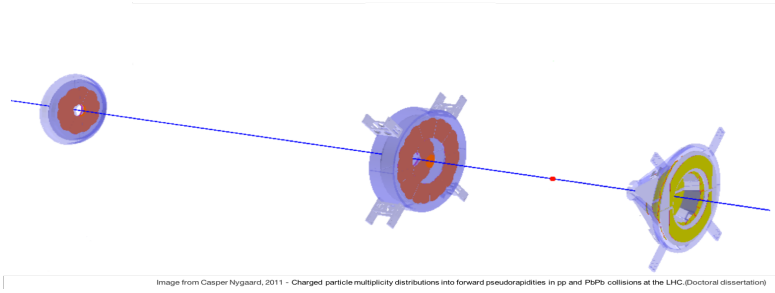


Image from Casper Nygaard, 2011 - Charged particle multiplicity distributions into forward pseudorapidities in pp and PbPb collisions at the LHC.(Doctoral dissertation)

Chapter 7

The FMD and SPD

The FMD and SPD are the main detectors used in this work, therefore they are presented in details in this chapter. The chapter contains the necessary tools to understand what they are made of and how they work, starting with basic concepts of energy loss in silicon and going through how the charged particles are detected by FMD and SPD.

7.1 Silicon Detectors

Silicon is the most used semiconductor material for charged particle detection. The biggest advantage is that it is widely available and it operates at room temperature [64]. In the next sections, the properties of semiconductors and how they are used to detect particles are presented.

7.1.1 Energy Loss in Silicon

There are different phenomena that can occur when a charged particle passes through matter. The particle loses energy when inelastic collisions with the atomic electrons occur, or it can be deflected from its incident direction when it is elastically scattered by the atomic nucleus. Other processes, like emission of Cherenkov radiation and bremsstrahlung can occur, though they are extremely rare in comparison to the other two, which occur many times per unit path length[64].The amount of energy transferred in each collision is a very small fraction of the particle's kinetic energy, however, many collisions occur in dense matter, therefore a considerable energy loss is observed even in thin layers.

Collisions occur with a certain probability with small fluctuations, being easily possible to use the mean energy loss per unit path lenght[64]. The quantum mechanical description of the energy loss, or also called the stopping power, was done by Bethe and Bloch, where they obtained the following expression:

$$-\frac{dE}{dx} = 4\pi N_a r_e^2 m_e c^2 \frac{Z}{A}\frac{z^2}{\beta^2}\left[\frac{1}{2}ln\left(\frac{2m_e c^2\gamma^2\beta^2 T_{max}}{I^2}\right) - \beta^2 - \left(\frac{\delta(\beta\gamma)}{2} + \frac{C}{Z}\right)\right] \tag{7.1}$$

where

N_A	Avogadro's number	6.0221415(10) x $10^{23} mol^{-1}$
r_e	Classic electron radius	2.817940325(28) fm
m_e	Electron mass	0.510998918(44) MeV
c	speed of light	
z	charge of the incident particle	e
Z	atomic number of the material	This case is $Z_{Si} = 14$
A	atomic mass of the material	This case is $A_{Si} = 28.0855$
β	p/m of the incident particle	
γ	$1/\sqrt{1-\beta^2}$ of the incident particle	
T_{max}	maximum energy transfer in a single collision	in MeV
I	Mean excitation energy	in eV
$\delta(\beta\gamma)$	Density effect correction to ionization energy loss	
C	shell correction	

T_{max} is the maximum energy transfer in one single collision, and it is produced by a head-on collision. For a particle with mass M, it is given by

$$T_{max} = \frac{2m_e c^2\beta^2\gamma^2}{1 + 2\gamma m_e/M + (m_e/M)^2} \tag{7.2}$$

The last terms in equation 7.1 represent the density, δ, and shell, C, corrections, which play an important role at high and low energies respectively. The shell correction comes from the assumption that the velocity of the incident particle is smaller than the orbital velocity of the atomic electrons or comparable. At low energies, the assumption of stationary electrons does not hold, making the correction necessary. The incident particle also has an electric field itself, which polarizes the atoms along its path when transversing a material. This polarization shields electrons far away from the particle path, making collisions with these electrons contribute less to the total energy loss. The effect increases as the energy of the particle increases, as well as the density of the material, hence the name "density effect"[64]. The densitity correction is necessary in this regime.

The equation describes well the energy loss of charged particles in a limited region of $0.1 \leq \gamma\beta \leq 1000$ for materials with intermediate Z, since, at the lower limit, the particle's velocity is comparable to the atomic electron's, making the assumption of stationary electrons break-down, and at the upper limit, radiative effects, such as Bremsstrahlung, are dominant.

As one can see in Figure 7.1, which shows the stopping power of μ in copper, the function has a minimum. At non-relativistic energies, the stopping power is dominated by the factor $1/\beta^2$ and decreases with increasing velocities. At certain velocity,

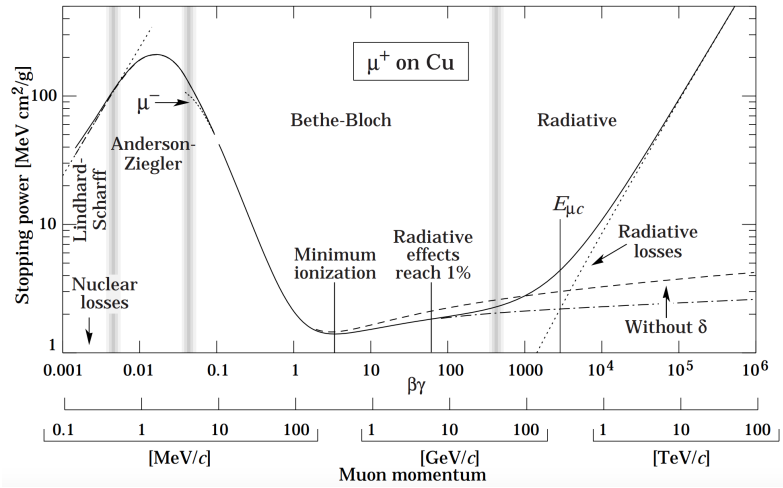


Figure 7.1: Stopping power for positive muons in copper as a function of $\beta\gamma$. Figure taken from [71].

a minimum is reached, and particles at this point are called *minimum ionizing particles*(MIP). This minimum is approximately the same for particles with the same charge, as show in Figure 8.1. Each different particle shows a distinct dE/dx curve at energies below the minimum ionizing point. This phenomenon is explored in the Particle Physics field to identify particles in this range.

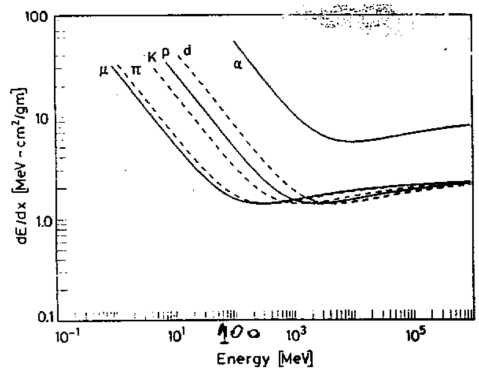


Figure 7.2: dE/dx for different particles as a function of energy. Figure taken from [64].

The approach described so far is valid only for heavy particles transversing matter, i.e particles heavier than electrons and positrons, like muons, pions, α-particles, etc. Due to their small masses, electrons(positrons) have additional effects in the energy loss mechanism, e.g. bremsstrahlung. The probability of bremsstrahlung increases

very fast with increasing energy, making the energy lost by radiation comparable or even greater than the energy lost by collisions. Therefore, the energy loss by electrons (positrons) can be described in two parts:

$$\left(\frac{dE}{dx}\right)_{tot} = \left(\frac{dE}{dx}\right)_{radiation} + \left(\frac{dE}{dx}\right)_{collisions} \tag{7.3}$$

where the collisions term is given by a modified Bethe-Bloch equation in order to account for the small masses.

7.1.2 Energy Straggling

Thus far, we have been talking about the mean energy loss of charged particles. However, the energy loss is a statistical process, which means that the energy lost by one particle depends on the amount of collisions suffered by it along its path, and it fluctuates around the mean value. Therefore, a beam of particles passing through a material with a specific thickness is going to show a distribution in energy, and this distribution depends on how thick the material is.

Thick Absorbers

When the material is relatively thick, it is expected that the number of collisions a particle transversing it will suffer is large. According to the Central Limit Theorem, the sum of N random variables, which follow the same distribution, goes to a Gaussian distribution when N is infinitely large. Thus, the energy loss distribution is given by[64]

$$\frac{dN}{d\Delta} \propto exp\left(\frac{-(\Delta - \overline{\Delta})^2}{2\sigma^2}\right), \tag{7.4}$$

where Δ is the energy loss in the absorber, $\overline{\Delta}$ is the mean energy loss and σ is the standard deviation.

Thin Absorbers

In case of thin absorbers, the central limit theorem does not hold since the number of collisions suffered by each particle is not large enough. Here, the small number of collisions makes it possible to large energy transfers to occur in one single collision, which generates a long tail in the high energy limit of the distribution. In this skewed distribution, the peak is no longer the mean, but the most probable value. The calculations in the regime was done by Landau, Symon and Vavilov for different regions of applicability [64], being characterized by the ratio

$$\kappa = \frac{\overline{\Delta}}{T_{max}}, \qquad \overline{\Delta} = \xi = 2\pi N_a r_e^2 m_e c^2 \rho \frac{Z}{A}\left(\frac{z}{\beta}\right)^2 x, \tag{7.5}$$

representing the ration between the mean energy loss calculated via the Bethe-Bloch equation and the maximum energy transfer.

Landau was the first to obtain the energy loss distribution for very thin absorbers, in this case $\kappa \leq 0.01$. In his theory, Landau made some assumptions:

- The maximum allowed energy transfer goes to infinity, which means that $\kappa \to 0$.
- Small energy transfers in distant collisions can be ignored, and each individual energy transfer is large enough to treat the target electrons as free.
- The incident particle maintain an approximately constant velocity.

From the above assumptions, Landau obtained the following formula:

$$\frac{dN}{d\Delta} = \frac{1}{\xi}\frac{1}{\pi}\int_0^\infty e^{-ulnu-u\lambda} sin\pi u du, \tag{7.6}$$

where

$$\lambda = \frac{1}{\xi}\left[\Delta - \xi\left(ln\xi - ln\epsilon + 1 - C\right)\right] \quad \text{and} \quad \ln\epsilon = \ln\frac{(1-\beta^2)\mathrm{I}^2}{2\mathrm{mc}^2\beta^2} + \beta^2 \tag{7.7}$$

In this equation, the value of ϵ is the minimum energy transfer described in the assumption number 2. This distribution has a long tail, as mentioned previously, and it is exemplified in Figure 7.3. Since the mean is displaced from the peak, it is more convenient to parametrize the distribution in terms of the most probable value (MPV), given by[57]

$$\Delta_p = \xi\left[ln\left(\frac{2mc^2\beta^2\gamma^2}{I}\right) + ln\frac{\xi}{I} + j - \beta^2 - \delta(\beta\gamma)\right], j = 0.2. \tag{7.8}$$

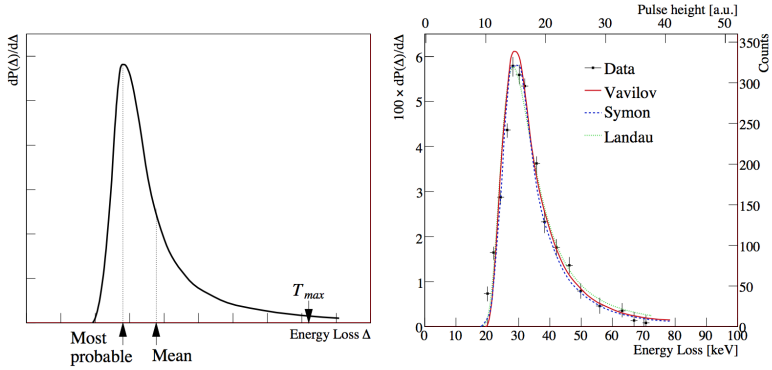


Figure 7.3: (a) Energy loss distribution in thin absorbers. (b) Calculations of the energy loss distribution by Landau, Syman, and Vavilov compared to data. Figure taken from[57].

Even though Landau's theory is a very good first order approximation, Vavilov and Shulek improved Landau's approach in order to account for excitations in the

inner shell of atoms. In their approach, the Landau distribution is convoluted with a Gaussian. Therefore, the total energy loss of N particles can be obtained by

$$F(\Delta) = \sum_{i=1}^{N} a_i F_n(\Delta) = \sum_{i=1}^{N} a_i (L_i(\Delta) \otimes G_i(\Delta))$$

$$F(\Delta) = \sum_{i=1}^{N} a_i \int_{-\infty}^{\infty} L(\Delta; \Delta', \xi_i) exp\left[-\frac{(\Delta_{p,i} - \Delta)^2}{2\sigma^2} \right] d\Delta' \tag{7.9}$$

Here, L_i and G_i denote Landau distribution with a most probable value $\Delta_{p,i}$ and a Gaussian distribution, respectively, and a_i is the weight of each Landau distribution convoluted with a Gaussian.

The energy deposited in thin absorbers can be described by 7.9, and that happens in silicon detectors used in high energy physics, like the Forward Multiplicity Detector, which is described in the next section.

7.2 Forward Multiplicity Detector

The Forward Multiplicity Detector was designed, built and tested at the Niels Bohr Institute, University of Copenhagen. It is a finely segmented silicon strip detector used to measure charged particles in the forward region. The motivation for building such a detector was to obtain a complete view of the particle production[57]. The specific physics goals of FMD required the following design considerations[57]:

- Broad η and full ϕ coverage.
- Fine η segmentation.

Since FMD was not designed considering any tracking capabilities, the detector was built based on silicon sensors organized in 5 rings grouped into 3 sub-detectors placed around the beam line: FMD1, FMD2, and FMD3. The first sub-detector, FMD1, consists of only one inner ring, whereas the latter two has 2 rings, an inner and an outer. The specific parameters of each ring is shown in Table 7.1.

Sub-detector	Ring	# ϕ segments	# r segments	z (cm)	r_{in}	r_{out}
FMD1	i	20	512	320.0	4.2	17.2
FMD2	i	20	512	83.4	4.2	17.2
	o	40	256	75.2	15.4	28.4
FMD3	i	20	512	-75.2	15.4	28.4
	o	40	256	-62.8	4.2	17.2

Table 7.1: Parameters of FMD sub-detectors [57].

7.2.1 Electronics

Every charged particle crossing the FMD is going to be detected starting at the sensors and following through the electronics until it gets to the data acquisition

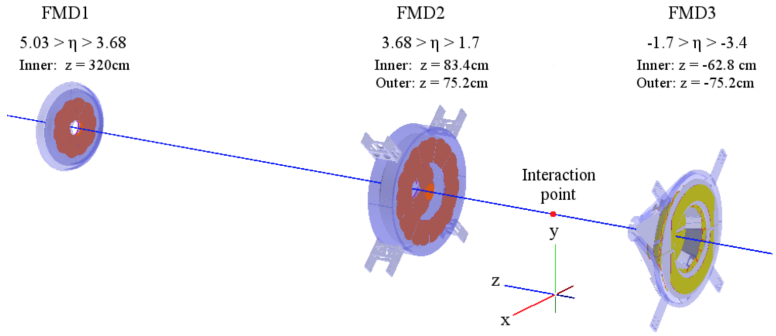


Figure 7.4: Overview of the FMD system. Figure taken from [70].

system. The electronics involved in the whole process is shown in Figure 7.5 and the most important pieces are described in this section.

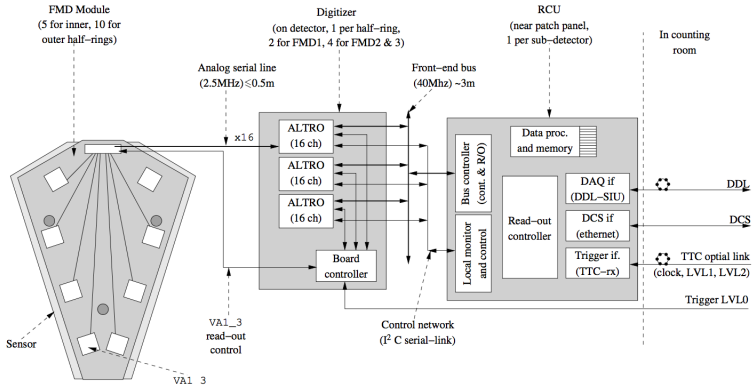


Figure 7.5: The Alice FMD electronics. Figure taken from [57].

Sensors

The basic units of the FMD rings are silicon sensors of two types, inner and outer, and each sensor is divided in sectors. The inner FMD rings are made by 10 sensors corresponding to 20 sectors, whereas the outer rings are made by 20 sensors and 40 sectors, as shown in Figure 7.6 and Figure 7.7.

As the Table 7.1 shows, each inner sensor sector is divided radially in 512 strips and the outer are divided in 256 strips. That means that each ring has 10240 channels and the whole FMD system has 51200 individual channels available for detection. The signal detected by the sensors is very small, therefore it is necessary to amplify it. For that purpose, a specifically designed hybrid card is glued to the back of the sensors, taking care of the shaping and pre-amplification of the signal. The $VA1_3$ chips on

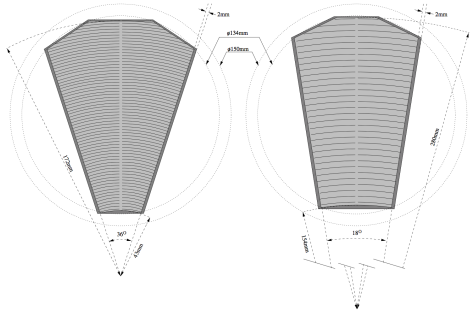


Figure 7.6: Inner (left) and outer (right) sensor types. Each sensor is divided azimuthally into 2 sectors. The dark gray area along the edges are used for bonding pads. Figure taken from[57].

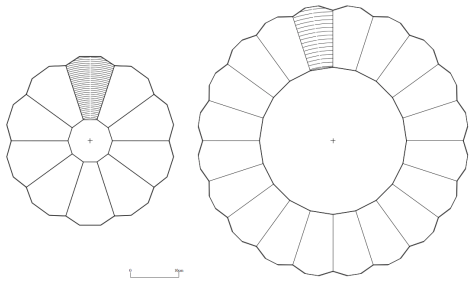


Figure 7.7: Arrangement of inner (left) and outer (right) sensor into rings. Figure taken from[57].

the hybrid cards are the responsible for amplification, and each one takes care of 128 strips.

The analog signal amplified and shaped by the $VA1_3$'s has to be converted to a digital signal in order to be stored in a computer. The FMD Digitiser (FMDD) is responsible for this conversion, making use of the ALTRO chip, which was developed by the ALICE TPC. The digital signal is sent to the read-out controller unit (RCU), where the read-out of the ALTRO's and triggers are handled, and, finally, the signal is pushed from the RCU to the data acquisition system for storage.

7.3 SPD

In section 6.3.1 the basic properties of the SPD were briefly mentioned. Therefore, by now, we already know that it is the innermost layer of the Inner Tracking System, it is used for vertexing and measurement of charged particles in the range $|\eta| < 2.0$. The figure 7.8 shows the position of the SPD in the ITS. Now, we are going to briefly

go through more specific details of the design and electronics of the SPD.

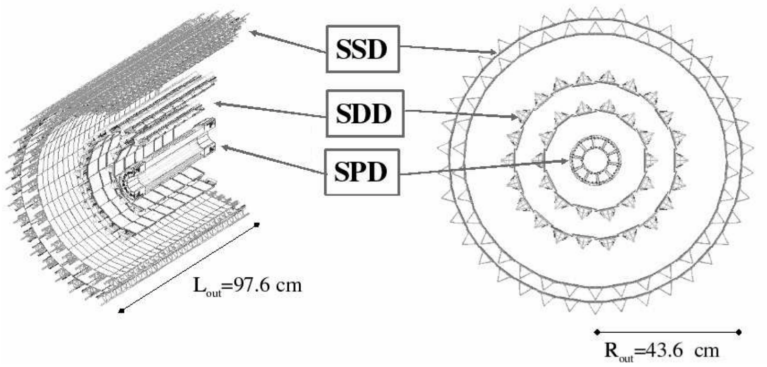


Figure 7.8: The position of the SPD in the ITS. Figure taken from[36].

As the FMD, the SPD also uses silicon as its fundamental material. The basic units of the SPD are ladders, which are made by a silicon pixel sensor matrix of 256 Œ 160 cells bonded to readout chips. The SPD has two barrels and both are subdivided in 10 lightweight carbon-fibre sector. Every 4 ladders form a stave, and every 2 staves make a sector in the inner layer, whereas four staves make an outer layer sector. In total, the SPD has 60 staves, 240 ladders and 1200 readout chips. The properties of the SPD layers are summarized in Table 7.2.

Layer	Radius (cm)	± z (cm)	# Area (m^2)	Channels
1	3.9	14.1	0.07	3276800
2	7.6	14.1	0.14	6553600

Table 7.2: Properties of the SPD layers[36].

Unlike FMD, the SPD does not register the energy deposition information when a charged particle crosses it. The SPD is a very finely segmented detector, which uses pixels, therefore it records only a hit in the pixel when the signal crosses the threshold. Nevertheless, a particle might fire more then one pixel, and the groups of hits in adjacent pixels are called clusters. The detector detects clusters in the inner and outer layers, and gathering the cluster plus the primary vertex information, it is possible to form *tracklets*, which point back to the interaction point. The definition of tracklets is that they are particles originating directly from the interaction point, meaning that they have very low contamination from other particles created by interaction with material and decay processes.



Chapter 8

Data Processing

In this work, specific software frameworks are used. This chapter is dedicated to present briefly each of them, including the framework used by ALICE to produce simulations. This chapter will give an overview of the most important analysis tools and, in the next chapter, the analysis methods are presented.

8.1 ROOT, AliRoot and AliPhysics

The most used tool is the ROOT[26] framework, which was created at CERN in order to be used for data processing purposes. ROOT is widely used within high-energy physics. It is written in object oriented language, C++, and supports all the steps of any analysis, from saving and accessing data, to data mining and visualization. Mathematical and statistical tools are also available, making the data analysis in this work completely possible through this framework.

AliRoot is the ALICE off-line project, which uses ROOT as basis. It contains a full implementation of the ALICE detector geometry, where it is possible to perform simulation and reconstruction. The analysis code is found under the AliPhysics framework, where every member of ALICE experiment can access and modify analysis codes though the git system.

8.2 Monte Carlo Simulations

Simulations have an important role in high energy physics, having a wide range of applications, e.g. design and optimization of detectors, background and efficiencies estimations, etc. The simulations are very useful as a theoretical basis to compare the data and obtain a better physical understanding of the results.

8.2.1 Generators

In Monte Carlo simulations, the event generator is the responsible for creating simulated particles according to a theoretical model of the collision. The transport code then propagates the simulated particles through a model of the detectors. In this work, two different event generators were used and they are described in section 5.4.

The generators output a stack of particles, which is used as an input to the transport code, like GEANT (GEometry And Tracking). During the simulation, for every event, one particle is taken from the top of the stack and is propagated through the detector. This particle might decay into new particles as a results of its interaction with the material. The decay products are therefore moved back to the stack and removed once their turn to be propagated comes. Every time a particle goes through some specific sub-detector, a routine is activated in order to calculate how the particle interacts with the active detector and store the hit information.

8.2.2 Digitisation

Just generating the events and propagating them through GEANT is not enough in order to have a simulation as close as possible to regular physics data. It is now necessary to process the generated hits and convert it into digits.

As known, FMD deals with the energy loss of particles passing through the strips. Therefore, the digitisation for FMD uses the integrated energy loss Δ_i for all the simulated particles that crossed the strip i in a given event to perform the conversion to ADC counts, or digits. In real data collected by the detector, some noise is also detected, thus it is also important to insert random noise into the simulation. The digits are obtained as

$$d_i = p_i + x + g_i \Delta_i C \tag{8.1}$$

where p_i is the pedestal mean and g_i is the gain of the strip, x is a random value obtained from a Gaussian, where the standard deviation is equal to the noise of the strip i. The constant C is a factor used to describe the relation between gain and digits. In order to understand this expression, one should define the pedestal and gain calibrations. Pedestal consists of control signals obtained when there are no particles crossing the detector. Electronic noise contribute to the pedestal signal, as well as the voltage kept in every trip in order to measure the particles. For FMD, the pedestal is obtained in separated calibration runs, where some events are measured when there is no beam in the LHC. The gain calibration is done using a pulser which is supposed to inject a certain amount of charge into FMD. The input charge is thus compared to the output data, which is then used to obtain the relative gain.

After obtaining the digits, they are encoded into ALTRO format, which is the same format as collected physics data. That ensures that both simulation and data are reconstructed in the same way.

8.3 Reconstruction

In this step, the raw data obtained from the detectors and converted into digits are reconstructed into data with the necessary information for analysis. Figure shows the reconstruction steps and, as one can see and as described previously, data and simulation end up in the same format, therefore the reconstruction algorithm for both situations is the same. This step varies according to the sub-detector and, here, we describe how it is done for FMD and summarize the process for SPD.

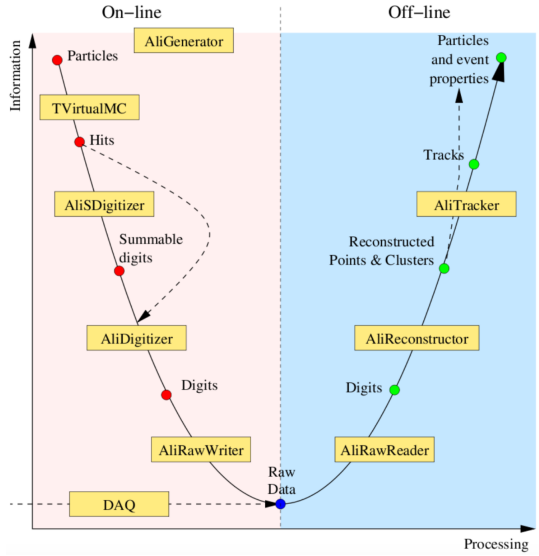


Figure 8.1: Reconstruction chain. Figure from [57].

The FMD deals with multiplicity of charged particles, therefore, in the process of reconstruction, the number of particles hitting the detector is estimated, independently if it comes from a primary or a secondary source. The starting point is the digits. For each of them, it is calculated the pedestal subtracted ADC counts, which contains the noise from the pedestal distribution. If this value is smaller than the noise of the strip by a factor f, this strip is considered empty, this cut is expressed as

$$c_i^{'} = \begin{cases} c_i - p_i, & \text{for } \; c_i > p_i + fn_i \\ 0, & \text{otherwise} \end{cases} \tag{8.2}$$

$c_i^{'}$ is the ADC value after the cut. f denotes how many standard deviations of the distribution are suppressed, and the values range from 2 to 4, where f = 3 suppresses around 99.4% since the distribution is approximately Gaussian.

The next step is to apply the gain factor g_i, obtained through the gain calibration, to the output of the previous step and the ADC counts then has to be converted into units of energy that are comparable to the MIP energy. This can be expressed as

$$E_{dep,i} = \frac{c_i^{'}1}{g_i C} E_{MIP} \tag{8.3}$$

$$\frac{E_{dep,i}}{E_{MIP}} = \frac{c_i^{'}1}{g_i C} \tag{8.4}$$

C is a global constant which means the proportionality factor between charge and the energy measured in the sensors. Eq.8.6 express the energy deposition in each strip in units of MIP energy.

The particles very often don't transverse the FMD in a normal angle, which means that particles traveling through a longer path, on average, are going to deposit more energy. The signal has to be corrected for the path length by

$$\frac{E_{dep,i}}{E_{MIP}} = \frac{E_{dep,i}^{'}}{E_{MIP}} \tag{8.5}$$

where

$$\theta = tan^{-1}\left(\frac{r_i}{z_i - v_z}\right), and \quad r_i = \sqrt{x_i^2 + y_i^2} \tag{8.6}$$

The output of the reconstruction for FMD is $\frac{E_{dep,i}}{E_{MIP}}$, the η and the coordinates of each strip. This information is stored in the Event Summary Data (ESD), which is the input used for the physics analysis.

Particles can deposit energy in many neighboring strips of the SPD layers, therefore the method of reconstruction of this detector is based on counting what is called "clusters", or group of pixels hit by particles. When clusters of two layers are combined with the information of the primary vertex is then possible to obtain what is called tracklets. Figure shows the process of tracklets obtaintion.

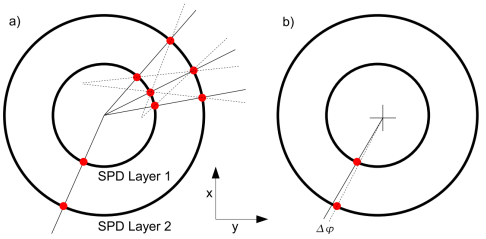


Figure 8.2: Left: The primary-vertex finding method. The vertex is found in the point where there are more intersections. Right: To define a tracklet, $\Delta\phi$ between the line defined by the vertex and the first layer cluster and the line defined by the second layer cluster should be under a threshold. Figure taken from [53].

The output of reconstruction is also stored in ESDs, which is the type of file that contains all the information originating from reconstruction. Another type of file of interest is used to store only the needed information of some very specific type of

analysis, making the size of the file smaller. These files are known as Analysis Object Data (AOD). The conversion from ESDs to AODs is made through the ALICE analysis framework, running the analysis trains, which contains the analysis manager responsible for handling the input/output and looping over the data. A visual representation of the analysis train of ALICE is shown in Figure 8.3

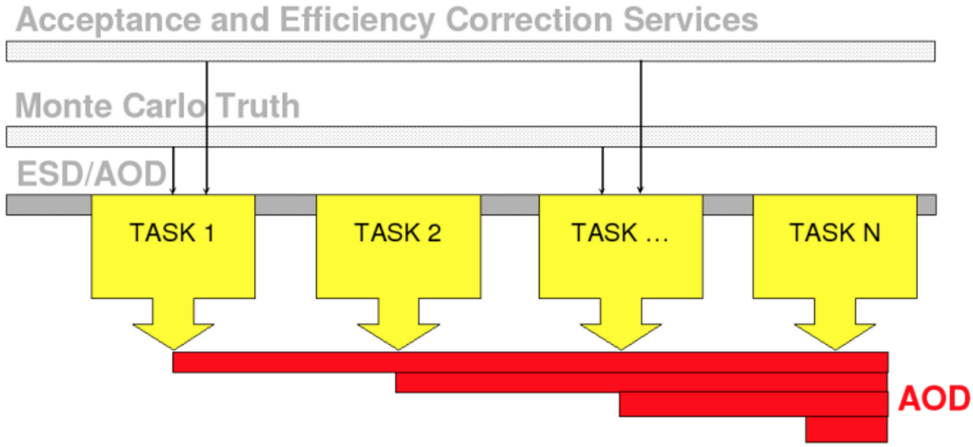


Figure 8.3: Schematic drawingf of the analysis train structure. Figure taken from [12].

8.4 GRID and AliEn

The amount of data collected by LHC is enormous - collisions happen around 1 billion times per second, which generates one petabyte of data per second! In 2017, CERN reached 200 petabytes of data stored in tape libraries at the Data Center. Taking into consideration this gigantic amount of data, a lot of resources are needed in order to store the data, redistribute it and, of course, analyze it.

The Worldwide LHC Computing Grid was created in 2002 to serve that purpose, integrating many regional computing facilities from universities collaborating on the LHC into a single computing Grid. Now, the Grid links computers across more than 40 countries and it is the world's largest computing grid.

The Grid is composed by four Tiers, which are called 0, 1, 2 and 3. The biggest center is located at CERN, being the tier 0, where it has the responsibility of storing the first copy of raw data and performing the first pass of reconstruction, passing the reconstructed output to Tiers 1. The Tiers 1 are computer centers capable of storing LHC data, which means that they have to be large centers. There are 13 computer centers as Tier 1, where their main task is to provide round-the-clock support for the Grid and also storing part of the raw and reconstructed data. Tiers 1 distribute data to Tiers 2 and store part of simulated data produced by them. The Tiers

2 are therefore smaller centers located in universities or scientific institutes. This structured can be accessed through the Tier 3, which are local devices - university clusters or even a personal PC.

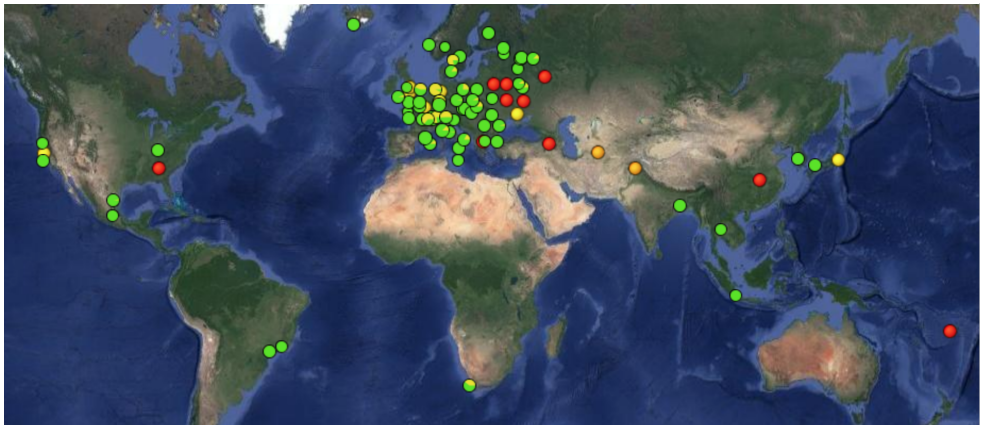
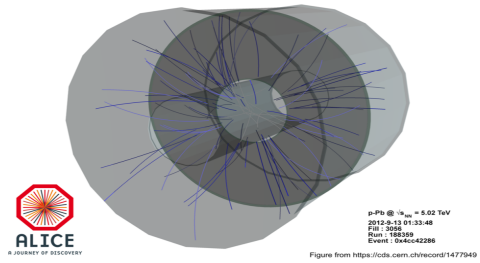


Figure 8.4: This figure shows the AliEn centers through the world. The circles correspond to the current activities of each center. Green means that there are jobs running on that site, yellow means server problems and red means service down or no jobs running. Figure taken from [31].

The system developed by the ALICE Offline Project to offer ALICE members access to the computing resources is called AliEn (ALICE environment)[27]. Through AliEn, users have always access to their data, jobs, package manager, and a resource that divide jobs and share the computing power among different centers. A monitoring service is also implemented using MonALISA (Monitoring A Large InfraStructure Architecture), where the user can see the status of the jobs, Monte Carlo productions and packages installation. The analysis framework is completely integrated to AliEn, which was used in this work.



Chapter 9

Analysis

The previous chapters depicted the theory behind Relativistic collisions, the experiment used to acquire the data, and how the raw data is processed. This chapter is dedicated to the description of the Physics analysis procedure done in this work.

The main steps of the analysis are:

Event Selection This part of the analysis selects the types of events to analyze by trigger, vertex, and centrality.

Particle Counting In this step, the number of charged particles is obtained using the FMD and SPD detectors. The process of converting the energy deposited in the FMD to N_{ch}, the specific cuts applied, and the hit merging algorithm are explained.

Unfolding Secondary particles, particles not originated from the initial collision, contribute considerably to the signal obtained by the detectors, especially for FMD, which covers large rapidities where the angles relative to the beam direction are small. Of great importance to this analysis, the unfolding procedure corrects the raw distributions for effects of secondary particles, acceptance gaps, and resolution, in order to obtain the "true" distribution.

Trigger Correction The efficiency of the used ALICE triggers may drop at low multiplicity. This step is dedicated to the corrections of the trigger and vertex evaluation inefficiencies in the low multiplicity bins.

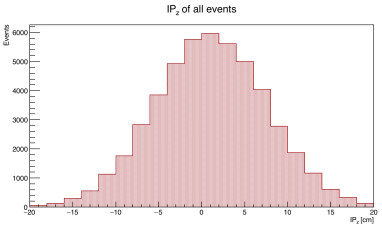


Figure 9.1: The measured z-vertex distribution for p-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV

.

9.1 Event Selection

This section presents the process for trigger, vertex, and centrality selection. The analysis is performed for p-Pb data obtained in 2016 for the center of mass energy of 5.02 TeV.

9.1.1 Trigger Selection

High energy physics experiments rely very much on triggering systems, normally multilevel trigger systems, which make a decision whether the event is worth storing or not. Online triggers define the signals acquired during the running time, assigning specific trigger words to every accepted event. The trigger words contain information on the conditions that were fulfilled for triggering, such as the desired physics of interest, the trigger type and the detectors that were read out. Offline triggers, on the other hand, are defined on the analysis of reconstructed events and can be the exact equivalent to the online triggers or add extra conditions to them.

The trigger choice plays a major role in the analysis, since the triggers are selected according to the type of physics being analyzed. In this work, the online trigger selection is done by the V0AND trigger, which requires a time coincidence between the A-side and the C-side of the V0 detector. The offline trigger of interest is the minimum bias (MB) trigger, also called V0AND, and it is equivalent to the online trigger being used.

During the event selection, one has to take into consideration the beam-gas interactions and the beam halo. Even though the beam pipes are maintained in a ultra-high vacuum condition, around 10^{-7} Pa and lower than 10^{-9} Pa close to the IPs, the system is not perfect, therefore there are still molecules in the beam pipe that may be hit by beam particles. This so-called beam gas can be detected and mistaken as a signal of a real collision.

The beams are formed by bunches of particles, and they are normally characterized by approximately a Gaussian distribution of particles. The beam core is where most of the particles are situated or, in terms of Gaussian distributions, within 3 standard deviations away from the mean (99% of the particles). The region $> 3\sigma$ is considered the beam halo, the tails of the distribution. The beam halo can also

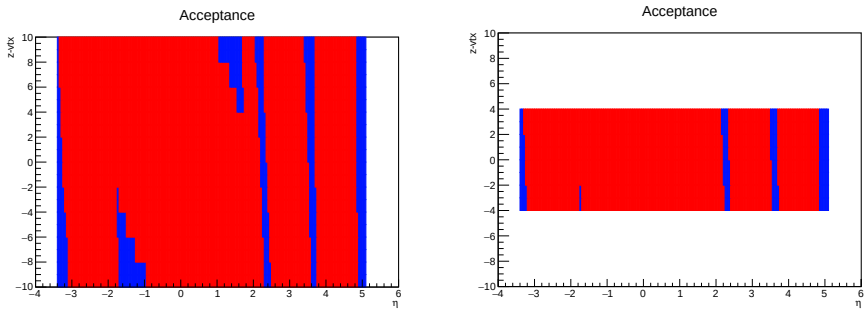


Figure 9.2: Acceptance of FMD and SPD: η-acceptance as a function of v_z vertex. The red areas denote the acceptance, and the blue areas denote no acceptance. Left: Selection in the range $-10 < v_z < +10$cm. Right: Cut to $-4 < v_z < +4$ cm, minimizing the blue areas.

interact with the other bunches or with the beam pipe and produce particles. The beam-halo and beam gas interactions originate through different processes, but their interaction with the beam is triggered in the same way. Demanding the time coincidence of the V0A and V0C signals is an efficient way to reject beam-halo and beam gas interactions.

9.1.2 Vertex Selection

The vertex determination is done by the SPD. As previously discussed, it is possible to deal with most of the effects of the beam gas and beam-halo during the trigger selection. To reduce the effect even more, vertex position conditions can be imposed. The maximum of the vertex distribution lies around the nominal interaction point (z $\sim$ 0). Selecting events in the range $-10 < v_z < +10$cm reduces the contamination from beam gas and beam halo, but a more careful selection has to be performed in this analysis, in order to maximize the η-coverage between FMD and SPD.

This is illustrated in Figure 9.2, where the red areas mark the combined acceptance of FMD and SPD. As it can be seen, it is possible to reduce the acceptance gaps choosing the interval $-4 < v_z < +4$ cm. The blue areas mark regions without acceptance, and this is corrected for later through the unfolding procedure, explained in the section 4.3.

It is possible to obtain the vertex distribution in the xy-plane, as shown in the figure 9.3. The measurements show that there is a shift in the vertex position from (0,0). This offset affects minimally the η of the particles. Furthermore the MC simulations used to perform the unfolding are anchored to specific runs, and have the same offset incorporated.

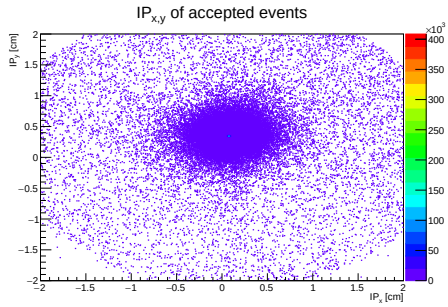


Figure 9.3: Measured xy-vertex distribution of $\sqrt{s_{NN}} = 5.02$ TeV.

9.1.3 Centrality Selection

It's very common that data from relativistic high-energy collisions is divided into different centrality classes. The centrality of an event can be determined according to its impact parameter b, which is the transverse distance between the center of one of the projectiles to the center of the other projectile, as shown in Figure 9.4. This means that very central collisions have small impact parameter, whereas peripheral collisions have large impact parameter. The amount of matter overlapping in a collision determines the number of produced particles.

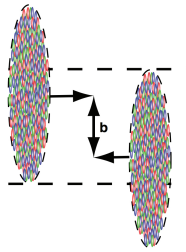


Figure 9.4: Representation of the impact parameter in a A-A collision[80].

However, the impact parameter is not a directly measurable parameter in an experiment, therefore theoretical models must be used to estimate geometrical quantities from experimental data.The Glauber model is then used for this purpose. The impact parameter can be related to the number of participants in a collision, N_{part}, and the number of binary collisions, N_{coll}. Through this model, one can obtain the probability distribution P_ν, where ν can be either N_{coll} or N_{part}. The variable ν is also not directly measurable, but it can be estimated through an observable, called centrality estimator E, through the conditional probability $P(E|\nu)$[6].

The determination of ν requires an observable which depends monotonically on ν. This work focuses on p-Pb collisions, where there is a very loose correlation between the impact parameter and the variable ν. Even though we still call it centrality

estimation, N_{coll} and N_{part} are the important variables. The typical observables used are, e.g. zero-degree spectators, slow-nucleons and multiplicity in the rapidity plateau. Centrality estimators using forward neutron energy and central multiplicity give agreeable results for N_{part} and N_{part}, and they are better explained in the following paragraphs[6].

Charged Particle Distribution Estimators

In this type of estimator, it is possible to obtain the geometry of the collision by fitting the distribution of N_{coll} obtained from the Glauber Model convolved with a negative binomial distribution to the measured multiplicity distribution in a certain η. The V0A estimator, used in this work, uses the amplitude measured by the V0 detector on the A-side, in the interval $2.8 < \eta < 5.1$. The centrality classes are then defined as percentiles of the multiplicity, as shown in Figure 9.5 [6].

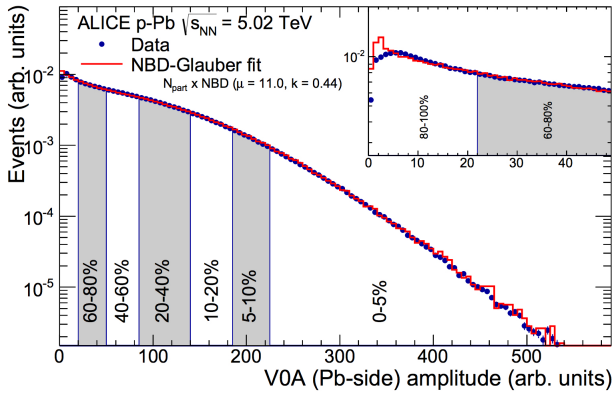


Figure 9.5: Distribution of V0A multiplicity, Glauber fits and centrality classes[6].

A similar procedure is employed for CL1 estimator, which measures the number of clusters in the outer layer of SPD, and V0M, which is a sum of the amplitudes measured by V0-A and V0-C. Even though multiplicity estimators are used, it is important to be aware that in p-Pb they impose some biases due to the multiplicity fluctuations. In p-Pb, the multiplicities used to select centrality have similar magnitude as the fluctuations, which is not the case for centrality selection in Pb-Pb. The V0A and V0M estimators are used in this work, and are compared with an estimator free of bias, described in the next section.

Zero Degree Energy Estimator

This estimator uses the energy measured by the zero degree calorimeter(ZDC) to estimate the centrality of an event, detecting the "slow" nucleons produced. The estimation of centrality can be done since the multiplicity of slow nucleons is monotonically related to N_{coll}. To translate the energy measured by ZDC into N_{coll} a model for the slow nucleon emission (SNM) is used[6].

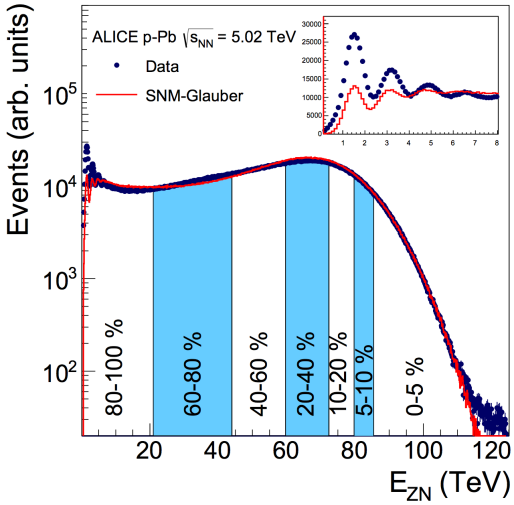


Figure 9.6: Distribution of the neutron energy measured in the Pb-remnant side and the centrality classes obtained by the method described in[6].

Figure 9.6 shows the distribution of the neutron energy measured in the ZN calorimeter and a SNM, coupled to the probability of N_{coll} calculated by the Glauber Model, fit to the data. The centrality of the event is defined dividing the spectrum in percentiles of the hadronic cross-section[6]. The ZNA is also used in this work as a bias-free estimator.

9.2 Particle Counting

To get the multiplicity distribution, one must count the particles transversing the detector. This processes encompasses many steps, starting from the energy deposition in each channel to the acquisition of the final number of charged particles. Each step is going to be described in the following sections.

9.2.1 Energy Distributions

As previously mentioned, in section 7.1.2, the energy loss distribution of particles traversing the strips of the FMD follows approximately a Landau distribution. In the reconstruction step, a near multiplicity is produced for each of the 51200 strips, which are stored in ESDs objects. Figure 9.7 shows the energy loss distributions obtained in every FMD ring for p-Pb collisions at 5.02 TeV. The distributions obtained from ESDs can be divided in 3 regions, as highlighted in Figure 9.7:

- The low energy loss region, where no particles deposited energy in the detector strip, shows a peak originated from pedestal remnants and noise accumulated during readout [36]

- Signals between the pedestal peak and the Landau peak, which originate from particles traversing multiple strips, hence sharing signal, which is explained in more details in section 9.2.2
- The Landau-shaped distribution. A Landau distribution convoluted with a Gaussian distribution is fitted to the energy loss distribution, which describes well the data.

According to the energy distribution, as one can see in Figure 9.7, some cuts have to be applied in order to remove the pedestal and shared signal. The hit sharing cuts are described in the next section.

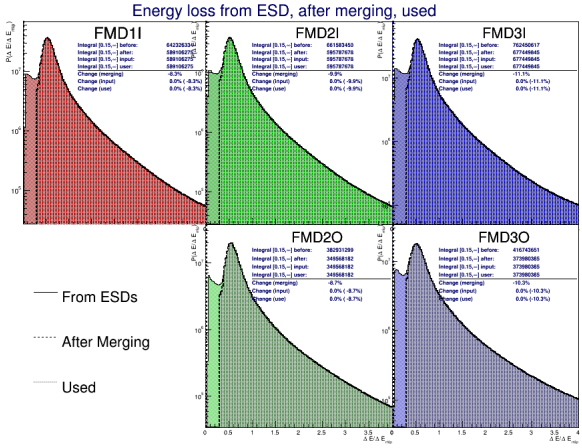


Figure 9.7: Energy loss distributions. The light colors represent the distributions obtained in the reconstruction, while the dark portion is obtained after the hit sharing algorithm is applied.

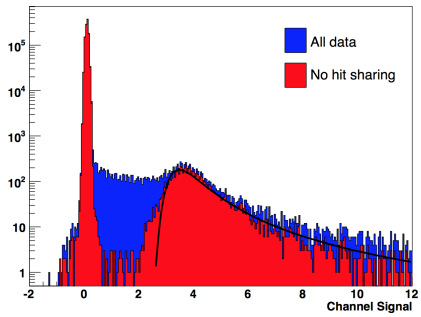


Figure 9.8: Signals from a single strip from FMD. The blue distribution includes all the signals detected in that strip, whereas the red one shows the distribution after removing the shared hits, a large set of signals that have energies between the pedestal peak and the maximum of the Landau distribution [36].

9.2.2 Hit Sharing

A particle traversing the FMD can, from time to time, hit and deposit energy in more than one strip. Consequently, one individual strip doesn't detect the full energy of the particle, whereas the neighboring strip detects the remaining of the signal. This concept is called *hit sharing*. The signal can be shared with up to three strips, even though it is a very rare event - a primary particle, for instance, will never hit more than two strips. The amount of energy detected by each strip depends on the distance Δr from the point at which the particle enter the strip to the boundary, however, this distance is not known, meaning that it is not possible to determine the shared energy a priori.[57]

To evaluate correctly the amount of particles crossing the detector, shared signals must be detected as such and merged into single signals. Even though Δr is not measurable, shared hits are possible to be spotted by the energy loss distribution, since it is expected that they have a signal below the Landau peak, Δ_p. An algorithm is used to loop over all the strips and verify pairs of strips that have signals less than Δ_p, merging them into one strip. The signals are added and attributed to the strip with the largest signal before merging, whereas the second strip is force to zero. In case of triple hits, the summed signal is attributed to the middle strip and the other two are forced to zero.

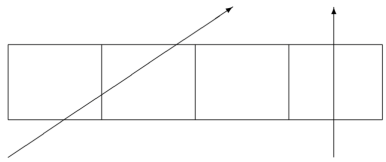


Figure 9.9: Hit sharing schematic. Figure taken from [57].

The algorithm uses two cuts, here called E_{low}, and E_{high}:

- E_{low}: Defined as the lowest energy that can be accepted as a component of a shared signal, being considered remnants of pedestal.
- E_{high}: The highest energy that can be accepted as a component of a shared signal. At some point, the value of the shared signal is similar to the lower part of the peak of the Landau distribution, and, from that point on, the deposits are considered to have originated from two different particles.

The algorithm is illustrated in Figure 9.10. Considering two adjacent strips, if $E_{low} < E_c < E_{high}$ for both, then the signals with be merged into a single signal, and $c = E_c$. If only one of the signals falls into the interval, and the other is above the higher limit, they are still going to be merged. In order to remove remnants of the pedestal peak, the lower limit is set as $0.15.\Delta E/\Delta E_{MIP}$ for the inner rings and $0.18.\Delta E/\Delta E_{MIP}$ for the outer rings.

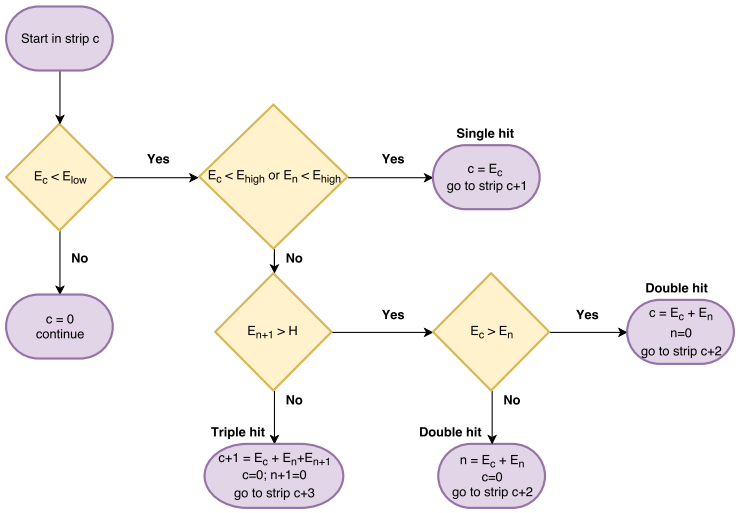


Figure 9.10: The diagram shows the hit merging algorithm.The subscript c indicates current strip and the subscript n indicates next strip. E_{low} and E_{high} are defined in the text.

9.2.3 Particle Counting Methods

Having the defined concepts in the last section, it is possible now to describe the process of particle counting. Here, the ESDs produced in the previous step are inspected, looping over all the strip shared corrected, and calculating the inclusive (primary + secondary) number of charged particles per ring. Two methods are implemented: A method based in Poisson statistics, and a method which used the fits to the energy loss distributions[8].

Poisson Method

To perform the particle counting, high occupancy has to be taken into account. In the case of hit merging, one particle traverses more than one strip, however it is possible that more than one particle traverses the same strip. For FMD, it is not possible to distinguish multiple slow particles depositing energy in a strip from a high energy particle, the particle counting method should contain a correction for that, in order to obtain the correct number of charged particles, N_{ch}.

The true occupancy in a region of FMD can be defined as $\mu = N_{ch}/N_{strips}$, where N_{strips} is the number of strips in the considered region and N_{ch} is the number of charged particles. The measured occupancy is defined as $\mu_{meas} = N_{hits}/N_{strips}$. Here, N_{hits} is the number of strips with $E_{strip} > E_{hit}$, which leads us to the definition of E_{hit}:

- E_{hit}: The threshold above which a merged signal is considered a hit. The low region of the energy loss distribution after merging presents left signals, which

could be caused by shared hits where the adjacent strips were not detected, or by non relevant physical particles.

Making use of Poisson statistics can take into consideration the high occupancy cases. If a region of FMD is considered, assuming that the number of charged particles falling into this region is randomly distributed, then a Poisson distribution is expected. Therefore, one can calculate the probability of finding N_{ch} particles in this region as:

$$P(n) = \frac{\mu^n e^{-\mu}}{n!}, \quad N_{ch} = n \tag{9.1}$$

The probability of having any number of hits in the considered region is the measured occupancy, which can be related to the true occupancy through:

$$\mu_{meas} = 1 - P(0) = 1 - e^{-\mu} \rightarrow \mu = \frac{1}{ln(1 - \mu_{meas})} \tag{9.2}$$

In consequence, to calculate the average number of particles hitting a strip:

$$\begin{aligned} < N_{ch} >= \frac{\sum_{n>0} nP(n)}{\sum_{n>0} nP(n)} &= \frac{e^{-\mu}}{1 - e^{-\mu}} \mu \sum \frac{\mu^n}{n!} \\ &= \frac{e^{-\mu}}{1 - e^{-\mu}} \mu e^{\mu} = \frac{\mu}{1 - e^{-\mu}} \end{aligned} \tag{9.3}$$

This is applied to all the regions of FMD, each one containing 256 strips. The final number of charged particles in each region is then $N_{ch} = N_{hits}. < N_{ch} >$.

Energy Fit Method

This method uses the weighted energy loss fits at the shared corrected energy loss of the strip t to evaluate the charged particle multiplicity for t. A single particle energy loss distribution distribution is given by

$$f(x; \Delta_p, \xi, \sigma') = \frac{1}{\sigma' \sqrt{2\pi}} \int_{-\infty}^{+\infty} dx' f'_L(x', \Delta_p, \xi) e^{\frac{-(x-x')^2}{2\sigma'^2}} \tag{9.4}$$

which is a Landau distribution, f'_L, convoluted with a Gaussian distribution, where Δ_p is the most probable value of energy loss, ξ is the width of the Landau distribution, $\sigma'^2 = \sigma^2 + \sigma_n^2$, with σ as the variance of the Gaussian distribution and σ_n as a parameter which models the noise in the detector.

For i particles, the equation 9.4 becomes

$$f(x; \Delta_p, \xi, \sigma') = f(x; \Delta_{p,i}, \xi_i, \sigma'_i), \tag{9.5}$$

where $\Delta_{p,i} = i[\Delta_p + \xi log(i)]$, $\xi_i = i\xi$, $\sigma_i = \sqrt{i}\sigma$, and $\sigma'^2_i = \sigma^2 + \sigma_i^2$.

The full energy loss distribution from 1+2+...+n particles is then written as

$$f_N(x; \Delta_p, \xi, \sigma', \mathbf{a}) = \sum_{i=1}^{N} a_i f_i(x; \Delta_p, \xi_i, \sigma'_i, a) \tag{9.6}$$

with $f_i(x; \Delta_p, \xi_i, \sigma_i', a)$ as the convolution of a Landau distribution with a Gaussian, and $\mathbf{a}$ is the weights vector for each f_i.

The function $Cf_j(x; \Delta_p, \xi_i, \sigma_i', a)$ is then fit to the energy loss distributions for each η bin, and the parameters Δ_p, ξ, σ and $\mathbf{a}$ are stored. To estimate the number of charged particles in each strip t for a given energy loss Δ_t, n_t is calculated

$$n_t = G_{N_{max}}(\Delta_t; \Delta_p, \xi, \sigma', \mathbf{a}) = \frac{\sum_{i=1}^{N_{max}} i a_i f_i(\Delta_t; \Delta_p, \xi, \sigma)}{\sum_{i=1}^{N_{max}} a_i f_i(\Delta_t; \Delta_p, \xi, \sigma)} \tag{9.7}$$

To obtain the final distribution, n_t is then summed for all strips within a (η, ϕ) bin

$$N_{ch}(\eta, \phi) = \sum_{t}^{t \in (\eta, \phi)} n_t \tag{9.8}$$

The energy fit method and the poisson method have a good agreement. Figure 9.11 shows the correlation between the N_ch calculated by both methods in each FMD ring for p-Pb collisions at 5TeV. The poisson method is the simplest of the two and therefore is the method used in this analysis.

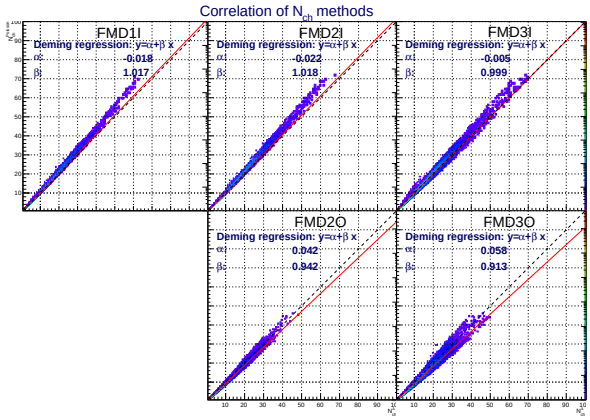


Figure 9.11: Correlation between methods. The energy fits are represented along the x axis and the poisson method along the y axis. The solid red line is the regression, and the dashed x=y line is plotted to guide the eye.

In the process of particle counting, not only primary particles are detected, but also secondary particles, which are considered particles originated by decays or by interaction with materials.

The goal of this work is to obtain the charged-particle multiplicity distribution of primary particles, meaning that the counted secondary particles can't be part of the distribution. The method used to correct the distribution for secondary particles and detector effects is described in the next section.

9.3 Unfolding

It is very common in high energy physics to measure the probability distribution of a certain variable. If all the conditions were ideal, the distributions obtained by direct measurements would be the *true* distribution, meaning that there is nothing else being detected but the variable of interest, and that the detectors have total acceptance and efficiency. That is not usually the case; secondary particles are also detected along with primary particles, the data is represented in histograms, containing statistical fluctuations of random data and the real detectors have limited acceptance and finite resolution, implying that the measured distribution is actually a distortion of the true distribution.

The process used to retrieve the true distribution given the observed measurements is called unfolding. It can be stated mathematically by the Fredholm integral as following [22]:

$$g(y) = \int A(x,y)f(x)dx \tag{9.9}$$

In the equation above, f(x) is the distribution of the true variable x, and g(y) is the measured distribution of the measured variable y. The effects of the detector is represented by A(x,y), which is called the resolution function. The response function maps the response of the detector in the variable y back to the true value x, therefore it has to be known in order to perform the unfolding.

For a physics analysis, the variables are going to be discretized in order to allow a numerical solution. We can rewrite equation 9.9 as a matrix equation:

$$\mathbf{y} = \mathbf{R}\mathbf{x} \tag{9.10}$$

The equation has to be solved for x, the true distribution, given y, and the response function is now the response matrix $\mathbf{R}$. Generally, $\mathbf{R}$ is estimated using Monte Carlo generators to simulate the signal and the detector response. An example of a response matrix, used in this work, is shown in Figure 9.12. To obtain the true distribution, the equation has to be solved by matrix inversion, which is not easily done. The process of unfolding is considered an ill-posed inverse problem because small variations in the measured distribution can cause large changes in the calculated true distribution, and the solution error is also large.

To overcome the inversion problem, there are different unfolding techniques, either using regularization or approximation methods. In this work, the Bayesian Unfolding method is used as the main unfolding method and the $\chi^2 -$ Minimization is used for comparison purposes.

9.3.1 Bayesian Unfolding

Iterative methods are normally preferred when it is necessary to deal with large dimensions. The Bayesian Unfolding is an iterative method based on the well-known

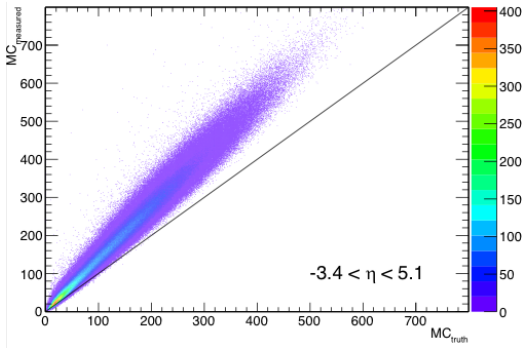


Figure 9.12: Response matrix obtained through a general purpose Monte Carlo production for p-Pb collisions at 5.02 TeV in the range $-3.4 < \eta < +5.1$.

Bayes' Theorem, which uses probability theory to obtain physical quantities from experimental data [43]:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{9.11}$$

where P(A) is the probability of the outcome A, P(B) is the probability of the outcome B, P(A|B) is the conditional probability of obtaining A when B is true and P(B|A) is the conditional probability of obtaining B when A is true.

We can rewrite equation 9.11 in more adequate terms to our case[42]. There are independent *causes* $(C_i, i = 1, 2, ..., n_C)$ that can produce some *effect* E. Therefore, Bayes' Theorem becomes:

$$P(C_i|E) = \frac{P(E|C_i)P_0(C_i)}{\sum_{l=i}^{n_C} P(E|C_l)P_0(C_l)} \tag{9.12}$$

If we have an event, represented by an effect E, the probability that the ith cause originated it $P(C_i|E)$ is proportional to the probability of the cause $P_0(C_i)$ times the probability of the cause actually producing the effect $P(E|C_i)$. It means that $P(C_i|E)$ depends on *prior* knowledge of the probability of the causes $P_0(C_i)$.

$P(C_i|E)$ is often called the smearing matrix $\mathbf{S}$, and $P(E|C_i)$ the response matrix. $P_0(C_i)$ is a *priori* guess that is provided to the algorithm, the prior, and it can be chosen arbitrarily. A good initial guess is the measured distribution itself. The prior affects mainly the number of iterations needed to the unfolding.

We can obtain the number of effects that can be assigned to each cause:

$$\mathbf{n(C_i)} = \sum_{j=1}^{n_E} \mathbf{n(E_j)S_j} \tag{9.13}$$

The number of true events can then obtained:

$$\mathbf{N_{true}} = \sum_{i=1}^{n_C} \mathbf{n(C_i)} \tag{9.14}$$

The final distribution of causes is:

$$\mathbf{P(C_i)} = \frac{\mathbf{n(C_i)}}{\mathbf{N_{true}}} \tag{9.15}$$

The distribution $\mathbf{P(C)}$ lies between $\mathbf{P_0(C)}$ and the true distribution. After every iteration, the process of unfolding uses the $\mathbf{P(C)}$ as the prior to the next iteration, until $\mathbf{P(C)} = \mathbf{P_0(C)}$. The number of iterations i is determined by the user, and the process of unfolding is stopped when i is reached. More details about the Bayesian Unfolding can be found in the Ref.[42].

In Figure 9.13, different number of iterations were used to unfold a MC reconstructed distribution. Because of how the Bayesian unfolding works, using the result of one iteration as the prior to the next, When the number of iterations is too low, the unfolding distribution might be biased towards the measured distributions and when the number of iterations is too high, the unfolded distribution might be biased towards the MC truth and not contain any trace of the measured data. Therefore, it is necessary to evaluate what is the optimal number of iterations.

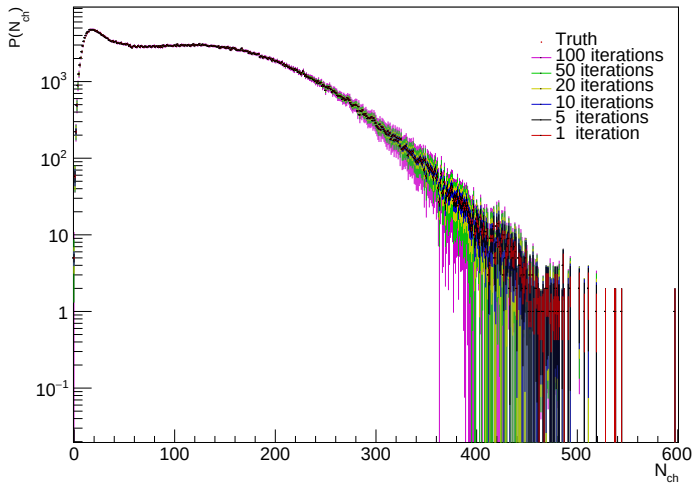


Figure 9.13: Distributions obtained using different number of iterations.

A closure test can evaluate if the method works properly. This test consists on preforming the unfolding procedure on reconstructed Monte Carlo simulations and comparing the result to the Monte Carlo truth distribution. Figure 9.14 shows the result using 100 iterations and Figure 9.15 show the results using 10 iterations.

The residuals show the difference between the unfolded distribution and the truth distribution and, as we repeat this same test for a set of different number of iterations, it is possible to see that 10 iterations give the best results.

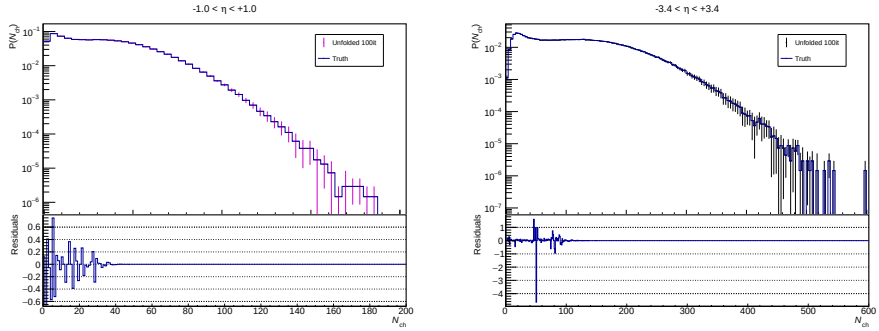


Figure 9.14: Closure test using 100 iterations, Left: Unfolding of MC reconstructed in a narrow η range. Right: Unfolding of MC reconstructed in a wide η range.

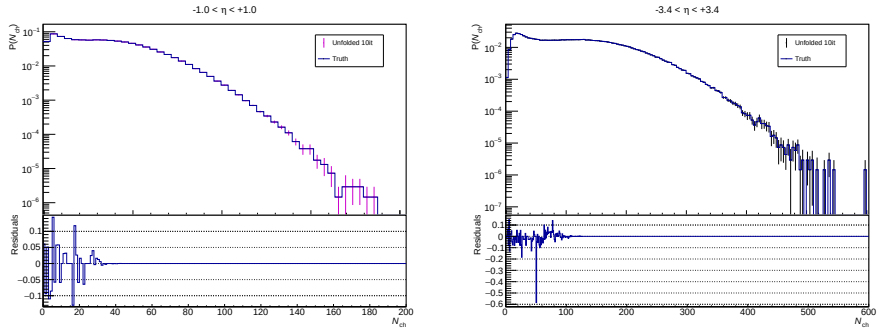


Figure 9.15: Closure test using 10 iterations, Left: Unfolding of MC reconstructed in a narrow η range. Right: Unfolding of MC reconstructed in a wide η range.

9.3.2 χ^2 – Minimization

This method, as the name suggests, uses the the χ^2 function to evaluate which spectrum best describes the measured distribution. The χ^2 function used in this case is the following

$$\hat{\chi}^2(U) = \sum_m \left(\frac{x_m - \sum_t R_{mt} U_t}{e_m} \right)^2, \tag{9.16}$$

where x_m is the measured points, R_{mt} is the response matrix, U_t is the corresponding unfolded points and e_m is the error on the measured. Combining 9.16 with 9.10 results in $\hat{\chi}^2(y) = 0$.

The χ^2-minimization is a numerical method, which uses the inverse of the response matrix. This has some consequences; as mentioned before, inverting a matrix can amplify the fluctuations in the result, thus not only the true distribution is, in fact, the only spectrum which minimizes the χ^2. Figure 9.16 shows an example of a χ^2-minimization unfolding by just inverting the response matrix and multiplying by the measured spectrum compared to the MC-truth distribution. To deal with this problem, a constraint is added to the $\hat{\chi}^2$ function, which works in favor of a certain shape of the unfolded spectrum. The new expression is given by

$$\chi^2(U) = \hat{\chi}^2(U) + wP(U) \tag{9.17}$$

The new term, P(U), called the regularization term, which depends solely on the unfolded spectrum, and w is a weight given to the regularization term. If w increases, χ^2 also increases, therefore the optimal value of w has to be evaluated. There are different regularization functions that can be used and a set of them were used in order to evaluate the best option. Three conventional options, a constant (Pol0), a linear (Pol1) and a logarithmic (Log) function, are given by the following expressions

$$P(U) = \sum_t \left(\frac{U'_t}{U_t}\right)^2 = \sum_t \left(\frac{U_t - U_{t-1}}{U_t}\right)^2, \tag{9.18}$$

$$P(U) = \sum_t \left(\frac{U''_t}{U_t}\right)^2 = \sum_t \left(\frac{U_{t-1} - 2U_t + U_{t+1}}{U_t}\right)^2, \tag{9.19}$$

$$\begin{aligned} P(U) &= P(\hat{U} := \ln U) \\ &= \sum_t \left(\frac{\hat{U}''_t}{\hat{U}_t}\right)^2 = \sum_t \left(\frac{\ln U_{t-1} - 2\ln U_t + \ln U_{t+1}}{\ln U_t}\right)^2, \end{aligned} \tag{9.20}$$

In order to estimate the best regularization and the best weight, some tests were performed. The results of the tests are shown in the next section, where the unfolding performance is analyzed.

9.3.3 Unfolding performance

At first, we can have an idea of how good is the unfolding performance just by looking at the final distribution. However, that's not a correct evaluation - a true performance check should be done in order to infer which parameters returns the optimal solution to our problem.

We can analyze the unfolding performance dividing the multiplicity distribution in 3 different regions, which can vary accordingly to the η range in question.

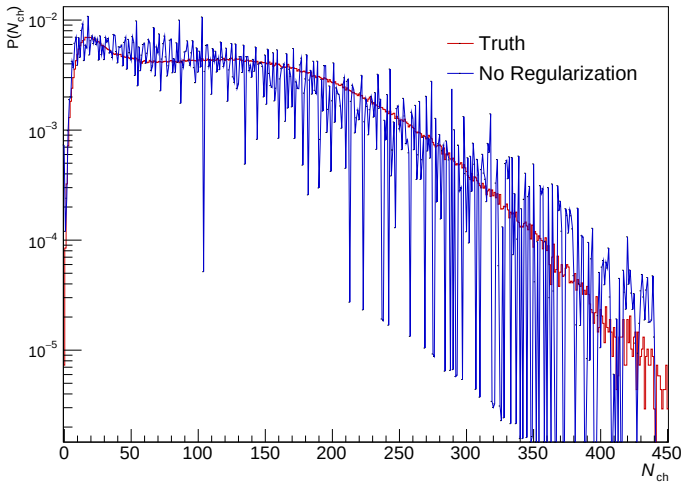


Figure 9.16: χ^2 - minimization unfolding closure test performed without regularization

Peak: The region where the P(N_{ch}) reaches its maximum and starts decreasing again.

Slope: The region where the slope does not change significantly.

Low Statistics: The region where the number of events is low (tail of the distribution).

It is possible to quantify the perfomance through the difference between the input distribution and the unfolded distribution in the following way

$$Q_1 = \frac{1}{t'' - t' + 1} \sum_{t=t'}^{t''} \left| \frac{T_t - U_t}{e_t} \right| \tag{9.21}$$

where T_t is the t^{th} bin of the truth, U_t is the same for the unfolded distribution and e_t is the error on the unfolded. This quantity is normalized accordingly to the degrees of freedom of the analyzed region. The sum is not put in quadrature in order to avoid influence of single-bin deviations.

Figure 9.17 shows the Q analysis for both methods. In the case of χ^2-minimization method, the regularization function with the best performance is the linear, eq. 9.19, and the best weight is found to be 10^5. Overall, the Bayesian method is more stable and easier to regularize, therefore is the main method used in this work.

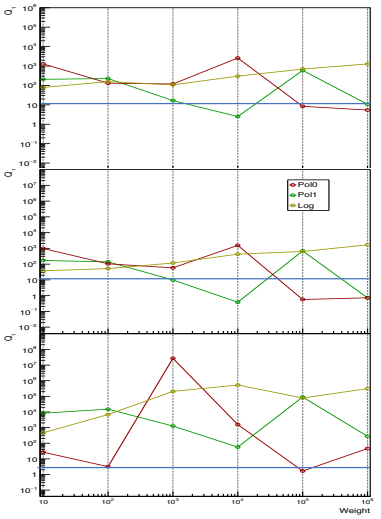


Figure 9.17: Q values for χ^2-minimization method using different weights and regularization functions for each region, in which the top plot corresponds to the peak, the middle plot corresponds to the slope and the bottom plot corresponds to the low statistics. The blue line represents the best value for the Bayesian method.

9.3.4 Problems in the Unfolding procedure

The unfolding procedure relies on the quality of the response matrix. As one can see in figure 9.12, as the multiplicity increases, the statistics decreases, which means that this lack of statistics affects the unfolding, preventing high unfolding quality at the tail region of the distribution. To overcome this problem, a specific type of Monte Carlo simulation is desirable, in which the production is flat in multiplicity (Flat MC), thereby enhancing the statistics for high multiplicities. The Flat MC does not have a physical meaning and it is tuned in such a way that the probability of having N_{ch} is the same up until a defined threshold. This production, however, was never done before for p-Pb collisions in the ALICE collaboration, which required some efforts to investigate different methods to obtain this tuning. The tests are still ongoing, which means that the production is not available for use by the handing-in date of this work. Therefore, it was necessary to obtain the final distributions using the response matrices available, which is going to limit the range of the analysis, decreasing the precision of the results in the tail of the distributions.

9.4 Trigger Correction

The unfolding process can deal with the acceptance gaps and secondary particles, but there are still trigger inefficiencies to take into account, which are mainly important

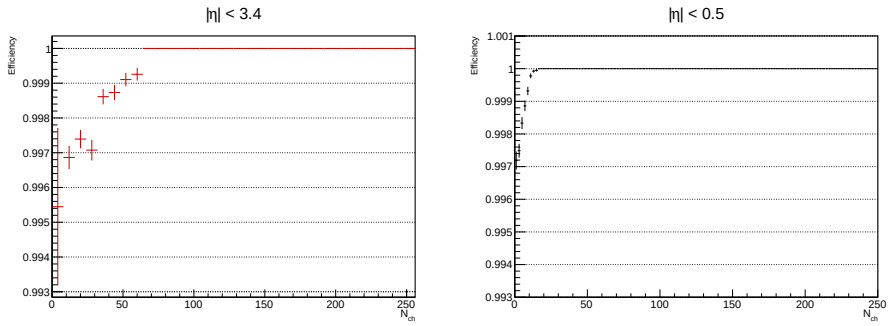


Figure 9.18: Trigger efficiency as a function of N_{ch}. Left: Efficiency calculated in the range $|\eta| < 3.4$. Right: Efficiency calculated in the range $|\eta| < 0.5$.

in the lower bins of the distribution, at the low multiplicity region. This is necessary since, in low multiplicity events, the probability of the triggers being fired not unity, generating inefficiencies. Simulations are done in order to perform the corrections, and the efficiency is calculated in the following way:

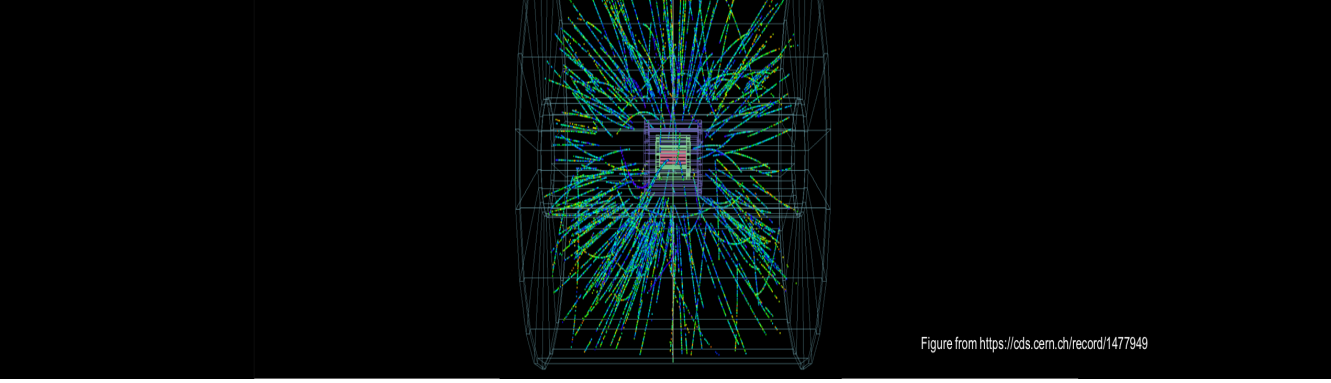
$$\epsilon = \frac{N_{MC}^{\mathrm{trig}\wedge \mathrm{v_z}}}{N_{\mathrm{Rec}}^{\mathrm{trig}\wedge \mathrm{v_z}}} \tag{9.22}$$

The numerator in the equation 9.22 represents "truth" events, obtained by simulation, in the selected trigger and vertex range $-4 < v_x < +4$. In the same way, the denominator represents reconstructed events that fulfilled the same requirements. The efficiency obtained for p-Pb collisions at 5.02 TeV is shown in Figure 9.18 and it is possible to see that the inefficiency in this collisions system is relatively low.

The correction to the distribution is given by:

$$U_{final} = \frac{U_{init}}{\epsilon} \tag{9.23}$$

The corrected multiplicity distribution is thus obtained dividing the unfolded distribution by the calculated efficiency.



Chapter 10

Systematic Uncertainties

In high energy physics, statistical and systematic uncertainties on measurements of physical quantities are generally provided. When we measure the same phenomena repeatedly, the set of observations will present fluctuations and the statistical uncertainties correspond to the range of the variation. The systematic uncertainties, on the other hand, come from the uncertainties associated with the apparatus, assumptions made by the researcher and different models used in the analysis. In this chapter, the systematic uncertainties that arose in the process of measuring multiplicity distributions are presented and discussed.

10.1 Minimum Bias Data

10.1.1 Sharing cuts in the raw distributions

As described before, some cuts are applied to FMD's counting methods. The variation of E_{high} and E_{hit} is one of the systematic checks necessary for measurements done by FMD. The study is made by raising the cuts to 1.5, lowering to 0.95 and obtaining the respective distributions for both cases. The standard procedure is to deduce how much the distributions vary when the cuts are changed and then propagate this difference to the final results. However, as mentioned in 9.3.4, the lack of statistics in the unfolding affects the final distribution, especially in the tail region. This results in statistical fluctuations that could be inserted in the systematic uncertainties and should be avoided. While the flat Monte Carlo simulations are not available, one way of avoiding the contribution of purely statistical fluctuations to the systematic uncertainties is to fit the unfolded distributions and use the difference of the fits as systematic uncertainties.

For distributions in wide eta ranges, it was possible to fit to the distribution double NBDs until a limited range of the distribution and a gaussian was the best fit to the tails. On the other hand, double NBDs can be fitted to the whole distribution in narrow η ranges, however the quality of the fit is quite poor, therefore a combination

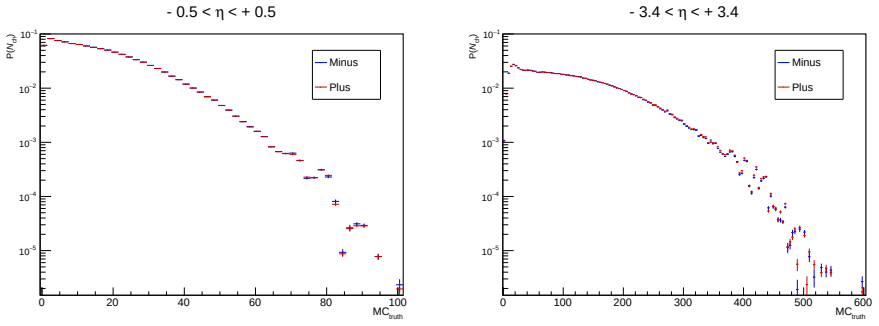


Figure 10.1: Multiplicity distributions obtained by lowering(blue) and increasing(red) the sharing cut. Left: Distributions in the range $|\eta| < 5.0$. Right: Distributions in the range $|\eta| < 3.4$.

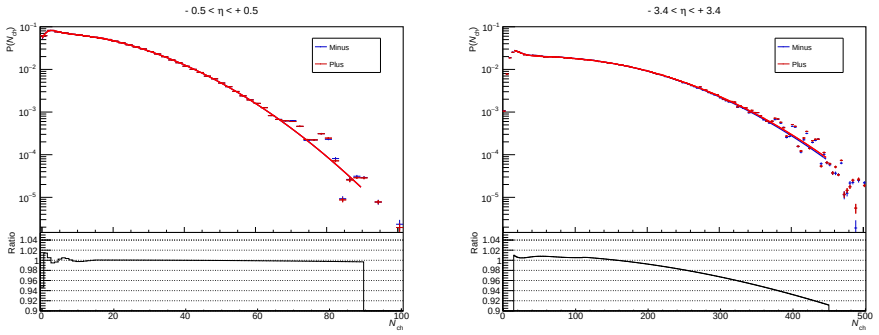


Figure 10.2: Fitted Multiplicity distribution obtained by lowering(blue) and increasing(red) the sharing cut. Left: Distributions in the range $|\eta| < 5.0$. Right: Distributions in the range $|\eta| < 3.4$.

of double NBDs and a gaussian was also used in these η ranges. Figure 10.3 shows the fits to the distributions and their respective ratios, which are used as uncertainties.

10.1.2 Vertex

As mentioned in the analysis chapter, the vertex selection is done to optimize the acceptance of the used detector. The vertex range used, - $4 < v_x < +4$, is thought to be optimal, however other options give different results and the their impact should be included in the final result. In order to evaluate the differences which come from the vertex selection, multiplicity distributions were obtained using different vertex ranges: - $4 < v_x < +1$, -1 $< v_x <$ +4 and -1 $< v_x <$ +1.

This systematic uncertainty is obtained in the same way as described in section 10.1.1, therefore the distributions were fitted and the distribution with different vertex selection were compared to the the standard vertex interval, namely - $4 < v_x <$

+4.

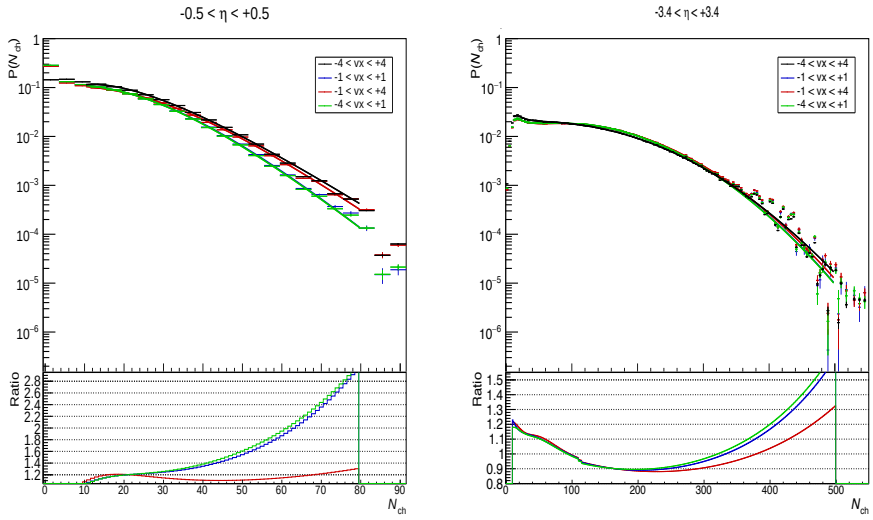


Figure 10.3: Fitted multiplicity distributions using different vertex selection (top) and the ratio of the distributions obtained in different vertex ranges to the distribution obtained in the standard vertex range (-4cm < vx < +4 cm). Left: $|\eta| < 5.0$. Right: Ratio for $|\eta| < 3.4$.

It is possible to note that, even after the unfolding, there is still a large difference when selecting different vertex intervals. In theory, this should not happen since the unfolding should account for acceptance gaps. Therefore, this needs to be investigated deeper in order to understand what is actually happening in the vertex selection and what is the extent of the corrections done through unfolding. In this work, this is still going to be used as systematic uncertainty, however further investigation is ongoing.

10.1.3 Material Budget

The interaction of charged particles with the detector depends on the amount of material within the detector, called the material budget. The accuracy of how this material budget is known and described affects the response matrix, or in other words, affects the final unfolded distribution. The estimation of the material budget systematic uncertainties is done using a special Monte Carlo production, where the material density in every component of ALICE is raised or lowered by a defined percentage. Comparing the results obtained unfolding the raw distributions with response matrices from this MC and the ones obtained unfolding the raw distributions with response matrices from a regular simulation gives the estimate of the systematic error.

The modification in question has to be done to a Monte Carlo production flat in multiplicity to amplify the statistics and not in a regular production, which also limits

the estimation of Material Budget in this work. However, the material budget is the largest source of systematic uncertainties and very important to be included, therefore an estimation of how much this could affect the distributions is done. Regular flat Monte Carlo productions and flat Monte Carlo productions with material budget increased are available for proton-proton collisions at 900 GeV. An estimation of the material budget contribution in multiplicity distributions in p-Pb collisions can be obtained by extrapolating the contribution in pp collisions, if the run conditions are similar enough. One way of analyzing if they are similar is to fit the response matrices and verify if their slope are similar. Figure 10.4 shows the fit to the response matrices and their slopes.

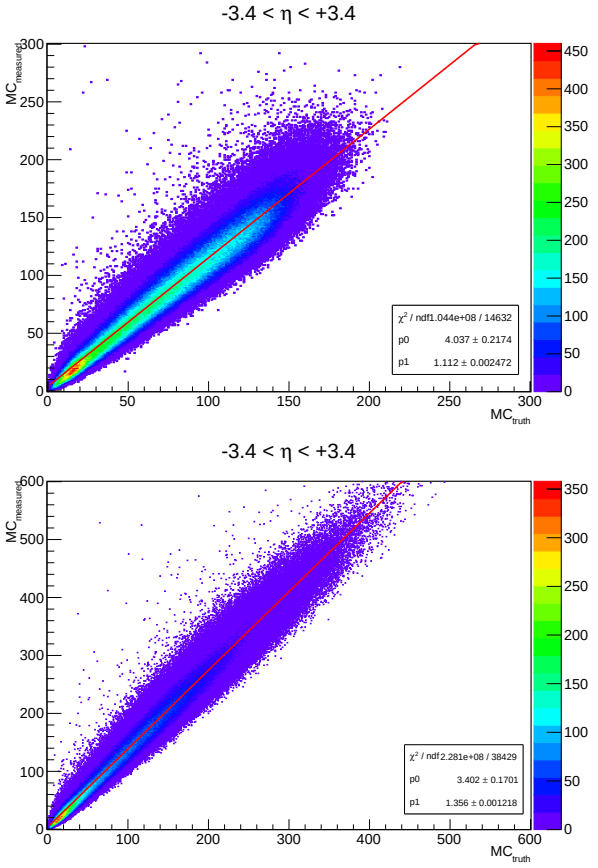


Figure 10.4: Fitted response matrices. Top: Response matrix obtained from a flat Monte Carlo production anchored to a pp collision run at 900GeV. Bottom: Response matrix obtained from a Monte Carlo production anchored to a pPb collision run at 5.02TeV.

The fits to the response matrices indicate that there is a difference of ∼18%

between the two systems in wide η range and of $\sim$15% in narrow η range. This difference is considerable, therefore the material budget estimation using this method is rough. However, since the material budget is the largest systematic contribution, it is still important to have an estimate. To calculate the uncertainty, a multiplicity distribution is calculated unfolding the raw distribution using a response matrix produced by a Monte Carlo simulation flat in multiplicity and another multiplicity distribution is calculated unfolding the raw distribution using a response matrix produced by a flat Monte Carlo simulation with material budget increased in 14%. The uncertainties are given by the difference between the two distributions, as shown in Figure 10.5.

The ratio is fitted and extrapolated to the range of p-Pb collisions and this extrapolation is used as the estimation of the material budget contribution. Figure 10.6 shows the fit to the ratio and the extrapolation.

10.2 Total Systematic Error

The total systematic error is thus obtained summing in quadrature all the described contributions as:

$$\delta_{tot}^2 = \delta_{cuts}^2 + \delta_{material}^2 + \delta_{vertex}^2, \tag{10.1}$$

where $\delta_i = 1 - |\frac{P(Nch,i)}{P(Nch,reference)}|$. Here, $P(N_{ch}, reference)$ is the reference distribution and $P(N_{ch}, i)$ is the distribution obtained varying the parameters. It is possible, then, to plot the total systematic uncertainties for different η region, as shown in Figure 10.7(blue curves).

It is possible to observe in Figure 10.7 that the material budget contribution becomes very relevant in wide eta ranges, where there are more particles and more material in their way, and the vertex selection contribution is more relevant in narrow eta ranges. For all of the systematic uncertainties, an accurate flat Monte Carlo production is desirable, especially in order to estimate precisely the contributions coming from the vertex selection and material budget. It is important to underline that the estimation made in this work is not sufficient for a publication of the results, and therefore is going to go through a refinement after the flat production is obtained.

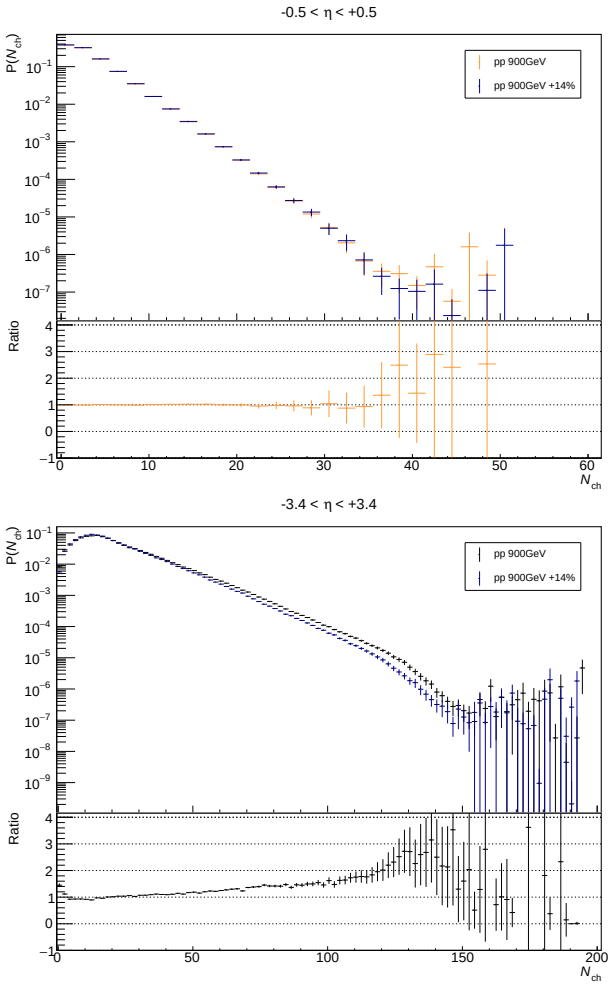


Figure 10.5: The top panels show the distribution obtained unfolding pp raw distributions using a response matrix generated from a regular flat Monte Carlo together with the distributions obtained unfolding the raw distributions using a response matrix generated from a flat Monte Carlo increasing the material budget in 14%. The bottom panels show the ratio of the distributions obtained with increased material budget to the regular distributions. Top: Narrow η range. Bottom: Wide η range.

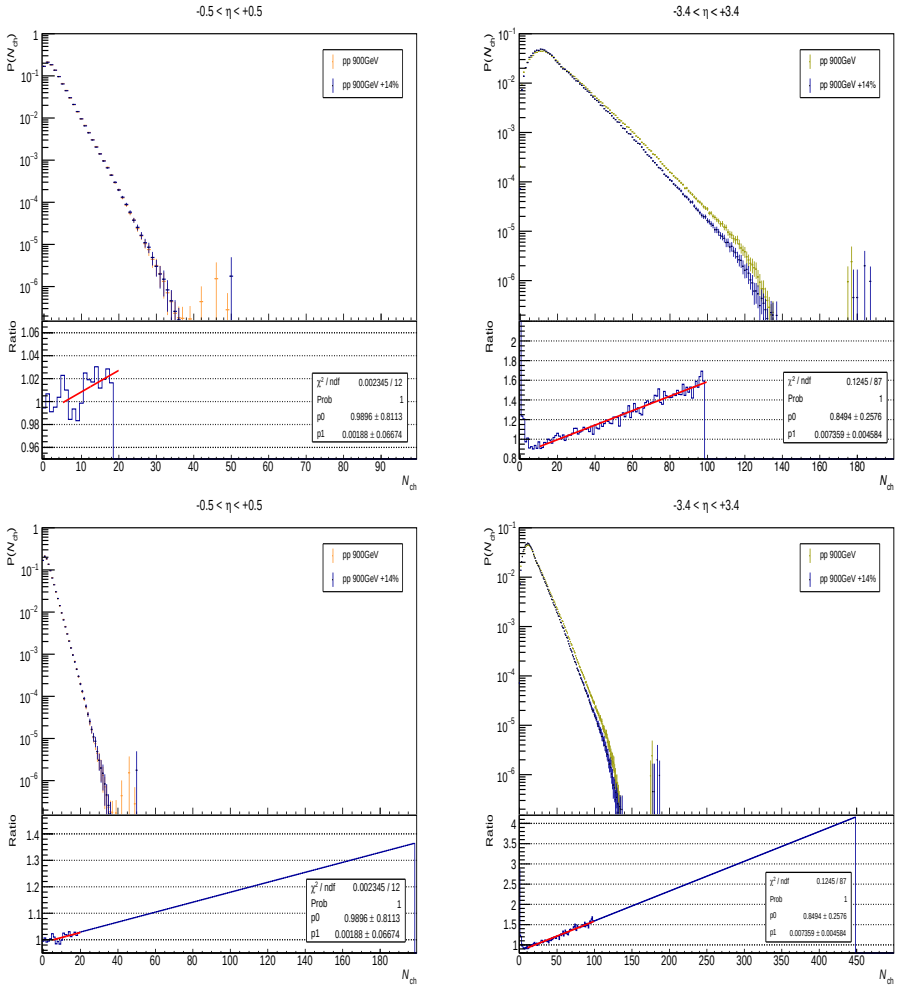


Figure 10.6: Multiplicity distributions calculated for pp collisions at 900 GeV with nominal material budget and material increased by 14%(top panel) and ratio of the multiplicity distributions (bottom panel). Top-left: Fit to the ratio in $\eta < 0.5$. Top-right: Fits to the ratio in $\eta < 3.4$. Bottom-left: Extrapolation to the fit in $\eta < 0.5$. Bottom-right: Extrapolation to the fit in $\eta < 3.4$.

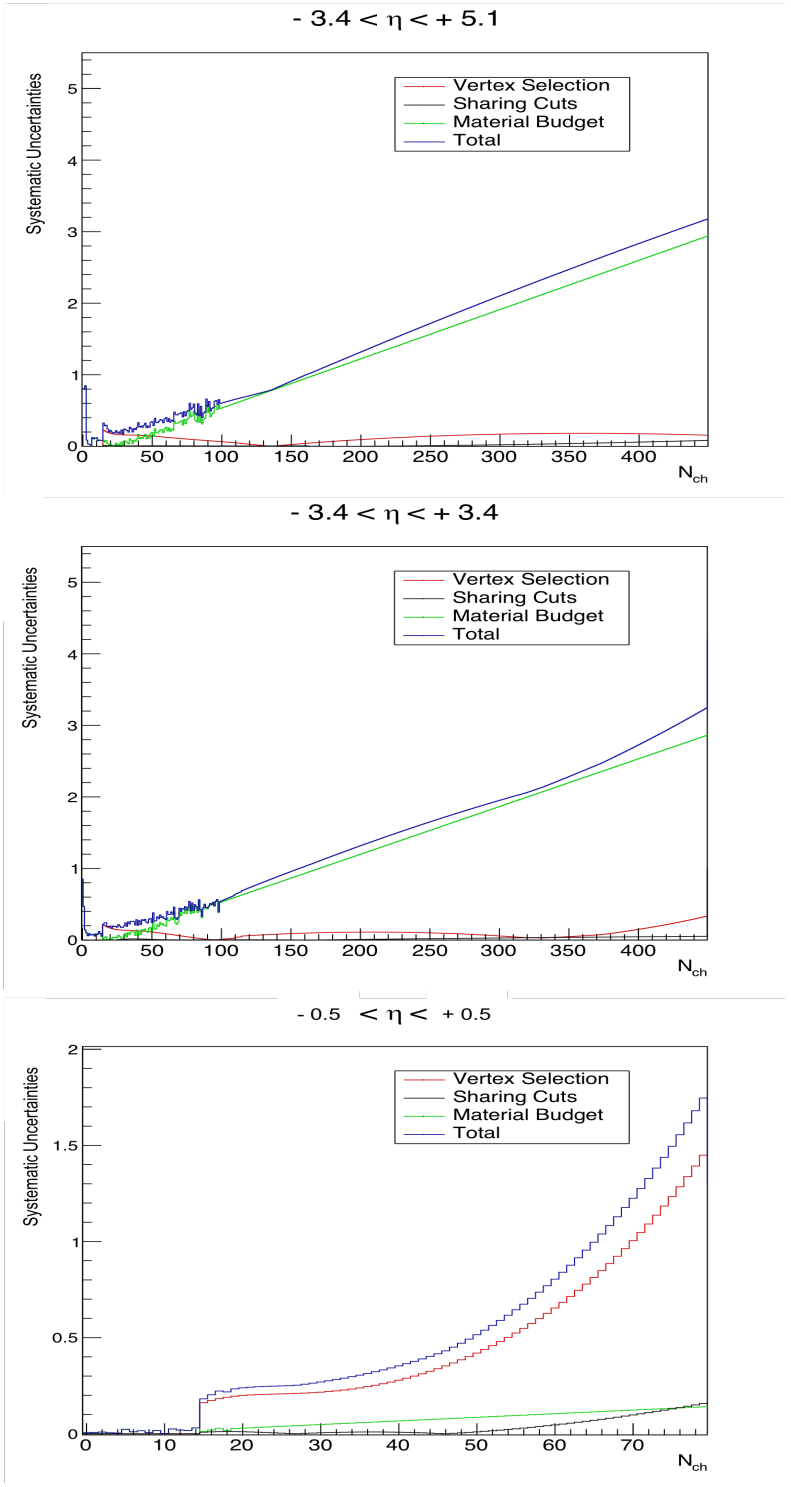
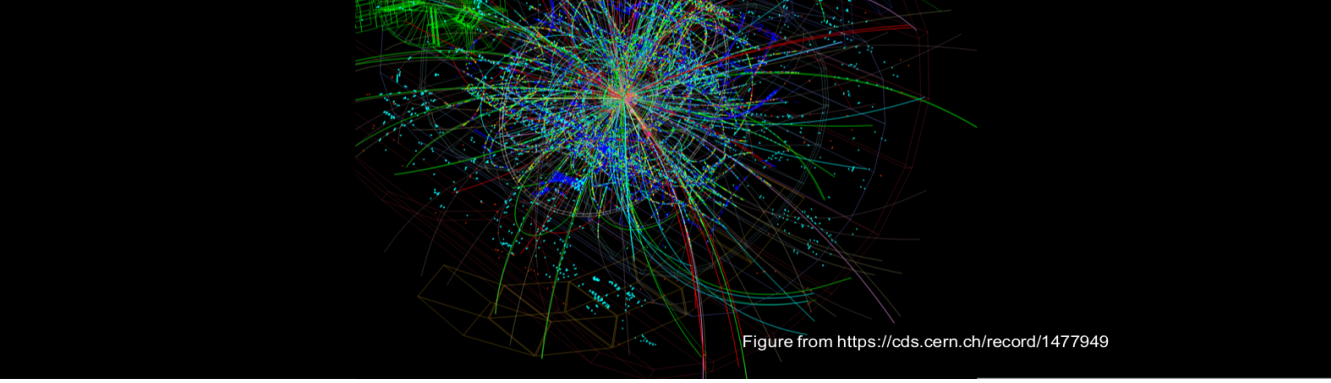


Figure 10.7: Total systematic uncertainties for $-3.4 < \eta < +5.1$, $|\eta| < 3.4$, $|\eta| < 0.5$.



Chapter 11

Results

‘

After the description of the theory, motivation and methods behind this work, it is finally the time to present the results obtained. The multiplicity distributions obtained at minimum bias are presented in the first section, followed by the distributions performing centrality selection in the second section. The third and fourth sections are dedicated to model comparisons, where we combine the experiment with the theory in order to bring a physical understanding of the observed results.

11.1 Minimum Bias Multiplicity Distributions in p-Pb Collisions

In order to obtain the final true multiplicity distribution, there are different analysis steps to be followed, as outlined in Chapter 9. The first distribution obtained after the event selection is called the raw distribution since it has to be corrected. The corrections applied to it are the unfolding procedure, which deals with acceptance gaps, limited resolution and secondary particles, and the trigger inefficiencies. After the corrections, the final result is called the true distribution. Figure 11.1 shows the true charged-particle multiplicity distributions in 8 different eta ranges. The distributions are unfolded and corrected for trigger inefficiencies, using the methods described in Chapter 9 and, here, are multiplied by multiples of 10 in order to show all the distributions in the same plot without overlapping.

The distributions in the η ranges $|\eta| < 0.5$, $|\eta| < 1.0$, $|\eta| < 1.5$ are obtained using solely the SPD, whereas the other five ranges are an overlap of FMD and SPD. It is possible to observe that FMD expands the probing region significantly, being able to detect twice as large multiplicity in comparison to the widest SPD-only region apart from providing large pseudorapidity coverage.

As described in Chapter 10, there are three different sources of systematic uncertainties, which have to be included as uncertainties in the final distribution. The

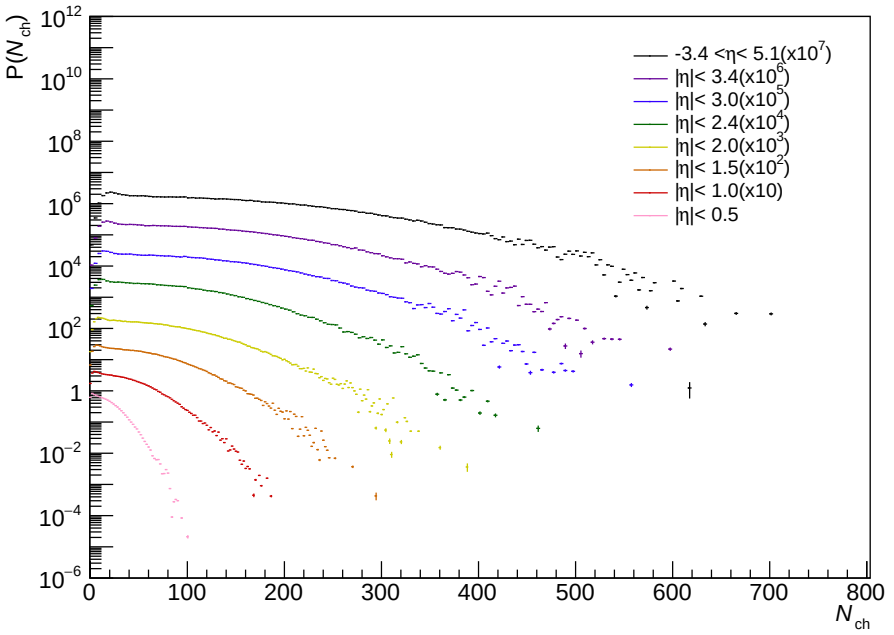


Figure 11.1: Unfolded and trigger-corrected charged-particle multiplicity distributions in 8 different η ranges for p-Pb collisions at $\sqrt{S_{NN}} = 5.02$ TeV.

distributions including systematic uncertainties are shown in Figure 11.2. In comparison to Figure 11.1, the range of the distributions in Figure 11.2 is much shorter. Due to the lack of statistics in the response matrix at high multiplicities, the true distributions also have a limited precision. Therefore, the distributions only have a relevant precision up to a certain N_{ch}, which varies accordingly to the η range. Figure 11.3 shows the selection criteria to cut the distributions, which is where the uncertainties reach approximately the same order of magnitude as the data, putting again into evidence the importance of having a flat high statistics Monte Carlo production to perform the unfolding and avoid information loss.

11.1.1 Comparison to Models

In this thesis, the charged-particle multiplicity distributions are compared to two Monte Carlo models, EPOS LHC and DPMJET, which are based on two different theories, both described in Section 5.4 and to a model based on the CGC framework.

In the EPOS LHC framework, the theory is based on the core-corona concept and treats parton evolution as "parton ladders", whereas DPMJET is based on dual parton model and deals with individual collisions using PHOJET. The differences in

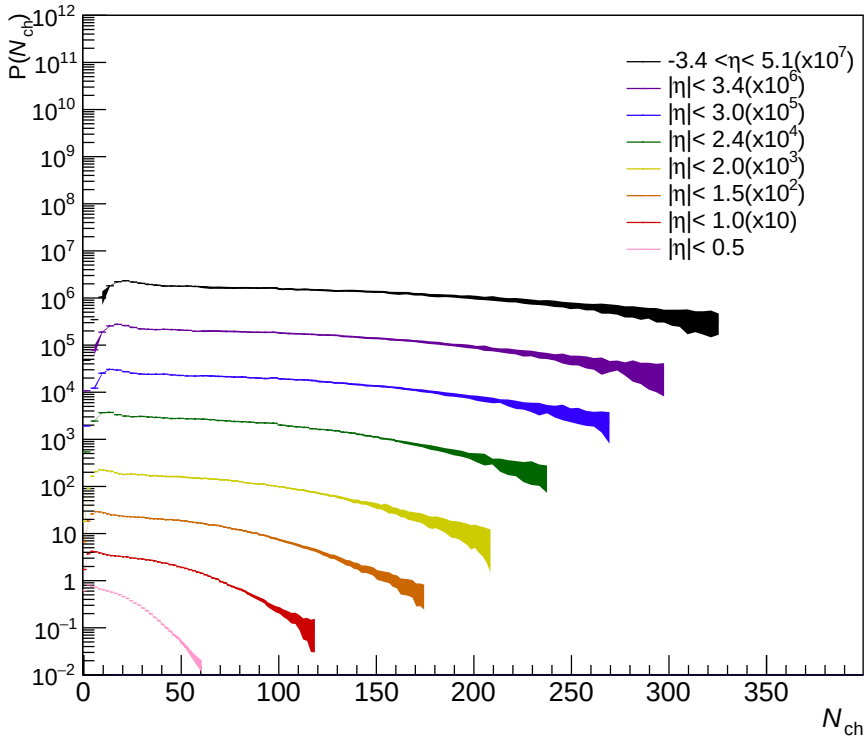


Figure 11.2: Unfolded and trigger-corrected charged-particle multiplicity distributions in 8 different η ranges including systematic uncertainties from Figure 11.1 for MB p-Pb collisions at $\sqrt{S_{NN}} = 5.02$ TeV.

the theory is reflected in the shape of the distributions, as shown in Figure 11.4

EPOS LHC presents two prominent regions, which can be interpreted as soft and semi-hard regions. DPMJET does not show the same behavior, which indicates that it is more dominated by semi-hard events, reflecting the wider tails. In Figure 11.5, the comparison of the measurements to the Monte Carlo models is shown, including the ratios $\frac{P(N_{ch},MC)}{P(N_{ch},Data}$ for five different η ranges.

The differences in the firsts bins can, as mentioned, be attributed to the proportion of soft and semi-hard effects in each model. These differences can be seen with more emphasis in Figure 11.6 for a wide η range. EPOS LHC seems to describe better the distributions in the first bins, whereas the tails are closer to the DPMJET description, indicating that the distributions might have a considerably higher contribution of soft events in the low multiplicity region and of semi-hard events in the high multiplicity region.

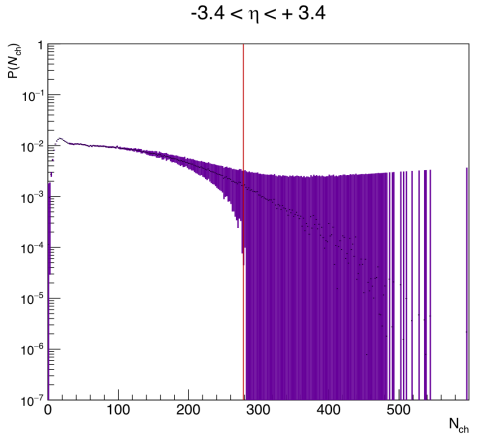


Figure 11.3: Unfolded and trigger-corrected charged-particle multiplicity distributions in $|\eta| < 3.4$ for MB p-Pb collisions at $\sqrt{S_{NN}} = 5.02$ TeV. The threshold where the distributions are cut is defined accordingly to the errors.

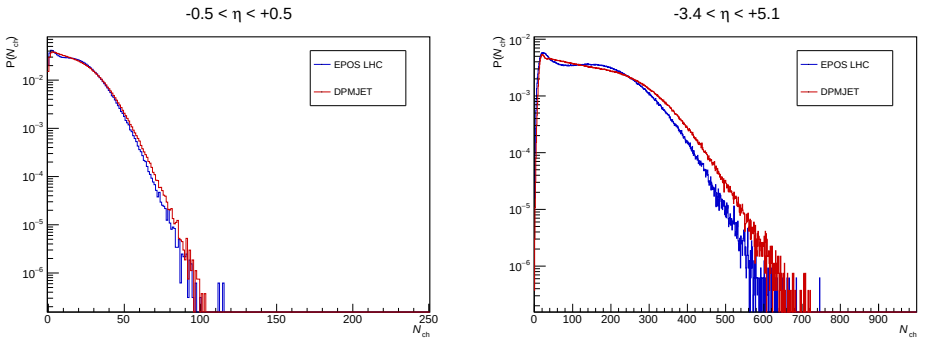


Figure 11.4: Charged-particle multiplicity distribution accordingly to EPOS LHC and DPMJET in $|\eta| < 0.5$ (left) and $-3.4 < \eta < +5.1$ (right) for p-Pb collisions at $\sqrt{S_{NN}} = 5.02$ TeV

It is possible to notice that, even though there are some differences between the two models, both of them can describe fairly well the data, especially in the five widest η ranges. There is an indication that the models start deviating from the data in the tails of the last three η ranges, $|\eta| < 1.5$, $|\eta| < 1.0$ and $|\eta| < 0.5$, however, DPJMET remains closer to the data and inside the error.

IP-Glasma

The distributions can be compared to the IP-Glasma model, described here [78], which is based on the Color Glass Condensate, explained in Section 3.1.1. The

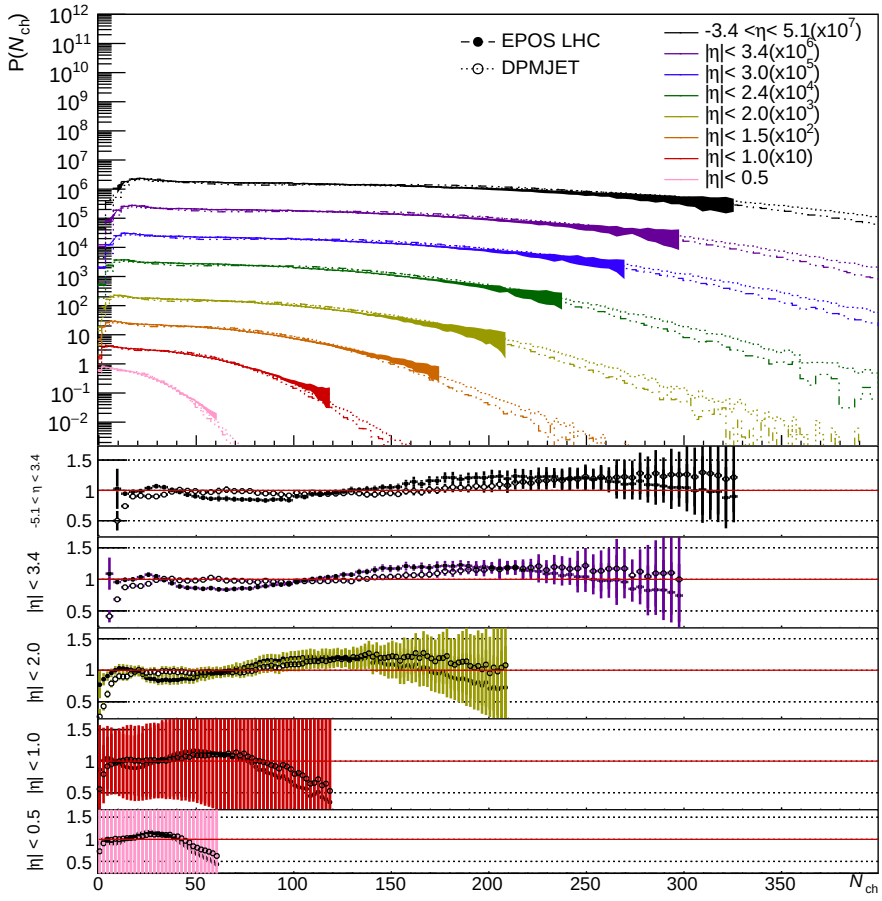


Figure 11.5: Comparison of the results with the calculated distributions using two different Monte Carlo models - EPOS LHC and DPMJET - for MB p-Pb collisions at $\sqrt{S_{NN}} = 5.02$ TeV.

comparison is done in the range $|\eta| < 0.5$, where the distributions are obtained using SPD clusters. Figure 11.7 show the comparison. In the distributions shown in green, the ration between the gluon saturation scale and density of color charge is fixed, in a way that no fluctuations are introduced. In the distribution shown in blue, the color charge density fluctuates around the mean, which makes it follows a Gaussian distribution with width of 0.09 [87].

The IP-Glasma model shown here presents no agreement with the data, and clearly underestimate the tails. It will be interesting to see if future improvements of the model can reproduce the data better.

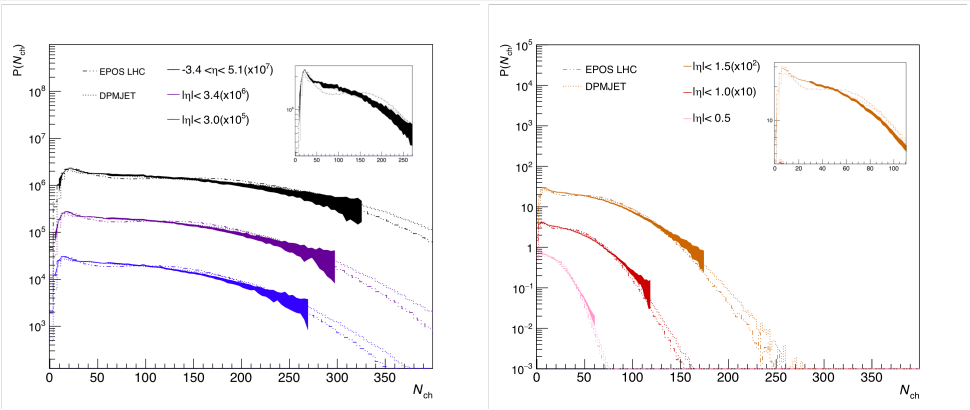


Figure 11.6: Closer view of the comparison data-MC models (EPOS LHC and DPMJET). The plot on the left shows closer the distributions in the three widest eta ranges and the plot on the left shows the distributions in the three narrowest eta ranges.

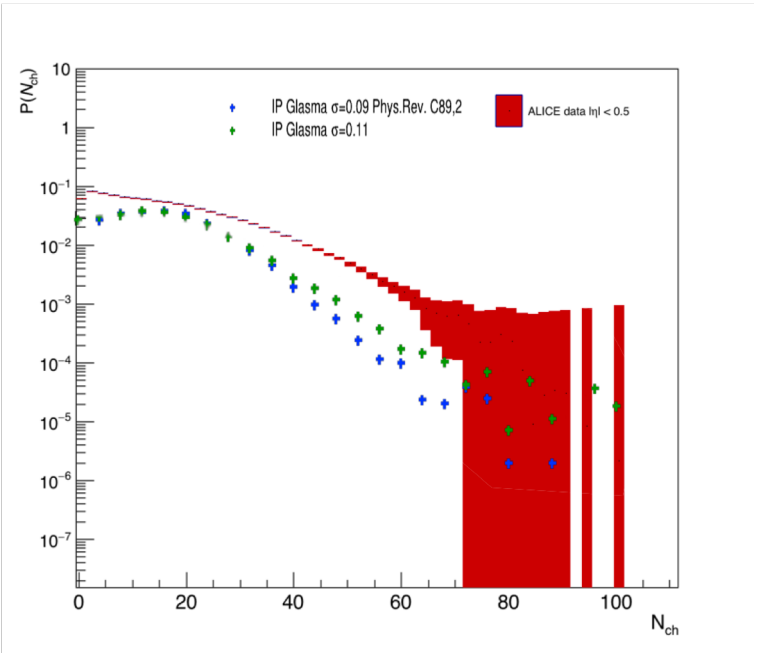


Figure 11.7: Charged-particle multiplicity distributions for p-Pb at $\sqrt{S_{NN}}$= 5.02 TeV compared to distributions from the IP-Glasma model. The IP-Glasma points are taken from [87].

11.2 Centrality Selection in p-Pb Collisions

Another way of analyzing the charged-particle multiplicity distributions is to divide them in different centrality classes. The concept of centrality and the different centrality estimators were defined in Section 9.1.3. In this work, two different centrality estimators were used to select the data in different centrality classes, one based on multiplicity, which uses the number of clusters in the outer layer of the SPD in the range $|\eta| < 1.4$ (CL1) and one based on zero-degree energy (ZNA). Monte Carlo studies were also performed in order to compare the performance of the two different models (DPMJET and EPOS LHC) and the estimators used. However, two different estimators had to be chosen to perform the MC studies, both based on multiplicity, namely V0A and CL1. This is due a current lack of centrality calibration for the ZNA estimator for ALICE in p-Pb collision Monte Carlo productions, which also limits the analysis of the data. Without the Monte Carlo calibration, it is not possible to divide the simulation in different centrality classes, therefore it is not possible to produce different response matrices and unfold the data. The Monte Carlo studies are shown in Appendix A.

11.2.1 Centrality Selection in Data and MC comparison

The analysis method to obtain charged-particle multiplicity distributions when dividing the data in centrality classes is still under development in ALICE. However, the preliminary results are presented here and the main issues are discussed. The data were selected in seven different centrality classes and. For the case of the selection using the CL1 estimator, the response matrices were also generated for the same centrality classes and the distributions were then unfolded. Figure 11.8 shows the data selected using the CL1 estimator compared to the Monte Carlo models EPOS LHC and DPJMET in the range $|\eta| < 0.5$ and 11.9 shows the distributions in the range $|\eta| < 3.4$.

From the comparison to the Models in different centrality classes, it is possible to see that the Monte Carlo models are in fair agreement with the data when it comes to the shape and position of the distributions, however, in both cases, the models overestimate the peaks and underestimate the tails.

It is also possible to plot the distributions for all the eight η ranges for one specific centrality class. Figure 11.10 shows the distributions obtained in the most central collisions, 00-00%, and 11.11 shows the distributions obtained in the most peripheral collisions, 80-100%.

In most central collisions, more particles are produced, according to the results obtained based on the CL1 centrality estimator. Furthermore, there is not a large overlap among the different η ranges, the widest η range has more particles and it covers a region from 300 charged-particles to 600 charged-particles, whereas the narrowest range has fewer particles covering a region from 10 to 100 particles. In Figure 11.10, it can be noted that there is an overlap among the η ranges.

The ZNA estimator is considerer bias-free, therefore would be the best choice to calculate the multiplicity distributions. However, it is still not possible to produce

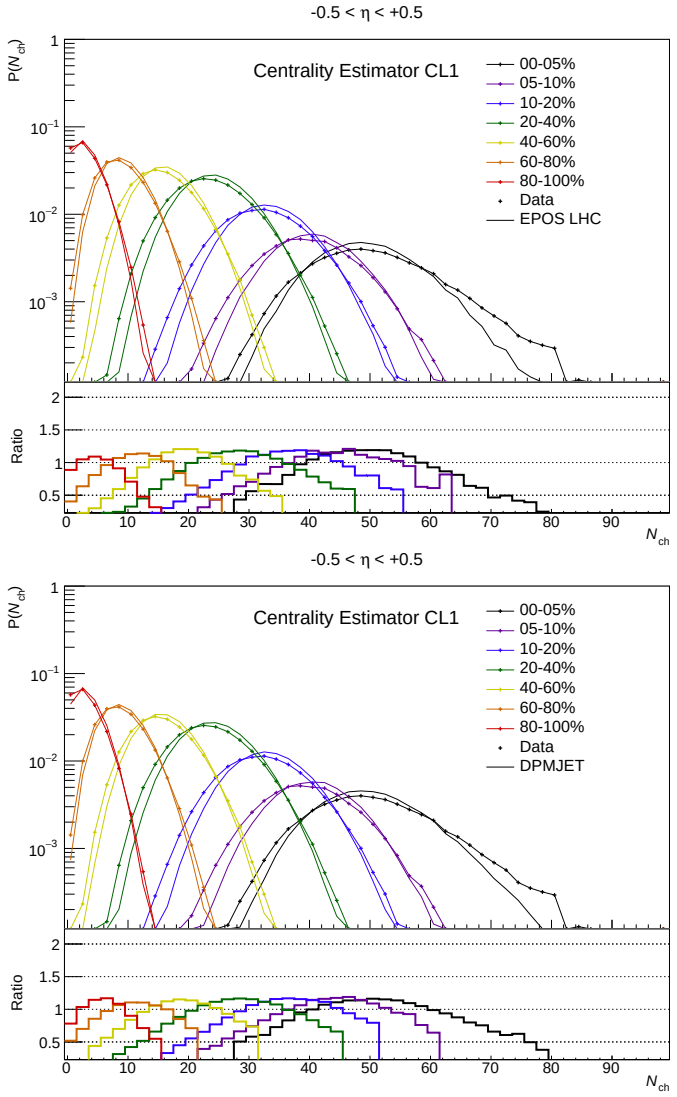


Figure 11.8: Comparison of charged-particle multiplicity distributions from p-Pb data in different centrality classes (using the estimator CL1) with the distributions produced using Monte Carlo simulations based on EPOS LHC (Top) and DPJMET (Bottom) in the pseudorapidity range $|\eta| < 0.5$. The bottom panels show the ratio of the model to the data.

response matrices for each centrality class and unfold the distributions. Nevertheless, the data can still be divided in centrality and the raw distributions are presented in Figure 11.12 , showing the results in two centrality classes, and in Figure 11.13,

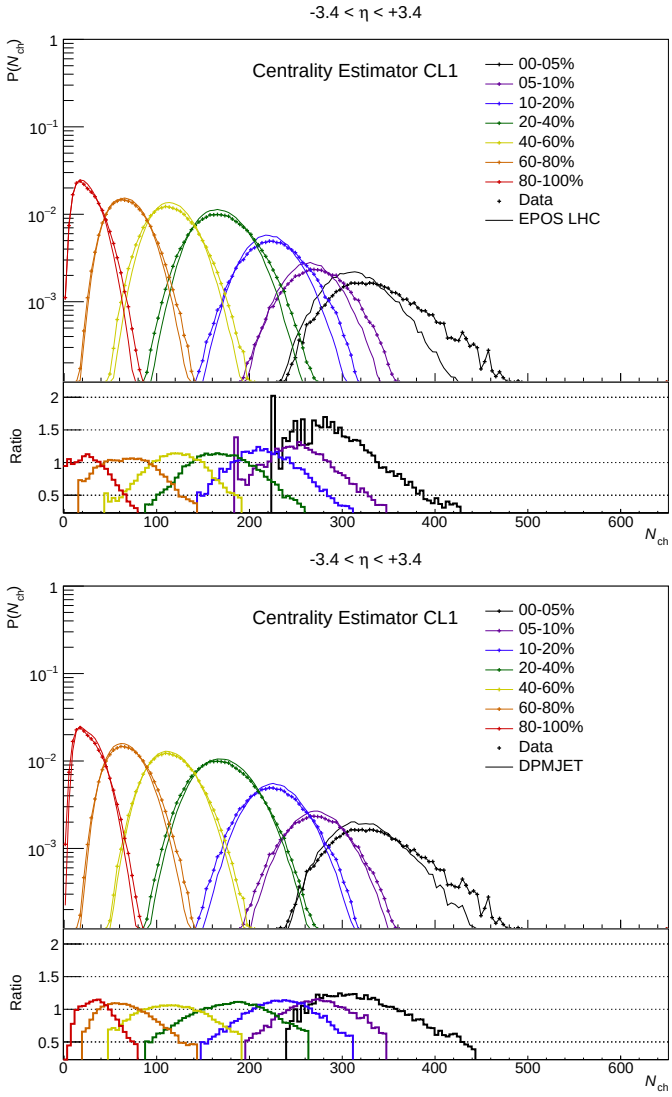


Figure 11.9: Comparison of charged-particle multiplicity distributions from p-Pb data in different centrality classes (using the estimator CL1) with the distributions produced using Monte Carlo simulations based on EPOS LHC (Top) and DPJMET (Bottom) in the pseudorapidity range $|\eta| < 3.4$. The bottom panels show the ratio of the model to the data.

showing the results in two η ranges.

In comparison to the centrality selection done by CL1, the distributions obtained using the ZNA estimator overlap considerably more, when looking at the distributions

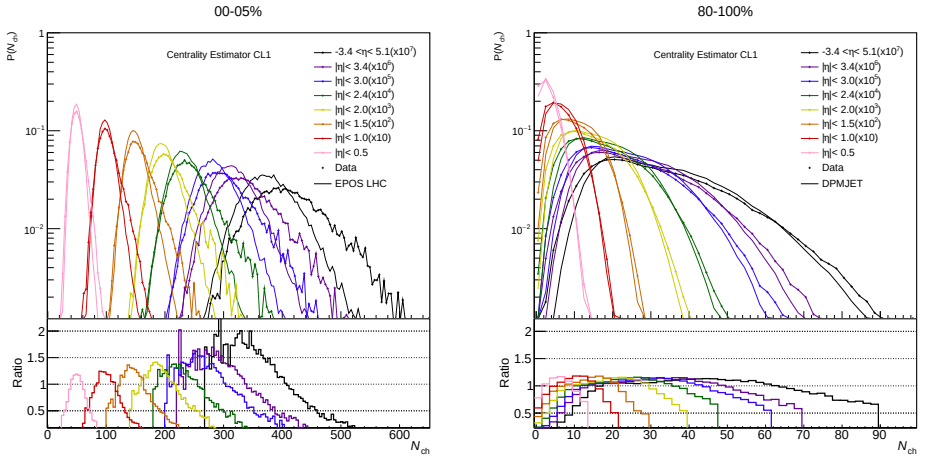


Figure 11.10: Comparison of charged-particle multiplicity distributions from p-Pb data in two different centrality classes (using the estimator CL1) , 00-05%(left) and 80-100%(right),with the distributions produced using Monte Carlo simulations based on EPOS LHC (Top) and DPJMET (Bottom) in the pseudorapidity range $|\eta| < 0.5$. The bottom panels show the ratio of the model to the data.

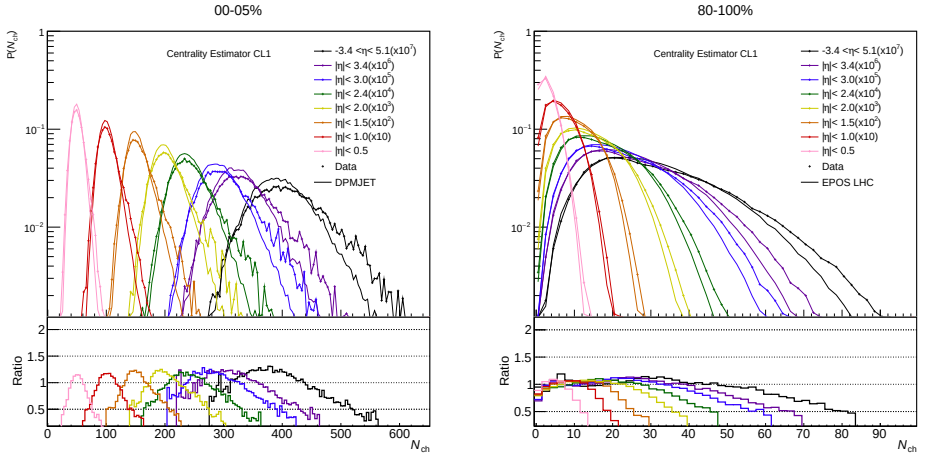


Figure 11.11: Comparison of charged-particle multiplicity distributions from p-Pb data in two different centrality classes (using the estimator CL1) , 00-05%(left) and 80-100%(right),with the distributions produced using Monte Carlo simulations based on EPOS LHC (Top) and DPJMET (Bottom) in the pseudorapidity range $|\eta| < 3.4$. The bottom panels show the ratio of the model to the data.

in the same η range, suggesting that peripheral events can produce up to 300 charged-particles and central collisions can originate $\sim$ 90 to $\sim$ 500 particles. In order to get

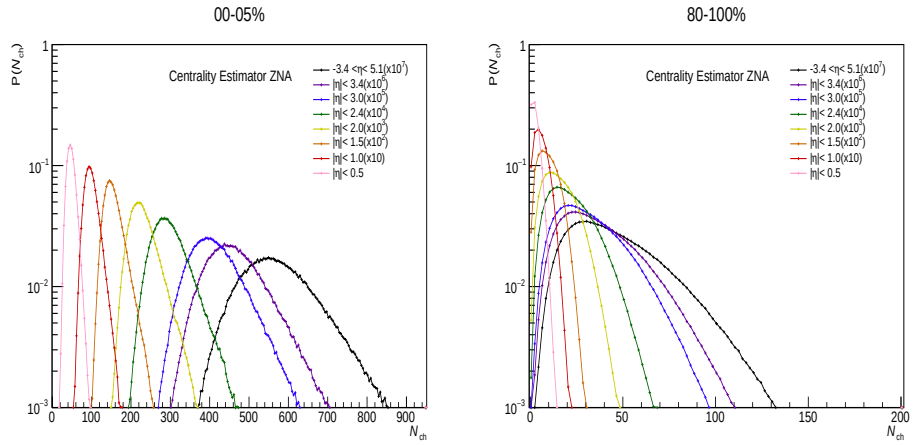


Figure 11.12: Charged-particle multiplicity distributions from pPb collisions data at $\sqrt{S_{NN}} = 5.02$ TeV in two different centrality classes (using the estimator ZNA), 00-05%(left) and 80-100%(right).

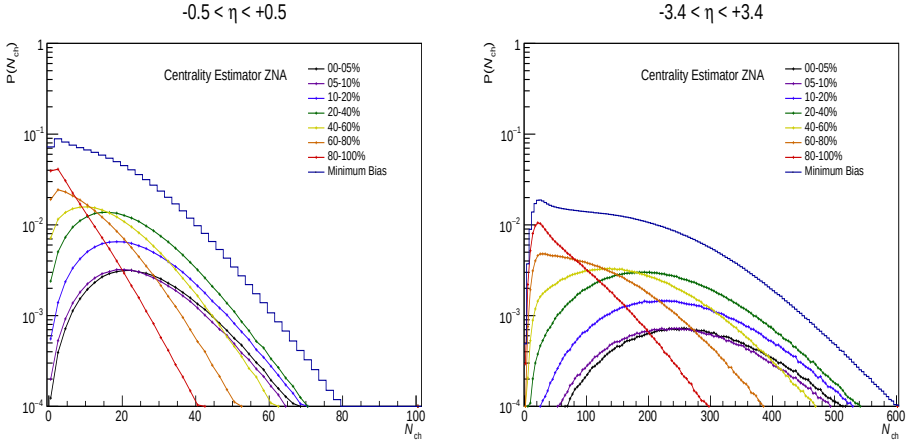


Figure 11.13: Charged-particle multiplicity distributions from pPb collisions data at $\sqrt{S_{NN}} = 5.02$ TeV in different centrality classes (using the estimator ZNA) in the ranges $|\eta| < 0.5$(left) and $|\eta| < 3.4$(right).

better and more precise results, the distributions should be corrected, which makes necessary the calibrations for the ZNA in Monte Carlo productions. Furthermore, a better understanding of the analysis methods and corrections that need to be applied to the final results is required. This is under development in ALICE and should be included in the future publication of this work. In this work, there is not a comparison of the data to the IP-Glasma model in different centrality classes since this calculation was not available.

Negative Binomial Distribution Fits

The charged-particle multiplicity distributions have been described by either a negative binomial distribution or double negative binomial distributions, as described in Chapter 5. It was tried to fit a double negative binomial distribution to the distributions, however, as described in Chapter 10, however it is not possible to fit the whole distribution with a double negative binomial distribution.

The fits in Chapter 10 were instead performed using a double negative binomial distribution convoluted with a gaussian. The impossibility of fitting the double negative binomial distributions to the multiplicity distributions could that the distributions can not be described by negative binomial distributions in this collision system. Proton-proton collisions have been described by double negative binomial distributions, however, pPb collisions are different since one has to take into account multiple collisions, which are expected to give additional contribution to the multiplicity, change its shape. This should be investigated further in order to find the best description of this system.

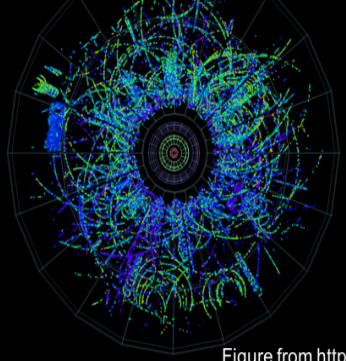


Figure from https://cds.cern.ch/record/1477949

Chapter 12

Summary

In 2013, LHC collected data of p-Pb collisions at $\sqrt{S_{NN}} = 5.02$ TeV in a test run for the first time. In 2016, data were collected with more statistics, which made many analysis possible and more accurate. Since then, signs of collectivity in small systems, pp and p-Pb collisions, were found, turning the detailed study of them necessary in order to better understand the origins of collectivity and of the Quark-Gluon Plasma. Each new result contributes to expand the overall knowledge of the systems and the charged-particle multiplicity distributions play an important role in order to provide constraints to theoretical models.

In this work, the theoretical background and motivation for obtaining charged-particle multiplicity distributions were discussed. The details of the analysis were presented as well as the experimental setup. The data collected with FMD and SPD were combined in order to obtain the $P(N_{ch})$ in eight different and increasing pseudorapidity ranges, from a very narrow interval $|\eta| < 0.5$ to a very wide and unique interval of $-3.4 < \eta < +5.1$. The results were compared to two Monte Carlo models, EPOS LHC and DPMJET, which are based on different theories. It was found that, overall, both models give a good account of the measured distributions. As mentioned throughout this thesis, the lack of a Monte Carlo production flat in multiplicity limits the unfolded distributions, the systematic uncertainties and, consequently, the estimations on how well the results agree to the models in the high multiplicity region. It is emphasized the importance of the flat production in order to avoid the loss of information, which can go up to $\sim 50\%$ of the possible true range of the distribution.

The results were also compared to the IP-Glasma model, showing a high disagreement. This disagreement might indicate the need of new models that are more accurate in describing LHC data and that can cover wider η ranges. It was also attempted to fit double negative binomial distributions to the charged-particle multiplicity distributions, however it is only possible to fit part of the distributions. A double negative binomial distribution convoluted with a gaussian distribution can be fitted to the measured distribution, which might be a result of the multiple collisions that the incoming proton undergoes when colliding with the lead nucleus.

Preliminary results of centrality selection of the data were also presented using one estimator based on the number of clusters in the outer layer of the silicon pixel detector (CL1) and one estimator based on zero-degree energy (ZNA). The data were divided in centrality classes using the CL1 centrality estimator, which is based on multiplicity, and unfolded with response matrices produced for each centrality range using the same estimator. The results were compared to the the distributions produced using EPOS LHC and DPJMET revealing that the models overestimate the peaks and underestimate the tails of the distributions in specific η ranges. The models, however, describe fairly well the shape and position of the distributions.

Further work has to be done to fill the gaps presented in this work. A flat Monte Carlo simulation has to be obtained in order to extend the range of the results and minimize the statistical errors in the high multiplicity region. This production is also going to make possible a better estimation of the systematic uncertainties, reducing the current total uncertainties. Besides the Flat Monte Carlo, it is also needed a calibration for using the ZNA centrality estimator in Monte Carlo simulations in order to be able to correct the raw distributions and estimate uncertainties. As previously mentioned, the analysis methods for obtaining charged-particle multiplicity distributions with centrality selection are still under development by myself and the Multiplicity Physics Analysis group from the ALICE collaboration.

The results in this work show interesting features, such as the impossibility of describing the distributions using the IP-Glasma model, which should be improved and support a larger pseudorapidity range, and the double negative binomial distributions. In the future, it is expected that the N_{ch} range of the distributions increase with an improved response matrix. Further investigation can lead to very exciting new results that might require new theories.

Bibliography

[1] Georges Aad et al. "Evidence for the spin-0 nature of the Higgs boson using ATLAS data". In: *Phys. Lett.* B726 (2013), pp. 120–144. DOI: 10.1016/j.physletb.2013.08.026. arXiv: 1307.1432 [hep-ex].

[2] K. Aamodt et al. "Higher harmonic anisotropic flow measurements of charged particles in Pb-Pb collisions at $\sqrt{s_{NN}}$=2.76 TeV". In: *Phys. Rev. Lett.* 107 (2011), p. 032301. DOI: 10.1103/PhysRevLett.107.032301. arXiv: 1105.3865 [nucl-ex].

[3] Ehab Abbas et al. "Centrality dependence of the pseudorapidity density distribution for charged particles in Pb-Pb collisions at $\sqrt{s_{\mathrm{NN}}}$ = 2.76 TeV". In: *Phys. Lett.* B726 (2013), pp. 610–622. DOI: 10.1016/j.physletb.2013.09.022. arXiv: 1304.0347 [nucl-ex].

[4] Betty Abelev et al. "Transverse momentum distribution and nuclear modification factor of charged particles in p-Pb collisions at $\sqrt{s_{NN}}$ = 5.02 TeV". In: *Phys. Rev. Lett.* 110.8 (2013), p. 082302. DOI: 10.1103/PhysRevLett.110.082302. arXiv: 1210.4520 [nucl-ex].

[5] S. Acharya et al. "Charged-particle multiplicity distributions over a wide pseudorapidity range in proton-proton collisions at $\sqrt{s}$ = 0.9, 7, and 8 TeV". In: *Eur. Phys. J.* C77.12 (2017), p. 852. DOI: 10.1140/epjc/s10052-017-5412-6. arXiv: 1708.01435 [hep-ex].

[6] Jaroslav Adam et al. "Centrality dependence of particle production in p-Pb collisions at $\sqrt{s_{\mathrm{NN}}}$= 5.02 TeV". In: *Phys. Rev.* C91.6 (2015), p. 064905. DOI: 10.1103/PhysRevC.91.064905. arXiv: 1412.6828 [nucl-ex].

[7] Jaroslav Adam et al. "Centrality dependence of the pseudorapidity density distribution for charged particles in Pb-Pb collisions at $\sqrt{s_{\mathrm{NN}}}$ = 5.02 TeV". In: *Phys. Lett.* B772 (2017), pp. 567–577. DOI: 10.1016/j.physletb.2017.07.017. arXiv: 1612.08966 [nucl-ex].

[8] Jaroslav Adam et al. "Centrality evolution of the charged-particle pseudorapidity density over a broad pseudorapidity range in Pb-Pb collisions at $\sqrt{s_{\mathrm{NN}}}$ = 2.76 TeV". In: *Phys. Lett.* B754 (2016), pp. 373–385. DOI: 10.1016/j.physletb.2015.12.082. arXiv: 1509.07299 [nucl-ex].

[9] J. Adams et al. “Evidence from d + Au measurements for final state suppression of high p(T) hadrons in Au+Au collisions at RHIC”. In: *Phys. Rev. Lett.* 91 (2003), p. 072304. DOI: 10.1103/PhysRevLett.91.072304. arXiv: nucl-ex/0306024 [nucl-ex].

[10] A Akindinov et al. “The multigap resistive plate chamber as a time-of-flight detector”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 456.1 (2000). Proceedings of the 5th Int. Workshop on Resistive Plate Chambers and Related Detectors, pp. 16–22. ISSN: 0168-9002. DOI: https://doi.org/10.1016/S0168-9002(00)00954-2. URL: http://www.sciencedirect.com/science/article/pii/S0168900200009542.

[11] ALICE Collaboration. [Online; accessed 24-October-2018]. URL: http://aliceinfo.cern.ch/Public/en/Chapter2/Chap2_dim_spec.html.

[12] ALICE Collaboration. [Online; accessed 24-October-2018]. URL: http://alice-offline.web.cern.ch/Activities/Analysis/AnalysisFramework/index.html.

[13] G J Alner et al. “A new empirical regularity for multiplicity distributions in place of KNO scaling”. In: *Phys. Lett. B* 160.CERN-EP-85-62 (Apr. 1985), 199–206. 18 p. URL: https://cds.cern.ch/record/159388.

[14] Martin B. R. *Nuclear and Particle Physics: An Introduction*. John Wiley Sons,Ltd, 2006. DOI: 10.1002/0470035471.

[15] Povh K. B. et al. *Particles and Nuclei: An Introduction to the Physical Concepts*. Springer-Verlag Berlin Heidelberg, 2015. DOI: 10.1007/978-3-662-46321-5.

[16] Gunnar S. Bali. “QCD forces and heavy quark bound states”. In: *Phys. Rept.* 343 (2001), pp. 1–136. DOI: 10.1016/S0370-1573(00)00079-X. arXiv: hep-ph/0001312 [hep-ph].

[17] Gordon Baym. “RHIC: From dreams to beams in two decades”. In: *Nucl. Phys.* A698 (2002), pp. XXIII–XXXII. DOI: 10.1016/S0375-9474(01)01342-2. arXiv: hep-ph/0104138 [hep-ph].

[18] S. Beole et al. “ALICE technical design report: Detector for high momentum PID”. In: *CERN-LHCC-98-19* (1998).

[19] Sylvia Berryman. “Democritus”. In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Winter 2016. Metaphysics Research Lab, Stanford University, 2016.

[20] P. Billoir and S. Qian. “Simultaneous pattern recognition and track fitting by the Kalman filtering method”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 294.1 (1990), pp. 219–228. ISSN: 0168-9002. DOI: https://doi.org/10.1016/0168-9002(90)91835-Y. URL: http://www.sciencedirect.com/science/article/pii/016890029091835Y.

[21] J. D. Bjorken. “Energy Loss of Energetic Partons in Quark - Gluon Plasma: Possible Extinction of High p(t) Jets in Hadron - Hadron Collisions”. In: (1982).

[22] Volker Blobel. “An Unfolding method for high-energy physics experiments”. In: *Advanced Statistical Techniques in Particle Physics. Proceedings, Conference, Durham, UK, March 18-22, 2002*. 2002, pp. 258–267. arXiv: hep-ex/0208022 [hep-ex]. URL: http://www.ippp.dur.ac.uk/Workshops/02/statistics/proceedings//blobel2.pdf.

[23] N. Bohr. “On the Constitution of Atoms and Molecules”. In: *Phil. Mag. Ser.6* 26 (1913), pp. 1–24. DOI: 10.1080/14786441308634955.

[24] M. Bondila et al. “ALICE T0 detector”. In: *IEEE Trans. Nucl. Sci.* 52 (2005), pp. 1705–1711. DOI: 10.1109/TNS.2005.856900.

[25] A Breakstone et al. “Charged multiplicity distribution in pp interactions at ISR energies”. In: *Phys. Rev. D* 30.CERN-EP-83-165 (Oct. 1983), 528–535. 24 p. URL: http://cds.cern.ch/record/150510.

[26] R. Brun and F. Rademakers. “ROOT: An object oriented data analysis framework”. In: *Nucl. Instrum. Meth.* A389 (1997), pp. 81–86. DOI: 10.1016/S0168-9002(97)00048-X.

[27] P. Buncic, A. J. Peters, and P. Saiz. “The AliEn system, status and perspectives”. In: *eConf* C0303241 (2003), MOAT004. arXiv: cs/0306067 [cs-dc].

[28] Tian Cao. *Conceptual Development of 20th Century Field Theories*. Cambridge University Press, 1997.

[29] A. Capella et al. “Dual parton model”. In: *Physics Reports* 236.4 (1994), pp. 225–329. ISSN: 0370-1573. DOI: https://doi.org/10.1016/0370-1573(94)90064-7. URL: http://www.sciencedirect.com/science/article/pii/0370157394900647.

[30] CERN. [Online; accessed 21-October-2018]. URL: https://home.cern/topics/large-hadron-collider.

[31] CERN. [Online; accessed 24-October-2018]. URL: http://alimonitor.cern.ch/map.jsp.

[32] Serguei Chatrchyan et al. “Observation and studies of jet quenching in PbPb collisions at nucleon-nucleon center-of-mass energy = 2.76 TeV”. In: *Phys. Rev.* C84 (2011), p. 024906. DOI: 10.1103/PhysRevC.84.024906. arXiv: 1102.1957 [nucl-ex].

[33] Serguei Chatrchyan et al. “Observation of long-range near-side angular correlations in proton-lead collisions at the LHC”. In: *Phys. Lett.* B718 (2013), pp. 795–814. DOI: 10.1016/j.physletb.2012.11.025. arXiv: 1210.5482 [nucl-ex].

[34] ALICE Collaboration. “Elliptic Flow of Charged Particles in Pb-Pb Collisions at $\sqrt{s_{NN}} = 2.76$ TeV”. In: *Phys. Rev. Lett.* 105 (25 Dec. 2010), p. 252302. DOI: 10.1103/PhysRevLett.105.252302. URL: https://link.aps.org/doi/10.1103/PhysRevLett.105.252302.

[35] ATLAS Collaboration. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: *Physics Letters* B716 (2012), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020. arXiv:1207.7214[hep-ex].

[36] The ALICE Collaboration. "The ALICE experiment at the CERN LHC". In: *JINST* 3 (2008), S08002. DOI: 10.1088/1748-0221/3/08/S08002.

[37] UA1 Collaboration. "A study of the general characteristics of proton-antiproton collisions at $\sqrt{s}$ =0.2 to 0.9 TeV". In: *Nuclear Physics B* 335.2 (1990), pp. 261–287. ISSN: 0550-3213. DOI: https://doi.org/10.1016/0550-3213(90)90493-W. URL: http://www.sciencedirect.com/science/article/pii/055032139090493W.

[38] UA1 Collaboration. "Charged particle multiplicity distributions in proton-antiproton collisions at 540 GeV centre of mass energy". In: *Physics Letters B* 123.1 (1983), pp. 108–114. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(83)90969-3. URL: http://www.sciencedirect.com/science/article/pii/0370269383909693.

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

[40] P. Cortese et al. "ALICE electromagnetic calorimeter technical design report". In: *CERN-LHCC-2008-014, CERN-ALICE-TDR-014* (2008).

[41] P. Cortese. "ALICE transition-radiation detector". In: (2001).

[42] G. D'Agostini. "A Multidimensional unfolding method based on Bayes' theorem". In: *Nucl. Instrum. Meth.* A362 (1995), pp. 487–498. DOI: 10.1016/0168-9002(95)00274-X.

[43] G. D'Agostini. "Improved iterative Bayesian unfolding". In: *ArXiv e-prints* (2010). arXiv: 1010.0632 [physics.data-an].

[44] G. Dellacasa et al. "ALICE technical design report of the inner tracking system (ITS)". In: *CERN-LHCC-99-12* (1999).

[45] G. Dellacasa et al. "ALICE technical design report of the photon spectrometer (PHOS)". In: *CERN-LHCC-99-04* (1999).

[46] G. Dellacasa et al. "ALICE technical design report of the time-of-flight system (TOF)". In: *CERN-LHCC-2000-012* (2000).

[47] R. Engel. "Photoproduction within the two-component Dual Parton Model: Amplitudes and cross sections". In: *Zeitschrift für Physik C Particles and Fields* 66.1 (Mar. 1, 1995), pp. 203–214. ISSN: 1431-5858. DOI: 10.1007/BF01496594. URL: https://doi.org/10.1007/BF01496594.

[48] Christian Wolfgang Fabjan, Yogendra Pathak Viyogi, and H de Groot. *ALICE Photon Multiplicity Detector (PMD): addendum to the Technical Design Report*. Technical Design Report ALICE. Submitted on 19 Sep 2003. Geneva: CERN, 2003. URL: https://cds.cern.ch/record/642177.

[49] Richard P. Feynman. "Very high-energy collisions of hadrons". In: *Phys. Rev. Lett.* 23 (1969). [,494(1969)], pp. 1415–1417. DOI: 10.1103/PhysRevLett.23.1415.

[50] C Fuglesang. "UA5 multiplicity distributions and fits of various functions". In: CERN-EP-89-135 (Oct. 1989), 19 p. URL: http://cds.cern.ch/record/202036.

[51] Kenji Fukushima and Tetsuo Hatsuda. "The phase diagram of dense QCD". In: *Reports on Progress in Physics* 74.1 (2011), p. 014001. URL: http://stacks.iop.org/0034-4885/74/i=1/a=014001.

[52] Alberto Giovannini and Roberto Ugoccioni. "Possible scenarios for soft and semihard components structure in central hadron hadron collisions in the TeV region: Pseudorapidity intervals". In: *Phys. Rev.* D60 (1999), p. 074027. DOI: 10.1103/PhysRevD.60.074027. arXiv: hep-ph/9905210 [hep-ph].

[53] Jan Fiete Grosse-Oetringhaus. "Measurement of the Charged-Particle Multiplicity in Proton-Proton Collisions with the ALICE Detector". PhD thesis. Munster U., 2009. URL: http://inspirehep.net/record/887184/files/CERN-THESIS-2009-033.pdf.

[54] Jan Fiete Grosse-Oetringhaus and Klaus Reygers. "Charged-Particle Multiplicity in Proton-Proton Collisions". In: *J. Phys.* G37 (2010), p. 083001. DOI: 10.1088/0954-3899/37/8/083001. arXiv: 0912.0023 [hep-ex].

[55] Rajan Gupta. "Introduction to lattice QCD: Course". In: *Probing the standard model of particle interactions. Proceedings, Summer School in Theoretical Physics, NATO Advanced Study Institute, 68th session, Les Houches, France, July 28-September 5, 1997. Pt. 1, 2.* 1997, pp. 83–219. arXiv: hep-lat/9807028 [hep-lat].

[56] Miklos Gyulassy et al. "Jet quenching and radiative energy loss in dense nuclear matter". In: (2003). DOI: 10.1142/9789812795533_0003. arXiv: nucl-th/0302077 [nucl-th].

[57] Christian Holm Christensen. "ALICE Forward Multiplicity Detector". PhD thesis. Copenhagen U., 2007. URL: http://inspirehep.net/record/886890/files/CERN-THESIS-2007-101.pdf.

[58] Edmond Iancu. "QCD in heavy ion collisions". In: *Proceedings, 2011 European School of High-Energy Physics (ESHEP 2011): Cheile Gradistei, Romania, September 7-20, 2011.* 2014, pp. 197–266. DOI: 10.5170/CERN-2014-003.197. arXiv: 1205.0579 [hep-ph]. URL: https://inspirehep.net/record/1113441/files/arXiv:1205.0579.pdf.

[59] Naoki Itoh. "Hydrostatic Equilibrium of Hypothetical Quark Stars". In: *Progress of Theoretical Physics* 44.1 (1970), pp. 291–292. DOI: 10.1143/PTP.44.291. eprint: /oup/backfile/content_public/journal/ptp/44/1/10.1143/ptp.44.291/2/44-1-291.pdf. URL: +%20http://dx.doi.org/10.1143/PTP.44.291.

[60] Vardan Khachatryan et al. "Charged particle multiplicities in pp interactions at $\sqrt{s} = 0.9$, 2.36, and 7 TeV". In: *JHEP* 01 (2011), p. 079. DOI: 10.1007/JHEP01(2011)079. arXiv: 1011.5531 [hep-ex].

[61] Vardan Khachatryan et al. "Observation of Long-Range Near-Side Angular Correlations in Proton-Proton Collisions at the LHC". In: *JHEP* 09 (2010), p. 091. DOI: 10.1007/JHEP09(2010)091. arXiv: 1009.4122 [hep-ex].

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

[63] T. Lappi and L. McLerran. "Some features of the glasma". In: *Nucl. Phys.* A772 (2006), pp. 200–212. DOI: 10.1016/j.nuclphysa.2006.04.001. arXiv: hep-ph/0602189 [hep-ph].

[64] William R. Leo. *Techniques for Nuclear and Particle Physics Experiments*. Springer-Verlag Berlin Heidelberg, 1994. DOI: 10.1007/978-3-642-57920-2.

[65] C. Lippmann. "The Time Projection Chamber for the ALICE Experiment". In: (2008). arXiv: 0809.5133 [nucl-ex].

[66] Joseph D. Lykken. "Beyond the Standard Model". In: *CERN Yellow Report CERN-2010-002, 101-109*. 2010. arXiv: 1005.1676 [hep-ph]. URL: http://lss.fnal.gov/cgi-bin/find_paper.pl?conf-10-103.

[67] P. Mermod. "Right-handed neutrinos: the hunt is on!" In: *Proceedings, Prospects in Neutrino Physics (NuPhys2016): London, UK, December 12-14, 2016*. 2017. arXiv: 1704.08635 [hep-ex].

[68] Swagato Mukherjee, Raju Venugopalan, and Yi Yin. "Real-time evolution of non-Gaussian cumulants in the QCD critical regime". In: *Phys. Rev. C* 92 (3 Sept. 2015), p. 034912. DOI: 10.1103/PhysRevC.92.034912. URL: https://link.aps.org/doi/10.1103/PhysRevC.92.034912.

[69] D Nouais et al. "The ALICE Silicon Drift Detector system". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 501.1 (2003). Proceedings of the 10th International Workshop on Vertex Detectors, pp. 119–125. ISSN: 0168-9002. DOI: https://doi.org/10.1016/S0168-9002(02)02020-X. URL: http://www.sciencedirect.com/science/article/pii/S016890020202020X.

[70] Casper Nygaard. "Charged particle multiplicity distributions into forward pseudorapidities in pp and PbPb collisions at the LHC". PhD thesis. Bohr Inst., 2011. URL: http://inspirehep.net/record/1231191/files/Casper_Nygaard.pdf.

[71] C. Patrignani et al. "Review of Particle Physics". In: *Chin. Phys.* C40.10 (2016), p. 100001. DOI: 10.1088/1674-1137/40/10/100001.

[72] Maxim Perelstein. "Introduction to Collider Physics". In: *Physics of the large and the small, TASI 09, proceedings of the Theoretical Advanced Study Institute in Elementary Particle Physics, Boulder, Colorado, USA, 1-26 June 2009*. 2011, pp. 421–486. DOI: 10.1142/9789814327183_0008. arXiv: 1002.0274 [hep-ph]. URL: https://inspirehep.net/record/844291/files/arXiv:1002.0274.pdf.

[73] MARTIN L. PERL, ERIC R. LEE, and DINESH LOOMBA. "A BRIEF REVIEW OF THE SEARCH FOR ISOLATABLE FRACTIONAL CHARGE ELEMENTARY PARTICLES". In: *Modern Physics Letters A* 19.35 (2004), pp. 2595–2610. DOI: 10.1142/S0217732304016019. eprint: http://www.worldscientific.com/doi/pdf/10.1142/S0217732304016019. URL: http://www.worldscientific.com/doi/abs/10.1142/S0217732304016019.

[74] T. Pierog et al. "EPOS LHC: Test of collective hadronization with data measured at the CERN Large Hadron Collider". In: *Phys. Rev.* C92.3 (2015), p. 034906. DOI: 10.1103/PhysRevC.92.034906. arXiv: 1306.0121 [hep-ph].

[75] Arthur M. Poskanzer and S. A. Voloshin. "Methods for analyzing anisotropic flow in relativistic nuclear collisions". In: *Phys. Rev.* C58 (1998), pp. 1671–1678. DOI: 10.1103/PhysRevC.58.1671. arXiv: nucl-ex/9805001 [nucl-ex].

[76] *QCD phase diagram*. https://www.bnl.gov/newsroom/news.php?a=24281. Accessed: 2018-09-09.

[77] F. Rimondi. "Multiplicity distributions in $\bar{p}p$ interactions at $\sqrt{s} = 1800$-GeV". In: *23rd International Symposium on Ultra-High Energy Multiparticle Phenomena Aspen, Colorado, September 12-17, 1993*. 1993, pp. 0400–404. URL: http://lss.fnal.gov/cgi-bin/find_paper.pl?conf-93-359.

[78] Bjoern Schenke, Prithwish Tribedy, and Raju Venugopalan. "Multiplicity distributions in p+p, p+A and A+A collisions from Yang-Mills dynamics". In: *Phys. Rev.* C89.2 (2014), p. 024901. DOI: 10.1103/PhysRevC.89.024901. arXiv: 1311.3636 [hep-ph].

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

[80] Raimond Snellings. "Collective Expansion at the LHC: selected ALICE anisotropic flow measurements". In: *J. Phys.* G41.12 (2014), p. 124007. DOI: 10.1088/0954-3899/41/12/124007. arXiv: 1408.2532 [nucl-ex].

[81] Raimond Snellings. "Elliptic Flow: A Brief Review". In: *New J. Phys.* 13 (2011), p. 055008. DOI: 10.1088/1367-2630/13/5/055008. arXiv: 1102.3010 [nucl-ex].

[82] *Standard Model of Elementary Particles*. https://commons.wikimedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg. Accessed: 2018-08-01.

[83] Mikhail Stephanov. "QCD Phase Diagram and the Critical Point". In: *Progress of Theoretical Physics Supplement* 153 (2004), pp. 139–156. DOI: 10.1143/PTPS.153.139. eprint: /oup/backfile/content_public/journal/ptps/153/10.1143/ptps.153.139/2/153-139.pdf. URL: +%20http://dx.doi.org/10.1143/PTPS.153.139.

[84] Arturo Tauro. "ALICE Schematics". General Photo. May 2017. URL: https://cds.cern.ch/record/2263642.

[85] J. J. Thomson. "On the structure of the atom: an investigation of the stability and periods of oscillation of a number of corpuscles arranged at equal intervals around the circumference of a circle; with application of the results to the theory of atomic structure". In: *Phil. Mag. Ser.6* 7.39 (1904), pp. 237–265. DOI: 10.1080/14786440409463107.

[86] *What Is Dark Matter?* https://www.nature.com/articles/d41586-018-05096-y. Accessed: 2018-09-16.

[87] Valentina Zaccolo. "Charged-Particle Multiplicity Distributions over Wide Pseudorapidity Range in Proton-Proton and Proton-Lead Collisions with ALICE". PhD thesis. Bohr Inst., 2015. URL: http://www.nbi.ku.dk/english/research/phd_theses/phd_theses_2015/valentina_zaccolo/.

[88] You Zhou. "Anisotropic flow measurements in Pb?Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV with ALICE". In: *J. Phys. Conf. Ser.* 779.1 (2017), p. 012055. DOI: 10.1088/1742-6596/779/1/012055. arXiv: 1609.04459 [nucl-ex].

Appendices

Appendix A

Monte Carlo Studies in Centrality Selection

The Monte Carlo studies are meant to study the differences between the two models, EPOS LHC and DPMJET, as well as the differences between the two centrality estimators, V0A and CL1. Centrality estimators based on multiplicity are not bias free, as mentioned in Section 9.1.3, however it is possible to compare the differences between the two estimators. Figure A.1 shows the multiplicity distributions produced using two different estimators in the range $|\eta| < 3.4$ and Figure A.2 in the range $|\eta| < 1.0$.

It is possible to observe that to evaluate the differences between the two models, it is necessary to take the used centrality estimator into consideration. When using the V0A estimator, both models seem to perform similarly in both η ranges with small differences in peripheral collisions, where it tends to show more particles in the distributions calculated using the EPOS LHC model than in those using DPMJET. In the scenario using the CL1 estimator, the differences between the two models are considerably higher in the widest η range. For the most peripheral collisions, the behavior is still the same, however it changed for mid and high-centrality collisions. In the narrowest η range, the models performed in a more similar way using the CL1 estimator.

Figure A.3 and A.4 present the distributions for all η ranges in a fixed centrality class. Both models, EPOS and DPMJET, perform in a similar way, as in the previous case. The differences become more pronounced accordingly to the centrality estimator, as one can see in the plots. Using the V0A, the differences are large in peripheral collisions and small in central collisions, whereas it is the opposite when CL1 is used.

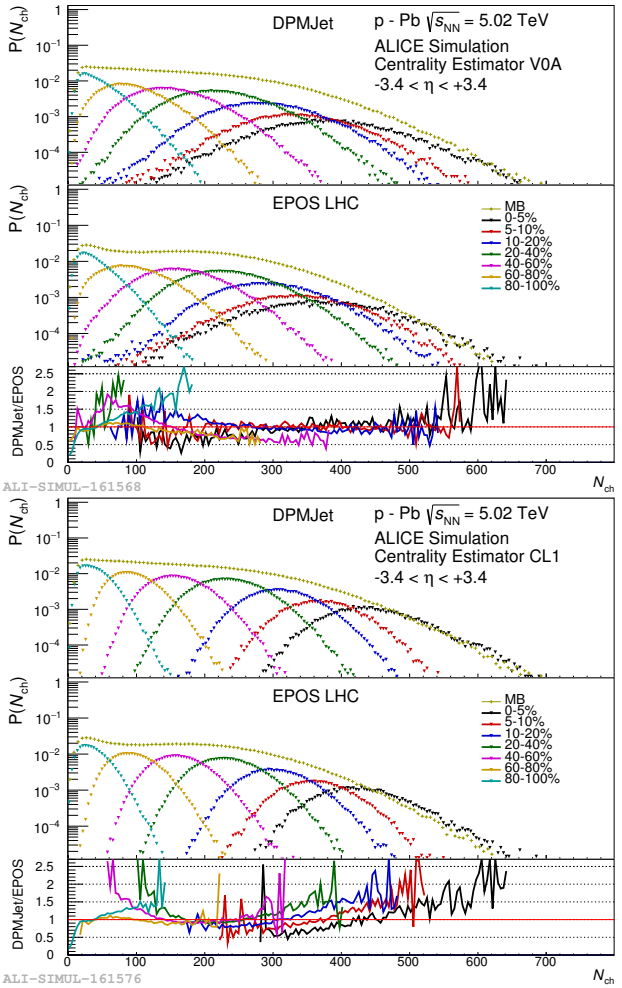


Figure A.1: Comparison between Charged-particle multiplicity distributions according to EPOS LHC and DPMJET divided in centrality classes using the estimators V0A (top plots) and CL1 (bottom plots) in $|\eta| < 3.4$.

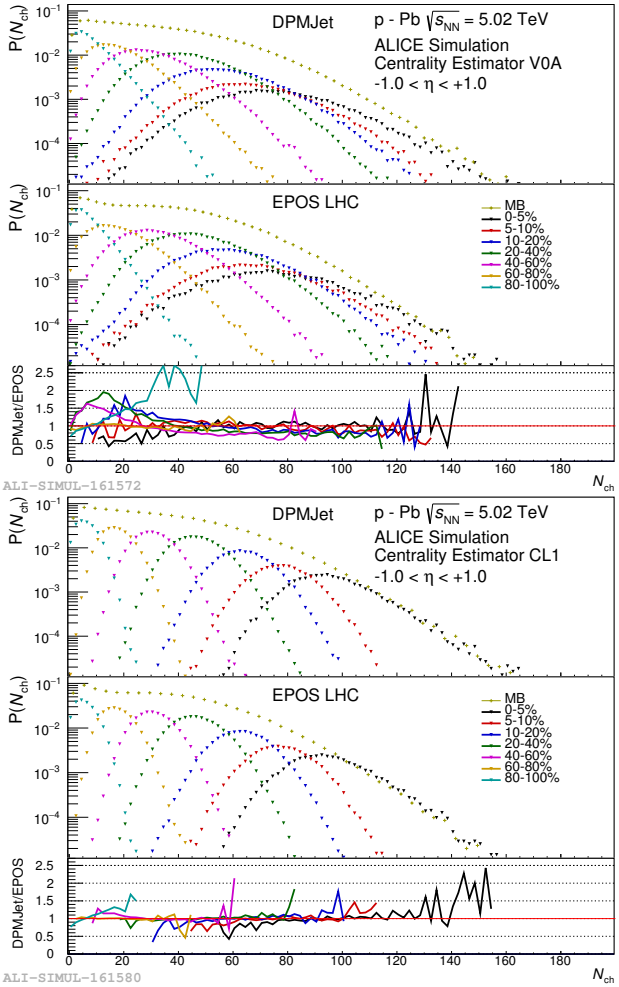


Figure A.2: Comparison between Charged-particle multiplicity distributions according to EPOS LHC and DPMJET divided in centrality classes using the estimators V0A (top plots) and CL1 (bottom plots) in $|\eta| < 1.0$.

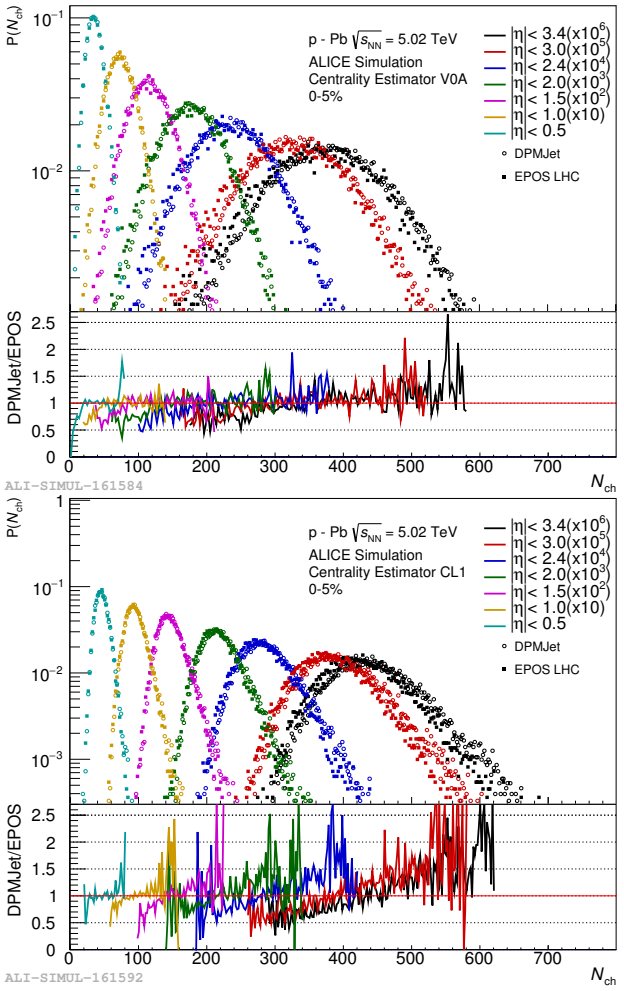


Figure A.3: Comparison of Charged-particle multiplicity distributions according to EPOS LHC and DPMJET divided in centrality classes using the estimators V0A (top plots) and CL1 (bottom plots) in central collisions (00-05%).

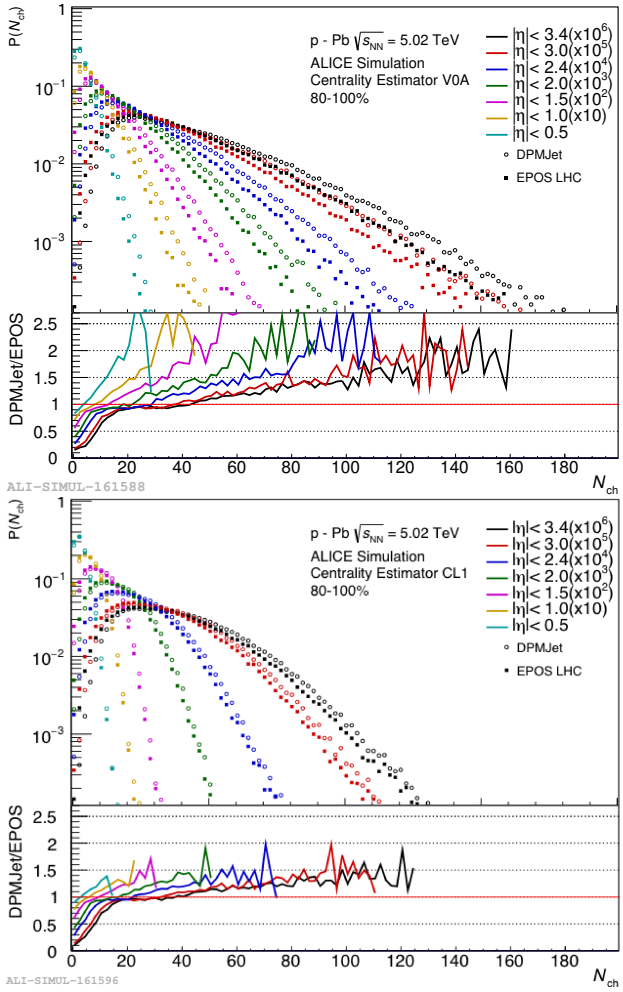


Figure A.4: Comparison of Charged-particle multiplicity distributions according to EPOS LHC and DPMJET divided in centrality classes using the estimators V0A (top plots) and CL1 (bottom plots) in peripheral collisions (80-100%).